



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN  
IL FUTURO DELLA TERRA, CAMBIAMENTI CLIMATICI E SFIDE  
SOCIALI

Ciclo 38

**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

IMPROVING PREDICTABILITY OF DEEP CONVECTIVE STORMS WITH  
ICON HECTOMETRIC-SCALE ENSEMBLES

**Presentata da:** Matteo Siena

**Coordinatore Dottorato**

Silvana Di Sabatino

**Supervisore**

Silvana Di Sabatino

**Co-supervisor**

Marco Briscolini

Chiara Marsigli

Paolo Ruggieri

Thomas Gastaldo

Esame finale anno 2026

# ABSTRACT

---

Extreme precipitation events represent an increasing global challenge, particularly affecting the Mediterranean basin, where their frequency and intensity have been rising in recent decades. Accurately forecasting these events remains a key objective for numerical weather prediction, yet operational models still struggle to reproduce the magnitude, spatial organization, and timing of intense convective systems. These limitations mainly stem from the relatively coarse grid spacing and the continued reliance on convection parameterizations, which hinder the realistic simulation of localized and short-lived storms.

This first part of this work explores the potential of hectometric-scale modelling, employing the ICON model with nested simulations characterized by a grid spacing as fine as 500 m approximately at the innermost domain to reproduce the devastating floods that impacted Italy's Marche region in September 2022. The study combines sensitivity experiments and ensemble simulations to assess the impact of enhanced resolution and improved turbulence representation. The results reveal substantial improvements in the structure, amplitude, and spatial accuracy of the simulated precipitation fields when convection is treated explicitly and the three-dimensional Smagorinsky turbulence scheme is adopted. The ensemble analysis shows that even at this stage, without changes in microphysical parameterizations, high-resolution simulations already provide more realistic precipitation distributions and more skillful forecasts. Moreover, the results highlight the crucial role of initial and boundary conditions in controlling forecast skill. Some ensemble members clearly outperform others due to a more accurate depiction of key mesoscale features, such as low-level wind convergence, moisture transport, and instability patterns. This emphasizes that ensemble spread at hectometric scales reflects not only model uncertainty but also the system's intrinsic sensitivity to initial perturbations.

In the second part of the thesis, the focus shifts to the role of cloud microphysics and the representation of hydrometeor species. Perturbations are introduced in the two-moment microphysics scheme to assess their influence on ensemble spread, convective intensity, and precipitation type. This analysis is



complemented by a second case study centred on a severe convective outbreak over southern Germany, characterized by intense hail formation. The event provides an ideal testbed for evaluating the model's capability to reproduce microphysical processes and solid precipitation at hectometric resolution. The comparison between the two cases reveals how the combined effect of high resolution, advanced turbulence, and microphysics parameterizations enhances the realism of both rainfall and hail forecasts, while also improving the ensemble's representation of uncertainty.

Overall, this study underlines the added value of hectometric-scale ensemble modelling in capturing the complexity of convective processes and demonstrates how advanced physical formulations, when coupled with explicit convection and careful ensemble design, can substantially improve the predictability of extreme rainfall and hail events. Such developments are essential to strengthen early warning capabilities and inform risk management strategies in a changing climate.

# Contents

|                                                                                                      |           |
|------------------------------------------------------------------------------------------------------|-----------|
| <b>List of Figures</b>                                                                               | <b>6</b>  |
| <b>List of Tables</b>                                                                                | <b>10</b> |
| <b>1 Introduction and motivation</b>                                                                 | <b>13</b> |
| 1.1 Extreme precipitation predictability in a changing climate . . . .                               | 13        |
| 1.1.1 The use of ensembles and high-resolution to tackle the<br>uncertainty . . . . .                | 14        |
| 1.1.2 The impact of deep convection and the role of early warn-<br>ing systems . . . . .             | 14        |
| 1.2 The importance of high-resolution NWP models . . . . .                                           | 17        |
| 1.2.1 Reference CP systems in research and operational activities                                    | 21        |
| 1.2.2 Summary and choice of the model . . . . .                                                      | 23        |
| 1.3 An example of challenges in storm predictability: The sinking of<br>the Bayesian yacht . . . . . | 24        |
| 1.4 Context and open challenges . . . . .                                                            | 28        |
| 1.5 The Marche 2022 flooding event as a motivation . . . . .                                         | 29        |
| 1.6 Thesis outline and objective . . . . .                                                           | 30        |
| <b>2 The ICON model for high-resolution simulations</b>                                              | <b>33</b> |
| 2.1 ICON Model Description . . . . .                                                                 | 33        |
| 2.1.1 Model physics and governing equations . . . . .                                                | 33        |
| 2.1.2 The Model Reference State and simplifying assumptions                                          | 38        |
| 2.1.3 Physical Parameterization Schemes . . . . .                                                    | 42        |
| 2.2 Initial and Boundary Conditions . . . . .                                                        | 43        |
| 2.2.1 Initial Conditions . . . . .                                                                   | 43        |
| 2.2.2 Boundary conditions and external parameters . . . . .                                          | 44        |
| 2.3 ICON nesting strategy . . . . .                                                                  | 45        |
| 2.4 ICON Application for High-Resolution Case Studies . . . . .                                      | 47        |

|          |                                                                                              |            |
|----------|----------------------------------------------------------------------------------------------|------------|
| 2.4.1    | Model Configuration and Computational Considerations                                         | 48         |
| <b>3</b> | <b>Preliminary analysis of a high impact convective storm with ICON high-res simulations</b> | <b>51</b>  |
| 3.1      | The unprecedented Marche 2022 flooding event . . . . .                                       | 51         |
| 3.1.1    | Predictability and preliminary sensitivity tests . . . . .                                   | 55         |
| <b>4</b> | <b>The role of hectometric-scale ensembles</b>                                               | <b>65</b>  |
| 4.1      | Sensitivity to initial conditions . . . . .                                                  | 65         |
| 4.1.1    | Ensemble forcing data . . . . .                                                              | 66         |
| 4.1.2    | LOWRES ensemble (2.2 km) . . . . .                                                           | 66         |
| 4.1.3    | COSMO_tke ensemble (550 m) . . . . .                                                         | 67         |
| 4.1.4    | SMAGO_500m ensemble (550 m) . . . . .                                                        | 67         |
| 4.2      | The forecasting of wind and water vapour . . . . .                                           | 68         |
| 4.3      | Validation of precipitation ensembles . . . . .                                              | 74         |
| 4.3.1    | Ensemble statistics . . . . .                                                                | 74         |
| 4.3.2    | Overlap ratio . . . . .                                                                      | 77         |
| 4.3.3    | SAL score . . . . .                                                                          | 79         |
| 4.4      | Final considerations and summary . . . . .                                                   | 90         |
| <b>5</b> | <b>Impact of microphysics perturbations</b>                                                  | <b>93</b>  |
| 5.1      | 2-MOM microphysics scheme . . . . .                                                          | 94         |
| 5.2      | Italian ensemble . . . . .                                                                   | 95         |
| 5.2.1    | Member analysis . . . . .                                                                    | 98         |
| 5.2.2    | SAL analysis . . . . .                                                                       | 102        |
| 5.3      | German case study and the MOSES campaign . . . . .                                           | 109        |
| 5.3.1    | A record year for hail . . . . .                                                             | 110        |
| 5.3.2    | Observation and description of the event . . . . .                                           | 111        |
| 5.3.3    | German ensemble setup . . . . .                                                              | 112        |
| 5.3.4    | Ensemble analysis . . . . .                                                                  | 114        |
| 5.3.5    | Domain-averaged microphysical processes . . . . .                                            | 116        |
| 5.4      | Comments on the correlation between the two events . . . . .                                 | 118        |
| <b>6</b> | <b>Conclusion and future perspectives</b>                                                    | <b>125</b> |
| 6.1      | The dominance of IC/LBC uncertainty . . . . .                                                | 125        |
| 6.2      | Added value of hectometric-scale ensembles . . . . .                                         | 127        |
| 6.3      | Turbulence closures and microphysics uncertainty . . . . .                                   | 128        |

|          |                                                           |            |
|----------|-----------------------------------------------------------|------------|
| 6.4      | Contrasting hazards: floods versus hail . . . . .         | 129        |
| 6.5      | Implications for forecasting and early warnings . . . . . | 130        |
| 6.6      | Future outlooks and generalized conclusions . . . . .     | 130        |
| <b>A</b> | <b>Parameters description</b>                             | <b>135</b> |
|          | <b>Acknowledgements</b>                                   | <b>161</b> |

# List of Figures

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |    |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | The streets of Senigallia (Marche, Italy) following the flooding caused by the overflowing of the Misa river after the huge rainfall that hit central Italy on 15th September 2022. . . . .                                                                                                                                                                                                                                                                                                                                   | 16 |
| 1.2 | (a), (b) show respectively the Probability of maximum wind gusts at 10 m stronger than 24.5 m/s and the maximum wind themselves as forecasted by the ICON-EPS system, while (c) and (d) illustrate the wind gusts at 10 m as forecasted by the deterministic run of the model ICON-2I operationally run at ARPAE-SIMC for the bayesian sinking. Initialization and valid time are reported on the figures. Pictures produced by the author of this thesis and co-author Thomas Gastaldo, from De Martin et al., 2025. . . . . | 26 |
| 3.1 | Synoptic setup at the beginning of the event, temperature at 850 hPa and geopotential height at 500 hPa are shown. . . . .                                                                                                                                                                                                                                                                                                                                                                                                    | 53 |
| 3.2 | Hourly and cumulative precipitation collected by the weather station in Cantiano on 15/09/2022. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                       | 54 |
| 3.3 | (a) 24h total cumulated precipitation as observed by the Italian radar-based network adjusted with rain gauges data. (b), (c), (d), (e) 3-hourly cumulated rainfall at different times from 12:00 UTC to 24:00 UTC. . . . .                                                                                                                                                                                                                                                                                                   | 56 |
| 3.4 | Domain of the simulations. Parent 2.2 km horizontal domain including the two nests used for the sensitivity tests and the high-resolution ensembles. . . . .                                                                                                                                                                                                                                                                                                                                                                  | 58 |
| 3.5 | Comparison on the whole Italian domain of 24h total precipitation field from 00:00 UTC 15/09/2022 to 00:00 UTC 15/09/2022 between (a) Observed radar-rain gauge network (b) IFS forecast data, (c) ICON control run at 2.2 km with COSMO operational setup. . . . .                                                                                                                                                                                                                                                           | 60 |

|      |                                                                                                                                                                                                                                                                                                                                            |    |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.6  | Comparison on the flooding area of 24h total precipitation field from 00:00 UTC 15/09/2022 to 00:00 UTC 15/09/2022 between (a) Observed radar network (b) IFS forecast data, (c) ICON run at 2.2 km with COSMO operational setup (d) ICON run at 550 m with Smagorinsky turbulence scheme. . . . .                                         | 61 |
| 4.1  | LOWRES ensemble without nested domains. Accumulated 24h total precipitation from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC. . . . .                                                                                                                                                                                                     | 70 |
| 4.2  | Accumulated 24h total precipitation over the 550 m grid from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC for the COSMO_tke high-resolution ensemble with double nesting. . . . .                                                                                                                                                          | 71 |
| 4.3  | Accumulated 24h total precipitation from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC over the 550 m grid for the SMAGO_500m high-resolution ensemble with double nesting and activation of Smagorinsky turbulence for the innermost domain only. . . . .                                                                                  | 72 |
| 4.4  | Zonal wind intensity and direction at 850 hPa as forecasted by the two members m1 (a,c) and m2 (b,d) at 16:00 UTC and 18:00 UTC respectively at 550 m resolution. Member m2 exhibits stronger and more organized southerly inflow and convergence. . . . .                                                                                 | 73 |
| 4.5  | Specific humidity at 850 hPa geopotential height for the members m1 (a) and m2 (b) at 18:00 UTC of 15/09/2022. . . . .                                                                                                                                                                                                                     | 73 |
| 4.6  | Ensemble statistics of the 10-members IC-perturbed ensemble for the LOWRES nest for the Marche flooding that happened on 15th September 2022 . . . . .                                                                                                                                                                                     | 76 |
| 4.7  | Same as figure 4.6, but for the COSMO_tke ensemble. . . . .                                                                                                                                                                                                                                                                                | 76 |
| 4.8  | Same as figure 4.6, but for the SMAGO_500m ensemble. . . . .                                                                                                                                                                                                                                                                               | 77 |
| 4.9  | Overlap ratio between forecasted and observed total precipitation field for varying threshold and ensemble member. The three ensembles are compared. . . . .                                                                                                                                                                               | 79 |
| 4.10 | SAL scores for the three different ensembles, LOWRES (a), COSMO_tke (b), SMAGO_500m (c), at 10 mm threshold. The grey box represents the interquartile range. Dotted lines represent the median values of A and S, their colour shows the corresponding median value of L, also pointed out by the black line into the L colorbar. . . . . | 84 |
| 4.11 | Same as figure 4.10 but for the 20 mm total precipitation threshold. . . . .                                                                                                                                                                                                                                                               | 85 |

|      |                                                                                                                                                                                                                                                                                                                   |     |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.12 | Same as figure 4.10 but for the 40 mm total precipitation threshold.                                                                                                                                                                                                                                              | 86  |
| 4.13 | Same as figure 4.10 but for the 60 mm total precipitation threshold.                                                                                                                                                                                                                                              | 87  |
| 4.14 | Same as figure 4.10 but for the 80 mm total precipitation threshold.                                                                                                                                                                                                                                              | 88  |
| 4.15 | Same as figure 4.10 but for the 100 mm total precipitation threshold.                                                                                                                                                                                                                                             | 89  |
| 5.1  | Ensemble statistics of the 36-members ensemble for the Marche flooding that happened on 15th September 2022. . . . .                                                                                                                                                                                              | 98  |
| 5.2  | Ensemble members 19, 20, 21 and 22 from the microphysics Marche ensemble . . . . .                                                                                                                                                                                                                                | 100 |
| 5.3  | Ensemble members 23, 24, 25, 26 from the microphysics Marche ensemble . . . . .                                                                                                                                                                                                                                   | 101 |
| 5.4  | Ensemble members 27, 28, 29, 30 from the microphysics Marche ensemble . . . . .                                                                                                                                                                                                                                   | 101 |
| 5.5  | SAL scores for the 10 mm and 20 mm thresholds for the 2-MOM Italian ensemble. The grey box represents the interquartile range (IQR). Dotted lines represent the median values of A and S, their colour shows the corresponding median value of L, also pointed out by the black line into the L colorbar. . . . . | 106 |
| 5.6  | Same as figure 5.5 but for the 40 mm and 60 mm precipitation thresholds. . . . .                                                                                                                                                                                                                                  | 107 |
| 5.7  | Same as figure 5.5 but for the 80 mm and 100 mm precipitation thresholds. . . . .                                                                                                                                                                                                                                 | 108 |
| 5.8  | 24h Observed total precipitation from 14/08/2023 00:00 UTC to 15/08/2023 00:00 UTC as retrieved from the RADOLAN network.                                                                                                                                                                                         | 112 |
| 5.9  | Domains of the german case simulations. . . . .                                                                                                                                                                                                                                                                   | 113 |
| 5.10 | Ensemble statistics of the 36-members ensemble for the Villingen-Schwenningen hail event that happened on 14th august 2023. .                                                                                                                                                                                     | 114 |
| 5.11 | Domain-averaged cumulative hail precipitation [mm] for each ensemble member. . . . .                                                                                                                                                                                                                              | 119 |
| 5.12 | Domain-averaged cumulative graupel [mm] for each ensemble member. . . . .                                                                                                                                                                                                                                         | 120 |
| 5.13 | Domain-averaged riming rate on hail particles (kg/kg/s) for each ensemble member. . . . .                                                                                                                                                                                                                         | 121 |
| 5.14 | Domain-averaged nucleation rate from cloud water to ice (kg/kg/s) for each ensemble member. . . . .                                                                                                                                                                                                               | 122 |

|                                                                                                 |     |
|-------------------------------------------------------------------------------------------------|-----|
| 5.15 Domain-averaged melting rate of cloud water (kg/kg/s) for each<br>ensemble member. . . . . | 123 |
|-------------------------------------------------------------------------------------------------|-----|



# List of Tables

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.1 | Summary of ensemble performance across different precipitation thresholds using SAL score analysis. HR = High Resolution. . .                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 83  |
| 5.1 | Adimensional factors for the perturbed microphysical parameters used in the 2-MOM Italian ensemble experiment at 550 m resolution. INP and CCN represent scaling factors for particle number concentrations (for CCN there are 4 values: 6 clean maritime air, 7 intermediate, 8 continental, 9 polluted), NU defines the shape parameter of the hydrometeor particle size distribution, and GR is a scaling factor that modifies the velocity of graupel sedimentation. The full ensemble consists of all possible combinations of these values (total of 36 members). . . . . | 96  |
| 5.2 | List of microphysics ensemble members and corresponding Double-moment schemes (2-MOM) microphysical parameter values. .                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 97  |
| 5.3 | Qualitative summary of SAL score tendencies for the 2-MOM Italian ensemble at different precipitation thresholds. Arrows indicate tendency: $\uparrow$ overestimation, $\downarrow$ underestimation, $\approx$ close to observations. . . . .                                                                                                                                                                                                                                                                                                                                   | 104 |
| A.1 | Main ICON configuration parameters, whose description can be found in paragraph A, for the three nested domains used in all experiments. Domain 1 = outer (2.2 km), Domain 2 = intermediate (1.1 km), Domain 3 = inner (550 m). . . . .                                                                                                                                                                                                                                                                                                                                         | 135 |

# List of Abbreviations

**1-MOM** Single-moment schemes.

**2-MOM** Double-moment schemes.

**AROME** Applications of Research to Operations at Mesoscale.

**COSMO** Consortium for Small-scale Modelling.

**CPM** convection-permitting modelling.

**DKRZ** Deutsches Klimarechenzentrum.

**DWD** Deutscher Wetterdienst.

**ECMWF** European Centre for Medium-Range Weather Forecasts.

**EnKF** Ensemble Kalman Filter.

**IC** Initial Conditions.

**ICON** Icosahedral Nonhydrostatic model.

**ICON-LAM** ICON limited-area mode.

**IFS** Integrated Forecasting System.

**IMKTRO** Institute of Meteorology and Climate Research.

**IQR** Interquartile Range.

**KENDA** Kilometre-scale ensemble data assimilation for the COSMO model.

**KIT** Karlsruhe Institute of Technology.

**LBC** Lateral Boundary Conditions.

**LETKF** Local Ensemble Transform Kalman Filter.

**LLJ** Low-level jet.

**MCSs** Mesoscale Convective Systems.

**MOSES** Modular Observation Solutions for Earth Systems.

**MPI-M** Max Planck Institute for Meteorology.

**NWP** Numerical Weather Prediction.

**SAL** Structure, Amplitude and Location metrics.

**SPPT** Stochastically perturbed parametrisations technique.

**WRF** Weather Research and Forecasting.

# 1 | Introduction and motivation

## 1.1 | Extreme precipitation predictability in a changing climate

Extreme precipitation events represent an urgent challenge worldwide, with the Mediterranean basin experiencing an increasing frequency of exceptionally intense and devastating floods (Cui et al., [2024](#); Jézéquel et al., [2025](#)). These events are exacerbated by record-high sea surface temperatures, which provide additional moisture and instability to the atmosphere, fuelling deep convection (Pastor et al., [2015](#); Strobach et al., [2024](#)). The catastrophic floods of early autumn 2024 in Eastern Europe, Italy, and Spain (Cadena SER, [2024](#); Reuters, [2024](#)) have underscored the need for improved numerical weather prediction systems capable of accurately forecasting both the intensity and precise location of extreme rainfall events.

Traditional numerical models struggle to capture the fine-scale processes that drive convective storms, often underestimating peak precipitation or misplacing the most intense convective cores (Capecchi, [2021](#); Hernández et al., [2025](#); Prein et al., [2015](#)). This limitation arises due to the coarse resolution of many operational global-scale weather models, which still rely on grid spacings of several kilometers and require deep convection to be parameterized rather than explicitly resolved. While recent advances in high-resolution convection-permitting modelling (CPM) have significantly improved precipitation forecasts (Clark et al., [2016](#); Fosser et al., [2024](#)), their grid spacing is still often limited to around 2–4 km. This resolution may not be sufficient to accurately reproduce the dynamics of highly localized and short-lived extreme convective storms (Hanley et al., [2015](#)) as for one of the case analysed in this work, that is the devastating flood that hit the Marche region in Italy in September 2022 (Marche Floods ArcGIS StoryMaps, [2023](#)).

### **1.1.1 The use of ensembles and high-resolution to tackle the uncertainty**

The predictability of extreme convective precipitation remains one of the most challenging aspects of modern meteorology. Indeed, the atmosphere is a highly chaotic dynamical system and it is characterized by a high sensitivity to initial conditions, as it was demonstrated by Lorenz in his work on deterministic chaos. It follows that small perturbations in the atmospheric state can grow very rapidly with time. This aspect, then, leads to inherent loss of deterministic predictability that neither strong and structural advances in numerical models formulation or observational networks can effectively overcome (Lorenz, 1963, 1969). Moreover, despite substantial advances in both deterministic and ensemble-based Numerical Weather Prediction (NWP) systems, the chaotic nature of the atmosphere, coupled with the strong non-linearity of convective processes, limits the lead time at which such events can be reliably predicted (Richardson, 2000). This uncertainty is further compounded by the rapid growth of small-scale perturbations in the initial conditions, which can have a disproportionate impact on the timing, location, and intensity of convective storms. In recent years, the adoption of convection-permitting ensembles has significantly improved short-range forecasts of severe weather. By explicitly representing deep convection, these ensembles are able to provide probabilistic guidance not only on the likelihood of occurrence but also on the potential severity and spatial distribution of high-impact events. This probabilistic information is a cornerstone of modern early warning systems, enabling forecasters and decision-makers to balance false alarms against missed events, and to better communicate risk to the public (Cloke & Pappenberger, 2009).

### **1.1.2 The impact of deep convection and the role of early warning systems**

Early warning systems benefit greatly from high-resolution ensemble approaches, as they allow for the quantification of forecast confidence and the identification of scenarios leading to extreme impacts. However, the effectiveness of such systems depends critically on the integration of ensemble outputs with observational data, rapid-update cycles, and impact-based forecast frameworks. For

localized storms that can potentially produce flash floods, reliable warnings require both timely detection of convective initiation and accurate projection of storm evolution, tasks that remain difficult even at hectometric resolutions (Doswell et al., 1996; Golding, 2009). To this issue, here the potential of hectometric scale ensemble forecast is explored at different horizontal scales using simulations performed with the Icosahedral Nonhydrostatic model (ICON).

Such high-resolution simulations, achieved by nesting approaches, aim to resolve fine-scale convective processes more accurately, including updraft and downdraft dynamics, microphysical interactions, and storm organization mechanisms. The ability to explicitly resolve these features is crucial for improving the predictability of extreme rainfall events and mitigating their societal impacts (Lean et al., 2024). Despite the scientific advances offered by hectometric-scale modelling, several challenges remain. The high computational cost, the handling of initial and boundary conditions, and the need for advanced post-processing and verification methods are critical factors in determining its operational feasibility. However, the potential benefits of more accurate and reliable severe weather forecasts could lead to a paradigm shift in the prediction of localized hazards.

Deep convection, moreover, is a key driver of extreme weather phenomena, particularly flash floods, severe winds, hail, and tornadoes. Its formation and evolution depend on the interplay of thermodynamic instability, moisture availability, and dynamic forcing, often concentrated in mesoscale environments (Houze, 2004; Weisman & Klemp, 1984). In the Mediterranean region, deep convection is frequently triggered by the convergence of moist, warm air from the sea with cooler, drier air masses from continental Europe or upper-level intrusions, creating strong uplift and rapid cloud development.

Once initiated, deep convective systems can organize into a variety of structures, including supercells, multicell clusters, and Mesoscale Convective Systems (MCSs), each with distinct hazards and life cycles (Markowski & Richardson, 2010). Orographic effects, such as those of the Apennine chain, further modulate convection by enhancing lifting on the windward slopes and influencing the downstream evolution of precipitation systems. This orographic modulation can lead to quasi-stationary convection, where repeated cell regeneration over the same area results in extreme local rainfall totals, as in the September 2022 Marche flood.

From a forecasting perspective, the impacts of deep convection are challenging to anticipate because small errors in simulating storm initiation or movement can lead to large discrepancies in predicted rainfall accumulations (Doswell et al., 1996). Microphysical processes add another layer of complexity: the growth, melting, and evaporation of hydrometeors strongly influence rainfall intensity and distribution. At hectometric resolution, models can better resolve the turbulent exchanges and convective updrafts that sustain heavy precipitation, but even so, uncertainties remain in representing cloud–precipitation conversion rates and storm longevity.

Given the societal impacts of deep convection, ranging from damage to infrastructure and agriculture to threats to human life, improving our predictive capabilities is a pressing priority. This requires not only advances in NWP resolution and physical parameterizations but also the continuous refinement of verification metrics and data assimilation strategies to better capture the rapid, multiscale processes that define these events.



Figure 1.1: The streets of Senigallia (Marche, Italy) following the flooding caused by the overflowing of the Misa river after the huge rainfall that hit central Italy on 15th September 2022.

## 1.2 | The importance of high-resolution NWP models

The increasing frequency of extreme high-impact precipitation events has driven the demand for improving accuracy of weather forecasts, thus the recent development and operational use of NWP models with ever finer spatial resolution was essential to achieve this task (Caldas-Alvarez et al., 2023). As extreme phenomena such as heatwaves, intense rainfall, and convective storms become more frequent and persistent, NWP systems need to resolve the underlying meteorological processes at increasingly small scales (Torma & Giorgi, 2024). This is particularly true in regions such as Europe, where strong orographic gradients, coastlines, and land–sea contrasts exert a pronounced influence on convection, turbulence, and microphysics (Médus et al., 2022). In such environments, hectometric-scale modelling becomes an essential tool to better represent the physical mechanisms that drive hazardous weather.

Moving towards high-resolution simulations allows for a more realistic depiction of atmospheric processes: the structure and intensity of up- and downdrafts, the dynamics of cold pools, mesoscale organization of convective systems, and the microphysical evolution of hydrometeors (Henderson et al., 2021). By reducing reliance on deep-convection parameterizations and resolving a larger fraction of the convective spectrum, models can improve the timing and spatial structure of rainfall features. However, pushing to hectometric grid spacing (Glazer et al., 2025; Lean et al., 2024) also brings new challenges, including stricter numerical stability requirements, more pronounced sensitivity to terrain representation, and the need for carefully tuned sub-grid closures (Bryan et al., 2003; Clark et al., 2016; Johnson et al., 2018).

### From kilometre to hectometric scales and the “grey zone”

CPMs with grid spacings of 1–4 km have already demonstrated substantial benefits for forecasting convective precipitation. However, many Mediterranean flash-flood cases are controlled by physical processes occurring on scales below one kilometer, such as the sharpening of convergence lines, narrow orographic ascent, or the propagation of cold pools.



This introduces the so-called “grey zone” of convection: a resolution range of the same order of these models, where deep convection schemes are often deactivated, since convection begins to be explicitly simulated, but the model grid is still too coarse to fully resolve convective updrafts (Honnert, 2016; Honnert et al., 2016). In this regime, mid- and deep convection parameterizations are usually switched off, while shallow convection may remain parameterized. The grey zone is therefore a hybrid representation, in which neither parameterizations nor explicit resolution provide a fully consistent description of convective processes, often leading to enhanced sensitivity to model choices and turbulence closure (Clark et al., 2016).

Pushing further into the *hectometric scale* (100–900 m) allows a growing portion of the convective dynamics to be explicitly represented (Lean et al., 2024; Nuissier et al., 2020). This includes the organization of convective cells and the fine-scale interactions with land–sea contrasts and orographic features. Grid spacings in the range of 500 m are particularly relevant for current research, as they represent a transition between conventional kilometre-scale CPMs and large-eddy simulations. While this regime can yield significant gains in realism, it also amplifies model sensitivity to the choice of turbulence schemes, surface representation, and boundary-layer processes, and comes at a steep computational cost. Moreover, hectometric domains are typically embedded within larger, coarser convection-permitting parents. To avoid artefacts at nest boundaries, best practice involves: (i) ensuring temporal consistency in Lateral Boundary Conditions (LBC) updates, (ii) applying consistent orography filtering across all nests, and (iii) using two-way nesting only when the feedback is expected to be beneficial and numerically stable. Phenomena such as coastal jets, mountain waves, and valley-channelled flows are especially sensitive to the interplay between LBC update frequency, numerical diffusion, and topographic representation (Ahrens & Leps, 2021).

## **Turbulence and diffusion at CP scales**

At convection-permitting resolutions, turbulence schemes directly influence the moisture distribution, the intensity of cold pools, and the ability of storms to remain organized. Common approaches include boundary layer closures for turbulent kinetic energy, which are robust and computationally efficient, and

three-dimensional Smagorinsky type schemes (Deardorff, 1980; Smagorinsky, 1963), which tend to better capture shear-driven, anisotropic mixing and can produce more LES-like behaviour at hectometric grid spacings.

In complex terrain, the interaction between the dynamical core, diffusion settings, and orography filtering can strongly affect both windward precipitation enhancement and leeward suppression or displacement (Chow et al., 2019).

### **Microphysics: single vs double moment schemes**

The representation of hydrometeors is crucial for simulating intense convective rainfall. Single-moment schemes (1-MOM) predict only the total mass of each species, while 2-MOM also predict number concentrations, enabling the size distribution to evolve dynamically. This additional information allows for a more realistic treatment of autoconversion, accretion, evaporation, and melting processes (Milbrandt & Yau, 2005; Morrison et al., 2009).

At hectometric resolution, 2-MOM schemes often provide improvements in storm timing, localization, duration and intensity, but they also require greater computational resources and can be more sensitive to assumptions about aerosols and particle shapes (Madhulatha et al., 2023; Taufour et al., 2024).

### **Convective-scale ensembles and uncertainty representation**

Even at 550 m resolution, a single deterministic forecast cannot capture the full range of uncertainties associated with convective extremes: thus, convective-scale ensembles tackle this limitation by introducing spread through initial and boundary condition perturbations, stochastic physics, or multi-physics approaches that vary turbulence and microphysics schemes like in this work.

The benefits of probabilistic guidance for high-impact weather are well documented (Bowler et al., 2008; Gebhardt et al., 2011). At hectometric scales, however, the high computational cost may lead to the specific request of efficient ensemble design. Thus, combining Initial Conditions (IC)/LBC perturbations with targeted physics perturbations, eventually, may be more effective than a large ensemble relying solely on IC variation.

## Data assimilation at convective scales

The assimilation of high-frequency observations, which include radar reflectivity and Doppler wind data, satellite radiances, lightning detections, and dense surface network measurements, has emerged as a cornerstone for improving the initialization and short-term evolution of convection systems in CPMs (Gustafsson et al., 2018). The flexibility of radar and surface observations, particularly those updated on rapid cycling schemes, helps to capture the evolving moisture and dynamic gradients central to convective initiation. Ensemble-based methods, such as the Ensemble Kalman Filter (EnKF), have demonstrated their value in convective-scale data assimilation (Gastaldo et al., 2021; Schraff et al., 2016). Early experiments assimilating simulated Doppler radar observations showed that the EnKF can reliably estimate the state of a supercell, including unobserved variables like updraft strength and temperature perturbations, thereby supporting its use in operational nowcasting systems. Lightning data assimilation has also gained traction. The implementation of a 1D and 4D-Var approach within the Weather Research and Forecasting (WRF) model data assimilation system has achieved remarkable improvements in precipitation forecasts, reducing root mean square errors during the assimilation window and maintaining gains for several hours afterward. In parallel, variational methods remain a core component of operational convective-scale assimilation. The European Centre for Medium-Range Weather Forecasts (ECMWF), for instance, has adopted 4D-Var for decades and more recently enhanced its performance by coupling it with ensemble information in hybrid 4D-EnVar (Barker & Clayton, 2011). This framework allows the assimilation of a wide variety of high-resolution satellite and conventional observations while maintaining flow-dependent background error covariances essential for mesoscale structure representation (Rabier et al., 2000).

## Verification for CP models: from points to objects

Traditional verification scores based on point-by-point comparisons often penalize even small spatial mismatches between observed and simulated convective features, which can be misleading in high-resolution forecasts (Ebert, 2008). To better account for the inherent spatial uncertainty in convection, more suitable

approaches include methods that consider the spatial scale of the error or the overall structure of the precipitation field. For example, neighbourhood-based metrics such as the fractions skill score (Roberts & Lean, 2008) evaluate model skill over areas rather than individual grid points, while object-based techniques (Davis et al., 2009) allow errors to be broken down into components like for the Structure, Amplitude and Location metrics (SAL), e.g. Wernli et al. (2008). For hydrological applications, these methods can be further complemented by diagnostics focused on exceedance areas, event duration, or peak intensity.

## **Operational implications and open challenges**

Despite the clear advances offered by convection-permitting models, several challenges remain: timing errors in convective initiation, sensitivity to surface states and coastal effects, biases in cold pool dynamics that influence back-building or deep convection, and persistent difficulties in leeward precipitation placement over complex terrain. In addition, the computational and data handling demands are substantial. Ultimately, a careful choice of model setup, balancing timestep, diffusion, orography smoothing, turbulence representation, and microphysics complexity, offers the best path to improving forecasts for the most dangerous convective events.

### **1.2.1 Reference CP systems in research and operational activities**

In the last decade, several convection-permitting NWP models have become reference tools for both research and operational forecasting of extreme events. Here the main characteristics of the most used models in current activities are described.

#### **AROME**

The Applications of Research to Operations at Mesoscale (AROME) model is a non-hydrostatic, convection-permitting model developed by Météo-France, designed specifically for fine-scale weather prediction with a typical grid spacing

of 1.3 km over France and other domains (Seity et al., 2011). AROME couples state-of-the-art physical parameterizations for microphysics, turbulence, and surface exchanges, and is used operationally by Météo-France and several European national meteorological services, including in the HARMONIE–AROME accord for pan-European domains (Bengtsson et al., 2017). AROME has shown significant improvements in forecasting precipitation extremes and convective phenomena compared to coarser models (Sahlaoui et al., 2020).

## WRF

The WRF Model is one of the most widely used community models for mesoscale and high-resolution meteorological research (Powers et al., 2017). Developed by NCAR and partners, WRF supports a wide variety of physical schemes, flexible domain nesting, and grid spacings down to hectometric scale. It is highly configurable, supporting research applications from urban weather to tropical cyclones, and is used operationally and for ensemble prediction in many research and weather centers worldwide (Skamarock et al., 2019). Its extensive user community and broad validation have made WRF a benchmark for intercomparison and development of new physical parameterizations.

## ICON

The ICON model is the current state-of-the-art numerical weather prediction and climate modeling system developed jointly by the Deutscher Wetterdienst (DWD) and the Max Planck Institute for Meteorology (MPI-M), in close collaboration with the Karlsruhe Institute of Technology (KIT) and the Deutsches Klimarechenzentrum (DKRZ) (Prill et al., 2022; Zängl et al., 2015). ICON employs an icosahedral triangular grid that avoids the pole problem of latitude–longitude grids, allowing for efficient simulations across scales with seamless nesting. Its non-hydrostatic dynamical core and advanced physics parameterizations make it suitable for both global climate applications and convection-permitting regional forecasts. Since 2015, ICON has been the operational forecasting model at DWD, replacing COSMO, and is also employed for ensemble prediction and climate research worldwide. In particular, convection-permitting configurations with grid spacings down to 1–3 km are now routinely

applied for short-range forecasting over Germany and Europe, showing substantial improvements in the representation of deep convection and precipitation extremes compared to parameterized approaches (Heinze et al., 2017; Singh et al., 2021).

## UKV

The UK Met Office operates the UM (Unified Model) in several configurations, from the global to the UKV domain (1.5 km resolution) for the British Isles and a global convective-scale configuration. The UKV uses the Unified Model dynamical core and advanced microphysics, turbulence, and land-surface schemes, providing detailed forecasts of convective precipitation and severe weather (Petch et al., 2020).

In summary, the widespread adoption of high-resolution in convection-permitting NWP models such as AROME, WRF, COSMO, and UKMO has been a real game changer for our ability to predict extreme precipitation and convective storms. By better capturing the actual physics of the atmosphere and explicitly simulating convective processes, these models have made forecasts of high-impact events much more reliable than in the past, especially in challenging environments like the Alps or the Mediterranean (Luu et al., 2022). Of course, there are still many open questions and practical challenges, particularly when it comes to the huge computational resources required, the handling of uncertainties, and the post-processing of massive high-resolution datasets.

### 1.2.2 Summary and choice of the model

Despite all this progress, there's always a need to push things a little further, both to improve operational forecasts and to deepen our understanding of the mechanisms behind severe weather. That's exactly why ICON, a natural evolution of the COSMO model, comes into play in this work (Prill et al., 2022). The ICON model, whose detailed description is given in chapter 2, stands out for its flexibility, modern physical parameterizations, and, above all, its ability to run simulations at hectometric scales over complex domains.

Moreover, the choice was also influenced by the collaboration with ARPAE-

SIMC (The Hydro-Meteo-Climate Service of ARPAE Emilia-Romagna) and ITALIAMETEO (the Italian national agency for meteorology and climate) that started using the ICON model at high-resolution for their operational activities since 2023.

