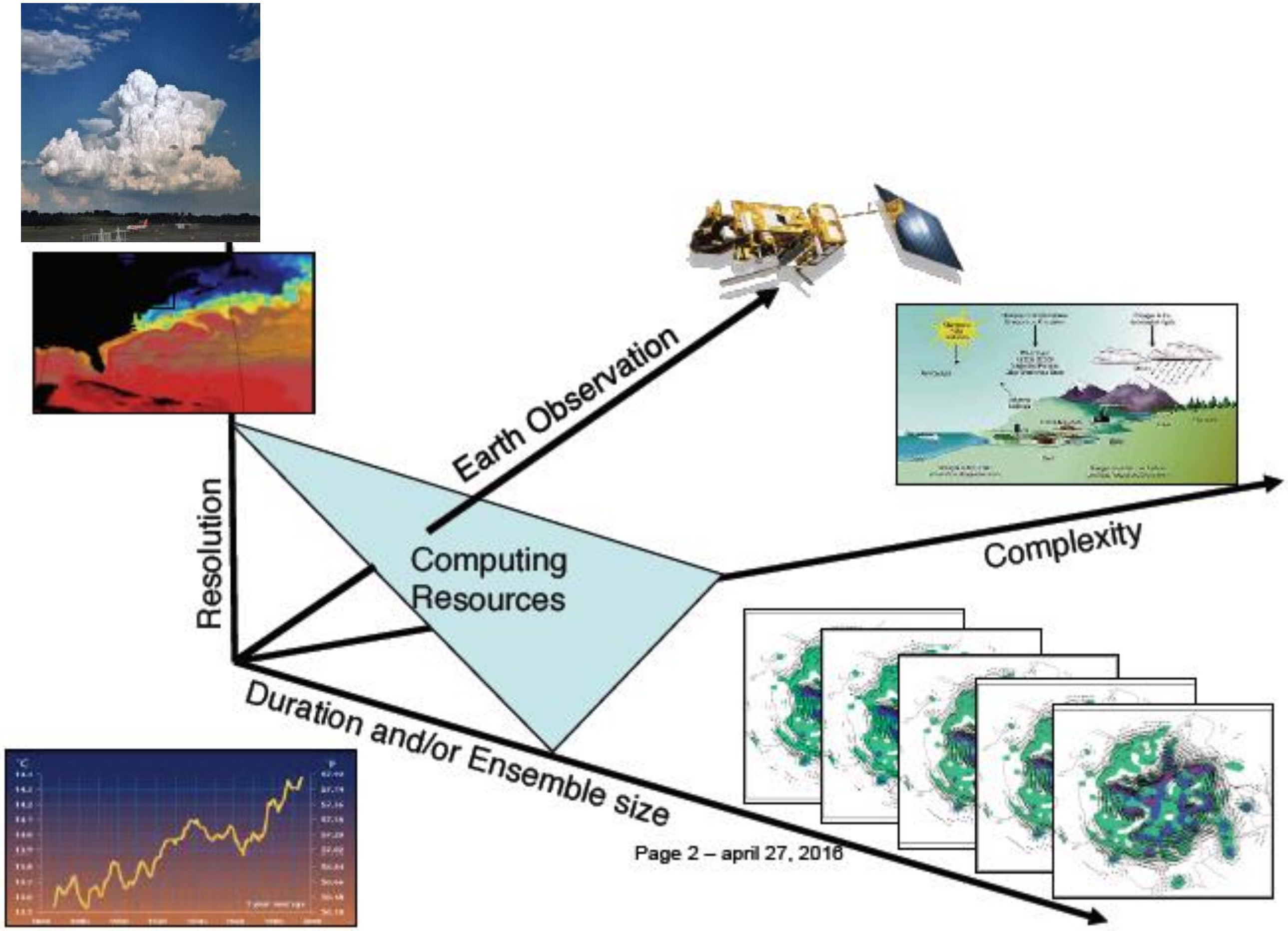
Imprecise Complexity

Tim Palmer, Hannah Christensen, Andrew Dawson, Peter Düben, Stephen Jeffress, Stephan Juricke, Dave Macleod, Kristian Strommen, Tobias Thornes, Antje Weisheimer

Atmospheric, Oceanic and Planetary Physics, University of Oxford


Earth Observation
Complexity
Resolution
Computing
Resources
Duration and/or Ensemble size
Page 2 – april 27, 2016

Strategy

Update July 2016: The Council has approved the new strategy for the period 2016–2025.

Goals by 2025

To provide forecast information needed to help save lives, protect infrastructure and promote economic development in Member and Co-operating States through:

Research at the frontiers of knowledge to develop an integrated global model of the Earth system to produce forecasts with increasing fidelity on time ranges up to one year ahead. This will tackle the most difficult problems in numerical weather prediction such as the currently low level of predictive skill of European weather for a month ahead.

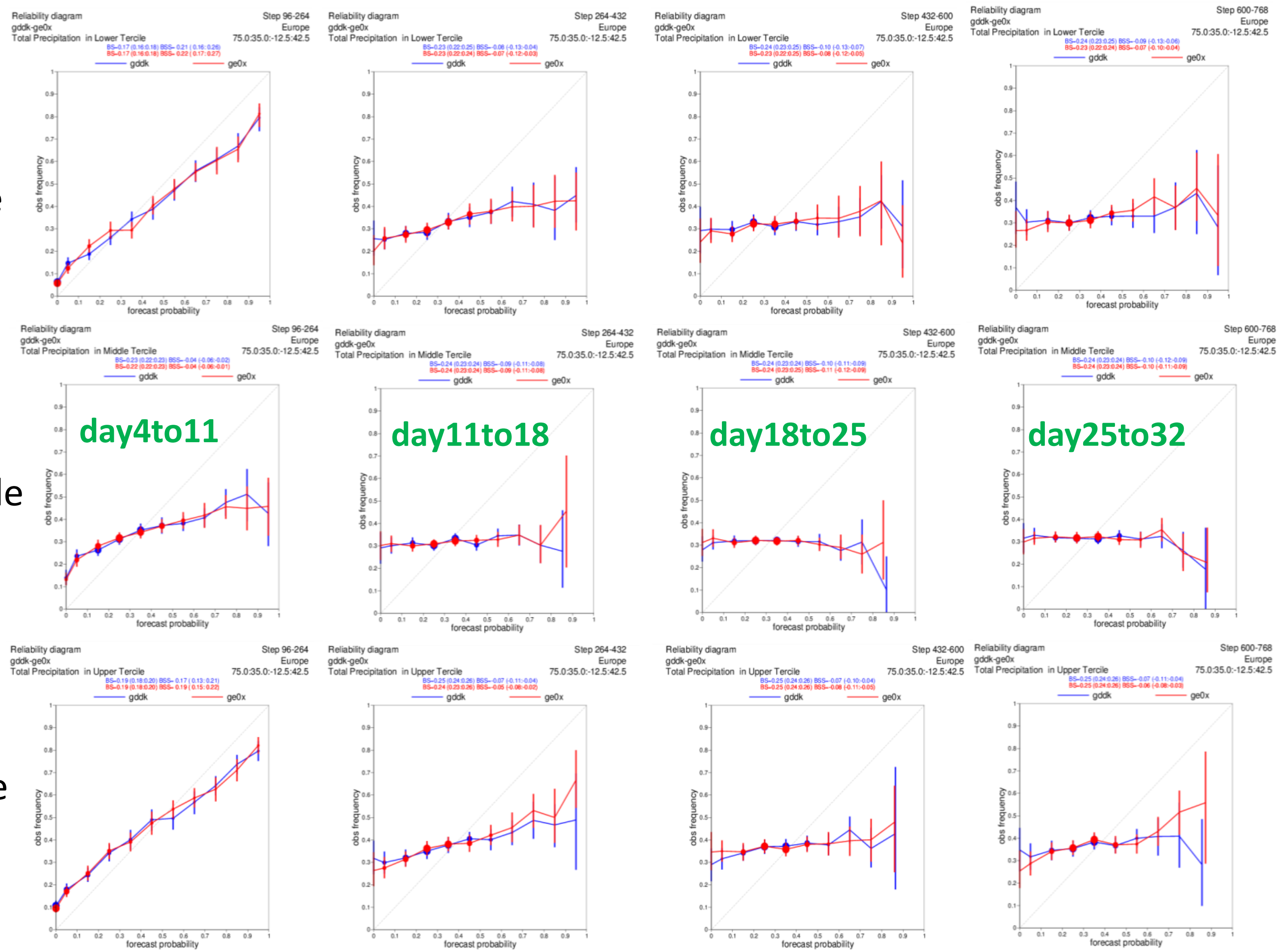
Operational ensemble-based analyses and predictions that describe the range of possible scenarios and their likelihood of occurrence and that raise the international bar for quality and operational reliability. Skill in medium-range weather predictions in 2016, on average, extends to about one week ahead. By 2025 the goal is to make skilful ensemble predictions of high-impact weather up to two weeks ahead. By developing a seamless approach, we also aim to predict large-scale patterns and regime transitions up to four weeks ahead, and global-scale anomalies up to a year ahead.

Antje Weisheimer

lower tercile

middle tercile

upper tercile



Reliability of precip forecasts over Europe in the monthly forecasting system (T399-T255)

PV Gradient Across Tropopause

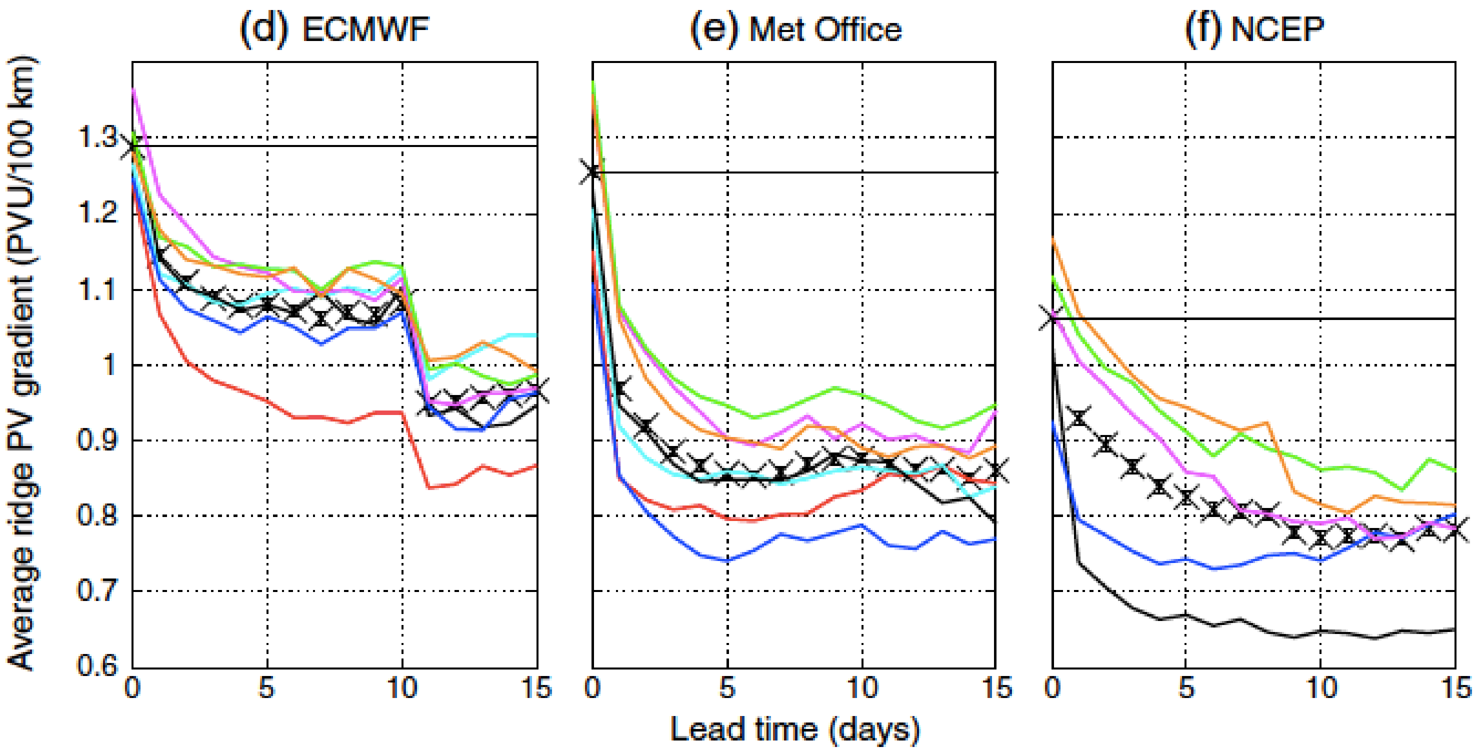


Figure 5. (a–c) Average ridge area and (d–f) the isentropic PV gradient flanking ridges as a function of forecast lead time for ECMWF, Met Office, and NCEP. Black markers with error bars (standard errors) are averages over all winter seasons with horizontal lines extending across all lead times from the analysis values. Colored lines are averages for the individual seasons where red is 2006/2007, cyan is 2007/2008, black is 2008/2009, blue is 2009/2010, magenta is 2010/2011, green is 2011/2012, and orange is 2012/2013. Note (as an example) that a fraction of the Northern Hemisphere of 0.05 is equivalent to an area of 1.275×10^7 km^2.

