Challenges and Controversies in Multiscale Fluid Dynamics Algorithms for Numerical Weather Prediction and Climate Modeling

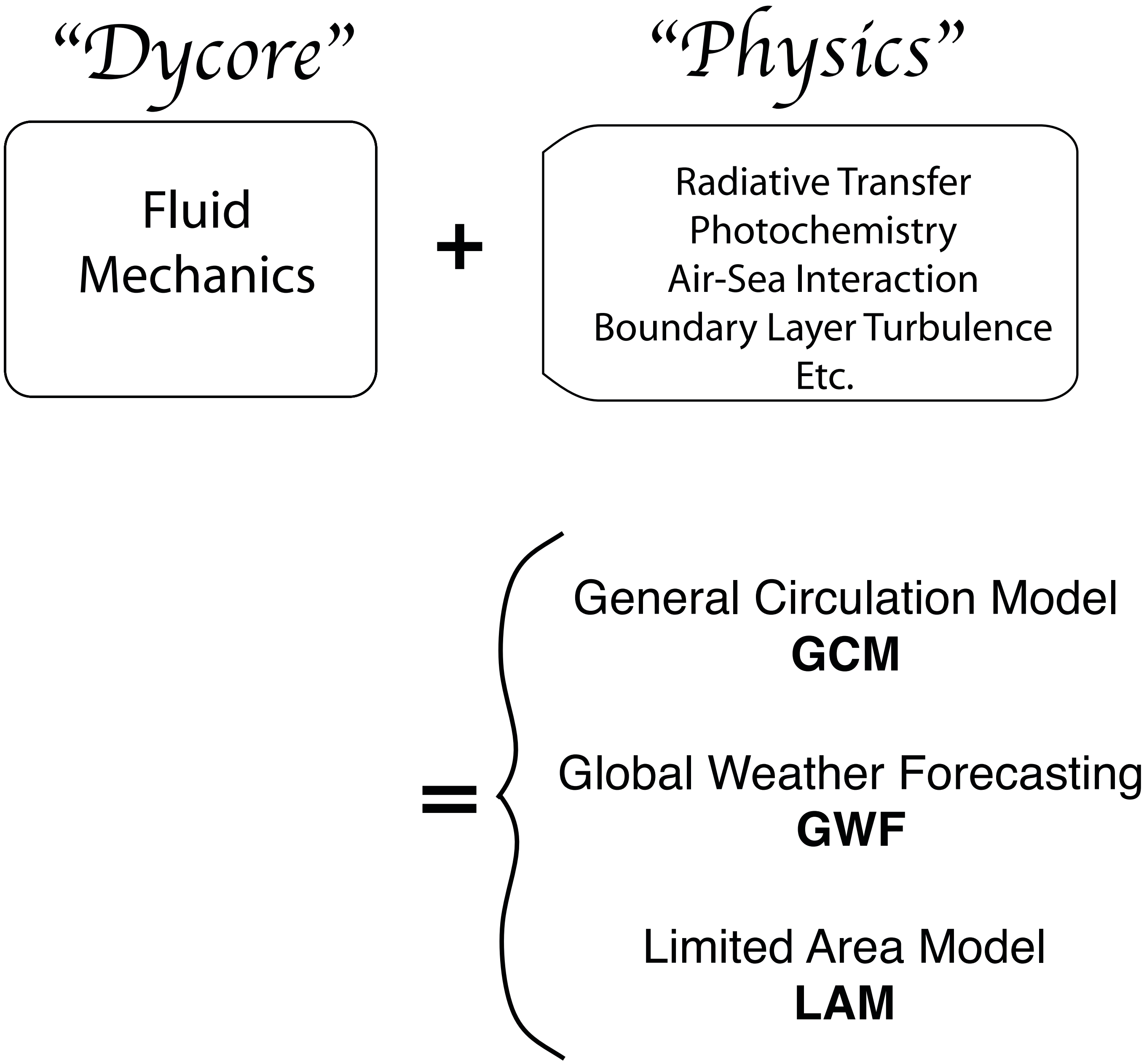
John P. Boyd

University of Michigan

July 10, 2008

Dycores & Physics

Models Fall Into 3 main groups.
Each has two major parts.



GCM, GWF and LAM Differences

- GCMs are run at LOW RESOLUTION [compared to forecasting codes] for LONG TIMES (years) from ARBITRARY INITIAL CONDITIONS to reach a STATISTICAL EQUILIBRIUM
- GWF run at HIGH RESOLUTION for 5–10 DAYS from OBSERVED INITIAL CONDITIONS to make a DETERMINISTIC FORECAST
- LAM are run on SMALL SPATIAL DOMAINS at VERY HIGH RESOLUTION for a FEW HOURS, NESTED in a GLOBAL MODEL. TARGET is SEVERE WEATHER: Hurricanes, typhoons, squall lines, etc.

Why Forecasting is Not a Pure IVP

Observations must be INITIALIZED

- Atmosphere is close to a "slow manifold" which is free of gravity waves
- Observations, due to instrumental and interpolation errors, violate this balance
- Data must be adjusted ("INITIALIZED") back to the slow manifold
- Uninitialized data triggers spurious gravity waves and lots of spurious rain
- Several practical strategies in operational use
- Slow manifold doesn't actually exist because of "beyond-all-orders" terms in the small parameter, $\epsilon =$ Rossby number;
- Math Challenge: Hyperasymptotics of "fuzzy, slow-with-a-little-fast manifold" are still confusing
- Math Challenge: PDE compatibility conditions are ignored

Data Assimilation

- Data from satellites and other observing systems can be fed into the forecast while the model is running. This is called “data assimilation”
- Radiosonde balloons collect data SYNCHRONOUSLY over the entire globe
- Satellite and other continuously-measuring systems collect ASYNCHRONOUS data — completely ignored in pure initial value approach.
- Because the assimilated data often consists of a single field, it must be adjusted towards the slow manifold, too
- Methods borrowed from optimization and control theory
- Many variants with different computational expenses
- Very active research frontier

Ocean Models

Generally similar to atmospheric models but

- Some physics can be omitted (CLOUDS, PHOTOCHEMISTRY, MOST RADIATIVE TRANSFER)
- Some physics added (SALT, BIOLOGICAL MODELS)
- MULTIPLY-CONNECTED FLUID DOMAIN (Islands, continents)
- VERY SPARSE OBSERVATIONS
- Characteristic scale of energy-containing vortices ("Rossby Radius of Deformation"), is TEN TIMES SMALLER in OCEAN
- No FORECASTING MODELS [except for surface wave height, hurricane storm surges, etc.]

Peculiarities: Differences from Other CFD

- "Physics" requires more code than the dycore
  Each subfield is a separate research area
- Strong HEIGHT/HORIZONTAL anisotropy
  Often use wildly different methods in height versus latitude-longitude
- Best models use 1000 x 1000 horizontal grid but only 100 levels. Ironically, the 25 km lat/long grid spacing is HUGE compared to vertical spacing of $O(0.5)$ km
- Many vertical coordinates: p, log-pressure $z*$, height z, pressure-over-surface-pressure σ, potential temperature θ, hybrid coordinates, hydrostatic height π^*, vertically-Lagrangian.
- NEVER solve the full Navier-Stokes; various approximations ("hydrostatic", "shallow atmosphere", etc.) plus replacement of molecular viscosity by implicit and explicit computational damping and turbulence models.
- Wave-y, not shock-y

Trends

1. Model convergence: GCM, WFM, LAM are more and more similar and sometimes the same
2. Hydrostatic ⇒ Nonhydrostatic
3. Eulerian advection ⇒ Semi-Lagrangian & Discontinuous Galerkin
4. Shallow atmosphere ⇒ Deep atmosphere
5. Ensemble Forecasting
6. Data Assimilation (in Forecasting)
7. Coupled ocean/atmosphere climate models

Fundamental Limitations

Definition 1 (Arithmurgy) *"Arithmurgy" is a synonym, literally meaning "number-working", for the craft that is now usually called "computational science and engineering". From the Greek* $\alpha\rho\iota\theta\mu o\sigma$, *"number", and* $\epsilon\rho\gamma o\sigma$, *"working".*

Definition 2 (Log Law of Arithmurgy) *Insight grows logarithmically with the number of floating point operations. The growth is sometimes punctuated by jumps ("saltation") so the curve of insight versus floats resembles a flight of stairs with a logarithmic trend.*

Corollary 1 (Linearity of Progress) *Insight in a scientific or engineering problem grows* LINEARLY *with* TIME *even when Moore's Law is true.*

Limitation on Forecasting

- Ed Lorenz argued in 1963, and later studies with real models have confirmed: errors grow EXPONENTIALLY fast with time.
- CHAOS theory justifies
  LOG LAW of ARITHMURGY for
  WEATHER FORECASTING