By choosing ICON, we can take advantage of its particular grid structure and ensemble capabilities to really dig into the fine-scale processes driving convective storms, something that is essential for both research and practical forecasting in regions such as central Italy and, overall, Europe.

In this thesis, the use of ICON allows us to explore not just the predictability of extreme events, but also the range of uncertainties and sensitivities that exist at these scales. Ultimately, pushing the boundaries with tools like ICON means getting closer to forecasts that truly reflect the complexity of the atmosphere, that is an essential step for reducing risk and improving warnings as our climate keeps changing.

### 1.3 | An example of challenges in storm predictability: The sinking of the Bayesian yacht

The sinking of the Bayesian yacht on 19th August 2024 in Italy offshore of Porticello, Sicily, is an explicative case of the forecasting difficulties and potential risks that occur when powerful convective storms form over the Mediterranean sea. Here we report a brief analysis of the event as a background for this study to highlight the main features and challenges of high-resolution modelling for convective storms. On that day, around 2 a.m., a thunderstorm producing observed extremely high wind gusts up to 24 m/s hit the yacht, leading to the sinking of the ship and to the death of seven of the 22 people that were on board. Unfortunately, no direct witnesses of the event were present, due to the fact that it happened during night hours off the shore.

Moreover, there wasn't a good radar coverage of the area at that time, since the only radar placed in Sicily was not working and the storm formed over an area of the Tyrrhenian sea that is uncovered by the radar observing system itself.



This led the author of this thesis and other colleagues to investigate the potential predictability of this event and making an effort to understand whether the sinking was due to a downburst or a tornado forming from the thunderstorm (De Martin et al., 2025).

Specifically, according to the collected material and observed data, it seems that the sinking was likely due to a downburst, because convective storms had a duration shorter than one hour; also the path of the storm, that was pretty linear, was not in agreement with that of a right mover supercell. In addition, the videos taken along the coast by some witnesses did not show any evidence of rotating winds or waterspouts. Lastly, no increase in pressure or drop in potential temperature were recorded by the weather stations that were present nearby.

### **Predictability of the event**

In the days and hours leading up to the Porticello windstorm, the operational ICON suite at ARPAE-SIMC offered a picture that, even though with uncertainty, already hinted at the right hazard in the right marine sector. Specifically, this study analyses the deterministic and ensemble configurations at 2.2 km horizontal resolution (ICON-2I and ICON-2I-EPS). These configurations use IFS boundary conditions and analyses generated by the Kilometre-scale ensemble data assimilation for the COSMO model (KENDA) system which uses a Local Ensemble Transform Kalman Filter (LETKF) (Hunt et al., 2007; Schraff et al., 2016). The data assimilation routinely incorporates conventional observations, as well as radar reflectivity and radial wind data from the Italian radar network (Gastaldo et al., 2021, 2022). Also, the runs exploited the ECMWF Integrated Forecasting System (IFS) model for lateral boundary conditions.

Seen in sequence, the forecasts tell a coherent story. The ICON-2I-EPS cycle of 21:00 UTC on 17th August 2024, roughly twenty-nine hours before the sinking, already lifts a flag over the southern Tyrrhenian: probabilities in the 10%-25% range for gusts stronger than 24.5 m/s, with very few members pushing beyond 32.7 m/s in the 01:00–02:00 UTC slot on 19th August. This is not yet a precise forecast, but it frames the problem: a convective wind threat in the correct corridor and time window. The deterministic run initialized at 00:00 UTC on 18th August then adds narrative detail, marching a storm eastward across the



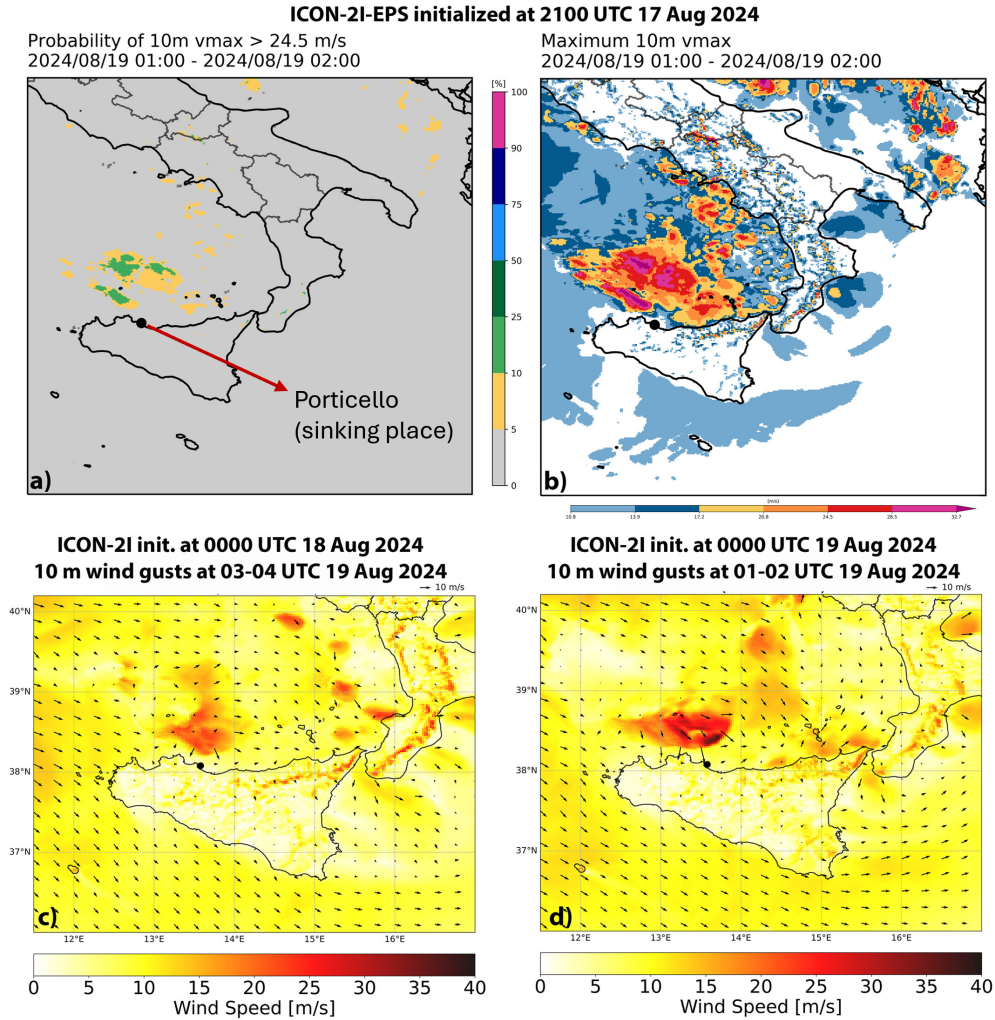


Figure 1.2: (a), (b) show respectively the Probability of maximum wind gusts at 10 m stronger than 24.5 m/s and the maximum wind themselves as forecasted by the ICON-EPS system, while (c) and (d) illustrate the wind gusts at 10 m as forecasted by the deterministic run of the model ICON-2I operationally run at ARPAE-SIMC for the bayesian sinking. Initialization and valid time are reported on the figures. Pictures produced by the author of this thesis and co-author Thomas Gastaldo, from De Martin et al., [2025](#).

same sector and driving a gust front toward Porticello a couple of hours late relative to observations, with peaks near 25 m/s. Finally, the 00:00 UTC 19th August cycle, too close to be actionable as advance warning, locks onto the synoptic-mesoscale phase with sharper timing and stronger offshore gusts (locally up to 40 m/s).

From the perspective of this thesis, the sequence is illustrative of how predictability typically manifests for short-duration but high-impact Mediterranean windstorms. First, ensemble guidance widens awareness: it quantifies a non-negligible chance of severe convective winds even when exact track and timing remain elastic. Second, a single high-resolution deterministic run clarifies the scenario but may carry modest phase errors and non realistic extremes. Third, a closer initialization can reconstruct the parent storm credibly, yet with limited lead-time value for warnings. Put together, these ingredients would have supported impact-oriented messaging for severe convective winds over the southern Tyrrhenian. Moreover, it highlights the necessity of improving the population's awareness towards the risks represented by these powerful storms, also by developing instruments and frameworks that could produce accurate early warning systems. Indeed, even though the ICON model used at ARPAE-SIMC was able to reproduce the storm in advance, at least 24h in advance, the warning in play was not enough to avoid such a high number of casualties: on 19th August 2024 a general thunderstorm alert was in force from the Italian Civil Protection for the northern Sicilian coastline, signalling that thunderstorms were expected in the area. Nevertheless, as proved by this event, issuing precise, location-specific warnings for severe convective storms remains a well-known challenge across Europe (Doswell, 2015). This accident therefore highlights the need for a dedicated severe-storm warning framework, as advocated by Miglietta and Rotunno, 2016, designed to cover both offshore and inland environments. In particular, an early-warning chain leveraging real time radar data could plausibly have played a crucial role in preventing loss of lives. Indeed, satellite analysis indicated that the storm initiated about three hours before the sinking and followed a coherent trajectory throughout its lifetime, leaving a sufficient window for alerting.

## 1.4 | Context and open challenges

Over the past decade, CPM has reshaped our ability to represent extreme precipitation and convective storms. Operating at kilometer-scale resolutions, these models have substantially reduced the reliance on deep-convection parameterizations and provided new insight into storm initiation, mesoscale organization, and rainfall extremes. Yet, as the preceding sections have outlined, major challenges remain when forecasts aim at the localized, high-impact events that most affect society.

A first source of difficulty lies in the so-called grey-zone described in section 1.2 where key atmospheric processes are neither fully resolved nor entirely subgrid. At hectometric resolutions, turbulence and microphysics parameterizations interact in scale-dependent ways with the dynamics explicitly captured by the model. Schemes originally developed for coarser scales may become overly diffusive or misrepresent the anisotropy of turbulent mixing, while more advanced closures introduce their own sensitivities and computational demands. Microphysics adds further complexity: the choice between single- and double-moment schemes, assumptions about particle size distributions, and sensitivity to aerosols can all substantially alter the life cycle of storms. These issues illustrate how progress in resolution does not eliminate the need for parameterizations but instead transforms the problem into one of scale-awareness and consistency across regimes (Morrison et al., 2009; Seifert & Beheng, 2006).

Complex terrain represents another major source of uncertainty: indeed, orographic forcing can enhance precipitation on the windward side of mountains while suppressing or displacing it leeward, and these effects are strongly modulated by diffusion settings, orography filtering, and the interaction between resolved and parameterized flows. Even small shifts in storm tracks or convergence lines can determine whether flooding occurs in one valley or in another. This sensitivity is particularly critical in Mediterranean regions such as central Italy, where steep gradients and coastal interactions combine to produce flash floods that are both severe and difficult to anticipate (Dallan et al., 2023).

Finally, the representation of uncertainty remains a central obstacle. Deterministic forecasts, even at 500 m resolution, cannot capture the full spectrum of possible outcomes in a convective environment. Ensemble approaches are

indispensable for assessing predictability, yet their cost grows dramatically with resolution. Designing efficient ensemble strategies, balancing initial and boundary-condition perturbations with stochastic or multi-physics methods, has therefore become a crucial research question (Frogner et al., [2022](#)). At the same time, verifying such high-resolution ensembles requires object-based approaches capable of distinguishing timing, structure, and displacement errors, rather than relying solely on point-wise metrics.

These scientific and practical challenges define the broader context of this thesis. The work presented here seeks to clarify how far hectometric-scale ensembles can improve the predictability of extreme convective events, to disentangle the roles of resolution, physical parameterizations, and perturbations, and to test these questions in real-world cases such as the catastrophic Marche 2022 flood and other high-impact events analysed in this thesis like the hail event that hit the Baden-Württemberg region on 14th august 2023. In this way, the thesis positions itself at the intersection of model development, process understanding, and the urgent societal need for more reliable warnings of flash flood-producing storms.

## **1.5 | The Marche 2022 flooding event as a motivation**

The devastating flood that struck the Marche region of central Italy on 15 September 2022 represents a paradigmatic example of the challenges outlined above. Within just a few hours, a quasi-stationary convective system released more than 400 mm of rainfall in narrow valleys of the Apennine foothills, causing catastrophic flooding, widespread damage, and tragic loss of life. The event was characterized by a V-shaped mesoscale convective system that repeatedly regenerated along the orographic boundary, leading to intense rainfall over the same catchments. Despite its extraordinary severity, the event was not well anticipated by operational forecasts, which underestimated both the intensity and the persistence of the convection (Marche Floods ArcGIS StoryMaps, [2023](#)).

This case highlights several open problems in convective-scale predictability. The storm initiation was strongly conditioned by mesoscale moisture convergence and orographic lifting, processes that are highly sensitive to model

resolution, diffusion settings, and initial and lateral boundary conditions. The persistence of the system depended on the interplay between different air masses convergence, storm regeneration, coastal and turbulent interactions, all of which remain difficult to capture even at kilometer-scale resolutions. Moreover, the displacement of precipitation towards leeward valleys illustrated the acute sensitivity of outcomes to orographic representation and microphysical feedbacks. In operational models, these mechanisms contributed to errors in both the timing and placement of the heaviest rainfall.

For these reasons, the Marche 2022 flood provides a compelling motivation for the present thesis. It embodies the scientific and societal stakes of improving convective-scale forecasts: on the one hand, the need to better understand the physical mechanisms that sustain such extreme storms; on the other hand, the practical need to provide more reliable warnings in regions highly exposed to flash-flood risk. By using the ICON model at hectometric resolution, and by employing ensemble strategies that account for ICs, LBCs, and physics uncertainties, this work aims to shed light on the predictability limits of such events and to explore modelling approaches that could enhance preparedness for future extremes.

## 1.6 | Thesis outline and objective

This thesis investigates the predictability of extreme convective precipitation over complex terrain using hectometric scale ensembles with the ICON model. The aim of the work is trying to capture the ingredients that play a crucial role in the correct prediction of extreme precipitation, from initial and boundary conditions, to the use of the correct physical parameterizations for turbulence and convection, combined to very high resolution grids with ensemble approaches, that are essential for forecasting such complex and chaotic events. This is why the work combines sensitivity experiments and multi-ensemble design to disentangle the roles of resolution, initial and boundary conditions uncertainty, turbulence closures, and microphysical assumptions. Two case studies that present many similarities and differences are analysed: (i) a large, long lived flood-producing event over central Italy (Marche, September 2022) in complex terrain, and (ii) a shorter lived, hail-prone thunderstorm over southwestern Germany (August 2023) in the Black Forest–Swabian Jura region.

The following sections of this thesis are framed as follows:

- **Chapter 2: The ICON model for high-resolution simulations**

Introduces the ICON model and its configuration for limited-area, convection-permitting simulations: descriptions of the unstructured icoashedral grid, handling of initial and boundary conditions, two-way nesting strategies adopted for hectometric-scale domains can be found here. The chapter details the physical parameterizations used, including the Smagorinsky turbulence closure, and compares single- and double-moment microphysics schemes. Special attention is given to computational aspects and the practical implementation of ICON for the case studies examined in this work.

- **Chapter 3: Studying flooding meteorological events (Marche 2022)**

Focuses on the devastating September 2022 flood in central Italy: it reconstructs the large-scale and mesoscale meteorological setup that led to the event and presents radar and in-situ observations. The chapter describes the design of preliminary high-resolution sensitivity tests compared to operational runs from a deterministic 2.2 km setup down to 550 m resolution, to investigate the impact of nesting, orography, and turbulence closures on the predictability of this exceptional storm.

- **Chapter 4: The role of hectometric-scale ensembles**

Explores the added value of ensemble forecasting at hectometric scales. Different ensemble configurations are analysed, from a 2 km resolution setup to double-nested 500 m ensembles, using both the 1D COSMO-type TKE and the 3D Smagorinsky turbulence schemes. Verification includes spatial and object-based methods (such as the Overlap Ratio and SAL scores) across multiple thresholds, complemented by diagnostics of low-level convergence and moisture transport. The results highlight where resolution and turbulence representation enhance the characteristics of precipitation features, while also quantifying the dominant contribution of initial and boundary condition uncertainty.

- **Chapter 5: Impact of microphysics perturbations**

Focus shifts from initial conditions perturbations to *physics uncertainty*: constructs a 36-member ensemble with the ICON double moment microphysics scheme at 500 m, perturbing key microphysical controls. For the

Marche case, the ensemble statistics and SAL reveal systematic tendencies. The chapter then applies the same approach to a German hail case (Villingen–Schwenningen, 14 August 2023), showing how microphysical choices modulate hail-relevant processes. A domain averaged process-rate analysis connects ensemble skill to the intensity and timing of hail growth patterns.

- **Chapter 6: Conclusion and future perspectives**

Summarises the main outcomes of the study, discussing the dominant role of ICs uncertainty, the benefits of hectometric resolution and advanced turbulence closures, and the influence of microphysics on convective predictability. It concludes with implications for ensemble design, forecasting systems, and future research directions toward next-generation convection-permitting simulations.





## 2 | The ICON model for high-resolution simulations

### 2.1 | ICON Model Description

The ICON model (version 2.6.6) is a non-hydrostatic NWP system developed collaboratively by the DWD, the MPI-M, the DKRZ, the KIT and the Swiss National Supercomputing Centre (Prill et al., [2022](#); Zängl et al., [2015](#)). ICON operates on an unstructured icosahedral grid, allowing seamless scalability from global to hectometric-scale simulations, making it especially well-suited for high-resolution convection-permitting forecasts. The governing equations describe a two-component system consisting of dry air and water (Gassmann & Herzog, [2008](#)), with prognostic variables including horizontal and vertical wind components, virtual potential temperature, total air density, and mass fractions of different water species (e.g., cloud water, ice, rain).

The flexibility of ICON's grid structure and its ability to perform nested simulations at increasing resolutions have proven particularly effective for studying localized and extreme convective weather phenomena. Recent studies demonstrated ICON's capability in capturing critical meteorological processes, providing valuable insights into extreme precipitation events and their predictability (Heinze et al., [2017](#)).

#### 2.1.1 Model physics and governing equations

The ICON model is characterized by three principal blocks that describe the Dynamics, Tracer advection and model's Physics: according to Prill et al. ([2022](#)), the dynamical core integrates the discrete equations for fluid motions forward in time, then, the Tracer advection, which is the numerical advection scheme, essential to describe the transport of variables such as humidity and cloud



water, is being coupled along with the dynamical core to the parameterization schemes that are crucial for representation of convective processes, especially for scales too small to be resolved.

The parameterization schemes play an important role in this study, indeed these schemes are analysed to understand how they interact with convection, orography and high resolution when it comes to get a proper forecast of extreme precipitation.

As mentioned in paragraph 2.1, the model is built on prognostic variables as described by Gassmann and Herzog, 2008, in order to get a 2-component system of dry air and water. Water, indeed, is then described and allowed to be present in all the three phases: solid, liquid and gaseous (vapour). It is interesting to notice that precipitating particles are then parameterized in different ways, with two main available schemes, the single and double moments schemes. The latter ones describes also the number concentration of hydrometeors, but a more detailed description can be found in paragraph 2.1.3.

Anyway, the equations that describe momentum, mass and energy of the mixture constitute a sum of the constituent-specific equation sets. Additionally, the equation set for the mixture is derived by introducing a reference velocity in the governing equations (Wacker & Herbert, 2003), which then defines the specific form of the governing equations themselves.

Thus, first of all the reference velocity is being defined as:

$$\mathbf{v}_b = \frac{\sum_k \rho_k \mathbf{v}_k}{\sum_k \rho_k}$$

where

- $\rho_k$  is the partial density of constituent  $k$
- $\mathbf{v}_k$  is the advective velocity of constituent  $k$

Hereafter, a density weighted averaging is applied, where every specific field  $\phi$  is being decomposed into a density-weighted mean and a deviation, allowing a separation of perturbations from the mean flow. This procedure is taken from Hesselberg (1925).

Given a generic scalar variable  $\phi$ , its mass-weighted average is defined as

$$\hat{\phi} = \frac{\overline{\rho\phi}}{\bar{\rho}}, \quad (2.1)$$

where  $\bar{\rho}$  is the Reynolds-averaged total density, and the overbar denotes a standard Reynolds average. The variable  $\hat{\phi}$  thus represents the density-weighted mean value of  $\phi$ , and this decomposition naturally leads to an expression of the form  $\phi = \hat{\phi} + \phi''$ , where  $\phi''$  represents the fluctuation relative to the mass-weighted mean.

This formulation is particularly relevant when dealing with turbulent flows and subgrid-scale processes, since it separates the mean state from unresolved fluctuations, allowing for a cleaner coupling with turbulence and cloud microphysics parameterizations. A more detailed theoretical foundation for this approach is presented in Zdunkowski and Bott (2003).

The ICON model solves the Hesselberg-averaged equations under the shallow-atmosphere approximation. The prognostic variables include the horizontal velocity component normal to the triangular mesh edges, the vertical wind velocity, the virtual potential temperature, the Exner pressure, the total air density, and the mass fractions of each hydrometeor species. The resulting equation set reads as follows:

$$\frac{\partial \hat{v}_n}{\partial t} + \frac{\partial \hat{K}_h}{\partial n} + (\zeta + f)\hat{v}_t + \hat{w} \frac{\partial \hat{v}_n}{\partial z} = -c_{pd}\hat{\theta}_v \frac{\partial \hat{\pi}}{\partial n} - F(v_n), \quad (2.2)$$

$$\frac{\partial \hat{w}}{\partial t} + \hat{v}_n \cdot \nabla \hat{w} + \hat{w} \frac{\partial \hat{w}}{\partial z} = -c_{pd}\hat{\theta}_v \frac{\partial \hat{\pi}}{\partial z} - g, \quad (2.3)$$

$$\frac{c_{vd}c_{pd}\hat{\rho}_0}{R_d} \frac{\partial \hat{\pi}}{\partial t} = c_{pd}\hat{\pi} \frac{\partial \hat{\rho}\hat{\theta}_v}{\partial t} = -c_{pd}\hat{\pi} \nabla \cdot (\hat{\rho}\hat{\theta}_v\hat{v}) + \hat{Q}, \quad (2.4)$$

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho}\hat{v}) = \sum_k \sigma_k^{\text{conv}}, \quad (2.5)$$

$$\frac{\partial (\bar{\rho}q_k)}{\partial t} + \nabla \cdot (\bar{\rho}q_k\hat{v}) = -\nabla \cdot (\bar{J}_k^r + \overline{\rho q_k'' v_r''}) + \bar{\sigma}_k, \quad (2.6)$$

Equations (2.2)–(2.6) describe the fundamental dynamics of the atmosphere in the model. Equation (2.2) represents the horizontal momentum along the grid edges. In this equation,  $\hat{v}_n$  is the component of the wind velocity normal to the edge, while  $\hat{v}_t$  is the tangential component. The term  $\hat{K}_h$  corresponds to the horizontal kinetic energy along the edge, and  $(\zeta + f)\hat{v}_t$  accounts for the effects of vorticity  $\zeta$  and the Coriolis force  $f$  on the flow. Vertical motions, represented by  $\hat{w}$ , advect horizontal momentum along the vertical direction. The pressure gradient force is expressed through the Exner pressure  $\hat{\pi}$  and the virtual potential temperature  $\hat{\theta}_v$ , multiplied by the specific heat at constant pressure for dry air,  $c_{pd}$ . Finally,  $F(v_n)$  summarizes frictional or subgrid-scale forces acting on the flow.

Equation (2.3) governs the vertical momentum, describing how the vertical velocity  $\hat{w}$  evolves in time under vertical advection, vertical pressure gradients, and gravity  $g$ .

The thermodynamic evolution of the atmosphere is captured by Equation (2.4), which links the time variation of the Exner pressure  $\hat{\pi}$  to changes in the virtual potential temperature  $\hat{\theta}_v$  and the mean air density  $\bar{\rho}$ . The constants  $c_{vd}$  and  $R_d$  denote the specific heat at constant volume and the gas constant for dry air, respectively. Diabatic processes, such as heating or cooling, are included via the term  $\hat{Q}$ .

Mass conservation is ensured by Equation (2.5), which states that the change in mean air density  $\bar{\rho}$  is balanced by the divergence of the mass flux  $\bar{\rho}\hat{w}$ , with additional contributions from convective mass sources or sinks represented by  $\sigma_k^{\text{conv}}$ .

Finally, Equation (2.6) describes the evolution of the mass fractions  $q_k$  of different atmospheric constituents, such as water vapor, cloud water, and hydrometeors. The transport of these species is influenced both by resolved advection and subgrid-scale fluxes, including turbulent transport  $\overline{J_k^r}$  and subgrid-scale covariance terms  $\overline{\rho q_k'' v_r''}$ . Source and sink terms  $\overline{\sigma_k}$  account for microphysical processes, such as condensation, evaporation, sedimentation, or phase changes among hydrometeors.

In summary, these equations define the key prognostic variables of the model: the horizontal and vertical velocities ( $\hat{v}_n, \hat{v}_t, \hat{w}$ ), the Exner pressure  $\hat{\pi}$ , the mean air density  $\bar{\rho}$ , and the mass fractions  $q_k$ . Moreover, they describe the

coupled evolution of momentum, thermodynamics, and composition in the atmosphere.

The total density of the air mixture is obtained by summing the partial densities of all components:

$$\bar{\rho} = \sum_k \rho_k, \quad (2.7)$$

and each constituent has its own mass fraction:

$$q_k = \frac{\rho_k}{\bar{\rho}}. \quad (2.8)$$

In the ICON model, the index  $k$  is used to identify the individual components of the atmospheric mixture, which includes both dry air and hydrometeors. They are, specifically:

- $k = d$  dry air,
- $k = v$  water vapor,
- $k = c$  cloud water,
- $k = i$  cloud ice,
- $k = r$  rain,
- $k = s$  snow,
- $k = g$  graupel.
- $k = h$  hail (included in the 2-moment microphysics scheme).

It is worth noting that the dry air mass fraction  $q_d$  is not treated prognostically. Instead, it is diagnosed from the other species to ensure that the sum of all mass fractions remains equal to one. This constraint reflects the fundamental physical requirement of mass conservation within the mixture and allows for consistent thermodynamic diagnostics.

### 2.1.2 The Model Reference State and simplifying assumptions

Atmospheric dynamics are typically characterized by relatively small variations in thermodynamic variables with respect to a large-scale background state. For this reason, and similarly to other numerical models, ICON adopts a formulation based on an atmospheric reference state. In this framework, thermodynamic variables are expressed as the sum of a time-invariant, horizontally homogeneous reference profile and a local deviation from it.

Any thermodynamic grid-scale variable  $\hat{\psi}(\lambda, \phi, z, t)$  can be written as:

$$\hat{\psi}(\lambda, \phi, z, t) = \psi_0(z) + \psi'(\lambda, \phi, z, t), \quad (2.9)$$

where the subscript 0 denotes the reference-state component and the prime denotes the perturbation (or grid-scale deviation).

Accordingly, the prognostic thermodynamic variables are expressed as:

$$\bar{\rho}(\lambda, \phi, z, t) = \rho_0(z) + \rho'(\lambda, \phi, z, t), \quad (2.10)$$

$$\bar{\pi}(\lambda, \phi, z, t) = \pi_0(z) + \pi'(\lambda, \phi, z, t), \quad (2.11)$$

$$\hat{\theta}_v(\lambda, \phi, z, t) = \theta_{v0}(z) + \theta'_v(\lambda, \phi, z, t). \quad (2.12)$$

The reference profiles  $\rho_0$ ,  $\pi_0$ , and  $\theta_{v0}$  are related through the equation of state and are hydrostatically balanced. Specifically, the hydrostatic balance is given by:

$$\pi_0 = \left( \frac{R_d \rho_0 \theta_{v0}}{p_0} \right)^{\frac{R_d}{c_{pd}}}, \quad \frac{d\pi_0}{dz} = -\frac{g}{c_{pd} \theta_{v0}}. \quad (2.13)$$

These reference profiles can be obtained by integrating Eq. (2.13), provided that appropriate boundary conditions are given. Notably, ICON uses the same reference state as the COSMO model.

In models such as ICON, the local atmospheric state may deviate significantly from a horizontally homogeneous reference. However, ICON does not apply any linearisation of the governing equations around the reference state, as it is often done in models based on the anelastic approximation. Instead, the primary motivation for introducing the reference state in ICON lies in simplifying the vertical pressure gradient term, by removing the contribution from the

base-state pressure gradient in the equation of motion:

$$c_{pd}\hat{\theta}_v \frac{\partial \bar{\pi}}{\partial n} = c_{pd}\hat{\theta}_v \frac{\partial \pi'}{\partial n}. \quad (2.14)$$

This formulation leads to a more efficient numerical scheme, particularly in the computation of pressure gradients along sloping coordinate surfaces. That said, the practical impact in ICON is limited because, by design, ICON evaluates horizontal pressure gradients on horizontal model surfaces (constant height), not along terrain-following layers. This avoids the typical errors seen in sloping-coordinate systems.

With this reference-state formulation in place, the vertical momentum equation simplifies. The vertical acceleration term in Eq. 2.3 becomes:

$$-c_{pd}\hat{\theta}_v \frac{\partial \bar{\pi}}{\partial z} - g = -c_{pd}(\theta_{v0} + \theta'_v) \frac{\partial(\pi_0 + \pi')}{\partial z} - g \quad (2.15)$$

$$= -c_{pd}\theta_{v0} \left( \frac{d\pi_0}{dz} + \frac{\partial \pi'}{\partial z} \right) - c_{pd}\theta'_v \left( \frac{d\pi_0}{dz} + \frac{\partial \pi'}{\partial z} \right) - g \quad (2.16)$$

$$= -c_{pd}\theta_{v0} \frac{\partial \pi'}{\partial z} - c_{pd}\theta'_v \left( \frac{d\pi_0}{dz} + \frac{\partial \pi'}{\partial z} \right). \quad (2.17)$$

Inserting the hydrostatic balance relation  $\frac{d\pi_0}{dz} = -\frac{g}{c_{pd}\theta_{v0}}$  into the equation gives:

$$-c_{pd}\hat{\theta}_v \frac{\partial \bar{\pi}}{\partial z} = -c_{pd}\theta_{v0} \frac{\partial \pi'}{\partial z} - c_{pd}\theta'_v \left( \frac{\partial \pi'}{\partial z} - \frac{g}{c_{pd}\theta_{v0}} \right), \quad (2.18)$$

and hence the full vertical momentum equation becomes:

$$\frac{\partial \hat{w}}{\partial t} + \hat{v}_h \cdot \nabla \hat{w} + \hat{w} \frac{\partial \hat{w}}{\partial z} = -c_{pd} \left( \theta_{v0} \frac{\partial \pi'}{\partial z} + \theta'_v \left( \frac{\partial \pi'}{\partial z} - \frac{g}{c_{pd}\theta_{v0}} \right) \right). \quad (2.19)$$

The perturbation fields  $\pi'$  and  $\theta'_v$  are computed directly from the prognostic equations, and the decomposition allows the hydrostatic contribution to cancel exactly with gravity, improving the balance and conditioning of the numerical scheme.

In the current version of the ICON model, certain simplifying assumptions are introduced concerning the treatment of mass tendencies associated with precipitation and evaporation processes. Specifically, the model does not explicitly account for atmospheric mass loss or gain due to these processes.

The term  $\sum_k \sigma_k^{\text{conv}}$  on the right-hand side of Eq. 2.5, which would represent the net mass contribution from convective parameterizations (including precipitation), is neglected. As a result, the ICON model does not include any source or sink terms in the mass continuity equation that arise from convective precipitation or redistribution of mass by convective updrafts or downdrafts. Accordingly, the continuity equation for total air density simplifies to:

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho} \hat{\mathbf{v}}) = 0. \quad (2.20)$$

At the surface, the boundary condition for the vertical mass flux is approximated as:

$$\bar{\rho} \hat{w}|_b = \sum_k \bar{\rho}_k \hat{w}_k|_b = 0, \quad (2.21)$$

which, when expressed in terrain-following coordinates, corresponds to:

$$\bar{\rho} \hat{w}|_b = -\bar{\rho} \hat{\mathbf{v}} \cdot \nabla h,$$

where  $h$  denotes the terrain height. Under this assumption, the atmospheric column is effectively treated as a closed system with respect to total mass.

Note that some mass fluxes associated with microphysical and surface processes are still retained in the model. In particular:

- Sedimentation fluxes of hydrometeors, denoted by  $\bar{\mathbf{F}}_{k,\perp}$ , represent the downward transport of condensate species (such as cloud water, rain, snow, graupel, and hail) due to gravitational settling. The subscript  $\perp$  indicates that these fluxes are evaluated perpendicular to the terrain-following surface, i.e., along the local vertical relative to the ground. This ensures that the flux correctly accounts for the orientation of sloping terrain.
- Surface evaporation fluxes, denoted by  $E_{k,\perp}|_b$ , represent the mass of condensate species evaporating or sublimating at the surface and entering the lowest model layer. Similarly, the  $\perp$  subscript indicates the component of the flux perpendicular to the terrain.

To satisfy the surface boundary condition in Eq. (2.21), which requires the net vertical mass flux of the total air mixture at the surface to vanish, it is

implicitly assumed that any local gain or loss of mass due to these processes is compensated by a fictitious vertical flux of dry air. This approach maintains the total mass balance in the atmospheric column while allowing sedimentation and surface evaporation to affect the individual constituent equations.

Away from the surface, diffusion fluxes for all hydrometeors (except dry air) are ignored. However, diffusive effects are still considered for all other species. The total mass continuity equation remains in the form of Eq. (2.20), provided that:

$$\sum_k \bar{J}_k^r = 0,$$

i.e., the sum of all diffusive mass fluxes is zero. This ensures that any sedimentation flux of a hydrometeor species (e.g., snow) is balanced by an artificial upward flux of dry air, such that the total mass within a control volume moving with the barycentric velocity remains conserved.

These assumptions can be summarized as follows:

- The model conserves the total air mass, rather than the dry air mass.
- Precipitation-related sinks and evaporation-related sources are excluded from the total mass budget.
- Changes in total mass due to phase change processes do not influence the surface pressure.

This last point becomes clearer when examining the hydrostatic pressure tendency equation, and incorporating the simplified continuity and boundary conditions from Eqs. (2.20) and (2.21), as also discussed in Wacker and Herbert (2003). Starting from the hydrostatic relation:

$$\frac{\partial \pi_0}{\partial z} = -\frac{g}{c_{pd}\theta_{v0}}, \quad (2.22)$$



and integrating vertically, we obtain the time derivative of surface pressure:

$$\frac{\partial p_s}{\partial t} = \frac{\partial}{\partial t} \left( \int_{z_s}^{z_T} \bar{\rho} dz \right) \quad (2.23)$$

$$= - \int_{z_s}^{z_T} \nabla \cdot (\bar{\rho} \widehat{\mathbf{v}}) dz \quad (2.24)$$

$$= - \nabla \cdot \int_{z_s}^{z_T} (\bar{\rho} \widehat{\mathbf{v}}) dz + \bar{\rho} \widehat{w}|_{z_T} - \bar{\rho} \widehat{w}|_{z_s} \quad (2.25)$$

$$= - \nabla \cdot \int_{z_s}^{z_T} (\bar{\rho} \widehat{\mathbf{v}}) dz + g \sum_k \bar{J}_k^r|_{z_T} - g \sum_k \bar{J}_k^r|_{z_s} \quad (2.26)$$

$$= - \nabla \cdot \int_{z_s}^{z_T} (\bar{\rho} \widehat{\mathbf{v}}) dz + g \sum_k (\bar{J}_k^r|_{z_T} - \bar{J}_k^r|_{z_s}), \quad (2.27)$$

where  $z_T$  and  $z_s$  denote the top and bottom model levels respectively. The convective surface precipitation flux is related to the total vertical flux by:

$$P_{\text{conv}}^{\text{tot}}|_b = \int_{z_s}^{\infty} \sum_k \sigma_k^{\text{conv}} dz. \quad (2.28)$$

Hence, the dynamical effects associated with precipitation or evaporation sources and sinks, particularly their feedback on surface pressure, are neglected. This effectively means that changes in total mass resulting from microphysical processes are not allowed to impact the large-scale pressure field.

### 2.1.3 Physical Parameterization Schemes

The ICON model incorporates several state-of-the-art physical parameterizations tailored for high-resolution convection-permitting simulations:

- **Radiation:** The ecRad radiation scheme (Hogan & Bozzo, 2018) offers improved accuracy and computational efficiency compared to traditional schemes such as RRTM (Mlawer et al., 1997). It explicitly represents both shortwave and longwave radiation.
- **Cloud Microphysics:** ICON includes both single-moment and double-moment microphysics schemes. The single-moment scheme predicts cloud water, rainwater, cloud ice, and graupel, whereas the double-moment scheme additionally predicts number concentrations of the same

hydrometeors also counting hail compared to the single-moment schemes. Both schemes have been extensively tested to assess their impact on the predictability of deep convective storms in this work.

