Quarterly Journal of the Royal Meteorological Society

Perturbation growth at the convective scale for CSIP IOP18

Journal:	*QJRMS*
Manuscript ID:	QJ-09-0025.R2
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	16-Dec-2009

Complete List of Authors:	Leoncini, Giovanni; University of Reading, Meteorology Plant, Robert; University of Reading, Meteorology Gray, Suzanne; University of Reading, Meteorology Clark, Peter; Met Office

Keywords:	convective-scale forecasting, quantitative precipitation forecasting, root-mean-square error, ensemble forecasting

Q. J. R. Meteorol. Soc. **00**: 1–27 (2009)

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.000

Perturbation growth at the convective scale for CSIP IOP18

G. Leoncini*, R. S. Plant, S. L. Gray and P. A. Clark
Department of Meteorology, University of Reading, Reading, UK

Abstract: The Met Office Unified Model is run for a case observed during Intensive Observation Period 18 (IOP18) of the Convective Storms Initiation Project (CSIP). The aims are to identify the physical processes that lead to perturbation growth at the convective scale in response to model-state perturbations and to determine their sensitivity to the character of the perturbations. The case is strongly upper-level forced but with detailed mesoscale/convective-scale evolution that is dependent on smaller-scale processes. Potential temperature is perturbed within the boundary layer. The effects on perturbation growth of both the amplitude and typical scalelength of the perturbations are investigated and perturbations are applied either sequentially (every 30 min. throughout the simulation) or at specific times.

The direct effects (within one timestep) of the perturbations are to generate propagating Lamb and acoustic waves and produce generally small changes in cloud parameters and convective instability. In exceptional cases a perturbation at a specific gridpoint leads to switching of the diagnosed boundary-layer type or discontinuous changes in convective instability, through the generation or removal of a lid. The indirect effects (during the entire simulation) are changes in the intensity and location of precipitation and in the cloud size distribution. Qualitatively different behaviour is found for strong (1 K amplitude) and weak (0.01 K amplitude) perturbations, with faster growth after sunrise found only for the weaker perturbations. However, the overall perturbation growth (as measured by the root-mean-square error of accumulated precipitation) reaches similar values at saturation, regardless of the perturbation characterisation.

KEY WORDS convective-scale forecasting; quantitative precipitation forecasting; root-mean-square error; ensemble forecasting

Received December 16, 2009

1 Introduction

Severe rainfall from convective events is the leading cause of floods and flash floods over the summer months in the UK (Hand *et al.*, 2004). The high societal impact of such floods means that accurate forecasting of severe convective events could greatly improve flood forecasting and specifically flash-flood forecasting, as highlighted by Collier (2006).

Increased computational power has recently made numerical weather prediction possible over large domains with grid spacings that allow convection to be, at least partially, resolved. For example, the Met Office, at the time of writing, runs operationally at 4 km grid spacing over the entire UK and is trialling a 1.5 km grid spacing, on a similar domain. Also, the National Center for Environmental Prediction has been running the WRF-ARW model at 4 km since 2003 (e.g. Weisman *et al.*, 2008). While

*Correspondence to: G. Leoncini, Department of Meteorology, University of Reading, PO Box 243, Reading, Berkshire, RG6 6BB, UK. E-mail: g.leoncini@reading.ac.uk

such grid spacings are not sufficient to resolve the individual convective elements properly (e.g. Bryan *et al.*, 2003) such 'convection permitting' simulations are generally able to describe convective phenomena more realistically than simulations with $\sim$10 km grid spacing (e.g. Weisman *et al.*, 2008).

The predictability of the atmosphere at the convective scale is different from that at the synoptic scale: error growth rates are around 10 times larger and the tangent linear approximation breaks down within a couple of hours (Hohenegger and Schar, 2007a). This rapid loss of linearity implies a fundamental qualitative difference between convective-scale and synoptic-scale forecasting. Poor convective-scale predictability is most likely due to the significant nonlinearities of the atmosphere at smaller scales: microphysics, turbulence, radiation and flow dynamics are strongly coupled and can act to amplify both model and observation uncertainties. This makes ensemble prediction systems particularly valuable because they provide a measure of confidence in the forecast, but at the same time it renders the large-scale methodologies for perturbation generation less likely to be effective (Hohenegger and Schar, 2007a).

Despite these difficulties the research into ensemble prediction systems at the convective scale is a developing field, and Kong *et al.* (2006, 2007) described a first attempt to design an ensemble prediction system for a full-physics numerical model using operational initial conditions. Specifically they tested different methodologies over three nested grids with 24, 6 and 3 km grid spacing, applying the scaled lagged average forecasting technique (Ebisuzaki and Kalnay, 1991) to a tornadic storm. They found that the associated perturbations grew too slowly and produced little spread. However, the spread improved significantly when the perturbations, derived from the difference between a previous forecast and verifying analysis, were scaled based on their amplitude, rather than by using the age of the forecast that generated them. Moreover, they point out that although the radar reflectivity patterns had greater spatial fidelity for the ensemble members with 3 km grid spacing, conventional skill scores (root mean square error, Brier score etc) do not always reflect such improvements. Lean *et al.* (2008) also found that simulations with 1 and 4 km grid spacings often give more realistic-looking precipitation fields (compared to those from simulations with 12 km grid spacing) and showed that a scale-selective precipitation verification technique can be used to demonstrate the improved performance.

Other studies (e.g. Zhang *et al.*, 2003; Walser *et al.*, 2004; Hohenegger *et al.*, 2008a,b) have shown that ensembles of convection-permitting simulations generated by perturbing the initial conditions or varying the lateral boundary conditions (LBCs) can be used to investigate the predictability of specific events. Gebhardt *et al.* (2009) ran the COSMO-DE model with 2.8 km grid spacing using different LBCs and varying parameters for a few physics schemes. Their results show how the different physics determines the spread for the first few hours, while the LBCs become more important later. Hohenegger and Schar (2007b) determined that fast, domain-wide perturbation growth in their simulations occurred due to the propagation both of small amplitude, fast acoustic waves (and/or numerical noise) and of large amplitude, slower gravity waves. This then leads to triggering and/or error growth in regions of moist convective instability. The following conclusions emerge from the cited studies (all of which directly or indirectly address the feasibility of ensemble prediction systems at convective scales): a) moist convection and nonlinearities in general strongly

DOI: 10.1002/qj

favour rapid error growth with typical time scales of the order of an hour, b) the presence of moist convection alone does not necessarily imply low predictability because of a strong dependence upon the weather regime and, c) model and LBC uncertainties also affect predictability; model uncertainties seem to dominate for the first few hours and LBC uncertainties after that.

There are two goals of this study: first, to investigate the use of a simple technique for perturbing the model state (perturbations of potential temperature) and second, to determine the cause, or causes, of the resulting perturbation growth at the convective scale for a convective event over the United Kingdom. The term "*perturbation growth*" here indicates the divergence of the ensemble members from the control run as a result of perturbations, rather than the divergence from observations for which the term "*error growth*" would be more appropriate. A verification against observations for several cases would be necessary to test the effectiveness of the technique for numerical weather prediction purposes and is outside the scope of this study.

The paper is structured as follows. The main features of the convective event, the model used and the control run are presented in Section 2. The perturbation strategy and characteristics are described in detail in Section 3 and a description of the diagnostics used is given in Section 4. Results are presented in Sections 5 and 6 and a summary and conclusions are provided in Section 7.

2 Case Overview

2.1 The case

The Convective Storm Initiation Project (CSIP Browning *et al.*, 2007) was carried out during June-August 2005. The objective was to improve understanding of the mechanisms that determine precisely when and where deep convective clouds initiate. The observational focus was on Southern England, and an overview of all 18 Intensive Observing Periods (IOPs) can be found in Browning and Morcrette (2006). IOP 18, which occurred on 25 August 2005, was chosen for this study because the convection was primarily forced by a large-scale upper level trough (suggesting predictability in the synoptic-scale forecast), but the evolution of the intense convective storms was dependent on secondary convective initiation driven by internal dynamics arising from cold downdraughts (suggesting that the details of the convective evolution will be sensitive to model state perturbations).

The main features of the synoptic scale weather for that day were well forecast (Clark and Lean, 2006) and are the cold front over the western edge of the European continent and the centre of the associated low pressure system to the north of the British Isles yielding westerly flow over the UK. Southern England lay below a tropopause fold running roughly along the southern coast of England. This led to widespread scattered convection not only over land but also over the surrounding seas. A squall line developed from a line of showers at 1015 UTC and formed a distinct arc by 1130 UTC; precursor cells formed at about 0815 UTC near the Bristol Channel. The squall line travelled East South-East to reach the East Coast of Southern England at about 1400 UTC. A radar analysis of the rain rates at 1000 UTC is shown in Figure 5(d). A more comprehensive description of the synoptic and mesoscale observations can be found in Browning and Morcrette (2006) and in Clark et al. (submitted); the latter also includes a detailed analysis of the squall line.

2.2 Model and model set up

Version 6.1 of the Met Office Unified Model (MetUM) was used in this study. The model solves non-hydrostatic,

Prepared using qjrms3.cls
DOI: 10.1002/qj

deep-atmosphere dynamics using a semi-implicit semi-Lagrangian numerical scheme (Davies *et al.*, 2005). The horizontal grid is rotated in latitude/longitude with Arakawa C staggering. The vertical grid is terrain following with a hybrid-height vertical coordinate and Charney-Phillips staggering. For this study, the model is run with 38 vertical levels and a horizontal grid spacing of 4 km one-way nested within a domain with 12 km grid spacing. The model is currently run operationally at these resolutions (albeit over a larger domain than that used here). The 4 km grid spacing domain is centred over the UK, has 288×360 grid-points and is the focus of this study.

The results presented here are based on a slightly cropped domain (as shown in Figure 5 for example), which has been stripped of 25 grid points on each side in order to avoid any spin-up effects associated with the forced lateral boundaries. The LBCs for the 4 km simulations were provided by a 12 km simulation which in turn used LBCs from the operational global model. The 4 and 12 km limited area simulations were started at 0100 UTC on 25 August 2005. The 0100 UTC initial conditions were obtained from the operational global simulation and thus incorporated the operational data assimilation that was completed prior to this time. No additional data assimilation was performed for the 4 km simulations or during the 12 km forecast and therefore the runs here were started at 0100 UTC (rather than at a later time) in order to allow the spin-up stage of the evolution to be completed before sunrise (which occurred at 0500 UTC).

The MetUM has a comprehensive set of parameterisations, including a surface layer scheme (Essery *et al.*, 2001), radiation scheme (Edwards and Slingo, 1996) and mixed-phase cloud microphysics scheme (Wilson and Ballard, 1999). The convection and boundary layer parameterisations are key to this study and so briefly described here. The convection scheme of Gregory and Rowntree (1990) is used for both the 12 and 4 km grid-spacing simulations, but with a modification developed by Roberts (2003) applied at the higher resolution. The Gregory and Rowntree (1990) scheme has a trigger dependent on the initial parcel buoyancy and a mass-flux determined by a specified timescale for adjustment of Convective Available Potential Energy (CAPE), typically 30 minutes. The Roberts (2003) modification avoids the accumulation of high values of CAPE at the gridscale (which can lead to unphysical "gridpoint storms") by specifying the CAPE adjustment timescale as an increasing function of the CAPE. This allows the model to resolve explicitly most of the deep convection, with the parameterisation scheme dealing mainly with shallow convection. This modification was specifically designed for the 4 km grid-spacing configuration of the MetUM and has proved reasonably successful (Lean *et al.*, 2005; Roberts and Lean, 2008).

Seven types of boundary layers are identified in the boundary layer parameterisation scheme: stable, stratocumulus over stable, well mixed, decoupled stratocumulus over cumulus, decoupled stratocumulus not over cumulus, cumulus capped and shear driven boundary layer. The first six of these are described by Lock *et al.* (2000) with the shear-driven type being a more recent addition. The categorisation of each grid column is based on the adiabatic ascent of a parcel (rising from 10 m above the ground) and on its descent from cloud top. To avoid oversensitivity to grid-level noise, a constant 0.4 K is added to the temperature in addition to a locally derived buoyancy- and stability-dependent perturbation before calculating the parcel ascent. The boundary layer type affects the calls made to other parameterisations (e.g. entrainment and convection) as well as the calculation of turbulent viscosity coefficients for boundary-layer mixing.

