

Unterstützung der CMIP6⁺-Aktivitäten in Deutschland

Mi. 18. April 2018, AWI Bremerhaven
10:00 – ca. 16:00

Teilnehmer: Lisa Bock (DLR), Bernadette Fritsch (AWI), Barbara Früh (DWD), Heinz-Dieter Hollweg (DKRZ), Markus Kunze (DLR), Kameswarrao Modali (MPI-M), Stephanie Legutke (DKRZ), Holger Pohlmann (MPI-M), Thomas Schartner (FUB), Martin Schupfner (DKRZ), Tido Semmler (AWI), Christian Steger (DWD), Martina Stockhause (DKRZ), Fabian Wachsmann (DKRZ), Kalle Wieners (MPI-M), ??

Tido Semmler (AWI) hat die Teilnehmer begrüßt und die Veranstaltung eröffnet.

Der erste Vortrag kam aus dem 2. Verbund 'Chemie', der durch Markus Kunze vertreten war.

Stand der EMAC-AERChemMIP Aktivitäten

Das Teilprojekt ? (TP?) liegt gut im Zeitplan. Die Meilensteine (MS) sind wie geplant erreicht. Der Spinup ohne Chemie (oC) ist abgeschlossen, und der piControl-Lauf (oC) wurde gestartet. Nur dieses Experiment muss sowohl mit als auch oC durchgeführt werden. Die Berücksichtigung der Chemie, insbesondere der Aerosole wurde erläutert. Stratosphärische Aerosole können mit der Modellversion nicht berücksichtigt werden. Die Simulationen werden mit dem EMAC v2.53 mit T42L47/GR15L40-Gittern auf nur 10 Knoten durchgeführt. Es kann also erwartet werden, dass mehrere Simulationen gleichzeitig durchgeführt werden können. Die Skripte zu Postprozessing und ‚CMORisierung‘ werden gemeinsam mit dem DICAD-TP1 erzeugt.

Diskussion: Die externen Inputdaten aus dem obs4MIPs-Projekt werden vom MPI-M auf die Gitter T63 und T127 interpoliert. Für die Interpolation auf T42 hat das MPI-M Unterstützung angeboten.

Es wurde in diesem Zusammenhang berichtet, dass die WG-I ‚literature dead line‘ auf den Sommer 2020 verschoben wurde.

Für das Postprozessieren der Daten wurde angemerkt, dass die Diagnostiken, die für den ‚expr‘ Operator in der Mapping Table formuliert sind, sehr lang sind und es überprüft werden muss, ob die CDOs diese überhaupt verarbeiten können. Das DKRZ wird Tests durchführen, nachdem der Pfad zu den Daten bereitgestellt wird und die Mapping-Tabelle zur Verfügung steht.

Die nachfolgenden Vorträge kamen aus dem DICAD-Verbund. Über den Stand der CMIP6⁺-Experimente mit MPI-ESM1.2 (WP 1.1-2 und WP 3.1) wurde von Stephanie Legutke (DKRZ) berichtet.

Stand der DECK-, historical-, amip- und ScenarioMIP-Experimente mit MPI-ESM1.2 und Schnittstelle: Integration <-> Datenaufbereitung

Zunächst wurde die Historie der Freigabe des CMIP6-Datenrequests (DReq), der externen Inputdaten, sowie der Freigabe des Modellquellcodes zusammengefasst. Schon vor Beginn des Vorhaben war beschlossen worden, dass das BMBF-geförderte Vorhaben ‚MiKlip‘ dieselbe MPI-ESM1.2-Version benutzen würde, die auch für die CMIP6-Experimente in dieser Auslösung (‚HR‘: T127L95/TP04L40)

verwendet wird (<https://code.mpimet.mpg.de/projects/mpi-esm-users/files> als 'mpiesm-1.2.01.tar.gz' unter 'mpiesm-1.2.01'; Anmeldung über <https://www.mpimet.mpg.de/en/science/models/mpi-esm/users-forum/>; wenn ein SVN-Account mit den entsprechenden Genehmigungen vorliegt unter <https://svn.zmaw.de/svn/cosmos/tags/mpiesm-1.2.01>).

