

Protokoll DICAD 30-Monatstreffen

1. Stephan Kindermann, DKRZ:

Der Datenbereitstellungsworkflow am DKRZ

Katharina ist für das ESGF Portal und die Publikation zuständig, HDH für die Qualitätskontrolle und Replikation. Ein Korrigendum für im ESGF publizierte Files sollte über den CMIP6 Errata Dienst dokumentiert werden. Dies bietet Datennutzern größtmögliche Transparenz.

Der Speicherplatz des ESGF CMIP6-Datapools wurde von 5Pb auf 2Pb verringert. Ein Steuergremium aus Michael L. und Vertreter der Partner entscheidet, welche Daten in den Datenpool kommen.

Der CMIP6 Datenpool ist auf der Mistral unter /work/ik1017, alle Mistral-Nutzer können darauf zugreifen.

Diskussion:

Stephan: Der Datenpool ist wahrscheinlich verkleinert worden, weil zum Zeitpunkt der Projektbegutachtung noch keine CMIP6 Daten vorlagen und damit der Speicherbedarf in 2019 niedriger eingeschätzt wurde. Eine Erhöhung des Kontingentes in der Projektbeantragung für 2020 auf die ursprünglich vorgesehenen 5PB wird vorgesehen.

Christian: Können Daten des DWD ESGF-Knotens im Datenpool repliziert werden, wenn lokaler Zugriff möglich sein soll?

2. Heinz-Dieter Hollweg, DKRZ:

Replikation und QA:

Es gibt eine Webseite mit Infos welche CMIP6-Daten bereits publiziert wurden (siehe Folien).

Diskussion:

Christian: Kann es für CORDEX ein ERRATA geben? Hintergrund ist, dass im Datenpool fehlerhafte Daten vorkommen können, weil andere Institute eine andere (oder keine; es ist nicht transparent welcher Datenknoten was macht) Qualitätssicherung durchführen.

Stephan: Das gegenwärtige Errata System ist für CMIP6 entwickelt worden und nutzt die CMIP6 PID Infrastruktur, für eine Anpassung auf existierende CORDEX Daten ist aufwändig - jedoch kann für (zukünftig) im Rahmen von CMIP6 publizierte CORDEX Daten der CMIP6 Errata genutzt werden.

Martina: Welche Daten sind bis jetzt repliziert?

HDH: Die, die leicht über synda zu replizieren sind.

Stephanie: Gibt es dabei eine Priorisierung?

HDH: Nach Möglichkeit parallel. ESMVal fordert bestimmte Daten an, die bevorzugt werden.

Fabian: Vorschlag: Der Datenbereiter soll sich bei Publikationsanfrage entscheiden können zwischen Check nach Projektanforderung und zeitintensiveren, dafür aber vollständigem QA-Check mit Test auf Ausreißer. Dies könnte dokumentiert werden.

3. Johann Jungclaus, MPI-M:

Status der Rechnungen am MPI-M

Testsimulationen mit dem ICON-ESM Modell nach Ruby-O Konfiguration sind angelaufen. Die horizontale Auflösung entspricht in der Atmosphäre 160km, 40km im Ozean. Historical Simulationen mit dem MPI-ESM1-2-HR sind abgeschlossen, aber das Jahr 2014 muss nochmal überprüft werden. Amip Simulationen sind fast abgeschlossen - es bedarf noch einer Überprüfung.

Diskussion:

Martin: Die Szenarien branchen nach 2014. Die Überprüfung muss abgeschlossen sein um mit den Szenarien zu beginnen.

Christian: Es gab die Anfrage nach mehreren Realisationen für die Szenarien. Wie akut ist diese?

Johann: Der Wunsch ist da. Für die Umsetzung fehlen Ressourcen.

Stephanie/Johann: Zuerst sollen die Aufgaben erledigt werden, die im Projekt versprochen wurden. Danach kann man sich Gedanken über zusätzliche Realisationen machen.

Vereinbarung: MPI/DKRZ kontaktieren DWD wenn dieses Thema wieder aktuell wird und MPI/DKRZ selbst besser abschätzen können welchen Beitrag sie bzgl. zusätzlicher Realisationen leisten könnten.

[Anm: Werden das bei Bedarf und nach Möglichkeit in den nächsten RZ-Antrag aufnehmen.]

Johann: Technischer Zugriff auf ICON Daten sollte möglich sein, um das Post-Processing zu testen. ICON wird nur ein limitiertes Set an Variablen für CMIP6 veröffentlichen.

Stephanie: Neustart AMIP nötig wegen Energieausgleich?

Johann: Sollte nicht nötig sein.

4. Vera Maurer, DWD:

Simulationen mit ICON-EUclim

Testdaten von Icon-A (ICON mit ECHAM-Physik) liegen auf der Mistral und können für Tests des Post-Processing verwendet werden. Die Meilensteine des Arbeitspaketes sind weitestgehend erfüllt. Der akkumulierte/gemittelte Output muss noch ermöglicht werden.

Diskussion:

Stephanie: Haben regionale CORDEX-Simulationen im Vergleich zu genesteten ICON-EUclim Simulationen noch einen signifikanten Benefit?

Christian: CORDEX behandelt ein andere Fragestellung: Bspw. werden mehrere globale Modelle als Forcing genutzt.

