Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

IL FUTURO DELLA TERRA, CAMBIAMENTI CLIMATICI E SFIDE SOCIALI

Ciclo XXXV

Settore Concorsuale: 02/C1 – Astronomia, Astrofisica, Fisica della terra e dei pianeti

Settore Scientifico Disciplinare: FIS/06 – Fisica per il sistema terra e il mezzo circumterrestre

SPHERA, A NEW ITALIAN CONVECTION-PERMITTING REGIONAL REANALYSIS: IMPROVING HEAVY-RAINFALL REPRESENTATION AND DESCRIBING HAIL-PRONE ENVIRONMENTS

Presentata da: Antonio Giordani

Coordinatore Dottorato

Prof. Dr. Silvana Di Sabatino

Supervisore

Dr. Tiziana Paccagnella

Co-supervisori

Prof. Dr. Silvana Di Sabatino

Dr. Ines Maria Luisa Cerenzia

Esame finale anno 2023

Abstract

Deep moist convection (DMC) is an atmospheric phenomenon leading to high-impact severe weather events such as heavy rainfall, wind gusts, flash floods, tornadoes, or hail. These events produce critical damages to the environment, infrastructures, and properties globally every year, generating enormous economic losses, and threatening human and animal health. From a climatological perspective, anthropogenic global warming is expected to induce increases in the frequency and severity of extreme events, hence there is an urgent need to advance their understanding. The occurrence of extreme events can be simulated through numerical weather prediction models. However, convection is characterized by chaotic dynamics, often implying reduced predictability of the associated processes. Additionally, models employ physical parameterization schemes to account for sub-grid processes such as convection initiation, which, however, still have substantial limitation. DMC parameterizations can be switched off in numerical simulations with refined spatial resolutions reaching the convection-permitting scales. These allow for the explicit representation of most convective motions, thus producing notable improvements. To enhance understanding and representation of the present and future atmospheric extremes, in-depth knowledge of the past is also necessary. Backward investigations of weather and climate are possible through atmospheric reanalyses obtained by combining numerical simulations with the available observations and representing the most complete picture of the past atmosphere.

For these reasons, SPHERA, a new convection-permitting regional reanalysis over Italy, has been produced at ARPAE-SIMC. SPHERA is obtained by dynamically downscaling the global reanalysis ERA5 by the model COSMO ran at 2.2 km grid spacing, and by continuously ingesting regional observations through a nudging data assimilation scheme over 1995-2020. SPHERA has been conceived to enhance the description of the past atmosphere over Italy and neighboring countries and is expected to improve the representation of weather processes linked to DMC. Hence, this thesis aims to verify the expectations placed on SPHERA by analyzing two weather phenomena that are particularly challenging to simulate: heavy rainfall and hail.

Concerning the reproduction of precipitation, a quantitative statistical analysis over Italy during 2003-2017 is carried out to compare the performance of SPHERA with its driver ERA5, particularly in case of heavy accumulations. Considering the national network of rain-gauge observations as a reference, several characteristics are investigated, such as the geographical distributions of precipitation, the ability to reproduce daily and hourly accumulations, and the daily cycle of summer rainfall. Additionally, two extreme precipitation events are deeply analyzed to investigate the performance of the reanalyses in specific conditions. The results indicate a quantitative added skill of SPHERA in representing precipitation, especially for moderate to severe and rapid accumulations in various terms: better adherence to the intensities of the observed states, higher detailing of the spatial fields, and more precise temporal matching.

These encouraging enhancements promoted the use of SPHERA for the investigation of hailstorms occurrence. Hail is, in fact, one of the most hazardous perils related to DMC. However, the scientific understanding of hail still needs improvement owing to intrinsic difficulties in observing and simulating its occurrence. Hence, the combination of multiple information is crucial to reduce the uncertainties associated with hailstorms. For this reason, a technique is proposed to combine a set of hail-favoring environmental predictors from SPHERA with observations of ESWD surface-based hail reports and satellite overshooting top detections, used as a proxy for severe convective thunderstorms. To obtain a proxy specifically for hail, a filter based on the environmental conditions surrounding the overshooting tops is developed to retain only the occurrences with the potential for hailstorm formation. The procedure is applied to the extended summer season (April-October) of 2016-2018 over the whole SPHERA spatial domain. The results show maximum hail likelihood over pre-Alpine regions and the northern Adriatic sea around 15 UTC in June-July, in agreement with recent European hail climatologies. Furthermore, the match with hail reports shows a better ability to detect occurred hailstorms compared to a previous study considered as the primary reference. Moreover, the method demonstrates enhanced performance in case of severe hail occurrences and can separate between ambient characteristics depending on hail severity.

The results presented in this thesis prove the added value in the description of DMC processes brought in by the new convection-permitting reanalysis SPHERA and ultimately indicate it as a valuable tool for advancing the representation of the associated weather dynamical features.

Contents

List of Symbols

Symbol	**Description**	**Units**
CAPE	Convective Available Potential Energy	J kg^{-1}
DLS	Deep Layer Shear	m s^{-1}
ΔT	Difference between IRBT and tropopause temperature	K
g	Gravitational acceleration	m s^{-2}
H_0	Freezing level height	m
IRBT	InfraRed Brightness Temperature	K
K	K index	°C
LZB	Level of Zero Buoyancy	m
p_{xx}	Percentile xxth of the distribution	various
SLI	Surface Lifted Index	°C
Td_{xxhPa}	Dew-point temperature at the isobaric level xx hPa	°C
T_{xxhPa}	Temperature at the isobaric level xx hPa	°C
$T_{v_{envir}}$	Virtual temperature of the environment	°C
$T_{v_{parc}}$	Virtual temperature of the parcel	°C
v_{500hPa}	Horizontal wind field at 500 hPa	m s^{-1}
v_{10m}	Horizontal wind field at 10 m	m s^{-1}

List of Abbreviations

Abbreviation	Description
AR6	Sixth Assessment Report
ARCIS	ARchivio Climatologico per l'Italia Centro Settentrionale
ARPAE-SIMC	Hydro-meteo-climate service of the Regional Agency for Prevention, Environment and Energy of Emilia Romagna region, Italy
CDF	Cumulative Density Function
CEST	Central European Summer Time
COSMO	COnsortium for Small-scale MOdelling
CP	Convection Permitting
DJF	December January February
DMC	Deep Moist Convection
DPCN	National Department of Civil Protection
ECMWF	European Centre for Medium-range Weather Forecasts
EPE	Extreme Precipitation Event
ERA5	ECMWF ReAnalysis 5th generation
ESWD	European Severe Weather Database
ESWD-S	Subset of ESWD reports having information on maximum hailstone Size
FAR	False Alarm Ratio
IPCC	Intergovernmental Panel on Climate Change
IQR	InterQuartile Range
IR	InfraRed
JJA	June July August
MAM	March April May
MCS	Mesoscale Convective System
MSG	Meteosat Second Generation
NWP	Numerical Weather Prediction
OT	Overshooting Top
POD	Probability Of Detection
QC	Quality Control level
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SON	September October November
SPHERA	Special Project: High rEsolution ReAnalysis over Italy
SRI	Surface Rainfall Intensity
SR	Success Ratio
TA	Temporal Accuracy
TS	Threat Score
UTC	Universal Time Coordinated

List of Figures

List of Tables

Chapter 1

Introduction and motivation

1.1 Extreme weather events, climate change, and deep moist convection

The IPCC (Intergovernmental Panel on Climate Change), in the Sixth Assessment Report (AR6), defines an extreme weather event as “an event that is rare at a particular place and time of year” and an extreme climate event as “a pattern of extreme weather that persists for some time, such as a season” (Seneviratne et al., 2021). Examples of extreme weather events are floods, severe storms, heatwaves, or droughts. Extreme weather events may produce severe impacts having socio-economical consequences on our everyday life, affecting our food or water supplies, damaging our infrastructures, and putting our and animal lives and whole ecosystems in danger (e.g., Paul et al., 2018). Hence, these meteorological events constitute some of the most threatening natural hazards which potentially may happen simultaneously and combine, producing multiple concomitant extremes of different nature and ultimately leading to even stronger impacts (e.g., as happened with Hurricane Irma in the Caribbean and the southeastern United States in September 2017 – Emrich et al., 2019). Additionally, it has been demonstrated by numerous research efforts in the last years, and documented in the AR6, that a changing climate leads to changes in the frequency, intensity, duration, spatial extent, and timing of weather and climate extremes and that the associated risks will continue to increase as the global mean temperature rises, possibly resulting in unprecedented extremes (e.g., Guerreiro et al., 2018; Wuebbles et al., 2017; Forzieri et al., 2016; Quevauviller and Gemmer, 2015). The reports “Explaining Extreme Events from a Climate Perspective” published annually by the American Meteorological Society since 2012 (Herring et al., 2021), found a substantial link between extreme events and climate change for 70% of the studies included in their analyses. The Germanwatch Global Climate Risk Index 2021 (Eckstein et al., 2021) estimated that between 2000 and 2019, more than 475,000 people lost their lives globally due to more than 11,000 extreme weather events, which produced losses for US$ 2.56 trillion. Within high-impact natural hazards, hydro-meteorological disasters (i.e., meteorological, hydrological, or climatological events) account for over 87% of the damages globally, including casualties, infrastructural and economic losses, and disruption to everyday life (Porcù et al., 2019).

Among hydro-meteorological hazards, severe convective storms, accompanied by phenomena such as tornadoes, large hail, wind gusts, or extreme precipitation, often causing flash floods, represent one of the main risks to society (Brooks et al., 2003). For example, in the 40 years from 1980 to 2019, the United States registered US$ 258 billion in damages from weather- and climate-related disasters (NCEI, 2020), ∼44% of which resulted from severe convective storms (Gensini, 2021). Moreover, the frequency and intensity of severe convective thunderstorms are expected to increase in response to anthropogenic-induced climate change (Allen, 2018; Trapp et al., 2019, 2007), even if recent investigations suggest that possible future shifts may not be straightforward and homogeneous over the planet (Taszarek et al., 2021; Raupach et al., 2021). Severe convective storms may form under favorable conditions for the instauration of deep moist convection (DMC). Generally, atmospheric convection is referred to as the process by which heat is transported by the vertical flow component associated with buoyancy. DMC occurs when air is lifted to saturation and achieves positive buoyancy with the potential to rise vertically to great heights, implying the convective overturning of most of the troposphere and causing water condensation (Markowski and Richardson, 2011). This process causes the release of large amounts of energy owing to phase changes of water, which often result in hazardous weather events (Doswell, 2001). Three main ingredients are required for the instauration of DMC and the associated development of severe thunderstorms: atmospheric instability, sufficient low-level moisture, and a lifting mechanism.

1.2 Extremes of hydrometeors

A hydrometeor is defined as “any product of condensation or deposition of atmospheric water vapor, whether formed in the free atmosphere or at the Earth’s surface” (American Meteorological Society, 2023). When referring to extremes in this context, two of the most prominent meteorological phenomena outstanding in terms of global entities of the produced impacts and risks are severe precipitation and hail.

1.2.1 Heavy rainfall

Extreme precipitation events (EPEs) cause large socio-economic losses worldwide (e.g., Lugo, 2018; Ackerman, 2017). They consist of fast accumulations of liquid precipitative water exceeding normal (i.e., climatological) conditions and reaching a hundred or even several hundreds of millimeters even in a few hours. The major hazard posed by EPEs is their potential to trigger flash floods or landslides, the former being considered the deadliest hazard associated with convection worldwide (Markowski and Richardson, 2011). Densely populated areas are the most vulnerable in this sense, especially where drainage systems may not be sufficient (Dale et al., 2015). Further, excessive amounts of rainfall can degrade the water quality of rivers, lakes, or seas, subsequently putting into danger aquatic ecosystems and, hence, human and animals health (Aguilera et al., 2019; Puczko and Jekatierynczuk-Rudczyk, 2020; Fukushima et al., 2021). Some recent examples of catastrophic EPEs worth mentioning include: the flood in central Europe in July 2021, causing hundreds of fatalities and estimated damages of EUR 32 billion (Mohr et al., 2022), the record-breaking heavy rainfalls in eastern China in June 2015, which produced widespread flooding affecting 5,380,000 people and causing the collapse of 46,600 houses (Wang and Gu, 2016), the torrential precipitation over southeastern Brazil in January 2020, leading to cascading effects of flooding and landslides, causing more than 90,000 people to lose their houses, and estimated to be >70% more likely to occur as a consequence of human-induced climate change (Dalagnol et al., 2022), or the multiple disastrous floods in Pakistan from June to October 2022 which inundated a third of the country, killed more than 1,700 people and affected 33 million, and caused over US$ 30 billion of damage (Nanditha et al., 2022; Bhutta et al., 2022).

As EPEs are the product of the interaction among several physical processes involving broad spatio-temporal scales, the complexity of the resulting impacts is strongly dependent on the local distribution of precipitation (e.g., Müller et al., 2018). Rainfalls can be classified into two categories based on the originating mechanisms of the associated clouds (Bližňák et al., 2021): stratiform precipitation, affecting widespread areas and lasting from several hours to several days (e.g., Minářová et al., 2017; Houze Jr, 2014), and convective precipitation, produced by convective thunderstorms and usually confined over smaller areas, lasting from minutes to several hours, and possibly resulting in sudden heavy accumulations potentially producing flash floods (e.g., Grazzini et al., 2020b,a). The latter kind of precipitation is generally mostly prone to producing EPEs (Semie and Bony, 2020; Gimeno et al., 2022). Furthermore, recent findings indicate that, as a consequence of anthropogenic global warming, the ratio between stratiform and convective precipitation has been changing in the last decades in favor of convective rainfalls (Chernokulsky et al., 2019; Ye, 2018; Han et al., 2016). As previously mentioned, this is linked to the growing evidence that extreme weather conditions under global warming are likely to strengthen in intensity and frequency of occurrence (Seneviratne et al., 2021). For EPEs, this translates into an expected shift towards higher accumulations, especially in those regions with high moisture availability, with a consequent increase in the frequency and severity of flood events (Tabari et al., 2020; Alfieri et al., 2015).

Multiple factors related to the atmospheric and geographical characteristics of a specific region play a central role in the development of EPEs, a necessary precondition of which is moisture availability in the lower troposphere and related mechanisms for its transport (Lavers and Villarini, 2015). Other elements are the thermal and moisture stratification of the atmosphere, the presence of convection and its organization, the interaction with topography and available mechanisms for the airflow ascent, or the proximity to water bodies (Grazzini et al., 2020a). Hence, due to the high complexity characterizing EPEs formation and the multiple spatio-temporal scales involved in their dynamical evolution, it is still challenging to skillfully forecast and observe these phenomena despite the significant improvements in the last decades. Concerning the observations, precipitation has been historically recorded through rain-gauge sensors deployed over a certain region to form a network. These have provided tremendous insights into regional hydrology and its long-term temporal characterization. However, this approach is mainly limited by the relatively small reference areas covered by the sensors and the partial information they provide, owing to their sparseness and inhomogeneous spatial density (Paz et al., 2020). This constitutes a serious limitation, especially in the case of EPEs characterized by peaked spatial localization. Remote sensing methods, whether through radar networks or satellite instruments, fill this gap thanks to their higher spatial resolution, homogeneity, and wider coverage. While radar coverage extends to several hundred kilometers, satellites can potentially encompass the entire Earth. However, also in this context, some limitations emerge, such as the sensitivity of the sensors to capture and distinguish all types of precipitation or the influence of complex orography that may block electromagnetic signals (Falconi and Marzano, 2019; Herold et al., 2017). For any kind of observation, numerous methods for post-processing the data to provide reliable estimates have been developed over the years (e.g., Lyra et al., 2018; Bhatti et al., 2016; Vulpiani et al., 2012), and the combination of different data sources represents a promising avenue of research (Ochoa-Rodriguez et al., 2019; Manz et al., 2016; Ma et al., 2021) that still needs further investigation.

Indeed, a deeper level of understanding of EPEs is envisaged in various directions, such as: i) better investigating the local-scale implications of global warming on atmospheric moisture availability and EPEs formation, ii) improving the techniques for the attribution of EPEs variability to human-induced warming to help decision-makers in better managing the risks and to reduce societal and ecosystemic vulnerability to EPEs, iii) reducing the uncertainty lying in the numerical simulations of, especially, severe convective rainfall events mainly owing to the use of parameterization schemes, which reveal a further need, i.e., iv) better understanding and representing the physical processes behind convective mechanisms for example through simulations with enhanced spatio-temporal resolutions, or via the assimilation of water vapor flux observations in model simulations (Gimeno et al., 2022).

1.2.2 Hail

Another hazardous weather phenomenon associated with precipitative hydrometeors is hail. Hailstones are more or less regular-shaped lumps of ice of diameter 5 mm or more produced by convective thunderstorms, which, in these cases, are also referred to as hailstorms. Hailstorms produce yearly billions of dollars of damage to crops, buildings, and properties worldwide, whose effects are proportional to hailstones size and frequency, resulting in severe economic and insurance losses. The extent of the impacts can range from the complete destruction of crops (Zhou et al., 2016) to damage to roofing, windows, or walls of buildings (Paterson and Sankaran, 1994) or to vehicle paneling and windows (Hohl et al., 2002). Additionally, although more rarely, very large hailstones (i.e., exceeding 5 cm diameters) may cause injury or death to animals or humans (e.g., Calianese et al., 2002). A striking case, in this sense, is the hailstorm that struck Moradabad, India, on the 30th of April 1888, which recorded the highest mortality rate globally of 246 people deceased as a result of hailfall (Cerveny et al., 2017). The damages that hail can produce depend on hailstone size: while for crops or trees, damages owing to small hail of 2-3 cm is most frequently reported, damages to properties such as buildings or vehicles are typically caused by larger hailstones of 4-6 cm or more (Púčik et al., 2019). Additionally, large accumulations of small hail (even <2 cm diameters) on the ground constitute a further threat (Kumjian et al., 2019). Although generally not as catastrophic as larger hail, deep small hail accumulations pose substantial risks to life and property, causing motor vehicle accidents, road closures, urban flooding and associated water rescues (Kalina et al., 2016; Schlatter and Doesken, 2010), or even reduced visibility owing to the formation of dense "hail fog" (Ward et al., 2018). With deep hail accumulations reaching 15-60 cm in 30 minutes, these kinds of hailstorms may produce significant economic losses (Grahame et al., 2009). Generally, individual hailstorm events can cause losses exceeding US$ 1 billion (Gunturi and Tippett, 2017), as reported in case-study analyses in Europe (Kunz et al., 2018), United States (Changnon and Burroughs, 2003), and Australia (Yeo et al., 1999).

Europe also boasts one of the costliest events in modern world history, i.e., the hailstorm that struck Munich, Germany, on the 12th of July 1984, damaging 70,000 houses and 200,000 automobiles, injuring 400 people, and producing damages for US$ 2 billion (Höller and Reinhardt, 1986). In the United States, hail damage is the costliest natural hazard in terms of insured losses to properties and agriculture and shows a tendency to increase (Changnon, 2009). In comparison, hailstorms in Europe (Punge and Kunz, 2016) generally tend to be less frequent and severe owing to the different orientation of large-scale mountain ranges (i.e., the Rocky mountains running north-south vs. the Alps running east-west) and the associated circulation patterns (Brooks et al., 2003).

Hail can form in severe DMC thunderstorms characterized by strong updrafts containing large amounts of supercooled liquid content, with high-reaching cloud tops and a sufficient lifetime for the accretion of the hailstones (Houze Jr, 2014; Knight and Knight, 2001). Hence, hail-prone storms are usually highly-organized convective systems such as multicells, mesoscale convective systems, or supercells (Dennis and Kumjian, 2017).

Anthropogenic global warming is expected to induce changes in the environments associated with hailstorm developments, mainly in the form of increased low-level moisture and hence convective instability (Trapp et al., 2007; Brooks, 2013; Rasmussen et al., 2020), higher freezing level altitudes (Xie et al., 2008; Mahoney et al., 2012; Dessens et al., 2015), and enhanced vertical wind shear (Trapp et al., 2007; Brooks, 2013; Brimelow et al., 2017). These conditions will generally increase the likelihood of hailstorm formation and the development of larger hailstones within them (Trapp et al., 2019; Brimelow et al., 2017; Dessens et al., 2015). However, due to the complexity of the formation mechanisms controlling hail and the crucial dependency on local geographical features, the response of hailstorms to warming is highly heterogeneous (Raupach et al., 2021). Furthermore, given the numerous limitations imposed by the general lack of direct hail observations, the incomplete understanding of microphysical processes controlling hail formation and growth, and the insufficient level of detail of the models numerically simulating the process, a high degree of uncertainty still permeates hailstorm response to climate change (Seneviratne et al., 2021; Allen et al., 2020).

The low probability of occurrence of hailstorms at a certain location makes observing hail, both with direct or indirect methods, still a major challenge worldwide (Allen et al., 2020; Prein and Holland, 2018). Indeed, a widespread, standardized, and operational surface hail observing system is still missing, and the existing methodologies are highly variable on a country-to-country basis. Atmospheric weather stations do not usually include dedicated sensors for hail detection (except for China - Li et al., 2018b), and the few manned stations providing them, which often are operative only during daytime, are too scarce to capture highly localized hailstorm events reliably (Punge and Kunz, 2016). Automatic hail recorders also exist (Löffler-Mang et al., 2011), a network of which has been recently deployed over Switzerland (Kopp et al., 2022). However, the elevated costs of these instruments prevent their systematic deployment over extended areas. Specific in-situ hailpad networks have been deployed in several parts of the world (e.g., Changnon Jr, 1970; Xie et al., 2008; Palencia et al., 2010; Dessens et al., 2015; Sánchez et al., 2009; Giaiotti et al., 2003; Manzato, 2012). Hailpad measurements can provide precious homogeneous long-term records for hail, including important information, such as the hailstone size spectrum. However, they are characterized by serious limitations such as the restricted area of extension implying only regional coverage possible, the considerable costs for their maintenance owing to their non-automatization, or the often low spatial density which could likely cause, for example, to miss the largest hailstones (Changnon, 1977; Smith and Waldvogel, 1989). An additional source of hail data comes from event-based reports collected by multiple sources such as storm spotters, voluntary observers, or media. This practice has been adopted for several years in different regions of the world, such as the United States (the Storm Prediction Center Severe Weather Database - Allen et al., 2015), Australia (the Bureau of Meteorology Severe Thunderstorm Archive - Allen and Allen, 2016), or Europe (the European Severe Weather Database, presented in detail in Section 4.1.1 - Dotzek et al., 2009). However, these data are also affected by a number of limitations, such as the non-uniform spatial coverage of observations being biased towards densely populated areas, the possible errors in reporting the largest hailstones, which could easily be missed owing to the limited spatial coverage of severe hailstorms, or the potential underreporting in cases of non-damaging hailstorms which are generally less interesting for the public to be reported (Allen et al., 2020). In any case, these data represent one of the most valuable sources of direct observations for hail and other convection-related events. Further, their reliability has the potential to increase in the future owing to the enhanced cooperation among various institutes allowing the establishment of standardized reporting practices and increasing the spatial homogeneity of the reports (Púčik et al., 2019), as well the growing interest of the general public for severe weather reporting also thanks to the introduction of specific mobile apps for the purpose (e.g., Holzer et al., 2017). Moreover, these data are essential to validate hail proxies from remote sensing observations. Indeed, a further important source of hail data comes from radar or satellite observations from which it is possible to indirectly estimate the occurrence of hail with more spatial homogeneity and higher temporal frequency than reports or hailpad measurements (e.g., Murillo and Homeyer, 2019; Gobbo et al., 2021; Mecikalski et al., 2021). Radar reflectivity information constitutes the most powerful instrument for detecting hail and estimating its size and probability of occurrence (Fluck et al., 2021; Nisi et al., 2020), even if limitations exist in the correct inference of hailstone dimensions (Ortega, 2018). Satellite observations cover larger regions of the world than radars (Cecil and Blankenship, 2012). The information relying on passive instruments measuring the outgoing electromagnetic radiation on the infrared or visible spectrum has been used to develop hailstorm detection methods (e.g., Melcón et al., 2016). In this context, a prominent approach for detecting severe convective updrafts is by identifying overshooting cloud tops from geostationary satellite imagery. The first version of the automatic identification algorithm was developed by Bedka et al. (2010) and subsequently refined and optimized by Bedka and Khlopenkov (2016) and Khlopenkov et al. (2021). Hence, constituting a promising proxy for severe convective storms identification, overshooting top detections have been employed in several studies specifically for investigating hail (Setvák et al., 2013; Mikuš and Mahović, 2013; Punge et al., 2017; Bedka et al., 2018; Punge et al., 2021). A detailed description of this technique is reported in Section 4.1.2. Another method for satellite-based hail detection, originally developed by Cecil (2009), exploits passive microwave imagery through which it is possible to detect areas of low apparent brightness temperatures as a consequence of radiative scattering owing to ice hydrometeors (e.g., Laviola et al., 2020). In any case, proxy data, based on empirical relationships generally hardly valid in multiple different contexts, are usually insufficient to characterize a challenging phenomenon to observe such as hail (Allen et al., 2020). Hence, direct and indirect hail observations are often combined and coupled with numerical simulations for hail estimation. For example, in the insurance sector dealing with hail-related damages, there is an impelling urge to enhance hail risk understanding and estimation to assess the possible average loss by region and time of the year (Martius et al., 2018). For this reason, some firms started to combine model data with supplemental information of various nature as proxies for hail occurrence (e.g., lightning detections, satellite- and radar-based detection algorithms, and event-based reports) to overcome observational limitations. Indeed, assessments performed by combining multiple data sources have the potential to reduce the biases of single-sourced records and overall account for knowledge gaps.

For a comprehensive review of the actual state of hail understanding, including the microphysical processes for hailstones formation and growth in a thunderstorm, the available observations worldwide, the numerical methods employed for hail estimation, the associated impacts, and possible future directions for their better understanding, the reader is referred to Allen et al. (2020).

1.3 Numerical modeling of extreme rainfall and hail

DMC processes, including EPEs and hailstorms, are characterized by extremely chaotic and non-linear behaviors, growing from the micro- and reaching even the meso- or synoptic-scales. Indeed, the dynamics controlling mesoscale and local-scale convective atmospheric processes are notoriously characterized by intermittency, characteristic of the non-linear evolution of chaotic systems (Doswell, 2001). Hence, the sensitive dependence on initial conditions characterizing chaotic dynamics (Lorenz and Haman, 1996) is also central to the development of DMC. This implies that, for example, the difference between a major outbreak of severe convective thunderstorms and a complete lack of convection might be imputable to a minimum difference in the initial state of convection inhibition. Furthermore, the microphysical processes occurring in convective clouds have a much higher level of complexity than stratiform counterparts owing to the strong vertical motions supporting mixed-phase water processes and the formation of a broad spectrum of hydrometeor types (Pruppacher et al., 1998). The non-linearity in the evolution of such atmospheric processes is the historical reason behind the substantial difficulties in understanding and forecasting DMC.

Since the earliest research days investigating atmospheric convection (e.g., Stommel, 1947), a long way has come in the representation of DMC through numerical weather prediction (NWP) models, with subsequent dynamical and physical refinements with increased complexity (Dudhia, 2014). Indeed, continuous improvements have been made over the years owing to the advancement in the physical understanding of the processes, combined with the advances in computer resources. This made evident that a crucial necessity for a reliable representation of specifically, but not exclusively, DMC is the need for a fine spatio-temporal resolution in the models (Wilhelmson and Wicker, 2001; Bryan et al., 2003; Wu and Arakawa, 2014; Clark et al., 2016). To the present moment, the investigation of extreme weather events linked to DMC, such as heavy rainfalls or hailstorms, with numerical models with high spatio-temporal resolutions is a primary field of research, as demonstrated by numerous recent studies (e.g., Tu et al., 2022; Chen et al., 2021; Lupo et al., 2020; Trapp et al., 2019; Malečić et al., 2022a,b; Tiesi et al., 2022). However, the high complexity characterizing DMC and severe thunderstorm formation and dynamical evolution substantially limit their predictability. The direct simulations of the storms through very high-resolution numerical models constitute a possible approach (e.g., Mahoney et al., 2012; Brimelow et al., 2017), but the associated high computational costs still prohibit anything beyond regional evaluations or case-by-case analyses. To address this issue, a common practice developed over the years is to investigate the simulated environmental proxies favorable for the development of DMC and associated hazards in a climatological sense (Johns and Doswell III, 1992; Brooks et al., 2003). This approach has demonstrated usefulness in identifying the atmospheric processes favoring the occurrence of phenomena associated with DMC events, such as hail, wind gusts, or tornadoes. This led the forecasters to improve their ability to anticipate severe thunderstorms (Thompson et al., 2003; Hitchens and Brooks, 2014; Tippett et al., 2014) and to develop extended climatologies on the global (Riemann-Campe et al., 2009; Chen et al., 2020; Prein and Holland, 2018) or regional scale (Gascón et al., 2015; Púčik et al., 2017; Li et al., 2018a; Tang et al., 2019; Taszarek et al., 2021). Most of these studies rely on atmospheric descriptions retained by reanalysis datasets instead of mere numerical models. As will be discussed in detail in the next Section, reanalyses provide a more comprehensive description of the past atmospheric states. Indeed, the investigation of past meteorological conditions leading to DMC and the associated events is crucial to enhance their understanding and increase their predictability in the present and future evolutions of the atmosphere.

1.3.1 The role of atmospheric reanalyses

Atmospheric reanalyses constitute the most common way to investigate past meteorological states and respond to the necessity of homogeneous spatio-temporal meteorological data by combining numerical weather predictions and observations. In these datasets, data assimilation, i.e., the blending process between atmospheric observations and simulations, is used to ingest real-state information and physically constrain model data. Atmospheric reanalyses extend from several years to a few decades. ERA5 (the fifth generation reanalysis, Hersbach et al., 2020) produced at ECMWF (European Centre for Medium-range Weather Forecasts) is widely considered state-of-the-art among reanalyses covering the whole globe. Anyhow, global datasets, including ERA5, are still characterized by coarse horizontal grid spacings spanning from 125 km to 31 km and with temporal frequencies varying from 6 to 1 hour. As previously stated, finer spatial grids are essential to describe small-scale and rapidly-evolving features characterizing convective processes with sufficient detail. In fact, physical parameterization schemes are necessary to account for sub-grid convective processes in large-scale models, but these are a major source of errors and inaccuracies in the simulations (Prein et al., 2015; Jones and Randall, 2011; Arakawa et al., 2011). Hence, even if ERA5 horizontal grid spacing (31 km) constitutes a significant leap forward compared to its predecessors, the need for DMC parameterization at this spatial resolution discourages its direct use for regional- or local-scale applications. This holds especially for investigating precipitative phenomena as revealed by recent research (Bandhauer et al., 2022; Jiang et al., 2021; Singh et al., 2021). This difficulty in reliably parameterizing DMC emerges from the interplay of scales involved in the processes responsible for triggering and developing convection ranging from micro- to synoptic scales.

The necessity of a superior level of precision for describing the phenomena associated with DMC led to the development of higher resolutions limited-area reanalyses. These are usually obtained by downscaling a global dataset to produce a finer displacement of the data over a smaller spatial domain (on the continental or regional scale). The higher level of detail gained in this way has been widely demonstrated to produce more realistic simulations of convective thunderstorms (Liu et al., 2017; Trapp et al., 2019; Prein et al., 2017; Hoogewind et al., 2017; Trapp and Hoogewind, 2016). In dynamical downscaling, the boundary conditions of the global driving model force a limited-area system to produce atmospheric fields at the desired grid spacing (e.g., Castro et al., 2005). However, the downscaling process could generate relevant small-scale variabilities (constrained to the larger-scale state but not to the observations), which can be detrimental to the finer-scale estimates (Desamsetti et al., 2019; Simon et al., 2013; Giorgi, 1990). The assimilation of regional observations is used to minimize coarse-scale driven variabilities and improve the quality of high-resolution estimates. Regional reanalyses are produced by coupling a downscaling strategy with a data assimilation method, which has proven to be superior to a mere downscaling (i.e., hindcast) of a global reanalysis (e.g., Bollmeyer et al., 2015; Jermey and Renshaw, 2016). Several regional reanalyses have been produced in recent years on the continental or national scales. Some examples are: CNRR (China Regional Reanalysis project, Zhang et al., 2017) having 18-km grid spacing, ASRv2 (Arctic System Reanalysis version 2, Bromwich et al., 2018) with 15-km grid spacing, BARRA (Bureau of Meteorology Atmospheric high-resolution Regional Reanalyses for Australia, New Zealand, and Southeast Asia, Su et al., 2019), or CARRA, the Arctic Regional Reanalysis produced by the Copernicus Climate Change Service (C3S) at a very fine grid spacing of 2.5 km (Køltzow et al., 2022). In Europe, the interest has been particularly high, as demonstrated by the numerous datasets recently developed: EURO4M (European Reanalysis and Observations for Monitoring, Klein Tank, 2010) and its continuation UERRA (Uncertainties in Ensembles of Regional Reanalyses, Unden et al., 2016), COSMO-REA6 (the Consortium for Small-scale MOdeling regional reanalysis, Bollmeyer et al., 2015) covering part of the European CORDEX domain (Coordinated Regional Downscaling EXperiment, `www.cordex.org`) with a grid spacing of 6 km, the new ensemble-based CERRA (the Copernicus European Regional ReAnalysis, El-Said et al., 2022) covering Europe and North Africa at 11-km horizontal grid spacing (while 5.5 km is used for its deterministic counterpart), or MERIDA (MEteorological Reanalysis Italian

DAtaset, Bonanno et al., 2019) covering Italy with a 7 km grid. The increased detail of these datasets provides added value, especially in the representation of precipitation (Bollmeyer et al., 2015; Jermey and Renshaw, 2016) and in the case of complex orography (Isotta et al., 2015). However, these grid spacings are still too coarse to permit the explicit representation of DMC, which is paramount for substantially improving the simulations.

1.3.2 Benefits of convection-permitting simulations

The explicit representation of DMC can be achieved through the so-called convection-permitting (CP) scales, implying horizontal grid spacings of few km, commonly less than 5. Indeed, at these spatial scales, numerical parameterization schemes for convection may be switched off as DMC starts to be resolved explicitly (Weisman et al., 1997). Furthermore, the enhanced spatial resolution in CP settings, besides allowing for a more realistic representation of convective processes (Leutwyler et al., 2017; Prein et al., 2015), also permits enhanced descriptions of the topographical features of the surface boundary of the model, implying improved representation of complex mountainous terrains or heterogeneous surfaces, like coastal or urban regions (Trusilova et al., 2013; Prein et al., 2013b). Anyhow, the fine grid spacings of CP simulations require computational costs that are still very demanding to cover extended/global domains (e.g., Schär et al., 2020). For this reason, CP reanalyses are limited to small spatial domains on the order of the national- and below the continental scales. Recently produced CP reanalyses over Europe include: COSMO-REA2 (nested in COSMO-REA6, Wahl et al., 2017) covering Central Europe at 2-km grid spacing, MERA (Met Éirann ReAnalysis, Gleeson et al., 2017) over Ireland and the United Kingdom with a 2.5-km grid, and the still-in-development IBERA (high-resolution regional reanalysis for Iberian Peninsula and Balearic Islands, Calvo Sánchez et al., 2021) developed by AEMET (Agencia Estatal de METeorología) at 2.5-km grid spacing.

When coming to precipitative phenomena, an improved description through CP models compared to larger-scale counterparts is obtained in terms of numerous aspects: enhanced representation of the summer diurnal cycle of precipitation (Fosser et al., 2015; Brisson et al., 2016), better accordance with the observed intensities of the most severe precipitation events (Prein et al., 2013a; Fosser et al., 2015), lower biases of average and extreme rainfalls (Pal et al., 2019), and smaller frequency biases of weak precipitation events (Berg et al., 2013). Similarly, CP regional reanalyses/hindcasts better describe the spatial variability on the local scale of precipitation and the frequency of heavy rainfalls (Wahl et al., 2017; Capecchi et al., 2022). Furthermore, CP simulations successfully represent organized convective structures and the formation of isolated convective cells and self-regenerating thunderstorms (Clark et al., 2016), as well as orographically-driven convection thanks also to the enhanced detailing of the topography and surface heterogeneities (Kirshbaum et al., 2018; Weusthoff et al., 2010). Considering specific severe weather phenomena linked to DMC, CP simulations have been shown to reliably reproduce proxies for hazardous convective events such as tornadoes or large hail on a climatic term (Gensini and Mote, 2014), to skillfully detect tornadoes path lengths (Clark et al., 2013), or to provide reasonable predictions of surface hail size (Adams-Selin and Ziegler, 2016; Labriola et al., 2019; Gagne II et al., 2019; Manzato et al., 2020). Additionally, the simulation of hailstorms in CP-scale frameworks have demonstrated to well resemble hail observations through radar reflectivity estimates or in-situ measurements (Malečić et al., 2022b) and also to provide future climatological changes in hailstorm seasonality and regionality under global warming conditions (Trapp et al., 2019; Mahoney et al., 2012).

1.3.3 The new CP reanalysis SPHERA

The need for a better description of the atmospheric processes related to DMC prompted the development of a new CP regional reanalysis: SPHERA (High rEsolution ReAnalysis over Italy). SPHERA has been produced by ARPAE-SIMC (the hydro-meteo-climate service of the Regional Agency for Prevention, Environment and Energy of Emilia Romagna region, Italy) with computational resources provided by ECMWF (the European Center for Medium-Range Weather Forecasts). Obtained through a dynamical downscaling of the global reanalysis ERA5, SPHERA is driven by the non-hydrostatic limited-area model COSMO, run deterministically with a 0.02° x 0.02° grid spacing (i.e., approximately 2.2 km x 2.2 km) over 65 vertical levels, and assimilates regional in-situ observations with a continuous nudging scheme during the forward model integration. The covered spatial domain is centered over south-central Europe and coincides with that of COSMO-2I, including Italy, the surrounding seas, and part of the neighboring states (Figure 2.1). SPHERA produces outputs at hourly temporal frequency over 26 years (1995-2020), but future extensions of the dataset are possible. The details of SPHERA production and the necessary preliminary experimentations to define its optimal setup have been investigated by Cerenzia et al. (2022). The main aims that led to the development of this new limited-area reanalysis are multiple: to have a high-resolution, space- and time-consistent description of the climatic characteristics during the past decades over the region, to provide an accurate and long-term estimate for the COSMO model to be used as reference (whether for climate change scenarios investigations or for shorter-term applications, calibrations, and comparisons with the operational setting of the model COSMO employed at ARPAE-SIMC for issuing weather forecasts and alerts), and, ultimately, to supply a novel tool to investigate extreme weather phenomena, whether on a long-term statistical term or for detailed investigations of specific events, whose understanding and reproduction could potentially benefit by the improvements introduced with the new weather archive data. Chapter 2 reports a comprehensive presentation of SPHERA.

It is worth noting that other datasets sharing similar characteristics have been recently produced. Indeed, when coming to Italy, the considerable interest in developing highly-resolved re-forecast datasets is demonstrated by the recent production of two CP regional hindcasts. These are also obtained by downscaling ERA5 using the COSMO model at 2.2-km grid spacing (Raffa et al., 2021; Reder et al., 2022) or the model MOLOCH at 2.5 km (Capecchi et al., 2022). The main difference with respect to SPHERA is that in neither case is the additional assimilation of regional observations included in the development of the datasets. Additionally, the production and comparison of similar but independent CP reforecast datasets is useful and encouraged owing to the high model-dependency of the simulations at these spatial scales (Prein et al., 2015), which potentially could further improve when considering multiple model settings together for the representation of high-impact weather occurrences (Coppola et al., 2020).

1.4 Goals and outline of the thesis

This thesis aims to assess the potential enhancements in the representation of DMC processes over Italy and neighboring countries and seas brought in by the newly-developed high-resolution regional reanalysis SPHERA, representing a state-of-the-art approach in NWP modeling. To accomplish this task, the analyses for two individual but closely related aspects are presented:

1. The statistical validation of SPHERA and a quantitative comparison in the performance simulation with its driver ERA5 for the representation of precipitation over Italy for a 15-year period, with particular focus towards heavy-rainfall occurrences, with the intent to respond to the scientific question: ***to what extent do the improvements of CP reanalysis simulations outperform coarser counterparts in representing precipitation and, specifically, severe precipitation events?***

2. The investigation of hailstorm occurrences over south-central Europe through the development of a method for the combination of high-resolution environmental predictors, simulated by SPHERA, with observative proxies for hail, such as satellite thunderstorm detections and surface-based hail reports, over

> three years to answer the question: ***can the combination of CP-simulated ambient predictors with observational proxies for hail enhance the ability to detect hailstorm events and improve the understanding of the necessary convective environmental conditions for their development?***

As extensively reported in the previous paragraphs, enhancing the representation of heavy precipitation is of primary importance in the context of NWP modeling and climate monitoring. Recent research employing similar CP reanalysis to SPHERA over Ireland (MERA - Gleeson et al., 2017) and Germany (COSMO-REA2 - Wahl et al., 2017) reported multiple benefits when comparing their simulated precipitation fields with coarser reanalysis datasets (i.e., ERA-Interim for the former and ERA-Interim, HIRLAM, and COSMO-REA6 for the latter). The found enhancements regard particularly the local scale of precipitation in terms of a more realistic representation of its spatial variability and temporal frequency, especially of summer convective rainfall, improved adherence to the observations in case of daily- and even more for sub-daily precipitation frequencies, and a higher ability to capture rainfall extremes in terms of intensity and spatial location. For this reason, similar improvements are expected with SPHERA from the comparison with ERA5. To quantify the potential added value of SPHERA, a long-term statistical assessment is performed and reported in Chapter 3 of this thesis, which also constitutes the first extensive validation of the novel dataset. After presenting the independent observational data used in the comparison of the two reanalyses, and the associated statistical methods employed to quantify the performance in representing precipitation, a preliminary analysis is dedicated to the identification of the optimal spatial scale to be used for the performance assessment. Then, an extensive comparison is performed over 15 years considering the statistical proximity to rain-gauge records of, separately, average and maximum precipitation estimates. Various aspects are investigated, such as the adherence to the observed spatial distributions, the representation of daily and hourly accumulations, and the daily cycle of summer precipitation. Finally, to deeply understand the abilities of the reanalyses in simulating extreme precipitation occurrences, detailed analyses for two EPEs case studies are presented.

The intrinsic difficulties encountered in the observation and simulation of hail call for new attempts at improving its understanding. Indeed, as reminded above, both observations and numerical simulations lack sufficient reliability to be standalone methods for accurate retrievals of hail information. Hence, a combination of various data sources is often applied to fill the gaps of their individual uncertainties. In recent years, much attention was devoted to the investigation of hailstorms through the identification of hail-favoring environmental conditions. These consist of proximity-soundings atmospheric parameters statistically associated with hailstorm formation, whether obtained through rawinsonde or model/reanalysis data. Reanalyses play a major role in this context, given the spatial homogeneity and long-term records they provide. However, reanalysis estimates need to be coupled with the available hail observations to retrieve the most skillful environmental hail proxies. Numerous efforts recently were devoted to the development of statistical models for estimating hail hazard by coupling large-scale reanalysis/regional climate models with lightning data (Rädler et al., 2018), surface-based reports (Prein and Holland, 2018; Torralba et al., 2022), or combinations of lightning, insurance loss, severe weather reports, and 3-D radar data (Mohr et al., 2015b; Taszarek et al., 2020). Additionally, another perspective exists for the usage of numerically simulated ambient ingredients associated with favorable hailstorm conditions. Such an approach involves their coupling in "calibration mode" with observational hail proxies, such as remote-sensing thunderstorm detections, to retain only those occurrences that are likely to have produced hail. In this context, satellite detections such as microwave scattering information (Bruick et al., 2019) or radar-based reflectivity estimates (Smith et al., 2012) have been considered. A particularly promising satellite-based proxy for hail is given by overshooting cloud top detections, which, when coupled with reanalysis-based predictors, have demonstrated to be a valuable tool for the spatio-temporal characterization of hail hazard (Punge et al., 2017; Bedka et al., 2018; Punge et al., 2021). However, these investigations rely on environmental conditions described with coarse global reanalysis such as ERA-Interim or ERA5, and, to the author's knowledge, no similar studies have been performed so far in this field by employing CP simulations. Moreover, due to the scarcity of studies involving environmental descriptions performed at increased resolutions, which potentially allow for a better representation of hail hazard, this direction has been indicated as a promising avenue of research (Allen et al., 2020; Raupach et al., 2021). This prompted the development of a methodology to couple a selected set of SPHERA environmental predictors for hail with satellite overshooting top detections from Meteosat Second Generation (MSG) SEVIRI instrument and surface-based hail reports. The analysis is applied for three years over south-central Europe to assess its value in representing hailstorm likelihood of occurrence and the related ambient signatures and is presented in Chapter 4 of this thesis. The main ideas for this investigation were developed during the Ph.D. mandatory period abroad, spent at the Atmospheric Risks group of the Institute of Weather and Climate research, Department of Tropospheric research (IMK-TRO), at the Karlsruhe Institute of Technology (KIT) in Karlsruhe, Germany. This opportunity was possible thanks to the collaboration and supervision of Prof. Dr. Michael Kunz, Dr. Heinz Jürgen Punge, and Dr. Kristopher M. Bedka from NASA Langley Research Center (who gently provided satellite data).

The remainder of the thesis is organized as follows: Chapter 2 presents the new reanalysis SPHERA including the technical details of the modeling setup, the data assimilation procedure adopted, and a brief summary of the preliminary investigations performed to define the optimal setup for its production. The comparative analysis with ERA5 for precipitation estimates is reported in Chapter 3, which includes a description of the data and statistical methods employed, the preliminary sensitivity analysis to define the optimal spatial configuration for the comparison, the long-term statistical analysis, the investigation of two heavy-rainfall case-studies, and a summary and discussion of the results through critical comparisons with similar works. Chapter 4 presents the analysis for hailstorm identification and associated environmental description, starting from presenting the satellite and in-situ observations included in the analysis, the description of the method developed to filter satellite detections based on the environmental signatures described with SPHERA coupled with hail reports, the resulting hailstorm spatio-temporal characterization over south-central Europe for three years, the investigation devoted to the separation of ambient conditions for different types of hailstorms, and a summary and discussion of the results including comparisons with similar research as well as possible further developments to enhance the method. Finally, the overall conclusion of the thesis and possible future developments are drawn in Chapter 5.

Chapter 2

SPHERA: High rEsolution ReAnalysis over Italy

In this Chapter the technical configuration of SPHERA reanalysis is presented with details of the driving numerical model, the modality by which the downscaling from ERA5 is operated, and the implemented data assimilation technique.

