# Advanced Lectures 1988

## NWP - A forward look to the 1990s



...uarters, Bracknell

# Advanced Lectures 1988

## NWP - A forward look to the 1990s

M.J.P. Cullen

A.C. Lorenc

G.J. Shutts

Articles not accessed separately

# NWP - A forward look to the 1990s

# OPPORTUNITIES FOR DEVELOPMENTS IN NWP OVER THE NEXT 10 YEARS

**M J P Cullen**

**Met O 11, Meteorological Office, Bracknell**

## 1.1 INTRODUCTION

A first step in forecasting the likely developments in NWP over the next 10 years is to look at the last 10. Forecasts of the general synoptic pattern over the U.K. are now as good for 5 days ahead as they were for 3 days ahead 10 years ago. The use of a global model has led to forecasts being made for all parts of the world. The development of the mesoscale model has brought respectable attempts at forecasts of cloud and surface weather. Pressure for development of commercial activities has led to attempts to forecast a much wider range of quantities from models, supplemented where necessary by statistical post-processing. Developments in the military requirement are towards demands for more detailed local forecasts. All these developments have hinged on the use of higher resolution and greater areal coverage in forecast models, and on more complete representations of physical processes. These greatly increase the computer power needed to run forecasts, and have been made possible by the advances in computer technology. The available data has changed in character, with increased amounts of satellite data but at best maintenance of the conventional data network.

These trends are likely to continue over the next 10 years. Faster computers are developed or planned, and further increases in resolution and sophistication of models will be possible. Some new problems are likely to become more prominent as NWP becomes more ambitious and attempts to predict more detail of the atmospheric behaviour. The principles on which models have been built since the 1960's were designed for low resolution models with grid lengths of several hundred km. The separation between resolved and unresolved motions becomes much harder to formulate when the grid length is tens of km. Direct predictions of clouds will become more important, and may prove very difficult. More imaginative ways of making use of all available data will need to be found.

## 1.2 DEVELOPMENTS IN REQUIREMENTS

### 1.2.1 Global forecast requirements

The military requirements in the South Atlantic brought forward the use of a global model in the Meteorological Office. A major influence in the next 10 years will be the ACMAD centre for forecasting precipitation over Africa. Other existing requirements include forecasts for oil drilling in the Indian Ocean, West Pacific and off New Zealand, for predicting coffee growth in Brazil and cocoa growth in Africa, and for predicting tourist weather worldwide. The ship routeing requirement will contine. Another important activity is prediction of extreme events, such as tropical storms.

The present emphasis on developing the commercial side of the Office's activities will lead to demands for an even wider range of direct output products from the model, such as cloud and surface weather worldwide up to a week ahead. This will require improvements to the treatment of surface processes, as well as higher resolution modelling of the atmosphere.

A major new requirement will be forecasting of the ocean circulation for the Navy. This will require inpout of surface fluxes from the global NWP model.

### 1.2.2 British Isles area forecast requirements

The military requirements for forecasts for low flying in remote exercise areas and for detailed forecasts of local weather such as cloud and fog near bases and in land exercise areas will increase. More specialised requirements include forecasts of radio propagation. Forecasts will be required over Germany as well as the U.K. and possibly in other NATO exercise areas.

The need for numerical forecast products to support advisory services is growing rapidly. For example, predictions of conditions for crop diseases are now being made using model output. Some require forecasts for the U.K. up to a week ahead, and may have to be met by extended runs of local models driven by boundary conditions from the global model. Others involve very short range forecasting, where the need is to match the skill of nowcasting methods over the first three hours and extend the range of such detailed forecasts out to 24 hours. Meeting these requirements will involve further development of the mesoscale model.

## 1·3 OPPORTUNITIES GIVEN BY DEVELOPMENTS IN COMPUTERS

Fig. 1 shows the past trend in computing speed and the outlook for the next 10 years. Each point on the 'future' graph corresponds to a machine already at least being planned. By the mid 1990's a speedup of a factor of 40 over the ETA 10 just installed in the Office should be possible. This will allow greatly increased model resolution. Since the extra computing cost partly goes into taking shorter timesteps and thus only partly into storing more data, increases in storage should at least match the new requirements.

On such computers it will be possible to run a global model with a 35 km grid in the horizontal and 65 levels in the vertical, assuming that the present ratio of horizontal to vertical resolution is maintained. With this high a resolution, it is unlikely that the maintenance of a separate fine mesh model will be justified, and that the need for 'quick-look' runs with an early data cut-off will be met by extra (? hemispheric) runs of the global model. The physical processes in such a model will be chosen to give the best representation of the global climate, and the model will also be used with a lower resolution for climate integrations. Current estimates of the resulting likely improvements in forecasting the general synoptic pattern suggest that the skill of current 5 day forecasts should be extended to 6 or 7 days. The model will be used to generate detailed forecast products, such as cloud forecasts, out to this range. The improvements in these derived products should be large and will represent the main part of the case for acquiring the new computers. The forecasts provided by the existing mesoscale model give an idea of what will be possible, as described in Lecture 3.

The mesoscale model itself will be enhanced in areal coverage to include West Germany, and the total area will probably be 3 times the area proposed for the ETA 10. Cloud forecasts are perhaps the most important product, and will require much higher vertical resolution, perhaps 100 levels. Local forecasts are also very important, and a detailed representation of the topography is essential. One way of providing this is to use much higher horizontal resolution, perhaps 1 km, in the bottom 2 km of the atmosphere. Such extra resolution would probably only be used over land areas where products were required.

## 1.4 SCIENTIFIC DEVELOPMENTS

### 1.4.1 Use of data

Developments required in analysis techniques are described in the next lecture. The use of a 35 km grid in the global model does not, of course, mean that observations are required every 35 km. This resolution is required for accurate numerical treatment of larger scale disturbances. Even if the data coverage remains at its present level, there will still be improvements to the forecasts because the model will be able to 'remember' observations from previous times more accurately and provide a better first guess for the analysis. This will allow more effective automatic quality control of information, as well as easing the job of the analysis itself.

### 1.4.2 Representation of physical processes

At present the global model contains representations of large-scale rain, convection, surface processes (assuming climatological moisture availability), and radiation (based on climatological heating and cooling rates). It also includes parametrizations of the effects of sub-grid-scale dynamics such as gravity waves. Much effort has been put into replacing the 'climatological' parts of these schemes by actual or forecast values. It is likely that within the next few years all the methods used or being developed for climate integration will be used in the global forecast model. These include:

Soil temperature and hydrological model.

Effects of vegetation and land use on surface fluxes.

Explicit predictions of atmospheric aerosol, cloud water and ice.

Use of model cloud and atmospheric composition predictions in radiation calculations.

Many of the above processes are already represnted in the mesoscale model. Further developments in the mesoscale model will continue to be specialised to the needs of forecasting near the U.K., rather than the need to represent the global climate. Representations of cloud properties will continue to be based on observed studies of clouds near the U.K. The use of extra resolution near the ground will require much more detailed reprentation of factors important in predicting surface temperature, visibility, dew, frost and fog.

### 1·4·3 Representation of atmospheric dynamics

Present models use space and time-averaged versions of the fundamental equations of motion. The effects on unresolved scales are taken into account be a 'sub-grid' model. This model must also maintain smoothness of the solution on the computing grid, so that the resulting problem can be accurately solved numerically. Fig. 2 shows a diagram due to J. Smagorinsky of the typical space and time scales of atmospheric motions. The current global model can treat wavelengths of over 600 km and periods of more than 12 hours accurately. It is seen that filtering on these scales will include only motions associated with synoptic meteorology. The mesoscale model treats wavelengths of over 60 km and periods of more than 1 hour accurately. Filtering on these scales includes a much wider variety of motions.

In the future the global model will also be able to resolve partially a wider class of motions. Though the extra resolution will increase the accuracy of treatment of synoptic weather systems and their associated structure such as fronts, it will also result in partial and probably inaccurate treatment of gravity waves. This information may contaminate the input to the parametrization schemes and make the derivation of forecast products more difficult. A method of selectively filtering unwanted classes of motion without smoothing other motions will have to be developed. Such work is described in lectures 4 and 5.

The use of dynamical filtering may give an opportunity to enhance predictions of the synoptic pattern by more than the day or two suggested above. Predictability is limited by lack of knowledge of the initial state and by the interaction of unresolved processes with resolved processes. The use of simple space-time averaging may not be the best way to reduce the impact of the second limitation. If a subset of the dynamics which is only weakly coupled to other motions can be identified, then the use of dynamical filtering based on this subset should enhance predictability. Any such improvement will lead to greater ability of the model to remember data, and therefore a more accurate initial state for the forecasts. Theoretical predictability studies mostly suggest a 2-week limit for synoptic prediction rather than the one week which is our current target.

## BIBLIOGRAPHY

Further reading on the topics introduced above is given under the subsequent lectures, which treat them in more detail.

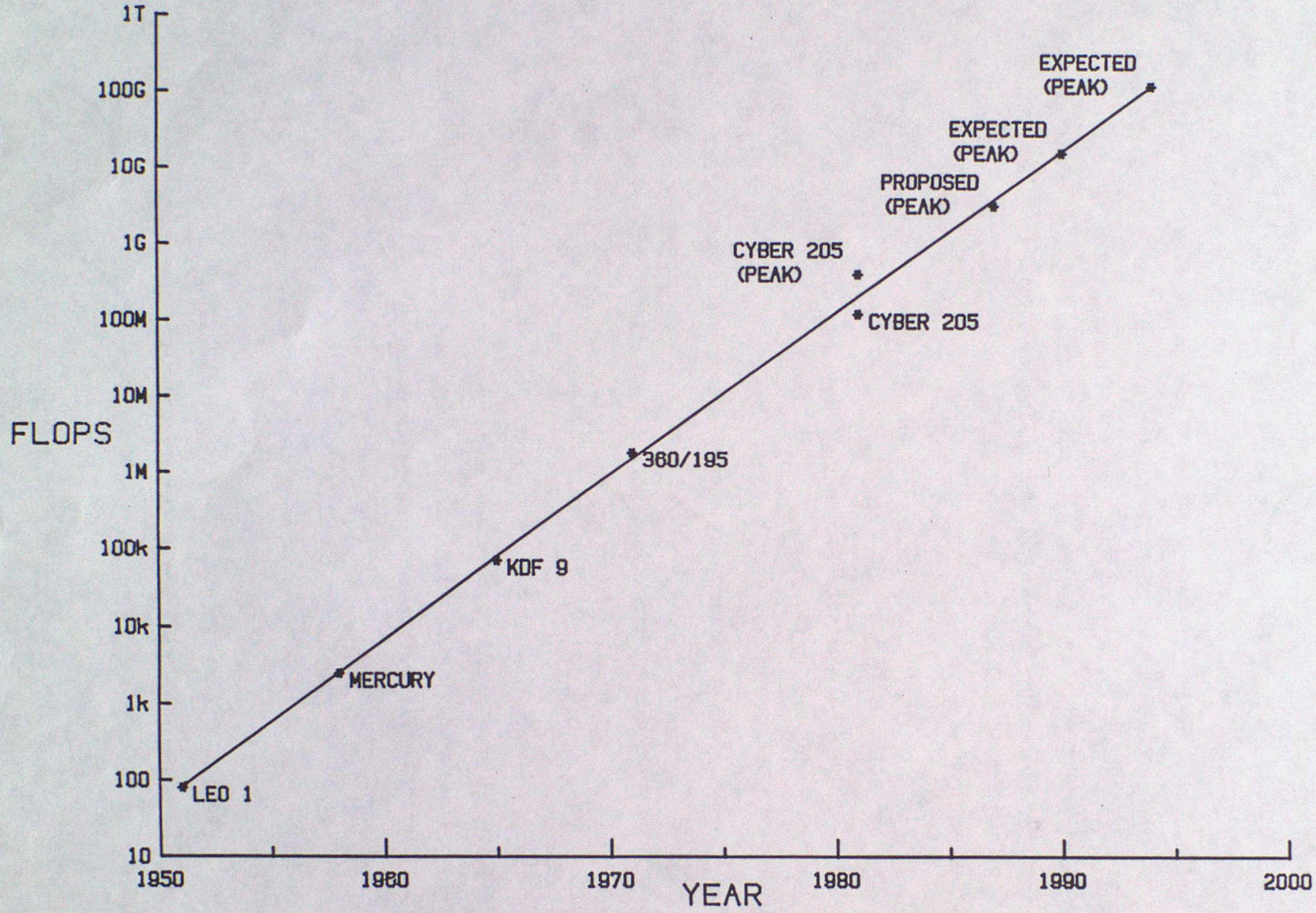

FLOPS
1T
100G
10G
1G
100M
10M
1M
100k
10k
1k
100
10
1950
1960
1970
1980
1990
2000
YEAR
LEO 1
MERCURY
KDF 9
360/195
CYBER 205
CYBER 205
(PEAK)
PROPOSED
(PEAK)
EXPECTED
(PEAK)
EXPECTED
(PEAK)

