1. Working group 1: Issues in convection

Participants: George Craig (Chair), Klaus Fraedrich, Jhonny Lin, Mitch Moncrieff, Robert Plant, David Thomson, Adrian Tompkins (rapporteur)

1.1. Outstanding problems

Several outstanding problems with current convective schemes in general, and the ECMWF IFS in particular, were identified, and the way stochastic physics could potentially tackle these were discussed.

1.1.1. Propagation

Convection schemes are generally unable to represent the propagation and organisation of convection since they are column based, and the relationship to the large-scale circulation is non-trivial. Examples in the IFS include a lack of convective spatial variability over the Tropical Western Pacific, and the apparent inability to represent the Eastward propagation of convective from the Rockies.

Stochastic physics could approach this problem in two ways. Cellular automaton (CA) based schemes enable communication between cells, and can mimic in a simple way the propagation of convective envelops at a variety of spatial scales. The communication based on ‘nearest neighbour’ relationships is ad-hoc and resolution dependent, but there is no reason why future schemes could not move towards automaton-type schemes based more on meteorological phenomena. One example raised was the communication and triggering of convective via coldpool dynamics, where the communication between grid would be via a simple coldpool propagation model. Other physical phenomena could be similarly modelled.

Additionally, it would be helpful not to restrict the parametrization to the large-scale model grid, but instead to allow the application of the stochastic parameterization on a different filter-scale, coarser than the model grid scale, with the parameterization representing the amount of convective fluctuations about the mean state expected on this scale.

The issues that would need attention would be

- what are the physical phenomena that are missing or that we are trying to represent by propagation-type schemes?
- what is the correct filter scale for convective parameterization?

1.1.2. Physical and Numerical Sources of Variability

Many studies have shown that the application of a “trigger decision” to convective activity is problematic. This combined with numerical artifacts of the implementation implies that convective schemes are inherently “noisy” even when forced by time-invariant forcing. One issue is whether, the nature (i.e. magnitude, spectra...) of this noise is important for the large scales dynamics. A convective scheme that behaves “smoothly” under constant forcing could be used as a basis for a stochastic convective parametrization which then parameterizes the convective “noise” of the correct characteristics, for example, with convective temporal fluctuations that are filter-scale dependent.

- What happens if “parametrized” convective noise is added to a “numerically” noisy convection scheme?

1.1.3. Initiation of convection

Convective initiation is related to the sub-grid variability of many variables such as temperature, humidity, orography, and surface fluxes. A stochastic scheme could invoke more physically based stochastic convective initiation by using, for example, information from joint probability distributions (JPDFs) of temperature, humidity and velocity provided by statistically based cloud schemes, or variability of surface fluxes from tiled land-surface schemes.

- How to couple information on unresolved variability present in different parametrizations?

1.1.4. Large-scale variability

A third more general defect of models, and the IFS in particular, is the general loss of large-scale variability in the medium range; for example the inability to sustain the principal components of the Madden-Julian Oscillation.

Is it not yet demonstrated that the addition of filter-scale convective variability directly leads to an improved representation of large-scale wave activity; indeed, the physical mechanisms by which this could occur are not understood. In fact, stochastic parametrizations of convection could provide a useful experimental framework to gain an understanding of this scale interaction. In parallel, stochastic parametrizations can be developed that directly force these large-scale wave patterns in the dynamical and/or thermodynamical fields, the example being the MJO-type forcing demonstrated in a cellular-automaton type scheme. If successful, these could lead to an understanding of the processes necessary in more physically based schemes.

- Do problems occur when forcing certain propagation characteristics on to a model which would naturally propagate these waves at a different phase velocity?

1.1.5. Humidity

Convective schemes appear to lack sensitivity to humidity, in particular to mid-tropospheric humidity. This could be related to lack of knowledge of the subgrid-scale thermodynamic variability. Stochastic models can take this into account by allowing “bulk-mass” plume, or even multiple stochastic plumes, to appreciate the sub-grid scale distributions of humidity, (in addition to other variables such as cloud water, temperature, vertical velocity and so on). Each cloud in a stochastic model would essentially rise through a different portion of the grid.

Here the issues are:

- How are correlations between temperature and humidity to be dealt with?
- If a stochastic scheme allows for sub-filter scale convective organisation, how would this be taken into account? Observations often show deep convection growing through the locally pre-moistened environments of previous shallow cumuli.
- How the scheme would feed back onto the sub grid humidity distributions, particularly if the convective organisation is parametrized, necessitation a treatment of the convection-humidity distribution overlap

1.2. Components of a future stochastic convection scheme

From the preceding discussion, it is clear that a stochastic convection scheme is desirable, but it is equally clear that there is considerable uncertainty as to how such a scheme should be constructed. However, there are certain characteristics that a future stochastic convection scheme would almost certainly possess.