- **Convection:** A bulk mass flux convection scheme parameterizes shallow, mid-level, and deep convection. For convection-permitting simulations at resolutions below a few kilometers, deep and mid-level convection schemes are typically disabled, allowing the explicit resolution of convective processes while shallow convection remains parameterized (Baldauf et al., 2011).
- **Turbulence and Diffusion:** ICON employs two turbulence schemes:
  - A Turbulent Kinetic Energy (TKE) scheme suitable for surface layer and free tropospheric turbulence in a 2D framework (Raschendorfer, 2001).
  - A Large Eddy Simulation (LES) approach based on a 3D Smagorinsky-type closure (Smagorinsky, 1963), particularly effective at hectometric scales where explicit resolution of turbulent eddies is required.

## 2.2 | Initial and Boundary Conditions

### 2.2.1 Initial Conditions

ICON can be run by exploiting initial and boundary conditions from different centres.

The first one, straight-forward way to initialize ICON is by using DWD's products. In this way, the model is directly initialized by using analysis data in the same format and grid of the model itself, usually stored in grib2 files. These data are updated every 3 hours and, if we consider the deterministic data, they are generated by means of a hybrid Ensemble Variational Data assimilation approach (Hamrud et al., 2014), while ensemble data exploit a LETKF approach (Bonavita et al., 2015; Hunt et al., 2007).

In addition, at the convective-scale, the limited-area version of ICON can also be initialized from analyses produced by the KENDA system (Schraff

et al., 2016), which provides ensemble-consistent initial conditions through a LETKF-based framework specifically designed for high-resolution regional forecasting.

The second way is by exploiting data analysis product from the IFS model of ECMWF. A crucial point when using IFS data is the conversion of variables: first of all, the names of the main retrieved atmospheric variables are quite different, that is why here a summary of the converted names used to run the ICON model from IFS variables is reported. Thus, it is essential to change the name of the IFS variables into the ICON short names.

### 2.2.2 Boundary conditions and external parameters

Limited-area simulations using the ICON-LAM framework require careful specification of LBCs and surface parameters to ensure dynamical consistency and realistic surface-atmosphere interactions. Since the limited-area domain does not cover the entire globe, the evolution of atmospheric variables along its lateral boundaries must be continuously constrained by external data. These boundary fields, available at discrete time steps (e.g., every 1, 3 or 6 hours), include wind components, temperature, humidity, surface pressure, and other relevant variables. To integrate them into the ICON-LAM simulation, the boundary data must first be remapped onto the high-resolution limited-area grid and vertically interpolated to the model's non-hydrostatic vertical coordinate. This is typically achieved using the `iconremap` and `int2icon` tools, which ensure consistent alignment of spatial and temporal dimensions across nested domains.

In addition to the dynamic LBCs, ICON requires a set of external parameter files that describe static surface properties such as orography, land-sea mask, vegetation type and coverage, soil categories, albedo, surface roughness, and others. These parameters, while temporally invariant during the simulation, are critical for representing land-surface interactions, radiative fluxes, and turbulent energy exchanges. They are generated using the `iconextpar` tool, which processes global high-resolution datasets (e.g., GLOBE for topography or ECOCLIMAP for vegetation) and interpolates them onto the ICON grid. Special attention is given to consistency among these surface fields, particularly at coast-line transitions and mountainous regions, as mismatches between orography

and pressure fields can induce spurious circulation patterns. Moreover, some land-surface properties may influence initial fields in the atmosphere through soil moisture or snow cover, making the preprocessing of external parameters a fundamental part of any limited-area ICON simulation. These files, once generated, are typically stored in NetCDF format and remain unchanged across simulation time steps, unless specific surface updates (e.g., snow assimilation or sea surface temperature changes) are applied during long-term runs.

The first ICON simulations carried out during this study use deterministic and ensemble forecast data from IFS, typically at a horizontal resolution of about 9 km, retrieved from ECMWF. These data provide IC and hourly updated LBCs. Given ICON's icosahedral grid, horizontal regridding is performed using the *iconremap* tool from the DWD icontools package (Prill, 2020), which applies Radial Basis Function-based interpolation from the regular IFS grid onto ICON's triangular grid. Vertical interpolation is managed internally by the ICON model. ICON-EU initial and boundary conditions were used for latter simulations in chapter 5, where the microphysics schemes were perturbed: in this case the initialization procedure involves a horizontal interpolation of prognostic and diagnostic fields from the coarser ICON-EU grid (6.5 km) onto the high-resolution limited-area grid (2.2 km). Since both configurations share the same icosahedral grid structure and variable definitions, no grid-type transformation or variable conversion is required.

## 2.3 | ICON nesting strategy

ICON supports both one-way and two-way nesting configurations, enabling simulations at multiple spatial resolutions within the same framework. Nesting allows high-resolution domains (children) to be embedded within coarser-resolution domains (parents), thereby improving the representation of local processes without the computational cost of refining the entire model grid.

In a typical nesting setup, each parent-child pair interacts via three key zones: a boundary interpolation zone, a nudging zone (in the case of one-way nesting), and a feedback zone (used in two-way nesting). Lateral boundary conditions for the nested domain are obtained by interpolating prognostic variables from the parent domain into the boundary interpolation zone. This

ensures that the child domain evolves in consistency with the larger-scale flow.

For two-way nesting, feedback from the child to the parent domain is activated by enabling the corresponding namelist switch (`lfeedback = .TRUE.`). In this configuration, that is the one adopted in this work, the model state on the parent domain is relaxed towards the updated, higher-resolution fields of the child at every fast physics time step. This feedback loop improves the parent simulation by incorporating finer-scale features resolved only on the nested grid. The feedback is applied to specific prognostic variables such as horizontal and vertical wind components, virtual potential temperature, density, and water vapor content, while surface fields and precipitating hydrometeors are typically excluded.

Technically, nesting in ICON is realized via a recursive time-stepping sequence, in which the global or outermost domain is advanced first. Boundary data are then passed down to each nested domain in succession. If multiple nested levels are defined, each level operates on a reduced time step (usually halved at each level), ensuring numerical stability. After advancing all nested domains, feedback is passed upward from the innermost to the outermost grid.

While one-way nesting is also possible and often used in ICON limited-area mode (ICON-LAM), its key difference lies in the origin of lateral boundary conditions: these are externally provided, typically from a global model or coarser regional simulation. In ICON-LAM, boundary updates occur at predefined intervals via input files, and interpolation is performed in time and space. The lack of feedback in this configuration makes it simpler but potentially less consistent at the domain boundaries compared to the two-way nesting approach.

Overall, ICON's nesting capabilities are designed to balance accuracy and computational efficiency, offering a flexible and scalable solution for multi-resolution weather and climate simulations.

## 2.4 | ICON Application for High-Resolution Case Studies

ICON's versatility has been crucial in investigating extreme precipitation events through high-resolution simulations and ensemble forecasting approaches. For instance, initial sensitivity tests and a subsequent ensemble of 10 ICON members were conducted to analyze the catastrophic flooding event that occurred in the Marche region in 2022. This initial ensemble utilized the first ten ECMWF ensemble members as initial conditions, applying double nesting starting from a 2.2 km horizontal resolution and refining down to 550 meters. For simplicity and readability, this resolution is occasionally referred to as "500 m". This approximation does not imply the use of different model configurations and has no impact on the interpretation of the results. This approach allowed for a detailed investigation of the mesoscale and microscale dynamics associated with the flood, demonstrating ICON's capability to capture fine-scale convective processes crucial for forecasting extreme precipitation events.

Further exploring ICON's potential, another ensemble consisting of 36 members was conducted for the same Marche event. For this ensemble, initial conditions derived from ICON itself were used, and parameters within the double-moment microphysics scheme were perturbed. Such microphysical perturbations were essential to assess the sensitivity of convective precipitation forecasts to microphysical processes and helped quantify uncertainties related to cloud and precipitation microphysics in high-resolution modeling frameworks (Seifert & Beheng, [2006](#)).

Additionally, ICON's applicability extends beyond the Italian context. Similar sensitivity tests and microphysical perturbation studies were successfully performed for a high-impact convective event that happened in Germany near Villingen-Schwenningen in August 2023 leading to the formation of a very localized and powerful storm that produced hail, using a 36-member ensemble at hectometric scales. These studies underline ICON's robust capabilities in exploring convective dynamics and uncertainties, enhancing the predictability and understanding of severe weather events across diverse geographical settings.

### 2.4.1 Model Configuration and Computational Considerations

The simulations conducted here are characterized by 65 vertical levels for all the tests and ensembles, striking a balance between computational efficiency and the adequate representation of tropospheric dynamics and moisture transport. The model top was set at 22 km, sufficient to capture relevant stratospheric interactions and their potential influences on tropospheric processes. A two-way nesting approach is applied, where the resolution of the nested domains is automatically defined by the ICON web-based tool by halving the horizontal resolution of each parents' domain. Thus, starting from a 2.2 km domain, two nested domains are obtained for each of the tests, with 1.1 km and 550 m resolution, respectively. Moreover, when employing the nested domains, a list of defined parameters is fixed, with some differences in some cases from domain to domain, while varying few other parameters depending on the analysed case study.

In Appendix A, the reader can find the table A.1 that reports a short list of the most important parameters as defined in the study with also an overall description of the parameters themselves.

Regarding the parameterization of turbulence, a switch was performed from the one-dimensional COSMO turbdiff/turbtrans with TKE closure turbulence scheme (corresponding to the value of 1 in the parameters' namelist) to the three-dimensional Smagorinsky scheme (corresponding to the value of 5) for the higher resolution domain depending on the experiments. Another important consideration concerns the parameter that defines the maximum height over which the topography smoother operates (`hgtdiff_max_smooth_topo`), that is, the maximum elevation (in meters) up to which the smoothing of the topography is applied. This parameter is essential, as it can lead to significant changes in precipitation forecasts when moving to high-resolution simulations over complex terrains, especially in cases of intense convection and turbulence.

This configuration exemplifies ICON's ability to balance computational feasibility with detailed process representation, confirming its suitability and effectiveness in operational and research settings focused on severe convective storms and extreme precipitation events.







## **3 | Preliminary analysis of a high impact convective storm with ICON high-res simulations**

### **3.1 | The unprecedented Marche 2022 flooding event**

On September 15th and 16th, 2022, a devastating flood struck the Marche region in central Italy, causing at least 2 billion euros worth of damage and resulting in 13 fatalities (Morelli et al., [2023](#)). The flooding was triggered by a powerful, self-regenerating V-shaped storm, which intensified as it interacted with a complex convergence of extremely warm and humid air masses from the subtropics and much colder, more stable air masses from the Atlantic Ocean and the Gulf of Lion. These systems, known for producing flash floods and generating significant hydrogeological risks, are relatively common in the Mediterranean basin (Amengual, [2022](#); Rebora et al., [2013](#)), particularly due to the prolific moisture source provided by the Mediterranean Sea and the latent heat released by the high temperatures of the sea surface. The Mediterranean Sea, in particular, has reached record-high temperatures in recent years (Marullo et al., [2023](#)), which in turn have made storms like this one even more intense. Indeed, September 2022 was marked by a series of powerful storms following an exceptional period of record-high temperatures in both the Atlantic Ocean and the Mediterranean Sea during the summer months, especially in August (NOAA National Centers for Environmental Information, [2022](#)).

The interaction between warm and moist air from the subtropics and colder air masses from the Atlantic was a key component of the storm's strength. This interaction was further intensified by the presence of a strong Low-level jet

(LLJ) transporting moisture-rich air towards the Apennine foothills, favouring convective development (Raveh-Rubin & Wernli, 2016). These mechanisms contributed to the persistence of intense precipitation over a localized area, a characteristic often associated with V-shaped storms. Extreme meteorological events like this are becoming increasingly common as the global climate continues to warm, exacerbating the energy available to weather systems and contributing to the persistence of blocking patterns, which often lead to extreme weather conditions in certain areas (Kautz et al., 2022). In fact, these blocking regimes have been shown to increase the frequency of extreme weather events, especially in areas that are already vulnerable to heavy precipitation and flooding.

Moreover, the summer of 2022 saw the Azores High shift significantly eastward, positioning itself between Spain and Morocco. This anomaly in the typical positioning of the Azores High meant that it persisted in a location that trapped warm and humid air across much of southern Europe, setting up a prolonged period of high pressure in the region. This setup favored the build-up of intense thermal gradients, which helped fuel storms in the central Mediterranean region, with moisture and energy being funnelled into the storm systems. The persistent positioning of the Azores High across southern Europe, combined with the cut-off low over Spain, contributed to the strong south-westerly winds that brought moist, warm air into the region, creating the perfect conditions for storm formation and subsequent flooding. Previous studies have highlighted similar atmospheric configurations leading to flash flood events in the Mediterranean. For instance, the October 2018 flood in Mallorca, Spain, and the devastating floods in southern France in October 2020 were both associated with persistent moisture convergence, strong low-level jets, and the presence of isolated low-pressure systems (Mastrantonas et al., 2020). These patterns suggest that the 2022 Marche event was not an isolated case but part of a broader trend of increasingly frequent extreme precipitation events in the region.

In Figure 3.1, the synoptic situation over Europe at 14:00 UTC on the day of the event is shown, specifically at the storm's onset, as retrieved from ERA5 reanalysis data (Hersbach et al., 2020). This figure shows temperature at 850 hPa and geopotential height at 500 hPa. We can see that a strong Azores High was positioned off the Iberian Peninsula, with a classic cut-off low centred over

Spain. This atmospheric setup generated strong south-westerly winds, which brought significant amounts of heat and moisture from the subtropics and the Mediterranean Sea. These winds impacted the central portion of the Apennine Mountain range almost perpendicularly, causing them to interact with other air masses coming from Corsica and the Gulf of Lion, which enhanced cloud formation and precipitation across the region. The region's complex orography, characterized by numerous hills and valleys in the areas just before the mountains, played a crucial role in enhancing convection and turbulence. The convergence of these warm and cold air masses, combined with the complex terrain, created significant atmospheric instability, which triggered the rapid development of intense thunderstorms, resulting in heavy and sustained precipitation over the Marche region (Tartaglione, 2025). As noted in several studies, these types of events are closely linked to specific synoptic dynamics, including the interaction of various air masses and complex orographic conditions, both of which enhance precipitation intensity and contribute to the severity of flash flooding (Amengual, 2022).

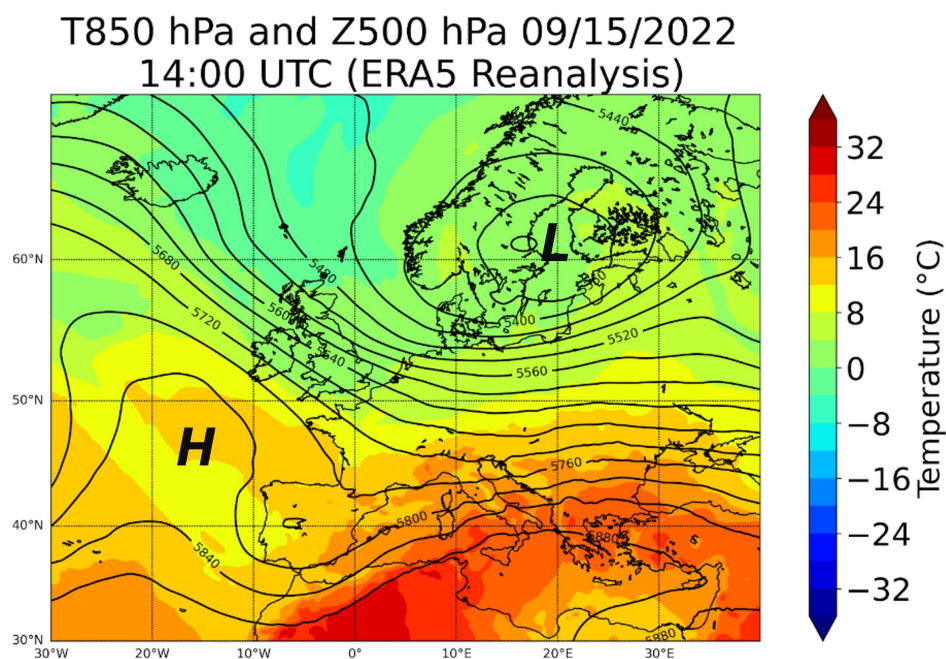


Figure 3.1: Synoptic setup at the beginning of the event, temperature at 850 hPa and geopotential height at 500 hPa are shown.

Despite the favourable conditions for the development of intense events, the main operational models struggled to accurately forecast the intensity and

localization of the storms. These models primarily underestimated the total precipitation amounts and wind intensity, which contributed to the surprise and severity of the flooding event. The observational network in the region most affected by the rainfall was relatively sparse in terms of the variety of observed variables. However, reliable land-based rain gauges, such as the one at Cantiano (whose location is shown in Figure 3.4), operated by the National Civil Protection Department of Italy, recorded over 400 mm of rainfall in just 8 hours, from 12:00 UTC to 20:00 UTC (Figure 3.2).

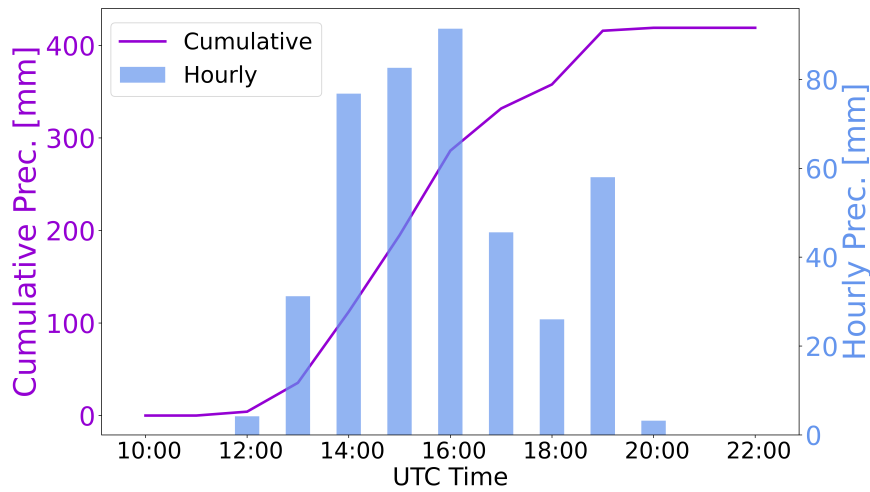


Figure 3.2: Hourly and cumulative precipitation collected by the weather station in Cantiano on 15/09/2022.

The spatial distribution of the observed precipitation, derived from the Italian radar network and adjusted with rain gauge data, is shown in Figure 3.3. The 24-hour cumulative rainfall (panel a) reveals the remarkable localization and intensity of the event, with a distinct V-shaped structure extending from northern Umbria to the central Marche region. The maximum precipitation core, exceeding 400 mm, is clearly centred along the valleys near the Apennine foothills, where orographic enhancement and low-level convergence likely played a decisive role.

The three-hourly accumulations (panels b–e) provide a clearer picture of the temporal evolution of the event. During the early afternoon (12–15 UTC), the first convective cells formed over the inland areas of Umbria and southern Marche, aligned along a quasi-stationary convergence zone. Between 15 UTC and 18 UTC, convection rapidly intensified and organized into a self-regenerating V-shaped system, with the apex located over the Cesano and

Misa river basins. This phase corresponded to the period of most intense rainfall observed at Cantiano and other nearby stations. In the following hours (18–21 UTC), the system remained nearly stationary, continuously feeding on moist air advected from the south-west by the low-level jet, which sustained intense and persistent precipitation over the same area. Finally, between 21 UTC and 24 UTC, the convective activity gradually weakened and shifted slightly north-eastward toward the Adriatic coast, where it eventually dissipated. Such sequence shows the slow-moving and quasi-stationary nature of the storm, typical of V-shaped convective systems, and confirms the crucial role of local orography and mesoscale convergence in maintaining such high precipitation rates over a limited area.

In conclusion, this event highlighted the growing difficulty for operational models to capture the physics of self-regenerating systems, especially when these systems interact with complex orographic features. It also underscored the importance of improving forecasting methodologies for extreme weather events in an era of rapidly changing climate patterns. In particular, the anomalous positioning of the Azores High, which was more eastward than usual, played a key role in shaping the conditions that led to this catastrophic flooding. Given that similar conditions are likely to occur more frequently as a result of climate change, these findings emphasize the need for more accurate and timely forecasting models to better predict extreme weather events and mitigate their impacts.

### **3.1.1 Predictability and preliminary sensitivity tests**

The case of Marche flooding was one illustrative case of how difficult it is to correctly forecast convective powerful storms like this. This is why some preliminary sensitivity tests were done first.

To explore the feasibility of high resolution runs, it was selected a first parent domain characterized by a horizontal resolution of 2.2 km including Italy and the neighbouring countries. This domain spans from 33.7° to 49° N in latitude and from 3° to 22° E in longitude. Then, by using the ICON grid file server, two nested domains were selected with a resolution of 1.1 km and 550 m respectively. The choice of these domains was a compromise between good representation of the territory, without taking too few grid points, and not

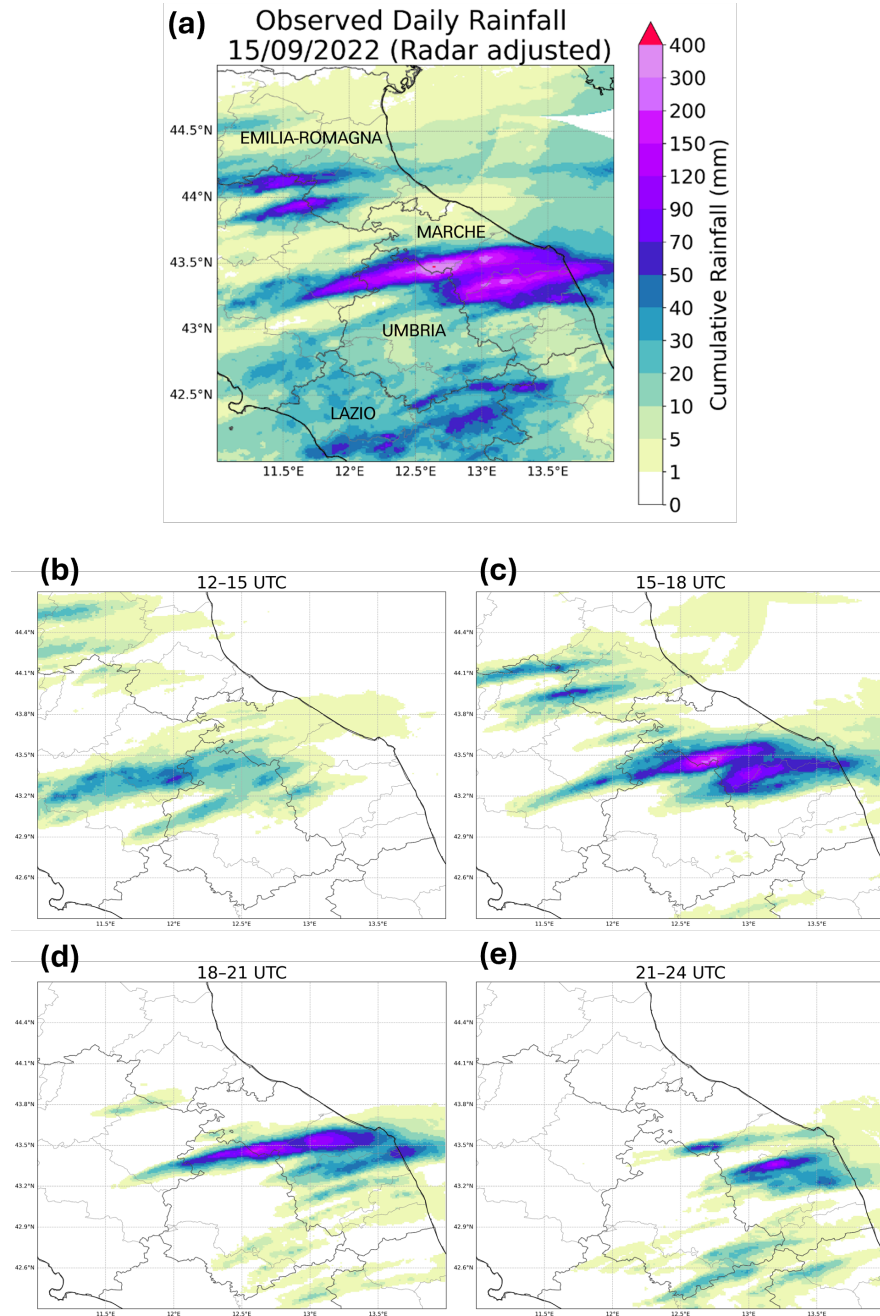


Figure 3.3: (a) 24h total cumulated precipitation as observed by the Italian radar-based network adjusted with rain gauges data. (b), (c), (d), (e) 3-hourly cumulated rainfall at different times from 12:00 UTC to 24:00 UTC.



cutting the mountains, since this may introduce errors and miscomputations while the model runs, especially in turbulent regimes. Considering the case study, with shifting south-westerly flux, it was decided to include the Corsica island in the first nested domain at 1.1 km, together with all the northern and central portions of the Apennine mountain range.

It is worth noticing that ICON was still not operational in Italy at that time, indeed the parameter namelist exploited the settings that were used at the regional meteorological agency of Emilia-Romagna ARPAE for the lately dismissed Consortium for Small-scale Modelling (COSMO) model (Marsigli et al., 2005; Schättler et al., 2021; Vela et al., 2013), which provided the basis for the physics package of the ICON model with ICs and LBCs taken from IFS (Iriza et al., 2024; Pham et al., 2021; Prill et al., 2022). The operational setup recalled the one of Domain 1 as in table A.1. Under these conditions, sensitivity tests were made, that highlighted interesting and promising features for improvement of extreme forecasting with convection-permitting models. After having identified the proper setup for the test simulations the high-resolution results were compared with IFS forecasts data considered as the benchmark.

### Whole domain comparison

Figure 3.5 shows the 24-hour accumulated precipitation over the entire Italian domain, from 00 UTC 15/09/2022 to 00 UTC 16/09/2022, comparing the observed radar network (panel a), IFS forecast data (panel b), and ICON at 2.2 km using the COSMO operational setup for the turbulence scheme (panel c). It is worth noting that the observed radar–rain gauge composite (panel a) does not cover the entire model domain, as the Italian radar network only provides reliable data over most of the national territory and nearby areas. Peripheral regions, as well as the covering of some areas of the seas and neighboring countries, remain unobserved and are plotted in gray. This results in a smaller spatial coverage compared to model fields.

The observations (panel a) reveal the extreme and highly localized nature of the event. The heaviest precipitation was concentrated almost exclusively over Marche and Umbria, with localized maxima exceeding 400 mm around Cantiano, and several bands above 200 mm extending along the Apennine ridge.



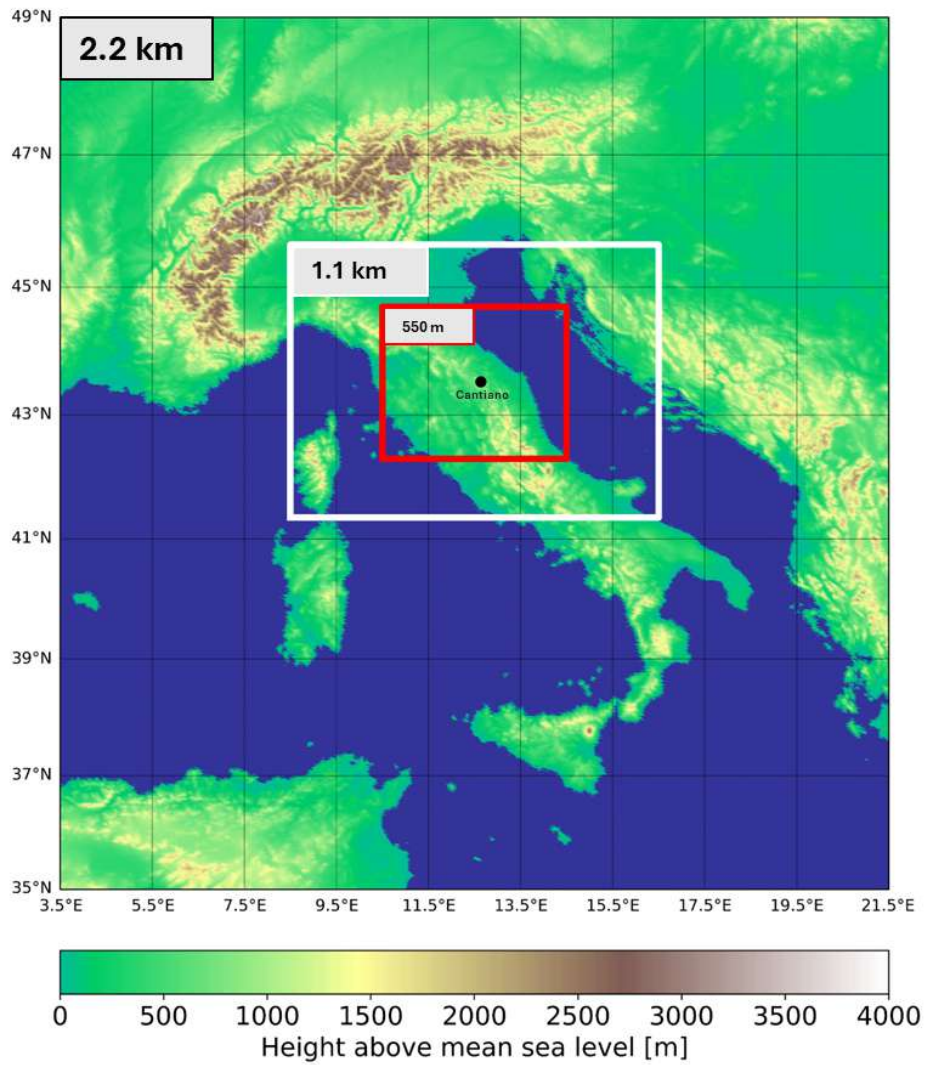


Figure 3.4: Domain of the simulations. Parent 2.2 km horizontal domain including the two nests used for the sensitivity tests and the high-resolution ensembles.

In contrast, the rest of Italy experienced relatively modest or scattered rainfall. This sharp spatial confinement clearly reflects the convective character of the storm and the strong orographic enhancement in central Italy. Readers can visually identify in the figure the intense “core” of the storm over Marche/Umbria, with the surrounding areas remaining mostly dry, highlighting the extreme localization of the event.

The IFS forecast (panel b) captures the general precipitation pattern across Italy, but misplaces the location of the heaviest rainfall. Its maxima occur primarily over Tuscany and Latium, reaching up to about 180 mm, while underestimating the extreme amounts over Marche and Umbria. In the figure, this is visible as pronounced peaks in the wrong regions, illustrating how the model spreads heavy precipitation over a broader area than observed. This mismatch highlights the limitations of coarse-resolution or global models in forecasting the precise location and magnitude of small-scale, high-intensity convective events.

The ICON simulation at 2.2 km (panel c), using the COSMO operational setup for turbulence, reproduces the general precipitation distribution in northern and southern Italy, but, similar to IFS, fails to capture the intense central storm. Only a weak precipitation signal appears over the flooded area. In the figure, one can see that the signal in Marche/Umbria is weak and diffuse, confirming that this first test with ICON and the operational COSMO setup is insufficient to reproduce the extreme event. This outcome emphasizes not only the importance of high spatial resolution and suitable turbulence and convection parameterizations, but also the critical role of accurate initial conditions.

These visual insights at the national scale motivated us to focus on the area directly affected by the flood. In the following section, here high-resolution simulations with double nesting down to 550 m are performed, which allow a more realistic representation of the thunderstorm and its associated precipitation over Marche and Umbria.

### **Focus domain comparison**

In Figure 3.6b we can observe the same simulations and observations as above, but with a focus on the area that was selected for the innermost domain. The IFS model is able to identify two interesting and distinguished areas of high

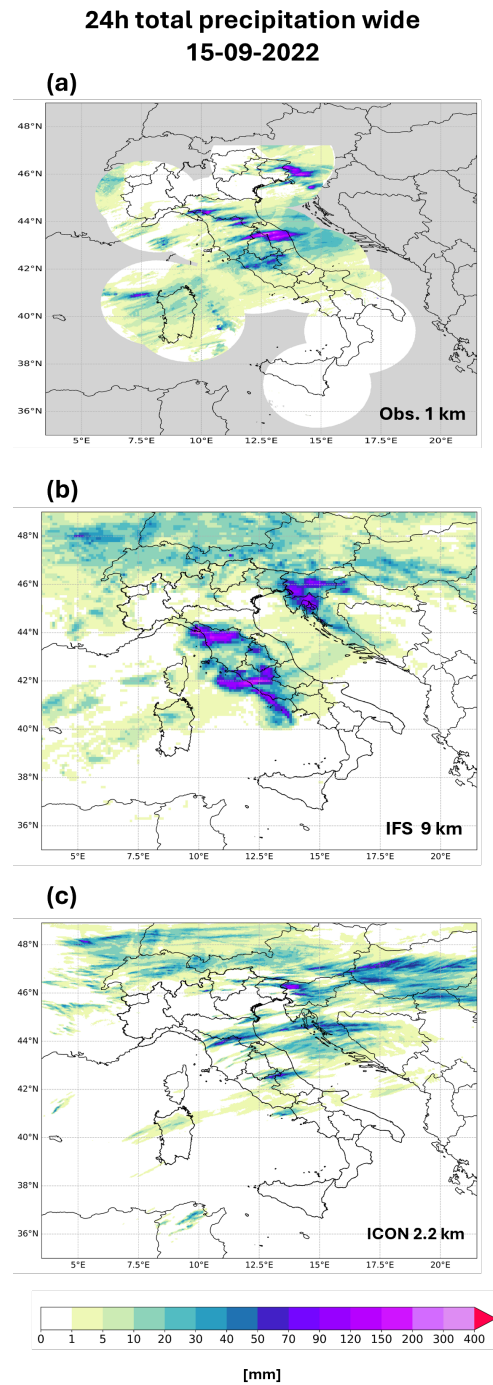


Figure 3.5: Comparison on the whole Italian domain of 24h total precipitation field from 00:00 UTC 15/09/2022 to 00:00 UTC 15/09/2022 between (a) Observed radar-rain gauge network (b) IFS forecast data, (c) ICON control run at 2.2 km with COSMO operational setup.

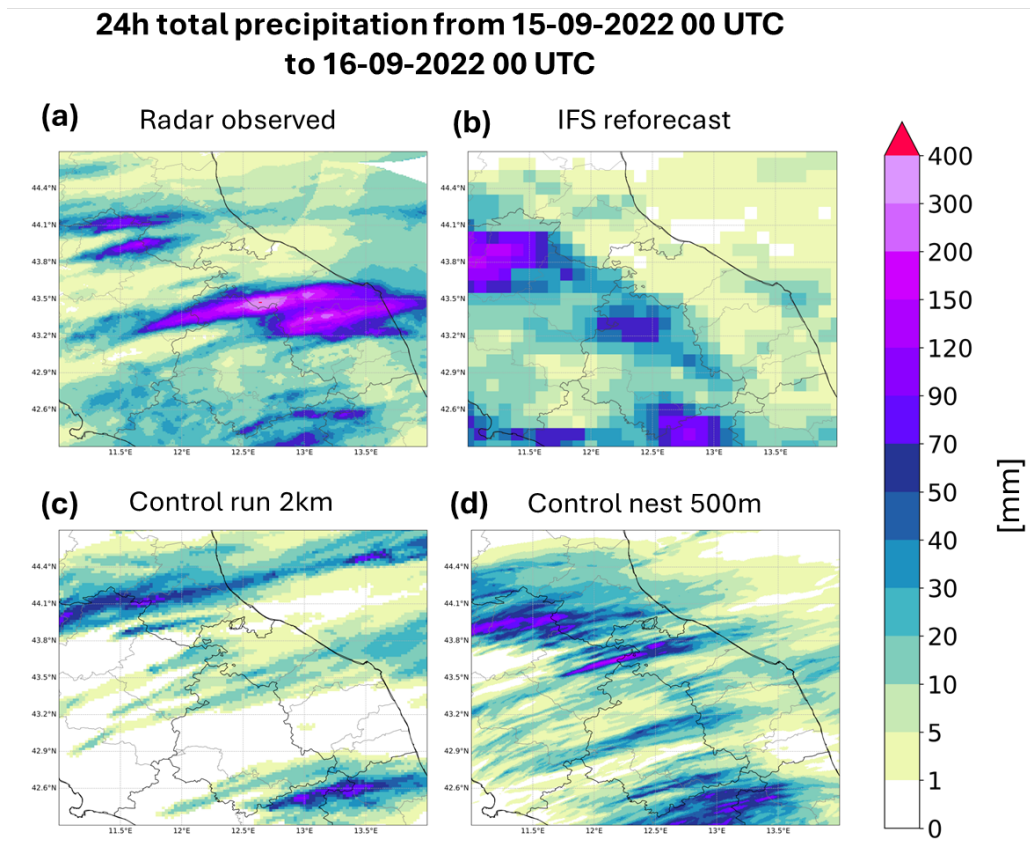


Figure 3.6: Comparison on the flooding area of 24h total precipitation field from 00:00 UTC 15/09/2022 to 00:00 UTC 15/09/2022 between (a) Observed radar network (b) IFS forecast data, (c) ICON run at 2.2 km with COSMO operational setup (d) ICON run at 550 m with Smagorinsky turbulence scheme.

precipitation as it was also shown in the whole domain comparison, in the centre—left and bottom-centre portions of the domain (respectively Tuscany and Latium) where precipitation actually reached high values, but they were overestimated by the model, that reported a maximum peak in the forecast of 131 mm over the focus area. At the same time, there was a huge underestimation of the exceptional rainfall between Umbria and Marche, with forecasted values not exceeding 70 mm in the overall area interested by the flooding event. Moreover, the model failed in capturing most of the rainfall in the leeward side of the mountains.

