

# *Downstream influence of mesoscale convective systems: part 2, influence on ensemble forecast skill and spread*

Article

Accepted Version

Clarke, S. J., Gray, S. L. ORCID: <https://orcid.org/0000-0001-8658-362X> and Roberts, N. M. (2019) Downstream influence of mesoscale convective systems: part 2, influence on ensemble forecast skill and spread. Quarterly Journal of the Royal Meteorological Society, 145 (724). pp. 2953-2972. ISSN 0035-9009 doi: <https://doi.org/10.1002/qj.3613> Available at <https://centaur.reading.ac.uk/84958/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/qj.3613>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Downstream influence of Mesoscale Convective Systems: Part 2, influence on ensemble forecast skill and spread

S. J. Clarke<sup>1</sup> | S. L. Gray<sup>1</sup> | N. M. Roberts<sup>2</sup>

<sup>1</sup>Department of Meteorology, University of Reading, UK

<sup>2</sup>MetOffice@Reading, Reading, UK

## Correspondence

S. J. Clarke, School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK

Email: s.j.clarke@leeds.ac.uk

## Funding information

Natural Environmental Research Council PhD studentship, Award reference: 101110, with Met Office CASE award

Ensemble forecasts are run operationally to determine forecast uncertainty arising from initial condition, model physics and boundary condition uncertainty. However, global configuration ensembles, which use a convection parametrization scheme, may miss uncertainty because of the misrepresentation of intense convection by such schemes. Here the impacts of the misrepresentation of Mesoscale Convective Systems (MCSs) on downstream ensemble forecast skill and evolution are determined for a case study. MCS perturbations (calculated from the difference between output from convection-parametrizing and convection-permitting Met Office model configurations) are added to six members of a global configuration ensemble created by downscaling forecasts from the global version of the Met Office Global and Regional Ensemble Prediction System.

For the first 36 h differences grow on the convective scale related to the MCSs leading to systematic deepening of a developing UK cyclone, although there is damping of the perturbations found in root mean square difference calculations between the forecasts with and without the perturbations (particularly in mean sea level pressure). Subsequently, differences grow rapidly onto the synoptic scale and by five days impact the entire northern hemisphere. The MCS perturbations can have systematic effects on the en-

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.3613

semble forecasts (e.g. a systematic displacement of a downstream cyclone is found), but for this case, there is no discernible change in forecast skill as measured by root mean square error of the ensemble means and the effects of the MCS perturbations are smaller than those generated by the initial condition perturbations. The spread of the combined ensemble (the two ensembles with and without the MCS perturbations) is larger than that of the individual ensembles. Thus, perturbing convection-parametrizing models to include PV anomalies associated with MCSs represented in convection-permitting forecasts, or idealised representations of them, produces alternative realisations to those generated by initial condition perturbations and has the potential to be useful operationally.

#### KEYWORDS

potential vorticity, convection-permitting, forecast error, convection, MOGREPS-G, Met Office Unified Model, MCS

## 1 | INTRODUCTION

Mesoscale convective systems (MCSs) are an amalgamation or organisation of individual thunderstorms and were defined broadly by Houze (2004) as cumulonimbus cloud systems that produce a contiguous precipitation area  $\sim 100$  km or more in at least one direction. They can be associated with severe local weather such as heavy rain, hail and strong winds. However, they can also impact the predictability of weather downstream. For example, Rodwell et al. (2013) linked six-day forecast busts for Europe to MCSs over North America. Previous studies have demonstrated that MCS events can be poorly represented by numerical weather prediction (NWP) models, particularly those with a grid coarse enough to require a convective parametrization scheme. Although current global configuration NWP models, with grid-spacings of about 10 km may be able to represent the structure of a mature MCS system, such models require a parametrization scheme for deep convection to represent the initiation and growth of MCS systems and average (steady-state) effect of the associated individual convective up- and downdraughts.

Potential vorticity (PV) is a useful diagnostic for the characterization of MCS structure since diabatic processes such as the mid-tropospheric heating in MCSs are a source of PV. Done et al. (2006) found that lenses of negative PV, associated with anticyclonic circulations, developed near the tropopause in NWP simulations of MCS events that represented convection explicitly whereas these did not develop in otherwise equivalent simulations in which convection was parametrized. Similarly, Chagnon and Gray (2009) illustrated and explained the larger amplitude horizontal variations in PV found in higher resolution (1-km grid spacing) convection-permitting simulations compared to coarser resolution (12-km grid spacing) convection-parametrizing simulations. Gray (2001) perturbed convection-parametrizing simulations of four cases studies with idealised PV structures chosen to resemble those associated with MCSs and found some reduction in mean sea-level pressure forecast errors. One proposed reason for the typically insufficient ensemble spread in operational ensemble prediction systems is a failure to properly represent the characteristics of MCSs, which

also have inherently low predictability due to their evolution from sub-grid scales in convection-parametrizing models: Shutts (2017) proposed that vorticity and divergence perturbations associated with MCSs are one of the dominant sources of the random model error that is alleviated by stochastic physics schemes in ensemble prediction systems. An insufficiently simulated upscale error growth has also been noted by Sun and Zhang (2016) as being one of the reasons for the underdispersion issue in ensemble predictions. The aim of this study is to assess the impact of the misrepresentation of MCSs by global (convection-parametrizing) models on downstream ensemble forecast skill and spread.

Latent heat release associated with deep convection is known to be an important mechanism for upscale error growth from initially-localised convective scales that can influence the large-scale flow (Zhang et al., 2002, 2003; Zhang, 2005; Zhang et al., 2007; Tan et al., 2004; Hohenegger and Schar, 2007a; Selz and Craig, 2015b; Bei and Zhang, 2007). A consequence of this mechanism is that mesoscale and synoptic-scale predictability can be limited by small-scale, small-amplitude initial errors where moist convection occurs. Errors occurring on the convective scale associated with moist processes grow much faster than errors on the synoptic scale, with saturation occurring after approximately 24 hours (Hohenegger and Schar, 2007a). A three-stage upscale error growth conceptual model has been proposed: see Zhang et al. (2007) for this conceptual model developed from results from idealised moist baroclinic wave experiments and Selz and Craig (2015b) for case studies that illustrate the conceptual model. These three stages of error growth are (i) rapid growth over the convective scale of small-scale errors associated with moist processes, (ii) a change in the character of the errors from that of convective-scale unbalanced motions to one more closely related to large-scale balanced (synoptic-scale) motions, and (iii) growth of the large-scale (balanced) components of the errors with the background baroclinic instability. The last of these stages can impact on downstream forecast predictability. Similarly, upscale error growth has been shown to occur in convection-permitting global medium-range models for initially small differences in moist convection, resulting in a loss of predictability at global scales (Mapes et al., 2008). Rodwell et al. (2013) found that six-day European busts were associated with a common initial condition (IC) that was evident in a composite of 584 events of high convective available potential energy (CAPE) to the east of a trough over the Rockies in North America, a situation frequently leading to MCSs; thus poor forecasts of MCSs can cause upscale error growth that impacts the synoptic-scale forecasts days after the MCS itself, leading to forecast busts downstream. The Rockies trough was part of a Rossby wave train that stretched across the North Atlantic, providing a pathway for forecast differences over North America to influence Europe. Two possible causes of the reduced forecast predictability were explored by Rodwell et al. (2013): initial condition (IC) uncertainty and model error. They found a strong sensitivity to ICs in a case study of a specific bust event and a strong relationship between the IC perturbation over North America and six-day forecast error over Europe for a given member of their operational ensemble. They also inferred a possible cause for model physics error in this case study since the diabatic and frictional processes acting in the MCSs were diagnosed to modify the Rossby wave structure, slowing the eastward propagation of the wave structure and opposing the adiabatic advection term. Negative PV tendencies on the 330-K isentropic surface above the region of strong positive CAPE anomalies were speculated to have arisen because of the PV "destruction" above the convective heating maximum. Tropopause-level PV anomalies associated with tropical cyclones undergoing extratropical transition have similarly been shown to slow down the propagation of Rossby waves through amplification of the tropopause (Grams and Archambault, 2016; Quinting and Jones, 2016; Riemer et al., 2008). In our study we similarly contrast these causes of reduced forecast predictability by comparing two ensembles of simulations both generated using operational IC perturbations: an ensemble with additional perturbations compensating for the poor representation of the MCSs in the forecasts (so-called MCS perturbations) and an ensemble without these perturbations.

As stated above, the aim of this study is to investigate the possible impacts of poor representation of MCSs on ensemble global NWP model forecasts with parametrized convection. In particular, the impact of MCS perturbations,

inserted to represent the poorly-represented MCSs, are contrasted with just the impact of IC perturbations used operationally to generate ensemble forecasts. The MCS perturbations are derived from the difference between simulations of the MCS event from a limited area convection-permitting configuration and the global configuration of the same NWP model (the Met Office Unified model: MetUM). A case study from July 2012, chosen as a typical example of European MCSs and which included a MCS that tracked over the UK, is used. In the companion study by (Clarke et al., 2019, hereafter 'Part 1'), it has been shown that the PV structure associated with the UK-tracking MCS in the convection-permitting simulation output is more intense, even after coarse-graining to the resolution of a convection-parametrizing simulation, than the equivalent PV structure in the output of that convection-parametrizing simulation. Here the derived MCS perturbations are added to an ensemble of convection-parametrizing model simulations to determine the impact on ensemble forecast skill and spread. We use an ensemble to examine the impact of the MCSs on the spread of forecasts because the inherent low predictability of MCSs implies that improved representation of MCS characteristics would not necessarily lead to an improvement in a single (i.e. deterministic) forecast.

A brief summary of the case study is given in Sec. 2. The MetUM and configurations used, the generation of the MCS perturbations and the diagnostics used to evaluate the ensemble are described in Sec. 3. An overview of the systematic downstream impact of the MCS perturbations in the ensembles is given in Sec. 4, Sec. 5 focusses on the impact on forecast skill and Sec. 6 describes the comparison of the impacts of the IC and MCS perturbations. Finally, Sec. 7 contains the conclusions.

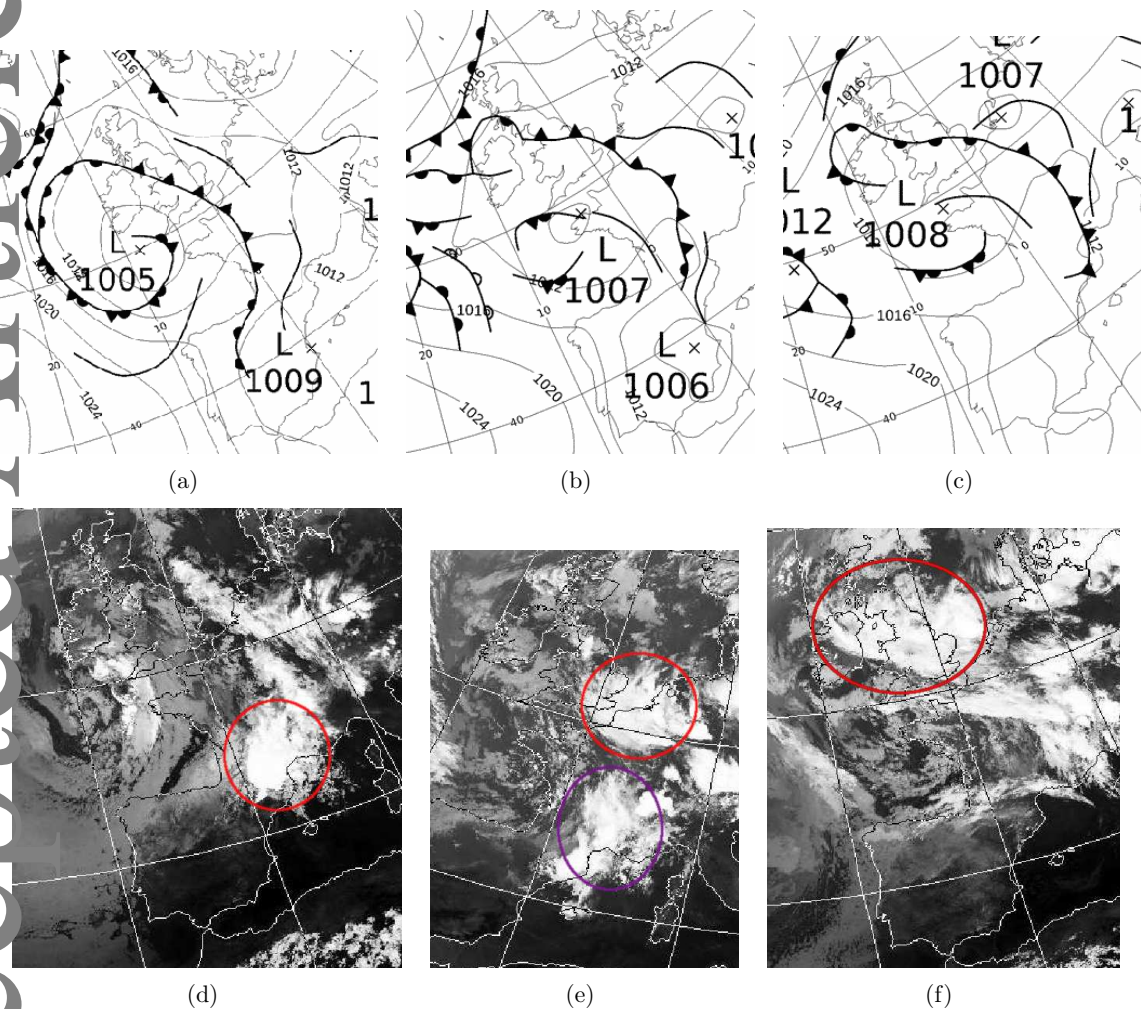
## 2 | OVERVIEW OF THE MCS CASE STUDY: 5–6 JULY 2012

