Assessment of the Weather Research and Forecasting (WRF) model for simulation of extreme rainfall events in the upper Ganga Basin

Ila Chawla¹, Krishna K. Osuri³,⁴, Pradeep P. Mujumdar¹,², and Dev Niyogi⁴,⁵

¹Department of Civil Engineering, Indian Institute of Science, Bangalore, 560012, India
²Divecha Centre for Climate Change, Indian Institute of Science, Bangalore, 560012, India
³Department of Earth and Atmospheric Sciences, NIT Rourkela, Odisha, 769008, India
⁴Department of Agronomy–Crops, Soils, Water Sciences, Purdue University, West Lafayette, IN 47907, USA
⁵Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA

Correspondence: Pradeep P. Mujumdar (pradeep@iisc.ac.in)

Received: 27 August 2017 – Discussion started: 30 August 2017
Accepted: 29 December 2017 – Published: 8 February 2018

Abstract. Reliable estimates of extreme rainfall events are necessary for an accurate prediction of floods. Most of the global rainfall products are available at a coarse resolution, rendering them less desirable for extreme rainfall analysis. Therefore, regional mesoscale models such as the advanced research version of the Weather Research and Forecasting (WRF) model are often used to provide rainfall estimates at fine grid spacing. Modelling heavy rainfall events is an enduring challenge, as such events depend on multi-scale interactions, and the model configurations such as grid spacing, physical parameterization and initialization. With this background, the WRF model is implemented in this study to investigate the impact of different processes on extreme rainfall simulation, by considering a representative event that occurred during 15–18 June 2013 over the Ganga Basin in India, which is located at the foothills of the Himalayas. This event is simulated with ensembles involving four different microphysics (MP), two cumulus (CU) parameterizations, two planetary boundary layers (PBLs) and two land surface physics options, as well as different resolutions (grid spacing) within the WRF model. The simulated rainfall is evaluated against the observations from 18 rain gauges and the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis (TMPA) 3B42RT version 7 data. From the analysis, it should be noted that the choice of MP scheme influences the spatial pattern of rainfall, while the choice of PBL and CU parameterizations influences the magnitude of rainfall in the model simulations. Further, the WRF run with Goddard MP, Mellor–Yamada–Janjic PBL and Betts–Miller–Janjic CU scheme is found to perform “best” in simulating this heavy rain event. The selected configuration is evaluated for several heavy to extremely heavy rainfall events that occurred across different months of the monsoon season in the region. The model performance improved through incorporation of detailed land surface processes involving prognostic soil moisture evolution in Noah scheme compared to the simple Slab model. To analyse the effect of model grid spacing, two sets of downscaling ratios – (i) 1:3, global to regional (G2R) scale and (ii) 1:9, global to convection-permitting scale (G2C) – are employed. Results indicate that a higher downscaling ratio (G2C) causes higher variability and consequently large errors in the simulations. Therefore, G2R is adopted as a suitable choice for simulating heavy rainfall event in the present case study. Further, the WRF-simulated rainfall is found to exhibit less bias when compared with the NCEP FiNaL (FNL) reanalysis data.
1 Introduction

Indian summer monsoon rainfall (ISMR) is often associated with very heavy (124.5 to 244.4 mm day$^{-1}$) to extremely heavy (more than 244.5 mm day$^{-1}$) rainfall (Indian Meteorological Department, Terminologies and Glossary; http://imd.gov.in/section/nhac/termglossary.pdf), particularly during June to September months. The extremely heavy rainfall events usually occur due to the presence of organized meso-convective systems (MCSs) embedded in large-scale monsoonal features such as offshore troughs and vortices, depressions over the Bay of Bengal and Arabian Sea, and mid-tropospheric cyclones (Sikka and Gadgil, 1980; Webster et al., 1998; Fasullo and Webster, 2003).

Extremely heavy rainfall on shorter timescales are particularly difficult to predict in mountainous terrains and continue to be a challenge to operational and research communities (Das et al., 2008; Li et al., 2017). Global models have been employed in several studies to understand the large-scale circulation pattern and for quantitative analysis of the monsoon rainfall, but due to their coarse resolution, they are unable to represent the local to regional characteristics of monsoon rainfall. Regional models, however, can explicitly simulate the interactions between the large-scale weather phenomenon and regional topography, making the climate simulations reliable (Gadgil and Sajani, 1998; Ratna et al., 2011; Srinivas et al., 2013). Furthermore, regional models have a better representation of convection, thus offsetting one of the major sources of errors and uncertainties in the global models. Therefore, regional models become a preferred choice to study seasonal monsoon rainfall.

The advanced research version of the Weather Research and Forecasting model (hereafter referred to as the WRF model) is a regional popular community model that is widely used for both studying as well as forecasting a variety of high-impact meteorological events, such as rainfall (Vaidya and Kulkarni, 2007; Deb et al., 2008; Kumar et al., 2008; Chang et al., 2009; Routray et al., 2010; Mohanty et al., 2012), tropical cyclones (Raju et al., 2011; Routray et al., 2016; Osuri et al., 2017b) and thunderstorms (Madala et al., 2014; Osuri et al., 2017a). Several works are reported in the literature which have considered the WRF model over the Himalayan region. Kumar et al. (2012) used the WRF model to simulate the cloudburst event of 2010 in the Leh area over the north-western Himalayan belt, while Kumar et al. (2014) and Thayyen et al. (2013) used the WRF model to gain insight into the atmospheric processes and the MCSs that led to the 2010 Leh event. Similarly, Chevuturi et al. (2015) simulated the heavy precipitation event of September 2012 in the central Himalayas using the WRF model. Medina et al. (2010) used the WRF model to understand how topography and land surface conditions affect the extreme convection in western and eastern Himalayas. Particularly for the 2013 heavy rainfall episode in the Uttarakhand region, the WRF model is used in several studies, including those by Kotal et al. (2014), Vellore et al. (2016) and Hazra et al. (2017), to understand the physical processes leading to the event. Shekhar et al. (2015), Chevuturi and Dimri, (2016), and Dimri et al. (2016) performed in-depth synoptic analysis of the June 2013 heavy rainfall event using the WRF model. Rajesh et al. (2016) presented the role of land surface conditions in simulating the heavy rainfall event. Therefore, from the existing literature, it can be established that the regional model performs considerably well over the region. However, finding the optimal set of physics parameterization schemes (along with the selection of appropriate model grid spacing and resolution) to simulate heavy to extremely heavy rainfall events, and understanding the effect of the combination of different parameterization schemes on rainfall estimates over the Indian monsoon region is still an active area of research.

