

UNIVERSITÄT WIEN

KLIMAMODELLIERUNG ÜBUNG

PROJEKTDOKUMENTATION

Analog Methode

Author:

Viktoria PRUCKNER

Hanna ROHRINGER

Susanna WINKELBAUER

Leiter:

Georg SEYERL

Sebastian LEHNER

25. Juli 2019

Inhaltsverzeichnis

1	Einleitung	2
2	Durchführung	2
2.1	Daten Download	2
2.2	Preprocessing	3
2.3	Anwendung Analog-Methode	3
2.4	Validierung	4
3	Ergebnisse	5
4	Schlussfolgerung	9

1 Einleitung

Mithilfe der Reanalysedaten der japanischen meteorologischen Agentur JRA-55 (Japanese 55-year Reanalysis) und Beobachtungsdaten der Zentralanstalt für Meteorologie und Geodynamik (ZAMG) soll eine Kreuzvalidierung durchgeführt werden.

Da in der Datenanalyse oft mit zeitlich und räumlich sehr großen Datenmengen gearbeitet wird, wenden wir eine EOF (Empirical Orthogonal Function) - Analyse an um die Datendimension zu verringern. Ziel hierbei ist es den gesamten Datensatz durch so wenig EOFs wie möglich möglichst genau darzustellen ohne wesentliche Informationen zu verlieren.

Wir wollen aus GCMs Informationen über das Klima auf der regionalen Skala ableiten. Dafür wenden wir statistisch empirisches Downscaling an und suchen uns Transferfunktionen von der LASC (Large Scale) auf die LOSC (Local Scale).

Als Downscalingverfahren verwenden wir die Analogmethode, welche auf Basis der Hauptkomponentenanalyse arbeitet und statistisch ähnliche Verteilungsmuster in den großräumigen meteorologischen Feldern sucht. Durch Minimierung der Distanz der PCs wird das ähnlichste Feld determiniert und angenommen, dass diese Ähnlichkeit von der großen Skala ebenso auf die lokalskaligen Wettererscheinungen übernommen werden kann.

2 Durchführung

2.1 Daten Download

Die Reanalysedaten JRA-55 werden über das Research Data Archive des National Center of Atmospheric Research (NCAR) heruntergeladen.

Folgende Parameter werden in 6 stündlicher Auflösung über eine definierte Domäne von 10° West bis 25° Ost und 32.5° Nord bis 67.5° Nord über den gesamten verfügbaren Zeitraum (1958 - 2018) heruntergeladen:

- Luftdruck am Boden (SLP)
- spezifische Feuchte in 700 hPa (SPFH)

- relative Feuchte in 700 hPa (RH)

Die SPARTACUS Beobachtungsdaten von der ZAMG liegen auf einem 1x1km Gitter mit täglicher Auflösung vor und beschreiben die räumliche Verteilung der Lufttemperatur und Niederschlagssumme in Österreich.

2.2 Preprocessing

Da wir auf täglicher Basis arbeiten, müssen unsere Reanalysedaten zu aller erst auf den Tag gemittelt werden.

Außerdem interessieren wir uns nur für die Veränderungen, also für die Stationen wo sich viel tut, deshalb berechnen wir die Anomalien indem wir die Stationsmittelwerte von den Stationsdaten abziehen. Hierzu wird zum jeweiligen Target Day ein Zeitfenster von ± 10 Tagen bei der Mittelwertbildung berücksichtigt und der Mittelwert für jeden Tag des Jahres berechnet, sodass 365 Mittelwerte für jeden Gitterpunkt resultieren. Diese Mittelwerte werden anschließend von den Stationswerten abgezogen um die Anomalien zu erhalten.

Um mögliche Probleme wegen der unterschiedlichen Größenordnungen der verschiedenen Parameterfelder zu vermeiden werden die Anomalien normiert. Hierzu wird die Standardabweichung für jedes Zeitfenster TD ± 10 Tage über die gesamte räumliche Dimension berechnet und die Anomalien durch die jeweilige Standardabweichung dividiert.

2.3 Anwendung Analog-Methode

Mit den normierten Anomalien kann nun die eigentliche Analog Methode durchgeführt werden.

