



universität
wien

MAGISTERARBEIT / MASTER'S THESIS

Titel der Magisterarbeit / Title of the Master's Thesis

„Zeitreihenanalyse von globalen Temperaturdaten unter
zeitlichem und geografischem Aspekt“

verfasst von / submitted by

Levin Tröster, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magister der Sozial- und Wirtschaftswissenschaften (Mag. rer. soc. oec.)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 951

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Magisterstudium Statistik

Betreut von / Supervisor:

ao. Univ.-Prof. Dipl.-Ing. Dr. Erhard Reschenhofer

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne Benutzung anderer als die angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Wien, am 28.09.2022

Levin Tröster

Inhalt

Abstract	7
1 Einführung.....	8
2 Daten.....	9
3 Methoden	15
3.1 Hodrick-Prescott-Filter	15
3.2 Imputation fehlender Werte	23
3.3 Testen auf Stationarität.....	23
4 Fazit.....	31
Literatur	33

Abstract

Seit vielen Jahren ist der Klimawandel eines der wichtigsten Themen unserer Gesellschaft, da er eine potentiell enorme Bedrohung für unseren Planeten darstellt und uns damit alle betrifft. Für viele Menschen besteht kein Zweifel, dass der Klimawandel existiert, einige verleugnen ihn aber noch. Diese Masterarbeit soll anhand von statistischen Analysen von Temperaturzeitreihen zeigen, dass die Luft sowohl über den Landesoberflächen, als auch über den Meeresoberflächen in den vergangenen 140 Jahren merkbar wärmer geworden ist und demnach auch, dass sie sich in den kommenden Jahren vermutlich weiterhin erwärmen wird und dadurch eine ernstzunehmende Gefahr für das Leben auf der Erde vorliegt. Es werden dabei mögliche Unterschiede des Verhaltens der Temperaturen in den einzelnen Monaten sowie den verschiedenen Regionen der Erde – von den Polen bis zu dem Äquator - untersucht. Zunächst werden die Zeitreihen auf verschiedene Arten visualisiert. Durch Anwenden des Hodrick-Prescott-Filters werden bereits steigende Trends erkennbar gemacht, die anschließend mithilfe von Inferenzstatistik bestätigt werden sollen. Es wird für jede Zeitreihe der fraktionale Differenzenparameter d geschätzt und basierend auf diesem Wert festgestellt, ob die jeweilige Zeitreihe stationär ist oder nicht. Kann keine Stationarität nachgewiesen werden, so spricht das für eine Steigung der Temperatur und somit für den Klimawandel.

1 Einführung

Die Änderung des Klimas auf der Erde beschäftigt schon lange viele Menschen. Demnach ist es nicht verwunderlich, dass nicht nur in der heutigen Zeit, in der dieses Thema mehr diskutiert wird als kaum ein anderes, sondern schon vor vielen Jahren Forschungen in diesem Bereich angestellt wurden. Wie jede Wissenschaft hat sich die Statistik im Laufe der letzten Jahrzehnte weiterentwickelt, es sind neue Erkenntnisse erlangt sowie fortgeschrittenere Methoden erarbeitet worden. Frühe Analysen von Temperaturzeitreihen basierten beispielsweise auf einfacher linearen Regression, in der die Steigung der Kurven geschätzt wurde, wie es *Grenander & Rosenblatt (1957)* gemacht haben. Eine andere Vorgehensweise wählten *Dickey & Fuller (1979)*, die die Zeitreihen mit ihrem bis heute etablierten *augmented Dickey-Fuller-Test (ADF)* analysiert haben, welcher auf Vorhandensein einer Einheitswurzel testet. Sie kamen zu dem Ergebnis, dass die beobachtete Zeitreihe eine solche besitzt, also dass sie $I(1)$ ist, was bedeutet, dass nach einmaligem Differenzieren Stationarität erreicht wird. Das gleiche Resultat erlangten auch *Kwiatkowski et al. (1992)* durch Anwendung des ebenfalls renommierten Kwiatkowski-Phillips-Schmidt-Shin-Tests (KPSS).

Eine noch stärker Aussage konnten *Phillips & Perron (1988)* tätigen. Mit ihrem Phillips-Perron-Test, der auf dem vorangegangenen Dickey-Fuller-Test basiert, haben sie festgestellt, dass die Temperaturzeitreihen $I(0)$ und somit stationär sind. Zusammen mit Peter Schmidt kam Phillips 4 Jahre darauf zur selben Erkenntnis (*Schmidt & Phillips 1992*). Ein von *Vogelsang (1998)* eingeführter, robuster Trend-Test war die Grundlage für die Analyse von *Fomby & Vogelsang (2002)*, die damit bereits eindeutige Ergebnisse erhielten, welche für einen positiven linearen Trend in globalen Temperaturdaten sprechen. Um herauszufinden, welchen Einfluss der Mensch auf die Veränderung der Temperaturen hat, haben *Estrada et al. (2013)* die Brüche in den Zeitreihen betrachtet und sind zu dem Schluss gekommen, dass ein deterministischer Trend vorhanden ist, was für den menschengemachten Klimawandel spricht. Ein ähnliches Resultat stammt von *Estrada et al. (2017)*, die ebenfalls Strukturbrüche in den Trendfunktionen der Daten entdecken konnten.

In dieser Arbeit werden Zeitreihen betrachtet, die durch einen ARFIMA-Prozess erklärt werden können. Diese Prozesse enthalten den sogenannten fraktionalen Differenzenparameter d , der in Abschnitt 3.3 geschätzt wird, um so auf Stationarität der Zeitreihen zu testen. Die Nullhypothese lautet, dass das d einen Wert annimmt, der kleiner als 0.5 ist. Nachdem das erfüllt sein muss, damit das Integral der Spektraldichte des Prozesses existiert, wird unter der Nullhypothese also angenommen, dass die Zeitreihe stationär ist. Kann diese Hypothese verworfen werden, hat man also einen Nachweis für die Nichtstationarität gefunden und kann behaupten, dass die Zeitreihe einen Trend aufweist. Dieses Verfahren wird in weiterer Folge auf mehrere Temperaturzeitreihen angewendet, womit festgestellt werden soll, in welchen Regionen der Erde und in welchen Monaten eine starke Veränderung in den vergangenen 140 Jahren stattgefunden hat.

2 Daten

Für die Analyse werden zwölf Zeitreihen aus dem „NOAA Merged Land Ocean Global Surface Temperature Analysis“ Datensatz (NOAAGlobalTemp) von der National Oceanic and Atmospheric Administration (NOAA) verwendet, die jeweils für einen bestimmten geografischen Bereich durchschnittliche Temperaturanomalien der Land- und Meeresoberfläche beinhalten, wobei die Anomalien auf der Klimatologie von 1971 bis 2000 basieren. Inkludiert sind die Temperaturdaten von jedem Monat seit Jänner 1880 bis Dezember 2021. Die unterschiedlichen geografischen Bereiche sind hier jeweils die Flächen zwischen zwei Breitenkreisen. Auf diese Weise wurde die Erde in sechs Abschnitte gegliedert, nämlich 90°S bis 60°S, 60°S bis 30°S, 30°S bis 0°N, 0°N bis 30°N, 30°N bis 60°N und 60°N bis 90°N. In Abbildung 1 sind die kompletten Zeitreihen veranschaulicht.

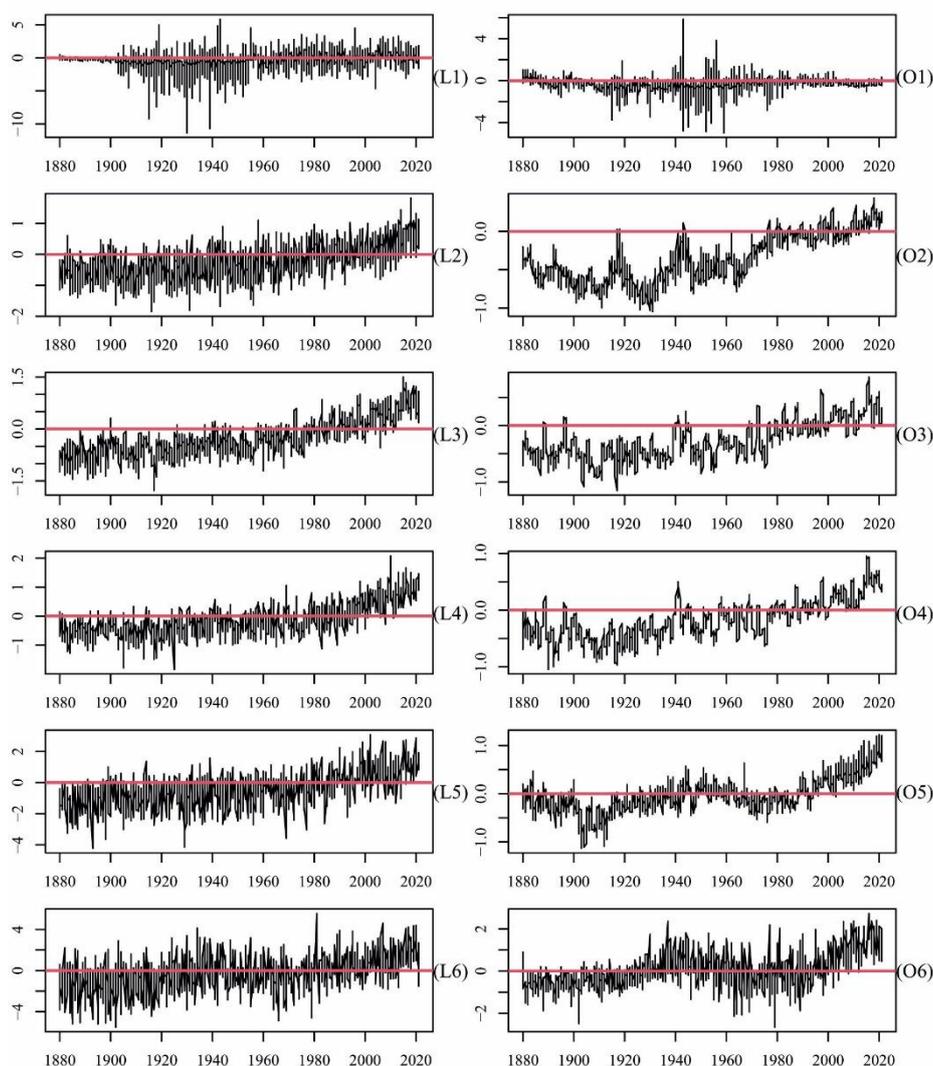


Abbildung 1. Monatliche Temperaturanomalien von Jänner 1880 bis Dezember 2021. Die mit „L“ markierten Zeitreihen stellen die Landoberflächentemperatur dar, jene mit einem „O“ die Meeresoberflächentemperatur. Die Spalten zeigen die geografischen Zonen von Süd bis Nord mit (1) = 90°S-60°S bis (6) = 60°N-90°N

Auffällig ist die Zeitreihe für Landoberflächen im Bereich 90°S - 60°S (s. (L1) in Abbildung 1). Die Werte in den ersten 20 Jahren weisen eine unnatürlich geringe Varianz auf, was darauf deutet, dass es sich hier um Schätzwerte handelt, die aufgrund von fehlenden Daten eingefügt worden sind. Dennoch ist die Zeitreihe in diesem Zeitraum nicht ganz vollständig, es fehlen die Daten für 18 Monate. In Abschnitt 3.2 wird beschrieben, wie jene fehlenden Werte imputiert wurden.

