Large-scale drivers of the mistral wind: link to Rossby wave life cycles and seasonal variability

Yonatan Givon*[1], Douglas Keller Jr.[2], Vered Silverman[1], Romain Pennel[2], Philippe Drobinski[2], Shira Raveh-Rubin[1]

1. Department of Earth and Planetary Sciences, Weizmann Institute of Science

2. Laboratoire de Météorologie Dynamique - IPSL, École Polytechnique, Institut Polytechnique de Paris, ENS, PSL Research University, Sorbonne Université, CNRS, Palaiseau France

Correspondence to: Yonatan Givon (yonatan.givon@weizmann.ac.il)

10   **Abstract.** The mistral is a northerly low level jet blowing through the Rhône valley in southern France, and down to the Gulf of Lions. It is co-located with the cold sector of a low level lee-cyclone in the Gulf of Genoa, behind an upper level trough north of the Alps. The mistral wind has long been associated with extreme weather events in the Mediterranean, and while extensive research focused on the low-tropospheric mistral and lee-cyclogenesis, the different upper-tropospheric large- and synoptic-scale settings involved in producing the mistral wind are not generally known. Here, the isentropic potential vorticity
15   (PV) structures governing the occurrence of the mistral wind are classified using a self-organizing map (SOM) clustering algorithm. Based upon a 36-year (1981-2016) mistral database and daily ERA-Interim isentropic PV data, 16 distinct mistral-associated PV structures emerge. Each classified flow pattern corresponds to a different type or stage of the Rossby wave life-cycle, from broad troughs, thin PV streamers, to distinguished cut-offs. Each of these PV patterns exhibit a distinct surface impact in terms of the surface cyclone, surface turbulent heat fluxes, wind, temperature and precipitation. A clear seasonal
20   separation between the clusters is evident and transitions between the clusters correspond to different Rossby wave-breaking processes. This analysis provides a new perspective on the variability of the mistral, and of the Genoa lee-cyclogenesis in general, linking the upper-level PV structures to their surface impact over Europe, the Mediterranean and north Africa.

1. Introduction:

The mistral is a northerly gap wind regime located at the Rhône valley in southern France. The Rhône valley separates the
25   Massif Central from the Alpine ridge by a ~50 km wide canyon, channelling northerly winds into the Gulf of Lions (GOL) in

the Mediterranean. The mistral winds yield the potential to deliver extreme weather impacts such as wildfires (Ruffault, et al., 2017), heavy precipitation (Berthou, et al., 2014; 2018) and direct wind damage (Obermann et al., 2018). It poses a frequent threat to agriculture, and is one of the most renowned weather phenomena in France. The mistral is seen as the primary source of severe storms and Mediterranean cyclogenesis (Drobinski et al., 2005) and is recognized as the most dangerous wind regime

30   in the Mediterranean (Jiang et al., 2003). The mistral outflow, composed of continental air-masses, picks up moisture at intense evaporation rates over the GOL, before flowing towards the lee-cyclone in the Gulf of Genoa (see Fig. 1) and further de-stabilizing it. Indeed, precipitation response to the mistral can be seen at the Dolomite Mountains east to Italy, where the mistral outflow is often headed, and the warm front of the Genoa low is often active. Rainaud et al. (2017) related strong mistral events to heavy precipitation events occurring along the European-Mediterranean coast line. This relationship is

35   manifested mainly by the re-moistening during mistral events, and the following flow of this moist air towards the European mountain slopes scattered along the coast. Classified as a dry air outbreak (Flamant, 2003), the mistral brings relatively dry and cold continental air masses into the Mediterranean Sea interface, resulting in massive air-sea heat exchanges. Berthou et al. (2014) found a significant sea surface cooling in the GOL in response to strong mistral events, which in turn weakened the following precipitation event occurring in southern France. The winter mistral often reduces SST (Li et al., 2006; Millot, 1979),

40   potentially destabilizing the water column, and indeed Schott et al. (1996) reported the initiation of deep convection in the GOL in response to strong surface cooling generated by a severe mistral case that took place in February 18-23, 1992. This atmosphere-ocean coupling (Lebeaupin-Brossier and Drobinski , 2009), in which the mistral lowers the SST in the GOL, destabilizing the water column and potentially triggering oceanic deep convection, might be viewed as an altitude-crossing mechanism, i.e., a pathway relating different atmospheric levels, in which anomalies from the tropopause (i.e., upper level

45   potential vorticity anomalies) modify the tropospheric flow via the mistral wind, with impacts all the way to the land or sea surface, and further down, essentially to the bottom of the Mediterranean sea, in cases of the onset of deep convection.

The mistral is dynamically connected to lee-cyclogenesis in the Gulf of Genoa (Drobinski et al., 2017; Guénard et al., 2005; Speranza et al., 1985; Tafferner, 1990; and others). The leading theories depicting the Alpine lee cyclogenesis process were recently reviewed (Buzzi et al., 2020). This process is characterized by an extra-tropical, baroclinic wave disturbance

50   composed of an upper level trough (at times accompanied by a surface low ahead), propagating towards the Alps. While the upper level trough propagates over the ~2.5 km mountain peaks of the Alps, at lower levels, the flow is blocked by the Alps, stalling the upper trough above the mountain range. A surface pressure dipole forms across the Alps, as low level cold air steadily accumulates on the wind-side of the Alps, and a mountain wake dominates the lee-side. Thus, a lee-cyclone is born, often in a phase-lock with its parent trough aloft. This initial stage of the lee cyclone is characterized with rapid deepening

55   rates, attributed primarily to geostrophic adjustment processes. As the jet streak propagates over the Alps, the underlying mass fluxes are deformed drastically, impairing geostrophic balance. A strong ageostrophic circulation across the orographic obstacle is generated, manifested by relatively strong upward/downward (~10 hPa/h, Jiang et al., 2003) motions in the wind/lee-side of the Alps, respectively. Both the upper trough and the surface low amplify rapidly, often forming a 'potential

vorticity (PV) tower' (Čampa and Wernli 2012; Rossa et al., 2000), virtually to the point where the upper level wave amplitude growth forces a wave-breaking. This initial stage is usually followed by a second baroclinic stage, in which the deepening rate is dialed down (Buzzi et al., 2020). This cyclogenesis process induces, and is enhanced by, a downslope, northerly jet centered at the Rhône valley, carrying the amounting cold air across the Alps and into the Mediterranean. This mid-level jet is the manifestation of the mistral wind. The mistral is accelerated through the Rhône valley, and more so as the valley opens to form a delta, apparently due to reduced surface drag (Drobinski et al., 2017). The mistral is then tilted eastward towards the Genoa low and surges across the GOL. In his study, Tafferner (1990) demonstrated the decisive impact of topography on Genoa lee-cyclogenesis using numerical models, and established that in the absence of topography, the Genoa cyclone may not form at all, and is instead replaced by a slowly-deepening surface low further to the east. He also points out that the main role of the jet-streak, recognized primarily as the Tramontane and Mistral winds, is in advecting high-PV air masses into the cyclogenesis region. One can presume that in the absence of topography, northerly flow in the GOL is not expected to accelerate as much, and mistral wind is expected to dramatically weaken. Mattocks and Bleck (1986) designed a numerical Quasi-Geostrophic (QG) experiment, and separately examined the role of topography and of the upper level PV anomaly in lee-cyclogenesis, and established that both are necessary in order to produce the rapid deepening stage of the lee-cyclogenesis reported by observations. Tsidulko and Alpert (2001) also examined the different contributions of topography and vorticity advection to Alpine cyclogenesis, and emphasized the synergetic nature of the PV-mountain interaction crucial for reproducing observed deepening rates. Thus, the mistral is understood as an integral part of the Alpine cyclogenesis process, making its connection to topography crucial. In the work of Guénard et al. (2005), the mistral structure appears to vary with response to gravity wave breaking, and a thematic separation between deep and shallow mistral is suggested, each responding to a different hydraulic state of the flow, ranging between flow splitting, mountain wake and gravity wave breaking. Furthermore, hydraulic jumps have been diagnosed at the upper edge of the boundary-layer during mistral events (Drobinski et al, 2001 a&b; Jiang et al., 2003), adding a strong boundary layer impact to the complicated cyclogenesis picture.

Past mistral-related studies are mostly localized (Drobinski et al., 2001a & b; Guénard et al., 2005), case-focused (Buzzi et al. 2003; Jiang et al 2003; Plačko-Vršnak et al., 2005), or very generalized (Buzzi et al.; 1986; Smith 1986; Tafferner 1990). However, little is known about the large-scale or synoptic state that induces, maintains and ends the mistral, beyond the mere presence of an upper level trough and a surface cyclone south of the Alps. The spatial and temporal variability within the mistral period remains poorly predicted, as is the mistral intensity and duration (Guénard et al., 2005). A systematic climatological classification of the large scale flow patterns associated with mistral events has, to the best of our knowledge, never been attempted. Therefore, this study is designed to address the following questions:

(i) Which upper-tropospheric features occur over the North East Atlantic and Europe during mistral days? Can they be classified into a coherent sequence of distinct, reccurring, dynamically significant flow patterns?

(ii) How do the mistral upper-tropospheric flow types vary throughout the year?

(iii) Does the impact on the hydrological cycle vary according to the upper-tropospheric flow type?

(iv) Are there typical life cycles of the flow types? Are some flow types more persistent than others?

Addressing these questions will enhance our understanding of mistral variability and predictability. Here, we address these questions by classifying the regional upper-tropospheric PV distribution during mistral periods, and quantify their surface impact in terms of precipitation and surface turbulent heat fluxes. Relying on the PV perspective, we speculate that the upper-level PV field bears the largest influence upon the mistral compared to other atmospheric parameters. Being a conservative quantity, PV can potentially reveal consistent yet unique flow patterns producing the mistral wind. Moreover, the fine-structures typical to isentropic PV surfaces allows one to identify a variety of different flow structures, as opposed to smooth wave-like formations presented by the geopotential field.

The climatological mistral database and the clustering approach are detailed in Section 2. The resulting flow types and the corresponding impact are presented along with three illustrative cases in Section 3. The findings are then summarized in Section 4.

2. Methods:

This study is based on climatological data for the 36-year period of 1981-2016. Objective mistral criteria are applied to identify mistral days and create a mistral database for its subsequent classification. The classification is performed by a self-organizing map algorithm (SOM; see Section 2.2) applied to upper-tropospheric isentropic PV during mistral days, and is followed by a subsequent persistence-transition analysis of the resulting clusters.

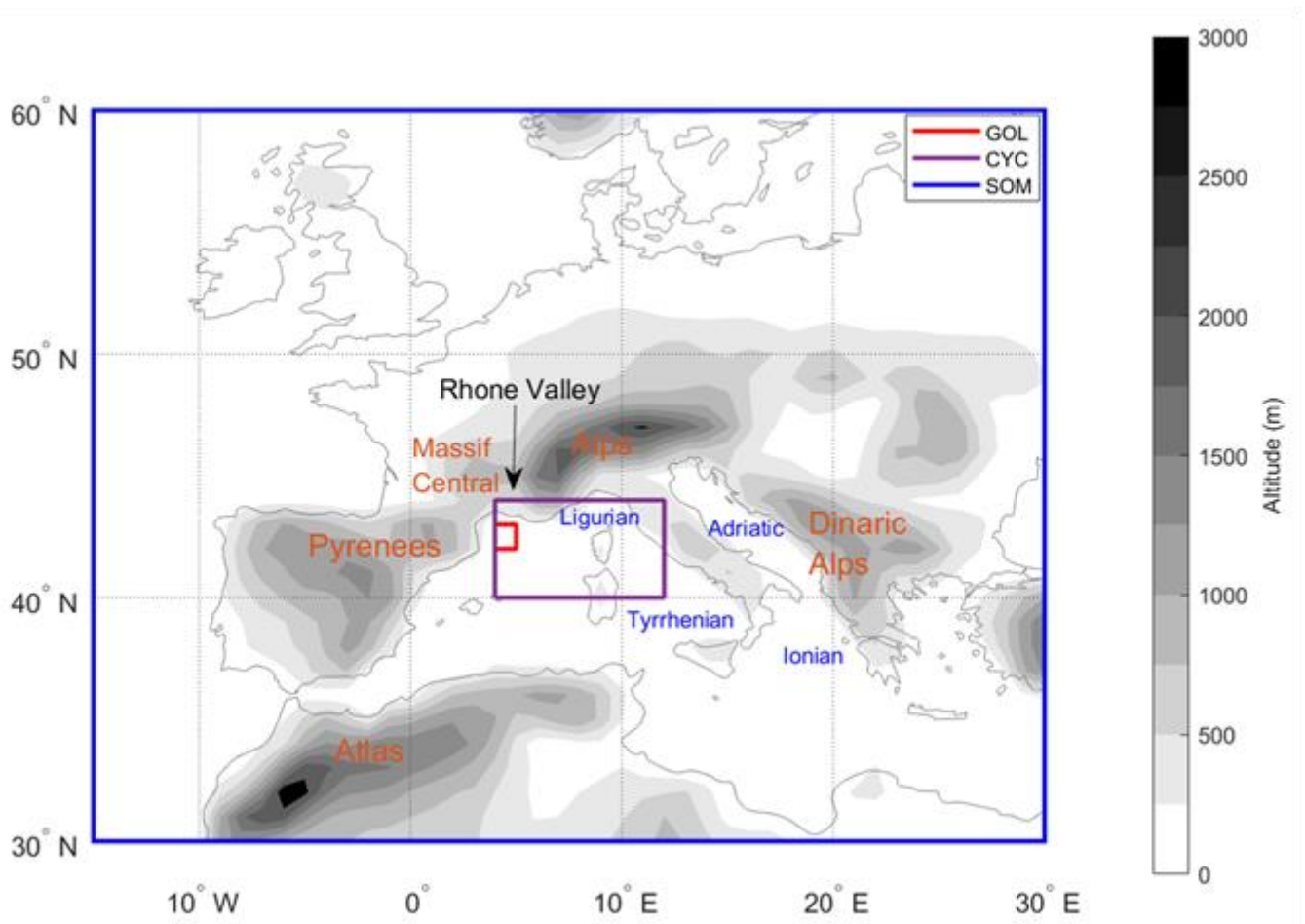



Figure 1: Geographical domains used in the analysis (See Methods section). The whole domain is used as SOM analysis input; the
110 **subregion where the Genoa cyclones are detected (CYC), and the subregion in which a northerly flow is required by the mistral**
definition (GOL) are marked in purple and red boxes, respectively. The topographic height from ERA Interim is shown in shading
(m).

