



HAL
open science

Evaluation of Alpine-Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms

Sebastian K. Müller, Cécile Caillaud, Steven Chan, Hylke de Vries, Sophie Bastin, Ségolène Berthou, Erwan Brisson, Marie-Estelle Demory, Hendrik Feldmann, Klaus Goergen, et al.

► To cite this version:

Sebastian K. Müller, Cécile Caillaud, Steven Chan, Hylke de Vries, Sophie Bastin, et al.. Evaluation of Alpine-Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms. *Climate Dynamics*, 2023, 61, pp.939-957. <10.1007/s00382-022-06555-z>. <insu-03863724>

HAL Id: insu-03863724

<https://insu.hal.science/insu-03863724v1>

Submitted on 9 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 Evaluation of Alpine-Mediterranean
 2 Precipitation Events in
 3 convection-permitting Regional Climate
 4 Models using a Set of Tracking Algorithms

5 Sebastian K. Müller^{1*†}, Cécile Caillaud^{2†}, Steven
 6 Chan^{3†}, Hylke de Vries^{4†}, Sophie Bastin⁵, Segolene
 7 Berthou⁶, Erwan Brisson⁷, Marie-Estelle Demory⁸, Hendrik
 8 Feldmann⁹, Klaus Goergen¹⁰, Stergios Kartsios, Petter
 9 Lind¹², Klaus Keuler¹³, Emanuela Pichelli¹, Mario
 10 Raffa¹⁴, Merja H. Tölle¹⁵ and Kirsten Warrach-Sagi¹⁶

11 ^{1*}ESP, Earth System Physics, ICTP, The Abdus Salam
 12 International Centre for Theoretical Physics, Strada Costiera 11,
 13 Trieste, 34100, Friuli-Venezia-Giulia, Italy.

14 ²CNRM, Université de Toulouse, CNRS, Météo-France.

15 ³School of Engineering, University of Newcastle.

16 ⁴KNMI, Royal Netherlands Meteorological Institute.

17 ⁵IPSL, Institut Pierre Simon Laplace.

18 ⁶MOHC, Met Office Hadley Centre.

19 ⁷GUF, Goethe University Frankfurt.

20 ⁸ETHZ, ETH Zürich.

21 ⁹KIT, Karlsruhe Institute of Technology.

22 ¹⁰FZJ, Institute of Bio- and Geosciences (IBG-3, Agrosphere),
 23 Forschungszentrum Jülich.

24 ¹¹AUTH, Aristotle University of Thessaloniki.

25 ¹²SMHI, Swedish Meteorological and Hydrological Institute.

26 ¹³BTU, Brandenburg University of Technology.

27 ¹⁴CMCC, Euro-Mediterranean Center on Climate Change.

28 ¹⁵Universität Kassel, Center for Environmental Systems Research
 29 (CESR), Kassel Institute for Sustainability.

30 ¹⁶UHOH, University of Hohenheim.

Contributing authors: sebastiank.mueller@posteo.net;

†These authors contributed equally to this work.

Keywords: convection-permitting modelling, heavy precipitation events, tracking algorithms

We here inter-compare four different tracking algorithms by applying them onto the precipitation fields of an ensemble of convection-permitting regional climate models (cpRCMs) and on high-resolution observational datasets of precipitation. The domain covers the Alps and the northern Mediterranean and thus we here analyse heavy precipitation events, that are renowned for causing hydrological hazards. In this way, this study is both, an inter-comparison of tracking algorithms as well as an evaluation study of cpRCMs in the Lagrangian frame of reference.

The tracker inter-comparison is performed by comparison of two case studies as well as of climatologies of cpRCMs and observations. We find that that all of the trackers produce qualitatively equal results concerning characteristic track properties. This means that, despite of quantitative differences, equivalent scientific conclusions would be drawn. This result suggests that all trackers investigated are reliable analysis tools of atmospheric research.

With respect to the model ensemble evaluation, we find an encouraging performance of cpRCMs in comparison to radar-based observations. In particular prominent hotspots of heavy precipitation events are well-reproduced by the models. In general most characteristic properties of precipitation events have positive biases. Assuming the under-catchment of precipitation in observations in a domain of such complex orography, this result is to be expected. Only the mean area of tracks is underestimated, while their duration is overestimated. Mean precipitation rate is estimated well, while maximum precipitation rate is overestimated. Furthermore, geometrical and rain volume are overestimated. We find that models overestimate the occurrence of precipitation events over all mountain chains, whereas over plain terrain in summer precipitation events are seen underestimated. This suggests that, despite the convection-permitting resolution, thermally driven thunderstorms are either not triggered or their dynamics still under-resolved. Eventually we find that biases in the spatio-temporal properties of precipitation events appear reduced when evaluating cpRCMs against Doppler radar-based and rain gauge-adjusted observational datasets of comparable spatial resolution, strengthening their role in evaluation studies.

1 Introduction

Moist deep convection in the atmosphere (Stevens, 2005), manifesting in storms of all scales, is responsible for the most severe precipitation events. However, as convection is by nature scarce in space and time, it is challenging to describe its properties, being fluxes of heat, momentum and water, appropriately. It is thus common to estimate convective heavy or extreme precipitation by e.g. the 99th or 99.9th percentile of hourly precipitation rate, or, to apply thresholds on precipitation rate and the frequency of their exceeding is then representative of the frequency of extreme events (Ban et al., 2020; Pichelli et al., 2021). However, statistical analyses in the Eulerian frame of reference remain limited to the description of the time series of each grid cell separately, and no information is retained about the underlying events and their spatial structure. Instead, information about the convective events themselves can be yielded through the application of a tracking algorithm, here referred to simply as tracker. By this, precipitation events are identified and tracked in time, meaning the analysis is transferred from the Eulerian into the Lagrangian frame of reference. Through the use of trackers the precipitation events themselves and their properties are in focus, rather than conditions at specific locations. Many studies (Prein et al., 2017a; Crook et al., 2019; Purr et al., 2019; Guo et al., 2022) have shown the benefit of this idea.

Tracking algorithms were originally developed in order to evaluate precipitation events in numerical weather predictions (Davis et al., 2006a,b; Wernli et al., 2008; Johnson et al., 2013). By the use of a tracker modelled precipitation objects can be compared against observations regarding their spatial structure, intensity, propagation and location. Any model output or observational field that is typically associated with a precipitation event may serve as the tracker input field. Although typically precipitation rate itself is used, also indirect proxies, like outgoing longwave radiation (Morel and Senesi, 2002a; Chen et al., 2019), or mid-level vertical velocity or vorticity, are suitable. On even finer resolutions individual updrafts of convective systems can be analyzed, along with their merging and splitting dynamics (Moseley et al., 2013). Another important application of tracking algorithms is the detection and observation of tropical and extra-tropical cyclones (Neu et al., 2013). Furthermore also droughts are operationally monitored using trackers (Abatzoglou et al., 2017).

The dynamical downscaling approach allows for investigating the impact of climate change on local scales and to derive actionable information for a variety of sectors (Giorgi, 2019, 2020). During the last decades convection-permitting Regional Climate Models (cpRCMs) were established, solving the non-hydrostatic equations of the atmosphere on grids with horizontal grid spacing smaller than 4 km and allowing to turn off flawed parameterizations of deep convection (Prein et al., 2015). At first, year-long integrations of distinct regions were realized (Grell et al., 2000), then decade-long integrations (Rasmussen et al., 2011) and decade-long integrations of entire continents (Prein et al., 2017a), and recently robust ensembles of cpRCMs (Coppola et al., 2020;

112 [Kendon et al., 2021](#)) have been achieved. cpRCMs brought great advances and
 113 still need further exploration of their capabilities ([Lucas-Picher et al., 2021](#)):
 114 significant added value ([Rummukainen, 2016](#); [Ciarlo et al., 2020](#)) lies in the
 115 representation of precipitation, in particular regarding its diurnal cycle, inten-
 116 sity and extremes ([Ban et al., 2015](#); [Kendon et al., 2017](#); [Fumière et al., 2020](#);
 117 [Reder et al., 2022](#)) and over complex orography ([Reder et al., 2020](#); [Adinolfi](#)
 118 [et al., 2020](#)). However, due to a more realistic orography, improvements are also
 119 found concerning surface temperature ([Hohenegger et al., 2008](#)) and mesoscale
 120 wind systems ([Belušić et al., 2018](#)). Through the application of trackers on
 121 cpRCMs, the climate change signal of convective storms can be analyzed in
 122 great detail ([Prein et al., 2017a,b](#); [Purr et al., 2019](#)). Along with advances
 123 in model development, novel observational precipitation datasets, based on
 124 Doppler radar measurements, providing comparable spatial and temporal res-
 125 olution emerged and allow for a rigorous evaluation of cpRCMs. Still, their
 126 impact on the evaluation of cpRCMs must be considered carefully ([Prein and](#)
 127 [Gobiet, 2017d](#)).

128 This present study uses an ensemble of cpRCMs, developed by the
 129 CORDEX - Flagship Pilot Study on *Convective phenomena at high resolu-*
 130 *tion over Europe and the Mediterranean* (CORDEX-FPSCONV, [Coppola](#)
 131 [et al. \(2020\)](#)), conducted on a domain covering the Alps and the northern
 132 Mediterranean Sea. This region is renowned for its severe precipitation events
 133 ([Drobinski et al., 2014](#)) and for being a climate change hotspot ([Giorgi, 2006](#)).
 134 Several research groups set up trackers in order to evaluate the behaviour
 135 of cpRCMs in simulating precipitation events and storms. This study takes
 136 advantage of this opportunity and carries out an inter-comparison study on
 137 a set of four trackers. They will be applied both to the model ensemble’s
 138 *evaluation* runs, driven by ERA-Interim reanalysis data, and to a composite
 139 of high-resolution observational datasets. Thus the scientific objective of this
 140 study is two-fold:

- 141 1. Tracker-Inter-comparison: we inter-compare four different trackers in order
 142 to find out about their reliability: do different trackers yield the same
 143 scientific conclusions?
- 144 2. Model-Evaluation: we evaluate an ensemble of convection-permitting
 145 regional climate models against observations in the Lagrangian frame of
 146 reference by using the trackers, and focus on the following aspects:
 - 147 • how good are cpRCMs at simulating the basic properties of precipitation
 148 events (intensity, spatial and temporal scales, rain volume)?
 - 149 • how good are cpRCMs at simulating the spatial patterns and the annual
 150 cycle of basic properties of precipitation events?

151 The paper is organized as follows. In section 2 we introduce the model
 152 ensemble and observational datasets used, as well as the two historic events
 153 that serve as case studies. In section 3 we explain the workflow of the tracking
 154 algorithms and motivate the setup chosen in order to identify the precipitation
 155 events of interest. In section 4 we present our results concerning the tracker

156 inter-comparison and in section 5 we present results on the model ensemble
157 evaluation. Finally we summarize our findings and give conclusions in section
158 6.

159 **2 Model Ensemble, Observational Datasets and** 160 **Historical Events**

161 In this section we briefly describe the CORDEX-FPSCONV ensemble of
162 cpRCMs as well as the composite of datasets of precipitation measurements
163 that in this study serve the input fields for the tracker analyses. Moreover we
164 here introduce two historical heavy precipitation events, that we use as case
165 studies for the tracker inter-comparison.

166 **2.1 Ensemble of convection-permitting Regional Climate** 167 **Models**

168 The CORDEX-FPSCONV community produced a first-of-its-kind ensemble
169 of cpRCMs for the domain studied herein, that is covering the Alps and the
170 northern Mediterranean (Coppola et al., 2020). Its Eulerian evaluation of pre-
171 cipitation is found in (Ban et al., 2020) and (Pichelli et al., 2021), which we here
172 build upon and extend into the Lagrangian frame of reference. Importantly
173 both studies demonstrate how the cpRCMs reduce model biases in compari-
174 son to the driving RCMs. Therein may also be found detailed information on
175 the models. We here analyze the *evaluation* runs, whose boundary conditions
176 are derived from ERA-Interim reanalysis data, through intermediate driving
177 simulations at coarser resolution (RCM) (Ban et al., 2020). The ensemble con-
178 tains several members using the COSMO-CLMcom and WRF model, which
179 differ in their nesting strategy and physics parameterizations respectively. The
180 model ensemble is summarized in Table 1.

Institute	cpRCM	dx(cpRCM)[km]	Driving RCM	dx(RCM)[km]	RCM domain
AUTH	WRF381BJ (A)	3	WRF	15	EURO-CORDEX
FZJ	WRF381BB	3	WRF	15	EURO-CORDEX
IPSL	WRF381BE (A)	3	WRF	15	EURO-CORDEX
UHOH	WRF381BD	3	WRF	15	EURO-CORDEX
BTU	COSMO-CLM (B)	3	COSMO-CLM	12	EURO-CORDEX
CMCC	COSMO-CLM (B)	3	COSMO-CLM	12	EURO-CORDEX
GUF	COSMO-CLM (B)	3	COSMO-CLM	12	Med-CORDEX
JLU	COSMO-CLM (B)	3	ERAINT	-	-
KIT	COSMO-CLM (B)	3	COSMO-CLM (B1)	25	Europe
ETHZ	COSMO-pompa_5.0 (C)	2.2	COSMO-CLM	12	Europe
CNRM	CNRM-AROME41t1 (C)	2.5	CNRM-ALADIN62 (C1)	12	Med-CORDEX (spectral nudging)
HCLIM-Com	HCLIM38-AROME (D)	3	ALADIN62	12	Europe
KNMI	HCLIM38-AROME (D)	2.5	RACMO	12	Europe
ICTP	RegCM4 (E)	3	RegCM4 (A)	12	Europe
UKMO	UM (F)	2.2	ERAINT	-	-

Table 1 Summary of numerical models used in this study. "dx" denotes the grid spacing of the respective model. Model documentation references: (A) (Powers et al., 2017); (B) (Rockel et al., 2008; Baldauf et al., 2011); (B1) (Keuler et al., 2016); (C) (Caillaud et al., 2021); (C1) (Nabat et al., 2020); (D) (van Meijgaard et al., 2008; Van Meijgaard et al., 2012; Belušić et al., 2020); (E) (Coppola et al., 2021); (F) (Chan et al., 2020)

182 Prior to the tracker analysis we remapped each of the models from their
 183 native grid onto the analysis grid *ALP-3i*, using distance weighted average
 184 remapping. It is a "regular lat-lon grid", spanning in longitude from 1°E to
 185 17°E in 582 grid cells, and in latitude from 40°N to 50°N in 364 grid cells.
 186 This results in a grid spacing of 0.0275° in both latitude and longitude, which
 187 translates on average to about 3 km.