In dieser Auswertung wird ein besonderer Fokus sowohl auf das Verhalten der Temperaturen in den einzelnen Monaten, als auch in den verschiedenen geografischen Zonen gelegt. Die Zeitreihen werden in weiterer Folge also jeweils separat in diesen Kategorien dargestellt und betrachtet. Abbildungen 2-5 bieten einen Überblick darüber, wie unterschiedlich sich die Daten in den einzelnen Kategorien verhalten. Auf den ersten Blick scheint es so, als habe der geografische Aspekt einen größeren Einfluss auf die Temperatur als der zeitliche Aspekt, da sich die Kurven in Abbildung 4 und Abbildung 5 stärker voneinander unterscheiden als in Abbildung 2 und Abbildung 3. Während in den Regionen nahe an den Polen starke Schwankungen von bis zu 15°C und mehr beobachtet wurden (s. Abbildung 4 (a)), kam es bei den Regionen rund um den Äquator teilweise nur zu einer Gesamtschwankung von 2°C (s. Abbildung 5 (c)). Tatsächlich hat aber auch die Jahreszeit einen großen Einfluss auf die Temperatur, auch wenn es etwas weniger offensichtlich ist. In Abbildung 4 ist deutlich zu erkennen, dass im südlichsten Teil der Erde (a) die Temperaturen am stärksten in den Monaten April bis August (Grün- und Blautöne) schwanken und in den Regionen, die dem Nordpol am nächsten gelegen sind (f), in den Monaten September bis März (Gelb- und Rottöne). Mit anderen Worten kommt es in den kalten Wintermonaten zu größeren Abweichungen als in den warmen Sommermonaten. Dieses Phänomen ist auch in den nach Monaten unterteilten Grafiken zu erkennen. In Abbildung 2 ist die Zeitreihe 90°S - 60°S in schwarz dargestellt und die Zeitreihe 60°N - 90°N in violett. Erstere weist in den Monaten, die man auf der Nordhalbkugel zum Sommer zählt, mehrere Ausreißer aus, wohingegen die violette Kurve ähnlich wie die anderen keine große Varianz besitzt. Genau umgekehrt ist es in den – aus der Sicht der Nordhalbkugel – Wintermonaten. Hier fällt die violette Kurve durch ihre verhältnismäßig großen Abweichungen auf und die schwarze Kurve verläuft weniger extrem.

Die Differenzen zwischen den Landoberflächen- und den Meeresoberflächentemperaturen sind dagegen tatsächlich eher gering. Vergleicht man jeweils Abbildung 2 mit Abbildung 3 und Abbildung 4 mit Abbildung 5, erkennt man, dass sich wieder hauptsächlich die Varianzen der Zeitreihen unterscheiden. Die Kurven in den Grafiken mit den Landoberflächen bewegen sich in einem größeren Wertebereich als jene bei den Meeresoberflächen. Die Luft über den Ozeanen der Erde schwankt also weitaus weniger stark als die Luft über den Landmassen.

Diese Zusammenhänge lassen sich anhand von Abbildungen 2-5 gut erkennen, allerdings ist es aufgrund der Schwankungen nicht ganz einfach den Trend der Kurven abzulesen. Aus diesem Grund werden die Zeitreihen im nächsten Abschnitt geglättet, woraufhin man bereits zu genaueren Erkenntnissen über die Forschungsfrage, nämlich ob es in den verschiedenen Regionen der Erde über die letzten 140 Jahre zu einer signifikanten Steigerung der Temperatur gekommen ist, kommen kann.

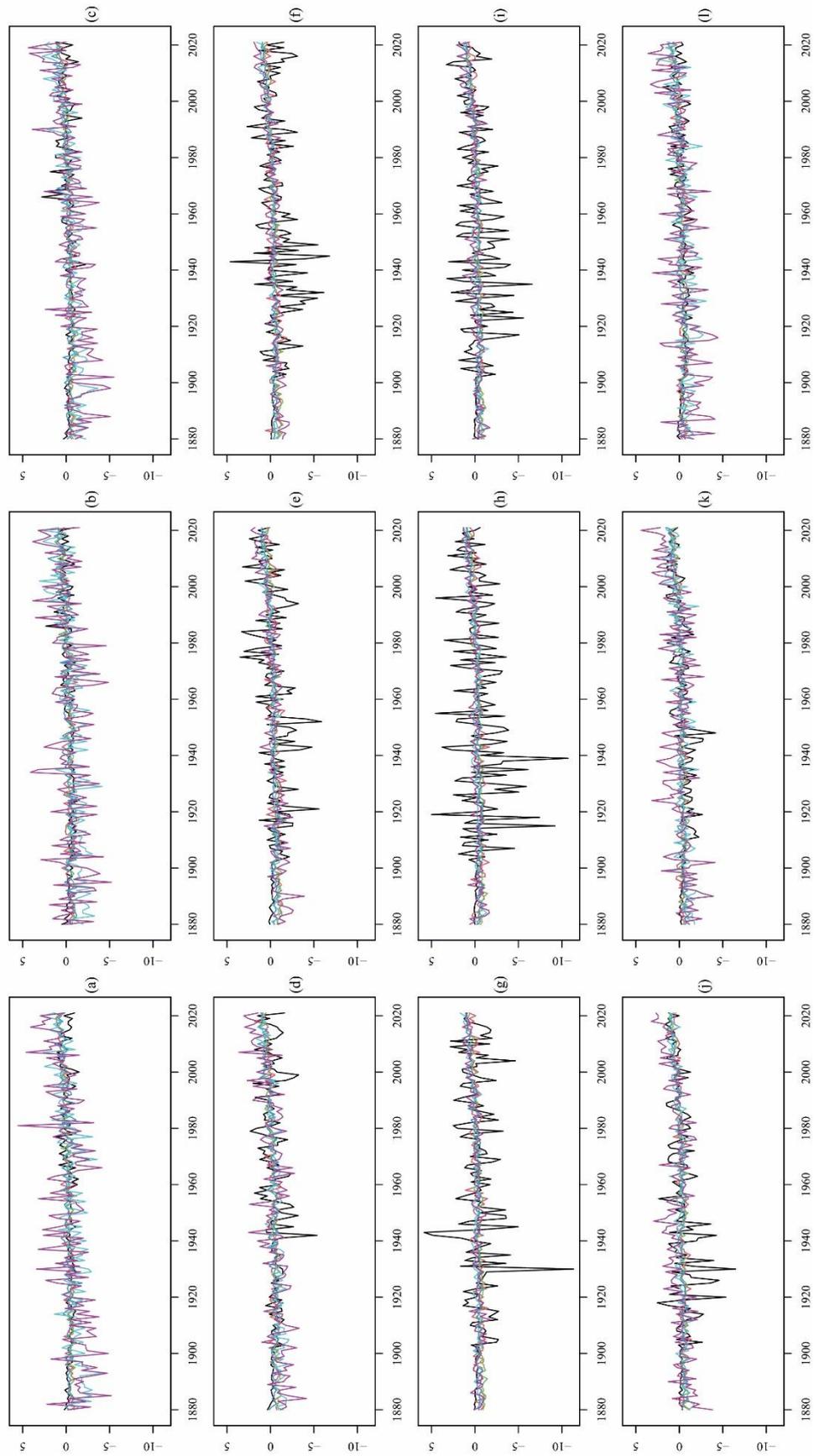


Abbildung 2. Anomalien Der Landoberflächentemperatur für (a) Januar, (b) Februar, ..., (l) Dezember. Die Linien stellen die geogr. Zonen dar.

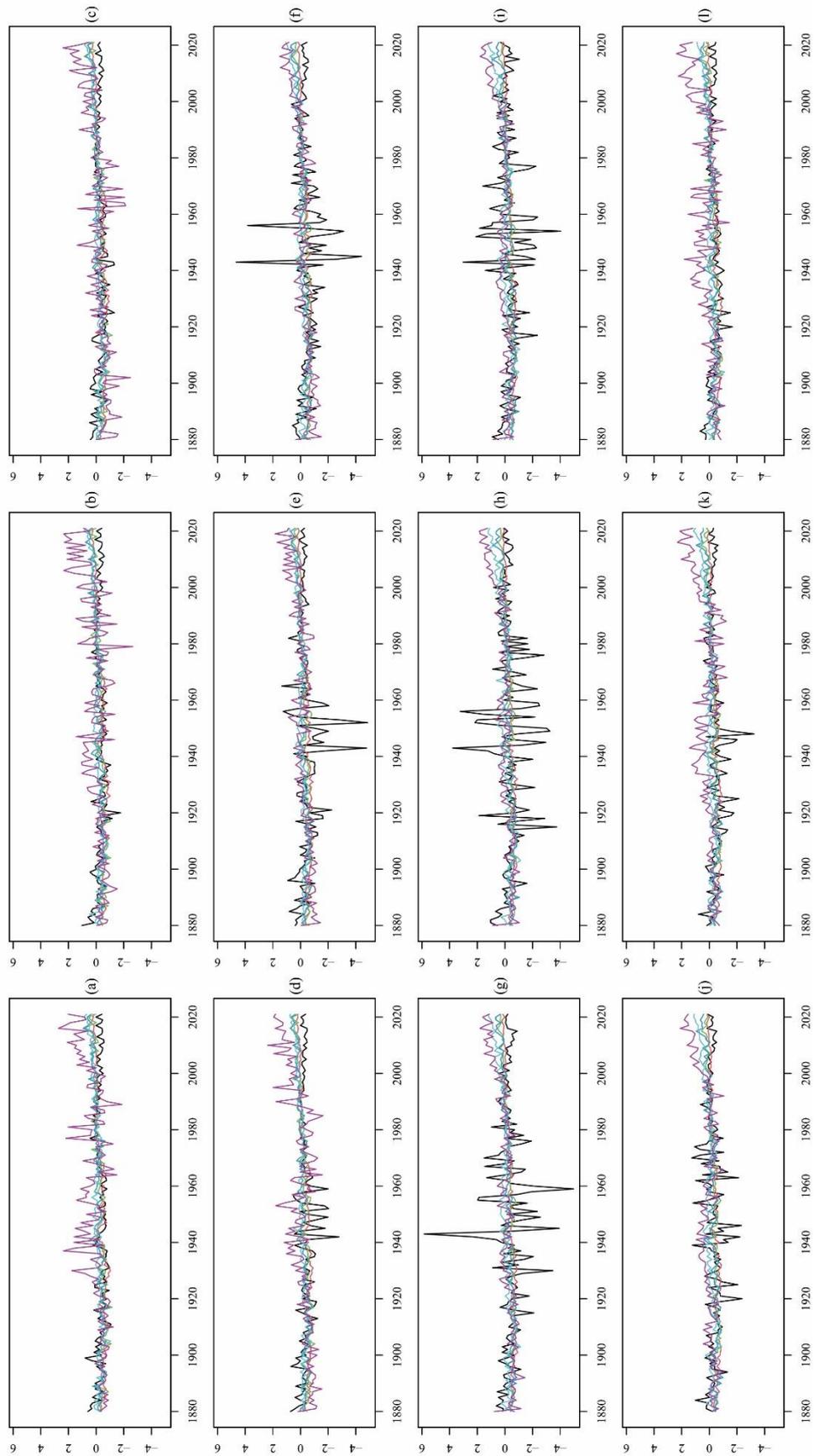


Abbildung 3. Anomalien der Meeresoberflächentemperatur für (a) Januar, (b) Februar, ..., (l) Dezember. Die Linien stellen die geogr. Zonen dar.