SPHERA (Special Project: High rEsolution ReAnalysis over Italy) is the new high-resolution regional reanalysis dataset produced at ARPAE-SIMC. SPHERA covers Italy, part of the neighboring countries, and the surrounding seas (Figure 2.1) at the CP horizontal grid spacing of 0.02°x 0.02°(approximately 2.2 x 2.2 km). Table 2.1 summarizes SPHERA main features, including the list of parameterizations employed to account for sub-grid physical processes. SPHERA spans 26 years (from 1995 to 2020) at hourly temporal frequency, and future extension of the dataset will be possible. The non-hydrostatic limited area model COSMO (`https://www.cosmo-model.org/` - Schättler et al., 2018; Baldauf et al., 2011), developed by the European Consortium for Small-Scale Modeling, drives SPHERA simulations. COSMO is a Numerical Weather Prediction (NWP) model used operationally in several European and extra-European countries belonging to the eponymous consortium (i.e., Germany - Baldauf et al. (2011); Greece - Avgoustoglou et al. (2018); Italy - Marsigli et al. (2005); Israel - Hochman et al. (2018); Poland - Starosta and Wyszogrodzki (2016); Romania - Dumitrache et al. (2011); Russia - Rivin et al. (2015); Switzerland - Klasa et al. (2018)). COSMO employs the primitive non-hydrostatic thermo-hydrodynamical equations for a compressible flow in a moist atmosphere. The three-dimensional equations are defined in advective form and numerically solved in SPHERA through a third-order Runge-Kutta split-explicit scheme (Wicker and Skamarock, 2002). The grid is defined in a rotated geographical coordinate system as an Arakawa C-grid with Lorenz vertical staggering. The vertical coordinate, extending up to 22,000 m a.s.l. over 65 levels, of which 19 in the lowest 1000 m, is defined through the terrain-following transformation of Gal-Chen

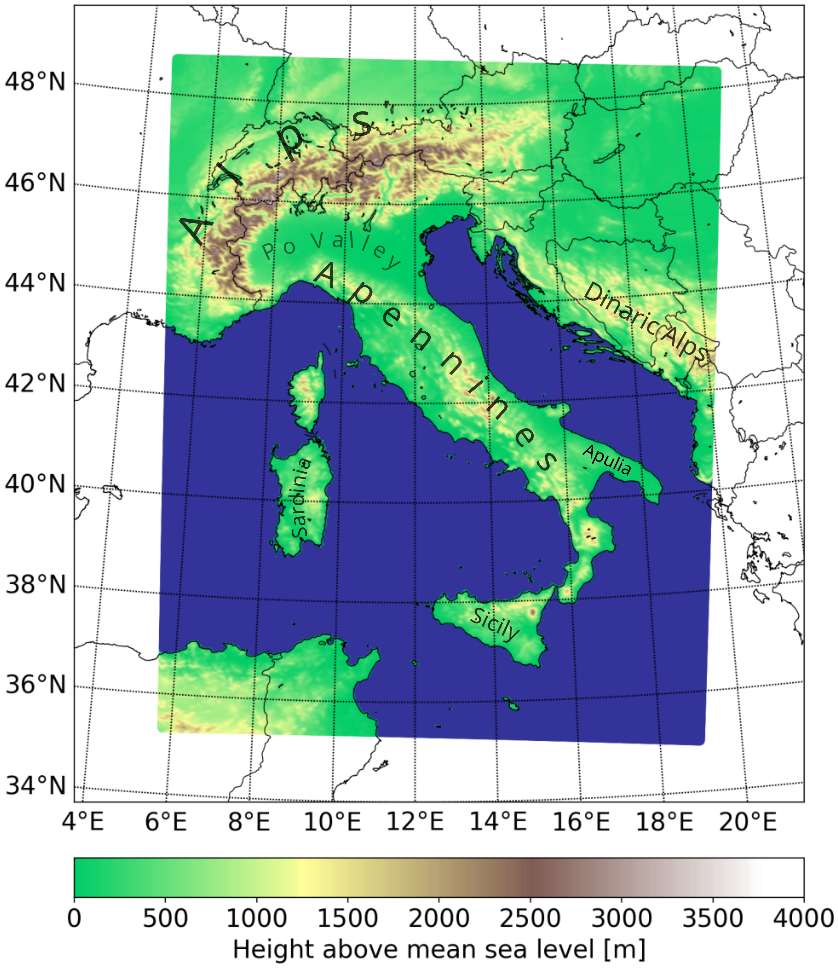

Figure 2.1: The spatial domain and model orography of SPHERA.

and Somerville (1975). In the soil, seven vertical levels are considered at depths of 0.005, 0.02, 0.06, 0.18, 0.54, 1.62, 4.86, and 14.58 m below the surface to account for near-surface, root-level, and deep-soil layer processes interactions with the atmosphere. The spatial domain is obtained by rotating the South Pole to W 10°, S 43°, and the horizontal grid mesh consists of 576 and 701 grid points in the west-east and south-north direction, respectively, with 0.02° spacing. The physical package included in the model is formed by: a radiation scheme (δ-two-stream scheme - Ritter and Geleyn, 1992), a turbulence scheme (prognostic turbulent kinetic energy closure at 2.5 level - Raschendorfer, 2001), a multi-layer soil scheme (based on the direct numerical solution of the heat conduction equation - Jacobsen et al., 1982), a multi-layer transfer scheme (Doms et al., 2018), a grid-scale cloud and precipitation scheme with prognostic cloud water, cloud ice and graupel, and a statistical scheme for sub-grid clouds (Sommeria and Deardorff, 1977), a two-layer bulk lake scheme (Mironov, 2008), and a shallow-convection scheme (reduced from Tiedtke, 1989). An extensive description of the parameterization schemes implemented in COSMO can be found in Schättler et al. (2018). A critical process not parameterized in the configuration used in SPHERA is deep moist convection, thanks to the fine grid spacing allowing its explicit representation (or at least for a large part of the associated convective motions). The reader is referred to Cerenzia et al. (2022) for a detailed description of the modeling framework adopted. SPHERA is initialized through dynamically downscaling the fifth-generation global reanalysis ERA5 produced at ECMWF (Hersbach et al., 2020), and continuously nudged towards the observed state during the forward model integration by assimilating regional observations. ERA5 is based on the Integrated Forecasting System (IFS) Cy41r2, producing deterministic hourly three-dimensional outputs at the spectral resolution TL639 (i.e., approximately 31 km grid spacing at mid-latitudes) over 137 vertical levels. The initialization process includes a pre-processing step by which ERA5 is interpolated to the COSMO grid with the procedure described in Schättler and Blahak (2017).

Preliminary analyses have been carried out to define the optimal setup for the production of SPHERA (Cerenzia et al., 2022) pertaining to the boundary conditions inherited from ERA5. As regards the nesting modality to downscale the lateral boundary condition to SPHERA, which is updated every hour, a 1-step nesting has been chosen. Indeed, this setup has proven more skillful than nesting passing through an intermediate grid. Additionally, the identification of the optimal bottom boundary conditions defining the deep-soil temperature has been investigated to avoid spurious drifts during the forward model integration caused by excessively warm or cold soils. The three-yearly running mean of ERA5 bottom soil temperature with a temporal delay showed the best adherence to the observed counterpart and hence was included in SPHERA.

For computational reasons, the production of SPHERA has been performed through six 4-year streams, each preceded by six months of rerun to account for the soil spin-up. The streams are formed by COSMO 24-hour long runs, each providing the initial condition for the subsequent one. The runs are then all stitched together to create a continuous hourly series. The boundary conditions imposed include a non-penetrative and free-slip condition at the model top, and free-percolation at the soil bottom for depths below 2.43 m. The data assimilation technique implemented in COSMO is a continuous nudging of in-situ observations, similar to COSMO-REA6 (Bollmeyer et al., 2015). The continuous nudging is based on the Newtonian relaxation principle and aims to dynamically adjust the model towards the prescribed observations within a predetermined time window (Stauffer and Seaman, 1990). This is done by inserting an additional term (so-called nudging term) in the prognostic model equations, which is proportional to the spatio-temporal misfit between the observations and the model, and that continuously adapts the simulation towards the observed values during the forward integration of the model. The nudging term

Initial conditions	ERA5
Boundary conditions	ERA5, updated every hour
Nesting modality	1-way nested in ERA5
Sea surface temperature	Interpolated from ERA5 every day
Deep soil temperature	Parameterized from ERA5 soil temperature after Cerenzia et al. (2022)
Assimilated observations	SYNOP and SHIP (not temperature at 2m and precipitation for both), TEMP, PILOT, and AIREP
Code version	INT2LM 2.04 (pre-processing), COSMO 5.05 in double precision
Spatial domain	Approximately 35°N, 5°E; 49°N, 20°E (estimated from the rotated domain, Figure 2.1)
Spatial resolution	0.02° (∼2.2km) horizontal (576x701 grid cells), 65 vertical levels terrain-following (0-22 km), 7 soil levels (0-14.58 m)
Temporal frequency	1 h
Temporal coverage	1995-2020
Physical schemes:	
Radiation scheme	δ-two-stream scheme after Ritter and Geleyn (1992)
Turbulence scheme	Prognostic turbulent kinetic energy closure at level 2.5 after Raschendorfer (2001)
Land-surface scheme	Multi-layer soil after Jacobsen et al. (1982)
Transfer scheme	Surface layer scheme coupled with the turbulence scheme (Doms et al., 2018)
Convection scheme	Only shallow convection with reduced Tiedtke (1989)
Microphysics scheme	Grid-scale cloud and precipitation scheme (3-categories ice scheme) and statistical scheme for sub-grid clouds after Sommeria and Deardorff (1977)
Subgrid-scale orography scheme	Lott and Miller (1997)
Lake scheme	Two-layer bulk model after Mironov (2008)
External parameters:	
Orography	Global Land One-km Base Elevation Project (Hastings et al., 1999)
Land cover	Global Landcover 2000 Database (Mayaux et al., 2006)
Soil type	Digital soil map of the World (FAO/UNESCO)

Table 2.1: Main technical characteristics of SPHERA.

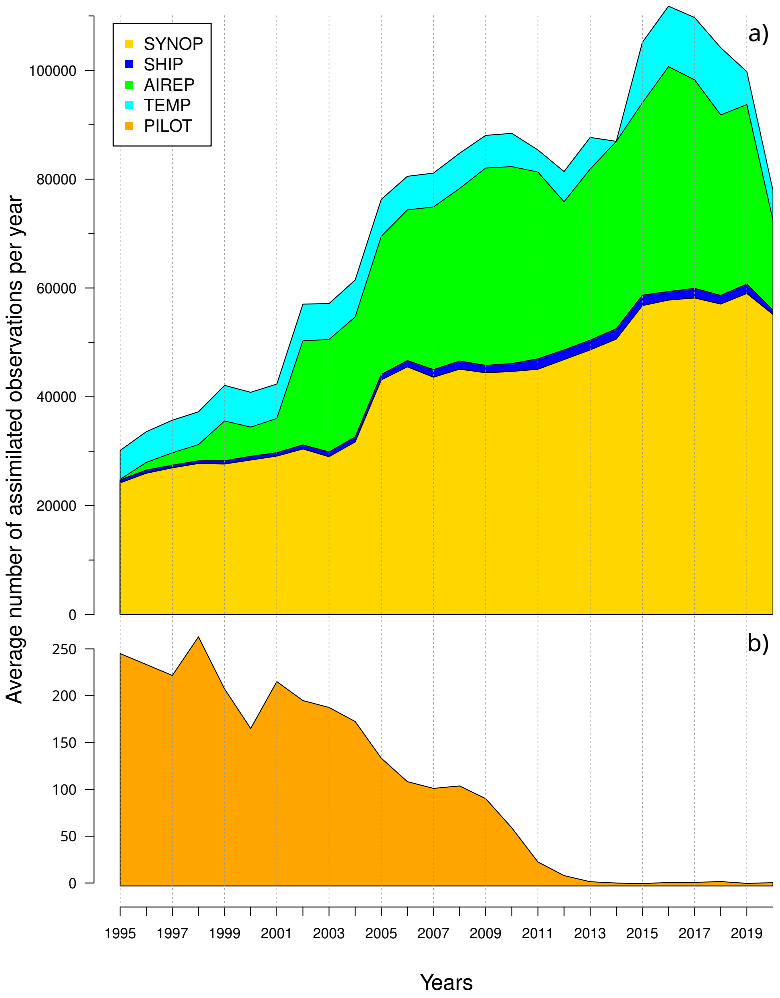

Figure 2.2: Average number of assimilated observations in SPHERA per year from 1995 to 2020. The assimilated observational datasets are reported in different colors. Panel a): SYNOP (in yellow), SHIP (in blue), AIREP (in green), and TEMP (in light blue). Panel b): PILOT (in orange), separated from the rest of the assimilated data given its lower numerosity.

always remains smaller than the largest term of the dynamics in the model equations in order to relax the model fields towards the observations without significantly disturbing their dynamic balance. The reader is referred to Schraff and Hess (2013) for a detailed description of the nudging method implemented in COSMO. The set of observational data nudged in SPHERA comes from the ECMWF catalog. It comprises near-surface observations over land and sea (SYNOP and SHIP), radiosounding and radar profilers data (TEMP and PILOT), and aircraft reports (AIREP). The meteorological variables assimilated include meridional and zonal wind speed components, pressure, air humidity, and temperature, except for temperature at 2 meters. All the assimilated observations are operationally quality-checked before the assimilation process and thus accepted or rejected following the standard procedure implemented in COSMO (details are reported in Schraff and Hess, 2013).

Figure 2.2 reports the average number of accepted observations per year for each assimilated dataset. Notably, their number increased almost constantly during the 26 years of SPHERA, from approximately 30 thousand in 1995 to roughly 100 thousand in 2019. This may have introduced artificial trends in the reanalysis estimates, owing to the inevitable increase of their quality with time due to the constraint of the numerical simulations with a larger number of observations. However, it is believed that an increasing quality is preferable to the maintenance of temporal homogeneity of the assimilated data. A sensitivity analysis to quantify the impact of the number of assimilated observations over time would shed more light on the actual weight of this effect. However, this would have required a substantial effort beyond the scope of the present thesis. During the last year of the dataset (2020), the number of accepted records remarkably reduced for SYNOP, SHIP, TEMP, and especially AIREP observation types (Figure 2.2a). This decrease is due to the lower number of available observations (and not to an increased data rejection), which have possibly originated from the globally reduced observational activity during the first year of the COVID-19 pandemic spreading (Riishojgaard, 2020). With respect to PILOT records (Figure 2.2b), a consistent decrease in the number of data with time is noted, approaching 0 after 2012. Also in this case, this is not a consequence of an increased rejection ratio in the assimilation process and is marginal compared to the much larger numerosity of the other types of assimilated data in SPHERA considering the whole 26-year period.

Chapter 3

Improving heavy-rainfall representation with SPHERA

This Chapter is devoted to the comparative analysis between SPHERA and its driver ERA5 to assess their ability in representing precipitation over a 15-year period (2003-2017) over Italy, and to highlight any possible added value owing to the CP description of the novel reanalysis. The simulations are evaluated against the independent observative dataset of the national rain-gauges network, maintained by the National Department of Civil Protection (DPCN), in terms of spatial distribution and daily and hourly precipitation accumulations, focusing on heavy rainfall events. The statistical assessment is performed with objective verification methods, considering various skill scores to evaluate different aspects of the simulations. As precipitation is a discontinuous atmospheric variable that may be characterized by strong local forcings and rapid temporal evolution, especially in the case of convective-driven events, a fuzzy methodology allowing for a certain degree of displacement between simulations and observation estimates is adopted. For this purpose, a preliminary sensitivity analysis to identify the optimal upscaling spatial dimension is performed. Finally, to better understand the performances of the two reanalyses in representing precipitation in specific conditions, two case studies of severe-precipitation events are additionally analyzed in detail.

The Chapter is organized as follows: Section 3.1 presents the observational data and the statistical methods considered for the verification performance analysis; Section 3.2 reports the sensitivity analysis for identifying the optimal dimension to upscale the precipitation fields, the results of the statistical verification in terms of spatial distribution, daily and hourly rainfall accumulations and the representation of the diurnal cycle of summer precipitation are presented in Section 3.3, while the analyses for the two heavy precipitation events are described in Section 3.4; finally, the findings are discussed and compared with similar works in Section 3.5.

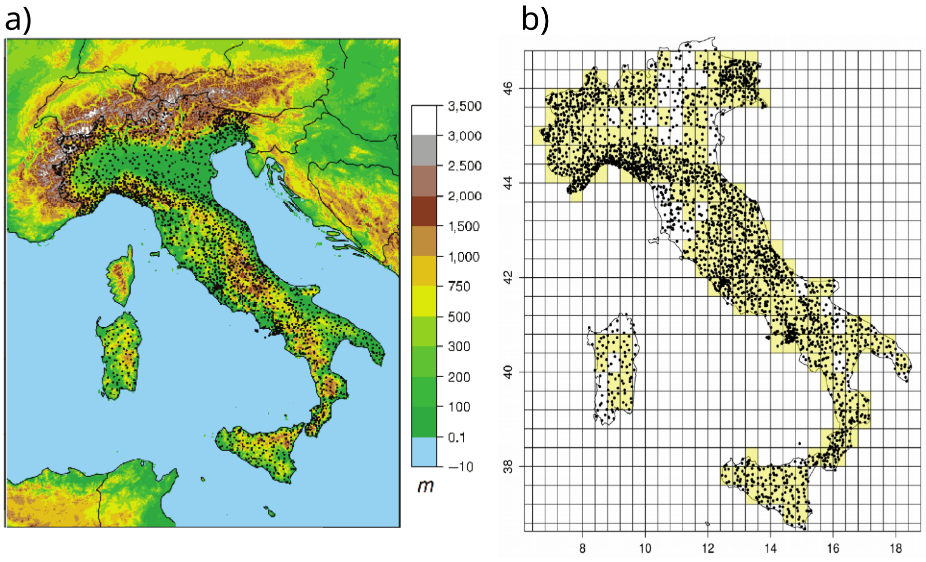

Figure 3.1: Dewetra rain-gauge observations. Panel a): spatial distribution of their location. Panel b) example of spatial upscaling operated in the performance process, with yellow boxes containing a sufficient number of pluviometers to represent the observed state (i.e., $\geqslant 5$), conversely to white boxes, which are then discarded from the analysis

3.1 Data and validation strategy

3.1.1 Rain-gauge observations

The performance of SPHERA and ERA5 is assessed through a comparison with rainfall rain-gauges observations. These pertain to the national composite of meteorological stations available on Dewetra, the Italian Civil Protection database of pluviometers maintained by the DPCN (Italian Civil Protection Department & CIMA Research Foundation, 2014), which contains the entirety of the Italian regional networks for precipitation and covers the period 2003-2017. These observations are independent of both ERA5 and SPHERA as they are assimilated in neither reanalyses. Dewetra contains a set of ground pluviometers supplying data at hourly temporal frequency, whose spatial distribution is reported in Figure 3.1a.

The number of rain gauges has significantly increased from approximately 1500 in 2003 to 6200 in 2017. Pluviometric data have been quality-checked operatively by the Department of Civil Protection. Additional controls have been performed for the most extreme rainfalls recorded (i.e., daily accumulations exceeding 500 mm). Anyhow, rain-gauge data are always affected by limitations, such as spatial inhomogeneity due to the low density of pluviometers in some areas or the sensors inadequacy to detect snowfall precipitation properly. Over the Italian territory, the former is a relevant issue in southern regions (where the situation has improved significantly in recent years), and in mountainous terrains (Crespi et al., 2018). The second issue pertains more to the northern and central regions during winter, particularly over mountains. Nevertheless, given the comparative nature of this analysis, these limitations should not be a major issue for assessing the added benefits of one reanalysis over the other with respect to their long-term statistical performance. On the other hand, for the case studies reported in Section 3.4, the results may be substantially affected by the spatial inhomogeneity of Dewetra pluviometers. For this reason, only for these specific analyses, two additional sources of rainfall observations to compare with the simulations are considered, namely: ARCIS (ARchivio Climatologico per l'Italia Settentrionale) and SRI (surface rainfall intensity) estimates. ARCIS is the high-resolution gridded precipitation analysis of pluviometric data covering north-central Italy at roughly 5-km spacing (Pavan et al., 2019). The analysis dataset is available over 1961-2015, including data from 1761 rain gauges, which are systematically controlled for quality, time consistency, synchronicity and statistical homogeneity, and then interpolated to a common grid using a modified Shepard scheme. SRI estimates are obtained by blending rain-gauge observations with radar reflectivity measures to overcome individual data limitations. The merging can be done with various methods, one of those providing the most accurate estimates is the kringing with external drift method (Goudenhoofdt and Delobbe, 2009) which is employed in this analysis. Indeed, the merging of these two different sources of rainfall information can lead to substantial improvements in the accuracy of the estimated precipitation state used for reference (e.g., Gabriele et al., 2017).

3.1.2 Validation strategy

A fuzzy verification approach (Ebert, 2008) is used to compare the precipitation performances of SPHERA and ERA5. With a fuzzy method, the matching conditions between the model and observations are relaxed, and the forecast is required to be in approximate agreement with the observation (by being close in space, time, or other quantities). This is alternative to traditional nearest-point verifications, seeking exact matches between forecast/observation pairs. Fuzzy methods are preferred when considering high-resolution models for which matching the observations with absolute precision is too difficult, due to the increased risk of "double-penalty" issues (i.e., to correctly forecast a situation but being offset from the observation). Indeed, small displacements in this context doubly penalize the simulation by producing a false alarm and missing the observation. Additionally, given the sparse nature of rain-gauge observations, a spatial aggregation is necessary to objectively assess the quality of the model simulation (Bollmeyer et al., 2015) and to compensate for representativeness limitations that may affect individual point observations (e.g., Weusthoff et al., 2010). These issues are particularly relevant in the representation of highly localized phenomena, such as severe convective precipitations, characterized by low spatio-temporal predictability. Indeed, their description needs high-resolution simulations, but the outcomes would be strongly penalized if traditional point-to-point comparisons were used (Lanciani et al., 2008).

With fuzzy methods, a spatial neighborhood to assess the closeness between forecasts and observations is defined around a certain point of interest. This is achieved by systematically upscaling (or boxing) forecasts and observations to a common coarser grid (compared to the respective original ones). The verification domain is subdivided into boxes of the same size, each containing a certain number of reanalysis and observation points (e.g., Weygandt et al., 2004). This process is schematically represented in Figure 3.1b. The upscaled distributions are then aggregated to provide a single forecast-observation pair for each box. The comparison between the aggregated data pairs is carried out considering both the mean and maximum (or 95th percentile) of their distributions within each box of the upscaled domain (similar to Marsigli et al., 2008). This is needed to evaluate different characteristics of the precipitation distributions. The comparison between the mean values can indicate the average skill of a reanalysis in representing precipitation. In contrast, comparing the maxima (or the 95th percentiles) can assess the ability to describe intense and localized rainfalls by considering the most extreme events (or the tail of extremes, respectively). Particularly, from the latter analysis, a high-resolution system is expected to reveal its added value over its coarser counterpart (e.g., Klasa et al., 2018). These distribution parameters have different impacts on identifying a useful forecast (i.e., a positive match between reanalysis and observations within a box). The matching between maxima is more generous than the averages: for the former case, a useful forecast implies its maximum in the neighborhood to be close to the observative counterpart, i.e., the match relies only on one value of each distribution. For the averages, the closeness criterion is more restrictive since their values are calculated from the whole distributions found in the neighborhood (Ebert, 2008). For this reason, comparing the performances obtained through different statistical aggregations is also useful to understand whether possible deviations from the observed states are due to the method used for aggregating the meteorological fields. Moreover, an observational mask is applied to retain only the grid cells where appropriate closeness between reanalysis-observations is met (similar to Marsigli et al., 2008). This is done to minimize the errors owing to the heterogeneous pluviometers distribution and the larger spatial domain covered by the reanalysis. In particular, for any dimension of the upscaled grid hereafter considered (see Section 3.2), boxes containing less than five rain gauges are withdrawn to avoid errors caused by insufficient sampling of the observed distribution (represented as white boxes for the exemplary upscaling shown in Figure 3.1b). This number has been defined for the operational verification of numerical forecasts at

		Observed		
		Yes	No	
Forecast	Yes	a Hits	b False alarms	a+b forecast yes
	No	c Misses	d Correct negatives	c+d forecast no
		a+c observed yes	b+d observed no	a+b+c+d total

Table 3.1: 2x2 contingency table representing the relationship between forecast/observation pairs for the dichotomous non-probabilistic verification related to the occurrence of an event.

ARPAE as the minimum required to allow a representation of precipitation suitable for this work. The upscaled grid neighborhood (or box) size defines the allowed spatial displacement between forecast and observations. The most appropriate definition of this horizontal scale is not trivial. It depends on various factors (e.g., the particular features investigated through the simulations, the time resolution, the meteorological situation, and the model itself, Gallus Jr, 2002; Ebert, 2008). Therefore, a preliminary sensitivity analysis on the neighborhood size is performed to investigate how reanalysis skill scores vary with grid size and to find the optimal grid spacing for the upscaling (described in Section 3.2).

The performance assessment of SPHERA and ERA5 is carried out with a categorical approach. A 2x2 contingency table (Table 3.1) is obtained by classifying precipitation occurrences in each box as yes/no events based on the exceedance of a rainfall threshold and aggregating the entire spatial domain considered. The reader is referred to Wilks (2019) for a detailed description of these statistical methods. A contingency table is built for each threshold considered from the joint distributions of forecasts (i.e., reanalysis estimates, in our case) and observations, determining the frequency of forecast/observation pairs falling within each of the four possible categories (a, b, c and d in Table 3.1). With the contingency table, a complete representation of the joint distribution between the binary observations and the binary outcome of the simulations (i.e., event observed/forecasted or not observed/not forecasted) is gained. From the results of the contingency tables, it is possible to calculate a variety of scalar attributes assessing different aspects of the simulations under exam. Particularly, in the context of the comparative performance analysis carried on, we considered the following dichotomous scores (considering the terminology defined in Table 3.1):

- Probability of Detection (POD), or hit rate: describes the fraction of observed "yes" that has been correctly forecasted, ranges between 0 (i.e., worst possible forecast: not even one observed event has been forecasted) and 1 (i.e., perfect forecast: all the observed events have been correctly simulated). It is defined as
$$POD = \frac{a}{a+c}$$
- False Alarm Ratio (FAR): describes the fraction of the simulated "yes" events that actually did not occur, ranges between 0 (i.e., perfect forecast: the simulations did produce not even one false detection) and 1 (i.e., worst possible forecast: all the simulations produced false event detections). It can be expressed as the Success Ratio (SR) through the relation SR = 1 - FAR, which is alternatively used (e.g., in the abscissa of the performance diagram). It is defined as
$$FAR = \frac{b}{a+b}$$
- Threat Score (TS) or Critical Success Index (CSI): measures the fraction of observed and/or forecast events that were correctly predicted, ranges between 0 (i.e., worst possible forecast: not even one forecast was correctly simulated) and 1 (i.e., perfect forecast: all the forecasts were correctly predicted without missing any observed event and producing any false alarm). TS can be expressed as a non-linear combination of POD and FAR, and this property is exploited in the performance diagram where POD is plotted against SR. It is defined as
$$TS = \frac{a}{a+b+c} = \left(\frac{1}{POD} + \frac{1}{1-FAR} - 1\right)^{-1}$$
- Frequency Bias: evaluates the deviation of the simulated- from the observed frequency of event occurrences, and it is defined as their ratio. With the frequency bias, it is possible to evaluate the tendency of the forecasting system to underestimate (BIAS < 1) or overestimate (BIAS > 1) the number of observed events. It ranges from 0 to infinity, and the score for the perfect forecast is 1 (i.e., equal frequency of occurrence between forecasts and observations). It is defined as
$$BIAS = \frac{a+b}{a+c}$$

The skill scores are calculated for the average and maximum values of the daily precipitation distributions, which are both needed to get a more complete assessment of the performance. Conversely, since the main interest of the present work is investigating precipitation extremes for rapidly-evolving cases, only maxima are considered for the hourly precipitation analysis. The resulting skill scores are aggregated over the entire verification domain and different temporal periods and reported through the performance diagram. The performance diagram (e.g., Figure 3.5) is a way to summarize the information of the four skill scores by exploiting their geometrical relationships, hence permitting a more comprehensive evaluation of the performance. A good forecast would lie in the top right corner of the diagram, where POD, SR, TS, and BIAS approach unity. At the same time, deviations in a particular direction would indicate relative differences in POD and SR and, thus, in TS and BIAS. The variability in the scores is estimated with a bootstrap resampling technique of 1000 new data samples and shown in the diagram as "cross-hairs". The reader is referred to Roebber (2009) for a detailed description of the performance diagram.

3.2 Sensitivity analysis for the upscaling

This section is devoted to the sensitivity analysis of the neighborhood size to find the optimal upscaling configuration and is performed on a sample period of three years (2015-2017). The maxima of SPHERA and Dewetra daily precipitation distributions are compared within boxes of different sizes before extending the verification to the entire 15-year period. The box dimensions cover a broad range from 0.14° x 0.14° to 1.85° x 1.85°, corresponding approximately to grids from 15 km x 15 km to 200 km x 200 km. Each of these grid spacings allows an adequate sampling of the number of grid points of SPHERA contained in the box, ranging from approximately 46 grid points (for the finest box of 15 km) to roughly 8264 grid points (for the coarsest box of 200 km).

Figure 3.2 reports the dependency of TS, POD, FAR, and frequency bias scores on the upscaled grid spacing and the daily precipitation intensity (with thresholds ranging from 1 to 150 mm/day). The results pertain to seasonal averages over summer (JJA) and winter (DJF). Two distinct effects contribute to a joint deterioration of the scores: the increase of precipitation intensity and the refinement of the upscaled grid spacing. The decrease of POD (panels c-d) and TS (a-b) and the increase of FAR (e-f) and frequency bias (g-h) for growing rainfall thresholds reflect the rising difficulty in detecting occurred events. As precipitation intensifies, the number of observations correctly simulated reduces, more false detections are produced, and the most extreme and rare events are overestimated. This positive frequency bias is reasonably dependent on the inadequate spatial sampling of the sensor networks to detect precipitation extremes. Indeed, if the mean daily precipitation is considered (Figure 3.7b), the overestimation disappears from rainfalls $\geqslant$20 mm/day. This observative limitation also plausibly explains the large number of false alarms detected, with FAR ranging 0.4-1.0 in JJA and 0.3-0.8 in DJF.

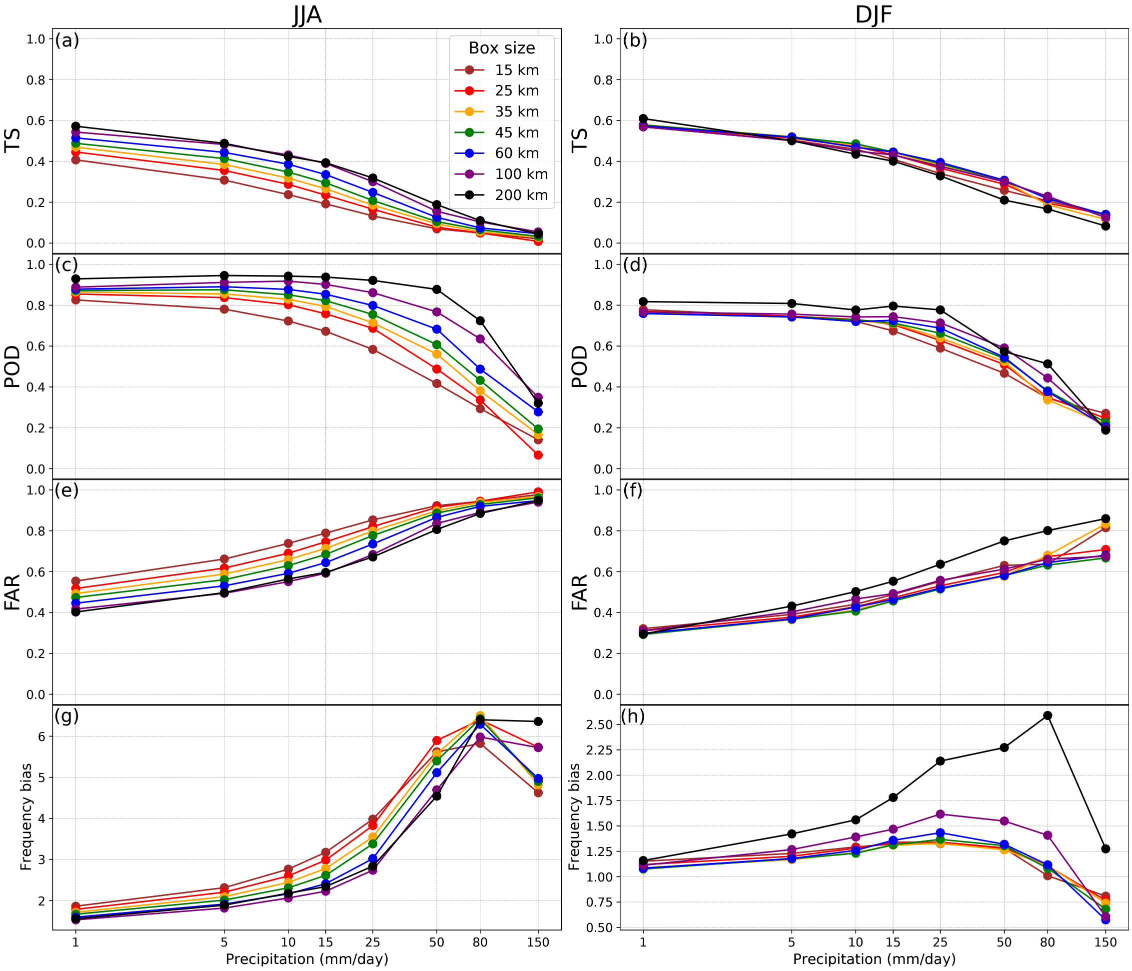

Figure 3.2: Seasonal averages of dichotomous scores over 2015-2017 obtained from the maximum values of the upscaled rainfall distributions within boxes of different sizes (indicated with different colors). From top to bottom: TS (panels a and b), POD (panels c and d), FAR (panels e and f), and frequency bias (panels g and h, note the different scale on the y-axis), as functions of daily precipitation threshold, on the left column for the summer season (JJA) on the right column for the winter season (DJF).

The performance degradation owing to the upscaled grid spacing is expected due to the rising impact of short-lived and localized events with low predictability. These are responsible for most spatio-temporal double-penalty errors, particularly in JJA (Marsigli et al., 2008). If the grid spacing is enlarged, the increased rain-gauge quantity per box and the removal of numerous small-scale occurrences affected by phase errors produce a better performance (Weygandt et al., 2004). However, simultaneously, the simulations lose sharpness due to the removal of actually occurred heavy rainfall events as well (Roberts and Lean, 2008). The performance improvement for coarser box sizes is enhanced in JJA when convective rainfall is most likely. Figure 3.2 shows a larger spread among different upscalings for all skill scores in JJA than in DJF. POD in JJA shows the strongest dependence on the grid size, especially for intense rainfall accumulations (⩾25 mm/day, Figure 3.2c). For example, in the case of 50 mm/day, POD ranges from approximately 0.4 (15 km box) to 0.9 (200 km box), while in DJF, the spread never exceeds 0.2 points. This effect is most likely imputable to different precipitation predictability during the two seasons: the detection of localized summer convective rainfalls is more challenging than the widespread frontal-like winter counterparts (e.g., Crespi et al., 2018). This seasonal dependency is also reflected by the systematic lower risk of committing false alarms in DJF, as expressed by the lower FAR scores (panel f) obtained for each threshold and grid size.

Finally, there is a further effect influencing the performance of the simulations and connected to the upscaling neighborhood size, which is linked to the typical spatial length of precipitation. This effect causes the opposite behavior in the performance on a seasonal basis, particularly visible for coarser boxes in terms of FAR and frequency bias (black and purple lines in Figure 3.2). In DJF, the worst performance is obtained with the 200-km grid box, while in JJA, with the 25-km one. Indeed, even if the risk of missing events is reduced with coarser grids due to the more generous matching condition, the likelihood of mixing different unrelated events falling within the same box is higher. This is particularly detrimental for the 200-km box case, likely because this spatial resolution may exceed the characteristic length scale of precipitation: while in JJA, typical convective rainfalls are associated with horizontal scales on the order of a few kilometers, in DJF, the horizontal extension of stratiform-like precipitation may vary from tens to few hundred of kilometers (Houze Jr, 2014). However, this length-scale issue does not seem to affect summer precipitation, most likely due to the higher impact of double penalties in localizing convective showers and the undersampling of the observed state in JJA, which mask this secondary effect.

The deviation between the normalized rainfall distributions of SPHERA and Dewetra as a function of rainfall intensity for different box sizes is also analyzed (Figure 3.3). The aim is to identify the spatial dimension that minimizes the difference between the two distributions, with particular emphasis on severe precipitation, to let the reanalysis maintain its resolution, i.e., providing predictions able to correctly distinguish situations with distinctly different frequencies of occurrence (Lewis et al., 2015; Murphy, 1993). For weak to moderate precipitation (up to 50-55 mm/day), the grid size best minimizing the difference is the coarsest one (200 km). In comparison, the others worsen almost linearly by reducing the box size. A reversal trend is evident mainly for intense rainfalls exceeding 60 mm/day. Indeed, the coarsest horizontal grid spacings (200 and 100 km) are always associated with the highest deterioration of the similarity between SPHERA and Dewetra. This excessive upscaling causes the removal of extreme rainfall occurrences, which are as-

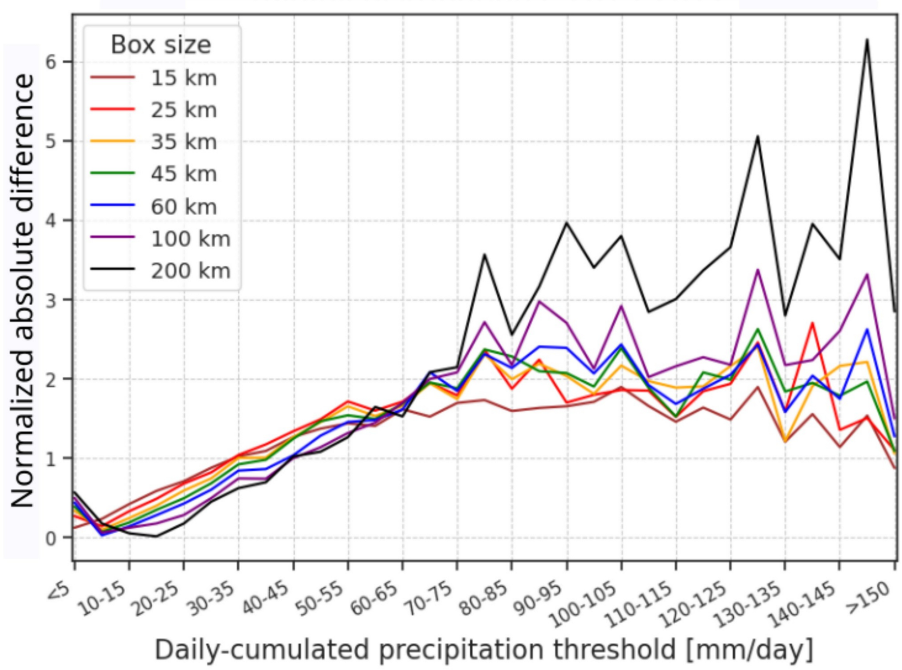

Figure 3.3: Normalized absolute difference between the number of precipitation events per daily threshold relative to the rainfall distributions of SPHERA and Dewetra observations for the period 2015-2017 for different horizontal resolutions of the upscaled domain (in different colors).

sociated with small-scale processes, inevitably smoothed out at these grid spacings. Moving to finer upscalings, the differences in this rainfall range are not prominent, except for the finest 15-km box always producing the most effective minimization. However, this grid spacing is excluded from the present analysis (as well as the 25-km box) since one of its main goals is the long-term comparison between SPHERA and its driver ERA5. ERA5 has a native grid spacing of 31 km. Therefore, upscaling to a 15-km grid (or 25-km grid) would result in a downscaling for ERA5, causing an underpopulation of the boxes containing less than one native ERA5 point. Considering the remaining box sizes (i.e., 35, 45, and 60 km), no decisive improvements of one over the other emerge at any intensity regime (Figures 3.2 and 3.3). However, boxes of 60 km contain enough grid points of ERA5 for its consistent representation (approximately 4 per box). For this reason, the 60-km grid spacing is chosen to perform the reanalysis evaluation in the next section. For any other application non-constrained by a comparative analysis, finer grid boxes should be considered (e.g., 15-25 km) for the investigation of severe precipitation events.

3.3 Evaluation of precipitation

This section presents the performance verification of SPHERA and ERA5 rainfall estimates against Dewetra pluviometric data over 15 years (2003-2017). Section 3.3.1 compares daily spatially-distributed rainfall fields when focusing on severe precipitation events. In Section 3.3.2, daily-accumulated precipitation is the subject of the fuzzy verification performed with the optimal upscaling resolution of 60 km. This analysis extends to hourly accumulations in Section 3.3.3. Finally, the mean diurnal precipitation cycle during the summer season is analyzed in Section 3.3.4.

3.3.1 Spatial distribution

A fundamental feature to investigate is the reanalysis ability to represent the "correct climate" in terms of the spatial distribution of precipitation. This section reports the analysis of the spatially-distributed daily observed state aggregated over 2003-2017, on the annual and seasonal terms (JJA and DJF), and the relative deviations of the reanalyses. In this context, the 95th percentile of the boxed distributions is chosen for the spatial intercomparison in order to focus on intense precipitation occurrences. Only in this section the precipitation fields are upscaled over a common grid of 31-km spacing, roughly corresponding to the ERA5 native grid. Indeed, the loss of detail resulting from the coarser optimal grid of 60 km (defined in Section 3.2 when aggregating the precipitation fields over the whole domain) would be too much for this spatial analysis.

Figure 3.4 (panels a,b,c) reports the observed 95th percentiles of daily rainfall distribution. Italy's wettest regions are visible from the annual plot: the northern Apennines and the far-eastern Alps, with daily accumulations exceeding 10 mm/day (the reader is referred to Figure 2.1 for geographical references). The western Alps and the south-central Apennines are other areas particularly prone to precipitation, with daily rainfalls of 5-8 mm/day. These precipitation patterns are mainly linked to the orographic enhancement of precipitation characterizing the Italian region (Napoli et al., 2019), of which an emblematic case study is reported in Section 3.4.1. The drier regions, presenting daily accumulations below 3 mm/day, are found over the plains and hilly areas, especially over the Po valley, the Sardinia and Sicily islands, and the Apulian peninsula. The observed patterns are consistent with several recent rainfall climatologies obtained with higher resolution datasets (Pavan et al., 2019; Longobardi et al., 2016; Crespi et al., 2018) and reflect the impact of the very complex Italian topography on the spatial behavior of precipitation. Therefore, the upscaling procedure does not affect the observed precipitation patterns estimated with Dewetra.

Considering reanalysis estimates on the annual term, ERA5 (Figure 3.4g)

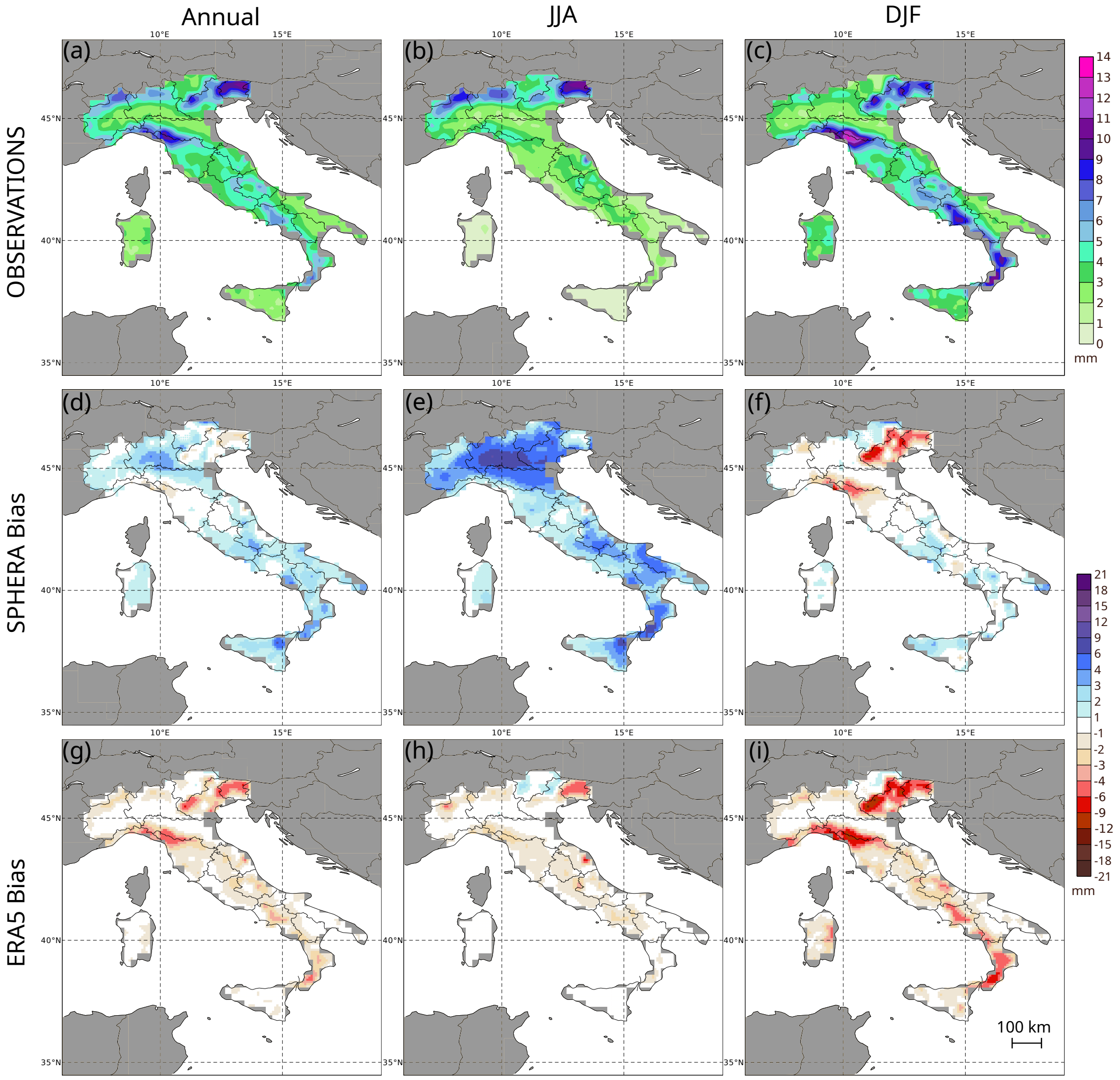

Figure 3.4: Observed (top row - panels a,b,c) average of the 95th percentile of daily precipitation distributions for different temporal aggregations over the years 2003-2017: annual (left column), summer (JJA, central column) and winter (DJF, right column). On the central (panels d,e,f) and bottom (panels g,h,i) rows are reported the daily deviations from the observed spatial distributions for SPHERA and ERA5, respectively. All three datasets are upscaled over a common grid of approximately 31-km horizontal grid spacing.

presents a heterogeneous dry bias, peaked over the wettest regions (i.e., primarily mountainous areas) and dampened over the plains, with an average relative bias per box of -1.1 mm/day. SPHERA also shows a heterogeneous distribution of the annual bias (panel d), but mainly linked with an overestimation of rainfall intensity, particularly over the Po valley and in southern Italy, with an average bias per box of 1.4 mm/day.

