Use of cluster analysis to identify the meteorological patterns associated with hail phenomena in Campania (Southern Italy)

Author: Sabrina Veltri[1], sabrina.veltri@studenti.unisalento.it

Supervisor: Vincenzo Capozzi[2]

[1] Università del Salento, [2] Università degli studi di Napoli "Parthenope"

Abstract

Hail is one of the most diffuse and harmful meteorological hazard, causing damages to crops, transport activities, buildings and infrastructure. Therefore, improving hail forecast is a crucial task to preserve assets and human life, especially in metropolitan areas, whose infrastructure and settlements are more and more vulnerable to severe convective events. The aim of this work is to analyze the relationship between hailstorm occurrence and atmospheric circulation patterns in Naples metropolitan area (Campania Region) over the period from 1988 to 2021. To achieve this goal, multivariate statistical techniques, including Principal Components Analysis and Cluster Analysis, have been applied to some meteorological fields obtained from ERA5 reanalysis, which describe the state of the atmosphere at different levels and, thus, help in the evaluation of thermal and dynamic instability. The selected fields are, (i) the 500-hPa geopotential height and temperature, (ii) the 850-hPa specific humidity, (iii) the 700-hPa vertical velocity and (iv) the 1000-hPa horizontal winds. The atmospheric pattern classification has been carried out after partitioning the available hail events according to their period of occurrence, i.e. warm season (May to September) and cold season (October to April). The Cluster Analysis allowed identifying different circulation types, which have been detailed analyzed and discussed. Moreover, two case studies related to the hailstorms occurred on January 30, 2015 and August 25, 2021 have been examined, in order to investigate the synoptic and mesoscale conditions and to validate the results obtained from Cluster Analysis.

Keywords: hailstorm, cluster analysis, atmospheric circulation pattern, Naples urban area

1. Introduction

It is widely accepted that one of the main consequences of global warming phenomenon is the increase in the frequency of occurrence of severe convective events, such as hailstorms, convective outbreaks, windstorms and heavy rainfall. Among these, hail precipitation is one of the most diffuse hazard, having adverse effects on socio-economic fabric, exacerbated, in many contexts, by the increasing anthropogenic pressure. There are several issues that prevent scientists to accurately forecast hailfall for a given location and time. Such limits in hail forecasting deal with the very limited extension of spatial and temporal scales of hail phenomena, as well as with the complex interaction between orographic features and air masses, which often has a pivotal role in the hail formation processes. In addition, many countries still do not have a national hail-monitoring network. Indeed, hail is monitored using different tools, which results in information not always comparable. Moreover, in the context of climate change, one of the greatest uncertainties lies in the evaluation of changes in the intensity and frequency of hailstorms. According to the projected climate change scenarios, if global warming increases, some extreme events will become more frequent, and there will likely occur events with increased severity, duration and/or spatial extension unprecedented in the observational record, due to the change of regional atmospheric circulation (IPCC, 2021). Several previous studies devoted a great attention to hail events and generally followed different research paths: (i) the investigation about the relationship between hail occurrence and atmospheric pattern over several regions of Europe (Aran et al. 2011; Merino et al. 2013; Simeonov et al. 2003); (ii) the analysis of the influence of climate change on hail events, using climatological data (Sanchez et al. 2017; Giaiotti et al. 2003; Eccel et al. 2012).

Within this framework, the purpose of this study is to identify the atmospheric circulation patterns that trigger the hail event in the Naples metropolitan area (Campania Region, Southern Italy). To pursue this aim, a well-populated dataset of severe convective events including on-ground observation of hailfall occurrence and reanalysis data has been investigated through popular statistical techniques, such as the Cluster Analysis.

This work has been structured as follows. Section 2 describes the dataset and methodology. Section 3 presents the results of both synoptic and mesoscale classification. Finally, in Section 4 two case studies are discussed, whereas the conclusions and the future perspectives are provided in Section 5.