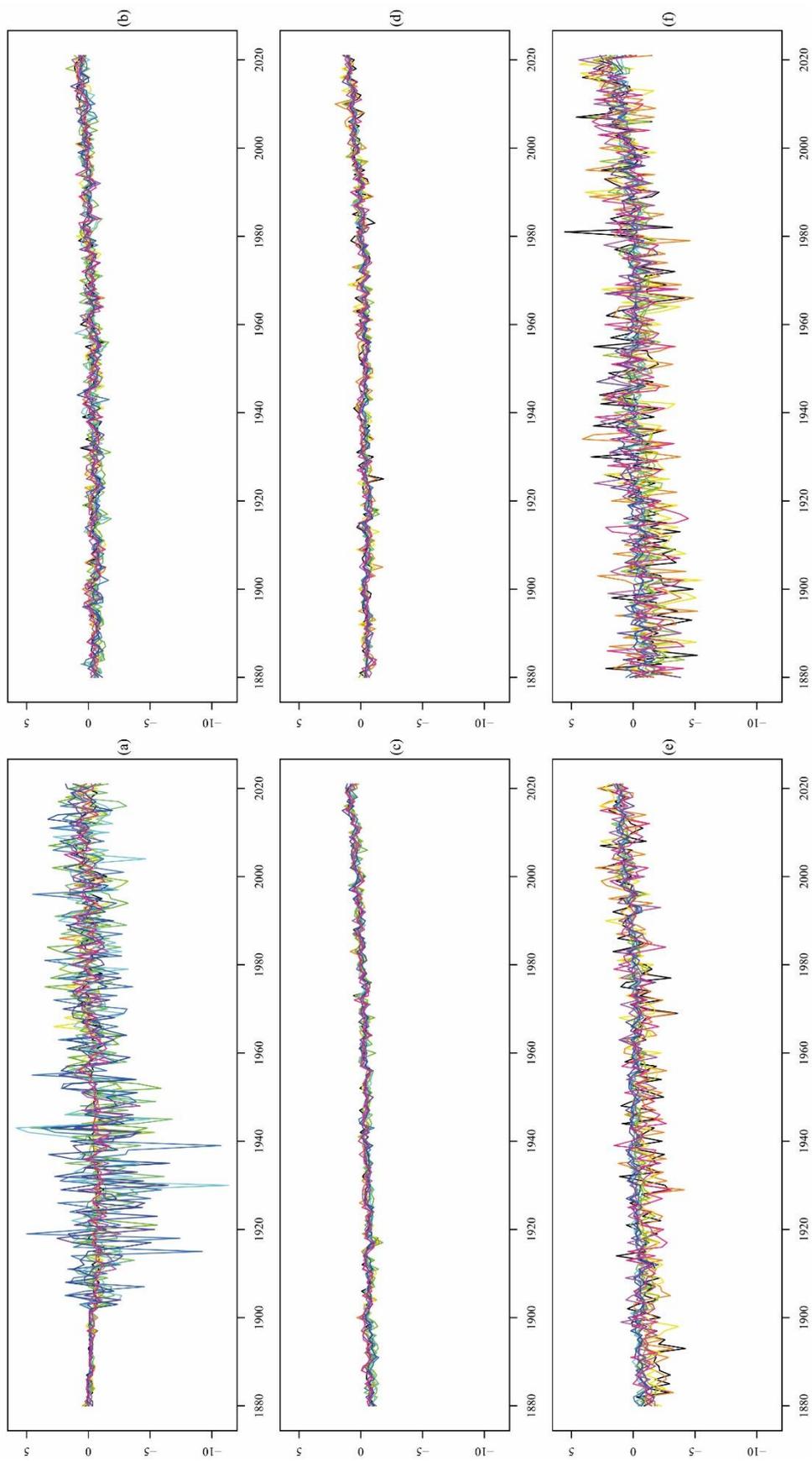


Abbildung 4. Anomalien der Landoberflächentemperatur für (a) 90°S - 60°S , (b) 60°S - 30°S , ..., (f) 60°N - 90°N . Die Linien stellen die 12 Monate dar.

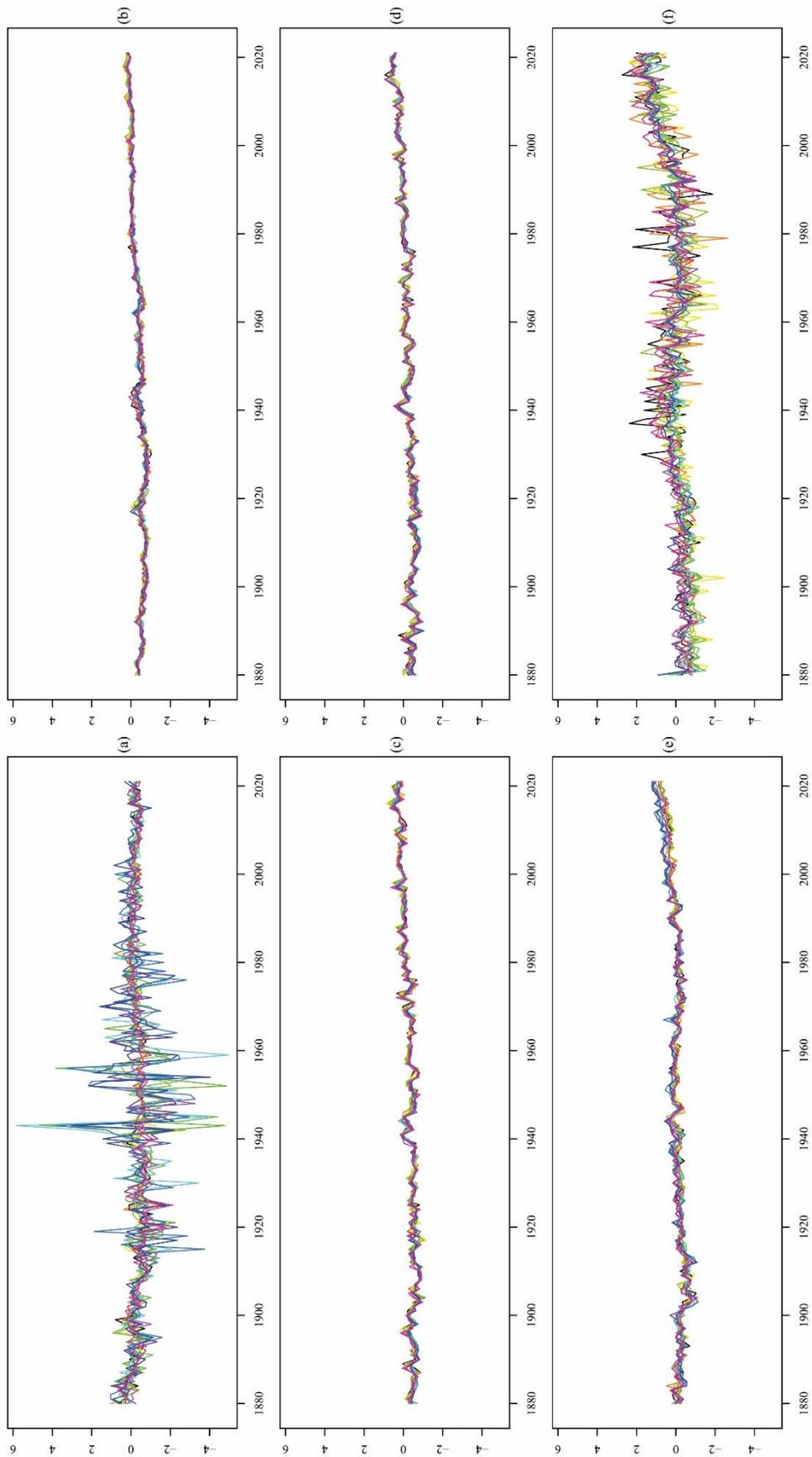


Abbildung 5. Anomalien der Meeresoberflächentemperatur für (a) 90°S - 60°S , (b) 60°S - 30°S , ..., (f) 60°N - 90°N . Die Linien stellen die 12 Monate dar.

3 Methoden

3.1 Hodrick-Prescott-Filter

Oft ist es schwierig bei langen Zeitreihen mit freiem Auge einen Trend abzulesen. Zu den Möglichkeiten das zu erleichtern gehört das Glätten der Kurve, also das Entfernen von kleineren Schwankungen. Hier wird eine der bekanntesten Methoden verwendet, namentlich der Hodrick-Prescott-Filter, benannt nach Robert J. Hodrick und Edward C. Prescott, die diesen Filter in den späten 1990er Jahren erstmals in ihren Publikationen verwendeten (Hodrick und Prescott; 1997). Die geglättete Zeitreihe $(\tau_t)_{t=1,\dots,T}$ erhält man durch das Minimieren des Ausdrucks

$$\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \cdot \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \quad (1)$$

wobei $(y_t)_{t=1,\dots,T}$ die ursprüngliche Zeitreihe und λ einen Parameter zur Steuerung des Glättungsgrades bezeichnet. Dieser Ausdruck besteht aus zwei Termen, und zwar einerseits aus

$$\sum_{t=1}^T (y_t - \tau_t)^2 \quad (2)$$

und andererseits aus

$$\lambda \cdot \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \quad (3)$$

Term (2) erfüllt den Zweck, dass die neue Zeitreihe τ_t möglichst nahe an der originalen Zeitreihe y_t ist, minimal wäre dieser Ausdruck für $\tau_t = y_t$ für alle $t \in (1, \dots, T)$. Auf der anderen Seite versucht Term (3) eine Zeitreihe zu konstruieren, die eine möglichst kleine Varianz besitzt, also idealerweise mit perfekt linearem Trend. Für den Fall $(\tau_{t+1} - \tau_t) = (\tau_t - \tau_{t-1})$ für alle $t \in (2, \dots, T-1)$ wäre (3) also gleich Null und somit minimal. Mit λ kann man demnach bestimmen, wie stark Term (3) ins Gewicht fallen soll, oder mit anderen Worten, wie wichtig eine konstante Steigung und wie vernachlässigbar die Entfernung zur ursprünglichen Zeitreihe ist. Je größer λ gewählt wird, umso mehr nähert sich die geglättete Kurve der klassischen Regressionsgerade an.

Da in dieser Auswertung weniger die kleineren Änderungen in den Zeitreihen, sondern mehr der grobe Trend der Temperaturen relevant ist, wurde mit $\lambda = 25000$ ein sehr großer Wert gewählt. In Abbildungen 6-9 sind die Ergebnisse des Hodrick-Prescott-Filters zu sehen.

