Homogeneity Testing of Sunspot Numbers

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Thomas K. Friedli

von Ursenbach BE

Leiter der Arbeit:

Prof. Dr. J. Hüsler

Institut für mathematische Statistik und Versicherungslehre

Homogeneity Testing of Sunspot Numbers

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Thomas K. Friedli

von Ursenbach BE

Leiter der Arbeit:

Prof. Dr. J. Hüsler

Institut für mathematische Statistik und Versicherungslehre

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Der Dekan:

Bern, den 16. Juni 2005

Prof. Dr. P. Messerli

Inhaltsverzeichnis

1 Einleitung **1**
- 1.1 Problemstellung . . . 2
- 1.2 Zielsetzung . . . 5
- 1.3 Vorgehen . . . 6
- 1.4 Inhaltsübersicht . . . 8
- 1.5 Verdankungen . . . 8

2 Strukturelle Zustandsraummodelle **9**
- 2.1 Lineare Gauss'sche Zustandsraummodelle . . . 9
- 2.2 Darstellung struktureller Modellkomponenten . . . 10
  - 2.2.1 Lokal linearer Trend . . . 10
  - 2.2.2 Stochastischer Zyklus . . . 11
  - 2.2.3 Erklärende Variablen . . . 12
  - 2.2.4 Interventionen . . . 12
- 2.3 Modellschätzung . . . 13
  - 2.3.1 Kalman Filter . . . 13
  - 2.3.2 Kalman Smoother . . . 13
  - 2.3.3 Disturbance Smoother . . . 14
  - 2.3.4 Fehlende Werte . . . 14
  - 2.3.5 Parameterbestimmung . . . 15
  - 2.3.6 Initialisierung . . . 15
- 2.4 Modellvalidierung . . . 15

3 Darstellung der Sonnenaktivität **17**
- 3.1 Der Anblick der Sonnenoberfläche . . . 17
- 3.2 Quantifizierung der Sonnenaktivität . . . 19
- 3.3 Klassische Darstellungsverfahren . . . 21
- 3.4 Zustandsraummodell für die Monatsmittel der Wolfschen Reihe . . . 23
- 3.5 Vergleich mit klassischen Darstellungsverfahren . . . 26
- 3.6 Zustandsraummodell für die Jahresmittel der Wolfschen Reihe . . . 29

4 Homogenitätstests **33**
- 4.1 Das Change - Point Problem . . . 33
- 4.2 Likelihood Ratio Testgrössen . . . 34
- 4.3 Der Alexandersson Test . . . 39
  - 4.3.1 Definition der Testgrösse . . . 39

4.3.2 Der Alexandersson Test als Likelihood Ratio Test . . . 40
4.4 Der Buishand Test . . . 41
4.4.1 Definition der Testgrösse . . . 41
4.4.2 Der Buishand Test als Likelihood Ratio Test . . . 43
4.5 Worsley's Likelihood Ratio Test . . . 43
4.6 Kritische Grenzen . . . 44
4.6.1 Bonferroni Approximation . . . 44
4.6.2 Asymptotische Verteilung . . . 45
4.6.3 Simulation . . . 46
4.6.4 Vergleich . . . 54
4.6.5 Empirische Approximation . . . 59
4.6.6 Buishand Test . . . 66
4.7 Verhalten unter H_0 . . . 69
4.7.1 Autokorrelation . . . 70
4.7.2 Buishand Test . . . 75
4.8 Verhalten unter H_1 . . . 78
4.8.1 Güte . . . 78
4.8.2 Schätzung von Bruchepoche und Sprunghöhe . . . 79
4.8.3 Autokorrelation . . . 79
4.8.4 Buishand Test . . . 83
4.9 Diskussion . . . 83

5 Verbesserte Homogenitätstests 87
5.1 Modifizierte Teststatistiken . . . 87
5.2 Kritische Grenzen . . . 92
5.3 Güte . . . 92
5.4 Sprunghöhen und Bruchepochen . . . 92
5.5 Diskussion . . . 94

6 Homogenitätsprüfung 97
6.1 Schockmodell . . . 97
6.2 Einfache Versatzinhomogenitäten als Schockmodell . . . 98
6.3 Likelihood Ratio Testgrössen . . . 98
6.4 Modifizierte Likelihood Ratio Testgrössen . . . 100
6.5 Diskussion . . . 103

7 Homogenität der Wolfschen Reihe 105
7.1 Vergleichsreihen . . . 105
7.2 Die kurzfristige Homogenität zwischen 1945 und 1995 . . . 106
7.3 Die langfristige Homogenität zwischen 1700 und 2005 . . . 116

8 Schlussfolgerungen und Ausblick 121

A Zufallszahlengeneratoren 123
A.1 Ox-Programmcode . . . 123

B Testreihe Z80 133

C Monte - Carlo Experimente **135**
C.1 Likelihood Ratio Test 1 136
C.2 Likelihood Ratio Test 2 150
C.3 Likelihood Ratio Test 3 154
C.4 Likelihood Ratio Test 4 158
C.5 Likelihood Ratio Test 5 166
C.6 Likelihood Ratio Test 6 168
C.7 Likelihood Ratio Test 7 170
C.8 Alexandersson Test 172
C.9 Buishand Test 175
C.10 Likelihood Ratio Test 8 184
C.11 Likelihood Ratio Test 9 192
C.12 Likelihood Ratio Test 10 200

Literaturverzeichnis **203**

Curriculum vitae **215**

Johann Rudolf Wolf (1816-1893). Gründer und erster Direktor der Eidg. Sternwarte in Zürich.

1. Einleitung

Im Winter 1843 entdeckte der Apotheker Heinrich Samuel Schwabe (1789 - 1875) aus Dessau bei der Durchsicht seiner seit 1826 beinahe täglich durchgeführten Sonnenbeobachtungen, eine rund zehnjährige Periode in der Häufigkeit der Sonnenflecken (Schwabe 1844). Zwar widersprach dies der damals gängigen Lehrmeinung, wonach die Zahl der Sonnenflecken sowie die Häufigkeit ihres Auftretens gänzlich regellos seien, doch blieb eine Reaktion der Fachwelt weitgehend aus, wohl weil ein sicherer Nachweis einer periodischen Schwankung der Sonnenaktivität mit dem damals vorliegenden knappen Beobachtungsmaterial nicht zu leisten war. – Erst als 1852 der englische Generalmajor Edward Sabine (1788 - 1883), der Genfer Astronom Alfred Gautier (1793 - 1881) und der Berner Astronom und Wissenschaftshistoriker Johann Rudolf Wolf (1816 - 1893) die Parallelität zwischen dem Schwabeschen Sonnenfleckenzyklus und der Variation der Magnetnadel entdeckten (Wolf 1857, 1890) und darin einen Kausalzusammenhang zwischen Sonnenflecken und Erdmagnetismus vermuteten, war das Echo unter den Gelehrten auf einmal gross: Schwankt die Sonnenfleckenaktivität tatsächlich periodisch und kann der Zusammenhang mit den magnetischen Variationen über einen längeren Zeitraum hinweg gesichert werden (Friedli und Keller 1993)?

Während sich Sabine und Gautier in der Beantwortung dieser Frage nicht weiter engagierten, legte Wolf noch im selben Jahr eine positive Antwort vor (Wolf 1852), doch stützte er sich hierbei nur auf einige wenige, sicher bestimmte Epochen der Sonnenfleckenaktivität und der erdmagnetischen Variationen. – Erst ein Vierteljahrhundert später konnte Wolf eine zuverlässige Darstellung des Verlaufs der Sonnenfleckenaktivität publizieren (Wolf 1877): Aus alten Aufzeichnungen bestimmte er einen von ihm selbst entwickelten Sonnenaktivitätsindex – die sog. Sonnenfleckenrelativzahl R

$$R = k\,(10\,g + f)$$

worin g die Anzahl Sonnenfleckengruppen, f die Anzahl Einzelflecken in g und k ein von Beobachter und Instrument abhängiger Reduktionsfaktor darstellen – welcher nun ab 1849 für alle Tage, ab 1749 für alle Monate und ab 1700 für alle Jahre vorliegt: die Wolfsche Reihe (Abb. 1.1).

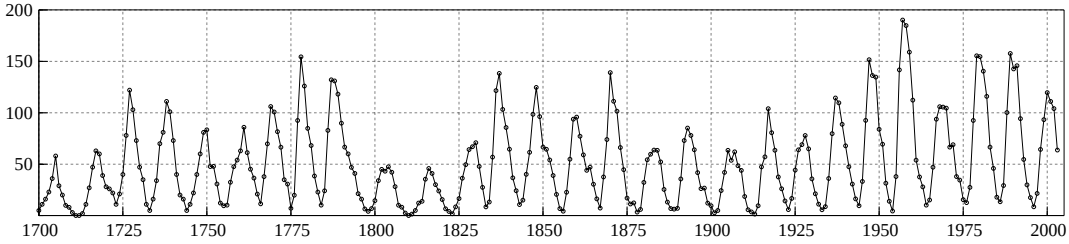


Abbildung 1.1: Die Jahresmittel der Wolfschen Sonnenfleckenrelativzahlen R_I 1700-2003

Wolf beschränkte sich jedoch nicht auf das Sammeln und Auswerten alter Aufzeichnungen, vielmehr beobachtete er selber 47 Jahre lang möglichst täglich die Sonne. 1863 gründete er die Eidgenössische Sternwarte in Zürich (Waldmeier 1963) und ermunterte zahlreiche befreundete Berufs- und Amateurastronomen aus aller Welt, ihm regelmässig ihre Fleckenzählungen mitzuteilen. So entstand ein internationales Netz erfahrener Sonnenbeobachter mit deren Hilfe die Sonnenaktivität beginnend mit 1849 lückenlos für jeden Tag dokumentiert werden konnte, wobei alle Beobachtungen auf Wolfs eigene Zählweise an einem 80 mm Fraunhofer Refraktor mit 1100 mm Brennweite und 64-facher Vergrösserung – dem sogenannten Wolfschen Normalrefraktor (Abb. 1.2) – reduziert wurden (Waldmeier 1961).

Die täglichen Sonnenfleckenbeobachtungen am Normalinstrument wurden von Wolfs Nachfolgern bis zur Schliessung der Eidgenössischen Sternwarte im Jahre 1980 kontinuierlich weitergeführt und die Auffassungs- und Zähltechnik sorgfältig von Generation zu Generation weitergegeben (Waldmeier 1948). Seit 1981 ist das Observatoire Royal de Belgique in Uccle für die Berechnung der internationalen Wolfschen Sonnenfleckenrelativzahlen verantwortlich, wobei die Konstanz des Skalenfaktors dadurch gesichert werden soll, dass statt den Zählungen am Wolfschen Refraktor die Beobachtungen an der ehemaligen "Specola Solare dell'Osservatorio Astronomico Federale" in Locarno-Monti als Normalbeobachtungen verwendet werden (Berghmans et al. 2002). Die Sonnenfleckenbeobachtungen am Wolfschen Normalrefraktor konnten jedoch auch nach der Schliessung der Eidgenössischen Sternwarte unterbruchslos fortgesetzt werden: Bis 1995 führte H.U. Keller im Auftrag des Bundesamtes für Übermittlungstruppen (BAUEM) die Bestimmung der Zürcher Sonnenfleckenrelativzahl am originalen Fraunhoferrefraktor fort (Keller und Friedli 1995); 1996 musste das historische Instrument allerdings infolge der Gesamtrenovation der Sternwarte (Friedli et al. 1998) von deren Dachterrasse entfernt werden und wird seither von mir zur regelmässigen Bestimmung der Wolfschen Sonnenfleckenrelativzahl eingesetzt (Friedli 1997).

1.1 Problemstellung

Die Bedeutung der Wolfschen Reihe ist im Laufe der Generationen stetig gewachsen: Wurde sie zu Beginn vorwiegend zur Analyse der Sonnenaktivität sowie zur Erforschung der solar-terestrischen Beziehungen verwendet, entstand in der Mitte des letzten Jahrhunderts nach der Entdeckung des solaren Einflusses auf die Qualität der Kurzwellenausbreitung zusätzlich das praktische Bedürfnis, die Sonnenaktivität zumindest über einen Zeitraum von wenigen Monaten hinweg zuverlässig vorhersagen zu können, wozu die Wolfsche Reihe die längste und sicherste Datengrundlage zu liefern schien. So wurden denn von Astronomen, Statistikern und Informatikern Verfahren zur Beschreibung und Prognose der Sonnenaktivität entwickelt (Hathaway et al. 1999) und mit den im Laufe der Jahre gestiegenen Bedürfnissen zu einer umfassenden kurz- und mittelfristigen Raumwettervorhersage ausgebaut (McIntosh und Dryer 1972, Hruska et al. 1994, Williamson et al. 2000).

Im Zuge der Erforschung und Nutzung des erdnahen Weltraumes ist zudem die immer besser begründete Gewissheit entstanden, dass die Sonnenaktivität nicht nur den interplanetaren Raum, die Magnetosphäre und die Hochatmosphäre, sondern auch die Stratosphäre und die Troposphäre beeinflusst und somit direkte Auswirkungen auf das irdische Wetter und Klima hat (Pepin et al. 1980, Schove 1983, Sonett et al. 1991, Hargreaves 1992, Friis-Christensen et al. 2000). Insbesondere unter Klimatologen werden die möglichen Ursachen des beobachteten raschen Klimawandels kontrovers diskutiert: Wie wirken sich die diversen anthropogenen und natürlichen Einflussfaktoren auf das heutige und das zukünftige Klima aus (Houghton et al. 2001, Rind 2002)?

Abbildung 1.2: Der von Rudolf Wolf 1855 in Betrieb genommene Fraunhofersche 80 / 1100 mm Refraktor mit 64-facher Vergrösserung auf der südlichen Beobachtungsterrasse der Eidg. Sternwarte in Zürich. Am Tubusende ist das Merzsche Polarisationshelioskop zur stufenlosen Dämpfung des Sonnenlichts zu erkennen, welches ein gefahrloses Betrachten der Sonnenphotosphäre ermöglicht. Aufnahme zwischen 1931 und 1956.

Die Wolfsche Reihe spielt in dieser Diskussion eine Schlüsselrolle (Hoyt und Schatten 1997, Benestad 2002). Leider sind die täglichen Sonnenfleckenrelativzahlen vor 1818 und damit der Verlauf der Sonnenaktivität während klimatologisch so bedeutungsvoller Perioden wie dem Maunder Minimum (1645-1715) und dem Dalton Minimum (1795-1825) weitgehend unbekannt. Dies veranlasste mehrere Autorengruppen in den letzten 10 Jahren systematisch nach bisher unbekannten Sonnenfleckenbeobachtungen zu suchen (Hoyt und Schatten 1994) und die Wolfsche Reihe über 1700 hinaus bis auf 1610 zurück zu verlängern (Abb. 1.3).

Die Qualität dieser Rekonstuktion ist jedoch zweifelhaft, da sowohl die Tauglichkeit des Reduktionsverfahrens wie die langfristige Homogenität der berechneten Reihe unbekannt sind. Rudolf Wolf optimierte seine Rekonstruktionstechnik auf eine ununterbrochene Abfolge von langjährigen Standardbeobachtern deren wichtigster er selber war. Vor 1818 waren für Wolf keine solchen bekannt und demzufolge sank die Sicherheit der beobachteten Monatsmittel der Sonnenfleckenrelativzahlen rasch (Wolf 1880). Dank den Anstrengungen der letzten 10 Jahre kennen wir nun für die kritischen Jahre zwischen 1610 und 1849 zwar viele neue Einzelbeobachtungen, aber keine neuen Standardbeobachter, welche sich an die bekannten anschliessen liessen.

Hoyt und Schatten (1998) änderten deshalb kurzerhand das Wolfsche Auswerteverfahren, indem sie vom Konzept der Standardbeobachter abrückten und die rekonstruierte Reihe als skaliertes, gewichtetes Mittel der vorhandenen Einzelbeobachtungen modellierten. Es ist jedoch seit längerem bekannt, dass Änderungen in der Berechnungsweise der Sonnenfleckenrelativzahlen zu systematischen Abweichungen führen können, was anhand der seit Dezember 1944 vorliegenden amerikanischen Sonnenfleckenrelativzahlen R_A leicht gezeigt werden kann (Abb. 1.4), resultierten doch schon nach wenigen Jahren grosse Differenzen gegenüber der Wolfschen Reihe (Shapley 1949), worauf die Amerikaner ihr Auswerteverfahren neu skalierten mussten (Taylor 1985). Trotzdem traten auch in den folgenden Jahren immer wieder systematische Abweichungen auf (Ruf 1977).

Es erscheint mir daher offensichtlich, dass es momentan für eine zuverlässige Rekonstruktion der Sonnenfleckenaktivität bis ins Entdeckungsjahr 1610 zurück primär nicht an zusätzlichen Beobachtungen fehlt, sondern vielmehr an einer statistisch validen Auswertungsmethode, welche es erlaubt, die mehr oder weniger kurzen Serien von Einzelbeobachtungen zu einer gemeinsamen, langfristig homogenen Reihe von täglichen Sonnenfleckenrelativzahlen zusammenzufassen.

Da die der Wolfschen Reihe zugrunde liegenden Originalbeobachtungen gegenwärtig elektronisch weitgehend unzugänglich sind, ja einige Teile gar als verschollen gelten (Friedli 2001), liegt eine statistische Neubearbeitung der Wolfschen Reihe in weiter Ferne.

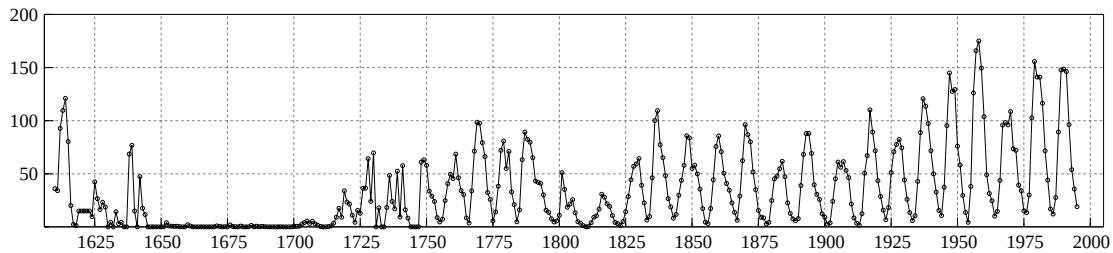


Abbildung 1.3: Die Jahresmittel der gruppierten Sonnenfleckenzahlen R_g 1610-1995

1.2 Zielsetzung

Gegenstand der vorliegenden Arbeit ist die Homogenitätsprüfung ausgewählter Teile der Wolfschen Reihe, insbesondere der Periode zwischen 1945 und 2003 mit dem auf den 1. Januar 1981 erfolgten Wechsel des Standardbeobachters. Hierzu wird ein statistisches Modell entwickelt, welches für die Darstellung der Sonnenfleckenaktivität geeignet ist und gleichzeitig eine valide Homogenitätsprüfung der rekonstruierten Beobachtungsreihe ermöglicht.

Damit wird ein Instrumentarium geschaffen, welche es erlaubt, auch andere – beispielsweise klimatologisch relevante – langjährige Beobachtungs- und Messreihen in verhältnismässig einfacher Weise auf ihre Homogenität zu prüfen (Friedli und Schuepbach 2001).

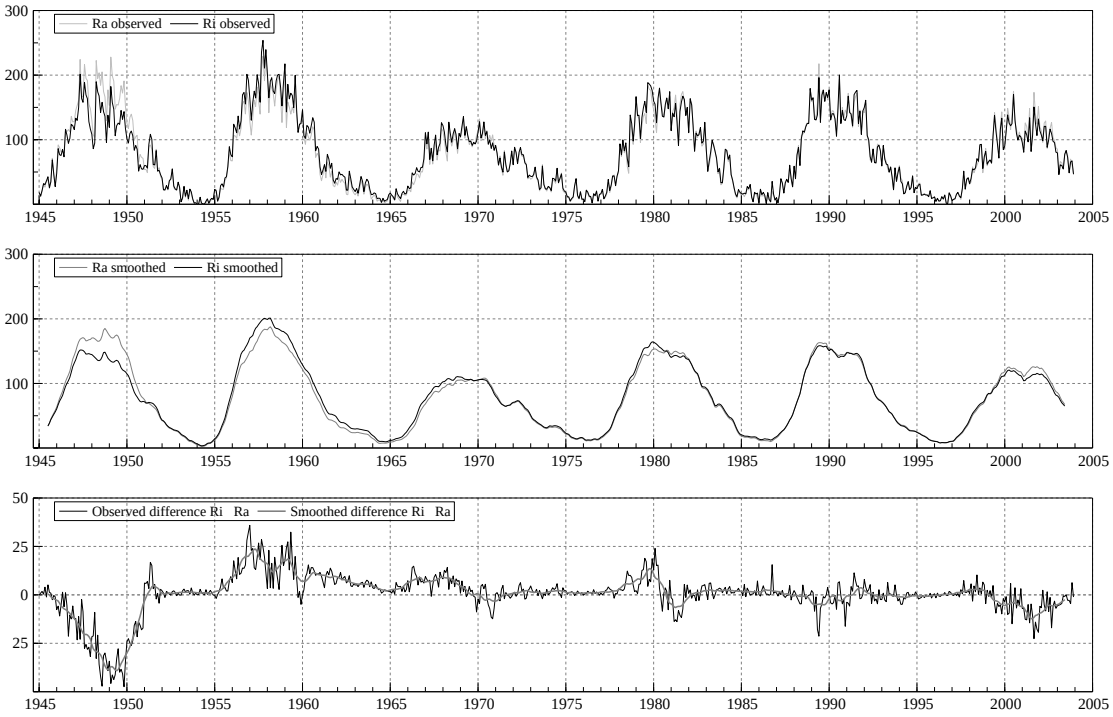


Abbildung 1.4: Verläufe der beobachteten und der ausgeglichenen Monatsmittel der internationalen Sonnenfleckenrelativzahlen R_I und der amerikanischen Sonnenfleckenrelativzahlen R_A zwischen 1945 und 2003. Beide Datensätze beruhen auf den Zählungen von erfahrenen Sonnenbeobachtern und sind vollständig. Sie unterscheiden sich lediglich in der Zusammensetzung der Beobachter und Instrumente sowie in der Art und Weise wie sie reduziert wurden. Daher wird erwartet, dass zwischen den beiden Reihen keine nennenswerten Unterschiede bestehen. – Wie in der untersten Grafik gezeigt, traten aber schon in den ersten Jahren grosse systematische Differenzen auf, welche auch nach der Reskalierung des amerikanischen Auswerteverfahrens nicht ganz eliminiert werden konnten, sich überdies langfristig selbst zu verändern scheinen und gegenwärtig beinahe ins Gegenteil umgeschlagen sind.

1.3 Vorgehen

Die Homogenitätsprüfung langer Beobachtungsreihen hat insbesondere in der Klimatologie eine lange Tradition, da schon früh erkannt wurde, dass wissenschaftlich relevante Resultate nur aus denjenigen Beobachtungsreihen gewonnen werden können, deren Messungen unbeeinflusst von möglichen Verzerrungen durch Standort, Messinstrument und Auswertung allein auf den Einflüssen von Wetter und Klima beruhen (Conrad und Pollak 1950).

Bedingt durch die in der zweiten Hälfte des letzten Jahrhunderts immer lauter geführten Diskussionen um mögliche Klimaveränderungen wurden die zahlreich vorhandenen klimatologischen Messreihen – allen voran die Temperatur- und Niederschlagsmessreihen – im Laufe der Jahre immer peinlicheren Qualitätskontrollen unterzogen, um allfällige Veränderungen der klimatischen Verhältnisse von mess- und verarbeitungstechnischen Artefakten sowie von lokalklimatischen Eigenheiten unterscheiden zu können (WMO 1986, Herzog und Müller-Westermeier 1998). Einen repräsentativen Überblick über die von öffentlichen Wetterämtern routinemässig eingesetzten Homogenisierungsverfahren geben Peterson et al. (1998). Die im Rahmen des Nationalen Forschungsprogramms "Klimaänderungen und Naturkatastrophen" (NFP 31) auf die schweizerischen klimatologischen Messreihen des 19. und 20. Jahrhunderts angewandten Homogenisierungsverfahren sind in Bosshard (1996), Baudenbacher (1997) und Gisler et al. (1997) näher beschrieben.

Die erfolgreichsten dieser Homogenisierungsverfahren beruhen auf der Differenzenbildung mit benachbarten, klimatologisch ähnlichen Messstationen und deren Prüfung auf Zufälligkeit anhand eines geeigneten statistischen Homogenitätstests, wie etwa dem Standard Normal Homogeneity Test (SNHT) nach Alexandersson (1986) oder dem Test nach Buishand (1982). Für die langjährigen, quasistationären Temperatur- und Niederschlagsmessreihen funktioniert dies auf den ersten Blick ganz gut. Offensichtliche Mängel in diesem Verfahren treten erst dann zu Tage, wenn relativ kurze, mit fehlenden Werten durchsetzte und nichtstationäre Messreihen homogenisiert werden sollen. Zanis et al. (1999) beispielsweise untersuchten die Homogenität der Ozonmessreihe auf der hochalpinen Messstation JFJ auf dem Jungfraujoch in einer Höhe von 3580 Meter über Meer (Abb. 1.5). Die Differenzenreihe mit der "benachbarten" Messstation ZUG auf der Zugspitze erwies sich nach Prüfung mit dem Alexandersson Test und dem Buishand Test als statistisch inhomogen.

Gegen eine derartige Interpretation spricht, dass sich die beobachteten Differenzen zwischen JFJ und ZUG auch durch reale Unterschiede in den grossräumigen Zirkulationsverhältnissen der Atmosphäre erklären lassen, während sich instrumentell bedingte Inhomogenitäten vorwiegend durch sehr kurzfristige, quasi sprunghafte Veränderungen in den Differenzen ausweisen würden. Schuepbach und Friedli (1998) wiesen denn auch darauf hin, dass statistische Homogenitätstests nicht unbesehen auf jede x-beliebige Differenzenreihe angewendet werden dürfen, dass vielmehr für jede Beobachtungsreihe ein geeignetes – gegebenenfalls nichtstationäres – Modell angepasst werden muss, welches dann im Rahmen einer kritischen Modelldiagnose einer statistischen Homogenitätsprüfung unterzogen werden kann. Das in der Folge von Schuepbach und Friedli (1998) an die Ozonmessungen von JFJ angepasste ARIMAX Modell erwies sich allerdings als wenig praxisgerecht, da zum einen in ARIMAX Modellen die Identifizierung möglicher Inhomogenitäten nur mit grosser Erfahrung, wenn überhaupt, möglich ist und zum anderen die Klimatologen mehr an den Verläufen einzelner Komponenten, wie dem langfristigen Trend oder dem Jahresgang, interessiert sind. Friedli und Schüpbach (2001) schlugen deshalb vor, für die Darstellung klimatologischer Messreihen strukturelle Zustandsraummodelle einzusetzen. Die Homogenitätsprüfung kann so nicht nur auf die

beobachtete Reihe als Ganzes, sondern auf jede Komponente einzeln angewandt werden. Schuepbach et al. (2001) schätzten daraufhin für die beiden Messstationen Jungfraujoch und Zugspitze passende strukturelle Zustandsraummodelle und zeigten, dass die in Zanis et al. (1999) entdeckten Inhomogenitäten durch einen leicht unterschiedlichen Mittelwertsverlauf der beiden Messreihen vorgetäuscht werden.

In der Folge wurde dieses zustandsraumgestützte Homogenisierungsverfahren im Rahmen des Forschungsprojekts TROTREP zur Exploration und Qualitätskontrolle von 52 europäischen Ozonmessstationen eingesetzt (Monks et al. 2003) und floss auch in zwei Machbarkeitsstudien zur Datenvervollständigung und Homogenisierung der Schadstoffreihen des Nationalen Beobachtungsnetzes für Luftfremdstoffe (NABEL) ein (Thöni et al. 2001, Krähenbühl et al. 2004).

Während somit die klassische Homogenisierungsstrategie erfolgreich verbessert und die rechnerische Homogenisierung mittels Schätzung geeigneter erklärender Variabler in Zustandsraummodellen auf eine solide statistische Basis gestellt werden konnte, blieb die eigentliche Homogenitäts*prüfung*, also die Identifizierung statistisch signifikanter Inhomogenitätsstellen, auf dem Stand der qualitativen Beurteilung von Modellresiduen stehen. Diese Lücke soll in der vorliegenden Arbeit geschlossen werden, indem statistisch valide Tests vorgestellt und entwickelt werden, welche in Zustandsraummodellen im Rahmen der Modellvalidierung eine sichere Identifizierung von Inhomogenitätsstellen erlauben und somit erstmals auch eine statistische Homogenitätsprüfung der Wolfschen Reihe ermöglichen.

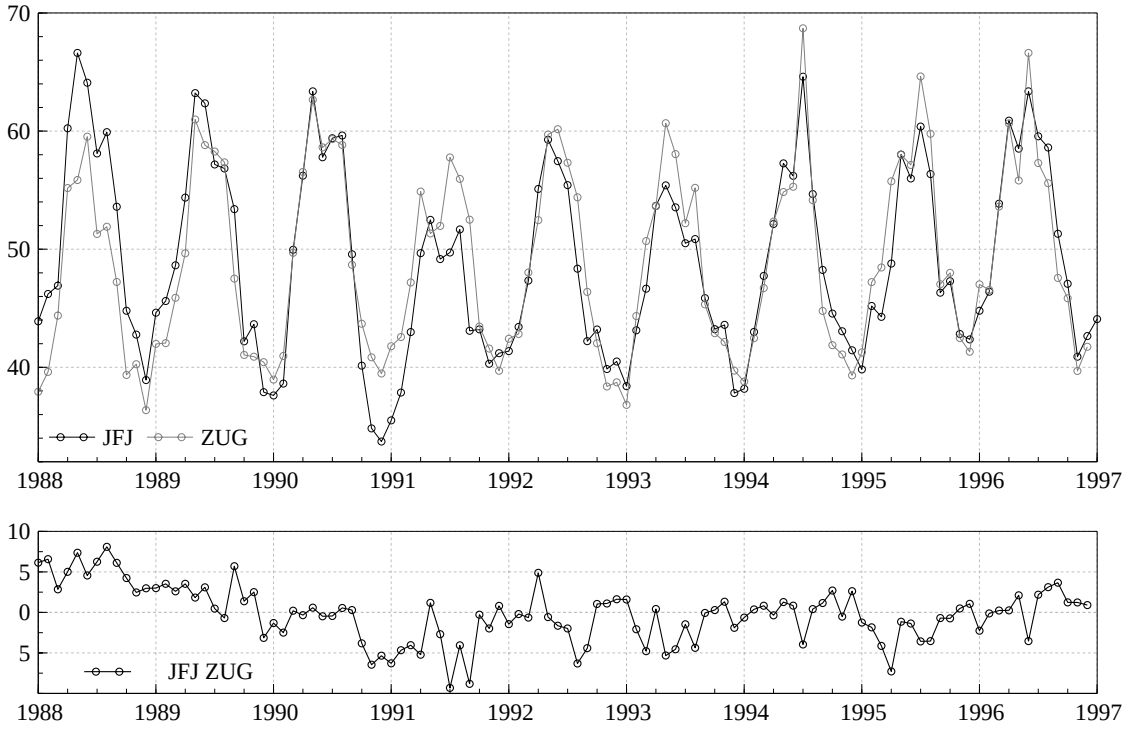


Abbildung 1.5: Monatsmittel der Ozonkonzentration (in ppbv) auf den Messstationen Jungfraujoch (JFJ) und Zugspitze (ZUG) zwischen 1988 und 1996 (oben) sowie deren Differenzenreihe (unten).

1.4 Inhaltsübersicht

Im nächsten Kapitel werden strukturelle Zustandsraummodelle vorgestellt. In Kapitel 3 wird ein für die Homogenitätsprüfung der Sonnenaktivität geeignetes Zustandsraummodell entwickelt und mit den bisher üblichen Darstellungsformen verglichen. In Kapitel 4 wird die Tauglichkeit der gängigsten klimatologischen Homogenitätstests geprüft. Einige der gefundenen Schwächen werden in Kapitel 5 korrigiert. In Kapitel 6 wird gezeigt, wie die korrigierten Tests zur Validierung von Zustandsraummodellen eingesetzt werden können. In Kapitel 7 wird die Homogenität einzelner Teile der Wolfschen Reihe geprüft. Die Diskussion der Ergebnisse erfolgt in Kapitel 8. Der Anhang enthält neben den Resultaten der Monte Carlo Experimente zur Validierung der Homogenitätstests auch das Literaturverzeichnis, sowie meinen Lebenslauf.

1.5 Verdankungen

Meinem Doktorvater Prof. Dr. Jürg Hüsler vom Institut für mathematische Statistik und Versicherungslehre (IMSV) der Universität Bern danke ich für seine vorbehaltlose Unterstützung und seine verständnisvolle Geduld.

Herrn Prof. Dr. Heinz Wanner vom Geographischen Institut der Universität Bern (GIUB) danke ich für die Übernahme des Koreferats.

Frau Prof. Dr. Evi Schüpbach früher Leiterin der Forschungsgruppe Climate and Background Ozone (CABO) des GIUB, jetzt Direktorin des Institut Universitaire Kurt Bösch (IUKB) in Sitten danke ich für die fruchtbare Zusammenarbeit auf dem Gebiet der statistischen Analyse klimatologischer Zeitreihen sowie ihren stets gewährten freundschaftlichen Rat.

Herrn H.U. Keller, während 20 Jahren Sonnenbeobachter an der ehem. Eidgenössischen Sternwarte in Zürich, danke ich für die erschöpfende Vermittlung der Geheimnisse um die Zürcher Schule der Sonnenfleckenrelativzahlbestimmung. Dr. Beat Glaus, ehem. Leiter der Wissenschaftshistorischen Sammlungen der ETH Bibliothek in Zürich, danke ich für die unbürokratische Öffnung der Sternwartenarchive sowie sein engagiertes Mitwirken bei der Rettung der Wolfschen Originalinstrumente zur Sonnenfleckenrelativzahlbestimmung. Frau Gaby Aebersold danke ich für die elektronische Erfassung einiger der historischen Zürcher Sonnenfleckenbeobachtungen.

Frau Prof. Dr. Karin Kovar von der Abteilung Biotechnologie der Hochschule Wädenswil sowie den Assistenten David Langenegger und Markus Keller danke ich für ihr wohlwollendes Verständnis, welches sie während der Schlussphase dieser Arbeit den zumindest inneren Absenzen ihres Statistikers entgegenbrachten.

Meinen Kollegen Manuela Schaller, Dr. Daniel Dietrich und PD Dr. Enkelejd Hashorva danke ich für jahrelange freundschaftliche Zusammenarbeit sowie “meinen” Diplomanden Manuela Schüttel, Mike Mitter und Niklaus Steiner für das tatkräftige Mitwirken bei der statistischen Analyse klimatologischer Zeitreihen.

Ohne die Hilfe meiner Familie wäre diese Arbeit nicht zu Stande gekommen. Ich danke meinen Eltern, meinen Schwiegereltern, meinen beiden Töchtern sowie insbesondere meiner Frau Maria für ihre Anteilnahme und ihre Geduld.

Diese Arbeit wurde vom Schweizerischen Nationalfonds unterstützt.

2. Strukturelle Zustandsraummodelle

Zustandsraummodelle bieten eine vereinheitlichte Sichtweise einer sehr breiten Klasse von Zeitreihenmodellen. Beginnend mit den zukunftsweisenden Arbeiten von Kalman (1960) sowie von Kalman und Bucy (1961) fanden Zustandsraummodelle in den Naturwissenschaften und in der Industrie rasch breite Anwendung, etwa in der Bahnverfolgung künstlicher Erdsatelliten oder in der chemischen Prozesskontrolle, während ihr statistisches Potential erst in den 1980er Jahren erkannt wurde. Heute zählen Zustandsraummodelle zu den Standardmethoden der Zeitreihenanalyse. Eine Einführung in die Möglichkeiten der Zustandsraummodellierung findet sich u.a. in Shumway und Stoffer (2000), Brockwell und Davis (2003) sowie in Chatfield (2004). Eine ausführliche Diskussion bieten Jazwinski (1970), Anderson und Moore (1979), Young (1984), Harvey (1989), Kitagawa und Gersch (1996), West und Harrison (1997) sowie Fahrmeir und Tutz (2001).

Im vorliegenden Kapitel werden univariate lineare strukturelle Zustandsraummodelle vorgestellt. In den ersten Abschnitten wird der Aufbau eines linearen Zustandsraummodells sowie die rekursive Formulierung ihrer strukturellen Komponenten besprochen, wobei wir uns auf diejenigen Komponenten beschränken, welche zur Modellierung der Wolfschen Reihe relevant sind. In den nachfolgenden Abschnitten wird die Schätzung und Validierung von linearen Gauss'schen Zustandsraummodellen gezeigt. Die numerischen Berechnungen wurden teils mit dem kommerziell vertriebenen Softwarepacket STAMP 6.2 von Koopman, Harvey, Doornik und Shephard (2000) ausgeführt, teils mit Programmen, welche mit der C-Library SsfPack 2.3 von Koopman et al. (1999) im Rahmen der Diplomarbeiten von Schüttel (2002) und Mitter (2004) in der objekt-orientierten Matrizensprache Ox von Doornik (1998) geschrieben wurden, teils mit dem im Rahmen der Diplomarbeit von Steiner (2004) korrigierten S-Plus Modul S+Finmetrics/SsfPack. Wir halten uns im Folgenden weitgehend an die Notation von Durbin und Koopman (2001) sowie von Zivot und Wang (2003).

2.1 Lineare Gauss'sche Zustandsraummodelle

In einem univariaten linearen Gauss'schen Zustandsraummodell wird die Messung y_t zum Zeitpunkt t mit folgender **Beobachtungsgleichung** (observation equation) modelliert:

$$y_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + G_t \epsilon_t. \tag{2.1}$$

worin y_t die beobachtete univariate Zeitreihe, $\boldsymbol{\alpha}_t$ den unbeobachteten Zustandsvektor sowie ϵ_t den stochastischen Beobachtungsfehlerprozess bezeichnen. Diese Gleichung erinnert an ein klassisches lineares Regressionsmodell mit dem Zustandsvektor $\boldsymbol{\alpha}_t$ als dem Vektor der Regressionskoeffizienten mit den bekannten Designmatrizen $\mathbf{Z}_t$ und G_t. Im Unterschied zu einem linearen Regressionsmodell darf sich in einem Zustandsraummodell der Vektor der Regressionskoeffizienten allerdings zeitlich verändern. Diese Änderung des Zustandes $\boldsymbol{\alpha}_t$, d.h. die Dynamik des Systems, wird in der

Zustandsgleichung (state oder transition equation) beschrieben:

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T}_t \boldsymbol{\alpha}_t + \mathbf{H}_t \boldsymbol{\eta}_t. \tag{2.2}$$

Für die beiden Fehlerprozesse ϵ_t und $\boldsymbol{\eta}_t$ gilt

$$\begin{bmatrix} \epsilon_t \\ \boldsymbol{\eta}_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \end{bmatrix} \right) \tag{2.3}$$

worin $\mathbf{I}_r$ eine Identitätsmatrix mit Rang r darstellt. Für den Anfangszustand $\boldsymbol{\alpha}_1$ gilt

$$\boldsymbol{\alpha}_1 \sim N\left(\mathbf{a}_1, \mathbf{P}_1\right). \tag{2.4}$$

Sind die Matrizen $\mathbf{Z}_t, \mathbf{T}_t, G_t$ und $\mathbf{H}_t$ alle zeitunabhängig (also $\mathbf{Z}_t = \mathbf{Z}, \mathbf{T}_t = \mathbf{T}$ usw. für alle t), spricht man von einem zeitinvarianten Modell. Üblicherweise enthalten zudem einige oder alle der Matrizen $\mathbf{Z}_t, \mathbf{T}_t, G_t$ und $\mathbf{H}_t$ unbekannte Parameter, welche aus den Beobachtungen y_t geschätzt werden müssen.

2.2 Darstellung struktureller Modellkomponenten

In einem strukturellen Komponentenmodell wird der Zustandsvektor als Summe einzelner Teilmodelle dargestellt, welche gewöhnlich eine leicht interpretierbare fachwissenschaftliche Bedeutung besitzen. So besteht das von Harvey (1989) in extenso diskutierte elementare strukturelle Modell aus einer stochastischen Trendkomponente μ_t, einer stochastischen Saisonkomponente γ_t sowie einer stochastischen Fehlerkomponente ϵ_t:

$$y_t = \mu_t + \gamma_t + \epsilon_t. \tag{2.5}$$

Im Folgenden werden nur diejenigen strukturellen Komponenten besprochen, welche aus meiner Sicht für die Modellierung der Wolfschen Reihe relevant sind. Dies sind lokal lineare Trends, stochastische Zyklen, erklärende Variablen und Interventionen. Ausführlichere Beschreibungen struktureller Modellkomponenten finden sich in Harvey und Shephard (1993), Harvey und Koopman (1996), Koopman et al. (1999), Koopman, Harvey, Doornik und Shephard (2000), Durbin und Koopman (2001), Koopman et al. (2002) sowie in Zivot und Wang (2003).

2.2.1 Lokal linearer Trend

Ein lokal lineares Trendmodell besteht aus einem lokalen stochastischen Niveau μ_t und einer lokalen stochastischen Steigung ν_t und kann rekursiv geschrieben werden als:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t & \xi_t &\sim N(0, \sigma_\xi^2) \\ \nu_{t+1} &= \nu_t + \zeta_t & \zeta_t &\sim N(0, \sigma_\zeta^2) \end{aligned} \tag{2.6}$$

Eingesetzt in (2.1) und (2.2) ergibt sich mit $\boldsymbol{\alpha}_t = [\mu_t, \nu_t]'$ folgendes Zustandsraummodell:

$$\begin{aligned} y_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \sigma_\epsilon \end{bmatrix} \varepsilon_t \\ \boldsymbol{\alpha}_{t+1} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \sigma_\xi & 0 \\ 0 & \sigma_\zeta \end{bmatrix} \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}. \end{aligned} \tag{2.7}$$

Trotz dieser einfachen Definition ist das Modell (2.7) sehr flexibel und enthält einige bekannte Trendmodelle als Spezialfälle (Tabelle 2.1), so etwa den smooth Trend, welcher als nichtparametrischer Kernschätzer mit rekursiv geschätztem, lokal veränderlichem Kern aufgefasst werden kann (Harvey und Koopman 2000, Brown und de Jong 2001).

Modell	σ_ϵ	σ_ξ	σ_ζ
Constant level	$*$	0	-
Local level	$*$	$*$	-
Random walk	0	$*$	-
Global linear	$*$	0	0
Local level mit konstanter Steigung	$*$	$*$	0
Random walk mit konstantem Drift	0	$*$	0
Smooth trend	$*$	0	$*$
Lokal linear	$*$	$*$	$*$
ARIMA(0,2,0)	0	0	$*$

Tabelle 2.1: Einige bekannte Spezialfälle des lokal linearen Trendmodells. Ein $*$ steht für einen beliebigen positiven Wert. Ein $-$ bedeutet, dass die Steigungskomponente in Modell (2.6) fehlt.

2.2.2 Stochastischer Zyklus

Ein stochastischer Zyklus ψ_t kann rekursiv als Superposition einer Sinus- und Cosinusfunktion sowie zweier voneinander unabhängiger stochastischer Fehlerprozesse formuliert werden:

$$\begin{aligned} y_t &= \psi_t + \varepsilon_t \\ \psi_{t+1} &= \rho(\psi_t \cos\lambda_c + \psi_t^* \sin\lambda_c) + \chi_t \\ \psi_{t+1}^* &= \rho(-\psi_t \sin\lambda_c + \psi_t^* \cos\lambda_c) + \chi_t^* \end{aligned} \tag{2.8}$$

wobei ρ ein Dämpfungsfaktor mit $0 < \rho \leq 1$ und λ_c die in Radians gemessene Frequenz mit $0 < \lambda_c < \pi$ darstellen. Die Periode c des Zyklus ist $2\pi/\lambda_c$. Für die Fehlerprozesse χ_t und χ_t^* gilt:

$$\begin{bmatrix} \chi_t \\ \chi_t^* \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_\psi^2(1-\rho^2)\mathbf{I}_2 \right).$$

Ist $\rho = 1$ wird der Zyklus deterministisch. Das Zustandsraummodell lautet mit $\boldsymbol{\alpha}_t = [\psi_t, \psi_t^*]'$:

$$\begin{aligned} y_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \sigma_\epsilon \end{bmatrix} \varepsilon_t \\ \boldsymbol{\alpha}_{t+1} &= \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \sigma_\psi\sqrt{1-\rho^2} & 0 \\ 0 & \sigma_\psi\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \chi_t \\ \chi_t^* \end{bmatrix} \end{aligned} \tag{2.9}$$

Ist $\rho < 1$ und λ_c gleich 0 oder π, reduziert sich der stochastische Zyklus zu einem AR(1) Prozess.

2.2.3 Erklärende Variablen

Mittels geeignet formulierten erklärenden Variablen lassen sich zahlreiche Zeitreihenmodelle als Zustandsraummodelle darstellen, so etwa die klassischen ARMA und ARIMA Modelle sowie Regressionsmodelle mit ARMA Fehlerstrukturen. Ein klassisches multiples lineares Regressionsmodell mit k erklärenden Variablen x_i und konstanten Koeffizienten β_i besitzt mit $\boldsymbol{\alpha}_t = [\beta_1, \ldots, \beta_k]'$ die Zustandsraumdarstellung

$$\begin{aligned} y_t &= \begin{bmatrix} x_{1,t} & \cdots & x_{k,t} \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \sigma_\epsilon \end{bmatrix} \varepsilon_t \\ \boldsymbol{\alpha}_{t+1} &= \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \boldsymbol{\alpha}_t \end{aligned} \tag{2.10}$$

wobei zur Modellierung des Intercepts im Bedarfsfall $x_{1,t} = 1$ gesetzt werden kann. Werden die Regressionskoeffizienten in (2.10) statt als Konstanten als Random Walks modelliert, lassen sich auch Regressionsmodelle mit zeitlich veränderlichen Koeffizienten in Zustandsraumform bringen (Durbin und Koopman 2001, Zivot und Wang 2003).

2.2.4 Interventionen

Inhomogenitäten äussern sich in strukturellen Zustandsraummodellen als sprunghafte Veränderungen einzelner Modellkomponenten. Mittels geeignet formulierten erklärenden Variablen können diese Interventionen modelliert und eliminiert werden. Interventionsvariablen stellen daher das wichtigste Werkzeug zur Homogenisierung von langen Beobachtungs- und Messreihen mittels strukturellen Zustandsraummodellen dar. Interventionsvariablen besitzen eine Epoche sowie eine Signatur. Erstere bezeichnet den ersten beeinflussten Zeitpunkt, letztere modelliert die Form der Beeinflussung. In Abbildung 2.1 sind häufig verwendete Signaturen dargestellt. Über den geschätzten Regressionskoeffizienten kann die Signifikanz des Interventionseffekts geprüft sowie seine Relevanz beurteilt werden. Die Identifizierung der Interventionsepochen wird in Kapitel 6 diskutiert.

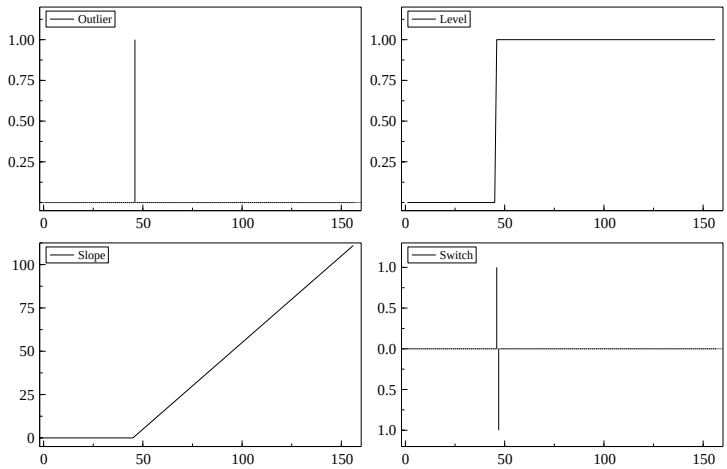


Abbildung 2.1: Verlauf von vier erklärenden Dummyvariablen zur Modellierung eines Ausreissers, einer Sprungintervention, einer Steigungsintervention sowie einer Switchintervention.

2.3 Modellschätzung

Ist eine passende Zustandsraumdarstellung gefunden, können die unbekannten Modellkomponenten mit dem Kalman Filter und Smoother (KFS) extrahiert werden (Durbin und Koopman 2001).

2.3.1 Kalman Filter

Der Kalman Filter berechnet aus den Beobachtungen $\mathbf{Y}_t = \{y_1, y_2, \ldots, y_t\}$ die Kleinstquadrateschätzer der Zustände $\mathbf{a}_{t+1} = \mathrm{E}(\boldsymbol{\alpha}_{t+1}|\mathbf{Y}_t)$ sowie deren Varianzen $\mathbf{P}_{t+1} = \mathrm{Var}(\boldsymbol{\alpha}_{t+1}|\mathbf{Y}_t)$:

$$\begin{aligned}\mathbf{a}_{t+1} &= \mathbf{T}_t\mathbf{a}_t + \mathbf{K}_t\upsilon_t \\ \mathbf{P}_{t+1} &= \mathbf{T}_t\mathbf{P}_t\mathbf{L}_t' + \mathbf{H}_t\mathbf{H}_t'\end{aligned}$$

mit

$$\begin{aligned}\upsilon_t &= y_t - \mathbf{Z}_t\mathbf{a}_t \\ F_t &= \mathbf{Z}_t\mathbf{P}_t\mathbf{Z}_t' + G_tG_t' \\ \mathbf{K}_t &= \mathbf{T}_t\mathbf{P}_t\mathbf{Z}_t'F_t^{-1} \\ \mathbf{L}_t &= \mathbf{T}_t - \mathbf{K}_t\mathbf{Z}_t\end{aligned} \tag{2.11}$$

Erstere bezeichnen den Einschrittprognosefehler υ_t, dessen Varianz F_t sowie den Kalman Gain $\mathbf{K}_t$.

2.3.2 Kalman Smoother

Sind alle Beobachtungen $\mathbf{Y}_n = \{y_1, y_2, \ldots, y_n\}$ bekannt, können mittels folgenden Rückwärtsrekursionen die geglätteten Zustände $\hat{\boldsymbol{\alpha}}_t = \mathrm{E}(\boldsymbol{\alpha}_t|\mathbf{Y}_n)$ sowie deren Varianzen $\mathbf{V}_t = \mathrm{Var}(\boldsymbol{\alpha}_t|\mathbf{Y}_n)$ berechnet werden:

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_{\boldsymbol{t}} &= \mathbf{a}_t + \mathbf{P}_t\mathbf{r}_{t-1} \\ \mathbf{V}_t &= \mathbf{P}_t - \mathbf{P}_t\mathbf{N}_{t-1}\mathbf{P}_t\end{aligned}$$

mit

$$\begin{aligned}\mathbf{r}_{t-1} &= \mathbf{Z}_t'e_t + \mathbf{T}_t'\mathbf{r}_t \\ \mathbf{N}_{t-1} &= \mathbf{Z}_t'F_t^{-1}\mathbf{Z}_t + \mathbf{L}_t'\mathbf{N}_t\mathbf{L}_t \\ e_t &= F_t^{-1}\upsilon_t - \mathbf{K}_t'\mathbf{r}_t \\ D_t &= F_t^{-1} + \mathbf{K}_t'\mathbf{N}_t\mathbf{K}_t\end{aligned} \tag{2.12}$$

worin die Glättungsfehler e_t als sog. smoothations bekannt sind (de Jong und Penzer 1998).

2.3.3 Disturbance Smoother

Sind die geglätteten Zustände $\hat{\boldsymbol{\alpha}}_t$ bekannt, so können die geglätteten Fehlerprozesse $\hat{\epsilon}_t = \mathrm{E}(\epsilon_t|\mathbf{Y}_n)$ und $\hat{\boldsymbol{\eta}}_t = \mathrm{E}(\boldsymbol{\eta}_t|\mathbf{Y}_n)$ mittels der Beziehungen

$$\begin{aligned} \hat{\epsilon}_t &= y_t - \mathbf{Z}_t\hat{\boldsymbol{\alpha}}_t \\ \hat{\boldsymbol{\eta}}_t &= \hat{\boldsymbol{\alpha}}_{t+1} - \hat{\boldsymbol{\alpha}}_t \end{aligned} \tag{2.13}$$

einfach berechnet werden. Wie Koopman (1993) gezeigt hat, können die geglätteten Fehlerprozesse $\hat{\epsilon}_t = \mathrm{E}(\epsilon_t|\mathbf{Y}_n)$ und $\hat{\boldsymbol{\eta}}_t = \mathrm{E}(\boldsymbol{\eta}_t|\mathbf{Y}_n)$ mittels den nachfolgenden Rekursionen auch ohne Kenntnis der geglätteten Zustände $\hat{\boldsymbol{\alpha}}_t$ direkt berechnet werden:

$$\begin{aligned} \hat{\epsilon}_t &= G_t G_t' e_t \\ \mathrm{Var}(\epsilon_t|\mathbf{Y}_n) &= G_t G_t' - G_t G_t' D_t G_t G_t' \\ \hat{\boldsymbol{\eta}}_t &= \mathbf{H}_t \mathbf{H}_t' \mathbf{r}_t \\ \mathrm{Var}(\boldsymbol{\eta}_t|\mathbf{Y}_n) &= \mathbf{H}_t\mathbf{H}_t' - \mathbf{H}_t\mathbf{H}_t'\mathbf{N}_t\mathbf{H}_t\mathbf{H}_t'. \end{aligned} \tag{2.14}$$

2.3.4 Fehlende Werte

Fehlen einige der Beobachtungen y_t werden an diesen Stellen $\upsilon_t = 0$, $F_t^{-1} = 0$ und $\mathbf{K}_t = 0$ gesetzt, wodurch sich für den Kalman Filter folgende modifizierten Gleichungen ergeben:

$$\begin{aligned} \mathbf{a}_{t+1} &= \mathbf{T}_t\mathbf{a}_t \\ \mathbf{P}_{t+1} &= \mathbf{T}_t\mathbf{P}_t\mathbf{T}_t' + \mathbf{H}_t\mathbf{H}_t'. \end{aligned} \tag{2.15}$$

In den Kalman Smoother Gleichungen modifizieren sich:

$$\begin{aligned} \mathbf{r}_{t-1} &= \mathbf{T}_t'\mathbf{r}_t \\ \mathbf{N}_{t-1} &= \mathbf{T}_t'\mathbf{N}_t\mathbf{T}_t. \end{aligned} \tag{2.16}$$

Für den Disturbance Smoother werden

$$\begin{aligned} \hat{\epsilon}_t &= 0 \\ \mathrm{Var}(\epsilon_t|\mathbf{Y}_n) &= G_t G_t', \end{aligned} \tag{2.17}$$

während $\hat{\boldsymbol{\eta}}_t$ und $\mathrm{Var}(\boldsymbol{\eta}_t|\mathbf{Y}_n)$ nur insofern beeinflusst werden, als $\mathbf{r}_t$ und $\mathbf{N}_t$ nun andere numerische Werte annehmen, als wenn y_t nicht fehlen würde. Diese einfache und effiziente Behandlung fehlender Werte ist einer der Vorzüge, welche den Kalman Filter und Smoother als Schätzer so attraktiv machen (Harvey 1989, Durbin und Koopman 2001).

2.3.5 Parameterbestimmung

Gewöhnlich enthalten die Systemmatrizen eine ganze Reihe unbekannter Parameter $\boldsymbol{\phi}$, welche mittels Maximum Likelihood geschätzt werden können. Für univariate Gauss'sche Zustandsraummodelle lautet die Likelihood Funktion:

$$L(\boldsymbol{\phi}) = p(y_1, \ldots, y_n) = \prod_{t=1}^{n} p(y_t|\mathbf{Y}_{t-1}) \tag{2.18}$$

mit $p(y_1|\mathbf{Y}_0) = p(y_1)$. Die einzelnen Terme sind durch

$$p(y_t|\mathbf{Y}_{t-1}) = \frac{1}{\sqrt{2\pi|F_t|}} \exp\left\{-\frac{1}{2}\upsilon_t' F_t^{-1} \upsilon_t\right\} \tag{2.19}$$

gegeben, deren Parameter υ_t und F_t^{-1} sich mit dem Kalman Filter einfach berechnen lassen. Daraus ergibt sich die Log-Likelihood Funktion:

$$l(\boldsymbol{\phi}) = -\frac{1}{2}\left(n \log 2\pi + \sum_{t=1}^{n} \log|F_t| + \sum_{t=1}^{n} \upsilon_t' F_t^{-1} \upsilon_t\right) \tag{2.20}$$

welche nach Schweppe (1965) die *prediction error decomposition* genannt wird (Harvey 1989). STAMP 6.2 und SsfPack 2.3 benutzen zur Maximierung der Log-Likelihood Funktion das Broyden-Fletcher-Goldfarb-Shanno Verfahren (BFGS), eine Quasi-Newton Optimierungsmethode welche in Schüttel (2002) ausführlich beschrieben ist.

