Point 1: Regarding the question raised by Y. Zhang and Referee #2, the following sentences and references are added in the manuscript (see line 26-31, page 3) to address the “the start time of the model and spin-up” issue.

“The spin-up period is necessary for WRF model to develop the smaller scale convective features and the widely used lengths are 6 hours (Givati et al, 2012), 12 hours (Hu et al, 2010) and 24 hours (Wang et al, 2012). Different spin-up lengths were tried for the six storm events in this study, whereas results did not show obvious differences regarding the simulated rainfall. In order to improve the calculation efficiency for further hydrological use (i.e., flood warning), a 6h period is chosen to spin-up the model. That is to say the start of the model integration is 6 h earlier than the storm start time and the end time of the model integration is consistent with the storm end time.”

Givati, A., Lynn, B., Liu, Y., and Rimmer, A.: Using the WRF Model in an Operational Streamflow Forecast System for the Jordan River, J.Appl. Meteorol. Clim., 51, 285-299, doi:10.1175/JAMC-D-11-082.1, 2012.

Hu, X. M., Nielsengammon, J. W., and Zhang F.: Evaluation of three planetary boundary layer schemes in the WRF model. J. Appl. Meteorol. Clim., 49, 1831-1844, doi:10.1175/2010JAMC2432.1, 2010.

Wang, S., Yu, E., and Wang, H.: A simulation study of a heavy rainfall process the Yangtze River valley using the two-way nesting approach. Adv. Atmos. Sci., 29, 731-743, doi:10.1007/s00376-012-1176-y, 2012.

Point 2: Regarding the question raised by Y. Zhang, the following sentences are added in the manuscript (see line 24-25, page 3) to address the “integration step of WRF model” issue.

“The integration step of WRF follows the ‘6×dx’ rule where dx is the grid spacing, and the integration step is 6s for innermost domain (Skamarock and Klemp, 2008).”

Point 3: Regarding the question raised by Y. Zhang, the following sentences are added in the manuscript (see line 12-13, page 6) to address the “insignificant precipitation” issue.

“It should be mentioned that the insignificant precipitation (less than 0.1 mm/h) is regarded as no rain.”

Point 4: Regarding the question raised by Y. Zhang, two references are added in the manuscript (see line 22-23, page 3) to support the “40 vertical layers with 1 km horizontal resolution” issue.

Aligo E.A., GallusW.A., Segal M., 2009. On the impact of WRF model vertical grid resolution on Midwest summer rainfall forecasts. Weather Forecast. 24, 575-594.

Qie X., Zhu R., Yuan T., Wu X., Li W., Liu D., 2014. Application of total-lightning data assimilation in a mesoscale convective system based on the WRF model. Atmos. Res. s145–146, 255-266.

Point 5: Regarding the question raised by Professor Chu, two references are added in the manuscript (see line 6-7, page 6) to support the “Thiessen polygon method” issue.

Sivapalan, M., and Blöschl, G.: Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves, J. Hydrol., 204, 150-167, doi:10.1016/S0022-1694(97)00117-0, 1998.

Jarvis, D., Stoeckl, N., and Chaiechi, T.: Applying econometric techniques to hydrological problems in a large basin: Quantifying the rainfall–discharge relationship in the Burdekin, Queensland, Australia, J. Hydrol., 496, 107–121, doi: 10.1016/j.jhydrol.2013.04.043, 2013.

Point 6: Regarding the suggestion raised by Professor Chu and Referee #2, the following sentences and references are added in the manuscript (see line 25-33, page 12 and line 1-2, page 13) to discuss the importance of the rainfall forecasts for the flood forecasts.

“Ultimately, the main goal of rainfall forecasts is to obtain efficient flood forecasts. The peak flood, flood peak appearance time, flood process are all significantly influenced by the rainfall accumulations and the spatiotemporal distribution of the rainfall (Schellekens et al, 2011; Cane et al, 2013; Fan et al, 2015). Event V which occurred on 21 July 2012 has caused the greatest flood during the past 10 year in Jing-Jin-Ji (Beijing-Tianjin-Hebei) area and received widespread attention in China. The 24 h rainfall accumulation was 155.43 mm in Zijingguan catchment and the peak flow reached 2580 m^3/s at the catchment outlet. In such cases, accurate rainfall simulations and predictions can do great help to flood warning. However, to analyze the usefulness of the WRF simulations to flood warning, the rainfall-runoff transformation processes should be further considered. This will involve many uncertainties, such as the choice of the rainfall-runoff model, the data used for model calibration, and the involvement of a real-time updating scheme, etc., which also have considerable impact on the accuracy of the flood forecasting results. The exploration of different parameterisations from the flood warning purposes is an important issue and worthy to be discussed in further study.”

Schellekens, J., Weerts, A. H., Moore, R. J., Pierce, C. E., and Hildon, S.: The use of MOGREPS ensemble rainfall forecasts in operational flood forecasting systems across England and Wales, Adv. Geosci., 29, 77-84, doi:10.5194/adgeo-29-77-2011, 2011.

