

Protokoll: Projekt CMIP6-DICAD – Halbjahrestreffen am 24.01.2017

12:00 Martin Schupfner (DKRZ) – Begrüßung

12:10 Martina Stockhause (DKRZ) – Langzeitarchivierung

Diskussion:

Es wurde die Frage gestellt, welches die Änderungen im Verfahren bei der Langzeitarchivierung im Vergleich zu CMIP5 sind (JJ).

Vorschlag SL: MSt denkt darüber nach, ob eine stichwortartige kurze Darstellung der Unterschiede im Projekt-Redmine abgelegt werden kann.

Es wurde um nähere Erläuterungen zu dem geplanten CMIP Data Pool gebeten (JJ). Die ETH Zürich hat für AR5 WGI ein Repository mit einem aktuellen CMIP5-Datensubset zur Verfügung gestellt. Die Zusammenstellung der Inhalte erfolgte nach Nutzerwünschen. Die wesentlichen Ursachen für den Aufbau dieses Repositories waren: ein vereinfachter skriptbasierter Zugang zu den Daten ohne OpenID sowie die Verteilung der aktuellen Datensatzversionen. Diese Funktionalität wird das IPCC DDC am DKRZ in Abstimmung mit allen IPCC AR6 WGs zur Verfügung stellen. Genauer zum CMIP Data Pool wird bis zum Vortrag von SK verhandelt.

Anmerkung MSt: Eine Darstellung des geplanten CMIP Data Pools ist im Redmine vorhanden unter https://redmine.dkrz.de/projects/dkrz_cdp und eine Darstellung der Data Citation und Langzeitarchivierung unter <http://cmip6cite.wdc-climate.de>.

12:30 Stephan Kindermann (DKRZ) – ESMValTool in der DKRZ-Infrastruktur

Diskussion:

Da GitHub ein US-amerikanisches Unternehmen ist, wurde die Frage nach der Rechtssicherheit bei der Verwendung von GitHub gestellt (BF).

SK: Es gab Diskussionen bzgl. des Urheberrechts der Software. Es gibt deshalb private Repositories für die Entwickler, aus denen Code bei Bedarf in die öffentlichen Repositories verschoben wird. Ziel war ein offenes Tool, das bereits vorhandene Tools auf GitHub nutzen kann, sodass diese nicht zusätzlich entwickelt werden müssen. Als Lizenz wählte man eine GPL-Lizenz mit Einschränkungen.

SL: Bezog sich die Frage auf die Verlässlichkeit der Pflege und Weiterentwicklung der GitHub-Software, oder ging es um Datenschutzkriterien? Oder um die Möglichkeit, dass man sich über GitHub einen Spion in Form eines Virus oder eines Wurms in Haus holt?

13:00 Lisa Bock (DLR) – Monitoring mit dem ESMVal-Tool

Diskussion:

Das AWI plant, das ESMValTool-Monitoring für die AWI-CM basierten CMIP6-Experimente zu nutzen (TS).

Das sollte problemlos möglich sein, da nur die Anpassungen des ESMValTools an die AWI-CM-Variablenamen erforderlich ist. Gespräche zwischen potentiellen Nutzern des ESMValTools werden zu einem anderen Zeitpunkt bilateral fortgeführt.

Die Frage, wie aufwändig (RZ und Memory) die Diagnostiken sind, und ob es negative Auswirkungen auf die Simulationsläufe geben könnte (ML), wurde dahingehend beantwortet, dass im Rahmen des Experimente-Monitorings nur die neu hinzugekommenen Daten-Chunks verarbeitet werden. Zudem kann die Verarbeitung parallel durchgeführt werden. Sehr zeitaufwändige Diagnostiken können in größeren Zeitabständen durchgeführt (nicht für jedes neue Daten-Chunk) werden. Das Monitoring-Tool wird zwar am DKRZ entwickelt, ist aber rechnerunabhängig, und kann ohne großen Aufwand auch auf den DWD-Rechnern oder anderen Plattformen benutzt werden.

Bei dem ESMValTool-Monitoring handelt es sich um eine Visualisierung der Ergebnisse eines laufenden Experiments. Es ist den Modellierern überlassen, wie sie reagieren, wenn das ESMValTool auf Fehlentwicklungen hinweist. Soweit es um Experimente geht, für die begrenzte oder auch gemeinsame RZ zur Verfügung steht, muss der RZ-Aspekt natürlich einbezogen werden. Auch die Frage, ob für das CMIP-Archiv aufbereitete oder Modellrohdaten als Input für das ESMValTool gewählt werden sollten (CS), bleibt den Modellierern überlassen. Beides ist möglich. Die erste Option erfordert weniger Anpassung. Bei der zweiten Option wird auch gleich die Formatierung überwacht.