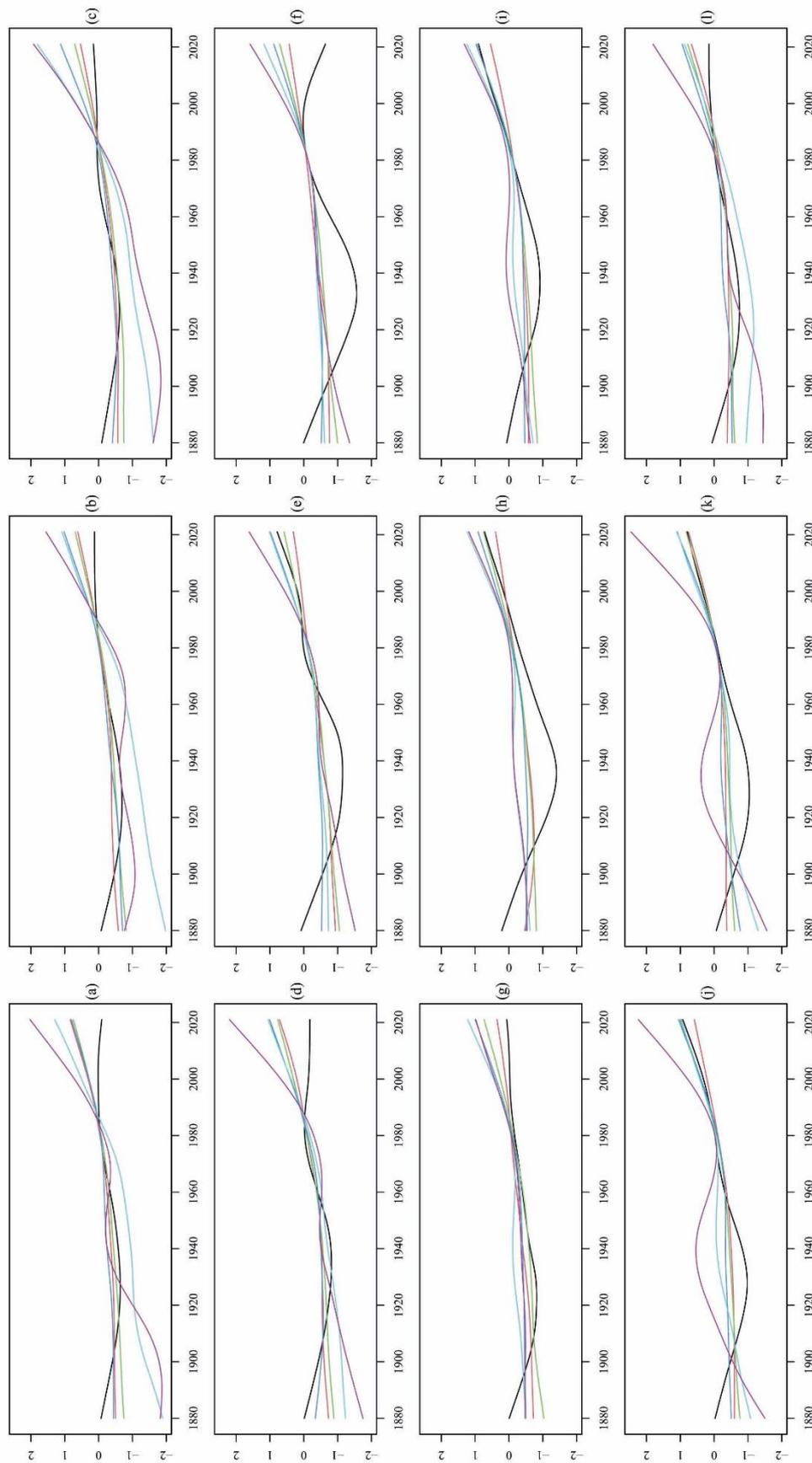


Abbildung 6. Anomalien der Landoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda = 25000$ für (a) Januar, (b) Februar, ..., (l) Dezember. Die Linien stellen die geogr. Zonen dar.

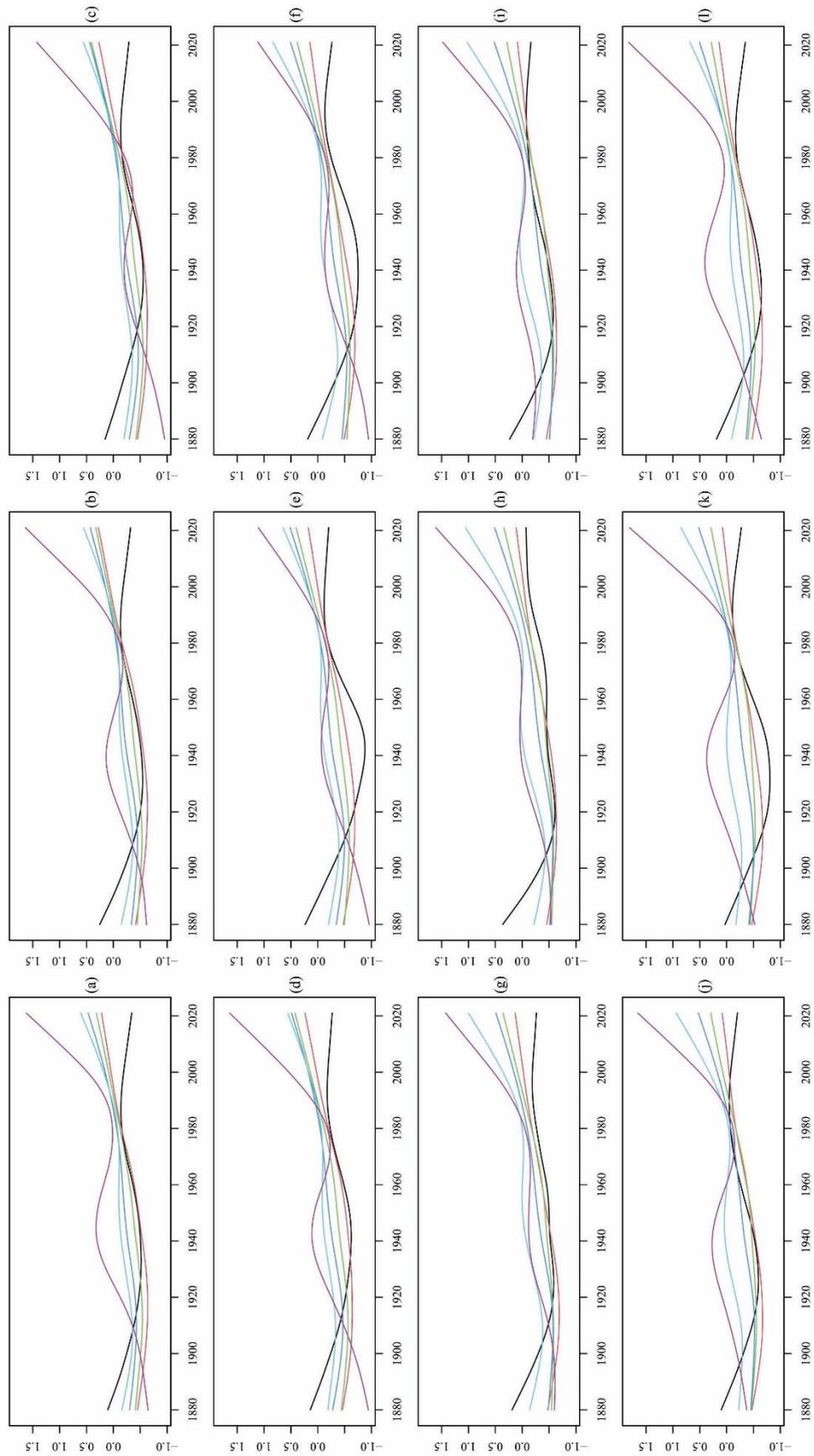


Abbildung 7. Anomalien der Meeresoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda=25000$ für (a) Januar, (b) Februar, ..., (l) Dezember. Die Linien stellen die geogr. Zonen dar.

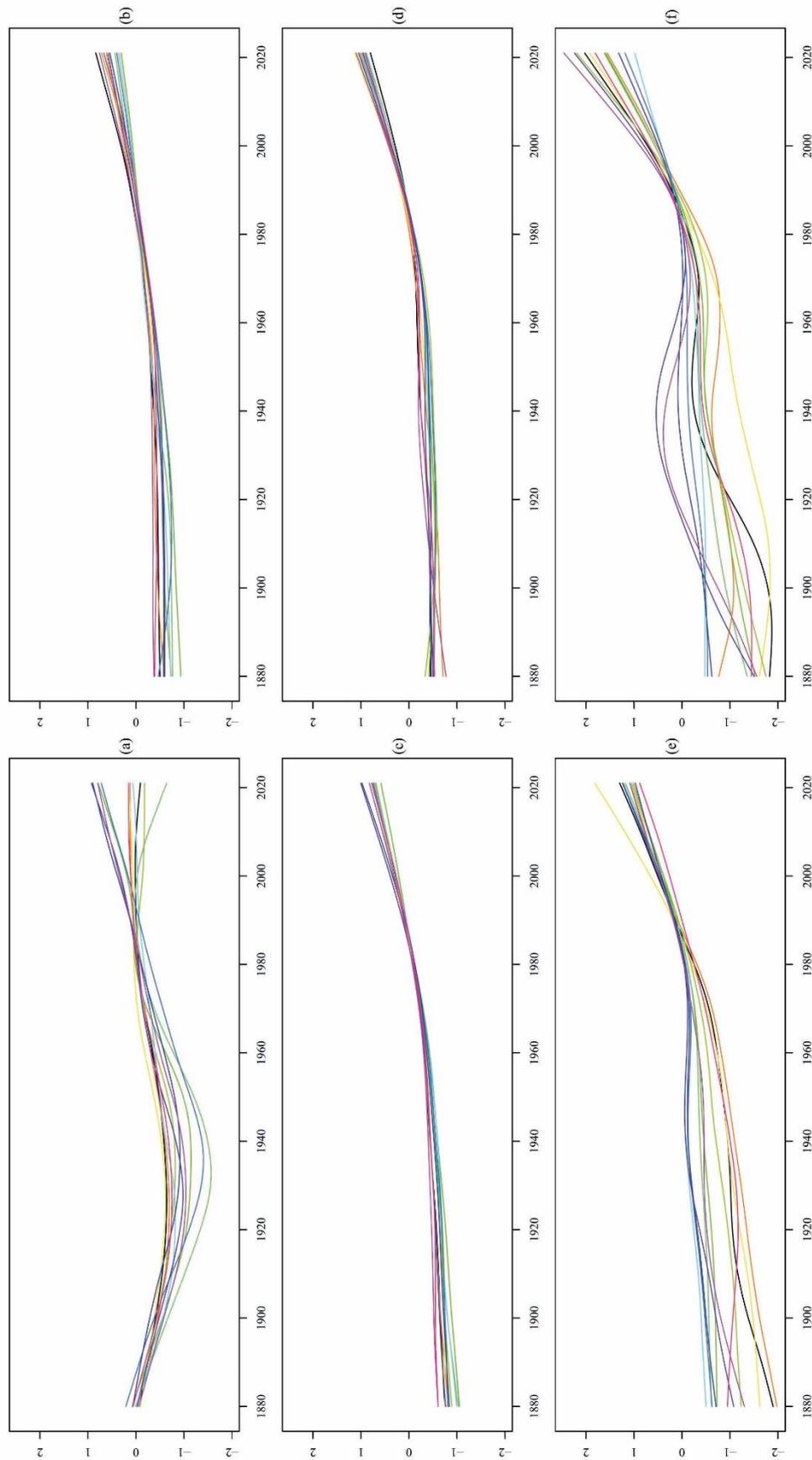


Abbildung 8. Anomalien der Landoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda=25000$ für (a) $90^{\circ}\text{S}-60^{\circ}\text{S}$, (b) $60^{\circ}\text{S}-30^{\circ}\text{S}$, ..., (f) $60^{\circ}\text{N}-90^{\circ}\text{N}$. Die Linien stellen die 12 Monate dar.