Figure 1

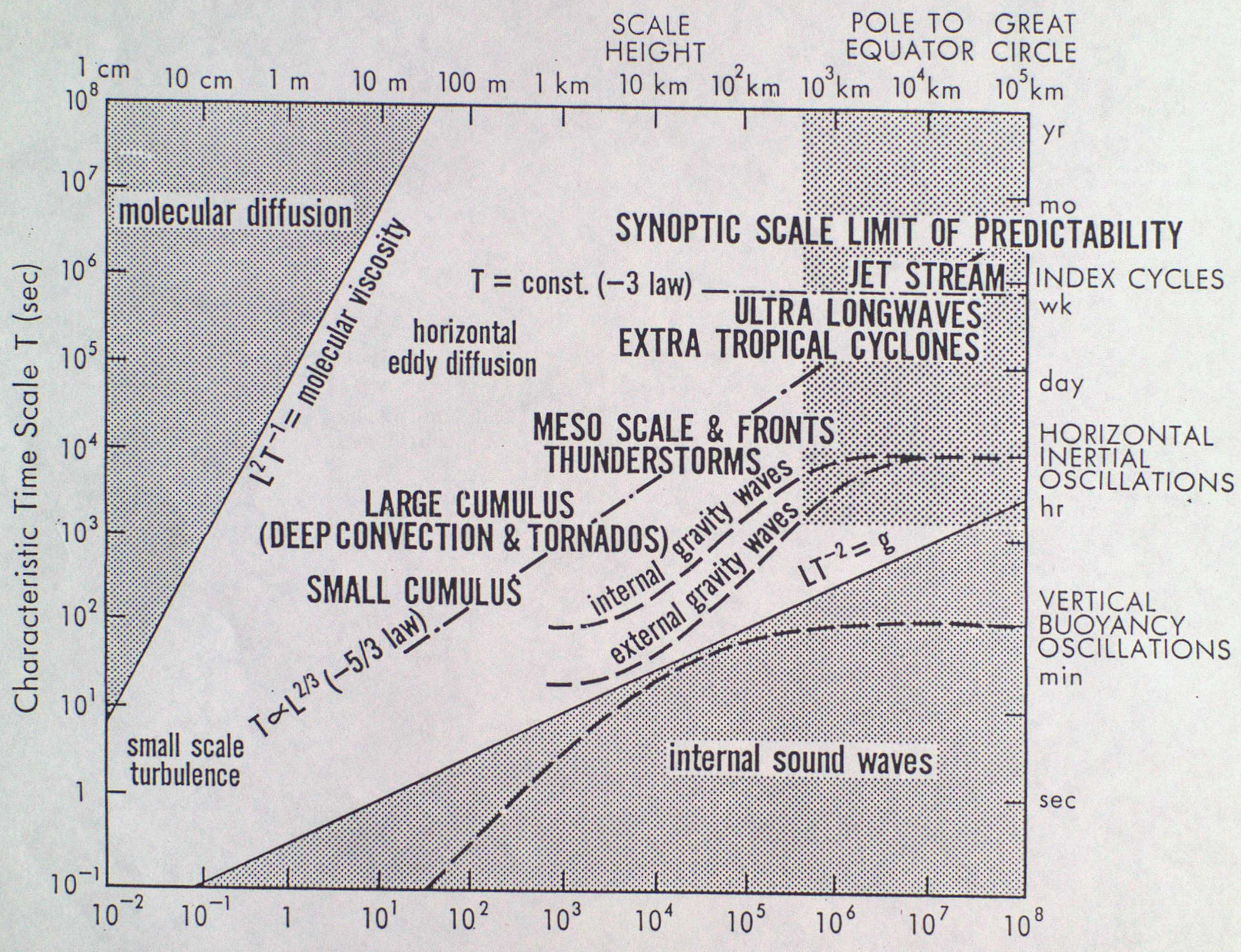

SCALE HEIGHT
POLE TO EQUATOR
GREAT CIRCLE
1 cm
10 cm
1 m
10 m
100 m
1 km
10 km
$10^2$km
$10^3$km
$10^4$km
$10^5$km
Characteristic Time Scale T (sec)
$10^8$
$10^7$
$10^6$
$10^5$
$10^4$
$10^3$
$10^2$
$10^1$
1
$10^{-1}$
$10^{-2}$
$10^{-1}$
1
$10^1$
$10^2$
$10^3$
$10^4$
$10^5$
$10^6$
$10^7$
$10^8$
yr
mo
INDEX CYCLES
wk
day
HORIZONTAL INERTIAL OSCILLATIONS
hr
VERTICAL BUOYANCY OSCILLATIONS
min
sec
molecular diffusion
$L^2T^{-1}$ = molecular viscosity
SYNOPTIC SCALE LIMIT OF PREDICTABILITY
T = const. (−3 law)
JET STREAM
ULTRA LONGWAVES
EXTRA TROPICAL CYCLONES
horizontal eddy diffusion
MESO SCALE & FRONTS
THUNDERSTORMS
LARGE CUMULUS
(DEEP CONVECTION & TORNADOS)
SMALL CUMULUS
internal gravity waves
external gravity waves
$LT^{-2} = g$
$T \propto L^{2/3}$ (−5/3 law)
small scale turbulence
internal sound waves

Figure 2

# NEW METHODS OF USING DATA FOR NWP

A C Lorenc

Met O 11, Meteorological Office, Bracknell

## 2.1 INTRODUCTION

The amount of data available from remote sensing instruments, the processing power of computers, and the initial data requirements of NWP forecasts are all increasing rapidly. We cannot be sure which will be more rapid; my personal views are:

(a) *NWP forecast models will expand to fill the available computers.*

(b) *We will never have enough observations.*

Because of (a), design of an analysis scheme for NWP will have to continue to compromise between the theoretically optimal and the practically achievable. (b) is saying that the model's initial data requirements will remain underdetermined by the observations alone. It has long been recognized that because of this *one has to utilize the physical relationships between the quantities. Even to construct a coherent picture of the total state of the atmosphere out of scattered observations, one has to use, to a large extent, dynamical-physical methods.* (Bjerknes, 1904).

Equations and notation in this lecture will follow those in a survey of analysis methods for NWP published by Lorenc (1986), hereafter refered to as L86. One of the themes of L86 was that most analysis methods can be thought of as approximations to an optimal analysis equation. Differences in terminology and generality of definition of 'optimal' can obscure this, as can the different names and buzz-words used. In section 2 I set out and discuss the 'optimal' equation, and its implications. Of course to do this without long-winded qualifications, I use my own prefered terminology, and apply my own understanding of 'optimal'. In section 3 I discuss some of the methods currently under development, trying to relate them to section 2. Finally in section 4 I mention briefly a few topics which warrant fuller discussion than can be given in this one lecture.

## 2.2 DERIVATION OF BASIC ANALYSIS EQUATION

### 2.2.1 Optimal analysis

Our objective in analysis for NWP is to find initial conditions for our forecast model which fit the observations, and which are consistent with our prior knowledge as to the behavior of the atmosphere. (We might also have requirements imposed by characteristics of the NWP model. Procedures to satisfy these I class as initialisation techniques. As the verisimilitude of the NWP model increases, the need for initialization of a good analysis should decrease.) We need to find a single specification of optimal initial conditions, whereas our knowledge in fact only gives us a range of possibilities. We can use Bayesian arguments to find the most likely. This can be done mathematically by a variational minimization of a penalty function which is the sum of penalties measuring the fit to the observations, and the consistency with our prior knowledge. If the error distributions are Gaussian, then these penalties are quadratic, and we are finding the minimum variance solution.

### 2.2.2 Fit to observations

In order to compare an analysis specification, or model state ($\mathbf{x}$) with the observations ($\mathbf{y}_o$), to measure the fit between them, we need to be able to convert from one to the other. Where a direct conversion is possible, it is simpler to convert the observations towards the atmospheric parameters used to define a model state. We assume that this has already been done in $\mathbf{y}$. However there remain some observations, particularly from remote sensing, for which a unique conversion is impossible. Since our NWP models are becoming more realistic and comprehensive in their representation of the atmosphere, the transformation ($\boldsymbol{K}$) from $\mathbf{x}$ to $\mathbf{y}$ is in most cases more nearly unique. If the probability distribution functions of observational errors, and of errors in $\boldsymbol{K}$, are approximately Gaussian, then the fit to the observations is measured by: $(\mathbf{y}_o-\boldsymbol{K}(\mathbf{x}))^*(\mathrm{O}+\mathrm{F})^{-1}(\mathbf{y}_o-\boldsymbol{K}(\mathbf{x}))$. $\boldsymbol{K}$ is the forward process, or generalized interpolation from the model to the observations. The analysis problem can be thought of as finding the generalized inverse of $\boldsymbol{K}$, to interpolate from observations to model.

### 2.2.3 Fit to background

If we had no observations, we would presumably still have a best estimate of the most likely model state, we call this the background $\mathbf{x}_b$. This will of course be consistent with our prior knowledge about atmospheric structures. This prior knowledge will also tell us that some modes of perturbation about $\mathbf{x}_b$ are more likely than others. If this prior knowledge is sufficiently linear and Gaussian, deviations from our prior knowledge can be measured by: $(\mathbf{x}_b-\mathbf{x})^*\mathrm{B}^{-1}(\mathbf{x}_b-\mathbf{x})$.

Our basic optimal analysis problem is thus simply to minimize a penalty functional (J) given by the sum of these two terms:

$$\mathrm{J} = (\mathbf{y}_o - K(\mathbf{x}))^*(\mathrm{O}+\mathrm{F})^{-1}(\mathbf{y}_o - K(\mathbf{x})) + (\mathbf{x}_b-\mathbf{x})^*\mathrm{B}^{-1}(\mathbf{x}_b-\mathbf{x}) \qquad (2.1)$$

### 2.2.4 Generality of equation

Note that I have been deliberately vague about the analysis specification, or model state $\mathbf{x}$. In its simplest traditional form it is a gridpoint field. For multivariate three-dimensional analysis it is many fields. It does not have to be represented by gridpoints (eg spectral representation is possible). It does not have to be for a single time. In each of these cases the analysis equation remains the same, although of course the meanings of $\mathbf{x}$, $K$, and B will change.

The time aspect is worth discussing more fully. Since observations are distributed in time, one way of helping reduce the underdeterminacy is to use a four-dimensional distribution of them. This alone does not help, since we have to make $\mathbf{x}$ four-dimensional as well. However we have prior knowledge about the way the atmosphere behaves; this enables us to say that most four-dimensional states are very unlikely. If we use the NWP model itself in some way to express our knowledge of atmospheric evolution, then the analysis technique is called <u>four-dimensional data assimilation</u>.

## 2.3 ANALYSIS METHODS AND APPROXIMATIONS

### 2.3.1 OI

If the error probability functions are Gaussian, or nearly so, then **O**, **F**, and B can be represented by matrices which are independent of $\mathbf{x}$ (although they might be functions of $\mathbf{y}_o$ and $\mathbf{x}_b$). If $K$ is approximately linear, its variations around $\mathbf{x}_b$ can be represented by:

$$K(\mathbf{x}) \simeq K(\mathbf{x}_b) + \mathrm{K}(\mathbf{x}-\mathbf{x}_b)$$

where K is similarly a matrix which is independent of $\mathbf{x}$. Then an explicit solution to the analysis equation can be derived (L86, eqn 28):

$$\mathbf{x}_a = \mathbf{x}_b + \mathrm{BK}^*(\mathrm{KBK}^*+\mathrm{O}+\mathrm{F})^{-1}(\mathbf{y}_o - K(\mathbf{x}_b)) \qquad (2.2)$$

The behavior of the elements of B can be modelled by a continuous covariance function, and this same function can be used to evaluate the interpolated covariances implied by multiplication by **K**. Then $\mathbf{BK}^*$ and $\mathbf{KBK}^*$ can be evaluated directly, and (2.2) is the OI equation. Note that for it to be optimal (apart from the approximations already made), then all the relevant observations must be used simultaneously in $\mathbf{y}_o$, giving a large computational task, proportional to $N_y{}^3$, in solving (2.2). If the number of data $N_y$ used in the analysis of one region is restricted, then the data selection algorithm is all important, and the analysis is not necessarily mathematically optimal. If $N_y$ is large (The ECMWF operational analysis scheme already solves systems with several hundred data; no doubt this will increase further in the future), then the covariance function model is the key of the behavior of the system. Much work has been done in this field, but much remains to be done in the next few years. For instance the dependence of the function on $\mathbf{x}_b$, both for its meteorological situation and for its previous observational support, needs refining. Covariance models suitable for the mesoscale need to be derived.

### 2.3.2 Iterative solutions

There are several iterative solutions to the problem of finding the **x** which minimizes (2.1). Iterative methods of finding the minimum of a functional are called descent algorithms. Advantages of iterative methods are:

(a) They can be computationally cheaper.

(b) The need for data selection algorithms can be avoided.

