Stochastic parametrization of multiscale processes using a dual-grid approach

BY GLENN SHUTTS¹,*, THOMAS ALLEN¹ AND JUDITH BERNER²

¹ Met Office, FitzRoy Road, Exeter EX1 3PB, UK
² European Centre for Medium Range Weather Forecasts (ECMWF), Shinfield Road, Reading RG2 9AX, UK

Some speculative proposals are made for extending current stochastic sub-gridscale parametrization methods using the techniques adopted from the field of computer graphics and flow visualization. The idea is to emulate sub-filter-scale physical process organization and time evolution on a fine grid and couple the implied coarse-grained tendencies with a forecast model. A two-way interaction is envisaged so that fine-grid physics (e.g. deep convective clouds) responds to forecast model fields. The fine-grid model may be as simple as a two-dimensional cellular automaton or as computationally demanding as a cloud-resolving model similar to the coupling strategy envisaged in ‘super-parametrization’. Computer codes used in computer games and visualization software illustrate the potential for cheap but realistic simulation where emphasis is placed on algorithmic stability and visual realism rather than pointwise accuracy in a predictive sense. In an ensemble prediction context, a computationally cheap technique would be essential and some possibilities are outlined. An idealized proof-of-concept simulation is described, which highlights technical problems such as the nature of the coupling.

Keywords: stochastic; backscatter; parametrization; games

1. Introduction

The inclusion of stochastic physical parametrization terms into ensemble prediction systems was motivated by insufficient member spread (relative to forecast error) and justified by the inherent statistical nature of the physical processes being parametrized (Buizza et al. 1999; Palmer 2001). Buizza et al. (1999) devised a simple technique for perturbing the physical parametrization tendencies using a latitude/longitude ‘tile’ pattern of random numbers. The random number associated with each tile was chosen to lie between 0.5 and 1.5 with uniform probability density function and the parametrization tendency for any grid point lying within the tile is multiplied by it. Positive impact on both spread and skill was found to be optimized for 10 degree tiles and a 6 hourly update frequency. The resulting improvements to spread and skill were modest and other approaches have been investigated.

* Author for correspondence (glen.shutts@metoffice.gov.uk).

One contribution of 12 to a Theme Issue ‘Stochastic physics and climate modelling’.
Shutts (2005) proposed a kinetic energy backscatter scheme that parallels the use of stochastic backscatter in large eddy simulation of turbulence (Mason & Thomson 1992) but computes a generalized dissipation rate function composed of contributions from numerical diffusion, mountain drag and deep convection. The underlying pattern of the two-dimensional stream function forcing arising from this was determined from a cellular automaton (CA) that had spatial characteristics reminiscent of mesoscale cellular convection. A revised version of this scheme replaces the CA with a spectral pattern generator in which each spectral component is evolved in time as a first-order autoregressive process with prescribed power spectrum (Berner et al. in preparation). These schemes should be contrasted with Mason & Thomson’s backscatter forcing pattern that was determined by smoothing the curl of a three-dimensional vector field whose components were random numbers updated at every other time step. No attempt was made to mimic the statistical properties of the filter-scale eddies.

It would seem to be desirable for the forcing patterns used in stochastic parametrizations to have the same statistical properties and autocorrelation functions as model error fields (e.g. the statistical properties of the difference field obtained by subtracting the fields from the simulations made with 100 and 10 km horizontal resolution). However, the patterns generated from an algorithm driven by random numbers, while being plausible as ‘model error’ fields, may or may not be likely, given the local state of the forecast model flow. Stochastic backscatter schemes are only flow dependent to the extent that the magnitude of the forcing pattern is scaled by the square root of the dissipation rate of the resolved flow. The random character of the error pattern does not account for its likelihood, given the form of the background flow. The impact of this assumption is hard to know since uncommon error patterns may grow rapidly and common error patterns may decay. In addition, no attempt is currently made to advect the error pattern with the flow. This might prove to be important in getting the appropriate energy transfer between the sub-filter-scale eddy forcing and the explicitly resolved flow.

To address these issues, one could attempt to characterize the properties of the model error fields (obtained from the differences between high-resolution ‘truth’ runs and operational forecast model runs) by space–time correlation functions and devise a forcing function with matching statistical behaviour (Hermanson 2006). Of course, there may still be substantial residual error in the high-resolution truth simulations. For instance, deep convection is neither resolvable nor parametrizable on a 10 km horizontal grid and one must expect important errors in the treatment of mesoscale convective systems and their subsequent upscale impacts.