Cane, D., Ghigo, S., Rabuffetti, D., and Milelli, M.: Real-time flood forecasting coupling different postprocessing techniques of precipitation forecast ensembles with a distributed hydrological model. The case study of may 2008 flood in western Piemonte, Italy, Nat. Hazard. Earth Sys., 13, 211-220, doi:10.5194/nhess-13-211-2013, 2013.

Fan, F. M., Collischonn, W., Quiroz, K. J., Sorribas, M. V., Buarque, D. C., and Siqueira, V. A.: Flood forecasting on the Tocantins River using ensemble rainfall forecasts and real-time satellite rainfall estimates, J. Flood Risk Manag., 9, 278-288, 2015.

Point 7: Regarding the suggestion raised by Professor Chu, the equation (7)-(9) are modified as:

$$RMSE = \frac{\sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(P_j - Q_j\right)^2}}{\frac{1}{M}\sum_{j=1}^{M}Q_j} \times 100\% \tag{7}$$

$$MBE = \frac{\frac{1}{M}\sum_{j=1}^{M}\left(P_j - O_j\right)}{\frac{1}{M}\sum_{j=1}^{M}Q_j} \times 100\% \tag{8}$$

$$SD = \frac{\sqrt{\frac{1}{M-1}\sum_{j=1}^{M}\left(P_j - O_j - MBE\right)^2}}{\frac{1}{M}\sum_{j=1}^{M}Q_j} \times 100\% \tag{9}$$

Point 8: Regarding the question raised by Referee #1 and Referee #2, the following sentences and references are added in the manuscript to address the “cumulus parameterizations” issue according to the new Table 1.

Table 1. The constitution of the WRF physical ensemble

Member ID	Microphysics	PBL	Cumulus parameterisation
1	Lin	YSU	KF
2	WSM6	YSU	KF
3	Lin	MYJ	KF
4	WSM6	MYJ	KF
5	Lin	YSU	GD
6	WSM6	YSU	GD
7	Lin	MYJ	GD

8	WSM6	MYJ	GD
9	Lin	YSU	BMJ
10	WSM6	YSU	BMJ
11	Lin	MYJ	BMJ
12	WSM6	MYJ	BMJ
13	***Lin***	***YSU***	/
14	***WSM6***	***YSU***	/
15	***Lin***	***MYJ***	/
16	***WSM6***	***MYJ***	/

Line 11-18, page 4: "Some studies indicated that the cumulus parameterizations may be invalid with fine horizontal resolutions, while the threshold of the resolution is unknown (Argüeso et al, 2011; Evans et al, 2012; Pei et al, 2014). Many studies use cumulus parameterizations with the about 1km resolution for weather simulation. For example, Shepherd et al (2016) explored the sensitivity of hurricane track to four cumulus parameterizations, including KF, BMJ, G-3 and TD, with the nested domains 1.33km, 4km and 12km. Remesan et al (2015) studied the WRF model sensitivity to the choice of parameterizations: 4 nested domains (1km, 3km, 9km and 27km) are used and the cumulus parameterizations of GD, BMJ, KF1 and KF2 are investigated. In order to make the study more rigorous, member 13, 14, 15 and 16 are also tested and compared with the members containing cumulus parameterizations."

Line 8-14, page 8: "According to Table 5, the four members without cumulus parameterization have quite different performance for different event. For example, member 15 performs the best for event IV, nevertheless, it performs the worst for event V. Comparing with the members containing cumulus parameterization, member 13, 14, 15 and 16 have no significant advantages or significant disadvantages for rainfall simulation. Taking event I as an example, the best one (member 16) in the four members without cumulus parameterization ranks fourth in the 16 members, whereas the worst one (member 13) ranks twelfth. However, few members without cumulus parameterization rank in the top 4, which means that it is necessary to use cumulus parameterization for the simulation of rainfall accumulation."

Line 24-28, page 9: "From Fig. 4 and 5, it can be easily found that few values of the indices for member 13, 14, 15 and 16 are out of the range of the values for other 12 members, which indicates that there are always some members performing better than the four members without cumulus parameterization. It is helpful to use appropriate cumulus parameterization for the simulation of the spatial rainfall distribution."

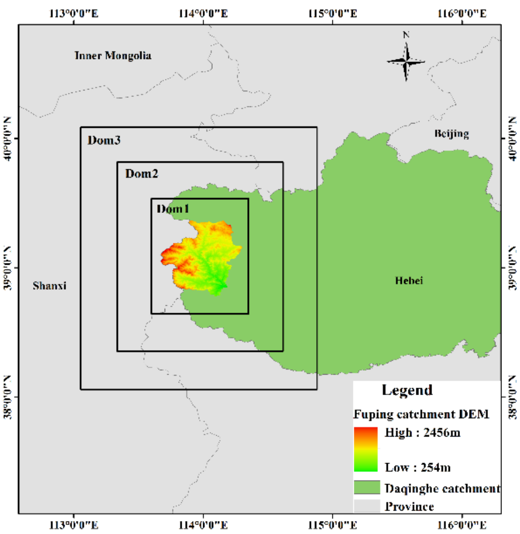
Line 31-21, page 10 and line 1-2, page 11: "From Fig. 7 and 8, the same as the conclusions in the spatial dimension, most values of the indices for member 13, 14, 15 and 16 are in the range of the values for other 12 members, which indicates that there are always some members performing better than the four members without cumulus parameterization. It is also necessary to use cumulus parameterization for the simulation of the temporal rainfall distribution."