Earlier studies such as those by Krishnamurthy et al. (2009), Misinis and Zhang, (2010), Rauscher et al. (2010), Mohanty et al. (2012), and Chevuturi et al. (2015) indicated that heavy rainfall predictions can be improved through ensemble model techniques and fine grid resolution. However, the influence of the interaction between model parameterization schemes on mesoscale rainfall simulations over India is still an understudied issue. In particular, heavy rain simulation studies have reviewed the impact of individual parameterization options such as the microphysics (MP) scheme (Rajeevan et al., 2010; Raju et al., 2011; Kumar et al., 2012), cumulus (CU) parameterization scheme (Deb et al., 2008; Mukhopadhyay et al., 2010; Srinivas et al., 2013; Madala et al., 2014), planetary boundary layer (PBL) scheme (Li and Pu, 2008; Hu et al., 2010; Hariprasad et al., 2014) and land surface model (LSM) options (Chang et al., 2009). However, the ensemble analysis that reviews the relative impact of different configurations and the associated variability (uncertainty) is lacking. It is important to study the impact of different parameterizations in an ensemble mode because it is often likely that the performance of one scheme depends on other model configurations considered. For example, the conclusions regarding which CU scheme performs best would be intimately tied to the choice of the MP or land surface options considered in conducting the numerical experiments. With this perspective, this paper seeks to assess the sensitivity of the WRF model in predicting heavy to extremely heavy rainfall episodes over the Ganga Basin in the foothills of the Himalayas. Specific tasks undertaken in this work are as follows: (i) quantitative verification of the WRF model to simulate an extremely heavy rainfall event; (ii) assessment of the sensitivity of the model-simulated rainfall to different parameterization options, downscaling ratios and land surface models; (iii) validation of the selected configuration for other rainfall events over the region; and (iv) comparison of the WRF-simulated rainfall with the global reanalysis data to investigate the impact of local versus global factors on rainfall simulations. A related objective is to provide suitable recommendations on a possible optimal choice for model configuration to simulate such heavy rainfall events in the region.
I. Chawla et al.: Assessment of the Weather Research and Forecasting model

Figure 1. Topography of the study region (shown with black outline) as represented in the WRF model for (i) Domain 1–27 km grid spacing, (ii) Domain 2a–9 km grid spacing (downscaling ratio–1:3) and (iii) Domain 2b–3 km grid spacing (downscaling ratio–1:9). Locations of the rain gauge stations within the UGB are presented as black dots in (iii).

Description of the heavy rainfall event

The 2013 summer monsoon had a normal onset but the trough advanced rapidly, covering the whole of India by mid-June, instead of mid-July (Ray et al., 2014). This large-scale setting is thought to have created a platform for interaction between two synoptic-scale events – a north-west moving depression from the Bay of Bengal and a pre-existing westerly trough in the mid-troposphere. Meteorological studies conducted over the region (Kotal et al., 2014; Ray et al., 2014; Chevuturi and Dimri, 2016; Rajesh et al., 2016) established that there was a monsoon low-pressure system during this period. The longitudinal time section for 850 hPa geopotential height along with anomaly averaged over 20–26° N showed a high negative anomaly on 14 June, which migrated to the west, moving over 75° E by 17 June. The meridional wind anomaly within the belt of 35–45° N showed a westerly wave, moving from 10° E on 12 June to 70° E on 17 June. These two anomalies are found to be in phase, consequently causing interaction between the eastward-moving trough in the mid- to upper troposphere and westward-moving monsoon low in the lower troposphere. The monsoon low provided the moisture feed and the upper-level westerly trough provided the divergence to lift the moisture. This whole system eventually led to an unanticipated heavy rainfall during 15–18 June 2013 in the Kedarnath valley and adjoining areas in the state of Uttarakhand, India (Kotal et al., 2014; Ray et al., 2014; Chevuturi and Dimri, 2016; Rajesh et al., 2016).

The region received rainfall greater than 370 mm in 1 day (16–17 June 2013), which is 375 % above the daily normal rainfall (65.9 mm) during the monsoon season (Ray et al., 2014). Consequently, heavy floods occurred in the region, causing unprecedented damage to life and property.

The synthesis of the synoptic setting of the event has been carried out in a number of studies such as Dube et al. (2014), Kotal et al. (2014), Ray et al. (2014), Shekhar et al. (2015), Chevuturi and Dimri, (2016), and Rajesh et al. (2016), but the mesoscale assessment pertaining to the simulation of this rainfall event is still lacking. Therefore, the present study emphasizes quantitatively evaluating and conducting the sensitivity analysis of the WRF model in predicting extreme rainfall. The ability of the WRF model to simulate heavy rainfall events is further verified by considering additional episodes (apart from the June 2013 event, details of which are presented in Sect. 2.1) that occurred within this region across different monsoon months.

A study region comprising of the upstream part of the Ganga Basin in India, referred as Upper Ganga Basin (UGB) hereafter, is selected for the analysis in this paper. Figure 1 presents the topography of the UGB as described for the three domains of the WRF model (Domain 1, Domain 2a and Domain 2b, with 27, 9 and 3 km grid resolution respectively) along with the 18 rain gauge stations located within the region. The region is of social, cultural and economic importance to India, further making this study necessary.

2 Data and experimental setup

2.1 Observed data

Figure 2 presents daily and cumulative rainfall data from 15 to 18 June 2013 obtained from the Indian Meteorological Department (IMD) and the literature (Ray et al., 2014) for the 18 official rain gauges located within the UGB.

It is noticed that the north-west part of the region received higher rainfall compared to the north-east, with stations such as Tehri and Dehradun showing 327 and 210 % (respectively) more rainfall than their historic means. A few stations like Chamoli in the north-east region received 250 mm of cumulative rainfall over the 4-day period, which is 144 % higher than the historic mean. In general, most of the stations in the southern part of the basin, which are located at a lower elevation, recorded relatively less rainfall with a cumulative range of 445 mm, in comparison to the northern part (at higher altitude), which had a rainfall range of 515 mm. Additionally, three stations in the south-east region, i.e. Mukteshwar, Haldwani and Nainital received extremely heavy rainfall with a cumulative average of 498 mm. From the above analysis, it is evident that the system moved from east to west with two distinct regions in the UGB – south-east and north-west – receiving extremely heavy rainfall.

The region has complex topography and a limited number of rain gauges because of the difficulty in operating a network in this region. To further capture the spatial variability in rainfall, the Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA) 3B42RT (version 7) product, which is available at 0.25° resolution (TRMM) Multi-Satellite Precipitation Analysis (TMPA) 3B42RT (version 7) product, which is available at 0.25° resolution on a daily scale, is analysed (Fig. 3). It is to be noted that, since the focus area for the analyses is the upstream region of the UGB (Figs. 1iii and 2), results are presented with respect to the geographical extent of Domain 2b throughout this paper. From Fig. 3, it can be noted that the TMPA data are able to capture the spatial variability in the rainfall – with distinct clusters corresponding to heavy rainfall in the north-west and south-east regions of the study area. However, the rainfall amount is significantly underestimated by the TMPA product, with the maximum value of 265 mm against the recorded 650 mm. This under-reporting for gridded satellite products versus rain gauges in the ISMR region is a well-known feature (Rahman et al., 2009; Mishra and Srinivasan, 2013; Kneis et al., 2014; Bharti et al., 2016). The TMPA estimates are verified against the IMD station observations for a baseline quality check. Mean absolute error (MAE), root mean square error (RMSE) and bias (β) are computed using the nearest neighbourhood mapping approach and are presented in Table 1.

