High resolution global climate modelling; the UPSCALE project, a large simulation campaign

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.5194/gmdd-7-563-2014
To link to this article DOI: http://dx.doi.org/10.5194/gmd-7-1629-2014

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Geoscientific
Model Development

Open Access

High-resolution global climate modelling: the UPSCALE project, a large-simulation campaign

M. S. Mizielinski[1], M. J. Roberts[1], P. L. Vidale[2], R. Schiemann[2], M.-E. Demory[2], J. Strachan[3,*], T. Edwards[4], A. Stephens[5], B. N. Lawrence[2,5], M. Pritchard[5], P. Chiu[5], A. Iwi[5], J. Churchill[6], C. del Cano Novales[6], J. Kettleborough[1], W. Roseblade[1], P. Selwood[1], M. Foster[1], M. Glover[1], and A. Malcolm[1]

[1]Met Office Hadley Centre, FitzRoy Road, Exeter, UK
[2]National Centre for Atmospheric Science (NCAS), Dept. of Meteorology, University of Reading, Reading, UK
[3]Willis Research Fellow, Knowledge Transfer Partnership, University of Reading, Reading, UK
[4]CRAY Inc., Centre of Excellence for HECToR, University of Edinburgh, Edinburgh, UK
[5]STFC Centre for Environmental Data Archival, Rutherford Appleton Laboratory, Didcot, UK
[6]STFC Scientific Computing Department, Rutherford Appleton Laboratory, Didcot, UK
[*]currently at: Met Office Hadley Centre, FitzRoy Road, Exeter, UK

Correspondence to: M. S. Mizielinski (matthew.mizielinski@metoffice.gov.uk)

Abstract. The UPSCALE (UK on PRACE: weather-resolving Simulations of Climate for globAL Environmental risk) project constructed and ran an ensemble of HadGEM3 (Hadley Centre Global Environment Model 3) atmosphere-only global climate simulations over the period 1985–2011, at resolutions of N512 (25 km), N216 (60 km) and N96 (130 km) as used in current global weather forecasting, seasonal prediction and climate modelling respectively. Alongside these present climate simulations a parallel ensemble looking at extremes of future climate was run, using a time-slice methodology to consider conditions at the end of this century.

These simulations were primarily performed using a 144 million core hour, single year grant of computing time from PRACE (the Partnership for Advanced Computing in Europe) in 2012, with additional resources supplied by the Natural Environment Research Council (NERC) and the Met Office. Almost 400 terabytes of simulation data were generated on the HERMIT supercomputer at the High Performance Computing Center Stuttgart (HLRS), and transferred to the JASMIN super-data cluster provided by the Science and Technology Facilities Council Centre for Data Archival (STFC CEDA) for analysis and storage.

In this paper we describe the implementation of the project, present the technical challenges in terms of optimisation, data output, transfer and storage that such a project involves and include details of the model configuration and the composition of the UPSCALE data set. This data set is available for scientific analysis to allow assessment of the value of model resolution in both present and potential future climate conditions.

1 Introduction

The development of the Met Office Unified Model™ (MetUM) in recent years has yielded a traceable hierarchy of model resolutions from the N96L85 grid[1], with 130 km horizontal resolution at 50° N and 85 vertical levels spanning the lower 85 km of the atmosphere, used in standard climate simulations to N512L70, 25 km at 50° N with 70 levels again spanning 0–85 km, used in global weather forecasting. This hierarchy, in which all but a few parameters are identical across configurations, allows the impact of resolution to be studied and understood.

---

[1]"Nx" denotes a global latitude–longitude grid of 1.5x by 2x points.

Table 1. Facilities used in UPSCALE.

Facility	Location	Manufacturer and model	Specification
HERMIT	HLRS[a], Germany	Cray XE6	113 664 cores (AMD Interlagos 2.3 GHz)
JASMIN	RAL[b], UK	Panasus ActiveStor 11, Dell R610 cluster	4.6 PB storage, 96 cores (Intel Xeon 3.5 GHz)
HECToR	EPCC[c], UK	Cray XE6	90 122 cores (AMD Interlagos 2.3 GHz)
MONSooN	Met Office, UK	IBM Power 775	5120 cores (POWER7, 3.8 GHz)

[a] High Performance Computing Center Stuttgart (HLRS) at the University of Stuttgart. [b] Rutherford Appleton Laboratory.[c] Edinburgh Parallel Computing Centre.

The role of resolution in different physical processes in the climate system is not necessarily the same (Roberts et al., 2009; Demory et al., 2014; Schiemann et al., 2014). For example, due to the local Rossby radius of deformation a 1/3° resolution ocean model cannot resolve the most important processes, eddies, while at 60 km the atmosphere can (Roberts et al., 2009; Shaffrey et al., 2009; Demory et al., 2014). Coarse-resolution simulations can produce representative data for global mean properties, but their limitations for studying regional effects and temporal variability are becoming more obvious (Roberts et al., 2009; Shaffrey et al., 2009; Scaife et al., 2011; Delworth et al., 2012; Kinter et al., 2013). Recent work by Demory et al. (2014) demonstrates that the energy budgets in an ensemble of different resolution versions of the HadGEM3 (Hadley Centre Global Environment Model 3) and HadGEM1 are very consistent, but moisture transport and the balance of evaporation and precipitation over land, critically important for climate impacts, only converges at resolutions finer than 60 km (N216 and above). Strachan et al. (2013) have shown that average tropical cyclone numbers can be well represented at resolutions of around 130 km, but grids finer than 60 km are needed to represent the inter-annual variability of cyclone counts properly, while accurate intensity simulation requires much higher resolution. An understanding of the dependence of different processes on resolution is vitally important both for determining critical resolution thresholds for model configuration, and for producing credible and useful information on future weather and climate. The construction of a traceable hierarchy of model resolutions is a necessary precondition for gaining this understanding.

