A Comparison of Perturbed Initial Conditions and Multiphysics Ensembles in a Severe Weather Episode in Spain

FRANCISCO J. TAPIADOR

Department of Environmental Sciences, University of Castilla-La Mancha, Toledo, Spain

WEI-KUO TAO AND JAINN JONG SHI

Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland

CARLOS F. ANGELIS

Instituto Nacional de Pesquisas Espaciais, Centro de Previsao de Tempo e Estudos Climaticos, São Paulo, Brazil

MIGUEL A. MARTINEZ, CECILIA MARCOS, AND ANTONIO RODRIGUEZ

Agencia Estatal de Meteorologia, Madrid, Spain

ARTHUR HOU

Goddard Modeling Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

ABSTRACT

Ensembles of numerical model forecasts are of interest to operational early warning forecasters as the spread of the ensemble provides an indication of the uncertainty of the alerts, and the mean value is deemed to outperform the forecasts of the individual models. This paper explores two ensembles on a severe weather episode in Spain, aiming to ascertain the relative usefulness of each one. One ensemble uses sensible choices of physical parameterizations (precipitation microphysics, land surface physics, and cumulus physics) while the other follows a perturbed initial conditions approach. The results show that, depending on the parameterizations, large differences can be expected in terms of storm location, spatial structure of the precipitation field, and rain intensity. It is also found that the spread of the perturbed initial conditions ensemble is smaller than the dispersion due to physical parameterizations. This confirms that in severe weather situations operational forecasts should address moist physics deficiencies to realize the full benefits of the ensemble approach, in addition to optimizing initial conditions. The results also provide insights into differences in simulations arising from ensembles of weather models using several combinations of different physical parameterizations.

1. Introduction

Localized heavy precipitation episodes are common in the Mediterranean Sea region. Providing better forecasts for such episodes is fundamental to deliver early warnings and thus to minimize personal losses. However, the task of forecasting and simulating precipitation is plagued with difficulties. First, precipitation is the consequence of several atmospheric factors including 3D temperature structure (Mahrt 2000), evaporation (Choudhury 2000), convective available potential energy (CAPE; Lorenz 1978), availability of condensation nuclei (Hudson and Li 1995), and orography modeling (Roe 2005). Uncertainties in the measuring of any of those processes propagate to the precipitation forecast, affecting model skill.

Second, all weather forecasts are affected by the defining characteristic of nonlinear systems, namely sensitivity to initial conditions (SIC). Two twin (arbitrarily close) initial conditions will soon generate different forecasts, so

Corresponding author address: Francisco J. Tapiador, University of Castilla-La Mancha, Institute of Environmental Sciences. Avda. Carlos III s/n, 45071 Toledo, Spain.
E-mail: francisco.tapiador@uclm.es

regardless of how accurate the estimates of the actual atmospheric state are, the forecasts are going to be different. This impedes generating deterministic forecasts, which have been replaced by probabilistic forecasts.

The third major factor affecting precipitation forecast skill is the incomplete treatment of some key physical processes in numerical models. Even if perfect measurements were available, both numerical weather prediction (NWP) models and regional climate models have to include parameterizations to account for unresolved physics at the model resolution (Lenderink et al. 2007). Processes such as small-scale turbulence, which operates at the Kolmogorov scale, cannot be formalized as explicit equations into the model as that would require unrealistic computing power. (In this particular example, eddies in the atmosphere are as small as a few millimeters in diameter, so mesoscale models have to use Reynolds averaging to make the problem numerically tractable.) In the case of precipitation, estimates are directly dependent on at least the parameterizations used for turbulence, radiation, land physics, cloud microphysics, and convection (Khain et al. 2005).

While the first problem is difficult to address, the second one, the effects of SIC into the forecast, can be mitigated by running the model with several different initial conditions. The spread of the ensemble provides an indication of the predictability of the atmospheric situation and of the uncertainty of the forecast, whereas the ensemble mean can be considered as the least biased estimate of the true state of the atmosphere. This perturbed initial conditions (PIC) ensemble approach is routinely employed by major operational centers such as the National Centers for Environmental Prediction (NCEP) or the European Centre for Medium-Range Weather Forecasts (ECMWF).

The effect of parameterizations is also approached in a similar way by running several models with different configurations. Providing that each individual model is better than any other in at least one aspect, it is accepted that the resulting multiphysics parameterization (MPP) ensemble is more skillful than the individual runs.

The sensitivity of modeled precipitation to both PIC and MPP ensemble approaches has been explored by several authors (Schwartz et al. 2010; Gallus and Bresch 2006; Jankov et al. 2005, to name but a few). Ensembles built using only initial perturbations generally have insufficient dispersion (Jankov et al. 2005), but large differences have been reported depending on parameterizations, dynamical cores, season, and location. Therefore, the objective of this paper is to compare both PIC and MPP strategies for a limited set of parameterizations, a single dynamical core, one episode, and in Spain. To that end, two sets of simulations—the PIC and MPP runs—were performed.

In the case of the MPP runs, the number of possible combinations between parameterizations is very high, so to focus on the role of microphysics, land model, and cumulus parameterizations, a limited set of options was explored. The use of different schemes for the planetary boundary layer (PBL) also has a significant effect on precipitation forecasts (Shin and Hong 2011; Parodi and Tanelli 2010), so here a single scheme is used to isolate the effects of microphysics, land model and cumulus parameterizations. For the same reason, the same numerical core was used in all the simulations, as it is known that differences in the numerics of the dynamic cores result in small differences in the initializations.

There has been several previous works on PIC and MPP parameterization ensembles using the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008). Clark et al. (2008) found among other results that the ensemble variance in their MPP ensemble was greater than that in a PIC ensemble during the first 24 h of the forecast. In another experiment comparing MPP and a stochastic kinetic-energy backscatter scheme, Berner et al. (2011) argued for a combination of both approaches. Vich and Romero (2010) followed a different approach using a MPP superensemble and ECMWF forecasts as initial data, obtaining good forecast skill for heavy precipitation episodes in the Mediterranean.

In this paper, the comparison is done within the constraints posed by limited validation data, as described in the next section below.

2. Data and experimental design

While the comparison of the MPP and PIC ensembles is the primary focus of this paper, ground radar and satellite rainfall estimates are used to provide a qualitative estimate of the simulations rainfall performance. No quantitative validation is attempted because of the lack of quality data in the study area, as described below.

a. Ground radar data

The available radar data for this study corresponds to the standard operational product from the Spanish Meteorological Agency (AEMET) network of 15 radars covering continental Spain and the Balearics. In this product, the raw reflectivity from the retrievals is calibrated, quality controlled, and filtered to provide the best-guess surface rainfall estimate. A fixed, Marshall and Palmer reflectivity–rainfall (Z–R) relationship is used across the year regardless of the type of precipitation, so this product is only useful to provide a qualitative estimate of actual rainfall rates. While the available single-polarization radars can provide an adequate estimate of the storm location and the spatial structure of precipitation fields,

Ground radar precipitation
(24h accum.)