From the perspective of physical parametrization (in particular, the parametrization of deep convection and gravity wave drag), the phenomena being represented comprise scales that span the whole range between the sub-gridscale and resolved scales of the model. For instance, deep tropical convective systems are driven from the convective updraught scale (less than 1 km) but have mesoscale outflows extending over hundreds of kilometres. The theory of convective parametrization is predicated on the assumption that an ensemble of clouds exist within each grid column of the model so that the predicted tendencies of temperature, moisture and momentum are ensemble averages. Shutts & Palmer (2007) quantified the statistical fluctuations of convective
warming in a big-domain cloud-resolving model simulation by computing the probability distribution functions of coarse-grained apparent warming, conditioned on a measure of the parametrized convective warming. They showed that the variance of the convective warming is typically greater than the mean when coarse grained to $120 \times 120$ km grid boxes. It was also found that the variance tends to increase in proportion to the mean, as implied by the Buizza et al. stochastic physics ansatz.

Randall et al. (2003) proposed a ‘super-parametrization’ methodology (now named multiscale modelling framework) that avoids assuming statistical equilibrium (and the use of the ensemble average) and replaces it by a brute-force approach. In its original inception (Grabowski & Smolarkiewicz 1999), this involved embedding the two-dimensional cloud-resolving models within each grid box of a coarse grid climate model and using the net vertical fluxes from the two-dimensional models as a replacement for convective parametrization. The technique has been subsequently demonstrated to provide superior treatment of the Madden–Julian oscillation and other low-frequency variabilities, but this comes at a substantial computational cost.

In the very different world of computer games technology, there is an inexorable drive towards greater graphical realism and interactivity with a simulated natural environment. The Hollywood film industry is increasingly using computer-generated imagery for special effects involving fluid motion and cloud/smoke rendering. For instance, flight simulators need to render smoke and cloudscapes at useable interactive frame rates for an immersive user experience (e.g. 40 frames per second). Even more remarkable is the development of fluid and particle motion simulators that can simulate the growth of clouds and describe the interaction of a moving aircraft with the cloud itself. This is possible through a combination of powerful graphics processor technology, highly simplified physics and efficient coding techniques (e.g. Harris 2003). From the perspective of computer visualization, predictive accuracy (in the sense of forecasting where a cloud will be at a particular time and what form it might take) is not required, but only visual plausibility. However, code stability and speed are paramount to the games programmer. Interestingly, the semi-Lagrangian advection algorithm used in numerical weather prediction (NWP; e.g. Staniforth & Coté 1991) has been adopted in real-time fluid simulation for visualization owing to its stability with respect to the choice of time step (Stam 2003). Much of the computational overhead in fluid simulation comes from the pressure solver that enforces the continuity of mass. Considerable savings can be obtained by reducing the accuracy requirement, e.g. by reducing the number of iterations in an iterative elliptic solver.

It has been proposed by Shutts & Allen (2007) that the games programmer’s approach to representing natural phenomena could be used to devise fast stochastic pattern generators that couple directly with NWP model flows (e.g. by replacing the stream function forcing scheme with a two-dimensional barotropic fluid solver of much higher resolution than the forecast model). Also, they envisage an efficient—though highly simplified—cloud-resolving model that could be coupled with the forecast model as a dual-grid system. The simulated cloud systems on the finer grid would have realistic spatial structure and temporal evolution and this would be passed to the coarser (NWP) model through coarse-grained fluxes and relaxation terms. This dual-grid approach
applied to each member of an ensemble prediction system would represent a natural extension of current stochastic forcing techniques while at the same time opening the door to physical parametrization that breaks free from the vertical column paradigm.

2. Pattern generators for stochastic parametrization

(a) Microscale models

The use of CA was suggested by Palmer (1997) as a means of describing near-grid scale variability associated with unresolved physical processes such as deep convection (see Gardner (1970) for a discussion of John Conway’s ‘Game of life’). He envisaged a probabilistic CA in which the state of a cell (alive or dead) is governed by a probability distribution function that depends on the state of the cell’s nearest neighbours and is also a function of an associated specification of topographic data, i.e. land/sea, orographic height and land type. Organized convection could be modelled on this fine grid of cells by making convection (represented by a living cell) more likely if the neighbouring cells were already convecting—an assumption that causes the CA convection to cluster. The topographic data could be used to make the state of a cell more likely to be alive (i.e. convecting) over land in the daytime, crudely representing the effect of surface solar heating. Some measure of gross moist atmospheric instability (e.g. convective available potential energy or CAPE) could be used to control the probability of convection so that the probability is zero (or very small) when the CAPE is zero or negative. Considerable interest exists in the methods for designing probabilistic CAs to fit complex natural processes, but their application in meteorology is still largely untested.

