BINGO

a better future under CLIMATE CHANGE

BRINGING INNOVATION TO ONGOING WATER MANAGEMENT

D2.5

Ensembles for present climate extremal episodes downscaled to 7km/6h (3-1km/1h); maps of return levels for Cyprus research site

December 2016

www.projectbingo.eu

The BINGO project has received funding from the European Union's Horizon 2020 Research and Innovation programme, under the Grant Agreement number 641739.

Horizon 2020 Societal challenge 5:
Climate action, environment, resource efficiency and raw materials

BINGO

Bringing INnovation to onGOing water management – a better future under climate change

Grant Agreement n° 641739, Research and Innovation Action

Deliverable number:	**D2.5**
Deliverable name:	**Ensembles for present climate extremal episodes downscaled to 7km/6h (3-1km/1h); maps of return levels for Cyprus research site**
WP / WP number:	WP2: Climate predictions and downscaling to extreme weather
Delivery due date:	Project month 18 (31/12/2016)
Actual date of submission:	14/12/2016
Dissemination level:	Public
Lead beneficiary:	CYI
Responsible scientist/administrator:	George Zittis (CYI)
Estimated effort (PM):	17
Contributor(s):	Adriana Bruggeman (CYI), Corrado Camera (CYI), Panos Hadjinicolaou (CYI)
Estimated effort contributor(s) (PM):	3
Internal reviewer:	Komlan Agbeko Kpogo Nuwoklo (FUB), Bernard Voortman(KWR)

Changes with respect to the DoA
Maps of return levels for daily precipitation events for Cyprus are presented in D2.4.

Dissemination and uptake
Users of hydrological impact/resources models for following WP's (e.g. WP3).

Short Summary of results (<250 words)
This deliverable consists of data for present climate extremal pattern episodes (15 days duration) downscaled to 12km and 4km/6h (20 ensemble members) and to 1km/1h (5 ensemble members) for the Cyprus research site. Five extreme rainfall events over Cyprus were identified from observations and were dynamically downscaled from the ERA-Interim (EI) dataset with the Weather Research and Forecasting model (WRF), for 15-day periods centered around the peak of each event. For validation we used a 1-km gridded observational dataset over Cyprus. Additionally, we explored the potential added value gained from the EI downscaling with WRF. This was done in both terms of timing and rainfall amounts for each horizontal resolution nest of the downscaling process. Simulations with WRF captured rainfall over the eastern Mediterranean reasonably well for three of the five selected extreme events. For these three cases, the higher spatial resolution WRF simulations were found to improve the ERA-Interim precipitation amounts, which strongly underestimated these rainfall extremes over Cyprus. The best model performance was obtained for the January 1989 event, which was simulated with an average negative bias of 4% and a modified Nash-Sutcliff efficiency of 0.72 for a 5-member ensemble of the 1-km simulations. Our 1-km simulations overall indicate a higher added value, especially over regions of high-elevation. Interestingly, for some cases the intermediate 4-km nest was found to outperform the 1-km simulations for low-elevation coastal parts of Cyprus.

Evidence of accomplishment
This report, peer-reviewed book chapter published and manuscript under submission.

Zittis G., Hadjinicolaou P., Bruggeman A., Camera C., Lelieveld J. (2017). High-resolution simulations of recent past extreme precipitation events over Cyprus. In: Perspectives on Atmospheric Sciences, Springer Atmospheric Sciences, ISBN: 978-3-319-35095-0

Zittis G., Hadjinicolaou P., Bruggeman A., Camera C., Lelieveld J. (2016). High-resolution simulations of extreme precipitation events of the recent past over the eastern Mediterranean. Atmospheric Research (under submission)

TABLE OF CONTENTS

1. INTRODUCTION

The objective of this deliverable is to dynamically downscale recent past extreme precipitation events for the research site of Cyprus. Output of this task, in the form of an ensemble set, will be further used to drive hydrological impact models. As a downscaling tool, we used the Weather, Research and Forecasting Model (WRF) forced by the ERA-Interim reanalysis data. The latter were downscaled in 12-, 4- and 1-km resolutions over the broader region of the eastern Mediterranean and eventually Cyprus according to Figure 1.

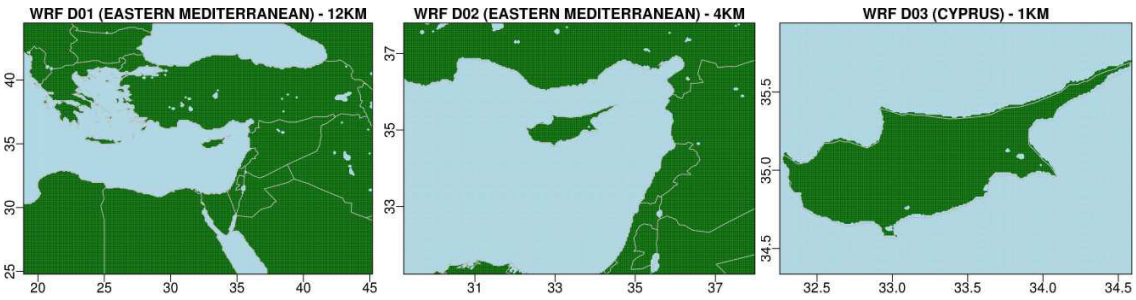

Figure 1 – Simulation domains of the eastern Mediterranean region centered over Cyprus for the 12 km (WRF-D01 – left panel), 4 km (WRF-D02 – middle panel) and 1 km (WRF-D03 – right panel) simulations.

Our ensemble was formulated by utilizing different WRF options for convection and microphysical processes parameterizations, the most critical model components for rainfall generation. Since there is no universal optimum setup, we test a large number of model configurations in order to investigate which ones are more skilful in reproducing five past extreme precipitation events over the region of interest. Each combination of four cumulus and five microphysics schemes was utilized while the rest of the model's physics was identical. This gave a total of 20 WRF configurations. All simulations have a length of 15 days with the peak of the events more or less centered within this period. This might not be the optimum for other type of applications, such as weather forecasting. However, we wanted to investigate the model's skill more from a climate mode perspective, i.e. to reduce the dependence on the initial conditions.

As already described in the previous paragraph, convection of warm air that is related to thunderstorms is parameterized since it naturally occurs in much smaller scales than the typical resolution of global or regional models. This was the case for our 12- and 4-km simulations. Nevertheless, an increasing number of studies for various locations and based on different models indicate a more realistic representation of precipitation

processes in high-resolution convection-permitting simulations (Prein et al., 2013; Cassola et al., 2015; Davolio et al., 2015; Fosser et al., 2015; Meredith et al., 2015; Prein et al., 2015). The spatial resolution of the latter is fine enough to explicitly resolve convective processes. Therefore, after applying an objective ranking scheme for identifying the best WRF configurations at 4-km resolution, we further downscale the five best configurations to 1-km horizontal resolution to explore the advances of convection-permitting simulations during the selected case studies.