An overview of the case study is presented in Part 1 and hence only a brief summary is included here. The MCSs developed in a synoptic environment termed the modified Spanish Plume by Lewis and Gray (2010). Figure 1(a) shows the mature low to the west of the UK at about the time of convective initiation of a MCS (hereafter MCS-A) over southern France early on 5 July (Fig. 1(d)). This MCS tracked towards the UK and by early evening on the same day lay along a cold front in a slack low pressure region (Fig. 1(b)) with the cloud shield spanning from southeast England to the near continent (Fig. 1(e)). Further convection (which developed into a second MCS, hereafter MCS-B) formed over southern France at this time. By early on 6 July (Figs. 1(c, f)) the cloud deriving from MCS-A had moved to cover most of the UK with the associated front now marked as a warm front.

## 3 | MET OFFICE UNIFIED MODEL

### 3.1 | Model and model configurations used

This study uses the same version (version 8.2) of the MetUM and identical configurations to those used in Part 1: the convection-permitting Euro4 and convection-parametrizing Global model configurations. The Euro4 is a 4.4-km grid spacing configuration with a domain that covers much of Europe and extends to the North Atlantic on a rotated latitude-longitude grid with 70 vertical levels and a lid at 40 km. The Global configuration has 25-km grid spacing. Additionally here we use the global model version of the Met Office global and regional ensemble prediction system (MOGREPS-G); this has 60-km grid spacing. Both the Global configuration and MOGREPS-G have the same 70 vertical levels with a lid at 80 km. All three of these model configurations were operational, with these grid spacings, at the Met Office at the time of the case study (2012). Convection is parametrized in the Global and MOGREPS-G configurations using the Gregory and Rowntree (1990) scheme. The resolution of the Euro4 configuration lies in the convective "grey zone" for



**FIGURE 1** Met Office synoptic analysis charts for (a) 00 UTC 5 July, (b) 18 UTC 5 July 2012 and (c) 00 UTC 6 July 2012 (Archived by [www.wetter3.de](http://www.wetter3.de), Crown copyright) and infra-red satellite images for (d) 0230 UTC 5 July, (e) 0336 UTC 5 July and (f) 0220 UTC 6 July 2012 (Courtesy of Dundee satellite receiving station). The red and purple ellipses mark the locations of MCS-A (red) and MCS-B (purple) discussed in the text in the satellite imagery.

which smaller showers will need to be parametrized but larger and more intense showers can be represented explicitly. To achieve this the Euro4 configuration uses a version of the Gregory and Rowntree (1990) convection parametrization that has been adapted for convective grey-zone resolutions such that the activity of the scheme is restricted through a modified CAPE closure (Roberts, 2003). Despite this modification and consequent more realistic representation of convective showers, this version of the Euro4 configuration typically produces showers that are too few, too intense, too organised and which initiate too late (Lean et al., 2008; Clark et al., 2016). The Euro4 simulations used boundary conditions and ICs from the associated Global configuration simulation. Global configuration simulations either used ICs from operational analyses or from a MOGREPS-G simulation as appropriate.

## 3.2 | Perturbation insertion method

As described in Part 1, MCS perturbations are calculated from the difference between output variables from corresponding Global and Euro4 configuration simulations with the Euro4 output reconfigured to the Global configuration vertical levels and coarse-grained to 25-km grid spacing (the grid-spacing of the Global configuration). The coarse-graining is performed by averaging the Euro4 output to the resolution of the Global configuration although the procedure is complicated by the rotated latitude-longitude grid used by the Euro4 model (as explained in Section 3.2 of Part 1). The fields used are the zonal ( $u$ ) and meridional ( $v$ ) components of wind, and potential temperature ( $\theta$ ). Using these three variables a so-called 'MCS perturbation' file is created. The perturbations are calculated over a limited area, chosen to encompass the region of western Europe covered by the two MCSs that developed in this case (40–54°N, 8.0°W–7.8°E). This region will also include perturbations that are attributable to synoptic-scale divergence between the convection-parametrizing and the convection-permitting simulations; however, these perturbations are relatively small as expected after only 12 hours of model forecast (see Fig. 5(c,d) in Part 1 which shows the differences between the convection-parametrizing and the convection-permitting deterministic simulations after 15 hours of forecast). The dependence of the impact of MCS perturbations on the types of perturbations (perturbations inserted over all levels, near-tropopause levels only, mid-troposphere levels only, and near-tropopause levels with amplitude increased by a factor of three) was examined in Part 1. Although perturbations inserted at all levels had greater impact on the downstream Rossby wave structure than those inserted over a limited range of levels, the near-tropopause perturbations had a greater impact than the mid-troposphere perturbations (comparing just the un-amplified perturbations). Here only the perturbations inserted at all levels are examined.

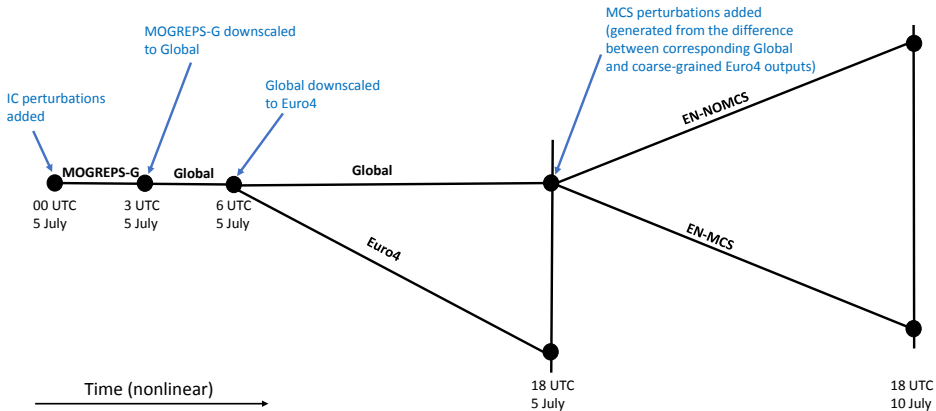
The MCS perturbation file is formatted as a unified model fields file (also known as an ancillary file) and inserted into the model simulation using the model's incremental analysis update (IAU) scheme (Clayton, 2012). The IAU scheme is designed to accept updates to the three fields used as well as to other basic variables, but cannot accept updates to PV directly; hence the associated circulation and temperature fields are updated as a surrogate to updates in PV. Pairs of Global configuration simulations with and without the inclusion of the MCS perturbation files (one pair for each Global configuration simulation downscaled from the MOGREPS-G ensemble members) were run for five days starting at 18 UTC 5 July 2012, when MCS-A was mature, and the differences between the two simulations analysed.

The IAU scheme allows increments to be added gradually over several timesteps with the fraction added at each timestep being determined by a (discrete) IAU weighting function (Bloom et al., 1996). The gradual insertion prevents some of the spurious noise that would otherwise be created. Here the increments are added uniformly over a 60-minute period at every 10-minute timestep. The IAU scheme is also used (here and operationally) to insert the IC perturbations into the MOGREPS-G simulations.



### 3.3 | Ensemble generation

To investigate the systematic impact of the MCS perturbations on flow evolution, five operational MOGREPS-G ensemble IC perturbation files (chosen randomly from the 24 members used operationally at 00 UTC 5 July 2012: members 1, 5, 12, 15 and 21) and the operational MOGREPS-G start dump for 00 UTC 5 July 2012 are used to generate six 60-km grid spacing MOGREPS-G simulations (a Control and the five members with IC perturbations). These simulations are run for 3 hours (until 03 UTC) and then reconfigured to the Global configuration resolution (25-km grid spacing). The experimental design can be seen in Figure 2. These six Global configuration simulations are run from 03 UTC 5 July for 15 hours (until 18 UTC) and start dumps are produced at 06 and 18 UTC 5 July. The six start dumps at 06 UTC are each used to initialise Euro4 configuration simulations which are then run from 06 UTC to 18 UTC 5 July 2012. An MCS perturbation file is calculated for all model levels for each of the six ensemble members using the difference between the 18 UTC output of each Global simulation and the corresponding coarse-grained output from the Euro4 simulation from the same MOGREPS-G simulation; hence, six different MCS perturbation files are created. The MCS perturbations have a similar magnitude to the IC perturbations added to the MOGREPS-G simulations (up to  $15 \text{ ms}^{-1}$  for the  $u$  wind component and 4 K for  $\theta$  (not shown)), but the IC perturbations occur across the whole globe. Finally, two corresponding six-member, five-day ensemble Global configuration forecasts are generated. These ensembles are initiated from the previously-generated start dumps valid at 18 UTC 5 July: one with the individual MCS perturbations added to each ensemble member (termed EN-MCS) and one without any MCS perturbations (termed EN-NOMCS).



**FIGURE 2** Schematic illustrating the generation of the two corresponding ensemble forecasts with and without MCS perturbations.

### 3.4 | Ensemble diagnostic methods

Two diagnostics used to evaluate the ensemble forecasts are described here: root mean squared difference (RMSD) and correspondence ratio (CR).

For this study the RMSD is a measure of either the differences between (i) a given member of an ensemble and the mean of that ensemble or (ii) a given member of one ensemble and the corresponding member of the other ensemble.

RMSD is calculated for a chosen field and for a given ensemble member as

$$\sqrt{\frac{\sum_1^N (x_i - y_i)^2}{N}},$$

where  $N$  is the number of points in the domain,  $x_i$  is the value of the field (at point  $i$ ) in the forecast from the given ensemble member and  $y_i$  is the value of the field from either the ensemble mean forecast or corresponding ensemble member from the other ensemble. Root mean square error (RMSE) is calculated in the same way as RMSD, but for the differences between a given member of an ensemble and an analysis.