Gray et al, GRL, 2014

Harvey, Methven and Ambaum JFM 2016.

RW propagation – smooth PV step

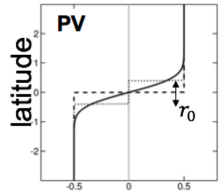
www.ncas.ac.uk

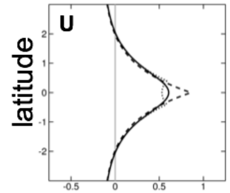
$$c_m \approx \int (U_s(y) - \phi_s(y))w(y)\,dy$$

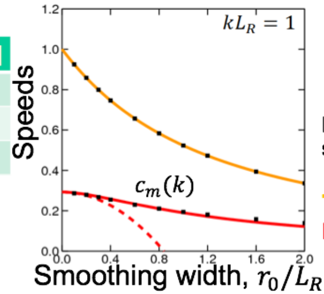
- Putting some numbers in…
  - $r_0 = \dfrac{\Delta PV}{\max PV_y} \approx \begin{cases} 308\ km\ (analysis) \\ 381\ km\ (forecast) \end{cases}$
  - Also use: $L_R = 700\ km\ , \Delta PV = 4PVU$
  - Result:

	U_{max} [m/s]	c_m [m/s]	$c_{m,g}$ [m/s]
Sharp PV front:	70.0	20.5	45.3
$r_0 = 308\ km$:	50.9	17.7	38.2
$r_0 = 381\ km$:	47.6	16.8	36.0

  - Typical phase error after 5 days = O(400km)





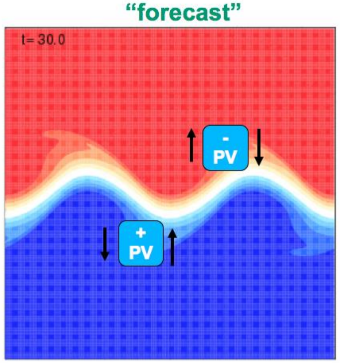


Amplitude error – what's its cause?

www.ncas.ac.uk

- Advection on the jet flank leads to a systematic accumulation of PV
  - The induced circulation acts to damp the Rossby wave
- Consistent with meridional dispersion of wave activity
- Diffusion at small scales acts to damp large scale waves via systematic accumulation of PV relative to the ridges and troughs

"forecast"



Multiple Flow Equilibria in the Atmosphere and Blocking[1]

JULE G. CHARNEY[2]

Massachusetts Institute of Technology, Cambridge 02139

JOHN G. DEVORE[3]

University of California, Los Angeles 90024

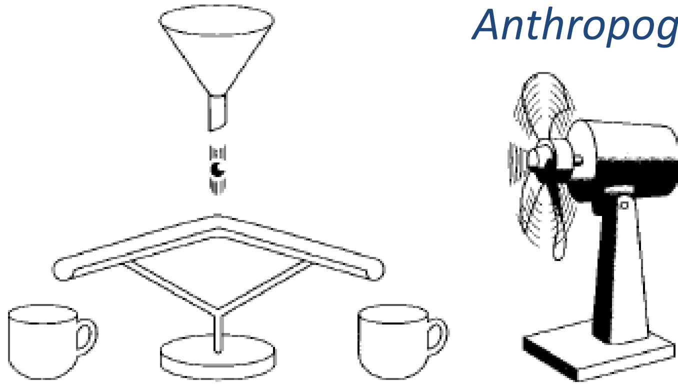
(Manuscript received 22 September 1978, in final form 28 February 1979)

ABSTRACT

A barotropic channel model is used to study the planetary-scale motions of an atmosphere whose zonal flow is externally driven. Perturbations are induced by topography and by a barotropic analogue of thermal driving. The use of highly truncated spectral expansions shows that there may exist a multiplicity of equilibrium states for a given driving, of which two or more may be stable. In the case of topographical forcing, two stable equilibrium states of very different character may be produced by the same forcing: one is a "low-index" flow with a strong wave component and a relatively weaker zonal component which is locked close to linear resonance; the other is a "high-index" flow with a weak wave component and a relatively stronger zonal component which is much farther from linear resonance. It is suggested that the phenomenon of blocking is a metastable equilibrium state of the low-index, near-resonant character. The existence of the two types of equilibria has been confirmed by numerical integration of a grid-point model with many more degrees of freedom than the spectral model.

It has also been found spectrally and for a grid-point model that oscillations may occur when one of the equilibrium states is stable for the lowest order spectral components but unstable for the next higher order components. The oscillation apparently is due to a barotropic instability of the topographic wave of the kind discussed by Lorenz and Gill.

Thermal forcing also produces multiple, stable equilibria in a spectral model but confirmation with a grid-point model has so far not been obtained.



Model

A Nonlinear Dynamical Perspective on Climate Prediction

T. N. Palmer

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom

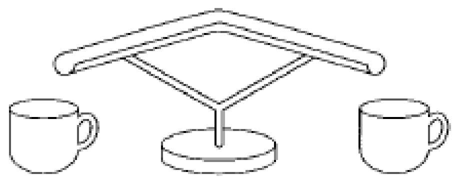
(Manuscript received 7 October 1997, in final form 26 February 1998)

ABSTRACT

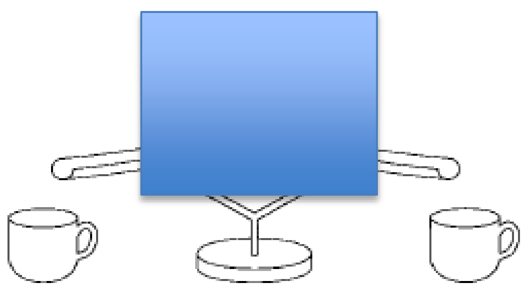
A nonlinear dynamical perspective on climate prediction is outlined, based on a treatment of climate as the attractor of a nonlinear dynamical system *D* with distinct quasi-stationary regimes. The main application is toward anthropogenic climate change, considered as the response of *D* to a small-amplitude imposed forcing **f**.

Model can simulate "Blocked" and "Zonal" regimes. The simulated unforced frequency of occurrence of both regimes is correct. However, the model response to external forcing is *completely* incorrect.

How could we know that the response to forcing was incorrect?



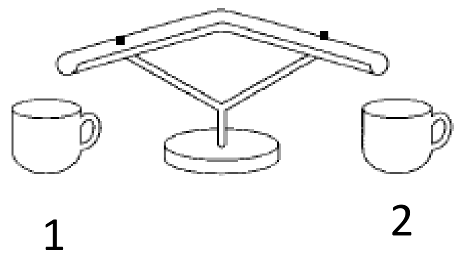
Model

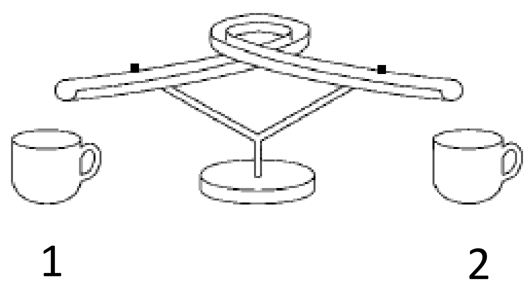


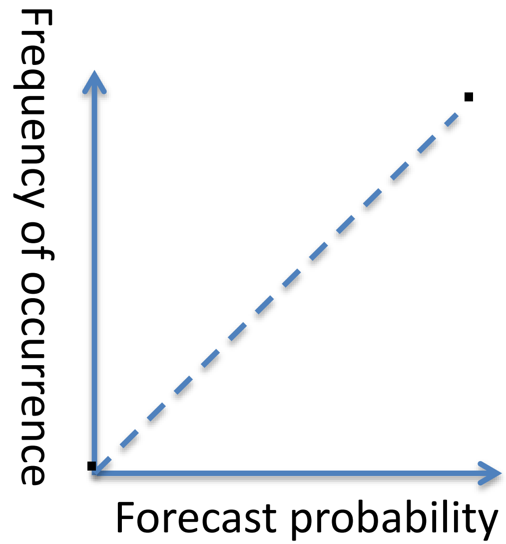
Real World

Look at reliability of initial value predictions.

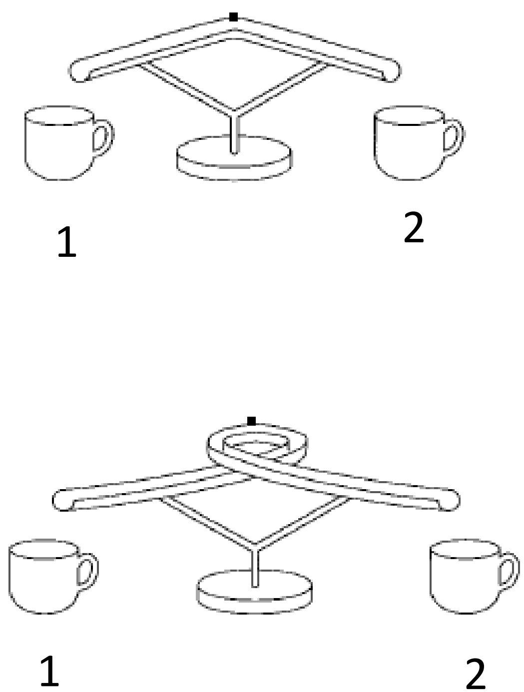
Probability of Occurrence
of Regime 1.

