Journal of Hydrometeorology

The DOI for this manuscript is doi: 10.1175/JHM-D-15-0098.1

Viterbo, F., J. von Hardenberg, A. Provenzale, L. Molini, A. Parodi, O. Sy, and S. Tanelli, 2016: High-resolution simulations of the 2010 Pakistan flood event: sensitivity to parameterizations and initialization time. J. Hydrometeor. doi:10.1175/JHM-D-15-0098.1, in press.

High-resolution simulations of the 2010 Pakistan flood event: sensitivity to parameterizations and initialization time

Francesca Viterbo*

Institute of Atmospheric Sciences and Climate (ISAC-CNR),Torino, Italy and University of Genova, Genova, Italy

Jost von Hardenberg

Institute of Atmospheric Sciences and Climate (ISAC-CNR),Torino, Italy

Antonello Provenzale

Institute of Geosciences and Earth Resources, National Research Council, Pisa, Italy

Luca Molini, and Antonio Parodi

CIMA Foundation, Savona, Italy

Ousmane O. Sy, and Simone Tanelli

Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA, USA

**Corresponding author address:* Institute of Atmospheric Sciences and Climate (ISAC-CNR), Corso Fiume 4, 10133, Torino, Italy

E-mail: f.viterbo@isac.cnr.it

ABSTRACT

Estimating the risk of flood-generating precipitation events in high mountain regions with complex orography is a difficult but crucial task. Quantitative precipitation forecasts (QPF) at fine resolution are an essential ingredient to address this issue. Along these lines, here we explore the ability of the WRF (Weather Research Forecasting) model, operated at 3.5 km grid spacing, to reproduce the extreme meteorological event that led to the 2010 Pakistan flood and produced heavy monsoonal rain in the Indus basin. The model results are compared with Tropical Rainfall Measuring Mission (TRMM) rainfall estimates, the available ground measures and radar observations from the CloudSat mission. In particular, we analyze the sensitivity of the WRF simulations to the use of different convective closures [explicit and Kain-Fritsch (KF)] and microphysical parameterizations [WRF Single-Moment 6-Class Scheme (WSM6) and Thompson]. The impact of using different initial conditions, associated with a different initialization day, is also examined. The use of the new generation DS3 NEOS^3 radar simulator allows a more accurate and extensive representation of the mesoscale processes and of the interaction with the complex orography. The results reported here indicate that the quality of the large scale initial conditions are a prominent factor affecting the possibility to retrieve a realistic representation of this event, when using a non-hydrostatic regional model.

1. Introduction

In 2010, Pakistan experienced a major flood event that started in late July and was triggered by persistent heavy monsoonal rains. Nearly one-fifth of the entire territory of Pakistan was submerged during the floods (Houze Jr et al. 2011) and the UN Secretary General Ban Ki-Moon, at the 19th August 2010 General Assembly, defined the consequences of this event as a global disaster.

The meteorological conditions that led to the 2010 Pakistan flooding were rather special when compared to the standard summer monsoon season (Webster et al. 2011; Houze Jr et al. 2011; Rasmussen et al. 2014).

The predictability of such unusual conditions and of this specific event from planetary and large scale synoptic conditions down to the mesoscale storm structures was explored in Rasmussen et al. (2014), analyzing ECMWF (European Centre for Medium-Range Weather Forecasts) ensemble forecasts: the synoptic pattern largely responsible for the conditions that generated the Pakistan flooding event in 2010 could be predicted over a week in advance with significant confidence (as stated also by the study of Webster et al. (2011)). However, the complex topography of the region also played a significant role in the mesoscale development of the event and in determining the detailed rainfall distribution over the area (Rasmussen et al. 2014). In particular, the presence of the Hindu-Kush Karakoram Himalaya (HKKH) range is a potential source of severe uncertainty in numerical simulations and forecasts, and cannot be properly captured by coarse grid spacing General Circulation Models (GCMs). ECMWF and GFS products are available at grid spacing between 0.5°and 0.75°and, even if the precipitation forecast was predictable with reasonably good skills, convective features of the event and orographic characteristics act on scales finer than the GCM pixel resolution and could not be appreciated. Different available forecast and remote observation products reproduced daily rainfall estimate on July 2010 flood, strongly influenced by their resolutions in capturing the magnitude and the features of precipitation.

In such a complex topography, areas separated by a relatively limited horizontal distance may exhibit a large variability of the spatial-temporal rainfall properties, which are affected by steepness, altitude, temperature and small-scale orographic characteristics (Anders et al. 2006). Then a higher level of detail is needed to describe the small scale features of the event. For this purpose, is used the WRF model, operated at 3.5 km, to investigate the ability to represent the small scale atmospheric processes responsible for the event and we have focused part of the analysis on the vertical structure. The non-hydrostatic characteristic of the WRF model permits to calculate the vertical accelerations and motion explicitly, without determining them diagnostically from the horizontal divergence (as hydrostatic GCMs does). The use of a non-hydrostatic approach permits to obtain simulations at higher spatial and temporal resolutions and it is generally applied when the scale of the phenomena is similar to the height scale, such as mesoscale and convective storms. In addition to that, higher resolution results also in more finely resolved orographic features of the simulations. In the evolution of the dynamic of the model, differences in terms of circulation, due to more finely or coarsely resolved orography, play an important role: in particular, the presence of valleys and ridges results in different local circulations (see also Yu and Teixeira (2014), Flesch and Reuter (2012), Jung et al. (2012), etc.). Even if the predictability of the event from the large scales was demonstrated by Webster et al. (2011), small differences in the local circulation and interaction with the orographic features of the region could produce different results at the mesoscale. Sub-grid-scale parameterizations and initial conditions can play different roles in determining the predictability of the event at different scales.

In a recent paper, Ushiyama et al. (2014) discussed forecasts of the 2010 Pakistan flood event provided by the WRF (Weather Research and Forecasting) model at 5 km resolution (KF cumulus parameterization and WRF single moment 3-class microphysics) forced by the NCEP-GFS (Global Forecast System by the National Centers for Environmental Prediction). They show that the dynamically downscaled forecasts predicted reliable amounts of rainfall in the Kabul River basin one day ahead of the rainfall onset, and predicted a high probability of heavy rainfall three days ahead. In this work we adopt a finer and cloud permitting grid spacing (3.5 km versus 5 km) in the innermost WRF domain, which is 7 times wider (3807 km x 2643 km vs 1245 km 1125 km) than the one used by Ushiyama et al. (2014). This allows us to better capture, also at cloud permitting resolution, the interaction between the mesoscale circulation and the synoptic situation, over the considered complex topography area, whose role has been crucial for the spatio-temporal evolution of this case study. To do that, a finer vertical grid spacing (42 vs 28 vertical levels) is needed to capture more accurately the topographic role on the the spatio-temporal evolution of this case study. Two different microphysics schemes (WSM6 and Thompson) instead of only one (WSM3) have been adopted, as well three different convection parameterization approaches (Kain-Fritsch, Betts-Miller Janjic, and explicit) versus one (Kain-Fritsch).

Again, using the WRF model, Ullah and Shouting (2013) showed that a high mid-tropospheric potential vorticity anomaly led to the development of a strong mesoscale convective vortex and to large scale cyclonic circulation over Pakistan during the summer monsoon of 2010. The symmetric instability consequent to the negative moist potential vorticity anomaly significantly enhanced the vertical ascending and precipitation in the convective area (Ullah and Shouting 2013). In such applications, however, the details of the parameterizations, boundary and initial conditions adopted in the mesoscale model play a crucial role, and the sensitivity of the results to these factors need to be addressed carefully and better understood, especially in the case of such a high impact weather event (HIWE) over an extremely complex topography area. To address these issues, we analyze the role of different convection and microphysics parameterizations, and we investigate the sensitivity to the choice of the initial conditions of WRF simulations performed at cloud-permitting resolution for the most intense days of the 2010 Pakistan flood (July 26th - 31st, 2010).

The outputs of the model in terms of daily rainfall are compared with estimates provided by the TRMM satellite (Kummerow et al. 1998) and by raingauge stations. We also investigate the vertical structure of the atmosphere by means of CloudSat observations, comparing them with the WRF simulations using the DS3 (Distributed Simulation and Stimulation System) simulator (Tanelli et al. 2002) included in the NEOS^3 [NASA Earth Observing System Simulation Suite, Tanelli et al. (2012)]. In the analysis presented here and discussed hereafter, the test case of the 2010 Pakistan flood can be considered as an HIWE case study where the ability of numerical weather models is seriously challenged.

2. Event Overview

In early July 2010, a strong ridge of high pressure began to develop near the Ural Mountains in Russia, creating an "Omega" shaped blocking pattern over Europe throughout all western Russia that lasted for at least two months. This high pressure center created an abnormally active jet stream riding around the perimeter of the blocking into western Pakistan, acting as a carrier of hot and moist air and creating a "supercharged monsoon" associated with unstable atmospheric conditions (Hong et al. 2011). The interaction between strong tropical monsoon surges and extratropical disturbances downstream of the blocking became crucial in triggering the flood (Hong et al. 2011). In normal monsoonal events, the low-level moisture flow originates predominantly from the Bay of Bengal, with smaller contributions from the Arabian Sea (Houze Jr et al. 2011). In this case, however, the low-level anomaly in the moisture flux introduced by the indirect contribution of a La Niña phase in south and southeast Asia, had a strong effect in weakening the eastward moisture transport and in helping to enhance the moisture transport and convergence in the northern Arabian Sea and Pakistan (Hong et al. 2011). In figure 1, and figure 2 panels a5, b5 and c5 show large scale fields of geopotential, temperature and specific humidity of the ERA-Interim reanalysis (Dee et al. 2011a) at 500 hPa, in comparison with the results obtained in the WRF runs (discussed more deeply in section 5). This situation represented by the reanalysis resulted in an unusual displacement of the heavy monsoonal stratiform precipitation patterns, which are typical for the wetlands in northeastern India and Bangladesh, towards the arid mountainous region of northern Pakistan. This anomalous flow extended also to lower levels, carrying moisture towards the Himalayan barrier and leading to a favourable environment for the mesoscale rain systems (Hong et al. 2011). The European blocking acted on the persistence of this event. Moist air was blocked inside a mountain region of usually dry air, leading to the anticipation of saturation conditions. This caused a less convective vertical growing of the cells and a more stratiform horizontal extension due to upslope flow, respect to what happens in normal monsoonal events in that mountain region (Houze Jr et al. 2011).

The most consistent heavy rainfall event occurred in late July, from 27th to 30th. Galarneau Jr et al. (2012) gives a good description of the developing of convection, analyzing Meteosat-7 and TRMM images. From late July 27th to 06 UTC of July 28th an intense convective event with evidence of possible widespread stratiform precipitation started to interest southwest Pakistan (Houze Jr et al. 2011; Galarneau Jr et al. 2012). Then the rainfall moved towards the high-mountain region in northern Pakistan and persised over the same region for nearly 24 hours from 12 UTC of July 28th to 12 UTC of July 29th, with a continuous redeveloping of convection. The extremely moist environment increased precipitation efficiency and mitigated the cold pool development that could propagate the convection away from mountains. Finally, on July 30th, only light rain persisted over northern Pakistan area and the highest precipitation shifted over west-central India (Galarneau Jr et al. 2012).

In conclusion, the event was characterized by a close interaction between larger and smaller scales and by a strong orographic component (Rasmussen et al. 2014).

3. Experimental set up

a. The WRF model

The WRF model (version 3.3.1) is a fully compressible, non-hydrostatic, scalar variable-conserving mesoscale model (Skamarock et al. 2008).

The runs analyzed in this study are performed using two domains (Fig. 3): an external domain (d01), extending in the range 2.59°N-55.52°N and 50.69°W-96.11°W, resolved at 14 km, and an internal domain (d02), extending in the range 10.12°N-49.84°N and 57.08°W-90.02°W, resolved at 3.5 km grid spacing. The grid spacing adopted for the innermost domain already belongs to the so-called cloud-permitting range and represents a good compromise between computational performances and capability of representing the key details of the complex topography of the HKKH range, so important for the case study under examination. This choice is improving what has been done in literature till now: to provide some examples of the state of the art, Ushiyama et al. (2014), Ahasan and Khan (2013), Ullah and Shouting (2013) indeed adopted similar small domain grid spacing (respectively 5 km, 3km and 3 km) but on definetely smaller domains (Kabul river basin and a smaller windows in north-west Pakistan), for the same event. Also Maussion et al. (2011) performed simulations in the area at 2 km, but in the Tibetan Plateau (with less steep local orography) and again over a less extended domain.

Fig. 3 shows the two nested computational domains and the orography of the region, obtained from the ETOPO1 dataset (Amante and Eakins 2009). A two-way nesting mode is used to couple the two grids. The vertical dimension is discretized with 42 levels. The turbulent parameterization is the Yonsei University scheme (Hong et al. 2006).

The radiation scheme adopted is the rapid radiative transfer model (RRTM) scheme for longwave parameterization (Mlawer et al. 1997), and the Goddard scheme for shortwave parameterization (Chou and Suarez 1999).

