communications
earth & environment

ARTICLE

https://doi.org/10.1038/s43247-022-00557-9 OPEN

Compound flood hazard at Lake Como, Italy, is driven by temporal clustering of rainfall events

Fabiola Banfi [1✉] & Carlo De Michele [1✉]

Lake floods occur when the water level in the lake exceeds a threshold causing inundation of neighbouring shorelines. Despite the potential impacts of this type of flood on neighbouring settlements, the mechanisms and drivers that govern when lake floods occur, and particularly how they result from compound factors, remains poorly understood. Here we compile and analyze meteorological and historical data on lake floods at Lake Como (northern Italy) between 1980 and 2020. We identify seven modes of lake floods with climate-based drivers. In 70% of cases, floods are associated with a temporal clustering of rainfall. This was also the predominant trigger of the seven most severe floods. To a lesser extent, floods were driven by a single rainfall event over a water level previously increased by rainfall and/or melting. We conclude that lake floods represent a clear example of the potential for compound mechanisms to govern and exacerbate hazards.

[1] Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy. ✉email: fabiola.banfi@polimi.it; carlo.demichele@polimi.it

Traditionally, extreme climate-based events have been studied adopting a univariate approach, with a singular driver triggering a singular hazard[1,2]. More and more catastrophic events, however, fail to be represented by this simple assumption, showing a widespread compoundness and connectivity. Understanding which factors combine together in order to generate an impact is of great importance to reduce it, both in the present and in future scenarios[3,4]. The change in the frequency or magnitude of a natural hazard, in fact, may be related to changes in the frequency or magnitude of each of the factors involved, as well as their dependence.

An important step in this regard is the introduction of the concept of compound climate-based events. It was introduced by the Intergovernmental Panel on Climate Change (IPCC)[5] and then refined and extended by Leonard et al.[1] and Zscheischler et al.[6], that define them as "the combination of multiple drivers and/or hazards that contributes to societal or environmental risk". The innovative idea in the latter is to categorize compound events in four typologies[6]: (a) multivariate compound events, in which the extreme impact is due to multiple drivers/hazards in the same area, (b) spatially compounding events, due to multiple drivers/hazards in spatially connected locations, (c) temporally compounding events due to the temporal succession of hazards/drivers, and (d) preconditioned compound events in which the presence of a preconditioning variable increases the resulting impact.

Here, we focus on lake floods, i.e., floods due to the increase of water level in the lake above a threshold, causing inundation of neighbouring settlements. The choice was motivated by the intuition that this typology of flood may be a good example of compound event and by the lack of a comprehensive literature about lake floods. Floods are usually classified in three main typologies, i.e., fluvial, pluvial, and coastal. Lake floods, on the contrary, are less acknowledged, despite the importance of lakes for humans and natural ecosystems and the severe impacts associated. As an example, the Poyang Lake, the largest freshwater lake in China, is also one of the areas most frequently subjected to flooding in the Country. Here, extreme flood events need the contemporaneous occurrence of intense rainfall, high lake inflows, and a high water level in the Yangtze River that prevents water outflow from the lake[7]. The summer 2020 lake flood event, triggered by repeated heavy rainfalls, caused direct economic losses of about 24 billion yuan. It affected more than 7 million people and more than 185,000 ha of crops[7] and it caused indirect impacts like the increase in the risk of transmission of intestinal schistosomiasis[8]. Lake Champlain, situated at the boundary between Canada and USA, is subjected to recurrent flooding. Generally, floods are the result of rapid snowmelt and heavy rainfall in the late winter or spring, with some local increase in severity due to wind. In 2011, the Lake experienced the highest flood in the 100 year historic record. It lasted for 67 days, and it caused more than \$82 million (\$2018 US) damage[9].

In contrast to riverine floods, lake floods are characterized by durations that may also reach months. The higher persistence of flooded water increases damage to materials, buildings, and infrastructures[10], or to agriculture, preventing crops from reaching maturity[9]. Longer durations are also likely to deteriorate aquatic vegetation[11] or they may enhance some health issues like mould exposure[9]. Lake flood impacts may also be aggravated by the several services that lakes provide to the community and to the natural environment. An in-depth analysis of this typology of events is also important in the context of climate changes, that are likely to increase lake levels in some part of the globe, like North America, Siberia, Tibetan Plateau, or the Amazon basin[12]. In addition, Shugar et al.[13] found that global volume of glacial lakes increased by 48% in the period 1990–2018. Since this trend may be related both to human and natural factors, a better understanding of the variables playing a role is important to understand and predict the trend in water lake levels.

A compound perspective is already being applied in the study of river floods. Multivariate compound flooding, i.e., due to the joint occurrence of high river flow and seiching or storm surge was modelled via copulas by several authors[14–17]. Statistical approaches were also adopted to study spatially compounding fluvial events, for example modelling joint flood occurrence at the confluence of multiple rivers[18], or preconditioned compound events, conditioning joint distributions on antecedent moisture conditions generated from predecessor rain events[19–22]. Bevacqua et al.[23], instead, define a river level predictor via a regression model that includes the water input over a catchment in the previous 10 days to account for antecedent catchment wetness. Temporally compounding events, like the exacerbation of floods due to multiple rainfall events in close succession, was also considered in literature[24]. In some cases, the statistical models were combined with hydraulic models to estimate flood hazard[25,26]. A compound perspective may equally benefit the study of lake floods. In order for the lake level to considerably increase the inflows must be higher than the capacity of the outlet stream to balance them (e.g., due to the compounding of spring precipitation and melting of snow[27]) or outflow must be inadequate to maintain equilibrium on the lake level (e.g., during periods of high discharge in the outlet stream[7]). Hence, several different factors can combine and contribute to the water level increase.

In the present study, we focused on the town of Como (Northern Italy), set on the banks of the homonymous lake. Lake Como, with a total drainage area of 4543 km^2, that includes the basin of the upper Adda river and the Mera river, splits, in the southern part, in two branches: the eastern branch, where the Adda leaves the lake, and the western branch without any outlet, where Como is located (Panel (a) and (b) of Fig. 1). The absence of any outlet results in frequent flooding of the town[28], with consequent great damages[29].

Here, we analyzed and reinterpreted in a compound framework a series of lake floods that affected the town in the last 40 years. We identified four different types of climate-based drivers: snowmelt, multiple sparse rainfall events, temporal clustering of rainfall and single rainfall events, and we answered the following questions: (1) to what extent the multi-hazard drivers are related and influence each other; (2) how many of the events are compound; (3) which are the elements that characterize the most severe events. Eventually, we looked if compoundness was a characteristic of drivers alone or if it was manifested in hazards as well. Answering these questions brought us to the construction of seven modes of lake floods that summarize the mechanisms and the connections leading to the flooding of Como. Once the drivers were identified, we moved to the investigation of the weather circulations associated with the most important triggers, thus understanding the most critical meteorological situations.