TMPA data are observed to behave differently for different ranges of rainfall values over the study domain. TMPA overestimated the rainfall at stations with cumulative rainfall less than 200 mm (e.g. Pauri and Almora). In contrast, rainfall at stations receiving more than 250 mm of cumulative rainfall is underestimated. Stations that received rainfall of 200–250 mm are well represented in the TMPA data (e.g. Mussorie, Pithoragarh and Champawat). From the analysis, it could be inferred that in the TMPA data rainfall values are clustered towards the mean value. Errors noticed in the TMPA data could be attributed to two factors: first, large spatial coverage and coarse resolution of the TMPA data, and second, for comparison with the observed data a simple approach of selecting the nearest grid point is implemented.

In addition to June 2013 case, five additional heavy to extremely heavy rainfall events are also considered in the present study for the analysis, details of which are presented in Table 2. Rainfall from the IMD gridded data at 0.25° resolution (Pai et al., 2014) is considered as the observed data for these events.

It is to be noted that on 13–14 September 2012, a cloud-burst event was reported in the region and the total amount of rainfall on 14 September was recorded as approximately 210 mm (Chevuturi et al., 2015). This event is significantly underestimated in the IMD gridded data, indicating that caution must be exercised while using the data for applications involving heavy rainfall events, such as flood modelling and validating the rainfall simulations from the mesoscale models. Figure 4 presents the spatially averaged daily and cumulative rainfall received during different events (as specified in Table 2).

<table>
<thead>
<tr>
<th>Station</th>
<th>Mean absolute error (mm)</th>
<th>Root mean square error (mm)</th>
<th>Bias (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uttarakashi</td>
<td>64</td>
<td>97</td>
<td>−40</td>
</tr>
<tr>
<td>Tehri</td>
<td>63</td>
<td>83</td>
<td>−44</td>
</tr>
<tr>
<td>Mussorie</td>
<td>75</td>
<td>94</td>
<td>−6</td>
</tr>
<tr>
<td>Dehradun</td>
<td>126</td>
<td>191</td>
<td>−70</td>
</tr>
<tr>
<td>Joshimath</td>
<td>51</td>
<td>55</td>
<td>−21</td>
</tr>
<tr>
<td>Chamoli</td>
<td>44</td>
<td>55</td>
<td>−27</td>
</tr>
<tr>
<td>Rudraprayag</td>
<td>48</td>
<td>53</td>
<td>−22</td>
</tr>
<tr>
<td>Pauri</td>
<td>34</td>
<td>44</td>
<td>37</td>
</tr>
<tr>
<td>Deoprayag</td>
<td>95</td>
<td>98</td>
<td>−56</td>
</tr>
<tr>
<td>Hardwar</td>
<td>83</td>
<td>114</td>
<td>−76</td>
</tr>
<tr>
<td>Roorkee</td>
<td>64</td>
<td>82</td>
<td>−39</td>
</tr>
<tr>
<td>Ranikhet</td>
<td>71</td>
<td>77</td>
<td>27</td>
</tr>
<tr>
<td>Almora</td>
<td>26</td>
<td>30</td>
<td>63</td>
</tr>
<tr>
<td>Mukteshwar</td>
<td>90</td>
<td>118</td>
<td>−60</td>
</tr>
<tr>
<td>Nainital</td>
<td>86</td>
<td>111</td>
<td>−67</td>
</tr>
<tr>
<td>Haldwani</td>
<td>115</td>
<td>157</td>
<td>−80</td>
</tr>
<tr>
<td>Pithoragarh</td>
<td>56</td>
<td>69</td>
<td>−6</td>
</tr>
<tr>
<td>Champawat</td>
<td>100</td>
<td>126</td>
<td>1</td>
</tr>
</tbody>
</table>

* Bias (%) = (Station data – TMPA data) / Station data × 100.

2.2 Model configuration and experimental setup

The simulation experiments in this paper are conducted using the WRF model, version 3.8. WRF is a numerical weather model.
prediction (NWP), non-hydrostatic, mesoscale model, available with several advanced physics and numerical schemes, designed for better prediction of atmospheric processes. The model description and updates can be found from Skamarock et al. (2005) and the WRF user web page (http://www2.mmm.ucar.edu/wrf/users/).

The WRF model utilizes large-scale atmospheric forcing as input for initialization and lateral boundary condition. These large-scale conditions are regridded by the model domain, considering the grid spacing and local topographical as well as other terrain conditions. As is common for most WRF studies over the Indian region, the National Centers for Environmental Prediction (NCEP) global FNL (FNL) analysis dataset, based on Global Data Assimilation System (GDAS) with Global Forecast System (GFS), is considered. The FNL data are available at a coarse resolution of 1° × 1°, at 6 h intervals – 00:00, 06:00, 12:00 and 18:00 UTC (Coordinated Universal Time) – and is used to provide initial and boundary conditions to the model. The lateral boundary conditions in the WRF model are updated at 6 h inter-
vals. Considering the short duration of the run, the model was forced with fixed sea surface temperature (SST) throughout the integration, and no regional data assimilation is carried out. The land surface boundary conditions are taken from the Moderate Resolution Imaging Spectroradiometer (MODIS) International Geosphere–Biosphere Programme (IGBP) 21-category land use and land cover fields that are available with a horizontal grid spacing of 10 min. Three telescopically nested domains are used in this study – the parent domain (Domain 1) is fixed between 60 and 100° E with grid spacing of 27 km; the first nested domain (Domain 2a) covers 70–85° E, 22–37° N with 9 km grid spacing and is indicative of “global to regional scale” (G2R) downscaling, and the second nested domain (Domain 2b) covers 76–81.5° E and 28.5–34° N at 3 km grid spacing (Fig. 1), for global to convective scale (G2C) downscaling (Trapp et al., 2007). The parent domain provides lateral boundary conditions to the inner domains, resulting in the downscaling ratios for simulations as 1 : 3 and 1 : 9. The three domains use 30 vertical pressure levels, with the top fixed at 50 hPa. The model time steps were a function of grid spacing: 135, 45 and 15 s respectively for the three domains.