The CA used in the kinetic energy backscatter scheme of Shutts (2005) (i.e. CA backscatter scheme or CABS) was inspired by a CA of generations class known as ‘forest fire’ or ‘prairie on fire’. In it each cell can have a certain number of lives (say NLIves, which was equal to 32 in the backscatter scheme) and failure to meet a survival condition causes this to be decremented by 1. A new cell is born provided one out of a set of birth rules is satisfied. The survival condition for a cell is that either three, four or five of its neighbouring eight cells are alive: the birth condition is that two or three of a dead cell’s neighbours are alive and have not lost any lives. In medical speak, any cell that loses a life becomes infertile and reduces the space available for new cells until it loses all of its lives. The value of NLIves and a notional time step between successive CA states effectively puts a time scale into the CA’s evolution. Since sets of living cells can generate new cells on their leading boundary at each step, a maximum propagation speed is determined by the cell’s physical size. This combined with the value of NLIves implies a length scale. Therefore, one has a limited amount of control over the pattern’s space and time autocorrelation scales (figure 1). In principle, NLIves could be made a function of the forecast model’s state so that the autocorrelation scales vary. The CA’s grid in CABS was regular in latitude and longitude and so the spatial autocorrelation scales become smaller with increasing latitude.

For each member of an ensemble forecast, the initial state of the CA was seeded differently, but from then on the CAs are deterministic. Some provision was made in the code for random seeding of the CAs during the course of the
forecast but never tried. This would have been desirable since the CA described above is overly periodic in its temporal behaviour. Further development at the European Centre for Medium Range Weather Forecasts (ECMWF) and the Met Office shifted to the use of a spectral forcing approach that was easier to implement and whose behaviour was better understood (see §2b). The CA approach has yet to be thoroughly tested, and it is now believed that Palmer’s original probabilistic CA proposal offers the most promise.

There are many other types of ‘microscopic model’ that could be used to generate patterns that mimic sub-filter-scale atmospheric flow phenomena. For instance, Khouider et al. (2003) developed a stochastic ‘spin-flip’ model of tropical convection where convective inhibition (CIN) is modelled on a fine grid by a set of interaction rules governing the way the sites of CIN interact with each other and with a larger scale model. Coupling between the coarse-grained variables and the microscopic states is achieved through an interaction potential. In an idealized model of tropical circulation, they were able to demonstrate plausible effects of the stochastic convection on climatology.

Another technique that could be adapted to suit the needs of stochastic parametrization is the popular lattice Boltzmann method where collections of particles on a lattice have a probability distribution function governing their velocities and collisions taking place between the particles at the neighbouring grid points (see Chen & Doolen 1998). In classical applications, the Navier–Stokes equations are obtained for time-averaged behaviour, but the method is amenable to more complex descriptions of the interaction between microscopic and macroscopic behaviours and is well suited to multiphase problems and those with complex geometry (e.g. flow through porous media). Given that the evolution of model flow only involves the nearest neighbour interactions, the method is well suited to parallelization on supercomputers.

The smoothed particle hydrodynamics method for representing fluid motion (see the review by Monaghan 1992) is another potential candidate for fluid pattern generation in stochastic backscatter schemes or direct coupling to a forecast model. In this approach, the fluid is represented by a set of discrete lumps of fluid characterized by their mass, position, velocity and thermodynamic energy or entropy variable. Continuous field descriptions of the flow are obtained using a weighting kernel through which the Newtonian laws of motion for each

Figure 1. (a–c) Patterns generated by the CA used by Shutts (2005) with, from left to right, NLIVES = 10, 50 and 100. The grey scale intensity is a measure of the number of lives remaining with NLIVES being white and zero being black. The resulting animated pattern is reminiscent of convective cloud organization in a cold air outbreak over the sea.
lump can be determined. The method has the advantage of not needing a mesh for computations, and continuity of mass (for the fluid and any other chemical species or conserved property) is implicit in the formulation.

Lastly, the pattern generator could consist of a cloud of the two-dimensional monopolar or dipolar point vortices whose individual motions are controlled not only by self-interaction (i.e. the fluid velocity due to each vortex advecting every other vortex) but also by the flow at some level in a forecast model (e.g. Bühler 2002). The point vortex field could be regarded as a type of model error for the vorticity field at, say, jet stream level and a coarse-graining procedure could be defined for adding velocity increments to the forecast model. The introduction and removal of point vortices together with their strength could be governed by a probability distribution function that depends on the local model dissipation rate.