Historical perspective of lake floods in Como

The first flood, affecting the town of Como, for which we have a consistent documentation, occurred in September 1431, while a previous flood event, triggered by an earthquake, is reported in 1255 (all information about historical flooding were collected from Poggi[30] and Ricci[31]). In the subsequent centuries the flood of June 1673 is remembered among the most severe; its severity was increased by the contemporaneous flooding of the Cosia stream, a small river that flows into the Lake and drains water from the area around Como. In Fig. 2, the reconstructed extent of

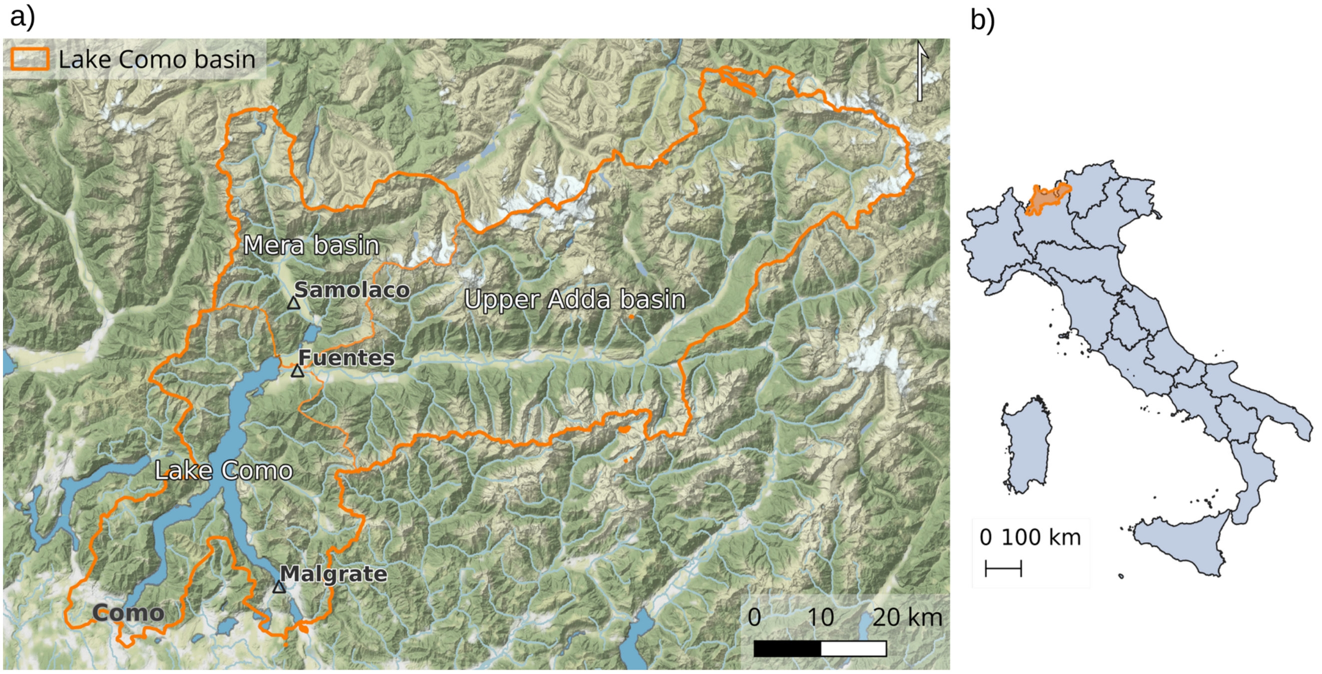

Fig. 1 Study area. a Lake Como and its drainage area. The main sites and gauges' locations named in the paper are also reported. Source of the basemap: map tiles by Stamen Design, under CC BY 3.0, and data by OpenStreetMap, under ODbL. **b** Location of Lake Como basin in Italy. The boundaries of the Italian regions were obtained from ISTAT (Istituto Nazionale di Statistica; https://www.istat.it/it/archivio/222527).

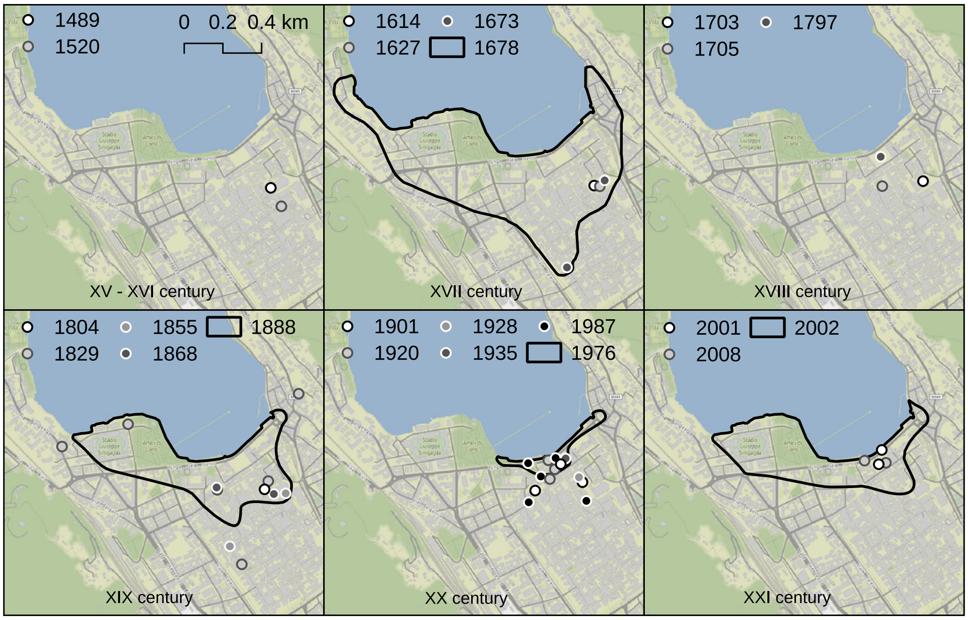

Fig. 2 Spatial extent of some flood events affecting the town of Como. Information about the impact of some flood events. For some events the reconstructed flooded area is reported, while for others only locations that are known to have been reached by water. Flooded data are provided by Regione Lombardia with CC-BY-NC-SA 3.0 Italy license and can be downloaded from the Lombardy Geoportal. Source of the basemap: map tiles by Stamen Design, under CC BY 3.0, and data by OpenStreetMap, under ODbL.

the flooded area of this event is reported, along with information about the extent of other lake flood events that affected the town.

In the next century, we have some information about a summer flood in 1703, that lasted for more than a month. The flood in 1792, instead, resulted in one-third of the town covered by water. The occurrence of the event has been associated with a rapid melting of snow, triggered by the rising of temperature due to Scirocco wind blowing over the mountainous part of the basin.

The XIX century was characterized by the occurrence of several floods of considerable extent, with at least ten events with maximum water level above 3 m (in 1955 the flooding of the town started when the level of the hydrometer was higher than 1.86 m). The September 1829 event, with a peak of 3.95 m above the zero of the hydrometer, flooded 458 houses, 45 factories, and 413 shops which corresponded to half of the houses and two third of shops in the town. Water levels in the town reached up to 1.94 m in the most depressed point of Piazza Volta, a square located in the centre of Como. The contamination of the water in wells with muddy water further aggravated the situation. The flood of September 1888 is, instead, remembered for the rapidity with which the level rose, passing from 0.9 m to 2.12 m in one single day and reaching a peak of 3.73 m five days later.

The second half of the XX century corresponded with the beginning of the lake regulation with the operation of Olginate dam, and also to the onset of the problem of the subsidence at Como. The natural subsidence rate of few mm/year increased to 10–20 mm/year in the period 1950–1975, with peaks higher than 20 mm/year[32]. The reason of this change was attributed to human activities, in particular to water withdrawal from the deep aquifer. Between 1975 and 1980 the subsidence rate started decreasing and in the recent period it returned to its natural value[32]. The floods that affected the town only marginally, associated with water levels between 1.2 and 1.3 m, can be in part reconducted to this phenomenon. Ten important floods occurred between 1946 and 1999 with levels also higher than 2.6 m and the threshold level above which flooding of Como occurs was not exceeded in only six years.