WRF Domain
25 km grid space runs

Satellite precipitation
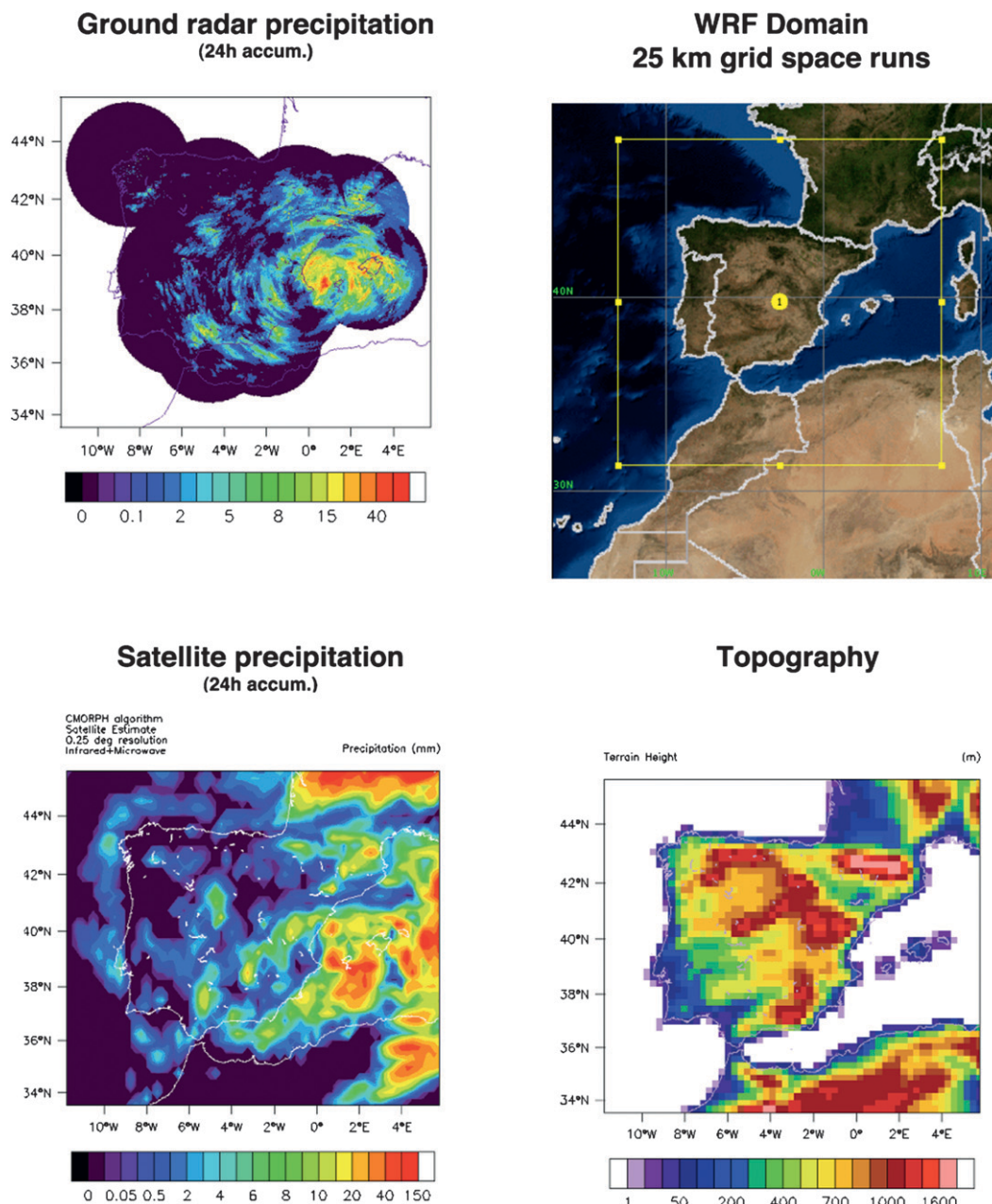### (24h accum.)

Topography

FIG. 1. Ground radar precipitation for 11 May 2008, WRF simulation domain, satellite precipitation estimate, and model topography.

their quantitative estimates are subject to many error sources and uncertainties. To name but one, the parameterization of the raindrop size distribution greatly affects the retrieval of precipitation from measured reflectivities.

The 15-min radar data were aggregated to a 24-h accumulation (Fig. 1). As some individual radars did fail for some minutes along the episode, a small bias is expected in the daily totals. Radar data do not compare well with gauge data for this particular episode. Thus, Barcelona

(41.28°N, 2.06°E) recorded 104.9 (radar ~ 15) mm, Valencia (39.5°N, 0.46°W) 30.4 (25) mm, Reus (41.15°N, 1.16°E) 57.9 (10) mm, Palma (39.55°N, 2.73°E) 13.9 (15) mm, and Valladolid (41.65°N, 4.76°W) 17.7 (5) mm for 11 May 2008. Most of those values (notably Barcelona) are at variance with radar estimates, illustrating either the effects of local factors in ground-measured precipitation, the limited quality of the radar estimates, or both issues. Given those shortcomings, radar data are used only qualitatively.
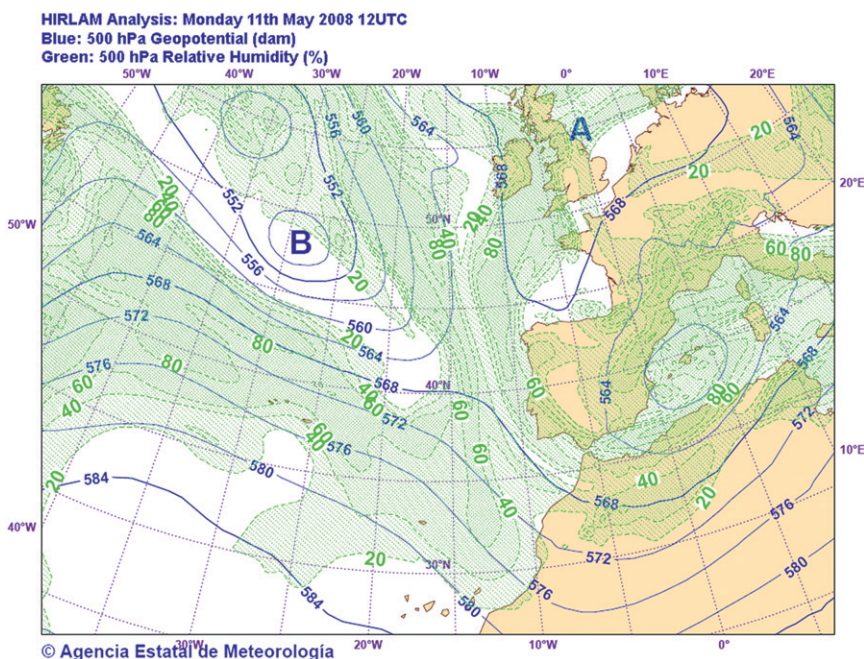
FIG. 2. Synoptic situation for the case study (11 May 2008). Here, B indicates a low center and
A indicates a high center.

b. Satellite data

Satellite data were used to have another independent estimate of precipitation to qualitatively cross compare with models and radar estimates. While satellite precipitation products are algorithm dependent and cannot be used for validation, they are informative of the large-scale spatial distribution of precipitation and thus can give an impression on model performance. The use of satellite products is also justified in this case as it allows one to extend the comparisons beyond the limited coverage of the ground radar.

The precipitation algorithm chosen was the Climate Prediction Center morphing technique (CMORPH; Joyce et al. 2004). The rationale of this algorithm consists in advecting microwave rainfall estimates along infrared-derived cloud motion winds. Those winds are then used to propagate the precipitation estimate provided by microwave (MW) sensors, thus generating a merged, high-temporal-resolution product. In this way, the more-frequent IR data are used to fill the gaps between more-direct MW overpasses. The spatial resolution of the NCEP's CMORPH product used here is 0.25°.

As with the radar operational data, the CMORPH satellite estimates are accumulated into daily totals for the qualitative comparison (Fig. 1). As in Gallus and Bresch (2006), it was found that correlation coefficients with these data did not appear to add meaningful insight into the sensitivity of forecasts and are not used.

Quantitative comparison with available ground data epitomizes the limitations of satellite algorithms, especially in coastal cities such as Barcelona, Valencia, Reus, and Palma, where the satellite greatly underestimates precipitation.

c. Model setup

The WRF model (Skamarock et al. 2008) was used to carry out the simulations. The version used corresponds to 2.2.1, which was modified to include the National Aeronautics and Space Administration Goddard Space Flight Center's microphysics. WRF model has the key advantage of having been extensively tested over a variety of locations and has been validated over many meteorological regimes and conditions. Another reason to prefer this model over the plethora of NWP models currently available is the vast number of scientists collaborating in its improvement, which results in an extensive set of well-documented, stable, and fully tested parameterizations.

The case selected (11 May 2008) is one with both high rain rates and full radar coverage of the episode. The synoptic situation (Fig. 2) corresponds to the dominance of an upper-level high over the British Isles and large portions of central Europe. A large storm developed over the east coast of Spain when a weakening upper-level cutoff low over the Balearic Islands remained almost stationary. At its southern flank, a jet streak with

40 m s^{-1} at 300 hPa moved across northern Algeria (not shown). Underneath the upper low, steep midlevel lapse rates and rich boundary layer moisture led to instability at the southwestern flank of the upper high. Some multicells developed over the east coast of the Iberian Peninsula and the Balearic Islands where large hail was reported.

The comparisons were performed on 11 May 2008 when good quality radar data were available. To allow for sufficient spinup time, the model was initialized 24 h earlier (on 10 May). This 24-h period is more than double the time WRF requires to generate a realistic kinetic energy spectra (Skamarock et al. 2008). The next 24 h (11 May 2008) was those actually used for comparison. The grid space of the simulations is 25 km.

The initial and boundary conditions for the simulations were from the Global Forecast System. A single domain (Fig. 1) was designed to minimize boundary issues. Given the computational burden of high-resolution ensemble simulations, the runs were carried out at the Barcelona Supercomputing Center's Mare Nostrum computer (94.21 Teraflops; flops = floating-point operations per second) using a parallel architecture on 256 processors. The use of more processors (512) did not show good scaling properties.

d. Model parameterizations

Four sets of parameterizations: microphysics, cumulus scheme, land physics, and radiation were analyzed. The MPP simulations were run for all possible combinations except for those not allowed by the model, resulting in 56 different runs. The PIC ensemble was made of 49 members. The rationale of the parameterizations is described in the following sections. The reader is directed to Skamarock et al. (2008) for a full account of the WRF physics.

