



## Collaborative Climate Community Data and Processing Grid (C3-Grid)

### Workflow-Dokumentation *cape*

---

Arbeitspakete:	Workflows (übergreifend AP1,3,4,5,6,7)
Autoren:	Diagnose Cluster (FUB, ZAIK, IGMK)
Version:	Generation 1 (Stand: Jul.2008)
Veröffentlichungsdatum:	Juli 2008
Koordination:	Universität zu Köln (ZAIK, IGMK), FUB
Partner:	AWI, ZIB, DKRZ, M&D, MPI/IFM-Geomar, Uni-Dortmund
Ansprechpartner:	T. Brücher, V. Achter
Email:	bruecher@uni-koeln.de

---

## Inhalt

Motivation

1 Einleitung

2 Einbindung des Diagnose-Tools „cape“ in das C3-Grid

3 Modularisierung von “cape”

4 Das Tool: *cape*

4.1 Methode

4.2 Variablen

4.3 Optionale Steuerung

5 Entwicklungsausblick

## **Motivation**

Durch die Erzeugung und Haltung von unterschiedlichen Datensätzen an ebenso verschiedenen Rechenzentren nehmen technische Aufgaben des Datenmanagements sehr viel Zeit des Wissenschaftlers in Anspruch. Dieses ist unter anderem durch ein unterschiedliches Datenhaltungskonzept, unterschiedliche Datenformate und ein uneinheitliches Rechtesystem bedingt. Damit dem Wissenschaftler mehr Zeit für die wesentlichen Aufgaben seines Spezialgebiets zur Verfügung stehen, ist eine einheitliche Schnittstelle, die die technischen Herausforderungen übernimmt, wünschenswert.

C3-Grid schafft, über diese einheitliche Schnittstelle zu den Datensätzen hinaus, die Kopplung von Diagnosetools an die verteilt liegenden Datensätze, so dass automatisiert Standardfragen / -untersuchungen schnell auf unterschiedlichste Beobachtungs- oder Modelldaten angewendet werden können.

## 1 Einleitung

Im C3-Grid werden diverse Tools zur Verfügung gestellt, die meteorologische Fragestellungen beantworten. Dieses sind unter anderem die Programme „cape“ und „qflux“, die sich dem Feuchtehaushalt der Atmosphäre widmen. „cape“ (convective available potential energy) liefert eine Maßzahl zur Stabilität der Atmosphäre, während „qflux“ den integrierten Feuchtefluß der Atmosphäre berechnet.

Beide Tools benötigen neben der Temperatur und dem horizontalen Windfeld den Feuchtegehalt (egal, ob relative oder spezifische Feuchte) auf dem Druckniveau der Atmosphäre für die zu untersuchende geographische Region.

Sollen mit diesen Diagnosetools klimatologische Fragestellungen beantwortet werden, so steigt die Anzahl der Einzelschritte enorm an, da auf Grund unterschiedlicher Datenanbieter, -formate und -quellen sowohl der Zugang, das Diagnosetool als auch die graphische Aufbereitung immer neu angepasst werden muss. Weiterhin stellen die zur Verfügung stehenden Ressourcen (z. B. Speicherplatz) oft Limitierungen dar.

## 2 Einbindung des Diagnose-Tools „cape“ in das C3-Grid

Der Workflow besteht aus einer Kombination aus dem Workflow des Datenproviders (Preprocessing) und dem Workflow des Serviceproviders (Feuchtetransport-Analyse) und lässt sich auf verschiedene Datensätze anwenden.

Im trivialen Fall umfasst der Workflow nur die Selektion einer Datenquelle und der Raum-Zeit Parametern. In einer späteren Version wird zusätzlich die Auswahl diverser Optimierungen (Einbindung von Bodendatensätzen zur genaueren Diagnose) bzw. physikalischer Formulierungen (unterschiedliche Formulierungen in der Berechnung von cape) wählbar sein. Diese Selektionen werden im Portal durchgeführt, die anschließende Datenbereitstellung wird durch den Datenprovider vorgenommen und beim Serviceprovider werden die diagnostischen Parameter berechnet. Abschließend stehen die Ergebnisse im Portal zum Download bereit.

Der Datentransfer zwischen den einzelnen Partnern wird durch den Daten Management Service (DMS) realisiert und durch den Scheduler gesteuert<sup>1</sup>. In der aktuellen Version (Generation 1) sind diese Schnittstellen nicht mehr fest verdrahtet, sondern der WF liegt in einer modularisierten, jedoch fest vorgeschriebenen Prozesskette vor. Das bedeutet, dass die wesentlichen, inhaltlich trennbaren Abschnitte separat voneinander durch den Scheduler gesteuert werden. Die prinzipielle Teilung dieser Prozesskette sieht wie folgt aus:

- a) Datenaufbereitung / Datenselektion
- b) Berechnung des / der Diagnoseparameter
- c) Statistische Weiterverarbeitung der Ergebnisdaten
- d) Visualisierung der Ergebnisse

### **3 Modularisierung von Workflows:**

Bei dem hier verwendeten Ansatz zur Modularisierung von Workflows sind die separaten Abschnitte so zu formulieren, daß kein „Wissen“ der Module untereinander vorausgesetzt werden darf. Das bedeutet, dass der Namensraum zum einen durch die global definierten Umgebungsvariablen bekannt ist, zum anderen wird jedem Modul der komplette Satz an notwendigen Parametern übergeben. Via Ansteuerung eines Parsers zur Übergabe der Parameter an die Prozesskette werden alle Module nacheinander – durch den Scheduler auf deren Abhängigkeit geprüft und deren Ausführung kontrolliert – ausgeführt. Der Ablauf ist sequentiell.