- A stochastic scheme is not an add-on to a deterministic scheme. This requires a re-thinking of the parametrization problem, and therefore must be physically-based and not an ad hoc noisemaker added to a conventional model.

- Such schemes will probably be non-local, with input and output based on physical scales of phenomena - not simply the model grid sizes.
- Such schemes may interact directly with other parametrizations, not just through resolved variables
- Such schemes will not be based on equilibrium assumptions, but will rather be a prognostic model with a ’memory’. Note that this memory may reside in the boundary layer (e.g. cold pool dynamics) or other parametrization, rather than in the convection scheme itself.

1.3. Validation

In order to test and validate stochastic convective schemes a rigorous set of benchmarks and validation exercises is essential. This will have to include tests of the schemes in isolation, and can not rely on evaluation only on the basis of the full model system, and without regard to the physical processes. Without this, given the current, very poor, understanding of convective phenomena, there is a high risk of getting a right answer for some particular diagnostic, but for the wrong reason, later preventing or impeding improvement in other tests.

Without knowing the nature of the parametrization to be tested, it is not reasonable to design a validation strategy, but there are a wide range of tools and data sets which are likely to be of use.

- The ensemble “spread”, and how it is maintained throughout the medium forecast range.
- Impacts for the longer term climate, and effects on the seasonal forecast system, (for example, the lack of variability in surface winds).
- Idealized tests with uniform forcing, where the correct “answer” is known. One possibility could involve aqua-planet runs with fixed forcing.
- The use of CRMs and/or regional models to provide benchmark test with which to compare stochastic schemes at various spatial scales should be encouraged. For instance these could provide examples of well-resolved convective characteristics (e.g. spacing statistics) produce benchmark examples of propagating systems.
- One-off (very expensive) high resolution GCM runs (such as recently conducted at ECMWF) provide a useful alternative, since they use the same basic model framework with identical parametrizations of other processes (e.g. microphysics).
- Case studies of particular difficult convective phenomena which it is believed stochastic physics approaches can benefit. Examples include the MJO or propagating organised convection over the continents. Note that the use of the term case studies does not imply individual events, but rather individual phenomena where the processes can be linked to statistically significant diagnostics.

1.4. Recommendations

i) Continue development of stochastic convective parameterization. There are benefits to be obtained immediately, even with the current, if inadequate, level of understanding.

ii) ECMWF’s recent efforts to develop a version of the IFS code which allows a separate filter scale for the physics are a welcome development. ECMWF should make full use of this new version to investigate such issues as the effect of parametrization noise added on scales larger than the grid scale.

iii) Make efforts to compare ECMWF's approach to stochastic convection with others, either in the ECMWF system, or preferably in other systems with cleaner problems. This will require new collaborative projects with outside institutions.

iv) Identify and explore the physical basis of the stochastic scheme based on Cellular automaton, bearing in mind that this will lead to a need for consistent modification of other parametrizations.

v) Increase validation work in the context of the Ensemble Prediction System, in addition to the deterministic model.

vi) Target improvements in statistics of specific physical phenomena, and attempt to identify the pathway by which the stochastic variability influences the relevant validation scores.

2. Working group 2: Evidence from observation and direct numerical simulations

Participants: Glenn Shutts (Chair), Judith Berner (rapporteur), Harm Jonker, Adam Monahan, Falk Niehörster, Ceçile Penland, Steve Thomas

2.1. Introduction

Field observational programmes, routine operational data and direct numerical simulation (DNS) can provide datasets from which the influence of unpredictable sub- or near-gridscale phenomena can be assessed and turned into stochastic numerical algorithms. We consider some of the methodologies available to extract the required statistical characteristics of these parameterizable processes.