2.3.6 Initialisierung

Die bisher vorgestellten Algorithmen gelten unter der Voraussetzung, dass die Parameter $\mathbf{a}_1$ und $\mathbf{P}_1$ der Verteilung $\boldsymbol{\alpha}_1 \sim N(\mathbf{a}_1, \mathbf{P}_1)$ gegeben sind. Im Allgemeinen ist jedoch über einige so genannt *diffuse* Elemente von $\boldsymbol{\alpha}_1$ nichts bekannt. Der Kalman Filter und Smoother muss in diesem Fall geeignet initialisiert werden. Hierzu sind mehrere Verfahren verfügbar:

Während in STAMP 6.2 die von Rosenberg (1973) entwickelte und von de Jong (1988, 1991) verfeinerte *augmentation* Methode implementiert ist (Koopman, Harvey, Doornik und Shephard 2000), wird in SsfPack 2.3 eine approximative *infinity* Methode verwendet (Koopman et al. 1999), welche in SsfPack 3 durch eine exakte Variante ersetzt worden ist (Koopman et al. 2002). Eine ausführliche Beschreibung der exakten diffusen Initialisierungsmethode findet sich in Steiner (2004).

2.4 Modellvalidierung

Nach Gleichung (2.3) sollten die beiden Fehlerprozesse ϵ_t und $\boldsymbol{\eta}_t$ unabhängig und normalverteilt sein. Harvey und Koopman (1992) haben jedoch gezeigt, dass die mit dem Disturbance Smoother berechneten Schätzer $\hat{\epsilon}_t$ und $\hat{\boldsymbol{\eta}}_t$ je nach Zustandsraummodell mehr oder weniger stark auto- und kreuzkorreliert sind, so dass eine Modellvalidierung anhand der geglätteten Fehlerprozesse $\hat{\epsilon}_t$ und $\hat{\boldsymbol{\eta}}_t$ nicht durchführbar ist. In einem korrekt spezifizierten linearen Gauss'schen

Zustandsraummodell mit bekannten Parametern sollte nach Harvey (1989) allerdings gelten:

$$v_t \sim N(0, F_t), \qquad t = d+1, \ldots, n. \tag{2.21}$$

wobei d die Anzahl diffus initialisierter Elemente von $\boldsymbol{\alpha}_1$ ist. Trotzdem die Modellparameter gewöhnlich unbekannt sind und mittels Maximum Likelihood geschätzt werden müssen, wird anhand geeigneter Visualisierungen und statistischer Tests geprüft, ob die standardisierten Innovationen normalverteilt, homoskedastisch und unkorreliert sind sowie einen konstanten Mittelwert aufweisen. Einen Überblick über die für die Validierung von Zustandsraummodellen üblichen Diagnoseverfahren geben Harvey (1989) und Schüttel (2002).

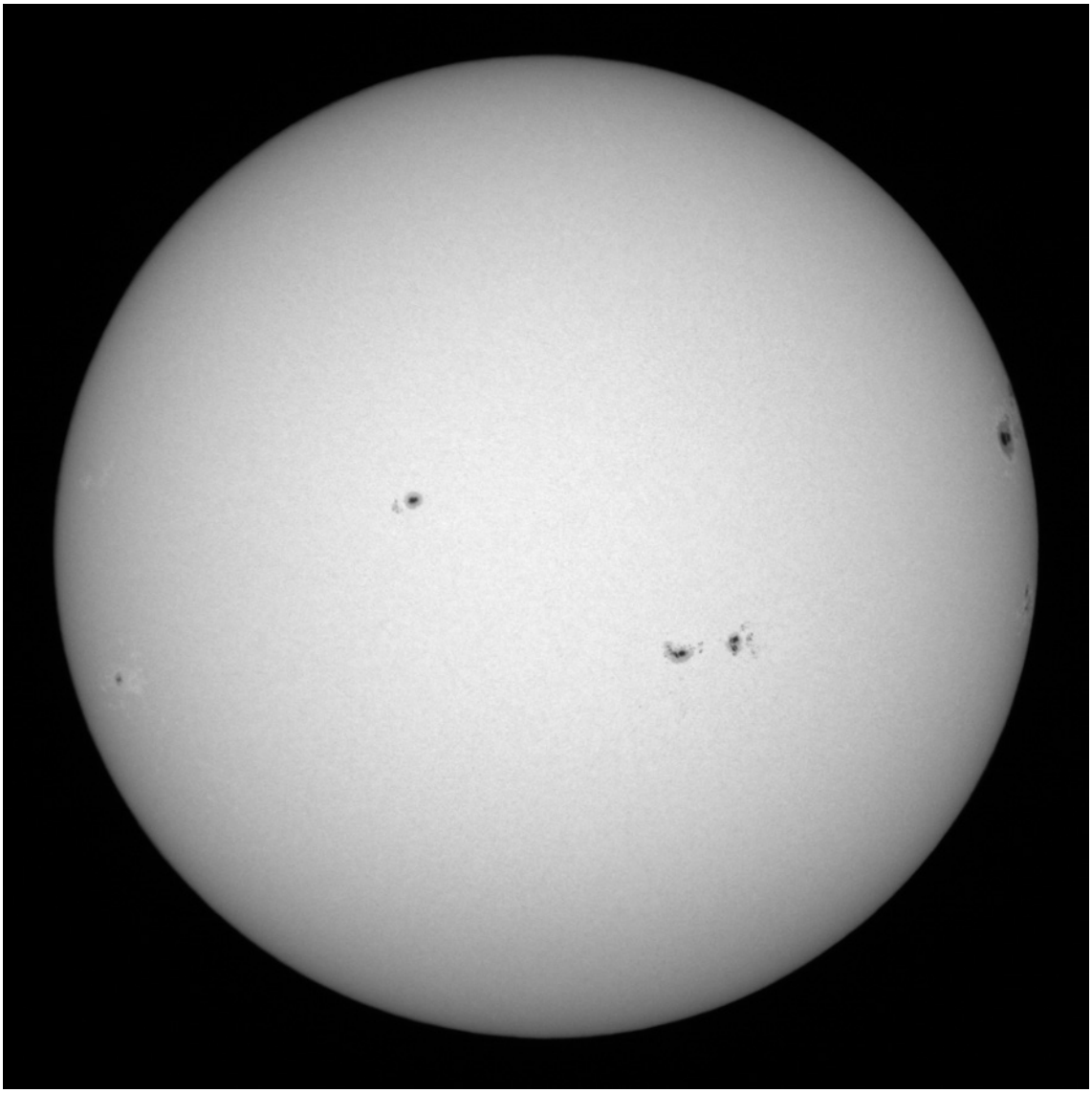


Abbildung 2.2: Sonnenaktivität am 17. Juni 2004. Aufnahme 10:15 UT am TeleVue NP101 Refraktor des Autors mit einer Canon EOS 300D Digitalkamera. In den fünf sichtbaren Sonnenfleckengruppen konnten insgesamt 78 Einzelflecken gezählt werden. Belichtungszeit 1/4000 sec.

3. Darstellung der Sonnenaktivität

In diesem Kapitel soll eine für die Homogenitätsprüfung brauchbare Darstellung der Sonnenaktivität gefunden werden. Hierzu werden in den ersten Abschnitten die Phänomene der Sonnenoberfläche – insbesondere die Flecken und Fackeln – beschrieben sowie die Quantifizierung und Darstellung der Sonnenaktivität durch die ausgeglichenen Monatsmittel der Wolfschen Sonnenfleckenrelativzahl vorgestellt. In den restlichen Abschnitten werden sodann strukturelle Zustandsraummodelle entwickelt und mit anderen Darstellungsformen der Wolfschen Reihe verglichen.

3.1 Der Anblick der Sonnenoberfläche

Betrachten wir die Sonne durch ein mit einem ausreichenden Blendschutz versehenen Fernrohr, so sehen wir die sog. Photosphäre der Sonne, diejenige unterste, nur rund 200 km mächtige Schicht der Sonnenatmosphäre aus welchem der Löwenanteil des sichtbaren Lichtes emittiert wird. Trotzdem diese Schicht so dünn ist, dass sie von der Erde aus nur eine vertikale Ausdehnung von 0.2 Bogensekunden besitzt und tangential betrachtet somit einen scharfen Sonnenrand vortäuscht, fällt die Lichtemission doch stark mit der Höhe ab, so dass wir bei schrägen und daher weniger tief eindringenden Sehstrahlen rasch in kühlere, weniger helle Schichten der Sonnenatmosphäre blicken. Dieser Effekt ist als sog. Randverdunkelung bekannt und kann mit jedem Fernrohr leicht beobachtet werden (Abb. 2.2). Bei stabilen Luftverhältnissen und ausreichender Instrumentenleistung können wir auf der Sonnenscheibe zudem eine körnige Struktur feststellen: die sog. Granulation. Die einzelnen Granula sind polygonal geformte Konvektionszellen mit einer mittleren Lebensdauer von wenigen Minuten und einer mittleren Ausdehnung von rund 700 km. Zwischen den Granulationszellen befindet sich das dunkle Netzwerk der intergranularen Zwischenräume (Abb. 3.1).

Ist die Sonne aktiv, so sehen wir in zwei mehr oder weniger breiten Gürteln nördlich und südlich des Sonnenäquators Aktivitätsgebiete mit dunklen Flecken und hellen Fackeln. Sonnenflecken sind kühlere Gebiete auf der Sonnenoberfläche und markieren die Durchstosspunkte gewaltiger Magnetfelder, welche sich zu magnetischen Flussröhren verdreht aus dem Sonneninnern bis hoch in die äusseren Atmosphärenschichten der Sonne erstrecken. In den meisten Fällen sind mehrere Sonnenflecken sichtbar, welche sich innerhalb eines Aktivitätsgebietes zu charakteristischen bipolaren Gruppen anordnen. Die beiden Hauptflecken einer Gruppe haben entgegengesetzte magnetische Polarität. Ihre Verbindungslinie ist ungefähr Ost-West orientiert und leicht äquatorwärts geneigt. Grössere Flecken sind von einem hofähnlichen Halbschatten, der sog. Penumbra, umgeben, welche auch mehrere Kernflecken, die sog. Umbrae, umschliessen kann. Der Anblick der Penumbrae ist gewöhnlich ohne weitere Struktur. Nur bei ruhiger Luft und genügender Instrumentenleistung ist zu erkennen, dass die Penumbren aus einzelnen helleren und dunkleren Lanzetten bestehen.

In grösseren Sonnenflecken sehen wir hin und wieder Stellen, welche gleich hell sind wie die ungestörte Photosphäre. Dies sind sog. Lichtbrücken. Sie treten gelegentlich als isolierte helle Inseln in der Penumbra oder Umbra in Erscheinung. Meistens zerschneiden sie jedoch als helle Zungen Teile der Penumbra oder der Umbra oder teilen gar ganze Flecken in mehrere disjunkte Teile.

Die Aktivitätsgebiete bewegen sich in rund 14 Tagen mit der allgemeinen Sonnenrotation von Ost nach West. Nähert sich ein grösserer Sonnenfleck dem Sonnenrand, so scheint sich die Umbra gegenüber der Penumbra zum Scheibenzentrum hin zu verschieben (sog. Wilson Effekt). Dies zeigt, dass die Photosphärenmaterie in der Umbra durchsichtiger ist als in der Penumbra und in der ungestörten Photosphäre und wir somit in den Zentren der schwarzen Kernflecken tiefer in das Soneninnere schauen können als in der ungestörten Photosphäre. In der Nähe des Sonnenrandes sehen wir auch die hellen Lichtadern der Sonnenfackeln, welche die Aktivitätsgebiete umgeben und in der Scheibenmitte von tieferliegenden Photosphärenschichten überstrahlt werden.

Die Aktivitätsgebiete beschränken sich nicht nur auf die Photosphäre der Sonne, sondern erreichen über die Chromosphäre auch die Sonnenkorona. Von den in diesen Schichten beobachtbaren Phänomenen, wie den chromosphärischen Fackelgebieten, den Protuberanzen, den Eruptionen und den gelegentlichen koronalen Massenauswürfen sehen wir in der Photosphäre nichts, lediglich die intensivsten Sonneneruptionen können in ganz seltenen Fällen auch im Intergrallicht als gelbliche Leuchterscheinungen wahrgenommen werden (sog. Weisslichtflares).

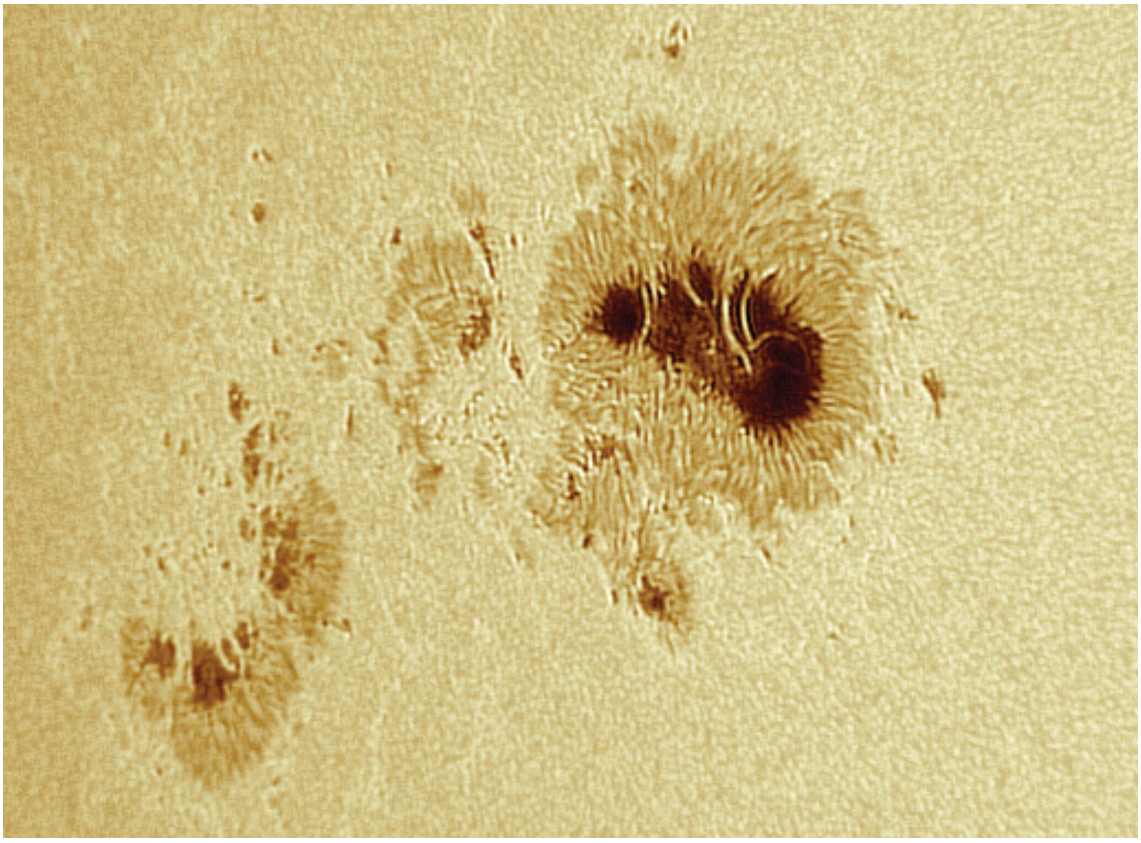


Abbildung 3.1: Detailansicht einer Sonnenfleckengruppe vom 19. Juni 2004. Neben der Granulation sind Feinstrukturen der Umbrae und Penumbrae zu erkennen, insbesondere Lichtbrücken und Penumbrafilamente. Aufnahme mit einem Astrophysics 130 mm EDF f/6 Refraktor.

3.2 Quantifizierung der Sonnenaktivität

Schon den Entdeckern der Sonnenflecken - dem Engländer Thomas Harriot (1560-1621), dem Friesen Johannes Fabricius (1587-1615), dem Italiener Galileo Galilei (1564-1642) sowie dem Deutschen Jesuit Christoph Scheiner (1575-1650) - fiel auf, dass die Zahl der Sonnenflecken starken Schwankungen unterworfen ist (Friedli 1984). Während gut zweier Jahrhunderte glaubte man an die Regellosigkeit dieser Schwankungen (Wolf 1890). Erst Heinrich Samuel Schwabe entdeckte aufgrund jahrelanger Beobachtung eine etwa zehnjährige Aktivitätsperiode (Schwabe 1844). Dies gelang ihm, weil er von Anfang an systematische Zählungen von Sonnenfleckengruppen vornahm und einen ersten Aktivitätsindex einführte: die Jahressumme auf der Sonnenscheibe neu entstandener oder durch die Rotation sichtbar gewordener Fleckengruppen (Schwabe 1843). Dieser hatte natürlich den gewichtigen Nachteil, dass die Beobachtungsregister praktisch lückenlos sein mussten, sollten die Jahressummen untereinander vergleichbar sein; eine für einen Einzelbeobachter schon aus meteorologischen Gründen unerfüllbare Bedingung. Rudolf Wolf führte deshalb für private Zwecke schon 1850, allgemein jedoch erst 1856, einen täglichen Index der Sonnenfleckenaktivität ein, die sog. Sonnenfleckenrelativzahl r,

> "für deren Berechnung ich (Wolf) mich auf die Überlegung stützte, dass die wechselnde Tätigkeit der Sonne zunächst der Anzahl g der gleichzeitig vorhandenen Gruppen proportional sein werde, in untergeordneter Weise aber auch in der Grösse der Gruppen ihren Ausdruck finde, welche durch die in Teilen der ganzen Sonnenoberfläche ausgedrückte Flächensumme f'' sämtlicher Flecken oder die diese erfahrungsweise durchschnittlich nahezu ersetzende Anzahl f dieser Flecken repräsentiert werden könne. Durch einen glücklichen Griff schon 1850 für die Gewichte der g und f die Zahlen 10 und 1 wählend, erhielt ich so zur Berechnung der Relativzahlen die bequeme Formel:
>
> $$r = k \cdot (10 \cdot g + f)$$
>
> in welcher k ein mit Beobachter und Instrument wechselnder, aus korrespondierenden Beobachtungen zu bestimmender Faktor ist, welchen ich für mich und die Vergrösserung 64 eines Fraunhofer'schen Vierfüssers gleich der Einheit nahm (Wolf 1890). Allerdings hätte ich den so erhaltenen Relativzahlen unbedingt solche vorgezogen, die dem Gesamtareal der Flecken proportional gewesen wären, — aber es hätte mich zu weit geführt, die zu deren Ermittlung nötigen Messungen und Rechnungen zu machen; denn ich hätte offenbar zu diesem Zwecke nicht nur die Dimensionen der Flecken abmessen oder abschätzen müssen, sondern wäre genötigt gewesen, auch die Position jedes Fleckens auf der Sonne zu bestimmen, um mit ihrer Hilfe die verkürzt erscheinenden Dimensionen zu reduzieren" (Wolf 1856), doch "glaube ich, durch die oben vorgenommene Weise ein möglichst gutes Surrogat dafür zu erhalten" (Wolf 1851).

Die Zählweise von Sonnenfleckengruppen und Einzelflecken wurde im Laufe der Jahre mehrfach geändert (Friedli 1988), insbesondere führte Wolfs Nachfolger Alfred Wolfer (1854-1931) eine Gewichtung der Einzelflecken ein, welche sich jedoch bei den zahlreichen auswärtigen Beobachtungsstationen und Amateursonnenbeobachtern nie durchsetzen konnte und nur für die Beobachter der Eidgenössischen Sternwarte in Zürich und ihrer Aussenstationen in Arosa und Locarno Verwendung fand. Friedli (1988, 1989a, 1989b, 1999) dokumentierte die vorhandenen Flecken- und

Gruppendefinitionen sowie die Zürcher Zählpraxis und erarbeitete eine einheitliche Definition der Wolfschen Relativzahl, welche konsistent ist mit den klassischen Zählmethoden Rudolf Wolfs und der massgebenden Zürcher Beobachter (Abbildung **??**). Darüberhinaus ist sie auch geeignet für die automatische Bestimmung der Wolfschen Relativzahl, wie sie von modernen Überwachungsstationen benötigt wird (Friedli 2001).

Die Sonnenfleckenrelativzahl hat keine offizielle Einheitsbezeichnung. Im Gedenken an Rudolf Wolf sowie zur Kennzeichnung der Sonnenfleckenrelativzahl als konstruiertem Aktivitätsindex und nicht als "Zahl der Sonnenflecken" schlage ich vor, als Einheit [Wolf] zu setzen.

$$
\begin{aligned}
r_t : \Gamma_t \times \Phi_t &\longrightarrow \mathbb{N}_0 \\
(\gamma_t\,, \varphi_t) &\longmapsto 10 \cdot \Bbbk_1(\gamma_t) + \Bbbk_2(\varphi_t)
\end{aligned}
$$

$\Phi_t := \{\varphi \in \mathbb{N}_0 \,|\, \varphi = \text{card } F_t\}$

$F := \{f \,|\, f \text{ ist ein Sonnenfleck}\}$

Ein Gebiet der Photosphäre der Sonne heisst Sonnenfleck $f \in F$, falls

S1: dessen Farbe schwarz, nicht nur gräulich ist,
S2: dessen Lebensdauer mindestens 30 Minuten beträgt,
S3: dessen Fläche 1 MH übersteigt und
S4: dessen Fläche zusammenhängend ist.

$\Gamma_t := \{\gamma \in \mathbb{N}_0 \,|\, \gamma = \text{card } G_t\}$

$G := \{g \,|\, g \text{ ist eine Sonnenfleckengruppe}\}$

Eine Ansammlung von Sonnenflecken $f \in F$ heisst Sonnenfleckengruppe $g \in G$, falls

G1: sie räumlich isoliert ist,
G2: sich eigenständig entwickelt und
G3: sie demselben magnetischen Aktivitätsgebiet angehört.

t	:=	Beobachtungsepoche
$\mathbb{N}_0$	:=	$\{n \,\|\, n \in \mathbb{N} \vee n = 0\}$
card X	:=	Anzahl Elemente der Menge X
$\Bbbk_1, \Bbbk_2$	:=	solarstatistische Kalibrierungsmodelle

Abbildung 3.2: Definition der Wolfschen Sonnenfleckenrelativzahl

3.3 Klassische Darstellungsverfahren

Zu Beginn wurden die beobachteten Sonnenfleckenrelativzahlen von Rudolf Wolf lediglich gesammelt und zu einer möglichst lückenlosen Reihe zusammengefügt. Der Aktivitätsverlauf wurde anhand der Monatsmittel, die Epochen der Maxima und Minima anhand der Jahresmittel bestimmt.

Erst in den frühen 1870er Jahren verfügte Wolf dank der tatkräftigen Mithilfe befreundeter Astronomen und Amateursonnenbeobachter über genügend Material, um den Verlauf der Sonnenaktivität bis ins 18. Jahrhundert zurück zuverlässig rekonstruieren zu können. Um eine von zufälligen Schwankungen und kurzfristigen Variationen weitgehend freie Darstellung des mittleren Verlaufs der Sonnenaktivität zu erhalten, berechnete Wolf 1873 nach der Formel

$$R_m = \frac{1}{2}\left(\frac{1}{12}\sum_{j=-6}^{5} r_{m-j} + \frac{1}{12}\sum_{j=-5}^{6} r_{m-j}\right) \tag{3.1}$$

geglättete Monatsmittel der Relativzahl (Abb. 3.2), welche er zuerst *Normalzahlen* (Wolf 1873),

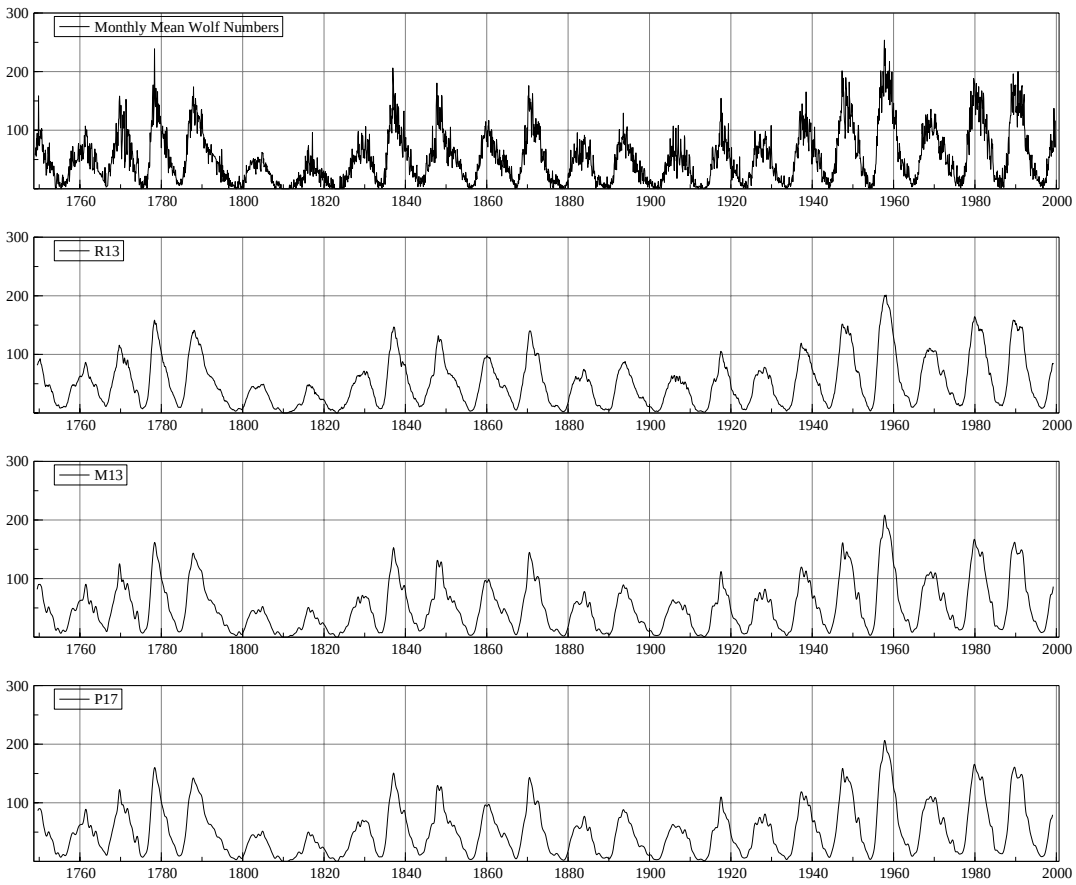


Abbildung 3.2: Beobachtete und mittels der klassischen Verfahren nach Wolf (R13), Meeus (M13) und Karkoschka (P17) ausgeglichene Monatsmittel der Wolfschen Sonnenfleckenrelativzahl.

später *ausgeglichene* Relativzahlen nannte (Wolf 1875). 1877 legte Wolf dann eine mit 1749 beginnende Reihe von beobachteten und ausgeglichenen Monatsmitteln der Sonnenfleckenrelativzahlen vor,

> "der ich wohl Andere bei Benutzung derselben **meinen** Namen beizulegen bitten darf" (Wolf 1890), um "mir so für die Hingabe des Kapitals wenigstens eine bescheidene Rente zu sichern" (Wolf 1877).

Nach heutigem Verständnis entspricht das Wolfsche Glättungsverfahren einem moving average Filter bzw. einem Kernschätzer der Form

$$R_m = \sum_{j=-q}^{q} w_j \cdot r_{m-j} \tag{3.2}$$

wobei die Gewichte w_j für das Wolfsche Verfahren nach Ladiray und Quenneville (2001) folgende Werte besitzen:

$$[13]; \frac{1}{24}\{1, 2, 2, 2, 2, 2, 2\}. \tag{3.3}$$

Izenman (1985) machte darauf aufmerksam, dass Wolf's Formel (3.1) eine der ersten bekannten nicht-ökonometrischen Anwendungen eines moving average Filters darstellt, älter jedenfalls als diejenigen in Klein (1997) hierfür angeführten Arbeiten von Poynting (1877, 1884), Jevons (1878) und Stewart (1883). Wolf's Formel ist auch unter dem Namen "centered 12-term moving average" bekannt und wird bis heute zur Glättung langer, vorwiegend ökonometrischer Zeitreihen eingesetzt. Insbesondere ist sie im bekannten X-11 Verfahren des US Bureau of the Census implementiert und wird dort zur Extraktion zyklischer Trendkomponenten verwendet (Ladiray und Quenneville 2001).

In Amateurkreisen wurden in der zweiten Hälfte des letzten Jahrhunderts mehrere Versuche unternommen, eine Formel zu finden, welche eine bessere Darstellung des mittleren Verlaufs der Sonnenaktivität liefert. Nach einem Vorschlag von Meeus (1958) sollte der trapezförmige Kern der Wolfschen Ausgleichsformel durch einen glockenförmigen der Form

$$[13]; \frac{1}{81}\{1, 3, 5, 7, 9, 10, 11\} \tag{3.4}$$

ersetzt werden. Zwar fanden die mit diesem Kern berechneten ausgeglichenen Monatsmittel dank der bekannten Veröffentlichung (Meeus 1983) eine gewisse internationale Verbreitung, doch wurde die neue Reihe ausser von Meeus (1997, 2002) kaum benutzt, obwohl sich aus deren Verlauf z.T. relevante Abweichungen in den bestimmten Epochen der Zyklusmaxima und -minima ergeben.

Ein ähnliches Schicksal widerfuhr der von Karkoschka (1979) entwickelten sog. P17-Mittelung, welche als Glättungskern ein Polynom sechsten Grades der Form

$$w_j = \frac{g_j}{\sum_{i=-8}^{8} g_j} \quad \text{mit} \quad g_j = \left(1 - \left(\frac{j}{9}\right)^2\right)^3 \tag{3.5}$$

verwendet und seit ihrer Einführung vom deutschen SONNE-Netz – der weltweit grössten Gruppe von Amateursonnenbeobachtern – regelmässig veröffentlicht wird (Reinsch et al. 1999).

Die Verläufe der nach den Verfahren von Wolf, Meeus und Karkoschka ausgeglichenen Monatsmittel der Sonnenfleckenrelativzahl sind in Abbildung 3.2 wiedergegeben.

3.4 Zustandsraummodell für die Monatsmittel der Wolfschen Reihe

Wie in Kapitel 2.2.1 diskutiert, entspricht ein strukturelles Zustandsraummodell mit einem glatten Trend einer nichtparametrischen Kernschätzung. Der ausgeglichene Verlauf der Sonnenaktivität könnte daher durch ein strukturelles Zustandsraummodell extrahiert werden, indem die Monatsmittel der Wolfschen Reihe als Superposition einer glatten, langsam variierenden Trendkomponente und einer irregulären Restkomponente modelliert werden:

$$y_t = \mu_t + \epsilon_t. \tag{3.6}$$

In Abbildung 3.3 sind die Verläufe der beiden extrahierten Komponenten wiedergegeben. Der mittlere Aktivitätsverlauf ist auf den ersten Blick demjenigen der klassischen Glättungsverfahren ähnlich. Die Restkomponente enthält aufgrund der einfachen Dekompositionsstruktur alle in der glatten Trendkomponente nicht modellierten kurzfristigen Aktivitätsschwankungen und Beobachtungsfehler. Die Validierung anhand der in Abbildung 3.4 zusammengestellten Grafiken ergibt denn auch, dass die Verteilung der Innovationen eher einer t-Verteilung mit wenigen Freiheitsgraden denn einer Normalverteilung entspricht und dass die Innovationen autokorreliert sind. Die langfistige Konstanz des Mittelwerts sowie die Homoskedastizität der Innovationen scheinen jedoch gemäss dem Cusum-Plot der Innovationen bzw. dem Cusum-Plot der Innnovationsquadrate gegeben.

Eine mögliche Verbesserung des strukturellen Modells (3.6) besteht in der Mitberücksichtigung einer zyklischen statt einer irregulären Komponente:

$$y_t = \mu_t + \psi_t. \tag{3.7}$$

In Abbildung 3.5 sind die Verläufe der extrahierten Komponenten dargestellt. Die geschätzte zyklische Komponente ist von Auge von der irregulären Restkomponente des Modells (3.6) fast nicht zu unterscheiden. Auch die Validierungsgrafiken in Abbildung 3.6 unterscheiden sich kaum von denjenigen in Abbildung 3.4, jedoch sind die kurzfristigen Autokorrelationen verschwunden und die globale Modellanpassung (Tabelle 3.1) ist besser.

Modell	PEV	AIC	σ_ζ^2	σ_ϵ^2	σ_ψ^2	ρ	c	A
(3.6)	282.37	5.64	0.59	203.27	-	-	-	-
(3.7)	255.94	5.55	0.38	-	201.77	0.35	26.19	8.97

Tabelle 3.1: Geschätzte Modellparameter und Komponenten der beiden Zustandsraummodelle (3.6) und (3.7). PEV bezeichnet die Varianz der Innovationen, AIC das Akaike Informationskriterium, c die Periodenlänge und A die Amplitude des stochastischen Zyklus.

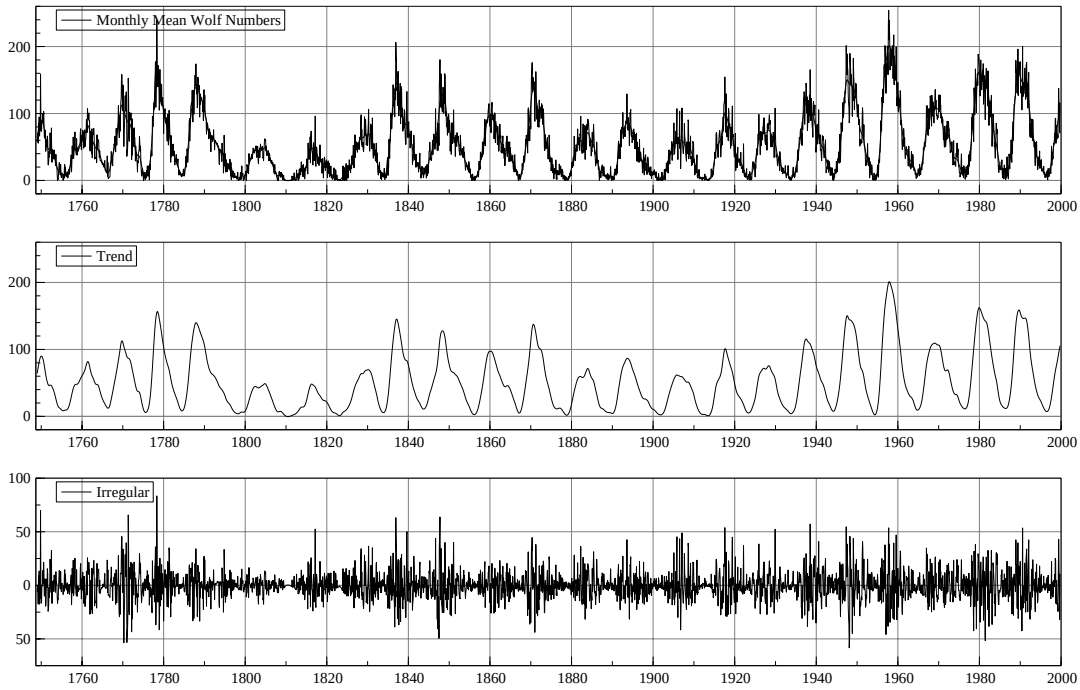


Abbildung 3.3: Dekomposition der Monatsmittel der Sonnenfleckenrelativzahl in einen stochastischen Trend und eine irreguläre Komponente.

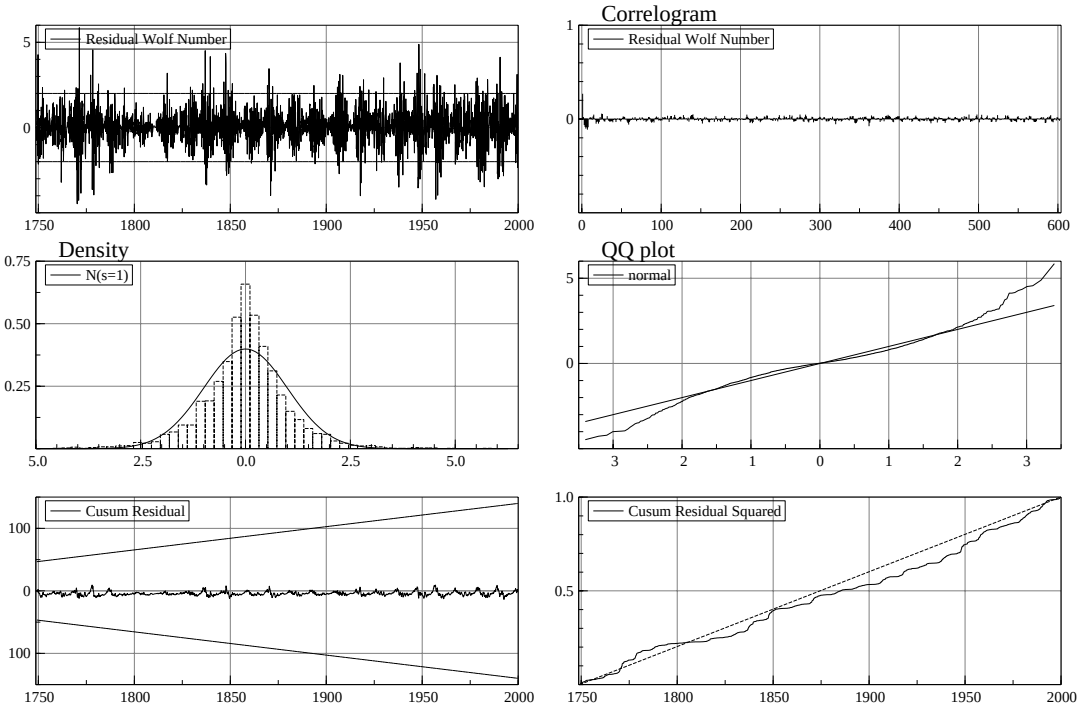


Abbildung 3.4: Diagnostik zu den Innovationen des Modells in Abbildung 3.3. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

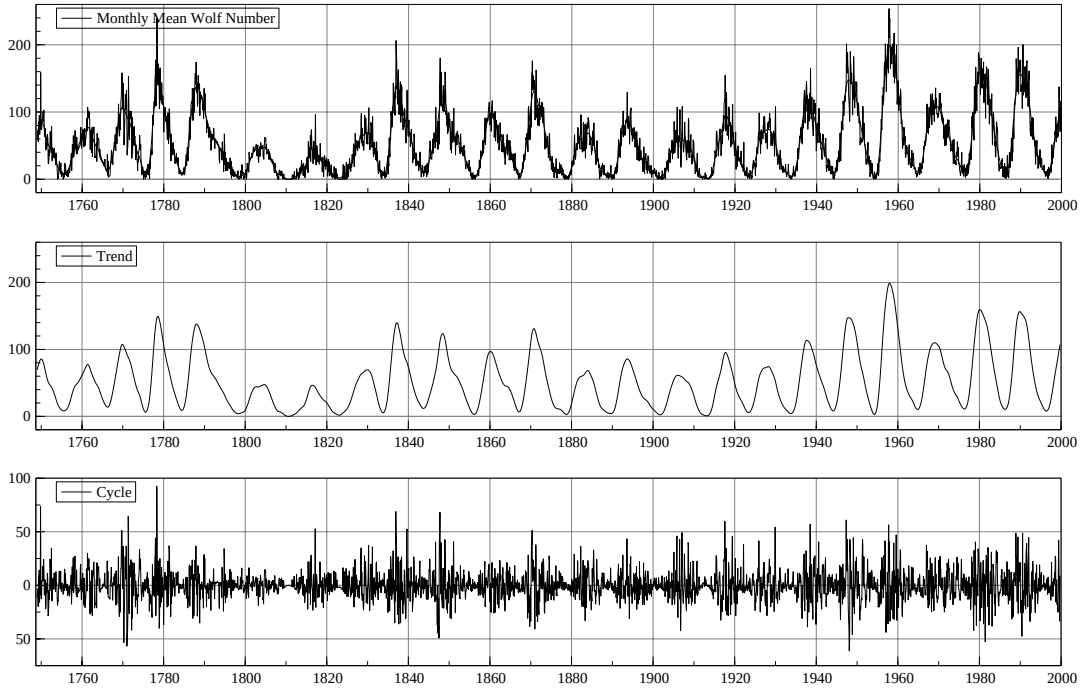


Abbildung 3.5: Dekomposition der Monatsmittel der Sonnenfleckenrelativzahl in einen stochastischen Trend und einen stochastischen Zyklus.

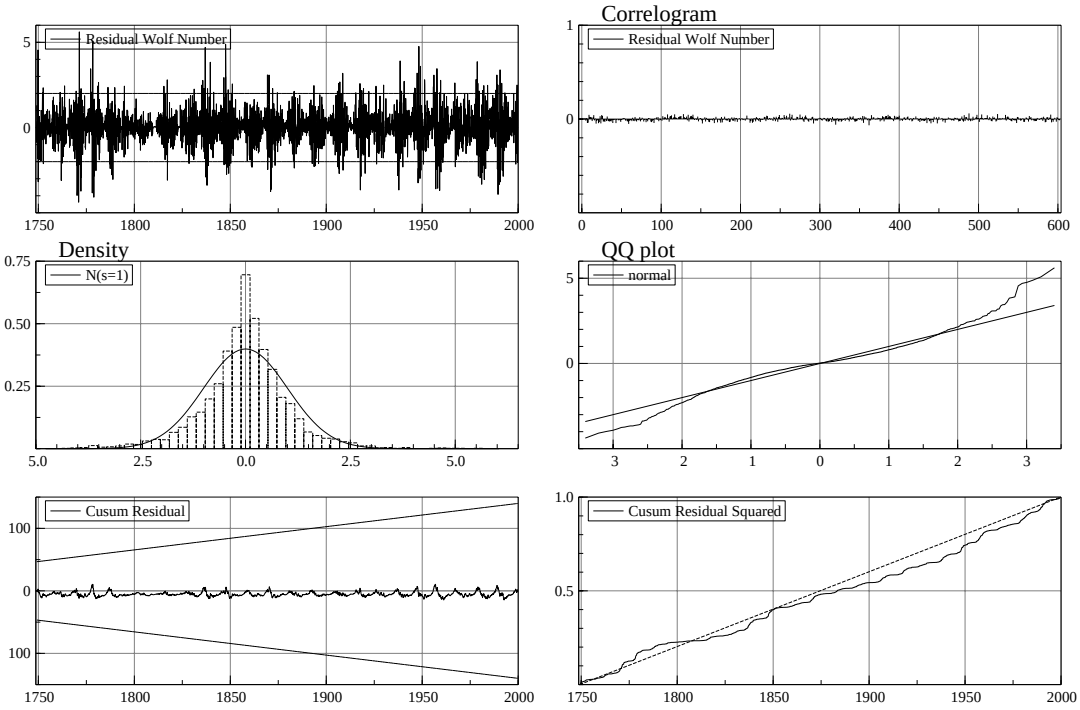


Abbildung 3.6: Diagnostik zu den Innovationen des Modells in Abbildung 3.5. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

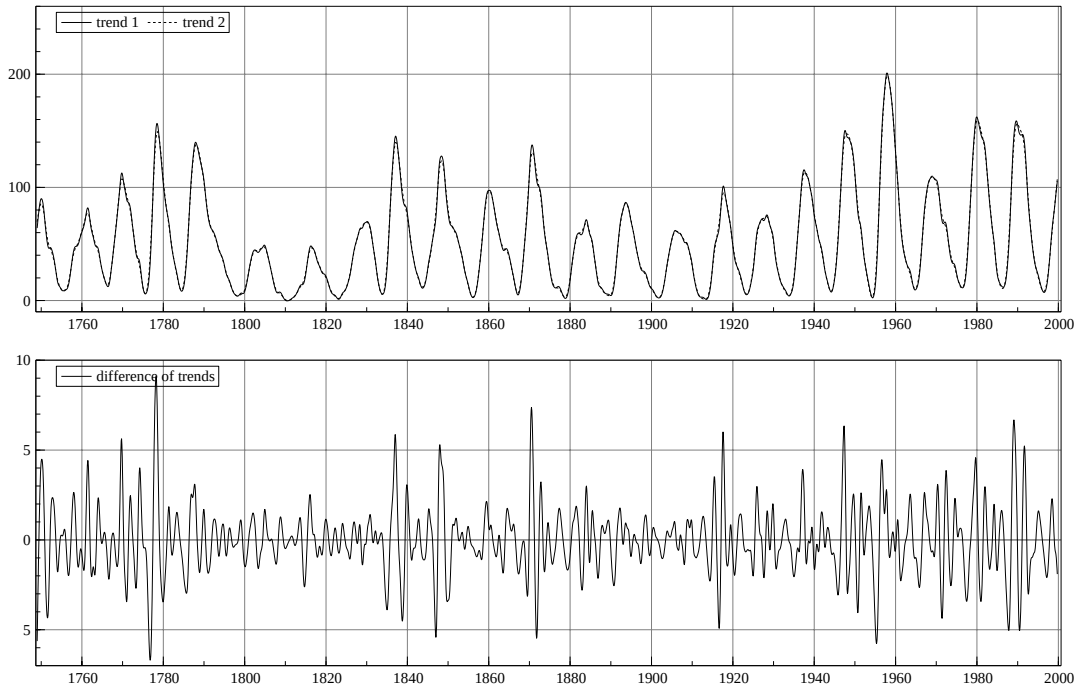


Abbildung 3.7: Differenz der geschätzten Trendkomponenten der beiden Modelle (3.6) und (3.7). Oben die beiden Trendverläufe, unten deren Differenz.

Ein Vergleich der beiden extrahierten Trendkomponenten in Abbildung 3.7 zeigt, dass die Differenzen in der Grössenordnung einiger weniger Sonnenflecken sind, allerdings ist der Trendverlauf im Modell (3.7) deutlich glatter, als derjenige in Modell (3.6), was bei raschen Aktivitätswechseln zu systematischen Abweichungen führt, so etwa bei den Maxima der Zyklen 3, 8, 9, 11, 15, 18 und 22.

3.5 Vergleich mit klassischen Darstellungsverfahren

Betrachten wir den in Abbildung 3.8 am Beispiel des 14. Sonnenaktivitätszyklus wiedergegebenen, nach dem Wolfschen Verfahren berechneten Verlauf der ausgeglichenen Monatsmittel der Relativzahl etwas genauer, so erkennen wir, dass dieser keineswegs ideal glatt ist. Auch fallen die lokalen Extremalstellen der ausgeglichenen Kurve nicht mit denjenigen der beobachteten zusammen, vielmehr korrespondiert meistens ein lokales Maximum der beobachteten Mittelwerte gerade mit einem lokalen Minimum der ausgeglichenen Mittelwerte und umgekehrt.

Wie aus Abbildung 3.9 ersichtlich, vermögen die nach den Verfahren von Meeus (1958) und Karkoschka (1979) berechneten ausgeglichenen Monatsmittel beide Missstände bis zu einem gewissen Grad zu korrigieren. So sind beide Verläufe viel glatter als die R13-Mittel des Wolfschen Verfahrens. Auch stimmen die lokalen Maxima und Minima der ausgeglichenen Monatsmittel mit den lokalen Maxima und Minima der beobachteten Werte überein. Leider zeigen jedoch beide Kurven ein leichtes Undersmoothing, insbesondere in den Jahren 1908 und 1909.

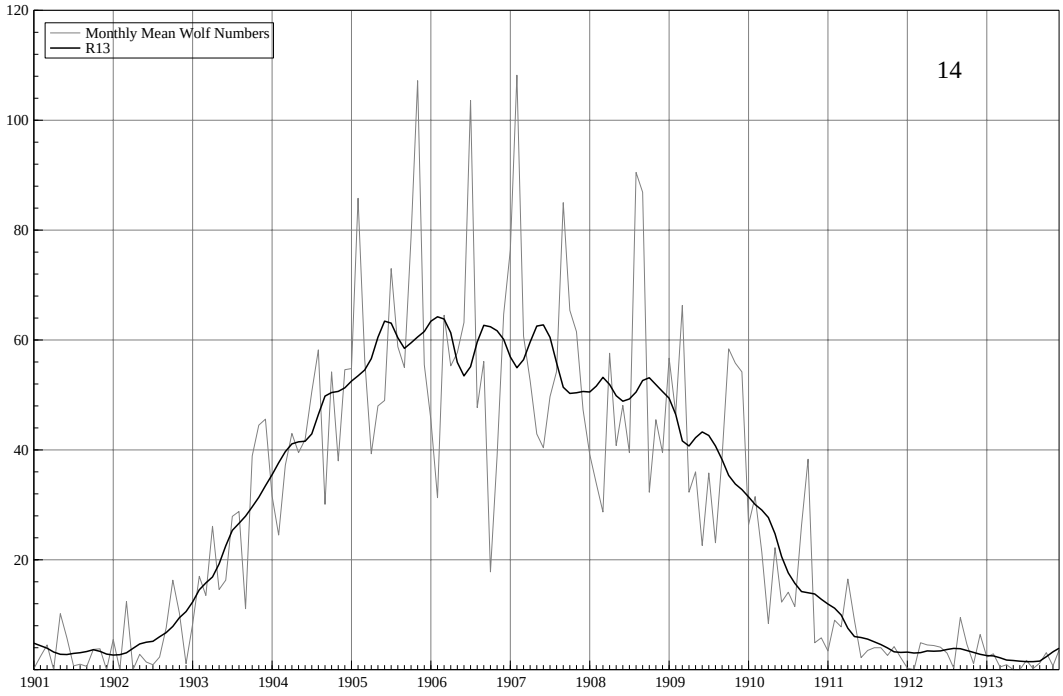


Abbildung 3.8: Verlauf der beobachteten Monatsmittel der Sonnenfleckenrelativzahl während Zyklus Nr. 14 sowie ausgeglichene Monatsmittel nach Wolf (R13).

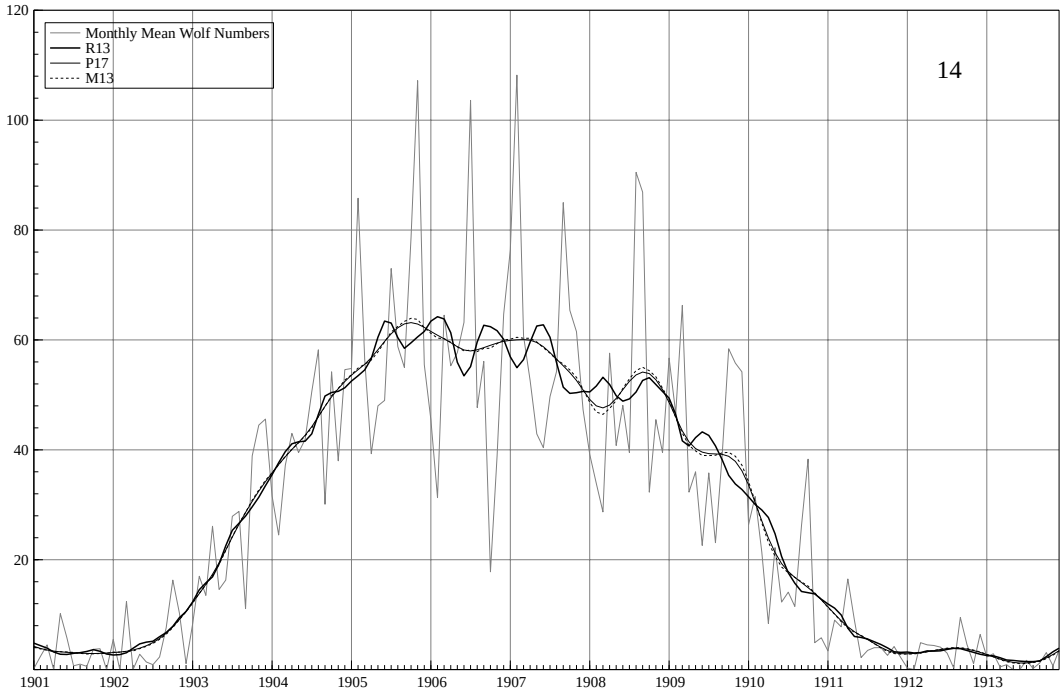


Abbildung 3.9: Vergleich der ausgeglichenen Monatsmittel nach Wolf (R13), Meeus (M13) und Karkoschka (P17).

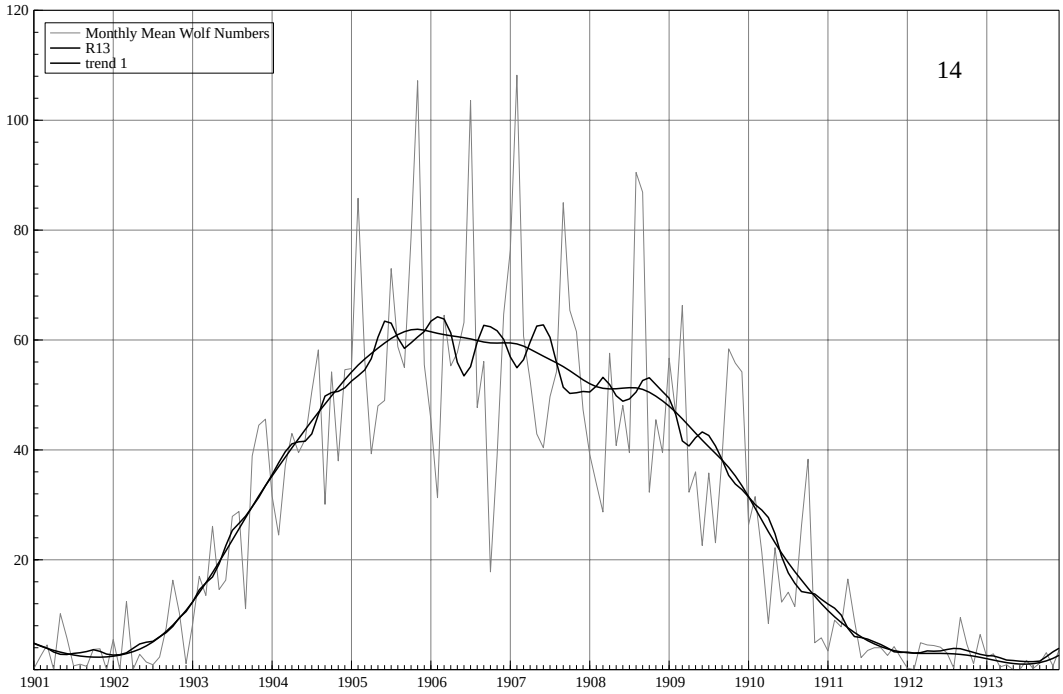


Abbildung 3.10: Vergleich der ausgeglichenen Monatsmittel nach Wolf (R13) mit der stochastischen Trendkomponente des strukturellen Zustandsraummodells (3.6).

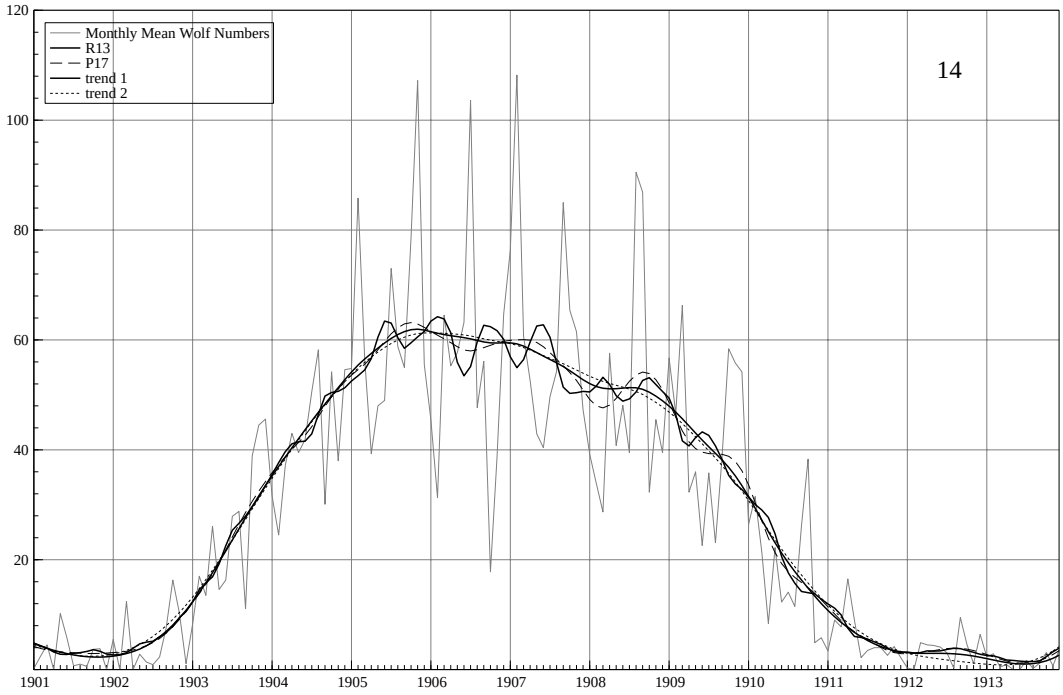


Abbildung 3.11: Vergleich der ausgeglichenen Monatsmittel nach Wolf (R13) und Karkoschka (P17) mit den stochastischen Trendkomponenten der Zustandsraummodelle (3.6) und (3.7).