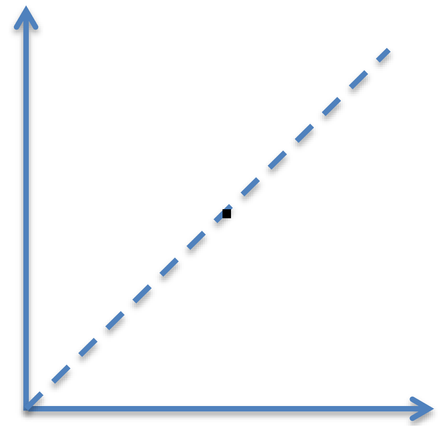




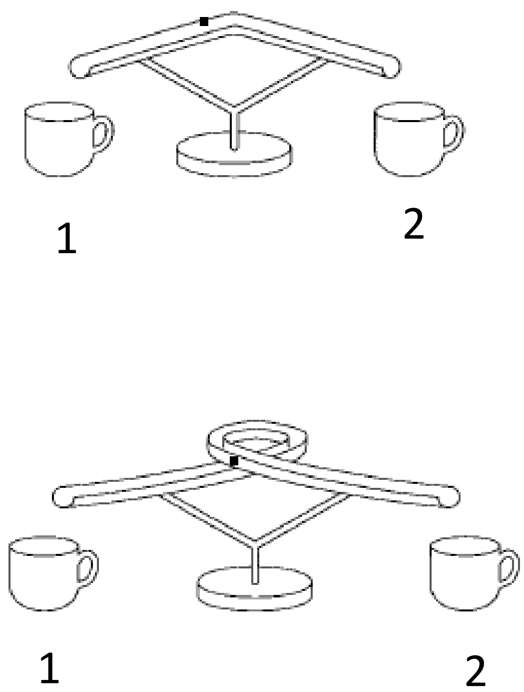


Probability of Occurrence of Regime 1.

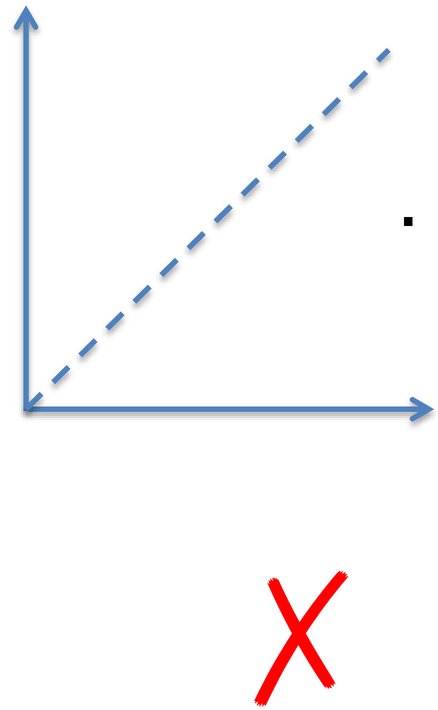




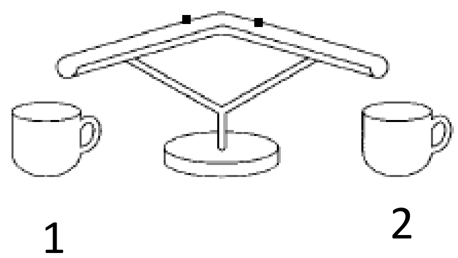
✓

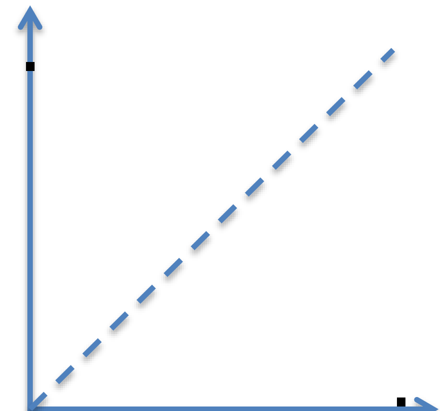


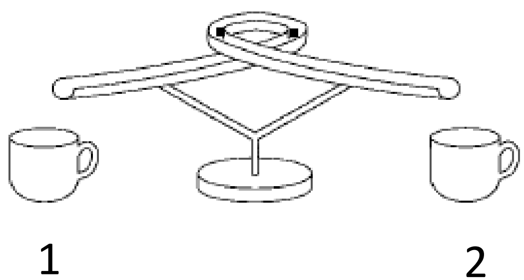
Probability of Occurrence
of Regime 1.



Probability of Occurrence of Regime 1.

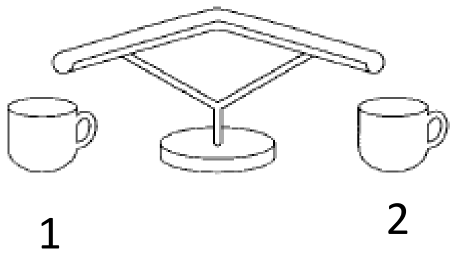


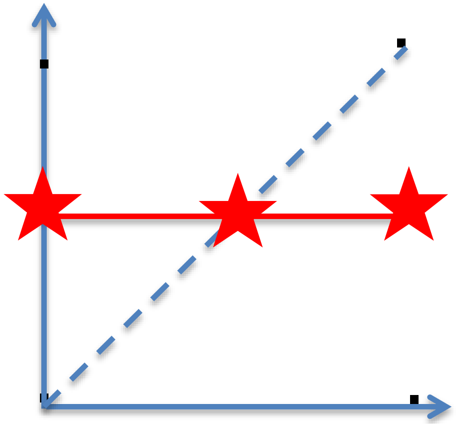


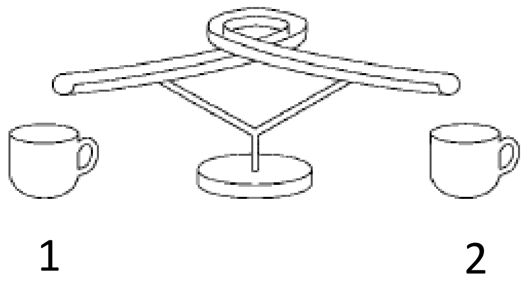


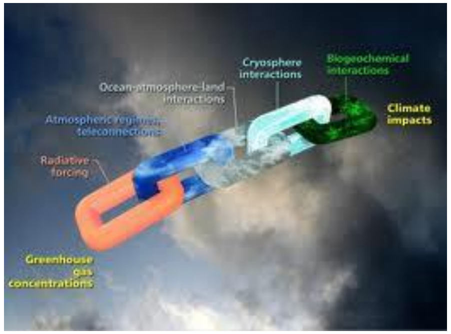
Putting it all together

Probability of Occurrence
of Regime 1.









One cannot be confident that the response to forcing is correct if the corresponding reliability diagram is flat.

Increased resolution likely to be crucial for ECMWF’s strategic goals.

Increased forecast reliability should be considered a key measure for assessing confidence in climate projections on longer timescales

How can this be achieved ***<u>and</u>*** develop Earth-System complexity?

Is there a cheaper way to represent Earth-System complexity?

TECHNICAL MEMORANDUM

598

Stochastic Parametrization and Model Uncertainty

Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer, A. Weisheimer

Research Department

October 8, 2009

This paper has not been published and should be regarded as an Internal Report from ECMWF. Permission to quote from it should be obtained from the ECMWF.

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

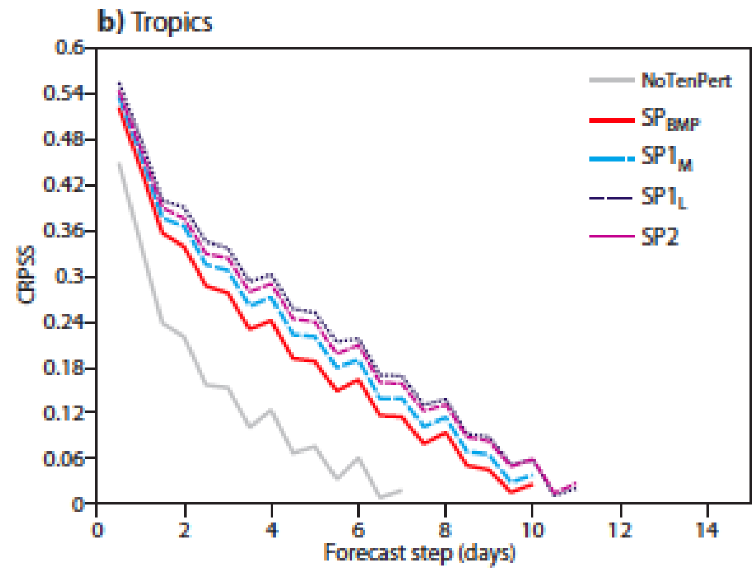


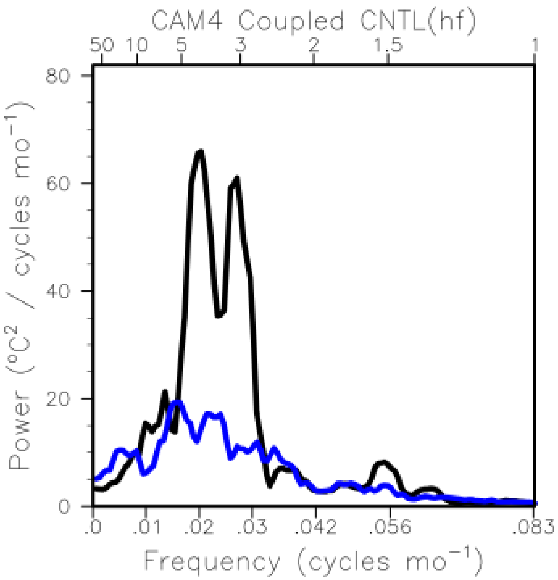
Figure 3: Continuous Ranked Probability Skill Score for 850 hPa temperature.

Technical Memorandum No. 598 7

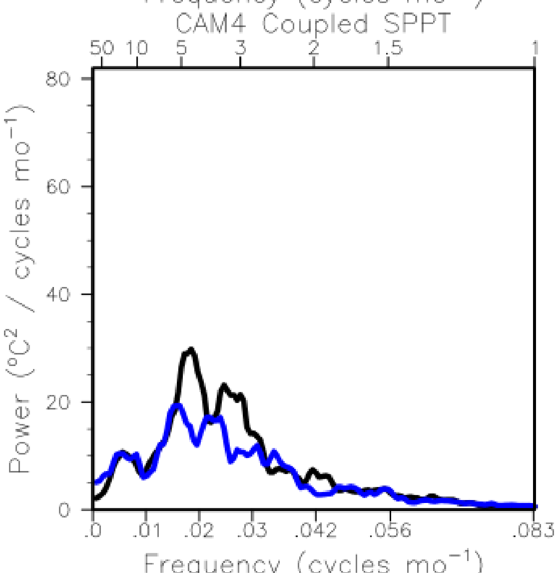
Irreducible Uncertainty in sub-grid representations

Hannah Christensen

Stochastic parametrisation can also improve NCAR climate model El Nino climatology