2.1. Data and Mistral Criteria

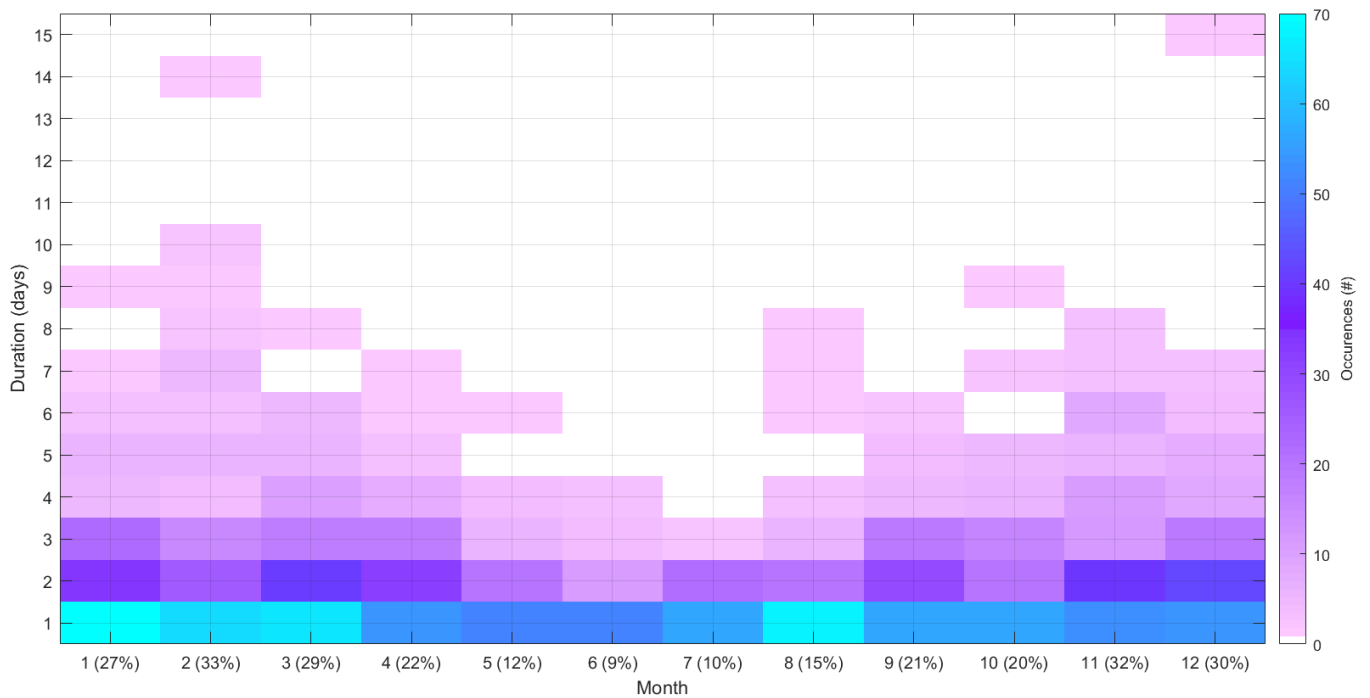
The data used to objectively define mistral days was obtained by a combination of the ERA-Interim reanalysis (Dee et al.,
115 2011) of the European Centre for Medium-range Weather Forecasts (ECMWF), and the regional climate model WRF-ORCHIDEE, downscaled from ERA-Interim, for the years 1981-2016. The WRF-ORCHIDEE model, at 20-km resolution, was performed by the IPSL Regional Climate Model (RegIPSL). The modelling framework is similar to Lebeaupin-Brossier et al. (2013) but uses the land surface model ORCHIDEE instead of DIFF (Drobinski et al., 2012; Guion et al., 2021).

To identify mistral days, first, a Genoa cyclone database is defined based on the presence of a cyclone in the CYC domain

120 (Fig. 1) in ERA-Interim, using the sea-level pressure field at 1-degree horizontal resolution and 6-hourly time intervals. The cyclone masks are identified in ERA Interim, at every 6-h time step, as the area within the outermost closed contour around local minima of the sea-level pressure field using 0.5-hPa intervals, adapted from Wernli and Schwierz (2006). To fulfill the cyclone criteria, a day should have at least one time-step in which a cyclone is identified. Then, wind direction and speed criteria were applied to days when a Genoa cyclone is detected. This coupling between the wind speed in the GOL and cyclone

125 in the Genoa region is considered the fundamental of lee-cyclogenesis in the Alps (Drobinski et al., 2005; Flamant, 2003; Lebeaupin Brossier et al., 2013; Burlando, 2009), and reassures that the mistral set of events under consideration is analogous to events of lee-cyclogenesis, filtering out northerlies stemming from sea-breeze dynamics (Drobinski et al., 2018). The wind criteria are defined as follows, using the daily WRF-ORCHIDEE model data that allows higher spatial resolution: NW to NE (i.e., +-45º) wind direction at 900 hPa and 10-m wind speed of at least 2 m/s, averaged in the GOL domain (Fig. 1). The

130 objective identification yielded 2734 mistral days, comprising 21% year-round frequency, in agreement with Burlando (2009). Consecutive mistral days were grouped into mistral events. The identified mistral events duration and monthly frequency are shown in Fig. 2. Mistral events peak typically in Jan-Feb at 30% frequency, while the mistral is less frequent in summer (~10%). Most mistral events last a single day. Duration of more than 4-8 days occurs exclusively in the autumn and winter months. This distribution generally agrees with the climatological lifetime properties of the Genoa low (e.g., Campins et al.

135 2011). The mistral identification method is robust with regard to variations in the wind speed threshold, and shows large sensitivity to the allowed wind direction. The additional criterion for the presence of a cyclone constrains the mistral frequency to the range of 20-30%, depending almost entirely on the opening angle criteria.

140 **Figure 2: Climatological monthly frequency of mistral duration (days). Shading represents the number of events within the division, for the entire time period. The percentages on the x-axis show the monthly frequency of mistral days.**

2.2. Isentropic PV Classification

2.2.1 Approach

To understand the variability of the synoptic environment during mistral, isentropic PV was classified during all identified
145 mistral days. The classification was achieved by the self-organizing map (SOM) algorithm, reviewed by Sheridan and Lee (2011) and Liu and Weisberg (2011) in the context of synoptic meteorological classification. In contrast to more conservative clustering methods, the SOM encompasses the full continuum of the system's variability, rather than only the dominant mean states. According to Sheridan and Lee (2011), for a successful SOM clustering analysis, the variance of the analysed field can be well represented by a sequence of consistent system states, each of which corresponding to an objectively related group of
150 samples. The choice of isentropic PV as the input field for the SOM analysis stemmed from both the natural advantages of using a conserved quantity, and its wide range of manifestations relevant for the mistral wind. Isentropic PV enables to depict fine structures such as PV streamers and cut offs attributed to Rossby wave-breaking, compared to the smoother geopotential height field. The main guidelines for a successful clustering analysis were a clear seasonal separation, a large enough number of members for each cluster, and a captured PV distribution that will emphasize the fine structures that are often averaged out

of long-term PV composites. Huang et al., (2017) used a SOM algorithm combined with Hierarchical Ascendant Classification (HAC) (Jain and Dubes, 1988) to classify winter 300-K isentropic PV regimes over East Asia and related the classes to cold surges occurring in Eastern Asia. Here, the data used for the clustering analysis is the ERA-Interim daily mean vertically-averaged upper-tropospheric isentropic PV field, within the SOM domain defined by the blue box in Fig. 1. Isentropic PV is averaged between the 320-340 K isentropic levels, with intervals of 5 K. This averaging is meant to overcome the seasonal temperature variation and allow a year-round analysis (Wernli and Sprenger, 2007). As the SOM algorithm is optimized for input values ranging from 0 to 1, the vertically averaged isentropic PV distribution is rescaled by the ratio between the mean PV and the mean standard deviation, specifically:

$$PV_{norm} = PV_{(x,y,t)} * (\frac{\overline{PV}}{\overline{STD}})$$

Where the overbar denotes spatio-temporal average, thus the ratio on the right hand side is a constant equal to 0.0732, so that PV_{norm} is roughly distributed between 0 and 1, as required by the SOM algorithm. Here, the SOM analysis is performed for the domain between 30-60 N and -15-30 E, a compromise between the largest domain centered at the GOL that preserved a seasonal signal, and the preserving of the identified flow patterns captured by its smaller alternatives.

The SOM attribute each mistral day to one cluster. Using ERA-Interim data, we further examined the mean surface conditions for each of the resulting clusters of mistral days, to unravel the possible link between the different PV-based clusters and the surface flow, surface turbulent heat fluxes, temperature and precipitation. Transitions among clusters during mistral were then systematically examined to reveal recurring transition sequences and their seasonal variability. Finally, we demonstrate the representativeness of the clusters and transition paths to individual cases.

2.2.2. SOM Set-up and Validation

The SOM Algorithm is provided by Mathworks:
 (https://www.mathworks.com/help/deeplearning/gs/cluster-data-with-a-self-organizing-map.html). The setting used in the present study yielded a shallow neural network with a single layer, using the sigmoid activation function, in hexagonal grid (Fig. A1). The hexagonal grid setting implies that each cluster interacts with the 6 surrounding clusters in terms of similarity, however the results are displayed on a rectangular grid to simplify the presentation. The choice of the map dimension (and therefore number of clusters) was set as the number beyond which the classification method begins to deteriorate, or does not add relevant information (i.e., near empty clusters or highly similar ones). Here we aimed to classify the mistral events into clusters that pose dynamically-meaningful PV distributions, representative of their daily individual members, and with a considerable annual mean frequency (i.e., on the order of 5%). Eventually, the number of clusters was set to 16 in a 4x4 configuration (Fig. A1), which satisfied these demands.

185 The SOM algorithm learning process optimizes a chosen function indicating each cluster's inner variance, and despite the wide variety of relevant functions, well-performed SOM processes are usually quite indifferent to the chosen function (Sheridan and Lee, 2011). This statement is usually true for other chosen parameters such as the neighborhood size and calculation time steps, within a reasonable range of values (e.g., Cassano et al., 2006; Johnson et al., 2008). For the present study we chose the intuitive MSE function as the optimization parameter with the initial neighborhood distance set to 2 and number of training

190 steps for initial covering of the input space set to 365. Indeed, the identified patterns were only weakly affected by the choice of these parameters. Furthermore, the SOM readily reproduced similar average patterns for several different mistral datasets. For example, the process was repeated for a subset of mistral dates in which single-day mistral events were removed and events separated by a single day were joined. This modified subset included 248 days less than the original, reducing the sample size by over 10%, yet the identified patterns were nearly identical.

195 Statistical significance is assessed by the student's t-test between each cluster and the total averaged mistral flow. Following Wilks et al. (2016), an additional criterion was added to the statistical test to account for the multiple testing problem. Specifically, the Walker criteria was applied, where a threshold of $\alpha_{walker} = 1 - (1 - \alpha_0)^{N_0^{-1}}$ is set on the p-value obtained by the t-test. α_0 is the required significance level (e.g. 0.05 for 95% confidence) and N_0 the number of individual t-tests, or in this case, the number of grid points in the domain of interest. If the p-value of an individual t-test is larger than α_{walker}, the

200 null hypothesis cannot be rejected at a level of α_0.

3. Results and Discussion:

3.1 PV Distribution Clusters

The classification of isentropic PV resulted in 16 clusters, set in a 4X4 hexagonal grid, where similar clusters are placed closer together in the SOM space. The identified mean PV patterns defining each cluster are displayed in Fig. 3, alongside the mean

205 500-hPa geopotential heights. The panel order represents each cluster location in the SOM space, i.e., the least similar clusters are placed farther apart in the SOM space, and more similar ones are placed adjacent to one another. Some clusters correspond to exotic upper-tropospheric PV structures, such as thin streamers and cut-offs, attributed to different Rossby-wave breaking (RWB) events. Dotted regions indicate statistical significance, i.e., the main features by which the classification process is established. For instance, a high PV tongue to the west of a cut-off appears to define cluster 8, while a westward stretching PV

210 streamer defines cluster 9, suggesting these clusters correspond to cyclonic or anticyclonic RWB lifecycles (Thorncroft et al., 1993), respectively. A southwesterly-oriented cut-off defines cluster 5, while a northerly thin streamer defines cluster 2, and so on. Together, the clusters illustrate a thematic separation of the PV continuum responsible for mistral events, and one can easily envision how the waves propagate by switching from one cluster to another.

Upstream of the high-PV anomaly, most clusters exhibit an amplified ridge in the upper troposphere over the Atlantic, a common precursor for intense Mediterranean cyclones (Raveh-Rubin and Flaounas 2017). As expected, the primary mode of variance is in the seasonal cycle, manifested by the meridional shift of the dynamical tropopause (subsection 3.2). Another mode apparently picked up by the SOM is evident when comparing clusters 12 and 16 to 11 and 15. Evidently, the SOM was able to distinguish between mountain-passed PV anomalies (11 and 15) and blocked ones (12 and 16), and the impact on the trough properties is evident by the tilting of the trough axis upon the passage across the Alps from NE to N, respectively. The standard deviation (STD) for the PV distribution within the clusters is presented in Fig. A2 in the appendix, emphasizing the different active regions between the clusters.

Note that the composites presented on Fig. 3 each constitute ~100 days. While some variance within the clusters is inevitable, these composites help to illustrate the SOM clustering process, especially in statistically significant regions. Detailed features of the actual patterns, as picked up by the SOM, can be better understood by carefully examining the cluster members with respect to their mean values. Such examples are provided in Sect. 3.6, aiding in establishing the cluster features, detailed in Table 1.