2.2. General considerations

2.2.1. Conditional distributions

Spatial and temporal scale separations are important to the validity of stochastic parameterizations, and certainly the shape of these spectra enters into the development of such parameterization schemes. Nevertheless, knowledge of the spectra of the quantities to be parameterized is not the only requirement; a realistic description of their distributions is also necessary. In light of this, we recommend that parallel runs of fine- and coarse-resolution models be run to estimate conditional distributions. Given sufficiently large datasets, powerful computational tools (e.g. Markov Chain Monte Carlo analysis) exist to estimate these conditional distributions. It should be noted that these distributions cannot be expected to depend only on the resolved variables at the space and time point under consideration; spatial and temporal correlations may be important. Further, we recommend the investigation of the best methods for ensuring that distributions provided to models reproduce dynamical processes represented in observed distributions. The general comments just made may be applied in particular to satellite data. For example, we ask: how are the coarse-grained moisture fields related to the fine-scale structure estimated from satellite retrievals? Can these relations form the basis for devising a distributional component to the convective parameterizations?

2.2.2. Implementation of results from coarse-graining

It is an outstanding issue how to best implement the results obtained from coarse-graining subgrid-scale models and from observational analysis into NWP. In addition to distribution-based extensions of conventional parameterizations, the potential of running stochastic-dynamic subgrid-models (cellular automata, lattice-Boltzmann models, spin-flip models) or mimicking the subgrid-state by neural networks should continue to be explored.

2.3. Specific issues

2.3.1. Geophysical turbulence

Gage and Nastrom (1986) collected observational data indicating that the energy spectrum of the Earth's atmosphere contains a `spectral kink' separating large and small scales. Their spectra cover scales ranging from 3 km to nearly 10,000 km. The observed spectrum is characterized by a –3 slope downscale enstrophy cascade at large scales and an -5/3 inverse energy cascade at small scales. Charney (1971) attributes the –3 slope at scales above 1000 km to quasi-geostrophic turbulence. The mesoscale dynamics follow a Kolmogorov –5/3 spectral slope. Two different mechanisms have been proposed to explain the observed mesoscale spectra. The first is strongly nonlinear and based on quasi-2D turbulence. Lilly (1983) postulates that it is due to stratified turbulence at small scales. The second mechanism is based on a weakly nonlinear wave theory involving the spectrum of internal waves.

At length scales below 1000 km, Lilly (1983) suggests that small scale sources of energy could be provided by thunderstorms. Small-scale shear instability may also contribute. He argues that only a small amount of this energy needs to inverse cascade in order to account for the observed mesoscale spectrum. Some of these types of atmospheric flows are nonhydrostatic, and therefore to reproduce the observed energy spectrum of the Earth's atmosphere might require running a global nonhydrostatic model with prescribed heat fluxes. The alternative theory postulates that the observed spectra are due to internal waves. This contribution is from modes not possessing PV, but not necessarily with high linear frequencies. If this is true for the atmosphere, then a 3D hydrostatic primitive equations model may be capable of correctly capturing the dynamics of the Earth's atmosphere. However, the time scales may be more restrictive as the explicit treatment of gravity waves could be important. There is tentative evidence suggesting that the spectral kink is visible in results from the GFDL-Princeton SkyHi model, Koshyk, Hamilton and Mahlman (1999). Their approach is to use explicit methods and resolve as much as possible.

2.3.2. Spectral transfer/backscatter

Although it is perhaps not straightforward, it may be useful to develop a methodology to diagnose the spectral transfer in GCMs, in terms of the mean direction (upscale/downscale) as well as in terms of the variability of the spectral flux of energy and scalar variance. The methodology could be based on recent work in the context of LES (e.g. Jonker), or observations (e.g. Lindborg). Due to heterogeneity issues, a wavelet basis may be more appropriate than a Fourier basis as was employed in the LES. Knowledge of the stochasticity of the upscale/downscale spectral transfer events at different scales within the GCM, may serve as a source of inspiration on how to implement subgrid scale stochastics (backscatter) most appropriately, as well as on quantitative issues. Comparison with LES and observations of dissipation rates and spectral fluxes, can be useful in this respect, also with regard to obtaining scaling laws for various meteorological conditions, and for identifying the impact of, for example, latent heat.

A critical point with regard to consistently implementing backscatter is the GCM property to be 'overly' dissipative at resolution scales; stochastic forcing at this scale to represent sub-grid processes, however cunning and/or sensibly based on physical arguments, may be washed away immediately.

Stochastic KE backscatter schemes that account for near-gridscale input of kinetic energy due to deep convection should be calibrated using a coarse-graining methodology applied to a cloud-resolving model. Fields (including advection tendencies) from CRM simulations with order 1 km grid spacing may be used to compute Reynolds' stresses appropriate to a coarse grid with order 100 km resolution. These would include not only the 'cumulus friction' effect associated with vertical momentum fluxes but also horizontal stress divergence associated with mesoscale convective storm organization and its interaction with the background flow. The spatial and temporal coherence of the effective vorticity forcing could then be used to calibrate convective backscatter schemes linked to the presence of CAPE. A similar procedure could be applied to determine the effective eddy forcing of temperature and humidity.

