Use of High Performance Computing (HPC) in "Weather" modelling

Alan Gadian, based at Leeds’ University and many others in NCAS Weather.

NCAS (National Centre for Atmospheric Science) is a NERC funded distributed Centre, with staff at many universities in the UK, and covers topics of “Composition”, “Climate” and “Weather”.

Scientists use HPC facilities predominantly under the NCAS computer allocation but also with peer reviewed grant allocations.

This talk will refer to some HPC activities in the “Weather”, but will predominantly focus on areas of my interest, rather than representing a complete review of HPC usage in this topic area.
Massively Parallel Questions for Weather

THE CHALLENGE OF MAKING THE MOST OF MASSIVELY PARALLEL COMPUTERS

Meteorology is vector process, (or so we have considered). We also have to consider the efficiency of the calculation. Should we bother with these new computers?

What has to be done if we do:

- Domain decomposition
- Load balancing algorithms
- Numerical algorithms
- Grid adaptivity, structured & unstructured, omega formulation, other schemes of limited iteration accuracy (toy games)

What about now?
Questions to be addressed by the *Weather* modelling community

The talk will address the issues:-

- why do we need fast computers? What is the scale of the problem?
- show some current HPC *Weather* modelling activities. What do we actually do with HPC computing?
- indicate some challenges for modelling *Weather*, both in the long term and in the short term. How is the change to massively parallel computers going to affect What are the likely future needs?

Contents:

- WRF “North Run”. Can this be a future process study model?
- UM modelling (Gravity Waves and Climate Research).
- LEM / BLASIUS parallelisation challenges.
- City Scale flows (Eulag, VHREM and CFD)
Do we believe the results?

The “Weather Research & Forecasting” model WRF has been constructed by the National Center for Atmospheric Research (NCAR) USA and is used all over the world, and is well supported.

There is now a chemistry version called WRF-CHEM.

Originally, designed as a local / mesoscale model it is now being modified to be a global model, with intentions to make it a climate model.

The “North Run” numerical is being setup up to run as a research activity in NCAS to assess the capabilities of the model to model atmospheric process over orography.

Can this model represent convection?

Is it realistic?

Does it work on HECToR?

Is it to be recommended?
WRF is a mesoscale model built at the National Center for Atmospheric Research. One project is to assess how this model represents meso-scale motions in the atmosphere, and performs as a forecasting model; weekly weekend forecasts will be posted on the NCAS web page. How does resolution affect accuracy?
Performance figures claimed by WRF. (Not published) It looks as though this may be just CPU, without IO
Simple WRF wall clock comparison for the “North Run”

Simple wall-clock comparison of times taken for a specific 3hr model “North Run” simulation (ignoring IO effects which dominates for large core jobs)
Current UM speed up figures.

UM version 6.1 performance

- perfect scaling
- 50% scaling
- N48 resolution
- N96 resolution
- N144 resolution
- N216 resolution

1 node on HPCx = 16 processors

UM atmosphere model resolutions:
- low: N48 -> N96 -> N144 -> N216
- high
Preliminary Assessment of Speed up of Um and WRF codes

Simple Summary of UM / WRF Speed – up Meteorological codes

Efficiency ~ Cost of using N processors (cores)

\[ \frac{N \times \text{cost of using 1 processor (core)}}{\text{raw wall clock comparison for WRF, no IO allowance}} \]

(raw wall clock comparison for WRF, no IO allowance. Other unknowns with the UM data)

WRF
For “North Run” size of job (approx 400*300*80 grid points)
4 times the 128 cores (512 cores), efficiency is ~ 80%
8 times the 128 cores (1024) efficiency is ~ 40%.

UM
This corresponds to the say UM Nx (where x is somewhere nearer N216 rather than N144)
Compared with 1 node (16 cores) on HPCx 400 cores efficiency is ~ 40%
There is potential for improvement with the UM, if UM figures are accurate.
**WRF**

The current aim is to **EVALUATE** the code

- a. use to produce a weekly forecast. (NCAS web-site)
- b. look at how well it scales on multi-processors
- c. use to look at boundary layer development, COPS.
- d. use to look at convection, valley flow. COPS

**Isosurface**

- defined on $W = 0.2\text{m/s}$ looking north from the southernmost face of the domain: 2d run, 16Z
- defined on $W = 0.2\text{m/s}$ looking north from the southernmost face of the domain: 3d run, 16Z
There is a continuing need and a very important need, to use the UK Meteorological Office Unified Model (UM) in “NCAS Weather” e.g. for

- Studies of Deep Convection in the UM
- Impact of Cloud Condensation Nuclei (CCN) sensitivity studies, using HaDGaM
Gravity waves, deep convection and cross-tropopause transport