188 2.2 High-resolution observational datasets of 189 precipitation

190 We use a composite of four observational datasets of precipitation covering
 191 France, Germany, Switzerland and Italy respectively, over a common time
 192 period from 2001 to end of 2009. Their original spatial resolution is comparable
 193 to that of the convection-permitting models, with native grid spacings ranging
 194 from 1 to 3 km, and their temporal resolution is hourly. Thus the observational
 195 datasets can be neatly compared to hourly precipitation rates of the models.
 196 All of the datasets except of one (GRIPHO) are based upon Doppler radar
 197 measurements and adjusted with rain gauge measurements. The spatial and
 198 temporal resolution of these datasets are the highest currently available for
 199 the respective regions. Still, Doppler radar observations are known to system-
 200 atically underestimate precipitation amounts over mountainous terrain, e.g.
 201 through the *shielding* effect (Germann et al., 2022), and underestimate partic-
 202 ularly heavy precipitation rates (Schleiss et al., 2020). Further, also rain gauges
 203 *under-catch* orographic precipitation and are moreover affected by windy condi-
 204 tions (La Barbera et al., 2002). Furthermore, interpolation methods used
 205 to map station data onto regular grids induce an underestimation of high
 206 intensities (smoothing effect) and an overestimation of low intensities (moist
 207 extension into dry areas) (Isotta et al., 2014). A brief summary of the individ-
 208 ual datasets, including their specific spatial resolution and references is given
 209 in Table 2.

Abbreviation	Coverage	dx [km]	Reference
COMEPHORE	France	1	(Tabary et al., 2012; Fumière et al., 2020)
RADKLIM_RW	Germany	1	(Winterrath et al., 2018)
RdisaggH	Switzerland	2	(Wüest et al., 2010)
GRIPHO	Italy	3	(Fantini, 2019)

Table 2 Summary of observational datasets of hourly precipitation rate used in this study. dx [km] denotes the original grid spacing.

210 Prior to the tracker analysis each of the datasets was remapped onto the
 211 analysis grid *ALP-3i*, again using distance weighted average remapping. Then
 212 we merged them and use their arithmetic mean for regions along the borders

213 of the nations, where measurements overlap. In this way, both the observations
214 and the models, were mapped onto the same grid.

215 ***GRIPHO over Italy and posteriori masking***

216 We here inform about two shortcomings of our analysis and show how we deal
217 with them when interpreting the results.

218 Firstly, the observational dataset covering Italy, GRIPHO, is based on
219 quality-controlled rain gauge measurements solely. The station density is
220 greater in the north than in the south of Italy, and on average it is estimated
221 to about 1 per $9 \times 9 \text{ km}^2$. It is then remapped onto a convection-permitting
222 grid with a 3 km grid spacing. In comparison to that, the other datasets are
223 based upon Doppler radar measurements and rain gauges and their original
224 spatial resolution is even finer than that of the analysis grid. We here must
225 expect differences in the spatio-temporal characterization of the precipitation
226 field observed by GRIPHO with respect to the other datasets. Nonetheless,
227 GRIPHO is the most accurate observational dataset available for Italy and in
228 particular the representation of extreme events was found improved (Fantini,
229 2019), especially over Northern Italy where the station density is higher and
230 where the most extreme precipitation events occur. In terms of domain com-
231 plexity though, we note that Italy is surrounded by the Mediterranean Sea
232 and the Alps, intersected by the Apennine Mountains and further shows both
233 steep coastlines and a large plain area (Po Valley). Due to this high degree
234 of complexity, which translates into very complex and local interactions, the
235 precipitation events are renowned for being particularly severe and their mod-
236 elling particularly challenging (see e.g. Morgan, 1973; Buzzi and Alberoni,
237 1992; Medina and Houze Jr, 2003; Rotunno and Houze, 2007; Panziera et al.,
238 2015; Miglietta et al., 2016; Pichelli et al., 2017).

239 Secondly, the observations do not cover the entire domain simulated by the
240 models, in particular not the Mediterranean sea. We consider this by *posteriori*
241 applying a mask onto the tracker analyses of models, meaning that only tracks
242 whose centroid is located within the domain of the observations are considered.
243 This implies that in models events entering the observational domain and here
244 particularly those making landfall, are expected to be overestimated in their
245 scales, but little in their averaged properties. Note further that for the Swiss
246 dataset RDisaggH, there is no data available for the period up to June 2003,
247 which is also accounted for through masking.

248 We account for both of these two shortcomings by presenting the relative
249 biases of the model ensemble not only for the entire domain of observations,
250 but also separately without GRIPHO as well as for GRIPHO exclusively, which
251 we consider representative of the most extreme precipitation events within the
252 domain. By this we account for and understand both, the specific model biases
253 due to the complex Italian domain as well as specific biases associated with the
254 GRIPHO drawbacks. Further, by doing so the overestimation of landfalling
255 tracks can be estimated, because by excluding GRIPHO we also exclude the
256 greatest part of the coastline.

2.3 Historical Heavy Precipitation Events

In the following we introduce the two historical heavy precipitation events that share these characteristics: both occurred along the Mediterranean coastline, both regions affected show steep orographic features and both happened in autumn. Coincidentally they both occurred along the same degree of latitude and the one happened just a little more than one year after the other.

2.3.1 Gard, France in September 2002

The first case study is a heavy precipitation event that occurred in south-eastern France, in the Gard region, during the 8th and 9th of September 2002 (Delrieu et al., 2005; Chancibault et al., 2006). Lasting more than a day, the event was particularly remarkable due to its rain amounts greater than 200 mm within 24 h spread over an area of 5500 km². The maximum rain rates of 600–700 mm observed locally by rain gauges are among the highest daily records in the region. The propitious slow-evolving synoptic-scale situation combined an upper-level south-westerly diffluent flow over south-eastern France with a moist and warm low-level south-easterly flow. The rain event can be characterized by three phases (Delrieu et al., 2005): at first a Mesoscale Convective System (MCS) developed over the Gard plains, second a displacement of the MCS toward the Cévennes mountain ridge took place and third, the passage of a cold front with embedded convection swept the convective activity out of the region. This catastrophic event resulted in 24 fatalities and an economic damage estimated at 1.2 billion € (Sauvagnargues-Lesage, 2004).

2.3.2 Carrara, Italy in September 2003

A second case study we carry out by looking at a heavy precipitation event that happened in Carrara, Italy, in September 2003, and which caused severe flash flooding and landslides. Cortopassi and Daddi (2008) investigated how the extensive quarrying activities of the region destabilize the terrain and promote hydro-geological hazards. It may be described as a landfalling convective system. A trough extending from a well structured low pressure centered over Northern Europe, advected hot and humid air from the Mediterranean sea and provided large-scale lifting. At the steep orography of the Apuan Alps the convective instability was triggered and the propagation of the storms was blocked. As a consequence, the region was affected by torrential rain, accumulating up to about 200 mm within a period of only 2 hours. The event claimed two fatalities and caused major damage to the local infrastructure.

3 Trackers

We here describe the basic functionality of the four tracking algorithms investigated in this study, which are referred to as MODE, OSIRIS, DYMECS and celltrack. Their functionalities are summarized in Table 5 and we provide a detailed description of each tracker in the appendix (Section 8). The trackers are completely independent developments and are here applied with setups that are as similar as possible, in order to compare the same precipitation objects. In Section 4, we compare the trackers individually against each other, while in Section 5, we evaluate the model ensemble by the mean of all four trackers, which we refer to as the "tracker ensemble mean".

The 1-hour accumulated precipitation fields (from observations or models) are used as input for all trackers. The principle operations of all trackers investigated include a first step of masking through a specified threshold, followed by a step of clustering in space to form objects and then tracking of those in time to form tracks. Prior to that the input field is smoothed in space. The treatment of cell merging and splitting is done by the allocation of *metatracks*, which can be understood as the smaller branches of merged or split tracks. It is a functionality that is not available in all trackers.

We designed the tracker setup such that it is able to identify precipitation events, that cause high impact weather situations like flash floods. For this reason we chose the (relatively large) precipitation threshold of 5 mm h^{-1} . On the other hand we want to investigate the small-scale isolated thunderstorms that cpRCMs are capable of resolving, and to this end, we chose a (relatively small) minimum space-time volume threshold of 100 cells. The input field is smoothed across 3×3 grid cells prior to the analysis.

The common tracker setup is summarized in Table 3:

spatial smoothing	3×3 grid cells
precipitation threshold	5 mm h ⁻¹
minimum space-time volume threshold	100 grid cells

Table 3 Summary of the tracker setup.

317

318

The characteristic track properties we are investigating in this study are defined in Table 4, with pr describing the precipitation field of a track:

Track Property	Definition
N_T	the number of tracks T identified
OF [time ⁻¹]	the <i>Occurrence Frequency</i> , defined as the number of tracks identified per unit time
OFD [time ⁻¹ area ⁻¹]	the <i>Occurrence Frequency Density</i> , defined as the number of tracks identified per unit time and unit area
\overline{pr}^c [mm h ⁻¹]	the <i>Mean Precipitation Rate</i> of a track, with c being the mean over all grid cells associated with a track
$\max^c(pr)$ [mm h ⁻¹]	the <i>Maximum Precipitation Rate</i> of a track, with \max^c being the maximum over all grid cells associated with a track
D [h]	the <i>Duration</i> of a track. (A track occurring only for a single time step will be attributed with 1 hour of duration.)
P_T [m ³]	the <i>Rain Volume</i> of a track, given by the integration of its precipitation field
\overline{A}^D [km ²]	the <i>Mean Area</i> of a track, averaged over its <i>Duration</i> , D
Vol [km ² h]	the <i>Space-Time Volume</i> of a track, given by integrating the area of all grid cells, in space and time

Table 4 Definitions of characteristic track properties.

319

Institute Tracker	space/time smoothing	metatrack at splitting/merging	statistics	boundary treatment	reference
ICTP MTD	yes/yes	no/no	original	smoothed	(Clark et al., 2014)
CNRM OSIRIS	yes/no	yes/yes off/off	smoothed	no	(Morel and Senesi, 2002a)
UKMO DYMECS	yes/no	yes/yes on/on	original	smoothed	(Stein et al., 2014)
KNMI celltrack	no/no	yes/yes off/off	smoothed	no	(Moseley et al., 2013) (Lochbihler et al., 2017)

Table 5 Summary of Trackers investigated in this study: "Institute" denotes the group executing the analysis using "Tracker", representing the abbreviation of the respective tracker. "space/time smoothing" denotes whether the tracker has an option for smoothing the input field prior to the analysis in space or time. "metatrack at splitting/merging" denotes whether the tracker assigns separate *metatracks* when tracks merge or split, with the second line indicating whether this functionality is switch on or off. Column "statistics" indicates whether the statistical properties of the tracks are derived from the original or from the smoothed input field. In "reference" the original description of the tracking algorithm is found.

320 4 Results on Tracker Inter-Comparison

321 In this section we inter-compare the four trackers in two steps: first, we compare
322 their performance at analyzing the two historic events, Gard 2002 and Carrara
323 2003 (subsection 4.1), and second, we compare their climatological properties,
324 derived from the tracker-analyses of the entire 9-year periods of observations
325 and model ensemble (subsection 4.2).

326 4.1 Tracker Inter-Comparison using Case Studies Gard 327 2002 and Carrara 2003

328 We apply the four trackers on the observational dataset and investigate only
329 the region and time periods of the respective historic events. In Figure 1, we
330 show for both events the accumulated total precipitation along with the loca-
331 tion of tracks and their respective rain volume. Note that we do not show
332 the full path of propagation for the tracks, as for the stationary precipitation
333 systems investigated here, the paths of propagation appear erratic and the
334 information added does not bring relevant insight. In general we recommend
335 that propagation features of multi-celled convective systems (e.g. distance trav-
336 elled or propagation velocity) must be interpreted with caution, as the correct
337 identification of their center is challenging. In Table 6, we summarize the
338 properties of all tracks associated with the two historic events.

339 We first note, that all of our trackers do identify both historic events and
340 attribute several tracks to them. OSIRIS and particularly DYMECS identify
341 more tracks than MTD and celltrack, which can be explained for DYMECS
342 by the allocation of *metatracks* in case of track merging or splitting. We find
343 that the number of tracks identified reflects in the sum of duration and sum of
344 mean track area. For each event, one main track is responsible for the major

Tracker	Gard 2002 / Carrara 2003						
	N_T [-]	$\max^T \overline{pr}^c$ [mm h ⁻¹]	$\max^T c(\text{pr})$ [mm h ⁻¹]	$\sum^T P_T$ [m ³ e6]	$\sum^T \overline{A}^D$ [km ² e3]	$\sum^T D$ [h]	$\sum^T \text{Vol}$ [km ² h e3]
MTD	3/4	11.0/9.0	96.0/57.6	5046/210	12/4	44/19	402/20
OSIRIS	10/7	8.8/6.9	85.0/52.0	3396/172	17/7	65/27	283/22
DYMECS	13/6	8.8/7.7	96.0/57.6	4476/240	25/7	141/20	335/24
celltrack	3/4	9.7/7.9	88.9/52.7	4587/190	7/5	45/19	415/22

Table 6 Averaged and integrated track properties, as defined in Table 4, associated with historic events Gard 2002 and Carrara 2003. $\max x^T$ and $\sum x^T$ denote the maximum value and summation of property x over all tracks T.

345 part of the rain volume and all the trackers agree well on these most severe
346 tracks. Focusing on intensities, lower intensities with OSIRIS can be explained
347 by the calculation of the diagnostics on the smoothed precipitation field.

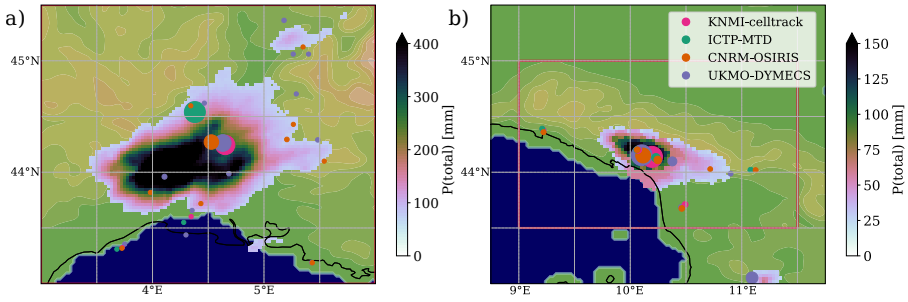


Fig. 1 Historical events Gard 2002 (left panel) and Carrara 2003 (right panel) using the observations. Shading illustrates accumulated total precipitation, $P(\text{total})$ [mm]. Filled circles indicate the location of the centroid of a track and their radius is proportional to their respective rain volume. Filled contours indicate the elevation of the model terrain in intervals of 250 m.

348 Still overall and as listed in Table 7, we find that all trackers agree on the following relations:

Compared to the Carrara 2003 case, the Gard 2002 case is greater in all scales: it is longer-lived, larger in area, and consequently also in Space-Time volume.

Compared to the Carrara 2003 case, the Gard 2002 case is more intense: i.e., its mean and maximum precipitation rates are greater.

As a consequence of both greater scales and intensity, the integrated rain volume of the Gard 2002 case is much greater than that of the Carrara 2003 case.

Table 7 Qualitative attribution of track properties to the two historical events, Carrara 2003 and Gard 2002.

349
350 Thus, all trackers describe the two events with equivalent track properties and
351 moreover the properties attributed and the relations found agree well with
352 the description of the events in literature: the smaller and less intense event,
353 Carrara 2003, is attributed with smaller spatial scales and less intensity than
354 the larger and more intense event, Gard 2002. Based upon these results, the
355 choice of the trackers seems irrelevant, meaning that from any of the trackers'
356 analyses, equivalent scientific conclusions would be derived. In other words,
357 with the differences being only of quantitative nature, the scientific conclusions
358 are found to be independent of the choice of the tracker.