Es wird eine multivariate EOF-Analyse für jeden Tag im Jahr erstellt um die Dimension der Daten zu verringern. Da es in xarray keine Funktion hierfür gibt müssen die Daten erst in iris cubes umgewandelt werden um dann die multivariate EOF-Analyse durchzuführen. Man erhält daraus die EOFs und PCs für jeden Tag. Die Anzahl der EOFs wird über die erklärte Varianz, welche auf 90 Prozent gesetzt wird, determiniert.

Durch Projektion der Reanalysedaten in den EOF-Raum erhält man die Pseudo

PCs. Jedes Feld kann nun als Linearkombination aus EOF mit zugehörigem (Pseudo) PC dargestellt werden.

Anschließend wird für jeden zeitlichen Punkt der Daten die Norm zwischen Pseudo PCs und PCs bestimmt. Hierzu werden die Pseudo PCs von den PCs abgezogen, quadriert und aufsummiert. Es resultiert eine Norm für jeden Zeitpunkt.

Von diesen Normen wird nun das Minimum gesucht (argmin Funktion), dieses entspricht dem Tag welcher dem gesuchten am ähnlichsten ist, also unserem Analogon.

2.4 Validierung

Die Transferfunktionen der Analog Methode müssen mit einem Teil der Daten kalibriert und mit einem anderen Teil validiert werden. Bei uns wird ein Jahr aus den Reanalysedaten nicht in die EOF-Berechnung einbezogen um damit die Validierung durchzuführen.

3 Ergebnisse

Es konnte mithilfe der Reanalysedaten und der multivariaten EOF Analyse ein Analogon für jeden Tag des gesamten Zeitraums gefunden werden. Das letzte Jahr (2018) wird zur Validierung verwendet und fließt nicht in die EOF-Analyse ein. Somit ergeben sich 21915 Tage mit zugehörigen Analogons. Ein Teil der gefundenen Tage ist in Abbildung 1 zu sehen. Die gesamte Liste steht im beigelegten Python Skript zur Verfügung.

Für den letzten Tag der Liste werden die Anomalien der drei Parameter P, SPFH

	search_day	found_day
0	1958-01-01	1977-01-10
0	1958-01-02	1976-12-24
0	1958-01-03	1976-12-24
0	1958-01-04	2010-12-29
...
0	2017-12-30	1966-01-01
0	2017-12-31	1991-01-08

Abbildung 1: Liste der gefundenen Analogons zu jedem untersuchten Tag.

und RH, die in die EOF Analyse einfließen, mit denen des Analogons verglichen(siehe Abbildung 2). Für alle drei Parameter lassen sich Ähnlichkeiten der Anomalien zwischen betrachtetem Tag und zugehörigem Analogon erkennen. Im Druckfeld sind, vor allem am 31.12.2017, das Islandtief und Azorenhoch sehr gut zu erkennen. Demzufolge gibt es an beiden Tagen ein Minimum im Norden Europas, einmal östlich und einmal westlich von Großbritannien. Nach Süden hin nimmt der Druck zu. Der höchste Druck liegt im Fall vom 31.12.2017 über Teilen Spaniens und westlich davon, im zugehörigen Anaologon sind die höchsten Druckwerte an der Nordküste Afrikas zu erkennen.

Die beiden Feuchtemaße zeigen an beiden Tagen ein Maximum über Skandinavien. Ein weiteres Maximum befindet sich im Osten Europas. Am 31.12.2017 liegt dieses Maximum etwas weiter im Norden, am 8.1.1991 etwas weiter im Süden. Am 31.12.2017 befindet sich ein Maximum über Frankreich, am 8.1.1991 liegt dieses Feuchtefeld etwas weiter im Norden, an der Südküste von Großbritannien. Das

Gebiet mit der geringsten Feuchtigkeit liegt am 31.12.2017 über dem Mittelmeerraum und erstreckt sich bis ins Landesinnere von Südosteuropa. Auch am 8.1.1991 ist diese Minimum im Süden Europas, wenn auch mit schwächerer Intensität, zu erkennen. An beiden Tagen fällt ein kleines Gebiet mit geringer Feuchtigkeit über Mitteleuropa auf, das in ihrer Lage und Ausprägung sehr starke Ähnlichkeiten aufweist.