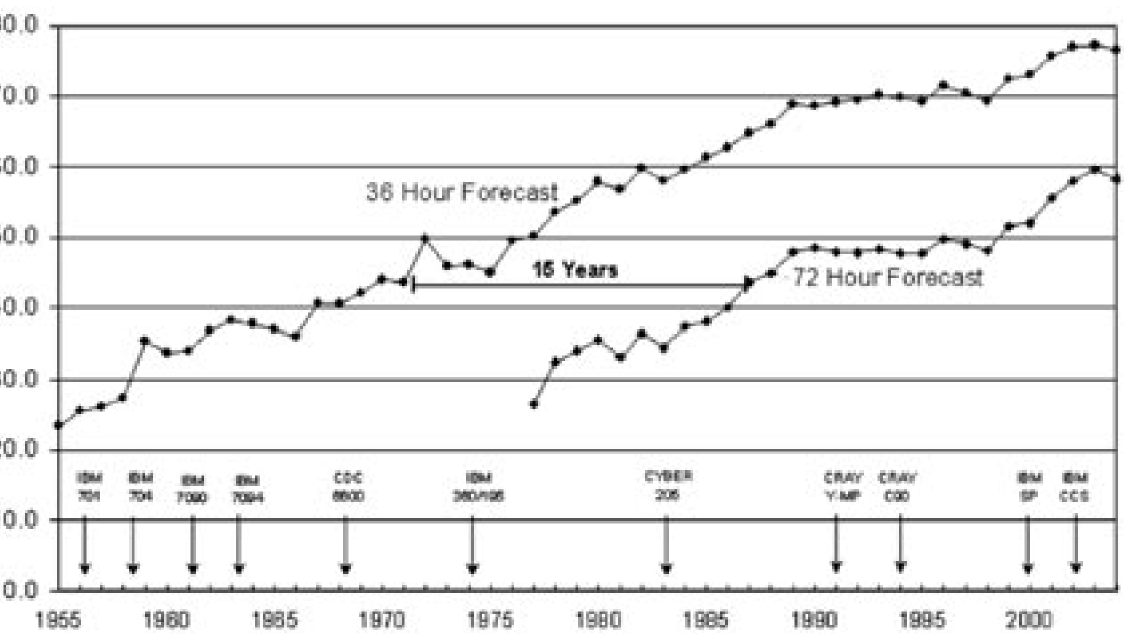


Figure 1: Fig. 4 of Peter Lynch, JCP (2008)

FORECASTING SKILL for a fixed time interval grows LINEARLY with TIME even though COMPUTING has grown EXPONENTIALLY

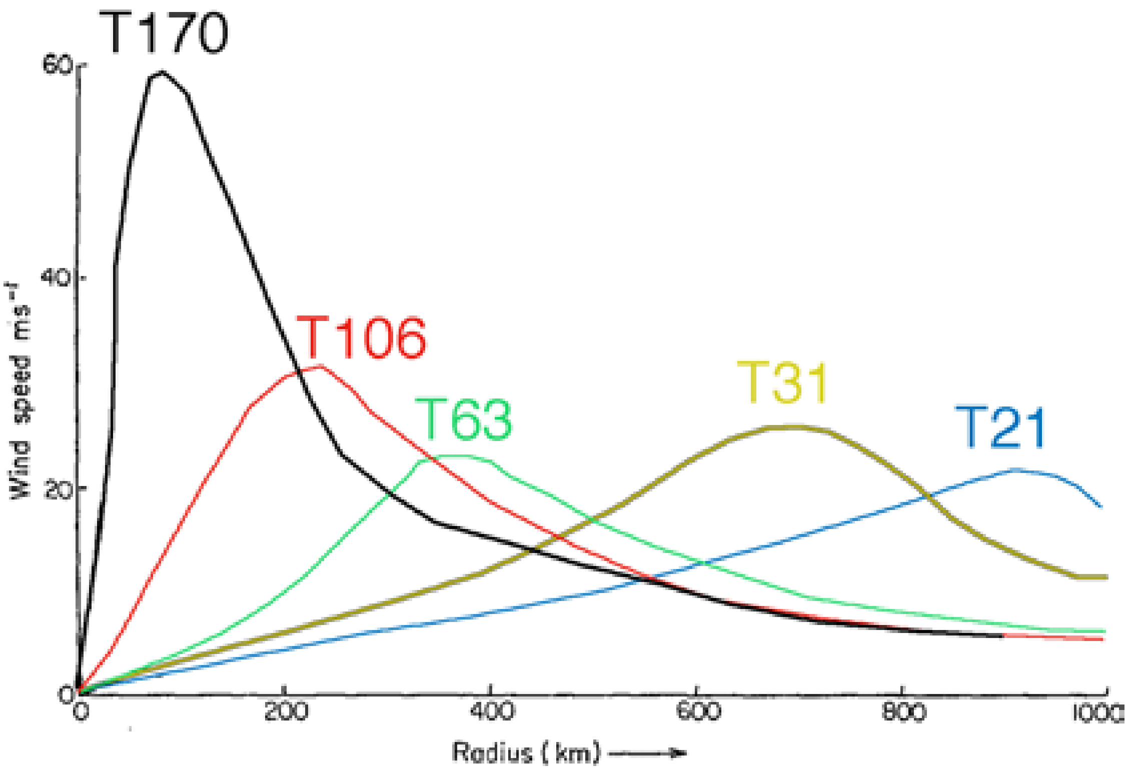


Figure 2: . From Krishnamurti and Oosterhof [?]

Radial distribution of tangential velocity at 850 mb height in Typhoon Hope at different resolutions. T_N denotes a triangular truncation with N^2 spherical harmonics. Even the T_{170} storm is somewhat wider than observed.

This illustrates how INSIGHT JUMPS as new phenomenon COME INTO FOCUS.

Here, the gross structure of a typhoon is well-represented at a resolution of T250, but distorted on coarser grids.

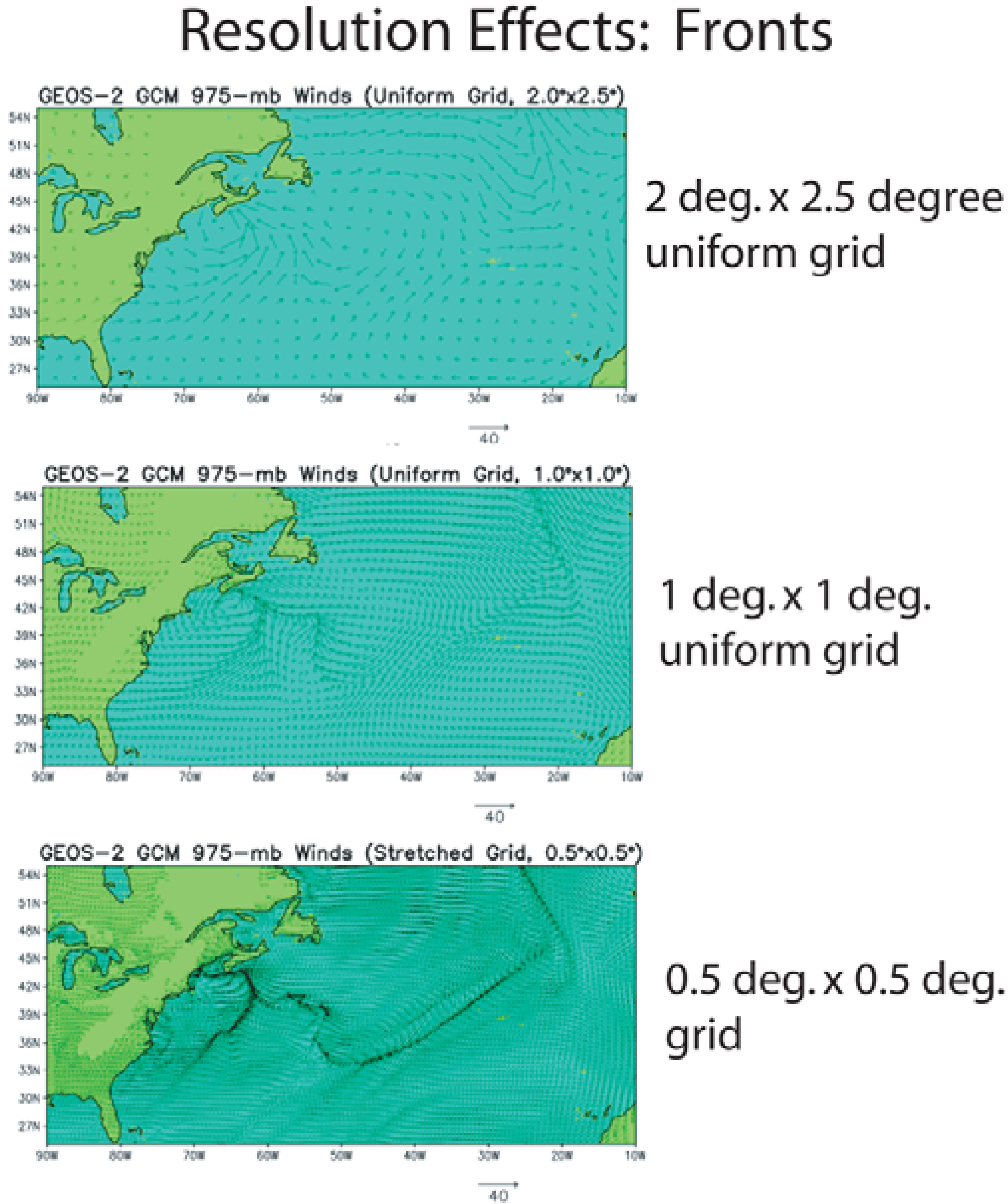


Figure 3: Atmospheric fronts as computed using the GEOS-2 model at three different resolutions. From Conaty *et al.* [**?**]. The black curves in the bottom figure show that synoptic fronts are resolved at the highest resolution.