Results

Path diagram of lake floods. Figure 3a shows the path diagram of a lake flood, generated by climate-based drivers. We considered four possible forcings: (a) single rainfall event, (b) snowmelt, (c) temporal clustering of rainfall events, (d) multiple sparse rainfall events. In the analysis of triggers, we did not explore the dependence between rainfall and snowmelt, nevertheless we pointed out in the scheme how the two variables may be related. Snowmelt events and rainfall ones can occur simultaneously due to a common cause, by chance, or because the latter triggers the former. The initial level of the lake acts as a precondition, increasing or decreasing the resulting peak level, and was therefore taken into account in order to understand floods occurrence. The presence of connected flood events was also explored, looking if the occurrence of a flood had an influence on the occurrence of the following one.

In the period 1981–2020, we identified 53 flood events: 36 events occurred between April and the beginning of August and 17 events occurred between late August and the beginning of December. In the following, we will refer to the former as spring events and to the latter as autumn events.

Based on the results of the present study, we categorized lake floods in seven different modes (referred respectively as A, B, ..., G) depending on the variables playing a role and the connections between drivers and hazards, and between hazards. From this, we found that clustering of rainfall events is needed to cause a flood without a high lake level as a precondition. In addition, it increases the probability of having multiple connected flood events.

The classification is reported in Fig. 3b. Category A and F include all cases in which temporal clustering of rainfall contributes alone for more than 80% of the increase in water level. When its contribution is lower, the event is inserted in the B or E categories. The same holds for category C or D, in which the single rainfall event contributes for less than 80% of the water level increase. The equivalent of type A but with a single rainfall event, that corresponds with the univariate case, was never observed. All individual flood events are classified as A, B, or C, while connected flood events as D, E, F, or G. Categories A, B, and C have a frequency of 18.9, 17, and 20.7%, respectively. Typology B is characterized by the presence of a cluster of rainfall events over an initial lake level above the average or/and increased due to a combination of rainfall and/or snowmelt. Typology C is similar to the B case, but where flood is triggered by a single rainfall event. The predominant typology in autumn is A that constitutes 9 out of 10 of the independent autumn flood events, while in spring the most frequent typologies are B and C among independent flood events. We can notice that typologies D, E, F, G are all combinations of the first three typologies, being respectively C + C, B + B, A + A, C + B. Their frequencies are, respectively, 3.8, 22.6, 3.8, and 13.2%.

Potential triggers of flood frequency and severity. Figure 4 shows that autumn events (indicated with a grey vertical bar) are almost all associated with the temporal clustering of rainfall events. On the contrary, spring events are mainly associated with an increase in the water level due to a combination of snowmelt

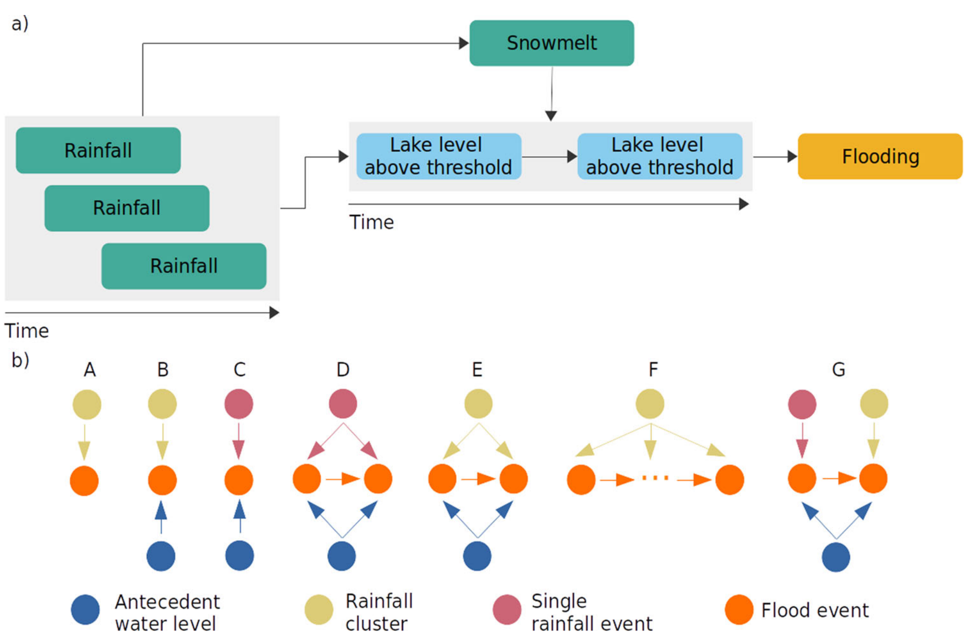

Fig. 3 Connections and variables in lake floods. a Path diagram of a lake flood event, climate-based; **b** seven (referred respectively as A, .., G) different modes of lake floods climate-based, obtained with different combinations of drivers and hazards.