The distribution of the observed 95th percentiles over JJA and DJF (panels b and c) reveals the seasonality of precipitation: summer rainfalls are mainly enhanced over the Alps, particularly over their far-Eastern region, presenting daily accumulations above 10 mm/day; winter precipitation is characterized by even more orographic enhancement: rainfall peaks extend across the entire Apennines with daily accumulations larger than 13 and 10 mm/day at their northern and southern ends, respectively. These patterns are due to the dominant mesoscale humid flows impinging over mountainous ranges typical of the cold season (e.g., Miglietta and Davolio, 2022; Khodayar et al., 2021; Krichak et al., 2015).

Moving to reanalysis estimates, in JJA ERA5 (Figure 3.4h) shows a similar bias distribution to that of the annual term with a mean bias per box of -1.7 mm/day. SPHERA (panel e) presents a higher wet bias, marked over the Po valley, southern Apennines, and eastern Sicily, with a daily average bias per box of 2.9 mm/day. During wintertime, ERA5 (panel i) notably underestimates precipitation, especially over the cited mountainous wet spots, with an average bias of -1.7 mm/day per box. Also SPHERA in DJF (panel f) shows a dry bias over part of the northern Italy orography (with an average bias per box of 0.1 mm/day). This indicates a difficulty in correctly representing dynamically-driven orographic precipitation in specific regions, even for a CP model. However, weaker and more spatially limited dry biases are obtained compared to ERA5. Furthermore, ERA5 always underestimates precipitation if the long-term average of the 95th percentile distributions is considered, suggesting a systematic inability to simulate severe rainfall events. This is typical of low-resolution models that employ convection-parametrizing schemes and lack sufficient detail in the representation of topography, which is crucial especially over complex terrains. On the other hand, the substantial overestimation committed by SPHERA in summer should be interpreted as something other than a systematic wet bias of the driving model. Suppose the observed state is not sampled as well as the reanalysis (as potentially could be the case for Dewetra as highlighted in Sections 3.1.1 and 3.2, especially in mountainous regions). In that case, an overestimation of the high-resolution rainfall simulations is likely when considering the tail of extreme values of its distribution (such as the 95th percentile). This hypothesized undersampling of the observed state is supported by the analysis of the distribution for average daily rainfall intensities in the next section: a systematic overprediction of SPHERA for intense rainfall rates is not found in that case. These results confirm and extend the preliminary equivalent analysis over 2015-2016 reported in Cerenzia et al. (2022).

3.3.2 Daily precipitation

The daily accumulated precipitation of the two reanalyses is assessed through performance diagrams (Figure 3.5). The performance diagram exploits the geometrical relationship of four dichotomous scores: the probability of detection POD (on the

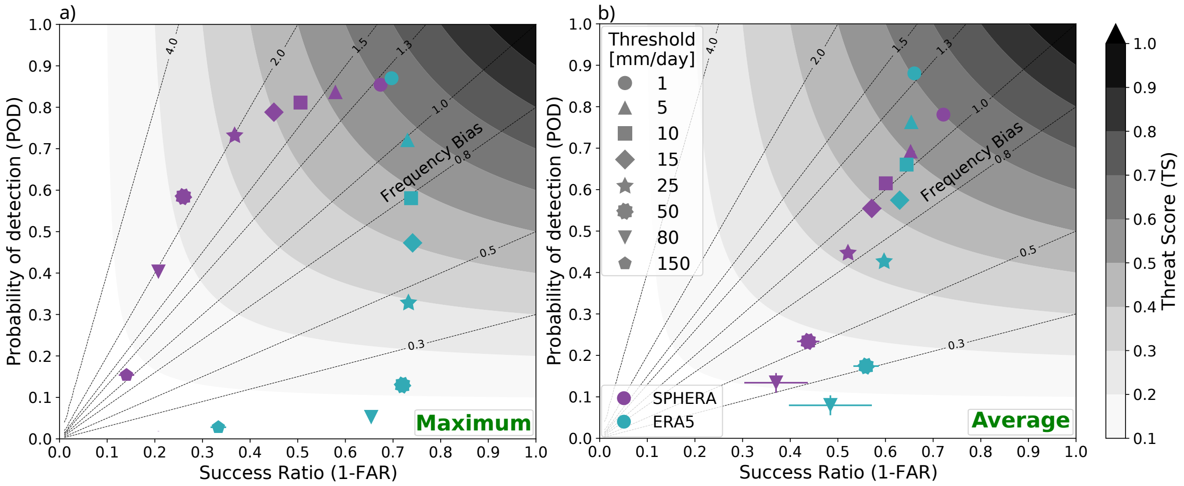

Figure 3.5: Performance diagrams for the aggregation of the scores over 2003-2017 for SPHERA (purple symbols) and ERA5 (turquoise symbols) when the maximum (panel a) and average (panel b) values over boxes of 60 km are considered. The threat score is indicated in different shades of gray, and the results pertaining to various daily-precipitation intensities (with thresholds ranging from 1 to 150 mm/day) are reported with different symbols. The frequency bias can be estimated from the deviations from the 45° black line, indicating an unbiased forecast (i.e., frequency bias = 1). Cross-hairs indicate the uncertainty related to the sampling variability of the data and are calculated from a bootstrap resampling of 1000 new samples.

y-axis), the false alarm ratio FAR (expressed on the x-axis as the success ratio SR=1-FAR), the threat score TS (indicated in the diagrams with gray shading contouring), and the frequency bias (shown by dashed diagonal lines). In the diagrams, the results of an accurate model would lie on the bisector line describing a null frequency bias, and a perfect simulation would lie on the top right corner of the diagram (i.e., POD, SR, and TS =1).

With respect to maximum precipitation (Figure 3.5a), ERA5 outperforms SPHERA in the range of weak rainfalls (1-10 mm/day), mainly due to fewer false alarms (i.e., higher SR score). As precipitation intensifies, the SR score stays almost constant for ERA5 (roughly 0.7) and is always higher than SPHERA counterpart, which gradually decreases from 0.7 to 0.2. Conversely, the hit rate of ERA5 (i.e., POD score), with a difference of more than 0.1 points for every consecutive threshold, decreases more rapidly than for SPHERA, which stays above 0.7 up to 25 mm/day. The result is a gradual increase in the score gap between the reanalyses in moderate-to heavy precipitation (25-80 mm/day). This suggests, on the one hand, a better ability of SPHERA to detect actually-occurred heavy rainfalls at the expense of a larger number of false detections. On the other hand, ERA5 shows a lower ability to simulate an adequate number of events for increasing rainfall intensity, but it keeps an almost constant skill to avoid false alarms due to the smaller sample of events

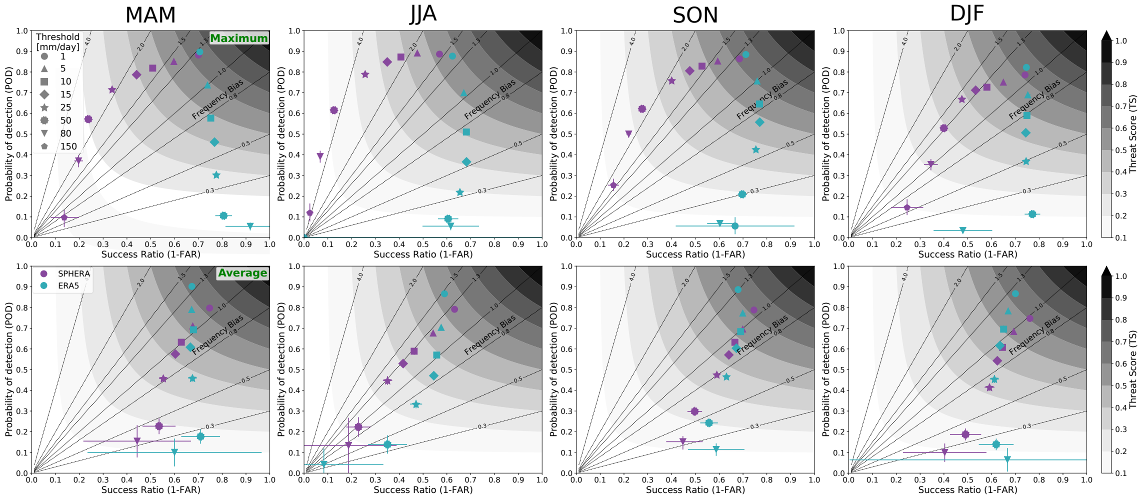

Figure 3.6: As Figure 3.5 but for seasonal aggregations over 2003 to 2017: MAM, JJA, SON, and DJF (columns from left to right) when maximum (upper row) and average (lower row) values over boxes of 60 km are considered.

simulated at increasing intensities. These outcomes are in line with the tendency of ERA5 to produce dry biases (as seen in Section 3.3.1), which are amplified for intense precipitation, as detected from the decreasing frequency bias from roughly 1.3 (1 mm/day) to <0.3 (80 mm/day). Conversely, SPHERA overestimates the number of events, with a frequency bias larger than 1 for the entire set of intensities, which is maximum (>2.0) for 50 mm/day, in line with the weakly wet annual bias found in Figure 3.4d.

The analysis of the distributions for average rainfall intensities (Figure 3.5b) confirms at first that the oversampling detected for maxima counterparts and for spatially-distributed fields is not linked to a systematic wet bias of the model driving SPHERA. In this case, a substantial reduction of the positive frequency bias of SPHERA is detected (always <1.3), which turns negative for rainfalls >15 mm/day and reaches a minimum of roughly 0.3 for 80 mm/day. This indicates a better agreement with the number of observed events (revealed by the lower deviation from the null frequency bias line at all intensities). Also ERA5 produces generally higher skill scores considering the averages. However, an increasing dry bias for intense accumulations, larger than SPHERA counterpart, is detected starting from 15 mm/day and reaches values <0.3 for 80 mm/day. This further proves the larger difficulty for ERA5 in producing a sufficient number of events even to match the average observed state. Both reanalyses present similar TS scores, suggesting comparable skills to represent the average daily precipitation for weak to moderate rainfalls. At the same time, the improvement of SPHERA is evident for the heaviest accumulations considered.

Figure 3.6 reports the seasonal analysis of the performance for maximum and average values in the boxes. The main differences compared to the annual aggregations are obtained in JJA and DJF, while MAM and SON are in line with the annual term performances (hence not discussed in the following). A general detriment of the scores is evident in JJA considering the maximum (upper row): ERA5 produces more false alarms and fewer hits, with SR and POD scores decreasing in the order of 0.1 points for most intensities. A performance deterioration in JJA is also detected for SPHERA in the form of a higher fraction of false detections (generally lower SR scores of 0.1 points than annual). Nevertheless, the ability to detect observed events is preserved and slightly enhanced compared to the annual term (POD scores systematically increased by roughly 0.1 points for most thresholds). This causes an exacerbation of the positive frequency bias at all intensities, which exceeds 4.0 for $\geqslant$50 mm/day.

As regards DJF, SPHERA performance shows a general opposite behavior compared to JJA (i.e., higher SR and lower POD and frequency bias at all intensities), as well as ERA5, even if less markedly (higher SR and lower POD). These results highlight the different seasonal predictability of precipitation. During summer, precipitation is mainly localized, short-lived, and linked to intense convective thunderstorms, making it typically more challenging to simulate compared to winter. In wintertime, stratiform rainfalls dynamically driven by large-scale frontal activity are more likely (Antolini et al., 2016; Houze Jr, 2014). Interestingly, the performance of SPHERA suggests the added ability of the high-resolution simulations, particularly in JJA when compared to ERA5, to reach heavy rainfall intensities at the expense of a high number of unrequited events (false alarms). As demonstrated previously, the latter is more likely a consequence of the undersampled observed state when considering extreme values (maxima) of precipitation rather than a deficit owing to the COSMO model. Indeed, moving to seasonal daily averages (Figure 3.6 - lower row), a decisive reduction of the frequency bias of SPHERA is detected for all seasons (in JJA confined between 1.0 and 1.3 for the majority of intensities), suggesting a good agreement with the mean observed precipitation state. For ERA5, similarly to the annual term, the bias (generally wet for intensities $\leqslant$10 mm/day and dry above) is reduced for the averages, but it always remains stronger than SPHERA counterpart (i.e., larger deviations from the 1:1 line at all intensities and seasons).

A complementary aspect in comparing simulated and observed precipitation series are the rates of daily rainfall distribution intensities, reported as histograms in Figure 3.7 for average and maximum values. Focusing on maxima (panel a), in the range below 5 mm/day, ERA5 overestimates while SPHERA underestimates the number of events. The trend is reversed for heavier intensities (i.e., $\geqslant$10-20 mm/day), with SPHERA overpredicting and ERA5 undersampling Dewetra. The maximum overestimation for SPHERA is detected in the range of 30-50 mm/day,

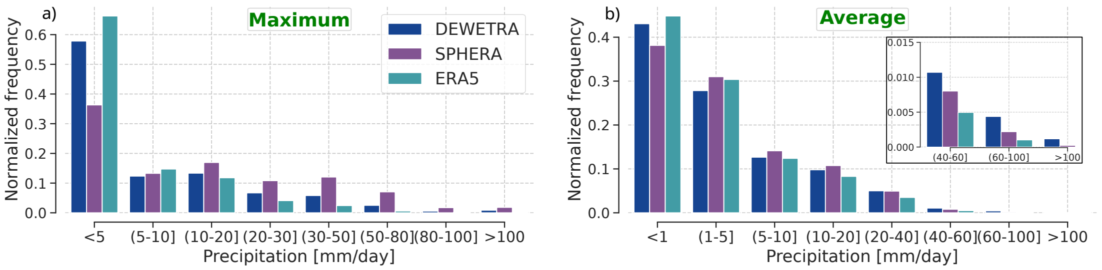

Figure 3.7: Normalized frequency histograms of the distributions of daily rainfall occurrences over 2003-2017 for SPHERA (in purple), ERA5 (in green), and Dewetra observations (in blue), when maximum (panel a) and average (panel b) values over boxes of 60 km are considered. For better visualization, the distributions of the average values for the highest thresholds are highlighted in a black-framed subplot.

decreasing then for higher thresholds but consistently producing more events than observed. The underestimation committed by ERA5 gets larger with rainfall intensity, producing almost no events above 80 mm/day. This supports the inadequacy of coarse and convection-parametrizing systems in representing severe precipitation occurrences. Conversely, SPHERA successfully simulates events at all rainfall intensities considered, even for the heaviest precipitation ranges ($\geqslant$ 80-100 mm/day), confirming the added value of CP models in this sense, as expected (e.g., Klasa et al., 2018). As previously discussed, SPHERA overestimating moderate- to heavy precipitation is more a throwback of the upscaling aggregation based on maxima rather than a systematic tendency of the reanalysis. If Dewetra undersamples precipitation maxima, representativity discrepancies may arise in datasets comparison when considering 60-km boxes. In this case, the number of SPHERA grid points falling within a box may be significantly larger than the corresponding number of rain gauges. This benefits the reanalysis to the disadvantage of the sparser pluviometric network when coming to detecting precipitation peaks. As proof of this, the frequency distribution analysis performed with the averages (Figure 3.7b) confirms the absence of the wet systematic bias in SPHERA: both reanalyses better agree with the observations, with SPHERA outperforming ERA5 starting from 5-10 mm/day. Moreover, ERA5 underestimates the number of rainfall events starting from 10-20 mm/day, although in a less pronounced way than maxima counterparts. This further confirms its increasing dry bias as precipitation intensifies.

3.3.3 Hourly precipitation

The seasonal performance of the sub-daily (hourly) time scale of precipitation is assessed through performance diagrams for maximum accumulations aggregated from 2003 to 2017 (Figure 3.8). A systematic worsening of the scores is evident compared to daily accumulation counterparts. Similarly to daily performance, the added value

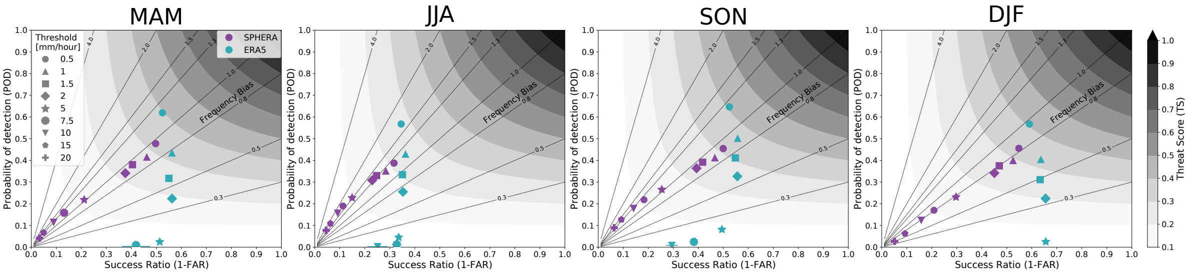

Figure 3.8: As Figure 3.6 but for hourly-accumulated precipitation, verified against Dewetra, when maximum values over boxes of 60 km are considered. The results pertain to hourly precipitation thresholds ranging from 0.5 to 20 mm/hour, reported with different symbols.

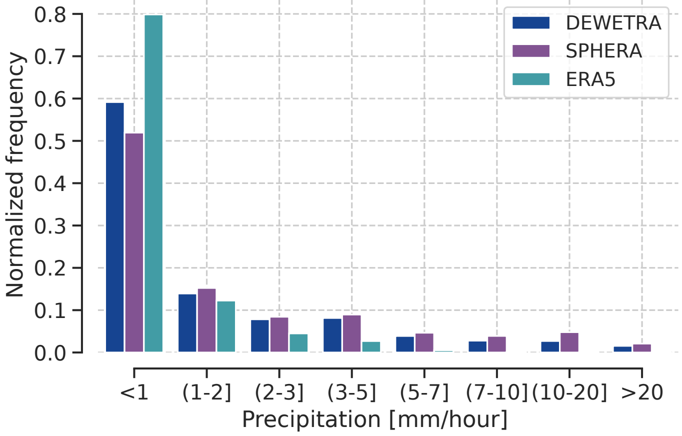

Figure 3.9: As Figure 3.7a, but for hourly rainfall occurrences.

of SPHERA over its driver emerges particularly in the larger fraction of hit events (starting from 2 mm/h for every season, but summer from 5 mm/h), as indicated by the higher POD, with gaps as large as 0.2 points more than ERA5. Further, as in the daily analysis, ERA5 systematically produces fewer false detections than SPHERA, as indicated by the larger SR values obtained for every season and threshold, but undersamples the number of hourly rainfalls as the intensity increases. The main difference with daily results is the strongly reduced overestimation of the number of events produced by SPHERA compared to the observations. Indeed, the frequency bias for hourly rainfall maxima is always close to 1 for every intensity and season: it is confined between 0.8 and 1.3 in MAM and SON, slightly dry in DJF (being arranged along the 0.8 line), and increasingly wet with rainfall intensity in JJA moving from roughly 1.0 to 2.0.

Figure 3.9 reports the histogram of the relative frequency distributions for hourly accumulations. The maximum hourly distributions reflect the behaviors detected for daily frequencies (Figure 3.7a). The overestimation committed by ERA5 in the number of low-precipitation events (<1 mm/h) is more pronounced, with the lowest bin populated by almost 80% of the entire sample of events, which is about

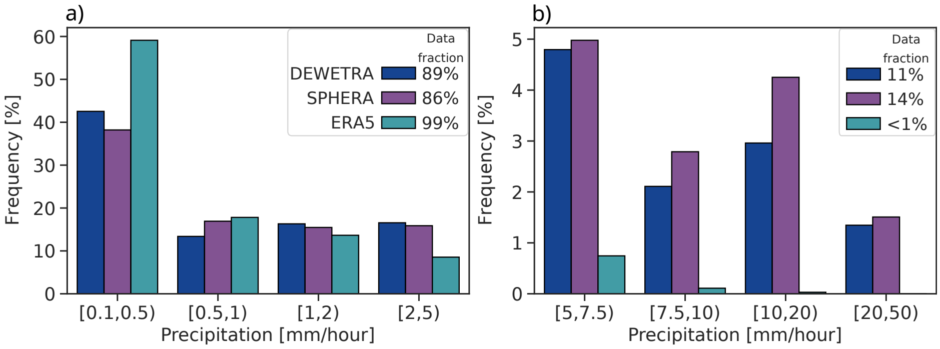

Figure 3.10: Relative frequency distributions of hourly rainfall occurrences for the summer months (JJA) aggregated over 2003-2017 for SPHERA (in purple), ERA5 (in green), and Dewetra (in blue). The distributions are obtained from the original unboxed horizontal resolutions of the datasets and are divided into two plots to separate lower-precipitation occurrences (panel a) from less frequent heavy rainfall rates (panel b).

20% more than observed. Starting from 1-2 mm/h, ERA5 frequency underestimation worsens as rainfall intensifies, producing an almost complete lack of occurrences for precipitation thresholds greater than 5-7 mm/h. Concerning SPHERA, despite the significant underestimation of roughly 10% in the number of weak rainfalls (<1 mm/h), for heavier accumulations the number of events is always in agreement with the observations, even in the range of severe events (>20mm/h). An overestimation with a frequency deviation roughly- or below 2% is present, with the maximum in the range of 10-20 mm/h. However, this effect is less pronounced than its daily counterpart in relative terms, indicating an improved skill in representing precipitation maxima at high temporal resolution.

As previously mentioned, when considering the maximum area-aggregated rainfall distributions over 60-km grids, important dynamical aspects of the simulations may be removed from the analysis, which is even more likely at hourly resolution. For this reason, it is worth investigating the frequency of hourly precipitation at the original resolutions of the datasets (i.e., 31 km for ERA5, 2.2 km for SPHERA, and scattered data points for pluviometers). It is worth highlighting the qualitative nature of this comparison given the numerous limitations arising from this approach (which were previously compensated by using a fuzzy verification method). Namely: the representativeness limitations of the uneven distribution of individual point observations, the non-fixed number of rain gauges in Dewetra over the years (increasing by a factor of 4), and the different sample sizes in grid points between the two reanalyses (as SPHERA is roughly 200 times denser than ERA5). The frequency distributions at the respective original resolutions of the datasets are reported in Figure 3.10. The results pertain to hourly summer rainfalls aggregated over 15 years when distinguishing between weak to moderate (Figure 3.10a) and severe (panel b) rainfalls to better visualize the less frequent severe events. Since

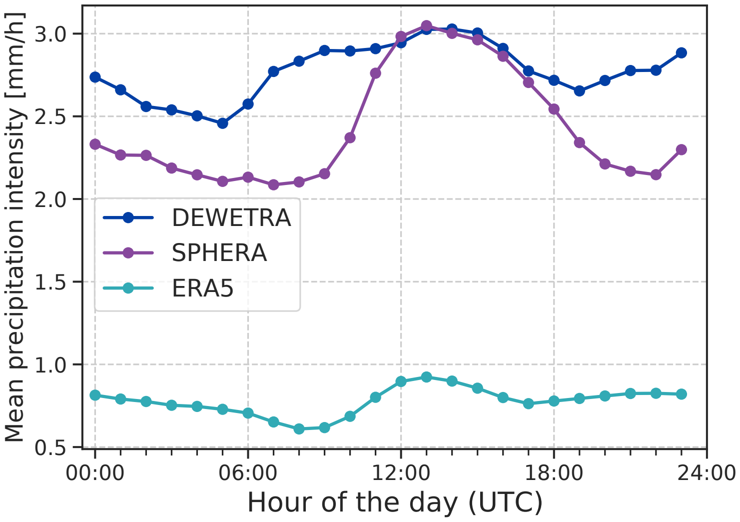

Figure 3.11: Mean diurnal cycle of hourly summer (JJA) precipitation intensity averaged over Italy during 2003-2017, considering only wet hours of the day. The time series are calculated from the datasets at their original resolutions.

numerical weather predictions frequently simulate substantial amounts of low rainfall occurrences, especially when their fields are not spatially aggregated, only wet hours (i.e., precipitation exceeding 0.1 mm/h, similar to Ban et al., 2014) are considered. The results confirm and strengthen the findings of the upscaled datasets: the substantial overestimation of weak rainfalls committed by ERA5 is evident, as well as its inefficiency in simulating heavier precipitation events, producing less than 1% of its occurrences, versus the 11% of the observed, for precipitation amounts greater than 5 mm/h. Additionally, SPHERA event distribution suits the observed counterpart well, without any systematic over- or underprediction. Indeed, SPHERA deviation from the observed frequency is always less than 5% in the range of weak-to moderate precipitation (0.1-5.0 mm/h). Moving to higher intensities, SPHERA generally overestimates Dewetra, with deviations in the frequency below 2%, resulting in a population of occurrences in the 5-50 mm/h range equivalent to the 14% of the total distribution, which is slightly larger than the observed 11%.

3.3.4 Diurnal cycle

A further relevant feature to assess is the ability to represent the diurnal cycle of summer precipitation. In JJA, daily variations of convective activity strongly control weather dynamics over Italy. Hence, we expect an enhanced description when employing CP simulations (Fosser et al., 2015; Brisson et al., 2016). The diurnal time series of hourly averaged rainfall during JJA for SPHERA, ERA5 and Dewetra, aggregated over 2003-2017, are reported in Figure 3.11. Land grid points only are selected for the reanalyses and averaged over Italy. At the same time, hourly unboxed rain-gauge data are aggregated with the same premises described at the end of Section 3.3.3. Only wet hours are considered also in this comparison. The observed cycle is characterized by higher rainfall intensities between 09 and 17 UTC, with the main peak at 13 UTC, and weaker rainfalls during the night and early morning, with a second lower peak between 23 and 00 UTC. ERA5 is unable to simulate the average observed precipitation rates, implying strong underestimations of the overall intensity and amplitude variation as a natural consequence of the coarse horizontal resolution (Bollmeyer et al., 2015), while the timing of the precipitation peak (13 UTC) is well captured. SPHERA diurnal cycle is more in line with the observations, with an optimal agreement during the rainiest hours of the day of the early afternoon (from 12 to 17 UTC) and a slight underestimation of the intensity (always $\leqslant$1 mm/h) for the rest. SPHERA well replicates rainfall temporal evolution. The main difference with the observed cycle is evident in the morning from 06 to 11 UTC when observations exhibit higher precipitation intensities. No substantial temporal shifts in the timing of the wettest and driest hours of the day are found for both reanalyses.

3.4 Heavy rainfall case studies

Besides the global statistical assessment for quantifying the overall ability of SPHERA and ERA5 to represent precipitation, it is relevant to investigate their simulation of specific meteorological conditions leading to extreme rainfall. For this reason, this section analyzes two relevant case studies of severe rainfall events in Italy: the severe orographic rainfalls leading to the flood of the Secchia river in January 2014 (Section 3.4.1) and the extreme precipitation associated with multiple mesoscale convective systems (MCS) over Sardinia in November 2013 (Section 3.4.2). For both events, an investigation of the dynamical configuration preceding the rainfall occurrence is presented to get a more comprehensive picture of the performance of the reanalyses.

3.4.1 Flood of the Secchia river (17-19 January 2014)

From the 17th to the 19th of January 2014, the passage of an upper-level trough over north-central Italy produced moderate to very intense rainfalls over the whole Po river valley. As a result, a large amount of rain fell over the Emilian catchments of the Secchia and Panaro rivers, two right tributaries of the Po river originating from the Apennines, thus leading to their flood. This event caused one fatality, the evacuation of approximately 10,000 people, and damages of roughly €500 million due to the failure of the Secchia river embankment (Porcù et al., 2019; D'Alpaos

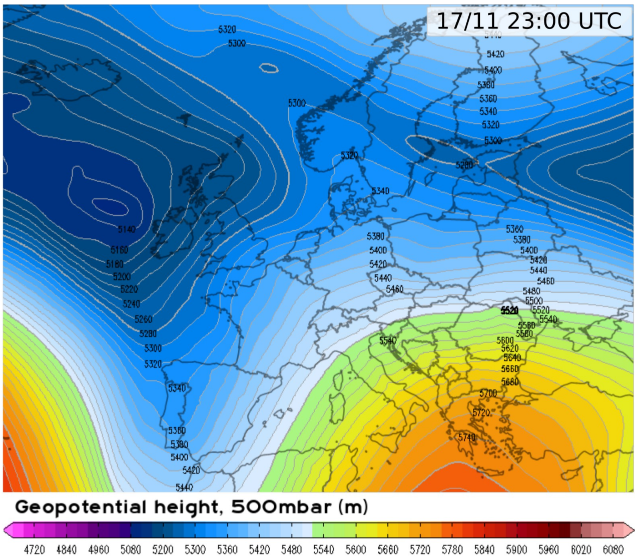

Figure 3.12: Geopotential height map at 500 hPa (from ERA5 reanalysis) over Europe and the northern Atlantic ocean on the 17th of January 2014 at 23 UTC (map courtesy of weather.us).

et al., 2014).

On the days preceding the event, the synoptic situation was characterized by an intensification of the subtropical jet stream, caused by the southward movement of a deep low-pressure trough centered on southern Iceland. From the geopotential height map at 500 hPa simulated by ERA5 during the night between the 17th and 18th of November (Figure 3.12), the deep upper-air trough extending down to northern Africa is well visible. Thanks to this baric configuration, the cold air masses carried southward by the trough were favored to ascend latitudinally over the Mediterranean and Tyrrhenian sea, where they received an important contribution of moisture. Subsequently, the polar and humid air masses, forced by the expansion of the anticyclonic ridge over the eastern Mediterranean sea, moved northeastward ahead of the cold front separating them from the warmer sub-tropical air, and reached the Italian coasts of Liguria and Tuscany. This is visible from Figure 3.13 reporting the geopotential height superposed to the horizontal wind field at the two isobaric levels of 950 hPa (Figure 3.13a) and 500 hPa (Figure 3.13b) as simulated by SPHERA on the 17th of January at 02 UTC. The intense flow blowing and extending throughout the vertical section of the atmosphere is evident, from the upper-levels (500hPa), with a north-eastern direction, to the lower-levels (950hPa), with a more northward direction. Near the surface, maximum wind intensity reaching 30-35 m/s was detected over the maritime sectors of the Thyrrenian and Ligurian seas, already from the early hours of the 17th. The sustained ventilation continued up to the evening of the 18th, with a progressive shifting of the jet-stream axis, and of the entire frontal system, towards south-east, as a result of the progression of the

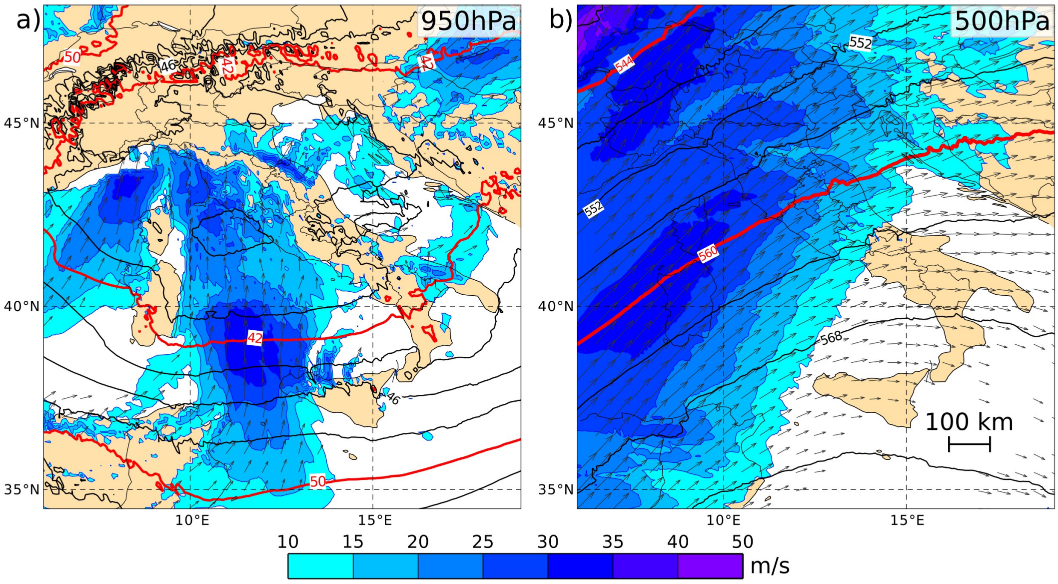

Figure 3.13: Geopotential height (black and red contours), in dam units, superposed to the wind field at 950hPa (panel a) and at 500hPa (panel b) simulated by SPHERA for the 17th of January 2014 at 02 UTC. The intensity and direction of wind speed fields are shown with blue color shading and black arrows, respectively (reported for an exceedance threshold of 10 m/s).

trough over the eastern Mediterranean (not shown).

On the 17th, the low-level flow bringing maritime moist air impacted over the Ligurian and Tuscanian coasts where it was forced to uplift by the presence of the northern Apennines. The result was the formation of intense and long-lasting orographic rainfalls starting over the mountains and extending on their leeward side over a large sector of the southern Po valley. Over the Tuscan-Emilian Apennines, high values of maximum wind speed were detected at mountainous meteorological stations over the Secchia, Panaro, and Reno catchments, in some cases exceeding values as large as 20 m/s for several consecutive hours (ARPAE, 2014). The associated precipitative event was active from the early morning of the 17th over the Apennine ridges, reaching heavy accumulations in a few hours, and intermittently persisting up to the evening of the 19th with a more widespread extension. Then, the decrease in the intensity of the moist air flow, and the counterclockwise rotation of the subtropical jet stream axis northward, led to the depletion of rainfalls over the area owing to their displacement towards southern Italy.

The accumulated precipitation in 72 hours simulated by SPHERA and ERA5 is compared to ARCIS analysis in Figure 3.14. The highest amounts of precipitation were reached over the Apennine sectors, especially over the Secchia river catchment (approximately around E 10.5°, N 44.3°). Here, upstream accumulations exceeded 250 mm/72h, affecting the overall river embankment stability and triggering its subsequent failure downstream. SPHERA simulated rainfall peaks reaching values

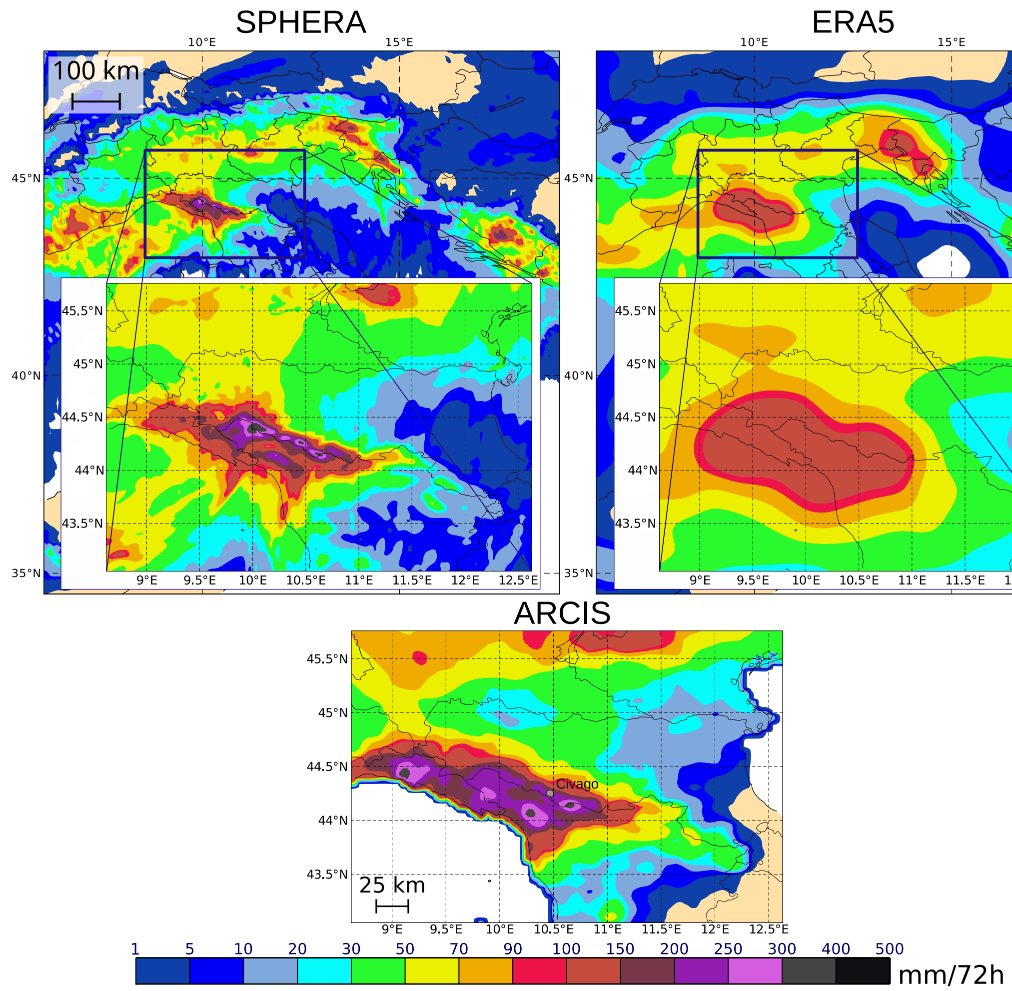

Figure 3.14: Maps of 72 hours-accumulated rainfall fields (from the 17th of January 00 UTC to the 19th of January 23 UTC) for SPHERA (upper-left panel), ERA5 (upper-right panel) over Italy, and ARCIS observative analysis (lower panel) over north-central Italy (the same sub-domain is reported also for the reanalyses maps). The solid gray dot represents the location of Civago where the ARPA rain gauge detected the maximum hourly amount of rainfall during the event (14.2 mm/h).

as high as 300 mm/72h over the river basin, in good agreement with the observed magnitude. In comparison, the rainfall field simulated by ERA5 did not exceed maximum accumulations of 150 mm/72h in the area. Furthermore, SPHERA spatial distribution presents complex details as a consequence also of the highly-resolved topography in the area. As a result, the localization of the most prominent rainfall peaks is sharp and in approximate (but non-exact) agreement with the respective observed pattern, with a tendency for displacing the wettest spots slightly towards north-east. On the other hand, the insufficient spatial detail of ERA5 produced a blunted, widespread, and smoothed rainfall field over the event area.

The observed maximum rainfall rates were recorded on the late morning of the 17th by a few isolated mountain rain gauges: the absolute maximum was detected

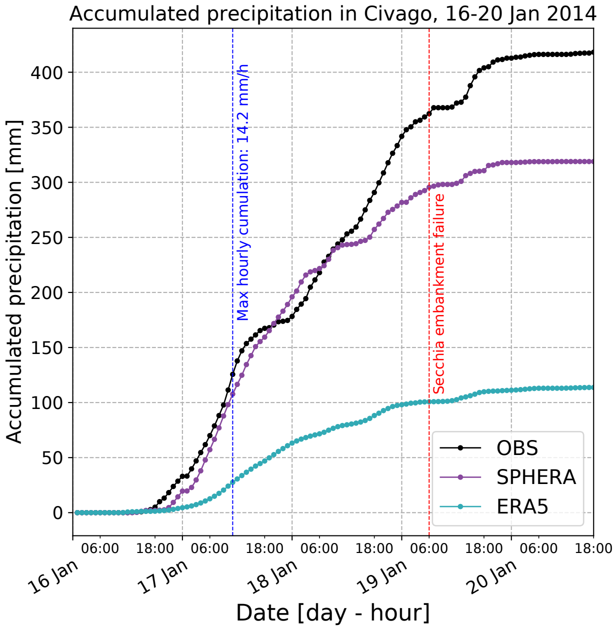

Figure 3.15: Accumulated precipitation for the period 16-20 January 2014, observed (black dots) and simulated by SPHERA (purple dots) and ERA5 (green dots) in the station of Civago. The vertical blue dashed line reports the timing of the observed hourly-cumulated precipitation peak, while the red dashed line indicates the timing of the Secchia river embankment failure downstream.

in Civago (1011 m a.s.l), located upstream in the Secchia river catchment, with 14.2 mm/h at 11 UTC. The most widespread and long-lived phenomena occurred during the evening and night of the 18th when almost the entirety of pluviometers located in the Panaro and Secchia rivers basins recorded more than 5 mm/h for at least three consecutive hours (Porcù et al., 2019). Figure 3.15 reports the temporal evolution of the heavy accumulation over Civago. A bilinear interpolation is applied to the reanalyses to perform a point-to-point comparison with pluviometric data. A total of 174.6 mm fell by the end of the first day of the event (17th). SPHERA well represents this steep accumulation trend by matching the observed peak time (between 9 and 12 UTC) and simulating 189.1 mm by the end of the 17th (i.e., roughly 8% more than observed). By the end of the third day, characterized by less intense and more spatially distributed rainfalls, SPHERA underestimates the total accumulation by 24%, producing 318.9 mm versus the 419.0 mm observed. The high-resolution simulation maintains a good match with rainfall timing throughout the event (as evident by the black and purple profiles in Figure 3.15, particularly during the intermittent phase of precipitation on the 19th of January). Moving to ERA5, its simulated field largely underestimates rainfall accumulation for the entire duration of the event. The underprediction is 113.8 mm (i.e., 65% less than observed) by the end of the 17th, increasing to 305.3 mm (i.e., 73%) by the end of the 19th. ERA5 also delays rainfall timing throughout the three days, indicating a lower ability to predict the correct temporal sequence during all the phases of the orographically-driven event.

This kind of extreme precipitation events, generated by the interaction with the orography of a strong and humid confluent flow ahead of a polar cold front, is frequent in the Mediterranean region during the winter season (Krichak et al., 2015). Evidence is provided by the observed spatial distribution in DJF (Figure 3.4c) and recent investigations (e.g., Grazzini et al., 2020a,b).

3.4.2 Cyclone Cleopatra and MCSs over Sardinia (18 November 2013)

During the fall season over the Mediterranean sea, a series of factors may favor the formation and organization of convective activity. Namely: the strong temperature gradients between the warm subtropical air and the colder northern air masses, the still-warm sea surface releasing large amounts of heat and moisture in the lower troposphere, and atmospheric instabilities caused by the more frequent changes in the baric configurations at these latitudes following the end of the warm season (Caillaud et al., 2021; Nieto et al., 2005; Jansa et al., 2000, 2001).

In this framework, the extratropical cyclone Cleopatra formed in November 2013, starting from a deep low-pressure trough in the westerlies and evolving into an upper-air cut-off low-pressure system centered in southern Europe. This system was well visible already from the 15th of November from the 500 hPa geopotential height map obtained with ERA5 (Figure 3.16a). On the following days, the cut-off system carrying the polar air mass moved towards south-west, reaching the southern Iberian peninsula (Figure 3.16b), where it stayed until the late afternoon of the 17th. In the meantime, its vertical extension reached the lower troposphere up to the formation of an enclosed low-pressure minimum at the surface over the Balearic sea. During the evening of the 17th, the cyclone started to move north-eastward, causing the advection of warm air coming from the subtropical continental African region towards the western Mediterranean, including Sardinia. This process was favored also by the presence of an intensification of the upper-air subtropical jet stream over northern Africa: a moderate jet streak was present (i.e., the fastest moving flow embedded in the jet stream with wind speed maxima >230 km/h), which inevitably interacted with the intense cyclonic circulation of the upper troposphere associated with the cut-off.