HDH: Führt die hohe Auflösung zu mehr konvektiver und weniger stratiformer Bewölkung?

Vera: Noch keine Zeit, genauer zu untersuchen.

5. Björn Brötz, DLR:

Stand der Entwicklung des ESMValTools

Das ESM-VAL Repository liegt auf Github und ist opensource. Es findet eine Weiterentwicklung in Richtung V2.0 statt. W3c prov standard? Die Dokumentation ist nun nicht mehr manuell sondern sphinx-basiert.

Es wird möglich sein, Quicklooks aus dem Jobsript zu erzeugen. Bis jetzt ist nur der passive Modus möglich, d.h. Quicklooks werden nach der Modellsimulationen erzeugt und nicht online. So wird es zur Zeit für EMAC verwendet.

Diskussion:

Martina: Werden die Diagnostiken des ESMVal so abgespeichert, dass sie leicht langzeitarchiviert werden können?

Björn: Dafür gibt es ein Portal.

Martina: Kann ausgehend von den Plots auf die zugrundeliegenden CMIP6-Daten zurückgeschlossen werden in DRS-Terminologie?

Björn: Sie werden mit dem Plot abgespeichert.

Diese Diskussion wurde bilateral in der anschließenden Kaffeepause fortgeführt.

6. Bianca, FUB:

Das Freva-Portal

Eine Data-Policy muss für CMIP6 noch entwickelt werden. Die angeforderten Meilensteine sind ansonsten erfüllt und das Freva-Portal ist einsatzbereit.

Diskussion:

Björn: Freva-Portal kommuniziert die Ergebnisse des ESMValTools. Für individuelle Diagnostiken und Analysen muss ESMValTool separat angewendet werden.

Björn: Es muss auf cmorisierten Daten aufgesetzt werden. Ein eigenes Interface je Modell ist zu aufwändig und Fehleranfälligkeit zu hoch.

7. Fabian Wachsmann, DKRZ:

CDO und der CMIP Standard

In CMIP6 konformen Daten werden die Versionen des Daten-Requests und von einem Dokument der Definition von globalen Attributen mit globalen Attributen referenziert. Diese beiden Dokumente können jedoch nicht einzeln betrachtet werden und Datenerzeuger müssen sich an MIP-Tabellen halten, da diese bei der QA als Input für PrePare dienen. Dies sollte in möglichen Nachfolgeprojekten beachtet werden.

Stdout und Stderr werden weiter überarbeitet, um das Post-processing über das Jobskript auch bei großer Anzahl von CMOR-Variablen einfach steuern zu können.

Die Operatoren zur Berechnung von Klimaextremindices werden überarbeitet. Bitte mit der Anwendung warten bis die Überarbeitung abgeschlossen ist.

Diskussion:

Christian: Kann ich cdo cmor 1.9.6 für die CMIP6 konforme Standardisierung hernehmen?

Fabi: Die Anwendung ist abhängig vom Format der Eingangsdaten. Eingangsdaten des MPI-ESM1-2-HR sind umfangreich getestet worden und können für fast alle CMOR-Variablen standardisiert werden. ICON Daten werde ich testen.

Fabi: Es fehlt eine Doku über Erstellung der MIP-Tabellen. Diese ist jedoch in Arbeit (Karl Taylor).

Stephanie: Support über Forum nötig (cdo cmor und WebGUI zusammen).

8. Martin Schupfner, DKRZ:

WebGUI und PostProcessing

WebGUI verfügbar um Variablen-Mapping und Skriptfragmente für das PostProcessing zu erstellen.

Diskussion:

Björn: Ist es möglich den Inhalt der Datenbank mit den Mappinginformationen zu bekommen?

Martin: Der Datenbankinhalt kann jederzeit in json- und Text-Format heruntergeladen werden.

9. Stephanie Legutke, DKRZ:

Nicht erreichte Meilensteine

Probleme bei der Absprache und bei der Organisation des Modelloutputs verzögern CMORisierung der Daten erheblich. Frist zum Veröffentlichen der Daten (Papers accepted Ende 2019) wird wohl verfehlt.

Abschließende Diskussion:

Johann: Michael Botzet muss sich einarbeiten und soll bei Bedarf auch Kalle Wieners vertreten.

Johann: Frist (Papers accepted Ende 2019) nicht unbedingt für verwendete Daten. Diese können auch noch später in Analysen/Veröffentlichungen und IPCC-Sachstandsbericht mit aufgenommen werden.

HDH/Stephan/Martina: DKRZ stellt den Data Pool den IPCC-Autoren zur Verfügung. Dies ist Teil der DDC (Data Distribution Centre) Support Aktivität, in der sich das DKRZ zusammen mit dem DDC-Partner CEDA (UK) engagiert. Besonders interessiert an diesem Angebot ist IPCC-WGI.

Martina: Es ist für die CMIP6-Infrastruktur (inkl. dem Data Citation Service) aufwändig wenn die DRS_id-Bestandteile wie z.B. Modellakronyme (source_ids) geändert werden.

Stephanie: Zwischenbericht je Partner bis 30.04.2019 abzugeben.

Baldiges Projekttreffen Ende Juni erwünscht. Online-Abstimmung folgt.