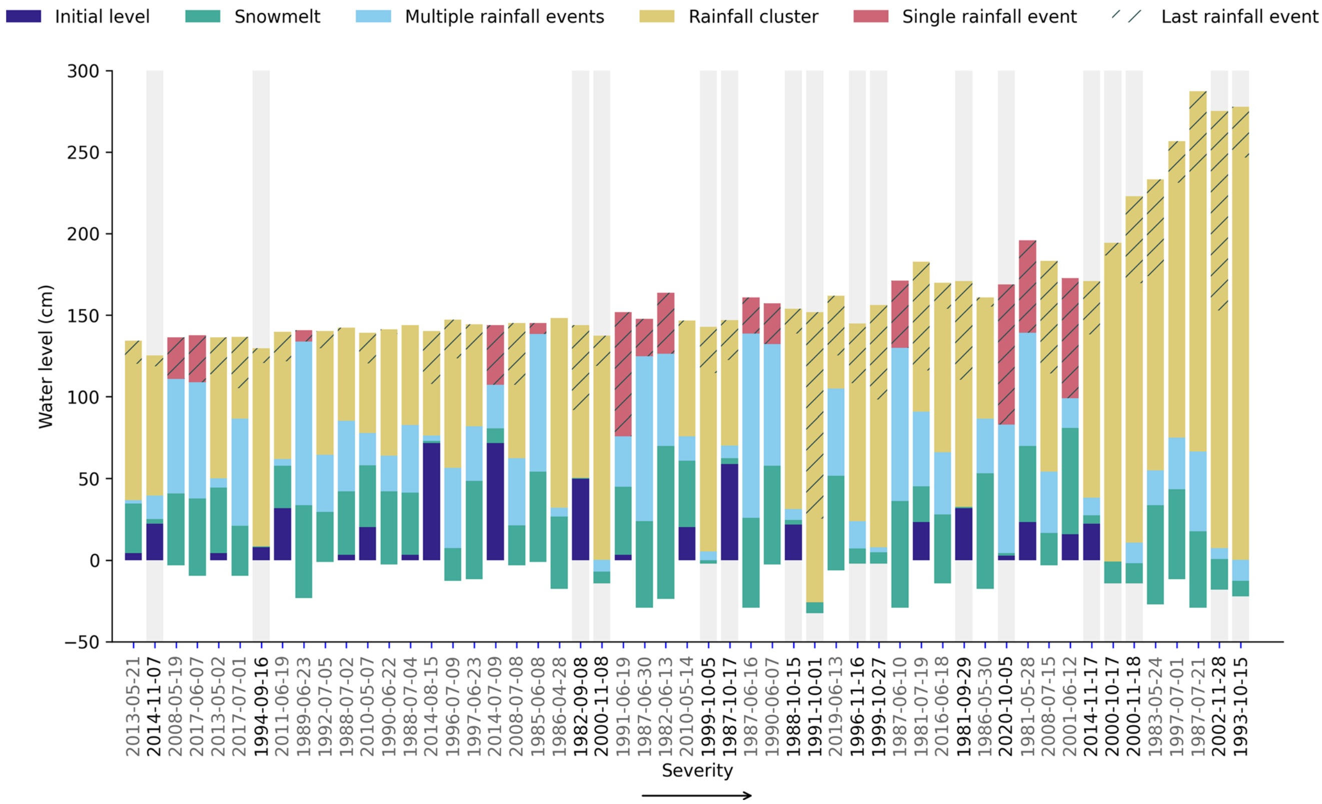

Fig. 4 Triggers of lake flood events. Barplot of the triggers of lake floods. The height of each block of a bar represents the increase of water level caused by that factor and the total height of a bar represents the difference between the water level at the flood peak and the minimum water level at the beginning of the season. Autumn events are indicated with a grey vertical bar. The position of the bottom of a bar with respect to the zero on the *y*-axis represents how much the initial water level was higher/lower than the average initial level, computed separately for spring and autumn events; if a bar starts below zero, it means that the initial level of the event was below the average one, if it is above zero, the difference between the initial level of the event and the average one is included in the initial level bar. The hatch with black lines identifies the contribution of the last rainfall event occurred before the date of the flood peak; this may be part of a rainfall cluster or an individual event. Events are in ascending order of severity, computed as the area contained between the water level during the duration of the flood event and the threshold level for flooding.

and multiple rainfall events; less intense rainfall events, also not clustered, are then enough to trigger flood events. In spring, the contribution of the temporal clustering of rainfall events to the final level is, in fact, not predominant with respect to the other contributions, in most of the cases. The only autumn event without temporal clustering of rainfall (i.e., 2020-10-05 flood) can be associated with the occurrence of two main rainfall events: one at the beginning of the season that increased the water level and one that triggered the flood event.

While the temporal clustering of rainfall events is less important in the onset of spring events than of autumn events, the three most severe spring events are the ones characterized by a higher contribution of this driver. Hence, the rainfall clustering appears to be the most important factor in determining the impact in both spring and autumn floods.

Temporal clustering of rainfall events and its significance as a driver. Given the importance of rainfall clustering in driving the severity of flood events, we here concentrate on this specific trigger. As reported in section 'Temporal clustering of rainfall events', the presence of temporal clustering of rainfall was assessed using two different precipitation data sets, ERA5-Land and E-OBS. As discussed in section 'Potential triggers of flood frequency and severity', the presence of the temporal clustering is a recurrent feature in both autumn and spring events. Using ERA5-Land data set, 64 and 94% of spring and autumn flood events are preceded by temporal clustering of rainfall events, while using E-OBS the percentages are 61 and 88% in spring and autumn. The different agreement of the two data sets may be in part due to the different approaches with which rainfall series was obtained from precipitation one. ERA5-Land, in fact, provides an estimation of snowfall along with precipitation, while for E-OBS the fraction of liquid precipitation was inferred by temperature data alone.

The significance of the temporal clustering of rainfall events as a driver of lake floods was assessed by reshuffling (1) the series of flood events; and (2) the series of rainfall events, as explained in section 'Temporal clustering of rainfall as driver of lake floods'. Considering a confidence level of 0.95, the association of rainfall cluster to flood events is significant for both data sets, the two resampling methods, and for spring, autumn, and all events.

Importance of compoundness in flood occurrence. Once the presence of compoundness in the generation of lake floods is assessed, we asked whether the water level increase due to the last rainfall event was enough to bring it above the threshold level also without the contribution of the other factors. Considering an average initial lake level, in none of the events was the last rainfall event alone enough to cause flooding. A compoundness of factors is, therefore, necessary for the occurrence of lake floods.

In order to further investigate the connections characterizing lake floods, we looked whether in case of multiple floods in the same year and season the first hazard influenced the following one (see section 'Compoundness of hazards'). In the 40 years

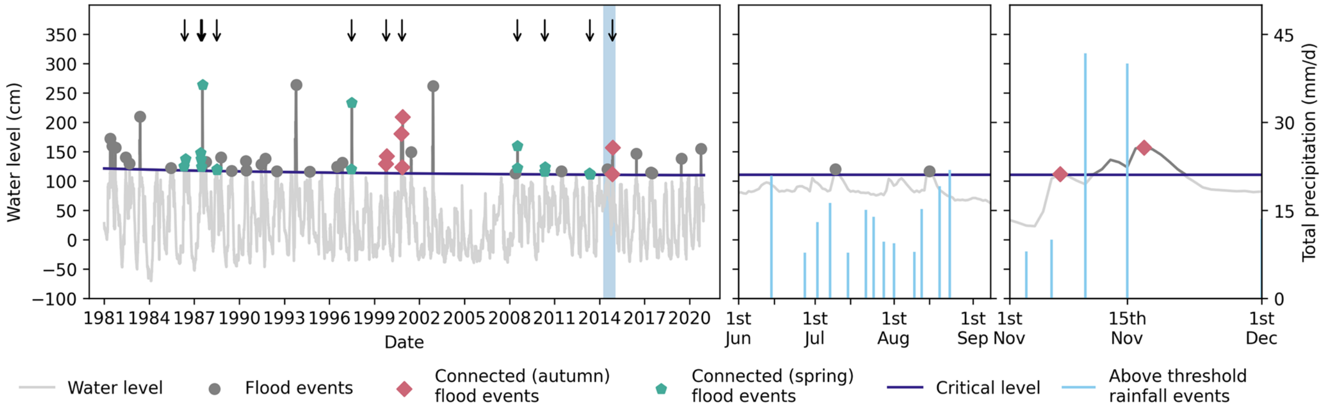

Fig. 5 Series of lake level and flood events. On the left, the lake level series with the maximum peak of flood events highlighted by dots. The purple line indicates the flooding level. In case of a connection between flood events, an arrow is drawn pointing to it. On the centre a zoom over the spring of 2014 during which two flood events occurred, but no influence of the first one on the second one was observed. On the right a zoom over the autumn of 2014 during which two flood events occurred, and a connection between them was identified.

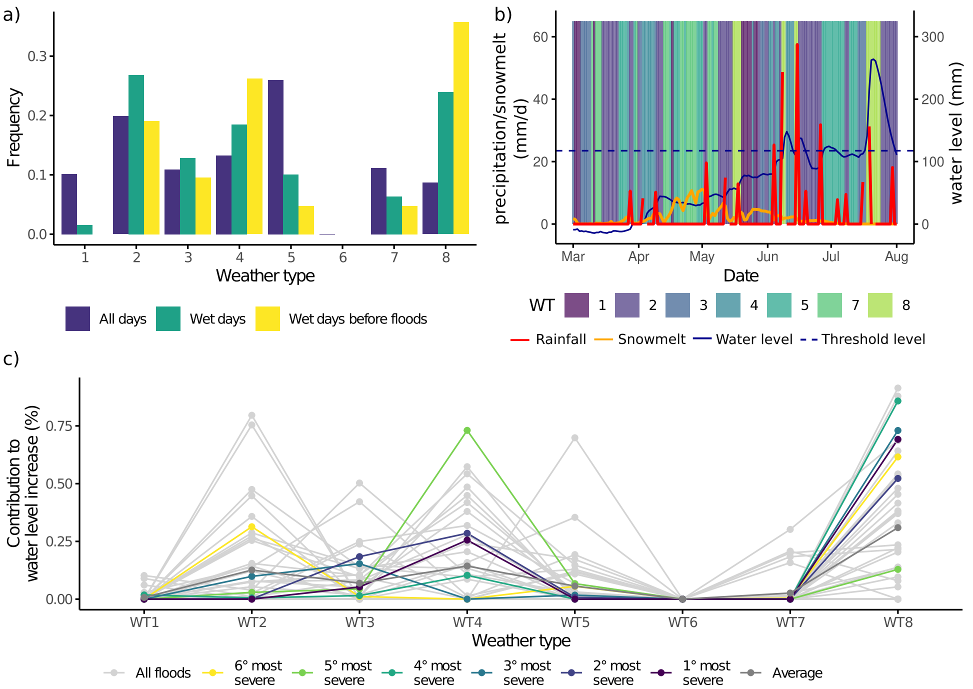

Fig. 6 Analysis of weather patterns associated with the rainfall events triggering lake floods. a Frequency of the weather circulation types, identified by the LaMMa Consortium, during all days in the record, wet days, and days in which the last rainfall event preceding a flood occurred. **b** Daily total rainfall and snowmelt over the basin and lake water level during the period preceding the flood of July 1987. The horizontal dashed blue line represents the water level above which flooding of Como occurs. The colours in the background identify the daily weather patterns. **c** Contribution of the eight weather types in the increase of water level due to rainfall clustering for each flood event. The six most severe events are highlighted along with the average frequency.