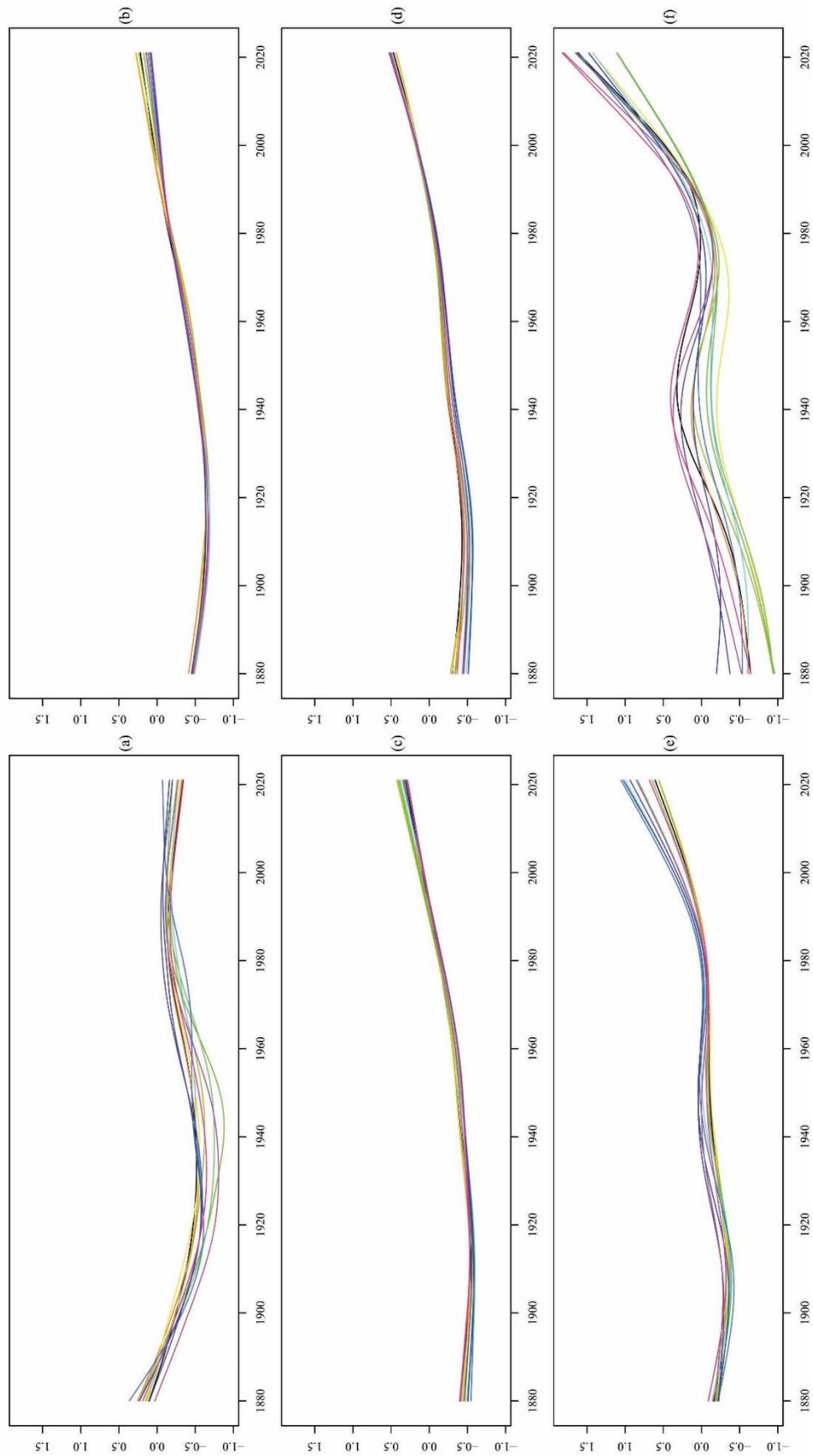


Abbildung 9. Anomalien der Meeresoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda=25000$ für (a) $90^{\circ}\text{S}-60^{\circ}\text{S}$, (b) $60^{\circ}\text{S}-30^{\circ}\text{S}$, ..., (f) $60^{\circ}\text{N}-90^{\circ}\text{N}$. Die Linien stellen die 12 Monate dar.

Nachdem die Kurven geglättet wurden, lassen sich direkt eindeutige Trends erkennen. Aus Abbildung 8 ist zu entnehmen, dass in den vergangenen Jahren bis auf die Region 90°S-60°S (a) überall eine Steigerung der Temperaturanomalien auftritt. Auch hier machen sich die regionalen Unterschiede bemerkbar. In den Zonen rund um den Äquator, (c) und (d), ist zuerst über viele Jahre hinweg fast gar kein Trend zu sehen. Erst um das Jahr 1980 bekommen die Kurven einen Knick und beginnen konstant zu steigen. Unterschiede zwischen den verschiedenen Monaten sind hier nicht zu erkennen, was aber zu erwarten ist; entlang des Äquators schwanken die Temperaturen über das Jahr gesehen schließlich generell kaum.

In den nördlicheren Gebieten, (e) und (f), sind nicht nur größere Differenzen unter den Monaten zu erkennen, sondern auch ein weitaus weniger geradliniger Verlauf der Zeitreihen. Genauer gesagt ist hier eine prägnante Ausbeulung rund um das Jahr 1940 zu sehen, also ein starker Anstieg der Temperaturanomalien gefolgt von einem ähnlich starken Rückgang. Bei den Meeresoberflächen, Abbildung 9 (f), ist dieser markante Trend sogar noch deutlicher ausgeprägt. Wie sich herausstellt, gibt es eine simple Erklärung dafür, nämlich die „Atlantische Multidekaden-Oszillation“ (AMO).

Bei der AMO handelt es sich um Schwankungen der Meeresoberflächentemperatur des (hauptsächlich nördlichen) Atlantiks. Einige Studien haben bereits gezeigt, dass die AMO auch einen starken Einfluss auf die Temperaturen in Europa und Nordamerika, sowie auf den globalen Klimawandel generell hat. Die hier zur Verfügung stehenden Daten bestätigen dies ebenfalls, wie im vorangegangenen Absatz bereits angedeutet. Abbildung 10 und Abbildung 11 zeigen nochmals die gleichen Temperaturdaten wie Abbildung 7 bzw. Abbildung 9, nur diesmal weniger stark geglättet ($\lambda=2500$) und inklusive der AMO-Zeitreihe in hellgrau, welche ebenfalls online von der NOAA zur Verfügung gestellt wird (<https://psl.noaa.gov/data/timeseries/AMO>). Sehr gut sieht man den Zusammenhang in Abbildung 10, da sich die lila Kurve (60°N-90°N) sehr ähnlich zu der AMO-Kurve verhält, vor allem bei der bereits erwähnten Ausbeulungen um das Jahr 1940. Abbildung 11 verdeutlicht wiederum, dass die AMO einen starken Einfluss auf die Temperaturen der Nordhalbkugel hat; je weiter es in den Süden geht umso weniger Einfluss ist erkennbar.

Stattdessen ist in der antarktischen Region (Abbildung 11 (a)) ein zu AMO fast komplett entgegengesetzter Zyklus zu wahrzunehmen. In der beinahe identen Zeitspanne tritt hier eine jahrzehntelange Ausbuchtung nach unten ein. Aufgrund dieser sich stark ähnelnden Trends, könnte man die Theorie aufstellen, dass AMO doch auch auf der Südhalbkugel einen Einfluss auf die Temperaturen hat, allerdings ist diese Theorie weder allgemein als richtig anerkannt, noch findet man Artikel, die sie versuchen zu bestätigen. Noch etwas naheliegender, wenngleich genau so schwierig nachzuweisen, ist die Theorie, dass die bekannten Oszillationen wie El Niño Southern Oscillation (ENSO), Interdecadal Pacific Oscillation (IPO), sowie der Southern Annual Mode (SAM) verantwortlich für diese großen Temperaturanomalien sind (Ejaz, Rahaman, Laluraj, Mahalinganathan und Thamban; 2022) – ENSO hat beispielsweise um 1940 eine ihrer ausgeprägtesten und längsten Wärmephasen seit 1880, was mit der Beobachtung der hier verwendeten Daten zusammenpasst.

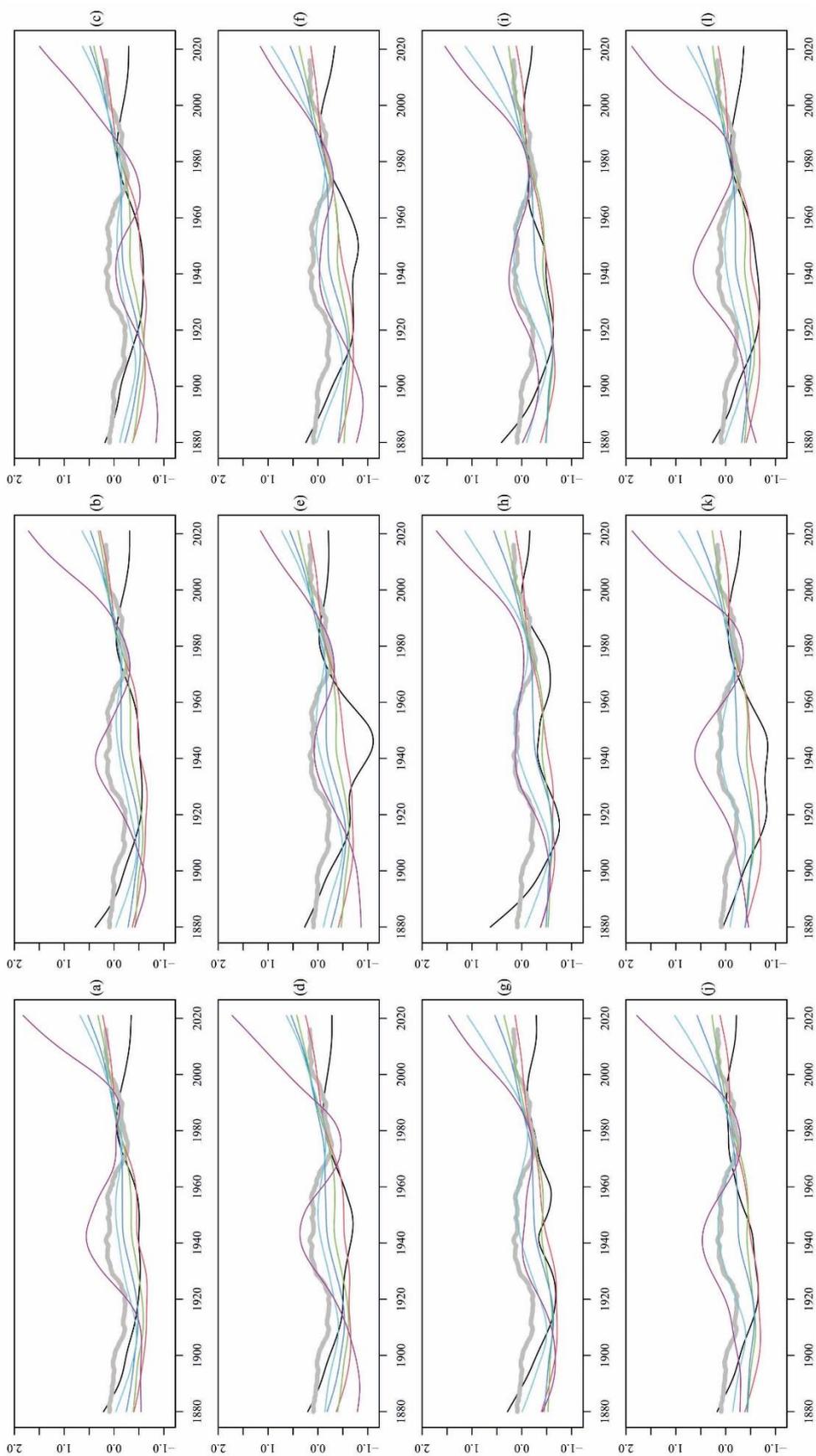


Abbildung 10. Anomalien der Meeresoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda=2500$ für (a) Januar, (b) Februar, ..., (l) Dezember. Die Linien stellen die geogr. Zonen dar. AMO-Zeitreihe in hellgrau zum Vergleich