Demgegenüber präsentiert sich der in Abbildung 3.10 gezeigte Verlauf der stochastischen Trendkomponente des strukturellen Zustandsraummodells (3.6) als nahezu ideal: Die Kurve glättet wie gewünscht die R13-Mittel quasi noch einmal aus und auch die wenigen vorhandenen Extremalstellen fallen mit denjenigen der beobachteten Monatsmittelwerte zusammen. Ein Vergleich mit dem in Abbildung 3.11 gezeigten Verlauf der stochastischen Trendkomponente des strukturellen Zustandsraummodells (3.7) zeigt, dass letztere tendenziell ein leichtes Oversmoothing aufweist, besonders in den Jahren 1902, 1905 und 1912.

Wir schliessen daraus, dass das einfache strukturelle Zustandsraummodell mit einer glatten Trend- und einer irregulären Restkomponente für die Darstellung und mithin für die Homogenitätsprüfung des mittleren Verlaufs der beobachteten Monatsmittel der Sonnenfleckenrelativzahl geeignet ist.

In Abbildung 6.5 ist der Verlauf der beobachteten Monatsmittel der Wolfschen Sonnenfleckenrelativzahl und der stochastischen Trendkomponente des strukturellen Zustandsraummodells (3.6) von 1749 bis 2004 wiedergegeben.

3.6 Zustandsraummodell für die Jahresmittel der Wolfschen Reihe

Wie in Kapitel 1.1 erwähnt, ist der in Abbildung 1.1 gezeigte Verlauf der Jahresmittel der Wolfschen Sonnenfleckenrelativzahl auch klimatologisch interessant. Allerdings stören die elfjährigen Aktivitätszyklen das Studium des langfristigen Aktivitätsverlaufs und damit die Diskussion eines möglichen *solar forcings* (Rind 2002). Ideal wäre ein langfristig glatter Verlauf.

Dieser kann durch ein strukturelles Zustandsraummodell extrahiert werden, indem die Jahresmittel der Wolfschen Reihe als Superposition einer glatten, langsam variierenden Trendkomponente und einer "weiter uninteressanten" zyklischen Restkomponente nach Formel (2.8) modelliert werden:

$$y_t = \mu_t + \psi_t. \tag{3.8}$$

In Abbildung 3.12 sind die Verläufe der beiden extrahierten Komponenten wiedergegeben. Der mittlere Aktivitätsverlauf zeigt die bekannten Maxima und Minima, insbesondere das Dalton-Minimum um 1800 herum, das moderne Maximum um 1950 sowie den raschen Aufstieg vom Maunder-Minimum ab 1700. Am Verlauf der zyklischen Komponente fällt auf, dass die Maxima und Minima der vorhandenen "stochastischen Schwebung" mit den Maxima und Minima der Trendkomponente zusammenfallen. Dies deutet auf eine Koppelung von Mittelwert und Variabilität hin und legt ein nichtlineares Zeitreihenmodell nahe. Durch die Dekomposition der Wolfschen Reihe in zwei unabhängige stochastische Komponenten kann ein solches allerdings elegant umgangen werden.

Eine eingehende Validierung anhand der in Abbildung 3.13 zusammengestellten Grafiken ergibt, dass die Innovationen abgesehen von einigen Ausreissern bei den Ausnahmezyklen 3, 8 und 19 recht gut normalverteilt sind. Auch die Homoskedastizität der Innovationen scheint gemäss dem Cusum-Plot der Innovationsquadrate gegeben. Signifikante Mittelwertsverschiebungen treten abgesehen von einer etwas problematischen Phase um 1800 gemäss dem Cusum-Plot der Innovationen nicht auf. Einzig das Korrelogramm zeigt eine persistente zyklische Autokorrelation der Innovationen mit einer Periode von 5-6 Jahren. Dies deutet auf eine Restvariabilität hin, welche durch keine der beiden vorhandenen Komponenten modelliert werden kann.

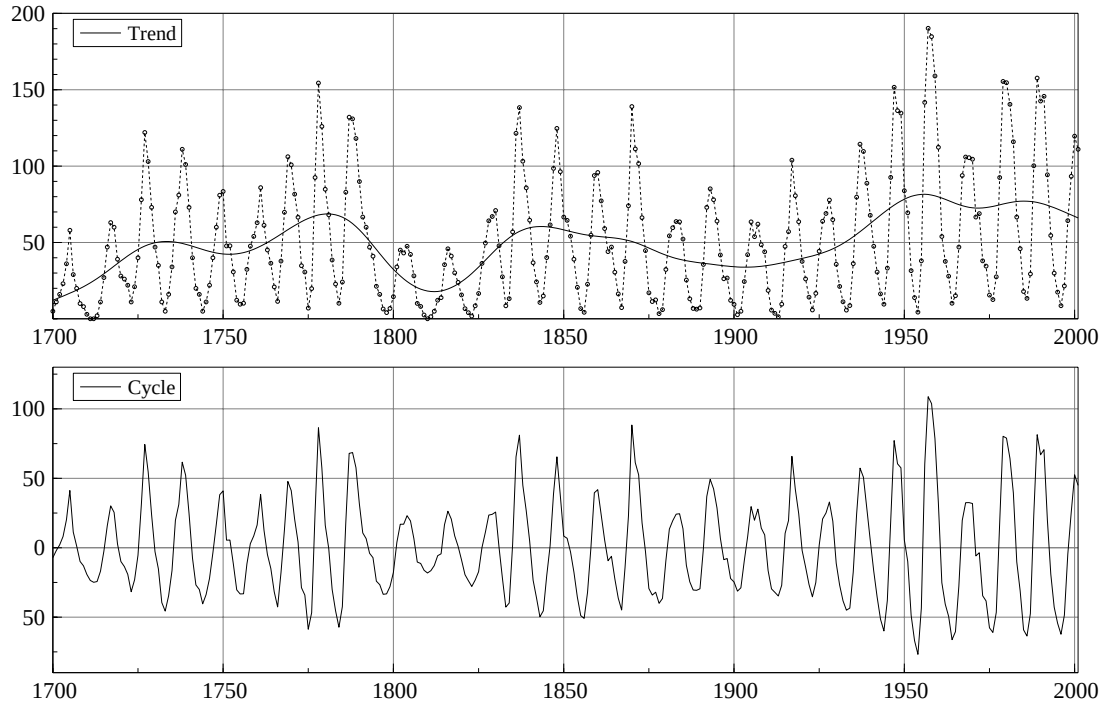


Abbildung 3.12: Dekomposition der Jahresmittel der Sonnenfleckenrelativzahl in einen stochastischen Trend und einen stochastischen Zyklus.

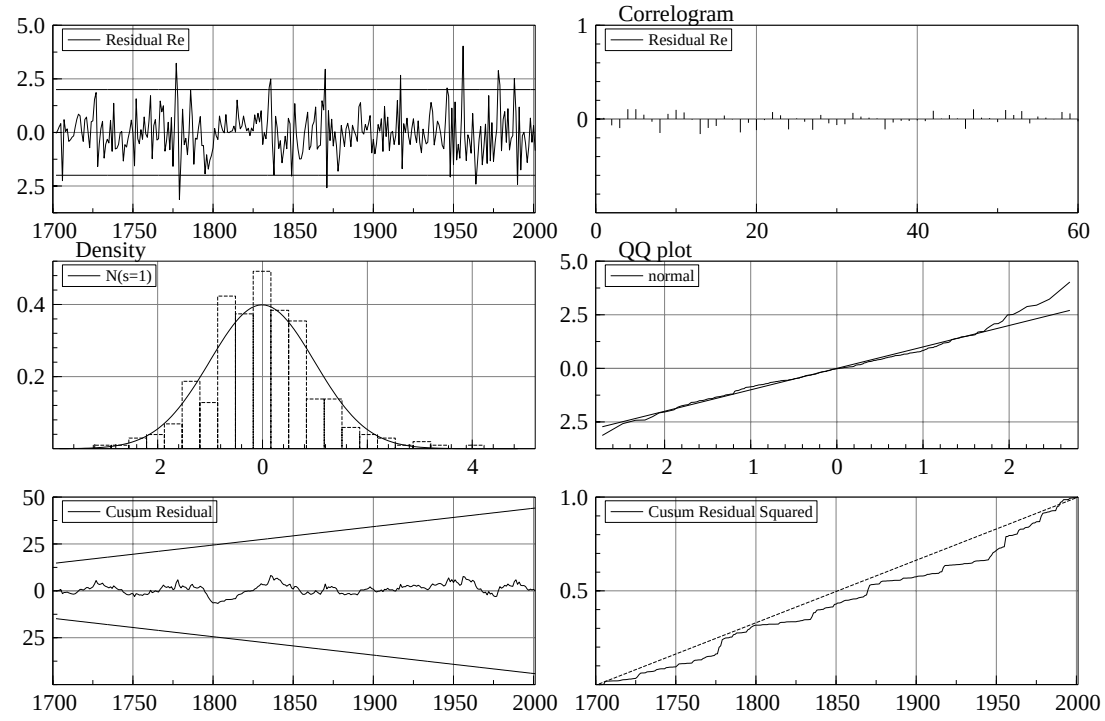


Abbildung 3.13: Diagnostik zu den Innovationen des Modells in Abbildung 3.12. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

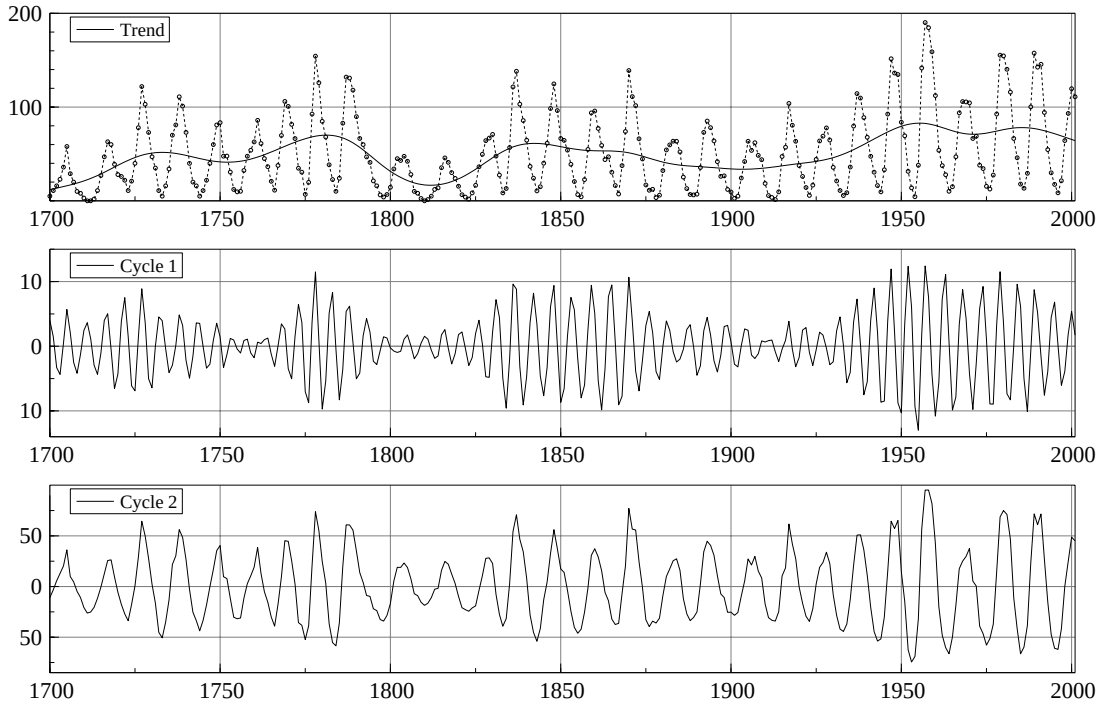


Abbildung 3.14: Dekomposition der Jahresmittel der Sonnenfleckenrelativzahl in einen stochastischen Trend und zwei stochastische Zyklen.

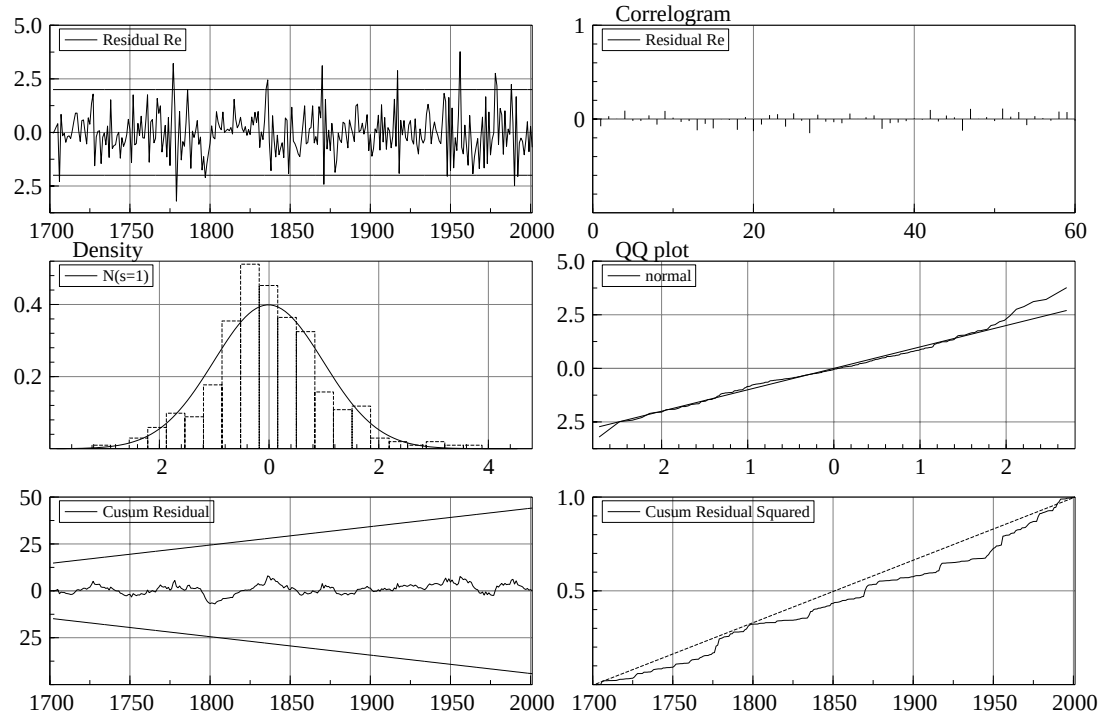


Abbildung 3.15: Diagnostik zu den Innovationen des Modells in Abbildung 3.14. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

Eine mögliche Verbesserung des Modells (3.8) besteht in der Berücksichtigung einer zusätzlichen zyklischen Komponente mit kürzerer Periode:

$$y_t = \mu_t + \psi_{1,t} + \psi_{2,t}. \tag{3.9}$$

In Abbildung 3.14 sind die drei extrahierten Komponenten dargestellt. Die zusätzliche zyklische Komponente erweist sich als eine harmonische des ersten Zyklus (Tabelle 3.2). Ihre Amplitude ist allerdings verschwindend klein, entspricht sie doch einer Aktivitätsschwankung von weniger als einer Sonnenfleckengruppe. Trotzdem hat sich ihre Mitberücksichtigung gelohnt: die Autokorrelationen in Abbildung 3.15 weisen nun ein zufälliges Muster auf.

Ein Vergleich der beiden extrahierten Trendkomponenten zeigt, dass die Mitberücksichtigung der zweiten zyklischen Komponente zu keinem wesentlich neuen Trendverlauf geführt hat: die Differenzen sind in der Grössenordnung eines einzelnen Sonnenflecks (Abb. 3.16). Trotzdem passt sich die Trendkomponente des zweiten Modells dem mittleren Verlauf der Zeitreihe besser an.

Das vorliegende Komponentenmodell (3.9) mit einem glatten, langsam variierenden Trend und zwei stochastischen Zyklen ist meines Wissens das erste strukturelle Zustandsraummodell dieser Art für die ansonsten oft analysierte Reihe der Jahresmittel der Wolfschen Sonnenfleckenrelativzahl (Tong 1990, Rao et al. 1997).

Modell	PEV	AIC	σ^2_ζ	$\sigma^2_{\psi,1}$	ρ_1	c_1	A_1	$\sigma^2_{\psi,2}$	ρ_2	c_2	A_2
(3.8)	257.91	5.59	0.26	141.54	0.94	10.50	46.29	-	-	-	-
(3.9)	244.30	5.55	0.31	108.89	0.95	10.64	45.97	5.84	0.93	5.37	4.96

Tabelle 3.2: Geschätzte Modellparameter und Kenngrössen der beiden Zustandsraummodelle (3.8) und (3.9). PEV bezeichnet die Varianz der Innovationen, AIC das Akaike Informationskriterium, c die Periodenlänge und A die Amplitude der stochastischen Zyklen.

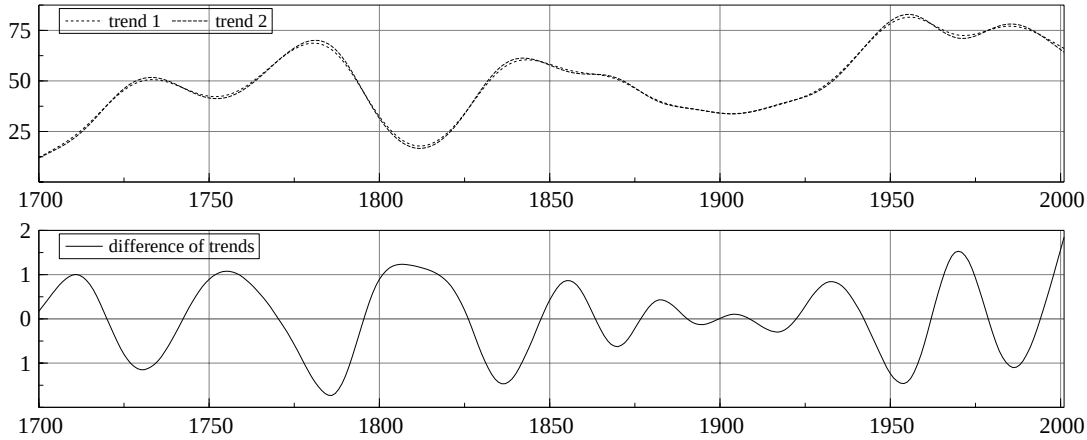


Abbildung 3.16: Differenz der geschätzten Trendkomponenten der beiden Modelle (3.8) und (3.9). Oben die beiden Trendverläufe, unten deren Differenz.

4. Homogenitätstests

In diesem Kapitel wird gezeigt, dass einige der gängigsten klimatologischen Homogenitätstests spezielle Likelihood Ratio Tests sind, beziehungsweise eng mit solchen verwandt sind.

Zusätzlich werden mit Hilfe von Monte-Carlo-Experimenten die kritischen Grenzen der Testgrössen, die Güte der Tests gegenüber einfachen Versatzinhomogenitäten sowie ihr Verhalten in autokorrelierten Datensätzen untersucht.

Als Alternativen betrachten wir lediglich einfache Mittelwertsverschiebungen in unabhängigen, normalverteilten Zufallsvariablen. Dagegen könnte berechtigterweise eingewendet werden, dass solche Testsituationen reichlich realitätsfremd sind, da klimatologische Beobachtungs- und Messreihen weder unabhängige noch normalverteilte Zufallsvariablen darstellen. Wir werden in Kapitel 6 jedoch zeigen, dass sich derartig einfache Testsituationen ergeben, falls wir als Zufallsvariablen spezielle Residuen von strukturellen Zustandsraummodellen betrachten.

4.1 Das Change - Point Problem

Seien in Anlehnung an Kapitel 1 von Csörgő und Horváth (1997) $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ unabhängige Zufallsvektoren in $\mathbb{R}^m$ mit den bekannten Verteilungsfunktionen $F(\boldsymbol{y}; \boldsymbol{\theta}_1, \boldsymbol{\eta}_1), \ldots, F(\boldsymbol{y}; \boldsymbol{\theta}_n, \boldsymbol{\eta}_n)$ worin $\boldsymbol{\theta}_i \in \Theta^{(1)} \subseteq \mathbb{R}^d$ und $\boldsymbol{\eta}_i \in \Theta^{(2)} \subseteq \mathbb{R}^p$, für alle $1 \leq i \leq n$.

Wir prüfen die Nullhypothese

$$H_0: \quad \boldsymbol{\theta}_1 = \cdots = \boldsymbol{\theta}_n \quad \text{und} \quad \boldsymbol{\eta}_1 = \cdots = \boldsymbol{\eta}_n$$

gegen die zweiseitige Alternativhypothese

$$\begin{aligned} H_1: \quad & \boldsymbol{\eta}_1 = \cdots = \boldsymbol{\eta}_n \quad \text{doch} \\ & \exists\, k^* \in \{1, \ldots, n-1\}, \quad \text{so dass} \\ & \boldsymbol{\theta}_1 = \cdots = \boldsymbol{\theta}_{k^*} \neq \boldsymbol{\theta}_{k^*+1} = \cdots = \boldsymbol{\theta}_n \end{aligned}$$

welche in der Literatur als “at most one change-point model” (AMOC) bekannt ist. $\boldsymbol{\eta}$ wird als fixer Störparametervektor betrachtet, während $\boldsymbol{\theta}$ unter der Alternative an der Stelle k^* sprunghaft ändert.

In der vorliegenden Arbeit betrachten wir lediglich univariate, unabhängig normalverteilte, reellwertige Zufallsvariablen $Y_1, \ldots, Y_n$ mit konstanter Varianz σ^2, wodurch sich obige Hypothesen wie folgt vereinfachen:

$$\begin{aligned} H_0: \quad & Y_i \sim \mathsf{IN}(\mu, \sigma^2) && 1 \leq i \leq n \\ H_1: \quad & \exists\, k^* \in \{1, \ldots, n-1\}, \text{ so dass} \\ & Y_i \sim \mathsf{IN}(\mu_1, \sigma^2) && 1 \leq i \leq k^* \\ & Y_i \sim \mathsf{IN}(\mu_2, \sigma^2) && k^*+1 \leq i \leq n \end{aligned}$$

mit $\mu_1 \neq \mu_2$.

4.2 Likelihood Ratio Testgrössen

Wäre die Epoche k^* der sprunghaften Veränderung bekannt, so läge ein Zweistichprobenproblem vor, welche wir beispielsweise mit einem Likelihood Ratio (LR) Test entscheiden könnten.

Die Likelihood Ratio lautet für $k^* = k$ im allgemeinen Fall

$$-2 \log \Lambda_k = 2\{L_k(\hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\eta}}_k) + L_k^*(\boldsymbol{\theta}_k^*, \hat{\boldsymbol{\eta}}_k) - L_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\eta}}_n)\} \tag{4.1}$$

mit den Maximum-Likelihood Schätzern

$$\begin{aligned} L_k(\hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\eta}}_k) &= \sup_{\theta,\eta} \sum_{i=1}^{k} \log f(\boldsymbol{y}_i; \boldsymbol{\theta}, \boldsymbol{\eta}) \\ L_k^*(\boldsymbol{\theta}_k^*, \hat{\boldsymbol{\eta}}_k) &= \sup_{\theta,\eta} \sum_{i=k+1}^{n} \log f(\boldsymbol{y}_i; \boldsymbol{\theta}, \boldsymbol{\eta}) \\ L_n(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\eta}}_n) &= \sup_{\theta,\eta} \sum_{i=1}^{n} \log f(\boldsymbol{y}_i; \boldsymbol{\theta}, \boldsymbol{\eta}) \end{aligned}$$

Für univariate, unabhängig normalverteilte, reellwertige Zufallsvariablen $Y_1, \ldots, Y_n$ mit bekannter konstanter Varianz σ^2 wird die Likelihood Ratio (4.1) zu

$$-2 \log \Lambda_k = 2\{k\bar{Y}_k^2/(2\sigma^2) + (n-k)\bar{Y}_k^{\circ 2}/(2\sigma^2) - n\bar{Y}_n^2/(2\sigma^2)\} \tag{4.2}$$

mit den Maximum-Likelihood Schätzern

$$\begin{aligned} \bar{Y}_k &= \frac{1}{k}\sum_{i=1}^{k} Y_i &&= \frac{1}{k} S(k) \\ \bar{Y}_k^{\circ} &= \frac{1}{n-k}\sum_{i=k+1}^{n} Y_i &&= \frac{1}{n-k}\left(S(n) - S(k)\right) \\ \bar{Y}_n &= \frac{1}{n}\sum_{i=1}^{n} Y_i &&= \frac{1}{n} S(n) \end{aligned}$$

wobei sich Gleichung (4.2) zu folgender Statistik umformen lässt:

$$\begin{aligned}
-2\log\Lambda_k &= \frac{1}{\sigma^2}\left\{\frac{1}{k}S(k)^2 + \frac{1}{n-k}\left(S(n)-S(k)\right)^2 - \frac{1}{n}S(n)^2\right\} \\
&= \frac{1}{\sigma^2}\frac{n}{k(n-k)}\left\{\frac{(n-k)}{n}S(k)^2 + \frac{k}{n}\left(S(n)-S(k)\right)^2 - \frac{k(n-k)}{n^2}S(n)^2\right\} \\
&= \frac{1}{\sigma^2}\frac{n}{k(n-k)}\left\{S(k)^2 - \frac{2k}{n}S(n)S(k) + \frac{k^2}{n^2}S(n)^2\right\} \\
&= \frac{1}{\sigma^2}\frac{n}{k(n-k)}\left\{S(k) - \frac{k}{n}S(n)\right\}^2 \\
&= \frac{1}{\sigma^2}\frac{n}{k(n-k)}\left\{\sum_{i=1}^{k}\left(Y_i - \bar{Y}_n\right)\right\}^2 \qquad (4.3)
\end{aligned}$$

Beispiele für die Likelihood Ratio von unabhängig exponential-, poisson- und binomialverteilten Zufallsvariablen finden sich in Csörgő und Horváth (1997, Kapitel 1.4).

Da die Sprungepoche k^* unbekannt ist, wird H_0 abgelehnt, falls die sog. "maximally selected" Likelihood Ratio Statistik

$$\boxed{Z_n = \max_{1\le k<n}\left(-2\log\Lambda_k\right)} \qquad (4.4)$$

gross wird. Je nachdem, ob in Gleichung (4.3) die Varianz σ^2 sowie der Mittelwert μ bekannt sind oder aus der Stichprobe geschätzt werden müssen, unterscheiden wir zwischen folgenden vier Testvarianten:

- LR1
$$\sqrt{Z_n} = \max_{1\le k<n}\left\{\frac{1}{\sigma}\left|\frac{1}{\sqrt{(n-k)}}\sum_{i=k+1}^{n}(Y_i-\mu)\right|\right\} \qquad (4.5)$$
$\mu =$ bekannt
$\sigma =$ bekannt

- LR2
$$\sqrt{Z_n} = \max_{1\le k<n}\left\{\frac{1}{s_k}\left|\frac{1}{\sqrt{(n-k)}}\sum_{i=k+1}^{n}(Y_i-\mu)\right|\right\} \qquad (4.6)$$
$\mu =$ bekannt
$$s_k^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k}(Y_i-\mu)^2 + \sum_{i=k+1}^{n}(Y_i-\bar{Y}_k^{\circ})^2\right) \qquad (4.7)$$
$$\bar{Y}_k^{\circ} = \frac{1}{n-k}\sum_{i=k+1}^{n}Y_i$$

- LR3
$$\sqrt{Z_n} = \max_{1\le k<n} \left\{ \frac{1}{\sigma} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n\right) \right| \right\} \tag{4.8}$$
$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$$
$$\sigma = \text{bekannt}$$

- LR4
$$\sqrt{Z_n} = \max_{1\le k<n} \left\{ \frac{1}{s_k} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n\right) \right| \right\} \tag{4.9}$$
$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$$
$$s_k^2 = \frac{1}{n-2}\left(\sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n} (Y_i - \bar{Y}_k^{\circ})^2 \right) \tag{4.10}$$
$$\bar{Y}_k = \frac{1}{k}\sum_{i=1}^{k} Y_i$$
$$\bar{Y}_k^{\circ} = \frac{1}{n-k}\sum_{i=k+1}^{n} Y_i$$

Während die Testvarianten LR1, LR2 und LR3 vor allem in der industriellen Qualitätskontrolle ihre Anwendung finden (Lauro et al. 2002), ist für die Homogenitätsprüfung klimatologischer Beobachtungs- und Messreihen einzig LR4 von praktischer Relevanz. In der Literatur sind für ihn denn auch weitere Formulierungsvarianten bekannt, von denen wir folgende drei weiter untersuchen wollen, welche sich von LR4 lediglich in der Berechnung der empirischen Varianz unterscheiden.

- LR5
$$\sqrt{Z_n} = \max_{1\le k<n} \left\{ \frac{1}{s_k} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n\right) \right| \right\} \tag{4.11}$$
$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$$
$$s_k^2 = \frac{1}{n}\left(\sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n} (Y_i - \bar{Y}_k^{\circ})^2 \right) \tag{4.12}$$
$$\bar{Y}_k = \frac{1}{k}\sum_{i=1}^{k} Y_i$$
$$\bar{Y}_k^{\circ} = \frac{1}{n-k}\sum_{i=k+1}^{n} Y_i$$

- LR6
$$\sqrt{Z_n} = \max_{1\leq k<n} \left\{ \frac{1}{s_k} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n\right) \right| \right\} \tag{4.13}$$

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

$$s_k^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2 \tag{4.14}$$

- LR7
$$\sqrt{Z_n} = \max_{1\leq k<n} \left\{ \frac{1}{s_k} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n\right) \right| \right\} \tag{4.15}$$

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

$$s_k^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2 \tag{4.16}$$

Erwähnenswert ist, dass nur LR1 und LR3 maximally selected Likelihood Ratio Statistiken im Sinne von Gleichung (4.4) sind: Ist die Varianz σ^2 zwar konstant jedoch unbekannt, sollte sie in Gleichung (4.1) korrekterweise als zu schätzender Störparameter eingeführt werden, was zu einer Likelihood Ratio von

$$-2 \log \Lambda_k = n \left\{ \log \hat{\sigma}_n^2 - \log \hat{\sigma}_k^2 \right\} \tag{4.17}$$

führt, mit der empirischen Varianz

$$\hat{\sigma}_k^2 = \frac{1}{n} \left(\sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n} (Y_i - \bar{Y}_k^{\circ})^2 \right).$$

Csörgő und Horváth (1997, Formeln 1.4.26-1.4.29) zeigen, dass (4.17) äquivalent zu (4.3) ist und dass die asymptotischen Verteilungen von (4.8), (4.11) und (4.13) dieselben sind. Wie leicht zu erkennen ist, sind auch die Verteilungen von (4.15) und (4.13) sowie von (4.9) und (4.11) dieselben.

Abbildung 4.1 zeigt den Verlauf der Testgrössen LR1 bis LR7 für eine simulierte Beobachtungsreihe $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 80$. Die Maxima der Testgrössen bei $k = 38$ markieren mögliche Bruchstellen. Hierzu müssten die Maxima der Testgrössen allerdings die von der Testvariante, der Länge der Beobachtungsreihen und dem Signifikanzniveau α abhängige kritische Grenze $C_{LR}(80, 0.05) \approx 3$ übertreffen, was nicht der Fall ist.

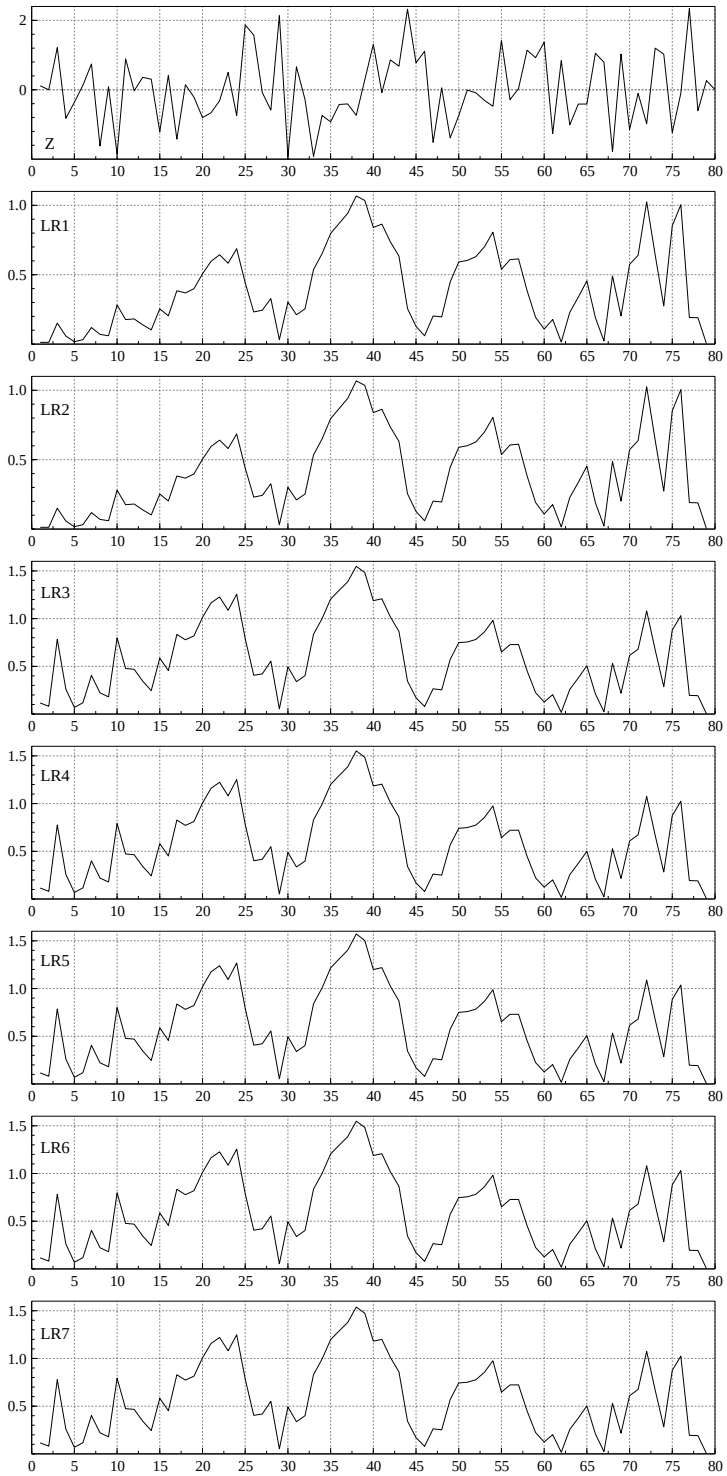


Abbildung 4.1: Simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 80$ mit korrespondierenden Werten der Testgrössen LR1 bis LR7. Die Maxima der Testgrössen bei $k = 38$ markieren mögliche Bruchstellen. Hierzu müssten die Maxima allerdings die kritische Grenze $C_{LR}(N,\alpha) = C_{LR}(80, 0.05) \approx 3$ übertreffen, was nicht der Fall ist.

4.3 Der Alexandersson Test

Der Klimatologe Hans Alexandersson entwickelte 1984 basierend auf Hawkins (1977) einen Test zur Homogenitätsprüfung von Niederschlagsdaten in Schweden (Alexandersson, 1986). Schon bald wurde dieser Test auch zur Homogenitätsprüfung anderer klimatologischer Messreihen verwendet. Seit den frühen neunziger Jahren ist der Test unter dem Namen "Standard-Normal-Homogeneity Test" (SNHT) einer der Standardtests zur Homogenitätsprüfung von klimatologischen Zeitreihen und wird von zahlreichen nationalen Wetterdiensten zur Homogenisierung ihrer Überwachungsreihen eingesetzt (Peterson et al. 1998).

4.3.1 Definition der Testgrösse

Sei $z_t = z_1, z_2, \ldots, z_n$ eine mittels dem empirischen Mittelwert $\bar{x}$ und der empirischen Standardabweichung s_x aus einer univariaten **Beobachtungsreihe** x_t standardisierte **Testreihe** wobei x_t nach Alexandersson (1986) je nach klimatologischer Messgrösse eine Reihe der Differenzen oder eine Reihe der Quotienten zwischen korrespondierenden Beobachtungen benachbarter Messstationen darstellt.

Der Alexandersson Test prüft die Hypothesen

$$\begin{aligned} H_0: &\quad Z_i \sim \mathsf{IN}(0,1), \forall i \\ H_1: &\quad \begin{cases} \text{Es gibt ein } k,\ 1 \leq k < n, \text{ so dass für } \mu_1 \neq \mu_2 \text{ gilt} \\ Z_i \sim \mathsf{IN}(\mu_1, 1) \text{ für } i \leq k \text{ und} \\ Z_i \sim \mathsf{IN}(\mu_2, 1) \text{ für } i > k. \end{cases} \end{aligned} \tag{4.18}$$

Als **Testgrösse** dient

$$\boxed{T_0 = \max_{1 \leq k < n} \{T_k\} = \max_{1 \leq k < n} \{k\bar{z}_1^2 + (n-k)\bar{z}_2^2\}} \tag{4.19}$$

mit

$$\bar{z}_1 = \frac{1}{k}\sum_{i=1}^{k} z_i \quad \text{und} \quad \bar{z}_2 = \frac{1}{n-k}\sum_{i=k+1}^{n} z_i$$

Überschreitet T_0 eine vom Umfang n der Testreihe Z_t und dem Signifikanzniveau α abhängige kritische Grenze $C_A(n, \alpha)$, so wird die Reihe als inhomogen betrachtet. Der Zeitpunkt k bei welchem die Testgrösse ihr Maximum annimmt, gilt als Schätzung der Bruchepoche, die Differenz der Mittelwerte vor und nach dem Bruchpunkt als Schätzung der Sprunghöhe.

Abbildung 4.2 zeigt den Verlauf der Testgrösse T_k für eine simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 80$. Das Maximum T_0 der Testgrössen T_k bei $k = 38$ markiert eine mögliche Bruchstelle. Hierzu müsste T_0 allerdings die kritische Grenze $C_A(80, 0.05) \approx 9.0$ übertreffen, was nicht der Fall ist.

4.3.2 Der Alexandersson Test als Likelihood Ratio Test

Zwar entwickelt Alexandersson (1986) im zweiten Abschnitt seiner Arbeit die Statistik (4.19) aus einem Likelihood Ratio Ansatz, doch lässt sich die Herleitung seiner Teststatistik (5) auf dem angegebenen Weg nur unvollständig nachvollziehen. Wir zeigen daher

Satz 4.1 *Unter den Modellannahmen (4.18) gilt:* $T_0 = Z_n$.

Beweis Durch elementare Umformungen erhalten wir aus (4.19)

$$
\begin{aligned}
T_0 &= \max_{1\le k<n} \left\{ k\bar{z}_1^2 + (n-k)\bar{z}_2^2 \right\} \\
&= \max_{1\le k<n} \left\{ k \left(\frac{1}{k}\sum_{i=1}^{k} z_i \right)^2 + (n-k) \left(\frac{1}{n-k}\sum_{i=k+1}^{n} z_i \right)^2 \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{k} \left(\sum_{i=1}^{k} z_i \right)^2 + \frac{1}{n-k} \left(\sum_{i=k+1}^{n} z_i \right)^2 \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{k} \left(\frac{1}{s_n}\sum_{i=1}^{k} (Y_i - \bar{Y}_n) \right)^2 + \frac{1}{n-k} \left(\frac{1}{s_n}\sum_{i=k+1}^{n} (Y_i - \bar{Y}_n) \right)^2 \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{s_n^2} \left\{ \frac{1}{k} \left(\sum_{i=1}^{k} Y_i - k\bar{Y}_n \right)^2 + \frac{1}{n-k} \left(\sum_{i=k+1}^{n} Y_i - (n-k)\bar{Y}_n \right)^2 \right\} \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{s_n^2} \frac{n}{k(n-k)} \left\{ \frac{n-k}{n} \left(S(k)^2 - \frac{2k}{n} S(n)S(k) + \frac{k^2}{n^2} S(n)^2 \right) \right. \right. \\
&\qquad \left. \left. + \frac{k}{n} \left((S(n) - S(k))^2 - 2\left(S(n) - S(k)\right) \frac{n-k}{n} S(n) + \frac{(n-k)^2}{n^2} S(n)^2 \right) \right\} \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{s_n^2} \frac{n}{k(n-k)} \left\{ S(k)^2 - \frac{2k}{n} S(n)S(k) + \frac{k^2}{n^2} S(n)^2 \right\} \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{s_n^2} \frac{n}{k(n-k)} \left\{ S(k) - \frac{k}{n} S(n) \right\}^2 \right\} \\
&= \max_{1\le k<n} \left\{ \frac{1}{s_n^2} \frac{n}{k(n-k)} \left\{ \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n \right) \right\}^2 \right\} \\
&= \left(\max_{1\le k<n} \left\{ \frac{1}{s_n} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n \right) \right| \right\} \right)^2 = Z_n
\end{aligned}
$$

□

In Alexandersson (1986) wird s_n als "sample standard deviation" bezeichnet, worunter üblicherweise die Quadratwurzel aus (4.16) verstanden wird. **Der Alexandersson Test ist somit äquivalent zum LR7 Test, wobei die Alexandersson Teststatistik das Quadrat der LR7 Teststatistik ist.**

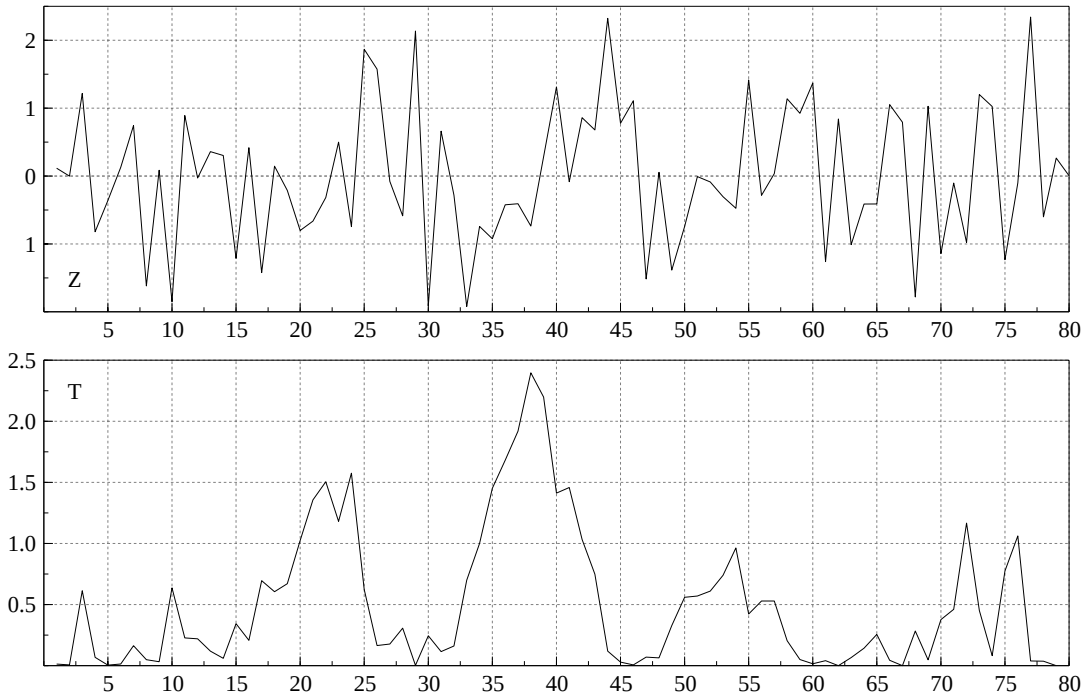


Abbildung 4.2: Simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 80$ mit korrespondierenden Werten der Testgrösse T_k nach Alexandersson. Das Maximum T_0 der Testgrössen T_k bei $k = 38$ markiert eine mögliche Bruchstelle. Hierzu müsste T_0 allerdings die kritische Grenze $C_A(N,\alpha) = C_A(80, 0.05) \approx 9.0$ übertreffen, was hier jedoch nicht der Fall ist.

4.4 Der Buishand Test

Der Klimatologe T.A. Buishand entwickelte 1982 einen Test zur Homogenitätsprüfung von Niederschlagsdaten in den Niederlanden (Buishand 1982).

4.4.1 Definition der Testgrösse

Der Buishand Test prüft die Hypothesen

$$
\begin{aligned}
H_0: \quad & Y_i \sim \mathsf{IN}(\mu, \sigma^2) && 1 \le i \le n \\
H_1: \quad & \exists\, k^* \in \{1, \ldots, n-1\}, \text{ so dass} \\
& Y_i \sim \mathsf{IN}(\mu, \sigma^2) && 1 \le i \le k^* \\
& Y_i \sim \mathsf{IN}(\mu + \Delta, \sigma^2) && k^* + 1 \le i \le n
\end{aligned}
$$

mit $\Delta \neq 0$.

Als **Testgrösse** dient

$$\boxed{Q = \max_{1 \le k < n} \{Q_k\} = \max_{1 \le k < n} \left\{ \left| \frac{1}{D_Y} \sum_{i=1}^{k} (Y_i - \bar{Y}) \right| \right\}} \tag{4.20}$$

mit

$$D_Y^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \quad \text{und} \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

Überschreitet Q eine vom Umfang n der Messreihe Y_t und dem Signifikanzniveau α abhängige kritische Grenze $C_B(n, \alpha)$, so wird die Reihe als inhomogen betrachtet. Der Zeitpunkt k bei welchem die Testgrösse ihr Maximum annimmt, gilt als Schätzung der Bruchepoche, die Differenz der Mittelwerte vor und nach dem Bruchpunkt als Schätzung der Sprunghöhe.

Abbildung 4.3 zeigt den Verlauf der Testgrösse Q_k für eine simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0, 1)$ der Länge $n = 80$. Das Maximum Q der Testgrössen Q_k bei $k = 38$ markiert eine mögliche Bruchstelle. Hierzu müsste Q allerdings die kritische Grenze $C_B(80, 0.05) \approx 11.0$ übertreffen, was nicht der Fall ist.

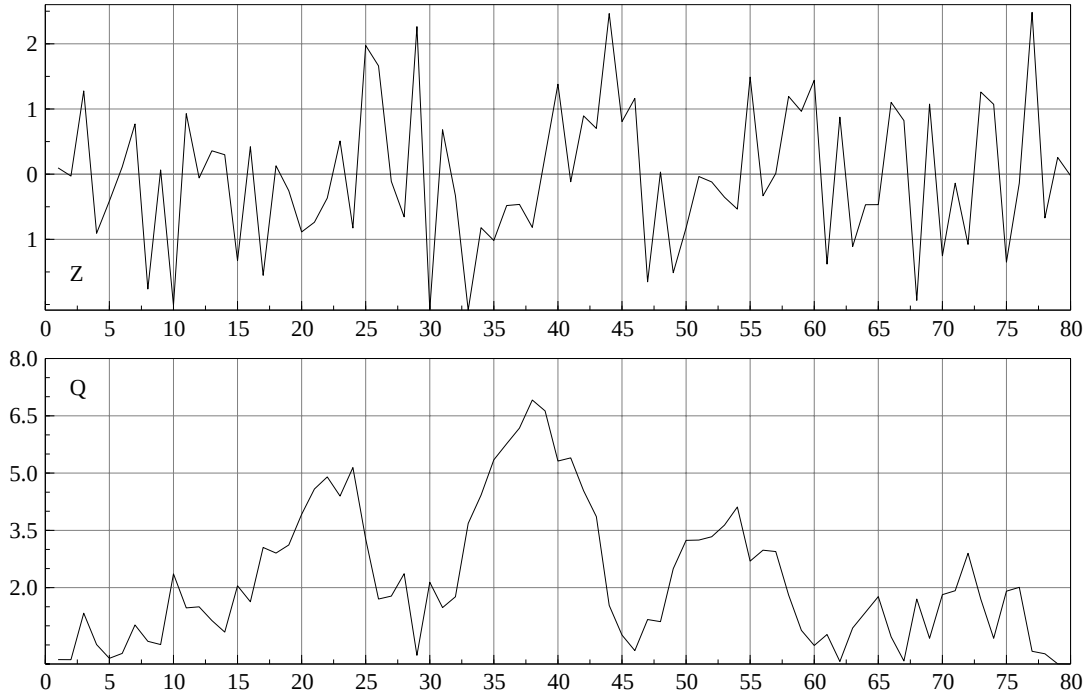


Abbildung 4.3: Simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0, 1)$ der Länge $N = 80$ mit korrespondierenden Werten der Testgrösse Q_k nach Buishand. Das Maximum Q der Testgrösse Q_k bei $k = 38$ markiert eine mögliche Bruchstelle. Hierzu müsste Q allerdings die kritische Grenze $C_B(N, \alpha) = C_B(80, 0.05) \approx 11.0$ übertreffen, was nicht der Fall ist.

4.4.2 Der Buishand Test als Likelihood Ratio Test

Aus dem Vergleich der Formeln (4.13) und (4.20) folgt:

$$\frac{1}{s_k}\left|\sqrt{\frac{n}{k(n-k)}}\sum_{i=1}^{k}\left(Y_i-\bar{Y}_n\right)\right| = \sqrt{\frac{n}{k(n-k)}}\,Q_k$$

Die Testgrösse Q_k ist also bis auf einen Vorfaktor identisch mit derjenigen des LR6 Tests. Allerdings ist dieser von k abhängig und daher ist

$$\max_{1\le k<n}\left\{\frac{1}{s_k}\left|\sqrt{\frac{n}{k(n-k)}}\sum_{i=1}^{k}\left(Y_i-\bar{Y}_n\right)\right|\right\} \neq \sqrt{\frac{n}{k(n-k)}}\max_{1\le k<n}\left\{Q_k\right\}.$$

Der Buishand Test ist somit kein "echter" Likelihood Ratio Test sondern lediglich eng mit diesen verwandt.

4.5 Worsley's Likelihood Ratio Test

In Buishand (1982) sowie in Herzog und Müller-Westermeier (1998) wird ein auf Worsley (1979) zurückgehender Test beschrieben, von welchem sich einfach zeigen lässt, dass es sich tatsächlich um einen "echten" Likelihood Ratio Test handelt.

Worsley (1979) prüft unter der Annahme, dass $\mu_1 \neq \mu_2$ die Hypothesen

$$\begin{aligned}
H_0: &\quad Y_i \sim \mathsf{IN}(\mu_1,\sigma^2) && 1\le i\le n\\
H_1: &\quad \exists\, k^*\in\{1,\ldots,n-1\}\text{, so dass}\\
&\quad Y_i \sim \mathsf{IN}(\mu_1,\sigma^2) && 1\le i\le k^*\\
&\quad Y_i \sim \mathsf{IN}(\mu_2,\sigma^2) && k^*+1\le i\le n
\end{aligned}$$

mit der Teststatistik

$$\boxed{W=\max_{1\le k\le n-1}\left\{\sqrt{(n-2)}\,|T_k|\,/S_k\right\}} \tag{4.21}$$

wobei

$$\begin{aligned}
T_k^2 &= (k(n-k)/n)\left(\bar{Y}_k-\bar{Y}_k^{\circ}\right)^2\\
S_k^2 &= \sum_{i=1}^{k}(Y_i-\bar{Y}_k)^2+\sum_{i=k+1}^{n}(Y_i-\bar{Y}_k^{\circ})^2\\
Y_k^{\circ} &= \frac{1}{n-k}\sum_{i=k+1}^{n}Y_i
\end{aligned}$$

Durch elementare Umformungen erhalten wir aus

$$\begin{aligned}
T_k^2 &= \frac{k(n-k)}{n}\left(\bar{Y}_k - \bar{Y}_k^\circ\right)^2 \\
&= \frac{k(n-k)}{n}\left(\frac{1}{k}\sum_{i=1}^{k} Y_i - \frac{1}{n-k}\sum_{i=k+1}^{k} Y_i\right)^2 \\
&= \frac{k(n-k)}{n}\left(\frac{1}{k}S(k) + \frac{1}{n-k}S(k) - \frac{1}{n-k}S(n)\right)^2 \\
&= \frac{k(n-k)}{n}\left(\frac{n}{k(n-k)}S(k) - \frac{k}{k(n-k)}S(n)\right)^2 \\
&= \frac{n}{k(n-k)}\left\{S(k)^2 - \frac{2k}{n}S(n)S(k) + \frac{k^2}{n^2}S(n)^2\right\} \\
&= \frac{n}{k(n-k)}\left\{S(k) - \frac{k}{n}S(n)\right\}^2 \\
&= \frac{n}{k(n-k)}\left\{\sum_{i=1}^{k}\left(Y_i - \bar{Y}_n\right)\right\}^2
\end{aligned}$$

und somit

$$W = \max_{1\le k\le n-1}\left\{\frac{1}{s_k}\left|\sqrt{\frac{n}{k(n-k)}}\sum_{i=1}^{k}\left(Y_i - \bar{Y}_n\right)\right|\right\}$$

mit

$$s_k^2 = \frac{1}{n-2}\left(\sum_{i=1}^{k}(Y_i - \bar{Y}_k)^2 + \sum_{i=k+1}^{n}(Y_i - \bar{Y}_k^\circ)^2\right)$$

was dem **LR4 Test** (4.9) entspricht.

4.6 Kritische Grenzen

Zur Berechnung kritischer Grenzen müssen die Verteilungen der Testatstistiken unter der Nullhypothese H_0 bekannt sein.

4.6.1 Bonferroni Approximation

Wird σ wie in LR1 und LR3 als bekannt vorausgesetzt oder wie in LR6 und LR7 nach Formeln des Typs (4.14) oder (4.16) geschätzt, so gilt für die einzelnen Likelihood Ratio Testgrössen:

$$T_k \sim N(0,1)$$

was beispielsweise anhand des LR1 Tests einfach demonstriert werden kann:

$$T_k = \frac{1}{\sqrt{(n-k)}}\sum_{i=k+1}^{n}\left(\frac{Y_i - \mu}{\sigma}\right) \sim N(0,1).$$

Wird σ jedoch nach den Formeln (4.7), (4.10) oder (4.12) geschätzt, so gilt für die Testgrössen (4.6), (4.9) und (4.11)

$$T_k \sim t(\nu)$$

was anhand der Herleitung von (4.21) gezeigt werden kann, entsprechen die W_k doch einer unverbundenen Zweistichproben t-Testgrösse zur Prüfung eines Mittelwertunterschiedes zwischen $\bar{Y}_k$ und $\bar{Y}_k^\circ$:

$$W_k = \sqrt{(n-2)}\,|T_k|\,/S_k \sim t(n-2).$$

Die kritischen Grenzen der Likelihood Ratio Tests können daher aufgrund der Bonferroni-Beziehung

$$P\left[\max_{1\le k\le n-1}|T_k| > c\right] \quad \le \quad (n-1)\,P\left[\,|T_1| > c\,\right]$$

durch entsprechende Quantile der Standardnormalverteilung bzw. einer t-Verteilung angenähert werden (Tabelle 4.1).

4.6.2 Asymptotische Verteilung

Nach Leadbetter et al. (1983, Theorem 1.5.3) ist die asymptotische Verteilung des Maximums einer Folge von unabhängigen, identisch standardnormalverteilten Zufallsvariablen eine Gumbelverteilung (Extremwertverteilung Typ I). Die T_k der Likelihood Ratio Tests sind jedoch nicht unabhängig, so gilt etwa für den LR1 Test (Antoch et al. 2002):

$$\rho_{T_k,T_l} = \sqrt{\frac{n-l}{n-k}}, \qquad k \le l$$

und für den LR4 Test (Hawkins 1977):

$$\rho_{T_k,T_l} = \sqrt{\frac{k(n-l)}{l(n-k)}}, \qquad k \le l.$$

Es zeigt sich, dass die Verteilungen der einzelnen Likelihood Ratio Teststatistiken nach geeigneter Normalisierung trotzdem gegen eine Extremwertverteilung konvergieren, wie den nachfolgenden Sätzen (4.22) und (4.23) sowie der untenstehenden Tabelle 4.1 entnommen werden kann.

Test	Bonferroni	Asymptotisch
LR1	$z_{1-\alpha/(2n-2)}$	(4.22)
LR2	$t_{1-\alpha/(2n-2)}(n-1)$	(4.22)
LR3	$z_{1-\alpha/(2n-2)}$	(4.23)
LR4	$t_{1-\alpha/(2n-2)}(n-2)$	(4.23)
LR5	$t_{1-\alpha/(2n-2)}(n-2)$	(4.23)
LR6	$z_{1-\alpha/(2n-2)}$	(4.23)
LR7	$z_{1-\alpha/(2n-2)}$	(4.23)

Tabelle 4.1: Formeln zur Berechnung approximativer kritischer Grenzen der Teststatistiken.