(b) Spectral method with time evolution by autoregressive process

The spectral stochastic backscatter scheme (SSBS) of Berner et al. (in preparation) used a pattern generator based on a truncated expansion in terms of spherical harmonics. If \( \psi(\lambda, \phi, t) \) is a two-dimensional pattern field on the sphere (with \( \phi \) and \( \lambda \) as the latitude and the longitude, respectively), each spectral coefficient \( (\psi_n^m) \) in the expansion is evolved according to the relation

\[
\psi_n^m(t + \Delta t) = (1 - \alpha)\psi_n^m(t) + g_n\sqrt{\alpha}\epsilon(t),
\]

where \( m \) and \( n \) refer to the zonal wavenumber and the spherical harmonic function degree, respectively; \( (1 - \alpha) \) is the damped, linear autoregressive parameter with \( \alpha \in [0,1] \); \( g_n \) is a wavenumber-dependent noise amplitude; and \( \epsilon(t) \) is a Gaussian white noise process (with zero mean) at time \( t \) with time step \( \Delta t \). The total non-dimensional wavenumber squared is \( n(n+1) \) and the function \( g_n \) can be chosen to provide the desired power spectrum. The autocorrelation time scale \( \tau \) is given by

\[
\tau = -\frac{\Delta t}{\ln(1 - \alpha)},
\]

which, if large relative to \( \Delta t \), can be approximated as \( \alpha \sim \Delta t/\tau \). The spectral function \( g_n \) is assumed to be given by a power law \( g_n = n^p \). In the experiments described by Berner et al. (in preparation), \( \alpha = 0.83 \) (corresponding to an autocorrelation time of approx. 6 hours) and \( p = -5/3 \). A backscatter stream function forcing is computed by multiplying this pattern by the square root of a local model dissipation rate (due to numerical diffusion, mountain wave drag and convection).

The SSBS has undergone extensive testing and tuning in the ECMWF ensemble prediction system at T255 resolution, and improvements in skill and member spread have been clearly demonstrated. Figure 2 (from Berner et al. in preparation) shows the positive impact of the SSBS and reduced initial perturbations on two probability skill scores: the ranked probability and ignorance skill scores. A small positive impact is seen every day in these 10-day forecasts and is particularly noticeable in the ignorance skill score (for details, see Roulston & Smith 2002). Improvements are more striking in the tropics where the ensemble system has insufficient spread and model dynamics is not active enough.
Figure 3 shows the impact of the SSBS on the systematic error in the Northern Hemisphere winter 500 hPa geopotential height field in a multi-year integration run at T95 resolution. It can be seen that the tendency of the forecasts without backscatter (figure 3b) has excessively low heights in Icelandic and Aleutian regions that are greatly reduced by stochastic backscatter along with a corresponding reduction in the positive height errors over the subtropics. This is matched by an increased frequency of synoptic blocking events near these regions, and this increase helps to improve the tendency of the model to underpredict blocking. The detailed mechanism by which backscatter improves the blocking frequency is not known, although one plausible explanation might be that the increase in small-scale eddy kinetic energy enhances the effectiveness of the eddy-straining process that maintains blocking flow patterns (Shutts 1983).

Figure 2. (a–c) Ranked probability skill score (10 categories) and (d–f) ignorance skill score for T255 forecast events in the Northern Hemisphere extratropics (20040501 00Z–20050426 00Z; 46 cases). In (d–f) anomalies exceed +1.5 s.d.; (a,d) 500 hPa geopotential height (z); (b,e) zonal wind (u) at 850 hPa; and (c,f) temperature (t) at 850 hPa. Solid line, no SSBS; dashed line, SSBS.
3. Computer animation and visualization techniques for fluid flow and cloud dynamics

In this section, we will compare the objectives of the computer visualization industry with those of the weather and climate forecasting business and argue that the techniques employed in the former can potentially be adopted in numerical weather prediction. This search for radical new approaches is borne out of a frustration with the conventional one-dimensional view of sub-gridscale parametrization in weather forecast models. In a general sense, these new approaches represent a desire to find better ways of deploying computer resource and a move away from the use of statistical equilibrium closure assumptions in parametrization. Advances in computer graphics hardware and algorithms provide a refreshing alternative source of inspiration for those involved in improving NWP model ‘physics’.

Since the earliest personal computer systems became available in the early 1980s, there has been a strong drive for physical realism in games and simulation software. This drive towards improving physical realism in computer games led Sony, along with IBM and Toshiba, to develop the Cell Broadband Engine for the PlayStation 3 games console. This processor will be used as the basis for the next Los Alamos supercomputer (Roadrunner).