Hourly precipitation output, from the "hydrologically-best" WRF configurations in the very-high 1-km grid spacing will be eventually used to force the hydrological impact models of WP3 for the research site of Cyprus.

2. DATA AND METHODOLOGY

2.1. Model and experimental setup

For the simulations of this study we used the version 3.7.1 of WRF model (Skamarock et al., 2008). It was driven by the ERA-Interim (EI) reanalysis dataset (Dee et al., 2011), which provide the initial and boundary conditions. The latter were updated every six hours. We used the one-way nesting option, downscaling EI (≈80 km) to 12 km (Domain 1: D01) and subsequently to 4 km (Domain 2: D02) and 1 km (Domain 3: D03), as seen in Figure 1. The first downscaling domain (D01) represents more or less the area overlapped from the different CORDEX domains that are relevant for the research site of Cyprus (EURO-CORDEX, MENA-CORDEX, MED-CORDEX). Preliminary test simulations of the two-way nesting and nudging options were found not to improve the results for this particular experimental setup and weather events. The model was set up with 40 vertical levels and a top at 50 hPa. To be useful for hydro-climatological applications, all model output was saved in hourly intervals.

2.2. Observational and reanalysis data

For the evaluation of the model output we used a daily 1 by 1 km gridded precipitation dataset (Camera et al. 2014), hereafter called OBS. This dataset was derived from a dense network of 145 rain gauges, which covers the area of the island that is under the jurisdiction of the Republic of Cyprus. For the purposes of this study, these gridded observations are considered to be the closest representation of reality. To perform a fair comparison between the three resolution nests we also averaged OBS to the 12- and 4-km grids.

To get a more representative picture of precipitation during the events in the broader region of the eastern Mediterranean we used three state-of-the-art, gridded daily observational datasets. These include version v.11.0 of the widely-used E-OBS dataset (Haylock et al. 2008), covering mainly Europe, and version V1101R1 of the APHRODITE's (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation) daily gridded precipitation (Yatagai et al., 2012). Both have a horizontal grid spacing of 25 X 25 km. The coverage of APHRODITE for the MENA region ends in 2007, thus, it was not available for the January 2010 extreme precipitation event. Additionally, we used the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) dataset (Funk et al. 2015). The latter is a

blended product of both satellite and station measurements and has a 5 X 5 km horizontal resolution, spanning a zone from 50°S to 50°N for all longitudes.

Since, for this part of the world, time and space consistent upper-air gridded observations either do not exist or are not publicly available, we utilized two reanalysis datasets in order to define the synoptic meteorological conditions during the selected extreme precipitation events. Besides the EI dataset, which was also used to provide initial and boundary conditions for our WRF simulations, we used geopotential values of the 500 hPa level derived from the NCEP-Reanalysis-2 dataset (Kanamitsu et al. 2002). The latter was provided by the Physical Science Division of the Earth System Research Laboratory/NOAA, from their web site at http://www.esrl.noaa.gov/psd/. The 500 hPa level is important in synoptic meteorology and weather forecasting since it is very near the level of non-divergence while it also represents the level where the about one half of the atmosphere's mass is below it and half is above it. We additionally used mean sea level pressure (MSLP) data from the same two reanalysis sources. These are considered to give a fair representation of the reality as both surface and upper air observations were assimilated during their simulations.

Finally, to investigate the vertical structure of the atmosphere during the events we used temperature (T) and dew temperature (Td) data from soundings. In meteorology, temperature profile is widely used for the study of atmosphere's vertical structure (e.g. definition of tropopause, identification of the planetary boundary layer etc.). Moreover, the coincidence of T and Td curves in this type of plots provides useful information on the available water content of the atmosphere and can indicate cloud or precipitation formation. Unfortunately, this kind of information was not available for Cyprus for the selected past events and therefore we used soundings from the adjacent station of Adana in south Turkey and only for the 1994 and 2010 cases. The latter were available from the Department of Atmospheric Science of the University of Wyoming (http://weather.uwyo.edu/ upperair/sounding.html).

2.3. Description of case studies

Five extreme precipitation events of the recent past (1980-2010) were selected based on precipitation and surface flow measurements on the northern side of the research site of Troodos Mountains (Table 1). The 1988 and 1989 events were marked by high surface runoff in the watersheds along the northern slopes of the Troodos. The 1994 event was the largest 2-, 3-, 4-, and 5-day rainfall over Cyprus for the period 1980-

2010; the 2005 event was a rare high summer rainfall, with extensive flooding in the capital city Nicosia; the 2010 event was the largest 1-day rainfall event over Cyprus.

Table 1 - Definition of the case studies and description of the extreme precipitation events.

	Simulation days	Season	Day of Peak	Average Precipitation (mm) over Cyprus during peak day	Max. point precipitation (mm) during peak day	Average Monthly Precipitation (mm) over Cyprus (1980-2010)
Case 1988	17/12-31/12	Wet	24/Dec	55	126	104
Case 1989	02/01-16/01	Wet	09/Jan	50.1	144.6	98.1
Case 1994	12/11-26/11	Wet	21/Nov	47.8	140.7	62.4
Case 2005	25/05-08/06	Dry	31/May	11.2	115.4	17.4
Case 2010	11/01-25/01	Wet	18/Jan	66.5	162.5	98.1

2.4. Selection of best configurations

To objectively assess if our WRF output is valid for hydro-climatological applications, we compared it with the high-resolution Cyprus observations by applying a set of relevant indices of extremes and efficiency metrics. These were applied on the daily model output and observations, for each model grid point over Cyprus, and were then averaged. Our selection of evaluation metrics is summarized below.

The absolute error of the Maximum 5-day Precipitation Total ($RX5D_{ERROR}$) (Frich et al. 2002) is computed as follows:

$$\mathrm{RX5D_{ERROR}} = \left|\mathrm{RX5D_{OBS}}\text{-}\mathrm{RX5D_{SIM}}\right| \qquad (1)$$

where *RX5D* is the maximum 5-day precipitation total during the two-week simulations and the subscripts OBS and SIM indicate the observed and simulated data. This five-day period is defined from the observations.

The absolute error of the Simple Daily Intensity Index (*SDII*) (Frich et al., 2002) is defined as:

$$\mathrm{SDII_{ERROR}} = \left|\mathrm{SDII_{OBS}}\text{-}\mathrm{SDII_{SIM}}\right| \qquad (2)$$

where SDII is the total precipitation over the period divided by the number of rainy days. Rainy days are defined as days with precipitation larger than 1mm.

The Percent Bias (*PBIAS*) measures the average tendency of the simulated values to be larger or smaller than their observed ones (Sorooshian et al. 1993):

$$\mathrm{PBIAS} = 100\frac{\sum_{i=1}^{N}(\mathrm{SIM_i}\text{-}\mathrm{OBS_i})}{\sum_{i=1}^{N}\mathrm{OBS_i}} \qquad (3)$$

where N is the number of simulation days times the number of grid cells.