Da MiKlip unter hohem Zeitdruck stand, sollten 500 Jahre des PiControl-Experiments, die DECK-Experimente 1pctCO2 und abrupt4xCO2, sowie 5 Realisationen des geplanten historical-Experiments in MiKlip durchgeführt werden. Diese Simulationen sind auch berechnet, allerdings ohne Rücksicht auf den CMIP6-Datenrequest (DReq), von dem zu der Zeit noch keine Version freigegeben war. Die Antriebsdaten für historical waren freigegeben, allerdings wurden die stratosphärischen Aerosole, wie schon auf dem letzten Projekttreffen berichtet, später durch eine neue Version ersetzt. Das soll für die noch zu berechnenden Realisationen des historical-Experiments auch so beibehalten werden. Grundsätzlich muss das CMIP6-Projekt Daten akzeptieren, die mit freigegeben Inputdaten oder DReq hergestellt wurden.

Es wurde erläutert, wie die projektkonforme Aufbereitung der Ausgabedaten jetzt unter Zuhilfenahme des gesamten DReqs durchgeführt werden kann. Die dazu erforderlichen Eingaben der MPI-M-Mitarbeiter in das in TP1 WP3 entwickelte Web-GUI wurden vor wenigen Wochen zur Verfügung gestellt, so dass die Aufbereitung jetzt beginnen kann.

Die noch ausstehenden Experimente werden gerechnet, sobald die auf dem gesamten DReq basierenden Namelists erzeugt sind.

Eine Wiederholung der schon berechneten Experimente und die dann noch benötigte Datenaufbereitung werden durch das MPI-M durchgeführt.

Diskussion: Die oben erwähnte ECHAM6-Version wurde versioniert, NACHDEM die Rechnungen sich als plausibel erwiesen hatten. Für das MPIESM-LR wird eine neuere Version von 2018 benutzt (ein BUG in MPIOM entfernt). Andere Modellvarianten laufen in anderen Konfigurationen (z. B. mit dynamischer Vegetation, oder mit dem zweiten Release der stratosphärischen Aerosole).

Christian Steger (DWD) hat i. V. von Vera Maurer (DWD), die verhindert war, über die im TP3 berichtet durchgeführten Arbeiten berichtet.

AMIP-Simulationen mit lokaler Gitterverfeinerung im CORDEX-Gebiet

Die Meilensteine AP1 der WP1.3 sind wie geplant erreicht. Das ICON-Modell ist im Klimamodus lauffähig. Ineinander geschachtelte Nester können mit jeweils einem Faktor 2 Verfeinerung sowohl in einer 1-Wege- als auch mit 2-Wege-Kopplung durchgeführt werden. Testläufe zeigen, dass im Vergleich zu globalen Simulationen in der Zielauflösung Rechenzeit (RZ) gespart wird. Für eine gute Performance sollte die Anzahl der Gitterpunkte der verschiedenen Nesting-domains ungefähr gleich sein. Im Fokus neuerer Tests steht die Frage, ob es performanter ist, mit einer feineren Anfangsauflösung Simulationen zu beginnen oder stattdessen öfter zu 'nesten'. Das Nesting ist gebunden an eine Halbierung der horizontalen Gitterweite (vertikales Skalieren funktioniert noch nicht). Mit einem Film wurde gezeigt, dass an den Rändern der Nester keine störenden Muster entstehen, wie sie in den CORDEX-Ergebnissen teilweise zu sehen sind. Dies ist einer der großen Vorteile der 2-Wege-Kopplung.

Diskussion: Die im Vorhaben geplanten Simulationen mit ICON-EU werden ausschließlich auf den Rechnern des DWD durchgeführt. Die Arbeitsteilung in Bezug auf die Durchführung der Szenarienprojektionen soll beibehalten werden wie abgesprochen.