2. Data and methods

The scarcity and unsatisfactory standardization of hail detection and monitoring systems is a critical issue in collecting information on hail events. For this reason, we needed to use different sources to build and broaden our dataset. First, we borrowed the data collected by Dr. Alberto Fortelli, who gathered the main features of hail and severe convective rainfall events occurred over Naples metropolitan area from 1988 to the end of the 20th century. Secondly, in order to improve this dataset, the information about thunderstorm and hail events provided by weather amateurs (http://www.campanialive.it/) have been taken into account. The volunteers' reports, although inevitability subjected to biases and uncertainties, have been widely used in some previous studies (e.g. Allen and Tippett, 2015, Schuster et al., 2005) and allowed us to extend our dataset up to 2015. Furthermore, another source ofprecious information has been the "European Severe Weather Database" (https://eswd.eu/), which gave us the opportunity to include in our work some recent severe events up to 2021. The dataset also includes data provided by ground observations of two meteorological stations based in Capri and Naples Capodichino, belonging to the network of "Meteorological Service of Military Italian Aeronautics". At the end of an exhaustive research, the thunderstorm events dataset, which includes both heavy rain and hailfall events, consists of 170 severe convective events, 85 of which were characterized by the presence of hail and will be called, henceforth, *hail days*. However, it is important to point out that our temporal series is not continuous and homogeneous. In fact, it presents some gaps, particularly between 2003 and 2009. This lack of data highlights the need to improve hail- monitoring network, as stated at the beginning of this section. The methodological approach to analyze our dataset must take into account the fact that not only that the duration of hailstorms is usually short but also that the area affected by hail precipitation is typically small. For these reasons, a multivariate statistical methodology is used in this study with the aim of building an atmospheric circulation catalogue, made of synoptic and mesoscalepatterns that could potentially cause hailfall over the examined area. Firstly, in order to obtain the meteorological configurations associated with hail days, we worked with ERA5 reanalysis developed by the European Medium Weather Forecast (ECMWF), which combines model data with observations from across the world into a globally complete and consistent data set using the laws of physics. In particular, the data used in this paper, available online at the site https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure- levels?tab=overview, provides hourly estimates for a large number of atmospheric, ocean-wave and land-surface quantities on pressure levels from 1979 onwards, with a horizontal resolution of 0.25 degrees and a vertical resolution of 37 pressure levels, from 1000 hPa to 1 hPa. The fields chosen to spatially characterize the atmospheric environment favorable for the formation of severe convection and hailstorms were: 500-hPa geopotential height and temperature (Z500 and T500, respectively), 850-hPa specific humidity (SH850), 700-hPa vertical velocity (VV700) and 1000-hPa horizontal winds (U1000 and V1000). These parameters are widely used in previous works(e.g. Aran et al. 2011; Merino et al. 2013), since they describe the dynamic and thermodynamic state of the atmosphere. Furthermore, horizontal components of wind at 1000 hPa were selected to take into account the mainflux at low levels. In addition, for better understanding the characteristics of air masses involved during cold season, we also chose temperature at 500 hPa. These fields were displayed over three domains: the first domain, extended from 29.5°N to 68.5°N in latitude and from 17.5°W to 21.5°E in longitude, was used for the evaluation of 500-hPa geopotential height and temperature, whereas the second one, extended from 34.6°N to 48.7°N in latitude and from 3.9°E to 21.5 °E in longitude, was used for 700-hPa vertical velocity and 850-hPa specific humidity. The last domain was defined over Campania region (39.5°N to 42°N in latitude and 13°E to 16.5°E in longitude) and it was selected for the 1000-hPa horizontal winds. This choice was made in line with the study purposes, in order to better visualize and analyze every field taken into account. Moreover, for each hail day, we decided to use the reanalysis data at 12:00 UTC as reference pattern. In order to obtain the atmospheric circulation patterns from our selected fields, multivariate analysis techniques were used, as many other authors have done (e.g. Aran et al. 2011; Merino et al. 2013; Esteban et al. 2006; Houssos et al. 2008). Firstly, a Principal Components Analysis (PCA) was carried out. PCA is one of the most important instruments for exploratory data analysis, which is an approach widely used to summarize the main characteristics of a determined dataset. More specifically, PCA is a dimensionality-reduction method that is used to reduce the dimensionality of large data sets by transforming a big set of variables into a smaller one, while maintaining a large representation of the variability contained in the original data. The procedures for the PCA consist in the estimation of eigenvalues and eigenvectors of covariance (S-mode) or correlation (R-mode) matrix in order to determine a new set of variables, called *principal components* of data. These are constructed as linear combinations of the initial variables, in such a way that the new variables, i.e. the principal components, are uncorrelated and most of the information is compressed in the first ones. In this work, it was performed a PCA in R-mode, using days as variables and reanalysis grid points as observations. Thisapproach allows standardizing variables so that each variable has unit variance, regardless of the physical dimension of the problem. In this way, all the variables have an equal weight. The criterion employed to determinethe optimal number of components is the Scree test. The latter can be defined as the procedure of finding statistically significant factors or components (using a Scree plot), i.e. the components that explains the most of variability (70-90%) in the data. In particular, we used the Scree plot to evaluate the cumulative portion of variance explained and we have selected the number of components that, together, reached an explained variance of at least 80%. Finally, we have applied the CA to the components extracted from PCA. The CA is a very commonly tool in the framework of atmospheric field that allows grouping a set of data in such a way that data in the same group, called *cluster*, have more similar characteristics to each other than to those belonging to the other groups. Various algorithms can achieve the goal of CA in many different ways. For this study, it has been chosen the non- hierarchical k-means algorithm method. It is an iterative, data-partitioning algorithm that classifies groups of data according to their similarities using the Euclidean distance, in a way that maximizes similarities within clusters and minimizes similarities among clusters. This method requires that the number of clusters (k) is pre-determined ahead of time, with a certain degree of subjectivity. However, it should always allow for a correct physical interpretation of the results.

3. Results and discussion

It is well known that the large-scale dynamic and processes that generate favorable conditions for hail events have a seasonal dependence. For this reason, the data set has been preliminary divided in two groups: a “warm season” sub-set, which includes the events that took place from May to September, and a “cold season” sub-set, which encompasses the events that took place from October to April. The seasonal distribution of the events is not homogeneous, since the events belonging to the cold season (109) are almost double the events that took place in the warm one (61). Moreover, the hail days in the cold season (63) appear to be almost three times those present in the warm season (22). This might be a consequence of the fact that during cold season mid-latitude low pressure systems are likely to occur due to the winter position and strength of Polar Front. In this scenario, severe weather events are often associated with the passage of a cold front or a cold occlusion over Campania region. During warm season, thunderstorms are often triggered by an upper level low, which generates a convective environment very prone to thunderstorm activity. Furthermore, during hot summer days the severe weather conditions could be also enhanced by low-level wind convergence lines, generated by the interaction between mesoscale circulation (i.e. sea breeze) and synoptic flow, as well as by local orographic features, which forces upward vertical motion. The monthly distribution of the events is presented in Table 1.

Table 1. Monthly distribution of all severe convective events and hail days observed in Naples metropolitan area for the 1988-2021 period.

Months	**Events**	**Hail days**
January	7	7
February	10	9
March	8	7
April	22	19
May	8	5
June	12	8
July	10	4
August	9	2
September	22	3
October	27	5
November	22	9
December	13	7
Total	**170**	**85**

Taking into account the 85 hail days, PCA and CA were separately applied to the warm and cold season. Theresults will be discussed in the next subsections.

3.1 Warm season

As a result of the PCA, the criterion established for the explained variance in Section 2 was satisfied by the first eight components for Z500 and T500, by the first fourteen for SH850, by the first eleven components for horizontal winds at 1000 hPa (HW1000) and by the first sixteen for VV700. Using these PCA components, we have applied the CA. Following standard procedures to assess cluster reproducibility and stability, we have performed several sensitivity experiments, by running the k-mean procedures using different randomly selected seed clusters. The final selection of the optimal number of clusters has been performed by manual inspection. According to this strategy, four different clusters were extracted. The dominant cluster, in terms of frequency of occurrence, is Cluster 1 ($CL1_W$), which is associated to 9 hail days. Cluster 3 ($CL3_W$) and Cluster 4 ($CL4_W$) have a similar number of elements, with 3 and 4 events, respectively, whereas six events belong Cluster 2 ($CL2_W$). It is worth noting that all the events of the third cluster took place in the last 10-years period, i.e. from 2011 onwards. The monthly distribution by cluster is depicted in the histogram reported in Figure 1.