---

<sup>1</sup> Da die Vorgehensweise analog zum Stormtrack-Workflow läuft, wird hier für eine graphische Darstellung des Ablaufs verzichtet und auf Abbildung 1 dessen Workflow-Dokumentation verwiesen.

## 4 Das Tool: *cape*

Das Diagnosetool „cape“ ist in einer PERL bzw. SHELL Umgebung eingebettet, die sich um die Analyse der Daten und Metadaten kümmert, den eigentlichen Programmaufruf samt Kommandozeilenparametern startet und ein Postprocessing welches aus einer einfachen statistischen Analyse besteht. Für die abschließende graphische Aufbereitung der Ergebnisse wird dynamisch ein Script erzeugt, das mit Hilfe von GrADS ein Bild im gif-Format erzeugt.

Die visuellen und numerischen Ergebnisdaten werden vom DMS an einen Webserver gesendet, auf den das Portal zugreift und die Ergebnisse dem Nutzer sichtbar macht. Die Ergebnisse können von dort über das Web auf den lokalen Klienten des Nutzers transferiert werden.

### 4.1 Methode

Die Berechnung der konvektiven, verfügbaren potentiellen Energie beruht auf einer vertikalen Integration des lokalen Auftriebs. Dabei startet die Berechnung beim Niveau des freien Auftriebs (LFC: level of free convection) und endet dort, wo das Paket nicht weiter aufsteigt (LZB: level of zero bouyancy). Dabei sind unterschiedliche Ansätze möglich, CAPE zu berechnen. Zum einen kann bei der Integration die Erhaltung reversibler Entropie angenommen werden, zum anderen die Erhaltung der Äquivalent-Potentiellen Temperatur.

Da bei der Integration der Temperaturunterschied des gedachten Luftpakets zur Umgebungstemperatur entscheidend ist, gehen in die Bestimmung der Temperaturunterschiede die Feuchtegrößen mit ein. Mit der Analyse von CAPE ist es möglich, die Stabilität der Atmosphäre zu bestimmen. CAPE wird in J/kg angegeben und Werte größer Null bedeuten, dass Luftmassen aufsteigen. Bei großen CAPE-Werten (von der Region abhängig) ist mit einer erhöhten Schauer- und Gewitteraktivität zu rechnen, so ist beispielsweise bei einem Sommergewitter CAPE hoch.

## 4.2 Variablen

Parameter:	Temperatur, relative oder spezifische Feuchte auf Druck-Niveaus, optional die Bodenparameter Temperatur, Taupunkt, Bodendruck
Geographische Region:	vom Nutzer frei wählbar (vertikal und horizontal)
Zeitraum:	vom Nutzer frei wählbar
Zeitliche Auflösung:	vom Nutzer frei wählbar

## 4.3 Optionale Steuerung

- Bestimmung des In- und Output Datenformats (grb oder NetCDF)
- Zuschaltung von Bodenparametern, damit insbesondere die Analysen im unteren Bereich der Atmosphäre optimiert / korrigiert werden können
- Festlegung der gewünschten physikalischen Formulierung für CAPE

## **5 Entwicklungsausblick**

Folgende Erweiterungen / Ergänzungen bieten sich an, den Funktionsumfang (aus meteorologischer Sicht) zu erweitern und mit erhöhten Funktionalität den Benutzerkreis des Diagnosetools zu vergrößern.

- Zuschaltung von Bodenparametern, damit insbesondere die Analysen im unteren Bereich der Atmosphäre optimiert / korrigiert werden können
- Erweiterung für Daten auf Modell-Niveau

Insbesondere die Ansteuerung optionaler Parameter bietet für das C3-Grid inhaltlich eine große Erweiterung, WFs zukünftig komplexer zu gestalten. Diese Funktionalität erwartet, dass Abhängigkeiten definiert werden können, die im Portal aufgelöst werden. So sollte zum Beispiel dem Nutzer angezeigt werden, daß eine Erweiterung der Standardberechnung in Abhängigkeit des selektierten Datensatzes möglich ist.

Die Erweiterung des C3Grids um eine frei kombinierbare Prozesskette auf Modulbasis stellt sicher, daß „cape“ auf Grund seiner Modularisierung mit weiteren Anwendungen

kombinierbar ist. Jedoch ist zu überprüfen, ob die aktuelle Version der vollen Flexibilität von zukünftigen Generationen gerecht wird, oder weitere technische Arbeiten an der Ansteuerung des Diagnosetools notwendig sind. Dieses ist insbesondere im Hinblick auf automatisierten Ensemble-Analysen (multi model multi member analysis) wichtig.