TerraFLOPS
Service for earth-system-research
Newsletter
Deutsches Klimarechenzentrum
Modelle & Daten

Issue Number 4 - August 2003

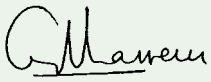
TerraFLOPS
Service für die Erdsystemforschung

New Head of M&D

(by Guy Brasseur) It is with pleasure that I have appointed Dr. Christoph Heinze as the new Project Manager of the „Model and Data" Group.

Christoph Heinze has been selected by an international search committee after a world-wide search.

Christoph Heinze, who until then worked at the National Environmental Research Institute of Denmark (Department of Marine Ecology), started his new job on 1 April, 2003. We wish Christoph Heinze all the best for his new tasks serving the national scientific community and continuing the very successful work of Dr. Ulrich Cubasch, who recently became Professor at the Free University of Berlin.



In this issue

World Data Center (WDC) for Climate established in Hamburg

February this year the ICSU Panel on World Data Centers sent the approval for the World Data Center for Climate (WDCC) in Hamburg. The WDCC is managed by M&D in close cooperation with the DKRZ. Operations started in April 2003 by publication of the WDCC homepage:

http://www.mad.zmaw.de/wdcc/

The homepage documents data content, data access, contacts and available services.

• **User Services:** A visiting scientist program exists. Data processing, copying and analysis facilities are available. Data are available on most media including CD-ROM, via Internet, and other media on request. On-line access via World-wide Web is the standard method. FTP access is possible on request.

• **Scientific Data Management:** The WDCC is aimed at collecting, scrutinizing, and disseminating data related to climate change on all time scales. Emphasis is on data products from climate modelling and related observational data. The WDCC focuses on geo-referenced data using the operational CERA data and information system. Input is accepted in electronic form, specifics have to be discussed with the WDCC staff.

• **Project Support:** Project data

➜ **Page 2**

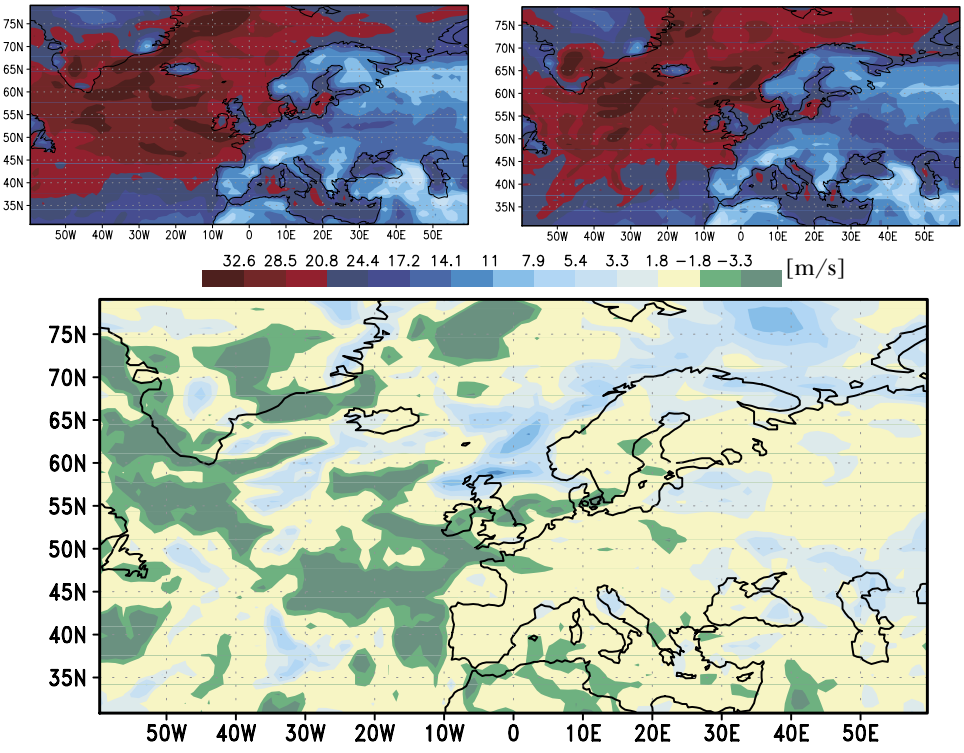


Figure 1: *Anomaly of January maximum wind speed in the North Atlantic – European sector as calculated by the ECHAM4-T106 climate model for two decades (1971-1980; upper leftt and 2041-2050; upper right) of the IS92a scenario.*

management will be accomplished at cost price basis. Details depend on the individual project and have to be discussed with the WDCC staff.