series of water level, we observed, in the autumn period, 67.5% of years without floods, 25% with one flood, 5% with two floods (i.e., 1999 and 2014), and 2.5% with three floods (i.e., 2000). Three years had, therefore, multiple events, and we found a connection between hazards in all three of them (Fig. 5). In the spring/summer period, the percentages are 45, 27.5, 22.5, 2.5, and 2.5% for zero to four flood events. Regarding years with multiple events, in 5 out of 9 of the years with two flood events we found a connection between hazards. In the one year with three events (i.e., 2008) we found a connection between two of the three events, and finally, in the one year with four events (i.e., 1987) we found a connection between the first and second event and between the third and fourth one.

Weather patterns driving lake floods. The frequency of the eight weather circulation types (WT1, WT2, WT3, WT4, WT5, WT6, WT7, and WT8), identified by the LaMMa Consortium (see section 'Weather patterns driving lake floods'), is reported in Fig. 6, distinguished in three cases: (a) all days, (b) wet days, and (c) days in which the last rainfall event preceding a flood occurred. We can observe important differences in the frequencies between the three cases. In particular, WT1, WT5, and WT7, while frequently observed in the record, are not associated with a high frequency of wet days. WT5, with a percentage of 25% between all days, is present only 5% of the time during wet days. These weather types are characterized by an anticyclonic circulation, that brings dry conditions over Italy.

Regarding wet days before floods events, the predominant weather types are WT8 and WT4, followed by WT2. WT8 is associated with a cyclonic circulation over west Europe with a ridge over the eastern Mediterranean, resulting in abundant precipitation over Northern Italy. WT4 is characterized by a subtropical high pressure that protects only southern and central regions of Italy, leaving northern Italy exposed to perturbations. Finally, WT2 is associated with a partial displacement of the Azores High Pressure to the Northern Atlantic Ocean, thus allowing maritime polar air masses to reach Central Europe and partly the Mediterranean area. These results are in line with the ones of Messeri et al.[33], that found WT8 to be the most impactful meteorological configuration for Italy, regarding flood and landslide events, and they found a moderate risk for Lombardy (i.e., the region where Lake Como is located) associated with WT4.

Given the importance of temporal clustering of rainfall in lake flood events, we focused in greater details on the weather patterns associated with this trigger. This is shown in Fig. 6c, where the increase of lake water level due to rainfall clustering is subdivided between the different weather types. Results are quite variable between the events, but we can still observe a predominance of WT8, WT4, and WT2. Important contributions from WT5 were also observed. Considering only the six most severe lake floods, WT8 is the most important weather type in all but one, with the highest contribution of WT4. From these results, a contribution of WT8 well above the average appears to be to configuration generating the most severe floods.

The example of the 1987-07-21 lake flood, the third most severe event in the record, is reported in Fig. 6b. This event is part of a more extended meteorological event that caused widespread flooding in Valtellina and the destructive Val Pola Landslide. The first part of July was characterized by an anticyclonic condition (WT5), that evolved into a blocked type situation (WT8). A low pressure area persisted over the English Channel until the 21th of July. The synoptic condition caused cold air from the Atlantic Ocean to flow towards the Central alpine region, and the thermal contrast with the warm and moist air, related with the persistence of the anticyclone in the preceding weeks, favoured an intense convective activity. The meteorological system resulted in maximum values of precipitation in the Alpine region, with lower precipitation in the Po valley[34]. The November 2000 and November 2002 floods are also part of extended meteorological events, documented in literature[35,36], due to their impacts on other Northern Italian regions or due to other natural hazards triggered. For example, the persistent precipitation that triggered the November 2002 flood, also triggered the Bindo-Cortenova deep landslide in the Como basin. These events were related with a persistent cyclonic circulation over Italy that resulted in a sustained moisture advection from the North Atlantic Ocean and in a northward flow towards the Alpine region[35,36].

Discussion

Here, we assessed a series of lake floods, affecting the town of Como, in a compound prospective. Respect to this, we have identified three fundamental typologies in lake flood occurrence: temporal clustering of rainfall events, rainfall clustering over an increased water level, single rainfall event over an increased water level. The univariate case, i.e., a single event without a previous increase in water level, was never observed, questioning, therefore, the classical approach used in flood frequency analysis. These three typologies, then, combine together in case of connected flood events. The presence of more than two connected flood events, however, was observed only when flood was triggered by a cluster of rainfall. Temporal clustering of rainfall emerged as an important factor in the severity other than in the occurrence of floods, both in spring and autumn. This poses rainfall cluster and its prediction as the most important factor in the reduction of the impacts associated with lake floods at Como.

From these results, we can conclude that all the events can be classified in one or more of the four typologies of compound events. The multivariate and temporal compoundness were found, respectively, in the co-occurrence of melting and rainfall events and in the temporal clustering of rainfall. The preconditioned typology can manifest both through the initial basin saturation, that increases rainfall-runoff ratio but was not considered here, and the presence of the snow cover that, when it melts, increases the total runoff. As a proxy for rain-on-snow events, we computed the contribution of melting to water level increase between the date of the last rainfall event and the date of the flood peak, and we found that it constituted a limited percentage both with the respect to the total contribution of melting and to the contribution of last rainfall event. Spatial compoundness can be important for lakes characterized by multiple inlet streams, since the resulting increase in lake level will likely be more severe if all the inflows reach peak flow close in time. In the case of Lake Como, only two inflows enter the lake, Adda and Mera rivers. For the flood events in correspondence of which inflow measurements were available (i.e., 25 events), we observed the occurrence of a peak in the water level in both Adda and Mera rivers. The peak occurred in the same day 44% of the time, with a 1 day delay 52% of the time, and with 2 days delay 8% of the time. We, therefore, did not investigate deeper this aspect, since it was a common feature of all events. However, further studies could explore whether the presence of concurrent peaks is required for the lake flood to occur or this situation commonly occurs, also in absence of flooding.

Eventually, we explored the weather patterns associated with both single rainfall events and rainfall clusters triggering lake floods. We found that the most critical synoptic situation is the one chracterized by a cyclonic circulation over west Europe with a ridge over the eastern Mediterranean. This is in line with the results found by Messeri et al.[33] for floods and landslides risk in Italy.

A limitation of this study is the rather simple approach used in order to associate the contribution of each trigger to flood events, in particular regarding the subdivision between rainfall events and snowmelt. The two drivers, in fact, often occur together, and in some cases, it is rainfall itself that triggers snowmelt, like in rain-on-snow events. An improvement of the work could be, for example, to use a hydrological model in order to subdivide between the two contributions. This would also permit the inclusion of basin saturation effects in the analysis. The amount of generated run-off from a rainfall event, in fact, that determines the water contributing to the increase in lake level, varies with the soil moisture of the basin.