The interaction of the flows is visible from the fields of wind speed and direction at 300 hPa simulated with ERA5 for two distinct times (Figure 3.17). In the early morning of the 17th (Figure 3.17a), the beginning of the interaction between the jet

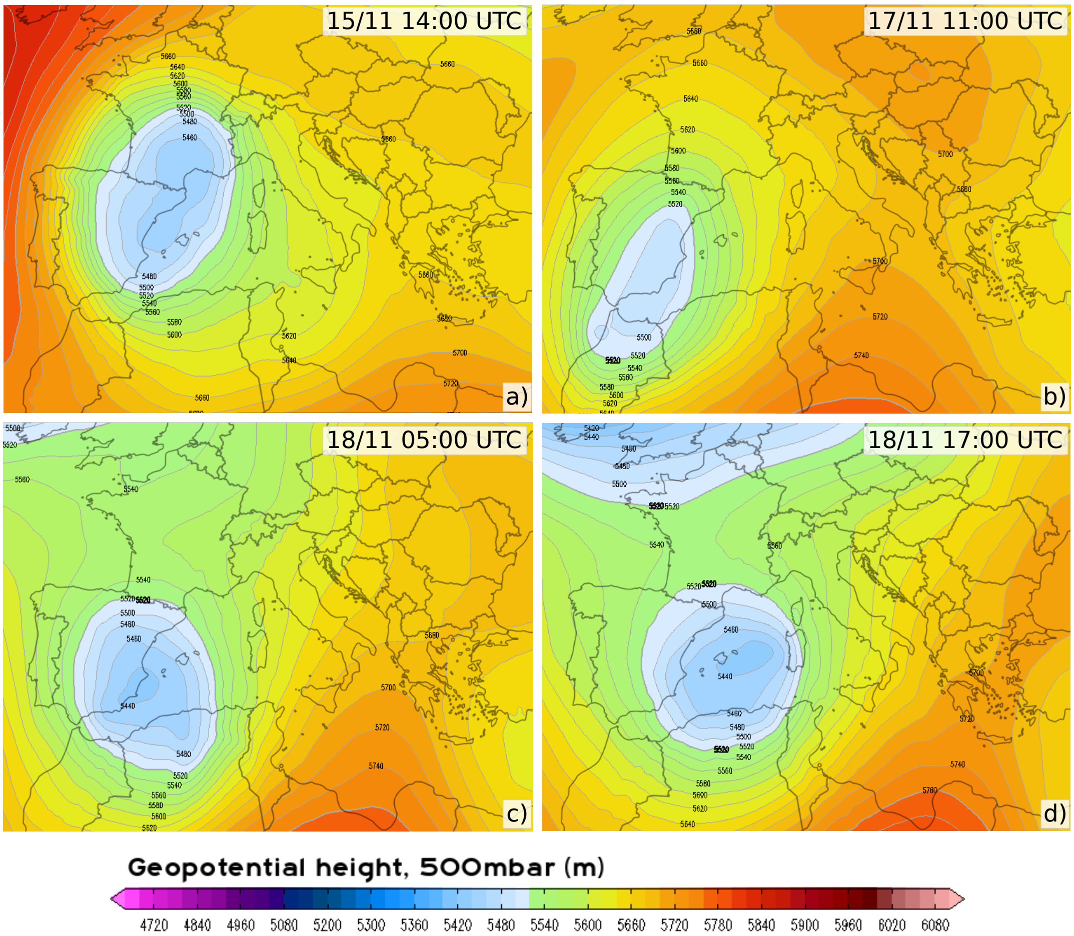

Figure 3.16: Geopotential height map at 500 hPa simulated with ERA5 over central-southern Europe and northern Africa. The four panels (a through d) show the evolution of the cyclone starting from the 15th of November 2013 at 14 UTC to the 18th of November at 17 UTC, respectively (maps courtesy of weather.us).

streak, located over northern Africa, and the cyclone, which circulation was flattened over the south-eastern Iberian peninsula, is evident. The interaction caused the formation of two distinct air flows presenting maximum wind speed, one following the eastward jet stream path and the other diverted towards the cyclonic circulation. By the early morning of the 18th (Figures 3.16c and 3.17b), the cyclone had moved towards north-east, and the meridional deviation of the jet streak was more developed and extended to larger sectors, causing the advection of warm subtropical air reaching the northern Mediterranean. The strong northward advection is well visible also from Figure 3.18 reporting the geopotential height superposed to the horizontal wind field simulated by SPHERA at 500hPa (right column) and at 950hPa (left column) at sequential temporal stages, from the evening of the 17th to the afternoon of the 18th of November 2013. The upper-level north-eastward ventilation started to reach Sardinia during the evening of the 17th as stronger pressure gradients approached the eastern Mediterranean (Figure 3.18b), while at the surface the wind

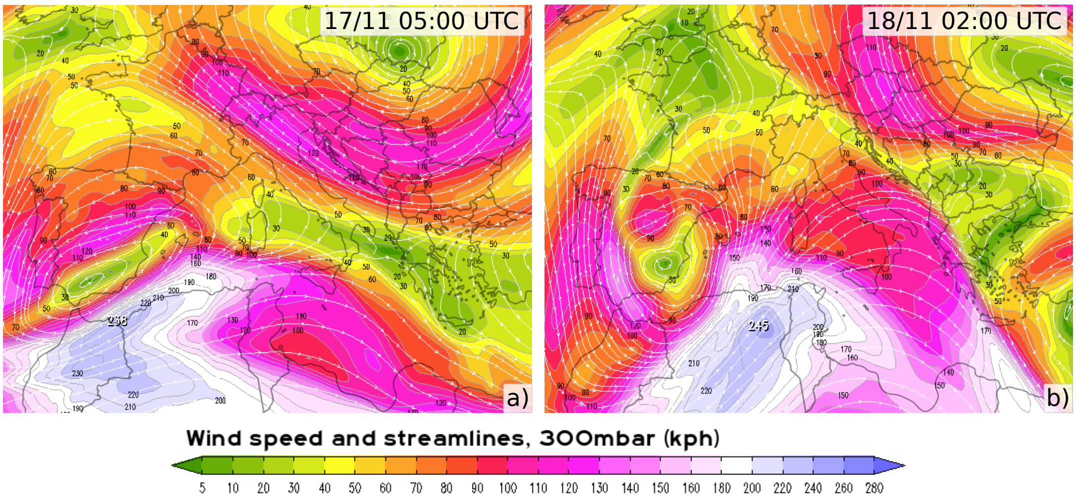

Figure 3.17: Wind speed at 300 hPa expressed in kilometers per hour (kph) and wind direction, represented as white streamlines, simulated with ERA5 over central-southern Europe and northern Africa. Panel a) for 17 November at 05 UTC, panel b) for the 18th of November at 02 UTC (maps courtesy of weather.us).

field remained weaker and less well established (Figure 3.18a). As the cyclone slowly moved eastward, the ventilation in the upper troposphere over Sardinia intensified (Figure 3.18d) and the flow aligned northward. At this stage, from the morning of the 18th, the vertical extension of the fast moving air flow reached the surface levels, where the further approach of the enclosed pressure minimum diverted the flow northwestward (Figure 3.18c), indicating the presence of a relevant wind shear between the surface and the upper atmosphere. This low-level Sirocco, pushed the warm African air over the Mediterranean sea surface, forcing the release of large amounts of moisture before reaching the Sardinian coasts. The large potential of these advected and humid air masses was triggered by the presence of the orography (presenting peaks as high as 1800 m a.s.l. in the central Gennargentu massif, approximately around E 9.5°, N 40°), that blocked the horizontal motion of the air flow by imparting a strong vertical acceleration to it. This blocking action is visible from the surface horizontal wind field (Figure 3.18c-e), characterized by strong gradients in the proximity of the eastern and south-eastern Sardininan coasts and in the central mountainous hinterland. The resulting vertical motions forced by the orography are visible from Figure 3.19a, showing the vertical wind speed component at 700 hPa obtained with SPHERA at 06 UTC of the 18th. Isolated peaks in the vertical wind speed are evident in the central-eastern sectors of Sardinia owing to the presence of the Gennargentu massif. Furthermore, the band structures oriented northeastward located on the northwestern part of the island (and also in southern Corsica) were due to the perpendicular orientation of the mountain ranges with respect to the air flow direction, resulting in strong vertical air lift extending from the central

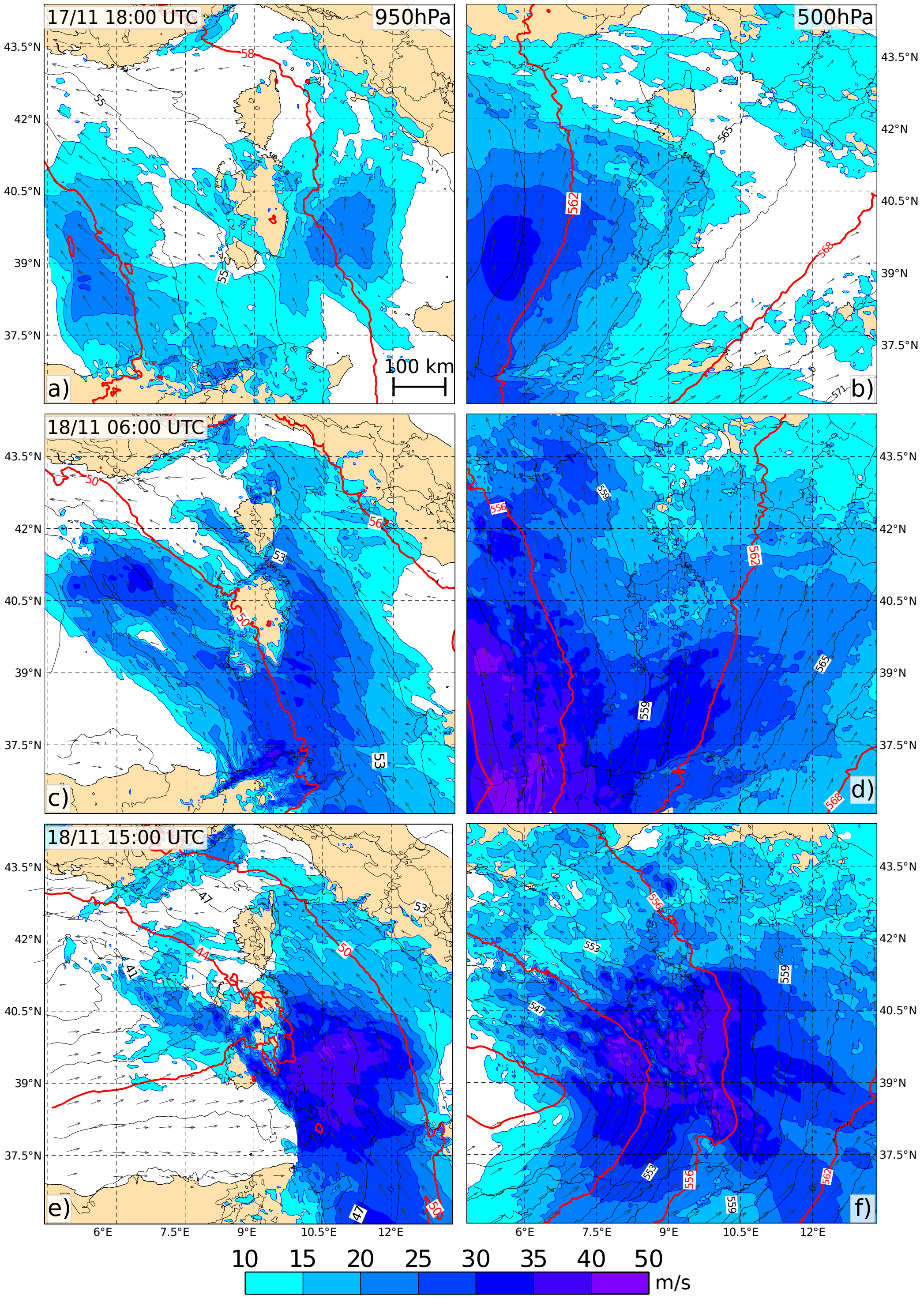

Figure 3.18: As Figure 3.13 but with the spatial domain centered over Sardinia and for three different times: at 18 UTC of the 17th of November (panels a and b), at 06 UTC (panels c and d) and at 15 UTC (panels e and f) of the 18th of November.

western- to the northern eastern coast. On the following temporal stages in the afternoon of the 18th (Figure 3.19b-e), stronger vertical motions organized along long narrow bands oriented towards north-west and moving northeastward are detected.

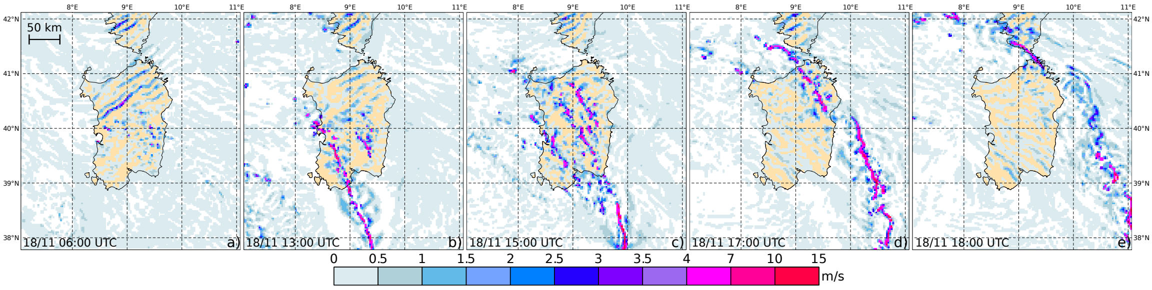

Figure 3.19: Vertical wind component field at 700hPa simulated by SPHERA at five sequential times, from 06 UTC (panel a) to 18 UTC (panel e) of the 18th of November.

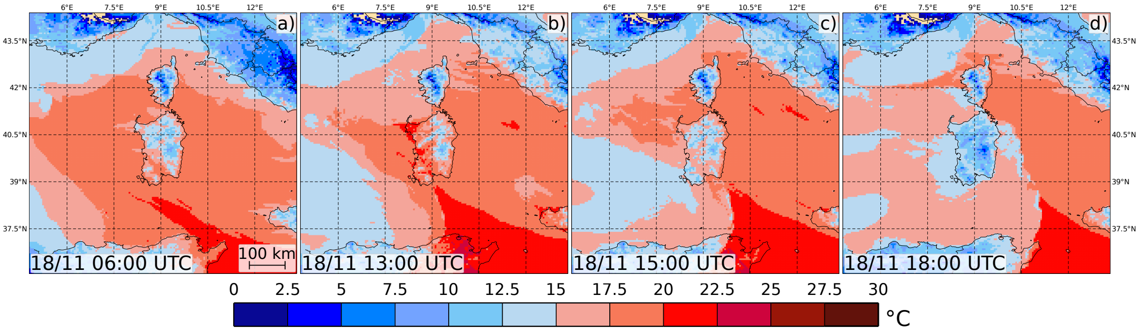

Figure 3.20: Temperature at 2 metres simulated by SPHERA at four sequential times, from 06 UTC (panel a) to 18 UTC (panel d) of the 18th of November.

These intense vertical velocities were caused by strong updrafts, owing to convective activity triggered by the front circulation, and organizing to form mesoscale convective systems (MCS), as discussed later. MCSs resulted by the combination of two processes: the aforementioned orographic lifting, and the collision of the two air masses at stake along the cyclonic cold front right over Sardinia (well visible also from Figure 3.18e). These are: the dry polar air associated to the cold front, slowly approaching from west and moving counterclockwise following the cyclonic motion, and the warm, humid and fast subtropical air advected from south-east, that is forced to lift over the colder air mass. The surface temperature field simulated by SPHERA gives an insight of the thermic differences between the two air masses (Figure 3.20). The low-level advection of warm air from southern latitudes was effective already at 06 UTC (Figure 3.20a). The arrival of the cyclone in the proximity of Sardinia is evident from the cooling of the surface air in the western Mediterranean, and from the strongest temperature gradients arranged along the line of the cold branch of the cyclonic front (Figure 3.20b-d).

This dynamical evolution resulted in extreme rainfalls causing extensive damage, loss of lives, and record-breaking accumulations (ARPAS, 2014). The approach of Cleopatra, and the interaction with the complex orography of Sardinia, led to the development on the 18th of November of multiple stationary MCSs over the island producing extreme precipitation. These are visible from the daily accumula-

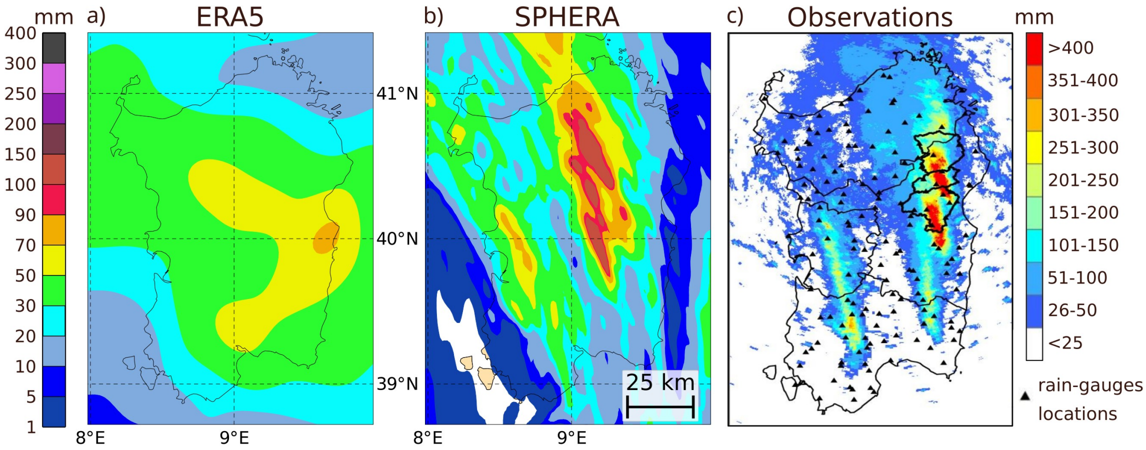

Figure 3.21: Daily accumulated precipitation on the 18th of November simulated by ERA5 (panel a), simulated by SPHERA (panel b), and obtained from the combination of radar estimates with pluviometric observations (SRI) (panel c - taken from Niedda et al., 2014). On panel c) the rain gauges over the territory are indicated with solid black triangles.

tion estimated through surface rainfall intensity (SRI) data obtained by combining radar with rain-gauge observations (Figure 3.21c). Figure3.21a-b report the daily accumulated rainfall fields simulated by ERA5 and SPHERA, respectively. The ability of the high-resolution reanalysis to simulate the convective band structures leading to severe rainfalls is evident. ERA5 produces a well-localized precipitation field. However, it results in excessively smoothed estimates with low spatial detail, and presenting daily accumulations below 90 mm. The accumulated intensities of SPHERA show lower underestimation than ERA5: the simulated values reach up to 150 mm versus the >400 mm detected, which can be expected when comparing reanalysis simulations to observational datasets in case of extreme convective rainfalls (Hu and Franzke, 2020).

To get better insights into the temporal evolution of the event, a comparison between SPHERA accumulated rainfall and SRI estimates at hourly temporal resolution is reported (Figure 3.22). During the early morning (at 5 UTC) of the 18th (Figure 3.22a), SRI fields reveal the initiation of moderate and localized rainfalls exceeding 12 mm/h in central-southern Sardinia. These are associated with the formation of the first convective cells starting to get organized along narrow bands aligned with the low-level flow direction (i.e., north-west) and resulting from the orographic lift of the advected subtropical air. The radar-based imagery then shows the permanence and development of the convective bands during the rest of the morning, gradually producing more sustained rainfalls with values larger than 25 mm/h (Figure 3.22b). Convective precipitation significantly intensified and extended starting from the early afternoon (14 UTC - Figure 3.22c), owing to the upscale growth of the MCSs, as a result of the gradual aggregation of the isolated convective cells,

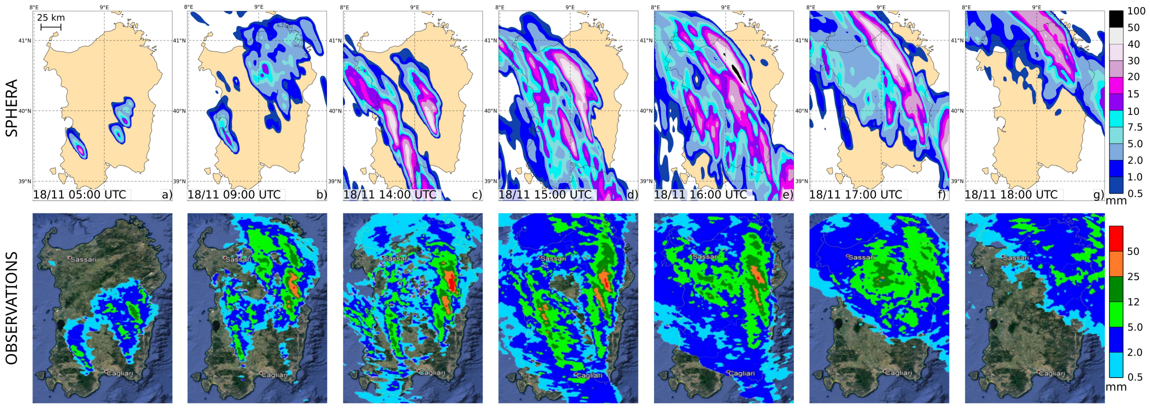

Figure 3.22: Hourly accumulated precipitation for progressive hours of the 18th of November (columns a through g) simulated by SPHERA (upper row) and estimated as SRI from radar/rain-gauge data (lower row).

with particular emphasis over central-eastern Sardinia. The peaked extension and strengthening of the MCSs derived from the uninterrupted orographic lift, coupled with the gradual approach of the strong thermal and pressure gradients linked with the arrival of the cyclonic front (not shown). This configuration triggered a decisive increase in the convective activity and organization, besides favoring strong stationarity and hence regeneration of the convective cells. The marked stationarity of the MCSs is visible from the comparison of SRI fields at 9 UTC (Figure 3.22b) with 14 UTC (Figure 3.22c), showing no significant spatial shifts of the heavy precipitation bands in five hours. The two most intense band structures located in central-eastern Sardinia were characterized by very similar geometries and dimensions, with a width in the range of 5-15 km, and presenting extreme spatial gradients in precipitation at their borders (approximately 200 mm/5km), as further proof of their sharped stationarity (Niedda et al., 2014). SPHERA successfully simulates the correct timing of hourly accumulations throughout the 18th of November, producing a remarkable agreement with the observed precipitation field (Figure 3.22c through g). Already from the early morning, the initiation of isolated and moderate rainfalls is satisfactorily simulated, forming along short bands that gradually organize and extend to larger areas. From 14 UTC, the simulated intensification and extension of the MCSs associated with the cold front passage are evident by the resulting heavier precipitation rates. However, SPHERA rainfall estimates are spatially shifted compared to the observations: the former are located more westward (of ∼20 km) and have their axis slightly tilted northwestward compared to the more northward-oriented radar-based structures. The cause is possibly a slightly different localization of the simulated surface low-pressure minimum of the cyclone (not shown): being further south and west than observed, it caused a slight tilt in the convective band axis orientation and a westward shift of their localization, respectively.

3.5 Summary and discussion

Rainfall is one of the most critical meteorological quantities to be estimated in numerical weather predictions and climate monitoring contexts. Severe rainfalls represent one of the most important causes of extended damages and societal costs in Europe (Spekkers et al., 2017; Rebora et al., 2013), hence the improvement in their representation is crucial (Feng et al., 2021). The assessment of the potential added value in the representation of extreme rainfalls as simulated by SPHERA, a new CP regional reanalysis over Italy, is the subject of the present investigation. The validation of precipitation simulated by SPHERA, and the comparison with its driver ERA5, is performed against the national rain-gauges network Dewetra for 2003-2017. The methodology consists of a spatial-neighborhood technique after assessing the optimal scale size to operate the upscaling aggregation. An excessive loss of detail in the representation of precipitation fields is detected, especially for heavy rainfall occurrences, when using grid boxes excessively wide (i.e., 100 or 200 km), as expected (Marsigli et al., 2008), while finer resolutions have proven to perform similarly to each other. Hence, a grid spacing permitting an adequate sampling of both SPHERA and ERA5 is chosen (i.e., 60 km).

The statistical analysis of precipitation permits assessing the added value of the CP system, particularly in terms of precipitation extremes for which higher benefits are expected. Concerning heavy-rainfall (i.e., the 95th percentile) geographical distributions, ERA5 locally underestimates precipitation intensity throughout the year, especially over the mountainous wettest regions (with a bias as large as -12 mm/day in DJF). SPHERA shows wet biases of up to 12 mm/day in JJA over the plains and dry biases of -9 mm/day in DJF. Moving to the analysis of daily-accumulated rainfalls, ERA5 better represents weak precipitation intensities with generally higher skill scores and weak wet biases. The benefits of the CP setting are evident for moderate and heavy accumulations in all seasons. In these cases, SPHERA successfully simulates severe rainfall occurrences, mainly related to warm-season locally-driven convective events, which ERA5 strongly underrepresents. The CP enhancement in skill scores shows a tendency to overpredict the number of events when considering daily distribution maxima. This is most likely due to undersampling of the observed state rather than a systematic deficit of the model, as proved by the almost null frequency bias obtained for daily average precipitation. In terms of hourly accumulation maxima, the accordance with the observed frequency stays stronger for SPHERA at all rainfall intensities. At the same time, ERA5 produces an insufficient number of events worsening with accumulation intensity. Further proof of the added value of SPHERA is the improved accordance with the observed average daily cycle of summer precipitation, compared to the consistent amplitude underestimation proper of ERA5.

Two case studies are considered to assess the ability of SPHERA to represent extreme precipitation and to further compare the reanalysis systems in describing specific severe-rainfall events. Both cases relate to extreme accumulations over a few hours resulting from orographic lifting and MCSs producing floods and extensive damage. ERA5 simulates blunted and smooth rainfall fields for both events, underpredicting the observed intense accumulations by 73% (when looking at the rainiest location for one of the events). SPHERA represents the precipitation fields with a higher level of detail in their spatial distribution, hourly-frequency timing, and intensity of extreme precipitation. Anyhow, an underestimation of 24% is detected at the same location (simulating an accumulated rainfall peak of 318 mm/72h vs. the observed 419 mm/72h). This result is expected when considering the simulation of heavy rainfall accumulations, even when performed with very fine spatial grid spacings (e.g., Buzzi et al., 2014, using 1.5-km CP simulations, obtained an underestimation of the 38% for a severe-precipitation event exceeding 500 mm/24h). The results suggest the potential of the CP system to reach high precipitation rates. In comparison, the stronger underestimation of the coarser ERA5 is most likely linked to the convection parameterization included in the driver model. In this case, smooth rainfall fields and a decreasing agreement with the observed state with increasing rainfall intensity are obtained, as expected when tailing the analysis towards the extremes (Bandhauer et al., 2022; Rivoire et al., 2021). The underestimation detected with SPHERA may potentially be linked to the spatial shifts in the localization of the rainiest areas and are acceptable in the context of high-resolution weather simulations. Indeed, these mismatches fall into the double-penalty class of occurrences, for which it is challenging to match the observations with absolute precision. The cause is the intrinsic chaotic behavior controlling DMC, which is responsible for the low predictability of the exact localization of the associated convective processes, justifying the use of fuzzy verification techniques for quantifying the performance of their simulations (Marsigli et al., 2021). This issue could be relevant in downstream modeling applications that require high spatial precision, such as in hydrological modeling (e.g., Lobligeois et al., 2014). Possible strategies to overcome spatial phase errors in high-resolution simulations when dealing with localized rainfalls are the additional assimilation of radar observations. This can be done by employing a latent heat nudging data assimilation scheme (Wahl et al., 2017), or by post-processing reanalysis data with optimal interpolation techniques (Bonanno et al., 2019).

These results prove the added value of SPHERA for describing medium-to-severe local precipitation events, owing to several improvements compared to its global driver. These are: the finer grid spacing allowing for an enhanced physical and microphysical description of the processes, and the explicit representation of DMC, the resulting higher level of topography detailing allowing for a better representation of atmosphere-land interactions, and, possibly, the better adherence to assimilated regional observations. Previous studies reported multiple benefits related to increased grid resolutions in numerical simulations, allowing an adequate representation of local dynamical features and forcings leading to or intensifying precipitation events (Capecchi, 2021; Cerenzia et al., 2020; Klasa et al., 2018; Wahl et al., 2017; Clark et al., 2016; Cassola et al., 2015; Davolio et al., 2015; Buzzi et al., 2014). Furthermore, the results obtained are in line with those of recently produced CP hindcasts over Italy, sharing similar characteristics with SPHERA, and obtained by downscaling ERA5 with the BOLAM/MOLOCH model (Capecchi et al., 2022) or COSMO model (Reder et al., 2022; Raffa et al., 2021). In Capecchi et al. (2022), similar wet frequency biases in reproducing the 90th percentiles of annual, daily, and hourly rainfalls are detected for MOLOCH simulations at 2.5-km grid spacing, as opposed to the dry biases detected with BOLAM ran at 7 km. Further, from the analysis of two severe-precipitation events, a higher level of detail in the spatial characterization and less deviation from maximum intensities is maintained with the CP hindcast, despite underestimating the most extreme rainfall observations (by 64% in one case). As regards the hindcast obtained by downscaling ERA5 at 2.2 km over 20 European cities (Reder et al., 2022), enhancements over its driver are obtained in terms of the spatial patterns of annual precipitation and the relative annual maxima, as well as the monthly and hourly cycles of precipitation. In fact, ERA5 tends to simulate precipitation fields that are too smooth to give an appropriate representation of rainfall extremes, in agreement with the results presented here. A significant difference between SPHERA and the CP-reforecasts mentioned above is the lack of the additional assimilation of regional observations in the latter. The assimilation procedure in the production phase of the dataset is known to substantially enhance the representation of the simulated atmospheric state (Bollmeyer et al., 2015; Bonanno et al., 2019), particularly for precipitation estimates (Gustafsson et al., 2018; Clark et al., 2016; Lewis et al., 2015). In SPHERA, this is accomplished by including several conventional observations on the near-surface (SYNOP and SHIP) or upper-air (TEMP, PILOT, and AIREP) through a continuous nudging scheme, permitting a better adherence of the past-weather simulations to the observed atmospheric state. The same observations are also assimilated in the global driver ERA5. However, the much finer grid spacing at which they are ingested in SPHERA (2.2 km vs. 31 km of ERA5) substantially impacts their representativeness. Moreover, it is worth noting that neither reanalyses assimilate precipitation observations. Further and deeper investigations are needed to assess the potential added value of the assimilation of regional observations in SPHERA through a quantitative comparison with similar CP datasets.

Concerning the observational data considered as the reference state, a crucial aspect emerging is the critical dependency of the performance of SPHERA and ERA5 on the quality of rain-gauge measurements. In fact, as reminded several times, sparse pluviometric data such as those of Dewetra, despite the enhanced spatial density over the years, may be affected by various representativity issues, namely: measurement errors, spatial inhomogeneities due to lower station coverage in some areas (i.e., especially in southern Italy or in mountainous regions, where additionally slightly different sensors positioning may cause significant underestimations in the recorded data, Crespi et al., 2018), and the non-constant number of the sensors during the years in exam which increased by roughly 400% (from 1500 in 2003 to 6200 in 2017). For these reasons, the results should be interpreted with caution, as the reference state is far from ideal. A possible way to partly overcome spatial inhomogeneities issues may be to consider multiple observative datasets, especially in the form of high-resolution gridded analyses such as GRIPHO (Fantini, 2019) over Italy or ARCIS (Pavan et al., 2019) over north-central Italy. The combined use of these datasets could enhance the uncertainty quantification in describing the actual atmospheric state. Anyhow, the data interpolation required to construct these gridded datasets may be an additional source of error due to the resulting intrinsic smoothing of the rainfall fields, which constitutes a major limit, particularly for the representation of extreme precipitation.

In conclusion, to respond to the first question posed in this thesis:

"To what extent do the improvements of CP reanalysis simulations outperform coarser counterparts in representing precipitation and, specifically, severe precipitation events?"

The CP reanalysis SPHERA demonstrated a quantitative added value over ERA5 in the representation of precipitation and severe precipitation over Italy in terms of various aspects: more variable and locally-driven spatial distributions of the 95th percentile of precipitation, higher statistical skill scores, especially for the occurrence of moderate- to intense accumulations, both on a daily and hourly frequency, and more detailed, higher-intensity reaching, and better spatio-temporally localized rainfall fields in case of specific extreme precipitation events.

Note: the analyses presented in this Chapter are the subject of a scientific manuscript published on the Quarterly Journal of the Royal Meteorological Society (Giordani et al., 2023).

Chapter 4

Describing hail-prone convective environments with SPHERA and overshooting top detections

This Chapter is dedicated to the analysis of hailstorms and the associated convective hail-favoring environments described with SPHERA numerical proxies, satellite thunderstorm detections, and crowdsourced surface-based hail observations.

The inspiration for this analysis relies on recent research led by Punge et al. (2017) and further applied by Bedka et al. (2018) and Punge et al. (2021). It represents an attempt to reduce the uncertainty around hailstorm events understanding and predictability by combining multiple direct and indirect data sources. A crucial innovation introduced by the present study is the description of hail-favoring environmental conditions with a CP reanalysis (SPHERA) instead of a coarser global dataset. Indeed, higher resolutions are expected to improve the representation of the atmospheric state thanks to the removal of physical parameterization schemes for describing DMC processes, which represent a dominant source of errors in the simulations. Furthermore, the updated automatic satellite-based algorithm for detecting thunderstorms of Khlopenkov et al. (2021) brings in a further novelty, which has already demonstrated its added value over the United States (Cooney et al., 2021).

The analysis is performed over three years (2016-2018), owing to the availability of the different datasets and to the increased numerosity of ESWD hail reports in the last years, for the extended summer season (April-October). The area considered comprises the entire SPHERA spatial domain (i.e., roughly 6-19°E, 35-49°N, see Figure 2.1). This region contains one of the main European hot spots for the development of severe convective thunderstorms and, hence, hailstorms (i.e., northern Italy, Taszarek et al., 2020). Furthermore, in the same area, the greatest likelihood of a worsening of hailstorm severity in the coming years as a consequence of anthropogenic global warming is expected (Raupach et al., 2021).

The Chapter is organized as follows: Section 4.1 presents the observational data considered; in Section 4.2, the methodology developed to couple reanalysis and hailstorm observations in the form of an environmental-based filter is reported; Section 4.3 shows the resulting satellite-based hailstorm characterization, while Section 4.4 is dedicated to the investigation of the environmental conditions identified with reanalysis proxies; finally, the results are summarized and discussed in Section 4.5.

4.1 Hail observations

4.1.1 The European hail-reports database ESWD

The European Severe Weather Database (ESWD, `https://www.eswd.eu/`; Dotzek et al., 2009) constitutes one of the primary sources of severe convective storm data in Europe and represents the biggest and only multinational European hail reports archive. Maintained by the European Severe Storm Laboratory (ESSL), the ESWD provides quality-checked data collected by networks of voluntary observers, meteorological services, weather enthusiasts, and news and media reports. The transversal cross-collaboration between multiple entities at the heart of this database was necessary to overcome the previous fragmentation of severe weather databases in Europe. Thanks to technological innovations and increasing public awareness of extreme meteorological events, reports have rapidly increased in number in recent years (Groenemeijer et al., 2017). The introduction of a mobile app, the European Weather Observer (EWOB; Holzer et al., 2017), made decisive contribution in this sense, which further enhanced the engagement with the general public for reporting severe weather occurrences in real- or near real-time. The encouraging added value given by crowdsourced severe-weather reporting prompted the development of similar modules in national weather services apps of Germany (DWD WarnWetter-App - Spitzer et al., 2022) and Switzerland (MeteoSwiss-App - Barras et al., 2019). Anyhow, despite the promising ameliorations, the ESWD database still suffers from serious deficits in data representativity, namely: the spatial-inhomogeneity bias in the localization of a larger part of reports towards the most populated areas (i.e., the main urban centers), especially in central Europe where the most active voluntary observer networks are located, and the short temporal series owing to the newness of the database. However, the ESWD forms the only reliable source of direct hail data for Europe, including information for any report such as location, date, hour (with an estimate of the temporal accuracy), and maximum size of hailstones. Further, an operational quality control procedure categorizes each report with different quality

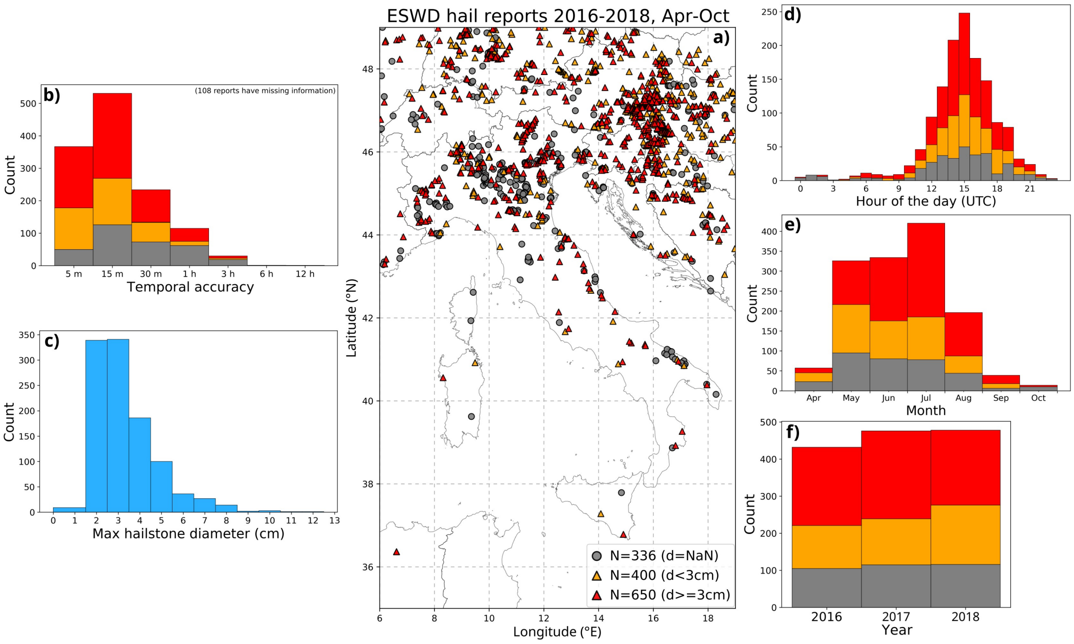

Figure 4.1: ESWD hail reports during April-October over 2016-2018. The reports are classified by distinguishing among three classes: reports with no information on hail size (in gray), small hail (maximum diameter <3 cm, in orange), and large hail (maximum diameter ⩾ 3 cm in red). a) spatial distribution, b) temporal accuracy distribution, c) maximum hailstone diameter distribution, d) number of reports per hour of the day (UTC), e) number of reports per month, and f) number of reports per year.

levels: QC0 (“as received”), QC0+ (“plausibility checked”), QC1 (“confirmed by reliable source”), and QC2 (“scientific case study”). Púčik et al. (2019) present the first ESWD-based hail climatology for 1990-2018, quantifying injuries and damages from hailstorms and noting an upward trend in the annual number of days with hail losses and their magnitude over Germany.

The set of ESWD hail reports considered in this work pertains to the extended summer season (April-October) during the years 2016-2018 over the whole SPHERA reanalysis domain and having a minimum quality level QC0+ (Figure 4.1). In total, 1386 hail reports were available, of which 1050 have information on hail size (i.e., the 76%). Hereafter we refer to small hail reports having maximum hailstone diameters <3 cm, large hail (⩾3 cm), and very large hail (⩾5 cm). They compose 38%, 62%, and 17% respectively of ESWD reports with information on hail size. Their spatial distribution (Figure 4.1a) denotes a strong inhomogeneity with a larger density of reports in northern Italy, south-western Austria, eastern Slovenia, and northern Croatia. On the other hand, the regions with the lowest hail reporting are central-southern Italy and islands, the Alpine arc, and southern Balkans (southern Croatia and Bosnia and Herzegovina). 90% of the reports have a temporal accuracy ⩽ 1h (Figure 4.1b), and they pertain to a minimum hailstone size of 2 cm (Figure 4.1c).

Their temporal distribution indicates the maximum likelihood for hail occurrence at 15 UTC (Figure 4.1d) in July (Figure 4.1e), with a similar number of reports (between 400 and 500) for each of the three years considered (Figure 4.1f).

4.1.2 Overshooting top detections

Remotely observing severe convective thunderstorms from space is possible through visible and infrared (IR) imagery from geostationary satellites. Particularly, severe thunderstorms extending throughout the troposphere are detectable in the IR as local cold spot anomalies sprouting from warmer areas. These are associated with Overshooting Tops (OT) resulting from intense updrafts protruding in the stable layer near the tropopause and penetrating the lower stratosphere, surrounded by a warmer cirrus anvil cloud confined below the last unstable level of free buoyancy (Adler et al., 1985). It is well known that thunderstorms presenting satellite OTs signatures have the potential to produce a variety of hazardous weather at the surface, such as tornadoes, heavy rainfall, large hail, and wind gusts, which are typically concentrated near OT regions (Reynolds, 1980; Brunner et al., 2007; Setvák et al., 2013; Bedka and Khlopenkov, 2016; Mikuš and Mahović, 2013; Mecikalski et al., 2021). Therefore, automatic OTs detection as a proxy to predict and investigate severe weather has become a common practice globally (de Coning et al., 2015; Hong et al., 2022). OTs can rapidly form and evolve within a thunderstorm and can exist for less than 15 minutes (even less than 5 minutes - Elliott et al., 2012) with a maximum diameter of roughly 15 km (Fujita, 1992; Brunner et al., 2007), and presenting typical temperatures between 190 and 215 K (Allen et al., 2020). The remote detection of OTs has been automated by Bedka et al. (2010) and has been considered to characterize OTs climatological distribution in North America (Bedka et al., 2010), Europe (Bedka, 2011), and Australia (Bedka et al., 2018). This OT detection algorithm relies on the comparison between clusters of cold pixels likely related to strong updrafts with warmer pixels consistent with the temperature in the anvil of the thunderstorm, as detected with IR satellite scans. A large temperature difference (> 6K) helps at separating actually occurred OTs from other non-convective clouds (e.g., cirrus) and is indicative of updraft penetration through the anvil of at least 1 km (Griffin et al., 2016). Recently, the automatic OT detection algorithm has been substantially improved by moving to a probabilistic approach (Bedka and Khlopenkov, 2016; Khlopenkov et al., 2021) conversely to the binary yes/no decisions based on predefined fixed thresholds in the identification of an OT. The statistical combination of tropopause-relative IR brightness temperature, the prominence of a candidate OT relative to the surrounding anvil, and the spatial uniformity of the area covered by the anvil delivers a 3-km gridded probabilistic

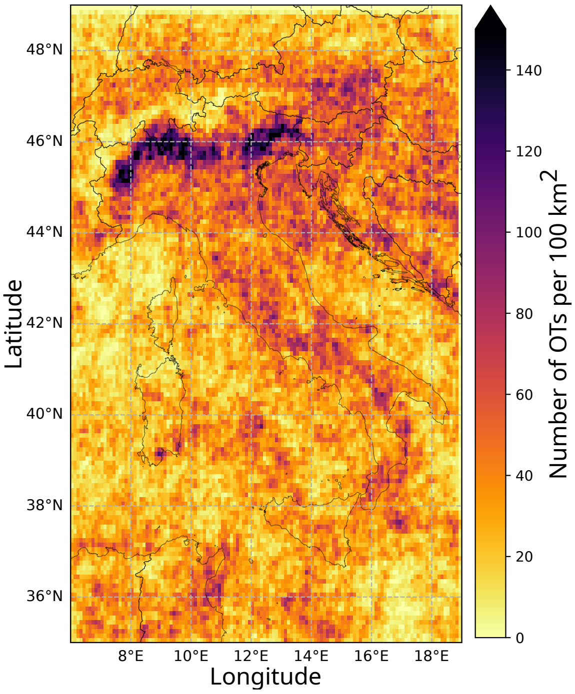

Figure 4.2: Number of overshooting tops detected per grid cell (on a 10-km regular grid) during April-October in 2016-2018.

OT estimate at 15-min resolution. The validation of this methodology (Khlopenkov et al., 2021; Cooney et al., 2021) revealed important ameliorations over the original OT-detection design of Bedka et al. (2010), especially in the form of a consistent reduction in the number of false OT detections.

Several studies reported the linking between large hail at the surface and OT intensity (Bedka, 2011; Punge et al., 2014; Proud, 2015; Jurković et al., 2015; Punge et al., 2017; Bedka et al., 2018; Punge et al., 2021; Scarino et al., 2022), indicating the appropriateness of OT observations as a proxy for hail occurrence. In this context, here are considered the cloud-top scans obtained through IR imagery from geostationary Meteosat Second Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) (Schmetz et al., 2002) between 2016 to 2018 in the extended summer months of April to October at a continuous temporal resolution of 15 minutes for developing a proxy for hail over Italy and the surrounding countries and seas. Only OTs detected with the Khlopenkov et al. (2021) algorithm having a probability >50% are considered, similarly to Punge et al. (2021). This constraint is derived by comparing OT detections with radar echo tops (Cooney et al., 2021) and demonstrated enhanced reliability as it indicates colder and more prominent anvil-relative tops. The resulting spatial distribution of the 633,141 OTs detected is reported over a 10-km regular grid in Figure 4.2. A generally higher number of OTs over land is observed, especially around the Alps, Apennines, and Dinaric Alps mountainous ranges.

4.2 Hail filter design

Although OT-producing thunderstorms can generate severe weather, many thunderstorms presenting OT signatures do not produce phenomena such as severe winds, heavy rainfall, or hail. In fact, intense updrafts reaching the tropopause and penetrating the stratosphere may be associated with convective environments not necessarily supportive of severe weather production. Although recent studies suggest links between convective storm severity and specific characteristics of the OTs, such as their area extension (Marion et al., 2019) or the temperature gradient between the OT top and the tropopause (Khlopenkov et al., 2021), satellite detections alone are not a sufficient proxy for investigating hail occurrence. Therefore, a way to discriminate between hail-producing and non-hail-producing OTs is necessary. The combination of satellite OT detections with storm environmental conditions offers a possible approach with the potential to analyze better the conditions leading to hail occurrence. In the methodology developed by Punge et al. (2017) for Europe, which has been similarly applied also in Australia (Bedka et al., 2018) and South Africa (Punge et al., 2021), the environmental conditions are described through reanalysis simulations. While global reanalyses such as ERA-Interim and, more recently, ERA5 have been considered in the former studies, the more detailed CP SPHERA reanalysis is used in this work to filter OTs not associated with hail. In fact, given the benefits shown by SPHERA over ERA5 in the description of severe rainfalls (Chapter 3), a better representation of hail-producing severe convective environments is expected compared to coarser global datasets. Section 4.2.1 reports the SPHERA parameters selected to describe environmental conditions, and the construction of the OT-filter is presented in Section 4.2.2.

4.2.1 SPHERA reanalysis data

For selecting the numerical proxies to best describe convective environments prone to hail development, we relied on common knowledge from previous literature (Jelić et al., 2020; Kunz, 2007; Kunz et al., 2020; Allen et al., 2020; Punge et al., 2017; Bedka et al., 2018; Prein and Holland, 2018; Punge et al., 2021). The selected parameters originate in operational forecasting and rely on direct or indirect relationships between hail observations and proximal atmospheric soundings. Their involvement can help in discriminating convective situations in which hail formation is improbable and are employed to filter out OT detections related to such occasions. They represent dynamic and thermodynamic characteristics of the atmosphere describing key ingredients for hail formation in a thunderstorm: atmospheric instability, storm organization, and freezing level altitude.

Atmospheric instability and the relative updraft strength of a thunderstorm are particularly relevant as strong updrafts are necessary for hail growth. Atmospheric instability can be quantified with numerous indices. We selected three parameters because only one is generally insufficient to comprehensively express the likelihood of thunderstorm development. This holds especially in Europe where thunderstorm environmental conditions may substantially differ from those found in the United States (Brooks, 2009; Taszarek et al., 2020, 2021), where the majority of the indices were developed. The chosen instability predictors represent some of the most used parameters for the purpose and show the highest skills for the prediction of severe thunderstorms (Kunz, 2007), and are: CAPE (Convective Available Potential Energy) in its Most Unstable (MU) formulation (to account for elevated convection and shallow boundary layers), K index and SLI (Surface Lifted Index). They are defined as:

$$CAPE = g \int_{z_i}^{LZB} \frac{T_{v_{parc}} - T_{v_{envir}}}{T_{v_{envir}}} dz$$

$$K = (T_{850hPa} - T_{500hPa}) + Td_{850hPa} - (T_{700hPa} - Td_{700hPa})$$

$$SLI = T_{500hPa_{envir}} - T_{500hPa_{parc}}$$

where g is the gravitational acceleration, z_i is the altitude in the lowest 300 hPa where the equivalent potential temperature is at its maximum (i.e., most unstable conditions), LZB is the level of zero buoyancy (or equilibrium level) where the virtual temperature of the parcel $T_{v_{parc}}$ equals the virtual temperature of the surrounding environment $T_{v_{envir}}$, T_{xxhPa} and Td_{xxhPa} are respectively the temperature and the dew-point temperature at the isobaric level xx hPa. CAPE represents the integrated amount of work over the vertical air column exerted by the upward buoyancy force over the air parcel, K index is a combination of the lapse rate, the low-level moisture content, and the moist layer depth, while SLI evaluates the temperature difference between the environment at 500 hPa and a parcel lifted dry adiabatically from the surface to the lifting condensation level and pseudo-adiabatically to 500 hPa.