Following the development of the first high-resolution global climate models at the Japanese Earth Simulator (Ohfuchi et al., 2004; Mizuta et al., 2006), investigations into the value of resolution have continued rapidly. High-resolution climate models require significant amounts of computer time and data storage, leading to episodic simulation campaigns, or "numerical missions" (Shaffrey et al., 2009; Navarra et al., 2010; Kinter et al., 2013), when resources can be obtained. These campaigns are characterised by short development and operational phases, followed by several years of work to extract scientific results from the data. Recent work on the MetUM (Malcolm et al., 2010; Selwood, 2012) has significantly improved its computational performance and scalability to the point where it is possible to conceive of running ensembles of multi-decadal climate simulations at weather forecast resolution. With this capability we successfully applied for a large amount of computing time from PRACE (the Partnership for Advanced Computing in Europe) to generate ensembles of atmosphere-only simulations for present and future climate conditions, at global weather forecast resolution to study extreme weather events and risks; the UPSCALE (UK on PRACE: weather-resolving Simulations of Climate for globAL Environmental risk) project.

The success of UPSCALE was made possible by two significant computing facilities: HERMIT and JASMIN. HERMIT is the Cray XE6 supercomputer at the High Performance Computing Center Stuttgart (HLRS), on which we were granted 144 million core hours during a single year by PRACE, and JASMIN is the super-data cluster (Lawrence et al., 2012) managed by the Science and Technology Facilities Council (STFC) Scientific Computing Department (SCD) on behalf of the Centre for Data Archival (CEDA), which hosts the 400 TiB[2] of data generated over the lifetime of UPSCALE along with analysis facilities. In addition support was provided by the UK supercomputers HECToR and MONSooN (Met Office NERC Supercomputing Node) along with the underlying network infrastructure provided by SuperJANET and GÉANT. Brief details of each facility are given in Table 1.

This paper has two main aims; to describe the important scientific and technical aspects of the execution of this project, and to provide a reference for users wanting to exploit the UPSCALE data set. Details of the model configuration are described in Sect. 2, while the ensemble of simulations performed and their output data are described in Sect. 3, with conclusions in Sect. 4. A significant supporting cast of people and organisations is noted in the acknowledgements.

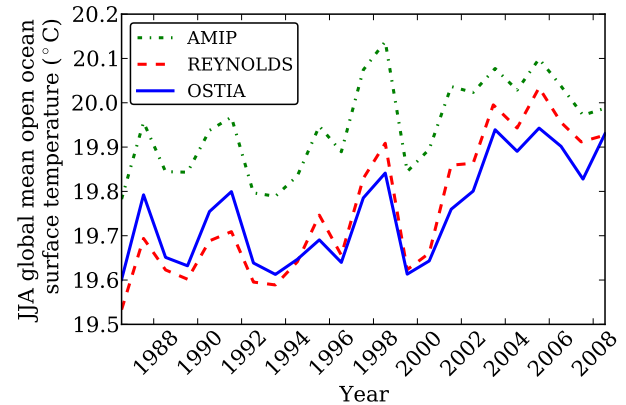[2] 1 tebibyte (TiB) = 2^{40} bytes = 1.1 terabytes (TB).
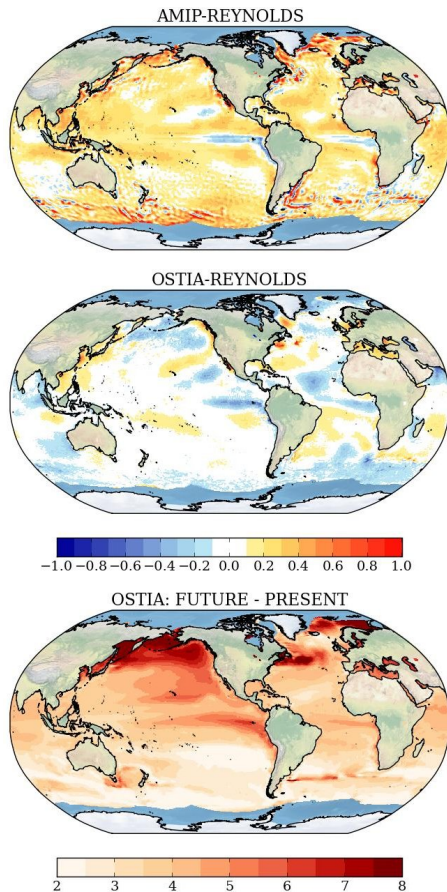
Figure 1. Spatial difference between 1986–2008 JJA mean SST in AMIP-II and Reynolds data sets (top) and OSTIA and Reynolds (middle). The bottom panel shows the future climate SST change applied, averaged over JJA. The colour bars are annotated in Kelvin.

Figure 2. Time series of global ocean mean JJA surface temperatures for the AMIP-II (dot-dashed green line), Reynolds (dashed red line) and OSTIA (solid blue line) data sets.

2 Model configuration

2.1 Science configuration

The UPSCALE ensemble of climate simulations are based upon the HadGEM3 Global Atmosphere 3 (GA3) and Global Land 3 (GL3) configurations of the MetUM and the Joint UK Land Environment Simulator (JULES) respectively, as documented in Walters et al. (2011). A core principle of development of the MetUM is the construction of a traceable hierarchy of model resolutions running from the coarse grids used in International Panel on Climate Change (IPCC) class climate models, typically around 130 km (at 50° N), to the finer grids used in global weather forecasting, around 25 km. The UPSCALE simulations use the same 25 km N512 grid used in the Met Office operational global weather forecasts, but with 85 vertical levels rather than 70, with the uppermost at 85 km.

There are very few differences in physical and dynamical settings in this model compared to lower-resolution counterparts, mostly related to numerical stability, which are noted in Table 2. We also apply diffusion to the vertical wind velocities in the upper five levels of the atmosphere to dissipate grid-scale artefacts in the stratosphere.