Tido Semmler (AWI) hat stellvertretend für Dmitry Sein, der verhindert war, berichtet über

Stand der AWI-CM Rechnungen

Das AWI-CM-Modell ist in 3 Auflösungen (LR, MR, HR) mit entsprechenden Auflösungen im Ozean und ECHAM6.3 auf den T63L47 und T127/L95 aufgesetzt und getestet. Die CMIP6-Variante hat einen Durchsatz von 4 Jahren pro Tag. Die HR-Variante soll auch am CMIP6-endorsed HighResMIP teilnehmen.

Die Datenaufbereitung für das ESGF wird für die FESOM-Daten mit selbstentwickelten Skripten ohne Verwendung der CMOR3-Bibliothek durchgeführt. Für ECHAM6, das in derselben Version im MPI-ESM1.2 verwendet wird, werden die in TP1 WP2-3 entwickelten Werkzeuge benutzt. Der Inhalt der Mapping- und Recipe-Datenbank für ECHAM6 wird allerdings zur Verfügung gestellt.

Diskussion: Es wurde diskutiert, ob Daten von mehreren Simulationen Eingang ins ESGF finden können, obwohl von Projektseite nur 1 Realisation verlangt wurde. Dem stehen eigentlich keine Aussagen des CMIP6-Panel oder -WIPs entgegen. Allerdings kann die Aufnahmefähigkeit des ESGFs sowohl in Bezug auf Plattenplatz als auch auf Personalkapazitäten beschränkt sein. Dann werden erfahrungsgemäß (CMIP5) die angeforderten Daten vorrangig behandelt. Es wurde geraten, sich an das ‚review board‘ zu wenden, das über diese Fragen beschließt. Auf dem Projekt-Redmine findet man (<https://redmine.dkrz.de/projects/cmip6-dicad-subproject/wiki/AP5>) Informationen über den Aufgabenbereich etc. der Gruppe. Das DKRZ hat noch mal darauf hingewiesen, dass eine Datenaufbereitung ohne Zuhilfenahme der CMOR3-Bibliothek ein großes Risiko birgt, dass die Daten nicht im ESGF akzeptiert werden können. Für das Konvertieren von Atmosphären-Daten schlägt das DKRZ vor, dass das AWI einmal den DKRZ-Workflow – ohne arbeitsintensive Schritte – durchläuft. Das AWI ist einverstanden, merkt jedoch an, dass die CDOs mit unstrukturierten Gittern deutlich langsamer arbeiten. Kalle bestätigt das, weist aber darauf hin, dass Uwe Schulzweida bereits an Verbesserungen arbeitet. Es wird nach einem Tool gefragt, das nur die Daten konvertiert, bei denen sich Änderungen im DReq ergeben haben. Das WebGUI bietet die Möglichkeit, sich die Variablen/MIP-Paare anzeigen zu lassen, die sich zwischen 2 bestimmten DReq-Versionen geändert haben.

Über die Entwicklung und Anwendung des CDO-cmor-Operators (AP3) hat Fabian Wachsmann (DKRZ) gesprochen, unter dem Titel

Stand des CDO CMOR Operators beispielhaft vorgeführt

Der entwickelte Operator kann wahlweise mit der CMOR2- oder CMOR3-Bibliothek gekoppelt werden. Damit können der CMIP5- oder CMIP6-Standard hergestellt werden, ohne dass am Aufruf des Operators Änderungen nötig sind. Die Konfigurationsdateien müssen natürlich dem jeweiligen Standard angepasst sein. Für die Erstellung des CMIP6-Standards kann ein im AP2 entwickeltes WEB-Interface benutzt werden, das auf den Daten mit dem Projekt-ControlledVocabulary (CV) basiert, und damit hilft, an dieser Stelle Fehler zu vermeiden. Für CMIP5 gibt es solche Dateien nicht.

Der Operator ist am DKRZ auf der Mistral verfügbar (in einem speziellen CDO-Toolkit). Demnächst soll die Installation mit Conda verfügbar sein, was die lokale Installation wesentlich erleichtert.