13:15 Thomas Schartner (FUB) – Freva im CMIP6-Projekt

Diskussion:

Es wurde gefragt, warum das ESMVal-Tool kein Plugin für Freva ist (BF). Die Antwort ist, dass die Plugin-Sektion in der Projektversion des Portals entfernt wurde. Auf dem Projekt-Portal sollen nur die Ergebnisse der ESMValTool-Diagnostiken für die Community sichtbar gemacht werden. Sie können dann von der Community auf Freva bewertet und diskutiert werden.

14:00 Stephan Kindermann (DKRZ) – ESGF Datenpool

Diskussion:

Es wurde nach der Möglichkeit gefragt, auf dem geplanten DWD-Knoten auch Nicht-CMIP-Daten zu veröffentlichen (Seasonals, Decadals und Reanalysen) (BF). Das ist grundsätzlich möglich, vorausgesetzt, es wird ein Index-Knoten eingerichtet, der unabhängig von den anderen zum CMIP-ESGF gehörenden Index-Knoten betrieben wird.

Grundsätzlich können neue Knoten zur CMIP-ESGF zugeschaltet werden (Frage LB), allerdings muss das mit einem kontrollierten Verfahren geschehen, um negative Auswirkungen auf den Rest der Föderation zu unterbinden.

Es wurde auch nach dem Stand der Diskussion über Prioritäten und Inhalte gefragt, für die auf dem Kick-Off-Treffen eine Gruppe gebildet wurde. ML (Vorsitz der Gruppe) hat berichtet, dass die Gruppe sich konstituiert hat, und zurzeit diese Fragen per e-mail diskutiert. Insgesamt stehen 5 PB Speicherkapazität für den Datenpool zur Verfügung. Davon sind 2 PB für die Erstpublikation am ESGF-Datenknoten des DKRZs vorgesehen. Wie die inhaltliche Gestaltung des Datenpools darüber hinaus aussehen soll, wird noch diskutiert. Dazu sollen Download-Statistiken für den DKRZ-Datenknoten erstellt werden. Priorität werden auch die vom ESMValTool verwendeten Daten erhalten.

Die Frage nach den Unterschieden zwischen QA und ESMVal-Monitoring wurde bis zum Vortrag über das QA vertagt.

Vorschlag SL: Zusammensetzung und Ergebnisse der Datenpoolgruppe werden im Projekt-Redmine dokumentiert.

14:30 Vera Maurer (DWD) – Klimasimulationen mit ICON

Diskussion:

Frage des DWD: Absprache zwischen MPI und DWD nötig bzgl. ‚expar‘-Tool (Nesting). Karl-Hermann Wieners (MPI-M) gibt positive Signale.

14:50 Johann Jungclaus (MPI-M) – Informationen über MPI-ESM1.2

Diskussion:

Es wurden mehrere Fragen über die Nutzbarkeit des Kohlenstoffkreislaufs (SH) und der dynam. Vegetation (BF) mit der HR-Version gestellt. Beides ist in der LR-Version möglich. Mit der HR-Version ist eine Simulation des Kohlenstoffkreislaufs aber auf Grund der damit verbundenen langen Einschwingphasen problematisch. Die Prioritäten am MPI-M liegen zurzeit bei anderen Aspekten.

Ansprechpartner für organisatorische Fragen bzgl. MPI-ESM1.2-HR ist Matthias Bittner, modellspezifische Fragen sind an Karl-Hermann Wieners zu richten.

Der DWD plant, einige Szenarien auf eigenen Plattformen durchzuführen. Die Liste der geplanten Experimente aus dem Projektantrag ist noch aktuell. Auch die Aufteilung der RZ und Experimente zwischen BMBF-CMIP6 (DKRZ Account 988) und dem MiKlip-Projekt vom Oktober 2016 ist noch gültig. Die personelle Verantwortung für die Durchführung liegt für die DECK- und historical-Experimente beim MPI, die für die Durchführung der Szenarien beim DKRZ (e-mail Marco Giorgetta, 1. Aug. 2016).