The Nash-Sutcliffe efficiency (*NSE*) determines the relative magnitude of the residual variance ("noise") compared to the measured data variance ("information"). The modified *NSE*, is not inflated by the squared values of the differences, because the squares in the original formula are replaced by absolute values (Krause et al., 2005). It ranges from –Infinitive to 1 (Perfect score =1). It is computed as follows:

$$mNSE = 1 - \frac{\sum_{i=1}^{N}|SIM_i\text{-}OBS_i|}{\sum_{i=1}^{N}|OBS_i\text{-}\overline{OBS}|} \qquad (4)$$

The Kling-Gupta Efficiency (*KGE*) was developed by Gupta et al. (2009) to provide a diagnostically interesting decomposition of the Nash-Sutcliffe efficiency, which facilitates the analysis of the relative importance of its different components (correlation, bias and variability) in the context of hydrological modelling (Zambrano-Bigiarini, 2015). Kling et al. (2012) proposed a revised version, which was used in this study, to ensure that the bias and variability ratios are not cross-correlated. Kling-Gupta efficiencies range from minus infinity to one. Essentially, the closer to one, the more accurate the model is. A more detailed description of this metric including the extended equations for calculation is presented in Zambrano-Bigiarini (2015).

The Modified Index of Agreement (*MIA*) is a standardized measure of the degree of the model prediction error. This index varies between 0 and 1. It was introduced by Willmott (1981) and refined by Legates and McCabe (1999) as follows:

$$MIA = 1 - \frac{\sum_{i=1}^{N}(OBS_i - SIM_i)}{\sum_{i=1}^{N}(|SIM_i - \overline{OBS}| + |OBS_i - \overline{OBS}|)} \qquad (5)$$

Except for the *RX5D-error*, all metrics are computed over the full 15 days of the simulation. To select the five best performing WRF configurations out of the twenty and further proceed with the finer resolution (1-km) simulations, we performed and objective ranking based on the 4-km results (WRF-D02). We ranked all WRF configurations for each of the metrics described in the previous paragraphs and for each case study independently. Then we averaged the ranks for each configuration. The five configurations with the highest average ranking values (i.e. closer to 1) are considered to be the most realistic for the selected case studies. We refer to those as the five best configurations.

2.5. Added value of dynamical downscaling

To quantify the added value (AV) of our WRF dynamical downscaling in comparison with the global EI rainfall, we applied the method suggested by Di Luca et al. (2015). They quantified the AV by comparing a distance metric (d) between the Global General Circulation Model (GCM) and the observations on one hand, and between the Regional Climate Model (RCM) simulation and the observations on the other, as follows:

$$AV = d(GCM, OBS) - d(RCM, OBS) \quad (6)$$

For our case the driving GCM is the ERA-Interim reanalysis dataset, while RCM refers to the WRF simulations for each of the three model nests. In order to have comparable results for all three WRF resolutions, we calculated the AV only over the five best-performing configurations that eventually drove the WRF-D03 1-km simulations. For this task, we used as distance metrics the six indices described in the previous section, applied similarly to the daily precipitation amounts. To compare AV results obtained from different climate statistics, we normalized the AV metrics by the sum of the RCM and GCM errors, according to Di Luca et al. (2016):

$$\widehat{AV}_d = \frac{d(GCM,OBS) - d(RCM,OBS)}{d(GCM,OBS) + d(RCM,OBS)} \quad (7)$$

The AV values vary between −1 and 1 with larger positive values suggesting smaller RCM errors than GCM errors and thus a higher added value from the RCM.

3. RESULTS

3.1. Selection of best configurations

Figure 2 summarizes the results of the ranking of the 20 WRF setups. This was based on the six skill metrics of the 4-km WRF-D02 simulations of all five extreme events. The ranking among all simulations is presented in the upper panel of this figure, while in the bottom panel we show the ranking of clusters of the same convection and microphysics parameterization schemes. Simulations that utilized the Grell 3D ensemble convection and Lin microphysics (WRF configurations: 1, 6, 11, 16, 17, 18, 19, 20) were found to give a poor reproduction of the five rainfall events, as was indicated by the performance metrics (not shown).

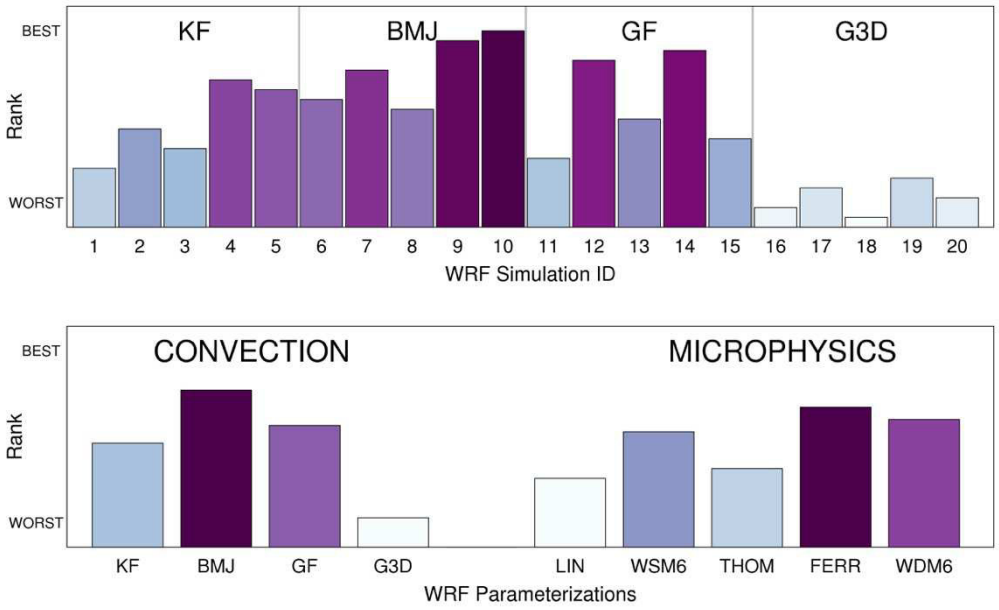

Figure 2 – Ranking results for the 4-km (D02) nest averaged for the five simulated extreme events. The upper panel shows all 20 WRF configurations, while in the bottom panel simulations are grouped according to their convection and microphysics parameterizations. Darker colours indicate a better ranking.

We found a clear advantage in the simulations of extreme precipitation when BMJ and GF convection schemes were used in a combination with Ferrier, WDM6 and WSM6 microphysics (WRF configurations: 7, 9, 10, 12 and 14). These are the five best model configurations that were used for the 1-km simulations. In agreement with our findings, a number of physics inter-comparison studies also indicated that the Bets-Miller-Janjic scheme outperforms in terms of precipitation simulation (Ishak et al., 2012; Evans et al., 2012; Ratna et al., 2014; Remesan et al., 2015). Similarly, the WRF Single and

Double Moment schemes (WSM, WDM) and Ferrier were found to be more accurate (Evans et al., 2012; Givati et al., 2012; Remesan et al., 2015; Kioutsoukis et al., 2016).