In Figure 3.6c we can appreciate the precipitation forecasted by the ICON model with the operational setup used for COSMO at that time, with 65 vertical levels, parameterized convection and 2.2 km horizontal resolution without nesting; also, the turbdiff/turbtrans 2D turbulence scheme is used. It is worth noticing that the ICON model was still not operational at ARPAE at that time, this is why the forecast with ICON by using the COSMO model settings that were used for the operational simulations at that time was reproduced. The model is able in this case to reproduce quite well the precipitation in the southern portion of the domain and in the northern regions, but still does not capture the thunderstorm responsible for the huge amounts of rain in central Italy, just showing a weak precipitation pattern that starts upwind of the Apennine and continues downwind towards the Adriatic sea. This is why it was decided to explore the role of different turbulence schemes and hectometric-scale simulations.

In Figure 3.6d, indeed, similar setup as for 3.6c but switch for the turbulence scheme is analysed: indeed, the Smagorinsky turbulence scheme is activated and the double nesting at 550 m resolution is used for the first time. This test gave us some preliminary useful insights: the total forecasted precipitation increases in the areas actually affected by rain as observed. A new intense precipitation pattern appears between 11.5° E and 13 ° E at a latitude included between 43.5° N and 44 ° N, suggesting that the model is now reproducing a thunderstorm close to the area hit by the flooding, even though not in the exact location and quite far from the total precipitation amounts.

Also, precipitation amounts increased in the central portion of the domain, leading to a better coverage of the observed field, giving an overall better precipitation forecast compared to the observations. This preliminary result

led us to set up ensemble simulations with double nesting down to 550 m resolution, in order to understand which features could be involved in the predictability of the event and how the interactions between high resolution, initial conditions and turbulence schemes could improve the forecast of a such unpredictable storm.





## 4 | The role of hectometric-scale ensembles

Despite the high computational cost that is usually needed to run them, hectometric-scale simulations present a huge opportunity for forecasting of deep convective events and, broader, for accurate and fine-scale predictions (Lean et al., 2024). Enhancing the resolution of the models is essential to produce adequate forecasts especially in a changing climate with soaring temperatures and frequency of extreme events. Indeed, instruments themselves are not perfect and characterized by internal limitations, moreover, human errors may add up when trying to register data, including time of recording and typing errors. Last, the overall coverage of the land surface by the instruments is pretty nonhomogeneous (Kidd et al., 2017), leading to many areas being uncovered, increasing the final level of uncertainty that plays a role when initializing the forecasts. This is why the first step was focusing on high-resolution simulations down to 550 m resolution.

However, sometimes high resolution is not enough: unpredictable storms highly depend on the weather situation and on the high sensitivity of the system to small uncertainties, that may derive also from errors on the initial conditions and model approximations. This led to the adoption an ensemble approach to evaluate the role of perturbations to the initial conditions in an innovative hectometric-scale setup for the Italian context, which gave us important information for future improvement of our meteorological forecasting network.

### 4.1 | Sensitivity to initial conditions



### 4.1.1 Ensemble forcing data

Initial and boundary conditions are taken from the members of the ECMWF IFS ensemble dataset: similarly to the data used for the sensitivity tests, the three-dimensional fields include dynamic and moisture-related variables, such as zonal and meridional wind components ( $u$  and  $v$ ), vertical wind component ( $w$ ), temperature ( $t$ ), specific humidity ( $q$ ), cloud liquid and ice water content ( $clwc$  and  $ciwc$ ), rain-origin water content ( $crwc$ ), and snow-origin solid water content ( $cswc$ ). These data are distributed across the model levels (from 31 to 137), allowing an accurate representation of the vertical structure of the atmosphere. For two-dimensional (2D) fields, the logarithm of surface pressure ( $lnsp$ ) was extracted. These data cover a spatial resolution of  $0.18^\circ$  by  $0.18^\circ$ , centered on a geographical area between  $52^\circ\text{N}$  -  $30^\circ\text{N}$  and  $-5^\circ\text{E}$  -  $25.1^\circ\text{E}$ , covering much of Europe.

Here the first 10 members from the IFS ENS which provides a diverse set of IC and BC were selected. Making use of the Stochastically perturbed parametrisations technique (SPPT) which adds random, stochastic perturbations to the tendencies of atmospheric variables (Buizza et al., 1999), highly uncertain physical processes like convection and cloud formation may benefit of a more realistic representation of possible atmospheric states. This approach ensures an equally probable set of scenarios.

### 4.1.2 LOWRES ensemble (2.2 km)

The first ensemble consists of a 2.2 km horizontal grid spacing resolution setup making use of the parameters' settings as shown in the second column (Domain 1) of table A.1. Since the nesting is not applied here, this experiment is called LOWRES (low-resolution) ensemble. This ensemble is considered as benchmark for the following high-resolution ones. As can be seen in Figure 4.1, the members show a large variability in the 24h total precipitation field at the local scale. The total precipitation amounts are particularly low compared to the observations, with just member 10 reporting accumulations exceeding 120 mm in a small area just a bit south and still upwind of the actual event.

### 4.1.3 COSMO\_tke ensemble (550 m)

Using the operational COSMO turbtran and turbdiff turbulence scheme, this first high-resolution ensemble was done with online two-way nesting between the domains: starting from the 2.2 km resolution domain, through an intermediate domain of approximately 1 km horizontal resolution, the 500 m scale was reached for the innermost grid (Figure 4.2). The spatial variability of the single members points out that the initial and boundary conditions played a crucial role in such event. Anyway, the total precipitation field improves when compared with the observation field in Figure 3.6(a), with member **m2** showing a very good match with the actual event, especially in the upwind side.

The spatial distribution of the thunderstorm in m2 recalls almost perfectly the precipitation distribution of the primary thunderstorm that hit the mountains, with broad precipitation bands exceeding 120 mm and peaks over 150 mm. Even though it is able to reproduce high precipitation in the downwind side, the second cell of the thunderstorm does still not appear, thus still there is overall underestimation of the event downwind.

### 4.1.4 SMAGO\_500m ensemble (550 m)

Finally, a last double nested ensemble was done with the activation of the Smagorinsky turbulence scheme at 550 m resolution, maintaining the COSMO turbdiff/turbtrans parameterizations for the coarser domains (2.2 km and 1.1 km) (Raschendorfer, 2016; Raschendorfer, 2001). Compared to the latter ones, the Smagorinsky turbulence scheme adopts a three-dimensional approach for the transport of moisture, while the operational COSMO\_tke closure runs in a one-dimensional approach. This, combined to a better representation of small turbulent eddies, seems to improve the formation of convective clouds and their interaction with the complex orography of the territory. Indeed, looking at Figure 4.3 we can see how the total forecasted precipitation increases in all the members, with a better match with the observed field, especially for the members that reproduce large precipitating bands.

The member m2, that is the one which better reproduced the precipitating field in the first ensemble test, still improves its spatial distribution of the rainfall, with larger bands of precipitation in the downwind side of the Apennines, that

appeared to be the main weak point of the model when forecasting this event. In fact, more rain starts to be forecasted and an important portion of it now falls in the leeward side, compared to the COSMO\_tke setup forecast. Still, the variability of the ensemble in terms of precipitating patterns does not change dramatically, with the initial conditions seeming to represent the main factor in determining the environment for the development of the devastating storm. Nonetheless, the combined role of the correct initial condition with the hectometric-scale ensemble and the activation of the Smagorinsky turbulence scheme led to an overall improvement of the predictability of this extreme precipitation event.

## 4.2 | The forecasting of wind and water vapour

In this section, the first two ensemble members are compared: they exhibited significantly different precipitation patterns, to identify the key atmospheric mechanisms responsible for these differences. As introduced in Section 3.1, low-level wind convergence emerged as a critical trigger for the initiation and maintenance of deep moist convection during the event. Member m2, which better reproduced the observed precipitation distribution, also provided a markedly different representation of the wind and humidity fields compared to member m1.

As illustrated in Figure 4.4, member m2 simulated substantially stronger low-level winds, particularly at 18:00 UTC, which marked the mature phase of the convective outbreak. These winds reached peak magnitudes of around  $30 \text{ m s}^{-1}$  at the geopotential height of 850 hPa over the central Apennines and maintained a coherent structure extending from the Tyrrhenian coast toward the inland mountain barrier. Crucially, a well-defined southerly flow originating from the central Mediterranean was evident, carrying warm and moist air masses toward the Italian peninsula. This inflow interacted with relatively cooler and drier air descending from the northwest, likely influenced by subsidence behind the Alpine ridge and air masses channeled from the Gulf of Lion.

This dynamic interaction led to the establishment of a marked wind convergence zone between  $42.5^\circ\text{N}$  and  $43.5^\circ\text{N}$ , particularly between  $12^\circ\text{E}$  and  $13^\circ\text{E}$ , which coincides with the regions of maximum precipitation accumulation in

both the forecasts and observations. The convergence was clearly identifiable from the curvature and deceleration of the wind vectors, forming a stagnation region favorable to sustained upward motion. The orographic forcing of the Apennines further enhanced this setup, promoting lifting at the leading edge of the moist inflow and anchoring convection to the terrain.

In contrast, member m1 predicted considerably weaker winds with less directional shear and a reduced southerly component. The near-surface and 850 hPa flow fields showed less curvature and little evidence of convergence, suggesting that this member failed to develop the mesoscale environment needed to support organized uplift. The wind fields remained mostly zonal and weakly divergent, leading to inefficient moisture accumulation inland and suppression of deep convective initiation.

The differences between the two members become even more evident in the specific humidity fields shown in Figure 4.5. At 850 hPa, member m2 produced a pronounced tongue of high humidity values (exceeding  $12\text{--}13\text{ g kg}^{-1}$ ) extending from the Tyrrhenian Sea into the central Apennine foothills. This moisture plume was strongly aligned with the southerly flow and converged into the low-level convergence zone described above. Such accumulation of moisture at low levels is a well-known prerequisite for convective destabilization, as it enhances moist static energy and reduces convective inhibition (Tomassini, 2020). The spatial overlap between high specific humidity and dynamically forced ascent regions is a clear indicator of the member's capacity to simulate a favorable pre-convective environment. Conversely, member m1 exhibited a more diffuse and disorganized humidity field. Although it captured general moisture advection from subtropical latitudes, the lack of coherent convergence prevented the development of narrow moist corridors and vertical humidity gradients. As a result, there was limited accumulation of water vapour inland and weaker boundary-layer moistening, both of which likely contributed to the failure to initiate deep convection inland. From a process-based perspective, the improved performance of member m2 stems from the synergistic interaction between dynamically driven convergence and thermodynamic moistening. The convergence of strong southerly winds with opposing northwesterly flows created a zone of horizontal mass accumulation and vertical lifting, while sustained low-level humidity transport ensured that sufficient water vapour was available to support convective cloud formation and precipitation production.

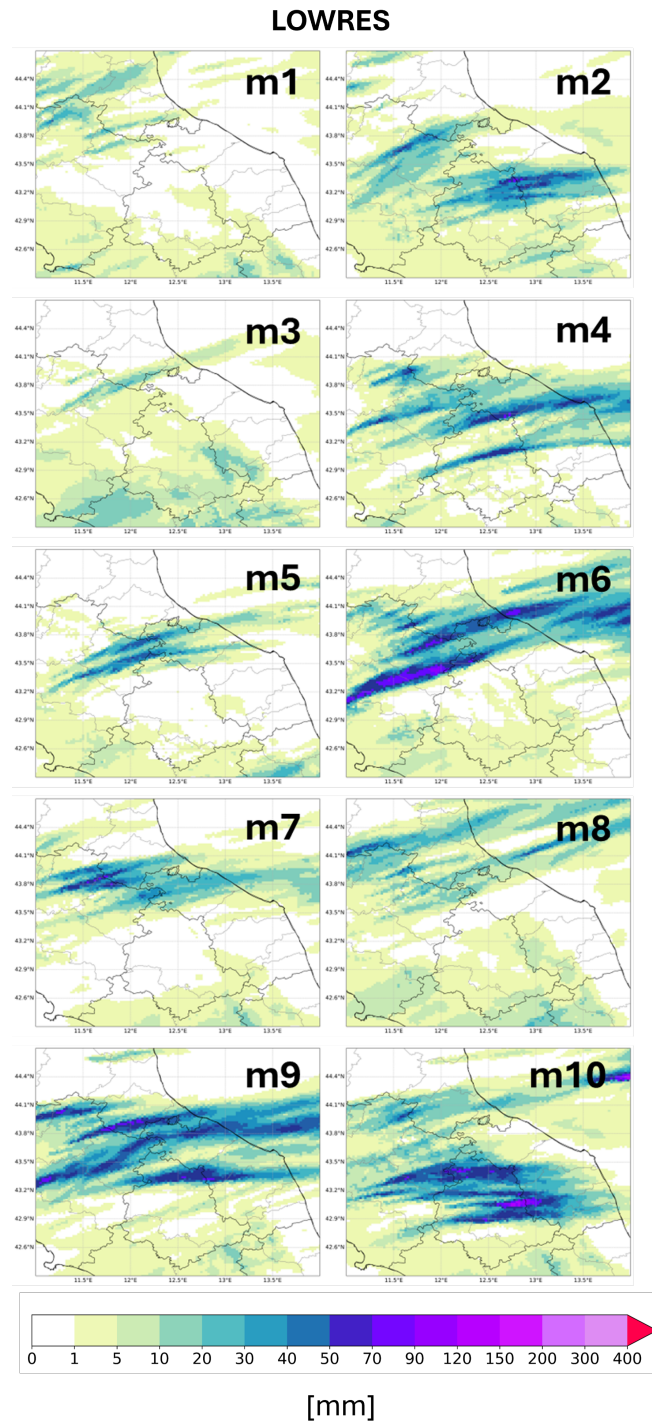


Figure 4.1: LOWRES ensemble without nested domains. Accumulated 24h total precipitation from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC.

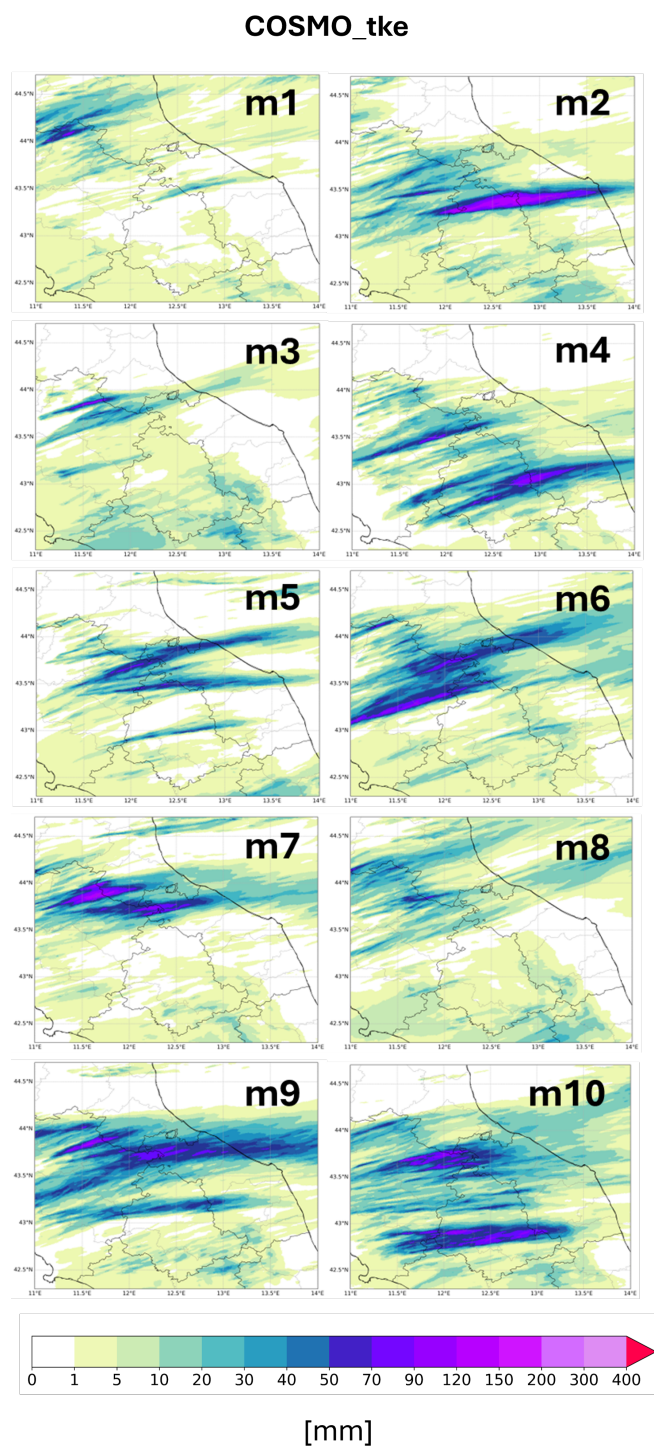


Figure 4.2: Accumulated 24h total precipitation over the 550 m grid from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC for the COSMO\_tke high-resolution ensemble with double nesting.



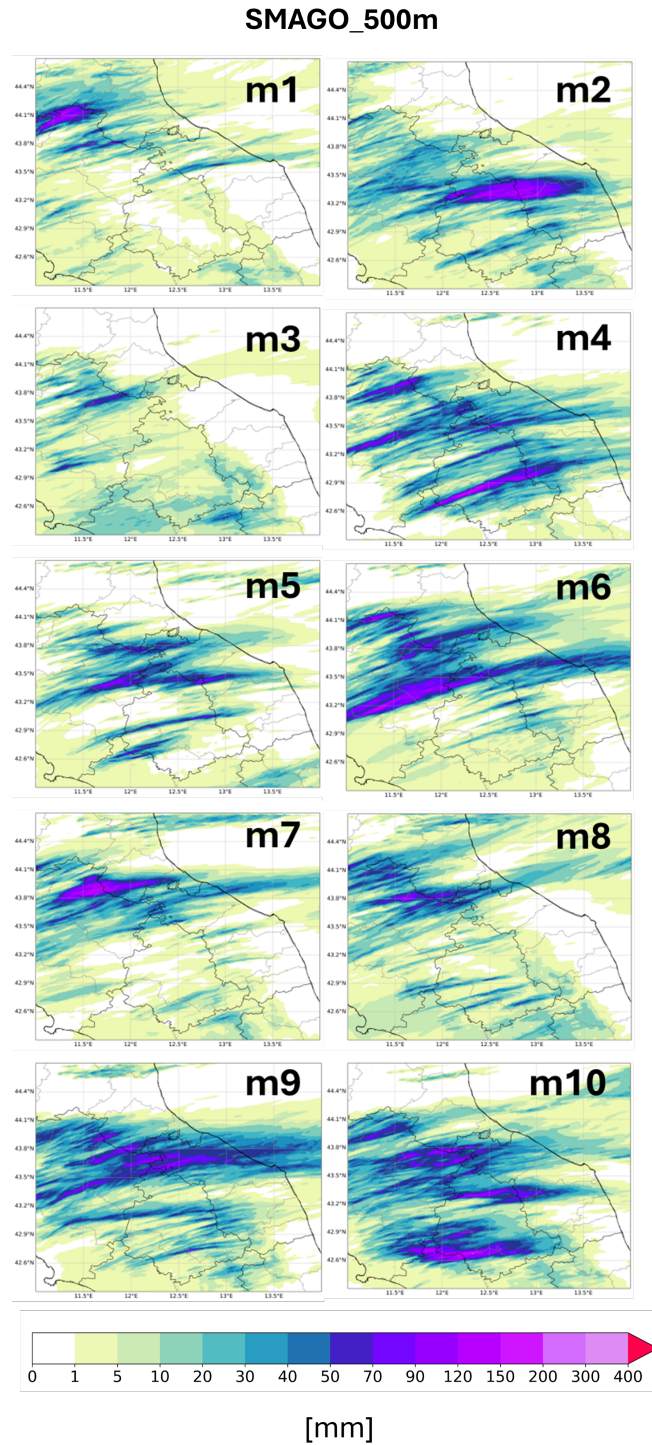


Figure 4.3: Accumulated 24h total precipitation from 15/09/2022 00:00 UTC to 16/09/2022 00:00 UTC over the 550 m grid for the SMAGO\_500m high-resolution ensemble with double nesting and activation of Smagorinsky turbulence for the innermost domain only.

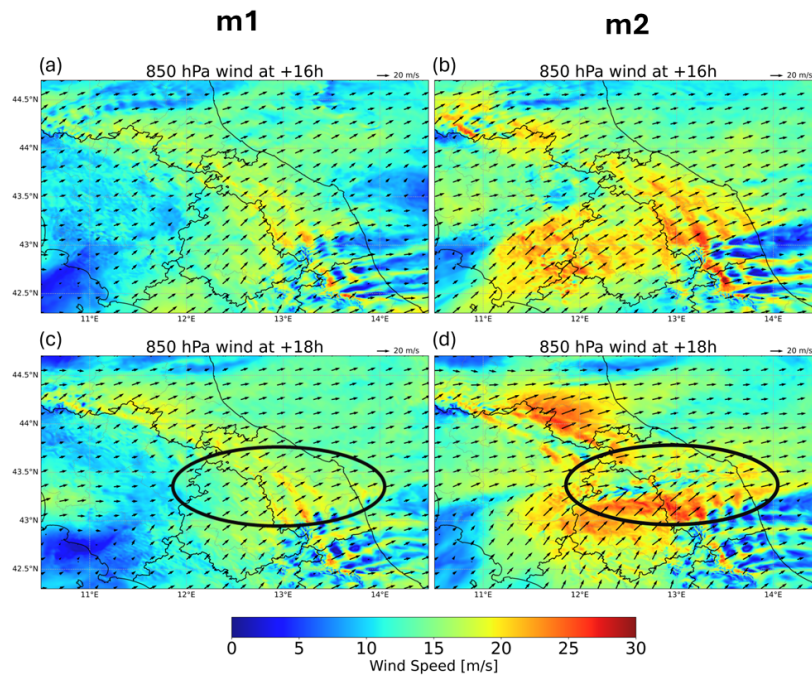


Figure 4.4: Zonal wind intensity and direction at 850 hPa as forecasted by the two members m1 (a,c) and m2 (b,d) at 16:00 UTC and 18:00 UTC respectively at 550 m resolution. Member m2 exhibits stronger and more organized southerly inflow and convergence.

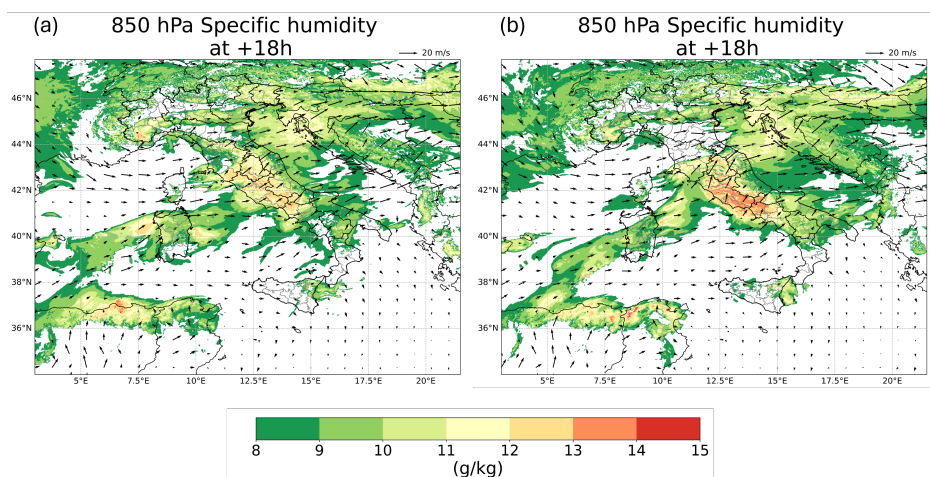


Figure 4.5: Specific humidity at 850 hPa geopotential height for the members m1 (a) and m2 (b) at 18:00 UTC of 15/09/2022.



## 4.3 | Validation of precipitation ensembles

### 4.3.1 Ensemble statistics

Figures 4.6–4.8 present the ensemble statistics for the three 10-member ensembles constructed from perturbed IFS initial conditions. These plots allow us to examine how the precipitation signal associated with the Marche flood case is represented in the different configurations, and to what extent the ensemble spread and percentiles provide useful information beyond the mean field.

Starting with the ensemble mean, all three configurations are characterized by a tendency to smooth out the extremes: the broad west–east precipitation band can still be recognized, but the maximum values barely reach 100 mm. This is a clear illustration of how averaging across displaced convective systems tends to erase the most relevant signals for risk assessment. While the mean may capture the background environment of widespread rainfall, it provides very little guidance on the potential severity of the event.

The lower bound of the distribution, represented by the 10th percentile, is essentially uninformative in this case. In all three ensembles, it highlights that a considerable fraction of members produce almost no relevant precipitation in the target area. This reinforces the idea that the system was highly sensitive to the perturbed initial conditions, with several members failing to initiate or sustain organized convection.

A more engaging picture emerges from the ensemble spread and the 99th percentile. The spread maps highlight how much members diverge in terms of location and intensity of the rainfall maxima. The LOWRES ensemble (Figure 4.6) shows comparatively modest spread, reflecting the limited variability allowed by the coarser resolution. COSMO\_tke (Figure 4.7) displays broader patterns of variability, with storm tracks shifting both northward and inland. The SMAGO\_500m ensemble (Figure 4.8), finally, shows the most pronounced spread, consistent with the increased degrees of freedom and sensitivity to the turbulence parameterization at hectometric scale. Specifically, this aspect highlights the high sensitivity of the 3D turbulence scheme to produce stronger storms when hectometric resolution is activated in combination with the correct initial conditions.

The 99th percentile maps are perhaps the most telling. In each configuration, maxima above 200 mm appear, roughly aligned with the observed flood-producing convective system. Yet the exact placement varies considerably: in LOWRES the core tends to be displaced a bit northward but still producing very low amounts, even though important peaks appear in the upwind side of central Italy Apennine range, in COSMO\_tke heavy precipitation often extends further inland into Umbria, with some precipitation starting to be advected over the Marche region, while in SMAGO\_500m the extreme values are more elongated and heavier and sometimes shifted, but still reproducing patterns close to observations. This variability underscores both the value and the challenge of using ensemble percentiles: they clearly demonstrate that catastrophic rainfall was a plausible outcome within the ensemble space, but they also reveal a high degree of uncertainty in its precise location and intensity.

Taken together, these ensemble statistics highlight a strong contrast. On one hand, the mean provides a diluted picture of little relevance for an event of this kind. On the other hand, the 99th percentile brings to light the potential for extreme rainfall, though at the cost of considerable spatial variability. For forecasters and decision-makers, this reinforces the importance of looking beyond the mean and considering ensemble spread and percentiles as the real carriers of information in high-impact convective events.

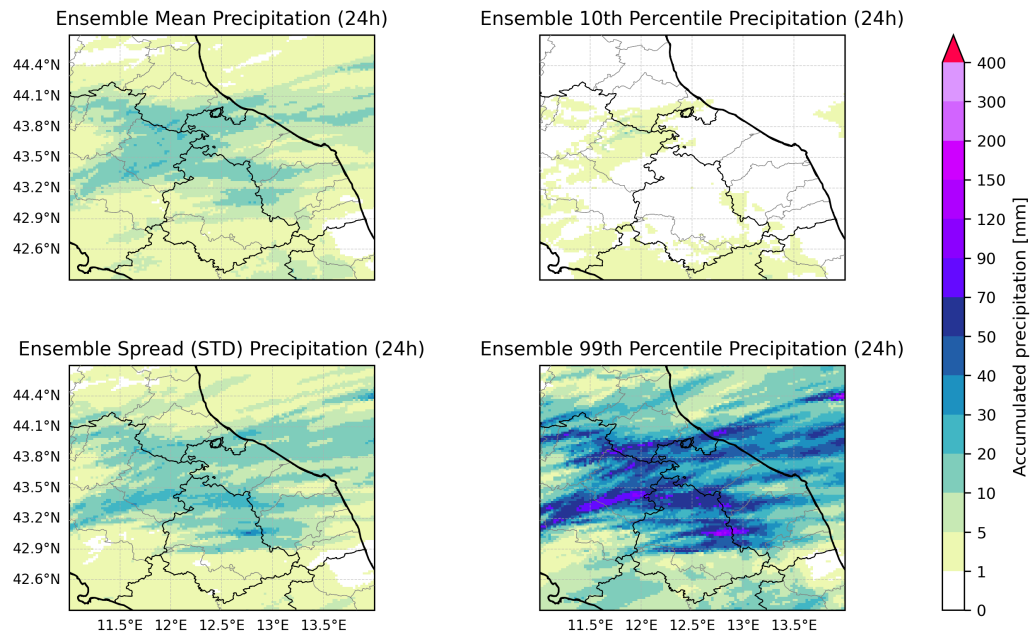


Figure 4.6: Ensemble statistics of the 10-members IC-perturbed ensemble for the LOWRES nest for the Marche flooding that happened on 15th September 2022

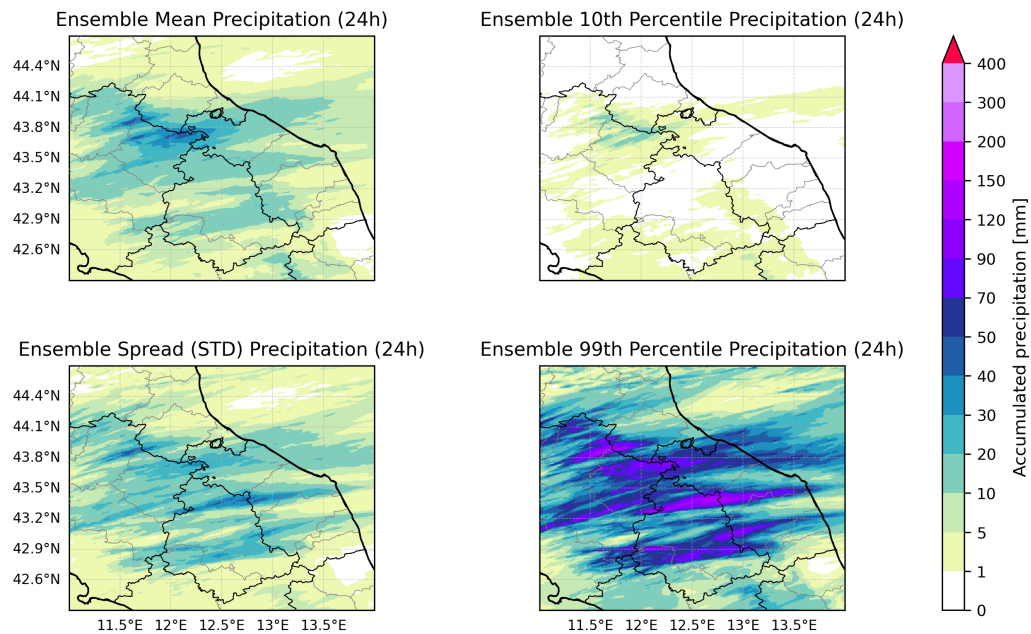


Figure 4.7: Same as figure 4.6, but for the COSMO\_tke ensemble.

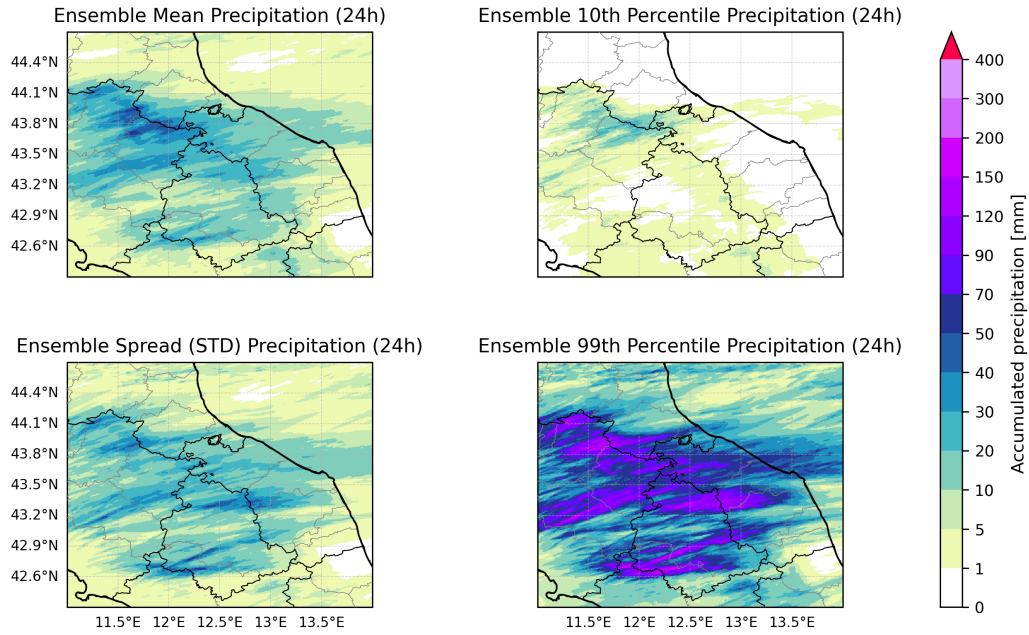


Figure 4.8: Same as figure 4.6, but for the SMAGO\_500m ensemble.

### 4.3.2 Overlap ratio

After the analysis of basic ensemble statistics, a simple yet effective version of the Overlap Ratio, designed to evaluate the spatial agreement between forecasted and observed precipitation fields is applied. This metric quantifies the number of grid points where both the forecast and the observation exceed a given precipitation threshold, providing a straightforward estimate of how well the model reproduces the location and extent of significant precipitation. Such spatial comparison techniques are widely used in meteorology and climate science to assess model skill in reproducing observed structures, particularly for high-impact weather events (Ebert, 2008; Li et al., 2015; Rezacova et al., 2015).

In Figure 4.9, the Overlap Ratio for six precipitation thresholds ranging from 10 mm to 100 mm of daily accumulated precipitation is displayed, after interpolating all forecasts onto the 1 km radar and rain-gauge based observational grid and selecting the area of the 550 m innermost domain. The analysis includes three ensemble configurations: the LOWRES control, the 550 m domain of the high-resolution COSMO\_tke, and the 550 m domain of the high-resolution SMAGO\_500m ensemble using the Smagorinsky turbulence closure scheme. A

first clear finding is that, already at the lowest thresholds, the SMAGO\_500m ensemble consistently outperforms the others in terms of spatial overlap with observations, with several members achieving values close to 40% of good spatial overlapping. This indicates that finer spatial resolution, coupled with more physically detailed turbulence parameterization, allows the model to better capture the spatial patterns of precipitation. Only a few low-resolution members, specifically m9 and m10, match or slightly exceed the SMAGO\_500m performance at higher thresholds, though this appears to be more the exception than the rule.

Particularly noteworthy is the behaviour of member m2, which stands out across all configurations. This member demonstrates a superior ability to replicate the event dynamics, including the convergence of low-level winds and the subsequent organization of convective activity. Its Overlap Ratio remains comparatively high even as the precipitation threshold increases, especially within the two high-resolution ensembles. While most members show a sharp decline in overlap beyond the 40 mm threshold, often dropping below 10% and approaching zero at 80 mm and 100 mm, m2 retains non-negligible agreement with the observations, suggesting a more realistic simulation of localized intense precipitation. These results highlight not only the added value of increased spatial resolution, but also the importance of ensemble diversity and member-specific performance in capturing extreme weather features. Metrics like the Overlap Ratio provide a practical and interpretable means of diagnosing such behaviour and can be complemented by more sophisticated object-based verification approaches such as the SAL metrics (e.g., Wernli et al., 2008) that was adopted in chapter 4.3 to further understand structural differences between modelled and observed fields.

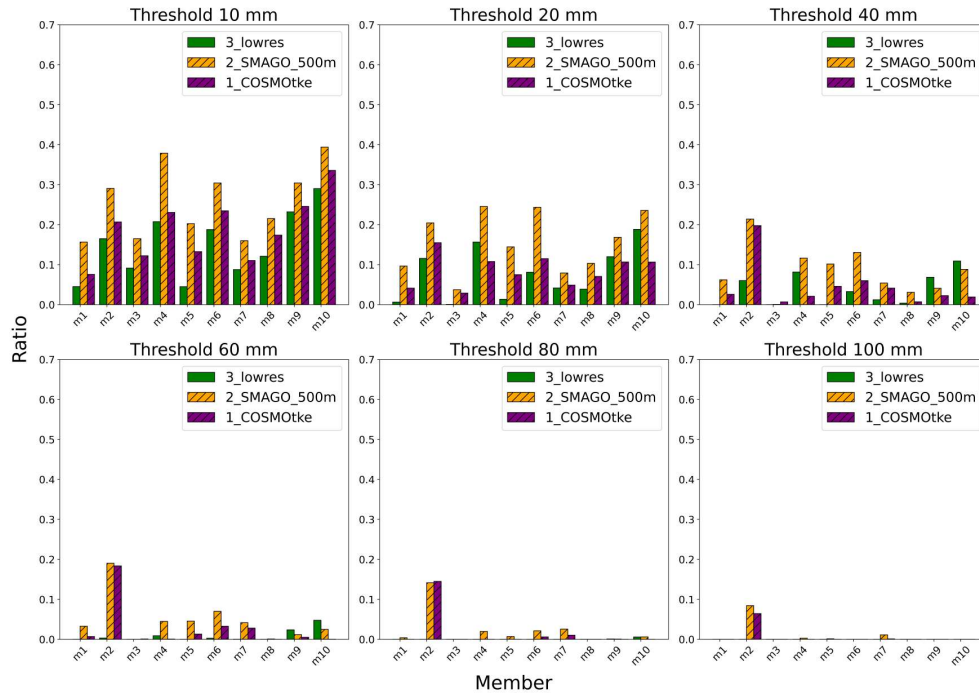


Figure 4.9: Overlap ratio between forecasted and observed total precipitation field for varying threshold and ensemble member. The three ensembles are compared.