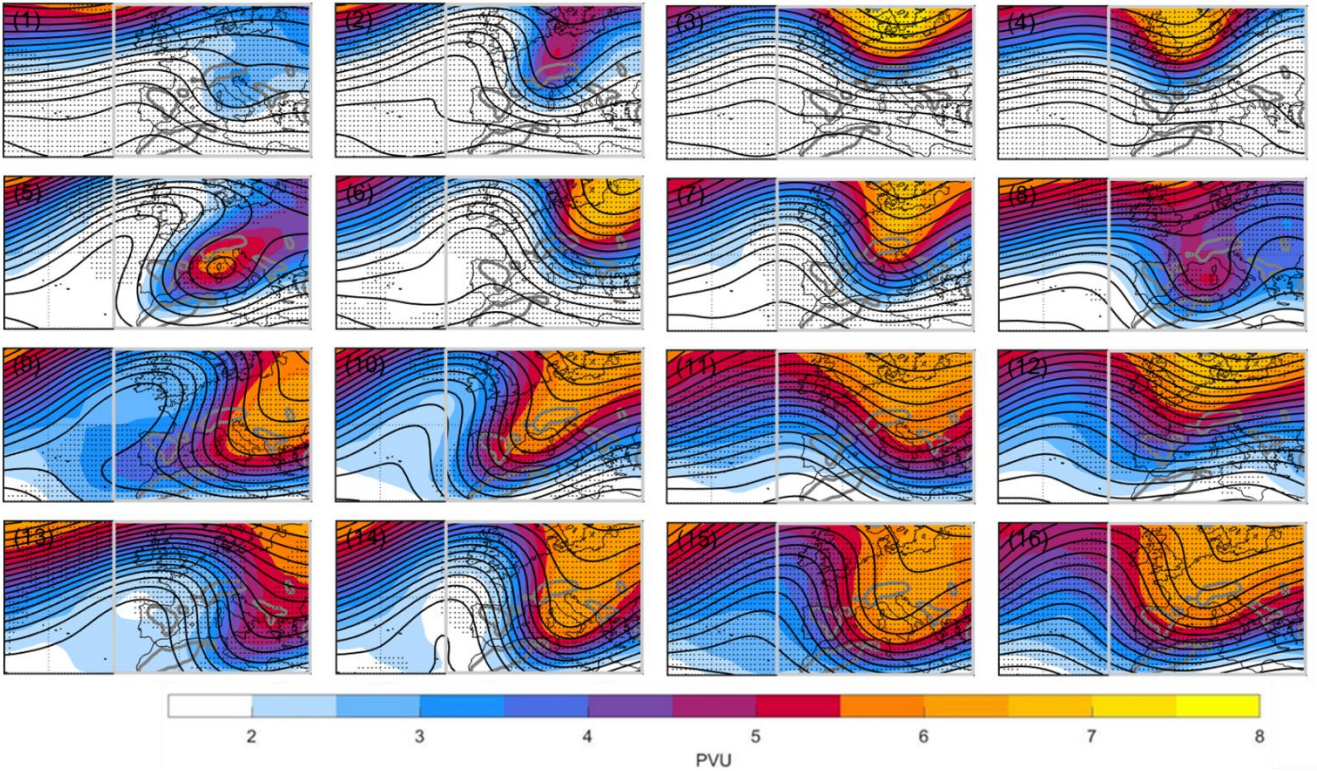



Figure 3: Daily mean vertically-averaged isentropic (320-340K) PV (PVU, color) and 500-hPa geopotential height (unlabeled black contours, 30 m intervals) of the SOM-identified clusters. The topographic height contour of 700 m is shown in dark grey. Dotted

regions indicate statistical significance >95% for the PV composites compared to the average mistral flow. The bright grey frame indicates the domain in which the SOM classification was done.

3.2 Seasonal Variation

The climatological monthly occurrence frequency of each cluster is displayed in Fig. 4, demonstrating the strong seasonal
235 affiliation of the clusters, i.e., all clusters have a clear seasonal peak in their occurrence. Clusters 1-4 occur mostly between June and October, while clusters 9-16 occur mainly between November and April. The low-PV background clearly dominates summer clusters (Fig. 3 panels 1-4) while much broader wave amplitudes constitute the winter clusters (9-16). In between, clusters 5-8 peak mainly in the transition seasons. Overall higher frequencies are obtained in the winter clusters, as expected by the larger frequency of mistral events appearing in winter.

240

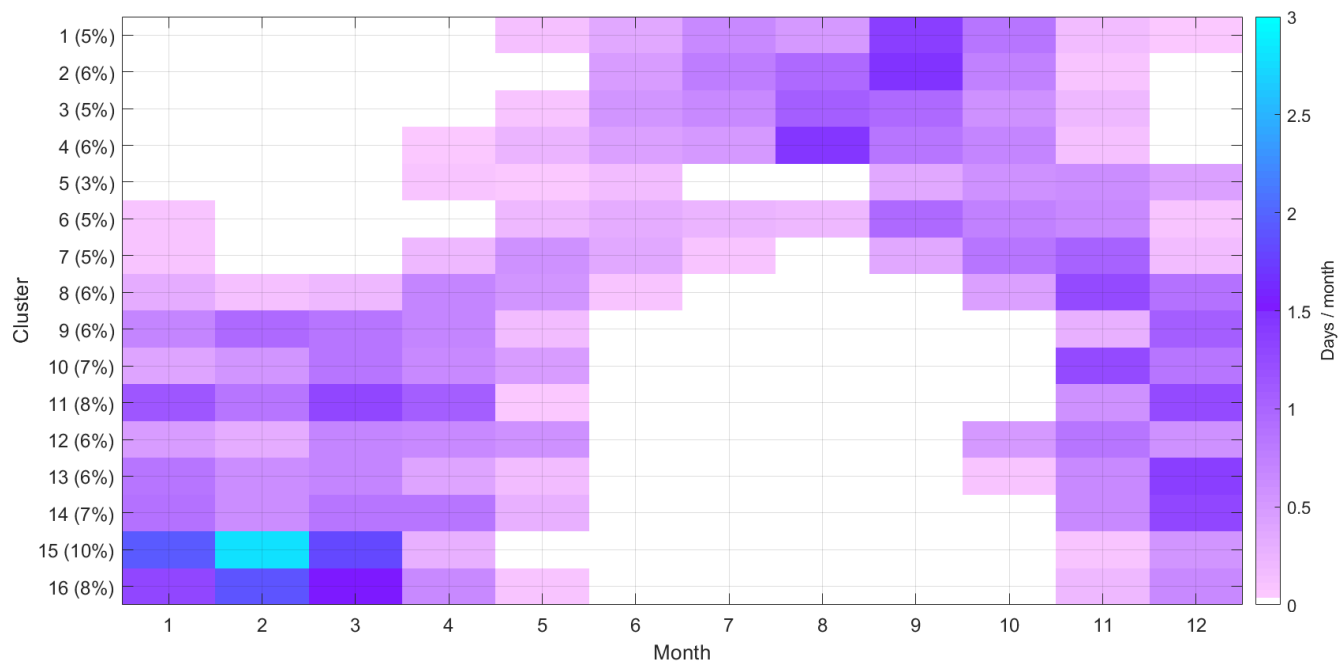



Figure 4: Monthly frequency of the classified mistral clusters (days per month). The percentage of mistral days corresponding to each cluster is shown on the y-axis (%).

3.3 Surface Circulation and Surface Impact

245 The surface impact of the differently classified mistral events is presented in terms composites of sea-level pressure (SLP) and precipitation (Fig. 5) and surface heat fluxes (SHF) along with 10-meter winds and 900 hPa equivalent potential temperature (Fig. 6).

The SLP patterns reveal the typical westward tilt with height, and suggest that some clusters favour a phase lock, usually corresponding to the deepest cyclones (e.g., clusters 5, 8, and 14-16). It is probable that each PV cluster is linked to a different

250 stage of the cyclone lifecycle, which can be centered to the east or west of Italy, or even south in the Ionian Sea, with varying depths. The cyclones are closest to the lee of the Alps in clusters 4, 8, 12 and 16, associating the right column in Fig. 3 to the initial stages of cyclogenesis, while the left column (clusters 1, 5, 9 and 13) likely correspond to the termination stage of the cyclone and their easternmost location. The anticyclone extending from the Atlantic is highly variable among the clusters in its strength, thereby affecting the surface pressure gradient, and its spatial extension towards Europe. At times, the high-

255 pressure system dominates the region (e.g., clusters 1, 6, 9, 13), such that a weak cyclone is sufficient for producing the strong mistral winds (see red arrows in Fig. 6). In other cases, the deep Mediterranean cyclone is the dominant feature (e.g., clusters 11, 12, 15, 16).

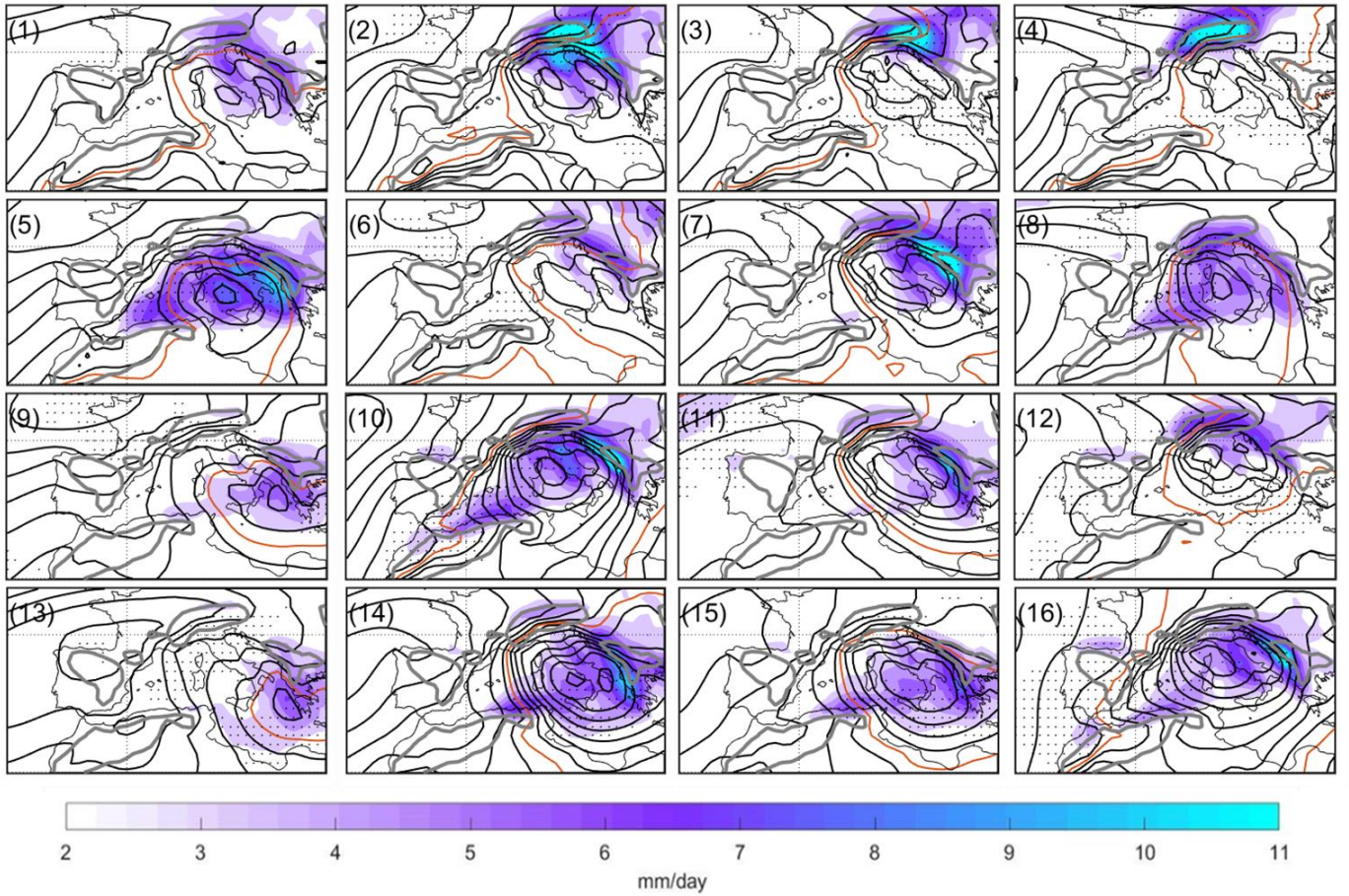



Figure 5: Same as Fig. 3, but for a smaller domain showing SLP (hPa, black contours) at 1-hPa intervals below 1015 hPa (colored line) and 2-hPa intervals above 1015 hPa. Daily accumulated precipitation (mm/day) is shaded. Dotted regions indicate statistical significance >95% for the precipitation composites compared to the average mistral flow.

The mean distribution of precipitation is unique for every cluster, with the location of the precipitation maxima differing among clusters more than across the seasons. In summer, precipitation varies sharply between the eastern and northern Alps (i.e., clusters 1-4), while in the winter, it is differently distributed between the Dolomite and Balkan Mountains, (i.e. clusters 12, 16) and the Alps (8), with notable precipitation occurring along the African shoreline as well (10, 14, 15 and more). Generally, precipitation is distributed along the northern and eastern sides of the cyclone, and roughly correlates with its intensity, which is consistent with previous work (Flaounas et al. 2015; Raveh-Rubin and Wernli 2015, 2016).

The surface heat flux pattern associated with each cluster is relatively localized. Most clusters exhibit the familiar heat loss hotspot in the GOL, however it can extend to different lengths south into the Mediterranean and is absent from several clusters (specifically, 8, 9, 12, and 13). The Bora winds are active together with the mistral when upper conditions allow for an easterly

flow towards the Adriatic Sea in central-eastern Europe (i.e., clusters 5, 9 13 and 14), generating heat loss hotspots in the Adriatic.