Tracker	Model Ensemble / Observations						
	N_T [year ⁻¹]	\overline{pr}^T [mm h ⁻¹]	$\overline{\max^c(pr)}^T$ [mm h ⁻¹]	\overline{P}_T^T [m ³ e6]	\overline{A}^D^T [km ² e3]	\overline{D}^T [h]	\overline{Vol}^T [km ² h e3]
MTD	2742/2536	8.3/7.8	21.3/17.3	74/56	1.2/1.6	4.4/3.1	8.3/6.8
OSIRIS	3347/2911	7.0/6.7	16.1/13.9	43/40	1.1/1.5	4.1/3.0	5.6/5.6
DYMECS	5095/3618	7.9/7.4	19.6/16.1	43/36	1.3/1.6	3.6/2.7	4.7/4.4
celltrack	2953/2550	7.7/7.3	17.2/14.7	72/54	1.1/1.6	4.5/3.1	8.8/7.1

Table 8 Climatology of track properties derived from both the model ensemble and the observations. In Table 4, definition of the respective properties are given.

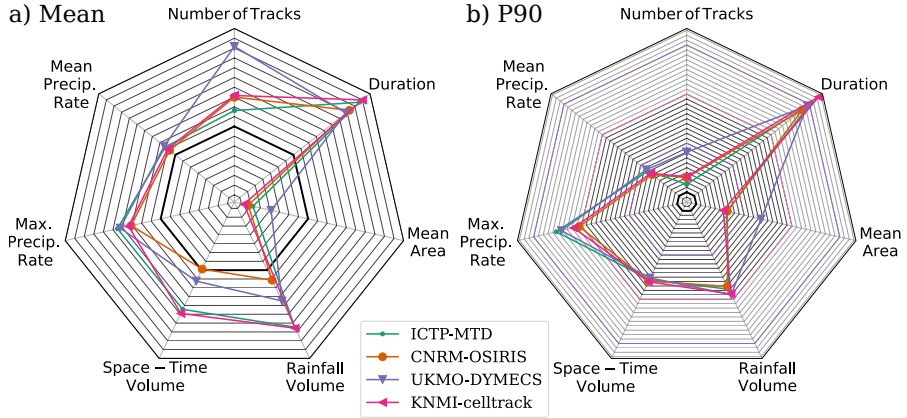


Fig. 2 Relative bias in characteristic of a) the mean and b) the 90th percentile of track properties of each tracker with respect to the observations $\frac{ModEns - Obs}{Obs}$ [%]. Black solid lines are increments of +5%, with the thick black line representing the tracker ensemble mean of the observations, i.e. 0%.

4.2 Tracker Inter-Comparison using the Climatologies of Model Ensemble and Observations

We continue the tracker inter-comparison over climatological scale through the ensemble of cPRCM simulations and the observations for the common time period 2001-2009. Table 8 shows the climatological means of characteristic track properties of each tracker and, in Figure 2, we show the relative biases of each tracker for the mean and 90th percentile of track properties with respect to the observations.

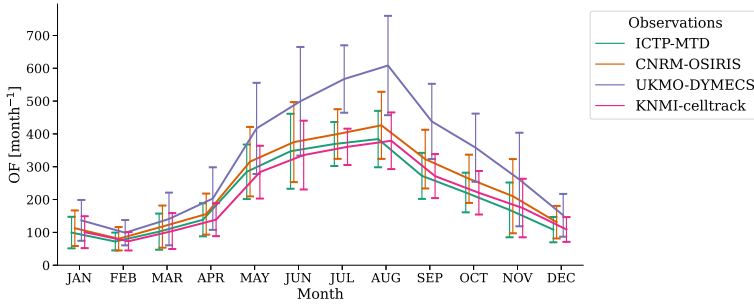


Fig. 3 Annual Cycle of track occurrence frequency, OF [month^{-1}], for all trackers analyzing the observations. Error bars indicate inter-annual variability by the temporal standard deviation.

367 With respect to characteristic track properties, for both their mean and
 368 90th percentile, all trackers identify the following qualitative biases, shown in
 Table 9, when comparing the model ensemble against observations:

underestimation	overestimation
mean area	track occurrence maximum precipitation rate mean precipitation rate space-time volume rain volume

Table 9 Qualitative biases of characteristic track properties that are consistent across all trackers.

369 This means that all trackers derive for all properties of precipitation events
 370 the same qualitative biases, but these can differ in magnitude.

371 Looking only at tracker results of the observation-based climatology (Table
 372 8), the characteristic track properties are overall similar between trackers. Par-
 373 ticularly mean area, mean and maximum precipitation rates and duration are
 374 estimated similarly by the trackers. Some pronounced quantitative differences
 375 can still be found and attributed to tracker characteristics: firstly, due to the
 376 allocation of *metatracks* at track merging and splitting, the number of tracks is
 377 highest with DYMECS, whereas the space-time volume is smallest. Secondly,
 378 because of the calculation of the characteristics from the smoothed field, the
 379 intensities are lowest with OSIRIS.

381 Figure 3 shows the annual cycle of track occurrence frequency identified
 382 in the observations. For all four trackers, we find that the distribution is uni-
 383 modal, with a peak in August and a minimum in February. Similarly to what
 384 we found for the two single historic events in Section 4.1, a tracker that allo-
 385 cates *metatracks* at splitting and merging (DYMECS) identifies more tracks
 386 in total than those that do not (MTD, celltrack and OSIRIS). It also shows
 387 greater inter-annual variability. Again, despite the differences of quantitative

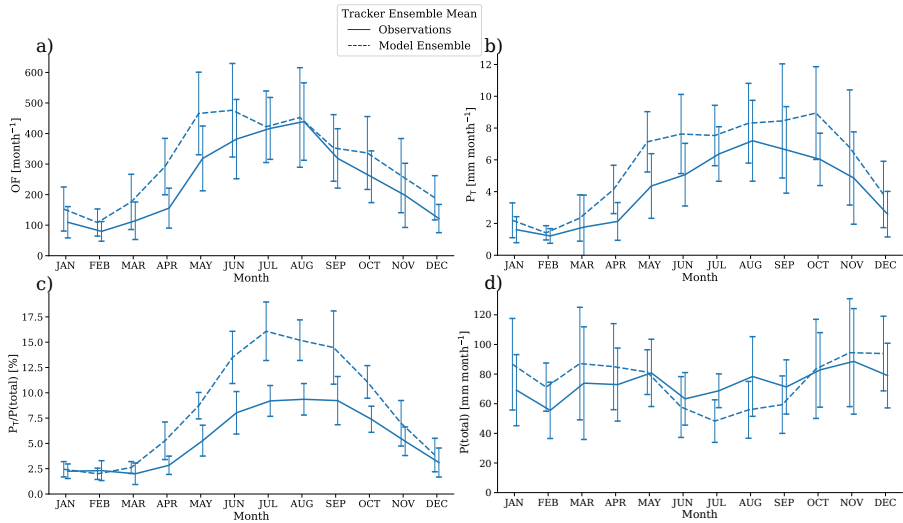


Fig. 4 Annual cycles for the tracker ensemble mean of both the model ensemble (dashed line) and observations (solid line), with panel a) showing track occurrence frequency, OF [month⁻¹], panel b) accumulated precipitation of tracks, P(tracks) [mm month⁻¹], panel c) heavy precipitation fraction, P(tracks)/P(total) [%], and panel d) accumulated total precipitation, P(total) [mm month⁻¹]. Error bars for the model ensemble indicate the temporal standard deviation of the model ensemble mean across years, and likewise for the observations error bars display inter-annual variability by the temporal standard deviation across years.

388 nature among trackers, the scientific conclusions when comparing climatolo-
 389 gies of model ensemble and observations are mainly independent of the choice
 390 of the tracker.

5 Results on Model Evaluation

In this section, we evaluate the representation of precipitation events in the cpRCM ensemble. To this end, we use the mean of the tracker analyses (tracker ensemble mean) and compare the entire 9-years periods of the model ensemble against the composite of observations.

In Figure 4a) we show the annual cycle of track occurrence for the tracker ensemble mean, of both the model ensemble and observations. We find that the number of events occurring in spring and fall is overestimated, whereas for July and August the occurrence frequency of tracks in the models is close to that of the observations. The annual cycle of the model ensemble shows two peaks, one in June and one in August, whereas the annual cycle of the observations is unimodal. With respect to the estimate of inter-annual variability, given by the temporal variance across years, we find that the model ensemble exceeds the observations.

Figure 4b) and d) show accumulated precipitation of tracks (P_T) and total accumulated precipitation ($P(\text{total})$), whereas panel c) shows their fraction. In this domain and time period there is no pronounced annual cycle found in $P(\text{total})$. If anything it is rather the models that show a dry summer w.r.t. a wet winter. In other words, the model ensemble overestimates $P(\text{total})$ in winter and underestimates it in summer. In contrast to that, P_T shows a strong seasonality, with the model ensemble showing a broad peak from May to November and the observations peaking from July to October. Here we find an overestimation of P_T throughout the whole year. Consequently their fraction, $P_T/P(\text{total})$, also shows strong seasonality, again with a peak in summer, and once more we identify a substantial overestimation by the model ensemble. Moreover we from this see that our setup chosen attributes only about 5% (observations) and 10% (model ensemble) of the annual precipitation amount to tracks. This overestimation was already of intense precipitation was already found in (Berthou et al., 2018; Meredith et al., 2020). Eventually we also see that this tracker setup serves well in identifying heavy precipitation events, as the fraction of precipitation amount identified is relatively low.

Figure 5 shows the track occurrence frequency density of the tracker ensemble mean for the model ensemble and observations, as well as their difference. Panel d) shows the normalized difference, i.e. as if there were as many tracks in model ensemble as in observations, and by this emphasizes qualitative differences. Moreover in panel e) we show the difference in track occurrence for different seasons and different elevations. It is evident from observations (Fig. 5b) that track occurrence is strongly correlated to the topography, meaning that the orographic forcing plays a major role for precipitation events to occur; this is well captured by the models as well (Fig. 5a). Prominent *hotspots* of heavy precipitation are the Julian Alps (North-East Italy), the Western Alps (especially the Italian side and the southern Maritime Alps between Italy and France) and the Massif Central (South France). Also Corsica and the Apennines can be identified as hotspots. However, also dry spots, located in the

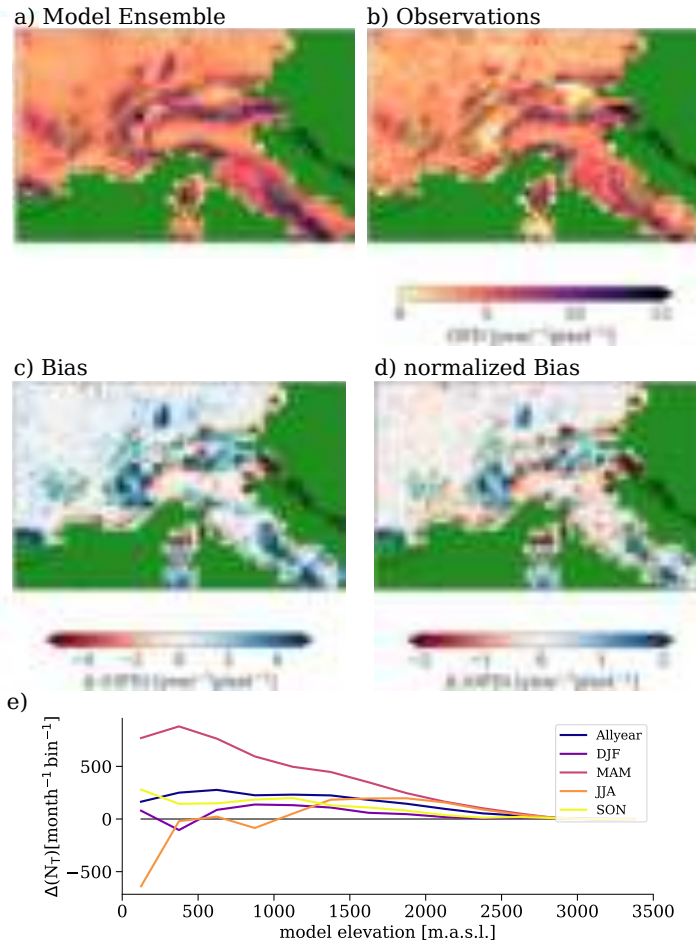


Fig. 5 Panel a) shows the track occurrence frequency density, OFD [$\text{month}^{-1} \text{pixel}^{-1}$], of the tracker ensemble mean for the model ensemble and panel b) shows the same for the observations. Panel c) shows their difference and panel d) show the difference, but with model ensemble and observations being normalized by their total number of tracks, respectively. A pixel is here defined as the reference area of $0.36^\circ \times 0.36^\circ$. The green iso-line shows the model elevation at 1000 m.a.s.l.. Panel e) shows the model-observation difference in track occurrence by model elevation.

435 interior of the Alps, like in Tyrol in Austria, or in the Western Alps are promi-
 436 nent in observations and re-produced well by the cpRCMs. Track occurrence
 437 appears overestimated over orography, particularly in the Maritime Alps, the
 438 Tyrolean Alps, the Apennines, the Black Forest and to a lesser extent, in the
 439 southern Massif Central. In contrast to this, in plains ahead of mountains,
 440 like in northern Italy, occurrence frequency is underestimated. We have seen
 441 already in the annual cycle of occurrence frequency (Figure 4 a)), that cpRCMs
 442 most strongly overestimate track occurrence in spring (MAM) and estimates
 443 OF well in summer (JJA). We now in panel e) of Figure 5 identify clearly that

444 cpRCMs in all seasons overestimate tracks above 1000 m.a.s.l., whereas in sum-
 445 mertime, below 1000 m.a.s.l. OF is underestimated. This behaviour may be
 446 explained through the following considerations: numerical models easily trig-
 447 ger convection through orographic lifting. However, they appear to struggle
 448 to trigger thunderstorms or to resolve complex thunderstorm dynamics over
 449 plain terrain in summer, even at convection-permitting resolution (see also
 450 [Craig et al., 2012](#); [Heim et al., 2020](#); [Prein et al., 2021](#)). On the other hand,
 451 observational datasets under-catch rainfall amounts in mountainous regions.
 452 Therefore, model performance over orography is expected to be better than
 453 it seems. This finding is in line with [Lundquist et al. \(2019\)](#), who propose
 454 that well-tuned cpRCMs may outperform observational datasets over complex
 455 mountain terrain, in terms of total precipitation amounts.

456 The statistics of characteristic track properties of the tracker ensemble
 457 mean, for the climatologies of both model ensemble and observations, are listed
 458 in [Table 8](#). Relative model biases of mean track properties are illustrated in
 panels a) and b) of [Figure 6](#) and are summarized in [Table 10](#).