(c) Nonlinearities in *K*, and non-Gaussian penalties can be treated. (The latter can be done by allowing **B** and **O** to be functions of the latest best estimate of **x**, rather than of $\mathbf{x}_b$.)

(d) Special structure in **K** and **B** can be used to implement multiplication by them, rather than the semi-empirical covariance function modelling used in OI.

(e) If the algorithm is carefully designed, we can make a virtue out of the slow convergence of some cases; often these cases are those where the observations were in error.

The [u+1]th iteration of one method is given by:

$$\mathbf{x}[u+1]=\mathbf{x}[u]+\mathbf{Q}\left\{\mathbf{W}\left(\mathbf{y}_o-K(\mathbf{x}[u])\right)+\mathbf{x}_b-\mathbf{x}[u]\right\} \qquad (2.3)$$

where $\mathbf{W}$ is a $N_x \times N_y$ matrix of weights, and $\mathbf{Q}$ is a $N_x \times N_x$ matrix of normalization factors. The optimal values for these, derived using a Newton descent algorithm for (2.1), are:

$$\mathbf{W} = \mathbf{BK}^*(\mathbf{O}+\mathbf{F})^{-1} \tag{2.4}$$

$$\mathbf{Q} = (\mathbf{WK}+\mathbf{I})^{-1} \tag{2.5}$$

Three approximations convert this iteration, which as it stands should converge to the optimal solution, into the <u>successive correction method</u>:

(a) As in OI, we model $\mathbf{BK}^*$ by a continuous covariance function, so that (2.4) gives us weights which depend on the distance between observations and grid points and on the observational errors.

(b) We approximate $\mathbf{Q}$ by the reciprocal of the sum of weights at each grid-point, rather than the $N_x \times N_x$ matrix inverse of (2.5). It is helpful to think of $\mathbf{Q}$ as a function of the data-density; this approximation makes it a local function at each grid-point. Note that $\mathbf{Q}$ only affects the rate of convergence of the iteration, not the final limit.

(c) We start the iteration from $\mathbf{x}_b$, and only continue for a finite number of iterations. In most cases then, $\mathbf{x}$ will not differ too much from $\mathbf{x}_b$, and the $\mathbf{x}_b-\mathbf{x}[u]$ term can be neglected.

(2.3) can be manipulated to avoid the need for the grid-point normalization of the forcing towards the observations, replacing it by a $N_y \times N_y$ matrix $\mathbf{Q}'$:

$$\mathbf{x}[u+1] = \mathbf{x}[u] + \mathbf{WQ}'\left(\mathbf{y}_o - K(\mathbf{x}[u])\right) + (\mathbf{WK}+\mathbf{I})^{-1}\left(\mathbf{x}_b-\mathbf{x}[u]\right) \tag{2.6}$$

$$\mathbf{Q}' = (\mathbf{KW}+\mathbf{I})^{-1} \tag{2.7}$$

We are still left with a normalization on the forcing towards the background; this is avoided however if we make approximation (c) above. There are indications that this iteration in its approximated form is better at handling variations in data-density; it is the form used for the Meteorological Office's <u>analysis correction</u> method. Alternatively, the background forcing can be incorporated into the observation term, giving an iteration which does converge towards the optimum:

$$\mathbf{x}[u+1]=\mathbf{x}[u]+\mathbf{WQ}'\left(\mathbf{y}[u]-K(\mathbf{x}[u])\right) \tag{2.8}$$

$$\mathbf{y}[u+1]=\mathbf{y}[u]+\mathbf{Q}'\left(\mathbf{y}[u]-K(\mathbf{x}[u])\right) \tag{2.9}$$

As in the successive correction method, any approximations made in $\mathbf{Q}'$ only affect the rate of convergence, not the final limit.

If we do not make the OI approximation, of modelling $\mathbf{BK^*}$ by a continuous covariance function, then other iterative methods can be derived, e.g.:

$$\mathbf{x}[u+1] = \mathbf{x}[u] + \mathbf{Q}\ \{\mathbf{BK^*}(\mathbf{O}+\mathbf{F})^{-1}(\mathbf{y}_o - K(\mathbf{x}[u])) + \mathbf{x}_b - \mathbf{x}[u]\} \tag{2.10}$$

$\mathbf{K^*}$ is the adjoint of $K$. If $K$ is simply an interpolation from grid points to observation positions, then its adjoint is simple. $\mathbf{B}$ is a matrix specifying our prior knowledge about likely modes of variation about $\mathbf{x}_b$. If this can be summarized as *smooth* and *balanced*, then multiplication by $\mathbf{B}$ can be approximated by a filter of rough and unbalanced increments. This can be constructed by using spectral model spherical harmonics, or by using normal modes.

The Newton descent algorithm, leaves us with some large computations in order to evaluate $\mathbf{Q}$:

$$\mathbf{Q} = (\mathbf{BK^*}(\mathbf{O}+\mathbf{F})^{-1}\mathbf{K}+\mathbf{I})^{-1} \tag{2.11}$$

However other numerical algorithms, such as the <u>conjugate gradient</u> method, enable us to find the minimum of (2.1) efficiently without knowing an accurate $\mathbf{Q}$. This iterative method is used in the <u>adjoint</u> method of four-dimensional data assimilation, described below.

### 2.3.3 <u>Four-dimensional data assimilation</u>

Since forecast models are not perfect, the constraint imposed on our four-dimensional analysis, that its time evolution should be consistent with the model's, should not be strictly enforced. Practical implementation of a optimal four-dimensional analysis scheme with such a constraint is severely limited by available computer resources. The storage and manipulation of high-resolution four-dimensional fields requires many more resources than running a NWP model, which only manipulates three-dimensional fields. The traditional NWP method of approximating four-dimensional data assimilation is the analysis-forecast cycle, in which observations are inserted using a purely three-dimensional analysis procedure into a background forecast from the results of the previous analysis. In this case our $\mathbf{x}_b$ summarises the information from all previous observations; it often contains more information than the current observations. This method does allow for defficiencies in our forecast model, since errors in $\mathbf{x}_b$ are represented by $\mathbf{B}$. However the approximations used to model $\mathbf{B}$, described below, and the fact that observations from different times are not analysed together, mean that the method is not optimal in its use of the four-dimensions information from the observations.

If we ignore imperfections in our forecast models, and impose them as a strong constraint, then many of the computational difficulties of a four-dimensional analysis can be avoided, since all possible four-dimensional fields are defined by their initial three-dimensional fields and the NWP model. We can thus reduce the analysis problem to that of finding the best initial state. $\boldsymbol{K}$ is extended to included the forecast from this initial state to the observation times, as well the interpolations in space to the observation positions. If $\boldsymbol{K}$ is linear, and matrix B is sufficiently small that it can be manipulated, then the optimal final state can be found explicitly, using a Kalman filter. Unfortunately it seems that, for NWP, B is too complex to make reasonable approximations to this feasible. However, since we can run the NWP model which forms the major component of $\boldsymbol{K}$, it is not too difficult also to run its adjoint, and evaluate explicitly $\mathbf{K}^*$. Since we are including more observations in the analysis, the background information is less important, and we can approximate (or even ignore) the background penalty. Then an iterative solution to the analysis equation (2.1) becomes feasible. However it is computationally expensive; each iteration requires an integration of the NWP model, the storage of all its results, and and integration of its adjoint. Thus the full analysis will consume at least an order of magnitude more computer resources than the NWP model used. If, as I suggested in the introduction, the operational NWP forecast model uses most of the available computer power, then a lower resolution version will have to be used for the operational analysis using this method.

## 2.4 SOME RELATED TOPICS

### 2.4.1 Quality control of observations

The Bayesian methods used to derive (2.1) can be used for observations which have a non-negligable probability of gross error; the probability distribution function becomes non-Gaussian and the penalty function non-quadratic. This invalidates explicit solutions such as OI. Iterative descent algorithms can also be unreliable, since there may be multiple local minima to J. Quality control thus needs a prior decision algorithm, before the analysis, in order to make J more nearly quadratic. These decisions can be taken objectively if we use the Bayesian formulation, and if we have an accumulated statistical knowledge about the distributions of errors for good observations and for those with gross errors.

### 2.4.2 Processing of satellite data

The equations derived in §2.2 and §2.3 are identical to those used in many inversion methods for satellite temperature retrievals. Many of the quality control problems can be treated by the techniques discussed in §2.4.1. Because of the complicated nonlinear form of $K$ for most satellite remote sensing data, we can say that in principle *the processing of remote sensing data for NWP should be integrate with the NWP analysis.*

### 2.4.3 Interactive, Man-machine analysis systems

The human is better than the computer at recognising complex structures, particularly in images, in term of prior conceptual models. If the information from such images is an important part of the input to the analysis, then an interactive analysis is indicated. The Meteorological Office has interactive systems in its *Frontiers* system, for processing radar images of rainfall, and in its *Mesoscale model* system, mainly for cloud images.

If the analysis input information comes from a large number of disparate sources, and has four-dimensional and balance constraints which are expressed in term of the NWP model, then the analysis system must use the full power of the computer, and results in an internally consistent set of initial conditions, for the NWP forecast which is to follow on the same computer. Man-machine interaction then has two bottlenecks to overcome:

(a) The human has to inspect the full multi-dimensional analysis fields, and understand their observation support through the analysis algorithm. He can then apply his better conceptual models, and his inspection of the imagery information inadequately used by the automatic analysis, to propose modification to the analysis fields.

(b) These modification need to be applied to the fields, while retaining the internal consistence and balance built in to the automatic system.

I remain unconvinced that development and routine running of such a system is a cost-effective use of manpower, for a large operational NWP system.

## BIBLIOGRAPHY

I have avoided giving any references in this lecture note; nearly all of those that I might have given can be found in my review article (L86):

Lorenc,A.C. 1986 'Analysis methods for numerical weather prediction'. *Quart. J. Roy. Met. Soc.*, 112, 1177-1194

# THE USE OF VERY HIGH RESOLUTION MODELS IN WEATHER FORECASTING

G J Shutts

Met O 11, Meteorological Office, Bracknell

## 3.1 INTRODUCTION

Noticeable improvements in the quality of weather forecasts over the last 15 years have resulted primarily from increases in the horizontal and vertical resolution of primitive equation models in combination with better analysis techniques and physical parametrizations. Rapid cyclogenesis may often be anticipated days in advance whereas the coarser resolution models of the Seventies rarely showed such skill. Indeed it has been the experience of US forecasters that major cyclonic events such the President's Day Storm could not be forecast adequately until the recent introduction of the NMC nested-grid model with inner gridlength of ≈90 km and improved boundary layer and other physical parametrizations (Sanders, 1987). Various sections of the research community had been under the impression that some 'missing physical ingredient' (such as a failure to deal with convective processes adequately) was responsible for the poor forecasts of rapid development. It is now recognised that this is not the case and that models comparable in resolution to the current Fine-Mesh have what it takes to forecast extremely deep depressions. This is not to say that the sub-synoptic detail is correct or that right balance between explicit and parametrized motions has been achieved.

Studies carried out at ECMWF have shown that horizontal gridlengths of < 100 km combined with an improved description of the orographic height profile are essential in the simulation of Alpine cyclogenesis. ECMWF have also recently carried out a ten day forecast with a T213 version of their global spectral model and found distinct benefits in the representation of a Mediterranean lee vortex, large-scale cloud patterns and local orographically-controlled winds (ECMWF Newsletter March 1988). Our own experience of the benefits and pitfalls of high resolution is derived from the Fine-Mesh model (at its operational resolution ≈ 75 km and a research version ≈ 37.5 km) , the ELF (Equatorial Latitude/longitude Fine-Mesh which has its poles displaced from the geographical poles so as to achieve an equatorial grid-point distribution in the region of interest - thereby minimising the convergence of meridians effect) run at 40km and the Mesoscale model which has a gridlength of 15 km. The skill of the operational Fine-Mesh model in handling explosive development gives some confidence in the quality of future global forecasts to be made on the ETA 10.

Our current global modal has had many notable successes in the forecasting of hurricanes and typhoons. There is plenty of evidence that models with horizontal resolution approaching ≃ 50 km are quite realistic in handling these tropical systems (eg. Dell' Osso and Bengtsson, 1985) - a rather surprising fact given that the diameter of hurricane ' eyes ' is frequently smaller than this gridlength. Much higher vertical resolution will help in the description of boundary layer processes and in the representation of thin stratocumulus layers though only if better parametrization of microphysical, turbulent and radiative processes can be achieved.

## 3.2 RESOLUTION, EDDY DIFFUSION AND BALANCE

It is worth considering briefly what aspects of the flow need to be better resolved in order to improve forecast dynamical evolution. Since the dominant horizontal scale of baroclinic disturbances (Rossby radius of deformation) is ≃ 1000 km , a gridlength of 200 km resolves the broad features of weather systems quite well. The effective Rossby radius of deformation in an explosively deepening baroclinic wave (in which the slantwise stability tends to zero) may be one half of the ' dry' value quoted above and a gridlength of 100 km or less appears to be sufficient to capture this type of development. Whether or not this zero slantwise stability limit is approached, much of the dynamical action in real weather systems is concentrated in frontal zones (mainly the cold front) on a scale of a few hundred kilometres and would require gridlengths of less than 40 km to be considered well-resolved. One consequence of insufficient resolution in fronts is that grid-scale smoothing by eddy diffusion consumes a lot more energy than it would otherwise do and with a gridlength of 200 km dissipation levels are likely to be a gross overestimate (note that the true level of kinetic energy dissipation within frontal zones is not actually known in any climatological sense).