### 4.3.3 SAL score

To better understand how ensemble and model performance depend on the intensity of precipitation, the SAL score (Wernli et al., 2008) was computed for six key 24h total precipitation thresholds: 10, 20, 40, 60, 80, and 100 mm (Figures 4.10–4.15). The SAL metric provides a comprehensive evaluation of spatial precipitation fields by decomposing model errors into three complementary components. The Amplitude (A) term measures domain-averaged differences in total precipitation amount between the forecast and the observation, indicating whether the model systematically overestimates or underestimates rainfall. The Structure (S) component quantifies how the internal organization of precipitation differs, describing whether simulated precipitation objects are too peaked or too flat compared to observations. Finally, the Location (L) term assesses the spatial displacement between observed and simulated precipitation features, combining both the distance between the centers of mass and the relative distribution of objects within each field.

Each component ranges from -2 to 2 (for A and S) and from 0 to 1 (for L), with zero representing a perfect match between model and observation. In the SAL diagrams shown in the following pictures, the grey boxes denote the Interquartile Range (IQR) range of the ensembles for the A and S values, providing a measure of spread and uncertainty among members, while the dotted lines mark the median values of the corresponding A and S components. The colour of the dotted lines represents the corresponding median of the L component, thus allowing a compact visualization of the SAL information on a two-dimensional plane.

- **10 mm threshold** (Figure 4.10):
  - All ensemble members, from every model configuration, are able to simulate objects above the 10 mm threshold.
  - The ensemble spread is large for all configurations, especially for S and A. This reflects substantial variability in how widespread, moderate precipitation is organized and quantified.
  - Members from the LOWRES ensemble tend to systematically underestimate both the structure and intensity of precipitation, leading to more negative values, and confirming a tendency for the low-resolution model to produce precipitation fields that are too diffuse and weak at this threshold.
  - COSMO\_tke and SMAGO\_500m ensembles already show some improvement: their members often cluster closer to the origin in SAL space, and some achieve lower L scores, indicating better spatial agreement with observations.
  - This highlights that at low thresholds, while all models can reproduce moderate precipitation, high-resolution and advanced turbulence schemes already offer a modest advantage in representing the structure and spatial distribution of rainfall.
- **20 mm threshold** (Figure 4.11):
  - At this slightly higher threshold, ensemble members from all configurations are still generally capable of capturing precipitation objects, but differences in performance begin to emerge more clearly.



- LOWRES continues to exhibit consistently negative S and A values, indicating an underestimation of both structure and intensity. Many members fall outside the shaded area representing the IQR.
  - COSMO\_tke and SMAGO\_500m show tighter clustering of points near the origin, particularly in A, suggesting improved representation of precipitation amount.
  - The L values remain lower in the high-resolution ensembles, confirming more accurate placement of rainfall features.
- **40 mm threshold** (Figure 4.12):
    - As the threshold increases further, fewer LOWRES members are able to simulate precipitation objects above 40 mm. The remaining points show increasingly poor SAL performance, especially in A.
    - COSMO\_tke and SMAGO\_500m retain a higher proportion of skillful members, with many falling within or close to the IQR zone in the SAL plots.
    - The S and A values continue to be better balanced, especially for the ensemble SMAGO\_500m, which shows a particularly compact distribution around (0, 0).
    - L scores remain consistently better in the high-resolution setup, suggesting that the enhanced grid spacing and turbulence representation support better placement of the precipitation features.
- **60 mm threshold** (Figure 4.13):
    - As the threshold increases, the number of members able to simulate objects above 60 mm declines, with the LOWRES ensemble being most affected: very few of its members reach this intensity, reflecting a lack of skill for intense convective rainfall.
    - COSMO\_tke and especially SMAGO\_500m retain a greater number of skillful members at this threshold. These members often cluster more tightly around  $(S, A) = (0, 0)$ , showing an improved ability to reproduce both the structure and amount of heavier precipitation.
    - The L scores are generally lower for COSMO\_tke and SMAGO\_500m members, indicating improved accuracy in the spatial placement of the most significant precipitation features.



- The IQR box narrows for all SAL components, particularly in the high-resolution ensembles, while differences between the three modelling approaches become more distinct and apparent.
- **80 mm threshold** (Figure 4.14):
  - At this threshold, the limitations of the LOWRES ensemble become even more evident, with almost all members failing to simulate objects above 80 mm.
  - COSMO\_tke continues to show a few skillful members, but the performance gap with SMAGO\_500m becomes more pronounced.
  - SMAGO\_500m members show a tight cluster near the skillful region, with several points maintaining near-zero or physically plausible S and A values, and some of the lowest L values across all configurations.
  - The 80 mm threshold serves as a prelude to the most extreme events, showing a drastic reduction in ensemble skill across all models except SMAGO\_500m.
- **100 mm threshold** (Figure 4.15):
  - At this extreme threshold, only a select few members, almost exclusively from SMAGO\_500m, and to a lesser extent COSMO\_tke, are able to simulate precipitation objects above 100 mm.
  - These high-resolution members maintain S and A values close to zero, or within a physically plausible range, indicating a realistic simulation of both the structure and magnitude of extreme rainfall.
  - Their L scores are among the lowest of all, reflecting a high degree of spatial accuracy in placing extreme precipitation features in close agreement with observations
  - LOWRES ensemble members virtually never reach this threshold, and if they do, their scores tend to be poor across all SAL components.
  - The spread among high-resolution members is still present, revealing the ongoing unpredictability inherent in convective extremes, but the median skill and frequency of accurate forecasts is far superior in SMAGO\_500m and COSMO\_tke compared to LOWRES.

Taken together, these results highlight that increasing the precipitation threshold may expose more clearly the differences in skill between the three modelling approaches. While all ensembles show similar but good skill at low thresholds, it is at the highest thresholds that the benefits of high resolution and improved physical schemes, especially as implemented in SMAGO\_500m, become more evident. Advanced configurations not only better capture the amount and structure of precipitation, but also its precise location, as quantified by this metric.

This validation underscores the importance of both model sophistication and ensemble diversity for forecasting convective extremes, and illustrates the diagnostic power of object-based verification approaches such as SAL for evaluating these challenging weather events.

| Threshold | LOWRES                                  | COSMO_tke                      | SMAGO_500m                     | Comment                                                                                               |
|-----------|-----------------------------------------|--------------------------------|--------------------------------|-------------------------------------------------------------------------------------------------------|
| 10 mm     | All members; large spread; $S, A \ll 0$ | Spread lower; some near origin | Clustered; better L            | All models capture moderate rain; high-resolution (HR) configs already improve structure and location |
| 20 mm     | Most members; $S, A$ still negative     | Tighter cluster in A; better L | Near (0,0); better consistency | Early divergence; HR models begin outperforming LOWRES                                                |
| 40 mm     | Few members skillful                    | Some skillful; A variable      | Compact near (0,0)             | HR members reproduce heavy rain better; LOWRES degrades                                               |
| 60 mm     | Very few reach threshold; poor SAL      | Mixed; some near origin        | Many skillful; good L          | Major performance gap emerges in favor of HR models                                                   |
| 80 mm     | Almost all fail; very poor scores       | Few pass; high spread          | Tight cluster; low L           | Only HR members simulate intense precipitation correctly                                              |
| 100 mm    | None or poor scores                     | Very few succeed               | Select members with good SAL   | Only SMAGO_500m consistently simulates extremes realistically                                         |

Table 4.1: Summary of ensemble performance across different precipitation thresholds using SAL score analysis. HR = High Resolution.

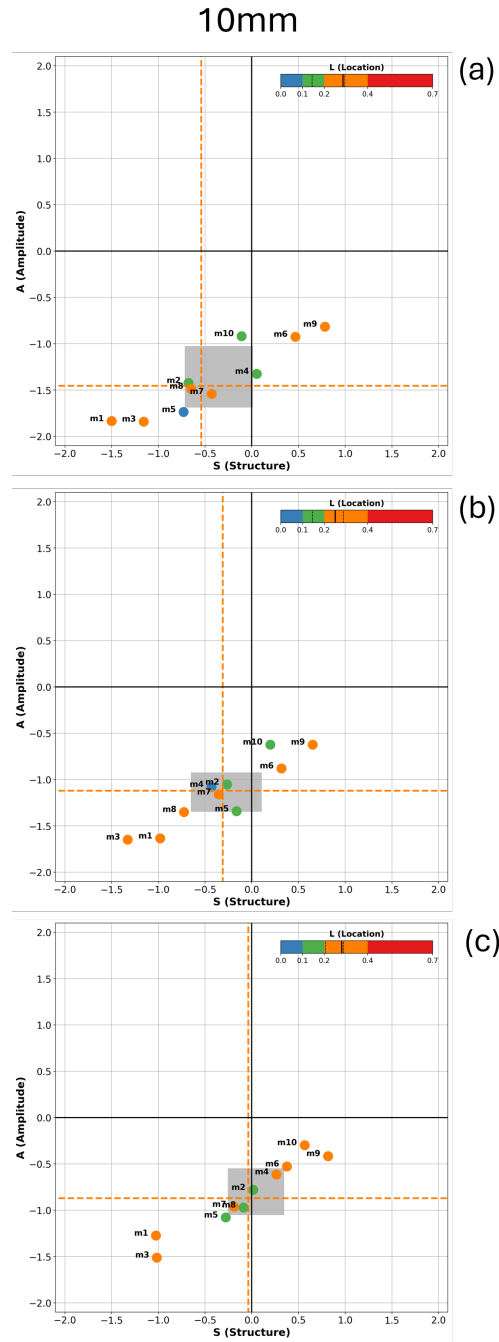


Figure 4.10: SAL scores for the three different ensembles, LOWRES (a), COSMO\_tke (b), SMAGO\_500m (c), at 10 mm threshold. The grey box represents the interquartile range. Dotted lines represent the median values of A and S, their colour shows the corresponding median value of L, also pointed out by the black line into the L colorbar.

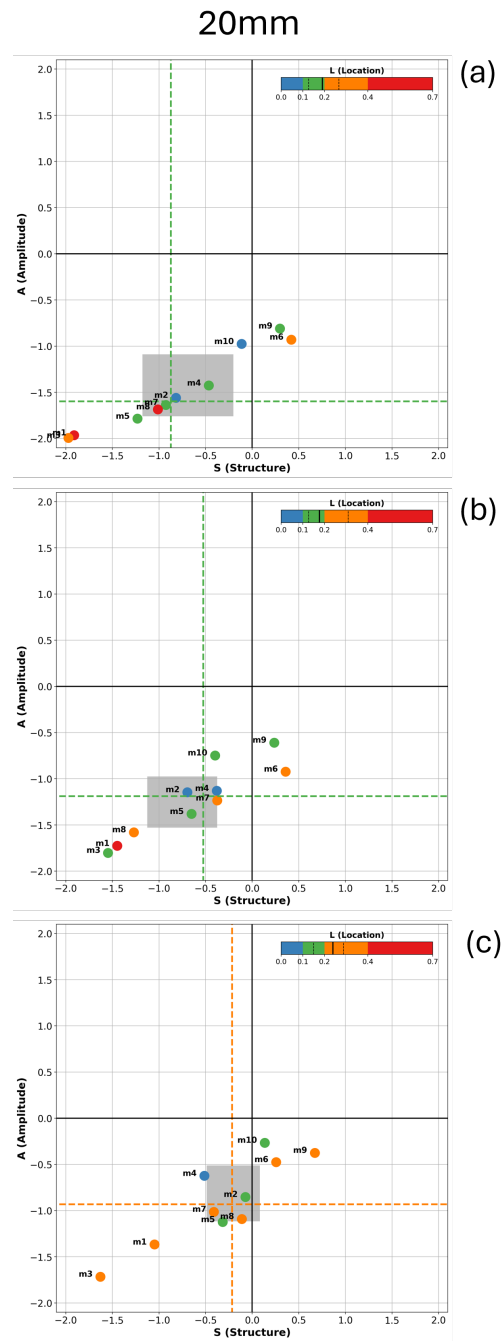


Figure 4.11: Same as figure 4.10 but for the 20 mm total precipitation threshold.

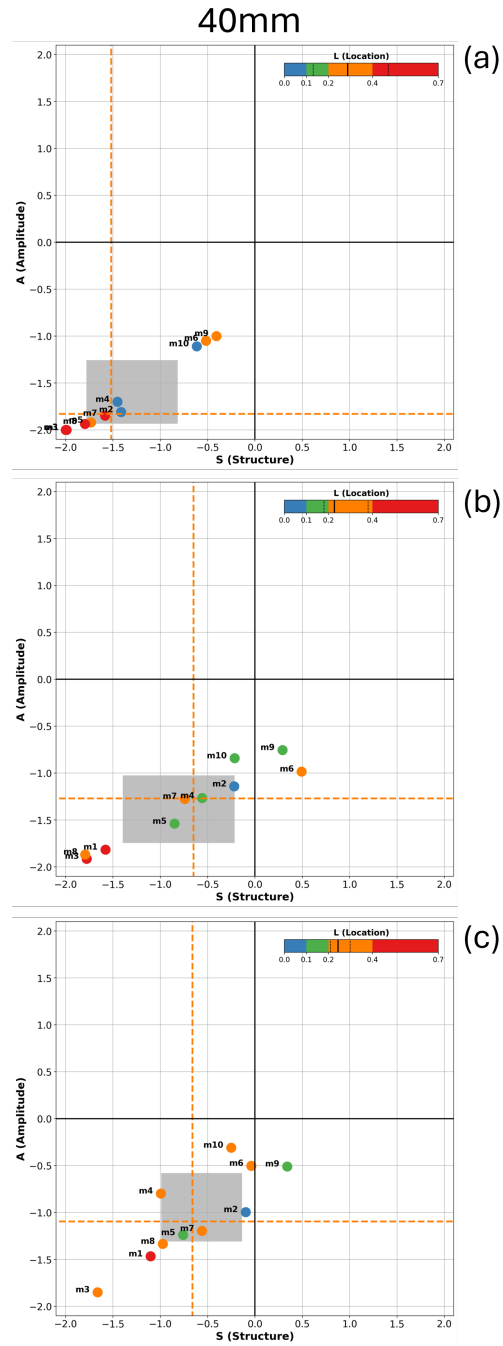


Figure 4.12: Same as figure 4.10 but for the 40 mm total precipitation threshold.

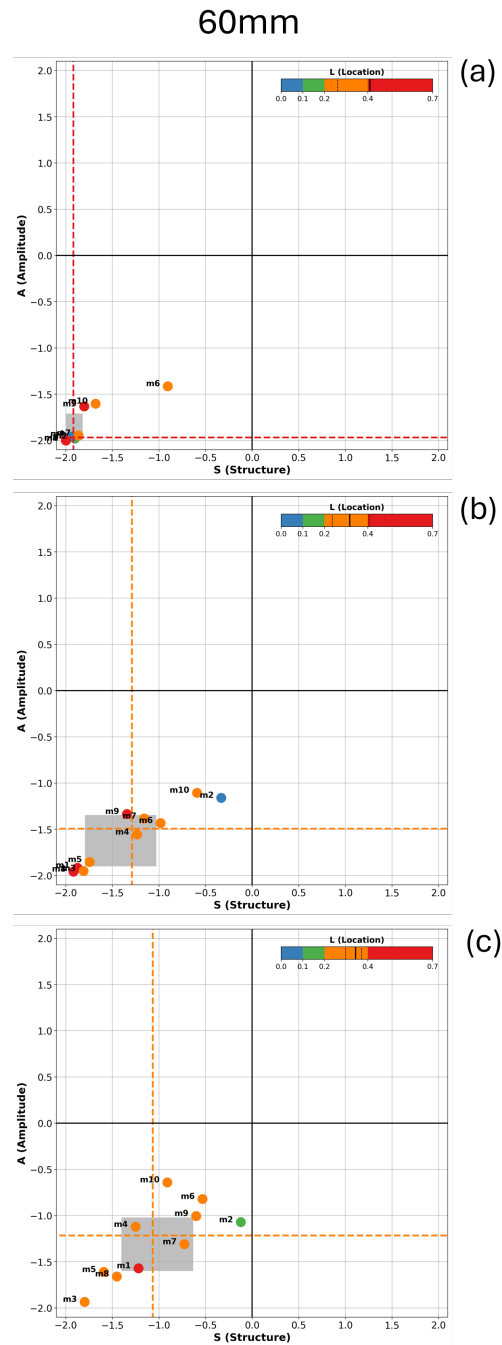


Figure 4.13: Same as figure 4.10 but for the 60 mm total precipitation threshold.

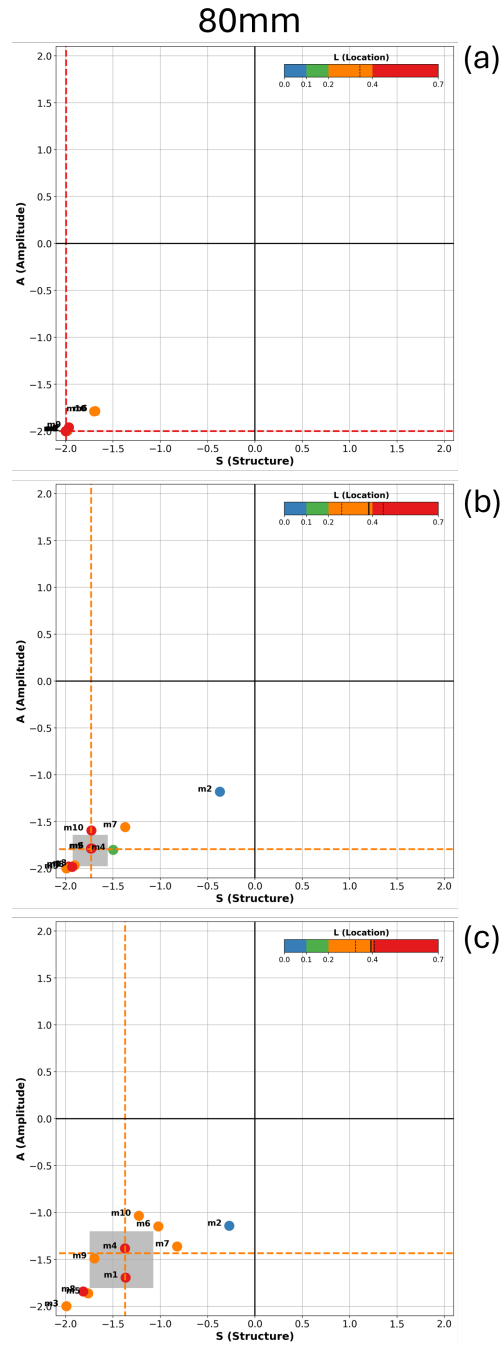


Figure 4.14: Same as figure 4.10 but for the 80 mm total precipitation threshold.

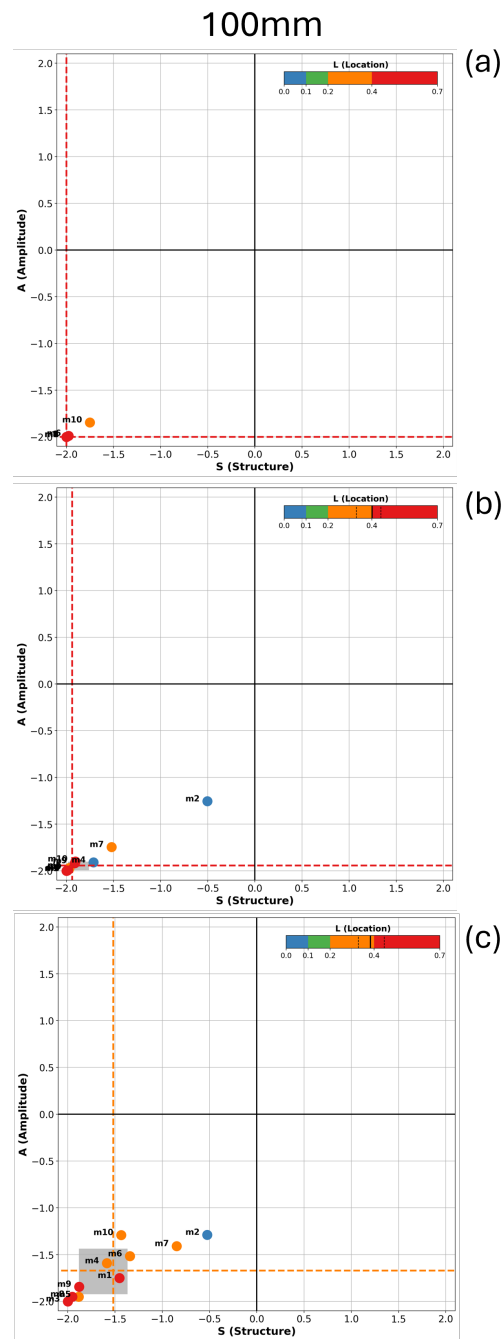


Figure 4.15: Same as figure 4.10 but for the 100 mm total precipitation threshold.



## 4.4 | Final considerations and summary

In this first part, the efforts focused on analysing a remarkable extreme precipitation event that struck central Italy in September 2022, leading to devastating floods and significant economic and human losses. The aim was to investigate the potential and the limitations of current numerical weather prediction systems in representing such high-impact convective events, especially in regions characterized by complex topography and intricate atmospheric dynamics.

The comparison between operational models, such as the IFS, revealed substantial shortcomings in both the intensity and spatial accuracy of predicted precipitation. Although this model was able to broadly identify areas at risk, it struggled in capturing the precise location of the heaviest rainfall. This study highlighted three main factors underlying these limitations: insufficient horizontal resolution, inadequacies in the parameterization of turbulence, and a pronounced sensitivity to initial and boundary conditions.

By employing high-resolution sensitivity experiments and ensemble simulations with the ICON model, this work demonstrates that moving towards hectometric grid spacings (sub-kilometer) results in a significant step forward in the simulation of convective processes. This approach enables a more explicit representation of small-scale turbulence, complex microphysical interactions, and the critical effects of orographic forcing, factors that proved decisive for accurately reproducing the observed precipitation patterns during the Marche flood event.

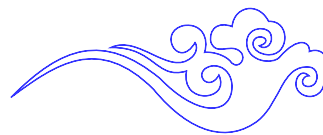
Ensemble analysis further clarified the crucial role of initial and boundary conditions. Only a limited subset of ensemble members was able to realistically simulate the key atmospheric processes driving the event, especially the low-level convergence between moist Mediterranean air and the intrusion of drier, cooler air from northern Europe. Notably, one particular ensemble member (m2) consistently performed best, accurately reproducing wind convergence patterns and the subsequent convective evolution. This finding underscores the importance of not only model resolution and physics, but also the accuracy of IC and LBC perturbations for the prediction of localized, high-impact convective events.

A key insight also emerged from the comparison of turbulence parame-

terizations. The adoption of the three-dimensional Smagorinsky closure at hectometric scales, as opposed to the more traditional one-dimensional TKE closure scheme, led to appreciable improvements in both the spatial distribution and intensity of simulated precipitation, especially in regions where orographic effects and local circulations play a dominant role.

Ultimately, still looking at the SAL scores, it is possible to notice that the spread, quantified by the interquartile range, appears rather limited for the COSMO\_tke ensemble. In particular, only one member (m2) shows very good skill at higher thresholds. These aspects suggest that the improvement mainly results from the increase in resolution, without any concurrent refinement in the physical parameterizations. Still, a noticeable gain in performance is already achieved at this stage, indicating that resolution alone can bring a clear benefit in reproducing the observed precipitation structures.

On the other hand, the SMAGO\_500m ensemble, where the treatment of turbulent processes has been significantly improved, shows a larger spread, with several members starting to perform similarly to m2. This behavior indicates that the model now manages to represent uncertainty within a more physically consistent and stable configuration. In other words, the ensemble appears to have entered a new region of the phase space, where variability and uncertainty are better captured and distributed among the members, leading to a more realistic ensemble representation.





## 5 | Impact of microphysics perturbations

During my stay in Karlsruhe in Spring 2025, I worked with the microphysics group of the Institute of Meteorology and Climate Research (IMKTRO) of the KIT headed by Professor Corinna Hoose. Specifically, my work encountered the collaboration of Dr. Christian Barthlott with whom we explored the potential of the ICON model release 2024.10 (ICON partnership, [2024](#)) at very high resolution using microphysics perturbations. Indeed, starting from the Italian Marche flooding we decided to do some preliminary test on the Horeka supercomputer from the KIT. Following Barthlott's findings on microphysics perturbations regarding the effect on the forecasting of hail and deep convective storms (Barthlott et al., [2024](#)), we wondered if it could be useful to focus on perturbing initial conditions, microphysics parameters, and turbulence schemes. The final choice was to produce a larger ensemble than the ones done in Bologna, exploiting ICON's initial and boundary conditions and analysing the role of microphysics perturbations. These perturbations were then only applicable if the double-moment microphysics scheme was activated, that is an advanced representation of how hydrometeors form into convective and in general precipitating clouds. In the next paragraph the preliminary tests are shown, carried out first of all by selecting the same high-resolution nested domains analysed in the first part of this thesis.

This first step was crucial: indeed, when using another HPC infrastructure, it is difficult to obtain the same results as obtained when using the DIFA OPH cluster in Bologna, because of the different configurations and packages, that can lead to very different output even if the version of the model being used is the same. To overcome this issue and in order to have more robust initial conditions without going through the process of remapping when running the new ensembles with the 2-MOM microphysics scheme, it was decided to exploit directly the initial and boundary conditions from the 6.5 km ICON-EU grid.

The new ensemble then returned very interesting and exciting results, showing how the role of microphysics parameters is essential when it comes to forecasting of extreme and unpredictable storms, especially in areas characterized by very complex orography. Then, to continue the comparison a new case-study for the German territory was selected to test the high-resolution ensemble approach with the same nesting technique and microphysics perturbations. With surprise, a match was discovered between the Italian and German case study regarding the members that produced the most reliable forecasts compared to the observations. Indeed, even though the two case studies were quite different in total precipitation, production of hail and persistence, both of them were characterized by strong convergence and deep convection over complex and mountainous territories, implying that high-resolution down to hectometric scale combined to a good representation of the characteristics of shape, dimensions and numbers of hydrometeors is essential to produce appropriate forecasts.

## 5.1 | 2-MOM microphysics scheme

In cloud-resolving models, the representation of microphysical processes is crucial for the accurate simulation of convective systems, especially in cases of intense precipitation. Many microphysical interactions depend nonlinearly on the size of hydrometeors, yet traditional bulk schemes often rely solely on prognostic equations for mass mixing ratios. In such single-moment approaches, the particle size is implicitly assumed to be tied to the total mass content, which can be a limiting approximation, particularly in rapidly evolving convective environments.

To overcome this limitation, double-moment schemes have been developed, in which both the mass mixing ratio and the number concentration of each hydrometeor category are predicted. These two quantities correspond to the zeroth and first moments of the particle size distribution, providing a more complete statistical description of cloud and precipitation particles (Morrison et al., 2009; Seifert & Beheng, 2006). In particular, knowing the number concentration allows the model to explicitly resolve variations in mean particle size independently of the total mass, which is essential for realistically simulating processes such as sedimentation, collection, and phase changes.

The ICON model includes such a double-moment microphysics option, which prognoses both specific mass and number concentrations for six hydrometeor classes: cloud water, rain water, cloud ice, snow, graupel, and hail. This level of detail is particularly beneficial at convection-permitting or convection-resolving resolutions (that is, grid spacings  $< 3$  km), where the model dynamics can explicitly resolve the strong updrafts required for the formation and growth of graupel and hail.

Given the complexity of the scheme, it incorporates parameterizations for nucleation processes (e.g., activation of cloud condensation nuclei and ice nuclei), particle geometry, fall speeds, and size-dependent microphysical interactions. However, it is important to note that many of these settings are hard-coded in the current ICON implementation and cannot be modified via the user namelists, thus requiring code-level customization for sensitivity experiments. Recent studies, moreover, compare the results of high-resolution simulations carried out with the two schemes, sometimes showing the better performance of the 2-MOM when it comes to very high-resolution runs (at scales lower than 1 km) in capturing the correct timing, localization and extension of convective storms (Barthlott et al., [2024](#)).

The choice of this scheme, selected for the following tests and simulations, is motivated by its ability to realistically represent the evolution of deep convective cores and associated precipitation structures, particularly those involving graupel and hail. The improved representation of particle size distributions and associated processes is expected to enhance the realism of the simulated storm structure and precipitation intensity, especially when compared to more traditional single-moment approaches.

## 5.2 | Italian ensemble

Under these assumptions, the study restarted from the storm that caused the flooding in the Marche region in September 2022 that has been analysed in the first part of the work. After implementing the same setup of the SMAGO\_500m high-resolution ensemble, that is the one with the Smagorinsky scheme activated at 550 m resolution, 4 different parameters were perturbed defining the cloud hydrometeors dimensions and accretion processes, such as the Ice

Nucleation Particle density (INP), the amount of Cloud Condensation Nuclei (CCN), the Shape parameter of the particle size distribution (NU) and the graupel sedimentation velocity (GR), as reported in the following table, leading to a total number of 36 ensemble members. For instance, when the scaling factor is lower than 1 for GR, it means that the graupel particle falls slower than the standard behaviour of the scheme used in the model, thus staying longer in the cloud; the opposite if the factor is greater than 1. The simulations were done on the HoreKa supercomputer of the KIT university and no changes to the ICs and LBCs taken from ICON-EU were applied, but only the mentioned key microphysical parameters were perturbed.

| Parameter                                  | Low | Medium | High |
|--------------------------------------------|-----|--------|------|
| <b>INP</b> (Ice Nucleating Particles)      | 0.1 | 1      | 10   |
| <b>CCN</b> (Cloud Condensation Nuclei)     | 6   | 8      | –    |
| <b>NU</b> (Shape parameter of PSD)         | –   | 1      | 4    |
| <b>GR</b> (Graupel sedimentation velocity) | 0.3 | 1      | 1.7  |

Table 5.1: Adimensional factors for the perturbed microphysical parameters used in the 2-MOM Italian ensemble experiment at 550 m resolution. INP and CCN represent scaling factors for particle number concentrations (for CCN there are 4 values: 6 clean maritime air, 7 intermediate, 8 continental, 9 polluted), NU defines the shape parameter of the hydrometeor particle size distribution, and GR is a scaling factor that modifies the velocity of graupel sedimentation. The full ensemble consists of all possible combinations of these values (total of 36 members).

As shown in Figure 5.1, the ensemble simulation reported useful insights about the total precipitation field that affected the flooded region, especially considering the ensemble mean and the 99th percentile, which shows cumulated values higher than 200 mm. The observing precipitation network reported extended bands over 200 mm with peaks of 400 mm between Umbria and Marche (shown in Figure 3.3). Moreover, it can be noticed that the members better reproducing the precipitating pattern were those characterized by low CCN values and low INP concentrations (members 19-30) as shown in table 5.2.

| #  | CCN | INP | GR  | NU |
|----|-----|-----|-----|----|
| 1  | 8   | 1   | 1.0 | 1  |
| 2  | 8   | 1   | 0.3 | 1  |
| 3  | 8   | 1   | 1.7 | 1  |
| 4  | 8   | 1   | 1.0 | 4  |
| 5  | 8   | 1   | 0.3 | 4  |
| 6  | 8   | 1   | 1.7 | 4  |
| 7  | 8   | 0.1 | 1.0 | 1  |
| 8  | 8   | 0.1 | 0.3 | 1  |
| 9  | 8   | 0.1 | 1.7 | 1  |
| 10 | 8   | 0.1 | 1.0 | 4  |
| 11 | 8   | 0.1 | 0.3 | 4  |
| 12 | 8   | 0.1 | 1.7 | 4  |
| 13 | 8   | 10  | 1.0 | 1  |
| 14 | 8   | 10  | 0.3 | 1  |
| 15 | 8   | 10  | 1.7 | 1  |
| 16 | 8   | 10  | 1.0 | 4  |
| 17 | 8   | 10  | 0.3 | 4  |
| 18 | 8   | 10  | 1.7 | 4  |
| 19 | 6   | 1   | 1.0 | 1  |
| 20 | 6   | 1   | 0.3 | 1  |
| 21 | 6   | 1   | 1.7 | 1  |
| 22 | 6   | 1   | 1.0 | 4  |
| 23 | 6   | 1   | 0.3 | 4  |
| 24 | 6   | 1   | 1.7 | 4  |
| 25 | 6   | 0.1 | 1.0 | 1  |
| 26 | 6   | 0.1 | 0.3 | 1  |
| 27 | 6   | 0.1 | 1.7 | 1  |
| 28 | 6   | 0.1 | 1.0 | 4  |
| 29 | 6   | 0.1 | 0.3 | 4  |
| 30 | 6   | 0.1 | 1.7 | 4  |
| 31 | 6   | 10  | 1.0 | 1  |
| 32 | 6   | 10  | 0.3 | 1  |
| 33 | 6   | 10  | 1.7 | 1  |
| 34 | 6   | 10  | 1.0 | 4  |
| 35 | 6   | 10  | 0.3 | 4  |
| 36 | 6   | 10  | 1.7 | 4  |

Table 5.2: List of microphysics ensemble members and corresponding 2-MOM microphysical parameter values.



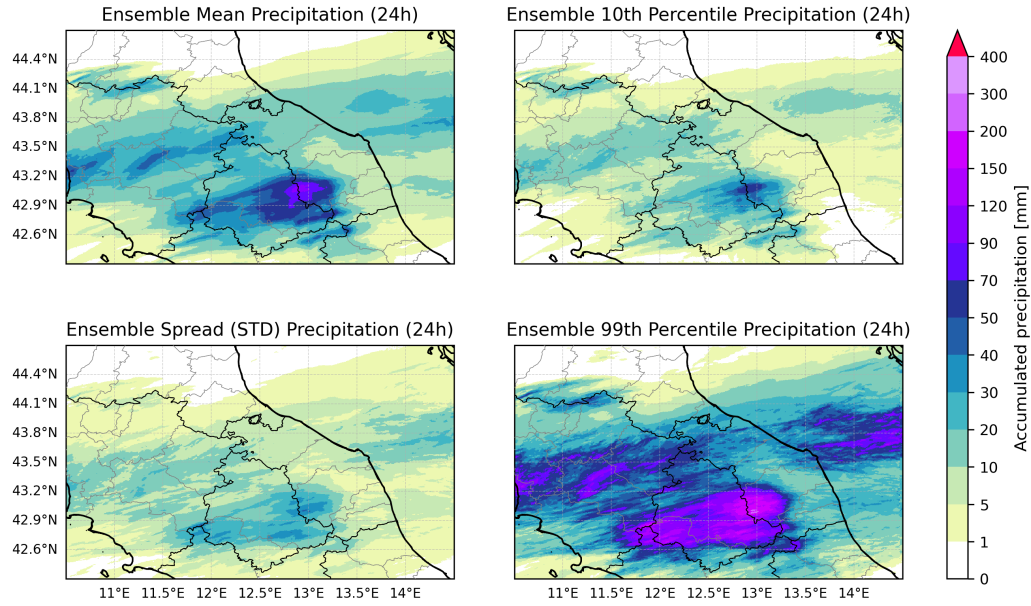


Figure 5.1: Ensemble statistics of the 36-members ensemble for the Marche flooding that happened on 15th September 2022.

### 5.2.1 Member analysis

During the analysis of the high-resolution microphysics-perturbed ensemble, several members clearly stood out for their skill in reproducing both the total rainfall and the spatial structure of the September 2022 Marche flood event. In particular, Figures 5.2, 5.3 and 5.4 display the simulated precipitation fields from members m19 to 30, those which most closely matched the observed event, both in terms of accumulated rainfall and the spatial pattern of the convective system.

When examining these 550 m resolution maps, the first thing that stands out is that all these members succeed in capturing the main band of heavy precipitation, with local maxima exceeding 200 mm in some of them, but with a noticeable southward displacement compared to the observations. However, even among the best-performing members, there remain some systematic discrepancies in the downwind (leeward) areas compared to observations. Specifically, in the valleys east of the Apennine ridge, where the flooding was most severe, the modelled precipitation maxima tend to be underestimated or slightly misplaced, usually shifted a few kilometers inland. This highlights how, even at hectometric resolution and with sophisticated microphysics, models

still face challenges in capturing the persistence and precise location of the most intense, stationary convective cells that are responsible for catastrophic impacts in these complex terrains.