2.2 Control run

The control 4 km grid-spacing run performed here captures all of the main features of the IOP, but the location and timing of specific features may differ slightly from the observations. In our control run the squall line originated from a cluster of showers that first formed around 0630 UTC over the Bristol Channel and then moved inland, intensifying at the right time and location. By 1030 UTC our simulation had a line of showers that did not extend far enough to the south (as in Clark and Lean, 2006), and which propagated more quickly than observed. More generally, in comparison with radar observations of rain rates, convective features that encompassed at least a few grid points were broadly consistent with the observations, both spatially and temporally (e.g. Figure 5). As the typical horizontal extent of the storms diminished in the later part of the afternoon the MetUM tended to underestimate both the size and intensity of particular features, but the total precipitation rates remained very realistic (not shown).

The evolution of CAPE, rainfall and boundary-layer types during the day are now presented; these are also used as diagnostics for the perturbation experiments. In this paper values quoted for CAPE are obtained from the integrated parcel buoyancy between the first model level (20 m above ground) and the (first) level of neutral buoyancy (LNB). Thus, they include the area of low-lying negative parcel buoyancy: i.e., the Convective INhibition (CIN). The domain-averaged CAPE increases through the simulations to a peak of around 270 Jkg^{-1} between 1100 and 1300 UTC. While this average value is moderate, maximum values can reach 1400 Jkg^{-1} in small areas close to the western coasts of the British Isles. Values of CIN are usually small, with a maximum domain-average during the day of 12 Jkg^{-1}, although occasionally over small areas values of 400 Jkg^{-1} may be reached.

The domain-averaged hourly accumulation for “rainy” grid-points in the control simulation is shown in Figure 1, together with the number of such points. Rainy points are here defined to be those with an hourly accumulation of at least 1 mm. The averaged accumulation peaks at 0700 UTC whereas the number of rainy grid-points peaks later at 1300 UTC. This indicates a transition from intense and localised precipitation to weaker but more-widely distributed precipitation. During the late afternoon and early evening the decrease both in the number of rainy points and in the averaged accumulation is associated with a reduction in CAPE throughout the domain. It is important to note that the convective parameterisation is responsible for less than 1% of the precipitation accumulated throughout the simulation. This suggests that a similar study at higher resolution will not have qualitatively different sensitivities. It also suggests that the results shown in this study may be relatively model independent, provided that the convection scheme is appropriately tuned to the resolution.

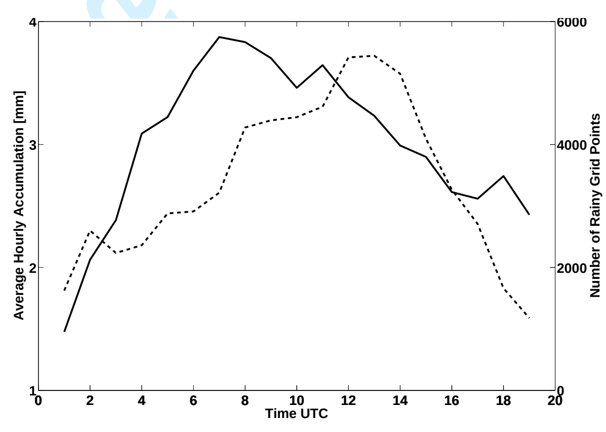


Figure 1. Number of grid points with an hourly accumulation of at least 1 mm (dashed line, right-hand axis) and the domain-averaged hourly precipitation accumulation (solid line, left-hand axis, between the time shown and the following hour) from such points. The total number of grid points analysed is 73780.

Figure 2 shows the evolution of the boundary layer

Prepared using qjrms3.cls
DOI: 10.1002/qj

types within the control simulation. The principal boundary layer types are cumulus-capped, stable, and well-mixed. Their evolution is characterised by two transition periods. The first transition occurs between 0600 and 0730 UTC in response to the increasing short-wave radiation after sunrise (which occured at 0500 UTC). During this first transition there is also a marked increase in the number of rainy grid points (Figure 1). It is marked by a strong decrease in the number of grid points with stable boundary layers and the development of more points with cumulus-capped states. The second transition reverses the changes seen in the first, and occurs between 1700 and 1830 UTC in the response to the diminishing short-wave radiation before sunset (which occurred at 1900 UTC). The number of grid-points classified as shear driven peaks during both transitions, so that this type is manifest as an intermediate state. In between the transitional periods, the percentage of cumulus-capped points exhibits a broad peak between 0900 and 1200 UTC, followed by a steady decrease thereafter. This is compensated for by slow increases in the proportions of the other boundary layer types. Of particular note is the increase in the proportion of stable boundary layers, which is attributable to the formation of cold pools.

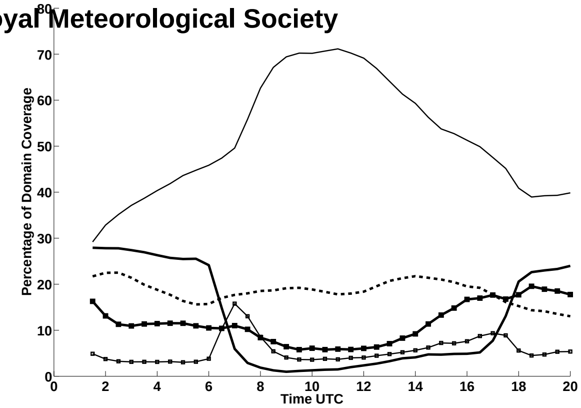


Figure 2. The percentage of the domain covered by the various boundary layer types: cumulus-capped (thin continuous line), stable (thick continuous line), well-mixed (thick dashed line), shear-driven (thin line with squares) and the sum of the three layer-cloud types (stratocumulus-over-stable, decoupled stratocumulus and decoupled stratocumulus-over-cumulus; thick line with squares).

3 Perturbation strategy

3.1 Overview

Uncertainty in model evolution can arise from numerous sources. Analysis uncertainty is inevitable and can, in principle in a variational system, be characterised in terms of the background and observation analysis covariance structures. In practice, however, errors in a given event may deviate from these statistical expectations. Analysis temperature uncertainties of order 1 K are common; this is probably dominated by observation representativity error in the boundary layer rather than observation errors *per se* (which are very often of the order of a few tenths of a degree). The former can be regarded as a measure of the variability on scales significantly smaller than those affected by the analysis system, typically 80 km.

A number of uncertainties can be classified as ‘model error’; turbulence parameterisations (e.g. boundary-layer or convective cloud) are designed to predict the equilibrium response of the parameterised process to a given model state; parameterisations are not perfect and the error is difficult to quantify. A related error, however, may be classed as ‘sampling error’. Even a system in equilibrium has high frequency variability; if we choose to study the system with averaging time less than that required to average this out, then we will see such variability. Furthermore, if we are forced to do this because other processes make the system state vary more rapidly than this time then parameterisation is not strictly valid but it may be reasonable to assume that the error is similar to the related sampling error in an equilibrium system.

If the parametrized system can be characterised approximately as the random superposition of a number

DOI: 10.1002/qj

of independent events, or coherent structures, with a characteristic time scale τ, and our 'sampling time' (i.e. the time over which the mean state varies significantly compared with the parametrized response) is T, then we can expect the relative error (i.e. the standard deviation divided by the mean) in the parameterisation to be of order $\sqrt{\tau/T}$ based on binomial or Poisson statistics. This is, of course, a crude estimate, but gives a realistic idea that, in practice, a lot of 'events' must be averaged out to yield a unique mean. Where convective cloud parameterisation is concerned, this is often expressed as the need to average 'many clouds'; for convection triggering the convective boundary layer is likely to be the most important parameterised component of the system, and 'events' may be thought of as buoyant thermals. Similarly, the spatial structure arising depends on the spatial structure of the 'events' but, if a number of independent events have contributed, this will tend to a Gaussian spatial structure.

Our motivation is thus working towards a stochastic parameterisation representing sampling error in the boundary-layer parameterisation; at this simple level, the stochastic forcing can be represented by its amplitude, a related timescale and a spatial scale. In a more complete formulation, at least of the convective boundary layer, one might envisage representing the characteristic eddy timescale in terms of the boundary-layer depth and the free convective velocity scale, itself a function of the surface buoyancy flux and boundary-layer depth (e.g. Garratt, 1992). Here we shall not attempt to do so, but instead note that it is reasonable to suppose that $\sqrt{\tau/T}$ is less than one and probably of order 0.1; if the eddy timescale is 5 minutes, then T would have to be over 8 hours for the relative error to be as small as 0.1. Given that boundary-layer heating rates of order $1\,\mathrm{Kh}^{-1}$ are a common occurence over land, the development of boundary-layer 'noise' with an amplitude of a few tenths of a Kelvin is certainly plausible. It is also reasonable to suppose that the horizontal covariance scale is related to (and larger than) the boundary-layer depth, i.e. a few km. Given these rough estimates, we have chosen to test the response of the model to precise choices of amplitude and spatial scale (Section 3.2).

Boundary-layer moisture structures with horizontal scales of order 10-20 km can be significant for convection, altering the characteristics of triggered storms, but smaller-scale temperature perturbations appear to be more important in determining triggering (e.g. Crook, 1996; Fabry, 2006). Implementation of a stochastic backscatter scheme in a cloud-permitting simulation has been shown to increase the temperature variance at the inversion level to be more consistent with those in higher resolution cloud-resolving simulations (Weinbrecht and Mason, 2008). This suggests that, as a first approximation, it is reasonable to implement stochastic forcing to the potential temperature near the top of the boundary-layer. The impact of moisture fluctuations may be the subject of a later study.

The temporal correlation of perturbations is an issue; use of perturbations fixed for a period of time raises the question of the need to advect them with the flow. Randomly evolving perturbations, for example through auto-regressive functions (e.g. Berner *et al.*, 2009) or cellular automata (e.g. Shutts, 2005), is a possible refinement; we have chosen a very simple approach of instantaneous perturbations applied either repeatedly (uncorrelated in time) with a constant frequency, *sequential perturbation experiments*, or at a specified time, *single perturbation experiments*.

Prepared using qjrms3.cls
DOI: 10.1002/qj

3.2 Perturbation structure

The random perturbation fields are constructed by the convolution of a random number field with a Gaussian kernel and applied at a specific model level (model level 8), which is at an average of 1280 m above ground (sensitivity tests to the height chosen are described below). At each horizontal grid point a random number is selected for the amplitude of a Gaussian distribution with a standard deviation $\sigma_{\rm gauss}$ (fixed for each experiment) that provides a horizontal lengthscale. The random numbers are generated with a specified seed and uniformly distributed between plus and minus unity so that the full two-dimensional perturbation field averages to zero. Gaussians are centred at each grid point within the domain and so, for $\sigma_{\rm gauss} > 0$ km, the complete perturbation at a particular point is comprised of the sum over all of the individual Gaussian distributions. This summed two-dimensional perturbation field is itself Gaussian distributed and at individual grid points may have an absolute value larger than unity. This field is then scaled to the desired amplitude before application to the model state. The scaling is based on the standard deviation $\sigma_{\rm pert}$ of the fully-summed perturbation field, and we shall henceforth refer to $A = 3\sigma_{\rm pert}$ as the chosen perturbation amplitude.

Three perturbation amplitudes are considered, specifically $A = 1$, 0.1 and 0.01 K. The largest value was chosen because it is representative of the parameterisation sampling error and is of the same order of magnitude as common analysis increments, as discussed in Section 3.1. Furthermore, we aim to test nonlinearities and to affect storm development directly by significantly altering the buoyancy of the underlying air. The tests by Kong *et al.* (2007) of methodologies for perturbing initial conditions also used increments in excess of 1 K. The smallest amplitude considered (0.01 K) provides an indicative bound on the potential predictability of a single convective-scale forecast; a practical system that can produce a model state to this level of accuracy is almost impossible to envisage. The 0.1 K perturbation amplitude is the most credible choice, being consistent with good surface temperature measurement errors (e.g Fabry, 2006) and typical turbulent fluctuations in the convective boundary layer (e.g. Stull, 1988, p358). Such perturbations are intended to be sufficient to change the location and timing of the triggering of moist convection.