The WDCC is part of the operational CERA (Climate and Environmental data Archiving and Retrieval) database system of M&D and DKRZ. Most data of CERA are available without restriction and are therefore part of the WDCC. Since there is no specific WDCC data portal available the standard CERA user interface (Fig 2)

http://mad.dkrz.de/java/CeraStart.html

is used for catalogue inspection and for climate data retrieval. The CERA data model as implemented by M&D contains two parts as realisation of the semantic climate data management (Fig.4). These two parts are the metadata or data catalogue and the application oriented storage of climate data. Metadata as well as climate data are stored in tables of a relational database system. The CERA metadata model (Fig.3) allows for data search according to discipline, keyword, variable, project, author, geographical region and time interval and additionally for climate data retrieval. The metadata model is complete with respect to specification of data processing without attaching the primary data. The climate data are archived within CERA for global climate model results as complete time series of global fields from individual variables. Most of these time series can be extracted from the WDCC and downloaded to the requesting client computer. The data download can be scaled from one table entry (one global field at a specific time step) up to one complete time series of the corresponding model experiment. The expected download size is indicated in the user interface.

Most of the total data volume of CERA and WDCC is primary data. Less than 2% are metadata. Primary data in CERA are structured into (climate model) experiments and datasets:

- Current database size is 14.5773 Terabyte
- Number of experiments:293
- Number o datasets: 19775
- Number of BLOBs within CERA at 22-APR-03 1007264826

Currently the WDCC main focus is set to climate model related data sets. Four different areas can be distinguished.

- **DKRZ: Climate model results** from global and regional climate models calculated at DKRZ in Hamburg are available from the WDCC. This includes present-day climate, paleo-climate simulations and IPCC-scenario runs for the future. The ECHAM-, OPYC-, LSG-, HOPE or REMO models generate most of these data.
- **International Data Centres:** Presently M&D administers two international data centres as part of the WDCC. The IPCC Data Distribution Centre (DDC) is run by the Climatic Research Unit, University of East Anglia (CRU) in the UK, by the Model and Data Group at the Max-Planck-Institute for Meteorology in Hamburg, and the Centre for International Earth Science Information Network (CIESIN) in the USA. The DDC contains: the results from climate change experiments; observed climate and environmental data; socio-economic datasets; and accompanying scientific advice and technical guidance material. The DDC has been founded 1997 in preparation of the Third Assessment Report (TAR) and it will continue during the compilation of the Fourth Assessment Report. The IPCC-Data Distribution Centre (DDC) in Hamburg as part of the WDCC contains the climate model data basis for the IPCC assessment reports. The IPCC DDC contains model results (monthly means) from different climate modelling centres (CCCma, CCSR/NIES, CSIRO,