CR is a measure of the spread (or equivalently dispersion) of an ensemble and takes a value from zero to unity with smaller values indicating larger spread. CR is used as in Gebhardt et al. (2011) to measure the fractional occurrence of an event in ensemble members where the event is defined by the exceedance of a given threshold for a particular variable (e.g. precipitation rates exceeding  $0.1 \text{ mm h}^{-1}$ ). CR is calculated as

$$CR = \frac{N(GP_{all})}{N(GP \geq 1)},$$

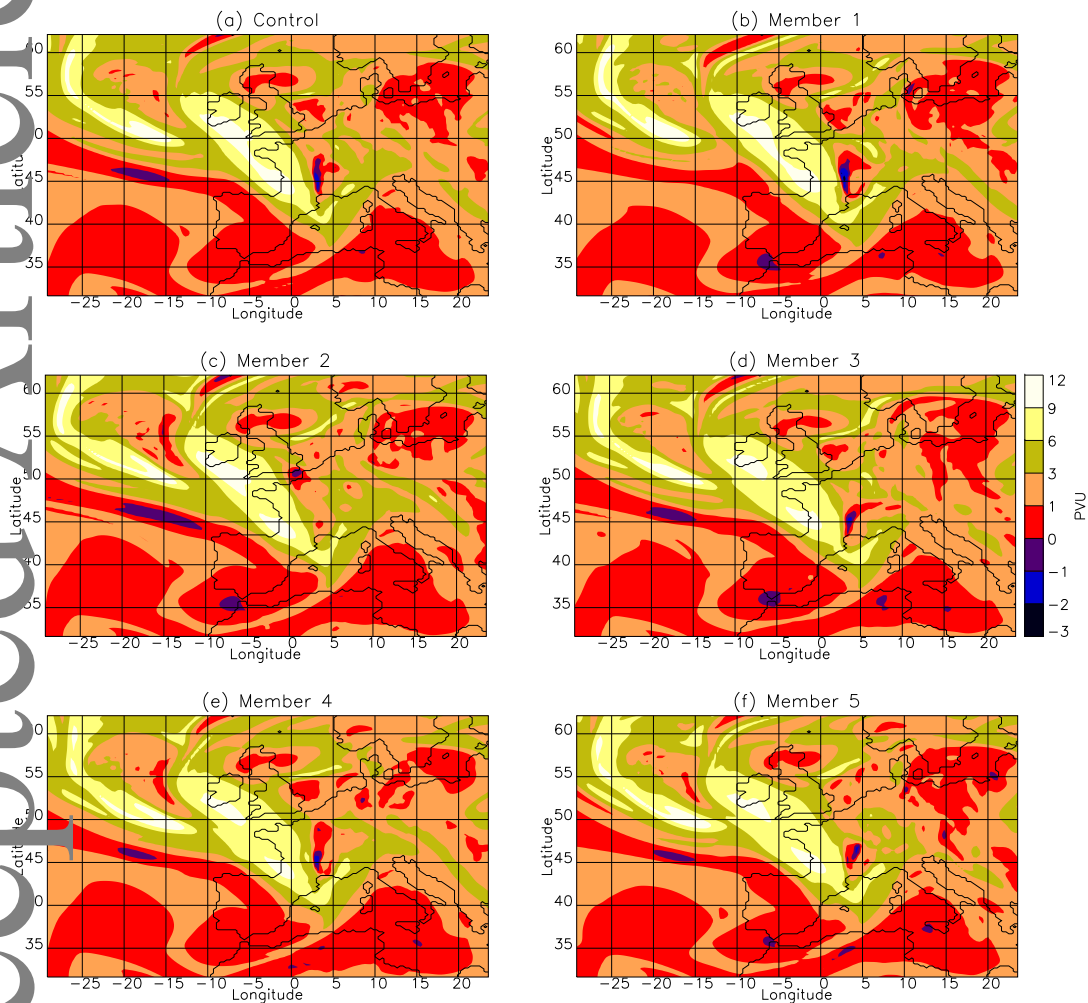
where  $N(GP_{all})$  is the number of grid points at which a given event is forecast by all members in the ensemble and  $N(GP \geq 1)$  is the number of grid points at which the event is forecast by at least one ensemble member. The CR was initially described by Stensrud and Wandishin (2000); however, they included an observation in their CR so their CR was a measure of forecast quality rather than forecast spread.

#### 4 | DOWNSTREAM INFLUENCE OF MCS PERTURBATIONS: OVERVIEW

The impact of the MCS perturbations on the different ensemble members are illustrated in this section at the short-range (over the first day) and medium-range (after five days). Part 1 also includes analysis of the impact of MCS perturbations at intermediate timescales and dynamical interpretation of the impact, though this is for global model forecasts initialised with the operational global model analysis rather than from MOGREPS-G forecasts as here. Fields of PV at 250 hPa (PV250 hereafter), precipitation rates and mean sea level pressure (PMSL) are shown to demonstrate the impact both at near-tropopause level and the surface.

The PV250 fields for the members of the EN-NOMCS ensemble are shown at 19 UTC 5 July (at the end of the hour over which the MCS perturbations were inserted in the EN-MCS ensemble) in Fig. 3. Here the differences arise primarily from the IC perturbations added to the MOGREPS-G simulations (at 00 UTC on the same day) from which the equivalent Global configuration simulations were downscaled. While obvious differences can be seen between the members in the details of the PV250 fields, the broad features are similar. A stratospheric PV filament (high PV values) associated with an upper-level trough extends from Ireland to the Mediterranean in all of the ensemble members. The observed MCSs (MCS-A and -B) are crossing into southeast England and in southeast France, respectively, at this time (see Fig. 1(e)). MCS-B is clearly associated with reduced values of PV250 in all members: erosion of the northern flank of the stratospheric filament is evident in all members and negative values of PV250 are present in all members except member 2. However, there is less evidence of MCS-A in this field. There are some indications of reduced PV250 over southeast England in all the ensemble members except for member 5, though they cannot be conclusively linked to MCS activity here. Negative PV250 values in the location of the observed convection over southeast England are only found for member 2.

At this time (19 UTC 5 July) clear signatures of the perturbations can be seen in the differences in the PV fields

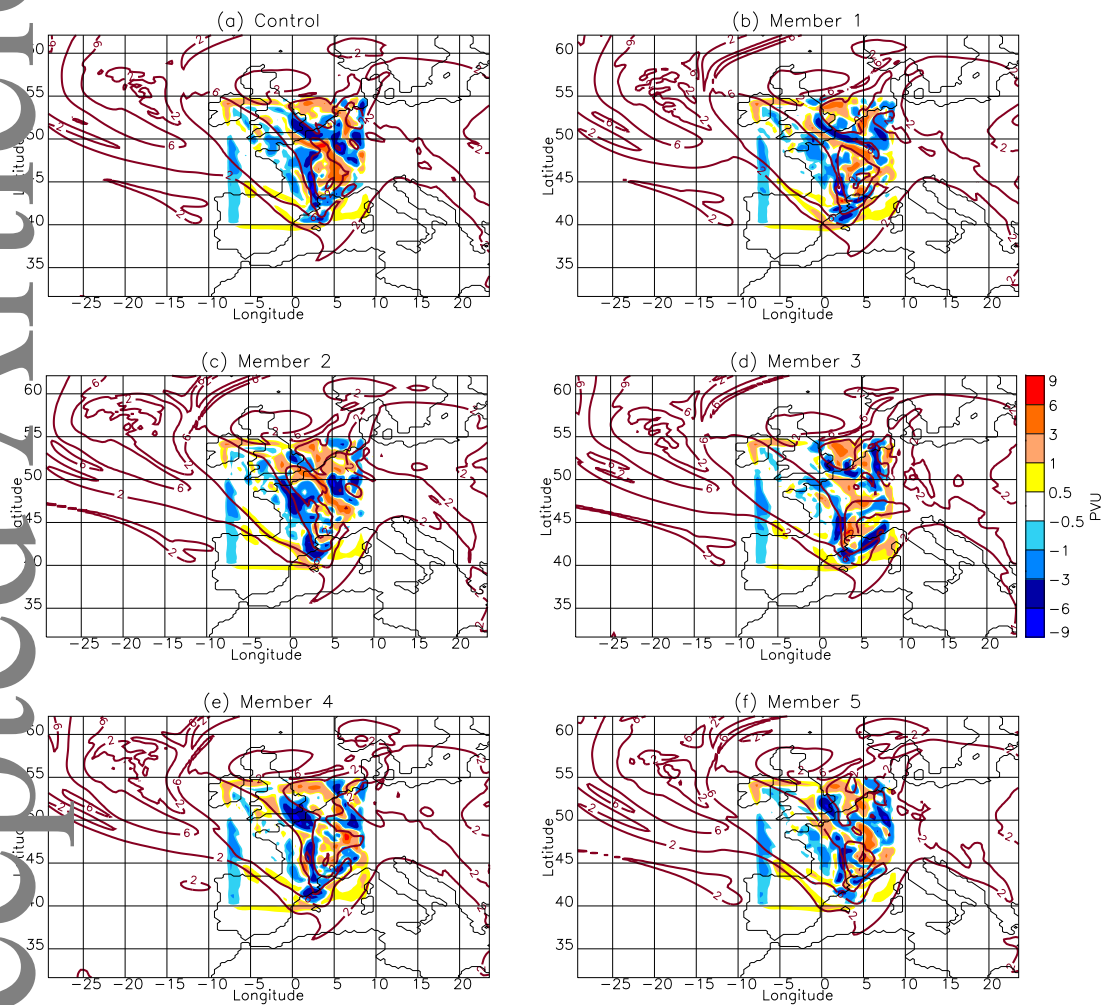


**FIGURE 3** PV250 at 19 UTC 5 July 2012 (1 h after the MCS perturbations were inserted in the EN-MCS members) for the following EN-NOMCS ensemble members: (a) Control, (b) member 1, (c) member 2, (d) member 3, (e) member 4 and (f) member 5.

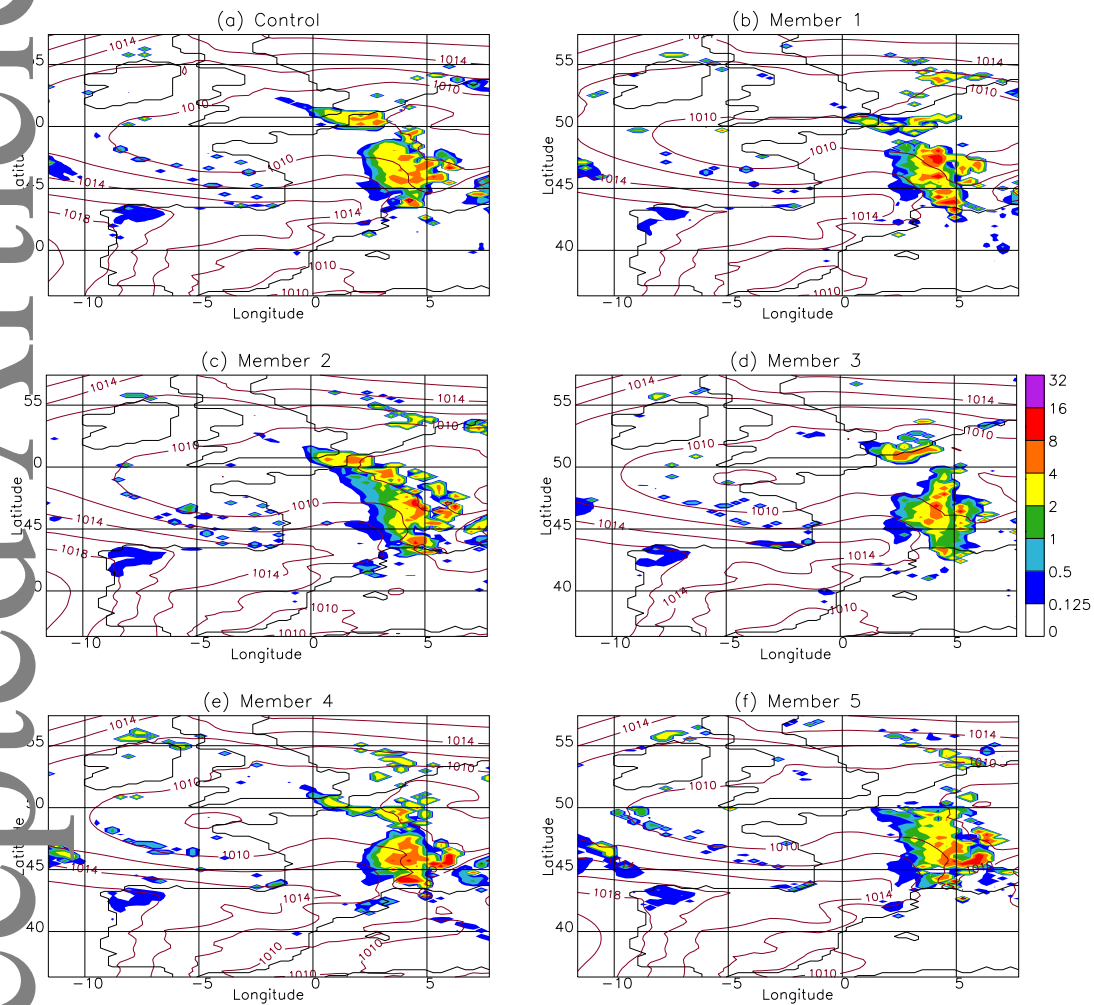
between the corresponding members of EN-NOMCS and EN-MCS (Fig. 4). These PV differences are mainly confined to the region where the perturbations are inserted and are different for each ensemble member because they are calculated from the difference between the outputs from the corresponding Euro4 and Global simulations downscaled from each MOGREPS-G ensemble member, where each ensemble member is initialised with different IC perturbations. The bands of differences along the western and southern edges of the perturbation region may be a consequence of small-amplitude, synoptic-scale differences between the Global and Euro4 simulations at this time (13 hours after the Euro4 simulations were initialised from the downscaled Global simulations). The magnitude of the PV anomalies added to each member of EN-MCS does not markedly differ (e.g.  $\theta$  perturbations at 250 hPa typically vary between  $-4$  and  $4^\circ\text{C}$  (not shown)), but the location of the anomalies varies depending on where MCSs were created in the Euro4 and Global simulations. Negative PV250 differences are found in the centre of the regions where the two MCSs are simulated, consistent with the larger amplitude negative PV anomalies in the Euro4 configuration output, compared to the Global configuration output, even after downscaling to the grid spacing of the Global configuration (see Part 1). The magnitude of these negative PV differences is such that absolute negative values of PV250 exist in the EN-MCS members at this time (compare Figs. 3 and 4).