A closer look at the spatial distribution reveals that members m20, m23, m26 and m29, for example, reproduce not only the main precipitation band but also the structure of the characteristic V-shape storm of the observed rainfall. These cases show a much greater spatial coherence and accuracy compared to the majority of the other ensemble members, managing to represent both the principal maximum and some of the smaller-scale features that are often missed in standard forecasts. Another notable feature is the ability of these ensemble members to simulate extreme rainfall totals that are very close to the observed peaks. This suggests that, for certain combinations of microphysical parameters (such as CCN, INP, shape parameter, and GR), the model is capable of realistically simulating the growth and maintenance of the most intense convective systems. Particularly impressive is the persistence of the simulated precipitation bands over the same area for several hours, which successfully captures the back-building, self-reinforcing nature of the actual event, a key factor that contributed to its destructive impact.

A useful comparison can be made with member m2, originating from the 10-member ensemble experiment in which the initial and boundary conditions were perturbed and driven by IFS data. Among that ensemble, m2 was the only member that realistically reproduced the event, successfully capturing the main convective band and the extreme rainfall over the flood-affected area. This suggests that, under favorable initial conditions, the model can spontaneously organize convection in a physically consistent way, even without microphysical parameter perturbations. In contrast, the microphysics-perturbed ensemble analysed here adopts a different strategy: all members share the same initial and boundary conditions, derived from ICON-EU, and only the microphysical parameters are varied. This setup naturally produces a more homogeneous ensemble, in which most members capture the large-scale precipitation pattern, but with varying accuracy in the spatial placement and leeward intensity of the rainfall maxima. Interestingly, while m2's realism primarily stems from favorable dynamical initialization, the best-performing microphysics members (m19-m30) achieve comparable skill through parameter-driven microphysical diversity. The comparison highlights that both approaches, perturbing the ini-

tial state or perturbing the microphysical scheme, can yield realistic outcomes, but through distinct physical pathways. Moreover, the persistence of systematic biases in the leeward areas across the microphysics ensemble indicates that, even at hectometric resolution and with refined physics, accurately simulating moisture resupply and convective stationarity remains a major modelling challenge in such complex orography.

In summary, the analysis of single members demonstrates that a high-resolution ensemble with advanced microphysical perturbations can approach the observed reality very closely, not just in rainfall totals but also in spatial organization. Still, certain “hotspots” remain difficult to predict accurately, emphasizing the importance of ensemble diversity, physical realism, and continual improvement in both microphysics and model resolution for reliable forecasting of convective extremes. This analysis reinforces the idea that, while model advances are steadily improving our ability to predict such events, the combination of complex terrain, subtle atmospheric interactions, and the inherent unpredictability of convection mean that some uncertainties will always persist.

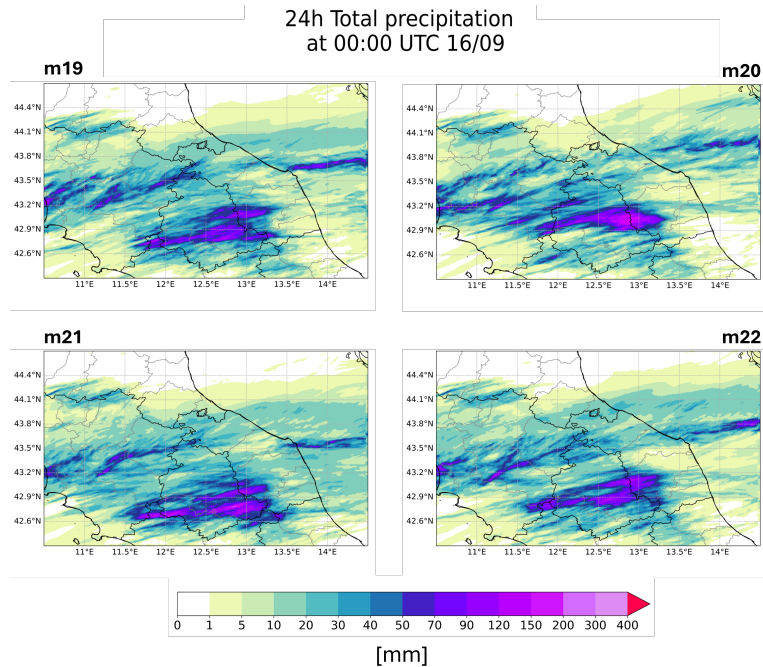


Figure 5.2: Ensemble members 19, 20, 21 and 22 from the microphysics Marche ensemble

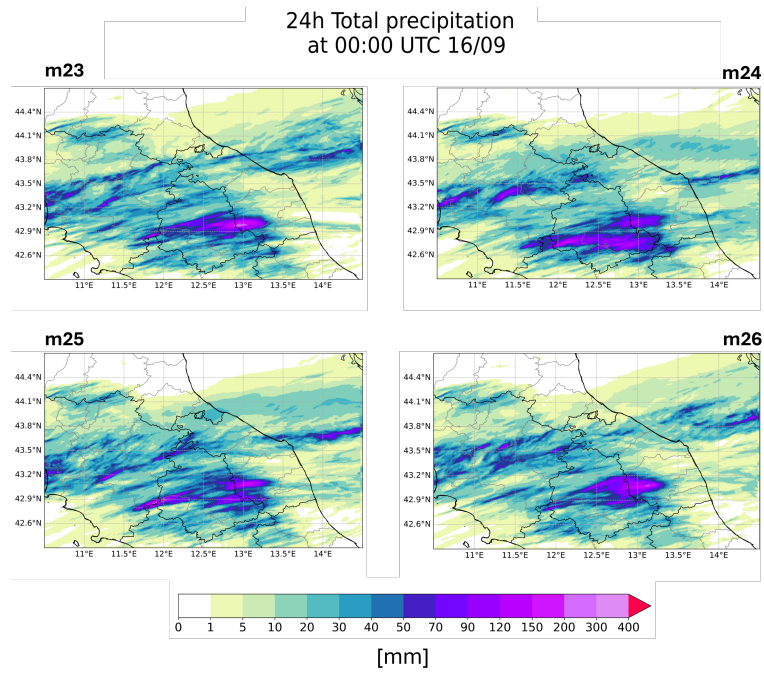


Figure 5.3: Ensemble members 23, 24, 25, 26 from the microphysics Marche ensemble

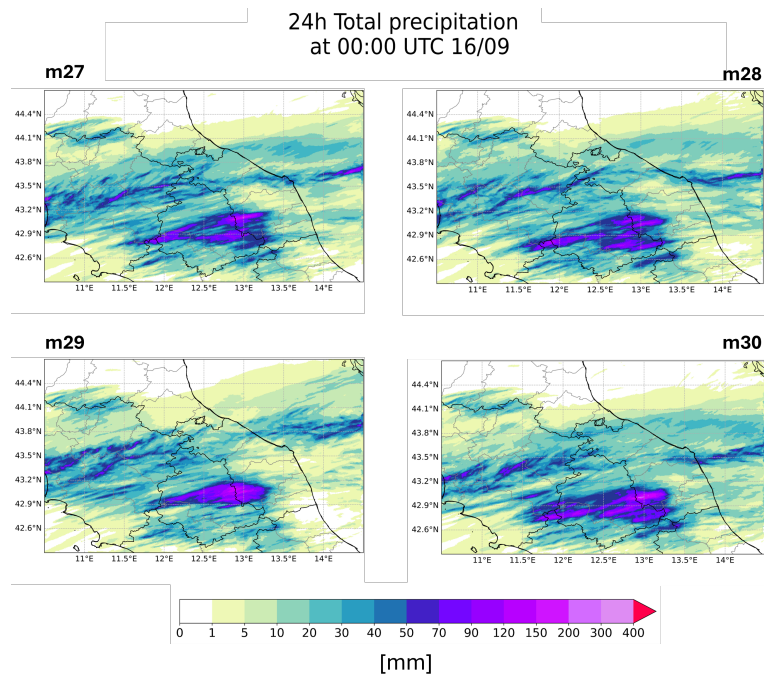


Figure 5.4: Ensemble members 27, 28, 29, 30 from the microphysics Marche ensemble

### 5.2.2 SAL analysis

The SAL score analysis was also extended to this 36-member high-resolution Italian ensemble.

Figures 5.5–5.7 summarise the results for thresholds of 10, 20, 40, 60, 80, and 100 mm. As in the previous experiments, increasing the threshold progressively filters out weaker precipitation features, focusing the evaluation on the model’s ability to represent the most intense rainfall objects.

- **Low thresholds (10 mm and 20 mm)** (Figure 5.5):
  - All 36 members produce precipitation objects above both thresholds, resulting in dense point clouds in SAL space.
  - The spread in S and A is relatively small compared to previous ensembles (LOWRES, COSMO\_tke, SMAGO\_500m), with most members clustered close to  $(S, A) \approx (0.5, -0.5)$  at 10 mm, indicating a slight tendency towards overly compact structures and a mild underestimation of total rainfall.
  - At 20 mm, the cluster moves closer to the origin, with S values ranging from slightly negative to slightly positive, and A values still mildly negative. This suggests improved balance between object size and intensity when weaker precipitation areas are excluded.
  - L scores are generally low (mostly below 0.2), indicating good spatial alignment between simulated and observed precipitation features.
- **Intermediate thresholds (40 mm and 60 mm)** (Figure 5.6):
  - As the threshold rises, fewer members retain precipitation objects above the limit, but the ensemble maintains a coherent clustering in SAL space.
  - At 40 mm, S values tend to become slightly negative for several members, reflecting a modest tendency to simulate more diffuse objects compared to observations, while A remains close to zero for many of the best members, indicating a good match in total amount for intense cores.

- The 60 mm threshold further narrows the ensemble spread, with the best-performing members (e.g., m23, m26, m29, m30) lying close to the origin in SAL space and showing low L scores, confirming their skill in reproducing both the structure and location of heavy rainfall bands.
- Members with higher L scores often correspond to spatial displacements of only a few kilometres, which is within the typical uncertainty range for convective systems in complex terrain.
- **High thresholds (80 mm and 100 mm)** (Figure 5.7):
  - At these extreme thresholds, only a subset of the ensemble retains precipitation objects, highlighting the difficulty of reproducing the most intense convective peaks.
  - At 80 mm, the spread in S increases slightly, with some members (notably m23, m26, m29, m30) positioned close to (0, -1.5) in A, indicating a moderate underestimation of amplitude despite realistic structures. These members also maintain among the lowest L scores, evidencing accurate spatial placement of the extreme rainfall maxima.
  - At 100 mm, the number of members further decreases, with the majority concentrated in the lower-left quadrant (negative S, strongly negative A). This reveals that even in the most skillful members, the highest observed peaks are generally underestimated, although their spatial location remains well captured in a few cases.
  - The persistence of low L values for certain members at these thresholds underscores the role of microphysical parameter choices in controlling both the magnitude and localization of extreme precipitation objects.

Overall, also referring to table 5.3, it is possible to say that the 2-MOM microphysics-perturbed ensemble shows less spread in SAL space at low and intermediate thresholds compared to the initial condition-perturbed ensembles analysed earlier. This suggests that, while microphysical variations significantly affect the magnitude and structure of extreme precipitation, they do not induce large variability in the positioning of rainfall objects as perturbations



| Threshold | S           | A              | L               | Main notes                                                                                                             |
|-----------|-------------|----------------|-----------------|------------------------------------------------------------------------------------------------------------------------|
| 10 mm     | ↑ (compact) | ↓ (slight)     | Low ( $< 0.2$ ) | All members active; cluster near $(0.5, -0.5)$ ; slight positive S, slight negative A; minor rainfall underestimation. |
| 20 mm     | $\approx$   | ↓ (slight)     | Low ( $< 0.2$ ) | Balanced S and A; closer to origin; reduced spread vs. 10 mm; better structure/intensity match.                        |
| 40 mm     | ↓ (diffuse) | $\approx$ to ↓ | Low–moderate    | Fewer members above threshold; best reproduce rain bands realistically; slight A underestimation.                      |
| 60 mm     | ↓           | $\approx$      | Low             | Cluster near $(S, A) \approx (-0.8, -1.2)$ ; m23, m26, m29, m30 with accurate L and strong cores.                      |
| 80 mm     | $\approx$   | ↓ (strong)     | Low             | Only top members; structure close to observed; amplitude underestimated; accurate location.                            |
| 100 mm    | ↓           | ↓ (strong)     | Low–moderate    | Very few exceed threshold; most underestimate peaks; some maintain good location; higher L for weaker runs.            |

Table 5.3: Qualitative summary of SAL score tendencies for the 2-MOM Italian ensemble at different precipitation thresholds. Arrows indicate tendency: ↑ overestimation, ↓ underestimation,  $\approx$  close to observations.

to the initial state. Nevertheless, the most skillful members at high thresholds, often those with lower CCN and INP values, consistently achieve the best compromise between S, A, and L, demonstrating the potential of targeted microphysical tuning for improving the representation of severe convective rainfall in high-resolution forecasts.



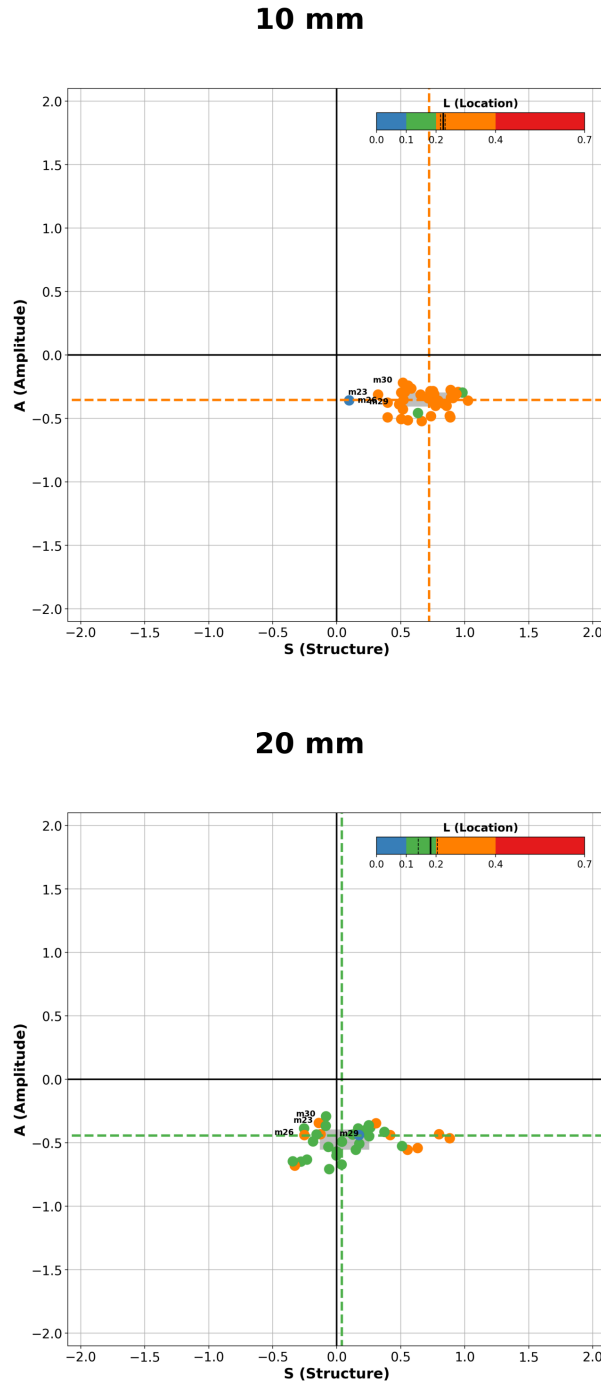


Figure 5.5: SAL scores for the 10 mm and 20 mm thresholds for the 2-MOM Italian ensemble. The grey box represents the interquartile range (IQR). Dotted lines represent the median values of A and S, their colour shows the corresponding median value of L, also pointed out by the black line into the L colorbar.

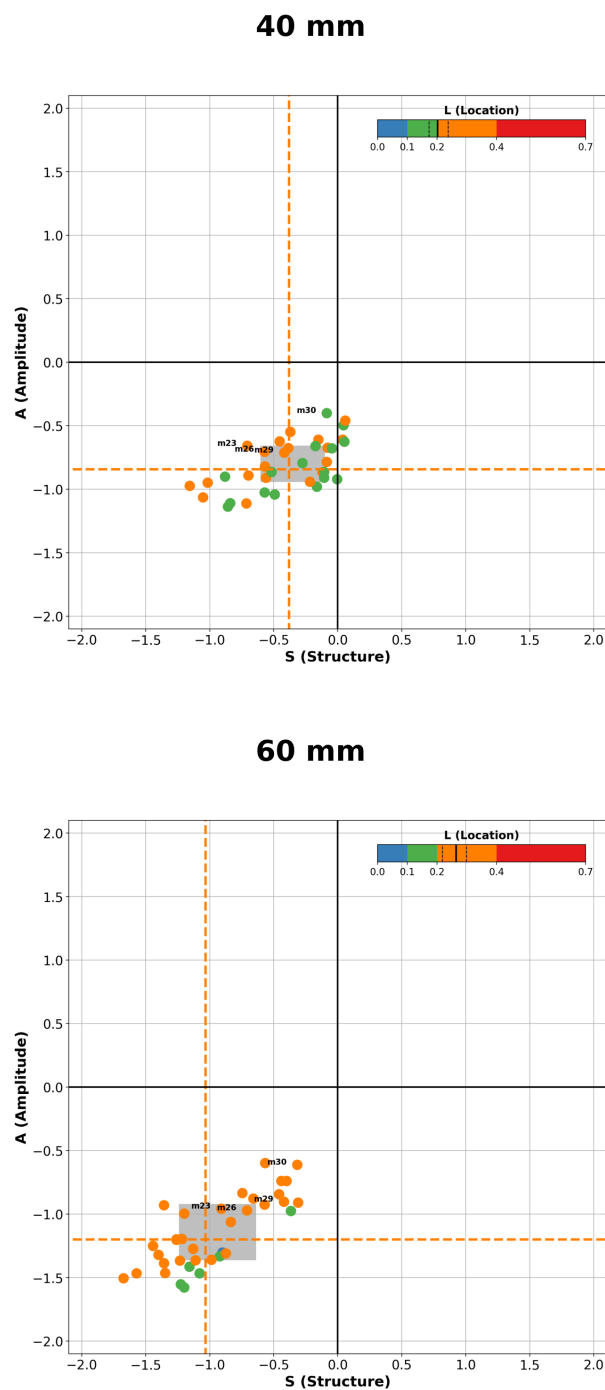


Figure 5.6: Same as figure 5.5 but for the 40 mm and 60 mm precipitation thresholds.

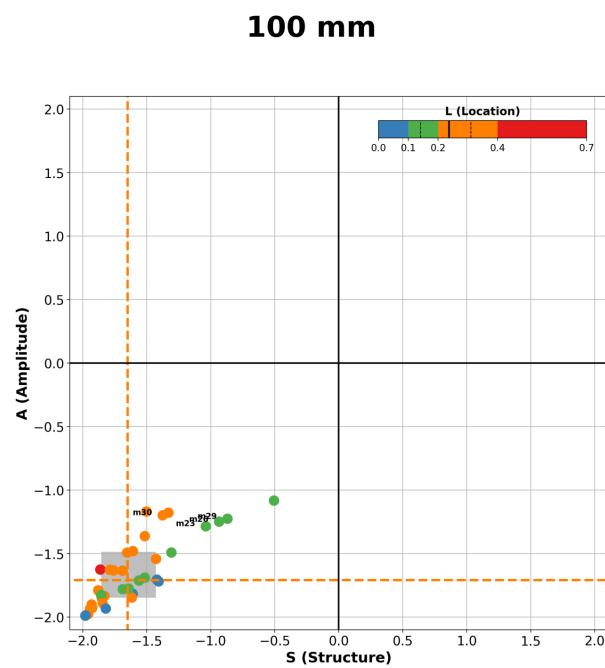
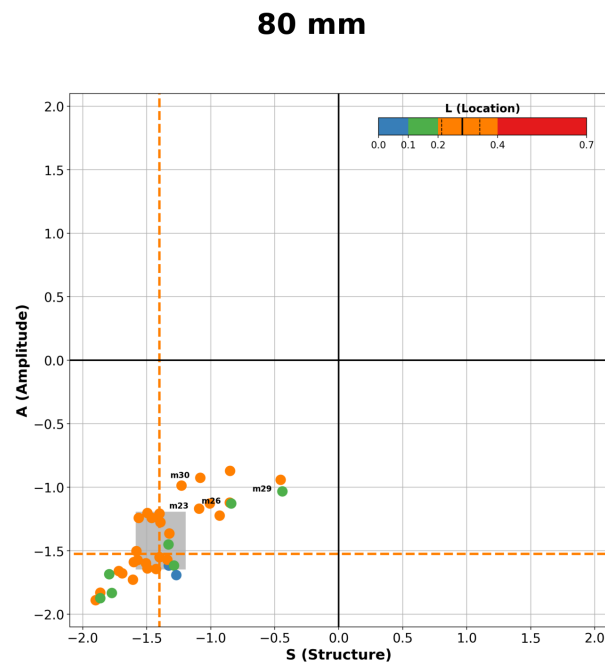


Figure 5.7: Same as figure 5.5 but for the 80 mm and 100 mm precipitation thresholds.

## 5.3 | German case study and the MOSES campaign

The Black Forest region, with its pronounced orography and strong land–atmosphere coupling, has long been recognised as a hotspot for convective storm development. Narrow valleys, forested slopes, and elevated plateaus provide multiple triggering mechanisms for deep convection, while mesoscale flow channeling often leads to rapid storm organisation and upscale growth. In recent years, this area has become the natural laboratory for the Helmholtz initiative Modular Observation Solutions for Earth Systems (MOSES), a research programme designed to capture so-called “event chains” of hydro-meteorological extremes from their atmospheric initiation through to surface impacts. MOSES combines stationary observatories with mobile modules such as the KITcube radar–lidar system, radiosonde launches, disdrometers, and hail sensors, allowing researchers to follow convective storms in unprecedented spatial and temporal detail (Kunz et al., 2022).

The Swabian MOSES campaigns in 2021 and 2023 specifically targeted the southeastern Black Forest, the Neckar Valley, and the Swabian Alb, providing a unique multi-institutional observational setup. Villingen-Schwenningen played a central role as a deployment site for KITcube and a dense network of supporting sensors, underlining its strategic location in a region where some of the most damaging hailstorms in Germany have been observed (GFZ German Research Centre for Geosciences, 2023; Helmholtz Centre for Environmental Research (UFZ), 2023; Karlsruhe Institute of Technology, 2023). The 2021 campaign demonstrated the exceptional variability and intensity of convective activity in this area, with detailed case studies revealing the importance of orographic lifting, boundary-layer processes, and aerosol conditions for storm initiation and hail growth (Barthlott et al., 2024; Kunz et al., 2022). These insights emphasised the Black Forest not only as a vulnerable region but also as an ideal testbed for investigating how local forcing interacts with larger-scale dynamics to produce severe weather.

Against this backdrop, the hailstorm of 14 August 2023 offers an interesting case for further study. According to local observational evidence, the storm peaked around late morning (11 UTC) and persisted for only a few hours, yet it was capable of producing hail of large size over Villingen-Schwenningen and surrounding areas. The short lived but highly damaging nature of this event

makes it especially relevant in the MOSES framework, where the goal is to capture rapid onset extremes with strong impacts on society and infrastructure.

### 5.3.1 A record year for hail

The year 2023 has been widely recognized as one of the most remarkable in recent history for severe hailstorms across Europe. According to the European Severe Storms Laboratory (ESSL), the number of very large hail reports (>5 cm) reached unprecedented levels, with several outbreaks producing giant hail exceeding 10 cm in diameter (ESSL, [2023](#), [2024](#)). Among the most striking examples is the new European record, with a hailstone of 19 cm measured in Azzano Decimo, northern Italy, on 24 July 2023 (The Guardian, [2023](#); Willis Towers Watson, [2024](#)). Other remarkable events occurred in Slovenia, Croatia, and Austria, where long-lived supercells generated destructive hail swaths with significant societal and economic impacts (Severe Weather Europe, [2023](#)).

Germany was also severely affected during the summer of 2023, with multiple outbreaks striking Bavaria, Baden-Württemberg, and parts of central Germany. Several events produced hailstones of 5–8 cm and extensive swaths of damage to cars, buildings, and crops, with insured losses running into billions of euros (ESSL, [2023](#)). These cases further underline Germany's exposure to organized convective systems and its relevance as a natural laboratory for studying hail processes in complex terrain, particularly in regions like the Black Forest and the Swabian Jura.

This unusual clustering of extreme hail events highlights both the vulnerability of central and southern Europe to convective hazards and the possible role of climate variability and change in shaping their frequency and intensity. The concentration of high-end hailstorms within a single summer has raised attention not only in the scientific community, but also in the insurance and risk-management sectors, where record-breaking losses were reported. In this context, the Swabian-Jura hailstorm of 14 August 2023, analysed in this study, should be viewed as part of a broader pattern of extraordinary convective activity during the summer of 2023.

### 5.3.2 Observation and description of the event

Figure 5.8 shows the 24-hour accumulated precipitation observed on 14 August 2023 during the Swabian-Jura hailstorm case. The precipitation field reveals a set of well-defined convective cores rather than a widespread mesoscale structure. The most intense rainfall is concentrated in two main nuclei: one located southwest of the Swabian Jura, with peak accumulations exceeding 60 mm, and another smaller but still distinct core further north. These convective cells are elongated along a southwest–northeast axis, consistent with the prevailing low-level convergence lines reported by the observational network.

Unlike the Marche flood event of September 2022, this case was characterized by a much shorter duration: convection developed rapidly during late morning and decayed after only a few hours, with no evidence of the long-lived, back-building structures that sustained the Italian event. Consequently, the overall precipitation amounts are more limited, rarely surpassing 60 mm, and the spatial footprint of the storm system is narrower. Nonetheless, the sharp gradients between convective maxima and surrounding low-precipitation areas highlight the highly localized impact typical of severe thunderstorms in complex terrain.

Complementary evidence from the MOSES field campaign (Handwerker et al., 2025) provides further insights into the dynamical and microphysical characteristics of the storm. Doppler lidar and GNSS measurements near Villingen-Schwenningen documented the development of a pronounced lee-side convergence zone east of the Black Forest, consistent with the triggering of deep convection by low-level flow splitting around the mountain ridge. Storm chasing teams deployed mobile probes and swarmsondes in the vicinity of the storm, capturing updrafts exceeding  $20\text{--}30\text{ m s}^{-1}$  and detecting hail at the surface, even though mostly of small to medium size. Two targeted radiosonde launches confirmed a highly unstable environment, with elevated CAPE and relatively weak convective inhibition, supporting rapid convective initiation.

Radar data from the RADOLAN system of the German Weather Service DWD, which combines radar measurements with ground-based rain gauge observations to provide high-resolution precipitation estimates, confirmed the presence of compact but intense reflectivity cores, consistent with the observed localized precipitation maxima. The observed pattern confirms that while the

event did not reach catastrophic rainfall totals, it remains meteorologically significant: strong wind convergence and intense but short-lived updrafts led to efficient precipitation production within a confined region. This illustrates a key contrast with the Marche case: here, predictability hinges not on prolonged organization of convection, but on correctly capturing the triggering and rapid intensification of isolated convective cells over an orographically sensitive area leading to relatively high hail production.

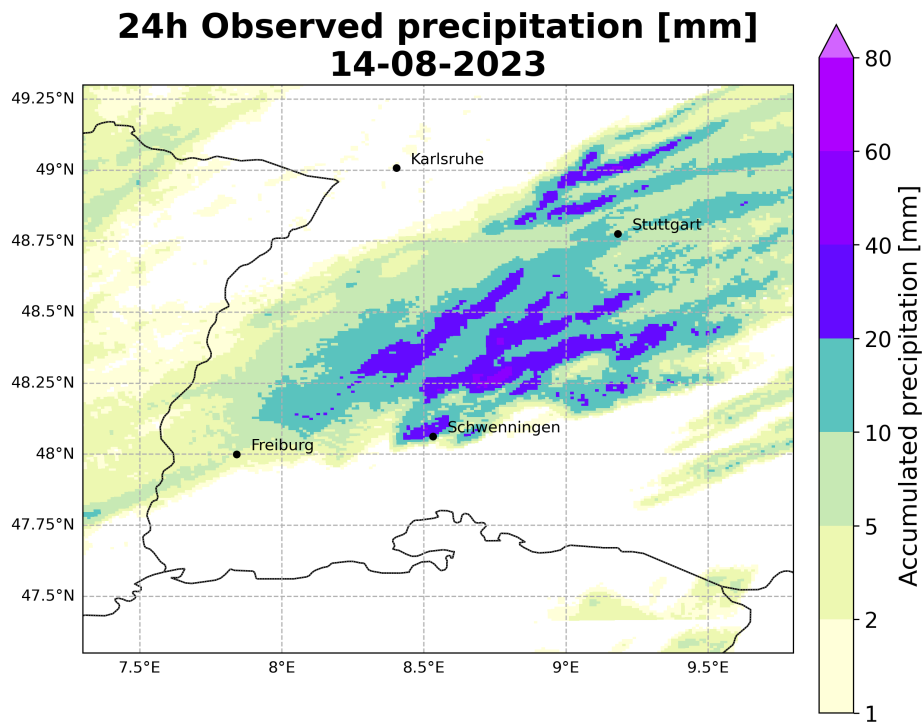


Figure 5.8: 24h Observed total precipitation from 14/08/2023 00:00 UTC to 15/08/2023 00:00 UTC as retrieved from the RADOLAN network.

### 5.3.3 German ensemble setup

To assess the role of microphysics in a different environment, another ensemble simulation with 36 member was produced. Here the case of the Villingen-Schwenningen hail event that took place on 14th august 2023 as reported in paragraph 5.3 is analysed.

As for the Italian microphysics ensemble, the same perturbation setup was adopted, by using the 2-MOM microphysics scheme that better describes the coalescence, formation and production of hydrometeors, especially hail. As

always, starting from a 2.2 km horizontal domain represented by all Germany and some neighbouring countries (that is the ICON-D2 domain used for operational forecasting at DWD), then applied the double nesting at 1.1 km and 550 m horizontal resolution respectively for the smaller domain. Also in this case, it was tricky to find the right way to cut the nested domain as in the southern part the Alps mountain range covers a large portion of the area (Figure 5.9).

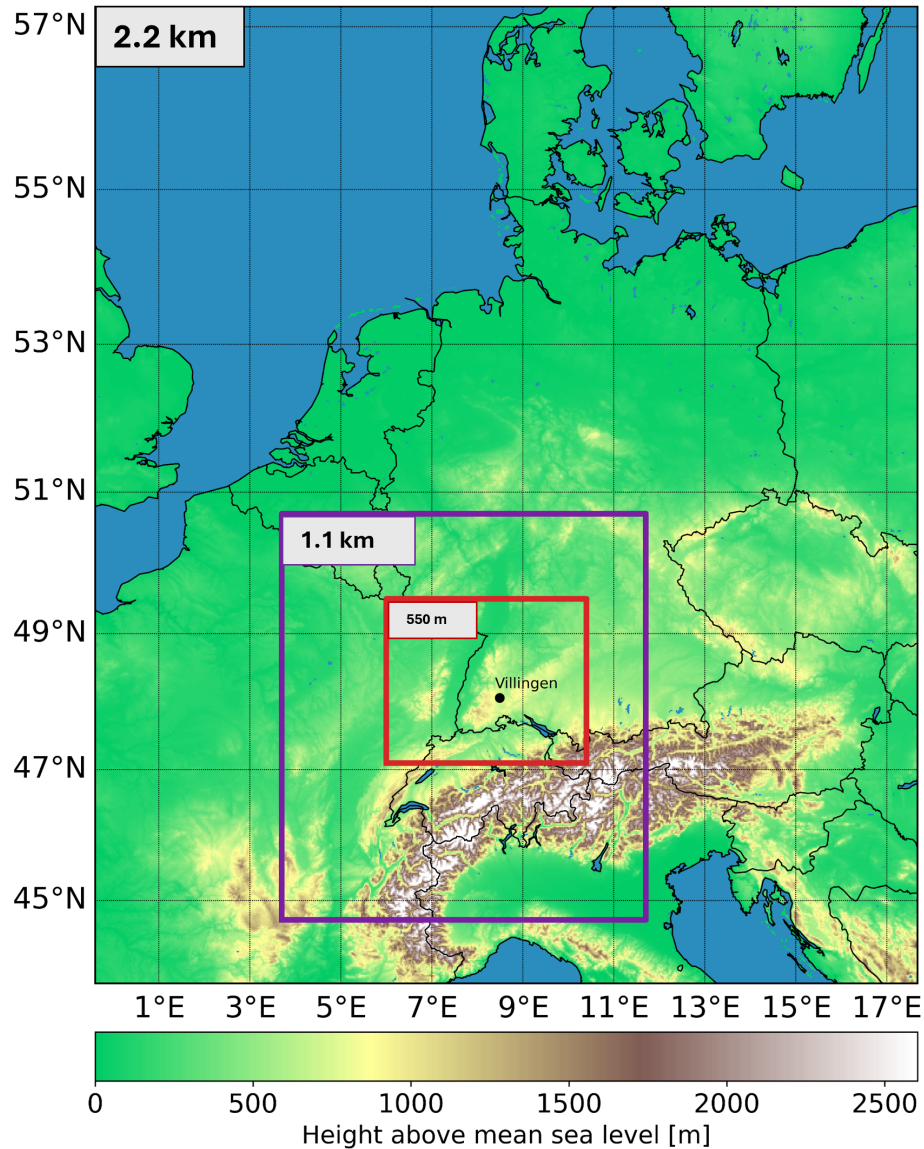


Figure 5.9: Domains of the German case simulations.

Then, it was adopted the same configuration used for the Italian case for most of the variables (see table A.1). In particular, the Smagorinsky turbulence scheme was selected for the smaller domain following the previous results



showing the goodness of this turbulence scheme for higher resolution domain.

Still, the same parameters as for the Italian case regarding the microphysics, that are CCN concentrations, INP, NU parameter and GR were perturbed as shown in table 5.1. Thus, as a first analysis here maps showing ensemble mean, spread and, again 10th and 99th percentile of the total precipitation are produced. They are reported in Figure 5.10.

### 5.3.4 Ensemble analysis

At first sight, we can observe the different nature of the event: compared to the Marche case, the Villingen-Schwenningen storm system is clearly more localized, both in space and in intensity. This is well captured by the ensemble results shown in Figure 5.10. The ensemble mean (top left panel) exhibits a relatively weak and spatially limited precipitation signal, with most grid points showing values below 10 mm. This reflects the short-lived and isolated nature of the observed convective cells, which do not appear as persistent or widespread systems within the ensemble mean field.

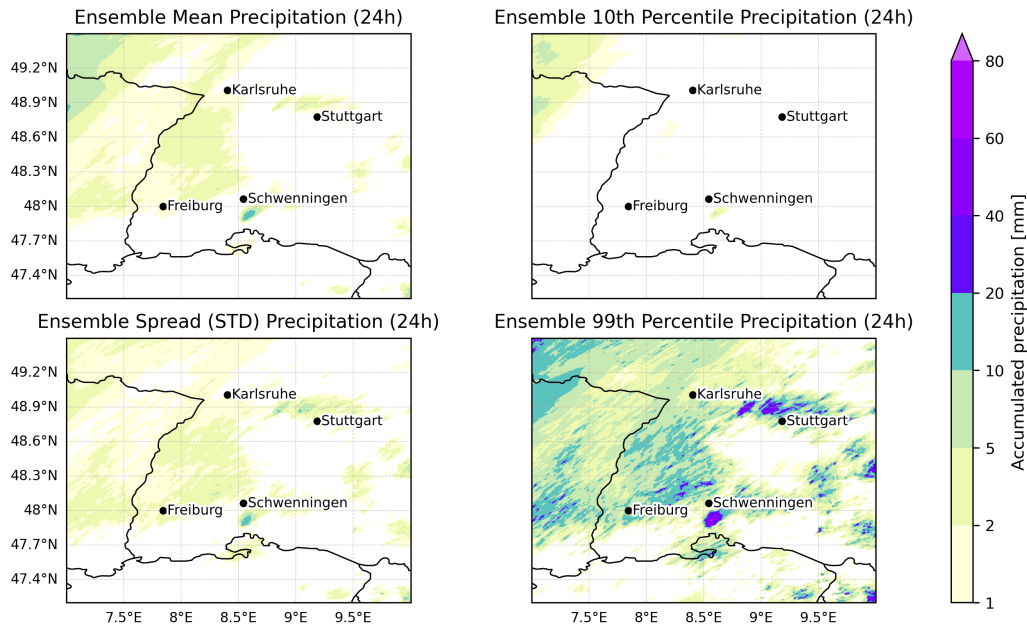


Figure 5.10: Ensemble statistics of the 36-members ensemble for the Villingen-Schwenningen hail event that happened on 14th August 2023.

The 10th percentile (top right panel) confirms this, as it shows very little precipitation across the entire domain. This suggests that in a significant portion

of ensemble members, the convection remained weak, consistent with the idea that the predictability of such events hinges on the correct simulation of local triggering mechanisms and small-scale convergence zones. Anyway, it seems that many members reproduced the hail cores, but few with the correct high intensity.