➜ Page 6

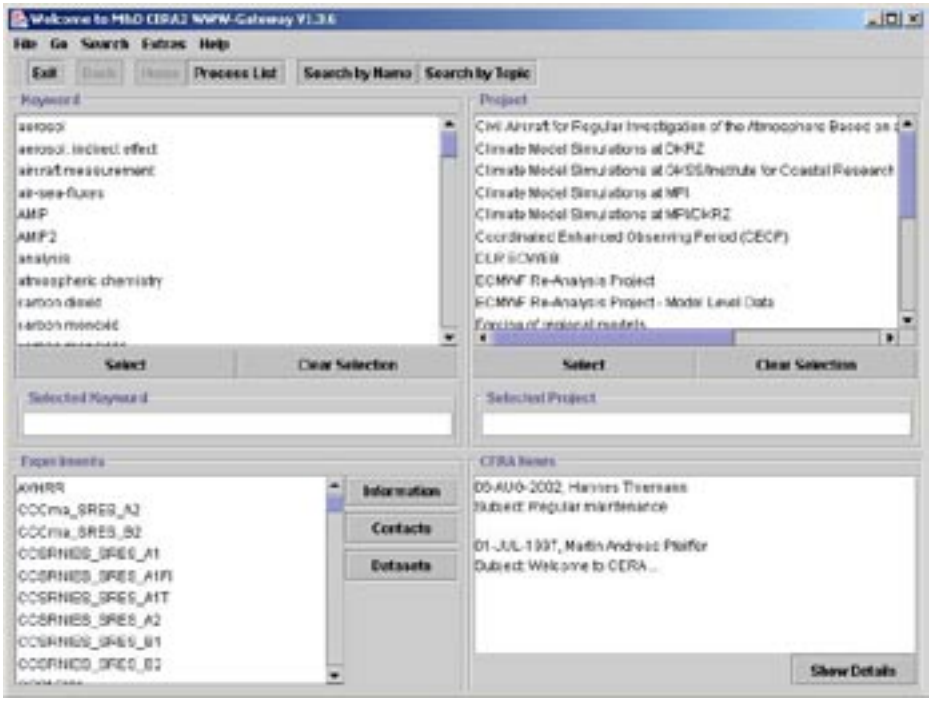


Figure 2: *The CERA graphical user interface is realised as signed Java applet and can be accessed with standard WWW-browsers. The GUI allows for catalogue inspection and climate data retrieval from the CERA database system.*

Report on the 2. WLA Workshop „Community models for climate research in Germany", 20.-21.2.2003 in Hamburg

The first WLA workshop on „Community models for climate research in Germany" was held at Bremerhaven in January 2001. In February 2003 a first follow-up workshop on this topic was carried out. In its first part, this symposium was dedicated to an overview on modelling activities in the German climate research community. Impressive results and developments in practically all relevant areas of earth system research were presented. The discussion of the models included atmosphere (physics, chemistry), ocean (physics, biogeochemistry, ecosystems), and the biogeochemistry of the land surface. The components ice sheet, seaice, ocean waves, and integrated assessment still need more consideration though they were discussed also during the meeting. Global as well as regional model versions were presented for each earth system sub-component. Next to general circulation models (grid point models, spectral models) also EMICS (earth system models of intermediate complexity) were presented. Model coupling was a further focus of the workshop. Extensive discussions on the development of standardised coupling interfaces within the PRISM project were carried out. Its final goal is the establishment of an infrastructure, which allows easy versatile crosswise coupling of different single component models with each other.

On the way to a well accepted common model platform for the German climate research community still some obstacles have to be overcome. The „Model and Data Group" will firmly further support this process. Among others, the following questions have to be answered and the associate problems have to be tackeled:

(1) How many parallel model developments for each component model should be supported ? We returned to the approach of the „minimal multitude". According to this approach, at present, not yet one single model for an earth system component can be selected as the „best" or „best suited" component, but an appropriate prioritised choice of models among the available initiatives must be made.

(2) Can the goal of a single central community model system be realised, a system which would be supported with highest priority ? In order to re-establish the German climate research more clearly as a „global player" and to keep that status, the focussed support of one single community model system makes sense and may even be necessary. If the selection of respective model components from the existing pool cannot be made ‚ad hoc', the standardised coupling of different models provides the alternative to let such a selection be the result of an evolutionary process. A further option would be the merging of parallel model developments in order to combine the advantages of the model components in question in one unified component.

(3) How can the regional climate modelling in Germany make progress ? A quite vivid discussion arose on the choice for the support of the potential model candidates for regional climate research.

In view of the limited personnel available, a decision on the priorities for the replacement or addition of well functioning model systems by necessary new developments (non-hydrostatic models) has to be made. A consensus about this has to be found within the modeller community in order to produce a competitive model for future reserach on one hand, and to continue the ongoing daily work. Moreover, a future regional model ideally should be fully embedded within an existing global model framework, so that nestings and zoomings on arbitrary regional sub-sets will become easily possible.

