

Übungen zu Klimamodelle Abschlussbericht

von

Salmi Michele und Thomas Aistleitner

1 Einführung

In der heutigen Zeit spielen Klimamodelle, sogenannte Global Circulation Models (GCM) eine wichtige Rolle in der Vorhersage für das zukünftige Klima. Hierbei gibt es viele verschiedene Herangehensweisen, vor allem in der Auflösung und in den Annahmen unterscheiden sich die Modelle. Einerseits gibt es weniger komplexe Simulationen, die sich einfacher Bilanzgleichungen bedienen und andererseits gibt es Modelle mit komplexer Parametrisierung der meteorologischen Größen. Das größte Problem der GCMs stellt jedoch die Vereinbarung der Ergebnisse mit der lokalen Skala (local scale) dar. Während großräumige Prozesse beziehungsweise auch Teleconnections im Klima (NAO, ElNino) mit GCMs besser berechnet werden können, ist es schwer, Prognosen auf die lokale Skala zu bringen. In diesem Sinne wurde eine Methode namens downscaling entwickelt, welche die globale Skala (global scale) mit der lokalen Skala (local scale) verknüpft. Hierbei gibt es wiederum zwei verschiedene Arten: statistisches und dynamisches downscaling. In der Übung Klimamodelle haben wir mit Hilfe der Analog Methode (statisches downscaling) versucht, ein Klimamodell zu erstellen.

Die Analog Methode beschreibt ein Verfahren, welches zum Beispiel täglich in der Wettervorhersage benutzt wird. Ein Meteorologe studiert täglich die Wetterkarten und schaut sich die Wetterlage in Europa an. Je länger sich dieser bereits in seinem Berufsleben mit dem Wetter beschäftigt, desto leichter ist es, eine ähnliche Wetterlage aus den aktuellen Karten mit einer aus der Vergangenheit zu verknüpfen. Dies stellt dann eine sogenannte Analogie dar und der Meteorologe weiß dann, mit welchen Wettererscheinungen er in gewissen Gebieten rechnen muss. In ähnlicher Weise funktioniert die Analog Methode auch in der Klimatologie, nur dass man hier mit Hilfe von Klimadaten aktuelle Prognosen vergleicht.

2 Reanalyse Daten

Am Anfang der Übung wurde jeder Gruppe ein gewisser Reanalyse Datensatz zugeordnet. Dieser Gruppe wurden die japanischen Reanalyse Daten (JRA-55) zugeordnet.

JRA-55

Die japanische 55-jährige Reanalyse oder auch JRA-55 wurde von der Japanischen Meteorologischen Agentur (JMA) als zweite japanische globale Reanalyse durchgeführt. Diese beginnt im Jahr 1958, ab diesem Jahr wurden regelmäßige Radiosondenbeobachtungen durchgeführt, und geht bis ins aktuelle Jahr. Im Gegensatz zur Vorgängerversion (JRA-25) werden hier fortschrittlichere Datenassimilations-schemata (4Dvar anstatt 3Dvar), eine höhere Modellauflösung sowie eine neue Variationskorrektur für Satellitendaten und mehrere zusätzliche Beobachtungsdatenquellen verwendet.

Bei den Reanalyse-Daten gibt es einen großen Pool an verfügbaren Daten, im Zuge dieser Übung mussten wir dann drei davon auswählen.

- SLP
- spezifische Feuchte 700hPa
- relative Feuchte 700hPa

	JRA-25	JRA-55
Years	26 years (1979–2004)	55 years (1958–2012)
Model version	as of Mar. 2004 operational	as of Dec. 2009 operational
Resolution	T106L40 (~120 km) (top layer at 0.4 hPa)	TL319L60 (~60 km) (top layer at 0.1 hPa)
Time integration	Eulerian	Semi-Lagrangian
Long-wave radiation	<i>Line absorption</i> Statistical band model <i>Water vapor continuum</i> e-type	<i>Line absorption</i> Table lookup + K-distribution <i>Water vapor continuum</i> e-type + P-type
Assimilation scheme	3D-Var (with T106 resolution)	4D-Var (with T106 inner model)

Abbildung 1: Vergleich JRA-25 und JRA-55

Wie in der Grafik ersichtlich, handelt es sich um Daten über 55 Jahre mit einer horizontalen Auflösung von 60km und einer zeitlichen Auflösung von 6h.