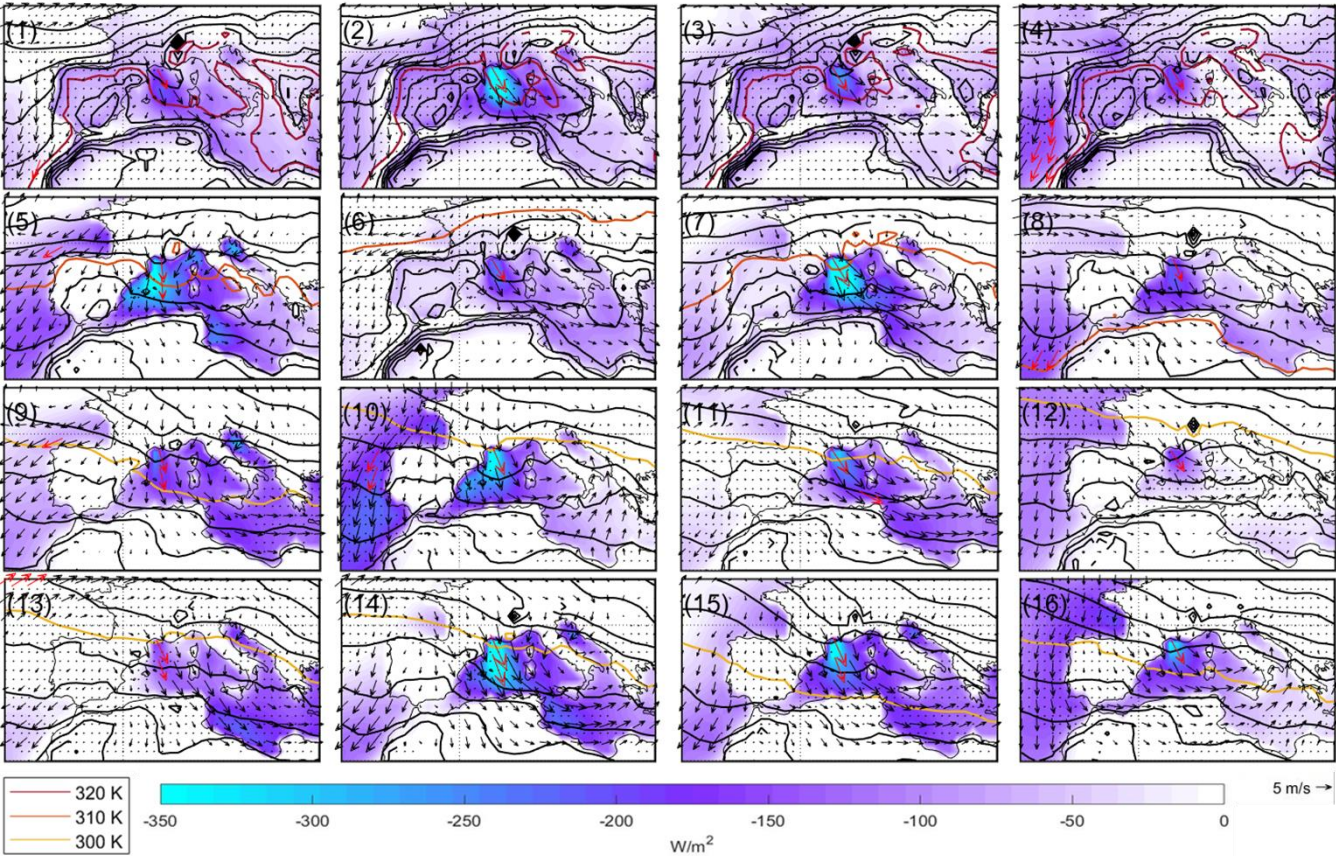



Figure 6: Same as 5 but for surface heat fluxes (sum of sensible and latent, W m-2) in color, equivalent potential temperature at the 900-hPa level in black contours at 2-K intervals and the 300, 310 and 320-K isotherms in color (see legend). Black arrows denote 10-meter wind vectors, where red arrows mark winds above the local 75th percentile. Dotted regions indicate statistical significance >95% for the surface heat fluxes composites compared to the average mistral flow.

The direction and horizontal extent of the mistral wind also differs between the clusters, with clusters 2, 7, 14, and 15 apparently delivering the strongest winds that also extend the furthest into the Mediterranean. The equivalent temperature field illustrates the cold and dry anomalies caused by the mistral. Differences are evident between the clusters, with some displaying a frontal deformation of the isotherms around the Gulf of Genoa. Interestingly, clusters without a marked cold/dry anomaly correspond to clusters with only weak fluxes, despite the strong surface winds (clusters 8, 9, 13).

Note that the statistical significance presented in Figs. 3, 5 and 6 is compared to the average mistral flow, thus highlighting the changes between the different mistral clusters, rather than deviations from climatology. E.g. referring to Fig. 6, the SHF

maxima in the GOL are often not statistically significant as it is a standard mistral feature (clusters 6, 11 and 16). However, statistical significance arises when the signal extends further south (clusters 2, 7, and 14), west (clusters 5 and 10) or if it is exceptionally weak (clusters 1, 12, 13). The clusters main characteristics are summarized in Table 1, along with the attribution of known mistral cases and/or Mediterranean cyclones to the PV-based clusters.

Cluster Name	PV feature location	Trough axis	Surface cyclone location and depth	Precipitation	SHF	Known example Storm/ Study, date
1. Summer cut-off	Northern Italy	N/NE	Tyrrhenian, shallow	E Alps-Dinaric, weak	Off-shore GOL, mild	Medicane Querida, 26-27/9/2006, Moscatello et al.(2008).
2. Summer streamer	Northern Europe	N	Ligurian/Adriatic, deep	E Alps-Dinaric, strong	GOL, strong	Medicane Qendressa, 6/11/2014, Bouin and Lebeaupin-Brossier (2020); Cioni et al. (2018).
3. Scandinavian trough	Scandinavia	N/NE	Adriatic, medium	E Alps, strong	GOL, weak	
4. North Sea trough	North sea	N/NW	Adriatic, shallow	NW Alps, strong	GOL, mild	
5. Genoa cut-off	Genoa	NE	Tyrrhenian, deep	Dinaric, strong	GOL, strong	Medicane Rolf, 7-8/11/2011, Dafis et al. (2018); Ricchi et al. (2017).
6. Autumn/Spring trough	NE Europe	N/NE	Ionian, shallow	Dinaric, weak	GOL, mild	
7. Autumn/Spring streamer	Scandinavia	N	Adriatic, deep	N Dinaric, strong	GOL, extreme	MAP IOP 15, 06/11/1999, Guenard et al. (2005).
8. CWB	Genoa	N/NW	Ligurian, deep	Broad distribution, weak	GOL, weak	Dust storm (Greece), 11-12/4/2005, Kaskaoutis et al. (2008).
9. Strong AWB	Eastern Europe	E	Ionian, deep	Dinaric, weak	GOL + Adriatic, weak	Medicane 15/12/2005, Dafis et al. (2020). 24/3/1998. Flamant (2003).
10. Mild AWB	Central and western Europe	NE	Ligurian, deep	Dinaric, strong	GOL, strong	Medicane 16/11/2007, Dafis et al. (2020). 05/03/1982, Speranza et al. (1985).
11. Relieved trough	Scandinavia-Dinaric	N	Adriatic, deep	Dinaric, strong	GOL, mild	15/3/2013, Drobinski et al. (2017).
12. Blocked trough	Scandinavia-Alps	NE	Ligurian, deep	Alps, weak	GOL, weak	4/11/1990, Drobinski et al. (2001a).
13. weak ACWB	Greece	E	Ionian, shallow	Ionian, weak	Ionian, mild	Medicane Celeno, 15/1/1996, Pytharoulis et al. (1999).
14.Sharp winter streamer	Central/Eastern Europe	N	Tyrrhenian/Ionian, deep	Dinaric, strong	GOL, extreme	13/2/2013, Drobinski et al. (2017).
15. Relieved broad trough	N. Italy	N/NE	Tyrrhenian, deep	Ionian, strong	GOL, strong	Medicane Leocossia, 25/1/1982, Ernst and Matson (1983).
16. Blocked broad trough	Central Europe	NE	Ligurian, deep	Dinaric, strong	GOL, mild	17/02/1992, Schott (1996).

Table 1 Summary of cluster names and main features. AWB/CWB refer to anticyclonic/cyclonic Rossby wave breaking, respectively.

3.4 Time Evolution of Mistral Events

Frequent cluster transitions and cluster persistence can be visualized using the transition probability matrix (TPM, Fig. 7). Column 0 shows the likelihood of a mistral event to end at any cluster with no further transition. The strong amplitude along the shifted-main diagonal suggests that every cluster has a tendency to sustain itself, at a varying likelihood. We interpret this feature as a persistence demonstration by the algorithm, as the time-scale for the evolution and migration of the PV structures driving the mistral events are often longer than a day. The fact that the PV-based SOM is able to consistently classify consequent events of slow-developing waves under the same classification reinforces the robustness of the method, given proper SOM constraints (such as the number of clusters), as only significant (by SOM interpretation) differences in the daily fields within a mistral event can force a transition. After Espinoza et al. (2012), a transition is deemed statistically significant if its frequency exceeds the 90th percentile of the corresponding transition frequency derived from a 1000 random redistributions of the original sequence. We constructed a reference random distribution by considering all mistral days (recall that the clustering is performed only for mistral days and not for all other days). While Huang et al., (2017) set their criteria by the 95th percentile, here the 90th percentile threshold was selected, considering that single-day mistral events represent roughly 50% of mistral events, and are of less interest from the dynamical perspective. Still, post-mistral days are accounted for as eligible transitions in the random distributions, represented by the 0 column in Fig. 7.

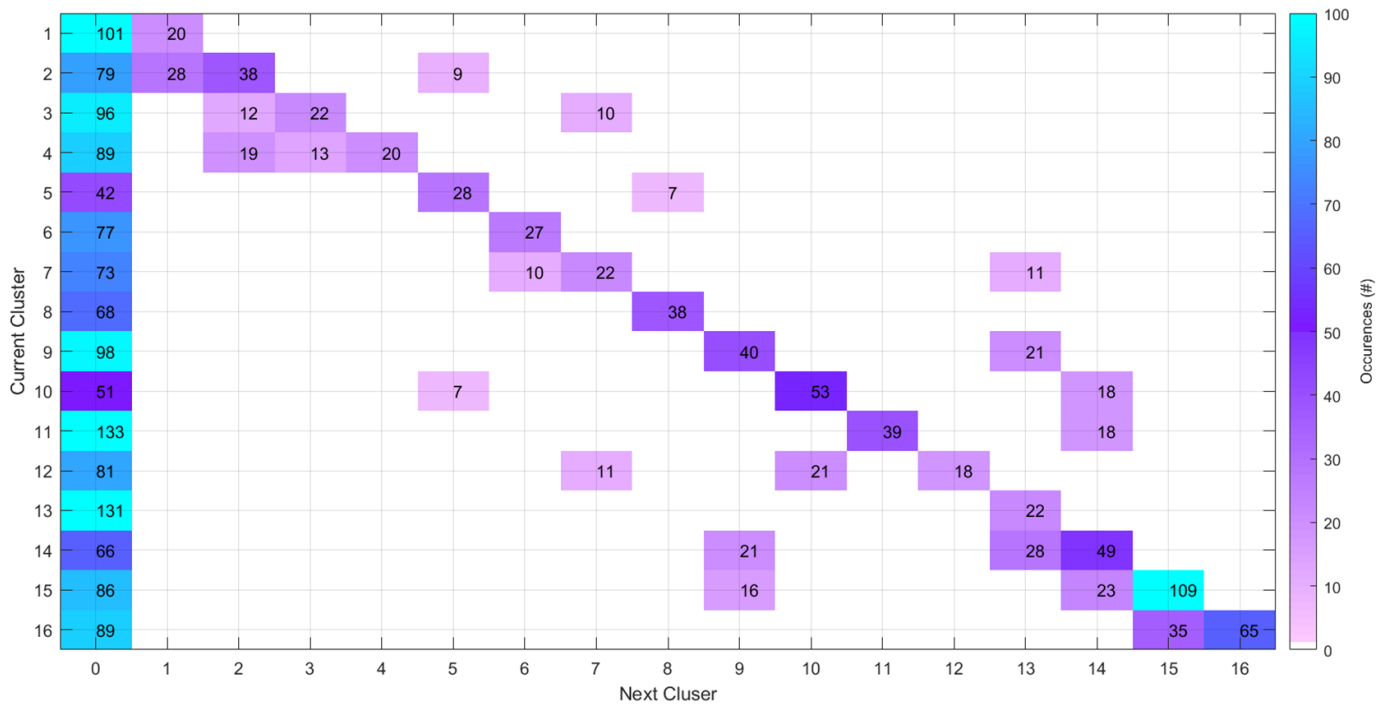



Figure 7: Transition probability matrix, based upon daily cluster transitions from the current cluster to the cluster on the next day (colored by number of days). Total days in each transition is given by the numbers in the corresponding rectangle. The 0 column represents the ending of mistral events, and the diagonal represents the probability to remain in the same cluster for the next day. Only statistically significant (90% confidence level) transitions occurring within mistral events are shown.

Note that the transitions are distributed mostly around the main diagonal and the 0 column, the latter is primarily due to frequent single-day mistral events. Nonetheless, some recurring cluster transitions are showing a considerable amplitude, such as transitions 14→9 and 2→5, and others. These transitions are made clearer when viewed separately for each season (Fig. 8). Very different amplitudes along the main diagonal and the 0 column suggest some clusters are only self-sustaining in certain seasons and are more likely to be the end of a mistral event in the other seasons (for example, cluster 5 in autumn compared to winter-spring). It is clear from the seasonal TPMs that some transitions are absent from certain seasons, whereas others may occur at any month of the year. Studied carefully, these transitions reveal many details about the development of upper level PV anomalies over the Alps. For instance, transition 12→ 7 (amplifying ridge behind streamer) and 10→ 5 (formation of Genoa cutoff) are more abundant in the transition seasons, while transition 14→ 9 (strong anticyclonic Rossby wave breaking, AWB) occurs in any season except the summer, suggesting this transition and several others are more resilient to the changes of the seasons. Some of these transitions, and indeed, some individual clusters, can be directly related to AWB (14, 15 leading to 9, and 13), cyclonic Rossby wave breaking (CWB, cluster 8), a cut-off low migrating into the domain from the north (2→ 1/5), or being cut out of a north-easterly streamer (10→ 5). Other transitions imply an equatorward stretching of a trough (12→

10), or the eastward propagation of a trough (4→ 3, 7→ 6), and so on. These features shed light on recurring wave-evolution processes, and allow one to easily access and investigate a large variety of rare yet dynamically similar events by selecting certain cluster transitions that are representative of certain wave lifecycles.

