

Qualitätskontrolle der IPCC-AR5 Daten

M. Stockhouse, M. Lautenschlager, H. Höck, H.-D. Hollweg, and F. Toussaint
Deutsches Klimarechenzentrum (DKRZ), WDC Climate, Hamburg, Germany

Die Erarbeitung des nächsten IPCC Zustandsberichts (IPCC-AR5) erfolgt auf Basis des Modellvergleichs CMIP5 (Climate Model Intercomparison Project). In CMIP5 sind langfristige und dekadische Simulationen für Erdsystemmodelle geplant, zu denen historische Läufe (mittleres Holozän bis letztes Millennium), vier Emissionsszenarien (hoch - RCP8.5, mittel - RCP6, niedrig - RCP4.5, peak im Jahr 2100 überschritten - RCP2.6) sowie die Untersuchung verschiedener Einflüsse auf das Erdsystem (Vulkanausbrüche, Aerosole, Wolken, komplexe Chemie) gehören. Von den entstehenden Daten werden ca. 10% als von zentralem Interesse eingestuft. Diese bilden die sogenannten Core-Daten, die im IPCC-AR5-Kontext bereitgestellt werden. Sie umfassen ca. 1,5 Millionen Parameter-Zeitserien und ein Datenvolumen von ca. 1,5 - 2 PB.

Da diese Core-Daten von wissenschaftlichen Interesse sind, wird für sie eine Qualitätskontrolle durchgeführt, die aus drei Stufen besteht:

- Stufe 1: Dateneingangskontrolle mit Überprüfung der Format- und Strukturvorgaben
- Stufe 2: Plausibilitätskontrollen der Dateninhalte sowie subjektive Daten- und Metadatenkontrollen
- Stufe 3: Prozess der Datenpublikation mittels STD-DOI (<http://www.std-doi.de>) über die TIB (Technische Informationsbibliothek) Hannover in Abstimmung mit den Datenautoren.

Zur Datenbereitstellung im IPCC-AR5-Kontext haben PCMDI, BADC und WDC Climate die ESGF (Earth System Grid Federation) gebildet, die eine verteilte Datenhaltung ermöglicht. Die Core-Daten werden zwischen den drei Partnern repliziert zum Schutz vor Datenverlust und für einen schnelleren/dezentralen Datenzugriff. Für die Durchführung und Organisation der Qualitätskontrollen bedeutet das, dass die Stufen 1 und 2 an verschiedenen Datenzentren durchgeführt und die Ergebnisse in einem Metadaten-Repository gesammelt werden. Von dort werden sie für die abschließende Datenpublikation (Stufe 3) abgerufen, die ausschließlich WDC Climate/DKRZ durchgeführt wird.

Die Konzeption der Qualitätskontrollen wurde vom WDC Climate entwickelt als Beitrag zur Förderation. Sie basiert auf eigenen Erfahrungen der Datenpublikation wie z.B. der Bereitstellung der auf IPCC-AR4 basierenden CCLM-Ergebnisse für KlimaZwei und KLIMZUG. Dieser Prozess der geographisch-verteilten Qualitätskontrolle für die IPCC-AR5 Core-Daten in den verschiedenen Stufen wird vorgestellt, inkl. der Inhalte der Qualitätskontrollen, der Organisation/Synchronisation des Ablaufs sowie der Zugriffsmöglichkeiten/-einschränkungen für Daten der verschiedenen Qualitätsstufen.