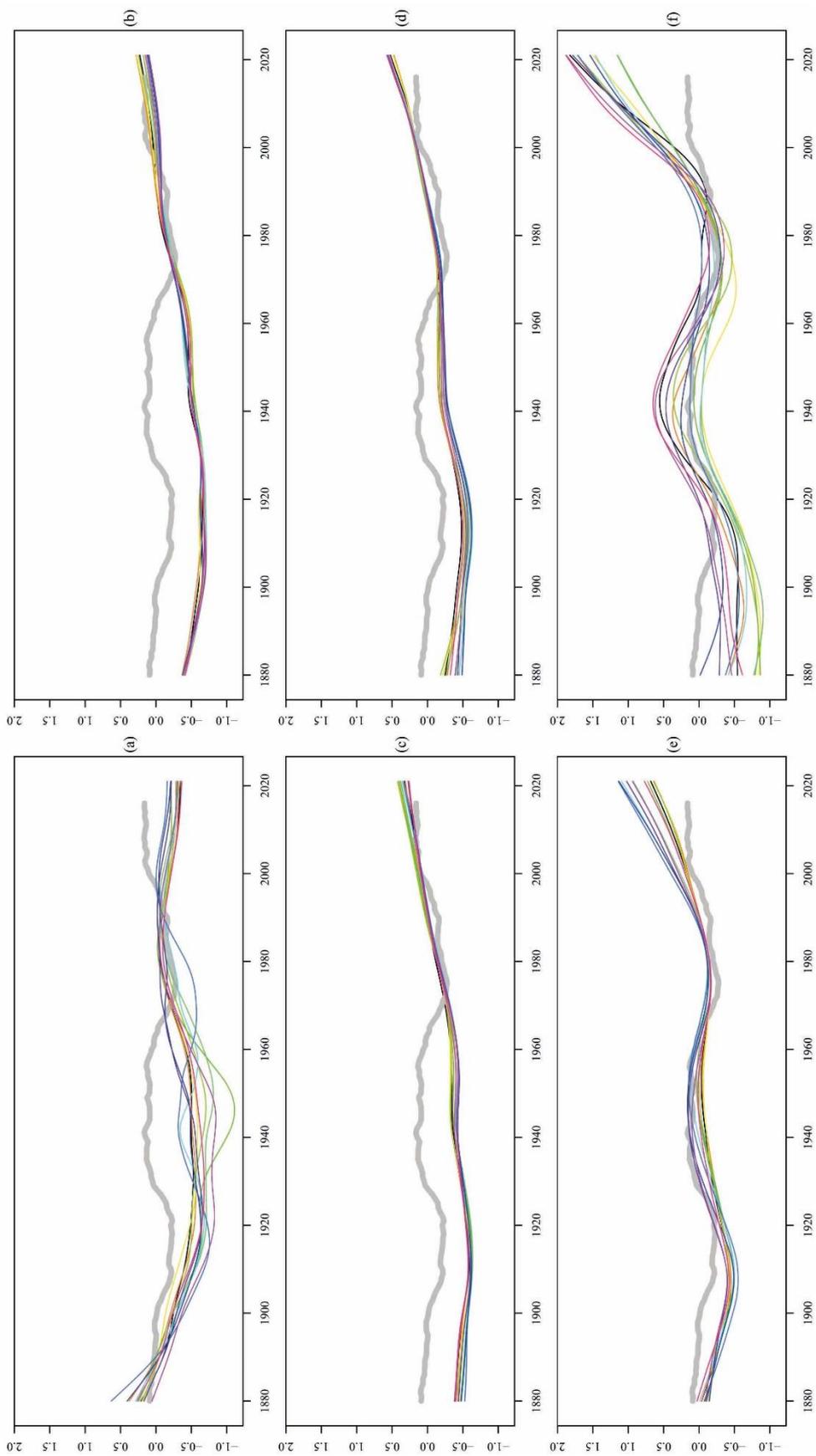


Abbildung 11. Anomalien der Meeresoberflächentemperatur nach Anwendung des Hodrick-Prescott-Filters mit $\lambda=2500$ für (a) 90°S-60°S, (b) 60°S-30°S, ..., (f) 60°N-90°N. Die Linien stellen die 12 Monate dar. AMO-Zeitreihe in hellgrau zum Vergleich

3.2 Imputation fehlender Werte

In Kapitel 2 wurde bereits erwähnt, dass die verwendeten Daten nicht ganz vollständig sind und die Werte für insgesamt 18 Monate in der Zeitreihe der Landestemperaturen im Bereich 90°S-60°S imputiert werden mussten, bevor eine Analyse möglich war. Für einen Monat, bei dem die Aufzeichnung fehlte, wurden die Daten desselben Monats in den darauffolgenden Jahren heran gezogen und der Mittelwert daraus gebildet. Dieser Mittelwert ist dann als neuer Wert eingesetzt worden. Angefangen wurde bei dem chronologisch gesehen spätesten fehlenden Wert (Oktober 1902) bis an den Anfang der Zeitreihe zum frühesten fehlenden Wert (September 1880).

Diese Vorgehensweise wurde gewählt, weil nicht in jedem Jahr das jeweils vorangegangene sowie das folgende Jahr zur Berechnung des Mittelwertes verwendet werden konnte, da es vorkommt, dass die Daten für einen bestimmten Monat bei zwei Jahren in Folge fehlt. Durch das Beginnen beim spätesten Fall und das Verwenden der beiden Folgejahre wurde dieses Problem jedoch umgangen und die Zeitreihe konnte komplettiert werden.

3.3 Testen auf Stationarität

Primäres Ziel dieser Arbeit ist es herauszufinden, ob es einen signifikanten Anstieg der Temperaturen auf der Erde seit 1880 gibt. Da also überprüft werden soll, ob die entsprechenden Zeitreihen einen gewissen Trend aufweisen, liegt es nahe, jene Zeitreihen auf Stationarität zu testen. Wenn man zeigen kann, dass sie nicht stationär sind, kann daraus geschlossen werden, dass die durchschnittlichen Lufttemperaturen auf unserem Planeten in den vergangenen Jahrzehnten tatsächlich gestiegen sind.

Die Analyse der Grafiken aus Kapitel 2 legt die Vermutung nahe, dass hier nicht-stationäre Zeitreihen vorliegen. Diese Vermutung gilt es nun mithilfe von statistischen Methoden zu bestätigen. Es gibt diverse Möglichkeiten eine Zeitreihe auf Stationarität zu testen, der Ansatz, der in dieser Ausarbeitung verfolgt wurde, erfordert es einen Blick auf die Spektraldichte von stochastischen Prozessen zu werfen. Bekanntlich gilt bei stationären Prozessen, dass das Integral

$$\int_0^{\varepsilon} f(\omega) d\omega$$

(4)

existiert, wobei $f(\omega)$ die Spektraldichte des Prozesses ist. In weiter Folge wird eine Methode beschrieben, wie die Hypothese, dass das Integral wohldefiniert ist, verworfen werden kann. Zunächst muss das klassische ARFIMA(p,d,q) Modell betrachtet werden, welches durch unten stehende Gleichung beschrieben werden kann.

$$x_t = \left(1 - \sum_{j=1}^p \varphi_j L^j \right)^{-1} (1 - L)^{-d} \left(1 + \sum_{j=1}^q \theta_j L^j \right) u_t$$

(5)

Hier sind die $(x_t)_{t=1,\dots,T}$ die Beobachtungen der Zeitreihe, die φ_j sind die AR(p)-Koeffizienten und die θ_j sind die MA(q)-Koeffizienten. Bei d handelt es sich um den oben angesprochenen fraktionalen Differenzenparameter. Relevant ist vor allem die Spektraldichte dieses Prozesses, die folgendermaßen definiert ist.

$$f(\omega) = \frac{\sigma^2}{2\pi} |1 - e^{-i\omega}|^{-2d} \left| \frac{1 + \sum_{j=1}^q \theta_j e^{-i\omega j}}{1 - \sum_{j=1}^p \varphi_j e^{-i\omega j}} \right|^2 \quad (6)$$

Betrachtet man die Spektraldichte nahe der Frequenz 0, lässt sich Ausdruck (6) vereinfacht ausdrücken als

$$f(\omega) \sim \frac{\sigma^2}{2\pi} |1 - e^{-i\omega}|^{-2d} \quad (7)$$

Des Weiteren lässt sich zeigen, dass

$$|1 - e^{-i\omega}|^2 = 4 \cdot \sin^2\left(\frac{\omega}{2}\right) \quad (8)$$

wodurch sich Formel (7) schreiben lässt als

$$f(\omega) \sim C \cdot 4^{-d} \cdot \sin^{-2d}\left(\frac{\omega}{2}\right) \quad (9)$$

Hier wurde der Ausdruck $\frac{\sigma^2}{2\pi}$ als Konstante C abgekürzt. Nachdem die Spektraldichte nun ausreichend vereinfacht wurde, kann der Logarithmus davon genommen werden, um den Parameter d auszudrücken.

$$\log f(\omega) \sim \log C \cdot 4^{-d} + d \left(-2 \log \sin\left(\frac{\omega}{2}\right) \right) \quad (10)$$

Nun ist man bereits in der Lage einen Schätzer für d zu konstruieren, und zwar durch Verwendung einer einfachen linearen Regression. Das entsprechende Modell ist in (11) dargestellt.

$$y_t = d_0 + d \cdot \left(-2 \log \sin\left(\frac{\omega_t}{2}\right) \right) + u_t \quad (11)$$

Bei $(y_t)_{t=1,\dots,k}$ handelt es sich um die Werte des log-Periodogramms an den ersten k Fourier-Frequenzen $(\omega_t)_{t=1,\dots,k}$, wobei k frei gewählt werden kann.

Das Schätzen von d ist essenziell, da der Parameter einen direkten Einfluss auf das Integral aus (4) hat. Setzt man den Ausdruck (9) in das Integral ein, so kann man es wie folgt schreiben

$$\int_0^\varepsilon f(\omega) d\omega \sim C \cdot 4^{-d} \cdot \int_0^\varepsilon \sin^{-2d} \left(\frac{\omega}{2} \right) d\omega \quad (12)$$

Da hier die ersten Fourier-Frequenzen betrachtet werden, also der Bereich in der Nähe von Frequenz 0, gilt, dass $\sin(\omega) \sim \omega$ und folglich Ausdruck (12) weitergehend vereinfacht werden kann.

$$\int_0^\varepsilon f(\omega) d\omega \sim C \cdot 4^{-d} \cdot \int_0^\varepsilon \left(\frac{\omega}{2} \right)^{-2d} d\omega \quad (13)$$

Für das Konstruieren eines Tests auf Stationarität fehlt nur noch der Blick auf eine letzte Eigenschaft der Fourier-Frequenzen: es gilt, dass das Integral

$$\int_0^\varepsilon \omega^{-2d} d\omega \quad (14)$$

nur dann existiert, wenn $d < \frac{1}{2}$. Demnach existiert das Integral der Spektraldichte nur dann, wenn der fraktionale Differenzenparameter kleiner als 0.5 ist. Mit anderen Worten formuliert heißt das, dass bei einer stationären Zeitreihe gelten muss, dass $d < \frac{1}{2}$ und dass ein Schätzer \hat{d} , der einen Wert größer als $\frac{1}{2}$ besitzt gegen Stationarität spricht.

In weiterer Folge wird für alle in dieser Arbeit betrachteten Zeitreihen der Differenzenparameter geschätzt und überprüft, ob sich dieser Schätzwert oberhalb oder unterhalb des Wertes 0.5 befindet. Können mehrheitlich große Werte von \hat{d} festgestellt werden, so wäre das ein klares Indiz für das Nichtvorhandensein von Stationarität und somit für einen relevanten Anstieg der Temperaturen.

Abbildungen 12 bis 15 zeigen die Ergebnisse der Schätzungen. Die Vermutungen, die in Kapitel 2 aufgestellt wurden, werden hier bestätigt. Mit Ausnahme der Zeitreihen, die die Temperaturen in der Nähe der Pole enthalten (s. Abbildung 14 und Abbildung 15, jeweils (a) & (f)), sind fast ausschließlich Schätzwerte über 0.5 vorhanden, die meisten davon sogar eindeutig. Und selbst in den ganz südlichen und den ganz nördlichen Zonen befindet sich das \hat{d} bei manchen Monaten für alle untersuchten Fourier-Frequenzen über der 0.5-Grenze, am stärksten zu sehen bei den Meeresoberflächentemperaturen in Bereich 60°N-90°N (Abbildung 15 (f)).