1) PRECIPITATION MICROPHYSICS

Precipitation microphysics is used to model the processes resulting in the several forms of precipitation (rain, snow, hail, etc.). We used the Kessler (Kessler et al. 2006), Purdue–Lin (Rutledge and Hobbs 1984) [implemented as in Chen and Dudhia (2001)], WRF single-moment 3-class (WSM3; Hong et al. 1998, 2004), WRF 5-class (WSM5; Hong et al. 1998, 2004), WRF 6-class (WSM6; Lin et al. 1983), Goddard (Shi et al. 2010), and Thompson (Thompson et al. 2004) schemes.

The Kessler scheme is a simple warm cloud scheme that includes water vapor, cloud water, and rain. The physics processes used are production, fall, and evaporation of rain; accretion and autoconversion of cloud water; and the production of cloud water from condensation.

The Purdue–Lin scheme (after the Purdue cloud scheme and Lin) takes a step forward by adding cloud ice, snow, and graupel. This scheme includes saturation adjustment and ice sedimentation.

The WSM schemes differ in the species of precipitation modeled. The WSM3 includes vapor, cloud water–ice and rain–snow with the differences between the liquid and the solid phase dependent on temperature. In WSM5, water and rain are treated separately from ice and snow. WSM6 differs from WSM5 in that it adds graupel to the other species, and in differences in the calculations, improved calculation of heat profile, optimized order of processes to decrease sensitivity to the time step of the model, and separated treatment of ice and water saturation processes.

The Thompson scheme uses six classes of moisture species and is deemed as suitable for a variety of situations. Known problems of this parameterization include overestimation of snow and graupel and underestimation of ice in outflow regions. This scheme includes improvements in the modeling such as the use of the Gamma distribution instead of the exponential and dependent intercept parameters for the raindrop size distribution.

The new Goddard bulk microphysics (Shi et al. 2010) was also incorporated into the simulations before been implemented in the next WRF version. This scheme has three different options: cloud ice and snow; cloud ice, snow, and graupel (used here); and cloud ice, snow, and hail.

2) CUMULUS PARAMETERIZATION

Cumulus parameterizations are intended to model sub-grid updrafts and downdrafts in the models. The cumulus schemes used in the MPP simulations were Kain–Fritsch (K–F; Kain and Fritsch 1993), Betts–Miller–Janjic (B–M–J; Betts and Miller 1986; Janjic 1994), and Grell–Devenyi (G–D; Grell and Devenyi 2002). We also used no cumulus option in a set of simulations to ascertain the actual effect of the parameterizations.

The WRF version of the K–F parameterization is a simple cloud model with moist updrafts and downdrafts. The B–M–J is a convective-adjustment type scheme. It makes cloud efficiency dependent on the entropy change, precipitation, and mean temperature of the cloud. It has been reported that this scheme may overpredict areas of light precipitation (Jankov et al. 2005). The G–D approach uses an ensemble of cumulus schemes for each grid box. The ensemble members differ in updraft and downdraft entrainment and detrainment parameters and in precipitation efficiency. The value back to the model is a weighted average of the ensemble. Here, we used the mean of the ensemble.

3) LAND PHYSICS

Land surface models (LSM) deal with surface heat and moisture fluxes. While land parameterization has

traditionally been considered to have little influence on the short-term forecast of a situation dominated by a synoptic weather event [see Chang et al. (2009) for an example], two different parameterizations were used to quantify the actual influence in this severe weather episode. The schemes chosen were relatively complex parameterizations: a 5-layer thermal diffusion model with layers 1, 2, 4, 8, and 16 cm thick, and the Noah LSM (Chen and Dudhia 2001) scheme consisting of a 4-layer soil temperature and moisture model with canopy moisture and snow cover prediction. The latest parameterization includes surface emissivity properties and provides sensible and latent fluxes to the PBL scheme.

4) RADIATION AND PBL

Radiation schemes model the interactions of solar (shortwave) and thermal (longwave) radiation with the atmosphere. The modeling of these processes in WRF follows the plane-parallel simplification. Goddard shortwave (Chou and Suarez 1994) and the Rapid Radiative Transfer Model (RRTM) longwave (Mlawer et al. 1997) schemes were used in the runs, with climatological profiles for carbon dioxide, ozone, and trace gases. Regarding the PBL parameterization, the Mellor–Yamada–Janjic (M–Y–J) scheme (Mellor and Yamada 1982; Janjic 1994) was used.

e. Perturbed initial conditions ensemble

The ensemble members were constructed by adding perturbations to a base initial condition field. In this experiment only wind, temperature, and moisture fields were perturbed. The model physics involved in the ensemble forecast error covariance estimation consists of Goddard microphysics and radiation scheme, Noah surface scheme, Yonsei University PBL scheme and K–F cumulus scheme. The perturbations were generated following the ensemble forecast error covariance procedure in an ensemble data assimilation system (EDAS) WRF-EDAS developed by Goddard and Colorado State University (Zupanski et al. 2011).

3. Results and discussion

Two sets of simulations starting 10 May 2008 were performed at 25-km resolution. One set of runs was done for the MPP ensemble and another for the PIC ensemble. The period of interest was the 24–48-h interval corresponding to 11 May 2008.

Figures 3 and 4 show the daily accumulations of the two types of experiments using a poststamp plot. The area shown corresponds to that of the satellite and radar in Fig. 1. In the multiphysics case (Fig. 3), a numerical

key is used to identify the setup. In the PIC case (Fig. 4), every individual plot represents a member of the ensemble. As those are indistinguishable, there is no need for labeling each single run.

The visual comparison of these stamp plots with radar and satellite estimates (Fig. 1) shows that the main features of the daily precipitation are captured in both MPP and PIC ensembles. The fingerprint of the storm is apparent in the east coast of Spain and corresponds to the approximate location given by the satellite and the radar. The precipitation system in southern France is also modeled by most of the simulations, and the same applies to the system off the coast of Algiers. Some known effects in the MPP runs (Jankov et al. 2005) are apparent in the plots, including that the areal coverage of the precipitation in the no-cumulus runs is quite different than the coverage in those using a cumulus parameterization. The former setup exhibits heavier intensities when compared to cumulus parameterizations. At 25-km resolution, convection is not explicitly resolved but is dependent on the convective scheme used. Although comparing cumulus parameterizations with explicit microphysics is not straightforward, contrasting the outputs with and without this parameterization may allow one to isolate the effects of this parameterization.

While the internal variability within the PIC ensemble seems lower than in the MPP ensemble, quantitative estimates are required to ascertain the full extent of such variability. Thus, three types of quantitative metrics were used: the storm center location, used here as a mean to gauge differences in locating precipitation patterns; the semivariogram, used to analyze the spatial variability of the simulations; and the mean, maximum, and standard deviation of the daily precipitation accumulations in the whole domain, which is a simple but effective method of characterizing each setup with a few aggregated values.

The combined use of three complementary metrics is justified by noting that two simulations can forecast two very different storms in terms of mean and minimum values but with the same center. The same applies to spatial variability: two forecasts possessing very different spatial structures and clustering patterns may well have the same mean and maximum values and place the storm in the same location. The combination of the three estimates increases confidence in the simulations presenting the same precipitation pattern if their three metrics are similar.