Satz 4.2 *Für Likelihood Ratio Testgrössen mit* **bekanntem** μ *gilt:*

$$\boxed{\lim_{n\to\infty} P\left(\sqrt{Z_n} \leq \frac{x+b_n}{a_n}\right) = \exp(-e^{-x})} \tag{4.22}$$

mit $a_n = \sqrt{2\log\log n}$ *und* $b_n = 2\log\log n + \frac{1}{2}\log\log\log n - \frac{1}{2}\log\pi.$

Beweis Darling und Erdős (1956, Theorem 2) und Csörgő und Horváth (1997, Theorem A.3.1).

Satz 4.3 *Für Likelihood Ratio Testgrössen mit* **unbekanntem** μ *gilt:*

$$\boxed{\lim_{n\to\infty} P\left(\sqrt{Z_n} \leq \frac{x+b_n}{a_n}\right) = \exp(-2e^{-x})} \tag{4.23}$$

mit $a_n = \sqrt{2\log\log n}$ *und* $b_n = 2\log\log n + \frac{1}{2}\log\log\log n - \frac{1}{2}\log\pi.$

Beweis Yao and Davis (1986, Theorem 2.1) und Csörgő und Horváth (1997, Theorem 1.3.1).

Nach Leadbetter et al. (1983, p.92) konvergieren die Verteilungen der Teststatistiken allerdings sehr langsam gegen die asymptotischen Extremwertverteilungen, so dass für die praktische Anwendung weniger konservative kritische Grenzen bestimmt werden müssen.

4.6.3 Simulation

Bevor wir mit der Simulierung kritischer Grenzen beginnen, schätzen wir anhand des Alexandersson Tests ab, wieviele Replikationen mit welchem Zufallszahlengenerator durchgeführt werden müssen.

Kritische Grenzen für kurze Zeitreihen mit n bis 1000 wurden von Alexandersson (1986) mittels Monte-Carlo-Experimenten bestimmt (Abbildung 4.4). Kritische Grenzen für längere Zeitreihen mit n zwischen 1'250 und 20'000 wurden von Baudenbacher (1997) basierend auf je 10'000 standardnormalverteilten Zufallsreihen simuliert (Tabelle 4.2).

Ein aufmerksamer Vergleich der in Abbildung 4.4 und in Tabelle 4.2 wiedergegebenen kritischen Grenzen ergibt, dass die von Alexandersson in Abbildung 4.4 angedeutete Annäherung der kritischen Grenzen an einen asymptotischen Grenzwert von Baudenbacher nicht bestätigt werden kann. Umgekehrt steigen die von Baudenbacher in Tabelle 4.2 simulierten kritischen Grenzen mit n nicht wie erwartet monoton an.

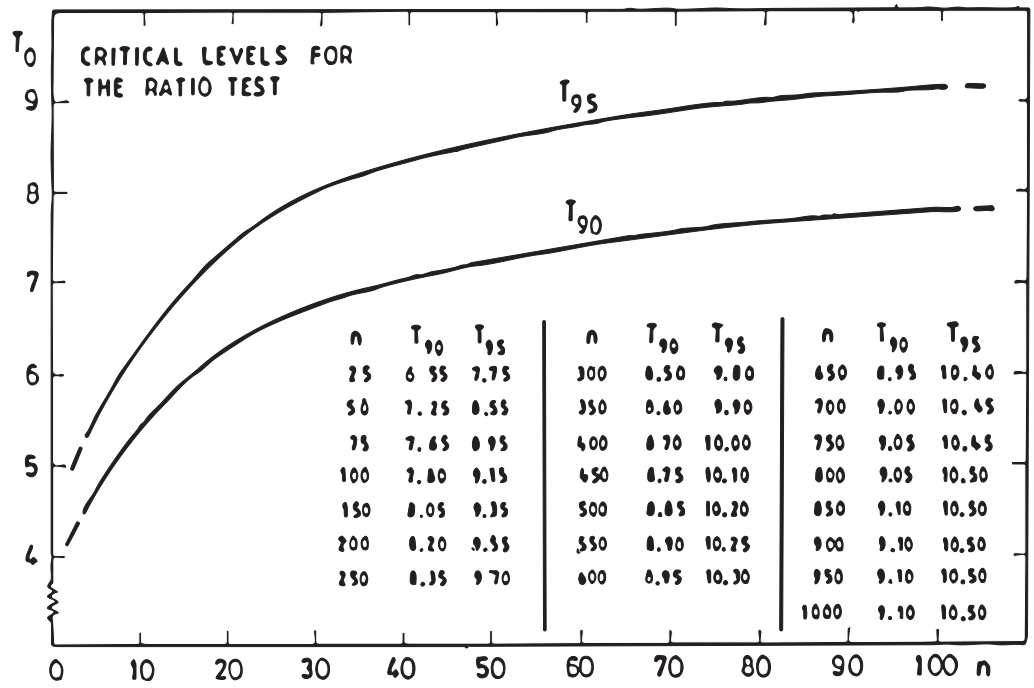


Abbildung 4.4: Kritische Grenzen des Alexandersson Tests.
Aus Alexandersson (1986).

n	T_{75}	T_{90}	T_{95}	T_{99}
1250	6.90	9.03	10.57	14.22
1500	7.02	9.12	10.64	13.96
1524	7.11	9.18	10.74	13.98
1750	7.11	9.23	10.74	14.04
2000	7.20	9.39	10.95	14.24
2500	7.19	9.29	10.78	14.38
3000	7.30	9.46	11.10	14.53
4000	7.41	9.54	11.09	14.27
5000	7.46	9.58	11.07	14.51
6000	7.46	9.58	11.15	14.45
7000	7.57	9.78	11.29	14.66
8000	7.64	9.83	11.33	14.62
9000	7.59	9.73	11.20	14.71
10000	7.72	9.85	11.51	14.74
15000	7.79	10.00	11.46	14.87
20000	7.91	10.11	11.59	14.95

Tabelle 4.2: Kritische Grenzen des Alexandersson Tests.
Aus Baudenbacher (1997).

Um uns ein unabhängiges Bild vom Verlauf des 95%-Quantils von T_0 für kurze Zeitreihen machen zu können, führten wir vier Monte-Carlo-Experimente mit verschiedener Anzahl Simulationen pro n durch, deren Ergebnisse in Abbildung 4.5 wiedergegeben sind.

Hierbei fällt sofort auf, dass für eine hinreichend genaue Schätzung der kritischen Grenzen offensichtlich eine mit n zunehmende, ziemlich grosse Anzahl von Replikationen notwendig ist.

Weniger offensichtlich ist, dass die simulierten kritischen Grenzen mit steigender Anzahl Replikationen nicht nur genauer werden, sondern auch tendenziell grösser. Dies deutet auf eine schiefe Verteilung der simulierten kritischen Grenzen für ein fixes N hin.

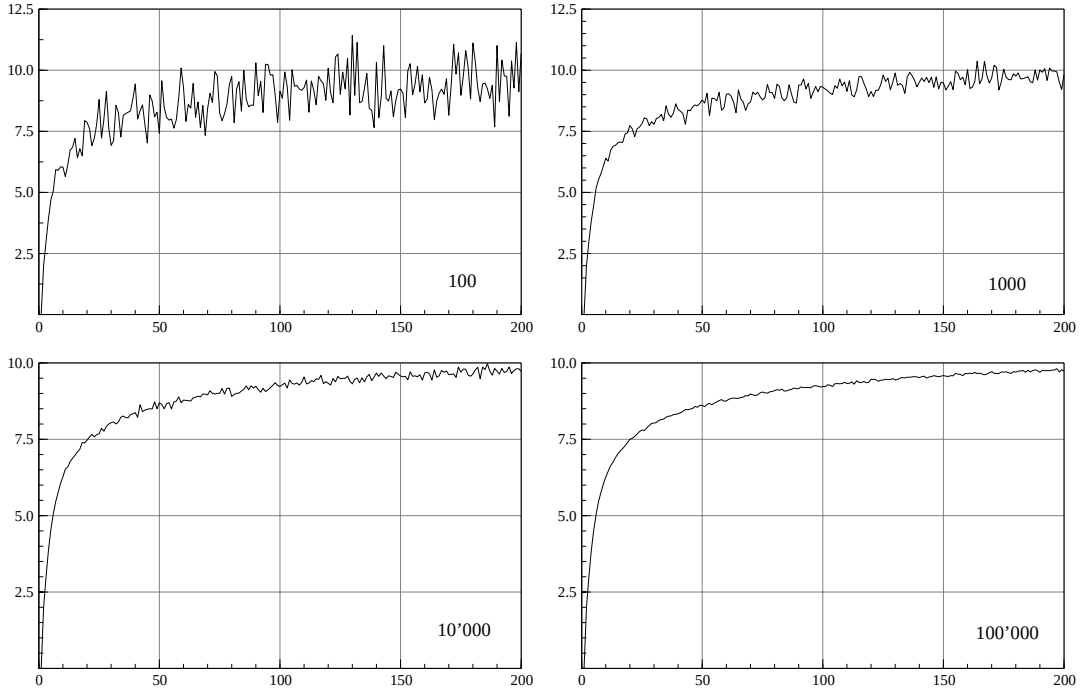


Abbildung 4.5: Vier Monte-Carlo-Simulationen des Verlaufs des 95%-Quantils von Alexandersson's T_0 für kurze Zeitreihen mit n zwischen 2 und 200. Für jedes n wurden je 100, 1000, 10'000 bzw. 100'000 Zeitreihen $Z_t \sim \mathsf{IN}(0, 1)$ simuliert und deren T_{95} berechnet.

Um einen detaillierteren Einblick in die Variabilität und Konvergenz von T_{95} zu erhalten, wurde für ein fixes $n = 100$ mittels vier Monte-Carlo-Experimenten die empirische Verteilung von 1000 unabhängigen Schätzungen von T_{95} bestimmt. Den Einzelschätzungen von T_{95} lagen pro Experiment je 100, 1000, 10'000 bzw. 100'000 Replikationen zugrunde. Die resultierenden vier Verläufe der Einzelschätzungen von T_{95}, ihre Verteilungen sowie Mittelwerte und Standardfehler können der Abbildung 4.6 entnommen werden.

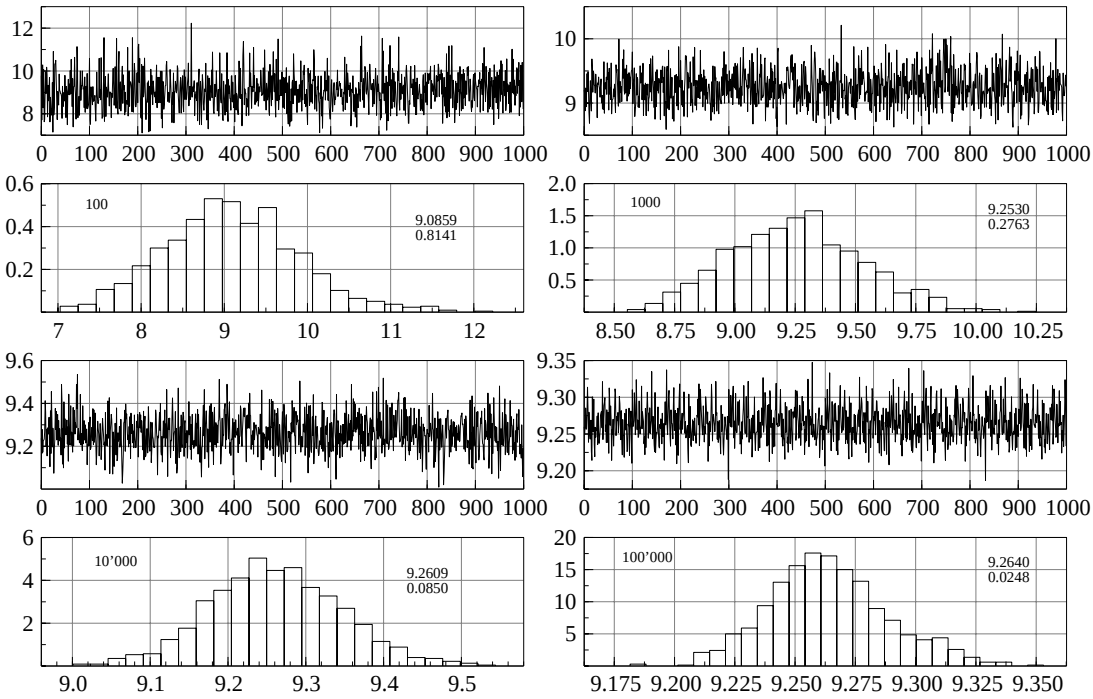


Abbildung 4.6: Vier Monte-Carlo-Experimente zur Verteilung des 95%-Quantils von Alexandersson's T_0. Für jedes Experiment wurden 1000 Mal je 100, 1000, 10'000 bzw. 100'000 Zeitreihen $Z_t \sim \mathsf{IN}(0, 1)$ der Länge $n = 100$ generiert und für jede Zeitreihe T_0 berechnet. Die pro Experiment resultierenden 1000 95%-Quantile von T_0 sind in den 4 Zeitreihendarstellungen wiedergegeben. Die korrespondierenden Histogramme zeigen die Verteilung der 1000 simulierten 95%-Quantile sowie deren empirische Mittelwerte und Standardfehler.

Es zeigt sich, dass die 1000 Einzelschätzungen von T_{95} für kleine Anzahl Replikationen in der Tat rechtsschief verteilt sind und mit steigendem Simulationsumfang nur sehr langsam konvergieren. So beträgt der relative Fehler des Mittelwertes der 1000 Einzelschätzungen von T_{95} für 10'000 Replikationen je Einzelschätzung immer noch rund 1%.

Ein in praxi wesentlich effizienterer Gebrauch der simulierten Zeitreihen – wurden doch zur Bestimmung der 1000 95%-Quantile von T_0 bei 100'000 Replikationen realiter 100'000'000 Einzelzeitreihen der Länge $n = 100$ simuliert! – ist möglich durch Schätzung von T_{95} und dessen Fehler aus den simulierten empirischen Verteilungsfunktionen von T_0, wie sie in Abbildung 4.7 für 1000 bzw. 10'000 Replikationen wiedergegeben sind. Hieraus lässt sich unschwer erkennen, dass für eine einigermassen zuverlässige Schätzung der hohen Quantile von T_0 aufgrund der langschwänzigen Form der Verteilungsfunktion in der Tat viele Replikationen notwendig sind.

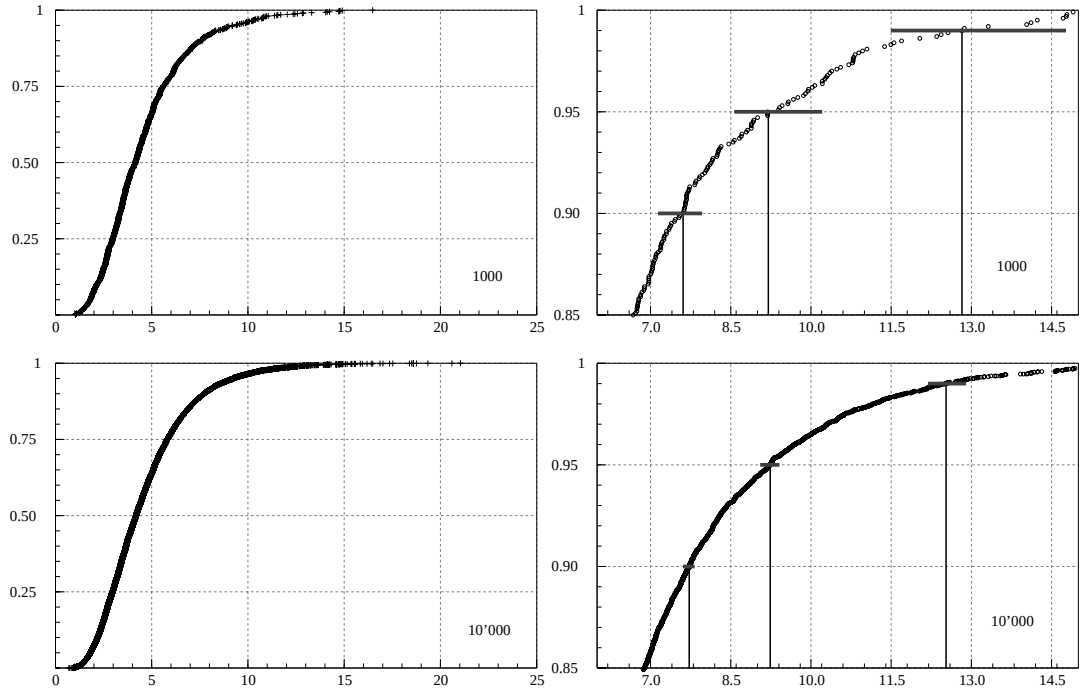


Abbildung 4.7: Die beiden Grafiken auf der linken Seite zeigen zwei Simulationen der empirischen Verteilungsfunktion von Alexandersson's T_0. Für jede Simulation wurden je 1000 bzw. 10'000 Zeitreihen $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 100$ generiert und deren T_0 berechnet. Auf der rechten Seite sind die korrespondierenden Ausschnitte der oberen Enden der simulierten empirischen Verteilungsfunktionen zu sehen, mit deren 90%, 95% und 99%-Quantilen sowie deren nichtparametrischen 95%-Vertrauensintervallen.

Nach Conover (1980; Kap. 3.2) lässt sich für ein Quantil T_p einer beliebigen, unbekannten Verteilungsfunktion ein **nichtparametrisches Vertrauensintervall** wie folgt konstruieren:

$$\begin{aligned} 1-\alpha &= P\left(T_{(l)} \le T_p \le T_{(u)}\right) = \sum_{i=l}^{u-1} \binom{n}{i} p^i (1-p)^{n-i} \\ \frac{\alpha}{2} &= P\left(T_{(l)} \ge T_p\right) = \sum_{i=0}^{l-1} \binom{n}{i} p^i (1-p)^{n-i} \approx \mathbf{\Phi}\left(\frac{l-np}{\sqrt{np(1-p)}}\right) \end{aligned}$$

woraus:

$$\begin{aligned} l &\approx \left\lfloor np - z_{1-\frac{\alpha}{2}}\sqrt{np(1-p)} \right\rfloor \\ u &\approx \left\lceil np + z_{1-\frac{\alpha}{2}}\sqrt{np(1-p)} \right\rceil \end{aligned}$$

Die Normalapproximation ist brauchbar, falls die Bedingung

$$np(1-p) \ge 10 \tag{4.24}$$

erfüllt ist (Hüsler und Zimmermann 2001, Kap. 12.1).

Somit überdeckt für hinreichend viele Replikationen n das nichtparametrische Vertrauensintervall

$$\boxed{\left(T_{\left(\left\lfloor np - z_{1-\frac{\alpha}{2}}\sqrt{np(1-p)}\right\rfloor\right)}, T_{\left(\left\lceil np + z_{1-\frac{\alpha}{2}}\sqrt{np(1-p)}\right\rceil\right)}\right)} \tag{4.25}$$

mit Wahrscheinlichkeit $1-\alpha$ das wahre Quantil T_p.

Abbildung 4.7 zeigt für 1000 bzw. 10'000 Replikationen die simulierten T_{90}, T_{95} und T_{99} sowie deren nach Formel (4.25) berechnete nichtparametrischen 95%-Vertrauensintervalle. In Tabelle 4.3 sind die Monte-Carlo-Schätzwerte von T_{90}, T_{95} und T_{99} sowie deren nichtparametrische 95%-Vertrauensintervalle für Replikationen zwischen 1000 und 10'000'000 zusammengestellt. Man beachte, dass bei 1000 Replikationen für T_{99} die Bedingung (4.24) knapp nicht erfüllt ist. Es zeigt sich, dass in praxi etwas mehr als 1'000'000 Replikationen vonnöten sind, um für $n = 100$ eine absolute Genauigkeit in der Grössenordnung von 0.01 zu erreichen. Ein Vergleich mit Abbildung 4.4 ergibt zudem, dass Alexandersson seinen Simulationen im besten Fall 10'000, wahrscheinlich jedoch nur 1000 Replikationen zu Grunde legte. Die publizierten kritischen Grenzen von Alexandersson (1986) sind daher wie die von Baudenbacher (1997) tendenziell **zu klein und zu ungenau.**

Replikationen	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
1000	7.1400	7.6083	7.9637	8.5649	9.2035	10.209	11.494	12.824	14.769
10'000	7.6067	7.7231	7.8233	9.0499	9.2375	9.4108	12.188	12.527	12.902
100'000	7.7910	7.8288	7.8692	9.2106	9.2598	9.3142	12.251	12.368	12.512
1'000'000	7.8423	7.8546	7.8671	9.2529	9.2692	9.2849	12.331	12.367	12.405
5'000'000	7.8519	7.8574	7.8627	9.2553	9.2626	9.2704	12.340	12.356	12.371
10'000'000	7.8522	7.8561	7.8599	9.2567	9.2619	9.2674	12.347	12.357	12.369

Tabelle 4.3: Kritische Grenzen des Alexandersson Tests für $n = 100$.

Schliesslich untersuchten wir, inwiefern die Schätzung der kritischen Grenzen von T_0 von der Wahl des Zufallszahlengenerators abhängt. In Abbildung 4.8 sind analog zu Abbildung 4.6 für sechs verschiedene Generatoren die Resultate der Simulation von T_{95} für $n = 100$ wiedergegeben.

Um zu prüfen, ob die mit den sechs Generatoren simulierten kritischen Grenzen im Rahmen einer absoluten Abweichung von maximal ± 0.05 äquivalent sind, stellten wir für die Differenz $\mu_i - \mu_i^\star$ zwischen dem Mittelwert μ_i eines einzelnen Generators und dem gemeinsamen Mittelwert $\mu_i^\star$ der fünf andern Generatoren nach Chow and Liu (1992) die beiden Hypothesen

$$H_0 \quad : \quad \mu_i - \mu_i^\star \leq -0.05 \ \vee \ \mu_i - \mu_i^\star \geq 0.05 \tag{4.26}$$

$$H_1 \quad : \quad -0.05 < \mu_i - \mu_i^\star < 0.05 \tag{4.27}$$

auf und konstruierten für die geschätzte Differenz $\bar{x}_i - \bar{x}_i^\star$ folgendes 90% Vertrauensintervall $\mathsf{CI}_{0.90,i}$

$$\mathsf{CI}_{0.90,i} = \left[\bar{x}_i - \bar{x}_i^\star - 1.645 \cdot \sqrt{s_{\bar{x}_i}^2 + s_{\bar{x}_i^\star}^2}\,,\ \bar{x}_i - \bar{x}_i^\star + 1.645 \cdot \sqrt{s_{\bar{x}_i}^2 + s_{\bar{x}_i^\star}^2}\right] \tag{4.28}$$

wobei $\bar{x}_i$ und $s_{\bar{x}_i}$ die in Tabelle 4.4 wiedergegebenen mit dem Generator i mittels 1'000'000 bzw. 5'000'000 Replikationen für $n = 100$ geschätzten kritischen Grenzen T_{95} sowie deren Standardfeh-

ler darstellen. Letztere berechnen sich wie die ebenfalls in Tabelle 4.4 wiedergegebenen geschätzten gemeinsamen Mittelwerte $\bar{x}_i^\star$ und Standardfehler $s_{\bar{x}_i^\star}$ nach den Formeln

$$
\begin{aligned}
s_{\bar{x}_i} &= (T_{(u)} - T_{(l)})/(2 \cdot 1.96) = (T_{(u)} - T_{(l)})/3.92 \\
\bar{x}_i^\star &= \sum_{\substack{j=1 \\ j \neq i}}^{6} \bar{x}_j / s_{\bar{x}_j}^2 \Bigg/ \sum_{\substack{j=1 \\ j \neq i}}^{6} 1/s_{\bar{x}_j}^2 \\
s_{\bar{x}_i^\star} &= \sqrt{\frac{1}{n(n-5)} \sum_{\substack{j=1 \\ j \neq i}}^{6} n_j(n_j-1)s_{\bar{x}_j}^2} = \sqrt{\frac{1}{5} \sum_{\substack{j=1 \\ j \neq i}}^{6} s_{\bar{x}_j}^2}.
\end{aligned}
$$

Um für einen Generator i die Nullhypothese (4.26) verwerfen zu können muss gelten:

$$
\boxed{\mathsf{CI}_{0.90,i} \in (-0.05, +0.05)} \tag{4.29}
$$

Da dies nach Tabelle 4.4 für alle i gilt, lässt sich schliessen, dass die sechs Zufallszahlengeneratoren im Rahmen einer maximalen Abweichung von ± 0.05 äquivalente Schätzungen der kritischen Grenzen liefern. Sollen die mit den sechs Generatoren simulierten kritischen Grenzen bis auf ± 0.01 äquivalent sein, so sind rund 10'000'000 Replikationen nötig, wie anhand von Tabelle 4.3 und Gleichung (4.28) abgeschätzt werden kann.

Replikationen	Generator	i	$\bar{x}$	$s_{\bar{x}}$	$\bar{x}^\star$	$s_{\bar{x}^\star}$	$\bar{x} - \bar{x}^\star$	$\mathsf{CI}_{0.90,l}$	$\mathsf{CI}_{0.90,u}$
1'000'000	PM	1	9.2649	0.0089	9.2563	0.0087	+0.0086	-0.0118	0.0290
1'000'000	GM	2	9.2510	0.0085	9.2591	0.0088	-0.0081	-0.0282	0.0120
1'000'000	LE	3	9.2557	0.0086	9.2581	0.0088	-0.0024	-0.0226	0.0178
1'000'000	PMS	4	9.2692	0.0082	9.2550	0.0088	+0.0142	-0.0056	0.0340
1'000'000	GMS	5	9.2526	0.0090	9.2586	0.0087	-0.0060	-0.0266	0.0146
1'000'000	LES	6	9.2507	0.0092	9.2589	0.0086	-0.0082	-0.0289	0.0125
5'000'000	PM	1	9.2608	0.0039	9.2607	0.0039	+0.0001	-0.0090	0.0092
5'000'000	GM	2	9.2588	0.0039	9.2612	0.0039	-0.0024	-0.0114	0.0066
5'000'000	LE	3	9.2575	0.0040	9.2614	0.0039	-0.0039	-0.0131	0.0053
5'000'000	PMS	4	9.2626	0.0039	9.2604	0.0039	+0.0022	-0.0070	0.0114
5'000'000	GMS	5	9.2641	0.0040	9.2601	0.0039	+0.0040	-0.0051	0.0131
5'000'000	LES	6	9.2606	0.0038	9.2608	0.0039	-0.0002	-0.0092	0.0088

Tabelle 4.4: Vergleich von sechs Simulationen des 95%-Quantils von Alexandersson's T_0 für $n = 100$ basierend auf je 1'000'000 bzw. 5'000'000 Replikationen mit verschiedenen Zufallszahlengeneratoren. Wiedergegeben sind die mittleren Schätzungen $\bar{x}$ von Alexandersson's T_{95} und deren Standardfehler $s_{\bar{x}}$ sowie die aus den gemeinsamen mittleren Schätzwerten $\bar{x}^\star$ und Standardfehlern $s_{\bar{x}^\star}$ nach Formel (4.28) berechneten gemeinsamen 90% Vertrauensintervalle $\mathsf{CI}_{0.90}$ für die geschätzte Differenz $\bar{x}_i - \bar{x}_i^\star$.

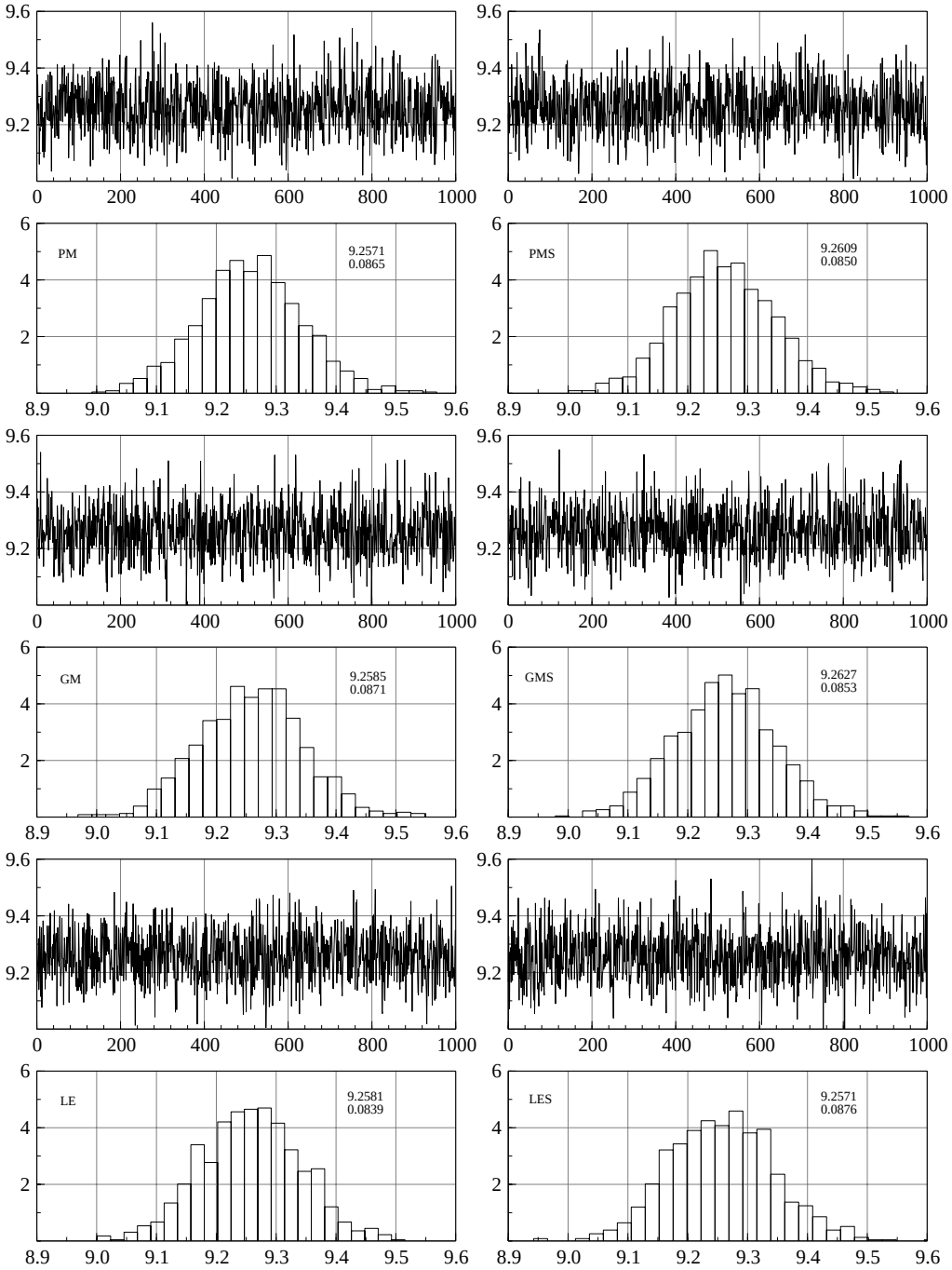


Abbildung 4.8: Simulation des 95%-Quantils von Alexandersson's T_0 mit sechs verschiedenen Zufallszahlengeneratoren. Für jede Simulation wurden 1000 Mal je 10'000 Zeitreihen $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 100$ generiert. Auf der linken Seite sind die Resultate der Generatoren von Park und Miller (PM), Marsaglia (GM) und L'Ecuyer (LE) wiedergegeben. Auf der rechten Seite finden sich die Ergebnisse der nach dem Verfahren von Knuth (1981) verbesserten Generatoren.

Wie im Anhang A beschrieben, wählten wir für alle unsere im Anhang C wiedergegebenen Monte-Carlo-Experimente den mittels dem Verfahren von Knuth (1981) verbesserten Generator von Park und Miller (1988), da dieser der schnellste der verbesserten Generatoren ist.

4.6.4 Vergleich

Simulierte kritische Grenzen der Likelihood Ratio Tests liegen vor von Hawkins (1977), Worsley (1979), Yao und Davis (1986), Antoch et al. (1997), Csörgő und Horváth (1997, p.25), Antoch et al. (2002) sowie von Alexandersson (1986) und Baudenbacher (1997). In Tabelle 4.5 sind ihre mit den Resultaten aus Anhang C vergleichbaren kritischen Grenzen zusammengestellt. Daraus ergibt sich primär die beruhigende Feststellung, dass die kritischen Grenzen grösstenteils übereinstimmen. Die Algorithmen der umfangreichen Monte-Carlo-Experimente scheinen daher korrekt implementiert zu sein. Zusätzlich ergeben sich folgende Feststellungen:

- Die von Hawkins (1977) publizierten kritischen Grenzen basieren nicht auf simulierten Zufallszahlenfolgen, sondern auf einem “exakten” Näherungsverfahren. Von diesem ist weder die Anzahl Iterationen, noch der Abbruchfehler bekannt.
- Die Anzahl Replikationen lag bei Worsley (1979) und bei Antoch et al. (1997) bei 10’000, bei Csörgő und Horváth (1997) sowie bei Yao und Davis (1986) bei 5000. Dies ist gemäss den Abschätzungen in Abschnittt 4.6.3 entschieden zu wenig und führt schon bei $n = 100$ zu einer Unsicherheit von ± 0.2.

Test	n	Quantil	A	B	C	D	E	F	G_l	G	G_u
LR1	100	90	2.603	-	-	-	-	-	2.6003	2.6027	2.6051
LR1	100	95	2.874	-	-	-	-	-	2.8693	2.8725	2.8755
LR1	100	99	3.410	-	-	-	-	-	3.4048	3.4106	3.4164
LR1	1000	90	2.804	-	-	-	-	-	2.8018	2.8041	2.8064
LR1	1000	95	3.066	-	-	-	-	-	3.0620	3.0652	3.0683
LR1	1000	99	3.586	-	-	-	-	-	3.5725	3.5787	3.5839
LR3	50	90	2.709	-	2.73	-	2.71	-	2.7103	2.7128	2.7150
LR3	50	95	2.960	-	2.98	-	2.94	-	2.9655	2.9684	2.9716
LR3	50	99	3.486	-	3.50	-	3.50	-	3.4805	3.4857	3.4911
LR3	100	90	2.809	-	-	-	2.82	2.84	2.8079	2.8101	2.8125
LR3	100	95	3.065	-	-	-	3.06	3.11	3.0625	3.0656	3.0687
LR3	100	99	3.563	-	-	-	3.58	3.58	3.5709	3.5773	3.5829
LR4	50	90	2.857	-	-	2.87	-	-	2.8611	2.8635	2.8662
LR4	50	95	3.157	-	-	3.16	-	-	3.1547	3.1580	3.1611
LR4	50	99	3.747	-	-	3.79	-	-	3.7584	3.7652	3.7723
LR4	100	90	2.891	-	-	-	-	-	2.8878	2.8903	2.8928
LR4	100	95	3.164	-	-	-	-	-	3.1611	3.1641	3.1671
LR4	100	99	3.696	-	-	-	-	-	3.7126	3.7189	3.7254
LR6	80	90	-	2.74	-	-	-	-	2.7700	2.7721	2.7741
LR6	80	95	-	3.00	-	-	-	-	3.0078	3.0105	3.0135
LR6	120	90	-	2.85	-	-	-	-	2.8223	2.8244	2.8266
LR6	120	95	-	3.08	-	-	-	-	3.0633	3.0659	3.0687
LR6	200	90	-	2.85	-	-	-	-	2.8843	2.8866	2.8888
LR6	200	95	-	3.10	-	-	-	-	3.1266	3.1293	3.1321

Tabelle 4.5: Vergleich der mittels 1’000’000 Replikationen simulierten mit den in der Literatur publizierten kritischen Grenzen einiger Likelihood Ratio Tests. A: Antoch et al. (2002), B: Antoch et al. (1997), C: Hawkins (1977), D: Worsley (1979), E: Csörgő und Horváth (1997), F: Yao und Davis (1986), G: Simulierte kritische Grenzen aus Anhang C.

In Abbildung 4.9 sind die Verläufe der kritischen Grenzen für Zeitreihen der Längen $n = 2$ bis $n = 200$ wiedergegeben. Deutlich ist das unterschiedliche Verhalten der kritischen Grenzen bei kleinen n zu erkennen: Wird σ wie beim LR3, LR6 und LR7 Test nach Formeln des Typs (4.14) oder (4.16) "global" geschätzt, so sind die kritischen Grenzen beginnend mit $n = 2$ monoton steigend. Wird σ hingegen wie beim LR2, LR4 und LR5 Test nach Formeln nach Formeln (4.7), (4.10) oder (4.12) "gepooled" geschätzt, sind die kritischen Grenzen beginnend mit $n = 2$ monoton fallend. Sie nähern sich bis ca. $n = 60$ rasch den "global" geschätzten kritischen Grenzen, ohne deren Verlauf jedoch zu kreuzen und steigen dann zusammen mit letzteren wieder monoton an. Auf den ersten Blick scheinen sich daher alle kritischen Grenzen demselben asymptotischen Wert anzunähern.

Aus Tabelle 4.1 wissen wir jedoch, dass die Teststatistiken der Likelihood Ratio Tests mit bekanntem μ gegen eine andere asymptotische Verteilung konvergieren, als diejenigen der Likelihood Ratio Tests mit unbekanntem μ. Dies ist in den Abbildungen 4.10 und 4.11 denn auch deutlich zu erkennen: Die kritischen Grenzen des LR1 und des LR2 Tests nähern sich einander mit steigendem n zwar rasch an, jedoch auf deutlich tieferem Niveau als diejenigen der Likelihood Ratio Tests mit unbekanntem μ. Auch die anderen Verläufe nähern sich einander mit steigendem n an, doch weisen die "globalen" und die "gepoolten" kritischen Grenzen unter einander kleinere Abstände auf als zwischen einander.

Nun haben Csörgő und Horváth (1997) gezeigt, dass die Verteilungen der Teststatistiken des LR3, LR4, LR5, LR6 und LR7 Tests und damit auch ihre kritischen Grenzen asymptotisch äquivalent sind. Es fragt sich nun, ab welchem n diese asymptotische Äquivalenz als praktisch realisiert

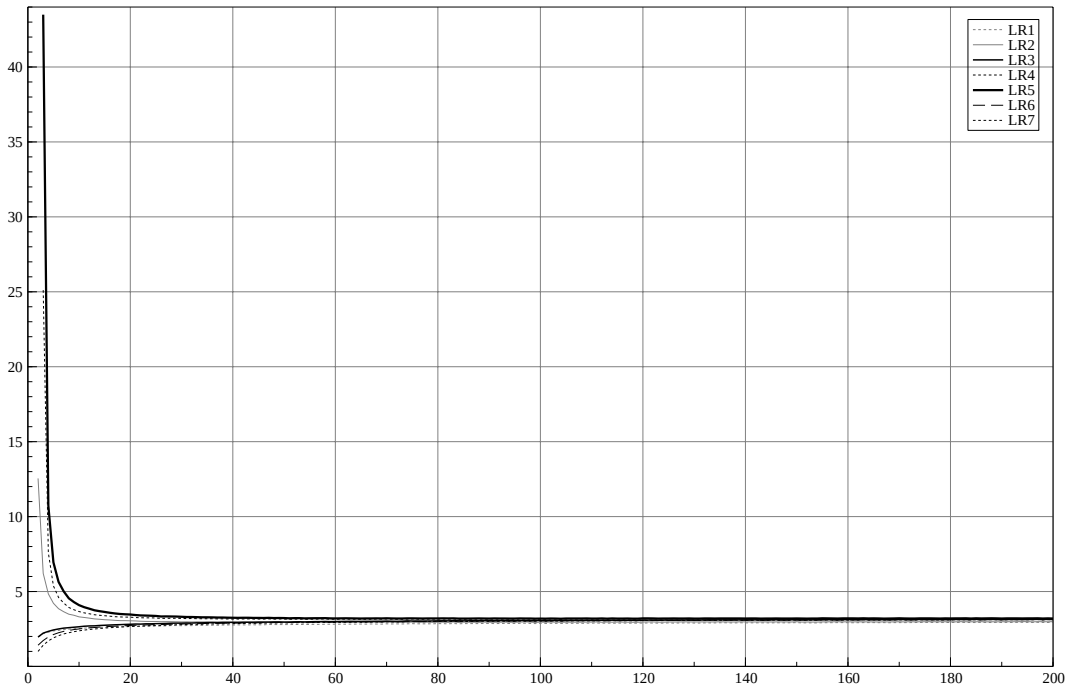


Abbildung 4.9: Verlauf der simulierten kritischen 95% Grenzen der Likelihood Ratio Tests für Zeitreihen der Längen $n = 2$ bis $n = 200$.

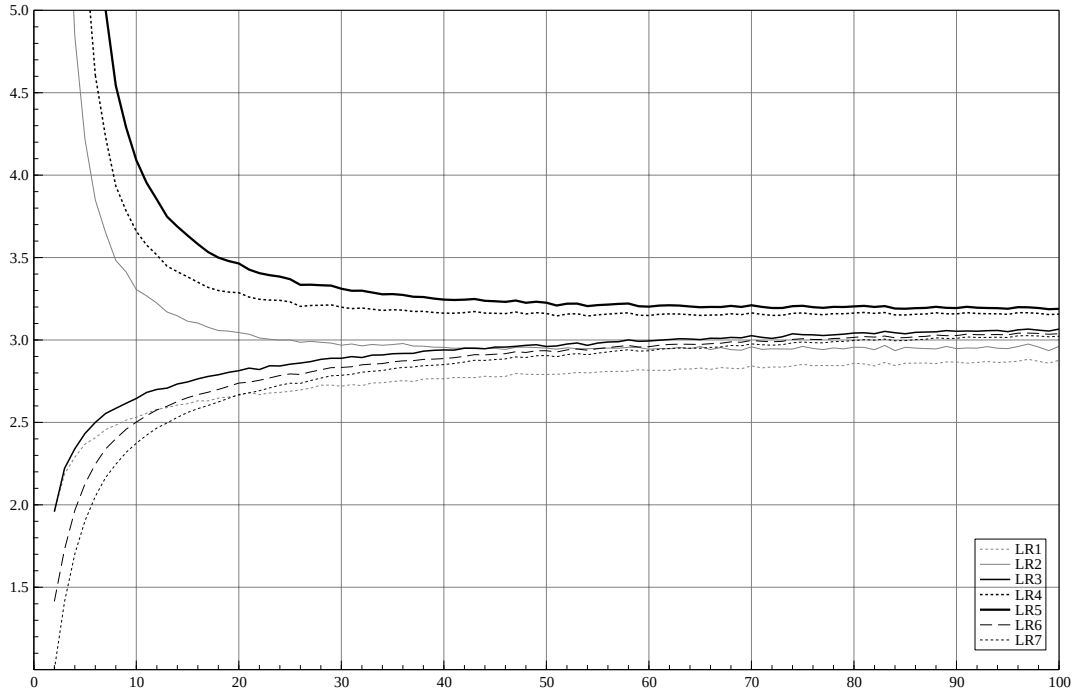


Abbildung 4.10: Verlauf der simulierten kritischen 95% Grenzen der Likelihood Ratio Tests für Zeitreihen der Längen $n = 2$ bis $n = 100$.

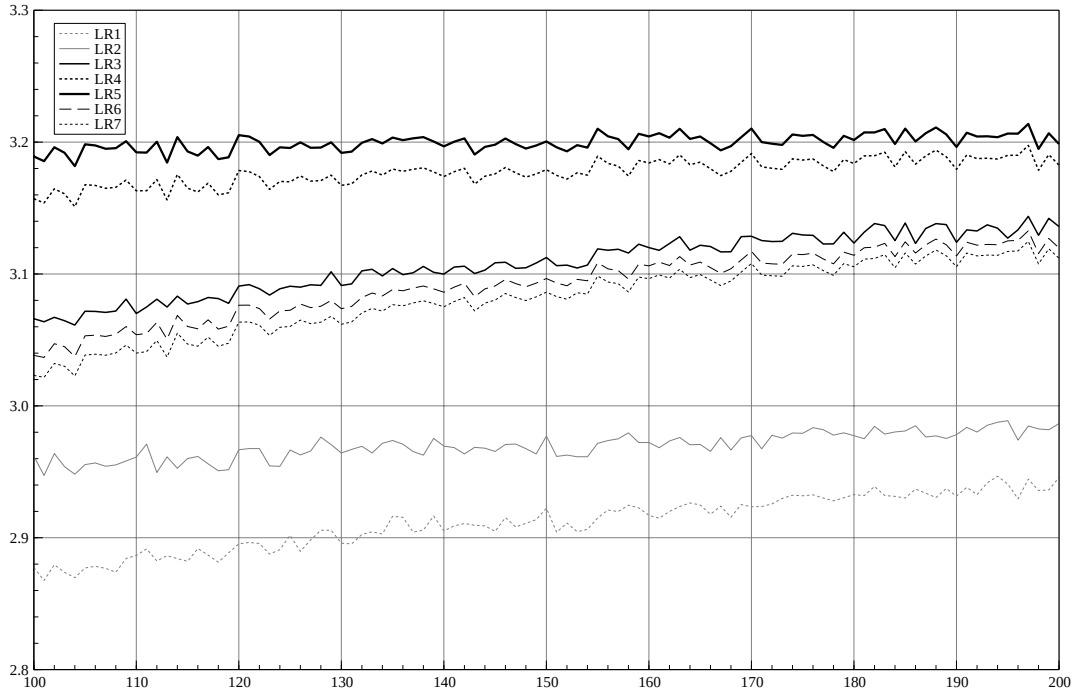


Abbildung 4.11: Verlauf der simulierten kritischen 95% Grenzen der Likelihood Ratio Tests für Zeitreihen der Längen $n = 100$ bis $n = 200$.

betrachtet werden kann. Hierzu wurde für den LR7 Test in Tabelle 4.6 berechnet, wie die Überschreitungswahrscheinlichkeit für das 95% Quantil der Testverteilung ändert, wenn die kritische Grenze leicht verschoben wird. Es zeigt sich, dass bei einer Änderung der kritischen Grenze um 1% die Überschreitungswahrscheinlichkeit um rund ein halbes Prozent ändert, während bei einer Änderung der kritischen Grenze um 10% von einer Einhaltung des α-Fehlers nicht mehr gesprochen werden kann.

Damit lässt sich nun aus den Abbildungen 4.10 und 4.11 herauslesen, dass die Differenzen innerhalb der "globalen" und der "gepoolten" Testgrössen ab ca. $n = 10$ unter 0.3, ab ca. $n = 150$ unter 0.03 und damit unter die "asymptotische Unterscheidbarkeit" fallen. Die Testgrössen der beiden Gruppen sind jedoch erst ab ca. $n = 1000$ so nahe beisammen, dass die Art und Weise der Schätzung von σ als irrelevant betrachtet werden kann.

Alle besprochenen Testvarianten besitzen daher für Zeitreihen mittlerer Länge ihre eigenen kritischen Grenzen, insbesondere auch diejenigen Testvarianten, welche sich nur in der Berechnung von s_k unterscheiden. Dem scheint in der statistischen Literatur - wahrscheinlich im vorschnellen Vertrauen auf deren asymptotische Äquivalenz - zu wenig Beachtung geschenkt worden zu sein: vom LR2 und LR5 Test wurden bisher gar keine kritischen Grenzen veröffentlicht und von den restlichen sind nur eine Hand voll Werte bekannt. Die einzigen längeren Tabellen kritischer Grenzen veröffentlichten Alexanderson (1986) und Baudenbacher (1997), allerdings mit den bereits im letzten Abschnitt besprochenen Mängeln.

Die Abbildungen 4.12 und 4.13 zeigen den Verlauf der asymptotischen Grenzwerte sowie der Bonferroni Approximationen für den LR3 sowie den LR4 Test. Hieraus wird deutlich, dass für die "gepoolten" Testgrössen die asymptotischen Grenzen für $n < 8$ zu liberal und für $n \geq 8$ zu konservativ sind, während für die "globalen" Testgrössen die asymptotischen Grenzen überall zu konservativ sind. Die Abweichungen der asymptotischen Grenzwerte von den simulierten sind überall grösser als 0.3. Dies gilt auch für den grössten in Anhang C simulierten kritischen Grenzwert bei $n = 6000$, beträgt dieser doch 3.38, während die asymptotische kritische Grenze bei 3.75 liegt, 0.37 höher. Die asymptotischen kritischen Grenzen sind daher zur Homogenitätsprüfung von Zeitreihen mittlerer Länge unbrauchbar.

Ähnliches gilt für die Bonferroni Approximationen, welche für den LR3 und den LR4 Test bis $n = 7$ brauchbare, für grössere n jedoch durchwegs unbrauchbar konservative kritische Grenzen liefern. Wie den Abbildungen 4.12 und 4.13 entnommen werden kann, werden diese für $n > 200$ bzw. $n > 130$ sogar noch konservativer als die asymptotischen kritischen Grenzen.

Δ	-0.3	-0.03	-0.003	-	0.003	0.03	0.3
Δ/T_{95}	-10%	-1%	-0.1%	-	0.1%	1%	10%
p	0.1124	0.0549	0.0509	0.0500	0.0490	0.0461	0.0213

Tabelle 4.6: Überschreitungswahrscheinlichkeit p des LR7 Tests bei $n = 100$ in Abhängigkeit der Verschiebung Δ der kritischen 95% Grenze.

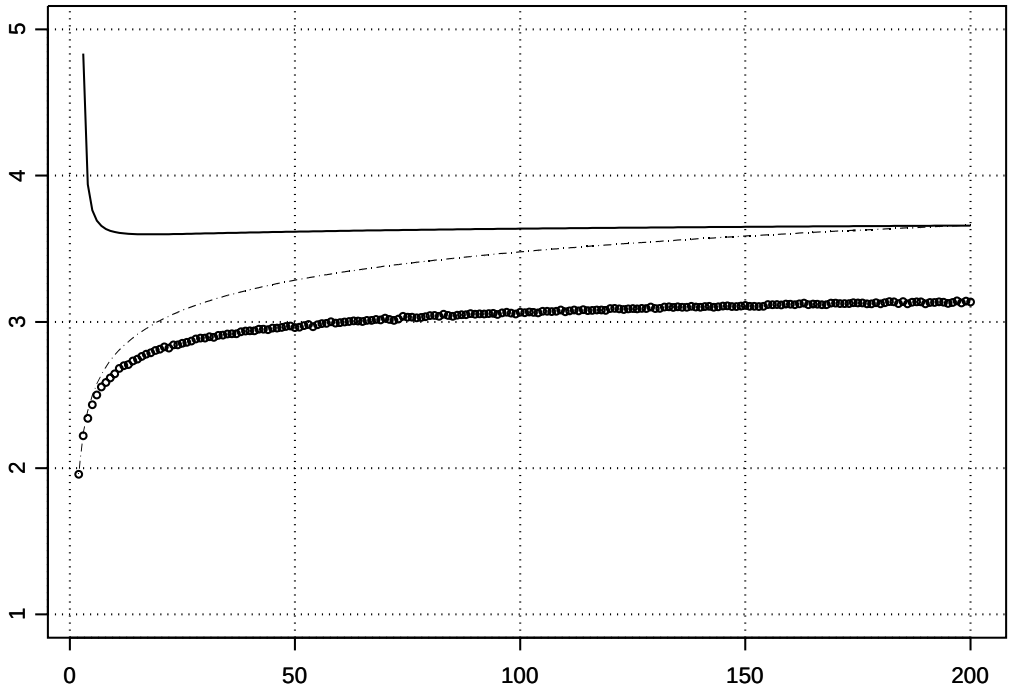


Abbildung 4.12: Verlauf der simulierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise), der Bonferroni Approximationen (strich-punktierte Kurve) sowie der asymptotischen kritischen Grenzen (ausgezogene Kurve).

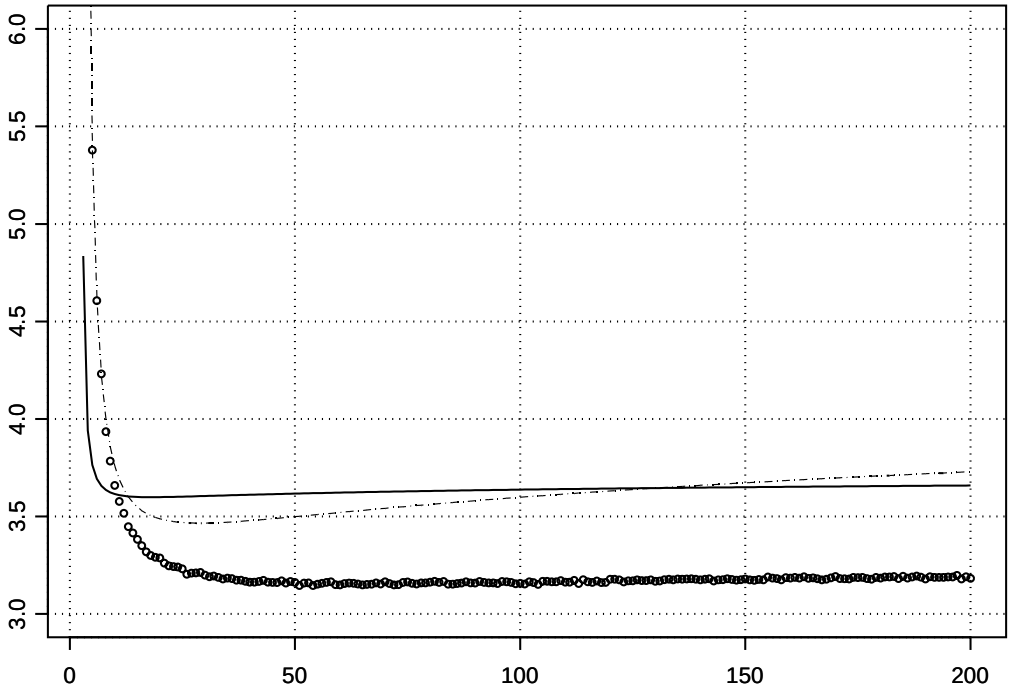


Abbildung 4.13: Verlauf der simulierten kritischen Grenzen des LR4 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise), der Bonferroni Approximationen (strich-punktierte Kurve) sowie der asymptotischen kritischen Grenzen (ausgezogene Kurve).

4.6.5 Empirische Approximation

Da für Zeitreihen mittlerer Länge weder die Bonferroni Approximationen noch die asymptotischen Verteilungen brauchbare Schätzungen der kritischen Grenzen liefern, müssen mit Hilfe der simulierten kritischen Grenzen empirische Näherungsformeln entwickelt werden.

Nach Yao und Davis (1986, p. 345) lässt sich die Verteilung der Testgrössen des LR3 Tests für Zeitreihen der Länge $20 \leq n \leq 50$ wie folgt annähern:

$$P(T_n \leq t) = \left[\Phi\left(t - \frac{\log\log\log n + \log 2}{\sqrt{2\log\log n}}\right)\right]^{2\log(n/2)} \tag{4.30}$$

Damit lautet das approximative p-Quantil der LR3 Testverteilung:

$$t_p = \Phi^{-1}\left(p^{\frac{1}{2\log(n/2)}}\right) + \frac{\log\log\log n + \log 2}{\sqrt{2\log\log n}} \tag{4.31}$$

In Abbildung 4.14 ist der Verlauf dieser approximativen kritischen Grenzen für den LR3 Test zusammen mit den simulierten und asymptotischen kritischen Grenzen sowie den Bonferroni Approximationen wiedergegeben. Überraschenderweise passen sich die Näherungswerte dem simulierten Verlauf der kritischen Grenzen bis auf relativ kleine n gut an, jedenfalls weit besser als dies aufgrund der von Yao und Davis (1986) gegebenen Einschränkung auf Zeitreihen der Länge $20 \leq n \leq 50$ vermutet werden könnte. Betrachten wir jedoch die in Abbildung 4.15 wiedergegeben Differenzen zwischen den simulierten kritischen Grenzen und den Näherungen nach Yao und Davis (1986) so erkennen wir, dass die Differenzen tatsächlich nur im angegebenen Intervall zwischen $20 \leq n \leq 50$ unter 0.05 bleiben. Aus der Abbildung 4.15 ist aber auch ersichtlich, dass die Differenzen nicht rasch ins Uferlose wachsen, sondern sich auf einen Wert um 0.058 einpendeln.

Die Approximation (4.31) kann daher für grössere n mit einem empirischen Korrekturterm verbessert werden:

$$t_{95} = \Phi^{-1}\left(0.95^{\frac{1}{2\log(n/2)}}\right) + \frac{\log\log\log n + \log 2}{\sqrt{2\log\log n}} - \frac{1}{20}\log\pi \tag{4.32}$$

Für $n = 6000$ ergibt sich mit dieser Korrektur ein Wert von 3.3737, welcher nur um 0.007 tiefer liegt als die simulierte kritische Grenze aus Anhang C. Wir schliessen daraus, dass die Formel (4.32) für grössere n bis mindestens $n = 6000$ brauchbare kritische Grenzen liefert.