These discussion items will be dealt with also on the forthcoming meeting of the scientific steering commitee in June. In any case, the existing working material is excellent and provides a solid basis for a comprehensive community model system, which will be competitive in Europe as well as worldwide for the solution of pressing scientific and infrastructural problems.

The „Model and Data Group" serves as a central node in order to bring this project to life. Progress will be made step by step in close coordination with the German climate research institutions.

[Christoph Heinze, M&D]

High-Resolution Modelling of Deep Water Formation in the Subpolar North Atlantic and its Role for Fluctuations of the Meridional Overturning Circulation

[J.-O. Beismann, L. Czeschel, Institut für Meereskunde Kiel]
http://www.ifm.uni-kiel.de/fb/fb1/tm/tm-e.htm
[René Redler, C&C Research Laboratories, NEC Europe Ltd., St. Augustin]
http://www.ccrl-nece.de/
[Klaus Ketelsen, Software Consultant, Berlin]
klaus.ketelsen@t-online.de

The subpolar North Atlantic is a key region for the large-scale circulation of mass and heat in the Atlantic Ocean: Warm surface waters that have been advected northward by the Gulf Stream – North Atlantic Current system are subject to vigorous cooling during winter which leads to an increase in density and to deep convective mixing of surface waters, thereby transforming these water masses to North Atlantic Deep Water (NADW). The southward transport of NADW forms the lower limb of the meridional overturning circulation (MOC), a key component of the Earth's climate system.

In the framework of "Sonderforschungsbereich 460" [1] intensive studies (both observational and theoretical) have been conducted with the aim to identify and understand the processes that govern the physical aspects of the MOC. Special emphasis has been put on MOC fluctuations in response to low-frequency (interannual to decadal) atmospheric variability and to variations in the dense overflows from the Arctic Ocean, on the export pathways of NADW and its components to the subtropics, and on the quantification of the oceanic uptake of anthropogenic trace gases (such as CO_2 and CFC). An important part of these studies is the FLAME [2] ("Family of Linked Atlantic Model Experiments") model hierarchy: Based on the MOM [3] code, both basin-wide and regional models of different horizontal resolution are being used to investigate the char-