- Present day forecasts have some accuracy to about FIVE DAYS
- Economic studies, such as by Sherden, show that forecasts are cost-effective in spite of the "CHAOS LIMIT"

- Improvements in computing can add a little to the FORECASTING WINDOW
- Because forecasts are dependent on OBSERVATIONAL DATA, it is unlikely that deterministic forecasts will ever extend beyond TWO WEEKS — even this may be UNACHIEVABLE
- Room for progress around the edges: hurricane forecasting, longer tornado warnings, more accurate local forecasts, etc.

Meteorology is not a field of BREAKTHROUGHS; it is a field of SLOW, INCREMENTAL PROGRESS

Limits to Climate Forecasting

- Climate equilibrium is controlled by STRONG, POORLY KNOWN FEEDBACKS
- Roe and Baker have argued theoretically that it will NEVER be POSSIBLE to know the feedbacks well enough to much reduce the odds of EXTREME TEMPERATURE RISE

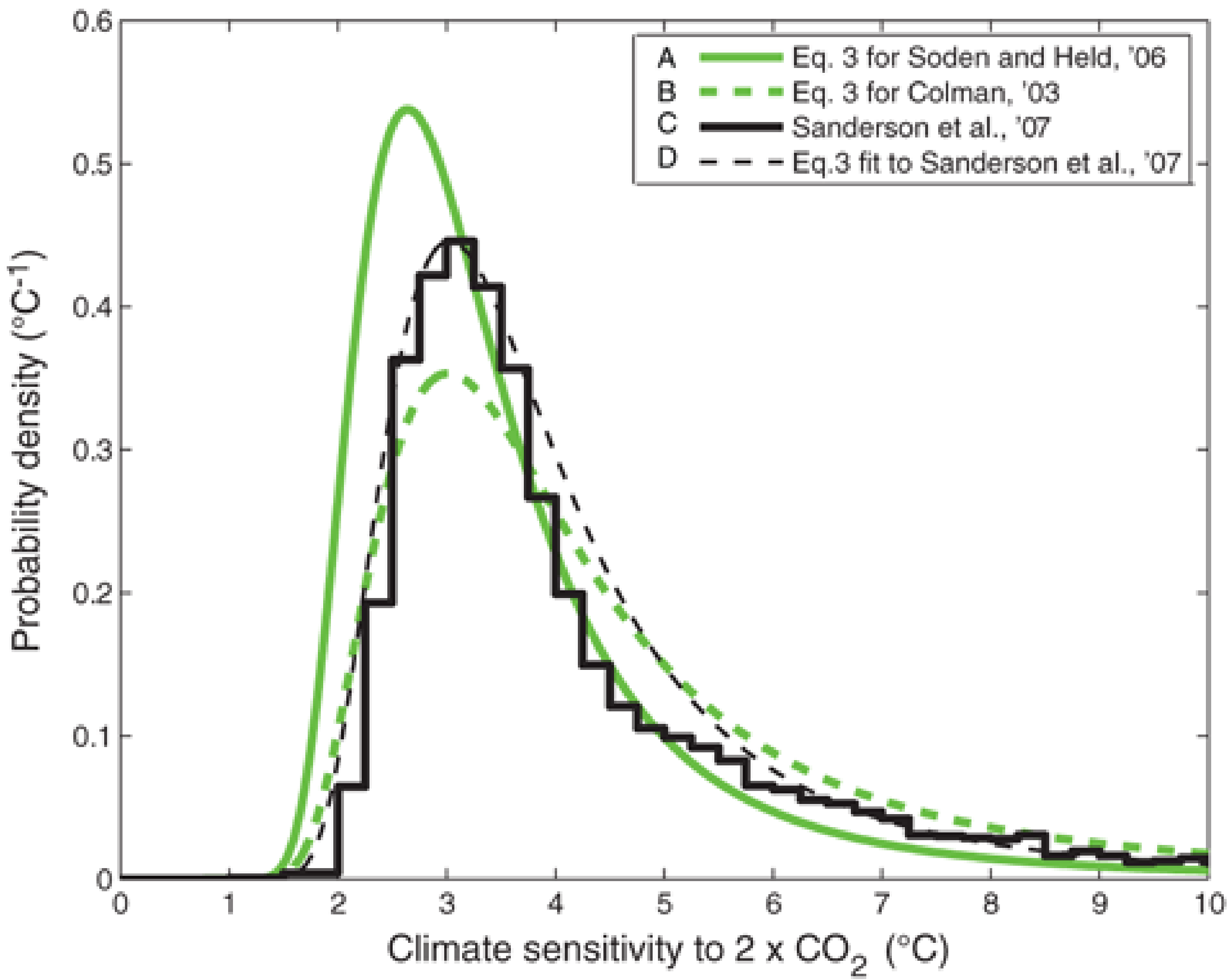


Figure 4: From Gerald Roe and Marcia Baker (2007)

- Mid-course corrections in a COUPLE of DECADES are POSSIBLE
- Maybe improvements in REGIONAL FORECASTING, TOO

"We now know that we can't forecast weather beyond a few days ahead. So, now let's do something even harder: forecast climate! Even so, it's worth a try because the potential payoff is so huge."

— Philip D. Thompson, explaining why he accepted the job of the founding director of NCAR's Climate Project (1976)

Special Difficulties of Spheres

- Science fiction writers have imagined artificial, toroidal planets.

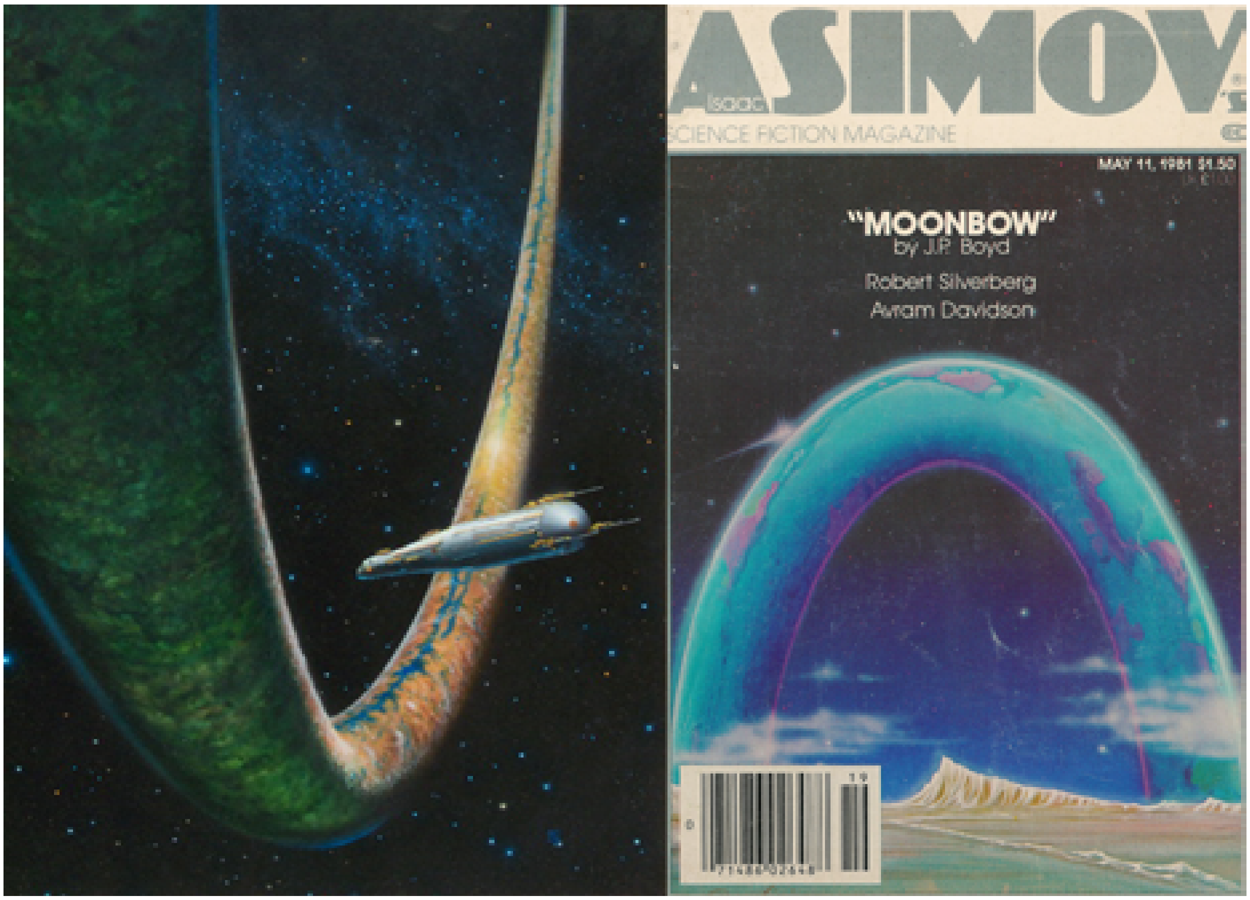
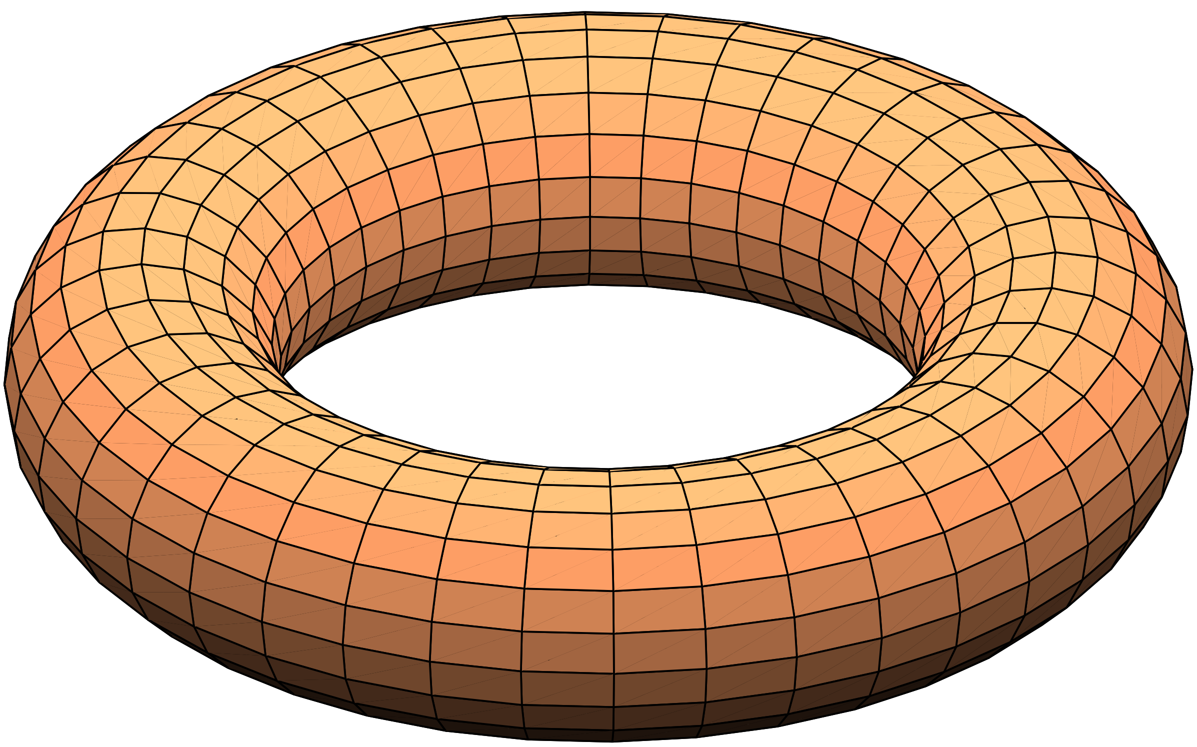