3.2. Added value of the dynamical downscaling

In Table 2 we present a summary of the performance metrics for the EI and three WRF nests versus the observations, averaged for each case study. For most case studies and performance metrics the dynamical downscaling with WRF improves the EI precipitation. Exceptions are all quantitative metrics (RX5D error, SDII error and PBIAS) for the December 1988 case and PBIAS for the January 2010 case. Interestingly, for these two events the efficiency metrics (mNSE, KGE and MIA) improved with the downscaling, despite the fact that the precipitation amounts did not. As the PBIAS metric indicates, our WRF simulations tend to underestimate the amount of precipitation during the majority of the simulations. Exception is the very localized June 2005 event, where the model overestimates precipitation in general. The efficiency and agreement indices have generally positive values that are closest to 1 for the January 1989 and November 1994 cases.

Table 2 – Statistical metrics for each case study between the Cyprus precipitation observations and EI, WRF-D01 (12 km), WRF-D02 (4 km) and WRF-D03 (1 km). Stars (*) indicate the cases where only the five-best configurations were considered.

	RX5D error (perfect=0)					SDII error (perfect=0)					PBIAS (perfect=0)				
Case	1988	1989	1994	2005	2010	1988	1989	1994	2005	2010	1988	1989	1994	2005	2010
EI	**34.7**	59.2	101	27.5	66.7	**10.1**	12.8	12.8	6.4	8.5	**2.2**	37.8	32.7	29.0	**46.5**
D01	56.6	50.8	97.2	22.8	76.5	12.8	8.3	12.1	5.6	9.1	48.7	31.6	16.7	137	62.1
D02	58.7	44.9	73.8	25.6	70.2	12.2	8.2	10.1	5.7	8.7	46.8	15.5	-9.6	186	57.6
D02*	58.9	38.0	57.3	21.7	72.2	12.3	7.0	8.6	5.0	8.7	57.8	16.9	16.6	32.9	62.7
D03*	40.7	**23.5**	**44.0**	**13.0**	**56.0**	10.9	**4.7**	**5.5**	**3.7**	**3.3**	51.4	**-9.6**	**4.4**	**22.3**	47.5
	mNSE (perfect =1)					**KGE (perfect=1)**					**MIA (perfect =1)**				
Case	1988	1989	1994	2005	2010	1988	1989	1994	2005	2010	1988	1989	1994	2005	2010
EI	0.02	0.31	0.32	0.06	0.13	0.12	0.29	0.34	**0.34**	0.24	0.50	0.58	0.58	0.54	0.57
D01	0.36	0.52	0.36	0.71	0.25	0.07	0.51	0.32	1.20	0.06	0.64	0.74	0.67	0.50	0.57
D02	0.36	0.50	0.38	1.05	0.26	0.08	0.50	0.41	1.65	0.03	0.63	0.74	0.68	0.51	0.58
D02*	0.41	0.53	0.46	0.14	**0.30**	0.06	0.55	0.51	0.41	0.04	0.65	0.76	0.71	**0.55**	**0.59**
D03*	**0.42**	**0.71**	**0.42**	**0.01**	0.24	**0.12**	**0.72**	**0.56**	0.39	**0.12**	**0.66**	**0.85**	**0.72**	0.54	**0.59**

The results of the added value quantification analysis, averaged over the five events, are presented in Table 3. We found that, on average, all three subsequent downscaling steps added some value to the EI precipitation. In general, this added value was found to be more significant in terms of model efficiency than for the precipitation amounts. Similar as other studies (Prein et al., 2013; Meredith et al., 2015), we found that the further increase of the horizontal resolution and the explicit treatment of convection processes, which was the case for the 1-km simulations, increased the added value.

Table 3 – Added value of the WRF dynamical downscaling, averaged for the five extreme precipitation events and for each distance and efficiency metric.

Added Value (AV)	RX5D error	SDII error	PBIAS	mNSE	KGE	MIA
EI vs WRF-12 km	0.08	0.01	-0.11	0.29	0.55	0.09
EI vs WRF-4 km	0.08	0.1	0.03	0.38	0.86	0.1
EI vs WRF-1 km	0.24	0.29	0.1	0.43	0.92	0.12

To explore the performance of the high-resolution simulations, we present a case by case comparison between the detailed gridded observations over Cyprus and the 4- and 1-km simulations (Figure 3). These maps display results averaged over the five best WRF configurations. For the 1988 event, the 1-km simulations (WRF-D03) were closer to the observations than the 4-km simulations (WRF-DO2), but still both runs strongly underestimated this event. For the January 1989 case, WRF-D03 is found to outperform WRF-D02, especially over the central Troodos Mountains, while it is found to overestimate the event for the western part of the island, which is dominated by the slopes of the Troodos mountains. This is also the case for the November 1994 event, but the D03 precipitation is closer to the observations for the central-east part of the Troodos Mountains. Noteworthy, for these two events WRF-D02 is found to perform better for the flat, southeast part of the island, where WRF-D03 overestimates precipitation.

For the late spring 2005 event the 4-km runs (WRF-D02) are found to better capture rainfall over the eastern part of the island, while on the contrary they missed the very localized precipitation center that was observed between the Nicosia and Kionia stations. The latter, slightly displaced was simulated only by the convection-permitting 1-km runs. Finally, the January 2010 event was not very well reproduced by either the

WRF-D02 or WRF-D03 simulations. However, the 4-km runs reproduced some rainfall over the southeast part of Cyprus that was absent in the 1-km runs. The added value of downscaling to convection-permitting simulations is therefore also found to be a question of location. Nevertheless, for these case studies, the improvement is greater over complex, high-elevation terrains.