Die Anpassung des CDO-Metadatenmodells wurde weitgehend durchgeführt. Insbesondere die Namen von Koordinaten werden nicht mehr geändert, und es gehen auch keine globalen Attribute mehr verloren. Allerdings ist fraglich, ob es erreichbar ist, Operationen auf den Daten durchzuführen, da damit einige Attribute zu ändern wären (Frequenz, MIP-Tabellen, etc.). Bei Bedarf werden weitere Angleichungen, insbesondere im Datenmodell, vorgenommen.

Diskussion: Es wurde gefragt, ob es nicht möglich sei, den cmor-Operator in das Standard-CDO-Operatorenpaket einzufügen. Dem spricht entgegen, dass der cmor-Operator ein umfangreiches eigenes Paket von Software-Bibliotheken beinhaltet (UDINITS, UUID,...), so dass die Kompilation immer mal wieder auf Schwierigkeiten stößt. Daneben benötigt der cmor-Operator, im Gegensatz zu den anderen Operatoren, eine Reihe von Konfigurations-Informationen, weshalb er per se mehr Nutzer-Unterstützung erfordert. Diese wird aber im Rahmen des MPI-M-CDO-Redmine-Supports von Fabian gegeben. Es besteht weiterhin Bedarf an Schulungen. Die nächste ist für den 23./24. Mai am GERICS im Rahmen von PRINCIPLES für CORDEX-Aktivitäten geplant. Restplätze können von Mitarbeitern im Vorhaben eingenommen werden. Darüber wird nach Anmeldeschluss (der noch nicht feststeht) auf dem Redmine des Vorhabens informiert.

Über das AP3 hat Martin Schupfner (DKRZ) berichtet mit dem Titel

Web-unterstützte Anpassung von Modelldaten an den CMIP6-Standard

Martin stellt die bekannten und die kürzlich hinzugefügten Funktionen des WebGUI vor (<https://c6dreq.dkrz.de>). Das WebGUI führt den Nutzer mittels mehrerer Web-Applikationen durch einen Arbeitsablauf zum Erzeugen der CMIP-konformen Daten. So können die Datenanforderung abgefragt, das zugehörige Datenvolumen berechnet und für jedes registrierte Modell Variablenzuordnungen zwischen CMIP6- und Modellvariablen angelegt werden. Darauf basierend können Skriptfragmente zur CMIP-Standard-Konvertierung und Modelldiagnostik sowie auch Konfigurationsdateien zur Nutzung des cdo-cmor-Operators automatisch erzeugt und heruntergeladen werden. Diese neue Funktion der Online-Skripterzeugung im WebGUI ist ab sofort in einer ersten Version verfügbar.

Diskussion: Bereits bestehende Variablenzuordnungen (etwa aus anderen Projekten) können nach Rückfrage in die Datenbank eingelesen werden. Anschließend ist über das WebGUI die weitere Bearbeitung und Ergänzung der Variablenzuordnung und die Online-Skripterzeugung möglich.

Heinz-Dieter Hollweg (DKRZ) berichtete über den Stand der Entwicklung der Qualitätsprüfung (AP4) von Projektdaten am DKRZ und die zurzeit geltenden Regeln.

Datenqualitätsprüfung

Die QA ist Teil der als DataChallenge (DC) bezeichneten Testreihe der ESGF-Entwickler und Anwender, die gegenwärtig stattfindet und mehrere Phasen umfasst.

Die Beschlusslage ist zurzeit, dass sich die Eingangskontrolle zum ESGF nur auf die Einhaltung des NetCDF/CF-Standards und die DRS-Struktur (Dateiname und Pfad, sowie NetCDF-Attribute und Variablennamen) bezieht. Diese werden mit dem CF-Checker und dem am PCMDI entwickelten PREPARE durchgeführt. Das am DKRZ entwickelte Tool untersucht darüber hinaus auch, ob die eigentlichen Daten der zu publizierenden Dateien verdächtige Werte enthalten. Dazu gehören beispielsweise _FillValue als Koordinatenwerte, NaN als Koordinaten- oder Zeitwerte, konstante Felder oder sich wiederholende Muster in den Feldwerten. Dieses QA-Tool wird in CMIP6 in erster Linie vor der LangzeitArchivierung (LZA) und der Vergabe der DOIs nach dem DataCite-Verfahren verwendet. Die Annotationen schon im ESGF publizierter Daten werden dann allgemein zugänglich sein.