Figure 5: Both pictures by Rick Sternbach

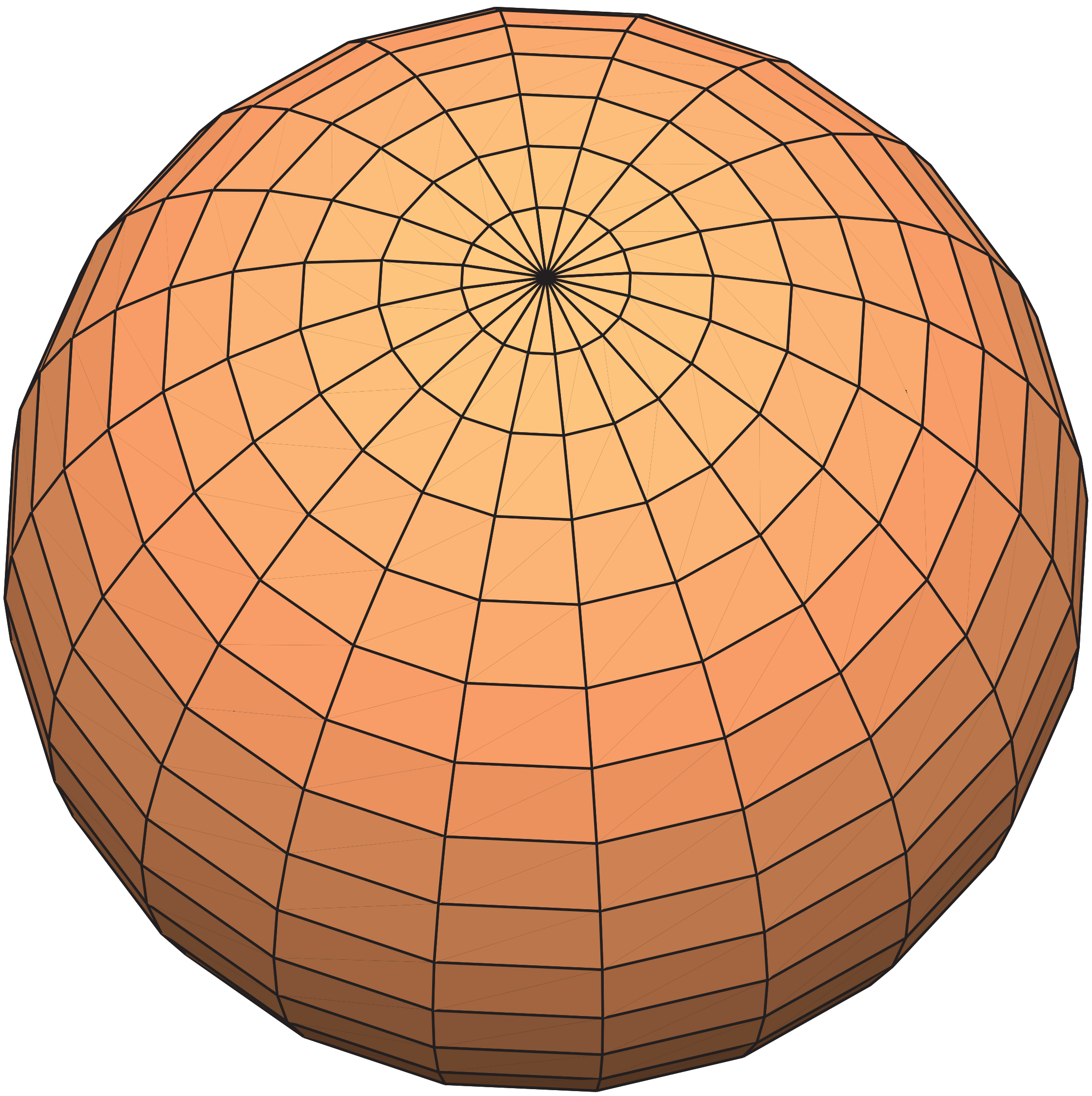
Why Meteorologists Love Doughnuts

- Meteorologists wish the earth were a bagel, too. One could then lay down a latitude-longitude grid with near-uniform grid squares.
- A spectral method could use a double Fourier series, and Fast Fourier Transform in both latitude and longitude.

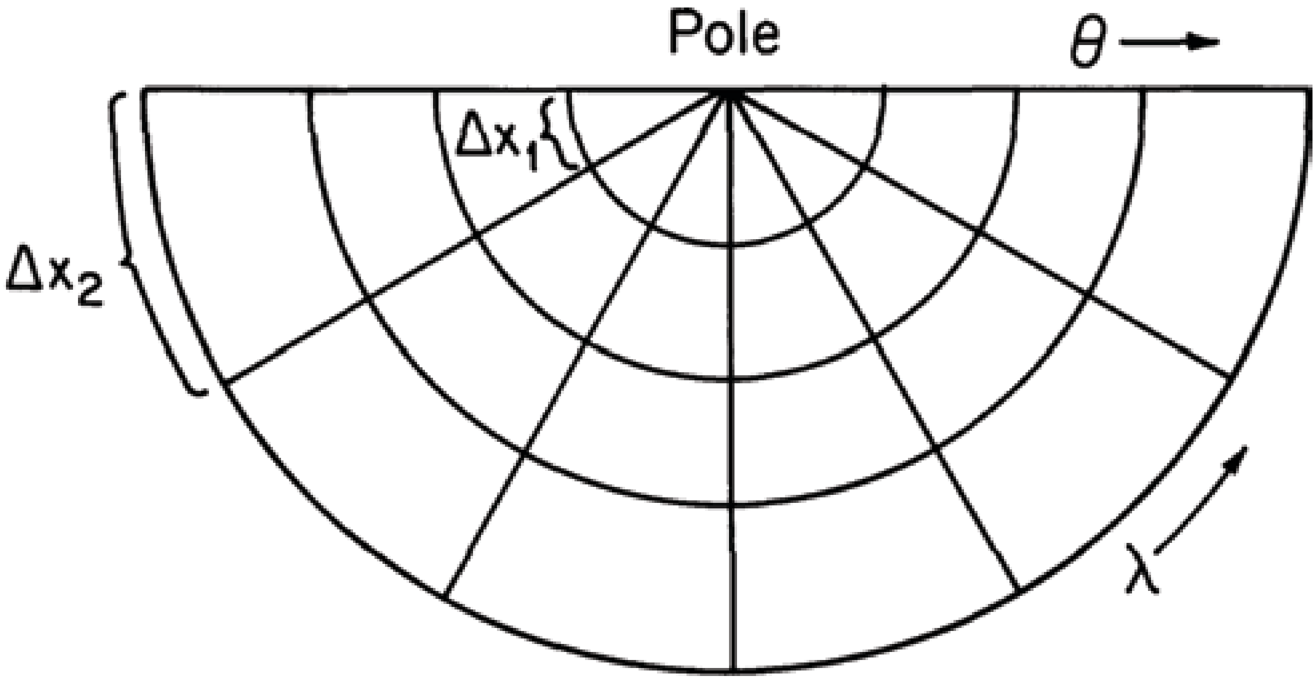


The "Pole Problem"

- The meridians of a latitude-longitude grid converge at both poles



- Because $\delta\lambda \to 0$, CFL limit goes to ZERO.
- Any reasonable timestep triggers instability at the poles, which spreads like a cancer to the entire globe.