Along with atmospheric instability, thunderstorms with an enhanced convective organization are more likely to produce severe weather such as hail. For this reason, in the filter, we included the deep layer shear DLS, a parameter typically used for investigating hail-favoring environments (e.g., Wellmann et al., 2020; Trapp et al., 2007). DLS represents the horizontal change in wind speed in the vertical between the surface (v_{10m}) and approximately 6 km above the ground (v_{500hPa}), which reads:

$$|\overrightarrow{DLS}| = |\overrightarrow{v_{500hPa}} - \overrightarrow{v_{10m}}|$$

Finally, a determinant factor influencing specifically hail, among other severe deep-convective phenomena, is the thermal characteristic of the vertical atmospheric profile that directly affects the phase state of the water content suspended in the air column. This critical factor is quantified through the freezing level height H_0, defined as the altitude of the 0 °C isotherm above mean sea level. In fact, the atmospheric thermal characteristics in different regions of the world imply thermodynamical differences in thunderstorms evolution (Ludlam, 1980), and hence explain, for example, the low hail frequency registered at latitudes near the tropics where surface atmospheric layers are generally warmer and the tropopause altitude is higher (Prein and Holland, 2018). This indicates that storms with high H_0 are less likely to produce hail on the surface, mainly because hailstones produced at higher altitudes could melt more easily before reaching the ground (Dessens et al., 2015). Indeed, the amount of moisture available below the freezing level height potentially influences the hydrometeor density in the thunderstorm and hence the hail growth rate (Allen et al., 2015; Johnson and Sugden, 2014). On the other hand, too low H_0 may limit the amount of supercooled water in the updraft necessary for hail growth (Prein and Holland, 2018). Recent findings also determined that anthropogenic-induced climate change could cause a general global increase in freezing level height, low-level moisture, and convective instability. This could influence the likelihood of hailstorms with a shift towards enhanced severity owing to increased hail melt, causing a larger average size of surviving hailstones, hence producing not unambiguous prediction in the change of their frequency due to the complex combination of these factors (Raupach et al., 2021).

While CAPE, SLI, and H_0 are direct outputs of SPHERA, DLS and K are computed from temperature and wind profile data simulated by SPHERA at the relative pressure levels. Every parameter is available at hourly frequency at the native high horizontal resolution of ∼2.2 km. However, the difficulties in simulating local and rapidly-evolving deep convective processes with exact precision, characterized by low predictability at this spatial resolution, could pose major limits in data representativity (see Section 3.2). For this reason, SPHERA fields are upscaled to a common coarser grid of 10 km to avoid possibly "noisy" reanalysis estimates and data representativity issues, similar to the precipitation analysis presented in Chapter 3.

4.2.2 OT-hail filter by conditions from SPHERA

The filter presented here intends to remove all OT detections for which hail presence is very unlikely based on the surrounding environmental conditions described by SPHERA. The methodology is mostly inspired by the work of (Punge et al., 2017), which has been applied similarly also in (Bedka et al., 2018) and (Punge et al., 2021).

Around each OT detection, SPHERA parameters are considered within a spatial

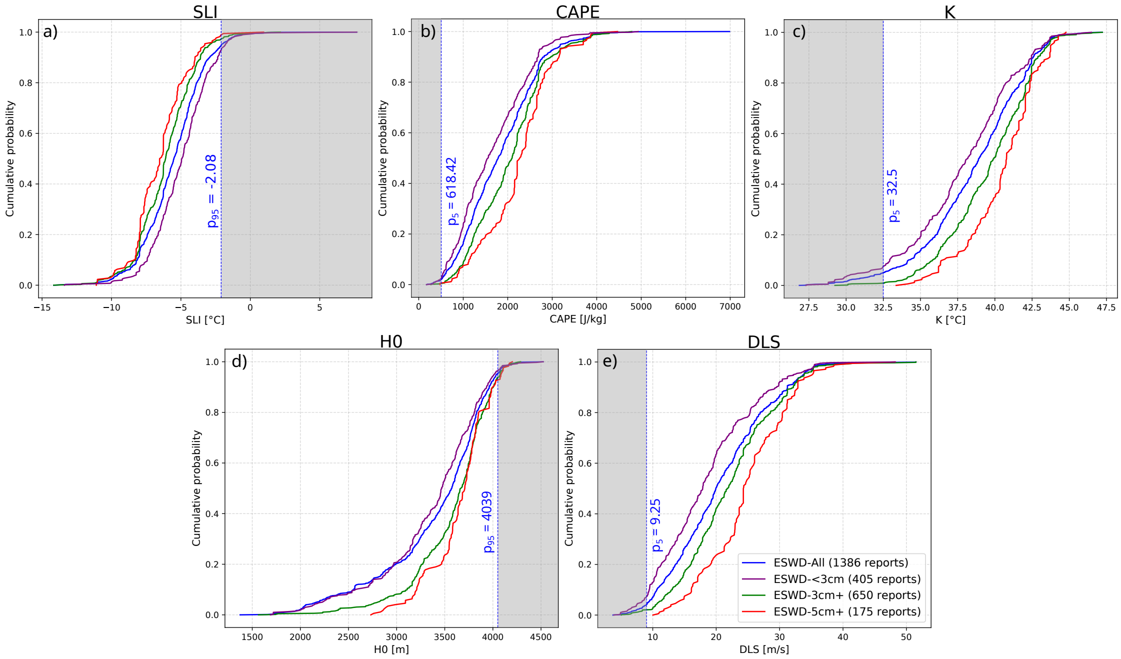

Figure 4.3: Cumulative density functions of the five parameters selected from SPHERA in the presence of ESWD hail reports in 2016-2018 during the extended summer season (April-October). Hail reports are divided into different classes: all reports (blue lines), small hail (<3 cm, purple lines), large hail (⩾3 cm, green lines), and very large hail (⩾5 cm, red lines). The blue dashed vertical lines indicate the thresholds selected for defining the OT hail filter reported in Table 4.1. The shadowed portion of the distributions indicates the range of values of the parameters where the filter is effective. a) SLI, b) CAPE, c) K, d) H_0, and e) DLS.

window of 0.63° x 0.63° (i.e., roughly 70 km x 70 km) over the three hours preceding the OT (including the hour at which the OT is issued). Indeed, the extremely localized and rapidly-evolving nature of hailstorms calls for some spatial relaxation in the matching criterion with numerical simulations. This is done to limit double-penalty issues owing to the incapacity of the models to predict the dynamical evolution of DMC processes with absolute precision (Ebert, 2008). Additionally, to take into account the pre-convective conditions included in the computation of instability parameters, a certain temporal window before the OT event is considered. Within this spatio-temporal neighborhood, the maximum (for CAPE, K, and DLS) or minimum (for SLI and H_0) are extracted.

To define the parameter thresholds for filtering the OTs, the same spatio-temporal analysis is applied to SPHERA parameters distributions in the presence of ESWD hail reports. In this case, we can be confident that hail actually occurred; further, we can also account for outliers in the ESWD-based parameters distributions due both to uncertainties in the timing of the reports and to the inaccuracies in the simulations. The thresholds are then defined as percentiles p_{th} of the distributions.

Variable	Threshold	OTs filtered	Fraction
SLI	$< p_{95} = -2.08$ °C	60,675	9.6%
CAPE	$> p_5 = 618.42$ J/kg	47,923	7.6%
K	$> p_5 = 32.5$ °C	44,639	7.1%
H_0	$< p_{95} = 4039$ m	45,394	7.2%
DLS	$> p_5 = 9.25$ m/s	38,744	6.1%
Full filter	All those above	163,510	25.9%

Table 4.1: Variables and thresholds used in the OT filter and relative number and fraction of OTs filtered.

After numerous tests, the 5th percentiles (for CAPE, K, and DLS) and the 95th percentiles (for SLI and H_0) of the distributions were selected. The ESWD-based cumulative density functions (CDFs) of every parameter are reported in Figure 4.3. To investigate the relationship between each parameter and hailstorm severity, the CDFs are reported for the distribution of the whole ESWD hail reports set and distinguishing among small hail, large hail, and very large hail. A general shift of the predictors towards severe-convective environments presenting higher instability (i.e., larger values of CAPE and K and lower values of SLI), enhanced organization (i.e., larger DLS) is detected for increasing hail severity, as well as generally higher freezing level heights H_0. This indicates the ability of the numerical proxies selected to distinguish between different convective environments prone to hail formation and suggest their appropriateness for the OT-filtering purpose. The shaded areas in Figure 4.3 indicate the tail of the CDFs (corresponding to the 5th or 95th percentile portions) where the filter is active.

The numbers and the fractions of OTs filtered by the full filter (i.e., applying the five conditions together) and by every single parameter are listed in Table 4.1. Singular parameter contributions to the filter vary from 6.1 to 9.6%. Since the impacts of different variables overlap with each other, the fraction of removed OTs with the full filter is lower than the sum of the singular filters and reaches 25.9%.

Figure 4.4 shows the spatially-distributed filtered fractions of OTs on a 10-km grid for the single-parameters filters (panels a-e) and for full filter (panel f). Instability parameters (CAPE, K, and SLI – panels a,b,c, respectively) filter mainly over the sea (especially in the southern and western Mediterranean) and the Alpine arc. The largest contribution over the sea is given by K, while over land, CAPE and SLI are more effective. This behavior may be linked to the formulation of K that, conversely to CAPE and SLI, explicitly considers the water vapor content in the atmospheric column (through Td), which weighs more over the sea. The identified values as thresholds for instability parameters responsible for these spatial distributions agree with previous research in this direction. For SLI, the imposed

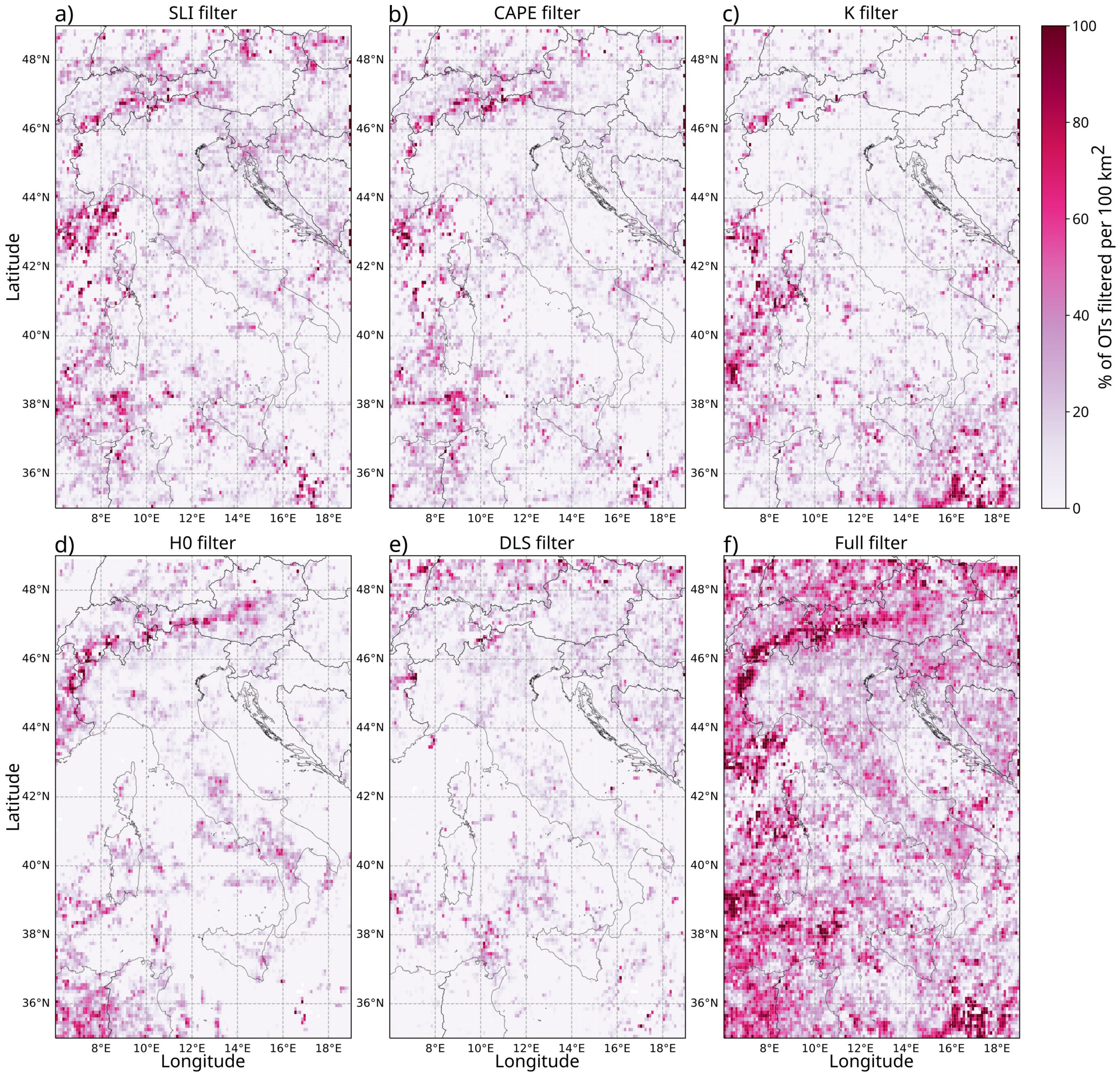

Figure 4.4: Percentage of filtered OTs per grid cell (on a 10-km regular grid) during April-October in 2016-2018. a) SLI, b) CAPE, c) K, d) H_0, e) DLS, and f) full filter.

threshold of -2.08 °C is in almost perfect agreement with the optimal threshold of -2.07 °C identified by Kunz (2007) for distinguishing between hail- from non-hail days in south-western Germany. Hence, the identified threshold can be considered an appropriate choice as an upper bound of SLI values distribution in light also of the mean value of -3.8 °C found in the presence of hailstorm events by Kunz et al. (2020). The 5th-percentile threshold for K (32.5 °C) well resembles the order of magnitude identified by Kunz (2007) for the distinction between widespread and isolated hail events (31.1 °C) or for separating between hail from non-hail days (38.9 °C, using a modified version of K where T and Td at 850hPa are substituted with an average from the surface to that pressure level). Torralba et al. (2022) also found a comparable median value of 30 °C for the K index for the occurrence of hailstorm events as described by ERA5 reanalysis. Regarding the most unstable CAPE (minimum conditions of 618.42 J/kg), Kunz (2007) found a higher optimal threshold of 1474 J/kg distinguishing hail- from non-hail days. Anyhow, the analysis from observational proximity soundings over central Europe provided a median value for most unstable CAPE in the case of small hail (<2 cm) of ∼250 J/kg, which increased to ∼700 J/kg when considering ESWD reports for larger hail of 2-5 cm diameter coupled with radiosounding data (Púčik et al., 2015). Additionally, using ERA5, Torralba et al. (2022) found CAPE values >500 J/kg to represent favorable atmospheric conditions for hail development. Further, given the less strict criteria used in previous similar works (i.e., thresholds of 0 or 100 J/kg in Punge et al., 2017, 2021, respectively), the identified criterion for CAPE is justified.

H_0 (Figure 4.4d) is filtering most OTs over lower latitudes (with fractions >60% in Tunisia and Algeria) and high-elevation terrains, especially over the whole Alpine arc, where almost 100% of OTs are removed. This enhanced removal is a direct consequence of the threshold value of 4039 m used for filtering, which, in the Alpine region, may be very close to the topographical surface (especially since the imposed topography boundary condition of SPHERA reaches elevations as high as 3950 m, see Figure 2.1). However, as seen by the OTs spatial distribution (Figure 4.2), the crest of the Alpine arc is the least active region of the domain in terms of OT formation. This is a consequence of the difficulties for deep-organized convective systems to develop in extremely complex terrains, in agreement with a recent climatology of lightning flashes and convective initiation over the Alpine area (Manzato et al., 2022). Furthermore, Jelić et al. (2020) found no evidence of hail observations over the north-eastern Adriatic region when considering ERA5 environments with H_0 exceeding 4000 m. Hence, it is believed that the limitation imposed by choice of the H_0 threshold is not detrimental to the analysis presented here. A proposal for a more sophisticated H_0-filtering could be topography-dependent, which, however, would require a deeper investigation beyond the scope of the present analysis.

DLS (Figure 4.4e) is the least filtering parameter, and its spatial distribution is mostly homogeneous, with enhanced filtering in the northern part of the domain (i.e., southern Germany and northern Austria). The 5th-percentile threshold of 9.25 m/s well represents the necessity of minimum storm organization conditions for hailstorm development. Indeed, the found value is close to the magnitude identified for the transition from single-cellular to multi-cellular convection (i.e., 10 m/s, Markowski and Richardson, 2011), which is also the minimum condition for the development of severe thunderstorms as defined in Taszarek et al. (2021). Further, the identified threshold well compares with the median value of DLS of ∼13 m/s in case of small hail (< 2cm) found from proximity soundings over central Europe, which increases to 15 m/s in case of larger hail of 2-5 cm diameter (Púčik et al., 2015), or with the average of 12.5 m/s found around hailstorm events by Kunz et al. (2020).

The combination of the five conditions (Figure 4.4f) leads to a spatial OT filter-

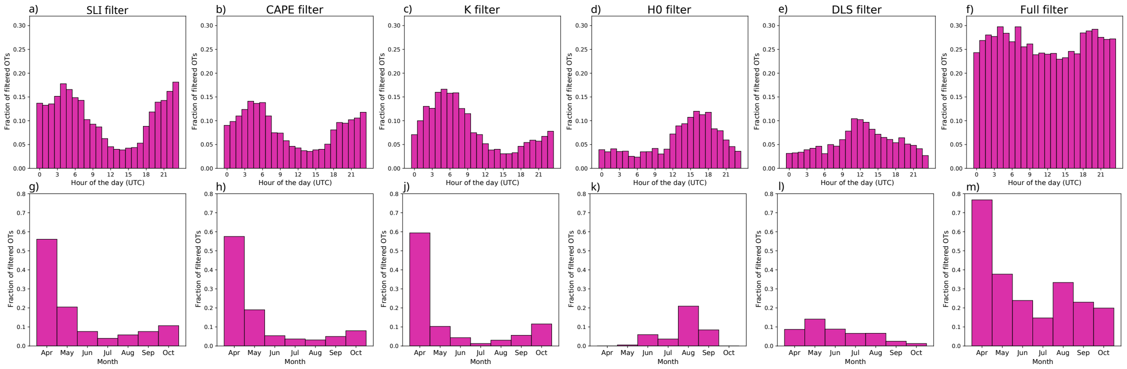

Figure 4.5: Percentage of filtered OTs aggregated over the spatial domain per hour of the day (top row) and per month (bottom row) in the period 2016-2018 considering singular parameters filters: a) & g) SLI, b) & h) CAPE, c) & j) K, d) & k) H_0, e) & l) DLS, and f) & m) the full filter.

ing which is maximum (∼100%) over the inner Alpine region and substantially high (∼60%) but locally variable over the western and southern Mediterranean sea and in the northern part of the domain (southern Germany and Austria). In contrast, the regions with the lowest OT removal (<20%) are the whole Po valley, the northern Adriatic and Thyrrenean seas, and the associated Italian and Croatian coastlines.

Figure 4.5 shows the fractions of filtered OTs by every parameter used in the filter, depending on the hour of the day (upper row) and the month of the year (lower row), and aggregated over the whole spatial domain. It should be kept in mind that the local time zone over the considered region, the Central European Summer Time (CEST), is two hours ahead of the UTC time zone. Instability parameters (CAPE, SLI, and K – panels a,b,c,g,h,j) filter mainly during the night and early morning (exceeding the 15% of OT removal around 3-6 UTC), in the Spring months of April (where more than the 50% are removed) and May, and in October. This reflects the lower likelihood of deep convection establishment at these times of the day and year. Conversely, their contribution to the filter is minimum when the increased heating of the lower troposphere enhances the potential for convective activity and reduces convective inhibition (Markowski and Richardson, 2011), i.e., in the central hours of the day (with less than 5% removal around 14-16 UTC) during JJA (particularly in July). No significant differences between the three parameters are detected. The only notable exception is the lower impact that K has in the late-evening/early-night hours (roughly around 20-23 UTC with slightly more than 5% OT removal), which may be explained again by the enhanced thermal inertia owing to atmospheric moisture explicit inclusion in the parameter leading to high values of K maintained in the hours after sunset.

H_0 (Figure 4.5d-k) contribution to the filter is roughly opposite to that of instability. The largest removal is found in the afternoon (∼12% at 16-18 UTC) and in late summer, especially in August (exceeding 20%), and to some extent in September. This seasonal variation may be linked to the warming of the lower troposphere, which reaches its peak in this region in August as a consequence of the annual cycle of solar insolation. Hence, the general high temperatures of the atmospheric profile cause a shifting of the freezing level height towards higher altitudes, for which it is more likely to exceed the threshold of 4039 m, causing an increased OT removal at this time of the year. On the other hand, the daily cycle in H_0-filter removal can not be generally related to the diurnal cycle of low-tropospheric heating. In fact, at altitudes of ∼4 km a.s.l., temperature changes are mainly driven by horizontal advective processes rather than vertical sensible heat fluxes. In any case, the largest fraction of H_0-driven OT removal is found over the Alpine crest (Figure 4.5d), where the atmospheric boundary layer could extend over 4 km altitude a.s.l., despite being very shallow, implying a direct diurnal impact on the H_0 variation.

The DLS filter (Figure 4.5e-l) also shows diurnal and seasonal variations, with higher OT removal rates around noon (∼12%) in May (∼15%), while the minimum is detected during the night (∼3% at 00 UTC) and in October when the filter is almost inefficient. DLS generally shows less abrupt increases in the daily and monthly variability compared to the other parameters and more gradual trends. The diurnal impact of the DLS filter is also roughly opposite to that of instability parameters. Similarly to H_0, its variation may not be directly related to low-layers diurnal heating: while the wind field close to the surface ($\overrightarrow{v_{10m}}$) could be affected by thermally-driven daily variations, its counterpart at ∼6 km a.s.l. ($\overrightarrow{v_{500hPa}}$) is much less influenced by daily variability. For this reason, a deeper investigation confined to the Alpine region is planned to shed more light on the diurnal behavior of H_0 and DLS filters.

The resulting full filter on the daily term (Figure 4.5f) leads to generally higher but less variable OT removal rates as an effect of the combination of the parameters, with a maximum removal reaching almost 30% in the morning and late afternoon (4-7 and 18-20 UTC respectively) and a minimum of ∼23% at 14 UTC. On the annual term (Figure 4.5m), the single parameters combination sums almost linearly due to their separate influence over different times of the year. This causes the highest OT removal in Spring (April with almost 80% and May almost 40%) followed by August with ∼35%. The highest likelihood for hail-favoring conditions standing to the filter is met in July and June, where minimum OTs removal of ∼13% and ∼22% are issued, respectively. This result is in good accordance with the ESWD observed distributions of hail reports over the years considered (Figure 4.1e), as well as with the 28-year ESWD-based hail climatology presented by Púčik et al. (2019) or with the radar-based hailstorm frequency over France and Germany estimated by Kunz et al. (2020).

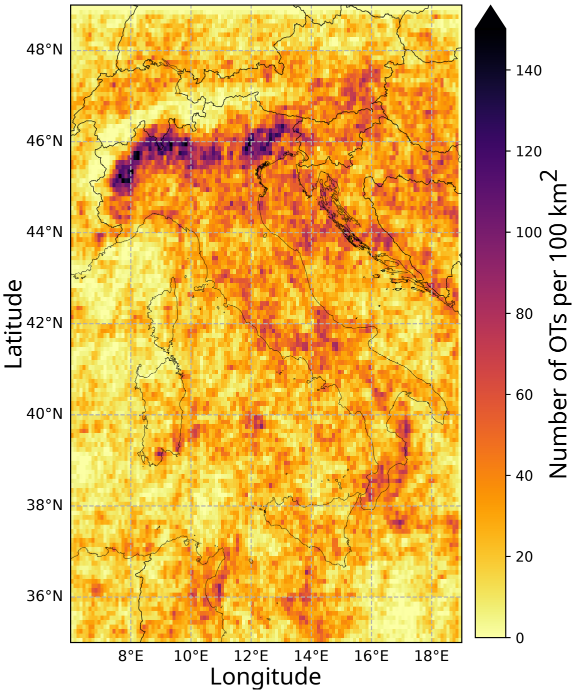

Figure 4.6: As Figure 4.2, but for OTs retained after the hail-specific filter.

4.3 The resulting hail proxy

This section focuses on the spatio-temporal characterization of the hail proxy obtained by the retained OTs after the environmental filtering and its comparison with ESWD hail reports.

4.3.1 Spatio-temporal characterization

Figure 4.6 shows the spatial distribution of the 468,890 OTs retained after hail-specific filtering over the three extended warm seasons considered (April-October, 2016-2018). Compared to the original distribution (Figure 4.2), the lower amount of OTs over the Alpine arc is evident, consequently to the highest OT removal detected (Figure 4.4f). Fewer OTs linked to hailstorm conditions are also detected over land at lower latitudes (Algeria and Tunisia), over the Mediterranean sea, throughout the Apennines, and in north-eastern continental areas (Austria, Slovenia, Croatia, and Bosnia). At the same time, the main hotspot for OT activity along the southern pre-Alpine region is well preserved after filtering. The resulting contrast between the minimum found in the inner-Alpine region and the increased hailing OT frequency in pre-Alpine areas is more pronounced than before filtering, in agreement with Punge et al. (2017), earliest hail researchers (Marbach, 1835), and recent radar-based hail climatology (Nisi et al., 2020).

The three years considered show significantly different hail-specific OT activity during the extended summer season (Figure 4.7), indicating a marked inter-annual

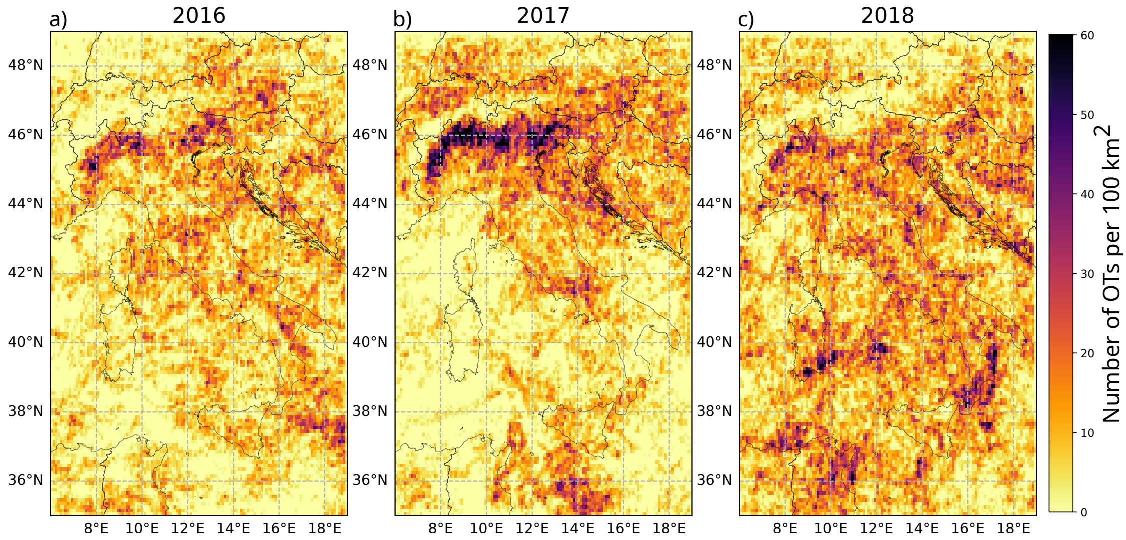

Figure 4.7: As Figure 4.6, but separating among a) 2016, b) 2017, and c) 2018.

variability. 2016 (panel a) shows the lowest annual number of hailing OTs (126,599) that are heterogeneously distributed over the region, while 2017 (panel b), with 143,469 detections, is characterized by a decisive increase over land areas, especially in the southern pre-Alpine region and northern Adriatic sea. Finally, 2018 reports the largest number of detected hail-prone OTs (198,823), most of which are linked to convection activity developed over the Mediterranean sea. In this context, it is important to mention that single events may significantly influence the overall annual OT spatial distribution. For example, it is the case of the Vaia storm, a high-impact event characterized by widespread convective activity producing extreme precipitation accumulation (up to 850 mm in three days), wind gusts exceeding 50 m/s, which damaged 41,000 ha of Alpine forest, and caused several landslides and floods (Giovannini et al., 2021). On the third day of this event, the 29th of October 2018, a total of 12,754 OTs were detected (i.e., almost 5% of the extended warm season of the year), the majority of which (6,735) over the sea.

The intra-annual variations in OT activity are investigated on a monthly basis, both in terms of the geographical distribution of hail-specific detections (Figure 4.8), as well through histograms of the spatial aggregation for the original, hail-filtered, and removed OT sets separately over land and sea areas (Figure 4.9). Very few OTs in Spring are detected, increasing from April to May, when the cool temperatures of the land and sea surface coupled with the limited availability of low-level moisture do not permit the establishment of significant convective motions. The number of OTs rapidly increases in June (characterized by the absolute maximum number of OTs) and July over continental areas, with a well-defined peak detected around the Alpine region. Starting from August and extending to September, a significant reduction in hail-filtered OT rate over land is evident (Figure 4.8e-f and Figure 4.9a), coupled with a gradual increase in thunderstorm development over the warm waters

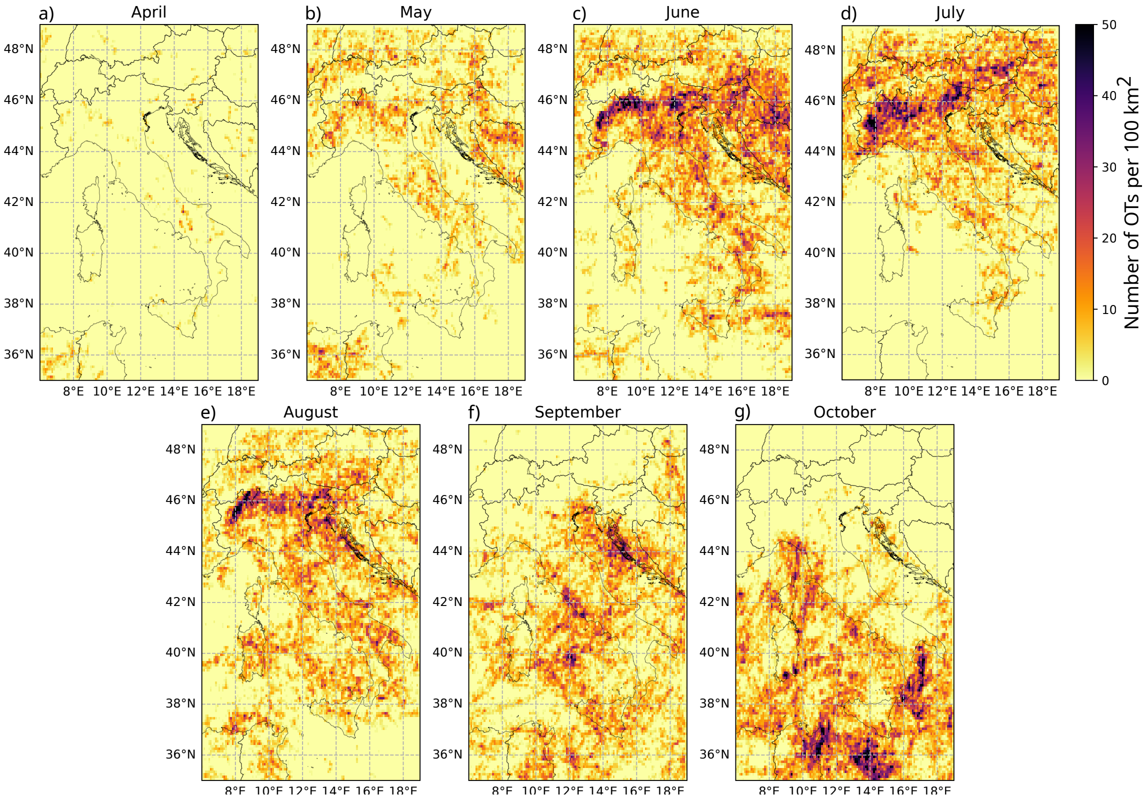

Figure 4.8: As Figure 4.6, but separating among a) April, b) May, c) June, d) July, e) August, f) September, and f) October.

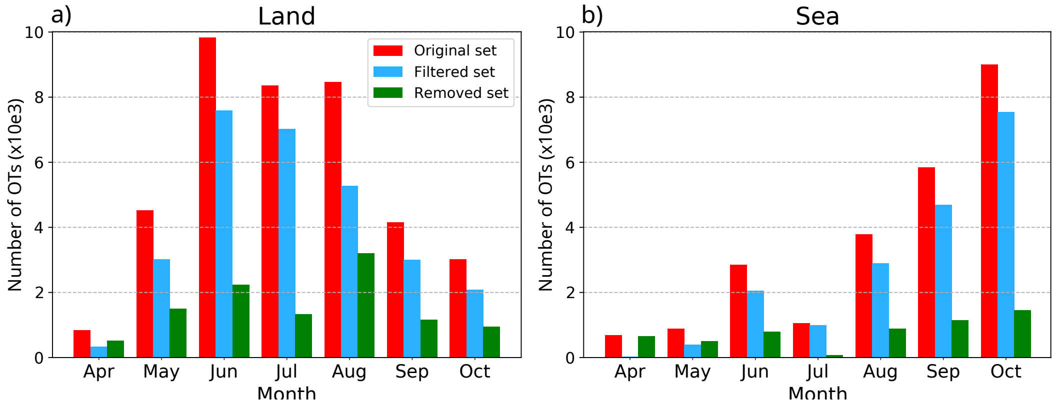

Figure 4.9: Histograms of the number of OTs detected per month separately over land (panel a) and sea (panel b) during 2016-2018 for the original set (in red), the hail-specific set (in light blue), and the removed detections after filtering (in green).

of the Tyrrhenian and Adriatic seas (Figure 4.9b). Finally, in October (Figure 4.8g) a further shift of convective thunderstorm activity towards lower latitudes of the southern Mediterranean sea is detected. This is linked to the increased cooling of the continental surface and the growing likelihood of mid-latitude cyclone formation in this region resulting from the maintenance of warm sea surface temperatures.

Hail-specific OTs are generally more frequent over land during daytime (from 8:00 to 19:45 UTC, Figure 4.10a) and over the sea during nighttime (from 20:00

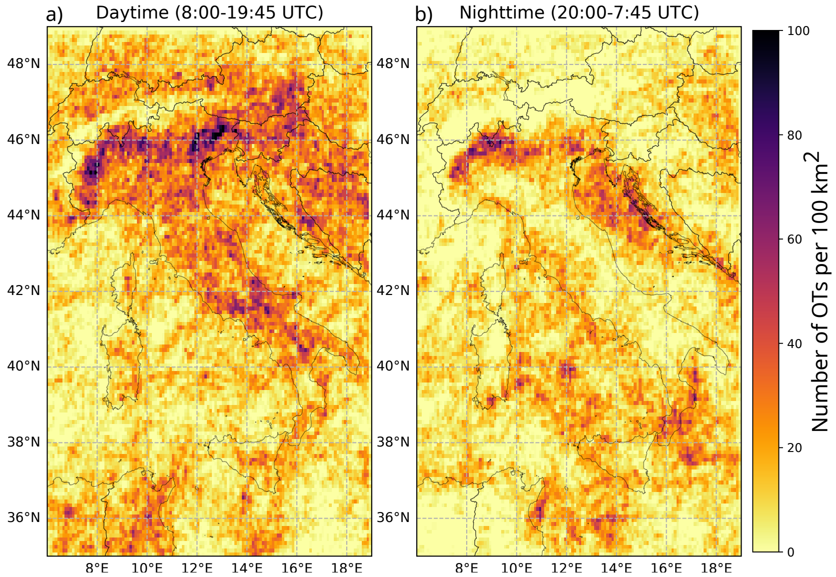

Figure 4.10: As Figure 4.6, but separating between a) Daytime (i.e., 8:00-19:45 UTC) and b) Nighttime (i.e., 20:00-7:45 UTC).

to 7:45 UTC, Figure 4.10b). The daytime bias is maximum over southern pre-Alpine areas and significantly pronounced over high-elevation terrains, especially in the eastern continental part of the domain (Austria, Slovenia, and the Balkans) and over the central-southern Italian peninsula. During the nighttime, the north Adriatic sea is an evident hotspot for OTs formation, with the maximum displaced along the western Croatian coast. This may be linked to the combination of the north-eastern mountains, supporting convective development, with local near-surface wind convergence, causing the formation and organization of convective cells over coastal areas later in the day (Jelić et al., 2020; Mikuš et al., 2012). A further relevant hotspot of nighttime OT activity is detected over the western pre-Alpine region along the Italian-Swiss border. This is linked to late-evening thunderstorm formation over the foothills, most likely imputable to katabatic winds interacting with thermally-driven Alpine Pumping circulation (Bica et al., 2007). This configuration produces local flow convergences, enhances vertical wind shear and orographic lifting, and ultimately promotes convection initiation over the region (Nisi et al., 2020).

Figure 4.11 shows the diurnal cycle of OT activity separating between land and sea areas. Over land, very infrequent OT detections are revealed during the night and early morning, with a rapid increase starting from 10 UTC and peaking at 15 UTC (i.e., 5 pm local time), as a consequence of the maximum diurnal heating of the near-surface troposphere and increased likelihood for atmospheric instability conditions. During the afternoon, a slightly more gradual decrease is detected. Over sea surface, OTs are more likely to form over the night and early morning (from 23 UTC to 9 UTC) compared to land, with a local maximum at 3 UTC. This is most likely linked to the north-eastern Adriatic hotspot of nocturnal thunderstorm gen-

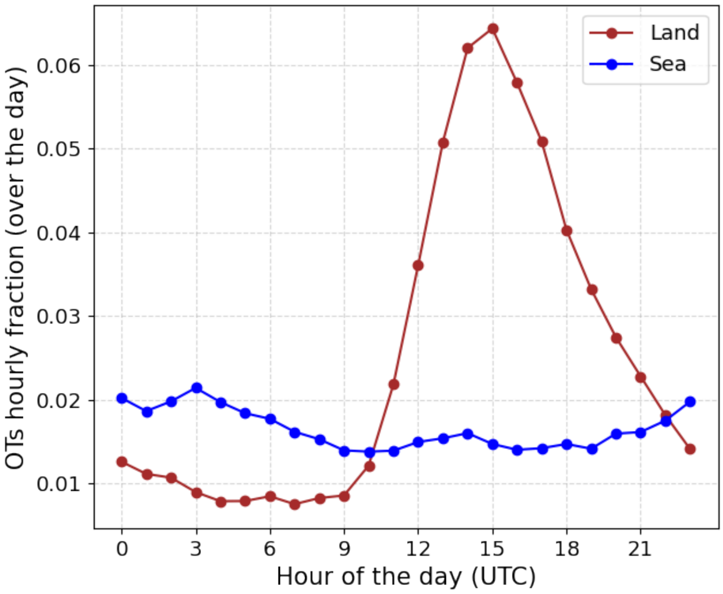

Figure 4.11: Hourly fraction of overshooting top detected after hail-specific filtering over 2016-2018, separating land (in brown) and sea surface (in blue) and aggregating over the whole spatial domain.

eration detected in Figure 4.10b. Then, a gradual decrease in marine OT activity is detected up to 12 UTC, when OT frequency slightly fluctuates during the afternoon and increases during the evening. The enhanced number of OTs over water during the last part of the day coincides with a decrease in land OTs. This pattern may be linked to a combination of several factors, such as: the advection towards water bodies of thunderstorms initially developed over land, the influence of land/sea breezes circulations responsible for coastal storm initiation, and the inertia exerted by warm waters and low-level moisture in late Summer-Fall to the maintenance of storms linked to mid-latitude cyclones.

These findings are in good agreement with the spatio-temporal distribution of the previous European OT characterization (over 2004-2009) of Bedka (2011).

4.3.2 Comparison of hail-specific OTs with ESWD reports

The aim of the comparison between OT detections and ESWD hail reports is to evaluate how often the formers are detected close to the spatio-temporal location of observed confirmed hailstorms. This ultimately gives a quantitative indication of the appropriateness of the reanalysis-OT-based hail proxy developed, which, if sufficiently according to reports, may increase forecaster confidence in predicting hailstorm occurrence, potentially on an operational basis. Unfortunately, a complete assessment of the method performance is not possible due to the limitations of the reports database, as previously mentioned in Section 4.1.1, which does not describe

	All reports		**QC1> reports**		**TA$\leqslant$ 1h reports**		**QC1> & TA$\leqslant$ 1h**	
	Filt OT	**Orig OT**	**Filt OT**	**Orig OT**	**Filt OT**	**Orig OT**	**Filt OT**	**Orig OT**
Number of ESWD rep.	1386		1059		1246		1009	
Hit ESWD rep.	871 62.8%	969 69.9%	677 63.9%	749 70.7%	807 64.8%	890 71.4%	658 65.2%	727 72.1%
% loss (Filt on Orig OT)	-7.1%		-6.8%		-6.6%		-6.9%	
Number of OTs	283,039	391,740	283,039	391,740	283,039	391,740	283,039	391,740
OTs hitting ESWD rep.	7,928 2.80%	9,356 2.30%	6,552 2.31%	7,578 1.93%	7,379 2.61%	8,604 2.20%	6,404 2.24%	7,384 1.89%
% gain (Filt on Orig OT)	+0.41%		+0.38%		+0.41%		+0.37%	

Table 4.2: Comparison between overshooting top detections (both for the original "Orig OT" and the hail-filtered "Filt OT" datasets, only over land) and four different sets of ESWD hail reports: all reports, only reports with quality level QC1 or superior (QC1>), only reports with TA$\leqslant$ 1h, and the combination of QC1> and TA$\leqslant$ 1h. For each comparison are reported the fraction of ESWD reports matching with OT detections (Hit ESWD rep. row), the difference in the ESWD hit rate between Filt OT and Orig OT (% loss row), the fraction of OTs hitting at least one ESWD report (OTs hitting ESWD row), and the difference in the OT hit rate between Filt OT and Orig OT (% gain row).

the real state in a sufficiently comprehensive way. In fact, while confirmed reports can give us precious information on the spatio-temporal occurrence of actually occurred hailstorms, the lack of reports does not necessarily imply the non-occurrence of events, as the reason may be linked to underreporting which is still a dominant issue of the ESWD database (Púčik et al., 2019). Hence, in statistical terms, this means that we can assess the hit and the miss rates in the contingency table of hailstorm events describing the joint distribution of the forecast (in this case, the reanalysis-OT combination) and observations (the hail reports) for the performance assessment. Contrariwise, it is not possible to evaluate crucial aspects such as the number of false alarms or the correct negatives.

The matching conditions for analyzing the correspondence between satellite detections and hail reports consist of a temporal window of ± 1.5 h around the OT detection time and 25 km from each OT location. This relatively sharp temporal window is considered given the high temporal accuracy below or equal to 1 h characterizing 90% of ESWD reports (Figure 4.1b). The 25-km distance criterion is retained from Bedka (2011) and Punge et al. (2017) and accounts for: a maximum storm motion of 60 km/h, possible errors in parallax correction in the post-processing of OT detections with heights differing from 14 km, and possible latitude/longitude erroneous reporting for ESWD events. The matching comparison is applied considering both the original and the hail-filtered OT datasets (over land only) to investigate possible differences owing to the hail-filter procedure. Different ESWD reports sets are employed: the whole set of 1386 reports considered up to now, only reports having Temporal Accuracy (TA) $\leqslant$ 1 h, only reports having quality level QC1 or superior, and the combination of these two conditions. The results of the comparisons are reported in Table 4.2.

An OT of the unfiltered dataset is found in the vicinity of 69.9% of the complete ESWD hail database, while the fraction of OTs associated with hail reports is 2.39%. For hail-filtered OTs, the hit rate slightly reduces to 62.8%, and the portion of OTs matching ESWD reports increases to 2.80%. With the refinement in the selection of ESWD reports by considering higher quality levels and temporal accuracy, the peak performance is obtained when considering both QC1> and TA$\leqslant$ 1 h conditions together. In this case, the hail reports hit rate increases to 72.1% and 65.2% for the original and hail-filtered OT sets, respectively. A lower fraction of OTs matching the reports is seen (2.26% and 1.89% for hail-filtered and original OTs, respectively) mainly due to a reduction of 27% in the number of hail reports considered (from 1386 to 1009) owing to the more stringent conditions applied. For all four comparisons, a loss of ~7% in report hit rate is obtained by comparing the hail-filtered to the original OT sets (corresponding to less than 100 ESWD reports). As regards the fraction of OTs found in the vicinity of reports, a gain of ~0.4% is always seen comparing filtered with unfiltered OT detections as a consequence of the smaller sample size of the former set.