With vast increases in microprocessor power, the physical models in games have come a long way from simply describing the parabolic path of a ball under the action of gravity. Flight simulators and Hollywood special effects need realistic and interactive descriptions of the movement of fluids and particulates together with convincing imagery of cloudscapes, smoke and even flowing water. Computer animation of fluid motion and cloud is now based on the same physical laws and equation sets which meteorologists and oceanographers use, and in many ways the constraints under which the games programmers operate are similar to those involved in NWP, i.e. the need for speed and stability. Whereas the forecast model has to at least keep ahead of (and in practice is approx. 100
times faster than the real phenomenon) real time, a cloud modeller might be happy just to keep pace. In a flight simulator game, one might hope to have realistic-looking three-dimensional cloud motion, evolution and illumination at 40 frames per second—something that current cloud-resolving models would not be capable of as a direct output. So why is this?

Part of the reason lies in a crucially different motivation: the games programmer only has to create the illusion of reality whereas the cloud modeller (and particularly the weather forecaster) has to focus on accuracy of representation. For instance, Stam (1999, 2003) uses the semi-Lagrangian advection algorithm with very large time steps to simulate neutrally buoyant fluid flow carrying smoke or cloud droplets. Since low-order interpolation of flow variables into the departure point acts to unnaturally smooth the fields, Stam has adopted the ‘vorticity confinement’ technique of Steinhoff & Underhill (1994; see also Steinhoff et al. 2006) to counteract spurious dissipation and maintain vorticity features close to the model’s gridscale. Steinhoff’s technique has even been used by the Hollywood film industry in the creation of realistic computer-generated smoke scenes. The simplest way of viewing vorticity confinement in two dimensions is the addition of an apparent force to the momentum equation that acts tangentially to the vorticity surface with strength proportional to the magnitude of the vorticity itself. Its effect is to maintain the strength of vorticity features (e.g. vortex sheets/tubes and fronts) close to the gridscale.

Inspired by the work of Stam, a flow past buildings code was written to confirm the extra speed attainable when the requirement for accuracy is relaxed. The code executes sufficiently quickly for the user to be able to interact with the display and place smoke tracers into the computational domain via mouse pointer actions. The authors were impressed by the efficiency and realism of the simulation and realized the potential for this approach to be used as a more sophisticated replacement for the CA pattern generator in stochastic parametrization.

An alternative to vorticity confinement occurred to the authors when making a more precise estimate of the numerical diffusion associated with the first-order semi-Lagrangian scheme. By subtracting this diffusive contribution from the equation of motion, the desired effect of reducing the numerical dissipation can be achieved. Although such a step is not entirely equivalent to using a higher order interpolant, it has a similar effect but at a much reduced computational cost.

The basic idea behind this method is to perform a ‘modified equation’ analysis (see Morton & Mayers 2005) for the first-order semi-Lagrangian advection scheme. When applied to the one-dimensional advection equation for $\phi$,

$$\phi_t + u\phi_x = 0,$$

one obtains

$$\hat{\phi}_t + u\hat{\phi}_x = \kappa\hat{\phi}_{xx} + R,$$

where $\hat{\phi}$ is the equivalent numerical solution; $R$ represents residual terms in the analysis; and the numerical diffusivity $\kappa$ is given by (see also McCalpin 1988)

$$\kappa = \frac{r(1-r)\cdot(\delta x)^2}{2\delta t},$$

where $\delta x$ and $\delta t$ are the horizontal grid spacing and the time step, respectively, and the parameter $r \in [0,1]$ is the (relative) distance of the departure point from the nearest grid point and is related to the fractional part of the Courant
number. The fact that $\kappa \geq 0$ is independent of the Courant number explains why the method is stable. Note that for non-constant advection, there is also a dispersive correction to (3.2), which is ignored since we are only interested in the dissipation. The fact that $r$ is calculated before the semi-Lagrangian advection is performed suggests that the numerical dissipation terms could be used in place of the vorticity confinement term. The ‘ill-posed’ nature of this anti-diffusion means that some care must be taken with the implementation of such a term in order to maintain numerical stability (e.g. by slightly reducing the anti-diffusivity so that net diffusivity is positive on average). However, the potential advantage of the method is that it can be applied to all fields. The beneficial effects of this method are shown in figure 4, which describes the vorticity field for flow over a series of boxes/buildings. The white streak indicates the dispersion of a passive tracer. The four simulations in this figure are as follows:

— the standard first-order scheme;
— the first-order scheme applied to momentum but with monotone cubic interpolation of the tracer;
— cubic interpolation of all variables; and
— the first-order scheme with the proposed anti-diffusion.

The simulation that used the anti-diffusion scheme was computationally as cheap as the standard first-order scheme and typically approximately four or five times faster than those using cubic interpolation.

As with vorticity confinement, the emphasis is on creating the effect of higher resolution without the attendant pointwise accuracy. The additional terms are like deterministic backscatter and probably do much to improve the statistical properties of the flow without attempting to have extra accuracy in the Lagrangian advection step.