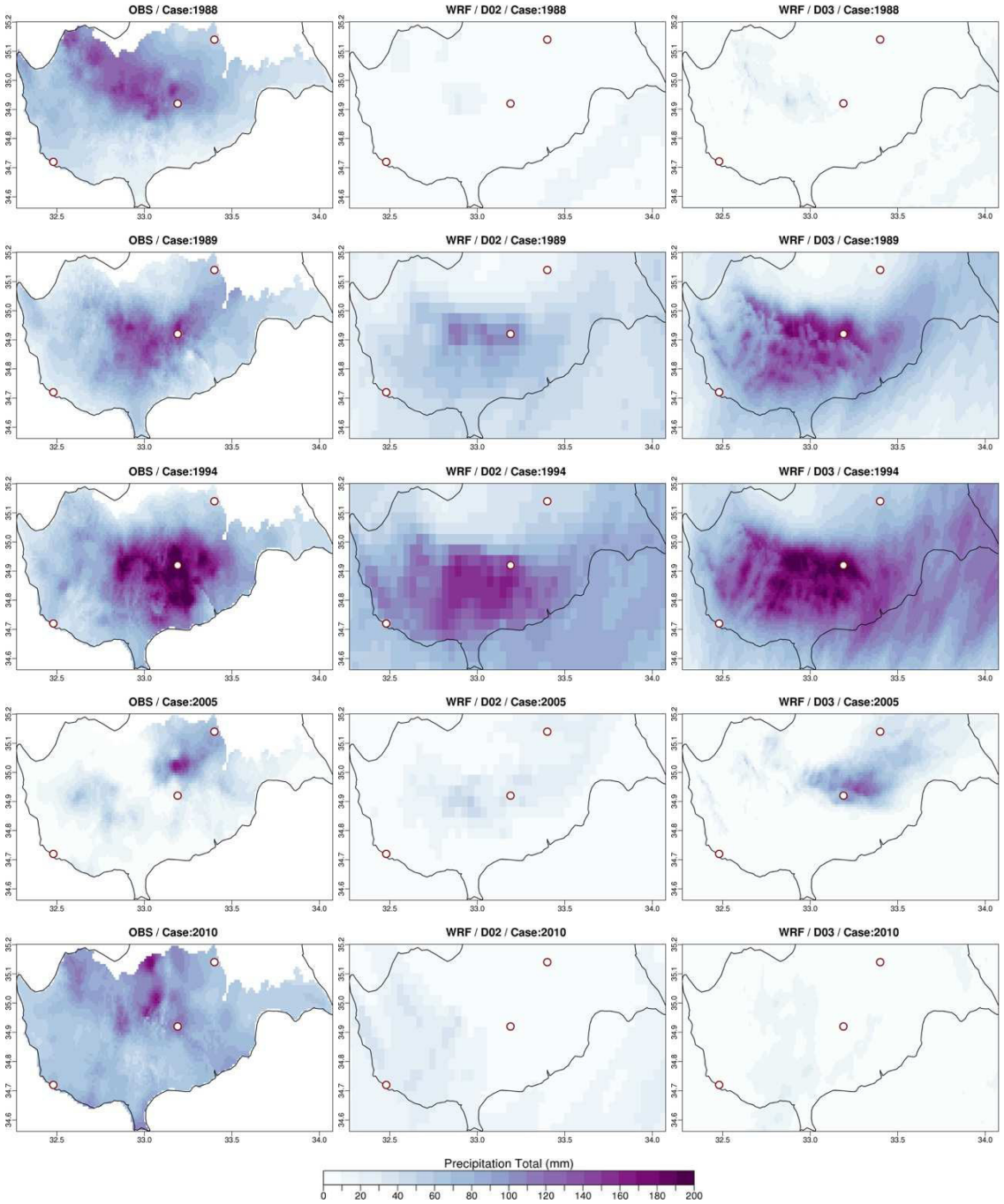

Figure 3 – Precipitation total over a 3-day period centered on the peak of each event for the high-resolution observations (left panels) the 4-km WRF-D02 (middle panels) and the 1-km WRF-D03 simulations (right panels). The latter as the ensemble mean of the five best WRF simulations.

3.3. Detailed analysis of November 1994 case

This extreme precipitation event occurred between the 20th and 22nd of November 1994. This is the only autumn event from the five selected case studies. According to the two reanalysis datasets (EI and NCEP), it was a result of an upper air trough of north-east to south-west direction extended from the southern Balkans towards the north coast of Africa. As depicted in the top row panels of Figure 4, it approached Cyprus from the west. Very similar conditions prevailed also during the 1988 and 1989 case studies (not shown).

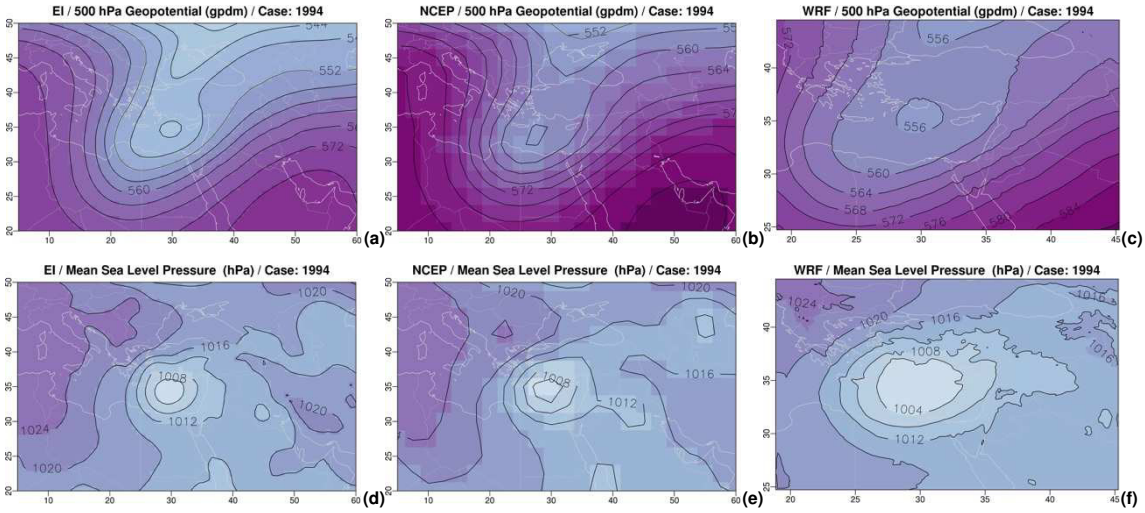

Figure 4 – Geopotential of the 500 hPa level (top row) and mean sea level pressure (bottom row), as daily average for the 23rd of November 1994 and based on the ERA-Interim (left column), NCEP reanalysis (middle column) and WRF 12-km simulations (right column).

A tear-off upper level low within the ridge was also formed over the eastern Mediterranean. Noteworthy, this is a pattern commonly connected with significant precipitation amounts over Cyprus. As a result, low pressure surface conditions prevailed ahead of the upper air trough (Figure 4 – bottom row). This surface depression that caused the extreme rainfall during the event was found to be well captured by our WRF simulations (Figures 4c and 4f).

Figure 5 depicts the observed and modeled vertical structure of the atmosphere in terms of temperature and dew temperature during the peak day of the event. Since there was no information available over Cyprus we used data from the adjacent station of Adana in south Turkey, which was also strongly affected by the event. For this particular case WRF (purple curves) is found skillful in reproducing the observed conditions (blue lines) for the whole extent of the atmosphere for which we have

information from the sounding. Noteworthy, for WRF we present the ensemble mean of the five best configurations of the 12-km simulations. The convergence of the temperature and dew point temperature curves around the 700-500 hPa levels indicates that high availability of water content is evident in both observations and model output, while other features such as the level of the tropopause are also accurately simulated.