While the configuration of the UPSCALE ensemble broadly follows the Atmospheric Model Intercomparison Project II (AMIP-II) standard there are a few deviations made for scientific reasons. One such deviation is the use of daily sea surface temperature (SST) and sea ice forcings, derived from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) product (Donlon et al., 2012), which has a native resolution of 1/20° and is a synthesis of satellite and in situ observations covering 1985 to the present day (where 1985–2008 is a reanalysis, see Roberts-Jones et al., 2012). OSTIA was chosen because of its finer resolution than other data sets, allowing a richer and more realistic representation of the ocean surface on the model grid. Figure 1 shows a comparison between OSTIA, Reynolds (Reynolds et al., 2002) and AMIP-II (Taylor et al., 2000) data sets, indicating that the latter is up to 0.4 K warmer than both Reynolds and OSTIA over large areas, including those important for tropical cyclone genesis. The global average AMIP-II SST is approximately 0.2 K warmer, see Fig. 2, with Reynolds and OSTIA agreeing well. The aerosol, ozone, solar variability, volcanic and other time-varying forcings are as defined by the AMIP-II protocols.

The design of the UPSCALE programme included two ensembles, each of five members, one simulating the present climate from 1985 to 2012 and the other looking at future climate change at the end of the 21st century using a timeslice methodology. The future climate simulations were configured with SST from the present climate runs plus the SST change between the 1990–2010 and 2090–2110 in the HadGEM2 Earth System runs under the IPCC Representative

Concentration Pathway (RCP) 8.5 climate change scenario (Collins et al., 2011; Jones et al., 2011). These SST changes were calculated for each month, interpolated in both space and time, and added to the daily varying OSTIA forcing data on the model grid. The increase in JJA SST forcing for the future climate ensemble is shown in Fig. 1, with a mean difference of just under 4 K.

Sea ice fractions for the future climate ensemble were regridded from the same HadGEM2 Earth System runs, but were interpolated from monthly to daily frequency. For regions of sea surface that lose sea ice coverage between the present and future climate scenarios, SST values were interpolated from the HadGEM2 results.

Other settings including CO_2, methane, nitrous oxide, CFC and HFC concentrations were adjusted accordingly, but do not vary with time in the future climate simulations. While the present-climate ensemble was completed in full, the climate change runs experienced significantly higher levels of numerical instability, making progress with these runs more problematic. As a result only three out of five runs were performed, owing to the excessive amount of user intervention required to deal with repeated grid point storms (see Sect. 2.3).

Additional suites of valuable scientific simulations were performed to further our understanding of the role of resolution vs. the role of other aspects of numerical simulation. This exercise included ensembles of present and future climate simulations at N216 (60 km) resolution on HERMIT and MONSooN and N96 (130 km) resolution on HECToR with parallel settings for both climate conditions. A set of N512 runs with an updated scientific configuration, Global Atmosphere 4 (GA4) (Walters et al., 2013), were performed for present climate conditions to explore a number of sensitivities. These sensitivities included entrainment rates and the dynamics and radiation time steps. The final set of runs performed, again using the GA4 configuration as a basis, was a perturbed initial conditions ensemble consisting of six simulations, each a year long, to expand the sample size in 1 year (2003) that produced particularly intense weather and climate events.

The settings of the GA4 configuration are described in Walters et al. (2013), but the major difference to our runs, based on GA3, is the use of the Reynolds SST climatology rather than OSTIA.

2.2 Technical configuration: optimisation and tuning

While the MetUM is designed to be portable to any computing platform, it is always necessary to test and optimise performance (Malcolm et al., 2010) when porting to new systems, and HERMIT was no exception. Alongside preliminary testing of the scientific behaviour of the MetUM, considerable effort was put into technical aspects of its configuration and the optimisation of its source code by T. Edwards (Cray Inc.), yielding significant performance benefits in our

production configurations. These optimisations were developed against the N512 GA3 present climate settings, but were applied to all runs on HERMIT, where possible.

2.2.1 Processor decomposition

Parallelisation within the MetUM has traditionally been achieved through the decomposition of the globe into rectangular latitude–longitude domains, each assigned to one Message Passing Interface (MPI) process. The haloes, used to supply the semi-Lagrangian advection scheme with departure point information, impose a minimum size on these domains and a maximum number of MPI processes, as haloes are not permitted to extend across multiple MPI tasks. Additional communication, required close to the poles where the longitudinal grid spacing falls below 100 m, is performed on demand by the advection routines. OpenMP threading directives have been introduced in recent versions of the MetUM, extending the ability to scale to larger processor counts. This hybrid parallelisation approach allows better performance, efficiency and greater scaling than either technique on its own could.

For this project a scan of around a hundred different decompositions of the latitude–longitude grid and threading combinations was performed, each test consisting of 2-day simulation with minimal data output and the same initial conditions. The decomposition of the latitude–longitude grid onto MPI processes was found to be important. Where different decompositions of a particular number of processors were tested, the best configuration could be up to 25 % more efficient than the worst. Decompositions where the longitude range is divided precisely onto an integer number of computing nodes yield the best performance, as the semi-Lagrangian advection scheme and numerical solver generate MPI traffic following the predominantly west to east atmospheric flow.

When using two OpenMP threads a sweet spot at $32 \times 72 = 2304$ MPI processes was found, yielding performance almost 20 % better than any similar configuration. This MPI decomposition is also the optimal configuration for four threads.

2.2.2 IO

For an IPCC-class resolution simulation the volume of data generated from an AMIP-II run is around 1 TiB, while at the N512 resolution the equivalent data set is 30 times larger. This data burden needs to be carefully considered and managed, and is the principal management issue for a climate project of this scale. These issues are not new – weather and climate simulations have been challenging the boundaries of IO speeds (Desgagné et al., 2006) and data storage (Ohfuchi et al., 2002; Hoffman, 2002; Sell, 2004) for well over a decade.