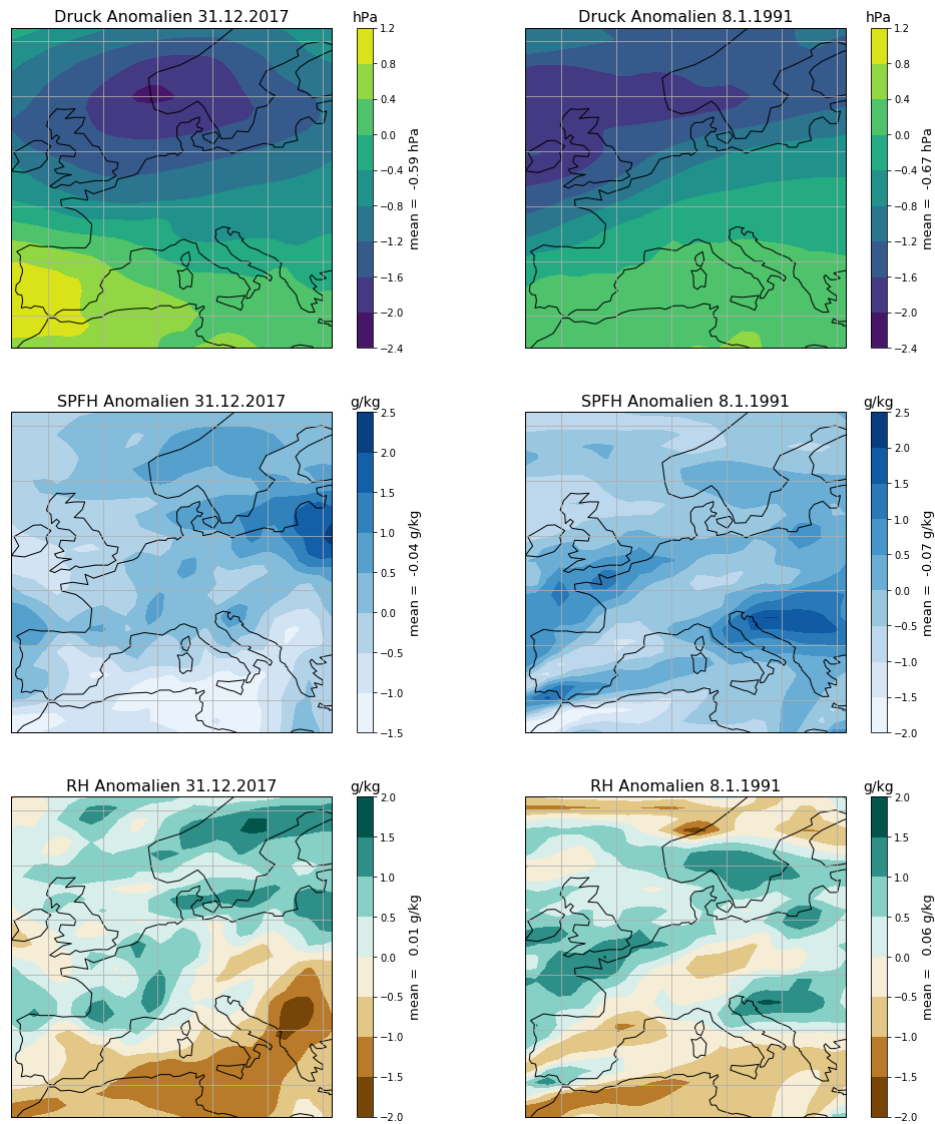


Abbildung 2: Anomalien der drei Parameter P, SPFH und RH am 31.12.2017 und dem dazugehörigen Analogon dem 8.1.1991.

Abbildung 3 zeigt die PCs und den ersten EOF des Druckfeldes. Es werden bei einer erklärten Varianz von 90 Prozent 49 EOF Felder und $49 \cdot \text{timerange}$ PCs, also 49 PCs pro Tag, gefunden. Das gesuchte Feld F kann nun als Linearkombination aus den EOFs und PCs wie folgt gebildet werden:

$$F = PC_1 * EOF_1 + PC_2 * EOF_2 + \dots + PC_i * EOF_i$$

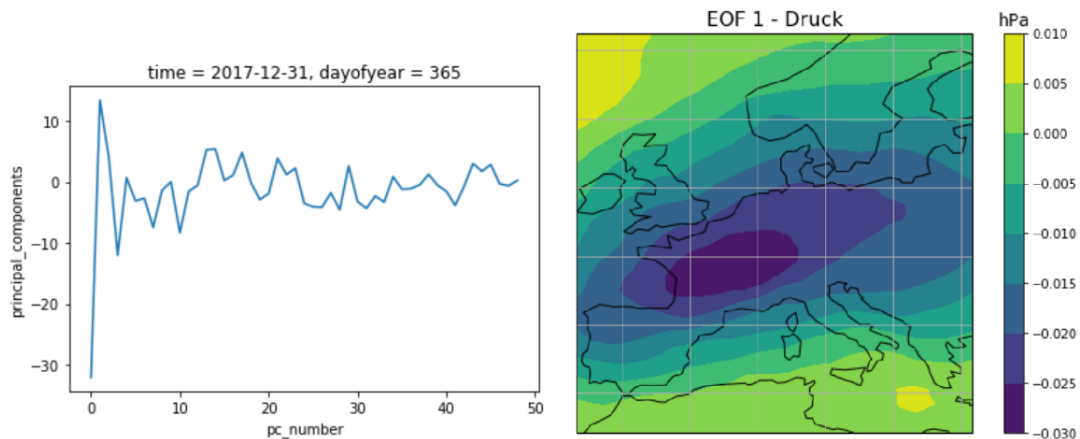


Abbildung 3: PCs vom 31.12.2017 (*links*) und Feld des ersten EOFs des Druckes (*rechts*).

Abbildung 4 zeigt eine Gegenüberstellung für alle drei Parameter von den Anomalien der tatsächlichen Felder mit den Anomalien der mit EOF und PC rekonstruierten Felder. Das rekonstruierte Druckfeld zeigt ein sehr ähnliches Muster wie das tatsächliche Feld. Wieder ist der steigende Druck von Nord nach Süd gut zu sehen. Auch im Mittel treten keine großen Unterschiede auf. So liegt zwischen tatsächlichem und rekonstruiertem Feld über die gesamte betrachtete Fläche im Mittel ein Unterschied von 0.02 hPa.

Auch die beiden Feuchtemaße zeigen zwischen tatsächlichen und rekonstruierten Feldern eine recht gute Übereinstimmung. Das Maximum im Nordosten Europas tritt jedoch etwas weiter westlich über Norddeutschland auf und auch das Feuchtigkeits Minimum liegt im rekonstruierten Feld weiter im Westen über dem Mittelmeer. Bei der relativen Feuchte werden im Mittel über die gesamte Fläche die selben Werte wie im tatsächlichen Feld erreicht, bei der spezifischen Feuchte werden diese jedoch um 0.17 g/kg überschätzt. Grund dafür ist das besonders stark ausgeprägte Maximum im Norden Deutschlands welches beim tatsächlichen Feld

größtenteils außerhalb der betrachteten Domäne liegt.