Of more immediate NWP interest would be to implement the type of stochastic backscatter scheme developed by Frederiksen and Davies (1997), which is based on EDQNM turbulence theory applied to the barotropic vorticity equation on the sphere. This would involve the introduction of a wavenumber-dependent backscatter term in spectral space and potentially would compensate for the energy-deficient spectral tail in most current forecast models. Such an approach would provide an interesting alternative to the heuristic approach based on backscattering dissipated energy with forcing patterns derived from cellular automata (i.e., CASBS).

2.3.3. Gravity waves and mountain-drag

Deep convection, mountains and jetstream instabilities launch gravity waves that propagate into the stratosphere and above depositing their pseudo-wave momentum. This process is not easily parameterized, as it does not occur within a single column. The path these wave packets take is sensitive to details of the flow and their amplitude highly dependent on the flow at the point of origin. This uncertainty could be addressed with a stochastic momentum forcing function with the influence of the source being distributed beyond the grid column in which the waves are forced.

Flow blocking effects are less obviously stochastic but given their very important role, it may still be important to consider their sensitivity to flow uncertainty.

2.3.4. Surface fluxes

Surface fluxes generally depend nonlinearly on surface wind speeds, so gridbox averaged fluxes will not generally equal the flux associated with the gridbox averaged wind speed. Furthermore, the gridbox averaged wind speed will not generally equal the gridbox averaged vector wind, which is what is predicted by the GCM.

This issue has been treated previously with e.g. gustiness corrections to bulk formulae. Surface flux calculations could potentially be improved with explicit parameterisations of the joint pdfs of surface fluxes conditioned on the resolved variables; appropriate pdfs will likely not depend only on the resolved variables at that particular time and place (i.e. the random process need not be "white"). These pdfs could also be conditioned on other parameterisations; in particular, higher gustiness would be expected during periods of strong convection.

Direct numerical simulation of the PBL is a natural tool for estimating these surface fluxes. Simulations should be carried out under a broad range of large-scale stability and wind shear profiles to establish the conditional pdfs. Furthermore, integrations of a coupled PBL/surface wave model would be appropriate for estimating conditional distributions in the marine boundary layer.

As far as possible, the pdfs arising from these LES studies should be compared to those obtained from observations. As direct measurements of surface fluxes are less common than those of surface wind, temperature, and humidity, only joint distributions of these quantities may be accessible from observations. These could then be used in association with standard bulk formulae to produce estimates of the conditional flux pdfs.

2.4. Recommendations

i) With regard to the low levels of energy in the tail of the energy spectrum, it is suggested that tests be carried out with the spectral backscatter scheme of Frederiksen based on EDQNM turbulence theory.

ii) Consider using the coarse-graining strategy to determine the statistics of an effective stochastic parameters function for use in the ensemble prediction system and seasonal forecasting. This is best achieved when the high resolution simulations have an order of magnitude finer resolution than the operational forecast model in which the effective forcing function would be used (e.g. CRM -> EPS models). It could also involve running the IFS at say T799 resolution and coarse-graining to T95. The conditional PDF of the effective forcing should be related to integral properties of the local flow such as CAPE, mean shear etc.

iii) Consider the use of CRM simulations to provide statistical information concerning the relationship between surface fluxes and coarse-grained model flow states e.g. relationships may exist between the variance of the surface flux and the amount of CAPE and lower tropospheric humidity.

iv) In addition to traditional analyses like spectra and spatio-temporal correlations, we recommend to develop a methodology to diagnose the spectral energy transfer, quantify the direction of spectral transfer and acquire insight into the variability of upscale/downscale events.

References

Charney, J.G. (1971) Geostrophic turbulence. J. Atmos. Sci., Vol. 28, No. 6, pp. 1087-1095.

Frederiksen, J.S. and Davies, A.G. (1997) Eddy Viscosity and Stochastic Backscatter Parameterizations on the Sphere for Atmospheric Circulation Models. J. Atmos. Sci., Vol. 54, No. 20, pp. 2475-2492.

Gage, K.S. and Nastrom, G.D. (1986) Theoretical Interpretation of Atmospheric Wavenumber Spectra of Wind and Temperature Observed by Commercial Aircraft During GASP. J.Atmos. Sci., Vol. 43, No. 7, pp. 729-740.