The computational speed of the MetUM on HERMIT makes IO a challenge; individual fields are output at frequencies from 3 h to 1 month, requiring data to be written

Table 2. Parameter differences between the GA3 standard and the configurations used here.

Parameter	Standard value	N512	N216	N96
Time step (s)	1200	600 or 450*	900	1200
CAPE threshold vertical velocity (m s^{-1})	0.3	0.4	0.4	0.4
CAPE closure timescale (s)	5400	3600	3600	3600
Targeted diffusion w threshold (m s^{-1})	0.3	1.0	1.0	0.5
ADI Pseudo Time step	8×10^{-4}	10^{-4}	3×10^{-4}	8×10^{-4}

* A shorter time step was used in some simulations to improve numerical stability, see Sect. 2.3 for details.
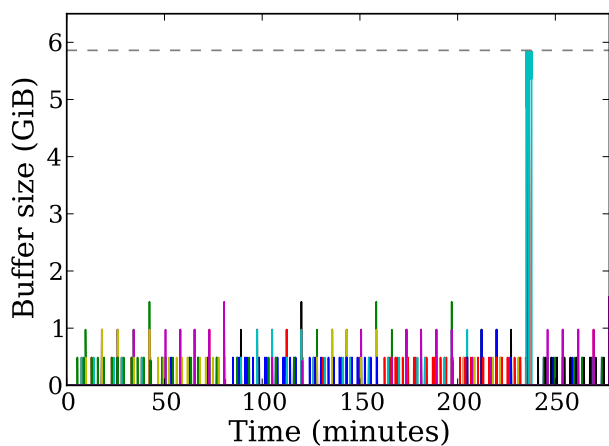
Figure 3. Buffer loading in a testing configuration of the MetUM. The buffer limit of 5.8 GiB (6000 MiB) is denoted by a dashed grey line.

to disk every real minute, with higher loads at the end of each simulated day and significantly more at the end of each simulated month. The MetUM can designate a subset of processes as IO servers to manage the writing of large volumes of data to disk, a common feature of modern high-resolution climate models (Madec, 2008; Dennis et al., 2012). These servers buffer and process the raw field data that are collected from the compute tasks, which allows a near-complete overlap between computation and disk IO, greatly improving the efficiency of the application. Our configuration uses 12 IO servers, one for each output stream plus one for the restart file, each with 6000 MiB of buffer space to maximise performance without triggering out-of-memory errors. The IO servers were located on separate nodes to the compute tasks to improve the coordination between the model grid and the decomposition of compute processes on individual nodes.

A time series of the volume of buffered data on each IO server is shown in Fig. 3, from which the regular peaks can be seen at the end of each model day, every five model days when output files are reinitialised, and at the end of each model month, when the combined volume of data exceeds the available buffer capacity and causes the simulation to briefly pause while IO tasks complete.

Given the IO loading described here it is important to tune the parameters of the underlying Lustre storage system on HERMIT – experimental tuning yielded optimal performance when the STRIPE_COUNT and STRIPE_SIZE attributes of the system were set to 8 and 16 MiB, respectively.

2.2.3 Segment sizes

Individual MPI processes decompose some of the larger computational tasks into smaller units, or "segments" of work, that can be processed independently by one of the OpenMP threads within each MPI task. The segment size denotes the number of grid points passed in each batch to the routine in question, with the results from each segment combined before proceeding to the next model physics component. Dividing the computational work into predefined segments allows the processor to make more efficient use of its memory cache and improve the overall run-time performance, with individual segments processed in parallel by OpenMP threads. The choice of segment size is fundamental to performance. Small segment sizes can incur unnecessary memory management overhead, time taken for data transfer between main memory and the CPU cache, while large segment sizes limit the benefit which can be obtained from parallel methods.

A profiling technique to find the optimal segment sizes was used, recording and playing back MPI communications, to analyse a small number of representative processes out of the thousands in the full simulation. This technique allowed all feasible segment size and OpenMP thread number combinations to be scanned in an efficient manner, and exposed an unexpected coupling between the segment size, number of OpenMP threads and run-time performance of these code kernels.

The results for the long-wave radiation routines are plotted in Fig. 4 along with the optimal segment sizes in Table 3. The dependence on segment size of the long-wave radiation routines using a single OpenMP thread is smooth, neglecting noise. However, when multiple threads are used a saw tooth pattern emerges in the dependency of performance on segment size, yielding significant performance differences for small changes in segment size. This saw tooth pattern arises from load balancing the segments of processing work within
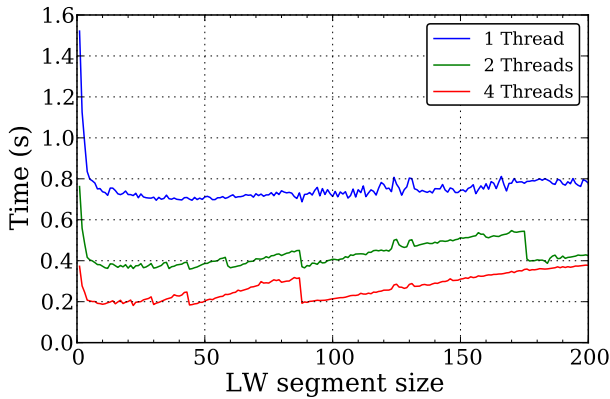
Figure 4. Variation in the time taken to complete the long-wave radiation calculation as a function of segment size for, from top to bottom, 1 (blue line), 2 (green line) and 4 (red line) threads.

the convection routines. As the segment size is increased the time taken for the routine to complete increases, as each segment occupies an OpenMP thread for a longer period of time, and if the number of segments does not divide equally into the number of threads some threads are under-worked. A sudden fall in the time taken for a routine to complete occurs when the number of segments divides equally into the number of available OpenMP threads.