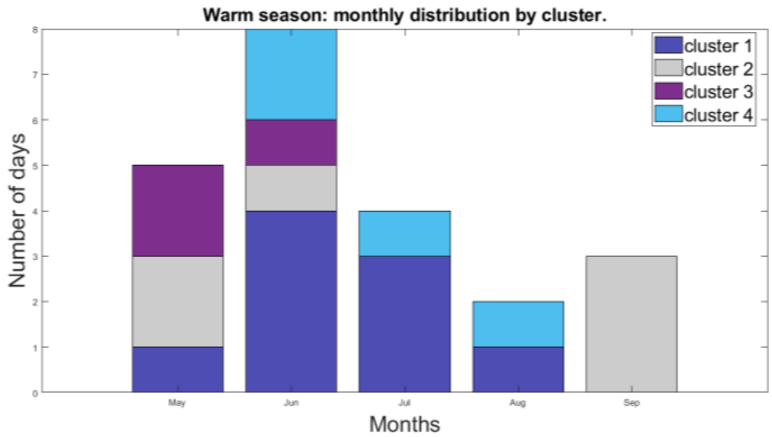


Figure 1. Frequency of occurrence of each cluster for every month of the warm season (May to September).

It can be seen that in June all the patterns are present, whereas in September it can be only find the meteorological configuration represented by $CL2_W$. Moreover, $CL1_W$ dominates in June and July. In the following paragraphs, the four cluster will be discussed in details.

3.1.1 Cluster 1 and 4

The description of the four circulation types starts from $CL1_W$ and $CL4_W$, which exhibit similar patterns. According to Z500 field, in $CL1_W$ (Fig. 2a, left panel) the synoptic environment area shows a closed anticlockwise circulation over Adriatic Sea's northern basin. This closed circulation resembles the final stage (i.e. cut-off low) of a Rossby wave evolution. The scenario described by $CL4_W$ (Fig. 2a, right panel) is similar, although the low-pressure area was located over central and northern Italy and is characterized by higher values of geopotential height. As a consequence, mid- and upper level winds blow from north-western and western directions over our study area. Furthermore, the mesoscale configuration at 1000 hPa is characterized by southwesterly winds over the western sectors of Campania, while in the eastern ones wind blows from north. This pattern determined a low-level wind convergence in the inland areas of the region of interest. With this configuration, it comes as no surprise that negative values of vertical velocity at 700 hPa can be observed in both clusters (Fig. 2c). Lastly, the air mass at 850 hPa presents elevated concentration of water vapor (exceeding 8 gkg^{-1}) in the mountainous areas of the Campania region (Fig.2d).