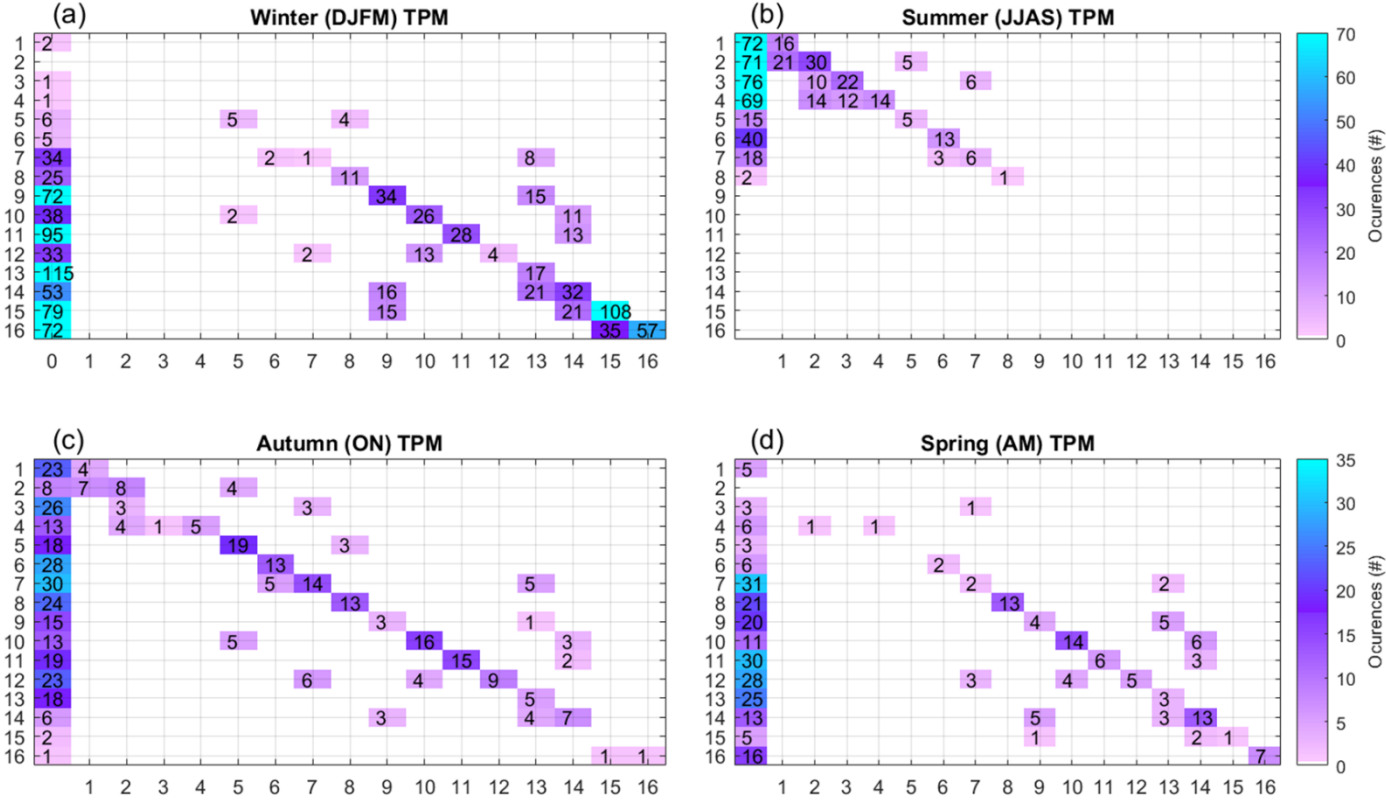



Figure 8: The same as 7, separating by the different seasons. Note the 2-month periods in autumn (October-November) and spring (April-May), compared to 4 months in winter (December-March) and summer (June-September), and the different color scales, accordingly.

Considering that many mistral events last more than a couple of days, it is insightful to examine the first event transitions, rather than all individual transitions without a time trace (as in Figs. 7 and 8). Therefore, we first identify the mistral initiation days, with respect to each cluster, and then display the first two transitions for every mistral event in the group. Despite containing tens of individual transition sequences at each mistral group, several dominating first transition sequences emerge (i.e., same sequence of days 1-3). Recurring transitions exhibit a strong seasonal dependency (Fig. 9). Qualitatively, the transition sequences seem to "push" the system towards seasonally dependent preferable clusters. Thus, an event initiating with an off-season cluster, say cluster 6 in winter, will be shifted towards the proper winter cluster 13, suggesting the broadening of the northerly streamer into a mature trough. Another example can be seen in the initial clusters 4 and 8, where

345 the first transitions are seasonally dependent. In winter time, mistral events are dominated by the transitions between clusters 9-11 and 14-16, with preferable paths illustrated by the thick blue lines in the corresponding panels. This view also highlights the directionality of the transitions. For instance, note that the clusters with largest numbers of initiated events that last 3 or more days are the blocked clusters 12 and 16, and that the "relieved" clusters rarely jump back to a blocked cluster, i.e., the transitions 11/15→12 are scarce. Considering the cluster configuration displayed in Fig. 3, the general direction of flow within

350 a mistral event is from right to left and from top to bottom, diverging mostly from clusters 8 and 12 and converging towards clusters 9, 14 and 15 (see Fig. S1 in supplementary material).

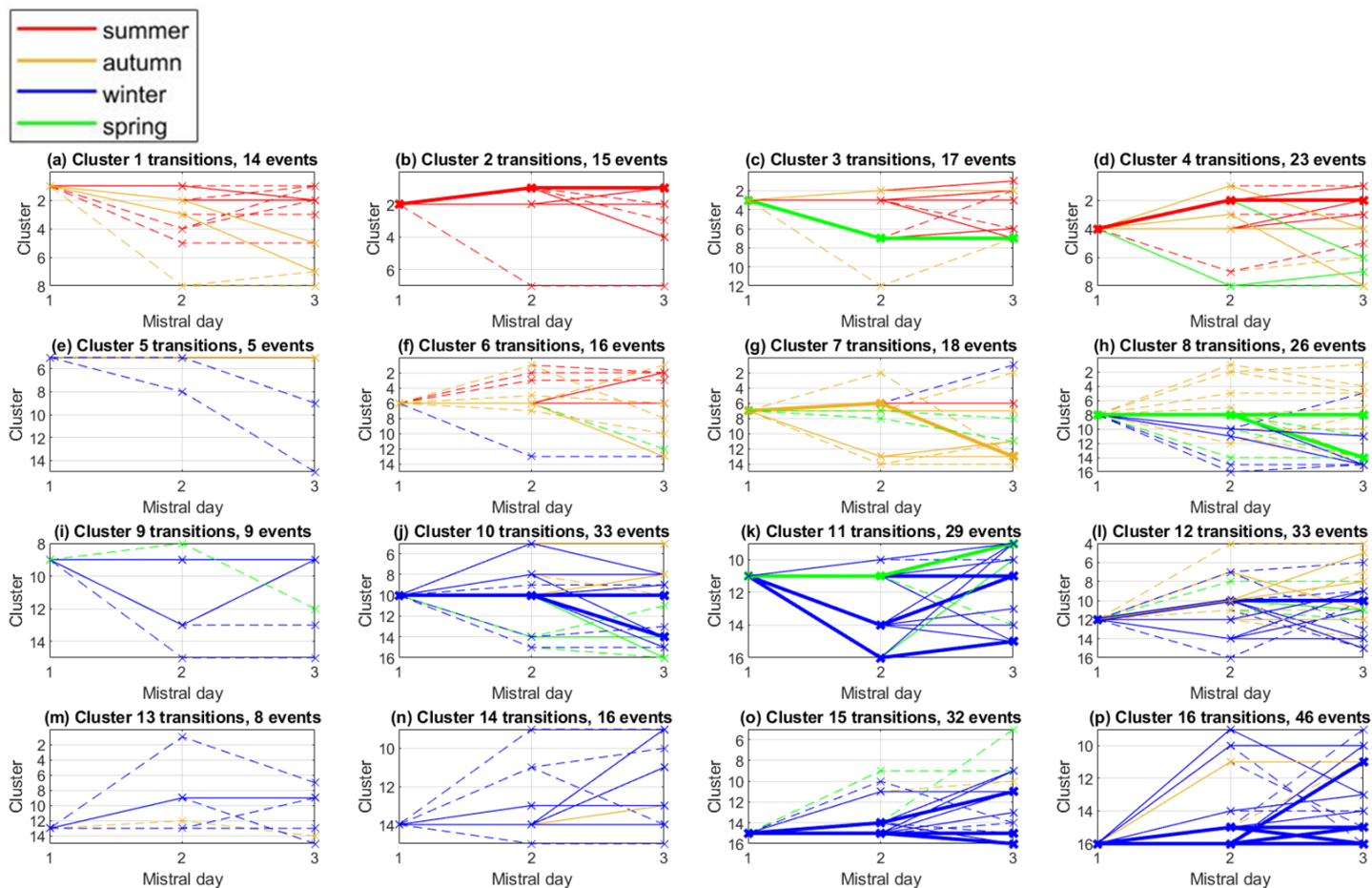



Figure 9: Transition sequences (days 1-3) for mistral events initiating with each cluster at mistral day 1 and lasting at least 3 days,

355 **at every season (using the same 2 and 4-months seasonal definitions as in Fig. 8). The x axis represents days passed from mistral initiation. Solid lines indicate transition sequences (days 1-3) that repeat 2-4 times throughout the time-period. Thick lines emphasize sequences occurring 5-10 times while dashed lines denote single-occurring sequences.**

Overall, the clustering analysis identified robust and distinct isentropic PV patterns, and the transition analysis offers additional perspective on the evolution of these upper level PV structures as they interact with the Alpine ridges during mistral events.

360 These TPMs, combined with the surface impact of each identified PV cluster, can potentially be utilized to improve weather predictions of both the mistral wind and the Genoa cyclone, and their impacts. In the following we examine three individual mistral events, and the ability of the cluster mean to represent their evolution in a meaningful way.

3.5 Three Illustrative Mistral Events

The qualitative view of the algorithm performance is presented by a comparison between the classified clusters and the

365 individual members for three mistral events. The large scale structure of the PV is, however, well captured by the SOM, as illustrated in the following cases.

3.5.1   Anticyclonic Wave Breaking Mistral, 16-20 January 1987

This winter mistral event initiated with a broad, mountain-passed winter trough that engulfed most of continental Europe.

370 Strong surface heat fluxes commenced on day 1 as the trough stretched over the Alps into day 2, with the transition $15 \rightarrow 14$, accompanied by a deepening of the cyclone, and intensified precipitation and GOL heat flux, as implied by the intense mistral events generally classified in cluster 14. Another transition back to 15, from day 2 to 3, implies another broadening of the trough, with a slight weakening of the mistral and persistence in cluster 15, a common scenario (Figs. 8 and 9). The wave is then deflected to the east of the Alps and breaks anti-cyclonically, as captured by the transition $15 \rightarrow 9$. On its last day at cluster

375 9, the cyclone is weakened, along with reduced associated surface heat flux and precipitation, as is indeed common for cluster 9.

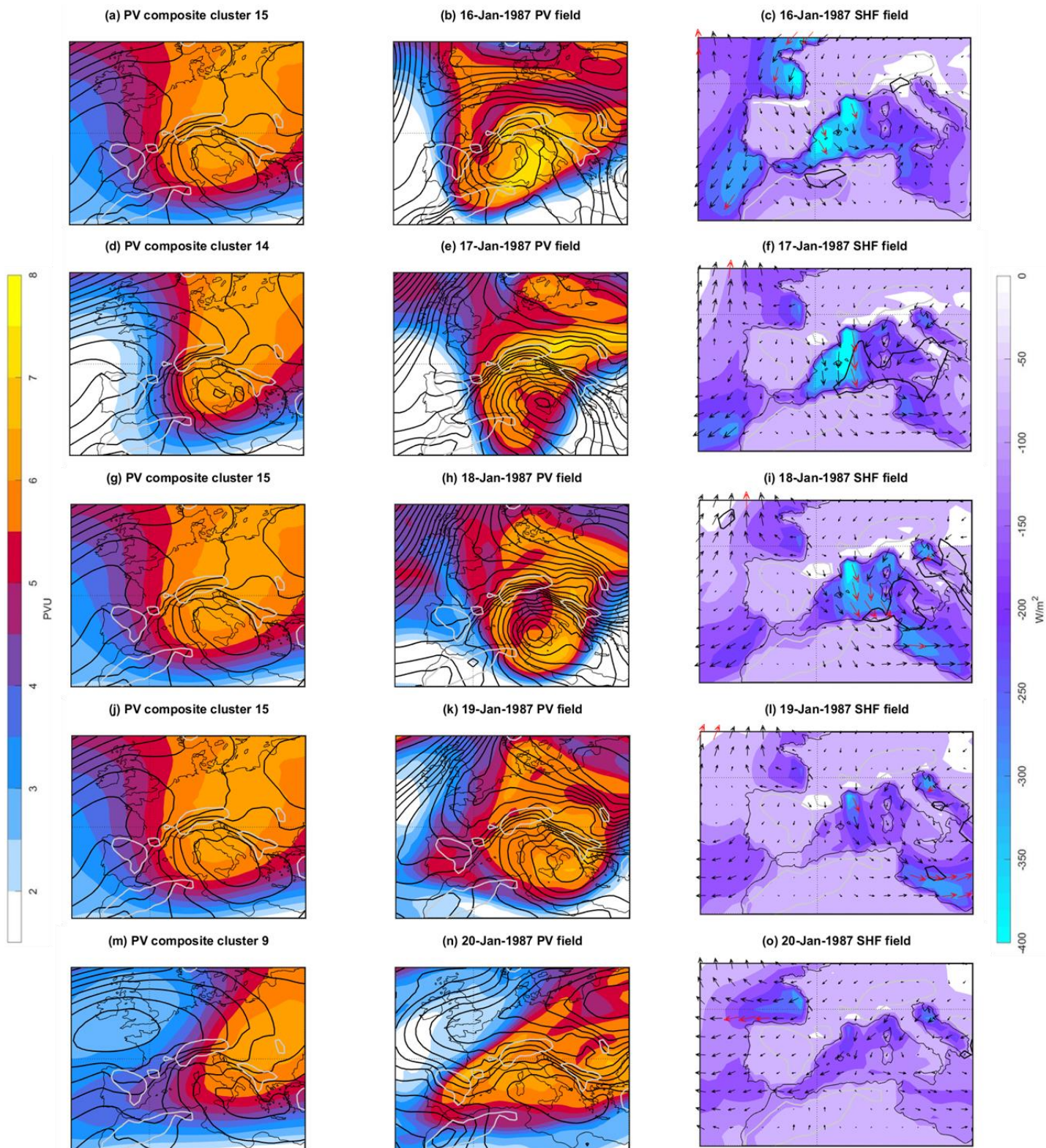



Figure 10. An illustrative AWB event identified by the transition 15→9. Shown are vertically averaged (320-340 K) daily PV distributions and SLP (black contours) throughout a mistral event (middle column), alongside their corresponding cluster composites (left column). Surface impact is shown on the right column in terms of the sum of sensible and latent heat fluxes (W m-2, shaded), 10-m wind vectors (the red vectors represent magnitude in the upper quartile) and daily accumulated precipitation (black contours, 10 and 30 mm/day)

3.5.2   Cyclonic Wave Breaking Mistral, 8-12 April 2005

This spring mistral event begins with a blocked trough and a weak cyclone in the lee of the Alps. The trough is stretched into a thin NE streamer, captured by transition 12→10, representing a common transition. Upon this transition which marks a first AWB, the SHF intensify dramatically, along with precipitation in the Adriatic region. The trough then further stretches and breaks cyclonically (10→8) to form a cut-off, as the cyclone attains a deep symmetric structure (Tous and Romero, 2013). Note that the streamer is channelled above the Rhône Valley just before breaking, illustrating the wrapping up of PV banners generated in the mistral region (Aebischer and Schär, 1998). The classification of 11/4/2005 to cluster 8, capturing the CWB pattern despite the PV streamer extending to the north-east rather than north-west as suggested by the composite, emphasizes the ability of the SOM to identify distinct geometrical features, rather than only geographical ones.