a. Precipitation centroid location

A sensible way of quantifying model differences is to compare the estimates of the precipitation centroid across runs. This is a broad comparison over an area of
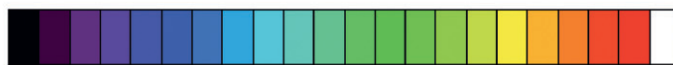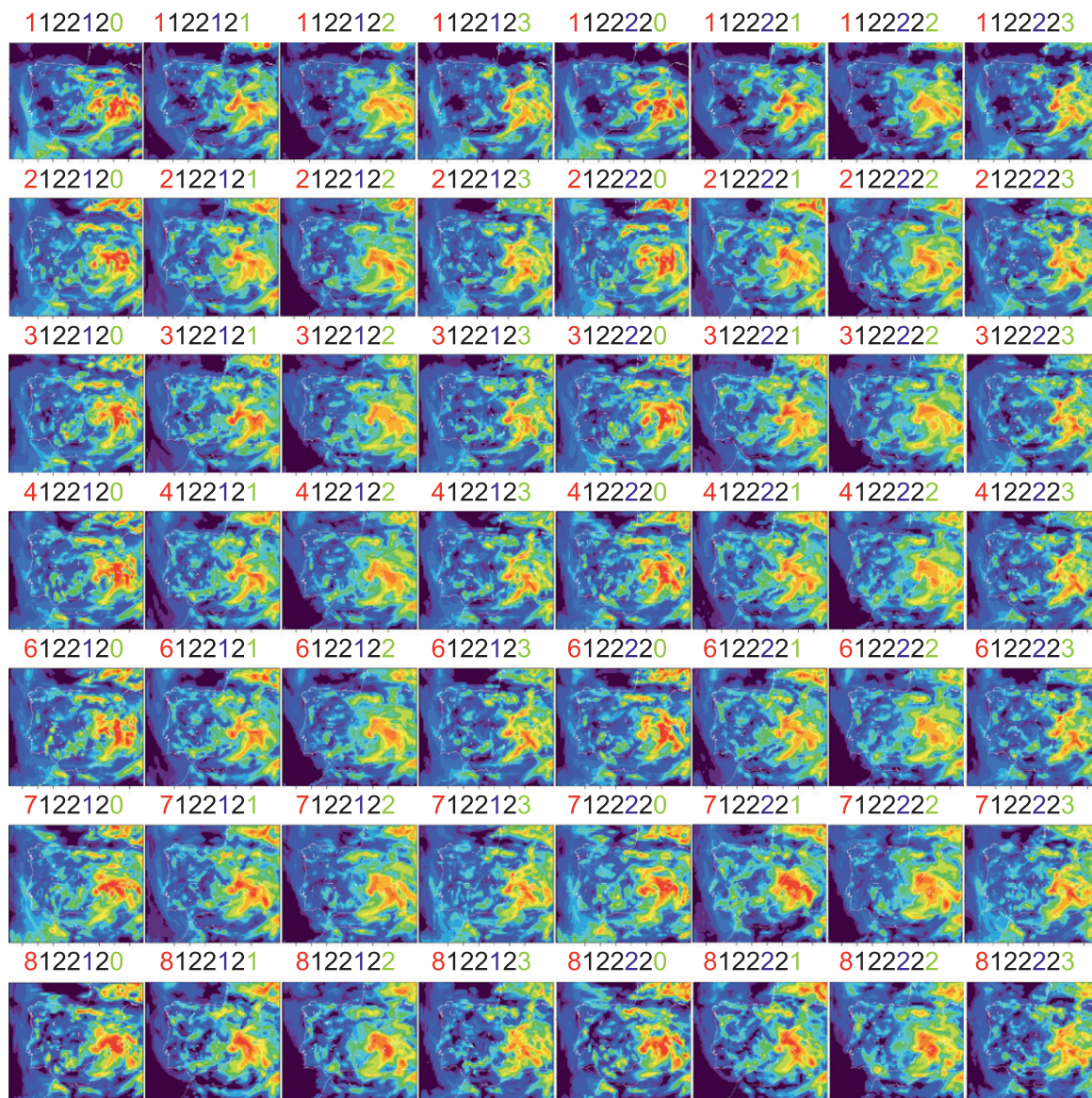
FIG. 3. Estimates of daily accumulation of precipitation for each one of the 56 MPP parameterizations. As described in the key below the coding is as follows: from left to right the seven digits correspond to 1) cloud microphysics (1: Kessler, 2: Lin et al., 3: WSM3 simple ice, 4: WSM5, 6: WSM6, 7: new Goddard, 8: Thompson et al.); 2) LW radiation (1: RRTM); 3) SW radiation (2: Goddard shortwave); 4) surface layer (2: Monin–Obukhov–Janjic); 5) land surface physics (1: thermal diffusion scheme, 2: Noah land surface model); 6) PBL scheme (2: M–Y–J); and 7) cumulus option [0: none, 1: K–F (new Eta), 2: B–M–J, 3: G–D]. The area covered is that of the satellite estimate in Fig. 1.
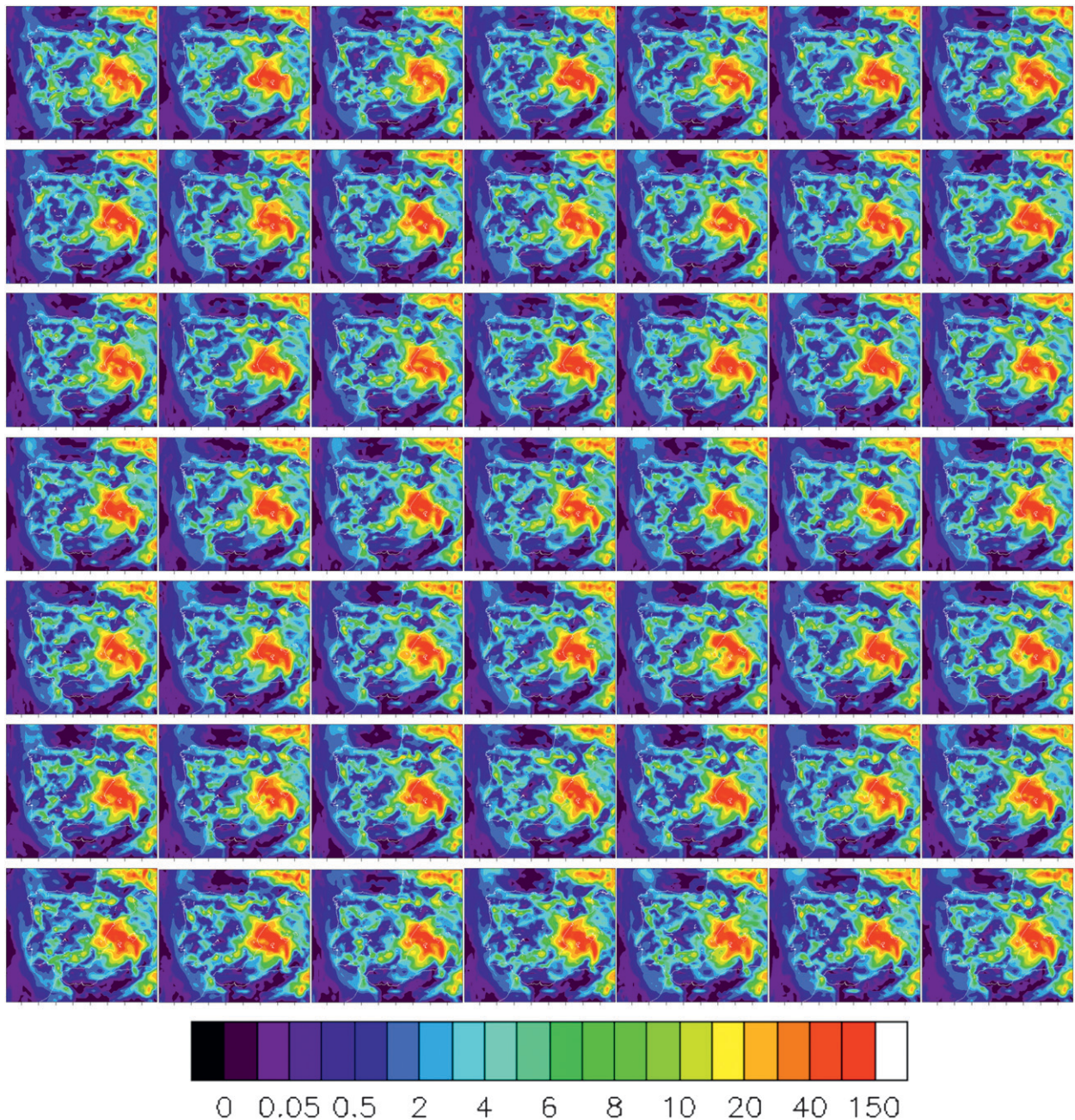
FIG. 4. Estimates of daily accumulation of precipitation for each one of the 49 ensemble members of the PIC ensemble.

hundreds of squared kilometers, thus being indicative of the ability of the individual simulations to capture the evolution of a mesoscale system.

The precipitation centroid was considered as a surrogate for the center of the storm. It was obtained from visual inspection of the radar data (39°N, 0°). This nominal centroid was used as the reference center for the WRF simulations. The centroid for each run was calculated not visually but as the weighted average of the

daily accumulated precipitation over a 4° × 4° window centered at 39°N, 0°. It should be noted that the aim here is to compare the dispersion of the storm center estimate across the runs so the actual method to derive the centroid is unimportant. Besides, several tests using various window sizes and centers gave minute absolute differences and no relative changes.