Für kleine n wurde mittels doppelt logarithmischer Transformation des Regressors eine leidliche Linearisierung des Verlaufs der kritischen Grenzen erreicht (Abb. 4.16), welche durch folgende Formel dargestellt werden kann:

$$t_{95} = (2.1668 \pm 0.0022) + (0.5865 \pm 0.0015) \cdot \log\log n \tag{4.33}$$

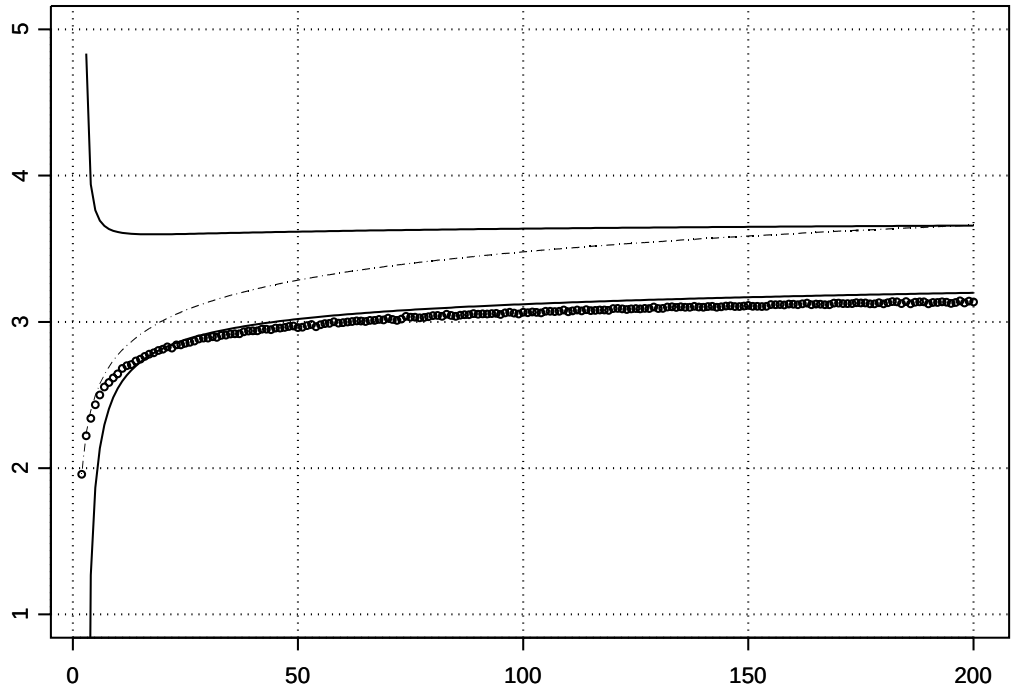


Abbildung 4.14: Verlauf der simulierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise), der Bonferroni Approximationen (strich-punktierte Kurve), der asymptotischen kritischen Grenzen (ausgezogene Kurve) sowie der von Yao und Davis (1986) gefundenen Näherungsformel (fette Kurve).

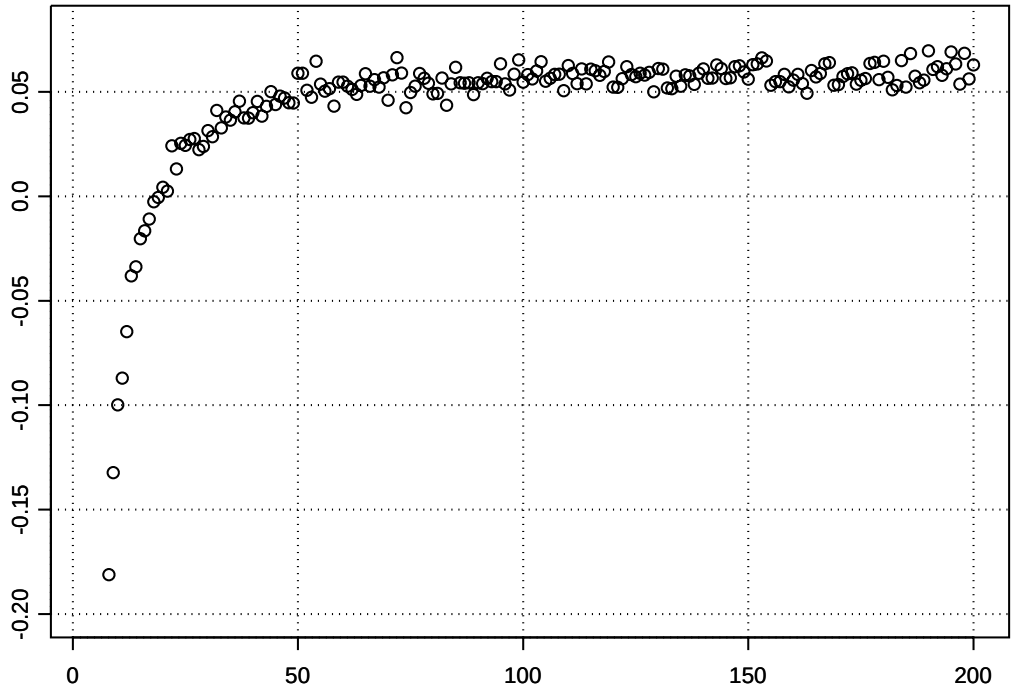


Abbildung 4.15: Differenzen zwischen den simulierten und den nach Yao und Davis (1986) approximierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$.

Wie den Abbildungen 4.17 und 4.18 entnommen werden kann, gelingt damit die Darstellung des Verlaufs der kritischen Grenzen des LR3 Tests zwischen $n = 2$ und $n = 200$ mit einem mittleren Fehler von 0.006. Aus dem Residuenplot 4.18 ist jedoch ersichtlich, dass die angepasste Kurve etwas weniger gekrümmt ist, als erforderlich. Für grössere n sind die Schätzungen der kritischen Grenzen nach Formel (4.33) daher eher zu konservativ. Vergleicht man, wie in Abbildung 4.19 dargestellt, den Verlauf der kritischen Grenzen nach Formel (4.32) mit demjenigen nach Formel (4.33), so gelangen wir zum Schluss, dass der Verlauf der kritischen Grenzen des LR3 Tests für $2 \leq n \leq 70$ am besten nach Formel (4.33) und für grössere n bis mindestens 6000 am besten nach Formel (4.32) dargestellt werden kann.

Man beachte, dass selbst für n = 1e+308 die Differenz zwischen den asymptotischen kritischen Grenzen und denjenigen nach Formel (4.32) grösser als 0.1 bleibt.

Für den LR1 Test existiert keine Näherungsformel für den Verlauf der kritischen Grenzen. Eine heuristische Anpassung der Formel (4.32) an den Verlauf der simulierten kritischen Grenzen führt zur Darstellung:

$$t_{95} = \Phi^{-1}\left(0.95^{\frac{1}{2\log(n)}}\right) + \frac{\log\log\log(2n)}{\sqrt{2\log\log n}} - \frac{1}{40}\log\pi \tag{4.34}$$

Der Verlauf der kritischen Grenzen nach Formel (4.34) ist zusammen mit den asymptotischen kritischen Grenzen und den Bonferroni Approximationen in Abbildung 4.22 dargestellt. Für $n = 1000$ ergibt sich nach Formel (4.34) ein Wert von 3.0661. Die Differenz zum simulierten Wert 3.0652 ist also lediglich 0.0009! Vergleicht man jedoch den Verlauf der kritischen Grenzen nach Formel (4.34) mit demjenigen der asymptotischen kritischen Grenzen nach Formel (4.22), so lässt sich feststellen, dass die beiden Verläufe für grosse n weitgehend parallel verlaufen. Die Formel (4.34) approximiert die kritischen Grenzen des LR1 Tests daher bestenfalls für kürzere Zeitreihen.

Erfreulicherweise gelingt die empirische Approximation der simulierten kritischen Grenzen mit Hilfe einer doppelt logarithmischen Transformation des Regressors wesentlich besser als beim LR3 Test. Wie aus den Abbildungen 4.21 bis 4.23 ersichtlich, stellt die Formel

$$t_{95} = (2.1327 \pm 0.0019) + (0.4847 \pm 0.0013) \cdot \log\log n \tag{4.35}$$

den Verlauf der simulierten kritischen Grenzen mit einem mittleren Fehler von 0.005 dar, wobei die Krümmung der Kurve perfekt modelliert wird. Für $n = 1000$ ergibt sich ein geschätzter Wert von 3.0693 nur 0.0041 verschieden vom simulierten Wert 3.0652.

Für die Darstellung der kritischen Grenzen des LR1 Tests für $2 \leq n \leq 1000$ und darüber hinaus, scheint also die Formel (4.35) allein zu genügen. Unterstellt man ihr allgemeine Gültigkeit, so folgt daraus, dass für n > 3e+10 die Differenz zwischen den simulierten kritischen Grenzen und den asymptotischen unter 0.03 fällt. Ab n > 9e+11 liefert die Formel allerdings Schätzungen der kritischen Grenzen, welche grösser sind als die asymptotischen.

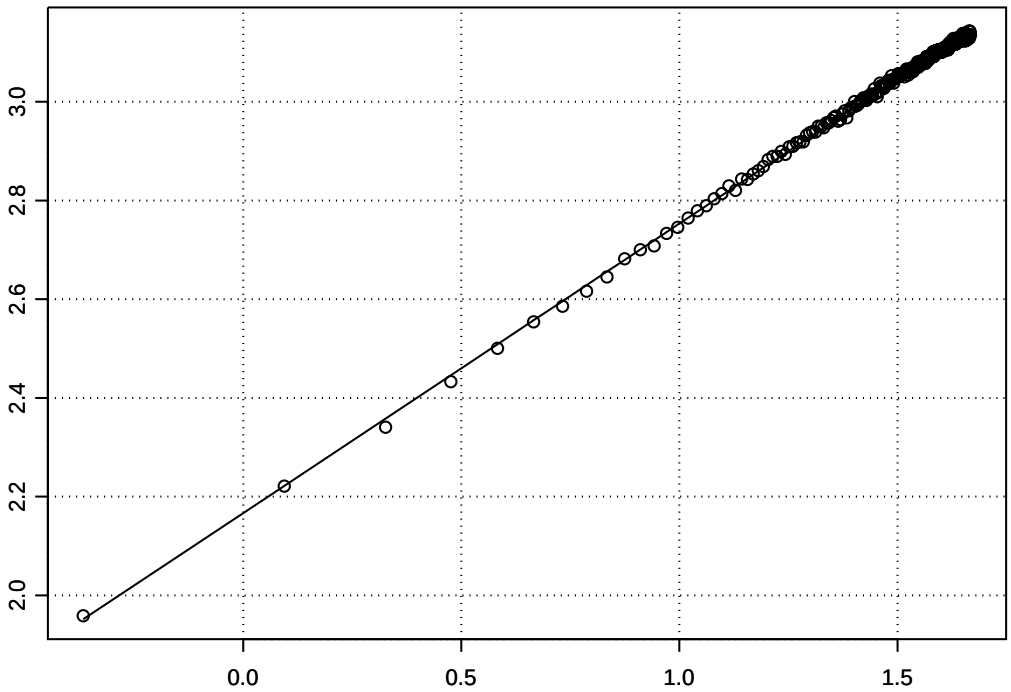


Abbildung 4.16: Verlauf der simulierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise) nach doppelt logarithmischer Transformation der Abszisse. Die mittels kleinsten Quadraten geschätzte Gerade entspricht der linearen Näherungsformel (4.33).

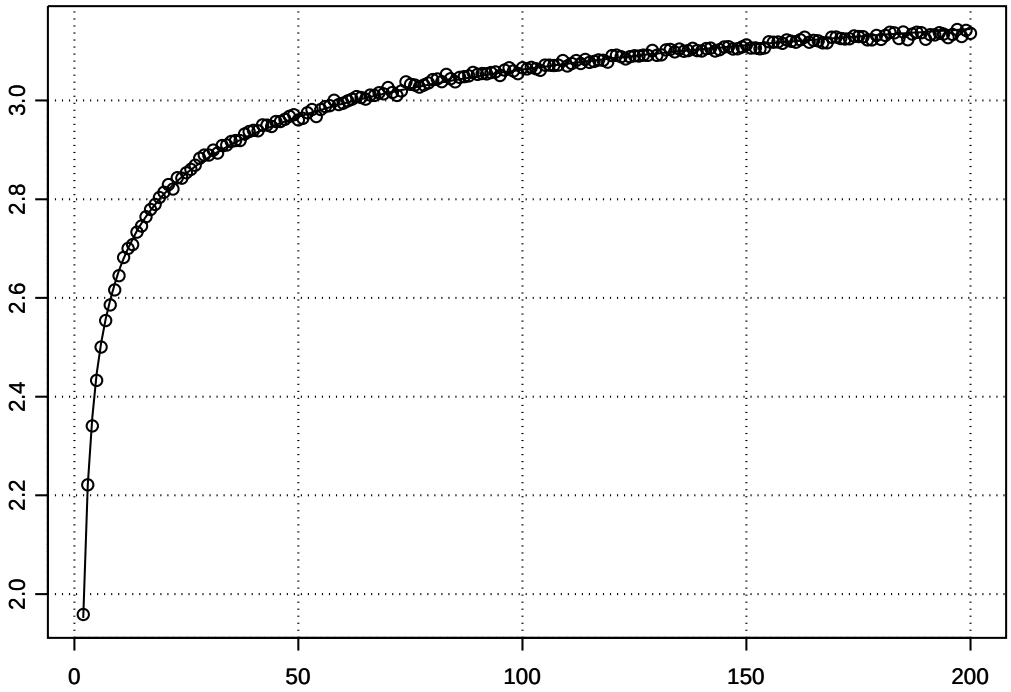


Abbildung 4.17: Verlauf der simulierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise) sowie der empirischen Schätzwerte nach Formel (4.33).

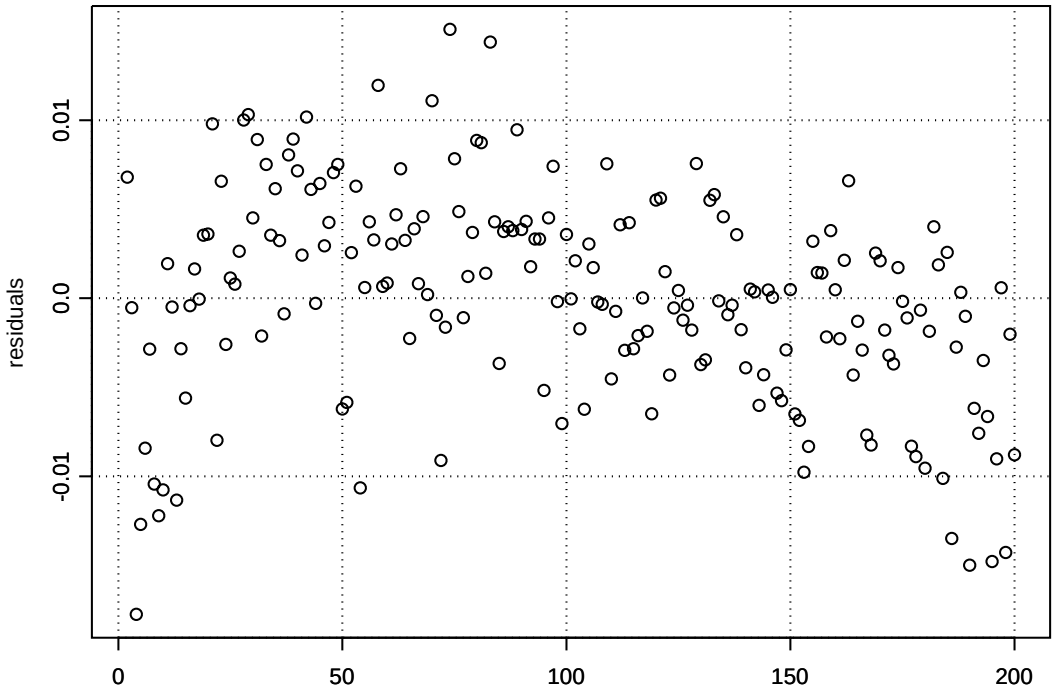


Abbildung 4.18: Verlauf der Residuen des Modells (4.33) mit deutlich erkennbarer Restkurvatur.

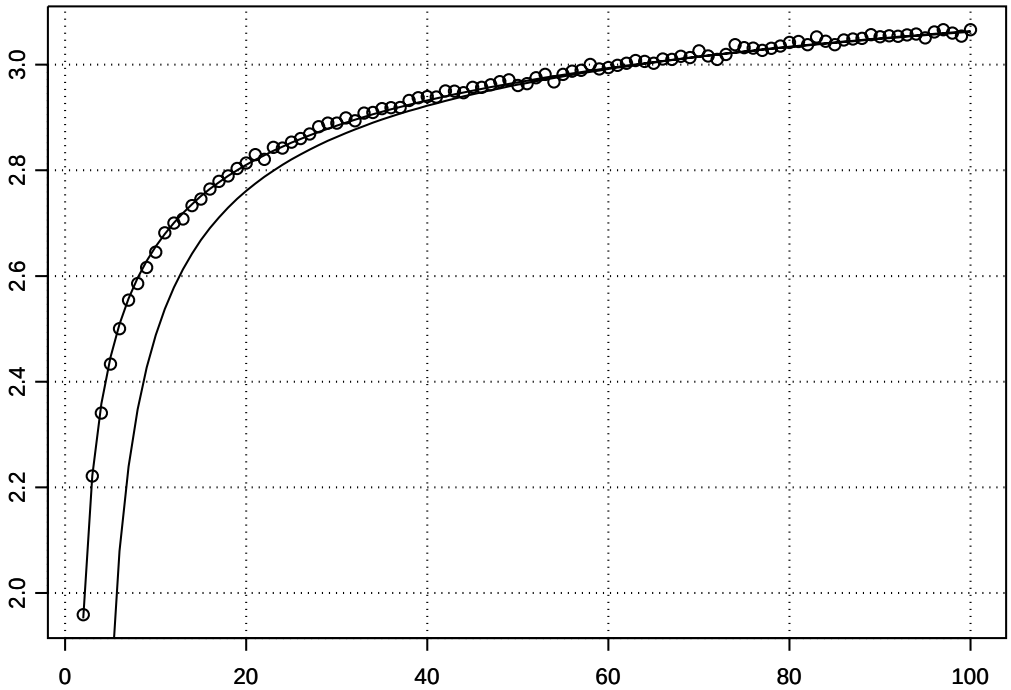


Abbildung 4.19: Verlauf der simulierten kritischen Grenzen des LR3 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise) sowie der empirischen Schätzwerte nach den Formeln (4.32) und (4.33).

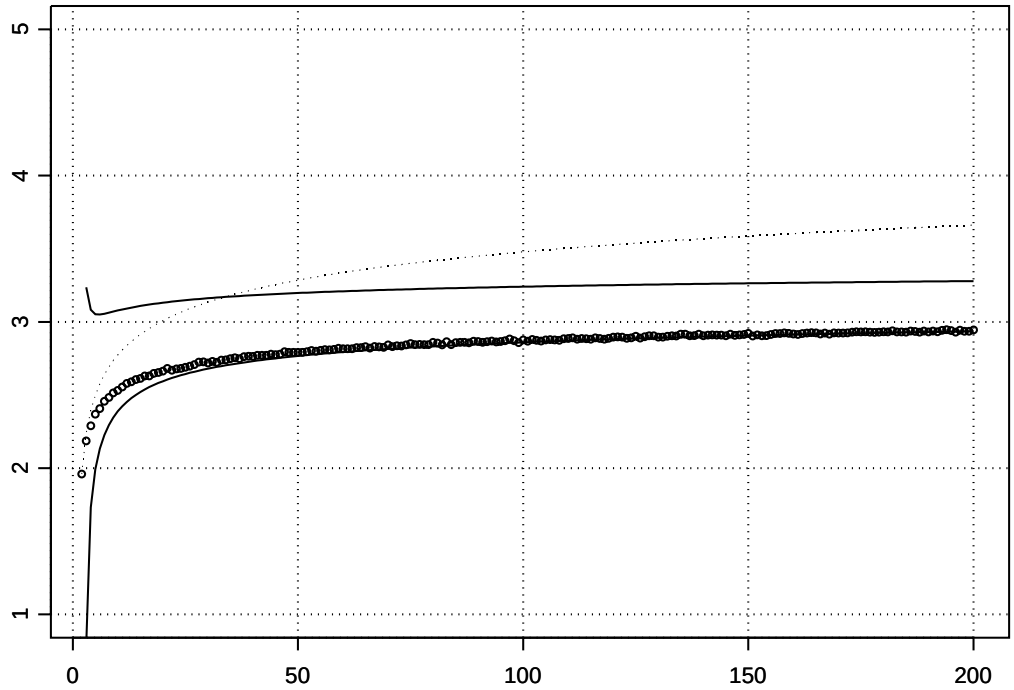


Abbildung 4.20: Verlauf der simulierten kritischen Grenzen des LR1 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise), der Bonferroni Approximationen (strich-punktierte Kurve), der asymptotischen kritischen Grenzen (ausgezogene Kurve) sowie der heuristischen Näherungsformel (4.34) (fette Kurve).

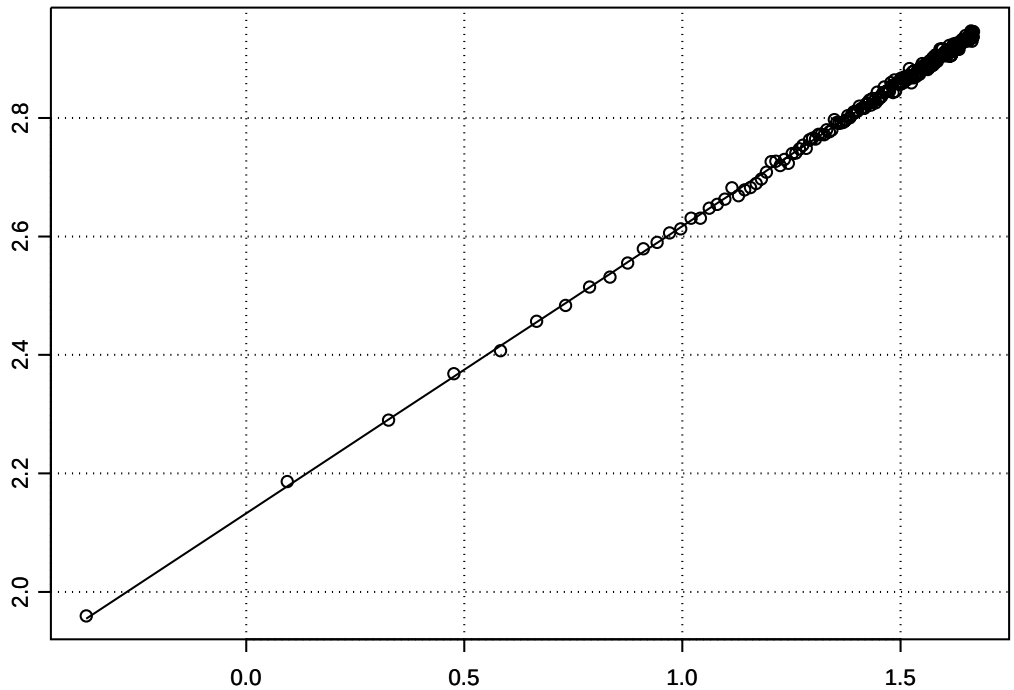


Abbildung 4.21: Verlauf der simulierten kritischen Grenzen des LR1 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise) nach doppelt logarithmischer Transformation der Abszisse. Die mittels kleinsten Quadraten geschätzte Gerade entspricht der linearen Näherungsformel (4.35).

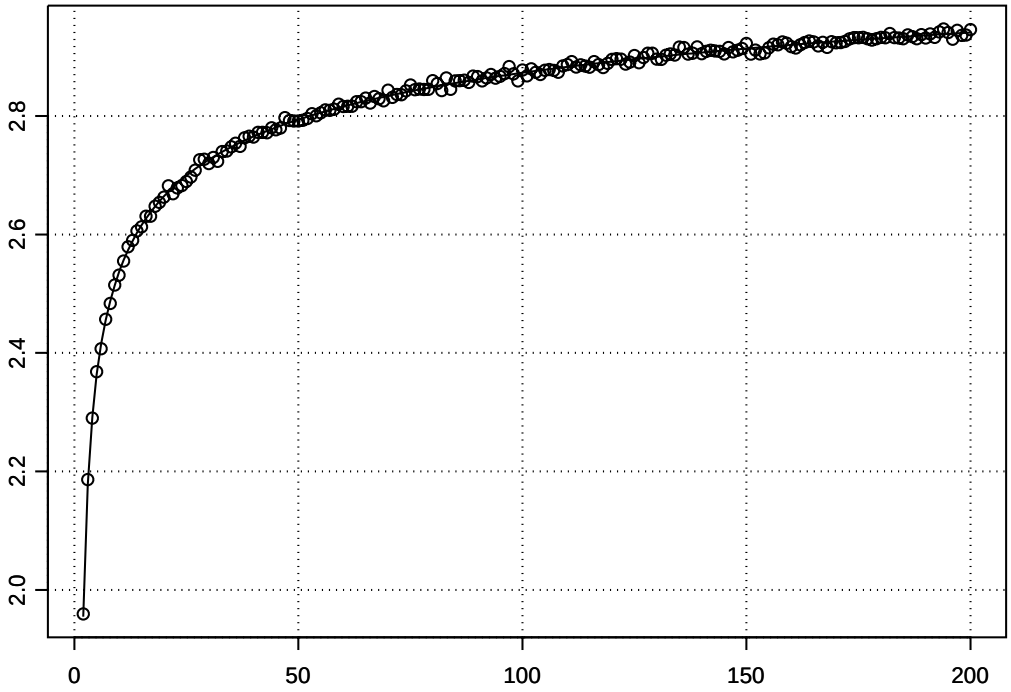


Abbildung 4.22: Verlauf der simulierten kritischen Grenzen des LR1 Tests für Zeitreihen der Länge $n = 2$ bis $n = 200$ (Kreise) sowie der empirischen Schätzwerte nach Formel (4.35).

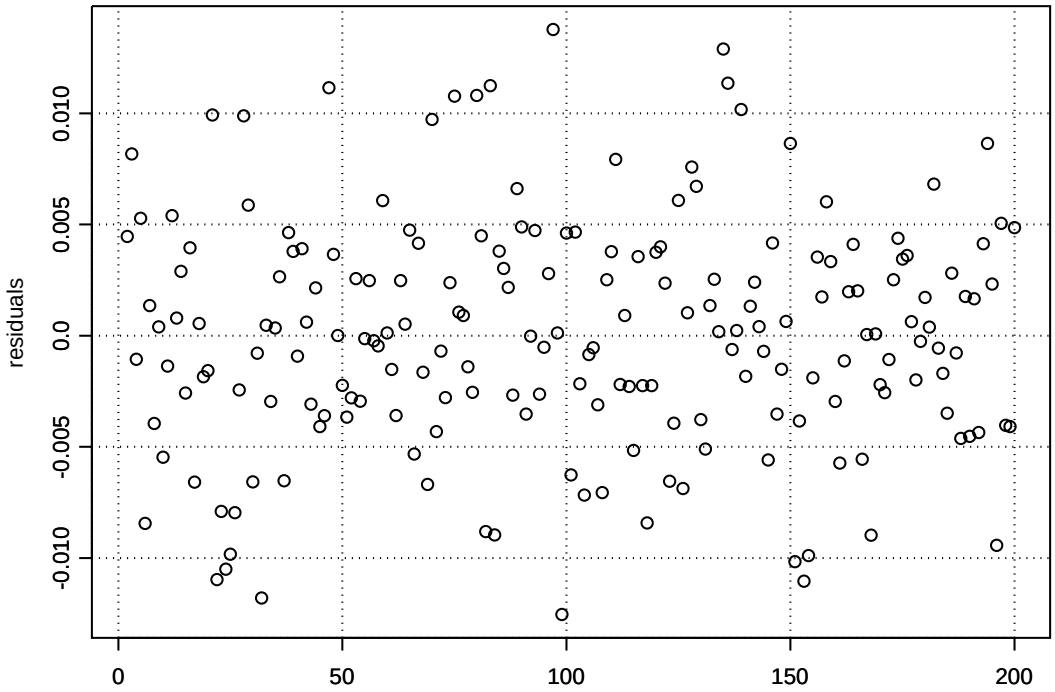


Abbildung 4.23: Verlauf der Residuen des Modells (4.35).

4.6.6 Buishand Test

Kritische Grenzen des Buishand Tests wurden von Buishand (1982) anhand von 19'999 Zufallsreihen bestimmt (Tab. 4.7). Erste Simulationsexperimente ergaben, dass die Genauigkeit der simulierten kritischen Grenzen mit derjenigen der Likelihood Ratio Tests vergleichbar ist. Die Resultate der Monte-Carlo Experimente sind in Anhang C zusammengestellt.

Ein Vergleich mit den von Buishand (1982) publizierten kritischen Grenzen zeigt eine gute Übereinstimmung mit den Resultaten der Simulationsexperimente (Tab. 4.7).

Gemäss den Angaben von Buishand (1982) konvergiert die transformierte Testgrösse $T = Q/\sqrt{n}$ für $n \to \infty$ gegen die Verteilung des Kolmogorov-Smirnov Anpassungstests.

In Abbildung 4.24 ist der Verlauf der transformierten simulierten kritischen Grenzen T_{95} dargestellt, welche auf den ersten Blick eine rasche Annäherung der simulierten kritischen Grenzen an den asymptotischen Wert nahelegt. Allerdings liegt $T_{95} = 1.3504$ für $n = 6000$ nur scheinbar nahe beim asymptotischen Grenzwert 1.36, dürfte doch für $n = 6000$ die Abweichung von 1.36 höchstens $0.03/\sqrt{6000} = 0.0004$ betragen!

Ähnlich wie bei den Likelihood Ratio Tests sind daher auch für den Buishand Test die asymptotischen kritischen Grenzen für Zeitreihen mittlerer Länge durchwegs zu konservativ, weshalb eine empirische Darstellung der simulierten kritischen Grenzen gefunden werden muss.

Mit Hilfe einer einfachen Wurzeltransformation des Regressors gelingt eine hinreichend gute Approximation der simulierten kritischen Grenzen. Wie den Abbildungen 4.25 bis 4.27 entnommen werden kann, stellt die Formel

$$Q_{95} = (-0.6339 \pm 0.0050) + (1.3592 \pm 0.0002) \cdot \sqrt{n} \tag{4.36}$$

den Verlauf der simulierten kritischen Grenzen mit einem mittleren Fehler von 0.02 dar. Aus dem Residuenplot 4.27 ist allerdings auch ersichtlich, dass die angepasste Kurve wahrscheinlich etwas zu wenig gekrümmt ist, so dass Schätzungen der kritischen Grenzen des Buishand Tests nach Formel (4.36) für $n > 6000$ tendenziell zu konservativ ausfallen.

n	T_{90}	T_{95}	T_{99}	T_{95}
10	1.05	1.14	1.29	1.14
20	1.10	1.22	1.42	1.21
30	1.12	1.24	1.46	1.24
40	1.13	1.26	1.50	1.26
50	1.14	1.27	1.52	1.27
100	1.17	1.29	1.55	1.30
∞	1.22	1.36	1.63	-

Tabelle 4.7: Transformierte kritische Grenzen $T = Q/\sqrt{n}$ des Buishand Tests. Die ersten drei Spalten stammen aus Buishand (1982), die letzte aus Anhang C.

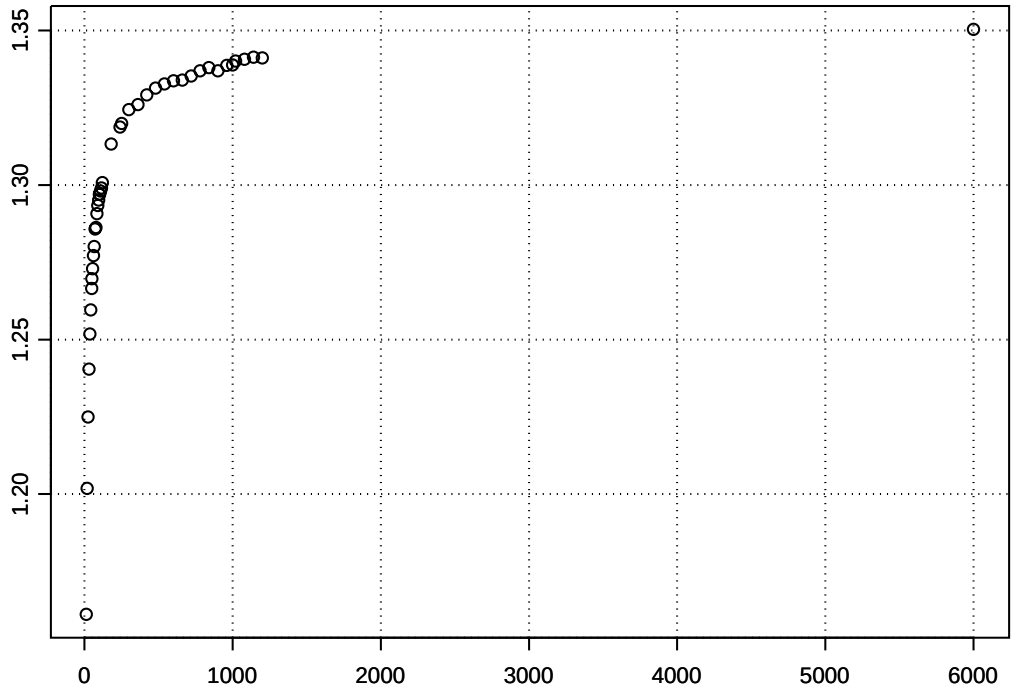


Abbildung 4.24: Verlauf der transformierten simulierten kritischen Grenzen $T = Q/\sqrt{n}$ des Buishand Tests für Zeitreihen der Länge $n = 2$ bis $n = 6000$ (Kreise). Der asymptotische Grenzwert liegt nach Buishand (1982) bei 1.36.

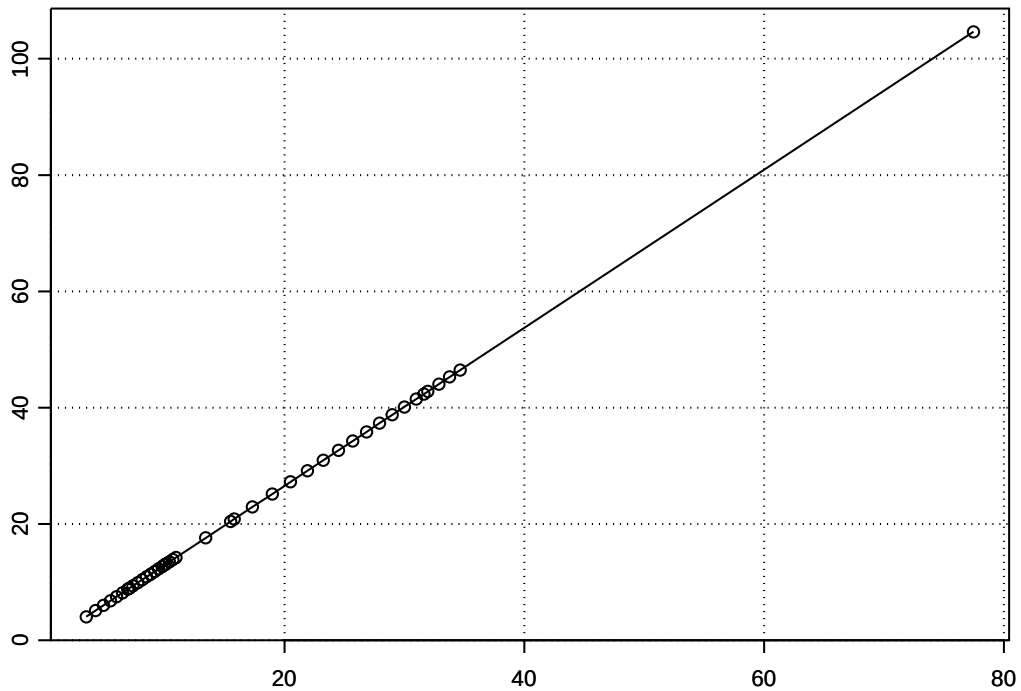


Abbildung 4.25: Verlauf der simulierten kritischen Grenzen des Buishand Tests für Zeitreihen der Länge $n = 2$ bis $n = 6000$ (Kreise) nach einer Wurzeltransformation der Abszisse. Die mittels kleinsten Quadraten geschätzte Gerade entspricht der linearen Näherungsformel (4.36).

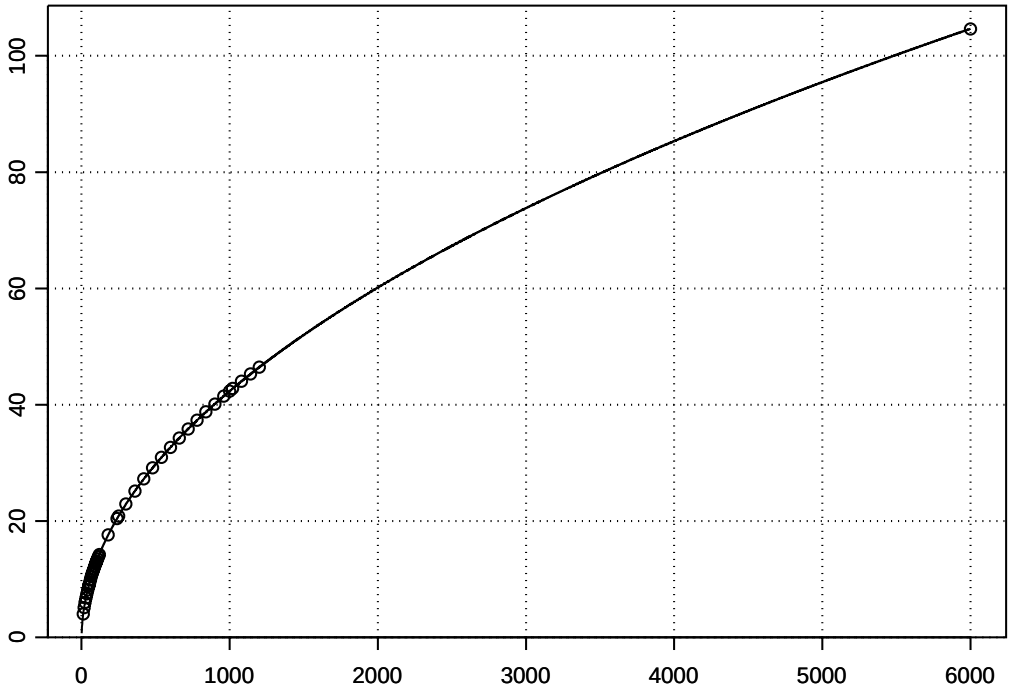


Abbildung 4.26: Verlauf der simulierten kritischen Grenzen des Buishand Tests für Zeitreihen der Länge $n = 2$ bis $n = 6000$ (Kreise) sowie der empirischen Schätzwerte nach Formel (4.36).

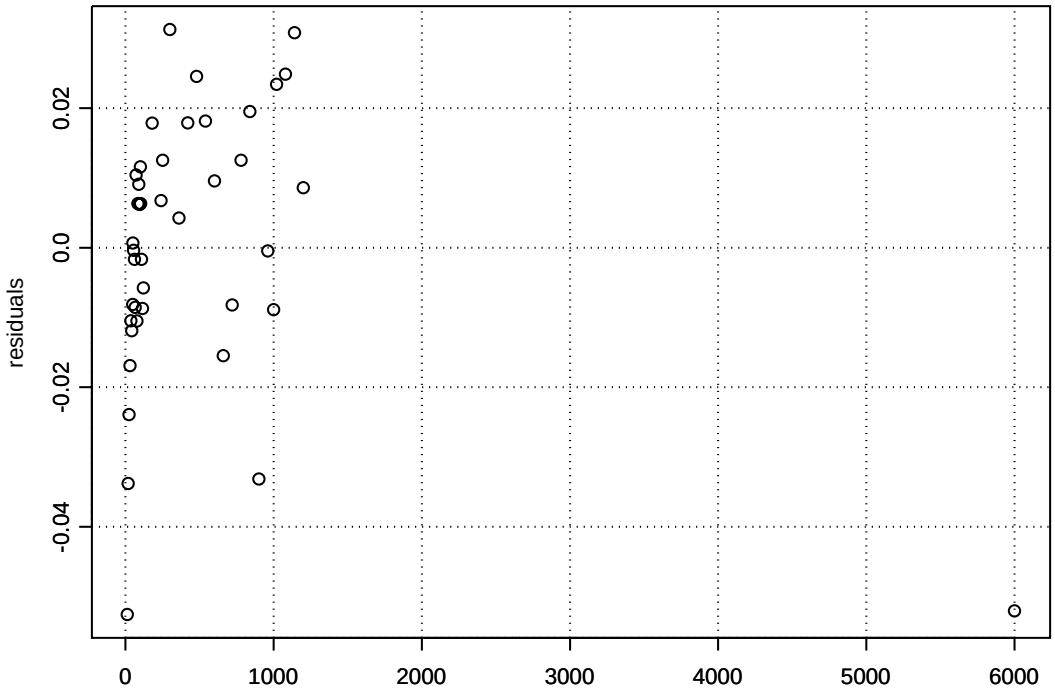


Abbildung 4.27: Verlauf der Residuen des Modells (4.36).

4.7 Verhalten unter H_0

Idealerweise sollte in einer zufälligen Beobachtungsreihe die Bruchwahrscheinlichkeit unabhängig von der Lage der Bruchstelle innerhalb der Zeitreihe sein, doch schon Hawkins (1977) zeigte, dass unter H_0 die Epochen, bei welchen T_0 angenommen wird, nicht uniform verteilt sind. Vielmehr ist die Wahrscheinlichkeit, dass die Maxima der Testgrösse an den Rändern der Zeitreihe angenommen werden, erheblich erhöht. Alexandersson (1986) schreibt sogar: "This is a disadvantage of the test, and it may be worth while to work out a 'better' test statistics that suppresses the ends of the series".

Abbildung 4.28 zeigt ein Beispiel einer derartigen Zeitreihe, bei welcher ein schwach signifikantes T_0 in der Nähe des Randes angenommen wird. Man erkennt, dass dies durch die letzten sechs Beobachtungen verursacht wird, welche zufälligerweise alle auf derselben (negativen) Seite des Mittelwertes liegen.

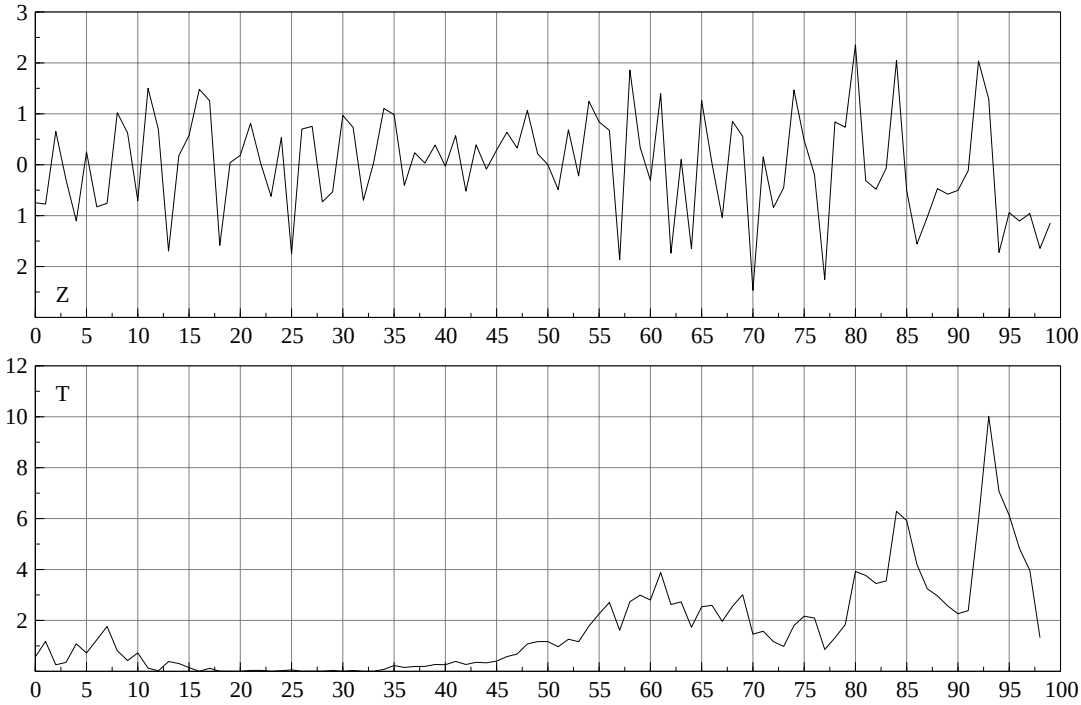


Abbildung 4.28: Simulierte Zeitreihe $Z_t \sim \mathsf{IN}(0,1)$ der Länge $n = 100$ mit einem schwach signifikanten Bruch in der Nähe des rechten Randes.

In Abbildung 4.29 ist für Zeitreihen mit Umfang $n = 50$, 100, 250 und 1000 dargestellt, wie oft das Maximum T_0 der Testgrösse T_k nach Alexandersson auf eine Epoche k fällt.

Die "Badewannen" sind bis auf Simulationsfehler symmetrisch bezüglich $n/2$, die einzelnen Kurven sind sich jedoch nicht ähnlich. Vielmehr konzentrieren sich die Häufigkeiten der Maxima mit steigendem n immer mehr an den Anfängen und an den Enden der Zeitreihen:

Satz 4.4 *Unter H_0 gilt:*

$$\frac{\hat{k}}{n} \xrightarrow{d} \xi_0 \tag{4.37}$$

mit

$$\hat{k} = \min\{k : Z_n = \max_{1 \le k < n} (-2 \log \Lambda_k)\}$$

wobei

$$P\{\xi_0 = 0\} = P\{\xi_0 = 1\} = \frac{1}{2}.$$

Beweis Csörgő und Horváth (1997, Theorem 1.6.1).

Nicht alle dieser Maxima markieren auch eine signifikante Bruchstelle. Hierzu muss die maximale Testgrösse die kritische Grenze übertreffen. In Abbildung 4.30 ist für Zeitreihen mit Umfang $n = 50, 100, 250$ und 1000 dargestellt wie oft nach dem Alexandersson Test ein signifikanter Bruch auf eine Epoche k fällt.

Die Verteilungen der Maxima- und Bruchhäufigkeiten sind für den LR3, LR4, LR5, LR6, LR7 sowie für den Alexandersson Test mit unbekanntem μ exakt dieselben, wie Tabelle C.1 entnommen werden kann. Für den LR1 und den LR2 Test mit bekanntem μ resultieren ebenfalls identische, jedoch "einseitige" Badewannen. In Abbildung 4.31 und 4.32 sind die zu den Abbildungen 4.29 und 4.30 korrespondierenden Verteilungen wiedergegeben.

Wir untersuchen das Verhältnis der signifikanten Brüche im Verhältnis zu den auftretenden Maxima. Hierbei fällt auf, dass der lokale Anteil signifikanter Maxima pro Bruchstelle von der Lage der Bruchstelle abhängig ist: für $k \to 0$ und $k \to n$ ist der Anteil signifikanter Brüche pro Bruchstelle kleiner als 5%, für $k \to 1/2$ ist der Anteil grösser (Abb. 4.33). Der Alexandersson Test und mit ihm auch der LR3, LR4, LR5, LR6 und LR7 Test sind somit an den Rändern der Zeitreihen lokal zu konservativ, in der Mitte der Zeitreihen lokal zu liberal. Dies drückt sich auch in lokal leicht unterschiedlichen kritischen Grenzen aus, wie in Abbildung 4.37 für Zeitreihen der Länge $n = 100$ gezeigt wird. Hierbei wird auch deutlich, dass die lokale Verteilung der Maxima in den Mitten der Zeitreihen langschwänziger ist als an deren Rändern. Ähnliches gilt auch für den LR1 und den LR2 Test (Abb. 4.35). Diese lokalen Betrachtungen dürfen jedoch nicht darüber hinweg täuschen, dass die Likelihood Ratio Tests gemessen am globalen Signifikanzniveau am Ende bzw. an den Rändern der Zeitreihen zu liberal und in den Mitten zu konservativ sind, wie den Abbildungen 4.34 und 4.36 entnommen werden kann.

4.7.1 Autokorrelation

Schon beim ersten flüchtigen Augenschein wird klar, dass die Likelihood Ratio Tests in autokorrelierten Datensätzen ganz andere Eigenschaften besitzen als in unabhängigen Datensätzen: Wie den Abbildungen im Anhang C entnommen werden kann, werden die "Badewannen"-Verteilungen der Maximumsepochen mit steigender Autokorrelation immer mehr zu "schiefen Gebirgen". Auch werden die in Anhang C wiedergegebenen kritischen Grenzen mit steigender Autokorrelation der Beobachtungen rasch grösser. Dadurch werden mit den originalen kritischen Grenzen umgekehrt auch fast alle auftretenden Maxima als signifikant betrachtet, wie den Abbildungen der Verteilungen der signifikanten Bruchstellen in Anhang C entnommen werden kann.

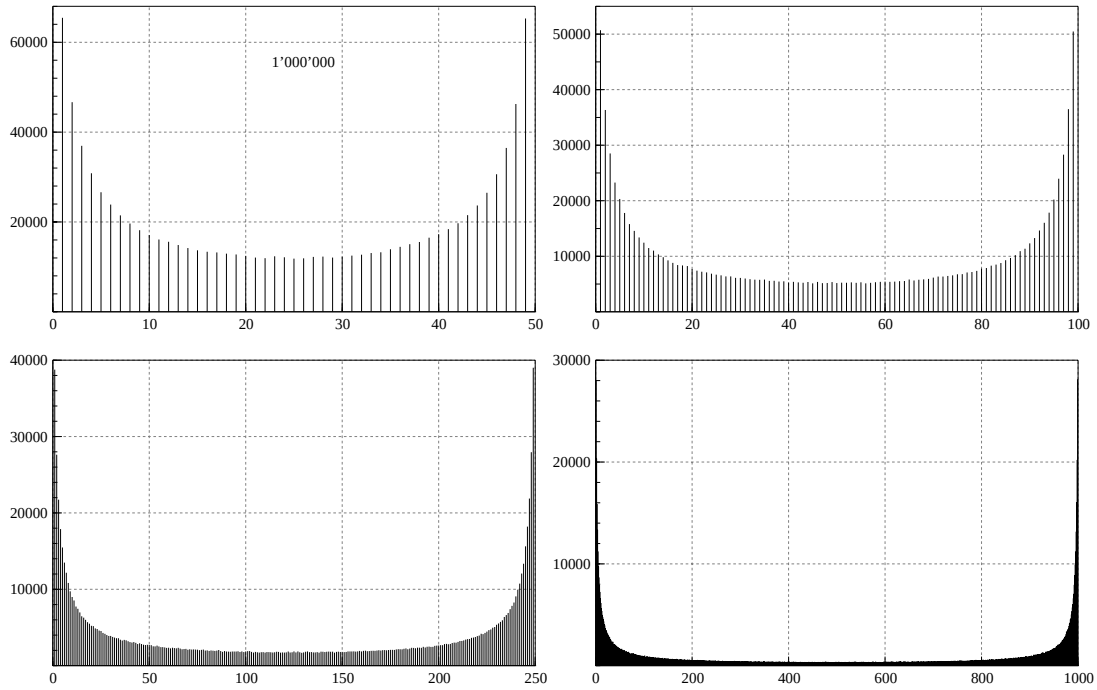


Abbildung 4.29: Häufigkeit mit der für 1'000'000 Replikationen das Maximum T_0 der Alexandersson Testgrösse bei der Epoche k (x-Achse) angenommen wird.

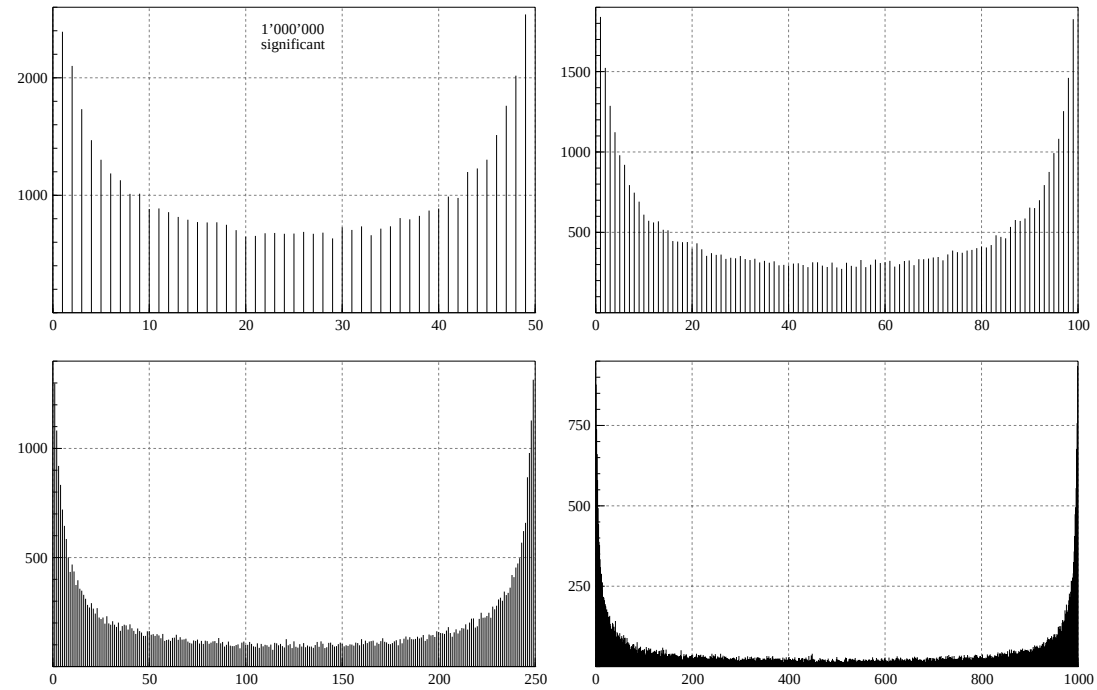


Abbildung 4.30: Häufigkeit mit der für 1'000'000 Replikationen bei der Epoche k (x-Achse) ein signifikanter Bruch entdeckt wird.

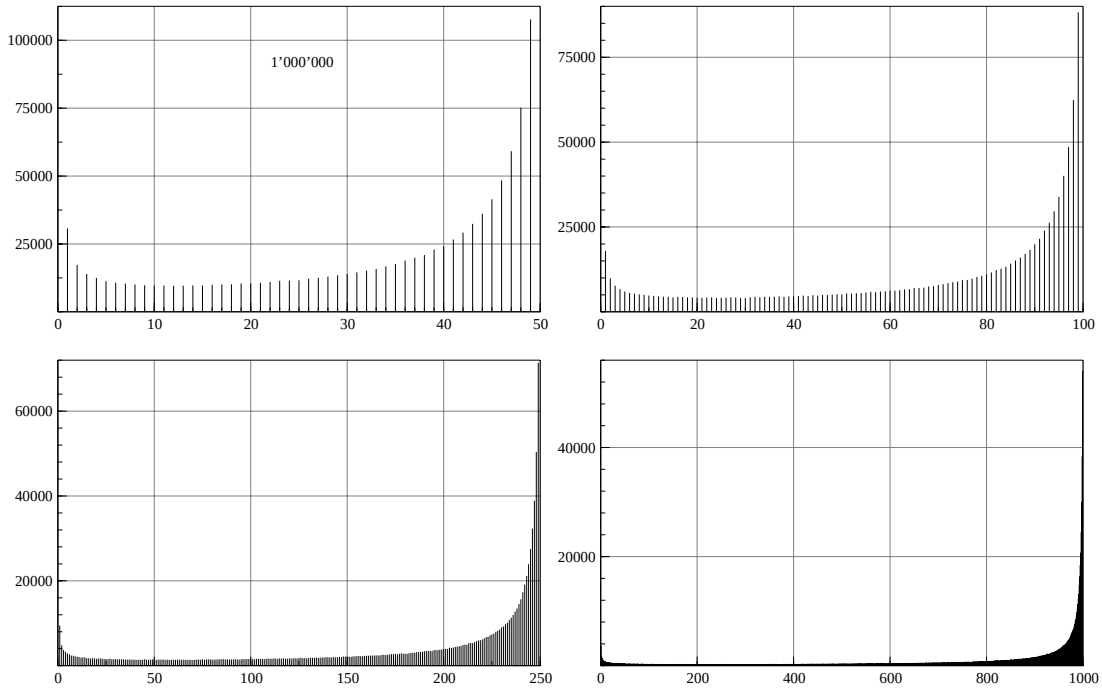


Abbildung 4.31: Häufigkeit mit der für 1’000’000 Replikationen das Maximum T_0 der LR1 Testgrösse bei der Epoche k angenommen wird.