In the baroclinic wave life-cycle experiments of Simmons and Hoskins (1978), *one half* of the available potential energy decrease over the course of their experiment was dissipated through eddy diffusion terms rather than released as kinetic energy. One would hope that as higher resolutions are achieved the rate of energy dissipation by eddy diffusion might approach some observable magnitude. At this point the diffusion (mainly in the vertical) would correspond to a real turbulent mixing process occurring in fronts rather than a numerical device for maintaining smooth fields. In practice, eddy diffusion is also required to damp out gravity waves produced through geostrophic adjustment at the scale of a few gridlengths : herein lies the principal snag with resolving mesoscale weather systems. Regions of thermal wind imbalance created by the intermittent modification of temperature in gridpoint columns through convective parametrization; by truncation errors in the vicinity of near-discontinuous flow structures (eg. fronts and inversions) and by flow over, rather than around, steep orography all tend to excite excessive gravity waves - diffusion is essential to remove them. We shall elaborate on these mechanisms in 3.3 and in Lecture 4.

One way to avoid the difficulties associated with spurious gravity waves is to integrate a dynamically-filtered set of equations, such as the semi-geostrophic equations, which do not support such motion types. These equations are well-known for their ability to form frontal discontinuities . The price to pay for meteorologically noise-free equations is *some loss of accuracy* through using the geostrophic wind instead of the actual wind in certain terms . But do the primitive equations of motion actually give us much more accuracy than integrating a balanced set in practice? Dr Cullen proposes that the generality of the primitive equations may be a penalty when gridscale motions of high Rossby number ( ie. >>1) are permitted. Some support for this suggestion now follows.

## 3.3 POTENTIAL SOURCES OF SPURIOUS GRAVITY WAVES

### 3.3.1 Parametrization of convection

Although convective updraughts are of small horizontal extent ( < 1 km) compared to current operational model gridlengths, they only represent a small part of convective circulations. Outflow anvils may spread over tens (or even hundreds) of kilometres associated with individual storms or in mesoscale convective complexes - particularly in the tropics. These are just the visible manifestation of a cloud-scale mass adjustment process which, for a single convective cell, takes place over a horizontal distance which scales as $(CAPE)^{1/2}/f$

(Shutts,1987). Most of the environmental adjustment is invisible and takes place over the scale of the Rossby radius of deformation based on the depth of convection. The problem for modelling convection begins at resolutions for which it is inappropriate to assume an ensemble of convective elements. At gridlengths of 20 km, the convective updraught (and downdraught) are still sub-grid scale yet the mass adjustment scale is bordering on the resolvable. A standard convective parametrization scheme which computes temperature and mixing ratio increments in a single column would not, at this resolution, give the correct spatial distribution of environmental modification associated with a single convective cell. The model is then forced to carry out the rest of the adjustment process with an extremely truncated set of internal gravity modes.

To demonstrate this problem we show results from three simulations of moist frontogenesis with a two-dimensional primitive equation model using horizontal gridlengths of 80, 40 and 20km (see Holt, 1985 for more detail). Figures 3.1 (a) and (b) show the initial potential temperature and humidity fields respectively. The moist boundary layer air on the warm side of the front is associated with CAPE values of about 290 J/Kg and frontogenesis is forced by pure deformation. The vertical velocity fields obtained after 16000 seconds for all three simulations are shown in Figs. 3.2 (a) - (c). Gravity wave motion excited by the switching on of the convection scheme causes a progressively more noisy vertical velocity field as the resolution increases. At the 20 km gridlength , unpleasant interactions between the convective parametrization scheme and the explicit dynamics are occurring as the upward motion due to the gravity wave destabilizes some columns and triggers convection. A considerably less noisy vertical velocity field results if the convective parametrization scheme is switched off though one should not infer that the dynamics are better handled. As mentioned earlier high levels of diffusion required to suppress this type of response have a detrimental effect on energy transformations and the ability of models to resolve near-discontinuous flow structures.

### 3.3.2 Flow splitting and the mountain barrier effect

The earth's orography clearly contains far more detailed structure than can be reasonably resolved in even the highest resolution models currently available. It is necessary to smooth out terrain height variations at the gridscale to avoid 'steps' - thereby misrepresenting the true barrier effect of , for example, the Alps and the Greenland plateau. The remaining sub-gridscale orographic height variation causes a diverse range of motions whose effect is parametrized as boundary layer friction and gravity wave drag.

An airstream with strong low-level static stability impinging on the Alps is usually *blocked* and air flows in a low-level jetstream around the mountains. In the Global and Fine-Mesh model there appears to be a tendency for air to go over the Alps. This often leads to a gravity wave train of unrealistic large amplitude and horizontal wavelength (Fig.3.3 , Cullen and Parrett, 1987). The steepness of real orographic barriers causes air below the mountain top height to be blocked and turn to the left under the influence of the Coriolis force. Air motion above is only slightly perturbed and flows over the mountain exciting some gravity wave motion, though not as much as implied by the Fine-Mesh model. This effect is called flow splitting and results in an intense shear layer at mountain top height and some way upstream. Lack of resolution causes model airflows to 'rollercoast' *over* a smoothed Alpine mountain profile thereby setting up gravity wave oscillations.

## 3.4 FLOW DETAIL PROVIDED BY THE MESOSCALE MODEL

To finish on a more optimistic note some examples of fine scale flow structures produced by the Mesoscale model will be presented. The model requires high levels of eddy diffusion and time smoothing to run on an operational basis. For instance a 150 km wavelength disturbance decays by a factor of 1/e in 4 hours due to horizontal diffusion and an oscillation of 10 minute period similarly decays in about 25 minutes. Nevertheless this does not prevent the model from representing extremely strong frontal contrast as can be seen from Figs. 3.4 (a), (c) and (d) . These diagnostics were obtained from a run of the Mesoscale model using an interpolated Fine-Mesh analysis at 12Z on October 15 1987 (The Storm) as initial conditions ; they verify 12 hours later. The potential temperature map at 310 m (Fig. 3.4 a ) shows a 8-10 K jump spread across a few gridpoints in the frontal zone and the vertical cross-section (Fig. 3.4 c) (along the line indicated in Fig. 3.4 b) confirms the existence of a marked frontal surface and near-discontinuity through the lowest 1.5 km. The component of the wind normal to the cross-section is shown in Fig. 3.4 (d) : note the very intense horizontal wind shear coincident with the temperature contrast.

Figs 3.5 (a) and (b) show vertical cross-sections of equivalent potential temperature and along-front wind obtained from a rerun of the model on the

Frontal Dynamics Project dropsonde area. This front occurred in the last of eight Intensive Observation Periods of FRONTS 87 and was perhaps the most intense of all. Notable features of the equivalent potential temperature field are the intense gradient marking the frontal zone and the boundary layer modification of the post-frontal air. The along-front wind shows a striking low-level jet at a height of 1 km. and a zone of strong vertical shear which coincides with the frontal surface. Further information on this case can be found in Shutts (1988).

## REFERENCES


Cullen, M.J.P. and C. Parrett 1987 'Mountain wave generation by models of flow over synoptic scale orography'. Met. O. 11 Tech. Note 252.

Dell'Osso, L. and L. Bengtsson 1985 'Prediction of a typhoon using a fine mesh NWP model'. Tellus, 37A, 97-105.

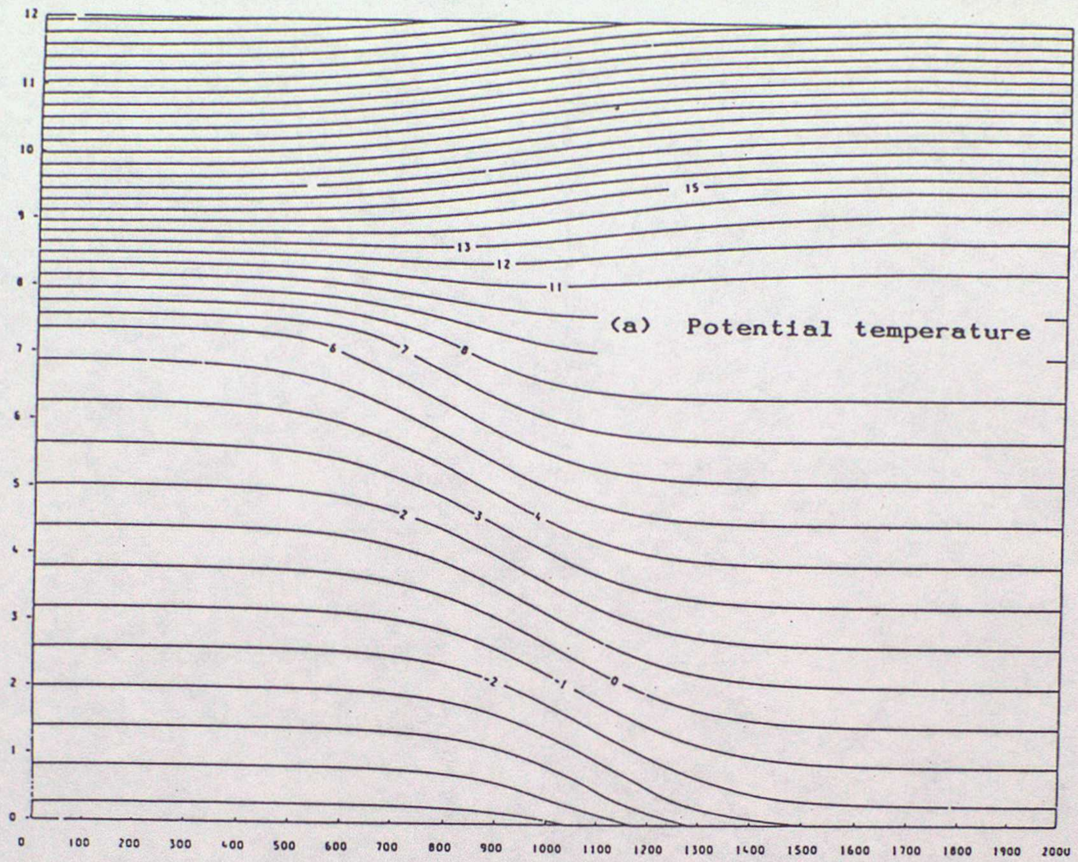
Holt, M.W. 1985 ' A shortcoming of the Operational Convection Scheme at Higher Resolution' Met.O.11 Tech. Note No.215.

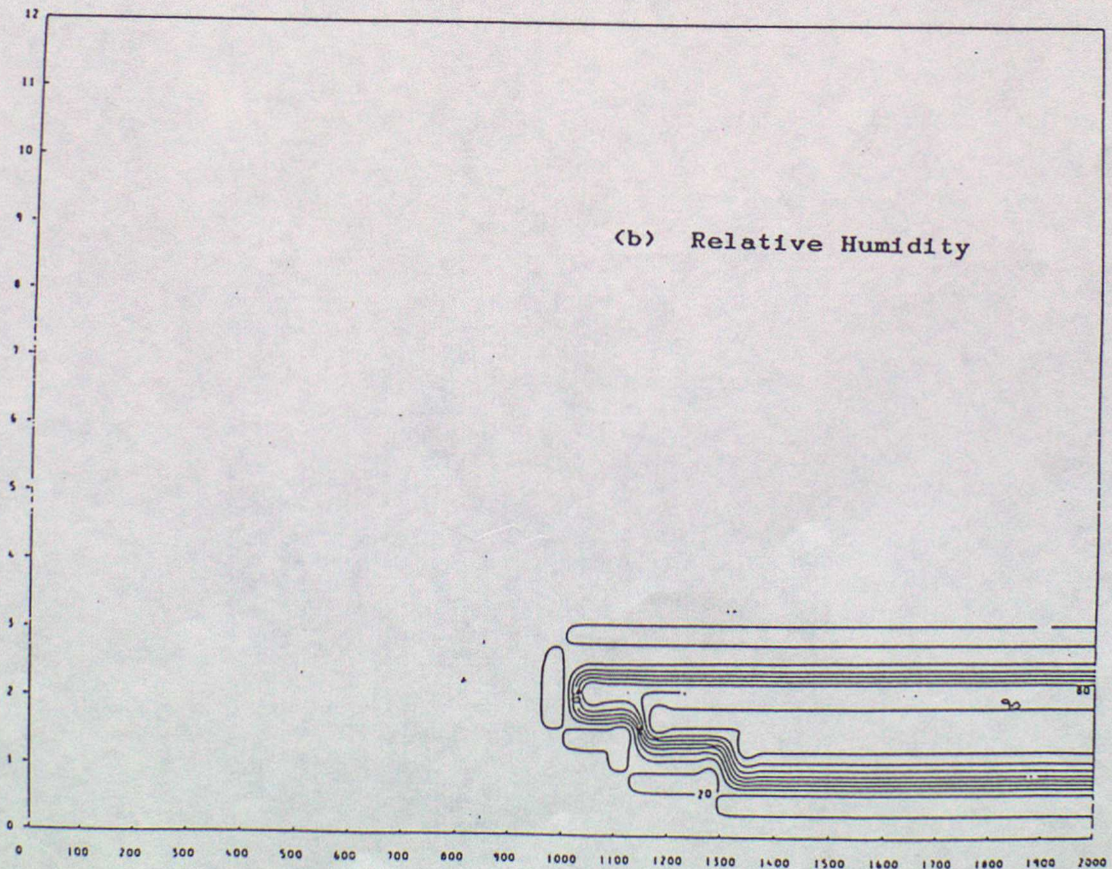
Sanders, F. 1987 ' Skill of NMC Operational Dynamical Models in Prediction of Explosive Cyclogenesis', Weather and Forecasting,2,322-336.