(a) PV composite cluster 12   (b) 08-Apr-2005 PV field   (c) 08-Apr-2005 SHF field

(d) PV composite cluster 10   (e) 09-Apr-2005 PV field   (f) 09-Apr-2005 SHF field

(g) PV composite cluster 10   (h) 10-Apr-2005 PV field   (i) 10-Apr-2005 SHF field

(j) PV composite cluster 8   (k) 11-Apr-2005 PV field   (l) 11-Apr-2005 SHF field

(m) PV composite cluster 8   (n) 12-Apr-2005 PV field   (o) 12-Apr-2005 SHF field

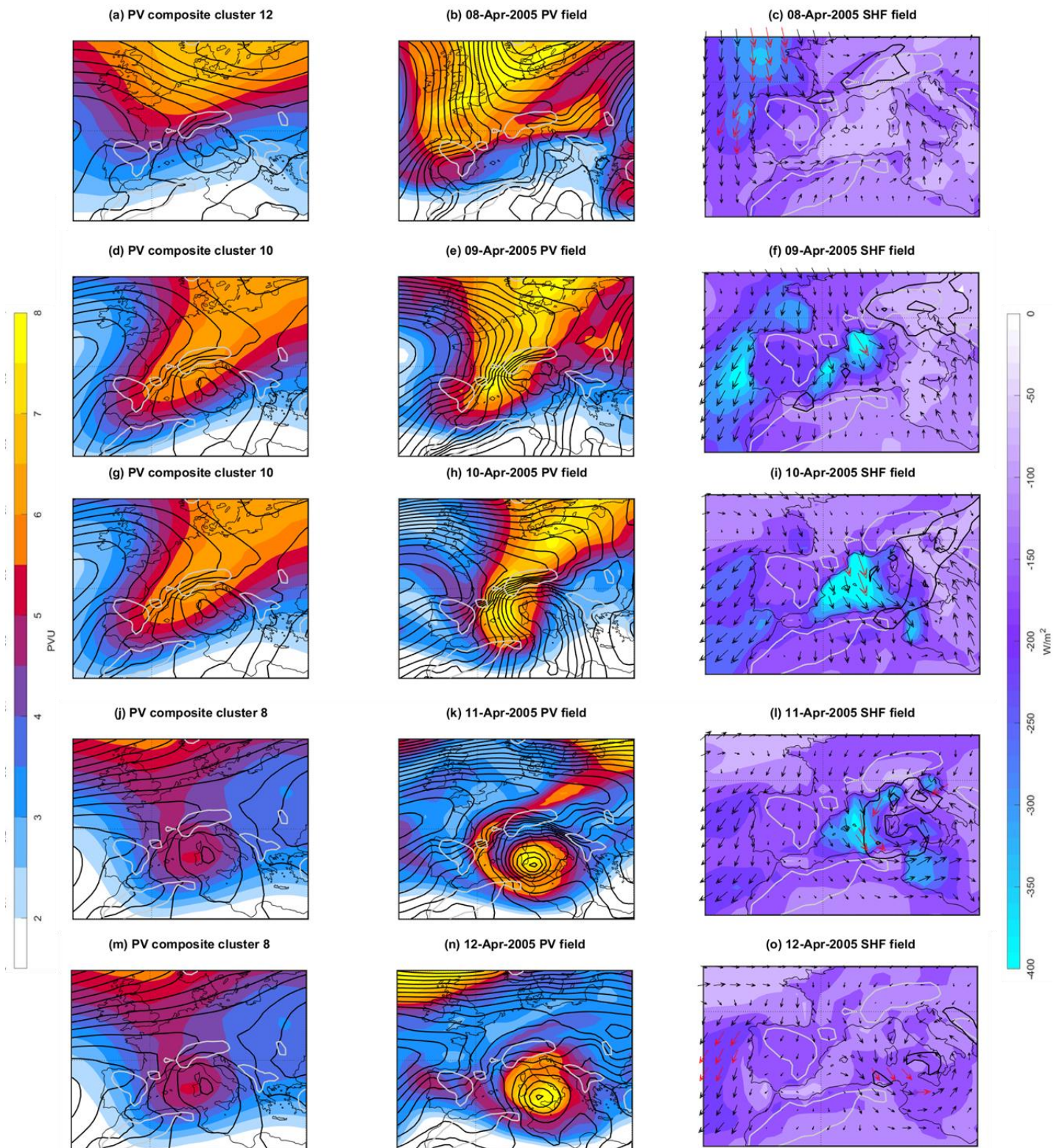



Figure 11: As in 10, but for a CWB event identified by the transition 10→8.

395

3.5.3 Cut-off Low Mistral, 25-31 August 1995

The summer mistral event is exceptionally long for the season. It begins with the weakest upper-level forcing recorded by the present analysis, with a 2→1 transition, demonstrating the cut-off of a 3-PVU northerly streamer over the Alps. The propagation of a second wave into the domain is identified as transition 1→3, and this second wave is again stretched

400   southward to form a summer streamer (3→2). This transition is accompanied by an intensification of the mistral, the deepening of a primary cyclone south-east of the Baltic Sea, and a lee cyclone in the Adriatic Sea. The streamer then tilts and breaks to form a cut-off low (2→5), with weakening of the mistral intensity. A noticeable characteristic is the persistent precipitation response of the days corresponding to cluster 2, centered just between the eastern Alps and northern Dolomites, as suggested by the precipitation composites (Fig. 5, cluster 2).

405

(a) PV composite cluster 2 (b) 25-Aug-1995 PV field (c) 25-Aug-1995 SHF field

(d) PV composite cluster 1 (e) 26-Aug-1995 PV field (f) 26-Aug-1995 SHF field

(g) PV composite cluster 3 (h) 27-Aug-1995 PV field (i) 27-Aug-1995 SHF field

(j) PV composite cluster 3 (k) 28-Aug-1995 PV field (l) 28-Aug-1995 SHF field

(m) PV composite cluster 2 (n) 29-Aug-1995 PV field (o) 29-Aug-1995 SHF field

(p) PV composite cluster 2 (q) 30-Aug-1995 PV field (r) 30-Aug-1995 SHF field

(s) PV composite cluster 5 (t) 31-Aug-1995 PV field (u) 31-Aug-1995 SHF field

Figure 12: As in 10, but for a summer cut-off formation event identified by the transitions 2→1 and 2→5

410   **4. Summary and Concluding Remarks**

This study examines systematically the large- and synoptic-scale drivers of the mistral wind during 1981-2016, by classifying the isentropic PV during mistrals. Mistral days are first identified objectively in a climatological dataset based on 20-km resolution WRF-ORCHIDEE simulation forced by ERA Interim, yielding 2734 mistral days, distributed among 1360 mistral events. Mistral occurs throughout the year, but occurs more often in winter, with more multi-day events, compared to summer.

415 A SOM clustering analysis then classified the PV distributions during mistral days, providing insight on the large-scale driving mechanisms of the mistral wind, and further served as a tool to examine different types of surface impact signatures. Referring to the questions posed in the introduction, here we summarize the main findings (see also Table 1).

(i) During objectively-identified mistral days, the daily mean, vertically averaged (320-340 K) isentropic PV distributions are classified into 16 distinct clusters according to their geometrical shapes. The emerging features vary among amplified Rossby

420 wave patterns ahead of an Atlantic ridge. Features include troughs, PV streamers with variable orientations indicative of cyclonic or anticyclonic Rossby-wave breaking CWB/AWB, and cut-off lows.

(ii) The clustering approach distinguishes between the seasons, including the transition seasons, suggesting that a limited range of PV features prevail in each season. In summer, clusters 1-4 suggest the dominance of either a cutoff, thin streamer or a trough over Scandinavia or the North Sea. In the extended winter season, several AWB scenarios prevail, as well as broad,

425 deep, southward-intruding troughs. In the transition seasons, troughs, streamers cut-off lows and CWB are the dominant features. Note that PV features (e.g., streamer or cutoff-low) exhibit different mean position, shape and magnitude across different seasons (see Table 1 for a detailed summary).

(iii) Each cluster indeed reveals a unique signature in terms of surface weather impact, i.e., surface cyclone and associated circulation, winds, sensible and latent heat fluxes, and precipitation. The Bora wind regime, for example, appears to co-occur

430 with the mistral within clusters 9 and 13, which is expected from a relatively eastern PV anomaly. The latent and sensible heat loss hotspots centered over the GOL strongly vary among the clusters, as does the wind speed and the extension of the mistral wind offshore. The highly variable mean intensity, size and location of the cyclones associated with the different clusters clearly demonstrate the impact of the upper tropospheric PV distribution on the surface pressure, with most cyclones under a phase lock with a parent trough, implying the lee-cyclogenesis effect. The deepest cyclones and strongest surface heat fluxes

435 occur with PV streamers indicative of AWB/CWB, a Genoa cutoff or a deep winter trough. The mean precipitation response strongly depends on the cluster and season, generally being correlated to the mean cyclone intensity and peaking to the north and northeast of the cyclone. At times, precipitation occurs also downstream of the mistral outflow along the north African

coast, particularly in clusters with strong surface fluxes. The latter is consistent with Rainaud et al. (2016) and Berthou et al. (2018) who demonstrate the remoistening of the dry mistral airmass and its downstream precipitation impact.

440 (iv) The evolution of the flow types during multi-day mistral events is examined through the analysis of transitions among clusters, while searching for recurring patterns. The dominant direction of transitions corresponds to the eastward drift of the waves (7→6, 16→15), while other transitions imply the stretching of the PV feature as it approaches the Alps (3→2/7, 12→10) or the formation of a cut-off (10→5). Long mistral events (>2 days) initiate mostly with the blocked clusters 12, 16 and 10, and tend to end with AWB/CWB events, i.e., cluster 8, 9 and 13, or with a cutoff (clusters 1, 5). The prior is mainly evident

445 by the non-uniform distribution of the mistral initiation days between clusters (Fig. 9) and the latter by the lack of transitions from these clusters (1,5,8,9,13) to other ones. Specifically, clusters 1, 8 and 13 do not transit to any other cluster (within the 90% confidence level, see Fig. 7), with clusters 5 and 9 only transiting to 8 and 13, respectively. Cluster 14 represents the strongest mistral events in terms of heat fluxes and wind speeds, while clusters 12 and 13 correspond to the weakest. Cluster 15 is the most persistent and abundant, whereas cluster 12 is the most transient.

450 The SOM performance is evaluated quantitatively by standard deviations within the cluster (Fig. A2) and qualitatively by comparing individual members to their corresponding cluster composites. It is evident that even highly non-linear Rossby wave breaking processes are well represented by their corresponding cluster composite means, while the inevitable case-to-case variability is manifested mainly along the high PV gradients, or in terms of the magnitude or location of the PV anomaly of cutoffs.

455 One primary conclusion implied by the present classification arises when attempting to examine the present clusters under the cyclone types as defined by Tibaldi and Buzzi (1978). According to the classic description of lee-cyclogenesis, the initial phase of the lee-cyclone is attributed mostly to a-geostrophic adjustment process, induced by the mountain, generating the cyclone rapid deepening rates in a still relatively barotropic environment. In the latter phase, thermal gradients are enhanced due to the advective nature of the Rossby wave, and baroclinic instability becomes the dominant mechanism influencing the cyclone, as

460 the low-level cold-pool air mass finally flows across the topography. Carefully examined, some clusters can be related to phase 1 or 2 of the lee cyclone, according to their representative thermal gradients, and PV maxima location relative to the Alps. The so called "blocked" modes, i.e. 12 and 16, appear to correspond to a much weaker mistral event in terms of wind speeds and GOL surface heat fluxes, when compared to their "relieved" companions, 11 and 15. This suggests that the maximum mistral related surface heat fluxes are attributed to the baroclinic phase of the cyclone, rather than the triggering phase. While further

465 confirmation of this perspective is required, it settles with the notion of the mistral as "one of many strong winds that manifest the penetration of cold air into the Mediterranean from the north" (Scorer et al., 1952), in the sense that a strong mistral indeed enables a significant penetration of polar air-masses into the Mediterranean, leading to the massive heat loss at the GOL occurring at post-blocked stages of the mistral event, rather than in the initial stages.

This aspect of the mistral and Genoa cyclogenesis as a synoptically-controlled phenomenon has the potential to improve
470 predictions of the mistral wind and Genoa cyclogenesis events, and deepen the understanding of synoptic-scale PV interactions
with topography. Furthermore, the systematic classification offers new insight into the variability of the mistral impact on air-
sea interaction in the region, with direct implications for understanding the seasonal build up and onset of deep convection in
the water column in autumn-winter.

Short summary

475 Mistral wind is a renowned phenomenon in the Mediterranean, yet, its large-scale controlling mechanisms have not been
systematically mapped. Here, using a new mistral database for 1981-2016, the upper-tropospheric flow patterns are classified
by a self-organizing map algorithm, resulting in 16 distinct patterns related to Rossby wave life cycles. Each pattern has unique
surface impact, having implications to understanding mistral predictability, air-sea interaction and their future projections.

Author contribution

480 YG conducted the analysis and wrote the initial manuscript draft and SRR supervised the research. SRR and PD conceptualized
the research and acquired funding. VS produced the cyclone mask dataset, while all other co-authors collaborated on the
interpretation of the results and reviewing of the manuscript.

Acknowledgements

This work is a contribution to the HyMeX programme (www.hymex.org) (Drobinski et al., 2014) and the Med-CORDEX
485 initiative (www.medcordex.eu) (Ruti et al., 2016). This article is based upon work from COST Action CA19109 "European
network for Mediterranean cyclones in weather and climate" supported by COST (European Cooperation in Science and
Technology, www.cost.eu). The Israel Meteorological Service and ECMWF are acknowledged for providing access to ERA-
Interim data. This work is supported by a research grant from the Benoziyo Endowment Fund for the Advancement of Science
and from the Weizmann - CNRS Collaboration Program. This work was supported by École Polytechnique, Institut
490 Polytechnique de Paris through the visiting scientist program. We would like to thank the two anonymous reviewers for their
constructive comments and suggestions.