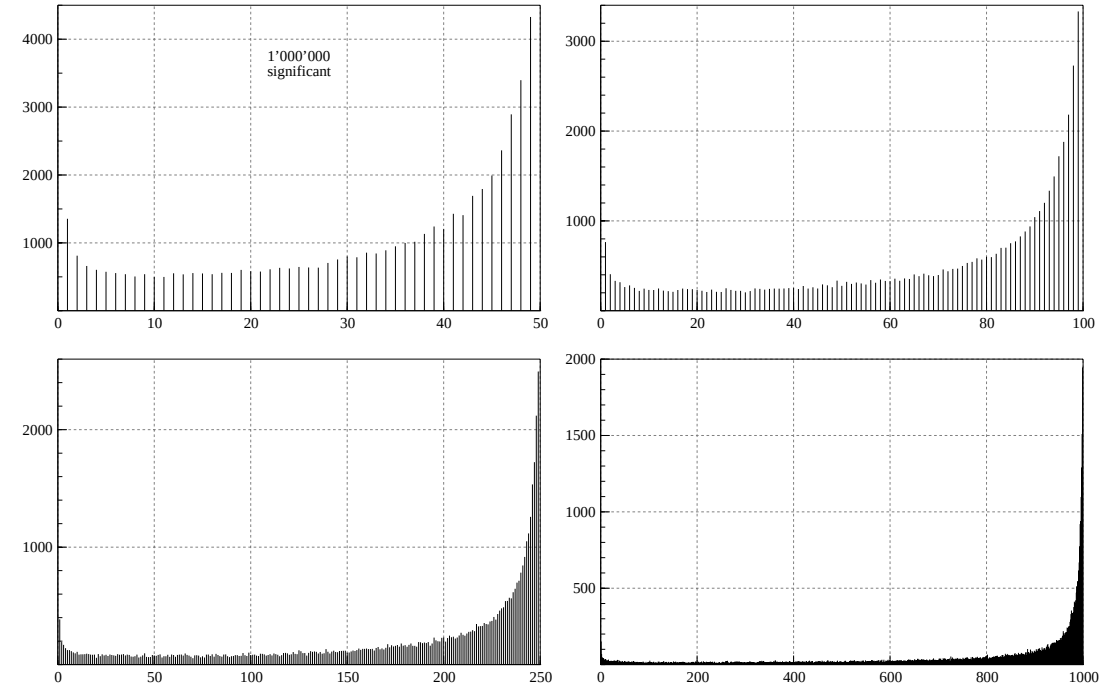


Abbildung 4.32: Häufigkeit mit der für 1’000’000 Replikationen bei der Epoche k ein signifikanter Bruch entdeckt wird.

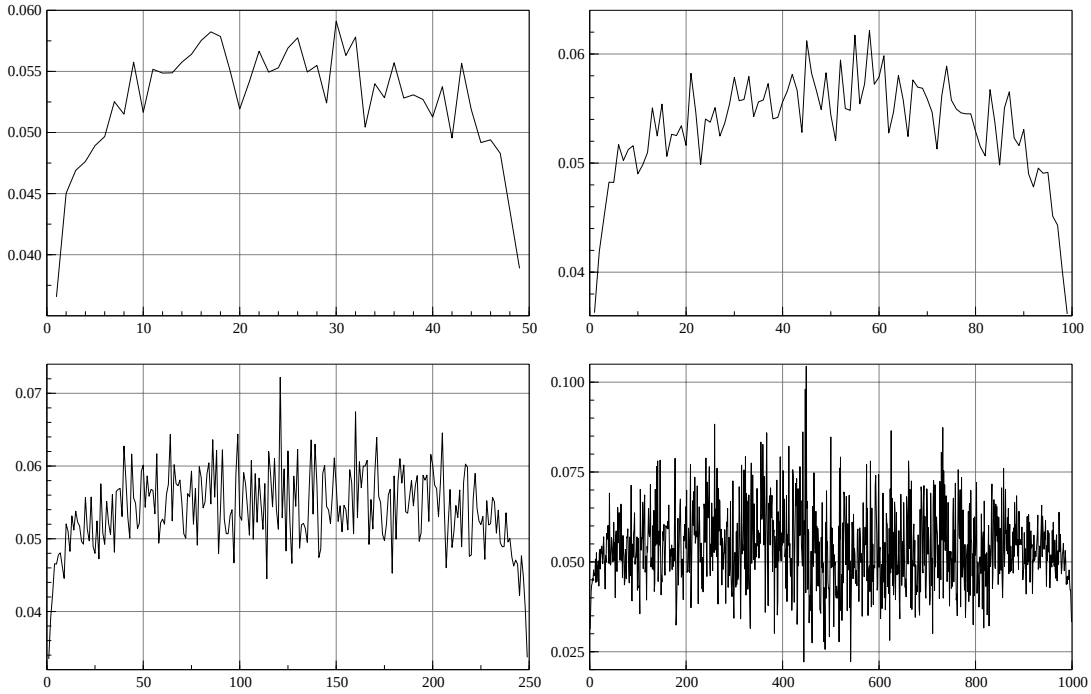


Abbildung 4.33: Lokaler Anteil signifikanter Maxima des Alexandersson Tests für 1’000’000 Replikationen.

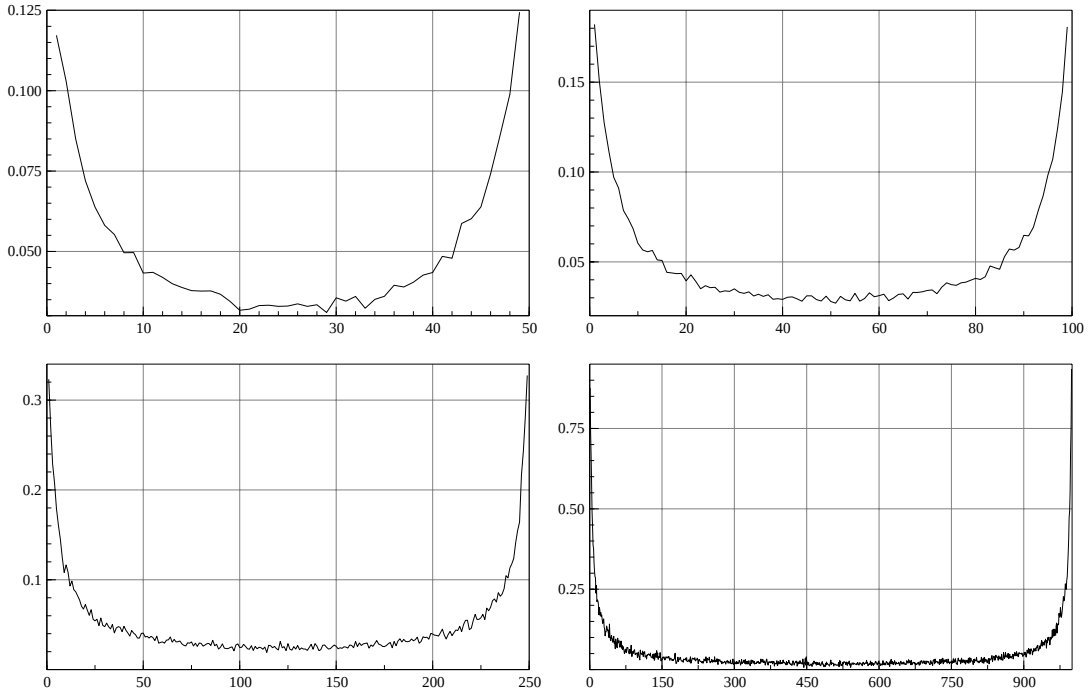


Abbildung 4.34: Lokale Realisierung des globalen Fehlers 1. Art des Alexandersson Tests für 1’000’000 Replikationen.

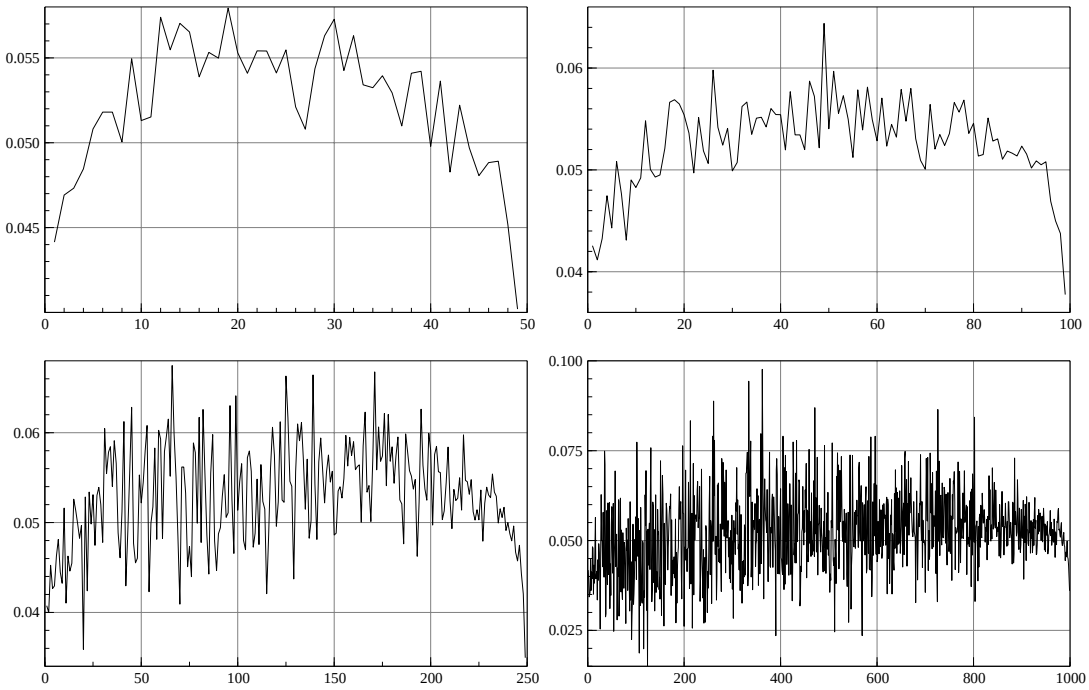


Abbildung 4.35: Lokaler Anteil signifikanter Maxima des LR1 Tests für 1’000’000 Replikationen.

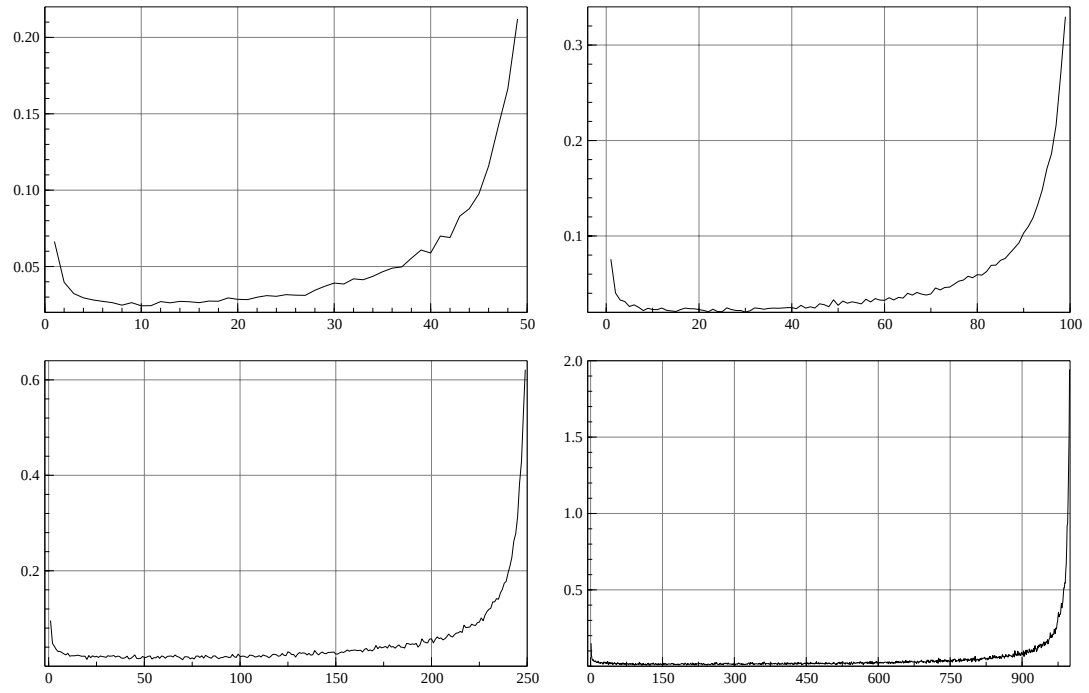


Abbildung 4.36: Lokale Realisierung des globalen Fehlers 1. Art des LR1 Tests für 1’000’000 Replikationen.

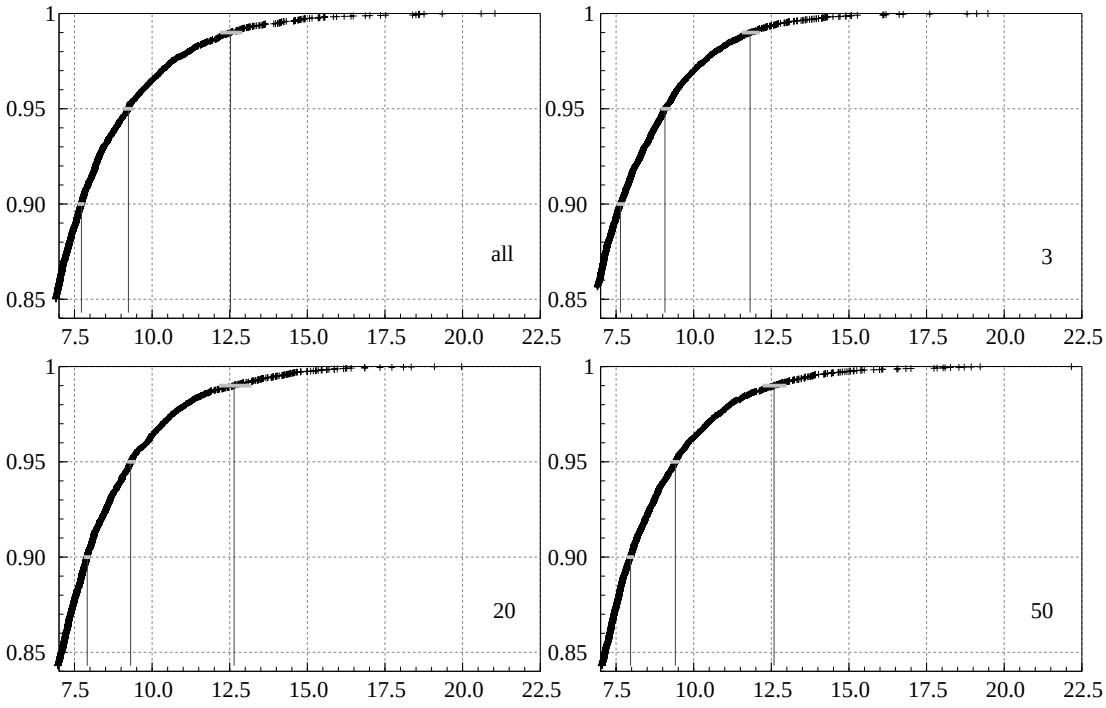


Abbildung 4.37: Simulation der kritischen Grenzen des Alexandersson Tests für drei verschiedene Epochen einer Zeitreihe der Länge $n = 100$ basierend auf je 10'000 dort realisierter Maxima.

4.7.2 Buishand Test

Die Verteilungen der Maxima- und Bruchhäufigkeiten sehen für den Buishand Test ganz anders aus, als diejenigen der Likelihood Ratio Tests, wie den Abbildungen 4.38 und 4.39 entnommen werden kann: Aus den "Badewannen" sind "Gebirge" geworden. Im Unterschied zu den Likelihood Ratio Tests zeigt der Buishand Test jedoch eine weit stärkere Abhängigkeit des Anteils signifikanter Maxima von der Epoche k, wie der Abbildung 4.40 entnommen werden kann: Für $k \to 0$ und $k \to n$ nimmt der Anteil signifikanter Maxima stark ab, ja an den 10% grössten und kleinsten k werden überhaupt keine signifikanten Brüche und beinahe keine Maximumsepochen mehr beobachtet. Für kleine und grosse k ist der Buishand Test daher insensitiv, während er in der Mitte der Zeitreihen mit einem Anteil signifikanter Maxima von ca. 8% sogar noch liberaler ist als die Likelihood Ratio Tests. Dies drückt sich auch in deutlich unterschiedlichen kritischen Grenzen aus, wie in Abbildung 4.42 für Zeitreihen der Länge $n = 100$ gezeigt wird, sind doch die kritischen Grenzen für $k = 3$ noch etwa halb so gross, wie diejenigen für $k = n/2$! Die lokale Realisierung des globalen Fehlers 1. Art findet sich in Abbildung 4.41.

In autokorrelierten Datensätzen steigen die kritischen Grenzen mit steigender Autokorrelation der Beobachtungen rasch an, wodurch beinahe alle auftretenden Maxima als signifikant betrachtet werden, wie den Abbildungen im Anhang C entnommen werden kann. Infolge des fehlenden von k abhängigen Vorfaktors der Testgrösse, bleiben die Verteilungen der Maxima- und Bruchepochen jedoch im Gegensatz zu denjenigen der Likelihood Ratio Tests unverzerrt.

Wie Friedli und Schüpbach (2001) gezeigt haben, ist der Buishand Test ebenso wie die Likelihood Ratio Tests zur Homogenitätsprüfung autokorrelierter Datensätze unbrauchbar.

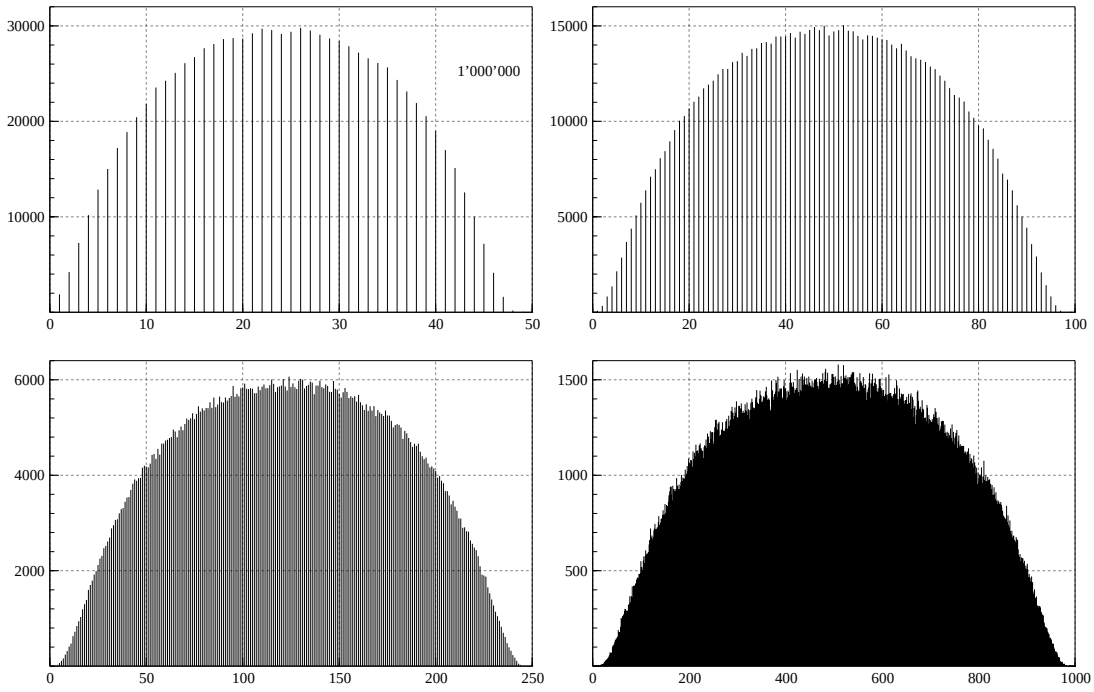


Abbildung 4.38: Häufigkeit mit der für 1’000’000 Replikationen das Maximum Q der Buishand Testgrösse bei der Epoche k angenommen wird.

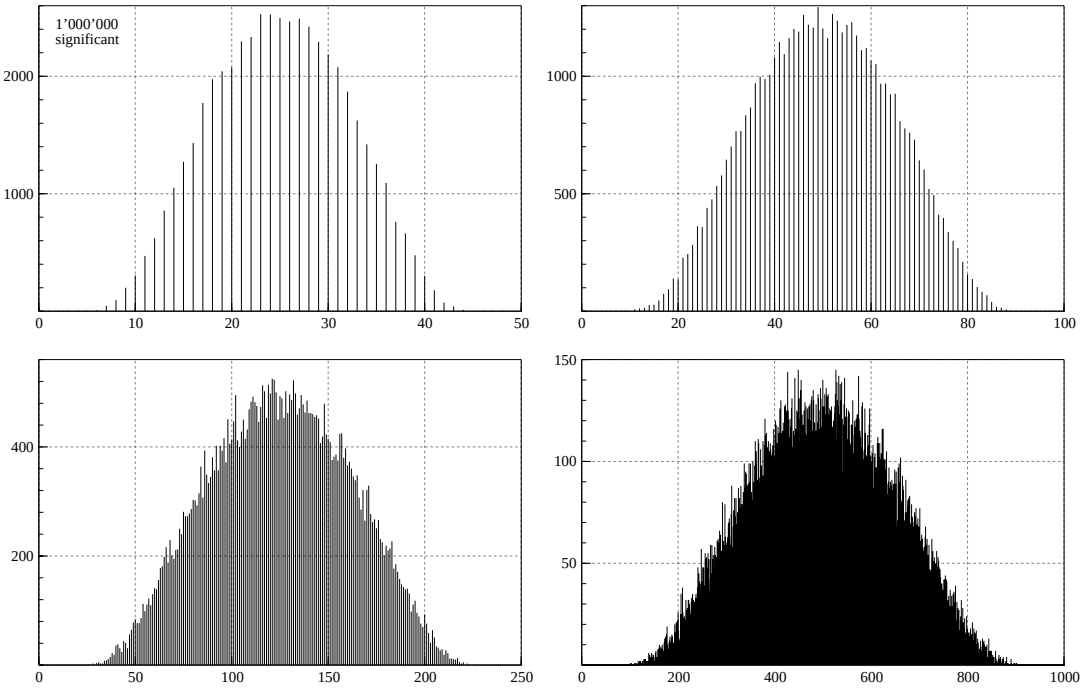


Abbildung 4.39: Häufigkeit mit der für 1’000’000 Replikationen bei der Epoche k ein signifikanter Bruch entdeckt wird.

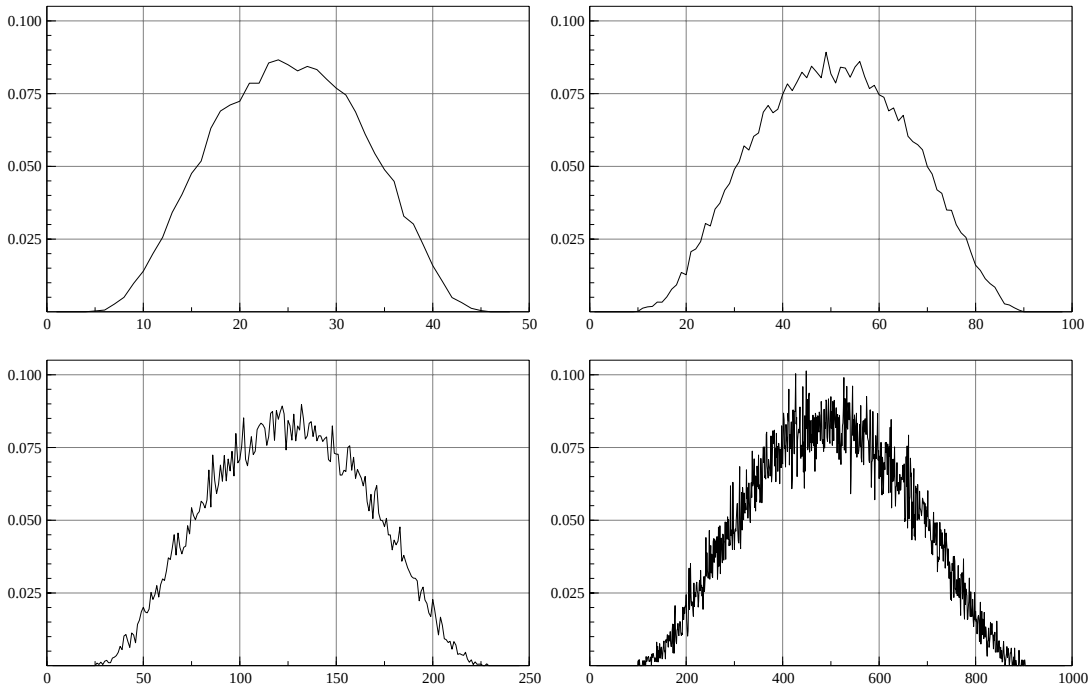


Abbildung 4.40: Lokaler Anteil signifikanter Maxima des Buishand Tests für 1’000’000 Replikationen.

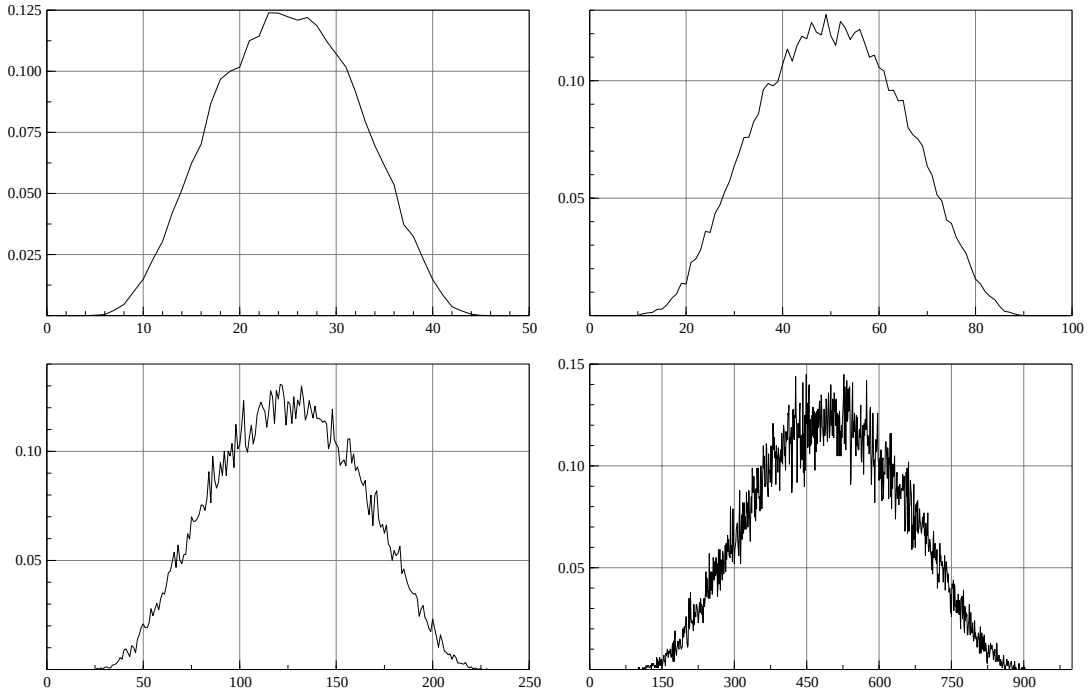


Abbildung 4.41: Lokale Realisierung des globalen Fehlers 1. Art des Buishand Tests für 1’000’000 Replikationen.

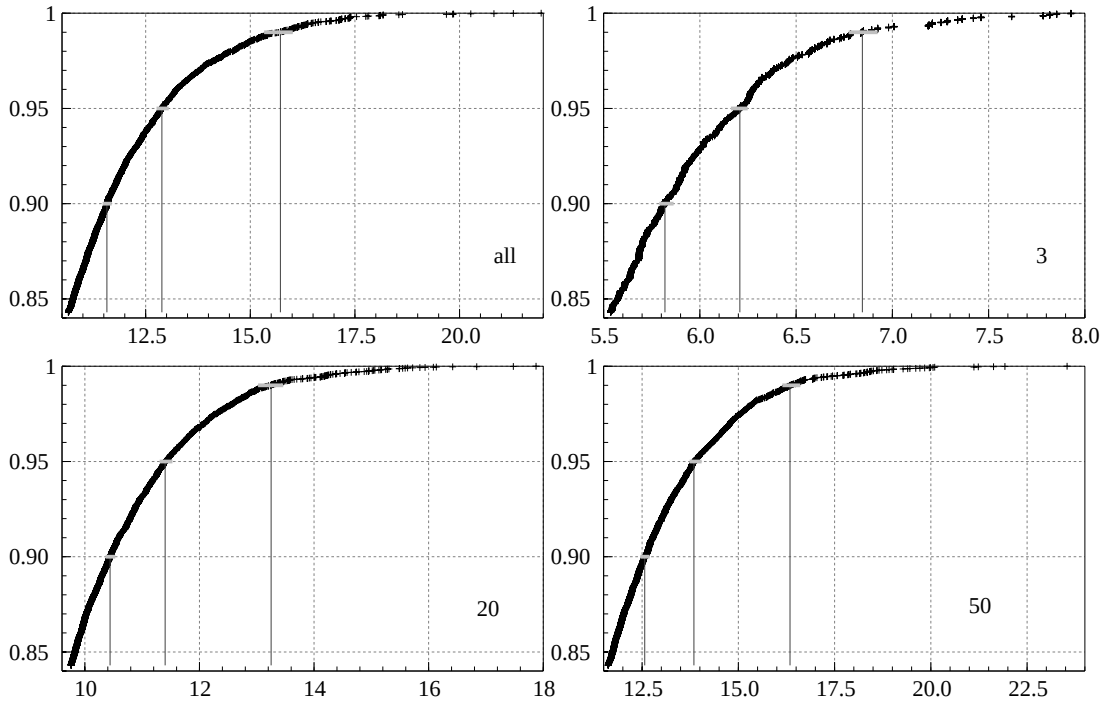


Abbildung 4.42: Simulation der kritischen Grenzen des Buishand Tests für drei verschiedene Epochen einer Zeitreihe der Länge $n = 100$ basierend auf je 10'000 dort realisierter Maxima.

4.8 Verhalten unter H_1

4.8.1 Güte

Die Resultate der Gütesimulationen sind im Anhang C zusammengestellt. Die Gütefunktion ist abhängig von der Anzahl geschätzter Parameter, von der Länge der Zeitreihe sowie von der Lage der Bruchstelle innerhalb der Zeitreihe. Für den LR4, LR5, LR6, LR7 und den Alexandersson Test sind die Gütefunktionen identisch (Abb. 4.43). Das heisst insbesondere, dass die Gütefunktion unabhängig ist von der Art und Weise wie s_k berechnet wird.

Dies steht im Widerspruch zu Antoch et al. (2002, p.16) und Csörgő und Horváth (1997, p. 32), welche zeigten, dass diejenigen Tests, deren s_k nach Formeln des Typs (4.7), (4.10) oder (4.12) geschätzt wurden, eine grössere Power besitzen als diejenigen Tests, deren s_k nach Formeln des Typs (4.14) oder (4.16) geschätzt wurden. Dies stimmt jedoch nur, wenn für die Güteberechnung die asymptotischen Grenzen nach den Formeln (4.22) und (4.23) verwendet werden. Da diese für Zeitreihen endlicher Länge zu konservativ sind, sind diejenigen Tests, deren Testgrössen betragsmässig am grössten ausfallen, auch diejenigen mit der grössten Macht. Die Grösse der Teststatistiken hängt aber direkt von s_k ab: je kleiner s_k, desto grösser T_k. Somit resultieren mit (4.10) die kleinsten, mit (4.14) die grössten s_k, so dass die T_k des LR4 Tests die grössten, die T_k des LR6 Tests die kleinsten Werte liefern. Da unsere Powerfunktionen jedoch für jeden Test mit den entsprechenen empirischen kritischen Grenzen gerechnet wurden und die Berücksichtigung verschiedener

s_k an der Lage der Maxima innerhalb der Zeitreihen nichts ändert, sind die Gütefunktionen für alle Tests mit denselben bekannten und unbekannten Parametern unabhängig von der Berechnungsart von s_k. Die Unterschiede der in Abbildung 4.44 zusammengestellten Gütefunktionen für einfache Versatzinhomogenitäten in der Mitte der Zeitreihen sind daher allein darauf zurückzuführen, dass jeweils die Mittelwerte und Standardabweichungen der Beobachtungen bekannt sind oder geschätzt werden müssen. Der mächtigste Test ist der LR1 Test mit bekanntem μ und σ. der schwächste der LR4 Test mit unbekanntem μ und σ. Muss μ geschätzt werden, sinkt die Power bedeutend mehr, als wenn σ unbekannt ist.

Die in Abbildung 4.45 zusammengestellten Gütefunktionen an den Rändern der Zeitreihen können nicht mehr miteinander verglichen werden, da die LR1 und LR2 Tests bei kleinen k aufgrund ihres "einseitigen" Verhaltens eine weit bessere Power aufweisen, als die LR3 und LR4 Tests mit ihrem "zweiseitigen" Verhalten.

Es zeigt sich, dass alle Tests unverfälscht sind, d.h. das Minimum der Gütefunktion wird bei Sprunghöhe 0 angenommen und liegt bei 5%. Auch sind die Tests konsistent in dem Sinne, dass bei $n \to \infty$ beliebig kleine Effekte nachgewiesen werden können. Für die Gütefunktionen des LR3 und des LR4 Tests scheint dies für $k \to 0$ auf den ersten Blick nicht erfüllt zu sein (Abb. 4.43) liegen doch die Gütefunktionen für die Zeitreihen der Länge 1000 bei $k = 3$ tiefer als diejenigen für die Zeitreihen der Länge 50. Dies ist jedoch damit zu erklären, dass $k = 3$ für Zeitreihen der Länge $n = 1000$ relativ gesehen viel randnaher ist als $k = 3$ für Zeitreihen der Länge $n = 50$. Da in unmittelbarer Randnähe die globalen kritischen Grenzen stark konservativ sind, ist dort der Powerverlust grösser als bei den kürzeren Zeitreihen. Dieser Effekt wird für $n \to \infty$ erst bei sehr grossen n aufgehoben, wenn sich die Bruchwahrscheinlichkeit nach (4.37) soweit ihrer Grenzverteilung angenähert hat, dass $k = 3$ nicht mehr in der Flanke der "Badewanne" liegt.

4.8.2 Schätzung von Bruchepoche und Sprunghöhe

Wie die Gütefunktionen hängen die in Anhang C wiedergegebenen Schätzungen der Bruchepochen und Sprunghöhen neben der Länge der Zeitreihe und der Lage der Bruchstelle innerhalb der Zeitreihe auch davon ab, welche Parameter bekannt sind und welche geschätzt werden müssen. Für den LR4, LR5, LR6, LR7 und den Alexandersson Test sind die Resultate identisch.

Für Sprunghöhen welche eine Power von mehr als 95% aufweisen, werden die Bruchepochen und die Sprunghöhen korrekt geschätzt. Für kleinere Sprunghöhen muss damit gerechnet werden, dass die Bruchepoche gegen die Mitte der Zeitreihe verfälscht, sowie die Sprunghöhe systematisch überschätzt wird.

4.8.3 Autokorrelation

Wie bereits in Abschnitt 4.7 diskutiert, sind die Likelihood Ratio Tests LR1, LR2, LR3, LR4, LR5, LR6, LR7 sowie der Alexandersson Test zur Homogenitätsprüfung autokorrelierter Datensätze unbrauchbar. Den dort angeführten Argumenten kann noch hinzugefügt werden, dass gemäss den im Anhang C exemplarisch für den LR1 Test wiedergegebenen Resultaten die Schätzung der Bruchepochen auch bei genügender Power systematisch verfälscht, die Sprunghöhen sogar massiv überschätzt werden.

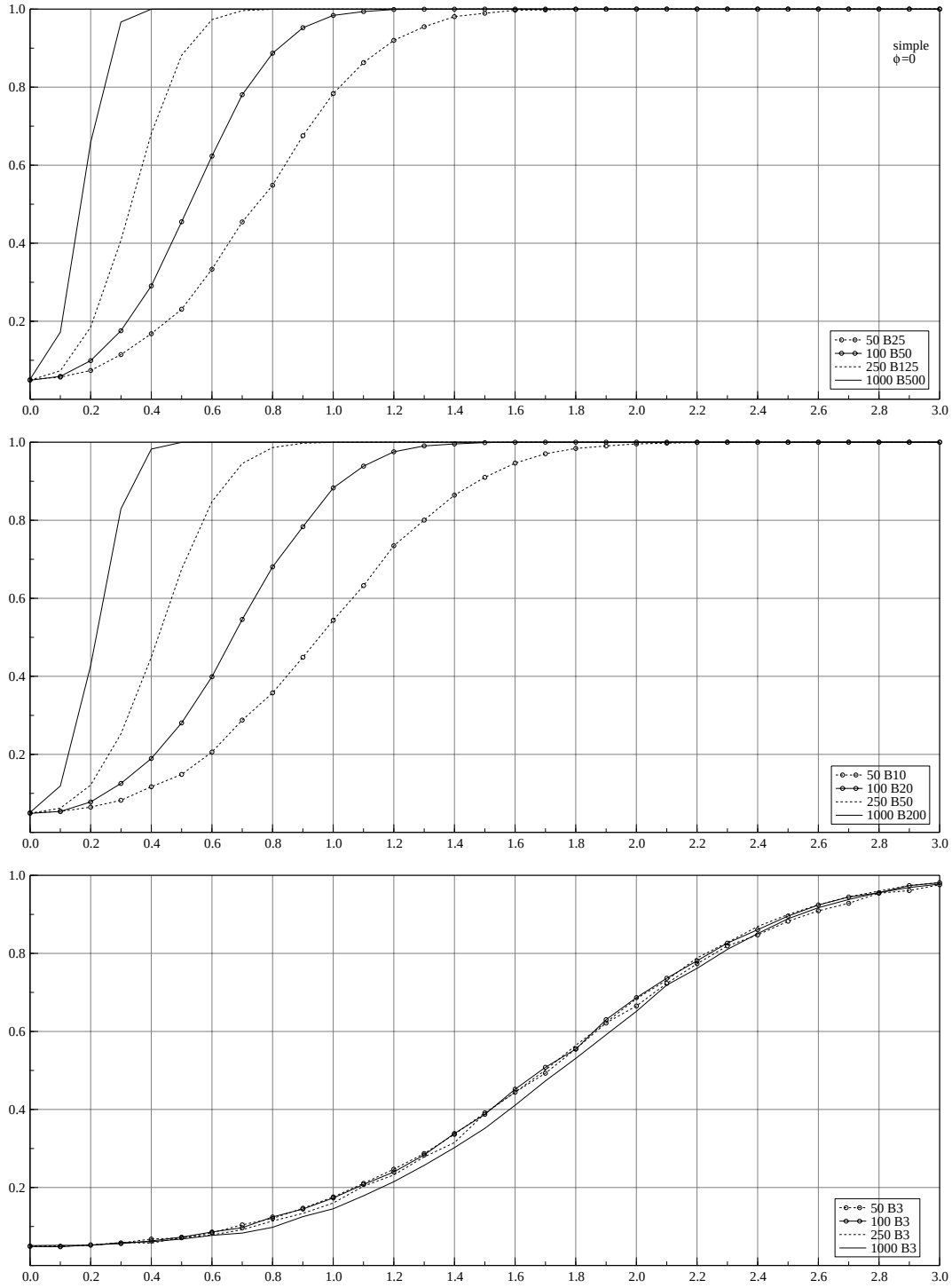


Abbildung 4.43: Güte des Alexandersson Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $n = 50$, 100, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

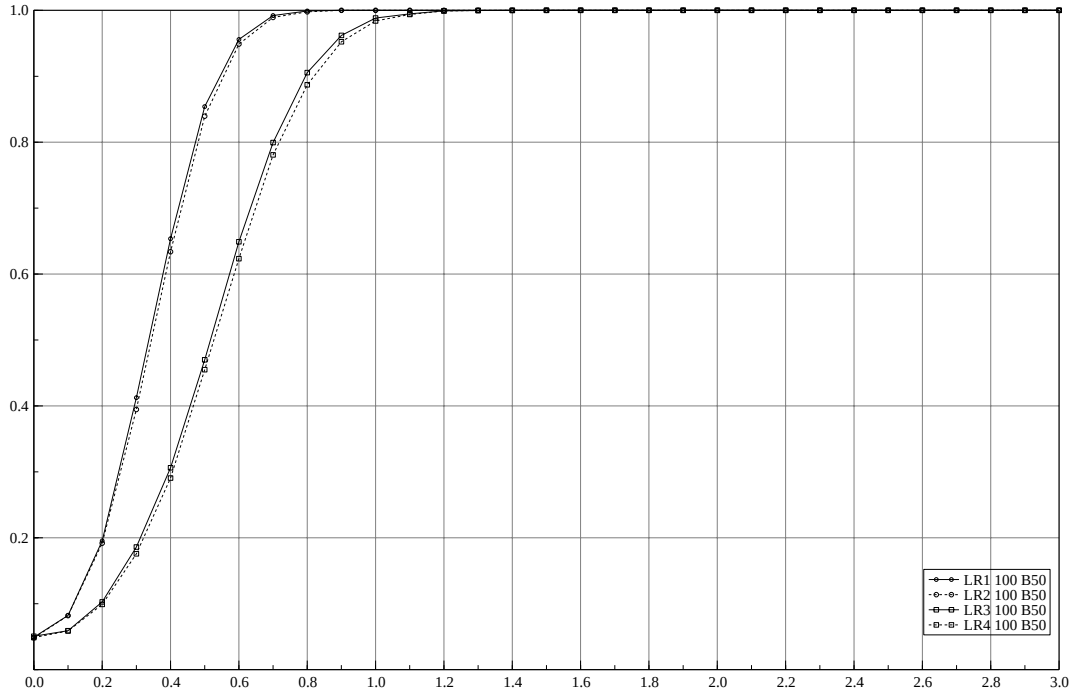


Abbildung 4.44: Vergleich der Powerkurven des LR1, LR2, LR3 und LR4 Tests für einfache Versatzinhomogenitäten bei Epoche $k = 50$ in Zeitreihen der Länge $n = 100$.

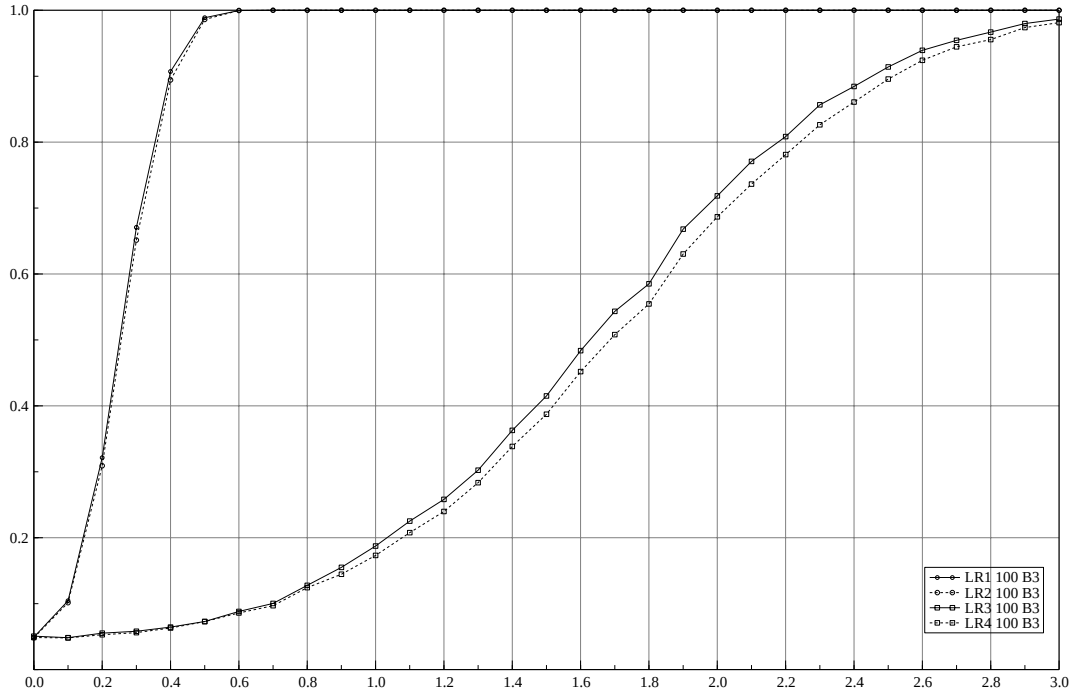
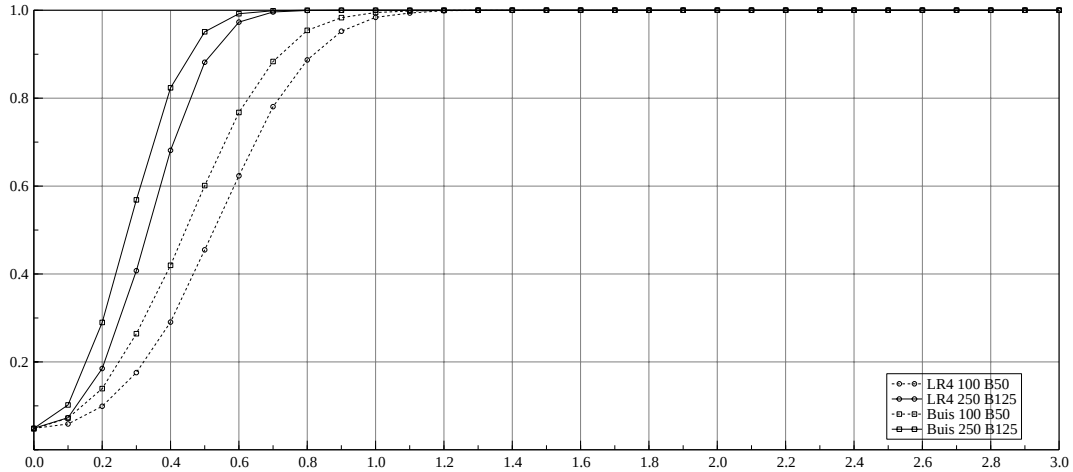
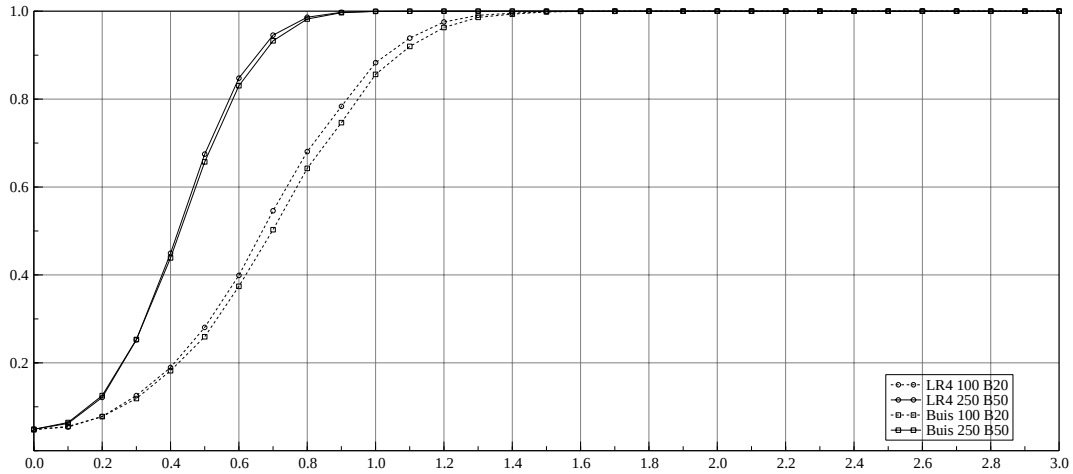
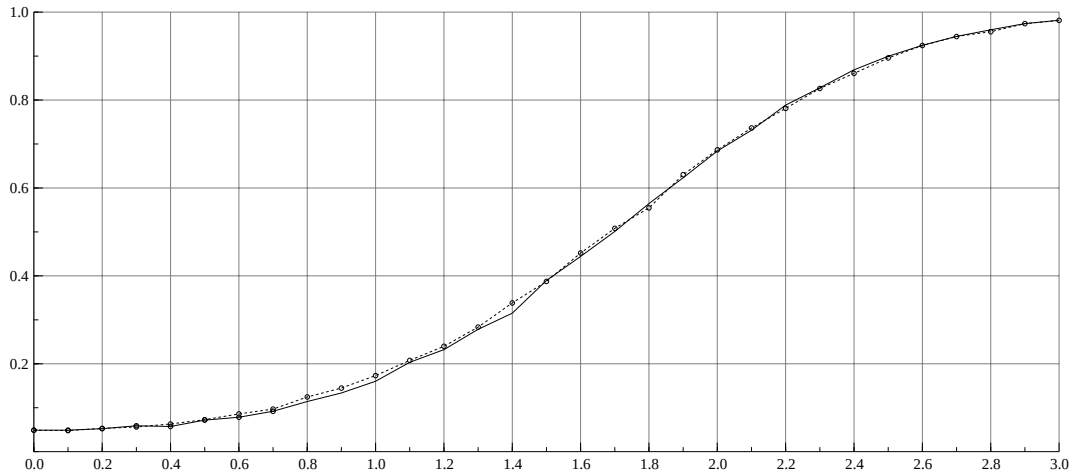


Abbildung 4.45: Vergleich der Powerkurven des LR1, LR2, LR3 und LR4 Tests für einfache Versatzinhomogenitäten bei Epoche $k = 3$ in Zeitreihen der Länge $n = 100$.


LR4 100 B50
LR4 250 B125
Buis 100 B50
Buis 250 B125


LR4 100 B20
LR4 250 B50
Buis 100 B20
Buis 250 B50



Daraus lassen sich folgende Empfehlungen ableiten:

- Der Buishand Test sollte zur Homogenitätsprüfung klimatologischer Zeitreihen nicht mehr verwendet werden, da seine kritischen Grenzen zu wenig genau bekannt sind und die Einhaltung des Fehlers 1. Art stark von der Lage der möglichen Bruchstelle innerhalb der Zeitreihe abhängt. Auch sind die Schätzungen der Bruchepochen und Sprunghöhen stärker verfälscht als bei den Likelihood Ratio Tests.

- Es existieren zur Prüfung einfacher Mittelwertsverschiebungen in unabhängigen, normalverteilten Zufallsvariablen nur vier verschiedene Varianten des Likelihood Tests, je nachdem ob die Parameter μ und σ bekannt sind oder aus der Zeitreihe bestimmt werden müssen.

  Als Schätzer genügen der arithmetische Mittelwert aller Beobachtungen sowie die empirische Standardabweichung über alle Beobachtungen.

  Die asymptotischen kritischen Grenzen sind für alle vier Testvarianten für alle praktisch relevanten Umfänge zu konservativ. Die kritischen Grenzen müssen daher simuliert werden.

  Hauptnachteil der Likelihood Ratio Tests ist die nichtuniforme Verteilung der Bruchhäufigkeiten entlang der Zeitreihe.

  Die in der Literatur insbesondere von Antoch et al. (2002) empfohlene Ausblendung gewisser Anteile der grössten und kleinsten Beobachtungen widerspricht den praktischen Anforderungen an einen Homogenitätstest. Zwar können hiermit die kritischen Grenzen genauer geschätzt und die Einhaltung des Fehlers 1. Art in den zentralen Abschnitten der Zeitreihen besser kontrolliert werden, doch dürften demgegenüber diejenigen Maxima der Teststatistik welche in diese meist besonders interessanten Regionen der Zeitreihen fallen nicht interpretiert werden.

5. Verbesserte Homogenitätstests

In diesem Kapitel wird ein Test konstruiert, welcher unter H_0 eine uniforme, d.h. von der Lage innerhalb der Zeitreihe unabhängige, Verteilung der Bruchhäufigkeit besitzt.

Wir beschränken uns auf den Fall mit unbekanntem Mittelwert μ und unbekannter Standardabweichung σ, da allein diese Testvariante zur Homogenitätsprüfung klimatologischer Beobachtungsreihen relevant ist.

Die den drei anderen Fällen äquivalenten Teststatistiken lassen sich auf analogem Wege verbessern.

5.1 Modifizierte Teststatistiken

Im letzten Kapitel wurde gezeigt, dass die Verteilung der Maximumshäufigkeit des LR4 Tests und diejenige des Buishand Tests ein komplementäres Verhalten aufweisen: während beim LR4 Test an den Enden der Zeitreihen eine starke Häufung der Maxima auftritt, fallen unter der Buishand Teststatistik beinahe keine Maxima auf diese Epochen.

Da sich gemäss dem Abschnitt 4.4.2 die beiden Teststatistiken nur durch eine andere Gewichtung unterscheiden, liegt es nahe, eine Einebnung der Häufigkeitsverteilung der Maxima dadurch zu erreichen, dass eine "mittlere" Gewichtung gesucht wird, welche die "Badewanne" des LR4 Tests mit dem "Gebirge" des Buishand Tests kombiniert.

Eine modifizierte Testgrösse kann daher wie folgt konstruiert werden:

$$T_k = \frac{1}{s_k}\left[\sqrt{\frac{n}{k(n-k)}} + c_n\right]\sum_{i=1}^{k}(Y_i - \bar{Y}_n). \tag{5.1}$$

In einem ersten - LR8 genannten - Versuch, wurden mit der Standardabweichung

$$s_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2} \tag{5.2}$$

für Zeitreihen verschiedener Länge dasjenige c_n gesucht, welches die Verteilung der Maximumshäufigkeiten möglichst flach werden lässt. In Tabelle 5.1 sind für Zeitreihen der Längen $n = 25$ bis $n = 1000$ die gefundenen c_n zusammengestellt. Wie der Abbildung 5.1 entnommen werden kann, eignet sich der Ansatz (5.1) zur Glättung der Verteilung der Maximumshäufigkeiten allerdings nur für Zeitreihen bis ca. $n = 200$. Für längere Zeitreihen bleibt eine Restkurvatur.

n	c_n
25	0.3325
50	0.2750
75	0.2425
100	0.2200
200	0.1700
250	0.1600
400	0.1350
1000	0.0985

Tabelle 5.1: Koeffizienten c_n in Formel (5.1) für verschiedene Zeitreihenlängen n.

Für die praktische Anwendung muss c_n für alle n bekannt sein. Wie aus der Abbildung 5.2 ersichtlich, können die empirisch gefundenen c_n für Zeitreihen bis $n = 1000$ durch die empirische Formel

$$c_n = (0.4336 \pm 0.0023) - (0.5101 \pm 0.0050) \log\log\log n \tag{5.3}$$

mit einem mittleren Regressionsfehler von 0.0021 approximiert werden. Hiermit konnten die in Anhang C zusammengestellten kritischen Grenzen simuliert werden.

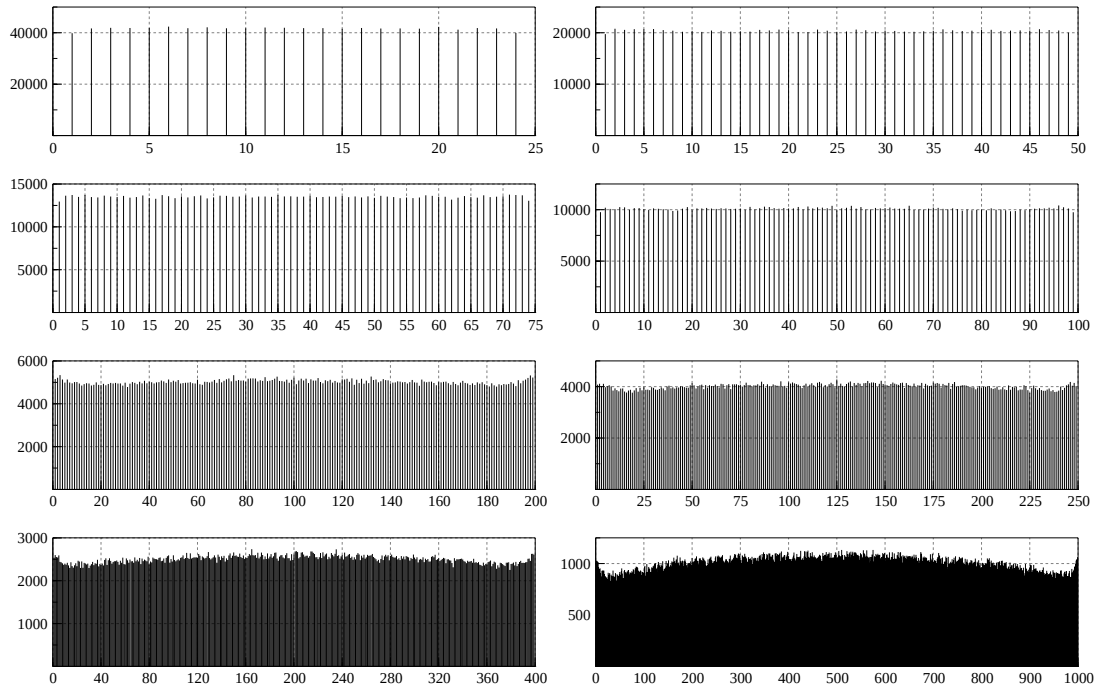


Abbildung 5.1: Vergleich der nach Formel (5.1) geglätteten Verteilungen der Maximumshäufigkeiten des LR8 Tests für 1'000'000 Replikationen.