To compare the simulations, the *relative storm center location* was used. This is defined as the meridional and
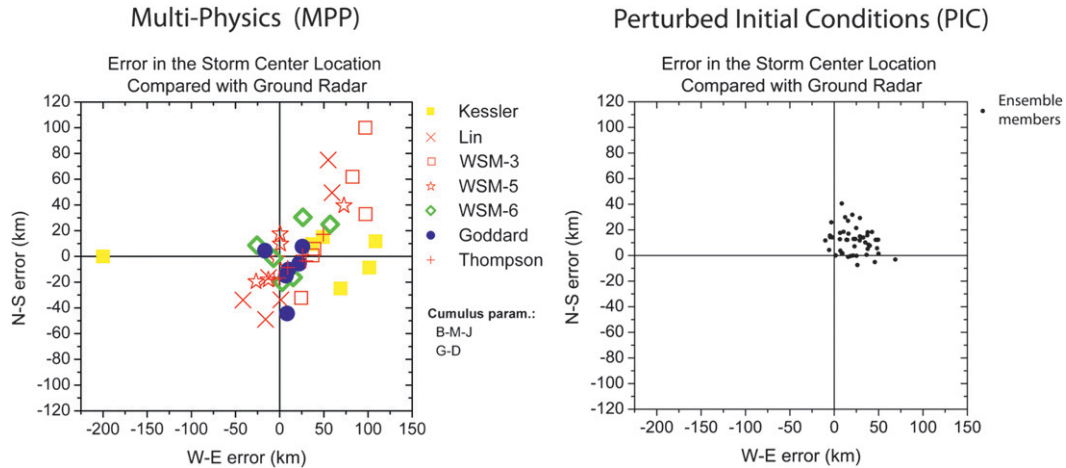
FIG. 5. Spread in the forecasting of the storm center location for the (left) MPP and (right) PIC ensembles.

zonal differences in kilometers from the calculated precipitation centroid to the nominal centroid.

Figure 5 shows the relative storm center location for the MPP (left) and the PIC (right) ensembles. The MPP plot shows that the new Goddard and the Thompson microphysics are closer than the others in estimating the nominal, visually identified storm center. Both microphysics are also consistent over the range of cumulus parameterizations. By contrast, Kessler microphysics presents highly variable estimates with differences as large as 200 km.

As mentioned above, the dispersion shown by the ensemble of parameterizations is as important as the individual performances in the ensemble approach. Leaving aside the Kessler parameterization, the remaining parameterizations give estimates that vary more than 100 km from the nominal value. The effect of the land parameterization is not included in Fig. 5 for clarity.

The contrast of these results with the PIC plots (Fig. 5, right) is apparent. Here, the estimates of the storm center are tightly clustered, which confirms the qualitative appraisal of fewer differences across simulations in the PIC ensemble than in the MPP ensemble. It is worth noting that this lower dispersion (high precision) may yield to severe forecasting errors if the center of the cluster completely miscalculates the storm center (low accuracy), and that the PIC ensemble approach would add no value in such a case.

b. Spatial structure

The empirical semivariogram is another complementary method to compare MPP and PIC simulations quantitatively. It expresses the change of the spatial variance with distance. The semivariance γ at d distance [$\gamma(d)$] is calculated as (Cressie 1993)

$$\gamma(d) = \frac{1}{2N(d)} \sum_{(i,j) \in N(d)} |p_i - p_j|^2,$$

where p_i and p_j are the precipitation at two different grid points at d distance, and $N(d)$ is the number of grid points at d distance. This function has been proved useful when comparing simulations in weather (Marzban et al. 2009) and climate (Tapiador 2010), as it provides a quantitative mean to gauge the spatial variability of the fields. Thus, spatially contrasting fields have a faster-growing semivariance with distance than less contrasted fields, which feature as a flatter line in a semivariogram plot.

Figure 6 illustrates such a clear difference between the MPP and PIC simulations. The figure also shows that in terms of the spatial structure of the precipitation fields, the semivariograms of the PIC simulation cluster around a run, so little new information in terms of spatial structure is added by this ensemble strategy. In contrast, the spread of the MPP semivariograms is far wider, which indicates more spatially different precipitation fields among MPP ensemble members. In other words, the MPP ensemble increases the variability of the members in terms of spatial structure. This is relevant as there is no a priori reason for this to happen. Precipitation fields from different ensemble strategies may well be different in terms of mean, maxima, or minima, but similar in terms of spatial structure (the converse is also true). Here, it is shown that there can be such a fundamental spatial difference between PIC and MPP ensembles.

c. Maxima, minima, and standard deviation

The analysis of the mean, maximum, and standard deviation of the mean daily precipitation runs confirms the
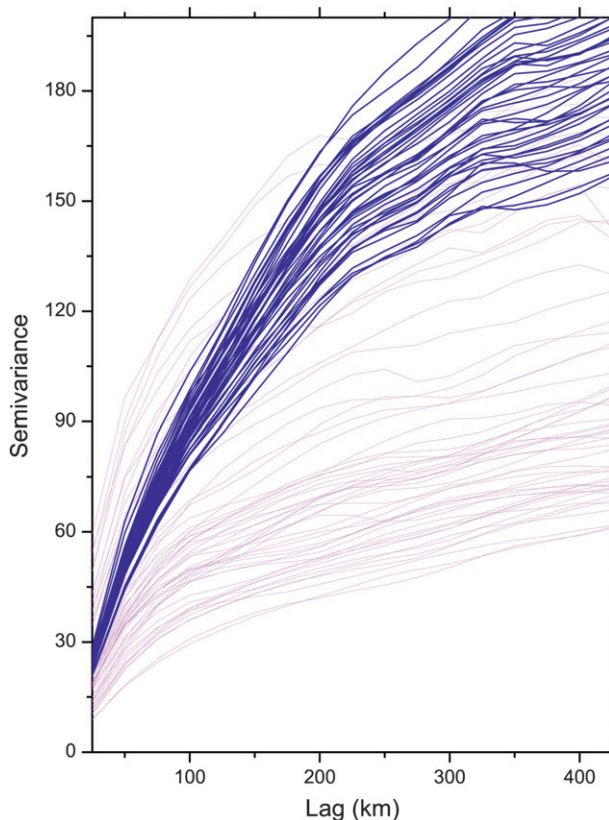
FIG. 6. Semivariogram of the PIC ensemble (blue lines) imposed over the semivariograms of the MPP ensemble (red lines).

previous observations also for these aggregated values. The 56 MPP estimates of the mean precipitation in the domain (Fig. 7, top-left) cluster around 3.5 mm day^{-1} (with a variability ranging from 2.25 to 4.5 mm day^{-1}), whereas the 49 runs of the PIC ensemble (Fig. 7, top-right) exhibit a much lower dispersion (from 3.9 to 4.25 mm day^{-1}). The maximum precipitation in the domain (Fig. 7, middle row) presents similar dispersion between the MPP and the PIC ensembles, while the standard deviation of the PIC ensemble (Fig. 7, bottom-right) shows in another way the same effect already seen in the semiovariogram (Fig. 6): the PIC ensemble tends to cluster around higher values than those in the MPP ensemble, thus indicating less global dispersion than in the MPP runs.

To relate those differences with different parameterization setups, the mean, maximum and standard deviation can be related to specific choices of the land use scheme (either thermal diffusion or Noah scheme), cumulus parameterization (none, K–F, B–M–J, or G–D) and microphysics (Kessler, Lin et al. WSM 3-class simple ice, WSM 5-class, WSM 6-class graupel, New Goddard, and Thompson et al.).

Figure 8 shows such dissagregated analysis. There is a clear effect in the three statistics of the cumulus parameterization and the microphysics, and to lower extent the land scheme. One overall observation is that for the same land and cumulus parameterizations, the mean and standard deviation of the precipitation fields given by Goddard microphysics closely follow Lin's. However, they differ in the maximum, which is generally much lower in Goddard's.

The three WSM microphysics methods behave quite similarly in most cases. Kessler microphysics tends to provide far lower mean, maximum, and standard deviation values, except when considering the maximum value with the K–F cumulus parameterization. This feature is more noticeable if Noah's land scheme is used (Fig. 8, middle), and is likely due to the warm-rain-only character of the K–F parameterization. It may be relevant, in terms of applying these results, to notice that an ensemble including models with Kessler microphysics would bias the mean precipitation toward lower values. The worst-case scenario in terms of expected differences is the combined use of Kessler and Lin microphysics with the wrong precipitation type. The performances of models including both types of microphysics critically depend on the ability to differentiate warm and cold clouds (Hong et al. 1999; Anagnostou and Kummerow 1997; Prabhakara et al. 2000; Steiner 2004; Tokay and Short 1996; Tokay et al. 1999) so a wrong choice could introduce large errors in the forecasts. The choice is less critical if Kessler microphysics is combined with schemes other than Lin's, but even in that case the differences are expected to be large compared with other combinations. In any case, it is critical to select the appropriated parameterizations for the case so the resulting ensemble is as realistic as possible.