Jeffrey M. Chagnon
(with Sue Gray)
University of Reading
Case study design

- Deep convection near UK
- Selected from summers 2004-2005
- Simulated using high resolution mesoscale version of the Met Office Unified Model (UM)
  - nonhydrostatic, compressible, horizontal grid lengths of 12, 4, 1 km (convective parameterization off a 1 km)
  - Passive tracer is initialised above the tropopause (mixing ratio = 1 in strat, 0 otherwise)
- Test the sensitivity of transport and gravity wave characteristics to model resolution.
Overview of 1km simulation

1 km simulation (7 hour loop)

- Precipitation (red > 30 mm/hr)
- Tracer beneath tropopause (black=0, white = 1)

Ahead of the lowered tropopause, convection forms and transports some stratospheric tracer into the upper troposphere.
Transport into the free troposphere

1 hour loop, mid troposphere

Vertical velocity (white = +1 m/s)
Mid troposphere

Tracer (red > 10\(^{-5}\)kg/kg)
Mid troposphere

The stronger cells mix stratospheric air downwards to mid tropospheric elevations.
Summary II

Lower stratosphere

The stronger cells mix stratospheric air downwards to mid tropospheric elevations. Above the tropopause, some of the penetrating updrafts deposit tracer-free tropospheric air into the lower stratosphere.
Some "climate scale simulations" on the effects of CCN (Laura Kettles)
Two major codes in use for process studies are also from UK Meteorological Office

1. Large Eddy Model code (LEM)

This model has good microphysics representation but has no orography, is bi periodic, and uses an fft solver for a helmholtz equation

2. BLASIUS

This model has orography, and some basic microphysics. It uses a terrain following co-ordinate system / transformation.

The challenge with these codes is that they were written for vector systems and do not work well on massively parallel systems. A plan is needed to overcome this issue.
New Developments – cfd (Ian Castro et al.)

NCAS has developed a partnership with CD-Adapco (Ian Castro) at Southampton to look at urban flows to develop CFD meshing software.

Simulations from the “DAPPLE” Marylebone Road site are produced and compared with observations. Below is a Google Earth and “model” image of the orography.

This software is currently licensed on HPCx, and is licensed to run on systems up to ~100 cores.

Some argue that engineering (CFD) codes deal poorly with dispersive waves (i.e., stably stratified non-hydrostatic flows) whereas they are well tuned for either neutrally (i.e. unstratified) flows or heated (i.e. Unstably stratified) flows. The jury is still out (in my opinion!)
Model development using massively parallel code from scratch

New developments include Imperial College ICON (Cut Cell model)

In Weather the “microscale model” project was developed to:

- investigate modelling approaches for very high-resolution studies;
- explore the applicability of full atmospheric models for flows on scales of $O(<100m)$;
- provide an accessible research model;
- provide a resource to help improve our understanding of small-scale dynamics/processes.

Note that a problem to be overcome is the development of efficient Helmholtz solver (exactly as in the UM .. John Thuburn, Exeter)
Klemp-Wilhelmson time-splitting method (for 1st-order stepping)

\[ \frac{\partial \phi}{\partial t} = s_\phi + f_\phi \]

Fast modes: \( \Delta \tau \)
Slow modes: \( \Delta t \)
where \( \Delta t = N \Delta \tau \)

Solutions at time

\[ n+\Delta \tau: \]
\[ \phi^{n+\Delta \tau} = \phi^n + \Delta \tau s_\phi^n + \Delta \tau f_\phi^n \]

\[ n+2\Delta \tau: \]
\[ \phi^{n+2\Delta \tau} = \phi^{n+\Delta \tau} + \Delta \tau s_\phi^{n+\Delta \tau} + \Delta \tau f_\phi^n \]
\[ = \phi^n + \Delta \tau \left( s_\phi^n + s_\phi^{n+\Delta \tau} \right) + 2\Delta \tau f_\phi^n \]
\[ \vdots \]

\[ N+\Delta t = n+N\Delta \tau: \]
\[ \phi^{n+N\Delta \tau} = \phi^{n+(N-1)\Delta \tau} + \Delta \tau s_\phi^{n+(N-1)\Delta \tau} + \Delta \tau f_\phi^n \]
\[ = \phi^n + \Delta \tau \left( s_\phi^n + s_\phi^{n+\Delta \tau} + \ldots + s_\phi^{n+(N-1)\Delta \tau} \right) + N\Delta \tau f_\phi^n \]
\[ \phi^{n+\Delta t} = \phi^n + \Delta \tau \left( s_\phi^n + s_\phi^{n+\Delta \tau} + \ldots + s_\phi^{n+(N-1)\Delta \tau} \right) + \Delta t f_\phi^n \]
Immersed boundary scheme (from LMz)