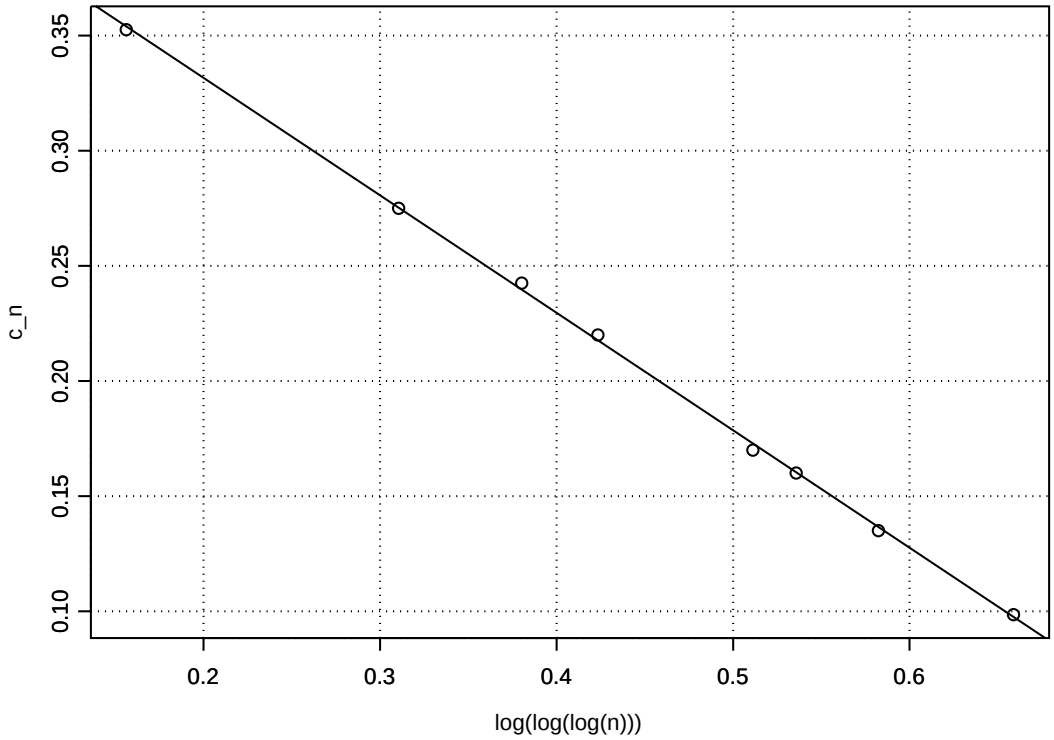


Abbildung 5.2: Verlauf der Koeffizienten c_n in Formel (5.1) in Abhängigkeit der dreifach logarithmierten Zeitreihenlängen n (Kreise) mit linearer Approximation nach Formel (5.3).

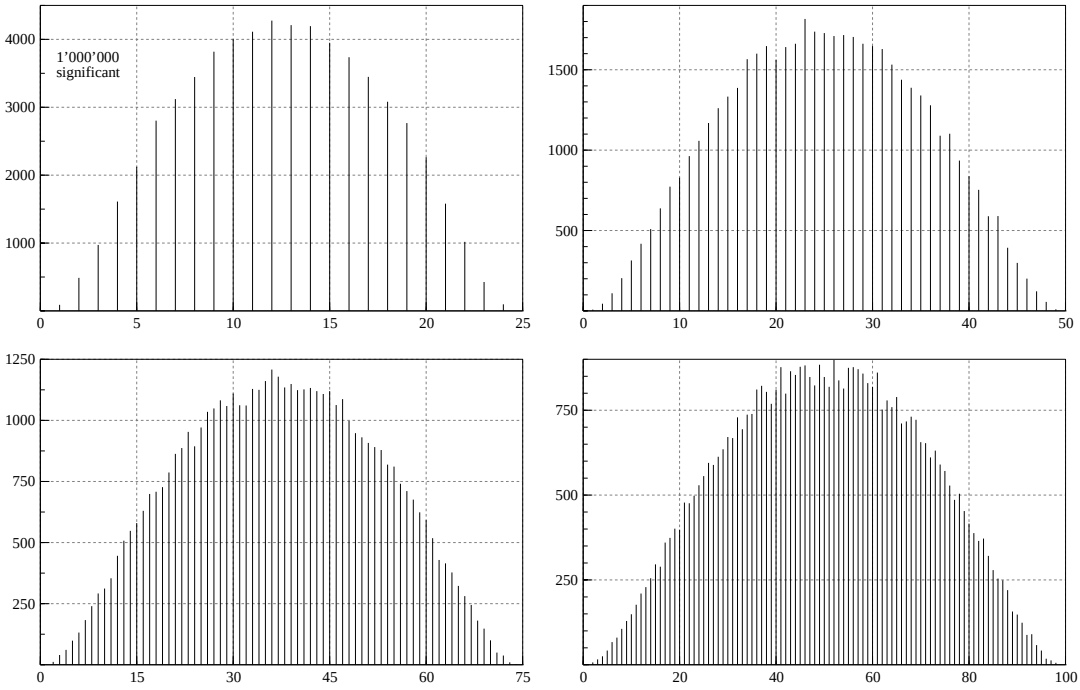


Abbildung 5.3: Verteilung der Bruchepochen der LR8 Testgrösse.

In Abbildung 5.3 sind die Häufigkeiten der signifikanten Bruchepochen dargestellt, woraus deutlich wird, dass mit der Glättung der Maximumshäufigkeiten leider nicht automatisch auch die Verteilung der Bruchhäufigkeiten geglättet wird.

Erste Simulationen ergaben, dass mit dem Ansatz (5.1) die Verteilung der Bruchhäufigkeiten nicht genügend eingeebnet werden kann. Wir wählten deshalb folgenden verbesserten Ansatz:

$$T_k = \frac{1}{s_k}\left[\sqrt{\frac{n}{k(n-k)}} + c_n\sqrt{\frac{k(n-k)}{n}} + d_n\right]\sum_{i=1}^{k}(Y_i - \bar{Y}_n). \tag{5.4}$$

In einem zweiten - LR9 genannten - Versuch wurden mit (5.2) diejenigen c_n und d_n bestimmt, welche die Verteilung der signifikanten Bruchepochen uniform werden lassen. In Tabelle 5.2 finden sich die entsprechenden Koeffizienten. Auch dieser Glättungsansatz eignet sich leider nur für Zeitreihen mit einer Länge von höchstens 200. Für längere Zeitreihen bleibt eine Restkurvatur in der Verteilung der signifikanten Bruchepochen.

Wie aus den Abbildungen 5.4 und 5.5 ersichtlich, können die empirisch gefundenen c_n und d_n für Zeitreihen der Längen $2 \leq n \leq 100$ durch die empirischen Formeln

$$c_n = (0.0556 \pm 0.0030) - (0.0361 \pm 0.0022)\log\log n \tag{5.5}$$
$$d_n = (-0.1031 \pm 0.0159) + (0.0319 \pm 0.0039)\log n \tag{5.6}$$

mit einem mittleren Regressionsfehler von 0.00059 bzw. 0.0041 approximiert werden. Hiermit konnten die in Anhang C zusammengestellten kritischen Grenzen simuliert werden.

In einem dritten - LR10 genannten - Versuch wurde mit der empirischen Standardabweichung

$$s_k = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2} \tag{5.7}$$

statt mit (5.2) gerechnet, was einer modifizierten LR7 Testgrösse entspricht. Die Verteilungen der Maximums- und Bruchhäufigkeiten sind wie erwartet identisch mit den korrespondierenden Verteilungen des LR9 Tests, da s_k an der Gewichtsstruktur innerhalb der Zeitreihe nichts ändert.

Wie bei den Likelihood Ratio Tests in Kapitel 4 ist das Verhalten der modifizierten Teststatistiken LR9 und LR10 unter H_0 und H_1 daher von der konkreten Art und Weise der Berechnung von s_k unabhängig.

n	c_n	d_n
25	0.0135	0.0000
50	0.0060	0.0230
75	0.0035	0.0300
100	0.0002	0.0470

Tabelle 5.2: Koeffizienten c_n und d_n in Formel (5.4) für verschiedene Zeitreihenlängen n.

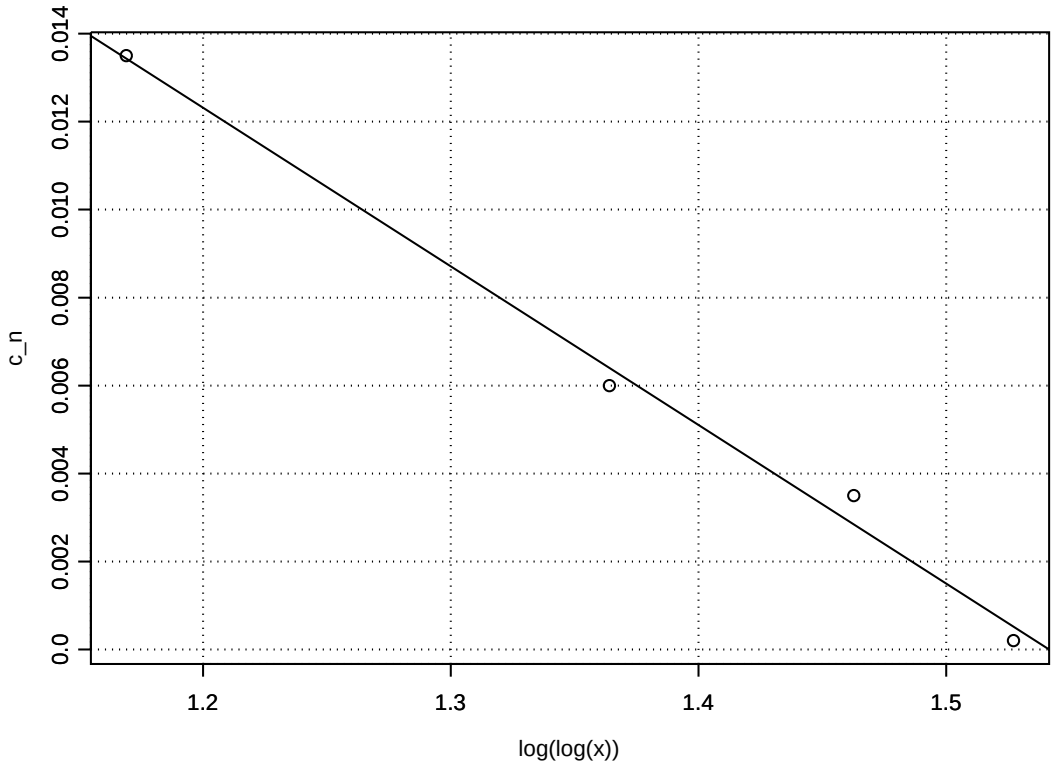


Abbildung 5.4: Verlauf der Koeffizienten c_n in Formel (5.4) in Abhängigkeit der zweifach logarithmierten Zeitreihenlängen n (Kreise) mit linearer Approximation nach Formel (5.5).

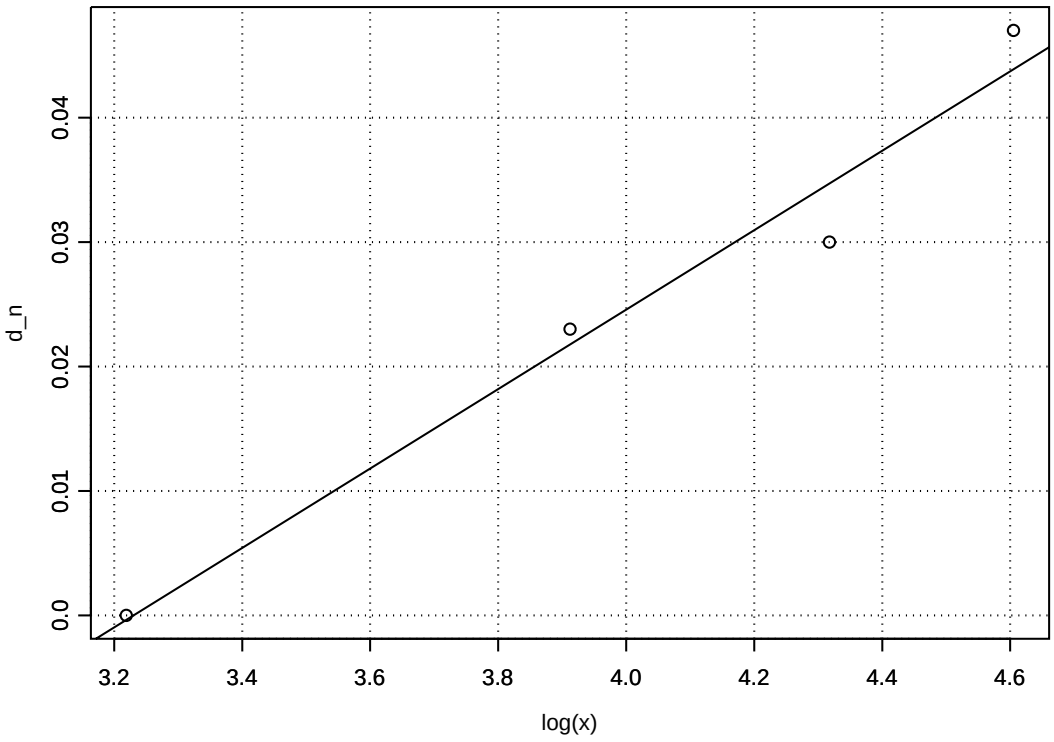


Abbildung 5.5: Verlauf der Koeffizienten d_n in Formel (5.4) in Abhängigkeit der logarithmierten Zeitreihenlängen n (Kreise) mit linearer Approximation nach Formel (5.6).

5.2 Kritische Grenzen

Die simulierten kritischen Grenzen finden sich in Anhang C. Die in den Abbildungen 5.6 und 5.7 wiedergegebenen Verläufe können mittels der empirischen Formeln

$$\text{LR9}: \qquad T_{95} = \Phi^{-1}\left(1 - \frac{0.05}{2(n-1)}\right)\frac{n}{p_n} \tag{5.8}$$

worin

$$\begin{aligned} p_n = \; & + \; (7.03\text{e-}1 \pm 0.68\text{e-}2) + (1.0470 \pm 0.0046) \cdot n \\ & - \; (1.203\text{e-}3 \pm 0.092\text{e-}3) \cdot n^2 + (5.89\text{e-}6 \pm 0.68\text{e-}6) \cdot n^3 \\ & - \; (7.69\text{e-}9 \pm 1.67\text{e-}9) \cdot n^4 \end{aligned}$$

und

$$\text{LR10}: \qquad T_{95} = \Phi^{-1}\left(1 - \frac{0.05}{2(n-1)}\right)\frac{n}{q_n} \tag{5.9}$$

worin

$$\begin{aligned} q_n = \; & + \; (1.79 \pm 0.13) + (9.73\text{e-}1 \pm 0.22\text{e-}1) \cdot n \\ & - \; (2.29\text{e-}3 \pm 1.22\text{e-}3) \cdot n^2 - (7.25\text{e-}5 \pm 3.07\text{e-}5) \cdot n^3 \\ & - \; (9.26\text{e-}7 \pm 3.97\text{e-}7) \cdot n^4 - (6.06\text{e-}9 \pm 2.76\text{e-}9) \cdot n^5 \\ & - \; (2.02\text{e-}11 \pm 0.98\text{e-}11) \cdot n^6 - (2.71\text{e-}14 \pm 1.38\text{e-}14) \cdot n^7 \end{aligned}$$

mit einem mittleren Fehler von 0.1758 bzw. 0.1802 approximiert werden. In Abbildung 5.8 sind die Anteile der signifikanten Bruchepochen für verschiedene Zeitreihenlängen wiedergegeben. Man erkennt, dass die modifizierten Likelihood Ratio Tests den Fehler 1. Art wie erwartet einhalten.

5.3 Güte

Die Gütefunktionen sind in Anhang C zusammengestellt, wobei diejenigen des LR9 und des LR10 Tests identisch sind. Wie der Abbildung 5.9 entnommen werden kann, ist die Power der modifizierten LR Tests für einfache Versatzinhomogenitäten erwartungsgemäss in der Mitte der Zeitreihen leicht besser als diejenige des klassischen Likelihood Ratio Tests wobei erfreulicherweise der Powerverlust an den Rändern weit weniger gross ist, als mit Blick auf die Güte des Buishand Tests an diesen Stellen befürchtet werden müsste.

5.4 Sprunghöhen und Bruchepochen

Die geschätzten Sprunghöhen und Bruchepochen der modifizierten Likelihood Ratio Tests finden sich in Anhang C und sind mit denjenigen des LR4 Tests vergleichbar, wobei die Ergebnisse der modifizierten Tests etwas besser sind als diejenigen des klassischen LR4 Tests.

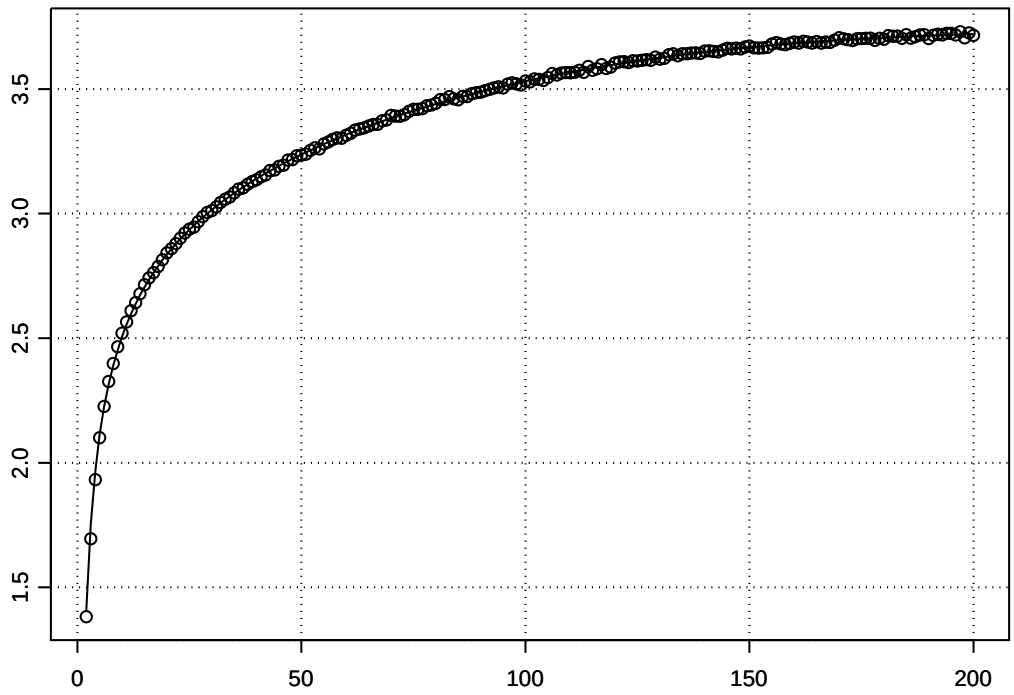


Abbildung 5.6: Verlauf der simulierten kritischen Grenzen des LR9 Tests (Kreise) mit approximierender Kurve nach Formel (5.8).

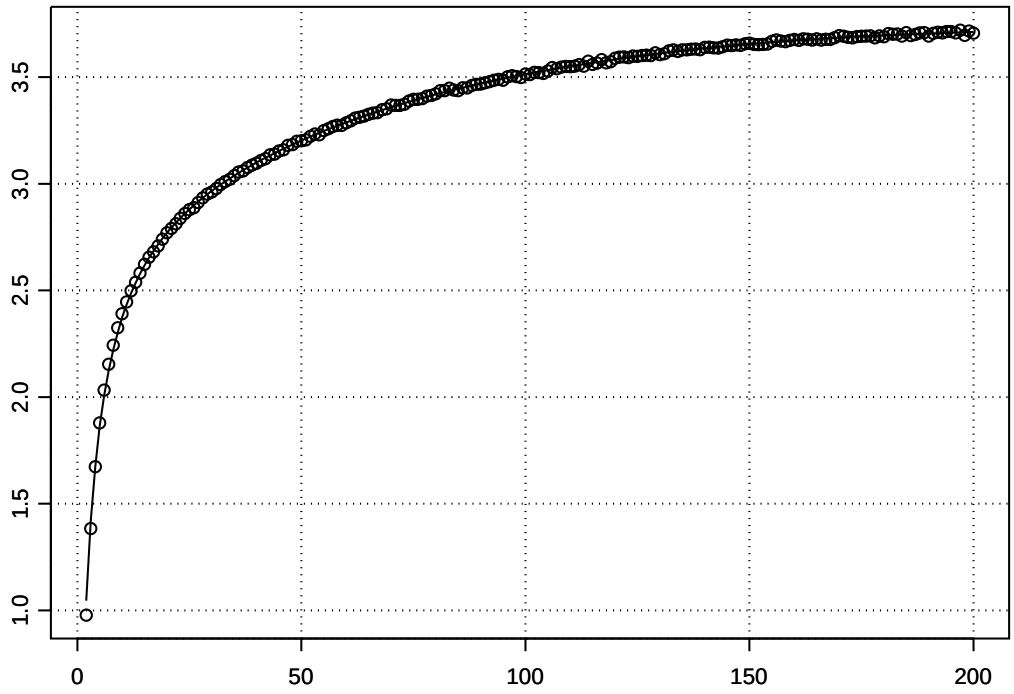


Abbildung 5.7: Verlauf der simulierten kritischen Grenzen des LR10 Tests (Kreise) mit approximierender Kurve nach Formel (5.9).

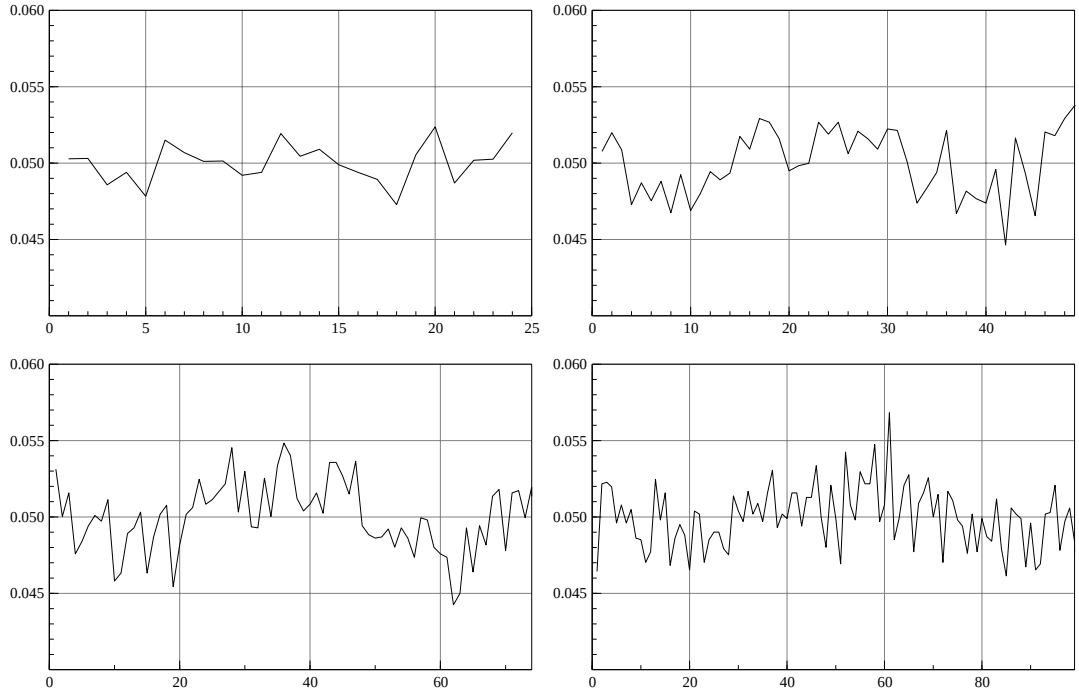


Abbildung 5.8: Lokale Realisierung des globalen Fehlers 1. Art der modifizierten Likelihood Ratio Tests für 1'000'000 Replikationen.

5.5 Diskussion

Aus den Resultaten der vorangehenden Abschnitte zeigt sich, dass mit Hilfe des Ansatzes (5.4) eine fast vollständig uniforme Verteilung der Häufigkeiten signifikanter Brüche pro Bruchepoche erreicht werden kann. Damit konnte der störendste der in Kapitel 4 diskutierten Mängel der klassischen Likelihood Ratio Tests behoben werden.

Leider gelingen die Modifikationen der Teststatistiken nur für vergleichsweise kurze Zeitreihen bis $n \leq 200$. Weitere Verbesserungen scheinen daher nötig. Allerdings halte ich die Suche nach passenderen Glättungsansätzen für die Homogenitätsprüfung klimatologischer Zeitreihen nicht für vordringlich (obwohl derartige Tests in der statistischen Qualitätskontrolle und Prozessüberwachung eingesetzt werden könnten), vielmehr scheint mir die Konstruktion eines modifizierten gleitenden MOSUM-Tests gewinnbringender (für MOSUM-Tests siehe Antoch et al. 2002), da hiermit Zeitreihen beliebiger Länge geprüft und auch mehrfache Versatzinhomogenitäten entdeckt werden könnten.

Mittels gleitenden MOSUM-Tests liessen sich zudem schwach bis mittelstark autokorrelierte Datensätze prüfen. Allerdings genügen hierzu die von Antoch et al. (1997) vorgeschlagenen Korrekturen an s_k keineswegs, da diese nur die kritischen Grenzen optimieren (welche ohnehin simuliert werden müssen), ohne an den Verteilungen der Maximums- und Bruchhäufigkeiten etwas zu ändern. Wie aus den bisherigen Resultaten hervorgeht, muss aber für jede Autokorrelationsstruktur ein eigener Glättungsansatz gefunden werden. Im Gegensatz zu den klassischen Likelihood Ratio Tests scheint dies jedoch für gleitende MOSUM-Tests realisierbar zu sein.

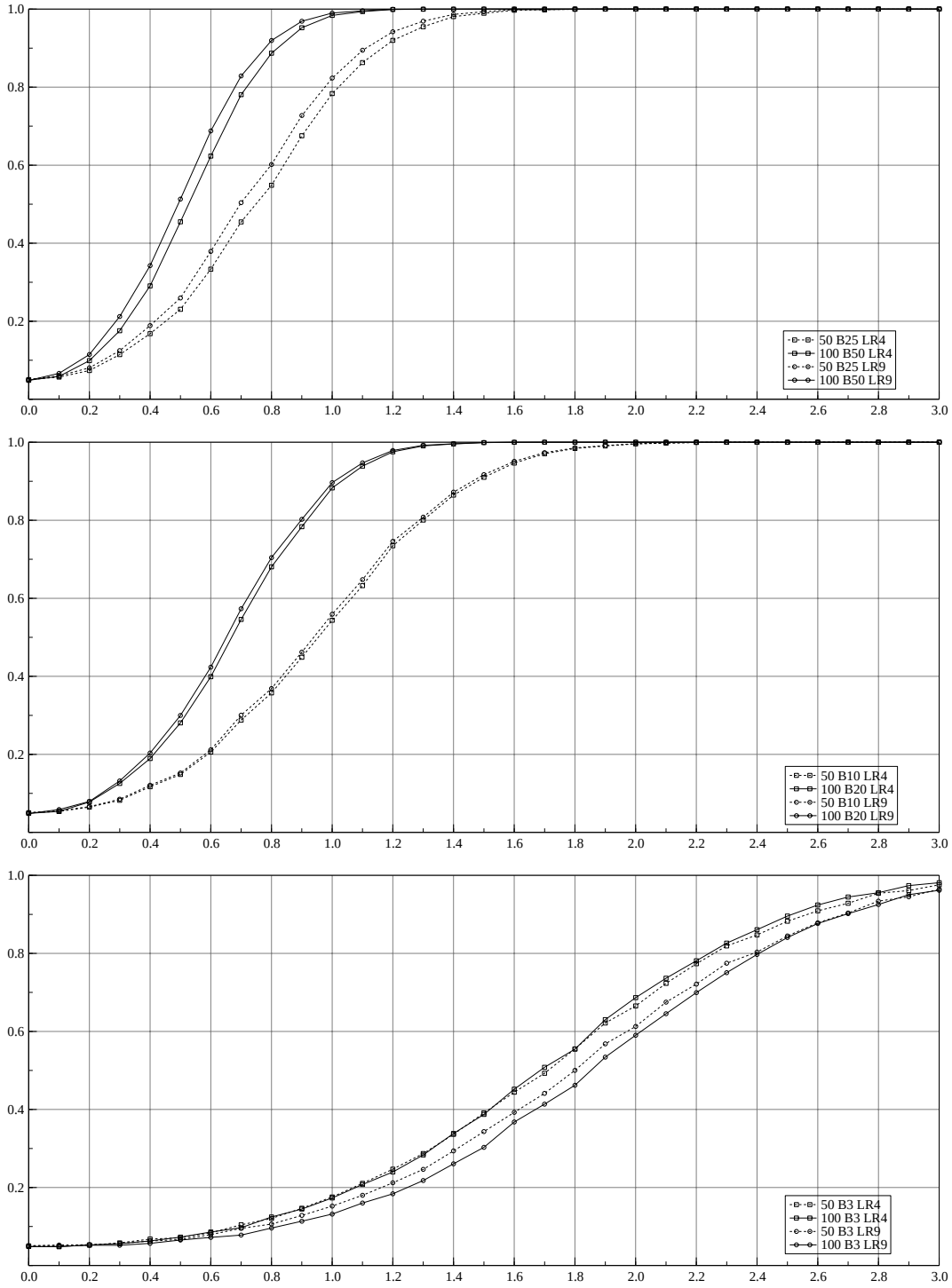


Abbildung 5.9: Vergleich der Gütefunktionen des LR9 Tests und des LR4 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 50$ und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

6. Homogenitätsprüfung

In diesem Kapitel wird gezeigt, wie die in Kapitel 4 und 5 diskutierten Homogenitätstests in den in Kapitel 2 vorgestellten univariaten strukturellen Zustandsraummodellen zur Identifizierung statistisch signifikanter Inhomogenitätsstellen eingesetzt werden können.

6.1 Schockmodell

Sprunghafte Veränderungen einzelner Modellkomponenten – sogenannte Schocks – lassen sich in strukturellen Zustandsraummodellen mittels erklärender Dummy-Variablen einfach modellieren. Nach de Jong und Penzer (1998) lautet eine effiziente Zustandsraumdarstellung eines univariaten strukturellen Schockmodells:

$$\begin{aligned} y_t &= X_t\lambda + \mathbf{Z}_t\boldsymbol{\alpha}_t + G_t\epsilon_t. \\ \boldsymbol{\alpha}_{t+1} &= \mathbf{W}_t\lambda + \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{H}_t\boldsymbol{\eta}_t. \end{aligned} \tag{6.1}$$

worin X_t und $\mathbf{W}_t$ bekannte, zeitabhängige sog. Schock Design Matrizen sind, während λ der zu schätzende unbekannte Schockeffekt darstellt. Die Auswirkung eines Schocks auf y_t wird durch die sogenannte Signatur $\mathbf{D}_t$ beschrieben, welche sich wie folgt aus den Schock Design Matrizen berechnen lässt:

$$\mathbf{D}_t = \begin{cases} \mathbf{0}, & t = 1, \ldots, k-1 \\ X_k, & t = k \\ \mathbf{Z}_t\mathbf{T}_{t-1,k+1}\mathbf{W}_k, & t = k+1, \ldots, n. \end{cases}$$

wobei $\mathbf{T}_{j,t} = \mathbf{T}_j \cdots \mathbf{T}_t$ für $j \geq t$ und $\mathbf{T}_{t-1,t} = \mathbf{I}_m$. $t = k$ bezeichnet die Bruchepoche.

Wird dem Schockmodell (6.1) ein lokal lineares Trendmodell nach Gleichung (2.6) zugrunde gelegt, besitzen die Schock Design Matrizen für einige bekannte Schocksignaturen die Form:

Ausreisser:	$X_k = 1,$	$\mathbf{W}_k = \mathbf{0}$
Versatz:	$X_k = 1,$	$\mathbf{W}_k = (1,0)'$
Knick:	$X_k = 0,$	$\mathbf{W}_k = (0,1)'$

Der unbekannte Schockeffekt λ und seine Varianz lassen sich nach de Jong und Penzer (1998) mit Hilfe des Schockkontrasts $\mathbf{s}_k$ und dessen Kovarianz $\mathbf{S}_k$ aus dem Output des Kalman Filter und Smoothers (KFS) des ungeschockten Nullmodells berechnen:

$$\hat{\lambda} = \mathbf{S}_k^{-1}\mathbf{s}_k \qquad\qquad \text{Var}(\hat{\lambda}) = \mathbf{S}_k^{-1}$$

worin

$$\mathbf{s}_k = \mathbf{X}_k'\mathbf{e}_k + \mathbf{W}_k'\mathbf{r}_k \tag{6.2}$$

$$\mathbf{S}_k = \mathbf{X}_k'\mathbf{F}_k^{-1}\mathbf{X}_k + \mathbf{Q}_k'\mathbf{N}_k\mathbf{Q}_k \tag{6.3}$$

mit $\mathbf{Q}_k = \mathbf{W}_k - \mathbf{K}_k\mathbf{X}_k$.

6.2 Einfache Versatzinhomogenitäten als Schockmodell

Eine einfache Mittelwertsverschiebung in einer unabhängig normalverteilten Folge von univariaten Zufallsvariablen y_t kann mit $\alpha_t = [\mu_t]$ wie folgt als strukturelles Schockmodell dargestellt werden:

$$\begin{aligned} y_t &= [1]\ \alpha_t + [\sigma_\epsilon]\ \varepsilon_t \\ \alpha_{t+1} &= [w_t]\ \lambda + [1]\ \alpha_t \end{aligned} \tag{6.4}$$

wobei w_t die Dummy-Darstellung

$$w_t = \begin{cases} 0 & t \neq k \\ 1 & t = k \end{cases} \tag{6.5}$$

besitzt, womit für die Schocksignatur D_t die gewünschte einstufige Treppenform resultiert. Die Abbildungen 6.1 und 6.2 zeigen für die im Anhang B wiedergegebene Testreihe Z80, welche bereits in den Kapiteln 4 und 5 zur Visualisierung der Testgrössen der dort diskutierten Homogenitätstests verwendet wurde, die Verläufe der beiden geschätzten Komponenten sowie die diagnostischen Grafiken des ungeschockten Nullmodells.

6.3 Likelihood Ratio Testgrössen

Nach de Jong und Penzer (1998) eignet sich zur Prüfung der Nullhypothese $\lambda = 0$ die Testgrösse

$$\rho_k^2 = \hat{\lambda}_k \text{Var}(\hat{\lambda}_k)^{-1}\hat{\lambda}_k = \mathbf{s}_k'\mathbf{S}_k^{-1}\mathbf{s}_k \tag{6.6}$$

welche dem Quadrat der 'usual regression t statistic' entsprechen und approximativ χ_p^2-verteilt sein soll, wobei p der Rang von $\mathbf{S}_k$ ist. Für das strukturelle Schockmodell (6.4) erhalten wir:

$$\rho_k^2 = r_k N_k^{-1} r_k \tag{6.7}$$

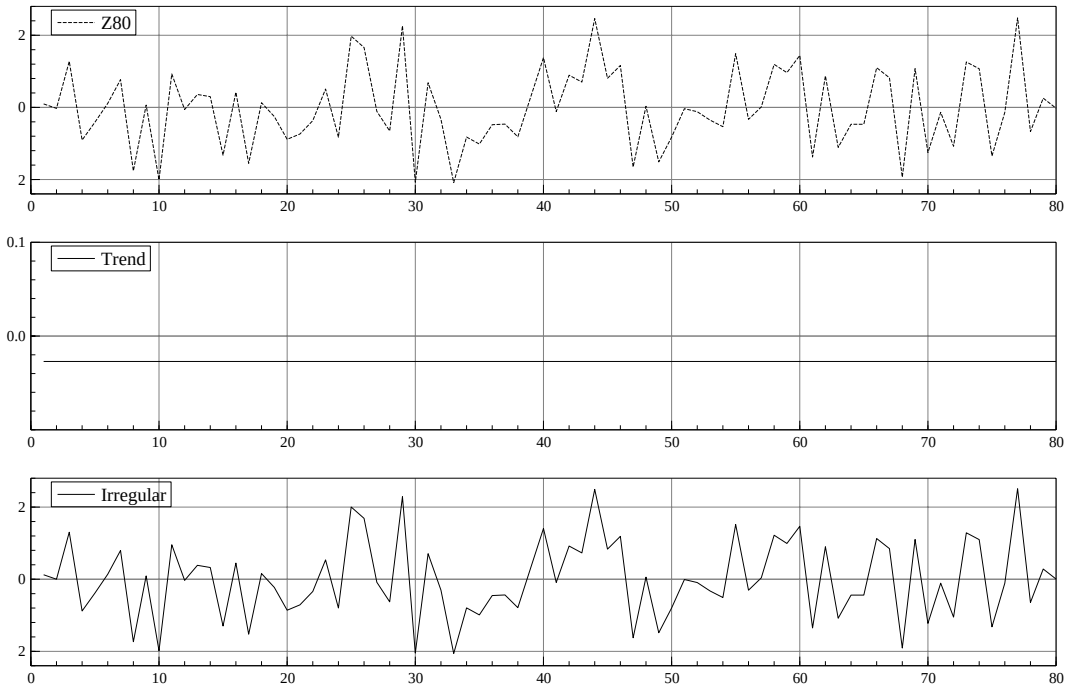


Abbildung 6.1: Dekomposition der Testreihe Z80 in einen lokalen Trend und einen Irregular.

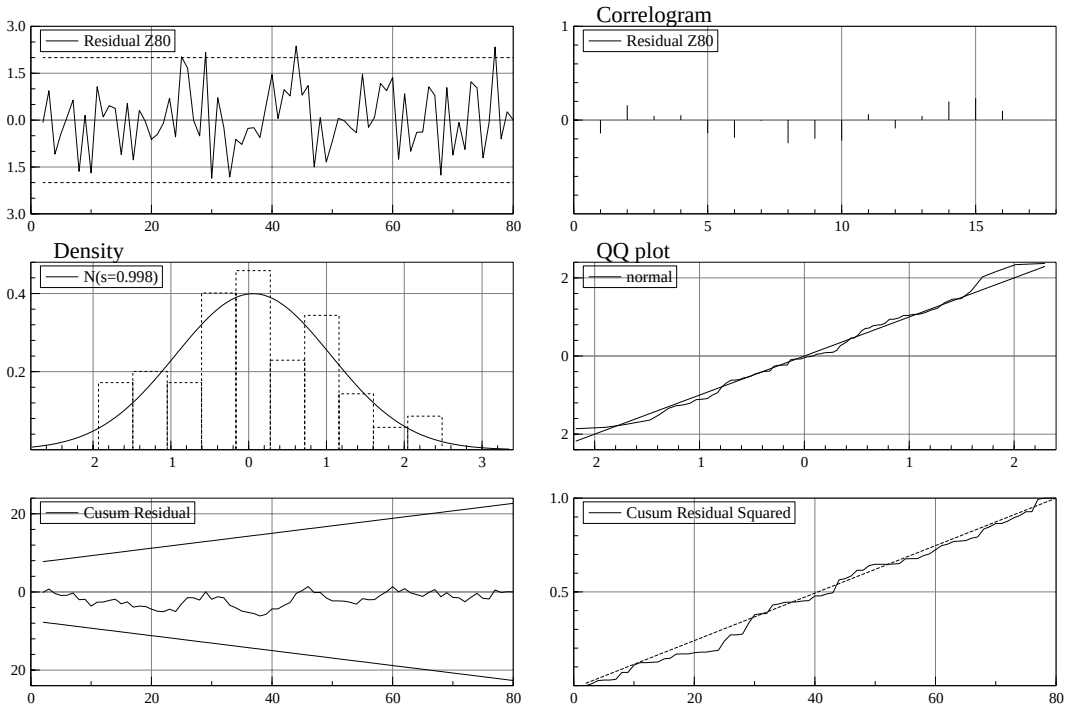


Abbildung 6.2: Diagnostik zu den Innovationen des Modells in Abbildung 6.1. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

In Abbildung 6.3 sind für die Testreihe Z80 die Verläufe von r_k und N_k wiedergegeben. Hieraus wird ersichtlich, dass

$$N_k = \frac{k(n-k)}{n(n-1)} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2. \tag{6.8}$$

Nach Durbin und Koopman (2001) ist

$$r_k = \sum_{i=k+1}^{n} (Y_i - \bar{Y}_n). \tag{6.9}$$

Da

$$\sum_{i=1}^{n} (Y_i - \bar{Y}_n) = 0$$

gilt:

$$r_k = \sum_{i=k+1}^{n} (Y_i - \bar{Y}_n) = -\sum_{i=1}^{k} (Y_i - \bar{Y}_n). \tag{6.10}$$

Somit ist:

$$\boxed{\rho_k^2 = r_k N_k^{-1} r_k = \left(\frac{1}{\hat{\sigma}_n} \left| \sqrt{\frac{n}{k(n-k)}} \sum_{i=1}^{k} \left(Y_i - \bar{Y}_n \right) \right| \right)^2} \tag{6.11}$$

worin

$$\hat{\sigma}_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2}.$$

Somit entspricht ρ_k^2 der Testgrösse des Alexandersson Tests sowie $\rho_k = \sqrt{\rho_k^2}$ der Testgrösse des LR7 Tests.

6.4 Modifizierte Likelihood Ratio Testgrössen

Die modifizierte LR7 Testgrösse LR10 ist gemäss Kapitel 5 definiert als:

$$T_k = \frac{1}{\hat{\sigma}_n} \left[\sqrt{\frac{n}{k(n-k)}} + c_n \sqrt{\frac{k(n-k)}{n}} + d_n \right] \sum_{i=1}^{k} (Y_i - \bar{Y}_n)$$

mit

$$\begin{aligned} c_n &= 0.0556 - 0.0361 \log\log n \\ d_n &= -0.1031 + 0.0319 \log n \\ \hat{\sigma}_n &= \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2}. \end{aligned}$$

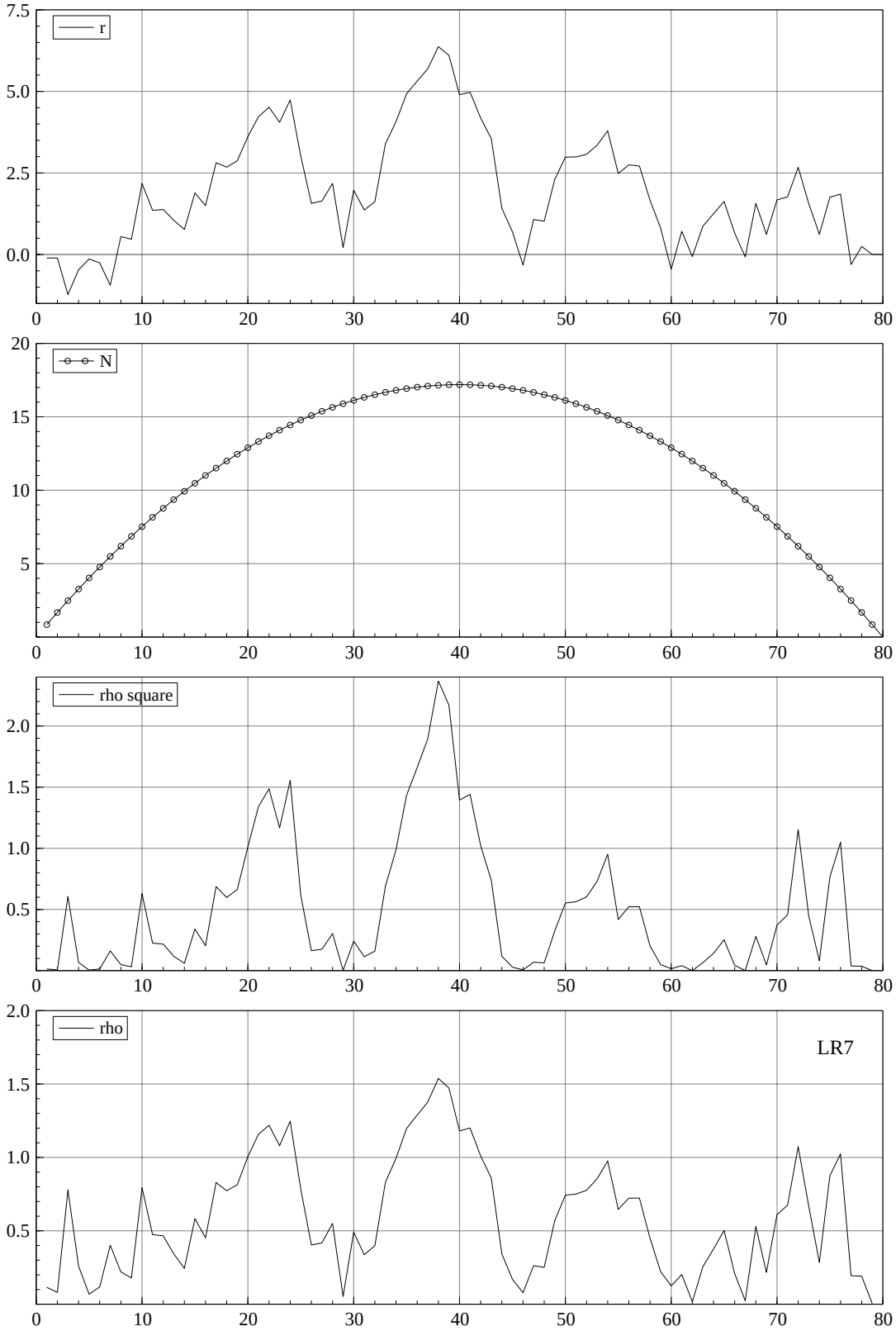


Abbildung 6.3: Verläufe zweier Grössen des KFS Outputs der Testreihe Z80 (obere zwei Grafiken) sowie der darauf berechneten Testgrössen (untere zwei Grafiken).

Somit lautet die **LR10 Teststatistik in rekursiver Form**:

$$\boxed{T = \max_{1 \le k < n} \{|R_k \; r_k|\}} \tag{6.12}$$

mit

$$R_k = \frac{1}{\hat{\sigma}_n} \left[\sqrt{\frac{n}{k(n-k)}} + c_n \sqrt{\frac{k(n-k)}{n}} + d_n \right]. \tag{6.13}$$

Abbildung 6.4 zeigt für die Testreihe Z80 die beiden Verläufe von R_k und der LR10 Testgrösse T_k.

Schliesslich sei darauf hingewiesen, dass bis auf den LR1 und den LR2 Test alle in dieser Arbeit diskutierten Homogenitätstests – insbesondere auch der Buishand Test – in der rekursiven Form (6.12) dargestellt werden können.

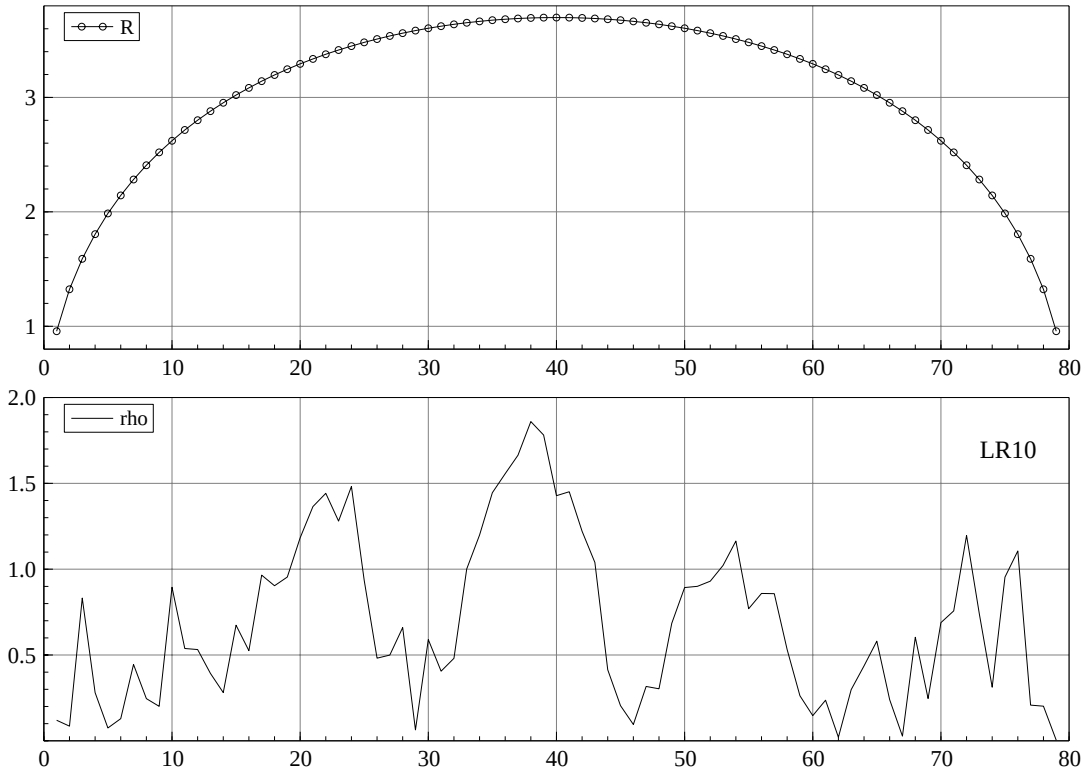


Abbildung 6.4: Verlauf der nach Formel (6.13) berechneten Gewichte R_k sowie der rekursiven LR10 Testgrösse für die Testreihe Z80.

6.5 Diskussion

Aus den Resultaten der vorangehenden Abschnitte zeigt sich, dass bis auf den LR1 und den LR2 Test alle in den Kapiteln 4 und 5 diskutierten Homogenitätstests mit Hilfe des Kalman Filter und Smoothers des Nullmodells berechnet werden können. Dies eröffnet der Diagnostik struktureller Zustandsraummodelle erstmals die Möglichkeit, sprunghafte Veränderungen der Trendkomponente quantitativ zu testen. Zwar wurde die Analogie zwischen den klassischen und den rekursiven Testgrössen nur für ein der klassischen Testsituation entsprechendes Zustandsraummodell gezeigt, doch finden sich in Steiner (2004) Simulationsstudien, denen entnommen werden kann, dass in komplexeren Zustandsraummodellen mit lokal linearen Trendkomponenten und zusätzlichen saisonalen Komponenten dieselben Teststatistiken und Gütefunktionen resultieren wie für das vorliegende local level Modell.

Allerdings muss daran erinnert werden, dass die Verteilung der Testgrösse ρ_k^2 für andere Formulierungen der Schock Design Matrizen ganz anders ausfällt. Wie Steiner (2004) zeigte, sind die Testgrössen für einen Ausreisser approximativ χ_1^2-verteilt (mit flachen Bruchhäufigkeitsverteilungen). Somit müssen die Verteilungen der Teststatistik ρ_k^2 für andere Signaturen, beispielsweise für einen Trendknick oder einen strukturellen Bruch in der Saisonfigur, fallweise bestimmt werden.

Auch gilt es zu bedenken, dass eine signifikante LR7 oder LR10 Teststatistik nur bei korrekt spezifiziertem Modell und korrekt geschätzten Hyperparametern als Hinweis auf eine signifikante Versatzinhomogenität an dieser Stelle interpretiert werden sollte: Wird beispielsweise σ_η^2 zu klein geschätzt – etwa weil in der Modelldefinition eine Komponente unberücksichtigt blieb oder weil die Maximum Likelihood Schätzung der Hyperparameter nicht wunschgemäss konvergierte – so wissen wir aus der Praxis, dass gewöhnlich zahlreiche "level shifts" auftreten, welche aber in der Regel auf das Oversmoothing der Trendkomponente zurückzuführen sind und nicht auf reale Strukturbrüche in der Zeitreihe.

Schliesslich sei darauf hingewiesen, dass die rekursiven Testgrössen ρ_k eng mit den CUSUM Testgrössen

$$C_k = \sum_{i=1}^{k} (Y_i - \bar{Y}_n). \tag{6.14}$$

verwandt sind, welche in den diagnostischen Plots struktureller Zustandsraummodelle gewöhnlich dargestellt werden (Koopman et al. 2000). Wie aus Gleichung (6.10) folgt und der Abbildung 6.2 entnommen werden kann, gilt $C_k = -r_k$. C_k ist somit bis auf die fehlende Skalierung identisch mit der Buishand Testgrösse. Im CUSUM Plot könnten deshalb auf den Höhen

$$T = \pm \frac{Q_{95}}{n} \sum_{i=1}^{n} (Y_i - \bar{Y}_n). \tag{6.15}$$

horizontale Geraden gezeichnet werden. Kreuzt C_k eine dieser beiden Geraden, so befindet sich an dieser Stelle eine Versatzinhomogenität. Dies steht im Widerspruch zu den wie in Abbildung 3.4 üblicherweise gezeichneten, trichterförmigen Geraden, welche auf Brown, Durbin und Evans (1975) zurückgehen und demselben Zweck dienen sollen: Überschreitet C_k diese Geraden, so ist an dieser Stelle der Mittelwert der Zeitreihe signifikant verschieden vom Anfangswert.

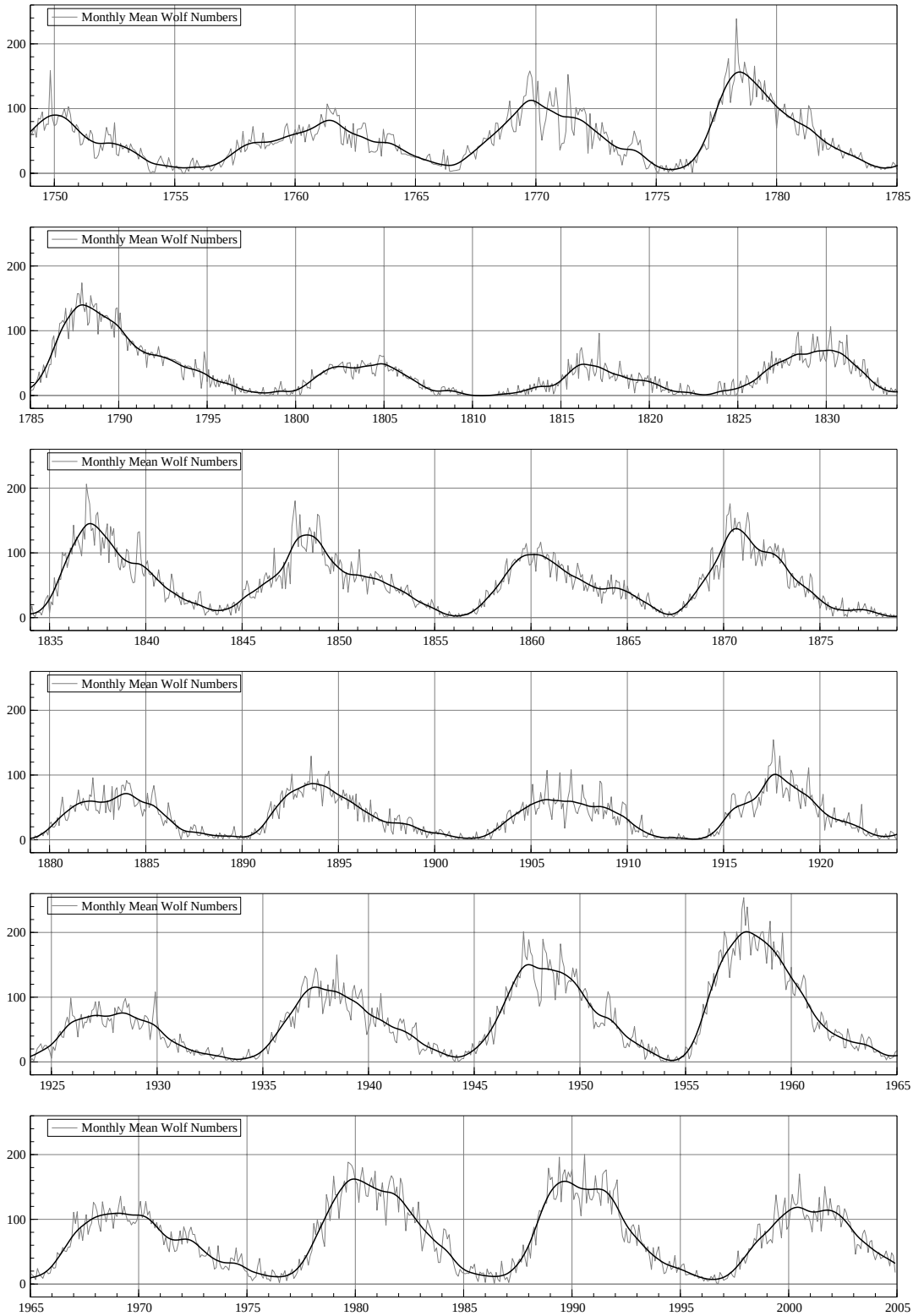


Abbildung 6.5: Verlauf der beobachteten Monatsmittel der Wolfschen Sonnenfleckenrelativzahl R_I von 1749 bis 2004 und der stochastischen Trendkomponente des Zustandsraummodells (3.6).