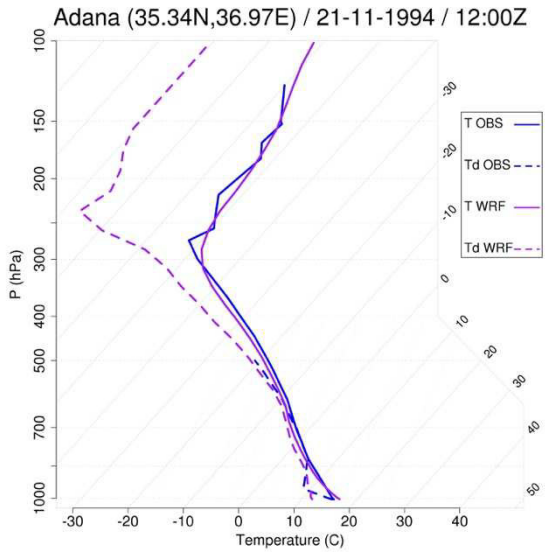

Figure 5 – Vertical structure of the atmosphere for temperature (T) and dew point temperature (Td) based on observations (blue) and WRF 12-km simulations (purple) for the November 1994 event. Radiosonde launched on Adana station in south Turkey.

This low-pressure system resulted in 3-day precipitation totals of more than 200 mm locally and it mostly affected south Turkey and the rest of the Levant coastal areas (Figure 6). CHIRPS precipitation data (Figure 6g-h) are found to be exaggerated, compared to the rest of the observational datasets, in most parts of the study region including Cyprus. However, CHIRPS does not indicate any precipitation over the eastern Mediterranean coast (Israel, Lebanon, Syria), which happens to be the case for E-OBS and, to a less extent, for APHRODITE (Figure 6c-f).

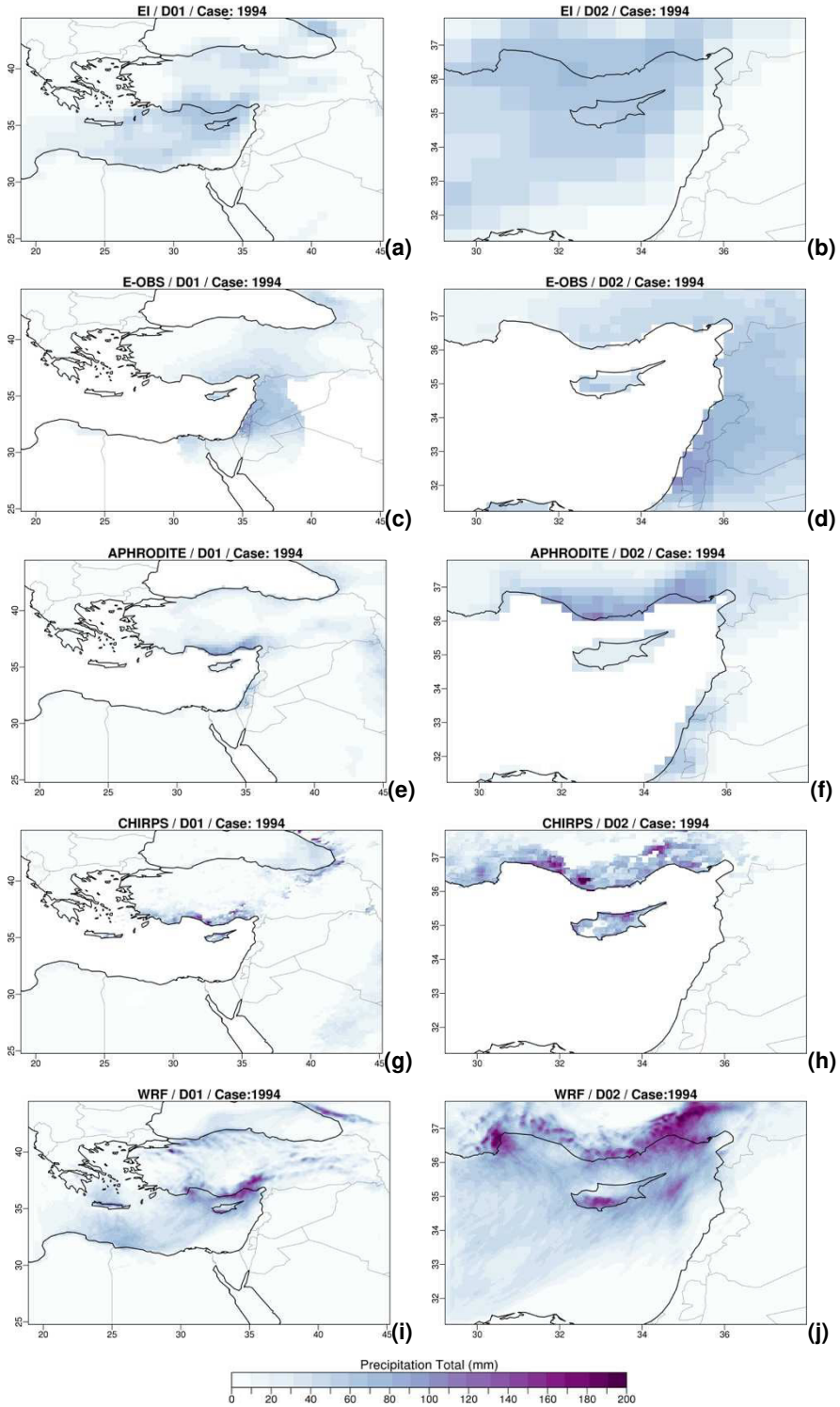

Figure 6 – Precipitation total over the 3-day period 20-22/11/1994 derived from EI (top row), E-OBS (second row), APHRODITE (third row), CHIRPS (fourth row) and ensemble mean of the five best WRF simulations (bottom row) for the 12-km (left panels) and 4-km (right panels) domains.

More regionalized information for Cyprus is presented in Figure 7. Precipitation of EI is found again very homogeneous and underestimates the observed rainfall for large parts of the island (Figure 7a). Generally, CHIRPS underestimates the event over most of the area of the island. On the other hand, our 1-km WRF-D03 simulation is performing relatively well in reproducing the precipitation amount during the November 1994 event (Figure 7d). However, some displacement of the location of the rainfall peaks is evident. In the gridded observations (Figure 7c) the event is found to be centered around Kionia station, however the simulation results in a westward shift of precipitation. Additional information for the whole extent of the simulation can be found in the indicative time-series plots of Figure 8 that indicate a reasonable correlation between the observed and modelled rainfall. A smaller precipitation event during the first days of the simulations, that was most pronounced in Pafos station (Figure 8 – left column), was also reasonably well reproduced by the model.