The variability in the maximum precipitation is large, and sometimes unexpected for some equally sensible combinations of the three parameterizations. Thus, the maximum precipitation (Fig. 8, middle) given by the Lin microphysics with the B–M–J cumulus and Noah's parameterizations sharply differs from the maximum given by the other microphysics with the same cumulus option. Figure 8 also shows that, as expected, the effects of the two land physics setups are small. The plots are almost symmetrical for the thermal diffusion and for the Noah scheme, although there are differences.

The larger effect on the maximum, though, comes from the cumulus parameterization. It has been also observed that the K–F scheme permits more grid-resolved precipitation to occur and results in both isolated heavier amounts and more finescale structure than the B–M–J scheme (Gallus 1999). In the experiment reported here, the B–M–J scheme generates lower maximum values

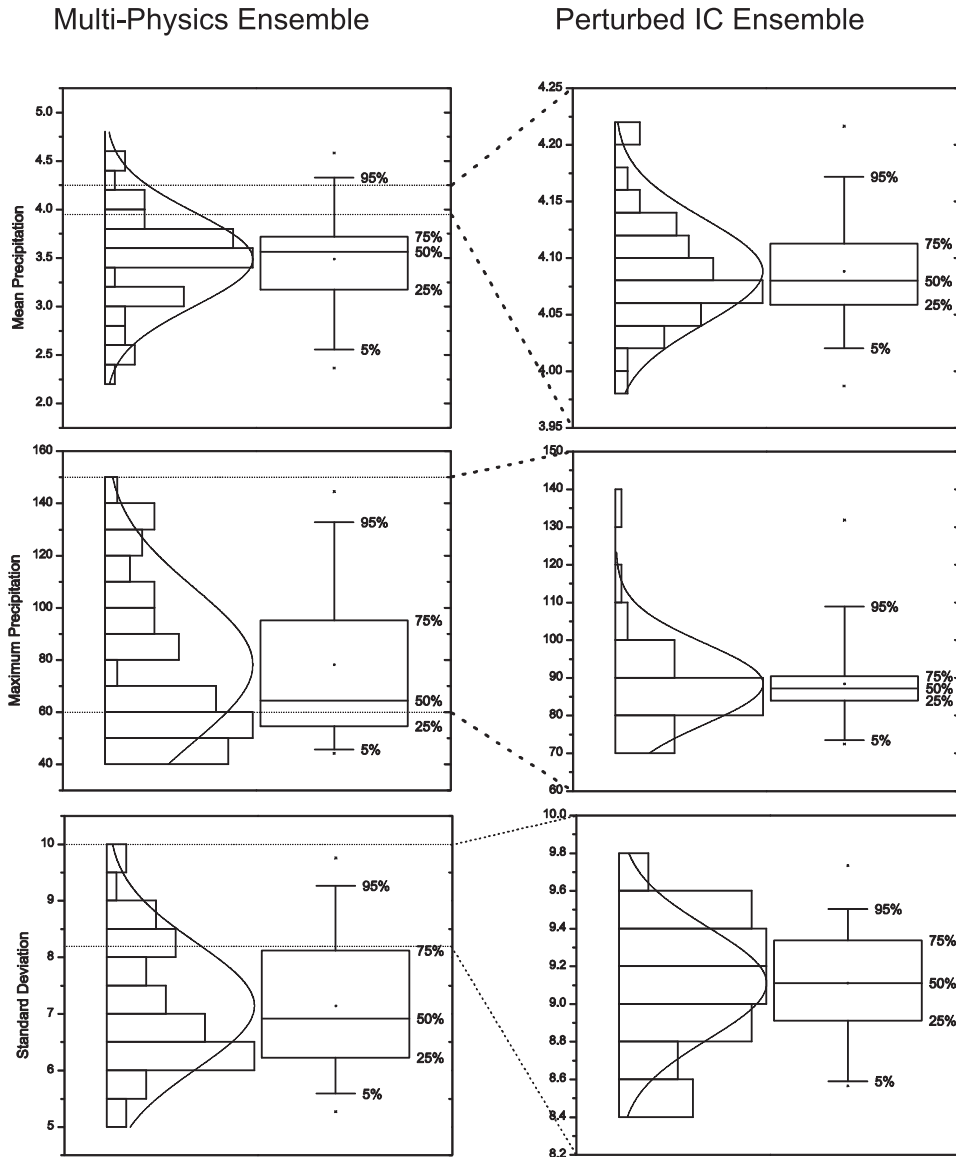## Multi-Physics Ensemble   Perturbed IC Ensemble

Fig. 7. Dispersion of the (left) MPP ensemble and (right) PIC ensemble for the (top) mean, (middle) maximum, and (bottom) standard deviation of the daily precipitation model estimates in the domain.

than the K–F and the G–D schemes (except if the Lin microphysics is used). The G–D scheme is the one that generates the lowest daily mean.

The new Goddard microphysics, on the other hand, gives a consistently higher mean and standard deviation than the rest of the microphysics with the exception of Lin's. This microphysics modeling, however, is not always the one generating the maximum values, which is indicative of the model generating wider regions with light rainfall.

It must be noted that the impact of the microphysical parameterizations might be underestimated, as convection

is not explicitly resolved at 25-km resolution and cumulus parameterizations has to be used at such grid space. Therefore, the conclusions in that regard should be made with caution.