The model configuration used default parameterization options following Osuri et al. (2012). For example, shortwave radiation is based on Dudhia (Dudhia, 1989) and longwave radiation is based on the Rapid Radiative Transfer Model (RRTM; Mlawer et al., 1997) scheme. Other physical parameterization options such as microphysics, cumulus parameterization schemes, planetary boundary layer and land surface models were selected as outlined ahead. There is currently no known unique configuration that can best simulate an extremely heavy rainfall event. Therefore, based on the literature (e.g. Kumar et al., 2008; Hong and Lee, 2009; Misenis and Zhang, 2010; Mukhopadhyay et al., 2010; Argüeso et al., 2011; Cardoso et al., 2013; Efstathiou et al., 2013), four MP schemes, two CU schemes, two PBL schemes and two LSMs are considered to obtain an ensemble of rainfall simulations. The two PBL schemes considered are the Yonsei University (YSU) scheme (Hong et al., 2006) and the Mellor–Yamada–Janjic (MYJ) scheme (Janjić, 2001). YSU is a non-local scheme, wherein fluxes are calculated at a certain height in the PBL considering the profile of the entire domain. The MYJ scheme, however, is a local scheme in which fluxes are calculated at various heights within the PBL and are related to vertical gradient in the atmospheric variables at the same height. Further details regarding the difference between the YSU and the MYJ schemes can be obtained from Misenis and Zhang (2010) and Efstathiou et al. (2013). The two CU schemes considered are the Kain–Fritsch (KF) scheme (Kain, 2004) and the Betts–Miller–Janjic (BMJ) scheme (Janjić, 1994, 2000). The KF scheme is both a shallow and deep convection scheme, wherein shallow convection is allowed for updrafts that do not reach minimum precipitating cloud depth. This scheme is based on entrainment and detrainment plume model with updrafts and downdrafts of mass flux. Potential energy is removed in the convective timescale within this scheme. Furthermore, it includes cloud, rain, snow and ice detrainment at cloud tops. BMJ also considers convection at both shallow and deep levels. However, there is no updraft and downdraft of mass flux and no cloud detrainment. Domain 2b is configured without any CU scheme, assuming MP can explicitly solve the convection at the finer resolution (Sikder and Hossain, 2016). The four MP schemes considered are the Purdue Lin (PLin) scheme (Lin et al., 1983; Chen and Sun, 2002), theEta Ferrier (Eta) scheme (NOAA, 2001), the WRF Single-Moment 6-class (WSM6) scheme (Hong and Lim, 2006) and the Goddard scheme (Tao et al., 1989). Both the PLin scheme and the WSM6 scheme are based on the parameterization from Rutledge and Hobbs (1984) and have 6-class microphysics,
which includes the mixing ratios of water vapour, cloud water, cloud ice, snow, rain and graupel. The main difference between these two schemes is related to the treatment of ice-phase microphysical processes. Details of the PLin and the WSM6 schemes are available in Hong et al. (2009). The Eta scheme was designed primarily for computational efficiency in NWP models, wherein the total condensate and the water vapours are directly advected into the model. The Goddard scheme is a slight modification from the PLin scheme for ice-water saturation. In general, all the MP schemes are known to influence the rainfall simulations at fine grid resolution by influencing the water phase component (Li et al., 2017). Since each physics scheme is associated with a distinct feature, it is important to examine the effect of their interactions on the rainfall simulations.

The sensitivity of various WRF configurations in simulating heavy rainfall events is assessed using the Noah LSM (Chen and Dudhia, 2001; Ek et al., 2003; Tewari et al., 2004). The Noah LSM is a community model that is included in the WRF suite with the prime aim of providing reliable boundary conditions to the atmospheric model. As a result, Noah LSM is a moderately detailed model, which includes a single canopy layer with the canopy resistance scheme of Noilhan and Planton (1989) and four soil layers (at 0.1, 0.3, 0.6 and 1.0 m), with a total soil depth of 2 m. The last soil layer of 1 m acts as a reservoir for drainage of water under gravity and the above three layers serve as root-zone depths. There is a provision in the model to allow default root-zone depths to be replaced with the actual values from the field, subject to data availability. In the Noah LSM, surface (skin) temperature is obtained using a single linearized surface energy balance equation, which effectively considers the ground and vegetation surface. Frozen soil parameterization based on Koren et al. (1999) and the surface runoff scheme of Schaake et al. (1996) are also included in this model. Soil moisture, soil temperature, water intercepted by the canopy and snow stored on the ground are also included as the prognostic variables in the model. More detailed information on the Noah LSM can be obtained from Ek et al. (2003).

To assess the effect of the land surface scheme on simulations, the Noah LSM is replaced with the simple five-layer soil model (Slab; Dudhia, 1996). In contrast to the relatively sophisticated Noah LSM, Slab is based on simple thermal diffusion in the soil layers that has constant soil moisture availability but a prognostic soil temperature term (Dardorff, 1978). Further differences between the two LSMs are presented in Sect. 3.1.3.

Table 3 provides the summary of the WRF physics schemes considered to simulate the extremely heavy rainfall events.

The ability of the WRF model configuration to simulate an extreme rainfall event is evaluated by comparing the simulated rainfall with the observations through indices such as scale error (SE), which is the ratio of standard deviation of model simulations to the observed standard deviation, and coefficient of variation (CV) in addition to MAE, RMSE and $\beta$.

3 Results and discussion

3.1 Sensitivity analysis

3.1.1 Verification of WRF simulations

Figure 5 presents cumulative rainfall for 15–18 June 2013 from 16 WRF simulations (4 MP, 2 CU and 2 PBL) corresponding to each of the three domains. From Fig. 5i to iii, it may be seen that the spatial pattern of rainfall appears to be sensitive to the microphysics, i.e. PLin, Eta and WSM6 MP schemes, while the amount of rainfall is more dependent on the PBL and CU scheme options. There is a considerable difference in the rainfall amount simulated with the Goddard MP scheme option compared to other MP schemes. Further, most of the model runs are able to reproduce the spatial gradient in the rainfall amount, which is perhaps primarily due to the topographical variation in the region. For locales below 1000 m, observations show distinctly lower rainfall compared to the high elevation regions (>1000 m). Further, distinct clusters corresponding to heavy rainfall event are observed in the north-east and north-west areas of the study region. These clusters are found to be consistent with the TMPA data; however, due to lack of surface rain gauge observations, the amount of rainfall in these regions could not be verified at this stage. Incidentally, the observed heavy rainfall event in the south-east part of the region is seen in a few WRF configurations, such as configuration (b) and (c). In general, WRF-simulated rainfall fields show a similar spatial pattern as that of the TMPA rainfall product. However, the magnitude of WRF rainfall is significantly high compared to TMPA and is attributed to the negative bias in TMPA for heavy rains. Figure 6 summarizes the comparison of WRF rainfall with rain gauge observations, accumulated over the 4-day period (15–18 June 2013) for the three domains. For comparison, grid points from the WRF domains closest to the gauge location are considered.

Figure 6 indicates that Domain 1 captures rainfall within the range of 150 to 400 mm for most of the WRF configurations. For Domain 2a and Domain 2b, an increase in the predicted rainfall amount is noted, particularly for small rainfall thresholds. Further, the WRF runs still under-predict extremely heavy rainfall and each of the configuration considered (across all the three domains) underestimated the rainfall amount more than 400 mm. However, the underestimation of rainfall is less in Domain 2b (G2C scale) compared to others, indicating the necessity of finer grid spacing as the first-order requirement for simulating the magnitudes of the extremely heavy rainfall events. The bias in the WRF simulations is typically due to a number of interactive factors: (i) scale feedback between mesoscale convection and...
Figure 5. Spatial plots showing rainfall estimates obtained for (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b. Arrows on the left indicate the PBL scheme, arrows on the right represent the CU scheme and the top arrows represent the MP scheme considered for the simulation runs. Subpanels (a) to (p) are the WRF configurations; for instance, (a)–(i) represent the WRF configuration with the YSU PBL scheme, KF CU scheme and PLin MP scheme for Domain 1. Refer to Appendix A (Table A1) for further details of WRF configurations (a) to (p).