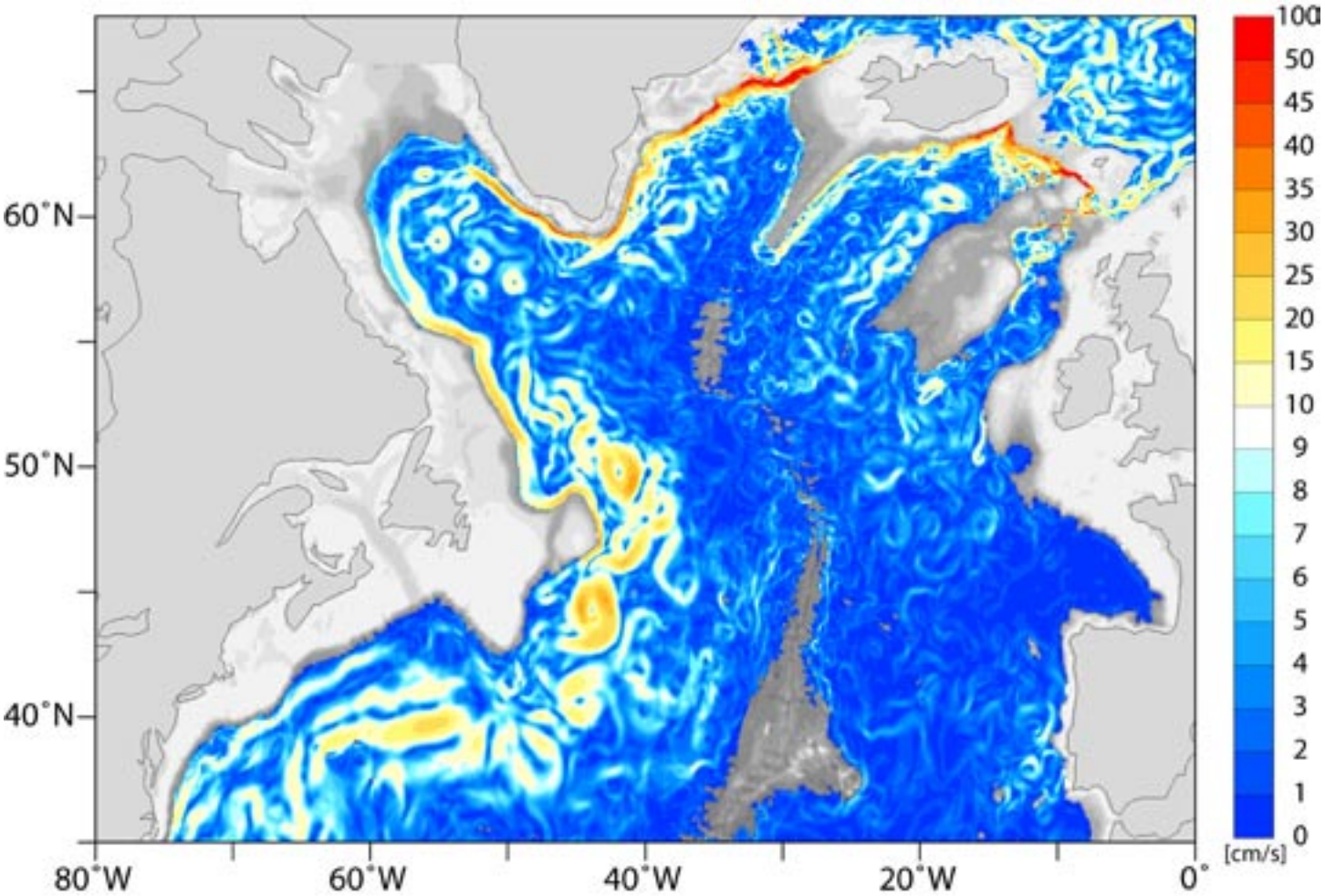


Figure 1: Magnitude of the horizontal velocity (cm/s) on the isopycnal surface $\sigma_0 = 27.85$ (FLAME 1/12° model).

acteristics of the MOC on different time scales. A non-eddy resolving version (4/3° resolution, corresponding to 74 km grid spacing at 60°N) is employed in climate change studies, eddy-permitting models (1/3°, 18.5 km) are used in long-term integrations for tracer uptake and in sensitivity studies, while the eddy-resolving configuration (1/12°, 4.6 km) enables us to analyse thermohaline processes in detail and serves in the interpretation of the (temporally and spatially inhomogeneous) observational data collected during the field experimens of SFB 460.

The model domain of the 1/12° model covers the region from 18°S to 70°N and from 100°W to 16°E. The model has 1021x1406 horizontal grid points and 45 vertical levels (approximately 65 million grid points). On the NEC SX6 system at DKRZ a 1-D horizontal domain decomposition is used, resulting in an inner loop length close to an integer multiple of the vector register size. Communication between processors is realized using MPI. Using state-of-the-art parameterisations for subgrid scale and bottom boundary layer processes, the model has a memory size of approximately 25 GB. Typical simulations span time periods of 5-10 years. Running the model on 2 nodes (16 processors), a 10 year simulation requires 30 days real time in dedicated use. The average performance of this configuration is 45 GFlops, corresponding to more than 35% of the theoretical peak performance. Based on these numbers, it is now conceivable to conduct not only idealized process and sensitivity studies, but also to carry out a "hindcasting" simulation of the period 1990-2003 including anthropogenic tracers, which will result in a unique data basis allowing for a detailed interpretation and synthesis of observational data.

Number of CPUs	GFlop/s
8	25.3
16	45.3
24	57.2
32	71.6

The table illustrates the scaling behaviour of our code in single- and multi-node applications. While the routines working on the 3-D fields scale perfectly (including the communication), total linear scalability is prevented by the 2-D conjugate gradient (CG) solver which is called once per time step. The CG solver requires additional communication in every iteration step, and the summation around islands creates both load imbalance and further communication.