Appendix:

A.1 The 4x4 hexagonal SOM map

The SOM neighbor distance is a measure of similarity between neighboring SOM clusters, specifically the mean squared error

495   (MSE) between the SOM weights corresponding to the Neuron that represents each cluster. The 4x4 hexagonal SOM map and
the neighbor distances are presented in Fig. A1. The smaller the neighbor distance the more similar the neighboring clusters.
This measure does not involve the frequency of transitions between clusters, discussed in section 3.4. The similarity map
exhibits a dark line of reduced similarity crossing the grid, defining the seasonal separation discussed in section 3.2. With that
said, there is a physical reason for similar clusters to appear consecutively, and a transition between two very different clusters,

500   e.g. 8→3, would seem very unlikely. Thus, the transition-season clusters are expected to provide a corridor through which
summer clusters can develop into winter clusters, or vice versa (especially cluster 7).

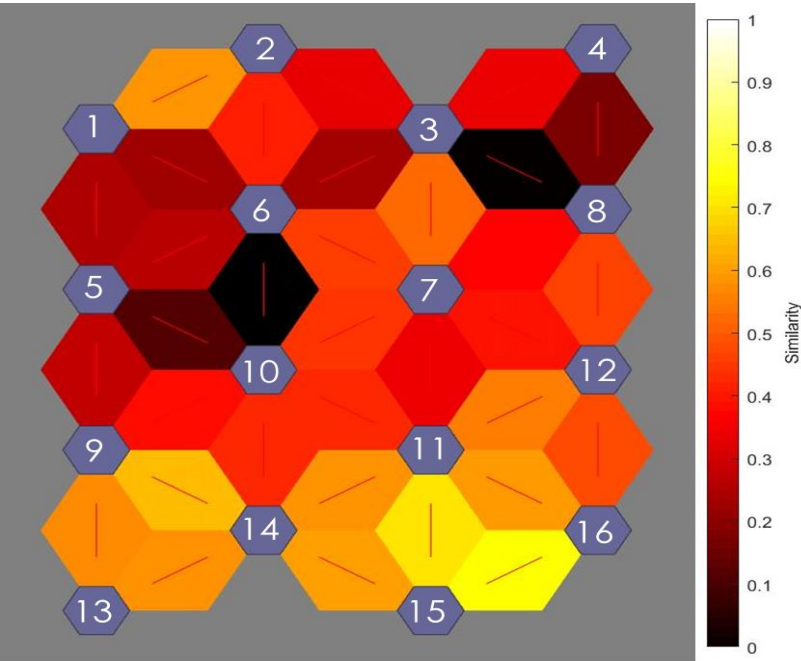



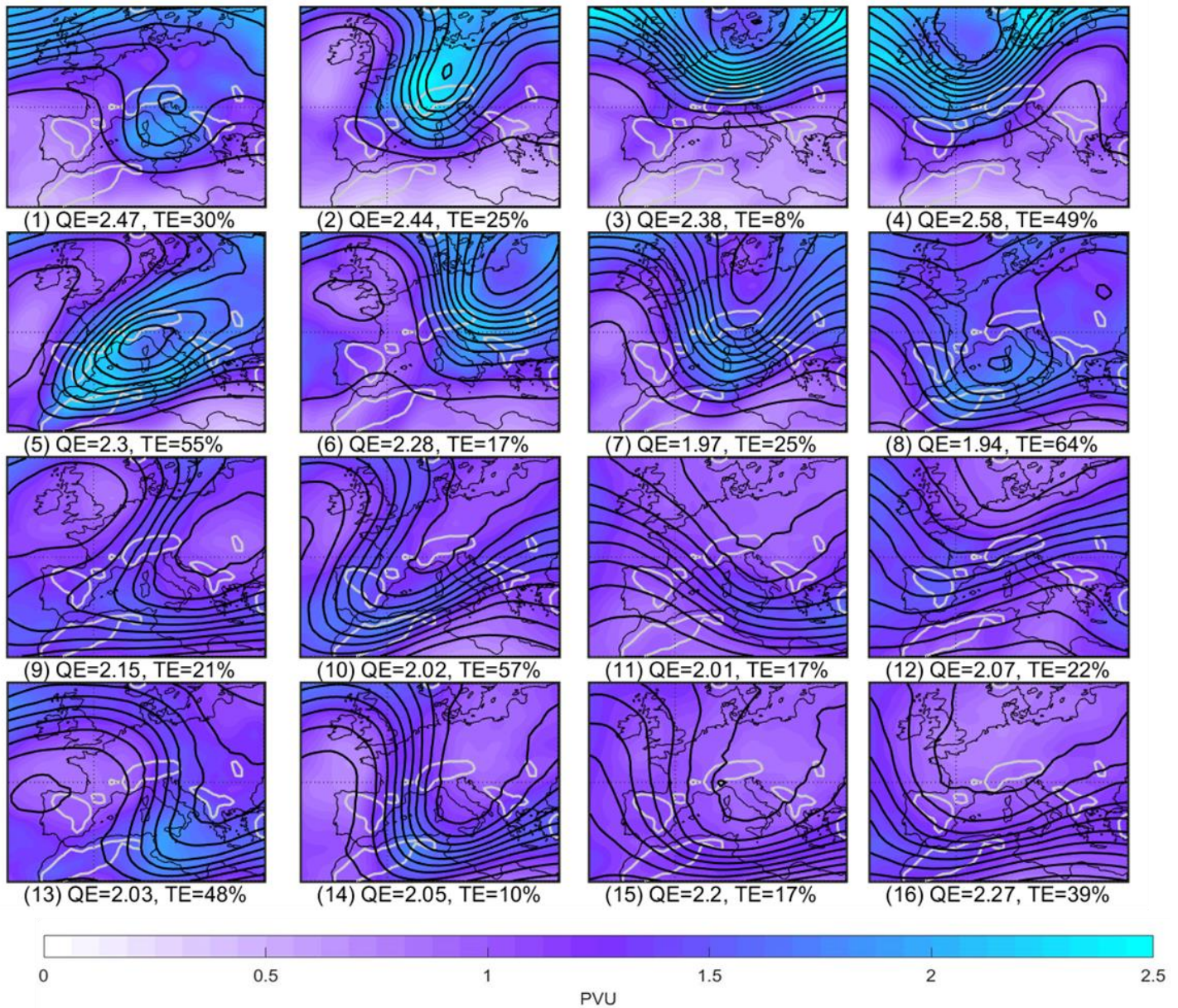
Figure A1: The 4x4 hexagonal SOM map with the numbered clusters and neighbor distances, or similarity between clusters (dimensionless) in colored connections. Brighter colors represent short distances, or more similar clusters, and dark colors indicate

505   **large distances, or dissimilarities.**

A.2 Intra-cluster variability

The variability among the members for each cluster can be quantified with the standard deviation (STD) map of the PV fields
(Fig. A2). The resulting patterns demonstrate the uncertainty in the magnitude and exact location of the identified PV feature.
As such, the largest STD is either along the boundary of the streamer (e.g., clusters 3, 6, 10, 14) or within a cut-off (1, 5, 8). It

510   is apparent that some clusters mean signal is reasonably well aligned with the individual members, while others exhibit larger
inner variance. For example, most members of cluster 9 fit right in the composite, as opposed to cluster 1, where the cut-off

magnitude is likely under-estimated by the averaged field, due to variations in location. The STD also reveals further details on some clusters, such as the hidden PV streamers apparently related to clusters 6 (towards the Atlas Mountains) and 3 (towards the Pyrenees). The mean STD for most clusters is on the order of 15%, however the major correspondence between the composite and its members is along the maximum PV gradient bands, and the quantitative inaccuracy is compensated by the qualitative description of the flow (i.e., Figs. 10-12). Compared against the 95% statistical significance (Fig. 3), focusing on cluster 5, the low STD region in the middle of the cut-off is translated into statistically significant signal, while its surrounding peak in STD is translated into non-significant regions, again demonstrating the variability regarding the extent of the identified cutoff. Similar links between the STD and statistical significance are evident in clusters 2, 7 and more. Quantization and topographical errors (QE and TE, respectively) are evaluated after Kiviluoto (1996) and are presented per cluster and averaged for the entire network in Fig. A2. QE is the average distance (in terms of the SOM optimization function, in this case MSE) between samples and their corresponding cluster. TE measures the continuity of the SOM space by evaluating the proximity of the first and second most important clusters, for each input sample. Ideally, both QE and TE should be minimal. While QE generally behaves similarly to the standard deviation within the clusters, large TEs usually imply a folding of the SOM space due to a hidden dimension of variability. This dimension of variability generates distant, yet similar clusters, as it is secondary to the main dimensions that make up the SOM space. The algorithm then attempts to reduce the distance along this extra dimension by folding the SOM space. In the present case, this hidden dimension appears to correspond to the dis-continuity and deformation of the PV feature, i.e., the extent of separation between the PV streamer and its parent reservoir and the amount of stretching of the PV streamer. This dimension of variability is secondary to both the meridional and zonal dimensions of variability. Thus, the largest TE is attributed to the cut-off clusters 5 and 8, suggesting a fold of the SOM space bridging the apparent distance between the two objectively similar clusters. The heavily-deformed cluster 10 exhibits a considerably large TE, possibly also due to the mentioned fold in the SOM-space.

535   **Figure A2: Standard deviation (STD) of isentropic PV within each cluster. The quantization and topographic errors (QE in PVU and TE in %, respectively) are noted at the bottom of each panel, following the cluster number. The average QE and TE for the network are 2.2 PVU and 30%, respectively.**

References

Aebischer, U., and Schär C.: Low-level potential vorticity and cyclogenesis to the lee of the Alps. J Atmos. Sci. 55.2: 186-207. https://doi.org/10.1175/1520-0469(1998)055<0186:LLPVAC>2.0.CO;2., 1998.

Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S., Béranger, K., and Lebeaupin-Brossier, C.: Prior history of mistral and tramontane winds modulates heavy precipitation events in southern France. Tellus A 66.1: 24064. https://doi.org/10.3402/tellusa.v66.24064, 2014.

Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S. Béranger, K., and Brossier, C. L.: Lagged effects of the Mistral wind on heavy precipitation through ocean-atmosphere coupling in the region of Valencia (Spain). Clim. Dynam., 51(3), 969-983.   https://doi.org/10.1007/s00382-016-3153-0, 2018.

Bouin, M.N. & Lebeaupin-Brossier, C.: Surface processes in the 7 November 2014 medicane from air-sea coupled high-resolution numerical modelling. Atmos. Chem. Phys., 20, 6861-6881, 2020.

Burlando, M.: The synoptic-scale surface wind climate regimes of the Mediterranean Sea according to the cluster analysis of ERA-40 wind fields. Theor. Appl. Climatol. 96, 69–83. https://doi.org/10.1007/s00704-008-0033-5, 2009.

Buzzi, A., Davolio, S. and Fantini, M.: Cyclogenesis in the lee of the Alps: a review of theories. B. of Atmos. Sci.& Technol. https://doi.org/10.1007/s42865-020-00021-6, 2020.

Buzzi, A., M. D'isidoro, and. Davolio. S.: "A case-study of an orographic cyclone south of the Alps during the MAP SOP." Q. J. Roy. Meteorol. Soc. 129.591: 1795-1818. https://doi.org/10.1256/qj.02.112, 2003.

Buzzi, A., and Speranza, A.: A theory of deep cyclogenesis in the lee of the Alps. Part II: Effects of finite topographic slope and height. J. Atmos. Sci., 43(23), 2826-2837. https://doi.org/10.1175/1520-0469(1986)043<2826:ATODCI>2.0.CO;2, 1986.

Čampa, J., and Wernli, H.: A PV perspective on the vertical structure of mature midlatitude cyclones in the Northern Hemisphere. J. Atmos. Sci, 69(2), 725-740. https://doi.org/10.1175/JAS-D-11-050.1, 2012.

Campins, J., Genovés, A., Picornell, M. A., and Jansà, A.: Climatology of Mediterranean cyclones using the ERA-40 dataset. Int. J. Climatol., 31(11), 1596-1614. https://doi.org/10.1002/joc.2183, 2011.

Cassano, E. N., Lynch, A. H., Cassano, J. J., and Koslow, M. R.: Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA. Clim. Res., 30(2), 83-97. httpd://doi.org/ 10.3354/cr030083, 2006.

575 Cioni, G., Cerrai, D., Klocke, D.: Investigating the predictability of a Mediterranean tropical-like cyclone using a storm-resolving model. Q. J. Roy. Meteorol. Soc., 144: 1598– 1610. https://doi.org/10.1002/qj.3322, 2018.

Dafis, S., Rysman, J. F., Claud, C., Flaounas, E.: Remote sensing of deep convection within a tropical-like cyclone over the Mediterranean Sea. Atmos. Sci. Lett. https://doi.org/10.1002/asl.823, 2018.

580

Dee, D. P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, .J, Bormann, .N, Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L. K., allberg P. K., Ohler, M., Matricardi. M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J. N.,
585 Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. Roy. Meteorol. Soc. 137: 553–597. DOI:10.1002/qj.828, 2011.

Drobinski, P., Alonzo, B., Basdevant, C., Cocquerez, P., Doerenbecher, A., Fourrié, N., and Nuret, M.: Lagrangian dynamics of the mistral during the HyMeX SOP2. J Geophys. Res.-Atmos., 122(3), 1387-1402.
590 https://doi.org/10.1002/2016JD025530, 2017.

Drobinski P., Anav A., Lebeaupin-Brossier C., Samson G., Stéfanon M., Bastin S., Baklouti M., Béranger K., Beuvier J., Bourdallé-Badie R., Coquart L., D'Andrea F., de Noblet-Ducoudré N., Diaz F., Dutay J. C., Ethe C., Foujols M. A., Khvorostyanov D., Madec G., Mancip M., Masson S., Menut L., Palmieri J., Polcher J., Turquety S., Valcke S., Viovy N.:
595 Model of the regional coupled earth system (MORCE): application to process and climate studies in vulnerable regions. Environ. Model Softw. 35:1–18. doi:10.1016/j.envsoft.2012.01.017, 2012.

Drobinski, P., Bastin, S., Guénard, V., Caccia, J. L., Dabas, A. M., Delville, P., Protat, A., Reitebuch, O., and Werner, C.: Summer mistral at the exit of the Rhône valley. Q. J. Roy. Meteorol. Soc., 131(605), 353-375.
600 https://doi.org/10.1256/qj.04.63, 2005.

