Model uncertainty quantification using ultra-efficient inexact hardware

Peter Düben, Tim Palmer

Department of Physics, University of Oxford

Contact: peter.dueben@physics.ox.ac.uk

1. Inexact hardware can reduce computing costs and therefore improve weather forecasts

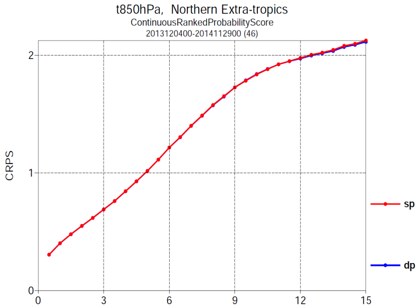
The use of inexact hardware

Inexact hardware trades an increase of performance and a reduction of power consumption against a reduced level of numerical precision. If savings due to reduced precision are re-invested into larger computing systems, resolution and ensemble sizes can be increased to improve forecasts. Inexact hardware is an emerging field of hardware development that promises significant savings.

Example 1: IFS in single precision

Ensemble forecasts and long-term simulations in double and single precision are almost identical for simulations at T399 resolution. We achieve ~40% speed-up on the Cray.

Düben and Palmer MWR 2015, Vana, Düben, Lang, Palmer, Leutbecher, Salmond, Carver submitted



Example 2: Reduced precision in a spectral dycore

We calculate weather forecasts with a spectral dynamical core (IGCM) in a "Held-Suarez world" and compare results against a high resolution truth. Real number precision is reduced by an emulator. Only 2% of the reduced precision simulation is calculated in double precision. We estimate savings for pruned hardware in cooperation with computer scientists (Krishna Palem Rice University, Christian Enz EPFL and John Augustine IITM).

Results show that it is much better to reduce numerical precision when compared with a reduction in resolution to save energy.

Resolution	Real number representation	Normalised energy demand	Mean error Z500 at day 2
235 km	64 bits	1.0	2.3
315 km	64 bits	0.47	4.5
235 km	**20 bits**	0.29	2.5

Düben et al. MWR 2015, Düben et al. DATE 2015

2. A study of inexact hardware can help to understand model error and model uncertainty

2.1 Inexact hardware and parameter uncertainty

If the optimal level of precision to represent physical parameters has been identified, this provides a minimal value of uncertainty for these parameters.

Table: Precision levels that can be used for parameters in a cloud resolving model in a superparametrised setup and the error in their representation in reduced precision.

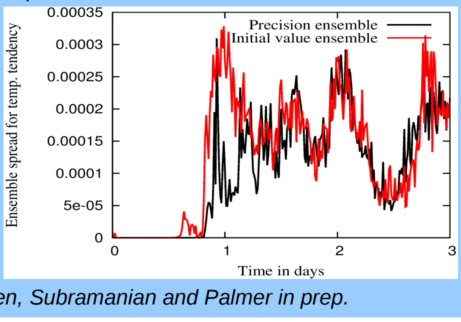
Parameter	Precision	Error
specific heat of air	16 bits	0%
gravitational acceleration	14 bits	0.6%
latent heat of condensation	13 bits	0.8%
latent heat of fusion	9 bits	21%
latent heat of sublimation	12 bits	1.4%
gas constant water vapour	9 bits	11%
diffusivity water vapour	10 bits	3.6%
thermal conductivity of air	10 bits	2.3%
dynamic viscosity of air	9 bits	11%

Düben, Subramanian and Palmer in prep.

2.2 Inexact hardware and ensembles

Rounding errors can also be used to generate ensemble simulations and it can be argued that rounding errors can be used to represent sub-grid-scale variability and model error. *Düben and Dolaptchiev TCFD 2015*

Figure: Tendency of temperature that feeds back from the cloud resolving model to the IFS single column model in a superparametrised setup. The simulations in the precision ensemble do only differ in the level of precision that was been used. The ensemble spread is comparable to the spread of an initial value ensemble.



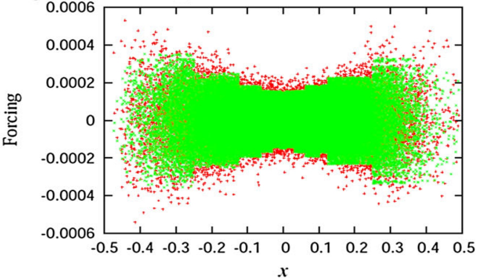
Düben, Subramanian and Palmer in prep.

2.3 Inexact hardware and stochastic parametrisation

The impact of rounding errors can be interpreted as a forcing term which is uncorrelated in space and time that is added to the differential equations. The impact of this forcing on model dynamics can be assumed to be small if the forcing is much smaller than the random forcing of a stochastic parametrisation scheme.

On the other hand, the magnitude of rounding errors can provide useful information for the development of stochastic parametrisation schemes as it provides a range of uncertainty.

Figure: We showed that rounding errors (green) can replace the random forcing of a stochastic parametrisation scheme (red) in a randomly forced Burgers equation model.



Düben and Dolaptchiev TCFD 2015

2.4 Inexact hardware and model development (co-designed)

A detailed analysis of the optimal level of precision will reveal (1.) parts of the model that can work with an extremely low level of precision and (2.) parts of the model in which precision is lost due to bad coding.

1. can be used to identify parts of the model that can be removed.

2. can be used to make the model more resilient against hardware errors.

Figure: We changed the cloud resolving model in a superparametrised setup following the results of a precision analysis. For the double precision setup redundant model parts could be removed such that performance is increased by ~10%. For the reduced precision setup we changed the 2D advection scheme to allow half precision. This was not possible in the original model setup.

Düben, Subramanian and Palmer in prep.