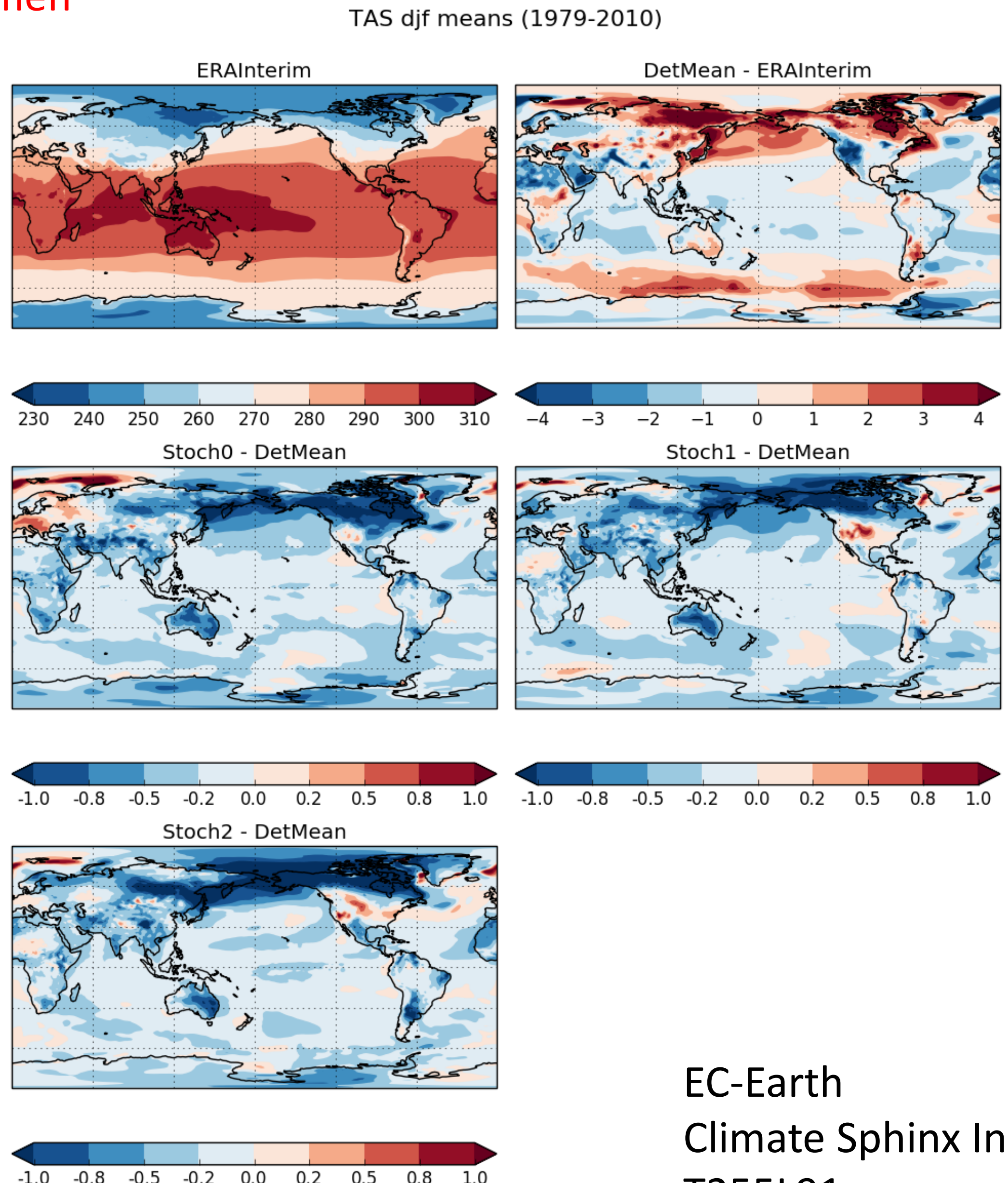
Nino3 variability without SPPT



Nino3 variability with SPPT

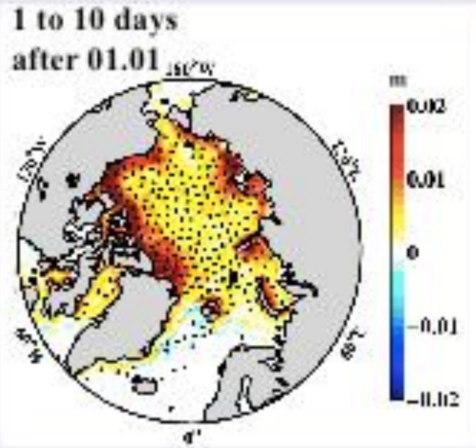
Christensen et al 2016

Kristian Strommen

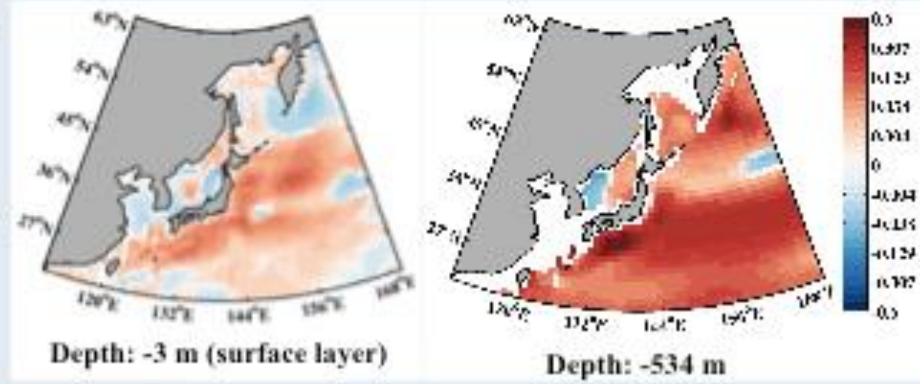


EC-Earth
Climate Sphinx Integrations
T255L91

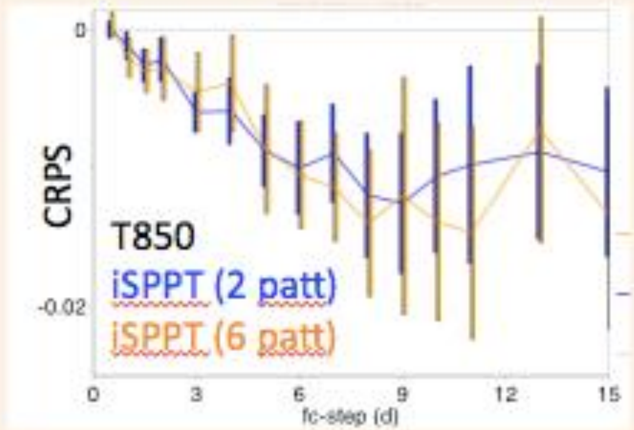
Primavera- Oxford
A Stochastic Package for EC-Earth



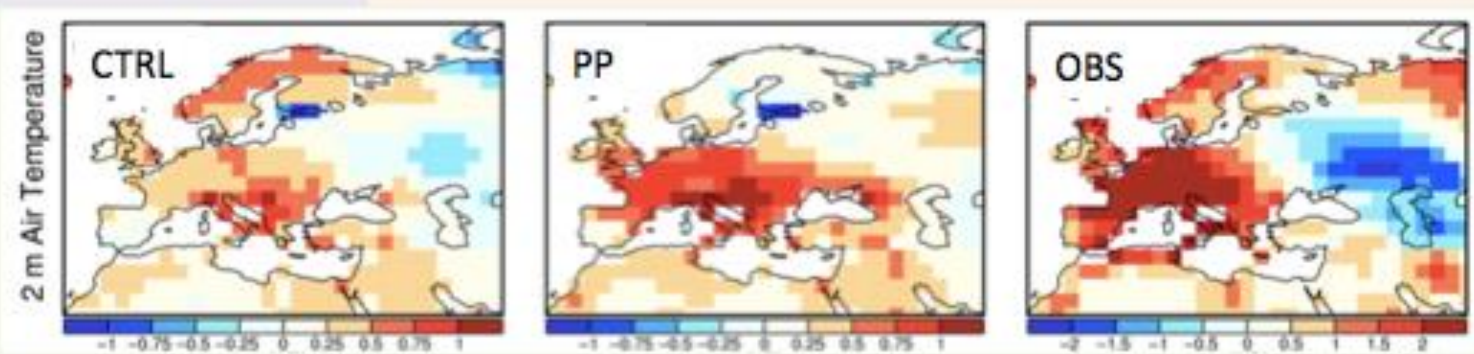
↑ Increase in ensemble spread sea ice thickness (Juricke et al, 2014)



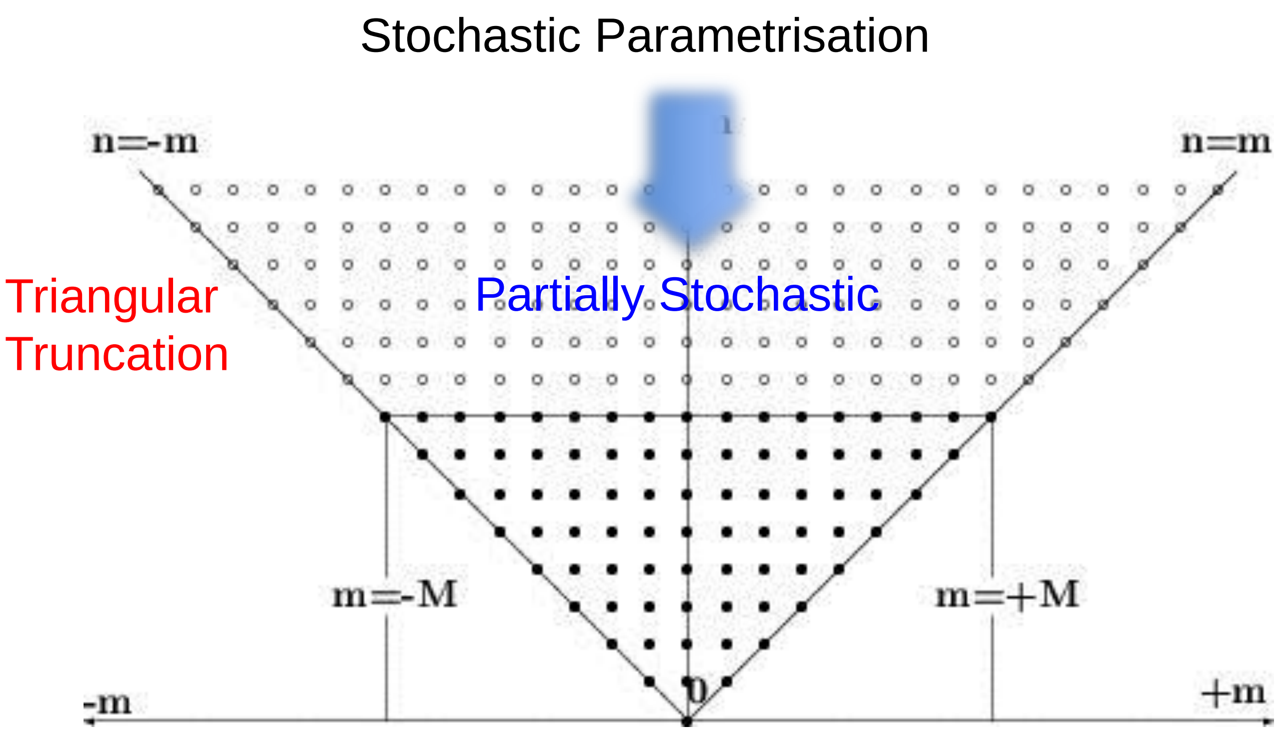
↑ Increase in inter-annual variance in ocean T in Kuroshio (Juricke et al, 2016)



← iSPPT improves CRPS tropical MW forecasts (Christensen et al, in prep)



↑ Perturbing land surface parameters improves simulation of 2003 European heatwave (MacLeod et al, 2016)