To explore the role of microphysics, runs at 3-km resolution were performed for several explicit cumulus schemes in the parent at the 9-km resolution and no convective parameterization in the inner domain (Fig. 9). The focus of this exercise was on the sensitivity of an advanced and complex microphysics parameterization to the cumulus parameterization in the parent domain. Therefore, a single microphysics scheme (Goddard's) was used.
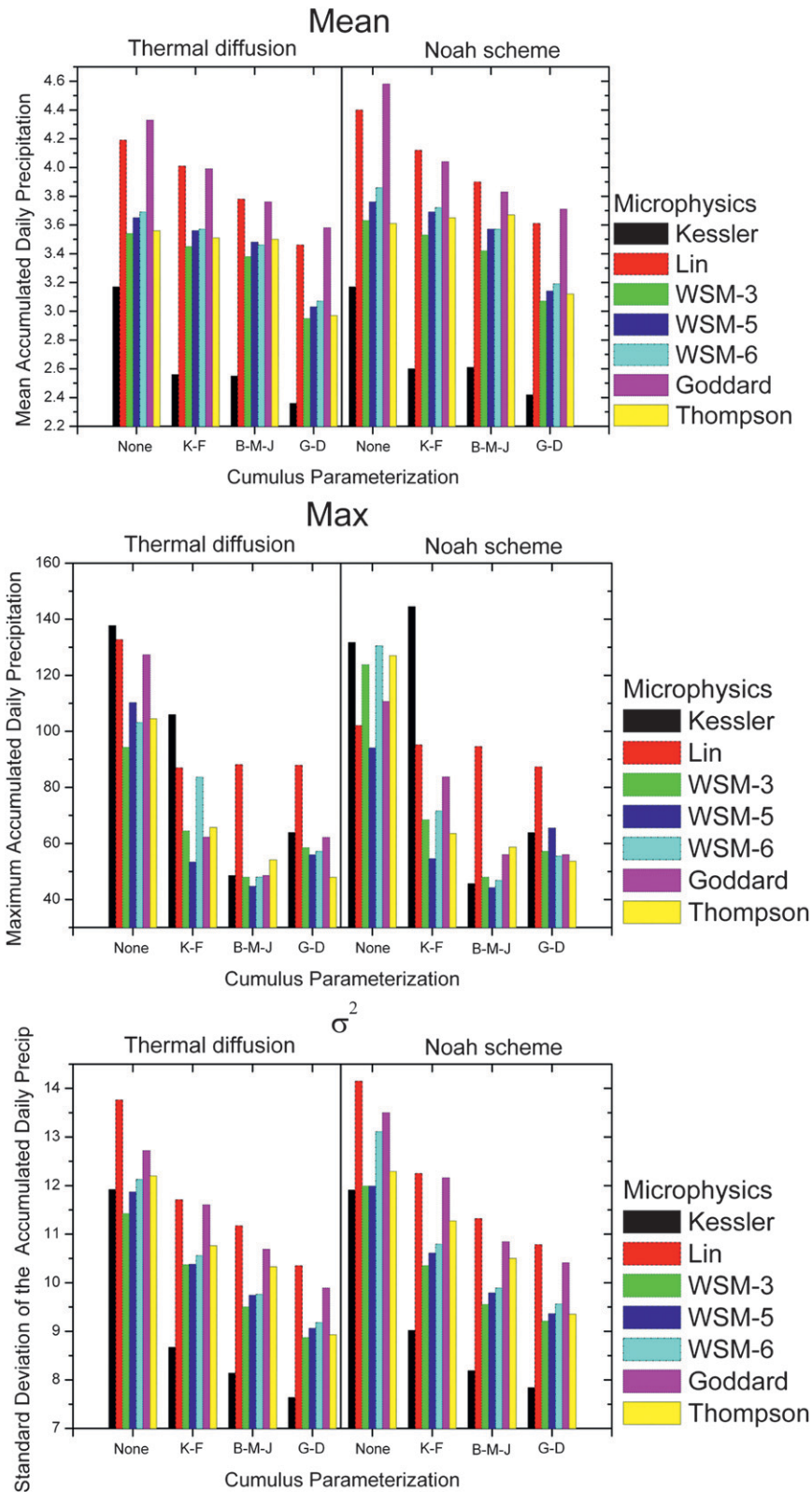
FIG. 8. Mean, maximum, and standard deviation of the daily precipitation estimates in the model domain, for land, cumulus, and microphysics parameterizations.
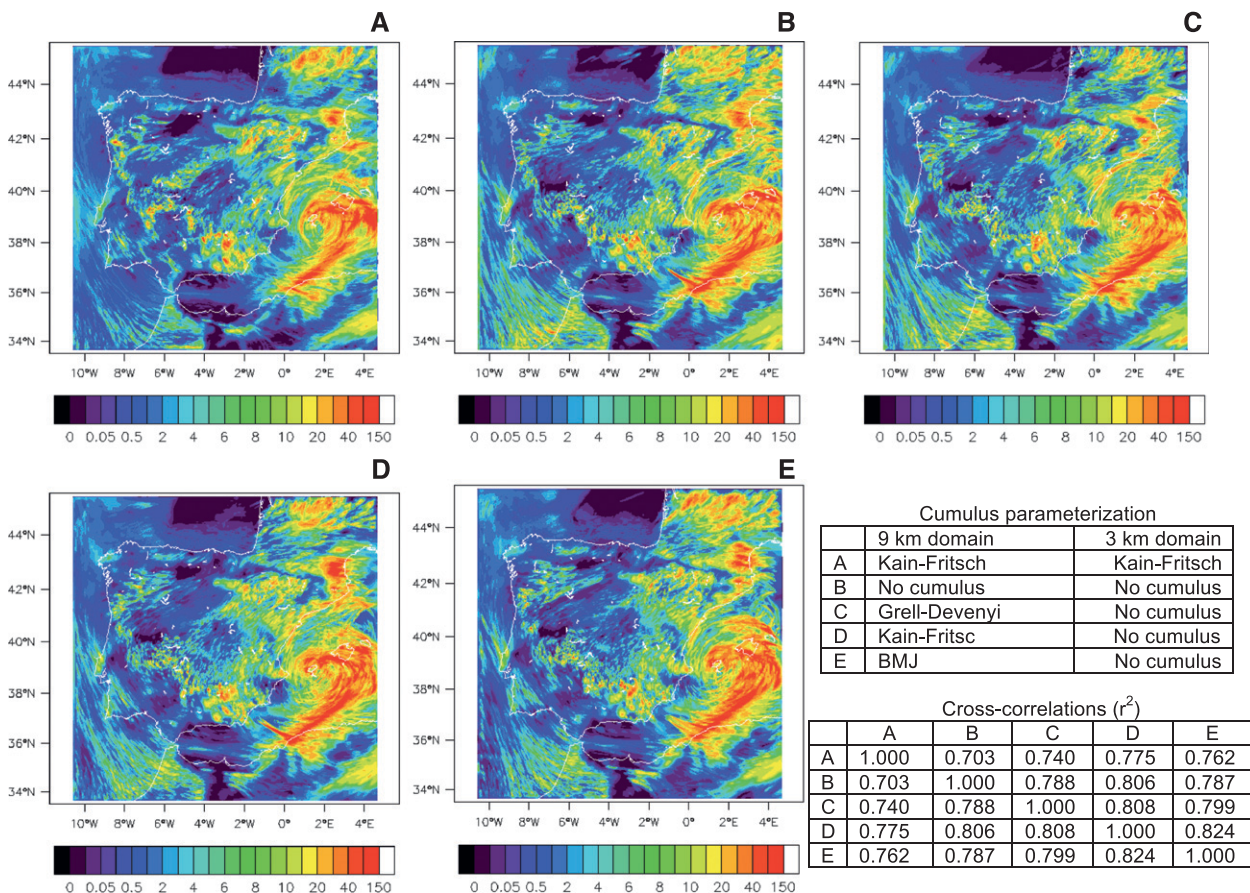
FIG. 9. WRF forecasts at 3-km grid space for the episode of interest (27-, 9-, and 3-km grid-spacing domains) for the five cumulus options in the 9- and 3-km domains listed in the lower right. The plots depict the daily accumulated precipitation for 11 May 2008 (cf. with equivalent 25-km-resolution runs in Figs. 3 and 4). The parameterizations used here are Goddard's (for microphysics), RRTM (LW radiation), Goddard's shortwave (SW radiation), Obukhov–Janjic (surface layer), thermal diffusion (land surface), and M–Y–J (PBL).

The results show that the spatial structure of the 3-km precipitation fields is consistent with that in the previous 25-km simulations (cf. Fig. 9 with Figs. 3 and 4). Cross correlations (Fig. 9, inset) and visual inspection show that no major changes appear in terms of the location of the main precipitating systems when the cumulus parameterization in the parent domain is varied. While a full analysis of the results of 3-km-grid simulations is beyond the scope of this paper, varying the microphysics schemes with several cumulus options in the parent domain will generate a new dataset worth of further analyses. Comparison of MPP and PIC ensembles at such fine spatial resolution is also a venue for future work. Note that in Fig. 9 the K–F cumulus parameterization for the 3-km-grid size domain is presented just as a reference, since at 3-km resolution explicit microphysics should be used. At such finescale, the model should be capable of resolving convective eddies and therefore there is little need for cumulus parameterization.

d. Cross-correlations for physics parameterizations

The relationships between the parameterizations in terms of the mean forecasted precipitation can be alternatively depicted as a correlation plot (Fig. 10), which shows the 3136 cross correlations corresponding to all the combinations of the parameterizations for the 25-km runs. The key of the plot is described in the inset of the figure: for each microphysics, the scheme of four cumulus parameterization is repeated for each one of the two land physics options, resulting in 56×56 correlations. Obviously, the resulting matrix is symmetrical and the main diagonal has a 1.0 r^2 value.

Figure 10 exhibits several patterns of interest. The basic repeating unit is a 4×4 square that corresponds to the cumulus parameterization. The main diagonals within each microphysics square is highly correlated whatever the cumulus parameterization (except the noncumulus option). The land physics, on the other hand, little affects
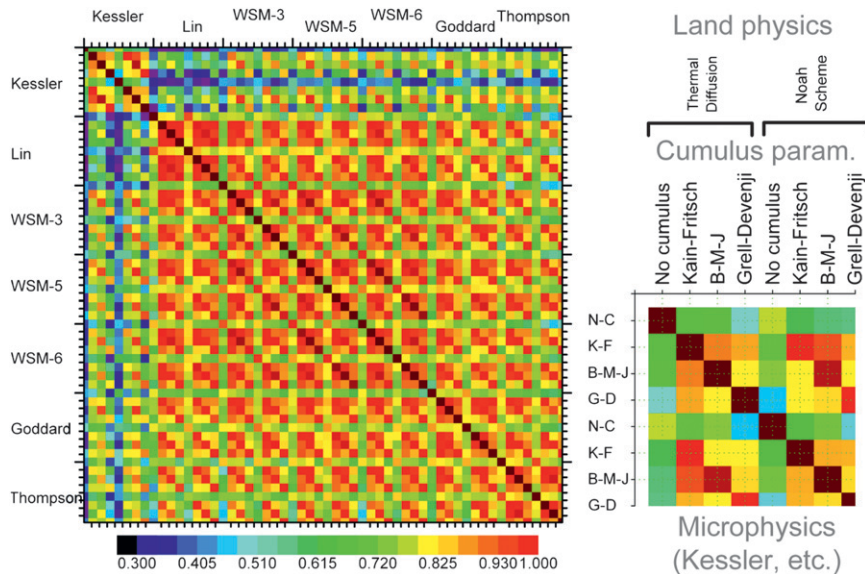
FIG. 10. Cross correlations for the 56 parameterizations of microphysics, land physics, and cumulus. The key of the main plot is given on the right. The values are the r^2 correlations of the mean accumulated precipitation over the simulation domain.

the correlations, as already expected from Fig. 8. There, it was also clear that the Kessler microphysics is loosely correlated with the rest of microphysics for this case, so one can expect large discrepancies in the correlation plot.