Three values of $\sigma_{\rm gauss}$ are considered: 24, 8 and 0 km. These values correspond to typical lengthscales in the full perturbation field of $\sim 6\sigma_{\rm gauss}$, as shown below. A standard deviation of 24 km was chosen to provide a perturbation lengthscale that is well resolved at the model grid spacing and larger than the typical smallest horizontal scale (80 km) in the 3DVAR system. The choice of 8 km provides an intermediate scale between the well-resolved and the grid-scale. The limiting case of a Gaussian for which the standard deviation tends to zero gives rise to spatially-uncorrelated grid-scale noise.

Figure 3 illustrates the effects of the perturbation on the power spectrum for potential temperature at 1000 UTC, but the conclusions drawn here hold equivalently at any time of the day. Figure 3(a) shows the normalised power spectra of the perturbation fields at this time. The power tends to a constant at around a wavelength of $6\sigma_{\rm gauss}$, or 144 and 48 km for the spectra for $\sigma_{\rm gauss} = 24$ and 8 km respectively. The ratio of spectra after and before the application of a perturbation to the model state also shows a peak in the added variance around these wavelengths (Figures 3(b) and 3(c)). For a perturbation amplitude of 1 K, the spectrum is significantly altered between wavelengths of $\sim \sigma_{\rm gauss}$ to $10\sigma_{\rm gauss}$. Small changes are discernible using the 0.1 K

Prepared using qjrms3.cls
DOI: 10.1002/qj

amplitude but are almost indistinguishable for 0.01 K. For spatially-uncorrelated perturbations ($\sigma_{\rm gauss} = 0$) the normalised power spectra for the perturbation fields is, as expected, approximately that for white noise (Figure 3(a)) and the relative magnitude of the added variance is much larger than for the correlated perturbation fields since the spectrum of the perturbed fields has much more power at near-gridscale wavelengths (Figure 3(d)).

The sensitivity of the results to the height of application of the perturbations was investigated by considering four different choices of perturbation height, all within the boundary layer. The perturbation used had an amplitude of 1 K and a standard deviation of 24 km. The root-mean-square error for the hourly accumulation of total precipitation exhibited very little sensitivity to the height of the perturbations, consistent with Lean (2006). No tests were performed applying perturbations above the boundary layer, since the aim is to perturb the triggering process (Section 3.2) and because Lean (2006) demonstrated that perturbations to potential temperature applied at 4500 m do not lead to significant perturbation growth in idealised convection-permitting simulations.

3.3 Frequency of perturbation

Over a model domain that encompasses the entire UK the phenomena that lead to the onset of convection and affect its development occur at different times of the day. Thus, perturbing the model state at a specified frequency throughout the simulation (sequential-perturbation experiments) is a simple and effective way to ensure that a perturbation has been applied prior to all potentially sensitive times. Perturbations applied at successive times during a simulation have no temporal correlation. We note that Grabowski *et al.* (2006) performed some analogous experiments in which the temperature and moisture in the lowest kilometre were randomly perturbed every 15 min, But their motivation was to trigger instabilities in order to inter-compare different cloud resolving models. Some single-perturbation experiments have also been performed to determine how the sensitivity of the simulation to perturbations changes during the day.

In the sequential-perturbation simulations the first perturbation is applied one perturbation period (i.e. 30 min in the standard set up) after the start of the simulation. This is done to allow some time for the simulation to adjust to a more balanced state from the interpolated lower-resolution initial conditions. Thus, such adjustment is considered to be a separate issue from the ongoing uncertainties that exist in the model state (see also Section 6.3).

Our default choice of the perturbation frequency is $2\mathrm{hr}^{-1}$ and represents a compromise between two considerations. On the one hand, a typical equilibrium timescale for a well-mixed boundary layer is of the order of 10 to 20 minutes (e.g. Nieuwstadt and Brost, 1986; Stull, 1988, p450), and the boundary layer would not be able to fully adjust to each perturbation if perturbations were applied too frequently. On the other hand, infrequent perturbations could result in the absence of strong perturbation growth during key transitions in the boundary layer structure: for example, from stable to cumulus-capped after sunrise (Figure 2). Application frequencies intermediate between these two limits are hypothesised to be likely to lead to similar levels of model spread.

The time evolution of the potential temperature field on the perturbed model level is now discussed with reference to Figure 4. A specific four hour period is plotted for clarity but the behaviour described is similar throughout the day. Shown are root-mean-square differences between the potential temperature at a current timestep and that at

DOI: 10.1002/qj

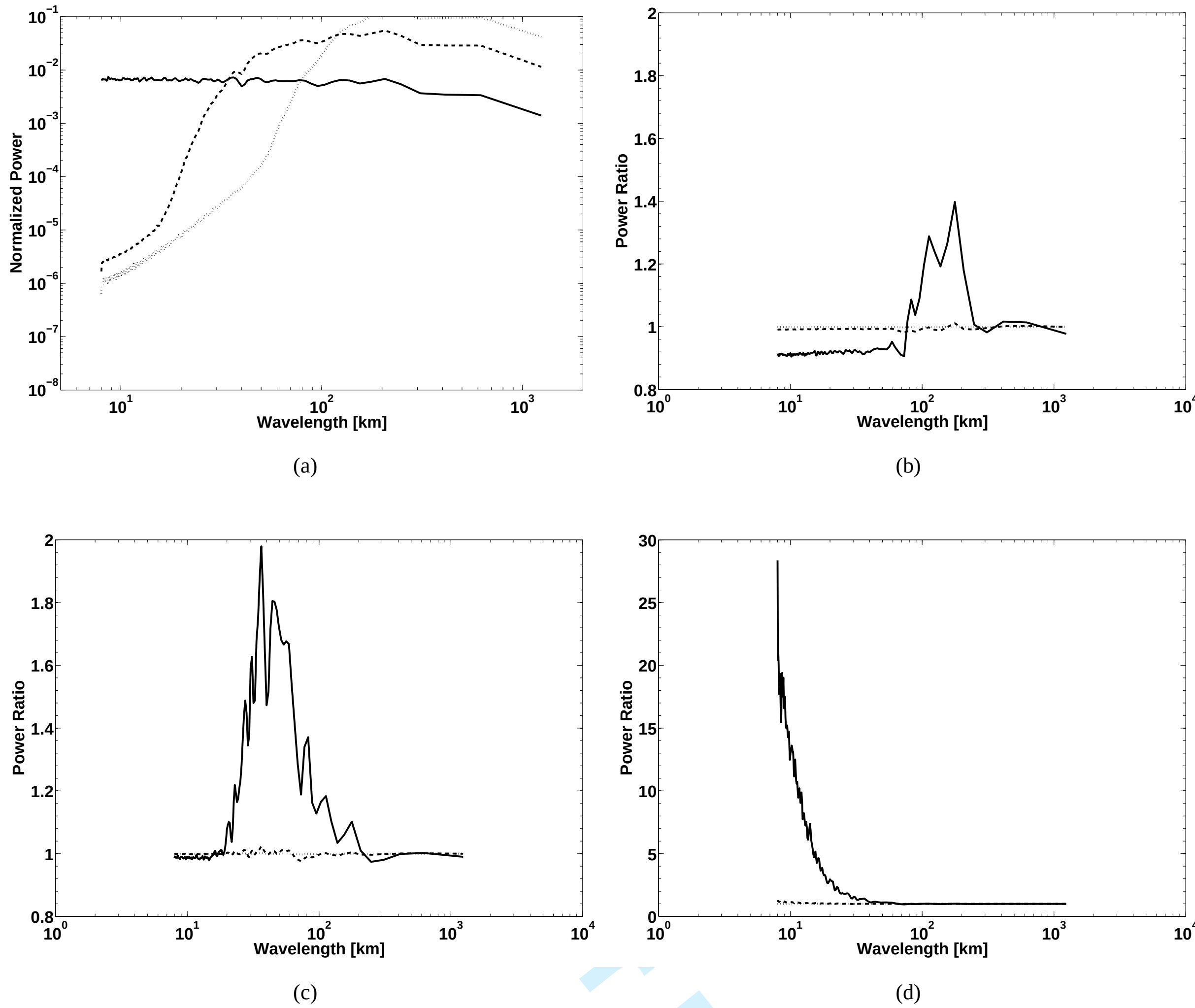


Figure 3. (a): Normalised power spectra for three example 1 K amplitude potential-temperature perturbation fields, with $\sigma_{\mathrm{gauss}} = 24$ km (dotted line), 8 km (dashed line) and 0 km (solid line). The spectra are normalised such that an integration over wavenumber produces unity. Note the log scales on both axes. (b), (c) and (d) show, for different values of σ_{gauss}, ratios of potential-temperature power spectra on model level 8. The ratio is the spectrum of the potential temperature immediately following the application of a perturbation to the corresponding spectrum just prior to the perturbation. It was computed at 1000 UTC and $\sigma_{\mathrm{gauss}} = 24$ km in (b), 8 km in (c) and 0 km in (d). Note the different y-axis scale for (d). In each case, perturbation amplitudes of 1 K, 0.1 K and 0.01 K are indicated by solid, dashed and dotted lines respectively. (Note that the lines for the smallest two perturbations are generally indistinguishable).

an earlier reference time for the control (unperturbed) simulation and two perturbed simulations. The two perturbed simulations were both perturbed with an amplitude of 1 K and a standard deviation σ_{gauss} of 24 km but one was perturbed every hour and the other every 30 min. For the control simulation the reference time is the start of the preceding hour; for the perturbed simulations the reference time is reset after two perturbation periods (i.e. every two hours for the simulation perturbed every hour and every hour for that perturbed every 30 min.). The model state at the reference time is taken after the introduction of perturbations.

The minima in root-mean-square differences reflect the evolution of the simulations over one timestep (i.e. they are calculated from the difference between the model state one timestep after the reference time and that at the reference time). Until the reference state is reset, the root-mean-square differences for the perturbed simulations increase for two reasons: first, perturbation growth

Prepared using qjrms3.cls
DOI: 10.1002/qj

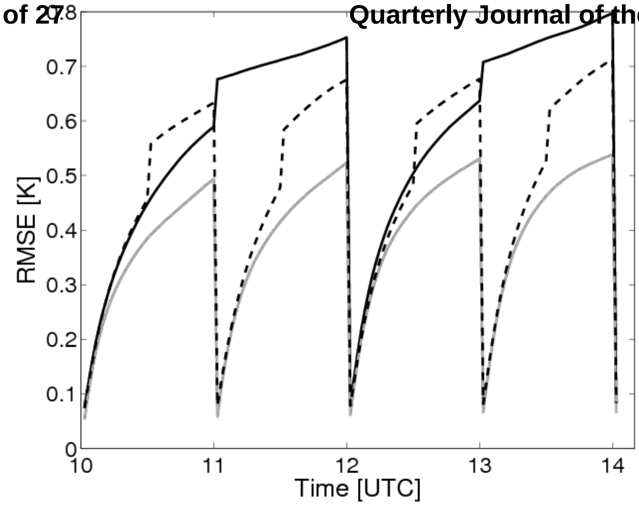


Figure 4. Root-mean-square difference between the potential temperature at the given time and that at a reference time earlier in the same simulation. The reference time is periodically reset as described in the main text. Calculations are performed on model level 8 at each timestep. The solid grey line is for the control simulation. The black lines are for perturbed simulations, with a perturbation amplitude of 1 K, σ_{gauss} of 24 km and application frequencies of $2\mathrm{hr}^{-1}$ (black dashed line) and $1\mathrm{hr}^{-1}$ (black solid line).

and second, the general evolution of the model. The root-mean-square difference for the control simulation gives an indication of the latter and hence the differences between the root-mean-square differences for the perturbed and control simulations give an indication of perturbation growth. The perturbation growth is clearly evident but of smaller magnitude than the model evolution suggesting that the changes to the model state induced by the perturbations are not unrealistic (i.e. they do not alter the main features of the event).