Figure 7 indicates that there is more error at the stations Dehradun and Haldwani, which received higher rainfall. The highest rainfall obtained in different WRF configurations for these stations was less than 500 mm and this underestimation is highlighted in the error statistics. The model results show large-scale processes within the model (Bohra et al., 2006), (ii) lack of local observations that can add mesoscale features (Osuri et al., 2012; Osuri et al., 2015), (iii) lack of proper land surface processes (Niyogi et al., 2006; Chang et al., 2009; Osuri et al., 2017a), and (iv) the inability of the model to fully resolve the complex topography (Argüeso et al., 2011; Cardoso et al., 2013; Chevuturi et al., 2015). To assess the performance of the WRF simulations, quantitative scores (MAE and RMSE) with respect to the observed data are computed for daily rainfall data, which is then averaged over the 4-day period. The results are shown in Fig. 7. The last column in the figure presents the spatially averaged values obtained for different model configurations.

Figure 7 indicates that there is more error at the stations Dehradun and Haldwani, which received higher rainfall. The highest rainfall obtained in different WRF configurations for these stations was less than 500 mm and this underestimation is highlighted in the error statistics. The model results show
Table 3. Configuration of the WRF model considered for simulation of rainfall.

<table>
<thead>
<tr>
<th>Model options</th>
<th>Dataset or value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domains</td>
<td>3</td>
</tr>
<tr>
<td>Grid resolution (spacing)</td>
<td>27; 9; 3 km</td>
</tr>
<tr>
<td>Downscaling ratio</td>
<td>1:3; 1:9</td>
</tr>
<tr>
<td>Projection system</td>
<td>Mercator</td>
</tr>
<tr>
<td>Land surface boundary condition</td>
<td>21-class MODIS</td>
</tr>
<tr>
<td>Initial conditions</td>
<td>NCEP FNL</td>
</tr>
<tr>
<td>Shortwave radiation scheme</td>
<td>MM5 shortwave or Dudhia</td>
</tr>
<tr>
<td>Longwave radiation scheme</td>
<td>Rapid radiative transfer model (RRTM)</td>
</tr>
<tr>
<td>PBL Schemes</td>
<td>1. Yonsei University (YSU)</td>
</tr>
<tr>
<td></td>
<td>2. Mellor–Yamada–Janjic (MYJ)</td>
</tr>
<tr>
<td>Cumulus schemes</td>
<td>1. Kain–Fritsch (KF)</td>
</tr>
<tr>
<td></td>
<td>2. Betts–Miller–Janjic (BMJ)</td>
</tr>
<tr>
<td>Microphysics schemes</td>
<td>1. Lin (Purdue)</td>
</tr>
<tr>
<td></td>
<td>2. Eta (Ferrier)</td>
</tr>
<tr>
<td></td>
<td>3. WSM6</td>
</tr>
<tr>
<td></td>
<td>4. Goddard</td>
</tr>
<tr>
<td>Surface layer option</td>
<td>Monin–Obukhov similarity theory</td>
</tr>
<tr>
<td>Land surface models</td>
<td>1. Simple 5-layer soil model (Slab)</td>
</tr>
<tr>
<td></td>
<td>2. Noah</td>
</tr>
</tbody>
</table>

Figure 6. Scatter plots between the rainfall data from the rain gauges and the WRF simulations for (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b for WRF configurations (a) to (p). Refer to Appendix A (Table A1) for the detailed list of the WRF configurations.

higher error and variability in the simulations for the northern part of the domain compared to the southern part. This is likely due to the complex terrain in the northern part of the domain.

To identify the “best” and the “worst” performing configurations, temporal errors (wherein MAEs across the 18 locations are summed up) and spatial errors (wherein MAEs obtained over the entire region for the 4 days are summed up) for all the three domains are obtained (Appendix B), which indicate that the configuration (b), with YSU PBL, KF CU and Eta MP, produces maximum error, whereas configuration (p), with MYJ PBL, BMJ CU and Goddard MP, gives minimum error.

To further assess the sensitivity of configuration (p) and configuration (b) in capturing the extreme rainfall events in the region, additional simulations pertaining to other heavy to extremely heavy rainfall events (as mentioned in Table 2) are conducted. Spatial plots showing the cumulative rainfall estimates obtained for the three domains in comparison to the observed IMD gridded data and the TMPA data are presented in Appendix C. To summarize the performance of configuration (p) and configuration (b) against the observations (IMD gridded data), spatio-temporal MAE values are computed, which are presented in Table 4.

From the analysis conducted over the additional rainfall events, it is noted that configuration (p) gives less error in
Figure 7. Root mean square error (top panels) and mean absolute error (bottom panels) computed temporally for (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b for WRF configurations (a) to (p). Refer to Appendix A (Table A1) for the detailed list of the WRF configurations.

Figure 8. Scale error (SE) in WRF configurations (a) to (p) for 18 locations in the UGB for (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b. Refer to Appendix A (Table A1) for the detailed list of the WRF configurations.
comparison to the configuration (b) for all the rainfall events. This makes configuration (p) with MYJ PBL, BMJ CU and Goddard MP the “best” at simulating the spatial and temporal variability of the extremely heavy rainfall over the upstream region of the UGB. Why this combination emerged as the best performing is an intriguing but difficult question to address at this stage. Note that the rainfall prediction is the combination of many nonlinear, interactive factors including the behaviour of each configuration and cannot be realistically studied with the sparse rainfall data and absence of vertical sounding observations. Some possible factors that could contribute would be that local boundary formulation in MYJ may be more appropriately capturing the vertical environment in the complex terrain compared to the non-local YSU scheme which seeks to simulate vertical mixing and boundary layer evolution using averaged and grid representative fields. With regard to the BMJ CU emerging in the top configuration, there are a number of studies for the ISMR where it has emerged as performing “overall best” (Vaidya and Singh, 2000; Ratnam and Kumar, 2005; Vaidya, 2006; Rao et al., 2007; Kumar et al., 2010; Mukhopadhyay et al., 2010; Srinivas et al., 2013; Sikder and Hossain, 2016). As for the MP scheme, there are limited studies in comparison to those that have studied the CU configuration for the ISMR. Further, the MP scheme performance has been evaluated for tropical cyclone cases because of the warm versus cold pool processes that are critical in the simulation of the cyclone intensity. Of those available in the literature, studies such as Sing and Mandal (2014) found that the Goddard scheme has a “slightly better” performance than other schemes. This conclusion is also supported by studies such as Choudhury and Das (2017) and has been used in hailstorm studies such as Chevuturi et al. (2014).