Die Daten wurden mit Hilfe eines Linux Skriptes von einem Server der UCAR runter geladen. Zuvor musste hier aber ein Account angelegt werden, um einen Zugang zu den Daten zu erhalten. Um die Größe der Daten bereits vor dem Herunterladen zu minimieren, wurden diese schon vorher auf die vorher bestimmten Breiten- und Längengrade geschnitten.

- Längengrad: λ : 10°W - 25°E
- Breitengrad: ϕ : 32.5°N - 67.5°N

Mit dieser Einschränkung wurde die Größe der Daten deutlich verringert und diese betrug somit nur noch etwa einen Gigabyte.

3 Aufbereitung der Daten

3.1 Anomalien

Zuerst wurden die Daten mit Hilfe von dem Python-Modul *xarray* geladen. Da die Daten im JRA-55 auf 6-Stundenbasis vorliegen, mussten diese zuerst auf Tagesbasis gemittelt werden. Anschließend wurde die Klimatologie der Daten berechnet. Hierbei wurde jeweils ein Tag des Jahres (365) ausgewählt und ein Fenster von ± 10 Tagen hinzugefügt, um den Pool an verfügbaren Tagen zu erhöhen.

Um nun die Anomalien zu bilden, wurden die Werte, die sich aus der Berechnung der Klimatologie ergaben, von den einzelnen Werte am Tag abgezogen. Außerdem wurde, um Fehler mit dem Datum und Schaltjahren zu vermeiden, der 29. Februar aus den Daten gelöscht.

3.2 Normieren

Um verschiedene Parameter vergleichen zu können, ist es sinnvoll die Anomalien zu normieren, sprich dimensionslos zu machen. In unserem Fall wurden drei verschiedene Parameter gewählt und diese mussten normiert werden. Dazu wird die neben der Klimatologie des Tages mit dem Zeitfenster auch die Standardabweichung berechnet. Anschließend dann die Anomalie durch die Standardabweichung dividiert und somit erhält man normierte Anomalien.

4 EOF und Analog-Methode

Wie bereits erwähnt, stellt die Analog-Methode ein Verfahren dar, welches versucht zukünftige Ereignisse mit Situationen aus der Vergangenheit darzustellen. Im Fall von diesem Klimamodell, waren es Daten aus Spartacus welche wir zur Verfügung hatten. Um nun die Ereignisse vergleichbar zu machen und auch das Problem zu vereinfachen, werden sogenannte EOF-Analysen verwendet.

4.1 EOF

Eine EOF-Analyse (Empirical Orthogonal Function) kommt oft in der Klimatologie zum Einsatz, um bestimmte Muster in den Daten zu erkennen (z.B: NAO) beziehungsweise kleine Störungen raus zu filtern. Die EOF's werden nun mit Hilfe eines Python-Paketes berechnet. Hierbei bestand ein Problem bei der Berechnung mit *xarray*. Da multivariate EOF's (EOF's mit mehreren Variablen) nicht berechnet werden konnten, mussten die Daten zuerst wieder in ein Paket namens *Iris* transferiert werden. Anschließend konnten die EOF's für jeden Tag des Jahres (DOY) und für das Zeitfenster bestimmt werden. Bei der Berechnung der EOF's wurde dann eine Bedingung bezüglich der Varianz gesetzt. Würde man eine Varianz von 100% nehmen, würden eine große Anzahl an EOF's benötigt. So wurden die EOF's solange berechnet, bis die Summe der Varianzen der verschiedenen EOF's 90% erreichen. (1.EOF 40% + 2.EOF 25% + 3.EOF 10% ... j 90%). In den Ergebnissen zeigte sich dann, dass es hier saisonale Unterschiede gibt.

Zu den EOF's wurden des Weiteren auch noch die sogenannte *principal components* (PC) berechnet, diese stellen Zeitreihen dar. Für jeden Zeitpunkt (Tag) und für jede EOF (1.,2.,3.) bekommt man einen Wert und dieser Wert gibt jeder EOF-Analyse für diesen bestimmten Tag einen bestimmte Rang.