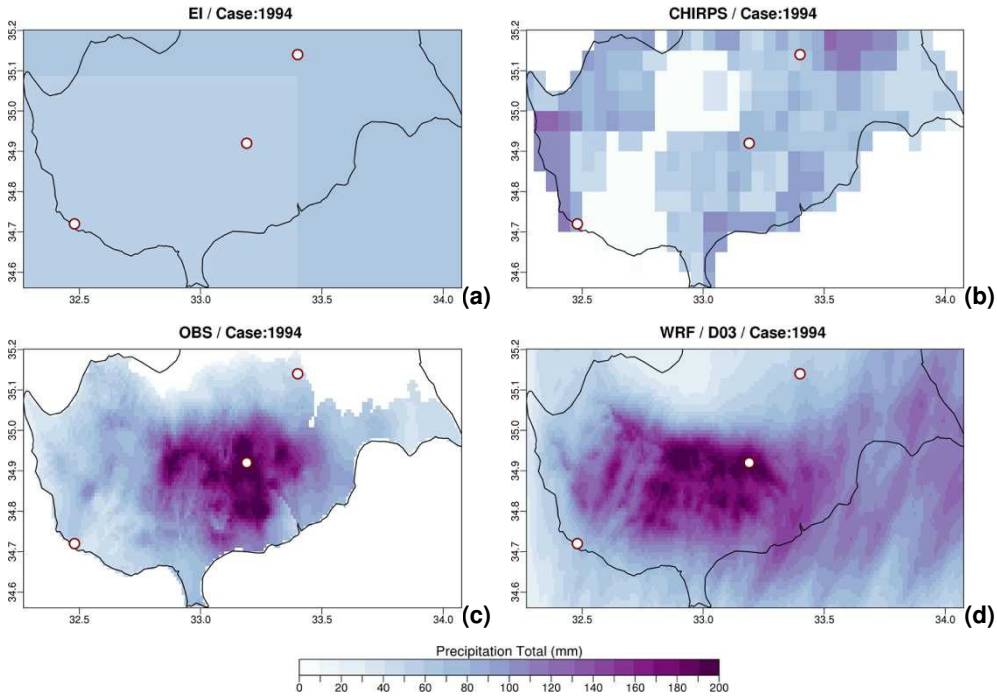

Figure 7 – Precipitation total over the 3-day period 20-22/11/1994 derived from EI (top left), CHIRPS (top right), high-resolution Cyprus observations (bottom left), and ensemble mean of the five best WRF simulations (bottom right) for the 1-km domain. Points indicate the locations of three weather stations for comparison. From west to east: Pafos, Kionia, Nicosia.

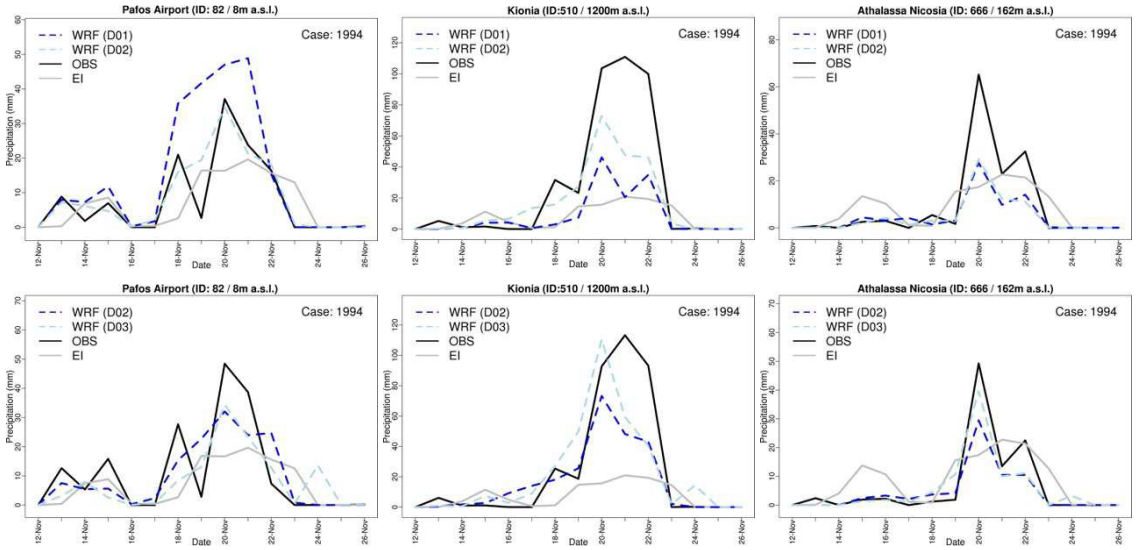

Figure 8 – Time-series of observed and modelled precipitation for the November 1994 extreme precipitation event interpolated in the 12km (top row) and 4km (bottom row) grids for the stations of Pafos, Kionia and Athalassa.

4. CONCLUSIONS AND DISCUSSION

We investigated the WRF model's sensitivity to 20 combinations of convection and microphysical parameterization schemes for five extreme precipitation events over the eastern Mediterranean region. After an objective ranking process, we concluded that the BMJ and GF convection schemes are found to reproduce these events most accurately, while the model's performance was highest when the Ferrier, WDM6 and WSM6 microphysics parameterizations were used. Noteworthy, our findings can be representative for the simulation of precipitation extremes over the region of interest. However, although we used a large number of parameterizations we did not exhaust the available options. Yet, besides the practical value for the BINGO project, this ranking can be still used as a reference for WRF users interested in similar type of environments and weather phenomena.

Our results show that dynamical downscaling with the high-resolution WRF model improves the EI rainfall patterns for the eastern Mediterranean region at least for three out of the five selected extreme precipitation events (1989, 1994 and 2005 cases) in terms of rainfall amounts and timing. Statistical analysis reveals that results of the two higher-resolution domains (4 and 1-km) are consistently closer to the observations Although the simulation was not nudged to the driving reanalysis dataset, for most of the tested events WRF was able to reproduce the synoptic conditions that resulted in the extreme precipitation. Moreover, the reproduced rainfall was generally very well synchronized with the observations, a fact that is not always expected in not-nudged hindcast simulations of such a length. Even for the events where the downscaling did not improve the EI precipitation total amounts (cases 1988 and 2010) the timing of the events was more accurate as indicated by the higher efficiency indices after the downscaling (not shown in this report).

As also suggested by literature, the added value of the downscaling is found higher between the EI and the 1-km nest. These convection-resolving simulations are found to outperform especially over the high elevation regions that are also more critical for the water resources of the island as the flow to the major dams depends on them. These regions are the ones most relevant for the scopes of the BINGO project also because flood events occurring in the mainland most of the times originate from these steeply sloping mountains. These 1-km output, derived from the identified as "hydrologically-best" WRF configurations, will provide the driving data for the hydrological/water resources models of following Work Packages (e.g. WP3).