Similar dependencies on segment size are seen in the short-wave radiation and convection routines, but as the volume of data processed is different in each case the optimal sizes are different.

2.2.4 Scaling

The scalability of the N512 configuration to higher core counts was investigated after the scientific configuration of the N512 resolution simulations was finalised. Short simulations of 2 model days with minimal IO were run for a range of MPI process and OpenMP thread combinations using up to 25 000 cores. The time taken per model time step was used to estimate simulation throughput, by accounting for initialisation times and IO costs, yielding the results shown in Fig. 5. The performance shows a good fit to Amdahl's law, despite the mixes of OpenMP and MPI, from which the fraction of the model code that is unparallelised is found to be in the range 3×10^{-4} to 5×10^{-4}.

The performance shown in Fig. 5 should be treated as the best possible level of performance for the MetUM on HERMIT. Analysis of the performance of successful job steps from production runs shows that the average model throughput was 5.0 months a day on 4600 cores, 10 % lower than shown, falling below 4.5 months a day at worst. Poor model throughput was particularly notable when the utilisation of HERMIT rose above 90 %. One possible explanation is connected with the distribution of allocated computing nodes on a busy system – the scheduler may allocate well separated

Table 3. Optimal segments sizes for different routines with different numbers of OpenMP threads.

Code	1 thread	2 threads	4 threads
LW radiation	53	30	22
SW radiation	30	44	22
Convection	137	60	12

nodes to a given job, impacting on MPI communication latency and reducing simulation throughput.

The frequency of IO within the MetUM can also lead to degraded performance under high system utilisation as competition for IO bandwidth during the writing of the end of month restart and output files slows progress. We are unable to determine which of these possible causes is contributing to the observed drop in model throughput.

2.3 Numerical stability issues

At resolutions above those used in IPCC-class climate runs, simulations such as the MetUM, see also Williamson (2013), are known to develop Grid Point Storms (GPS) where a grid cell size convective cell grows, typically over sharp orography, to the point where the numerical schemes in the dynamics routines break down. A GPS is characterised by a sudden growth in vertical wind-speed over a few hours to physically unreasonable values, affecting all other prognostic fields, leading to numerical failure of the model. Recent improvements in the MetUM have reduced the frequency of GPS at resolutions such as N216, but the frequency of occurrence in the GA3 present climate ensemble was around one failure every nine months, improving to one in 19 months in the GA4 configurations. The procedure for avoiding a GPS is described in Appendix A.

Members of the future climate ensemble initially demonstrated extremely poor numerical stability. This stability was significantly improved by reducing the time step of the simulations from 10 to 7.5 min at the expense of a 20 % reduction in performance.

The development of a new dynamical core for the MetUM, ENDGAME (Even Newer Dynamics for General Atmospheric Modelling of the Environment) (Walters et al., 2014) has eliminated the occurrence of GPS failures in all configurations currently in use, including a 5-year N512 future climate simulation. We expect numerical stability issues will therefore not have a significant impact on similar future projects.
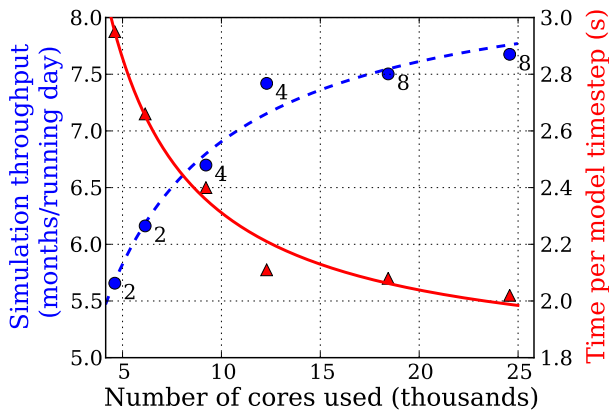
Figure 5. Simulation speed as a function of processor count. Red triangles show time per model time step, blue circles show a calibrated estimation of model throughput. The annotations show the number of OpenMP threads used and lines show least-squares fits to Amdahl's law.

3 Data

3.1 Data specification

The core set of output data used in all runs is an extension of those required for IPCC AMIP-II simulations, with additional fields used in assessment of MetUM global atmosphere configurations, including the tracking of cyclones. The full specification of the individual output fields is long, with more than 500 combinations of field and time and space sampling/averaging, and is therefore documented in the supplementary information attached to this paper.

3.2 Ensemble definition

The full list of simulations in the UPSCALE ensemble is shown in Table 4.

Initial conditions for the GA3 N512 simulations were taken from consecutive days of a testing configuration following a 5-year spin-up run starting from an N320 (40 km) resolution restart file from a previous configuration produced as part of the HadGEM3 development process. Such a period is necessary to allow land surface properties to acclimatise to the different resolution, a process that happens over a period of days to months in the atmosphere. This procedure was performed separately for the present climate and future climate scenarios, and initial conditions for coarser-resolution runs were obtained by regridding the N512 restart files.

The two long GA4 simulations, *xgxqr* and *xgxqx*, were initialised using the same conditions as the second member of the present climate ensemble, with all remaining GA4 runs using restart conditions taken from the 1.5× entrainment rate run *xgxqx*.

The six-member GA4 perturbed initial condition ensemble was initialised from restart files taken from *xgxqx*, with each member perturbed by randomly altering the lowest order bit in the potential temperature field.

3.3 Data management

The most time-consuming aspect of UPSCALE was the management of the output data. Each N512 ensemble member produced around 1 TiB of data each running day, which following a reduction in precision and format conversion produced more than 400 GiB of data for archiving. At the peak of the project, seven simulations were running at once generating more than 2 TiB per real day.

Housekeeping and monitoring tasks were largely automated via a suite of processes on a server attached to JASMIN, which also managed all data transfer tasks. Output data were transferred using gridFTP (Foster, 2006) between dedicated nodes on HERMIT, or HECToR, and JASMIN.