Shutts, G.J. 1987 ' Balanced flow states resulting from penetrative slantwise convection'. J. Atmos. Sci., 44, 3363-3376.

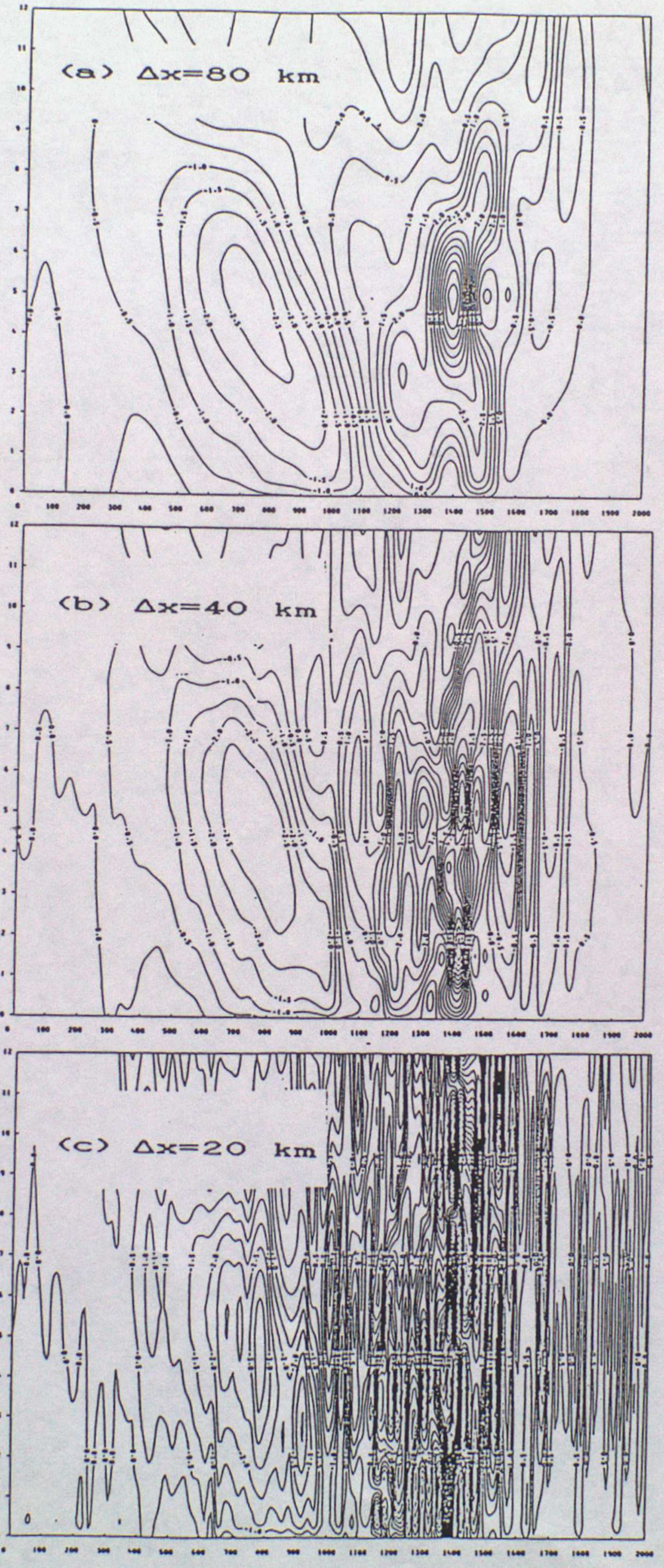
Shutts, G.J. 1988 ' QuickLook Atlas'. Mesoscale Frontal Dynamics Project Report No.6.

Simmons, A.J. and B.J. Hoskins 1978 'The life cycles of some nonlinear baroclinic waves ' J.Atmos. Sci., 35, 414-432.




Figs. 3.1 Initial fields for 2D primitive equation model




Figs. 3.2 Vertical velocity resulting ($cms^{-1}$) resulting from deformation.

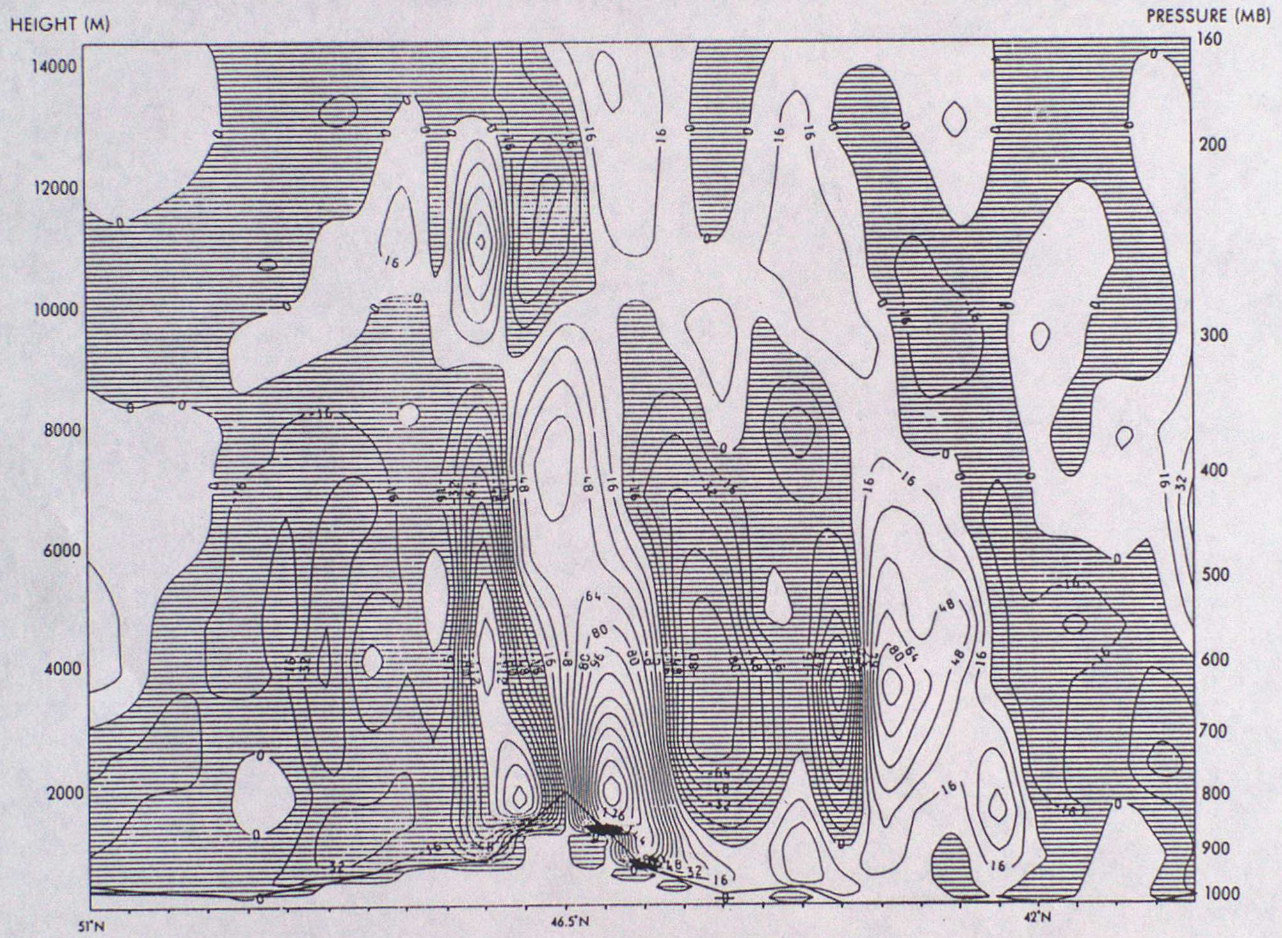
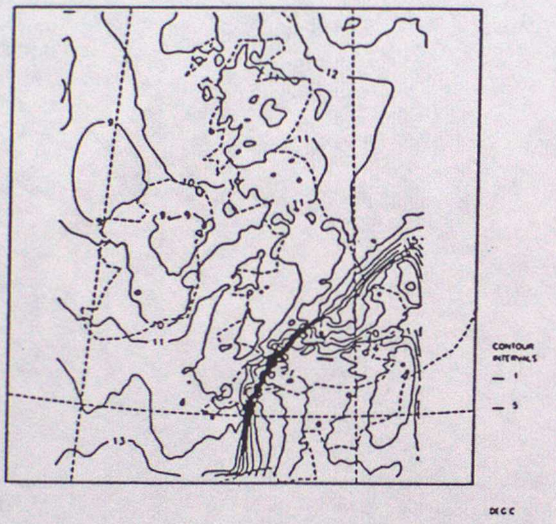
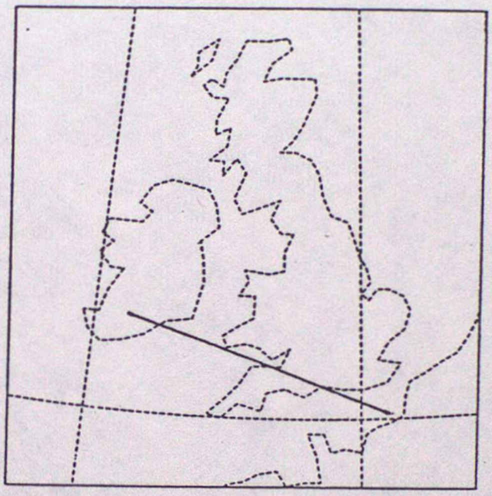



Fig. 3·3 Vertical velocity (mb./hr.) in a north-south cross-section across the Alps from an 18 hour Fine-Mesh forecast valid at 06Z on March 3 1984, (see Cullen and Parrett, 1987) .

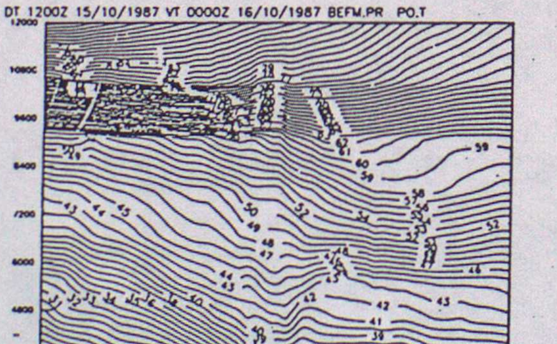
(a) Potential temperature at 310 m.




(b) Vertical cross-section line.

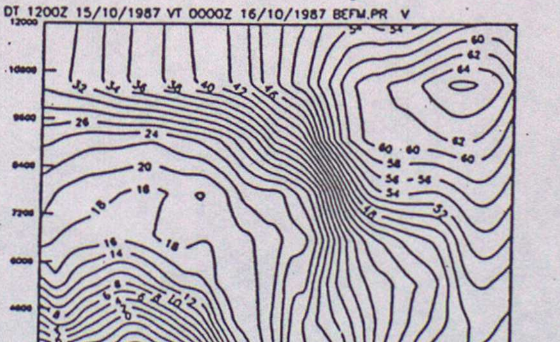


(c) Potential temperature.