Für die DC-Tests des Tools wurden von AP2-3 Testdaten zur Verfügung gestellt,

die mit den dort entwickelten Werkzeugen aufbereitet wurden. Das Protokoll der QA dieser Daten wurde beispielhaft gezeigt und erläutert.

Diskussion: Es wurde diskutiert, warum die in dem QA-Tool enthaltenen Checks der Variablenwerte keine Plausibilitätsprüfung der Werte enthalten. Diese Checks sind extrem schwierig, insbesondere, da sie auch auf Ergebnisse von Projektionen für einige 100 Jahre angewendet werden würden, über die der Natur nach nicht viel bekannt ist. Außerdem haben im CMIP5-Projekt die von der CMOR2-Bibliothek herbeigeführten Abbrüche für Unmut unter den Anwendern geführt. Es gibt aber im DReq für viele Variablen Schranken, deren Über- oder Unterschreitung protokolliert wird. Der eigene ESGF-Knoten des DWDs ist seit 16.4. in Betrieb. Zurzeit wird die am DKRZ entwickelte QA-Software installiert, damit die am DWD erzeugten Daten vor Ort qualitätsgesichert werden können.

Danach wurde von HDH noch i. V. von Stephan Kindermann (DKRZ) über die Arbeit in AP 5 berichtet.

DKRZ CMIP Datenpool und ESGF Knoten: Stand

Die Infrastruktur bezüglich des ESGF-Datenknotens am DKRZ, des MIP-Datenpools und des DKRZ-Langzeitarchivs sowie die Schnittstellen zum Mistral-Rechner wurden erläutert. Es wurde betont, dass allgemeine Anfragen, Wünsche etc. rund um den Datenpool an data-pool@dkrz.de zu richten sind. Für die Ablieferung von CMIP6-Daten zur Speicherung, ESGF Publikation und Archivierung ist esgf-publication@dkrz.de vorgesehen. Eine Liste der am häufigsten benutzten CMIP5-Variablen sowie eine ESMValTool-Variablenliste liegt vor. Die sich daraus ergebende aktuelle Volumeneinschätzung bleibt unverändert und dürfte ca. 2 Pbyte umfassen. Abschließend wurde berichtet, dass in der ersten Hälfte von 2018 die sogenannte „Data Challenge“ für die Erprobung der ESGF-Software und der Replikation stattfindet; in Phasen unterteilte Zeiträume mit sukzessiv anwachsenden Datenvolumina, die an den Tier1- und einem Tier2-Knoten installiert, publiziert und repliziert werden. Im Juni wird die Publikation von CMIP6-Daten beginnen.

Lisa Bock (DLR) hat, auch Björn Brötz (DLR) vertretend, über die Arbeit in WP6.1-3 berichtet

Supporting CMIP6 Simulations with the ESMValTool

Das ESMValTool kann benutzt werden, um ein Monitoring der Ergebnisse schon während der Modellläufe durchzuführen oder um nach Beendigung einer Simulation Modellergebnisse mit denen anderer CMIP6-Modelle zu vergleichen. Das Monitoring wird zurzeit in die Modellierumgebung von EMAC am DKRZ eingebaut. Bei diesem Modell ist es möglich auf den Rohdaten aufzusetzen, da deren Format in der Arbeitsgruppe bekannt ist. Normalerweise erfordert dies einige Anpassungen des ESMValTools, während die Verwendung von CMIP6-Protokolldaten ohne weitere Maßnahmen möglich ist.

Die im ESMValTool eingebauten Diagnostiken sind in Namelists gruppiert und können mit Namen angefordert werden, die auf vorherige Verwendung der Diagnostiken, meist in peer-reviewed Veröffentlichungen, verweisen (IPCC/AR5 Kap. 9, etc.). Es wurden fast alle vorhandenen Namelists mit den CMIP5 Daten durchgeführt, die Ergebnisse sind auf <http://cmip-esmvaltool.dkrz.de> dargestellt. Der Umfang des Tools wird kontinuierlich erweitert.