In almost all of the ensemble members, in both the EN-MCS and EN-NOMCS ensembles, precipitation was found to be occurring in the regions of both observed MCSs at 00 UTC 6 July, six hours after the insertion of the MCS perturbations in the EN-MCS ensemble (compare Figs. 5 and 6 with Fig. 1(f); note that the observed precipitation within the domain was concentrated over the southeast England and the English channel and in two regions in eastern France at 00 UTC (not shown)). The exception is ensemble member 5 in which very little precipitation occurred in the region of MCS-A at this time in both ensembles (though precipitation did occur in the region of MCS-B). Differences exist though in the precise location and intensity of the precipitation between ensemble members in a given ensemble (due to the IC perturbations) and between corresponding ensemble members from the EN-NOMCS and EN-MCS ensembles (due to the MCS perturbations) with MCS-A being noticeably more intense in the Control member and members 1 and 4 of EN-MCS than in the corresponding EN-NOMCS members. The MCS perturbations also affect the PMSL field (also shown in Figs. 5 and 6). The most noticeable difference is the reduction in PMSL to the east of MCS-A that occurs in all ensemble members after insertion of the MCS perturbations. The analysis has a minimum of 1007 hPa marked within a small cyclone (single closed PMSL contour) over the Netherlands at this time (Fig. 1(c)) and the EN-MCS members more closely agree with this analysis than the EN-NOMCS members (mean PMSL for the grid point at  $51.56^\circ\text{N}$ ,  $4.57^\circ\text{E}$  is 1007.11 and 1008.11 hPa for the EN-MCS and EN-NOMCS ensembles, respectively with the PMSL of each EN-MCS member always less than that of the corresponding EN-NOMCS member). Hence, perturbations leading to a more intense MCS structure are associated with a deeper surface cyclone.

The PV250 fields for the EN-NOMCS members are shown at the end of the five-day forecasts in Fig. 7. The differences are typical of those found for medium-range ensemble forecasts: for example, while all six members have a ridge over central Europe (straddling  $40^\circ\text{E}$ ), the shape and northwards excursion of the ridge differs between members. The differences in PV at the same time arising in each ensemble member from the insertion of the MCS perturbations five days previously are shown in Fig. 8. As for the experiments with different types of MCS perturbations examined in Part 1, the largest amplitude differences are found to lie in bands, often as positive/negative dipoles, along the tropopause boundary (2-PVU surface) in the ridge over central Europe. However, the precise location of the differences and their magnitudes vary between ensemble members. Recent error growth studies using PV-based analysis have shown differences in the representation of convection (particularly for extratropical cyclones, tropical cyclones and Rossby waves and warm-conveyor belts) project onto larger scales predominantly by displacing the sharp PV gradient associated with the upper-level midlatitude jet, similarly to in this study (Schäfler and Harnisch, 2015; Joos and Forbes, 2016; Grams and Archambault, 2016; Quinting and Jones, 2016). Where PV difference dipoles are aligned with the



**FIGURE 4** As for Fig. 3 but differences in PV250 (EN-MCS minus EN-NOMCS) overlain with contours of PV250 (in PVU) from the corresponding EN-NOMCS members.



**FIGURE 5** Total precipitation rates and PMSL contours at 00 UTC 6 July 2012 for each member of EN-NOMCS for (a) Control member and ensemble members (b) 1, (c) 2, (d) 3, (e) 4 and (f) 5. Contours of PMSL (in hPa, contour interval 2 hPa) are overlain.

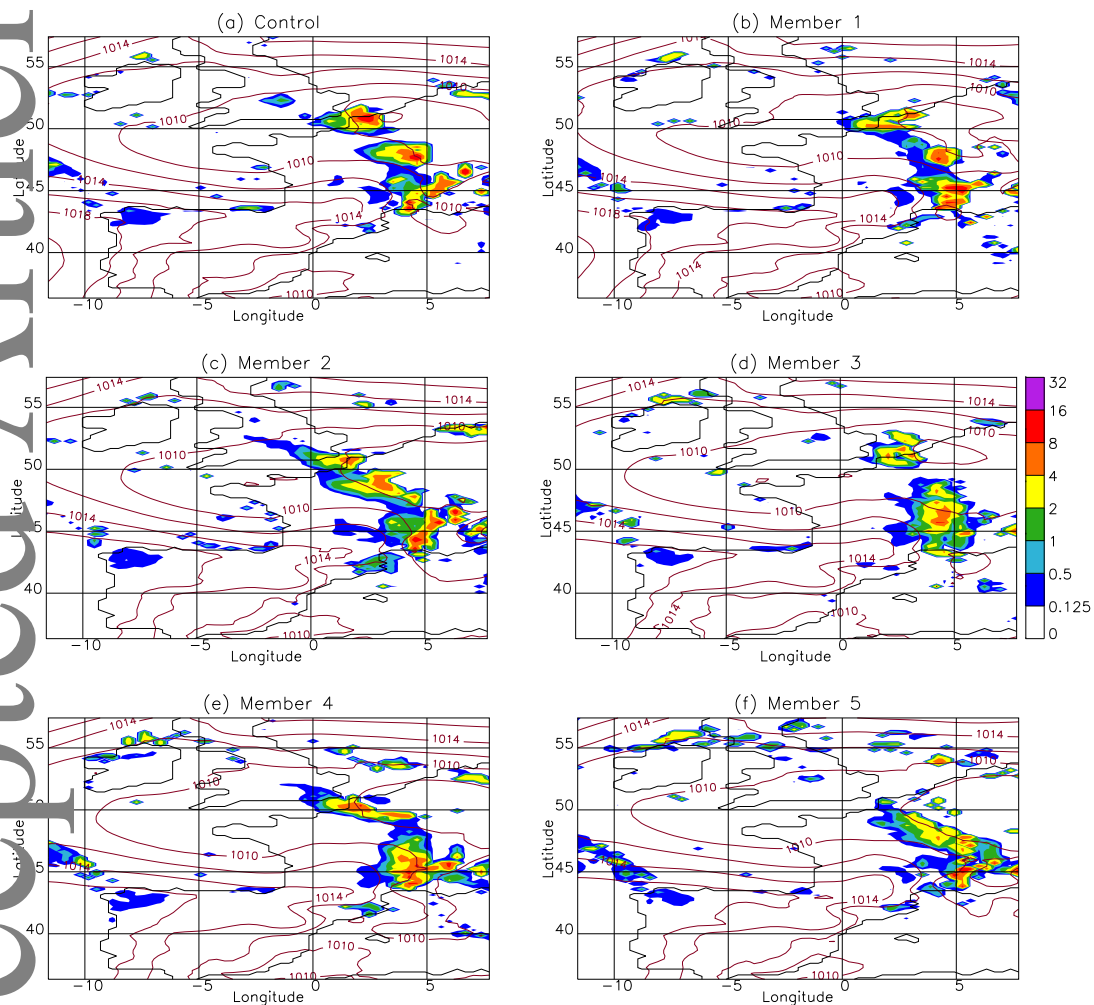


FIGURE 6 As Fig. 5, but for each member of EN-MCS.

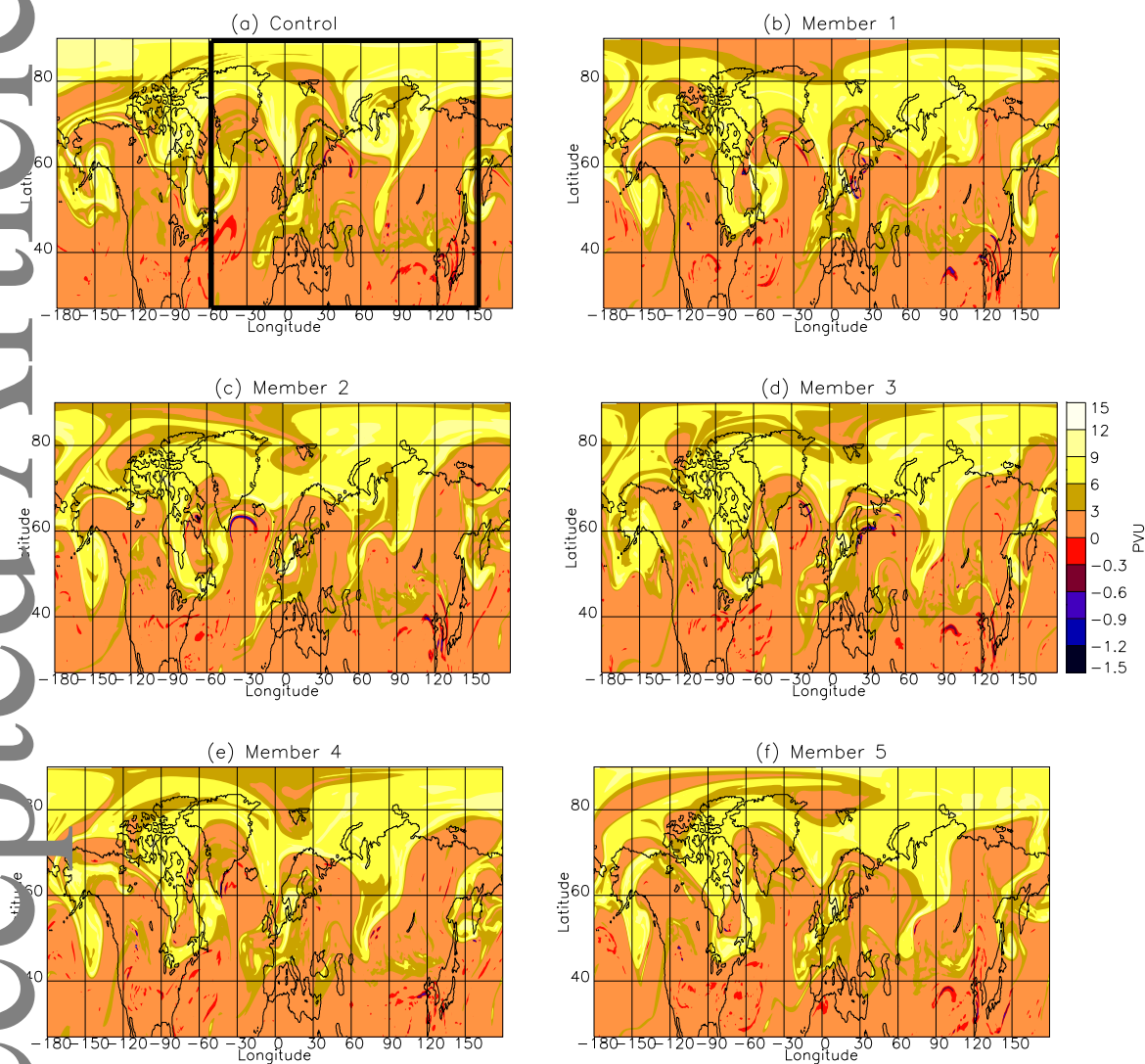
tropopause, and the tropopause location is not significantly changed by the MCS perturbations, these PV differences will be associated with modifications to the upper-tropospheric jet speed. Such changes are consistent with the previously documented ability of MCSs to act as a Rossby wave source region leading to the modification and generation of upper-tropospheric jet streaks that can then propagate downstream (e.g. Perkey and Maddox, 1985; Wolf and Johnson, 1995; Stensrud, 1996). The impact of differences in the representation of convection on upscale error growth has been investigated by Baumgart et al. (2019) by applying diagnostics based on PV and the envelope of Rossby waves to convection-parametrizing ensemble forecasts that use the Plant-Craig stochastic convection scheme (Plant and Craig, 2008). Similar to the findings in our study, they found the strong PV gradient associated with the tropopause is displaced by error growth associated with moist processes occurring below (with associated latent heat release) and upper-level divergence projects these errors effectively into the tropopause region. After two days, they found that error amplification at the large-scale was dominated by differences in the nonlinear near-tropopause dynamics.