- Finite difference models use STRONG POLAR FILTERING
- One big trend is the development of global models free from the "POLE PROBLEM"

Spherical Harmonics/Semi-Implicit: The Old Reliable

- Spherical harmonic kills the pole problem
- In a triangular truncation T_N, zonal wavenumber m, $[\exp(im\lambda)]$, is paired with $N-m$ latitudinal basis functions
- HIGH m spherical harmonics are STABLE because they have EXPONENTIALLY LITTLE AMPLITUDE near the POLES

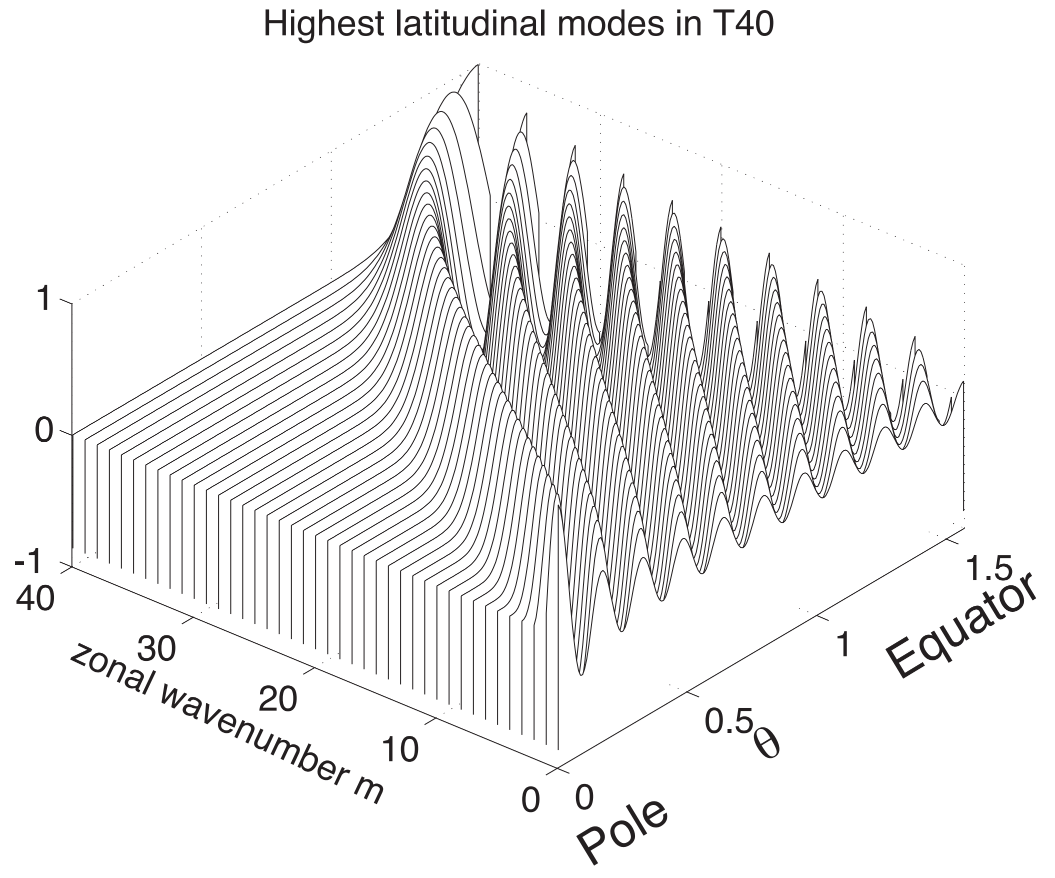


Figure 6: Waterfall plot of the highest latitudinal modes for a triangular truncation of 40.

Semi-Implicit Time-Integration

- Because the $Y_n^m(\lambda, \theta)$ are eigenfunctions of the horizontal Laplacian, SEMI-IMPLICIT time-marching is TRIVIAL
- Three-dimensional Poisson equation is therefore broked into N^2 one-dimensional BVPs in height only.
- Spherical harmonics are combined with FINITE DIFFERENCES in HEIGHT; 1D BVPs discretize as tridiagonal matrices.
- Why semi-implicit?

  IMPLICIT SLOWS the PHYSICS

  Advection must be accurate, so no reason not to treat it EXPLICTLY

  Gravity waves are FILTERED by INITIALIZATION and FAST; propagation described by LINEAR terms; no reason not to treat IMPLICITLY

Why Spherical Harmonics Are (Sort of) Obsolete

Legendre Transforms in Latitude for T_N truncation [N^2 basis functions] cost $O(N^3)$

Why Have Spectral Methods Hung On?

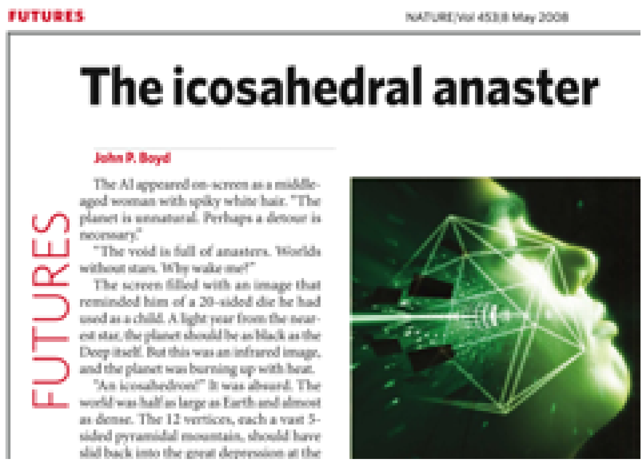
- Mid-life upgrade: SEMI-LAGRANGIAN ADVECTION
  with interpolation of trajectories by LOW DEGREE POLYNOMIALS
  (not SPECTRAL)
  (fringe benefit: fewer transforms)
- Cost of upgrades to the PHYSICS has risen FASTER THAN LINEAR in grid points, too

Current Status

- Japanese Meteorological Model is T959
  [about 25 km between grid points,
  about 1,000,000 points per vertical level]
- Mild ill-conditioning for $N > 2000$ or so.
- Replacements: "Let a Thousand Flowers Bloom!"

Icosahedrons & Triangles

Writers have imagined icosahedral planets, too.



FUTURES NATURE|Vol 453|8 May 2008

The icosahedral anaster

John P. Boyd

FUTURES

The AI appeared on-screen as a middle-aged woman with spiky white hair. "The planet is unnatural. Perhaps a detour is necessary."

"The void is full of anasters. Worlds without stars. Why wake me?"

The screen filled with an image that reminded him of a 20-sided die he had used as a child. A light year from the nearest star, the planet should be as black as the Deep itself. But this was an infrared image, and the planet was burning up with heat.

"An icosahedron!" It was absurd. The world was half as large as Earth and almost as dense. The 12 vertices, each a vast 5-sided pyramidal mountain, should have slid back into the great depression at the

The sphere can be triangulated by subdividing an icosahedron and projecting the result onto the sphere.

German DWD global model use a triangular finite element code.

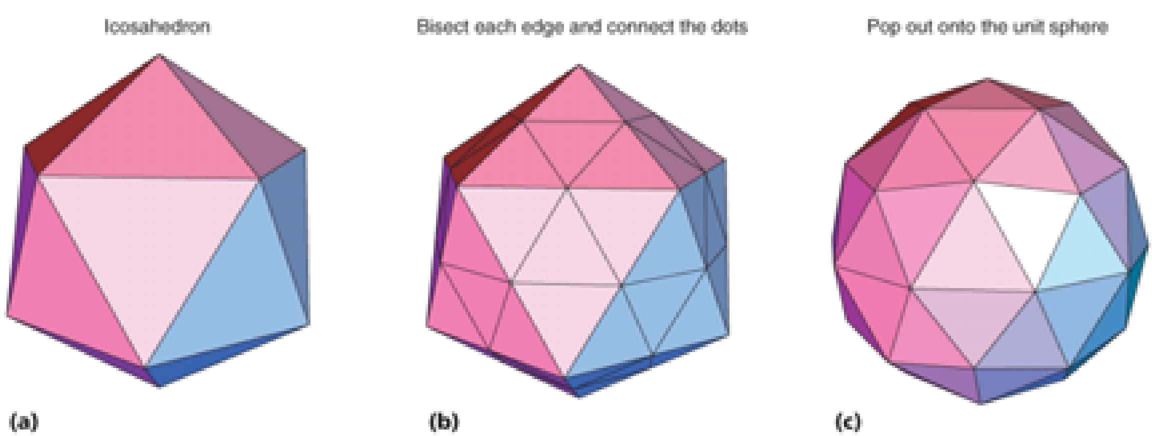


Figure 7: From Randall, Ringler, Heikes, Jones and Baumgardner (2003).

Hexagons and Pentagons

- The triangles can be grouped into hexagons, as in NIKAM and the CSU GCM.
- But the triangularization must always contain 12 pentagons, as first discovered in the *Challenger* oceanographic expedition in the 1870's from the hexagon-and-twelve-pentagon skeletons of *radiolaria*.