Um sich der anderen Frage dieser Ausarbeitung zuzuwenden, nämlich wie unterschiedlich der Temperaturverlauf zwischen den verschiedenen Monaten ist, muss ein Blick auf Abbildungen 12 und 13 geworfen werden. Große Differenzen erkennt man allerdings nicht unter den zwölf Grafiken, lediglich die in den vorherigen Abbildungen ebenfalls schon beobachteten, stärker variierenden Linien in schwarz (90°S - 60°S) und lila (60°N - 90°N). Erstere liegt bei den Meeresoberflächentemperaturen in den Monaten Juni bis Oktober (Abbildung 13 (6) - (10)) deutlicher unterhalb des Werts 0.5 als in den anderen Monaten, was dafür spricht, dass in den - in dieser Region - kalten Monaten der Temperaturanstieg weniger gut bzw. gar nicht nachgewiesen werden kann. Andere offensichtliche Diskrepanzen sind aber tatsächlich nicht feststellbar, es scheint also nicht der Fall zu sein, dass die globale Erderwärmung in bestimmten Monaten stärker ausgeprägt ist, als in anderen.

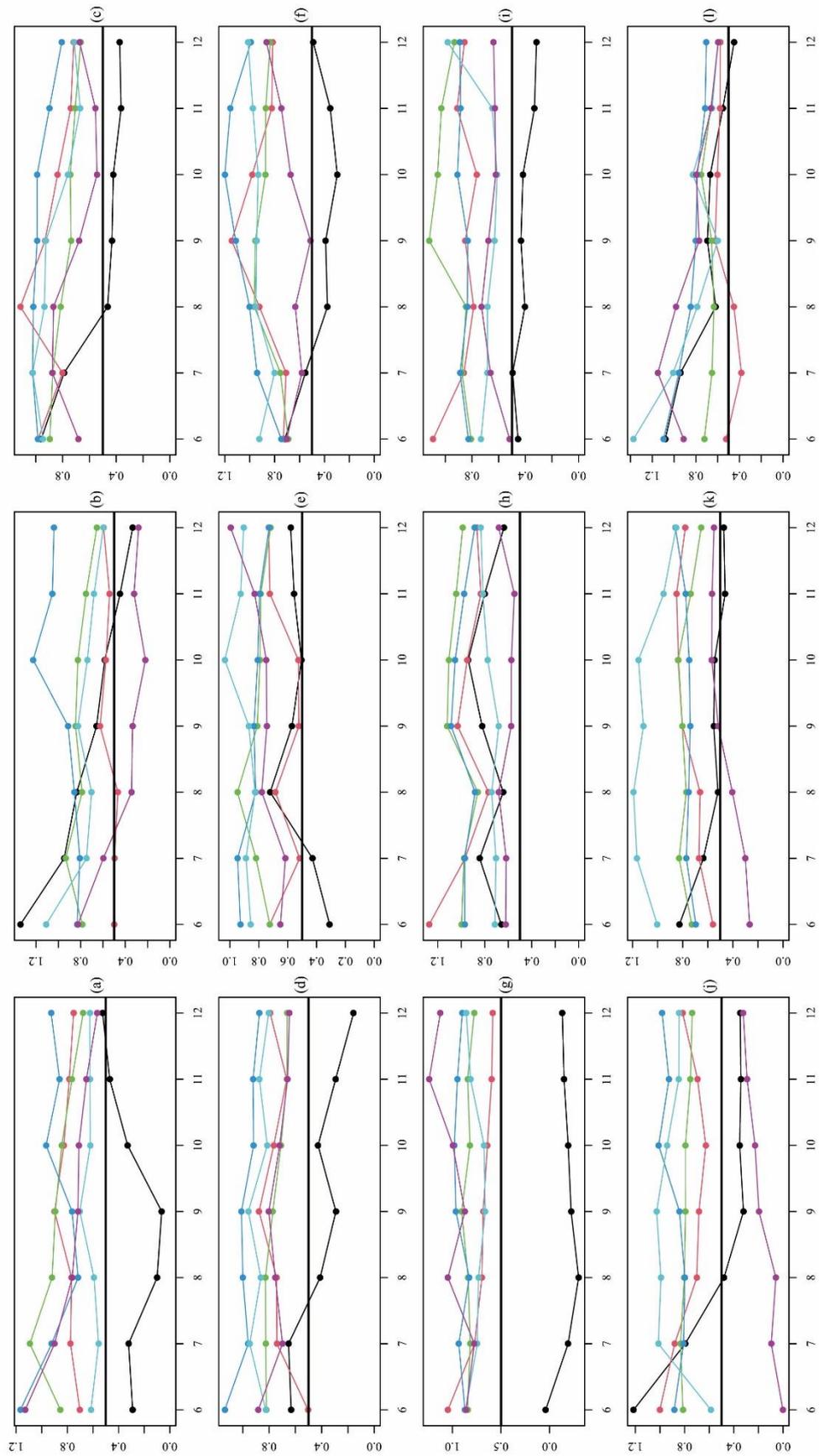


Abbildung 12. Schätzwerte von d für die Landoberflächentemperaturen bei Betrachtung der ersten 6 - 12 Fourier-Frequenzen für (a) Januar, (b) Februar, ..., (l) Dezember.

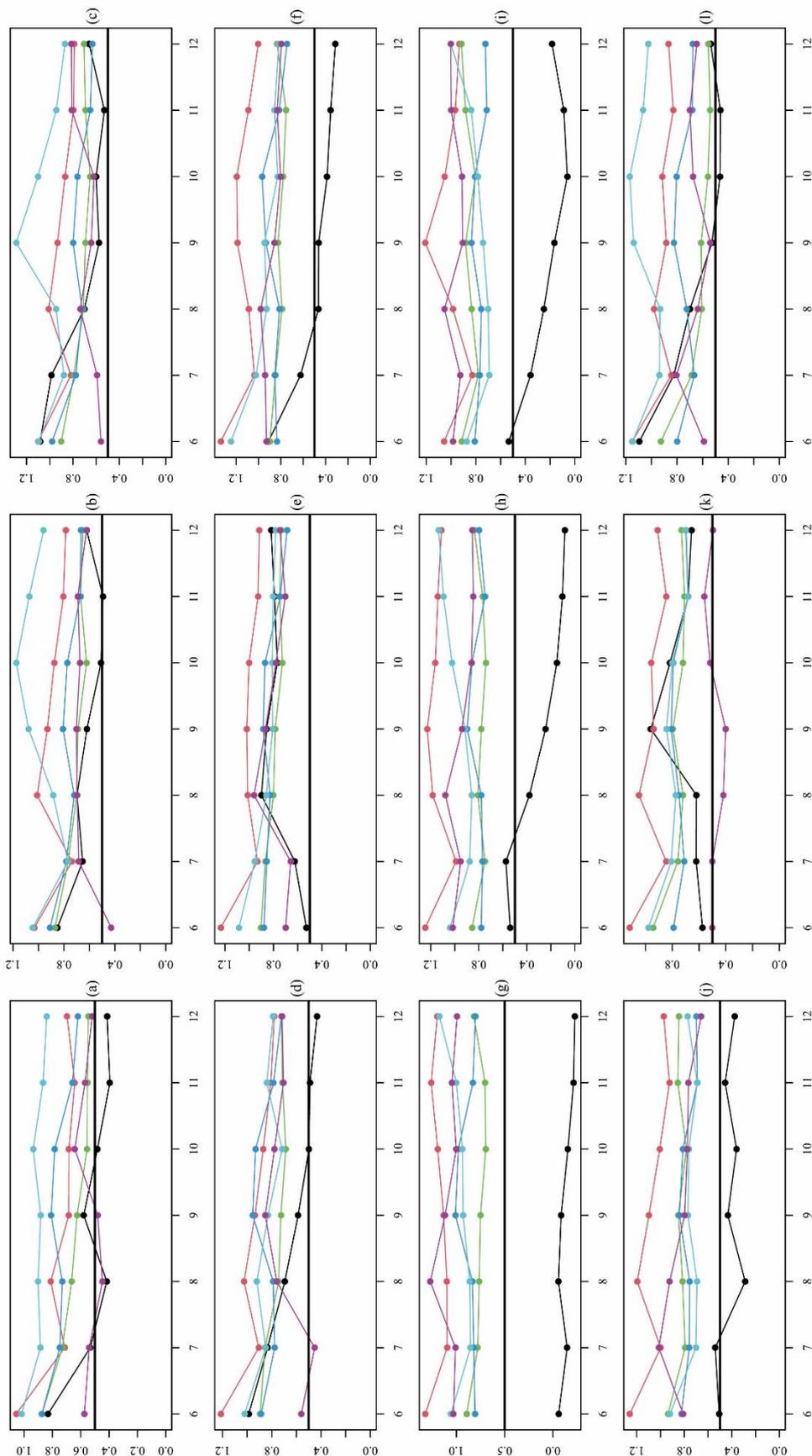


Abbildung 13. Schätzwerte von d für die Meeresoberflächentemperaturen bei Betrachtung der ersten 6 - 12 Fourier-Frequenzen für (a) Januar, (b) Februar, ..., (l) Dezember.

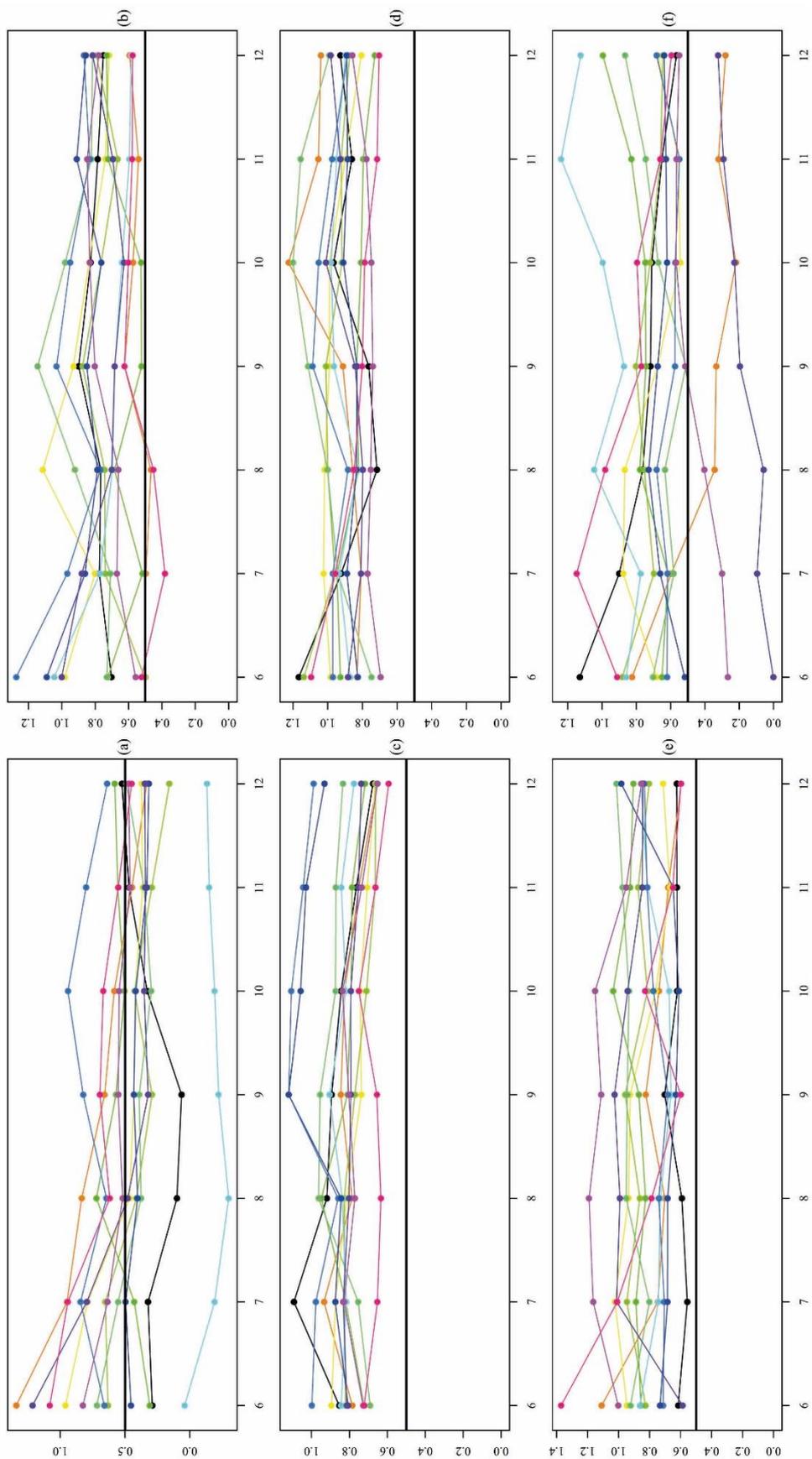


Abbildung 14. Schätzwerte von d für die Landoberflächentemperaturen bei Betrachtung der ersten 6 - 12 Fourier-Frequenzen für (a) 90°S-60°S, (b) 60°S-30°S, ..., (f) 60°N-90°N.