A simulation of one month produces approximately 1.7 GB of output data. Each processor writes a snapshot of the output variables of its subdomain at predefined time intervals. In a postprocessing step the data is recombined for the whole domain and (usually) averaged to monthly or seasonal means. For certain applications (e.g. off-line Lagrangian float diagnostics), however, high-frequency output fields are needed, which leads to the necessity to store between 200 and 500 GB of data permanently (for a simulation of 10 years, depending on the output frequency). The enormous computing power of the HLRE system thus implies the demand for sufficient storage capacities and for the installation of a machine on which postprocessing and analysis applications can be executed (including easy access to the storage system).

As an example for the quality of our simulations, Figure 1 shows the magnitude of the horizontal velocity on an isopycnal surface in the NADW range (σ_0=27.85; September monthly mean of the 16th year of simulation). The depth of this surface is 400 – 600 m in the region north of the Greenland-Iceland-Scotland ridge, 1000 – 1400 m in the boundary currents of the subpolar gyre, and it descends to depths greater than 2000 m in the interior and subtropical ocean. Our simulation reproduces with a high degree of realism the narrow structure of the boundary current system around Greenland and along the Labrador shelf. In the Labrador and Iceland basins, mesoscale vortices play an important role for lateral mixing of water masses. A very interesting feature of this experiment is the presence of an intense eddy field around the Grand Banks of Newfoundland. These eddies seem to be crucial in shaping the details of the pathways of NADW export, a question that has become a focus of the physical investigations in SFB 460. The collaboration with sea-going oceanographers will thus greatly benefit from the simulations carried out on the NEC SX6 system.

REFERENCES

[1] *www.ifm.uni-kiel.de/allgemein/research/projects/sfb460/sfb460-d.htm*
[2] *www.ifm.uni-kiel.de/fb/fb1/tm/research/FLAME/index.html*
[3] *www.gfdl.gov/MOM/MOM.html*

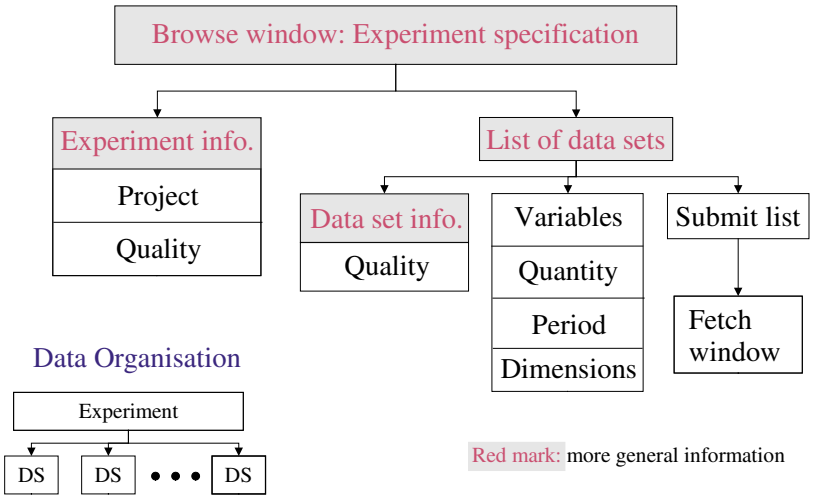


Figure 3: *Structure of the CERA metadata model as implemented by M&D in its climate database system*

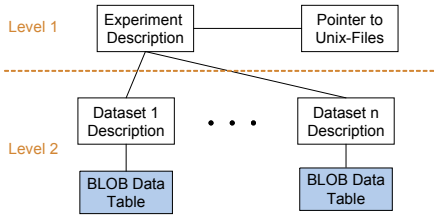


Figure 4: *The semantic oriented data management within CERA contains two levels of complexity. Level 1 includes the complete data description and the pointers to data files in the mass storage archive. Level 2 additionally contains database tables, which archive climate data with respect to their expected application. For climate model data these data tables contain e.g. complete time series of global fields from individual climate variables.*

GFDL, HADLEY, MPIfM , NCAR) based on IPCC-emission scenarios (IS92, SRES).