In complex orography areas, the high variabilities in elevation, surface slope and aspect lead to in a strong heterogeneity in solar radiation distribution and, by consequence they affect evapotranspiration, moist and heat fluxes and soil and air temperature (Chen et al. 2013). In this experiment, the land use dataset is derived from the United States Geological Survey (USGS) 24-category data at 30 arc-second resolution and the land surface model is the 5-layer thermal diffusion scheme from MM5. The experiment has been carried out in hindcast mode, with boundary and initial conditions provided by ERA-Interim reanalysis fields at the native resolution (0.75°) (Dee et al. 2011b) representing the latest global reanalysis produced by ECMWF.

b. Microphysical schemes and convective closures

The joint action of the complex topography (due to the presence of the Tibetan plateau and the HKKH range) and of the climatic features of a monsoon-influenced environment make the choice of the convective and mycrophysics parameterizations difficult (Sardar et al. 2012).

For the convective closure schemes, the choice of a 3.5 km horizontal resolution allows to explicitly resolve (albeit crudely) convective processes (Kain et al. 2006, 2008). A number of studies investigated numerical simulations in the so-called "grey zone" of spatial resolution, corresponding roughly to 1-5 km, to understand whether convective parameterization is needed at this resolution [e.g. Gerard (2007), Parodi and Tanelli (2010)]. Since no definite conclusion on this issue has been reached [e.g. Yu and Lee (2010)], in this study we opt for running simulations with either a parameterized (Kain and Fritsch 1990) or explicitly-resolved convection scheme in the d02 domain, while the outermost domain at 14 km adopts always parameterized convection (Kain and Fritsch 1990)). The choice of Kain-Fritsch as parameterized run is motivated by the results and recommendations of previous studies in the region (Ahasan and Khan (2013), Sardar et al. (2012)).

With regard to microphysics, the leading idea has been to compare the performances of a single-moment scheme, versus a double-moment one when modeling a severe rainfall event, over such an extremely complex topography area with a cloud-permitting grid spacing. For this reason, the single-moment WSM6 (Hong and Lim 2006) and the double-moment Thompson scheme schemes are selected.

The six-class WSM6 scheme (Hong and Lim 2006) extends the WSM5 scheme. In this scheme, a new method for representing mixed-phase particle fall speeds for snow and graupel has been introduced. The single fall speed assigned to both classes is weighted by their mixing ratios, and it is applied to both sedimentation and accretion processes (Dudhia et al. 2008).

The Thompson scheme (Thompson et al. 2008) presents a significant number of improvements in the physical processes modeling if compared to earlier single-moment approaches, and it takes advantage of results provided by more complex spectral/binned schemes that adopt look-up tables. The assumed snow size distribution depends on ice water content and temperature and it is represented as a sum of exponential and gamma distributions. Snow assumes a non-spherical shape with a bulk density that varies inversely with the diameter, as found in observations.

It is certainly true that using also the single moment WSM6 vs. the double WDM6 microphysics (Lim and Hong 2010) would have been a worth experiment to perform. However in this study we use WRF version 3.3.1 and WDM6 is a quite new entry in the microphysics parameterization portfolio, still subjected to testing and bug fixes.

In the external domain (d01, 14 km) we use the KF convective scheme and the same microphysics as in the interior domain.

4. Observational data

The orographic complexity of the region under study and the limited availability of meteorological observations in the area represent two of the main challenges in comparing model results with measured data.

The study of Palazzi et al. (2013) considered and compared different available datasets in the Hindu-Kush Karakoram Himalaya region and evaluated the capability of these observations in reproducing precipitation characteristics and trends. Andermann et al. (2011) produced also a similar study and gave an overview of gridded available precipitation datasets along the Himalaya front. These studies analyzed the differences between the available products, with similarities and discrepancies. Great caution should be used in comparing pixel values of station observations and remote sensing techniques, especially at high temporal resolution (Andermann et al. 2011), particularly when the resolution of observations is coarser than the spatial variability of rainfall. The study of Bytheway and Kummerow (2013) confirms the previous statement, investigating the uncertanties related to the TRMM 3B42 product at 3-h accumulation and 0.25 resolution. In their global study of TRMM 3B42 uncertanties over land, they conclude that differences in error characteristics are most prevalent at accumulations below 4mm/h. At accumulations higher than 10 mm/h, the uncertanties of the 3-hour product converge to values between 75% and 85%. They add that high uncertanties values are not surprising for fine temporal resolution data. At the daily scale, uncertainty estimates are grater than 100% for low intensity daily accumulations and decrease to 20% and 40% at higher daily rainfall rates (Bytheway and Kummerow 2013; Huffman 1997; Tian and Peters-Lidard 2010).

Because of the inaccessibility of mountain regions, raingauge stations are mainly located in valley floors (Fowler and Archer 2006) and, for this reason, regions above 5 km still remain poorly monitored (Palazzi et al. 2013). The available gauge observations in the area are scarce and largely biased by altitude, mainly due to technical reasons such as the difficulty to measure the snow water equivalent depth and the deflection of precipitation by winds (see for example Winiger et al. (2005); Anders et al. (2006); Barros et al. (2000)). On the other hand, remote observations provide spatially-complete coverage of precipitation estimates, but local conditions cannot be incorporated in the sensor algorithm, with potentially large errors within each point of the grid space (Andermann et al. 2011). In the work of Andermann et al. (2011) the authors stress the difficulties of TRMM-3B42, Global Satellite Mapping of Precipitation (GSMaP) and Climate Prediction CenterRainfall Estimates (CPC-RFE) to correctly describe the precipitation distribution at elevations higher than 1 km and to capture precipitation in areas of strong orographic effect. Nevertheless, in the comparison performed by Andermann et al. (2011), the TRMM 3B42 product results to have the smallest bulk error in the monsoon period. Another study by Prakash et al. (2015) has compared the real time TRMM Multisatellite Precipitation Analysis (TMPA)-3B42 and GSMaP estimates against gauge-based measures by the India Meteorological Department (IMD) at the daily scale, using 2000-2010 datasets. They found that these products are able to capture large scale spatial features of monsoon rainfall, but still have region-specific biases. Generally they found a TRMM 3B42 overestimation of 21% and a GSMaP underestimation of 22% over all India, with respect to raingauge based dataset. The largest difficulties in rainfall detection have been found in mountain regions of northeast India (Jammu and Kashmir regions) and in southern peninsular India. Even if their study is referred to Indian area, the Kasmir and are Jummu are neighbouring areas for northern Pakistan, characterized by similar features in terms of monsoon season and high topography.

Taking all this into account, the recommended approach in handling these datasets is a multi-sensor strategy where a collection of information is carefully evaluated, considering the uncertainties of each single dataset (Palazzi et al. 2013).

Gridded daily rainfall datasets are available from different remote sensing products (e.g. TRMM, GSMaP, etc.). Additionally we have also considered the new PERSIANN CDR dataset (for more information on this dataset the reader is referred to Ashouri et al. (2014)). The precipitation information provided by TRMM, GSMaP and PERSIANN estimates are coherent among each other and provides an encouraging signal on the quality of the satellite estimates available for this specific event.

The vertical structure of the atmosphere has been measured by the TRMM PR 2A25 overpasses and by the CloudSat product, with different times of passing (thus making not easy and immediate their comparison and joint analysis). The TRMM PR 2A25 tracks cut the study area in the south, in a region with only light precipitation; the CloudSat track, at the contrary, passes directly over the main system of interest.

We also have considered raingauge interpolated maps, to provide a source of ground based measurements, instead of only remote sensed estimates, in the daily rainfall comparison.

In this work we rely mainly on remotely-sensed data from TRMM 3B42 and on raingauge interpolated maps as quantitative precipitation estimate (QPE) data sources, while CloudSat data are used for vertical cross-sections.

A quick description of the observational datasets used in the paper is given in the following paragraphs.

a. TRMM

The TRMM 3B42 rain products are used as the main QPE source. The purpose of the 3B42 algorithm is to produce TRMM-adjusted merged-infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. The final gridded estimates have a daily temporal resolution and a 0.25°by 0.25°spatial resolution. Spatial coverage extends from 50° S to 50° N. Although the dataset of the TRMM 3B42 product has a 3-hourly temporal resolution, at the finer temporal scales the incidence of sampling errors can be large. For this reason, in our analysis we consider only daily cumulates. A more accurate description of the TRMM mission is given by Kummerow et al. (1998) and by the official NASA product site (`http://disc.sci.gsfc.nasa.gov/TRMM`).

To compare WRF and TRMM daily rainfall fields we have, at first, linearly interpolated WRF fields on a grid finer than the target one (1 km grid spacing), and then we have aggregated the pixels at the TRMM 0.25 resolution (see also Herrera et al. (2015)). We did this transformation in order to conserve the area between the two different grids (the WRF curvilinear and the TRMM linear grid). The fields obtained are focused on a geographic window centered on northern Pakistan (23°N to 40°N in latitude and 66°E to 78°E in longitude). This study area was characterized by heavy precipitation on July 28th and 29th.

The quantitative comparison between WRF and TRMM is computed using statistical scores derived both from the traditional calculation of percentiles (60th and 95th), root mean square error (RMSE), mean bias (MB), and from the Method for Object-based Diagnostic Evaluation (MODE). This latter method was developed at Research Application Laboratory NCAR/Boulder (USA) and intends to reproduce an human analysts evaluation of the forecast performance. In many cases the traditional scores penalize the performance of forecasts without identifying the cause of the poor performance. An object-based analysis becomes particularly relevant when the model is pushed towards high-resolution and the localization and the episodic characteristics of rain became more important in the verification process. The MODE analysis is performed using a multi-step automated process. A convolution filter and a threshold specified by parameters *r* and *t* are applied to the raw field to identify the objects. When the objects are identified, some attributes regarding geometrical features of the objects (such as location, size, aspect ratio and complexity) and precipitation intensity (percentiles, etc.) are measured. These attributes are used to merge objects within the same forecast/observation field, to match forecast and observed objects and to summarize the performance of the forecast by attribute comparison. Finally, the interest value combines in a total interest function all the attributes computed in the object analysis (as shown in Brown et al. (2007), equation 1)), providing an indicator of the overall performance of matching and merging between different observed and simulated objects. In our experiment we have empirically chosen the convolution disk radius and convolution threshold, so that this choice would recognize precipitation areas similar to what a human would identify. After a set of experiments, we fixed the value of the convolution radius to three grid points and the threshold of the convoluted field to 35 mm/day. More information about the MODE technique can be found in Davis et al. (2006a), Davis et al. (2006b) and Brown et al. (2007).

b. Raingauge stations

Raingauge stations data have been considered as an additional term of comparison for daily rainfall estimates. A set of 98 stations from the Pakistan Meteorological Department (PMD) monitoring network was collected and linearly interpolated over the focus area. Moreover, we have selected 90 stations that fall inside the geographic window of interest, we have compared the gauge measures with the nearest neighbour WRF grid point of the map comparison and we have calculated the associated MB and RMSE. The MB and RMSE calculated comparing with the raingauge dataset are obviously not comparable to the same statistics compared to the TRMM dataset. The raingauge evaluation is computed based on 90 grid points, while the MB and RMSE computed based on TRMM estimates represent a pixel comparison extended to all grid points in the geographic window. Additionally, the two products (raingauges and satellite products) are differently accumulated. The daily rainfall station data are accumulated from 03 UTC for the next 24 hours, so, great caution should be used when comparing them to TRMM data because a 3h offset has to be considered. Finally, the comparison is strongly influenced by the different nature of ground and satellite instruments and by their different weaknesses and strengths in measuring precipitation in areas with complex orography. Nevertheless, in an area of scarce observations, they provide an additional point for the discussion.

c. CloudSat

The CloudSat satellite mission was designed by NASA to measure the vertical structure of clouds from space and to improve global knowledge of cloud abundance, distribution, structure, and radiative properties. The CloudSat instrument was launched in April 2006, as a part of the A-Train satellite constellation. The Cloud Profiling Radar (CPR) installed on CloudSat is a millimeter-wavelength cloud radar that allows detection of cloud droplets and ice particles forming the cloud masses. The CPR operates at 94 GHz, which represents the best compromise between performance and spacecraft resources, to achieve sufficient cloud detection sensitivity (Tanelli et al. 2008). The data are given to the 2B-GEOPROF product, whose algorithm identifies those levels in the vertical column that contain significant radar echo from hydrometeors and provides an estimate of the radar reflectivity factor for each of these volumes. The CPR provides detailed information on the vertical structure of cloud systems and it represents a relevant source of information for the evaluation of climate and weather prediction models (for more information, see http://cloudsat.atmos.colostate.edu and Stephens et al. (2008)).

To compare model outputs with satellite estimates, it is necessary to have a simulator converting model quantities into equivalent radar reflectivities. The effects of instrumental sensitivity and attenuation by clouds and precipitation have also to be taken into account (Bony et al. 2009; Haynes et al. 2007). For this reason, the NASA Earth Observing System Simulators Suite ($NEOS^3$) includes the DS3 simulator (Tanelli et al. 2002), that provides forward simulation to evaluate cloud radar and other remote sensing products (Tanelli et al. 2011, 2012). Using this tool, the WRF outputs are compared to CloudSat observations considering the two available satellite tracks over northern Pakistan during the days of the event: granule 22608, recorded on July 28th around 21:00 UTC, and granule 22615, recorded on July 29th around 08:00 UTC. The CloudSat observation tracks are provided in Fig. 4 and Fig. 5 (blue lines). Since the granule 22615 of July 29 th misses the main observed precipitation core (see Fig. 4 panel b6 or Fig. 5 panel b7), the results in section 5 are discussed only for granule 22608.