The root-mean-square difference grows throughout the first perturbation and then jumps when a new perturbation is applied. This is followed by further growth until the reference time is reset. The size of the jumps relative to the growth indicates that the growth is driven more by the model's response to a perturbation than by the direct perturbation application itself (the relative size of the jumps also decreases for smaller amplitude perturbations (not shown)). During each hour shown the overall growth of root-mean-square differences within the two perturbation simulations are very similar (i.e. this growth is independent of the frequency of perturbations). This suggests that the model evolution is relatively insensitive to whether the perturbations are applied every hour or 30 min.

We have also examined the timestep-to-timestep evolution of the potential-temperature power spectrum between perturbation applications (not shown). Using a 1 K perturbation amplitude, we find that the signature of a perturbation at the perturbed scales decreases rapidly but that after 30 min. it remains perceptible. Hence, the perturbations do not dissipate entirely between applications.

3.4 Perturbation experiments

The experiments performed are summarised in Table I. Nine sequential-perturbation experiments were performed, with varying perturbation amplitudes, A, and standard deviations, σ_{gauss}. These experiments are labelled $\sigma r A s$, as a shorthand for $\sigma_{\mathrm{gauss}} = r$ km, $A = s$ K. A trailing asterisk indicates that, for those experiments, a set of six simulations were performed differing only in the set of random numbers generated. We will refer to such sets of simulations as *ensembles*. They allow us to compare the spread produced by varying the perturbation-field parameters to that produced by different realizations of the same perturbation process. Other than the ensemble simulations, all runs were performed with the same random-number sequence. The six-member ensembles were generated by using two additional seeds for the random number generator (to generate three members) and then reversing the signs of the perturbations generated by the resultant three random sequences (to generate the other three members) (as in Done *et al.*, 2008, for instance).

Eight single-perturbation experiments have also been performed. These allow us to investigate the importance of the sequential perturbation strategy, and to distinguish

DOI: 10.1002/qj

Run label	$\sigma_{\rm gauss}$ [km]	Amplitude A [K]	Application timing	Bias [$\times 10^{-2}$mm]	Max 3h RMSP increment [mm]
Control	–	–	never	0.0	–
σ24A1*	24	1	2 hr^{-1}	-6.37	2.07
σ8A1	8	1	2 hr^{-1}	-6.47	2.55
σ0A1	0	1	2 hr^{-1}	-3.92	2.13
σ24A0.1*	24	0.1	2 hr^{-1}	-1.07	1.41
σ8A0.1	8	0.1	2 hr^{-1}	0.67	1.12
σ0A0.1	0	0.1	2 hr^{-1}	1.13	1.16
σ24A0.01*	24	0.01	2 hr^{-1}	0.74	0.94
σ8A0.01*	8	0.01	2 hr^{-1}	-0.63	1.00
σ0A0.01*	0	0.01	2 hr^{-1}	-1.00	1.12
IC-1	24	1	IC	7.43	2.62
0700-1	24	1	0700	-0.08	1.52
0830-1	24	1	0830	0.28	1.47
1000-1	24	1	1000	-1.35	1.13
IC-0.01	24	0.01	IC	-0.49	0.84
0700-0.01	24	0.01	0700	1.11	0.95
0830-0.01	24	0.01	0830	-1.18	0.74
1000-0.01	24	0.01	1000	-0.09	0.74

Table I. List of simulations performed and their characteristics. The simulation labelling is explained in the main text. Characteristics shown are the standard deviation $\sigma_{\rm gauss}$, the perturbation amplitude A, the application timing, the bias in the domain-averaged precipitation accumulated during the perturbed simulations (the difference from the control simulation value of 2.127 mm), and the maximum three-hourly increment of RMSP (as defined in Section 4). For the starred simulations the values reported refer to a single member of the ensemble (that generated using the same random number sequence as used in the experiments for which an ensemble was not performed).

between direct and indirect effects of a perturbation (Section 4). For these experiments the standard deviation $\sigma_{\rm gauss}$ was fixed at 24 km and two perturbation amplitudes ($A = 0.01$ and 1 K) were considered. Four application times were tested: specifically in the initial conditions (0100 UTC) and at 0700, 0830 and 1000 UTC. These experiments are labelled in the form t-s. Here t indicates the application time in UTC (or else as IC for initial condition perturbations) and s is the perturbation amplitude A in K.

As an example of the impact of these perturbations, Figure 5 shows the precipitating cloud fields at 1000 UTC for the control run and for two sequential-perturbation experiments. Radar rain rates are also included for comparison. This snapshot shows that while the locations of individual clouds have changed, on the regional scale the cloud distribution remains realistic.

4 Diagnostics

Two types of diagnostic are described here. Diagnostics of the direct impact of the perturbations reveal the instantaneous response of the model and may be somewhat model specific (Section 5). Diagnostics of the indirect impact of the perturbations reveal overall the perturbation growth due both to the model evolution and to the sequential perturbations (Section 6). For both types of diagnostics the reference is the control run, rather than observations. Hence the term bias here refers to the difference of a perturbed simulation from the control simulation.

4.1 Direct effects

Four measures of direct effects (within one time step) are examined: the effects on convective instability, the boundary-layer types, total cloud water and cloud distributions, and the model adjustment to perturbations. All

Prepared using qjrms3.cls

DOI: 10.1002/qj

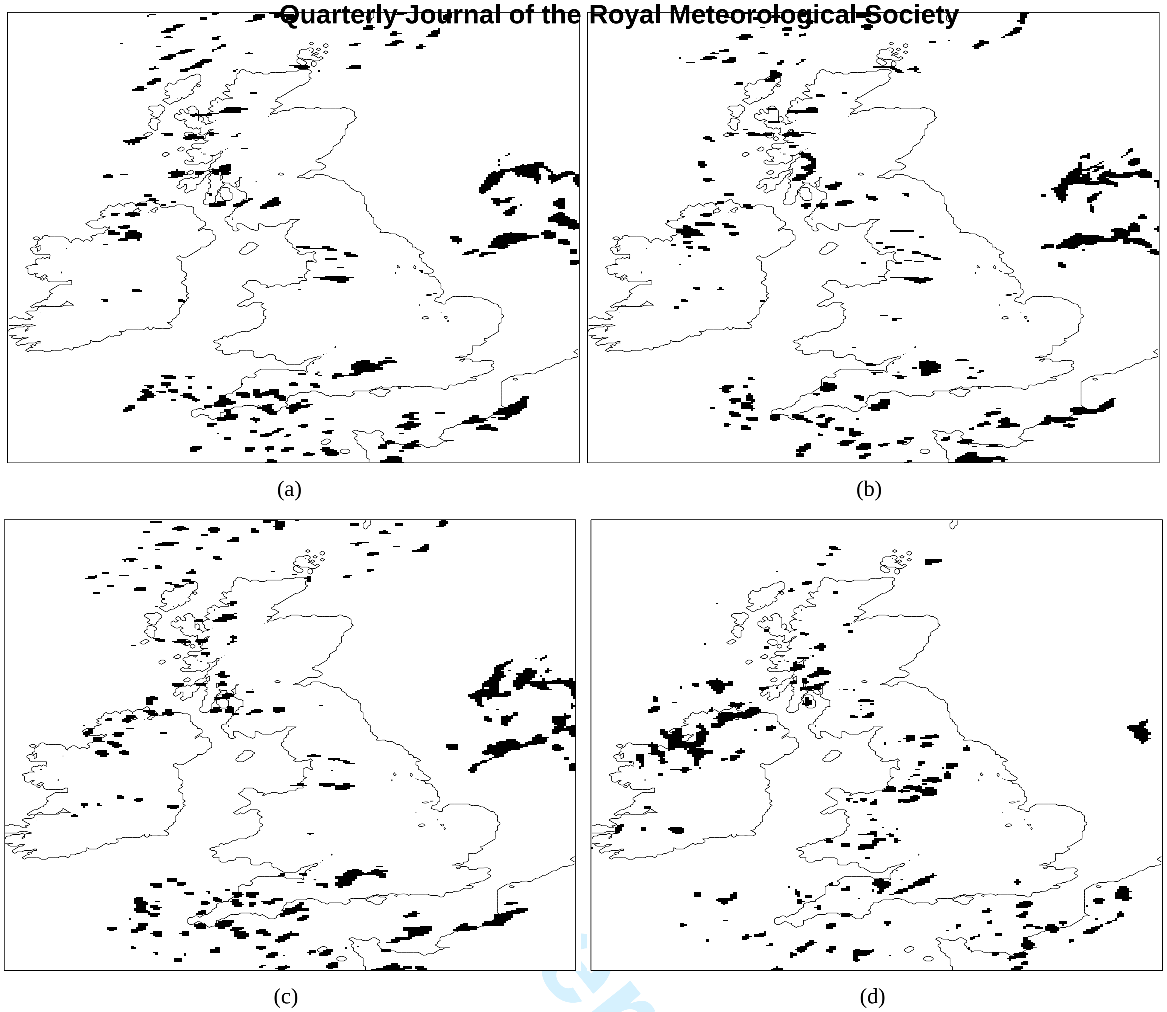


Figure 5. Precipitating clouds, defined as neighbouring grid points with a rain rate of at least 1 mmh^{-1} (see Section 4) for the (a) control, (b) σ24A1 and (c) σ0A1 simulations. (d) shows the precipitating clouds obtained from the observed radar rain rates. The radar data originally on the 5 km national grid have been interpolated to the UM grid used here. All panels are for 1000 UTC.

of the direct effects are assessed by comparing single-perturbation simulations with the control simulation, one timestep after the perturbation has been applied.

Convective instability is diagnosed using CAPE and CIN, defined as in Section 2.3. The adjustment mechanisms are assessed by comparing profiles of domain-averaged pressure, vertical velocity and total cloud water content conditioned on the sign of the imposed perturbation.

The boundary-layer type determined by the MetUM is affected both directly and indirectly by the applied perturbations: directly because the determination is based upon parcel ascent and indirectly because the perturbation can induce changes which later cause the boundary-layer type to switch, as described by Lean (2006). The direct aspect can be studied from the fraction of the domain that changes its boundary-layer type with respect to the control run on application of the perturbation.

Cloud distributions are affected directly because a change of potential temperature produces a change in relative humidity, and also indirectly because the evolution of existing clouds and future cloud triggering can be affected. We consider only the horizontal distribution of cloud i.e. each grid column is defined as either cloudy or not cloudy. Two cloud definitions are considered here, *precipitating cloud* and *non-cirrus cloud*. Precipitating

DOI: 10.1002/qj

clouds are constructed from grid points with rain rates of at least 1 mmh^{-1}. Tests have shown that the results presented in Sections 5.4 and 6.2 are not sensitive to modest variations of this threshold. Non-cirrus cloud is a more generic definition of cloudy air and is based on grid points with a vertically-integrated total-water path of at least 0.05 kgm^{-2}. The integration extends from the surface up to the first model level above the maximum height of the squall line, at 8 km. Thus, it excludes high-level layer cloud which is unlikely to be affected by the perturbations. Other definitions of cloud have also been tested with similar results.

For both definitions of cloudy grid points, the "clouds" themselves are constructed as connected clusters of such grid points. Connections are considered to occur if the associated grid boxes share either an edge or a corner. The area of each cloud is counted in grid boxes, and the cloud distribution statistics are determined from snapshots of the cloud field every 30 minutes during the simulation.

4.2 Indirect effects

Three measures of indirect effects are examined: the effects on boundary-layer types, cloud distributions and, the root-mean-square error of the hourly-accumulated precipitation (RMSP). The first two of these were described above.