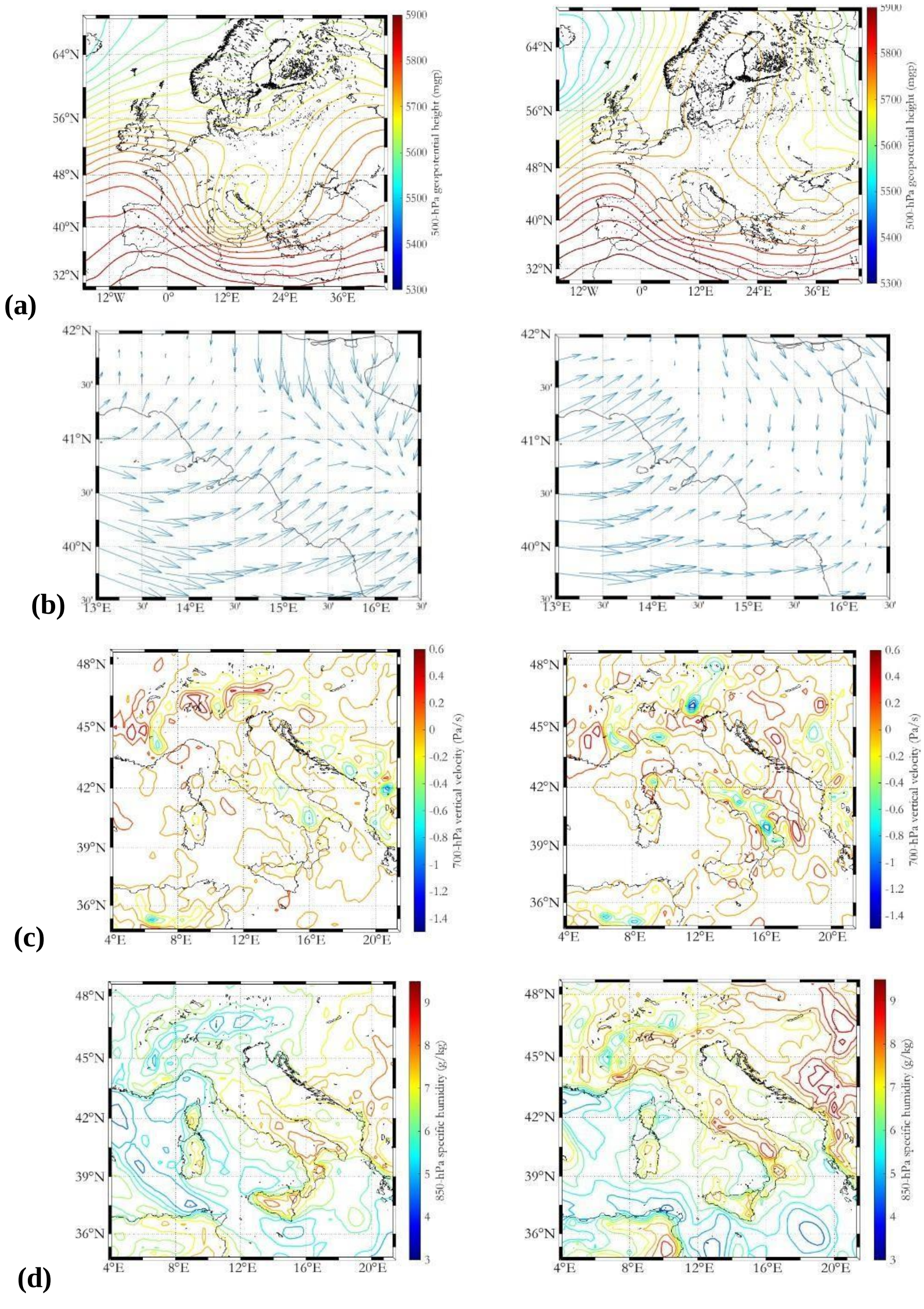
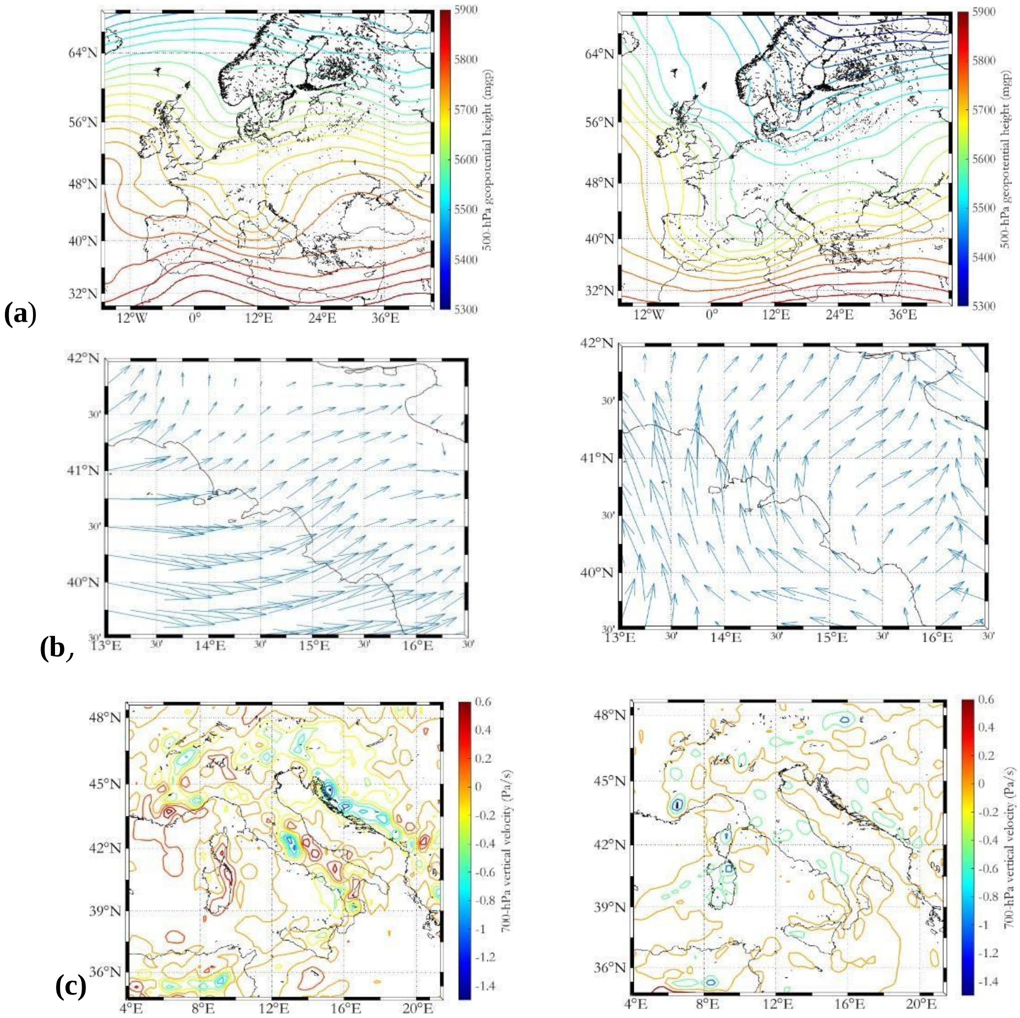


Figure 2 (a) Map of geopotential height at 500 hPa for $CL1_W$ (left) and $CL4_W$ (right). (b) Map of horizontal winds at 1000 hPa for $CL1_W$ (left) and $CL4_W$ (right). (c) Map of vertical velocity at 700 hPa for $CL1_W$ (left) and $CL4_W$ (right). (d) Map of specific humidity at 850 hPa for $CL1_W$ (left) and $CL4_W$ (right).

From this analysis, it emerges that hailstorms occurred during hail days within $CL1_W$ and $CL4_W$ could have been caused by the interaction between the synoptic low-pressure system and low-level convergence of winds, which forced warm and humid air mass to uplift over the steep orography of Campania.

3.1.2 Cluster 2 and 3

The synoptic pattern that provided a favorable environment for the onset of deep convection, in the case of $CL2_W$ and $CL3_W$, is characterized by the presence of a neutrally tilted trough, i.e. from north to south, over Italy, and a positive tilted trough, which extend from Northeastern Europe to southwest over Western Mediterranean basin, respectively (Fig.3a). The latter appears to be deeper than the former. Nevertheless, in both cases the study area is affected by upper level southwesterly winds. The mesoscale configuration at 1000 hPa is modulated by prevailing westerly ($CL2_W$) and southerly ($CL3_W$) winds from the sea, which back ($CL2_W$) and veer ($CL3_W$) becoming weaker once approaching the land, as it can be seen in Fig.3b. Moreover, the pattern described by $CL2_W$ shows negative mid-level values of vertical velocity over southwestern sectors of the domain of interest, along with relatively high concentration of water vapor at 850 hPa, between 7.5 and 8.5 gkg^{-1}, over the whole region under study. In contrast, no significant value of vertical velocity at 700 hPa was detected in $CL3_W$, expect for slightly negative values of vertical velocity over inland areas of Campania, because of the presence of mountainous areas (Fig.3c, right panel). Lastly, the air mass at 850 hPa of $CL3_W$ appears to be drier than that of $CL2_W$ (Fig.3d).