## 5 | DOWNSTREAM INFLUENCE OF MCS PERTURBATIONS: FORECAST SKILL

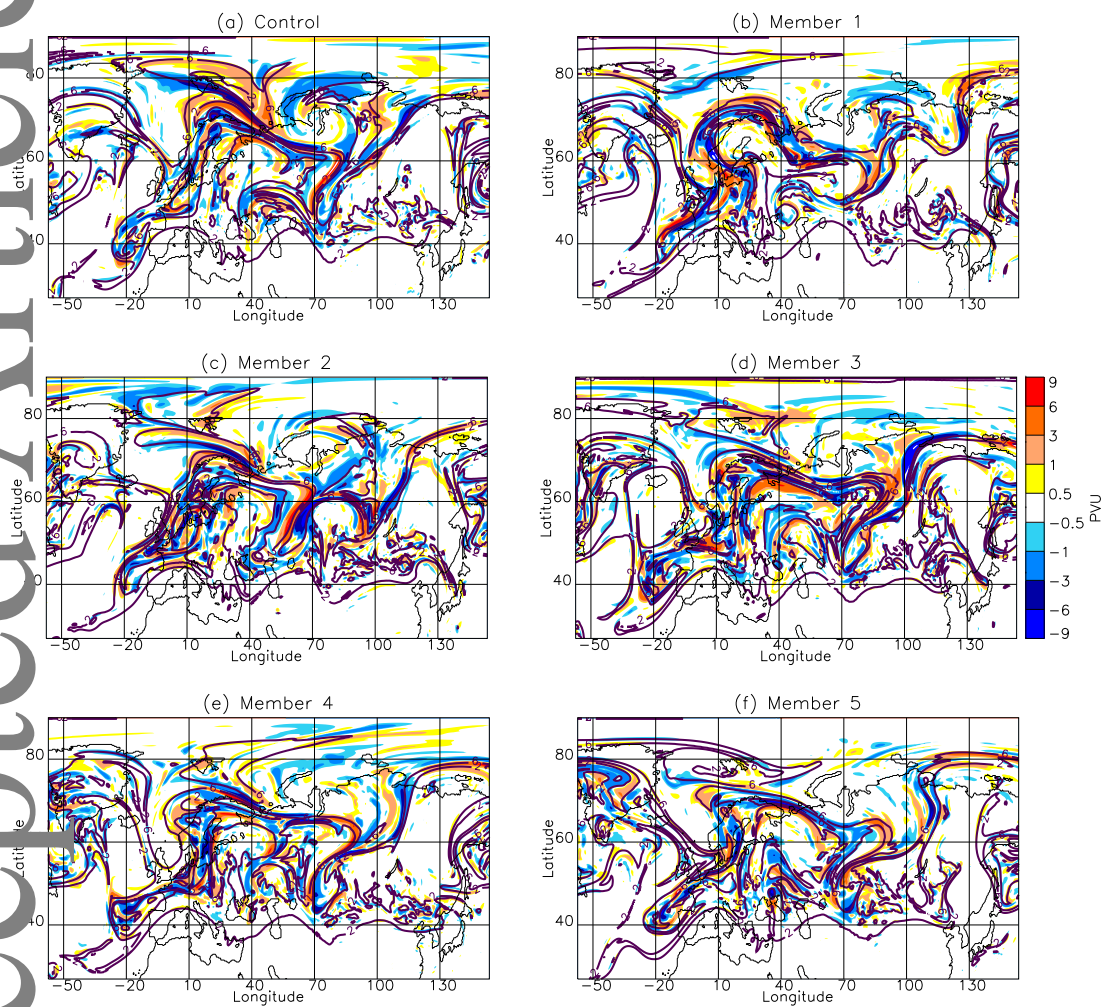
The MCS perturbations can have systematic impacts on medium-range downstream flow evolution despite the variability between ensemble members seen in Fig. 8. This is illustrated by examining a surface low pressure system that developed over northern Russia, downstream of the MCS perturbation location, five days after the MCS perturbations were inserted. This system was also analysed in Part 1 and was chosen because the MCS perturbations were found to lead to the largest differences in PMSL in this region (for the forecasts initialised with the global operational analysis). Figure 9 shows, for each ensemble member, a comparison between contours of PMSL from the given ensemble member in both EN-MCS and EN-NOMCS and from the Met Office global model analysis. For each ensemble member the amplitude and location of the low pressure system is more similar between the two forecasts (with and without the MCS perturbations) than between either forecast and the analysis: the analysed system is further to the east and more intense than the system in all ensemble members. This degree of disparity between the forecasts and the analysis is not unexpected at this five-day lead time. The forecast low pressure systems also vary more between ensemble members than between the corresponding ensemble members from EN-MCS and EN-NOMCS. Hence, the IC perturbations have more of an impact on the strength and location of the low pressure system than the MCS perturbations. Despite this, there is a systematic difference between the location of the low pressure system when comparing the corresponding ensemble members with and without the MCS perturbations: the low pressure system in the EN-MCS member with the perturbations is always further to the west or northwest. The MCS perturbations lead to a slight deepening of the low pressure system in half of the ensemble members (by up to 1 hPa) and a slight weakening in the other half. Finally, the poorest forecasts of the low pressure system occur in ensemble member 5 of both ensembles. This is the ensemble member in which there was no precipitation generated associated with MCS-A in both EN-MCS and EN-NOMCS. This suggests that the poor representation of the MCS in this ensemble member may have been a factor in the poorer forecast of the downstream low pressure system.

Despite this specific example, a consistent improvement or degradation in forecast skill through inclusion of the MCS perturbations is not seen when considered over a European subdomain (30–70°N and 50°W–100°E; note that this subdomain is much larger than the region over which the perturbations were added) and over the duration of the five-day forecast. Figure 10 shows the evolution of RMSE in PMSL for all the individual ensemble members and for the two ensemble means. The difference in RMSE between ensemble members from the same ensemble is far greater than that between corresponding ensemble members from the EN-MCS and EN-NOMCS simulations. At the end of the simulations, four of the simulations have been improved by the MCS perturbations (by between 1.1% (member





**FIGURE 7** PV250 at 18 UTC 10 July 2012 for each member of EN-NOMCS for (a) Control member, and ensemble members (b) 1, (c) 2, (d) 3, (e) 4 and (f) 5. The black box in panel (a) indicates the subdomain shown in the panels in Fig 8.



**FIGURE 8** Differences in PV250 18 UTC 10 July 2012 (EN-MCS minus EN-NOMCS) for (a) Control, (b) ensemble member 1, (c) ensemble member 2, (d) ensemble member 3, (e) ensemble member 4 and (f) ensemble member 5. PV250 from the corresponding EN-NOMCS member is overlain (2 and 6 PVU contours).

3) and 1.9% (Control member)), but the simulation for member 2 has been degraded by 4.6% and consequently the improvement in the ensemble mean is very modest: the RMSE lines for the two ensemble means (thin and thick dashed grey lines) are indistinguishable in Fig. 10 implying that there is no discernible ensemble mean change in forecast skill from the addition of the MCS perturbations. Members 2 of EN-MCS and EN-NOMCS also have the worst RMSE at five days. A low pressure system developed over Kazakhstan in member 2 at this time that was not present in any of the other ensemble members or in the analysis (not shown); the addition of the MCS perturbations deepened this low pressure system, enhancing the RMSE compared to the simulation without the perturbations. The lack of a consistent improvement in forecast skill over a broad region and five-day forecast is not unexpected when considering a single case study and bearing in mind the limitations in the predictability of MCS events. This result supports the ensemble approach taken in this paper which enables the impact of the representation of MCSs on forecast spread, as well as on deterministic forecast skill, to be assessed.

## 6 | DOWNSTREAM INFLUENCE OF MCS PERTURBATIONS: COMPARISON OF INITIAL CONDITION AND MCS PERTURBATIONS

Figure 11 shows the evolution of the RMSD over the European subdomain (30–70°N and 50°W–100°E) between corresponding members from the EN-MCS and EN-NOMCS ensembles for PV250 and PMSL. The overall growth of RMSD in both output fields for all ensemble members clearly demonstrates that the MCS perturbations change the forecast evolution. The RMSD evolution is similar for all ensemble members for a given field and there is a tendency for the ranking of RMSD values among the ensemble members to be similar for both fields e.g. the greatest RMSD values at the end of the forecasts are found for ensemble members 2 and 3, and ensemble member 5 generally has the lowest values over the forecasts for both fields. The relative lack of difference growth for ensemble member 5, and so more similar evolution of the forecasts with and without the MCS perturbations, is consistent with the lack of precipitation simulated in the location of MCS-A in this ensemble member (irrespective of the addition of the MCS perturbations). This finding of lack of difference growth where moist convection has not been simulated is consistent with findings from other error growth studies showing the importance of accurate representation of moist processes (e.g. Zhang et al., 2003; Tan et al., 2004). In particular, Zhang (2005) found that maximum error growth in an ensemble forecast with both unbalanced and balanced ICs was found to occur in areas of active convection. For PV250 RMSD has three stages of difference growth (Fig. 11(a)). For the first 36 h little growth occurs followed by approximately linear growth, out to at least 72 h, and then saturation. In contrast, for PMSL the RMSD declines for about the first 12 h, then plateaus before increasing approximately linearly until the end of the forecast (Fig. 11(b)). The initial reduction in PMSL RMSD (and lack of growth in PV RMSD) is attributed to the damping of the effect of the MCS perturbations on the respective fields, as would be expected in response to the insertion of unbalanced perturbations (this effect was also noted in Part 1). The subsequent growth is attributed to the projection of the perturbations onto growing modes at the synoptic scale (also found for the later stages of convective perturbation growth by Zhang et al. (2007) and Selz and Craig (2015b)). Saturation of the RMSD values for PV prior to those for PMSL (the latter not obviously saturating within the 120 h forecast) is attributed to the more complex smaller-scale filamentary structure of PV250: larger-scale, smoother fields require larger spatial displacement between forecasts to obtain a complete mis-match between the fields and so for saturation in RMSD. Here, consistent with previous studies, we found that in all ensemble forecasts the error growth is maximized in the vicinity of the strongest mean PV gradient (along the Rossby wave guide; Fig. 8) and/or over the area of active moist convection (since member 5 has smaller error growth), consistent with the expected reduced predictability in these regions.