5. Sensitivity experiments

a. Sensitivity to the convective and microphysical schemes

The four different configurations tested in this work are listed in Table 1. Figure 4 shows the precipitation fields produced by the WRF model for the different parameterization choices, compared with the TRMM estimates and raingauge observations, in experiments initialized on July 26th at 00 UTC.

When looking at Exp-WSM6 vs. KF-WSM6 (Fig. 4a1 vs. Fig. 4a2), we see that the KF scheme produces more precipitation and more organized patterns. This is also true for Exp-Thompson vs.

KF-Thompson (Fig. 4a3 vs. Fig. 4a4). Therefore, in general, it appears that the KF scheme tends to overestimate precipitation and to produce more organized rainfall patterns for our case.

The statistical evaluation computed for our experiment using traditional statistic and MODE verification analysis is reported in Table 2. The MODE values considered refer to the higher intensity object identified by the verification technique that matches with a corresponding object in TRMM. The white countours in Figure 4 represent the MODE objects. The percentile values indicates that all four configurations tested tend to overestimate the rainfall amount compared to the TRMM estimates, especially for 60th percentile on July 28th and 95th percentile on July 29th. The pecentile values confirm the tendency of the KF simulations to overestimate TRMM estimates. On July 28th all the values of the Exp-WSM6 configuration indicate good accordance with TRMM values. The rainfall intensity given by the percentiles and the localization of the object corresponding to the main precipitation core seem to be best represented by Exp-WSM6. On July 29th, on the contrary, the evaluation doesn't seems univocal: MODE statistical indicators have good agreement with TRMM in terms of total interest and geometric attributes of localization (centroid distance and area ratio) for Thompson microphysic configurations (Exp-Thom and KF-Thom); at the contrary MB and RMSE result the best for Exp-WSM6. All values on July 28th and (especially) on July 29th indicate that the worst results are seemingly obtained using the KF-WSM6 configuration, where the main precipitation core is misplaced and overestimated. As a word of caution, however, we note that the differences in score between the different configurations are not very large, and the highly fragmented appearance of the precipitation fields obtained with explicit convection does not match entirely the TRMM data. In addition to that, with equal convective scheme, the Thompson microphysics presents higher 95th percentile values. If we examine the results of the statistics calculated in comparison with the raingauge datasets (Table 3, fourth part), also the MB estimates on July 29th confirm the tendency of the Thompson microphysics to produce higher than observed rainfall amounts (even if closer to measured values than the other simulations). The WSM6 has been found to produce larger values of evaporation rate over the entire atmospheric column in Bryan and Morrison (2012) and in Morrison et al. (2015) with reference to highly-idealized settings with no orography, possibly explaining its reduced precipitation compared to the Thompson scheme. The raingauge statistics produce less underestimation (meaning higher precipitation values) for KF configurations on July 28th and the best RMSE for Ex-WSM6 on July 29th.

The analysis of cumulative distributions permits to understand the variability of the precipitation field and, in particular, the tail of the distribution gives an important information about the probability of exceedace of the highest values of the precipitation field. Figure 6 confirms that the Exp-WSM6 simulation produces results which are closer to the statistics of the TRMM estimates. While all the other schemes tend to overestimate the probability of extreme precipitation compared to TRMM, on July 28th the Exp-WSM6 configuration generates distributions which are fairly close to the observations. In this case, the main cause of discrepancy with TRMM (reflected in the statistical scores) is due to a misplacement of the precipitation structures, while intensity and frequency are properly reproduced. On July 29th, all schemes tend to significantly overestimate the observed precipitation. In Figure 4 panels a6 and b6 provides another term of comparison with raingauge ground measurements. Even if, as discussed above, raingauge station present several limitations, the QPE provided by raingauges is higher than TRMM estimates, providing support for the hypothesis of TRMM underestimation instead of a WRF model overestimation. Again, the Exp-WSM6 configuration is the closest to raingauge observations.

b. Sensitivity to the initialization date

Based on the above results, the configuration with explicit convection and the WSM6 microphysical scheme is selected for further sensitivity analysis. Using this configuration, we perform forecast experiments considering three different initialization days: July 24th (J24), July 26th (J26) and July 28th (J28), all at 00 UTC (Table 3). The different initialization experiments have been chosen considering initialization from 1 to 4 days in advance, every 48 hours. We choose this range, as a good compromise between possible required spin up time of the model and expected model time integration reliability.

The meteorological analysis is performed starting from the large scales (Fig. 1 and Fig. 2), down to the mesoscale fields for all the different simulations performed, in order to understand the inter-scale interplay of the phenomena. At larger scales, variables such as geopotential, temperature, water vapor mixing ratio are interpolated on the vertical 500 hPa isobar level and compared with the ERA-Interim reanalyses. The synoptic features that led to the severe 2010 events on Pakistan are well reproduced by the model for all its different initializations, if compared with the reanalysis. The geopotential underlines the presence of a strong high pressure block in the northern part of the domain. This blocking high, associated with the presence of smaller scale troughs in the Arabian sea and in the south of Afghanistan, led to the formation of a strong wind blowing from the Arabian sea to the northern part of Pakistan. From the examination of water vapor fields, there is a moisture transport associated to the south-westerly winds that brought a high water vapor quantity up to northern Pakistan. Another source of vapor is given by the moisture flux approaching from the Bay of Bengal. The moisture flux convergence supports the accumulation of moisture during the two days in which the maximum precipitation occurs (July 28th and July 29th). The different model runs exhibit similar large scale circulation, with small differences between the different simulations. Differences start to emerge when we look at the smaller scales, in which the role of orography (valley and ridges) starts to emerge because of the interaction with the small scale circulation. In this case small differences in moisture transport or in wind circulation reflect deeply the different distribution of the resulting precipitation fields.

Figure 5 shows the daily precipitation maps for the crucial days of the event (July 28th and 29th), reporting also TRMM observations and interpolated observations from the available raingauge stations. The J24 run simulates rather well the actual rainfall amounts until July 27th, and then downgrades as the event develops further. The J26 run offers a good performance even though the simulation is not able to correctly reproduce the patterns of the first days (July 26th and 27th - not shown). For July 29th, the J26 run (Fig. 5, panel b2) captures well the main rainfall core, even if it is more widespread and displaced slightly eastward with respect to the observations. The J28 run reproduces well the precipitation pattern on July 28th. On the following day, the J28 run (Fig. 5, panel b3) displays a very poor performance, especially when the rainfall pattern of the maximum core is considered. A strong orographic control on the QPF is evident: precipitation is confined to Pakistan plains by the local mountainous range and the most important precipitation core is completely missed. Because of the bad results of J28 for July 29th, we conclude from the map comparison that the J26 run provides a better forecast of the event.

The J28 run produces higher QPF during its first 24 hours of simulation (July 28th), and after that the precipitation rates decreases significantly: a possible explanation for this behavior is the dry-out of the atmospheric column caused by the high precipitation rates on the 28th, together with the lack of time for the moisture from the boundaries to gather in the domain in the following 24 hours. To test this latter possibility we reduce the dimension of the domain: the J28 simulation is run again in the 2-way nesting mode, but this time the original high-resolution domain d02 is downsized to the focus area (23°N to 40°N, 66°E to 78°E) (J28S run). The results obtained for July 29th with the smaller domain do not display any significant improvement, indicating that there must be other causes for the bad performance of the J28 run. In addition to that, the J28S, if compared with J28, shows no sensible dependence of WRF model on small perturbation of initial conditions over the time scale of the experiment.

Figure 7 compares the cumulative distributions of daily precipitation for the different initialization dates and for the two target forecast days. The comparison indicates that the J26 run shows a better agreement with the amplitude statistics of the TRMM data. In particular, on the 28th, all other runs (except J24) tend to overestimate the probability of exceedence of precipitation rates larger than about 100 mm/day. On the 29th, on the contrary, the runs started on the other initialization dates lead to an underestimation of precipitation over the area, even if they are still closer to the TRMM estimates.

Statistical evaluation for the different simulations are summarized in Table 4. The statistical scores partially confirm the previous analysis. The J24 simulation displays a good performance on July 28th and the worst performance on July 29th. On the second day the interest value of the MODE analysis is extremely low and the geometrical properties of the forecast-observed objects are highly unrelated. For July 28th, and July 29th J28 appears to have a good statistical evaluation, even if in the map comparison the pattern of the main precipitation core is totally missed. On the contrary of what observed in the map comparison, on July 29th the J28 run result in best values of interest and good percentile values. On the other hand, on July 29th, the 95th percentile confirms the J28 underestimation even if it is still the closest to TRMM values. If we consider that TRMM tend to underestimate in that area (as stated in the previous sections) and the information of the raingauges, we are more prone to penalize an underestimation of the model rainfall values. The TRMM tendency to underestimate, with respect to raingauges is evident from the comparison between MB related to TRMM and the one based on raingauges measures. Even if great caution should be given in the comparison, raingauge MBs are negative (meaning an undestimation of the model, with respect to the raingauges), while the MB of TRMM seems to indicate a general overestimation of the model respect to the satellite estimates. Barring that, the raingauge statistics are rather in accordance with what observed in the previous analysis. The MB and RMSE have best scores for J28 run on July 28$^{\text{th}}$. On July 29$^{\text{th}}$ the J28 has still the best RMSE evaluation, but the best MB is calculated for J26. A lower model underestimation is observed on July 29$^{\text{th}}$ for J26 simulation where the main precipitation pattern is simulated properly.

Nevertheless, the better performance of J26 in the map comparison with respect to J28 on July 29$^{\text{th}}$ is rather unexpected, as the J28 run misses the main precipitation pattern.

c. Sensitivity to initial conditions

The low QPF performances of the J28 run for the 29$^{\text{th}}$ July can be related to the role of the specific ERA-Interim initial conditions. In support of this initial conditions, the study of Ahasan and Khan (2013), which was initialized on the same day of J28, but with a NCEP reanalysis, produced a better rainfall distribution for July 29$^{\text{th}}$ (not shown). To test the sensitivity to initialization we perform a new run, initialized on 28$^{\text{th}}$ July 2010 at 00 UTC with a different set of initial conditions. Instead of using the ERA-Interim fields, we run J26 for 48 hours till July 28$^{\text{th}}$ at 00 UTC. Then all the microphysical variables deriving from the WRF dynamics (namely cloud water, rainwater, snow, cloud ice and graupel) are set equal to zero: this provides a set of initial conditions comparable with those provided by ERA-Interim (the same required by the WRF preprocessor WPS for ERA-Interim initialization). In ERA-Interim we do not have humid variables (microphysical variables), so we have tested the importance of this aspect, initializing the WRF restarted run in the same way. This new set of initial conditions is fed into the model and WRF is run for another 48 hours. In this way, we run a novel J28 experiment, initialized with the (partial) output of the

J26 run. As shown in Fig. 5 and Fig. 7, the results of the J28 restarted run (J28R) outperform the original J28 results: the main precipitation core is well modelled and none of the main precipitation structures is missed. The restarted run produces daily rainfall outputs which are similar to those of J26, providing a better estimate of the main precipitation patterns and positions. Since the only difference between J28 and J28R are the initial conditions, these result suggest that the initial conditions provided by ERA-Interim on 28^{th} July at 00 UTC are mainly responsible for the poor results provided by J28 on the 29^{th}.

To better understand the evolution of the J28 and J28R runs, we compare the surface temperature (Fig. 8) and moist transport (Fig. 9) at the initialization time (July 28^{th} at 00 UTC) and 24 hours later (July 29^{th} at 00 UTC), at the beginning of the most intense day of the event.

At each horizontal point (pixel), we define the moist transport as the vertically-integrated total moisture transport $\mathbf{F}$, [kg (m s)$^{-1}$] given by the product of the water vapor mixing ratio q [kg kg^{-1}] and the horizontal wind speed $\mathbf{V}$ [m s^{-1}]

$$\mathbf{F} = \int_{z_{Surf}}^{z_{Top}} \rho \, \mathbf{f} \, dz \qquad \text{where } \mathbf{f} = q\mathbf{V}.$$

At 00 UTC the J28R run is identical by construction to the J26 frame. Twenty-four hours later, we find that J28R and J26 present very similar precipitation, as shown in Fig. 5. Surface temperature and moisture transport fields are also very similar, so we choose not to show the J26 run in the comparison of Fig. 8 and Fig. 9, to make the comparison clear and straightforward.

The pixel-by-pixel differences for the temperature field at 2 meters between the ERA-Interim initialization (J28) and the (partial) WRF initialization (J28R), show strong temperature anomalies (Fig. 8). On July 28^{th}, the J28 field shows a positive bias of more than 3K in north-western Pakistan, near the Afghanistan border, and a negative bias eastwards (Fig. 8, panel a3). The warmer zone of the J28 run can create a stronger instability of the air masses, with a tendency to generate intense precipitation in the next 24 hours (on July 28th) and drier and colder atmospheric conditions on July 29th.

The moist transport reflects the temperature anomalies (Fig. 9). On July 28th, both initializations generated a moisture transport directed towards the orographic barrier, even if the transport of the J28 run is more concentrated (Fig. 9, panels a1 and a2). The J28R run presents broader region with large amount of moisture transport.