Koshyk, J.N. , Hamilton, K. and Mahlman, J.D. (1999) Simulations of the k-5/3 mesoscale spectral regime in the GFDL SKYHI general circulation model. Geophys. Res. Letts., Vol. 26, No. 7, 843-846.

Lilly, D.K. (1983) Stratified Turbulence and the Mesoscale Variability of the Atmosphere. J. Atmos. Sci.,Vol. 40, No. 3, pp. 749-761.

3. Working Group 2: The parameterisation problem in weather and climate models

Participants: Michael Ghil (chair), Pavel Berloff, Roberto Buizza (rapporteur), Leon Hermanson, Martin Miller, Tim Palmer, Ian Roulstone, Prashant Sardeshmukh and Paul Williams

3.1. Summary

- The study and parameterisation of stochastic processes is a promising and rapidly evolving area of research in the atmospheric and oceanic sciences.
- Several formulations of the general problem are useful at this point, and explorations are proceeding with many models and approaches.
- Promising avenues include (but are not restricted to) studying differences between high- and low-resolution models, reconsidering numerical and data-assimilation methods in this context, and studying a hierarchy of models and approaches for sub-grid-scale processes.

3.2. Why are we discussing stochastic processes now?

The introduction of stochastic processes into dynamic models of weather and climate is very timely: the 1990s have seen different groups exploring the use of stochastic processes in models of varying complexity and their effect on the resulting simulations and predictions.

3.3. What is the issue that we are trying to address?

3.3.1. General formulation of the problem

A possible framework to define stochastic parameterisations is the following:

$$\dot{X} = N(X) + N_1, \tag{1a}$$

$$\dot{X} = N(X) + M(X) + R, \tag{1b}$$

$$\dot{X} = N(X) + M(X) + B(X) \cdot F_S. \tag{1c}$$

In Eqs. (1a–c), X is the resolved model state vector, $N(X)$ is the resolved model tendency, $\dot{X}$ is the true tendency, and N_1 is the tendency error. Deterministic parameterisation involves approximating $N1$ as $M(X)$. The error of this approximation is represented by R. One can model R as a stochastic noise vector F_s multiplied by a matrix B, which may in general depend upon X. Thus, $M(X)$ represents the 'deterministic' parameterisation and $B(X) \cdot F_s$ the 'stochastic' parameterisation.

The analysis of tendency errors R can give us information about the representation of sub-grid processes, be it via deterministic or stochastic parameterisations. One way to estimate R is to run a numerical weather prediction (NWP) model at the current resolution, as well as at a much higher one, and to compare the wave-number spectra of the two systems. This approach can provide useful information on whether the remainder R may be well approximated by a truly stochastic process. More specifically, can we write R in a simple way, e.g. as it is done in the ECMWF ensemble prediction system?

A different formulation in terms of resolved and unresolved scales is the following:

$$\dot{x} = N(x, x) + M(x, y) + P(y, y) + F_S(x; y), \tag{2a}$$

$$\dot{y} = Q(x, y) + R(y, y) + G_S(x; y). \tag{2b}$$

In Eqs. (2a,b), x and y represent the resolved and unresolved scales in space and time, and subscript 'S' denotes stochastic processes. The terms F_s and G_s are of 'physical' nature, e.g., they represent cloud processes or radiative-convective interactions. All other terms on the right-hand-side are 'dynamical' in nature. It is clear that the 'physical' terms are more irregular in space and time, and are more apt to be well approximated by stochastic processes with short decorrelation times. This approach is less obvious for the 'dynamical' terms, but still worth exploring.

Insight can be gained by treating the problem formalized in Eqs. (2a,b) in a systematic, rigorous way, on the one hand, and by numerical experimentation on the other. To attack the problem of stochastic parameterisation, formulations (1a–c) and (2a,b) appear to be useful at this time. The most fruitful formulation of the problem will only become obvious when, and if, it is solved.

At this point, it is still open to discussion whether we need stochastic parameterisations in single simulations or forecasts. There is, however, evidence for the benefit of using stochastic schemes in ensemble systems.

3.3.2. Scale separation and treatment of the small scales

A key issue in stochastic parameterisation is that of scale separation, in space and time. There is considerable evidence for observed phenomena in the atmosphere lying along a diagonal in a graph whose axes are their spatial and temporal scales (Fig. 1). This association of scales is of considerable interest in defining resolved and un-resolved variables, but it is not accompanied by much evidence of a spectral gap. The absence of such a gap renders the problem at hand much more difficult.