Das ESMValTool ist auch auf dem ESGF-Datenknoten am DKRZ installiert, wo ein Synda-Dämon observiert, ob neue Daten eingegangen sind, die von den aktivierten

ESMValTool-Diagnostiken verwendet werden. Sobald dies geschieht, werden die entsprechenden Kommandodateien durchgeführt. Die Ergebnisse sind dann auf dem Freva-Portal zugänglich,

über das im Folgenden Thomas Schartner (FUB) berichtet hat.

Präsentation des CMIP-Data und ESMValTool Result Portals

Die aktuelle Internetpräsenz von Freva wird vorgestellt. Jeder auf Mistral registrierte User kann sich einloggen. Freva ist auch als Modul ladbar und verfügt über Autovervollständigung.

Der Result-Browser und verfügbare Search-Facets wurden vorgestellt.

Aktuell sind die Graphiken mit Wasserzeichen versehen, um Missbrauch vorzubeugen. Der Zugang zu Graphiken ist offen. Um die Kommentarfunktion zu nutzen, ist ein Login erforderlich.

Zum Schluss hat Martina Stockhause (DKRZ) zusammenfassend über die AP 7-8 geredet.

Bericht über AP7 und AP8

Auf Grund von krankheits- und schwangerschaftsbedingtem Ausfall der Mitarbeiterin, die in diesen APs beschäftigt war, haben sich die Aktivitäten auf AP7 beschränkt. Die Frühziterung ist jetzt vollständig in die ESGF-Publikation integriert. Deswegen wurden jetzt ein Entwickler-Bereich und ein Bereich für die DC-Tests eingerichtet. Die Anwenderoberfläche wurde insofern verbessert, dass jetzt von der source_id abhängige Angaben automatisch gesetzt werden. Es wurde auch beschlossen, dass ‚vorläufige‘ DOIs vergeben werden können, sobald Daten zur ESGF-Publikation eingereicht sind. Allerdings müssen vorher die nötigen DOI-Informationen in das WEB-GUI eingetragen sein.

Sonstiges

- Zwischenbericht?
Der Projektträger (PT) fragt nach einem unter den Teilprojekten abgestimmten neuen Arbeitsplan, wenn es gravierende Änderungen geben sollte. Wg. der Verzögerungen des Starts insbesondere der Szenarienprojektionen wurde von der DICAD-Koordination eine Anfrage an die Koordinatoren der Teilprojekte gestellt. Da es sich nur um eine Verschiebung der Experimente handelt, wurde die Notwendigkeit eines überarbeiteten Arbeitsplans verneint, bzw. gab es keine positive Antwort auf die Anfrage. TP2 (DLR) hat allerdings beim PT um eine Verschiebung des Fälligkeitstermins für 2 Meilensteine gebeten (nicht abgestimmt).
- RZ-Nachschatz?
Der am 30. April fällige RZ-Antrag für Jul-Dez 2018 wurde kurz besprochen. Von DWD-Seite besteht kein Interesse mehr, sich an diesem Antrag zu beteiligen. Auch soll keine RZ für neue Experimente beantragt werden, sondern nur für die Experimente, die für 2018 geplant waren, aber noch nicht durchgeführt wurden.
- nächstes Treffen?
Es wurde allgemein kritisiert, dass der Termin für dieses Treffen so kurzfristig festgesetzt wurde. Deshalb soll nach Verabschiedung des Protokolls innerhalb einer Woche über den Ort der nächsten Veranstaltung entschieden werden, und ein Doodle aufgesetzt werden. Der Doodle wird die Zeitspanne 1 Juli 2018 + 2 Monate +/- 2 Wochen abdecken.

15:55 Abschließendes Resümee und Verabschiedung

Trotz des verspäteten Beginns des Treffens, konnte es wie vorgesehen um 16:00 beendet werden. Tido Semmler dankt allen Beteiligten für ihre Zeit und die Beiträge....