(d) Wind component normal to

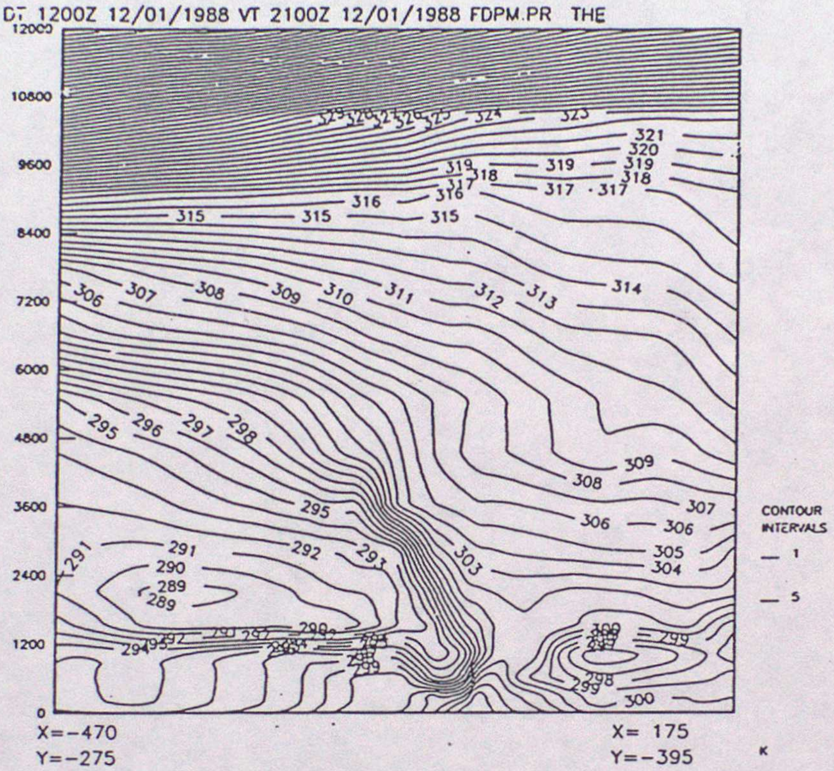
the plane of the picture.



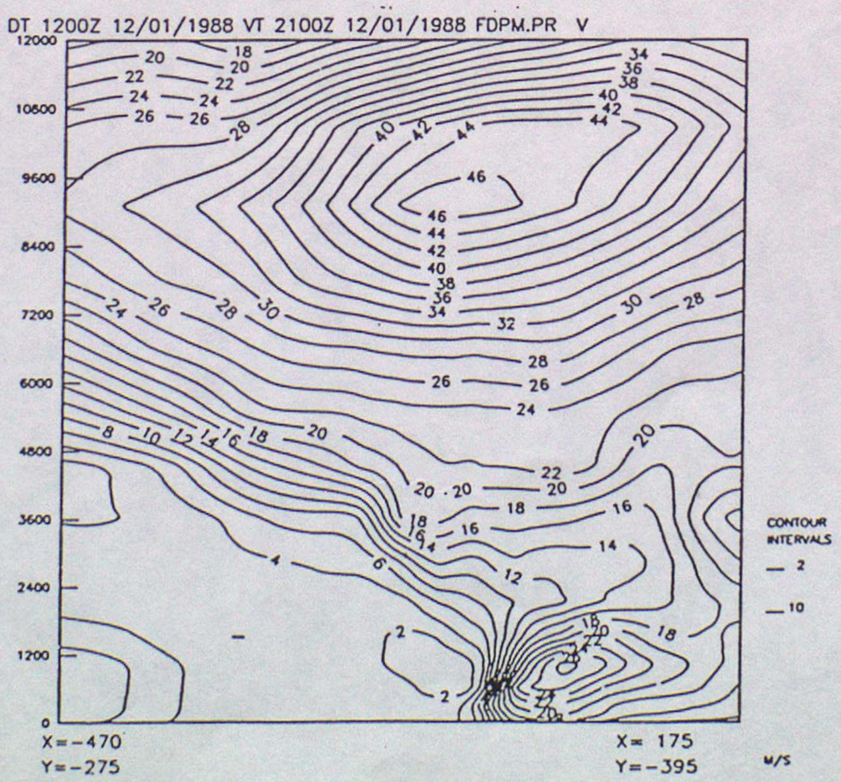

Figs.3.4 Mesoscale model maps (a and b) and vertical cross-sections (c and d) for 00Z Oct. 16 1987.

Figs. 3.5 Vertical Cross-sections from the Mesoscale model for IOP8.

(a) Equivalent Potential Temperature.




(b) Normal wind component.



# MESOSCALE DYNAMICAL PROCESSES

G J Shutts

Met O 11, Meteorological Office, Bracknell

## 4.1 INTRODUCTION

The earth's rotation exerts a powerful control on the sort of atmospheric weather systems which may be forecast with a computer model. Indeed, it is the existence of this rotational constraint that forces energy into the largest planetary scales and causes the dominant energy-containing motions to have time scales of the order of one week or longer. Baroclinic instability, with its shorter time scale ($\approx$ 1-2 days), is also sufficiently slowly evolving compared to $f^{-1}$ (inverse Coriolis parameter) for rotational effects to dominate. On the other hand, the smallest resolvable flow features in current operational models could have a Rossby number $R_o = U/(f\Delta x)$ of the order of unity (taking $U=10$ ms$^{-1}$, $f=10^{-4}$ and $\Delta x=100$ km where $\Delta x$ is the horizontal gridlength) though would, in general, be much less. We are therefore entering a new parameter regime in numerical weather prediction for which the dynamical stability afforded by rotation cannot be guaranteed.

By some definitions, weather systems which have $R_o \approx 1$ are called mesoscale (Emanuel, 1983). It should be noted however that scale analysis does not always provide an accurate measure of the ratio of the particle acceleration to the Coriolis force (the Rossby number). Highly anisotropic weather systems like fronts may have a width of $\approx$ 100 km and associated wind speed of 10 ms$^{-1}$ implying $R_o \approx 1$ yet the flow is mainly along the front and so the frontal width is an inappropriate length scale for measuring particle acceleration. We shall regard frontal zones as being mesoscale even though their true (Lagrangian) Rossby number may be much less than unity.

As alluded to in Lecture 3, we need to forecast the low Rossby number dynamics of frontal zones without permitting spurious high $R_o$ motions to develop at the grid scale. Semi-geostrophic (SG) theory is the (near) perfect mathematical description of highly anisotropic and slowly evolving mesoscale systems - particularly those which contain or have a tendency to form temperature and wind discontinuities. The original paper setting out the SG equations (Hoskins, 1975) focused mainly on the use of the geostrophic momentum coordinate transformation. In this Lecture, I will describe a Lagrangian model (the Geometric model) developed in Met. O. 11 which constitutes the practical realisation of the *extended* semi-geostrophic theory proposed by Cullen and Purser (1984). The model enables us to gain some new insights into mesoscale dynamics that are not easily obtainable by any other approach.

## 4.2 THE GEOMETRIC MODEL

### 4.2.1 Convective and inertial stability

There is a fascinating underlying mathematical structure to the Geometric model ( and Lagrangian view of SG theory) which is beyond the scope of this Lecture. Details can be found in the papers quoted in the first four references. Instead I will present a simple ' picture guide' to the Geometric model which requires little familiarity with conventional dynamical theory.

The simplest possible Geometric model is illustrated in Figure 4.1 (a). Given five isentropic (and incompressible) fluid blobs of known volume, what is the stable equilibrium configuration if this fluid is put into a tall, cylindrical container assuming no mixing ? The answer is, of course, that the potentially warmest fluid must be at the top and coolest at the bottom so as to give positive static stability everywhere. The interface heights are governed by the blob volumes and cross-sectional area of the container.

Now consider the barotropic fluid in the horizontal container in Fig. 4.1 (b). Each fluid blob is now characterised by a certain *absolute momentum* $M_i$ defined as $(fx+v)_i$ where, for instance, x could be the distance eastwards and v the meridional wind component. Such a quantity is conserved in purely two-dimensional motion so that x displacements change v through the Coriolis force. The well-known inertial stability requirement that the absolute vorticity should be $\geqslant 0$ (strictly this applies on an isentropic surface) translates into the requirement that M should increase monotonically in the x-direction. Again this stability condition along with mass continuity implies a unique equilibrium state.

Consider now the three blob problem (Fig. 4.2) where each blob has a different M and $\theta$ value (NB from now on we shall refer to the region occupied by a fluid blob in the equilibrium state as an 'element'). To meet both stability conditions and maintain thermal wind balance across the element boundaries there is again a unique arrangement for the fluid. The thermal wind balance condition is simply Margules' formula for the slope of a frontal discontinuity and may be expressed as:

$$\frac{dz}{dx} = \frac{-f\theta_0 [M]}{g[\theta]} \quad = \text{interface slope} \quad \text{---} \ (4.1)$$

where [ ] indicates the difference across the element boundary.

The foregoing concepts may be extended to flows represented by any number of elements. Figure 4.3 (a) shows the initial element configuration representing a horizontally-stratified flow at rest. Elements within the same horizontal layer have the same $\theta$ value and those in the same column have the same absolute momentum. We consider now the equilibrium state which results if element 30 is heated so that its potential temperature is between the corresponding values in the third and fourth layers up from the bottom. The Geometric model is able to find this state assuming conservation of $M_i$, $\theta_i$ and element areas $A_i$ for $i=1$ , 100. Fig. 4.3 (b) shows that in this new balanced state, element 30 has jumped upwards and has squeezed itself between the third and fourth layers up, adopting a lens-like shape. We could imagine that real penetrative convective mass transport might ultimately lead to mesoscale anvils of this shape after the geostrophic adjustment phase has completed. The ratio of the width to depth of this lens is, for a continuously stratified fluid, of the order of N/f where N is the buoyancy frequency.

4.2.2 Frontogenesis

The clearest way of visualising frontogenesis in a fluid bounded above and below by rigid horizontal surfaces is to study the zero potential vorticity model of Hoskins and Bretherton (1972) for which $\theta=\theta(M)$. The classical frontogenetic mechanism is that due to flow *deformation*. In this, air parcels suffer stretching in the along-front sense and compression in the direction of the horizontal temperature gradient. The area of cross-section of individual elements $A_i(t)$ decreases as their length is extended. It may be shown that the absolute momenta of the elements decreases in proportion to their area. $\theta$ is assumed to be conserved in this model. Fig. 4.4 (a) shows the initial model state with M and $\theta$ increasing monotonically from left to right. Eq. 4.1 demands that the slope of element interfaces should decrease in time in proportion to

[M] and therefore [A]. In general these interface slopes change at different rates so that a collision of the lines happens first on the upper and lower boundaries, Fig. 4.4 (b). As time proceeds these boundary discontinuities grow into the fluid as more interfaces collide (Fig. 4.4 (c)). This process, described by Cullen (1983), *is* frontogenesis and is limited in reality primarily by mixing as the local Richardson number falls below 1/4.

4.2.3 Mountain flow blocking

The technical problem of using the Geometric model in the presence of mountains is very difficult. This relates to the differing ways in which the inertial stabilty requirement can be met in the presence of an intruding solid boundary. This non-uniqueness is discussed fully in Cullen et al (1987). It does not mean that we have to abandon any hope of solving initial value problems but great care is required to follow the correct physical solution.

Consider the element representation depicted in Fig. 4.5 (a) which is based on a vertical cross-section across the Alps derived from the Fine-Mesh model. The shaded block represents a mountain barrier 2 km high. Elements with equal $\theta$ can be identified by their vertical inter-element boundaries so that elements 32 to 36 form a cold wedge approaching from the north if the mountain block is translated from right to left. Figs. 4.5 (b) and (c) show the balanced flow states at 6 and 12 hours later respectively. Note how the leading element (36) of the cold wedge is blocked by the mountain at 6 hours and subsequently splits into two elements (36A and 36B) so that at 12 hours 36A is ' banked ' up against the mountain and 36B has formed a downstream dome. In this phase of the solution, air in 36A jumps discontinuously to 36B as the mountain is forced further into the cold air. In reality, this transfer would take the form of an unbalanced downslope windstorm or Bora. Model energy is dissipated implicitly in this jump : air in the upper part of element 36A has high kinetic and potential energy due to the generation of a mountain-parallel jetstream and the upwelling of cold air. This disappears as air from the element is released from the mountain top and jumps downstream to the dome whilst conserving M. Another interesting feature of the model integration is the tendency of air with higher potential temperature ( in element 28 ) to be sucked down on to the leeside of the block. This ' Foehn ' effect does not require latent heat release or large amplitude gravity wave motion.

The above type of solution is very difficult to model with the primitive equations since it involves shear layers of small vertical scale and a synoptically-controlled energy dissipation mechanism. The kinetic energy released as air escapes over the mountain is in reality likely to give rise to gravity wave and local turbulence (perhaps associated with hydraulic jumps). Since the primitive equation model can only represent the latter process through eddy diffusion, it is unlikely to achieve the correct partitioning between gravity wave generation and turbulent dissipation. A useful analogy to consider is that of a weir in a river. The potential energy lost as water drops to its new level turns to kinetic energy which is very efficiently dissipated in turbulence at the foot of the weir. Nevertheless, a small proportion of the kinetic energy is converted to surface gravity waves which then propagate downstream. Any attempt to model such a process using the Navier-Stokes equations would need the turbulence to be explicitly resolved to ensure that excessive gravity waves are not generated out of undissipated kinetic energy. For further information on this weir analogy and a somewhat simpler element model of 2D flow over different shaped orographic ridges see Shutts (1987a).