The third measure, the RMSP, is a simple and widely-used error norm (e.g. Molteni *et al.*, 2001; Snyder and Zhang, 2003), here computed relative to the control simulation. It is convenient to adopt a slightly-different definition here, according to which its square is given by

$$\mathrm{RMSP}^2 = \frac{1}{N}\sum_{i=1}^{N}(p_i - c_i)^2 \qquad (1)$$

where the summation extends over those N grid points that are classified as "rainy" (Section 2.3) in either the perturbed or the control run. This restriction requires that either p_i (the hourly-accumulated precipitation in the perturbed simulation) or c_i (the same for the control simulation) is at least 1 mm. This allows a useful decomposition of the RMSP to be introduced below. Tests indicate that the conclusions are robust to changes in the threshold and the precipitation diagnostic (using instantaneous rain rates sampled every 30 min instead did not affect the conclusions). Such insensitivity may be due to the scattered nature of much of the convection in this case; at any time within the simulations there are multiple storms at different stages of their life cycles. Results also do not change when the RMSP is computed using averages over square boxes of a few grid points in width.

From both meteorological and hydrological perspectives it is important to know to what extent the perturbations tend to displace storms, alter their intensity and create or suppress new ones. A complete analysis of this issue could be provided only by keeping track of each storm at each timestep. This is beyond the scope of the current study, although we note that it may become a practical proposition in the future (Plant, 2008). Nonetheless, some insight into such issues can be provided by decomposing the squared RMSP into three components, from three types of points that contribute to the sum on the right-hand-side of Equation 1: specifically, those that are rainy only in the control run ($c_j > 1, p_j < 1$mm), those that are rainy only in the perturbed run ($c_k < 1, p_k > 1$mm), and those that are rainy in both runs ($c_l > 1, p_l > 1$mm). The contributing types will be referred to as CONTROL, PERTURBED and COMMON points respectively. Thus,

$$N = N_{\mathrm{CONTROL}} + N_{\mathrm{PERTURBED}} + N_{\mathrm{COMMON}} \qquad (2)$$

Prepared using qjrms3.cls
DOI: 10.1002/qj

and

$$
\begin{aligned}
\mathrm{RMSP}^2 &= \mathrm{MSP}_{\mathrm{CONTROL}} + \mathrm{MSP}_{\mathrm{PERTURBED}} \\
&\quad + \mathrm{MSP}_{\mathrm{COMMON}} \\
&= \frac{1}{N}\sum_{j=1}^{N_{\mathrm{CONTROL}}} c_j^2 + \frac{1}{N}\sum_{k=1}^{N_{\mathrm{PERTURBED}}} p_k^2 \qquad (3) \\
&\quad + \frac{1}{N}\sum_{l=1}^{N_{\mathrm{COMMON}}} (c_l - p_l)^2
\end{aligned}
$$

where j, k and l label gridpoints in the sets CONTROL, PERTURBED and COMMON respectively. Note that p_j and c_k are set to zero (as these points are not raining according to the above definition).

The decomposition distinguishes between changes in the intensity of storms and changes in their location; however, it does not distinguish between changes in location that are caused by differences in advection and changes that are caused by, say, the generation of new storms. Notice that for fixed values of each N, an increase in $\mathrm{MSP}_{\mathrm{CONTROL}}$ or $\mathrm{MSP}_{\mathrm{PERTURBATION}}$ would represent an increase in precipitation at the associated points, whereas an increase in $\mathrm{MSP}_{\mathrm{COMMON}}$ would represent increased difference between two simulations at the COMMON points.

5 Results: Direct effects

5.1 Perturbation effects on CAPE

On application of a perturbation, changes to the distribution of CAPE are minor. For example, averaged over the domain, the magnitude of the bias for all single-perturbation experiments is less than 0.5 Jkg^{-1}. For most grid points, the perturbation application level lies above the lifting condensation level. Thus, there may be a slight increase (decrease) in CAPE associated with a negative (positive) potential temperature perturbation, but there is no overall bias. The small bias that does occur is attributable to grid points at which the perturbation alters the CAPE by changing the upper limit of the vertical integral. At a small number of grid points (e.g. 0.6% of the domain for the 0700-1 experiment), the parcel ascent is always cooler than the environment, so that there is no LNB identified and the CAPE is considered to be null. For some of these points a negative perturbation introduces a lid that sets the LNB to the perturbation level, which can result in a negative contribution to the available energy of up to -60 Jkg^{-1}. At other points the opposite process may also happen, i.e. a very weak lid is removed, with positive changes to the CAPE. There are very few points where this occurs, but storms may be generated there, should a suitable trigger also exist.

5.2 Model adjustment to the perturbation

The first dynamical response to the random, imposed heating (cooling) consists of acoustic and Lamb waves which within minutes accomplish the required expansion (compression) (Chagnon and Bannon, 2005). Thus, if the vertical velocity is conditionally averaged over those grid points experiencing a positive potential-temperature perturbation (hereafter *positive points*), then a difference with respect to the control simulation is evident (Figure 6). The vertical profile is consistent with a Lamb wave in a non-isothermal atmosphere with a rigid lid top boundary condition (Lindzen and Blake, 1972). The increase in vertical velocity with respect to the control simulation is significant but small in absolute value and, because the wave propagates at the speed of sound, the associated parcel displacements are very small and unlikely to trigger new storms. However, the wave affects the entire three-dimensional domain. Each local maximum in the perturbation field is effectively a source of acoustic waves which take roughly 30 minutes to travel across the domain. These

DOI: 10.1002/qj

acoustic waves change the environment in which convection forms contributing to the divergence of the perturbed runs from the control. The magnitude of the difference between vertical velocities in the control and perturbed runs scales linearly with the perturbation amplitude and changes sign with it. Associated changes in the pressure profile were also detected (not shown).

It should also be pointed out that the effects of acoustic waves may be underestimated in the model, partly because the parameters used in the off-centering of the advection scheme are designed to damp them (Davies *et al.*, 2005), and partly because the relatively long timestep (100 s) will not properly resolve the fast acoustic waves.

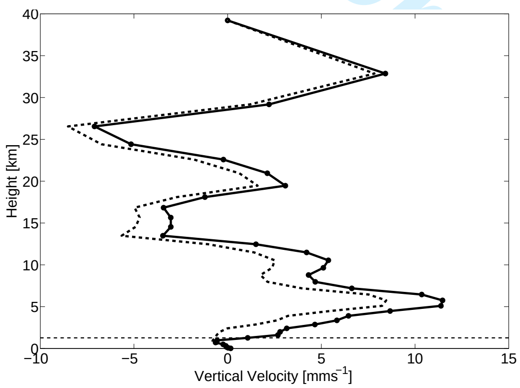


Figure 6. Vertical profiles of vertical velocity, computed one timestep after a perturbation application at 0700 UTC and averaged over those grid points experiencing a positive potential temperature perturbation. A dashed horizontal line marks the perturbation level. The solid line represents the 0700-1 run and the dashed line represents an average from the control run for the same points at the same time.

Using an individual Gaussian perturbation with an amplitude of slightly less than 1 K, Hohenegger and Schar (2007b) found that the acoustic wave response to perturbations may be responsible for error growth. Acoustic waves will also be generated by the convective storms themselves (Nicholls and Pielke, 2000) and can further accelerate error growth if the storms have been displaced (Section 6.4). It is also worth noting that analysis increments can be larger than 1 K (e.g. Kong *et al.*, 2007): such increments are not acoustically balanced and so may excite wave responses stronger than observed in this study.

5.3 Boundary layer changes

The perturbations introduced can directly change the boundary-layer types. The percentages of grid points in the domain that change type in the single perturbation simulations and the percentage coverage of each type in the control simulation at specific times are shown in Table II. The percentage change is defined as the percentage of half the number of grid points in the domain at which there was a change either to or from a given type. The effect is weak for $0.01\,$K amplitude perturbations, the largest change is 0.05%; for the $1\,$K perturbation amplitude the largest change is around 2%. These changes may be substantial for some of the types, e.g. 10% of stratocumulus over stable grid points change type in the 0700-1 simulation ($100 \times (0.31/3.21)$%); this can contribute to perturbation growth (as shown by Lean (2006)). Note that it is unlikely that grid points that change their boundary-layer type will immediately revert back to the original type on the following timestep; even non-growing potential temperature perturbations persist for 30 min. or more (Section 3.3), albeit with decaying amplitude.

A more detailed analysis of the switches in the boundary-layer types shows that there are switches to and from each type, with the exception of the stable boundary layer which only loses points on application of a perturbation, and only loses them to the stratocumulus over-stable-type. Thus the perturbation must generate stratocumulus at such points.

5.4 Total water path and cloud distribution changes

The total water (ice and liquid water) has been compared against the control run one timestep after the perturbation application. Other than at the perturbation level, changes

DOI: 10.1002/qj

Run label	Well mixed	Stable	Sc over stable	Decoupled Sc	Decoupled Sc over Cu	Cu capped	Shear dominated
Control at 0700	17.68	5.96	3.21	6.67	1.12	49.59	15.78
Control at 0830	18.66	1.27	0.26	5.71	1.55	67.13	5.41
Control at 1000	18.87	1.28	0.18	4.62	1.29	70.18	3.59
0700-1	0.92	0.31	0.31	0.98	0.20	1.12	0.81
0830-1	2.01	0.01	0.01	1.33	0.40	2.07	0.61
1000-1	2.21	0.04	0.04	1.81	0.30	2.07	0.47
0700-0.01	0.02	0.01	0.01	0.02	0.00	0.02	0.02
0830-0.01	0.04	0.00	0.00	0.03	0.01	0.03	0.01
1000-0.01	0.05	0.00	0.00	0.03	0.01	0.05	0.01

Table II. The first three rows show the domain cover, expressed as percentage, for each boundary-layer type of the MetUM at specific times within the control simulation. The subsequent rows show the percentage, for various single-perturbation experiments, of grid points in the domain that immediately changed boundary-layer type with respect to the control run on application of the perturbation. In calculating these figures a switch from type T_1 to type T_2 is counted as a switch under both T_1 and T_2. This sum is then divided by two in order to avoid double counting, and so that the sum of the percentages is the total percentage of the domain that changed boundary-layer type.

are negligible, but at that level evaporation (condensation) following positive (negative) potential-temperature perturbations can be clearly seen. For example, in the 0700-1 simulation, conditional averaging over positive points reveals an 19% reduction in total water (specifically, the contribution from that vertical layer to the total water path falls from 7.08$\times10^{-3}$ to 5.79$\times10^{-3}$ kgm^{-2}). There is perhaps a weak sensitivity of this effect to the time of the day, the corresponding reduction being 15% in the 1000-1 simulation. Such effects are slightly less consistent with a linear scaling than the pressure and vertical adjustments described in Section 5.2 due to the saturation process. An increase in potential temperature causes more condensate to evaporate than the equivalent decrease causes it to condense.

The immediate repercussions of the total water modifications on cloud distributions are quite small, both in terms of the cloud number and average size, for both of the cloud definitions. Changes in cloud number are generally 1% or less, and changes in the mean cloud size are even smaller. However, it is worth noting that the direct changes to cloud distributions that do occur have a different character from changes produced indirectly (cf. Section 6.2). For example, in the 0700-1 simulation, the perturbation produces a direct increase in both the cloud mean size and number for non-cirrus clouds, but the indirect effect is of an increased mean size being somewhat offset by a reduction in cloud number. The behaviour for the direct effect is not surprising given that the mean cloud cover is around 22% of the domain (Section 6.2). Thus, on application of the perturbation it is more likely that a grid point will become "cloudy" due to condensation rather than "non-cloudy" due to evaporation.

5.5 Comments on Gravity Waves

It is well known that a potential-temperature perturbation induces gravity waves whose characteristics depend on the static-stability of the background state, on the duration and intensity of the heating, and on the size and aspect ratio of the heated region (e.g. Chagnon and Bannon, 2005). However, examination of potential-temperature fields output at every timestep (from various sequential-perturbation simulations) did not show any coherent buoyancy-wave activity at the perturbation level. This is probably due to various contributing factors. One aspect is the horizontally heterogeneous shear and stability, which generate spatially incoherent responses, but also important is the very

DOI: 10.1002/qj

limited vertical extent and duration of the heating, respectively 320 m and 100 s. These values are small compared to idealised studies with horizontally homogeneous conditions (e.g. Chagnon and Bannon, 2005; Robinson *et al.*, 2008), and the variety of processes occurring is also likely to hide any gravity wave signal. Furthermore, the model timestep is too long to provide a good representation of the slowest gravity wave modes, with period $2\pi/N$, since N is of the order of 0.01 s^{-1}.