From this study, two main characteristics of lake flood events emerge: (1) the sequential or concurrent combination of different climate-based factors that drive the process of flooding and (2) the possible presence of connected flood events. Even though the presented study focuses on a specific case study, i.e., Como lake, the aim of the paper is, more broadly, to point the attention on lake floods and on the importance of investigating them in a compound perspective. These results are a first step in the study of this particular typology of floods, related to lake level fluctuations, and the presented analysis may be seen as a guideline to investigate them further. Generalizing it, we propose the following four steps:

1. Selection of potential drivers: this was done by looking at the existing case studies of lake floods and by looking at

similar events, i.e., river floods. We propose to test, as a first attempt, the following variables: rainfall, snowmelt, inflows discharge, and outflows discharge. For a more systematic collection of potential variables, data mining techniques on scientific literature can also be adopted. This may permit to understand if other influencing factors may play a role in other basins;
2. Investigation of drivers' compoundness: here the connections and dependencies inside and between drivers are investigated. For example, the occurrence of temporal clustering of rainfall, rain-on-snow events, or the dependence between inflows discharge;
3. Association between hazards and drivers: here the identified drivers are checked for their causal links with hazards. In this work, we looked if the drivers were statistically significant with a resampling procedure. Causal inference methods could be used to properly asses causal links[37] as well as factorial experiments with hydrological models;
4. Investigation of hazards' compoundness: here the connections between hazards are assessed. Starting from the work of De Angeli et al.[38] and Zscheischler et al.[6], we report three possible connections that may exist between lake floods: (a) hazards are triggered by the same cause, i.e., the same rainfall cluster (spurious dependence between the hazards), (b) a first hazard alters the likelihood/magnitude of a second hazard, i.e., by increasing the water level (disposition alteration; investigated in this work), and (c) the occurrence of one hazard changes the vulnerability of the exposed elements that are then affected by a second hazard, i.e., the second flood occurs when there is still water in the town, therefore prolonging the duration of the exposure of buildings to water (dependence at the impact level).
5. Attribution of synoptic conditions to the identified compound event: the previous steps permit to asses the whole compound event (typologies in Fig. 3b). This final step consists of associating to each event the meteorological situation in order to properly understand the associated risk.

Even though the presented steps are designed for lake floods, they may also benefit the investigation of other typologies of flood, providing a core guideline with which to address compound flooding.

The application of these steps at a global scale will permit to asses if and how the conclusions here presented may be generalized and if other influencing factors may play a role determining lake floods impacts. This would, for example, permit to make a classification of existing lakes depending on the mechanisms of occurrence of lake floods and to identify connections between water level fluctuations of lakes in different regions, due to connections in their triggers. These conclusions would help to manage the risk associated with lake floods. Eventually, this may help in assessing the evolution of this natural hazard under climate changes, by combining the changes in the identified driving factors and their dependence. Looking at the evolution of the drivers to predict changes in the hazard is particularly important in the context of lake floods, given that the majority of lakes are regulated, thus a possible trend in lake flood frequency may be masked by the adopted lake management.

Methods

Data. Precipitation series for each cell of the basin, feeding Lake Como, was obtained using two different data sets: (a) ERA5-Land[39] and (b) E-OBS[40]. The former is a reanalysis product while the latter was obtained interpolating observational data. Both data sets have a resolution of 9 km × 9 km and can be downloaded from Copernicus Climate Change Service (C3S) Climate Data Store. For the analysis, we considered the period 1981–2020. In order to make the comparison consistent, ERA5-Land grid was resampled on E-OBS grid using nearest neighbours. From the precipitation series, the series of rainfall was obtained (a) for ERA5-Land, subtracting the series of snowfall provided by the same data set, (b) for E-OBS, we set to zero all precipitation values corresponding to temperatures lower than 1 °C following the results of Jennings et al.[41]. The temperature series over the basin was obtained from E-OBS data set. To better understand the flooding drivers, we also used snowmelt and snow cover series of ERA5-Land.

The series of water level of Lake Como was obtained from the hydrometer of Malgrate, situated near the outlet of the lake. Data were provided by Consorzio dell'Adda. The reference level of the hydrometer for which flooding of Como starts changed over time, with a value of 1.86 m in 1955, 1.25 m in 1978, 1.20 m in 2003, and a current value of 1.10 m[42]. The threshold for all the other years was reconstructed from these known points with an interpolation.

The series of water level of the two tributaries of Lake Como were also used in the study, to look for the co-occurrence of peaks in the two inflows before lake floods occurrence. The data used for the Mera river were measured at the hydrometer of Samolaco, controlled by Consorzio dell'Adda. Data were available for the period 1994 to present, with a variable resolution, and they were aggregated to a daily time step. The data used for the Adda river were measured at the hydrometer of Fuentes, controlled by Arpa Lombardia. Levels were available from 1971 with a variable resolution, and they were aggregated to a daily time step. In both cases, missing values were not replaced, and daily aggregates were computed only if all the sub-daily values were available and of good quality.

Lake flood events and associated triggers. We identified flood events as all the exceedances of water level above the critical level, obtained as reported in section 'Data'. We separated events with an interval time of 1 day and we associated to each event the date of the maximum peak.

In order for a flood to occur, a water level gap needs to be filled, from the initial level up to the threshold level. This gap can be filled by multiple sparse rainfall events, melting, cluster of rainfall events, a single rainfall event, or a combination. Besides, an initial level higher, or lower, than the average may increase or decrease the size of the gap to be filled. For each event we identified the initial level given, (a) for spring events, by the lowest water level before the onset of melting, usually occurring between February and March, and, (b) for autumn events, by the lowest water level after the reduction of water level in August for agricultural needs. The average initial water level for the spring and autumn period was also calculated along with the difference, if positive, between the initial level of each event and the average one. To compute the increase of water level due to the different triggers, we first identified the presence of cluster before each event and the maximum time window before the event over which it was significant (see section 'Temporal clustering of rainfall events' for the details about the identification of clustering). Then, we computed the daily sum of melting and rainfall over the basin. For each day we computed the variation in water level. When this was positive the variation was assigned to melting and rainfall proportionally to their daily sum over the basin. If temporal clustering of rainfall was significant the increase due to rainfall was associated to rainfall cluster, otherwise to multiple sparse rainfall events. When negative, the variation was removed from the triggers proportionally to their current contribution. The variation due to the last rainfall event before the day of the flood was assigned to the category single rainfall event when it was not part of a rainfall cluster.

Temporal clustering of rainfall events. The temporal clustering of rainfall was identified adopting the methodology proposed by Bevacqua et al.[43], with some modifications. The approach requires a series of independent exceedances over a threshold, therefore the high frequency clustering was firstly removed running a (high frequency) declustering procedure[44]: (1) a series of exceedances over a threshold u is extracted from the rainfall series, (2) the exceedances separated by less than r days are clustered together, and (3) for each cluster only the first exceedance is retained and the other events are replaced with NA. From the series of independent events the associated probability of exceedance, p, was also computed, disregarding the days with a rainfall equal to NA, i.e., the days in which rainfall events were removed during the (high-frequency) delcustering procedure.

The value of r was set equal to 2 days following the results of Barton et al.[45] and the threshold u was set to the rainfall quantile with a probability of 80%. The threshold was chosen lower than the one typically adopted for the study of river floods because the two processes are different; lake flood occurs when the water level in the lake passes a given threshold, occurring when the inflows exceed the capacity of the outflows to balance them. Already for the study of rainfall-triggered landslides, Bevacqua et al.[43], adopted a threshold corresponding to a quantile level of 0.7, considering only wet days, moved by the consideration that the process is one of saturation, occurring over large windows. Therefore, repeated but less intense rainfall events may better explain this. In addition, we are here studying temporal clustering, and also less intense rainfall events, if clustered, may be able to cause extreme impacts.