The availability of JANET and GÉANT high speed network links between HERMIT at HLRS in Stuttgart (Germany) and JASMIN at the STFC Rutherford Appleton Laboratory (UK) made sustained data transfer rates of around 4 TiB per day routinely possible using gridFTP, with almost 100 MiB s^{-1} (equivalent to 8 TiB day^{-1}) possible for short periods. This data transfer rate was invaluable in maintaining progress of simulations on HERMIT, as restrictions on bandwidth would in turn have placed limits on the number of running simulations.

A second copy of the UPSCALE data set was made in the UK Met Office archives, with the full transfer of the data set from JASMIN taking approximately 10 months.

4 Conclusions

We have in this paper described the configuration and optimisation of the MetUM, the facilities and procedures behind implementing a large simulation campaign and composition of the UPSCALE ensemble. The success of the operational phase of this project has been contingent on a mix of computing facilities, such as HERMIT and JASMIN, with committed groups of experts who have worked on and supported aspects including extending and adapting the model configuration, data transfer and data hosting. UPSCALE, along with other simulation campaigns such as ATHENA (Kinter et al., 2013) and HiGEM (Shaffrey et al., 2009), demonstrates a clear and growing ability of the climate and weather science community to exploit the largest supercomputing facilities available.

There are several technical matters of note with important implications for future weather and climate projects on this scale. Within climate and weather science we strongly prefer bit-reproducibility, i.e. a given simulation configuration should evolve identically given the same initial conditions

Table 4. Specification of the runs in the UPSCALE data set.

Run identifiers	Resolution	Duration	Notes
GA3 Science, OSTIA SSTs and sea ice (Present Climate)			
xgxq[e[a]*,f,g,h*[b]*,i*[c]*]*	N512 (25 km)	Feb 1985–Dec 2011	
xgxq[o,p,q]	N216 (60 km)	Feb 1985–Dec 2011	
xhqi[j,k,l,o,n]	N96 (130 km)	Feb 1985–Dec 2011	
xgxqj[d]	N512	Jun–Sep: 1988, 1996, 1997, 1998, 2000, 2006, 2008	Time-step data over African and Indian monsoon regions
GA3 Science, OSTIA SSTs and sea ice + climate change signal (Future Climate)			
xgxq[k,l,m]	N512	Feb 1985–Dec 2011	
xgyi[d,e,f]	N216	Feb 1985–Dec 2011	
xhqi[r,s], *xgyip*	N96	Feb 1985–Dec 2011	
GA4 Science, Reynolds SSTs and sea ice (Present Climate)			
xgxqr, *xgxpr*[e]	N512	Feb 1985–Dec 2010	
xgxqs	N512	Sep 2002–Dec 2010	1 h radiation time step
xgxqt	N512	Sep 2002–Dec 2010	5 min time step, high convection limit
xgxqx	N512	Feb 1985–Dec 2010	1.5× Entrainment
xibd[a–f]	N512	Mar 2003–Feb 2004	Perturbed initial condition ensemble

[a] Run extended to August 2012. [b] Additional stratospheric diagnostics included in output data. [c] Additional land surface diagnostics included in output data. [d] Restart files for each season were taken from *xgxqg*. [e] *xgxqr* and *xgxpr* are two sections of the same run performed on HERMIT and MONSooN respectively. The notation *xxxx[a, b, c]* is used to denote ensemble members *xxxxa, xxxxb, xxxxc*.

and ancillary data every time it is run using a particular compiled executable and associated code libraries. As well as being convenient, this makes testing and finding coding faults much easier. Future computing architecture developments may render this preference unsustainable, with consequences for operating procedures. The maintenance of the bit reproducibility preference requires some care, both on the part of scientists using computing facilities and system administrators to keep a clear history of changes to shared code libraries. Supercomputer upgrade cycles can also be disruptive to scientific projects, with hardware alterations preventing data reproduction, therefore increasing the data volume generated with implications for storage costs.

Another non-trivial issue, that we see on many supercomputers, and which may become more significant as supercomputing moves towards the exascale, is that of hardware failures. On several occasions during operations on HERMIT we observed job-step failures that were not triggered by numerical instabilities (GPS) but included errors connected with MPI communications libraries or IO. With multiple jobs requiring a significant fraction of a busy system, it was not uncommon to see clusters of failures, as a faulty node, or network interconnect, was used by each ensemble member sequentially. When provided with information on suspicious computing nodes, the HLRS-Cray support teams reacted rapidly to remove, test and fix the components in question. This type of failure has been seen on many other HPC platforms, so future computing environments, and simulation codes, will need to become fault tolerant, possibly

quarantining or excluding compute nodes with questionable behaviour.

The scientific success of UPSCALE and future projects will be contingent on the exploitation of the data, for which petascale storage and analysis facilities, such as JASMIN, will play a bigger role than the computing platforms used to generate the data. The scale of the "Big Data" issues around simulation campaigns and comparable programmes such as CMIP5 should continue to drive the development and commissioning of substantial analysis facilities.

Alongside the computing and analysis facilities it is important to note that building UPSCALE required a significant level of leadership, commitment and coordination from many people involved. With current levels of available personnel, it would not be possible to repeat this project without compromising our ability to extract scientific value from the data. This, the lengths of available computing grants and supercomputing upgrade cycles, will continue to reinforce the episodic approach taken by us and others to projects of this scale.

Results from our initial analyses of the present and future climate ensembles are in preparation, considering the impact of model resolution on overall climate and climate variability (Vidale et al., 2014), and with specific focus on tropical cyclones (Roberts et al., 2014). We are already working with a number of groups to pursue further analyses, and would welcome approaches from interested scientists.