6 Results: Indirect effects

6.1 Boundary layer changes

The perturbations can influence the determination of the boundary-layer types throughout a simulation. Table III lists the time-averaged percentage coverage for each of the seven types, along with the changes to those values that occur in the sequential-perturbation simulations. Note that this diagnostic is not the same as that shown in Table II to illustrate the direct effect on the boundary layer.

For perturbation amplitudes of 0.01 and 0.1 K, the changes are small (less than 0.2% for all boundary-layer types). However, for the 1 K perturbation amplitudes larger changes are found (up to nearly 6% for the cumulus-capped type). The main change is a reduced coverage by the cumulus-capped boundary layer, mainly balanced by increases to the coverage of the well-mixed and the decoupled stratocumulus types. More detailed inspection of the changes in domain cover over the course of the simulations reveals that, for any given simulation, the changes that occur are of a similar character throughout the day (e.g., a modest reduction in the stable type is a consistent feature of σ0A1).

When the changes for the single-perturbation experiments are compared against those in the sequential-perturbation simulations with the same metric, the changes are found to be larger in the latter case. This indicates that repeated application of the perturbations is indeed more relevant for model divergence than an individual application, and highlights the importance of considering the uncertainty in the evolving model state.

6.2 Cloud distribution changes

The results for precipitating clouds can be seen in Figure 7. An inverse linear relationship is found between the time-averaged number of clouds in the model domain and their mean size. This is perhaps not surprising given that all simulations produce very similar amounts of total rainfall over the course of the day (Table I). The control simulation has an average of around 111 clouds with a mean size of about 22 grid boxes (covering around 2.7% of the domain). The mean cloud number and size are generally clustered around these values for most of the perturbed simulations, albeit with a tendency for slightly fewer, slightly larger clouds. The exceptions are the sequential 1 K and IC-1 simulations which span a wider range, with the IC-1 simulation being very close to the σ8A1 simulation. For these simulations, perturbations with smaller lengthscales give rise to smaller but more numerous clouds. This seems to indicate that the dynamics of the precipitating clouds is slightly altered by the perturbations.

Similar comments apply for the non-cirrus clouds (not shown), which are of course larger and more numerous covering on average around 22% of the domain. The IC-1 simulation and the sequential-perturbation simulations with 1 K perturbation amplitude again form a distinct subset, with similar sensitivity to the typical scalelength of the perturbation to that seen for the precipitating clouds. However, the relative changes to both cloud size are much larger than for the precipitating clouds, suggesting that non-cirrus cloud dynamics is altered more strongly by

Run label	Well mixed	Stable	Sc over stable	Decoupled Sc	Decoupled Sc over Cu	Cu capped	Shear dominated
Control	18.44	12.51	2.47	7.85	1.24	51.83	5.66
σ24A1	2.02	-0.73	0.80	2.11	0.81	-5.11	0.10
σ8A1	2.49	-0.26	0.33	1.78	0.54	-5.17	0.30
σ0A1	3.91	-0.58	0.52	1.43	0.45	-5.86	0.12
σ24A0.1	-0.04	-0.05	0.02	0.05	0.00	0.09	-0.06
σ8A0.1	0.03	-0.04	-0.02	0.15	0.00	-0.16	0.04
σ0A0.1	-0.06	-0.08	0.04	0.14	-0.01	-0.05	0.01
σ24A0.01	-0.02	0.02	-0.01	0.01	0.03	-0.03	0.00
σ8A0.01	-0.08	-0.02	-0.01	-0.01	-0.01	0.06	0.06
σ0A0.01	-0.09	0.02	-0.03	-0.06	0.00	0.15	0.01

Table III. Percentages of the domain covered by each of the MetUM boundary-layer types averaged over the entire duration of the simulations, using half hourly data. The first row shows the percentages for the control run. Subsequent rows show the differences between the first row and the corresponding value for the named sequential-perturbation simulation. Positive values for the changes indicate increased cover for that boundary-layer type in the perturbed simulation.

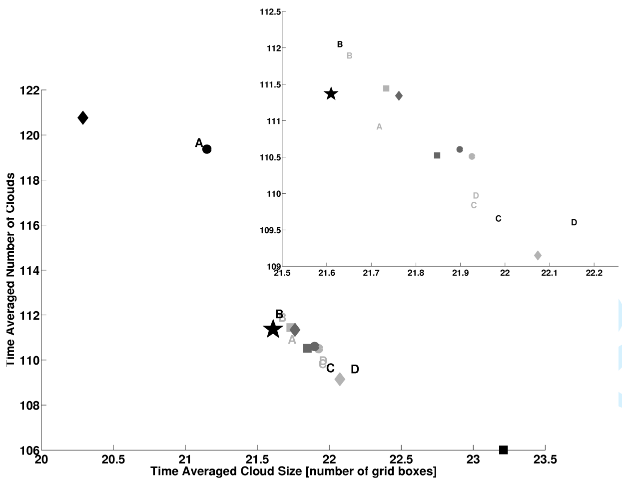


Figure 7. Mean precipitating cloud size and cloud number, averaged over the entire duration of the simulations using half hourly data and plotted as a point for each simulation listed in Table I. The control simulation is denoted by a black star. Simulations with perturbation amplitudes of 0.01, 0.1 and 1K are denoted by light grey, dark grey and black symbols respectively. For sequential-perturbation simulations, the symbols used are squares, circles and diamonds for $\sigma_{\mathrm{gauss}} = 24, 8$ and 0 km respectively. Where an ensemble exists its mean value is plotted. For single-perturbation simulations letters are used: 'A', 'B', 'C' and 'D' for perturbation application times of the initial time, 0700, 0830 and 1000 UTC respectively. The upper right plot is an expansion of the central area of the main plot.

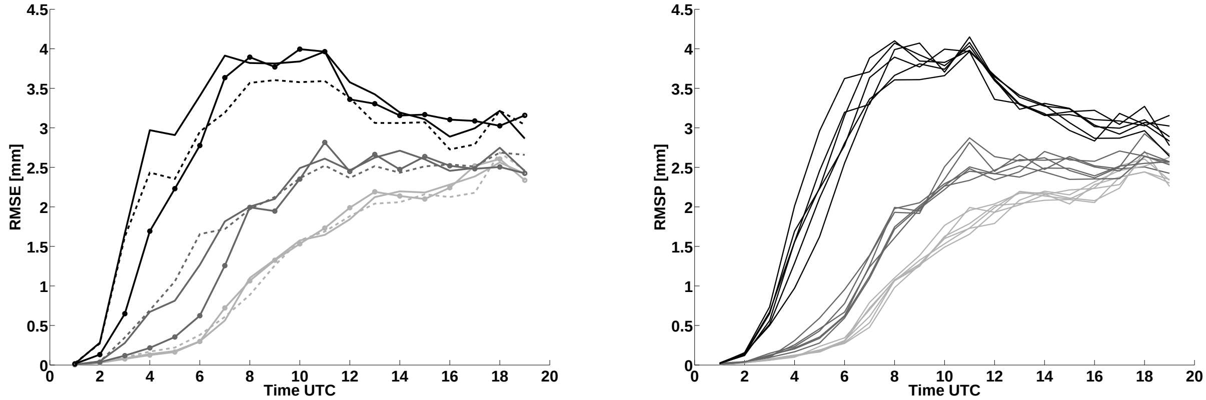
these perturbations. Regression analysis of the data from this subset and from all the other simulations separately both produce straight-line fits with high correlations and different slopes.

6.3 RMSP

Figure 8 shows the evolution of RMSP in the sequential-perturbation simulations. It is most responsive to the perturbation amplitude. When this is 1 K, the strongest RMSP growth occurs after the second perturbation application (at 0200 UTC) and the RMSP peaks between 0700 and 1100 UTC before levelling off at about 3 mm. In contrast, although the strongest growth is again seen at early times for perturbations of amplitude 0.1 K, there is no peak in the RMSP evolution, which levels-off at about 2.5 mm beyond about 1200 UTC. With a 0.01 K perturbation amplitude, the strongest growth is delayed to around 0600 UTC. A clear saturation phase is not seen for these simulations, although similar RMSP values to those achieved with the 0.1 K perturbations are reached at the end of the simulation time. If the RMSP is computed using average values over small square areas (up to 11 grid points in width), its absolute value is reduced, but the relative behaviour of the different simulations is essentially unmodified.

These results for the onset of strong RMSP growth show that smallest perturbations have little effect on the precipitation that occurs before sunrise (Figure 1). Only once surface heating begins and the boundary layer starts

Prepared using qjrms3.cls
DOI: 10.1002/qj



(a) (b)

Figure 8. Evolution of the RMSP for: (a) nine sequential-perturbation simulations with different perturbation amplitudes (0.01 K light grey, 0.1 K in dark grey and 1 K in black) and standard deviations $\sigma_{\rm gauss}$ (24 km as solid lines with filled circles, 8 km as solid lines and 0 km as dashed lines); and, (b) three ensembles with different perturbation amplitudes (0.01 K in light grey, 0.1 K in dark grey and 1 K in grey), each with $\sigma_{\rm gauss} = 24$ km. All times along the horizontal axis refer to the beginning of the hour of accumulation.

to change its structure (Figure 2) are such perturbations capable of stimulating strong RMSP growth. By contrast, the 1 K perturbations are powerful enough to produce strong growth almost from the outset. We note that an equivalent time lag between perturbation application and strong RMSP growth is not seen in those single perturbation runs with the perturbations applied after sunrise (Figure 9). Hence the delay in growth until after sunrise is not merely a result of weak perturbations requiring time to grow to a significant amplitude prior to perturbing precipitation.

Figure 8(b) shows the RMSP evolution for three ensembles with different perturbation amplitudes. Note that the spread of the ensembles increases with increasing perturbation amplitude (particularly during the strong growth phase). Systematic dependence on the perturbation standard deviation, $\sigma_{\rm gauss}$, is not obvious from Figure 8. In particular, for a given perturbation amplitude the spread for different $\sigma_{\rm gauss}$ is comparable to the spread within the six-member ensembles. Thus, the horizontal scale length of the perturbation does not strongly affect the RMSP. However, for the 1 K and 0.1 K amplitudes, there is some delay in the onset of the strongest growth for $\sigma_{\rm gauss} = 24$ km.

The above conclusions on RMSP growth are also confirmed by the final column of Table I. This lists the strongest perturbation growth in each simulation, as measured by the maximum three-hourly increment of RMSP. The RMSP increments are somewhat noisy and so it is convenient to apply a 1-2-1 filter to the increments prior to determining this maximum, but the relationship between RMSP growth and perturbation amplitude is readily apparent regardless of any filtering. The maximum growth is around twice as strong for the 1 K compared with the 0.01 K perturbation amplitude. Moreover, the variations in maximum growth within each of the ensembles in Figure 8(b) are found to be smaller than the differences between the ensemble-mean values. Thus, at least for $\sigma_{\rm gauss} = 24$ km (the only $\sigma_{\rm gauss}$ for which ensembles were performed), the amplitude of the perturbations affects the RMSP growth more strongly than the choice of random number sequence.

Prepared using qjrms3.cls
DOI: 10.1002/qj

The domain-averaged hourly accumulated precipitation for rainy points over the full duration of the control simulation is 2.127 mm. The biases from the control simulation for the perturbed simulations are listed in Table I. For most cases, the bias is two orders of magnitude smaller, demonstrating that the total rainfall in this case study is primarily dictated by the large-scale convective forcing. While the perturbations can alter the timing and location of particular storms, they do not affect the time-space averaged moisture budget.

The largest biases, of the order of a few percent, occur for a subset of simulations identified in Section 6.2 as leading to significant changes in cloud distribution: specifically, σ24A1, σ8A1, σ0A1 and IC-1. If a perturbation field is applied to the initial conditions, the rainfall increases throughout the course of the simulation that follows, whereas in the 1 K sequential-perturbation simulations total rainfall is reduced. The reduction occurs primarily between 0500 and 1500 UTC, albeit somewhat offset by a positive bias later (not shown). These results highlight the point that the model is sensitive to strong perturbations at early times, and also suggest that perturbations affecting the spin-up phase of the model can produce markedly different results to those applied later on.

