HiGEM Decadal Predictions: Final Report

Len Shaffrey, Lois Steenman-Clark, Grenville Lister (National Centre for Atmospheric Science), David Stevens (University of East Anglia), Jason Beech-Brandt, Tom Edwards (Cray Centre of Excellence)

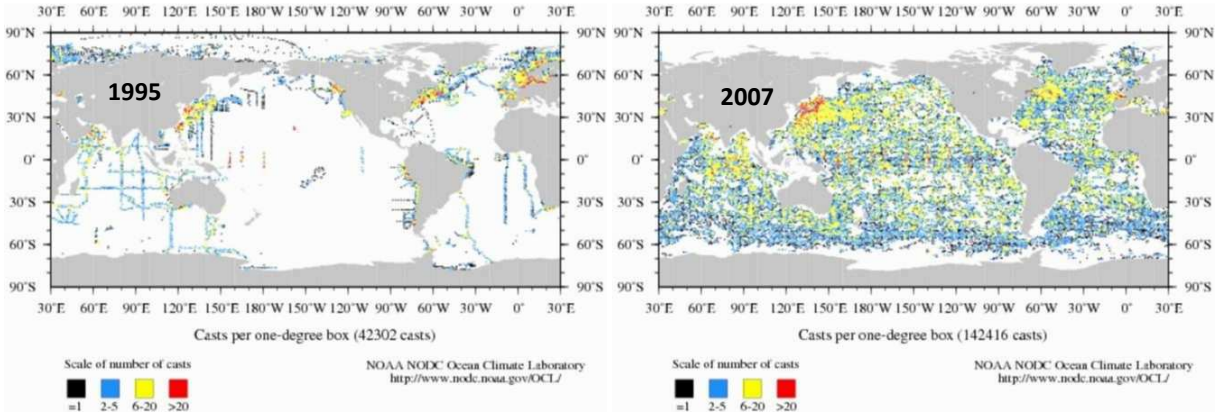

Fig 1: The location and density of ocean temperature and salinity profiles taken during 1995 (left) and 2007 (right). The dramatic increase in density and coverage of ocean observations is due to the introduction of Argo floats.

The scientific goal of this project is produce a set of decadal climate predictions using the HiGEM high-resolution climate model. HiGEM will be initialised with recent ocean observations using Argo floats. Argo floats have dramatically increased the density and coverage of the ocean observing network; this will allow the HiGEM model to be evaluated in greater detail than has been previously possible. The project has implemented a more flexible I/O strategy for HiGEM to facilitate the management of the large data volumes produced by the decadal predictions.

Introduction

Improving climate predictions is one of the key challenges facing climate science. A recently developed methodology for climate model prediction and evaluation is *decadal prediction*, where climate predictions are initialised using information about the current state of the ocean. At Reading University we are using a high-resolution climate model, HiGEM, to produce decadal predictions (every 5 years, from 1960 to 2005) for the forthcoming IPCC assessment report (IPCC AR5). This is a computationally demanding project that requires resources of a HPC platform such as HECToR. One of the limiting factors of the decadal predictions that are currently being produced for IPCC AR5 is that they are mostly for an era when ocean observations were sparse and had poor global coverage. The release of the Argo floats

has brought about a step change in the density and coverage of the ocean observing network (compare the coverage of ocean profiles taken in 1995 and 2007 in Figure 1). **The scientific goal of this proposal is to initialise decadal predictions using HiGEM for the Argo period (every year, from 2002 onwards). By using the Argo observations, the aim is to achieve a more constrained evaluation of the model than has previously been possible.**

HiGEM is based on the Met Office Unified Model (UM), and the data intensive nature of high-resolution climate modelling is hindered by the rigid I/O strategy that is presently implemented in the UM. **The project has implemented a more flexible, asynchronous I/O strategy for the UM, which will directly benefit both the Argo HiGEM decadal predictions, and UM modelling activities within the wider climate science community.**

Scientific Progress

Over the past year, scientists from National Centre for Atmospheric Science, the University of East Anglia and the Met Office Hadley Centre have been involved in the development of the HiGEM decadal prediction framework. The decadal predictions for the Argo period (every year, from 2002 to 2009) are now underway. Preliminary hindcasts of SST (sea surface temperature) for the North Atlantic subpolar gyre are shown in Fig. 2. These results will be become part of the IPCC AR5 archive.