Property	Relative Bias (all over)	Relative Bias (w/o GRIPHO- Italy)	Relative Bias (only GRIPHO-Italy)
N_T	+22 %	+43 %	-11 %
\overline{pr}^c	+6 %	+3 %	+12 %
$\max^c(\text{pr})$	+20 %	+13 %	+36 %
D [h]	+37 %	+29 %	+56 %
\overline{A}^D	-27 %	-24 %	-34 %
Vol	+8 %	+0 %	+29 %
P_T	+17 %	+8 %	+43 %

Table 10 Relative biases [%] of characteristic track properties, as defined in [Table 4](#), using the tracker ensemble mean, for the whole domain, as well as without GRIPHO and exclusively for GRIPHO.

459 We find that the biases of mean track properties are mostly positive (see
 460 [Figure 6b](#)) and [Table 10](#)). Biases of the 90th percentile of track properties are
 461 still much larger (see [Figure 6d](#)), suggesting that extreme events are strongly
 462 overestimated regarding their scales and intensity. Considering the complex
 463 orography of the domain investigated ([Rotunno and Houze, 2007](#)), in combi-
 464 nation with the known issue of under-catchment of orographic precipitation
 465 in observations, the positive biases were to be expected. Despite this, we find
 466 it important to note that mean precipitation rate of tracks is well-estimated
 467 (+6% all over the domain, +3% w/o GRIPHO-Italy). The absolute number of
 468 events, maximum precipitation rate, track duration and rainfall volume are
 469 considerably overestimated (>17%). Only the mean area of tracks is under-
 470 estimated. Model biases with respect to GRIPHO-Italy differ from those of
 471 the other datasets and regions qualitatively only in terms of number of tracks,
 472 showing here an underestimation. It is worth to note that the model spread
 473 is particularly large in terms of number of tracks ([Figure 6a](#)). For all other
 474

475 properties we find smaller biases over regions with radar-based datasets, i.e.
476 France, Germany and Switzerland (w/o GRIPHO-Italy) than over Italy (only
477 GRIPHO-Italy). It is particularly the spatio-temporal properties (duration,
478 mean area, space-time volume) and maximum precipitation rate, that show
479 the greatest differences. We assume that the bias reduction in spatio-temporal
480 properties, particularly in mean area, is associated with improvements that
481 the spatially continuous radar measurements ensure. Larger model biases over
482 Italy might be also attributed to some higher degree of complexity not well
483 captured by some or all cpRCMs within the ensemble. Certainly the optimal
484 spatial-temporal representation of precipitation events in radar-based datasets
485 constitutes an advantage in the evaluation of models in a context of Lagrangian
486 analysis. Our findings confirm the key role of observational datasets with com-
487 parable spatial resolution in evaluation studies of RCMs (Torma et al., 2015;
488 Prein and Gobiet, 2017d).

489 The probability density functions in Figure 6 give more detailed insight
490 into the models' behaviour. Looking at track duration, we find that cpRCMs
491 simulate precipitation events of temporal scales longer than 50 hours, that are
492 not found in observations. In turn simulated precipitation events are generally
493 too small regarding their area, whereas we see only a minor overestimation of
494 the distribution's tail. As a combination of the biases of duration and area,
495 the bias of geometrical volume is still positive. It is mostly the overestimation
496 of track duration that causes to the positive biases in geometrical volume and
497 also biases in rain volume are mostly found in the tails of the distribution. In
498 other words, cpRCMs are found to simulate precipitation events of large scales
499 that are not seen in observations. This finding is also reflected in the high
500 biases of the 90th percentile of track properties shown in Figure 6 c) and d).

501 We in Figure 8 (in the Supplementary Material Section 9) provide the relative
502 biases of mean properties for each model individually and we here would
503 like to address the 2 cpRCM families WRF and CCLM. While the WRF mod-
504 els differ in their physics parameterizations, the CCLM models differ only in
505 their nesting strategy. Among the WRF models the variability in mean biases
506 is considerably large, with e.g. the IPSL-WRF and UHOH being particularly
507 different. In turn the biases among the CCLM model family look much more
508 similar. We from this conclude that physics parameterizations have a large
509 effect on model behaviour and thus can generate greater ensemble variability
510 than differing nesting strategies.

511 The spatial mapping of model biases in Figure 7 allows us to further under-
512 stand impacts of the technical shortcoming mentioned in Section 2.2. Note that
513 we in Figure 9 (in the Supplementary Material Section 9) show the mapping of
514 the respective properties for observations and model ensemble. The *posteriori*
515 *masking* of tracks means that landfalling tracks are overestimated in their spa-
516 tial and temporal scales. In fact we find that along the coasts rain volume and
517 geometrical volume are overestimated, whereas the other averaged variables
518 appear unaffected. Over Italy we again find the pronounced underestimation of
519 track mean area and overestimation of track duration. We here can speculate

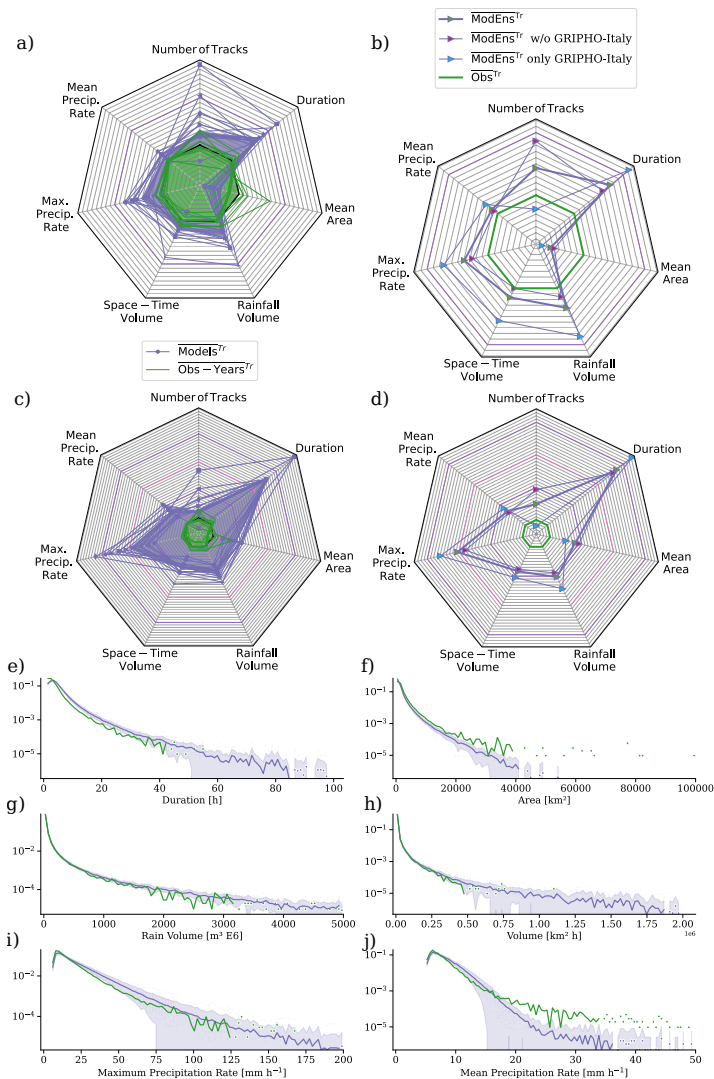


Fig. 6 Panels a) and c): The purple shaded area illustrates the relative bias of a) the tracker ensemble mean of the model ensemble mean and c) the 90th percentile, with respect to the observations: $\frac{ModEns^{Tr} - Obs^{Tr}}{Obs^{Tr}} [\%]$, while purple lines indicate the individual models and green lines individual years of the observations. Panels b) and d) shows also $\frac{ModEns^{Tr} - Obs^{Tr}}{Obs^{Tr}} [\%]$ for the mean and 90th percentile of characteristic properties, with the Italian dataset GRIPHO excluded as well as for GRIPHO only. Black solid lines are increments of +5%, with the thick black line denoting 0%. Panels e) to j): probability density functions of Duration [h], Area [km²], Rain Volume [m³ E6], Space-Time Volume [km² h], Maximum Precipitation Rate [mm h⁻¹] and Mean Precipitation Rate [mm h⁻¹].

520 that GRIPHO's interpolation method smoothens the field strongly, enlarging

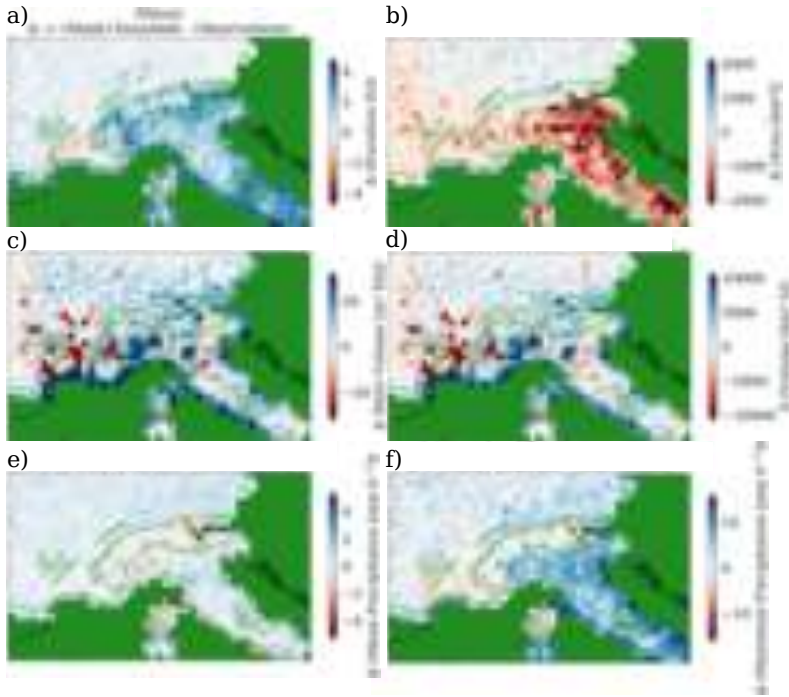


Fig. 7 The spatial biases of the tracker ensemble mean of the model ensemble w.r.t. observations. Panels a) to f): duration [h], (mean) area [km²], rain volume [m³ E6], (geometrical) volume [km² h], mean and maximum precipitation (rate) [mm h⁻¹]. Again a pixel is here defined as the reference area of $0.36^\circ \times 0.36^\circ$ and the green iso-line shows the model elevation at 1000 m.a.s.l..

521 the spatial extent of events. Also positive biases in mean and maximum pre-
 522 cipitation are pronounced over Italy, but are not dramatically different from
 523 the other sub-regions in the domain.

6 Summary and Conclusions

The present study has a two-fold scientific objective: on the one hand we provide an inter-comparison of tracking algorithms and on the other hand we present an evaluation of an ensemble of convection-permitting regional climate models in terms of Lagrangian precipitation events. We here summarize our findings and give conclusions.

With respect to the tracker inter-comparison (see Section 4) we were able to show through both, the comparison of two historic events and the comparison of climatologies of model ensemble and observations, that all trackers investigated produce equal relations of characteristic track properties and model biases (see Tables 7 and 9 and Figures 1, 2 and 3). Thus all trackers produce qualitatively equal results. In other words, differences among the trackers were found to be only of quantitative nature, which could be addressed to certain specifications of the algorithms. From this we infer that from each tracker analysis the equivalent scientific conclusion would be derived. This result suggests that all trackers investigated are reliable analysis tools of atmospheric research.

The choice of tracker depends here much on whether *metatracks*, allocated when tracks are splitting or merging, are of interest. Further code availability, portability and user support also play a major role.

We find that the setup chosen here, given through smoothing, precipitation rate and volume threshold (Table 3), identifies an abundance of precipitation events all over the domain, of which only a fraction would be considered an extreme event. In our analysis of two historical events, we see that those are represented by several tracks. We recommend to consider that a tracker would identify fewer or only a single track, if thresholds on precipitation rate and minimum volume were raised and smoothing strengthened. The choice of setup depends upon the user-specific application. Certainly though, the most intense events are retained. In turn, reducing thresholds and weakening smoothing will result in a setup that identifies more and greater tracks and a greater fraction of precipitation will be attributed to the events.

With respect to model evaluation (see Section 5) we summarize the following findings. Looking into the spatial representation of precipitation events (see Figure 5), we found that cpRCMs perform well in reproducing *hotspots* of heavy precipitation, which are generally associated with orographic features. At the same time though, cpRCMs appear to overestimate the occurrence of precipitation events over orography. However, the under-catchment of orographic precipitation in radar-based and rain gauge observations (Creutin et al., 1997; La Barbera et al., 2002; Prein and Gobiet, 2017d; Germann et al., 2022) suggests that cpRCMs perform better than it seems. The idea of cpRCMs outperforming observations in complex terrain, particularly in terms of total precipitation amounts, is strongly supported in Lundquist et al. (2019). In contrast to this, we found the occurrence of precipitation events underestimated over plain terrain and ahead of orographic features, particularly in summer. The same model behaviour was found by Prein et al. (2017a) for North America, where the occurrence frequency of MCSs was underestimated in the central

569 plains but overestimated over the Appalachians. We here assume that, despite
570 of the convection-permitting resolution, complex thunderstorms (e.g. super-
571 cells or squall lines) in plain terrain are either not triggered or their dynamics
572 still under-resolved (see also [Bryan and Morrison \(2012\)](#); [Pichelli et al. \(2017\)](#);
573 [Prein et al. \(2021\)](#)). Moreover, the correct prescription of sea surface temper-
574 atures is crucial for the intensity and evolution of characteristic landfalling
575 Mediterranean heavy precipitation events ([Lebeaupin et al., 2006](#)). Looking
576 into the seasonal representation of precipitation events, we find that cpRCMs
577 overestimate the occurrence of tracks and associated precipitation amounts
578 particularly in late spring (AMJ), and also in fall. In late summer months
579 (JAS) the domain-wide occurrence appears estimated well, as the overestima-
580 tion in regions over 1000 m.a.s.l. is compensated by the underestimation in
581 regions below 1000 m.a.s.l.

582 In terms of characteristic properties of precipitation events we found the
583 following biases (listed in [Table 10](#) and illustrated in [Figure 6](#)) and give refer-
584 ence to tracker studies using convection-permitting models. The occurrence
585 frequency of events is overestimated with respect to radar-based observations
586 (in line with [Clark et al. \(2014\)](#); [Prein et al. \(2017a\)](#); [Caillaud et al. \(2021\)](#))
587 and under-estimated over Italy, although the models spread is large around
588 this property. The mean area of tracks is underestimated (in line with ([Crook
589 et al., 2019](#)), but in contrast to [Caillaud et al. \(2021\)](#)), while their duration is
590 overestimated (in line with [Crook et al. \(2019\)](#); [Purr et al. \(2019\)](#)). Still, we
591 have identified that both of these biases are particularly pronounced over Italy.
592 In turn, the cpRCMs agree much better with the radar-based observational
593 datasets in terms of track area and duration. The tracks' space-time volume,
594 that is the combination of area and duration, as well as the rain volume,
595 are overestimated. However, we here find considerable impact by the differing
596 representation of landfalling tracks in models and observations, and exclud-
597 ing a major part of the coastline (the Italian sub-region) reduces the biases
598 much. Mean precipitation rates show only small positive biases, with cpRCMs
599 aligning again even better with radar-based observations. Maximum precipi-
600 tation rate is overestimated in models and here again biases are much reduced
601 when using radar-based observations as benchmark (in line with ([Davis et al.,
602 2006b](#); [Prein et al., 2017a](#); [Crook et al., 2019](#); [Caillaud et al., 2021](#)). Still
603 we find that cpRCMs simulate precipitation events of scales and intensities
604 that are not seen in observations, which means an overestimation of extreme
605 event properties. Overall, the results on cpRCM evaluation are encouraging.
606 On the one hand the (mostly) positive biases we find are to be expected, when
607 assuming underestimated precipitation amounts in observations in a region of
608 such complex orography. On the other hand we find that biases of the spatio-
609 temporal properties of precipitation events in cpRCMs appear much reduced
610 when using high-resolution observational datasets, based upon Doppler radar
611 measurements.