The RMSP of the single-perturbation simulations is shown in Figure 9 together with that of the σ24A1 and σ24A0.01 simulations for comparison. The behaviour for the single-perturbation simulations is broadly similar to that for the sequential perturbations. Of particular note is that the IC-1 and IC-0.0.1 simulations behave similarly in RMSP terms to their sequential-perturbation counterparts, σ24A1 and σ24A0.01 respectively; the difference between the single and sequential simulations lies within the spread of the ensemble generated through different random number realisations (Figure 8(b)). It should, however, be recalled that the IC-1 and σ24A1 simulations have different domain-averaged accumulated precipitation (Table I). This emphasises the importance of considering a range of diagnostics when assessing the impact of perturbations.

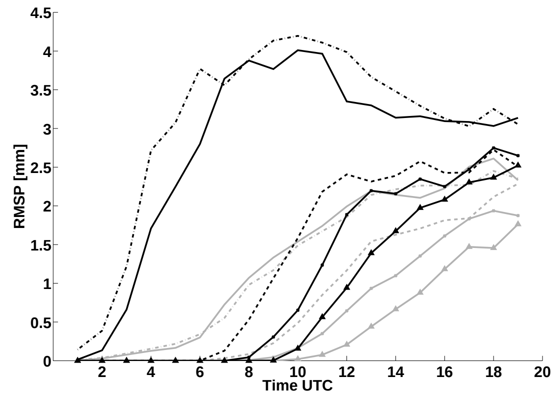


Figure 9. Evolution of the RMSP for six single-perturbation simulations with different perturbation amplitudes (0.01 K in light grey and 1 K in black) and application times (in the initial conditions as dot-dashed lines, at 0700 as dashed lines, at 0830 as solid lines with filled circles and at 1000 UTC as solid lines with filled triangles). $\sigma_{\rm gauss} = 24$ km in each case. Also shown are the RMSP for the sequential-perturbation simulations $\sigma 24A1$ (black solid line without symbols) and $\sigma 24A0.01$ (light grey solid line without symbols). Times along the horizontal axis refer to the beginning of the hour of accumulation.

The RMSP curves for the 0.01 K simulations show important changes with the time of perturbation application. The IC-0.01 run exhibits a clear change in growth rate after 0600 UTC. For the 0700-0.01, 0830-0.01 and 1000-0.01 simulations strong RMSP growth rates are achieved after around an hour, as opposed to five hours for both the sequential and IC perturbations. This is consistent with the hypothesised sensitivity of perturbation growth to the state of the boundary layer, as discussed earlier in this section.

In some cases with the 1 K perturbation amplitude the RMSP reaches a clear saturation level. In general though the later a single strong perturbation is applied, the less likely the RMSP is to reach saturation and the smaller the RMSP at the end of the simulation. The maximum

Prepared using qjrms3.cls
DOI: 10.1002/qj

growth rate of RMSP also reduces if single perturbations are applied later (Table I), clearly showing that perturbing early in the day is most effective in producing perturbation growth.

6.4 Intensity and displacement errors

A decomposition of the squared RMSP was presented in Section 4. The number of rainy points in the control simulation has been shown earlier, in Figure 1, and is equivalent to $N_{\mathrm{COMMON}} + N_{\mathrm{CONTROL}}$. Here we discuss the decomposition for the sequential-perturbation simulations based on Figure 10 which shows the three contributing MSP components on the left and the fractions of each type of rainy grid point (relative to the total number of rainy points) on the right.

Consider first the COMMON points, which are rainy in both the perturbed and control simulations. The precipitation intensity at such points is altered in all of the perturbed simulations but more strongly and more quickly for the larger perturbation amplitudes (Figure 10(a)). This contribution dominates the total MSP at early times. Consistent with this observation, and with Figure 8, the MSP contribution from COMMON points grows more slowly and reaches a peak at later times for decreasing perturbation amplitudes. For a perturbation amplitude of 1 K, the fraction of COMMON points decreases from the start of the simulations (Figure 10(b)). Indeed, by around 0700 UTC most of the rainy grid points in these simulations differ from those in the control simulation. Thus, the 1 K perturbations are extremely effective from the outset at both displacing storms and altering the intensity of COMMON storms. By contrast, at the same time in the simulations with weaker perturbation amplitudes the rain occurs in predominantly the same locations and at similar rates to the control simulation.

The points that are rainy only in the perturbed or control simulations are now considered. For the two smaller perturbation amplitudes these two sets of points exhibit broadly similar behaviour. The simulations with 0.01 K perturbation amplitude start to generate points with a different rain status to the control simulation (i.e. raining in the perturbed run but not in the control or vice versa) around 1-2 h after such points are generated by 0.1 K perturbation simulations (Figures 10(d) and 10(f)). Once produced though, the growth rates of the MSP contributions and of the fractions of those points are similar, so that the same timing difference remains perceptible throughout the remainder of the simulations. With these perturbation amplitudes, the MSP contributions from PERTURBED and CONTROL rainy points are roughly equal.

By contrast, the simulations with 1 K perturbation amplitude have a different pattern of behaviour for rainy but non-COMMON points. As seen in the comparison of 0.1 K and 0.01 K simulations, more storms are displaced earlier for a stronger perturbation amplitude. However, at early times the 1 K perturbations are more effective at triggering new storms than they are at suppressing storms seen in the control simulation. Thus, the fraction of PERTURBED points and their contribution to MSP grows rapidly up to around 0700 UTC (Figures 10(e), 10(f)). Beyond that time, the ability of the perturbations to trigger new storms, and the intensity of such storms, increases only slowly if at all. Interestingly, the growth of the fraction of CONTROL points, and their contribution to MSP, appear to stall at around the same time (0400-0600 UTC, Figures 10(c),10(d)), indicating a reduced ability of 1 K perturbations to alter existing storms. Thus, we can see that the period between 0600 and 0800 UTC is a critical one for the development of storms. It is during this time that perturbations of weaker amplitude first become

DOI: 10.1002/qj

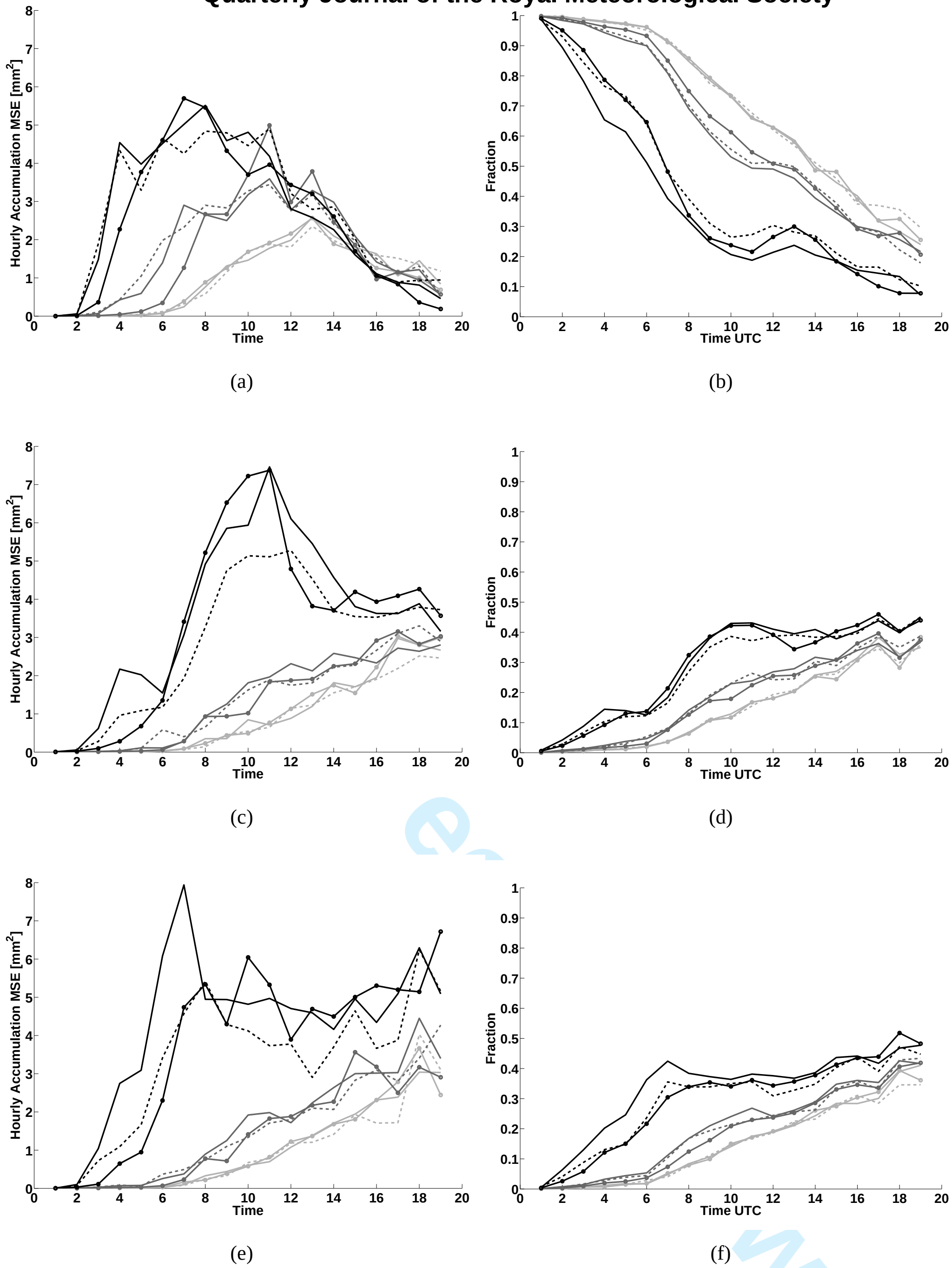


Figure 10. Panels in the left-hand column show the contributions to the squared RMSP from the grid points that are classified as (a) COMMON, (c) CONTROL and (e) PERTURBED. Such points, and their contributions to the squared RMSP, are defined in Section 4. Panels in the right-hand column show the fractions (b) N_{COMMON}/N, (d) N_{CONTROL}/N and (f) $N_{\mathrm{PERTURBED}}/N$. In all panels, results are shown for nine sequential-perturbation simulations with different perturbation amplitudes (0.01 K in light grey, 0.1 K in dark grey and 1 K in black) and standard deviations (24 km as solid lines with filled circles, 8 km as solid lines and 0 km as dashed lines).

effective at displacing storms. The 1 K perturbations meanwhile are extremely effective at producing additional storms, without greatly suppressing the triggering of storms in the control simulation (note the small MSP contribution from CONTROL points prior to 0700 UTC).

It was noted in Section 6.3 that until mid-afternoon the sequential-perturbation simulations with 1 K perturbation amplitude produce a little less rain in total than in the control simulation. Recalling this point, the results above contrasting CONTROL and PERTURBED points imply that

Prepared using qjrms3.cls
DOI: 10.1002/qj

up to 0700 UTC, the 1 K perturbations must be effective at reducing the strength of those storms that are in COMMON between the simulations. In essence, the strong perturbations produce more, but less intense, storms at this time.

For much of the morning the storms present only in the control simulation are strongly affected by the 1 K perturbations. From about 1000 UTC the fraction of such points remains constant or increases only slightly, whereas their contribution to the MSP decreases during the late morning and early afternoon. Therefore, on average the intensity of these storms decreases. Comparing Figures 10(e) and 10(c), we also note that the PERTURBED storms are stronger than the CONTROL storms in the afternoon and evening.

In general the MSP decomposition is only slightly sensitive to the standard deviation of the perturbations σ_{gauss}, particularly for the perturbation amplitude 0.01 K. However, for the larger amplitudes there are indications in the σ8A1 and σ8A0.1 simulations that the 8 km standard deviation is consistently the most effective at displacing the storms.