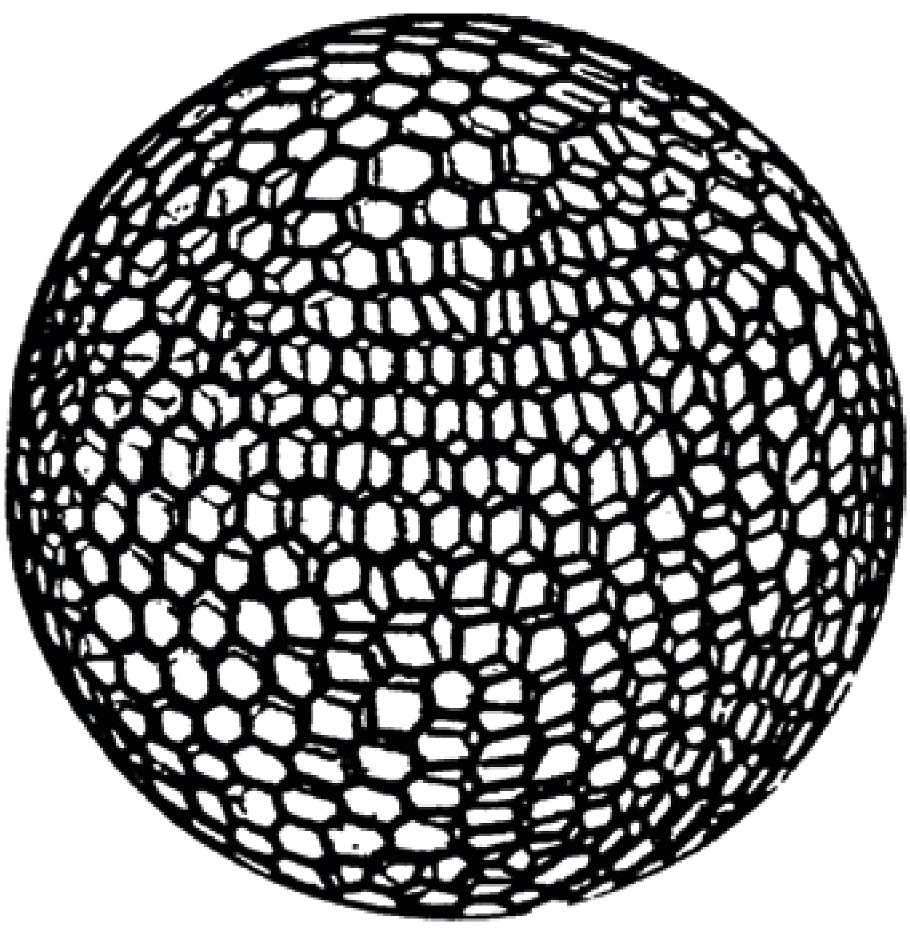


Figure 8: The silica skeleton of the *Radiolaria* diamaton.

The Cubed-Sphere

"The best way to handle spherical geometry is through the cube" — Mark A. Taylor (1998)

- A cube can be subdivided into rectangles, and the result projected on the sphere.
- The "cubed-sphere" has become standard for atmospheric spectral element models, and popular for finite volume methods, too.
- The blocks can vary in size for adaptation, as shown below.

Figure 9: Adapted from St-Cyr, Jablonowski, Dennis, Tufo and Thomas (2008).

Adaptive Mesh Refinement (AMR)

- Nested ("MACROADAPTIVE") models are routine for

  Hurricane/typhoon forecasting

  Regional/Local Area Models (LAM)

- Frontier: Is "MICROADAPTATION" with complicated fine grid topology useful?

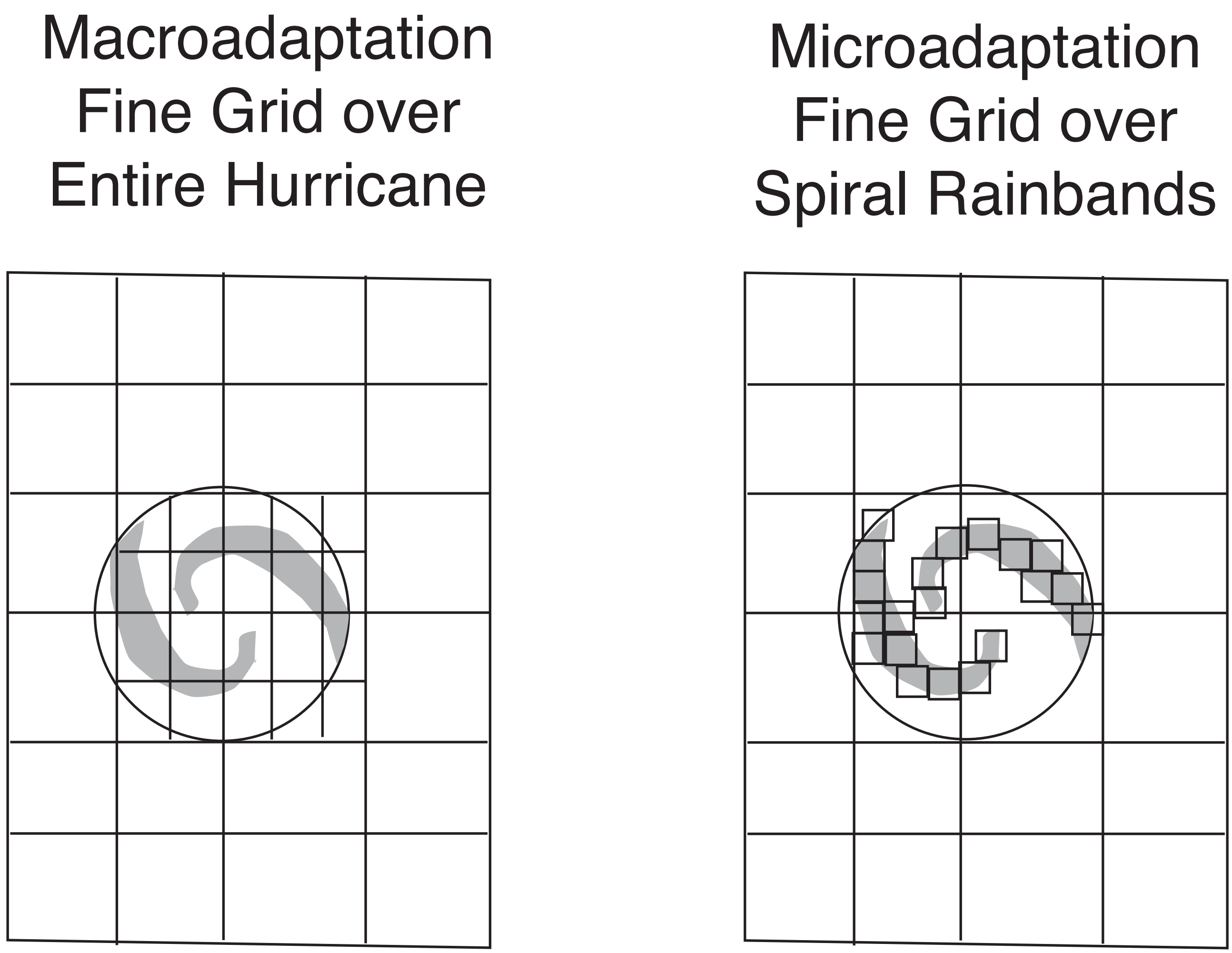


Figure 10: Vorticity gradient at $t = 25$. Where should the fine grid go?

Why AMR is Not a Slam Dunk

1. Non-uniform grid $\Rightarrow$ OVERHEAD, especially on MASSIVELY PARALLEL machines
2. Atmosphere and ocean are very WAVY
3. Complicated “physics” responds badly and unpredictably to RESOLUTION CHANGES

Candace Parker has more dunks than any other woman in college basketball history.

Figure 11: Candace Parker dunking in college.

Adaptive Mesh Refiner's Prayer:
Please God, No Waves!

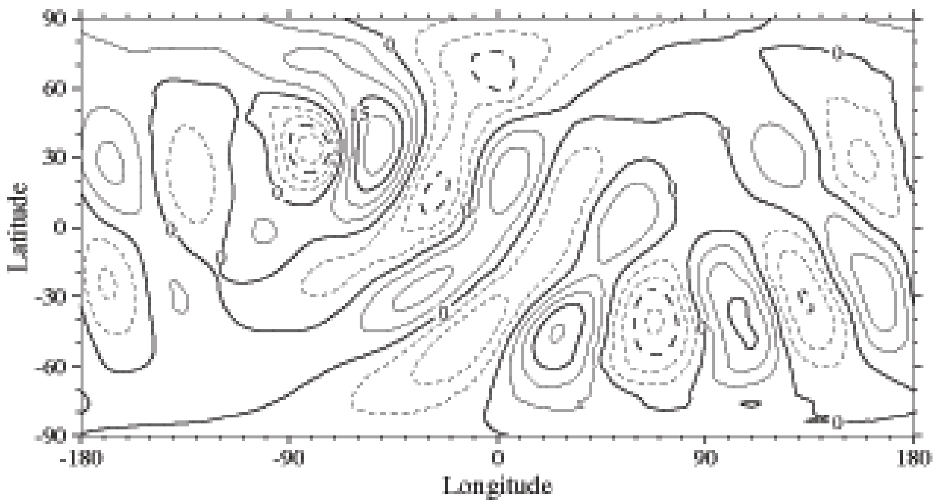


Figure 12: Isolines of geopotential for flow over a mountain (whose crest is marked by the black dot at 30^o N., 90^o W.). Image and adaptive model by Christiane Jablonowski.