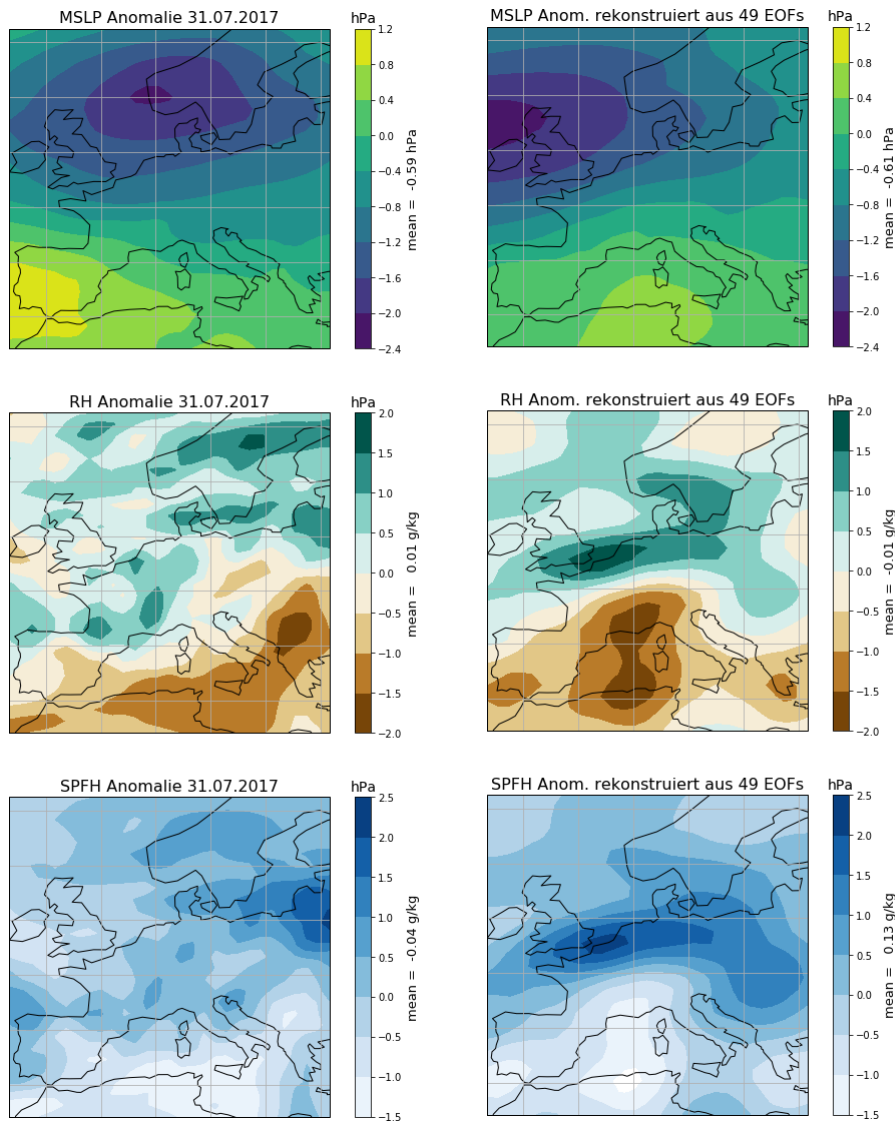


Abbildung 4: *links:* Anomalien von Druck sowie relativer und spezifischer Feuchte wie in Abb. 2, *rechts:* durch Linearkombination von PC und EOF rekonstruierte Felder.

Zur Validierung wird der Root Mean Square Error (RMSE) herangezogen. Tabelle 1 zeigt die berechneten RMSE Werte zwischen den tatsächlichen Feldern, den Analogons und den durch EOFs rekonstruierten Feldern.

Der RMSE zwischen tatsächlichem Feld(31.12.2017) und Analogon(8.1.1991) er-

reicht bei den einzelnen Parametern nicht unbedingt sein absolutes Minimum. So liegen beispielsweise die RMSE Werte beim Druck (MSLP) am Tag des Analogons mit 0.08 höher als an den unmittelbaren Tagen danach (am 9.1. RMSE von 0.03). Betrachtet man jedoch alle drei Parameter gleichzeitig so liegt ein RMSE von 0.05 vor, das absolute Minimum von allen betrachteten Tagen.

Auch beim durch die EOFs rekonstruierten Feld treten relativ niedrige RMSE Werte auf, einzig der Wert der spezifischen Feuchte (SPFH) erreicht knapp 0,2. Dies ist jedoch schon in der obigen Abbildung (Abb 4) gut zu sehen, hier weicht das rekonstruierte Feld des SPFH am stärksten vom tatsächlichen Feld ab.

RMSE	Analogon	rek. Feld
MSLP	0.08	0.02
RH	0.04	0.03
SPFH	0.03	0.17
Gesamt	0.05	0.07

Tabelle 1: Vergleich der RMSE Werte von tatsächlichem Feld zu Analogon bzw. rekonstruiertem Feld.

4 Schlussfolgerung

Nach anfänglichen Schwierigkeiten konnten wir die Analog-Methode erfolgreich umsetzen, was sich in den angestellten Gegenüberstellungen widerspiegelt. Ziel der Analog-Methode ist für jeden Tag im Jahr jenen Tag zu finden, an dem verschiedene meteorologische Felder ein ähnliches Verteilungsmuster aufweisen. Ein Vergleich der Anomalien des betrachteten Tages mit den Anomalien des errechneten Analogons weist eindeutig Ähnlichkeiten auf. Gleiches gilt auch für den Vergleich der Anomalien des tatsächlichen Feldes mit jenen Feldern, die sich aus der Linearkombination von PC und EOF ergeben. Zum Schluss wird die Ähnlichkeit der betrachteten Felder mit Hilfe des RSME quantifiziert. Die sehr geringen Werte des RSME lassen lediglich auf geringe Abweichungen schließen.