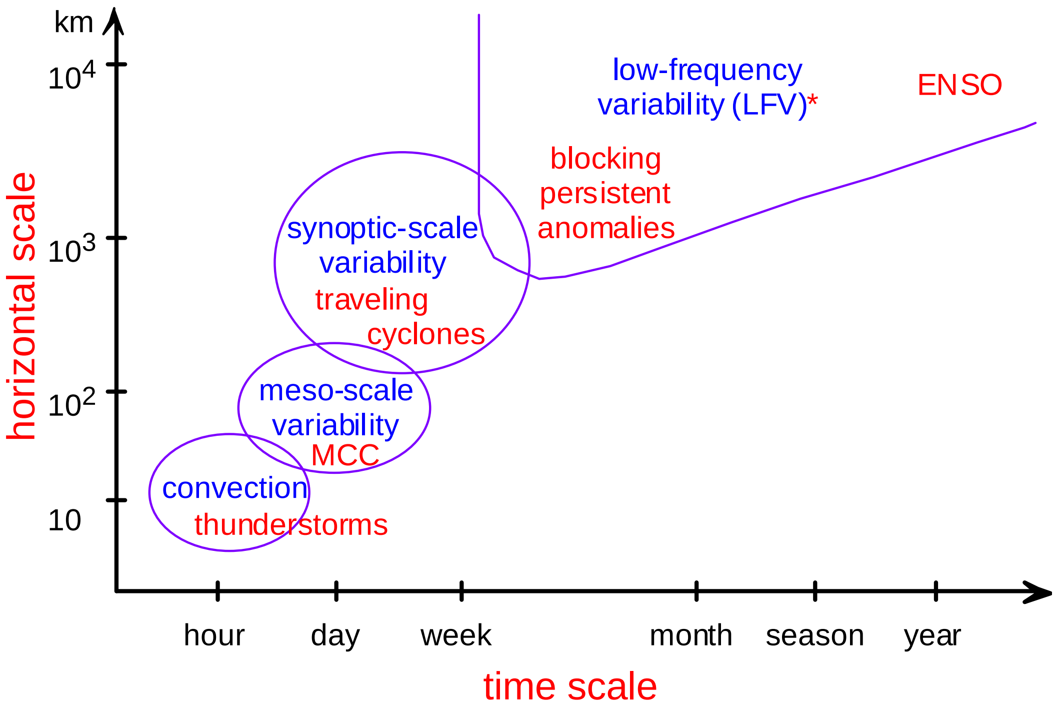

Figure 1. Characteristic time and spatial scales of observed atmospheric phenomena (loosely based on Fraedrich, K., and Boettger, H., 1978: Wavenumber-frequency analysis of 500 mb geopotential at 50-degrees North, J. Atmos. Sci., 35, 745-750).

Some recent work suggests that, instead of dividing the spectrum into resolved and unresolved scales, we should divide it into three parts: resolved, unresolved ('fast', stochastic, noise), and intermediate. The latter part may help mediate the interaction between the former two.

What are we trying to do?

- To improve the understanding and the prediction of weather and climate; and
- To describe the time evolution of the probability density function (PDF) of the system.

For simplicity, we refer to the unresolved processes also as 'processes in the little box'. It is worth exploring the behaviour in the little box with the full toolkit of dynamical systems. A schematic diagram of the possibilities appears in Fig. 2.

Boolean delay equations (BDEs) are a generalization of cellular automata (CAs) in which not all the delays between 'channels' are necessarily equal. They might be useful in dealing with sub-grid-scale processes with different characteristic times, such as cumulus clouds (fast) versus stratus cumulus (slow).