Noteworthy, for some parts of Cyprus precipitation was better simulated by the 4-km nest, highlighting the need for further research on the advantage of convection-permitting simulations.

Interestingly, we found large discrepancies between three state-of-the-art gridded observational datasets during the tested extreme precipitation events. Our findings indicate that E-OBS and APHRODITE precipitation is systematically lower than CHIRPS over the broader eastern Mediterranean region. Accordingly, for the case of Cyprus the two latter datasets strongly underestimate the extreme precipitation reported by the high-density station net-work. In this context of observational uncertainty, we recommend the use of multiple sources of observations when it comes to model validation.

These findings and a detailed description of the downscaling for all five extreme precipitation events for the BINGO research site of Cyprus are summarised in a manuscript that will be submitted to be considered for publication in Atmospheric Research (http://www.journals.elsevier.com/atmospheric-research).

BIBLIOGRAPHY

Camera C., Bruggeman A., Hadjinicolaou P., Pashiardis S., Lange M.A., 2014. Evaluation of interpolation techniques for the creation of gridded daily precipitation (1 × 1 km2); Cyprus, 1980–2010. *Journal of Geophysical Research, Atmospheres*, 119(2):2013JD020611+

Cassola F., Ferrari F., Mazzino A., 2015. Numerical simulations of Mediterranean heavy precipitation events with the WRF model: A verification exercise using different approaches. *Atmospheric Research,* 164–165, 210-225.

Davolio S., Silvestro F., Malguzzi P., 2015. Effects of increasing horizontal resolution in a Convection-Permitting model on flood forecasting: The 2011 dramatic events in Liguria, Italy. *Journal of Hydrometeorology*, 16(4), 1843-1856, doi: 10.1175/jhm-d-14-0094.1

Dee D. P., et al., 2011. The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553-597.

Di Luca A., de Elía R., Laprise R., 2015. Challenges in the quest for added value of regional climate dynamical downscaling. *Current Climate Change Reports* 1(1), 10-21, doi: 10.1007/s40641-015-0003-9

Di Luca A., Argüeso D., Evans J. P., de Elía R., Laprise, R., 2016. Quantifying the overall added value of dynamical downscaling and the contribution from different spatial scales. *Journal of Geophysical Research: Atmospheres*, 121(4), 1575-1590, doi: 10.1002/2015jd024009

Evans J., Ekström M., Ji F., 2012. Evaluating the performance of a WRF physics ensemble over South-East Australia. *Climate Dynamics*, 39(6):1241-1258, doi: 10.1007/s00382-011-1244-5

Frich P., Alexander L.V., Della-Marta P., Gleason B., Haylock M., Klein Tank A.M.G., Peterson T., 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, 19(3), 193-212.

Funk C., Peterson P., Landsfeld M., Pedreros D., Verdin J., Shukla S., Husak G., Rowland J., Harrison L., Hoell A., Michaelsen J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066, doi:10.1038/sdata.2015.66

Givati A., Lynn B., Liu Y., Rimmer A., 2012. Using the WRF model in an operational streamflow forecast system for the Jordan river. *Journal of Applied Meteorology and Climatology*, 51(2), 285-299.

Gupta H. V., Kling H., Yilmaz, K. K., Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2):80-91.

Haylock M.R., Hofstra N., Klein Tank A.M.G., Klok E.J., Jones P.D., New M., 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. *Journal of Geophysical Research, Atmospheres*, 113, D20119, doi:10.1029/2008JD10201

Ishak A. M., Bray M., Remesan R., Han, D., 2012. Seasonal evaluation of rainfall estimation by four cumulus parameterization schemes and their sensitivity analysis. *Hydrological Processes*, 26(7):1062-1078. doi:10.1002/hyp.8194

Kanamitsu M., Ebisuzaki W., Woollen J., Yang S.-K., Hnilo J. J., Fiorino M., Potter G. L., 2002. NCEP–DOE AMIP-II reanalysis (r-2). *Bulletin of the American Meteorological Society*, 83(11):1631-1643. doi: 10.1175/bams-83-11-1631

Kioutsioukis I., de Meij A., Jakobs H., Katragkou E., Vinuesa J.-F., Kazantzidis A., 2016. High resolution WRF ensemble forecasting for irrigation: Multi-variable evaluation. *Atmospheric Research*, 167, 156-174, doi: 10.1016/j.atmosres.2015.07.015

Kling H., Fuchs M., Paulin M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424-425, 264-277.

Krause P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, 5, 89-97.

Legates D.R., Mc Cabe G.J., 1999. Evaluating the Use of “Goodness-of-Fit” Measures in Hydrologic and Hydroclimatic Model Validation. *Water Resources Research*, 35, 233-241.

Meredith E.P., Maraun D., Semenov V.A., Park W., 2015. Evidence for added value of convection-permitting models for studying changes in extreme precipitation. *Journal of Geophysical Research, Atmospheres*, 120(24), 2015JD024238+.

Prein A.F., Gobiet A., Suklitsch M., Truhetz H., Awan N. K., Keuler K., Georgievski G., 2013. Added value of convection permitting seasonal simulations. *Climate Dynamics*, 41(9), 2655-2677, doi: 10.1007/s00382-013-1744-6

Prein A.F., Langhans W., Fosser G., Ferrone A., Ban N., Goergen K., Keller M., Tölle M., Gutjahr O., Feser F., Brisson E., Kollet S., Schmidli J., van Lipzig N.P.M., Leung R., 2015. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53(2), 323-361. doi:10.1002/2014rg000475

Ratna S., Ratnam J. V., Behera S. K., Rautenbach Ndarana T., Takahashi K., Yamagata T., 2014. Performance assessment of three convective parameterization schemes in WRF for downscaling summer rainfall over south Africa. Climate Dynamics, 42(11-12):2931-2953, doi: 10.1007/s00382-013-1918-2

Remesan R., Bellerby T., Holman I., Frostick L., 2015. WRF model sensitivity to choice of parameterization: a study of the 'York flood 1999'. *Theoretical and applied climatology*, 122(1-2):229-247, doi: 10.1007/s00704-014-1282-0

Skamarock, W.C. et al., 2008, A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-4751STR.

Sorooshian S., Duan Q., Gupta, V.K., 1993. Calibration of rainfall-runoff models: Application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research*, 29(4), 1185-1194, doi: 10.1029/92WR02617

Willmott C.J., 1981. On the Validation of Models. *Physical Geography*, 2, 184-194.

Yatagai A., Kamiguchi K., Arakawa O., Hamada A., Yasutomi N., Kitoh A., 2012. APHRODITE: Constructing a Long-Term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93(9), 1401-1415, doi: 10.1175/bams-d-11-00122.1

Zambrano-Bigiarini M., 2015. Goodness-of-fit functions for comparison of simulated and observed hydrological time series. A description of the HydroGOF R package (https://cran.r-project.org/web/packages/hydroGOF/hydroGOF.pdf)