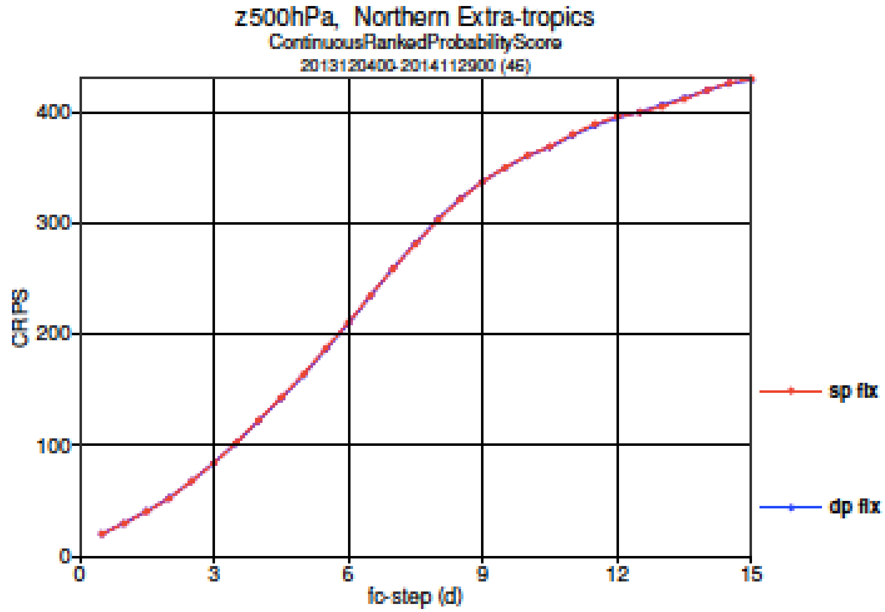


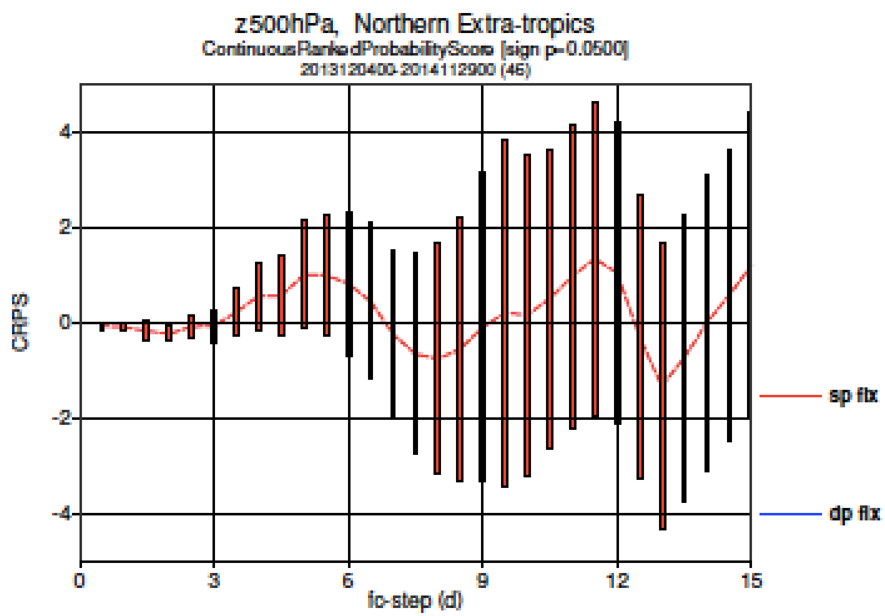
If parametrisation is partially stochastic, are we "over-engineering" our dynamical cores, parametrisations and Earth-System modules by using double precision bit-reproducible computations for scales near and below the truncation scale?

Are we making inefficient use of computing resources (i.e. energy) that could otherwise be used to increase resolution, better representation of physical processes, ensemble size?

Vana et al 2016

Single vs Double Precision in IFS

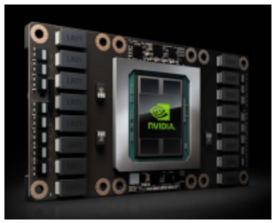




April 5, 2016

NVIDIA Unleashes Monster Pascal GPU Card at GTC16

Tiffany Trader



Earlier today (Tuesday) at the seventh-annual GPU Technology Conference (GTC) in San Jose, Calif., NVIDIA revealed its first Pascal-architecture based GPU card, the P100, calling it "the most advanced accelerator ever built." The P100 is based on the NVIDIA Pascal GP100 GPU — a successor to the Kepler GK110/210 — and is aimed squarely at HPC, technical computing and deep learning workloads.

Packing a whopping !
point performance, t
to date. And with 15
GPU that NVIDIA has
on TSMC's 16nm FinFET manufacturing process.

Floating point is cited as another critical resource. The three sizes – half-precision, single-precision and double-precision — all fit the IEEE standard. The peak speed of 5.2 teraflops double-precision performance doubles to 10.6 teraflops running in single-precision floating point mode. Double it again, and you get 21 peak teraflops of half-precision floating point performance — another first.

"GPUs have used half-precision for at least a dozen years as a storage mechanism to save space — for textures — but we've never built an arithmetic pipeline to implement the 16-bit floating point directly, we've always converted it," Nyland said. "What we've done is left it in its native size and then pair it

April 6, 2016

Europe's Fastest Supercomputer to Get Pascal GPU Upgrade

Tiffany Trader and John Russell

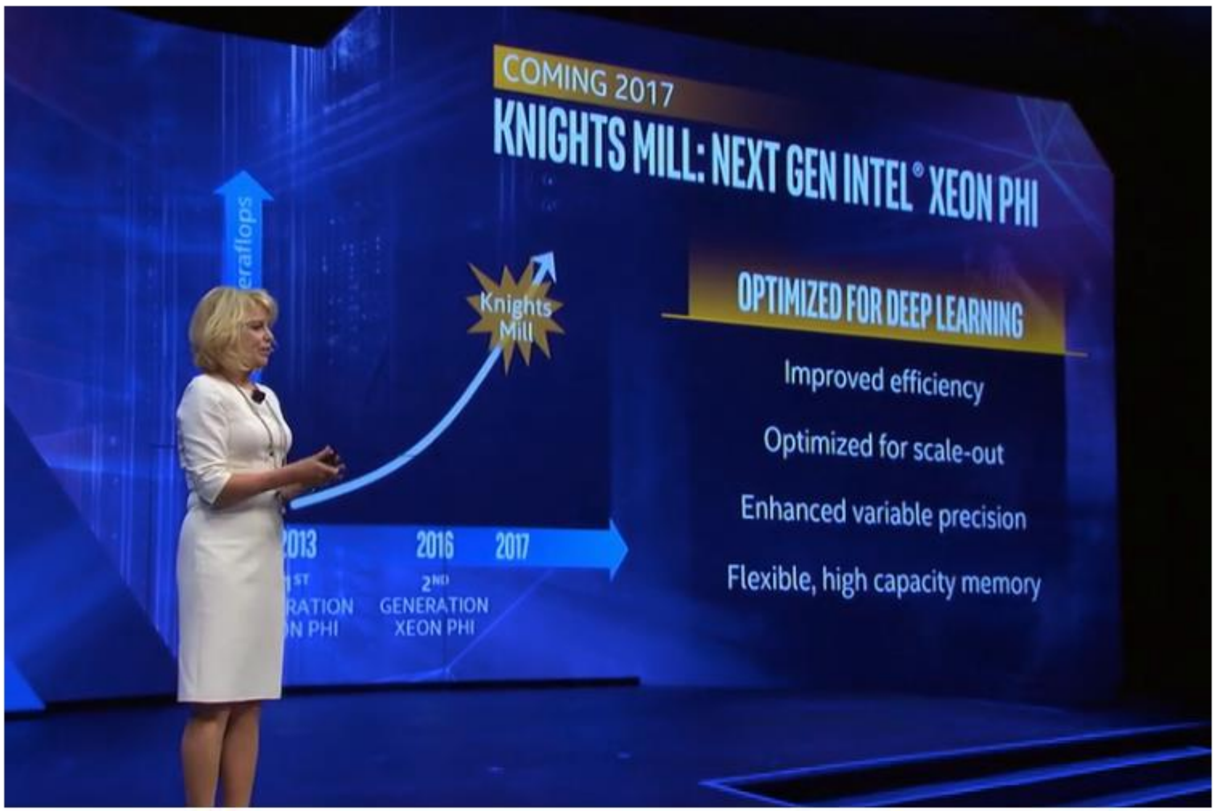
Already Europe's fastest supercomputer at 7.8 petaflops, the Piz Daint (hybrid CPU/GPU Cray XC30) at the Swiss National Computing Center (CSCS) will double its performance with a massive upgrade that involves switching to NVIDIA's newest Pascal GPU architecture and merging with Piz Dora (Cray XC40), a smaller CPU-based machine. The announcement was made at GTC16 yesterday. Last November Piz Daint placed seventh on the TOP500 list.

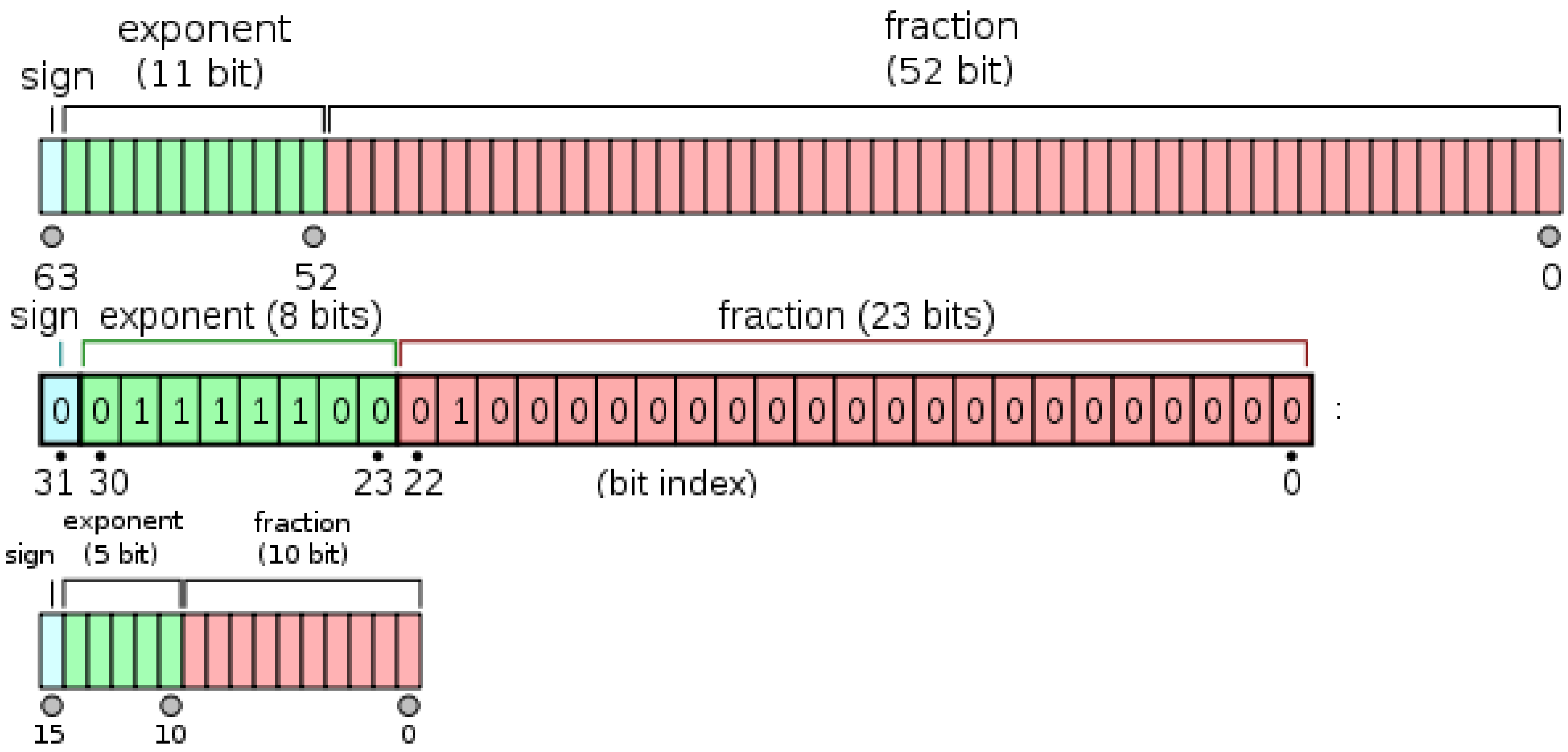
Plans call for 5,200 NVIDIA K20xs to be replaced by 4,500 Pascal GPUs – which version hasn't been decided. Also, the Intel processors will be upgraded from Sandy Bridge to Haswell architecture. When completed, the new combined system, all on a single fabric, will keep the Piz Daint name and provide