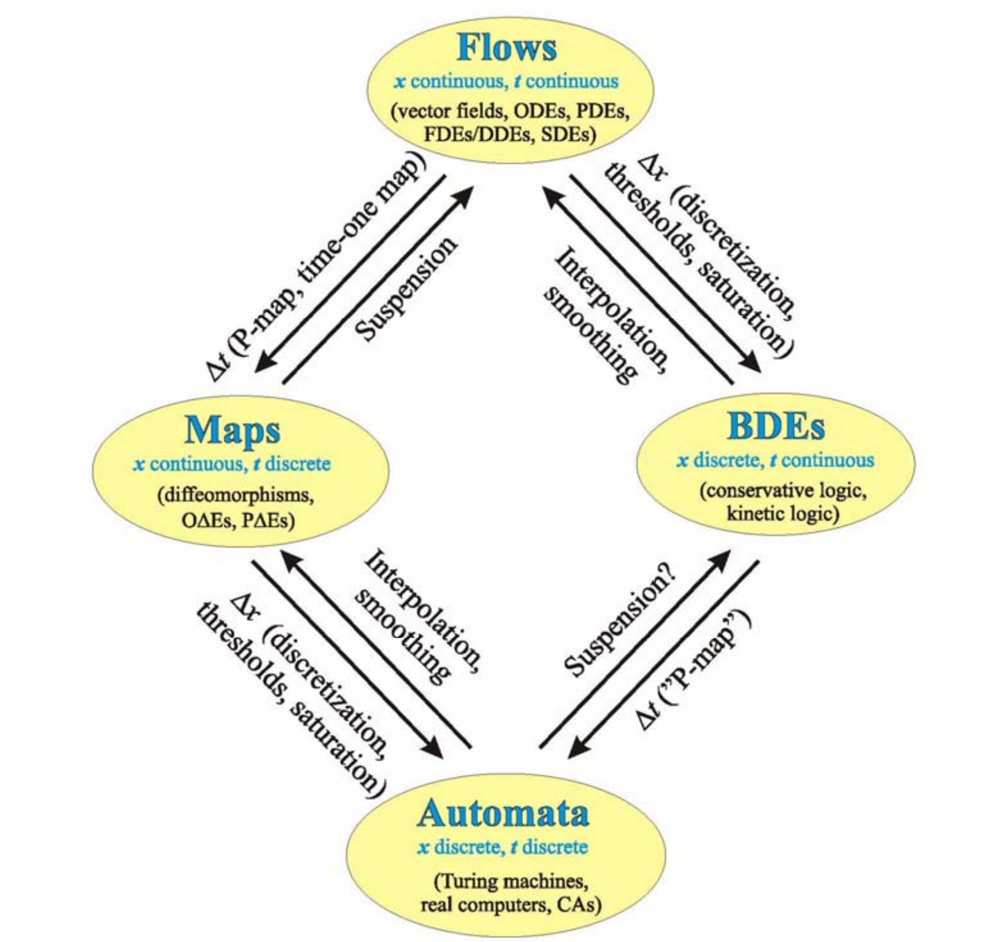

Figure 2. Schematic diagram of the four types of dynamical systems and the relations among them (after Zaliapin, I., Keilis-Borok, V., and Ghil, M., 2003: A Boolean delay equation model of colliding cascades. I: multiple seismic regimes. J. Stat. Phys., 111, 815-837).

Discrete-valued processes, like CAs or BDEs, are as valid as other parameterisation schemes. The former have shown some success in simulating the propagation of waves in the tropics, the latter in simulating ENSO and predicting earthquake clustering.

In investigating local phenomena, the choice of variables used to simulate stochastic processes is of paramount importance. In particular, should a stochastic variable be conditioned by the same large-scale variable, or by a different one, and if the latter, by which one? In other words, does the matrix B(X) in Eq. (1.c) have to be diagonal, and if not, what structure should it have?

The issue of stochastic parameterisation is closely intertwined with the more general issue of the representation of model error in the analysis problem. For instance, the analysis variables may include wind, temperature, humidity, or PV, divergence etc. Is this choice between primitive or derived variables helpful in the context of representing stochastic processes? Do all model variables need a stochastic element, and can we diagnose this element from analyses? How do we project information onto the variables that have been selected? Are non-local effects important, as in PV inversion with a non-local balance condition?

3.4. Specific issues

A non-exhaustive list of specific issues follows herewith.

3.4.1. Transport of chemical species

Material transport by sub-grid scale processes is a fertile ground for developing stochastic parameterisations. The transport of passive tracers, chemical components, and aerosols is among the processes that need to be parametrized. Both Lagrangian and Eulerian frameworks should be explored, using a hierarchy of models with different complexities.

3.4.2. Effect of stochastic parameterisations on ocean and coupled models

The ECMWF and several other major NWP centers are attempting to extend the range of their predictions to monthly, seasonal and interannual time scales. As the time scale of simulation and prediction is extended, the role of the oceans becomes more and more important.

The meso-scale oceanic eddies have significant dynamic and mixing effects on the large-scale flow and transport. The parameterisation of these effects is, therefore, an important issue. Mechanisms by which oceanic turbulence can affect coupled ocean–atmosphere variability also need to be explored. These issues are particularly relevant for the simulation and prediction of middle- and high-latitude climate.