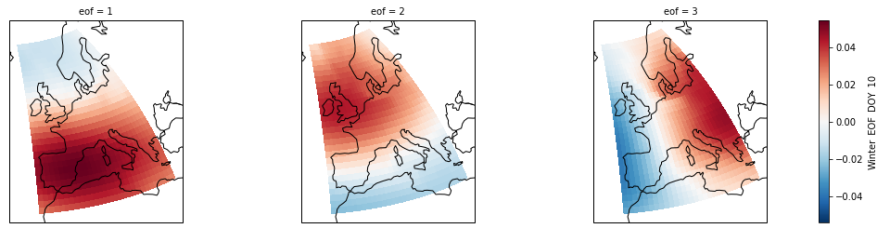


Abbildung 2: EOF Analyse für DOY10 (10.Jänner)

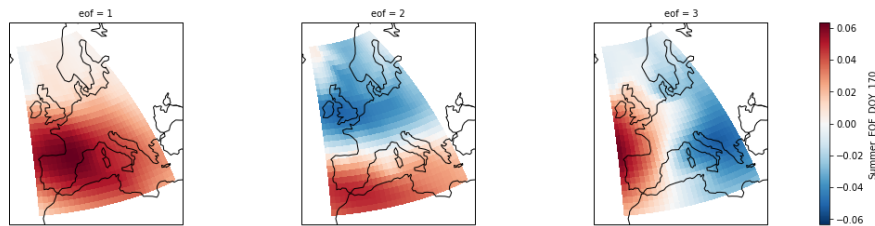


Abbildung 3: EOF Analyse für DOY170 (19.Juni)

4.2 Analog-Methode

Mit den EOF-Analysen sowie mit der Berechnung der PC's für jeden Tag kann nun die Analog-Methode angewandt werden. Dabei wird eine Schleife über alle Tage des Jahre gemacht und wieder darauf geachtet, dass kein Problem mit dem 29. Februar besteht. Dieser eine Tag wird nun mit den ± 10 Tagen des Datensatzes verglichen und sogenanntes Analog gesucht. Hierbei wird die Norm zwischen den PC's und den sogenannten Pseudo-PC's berechnet. Wenn die Norm klein ist, stellt dieser bestimmte Tag ein Analoga dar. Dieser Schritt wurde dann 10 mal wiederholt, um die 10 besten Analoga dieses bestimmten Tages zu finden. Als Ergebnis standen für jeden Tag des gesamten Zeitraumes 10 analoge Tage zur Verfügung.

Ein Problem stellte bei diesen Berechnungen immer der Arbeitsspeicher da. Aufgrund der großen Datenmenge, kam es bei den einzelnen Schritten oft zu einem *memory error*, sodass der Code oft umgebaut werden musste, damit die Berechnung funktionieren konnten.

5 Validierung

Mit den gewonnen analogen Tagen, wurde nun mit Hilfe des Spartacus Datensatzes eine Klimamodell programmiert.

Dabei stellt Spartacus einen gegitterten Beobachtungssatz dar und beschreibt die räumliche Verteilung von der täglichen Lufttemperatur sowie der Niederschlagssummen in Österreich. Diese Datenbank wurde von der ZAMG entwickelt und beinhaltet Daten von 1961 bis jetzt.

Um die Ergebnisse validieren zu können, kamen nun zwei verschiedene statistische Maße zur Anwendung. Einerseits einfach die Differenz (BIAS) zwischen den Beobachtungen und Vorhersagen, andererseits der Root Mean Square Error (RMSE). Für die Verifikation der Daten wurde nun ein Jahr ausgewählt und durch dieses iteriert. Als Beobachtung wurde immer ein Tag im Jahr gewählt und für die Vorhersagen die daran geknüpften Analoga.

Als Parameter wurde die Maximaltemperatur gewählt. Im Spartacus-Datensatz gibt es dazu zwei verschiedene, die gemittelten Maxima über Österreich und das Maximum an jedem Gitterpunkt. Da mit den bereits gemittelten Werte schneller gerechnet werden konnte, wurden diese ausgewählt. Bei den Ergebnissen zeigt sich, dass die Unterschieden zwischen den Maxima an den Gitterpunkten und den bereits Gemittelten nicht sehr groß waren.

6 Ergebnisse

In Abbildung 4 sind nun der RMSE sowie auch der BIAS der ersten drei Analoga dargestellt. Mehr Analoga hätten den Plot sehr unübersichtlich gemacht, deswegen wurden nur die ersten drei ausgewählt. In den Ergebnissen zeigt sich nun, dass es bei den Analoga an bestimmte Tagen große Unterschiede gibt. Vor allem im RMSE kommen diese Unterschiede aufgrund der Definition des RMSE (quadrieren der Werte) besser raus. Eine saisonale Abhängigkeit kann mit Hilfe dieser statistischen Parameter jedoch nicht festgestellt werden.