The M&D/MPIM contribution to CEOP (Co-ordinated Enhanced Observing Period) will be to incorporate the CEOP model output data into their CERA data base system through the application of the most current scientific data management techniques available and as a result they will be considered the Central CEOP Model output retention and handling Centre (CEOP-CMOC). The CEOP CMOC is part of the WDCC. M&D has responsibility for control of both the MOLTS and the gridded data products. Other mirror sites for either the MOLTS data and some or all of the gridded data that may be implemented possibly at UCAR/JOSS and the university of Tokyo will be established as official CEOP data centres that secondary or subordinate to MPIM. Such sites would be set up formally only through direct knowledge of and connection with M&D/MPIM.

• **Scientific Projects:** Data from scientific projects become part of the WDCC and are therefore accessible for the community after a validation period of the data. Currently the HOAPS (Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data) and the CARIBIC (Civil Aircraft for Regular Investigation of the Atmosphere Based on an Instrumentation Container) provide data to the CERA database. In co-operation with the BSH (Hamburg) observations from WOCE (World Ocean Circulation Experiment) are integrated into CERA.

• **Model like Observations** Reanalysis data from ECWMF (ERA15) and NCAR (NECP40) are available. The new ECMWF-Reanalysis Project (ERA40) is under way and we intend to include these data when they become available. Access to ECMWF data via the CERA database system is restricted to German scientists due to regulations of the European national weather services.

An example of the WDCC data content can be obtained from the IPCC scenario calculations of the ECHAM4 climate model. For the IPCC scenario IS92a („business as usual“) two decades have been recalculated with the high-resolution version T106. The data represent a spatial resolution of about 130 km. The two decades are 1971-1980 for the present and 2041-2050 for the CO2-doubling. Fig. 1 (cover page) shows the maximum January wind speed near the surface for the two decades in the North Atlantic – European sector and the anomaly between the two decades. The model shows an increase at the Norwegian coast and in central Europe. These data, the wind speed maximum in 6 hour time resolution, can be directly accessed from the WDCC because of the application oriented data storage. But even in this case of direct data access 250 Mbytes for each decade have to be processed.

For the future the data content of the WDCC is planned to be extended by instrumental observations and additional project data. The development of networking between the CERA database system and other data archives has been started. Then it is not necessary to copy data physically into a mirror section of the WDCC, as it has been done for the reanalysis data in the past. Networking allows to access climate data in remote archives while the physical location is transparent to the user. Additionally project data archive support is offered within the resource limits of the WDCC. Projects are welcome to use the existing facilities and experience under the regularities of the WDC system.

[Michael Lautenschlager. M&D, Director WDC]

International Supercomputing Conference (Heidelberg, June 2003)

Under the Headline "Focussing Todays HPC Key Application" this years International Supercomputing Conference in Heidelberg

http://www.isc2003.org/

saw a session as well as a panel discussion on the topic of "Earth System Modeling" - a fact which emphazises the importance of High Performance Computing (HPC) for Earth System Research and vice versa.

From the conference webpage:

The only laboratories available for Climate Change and Earth System Research are numerical models. To apply global high resolution models scientists have to use High Performance Computers. Traditionally and in modern times HPC systems and compute centres devoted to meteorological applications were high in the TOP500 ranking, just to mention a few:

- *NCAR's long line of Cray and IBM machines*
- *ECMWF systems from Cray, Fujitsu and most recently IBM*
- *The German Climate Compute Centre (**DKRZ**), as one of the very few compute centres devoted to climate research exclusively, with Cyber, Cray and NEC systems*
- *Last but in no sense least the Earth Simulator Project, employing at least for a few years to come an NEC system being the worlds most powerful machine*

The session was chaired by Reinhard Budich., MPI Hamburg, Speakers and Panelists included: Ulrich Cubasch, FU Berlin, Warren M. Washington, NCAR, Walter Zwieflhofer, ECMWF, and Tetsuya Sato, Earth Simulator Center.

The latest TOP500 lists our NEC SX-6 on rank 33nd showing that DKRZ features as one of the most powerfull HPC centers in germany and as one of the leading centers for earth system modelling worldwide. See:

http://.www.top500.org/

Meeting

Das Herbsttreffen 2003 des **Arbeitskreises Supercomputing** der Zentren für Kommunikation und Informationsverarbeitung in Lehre und Forschung e.V. (ZKI) findet am Deutschen Klimarechenzentrum (DKRZ) in Hamburg statt.