3.5. How can we approach the problem, and learn how to solve it?

Four possible approaches present themselves to our attention:

- Explore fully residual tendencies, and the associated wave-number spectra
- Use data-assimilation techniques
- Understand communication between grid boxes
- Investigate the use of numerical methods with stochastic and Monte-Carlo elements

3.5.1. Exploring fully residual tendencies

As we saw in Sections 1 and 2, both deterministic and stochastic processes contribute to the variability of an NWP or climate model's resolved scales, and thus to the accuracy and spread of an ensemble forecast. Two complementary temptations should therefore be avoided: either attempting to correct for deficiencies in a deterministic parameterisation scheme by introducing stochastic processes, instead of correcting the parameterisation itself, or, vice versa, to vainly attempt to correct a deterministic parameterisation instead of introducing well-designed stochastic processes. Is there a way to know a-priori whether a stochastic process has to be included in the model in order to improve its prediction, or can this only be evaluated post-factum?

This workshop has addressed, in particular, the way NWP models have separated the resolved from the unresolved scales. Given, for instance, the ECMWF T511L60 model, how can we improve the representation of the unresolved scales? An important way of finding out is to carry out simulations at much higher resolution and study the difference between the two.

3.5.2. Data assimilation methods

The rich body of literature, and considerable practical experience with data assimilation and reanalysis may provide some useful indications on ways to formulate and calibrate stochastic parameterisations. Building on the current 4D-Var approach at ECMWF, should one have a penalty term to simulate stochastic model errors?

Doing so would entail, in addition to the usual J_o, J_b and J_c terms, including stochastic elements into a weak constraint within the data-assimilation system. For example, using the system defined in Eqs. (1a–c), should we include terms like $[\dot{X} - N(X) - M(X)]^2$ in the cost function?

Issues of this type appear to be especially important for data assimilation in a probabilistic, ensemble prediction framework. More generally, the systematic inclusion of stochastic processes into NWP models suggests a new look at sequential-estimation methods for data assimilation. The Kalman filter and related methods address, in the most natural way, the presence of system noise, including the estimation of noise parameters.

3.5.3. Communication between grid boxes and other implementation issues

Recent work on super-parameterisation schemes and on CAs highlights the importance for parameterisation schemes to allow for communication between grid boxes. The CA approach has demonstrated its ability to create structures in the sub-grid-scale realm that span several boxes. If super-parameterisation of cloud systems could extend to communicating between grid boxes, like CAs, it would provide more physically based tendencies than the latter, while possibly generating meso-scale structures as well.

Other issues worth mentioning are thresholds (or triggers) and memory. Triggers risk causing stochastic noise in the model, which is undesirable when a separate stochastic implementation is present. CAs do not have such triggers. If convection had memory, then it would provide much more realistic tendencies and probably represent the diurnal cycle better. This would be a good addition to a super-parameterisation.

3.5.4. Numerical methods for stochastic problems

There is a rich tradition of using stochastic processes in numerical methods. Examples include Monte-Carlo methods for the calculation of multi-variate integrals or finite-difference methods for various flow problems, such as the Glimm method for the Riemann problem in 1-dimensional gas dynamics, or the Chorin method for the Navier-Stokes equations. More recently, symplectic methods have gained popularity and may be useful in getting the momentum transport right. In these methods, the skew-symmetry and non-commutativity of the flows can be crucial in the accuracy of the drift term. Stochastic perturbations are associated, in this framework, with Arnold diffusion.

3.6. Recommendations to ECMWF

- The study and parameterisation of stochastic processes in NWP and climate models is an important area for ECMWF to pursue.
- Run higher-resolution model versions and compare their tendencies to lower-resolution ones.
- Investigate a hierarchy of models and approaches.
- Interact more closely with the academic community on issues discussed in this workshop

Running and comparing model simulations with a different number of resolved scales is likely to provide insights on how to best approximate sub-grid-scale processes. Research on optimizing the simulation and parameterisation of these sub-grid-scale processes should explore and test a hierarchy of models and approaches. The issues of how to combine stochastic with deterministic representations of weather and climate processes are at the frontier of the field. It is of the essence, therefore, to interact with the academic community— in applied mathematics, statistical physics, and the geosciences at large[1] — in exploring these issues, finding solutions, and implementing the proposed solutions into operational models.

[1] In the UK, for example, the Data Assimilation Research Centre (DARC) and the Universities Weather Research Network (UWERN) are actively engaged in projects that may be relevant to work at the Centre.