The impact of the downscaling ratio on the rainfall simulations is addressed next. On comparing the simulations of the June 2013 event from G2R and G2C domains with the rain gauge data, it is noted that the former gives less error for most of the locations (Appendix D). The G2C scale has a large resolution (grid spacing) gap from the outer to the inner domain in comparison to G2R, which could result in less accurate initial and lateral boundary conditions and, consequently, more simulation errors in G2C. Another possibility is that the metric being used, which is the rainfall observation from in situ data, is itself more conservative with regards to the grid in which rainfall occurs in the coarser domain and may slightly favour the G2R. However, on reviewing the overall structure of rainfall fields and the amounts across the domain, results suggest that the G2R scale with a moderate downsampling ratio may be better suited for simulation of the extreme rainfall event as in the present case study. The results are found to be consistent with other studies, such as that of Liu et al. (2012), wherein the moderate ratio of 1 : 3 is found to perform best. However, it is to be noted that errors corresponding to the grid point nearest to the rain gauge are considered here for comparison. The result may vary upon selection of another grid point.

3.1.2 Impact of different parameterization schemes

Although configuration (p), with MYJ PBL, BMJ CU and Goddard MP, appears to be the “best” physics configuration for the study region, significant variability exists among the simulations pertaining to different configurations of the WRF model. This variability causes significant uncertainty across different runs, which is quantified through computation of SE and CV (Figs. 8 and 9) for the June 2013 event. Deviation in model simulations with respect to observed data provides the SE; however, CV gives variation within different model simulations.

Most of the model configurations have an SE value clustered around 1 (Fig. 8), indicating that the variability in simulated rainfall is similar to the observed rainfall. However, variability in the north-eastern part of the domain is observed to be high compared to others. The same is reflected in the CV plot (Fig. 9), wherein grid points around the Chamoli station (on the north-eastern side) have CV between 41 and 51%, whereas stations closer to Uttarakashi and Tehri have values ranging between 11 and 30%. Further, grid points closer to Dehradun have low CV values, which could be due to the models consistently underestimating the rainfall in this subdomain. The southern part of the region, which received low rainfall, also exhibited high variability. In general, it can be inferred that uncertainty in rainfall is greater in the north-eastern part compared to the north-west. The regions that received very high or very low rainfall during this period also displayed higher uncertainty. Uncertainty in rainfall simulations varies between the domains, with Domain 2b having maximum uncertainty. This could be attributed to high variability in the simulated values at higher spatial resolution.

Since consideration of different parameterizaiton schemes is the reason for variability in rainfall simulations, it is of interest to understand how the former influences the model output. For this, the average cumulative rainfall over the region and across different configurations is considered. The differences between various configurations are evaluated to assess the influence of PBL, CU and MP parameterization.
schemes on the rainfall simulations. Results for the same are presented in Fig. 10.

It is noticeable that, in general, the WRF configurations with the KF convective scheme produce rainfall of higher magnitudes. This result is consistent with previously conducted studies (Gallus, 1999; Fonseca et al., 2015; Pieri et al., 2015). For PLin MP, it is noted that considering YSU PBL along with KF CU scheme has a synergistic effect, leading to the maximum amount of rainfall over the region. This additive effect could be attributed to the YSU being a non-local scheme, making it suitable for convective, unstable PBL conditions (Bright and Mullen, 2002). Upon changing the PBL scheme (from YSU to MYJ), and maintaining the convective scheme as KF, a notable difference in the fields is simulated (as shown by red circles in Fig. 10). This difference obtained for changing the PBL (with PLin MP and KF CU) is found to be equivalent to the case when only the CU is changed (from KF to BMJ) under the YSU PBL scheme (shown by blue plus sign in Fig. 10). Similarly, the difference in rainfall obtained for two cases – changing the PBL (from YSU to MYJ) with PLin MP and BMJ CU; and changing the CU (from KF to BMJ) with PLin MP and MYJ PBL is also found to be approximately equal. This indicates that the average cumulative rainfall values obtained under two configurations – PLin MP, BMJ CU and YSU PBL and PLin MP, KF CU and MYJ PBL are almost equal. Further, BMJ CU, irrespective of the PBL scheme, results in less simulated rainfall across the region. For WSM6 MP, within the YSU PBL scheme, changing the CU (from KF to BMJ) parameterization produces significant variability (displayed by blue plus sign) in rainfall than changing the PBL scheme itself (from YSU to MYJ PBL with KF CU). However, with MYJ PBL, the effect of changing the CU scheme is insignificant (yellow star in Fig. 10). Furthermore, with BMJ the difference in rainfall produced due to changing the PBL is minimal. With Eta and Goddard MP, changing the PBL (irrespective of the CU scheme) produces a negligible difference in the rainfall simulations, as represented by red circles and purple diamonds in Fig. 10. However, changing the CU schemes (irrespective of the PBL condition) is seen to have a significant influence on rainfall. It can be concluded from this section that the relationship between PBL and CU is interlinked, wherein YSU and MYJ PBL complement (contradict) the effect of KF (BMJ) and BMJ (KF) CUs with regard to the quantity of rainfall. Overall, the choice of CU appears to have a significant impact on the simulation of rainfall over the region. This conclusion is consistent with the earlier studies such as those by Sikder and Hossain (2016), where they found ISMR to be more sensitive to CU than to MP.

3.1.3 Impact of land surface boundary condition

It is well established that the soil moisture plays a significant role in weather predictions (Chen and Brutsetaert, 1995; Betts et al., 1996, 1997; Entekhabi et al., 1996). Therefore, Slab and Noah LSMs, which differ significantly in two factors – (i) the soil depths along with the inclusion of land surface processes and (ii) the temporal evolution of soil moisture – are selected in the present study to assess the impact of land surface conditions on simulation of heavy rainfall events.

The Slab is a relatively simple LSM with five soil layers (at 1, 2, 4, 8 and 16 cm depths) and uses a thermal diffusion equation to compute surface fluxes based on a surface temperature and drag coefficient formulations. The Noah LSM is modestly detailed (compared to Slab) with four soil layers (at 10, 30, 60 and 100 cm depths) and explicit representation of land surface parameters, which includes the effect of soil moisture changes, snow cover, evapotranspiration and hydrologic processes such as runoff and drainage (to subsurface layers). Further, in the Noah LSM soil moisture and temperature are prognostically computed for each of the four soil layers, whereas in the Slab LSM only soil temperature is prognostic and moisture is considered as a constant value based on the land use. Slab lacks the feature of predicting the snow cover and does not capture the evaporation and runoff processes over the region.

To understand the influence of each LSM on the rainfall estimation, simulations using the Slab LSM are conducted for the “best” and the “worst” performing configurations from the Noah LSM case (configuration (p) and (b), respectively). Results comparing the two land-model-based runs are presented in Fig. 11. The top panel in Fig. 11i–iii presents the scatter plot of the cumulative rainfall obtained with Noah and Slab LSM runs versus the observed cumulative rainfall. Bottom panel (Fig. 11i–iii) presents the mean...
Figure 10. Difference in simulated rainfall due to PBL, CU and MP parameterization schemes corresponding to (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b over the UGB region.