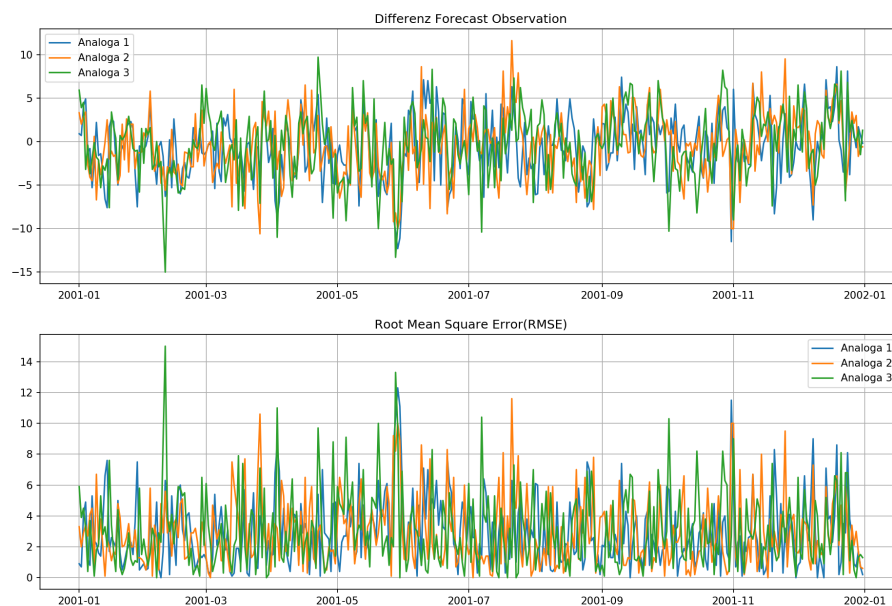


Abbildung 4: BIAS und RMSE für die Analoga

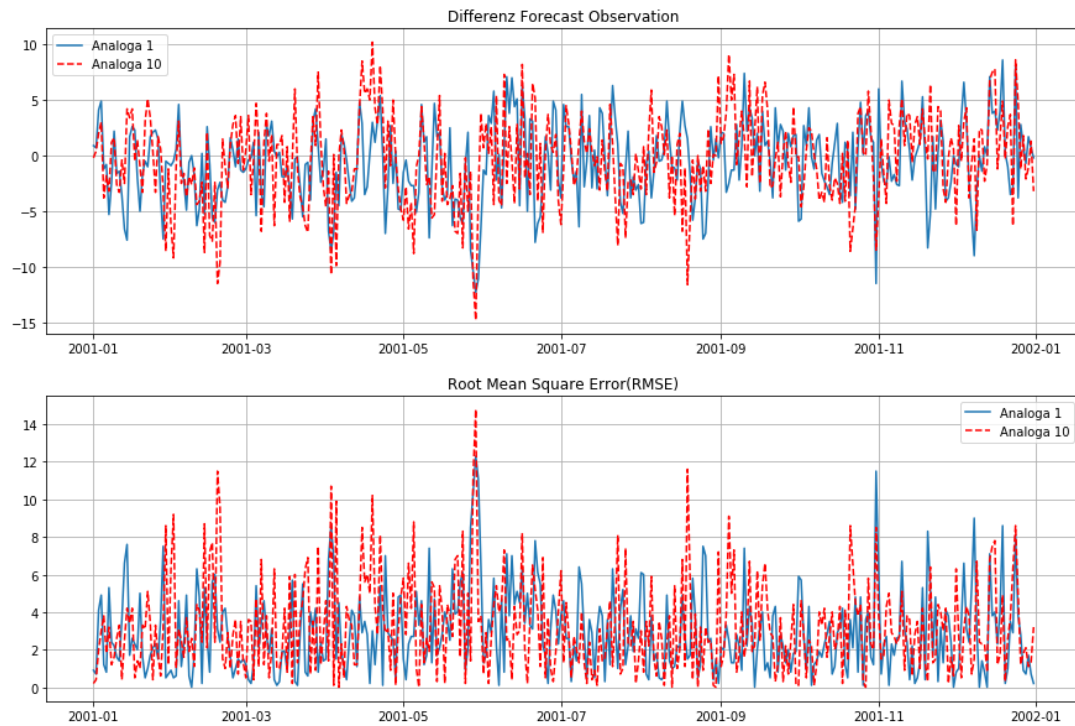


Abbildung 5: Differenz Analog 1 und 10

Die Differenz zwischen den verschiedenen Analoga ist am besten durch den direkten Vergleich zwischen dem 1. und dem 10. analogen Tag zu erkennen. Dies ist in Abbildung 5 zu sehen. Vor allem im RMSE zeigen sich deutliche Unterschiede.