7 Declarations

Acknowledgments. The authors gratefully acknowledge the WCRP-CORDEX-FPS on Convective phenomena at high resolution over Europe and the Mediterranean [FPSCONV-ALP-3] and the research data exchange infrastructure and services provided by the Jülich Supercomputing Centre, Germany, as part of the Helmholtz Data Federation initiative. This work has also been supported in part by the Horizon 2020 EUCP (European Climate Prediction System, <https://www.eucp-project.eu>) project. This EUCP project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 776613. This research is supported by the United Kingdom Natural Environment Research Council (NERC) Changing Water Cycle programme (FUTURE-STORMS; grant: NE/R01079X/1) AUTH acknowledges the support of the Greek Research and Technology Network (GRNET) High Performance Computing (HPC) infrastructure for providing the computational resources of AUTH-simulations (under project ID pr003005) and the AUTH Scientific Computing Center for technical support For JLU simulations computational resources were made available by the German Climate Computing Center (DKRZ) through support from the Federal Ministry of Education and Research in Germany (BMBF) and funding stems from the German Research Foundation (DFG) through grant nr. 401857120. The authors from FZJ gratefully acknowledge the computing time granted by the JARA Vergabegremium and provided on the JARA Partition part of the supercomputer JURECA at Jülich Supercomputing Centre at Forschungszentrum Jülich. The authors thank Meteo-France for providing the COMEPHORE dataset, MeteoSwiss for the RdisaggH dataset. They thank the German Weather Service for providing the RADKLIM dataset (RADKLIM Version 2017.002: Reprocessed gauge-adjusted radar data, one-hour precipitation sums (RW) DOI:10.5676/DWD/RADKLIM_RW_V2017.002). Hylke de Vries wishes to thank Geert Lenderink and Kai Lochbihler for discussions on the use of celltrack. ICTP also acknowledges the CETEMPS, University of L'Aquila, for allowing access to the Italian database of precipitation which GRIPHO is based on.

Data Availability. The observational datasets are available upon request through the institutions listed above. Tracking algorithms OSIRIS and DYMECS are available upon request from the developers respectively. MTD is available at <https://dtcenter.org/community-code/model-evaluation-tools-met> and celltrack at <https://github.com/lochbika/celltrack>. The entire tracker analysis is publicly available at <https://doi.org/10.5281/zenodo.6949615>.

Competing Interests. The authors have no relevant financial or non-financial interests to disclose.

References

- 653
- 654 J. T. Abatzoglou, D. J. McEvoy, and K. T. Redmond. The West Wide
 655 Drought Tracker: Drought Monitoring at Fine Spatial Scales. *Bulletin of the*
 656 *American Meteorological Society*, 98(9):1815–1820, Sept. 2017. ISSN 0003-
 657 0007, 1520-0477. doi: 10.1175/BAMS-D-16-0193.1. URL [https://journals.](https://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0193.1)
 658 [ametsoc.org/doi/10.1175/BAMS-D-16-0193.1](https://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0193.1).
- 659 M. Adinolfi, M. Raffa, A. Reder, and P. Mercogliano. Evaluation and Expected
 660 Changes of Summer Precipitation at Convection Permitting Scale with
 661 COSMO-CLM over Alpine Space. *Atmosphere*, 12(1):54, Dec. 2020. ISSN
 662 2073-4433. doi: 10.3390/atmos12010054. URL [https://www.mdpi.com/](https://www.mdpi.com/2073-4433/12/1/54)
 663 [2073-4433/12/1/54](https://www.mdpi.com/2073-4433/12/1/54).
- 664 M. Baldauf, A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and
 665 T. Reinhardt. Operational convective-scale numerical weather prediction
 666 with the COSMO model: Description and sensitivities. *Monthly Weather*
 667 *Review*, 139(12):3887–3905, 2011.
- 668 N. Ban, J. Schmidli, and C. Schär. Heavy precipitation in a changing climate:
 669 Does short-term summer precipitation increase faster? *Geophysical Research*
 670 *Letters*, 42(4):1165–1172, 2015. Publisher: Wiley Online Library.
- 671 N. Ban, J. Rajczak, J. Schmidli, and C. Schär. Analysis of Alpine precipita-
 672 tion extremes using generalized extreme value theory in convection-resolving
 673 climate simulations. *Climate Dynamics*, 55(1):61–75, 2020. Publisher:
 674 Springer.
- 675 A. Belušić, M. T. Prtenjak, I. Güttler, N. Ban, D. Leutwyler, and
 676 C. Schär. Near-surface wind variability over the broader Adriatic
 677 region: insights from an ensemble of regional climate models. *Climate*
 678 *Dynamics*, 50(11-12):4455–4480, June 2018. ISSN 0930-7575, 1432-0894.
 679 doi: 10.1007/s00382-017-3885-5. URL [http://link.springer.com/10.1007/](http://link.springer.com/10.1007/s00382-017-3885-5)
 680 [s00382-017-3885-5](http://link.springer.com/10.1007/s00382-017-3885-5).
- 681 D. Belušić, H. de Vries, A. Dobler, O. Landgren, P. Lind, D. Lindstedt,
 682 R. A. Pedersen, J. C. Sánchez-Perrino, E. Toivonen, B. van Uft, and oth-
 683 ers. HCLIM38: a flexible regional climate model applicable for different
 684 climate zones from coarse to convection-permitting scales. *Geoscientific*
 685 *Model Development*, 13(3):1311–1333, 2020. Publisher: Copernicus GmbH.
- 686 S. Berthou, S. Mailler, P. Drobinski, T. Arsouze, S. Bastin, K. Béranger, and
 687 C. Lebeaupin Brossier. Lagged effects of the mistral wind on heavy precipi-
 688 tation through ocean-atmosphere coupling in the region of valencia (spain).
 689 *Climate Dynamics*, 51(3):969–983, 2018.

- 690 P. Brousseau, Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast
691 of convective activity from the AROME-France system. *Quarterly Journal*
692 *of the Royal Meteorological Society*, 142(699):2231–2243, 2016. Publisher:
693 Wiley Online Library.
- 694 G. H. Bryan and H. Morrison. Sensitivity of a Simulated Squall Line to
695 Horizontal Resolution and Parameterization of Microphysics. *Monthly*
696 *Weather Review*, 140(1):202–225, Jan. 2012. ISSN 0027-0644, 1520-0493.
697 doi: 10.1175/MWR-D-11-00046.1. URL [http://journals.ametsoc.org/doi/](http://journals.ametsoc.org/doi/10.1175/MWR-D-11-00046.1)
698 [10.1175/MWR-D-11-00046.1](http://journals.ametsoc.org/doi/10.1175/MWR-D-11-00046.1).
- 699 A. Buzzi and P. P. Alberoni. Analysis and numerical modelling of a frontal
700 passage associated with thunderstorm development over the Po valley and
701 the Adriatic sea. *Meteorology and Atmospheric Physics*, 48(1-4):205–224,
702 1992. ISSN 0177-7971, 1436-5065. doi: 10.1007/BF01029569. URL [http://](http://link.springer.com/10.1007/BF01029569)
703 link.springer.com/10.1007/BF01029569.
- 704 C. Caillaud, S. Somot, A. Alias, I. Bernard-Bouissières, Q. Fumière, O. Lau-
705 rantin, Y. Seity, and V. Ducrocq. Modelling Mediterranean heavy pre-
706 cipitation events at climate scale: an object-oriented evaluation of the
707 CNRM-AROME convection-permitting regional climate model. *Climate*
708 *Dynamics*, 56(5):1717–1752, 2021. Publisher: Springer.
- 709 S. C. Chan, E. J. Kendon, S. Berthou, G. Fossier, E. Lewis, and H. J. Fowler.
710 Europe-wide precipitation projections at convection permitting scale with
711 the Unified Model. *Climate Dynamics*, 55(3):409–428, 2020. Publisher:
712 Springer.
- 713 K. Chancibault, S. Anquetin, V. Ducrocq, and G.-M. Saulnier. Hydrological
714 evaluation of high-resolution precipitation forecasts of the Gard flash-flood
715 event (8–9 September 2002). *Quarterly Journal of the Royal Meteorological*
716 *Society*, 132(617):1091–1117, Apr. 2006. ISSN 00359009, 1477870X. doi:
717 10.1256/qj.04.164. URL <http://doi.wiley.com/10.1256/qj.04.164>.
- 718 D. Chen, J. Guo, D. Yao, Y. Lin, C. Zhao, M. Min, H. Xu, L. Liu, X. Huang,
719 T. Chen, and P. Zhai. Mesoscale Convective Systems in the Asian Mon-
720 soon Region From Advanced Himawari Imager: Algorithms and Preliminary
721 Results. *Journal of Geophysical Research: Atmospheres*, 124(4):2210–2234,
722 Feb. 2019. ISSN 2169-897X, 2169-8996. doi: 10.1029/2018JD029707. URL
723 <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029707>.
- 724 J. M. Ciarlo, E. Coppola, A. Fantini, F. Giorgi, X. Gao, Y. Tong, R. H. Glazer,
725 J. A. T. Alavez, T. Sines, E. Pichelli, and others. A new spatially distributed
726 added value index for regional climate models: the EURO-CORDEX and the
727 CORDEX-CORE highest resolution ensembles. *Climate Dynamics*, pages
728 1–22, 2020. Publisher: Springer.

- 729 A. J. Clark, R. G. Bullock, T. L. Jensen, M. Xue, and F. Kong. Application of
730 object-based time-domain diagnostics for tracking precipitation systems in
731 convection-allowing models. *Weather and Forecasting*, 29(3):517–542, 2014.
- 732 E. Coppola, S. Sobolowski, E. Pichelli, F. Raffaele, B. Ahrens, I. Anders,
733 N. Ban, S. Bastin, M. Belda, D. Belusic, and others. A first-of-its-kind
734 multi-model convection permitting ensemble for investigating convective
735 phenomena over Europe and the Mediterranean. *Climate Dynamics*, 55(1):
736 3–34, 2020. Publisher: Springer.
- 737 E. Coppola, P. Stocchi, E. Pichelli, J. A. Torres Alavez, R. Glazer, G. Giu-
738 liani, F. Di Sante, R. Nogherotto, and F. Giorgi. Non-Hydrostatic RegCM4
739 (RegCM4-NH): model description and case studies over multiple domains.
740 *Geoscientific Model Development*, 14(12):7705–7723, Dec. 2021. ISSN 1991-
741 9603. doi: 10.5194/gmd-14-7705-2021. URL [https://gmd.copernicus.org/
742 articles/14/7705/2021/](https://gmd.copernicus.org/articles/14/7705/2021/).
- 743 P. F. Cortopassi and M. Daddi. Discariche Di Cava E Instabilita Dei Ver-
744 santi: Valutazione Preliminare Di Alcuni Fattori Significativi Nel Bacino
745 Marmifero Di Carrara (Italia) / Quarrywaste And Slope Instability: Pre-
746 liminary Assessment Of Some Controlling Factors In The Carrara Marble
747 Basin (Italy). *Italian Journal of Engineering Geology and Environment*,
748 (2008):99–118, Oct. 2008. doi: 10.4408/IJEGE.2008-01.S-08. URL <https://doi.org/10.4408/IJEGE.2008-01.S-08>.
- 750 G. C. Craig, C. Keil, and D. Leuenberger. Constraints on the impact of radar
751 rainfall data assimilation on forecasts of cumulus convection. *Quarterly
752 Journal of the Royal Meteorological Society*, 138(663):340–352, 2012.
- 753 J. Creutin, H. Andrieu, and D. Faure. Use of a weather radar for
754 the hydrology of a mountainous area. Part II: radar measurement val-
755 idation. *Journal of Hydrology*, 193(1):26–44, 1997. ISSN 0022-1694.
756 doi: 10.1016/S0022-1694(96)03203-9. URL [https://www.sciencedirect.com/
757 science/article/pii/S0022169496032039](https://www.sciencedirect.com/science/article/pii/S0022169496032039).
- 758 J. Crook, C. Klein, S. Folwell, C. M. Taylor, D. J. Parker, R. Stratton, and
759 T. Stein. Assessment of the representation of West African storm lifecycles in
760 convection-permitting simulations. *Earth and Space Science*, 6(5):818–835,
761 2019. Publisher: Wiley Online Library.
- 762 C. Davis, B. Brown, and R. Bullock. Object-Based Verification of Precipitation
763 Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas.
764 *Monthly Weather Review*, 134(7):1772–1784, July 2006a. ISSN 1520-0493,
765 0027-0644. doi: 10.1175/MWR3145.1. URL [http://journals.ametsoc.org/
766 doi/10.1175/MWR3145.1](http://journals.ametsoc.org/doi/10.1175/MWR3145.1).