Unlike the RMSD between corresponding members from the EN-MCS and EN-NOMCS ensembles, the RMSD increases fairly linearly with time when considering the members of the EN-MCS and EN-NOMCS ensembles compared to the ensemble mean of the combined ensemble (Fig. 12). The evolution of RMSD is more similar for corresponding members of EN-MCS and EN-NOMCS than for different members in a given ensemble. Hence, while the addition of the MCS perturbations changes the forecast of PV, these changes are smaller than those arising from the IC perturbations used to create the two downscaled ensembles (consistent with the finding in Sec. 5). The larger-scale IC perturbations causing greater differences between forecasts is consistent with Bei and Zhang (2007) who found that larger-scale initial uncertainties generally led to larger forecast divergence. The MCS perturbations lead to a systematic difference in RMSD over the forecast period for most ensemble members. However, that difference is a decrease in RMSD for some members (e.g. member 1) and an increase for other members (e.g. member 2). Hence, MCS perturbations do not systematically either enhance or diminish the RMSD for all ensemble members. The RMSD of the Control members of the two ensembles is smaller (at all times) than that of the other members implying that the Control members are most similar to the combined ensemble mean, consistent with the IC perturbations perturbing the ensemble about the Control member.

The correspondence ratio for PV250 and precipitation (two thresholds for both) is shown as a function of time for the six-member EN-MCS and EN-NOMCS ensembles individually as well as for the combined 12-member ensemble in Fig. 13(a, b: blue and black lines). Values of CR generally decrease as a function of time indicating that, as expected, the spread of the ensemble increases with time. For surface precipitation rates (Fig. 13(b)) a diurnal cycle (with minimum values at 18 UTC) is also evident; the spread in the ensemble increases in the afternoon as surface heating leads to increased convection. Gebhardt et al. (2011) also calculated precipitation CR values, though for 12-member ensembles created using convection-permitting model simulations (with 2.8-km grid spacing) with either perturbed lateral boundary conditions, model physics perturbations or a combination of both perturbation types and for a threshold of  $0.1 \text{ mm h}^{-1}$  (higher than those used here). They found a rapid drop in CR down to about 0.1 over the first 12 hours of the simulations followed by a slower reduction over the remaining 12 hours. For the same threshold for our simulations, CR is comparable at 0.07 after 12 hours for the combined 12-member ensemble (not shown) although the structure of the time evolution is different. The much faster drop in CR in the Gebhardt et al. (2011) simulations is likely a consequence of their use of a convection-permitting model: errors have been found to grow ten times faster in convection-permitting than in convection-parametrizing models (Hohenegger and Schar, 2007a). The CR values are higher for PV250 (Fig.13(a)) than for precipitation rates (for the chosen thresholds) for the same forecast times and there is no diurnal signal in the PV CRs. The larger values (smaller spreads) are likely a consequence of the smoother synoptic-scale structure of the PV field relative to that of the precipitation rate field (i.e. a complete mis-match between forecasts is less likely for PV, although the results are of course threshold dependent). The evolutions of the CRs with time (for a given field and threshold) are very similar for the EN-MCS and EN-NOMCS ensembles (dotted and dashed lines) indicating that the MCS perturbations do not strongly impact the ensemble spread of these individual ensembles. In contrast, the CRs for the combined ensembles (solid lines) are less than those of the individual ensembles (for both fields and all thresholds). Hence, the MCS perturbations do lead to a systematically-different evolution of the forecast implying that the MCS perturbations create different realisations of the forecast that could be useful for operational guidance for forecasters.

Although the analysis above shows that the addition of six additional ensemble members with MCS perturbations to an ensemble of six members created using different ICs does add some spread, the question arises of how big this additional spread is compared to that gained by, for example, a single additional ensemble member created through IC perturbations. The grey lines in Fig. 13(a, b) show the CR from all possible six-member ensembles created using matched pairs of members from the EN-MCS and EN-NOMCS ensembles (e.g. members 1, 2, and 3 or members 2, 5, and

6 from the two ensembles). CR is increased compared to the values from the six-member EN-MCS and EN-NOMCS ensembles, demonstrating that a six-member ensemble created solely from six IC perturbations has far greater spread than one created from three IC perturbations with the simulations run both with and without MCS perturbations; consistent with this, Figs. 10 and 12 show the greater impact of the IC, relative to MCS perturbations, on RMSE and RMSD respectively. Figure 13(c, d) illustrates the relative effects of the two types of perturbations in more detail (and the following conclusions are reached for both PV250 and precipitation rates fields, and for both thresholds used for each field). The CR at the end of the simulations for ensembles generated using 3–6 matched pairs of members from the EN-MCS and EN-NOMCS ensembles (total ensemble size of 6–12 members) can be compared with the equivalent CR values from the six-member EN-MCS and EN-NOMCS ensembles. As the number of matched pairs used increases, the ensemble spread increases (CR values decrease) as expected. The CR values for ensembles with ten ensemble members (five matched pairs) are less than those for the six-member EN-MCS and EN-NOMCS ensembles. So, the effect on the spread of doubling the ensemble members by adding MCS perturbations to all the simulations in the ensemble is somewhat equivalent to adding a single additional ensemble member generated using IC perturbations.

## 7 | CONCLUSIONS

Moist convection, such as in mesoscale convective systems, has long been known to be a source of error in NWP models due to the way convection is represented by parametrization schemes ultimately leading to a loss of predictability in the forecasts (see Sec. 1). The possible implications for ensemble weather forecasting of poor representation of MCSs by convection-parametrizing models have been explored here for a single case study using the convection-permitting Euro4 and convection-parametrizing Global configurations of the MetUM. In the companion paper (Clarke et al., 2019) which used the same case study, the PV anomalies associated with two MCSs, termed MCS-A and -B, that formed over France were shown to be more intense, even after coarse-graining to the same resolution, in the output from the Euro4 configuration simulation than in that from the Global configuration simulation. This result suggests that the poor representation of the MCSs by convection-parametrizing models can impact synoptic-scale medium-range flow evolution. This hypothesised impact was tested by perturbing Global configuration five-day forecasts with perturbations in winds and  $\theta$  (model-accepted surrogates for PV) derived by calculating the difference between the Euro4 and Global configuration forecasts in the region of the MCSs, so-called MCS perturbations. The impact was found to lead to synoptic-scale modification of the downstream flow evolution, extending across the entire hemisphere by five days, and so be a potential source of error in convection-parametrizing weather forecasts.

In the current paper, the systematic impact of the MCS perturbations and their impact on ensemble spread has been assessed by inserting them into six Global configuration simulations downscaled from simulations generated using the MOGREPS-G ensemble system with operational IC perturbations. Different MCS perturbations were used for each ensemble member calculated as the differences between outputs from the Global and downscaled Euro4 configuration simulations. Even convection-permitting models struggle to predict exactly where and how intense a given MCS will be although they can provide more realism than convection-parametrizing forecasts in the general region of likely MCS development. Hence, the insertion of MCS perturbations into a single deterministic convection-parametrizing forecast will not necessarily improve that forecast. However, we hypothesize that the use of these perturbations will provide an improved representation of the inherent uncertainty associated with an MCS (or MCSs) in an ensemble and test that hypothesis here. An advantage of this approach using physically-derived perturbations to improve probabilistic predictions by enhancing ensemble spread, compared to other statistically-derived approaches such as stochastic physics perturbations (e.g. Sanchez et al., 2015), is that it maintains physically realistic structures. Precipitation in

the region of the observed MCS-A occurred in five of the six ensemble members and in the region of MCS-B in all the members independent of the addition of MCS perturbations. However, the perturbations modified the structure and intensity of the precipitation patterns. After five days the PV250 differences due to the MCS perturbations lay along the tropopause consistent with findings from previous studies of error growth related to the representation of moist processes using PV-based analysis (Baumgart et al., 2019; Joos and Forbes, 2016; Schäfer and Harnisch, 2015), but the position of these differences varied between the ensemble members corresponding to the different Rossby wave structures generated by the IC perturbations. Despite this variability, the MCS perturbations have been shown to systematically shift the location of a surface low pressure system over northern Russia, downstream of the location where the MCS perturbations were inserted, to the west of northwest; this shift is consistent with other studies that have found a slowed eastwards Rossby waves propagation due to intense convection (Grams and Archambault, 2016; Ojinting and Jones, 2016; Riemer et al., 2008; Rodwell et al., 2013).

Member-member RMSD calculations reveal three stages of difference growth when considering PV250 (little growth spatially, approximately linear growth as differences spread onto the large scale, and then saturation). These stages are consistent with findings from previous studies of error growth (e.g Zhang et al., 2003; Tan et al., 2004; Selz and Craig, 2015c). A different evolution occurs when considering PMSL (difference damping, plateauing and then approximately linear growth until the end of the forecast). The initial damping of the PMSL differences is attributed to the model damping the unbalanced inserted MCS perturbations and the (presumed) later saturation to the larger-scale smoother nature of the field of PMSL compared to that of PV250. The PV250 RMSD between ensemble members and the corresponding ensemble mean reveals that the changes to the forecast arising from the IC perturbations dominate over those arising from the MCS perturbations. This is not surprising since the IC perturbations are inserted over the whole globe in contrast to the insertion of the MCS perturbations over a region extending from southern UK to northern Spain; future work could compare the impact of IC and MCS perturbations inserted over the same region. Similarly, the CRs for the fields of both PV250 and precipitation demonstrate that the MCS perturbations do not strongly impact ensemble spread (i.e. the ensemble spread is very similar for the two ensembles with and without the perturbations). However, the ensemble spread for the combined ensemble (members with and without MCS perturbations) is larger than that of the individual ensembles (i.e. has smaller CRs) implying that the MCS perturbations have led to a systematically different evolution of the forecast. Bei and Zhang (2007) found that larger-scale, larger-amplitude IC uncertainties generally led to larger forecast divergence than uncertainties of smaller scales and smaller amplitude which further explains the fact that the ensemble ICs have more of an impact on the differences between members than the MCS perturbations.

Here, physically-realistic MCS perturbations have been demonstrated to grow and impact downstream ensemble forecast evolution (in both surface and tropopause-level fields) for a five-day forecast. The MCS perturbations enhance the ensemble spread with the degree of enhancement from doubling the ensemble size by adding MCS perturbations to a forecast comparable to that from adding a single additional ensemble member generated through IC perturbations. Other studies have found an increase in ensemble spread occurs when using a stochastic convection scheme (e.g. Groenemeijer and Craig (2012) found an improvement using the scheme of Plant and Craig (2008). This scheme was also found to better represent the convective variability near the grid scale for deterministic forecasts (Selz and Craig, 2015a). The Global Met Office ensemble (MOGREPS-G) (Bowler et al., 2008) used to include a stochastic physics scheme called the stochastic convective vorticity (scv) scheme to represent PV anomaly dipoles similar to those typically associated with MCSs. This was implemented after Gray (2001) found that the PV signatures associated with MCSs were not well represented in the MetUM and this had an impact on the downstream forecast evolution. This PV dipole consisted of an upper-level anticyclone consisting of a zero-PV core and a mid-level cyclonic vortex; the scales of the vortices were determined by a randomised function (Bowler et al., 2008). The scv scheme was later removed and

replaced by a stochastic kinetic energy backscatter scheme (which uses wind perturbations) as it was found to provide little improvement to the skill of the ensemble. However, the scv did contribute to a slight increase in growth of the spread within the ensemble (Tennant et al., 2011). Shutts (2017) describes how this stochastic convective backscatter scheme (described fully in Shutts (2015)) could be tuned to better represent model wind error produced by an MCS at jet stream level for an ensemble forecast. We suggest that the scv scheme could have been more effective by using winds and temperatures associated with negative PV anomalies at upper-levels instead of zero-PV anomalies since we found in Clarke et al. (2019) that it is the negative PV anomalies at upper-level that have the largest impact on the downstream forecast. It is apparent that taking account of the unpredictability of MCSs (or other deep convection) and their misrepresentation by convection-parametrizing models, particularly at upper levels where the largest errors may develop due to divergent outflow from convection, is required to produce an ensemble with a large enough spread.