Bei den Ergebnissen zur erklärten Varianz zeigte sich aber ein Unterschied in der Saisonalität. Während die im Winter nur etwa 45 EOF benötigt werden, werden im Sommer über 70 EOF Analysen benötigt, um 90% der Varianz zu erklären.

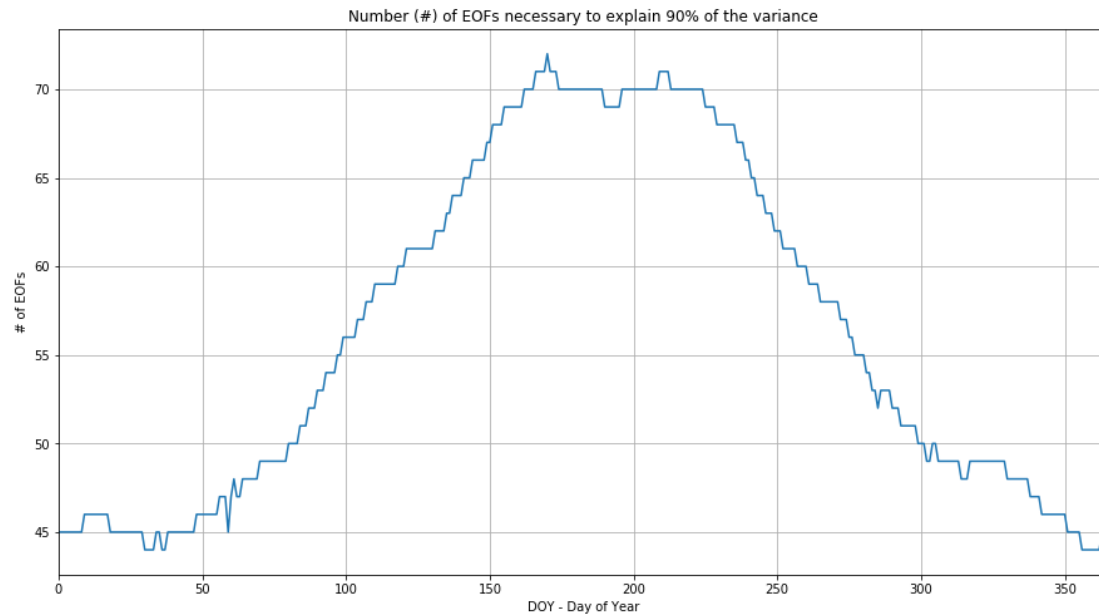


Abbildung 6: Anzahl EOF's um Varianz zu erklären

Wie weiter oben erwähnt, wurden auch aus dem Spartacus Datensatz zwei verschiedene Parameter für die Validierung gewählt. Am Ende zeigte sich aber, dass der Unterschied nicht sehr groß ist und deshalb der effizientere Parameter gewählt wurde.

	Analoga 1	Analoga 2	Analoga 3
RMSE Tx	3.92	3.40	3.55
RMSE Tx.mean	2.90	2.92	3.07
BIAS Tx	-0.62	-0.56	-0.43
BIAS Tx.mean	-0.62	-0.57	-0.44

Die Tabelle zeigt nun die Werte für den BIAS und RMSE für die beiden Parameter. Im BIAS gibt es nur wenig Differenzen und im RMSE ist es nur die Größenordnung die den Unterschied ausmacht,

Abschließend lässt sich sagen, dass sich die Analog-Methode für ein downscaling recht gut eignet. Jedoch muss der Code auf jeden Fall sehr effizient arbeiten, da

große Datenmengen verarbeitet werden müssen. Auch in der Wahl der Parameter zur Verifikation würden noch andere zur Verfügung stehen (Niederschlag), doch aufgrund von anderen, weitaus komplexeren Verifikationsmethoden, hätte dies den Rahmen dieser Arbeit aber übertroffen.