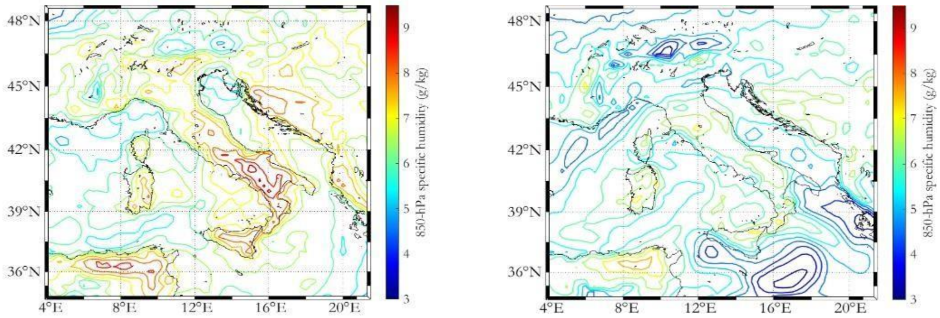
(d) 

Figure 3 (a) Map of geopotential height at 500 hPa for $CL2_W$ (left) and $CL3_W$ (right). (b) Map of horizontal winds at 1000 hPa for $CL2_W$ (left) and $CL3_W$ (right). (c) Map of vertical velocity at 700 hPa for $CL2_W$ (left) and $CL3_W$ (right). (d) Map of specific humidity at 850 hPa for $CL2_W$ (left) and $CL3_W$ (right).

Herein, the analysis point out that the cause of hailstorms in meteorological environments described by $CL2_W$ and $CL3_W$, appears to be due to the interaction between the large-scale pattern at upper level and moist air masses coming from Tyrrhenian Sea, advected towards Campania Region by western and southern winds.

3.2 Cold season

Firstly, the principal components extracted from PCA where 8 and 11 for Z500 and T500, respectively, 31 for SH850, 21 and 20 for U1000 and V1000 and 43 for VV700 parameter. From CA, using the same approach adopted for warm season, we have extracted three different clusters. According to the Fig. 4, which shows the frequency of occurrence of each cluster for every month of the cold season, the Cluster 2 ($CL2_c$) is the dominant one: it is associated to 34 hail days. Cluster 1 ($CL1_c$) and 3 ($CL3_C$) forced 11 and 18 events, respectively, instead. The monthly distribution by cluster is not uniform. Indeed, it can be seen that $CL3_C$ predominates during November and April, whereas in January and February the most frequent Cluster is $CL2_C$. Furthermore, the meteorological pattern described by $CL1_c$ is not present in January and February, as $CL3_c$ in December. All the clusters obtained from the CA will be discussed in details in the following paragraphs. As done before, clusters with similar synoptic fields will be presented together.

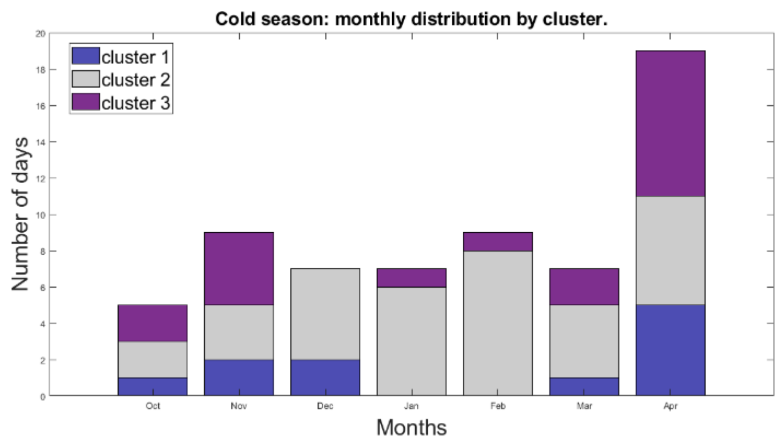


Figure 4. Frequency of occurrence of each cluster for every month of the cold season (October to April).

3.2.1 Cluster 1

The main features of the synoptic environment can be identified in the anomalous position of Azores anticyclone, which extended northwards to the North Atlantic and in the presence of a trough stretched from northern Europe to central Mediterranean basin (Fig. 5a). A closed anticlockwise circulation placed between Northwestern Italy and the Gulf of Genoa, characterized by a cold core at 500-hPa isobaric level can be easily detected from Fig.5b. Within this scenario, the Campania region is affected by southwesterly at mid and upper tropospheric levels, as well as in the low troposphere, as well highlighted by Fig. 5c. Moreover, updrafts at 700-hPa can detected over the entire region, while the air mass at 850 hPa shows relative elevated concentration of water vapor, exceeding 5 gkg^{-1} (Fig. 5d,e). It may be envisaged that this synoptic and mesoscale configuration promotes the development of convective cells along the relatively warm Tyrrhenian Sea that are advected towards coastal area of Campania Region by southwesterly winds.