The separate contribution of moisture fields and wind fields to total moist transport has been investigated in terms of horizontal and vertical distributions (not shown). The major contribution of the moisture flow to total moist transport is always more evident in the south west part of the domain, for both runs and days of the event. On the contrary, a predominant role of wind is apparent in the north east part of the study area, over the mountain region. Along the vertical, the highest moist transport occurs on lower levels, with a major contribution provided by water vapor, instead of wind (which contributes more significantly on higher levels).

On July 28th, the higher temperatures and the more intense transport are responsible for larger QPF exhibited by J28 run. The day after (Fig. 9, panels a3 and a4), the J28 run has completely lost the moisture transport contribution, while transport remains high for J28R. In the J28 run there is no moist convergence on July 29th (Fig. 9, panel a3), while in J28R the moist air is pushed towards the northern Pakistan orographic barrier producing heavy rain (Fig. 9, panel a4). All these factors concurred to create a more intense rainfall spell on the July 28th and a drier environment for the following day in the J28 run.

6. Qualitative and quantitative analysis of the vertical structure

Comparison of the surface precipitation patterns against TRMM has allowed to assess the overall performance of WRF for hydrological purposes. The comparison between the simulated CloudSat and CloudSat observations provides more insight into the ability of WRF to reproduce vertical profiles of cloud structure.

On July 28th at 21:00 (granule 22608) CloudSat passed directly over the system of interest. A comparison of simulated CloudSat using various assumptions and WRF experiments is shown in Figure 10. As a reference, the CloudSat L2B-GEOPROF is provided in the top panel (Fig.10, panel a). This graphs shows the surface clutter, when it is not attenuated by heavy precipitation above it like around 33°N.

It is evident that the changes in parameterizations and initial conditions result in major differences. These need to be interpreted in light of the temporal and spatial evolution of the system. Therefore we identify three salient features at the large scale (Fig. 10, panel i) and discuss how each experiment performed in that regard. First, the region of greatest hydrological importance in this portion of this event is the wide and persistent stratiform precipitation area between 33°and 35°N, which was for the most part generated by a relatively low convective plume (minimum IR brightness temperatures observed around 230 K) and advected moisture from the SE (hereinafter STR34N). Second, consider the organized convective towers along the southern part of the line of convergence, characterized by an anvil much less developed than what observed and top heights of the large hydrometeors (marking convective cores) barely reaching above 10 km, with corresponding IR in the 190 to 200K range (hereinafter CONV30N). It is important to note that at the time of the overpass, the line of convective activity curved to the SW around 29°N along the CloudSat ground track (blue line in Fig. 4 panel a6 and Fig. 5 panel a7), and therefore all convection occurring between 27°N and 29°N is not observed by CloudSat because it was to the west of the track. Such misplacement is noted here just to address a key feature, viz. the limited representativeness of nadir curtains when interpreting Figure 10: one should not conclude that a configuration did or did not produce convection according to observations only focusing on these data. The Geostationary imagery should always be consulted when interpreting these observations to provide the context that is lacking from the nadir-only profiles. All considerations expressed hereinafter were always developed in this context. The third feature considered is the long outflow associated with STR34N over the Karakoram range and the Taklimakan desert (latitude from 33°N to 35 °N) resulting for the most part in snowfall to the surface, but with the zero isotherm in close proximity to the prevailing ground altitude of the desert.

Panel b shows that the Exp-WSM6 experiment, initialized on J24, essentially failed to generate precipitation between 33°N and 35°N, as also shown in Figure 5. The CONV30N structure was much suppressed and disorganized, however a remnant plume did produce snowfall over Karakoram, albeit with cloud top heights 3 km lower than observed. Panel c shows the product of the same configuration but initialized on J26: in this case all three elements are captured to some extent, however the stratiform region is spatially much less extensive, the convective region extends more to the north, and most importantly exhibits notably deeper towers than observed (topping at 15-16 km). This comparison confirms that this configuration, while it achieved among the best statistical scores in total precipitation patterns, doesn't necessarily capture a realistic partitioning in convective vs. stratiform precipitation.

In order to assess the sensitivity of the forward simulations to assumptions independent of the bulk-hydrometeor quantities produced by these single-moment schemes, a series of tests using the same WRF output as input to the CloudSat simulation are performed: the assumptions on particle size distribution (PSD) and mass-size (m-D) relationship for the hydrometeor species are swapped between those assumed internally in the WSM6 scheme and those assumed in the Thompson scheme, plus a third set adopted in airborne precipitation radar microphysical retrievals. In each case the entire set of micropysical assumptions was swapped, and for all of them T-Matrix calculations (Mishchenko and Travis 1998) were used to calculate the scattering properties of the hydrometeor species according to the internal assumptions within each module. Oblate spheroids were adopted for raindrops (Beard and Chuang 1987) and snowflakes (Matrosov et al. 2008), and spheres for all other particles.

One example of these tests is shown in panel d (where both the PSD and m-D assumptions of Thompson are applied to the bulk quantities generated by Exp-WSM6 J26). Visual comparison of panels c and d confirms the intuition that at the level of assessment of the general aspect of cloud and precipitation systems, the microphysical assumptions made during the radar simulations are of second-order importance compared to the microphysical assumptions made in the CRM simulations. While the microphysical assumptions at the radar simulation stage change by several dB the observed reflectivities on various portions of the profile, but such change is indeed not sufficient to alter the visual interpretation of the general aspect of the systems other than in a small minority of locations. For example the only striking difference can be noticed in the rain portion between 33N and 35N where the Thompson microphysical assumptions generate reflectivities lower than the WSM6 by more than 10 dB. This particular difference is due to the fact that for low water contents WSM6 still assumes raindrops of about 1 mm on average, while the Thompson parameterization results in drop sizes smaller than 0.5 mm (notably, this change was explicitly targeted in that module to better reproduce mid-latitude light precipitation and drizzle, Thompson et al. (2008)). Therefore at W band, although the water content is identical (because it comes from the same WRF run), the 0.5 mm particle will be in a Rayleigh scattering regime, unlike the 1 mm, which in turns explains the large differences observed in the radar returns. Overall, an investigation focusing on quantitative retrievals of precipitation must indeed account for them, and the uncertainties within, but when CloudSat data are only used to validate the structure of the observed systems assumptions on PSD and scattering models, they become of secondary importance. These tests – performed on each one of the WRF experiments – served to eliminate one possible source of ambiguity in the interpretation of the simulated results. Along the same lines, we note that the DS3 simulator has a relatively basic representation of multiple scattering effects, particularly when compared to the advanced simulator DOMUS (Battaglia and Tanelli (2011)), which is included in NEOS[3]. Nonetheless, it was found that the DS3 simulations yielded a more direct interpretation in regards to the nature of the problem. Absence of multiple-scattering effects is for example evident in the deep convective storm modeled at 29N in this simulation where the single-scattering signal is completely attenuated instead of showing the typical stretched echo of multiple scattering all the way to the surface and beyond (see Battaglia et al. (2010) for a comprehensive review on multiple-scattering).

Panel f shows the product of the KF-WSM6 J26 experiment. Despite identical synoptic conditions and microphysical parameterization, this experiment generates much more developed anvils around CONV30N. However it fails to capture the stratiform region of greatest interest (STR34N). The low statistical scores quantify the fact that this experiment overestimated precipitation in CONV30N and underestimated it in STR34N. Panels g and h show the products of KF-Thompson J26 with Thompson and WSM6 assumptions in the radar simulations, respectively. The higher propensity of this microphysical parameterization to produce anvils and resulting stratiform rain is manifest in both cases. The Thompson scheme, unlike a simple single moment scheme, explicitly predicts the mixing ratio and the number concentration of cloud ice (Thompson et al. 2008). In this scheme the rain size distribution significantly shifts depending on whether the rain appears to originate from melted ice versus rain produced by collision/coalescence (warm rain). As evident from Fig. 10 (panels g,h,i), the largest reflectivity factors are usually observed above the line of melting level and the volume above this level is significantly enhanced in the Thompson scheme simulations. Consequently, it generates convection even deeper than WSM6, and produces wider anvils. The latter aspect is more in line with observations, but combined with the former it results in an overestimation with respect to TRMM products (Fig. 6). Comparison to the CloudSat reflectivities in the rain portion shows much smaller values in the model than in the observation: this is likely due to the aforementioned assumption of small raindrops in Thompson and the absence of significant multiple-scattering contribution in the simulation. Small drops result in unattenuated reflectivities that are possibly biased low, and if the water contents are overestimated the specific attenuation can be larger than observations (it is almost independent on drop size), these two factors, combined with the absence of multiple-scattering stretched echo generated in the ice region above, provide a framework to explain this particular difference. Notably the model runs used to generate these simulated CloudSat products apparently extend the region with precipitation more southward than observations. This is because the line of convergence mentioned before did not bend SW at 29 °N as in reality, once again reflecting the great importance of the choice of microphysical parameterizations not only in the resulting storm structures, but also in the large scale patterns.

Panel j shows the product of Exp-WSM6 J28. In this experiment the entire set of features is moved northward, the region of highest accumulation on July 28$^{\text{th}}$ is captured better than the other cases, but not because of an improved skill in capturing the nature of the process (which is entirely convective at this time with no significant anvil).

Finally, a perturbation analysis was applied to the simulations to examine the importance of shifting the simulated track relative the core of the precipitation and the importance of the timing of the WRF run. This is particularly relevant when studying convection, for which location and timing of occurrence are fundamental. To this end, we looked at the satellite simulated overpasses shifted by 0.8 °to the East/West with respect to the center of the main precipitation core and using hourly WRF runs in a range of plus and minus three hours about 21 UTC (time range comparable with the time evolution of these cloud structures). To serve as a reference, the actual CloudSat overpass was also overlaid on a map of IR temperature from Geostationary satellites. Perturbing the simulated tracks did not reveal any significant improvement and if anything led to sometimes missing the main core of the precipitation. For this reason, we focused on the nominal CloudSat track. Regarding the timing analysis, we have focused our investigations on J26 and J28 (Exp-WSM6) simulations. For each of these hourly simulations, the simulated radar products were compared to the CloudSat measurements in terms of their contour-frequency-by-altitude diagram (CFAD). Namely, we considered the a vector consisting of the vertical profiles of the 10%, 50% and 90% quantiles of the simulated CFADs and compared them to those of the CloudSat data. For instance, the correlation coefficient between the CFAD of CloudSat measurements and that of the simulated results at 21UTC is equal to 85% for J26 (Fig.10 panel c) and 93% for J28 (Fig.10 panel j). The strongest correlations to the measurements are observed with the WRF products at 23 UTC for both J26 and J28, with correlation coefficients of 96% in both cases. The corresponding resulting radar cross sections are depicted in Fig.10 (panels e and k) for the J26 ad J28 runs are 23 UTC. For the J26 case (Panel e), a comparison with the results at 21 UTC (panel c) shows that the clouds and precipitation have moved to the North, as evidenced by the convective cell around 30°N in the CONV30N region. Furthermore, similarly to the CloudSat measurements, the top of the cells is lower at 23 UTC than at 21 UTC, which explains the slightly larger correlation between CFADs. Nonetheless, there is still a strong resemblance between the features at both instants, e.g. in the STR34N region where the precipitation in still disorganized. Similar observations can be made for the Exp-WSM6 case initialized on J28 (panels j and k) where results at both times capture the persistent precipitation in the mountains (north of 35°N). One can note also the lesser impact of attenuation on the measurements at 23 UTC in the STR34N region (around 34°N) owing to the lower levels of the clouds at that time.

7. Summary and Conclusions

In this paper we have performed WRF non-hydrostatic simulations at 3.5 km of the HIWE that led to the Pakistan flood in July 2010. We have tested the ability of the modelling system to reproduce the observed precipitation rates and patterns, and we have analyzed the model sensitivity to different microphysics and convection parameterizations and different initializations.

Explicit convection and the WMS6 microphysical scheme turned out to provide a better match in terms of rainfall amount, patterns and localization when compared to other choices.

Using this configuration, we varied the initialization day to determine the dependence of the model results on the choice of initial and boundary conditions. Even though model outputs are usually more reliable in the first days of the simulation, the J28 run (initalized on July 28th) performed poorly on July 29th, especially when compared to a run initialized on July 26th. This uncommon behavior motivated an additional set of experiments. A new model run (J28R) was initialized on July 28th with the inputs provided by a WRF simulation started on July 26th, with all variables related to clouds and vertical velocities set to zero to be consistent with a standard large scale inizialization. This novel run outperformed the original J28 run initialized with ERA-Interim fields, both in terms of rainfall localization and patterns, as well as of daily accumulation, indicating that the initial conditions are a crucial factor in order to obtain a satisfying representation of the event.

The joint use of CloudSat observations and simulated cloud radar profiles allowed to investigate further the skill of each experiment in capturing the most important aspects of the observed vertical structure of this event. In this regard, the Thompson microphysics produces more stratiform precipitation and more organized precipitation patterns than the WSM6, in line with the observations. Both microphysical parameterizations produce convective activity deeper and more intense than observed. Since Thompson also produces more extensive widespread precipitation from the outflow, it results in an overestimation of the total precipitation. The striking differences in cloud structure resulting from the different microphysical and cumulus parameterizations, even when the same synoptic conditions are adopted, reinforce the assessment that performance of models in reproducing QPE estimated from observations cannot be limited to a few exercises with different models, resolutions or initial conditions. Notably, the principal differences resulting from the adoption of different parameterizations within a particular model (in this case WRF) are consequence of their resulting macroscopic distributions of the bulk quantities of the various hydrometeors and of the different latent heating profiles and they can radically change the final output of the model given equal initial conditions and resolutions.