Intel Unveils Plans for Knights Mill, a Xeon Phi for Deep Learning

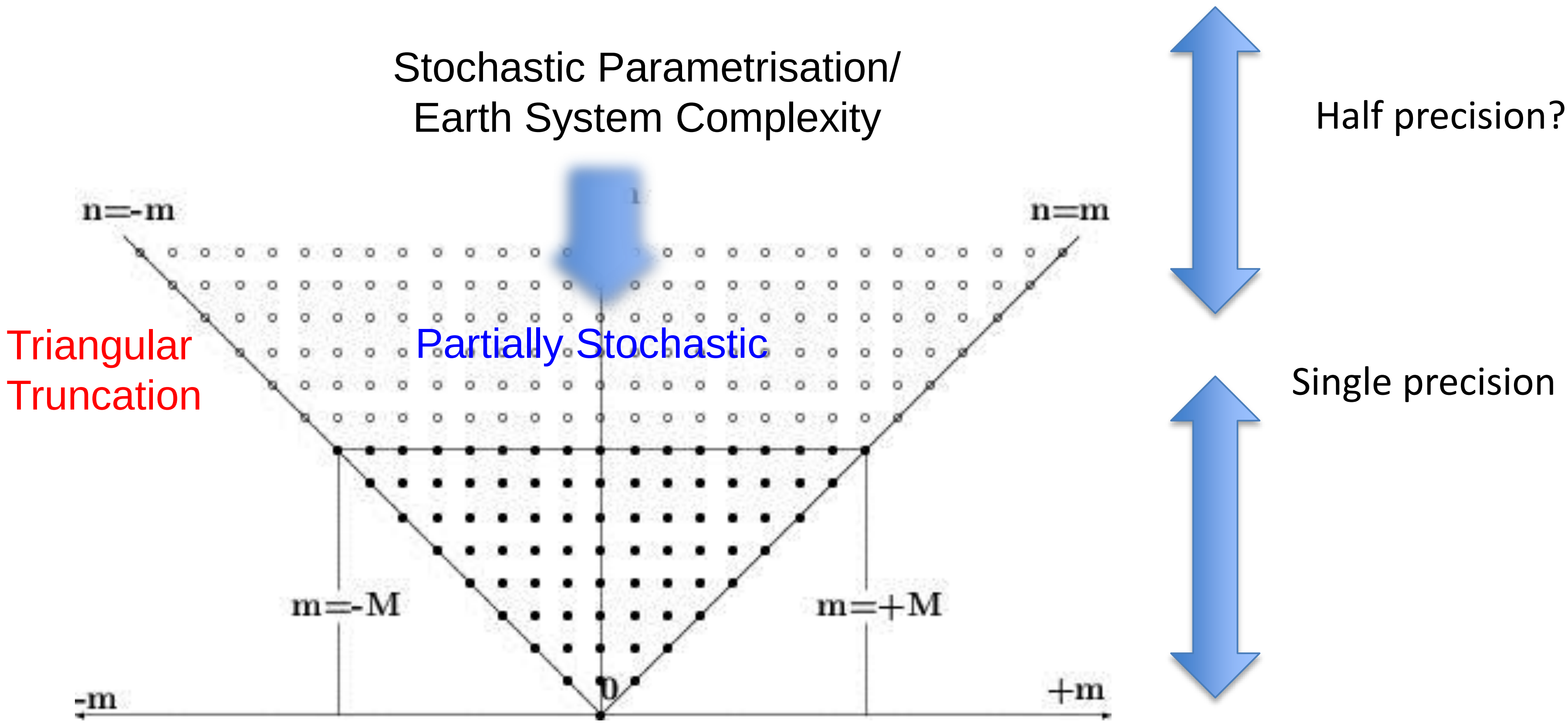
Michael Feldman, Aug. 18, 2016, 1:33 a.m.

At the Intel Developer Forum (IDF) this week in San Francisco, Intel revealed it is working on a new Xeon Phi processor aimed at deep learning applications. Diane Bryant, executive VP and GM of Intel's Data Center Group, unveiled the new chip, known as Knights Mill, during her IDF keynote address on Wednesday.



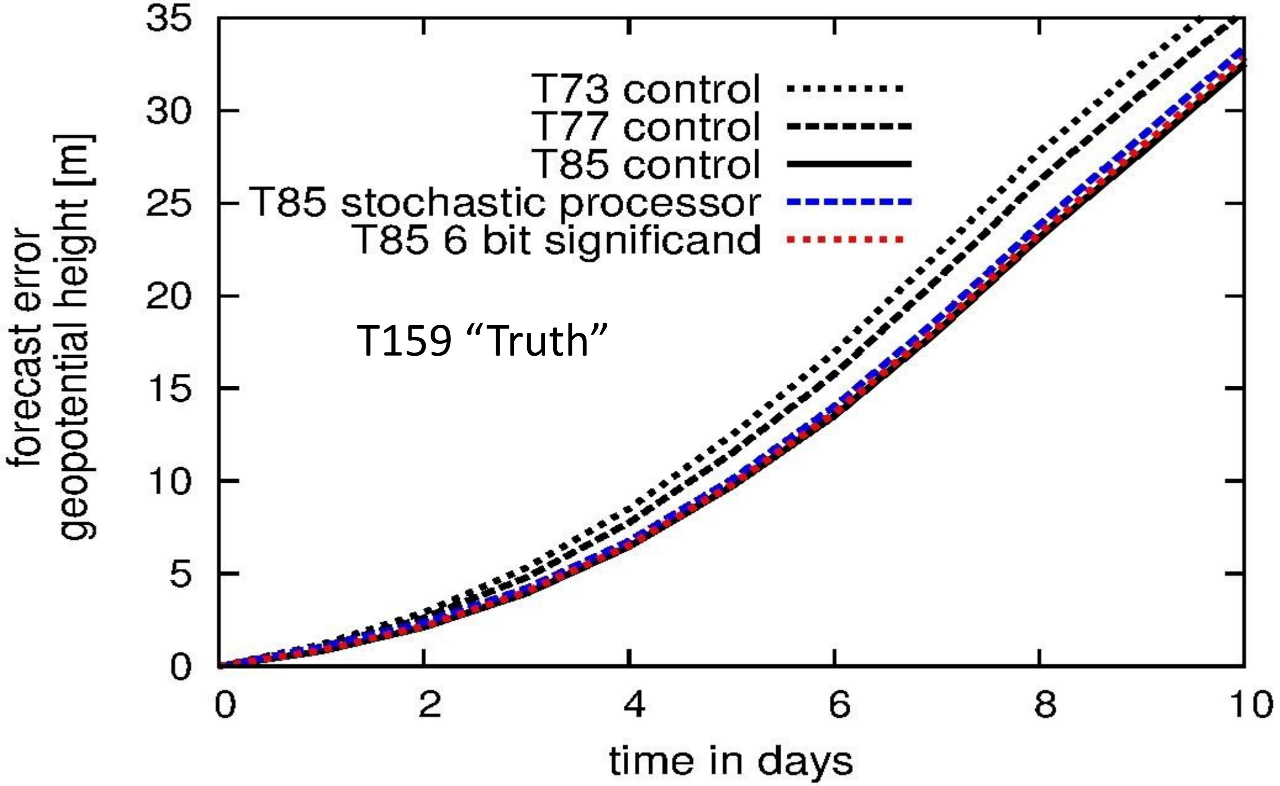


Reduced precision allows computations to proceed more quickly, and data to be moved with less energy overhead.


Stochastic Parametrisation/
Earth System Complexity
Half precision?
n=-m
n=m
Triangular
Truncation
Partially Stochastic
Single precision
m=-M
m=+M
-m
0
+m

More accurate "weather forecasts" with less precision Reading Spectral Model

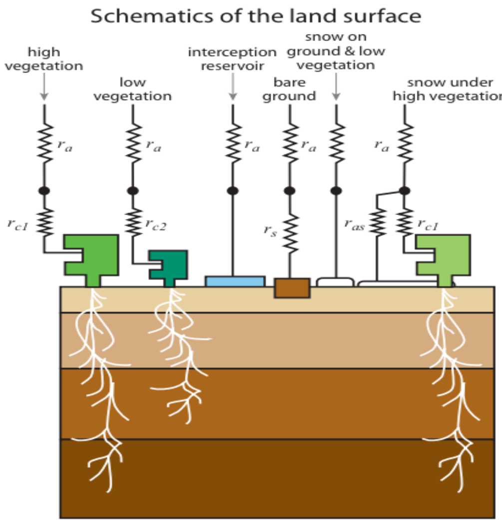
Düben and Palmer, 2014. Monthly Weather Review



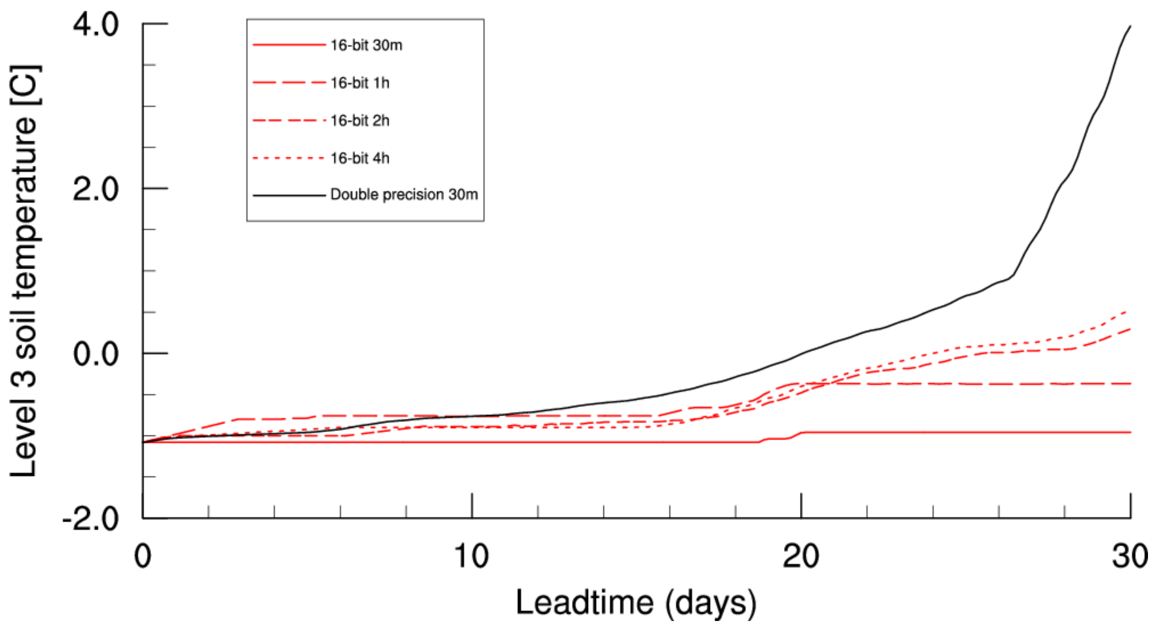
The stochastic chip / reduced precision emulator is used on 50% of numerical workload:
All floating point operations in grid point space
All floating point operations in the Legendre transforms between wavenumbers 31 and 85.
T85 cost approx that of T73

The representation of physical and biogeochemical processes in Earth-System Modules are highly parametrised and hence uncertain.

Could computations in Earth-System modules be performed at half precision?