To give an objective evaluation of these results, it is worth a comparison with a similar analysis previously performed by Punge et al. (2017). However, given the several substantial differences with their research design, the qualitative nature of this comparison must be pointed out. Particularly, the main dissimilarities of their study with the investigation here reported are: a larger spatial domain covering all of Europe (extending from England to Russia and from Norway to Egypt) over ten years (2004-2014), the employment of a previous non-probabilistic version of the OT detection algorithm from satellite infrared imagery, and the substantially coarser description of hail-favoring convective environments owing to ERA-Interim reanalysis (having a horizontal grid spacing of roughly 80 km). Using a spatio-temporal window of 25 km and ± 1h around each OT detection, they found a hit rate decreasing from 40.3% to 39.7% for the unfiltered and filtered OT datasets, respectively, for a subset of 2475 ESWD reports with quality level QC1 or superior and TA $\leqslant$ 15 min. This indicates that, considering the hail-filtered OT set performance obtained in the present analysis for the most similar conditions (i.e., QC1> & TA $\leqslant$ 1h - Table 4.2), the ESWD hail reports hit rate exceeds 25% more than what obtained by Punge et al. (2017). The main source for this improvement is likely imputable to the enhanced ability to detect overshooting top occurrences and avoid false detections, thanks to the probabilistic algorithm developed by Khlopenkov et al. (2021). As a possible indication of this, the fraction of OTs found in the vicinity of hail reports in Punge et al. (2017) went up from 0.67% before to 0.84% after the filter, conversely to the values of 1.89% and 2.26%, respectively found in the present analysis (Table 4.2). However, even if an increase in the fraction of OTs associated with hail reports is detected, its absolute value remains low. This is most likely imputable to under-reporting issues characterizing the ESWD dataset more than the impact given by possible OT false detections. Indeed, under-reporting constitutes the main limitation of the observed state estimated through ESWD (Púčik et al., 2019) and weighs more than possible false alarms in the matching assessment with the hail proxy. However, despite this limitation, ESWD represents the best available dataset to estimate the real state of hail occurrence.

4.4 Separating hailstorm environmental conditions

Given the promising performance shown in the detection of hail events with the reanalysis-OT-based approach, it is worth investigating the relative environmental characterization during the occurrence of ESWD hailstorm events with the five reanalysis parameters considered for the filtering. In the following, are presented the behaviors and interrelations among the numerical convective descriptors in the presence of hit or missed ESWD reports considering the least stringent conditions for matching hail-filtered OT detections (i.e., the first column in Table 4.2). Furthermore, to analyze the relationship between environmental conditions and hail severity, only the subset of 1050 ESWD hail reports having information on hailstone size are considered (in the following referred to ESWD-S), and results are presented by separating between small (<3 cm) and large hail ($\geqslant$ 3 cm). The reason behind these separations is to investigate the appropriateness of the hail-filtering approach for different hailstorm conditions and to understand the causes for its inability to detect certain hail events (i.e., missed ESWD reports). Additionally, the thermal characteristics of satellite OT detections in the presence of ESWD-S hail reports are analyzed.

The cumulative density functions of SPHERA convective parameters in the presence of ESWD-S reports, separated into four categories based on matching and hail severity conditions, are reported in Figure 4.12. Among the ESWD-S subsample, 66% are successfully detected with the filter method. Out of them, the majority (68%) pertain to large hail, while missed reports are almost equally distributed between the two hailstone size categories. Moving from small to large hail, a tendency of the distributions to shift towards parameter values with higher potential for severe convection (i.e., larger CAPE, K, and DLS and smaller SLI) and higher

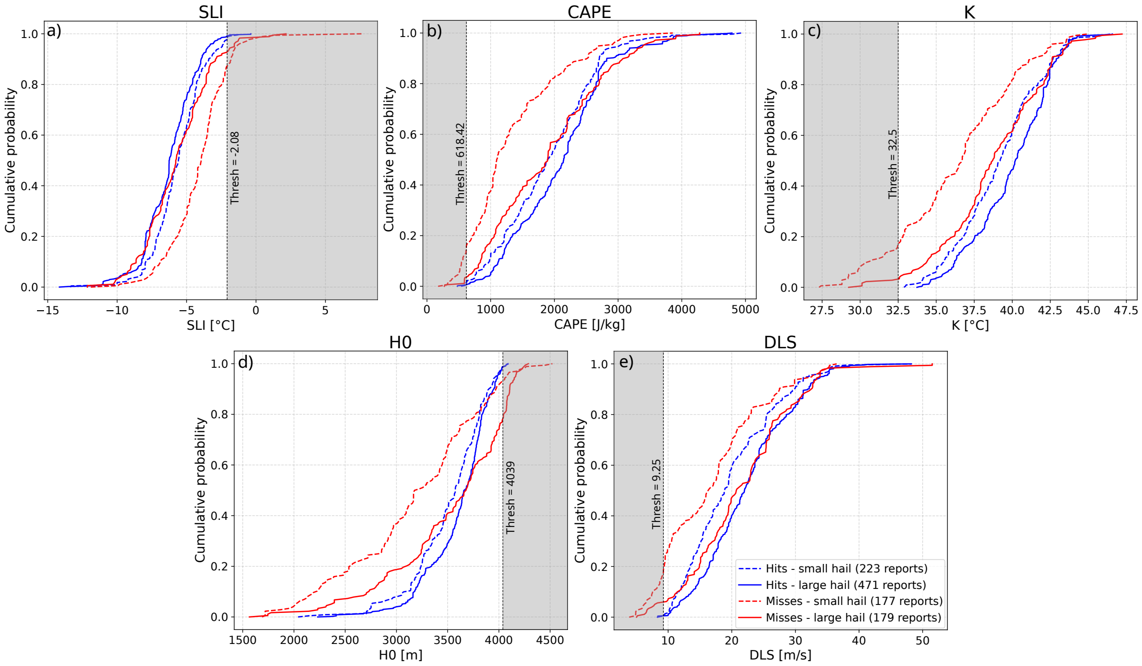

Figure 4.12: Cumulative density functions of the five parameters selected from SPHERA in the presence of ESWD-S hail reports. The same criteria described in Section 4.2.2 apply to spatio-temporally aggregate the parameters in the vicinity of hail reports. Reports are divided into small hail (< 3 cm, dashed lines) and large hail ($\geqslant 3$ cm, solid lines) when hit (in blue) or missed (in red) by the hail-specific OT dataset (Table 4.2 - "All reports" column). The black dashed vertical lines indicate the thresholds used for the filtering defined in Table 4.1. The shadowed portion of the distributions reports the range of values when the filter is active. a) SLI, b) CAPE, c) K, d) H_0, and e) DLS.

freezing level heights H_0 is detected. Indeed, the hit–large hail class (Figure 4.12 - solid blue lines) shows this systematic trend for every parameter and almost the entire range of the distributions, but always being very close to the curves pertaining to hit–small hail (dashed blue lines) and missed–large hail (solid red lines). The most evident separation for the whole set of parameters emerges for the missed–small hail class (dashed red lines), showing cumulative density curves significantly shifted towards values with less potential for convection and lower freezing level heights. Interestingly, only for the 2.5% of hit ESWD-S reports at least one parameter falls in its filtered data range (i.e., shadowed areas in Figure 4.12), while the fraction increases to 33.6% for missed reports.

To deeper understand the relationships between ambient convective descriptors and different hailstorm severity and matching conditions, the phase spaces of the parameters for the four samples of hail reports are considered in the form of bivariate histograms. Figure 4.13 shows the joint distributions of H_0 and K index for the four hailstorm classes. In all cases, a joint increase of freezing level height with atmospheric instability and low-level moisture content is noted, suggesting a positive linear relationship between H_0 and K. This is particularly evident for missed hail reports (panels c and d). In fact, the distributions for hits (panels a and b) are more compact, as indicated by the sharper interquartile (IQR) ranges of both parameters, reported with blue dashed lines. Further, they are shifted towards higher values and do not show strong differences between small and large hail cases. On the other hand, the missed reports counterparts show distributions covering wider ranges and extending towards lower H_0 and K, with a significant difference between large hail (median values of 38.7°C for K and 3678.6 m for H_0) and small hail (36.7°C for K and 3176.7 m for H_0). This suggests that hail, in case of missed reports, especially for small hailstones, tends to form in environments with lower instability and freezing level height. Furthermore, the tendency for missed–small hail parameters distribution to populate the lowest range of values, extending to freezing levels below 2000 m and K <30 °C, is linked to convective environments in which the overall cool thermal conditions of the atmosphere limit the presence of supercooled liquid water in the updraft, ultimately blocking hail growth (e.g., Dessens et al., 2015). Appendix A reports the joint distributions of the parameters for the hit/miss small/large hail separation for the cases not treated here to avoid an excessive number of figures in the main text. Similar trends to that of Figure 4.13 are detected for the relationships between H_0 and the other atmospheric instability parameters (CAPE and SLI, Figure A.1 and A.2 respectively).

A more dispersed and less explicit relationship emerges from the analysis of the joint distributions between H_0 and DLS (Figure 4.14). In all four classes, DLS covers a broad spectrum with IQR ranges exceeding 10 m/s, in agreement with the difficulty in separating hailstorm events based on their hailstone size using DLS, also found by Kunz et al. (2020). Hit reports (Figure 4.14a-b) show median DLS values of 22.0 m/s and 18.9 m/s for large and small hail, respectively. Moving to missed reports (Figure 4.14c-d), the central values of the distributions decrease to 20.7 m/s and 16.3 m/s respectively. The spread of H_0, as seen in Figure 4.13, is more circumscribed for hit reports, with IQR extending to 367.4 m for large hail and 501.9 m for small hail, and decisively increasing to 765.9 m and 782.3 m for missed reports of large and small hail, respectively. As suggested from the results of Figure 4.12, the most diverse parameter distributions are obtained in the case of missed–small hail class. This is confirmed by the corresponding bi-variate histogram of H_0 and DLS (Figure 4.14d), for which, despite the large spread in both parameters, the central values of their distributions indicate environments associated with lower freezing level altitudes and less storm organization (with p_{25} of DLS being 9.60 m/s, very close to the value defined as the threshold of 9.25 m/s) compared to the other three classes of events considered.

Finally, the relationship between atmospheric instability and storm organiza-

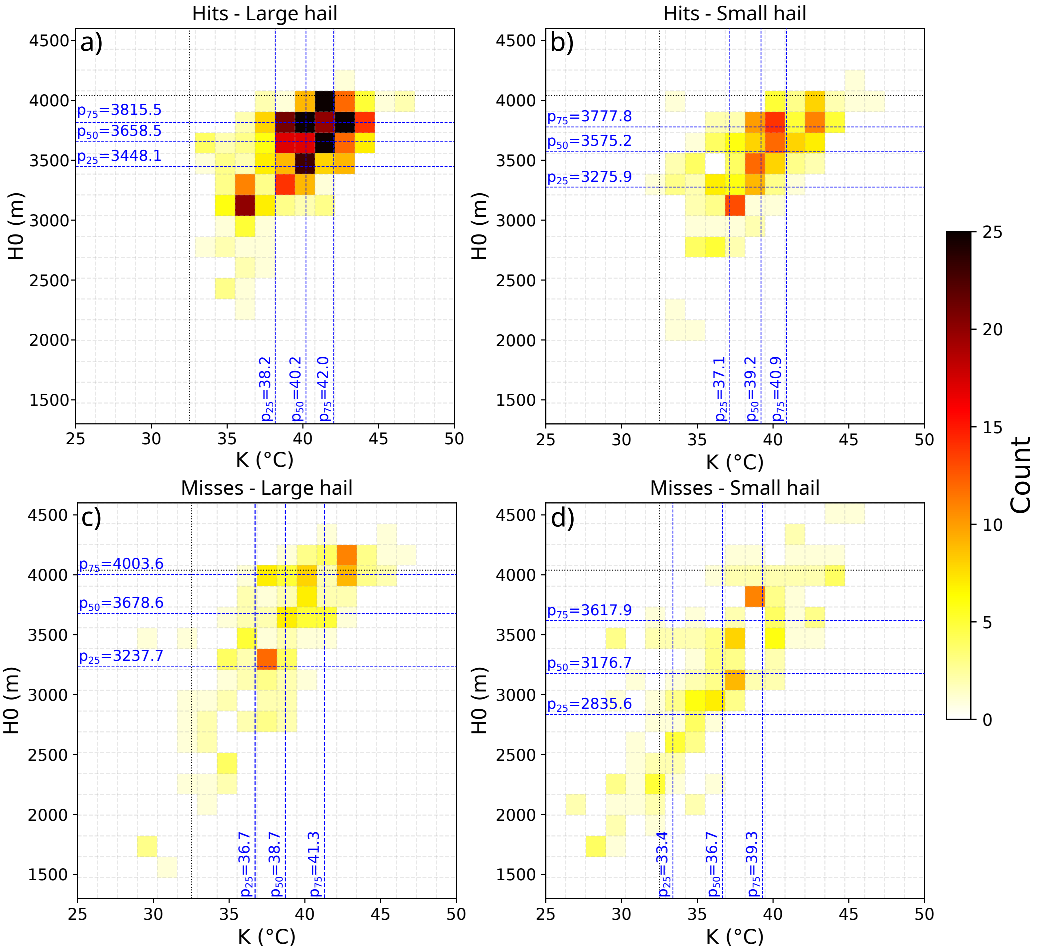

Figure 4.13: Bi-variate histogram distributions of H_0 vs. K index in the presence of ESWD-S hail reports for the separation considered in Figure 4.12: a) hits–large hail, b) hits–small hail, c) misses–large hail, and d) misses–small hail. The blue dashed vertical and horizontal lines represent the median (p_{50}) and the interquartile (IQR) range values (p_{25} and p_{75}) of the distributions. The black dotted lines report the thresholds used for the filter defined in Table 4.1.

tion is analyzed by reporting the phase spaces of CAPE versus DLS for the four hail classes (Figure 4.15). Also in this case, a significant spread in the parameters distributions and a not-so-direct link between CAPE and DLS emerges. CAPE, for the hit–large hail cases (Figure 4.15a), shows a median value of 2146.4 J/kg with an IQR spread of 1065.1 J/kg, which stays roughly equal when moving to hit–small hail Figure 4.15b) while the median decreases to 1892.7 J/kg. The CAPE distribution for missed–large hail (Figure 4.15c) remains similar to that of hit–small hail, with a very closer median and an enhanced IQR range extending to 1360.2 J/kg. Finally, a decisive shift towards lower CAPE values is detected for missed–small hail (Figure 4.15d) cases, presenting a central value of 1094.4 J/kg. The reasoning behind DLS behavior is the same as that discussed in Figure 4.14. Hence, also in this case, the most different conditions emerge for the missed–small hail class, characterized

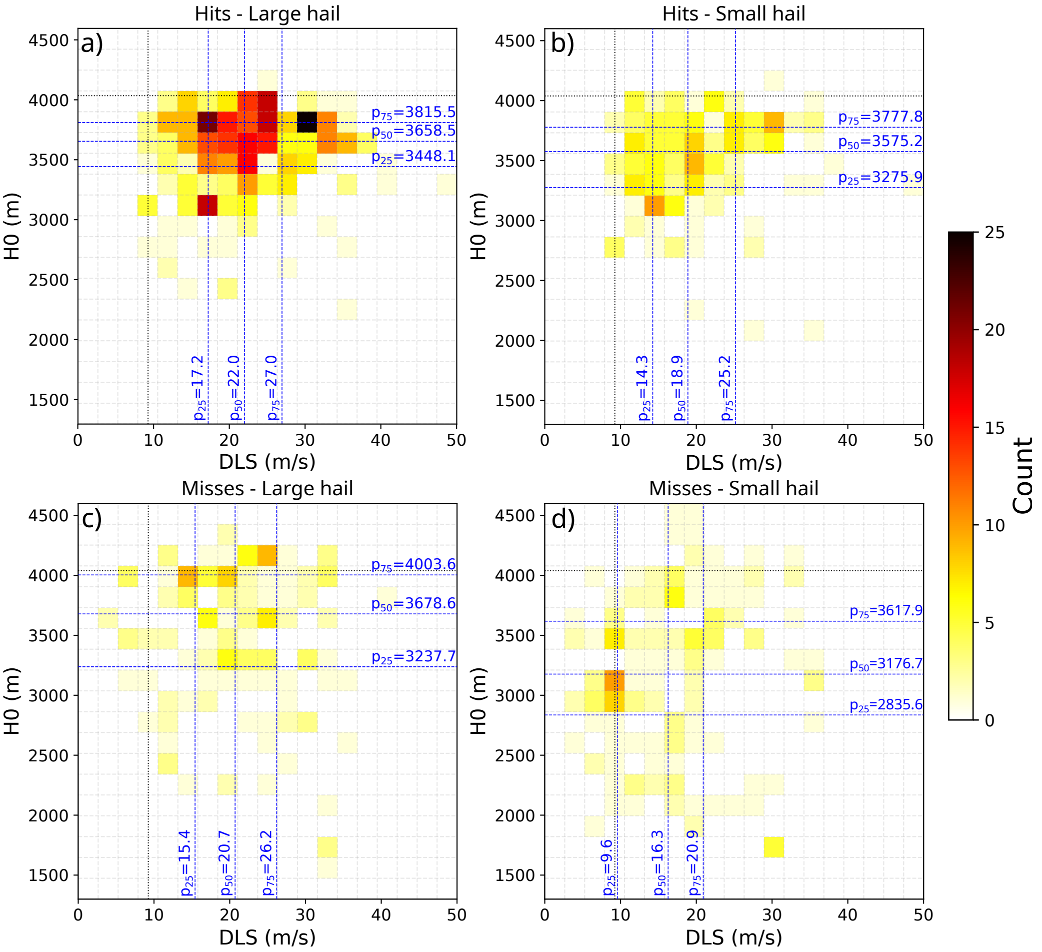

Figure 4.14: As Figure 4.13, but for H_0 and DLS.

by generally pronounced low-CAPE and low-DLS environments. Similar degrees of dispersion emerge also from the joint distributions of K and SLI with DLS (Figure A.3 and A.4, respectively), as well as an evident shift of the central value of the distributions towards conditions with less potential for DMC for the misses-small hail class.

A factor playing a central role in the identification of an overshooting top from satellite scans data is the thermal characteristic of the cloud top where the OT can be found. Previous research showed how OTs linked to deep convective clouds can be detected as cold pixels in visible and infrared satellite imagery scans (e.g., Morel and Senesi, 2002; Mikuš and Mahović, 2013). These cold spots are associated with small and sharp infrared brightness temperature (IRBT) minima that are near to or colder than the tropopause temperature associated with the anvil cirrus cloud. For this reason, a critical variable included in the Khlopenkov et al. (2021) algorithm for automatic OT detection employed in this study is the temperature difference ΔT between infrared brightness and tropopause. A large ΔT (> 6 K) indicates a strong penetration of the updraft through the anvil of at least 1 km (Griffin et al., 2016). Consequently, these observed thermal variables are worth investigating to get a more

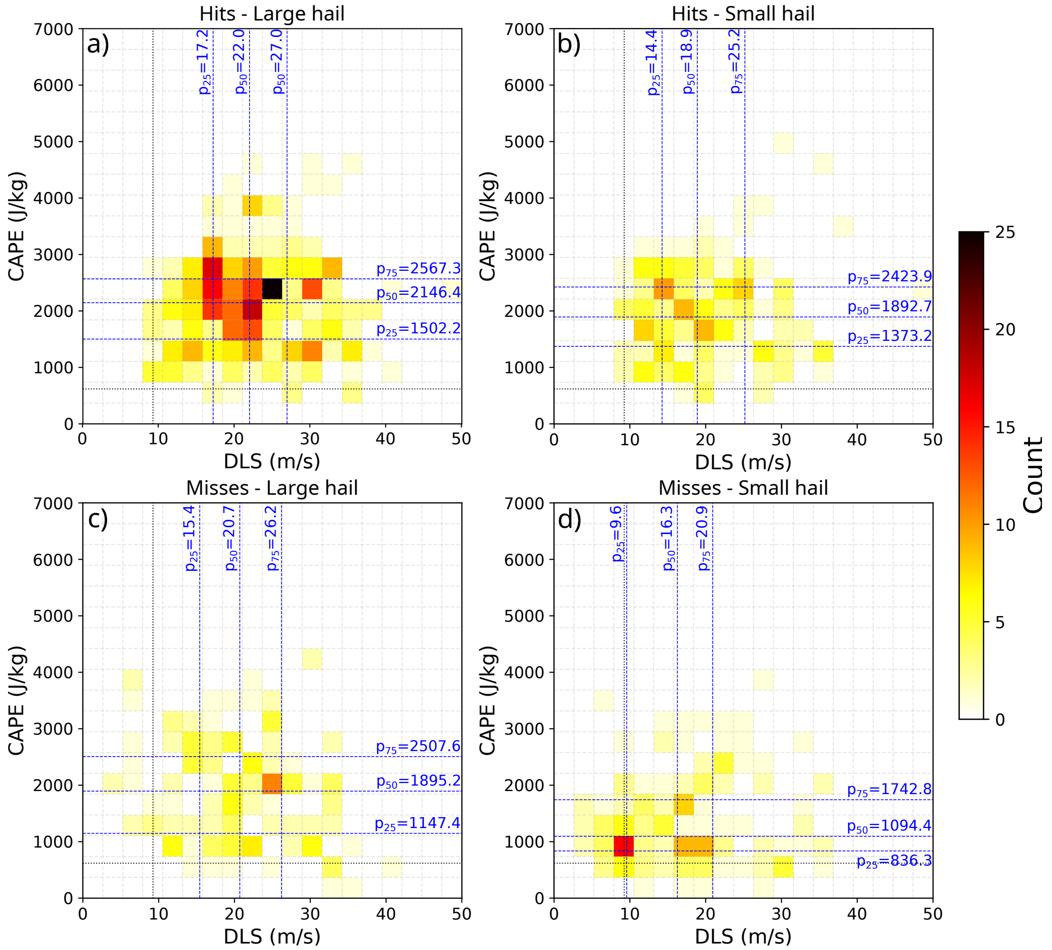

Figure 4.15: As Figure 4.13, but for CAPE and DLS.

complete picture of the environmental conditions established during hailstorms. To do so, all satellite detections are considered (for any OT probability of occurrence that up to now has been constrained to be >0.5) in the presence of hit or missed ESWD and ESWD-S reports datasets, when matching the hail-specific OT dataset and separating among small, large, and very large hail.

Figure 4.16 shows the distributions of IRBT and ΔT minima in separate presence of hit and missed hail reports, where the continuous probability density functions (dashed curves) are calculated with a kernel density estimate method. Hit reports present sharp IRBT minima distributions (Figure 4.16a) centered towards mean values of ∼211 K and rarely exceeding warmer temperatures than 224 K. No significant deviations between different hail severity subsamples are detected, only a slight opposite behavior than expected with the average IRBT for small hail (210.98 K) being 0.36 K colder than large hail counterpart (211.34 K), most likely imputable to the excessively small data sample considered. Considering ΔT minima for hits (Figure 4.16c), an almost total absence of positive values is noted, meaning that IRBT temperature is almost always colder than tropopause anvil temperature, as expected from severe thunderstorms producing prominent overshooting tops. Fur-

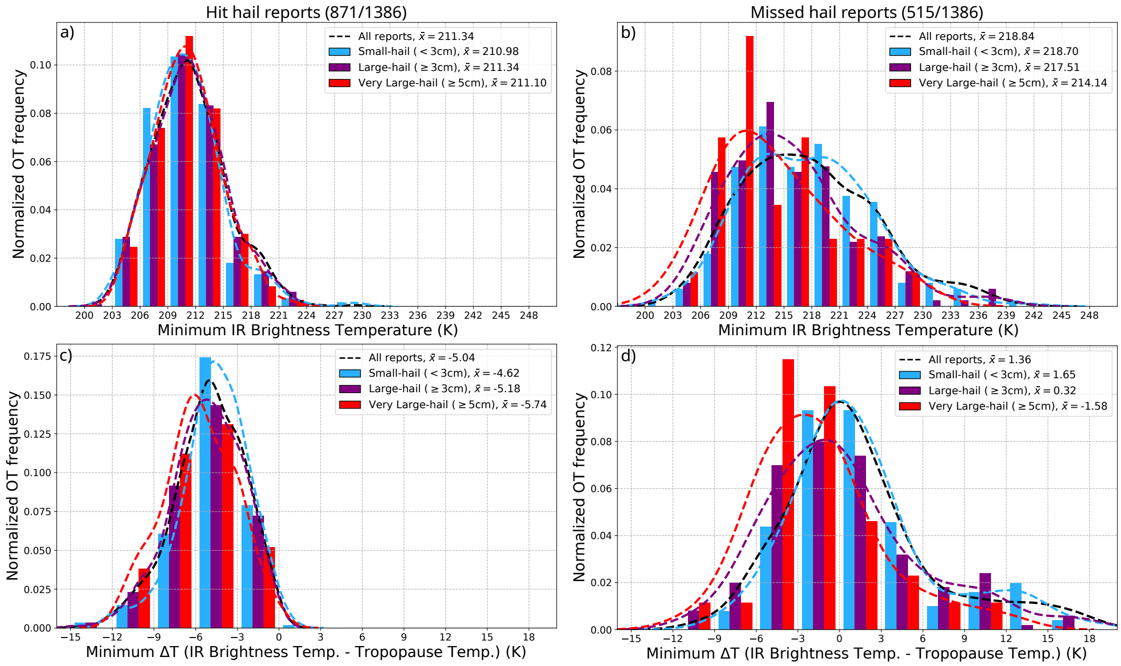

Figure 4.16: Normalized distributions of minimum IRBT (top row) and ΔT (bottom row) of OT detections in the presence of ESWD hail reports. The histogram bars quantify the normalized frequency of OTs in the presence of the ESWD-S subset for small hail (light blue), large hail (purple), and very large hail (red). Legends report the mean values for every distribution. The kernel-density estimated probability density functions are reported with dashed curves in respective colors, additionally, the density functions for the whole ESWD set (including all reports, also those without maximum hailstone size information) are displayed with black dashed curves. Hail reports are divided into hit (panels a and c) and missed (panels b and d) when matching with the hail-specific OT dataset.

ther, a slightly more pronounced separation among hail severity classes is evident, with an average ΔT difference between small and very large hail exceeding 1 K (i.e., -4.62 K for small hail vs. -5.74 K for very large hail). Moving to missed reports, more blunted and warmer IRBT minima distributions are detected, which extend to temperatures as high as 240 K (Figure 4.16b). These generally warmer conditions are also reflected by the mean values, which suggest a more pronounced separation among hail severity classes. Indeed, the minimum IRBT for very large hail is on average ∼4.5 K colder than small hail counterpart (i.e., 214.14 K for very large vs. 218.70 K for small hail). The analysis of ΔT minima (Figure 4.16d) confirms and strengthens these results: the majority (i.e., 56%) of the "All reports" distribution presents positive ΔT, holding especially in the case of small and large hail for which the positive portions are 56% and 48%, respectively, that reach and exceed values as large as 15 K. These conditions indicate tropopause temperatures substantially colder than those of the detected OT, thus suggesting not particularly prominent protrusions over the cloud top anvil. This is also reflected by the ΔT minima average, that for small hail (1.65 K) is more than 3 K warmer than very large hail counterpart (-1.58 K), i.e., more than three times the difference detected for hit reports.

4.5 Summary and discussion

4.5.1 Summary of the results

This work presents an enhanced hailstorm identification approach based on a combination of CP reanalysis proxies from SPHERA, satellite overshooting top detections, and crowdsourced ESWD hail reports. The analysis is performed over three years (2016-2018) during the extended summer season (April-October) to investigate the environmental conditions necessary for hailstorm development and, ultimately, the potential to describe hail hazard over Italy and surrounding countries and seas. For this purpose, a filter to identify OTs potentially linked to hailstones formation in a thunderstorm, taking into account their surrounding environmental conditions, is developed. This filtering technique considers five numerical proxies describing convective potential and quantifying key ingredients for hail formation in a thunderstorm that are: atmospheric instability and low-level moisture content, the degree of organization of the storm, and the altitude of the freezing level. The identified thresholds for the filter show good agreement with minimum conditions for severe convection development and agree with previous findings in this context (Kunz, 2007; Púčik et al., 2015; Taszarek et al., 2020; Jelić et al., 2020; Kunz et al., 2020). The spatio-temporal analysis of the singular parameter contributions to the filter highlights their different impacts in the identification of hail-prone conditions. Their combination shows enhanced non-hailing OT removal over high-elevation terrains during the early morning and afternoon of Spring and late Summer/Fall months. The resulting OT-hail proxy characterization delivers a maximum hail potential over northern Italy pre-Alpine areas in the summer months of June and July peaking at 15 UTC, but with strong intra-annual and inter-annual variability. Significant differences are also detected considering daytime and nighttime periods separately, with enhanced OT frequency over land during the day and a more important contribution over the sea at night. From the statistical matching comparison with the ESWD database, a hail-filtered OT is found in the vicinity of 65.2% of reports, exceeding 25% more than what previously found with a similar but less sophisticated methodology (in terms of OTs detection and reanalysis proxies – Punge et al., 2017). Finally, investigating the relationships between convective parameters and hail reports separating between hit/misses and small/large hail, enhanced appropriateness of the developed method in case of severe hailstorm events is suggested. Indeed, the majority (68%) of correctly identified reports are linked to storms producing hailstones with diameters exceeding 3 cm. These show comparable environmental signatures to that of hit–small hail and missed–large hail samples. The class of reports presenting the most distinct conditions from the others is that of missed reports with hailstones having small diameters (<3 cm), accounting for the 17% of the sample. In this case is found a general tendency for hail to form in ambients characterized by systematically lower instability, less storm organization, lower freezing level heights, and presenting warmer cloud top temperature signatures, making it more challenging to be detected with the developed methodology.

4.5.2 Comparison with similar studies

The obtained OT-hail proxy characterization shows good agreement with recent and extended hail climatologies over the region (Punge and Kunz, 2016). No particularly favorable conditions for hail development are detected in southern Germany, which is instead considered a main hotspot in Europe for hail hazard (Fluck et al., 2021; Punge et al., 2017, 2014). The reason is most probably the limited temporal extent of the analysis performed which may produce a significant deviation from the long-term statistical tendency. Pre-Alpine regions over Switzerland, Austria, and Slovenia show an enhanced likelihood for hail formation as described by the OT-hail proxy. In contrast, the most prominent minimum for hail likelihood over the whole spatial domain is found throughout the inner Alpine arc. These results agree with a 15-year radar-based characterization over the western Alps (Nisi et al., 2018, 2020), which found evidence of the nighttime hailstorm hotspot confined on the southern Alpine foothills area along the Swiss-Italian border also detected in the present analysis. Good accordance is found with the Austrian hail hazard map based on radar and point observations developed by Svabik et al. (2013). Particularly, the most hail-prone eastern Alpine region, according to the resulting OT-hail frequency, is detected over south-eastern Austria and Slovenia, even if, over the latter, a climatological regime characterized by considerable variability in hail occurrence has been recently highlighted (Stržinar and Skok, 2018). Close to this area, also the northern Adriatic shows high potential for hailstorm development, representing the primary marine hotspot over the whole region, particularly enhanced during nighttime along the Croatian coastline. A similar pattern is detected by Jelić et al. (2020), who, considering 60 years of data, found significant spatial variability in hail occurrence with enhanced likelihood in central summer months around 14:00-16:00 CEST in the hinterland and a significant peak in hail probability also in Fall months at 8:00-12:00 CEST over the coastal area. Over the Italian peninsula, the most favorable conditions for hail development are found throughout the northern pre-Alpine regions, where the absolute maximum in OT activity over the whole spatial domain is detected. Nevertheless, the potential for hailstorm formation is met more or less in all Italian regions. This is in agreement with several regional radar-based hail climatologies over north-western Italy (Davini et al., 2012), north-eastern Italy (Sartori, 2012), especially over the easternmost Friuli Venezia Giulia region as described by the hailpad-based climatology of Giaiotti et al. (2003). Additionally, the first developed national hail climatology (Baldi et al., 2014) found evidence of non-homogeneous spatial distributions in hail frequency of occurrence over Italy, with local maxima over the northern regions and minor hot spots in central and south-western coastal areas, in agreement also with Taszarek et al. (2020). A recently developed Italian hail characterization obtained by combining ESWD reports and regional hailpad data with machine-learning-derived optimal large-scale meteorological variables from ERA5 over 1979-2020 (Torralba et al., 2022) further confirms these tendencies. Moreover, the monthly hail distribution obtained with their method, confirms our detected tendency of an increasing likelihood for hailstorm events in Spring. The maximum detected in June and July over northern regions, then decreasing and shifting towards lower latitudes along central-southern coastlines in September and October. Good temporal matching is also found compared with the recent ESWD-based hail statistics over 1990-2018 (Púčik et al., 2019).

The results of the analysis investigating ambient characteristic separations and interrelations based on hail severity and OT-matching condition agree with previous environmental analyses. The conditions identified by Torralba et al. (2022) for the occurrence of hailstorm events over Italy, as described by ERA5 reanalysis proxies when separating between all and severe (i.e., diameters >2.5 cm) hail, include CAPE values exceeding 900 and 1000 J/kg respectively for the two classes, and K index with a median of 30 °C in case of all hail, with a significant increase detected considering only severe hail. This suggests a good agreement with the necessary enhanced instability for severe hail development also detected in this work and other previous research (e.g., Marcos et al., 2021; Púčik et al., 2015). Regarding the remaining parameter quantifying atmospheric instability (SLI), Kunz et al. (2020), considering ERA-Interim reanalysis estimates, found only a slight difference between small and large hail conditions (∼-3.7 °C and -4.5 °C respectively). In the same study, a more decisive separation emerged for deep layer wind shear (DLS), which, however, is more linked to the length of the storm tracks included in their analysis than to the relative hailstone size. Indeed, distinguishing between hailstones with diameters <3 cm and ⩾5 cm, they obtained similar scattered distributions showing DLS ranging ∼0-30 m/s for the former and ∼5-30 m/s for the latter, confirming the high degree of dispersion in the relationship between hail severity and storm organization detected in the present work. This result is also confirmed in the analysis with proximity radiosounding data performed by Púčik et al. (2015), where a higher degree of separation using DLS emerges, differentiating between severe (2-5 cm) and extremely severe (⩾5 cm) hail compared to the less distinct counterpart between non-severe (⩽2 cm) and severe hail. This indicates that small and medium hail can also form in weakly organized storms, while very large hail exceeding 5 cm requires strongly sheared environments which are almost exclusively related to supercell thunderstorms (Púčik et al., 2021; Blair et al., 2017). The joint distribution of instability and organization parameters (CAPE and DLS) has been used in numerous studies as a proxy for severe thunderstorms. Specifically for hail, Púčik et al. (2015) found wide ranges in the joint distribution but with a clear lack of severe events in low-CAPE and low-DLS environments, which instead tend to concentrate on the opposite high-CAPE and high-DLS range, similarly to the tendency found in this analysis. Likewise, Taszarek et al. (2020) also found more stringent ambient conditions with increasing hail severity towards higher values in CAPE and DLS phase spaces. This is systematically more evident in the United States than in Europe, where the distribution of CAPE is shifted towards lower values. The freezing level height characterization is also in accordance with similar studies: Jelić et al. (2020) found an upper limit for H_0 of ∼4000 m above which they observed no hail, similar to the upper ranges of the distributions found in this analysis for which hail events exceeding this altitude (which is also close to the threshold of 4039 m selected for the filter) are found only rarely and mainly in case of missed reports. The separation between all and severe hail conditions, as described by Torralba et al. (2022), yielded a median H_0 of ∼3400 m for the former and increased to ∼3500 m for the latter, confirming the tendency identified in the present analysis. Finally, the resulting joint distribution of freezing level height and instability, indicating a positive relationship for which large hail can form in environments with low instability if H_0 is low while larger instability is needed if H_0 is higher, has been detected also by Prein and Holland (2018).

The proposed approach mainly relies on the exclusion of conditions for which hail is very unlikely to form by the application of the environmental filtering procedure on satellite OT data. Other possible approaches exist more focusing on the identification of hail occurrences without excluding unlikely situations but by training a model to estimate favorable ambient conditions for hail on a climatological timescale through, for example, logistic regression models (Rädler et al., 2018; Mohr et al., 2015b). With these approaches, a quantification of hail potential over past and future timescales is possible over broad spatial domains and without the necessity of direct hail observations. However, this missing constraint may produce significant discrepancies (Mohr et al., 2015a).

Additionally, in recent years the employment of machine learning techniques to objectively identify the predictors best performing for the description of hailstorm environments in a multivariate approach is receiving considerable interest (Scarino et al., 2022; Torralba et al., 2022; Gagne et al., 2017). These approaches allow greater dimensionality within forecast parameters, which could potentially reduce the overall uncertainty in the ambient description, but have also proven computationally expensive. Furthermore, despite the less influence on subjective choices for the identification of hailstorm environmental parameters, in the majority of these studies, the best-performing predictors coincide with the variables considered to characterize hail-favoring environments employed in the present and other similar works (Brooks et al., 2003; Kunz, 2007; Kunz et al., 2020; Punge et al., 2017; Bedka et al., 2018; Raupach et al., 2021; Punge et al., 2021).

4.5.3 Limitations and potential improvements

The method presented in this analysis has demonstrated its appropriateness for hailstorm identification over south-central Europe, particularly in case of severe conditions producing large hailstones (with diameters $\geqslant$3 cm). This is revealed by the majority of detected ESWD reports presenting large maximum hail diameters and sharper distributions of their thermodynamical and kinetic environmental characteristics. In the case of less severe hailstorms producing smaller hailstones, a larger degree of uncertainty is detected through less explicit OT signatures characterized by warmer cloud top temperatures and by environments showing systematically lower instability and shear and colder atmospheric profiles. Hence, the fixed thresholds introduced to filter the high-resolution reanalysis proxies could be inappropriate in these occasions. Anyhow, despite this limitation, since the main interest is to enhance the understanding of the generation processes and characteristics of damaging-producing hailstorms, our methodology has demonstrated to be appropriate for the scope.

The main limitation of this study is the very short time period of three years considered (2016-2018), which, as demonstrated by the detected large intra-annual variability, is insufficient for a robust assessment of hail likelihood of occurrence. However, it is sufficient for the purpose to present the novel methodology developed and assess its potentialities. Given the promising results demonstrated, a further extension of the analysis, including the years 2019 and 2020, is planned soon. During these two years, additional ~900 ESWD reports have been recorded, which could constitute a decisive increase in the robustness of the analysis, as well as improved reliability in the use of crowdsourced hail observations. Afterward, ideally, a backward extension down to 2010 is envisaged in order to present a minimum of 10-year hail statistics.

Another limitation is constituted by the imperfect OT observations detected with IR satellite instruments. Indeed, a large amount of missed ESWD reports in this study is linked with cloud top temperatures warmer than the tropopause, which are more difficult to detect with the satellite-based approach developed by Khlopenkov et al. (2021) applied to the MSG SEVIRI instrument. This difficulty, as noted by Cooney et al. (2021), may be linked with satellite scans performed when storm tops are not optically thick enough at the initial stages of the storm development, causing the radiation measurements to be performed in warmer and deeper regions of the cloud than the updraft top, likely owing to the insufficient density of the particles in this region (Sherwood et al., 2004). This results in a lower magnitude of the temperature difference between tropopause and cloud top, ultimately reducing the probability of OT detection and, consequently, causing to miss the event. Hence, possible ways to overcome this issue and improve the OT detection methodology are the inclusion of visible channel textures together with IR detections (Bedka and Khlopenkov, 2016), or the enhancement in the spatio-temporal resolution of the satellite scanning that in this study are limited to 3 km and 15 minutes, respectively. Particularly, an higher temporal frequency for detecting thunderstorm cloud tops, such as that possible with the Geostationary Operational Environmental Satellites (GOES) 16 or 17, is essential given the rapidity of their formation and dissipation, which may be even below 15 minutes (Elliott et al., 2012).

Other potential sources of improvements in hailstorm characterization may emerge by considering additional observational data types, given the necessity for as comprehensive an approach as possible. Indeed, several techniques exist for hail detection relying on different information, such as satellite microwave scattering data (Laviola et al., 2020) or ground-based hailpad networks (Manzato, 2012; Eccel et al., 2012). In this context, radar reflectivity observations inclusion could represent a significant improvement, constituting one of the best proxies for severe convective thunderstorm identification (Fluck et al., 2021; Nisi et al., 2020). In fact, the higher spatio-temporal resolution in the atmospheric description possible with radar data, compared to satellite detections which in this work are limited to a 15 min frequency, is crucial to achieve given the rapid and local-scale evolution of DMC phenomena. However, while European archives aggregating data from multiple national weather radar networks exist, such as the Operational Program on the Exchange of Weather Radar Information (OPERA – Saltikoff et al., 2019), at the present moment, they do not include Italy where, additionally, the regional fragmentation of the weather services maintaining the instruments makes even more challenging to retrieve historical series of radar products on a national basis for the purpose. Another potential source to enhance hailstorm detections and characterization could be automatic lightning detections presenting rapid intensification in the total flash rates (i.e., lightning jump, LJ - Schultz et al., 2009). LJs has, in fact, demonstrated to be a promising proxy for detecting severe thunderstorms (Taszarek et al., 2020; Farnell et al., 2017), particularly for hailstorms (Farnell et al., 2018; Jurković et al., 2015). However, although it was initially considered to include in the analysis lightning data from the Italian LAMPINET dataset (Biron, 2009), the identification of LJs, as defined by Schultz et al. (2009) (which is now the most widely accepted definition in the scientific community), has demonstrated to be impractical with LAMPINET owing to the very high sampling frequency on the order of 100 µs required for the purpose.

An important innovation introduced in the analysis here presented is the use of high-resolution numerical simulations to describe atmospheric ambient conditions, which is considered a promising avenue of development in hail research (Allen et al., 2020). This is opposed to coarser global datasets employed in similar studies up to now, whose simulations, through the use of physical parameterizations to account for deep convection, may be affected by significant errors and inaccuracies (Prein et al., 2015). In this work, the atmospheric characterization has been carried out employing the new regional reanalysis SPHERA, which, thanks to the high horizontal resolution of 2.2 km, allows the explicit representation of most deep convective motions. However, unfortunately, a quantification of the added value introduced by this innovation was not feasible, and, generally, no quantitative studies have been performed so far to assess the benefits of finer spatial grid spacings configurations over coarser datasets specifically for the reproduction of hailstorm environments. A possible contribution towards this direction can be found in the study of Kunz et al. (2020), who compared the ambient conditions favorable for hailstorm development simulated by ERA-Interim (with ∼80 km horizontal resolution) with those of coastDat-3 reanalysis (∼11 km) over central Europe for a 10-year period. The results suggest that no significant improvements emerge of one dataset over the other for the simulation of the dynamical parameters, while a higher model resolution is envisaged for estimating the thermodynamical signatures of hailstorm ambient conditions. Nevertheless, this investigation does not pertain to CP-scale simulations and has been limited in the comparison of only two specific numerical proxies (i.e., vertical lapse rate and deep layer shear). Hence, a possible future direction could be to conduct a sensitivity analysis on the driver dataset providing hail environmental predictors to better understand the role of CP simulations in this context.

In conclusion, to respond to the second question posed in the introduction of the thesis:

“Can the combination of CP-simulated ambient predictors with observational proxies for hail enhance the ability to detect hailstorm events and improve the understanding of the necessary convective environmental conditions for their development?”

The proposed methodology, relying on high-resolution ambient numerical predictors for hail development from SPHERA, coupled with satellite overshooting cloud top detections and surface-based hail reports, demonstrated its potential in the distinction between hailing and non-hailing convective conditions. In particular, the spatio-temporal characterization of the obtained hail-proxy has demonstrated to be in line with several recent hail climatologies developed with different techniques, its ability to match with direct hail reports has shown significant improvement compared to a previous similar estimate, and the description of hailstorm environmental conditions revealed separate ambient signatures associated with different hailstorm severity classes.

Chapter 5

Conclusion and future perspectives

The aim of this thesis, as stated in the Introduction (Section 1.4), was to investigate the potential benefits brought in by a new convection-permitting regional reanalysis centered over Italy, SPHERA, in the description of deep moist convective phenomena. Indeed, advancing the representation and understanding of extreme weather events associated with DMC is of crucial relevance, both from a scientific standpoint and to adopt successful strategies to mitigate their impacts and reduce the associated risks to human and animal health and to the environment. A possible road in this direction is the development of meteorological datasets at increased resolution and sophistication, permitting the enhancement in the representation of the physical atmospheric processes associated with DMC. For this reason, SPHERA, presented in Chapter 2, has been recently developed at ARPAE-SIMC through a dynamical downscaling of the global reanalysis ERA5, driven by the limited-area model COSMO at 0.02°x0.02° horizontal resolution, and continuously assimilating regional observations over 1995-2020. In this thesis, two main topics were developed to separately address the ability of the reanalysis to represent severe precipitation over Italy and to investigate the occurrence of hailstorm events and the associated atmospheric characteristics employing SPHERA in conjunction with multiple observations.

The first analysis, presented in Chapter 3, had the intention to answer the question:

"To what extent do the improvements of CP reanalysis simulations outperform coarser counterparts in representing precipitation and, specifically, severe precipitation events?"

To respond, a statistical analysis has been carried out to compare the simulated precipitation fields of SPHERA with those of its coarser driver ERA5 and evaluate both against the historical series of the Italian rain-gauge network over 2003-2017. Several aspects were taken into account to give a comprehensive picture of the performance of the reanalyses, such as the geographical distribution of daily precipitation on the annual and seasonal terms, the assessment of various aspects of the simulations through quantitative skill scores for daily and hourly rainfalls both for the maximum and average aggregations, and the daily cycle of summer precipitation. Moreover, a detailed analysis was dedicated to two specific extreme precipitation events, with an in-depth dynamical assessment of the atmospheric evolution leading to severe rainfalls, which constitute additional material for a comprehensive comparative evaluation. The results demonstrated a quantitative added value of SPHERA over ERA5 for the description of precipitation, particularly evident in the case of severe precipitation. Indeed, SPHERA has shown: more variable and locally-driven geographical distributions of the 95th percentile of precipitation, higher skill scores at all seasons, especially for the occurrence of moderate- to intense accumulations, both on a daily and hourly frequency, and a magnitude of the daily cycle of summer precipitation more adherent to the observed counterpart. For the two case-study analyses, SPHERA rainfall fields, compared to ERA5, were characterized by more spatial detailing with sharper distributions reaching higher intensities and better spatio-temporal localization of the related maximum accumulations.

The second analysis reported in this thesis and presented in Chapter 4 had the purpose of responding to the question:

"Can the combination of CP-simulated ambient predictors with observational proxies for hail enhance the ability to detect hailstorm events and improve the understanding of the necessary convective environmental conditions for their development?"