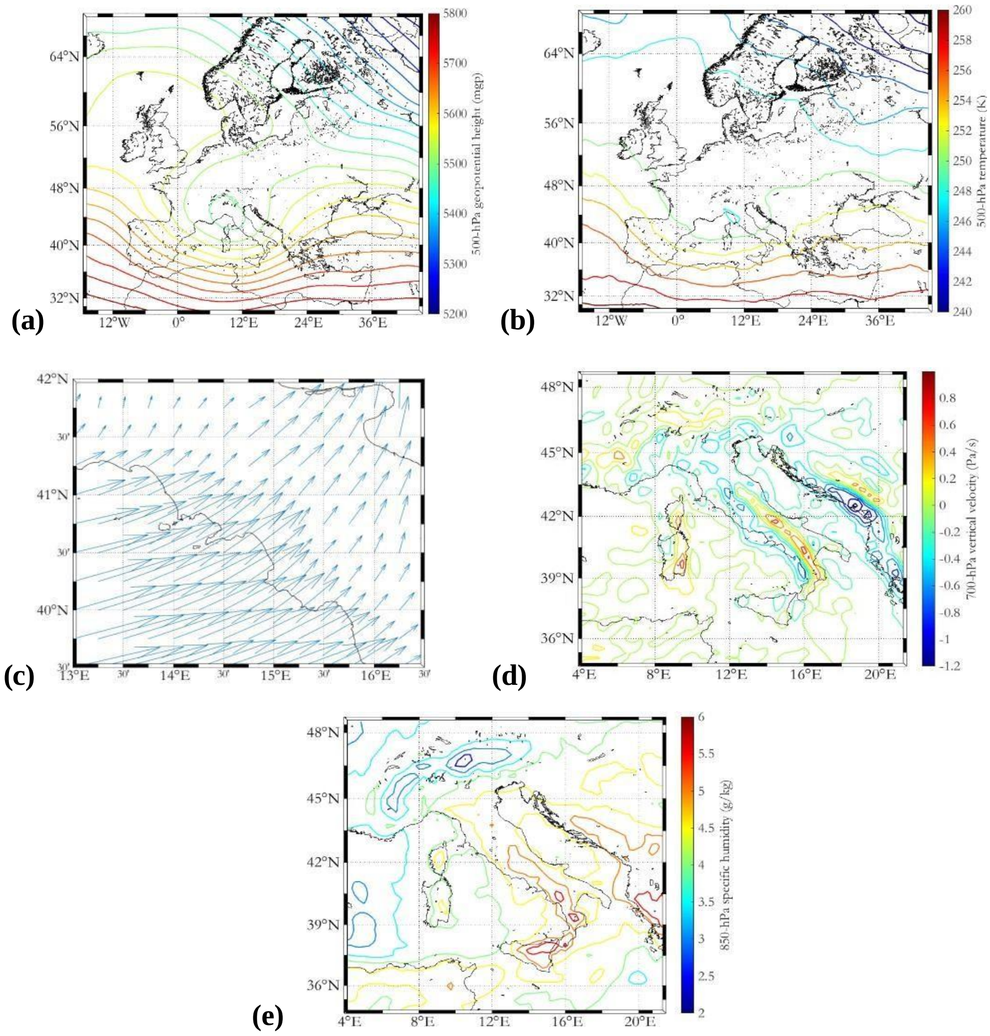


Figure 5 (a) Map of geopotential height at 500 hPa for $CL1_C$. (b) Map of temperature at 500 hPa for $CL1_C$. (c) Map of horizontal winds at 1000 hPa for $CL1_C$. (d) Map of vertical velocity at 700 hPa for $CL1_C$. (e) Map of specific humidity at 850 hPa for $CL1_C$.

3.2.2 Cluster 2 and 3

The synoptic pattern associated with hail events of $CL2_C$ and $CL3_C$ is, in both cases, a neutrally tilted trough elongated from northern Europe to Italy. The latitudinal extension of Z500 field suggests that this pattern represents, on average, the mature stage of the Rossby wave associated to the polar front oscillation, which results in strong thermal and pressure gradients. Despite Z500 values are similar, trough amplitude in $CL3_C$ is higher than $CL2_C$ (Fig.6a). In addition, there is an upper level thermal trough at 500 hPa, with temperatures lower that 248 K (-25°C) over Campania region, as it can be detected from Fig.6b. Similarly to the Cluster 1, this scenario promotes a southwestern flow along the entire tropospheric column, as well as negative values VV700, especially along coastal areas of Campania (Fig.6c,d). Lastly, the air mass at 850 hPa exhibit relatively high values of water vapor concentration, ranging between *4.0* and 4.5 gkg^{-1} along the coastline of the entire Campania region (Fig.6e).