Another way the games programmer has achieved speed is to simplify the equations of moist thermodynamics, cloud microphysics and radiative transfer. The use of neural networks for radiation schemes strongly suggests that current, rather expensive, schemes are overly complex and that the desired effect can be obtained via a simplified ‘function’. Similarly, it is not unreasonable to question the extent to which the complexity of cloud microphysics is necessary for the modelling of the overall statistical character of deep convection. Indeed, it is an interesting challenge to perform a kind of ‘algorithmic compression’ and see whether much simpler parametrization could suffice. The reason this is not quite the outrageous assertion that it seems to be is that cloud-resolving models (and the latest generation of ultra-high, limited-area forecast models) are frequently run with a resolution of the order of magnitude less than a true cloud-resolving model needs (e.g. 2 km horizontal grid length when 200 m is required). Much of the complexity in the description of cloud microphysics is due to excessive numerical diffusivity in the cloud scales. Shutts & Allen (2007) described a simple ‘fake’ cloud-resolving model that attempts to create highly simplified moist physics while permitting believable deep convective clouds.

It is convenient for our purposes to use the term emulator to describe a class of computer model that mimics a conventional numerical simulation (i.e. the simulator such as a cloud-resolving model) but employs a much higher level of simplification than one would regard as acceptable for a research study. The
emulator therefore imitates the simulator and the simulator models reality. For the purposes of parametrization, the emulator must at the very least be numerically stable and sufficiently accurate to produce plausible patterns for use in stochastic parametrization. At the other end of the spectrum, the emulator may use the same equations as the simulator and ultimately converge to it if the accuracy requirements are successively increased.

4. Proposal for a dual-grid class of stochastic parametrization

The combined stochastic pattern generator (or emulator) and the forecast model constitute a dual-grid system in which the former acts to force statistical fluctuations due to sub-filter-scale processes. The potential level of complexity of the emulator could range from a simple two-dimensional CA to a three-dimensional cloud-resolving model with simplified physics (similar to the super-parametrization concept of Randall et al. 2003). In the computationally cheapest scenario, all the conventional column-based parametrization schemes are retained and the emulator modulates their output. The computational demands of this kind of emulator are very small relative to those of the conventional forecast model, although still relatively high compared with the ‘pure’ dynamical core. By contrast, the computational burden of the cloud-resolving model limit would exceed that of the forecast or climate model. The role of the forecast model would then be to merely orchestrate sub-mesoscale processes in a fine-scale model and remove the need for many of the sub-gridscale parametrizations (e.g. deep convection and mountain drag). Wherever the computational demand lies, it is important to remember that a fine-scale emulator has substantial compromises in its pointwise accuracy relative to the coarser NWP or climate model. In effect, the combined system does its dynamics and physics at different resolutions and on different grids with the physics being computed at higher resolution but with degraded accuracy for speed.

On a related side, note the ECMWF have recently compared forecast accuracy in their forecasting system when dynamics and physical parametrizations are computed at different resolutions. Somewhat surprisingly, it was found that forecasts made with a higher resolution for the physics grid than the dynamics gave more forecast skills than the combination of higher resolution dynamics and coarser resolution physics (M. Hortal & D. Salmond, unpublished work). In the former case the physical parametrizations derive their input from interpolated dynamical fields whereas in the latter case it is obtained from coarse-grained fields. The idea of using different grids to compute dynamics and physics is similar to our proposed dual-grid system in which ‘the physics’ is played out on a fine grid using an emulator.

The coupling issue is now explored by considering the evolution of some scalar variable $q$ defined to be $q^F$ on the fine grid and $q^C$ on the coarse grid. If the scalar is advected by an incompressible three-dimensional vector velocity field $\mathbf{U}$ for which

$$\nabla \cdot \mathbf{U} = 0,$$

then $q^F$ satisfies

$$q_t^F + \nabla \cdot (\mathbf{U} q^F) = F_q,$$

where $F_q$ represents the forcing term. This equation describes the evolution of the scalar field $q^F$ under the influence of advection by the velocity field $\mathbf{U}$. The coupling between the two grids is achieved by modulating the output of the emulator $q^F$ to influence the forecast model, which in turn modulates the input to the emulator, creating a feedback loop that adjusts the emulator’s predictions to better match the observed patterns.
where \( F_q \) is some unspecified source term and the subscript \( t \) denotes differentiation. Let \( U \) and \( q^F \) be expanded so that

\[
U = \bar{U} + \tilde{U}
\]

and

\[
q^F = \bar{q}^F + \tilde{q}^F,
\]

with \( \bar{q}^F = 0 \).