Orography set-up:

- Cartesian grid, i.e. horizontal vertical levels
- define orography height at column-centres
- interpolate orography heights at column-corners, using a continuous piecewise bilinear function
- 3 types of grid-cell:
  - i) “pure Earth” cells
  - ii) “pure air” cells
  - iii) “cut-cells” treated with finite-volume method
- calculate “weights” associated with air surfaces/volumes in grid-cells
  - simple since bilinear function means intersections are straight lines
- “weight” = ratio of area/volume in air to area/volume of full grid-cell

From Steppeler et al. (2006)
Flow over a 3D bell-shaped hill:

**Vertical cross-section through centre of hill:**

- **Time = 0s**
- **Time = 180s**
- **Time = 360s**

**Horizontal cross-section at height 250m:**

- **Time = 0s**
- **Time = 180s**
- **Time = 360s**
Flow over a 3D bell-shaped hill:

Contours of vertical velocities - vertical slice through centre of hill

$xz$ plot of $W$, $J = 30$, $time (sec) = 1440.0$

Solid lines = upward flow; dashed lines = downward flow

CONTOUR FROM $-1$ TO $0.6$ BY $0.1$
Old code, with new numerical methods on a 5-sided building

There are questions, whether with good newer iterative methods, for the Helmholtz equation, and even with terrain following co-ordinate system, results can be obtained which are good.

Pentagon setup
- Boundary-fitted representation $z' = H(z-h(x,y))/(H-h(x,y))$
- $600 \times 600 \times 31 @ \Delta x = \Delta y = \Delta z = 2m$
- 7200 time steps @ $\Delta t = 0.05 s$
- Rigid upper boundary
- Eulerian option 2nd order in space and time
- sgs 1 ½ order closure
- Specified CD on building and stc
- Neutral with prescribed velocity profile from previous LES simulation (Moeng and Sullivan)
- NCAR supercomputer ("older" IBM MPI using 200 processors): 10 ½ hrs wallclock time for 7400 time steps (with 3 tracers)

Some results (cont): $w$ at $z=10m$ after 7200 time steps (6 min)

Some results: $w$ after 7200 time steps (6 min)
Interesting other cases

Set-up for Gillygate experiment:
(Dixon et al, 2005, Atmos Env)

These results use the Smolarkiewicz model, 256 processors on HPCx

- periodic domain, with uniform $z_0 = 0.1m$
- grid-boxes: $231 \times 261 \times 60$, for $dx = dy = dz = 1m$
- time: 1200 time steps, for $dt = 0.025s$
- model spin up ~ 30s - results computed for 30s run time – it is anticipated that the spin up and run times need to be longer for statistically significant outputs
- Rayleigh damping sponge above 50m
- Neutral, (constant potential temperature), $u_0 = 5ms^{-1}$ from right to left. The periodic boundaries develop a logarithmic type surface layer

Plan of Gillygate, York - Dashed outline defines the model domain. Two lamp-posts G3 and G4 are marked on the diagram
Easterly Wind over the Gillygate street canyon

Upper left frame is Vertical velocity.
All other frames are concentrations.
The two lamp-posts are marked.
Summary

- There is a great deal of HPC computing in the University Weather Sector.
- There is an ever increasing demand for resources.
- There are major challenges to cope with the arrival of massively parallel computing.
- A great deal of resource / effort will be needed to take advantage of these new computing resources.
- New approaches (not discussed) of grid adaptivity, unstructured grids, “toy games” approaches will be needed to improve our understanding and prediction of “weather”
1. **WRF**  Aim to **EVALUATE**
   a. use to produce a weekly forecast.
   b. use to look at boundary layer development, COPS.
   c. use to look at convection, valley flow. COPS

2. **EULAG**
   a. Street Canyon with Gal-Chen & Immersed Boundary lbc
   b. Convection over Hajar Mountains.
   c. SCMS, warm clouds

3. **VHREM**
   A new terrain intersecting, finite volume immersed lbc structured C grid with warm microphysics.
   All singing and dancing.
Some "climate scale simulations" on the effects of CCN (Laura Kettles)

<table>
<thead>
<tr>
<th>Layer Cloud Droplet Effective Radius ((\mu m)) at cloud top</th>
<th>Liquid Water Path (g m(^{-2}))</th>
<th>Outgoing Shortwave Radiation (W m(^{-2})) at TOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) control</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
</tr>
<tr>
<td>b) with (N_d = 375) cm(^{-3}) in Sc regions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
</tr>
<tr>
<td>c) with (N_d = 230) cm(^{-3}) in Sc regions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
<td><img src="http://www.ncas.ac.uk" alt="Image" /></td>
</tr>
</tbody>
</table>