- 767 C. Davis, B. Brown, and R. Bullock. Object-Based Verification of Precipitation
768 Forecasts. Part II: Application to Convective Rain Systems. *Monthly*
769 *Weather Review*, 134(7):1785–1795, July 2006b. ISSN 1520-0493, 0027-0644.
770 doi: 10.1175/MWR3146.1. URL [http://journals.ametsoc.org/doi/10.1175/
771 MWR3146.1](http://journals.ametsoc.org/doi/10.1175/MWR3146.1).
- 772 G. Delrieu, J. Nicol, E. Yates, P.-E. Kirstetter, J.-D. Creutin, S. Anquetin,
773 C. Obled, G.-M. Saulnier, V. Ducrocq, E. Gaume, and others. The catastro-
774 phic flash-flood event of 8–9 September 2002 in the Gard Region, France:
775 a first case study for the Cévennes–Vivarais Mediterranean Hydrometeorolo-
776 gical Observatory. *Journal of hydrometeorology*, 6(1):34–52, 2005.
- 777 P. Drobinski, V. Ducrocq, P. Alpert, E. Anagnostou, K. Béranger, M. Borga,
778 I. Braud, A. Chanzy, S. Davolio, G. Delrieu, C. Estournel, N. F. Boubrahmi,
779 J. Font, V. Grubišić, S. Gualdi, V. Homar, B. Ivančan-Picek, C. Kottmeier,
780 V. Kotroni, K. Lagouvardos, P. Lionello, M. C. Llasat, W. Ludwig, C. Lut-
781 off, A. Mariotti, E. Richard, R. Romero, R. Rotunno, O. Roussot, I. Ruin,
782 S. Somot, I. Taupier-Letage, J. Tintore, R. Uijlenhoet, and H. Wernli.
783 HyMeX: A 10-Year Multidisciplinary Program on the Mediterranean Water
784 Cycle. *Bulletin of the American Meteorological Society*, 95(7):1063–1082,
785 July 2014. ISSN 0003-0007, 1520-0477. doi: 10.1175/BAMS-D-12-00242.1.
786 URL <http://journals.ametsoc.org/doi/10.1175/BAMS-D-12-00242.1>.
- 787 A. Fantini. *Climate change impact on flood hazard over Italy*. PhD Thesis,
788 University of Trieste, 2019. URL <http://hdl.handle.net/11368/2940009>.
- 789 Q. Fumière, M. Déqué, O. Nuissier, S. Somot, A. Alias, C. Caillaud, O. Lau-
790 rantin, and Y. Seity. Extreme rainfall in Mediterranean France during the
791 fall: added value of the CNRM-AROME Convection-Permitting Regional
792 Climate Model. *Climate Dynamics*, 55(1):77–91, 2020. Publisher: Springer.
- 793 U. Germann, M. Boscacci, L. Clementi, M. Gabella, A. Hering, M. Sartori,
794 I. V. Sideris, and B. Calpini. Weather Radar in Complex Orography. *Remote*
795 *Sensing*, 14(3):503, Jan. 2022. ISSN 2072-4292. doi: 10.3390/rs14030503.
796 URL <https://www.mdpi.com/2072-4292/14/3/503>.
- 797 F. Giorgi. Climate change hot-spots. *Geophysical research letters*, 33(8), 2006.
798 Publisher: Wiley Online Library.
- 799 F. Giorgi. Thirty years of regional climate modeling: where are we and where
800 are we going next? *Journal of Geophysical Research: Atmospheres*, 124(11):
801 5696–5723, 2019. Publisher: Wiley Online Library.
- 802 F. Giorgi. Producing actionable climate change information for regions:
803 the distillation paradigm and the 3R framework. *The European Physi-
804 cal Journal Plus*, 135(5):435, May 2020. ISSN 2190-5444. doi: 10.1140/
805 epjp/s13360-020-00453-1. URL <https://link.springer.com/10.1140/epjp/>

806 [s13360-020-00453-1](#).

807 G. A. Grell, L. Schade, R. Knoche, A. Pfeiffer, and J. Egger. Nonhydrostatic
808 climate simulations of precipitation over complex terrain. *Journal of Geo-*
809 *physical Research: Atmospheres*, 105(D24):29595–29608, Dec. 2000. ISSN
810 01480227. doi: 10.1029/2000JD900445. URL [http://doi.wiley.com/10.1029/](http://doi.wiley.com/10.1029/2000JD900445)
811 [2000JD900445](http://doi.wiley.com/10.1029/2000JD900445).

812 Z. Guo, J. Tang, J. Tang, S. Wang, Y. Yang, W. Luo, and J. Fang.
813 Object-Based Evaluation of Precipitation Systems in Convection-Permitting
814 Regional Climate Simulation Over Eastern China. *Journal of Geophysical*
815 *Research: Atmospheres*, 127(1), Jan. 2022. ISSN 2169-897X, 2169-8996. doi:
816 10.1029/2021JD035645. URL [https://onlinelibrary.wiley.com/doi/10.1029/](https://onlinelibrary.wiley.com/doi/10.1029/2021JD035645)
817 [2021JD035645](https://onlinelibrary.wiley.com/doi/10.1029/2021JD035645).

818 R. M. Haralick and L. G. Shapiro. *Computer and robot vision*. Addison-Wesley
819 Pub. Co, Reading, Mass, 1992. ISBN 978-0-201-10877-4 978-0-201-56943-8.

820 C. Heim, D. Panosetti, L. Schlemmer, D. Leuenberger, and C. Schär. The
821 Influence of the Resolution of Orography on the Simulation of Orographic
822 Moist Convection. *Monthly Weather Review*, 148(6):2391–2410, June 2020.
823 ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-19-0247.1. URL [http://](http://journals.ametsoc.org/doi/10.1175/MWR-D-19-0247.1)
824 journals.ametsoc.org/doi/10.1175/MWR-D-19-0247.1.

825 C. Hohenegger, P. Brockhaus, and C. Schär. Towards climate sim-
826 ulations at cloud-resolving scales. *Meteorologische Zeitschrift*, 17(4):
827 383–394, Aug. 2008. ISSN 0941-2948. doi: 10.1127/0941-2948/2008/
828 0303. URL [http://www.schweizerbart.de/papers/metz/detail/17/56729/](http://www.schweizerbart.de/papers/metz/detail/17/56729/Towards_climate_simulations_at_cloud_resolving_sca?af=crossref)
829 [Towards_climate_simulations_at_cloud_resolving_sca?af=crossref](http://www.schweizerbart.de/papers/metz/detail/17/56729/Towards_climate_simulations_at_cloud_resolving_sca?af=crossref).

830 F. A. Isotta, C. Frei, V. Weilguni, M. Perčec Tadić, P. Lassègues, B. Rudolf,
831 V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, M. Munari, S. Micheletti,
832 V. Bonati, C. Lussana, C. Ronchi, E. Panettieri, G. Marigo, and
833 G. Vertačnik. The climate of daily precipitation in the Alps: development
834 and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge
835 data: CLIMATE OF DAILY PRECIPITATION IN THE ALPS. *Internation-*
836 *al Journal of Climatology*, 34(5):1657–1675, Apr. 2014. ISSN 08998418.
837 doi: 10.1002/joc.3794. URL [https://onlinelibrary.wiley.com/doi/10.1002/](https://onlinelibrary.wiley.com/doi/10.1002/joc.3794)
838 [joc.3794](https://onlinelibrary.wiley.com/doi/10.1002/joc.3794).

839 A. Johnson, X. Wang, F. Kong, and M. Xue. Object-Based Evaluation
840 of the Impact of Horizontal Grid Spacing on Convection-Allowing Fore-
841 casts. *Monthly Weather Review*, 141(10):3413–3425, Oct. 2013. ISSN
842 0027-0644, 1520-0493. doi: 10.1175/MWR-D-13-00027.1. URL [http://](http://journals.ametsoc.org/doi/10.1175/MWR-D-13-00027.1)
843 journals.ametsoc.org/doi/10.1175/MWR-D-13-00027.1.

- 844 E. J. Kendon, N. Ban, N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan,
845 J. P. Evans, G. Fossier, and J. M. Wilkinson. Do convection-permitting
846 regional climate models improve projections of future precipitation change?
847 *Bulletin of the American Meteorological Society*, 98(1):79–93, 2017. Pub-
848 lisher: American Meteorological Society.
- 849 E. J. Kendon, C. Short, J. Pope, S. C. Chan, J. Wilkinson, S. Tucker,
850 P. Bett, G. Harris, and J. Murphy. Update to ukcp local (2.2km) pro-
851 jections. Technical report, United Kingdom Met Office, Exeter, United
852 Kingdom, 7 2021. URL [https://www.metoffice.gov.uk/research/approach/
853 collaboration/ukcp/guidance-science-reports](https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports).
- 854 K. Keuler, K. Radtke, S. Kotlarski, and D. Lüthi. Regional climate change
855 over Europe in COSMO-CLM: Influence of emission scenario and driving
856 global model. *Meteorologische Zeitschrift*, 25(2):121–136, 2016. Publisher:
857 Schweizerbart.
- 858 P. La Barbera, L. Lanza, and L. Stagi. Tipping bucket mechanical errors
859 and their influence on rainfall statistics and extremes. *Water Science
860 and Technology*, 45(2):1–9, Jan. 2002. ISSN 0273-1223, 1996-9732. doi:
861 10.2166/wst.2002.0020. URL [https://iwaponline.com/wst/article/45/2/1/
862 9061/Tipping-bucket-mechanical-errors-and-their](https://iwaponline.com/wst/article/45/2/1/9061/Tipping-bucket-mechanical-errors-and-their).
- 863 C. Lebeaupin, V. Ducrocq, and H. Giordani. Sensitivity of torrential rain
864 events to the sea surface temperature based on high-resolution numerical
865 forecasts. *Journal of Geophysical Research: Atmospheres*, 111(D12), 2006.
- 866 K. Lochbihler, G. Lenderink, and A. P. Siebesma. The spatial extent of rainfall
867 events and its relation to precipitation scaling. *Geophysical Research Letters*,
868 44(16):8629–8636, 2017. Publisher: Wiley Online Library.
- 869 P. Lucas-Picher, D. Argüeso, E. Brisson, Y. Trambly, P. Berg, A. Lemonsu,
870 S. Kotlarski, and C. Caillaud. Convection-permitting modeling with regional
871 climate models: Latest developments and next steps. *WIREs Climate
872 Change*, 12(6), Nov. 2021. ISSN 1757-7780, 1757-7799. doi: 10.1002/wcc.731.
873 URL <https://onlinelibrary.wiley.com/doi/10.1002/wcc.731>.
- 874 J. Lundquist, M. Hughes, E. Gutmann, and S. Kapnick. Our Skill in
875 Modeling Mountain Rain and Snow is Bypassing the Skill of Our Obser-
876 vational Networks. *Bulletin of the American Meteorological Society*, 100
877 (12):2473–2490, Dec. 2019. ISSN 0003-0007, 1520-0477. doi: 10.1175/
878 BAMS-D-19-0001.1. URL [https://journals.ametsoc.org/view/journals/
879 bams/100/12/bams-d-19-0001.1.xml](https://journals.ametsoc.org/view/journals/bams/100/12/bams-d-19-0001.1.xml).
- 880 S. Medina and R. A. Houze Jr. Air motions and precipitation growth in Alpine
881 storms. *Quarterly Journal of the Royal Meteorological Society*, 129(588):
882 345–371, Jan. 2003. ISSN 1477870X, 00359009. doi: 10.1256/qj.02.13. URL

883 <http://doi.wiley.com/10.1256/qj.02.13>.

884 E. P. Meredith, U. Ulbrich, and H. W. Rust. Subhourly rainfall in a convection-
885 permitting model. *Environmental Research Letters*, 15(3):034031, 2020.

886 M. M. Miglietta, A. Manzato, and R. Rotunno. Characteristics and pre-
887 dictability of a supercell during HyMeX SOP1. *Quarterly Journal of the*
888 *Royal Meteorological Society*, 142(700):2839–2853, Oct. 2016. ISSN 0035-
889 9009, 1477-870X. doi: 10.1002/qj.2872. URL <https://onlinelibrary.wiley.com/doi/10.1002/qj.2872>.

891 C. Morel and S. Senesi. A climatology of mesoscale convective systems over
892 Europe using satellite infrared imagery. II: Characteristics of European
893 mesoscale convective systems. *Quarterly Journal of the Royal Meteorolog-*
894 *ical Society*, 128(584):1973–1995, July 2002a. ISSN 00000000, 00359009.
895 doi: 10.1256/003590002320603494. URL [http://doi.wiley.com/10.1256/](http://doi.wiley.com/10.1256/003590002320603494)
896 [003590002320603494](http://doi.wiley.com/10.1256/003590002320603494).

897 C. Morel and S. Senesi. A climatology of mesoscale convective systems over
898 Europe using satellite infrared imagery. I: Methodology. *Quarterly Journal*
899 *of the Royal Meteorological Society*, 128(584):1953–1971, July 2002b. ISSN
900 00000000, 00359009. doi: 10.1256/003590002320603485. URL <http://doi.wiley.com/10.1256/003590002320603485>.

902 G. M. Morgan. A General Description of the Hail Problem in the Po Val-
903 ley of Northern Italy. *Journal of Applied Meteorology*, 12(2):338–353, Mar.
904 1973. ISSN 0021-8952. doi: 10.1175/1520-0450(1973)012<0338:AGDOTH>
905 2.0.CO;2. URL [http://journals.ametsoc.org/doi/10.1175/1520-0450\(1973\)](http://journals.ametsoc.org/doi/10.1175/1520-0450(1973)012(0338:AGDOTH)2.0.CO;2)
906 [012\(0338:AGDOTH\)2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0450(1973)012(0338:AGDOTH)2.0.CO;2).

907 C. Moseley, P. Berg, and J. O. Haerter. Probing the precipitation life cycle by
908 iterative rain cell tracking. *Journal of Geophysical Research: Atmospheres*,
909 118(24):13–361, 2013. Publisher: Wiley Online Library.

910 P. Nabat, S. Somot, C. Cassou, M. Mallet, M. Michou, D. Bouniol,
911 B. Decharme, T. Drugé, R. Roehrig, and D. Saint-Martin. Modula-
912 tion of radiative aerosols effects by atmospheric circulation over the
913 Euro-Mediterranean region. *Atmospheric Chemistry and Physics*, 20(14):
914 8315–8349, 2020. Publisher: Copernicus GmbH.

915 U. Neu, M. G. Akperov, N. Bellenbaum, R. Benestad, R. Blender, R. Caballero,
916 A. Coccozza, H. F. Dacre, Y. Feng, K. Fraedrich, J. Grieger, S. Gulev,
917 J. Hanley, T. Hewson, M. Inatsu, K. Keay, S. F. Kew, I. Kindem, G. C.
918 Leckebusch, M. L. R. Liberato, P. Lionello, I. I. Mokhov, J. G. Pinto,
919 C. C. Raible, M. Reale, I. Rudeva, M. Schuster, I. Simmonds, M. Sinclair,
920 M. Sprenger, N. D. Tilinina, I. F. Trigo, S. Ulbrich, U. Ulbrich, X. L. Wang,

- 921 and H. Wernli. IMILAST: A Community Effort to Intercompare Extratrop-
 922 ical Cyclone Detection and Tracking Algorithms. *Bulletin of the American*
 923 *Meteorological Society*, 94(4):529–547, Apr. 2013. ISSN 1520-0477. doi:
 924 10.1175/BAMS-D-11-00154.1. URL [https://journals.ametsoc.org/doi/10.](https://journals.ametsoc.org/doi/10.1175/BAMS-D-11-00154.1)
 925 [1175/BAMS-D-11-00154.1](https://journals.ametsoc.org/doi/10.1175/BAMS-D-11-00154.1).
- 926 L. Panziera, C. N. James, and U. Germann. Mesoscale organization and struc-
 927 ture of orographic precipitation producing flash floods in the Lago Maggiore
 928 region: Orographic Convection in the Lago Maggiore Area. *Quarterly Jour-*
 929 *nal of the Royal Meteorological Society*, 141(686):224–248, Jan. 2015. ISSN
 930 00359009. doi: 10.1002/qj.2351. URL [https://onlinelibrary.wiley.com/doi/](https://onlinelibrary.wiley.com/doi/10.1002/qj.2351)
 931 [10.1002/qj.2351](https://onlinelibrary.wiley.com/doi/10.1002/qj.2351).
- 932 E. Pichelli, R. Rotunno, and R. Ferretti. Effects of the Alps and Apen-
 933 nines on forecasts for Po Valley convection in two HyMeX cases: Effects of
 934 Alps and Apennines on Po Valley Convection Forecasts. *Quarterly Journal*
 935 *of the Royal Meteorological Society*, 143(707):2420–2435, July 2017. ISSN
 936 00359009. doi: 10.1002/qj.3096. URL [https://onlinelibrary.wiley.com/doi/](https://onlinelibrary.wiley.com/doi/10.1002/qj.3096)
 937 [10.1002/qj.3096](https://onlinelibrary.wiley.com/doi/10.1002/qj.3096).
- 938 E. Pichelli, E. Coppola, S. Sobolowski, N. Ban, F. Giorgi, P. Stocchi,
 939 A. Alias, D. Belušić, S. Berthou, C. Caillaud, R. M. Cardoso, S. Chan,
 940 O. B. Christensen, A. Dobler, H. de Vries, K. Goergen, E. J. Kendon,
 941 K. Keuler, G. Lenderink, T. Lorenz, A. N. Mishra, H.-J. Panitz, C. Schär,
 942 P. M. M. Soares, H. Truhetz, and J. Vergara-Temprado. The first
 943 multi-model ensemble of regional climate simulations at kilometer-scale res-
 944 olution part 2: historical and future simulations of precipitation. *Climate*
 945 *Dynamics*, 56(11-12):3581–3602, June 2021. ISSN 0930-7575, 1432-0894.
 946 doi: 10.1007/s00382-021-05657-4. URL [https://link.springer.com/10.1007/](https://link.springer.com/10.1007/s00382-021-05657-4)
 947 [s00382-021-05657-4](https://link.springer.com/10.1007/s00382-021-05657-4).
- 948 J. G. Powers, J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O.
 949 Gill, J. L. Coen, D. J. Gochis, R. Ahmadov, S. E. Peckham, and others.
 950 The weather research and forecasting model: Overview, system efforts, and
 951 future directions. *Bulletin of the American Meteorological Society*, 98(8):
 952 1717–1737, 2017.
- 953 A. Prein, R. Rasmussen, D. Wang, and S. Giangrande. Sensitivity of orga-
 954 nized convective storms to model grid spacing in current and future climates.
 955 *Philosophical Transactions of the Royal Society A*, 379(2195):20190546,
 956 2021. Publisher: The Royal Society Publishing.
- 957 A. F. Prein and A. Gobiet. Impacts of uncertainties in European gridded pre-
 958 cipitation observations on regional climate analysis. *International Journal*
 959 *of Climatology*, 37(1):305–327, 2017d. Publisher: Wiley Online Library.