Figure 11. Scatter plot (top panels) and mean difference in rainfall (bottom panels) for the observed rainfall data and the WRF simulations (for (b) and (p) configurations) pertaining to Noah and Slab LSMs corresponding to (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b.
difference in the simulated and the observed rainfall values over the 4-day period for 18 locations along with the spatially averaged values (last columns in Fig. 11i–iii) within the region.

The Slab-LSM-based run significantly underestimates the rainfall in comparison to the Noah LSM. For example, the locations which recorded rainfall greater than 400 mm have the Slab-LSM-based simulated values in the range of 100–150 mm. As stated earlier, although the Noah LSM also underestimated the rainfall for such stations, the bias with the Noah LSM is significantly less than the Slab LSM (−26 % with the Noah LSM, in contrast to −64 % with the Slab LSM for Domain 2a and the (p) configuration). Further, the mean difference in rainfall obtained with Slab LSM is found to be higher in comparison to the Noah LSM. This is essentially due to significant underestimation of rainfall during 16 and 17 June 2013 by the Slab LSM.

In a number of studies, the differences in the surface energy fluxes simulated by the choice of different LSMS (i.e. Slab versus Noah) has been discussed (see Niyogi et al., 2016 for a review). The main reason being that the surface processes affect the boundary layer feedbacks which in turn create zones of mesoscale convergence that can affect the location and intensity of convection. These convective systems then contribute to the simulated rainfall. The results obtained in this study emphasize this feature with differences in the rain amounts and locations in response to the change in LSM. The better performance of using the Noah model could be attributed to the temporal evolution of soil moisture fields. Analysing the soil moisture in Slab and Noah models, the soil is noted to be relatively dry in Slab (soil moisture less than 0.05 m$^3$ m$^{-3}$) and the value is constant throughout the model run (since in the Slab model there is no prognostic soil moisture term). In the case of Noah, soil moisture varies in response to the rainfall and is found to vary between 0.25 and 0.45 m$^3$ m$^{-3}$. Higher surface moisture conditions improve mass flux, convective updrafts and diabatic heating in the boundary layer that contributes to low-level positive potential vorticity or convective potential, which leads to enhanced rainfall potential (Osuri et al., 2017a). The importance of representing soil moisture variability over India for extreme weather conditions is also highlighted through this work.

### 3.2 Comparison between rainfall from the WRF and the FNL dataset

Simulated rainfall from the WRF model runs is assessed with respect to the NCEP FNL reanalysis dataset. To achieve this, it is necessary to bring both the datasets to a common spatial resolution. Therefore, the WRF-simulated rainfall is up-scaled, through averaging of the grids, to match the resolution of NCEP FNL (1° × 1°). For the analysis, only simulations pertaining to the “best” performing configuration (p), are considered. Bias (β) in rainfall simulations from the two datasets corresponding to 18 rain gauge locations is obtained, results for which are presented in Fig. 12.

From Fig. 12 it can be observed that the NCEP FNL data overestimate the rainfall for most of the locations. Upon dynamic downscaling of the FNL data through the WRF model, rainfall simulations improved over the UGB region. Locations such as Mussorie, Pauri and Roorkee, which have shown β between 2.5 and 3.5 in the FNL data, reduced to 0–0.25 in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations. Dehradun along with three stations from the south-eastern region, (Mukteshwar, Haldwani and Nainital), which recorded heavy rainfall (Sect. 2.1), are observed to have a small bias in the FNL data, show similar small bias in the WRF simulations.

From the above analysis, it is evident that the WRF model can simulate extreme precipitation better than the reanalysis data. This can be attributed to increase in spatial resolution, and better representation of surface and meteorological features, with respect to the lateral boundary conditions as suggested in some of the previous works such as those by Argüeso et al. (2011), Mishra et al. (2014), Giorgi and Gutowski, (2015), and Singh et al. (2017).

### 4 Summary and conclusions

The main focus of this paper is to provide a general guideline for setting up the WRF model configuration to simulate heavy rainfall events. In this regard, sensitivity of the WRF model to predict heavy to extremely heavy rainfall events is examined through (a) quantitative verification of the rainfall simulated by the WRF model, (b) investigating sensitivity of
the simulated rainfall to different parameterization schemes, downscaling ratios and land surface models, (c) testing the selected scheme for other rainfall events, and (d) assessing the effect of local and global factors by comparing the simulated rainfall with a global reanalysis dataset.

For the analysis, an extremely heavy rainfall event, which occurred from 15 to 18 June 2013, over the Ganga Basin, in the foothills of the Himalayas in the Uttarakhand State of northern India is considered. Most of the studies conducted earlier over this region (Medina et al., 2010; Kumar et al., 2012; Thayyen et al., 2013; Kumar et al., 2014; Chevuturi et al., 2015; Shekhar et al., 2015; Chevuturi and Dimri, 2016; Rajesh et al., 2016; Hazra et al., 2017) are based on the general or default WRF configuration with WSM6 microphysics, a Kain–Fritsch cumulus parameterization scheme and planetary boundary layer of the Yonsei University scheme. In this paper, ensemble experiments are conducted using the WRF model with different grid spacing, four microphysics schemes, two cumulus parameterization schemes, two planetary boundary layer schemes and two land surface model conditions. The rainfall simulations are evaluated against the observed rain gauge data and the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis precipitation data. The WRF configuration with Goddard microphysics, a Mellor–Yamada–Janjic planetary boundary layer condition and a Betts–Miller–Janjic cumulus parameterization scheme is found to perform “best” in simulating an extremely heavy rain event of June 2013. The selected configuration is then verified by simulating several other heavy to extremely heavy rainfall events that occurred across different months in the monsoon season over the upstream region of the UGB. The results for the additional events indicate that the selected configuration (Goddard microphysics, Mellor–Yamada–Janjic planetary boundary layer condition and Betts–Miller–Janjic cumulus parameterization scheme) is indeed the “best” in simulating the spatial and temporal variability of the extremely heavy rainfall over the region. Therefore, through the exhaustive analysis conducted in this paper, the recommended WRF configuration for extreme rainfall simulations in the Himalayan region is Goddard microphysics, Mellor–Yamada–Janjic planetary boundary layer condition and Betts–Miller–Janjic cumulus parameterization scheme.

Although complex interactions are observed between different physics options, microphysics schemes can be seen to influence the spatial pattern of the rainfall, while the choice of cumulus scheme is found to modulate the magnitude of the simulated rainfall. Upon analysing the impact of downscaling ratios on rainfall simulations, it is concluded that downscaling from global to regional scale with a moderate downscaling ratio may give the least model errors and, thus, be considered as suitable for reproducing the extreme rainfall event. In addition to this, the effect of land surface models on rainfall simulations is also assessed in this paper. The Slab LSM significantly underestimates the rainfall values, and incorporating Noah helped improve the performance. The underperformance of the Slab model is attributed to dry soil conditions in the region for this LSM.