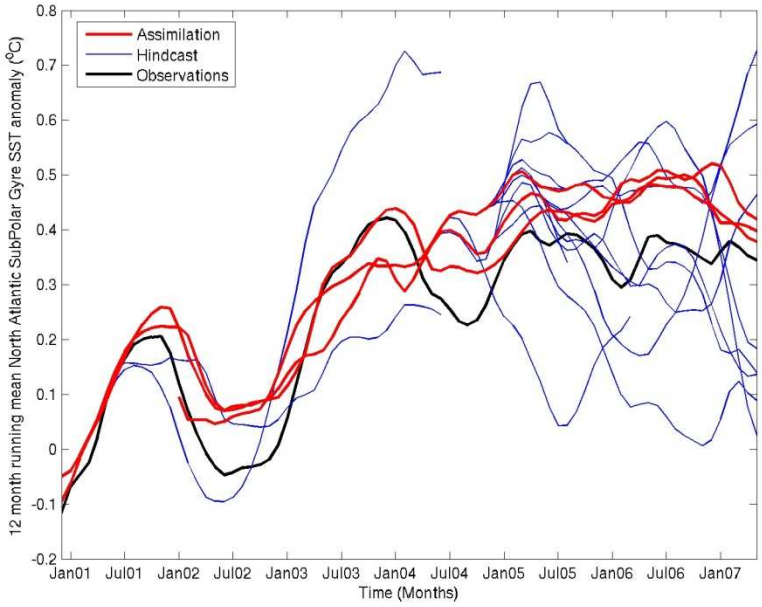

Figure 2: Preliminary results from the HiGEM decadal prediction system; forecasting the North Atlantic subpolar gyre sea surface temperatures for the Argo period. Black: observations (HadISST). Red: HiGEM assimilation integrations. Blue: HiGEM hindcasts.

Preliminary results are encouraging. HiGEM appears capable of capturing some of the observed variability in the North Atlantic Ocean, particularly during 2001 to 2003. These very early results suggest the model has multi-annual predictive skill in some cases. Evaluating the model against the high density

Argo ocean observations will allow us to determine whether the model is capable of capturing the processes that appear to be governing multi-annual variability in the North Atlantic Ocean. Over the next six months the hindcast set for the Argo period will be completed which will provide a novel and unique high-resolution climate model dataset to further investigate questions surrounding climate predictability and fidelity of climate models.

Technical Achievements

One of the difficulties in climate science is managing the large and complex datasets produced by climate models. The technical aim of this project was to implement a more flexible I/O strategy for the code upon which HiGEM is based, the Met Office Unified Model (UM). This has been achieved with the asynchronous I/O server initially developed by Cray. The asynchronous I/O server potentially allows increased performance of the UM, since I/O is handled by dedicated processors rather than processors involved in the model computation.

Results from a series of test runs are shown below in Figure 3, indicating the I/O server can effectively mitigate the I/O time in the Unified Model. One aspect of the project that we have been unable to implement is using the I/O servers for post-processing. Common UM post-processing tools have the potential to consume large amounts of resources on HECToR which are expected to be accelerated by a dedicated external server.

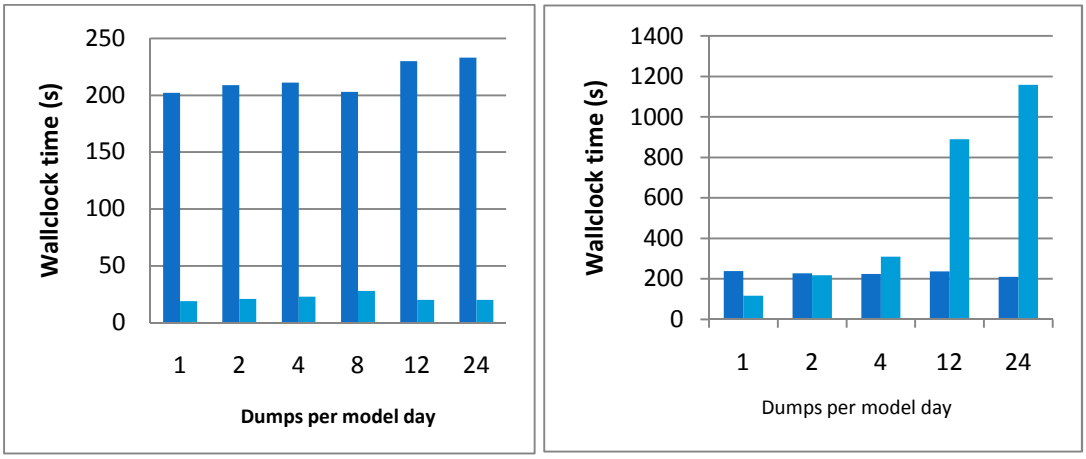

Figure 3. N216 atmosphere model on 512 processors with asynchronous I/O (left) and with 'single-writer' I/O (right). The performance is dominated by I/O at high data volumes for the 'single-writer' I/O. Dark blue columns denote model compute time, light blue model I/O time.

The asynchronous I/O server has also been successfully implemented in HiGEM. HiGEM is not as severely I/O dominated as the test cases above, and so the speed-up is more modest (5% to 10%). The

I/O server will be extremely valuable for other UM projects. For example, the I/O server is presently being implemented in the Cascade project. The aim of Cascade is to study organized convection in the tropical atmosphere using large domain cloud system resolving (approx. 1km) model simulations. The Cascade simulations are extremely I/O intensive. Figure 4. Shows some preliminary results from a Cascade UM configuration where speed-ups of over 50% are found using the I/O server. The technical outcomes of the HiGEM Cray CoE project are to be disseminated at the NCAS conference in July 2010.