- 960 A. F. Prein, W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen,
961 M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli,
962 N. P. M. Lipzig, and R. Leung. A review on regional convection-permitting
963 climate modeling: Demonstrations, prospects, and challenges. *Reviews of*
964 *Geophysics*, 53(2):323–361, June 2015. ISSN 8755-1209, 1944-9208. doi:
965 10.1002/2014RG000475. URL [https://onlinelibrary.wiley.com/doi/10.1002/](https://onlinelibrary.wiley.com/doi/10.1002/2014RG000475)
966 [2014RG000475](https://onlinelibrary.wiley.com/doi/10.1002/2014RG000475).
- 967 A. F. Prein, C. Liu, K. Ikeda, R. Bullock, R. M. Rasmussen, G. J. Holland,
968 M. Clark, and others. Simulating North American mesoscale convective
969 systems with a convection-permitting climate model. *Climate Dynamics*,
970 pages 1–16, 2017a. Publisher: Springer.
- 971 A. F. Prein, C. Liu, K. Ikeda, S. B. Trier, R. M. Rasmussen, G. J. Holland,
972 and M. P. Clark. Increased rainfall volume from future convective storms in
973 the US. *Nature Climate Change*, 7(12):880–884, 2017b. Publisher: Nature
974 Publishing Group.
- 975 C. Purr, E. Brisson, and B. Ahrens. Convective Shower Characteristics Simu-
976 lated with the Convection-Permitting Climate Model COSMO-CLM. *Atmo-*
977 *sphere*, 10(12):810, Dec. 2019. ISSN 2073-4433. doi: 10.3390/atmos10120810.
978 URL <https://www.mdpi.com/2073-4433/10/12/810>.
- 979 R. Rasmussen, C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, M. Tewari,
980 M. Barlage, J. Dudhia, W. Yu, K. Miller, K. Arsenault, V. Grubišić,
981 G. Thompson, and E. Gutmann. High-Resolution Coupled Climate Runoff
982 Simulations of Seasonal Snowfall over Colorado: A Process Study of Cur-
983 rent and Warmer Climate. *Journal of Climate*, 24(12):3015–3048, June
984 2011. ISSN 0894-8755, 1520-0442. doi: 10.1175/2010JCLI3985.1. URL
985 <http://journals.ametsoc.org/doi/10.1175/2010JCLI3985.1>.
- 986 A. Reder, M. Raffa, M. Montesarchio, and P. Mercogliano. Performance
987 evaluation of regional climate model simulations at different spatial and
988 temporal scales over the complex orography area of the Alpine region.
989 *Natural Hazards*, 102(1):151–177, May 2020. ISSN 0921-030X, 1573-0840.
990 doi: 10.1007/s11069-020-03916-x. URL [http://link.springer.com/10.1007/](http://link.springer.com/10.1007/s11069-020-03916-x)
991 [s11069-020-03916-x](http://link.springer.com/10.1007/s11069-020-03916-x).
- 992 A. Reder, M. Raffa, R. Padulano, G. Rianna, and P. Mercogliano. Character-
993 izing extreme values of precipitation at very high resolution: An experiment
994 over twenty European cities. *Weather and Climate Extremes*, 35:100407,
995 Mar. 2022. ISSN 22120947. doi: 10.1016/j.wace.2022.100407. URL [https://](https://linkinghub.elsevier.com/retrieve/pii/S2212094722000019)
996 linkinghub.elsevier.com/retrieve/pii/S2212094722000019.
- 997 R. E. Rinehart and E. T. Garvey. Three-dimensional storm motion detection
998 by conventional weather radar. *Nature*, 273(5660):287–289, May 1978. ISSN
999 0028-0836, 1476-4687. doi: 10.1038/273287a0. URL <http://www.nature>.

1000 [com/articles/273287a0](https://doi.org/10.1002/articles/273287a0).

1001 B. Rockel, A. Will, and A. Hense. The regional climate model COSMO-
1002 CLM (CCLM). *Meteorologische Zeitschrift*, 17(4):347–348, 2008. Publisher:
1003 Berlin: Borntraeger, c1992-.

1004 R. Rotunno and R. A. Houze. Lessons on orographic precipitation from the
1005 Mesoscale Alpine Programme. *Quarterly Journal of the Royal Meteorological*
1006 *Society*, 133(625):811–830, Apr. 2007. ISSN 00359009, 1477870X. doi: 10.
1007 1002/qj.67. URL <https://onlinelibrary.wiley.com/doi/10.1002/qj.67>.

1008 M. Rummukainen. Added value in regional climate modeling. *WIREs Climate*
1009 *Change*, 7(1):145–159, Jan. 2016. ISSN 1757-7780, 1757-7799. doi: 10.1002/
1010 wcc.378. URL <https://onlinelibrary.wiley.com/doi/10.1002/wcc.378>.

1011 S. Sauvagnargues-Lesage. Retour d’expérience sur la gestion de l’événement
1012 de Septembre 2002 par les services de Sécurité Civile. *La Houille Blanche*,
1013 90(6):107–113, Nov. 2004. ISSN 0018-6368, 1958-5551. doi: 10.1051/
1014 lhb:200406015. URL [https://www.tandfonline.com/doi/full/10.1051/lhb%
1015 3A200406015](https://www.tandfonline.com/doi/full/10.1051/lhb%3A200406015).

1016 M. Schleiss, J. Olsson, P. Berg, T. Niemi, T. Kokkonen, S. Thorndahl,
1017 R. Nielsen, J. Ellerbæk Nielsen, D. Bozhinova, and S. Pulkkinen. The
1018 accuracy of weather radar in heavy rain: a comparative study for den-
1019 mark, the netherlands, finland and sweden. *Hydrology and Earth System*
1020 *Sciences*, 24(6):3157–3188, 2020. doi: 10.5194/hess-24-3157-2020. URL
1021 <https://hess.copernicus.org/articles/24/3157/2020/>.

1022 T. Stein, R. Hogan, K. Hanley, P. Clark, C. Halliwell, H. Lean, J. Nicol,
1023 and R. Plant. The three-dimensional microphysical structure of convective
1024 storms over the southern United Kingdom. *Monthly Weather Review*, 142:
1025 3264–3283, 2014.

1026 B. Stevens. ATMOSPHERIC MOIST CONVECTION. *Annual Review*
1027 *of Earth and Planetary Sciences*, 33(1):605–643, May 2005. ISSN 0084-
1028 6597, 1545-4495. doi: 10.1146/annurev.earth.33.092203.122658. URL <https://www.annualreviews.org/doi/10.1146/annurev.earth.33.092203.122658>.

1030 P. Tabary, P. Dupuy, G. L’Hirou, C. Gueguen, L. Moulin, O. Laurantin,
1031 C. Merlier, and J.-M. Soubeyrou. A 10-year (1997–2006) reanalysis of quan-
1032 titative precipitation estimation over France: methodology and first results.
1033 *IAHS Publ*, 351:255–260, 2012.

1034 C. Torma, F. Giorgi, and E. Coppola. Added value of regional climate mod-
1035 eling over areas characterized by complex terrain-Precipitation over the
1036 Alps: ADDED VALUE OF RCM OVER COMPLEX TERRAIN. *Journal*
1037 *of Geophysical Research: Atmospheres*, 120(9):3957–3972, May 2015. ISSN

1038 2169897X. doi: 10.1002/2014JD022781. URL [http://doi.wiley.com/10.1002/](http://doi.wiley.com/10.1002/2014JD022781)
1039 [2014JD022781](http://doi.wiley.com/10.1002/2014JD022781).

1040 E. van Meijgaard, L. Van Ulft, W. Van de Berg, F. Bosveld, B. Van den Hurk,
1041 G. Lenderink, and A. Siebesma. *The KNMI regional atmospheric climate*
1042 *model RACMO, version 2.1*. KNMI De Bilt, The Netherlands, 2008.

1043 E. Van Meijgaard, L. Van Ulft, G. Lenderink, S. De Roode, E. L. Wipfler,
1044 R. Boers, and R. van Timmermans. *Refinement and application of a regional*
1045 *atmospheric model for climate scenario calculations of Western Europe*.
1046 Number KVR 054/12. KVR, 2012.

1047 H. Wernli, M. Paulat, M. Hagen, and C. Frei. SAL—A Novel Quality Mea-
1048 sure for the Verification of Quantitative Precipitation Forecasts. *Monthly*
1049 *Weather Review*, 136(11):4470–4487, Nov. 2008. ISSN 1520-0493, 0027-0644.
1050 doi: 10.1175/2008MWR2415.1. URL [http://journals.ametsoc.org/doi/10.](http://journals.ametsoc.org/doi/10.1175/2008MWR2415.1)
1051 [1175/2008MWR2415.1](http://journals.ametsoc.org/doi/10.1175/2008MWR2415.1).

1052 T. Winterrath, C. Brendel, M. Hafer, T. Junghänel, A. Klameth, K. Lengfeld,
1053 E. Walawender, E. Weigl, and A. Becker. RADKLIM Version 2017.002:
1054 Reprocessed gauge-adjusted radar data, one-hour precipitation sums (RW).
1055 *Deutscher Wetterdienst (DWD)*, 2018.

1056 M. Wüest, C. Frei, A. Altenhoff, M. Hagen, M. Litschi, and C. Schär. A gridded
1057 hourly precipitation dataset for Switzerland using rain-gauge analysis and
1058 radar-based disaggregation. *International Journal of Climatology*, 30(12):
1059 1764–1775, 2010. Publisher: Wiley Online Library.

1060 List of Figures

- 1061 1 Historical events Gard 2002 (left panel) and Carrara 2003 (right
1062 panel) using the observations. Shading illustrates accumulated
1063 total precipitation, $P(\text{total})$ [mm]. Filled circles indicate the
1064 location of the centroid of a track and their radius is propor-
1065 tional to their respective rain volume. Filled contours indicate
1066 the elevation of the model terrain in intervals of 250 m. 14
- 1067 2 Relative bias in characteristic of a) the mean and b) the 90th
1068 percentile of track properties of each tracker with respect to
1069 the observations $\frac{ModEns - Obs}{Obs}$ [%]. Black solid lines are incre-
1070 ments of $\pm 5\%$, with the thick black line representing the tracker
1071 ensemble mean of the observations, i.e. 0%. 15
- 1072 3 Annual Cycle of track occurrence frequency, OF [month^{-1}], for
1073 all trackers analyzing the observations. Error bars indicate inter-
1074 annual variability by the temporal standard deviation. 16

1075	4	Annual cycles for the tracker ensemble mean of both the model ensemble (dashed line) and observations (solid line), with panel a) showing track occurrence frequency, OF [month ⁻¹], panel b) accumulated precipitation of tracks, P(tracks) [mm month ⁻¹], panel c) heavy precipitation fraction, P(tracks)/P(total) [%], and panel d) accumulated total precipitation, P(total) [mm month ⁻¹]. Error bars for the model ensemble indicate the temporal standard deviation of the model ensemble mean across years, and likewise for the observations error bars display inter-annual variability by the temporal standard deviation across years.	17
1086	5	Panel a) shows the track occurrence frequency density, OFD [month ⁻¹ pixel ⁻¹], of the tracker ensemble mean for the model ensemble and panel b) shows the same for the observations. Panel c) shows their difference and panel d) show the difference, but with model ensemble and observations being normalized by their total number of tracks, respectively. A pixel is here defined as the reference area of 0.36° × 0.36°. The green iso-line shows the model elevation at 1000 m.a.s.l.. Panel e) shows the model-observation difference in track occurrence by model elevation.	19
1096	6	Panels a) and c): The purple shaded area illustrates the relative bias of a) the tracker ensemble mean of the model ensemble mean and c) the 90th percentile, with respect to the observations: $\frac{ModEns^{Tr} - Obs^{Tr}}{Obs^{Tr}} [\%]$, while purple lines indicate the individual models and green lines individual years of the observations. Panels b) and d) shows also $\frac{ModEns^{Tr} - Obs^{Tr}}{Obs^{Tr}} [\%]$ for the mean and 90th percentile of characteristic properties, with the Italian dataset GRIPHO excluded as well as for GRIPHO only. Black solid lines are increments of +5%, with the thick black line denoting 0%. Panels e) to j): probability density functions of Duration [h], Area [km ²], Rain Volume [m ³ E6], Space-Time Volume [km ² h], Maximum Precipitation Rate [mm h ⁻¹] and Mean Precipitation Rate [mm h ⁻¹].	22
1109	7	The spatial biases of the tracker ensemble mean of the model ensemble w.r.t. observations. Panels a) to f): duration [h], (mean) area [km ²], rain volume [m ³ E6], (geometrical) volume [km ² h], mean and maximum precipitation (rate) [mm h ⁻¹]. Again a pixel is here defined as the reference area of 0.36° × 0.36° and the green iso-line shows the model elevation at 1000 m.a.s.l..	23