Given the major difficulties still persisting for a reliable observation and simulation of hail events, the combination of different data sources is necessary to reduce the uncertainty permeating this hazardous phenomenon. For this purpose, a methodology has been developed relying on the numerical simulations of ambient hail-favoring conditions provided by SPHERA, overshooting top detections from satellite infrared imagery scans associated with thunderstorms presenting strong updrafts, and crowdsourced surface-based hail reports. The method, mainly inspired by previous investigations making use of coarser global reanalyses, consists of the identification of the favorable ambient conditions for the formation of hail, obtained by combining SPHERA predictors with hail reports information, with the intent to define an environmental filter to exclude the overshooting tops not associated with hail. The analysis was performed over 2016-2018 during the extended summer period (April-October) and considering the whole SPHERA spatial domain, which is centered over Italy and includes Switzerland, Austria, southern Germany, Slovenia, Croatia, and the Adriatic and Tyrrhenian seas. The resulting overshooting tops distribution has been tested as a proxy for hail, and the performed analysis indicated a positive outcome in this sense. Indeed, the spatio-temporal characterization of the proxy, indicating a maximum probability for hail occurrence in pre-Alpine areas, especially in northern Italy during June-July around 15 UTC, is in agreement with several hail climatologies over the area obtained with different methods and extending over longer temporal periods. Furthermore, a significant improvement was related to the matching between the hail proxy and surface-based hail reports, with a detected hit rate exceeding 25% more than what was obtained in the previous study taken as the main inspiration (Punge et al., 2017). Finally, from the analysis of the environmental conditions characterizing hail events, different signatures emerged from the separation between hit and miss events and different hail severity classes linked with small and large hailstones. Notably, the strongest separation was evident for the missed-small hail class characterized by systematically less instability and storm organization and lower freezing level altitudes, as well as by generally warmer cloud top temperatures making the associated overshooting tops hardly detectable with the automatic satellite algorithm employed.

The results of this thesis demonstrated promising enhancements in the representation of severe weather events related to DMC using CP simulations, particularly in the case of heavy precipitation and hailstorms, motivating potential future investigations of phenomena not treated here, such as wind gusts, tornadoes, or even heat waves, droughts or local heat island effects. However, there is still room for improvement. Indeed, even if reduced, relevant uncertainties in the numerical simulations are still present at these spatial scales. A major one is the inability to explicitly resolve shallow convection at the CP scales (Khairoutdinov and Randall, 2006), which may also be detrimental in the simulation of deep convection given the intrinsic bond between these two processes (Teixeira et al., 2008). Consequently, CP simulations are highly model-dependent, posing the problem of the robustness of the results. A way to overcome this issue is through multi-model ensemble-based approaches with which it would be possible to “advance parameterizations of unresolved physics and to assess the full potential of CP models” (Prein et al., 2015). In this context, the detailed analysis of specific severe weather events highlighted the necessity of multi-model approaches to reduce the uncertainty and deficiencies in local-scale detailing of deterministic products (Giovannini et al., 2021), as well as the benefits of running high-resolution CP ensembles in this framework (Capecchi, 2021). On a broader perspective, several recent European efforts demonstrated the potential that multi-model CP regional climate simulations have for numerous aspects: better understanding the response of convection extremes to human-induced climate change and providing critical added value to decision-makers due to the enhanced confidence in simulating convection extremes (Coppola et al., 2020), more realistic representation of heavy precipitation with a significant reduction of the hourly summer bias and a reduction of their temporal uncertainty (Ban et al., 2021), and improved representation of fine-scale details of seasonal, daily and hourly heavy rainfalls (Pichelli et al., 2021). Furthermore, ensemble-based approaches are extremely useful to quantify a crucial aspect that can not be estimated with deterministic counterparts alone, i.e. the uncertainty associated with the model driving the simulations. This may shed more light on the reasons producing the differences in the dynamical evolution of the models, that may also indirectly advantage a deeper comprehension of future scenario simulations driven with the models used to produce the reanalyses. In light of this, a possible promising direction to take in future studies would be to combine the efforts leading to the recent development of similar CP reanalysis/hindcast datasets over Italy (such as SPHERA or those developed by Capecchi et al., 2022; Reder et al., 2022), which could reveal of paramount importance to make further progress in improving extreme weather events representation. Hence, the proposal is to jointly develop the first Italian multi-model high-resolution reanalysis/hindcast ensemble. This tool is expected to better assess the uncertainty of past climate, with a particular focus on high-impact convective events such as extreme precipitation and hailstorms. Additionally, it would be possible to quantify the role of data assimilation in SPHERA simulations, given the lack of this component in similar datasets.

To conclude from a broader perspective, numerous European achievements have been recently made to improve our understanding of hydro-meteorological hazards, assess their risk, and mitigate their effects through disaster-risk reduction strategies (Shah et al., 2020). Among them, nature-based solutions play a central role. Compared to traditionally engineered risk-reduction strategies, nature-based solutions are designed to be long-lasting, cost-effective, and environmentally sustainable (Sahani et al., 2019). Several European projects aim at the investigation and enhancement of nature-based solutions, such as OPERANDUM (OPEn-air laboRAtories for Nature baseD solUtions to Manage hydro-meteo risks Mapping, Debele et al., 2019), PHUSICOS ("According to nature" in Greek, Baills et al., 2021), or RECONECT (Regenerating ECOsystems with Nature-based solutions for hydro-meteorological risk rEduCTion, Ruangpan et al., 2020). The effectiveness of nature-based solutions can be improved with a finer knowledge of the small-scale heterogeneities characterizing the spatio-temporal meteorological configurations (Qiu et al., 2019). This is possible through CP simulations able to better describe high-impact weather events, such as severe precipitation and wind gusts associated with DMC activity (Weisman et al., 1997). Finally, applications in climate monitoring can also benefit from high-resolution datasets, which can improve the understanding of climate change impacts on a regional basis, help identify the mechanisms accountable for local climatic features, and add information on the climatology of severe weather-related phenomena.

Appendix A

Additional figures

In the following are reported the remaining bi-variate histograms plots for the joint distributions of the parameters not treated in Section 4.4. In particular, Figure A.1 pertains to the H_0-CAPE distributions, Figure A.2 to H_0-SLI, Figure A.3 to K-DLS, Figure A.4 to SLI-DLS, Figure A.5 to CAPE-SLI, Figure A.6 to K-CAPE, and Figure A.7 to K-SLI.

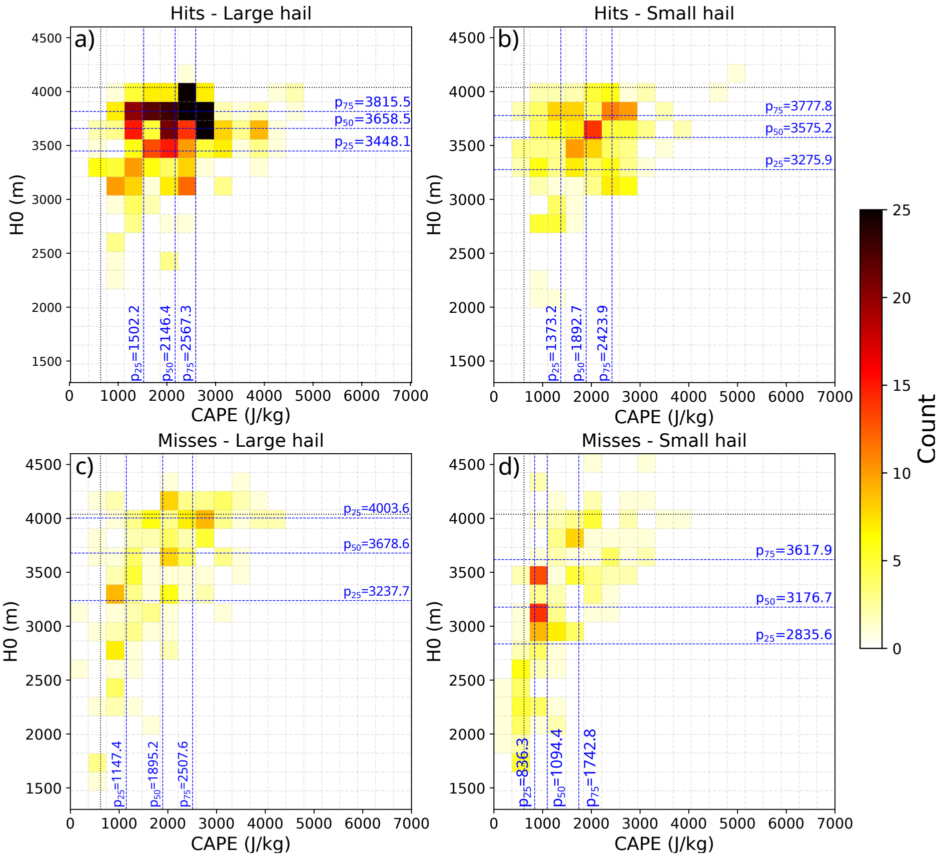

Figure A.1: As Figure 4.13, but for H_0 and CAPE.

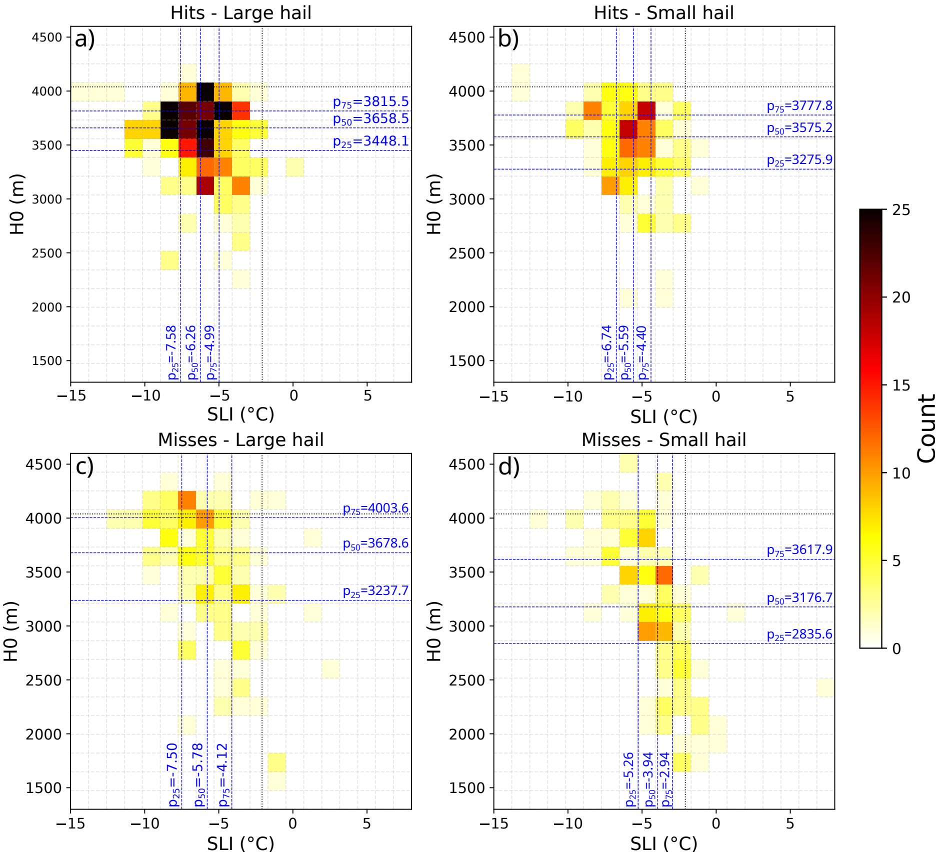

Figure A.2: As Figure 4.13, but for H_0 and SLI.

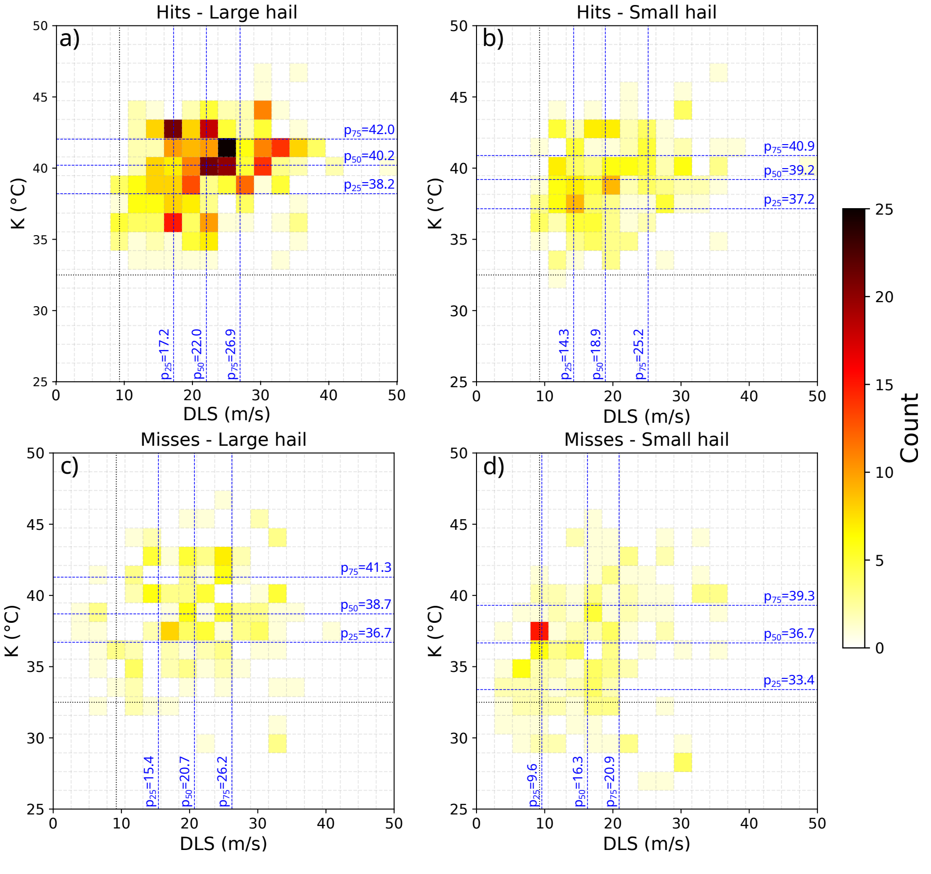

Figure A.3: As Figure 4.13, but for K index and DLS.

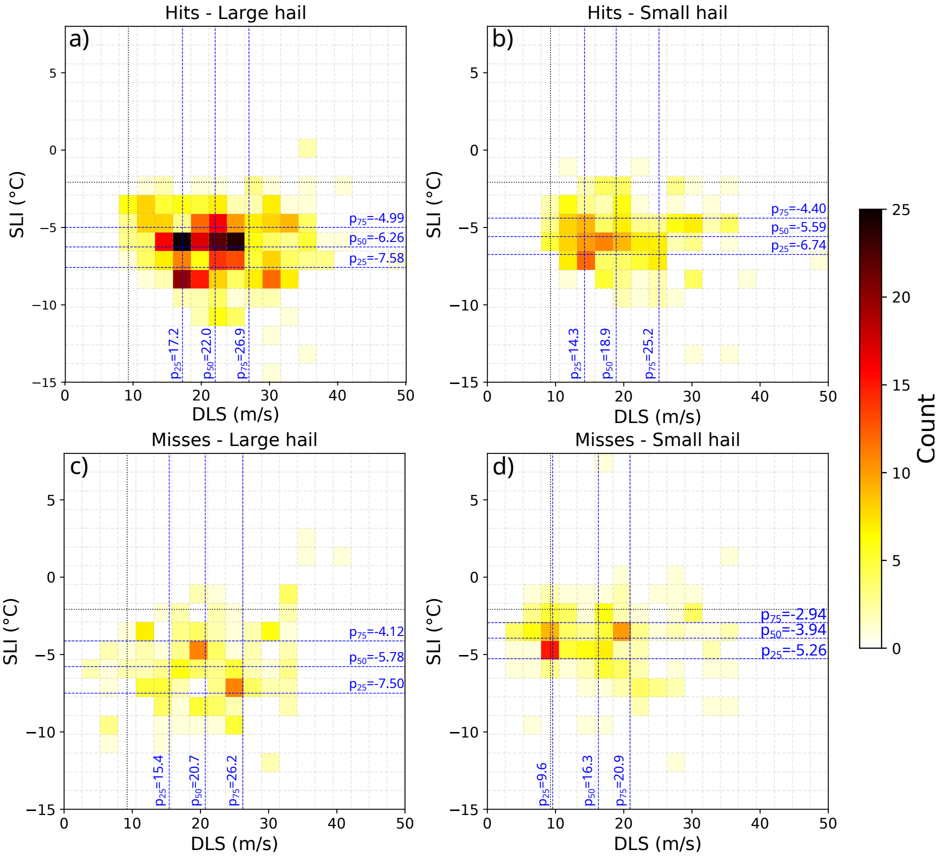

Figure A.4: As Figure 4.13, but for SLI and DLS.

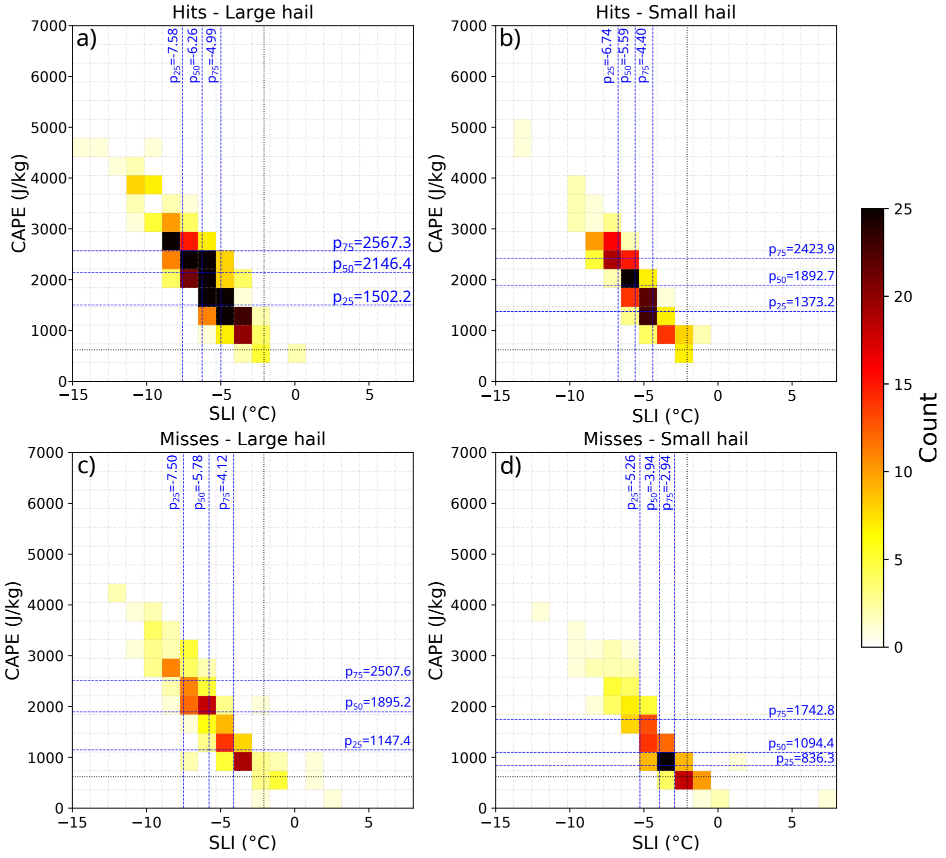

Figure A.5: As Figure 4.13, but for CAPE and SLI.

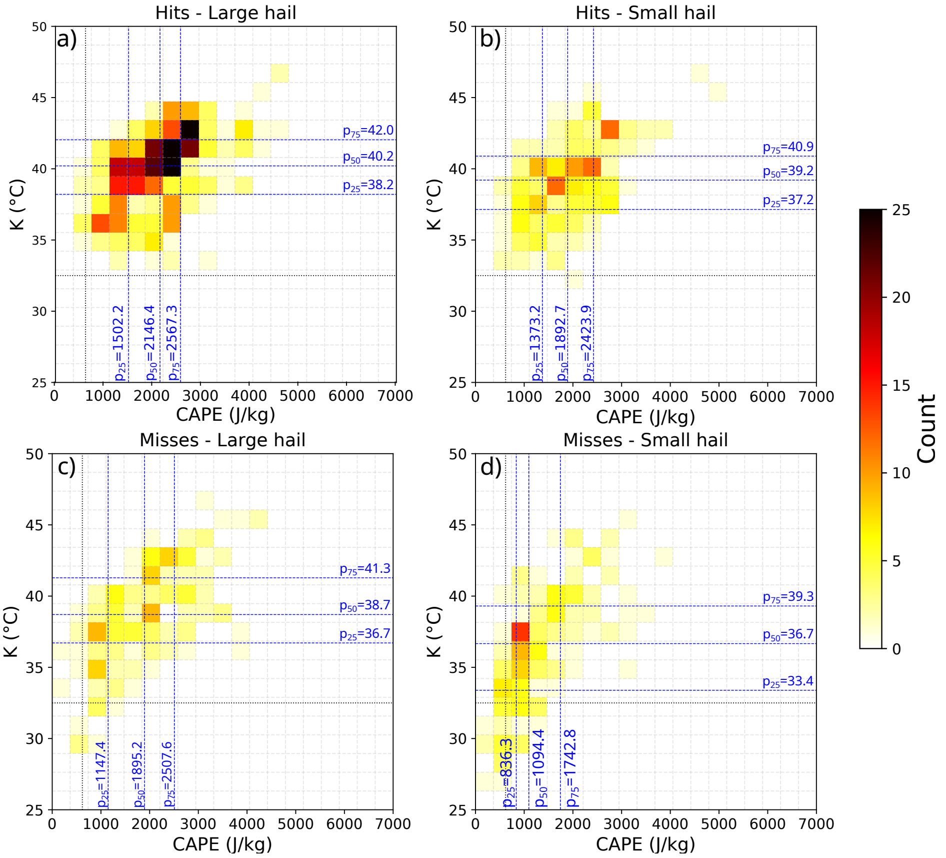

Figure A.6: As Figure 4.13, but for K index and CAPE.

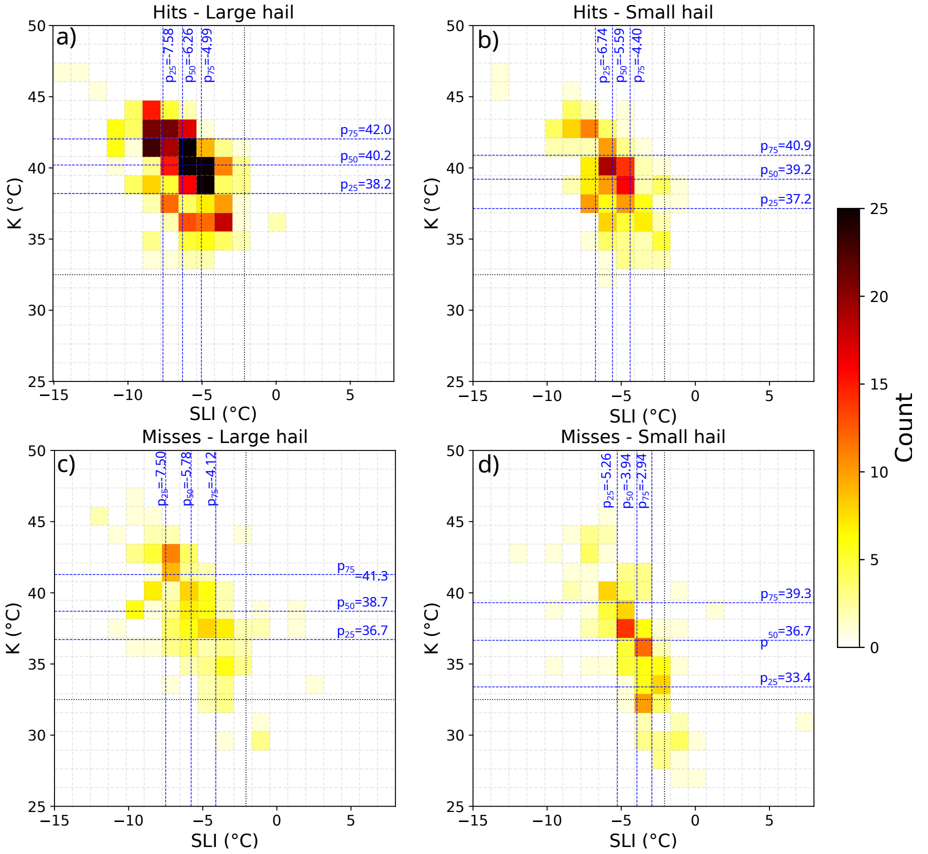

Figure A.7: As Figure 4.13, but for K index and SLI.

Appendix B

Hardware and software

The statistical analysis for precipitation was conducted on the ARPAE machines using internal scripts that are routinely used within the institute and are mainly based on Fortran (gathered in the Libsim library and accessible at the GitHub repository ARPA-SIMC/libsim). To handle the original reanalysis data extracted in grib file format, I employed `ecCodes` and `Magics`, two packages for decoding/encoding the data and visualize them, respectively, that were developed at ECMWF. All the rest of this thesis was thought, computed, and written on my laptop Asus S510U X510UNR, using Manjaro Linux v5.13.19-2. The pre-processing of satellite overshooting top detections data provided by NASA in NetCDF format was partly based on Python scripts developed by Heinz Jürgen Punge. All the rest of the scripts used to perform the analyses and visualize the results were written in Python v3.8.8. In particular, I used Spyder (v4.2.5) and Jupyter Notebook (v6.3.0) to write and run all the code. For data handling and calculations, I made use especially of the packages `pandas`, `geopandas`, `numpy`, and `rpy2`. All the plots were produced with `matplotlib` and `seaborn` and eventually post-processed with the vector graphics software Inkscape v1.2.1. I wrote the thesis in LaTeX using TeXstudio v4.3.1. The scripts developed for the analysis centered on hail are freely accessible at the GitHub repository agiord/hail-analysis.

Bibliography

Ackerman, F., 2017: *Worst-case economics: Extreme events in climate and finance*. Anthem Press.

Adams-Selin, R. D. and C. L. Ziegler, 2016: Forecasting hail using a one-dimensional hail growth model within WRF. *Monthly Weather Review*, **144 (12)**, 4919–4939.

Adler, R. F., M. J. Markus, and D. D. Fenn, 1985: Detection of severe midwest thunderstorms using geosynchronous satellite data. *Monthly weather review*, **113 (5)**, 769–781.

Aguilera, R., A. Gershunov, and T. Benmarhnia, 2019: Atmospheric rivers impact California's coastal water quality via extreme precipitation. *Science of the Total Environment*, **671**, 488–494.

Alfieri, L., P. Burek, L. Feyen, and G. Forzieri, 2015: Global warming increases the frequency of river floods in Europe. *Hydrology and Earth System Sciences*, **19 (5)**, 2247–2260.

Allen, J. T., 2018: Climate change and severe thunderstorms. *Oxford research encyclopedia of climate science*.

Allen, J. T. and E. R. Allen, 2016: A review of severe thunderstorms in Australia. *Atmospheric research*, **178**, 347–366.

Allen, J. T., I. M. Giammanco, M. R. Kumjian, H. Jurgen Punge, Q. Zhang, P. Groenemeijer, M. Kunz, and K. Ortega, 2020: Understanding hail in the Earth system. *Reviews of Geophysics*, **58 (1)**, e2019RG000 665.

Allen, J. T., M. K. Tippett, and A. H. Sobel, 2015: An empirical model relating US monthly hail occurrence to large-scale meteorological environment. *Journal of Advances in Modeling Earth Systems*, **7 (1)**, 226–243.

American Meteorological Society, 2023: Hydrometeor, Glossary of meteorology. `http://glossary.ametsoc.org/wiki/hydrometeor`.

Antolini, G., L. Auteri, V. Pavan, F. Tomei, R. Tomozeiu, and V. Marletto, 2016: A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961–2010. *International Journal of Climatology*, **36 (4)**, 1970–1986.

Arakawa, A., J.-H. Jung, and C.-M. Wu, 2011: Toward unification of the multiscale modeling of the atmosphere. *Atmospheric Chemistry and Physics*, **11 (8)**, 3731–3742.

ARPAE, 2014: Hydro-meteorological report of the event of 17-19 January 2014. Technical report of servizio idrometeoclima, ARPAE, Regional agency for environmental protection of Emilia Romagna.

ARPAS, 2014: Analysis of the weather event of November 18, 2013. Technical report, ARPAS, Regional agency for environmental protection of Sardinia.

Avgoustoglou, E., I. Matsangouras, I. Pytharoulis, N. Kamperakis, M. Mylonas, P. Nastos, and H. Bluestein, 2018: Numerical modeling analysis of the mesoscale environment conducive to two tornado events using the COSMO. Gr model over Greece. *Atmospheric Research*, **208**, 148–155.

Baills, A., M. Garcin, and S. Bernardie, 2021: Platform dedicated to nature-based solutions for risk reduction and environmental issues in hilly and mountainous lands. *Sustainability*, **13 (3)**, 1094.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the cosmo model: Description and sensitivities. *Monthly Weather Review*, **139 (12)**, 3887–3905.

Baldi, M., V. Ciardini, J. D. Dalu, T. De Filippis, G. Maracchi, and G. Dalu, 2014: Hail occurrence in Italy: Towards a national database and climatology. *Atmospheric research*, **138**, 268–277.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres*, **119 (13)**, 7889–7907.

Ban, N., et al., 2021: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part i: evaluation of precipitation. *Climate Dynamics*, **57 (1)**, 275–302.

Bandhauer, M., et al., 2022: Evaluation of daily precipitation analyses in E-OBS (v19. 0e) and ERA5 by comparison to regional high-resolution datasets in European regions. *International Journal of Climatology*, **42 (2)**, 727–747.

Barras, H., A. Hering, A. Martynov, P.-A. Noti, U. Germann, and O. Martius, 2019: Experiences with> 50,000 crowdsourced hail reports in Switzerland. *Bulletin of the American Meteorological Society*, **100 (8)**, 1429–1440.

Bedka, K., J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, 2010: Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients. *Journal of applied meteorology and climatology*, **49 (2)**, 181–202.

Bedka, K. M., 2011: Overshooting cloud top detections using MSG SEVIRI Infrared brightness temperatures and their relationship to severe weather over Europe. *Atmospheric Research*, **99 (2)**, 175–189.

Bedka, K. M., J. T. Allen, H. J. Punge, M. Kunz, and D. Simanovic, 2018: A long-term overshooting convective cloud-top detection database over Australia derived from MTSAT Japanese advanced meteorological imager observations. *Journal of Applied Meteorology and Climatology*, **57 (4)**, 937–951.

Bedka, K. M. and K. Khlopenkov, 2016: A probabilistic multispectral pattern recognition method for detection of overshooting cloud tops using passive satellite imager observations. *Journal of Applied Meteorology and Climatology*, **55 (9)**, 1983–2005.

Berg, P., S. Wagner, H. Kunstmann, and G. Schädler, 2013: High resolution regional climate model simulations for Germany: Part I — validation. *Climate dynamics*, **40 (1)**, 401–414.

Bhatti, H. A., T. Rientjes, A. T. Haile, E. Habib, and W. Verhoef, 2016: Evaluation of bias correction method for satellite-based rainfall data. *Sensors*, **16 (6)**, 884.

Bhutta, Z. A., S. Z. Bhutta, S. Raza, and A. T. Sheikh, 2022: Addressing the human costs and consequences of the Pakistan flood disaster. *The Lancet*, **400 (10360)**, 1287–1289.

Bica, B., et al., 2007: Thermally and dynamically induced pressure features over complex terrain from high-resolution analyses. *Journal of applied meteorology and climatology*, **46 (1)**, 50–65.

Biron, D., 2009: LAMPINET-lightning detection in italy. *Lightning: Principles, Instruments and Applications*, Springer, 141–159.

Blair, S. F., et al., 2017: High-resolution hail observations: Implications for NWS warning operations. *Weather and Forecasting*, **32 (3)**, 1101–1119.

Bližňák, V., M. Kašpar, M. Müller, and P. Zacharov, 2021: Analysis and verification of reconstructed historical extreme precipitation events in an hourly resolution. *Atmospheric Research*, **249**, 105 309.

Bollmeyer, C., et al., 2015: Towards a high-resolution regional reanalysis for the European CORDEX domain. *Quarterly Journal of the Royal Meteorological Society*, **141 (686)**, 1–15.

Bonanno, R., M. Lacavalla, and S. Sperati, 2019: A new high-resolution Meteorological Reanalysis Italian Dataset: MERIDA. *Quarterly Journal of the Royal Meteorological Society*, **145 (721)**, 1756–1779.

Brimelow, J. C., W. R. Burrows, and J. M. Hanesiak, 2017: The changing hail threat over North America in response to anthropogenic climate change. *Nature Climate Change*, **7 (7)**, 516–522.

Brisson, E., K. Van Weverberg, M. Demuzere, A. Devis, S. Saeed, M. Stengel, and N. P. van Lipzig, 2016: How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics? *Climate Dynamics*, **47 (9)**, 3043–3061.

Bromwich, D., et al., 2018: The arctic system reanalysis, version 2. *Bulletin of the American Meteorological Society*, **99 (4)**, 805–828.

Brooks, H. E., 2009: Proximity soundings for severe convection for Europe and the United States from reanalysis data. *Atmospheric Research*, **93 (1-3)**, 546–553.

Brooks, H. E., 2013: Severe thunderstorms and climate change. *Atmospheric research*, **123**, 129–138.

Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmospheric Research*, **67**, 73–94.

Bruick, Z. S., K. L. Rasmussen, and D. J. Cecil, 2019: Subtropical South American hailstorm characteristics and environments. *Monthly weather review*, **147 (12)**, 4289–4304.

Brunner, J. C., S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin, 2007: A quantitative analysis of the enhanced-V feature in relation to severe weather. *Weather and Forecasting*, **22 (4)**, 853–872.

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, **131 (10)**, 2394–2416.

Buzzi, A., S. Davolio, P. Malguzzi, O. Drofa, and D. Mastrangelo, 2014: Heavy rainfall episodes over Liguria in autumn 2011: numerical forecasting experiments. *Natural Hazards and Earth System Sciences*, **14 (5)**, 1325–1340.

Caillaud, C., S. Somot, A. Alias, I. Bernard-Bouissières, Q. Fumière, O. Laurantin, Y. Seity, and V. Ducrocq, 2021: Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model. *Climate Dynamics*, **56 (5)**, 1717–1752.

Calianese, E. J., J. K. Jordan, E. B. Curran, A. R. Moller, and G. Woodall, 2002: The Mayfest high-precipitation supercell of 5 May 1995—a case study. *Preprints, 21st Conf. on Severe Local Storms, San Antonio, TX, Amer. Meteor. Soc*, Vol. 105, 108.

Calvo Sánchez, F. J., et al., 2021: Prototype of a high resolution regional reanalysis for Iberian peninsula an Balearic islands (IBERA). *Poster - 1st ACCORD All Staff Workshop, 12-16 April 2021.*

Capecchi, V., 2021: Reforecasting two heavy-precipitation events with three convection-permitting ensembles. *Weather and Forecasting*, **36 (3)**, 769–790.

Capecchi, V., F. Pasi, B. Gozzini, and C. Brandini, 2022: A convection-permitting and limited-area model hindcast driven by ERA5 data: precipitation performances in Italy. *Climate Dynamics*, 1–27.

Cassola, F., F. Ferrari, and A. Mazzino, 2015: Numerical simulations of Mediterranean heavy precipitation events with the WRF model: A verification exercise using different approaches. *Atmospheric Research*, **164**, 210–225.

Castro, C. L., R. A. Pielke Sr, and G. Leoncini, 2005: Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *Journal of Geophysical Research: Atmospheres*, **110 (D5)**.

Cecil, D. J., 2009: Passive microwave brightness temperatures as proxies for hailstorms. *Journal of applied meteorology and climatology*, **48 (6)**, 1281–1286.

Cecil, D. J. and C. B. Blankenship, 2012: Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *Journal of Climate*, **25 (2)**, 687–703.

Cerenzia, I., et al., 2020: Forecast of precipitation for the 1994 flood in Piedmont: performance of an ensemble system at convection-permitting resolution. *Bulletin of Atmospheric Science and Technology*, **1 (3)**, 319–338.

Cerenzia, I. M. L., A. Giordani, T. Paccagnella, and A. Montani, 2022: Towards a convection-permitting regional reanalysis over the Italian domain. *Meteorological Applications*, **29 (5)**, e2092.

Cerveny, R. S., et al., 2017: WMO assessment of weather and climate mortality extremes: lightning, tropical cyclones, tornadoes, and hail. *Weather, climate, and society*, **9 (3)**, 487–497.

Changnon, S. A., 1977: The scales of hail. *Journal of Applied Meteorology and Climatology*, **16 (6)**, 626–648.

Changnon, S. A., 2009: Increasing major hail losses in the US. *Climatic Change*, **96 (1)**, 161–166.

Changnon, S. A. and J. Burroughs, 2003: The tristate hailstorm: The most costly on record. *Monthly Weather Review*, **131 (8)**, 1734–1739.

Changnon Jr, S. A., 1970: Hailstreaks. *Journal of Atmospheric Sciences*, **27 (1)**, 109–125.

Chen, J., A. Dai, Y. Zhang, and K. L. Rasmussen, 2020: Changes in convective available potential energy and convective inhibition under global warming. *Journal of Climate*, **33 (6)**, 2025–2050.

Chen, Y., A. Paschalis, E. Kendon, D. Kim, and C. Onof, 2021: Changing spatial structure of summer heavy rainfall, using convection-permitting ensemble. *Geophysical Research Letters*, **48 (3)**, e2020GL090 903.

Chernokulsky, A., F. Kozlov, O. Zolina, O. Bulygina, I. I. Mokhov, and V. A. Semenov, 2019: Observed changes in convective and stratiform precipitation in northern Eurasia over the last five decades. *Environmental Research Letters*, **14 (4)**, 045 001.

Clark, A. J., J. Gao, P. T. Marsh, T. Smith, J. S. Kain, J. Correia Jr, M. Xue, and F. Kong, 2013: Tornado pathlength forecasts from 2010 to 2011 using ensemble updraft helicity. *Weather and Forecasting*, **28 (2)**, 387–407.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: a step-change in rainfall forecasting. *Meteorological Applications*, **23 (2)**, 165–181.

Cooney, J. W., K. M. Bedka, K. P. Bowman, K. V. Khlopenkov, and K. Itterly, 2021: Comparing tropopause-penetrating convection identifications derived from NEXRAD and GOES over the contiguous United States. *Journal of Geophysical Research: Atmospheres*, **126 (14)**, e2020JD034 319.

Coppola, E., et al., 2020: A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. *Climate Dynamics*, **55 (1)**, 3–34.

Crespi, A., M. Brunetti, G. Lentini, and M. Maugeri, 2018: 1961–1990 high-resolution monthly precipitation climatologies for italy. *International Journal of Climatology*, **38 (2)**, 878–895.

Dalagnol, R., et al., 2022: Extreme rainfall and its impacts in the Brazilian Minas Gerais state in January 2020: Can we blame climate change? *Climate Resilience and Sustainability*, **1 (1)**, e15.

Dale, M., B. Luck, H. J. Fowler, S. Blenkinsop, E. Gill, J. Bennett, E. Kendon, and S. Chan, 2015: New climate change rainfall estimates for sustainable drainage. *Proceedings of the Institution of Civil Engineers-Engineering Sustainability*, Thomas Telford Ltd, Vol. 170, 214–224.

Davini, P., R. Bechini, R. Cremonini, and C. Cassardo, 2012: Radar-based analysis of convective storms over northwestern Italy. *Atmosphere*, **3 (1)**, 33–58.

Davolio, S., F. Silvestro, and P. Malguzzi, 2015: Effects of increasing horizontal resolution in a convection-permitting model on flood forecasting: The 2011 dramatic events in liguria, italy. *Journal of Hydrometeorology*, **16 (4)**, 1843–1856.

de Coning, E., M. Gijben, B. Maseko, and L. van Hemert, 2015: Using satellite data to identify and track intense thunderstorms in South and southern Africa. *South African Journal of Science*, **111 (7-8)**, 1–5.

Debele, S. E., et al., 2019: Nature-based solutions for hydro-meteorological hazards: Critical evaluation of risks and opportunities for OPERANDUM OALs. *OPERANDUM Deliverable 1.1*, URL `https://doi.org/10.13140/RG.2.2.29763.43048`.

Dennis, E. J. and M. R. Kumjian, 2017: The impact of vertical wind shear on hail growth in simulated supercells. *Journal of the Atmospheric Sciences*, **74 (3)**, 641–663.

Desamsetti, S., H. P. Dasari, S. Langodan, E. S. Titi, O. Knio, and I. Hoteit, 2019: Efficient dynamical downscaling of general circulation models using continuous data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **145 (724)**, 3175–3194.

Dessens, J., C. Berthet, and J. Sanchez, 2015: Change in hailstone size distributions with an increase in the melting level height. *Atmospheric Research*, **158**, 245–253.

Doms, G., et al., 2018: A description of the nonhydrostatic regional COSMO-model. Part II: physical parameterizations. *Deutscher Wetterdienst Rep. COSMO-Model*, 167.

Doswell, C. A., 2001: Severe convective storms—an overview. *Severe convective storms*, 1–26.

Dotzek, N., P. Groenemeijer, B. Feuerstein, and A. M. Holzer, 2009: Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. *Atmospheric research*, **93 (1-3)**, 575–586.

Dudhia, J., 2014: A history of mesoscale model development. *Asia-Pacific Journal of Atmospheric Sciences*, **50 (1)**, 121–131.

Dumitrache, R., L. Velea, C. Barbu, I. Ibanescu, and A. Lupascu, 2011: Preliminary results of COSMO model forecast for the Romanian territory case studies. *Romanian Reports in Physics*, **63 (1)**, 208–219.

D'Alpaos, L., A. Brath, V. Fioravante, G. Gottardi, P. Mignosa, and S. Orlandini, 2014: Relazione tecnico-scientifica sulle cause del collasso dell'argine del fiume Secchia avvenuto il giorno 19 gennaio 2014 presso la frazione San Matteo. *Report of the Emilia-Romagna Region.*

Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, **15 (1)**, 51–64.

Eccel, E., P. Cau, K. Riemann-Campe, and F. Biasioli, 2012: Quantitative hail monitoring in an alpine area: 35-year climatology and links with atmospheric variables. *International journal of climatology*, **32 (4)**, 503–517.

Eckstein, D., V. Künzel, and L. Schäfer, 2021: Global climate risk index 2021. *Who Suffers Most from Extreme Weather Events*, 2000–2019.

El-Said, A., P. Brousseau, M. Ridal, and R. Randriamampianina, 2022: A new temporally flow-dependent EDA estimating background errors in the new Copernicus European Regional Re-Analysis (CERRA). *Authorea Preprints.*

Elliott, M. S., D. R. MacGorman, T. J. Schuur, and P. L. Heinselman, 2012: An analysis of overshooting top lightning mapping array signatures in supercell thunderstorms. *Proc. 22nd Int. Lightning Detection Conf.*

Emrich, C. T., S. Alvarez, C. C. Knox, A. A. Sadiq, and Y. Zhou, 2019: Hurricane Irma and cascading impacts. *US Emergency Management in the 21st Century*, Routledge, 123–153.

Falconi, M. T. and F. S. Marzano, 2019: Weather radar data processing and atmospheric applications: An overview of tools for monitoring clouds and detecting wind shear. *IEEE Signal Processing Magazine*, **36 (4)**, 85–97.

Fantini, A., 2019: Climate change impact on flood hazard over Italy. *PhD Thesis, Università degli Studi di Trieste.*

Farnell, C., T. Rigo, and N. Pineda, 2017: Lightning jump as a nowcast predictor: Application to severe weather events in Catalonia. *Atmospheric Research*, **183**, 130–141.

Farnell, C., T. Rigo, and N. Pineda, 2018: Exploring radar and lightning variables associated with the Lightning Jump. Can we predict the size of the hail? *Atmospheric Research*, **202**, 175–186.

Feng, T., Z. Hu, S. Tang, and J. Huang, 2021: Improvement of an extreme heavy rainfall simulation using nudging assimilation. *Journal of Meteorological Research*, **35 (2)**, 313–328.

Fluck, E., M. Kunz, P. Geissbuehler, and S. P. Ritz, 2021: Radar-based assessment of hail frequency in Europe. *Natural Hazards and Earth System Sciences*, **21 (2)**, 683–701.

Forzieri, G., et al., 2016: Multi-hazard assessment in Europe under climate change. *Climatic Change*, **137 (1)**, 105–119.

Fosser, G., S. Khodayar, and P. Berg, 2015: Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Climate Dynamics*, **44 (1)**, 45–60.

Fujita, T. T., 1992: *Memoirs of an Effort to Unlock the Mystery of Severe Storms: During the 50 Years, 1942-1992*. University of Chicago.

Fukushima, T., T. Kitamura, and B. Matsushita, 2021: Lake water quality observed after extreme rainfall events: implications for water quality affected by stormy runoff. *SN Applied Sciences*, **3 (11)**, 1–15.

Gabriele, S., F. Chiaravalloti, and A. Procopio, 2017: Radar–rain-gauge rainfall estimation for hydrological applications in small catchments. *Advances in Geosciences*, **44**, 61–66.

Gagne, D. J., A. McGovern, S. E. Haupt, R. A. Sobash, J. K. Williams, and M. Xue, 2017: Storm-based probabilistic hail forecasting with machine learning applied to convection-allowing ensembles. *Weather and forecasting*, **32 (5)**, 1819–1840.

Gagne II, D. J., S. E. Haupt, D. W. Nychka, and G. Thompson, 2019: Interpretable deep learning for spatial analysis of severe hailstorms. *Monthly Weather Review*, **147 (8)**, 2827–2845.

Gal-Chen, T. and R. C. Somerville, 1975: On the use of a coordinate transformation for the solution of the Navier-Stokes equations. *Journal of Computational Physics*, **17 (2)**, 209–228.

Gallus Jr, W. A., 2002: Impact of verification grid-box size on warm-season QPF skill measures. *Weather and forecasting*, **17 (6)**, 1296–1302.

Gascón, E., A. Merino, J. Sánchez, S. Fernández-González, E. García-Ortega, L. López, and L. Hermida, 2015: Spatial distribution of thermodynamic conditions of severe storms in southwestern Europe. *Atmospheric Research*, **164**, 194–209.

Gensini, V. A. and T. L. Mote, 2014: Estimations of hazardous convective weather in the United States using dynamical downscaling. *Journal of Climate*, **27 (17)**, 6581–6589.

Gensini, V. V. A., 2021: Severe convective storms in a changing climate. *Climate Change and Extreme Events*, Elsevier, 39–56.

Giaiotti, D., S. Nordio, and F. Stel, 2003: The climatology of hail in the plain of Friuli Venezia Giulia. *Atmospheric Research*, **67**, 247–259.

Gimeno, L., R. Sorí, M. Vázquez, M. Stojanovic, I. Algarra, J. Eiras-Barca, L. Gimeno-Sotelo, and R. Nieto, 2022: Extreme precipitation events. *Wiley Interdisciplinary Reviews: Water*, **9 (6)**, e1611.