In addition to the sensitivity experiments, the WRF-simulated rainfall is also compared with the NCEP FNL reanalysis data. The NCEP FNL data are found to overestimate the rainfall, whereas the WRF-simulated rainfall exhibited less bias. The comparison results indicated that care must be taken while employing global datasets for regional analysis. Through this, it can be established that the rainfall values obtained from the high-resolution mesoscale model can be effectively used in hydrologic models for realistic streamflow estimates.

The analyses presented in this paper are subject to a few limitations: first, results are limited to the physics parameterization schemes considered in this paper, and may vary upon inclusion of other schemes; second, only two sets of downscaling ratios, i.e. 1:9 and 1:3, are tested in the current work. The sensitivity of simulations pertaining to other downscaling ratios should be tested in future; and third, only G2R and G2C sensitivity are assessed in this work.

Data availability. The meteorological data were purchased from the India Meteorological Department using the requisition form given at the following link: http://www.imd.gov.in/advertisements/20170320_advt_34.pdf (Pai et al., 2014). The station data for June 2013 event can be obtained from Ray et al. (2013). The following are the links for procuring the publicly accessible data: TMPA data – https://pmm.nasa.gov/data-access/downloads/trmm (Huffman, 2016); FNL data – https://doi.org/10.5065/D6M043C6 (NCEP, 2017). The source code for the WRF model, used in the study, can be obtained from http://www2.mmm.ucar.edu/wrf/users/downloads.html (Skamarock et al., 2005). The rainfall simulation obtained from the WRF model are available at https://github.com/ilachawla/Chawla_2018_HESS_Data (Chawla, 2018).
Appendix A

Table A1. List of different WRF configuration.

<table>
<thead>
<tr>
<th>WRF configuration</th>
<th>Microphysics scheme (MP)</th>
<th>Cumulus scheme (CU)</th>
<th>Planetary boundary layer scheme (PBL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>PLin</td>
<td>KF</td>
<td>YSU</td>
</tr>
<tr>
<td>b</td>
<td>Eta</td>
<td>KF</td>
<td>YSU</td>
</tr>
<tr>
<td>c</td>
<td>WSM6</td>
<td>KF</td>
<td>YSU</td>
</tr>
<tr>
<td>d</td>
<td>Goddard</td>
<td>KF</td>
<td>YSU</td>
</tr>
<tr>
<td>e</td>
<td>PLin</td>
<td>BMJ</td>
<td>YSU</td>
</tr>
<tr>
<td>f</td>
<td>Eta</td>
<td>BMJ</td>
<td>YSU</td>
</tr>
<tr>
<td>g</td>
<td>WSM6</td>
<td>BMJ</td>
<td>YSU</td>
</tr>
<tr>
<td>h</td>
<td>Goddard</td>
<td>BMJ</td>
<td>YSU</td>
</tr>
<tr>
<td>i</td>
<td>PLin</td>
<td>KF</td>
<td>MYJ</td>
</tr>
<tr>
<td>j</td>
<td>Eta</td>
<td>KF</td>
<td>MYJ</td>
</tr>
<tr>
<td>k</td>
<td>WSM6</td>
<td>KF</td>
<td>MYJ</td>
</tr>
<tr>
<td>l</td>
<td>Goddard</td>
<td>KF</td>
<td>MYJ</td>
</tr>
<tr>
<td>m</td>
<td>PLin</td>
<td>BMJ</td>
<td>MYJ</td>
</tr>
<tr>
<td>n</td>
<td>Eta</td>
<td>BMJ</td>
<td>MYJ</td>
</tr>
<tr>
<td>o</td>
<td>WSM6</td>
<td>BMJ</td>
<td>MYJ</td>
</tr>
<tr>
<td>p</td>
<td>Goddard</td>
<td>BMJ</td>
<td>MYJ</td>
</tr>
</tbody>
</table>
Appendix B

Table B1. Mean absolute error (MAE) values corresponding to different WRF configuration for the three domains.

| WRF configuration | Domain 1 | | | Domain 2a | | | Domain 2b | |
|---|---|---|---|---|---|---|---|---|---|
| | Temporal | Spatial | | Temporal | Spatial | | Temporal | Spatial | |
| a | 1028 | 228 | | 943 | 210 | | 1066 | 237 | |
| b | 1227 | 273 | | 1097 | 244 | | 1249 | 278 | |
| c | 960 | 213 | | 1048 | 233 | | 1077 | 239 | |
| d | 825 | 183 | | 852 | 189 | | 908 | 202 | |
| e | 877 | 195 | | 885 | 197 | | 995 | 221 | |
| f | 1068 | 237 | | 886 | 197 | | 1055 | 234 | |
| g | 887 | 197 | | 950 | 211 | | 979 | 218 | |
| h | 885 | 197 | | 976 | 217 | | 1010 | 225 | |
| i | 932 | 207 | | 939 | 209 | | 1024 | 228 | |
| j | 1161 | 258 | | 970 | 215 | | 1131 | 251 | |
| k | 894 | 199 | | 888 | 197 | | 987 | 219 | |
| l | 883 | 196 | | 855 | 190 | | 918 | 204 | |
| m | 890 | 198 | | 913 | 203 | | 974 | 217 | |
| n | 1012 | 225 | | 870 | 193 | | 976 | 217 | |
| o | 837 | 186 | | 863 | 192 | | 976 | 217 | |
| p | 740 | 164 | | 843 | 187 | | 886 | 197 | |

Figure B1. Mean absolute errors in space corresponding to different WRF configurations for (i) Domain 1, (ii) Domain 2a and (iii) Domain 2b. Blue dotted lines present the “worst” performing configuration, i.e. configuration (b), and red dotted lines show the “best” performing configuration, i.e. configuration (p).
Figure C1. Spatial plots presenting the rainfall simulations obtained across the three domains for the best and the worst configuration for heavy to extremely heavy rainfall events during (a) Event 1 (18–22 June 2008), (b) Event 2 (29 July–2 August 2010), (c) Event 3 (15–19 August 2011), (d) Event 4 (17–21 September 2010) and (e) Event 5 (11–15 September 2012).
Figure D1. Bar plot representing the mean absolute errors in simulating rainfall across the 18 rain gauge locations on global to regional (G2R) scale (1:3) and global to convection-permitting (G2C) scale (1:9).
Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “The changing water cycle of the Indo-Gangetic Plain”. It is not associated with a conference.

Acknowledgements. The work is part of the study supported by the US National Science Foundation (CAREER AGS0847472) and the Government of India/IOES National Monsoon Mission (grant no./project no.: MM/SERP/CNRS/2013/INT-10/002) at Purdue University. The second author gratefully acknowledges the financial support of ESSO, Ministry of Earth Sciences, Government of India and Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India.

Edited by: Ana Mijic
Reviewed by: two anonymous referees

References


Huffman, G.: TRMM (TMPA-RT) Near Real-Time Precipitation L3 3 hour 0.25 degree x 0.25 degree V7, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), available at: https://disc.gsfc.nasa.gov/datasetcollection/TRMM_3B42RT_7.html (last access: 2 November 2017), 2016.