Our study has used physically-realistic perturbations of wind and temperatures to account for differences in the PV field due to the misrepresentation of convection associated with MCSs by a convection-parametrizing ensemble. This approach could be implemented operationally using idealised MCS perturbations in regions where MCSs are likely to form (e.g. where there is large convective available potential energy) or have been diagnosed from observations such as satellite imagery, or 'real' MCS perturbations if convection-permitting simulations are available over suitable regional domains. For example, we believe there would be usefulness for ~2–3 day Western Europe ensemble forecasts in re-running a global ensemble (or a few members) with the inclusion of convectively-generated convection-permitting model perturbations over the US over a convective period (ahead of the next cycle). It would require a convection-permitting ensemble, downscaled within the global ensemble, to be run over the deep convective areas of the USA and a large number of test cases would be needed to examine the feasibility and benefit. Deterministic global weather forecasts are of course now being run at much higher resolution than the 25-km grid spacing used here: for example the ECMWF now runs its deterministic model with 9-km grid spacing although their ensemble uses a coarser 18-km grid spacing. Even global ensembles are being tested in research mode at sub-10km (Zhang et al., 2019). However, these resolutions are still very much in the convective "grey zone", even for MCSs, and it is likely that much of the convection will still be represented by a convection parametrization scheme and therefore be subject to the same limitations as at coarser resolutions. Larger storms or MCSs may be represented explicitly, but that would have to involve unphysically wide updraughts especially during early development. Synthetic imagery from the deterministic MetUM global model and the ECMWF model compared with satellite imagery (not shown) reveals that even at ~10-km grid spacing large storm anvil structures can be reduced or missing in the models. The benefit of using physically-realistic structures to account for model error introduced by convection is, as noted by Shutts (2015), to avoid some of the more speculative aspects of a stochastic physics parametrization.

The focus in this study has been on MCSs that affect France and the UK because of our access to the Met Office's Euro4 configuration – a convection-permitting operational NWP model configuration with a relatively large domain (for convection-permitting models) that covers this region of MCS initiation and development. An obvious limitation of this study is that we have used a single case study and specific model configurations. Although the Euro4 configuration is classed as convection permitting, the 4.4-km grid spacing is in the convective grey zone. While the precipitation field simulated at this resolution looks more realistic than that from a convection-parametrizing model, the showers are typically too few, too intense and too organised compared to observations (Clark et al., 2016). This raises the question of whether the MCS perturbations used in the experiments here are realistic and particularly whether they are too large in amplitude. However, MCSs in other parts of the globe are more frequent and intense than those over northern Europe. For example Augustine and Howard (1991) identified 58 and 44 mesoscale convective complexes (a more intense subset of MCSs) in 1986 and 1987, respectively, over the US and Rodwell et al. (2013) linked MCSs in this region to subsequent European forecast busts. Hence, the MCS perturbations used in this study are likely to be within the

range of reasonable magnitudes for such perturbations and an obvious future experiment is to repeat this study for more vigorous MCS cases over a different global region. Another question is the extent to which the Global simulations can retain the energy from the introduction of the unbalanced MCS perturbations given that they are inserted over a single hour. An alternative approach could be to insert appropriate MCS perturbations into the ensemble forecasts at regular intervals over a longer period e.g. spanning the time from soon after the initialisation of the Euro4 simulations at 0600 UTC 5 July (prior to this the MCS perturbations cannot be calculated) to the time of maturity of the MCSs at about 1900 UTC on the same day. This is beyond the scope of this study, but could provide benefit in that the Global model would start adjusting to the MCS perturbations earlier and repeated inclusion of MCS perturbations may help the Global model to maintain more of the energy in these perturbations, so reducing damping. However, although this approach may enhance the impact of the MCS perturbations on the downstream evolution, the spatial pattern of the response is unlikely to be altered much since previous studies have noted that forecast error caused by rapid upscale error growth on convective scales from moist convection is less sensitive to the scale and amplitude of the initial perturbations (e.g. Sun and Zhang, 2016; Zhang et al., 2007; Hohenegger and Schar, 2007b) than error growth in baroclinic systems where the growth rate is dependant on the structure and amplitude of the initial perturbation (Yan and Thorpe, 2006). We also found in Clarke et al. (2019) that increasing the amplitude of the perturbations had no impact on the spatial extent of the impact downstream, but did cause the differences to be larger in magnitude (similarly, also found in Zhang et al. (2003); see their Fig. 5).

## 8 | ACKNOWLEDGEMENTS

ERC funded the Natural Environment Research Council studentship 'Multiscale prediction and upscale impact of mesoscale convective systems' (Award reference: 1110110) with a CASE award from the Met Office. We acknowledge the use of the MONSoon system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, which is a strategic partnership between the Met Office and NERC. Operational Met Office analyses, and the initial and boundary condition files needed to run the MetUM simulations on MONsoon, were provided by the Met Office. The data used in this work is available by contacting the corresponding author and is subject to licensing constraints. We acknowledge the interesting questions and helpful comments of two anonymous reviewers.

## REFERENCES

- Augustine, J. A. and Howard, K. W. (1991) Mesoscale convective complexes over the United States during 1986 and 1987. *Mon. Wea. Rev.*, **119**, 1575–1589.
- Baumgart, M., Ghinassi, P., Wirth, V., Selz, T., Craig, G. and Riemer, M. (2019) Quantitative view on the processes governing the upscale error growth up to the planetary scale using a stochastic convection scheme. *In press, Mon. Wea. Rev.*
- Bei, N. and Zhang, F. (2007) Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the mei-yu front of china. *Q. J. R. Meteorol. Soc.*, **133**, 83–99.
- Bloom, S. C., L. L. Takacs, A. M. S. and Ledvina, D. (1996) Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.
- Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B. and Beare, S. E. (2008) The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **134**, 703–722.
- Chagnon, J. and Gray, S. L. (2009) Horizontal potential vorticity dipoles on the convective storm scale. *Q. J. R. Meteorol. Soc.*, **135**, 1392–1408.



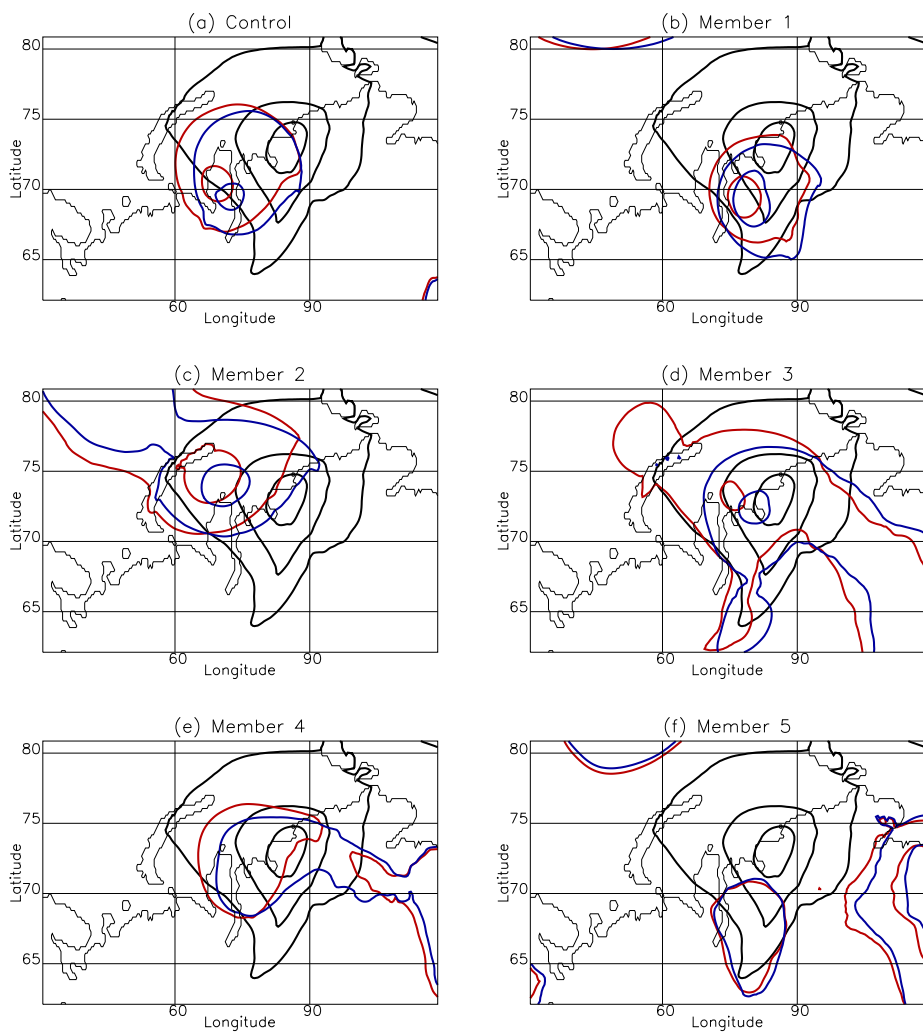
- Clark, P., Roberts, N., Humphrey, L., Ballard, S. P. and Charlton-Perez, C. (2016) Convection-permitting models: A step-change in rainfall forecasting. *Meteorol. Appl.*, **23**, 165–181.
- Clarke, S. J., Gray, S. L. and Roberts, N. M. (2019) Downstream influence of mesoscale convective systems: Part 1, influence on forecast evolution. *submitted to Q. J. R. Meteorol. Soc.*
- Clayton, A. (2012) Incremental analysis update (IAU) scheme. Unified model documentation paper no.31. *Tech. rep.*, Met Office, FitzRoy Road, Exeter, UK, EX1 3PB.
- Done, J. M., Craig, G. C., Gray, S. L., Clark, P. A. and Gray, M. E. B. (2006) Mesoscale simulations of organized convection: Importance of convective equilibrium. *Q. J. R. Meteorol. Soc.*, **132**, 737–756.
- Geleyn, J. E., Theis, S. E., Paulat, M. and Bouallègue, Z. B. (2011) Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmos. Research*, **100**, 168–177.
- Giam, C. M. and Archambault, H. M. (2016) The key role of diabatic outflow in amplifying the midlatitude flow: a representative case study of weather systems surrounding western North Pacific extratropical transition. *Mon. Wea. Rev.*, **144**, 3847–3869.
- Gray, M. E. B. (2001) The impact of mesoscale convective-system potential vorticity anomalies on numerical-weather-prediction forecasts. *Q. J. R. Meteorol. Soc.*, **127**, 73–88.
- Gregory, D. and Rowntree, P. R. (1990) A mass-flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506.
- Groenemeijer, P. and Craig, G. C. (2012) Ensemble forecasting with a stochastic convective parametrization based on equilibrium statistics. *Atmospheric Chemistry and Physics*, **12**, 4555–4565.
- Hohenegger, C. and Schar, C. (2007a) Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1793.
- (2007b) Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.*, **64**, 4467–4478.
- Houze, R. A. (2004) Mesoscale Convective Systems. *Rev. Geophys.*, **42**, 43pp.
- Joos, H. and Forbes, R. M. (2016) Impact of different ifs microphysics on a warm conveyor belt and the downstream flow evolution. *Q. J. R. Meteorol. Soc.*, **142**, 2727–2739.
- Lean, F. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R. and Halliwell, C. (2008) Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136**, 3408–3424.
- Lewis, M. W. and Gray, S. L. (2010) Categorisation of synoptic environments associated with mesoscale convective systems over the UK. *Atmos. Res.*, **97**, 194–213.
- Mapes, B., Tulich, S., Nasuno, T. and Satoh, M. (2008) Predictability aspects of global aqua planet simulations with explicit convection. *J. Meteor. Soc. Japan*, **86A**, 175–185.
- Perkey, D. J. and Maddox, R. A. (1985) A numerical investigation of a mesoscale convective system. *Mon. Wea. Rev.*, **113**, 553–566.
- Plant, R. S. and Craig, G. C. (2008) A stochastic parametrization for deep convection based on equilibrium statistics. *J. Atmos. Sci.*, **65**, 87–105.
- Quinting, J. F. and Jones, S. C. (2016) On the impact of tropical cyclones on Rossby wave packets: A climatological perspective. *Mon. Wea. Rev.*, **144**, 2021–2048.