Drobinski, P., Bastin, S., Arsouze, T., Beranger, K., Flaounas, E., & Stefanon, M.: North-western Mediterranean sea-breeze circulation in a regional climate system model. Clim. Dynam., 51(3), 1077-1093. https://doi.org/10.1007/s00382-017-3595-z , 2018.

605

Drobinski P., Ducrocq V., Alpert P., Anagnostou E., Béranger K., Borga M., Braud I., Chanzy A., Davolio S., Delrieu G., Estournel C., Boubrahmi N. F., Font J., Grubisic V., Gualdi S., Homar V., Ivancan-Picek B., Kottmeier C., Kotroni V., Lagouvardos K., Lionello P., Llasat M., Ludwig W., Lutoff C., Mariotti A., Richard E., Romero R., Rotunno R., Roussot O., Ruin I., Somot S., Taupier-Letage I., Tintore J., Uijlenhoet R., Wernli H.: HyMeX, a 10-year multidisciplinary program on
610   the Mediterranean water cycle. B. Am. Meteorol. Soc. 95:1063–1082. doi:10.1175/BAMS-D-12-00242.1, 2014.

Drobinski, P., Dusek, J., and Flamant, C.: Diagnostics of hydraulic jump and gap flow in stratified flows over topography. Bound.-Lay. Meteorol., 98(3), 475-495. https://doi.org/10.1023/A:1018703428762, 2001b.

615   Espinoza, J. C., Lengaigne, M., Ronchail, J., and Janicot, S.: Large-scale circulation patterns and related rainfall in the Amazon Basin: a neuronal networks approach. Clim. Dynam., 38(1-2), 121-140. https://doi.org/10.1007/s00382-011-1010-8, 2012.

Drobinski, P., Flamant, C., Dusek, J., Flamant, P. H., and Pelon, J.: Observational evidence and modelling of an internal
620   hydraulic jump at the atmospheric boundary-layer top during a tramontane event. Bound.-Lay. Meteorol., 98, 497–515. https://doi.org/10.1023/A:1018751311924, 2001a.

Ernst, J. A., & Matson, M.: A Mediterranean tropical storm?. Weather, 38(11), 332-337. https://doi.org/10.1002/j.1477-8696.1983.tb04818.x, 1983.
625

Flamant, C.: Alpine lee cyclogenesis influence on air-sea heat exchanges and marine atmospheric boundary layer thermodynamics over the western Mediterranean during a Tramontane/Mistral event. J. Geophys. Res.-Oceans, 108(C2). https://doi.org/10.1029/2001JC001040, 2003.

630   Flaounas, E., Raveh-Rubin, S., Wernli, H., Drobinski, P., and Bastin, S.: The dynamical structure of intense Mediterranean cyclones. Clim. Dynam., 44(9-10), 2411-2427. https://doi.org/10.1007/s00382-014-2330-2, 2015.

Guenard, V., Drobinski, P., Caccia, J. L., Campistron, B., and Bench, B.: An observational study of the mesoscale mistral dynamics. Bound.-Lay. Meteorol., 115(2), 263-288. https://doi.org/10.1007/s10546-004-3406-z, 2005.
635

Guion, A., Turquety, S., Polcher, J., Pennel, R., Bastin S. and Arsouze, T.: Droughts and heatwaves in the Western Mediterranean: impact on vegetation and wildfires using the coupled WRF-ORCHIDEE regional model (RegIPSL). Submitted to Clim. Dynam. (2021).

640   Hitchman, M. H., and Huesmann, A. S.: A seasonal climatology of Rossby wave breaking in the 320–2000-K layer. J. Atmos. Sci., 64(6), 1922-1940. https://doi.org/10.1175/JAS3927.1, 2007.

Huang, W., Chen, R., Wang, B., Wright, J. S., Yang, Z., and Ma, W.: Potential vorticity regimes over East Asia during winter. J. Geophys. Res.-Atmos., 122(3), 1524-1544. https://doi.org/10.1002/2016JD025893, 2017.

645

Jain, A. K., & Dubes, R. C.: *Algorithms for clustering data*. Prentice-Hall, Inc, 1988.

Jiang, Q., Smith, R. B., and Doyle, J.: The nature of the mistral: Observations and modelling of two MAP events. Q. J. Roy. Meteorol. Soc., 129(588), 857-875. https://doi.org/10.1256/qj.02.21, 2003.

650

Johnson, N. C., Feldstein, S. B., and Tremblay, B.: The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. J. Climate, 21(23), 6354-6371. https://doi.org/10.1175/2008JCLI2380.1, 2008.

655   Kaskaoutis, D. G., Kambezidis, H. D., Nastos, P. T., & Kosmopoulos, P. G.: Study on an intense dust storm over Greece. Atmospheric Environment, 42(29), 6884-6896. https://doi.org/10.1016/j.atmosenv.2008.05.017, 2008.

Kiviluoto, K.: Topology preservation in self-organizing maps. Proceedings of International Conference on Neural Networks (ICNN'96), pp. 294-299 vol.1. 10.1109/ICNN.1996.548907, 1996.

660   Kleinschmidt, E.:On the structure and origin of cyclones: Part II. Meteor. Rundsch., 3 (3/4), 54–61. [U. K. Met. Office Transl. 1701.], 1950.

Lebeaupin-Brossier, C., and Drobinski, P.: Numerical high-resolution air-sea coupling over the GOL during two tramontane/mistral events. J. of Geophys. Res.: Atmospheres, 114(D10). https://doi.org/10.1029/2008JD011601, 2009.

665

Lebeaupin-Brossier, C., Drobinski, P., Béranger, K., Bastin, S., and Orain, F.: Ocean memory effect on the dynamics of coastal heavy precipitation preceded by a mistral event in the northwestern Mediterranean. Q. J. Roy. Meteorol. Soc., 139(675), 1583-1597. https://doi.org/10.1002/qj.2049, 2013.

670   Li, L., Bozec, A., Somot, S., Béranger, K., Bouruet-Aubertot, P., Sevault, F., and Crepon, M.: Regional atmospheric, marine processes and climate modelling. In Developments in Earth and Environmental Sciences (Vol. 4, pp. 373-397). Elsevier. https://doi.org/10.1016/S1571-9197(06)80010-8, 2006.

Liu, Y., Weisberg, R. H. and J. I. Mwasiagi (Eds.):. A review of self-organizing map applications in meteorology and

675   oceanography. Self-Organizing Maps: Applications and Novel Algorithm Design, InTech publications, Rijeka, Croatia, 2011.

Mattocks, C., and Bleck, R.: Jet streak dynamics and geostrophic adjustment processes during the initial stages of lee cyclogenesis. Monthly weather review, 114(11), 2033-2056.

680   https://doi.org/10.1175/1520-0493(1986)114<2033:JSDAGA>2.0.CO;2, 1986.

Millot, C.: Wind induced upwellings in the GOL. Oceanologica Acta, 2(3), 261-274. https://archimer.ifremer.fr/doc/00122/23335/, 1979.

685   Moscatello, A., Marcello Miglietta, M., & Rotunno, R.: Observational analysis of a Mediterranean 'hurricane' over south-eastern Italy. Weather, 63(10), 306, 2008.

Obermann, A., Bastin, S., Belamari, S., Conte, D., Gaertner, M. A., Li, L., and Ahrens, B.: Mistral and Tramontane wind

690   speed and wind direction patterns in regional climate simulations. Clim. Dynam., 51(3), 1059-1076. https://doi.org/10.1007/s00382-016-3053-3, 2018.

Plačko-Vršnak, D., Mahović, N. S., and Drvar, D.: Case study on Genoa cyclone with mistral 13-15 February 2005. http://www.umr-cnrm.fr/icam2007/ICAM2007/extended/manuscript_180.pdf, 2005.

695

Pytharoulis, I.; Craig, G.C.; Ballard, S.P.: Study of the Hurricane-like Mediterranean Cyclone of January 1995. Phys. Chem. Earth. Elsevier B.V. 24 (6): 627–632. Bibcode:1999PCEB...24..627P. https://doi.org/10.1016/S1464-1909(99)00056-8 1999.

700   Rainaud R, Lebeaupin-Brossier C, Ducrocq V, Giordani H, Nuret M, Fourrié N, Bouin M-N, Taupier-Letage I, and Legain D.: Characterisation of air–sea exchanges over the Western Mediterranean Sea during the HyMeX SOP1 using the AROME-WMED model. Q. J. Roy. Meteorol. Soc. 142: 173–187. https://doi.org/10.1002/qj.2480, 2016.

Rainaud, R., Lebeaupin-Brossier, C. L., Ducrocq, V., and Giordani, H.: High-resolution air–sea coupling impact on two heavy precipitation events in the Western Mediterranean. Q. J. Roy. Meteorol. Soc., 143(707), 2448-2462. https://doi.org/10.1002/qj.3098, 2017.

Raveh-Rubin, S., and Flaounas, E.: A dynamical link between deep Atlantic extratropical cyclones and intense Mediterranean cyclones. Atmos. Sci. Lett., 18(5), 215-221. https://doi.org/10.1002/asl.745, 2017.

Raveh-Rubin, S., and Wernli, H.: Large-scale wind and precipitation extremes in the Mediterranean: a climatological analysis for 1979–2012. Q. J. Roy. Meteorol. Soc., 141(691), 2404-2417. https://doi.org/10.1002/qj.2531, 2015.

Raveh-Rubin, S., and Wernli, H.: Large-scale wind and precipitation extremes in the Mediterranean: dynamical aspects of five selected cyclone events. Q. J. Roy. Meteorol. Soc., 142(701), 3097-3114. https://doi.org/10.1002/qj.2891, 2016.

Ricchi, A.; Miglietta, M.M.; Barbariol, F.; Benetazzo, A.; Bergamasco, A.; Bonaldo, D.; Cassardo, C.; Falcieri, F.M.; Modugno, G.; Russo, A.; Sclavo, M.; Carniel, S. Sensitivity of a Mediterranean Tropical-Like Cyclone to Different Model Configurations and Coupling Strategies. Atmosphere-Basel 2017, 8, 92.

Rossa, A. M., Wernli, H., and Davies, H. C.: Growth and decay of an extra-tropical cyclone's PV-tower. Meteorol. Atmos. Phys., 73(3-4), 139-156. https://doi.org/10.1007/s007030050070, 2000.

Ruffault, J., Moron, V., Trigo, R. M., and Curt, T.: Daily synoptic conditions associated with large fire occurrence in Mediterranean France: evidence for a wind-driven fire regime. Int. J. Climatol., 37(1), 524-533. https://doi.org/10.1002/joc.4680, 2017.

Ruti P., Somot S., Giorgi F., Dubois C., Flaounas E., Obermann A., Dell'Aquila A., Pisacane G., Harzallah A., Lombardi E., Ahrens B., Akhtar N., Alias A,. Arsouze T., Raznar R., Bastin S., Bartholy J., Béranger K., Beuvier J., Bouffies-Cloche S., Brauch J., Cabos W., Calmanti S., Calvet J., Carillo A., Conte D., Coppola E., Djurdjevic V., Drobinski P., Elizalde A., Gaertner M., Galan P., Gallardo C., Gualdi S., Goncalves M., Jorba O., Jorda G., Lheveder B., Lebeaupin-Brossier C., Li L., Liguori G., Lionello P., Macias-Moy D., Onol B., Rajkovic B., Ramage K., Sevault F., Sannino G., Struglia M., Sanna A., Torma C., Vervatis V.: MED-CORDEX initiative for Mediterranean Climate studies. Bull Am Meteorol Soc. doi:10.1175/BAMS-D-14-00176.1, 2015.

Schott, F., Visbeck, M., Send, U., Fischer, J., Stramma, L., and Desaubies, Y.: Observations of deep convection in the GOL, northern Mediterranean, during the winter of 1991/92. J. Phys. Oceanogr., 26(4), 505-524. https://doi.org/10.1175/1520-0485(1996)026<0505:OODCIT>2.0.CO;2, 1996.

740

Scorer, R. S.: Mountain-gap winds; a study of surface wind at Gibraltar. Q. J. Roy. Meteorol. Soc., 78(335), 53-61. https://doi.org/10.1002/qj.49707833507, 1952.

Sheridan, S. C., and Lee, C. C.: The self-organizing map in synoptic climatological research. Prog. Phys. Geog., 35(1), 109-119. https://doi.org/10.1177/0309133310397582, 2011.

745

Smith, R. B. (1986). Further development of a theory of lee cyclogenesis. J. Atmos. Sci., 43(15), 1582-1602.

Speranza, A., Buzzi, A., Trevisan, A., and Malguzzi, P.: A theory of deep cyclogenesis in the lee of the Alps. Part I: Modifications of baroclinic instability by localized topography. J. Atmos. Sci., 42(14), 1521-1535. https://doi.org/10.1175/1520-0469(1985)042<1521:ATODCI>2.0.CO;2, 1985.

750

Tafferner, A.: Lee cyclogenesis resulting from the combined outbreak of cold air and potential vorticity against the Alps. Meteorol. Atmos. Phys., 43(1), 31-47. https://doi.org/10.1007/BF01028107, 1990.

755

Thorncroft, C. D., Hoskins, B. J., and McIntyre, M. E.: Two paradigms of baroclinic-wave life-cycle behaviour. Q. J. Roy. Meteorol. Soc., 119(509), 17-55. https://doi.org/10.1002/qj.49711950903, 1993.

Tous, M., and Romero, R.: Meteorological environments associated with medicane development. Int. J. Climatol., 33(1), 1-14. https://doi.org/10.1002/joc.3428, 2013.

760

Tsidulko, M., and Alpert, P.: Synergism of upper-level potential vorticity and mountains in Genoa lee cyclogenesis–a numerical study. Meteorol. Atmos. Phys., 78(3-4), 261-285. https://doi.org/10.1007/s703-001-8178-8, 2001.

765

Wernli, H., and Schwierz, C.: Surface cyclones in the ERA-40 dataset (1958–2001). Part I: Novel identification method and global climatology. J. Atmos. Sci., 63(10), 2486-2507. https://doi.org/10.1175/JAS3766.1, 2006.

Wernli, H., and Sprenger, M.: Identification and ERA-15 climatology of potential vorticity streamers and cutoffs near the extratropical tropopause. J. Atmos. Sci., 64(5), 1569-1586. https://doi.org/10.1175/JAS3912.1, 2007.

770

Wilks, D.: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. B. Am. Meteorol. Soc, 97(12), 2263-2273. https://doi.org/10.1175/BAMS-D-15-00267.1, 2016.