Now taking the coarse-grain average of equation (4.2) leads to

\[
\bar{q}_t^F + \nabla \cdot (\bar{U} \bar{q}^F) = \bar{F}_q - \nabla \cdot (\bar{U} \tilde{q}^F),
\]

Figure 4. A snapshot of the vorticity field from an interactive movie of flow past a series of buildings. 500 × 125 points were used and the location of buildings is enforced by masking the flow to zero. The white streak is a tracer field introduced at its leading edge on the left of the domain. On a 2.4 GHz Pentium 4 personal computer using OpenGL graphics, the animation was able to run at 15 frames per second. (a) Standard linear scheme, (b) using monotone cubic interpolation on tracer, (c) monotone cubic on all variables, and (d) linear interpolation with anti-diffusion terms.
where the overbars refer to coarse-grain spatial averages and the tildes are the deviation therefrom. Now define the difference \( Q \) between the coarse model \( q \) field and its coarse-grained equivalent in the fine model, i.e.

\[
Q = q^C - \tilde{q}^F. \tag{4.6}
\]

Now suppose that \( Q \) obeys a conservation equation of the form

\[
Q_t + \nabla \cdot (U^C Q) = F_Q, \tag{4.7}
\]

where \( F_Q \) is an unspecified forcing function. Substituting for \( Q \) from equation (4.6) gives

\[
q^C_t + \nabla \cdot (U^C q^C) - \tilde{q}^F_t - \nabla \cdot (U^C \tilde{q}^F) = F_Q,
\]

which can be written as

\[
q^C_t + \nabla \cdot (U^C q^C) - \tilde{q}^F_t - \nabla \cdot (\bar{U}^F \tilde{q}^F) - \nabla \cdot \left[ (U^C - \bar{U}^F) \tilde{q}^F \right] = F_Q. \tag{4.9}
\]

Finally, using equation (4.5), the above equation can be written as

\[
q^C_t + \nabla \cdot (U^C q^C) = F_Q + \overline{F_q} - \nabla \cdot (\bar{U} \tilde{q}^F) + \nabla \cdot \left[ (U^C - \bar{U}^F) \tilde{q}^F \right]. \tag{4.10}
\]

Equation (4.10) can be regarded as the coarse model equivalent of equation (4.2) but with the r.h.s. containing (from left to right) the source term for \( Q \), the coarse-grained source of \( q \), an eddy flux divergence of \( q \) on the fine grid and an advective correction term. Thus far, equation (4.10) is no more than mathematical manipulation motivated by the idea that the evolution equation for \( q \) obeys conservation style laws on both coarse and fine grids. A potentially important aspect of the coupling lies in the specification of the term \( F_Q \) that represents a form of model error source. One approach would be to represent \( F_Q \) as a first-order autoregressive process characterized by judiciously chosen mean and variance. This would suppress the natural tendency of \( q^C \) and \( \tilde{q}^F \) to
increasingly diverge in time and allow for some optimal level of departure between the two fields. Note that it is not desirable for force \( q^C \) to be exactly equal to \( \frac{F}{C^2} \) (i.e. \( Q=0 \)) because the fine-grid calculation is inaccurate at the small-scale end of the spectrum. By contrast, the coarse model uses accurate numerical techniques and in an NWP context ingests ‘observational truth’, particularly at the largest scales.

The coupling should therefore include scale-selective relaxation terms that relax the medium to large scales of the fine model towards the coarse model’s accurate \( q^C \) field as a one-way process (no corresponding relaxation terms would exist in the coarse model). At a practical level, this might be enforced through linear relaxation of fine model point values to those of the coarse model where the points coincide. The other points would be free to evolve without direct influence from the coarse model. At the near-grid scale end of the coarse model, coupling is achieved through the coarse-grained fine model fluxes, the coarse-grained source term \( \frac{F}{C^2} \) and the advective correction term of equation (4.10). These terms also represent a one-way transfer of information, though now from the fine model to the coarse model.

As a proof-of-concept exercise, a simple fluid dynamical problem has been modelled on a fine grid and coupled with a coarse grid, i.e. a dual-grid simulation. The fine grid has 10 times more resolution in both directions than the coarse model but uses the same time step, thereby implying a Courant number 10 times larger (see Shutts & Allen 2007 for details).

The simulated flow consists of three convective plumes ascending through an isentropic atmosphere at rest. The plumes are forced by three heat sources situated at the surface, each of which spans three grid points on the fine mesh. The heat sources are entirely sub-gridscale for the coarse model, and therefore have no direct influence on the coarse model fields in the absence of diffusion and a boundary-layer scheme. Coupling in the momentum and buoyancy equations is achieved using the terms equivalent to those on the r.h.s. of equation (4.10) with \( F_{Q}=0 \) and \( F_{q} \) corresponding to very weak relaxation towards the coarse model.