- Riemer, M., Jones, S. C. and Davies, C. A. (2008) The impact of extratropical transition on the downstream flow: An idealized modelling study with a straight jet. *Q. J. R. Meteor. Soc.*, **134**, 69–91.
- Roberts, N. M. (2003) The impact of a change to the use of the convection scheme to high-resolution simulations of convective events. forecasting research technical report 407. *Tech. rep.*, Met Office, FitzRoy Road, Exeter, EX1 3PB, United Kingdom. Available from the National Meteorological Library and Archive.
- Rodwell, M. J., Magnusson, L., Bauer, P., Betchtold, P., Bonavita, M., Cardinali, C. and Diamantakis, M. (2013) Characteristics of occasional poor medium-range weather forecasts for Europe. *Bull. Amer. Meteor. Soc.*, **94**, 1393–1405.
- Sanchez, C., Williams, K. D. and Collins, M. (2015) Improved stochastic physics schemes for global weather and climate models. *Q. J. R. Meteorol. Soc.*, **142**, 147–159.
- Schäfler, A. and Harnisch, F. (2015) Impact of the inflow moisture on the evolution of a warm conveyor belt. *Q. J. R. Meteorol. Soc.*, **141**, 299–310.
- Selz, T. and Craig, G. C. (2015a) Simulation of upscale error growth with a stochastic convection scheme. *Geophys. Res. Lett.*, **42**, 3056–3062.
- (2015b) Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Mon. Wea. Rev.*, **143**, 813–827.
- (2015c) Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Mon. Wea. Rev.*, **143**, 813–827.
- Shutts, G. (2017) Idealized numerical simulations of Mesoscale Convective Systems and their implications for forecast error. *Q. J. R. Meteorol. Soc.*, **143**, 1608–1619.
- Shutts, G. J. (2015) A stochastic convective backscatter scheme for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.*, **141**, 2602–2616.
- Stensrud, D. J. (1996) Effects of persistent, midlatitude mesoscale regions of convection on the large-scale environment during the warm season. *J. Atmos. Sci.*, **53**, 3503–3527.
- Stensrud, D. J. and Wandishin, M. S. (2000) The correspondence ratio in forecast evaluation. *Wea. Forecasting*, **15**, 593–602.
- Sun, Y. Q. and Zhang, F. (2016) Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73**, 1419–1438.
- Tan, Z.-M., Zhang, F., Rotunno, R. and Snyder, C. (2004) Mesoscale predictability of moist baroclinic waves: Experiments with parameterized convection. *J. Atmos. Sci.*, **61**, 1794–1804.
- Tennant, W. J., Shutts, G. J., Arribas, A. and Thompson, S. A. (2011) Using a stochastic kinetic energy backscatter scheme to improve mogreps probabilistic forecast skill. *Mon. Wea. Rev.*, **139**, 1190–1206.
- Wolf, B. J. and Johnson, D. R. (1995) The mesoscale forcing of a midlatitude upper-tropospheric jet streak by a simulated convective system. Part I: Mass circulation and ageostrophic processes. *Mon. Wea. Rev.*, **123**, 1059–1087.
- Zhang, F. (2005) Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon. Wea. Rev.*, **113**, 2876–2893.
- Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S., Chen, J. and Emanuel, K. (2019) What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091.
- Zhang, F., Snyder, C. and Rotunno, R. (2003) Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173–1185.

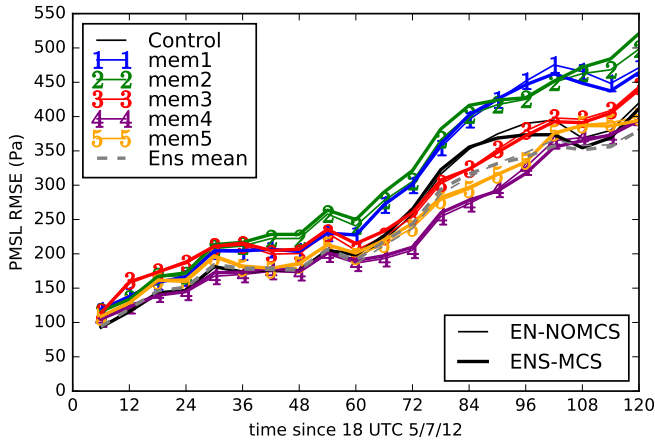
Zhang, F. C., Bei, N., Rotunno, R., Snyder, C. and Epifanio, C. (2007) Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, **64**, 3579–3594.

Zhang, F. C., Snyder, C. and Rotunno, R. (2002) Mesoscale predictability of the "surprise" snowstorm of 24-25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632.

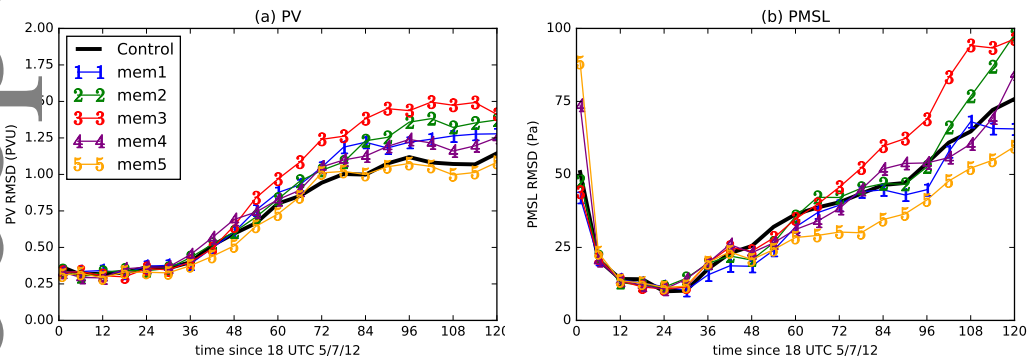
7 Lu, H. and Thorpe, A. (2006) Predictability of extratropical cyclones: The influence of initial condition and model uncertainties. *J. Atmos. Sci.*, **63**, 1483–1497.



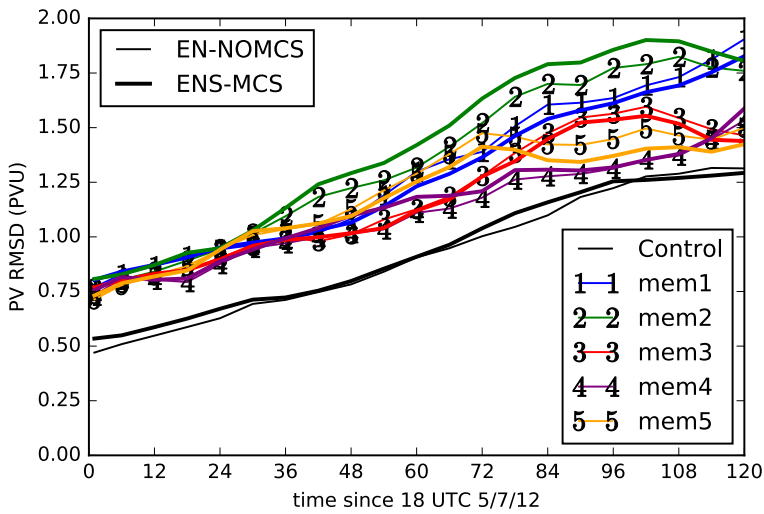
**FIGURE 9** PMSL over northern Russia at 18 UTC 10 July 2012 for ensemble member (a) Control, (b) 1, (c) 2, (d) 3, (e) 4 and (f) 5 from the EN-NOMCS (blue contours) and EN-MCS (red contours). Each panel also shows the PMSL from the analysis (black contours). Contours are plotted at 984, 992 and 1000 hPa though all three contours are only needed for the more intense analysis PMSL.



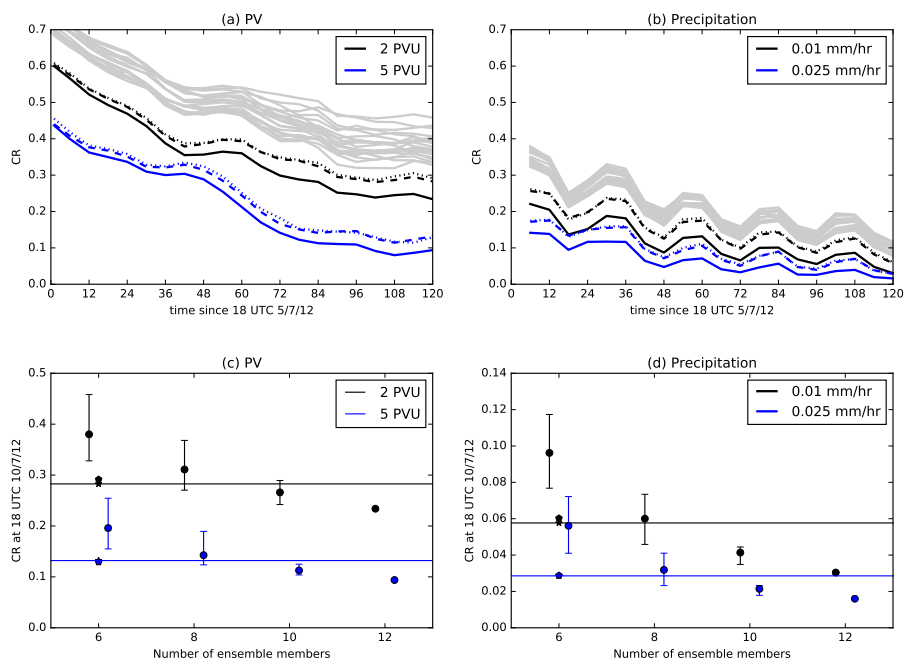
**FIGURE 10** RMSE (individual ensemble members and ensemble means compared to the six-hourly analyses) for PMSL from 0 UTC 6 July to 18 UTC 10 July 2012 calculated over the European sub-domain (30–70°N and 50°W–100°E) for each member (thin solid lines with number markers for ensemble members) and EN-MCS (corresponding thick lines without markers which can be difficult to distinguish due to overlapping). The ensemble means of the EN-NOMCS and EN-MCS ensembles are shown in thin and thick dashed lines, respectively.



**FIGURE 11** RMSD from 19 UTC 5 July to 18 UTC 10 July 2012 calculated over the European sub-domain (30–70°N and 50°W–100°E) for for each member of EN-MCS compared to the corresponding member of EN-NOMCS for (a) PV250 and (b) PMSL.



**FIGURE 12** RMSD for 250 hPa PV from 19 UTC 5 July to 18 UTC 10 July 2012 calculated over the European sub-domain (30–70°N and 50°W–100°E) for each member EN-NOMCS (thin lines with number markers for ensemble members) and EN-MCS (corresponding thick lines without markers) compared to the ensemble mean calculated from the combined ensemble.



**FIGURE 13** CR values calculated for the European subdomain ( $30\text{--}70^\circ\text{N}$  and  $50^\circ\text{W}\text{--}100^\circ\text{E}$ ). (a, b) CRs for (a) PV250 (thresholds of 2 and 5 PVU) from 19 UTC 5 July to 18 UTC 10 July 2012 and (b) surface precipitation rate (thresholds of 0.01 and 0.025  $\text{mm h}^{-1}$ ) from 0 UTC 6 July to 18 UTC 10 July 2012. CRs are shown for the EN-MCS (dashed lines) and EN-NOMCS ensembles (dotted lines) and the combined 12-member ensemble (solid lines) for both thresholds. CRs are also shown for all 20 possible combinations of six-member ensembles composed of three matched pairs of EN-MCS and EN-NOMCS members for the lower thresholds only (grey solid lines). (c, d) CR values at 18 UTC 10 July 2012 for (c) PV and (d) precipitation rates for different total ensemble sizes. Circle symbols with error bars show the mean and range of the CR values for ensembles comprised of different numbers of matched ensemble pairs (corresponding members from the EN-MCS and EN-NOMCS ensembles: 20, 15, 6 and 1 combinations for 3, 4, 5 and 6 matched pairs, respectively). Pentagon and star symbols show the CR for the six-member EN-NO-MCS and EN-MCS ensembles, respectively (these symbols often overlap), and horizontal lines are drawn across the panels at the CR values for the six-member EN-MCS ensemble (passing through the star symbols) to add interpretation. Note that x-axis locations of the data are offset slightly for the two different thresholds to avoid overlapping.