The spread (bottom left panel) highlights this uncertainty: although low in absolute terms, it is concentrated over the same regions where observed precipitation was most intense (see Figure 5.8), suggesting a high sensitivity to small perturbations in microphysical or dynamical processes. This is typical of highly convective and terrain-sensitive scenarios, where slight shifts in storm position or intensity can lead to markedly different rainfall outcomes.

Interestingly, the 99th percentile (bottom right panel) reveals that some members do indeed produce localized maxima that closely resemble the observed convective cores, especially near Schwenningen and north of Stuttgart, in both shape and intensity. In those areas, maximum values exceed 60 mm, in line with the observations. However, these high-precipitation members are rare and tend to overestimate spatial coverage when convection is triggered too aggressively. Nonetheless, this indicates that the model, when properly initialized and under favorable perturbation combinations, is capable of capturing the event's key features, especially the intense but short-lived nature of the rainfall.

Overall, these results underline a fundamental contrast with the Marche case: while that event was characterized by a large, persistent convective structure with a strong mesoscale organization, here we deal with small-scale, rapidly developing cells whose predictability relies more on precise timing and local convergence than on synoptic-scale dynamics. As such, traditional diagnostics such as total precipitation may not be sufficient to evaluate ensemble performance.

Given the observed high spatial variability and the importance of microphysical interactions in modulating hail production, in the following section you can find the analysis of microphysical process rates. In particular, there will be a focus on the processes related to hail formation and growth, such as accretion, riming and freezing, to understand which ensemble members exhibit stronger hail-related microphysics, and whether these correlate with the precipitation extremes seen in the 99th percentile. This process-level analysis

is crucial to explain why only some members correctly captured the intense localized precipitation while others failed to initiate convection altogether. It also provides insight into the specific role played by the perturbed parameters in the two-moment scheme, especially CCN and INP concentrations, which may influence hydrometeor size distributions and phase transitions critical to hail generation.

### 5.3.5 Domain-averaged microphysical processes

To further investigate the internal structure and physical drivers of the ensemble precipitation variability, the domain-averaged microphysical tendencies associated with hail growth and phase transitions are analysed. Figures 5.11, 5.12, 5.13, 5.14, and 5.15 show the evolution throughout 14 August 2023 of hail, graupel, and several relevant process rates, each averaged over the high-resolution (500 m) domain for all ensemble members.

Figure 5.11 reveals substantial inter-member variability in hail production. While many members produce negligible amounts, others (notably in groups 19–27 and 28–36) reach domain-mean values exceeding 0.2 mm. These members correspond closely to those responsible for the local maxima seen in the 99th percentile precipitation map (Figure 5.10), suggesting that they correctly captured the convective cores associated with hail. Moreover, their curves show a coherent and simultaneous onset of hail around 11 UTC, followed by a plateau or moderate increase, indicating not only more intense convection but also a realistic timing and limited duration, in agreement with observations.

The graupel accumulation (Figure 5.12) shows a similarly wide spread, though the relationship to hail is not always direct. In some limited cases graupel patterns weakly differ from hail production, indicating that convection did occur but followed a different microphysical pathway. This may be due to weaker updrafts, limited riming efficiency, or thermodynamic profiles that favoured the growth of graupel instead of hailstones.

The key process of hail growth, riming of supercooled droplets onto hail ( $\Delta q_{hr}$ ), is shown in Figure 5.13. Interestingly, members with higher hail accumulation exhibit markedly weaker riming rates, but still with increases between 11 and 13 UTC. This behaviour likely reflects the nature of the convec-

tive dynamics rather than a weaker riming process. In these members, intense but spatially confined updrafts promoted rapid growth of a limited number of hail embryos, resulting in an efficient conversion of supercooled liquid water into hail within narrow cores. Consequently, the riming process was highly effective locally but less widespread, leading to lower mean riming values when averaged over the domain. Moreover, once large hail particles formed, their subsequent growth may have proceeded mainly through accretion and freezing of raindrops, rather than by continuous riming of cloud droplets, further reducing the overall riming signal.

Importantly, these members also show a synchronous onset of riming, which matches both the observational timeline and the structure of the simulated hail fields. This suggests that the best-performing ensemble members not only captured the right physical mechanisms, but also reproduced the temporal evolution of the event correctly. After this initial burst, riming tends to stabilize, with a modest secondary peak around 23 UTC in a few members, probably linked to isolated late convective cells.

At the same time, the nucleation of ice (Figure 5.14) show a similar pattern as for the riming process. Almost all members show a gradual, sustained increase throughout the day, indicating that the initial seeding of ice crystals is not a limiting factor in this case. This uniformity also suggests that differences in INP concentrations alone do not explain the variability in hail production. Rather, it is the interaction of those early ice particles with localized updrafts and supercooled water that determines the eventual hail outcome. Anyway, the same members reporting elevated hail accumulation values, are characterised by low nucleation rates. This again points to a more localized and efficient microphysical pathway, where intense but confined convective cores rapidly activate and grow a limited number of ice particles. In this context, the lower average nucleation simply reflects a more focused initiation process consistent with the concentrated riming and hail growth previously discussed.

Lastly, the melting of cloud water (Figure 5.15) is weak and relatively uniform across the ensemble. These low values suggest that melting played a marginal role in this event, contributing mostly to stratiform rain or drizzle, but not affecting the convective hail dynamics directly.

In summary, this analysis confirms that there was a fraction of ensemble

members which simulated quite well the full chain of microphysical processes required for hail production, from ice nucleation to riming and sustained convective organization. Crucially, the best-performing members are also those showing the most realistic timing of initiation, intensity and duration of the event. This highlights that process-level metrics can offer valuable insight into ensemble skill for highly localized convective phenomena, complementing traditional metrics based on precipitation amount or spatial patterns.

## **5.4 | Comments on the correlation between the two events**

An important indication of potential generalizability emerges from the comparison between the Italian flood and the German hail case. Despite their different synoptic and hazard characteristics, both cases exhibit a consistent sensitivity to microphysical configurations, with ensemble members characterized by lower CCN and INP concentrations performing better against observations. This suggests that some microphysical controls may exert a systematic influence across different convective regimes, although their impact remains modulated by environmental conditions and storm type.



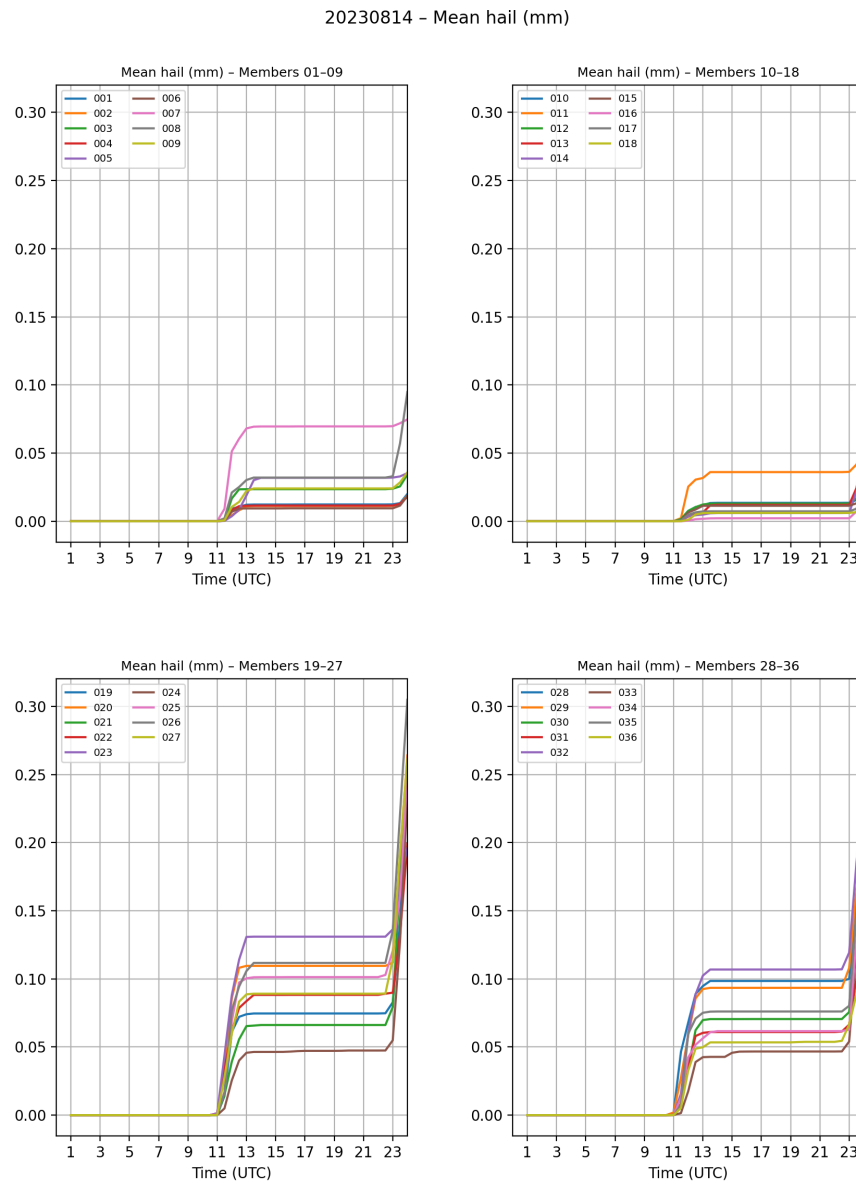


Figure 5.11: Domain-averaged cumulative hail precipitation [mm] for each ensemble member.

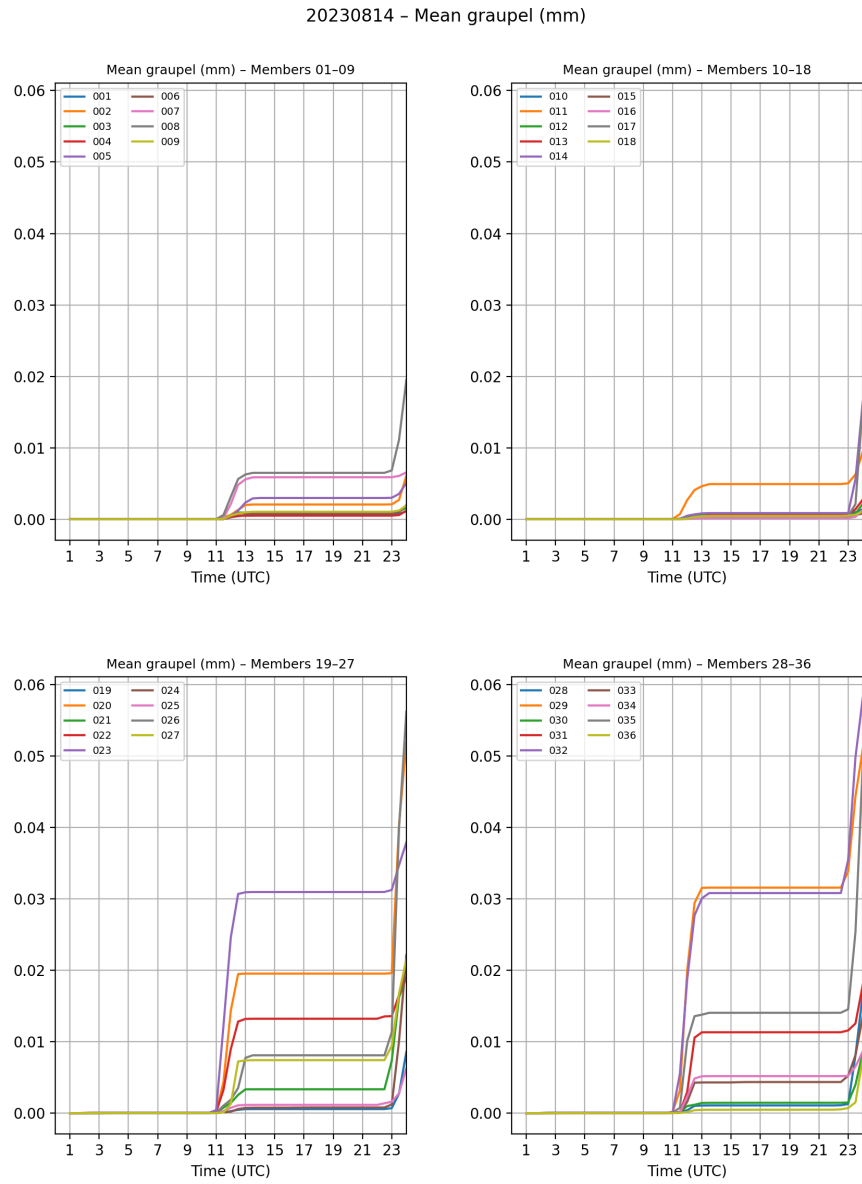


Figure 5.12: Domain-averaged cumulative graupel [mm] for each ensemble member.

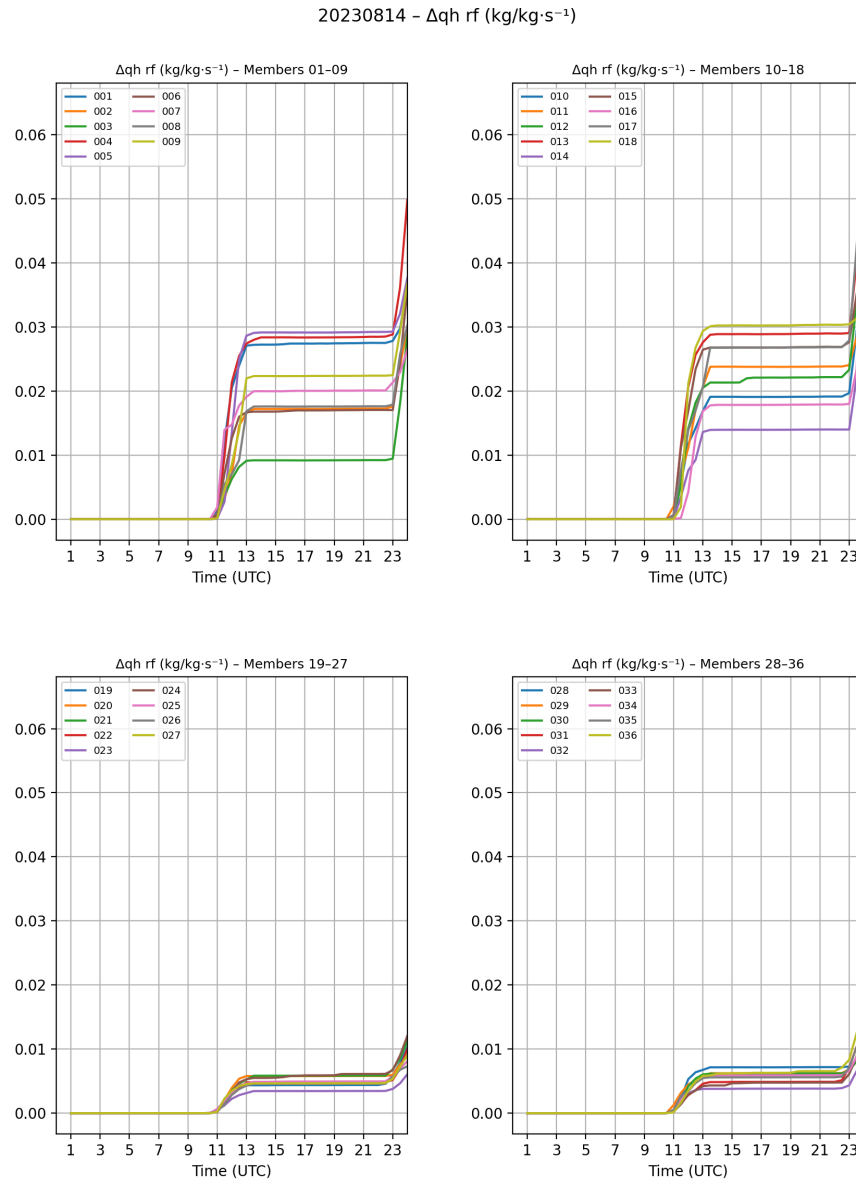


Figure 5.13: Domain-averaged riming rate on hail particles ( $\text{kg/kg/s}$ ) for each ensemble member.



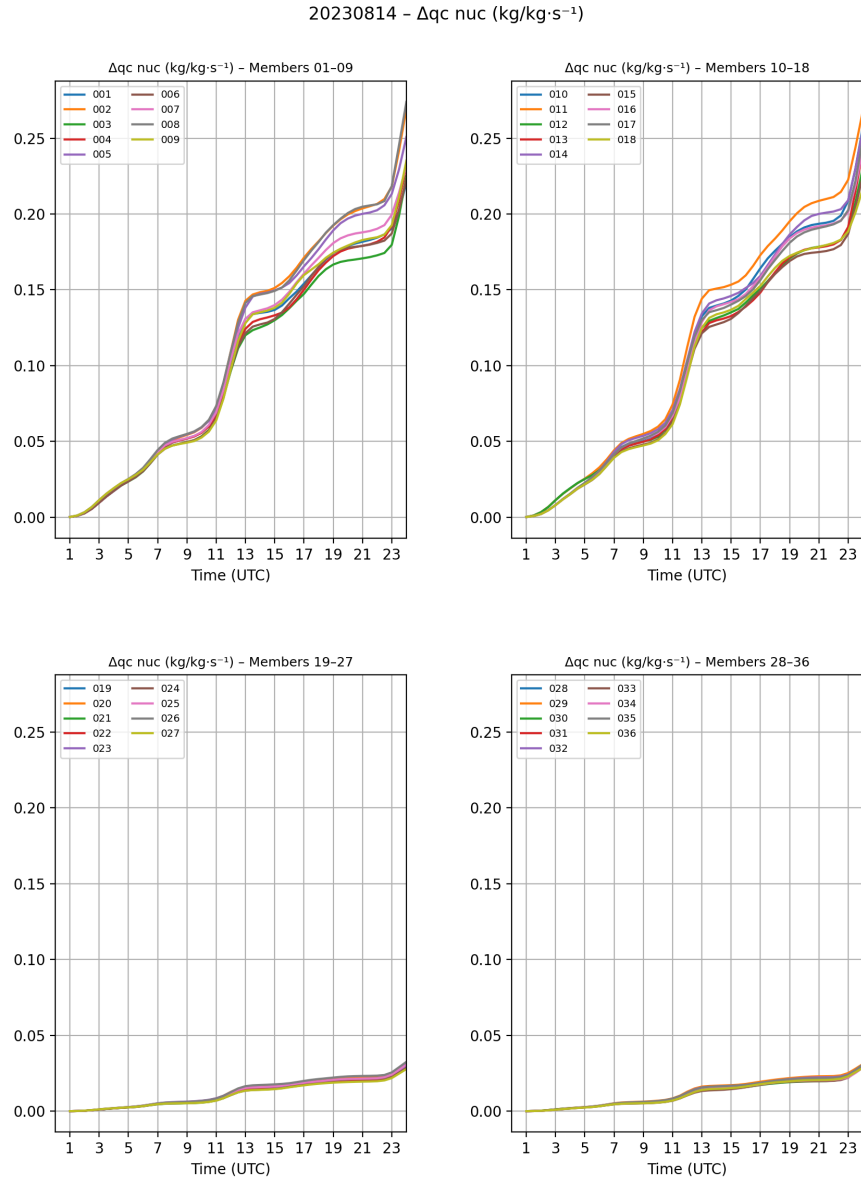


Figure 5.14: Domain-averaged nucleation rate from cloud water to ice (kg/kg/s) for each ensemble member.

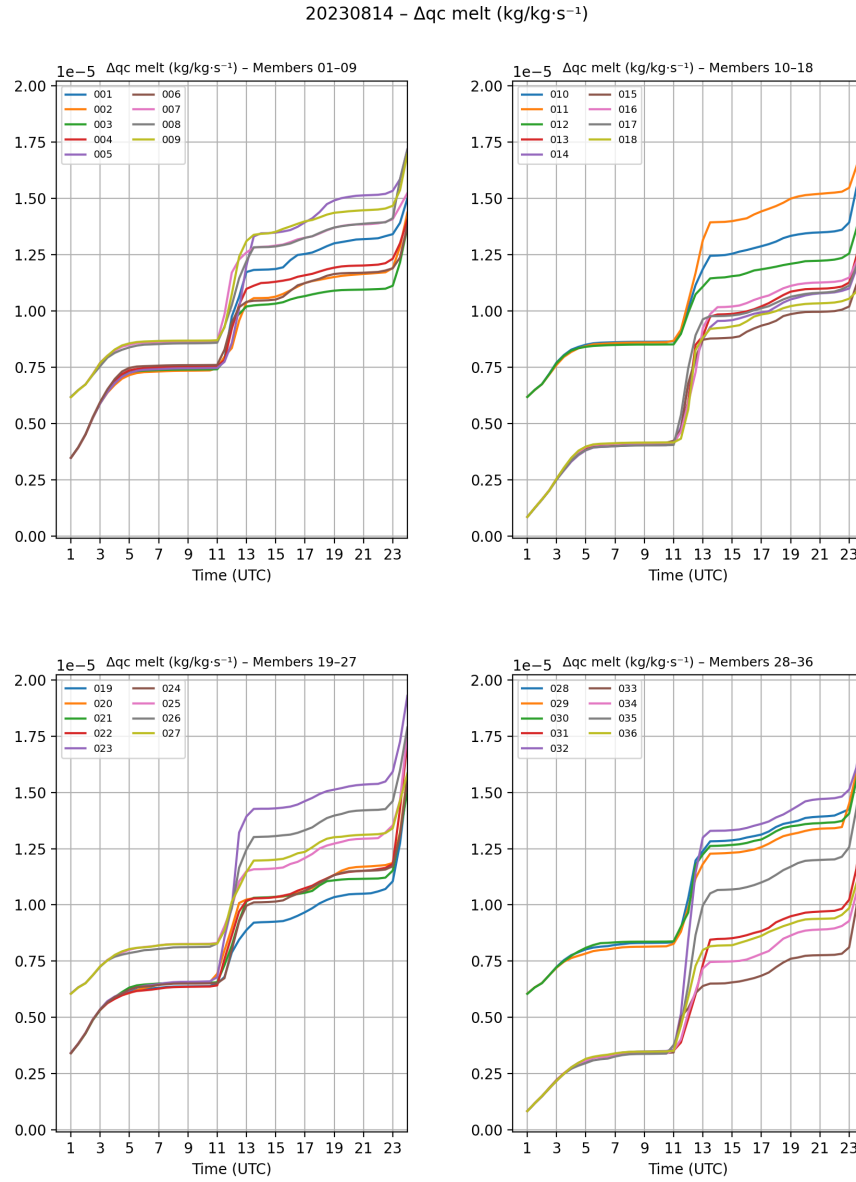


Figure 5.15: Domain-averaged melting rate of cloud water ( $\text{kg/kg/s}$ ) for each ensemble member.



## 6 | Conclusion and future perspectives

Forecasting extreme convective precipitation remains one of the grand challenges of atmospheric science and operational meteorology. As climate change is expected to intensify the frequency and impact of short-lived but devastating events such as flash floods and large hailstorms, the ability to anticipate their onset, evolution, and consequences becomes ever more critical. This thesis has explored the potential and limitations of high-resolution NWP using the ICON model, with a particular focus on hectometric-scale simulations and ensemble design. Two contrasting case studies, the catastrophic September 2022 flood in the Marche region of Italy and the localized hailstorm over Villingen-Schwenningen in August 2023, provided a fertile ground to test model resolution, turbulence and microphysics schemes, and different ensemble generation strategies.

The work presented here allows to draw several major conclusions that could help improve the scientific research towards better initialization and simulation of severe and unpredictable convective storms that can impact society in multiple ways.

### 6.1 | The dominance of IC/LBC uncertainty

One of the clearest outcomes of this thesis is that the predictability of extreme convective events is still fundamentally constrained by the uncertainty in the initial and boundary conditions. Even with hectometric grid spacings, turbulence closures tuned to resolve small-scale mixing, and advanced microphysical schemes, the forecasts remain strongly dependent on whether the large-scale and mesoscale precursors of the event are correctly represented in the initial state (Buizza et al., [1999](#); Richardson, [2000](#); Selz & Craig, [2015](#)).

This behaviour is well illustrated by the September 2022 Marche flood. In this case, most ECMWF ensemble members provided ICs and BCs that, once tuned for the nested ensemble simulations, struggled in capturing the sharp inland convergence zone that triggered the back-building convection. However, one member, namely **m2**, deviated significantly from the ensemble mean and happened to capture a stronger and more realistic precipitation pattern. The resulting 550 m nested forecast displayed a striking resemblance to the observed event: the position of the convergence line, the persistence of convective regeneration, and the intensity of the accumulated rainfall all aligned closely with radar and rain-gauge observations.

This example underscores a key message: hectometric resolution alone is not a guarantee of forecast success. Instead, the combination of *ensemble diversity* at the IC level and *physical realism* at the hectometric scale determines whether a forecast can reproduce an extreme event. In other words, the “raw material” provided by the initial state must already contain the seeds of the correct mesoscale evolution; the high-resolution nesting then acts to refine and amplify this signal, translating it into realistic convective structures and rainfall patterns. This is consistent with previous findings that convective-scale predictability is largely driven by upscale error growth from small perturbations in the initial state (Dey et al., 2016; Hohenegger et al., 2006).

In this sense, again, the results should be interpreted in the context of the intrinsic limits of atmospheric predictability. As discussed in section 1.1.1, the atmosphere is a chaotic dynamical system, and the rapid amplification of small-scale perturbations imposes a fundamental constraint on deterministic forecasts, particularly for convective-scale phenomena. Thus, the improvements obtained through increased horizontal resolution and advanced physical parameterizations do not eliminate uncertainty. Convection-permitting ensembles therefore emerge not as a means to overcome chaos, but as a strategy to represent it. The enhanced realism of storm structure and intensity documented in this work should thus be understood as an improvement in probabilistic and structural guidance, rather than as an extension of deterministic predictability. This perspective is essential when translating high-resolution ensemble output into information for impact-based forecasting and early warning systems. It follows that from an operational perspective, these results highlight the continued need for well-designed ensemble systems that adequately sample the

uncertainty in the larger-scale flow. Without such diversity, even the most advanced high-resolution models will be limited in their ability to anticipate high-impact extremes.

## 6.2 | Added value of hectometric-scale ensembles

While ICs and BCs uncertainties dominate the large-scale setup of the event's initiation, our results also show that hectometric resolution provides substantial added value once the correct mesoscale scenario is present. This added value manifests in several ways: more realistic structures of rainfall objects, improved timing of convective initiation and back-building, and reduced displacement errors in object-based scores such as SAL (Davis et al., 2009; Roberts & Lean, 2008).

Returning to the Marche case, the member **m2** demonstrates this synergy particularly well. At 2 km resolution, the forecast driven by m2 already suggested a higher risk of intense rainfall. When nested to 550 m and combined with the 3-D Smagorinsky turbulence closure, the same IC realization produced a forecast where convergence sharpening and localized convective regeneration were explicitly resolved. This resulted in a more faithful reproduction of the observed rainfall distribution and accumulation.

Such examples confirm that hectometric simulations are not simply “higher resolution versions” of the same forecast, but can qualitatively change the way of interaction between processes. By resolving narrower up and downdrafts, finer-scale turbulence, and more realistic microphysical feedbacks, they shift the balance between storm organization and dissipation. However, these benefits only emerge when the large-scale forcing is conducive to the event; otherwise, high resolution alone cannot correct an unfavourable initial condition.

In ensemble terms, this means that the optimal strategy is not to rely solely on a larger number of perturbed IC members at kilometer scale, nor exclusively on deterministic high-resolution runs. Instead, the key is to design ensembles that combine both: a sufficiently broad sampling of ICs uncertainty to ensure that “good” members like m2 are present, and targeted hectometric nests that can refine those promising scenarios into forecasts of genuine predictive value. This is where operational systems can benefit most, striking a balance between

ensemble spread and computational cost (Clark et al., 2016; Raynaud & Bouttier, 2017).

Anyway, although this thesis does not aim to define a universally optimal ensemble configuration, the results provide insight into how ensembles at hectometric resolution could be designed more effectively. The analysis indicates that physical realism, achieved through adequate resolution and carefully chosen parameterization schemes, represents a prerequisite for meaningful ensemble forecasts. Ensemble diversity should therefore be introduced in a targeted manner, focusing on physically relevant sources of uncertainty rather than on maximizing ensemble size. These findings support a regime-dependent approach to ensemble design, where ensemble structure is adapted to the dominant physical processes of the event under consideration.

## 6.3 | Turbulence closures and microphysics uncertainty

The sensitivity of hectometric simulations to physical parameterizations emerged as a second key result. Switching from a COSMO-type 1D TKE closure to a 3D Smagorinsky turbulence scheme altered the organization of convective lines, and the leeward displacement of rainfall, especially for the Italian case, leading to more realistic precipitating patterns. This confirms that turbulence closures strongly influence storm dynamics at the high resolution scales over complex terrains between LES and parameterized convection (Deardorff, 1980; Smagorinsky, 1963).

In addition, perturbations to microphysical parameters (e.g., CCN, INP, shape parameter of the PSD, graupel sedimentation velocity) introduced substantial ensemble spread, particularly in the German hail case. Such perturbations controlled not only precipitation totals but also the partitioning between graupel and hail, influencing the timing and intensity of hail growth pathways. This aligns with recent work showing that microphysics remains one of the dominant uncertainties in convection-permitting forecasts (Morrison et al., 2020).

## 6.4 | Contrasting hazards: floods versus hail

The two case studies analysed in this thesis, the catastrophic flood in the Marche region (September 2022) and the hailstorm over Villingen–Schwenningen (August 2023), illustrate the very different challenges posed by extreme convective hazards. Despite sharing common drivers such as orographic lifting and mesoscale convergence, the events diverged strongly in their spatio-temporal scales and impacts.

The Marche flood was characterised by a long-lived, back-building convective system producing sustained rainfall accumulations exceeding 400 mm in few hours. Its predictability hinged on the accurate simulation of low-level wind convergence, moisture transport and turbulence parameterization over complex terrain, with small errors in boundary-layer dynamics leading to large displacements of the precipitation. In this case, ensemble diversity in ICs and hectometric-scale refinement were both essential to reproduce the persistent regeneration of convective cells.

In contrast, the German hailstorm was short-lived and spatially confined, with peak precipitation below 70 mm but producing hailstones that caused damage. Here, predictability depended less on prolonged organisation and more on the correct timing and intensity of isolated convective updrafts. Microphysical parameter perturbations such as CCN and INP concentrations, strongly modulated whether hail growth pathways were activated. Ensemble spread therefore gave promising insights primarily from physics perturbations.

Taken together, these results highlight that extreme rainfall and hail, though both products of deep convection, require different forecasting emphases: floods are more sensitive to mesoscale dynamical predictability, while hail is more strongly tied to microphysical uncertainty and the representation of particle growth processes.



## 6.5 | Implications for forecasting and early warnings

The results of this thesis also have important implications for how we design forecasting systems and early-warning strategies. A first clear message is that ensemble prediction remains the essential backbone for severe-weather forecasting. Even the most advanced high resolution deterministic runs cannot capture the full range of possible evolutions of a convective storm. What is really needed is a combination: on the one hand, a sufficiently broad ensemble of perturbed ICs to ensure that the large-scale and mesoscale environment is well represented, and on the other hand, targeted high-resolution nests that can refine those favourable members into realistic storm structures.

However, among the different factors investigated, increasing horizontal resolution emerges as the most effective strategy to improve forecast quality, outweighing the benefits of increasing ensemble size alone. Indeed, a higher number of ensemble members becomes a meaningful strategy mainly once a physically adequate resolution is achieved and when initial condition uncertainty is the dominant error source. In contrast, the results of this work suggest that investing in higher resolution represents a more effective and robust pathway to improving forecast quality for high-impact convective events.

Ultimately, direct consequences can be imagined for early warning systems: probabilistic products from hectometric-scale ensembles can provide not only where and when storms are most likely, but also which type of impact is more probable, for example intense rainfall versus large hail. Integrating this kind of guidance with dense observations and nowcasting techniques could help to extend lead times and communicate risk more effectively, especially in areas with complex terrain where predictability is most limited.

## 6.6 | Future outlooks and generalized conclusions

The results of this thesis show both the potential and the limits of hectometric-scale simulations for forecasting extreme convective events. Looking ahead, several directions appear particularly important.

As highlighted by the Marche flood case, the accuracy of the data describing the large-scale flow provided to the model can make the difference between success and failure, regardless of how fine the resolution is. This suggests that future progress cannot rely on resolution alone, but will require ensemble systems that better capture mesoscale precursors and that can provide a sufficiently diverse set of members to feed high-resolution nests. Adaptive ensemble strategies, where computational resources are focused on the most promising scenarios, may offer a way forward.

A second lesson comes from the sensitivity to turbulence and microphysics. The experiments carried out here clearly show that hectometric forecasts depend strongly on how these processes are represented, and that even subtle changes in closures or parameter values can reshape storm organisation, rainfall persistence, or hail growth. This means that tuning high-resolution models is not a marginal task, but a central one: scale-aware turbulence schemes, more realistic microphysics, and systematic calibration against dense observations will all be necessary.

Finally, the question of resolution itself remains open. Horizontal grid spacings of 550 m already improved the representation of narrow convergence lines and convective regeneration, but the vertical discretisation still appears as a limiting factor. A finer layering in the boundary layer and in the upper troposphere could improve the depiction of turbulence, moisture transport, and hail growth pathways. Pushing both horizontal and vertical resolution a step further, even if only incrementally, may unlock important advances in forecast realism. Thus we need a careful combination: diverse and well-designed initial conditions, finely tuned turbulence and microphysics schemes, and gradual improvements in resolution. If pursued together, they could bring forecasting systems closer to anticipating not only where and when convective extremes will occur, but also how they will evolve and what kind of impact they will have.

A brief discussion could be done on how Artificial Intelligence could impact such kind of high-resolution weather forecasting research. Recent advances have demonstrated substantial potential for improving weather prediction at synoptic scales, with AI-based models achieving competitive performance with traditional numerical systems for large-scale atmospheric flow (Bi et al., 2023; Lam et al., 2023). However, current data-driven approaches operate

at relatively coarse resolutions and do not explicitly resolve convective-scale processes, highlighting the continued importance of physically based models and hybrid AI-NWP strategies for high-impact weather prediction (Rasp et al., 2020). However, since AI does not yet directly resolve convection, research on hectometric-scale physical models remains crucial.

As concerning the generalization of the results presented in this thesis, it must be interpreted in a regime-dependent sense. The analysed events are both characterized by strong mesoscale forcing, significant moisture convergence, and a pronounced interaction with complex terrain. Under these conditions, increasing horizontal resolution to hectometric scales and adopting advanced three-dimensional turbulence closures consistently improves the spatial organization and localization of intense precipitation. However, such improvements cannot be assumed to be universal. In weakly forced convective regimes, or in environments dominated by large-scale dynamics rather than orographic control, increased resolution and reduced diffusion may amplify small-scale noise and lead to less robust forecasts. Consequently, the findings of this work should be interpreted as particularly relevant for high-impact convective events in complex terrain, rather than as a universally optimal modelling strategy.

Beyond its methodological and meteorological scope, this work also fits into a much broader context of understanding and managing the complex relationship between the atmosphere, the climate system, and society. The doctoral course that I attended precisely aims at building these bridges, and the study of convective extremes represents one of the most direct and tangible ways in which atmospheric science intersects with societal concerns. Indeed, extreme rainfall, flash floods, and severe hailstorms are not abstract phenomena: they have profound implications for safety, infrastructure, and territorial planning; therefore, improving their predictability is not only a question of physical understanding, but also an essential step toward reducing vulnerability and increasing resilience in a changing climate. In this sense, the high-resolution modelling experiments and ensemble approaches explored in this thesis can be seen as part of a wider effort to link scientific progress with practical action.

At the same time, this work highlights how scientific disciplines cannot operate in isolation, since the improvement of convective forecasting requires not only better physical models, but also closer integration with observational networks, risk communication strategies, and the social sciences that study

how communities perceive and respond to warnings. Future advances will increasingly depend on the ability to connect these different perspectives, from atmospheric dynamics to human decision-making, within a coherent framework.





## A | Parameters description

Table A.1 reports the main configuration parameters adopted in these ICON simulations for the three nested domains. These parameters regulate both the numerical integration and the treatment of physical processes. They can be broadly divided into two categories: (i) parameters that define temporal intervals, i.e. how frequently a given model component is called; and (ii) parameters that control the choice and behavior of the physical parameterizations.

| Parameter               | Domain 1 | Domain 2 | Domain 3 |
|-------------------------|----------|----------|----------|
| start_time              | 0        | 900      | 1800     |
| num_lev                 | 65       | 65       | 65       |
| dtype                   | 20       | 10       | 5        |
| inwp_gscp               | 2/4      | 2/4      | 2/4      |
| inwp_conv               | 1        | 1        | 1        |
| grayzone_conv           | true     | true     | false    |
| shallow_conv            | false    | false    | true     |
| inwp_rad                | 4        | 4        | 4        |
| inwp_turb               | 1        | 1        | 1 (5)    |
| inwp_satad              | 1        | 1        | 1        |
| inwp_qvcd               | 0        | 0        | 0        |
| inwp_surf               | 1        | 1        | 1        |
| dt_rad                  | 720      | 720      | 720      |
| dt_conv                 | 120      | 120      | 60       |
| dt_eso                  | 120      | 120      | 120      |
| dt_gwd                  | 360      | 360      | 360      |
| lsmag_3d                | false    | false    | true     |
| veg_cycle               | 2        |          |          |
| lhdifff_smag_w          | false    | false    | true     |
| itycecanopy             | 2        |          |          |
| hgtdiff_max_smooth_topo | 750      | 400      | 200      |
| n_filter_smooth         | 1        | 1        | 1        |

Table A.1: Main ICON configuration parameters, whose description can be found in paragraph A, for the three nested domains used in all experiments. Domain 1 = outer (2.2 km), Domain 2 = intermediate (1.1 km), Domain 3 = inner (550 m).