Overall, we found that the simulation results are affected more significantly by the choice of the initialization day than by the parameterization schemes adopted. As expected, the largest errors are located near Himalayas and northern Pakistan, where the steep local orography affected the numerical integration.

All the study has dealt with the presence of the highest mountain topography of the world and the experiment of going to 3.5 kilometres resolution with a non-hydrostatic model has represented an instrument to understand the physical processes responsible of the tragic event. In particular we have found that ICs and BCs are a prominent factor affecting the results and that small variations in local atmospheric dynamics can produce very different results in complex orography areas. This study has investigated the event at different spatial and temporal scales, starting from the large scales, down to the mesoscale fields (section 5b and 5c) and vertical sections (section 6). The synoptic features of the different initializations in terms of geopotential, temperature and water vapor mixing ratio are pretty similar for all the runs and the WRF successfully reproduces the main large scale features responsible of the event. Moreover, the model, as expected, strongly reflects the large scale characteristics inherited by the coarsely resolved GCM. The highest differences are evident when the model is challenged to reproduce the smaller scale features. The different pattern results obtained for J26 and J28 run are a manifestation of this: the presence of a valley or of a ridge is capable of strongly influencing the simulation, producing different moisture transport and wind circulation that affect the resulting precipitation fields. As stated in Webster et al. (2011), the predictability of this event was evident from large scale models, but we agree with Rasmussen et al. (2014) that conclude that an higher degree of detail is needed to understand the anomalous convective features that led to the tragic flooding.

This work focuses on a specific extreme event, viz. the 2010 Pakistan flood, studied using the WRF model in cloud permitting mode and operated at 3.5 km in order to gain insight on the predictability of this flood event. While in general it can be difficult to make solid conclusions on the choice of any one or the other microphysics from individual case studies, nonetheless our results allow to draw some more general conclusions. In particular, they suggest that a careful choice of parameterization schemes and initialization day must always be adopted, because these factors can affect significantly the simulation. Configurations that at the large scale exhibit small differences, at the small scale start to produce very different precipitation amounts, patterns and circulations, especially over mountain terrain. The results presented here indicate that the reliability of the large scale fields used for initialization and boundary conditions remains an essential ingredient of the simulation, and that errors in the large scale fields can be propagated, or even amplified, in the outputs of high-resolution simulations. For all these reasons, we recommend a dual selection of both initial and boundary conditions and parameterization assumptions to propagate the model through this kind of events in complex topography areas, rather than an independent analysis of one or another. Inter-scales phenomena and orography interaction are thus predominant features in studying these particular processes over complex orography areas such HKKH.

This study intends to contribute to future studies in that area, and it highlights the complexity of studying an HIWE case study in a geographical area in which the ability of numerical weather models is seriously challenged.

Acknowledgments. This work was supported by the PNR 2011-2013 Project of Interest "Next-Data" and by the 2010-2011 PRIN Project "Innovative methods for water resources management and risk assessment under uncertainty" of the Italian Ministry for Education, University and Research.

The work of Dr. Tanelli and Dr. Sy was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This work was carried out in support to the CloudSat and GPM missions.

The authors are also grateful to the Pakistan Meteorological Department for the provision of the daily rainfall raingauges data, as well as to the Editor and the four anonymous reviewers for their constructive comments on the early version of the manuscript.

The numerical simulations were performed on the SuperMUC Petascale System of the LRZ Supercomputing Centre, Garching, Germany. Project-ID: pr45de.

References

Ahasan, M., and A. Khan, 2013: Simulation of a flood producing rainfall event of 29 July 2010 over north-west Pakistan using WRF-ARW model. *Nat. Hazards*, **69 (1)**, 351–363, doi:10.1007/s11069-013-0719-6.

Amante, C., and B. Eakins, 2009: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. *NOAA Technical Memorandum NESDIS NGDC-24*, **19**, doi:10.7289/V5C8276M.

Andermann, C., S. Bonnet, and R. Gloaguen, 2011: Evaluation of precipitation data sets along the himalayan front. *Geochemistry, Geophysics, Geosystems*, **12 (7)**.

Anders, A. M., G. H. Roe, B. Hallet, D. R. Montgomery, N. J. Finnegan, and J. Putkonen, 2006: Spatial patterns of precipitation and topography in the himalaya. *Geological Society of America Special Papers*, **398**, 39–53.

Ashouri, H., K.-L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat, 2014: Persiann-cdr: Daily precipitation climate data record from multi-satellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.*

Barros, A., M. Joshi, J. Putkonen, and D. Burbank, 2000: A study of the 1999 monsoon rainfall in a mountainous region in central Nepal using TRMM products and rain gauge observations. *Geophysical Research Letters*, **27 (22)**, 3683–3686, doi:10.1029/2000GL011827.

Battaglia, A., S. Tanelli, S. Kobayashi, D. Zrnic, R. J. Hogan, and C. Simmer, 2010: Multiple-scattering in radar systems: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **111 (6)**, 917–947.

Beard, K. V., and C. Chuang, 1987: A new model for the equilibrium shape of raindrops. *Journal of the Atmospheric sciences*, **44 (11)**, 1509–1524.

Bony, S., M. Webb, B. Stevens, C. Bretherton, S. Klein, and G. Tselioudis, 2009: The cloud feedback model intercomparison project: summary of activities and recommendations for advancing assessments of cloud-climate feedbacks. Availble online at: http://cfmip.metoffice.com/CFMIP2_experiments_March20th2009.pdf.

Brown, B. G., R. Bullock, J. H. Gotway, D. Ahijevych, C. Davis, E. Gilleland, and L. Holland, 2007: Application of the mode object-based verification tool for the evaluation of model precipitation fields. *AMS 22nd conference on weather analysis and forecasting and 18th conference on numerical weather prediction*, Vol. 25, 29.

Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140 (1)**, 202–225.

Bytheway, J. L., and C. D. Kummerow, 2013: Inferring the uncertainty of satellite precipitation estimates in data-sparse regions over land. *J. Geophys. Res.: Atmospheres*, **118 (17)**, 9524–9533.

Chen, X., Z. Su, Y. Ma, K. Yang, and B. Wang, 2013: Estimation of surface energy fluxes under complex terrain of Mt. Qomolangma over the Tibetan Plateau. *Hydrology and Earth System Sciences*, **17 (4)**, 1607–1618, doi:10.5194/hess-17-1607-2013.

Chou, M.-D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. *NASA Tech. Memo*, **104606**, 40.

Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. part i: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134 (7)**, 1772–1784.

Davis, C., B. Brown, and R. Bullock, 2006b: Object-based verification of precipitation forecasts. part ii: Application to convective rain systems. *Mon. Wea. Rev.*, **134 (7)**, 1785–1795.

Dee, D., and Coauthors, 2011a: The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597.

Dee, D. P., and Coauthors, 2011b: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137 (656)**, 553–597, doi:10.1002/qj.828.

Dudhia, J., S.-Y. Hong, and K.-S. Lim, 2008: A new method for representing mixed-phase particle fall speeds in bulk microphysics parameterizations. *Journal of the Meteorological Society of Japan*, **86A**, 33–44, doi:10.2151/jmsj.86A.33.

Flesch, T. K., and G. W. Reuter, 2012: Wrf model simulation of two alberta flooding events and the impact of topography. *J. Hydrometeor.*, **13 (2)**, 695–708.

Fowler, H., and D. Archer, 2006: Conflicting signals of climatic change in the upper indus basin. *J. Climate*, **19 (17)**, 4276–4293, doi:10.7289/V5C8276M.

Galarneau Jr, T. J., T. M. Hamill, R. M. Dole, and J. Perlwitz, 2012: A multiscale analysis of the extreme weather events over western russia and northern pakistan during july 2010. *Mon. Wea. Rev.*, **140 (5)**, 1639–1664.

Gerard, L., 2007: An integrated package for subgrid convection, clouds and precipitation compatible with meso-gamma scales. *Quart. J. Roy. Meteor. Soc.*, **133 (624)**, 711–730, doi:10.1002/qj.58.

Haynes, J., Z. Luo, G. Stephens, R. Marchand, and A. Bodas-Salcedo, 2007: A multipurpose radar simulation package: QuickBeam. *Bull. Amer. Meteor. Soc.*, **88 (11)**, 1723–1727, doi:http://dx.doi.org/10.1175/BAMS-88-11-1723.

Herrera, S., J. Fernández, and J. Gutiérrez, 2015: Update of the spain02 gridded observational dataset for euro-cordex evaluation: assessing the effect of the interpolation methodology. *Int. J. Climatol.*

Hong, C.-C., H.-H. Hsu, N.-H. Lin, and H. Chiu, 2011: Roles of European blocking and tropical-extratropical interaction in the 2010 Pakistan flooding. *Geophys. Res. Lett.*, **38 (13)**, doi:10.1029/2011GL047583.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, **42 (2)**, 129–151.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134 (9)**, 2318–2341, doi:tp://dx.doi.org/10.1175/MWR3199.1.

Houze Jr, R., K. Rasmussen, S. Medina, S. Brodzik, and U. Romatschke, 2011: Anomalous atmospheric events leading to the summer 2010 floods in Pakistan. *Bull. Amer. Meteor. Soc.*, **92 (3)**, 291–298, doi:http://dx.doi.org/10.1175/2010BAMS3173.1.

Huffman, G. J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteor.*, **36 (9)**, 1191–1201.

Jung, S.-H., E.-S. Im, and S.-O. Han, 2012: The effect of topography and sea surface temperature on heavy snowfall in the yeongdong region: A case study with high resolution wrf simulation. *Asia-Pacific Journal of Atmospheric Sciences*, **48 (3)**, 259–273.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47 (23)**, 2784–2802, doi:http://dx.doi.org/10.1175/1520-0469(1990)047⟨2784:AODEPM⟩2.0.CO;2.

Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Wea. Forecasting*, **21 (2)**, 167–181, doi:http://dx.doi.org/10.1175/WAF906.1.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23 (5)**, 931–952, doi:http://dx.doi.org/10.1175/WAF2007106.1.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15 (3)**, 809–817, doi:http://dx.doi.org/10.1175/1520-0426(1998)015⟨0809:TTRMMT⟩2.0.CO;2.

Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (ccn) for weather and climate models. *Mon. Wea. Rev.*, **138 (5)**, 1587–1612.

Matrosov, S. Y., M. D. Shupe, and I. V. Djalalova, 2008: Snowfall retrievals using millimeter-wavelength cloud radars. *Journal of Applied Meteorology and Climatology*, **47 (3)**, 769–777.

Maussion, F., D. Scherer, R. Finkelnburg, J. Richters, W. Yang, and T. Yao, 2011: WRF simulation of a precipitation event over the Tibetan Plateau, China-an assessment using remote sensing and ground observations. *Hydrology and Earth System Sciences*, **15**, 1795–1817, doi:10.5194/hess-15-1795-2011.

Mishchenko, M. I., and L. D. Travis, 1998: Capabilities and limitations of a current fortran implementation of the t-matrix method for randomly oriented, rotationally symmetric scatterers. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **60 (3)**, 309–324.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *J. Geophys. Res.: Atmospheres (1984–2012)*, **102 (D14)**, 16 663–16 682.

Morrison, H., A. Morales, and C. Villanueva-Birriel, 2015: Concurrent sensitivities of an idealized deep convective storm to parameterization of microphysics, horizontal grid resolution, and environmental static stability. *Mon. Wea. Rev.*, **(2015)**.

Palazzi, E., J. von Hardenberg, and A. Provenzale, 2013: Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *Journal of Geophysical Research: Atmospheres*, **118**, 85–100, doi:10.1029/2012JD018697.

Parodi, A., and S. Tanelli, 2010: Influence of turbulence parameterizations on high-resolution numerical modeling of tropical convection observed during the tc4 field campaign. *J. Geophys. Res.: Atmospheres*, **115 (D10)**, doi:10.1029/2009JD013302.

Prakash, S., A. K. Mitra, E. Rajagopal, and D. Pai, 2015: Assessment of trmm-based tmpa-3b42 and gsmap precipitation products over india for the peak southwest monsoon season. *International Journal of Climatology*.

Rasmussen, K. L., A. J. Hill, V. E. Toma, M. D. Zuluaga, P. J. Webster, and R. A. Houze, 2014: Multiscale analysis of three consecutive years of anomalous flooding in Pakistan. *Quart. J. Roy. Meteor. Soc.*, doi:10.1002/qj.2433.

Sardar, S., I. Ahmad, S. S. Raza, and N. Irfan, 2012: Simulation of south Asian physical environment using various cumulus parameterization schemes of MM5. *Meteor. Appl.*, **19 (2)**, 140–151, doi:10.1002/met.266.

Skamarock, W., and Coauthors, 2008: A description of the advanced research wrf version 3. *NCAR technical note NCAR/TN/u2013475*, 1–113, doi:10.5065/D68S4MVH.