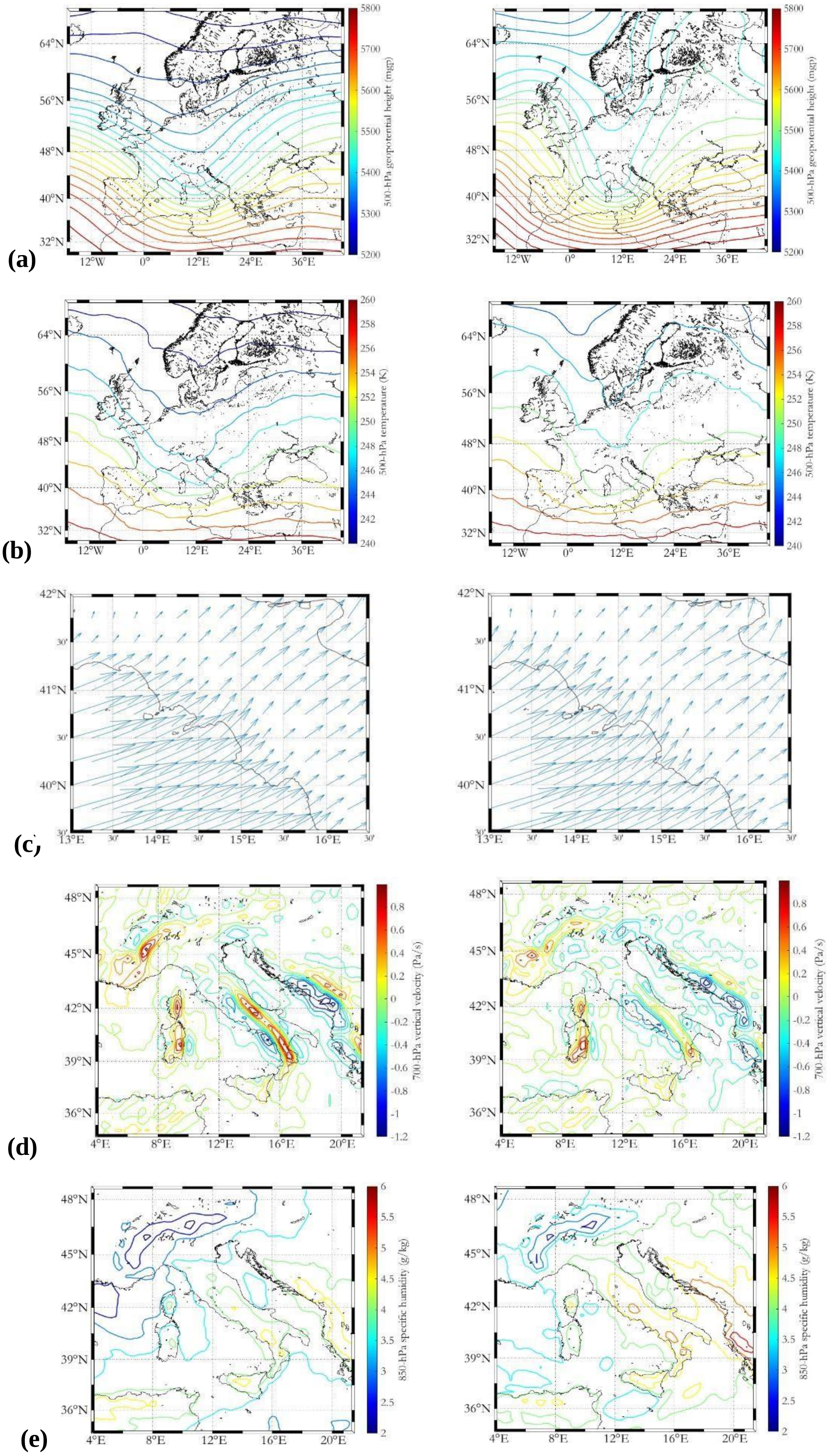


Figure 6 (a) Map of geopotential height at 500 hPa for $CL2_C$ (left) and $CL3_C$ (right) .(b) Map of temperature at 500 hPa for $CL2_C$ (left) and $CL3_C$ (right). (c) Map of horizontal components of wind at 1000 hPa for $CL2_C$ (left) and $CL3_C$ (right). (d) Map of vertical velocity at 700 hPa for $CL2_C$ (left) and $CL3_C$ (right). (e) Map of specific humidity at 850 hPa for cluster 2 (left) and $CL3_C$ (right).

From the analysis of the events occurred during cold season, it is clear that hailstorms were caused by an advection of upper level cold air mass, which appears to be an arctic maritime air mass in the case of $CL1_c$ and an arctic continental one in the case of $CL2_C$ and $CL3_C$. In both cases, the interaction between the air mass and the warm sea surface can lead to the formation of clouds that exhibits significant vertical development, thanks to the condensation of water vapor removed from the sea. These features, together with strong southwesterly winds, may be responsible for the onset of convective phenomena, which could lead to hailfall. However, it is necessary to point out that the analysis presented in this section did not take into account any instability parameter for the prediction of hailfall, since it would have gone beyond the purposes of this work. Indeed, as stated before, our aim was to build a preliminary catalogue of meteorological pattern associated with hail in order to improve forecasting of hail over Campania region. Finally, two hail events of our dataset will be analyzed, in order to validate the CA carried out in this paper.

4. Cases study

In this section, two hailstorm events, one occurred during warm season and the other occurred during cold season, are analyzed. Within the framework of our analysis, the former belongs to Cluster 1 (Section 3.1.1) and the latter to Cluster 2 (Section 3.2.2).

4.1 Warm season

On August 25, 2021, a hailstorm occurred in Naples between 2:30 and 4:00 AM and determined serious damage especially in the southern sectors of the metropolitan area, where some commercial activities were affected by floods, as reported by local newspapers. The cluster in which this event has been included, which is $CL1_W$, shows a cyclone over the Adriatic Sea's northern basin, which drives westerly winds towards the coast of Campania region. The specific pattern that forced this event is synthetized in Fig.7. More specifically, the upper level configuration (shown in Fig. 7a) presents some similarities with the one observed for $CL1_W$, showing a cyclonic circulation over northeastern Italy and Slovenia. Nevertheless, the Z500 values over Campania region are almost the same, as it can be seen from the map at 1:00 UTC sketched in Fig. 7a. Furthermore, the air mass at 850 hPa is characterized by the presence of an elevated concentration of water vapor, which is roughly $7 g kg^{-1}$ and thus, comparable with that detected in $CL1_W$. The most interesting feature is the presence of a cyclonic circulation of surface winds over the entire Campania's coast and the adjacent Tyrrhenian Sea, which is not present in the pattern obtained previously (Fig.7b). In addition, the pattern of VV700 is not characterized by any significant value, neither negative nor positive. Considering all these factors, it could be inferred that the strong hailstorm was caused by the interaction between upper level westerly flow and cyclonic surface winds circulation.

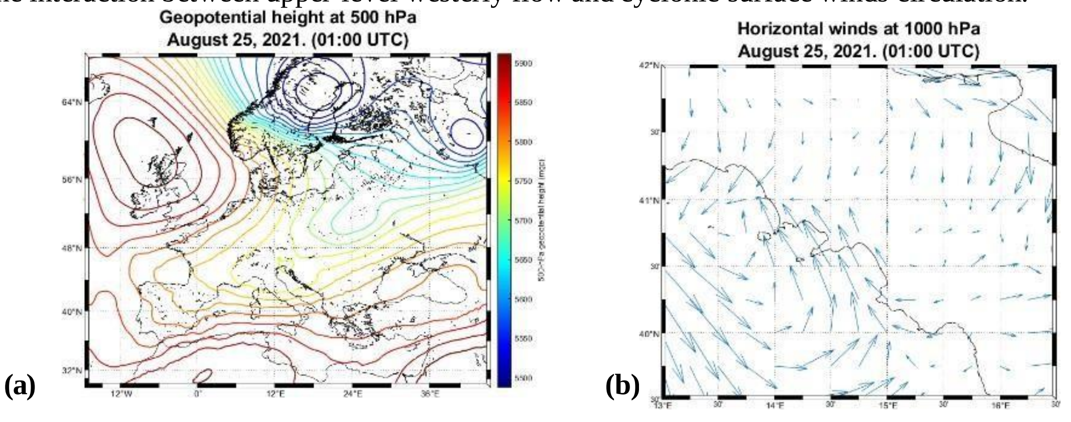
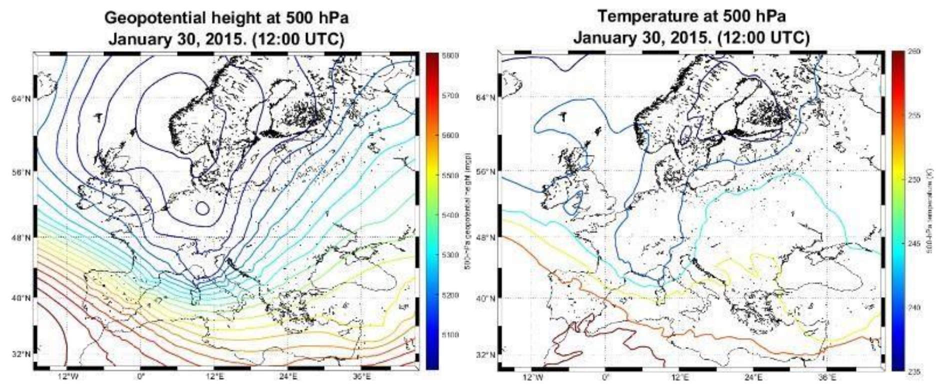


Figure 7 (a) Map of 500 hPa geopotential height on August 25, 2021 (01:00 UTC). (b) Map of horizontal components of wind at 1000 hPa on August 25, 2021 (01:00 UTC).