Dave Macleod, Peter Düben, Andrew Dawson



Clim Dyn (2016) 46:3865–3882
DOI 10.1007/s00382-015-2809-5

The reliability of single precision computations in the simulation of deep soil heat diffusion in a land surface model

Richard Harvey[1,2] · Diana L. Verseghy[1]

Abstract Climate models need discretized numerical algorithms and finite precision arithmetic to solve their differential equations. Most efforts to date have focused on reducing truncation errors due to discretization effects, whereas rounding errors due to the use of floating-point arithmetic have received little attention. However, there are increasing concerns about more frequent occurrences of rounding errors in larger parallel computing platforms (due to the conflicting needs of stability and accuracy vs. performance), and while this has not been the norm in climate and forecast models using double precision, this could change with some models that are now compiled with single precision, which raises questions about the validity of using such low precision in climate applications. For example, processes occurring over large time scales such as permafrost thawing are potentially more vulnerable to this issue. In this study we analyze the theoretical and experimental effects of using single and double precision on simulated deep soil temperature from the Canadian LAnd Surface Scheme (CLASS), a state-of-the-art land surface model. We found that reliable single precision temperatures are limited to depths of less than about 20–25 m while double precision shows no loss of accuracy to depths of at least several hundred meters. We also found that, for a given precision level, model accuracy *deteriorates* when using *smaller* time steps, further reducing the usefulness of single precision. There is thus a clear danger of using single precision in some climate model applications, in particular any scientifically meaningful study of deep soil permafrost must at least use double precision. In addition, climate modelling teams might well benefit from paying more attention to numerical precision and roundoff issues to offset the potentially more frequent numerical anomalies in future large-scale parallel climate applications.

Keywords Floating-point arithmetic · Numerical precision · Single precision arithmetic · Double precision arithmetic · Climate models · Permafrost · Land surface models · Deep soil processes

1 Introduction

Climate models use sophisticated numerical algorithms to solve the complex primitive equations of atmospheric and oceanic motions. These algorithms contain two well-known and unavoidable sources of errors: *truncation errors* (because computations must be completed in a finite time), which are caused by replacing the continuous time and space differentials of the original field equations with finite increments, and *rounding errors* (because computer memory is not infinite), which are caused by replacing real numbers of infinite precision with finite-sized com-

$$\frac{\P y}{\P t} = S(y, \underset{\text{Highly uncertain}}{f}, t)$$

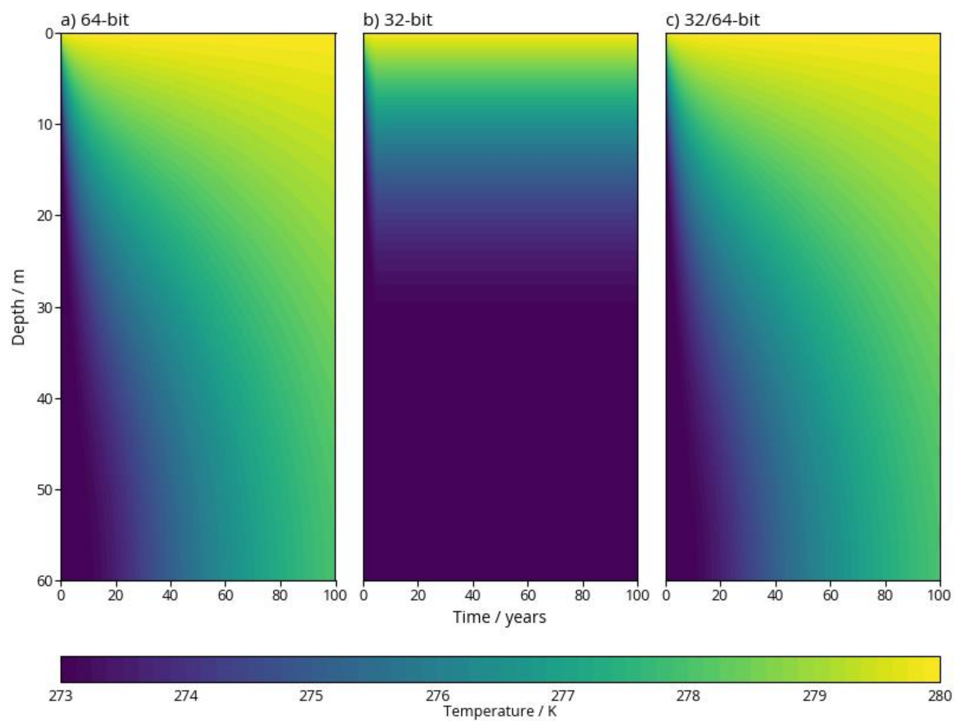
$$\underbrace{y(t^{n+1}) = y(t^n)}_{\text{Represent at high precision}} + \underbrace{\mathrm{D}t \cdot S(y(t^n), f(t^n), t^n)}_{\text{Compute (and retrieve fields from memory) at low precision}}$$

$$\frac{\partial T}{\partial t} = D \frac{\partial^2 T}{\partial z^2}$$

Andrew Dawson, Peter Düben

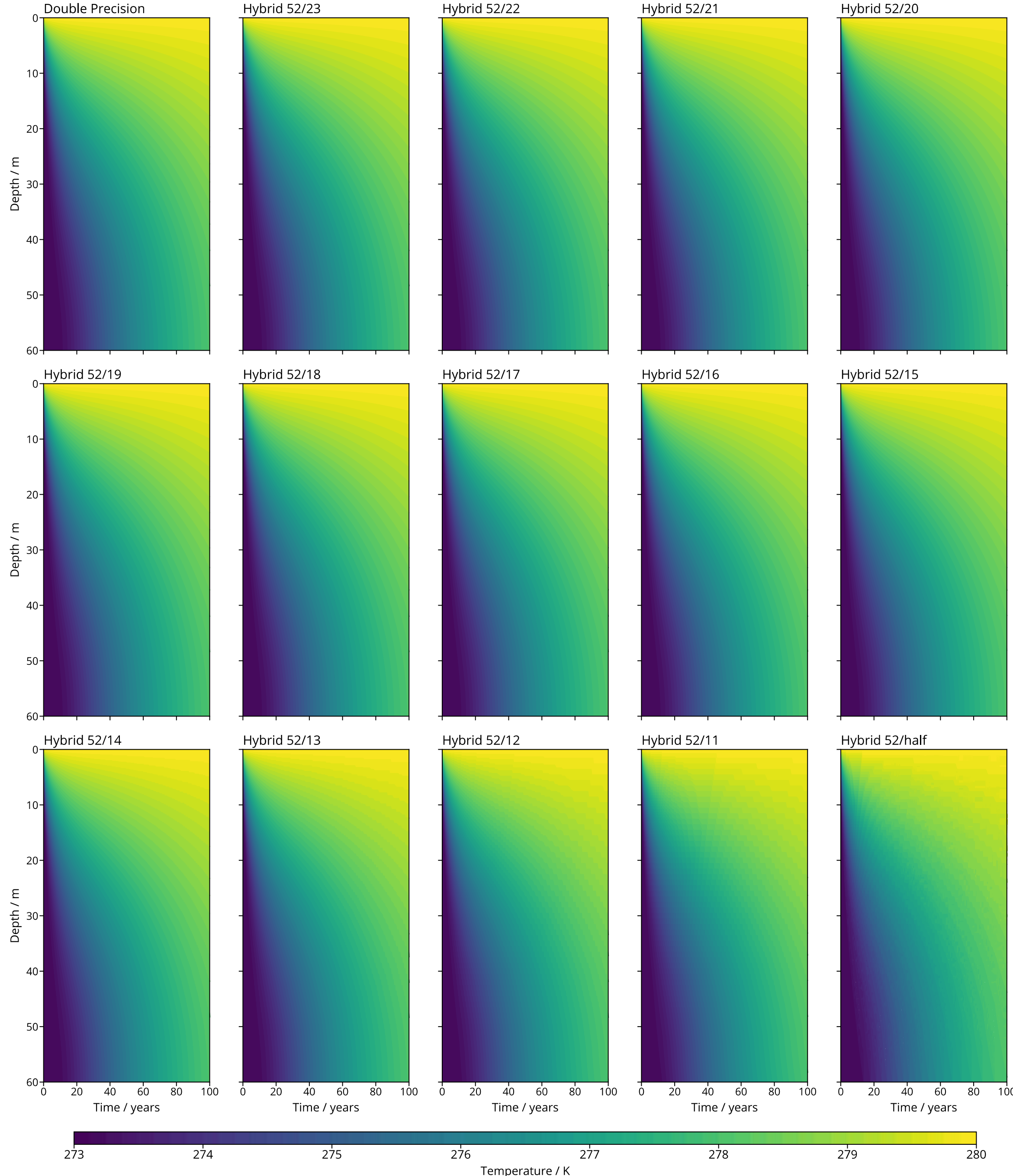
$$T_j^{n+1} = T_j^n + \underbrace{\Delta t \, D \, \frac{(T_{j+1}^n - 2T_j^n + T_{j-1}^n)}{(\Delta z)^2}}_{\text{32 bits}}$$