The presence of temporal clustering is assessed over a time window w, ending the day of the flood event, by testing whether the number of exceedances of the declustered series, inside w, follows a binomial distribution with size w_{eff} and probability p. If the number of exceedances is higher than what expected from the binomial distribution, with a 5% significance level, the null hypothesis or the absence of cluster in the window w is refused. Here, w_{eff} is the effective window that takes into account the effect of the (high-frequency) declustering on the low frequency clustering and it is equal to w minus the days with rainfall equal to NA. Here, we tested different time windows w varying from 4 days to 45 days. The minimum window is set to 4 days, because, after high-frequency declustering, events are distant at least 2 days, so a window larger than 3 days is required to have more than one event.

The presence of temporal clustering was tested on the rainfall series summed over the basin feeding Lake Como. In total we had a number of tests equal to the number of windows × number of events, i.e., 2226 tests. A multiple testing correction was therefore required to keep the overall significance at 0.05. In order to take into account the discreteness of the p-values we adopted the Benjamini–Hochberg (BH) procedure[46] applied on mid-p-values[47].

Temporal clustering of rainfall as driver of lake floods. Following Bevacqua et al.[43], the significance of temporal clustering of rainfall events as a driver of lake floods was assessed using the bootstrapping technique. Here in particular, we resampled the series of rainfall, and the series of flood events as follows: we resampled (a) the rainfall events between all the years but keeping each month separated; (b) we resampled flood events between all the years but keeping separate spring/summer (from April to August) and autumn (from September to December) periods. In both cases, we run 5000 simulations, and we computed the 0.95 quantile of the number of events with temporal clustering. We then compared it with the number of events with cluster without reshuffling, in order to test the significance of rainfall cluster as a driver of lake floods.

Compoundness of hazards. Multiple hazards can be connected in several ways: they can be triggered by a common cause, they can be a cascading of hazards when one triggers the other, or one hazard can alter the predisposition for the second one. Here, we focused on seasons with multiple flood events, and we looked for the presence of this last type of connection. To achieve this, we identified the cases in which the occurrence of the first hazard resulted in a higher level that persisted up to the occurrence of the following one. In particular, we identified a connection between two events, occurring in the same season and year, if the lake level after the first flood event did not lower to the pre-event lake level before the occurrence of the second flood event. The pre-event lake level was computed as the level of the lake the day of the beginning of the single rainfall event or cluster of rainfall triggering the first flood. As reported above, different types of connections may exist between the hazards, and for this two lake floods may or may not be linked depending on which aspect we are considering.

Combining all the information about drivers and hazards, it was possible to create for each flood event or group of flood events, when connected, a diagram showing the relations between the variables. A direct link connecting two consecutive floods was drawn in case of a relation, of the type specified above, between the hazards. Regarding drivers, we always reported once each driver, also when it was a trigger of more than one flood event. In this case, the measure in which the driver contributes to the occurrence of the different flood events may not be the same.

Weather patterns driving lake floods. To perform a meteorological analysis of the drivers, we used the classification of weather patterns and circulation types (WT) developed by the LaMMa Consortium using COST 733 methodology[48,49]. They identified eight classes, focusing on the main types of circulation over Italy. The definition of the eight WTs is reported below[33]:

- WT1: Marked northward expansion of the Azores anticyclone with blocked anticyclonic circulation over the North Atlantic and northerly winds over Italy.
- WT2: Moderate northward expansion of the Azores anticyclone with cyclonic circulation over south Scandinavia and northwesterly winds over Italy.
- WT3: Marked cyclonic circulation over Iceland with anticyclonic circulation over northern central Europe accompanied with increased precipitation over Italy, generated by intermittent Atlantic perturbations.
- WT4: Cyclonic circulation over the North Atlantic and cyclonic circulation over west Mediterranean Europe and central Mediterranean Europe with decreased precipitations over central Mediterranean Europe.
- WT5: Cyclonic circulation over the north-west Atlantic with marked anticyclonic circulation over west Mediterranean Europe and central Mediterranean Europe, inducing warm and dry conditions over Italy.
- WT6: Anticyclonic circulation over Iceland and cyclonic circulation over central Europe, with higher precipitation over Tuscany fuelled by intrusions of Arctic and polar continental air.
- WT7: Southwesterly flow over the North Atlantic with ridging over the British Isles towards Scandinavia, with easterly wind over central Mediterranean Europe resulting in very cold dry conditions.
- WT8: Cyclonic circulation over West Europe with a ridge over the eastern Mediterranean.

Daily weather types, from 1948 to 2010, are provided by Messeri et al.[33] and they were used here.

The investigation focuses on two aspects. First, we looked at the rainfall events closest to each floods and the associated weather type. The frequency of each weather type was compared to the same frequency computed considering days with precipitation above the threshold and all days in the record, both wet and dry. Second, we moved to temporal clustering of rainfall. For each event, we computed how much of the water level increase due to rainfall clustering was associated with each of the weather types. This, to understand the most important weather types during rainfall clusters and the ones causing the most severe flooding.

Data availability

ERA5-Land[39] and E-OBS[40] data sets are freely downloadable from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/), upon registration. The hydrometer at Malgrate and Samolaco are controlled by Consorzio dell'Adda and data are freely available upon request (https://www.addaconsorzio.it). The hydrometer at Fuentes is controlled by Arpa Lombardia. Historical data can be downloaded from https://idro.arpalombardia.it and are also available upon request. Daily weather types are provided by Messeri et al.[33] and available at https://github.com/meteosalute/weather_landslide. The dataset containing the main results of this study is available at https://doi.org/10.5281/zenodo.6838802.

Received: 4 January 2022; Accepted: 20 September 2022;
Published online: 08 October 2022

References

1. Leonard, M. et al. A compound event framework for understanding extreme impacts. *WIREs Climate Change* **5**, 113–128 (2014).
2. Zscheischler, J. et al. Future climate risk from compound events. *Nat. Clim. Change* **8**, 469–477 (2018).
3. Aghakouchak, A. et al. Climate extremes and compound hazards in a warming world. *Annu. Rev. Earth Planetary Sci.* **48**, 519–548 (2020).
4. Raymond, C. et al. Understanding and managing connected extreme events. *Nat. Clim. Change* **10**, 611–621 (2020).
5. Seneviratne, S. et al. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)* (eds Field, C.) 109–230 (Cambridge University Press, 2012).
6. Zscheischler, J. et al. A typology of compound weather and climate events. *Nat. Rev. Earth Environ.* **1**, 333–347 (2020).
7. Shankman, D. et al. Hydroclimate analysis of severe floods in china's Poyang lake region. *Earth Interactions* **16**, 1–16 (2012).
8. Xue, J.-B. et al. Potential impact of flooding on schistosomiasis in Poyang lake regions based on multi-source remote sensing images. *Parasites Vectors* **14**, 1–13 (2021).
9. ILCRRSB (International Lake Champlain-Richelieu River Study Board). The causes and impacts of past floods in the lake Champlain-Richelieu river basin: Historical information on flooding. Tech. Rep. https://ijc.org/en/lcrr/causes-and-impacts-past-floods-lake-champlain-richelieu-river-basin-historical-information (2019).
10. Grahn, T. & Nyberg, R. Damage assessment of lake floods: Insured damage to private property during two lake floods in Sweden 2000/2001. *Int. J. Disaster Risk Reduction* **10**, 305–314 (2014).
11. Wang, L., Han, Y., Yu, H., Fan, S. & Liu, C. Submerged vegetation and water quality degeneration from serious flooding in liangzi lake, china. *Front. Plant Sci.* **10**, 1504 (2019).
12. Ma, Y. et al. Increasing water levels of global lakes between 2003 and 2009. *IEEE Geosci. Remote Sensing Lett.* **17**, 187–191 (2020).
13. Shugar, D. H. et al. Rapid worldwide growth of glacial lakes since 1990. *Nat. Clim. Change* **10**, 939–945 (2020).
14. Ward, P. J. et al. Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. *Environ. Res. Lett.* **13**, 084012 (2018).
15. Bevacqua, E. et al. More meteorological events that drive compound coastal flooding are projected under climate change. *Commun. Earth Environ.* **1**, 1–11 (2020).
16. Jane, R., Cadavid, L., Obeysekera, J. & Wahl, T. Multivariate statistical modelling of the drivers of compound flood events in south Florida. *Nat. Hazards Earth Syst. Sci.* **20**, 2681–2699 (2020).
17. Saharia, A. M., Zhu, Z. & Atkinson, J. F. Compound flooding from lake seiche and river flow in a freshwater coastal river. *J. Hydrol.* **603**, 126969 (2021).
18. Bender, J., Wahl, T., Müller, A. & Jensen, J. A multivariate design framework for river confluences. *Hydrol. Sci. J.* **61**, 471–482 (2016).