Giordani, A., I. M. L. Cerenzia, T. Paccagnella, and S. Di Sabatino, 2023: SPHERA, a new convection-permitting regional reanalysis over Italy: Improving the description of heavy rainfall. *Quarterly Journal of the Royal Meteorological Society*, **149**, 781–808, doi:10.1002/qj.4428.

Giorgi, F., 1990: Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate*, **3 (9)**, 941–963.

Giovannini, L., S. Davolio, M. Zaramella, D. Zardi, and M. Borga, 2021: Multimodel convection-resolving simulations of the October 2018 Vaia storm over Northeastern Italy. *Atmospheric Research*, **253**, 105 455.

Gleeson, E., E. Whelan, and J. Hanley, 2017: Met Éireann high resolution reanalysis for Ireland. *Advances in Science and Research*, **14**, 49–61.

Gobbo, S., A. Ghiraldini, A. Dramis, N. Dal Ferro, and F. Morari, 2021: Estimation of hail damage using crop models and remote sensing. *Remote Sensing*, **13 (14)**, 2655.

Goudenhoofdt, E. and L. Delobbe, 2009: Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, **13 (2)**, 195–203, doi:10.5194/hess-13-195-2009, URL `https://hess.copernicus.org/articles/13/195/2009/`.

Grahame, N., B. Riddaway, A. Eadie, B. Hall, and E. McCallum, 2009: Exceptional hailstorm hits Ottery St Mary on 30 October 2008. *Weather*, **64 (10)**, 255–263.

Grazzini, F., G. C. Craig, C. Keil, G. Antolini, and V. Pavan, 2020a: Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques. *Quarterly Journal of the Royal Meteorological Society*, **146 (726)**, 69–85.

Grazzini, F., G. Fragkoulidis, V. Pavan, and G. Antolini, 2020b: The 1994 Piedmont flood: an archetype of extreme precipitation events in Northern Italy. *Bulletin of Atmospheric Science and Technology*, **1 (3)**, 283–295.

Griffin, S. M., K. M. Bedka, and C. S. Velden, 2016: A method for calculating the height of overshooting convective cloud tops using satellite-based IR imager and CloudSat cloud profiling radar observations. *Journal of Applied Meteorology and Climatology*, **55 (2)**, 479–491.

Groenemeijer, P., et al., 2017: Severe convective storms in Europe: Ten years of research and education at the European Severe Storms Laboratory. *Bulletin of the American Meteorological Society*, **98 (12)**, 2641–2651.

Guerreiro, S. B., R. J. Dawson, C. Kilsby, E. Lewis, and A. Ford, 2018: Future heat-waves, droughts and floods in 571 European cities. *Environmental Research Letters*, **13 (3)**, 034 009.

Gunturi, P. and M. Tippett, 2017: Managing severe thunderstorm risk: Impact of ENSO on US tornado and hail frequencies. *Technical Rep. Minneapolis: WillisRe.*

Gustafsson, N., et al., 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 1218–1256.

Han, X., H. Xue, C. Zhao, and D. Lu, 2016: The roles of convective and stratiform precipitation in the observed precipitation trends in Northwest China during 1961–2000. *Atmospheric Research*, **169**, 139–146.

Hastings, D. A., et al., 1999: The global land one-kilometer base elevation (GLOBE) digital elevation model, version 1.0. *National Oceanic and Atmospheric Administration, National Geophysical Data Center*, **325**, 80 305–3328, URL http://www.ngdc.noaa.gov/mgg/topo/globe.html.

Herold, N., A. Behrangi, and L. V. Alexander, 2017: Large uncertainties in observed daily precipitation extremes over land. *Journal of Geophysical Research: Atmospheres*, **122 (2)**, 668–681.

Herring, S. C., N. Christidis, A. Hoell, M. P. Hoerling, and P. A. Stott, 2021: Explaining extreme events of 2019 from a climate perspective. *Bulletin of the American Meteorological Society*, **102 (1)**, S1–S115.

Hersbach, H., et al., 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049.

Hitchens, N. M. and H. E. Brooks, 2014: Evaluation of the Storm Prediction Center's convective outlooks from day 3 through day 1. *Weather and Forecasting*, **29 (5)**, 1134–1142.

Hochman, A., P. Mercogliano, P. Alpert, H. Saaroni, and E. Bucchignani, 2018: High-resolution projection of climate change and extremity over israel using COSMO-CLM. *International Journal of Climatology*, **38 (14)**, 5095–5106.

Hohl, R., H.-H. Schiesser, and I. Knepper, 2002: The use of weather radars to estimate hail damage to automobiles: an exploratory study in Switzerland. *Atmospheric research*, **61 (3)**, 215–238.

Höller, H. and M. Reinhardt, 1986: The Munich hailstorm of july 12, 1984-convective development and preliminary hailstone analysis. *Beiträge zur Physik der Atmosphäre*, **59**, 1–12.

Holzer, A. M., P. Groenemeijer, K. Riemann-Campe, and B. Antonescu, 2017: Experience after 1 year of EWOB. *Proc., Ninth European Conf. on Severe Storms (ECSS 2017), Pula, Croatia, https://presentations. copernicus. org/ECSS2017/ECSS2017-188_presentation. pdf.*

Hong, Y., R. J. Trapp, S. W. Nesbitt, and L. Di Girolamo, 2022: Near global distributions of overshooting tops derived from Terra and Aqua MODIS observations. *Atmospheric Measurement Techniques Discussions*, 1–29.

Hoogewind, K. A., M. E. Baldwin, and R. J. Trapp, 2017: The impact of climate change on hazardous convective weather in the United States: Insight from high-resolution dynamical downscaling. *Journal of Climate*, **30 (24)**, 10 081–10 100.

Houze Jr, R. A., 2014: *Cloud dynamics*. Academic press.

Hu, G. and C. L. Franzke, 2020: Evaluation of daily precipitation extremes in reanalysis and gridded observation-based data sets over Germany. *Geophysical Research Letters*, **47 (18)**, e2020GL089 624.

Isotta, F. A., R. Vogel, and C. Frei, 2015: Evaluation of European regional reanalyses and downscalings for precipitation in the alpine region. *Meteorologische Zeitschrift*, **24**, 15–37.

Italian Civil Protection Department & CIMA Research Foundation, 2014: The Dewetra platform: a multi-perspective architecture for risk management during emergencies. *Information Systems for Crisis Response and Management in Mediterranean Countries: First International Conference, ISCRAM-med 2014, Toulouse, France, October 15-17, 2014. Proceedings 1*, Springer, 165–177.

Jacobsen, I. et al., 1982: A new economic method for the computation of the surface temperature in numerical models. *Beitr. Phys. Atmos.*, **55**, 128–141.

Jansa, A., A. Genoves, and J. A. Garcia-Moya, 2000: Western Mediterranean cyclones and heavy rain. Part 1: Numerical experiment concerning the Piedmont flood case. *Meteorological Applications*, **7 (4)**, 323–333.

Jansa, A., A. Genoves, M. A. Picornell, J. Campins, R. Riosalido, and O. Carretero, 2001: Western Mediterranean cyclones and heavy rain. Part 2: Statistical approach. *Meteorological Applications*, **8 (1)**, 43–56.

Jelić, D., O. A. Megyeri, B. Malečić, A. Belušić Vozila, N. Strelec Mahović, and M. Telišman Prtenjak, 2020: Hail climatology along the northeastern Adriatic. *Journal of Geophysical Research: Atmospheres*, **125 (23)**, e2020JD032 749.

Jermey, P. and R. Renshaw, 2016: Precipitation representation over a two-year period in regional reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **142 (696)**, 1300–1310.

Jiang, Y., K. Yang, C. Shao, X. Zhou, L. Zhao, Y. Chen, and H. Wu, 2021: A downscaling approach for constructing high-resolution precipitation dataset over the Tibetan Plateau from ERA5 reanalysis. *Atmospheric Research*, **256**, 105 574.

Johns, R. H. and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7 (4)**, 588–612.

Johnson, A. W. and K. E. Sugden, 2014: Evaluation of sounding-derived thermodynamic and wind-related parameters associated with large hail events. *E-Journal of Severe Storms Meteorology*, **9 (5)**.

Jones, T. R. and D. A. Randall, 2011: Quantifying the limits of convective parameterizations. *Journal of Geophysical Research: Atmospheres*, **116 (D8)**.

Jurković, P. M., N. S. Mahović, and D. Počakal, 2015: Lightning, overshooting top and hail characteristics for strong convective storms in Central Europe. *Atmospheric Research*, **161**, 153–168.

Kalina, E. A., K. Friedrich, B. C. Motta, W. Deierling, G. T. Stano, and N. N. Rydell, 2016: Colorado plowable hailstorms: Synoptic weather, radar, and lightning characteristics. *Weather and Forecasting*, **31 (2)**, 663–693.

Khairoutdinov, M. and D. Randall, 2006: High-resolution simulation of shallow-to-deep convection transition over land. *Journal of the atmospheric sciences*, **63 (12)**, 3421–3436.

Khlopenkov, K. V., K. M. Bedka, J. W. Cooney, and K. Itterly, 2021: Recent advances in detection of overshooting cloud tops from longwave infrared satellite imagery. *Journal of Geophysical Research: Atmospheres*, **126 (14)**, e2020JD034 359.

Khodayar, S., et al., 2021: Overview towards improved understanding of the mechanisms leading to heavy precipitation in the western mediterranean: lessons learned from hymex. *Atmospheric Chemistry and Physics*, **21 (22)**, 17 051–17 078.

Kirshbaum, D. J., B. Adler, N. Kalthoff, C. Barthlott, and S. Serafin, 2018: Moist orographic convection: Physical mechanisms and links to surface-exchange processes. *Atmosphere*, **9 (3)**, 80.

Klasa, C., M. Arpagaus, A. Walser, and H. Wernli, 2018: An evaluation of the convection-permitting ensemble COSMO-E for three contrasting precipitation events in Switzerland. *Quarterly Journal of the Royal Meteorological Society*, **144 (712)**, 744–764.

Klein Tank, A., 2010: EURO4M: Monitoring weather and climate extremes in Europe. *10th EMS Annual Meeting*, EMS2010–202.

Knight, C. A. and N. C. Knight, 2001: Hailstorms. *Severe convective storms*, Springer, 223–254.

Køltzow, M., H. Schyberg, E. Støylen, and X. Yang, 2022: Value of the Copernicus Arctic Regional Reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic. *Polar Research*, **41**.

Kopp, J., K. Schröer, C. Schwierz, A. Hering, U. Germann, and O. Martius, 2022: The summer 2021 switzerland hailstorms: weather situation, major impacts and unique observational data. *Weather*.

Krichak, S. O., J. Barkan, J. S. Breitgand, S. Gualdi, and S. B. Feldstein, 2015: The role of the export of tropical moisture into midlatitudes for extreme precipitation events in the Mediterranean region. *Theoretical and Applied Climatology*, **121 (3)**, 499–515.

Kumjian, M. R., Z. J. Lebo, and A. M. Ward, 2019: Storms producing large accumulations of small hail. *Journal of Applied Meteorology and Climatology*, **58 (2)**, 341–364.

Kunz, M., 2007: The skill of convective parameters and indices to predict isolated and severe thunderstorms. *Natural Hazards and Earth System Sciences*, **7 (2)**, 327–342.

Kunz, M., U. Blahak, J. Handwerker, M. Schmidberger, H. J. Punge, S. Mohr, E. Fluck, and K. M. Bedka, 2018: The severe hailstorm in southwest Germany on 28 July 2013: Characteristics, impacts and meteorological conditions. *Quarterly Journal of the Royal Meteorological Society*, **144 (710)**, 231–250.

Kunz, M., J. Wandel, E. Fluck, S. Baumstark, S. Mohr, and S. Schemm, 2020: Ambient conditions prevailing during hail events in central Europe. *Natural Hazards and Earth System Sciences*, **20 (6)**, 1867–1887.

Labriola, J., N. Snook, Y. Jung, and M. Xue, 2019: Explicit ensemble prediction of hail in 19 May 2013 Oklahoma City thunderstorms and analysis of hail growth processes with several multimoment microphysics schemes. *Monthly Weather Review*, **147 (4)**, 1193–1213.

Lanciani, A., S. Mariani, M. Casaioli, C. Accadia, and N. Tartaglione, 2008: A multiscale approach for precipitation verification applied to the FORALPS case studies. *Advances in Geosciences*, **16**, 3–9.

Lavers, D. A. and G. Villarini, 2015: The relationship between daily European precipitation and measures of atmospheric water vapour transport. *International Journal of Climatology*, **35 (8)**, 2187–2192.

Laviola, S., V. Levizzani, R. R. Ferraro, and J. Beauchamp, 2020: Hailstorm detection by satellite microwave radiometers. *Remote Sensing*, **12 (4)**, 621.

Leutwyler, D., D. Lüthi, N. Ban, O. Fuhrer, and C. Schär, 2017: Evaluation of the convection-resolving climate modeling approach on continental scales. *Journal of Geophysical Research: Atmospheres*, **122 (10)**, 5237–5258.

Lewis, H., et al., 2015: From months to minutes–exploring the value of high-resolution rainfall observation and prediction during the UK winter storms of 2013/2014. *Meteorological Applications*, **22 (1)**, 90–104.

Li, M., D.-L. Zhang, J. Sun, and Q. Zhang, 2018a: A statistical analysis of hail events and their environmental conditions in China during 2008–15. *Journal of Applied Meteorology and Climatology*, **57 (12)**, 2817–2833.

Li, X., Q. Zhang, T. Zou, J. Lin, H. Kong, and Z. Ren, 2018b: Climatology of hail frequency and size in China, 1980–2015. *Journal of Applied Meteorology and Climatology*, **57 (4)**, 875–887.

Liu, C., et al., 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, **49 (1)**, 71–95.

Lobligeois, F., V. Andréassian, C. Perrin, P. Tabary, and C. Loumagne, 2014: When does higher spatial resolution rainfall information improve streamflow simulation? an evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, **18 (2)**, 575–594.

Löffler-Mang, M., D. Schön, and M. Landry, 2011: Characteristics of a new automatic hail recorder. *Atmospheric research*, **100 (4)**, 439–446.

Longobardi, A., G. Buttafuoco, T. Caloiero, and R. Coscarelli, 2016: Spatial and temporal distribution of precipitation in a Mediterranean area (southern Italy). *Environmental earth sciences*, **75 (3)**, 1–20.

Lorenz, E. N. and K. Haman, 1996: The essence of chaos. *Pure and Applied Geophysics*, **147 (3)**, 598–599.

Lott, F. and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, **123 (537)**, 101–127.

Ludlam, F. H., 1980: *Clouds and storms: The behavior and effect of water in the atmosphere.* Pennsylvania State University Press.

Lugo, A. E., 2018: Characterization of ten extreme disturbance events in the context of social and ecological systems. *Biogeochemistry*, **141 (3)**, 385–400.

Lupo, K. M., R. D. Torn, and S.-C. Yang, 2020: Evaluation of stochastic perturbed parameterization tendencies on convective-permitting ensemble forecasts of heavy rainfall events in New York and Taiwan. *Weather and Forecasting*, **35 (1)**, 5–24.

Lyra, G. B., T. P. Correia, J. F. de Oliveira-Júnior, and M. Zeri, 2018: Evaluation of methods of spatial interpolation for monthly rainfall data over the state of Rio de Janeiro, Brazil. *Theoretical and Applied Climatology*, **134 (3)**, 955–965.

Ma, Y., X. Sun, H. Chen, Y. Hong, and Y. Zhang, 2021: A two-stage blending approach for merging multiple satellite precipitation estimates and rain gauge observations: An experiment in the northeastern Tibetan Plateau. *Hydrology and Earth System Sciences*, **25 (1)**, 359–374.

Mahoney, K., M. A. Alexander, G. Thompson, J. J. Barsugli, and J. D. Scott, 2012: Changes in hail and flood risk in high-resolution simulations over Colorado's mountains. *Nature Climate Change*, **2 (2)**, 125–131.

Malečić, B., M. T. Prtenjak, K. Horvath, D. Jelić, P. M. Jurković, K. Ćorko, and N. S. Mahović, 2022a: Performance of HAILCAST and the Lightning Potential Index in simulating hailstorms in Croatia in a mesoscale model–Sensitivity to the PBL and microphysics parameterization schemes. *Atmospheric Research*, **272**, 106 143.

Malečić, B., et al., 2022b: Simulating hail and lightning over the Alpine Adriatic region-a model intercomparison study. *Authorea Preprints*.

Manz, B., W. Buytaert, Z. Zulkafli, W. Lavado, B. Willems, L. A. Robles, and J.-P. Rodríguez-Sánchez, 2016: High-resolution satellite-gauge merged precipitation climatologies of the Tropical Andes. *Journal of Geophysical Research: Atmospheres*, **121 (3)**, 1190–1207.

Manzato, A., 2012: Hail in northeast Italy: Climatology and bivariate analysis with the sounding-derived indices. *Journal of Applied Meteorology and Climatology*, **51 (3)**, 449–467.

Manzato, A., V. Riva, A. Tiesi, and M. Marcello Miglietta, 2020: Observational analysis and simulations of a severe hailstorm in northeastern Italy. *Quarterly Journal of the Royal Meteorological Society*, **146 (732)**, 3587–3611.

Manzato, A., S. Serafin, M. M. Miglietta, D. Kirshbaum, and W. Schulz, 2022: A pan–Alpine climatology of lightning and convective initiation. *Monthly Weather Review*.

Marbach, G. O., 1835: *Populäres physikalisches Lexikon oder Handwörterbuch der gesammten Naturlehre für die Gebildeten aus allen Ständen: E-G. 2*, Vol. 2. Wigand, URL `http://books.google.de/books?id=yHNYAAAAYAAJ`.

Marcos, J., J. Sánchez, A. Merino, P. Melcón, G. Mérida, and E. García-Ortega, 2021: Spatial and temporal variability of hail falls and estimation of maximum diameter from meteorological variables. *Atmospheric Research*, **247**, 105 142.

Marion, G., R. J. Trapp, and S. W. Nesbitt, 2019: Using overshooting top area to discriminate potential for large, intense tornadoes. *Geophysical Research Letters*, **46 (21)**, 12 520–12 526.

Markowski, P. and Y. Richardson, 2011: *Mesoscale meteorology in midlatitudes*, Vol. 2. John Wiley & Sons.

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella, 2005: The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification. *Nonlinear Processes in Geophysics*, **12 (4)**, 527–536.

Marsigli, C., A. Montani, and T. Paccangnella, 2008: A spatial verification method applied to the evaluation of high-resolution ensemble forecasts. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, **15 (1)**, 125–143.

Marsigli, C., et al., 2021: Observations for high-impact weather and their use in verification. *Natural Hazards and Earth System Sciences*, **21 (4)**, 1297–1312.

Martius, O., A. Hering, M. Kunz, A. Manzato, S. Mohr, L. Nisi, and S. Trefalt, 2018: Challenges and recent advances in hail research. *Bulletin of the American Meteorological Society*, **99 (3)**, ES51–ES54.

Mayaux, P., et al., 2006: Validation of the global land cover 2000 map. *IEEE Transactions on Geoscience and Remote Sensing*, **44 (7)**, 1728–1739.

Mecikalski, J. R., T. N. Sandmæl, E. M. Murillo, C. R. Homeyer, K. M. Bedka, J. M. Apke, and C. P. Jewett, 2021: A random-forest model to assess predictor importance and nowcast severe storms using high-resolution radar–GOES satellite–lightning observations. *Monthly Weather Review*, **149 (6)**, 1725–1746.

Melcón, P., A. Merino, J. L. Sánchez, L. López, and L. Hermida, 2016: Satellite remote sensing of hailstorms in France. *Atmospheric Research*, **182**, 221–231.

Miglietta, M. M. and S. Davolio, 2022: Dynamical forcings in heavy precipitation events over italy: lessons from the hymex sop1 campaign. *Hydrology and Earth System Sciences*, **26 (3)**, 627–646.

Mikuš, P. and N. S. Mahović, 2013: Satellite-based overshooting top detection methods and an analysis of correlated weather conditions. *Atmospheric Research*, **123**, 268–280.

Mikuš, P., M. T. Prtenjak, and N. S. Mahović, 2012: Analysis of the convective activity and its synoptic background over Croatia. *Atmospheric Research*, **104**, 139–153.

Minářová, J., M. Müller, A. Clappier, S. Hänsel, A. Hoy, J. Matschullat, and M. Kašpar, 2017: Duration, rarity, affected area, and weather types associated with extreme precipitation in the Ore Mountains (Erzgebirge) region, Central Europe. *International Journal of Climatology*, **37 (12)**, 4463–4477.

Mironov, D. V., 2008: *Parameterization of lakes in numerical weather prediction: Description of a lake model.* DWD Offenbach, Germany.

Mohr, S., M. Kunz, and B. Geyer, 2015a: Hail potential in Europe based on a regional climate model hindcast. *Geophysical Research Letters*, **42 (24)**, 10–904.

Mohr, S., M. Kunz, and K. Keuler, 2015b: Development and application of a logistic model to estimate the past and future hail potential in Germany. *Journal of Geophysical Research: Atmospheres*, **120 (9)**, 3939–3956.

Mohr, S., et al., 2022: A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis. *Natural Hazards and Earth System Sciences Discussions*, 1–44.

Morel, C. and S. Senesi, 2002: A climatology of mesoscale convective systems over Europe using satellite infrared imagery. i: Methodology. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, **128 (584)**, 1953–1971.

Müller, M., V. Bližňák, and M. Kašpar, 2018: Analysis of rainfall time structures on a scale of hours. *Atmospheric Research*, **211**, 38–51.

Murillo, E. M. and C. R. Homeyer, 2019: Severe hail fall and hailstorm detection using remote sensing observations. *Journal of applied meteorology and climatology*, **58 (5)**, 947–970.

Murphy, A. H., 1993: What is a good forecast? an essay on the nature of goodness in weather forecasting. *Weather and forecasting*, **8 (2)**, 281–293.

Nanditha, J., et al., 2022: The Pakistan flood of August 2022: causes and implications. *Authorea Preprints.*

Napoli, A., A. Crespi, F. Ragone, M. Maugeri, and C. Pasquero, 2019: Variability of orographic enhancement of precipitation in the Alpine region. *Scientific reports*, **9 (1)**, 1–8.

NCEI, 2020: Billion-dollar weather and climate disasters: Overview. National Centers for Environmental Information, NOAA, URL: `https://www.ncdc.noaa.gov/billions/`.

Niedda, M., et al., 2014: Il ciclone Cleopatra del 18 novembre 2013 in Sardegna: Analisi e modellazione dell'evento di piena. *Quad. Idronomia Montana*, **32**, 47–58.

Nieto, R., et al., 2005: Climatological features of cutoff low systems in the Northern Hemisphere. *Journal of climate*, **18 (16)**, 3085–3103.

Nisi, L., A. Hering, U. Germann, and O. Martius, 2018: A 15-year hail streak climatology for the Alpine region. *Quarterly Journal of the Royal Meteorological Society*, **144 (714)**, 1429–1449.

Nisi, L., A. Hering, U. Germann, K. Schroeer, H. Barras, M. Kunz, and O. Martius, 2020: Hailstorms in the Alpine region: Diurnal cycle, 4d-characteristics, and the nowcasting potential of lightning properties. *Quarterly Journal of the Royal Meteorological Society*, **146 (733)**, 4170–4194.

Ochoa-Rodriguez, S., L.-P. Wang, P. Willems, and C. Onof, 2019: A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resources Research*, **55 (8)**, 6356–6391.

Ortega, K. L., 2018: Evaluating multi-radar, multi-sensor products for surface hailfall diagnosis. *E-Journal of Severe Storms Meteorology*, **13 (1)**.

Pal, S., H.-I. Chang, C. L. Castro, and F. Dominguez, 2019: Credibility of convection-permitting modeling to improve seasonal precipitation forecasting in the southwestern United States. *Frontiers in Earth Science*, **7**, 11.

Palencia, C., D. Giaiotti, F. Stel, A. Castro, and R. Fraile, 2010: Maximum hailstone size: Relationship with meteorological variables. *Atmospheric Research*, **96 (2-3)**, 256–265.

Paterson, D. A. and R. Sankaran, 1994: Hail impact on building envelopes. *Journal of wind engineering and industrial aerodynamics*, **53 (1-2)**, 229–246.

Paul, S. H., H. O. Sharif, and A. M. Crawford, 2018: Fatalities caused by hydrometeorological disasters in Texas. *Geosciences*, **8 (5)**, 186.

Pavan, V., et al., 2019: High resolution climate precipitation analysis for north-central Italy, 1961–2015. *Climate Dynamics*, **52 (5)**, 3435–3453.

Paz, I., I. Tchiguirinskaia, and D. Schertzer, 2020: Rain gauge networks' limitations and the implications to hydrological modelling highlighted with a X-band radar. *Journal of Hydrology*, **583**, 124 615.

Pichelli, E., et al., 2021: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. *Climate Dynamics*, **56 (11)**, 3581–3602.

Porcù, F., L. Aragão, et al., 2019: OPEn-air laboRAtories for Nature baseD solUtions to Manage hydro-meteo risks: Data record on extreme events by OAL and by hazard. *OPERANDUM Deliverable 4.2*, URL https://doi.org/10.3030/776848.

Prein, A., A. Gobiet, M. Suklitsch, H. Truhetz, N. Awan, K. Keuler, and G. Georgievski, 2013a: Added value of convection permitting seasonal simulations. *Climate Dynamics*, **41 (9)**, 2655–2677.

Prein, A. F. and G. J. Holland, 2018: Global estimates of damaging hail hazard. *Weather and Climate Extremes*, **22**, 10–23.

Prein, A. F., G. J. Holland, R. M. Rasmussen, J. Done, K. Ikeda, M. P. Clark, and C. H. Liu, 2013b: Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters. *Journal of climate*, **26 (13)**, 4848–4857.

Prein, A. F., R. M. Rasmussen, K. Ikeda, C. Liu, M. P. Clark, and G. J. Holland, 2017: The future intensification of hourly precipitation extremes. *Nature climate change*, **7 (1)**, 48–52.

Prein, A. F., et al., 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of geophysics*, **53 (2)**, 323–361.

Proud, S. R., 2015: Analysis of overshooting top detections by Meteosat Second Generation: a 5-year dataset. *Quarterly Journal of the Royal Meteorological Society*, **141 (688)**, 909–915.

Pruppacher, H. R., J. D. Klett, and P. K. Wang, 1998: *Microphysics of clouds and precipitation*. Taylor & Francis.

Púčik, T., C. Castellano, P. Groenemeijer, T. Kühne, A. T. Rädler, B. Antonescu, and E. Faust, 2019: Large hail incidence and its economic and societal impacts across Europe. *Monthly Weather Review*, **147 (11)**, 3901–3916.

Púčik, T., P. Groenemeijer, D. Rýva, and M. Kolář, 2015: Proximity soundings of severe and nonsevere thunderstorms in central Europe. *Monthly Weather Review*, **143 (12)**, 4805–4821.

Púčik, T., P. Groenemeijer, and I. Tsonevsky, 2021: Technical memorandum n 879: Vertical wind shear and convective storms. *ECMWF Technical Memoranda*.

Púčik, T., et al., 2017: Future changes in European severe convection environments in a regional climate model ensemble. *Journal of Climate*, **30 (17)**, 6771–6794.

Puczko, K. and E. Jekatierynczuk-Rudczyk, 2020: Extreme hydro-meteorological events influence to water quality of small rivers in urban area: a case study in Northeast Poland. *Scientific Reports*, **10 (1)**, 1–14.

Punge, H., K. Bedka, M. Kunz, and A. Reinbold, 2017: Hail frequency estimation across Europe based on a combination of overshooting top detections and the ERA-INTERIM reanalysis. *Atmospheric Research*, **198**, 34–43.

Punge, H., K. Bedka, M. Kunz, and A. Werner, 2014: A new physically based stochastic event catalog for hail in Europe. *Natural Hazards*, **73 (3)**, 1625–1645.

Punge, H. J., K. M. Bedka, M. Kunz, S. D. Bang, and K. F. Itterly, 2021: Characteristics of hail hazard in South Africa based on satellite detection of convective storms. *Natural Hazards and Earth System Sciences Discussions*, 1–32.

Punge, H. J. and M. Kunz, 2016: Hail observations and hailstorm characteristics in Europe: A review. *Atmospheric Research*, **176**, 159–184.

Qiu, Y., A. Ichiba, I. D. S. R. Paz, F. Chen, P.-A. Versini, D. Schertzer, and I. Tchiguirinskaia, 2019: Evaluation of low impact development and nature-based solutions for stormwater management: A fully distributed modelling approach. *Hydrology and Earth System Sciences Discussions*, 1–27.

Quevauviller, P. and M. Gemmer, 2015: EU and international policies for hydrometeorological risks: operational aspects and link to climate action. *Advances in Climate Change Research*, **6 (1)**, 74–79.

Rädler, A. T., P. Groenemeijer, E. Faust, and R. Sausen, 2018: Detecting severe weather trends using an additive regressive convective hazard model (AR-CHaMo). *Journal of Applied Meteorology and Climatology*, **57 (3)**, 569–587.

Raffa, M., A. Reder, G. F. Marras, M. Mancini, G. Scipione, M. Santini, and P. Mercogliano, 2021: VHR-REA_IT dataset: Very high resolution dynamical downscaling of ERA5 reanalysis over Italy by COSMO-CLM. *Data*, **6 (8)**, 88.

Raschendorfer, M., 2001: The new turbulence parameterization of LM. *COSMO newsletter*, **1**, 89–97.

Rasmussen, K. L., A. F. Prein, R. M. Rasmussen, K. Ikeda, and C. Liu, 2020: Changes in the convective population and thermodynamic environments in convection-permitting regional climate simulations over the United States. *Climate Dynamics*, **55 (1)**, 383–408.

Raupach, T. H., et al., 2021: The effects of climate change on hailstorms. *Nature reviews earth & environment*, **2 (3)**, 213–226.

Rebora, N., et al., 2013: Extreme rainfall in the Mediterranean: What can we learn from observations? *Journal of Hydrometeorology*, **14 (3)**, 906–922.

Reder, A., M. Raffa, R. Padulano, G. Rianna, and P. Mercogliano, 2022: Characterizing extreme values of precipitation at very high resolution: An experiment over twenty European cities. *Weather and Climate Extremes*, **35**, 100 407.

Reynolds, D. W., 1980: Observations of damaging hailstorms from geosynchronous satellite digital data. *Monthly Weather Review*, **108 (3)**, 337–348.

Riemann-Campe, K., K. Fraedrich, and F. Lunkeit, 2009: Global climatology of convective available potential energy (CAPE) and convective inhibition (CIN) in ERA-40 reanalysis. *Atmospheric Research*, **93 (1-3)**, 534–545.

Riishojgaard, L. P., 2020: Impacts of COVID-19 restrictions on observations and monitoring. *Boletín-Organización Meteorológica Mundial*, **69 (2)**, 16–19.

Ritter, B. and J.-F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Monthly weather review*, **120 (2)**, 303–325.

Rivin, G., et al., 2015: The COSMO-Ru system of nonhydrostatic mesoscale short-range weather forecasting of the Hydrometcenter of Russia: The second stage of implementation and development. *Russian Meteorology and hydrology*, **40 (6)**, 400–410.

Rivoire, P., O. Martius, and P. Naveau, 2021: A comparison of moderate and extreme ERA-5 daily precipitation with two observational data sets. *Earth and Space Science*, **8 (4)**, e2020EA001 633.

Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, **136 (1)**, 78–97.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Weather and Forecasting*, **24 (2)**, 601–608.

Ruangpan, L., Z. Vojinovic, S. Di Sabatino, L. S. Leo, V. Capobianco, A. M. Oen, M. E. McClain, and E. Lopez-Gunn, 2020: Nature-based solutions for hydro-meteorological risk reduction: A state-of-the-art review of the research area. *Natural Hazards and Earth System Sciences*, **20 (1)**, 243–270.

Sahani, J., et al., 2019: Hydro-meteorological risk assessment methods and management by nature-based solutions. *Science of the Total Environment*, **696**, 133 936.

Saltikoff, E., et al., 2019: Opera the radar project. *Atmosphere*, **10 (6)**, 320.

Sánchez, J., et al., 2009: Characterization of hailstone size spectra in hailpad networks in France, Spain, and Argentina. *Atmospheric Research*, **93 (1-3)**, 641–654.

Sartori, G., 2012: Eventi meteorologici estremi. *Dati e valutazioni sulla radicalizzazione del clima in Veneto. Rapporto Tecnico. Consiglio Regionale del Veneto.*

Scarino, B., K. Bedka, K. Itterly, C. Homeyer, and J. Allen, 2022: Deriving severe hail likelihood from satellite observations and model reanalysis parameters using a deep neural network. *EUMETSAT Convection Working Group and MTG 3T Forum.*

Schär, C., et al., 2020: Kilometer-scale climate models: Prospects and challenges. *Bulletin of the American Meteorological Society*, **101 (5)**, E567–E587.

Schättler, U. and U. Blahak, 2017: A description of the non-hydrostatic regional COSMO-model. Part V: Preprocessing: Initial and boundary data for the COSMO-model. *Deutscher Wetterdienst Rep. COSMO-Model*, 86.

Schättler, U., G. Doms, and C. Schraff, 2018: A description of the nonhydrostatic regional COSMO-model. Part VII: user's guide. *Deutscher Wetterdienst Rep. COSMO-Model*, 195.

Schlatter, T. W. and N. Doesken, 2010: Deep hail: Tracking an elusive phenomenon. *Weatherwise: THE POWER, THE BEAUTY, THE EXCITEMENT*, **63 (5)**, 35–41.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, **83 (7)**, 977–992.

Schraff, C. and R. Hess, 2013: A description of the nonhydrostatic regional COSMO-model. Part III: Data assimilation. *Deutscher Wetterdienst Rep. COSMO-Model*, 93.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *Journal of Applied Meteorology and Climatology*, **48 (12)**, 2543–2563.

Semie, A. G. and S. Bony, 2020: Relationship between precipitation extremes and convective organization inferred from satellite observations. *Geophysical Research Letters*, **47 (9)**, e2019GL086 927.

Seneviratne, S., et al., 2021: Chapter 11: Weather and climate extreme events in a changing climate. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, in Climate Change 2021: The Physical Science Basis*, (Cambridge University Press Cambridge, Cambridge, UK and New York, USA), 1513–1766.

Setvák, M., K. Bedka, D. T. Lindsey, A. Sokol, Z. Charvát, J. Šťástka, and P. K. Wang, 2013: A-train observations of deep convective storm tops. *Atmospheric research*, **123**, 229–248.

Shah, M. A. R., et al., 2020: A review of hydro-meteorological hazard, vulnerability, and risk assessment frameworks and indicators in the context of nature-based solutions. *International journal of disaster risk reduction*, **50**, 101 728.

Sherwood, S. C., J.-H. Chae, P. Minnis, and M. McGill, 2004: Underestimation of deep convective cloud tops by thermal imagery. *Geophysical Research Letters*, **31 (11)**.

Simon, T., D. Wang, A. Hense, C. Simmer, and C. Ohlwein, 2013: Generation and transfer of internal variability in a regional climate model. *Tellus A: Dynamic Meteorology and Oceanography*, **65 (1)**, 22 485.

Singh, T., U. Saha, V. Prasad, and M. D. Gupta, 2021: Assessment of newly-developed high resolution reanalyses (IMDAA, NGFS and ERA5) against rainfall observations for Indian region. *Atmospheric Research*, **259**, 105 679.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Weather and Forecasting*, **27 (5)**, 1114–1135.

Smith, P. L. and A. Waldvogel, 1989: On determinations of maximum hailstone sizes from hailpad observations. *Journal of Applied Meteorology (1988-2005)*, 71–76.

Sommeria, G. and J. Deardorff, 1977: Subgrid-scale condensation in models of nonprecipitating clouds. *Journal of Atmospheric Sciences*, **34 (2)**, 344–355.

Spekkers, M., V. Rözer, A. Thieken, M.-C. ten Veldhuis, and H. Kreibich, 2017: A comparative survey of the impacts of extreme rainfall in two international case studies. *Natural Hazards and Earth System Sciences*, **17 (8)**, 1337–1355.

Spitzer, A., H. Kempf, M. Jerg, and U. Blahak, 2022: DWD-crowdsourcing: User reports available on open data. Tech. rep., Copernicus Meetings.

Starosta, K. and A. Wyszogrodzki, 2016: Assessment of model generated wind energy potential in Poland. *COSMO News Letter*, **(16)**.

Stauffer, D. R. and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*, **118 (6)**, 1250–1277.

Stommel, H., 1947: Entrainment of air into a cumulus cloud. *J. Meteor*, **4 (3)**, 91–94.

Stržinar, G. and G. Skok, 2018: Comparison and optimization of radar-based hail detection algorithms in Slovenia. *Atmospheric research*, **203**, 275–285.

Su, C.-H., et al., 2019: BARRA v1. 0: the bureau of meteorology atmospheric high-resolution regional reanalysis for Australia. *Geoscientific Model Development*, **12 (5)**, 2049–2068.

Svabik, O., V. Meyer, L. Tüchler, and G. Zenkl, 2013: Hail risk areas in Austria, on the basis of reports 1971–2011 and weather radar images 2002–2011. *7th European Conference on Severe Storms. Helsinki, Finland*, 2.

Tabari, H., K. Madani, and P. Willems, 2020: The contribution of anthropogenic influence to more anomalous extreme precipitation in Europe. *Environmental Research Letters*, **15 (10)**, 104 077.

Tang, B. H., V. A. Gensini, and C. R. Homeyer, 2019: Trends in United States large hail environments and observations. *NPJ climate and atmospheric science*, **2 (1)**, 1–7.

Taszarek, M., J. T. Allen, H. E. Brooks, N. Pilguj, and B. Czernecki, 2021: Differing trends in United States and European severe thunderstorm environments in a warming climate. *Bulletin of the American Meteorological society*, **102 (2)**, E296–E322.

Taszarek, M., J. T. Allen, T. Púčik, K. A. Hoogewind, and H. E. Brooks, 2020: Severe convective storms across europe and the United States. Part II: ERA5 environments associated with lightning, large hail, severe wind, and tornadoes. *Journal of Climate*, **33 (23)**, 10 263–10 286.

Teixeira, J., et al., 2008: Parameterization of the atmospheric boundary layer: a view from just above the inversion. *Bulletin of the American Meteorological Society*, **89 (4)**, 453–458.

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Weather and Forecasting*, **18 (6)**, 1243–1261.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly weather review*, **117 (8)**, 1779–1800.

Tiesi, A., et al., 2022: Numerical simulation of a giant-hail-bearing Mediterranean supercell in the Adriatic sea. *Atmosphere*, **13 (8)**, 1219.

Tippett, M. K., A. H. Sobel, S. J. Camargo, and J. T. Allen, 2014: An empirical relation between US tornado activity and monthly environmental parameters. *Journal of Climate*, **27 (8)**, 2983–2999.

Torralba, V., R. Hénin, A. Cantelli, E. Scoccimarro, S. Materia, A. Manzato, and S. Gualdi, 2022: Modelling hail hazard over Italy with ERA5 large-scale variables. *Weather and Climate Extremes*, 100535.

Trapp, R. J., N. S. Diffenbaugh, H. E. Brooks, M. E. Baldwin, E. D. Robinson, and J. S. Pal, 2007: Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proceedings of the National Academy of Sciences*, **104 (50)**, 19 719–19 723.

Trapp, R. J. and K. A. Hoogewind, 2016: The realization of extreme tornadic storm events under future anthropogenic climate change. *Journal of Climate*, **29 (14)**, 5251–5265.

Trapp, R. J., K. A. Hoogewind, and S. Lasher-Trapp, 2019: Future changes in hail occurrence in the United States determined through convection-permitting dynamical downscaling. *Journal of Climate*, **32 (17)**, 5493–5509.

Trusilova, K., B. Früh, S. Brienen, A. Walter, V. Masson, G. Pigeon, and P. Becker, 2013: Implementation of an urban parameterization scheme into the regional climate model COSMO-CLM. *Journal of Applied Meteorology and Climatology*, **52 (10)**, 2296–2311.

Tu, C.-C., Y.-L. Chen, P.-L. Lin, and M.-Q. Huang, 2022: Analysis and simulations of a heavy rainfall event associated with the passage of a shallow front over Northern Taiwan on 2 June 2017. *Monthly Weather Review*, **150 (3)**, 505–528.

Unden, P., R. Renshaw, E. Bazile, M. Brunet, A. Kaiser-weiss, and A. Klein Tank, 2016: UERRA-uncertainties in ensembles of regional reanalyses. *EGU General Assembly Conference Abstracts*, EPSC2016–15 450.

Vulpiani, G., M. Montopoli, L. D. Passeri, A. G. Gioia, P. Giordano, and F. S. Marzano, 2012: On the use of dual-polarized C-band radar for operational rainfall retrieval in mountainous areas. *Journal of Applied Meteorology and Climatology*, **51 (2)**, 405–425.

Wahl, S., C. Bollmeyer, S. Crewell, C. Figura, P. Friederichs, A. Hense, J. D. Keller, and C. Ohlwein, 2017: A novel convective-scale regional reanalysis COSMO-REA2: Improving the representation of precipitation. *Meteorol. Z*, **26 (4)**, 345–361.

Wang, L. and W. Gu, 2016: The Eastern China flood of June 2015 and its causes. *Science Bulletin*, **61 (2)**, 178–184.

Ward, A., M. Kumjian, M. Bunkers, S. Bieda III, and R. Simpson, 2018: Using polarimetric radar data to identify potentially hazardous hail accumulations. *34th Conf. on Environmental Information Processing Technologies, Austin, TX*, Amer. Meteor. Soc., 11B. 1.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, **125 (4)**, 527–548.

Wellmann, C., A. I. Barrett, J. S. Johnson, M. Kunz, B. Vogel, K. S. Carslaw, and C. Hoose, 2020: Comparing the impact of environmental conditions and microphysics on the forecast uncertainty of deep convective clouds and hail. *Atmospheric Chemistry and Physics*, **20 (4)**, 2201–2219.

Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Monthly Weather Review*, **138 (9)**, 3418–3433.

Weygandt, S. S., A. F. Loughe, S. G. Benjamin, and J. L. Mahoney, 2004: Scale sensitivities in model precipitation skill scores during IHOP. *Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc. A*, Vol. 16.

Wicker, L. J. and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Monthly weather review*, **130 (8)**, 2088–2097.

Wilhelmson, R. B. and L. J. Wicker, 2001: Numerical modeling of severe local storms. *Severe convective storms*, Springer, 123–166.

Wilks, D. S., 2019: *Statistical methods in the atmospheric sciences*, Vol. 100. Elsevier Science.

Wu, C.-M. and A. Arakawa, 2014: A unified representation of deep moist convection in numerical modeling of the atmosphere. Part II. *Journal of the Atmospheric Sciences*, **71 (6)**, 2089–2103.

Wuebbles, D. J., D. W. Fahey, and K. A. Hibbard, 2017: Climate science special report: 4th national climate assessment, volume I, US. *Global Change Research Program, Washington, DC, USA*, 470.

Xie, B., Q. Zhang, and Y. Wang, 2008: Trends in hail in China during 1960–2005. *Geophysical Research Letters*, **35 (13)**.

Ye, H., 2018: Changes in duration of dry and wet spells associated with air temperatures in Russia. *Environmental Research Letters*, **13 (3)**, 034 036.

Yeo, S., R. Leigh, and I. Kuhne, 1999: The April 1999 Sydney hailstorm. *Australian Journal of Emergency Management, The*, **14 (4)**, 23–25.

Zhang, Q., Y. Pan, S. Wang, J. Xu, and J. Tang, 2017: High-resolution regional reanalysis in China: Evaluation of 1 year period experiments. *Journal of Geophysical Research: Atmospheres*, **122 (20)**, 10–801.

Zhou, J., M. J. Pavek, S. C. Shelton, Z. J. Holden, and S. Sankaran, 2016: Aerial multispectral imaging for crop hail damage assessment in potato. *Computers and Electronics in Agriculture*, **127**, 406–412.

Acknowledgments and credits

This thesis would not have been possible without the vital contributions of some people I was fortunate enough to meet during the journey of my doctoral studies and whom I wish to acknowledge.

I wish to thank my supervisors Tiziana Paccagnella and Silvana Di Sabatino for the trust and the guidance received throughout the whole path of my PhD studies. Particularly, I acknowledge Tiziana for the fruitful scientific discussions through which I could deeply expand my knowledge and improve the quality of the analysis contained in this thesis. I thank Silvana for the fundamental suggestions to best direct my efforts and improve the formal presentation of the scientific results through which this work could substantially elevate to a superior level.

Many thanks go also to the whole numerical modeling group of ARPAE-SIMC for having welcomed me as family and for the many crucial discussions on technical and scientific aspects of the analyses from which I could learn how to properly tackle the obstacles I found during my path. In particular, I wish to acknowledge Ines Cerenzia, co-supervisor of this thesis, for the friendship and the time dedicated to teaching me the proper tools to make this work possible, and Davide Cesari for being always available at helping to solve any kind of possible technical problem that could arise (which at the beginning were many!). I wish to acknowledge also the rest of the ARPAE team who gave me many precious and uncountable suggestions, as well the opportunity to see the tackled problems from many different angles and perspectives, who are (in random order): Thomas Gastaldo, Valentina Pavan, Virginia Poli, Pier Paolo Alberoni, Maria Stefania Tesini, Enrico Minguzzi, and Chiara Marsigli.

I received fundamental and decisively precious inputs also by Michael Kunz and Heinz Jürgen Punge, to whom go all my gratitude for welcoming me into their Atmospheric Risks research group at KIT, Karlsruhe, Germany during the four months I had the luck to spend there. I wish to acknowledge also Kristopher (Kris) Bedka from NASA for providing satellite data, as well as the crucial expertise and

time to discuss how to properly employ them.

I wish to acknowledge Silvio Davolio and Dino Zardi for revising the first draft of the thesis, whose comments, I believe, led to significant improvements in the clarity of the exposition of the scientific results presented.

I am convinced that one of the most important aspects for better understanding a scientific problem (and not only that) is to face it from different angles, and I am thankful to you all that gave me this opportunity.

I would like to thank also Valerie, a very special friend I met during my stay in Germany, for her precious contribution that gave me the chance to venture into new unexplored perspectives.

I acknowledge the OPERANDUM project for having granted me economical support throughout my PhD studies.

I want to express my gratitude also to my family who always supported my decision to engage in this adventure and for the love received.

Last but not least, I wish to thank Irene for daily teaching me how to breathe and conspire together.