7 Summary and conclusions

The processes leading to the growth of convective-scale model-state perturbations (specifically perturbations in potential temperature), and the sensitivity of the perturbation growth to the perturbation characteristics, have been investigated for a case study from the CSIP field campaign. The case was chosen because it was strongly upper-level forced but with detailed mesoscale/convective-scale evolution that was dependent on smaller-scale processes. The focus of this study is the identification of processes leading to perturbation growth – determination of the relative importance of these processes is left as future work.

The potential temperature was perturbed at a fixed model level within the boundary layer, usually a little above the lifting condensation level (sensitivity studies showed little sensitivity to the height of the perturbation). Various perturbation amplitudes and horizontal length-scales were considered, and perturbations were applied either once only (at various specific times) or else sequentially (applied every 30 min. throughout the run and uncorrelated in time). In all cases the perturbation fields generate alternative realisations of the flow that are consistent with the large-scale conditions (large transient changes in the model evolution are not created, nor do the changes in the convective-scale evolution significantly modify the large-scale conditions).

Diagnostics were carefully selected to elucidate both the direct (within one timestep) and indirect effects (as evolved by the model) of the perturbations on the model. Motivated in part by hydrological considerations we have also developed diagnostics to distinguish changes in precipitation intensity from changes in the location and distribution of clouds.

The direct effects of the perturbations on CAPE are small, except for a very few points where the strongest perturbations generate or remove a convective lid. These create the conditions for changes in storm location and so favour localised perturbation growth. Similarly, there are some direct, localised, effects on the condensate at the perturbed level. The perturbations also have a direct effect on the model's boundary-layer types, leading to a switching of the type at some grid points (at up to 7% of points in the domain for the largest amplitude perturbations; Table II). Such switches will change the model evolution by activating different parameterisations and causing different coefficients to be used within the parameterisations. On the larger scale, the direct effect is the generation of Lamb and

Q. J. R. Meteorol. Soc. 00: 1–27 (2009)
DOI: 10.1002/qj

acoustic waves that rapidly affect some model fields (e.g. pressure and vertical velocity) throughout the domain. Such waves will produce a slightly different environmental profile into which the convective plumes ascend.

Continued perturbation growth throughout the model integrations has been analysed in terms of the evolving changes to boundary-layer types, cloud distributions and root-mean-square error of the hourly-accumulated precipitation (RMSP). Overall, the amplitude of the perturbations is the main determinant of perturbation growth, although the perturbation lengthscale and single/sequential character do have a modulating role on some of the diagnostics.

There are various indications that qualitatively different perturbation growth behaviour occurs for strong (1 K) and weaker (0.1 K and 0.01 K) amplitude perturbations. In various respects, the effects of strong perturbations are not simply a more intense version of the effects seen in weaker perturbation simulations. Relevant indicators include the extent of boundary-layer switching, the cloud size and number, the timing of RMSP growth and various aspects of storm displacement and generation. For example, early in the day strong perturbations are highly effective at triggering different storms, but less effective at suppressing the storms found in the control simulation. The weaker perturbations do not generate or suppress storms (or result in significant perturbation growth) until later, a little after sunrise, but are then equally likely to induce either generation or suppression. Thus it appears that the impact of weaker perturbations applied before sunrise is to modify the environment into which the convective plumes will later rise rather than to immediately lead to perturbation growth, as measured by RMSP. Despite these (and other) important differences, the RMSP at the end of the day is similar for all perturbation amplitudes. This indicates, on the one hand, that the RMSP is a somewhat crude indicator of perturbation growth because it is not sensitive to important features but, on the other hand, that the nonlinearities of the atmosphere are such that the saturation level of perturbation growth is relatively independent of the perturbation amplitude.

The spread in RMSP due to changes in the horizontal lengthscale of the perturbations is similar to that generated by alternate realisations (different random number seeds) with identical perturbation characteristics. However, there are indications of systematic dependences on lengthscale for some aspects of timing of perturbation growth, storm displacement and generation. In addition, for the largest amplitude perturbations, smaller lengthscales result in more, but smaller, clouds.

Finally, some qualitative differences have also been found in the response to strong perturbations applied to the initial conditions. These differences are not apparent from the RMSP but can be seen in the cloud distributions and the sign of the small precipitation bias. At least for this case, the model may be sensitive to perturbations applied during spin-up, before it has balance-adjusted the initial conditions interpolated from a coarser grid.

8 Acknowledgements

This work has been funded through the NERC (Natural Environment Research Council) FREE (Flood Risk from Extreme Events) programme grant NE/E002137/1. We are grateful to the Met Office for making available the MetUM, and to NCAS (National Centre for Atmospheric Science) CMS (Computational Modelling Support) for providing computing and technical support.

References

Berner J, Shutts GJ, Leutbecher M, Palmer TN. 2009. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *J. Atmos. Sci.* **66**: 603–626.

Browning KA, Blyth AM, Clark PA, Corsmeier U, Morcrette C, Agnew JL, Ballard SP, Bamber D, Barthlott C, Bennett LJ, Beswick KM, Bitter M, Bozier KE, Brooks BJ, Collier CG, Davies F, Deny B, Dixon MA, Feuerle T, Forbes RM, Gaffard C, Gray MD, Hankers R, Hewison TJ, Kalthoff N, Khodayar S, Kohler M, Kottmeier C, Kraut S, Kunz M, Ladd DN, Lean HW, Lenfant J, Li Z, Marsham J, McGregor J, Mobbs SD, Nicol J, Norton E, Parker DJ, Perry F, Ramatschi M, Ricketts HMA, Roberts NM, Russell A, Schulz H, Slack EC, Vaughan G, Waight J, Watson RJ, Webb AR, Wareing DP, Wieser A. 2007. The Convective Storms Initiation Project. *Bull. Amer. Meteor. Soc.* **88**: 1939–1955.

Browning KA, Morcrette CJ. 2006. A summary of the Convective Storm Initiation Project Intensive Observation Periods. Technical Report 474, MetOffice R&D, URL `http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html`.

Bryan GH, Wyngaard J, Fritsch J. 2003. Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.* **131**: 2394–2416.

Chagnon JM, Bannon PR. 2005. Wave response during hydrostatic and geostrophic adjustment. Part I: Transient dynamics. *J. Atmos. Sci.* **62**: 1311–1329.

Clark PA, Lean HW. 2006. An overview of high resolution UM performance for CSIP cases. Technical Report 478, MetOffice R&D, URL `http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html`.

Collier CG. 2006. Flash flood forecasting: What are the limits of predictability? *Q. J. R. Meteorol. Soc.* **133**: 3–23.

Crook NA. 1996. Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.* **124**: 1767–1785.

Davies T, Cullen MJP, Malcolm AJ, Mawson MH, Staniforth A, White AA, Wood N. 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.* **131**: 1759–1782.

Done JM, Graig GC, Gray SL, Clark PA. 2008. Case-to-case variability of predictability of deep convection in a mesoscale model. *Q. J. R. Meteorol. Soc.* **submitted**.

Ebisuzaki W, Kalnay E. 1991. Ensemble experiments with a new lagged analysis forecasting scheme. Research activities in atmospheric and oceanic modeling. Technical Report 15, WMO.

Edwards J, Slingo A. 1996. Studies with a flexible new radiation code. Part I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.* **122**: 689–719.

Essery R, Best M, Cox P. 2001. MOSES 2.2 Technical Documentation. Technical Report 30, Hadley Centre, URL `http://www.metoffice.gov.uk/research/hadleycentre/pubs/HCTN/index.html`.

Fabry F. 2006. The spatial variability of moisture in the boundary layer and its effect on convection initiation: Project-long characterization. *Mon. Wea. Rev.* **134**: 79–91.

Garratt JR. 1992. *The atmospheric boundary layer*. Cambridge University Press. Pp. 316.

Gebhardt C, Theis SE, Paulat M, Bouallegue ZB. 2009. Uncertainties in cosmo-de precipitation forecasts introduced by model perturbation and variations of later boundaries. *Atmos. Res.* **submitted**.

Grabowski WW, Bechtold P, Cheng A, Forbes R, Halliwell C, Khairoutdinov M, Lang S, Nasuno T, Petch J, Tao WK, Wong R, Wu X, Xu KM. 2006. Daytime convective development over land: A model intercomparison based on LBA observations. *Q. J. R. Meteorol. Soc.* **132**: 317–344.

Gregory D, Rowntree PR. 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.* **118**: 1483–1506.

Hand WH, Fox NI, Collier CG. 2004. A study of the twentieth-century extreme rainfall events in the United Kingdom with implications for forecasting. *Meteorol. Appl.* **11**: 15–31.

Hohenegger C, Luthi D, Schar C. 2008a. Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.* **134**: 2095–2107.

Hohenegger C, Schar C. 2007a. Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.* **88**: 1783–1793.

Hohenegger C, Schar C. 2007b. Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.* **64**: 4467–4478.

Hohenegger C, Walser A, Langhans W, Schar C. 2008b. Cloud-resolving ensemble simulations of the August 2005 Alpine flood. *Q. J. R. Meteorol. Soc.* **134**: 889–904.

Kong F, Droegemeier KK, Hickmon NL. 2006. Multiresolution ensemble forecasts of and observed tornadic thunderstorm system. Part I: Comparison of coarse- and fine-grid experiments. *Mon. Wea. Rev.* **134**: 807–833.

Kong F, Droegemeier KK, Hickmon NL. 2007. Multiresolution ensemble forecasts of and observed tornadic thunderstorm system. Part II: Storm scale experiments. *Mon. Wea. Rev.* **135**: 759–782.

Lean HW, Ballard SP, Clark PA, Dixon MA, Li Z, Roberts

Prepared using qjrms3.cls
DOI: 10.1002/qj

RM. 2005. The summer 2004 reruns with the High Resolution Trial Model. Technical Report 466, MetOffice R&D, URL http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html.

Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, 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.

Lean PW. 2006. The predictability of convective storms over the ocean. Phd thesis, Department of Meteorology, University of Reading, UK.

Lindzen RS, Blake D. 1972. Lamb waves in the presence of realistic distributions of temperature and dissipation. *J. Geophys. Res.* **77**(12): 2166–2176.

Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB. 2000. A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.* **128**: 3187–3199.

Molteni F, Buizza R, Marisgli C, Montani A, Nerozzi F, Paccagnella T. 2001. A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments. *Q. J. R. Meteorol. Soc.* **121**: 2069–2094.

Nicholls ME, Pielke RA. 2000. Thermally induced compression waves and gravity waves generated by convective storms. *J. Atmos. Sci.* **57**: 3251–3271.

Nieuwstadt FTM, Brost RA. 1986. The decay of convective turbulence. *J. Atmos. Sci.* **43**: 532–546.

Plant RS. 2008. Statistical properties of cloud life cycles in cloud-resolving models. *Atmos. Chem. Phys. Discuss.* **8**: 20 537–20 564.

Roberts NM. 2003. Stage 2 report form the storm-scale numerical modelling project. Technical Report 407, MetOffice R&D, URL http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html.

Roberts NM, Lean HW. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.* **136**: 78–97.

Robinson FJ, Sherwood SC, Li Y. 2008. Resonant response of deep convection to surface hot spots. *J. Atmos. Sci.* **65**: 276–286.

Shutts GJ. 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.* **612**: 3079–3102.

Snyder C, Zhang F. 2003. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.* **131**: 1663–1677.

Stull RB. 1988. *An introduction to boundary layer meteorology*. Kluwer Academic Publishers, first edn. Pp. 670.

Walser A, Lüthi D, Schär C. 2004. Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.* **132**: 560–577.

Weinbrecht S, Mason PJ. 2008. Stochastic backscatter for cloud-resolving models. Part I: Implementation and testing in a dry convective boundary layer. *J. Atmos. Sci.* **65**: 123–139.

Weisman ML, Davis C, Wang W, Manning KW, Klemp JB. 2008. Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting* **23**: 407–437.

Wilson DR, Ballard SP. 1999. A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.* **125**: 1607–1636.

Zhang F, Snyder C, Rotunno R. 2003. Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.* **60**: 1173–1185.

Prepared using qjrms3.cls

DOI: 10.1002/qj