1116	8	Relative Biases of each model w.r.t. to the model ensemble mean	
1117		using the tracker ensemble mean: $\frac{Model^{Tr} - ModEns^{Tr}}{ModEns^{Tr}}$ [%]. Black	
1118		solid lines are increments of +5%, with the thick black line	
1119		representing the tracker ensemble mean of the model ensemble,	
1120		i.e. 0%.	4
1121	9	Spatial mapping of characteristic track properties using the	
1122		tracker ensemble mean, for the model ensemble (left column)	
1123		and the observations (right column), From top to bottom:	
1124		Duration [h], Area [km ²], Rain Volume [m ³ E6], (geometri-	
1125		cal) Volume [km ² h], Mean Precipitation Rate [mm h ⁻¹] and	
1126		Maximum Precipitation Rate [mm h ⁻¹].	5

1127 **List of Tables**

1128	1	Summary of numerical models used in this study. "dx" denotes the grid spacing of the respective model. Model documentation references: (A) (Powers et al., 2017); (B) (Rockel et al., 2008; Baldauf et al., 2011); (B1) (Keuler et al., 2016); (C) (Caillaud et al., 2021); (C1) (Nabat et al., 2020); (D) (van Meijgaard et al., 2008; Van Meijgaard et al., 2012; Belušić et al., 2020); (E) (Coppola et al., 2021); (F) (Chan et al., 2020)	6
1129	2	Summary of observational datasets of hourly precipitation rate used in this study. dx [km] denotes the original grid spacing.	7
1130	3	Summary of the tracker setup.	11
1131	4	Definitions of characteristic track properties.	11
1132	5	Summary of Trackers investigated in this study: "Institute" denotes the group executing the analysis using "Tracker", representing the abbreviation of the respective tracker. "space/time smoothing" denotes whether the tracker has an option for smoothing the input field prior to the analysis in space or time. "metatrack at splitting/merging" denotes whether the tracker assigns separate <i>metatracks</i> when tracks merge or split, with the second line indicating whether this functionality is switch on or off. Column "statistics" indicates whether the statistical properties of the tracks are derived from the original or from the smoothed input field. In "reference" the original description of the tracking algorithm is found.	12
1133	6	Averaged and integrated track properties, as defined in Table 4, associated with historic events Gard 2002 and Carrara 2003. $\max x^T$ and $\sum x^T$ denote the maximum value and summation of property x over all tracks T.	13
1134	7	Qualitative attribution of track properties to the two historical events, Carrara 2003 and Gard 2002.	14
1135	8	Climatology of track properties derived from both the model ensemble and the observations. In Table 4, definition of the respective properties are given.	15
1136	9	Qualitative biases of characteristic track properties that are consistent across all trackers.	16
1137	10	Relative biases [%] of characteristic track properties, as defined in Table 4, using the tracker ensemble mean, for the whole domain, as well as without GRIPHO and exclusively for GRIPHO.	20
1138			
1139			
1140			
1141			
1142			
1143			
1144			
1145			
1146			
1147			
1148			
1149			
1150			
1151			
1152			
1153			
1154			
1155			
1156			
1157			
1158			
1159			
1160			
1161			
1162			
1163			
1164			

1165 8 Tracker Descriptions

1166 We here provide detailed descriptions of the tracking algorithms and supple-
1167 mentary material.

1168 8.1 MTD

1169 The Method for Object-Based Diagnostic Evaluation (MODE) time domain
1170 tool (MTD) is part of the Model Evaluation Tools (MET, see [https://
1171 dtcenter.org/community-code/model-evaluation-tools-met](https://dtcenter.org/community-code/model-evaluation-tools-met)). It is developed,
1172 maintained and made freely available by the Developmental Testbed Center
1173 and used here by ICTP. The toolbox comprises various analysis tools developed
1174 for the evaluation of numerical atmospheric models. (Davis et al., 2006a,b) first
1175 introduce the basic methodology of MODE and demonstrate the advantages
1176 of an object-based evaluation in numerical weather predictions. Later on the
1177 tracking of objects in time was added and the capability of MTD in describing
1178 the characteristic properties of rain systems in both simulations and observa-
1179 tions on a continental scale was explored in (Clark et al., 2014). Finally (Prein
1180 et al., 2017a) applied MTD in order to identify mesoscale convective systems
1181 in convection-permitting climate simulations of North America.

1182 The algorithm can be summarized as follows:

- 1183 1. the field is smoothed by convolution in space, across a radius (here: 1 grid
1184 cell, which means smoothing across 3×3 grid cells) and in time across a
1185 number of time steps (here: $+0$).
- 1186 2. a precipitation threshold (here: 5 mm h^{-1}) is applied and thereafter only
1187 grid cells exceeding the threshold are considered
- 1188 3. adjacent cells in space and time are clustered to form objects.
- 1189 4. a minimum volume threshold is applied (here: 100 grid cells), meaning that
1190 all object that are too small will be dropped.

1191 The output of the analysis is a set of tracks, that represent precipitation events,
1192 along with information about their respective location, scale, intensity and
1193 propagation.

1194 The location of a track is given through the geometrical centroid across all
1195 grid cells associated with the track in space and time. Due to computational
1196 limitations the tracker only processes periods of 10 days at a time. Since we are
1197 here looking at events with time scales much shorter than that, we don't expect
1198 the analysis being deteriorated much due to this. The mask for comparison
1199 of observations against models was applied after the analysis. All statistical
1200 properties presented here are derived from the raw un-smoothed precipitation
1201 field, whereas the grid cells associated with the event are identified from the
1202 smoothed field. Along the boundaries and 1 smoothing radius inwards the
1203 input field is set to zero before applying the smoothing.

2 LIST OF TABLES

1204 **8.2 OSIRIS**

1205 The precipitating system detection and tracking algorithm used by CNRM
1206 is based on the algorithm developed at CNRM (Morel and Senesi, 2002a,b)
1207 applied in precipitation nowcasting at Meteo-France and for the evaluation
1208 of the numerical weather prediction model AROME (Brousseau et al., 2016).
1209 It has also been recently used in an evaluation study of CNRM-AROME on
1210 Mediterranean Heavy Precipitation Events (Caillaud et al., 2021). The 1-hour
1211 accumulated precipitation fields are used as input of the tool and the method
1212 can be summarised as follows:

- 1213 1. Smoothing : first, each grid cell is replaced by a weighted average of the
1214 3×3 adjacent grid cells and second, a Gaussian filter is applied with a small
1215 standard deviation (0.5) allowing for a slight smoothing;
- 1216 2. Detection of the precipitating systems every hour with a minimum surface
1217 of 20 km^2 and contiguous grid points exceeding several intensity thresholds
1218 (here: 5 mm h^{-1});
- 1219 3. Tracking of system trajectories by identifying links between systems at dif-
1220 ferent time steps according to overlapping and correlation conditions. The
1221 overlapping condition uses the velocity of the cells calculated between dif-
1222 ferent time steps with a minimum recovery rate of 15 %. The correlation
1223 condition is based on spatial correlation calculation between cells at differ-
1224 ent time steps present in a research box around the cell, with a minimum
1225 correlation of 0.4;
- 1226 4. Minimum volume threshold applied (here: 100 grid cells),
- 1227 5. Diagnostics: each cell is schematized as an ellipse: centre of gravity, length
1228 of the minor axis and the major axis, angle and the main characteristics
1229 of each trajectory can be calculated (location, duration, area, mean and
1230 maximum intensity, velocity, ...). The different characteristics are calculated
1231 on the smoothed field.

1232 A further description of the algorithm can be found in (Caillaud et al., 2021).

1233 **8.3 celltrack**

1234 The tracker (celltrack) used by KNMI is described in detail in (Lochbihler
1235 et al., 2017) and is inspired by the work of (Moseley et al., 2013). By default
1236 celltrack does not use prior smoothing of the input field. To make celltrack
1237 comparable to the other trackers, the input field was smoothed using a 3×3
1238 box-smoothing. This smoothed input field is used in all subsequent steps. First
1239 individual cells above a precipitation threshold (5 mm h^{-1}) are diagnosed, not
1240 considering a specific minimum area. These cells are subsequently linked into
1241 tracks. A six-fold iterative advection correction is implemented using advective
1242 velocities derived on a coarse-grained grid. After the linking, various track
1243 types can be diagnosed (e.g., single "clean" tracks, mergers, splits, etc) follow-
1244 ing a specific taxonomic classification (Lochbihler et al., 2017). The optional
1245 diagnostics of sub-cell and mainstream detection are not used in this study.
1246 The Fortran code is available on GitHub.

1247 **8.4 DYMECS**

1248 The precipitation system detection and tracking algorithm was originally
1249 developed for sub-hourly radar and forecast model precipitation data (Stein
1250 et al., 2014). Since then, it has been applied to hourly climate model data with
1251 resolutions as coarse as 25km (Crook et al., 2019).

1252 The algorithm is divided into two parts: the detection of objects-of-interest
1253 for each image and the tracking of these objects-of-interest between consecutive
1254 images. The detection algorithm is based on the "local table method" (Har-
1255 alick and Shapiro, 1992), labelling pixels-of-interest by line-by-line scanning.
1256 The tracking component is based on the windowed cross-correlation between
1257 consecutive precipitation images (Rinehart and Garvey, 1978). Windowed cor-
1258 relations between consecutive images are computed, and velocities between
1259 images estimated. Objects identified in the previous image are then moved by
1260 those estimated velocities. Areal overlap of objects between the two images
1261 are computed. If the object overlapping fraction exceeds 0.6, the overlapping
1262 objects are considered part of the same track, with splitting and merging
1263 allowed if there are multiple overlapping. If two or more objects can be traced
1264 back to a single object in the previous image, a split occurs with the object
1265 with higher area overlap retaining the same track identifier (i.e., metatrack)
1266 and the other objects labelled as new tracks. For the opposite case of two or
1267 more objects from the previous image tracing to a single object in the cur-
1268 rent image, a merge occurs, and retains the identifier with the track with the
1269 highest overlap; other merged objects and their identifiers cease to exist.

1270 Smoothing is not originally part of the algorithm. This is added for this
1271 study, using the same Gaussian blurring approach as in (Caillaud et al., 2021).
1272 Smoothing is applied to the detection phase only, affecting only pixel labelling
1273 without changing the underlying precipitation intensities.

1274 The code is written in MATLAB/OCTAVE and is available from the Met
1275 Office upon request.

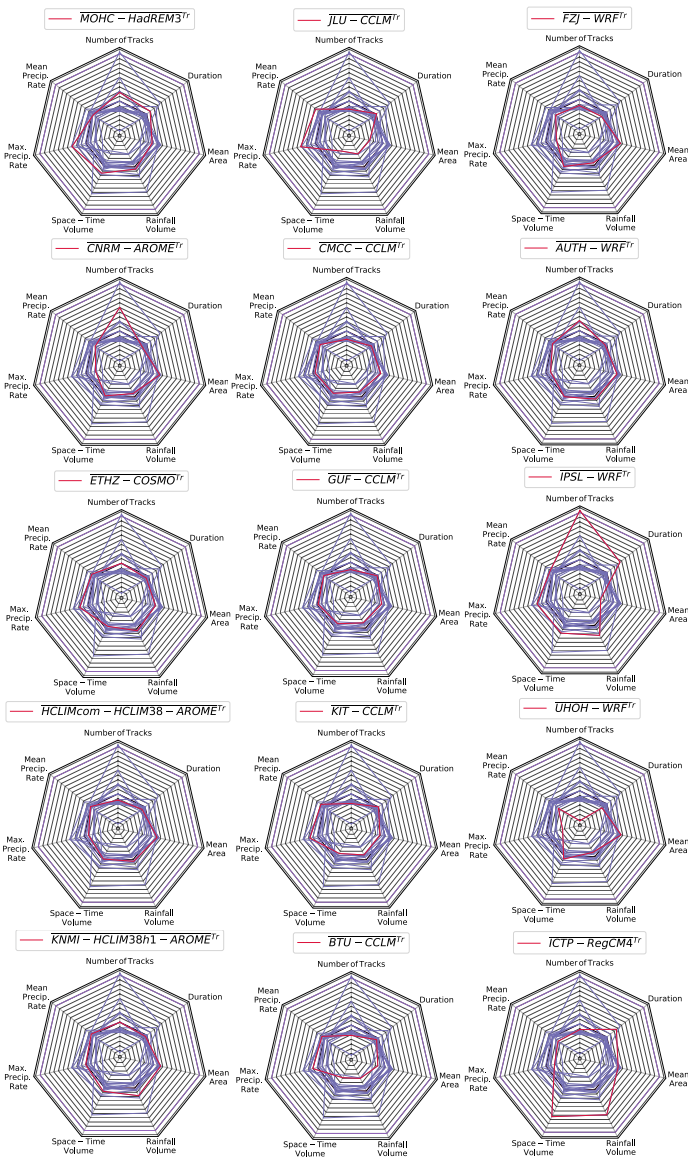


Fig. 8 Relative Biases of each model w.r.t. to the model ensemble mean using the tracker ensemble mean: $\frac{Model^{Tr} - ModEns^{Tr}}{ModEns^{Tr}} [\%]$. Black solid lines are increments of $\pm 5\%$, with the thick black line representing the tracker ensemble mean of the model ensemble, i.e. 0%.

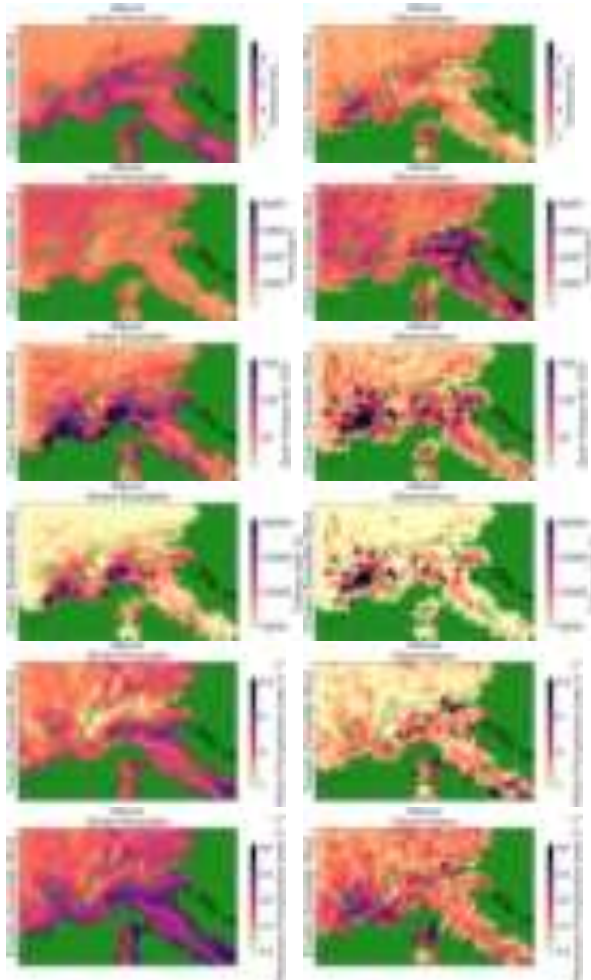


Fig. 9 Spatial mapping of characteristic track properties using the tracker ensemble mean, for the model ensemble (left column) and the observations (right column), From top to bottom: Duration [h], Area [km^2], Rain Volume [$\text{m}^3 \text{E}6$], (geometrical) Volume [$\text{km}^2 \text{h}$], Mean Precipitation Rate [mm h^{-1}] and Maximum Precipitation Rate [mm h^{-1}].