### 4.2.4 Axisymmetric flows

In order to study the influence of heat and momentum sources/sinks on quasi-balanced vortex flows, much use has been made of axisymmetric models. Most of the early models and theories of hurricane development assumed axisymmetry. The semi-geostrophic Geometric model discussed so far is valid for nearly straight flows in geostrophic and hydrostatic balance. Recently however, it has been extended to cover the case of flows with circular symmetry for which there is *gradient wind* and hydrostatic balance (Shutts et al, 1988). This requires the use of *angular* momentum instead of *absolute* momentum and a particular radial coordinate transformation which makes element interfaces straight. We have used this version of the Geometric model to study the balanced flow states which arise from penetrative convection in a vortical environment and will present one experiment here.

Consider the element distribution in radial cross-section depicted in Fig. 4.6 (a). The angular momentum and potential temperature values of elements are consistent with a horizontally-stratified atmosphere at rest with respect to a rotating system. The inner wall must be located at finite radius (r) otherwise a singularity is implied by non-zero angular momentum at r=0. The shaded elements are to be heated so as to convect to the upper part of the model. Fig. 4.6 (b) shows the resulting balanced state with surface pressure field set below. On removal of the shaded fluid, environmental air subsides and converges to fill the gap. The thick solid line represents a curved frontal discontinuity across which a marked change in tangential wind and pressure gradient has formed. Maximum tangential winds in this warm-cored vortex are to be found just outside of the front. We call this discontinuity line, 'the eyewall front' by analogy with the hurricane eye wall. The upper-level lens formed out of the shaded fluid is broader in extent and has weaker anticyclonic flow with very intense velocity shears near the outer rim.

## 4.3 CONCLUDING REMARKS

The coding of the Geometric model was carried out by S. Chynoweth as part of his Ph.D (Chynoweth, 1987) . It, in itself, constitutes an extremely challenging problem in computer science and deserves much attention. The computational expense of calculating the element geometry restricts us from using any more than about 1000 elements for most purposes though over 6000 elements were used in one geometrical construction described in Chynoweth's thesis. It has been possible however to devise analytic solutions (using conformal mapping methods) corresponding to similar idealised experiments to those discussed in this Lecture (eg. Shutts, 1987b). We can think of these as perfect resolution thought experiments in which there is no doubt about the appearance of discontinuities. These solutions, together with those from the Geometric model, provide a vivid dynamically-consistent picture of the formation of balanced mesoscale flow structures containing near-discontinuities.

## REFERENCES


Chynoweth, S — 1987 'The semi-geostrophic equations and the Legendre transformation'. Ph.D Thesis, Dept. of Mathematics, Univ. of Reading, England.

Cullen, M J P and R J Purser — 1984 'An extended Lagrangian semi-geostrophic theory of semi-geostrophic frontogenesis'. J. Atmos. Sci. , **41** , 1477-1497.

Cullen, M J P — 1983 ' Solutions to a model of a front forced by deformation'. Quart. J. Roy. Met. Soc., **109**, 565-573.

Cullen, M J P,S Chynoweth and R J Purser — 1987 ' On some aspects of flow over synoptic scale topography'. Quart. J. Roy. Met. Soc., **113**, 163-180.

Emanuel, K — 1983 'On the dynamical definition(s) of Mesoscale'. 'Mesoscale meteorology - Theories, Observations and Models', (Eds. D. Lilly and T. Gal-Chen), D.Reidel, 1-11.

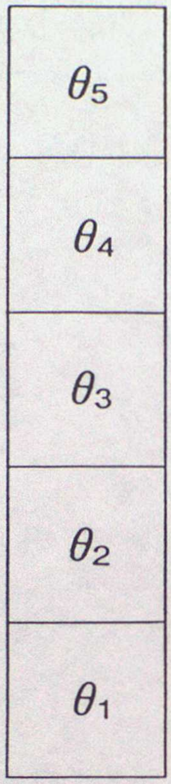
Hoskins, B J — 1975 'The geostrophic momentum approximation and the semi-geostrophic equations'. J. Atmos. Sci., **32** , 233-242.

Hoskins, B J and F P Bretherton — 1972 'Atmospheric frontogenesis models : mathematical formulation and solution'. J. Atmos. Sci., **29**, 11-37.

Shutts, G J — 1987a 'The semi-geostrophic weir : a simple model of flow over mountain barriers'. J. Atmos. Sci., **44**, 2018-2030.

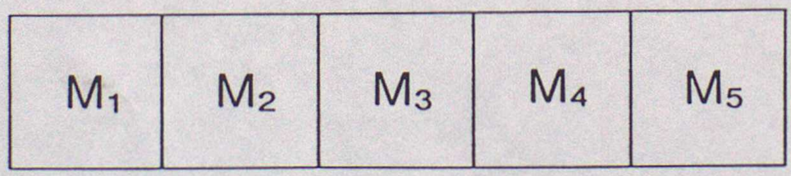
Shutts, G J — 1987b 'Balanced flow states resulting from penetrative slantwise convection'. J. Atmos. Sci. **44** , 3363-3376.

Shutts, G J , M Booth and J Norbury — 1988 'A Geometric model of balanced, axisymmetric flow with embedded penetrative convection'. J. Atmos. Sci.,**45**, (August).

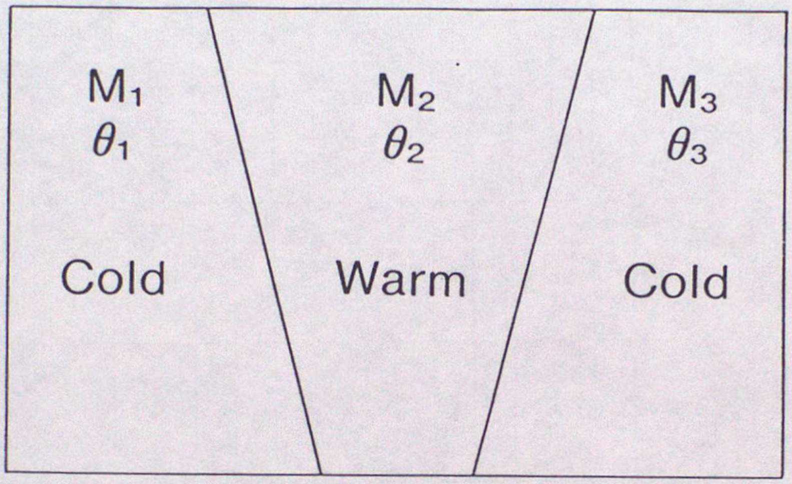



$$\theta_{i+1} > \theta_i$$
$$i = 1, 4$$

Figs. 4. 1 1D element models

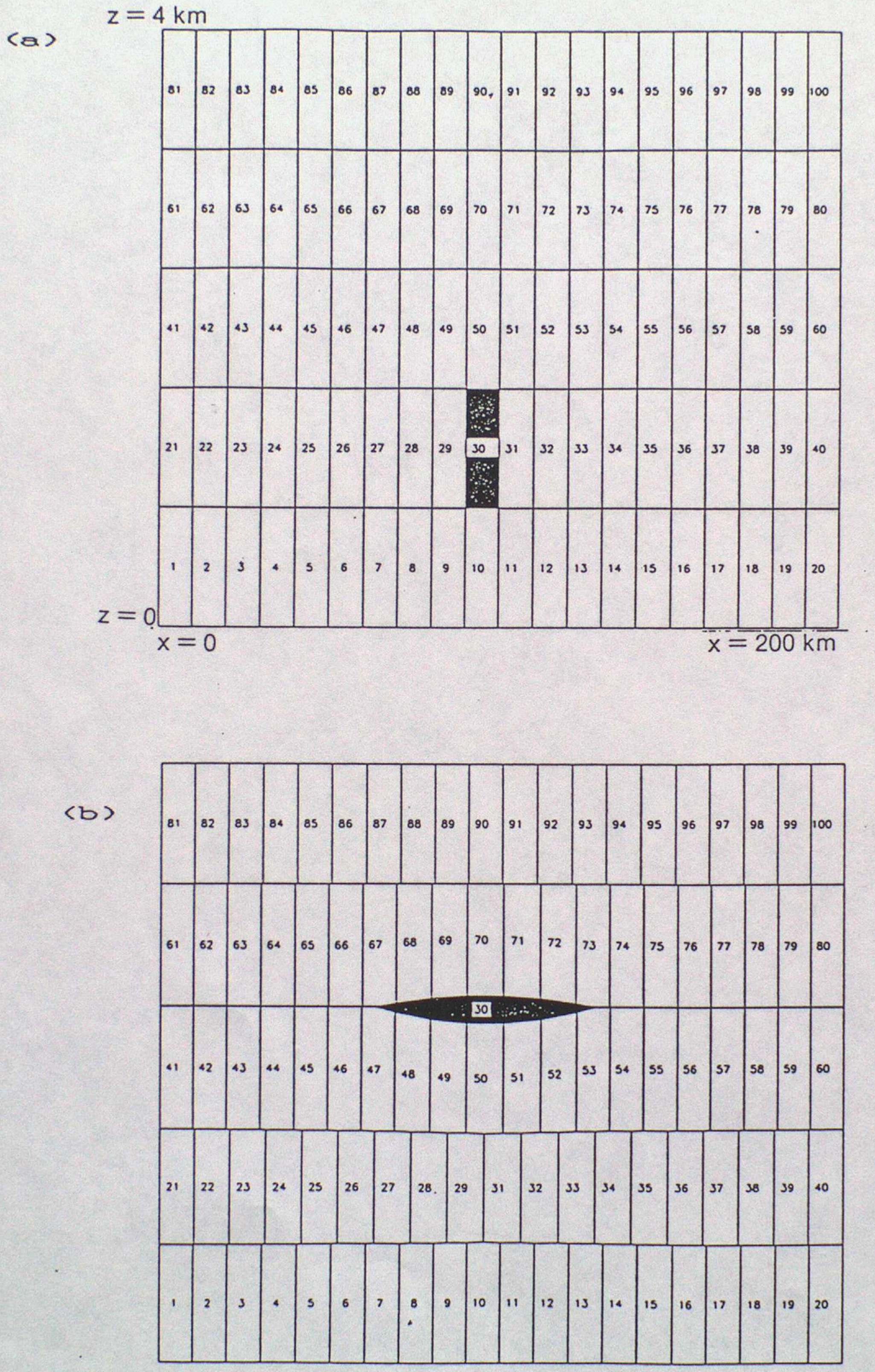



$$M_{i+1} > M_i$$
$$i = 1, 4$$



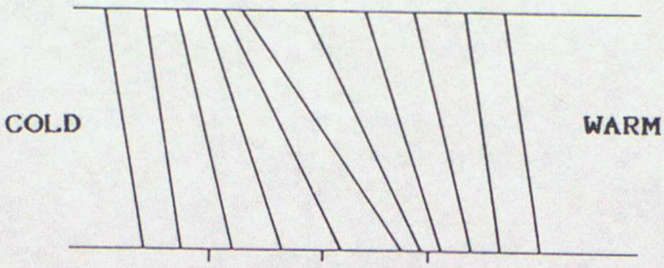

$$M_3 > M_2 > M_1$$
$$\theta_2 > \theta_1, \theta_3$$

Fig. 4.2 Three element model.



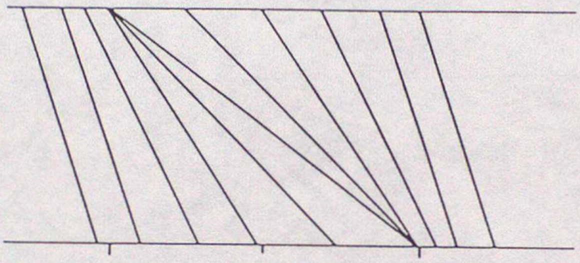

Figs. 4.3 (a) Initial element distribution for stratified flow at rest
(b) Final state after heating element 30.

(a)



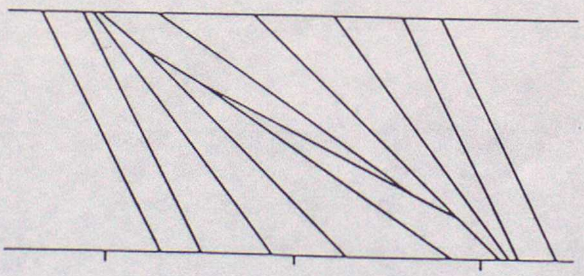

Typical initial distribution of segments.