Stephens, G. L., and Coauthors, 2008: Cloudsat mission: Performance and early science after the first year of operation. *J. Geophys. Res.: Atmospheres*, **113 (D8)**, doi:10.1029/2008JD009982.

Tanelli, S., S. L. Durden, E. Im, K. S. Pak, D. G. Reinke, P. Partain, J. M. Haynes, and R. T. Marchand, 2008: Cloudsat's cloud profiling radar after two years in orbit: Performance, calibration, and processing. *Geoscience and Remote Sensing, IEEE Transactions on*, **46 (11)**, 3560–3573.

Tanelli, S., E. Im, S. L. Durden, L. Facheris, and D. Giuli, 2002: The effects of nonuniform beam filling on vertical rainfall velocity measurements with a spaceborne Doppler radar. *J. Atmos. Oceanic Technol.*, **19 (7)**, 1019–1034, doi:10.1175/1520-0426(2002)0192.0.CO;2.

Tanelli, S., and Coauthors, 2011: NASA's integrated Instrument Simulator Suite for Atmospheric Remote Sensing from spaceborne platform (ISSARS). *Earth Sci.*

Tanelli, S., and Coauthors, 2012: Integrated instrument simulator suites for Earth Science. *SPIE Asia-Pacific Remote Sensing*, International Society for Optics and Photonics, 85 290D–85 290D, doi:10.1117/12.977577.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136 (12)**, 5095–5115, doi:http://dx.doi.org/10.1175/2008MWR2387.1.

Tian, Y., and C. D. Peters-Lidard, 2010: A global map of uncertainties in satellite-based precipitation measurements. *Geophys. Res. Lett.*, **37 (24)**.

Ullah, K., and G. Shouting, 2013: A diagnostic study of convective environment leading to heavy rainfall during the summer monsoon 2010 over Pakistan. *Atmos. Res.*, **120**, 226–239, doi:10.1016/j.atmosres.2012.08.021.

Ushiyama, T., T. Sayama, Y. Tatebe, S. Fujioka, and K. Fukami, 2014: Numerical simulation of 2010 Pakistan flood in the Kabul River Basin by using lagged ensemble rainfall forecasting. *J. Hydrometeor.*, **15 (1)**, 193–211.

Webster, P., V. Toma, and H.-M. Kim, 2011: Were the 2010 Pakistan floods predictable? *Geophys. Res. Lett.*, **38 (4)**, doi:10.1029/2010GL046346.

Winiger, M., M. Gumpert, and H. Yamout, 2005: Karakorum–hindukush–western himalaya: assessing high-altitude water resources. *Hydrological Processes*, **19 (12)**, 2329–2338.

Yu, C., and M. Teixeira, 2014: Impact of non-hydrostatic effects and trapped lee waves on mountain-wave drag in directionally sheared flow. *Quart. J. Roy. Meteor. Soc.*

Yu, X., and T.-Y. Lee, 2010: Role of convective parameterization in simulations of a convection band at grey-zone resolutions. *Tellus*, **62 (5)**, 617–632, doi:http://dx.doi.org/10.1175/JAS-D-12-0104.1.

LIST OF TABLES

TABLE 1: Experiment configurations.

Configuration	*Convective closure*	*Microphysics*
Exp-WSM6	Explicit	WSM6
KF-WSM6	Kain-Fritsch	WSM6
Exp-Thompson	Explicit	Thompson
KF-Thompson	Kain-Fritsch	Thompson

TABLE 2: Statistical score analysis for the different configurations for July 28th (upper panel) and for July 29th (lower panel). The first part of the table shows the values of MODE verification analysis of centroid distance, area ratio and and interest. The MODE evaluation refers to the highest intensity object identified in each run that matches with the corresponding TRMM object. The matched objects are shown in Fig.4. In the second part the different percentiles (median, 60th, 90th and 95th) are shown . In the third part are reported MB and RMSE. The fourth part of the table shows MB and RMSE calculated between raingauge station measures and associated nearest neighbour WRF grid point. The first three parts of the table use TRMM as reference dataset. The fourth part of the table shows MB and RMSE calculated between raingauge station measures and associated nearest neighbour WRF grid point.

July 28th	$Ex-WSM6$	$KF-WSM6$	$Ex-Thomson$	$KF-Thompson$	$TRMM$
$CENTROID\ DISTANCE$	601	1860	1934	1884	-
$AREA\ RATIO$	0.919	0.452	0.571	0.422	-
$INTEREST$	0.961	0.858	0.851	0.842	-
$PERCENTILE_{60}$	12.19	15.90	12.95	15.15	4.83
$PERCENTILE_{95}$	53.30	67.74	58.99	71.88	52.08
MB	3.73	8.43	5.28	8.03	-
$RMSE$	21.46	26.92	27.31	26.66	-
$MB_{raingauges}$	-20.34	-11.56	-14.60	-10.83	-
$RMSE_{raingauges}$	65.49	65.23	68.81	59.14	-

July 29th	$Ex-WSM6$	$KF-WSM6$	$Ex-Thomson$	$KF-Thompson$	$TRMM$
$CENTROID\ DISTANCE$	967	1208	472	551	-
$AREA\ RATIO$	0.567	0.599	0.544	0.529	-
$INTEREST$	0.914	0.899	0.946	0.940	-
$PERCENTILE_{60}$	3.63	6.55	3.62	5.87	1.04
$PERCENTILE_{95}$	69.99	62.38	83.57	94.04	44.70
MB	6.05	7.27	8.79	8.10.62	-
$RMSE$	30.42	38.12	40.94	40.60	-
$MB_{raingauges}$	-10.41	-10.97	0.44	14.94	-
$RMSE_{raingauges}$	62.54	87.48	96.60	93.46	-

TABLE 3: Summary of all the different runs performed in the second part of the experiment.

Run	*Day of initialization*	*Configuration*	*Initial conditions*	*Boundary conditions*	*Hig − res domain*
J26	July 26^{th}	Exp-WSM6	ERA Interim	ERA Interim	10N to 50N 60E to 90E
J24	July 24^{th}	Exp-WSM6	ERA Interim	ERA Interim	10N to 50N 60E to 90E
J28	July 28^{th}	Exp-WSM6	ERA Interim	ERA Interim	10N to 50N 60E to 90E
J28S	July 28^{th}	Exp-WSM6	ERA Interim	ERA Interim	23N to 40N 66E to 78E
J28R	July 28^{th}	Exp-WSM6	WRF J26 restarted at July 28^{th} 00 UTC	ERA Interim	10N to 50N 60E to 90E

TABLE 4: Statistical score analysis for the different initializations, for July 28th (upper panel) and for July 29th (lower panel). The first part of the table shows the values of MODE verification analysis of centroid distance, area ratio and and interest. The MODE evaluation refers to the highest intensity object identified in each run that matches with the corresponding TRMM object. The matched objects are shown in Fig.5. In the second part the different percentiles (median, 60th, 90th and 95th) are shown . In the third part are reported MB and RMSE. The fourth part of the table shows MB and RMSE calculated between raingauge station measures and associated nearest neighbour WRF grid point. The first three parts of the table use TRMM as reference dataset. The fourth part of the table shows MB and RMSE calculated between raingauge station measures and associated nearest neighbour WRF grid point.

July 28th	*J24*	*J26*	*J28*	*TRMM*
CENTROID DISTANCE	568	601	322	-
AREA RATIO	0.815	0.919	0.750	-
INTEREST	0.963	0.961	0.984	-
$PERCENTILE_{60}$	6.53	12.19	6.69	4.83
$PERCENTILE_{95}$	53.29	53.30	55.40	52.08
MB	0.77	3.73	1.96	-
RMSE	20.18	21.46	21.80	-
$MB_{raingauges}$	-20.22	-20.34	-9.10	-
$RMSE_{raingauges}$	58.41	65.49	56.31	-

July 29th	*J24*	*J26*	*J28*	*TRMM*
CENTROID DISTANCE	1544	967	633	-
AREA RATIO	0.558	0.567	0.924	-
INTEREST	0.659	0.914	0.957	-
$PERCENTILE_{60}$	3.28	3.63	2.80	1.04
$PERCENTILE_{95}$	36.25	69.99	39.01	44.70
MB	0.31	6.05	0.28	-
RMSE	26.35	30.42	19.24	-
$MB_{raingauges}$	-30.41	-10.41	-18.24	-
$RMSE_{raingauges}$	65.04	62.54	49.83	-

LIST OF FIGURES

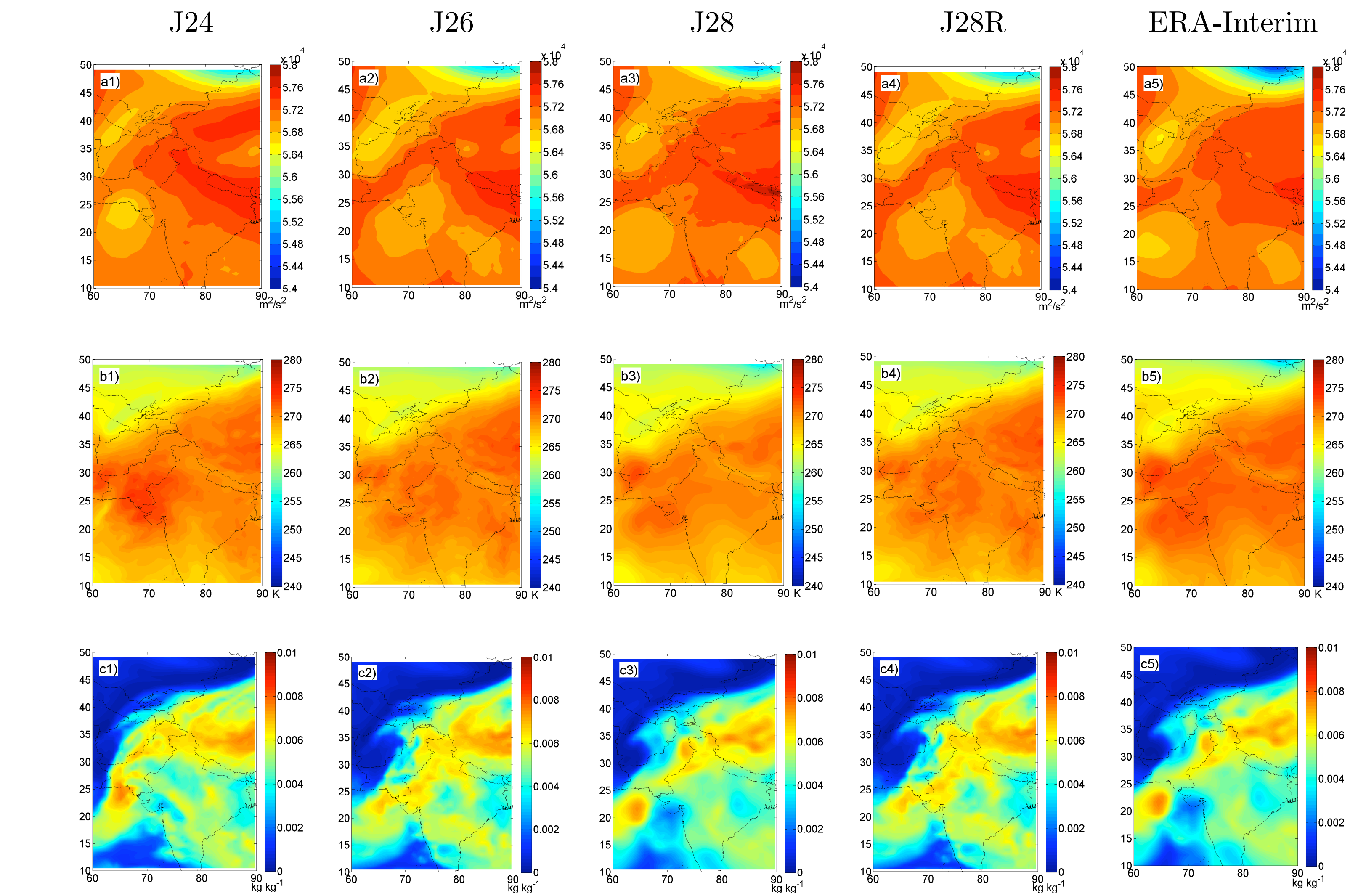

FIG. 1: Large scale circulation of geopotential (a panels), temperature (b panels) and specific humidity (c panels) at 500 hPa on July 28th at 00 UTC as simulated by WRF J24 (a1, b1, c1), J26 (a2, b2, c2), J28 (a3, b3, c3), J28R (a4, b4, c4) runs and by ERA-Interim reanalysis (a5, b5, c5).