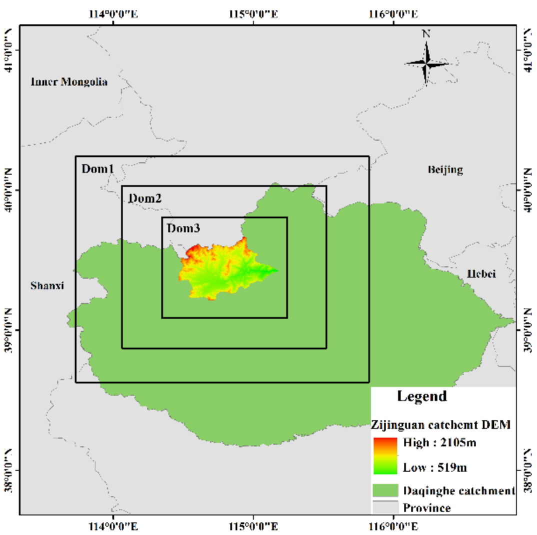
Line 11-13, page 12: "Additionally, it is necessary to use cumulus parameterization for the simulation of the rainfall accumulation and spatiotemporal rainfall distribution in the study area. However, the threshold of the horizontal resolution need to be further discussed for whether to use the cumulus parameterization."

Point 9: Regarding the suggestion raised by Referee #1, "ensemble N" should be modified as "the member N of the ensemble".

The nomenclature is revised accordingly in the manuscript.

Point 10: Regarding the suggestion raised by Referee #1, "Plots showing the domains and orography" should be added in the manuscript. The following two figures are added in the manuscript.





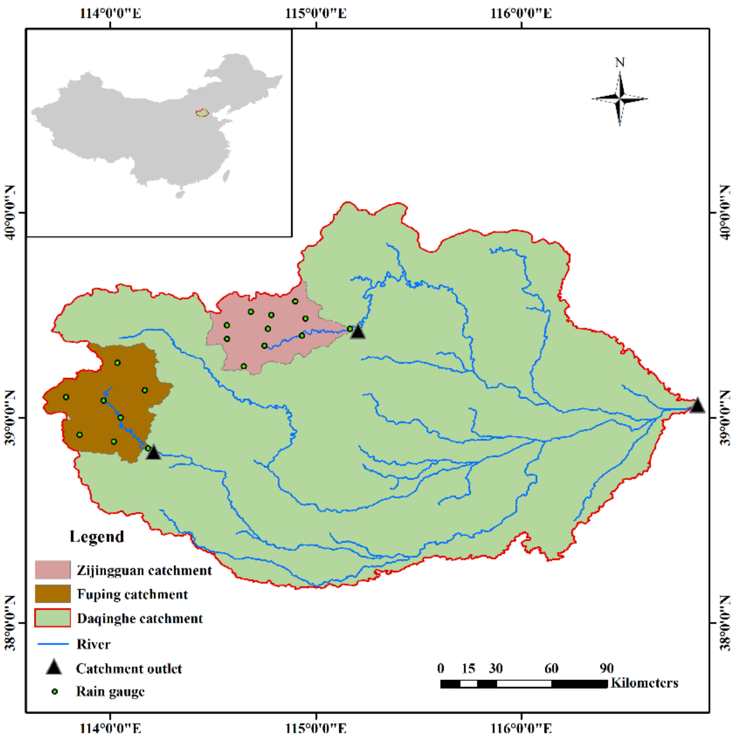
Point 11: Regarding the question raised by Referee #1, the climatic region should be

clarified clearly.

The two catchments in the study are located in semi-humid climatic region, which has been clarified clearly in the manuscript.

Point 12: Regarding the question raised by Referee #1, Fig 1 should show the location of the rain gauges.

The location of rain gauges are added in Fig 1.



Point 13: Regarding the question raised by Referee #1 and Referee #2, the following sentences are added in the manuscript (see line 15-21) to address the "Cv" issue.

"In order to learn the spatial and temporal evenness of the rainfall in the two catchments, both spatial and temporal Cv of the storm events from 1985 to 2015 are all calculated. In reality, rainfall in Northern China is much more uneven than the south and it is impossible to find absolute even rainfall in both space and time. So we chose a threshold of 5%, which is also considered in other statistical analyses in the same area, as the critical value to separate even and uneven rainfall events. With the threshold, we found the two critical values of 0.4 for the spatial Cv and 0.6 for the temporal Cv. That is to say, the storm events with the spatial Cv below 0.4 or with the temporal Cv below 1.0 account for 5% of the total storm events from 1985 to 2015 in the study area.”