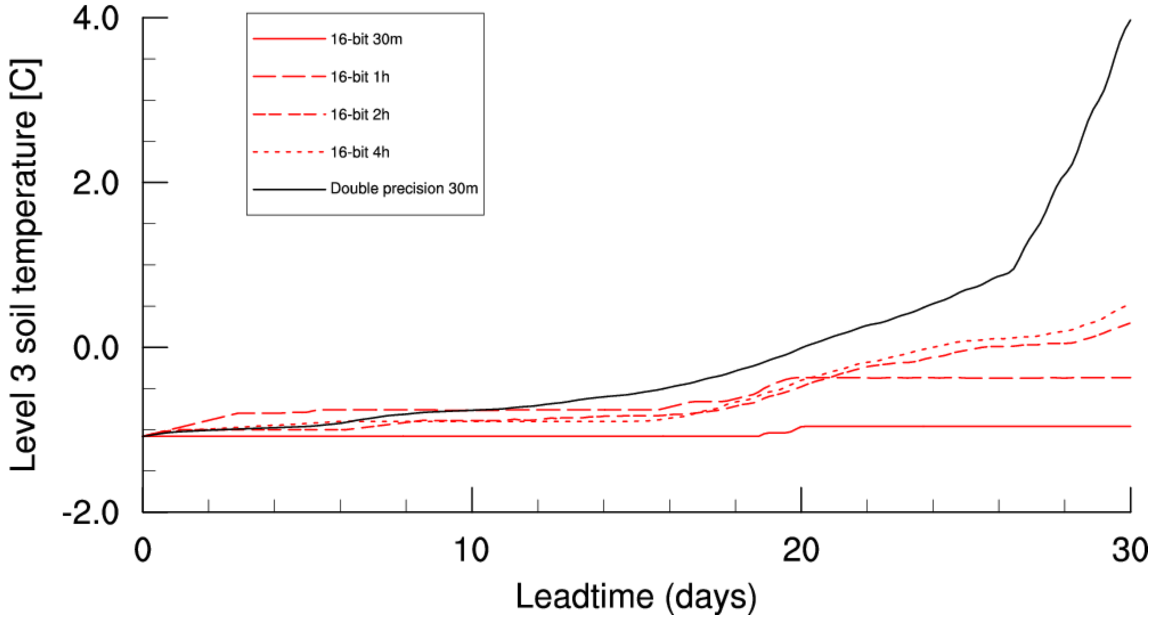
64 bits

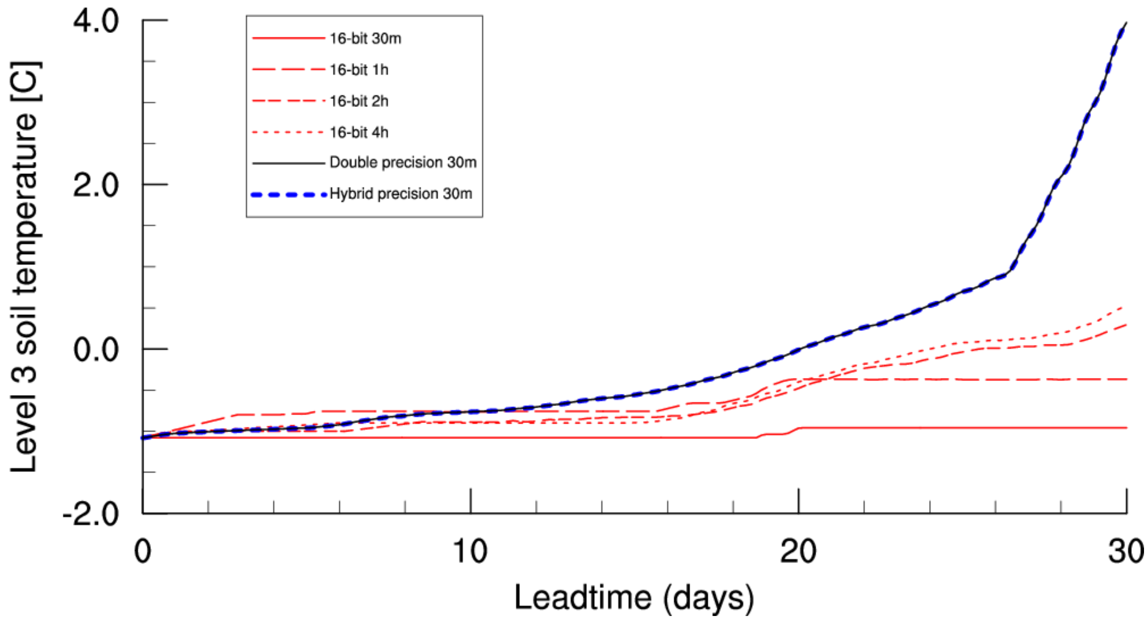


Andrew Dawson

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial z^2}$$

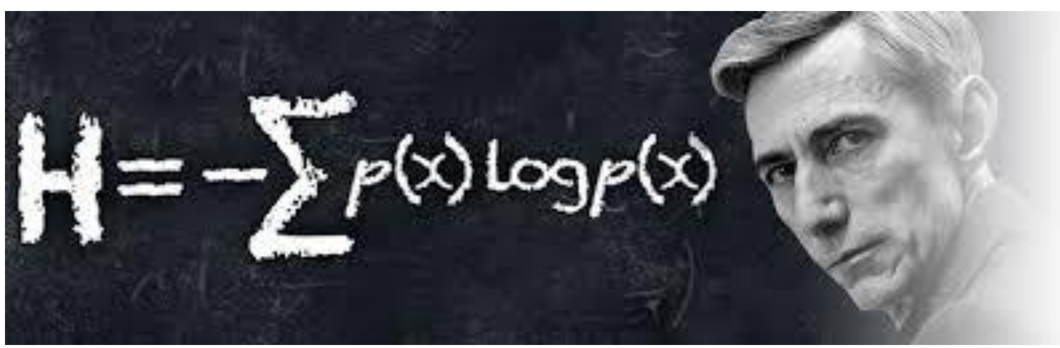






Dave Macleod,
Peter Düben,
Andrew Dawson

What is the real information content in each of the billions of bits that represent variables in a weather/climate model?



Stephen Jeffress

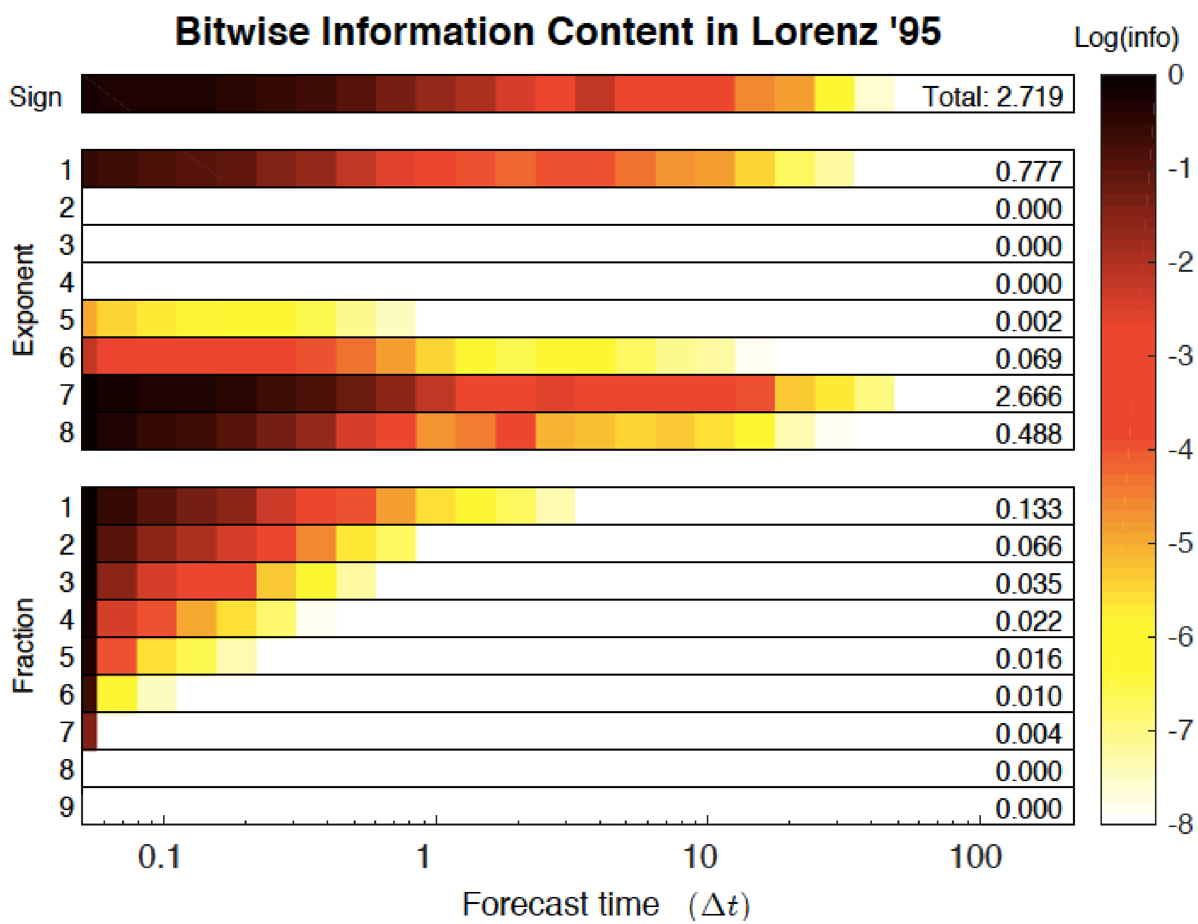
Entropy: $$H\left[X\left(t\right)\right] = -\sum_{i} P_i\left[\log\left(P_i\right)\right]$$

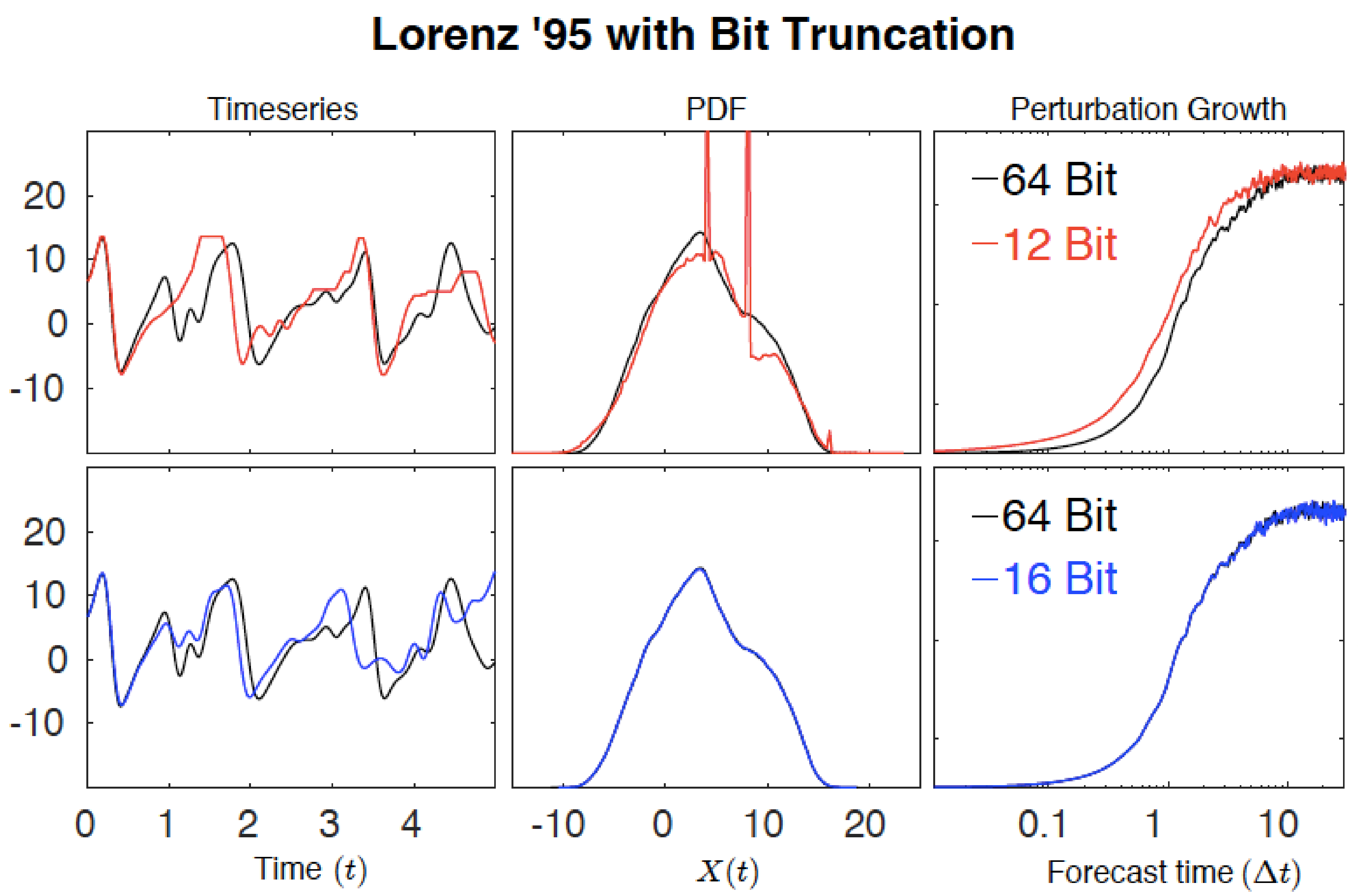
Bitwise information content at a single forecast time:

$$I_b\left(\Delta t\right) = H\left[X(t+\Delta t)\right]$$
$$-p_0 H\left[X(t+\Delta t) \mid b(t) = 0\right] - p_1 H\left[X(t+\Delta t) \mid b(t) = 1\right]$$

Total bitwise information content: $$J_b = \int_0^{\infty} I_b\left(\Delta t\right) d\Delta t$$

$$\dot{X}_k = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F$$




Lorenz '95 with Bit Truncation
Timeseries
PDF
Perturbation Growth
64 Bit
12 Bit
64 Bit
16 Bit
20
10
0
-10
5
3
1
0
1
2
3
4
-10
0
10
20
0.1
1
10
Time (t)
$X(t)$
Forecast time (Δt)

Conclusions

- Forecast reliability is poor beyond medium range.
- Improving forecast reliability is crucially important for both weather and climate communities
- Increased resolution likely to be top priority for ECMWF to reach strategic goals
- Given the inherent stochasticity of the sub-grid closure problem, we are very likely wasting computer resources integrating all model variables with 64 bits
- Earth-System modules are highly uncertain
- It is possible that small-scale variables in the dynamical core and many Earth-System modules could be computed at half-precision, making use of new mixed-precision machine-learning chips.
- Information theory could play an important role in determining the real information content in model variables.