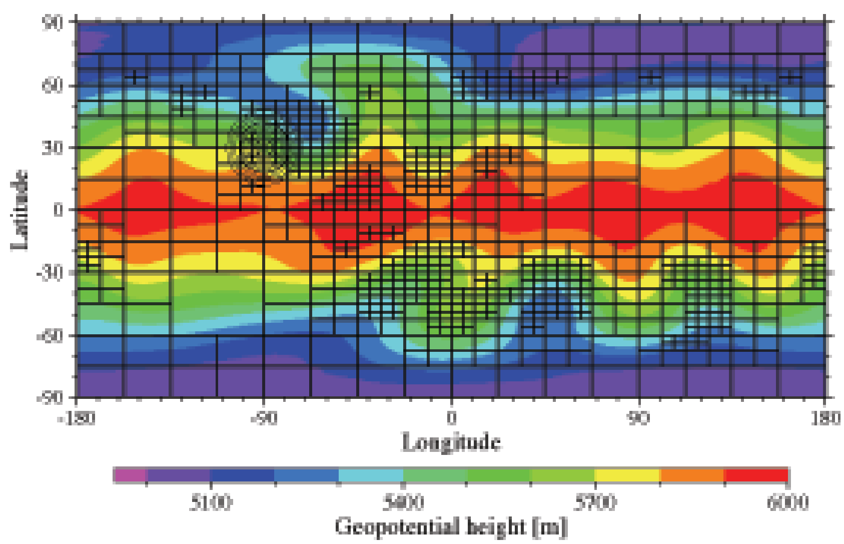


Figure 13: Same but showing the geopotential as false color with superimposed adaptive grid.

Spurious Reflections at Coarse-Fine Boundaries

"A compatibility problem exists at the interface where the two grid systems meet. For instance, a disturbance propagating from a coarse grid mesh (CGM) to the FGM may undergo false reflection back to the CGM and scattering into the FGM. On the other hand, a disturbance propagating from the FGM to the CGM may also experience false reflection back onto the FGM or aliasing as it enters the CGM. These interface-generated problems may lead to numerical instabilities that can seriously affect the forecast over the entire domain. This is known as the interface condition problem." pg. 1330

"An optimal interface procedure that eliminates this problem should have the following properties:

1. all resolvable waves propagate across interfaces smoothly with only minimal changes in amplitude and minimum reflection of energy and

2. mass, momentum and total energy exchanged between the two grid systems should be conserved.

Because of the numerical difficulties involved in the design of a nested grid, procedures which fully satisfy the foregoing two requirements have yet to be described in the literature." pg. 1330

From pg. 1330 of their article by Da-Lin Zhang, Hai-Ru Chang, Nelson L. Seaman, Thomas T. Warner and J. Michael Fritsch, *Monthly Weather Review*, **114**, 1330–1339 (1986).

Relaxation Zones

- A common strategy to minimize "resolution-change shock" is to include a "relaxation zone" around a nested, high resolution Limited Area Model where the coarse and fine grid solutions are SMOOTHLY BLENDED together
- Empirically effective, but mathematical justification is shaky and blending methods are crude

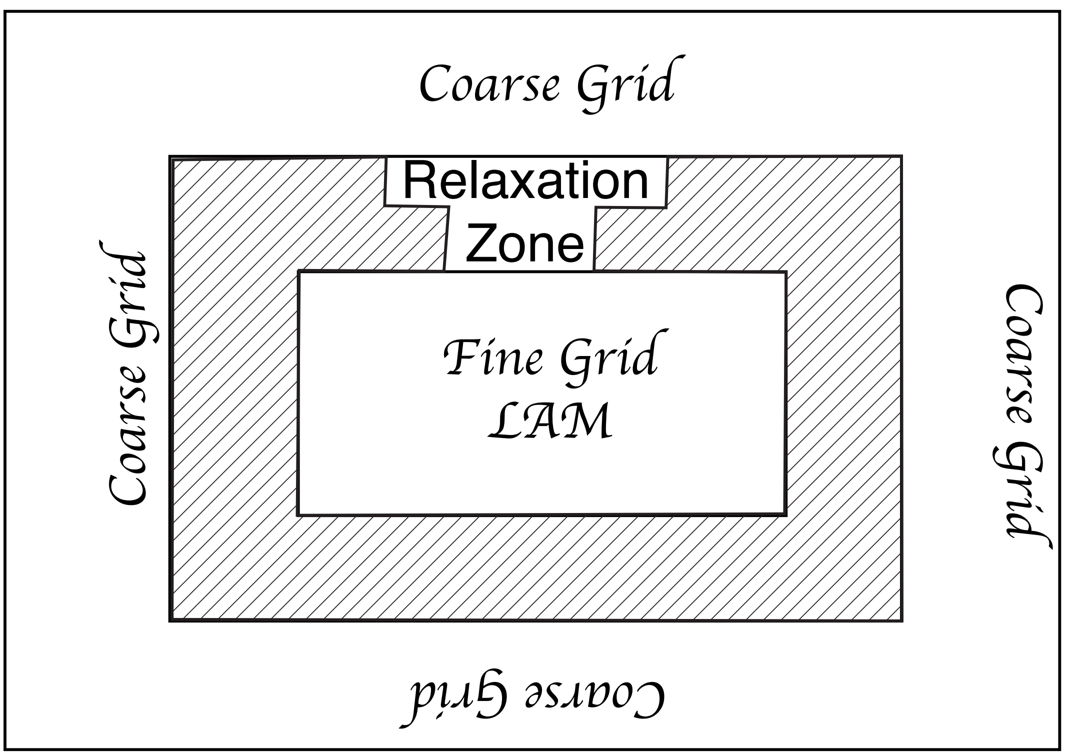
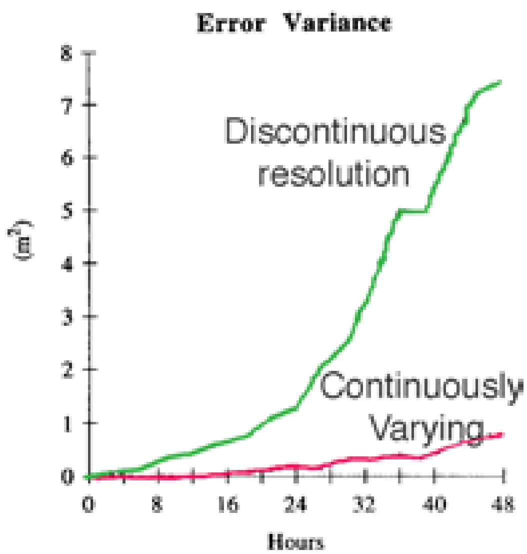


Figure 14: A hill much smaller than a grid box can trigger "orographic rain".

Continuously-Varying Resolution

- Microadaptation usually involves DISCONTINUOUS JUMPS in RESOLUTION
- Nested models can be blended with global models through a SMOOTH MAP of the SPHERE to the SPHERE, varying the RESOLUTION CONTINUOUSLY



- Gravel & Staniforth (1992) showed that continuously-varying resolution is best
- Higher resolution reduces grid-jump reflection: less energy in the short waves that feel the jump. [Long waves, resolved on both coarse and fine, are not reflected].

Geophysical AMR Prayer to St. Jude:
Deliver Us from the Physics

Example: moist convection parameterization

- Physical law: condensation only at 100 % relative humidity
- Meteorological reality: Local wet spots, such as "orographically-triggered showers", when the AVERAGE humidity is only 90 %