Termin:
30. September - 1. Oktober 2003

(Achtung: Der Termin liegt abweichend von den üblichen Gepflogenheiten des Arbeitskreises Dienstag und Mittwoch!

Tagungsort:
Geomatikum, Bundesstraße 55
D-20146 Hamburg, Hörsaal H5

Geplante Themen

- High Throughput Computing
- Filesysteme
- Archivierung
- nicht numerische Anwendungen

Genauere Informationen und Anmeldung unter

http://www.dkrz.de/zki.html

Climate Data News

- **NOAA AVHRR Data:** In cooperation with the DFD/DLR the CERA climate database system offers access to monthly mean SST data. The AVHRR satellite data are integrated into the local data catalogue and can be accessed via the project „DLR EOWEB". The data download is transparently linked to the DFD in Oberpfaffenhofen. This is a first example of cooperation between the WDC for Climate in Hamburg and the WDC RSAT in Oberpfaffenhofen. Presently the SST's are restricted to the Mediterranean sea. North Atlantic and the Atlantic around Madeira are going to be implemented.

- **ERA 40:** The new generation of ECMWF's reanalysis data for the last 40 years will be finalised soon. We started to extract data from the MARS archive in Reading. The status of the retrieval process is documented at

http://mad.zmaw.de/e40/index.html.

Presently monthly mean surface data are available in Hamburg. Integration into the CERA database system is under development. We are discussing with the ECMWF how to organise the data transfer from Reading to Hamburg. The entire amount of ERA 40 data has been estimated to 15 – 30 Tbyte depending on horizontal and vertical resolution.

- **DWD Station data:** The German weather service (DWD) started to open the German station data archive. The DWD entry point of the community climate data archive

http://www.mad.zmaw.de/Klimadaten/Klimaarchiv/ausgangsseite.html

offers under „Klimadaten-Online" access to time series of 44 German observation stations. The time series data are provided in station data standard format and are disseminated as daily data in ASCII format. The structure of the standard format is explained in the web pages.

OUR SERVICES

• Computer Time For Earth System Modelling:

DKRZ

Top level computing ressources and archive capacity are available to all interested research groups working on climate and earth system modelling. DKRZs scientific steering committee (WLA) selects admissible projects.

Information: *http://www.dkrz.de/projects*
Email: *projects@dkrz.de*
Responsible: *Joachim Biercamp (EXT -314)*

• Support for users:

Support for porting models to DKRZ computers, support for model optimization, visualisation services and general user support to all users.

Information: *http://www.dkrz.de/support*
Email: *beratung@dkrz.de*
Responsible: *Help Desk (EXT -275)*

• Community climate models:

M&D Modelle & Daten

M&D provides numerical models and diagnostic software as well as user support for their application and carries out comprehensive climate model runs.

Information: *http://www.mad.zmaw.de/ClimateModels*
Email: *model@dkrz.de*
Responsible: *Christoph Heinze (EXT -376)*

• Data Support:

Provision, maintenance and easy access of climate research relevant data of different research groups as well as post-processing tools and the support of users of the climate data base.

Information: *http://www.mad.zmaw.de/ClimateData*
Email: *data@dkrz.de*
Responsible: *Michael Lautenschlager (EXT -297)*

DKRZ
Deutsches
Klimarechenzentrum GmbH
Bundesstraße 55
D-20146 Hamburg, Germany

Phone: +49 40 41173-334
Email: sekretariat@dkrz.de
Web: http://www.dkrz.de

M&D Modelle & Daten
Gruppe Modelle & Daten
Max-Planck-Institut für Meteorologie
Bundesstraße 55
D-20146 Hamburg, Germany

Phone: +49 40 41173-397
Email: mad_office@dkrz.de
Web: http://www.mad.zmaw.de

Editors
Dr. Joachim Biercamp
Prof. Dr. Guy Brasseur
Dr. Hans Luthardt

Layout
Norbert P. Noreiks
Joachim Biercamp