- **Time-related parameters.**

- **start\_time**: defines the delay (in seconds) after which each nested domain becomes active with respect to the beginning of the simulation.
- **dtime**: sets the model integration time step in seconds.
- **dt\_rad**, **dt\_conv**, **dt\_eso**, and **dt\_gwd**: specify the call frequency (in seconds) of the radiation scheme, convection parameterization, surface exchange processes, and gravity-wave drag scheme, respectively.
- **num\_lev**: indicates the total number of vertical levels used in the model atmosphere.

- **Physical parameterizations.**

- **inwp\_gscp**: selects the microphysics scheme. In this setup, values of 2 and 4 correspond to the single-moment and double-moment microphysics schemes, respectively.
- **inwp\_rad**: chooses the radiation scheme. Here it is set to 4, activating ECRAD, the ECMWF radiation scheme recently introduced in ICON.
- **inwp\_conv**, **grayzone\_conv**, and **shallow\_conv**: control the representation of cumulus convection. These switches can activate the standard deep-convection parameterization, the “gray-zone” configuration for partially resolved convection, or the shallow-cumulus scheme. In this configuration, they act as simple on/off flags depending on the domain resolution.
- **inwp\_surf**: defines the surface coupling method, i.e. how the land surface scheme interacts with the atmosphere.

- **Turbulence and diffusion.**

- **inwp\_turb**: specifies the turbulence closure. A value of 1 activates the traditional one-dimensional COSMO **TURBDIFF/TURBTRANS** scheme, while a value of 4 selects the fully 3D **Smagorinsky** closure, in which turbulent mixing is parameterized through a 3D diffusion tensor.
- **lsmag\_3d**: activates explicit 3D Smagorinsky diffusion, further enhancing the representation of turbulent mixing.

- 
- **hdiff\_smag\_w**: controls the amount of horizontal diffusion applied to the vertical velocity component.
  - **Additional switches.**
    - **inwp\_satad**: activates the saturation adjustment process.
    - **inwp\_qvcd**: applies corrections to the water-vapour content.
    - **ityccanopy**: defines how the canopy layer is represented in the surface scheme.
    - **veg\_cycle**: specifies how vegetation cycles are treated. This parameter controls the representation of the seasonal vegetation cycle in ICON's land-surface scheme: when set to 1, the model employs a simplified, static vegetation cycle, where vegetation-related parameters such as leaf area index, fraction of green vegetation, and albedo remain constant throughout the simulation; when set to 2, a seasonally varying vegetation cycle is activated, in which vegetation properties evolve dynamically according to climatological or model-based annual cycles, leading to a more realistic depiction of vegetation variability and its feedbacks on surface fluxes.
    - **hgtdiff\_max\_smooth\_topo** and **n\_filter\_smooth**: control the orographic smoothing. The first sets the maximum vertical smoothing length (in meters), while the second defines the number of filter iterations applied.

Overall, this configuration ensures a consistent and physically coherent treatment of processes across all domains, while maintaining flexibility and dealing with the necessities of the available HPC resources without losing accuracy.





# Bibliography

Ahrens, B., & Leps, N. (2021). Sensitivity of convection permitting simulations to lateral boundary conditions in idealized experiments [Open Access, First published: 06 December 2021]. *Journal of Advances in Modeling Earth Systems*.  
<https://doi.org/10.1029/2021MS002519>.

Amengual, A. (2022). Hydrometeorological analysis of the 12 and 13 september 2019 widespread flash flooding in eastern spain. *Natural Hazards and Earth System Science*, 22, 1159–1179.  
<https://doi.org/10.5194/nhess-22-1159-2022>.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the cosmo model: Description and sensitivities. *Monthly Weather Review*, 139(12), 3887–3905.  
<https://doi.org/10.1175/MWR-D-10-05013.1>.

Barker, D. M., & Clayton, A. M. (2011). Hybrid variational–ensemble data assimilation [Available online]. *ECMWF Seminar on Data Assimilation for Atmosphere and Ocean*.  
<https://www.ecmwf.int/sites/default/files/elibrary/2012/7934-hybrid-variational-ensemble-data-assimilation.pdf>.

Barthlott, C., Czajka, B., Kunz, M., Saathoff, H., Zhang, H., Böhmmländer, A., Gasch, P., Handwerker, J., Kohler, M., Wilhelm, J., Wieser, A., & Hoose, C. (2024). The impact of aerosols and model grid spacing on a supercell

- storm from swabian mooses 2021 [First published: 18 March 2024]. *Quarterly Journal of the Royal Meteorological Society*.  
<https://doi.org/10.1002/qj.4687>.
- Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., & et al. (2017). The harmonie-arome model configuration in the aladin–hirlam nwp system. *Monthly Weather Review*, 145(5), 1919–1935.  
<https://doi.org/10.1175/MWR-D-16-0417.1>.
- Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 619, 533–538.  
<https://doi.org/10.1038/s41586-023-06185-3>.
- Bonavita, M., Hólm, E., Isaksen, L., & Fisher, M. (2015). The evolution of the ecmwf hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 142(698), 626–643.  
<https://doi.org/10.1002/qj.2652>.
- Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., & Beare, S. E. (2008). The mogreps short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 134(632), 703–722.  
<https://doi.org/10.1002/qj.234>.
- Bryan, G. H., Wyngaard, J. C., & Fritsch, J. M. (2003). Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, 131(10), 2394–2416.  
[https://doi.org/10.1175/1520-0493\(2003\)131<2394:RRFTSO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2).

- Buizza, R., Milleer, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ecmwf ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560), 2887–2908.  
<https://doi.org/10.1002/qj.49712556006>.
- Cadena SER. (2024). *Decenas de muertos y miles de millones en daños: 2024 es un año de impacto récord por las inundaciones en europa* [Accessed: 2025-07-17].  
<https://cadenaser.com/nacional/2024/10/30/decenas-de-muertos-y-miles-de-millones-en-danos-2024-es-un-ano-de-impacto-record-por-las-inundaciones-en-europa-cadena-ser/>.
- Caldas-Alvarez, A., Feldmann, H., Lucio-Eceiza, E., & Pinto, J. G. (2023). Convection-parameterized and convection-permitting modelling of heavy precipitation in decadal simulations of the greater alpine region with cosmo-clm. *Weather and Climate Dynamics*, 4, 543–565.  
<https://doi.org/10.5194/wcd-4-543-2023>.
- Capecchi, V. (2021). Reforecasting two heavy-precipitation events with three convection-permitting ensembles. *Weather and Forecasting*, 36(3), 769–790.  
<https://doi.org/10.1175/WAF-D-20-0130.1>.
- Chow, F. K., Schär, C., Ban, N., Lundquist, K. A., Schlemmer, L., & Shi, X. (2019). Crossing multiple gray zones in the transition from mesoscale to microscale simulation over complex terrain. *Atmosphere*, 10(5), 274.  
<https://doi.org/10.3390/atmos10050274>.
- Clark, P., Roberts, N., Lean, H., Ballard, S. P., & Charlton-Perez, C. (2016). Convection-permitting models: A step-change in rainfall forecasting. *Meteorological Applications*, 23(2), 165–181.  
<https://doi.org/10.1002/met.1538>.

- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3-4), 613–626.  
<https://doi.org/10.1016/j.jhydrol.2009.06.005>.
- Cui, W., Galarneau, T. J., & Hoogewind, K. A. (2024). Changes in mesoscale convective system precipitation structures in response to a warming climate. *Journal of Geophysical Research: Atmospheres*, 129(9), e2023JD039920.  
<https://doi.org/10.1029/2023JD039920>.
- Dallan, E., Marra, F., Fosser, G., Marani, M., Formetta, G., Schär, C., & Borga, M. (2023). How well does a convection-permitting regional climate model represent the reverse orographic effect of extreme hourly precipitation? *Hydrology and Earth System Sciences*, 27(5), 1133–1149.  
<https://doi.org/10.5194/hess-27-1133-2023>.
- Davis, C., Brown, B., Bullock, R., & Halley-Gotway, J. (2009). The method for object-based diagnostic evaluation (mode) applied to numerical forecasts from the 2005 nssl/spc spring program. *Weather and Forecasting*, 24(5), 1252–1267.  
<https://doi.org/10.1175/2009WAF2222241.1>.
- De Martin, F., Miglietta, M. M., Gastaldo, T., Martinazzo, M., Pavan, F., Siena, M., & Di Sabatino, S. (2025). The bayesian sinking in porticello: A predictable convective windstorm? *Weather*.  
<https://doi.org/10.1002/wea.7715>.
- Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorology*, 18, 495–527.  
<https://doi.org/10.1007/BF00119502>.

- Dey, S. R. A., Plant, R. S., Roberts, N. M., & Migliorini, S. (2016). Assessing spatial precipitation uncertainties in a convective-scale ensemble. *Quarterly Journal of the Royal Meteorological Society*, 142(701), 2935–2948.  
<https://doi.org/10.1002/qj.2893>.
- Doswell, C. A., Brooks, H. E., & Maddox, R. A. (1996). Flash flood forecasting: An ingredients-based methodology. *Weather and Forecasting*, 11(4), 560–581.  
[https://doi.org/10.1175/1520-0434\(1996\)011<0560:FFFAIB>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0560:FFFAIB>2.0.CO;2).
- Doswell, C. A. (2015). Severe convective storms in the european societal context. *Atmospheric Research*, 158–159, 210–215.  
<https://doi.org/10.1016/j.atmosres.2014.08.007>.
- Ebert, E. (2008). Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. *Meteorological Applications*, 15, 15–64.  
<https://doi.org/10.1002/met.25>.
- ESSL. (2023). Hailstorms of 2023 in europe [Accessed: 17 August 2025].  
<https://www.essl.org/cms/hailstorms-of-2023/>.
- ESSL. (2024). *Essl annual report 2023* (tech. rep.) (Accessed: 17 August 2025). ESSL.  
[https://www.essl.org/cms/wp-content/uploads/ESSL-AnnualReport2023\\_final\\_fully\\_approved\\_version.pdf](https://www.essl.org/cms/wp-content/uploads/ESSL-AnnualReport2023_final_fully_approved_version.pdf).
- Fosser, G., Gaetani, M., Kendon, E. J., et al. (2024). Convection-permitting climate models offer more certain extreme rainfall projections. *npj Climate and Atmospheric Science*, 7, 51.  
<https://doi.org/10.1038/s41612-024-00600-w>.

- Frogner, I.-L., Andrae, U., Ollinaho, P., Hally, A., Hämäläinen, K., Kauhanen, J., Ivarsson, K.-I., & Yazgi, D. (2022). Model uncertainty representation in a convection-permitting ensemble — spp and sppt in harmoneps. *Monthly Weather Review*, 150(4), 775–795.  
<https://doi.org/10.1175/MWR-D-21-0099.1>.
- Gassmann, A., & Herzog, H.-J. (2008). Towards a consistent numerical compressible non-hydrostatic model using generalized hamiltonian tools. *Quarterly Journal of the Royal Meteorological Society*, 134(635), 1597–1613.  
<https://doi.org/10.1002/qj.297>.
- Gastaldo, T., Poli, V., Marsigli, C., Cesari, D., Alberoni, P. P., & Paccagnella, T. (2021). Assimilation of radar reflectivity volumes in a pre-operational framework. *Quarterly Journal of the Royal Meteorological Society*, 147(735), 1031–1054.  
<https://doi.org/10.1002/qj.3957>.
- Gastaldo, T., Poli, V., Paccagnella, T., & Alberoni, P. P. (2022). Direct assimilation of radar reflectivity volumes in cosmo-letkf. In *Precipitation science* (pp. 757–786). Elsevier.  
<https://doi.org/10.1016/B978-0-12-822973-6.00024-X>.
- Gebhardt, C., Theis, S. E., Paulat, M., & Ben Bouallègue, Z. (2011). Uncertainties in cosmo-de precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmospheric Research*, 100(2-3), 168–177.  
<https://doi.org/10.1016/j.atmosres.2010.12.008>.
- GFZ German Research Centre for Geosciences. (2023). Swabian moose 2023 – measurement campaign on hydro-meteorological extremes [Accessed: 17 August 2025].  
<https://www.gfz.de/en/section/space-geodetic-techniques/projects/s>

- wabian-moses-2023-measurement-campaign-on-hydro-meteorological-extremes.
- Glazer, R. H., Grimmond, S., Blunn, L., Fenner, D., Lean, H., Christen, A., Morrison, W., & Looschelders, D. (2025). Hectometric-scale modelling of the urban mixed layer evaluated with a dense lidar-ceilometer network [Preprint]. *EGUsphere*.  
<https://doi.org/10.5194/egusphere-2025-2064>.
- Golding, B. W. (2009). Long lead time flood warnings: Reality or fantasy? *Meteorological Applications*, 16(1), 3–12.  
<https://doi.org/10.1002/met.123>.
- Gustafsson, N., Janjic-Pfander, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., 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.  
<https://doi.org/10.1002/qj.3179>.
- Hamrud, M., Bonavita, M., & Isaksen, L. (2014). *Enkf and hybrid gain ensemble data assimilation* (ECMWF Technical Memorandum No. 747). European Centre for Medium-Range Weather Forecasts (ECMWF). Reading, UK.  
<https://doi.org/10.1175/MWR-D-14-00333.1>.
- Handwerker, J., Barthlott, C., Bauckholt, M., Belleflamme, A., Böhmländer, A., Borg, E., Dick, G., Dietrich, P., Fichtelmann, B., Geppert, G., et al. (2025). From initiation of convective storms to their impact – the swabian mooses 2023 campaign in southwestern germany [Provisional acceptance]. *Frontiers in Earth Science*, 13.  
<https://doi.org/10.3389/feart.2025.1555755>.



- Hanley, K. E., Plant, R. S., Stein, T. H., Hogan, R. J., Nicol, J. C., Lean, H. W., et al. (2015). Mixing-length controls on high-resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, 141, 272–284.  
<https://doi.org/10.1002/qj.2356>.
- Heinze, R., Dipankar, A., Henken, C. C., Moseley, C., Sourdeval, O., Trömel, S., Xie, X., Adamidis, P., Ament, F., Baars, H., et al. (2017). Large-eddy simulations over germany using icon: A comprehensive evaluation. *Quarterly Journal of the Royal Meteorological Society*, 143(702), 69–100.  
<https://doi.org/10.1002/qj.2947>.
- Helmholtz Centre for Environmental Research (UFZ). (2023). Swabian mooses 2023 campaign overview [Accessed: 17 August 2025].  
<https://www.ufz.de/moses/index.php?en=44877>.
- Henderson, D. S., Otkin, J. A., & Mecikalski, J. R. (2021). Evaluating convective initiation in high-resolution numerical weather prediction models using GOES-16 infrared brightness temperatures. *Monthly Weather Review*, 149(4), 1183–1201.  
<https://doi.org/10.1175/MWR-D-20-0272.1>.
- Hernández, K. S., Henao, J. J., Gómez-Ríos, S., Robledo, V., & Mejía, J. F. (2025). Spatio-temporal representation of mesoscale convective systems in convection-permitting simulations over northwestern south america: Insights into rainfall overestimation. *Journal of Geophysical Research: Atmospheres*, 130(7), e2024JD042289.  
<https://doi.org/10.1029/2024JD042289>.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz, S. J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita,

- M., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.  
<https://doi.org/10.1002/qj.3803>.
- Hesselberg, A. (1925). Die gesetze der ausgeglichenen atmosphärischen bewegungen. *Beiträge zur Physik der freien Atmosphäre*, 12, 141–160.
- Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ecmwf model. *Journal of Advances in Modeling Earth Systems*, 10(8), 1990–2008.  
<https://doi.org/10.1029/2018MS001364>.
- Hohenegger, C., Lüthi, D., & Schär, C. (2006). Predictability mysteries in cloud-resolving models. *Monthly Weather Review*, 134(7), 2095–2107.  
<https://doi.org/10.1175/MWR3176.1>.
- Honnert, R. (2016). Representation of the grey zone of turbulence in the atmospheric boundary layer. *Advances in Science and Research*, 13, 63–67.  
<https://doi.org/10.5194/asr-13-63-2016>.
- Honnert, R., Couvreux, F., Masson, V., & Dávid, L. (2016). Sampling the structure of convective turbulence and implications for grey-zone parametrizations. *Boundary-Layer Meteorology*, 160(1), 133–156.  
<https://doi.org/10.1007/s10546-016-0130-4>.
- Houze, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, 42(4), RG4003.  
<https://doi.org/10.1029/2004RG000150>.

- Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter. *Physica D: Nonlinear Phenomena*, 230(1–2), 112–126.  
<https://doi.org/10.1016/j.physd.2006.11.008>.
- ICON partnership. (2024). *Icon release 2024.10*. World Data Center for Climate (WDCC) at DKRZ.  
<https://doi.org/10.35089/WDCC/IconRelease2024.10>.
- Iriza, A., Dumitrache, R., Maco, B., Huștiu, M., Fundel, F., Rieger, D., & Potthast, R. (2024). Comparison of cosmo and icon-lam high-resolution numerical forecast for romanian territory: Case studies and evaluation. *Atmósfera*, 38, 421–451.  
<https://doi.org/10.20937/ATM.53305>.
- Jézéquel, A., Faranda, D., Drobinski, P., & Lionello, P. (2025). Extreme event attribution in the mediterranean. *International Journal of Climatology*, E8799.  
<https://doi.org/10.1002/joc.8799>.
- Johnson, M., Jung, Y., Dawson, D. T., Supinie, T., Xue, M., Park, J., & Lee, Y.-H. (2018). Evaluation of unified model microphysics in high-resolution nwp simulations using polarimetric radar observations. *Advances in Atmospheric Sciences*, 35(7), 771–784.  
<https://doi.org/10.1007/s00376-017-7177-0>.
- Karlsruhe Institute of Technology. (2023). Measurement campaigns – swabian mooses 2023 [Accessed: 17 August 2025].  
<https://www.atmohub.kit.edu/english/684.php>.

- Kautz, L. A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., & Woollings, T. (2022). Atmospheric blocking and weather extremes over the euro-atlantic sector – a review. *Weather and Climate Dynamics*, 3(1), 305–336. <https://doi.org/10.5194/wcd-3-305-2022>.
- Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the earth’s surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98(1), 69–78. <https://doi.org/10.1175/BAMS-D-14-00283.1>.
- Kunz, M., Abbas, S. S., Bauckholt, M., Böhmländer, A., Feuerle, T., Gasch, P., Glaser, C., Groß, J., Hajsek, I., Handwerker, J., et al. (2022). Swabian moose 2021: An interdisciplinary field campaign for investigating convective storms and their event chains. *Frontiers in Earth Science*, 10, 999593. <https://doi.org/10.3389/feart.2022.999593>.
- Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382(6677), 1416–1421. <https://doi.org/10.1126/science.adi2336>.
- Lean, H. W., Theeuwes, N. E., Baldauf, M., Barkmeijer, J., Bessardon, G., Blunn, L., et al. (2024). The hectometric modelling challenge: Gaps in the current state of the art and ways forward towards the implementation of 100-m scale weather and climate models. *Quarterly Journal of the Royal Meteorological Society*, 150(765), 4671–4708. <https://doi.org/10.1002/qj.4858>.
- Li, J., Hsu, K., AghaKouchak, A., & Sorooshian, S. (2015). An object-based approach for verification of precipitation estimation. *International Journal of*

*Remote Sensing*, 36(2), 513–529.

<https://doi.org/10.1080/01431161.2014.999170>.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2), 130–141.

[https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, 26(4), 636–646.

[https://doi.org/10.1175/1520-0469\(1969\)026<0636:APARBN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)026<0636:APARBN>2.0.CO;2).

Luu, L. N., Vautard, R., Yiou, P., & Soubeyroux, J.-M. (2022). Evaluation of convection-permitting extreme precipitation simulations for the south of france. *Earth System Dynamics*, 13(3), 687–703.

<https://doi.org/10.5194/esd-13-687-2022>.

Madhulatha, A., Dudhia, J., Park, R.-S., Bhan, S. C., & Mohapatra, M. (2023). Effect of single and double moment microphysics schemes and change in cloud condensation nuclei, latent heating rate structure associated with severe convective system over korean peninsula. *Atmosphere*, 14(11), 1680.

<https://doi.org/10.3390/atmos14111680>.

Marche Floods ArcGIS StoryMaps. (2023). *The marche 2022 floods* [Accessed: 2025-07-17].

<https://storymaps.arcgis.com/stories/2b53ae1dbde44cd19dae9ed3a22b8aed>.

Markowski, P., & Richardson, Y. (2010). *Mesoscale meteorology in midlatitudes*. Wiley-Blackwell.

<https://doi.org/10.1002/9780470682104>.

- Marsigli, C., Boccanera, F., Montani, A., & Paccagnella, T. (2005). The cosmo-leps mesoscale ensemble system: Validation of the methodology and verification. *Nonlinear Processes in Geophysics*, 12, 527–536.  
<https://doi.org/10.5194/npg-12-527-2005>.
- Marullo, S., Serva, F., Iacono, R., Napolitano, E., di Sarra, A., Meloni, D., Monteleone, F., Sferlazzo, D., De Silvestri, L., de Toma, V., Pisano, A., Bellacicco, M., Landolfi, A., Organelli, E., Yang, C., & Santoleri, R. (2023). Record-breaking persistence of the 2022/23 marine heatwave in the mediterranean sea. *Environmental Research Letters*, 18(11), 114041.  
<https://doi.org/10.1088/1748-9326/ad02ae>.
- Mastrantonas, N., Herrera-Lormendez, P., Magnusson, L., Pappenberger, F., & Matschullat, J. (2020). Extreme precipitation events in the mediterranean: Spatiotemporal characteristics and connection to large-scale atmospheric flow patterns. *International Journal of Climatology*, 41(4), 2710–2728.  
<https://doi.org/10.1002/joc.6985>.
- Médus, E., Thomassen, E. D., Belušić, D., Lind, P., Berg, P., Christensen, J. H., Christensen, O. B., Dobler, A., Kjellström, E., Olsson, J., & Yang, W. (2022). Characteristics of precipitation extremes over the nordic region: Added value of convection-permitting modeling. *Natural Hazards and Earth System Sciences*, 22, 693–711.  
<https://doi.org/10.5194/nhess-22-693-2022>.
- Miglietta, M. M., & Rotunno, R. (2016). An ef3 multivortex tornado over the ionian region: Is it time for a dedicated warning system over italy? *Bulletin of the American Meteorological Society*, 97(3), 337–344.  
<https://doi.org/10.1175/BAMS-D-14-00227.1>.

- Milbrandt, J. A., & Yau, M.-K. (2005). A multimoment bulk microphysics parameterization. part i: Analysis of the role of the spectral shape parameter. *Journal of the Atmospheric Sciences*, 62(9), 3051–3074.  
<https://doi.org/10.1175/JAS3534.1>.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, 102(D14), 16663–16682.  
<https://doi.org/10.1029/97JD00237>.
- Morelli, S., Bonì, R., Guidi, E., De Donatis, M., Pappafico, G., & Francioni, M. (2023). L'alluvione delle marche del 15 settembre, cause e conseguenze. In C. Cencetti & L. Di Matteo (Eds.), *La dinamica fluviale. la conoscenza del fiume per la pianificazione e la salvaguardia del territorio* (pp. 136–147, Vol. 24).
- Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, 137(3), 991–1007.  
<https://doi.org/10.1175/2008MWR2556.1>.
- Morrison, H., van Lier-Walqui, M., Fridlind, A. M., Grabowski, W. W., Harrington, J. Y., Hoose, C., Korolev, A., Kumjian, M. R., Milbrandt, J. A., Pawlowska, H., et al. (2020). Confronting the challenge of modeling cloud and precipitation microphysics. *Journal of Advances in Modeling Earth Systems*, 12(8), e2019MS001689.  
<https://doi.org/10.1029/2019MS001689>.
- NOAA National Centers for Environmental Information. (2022, September). *Monthly global climate report for august 2022* [Published online September

- 2022, accessed January 20, 2025].  
<https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202208>.
- Nuissier, O., Duffourg, F., Martinet, M., Ducrocq, V., & Lac, C. (2020). Hectometric-scale simulations of a mediterranean heavy-precipitation event during the hydrological cycle in the mediterranean experiment (hymex) first special observation period (sop1). *Atmospheric Chemistry and Physics*, 20, 14649–14667.  
<https://doi.org/10.5194/acp-20-14649-2020>.
- Pastor, F., Valiente, J. A., & Estrela, M. J. (2015). Sea surface temperature and torrential rains in the valencia region: Modelling the role of recharge areas. *Natural Hazards and Earth System Sciences*, 15, 1677–1693.  
<https://doi.org/10.5194/nhess-15-1677-2015>.
- Petch, J., Short, C., Best, M., McCarthy, M., Lewis, H., Vosper, S., & Weeks, M. (2020). Sensitivity of the 2018 uk summer heatwave to local sea temperatures and soil moisture. *Atmospheric Science Letters*, 21.  
<https://doi.org/10.1002/asl.948>.
- Pham, T. V., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., Rieger, D., Zängl, G., & Früh, B. (2021). ICON in climate limited-area mode (ICON release version 2.6.1): A new regional climate model. *Geoscientific Model Development*, 14, 985–1005.  
<https://doi.org/10.5194/gmd-14-985-2021>.
- Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., Peckham, S. E., et al. (2017). The weather research and forecasting (wrf) model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, 98(8),



1717–1737.

<https://doi.org/10.1175/BAMS-D-15-00308.1>.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., & Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53(2), 323–361.  
<https://doi.org/10.1002/2014RG000475>.

Prill, F., Reinert, D., Rieger, D., & Zängl, G. (2022). *Icon tutorial: Working with the icon model* [Available from the DWD ICON documentation resources].

Prill, F. (2020). Dwd icon tools: User's guide [DWD, Offenbach, Germany]. *Deutscher Wetterdienst*.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., & Simmons, A. (2000). The ecmwf operational implementation of four-dimensional variational assimilation. part i: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society*, 126(564), 1143–1170.  
<https://doi.org/10.1002/qj.49712656415>.

Raschendorfer, M. (2016). New features of the common turbulence parameterization for cosmo and icon. Retrieved October 13, 2025, from  
[https://download.dwd.de/pub/DWD/Forschung\\_und\\_Entwicklung/CUS2016\\_presentations\\_PDF/04\\_Planetary\\_Boundary\\_Layer/02\\_Raschendorfer/Common\\_TURBDIFF\\_Raschendorfer\\_Mar\\_2016.pdf](https://download.dwd.de/pub/DWD/Forschung_und_Entwicklung/CUS2016_presentations_PDF/04_Planetary_Boundary_Layer/02_Raschendorfer/Common_TURBDIFF_Raschendorfer_Mar_2016.pdf).

- Raschendorfer, M. (2001). The new turbulence parameterization of Im. *COSMO Newsletter*, 1, 89–97.  
[www.cosmo-model.org](http://www.cosmo-model.org).
- Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2020). Weatherbench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, 12(11), e2020MS002203.  
<https://doi.org/10.1029/2020MS002203>.
- Raveh-Rubin, S., & Wernli, H. (2016). Large-scale wind and precipitation extremes in the mediterranean: Dynamical aspects of five selected cyclone events. *Quarterly Journal of the Royal Meteorological Society*, 142, 3097–3114.  
<https://doi.org/10.1002/qj.2891>.
- Raynaud, L., & Bouttier, F. (2017). The impact of horizontal resolution and ensemble size for convective-scale probabilistic forecasts. *Quarterly Journal of the Royal Meteorological Society*, 143(703), 3037–3047.  
<https://doi.org/10.1002/qj.3159>.
- Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., & Parodi, A. (2013). Extreme rainfall in the mediterranean: What can we learn from observations? *Journal of Hydrometeorology*, 14(3), 906–922.  
<https://doi.org/10.1175/JHM-D-12-083.1>.
- Reuters. (2024). *What caused deadly floods in spain? the impact of dana explained* [Accessed: 2025-07-17].  
<https://www.reuters.com/business/environment/spains-deadly-dana-weather-phenomenon-its-links-climate-change-2024-10-30/>.

- Rezacova, D., Szintai, B., Jakubiak, B., Yano, J., & Turner, S. (2015). Verification of high-resolution precipitation forecast with radar-based data. In *Parameterization of atmospheric convection* (pp. 173–214). World Scientific.  
[https://doi.org/10.1142/9781783266913\\_0024](https://doi.org/10.1142/9781783266913_0024).
- Richardson, D. S. (2000). Skill and relative economic value of the ecmwf ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 126(563), 649–667.  
<https://doi.org/10.1002/qj.49712656313>.
- Roberts, N. M., & Lean, H. W. (2008). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 136(1), 78–97.  
<https://doi.org/10.1175/2007MWR2123.1>.
- Sahlaoui, Z., Mordane, S., Wattrelot, E., & Mahfouf, J.-F. (2020). Improving heavy rainfall forecasts by assimilating surface precipitation in the convective scale model arome: A case study of the mediterranean event of november 4, 2017. *Meteorological Applications*, 27, e1860.  
<https://doi.org/10.1002/met.1860>.
- Schättler, U., Doms, G., & Schraff, C. (2021). *A description of the nonhydrostatic regional cosmo model: Cosmo user's guide, part vii* (tech. rep.) (Available at: <https://www.cosmo-model.org>). Consortium for Small-scale Modeling (COSMO).
- Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periañez, A., & Potthast, R. (2016). Kilometre-scale ensemble data assimilation for the cosmo model (kenda). *Quarterly Journal of the Royal Meteorological Society*, 142(696), 1453–1472.  
<https://doi.org/10.1002/qj.2748>.

- Seifert, A., & Beheng, K. D. (2006). A two-moment cloud microphysics parameterization for mixed-phase clouds. part 1: Model description. *Meteorology and Atmospheric Physics*, 92, 45–66.  
<https://doi.org/10.1007/s00703-005-0112-4>.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., & Masson, V. (2011). The arôme-france convective-scale operational model. *Monthly Weather Review*, 139(3), 976–991.  
<https://doi.org/10.1175/2010MWR3425.1>.
- Selz, T., & Craig, G. C. (2015). Upscale error growth in a high-resolution simulation of a summertime weather event over europe. *Monthly Weather Review*, 143(3), 813–827.  
<https://doi.org/10.1175/MWR-D-14-00140.1>.
- Severe Weather Europe. (2023). *A historic long-lived supercell storm with giant 14 cm hailstones tracks for more than 1200 km across five european countries leaves severe wind and hail damage along its path* [Accessed: 17 August 2025].  
<https://www.severe-weather.eu/global-weather/record-giant-hail-severe-weather-outbreak-europe-summer-2023-mk/>.
- Singh, S., Kalthoff, N., & Gantner, L. (2021). Sensitivity of convective precipitation to model grid spacing and land-surface resolution in icon. *Quarterly Journal of the Royal Meteorological Society*, 147(738), 2709–2728.  
<https://doi.org/10.1002/qj.4046>.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., & Huang, X.-Y. (2019). *A description of the advanced research wrf model version 4* (tech. rep.). NCAR Technical Note NCAR/TN-556+STR.  
<https://doi.org/10.5065/1dfh-6p97>.

- Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. the basic experiment. *Monthly Weather Review*, 91(3), 99–164.  
[https://doi.org/10.1175/1520-0493\(1963\)091<0099:GCEWTP>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2).
- Strobach, E., Mishra, A. K., Jangir, B., Ziv, B., Sun, R., Siegelman, L., Meroni, A. N., & Klein, P. (2024). Intensification of a rain system imparted by mediterranean mesoscale eddies. *Scientific Reports*, 14, 26810.  
<https://doi.org/10.1038/s41598-024-76767-2>.
- Tartaglione, N. (2025). Unprecedented flooding in the marche region (italy): Analyzing the 15 september 2022 event and its unique meteorological conditions. *Meteorology*, 4(1), 3.  
<https://doi.org/10.3390/meteorology4010003>.
- Taufour, M., Pinty, J.-P., Barthe, C., Vié, B., & Wang, C. (2024). Lima (v2.0): A full two-moment cloud microphysical scheme for the mesoscale non-hydrostatic model meso-nh v5-6. *Geoscientific Model Development*, 17, 8773–8798.  
<https://doi.org/10.5194/gmd-17-8773-2024>.
- The Guardian. (2023). *Weather tracker: New european record set as 19 cm hailstone found in italy* [Accessed: 17 August 2025].  
<https://www.theguardian.com/environment/2023/jul/31/weather-tracker-new-european-record-set-as-19cm-hailstone-found-in-italy>.
- Tomassini, L. (2020). The interaction between moist convection and the atmospheric circulation in the tropics. *Bulletin of the American Meteorological Society*, 101(8), E1378–E1396.  
<https://doi.org/10.1175/BAMS-D-19-0180.1>.

- Torma, C. Z., & Giorgi, F. (2024). Convection permitting regional climate modelling over the carpathian region. *Earth Systems and Environment*, 9, 475–490. <https://doi.org/10.1007/s41748-024-00467-0>.
- Vela, N., Cremonini, R., Oberto, E., & Giorcelli, M. (2013). Cosmo model validation using the italian radar mosaic and the rain gauges estimated precipitation [Available from: [https://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter13/cnl13\\_11.pdf](https://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter13/cnl13_11.pdf)]. *COSMO Newsletter*, (13).
- Wacker, U., & Herbert, F. (2003). Continuity equations as expressions for local balances of masses in cloudy air. *Tellus A: Dynamic Meteorology and Oceanography*, 55(3), 247–254. <https://doi.org/10.1034/j.1600-0870.2003.00019.x>.
- Weisman, M. L., & Klemp, J. B. (1984). The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Monthly Weather Review*, 112(12), 2479–2498. [https://doi.org/10.1175/1520-0493\(1984\)112<2479:TSACON>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2479:TSACON>2.0.CO;2).
- Wernli, H., Paulat, M., Hagen, M., & Frei, C. (2008). Sal - a novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, 136, 4470–4487. <https://doi.org/10.1175/2008MWR2415.1>.
- Willis Towers Watson. (2024). *Gargantuan hail in northern italy: Natural climate variability or climate change?* [Accessed: 17 August 2025]. <https://www.wtwco.com/en-us/insights/2024/04/gargantuan-hail-in-northern-italy-natural-climate-variability-or-climate-change>.

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The icon (icosahedral nonhydrostatic) modelling framework of dwd and mpi-m: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141(687), 563–579.  
<https://doi.org/10.1002/qj.2378>.

Zdunkowski, W., & Bott, A. (2003). *Dynamics of the atmosphere: A course in theoretical meteorology* (1st). Cambridge University Press.

# Acknowledgements and credits

For this long and winding journey, I would first like to thank my supervisors, **Silvana and Paolo**, who have been able to gradually guide me toward a project that truly matched my curiosity and interests, while never losing sight of the importance of developing something with a meaningful impact on the scientific research community in the field of meteorology.

A heartfelt thank you goes to **Thomas and Chiara**, whose guidance has been essential at every single stage since the initial setup of the ICON model experiments. Their patience, experience, practical help, and constant support have been truly fundamental pieces in the success of this work. In this context, I would also like to express my gratitude to all the staff at **ARPAE-SIMC** and **ItaliaMeteo**, with whom I had the pleasure to collaborate and who kindly provided both physical and technical spaces that greatly facilitated the development and completion of this research.

I would also like to acknowledge **ATOS Italia S.p.A.** for supporting this Ph.D. through its funding, and **CINECA** for the technical assistance provided.

To **Christian**, for all the support during my research stay in Karlsruhe, from the very moment he welcomed me to the Department of Physics at the South Campus of Karlsruhe, showing me from the department's rooftop the view over the city that, in just a few months, would make me grow immensely, both professionally and personally. He has always been available to guide me, technically and humanly, throughout the complex and fascinating research that has greatly enriched the outcome of this thesis.



I would also gratefully acknowledge the computing time provided on the high-performance computer **HoreKa** by the **National High-Performance Computing Center at KIT (NHR@KIT)**. This center is jointly supported by the Federal Ministry of Education and Research and the Ministry of Science, Research and the Arts of Baden-Württemberg, as part of the National High-Performance Computing (NHR) joint funding program. HoreKa is partly funded by the **German Research Foundation (DFG)**.

To all of you from **Room 171**: this journey has been far more enjoyable and bearable thanks to the laughter, the wandering around, and the work we shared together.

**To Sara and Chiara**, among the few certainties of these three years. It is hard to realize how rare it is to find people so genuinely kind and understanding, with whom to share time and everyday spaces in such a hectic city.

**To my parents**, who have always made sure I never lacked anything, especially their support in the darkest and loneliest moments. Despite our differences and misunderstandings, I know I will always find a safe harbor with them.

**Ciao Bea**, you will probably never be able to read these words, but if today I know how to love the other, it is only thanks to you.

And now, I'd say it's time to go dance some tango.