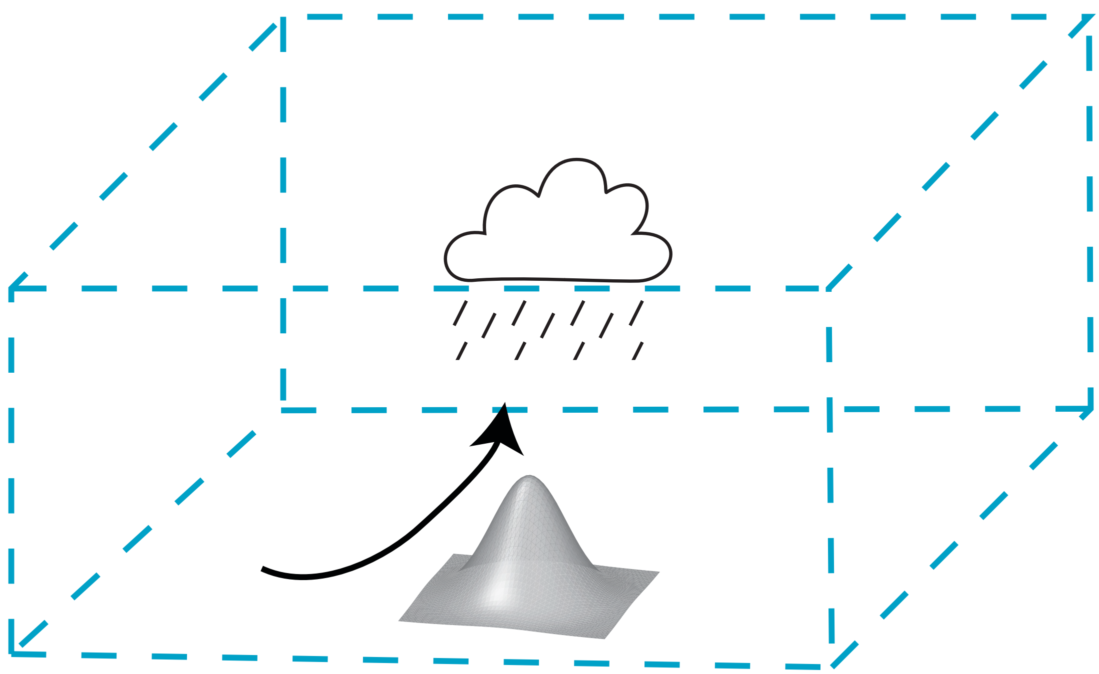
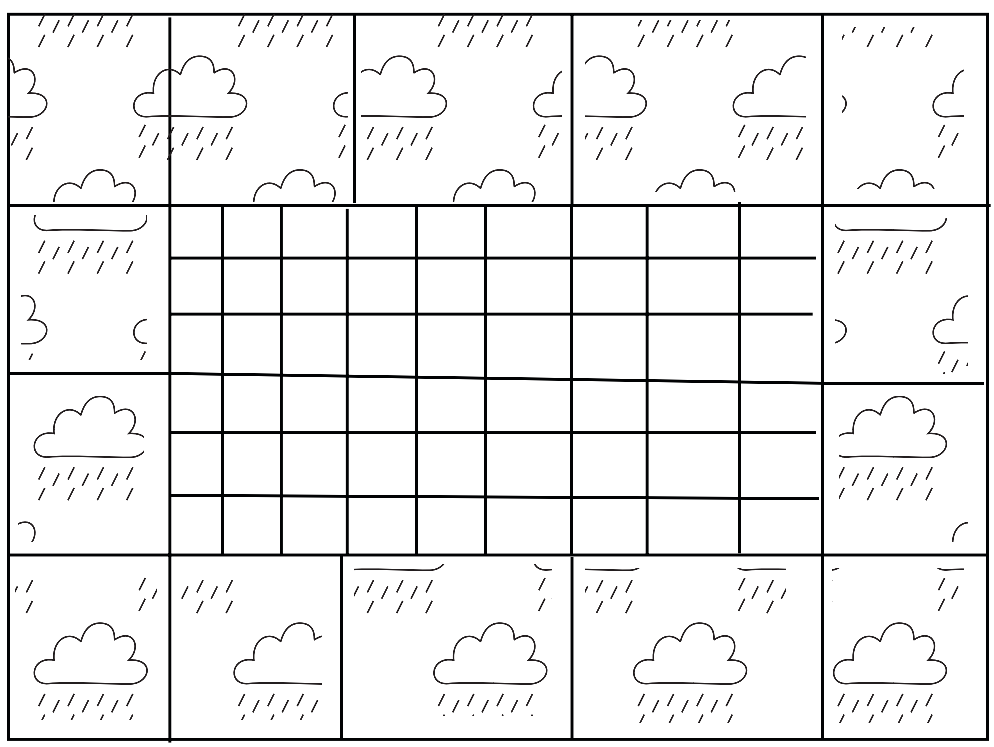


Figure 15: A hill much smaller than a grid box can trigger "orographic rain".

Moist Convective Parameterization -2

- Parameterization triggers condensation when $q = q_{trigger}$
- $q_{trigger}$ should be LARGER for SMALLER boxes
- Bad choices can give a HOLLOW RECTANGLE of RAIN in COARSE GRID BOXES surrounding a cluster of FINE BOXES which are CLOUD-FREE



"A more fundamental concern is the validity and consistency of grid-scale-dependent parameterizations , particularly in nested-grid and adaptive calculations... convective parameterizations will produce different convective realizations (time, location, and extent) solely due to a change in grid scale (Zhang and Fritsch, 1988). It is also unclear at what grid scale the convective parameterizations become unnecessary and even detrimental." — William Skamarock & Joseph Klemp pg. 798 of their article in *Monthly Weather Review*, **121**, 788–204 (1993).

- To minimize overhead, adaptive models are often low order.

"We ... show that the requirements on the AMR scheme to be cheaper than a high order scheme are unrealistic for most computational scenarios." — Leland Jameson in the abstract of his paper "AMR vs High Order Schemes", *J. Sci. Comput.*, **18**, no. 1, 1–24 (2003).

- Spectral elements are HIGH ORDER & ADAPTIVE
- Parameterization problem remains
- Physics of the atmosphere is strongly MULTI-SCALE, so AMR will be a big research area.

Challenges

1. Finite Volume vs. Spectral vs. Spectral Element vs. Finite Difference
2. Best Nearly-Uniform Grid for the Sphere
3. Macro-adaptation ("Nesting") versus Micro-adaptation
4. Importance of Discrete Conservation of Mass, Energy, Enstropy, Etc.
5. Better subgridscale parameterizations in both "dycore" and "physics"
6. Better ocean/lower atmosphere/upper atmosphere coupling
7. Is theory dead?
8. Incorporating PDE Compatibility Conditions
9. Steep topography: the mountainous difficulties of mountains

References

Boyd, J. P., 1981: The Moonbow, *Isaac Asimov's Science Fiction Magazine*, **5**, No. 5 (May 11), 18-37.

Boyd, J. P., 1994: The slow manifold of a five mode model, *J. Atmos. Sci.*, **51**, 1057–1064.

Boyd, J. P., 1998: *Weakly Nonlocal Solitary Waves and Beyond-All-Orders Asymptotics*, Kluwer, Amsterdam, 608 pp.

Boyd, J. P. and N. Flyer, 1999: Compatibility conditions for time-dependent partial differential equations and the the rate of convergence of Chebyshev and Fourier spectral methods", *Comput. Meths. Appl. Mech. Engrg.*, **175**, 281–309. Errata: in Eq.(22), the square root should be in front of the integral, not in the exponential.,

Boyd, J. P., 2001: *Fourier and Chebyshev Spectral Methods*, 2d. ed., Dover, Mineola, New York, 660 pp.

Boyd, J. P., 2008: Icosahedral Anaster, *Nature*, **453**, 8 May, 256.

Conaty, A. L. and J. C. Jusem, L. Takacs, D. Keyser and R. Atlas, 2001: The structure and evolution of extratropical cyclones, fronts, jet streams and tropopause in the GEOS general circulation model. *Bull. Amer. Meteor. Soc.*, **82**, 1853–1867.

Daley, Roger, 1993: *Atmospheric Data Analysis.* Cambridge University Press, 471 pp. [data assimilation]

Giraldo, F. X., and M. Restelli: 2008: A study of spectral element and discontinuous Galerkin methods for the Navier-Stokes equations in nonhydrostatic mesoscale atmospheric modeling: Equation sets and test cases. *J. Comput. Phys.*, **227**, 3849–3877.

Gravel, S. and A. Staniforth, 1992: Variable resolution and robustness. *Monthly Weather Rev.*, **120**, 2633–2640.

Haidvogel, D. B., and A. Beckmann, 1999: *Numerical Ocean Circulation Modeling*, World Scientific Publishing Company, 300 pp.

Hesthaven, J. S., and T. Warburton, 2007: *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications* , Springer, 500pp.

Jablonowski, C., and M. Herzog and J. E. Penner, R. C. Oehmke, Q. F. Stout, B. van Leer, and K. G. Powell, 2006: Block-structured adaptive grids on the sphere: Advection experiments. *Monthly Weather Rev.*, **134**, 3691–3713.

Jameson, L., 2003: AMR vs High Order Schemes, *J. Sci. Comput.*, **18**, no. 1, 1–24 (2003).

Kalnay, E., 2002: *Atmospheric, data assimilation and predictability.* Cambridge University Press, 364 pp.

Krishnamurti, T. N., 1995: Numerical weather prediction. *Ann. Revs. Fluid Mech.*, **27**, 195–224.

Lorenz, E. N., 1963: Deterministic on-periodic flow. *J. Atmos. Sci.*, **20**, 130–141.

Lynch, P., 2008: The origins of computer weather prediction and climate modeling. *J. Comput. Phys.*, **227**, 343–3444.

Nair, R. D., and C. Jablonowski, 2008: Moving vortices on the sphere: A test case for horizontal advection problems. *Monthly Weather Review*, **136**, 699–711.

Skamarock, W. C., and J. B. Klemp, 1993: Adaptive grid refinement for 2-D and 3-D nonhydrostatic atmsopheric flow. *Monthly Weather Review*, **121**, 788–204.

Zhang, D.-L., and H.-R. Chang, N. L. Seaman, T. T. Warner and J. M. Fritsch, 1986: A two-way interactive nesting procedure with variable terrain resolution. *Monthly Weather Review*, **114**, 1330–1339.

Definition 3 (Mokita) *"Mokita" is a word in a language of Papua New Guinea for "knowledge that we all know but agree never to talk about".*

John P. Boyd Dept. of Atmospheric, Oceanic and Space Science University of Michigan 2455 Hayward Avenue Ann Arbor MI 48109-2143 jpboydATumich.edu