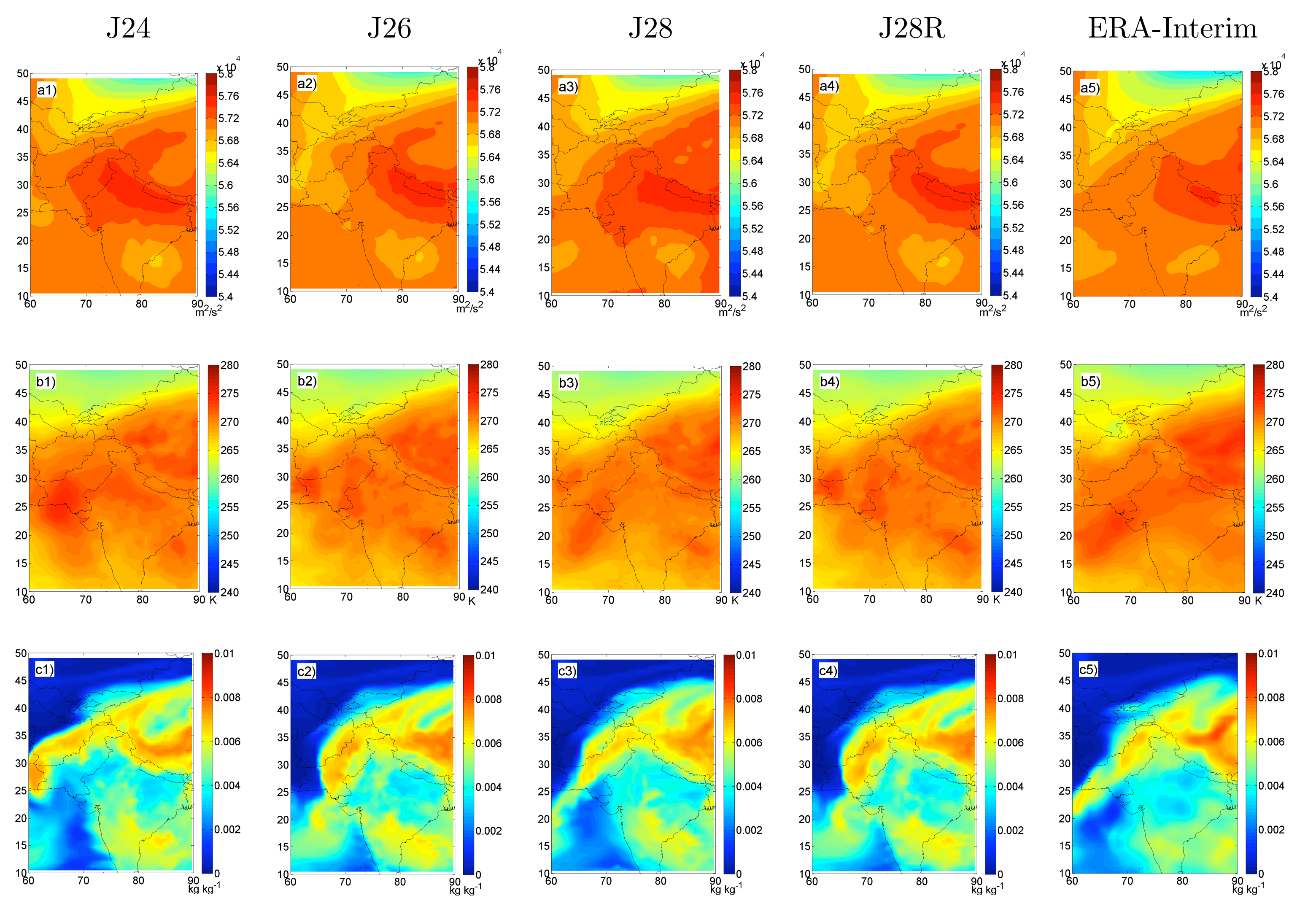

FIG. 2: Large scale circulation of geopotential (a panels), temperature (b panels) and specific humidity (c panels) at 500 hPa on July 29th at 00 UTC, as simulated by WRF J24 (a1, b1, c1), J26 (a2, b2, c2), J28 (a3, b3, c3), J28R (a4, b4, c4) runs and by ERA-Interim reanalysis (a5, b5, c5).

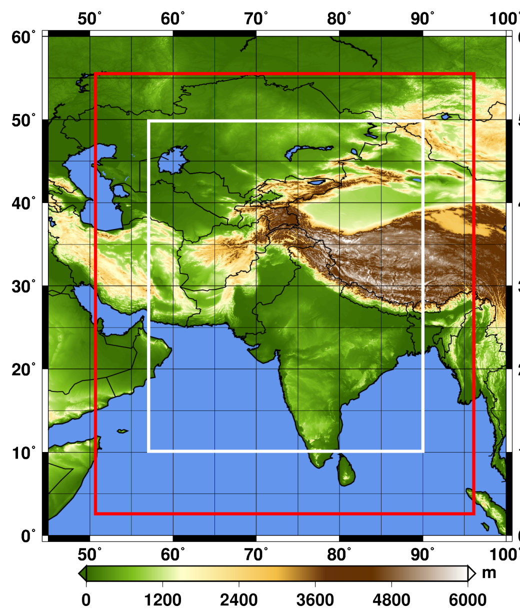

FIG. 3: The two nested domains used for the simulations: external domain d01 (red box) resolved at 14 km resolution and inner domain d02 (white box) resolved at 3.5 km. The color levels report the orography of the region, provided by the ETOPO1 dataset.

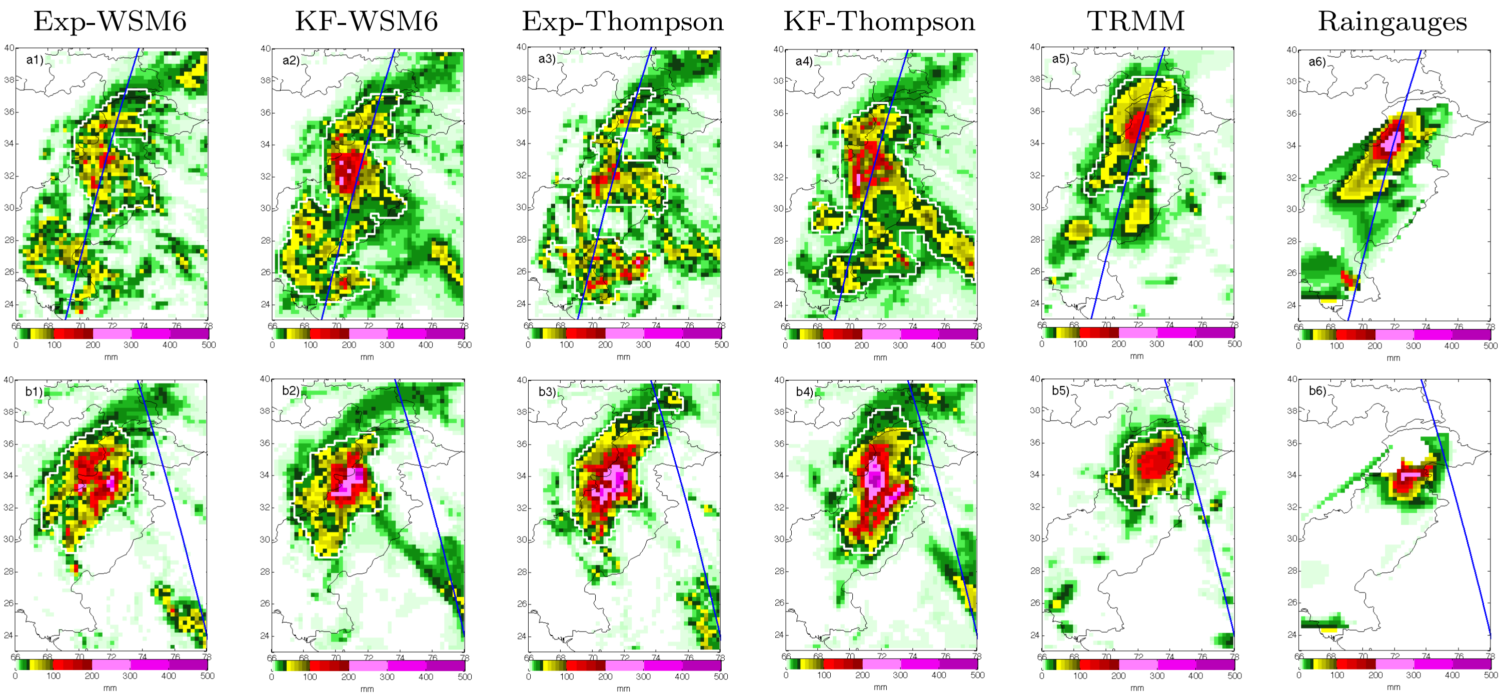

FIG. 4: WRF Quantitative Precipitation Forecasts and TRMM daily rainfall. From left to right: Exp-WSM6 (a1, b1), KF-WSM6 (a2, b2), Exp-Thompson (a3, b3), KF-Thompson (a4, b4), TRMM (a5, b5) and raingauge observations (a6, b6). All fields have been aggregated at 0.25°resolution in the study area. The top row refers to July 28th 2010 (a) and the bottom row refers to July 29th (b). The blue lines represent CloudSat tracks and the white contour represent the object identified by MODE analysis.

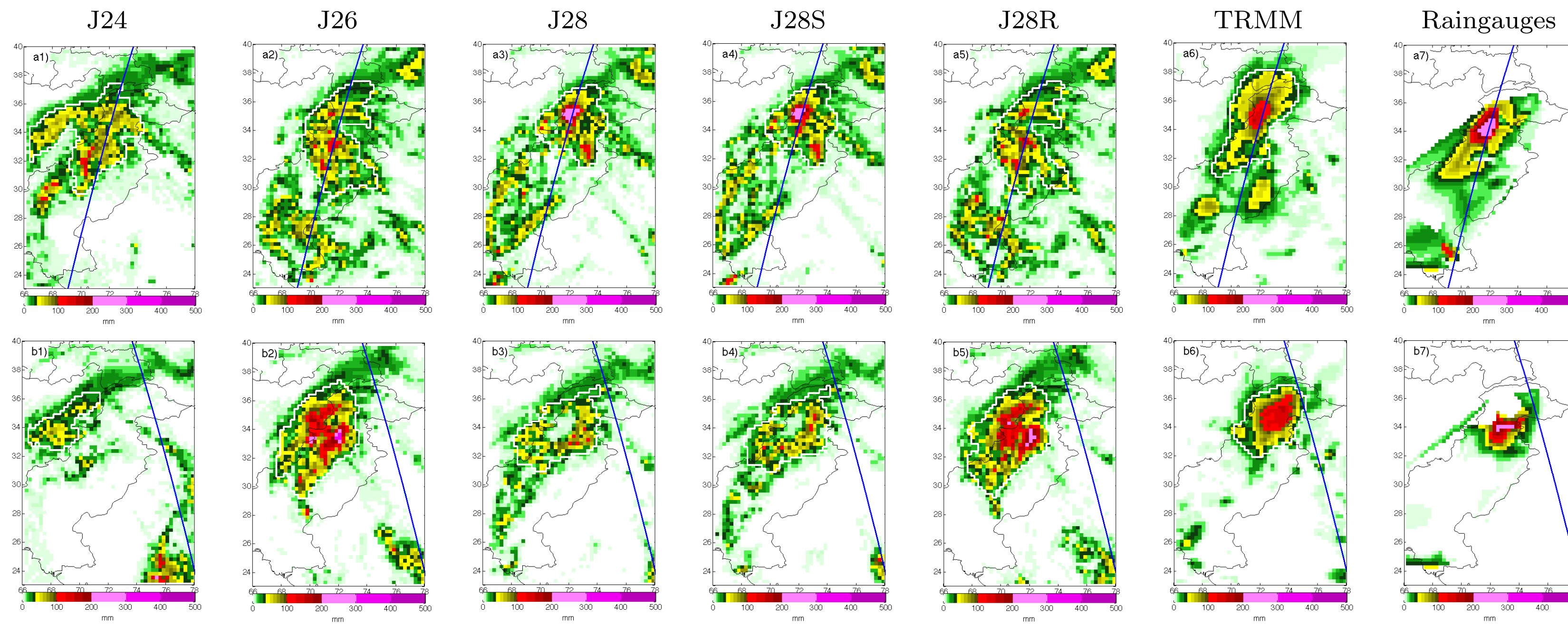

FIG. 5: First row: 24-hr rainfall cumulates on July 28th given by: J24 (a1), J26 (a2), J28 (a3), J28S (a4), J28R (a5), TRMM (a6) and raingauge stations (a7). Second row: 24-hr rainfall accumulation on July 29th given by: J24 (b1), J26 (b2), J28 (b3), J28S (b4), J28R (b5), TRMM (b6) and raingauge stations (b7). All rainfall fields have been aggregated at 0.25°horizontal resolution. The blue lines represent CloudSat tracks and the white contour represent the object identified by MODE analysis.

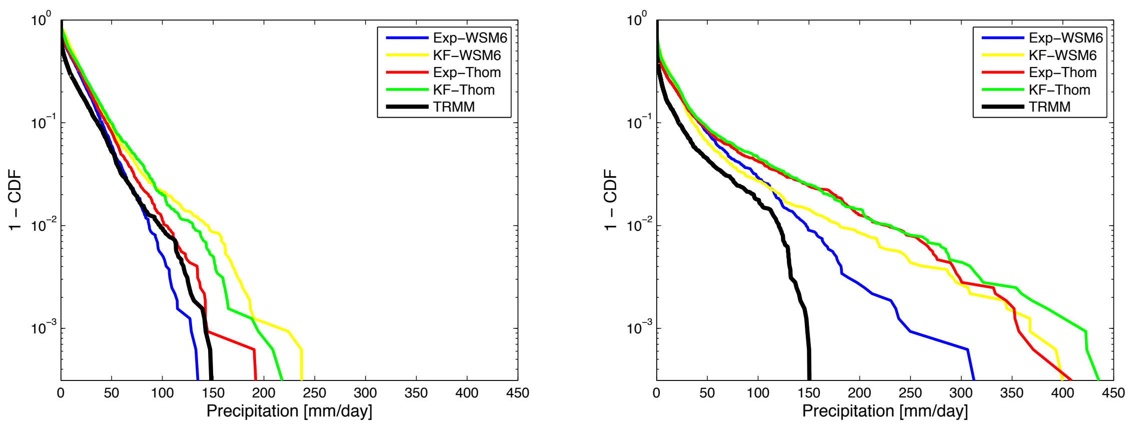

FIG. 6: Comparison between probabilities of exceedence (1-CDF) for daily rainfall from WRF simulations and TRMM estimates, for July 28th (left panel) and July 29th (right panel). Spatial resolution is 0.25°and the results refer to the whole study area.