(b)



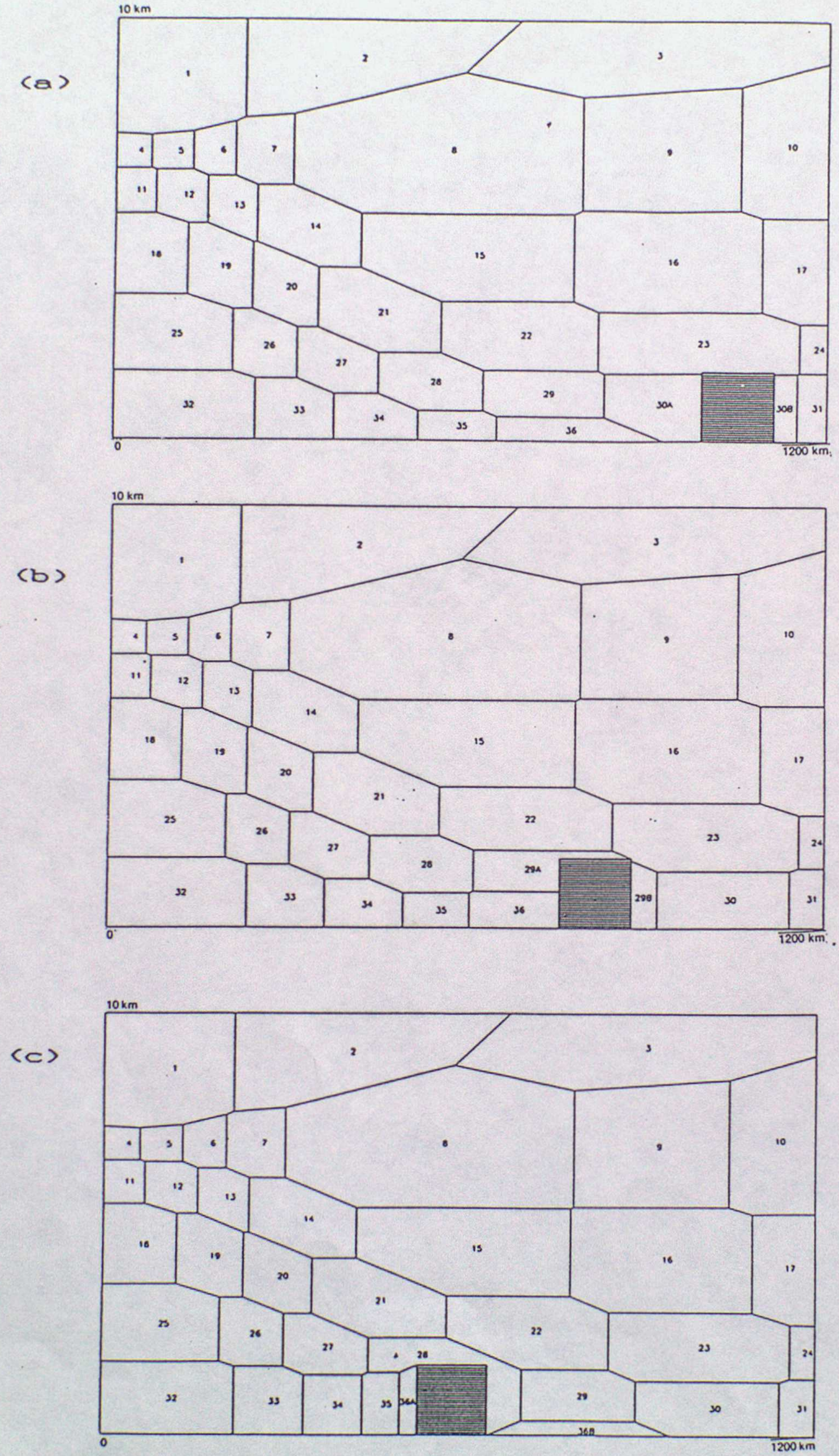
Distribution of segments at the formation of a discontinuity.

(c)

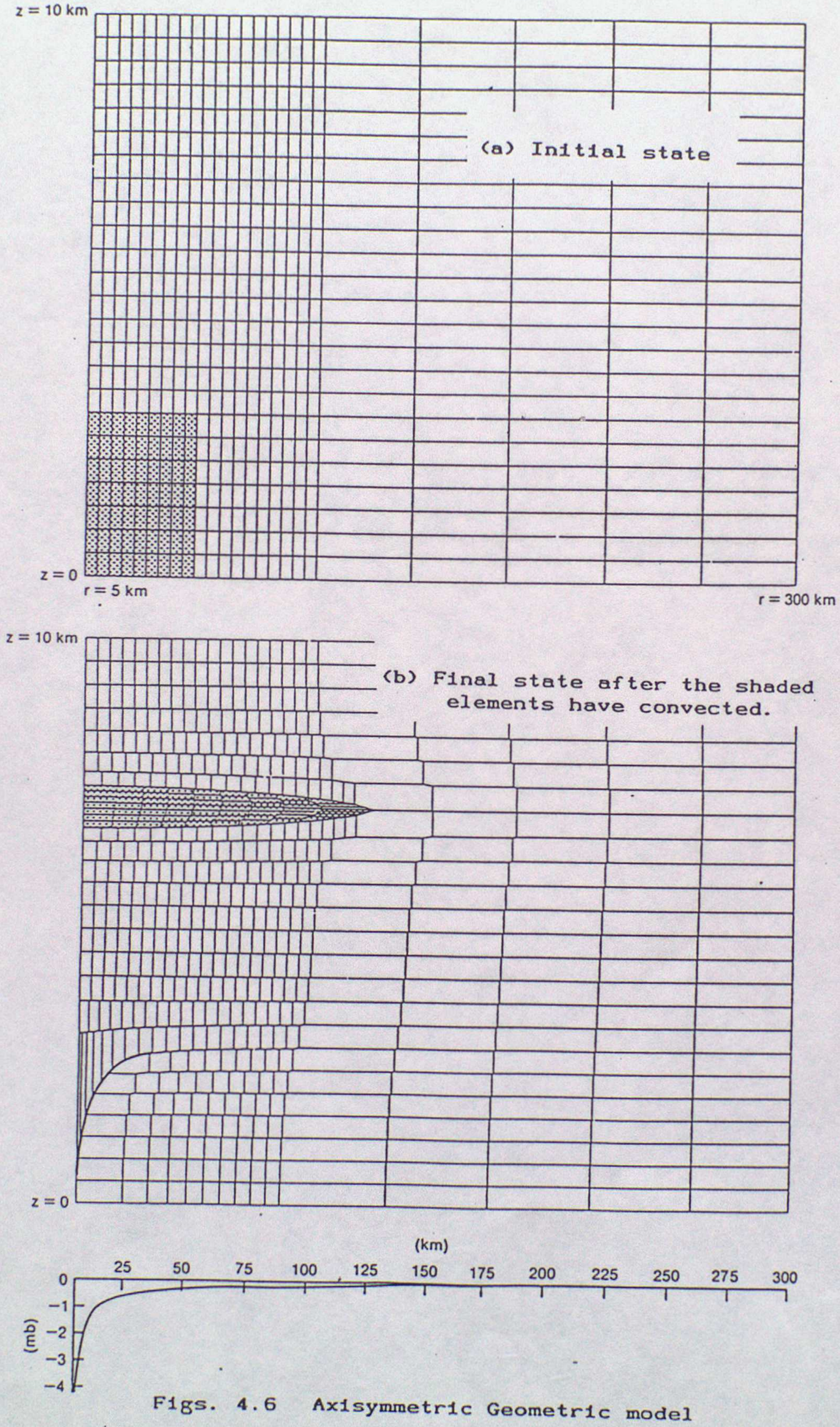


Evolution of the discontinuity.

Figs. 4.4 Three stages of frontogenesis under the action of deformation. The areas of elements in (b) and (c) have been rescaled to their original value.




Figs. 4.5 Three stages in the evolution of frontal flow over a rectangular orographic barrier.




Figs. 4.6 Axisymmetric Geometric model

# THE USE OF CONCEPTUAL AND DYNAMICAL UNDERSTANDING IN THE DESIGN OF NWP SYSTEMS

**M J P Cullen,**

**Met O 11, Meteorological Office, Bracknell.**

## 5.1 INTRODUCTION

In this lecture I describe a method of designing an NWP model which attempts to simulate the restricted range of atmospheric motions decribed in Lecture 4. This is an illustration of how a conceptual understanding of how the atmosphere works can be converted into a model design, and therefore translated into more accurate operational weather forecasts. This particular model has only undergone limited testing to date, and so its usefulness is not yet known. The techniques used to design it could be applied to other conceptual models which will be developed in the future.

These notes only summarise the main points of the lecture, since all the detail is contained in the papers listed in the bibliography.

## 5.2 MODEL FORMULATION

The dynamical filtering is done by explicitly solving the Lagrangian form of the semi-geostrophic equations described in Lecture 4. It was shown there that most motions of direct relevance to weather forecasting can be described in at least a simplified way, and that many of the motions excluded are too complex to be treated accurately in a NWP model. In particular, the interaction between the simplified dynamics and physical processes can be treated quite rigorously.

## 5·3 MODEL DESIGN

At present it does not appear practicable to use the Lagrangian method described in Lecture 4 in an operational model, despite its clear advantages in describing convection and its ability to achieve good results up to the limit of vanishing potential vorticity. A conventional finite difference method is therefore used, and the parametrizations of physical processes are also conventional. What is novel is the way these are linked together.

The dynamical equations determine the total velocity field implicitly, and so an implicit numerical method must be used. The diagnostic equation which has to be solved (effectively a three-dimensional Sawyer-Eliassen equation) has coefficients which depend on the local atmospheric state. The very efficient constant coefficient methods of solution used in semi-implicit models such as the ECMWF model or Meteorological Office mesoscale models cannot be used. At present an alternating direction method is used to break the problem into two-dimensional problems on vertical cross-sections. These two-dimensional problems are solved by block tridiagonal methods.

The equations become very ill-conditioned as the potential vorticity approaches zero, and cannot be solved if the potential vorticity is negative. Convection schemes in conventional models are designed to maintain moist static stability but not positive potential vorticity. No method has yet been found of adjusting a field locally to maintain positive poential vorticity which does not also have unacceptable side effects. The method adopted is therefore to enforce moist vertical stability by using a conventional convection scheme. This will replace the instantaneous jump of fluid in the geometric model with a representation of real convective clouds. The adjustment has to be made in a timestep, which may be unrealistically fast. Use of a one hour timestep allows the simplified dynamics to be adequately treated while not distorting real convective timescales too much. The same problem arises in the operational fine mesh model, where the convection scheme is forced to stabilise the atmosphere in a 7½ minute timestep, much faster than real convection can operate.

The model must also enforce horizontal inertial stability. This is done by a simple mixing scheme. The real processes being parametrized here include unbalanced downslope winds. A reduced version of the Sawyer-Eliassen equation is then solved to estimate the ageostrophic wind, and this is then substituted into the full semi-geostrophic equations including moisture. The full equations will not be satisfied exactly and the estimate must then be improved iteratively. The iteration may not converge if the moist potential vorticity is negative, and further adjustments will be needed to the data, which parametrize moist slantwise convection. The iteration also includes the alternating-direction iteration between north-south and east-west solutions. Thus, though a long timestep is used. most of the calculations in it have to be repeated a number of times to ensure adequate coupling between the directions.

The above shows how moist processes become a fully integrated part of the model, rather than separate 'add-on' routines. A similar argument applies to boundary layer processes. In the boundary layer the semi-geostrophic equations are reformulated so that friction becomes part of the definition of the balanced solution. Because the balance is now between three forces, there are two modes of solution, one of which decays very fast as the friction tends to zero. The dependence of the friction on the total wind has to be built in to the equation which estimates the ageostrophic wind. The boundary layer parametrization is thus now an integral part of the solution procedure.

## 5·4 OTHER ASPECTS

In such a model, the motions not represented explicitly should be parametrized. The gravity wave drag schemes developed for the existing models may well be appropriate. In a semi-geostrophic model a wave free basic state is provided as input to the scheme, which is consistent with the way such schemes are designed.

The use of a balanced model has advantages for data assimilation, where it is normally necessary to use explicit or implicit balance relationships to spread observed information about one variable into the analysis of other variables. This suggests that such a model might be useful for data assimilation even if found inadequate for forecasting. Another use is in calculating the dynamical response to inputs from the parametrizations schemes, such as the convection scheme.

Other conceptual models of atmospheric behaviour may be developed. They are all likely to express some coupling between quantities defining the state of the atmosphere. This coupling has to be built into a numerical model by using an implicit integration scheme. A variable coefficient implicit method such as outlined here is therefore a key tool.

## BIBLIOGRAPHY

General philosophy:

M.J.P. Cullen, J. Norbury, R.J. Purser and G.J. Shutts (1987)
*'Modelling the quasi-equilibrium dynamics of the atmosphere'*, Q.J.Roy.Met. Soc., **113**, 735-757.

Finite difference scheme:

M.J.P. Cullen (1988)
*'Implicit finite difference methods for modelling discontinuous atmospheric flows'*, Met O 11 Sci. Note no. 7.

Boundary layer scheme:

M.J.P. Cullen (1988)

*'On the incorporation of atmospheric boundary layer effects into a balanced model'*, Met O 11 Sci. Note no. 6.

Three dimensional model:

M.J.P. Cullen and M.H.Mawson (1988)

*'A fully implicit scheme for the three-dimensional quasi-equilibrium equations'*, Met O 11 Working Paper no. 85.