Appendix A: Grid point storm avoidance procedure

In order to progress model integrations past GPSs, parameter perturbations were applied for a limited period of time, typically 1 month. Following a GPS failure the model was restarted from the restart file for the beginning of the month in which the failure occurred, unless the time series of the maximum vertical velocity in the atmosphere (or the dynamics solver iteration count) suggest that the GPS was already spinning up, in which case the previous months restart file was used. When the model was restarted the targeted diffusion of vertical velocity threshold was reduced from 1.0 to $0.5 \, \mathrm{m \, s^{-1}}$ for a single month before restoring it to the default value. If such a perturbation was not sufficient to avoid the failure point, the number of convection calls per time step was increased from two to three. If neither of these methods evaded the GPS, then the model was restarted from the previous restart file with both parameters perturbed, with defaults restored after 1 month.

All changes to the parameters were logged and time series of their values are available to users of the UPSCALE data set.

Edited by: R. Redler

References

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geosci. Model Dev., 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., Dixon, K., Griffies, S. M., Lee, H.-C., Pacanowski, R. C., Vecchi, G. A., Wittenberg, A. T., Zeng, F., and Zhang, R.: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model, J. Climate, 25, 2755–2781, doi:10.1175/JCLI-D-11-00316.1, 2012.

Demory, M.-E., Vidale, P.-L., Roberts, M., Berrisford, P., Strachan, J., Schiemann, R., and Mizielinski, M. S.: The role of horizontal resolution in simulating drivers of the global hydrological cycle, Clim. Dynam., 42, 2201–2225, doi:10.1007/s00382-013-1924-4, 2014.

Dennis, J. M., Edwards, J., Loy, R., Jacob, R., Mirin, A. A., Craig, A. P., and Vertenstein, M.: An application-level parallel I/O library for Earth system models, Int. J. High Perform. C., 26, 43–53, doi:10.1177/1094342011428143, 2012.

Desgagné, M., McTaggart-Cowan, R., Ohfuchi, W., Brunet, G., Yau, P., Gyakum, J., Furukawa, Y., and Valin, M.: Large atmospheric computation on the earth simulator: The LACES project, Scientific Programming, 14, 13–25, 2006.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sens. Environ., 116, 140–158, doi:10.1016/j.rse.2010.10.017, 2012.

Foster, I.: Globus toolkit version 4: software for service-oriented systems, J. Comput. Sci. Technol., 21, 513–520, available at: http://www.globus.org/alliance/publications/papers/IFIP-2006.pdf (last access: 14 January 2014), 2006.

Hoffman, G.: From Megaflops to Teraflops – The 10th ECMWF workshop Realizing Teracomputing, Proceedings of the Tenth ECMWF Workshop on the Use of High Performance Computing in Meteorology, World Scientific Publishing, 496–203, 2002.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geosci. Model Dev., 4, 543–570, doi:10.5194/gmd-4-543-2011, 2011.

Kinter, J. L., Cash, B., Achuthavarier, D., Adams, J., Altshuler, E., Dirmeyer, P., Doty, B., Huang, B., Jin, E. K., Marx, L., Manganello, J., Stan, C., Wakefield, T., Palmer, T., Hamrud, M., Jung, T., Miller, M., Towers, P., Wedi, N., Satoh, M., Tomita, H., Kodama, C., Nasuno, T., Oouchi, K., Yamada, Y., Taniguchi, H., Andrews, P., Baer, T., Ezell, M., Halloy, C., John, D., Loftis, B., Mohr, R., and Wong, K.: Revolutionizing climate modeling with

project Athena: a multi-institutional, international collaboration, B. Am. Meteorol. Soc., 94, 231–245, doi:10.1175/BAMS-D-11-00043.1, 2013.

Lawrence, B. N., Bennett, V., Churchill, J., Juckes, M., Kershaw, P., Oliver, P., Pritchard, M., and Stephens, A.: The JASMIN super-data-cluster, ArXiv e-prints, available at: http://adsabs.harvard.edu/abs/2012arXiv1204.3553L (last access: 14 January 2014), 2012.

Madec, G.: The NEMO Team, 2008: NEMO ocean engine, Institut Pierre-Simon Laplace Note du Pôle de Modélisation, 27, 2008.

Malcolm, A., Selwood, P., and Glover, M.: Scalability of the Met Office Unified Model, The 14th ECMWF Workshop on High Performance Computing in Meteorology, available at: http://www.ecmwf.int/newsevents/meetings/workshops/2010/high_performance_computing_14th/presentations/Malcolm.pdf (last access: 14 January 2014), 2010.

Mizuta, R., Oouchi, K., Yoshimura, H., Noda, A., Katayama, K., Yukimoto, S., Hosaka, M., Kusunoki, S., Kawai, H., and Nakagawa, M.: 20-km-Mesh Global Climate Simulations Using JMA-GSM Model: -Mean Climate States-, J. Meteorol. Soc. Jpn., 84, 165–185, 2006.

Navarra, A., Kinter, J. L., and Tribbia, J.: Crucial experiments in climate science, B. Am. Meteorol. Soc., 91, 343–352, doi:10.1175/2009BAMS2712.1, 2010.

Ohfuchi, W., Enomoto, T., Takaya, K. and Yoshioka, M.: 10-km mesh global atmospheric simulations, Realizing Teracomputing: Proceedings of the Tenth ECMWF Workshop on the Use of High Performance Computing in Meteorology, World Scientific Publishing, 47–57, 2002.

Ohfuchi, W., Nakamura, H., Yoshioka, M. K., Enomoto, T., Takaya, K., Peng, X., Yamane, S., Nishimura, T., Kurihara, Y., and Ninomiya, K.: 10-km mesh meso-scale resolving simulations of the global atmosphere on the Earth Simulator: Preliminary outcomes of AFES (AGCM for the Earth Simulator), Journal of the Earth Simulator, 1, 8–34, 2004.

Reynolds, R., Rayner, N., Smith, T., Stokes, D., and Wang, W.: An improved in situ and satellite SST analysis for climate, J. Climate, 15, 1609, doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2, 2002.