It should be emphasized at this point that exactly the same horizontal grid, vertical levels, and dynamics were employed for all the MPP and PIC simulations, with the only difference in the MPP being in some physical parameterizations. The spread of the MPP ensemble would be even more marked if several models mixing hydrostatic and nonhydrostatic models, different numerical schemes, semiempirical adjustments, and software precision were to be used. The same applies if the rest of the parameterizations were also modified. Thus for instance, the PBL scheme is known to affect precipitation as convection starts when heat and moisture in the cumulus scheme reach a given value. In the MPP simulations presented here, the M–Y–J PBL scheme (Mellor and Yamada 1982; Janjic 1994) was held constant. Varying this parameterization would certainly result in a further increase of the ensemble dispersion, making stronger the case for using MPP ensembles in addition to PIC ensembles.

4. Summary

Several simulations with varied microphysics, cumulus, and land parameterizations (MPP runs) have been run for a severe weather case study over Spain and compared with those from a physically perturbed ensemble (PIC runs). It has been shown that the dispersion of the MPP ensemble is larger than that of the PIC ensemble in

terms of storm center location; spatial structure of the precipitation; and maximum, mean, and standard deviation of the daily precipitation estimates. It should not be forgotten that the point of designing an ensemble is to provide a capture of the uncertainty by running *equally likely* realistic setups, and that the improvement in the forecast is contingent upon the spread of the ensemble being comprehensive enough. This is of the interest to operational early warning forecasters because, apart from obtaining better forecasts through the ensemble mean, the dispersion of the ensemble provides an indication of the uncertainty of the alerts.

These results suggest that in addition to optimizing initial conditions, as done in most operational forecasts, one must address physics deficiencies by using the model as a weak constraint (Hou and Zhang 2007) in order to provide better forecasts of extreme precipitation events. The simulations also show that there are large differences in storm location and rain intensity depending on the parameterizations. In the case of the modeling of precipitation extremes, the differences may be as large as 100% depending on the modeling done, so attention should be paid to how multimodel ensembles are built if changes in extremes or return periods are the variable of interest.

The conclusions described here are based on only one case. However, the focus of the research was to analyze the dispersion of the estimates and not the particular performances of the parameterizations; so it is likely that the differences observed here are not a peculiarity of the case chosen for study. The province of future work will be

to investigate the effects of the parameterizations in long-term simulations at higher spatial resolution.

REFERENCES

Anagnostou, E. N., and C. Kummerow, 1997: Stratiform and convective classification of rainfall using SSM/I 85-GHz brightness temperature observations. *J. Atmos. Oceanic Technol.,* **14,** 570–575.

Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.,* **139,** 1972–1995.

Betts, A., and M. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.,* **112,** 2949–2972.

Chang, H.-I., D. Niyogi, A. Kumar, C. Kishtawal, J. Dudhia, F. Chen, U. Mohanty, and M. Shepherd, 2009: Possible relation between land surface feedback and the post-landfall structure of monsoon depressions. *Geophys. Res. Lett.,* **36,** L15826, doi:10.1029/2009GL037781.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585.

Chou, M., and M. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, 85 pp.

Choudhury, B. J., 2000: Seasonal and interannual variations of total evaporation and their relations with precipitation, net radiation, and net carbon accumulation for the Gediz basin area. *J. Hydrol.,* **229** (1–2), 77–86.

Clark, A. J., W. Gallus, and T.-S. Chen, 2008: Contributions of mixed physics versus perturbed initial/lateral boundary conditions to ensemble-based precipitation forecast skill. *Mon. Wea. Rev.,* **136,** 2140–2156.

Cressie, N., 1993: *Statistics for Spatial Data.* Wiley-Interscience, 928 pp.

Gallus, W., 1999: Eta simulations of three extreme precipitation events: Sensitivity to resolution and convective parameterization. *Wea. Forecasting,* **14,** 405–426.

——, and J. Bresch, 2006: Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts. *Mon. Wea. Rev.,* **134,** 2632–2641.

Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.,* **29,** 1693, doi:10.1029/2002GL015311.

Hong, S. Y., H. M. H. Juang, and Q. Y. Zhao, 1998: Implementation of prognostic cloud scheme for a regional spectral model. *Mon. Wea. Rev.,* **126,** 2621–2639.

——, J. Dudhia, and S. H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.,* **132,** 103–120.

Hong, Y., C. D. Kummerow, and W. S. Olson, 1999: Separation of convective and stratiform precipitation using microwave brightness temperature. *J. Appl. Meteor.,* **38,** 1195–1213.

Hou, A. Y., and S. Q. Zhang, 2007: Assimilation of precipitation information using column model physics as a weak constraint. *J. Atmos. Sci.,* **64,** 3865–3879.

Hudson, J. G., and H. G. Li, 1995: Microphysical contrasts in Atlantic stratus. *J. Atmos. Sci.,* **52,** 3031–3040.

Janjic, Z., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.,* **122,** 927–945.

Jankov, I., W. Gallus, M. Segal, B. Shaw, and S. Koch, 2005: The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Wea. Forecasting,* **20,** 1048–1060.

Joyce, R., J. E. Janowiak, P. A. Arkin, and P. P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.,* **5,** 487–503.

Kain, J., and J. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 165–170.

Kessler, M. A., R. S. Anderson, and G. M. Stock, 2006: Modeling topographic and climatic control of east-west asymmetry in Sierra Nevada glacier length during the Last Glacial Maximum. *J. Geophys. Res.,* **111,** F02002, doi:10.1029/2005JF000365.

Khain, A., D. Rosenfeld, and A. Pokrovsky, 2005: Aerosol impact on the dynamics and microphysics of deep convective clouds. *Quart. J. Roy. Meteor. Soc.,* **131,** 2639–2663.

Lenderink, G., A. van Ulden, B. van den Hurk, and E. van Meijgaard, 2007: Summertime inter-annual temperature variability in an ensemble of regional model simulations: Analysis of the surface energy budget. *Climatic Change,* **81,** 233–247.

Lin, Y.-L., R. Farley, and H. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092.

Lorenz, E., 1978: Available energy and the maintenance of a moist circulation. *Tellus,* **30,** 15–31.

Mahrt, L., 2000: Surface heterogeneity and vertical structure of the boundary layer. *Bound.-Layer Meteor.,* **96** (1–2), 33–62.

Marzban, C., S. Sandgathe, H. Lyons, and N. Lederer, 2009: Three spatial verification techniques: Cluster analysis, variogram, and optical flow. *Wea. Forecasting,* **24,** 1457–1471.

Mellor, G., and T. Yamada, 1982: Development of a turbulent closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.,* **20,** 851–875.

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

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.,* **115,** D00J14, doi:10.1029/2009JD013302.

Prabhakara, C., R. Iacovazzi, J. A. Weinman, and G. Dalu, 2000: A TRMM microwave radiometer rain rate estimation method with convective and stratiform discrimination. *J. Meteor. Soc. Japan,* **78,** 241–258.

Roe, G. H., 2005: Orographic precipitation. *Annu. Rev. Earth Planet. Sci.,* **33,** 645–671.

Rutledge, S., and P. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation

development in narrow cold-frontal rainbands. *J. Atmos. Sci.,* **41,** 2949–2972.

Schwartz, C., J. Kain, S. Weiss, and M. Xue, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting,* **25,** 263–280.

Shi, J., and Coauthors, 2010: WRF simulations of the 20–22 January 2007 snow events over eastern Canada: Comparison with in situ and satellite observations. *J. Appl. Meteor. Climatol.,* **49,** 2246–2266.

Shin, H. H., and S. Hong, 2011: Intercomparison of planetary boundary-layer parameterizations in the WRF model for a single day from CASES-99. *Bound.-Layer Meteor.,* **139,** 261–281.

Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, 125 pp.

Steiner, M., 2004: Comments on "Modeling of convective–stratiform precipitation processes: Sensitivity to partitioning methods." *J. Appl. Meteor.,* **43,** 958–961.

Tapiador, F. J., 2010: A joint estimate of the precipitation climate signal in Europe using eight regional models and five observational datasets. *J. Climate,* **23,** 1719–1738.

Thompson, G., R. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132,** 519–542.

Tokay, A., and D. A. Short, 1996: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor.,* **35,** 355–371.

——, ——, C. R. Williams, W. L. Ecklund, and K. S. Gage, 1999: Tropical rainfall associated with convective and stratiform clouds: Intercomparison of disdrometer and profiler measurements. *J. Appl. Meteor.,* **38,** 302–320.

Vich, M., and R. Romero, 2010: Multiphysics superensemble forecast applied to Mediterranean heavy precipitation situations. *Nat. Hazards Earth Syst. Sci.,* **10,** 2371–2377.

Zupanski, D., S. Q. Zhang, M. Zupanski, A. Y. Hou, and S. H. Cheung, 2011: A prototype WRF-based ensemble data assimilation system for dynamically downscaling satellite precipitation observations. *J. Hydrometeor.,* **12,** 118–134.