Figure 5 shows the buoyancy on both coarse and fine-mesh models at some time after the plume has ascended up to the model’s lid. From this figure, it can be seen that, even though the fine mesh solution would be considered highly inaccurate, many of the actual features of an accurately converged solution are present. It can also be seen that the coarse model that only sees the plumes through the coupling to the fine grid reproduces most of the features of the rising plume. It is important to note that, apart from the relaxation term (which could be set to zero in this instance), there are no tunable parameters in this form of parametrization. It is unlikely that a column-based parametrization could generate the upscale influence found in the coupled grid simulation and the additional computational expense, provided not too prohibitive, would be worthwhile.

The feasibility of the dual-grid approach will depend primarily on the speed achievable for the fine-grid model and on the nature of the coupling between the two grids. In the context of ensemble prediction systems, the fine-grid model would necessarily have to be simple because the computational burden would be too high for each forecast member to have a three-dimensional cloud-resolving model associated with it. More realistically, one might consider using a fast two-dimensional fluid emulator (with stochastic forcing) as a fine-grid model associated with each member and relax the large scales of the fine model towards the forecast.
model while relaxing the small scales of the forecast model towards the fine model. In each of these cases, the relaxation process is one way so that, for instance, the forecast model does not receive large-scale information from the fine model.

5. Discussion and conclusion

It is proposed to use techniques similar to those employed by computer games and visualization programmers for emulating the statistical behaviour of some sub-gridscale physical processes. These might be as simple as using a probabilistic CA to describe patterns of convective cloud organization to use as driver in a stochastic parametrization scheme or as ambitious as a three-dimensional cloud system emulator that directly couples with a forecast model (in a way analogous to the super-parametrization concept except with much lower computational cost than a conventional cloud-resolving model). In this latter limit, conventional column-based parametrization would be made redundant, though realistically one should expect that some blend of the emulator and conventional parametrization would be necessary in practice, depending on the computer resources available. It would also be possible to use the two-dimensional convection emulators along each latitude circle. This would resemble the use of a horizontally anisotropic grid (e.g. Shutts 2006) but would have less computational cost.

In the context of ensemble prediction, each forecast member would see a different evolving pattern of stochastic forcing resulting from the underlying stochastic nature of the pattern maker or emulator. The current ECMWF spectral backscatter scheme (SSBS) uses a first-order autoregressive model to describe the time evolution of stream function forcing spectral amplitudes. An interesting next step would be to replace the autoregressive model with a high-resolution barotropic fluid emulator that itself is stochastically forced. The largest scales of the fluid emulator would be relaxed towards the jet stream level flow in the forecast member. In this way, the fluid emulator acts as an elaborate flow-dependent dynamical filter for the random numbers used to make it stochastic. The nature of the coupling between the emulator and the forecast model would probably be critical to the success of this approach.

As computer resources become available, the fluid emulator would become three-dimensional and include a more realistic description of the physics (though still an emulator in the sense of computer games and visualization). Ultimately, one envisages the continual refinement of the emulator until it takes over the role of the NWP model, thereby defining a convergent process. At this stage, one must regard the dual-grid strategy defined here as rather speculative.

Although the stochastic models have been discussed here as if they were merely devices to achieve better forecast model skill scores, it is important to not lose sight of their role in representing observable and quantifiable atmospheric processes. Focusing alone on simple global model metrics of success without studying the nature of the stochastic forcing functions that are implemented in a forecast model will hinder progress. Like any other parametrization scheme, they need to be validated against observations and more comprehensive higher resolution models (e.g. cloud-resolving models).
Apart from the barotropic fluid emulator, future work will be directed at the design of a fast and stable deep convection emulator that can be coupled with a forecast model and hopefully speed up the replacement of column-based convection parametrization.

Thanks to Clive Temperton who alerted the first author to the paper by McCalpin and to Chris Smith for useful discussions. Reproduced with the permission of Her Majesty’s Stationary Office. © Crown Copyright 2008.

References


Palmer, T. N. 1997 On parametrizing scales that are only somewhat smaller than the smallest resolved scales with application to convection and orography. In Proc. ECMWF Workshop on New Insights and Approaches to Convective Parametrization, Reading, UK, 4–7 November 1996.

Palmer, T. N. 2001 On parametrizing scales that are only somewhat smaller than the smallest resolved scales, with application to convection and orography. In ECMWF Workshop on New Insights and Approaches to Convective Parametrization, Shinfield, Reading, UK, 4–7 November 1996.


Phil. Trans. R. Soc. A (2008)