4.2 Cold season

The event under investigation occurred on January 2015 in the heart of winter season and lasted three days, from the 29th to the 31st of January. In the late morning of the central day of the event, a hailstorm affected the metropolitan area of Naples. For this reason, the synoptic maps of the 30th of January at 12:00 UTC are presented in the following, in order to give a general description of the main meteorological conditions that caused the hailstorm and to compare them with the pattern of $CL2_C$, in which this event has been included, as well. The whole event was determined by an advection of arctic maritime air mass that approached Italy on the 29th of January: on January 30, a trough deepened and widened, affecting the entire Italian peninsula, as shown in the charts of the Z500 and T500 presented in Fig. 8a,b.



(a) (b)

Figure 8 (a) Map of 500 hPa geopotential height on January 30, 2015 (12:00 UTC). (b) Map of 500 hPa temperature on January 30, 2021 (12:00 UTC).

It is evident that the region of Campania, particularly our area of interest, is in the east region of the trough, which is known to be an area of fast, divergent upper-troposphere winds, which triggered the genesis of a cyclonic area. The low-pressure system, along with the cold temperature at 500 hPa and the presence of westerlies surface winds (not shown), lead to pronounced atmospheric instability and, therefore, caused a hailstorm. As concern the maps of Z500, it appears that there is no close circulation associated with the trough of the pattern given by the cluster analysis. Indeed, it appears to be an open Rossby wave, in agreement with the evidences provided by the reference cluster for this event. However, it should be noted that Z500 and T500 are higher than those characterizing $CL2_c$. Small differences are also present in the values of vertical velocity and specific humidity. Overall, considering the fact that cold season is affected by a great climate variability, it may be concluded that the meteorological pattern associated with this event validate our cluster.

5. Conclusions

This work aimed to acquire a better knowledge of the relation between hailstorm occurrence and atmospheric circulation over Naples metropolitan area using multivariate statistical techniques as Principal Component Analysis and Cluster Analysis. However, the greatest problem that emerged while carrying out this study is the lack of a national hail detection network. Indeed, despite several efforts in searching information on severe convective events over the region under investigation, the data sample used is small, consisting only of 85 hail events. This also prevented finding climatological trends in hail precipitation. Nevertheless, by analyzing the meteorological fields extracted from ERA5 reanalysis data set, four and three patterns were obtained for the hail events occurred during warm season (May to September) and those that took place during cold season (October to April). The results show that hailstorms occurred during warm season are caused by closed cyclonic circulation (upper level or cut-off), which enhanced the convective instability in conjunction with the advection, at low atmospheric levels, of warm and moist winds from the sea and/or low-level convergence between synoptic scale flow and mesoscale circulation (sea breeze), resulting in a strong convective activity mainly in the inland sectors of the study region. It is important to highlight that, according to the preliminary findings of this study, the atmospheric configuration represented by the third cluster is only present from 2011 onwards. During cold season, the hailstorms in Naples urban area are generally triggered by wide trough elongated from northern Europe to central and western Mediterranean basins, which determine an advection of arctic cold air masses. The latter, crossing the relatively warm Mediterranean Sea, enhance the convective instability, causing the formation of thunderstorms events that mainly affect the coastal sector of Campania Region. Lastly, two case studies occurred in warm and cold season, respectively, were examined to validate the Cluster Analysis previously done. Overall, it appears that the circulation pattern of the two events successfully reproduced the configurations of the cluster in which they were included.

Future works are needed to assess the significance of the result obtained for $CL3_w$, which is only present in the last decade. Moreover, it is evident that additional efforts are required to include in the cluster analysis some instability indices that can be used as hail forecast parameters. Finally, future works should also devoted to collect more information about past severe convective events, in order to build up a more robust and less biased dataset.

Acknowledgements

Special thanks to Dr. Alberto Fortelli from University of Naples "Federico II" who allowed us to use his meteorological diaries and to Military Italian Aeronautics for SYNOP data, which were of key importance for the development of this work.

References

Allen, J T, and Tippett M K, (2015) The characteristics of United States hail reports: 1955–2014.Electronic J. Severe Storms Meteor. 10: 1–31.

Aran M, Pena J C and Torà M (2011) Atmospheric circulation patterns associated with hail events in Lleida (Catalonia). Atmos. Res. 100:428–438.

Eccel E, Cau P, Riemann-Campe K and Biasioli F (2012) Quantitative hail monitoring in an alpine area:35-year climatology and link with atmospheric variables. Int. J. Climatol. 32:503–517. doi:10.1002/joc.2291

Esteban P, Martìn-Vide J and Mases M (2006) Daily atmospheric circulation catalogue for western Europe using multivariate techniques. Int. J. Climatol. 26:1501-1515.

Giaiotti D, Nordio S and Stel F (2003) The climatology of hail in the plain of Friuli Venezia Giulia. Atmos.Res. 67-68:247–259.

Houssos E E, Lolis C J and Bartzokas A (2008) Atmospheric circulation patterns associated with extreme precipitation amounts in Greece. Adv. Geosci. 6:5-11.

IPCC (2021) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. In Press.

Merino A, Garcìa-Ortega E, Lòpez L, Sànchez J L and Guerrero-Higueras A M (2013) Synoptic environment, mesoscale configurations and forecast parameters for hailstorms in Southwestern Europe. Atmos. Res. 122:183-198.

Sànchez J L, Merino A, Melcòm P, Garcìa-Ortega E, Fernàndez-Gonzàlez S, Berthet C and Dessens J (2017) Are meteorological conditions favoring hail precipitation change in Southern Europe? Analysis of the period 1948-2015. Atmos. Res. 198:1–10.

Schuster S S, Blong R J and Speer M S (2005) A hail climatology of the greater Sydney area and New South Wales, Australia. Int. J. Climatol., 25:1633-1650.

Simeonov P and Geogiev C G (2003) Severe wind/hail storms over Bulgaria in 1999-2001 period: synoptic- and meso-scale factors for generation. Atmos. Res. 67-68:629-643.