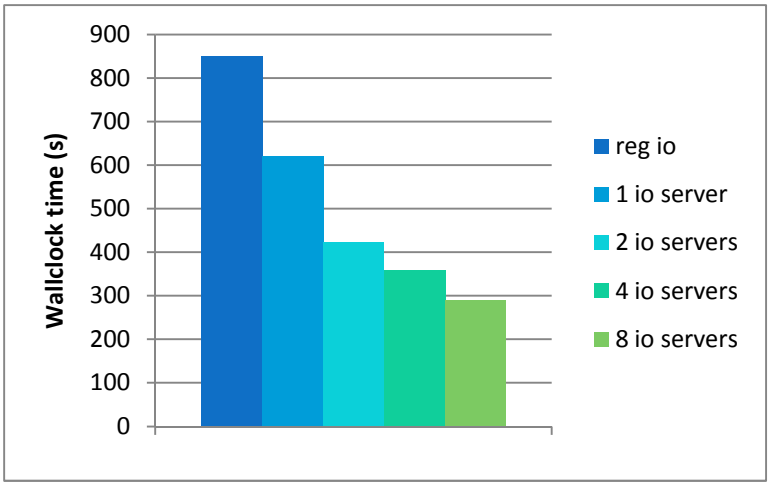

Figure 4. Wallclock time for the Cascade African LAM configuration using different numbers of processors for the I/O server. 'reg io' refers to the performance of the 'single-writer' model. The compute processors are throttled in all cases except for 8 I/O servers, in which case the I/O slaves are consuming data faster than the compute processors can produce it. There is no further performance improvement for ten I/O slaves.

Outcomes and Future Directions

Scientific Progress: The production of the Argo decadal predictions has begun as scheduled (June 2010). Preliminary results are very encouraging suggesting the model may have multi-annual predictive skill in the North Atlantic.

Technical Outcomes: Final issues with The I/O server code have been resolved and the code has been implemented in HiGEM. The I/O server is also being implemented in projects such as Cascade, where large speed-ups may be possible.

Remaining work and Future Directions: The Gantt chart below outlines the remaining tasks for the Argo decadal prediction project, primarily completing the production of the HiGEM Argo decadal predictions. Future directions may include the implementing post-processing on a remote platform by extending the I/O server code.

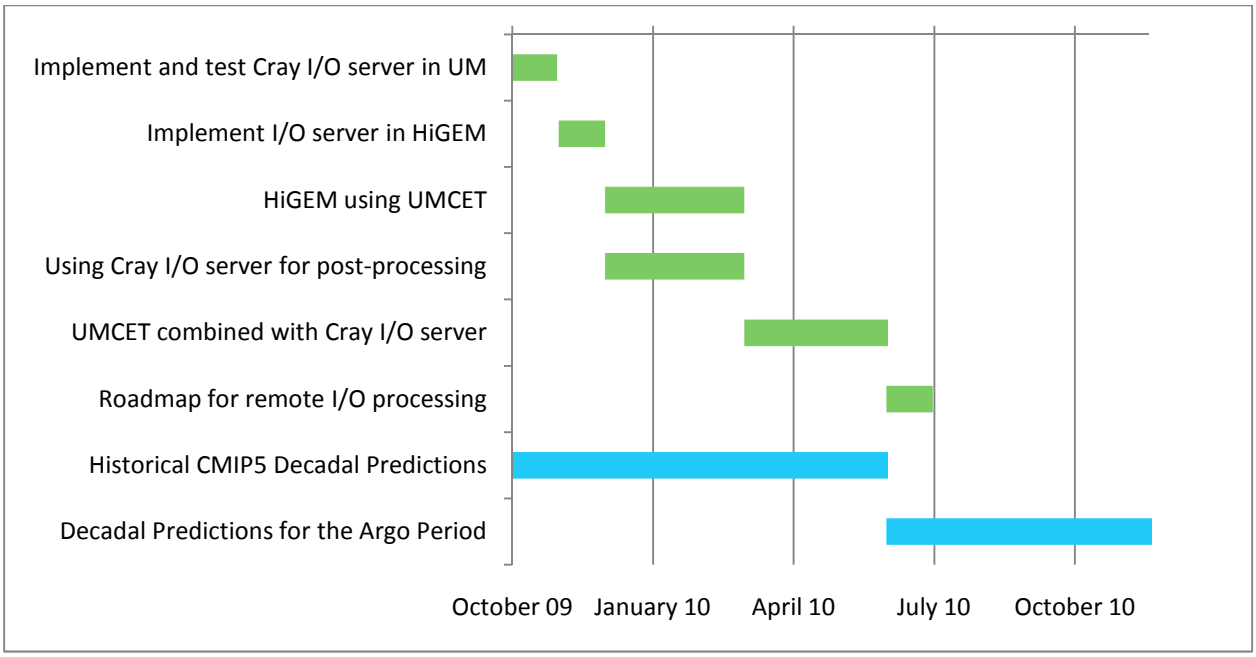

Fig 5: Gantt chart outlining the HiGEM Argo decadal prediction project. Technical development tasks are in green; scientific tasks in blue. The project is scheduled to complete at the end of 2010.