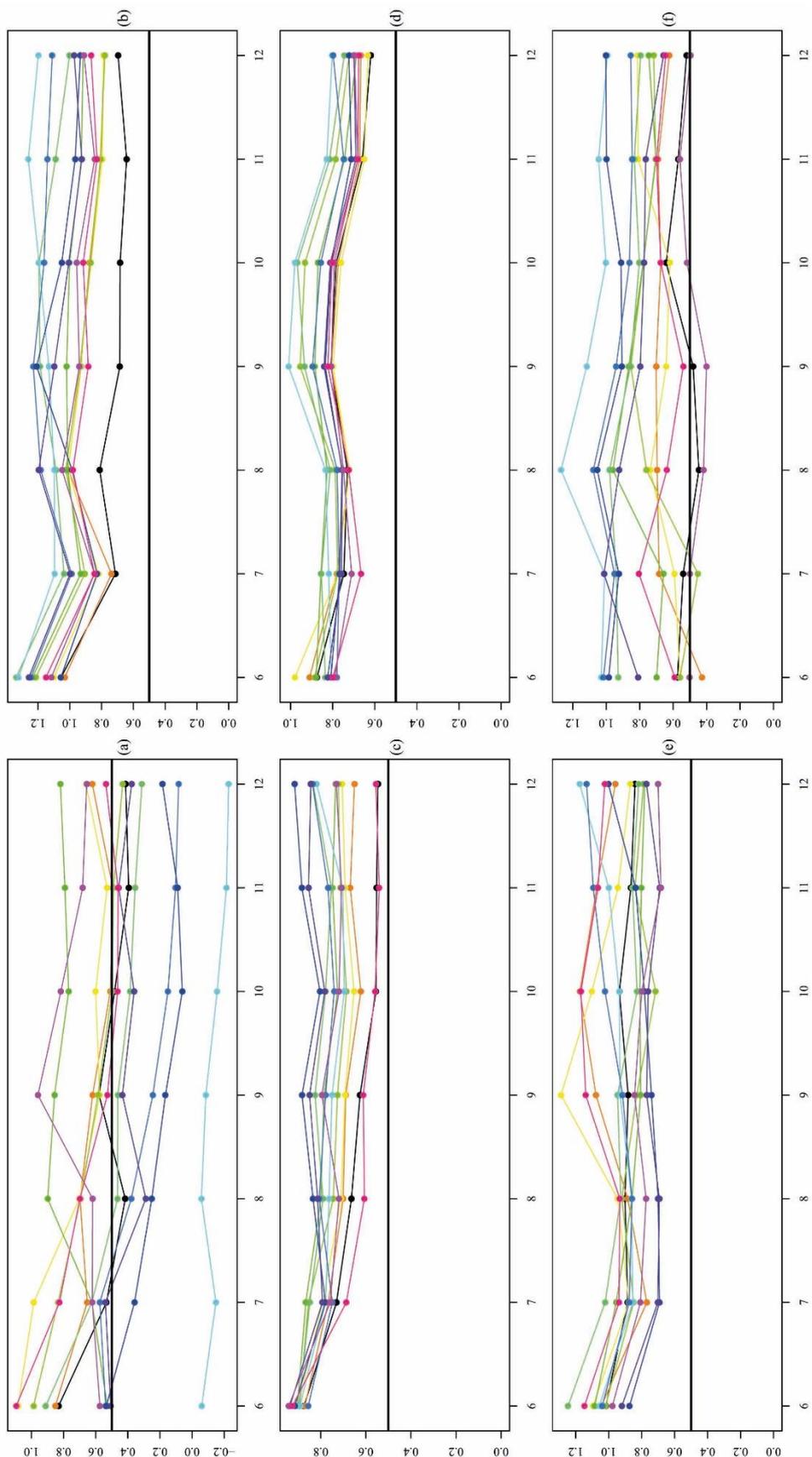


Abbildung 15. Schätzwerte von d für die Meeresoberflächentemperaturen bei Betrachtung der ersten 6 - 12 Fourier-Frequenzen für (a) $90^{\circ}\text{S}-60^{\circ}\text{S}$, (b) $60^{\circ}\text{S}-30^{\circ}\text{S}$, ..., (f) $60^{\circ}\text{N}-90^{\circ}\text{N}$.

4 Fazit

Das Ziel dieser Arbeit war es den Klimawandel durch Anwendung von statistischen Methoden nachzuweisen. Für diesen Zweck wurden von der NOAA zur Verfügung gestellte Zeitreihen herangezogen, die die monatlichen Durchschnittstemperaturen auf der Erde seit 1880 beinhalten. Es wurden sechs verschiedene Zeitreihen betrachtet, jede repräsentativ für eine bestimmte Zone unseres Planeten. Diese Zonen sind definiert durch die Breitengrade, beginnend bei 90°S (Südpol) und in 30°-Schritten bis zu 90°N gehend. Es sind die Temperaturanomalien (basierend auf den Jahren 1971 – 2000) für die Monate und Regionen getrennt betrachtet worden, um mögliche Unterschiede zwischen ihnen zu erkennen.

Zunächst wurden die Zeitreihen grafisch als klassisches Liniendiagramm dargestellt, um einen ersten Überblick über die Daten zu erhalten. Da die Kurven, wie für Temperaturdaten üblich, viele kleine Schwankungen aufweisen, mussten sie erst noch geglättet werden, damit ein zumindest grober Trend ausgemacht werden konnte. Als Methode wurde der Hodrick-Prescott-Filter gewählt. Das hatte zur Folge, dass selbst mit freiem Auge ein leicht steigender Trend in dem Großteil der Zeitreihen zu sehen war, wodurch bereits die Vermutung aufgestellt werden konnte, dass es in den vergangenen Jahren tatsächlich zu einer Erhöhung der durchschnittlichen Temperaturen gekommen ist. Es wurden auch andere auffällige Verhalten der Zeitreihen erkannt, wie etwa die auf den ersten Blick genau entgegen der AMO wirkenden, jahrzehntlangen Oszillation in der antarktischen Region, die im Gegensatz zur AMO selbst in der Literatur nicht sehr gut dokumentiert zu sein scheint.

Schließlich wurden die eben erwähnten, steigenden Trends in den Zeitreihen auch noch mittels statistischer Verfahren überprüft. Zu Beginn musste eine kleine Menge an fehlenden Werten imputiert werden um mit der eigentlichen Analyse starten zu können. Aufgrund der geringen Anzahl und wegen der Tatsache, dass sich diese fehlenden Werte nur in der Zeitreihe für die Region 90°S-60°S und zwar gleich am Anfang in den Jahren 1880 – 1902 befunden haben, wurde eine simple Methode verwendet. Es wurde jeweils der Mittelwert der Temperaturen des entsprechenden Monats in den beiden folgenden Jahren eingesetzt. Daraufhin ist für jede Zeitreihe der fraktionale Differenzenparameter d geschätzt worden, da anhand dieses Schätzwertes eine Aussage über die Stationarität der jeweiligen Zeitreihe getroffen werden kann. Wie in Abschnitt 3.3 beschrieben, muss eine stationäre Zeitreihe ein d mit Wert 0.5 oder kleiner besitzen, da ansonsten eine zentrale Annahme über die Spektraldichte der Zeitreihe verletzt wäre.

Die Ergebnisse fielen größtenteils sehr eindeutig aus. Selbst unter Berücksichtigung mehrerer verschiedener Fourier-Frequenzen lagen die Schätzwerte sowohl für die Land- als auch für die Meeresoberflächentemperaturen in den Bereichen 30°S-0°N, 0°N-30°N und 30°N-60°N allesamt über der 0.5-Grenze und im Bereich 60°S-30°S war das auch bis auf sehr wenige Ausnahmen der Fall. Bei den Meeresoberflächentemperaturen in der nördlichsten Zone (60°N-90°N) konnten auch größtenteils Schätzer, die größer oder zumindest nur knapp kleiner als 0.5 sind, berechnet werden. Ansonsten waren die Ergebnisse an den Polen um einiges weniger eindeutig zugunsten der Nichtstationarität.

Daher lässt sich feststellen, dass seit 1880 in der Region 60°S-60°N unserer Erde ein deutlicher Temperaturanstieg sowohl am Land, als auch über dem Meer stattgefunden

hat. In den Regionen südlicher und nördlicher davon lässt sich das anhand dieser Analyse nicht mit Sicherheit sagen, wenngleich die Tendenz auch hier in Richtung Nichtstationarität zeigt. Des Weiteren konnten keine merkbaren Unterschiede zwischen den Temperaturverläufen in den einzelnen Monaten festgestellt werden. In jedem davon lässt sich der Klimawandel relativ eindeutig untermauern, ohne dass es in manchen zu jeweils mehr oder weniger eindeutigen Ergebnissen gekommen ist.

Literatur

- Dickey, David A., and Wayne A. Fuller. 1979. *Distribution of the estimators for autoregressive time series with a unit root*. Journal of the American Statistical Association 74: 427–31.
- Ejaz, Tariq, Walihur Rahaman, C. M. Laluraj, Kanthanathan Mahalinganathan, and Meloth Thamban. 2022. *Rapid Warming Over East Antarctica Since the 1940s Caused by Increasing Influence of El Niño Southern Oscillation and Southern Annular Mode*. Frontiers in Earth Science 10: 799613.
- Estrada, Francisco, Luis Filipe Martins, and Pierre Perron. 2017. *Characterizing and attributing the warming trend in sea and land surface temperatures*. *Atmósfera* 30: 163–87.
- Estrada, Francisco, Pierre Perron, and Benjamín Martínez-López. 2013. *Statistically derived contributions of diverse human influences to twentieth-century temperature changes*. Nature Geoscience 6: 1050–55.
- Fomby, Thomas B., and Timothy J. Vogelsang. 2002. *The application of size-robust trend statistics to global-warming temperature series*. Journal of Climate 15: 117–23.
- Grenander, Ulf, and Murray Rosenblatt. 1957. *Statistical Analysis of Stationary Time Series*. New York: John Wiley and Sons.
- Hodrick, Robert, and Edward C. Prescott. 1997. *Postwar U.S. Business Cycles: An Empirical Investigation*. Journal of Money, Credit, and Banking 29 (1): 1–16.
- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. *Testing the null hypothesis of stationarity against the alternative of a unit root*. Journal of Econometrics 54: 159–78.
- Phillips, Peter C. B., and Pierre Perron. 1988. *Testing for a unit root in time series regression*. Biometrika 75: 335–46.
- Schmidt, Peter, and Peter C. B. Phillips. 1992. *Lm tests for a unit root in the presence of deterministic trends*. Oxford Bulletin of Economics and Statistics 54: 257–87.
- Vogelsang, Timothy J. 1998. *Trend function hypothesis testing in the presence of serial correlation*. Econometrica 66: 123–48.