19. De Michele, C. & Salvadori, G. On the derived flood frequency distribution: Analytical formulation and the influence of antecedent soil moisture condition. *J. Hydrol.* **262**, 245–258 (2002).
20. Zhong, M. et al. A copula-based multivariate probability analysis for flash flood risk under the compound effect of soil moisture and rainfall. *Water Resour. Manag.* **35**, 83–98 (2021).
21. Khatun, A., Ganguli, P., Bisht, D. S., Chatterjee, C. & Sahoo, B. Understanding the impacts of predecessor rain events on flood hazard in a changing climate. *Hydrol. Processes* **36**, e14500 (2022).
22. Manoj J, A., Guntu, R. K. & Agarwal, A. Spatiotemporal dependence of soil moisture and precipitation over India. *J. Hydrol.* **610**, 127898 (2022).
23. Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M. & Vrac, M. Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrol. Earth Syst. Sci.* **21**, 2701–2723 (2017).
24. Kopp, J., Rivoire, P., Ali, S. M., Barton, Y. & Martius, O. A novel method to identify sub-seasonal clustering episodes of extreme precipitation events and their contributions to large accumulation periods. *Hydrol. Earth Syst. Sci.* **25**, 5153–5174 (2021).
25. Serafin, K. A., Ruggiero, P., Parker, K. & Hill, D. F. What's streamflow got to do with it? a probabilistic simulation of the competing oceanographic and fluvial processes driving extreme along-river water levels. *Nat. Hazards Earth Syst. Sci.* **19**, 1415–1431 (2019).
26. Moftakhari, H., Schubert, J. E., AghaKouchak, A., Matthew, R. A. & Sanders, B. F. Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Adv. Water Resour.* **128**, 28–38 (2019).
27. Riboust, P. & Brissette, F. Analysis of lake champlain/richelieu river's historical 2011 flood. *Can. Water Resour. J./Revue canadienne des ressources hydriques* **41**, 174–185 (2016).
28. Bajni, G., Apuani, T. & Beretta, G. P. Hydro-geotechnical modelling of subsidence in the como urban area. *Eng. Geol.* **257**, 105144 (2019).
29. Guariso, G., Rinaldi, S. & Soncini-Sessa, R. The management of lake como: A multiobjective analysis. *Water Resour. Res.* **22**, 109–120 (1986).
30. Poggi, C. & Cantù, C. Como e le Piene più Famose del Lago (Enzo Pifferi editore, Como, Italy, 2000). [In Italian].
31. Ricci, F. La Terra Segnata. Calamità Naturali Storiche nella Provincia di Como (Editoriale, Como, Italy, 2001). [In Italian].
32. Comerci, V. et al. Land subsidence and late glacial environmental evolution of the como urban area (northern Italy). *Quat. Int.* **173–174**, 67–86 (2007).
33. Messeri, A. et al. Weather-related flood and landslide damage: A risk index for Italian regions. *PLoS One* **10**, 1–17 (2016).
34. Buizza, R. & Pelosini, R. Rainfall prediction with a limited-area model during an intense storm over northern Italy. *Il Nuovo Cimento C* **17**, 665–681 (1994).
35. Winschall, A., Pfahl, S., Sodemann, H. & Wernli, H. Impact of north Atlantic evaporation hot spots on southern alpine heavy precipitation events. *Quart. J. R. Meteorol. Soc.* **138**, 1245–1258 (2012).
36. Pinto, J. G. et al. Identification and ranking of extraordinary rainfall events over northwest Italy: The role of Atlantic moisture. *J. Geophys.Res.: Atmos.* **118**, 2085–2097 (2013).
37. Runge, J. et al. Inferring causation from time series in earth system sciences. *Nat. Commun.* **10**, 1–13 (2019).
38. De Angeli, S. et al. A multi-hazard framework for spatial-temporal impact analysis. *Int. J. Disaster Risk Reduction* **73**, 102829 (2022).
39. Muñoz Sabater, J. Era5-land hourly data from 1981 to present. copernicus climate change service (c3s) climate data store (cds). https://doi.org/10.24381/cds.e2161bac (2019).
40. Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M. & Jones, P. D. An ensemble version of the e-obs temperature and precipitation data sets. *J. Geophys. Res.: Atmos.* **123**, 9391–9409 (2018).
41. Jennings, K. S., Winchell, T. S., Livneh, B. & Molotch, N. P. Spatial variation of the rain–snow temperature threshold across the northern hemisphere. *Nat. Commun* **9**, 1148 (2018).
42. Consorzio dell'Adda. Gli effetti della regolazione sulle portate dell'adda e sulle piene del lago di como. Tech. Rep. Pubblication n° 12, Milano (2003). [In Italian] https://www.addaconsorzio.it/wp-content/uploads/2021/07/pubblicazione12.pdf.
43. Bevacqua, E. et al. Guidelines for studying diverse types of compound weather and climate events. *Earth's Future* **9**, e2021EF002340 (2021).
44. Coles, S. *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001).
45. Barton, Y. et al. Clustering of regional-scale extreme precipitation events in southern switzerland. *Mont. Weather Rev.* **144**, 347–369 (2016).
46. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.: Ser. B (Methodological)* **57**, 289–300 (1995).
47. Heller, R. & Gur, H. False discovery rate controlling procedures for discrete tests (2011). Preprint at https://arxiv.org/abs/1112.4627.
48. Philipp, A., Beck, C., Huth, R. & Jacobeit, J. Development and comparison of circulation type classifications using the cost 733 dataset and software. *Int. J. Clim.* **36**, 2673–2691 (2016).
49. Salinger, M. J. et al. Seasonal differences in climate in the chianti region of Tuscany and the relationship to vintage wine quality. *Int. J. Biometeorol.* **59**, 1799–1811 (2015).

Acknowledgements
We would like to acknowledge Consorzio dell'Adda and Arpa Lombardia for the hydrological data used in this study. In addition, we thanks Joaquim G. Pinto for its input on the meteorological interpretation of the results. The work was partially funded by the Italian Ministry of Research through the PRIN project RELAID.

Author contributions
Fabiola and Carlo conceived the idea. Fabiola collected data and made the statistical analyses. Fabiola and Carlo discussed the results. Fabiola prepared figures and wrote the first draft of the manuscript. Carlo has reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Fabiola Banfi or Carlo De Michele.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Joy Merwin Monteiro, Joe Aslin