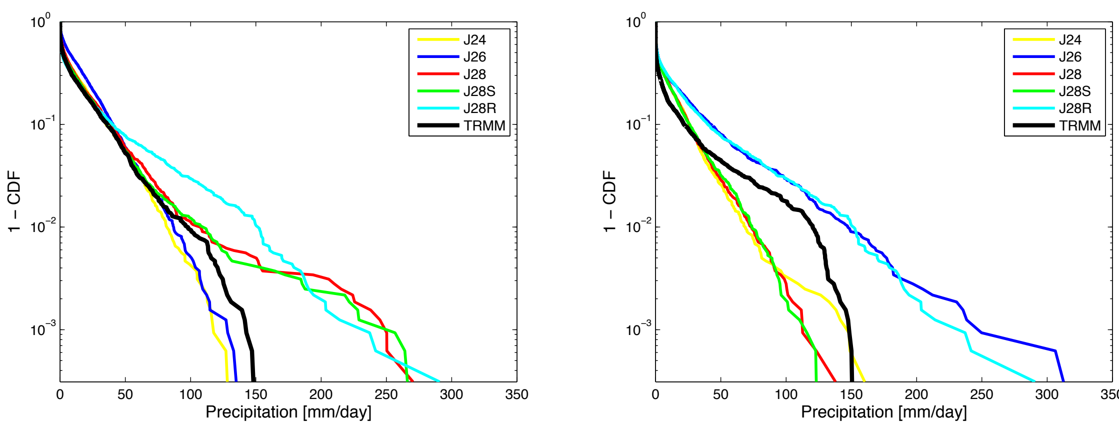

FIG. 7: Comparison between probabilities of exceedence (1-CDF) obtained from WRF using different initialization days and those derived from TRMM estimates. Left panel: July 28th; right panel: July 29th. The spatial resolution is 0.25°and the results refer to the whole study area.

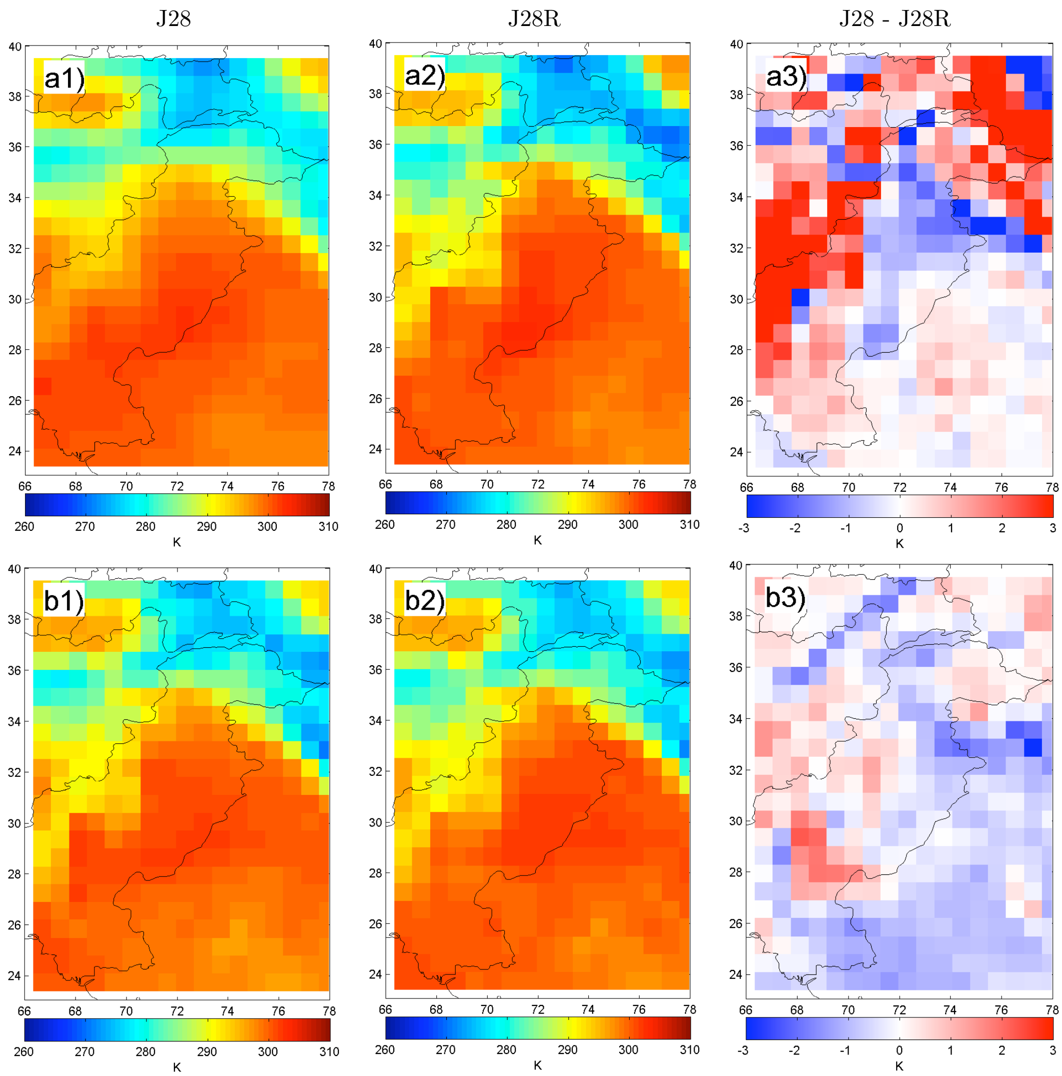

FIG. 8: Surface temperature at the time of initialization (28th at 00 UTC) and on 29th at 00 UTC for the J28 and J28R runs. Upper row: Temperature field at 2m in the J28 run on July 28th at 00 UTC (a1); the same for the J28R run (a2); pixel-by-pixel difference between these two temperature fields (a3). Bottom row: Temperature field at 2m for the J28 run on July 29th at 00 UTC (b1); the same for the J28R run (b2); pixel-by-pixel difference between these two temperature fields (b3). Temperature fields are plotted at 0.75° horizontal resolution.

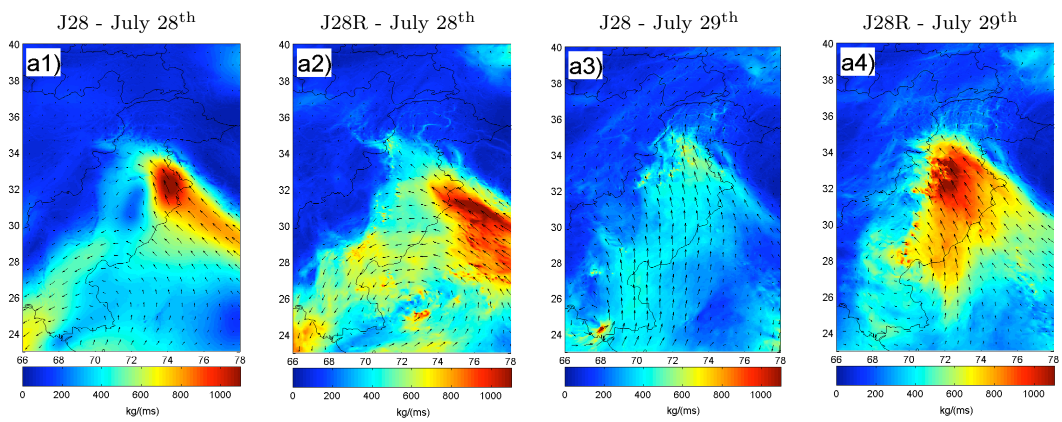

FIG. 9: Moisture transport field for the J28 run on July 28$^{\text{th}}$ at 00 UTC (a1); the same for the J28R run (a2); moisture transport for the J28 run on July 29$^{\text{th}}$ at 00 UTC (a3); the same for the J28R run (a4). Moisture transport fields are plotted at the resolution of WRF simulations (3.5 km). The colors indicate the intensity and the vectors rapresent the directions of the moist transport.

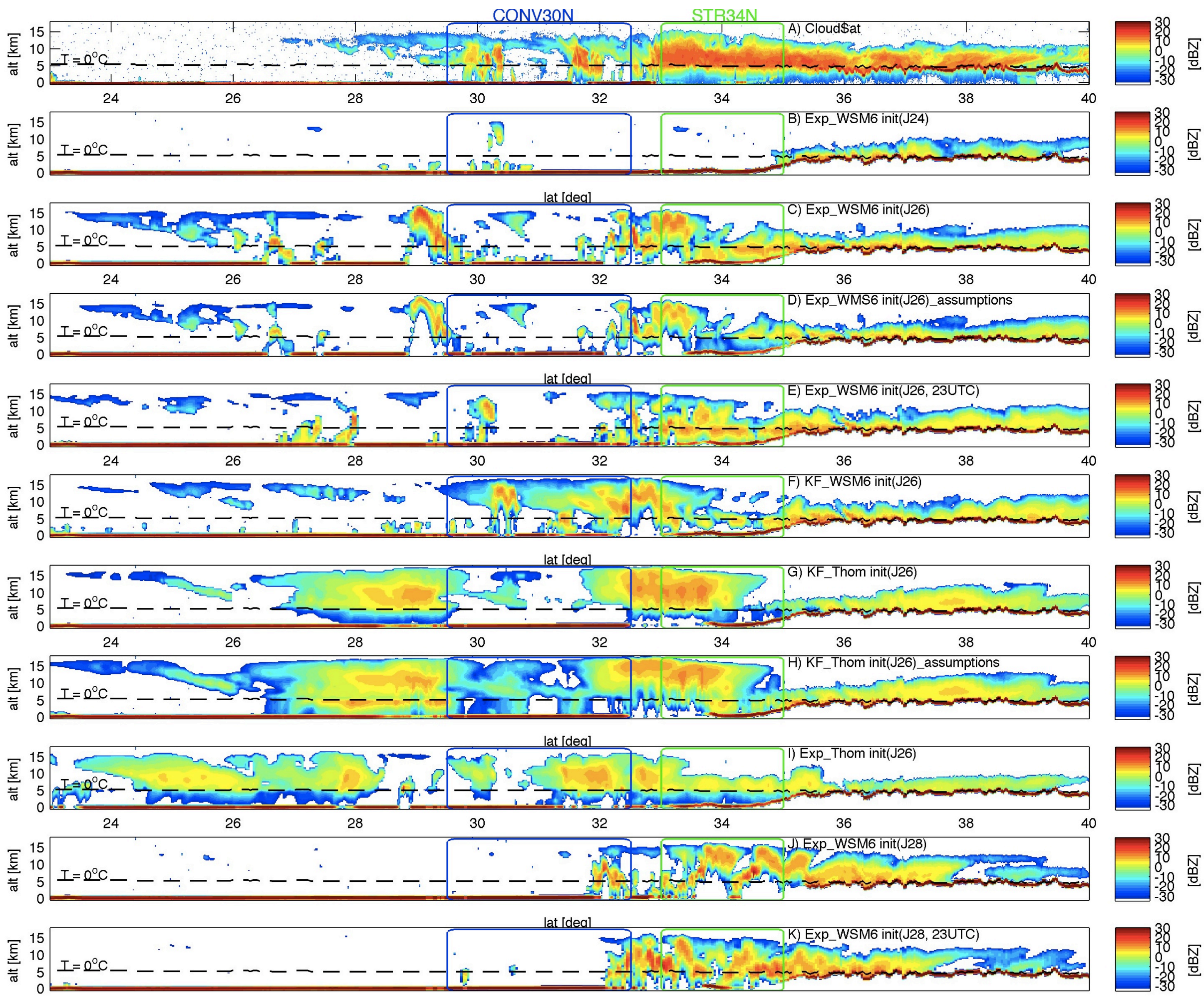

FIG. 10: Vertical structure of the atmosphere on July 28th at 21 UTC. From the upper to the lower panel: CloudSat observation (Granule 22608) (a) and DS3 CloudSat simulations for Exp-WSM6 initialized on J24 (b), Exp-WSM6 initialized on J26 (c), Exp-WSM6 initialized on J26 with different microphysical assumptions (d), Exp-WSM6 at 23 UTC initialized on J26 (e), KF-WSM6 initialized on J26 (f), KF-Thompson initialized on J26 (g), KF-Thompson initialized on J26 with different microphysical assumptions (h), Exp-Thompson initialized on J26 (i), Exp-WSM6 initialized on J28 (j), Exp-WSM6 at 23 UTC initialized on J28 (k).