Roberts, M. J., Clayton, A., Demory, M.-E., Donners, J., Vidale, P. L., Norton, W., Shaffrey, L., Stevens, D. P., Stevens, I., Wood, R. A., and Slingo, J.: Impact of resolution on the Tropical Pacific Circulation in a matrix of coupled models, J. Climate, 22, 2541–2556, doi:10.1175/2008JCLI2537.1, 2009.

Roberts, M. J., Vidale, P.-L., Mizielinski, M. S., Demory, M.-E., Schiemann, R., Strachan, J., Hodges, K., Camp, J., and Bell, R.: Tropical cyclones in the UPSCALE ensemble of high resolution global climate models, J. Climate, in revision, 2014.

Roberts-Jones, J., Fiedler, E. K., and Martin, M. J.: Daily, global, high-resolution SST and sea ice reanalysis for 1985–2007 using the OSTIA system, J. Climate, 25, 6215–6232, doi:10.1175/JCLI-D-11-00648.1, 2012.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K.: Improved Atlantic winter blocking in a climate model, Geophys. Res. Lett., 38, L23703, doi:10.1029/2011GL049573, 2011.

Schiemann, R., Demory, M.-E., Mizielinski, M. S., and Roberts, M. J., Shaffrey, L. C., Strachan, J., and Vidale, P. L.: The sensitivity of the tropical circulation and

Maritime Continent precipitation to climate model resolution, Clim. Dynam., 42, 2455–2468, doi:10.1007/s00382-013-1997-0, 2014.

Sell, W.: Distributed Data Management at DKRZ Use of High Performance Computing in Meteorology, World Scientific Publishing, 2004, 108–130

Selwood, P.: The Met Office Unified Model I/O Server, ENES Workshop: Scalable IO in climate models, available at: https://verc.enes.org/computing/hpc-collaborations/parallel-i-o/workshop-scalable-io-in-climate-models/presentations/Unified_Model_IO_Server_Paul_Selwood.ppt (last access: 14 January 2014), 2012.

Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., Jrrar, A., Stevens, D. P., Woodage, M. J., Demory, M. E., Donners, J., Clark, D. B., Clayton, A., Cole, J. W., Wilson, S. S., Connolley, W. M., Davies, T. M., Iwi, A. M., Johns, T. C., King, J. C., New, A. L., Slingo, J. M., Slingo, A., Steenman-Clark, L., and Martin, G. M.: UK HiGEM: the new UK High-resolution Global Environment Model – model description and basic evaluation, J. Climate, 22, 1861–1896, doi:10.1175/2008JCLI2508.1, 2009.

Strachan, J., Vidale, P. L., Hodges, K., Roberts, M., and Demory, M.-E.: Investigating global tropical cyclone activity with a hierarchy of AGCMs: the role of model resolution, J. Climate, 26, 133–152, doi:10.1175/JCLI-D-12-00012.1, 2013.

Taylor, K., Williamson, D., and Zwiers, F.: The sea surface temperature and sea-ice concentration boundary conditions for AMIP II simulations, PCMDI Rep. 60, Tech. Rep. 60, PCMDI, 25 pp., available at: http://www-pcmdi.llnl.gov/publications/ab60.html (last access: 14 January 2014), 2000.

Vidale, P.-L., Roberts, M. J., Mizielinski, M. S., Schiemann, R., Demory, M.-E., Strachan, J., Shaffrey, L. C., Zappa, G., and Tsushima, Y.: Weather-permitting climate models, J. Climate, in preparation, 2014.

Walters, D. N., Best, M. J., Bushell, A. C., Copsey, D., Edwards, J. M., Falloon, P. D., Harris, C. M., Lock, A. P., Manners, J. C., Morcrette, C. J., Roberts, M. J., Stratton, R. A., Webster, S., Wilkinson, J. M., Willett, M. R., Boutle, I. A., Earnshaw, P. D., Hill, P. G., MacLachlan, C., Martin, G. M., Moufouma-Okia, W., Palmer, M. D., Petch, J. C., Rooney, G. G., Scaife, A. A., and Williams, K. D.: The Met Office Unified Model Global Atmosphere 3.0/3.1 and JULES Global Land 3.0/3.1 configurations, Geosci. Model Dev., 4, 919–941, doi:10.5194/gmd-4-919-2011, 2011.

Walters, D. N., Williams, K. D., Boutle, I. A., Bushell, A. C., Edwards, J. M., Field, P. R., Lock, A. P., Morcrette, C. J., Stratton, R. A., Wilkinson, J. M., Willett, M. R., Bellouin, N., Bodas-Salcedo, A., Brooks, M. E., Copsey, D., Earnshaw, P. D., Hardiman, S. C., Harris, C. M., Levine, R. C., MacLachlan, C., Manners, J. C., Martin, G. M., Milton, S. F., Palmer, M. D., Roberts, M. J., Rodríguez, J. M., Tennant, W. J., and Vidale, P. L.: The Met Office Unified Model Global Atmosphere 4.0 and JULES Global Land 4.0 configurations, Geosci. Model Dev., 7, 361–386, doi:10.5194/gmd-7-361-2014, 2014.

Walters, D. N., Williams, K. D., Boutle, I. A., Bushell, A. C., Edwards, J. M., Field, P. R., Lock, A. P., Morcrette, C. J., Stratton, R. A., Wilkinson, J. M., Willett, M. R., Brooks, M. E., Copsey, D., Earnshaw, P. D., Harris, C. M., Manners, J. C., MacLachlan, C., Palmer, M. D., Roberts, M. J., and Ten-

nant, W. J.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1.0 configurations, in preparation, 2014.

Williamson, D. L.: The effect of time steps and time-scales on parameterization suites, Q. J. Roy. Meteor. Soc., 139, 548–560, doi:10.1002/qj.1992, 2013.