Die Absprachen zwischen DWD und DKRZ wären also über die Durchführung der verschiedenen Realisationen zu treffen. Die am DKRZ beantragte RZ ist davon nicht betroffen. Ergebnisse der Absprachen (per e-mail) werden im Projekt-Redmine abgelegt.

15:40 Martin Schupfner (DKRZ) – Datenanforderung

Diskussion:

Es wurde die Existenz und Notwendigkeit von Mapping-Tabellen (Variablenzuordnung zwischen CMOR- und Modellvariablen) diskutiert. Einiges dazu ist schon vorhanden. Martin und Kalle arbeiten an einem Tool, das die Erstellung von Mapping-Tabellen unabhängig vom ESM unterstützt.

16:00 Martin Schupfner (DKRZ, i.V. Fabian Wachsmann) – cdo und CMIP

Diskussion:

Einen Releasetermin für ‚cdo cmor‘ gibt es noch nicht. Zunächst muss die Formatierung sämtlicher CMIP5-Variablen getestet sein. Dazu wird noch die CMOR2-Bibliothek benutzt. Anschließend wird der CMIP6-Standard verwendet, mit Übergang zu CMOR3, und Tests mit möglichst allen CMIP6-Variablen durchgeführt. Beta-Tester werden die Projektpartner mit eigenen Modellen sein. Es ist geplant, ‚cdo cmor‘ im Post-Processing-Workflow am DKRZ einzusetzen. Das CMIP-konforme Prozessieren mit cdo wird nach und nach erweitert.

16:20 Heinz-Dieter Hollweg (DKRZ) – QA

Diskussion:

Die Fragen bezogen sich in der Mehrheit auf die Unterschiede zwischen den diversen Tools zur Qualitätssicherung von CMIP6-Modellergebnissen. Grundsätzliche Unterschiede bestehen zwischen dem CMOR-Checker und dem am DKRZ entwickelten QA-Tool samt CF-Checker auf der einen Seite, und dem ESMValTool auf der anderen Seite.

Das ESMValTool als Monitoringtool ermöglicht den beteiligten Wissenschaftlern, die Qualität ihrer Modellergebnisse insbesondere mit Blick auf andere ESMs und Beobachtungsdaten schon während der Experimentdurchführung zu beurteilen. Auch als Diagnostiktool nach Beendigung eines Experiments dient es zur Beurteilung der physikalischen Qualität der Ergebnisse (z. B. Realitätsnähe).

Dagegen überprüfen die anderen Tools technische Anforderungen an die Ergebnisdaten (Metadaten, Genauigkeit, Einheiten, Vollständigkeit der Zeitachse, etc.). Der CMOR-Checker ist in HDHs QA-Tool integriert. Der CF-Checker ist ein Bestandteil des Tools. Mit dem QA-DKRZ Software-Paket werden viele zusätzliche technische Tests durchgeführt, deren Ergebnisse über die Aufnahme im ESGF entscheiden.

17:00 Martin Schupfner (DKRZ) – Organisatorisches

- Die Zwischenberichte für das BMBF sind von jedem Partner im Projekt selbst anzufertigen und einzuschicken. Die nächsten sind zum 1. April 2017 fällig.
- Das BMBF verlangt eine Internetpräsenz des Projekts.. Ein erster Entwurf ist unter <https://dkrz.de/c6de> zu besichtigen. Alle Projektpartner möchten erwägen, Inhalte dort einzustellen, unter Beachtung der unterschiedlichen Funktionen des Projekt-Redmines (projektintern) und der öffentlichen Seiten <https://dkrz.de/c6de>. Die Seiten der Internetpräsenz werden mit der Plone-Software erstellt. Da die Bedienung von Plone gewöhnungsbedürftig ist, können die Inhalte auch ans DKRZ geschickt werden.
- Nächstes Projekttreffen (Anlass: 12-Monattreffen) würde zeitlich in die Nähe des DKRZ Nutzerworkshops passen (Montag-Nachmittag 9. Oktober 2017 bis Dienstag-Vormittag 10. Oktober 2017).
Projekttreffen (vorgemerkt): 10. Oktober 2017 – Ein halber Tag Schulungen (Post-Processing Skripte und cdo cmor) im Rahmen des DKRZ-Nutzerworkshops.
11. Oktober 2017 Vormittags: Allgemeines Projekttreffen mit Diskussionsrunde.
Es muss noch besprochen werden, ob es sich um ein gemeinsames DICAD-CHEMIE-Treffen handeln soll.

17:10 Michael Lautenschlager (DKRZ) - Verabschiedung