7. Homogenität der Wolfschen Reihe

Nachdem in Kapitel 3 ein strukturelles Zustandsraummodell zur Darstellung des Verlaufs der Sonnenfleckenaktivität und in Kapitel 6 valide Homogenitätstests entwickelt wurden, wenden wir uns in diesem Kapitel der Homogenitätsprüfung der Wolfschen Reihe zu, wobei wir uns hauptsächlich auf den Zeitraum zwischen 1945 und 1995 konzentrieren, für welchen geeignete Vergleichsreihen vorhanden sind.

7.1 Vergleichsreihen

Zur Homogenitätsprüfung der Reihe der Zürcher Sonnenfleckenrelativzahlen R_Z stehen nur wenige Vergleichsreihen zur Verfügung. Die einzige "echte" Vergleichsreihe ist diejenige der amerikanischen Relativzahlen R_A welche seit Dezember 1944 aus den Beobachtungen mehrheitlich nordamerikanischer Observatorien und Amateurastronomen berechnet werden (Shapley 1946). Wie bereits in Kapitel 1 erwähnt, wurde die Eidgenössische Sternwarte 1980 geschlossen. Die Berechnung von Standardrelativzahlen sowie das im Auftrag der Internationalen Astronomischen Union (IAU) geführte "World Data Center for Sunspot Numbers" wurden an das neugegründete "Sunspot Index Data Center" (SIDC) in Bruxelles abgetreten. Als Standardbeobachter wurde zwecks Bewahrung der Wolfschen Skala die ehemalige Aussenstation der Eidgenössischen Sternwarte in Locarno gewählt. Die neue Reihe erhielt die offizielle Bezeichnung internationale Sonnenfleckenrelativzahl R_I (Berghmans et al. 2002). Weniger bekannt ist, dass in Zürich die täglichen Zählungen am Fraunhoferschen Normalrefraktor sowie das tägliche Zeichnen der Sonnenphotosphäre am Zeiss-Coudé Refraktor in Zürich von H.U. Keller, dem ehemaligen Observator der Eidgenössische Sternwarte, im Auftrag des Bundesamtes für Übermittlungstruppen (BAUEM) unterbruchslos weiterführt werden konnte. Erst 1995 mussten die Beobachtungsinstrumente im Zuge der Gesamtrenovation des Sternwartengebäudes von der Dachterrasse und aus der Kuppel entfernt werden (Friedli et al. 1998). Die möglichst tägliche Bestimmung der Wolfschen Sonnenfleckenrelativzahl am Fraunhoferschen Normalrefraktor wird seither von mir vorgenommen. Die von H.U. Keller in den Jahren 1981 bis 1995 bestimmten Zürcher Sonnenfleckenrelativzahlen R_Z können denjenigen von Waldmeier und Zelenka unterbruchslos hinzugefügt werden, so dass hiermit die Homogenität der R_I kontrolliert werden kann. Der Verlauf der offiziellen Reihe der monatlichen Sonnenfleckenrelativzahlen R_I ist in Abbildung 6.5 wiedergegeben.

Eine wichtige Ergänzung dieser drei Reihen R_Z, R_A und R_I stellen die Zürcher Gruppenzahlen g_Z dar, auf welchen die Zürcher Sonnenfleckenrelativzahl R_Z basiert. Diese wurden seit 1945 von Waldmeier in den jährlichen Sonnenaktivitätsberichten veröffentlicht und für diese Arbeit elektronisch erfasst. Die von H.U. Keller ab 1991 nach demselben System berechneten Gruppen- und Relativzahlen wurden von Keller und Friedli (1995) veröffentlicht und liegen ebenfalls in elektronischer Form vor. Eng verwandt mit dieser Reihe ist die von Hoyt und Schatten (1998) vorgelegte

Reihe sogenannter gruppierter Sonnenfleckenrelativzahlen R_g welche eine nur auf den Gruppenzahlen basierende Neurekonstruktion der Sonnenfleckenaktivität insbesondere während dem 17. und 18. Jahrhundert darstellt.

7.2 Die kurzfristige Homogenität zwischen 1945 und 1995

In Abbildung 7.1 sind die Monatsmittel von R_Z sowie R_A und in Abbildung 7.2 die korrespondierenden Tageswerte dargestellt. Neben den beobachteten Werten und den Differenzen der beiden Beobachtungsreihen sind auch die mittleren Trendverläufe wiedergegeben. Diese wurden - wie für alle in diesem Kapitel gezeigen Zeitreihen – mit einem in Kapitel 2.2.1 diskutierten strukturellen Zustandsraummodell der Form

$$\begin{aligned} y_t &= [1 \quad 0]\ \boldsymbol{\alpha}_t + \ [\sigma_\epsilon]\ \varepsilon_t \\ \boldsymbol{\alpha}_{t+1} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} 0 & 0 \\ 0 & \sigma_\zeta \end{bmatrix} \begin{bmatrix} 0 \\ \zeta_t \end{bmatrix}. \end{aligned} \tag{7.1}$$

berechnet, wobei $\boldsymbol{\alpha}_t = [\mu_t, \nu_t]'$ und y_t fallweise Tageswerte, Monatsmittel oder deren Differenzen darstellen. Alle Berechnungen erfolgten mit Programmen, welche mit der C-Library SsfPack 2.3 von Koopman et al. (1999) in der objekt-orientierten Matrizensprache Ox von Doornik (1998) geschrieben wurden und welche auch eine Homogeniätsprüfung nach dem Alexandersson Test, bzw. dem LR10 Test erlauben.

Bereits der erste flüchtige Blick auf die abgebildeten Monatsmittel und Tageswerte zeigt, dass gleich zu Beginn der Zeitreihen massive systematische Differenzen zwischen R_Z und R_A auftraten. Shapley (1949) und Gleissberg (1952) führten die doch erhebliche Überschätzung der Sonnenaktivität durch R_A auf die mangelnde Erfahrung der nordamerikanischen Beobachter und ein zu kurzes Kalibrierungsintervall mit Zürich zurück. Das Auswerteverfahren, welches ursprünglich eine ungewichtete Mittelung aller auf Wolfsche Skala kalibrierten Einzelwerte umfasste, wurde dahingehend geändert, dass Beobachter mit mehr Erfahrung einen bestimmenden Einfluss auf den langfristigen Verlauf der berechneten Reihe erhielten. Zusätzlich wurde die Bestimmung des k-Faktors aus N korrespondierenden Relativzahlbestimmungen R_i des Beobachters i mit Zürich optimiert, welcher seither nach der interessanten Formel

$$\log k_i = \frac{1}{N} \left(\sum_{j=1}^{N} \log R_{Z,j} - \sum_{j=1}^{N} \log R_{ij} \right) \tag{7.2}$$

berechnet wird. Dies entspricht einer von Wold (1940) beschriebenen Schätzung des Intercepts in einem linearen error-in-variables Modell mit bekannter Steigung nach vorangehender Varianzstabilisierung mittels zweiseitiger Logarithmustransformation. Dieses von Shapley (1949) vorgeschlagene innovative Verfahren wurde 1951 erfolgreich neu kalibriert. Allerdings gab es während der Maximumsphasen der darauffolgenden Zyklen immer wieder mehr oder weniger grosse Abweichungen, am stärksten während des extrem hohen Zyklus Nr. 19, welcher 1954 einsetzte.

In beiden Darstellungen fällt auf, dass die während den Maximumsphasen auftretenden Abweichungen ab 1980 offenbar ein anderes Vorzeichen besitzen. *Eine Homogenitätsprüfung anhand des Alexanderson Tests ergab, dass im Mai 1980 eine signifikante Versatzinhomogenität aufgetreten ist.* Wie der Abbildung 7.3 entnommen werden kann, liegen die Tageswerte und Monatsmittel der Zürcher

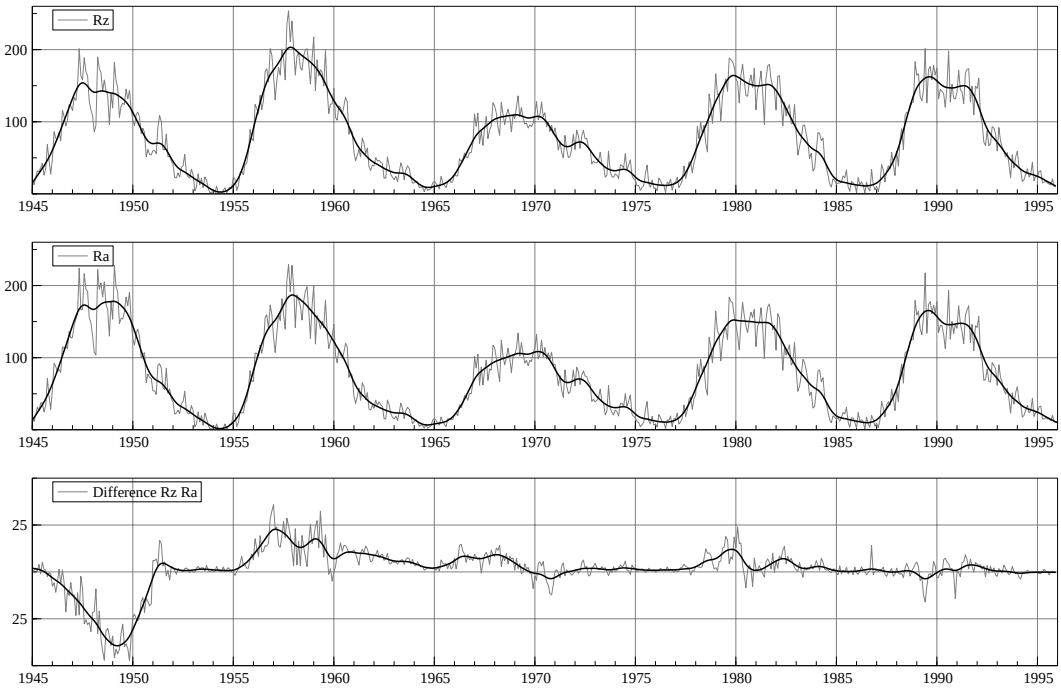


Abbildung 7.1: Verlauf der beobachteten Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der amerikanischen Sonnenfleckenrelativzahl R_A sowie ihrer Differenzen von 1945 bis 1995 und der stochastischen Trendkomponenten der Zustandsraummodelle (7.1).

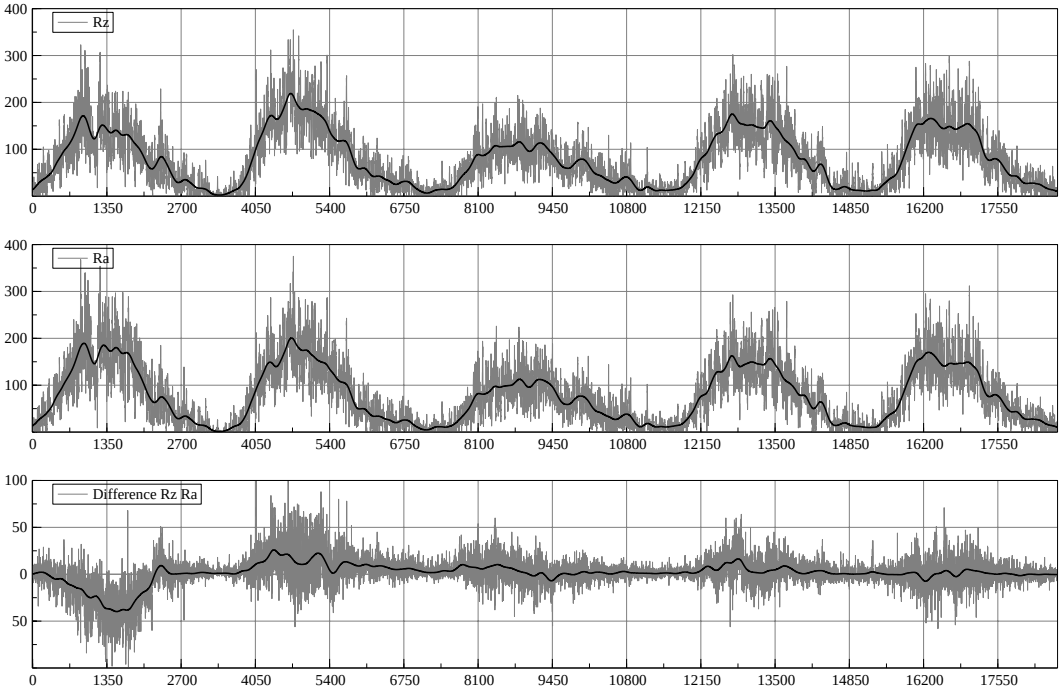


Abbildung 7.2: Verlauf der beobachteten Tageswerte der Zürcher Sonnenfleckenrelativzahl R_Z und der amerikanischen Sonnenfleckenrelativzahl R_A sowie ihrer Differenzen von 1945 bis 1995 und der stochastischen Trendkomponenten der Zustandsraummodelle (7.1).

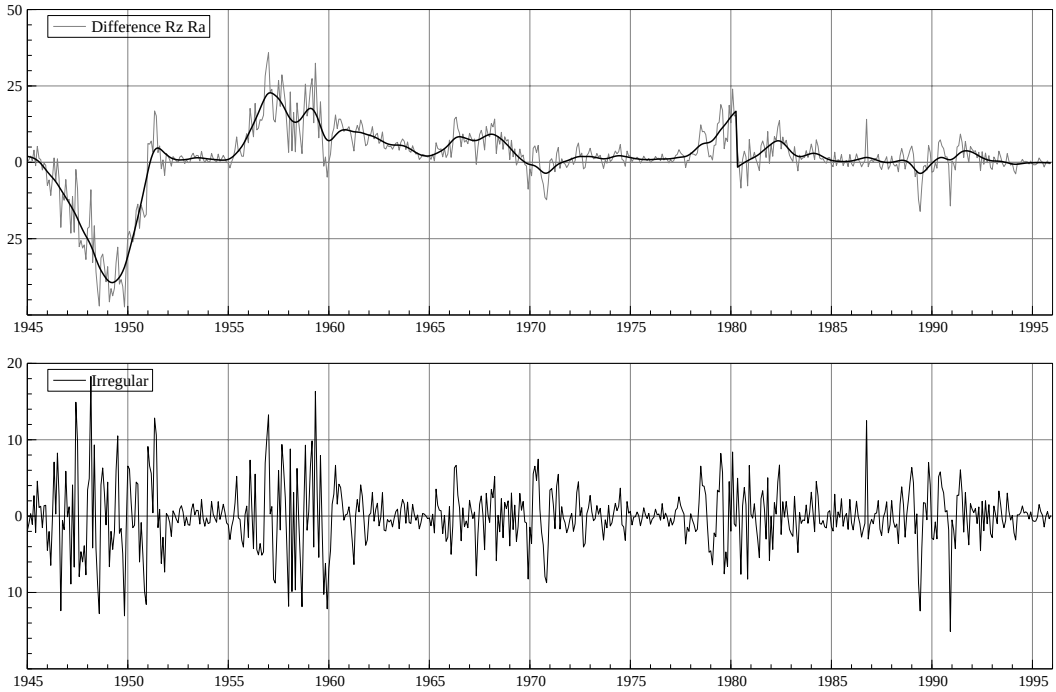


Abbildung 7.3: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der amerikanischen Sonnenfleckenrelativzahl R_A von 1945 bis 1995 mit signifikanter Versatzinhomogenität um -19 ± 3.8 im Mai 1980.

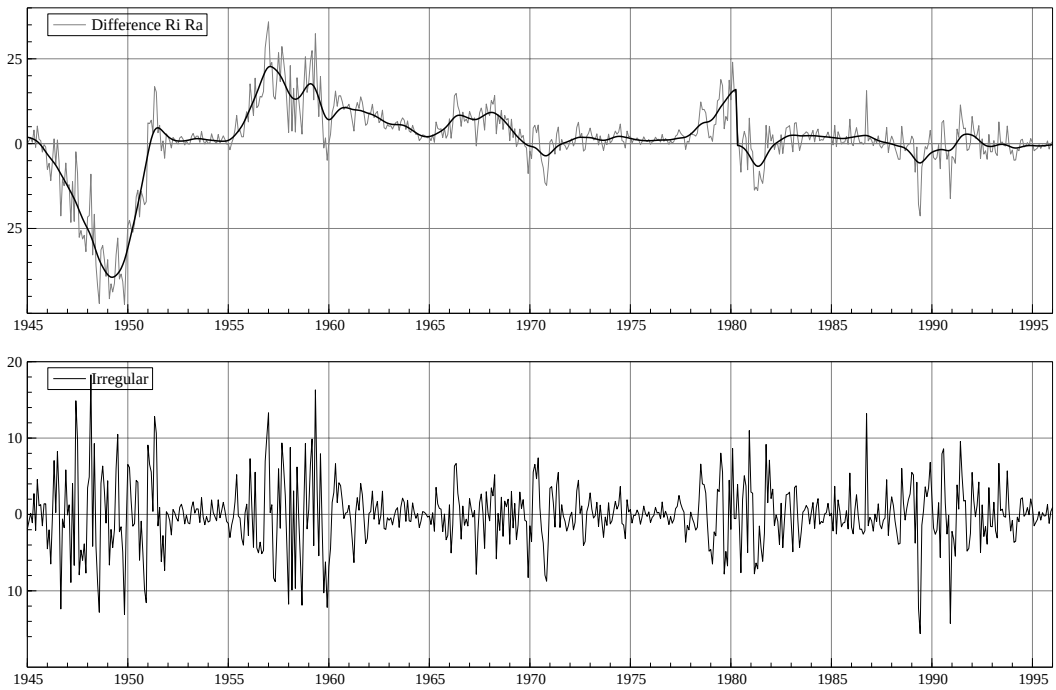


Abbildung 7.4: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der internationalen Sonnenfleckenrelativzahl R_I von 1945 bis 1995 mit signifikanter Versatzinhomogenität um -17 ± 3.9 im Mai 1980.

Sonnenfleckenrelativzahlen ab Juni 1980 signifikant tiefer als diejenigen davor. Dieser Bruch ist offenbar auf den Wechsel des Standardbeobachters zurückzuführen und liegt genau in der Mitte zwischen dem Ende der Waldmeierschen Beobachtungen auf Ende September 1979 und dem Beginn der Kellerschen Standardbeobachtungen im Januar 1981 (dazwischen war A. Zelenka für 14 Monate offiziell Hüter der Wolfschen Einheit). Interessanterweise zeigt sich dieser Bruch sowohl beim Vergleich von R_A mit R_Z wie beim Vergleich von R_A mit R_I, wie der Abbildung 7.4 entnommen werden kann. Der Bruch könnte daher weniger in der Einführung der beiden neuen Standardbeobachter als im Wegfall des alten - Waldmeier - begründet liegen.

Zwar wurde in den von der Eidgenössischen Sternwarte veröffentlichten Jahresberichten zur Sonnenaktivität jeweils eine eindrucksvolle Liste mitarbeitender Observatorien und Amateurastronomen veröffentlicht, doch darf nicht vergessen werden, dass üblicherweise 90 - 95% der Tage durch Zürcher Beobachtungen gedeckt waren und die auswärtigen Stationen nur für die verbleibenden Lückentage Verwendung fanden. Die Zürcher Beobachtungen wurden jedoch von Waldmeier dominiert: nicht nur lieferte er jeweils mindestens die Hälfte der definitiven Tageswerte, sondern bestimmte mit seinen Beobachtungen auch die Reduktionsfaktoren seiner Zücher Mitarbeiter. Eine allmähliche Änderung seiner Zählweise hätte sich also direkt auf die ganze Reihe übertragen. Dass eine solche allmähliche Änderung nicht ganz ausgeschlossen werden kann, hat gute Gründe: zum Einen war Waldmeier vor seiner Ernennung zum Direktor der Eidgenössischen Sternwarte jahrelang der Leiter der Aussenstation in Arosa und hatte kaum Beobachtungserfahrung am Wolfschen Normalrefraktor in Zürich, zum Anderen hat Waldmeiers Vorgänger William Brunner nach seiner Emeritierung nicht mehr weiterbeobachtet und auch dessen langjähriger Assistent hat die Eidgenössische Sternwarte schon ein Jahr nach Waldmeiers Amtsantritt verlassen. Das neue Beobachtungsteam in Zürich war also relativ unerfahren und musste zudem noch während der Minimumsphase beginnen. Erschwerend kam hinzu, dass die beiden nachfolgenden Zyklen die intensivsten je direkt beobachteten waren, mit bis zu 100 Einzelgruppen pro Sonnenrotation in den Maximumsphasen. Waldmeier hat denn auch selber befürchtet, sein als konstant angenommener Skalenfaktor könnte variieren. In der Folge versuchte er die Relativzahlen anhand objektiv messbarer Kenngrössen zu kalibrieren. Hierzu untersuchte er die Zusammenhänge zwischen der Zürcher Sonnenfleckenrelativzahl und den Greenwicher Fleckenflächenmessungen, zwischen der Zürcher Sonnenfleckenrelativzahl und den Messungen des solaren Radioflusses bei 10.7 cm Wellenlänge, sowie zwischen der Zürcher Gruppenzahl und der Zürcher Sonnenfleckenrelativzahl (Waldmeier 1968, 1971, 1978), ohne jedoch irgendwelche Inhomogenitäten entdecken zu können.

Da sowohl R_A wie R_I bis in die Gegenwart fortgeführt wurden, sei an dieser Stelle auch ein kurzer Blick auf die neueste Aktivitätsentwicklung geworfen. Wie der Abbildung 7.5 entnommen werden kann, ist der 23. Sonnenaktivitätszyklus deutlich tiefer ausgefallen als die beiden vorangehenden. Damit wird übrigens die seit 1855 gültige Regel, wonach die Maximumshöhe geradzahliger Zyklen geringer ausfällt als die Maximumshöhe ungeradzahliger Zyklen, erstmals verletzt. Worauf dies zurückzuführen ist, ist unbekannt. Aus der Abbildung 7.5 ist jedoch auch ersichtlich, dass ungefähr ab Mitte 1999 ein deutlicher Abfall der internationalen Sonnenfleckenrelativzahlen R_I gegenüber den amerikanischen Sonnenfleckenrelativzahlen R_A einsetzt, wodurch sogar die Maximumsepoche des 23. Zyklus zwischen den beiden Reihen um fast zwei Jahre differiert. Eine Homogenitätsprüfung mittels dem Alexandersson Test ergibt einen schwach signifikanten Bruch im Juni 1999 mit einer Sprunghöhe von -8 ± 3.9. Da auch andre Vergleichsreihen - beispielsweise das deutsche SONNE-Netz - dasselbe Verhalten gegenüber R_I zeigen, muss die Ursache dieser Inhomogenitäten in R_I zu finden sein.

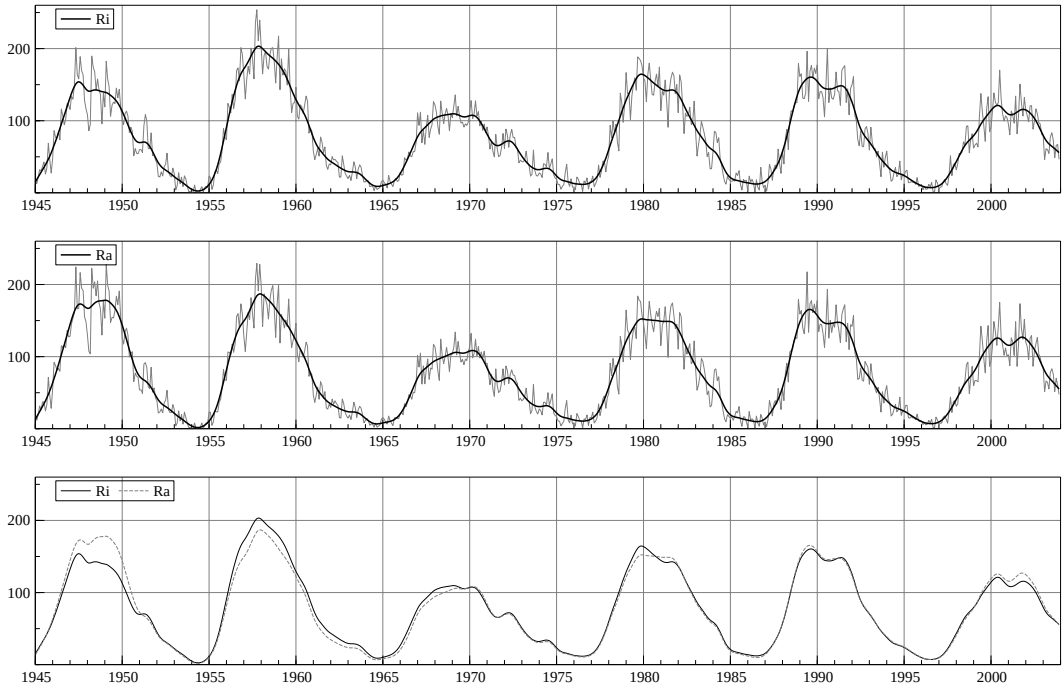


Abbildung 7.5: Verlauf der beobachteten Monatsmittel der internationalen Sonnenfleckenrelativzahl R_I und der amerikanischen Sonnenfleckenrelativzahl R_A von 1945 bis 2003.

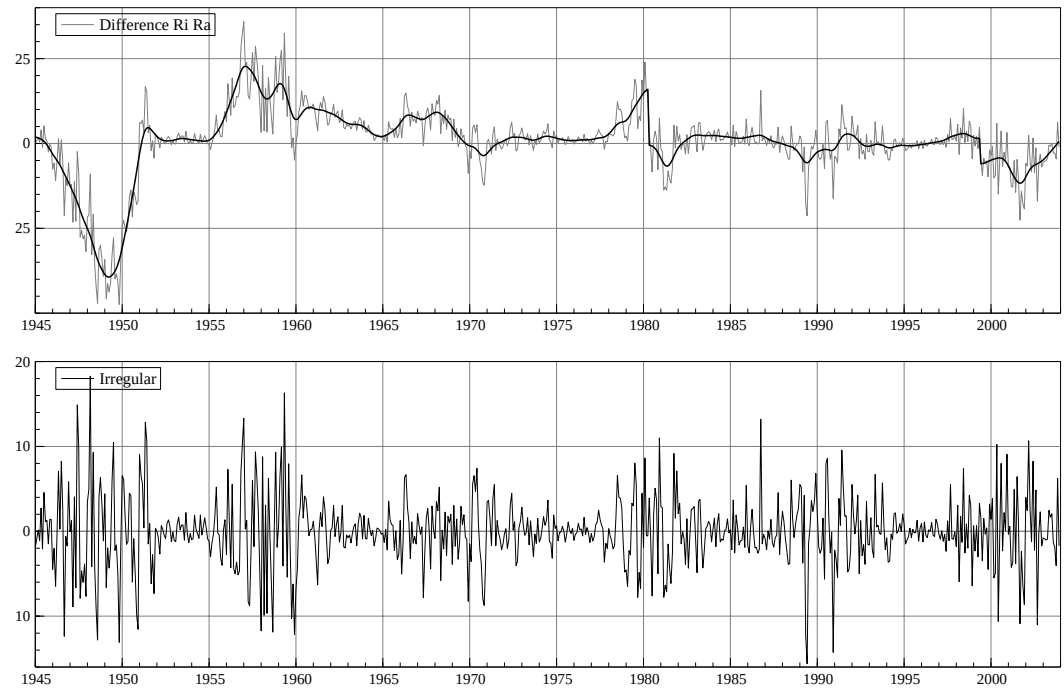


Abbildung 7.6: Verlauf der beobachteten Differenzen der Monatsmittel der internationalen Sonnenfleckenrelativzahl R_I und der amerikanischen Sonnenfleckenrelativzahl R_A von 1945 bis 2003 mit signifikanten Versatzinhomogenitäten im Mai 1980 (-17 ± 3.9) und im Juni 1999 (-8 ± 3.9).

Schliesslich sei darauf hingewiesen, dass die beiden Reihen R_Z und R_A im betrachteten Zeitraum vollständig sind und durch mehr oder weniger disjunkte Beobachtergruppen bestimmt wurden. Die in Abbildung 7.7c dargestellte trendbereinigte Differenzenreihe kann daher durchaus als Mass für die Aufassungs- und Kalibrierungsfehler betrachtet werden, welche den Tageswerten der Wolfschen Reihe anhaften. Wie aus Abbildung 7.7e ersichtlich, beträgt der absolute Fehler im Mittel etwa ± 10, der relative Fehler also bestenfalls 10%. Auf Tageswerten basierende Auswertungen - etwa die Suche nach chaotischem Verhalten - müssen also mit erheblichem Messrauschen rechnen, welches zudem in den weniger vollständig und naturgemäss gröber kalibrierten Bereichen der Wolfschen Reihe im 18. und 19. Jahrhundert noch merklich grösser sein dürfte.

In Abbildung 7.8 sind die Zürcher Relativzahlen R_Z und die korrespondierenden Zürcher Gruppenzahlen g_Z dargestellt. Aus diesen beiden Reihen wurden nach der Formel

$$f_Z = \frac{5}{3} R_z - 10\, g_Z \tag{7.3}$$

Zürcher Fleckenzahlen f_Z berechnet, welche ein Standardbeobachter mit konstantem k-Faktor 0.6 hätte beobachten können. In Abbildung 7.8 ist zusätzlich das Verhältnis f_Z/g_Z wiedergegeben.

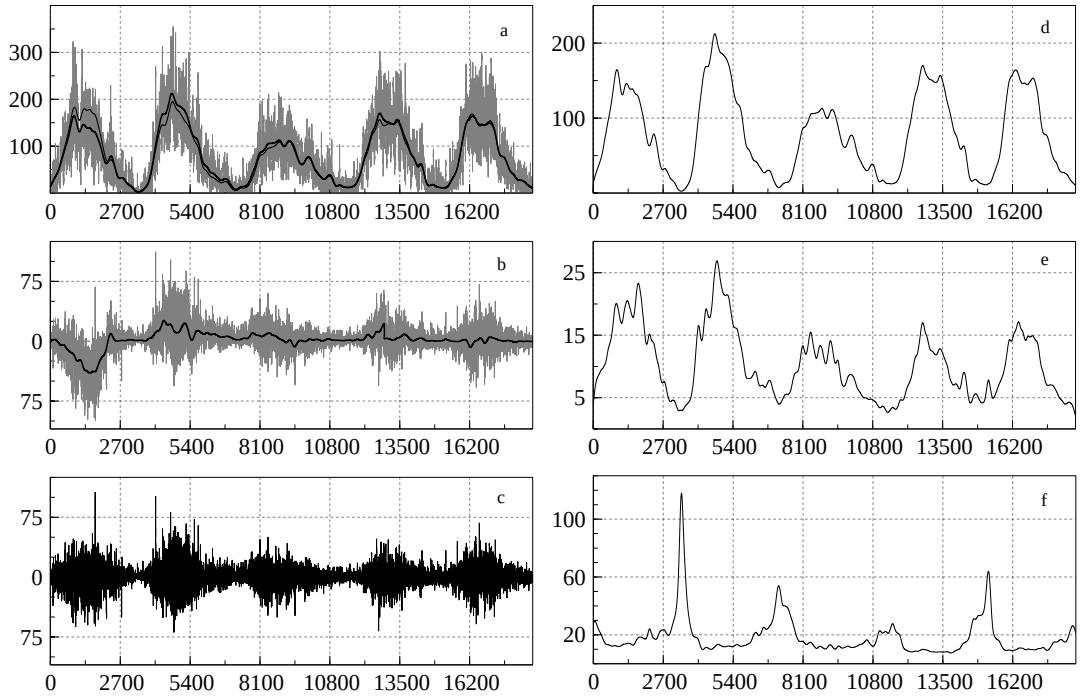


Abbildung 7.7: Vergleich der Zürcher Sonnenfleckenrelativzahl R_Z und der Amerikanischen Sonnenfleckenrelativzahl R_A zwischen dem 1. Januar 1945 und dem 31. Dezember 1995. (a) Tageswerte von R_Z sowie strukturelle Trendkomponenten von R_Z und R_A (dünne Linie). (b) Differenz der Tageswerte von R_Z und R_A sowie strukturelle Trendkomponente. (c) Differenz der Tageswerte abzüglich der strukturellen Trendkomponente. (d) Geglätteter mittlerer Verlauf der R_Z. (e) Geglättete mittlere Streuung von (c). (f) Relative mittlere Streuung in Prozenten von R_Z.

Zwar ist dieses Verhältnis nicht konstant, doch scheinen im betrachteten Zeitraum keine systematischen Abweichungen oder plötzliche Veränderungen aufzutreten. Insbesondere unterscheidet sich das Verhalten von f_Z/g_Z in den Zyklen Nr. 18 und Nr. 19 zwischen 1945 und 1965 nicht wesentlich von demjenigen in den drei nachfolgenden Zyklen. Die in der Abbildung 7.8 erkennbaren grossen Verhältniszahlen treten immer dann auf, wenn bei allgemein geringer Sonnenaktivität eine oder zwei extrem fleckenreiche Einzelgruppen auftreten, was vor allem bei sinkender Sonnenaktivität und während den Minimumsphasen häufiger der Fall ist. Gleiches gilt für das in Abbildung 7.8 dargestellte und von Waldmeier (1968) untersuchte Verhältnis R_Z/g_Z welches langfristig 12 beträgt. Allerdings gibt es charakteristische, von der Zyklusphase abhängige Abweichungen, da der relative Anteil der Gruppenzahl an der Relativzahl im Minimum grösser ist als im Maximum.

Dies ist auch der Hauptgrund, weshalb die von Hoyt und Schatten (1998) vorgelegte gruppierte Sonnenfleckenrelativzahl R_g nicht als Ersatz für die Wolfsche Sonnenfleckenrelativzahl R_Z betrachtet und somit strenggenommen auch nicht zur Homogenitätsprüfung hinzugezogen werden kann. Allerdings ist der absolute Unterschied doch relativ gering, insbesondere sollten keine systematischen Unterschiede zwischen den langfristigen Verläufen von R_{gz} und R_Z auftreten - solange das Verhältnis von Gruppenzahl zu Relativzahl langfristig konstant bleibt.

Die gruppierte Zürcher Sonnenfleckenrelativzahl R_{gz} kann nach der Formel

$$R_{gz} = 12.08\ g_Z \tag{7.4}$$

aus den Zürcher Gruppenzahlen g_Z berechnet werden, wobei der Faktor 12.08 nach Hoyt und Schatten (1998) dem mittleren Verhältnis von Gruppenzahl und Relativzahl des Royal Greenwich Observatoriums zwischen 1874 und 1976 entspricht. In Abbildung 7.9 ist die Differenz zwischen R_Z und R_{gz} wiedergegeben. Sie stellt quasi die sachliche Differenz zwischen den beiden Relativzahldefinitionen dar und zeigt, dass R_Z und R_{gz} systematische Unterschiede aufweisen und nicht ausgetauscht werden können.

Die Qualität der Rekonstruktion von R_g kann anhand von R_{gz} überprüft werden, da diese Verläufe - oder diejenigen von g und gz - identisch sein sollten. Im Gegensatz zu den Zürcher Relativzahlbestimmungen R_Z können die Zürcher Gruppenzahlen g_Z objektiv überprüft werden: es liegen für jeden Tag Zeichnungen der Sonnenoberfläche vor, aus denen die Lage und Form jeder einzelnen Fleckengruppe entnommen werden kann. Fehlt in Zürich eine Beobachtung, so ist sie durch vergleichbare Zeichnungen ausländischer Stationen ersetzt worden. Die von der Eidgenössischen Sternwarte konstruierten heliographischen Übersichtskarten und Entwicklungstabellen stellen somit ein zuverlässiges Verzeichnis der vorhandenen Sonnenfleckengruppen dar.

Die in Abbildung 7.11 dargestellten Abweichungen zwischen g und gz wären also sorgsam zu analysieren. Dies ist jedoch bisher nicht geschehen. Erstaunlicherweise basiert g im 20. Jahrhundert mehrheitlich auf den photographischen Beobachtungen des Royal Greenwich Observatoriums, welche allerdings bekanntermassen systematische Unterschiede in den erfassten Gruppenzahlen aufweisen. Neben dieser Reihe finden sich nur noch einige wenige, ansonsten ziemlich unbekannte Serien. Die offiziellen Zürcher Gruppenzahlen g_Z und die meisten der hierfür verwendeten auswärtigen Stationen sind jedoch auffälligerweise nur bis 1947 in der Rekonstruktion von Hoyt und Schatten (1998) berücksichtigt. Und genau dann beginnen die grossen systematischen Abweichungen zwischen g und g_Z.

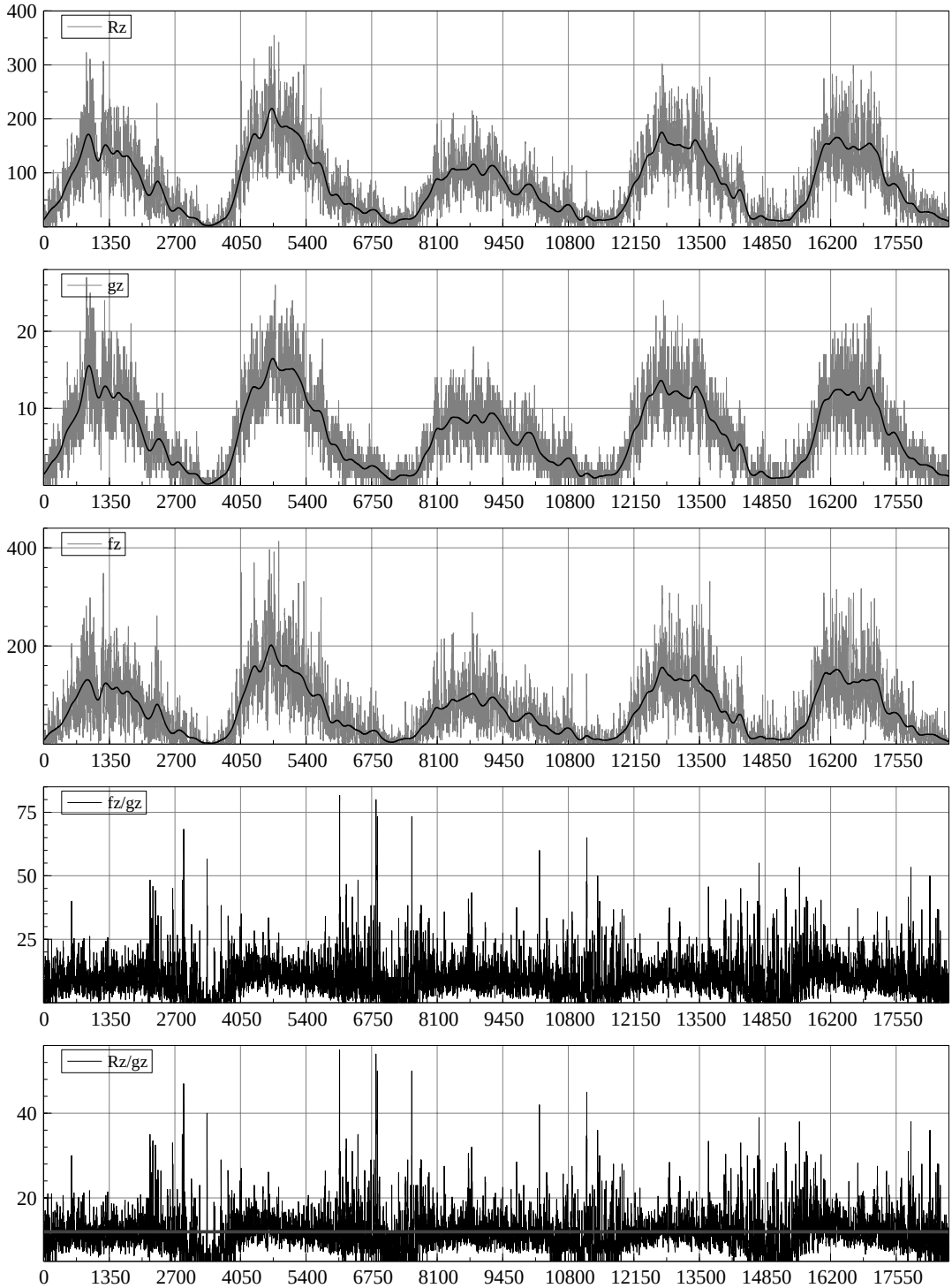


Abbildung 7.8: Verlauf der beobachteten Tageswerte der Zürcher Sonnenfleckenrelativzahl R_Z von 1945 bis 1995 sowie der Gruppenzahlen g_Z, der aus Formel (7.1) berechneten Fleckenzahlen f_Z und der Verhältnisse f_Z/g_Z und R_Z/g_Z. Letzteres beträgt langfristig ungefähr 12, was durch die fette horizontale Linie angedeutet werden soll.

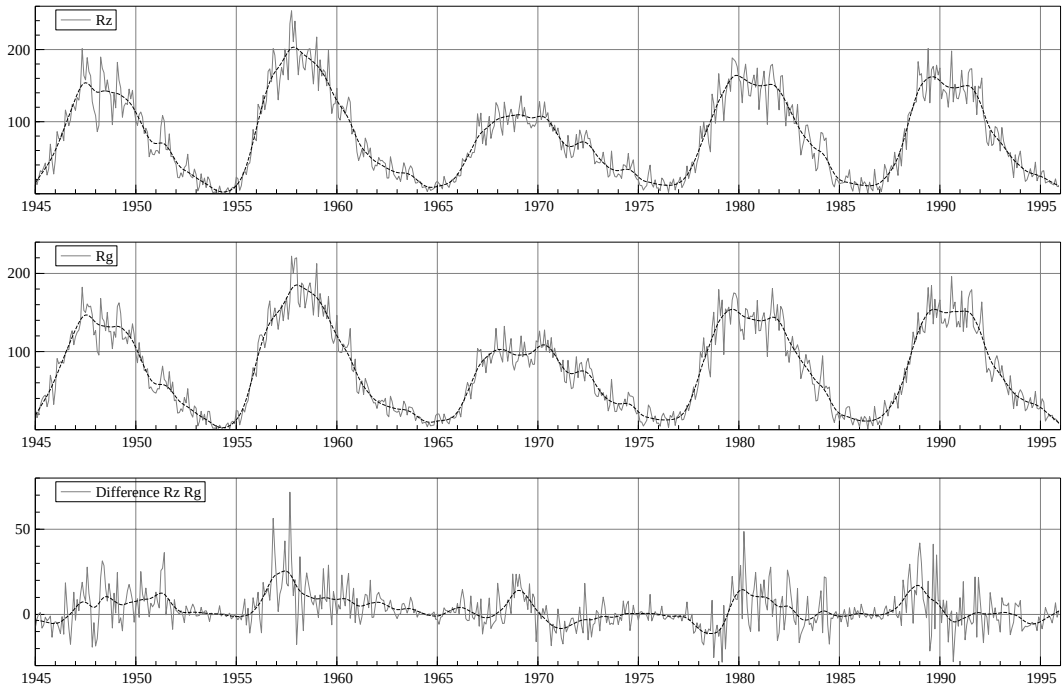


Abbildung 7.9: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der gruppierten Sonnenfleckenrelativzahl R_g von 1945 bis 1995.

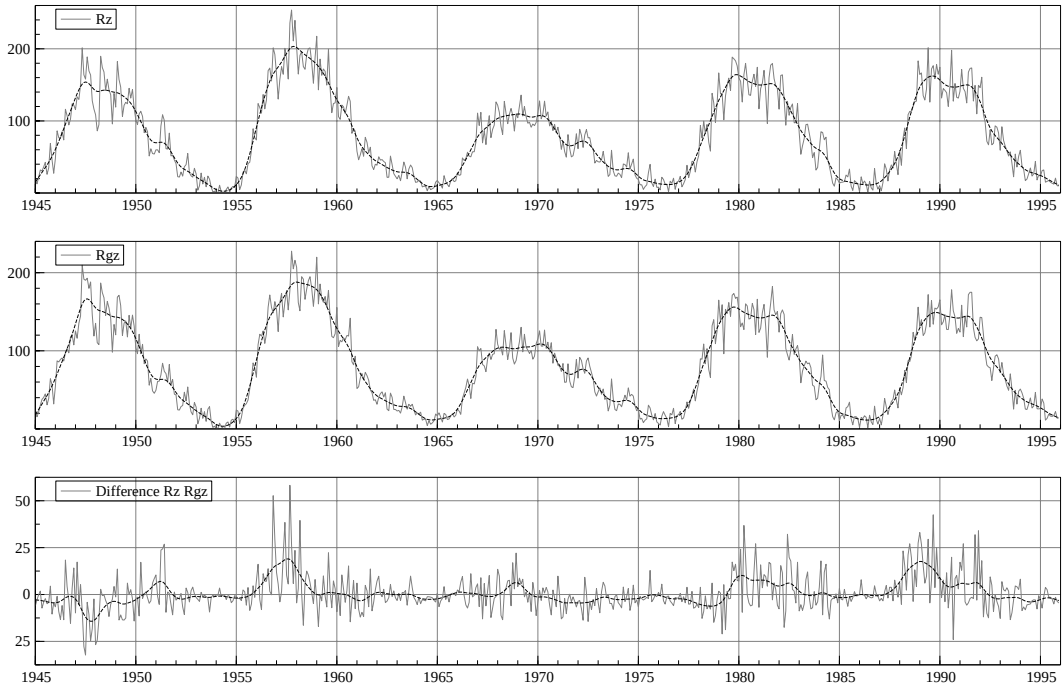


Abbildung 7.10: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der Zürcher gruppierten Sonnenfleckenrelativzahl R_{gz} von 1945 bis 1995.

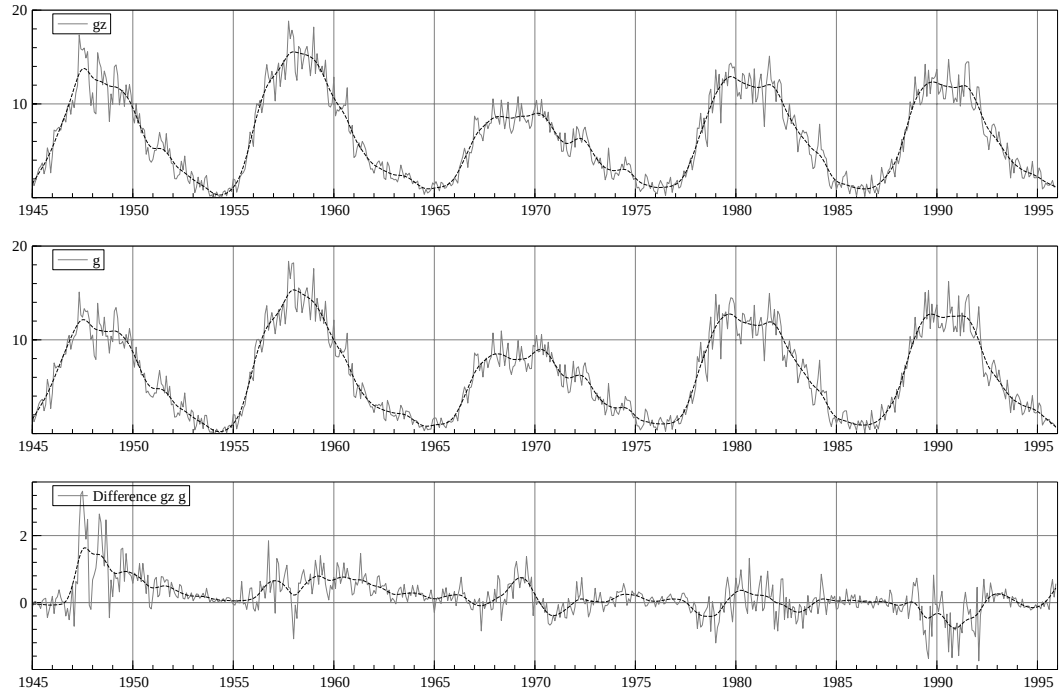


Abbildung 7.11: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckengruppenzahl g_Z und der gruppierten Sonnenfleckengruppenzahl g von 1945 bis 1995.

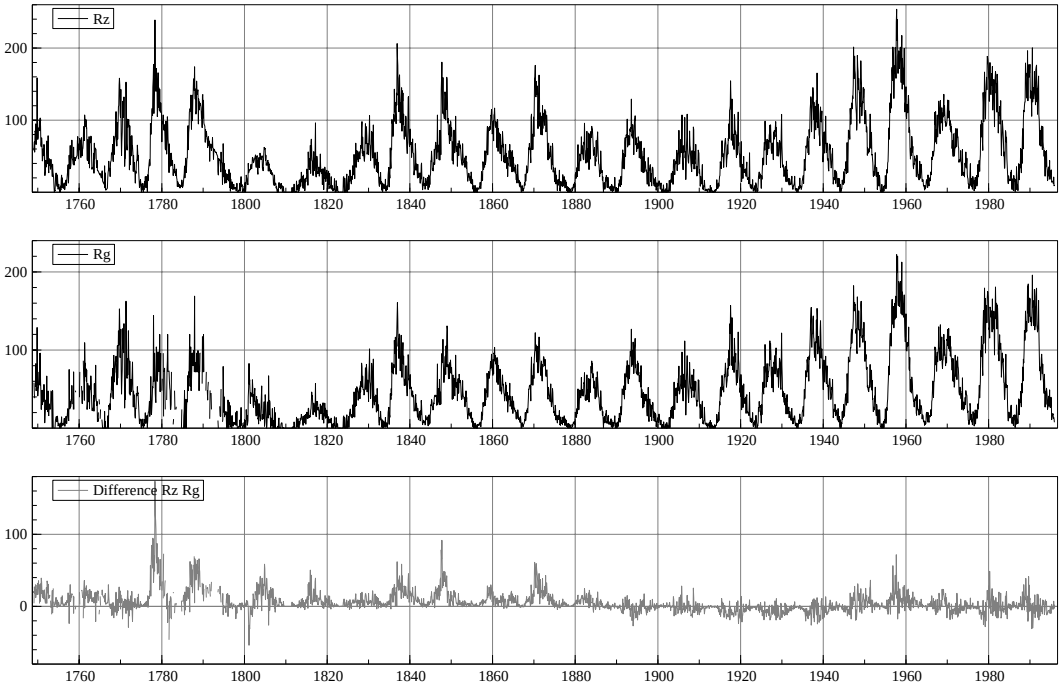


Abbildung 7.12: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der gruppierten Sonnenfleckenrelativzahl R_g von 1749 bis 1995.

7.3 Die langfristige Homogenität zwischen 1700 und 2005

Betrachten wir die in Abbildung 7.12 wiedergegebene Reihe der Differenzen der Monatsmittel von R_Z und R_g zwischen 1749 und 1995, wird sofort klar, dass die auftretenden Unterschiede zwischen R_Z und R_g keine Inhomogenitäten im engeren Sinne darstellen, sondern Rekonstruktionsunterschiede wiedergeben: Dass der in R_Z zweithöchste Sonnenaktivitätszyklus Nr. 3 um 1780 in R_g klein wird, die Aktivitätszyklen im Dalton Minimum um 1800 noch kleiner und kantiger werden etc. hat seine Ursachen wahrscheinlich in den stark unterschiedlichen Rekonstruktionsverfahren beider Reihen. Es könnte sich aber auch um systematische Unterschiede im Verhältnis R/g handeln, was leider nicht überprüft werden kann, da Hoyt und Schatten (1998) leider nur die Gruppenzahlen g rekonstruierten und in ihrer systematischen Neuerfassung alter Sonnenfleckenbeobachtungen die f unberücksichtigt liessen.

Trotzdem fallen einzelne Inhomogenitäten sofort ins Auge: zum einen scheint da kurz vor dem Beginn der systematischen Beobachtungen von Rudolf Wolf eine Inkonsistenz in der Wolfschen Reihe vorzuliegen, wie bereits aus dem in Abbildung 6.5 ersichtlichen ungewöhnlichen Anstiegsverhalten des 9. Aktivitätszyklus vermutet und in Abbildung 7.13 deutlicher festgestellt werden kann.

Ein Blick in die Wolfschen Originalregister (Wolf 1878) ergibt, dass 1847 die Standardbeobachtungen von Heinrich Samuel Schwabe eventuell mit einem falschen Reduktionsfaktor kalibriert wurden, jedenfalls lässt das Verhalten der Reduktionsfaktoren der anderen Beobachter in den Jahren davor und danach dies vermuten. Auch sei darauf hingewiesen, dass zu Beginn des Jahres 1849 - dem Beginn der Wolfschen Beobachtungsreihe - ein weiterer markanter Bruch auftritt. Hingegen konnten wir für die um 1870 auftretenden, in Abbildung 7.14 im Detail gezeigten, Differenzen zwischen R_Z und R_g keine offensichtlichen Inkonsistenzen in den Originalregistern finden.

Auch sei auf den eigentümlichen Umstand hingewiesen, dass zwischen 1894 und 1945 die Differenzen zwischen R_Z und R_g vorwiegend negativ, vorher (und nachher) vorwiegend positiv sind. Diese Periode fällt genau zwischen die beiden Wechsel der Standardbeobachter von Wolf zu Wolfer (1894) und von Brunner zu Waldmeier (1945) und könnte daher sehr wohl in Inhomogenitäten der Wolfschen Reihe begründet sein: wie erwähnt, ist es wahrscheinlich, dass während den Maximumsphasen die Waldmeierschen Zahlen tendenziell zu gross sind. Umgekehrt deutet einiges darauf hin, dass Wolfer (1895) mit dem von ihm bestimmten $k = 0.6$ etwas zu stark korrigiert hat, so dass seine Zahlen und diejenigen seines Nachfolgers Brunner systematisch zu tief ausgefallen sind.

Betrachten wir schliesslich die in Abbildung 7.15 gezeigten Jahresmittel zwischen 1700 und 1995, so ist hieraus für die bereits diskutierten Jahre zwischen 1749 und 1995 nichts neues hinzuzufügen. Was die Jahre zwischen 1700 und 1749 betrifft, so frage ich mich allerdings, was die Neurekonstruktion der Sonnenaktivität durch Hoyt und Schatten (1998) letztlich gebracht hat, scheint doch die Qualität der Rekonstruktion vor 1749 förmlich zusammenzubrechen: wie der Abbildung 7.16 entnommen werden kann, sind gerade für die so wichtigen Anschlussjahre vor 1749 (und vor 1849) keine neuen Beobachter bekannt geworden. Auch scheinen die elfjährigen Aktivitätszyklen (falls überhaupt von solchen überhaupt noch gesprochen werden darf) ganz anderen Gesetzmässigkeiten zu gehorchen, als die nachfolgenden.

Dies schlägt sich bis zu einem gewissen Grad auch im ganzen Verlauf von R_g nieder, zeigt doch die in Abbildung 7.17 wiedergegebene Zustandsraumdarstellung von R_g ein ganz anderes Verhalten, als die in Abbildung 3.14 gezeigte Zustandsraumdarstellung der originalen Wolfschen Reihe.

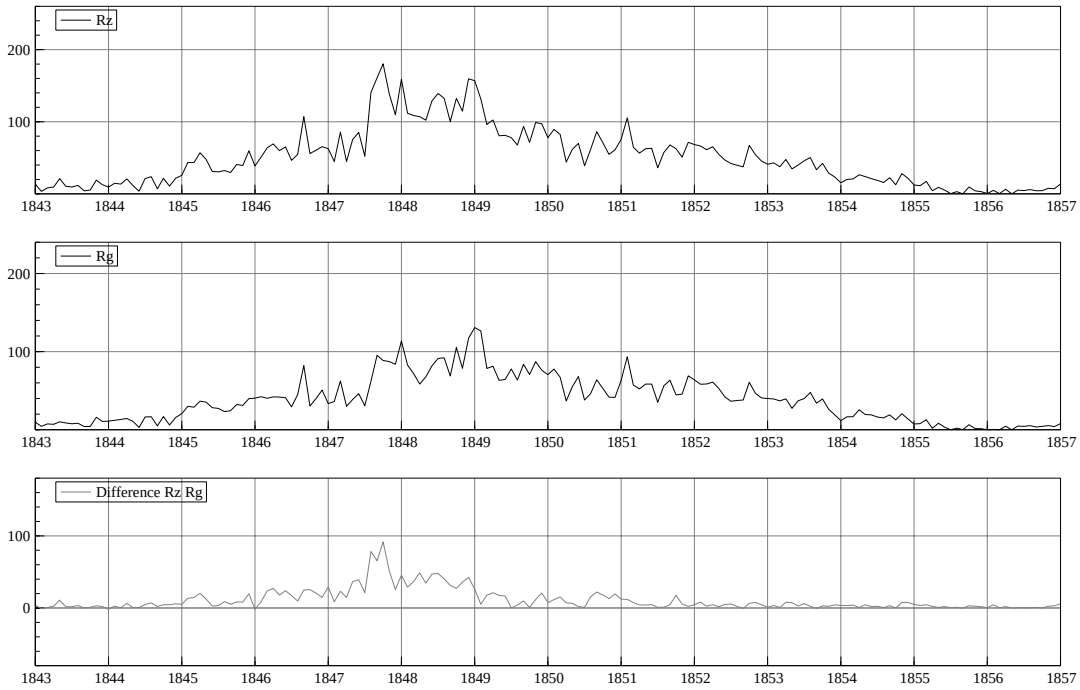


Abbildung 7.13: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der gruppierten Sonnenfleckenrelativzahl R_g von 1843 bis 1856.

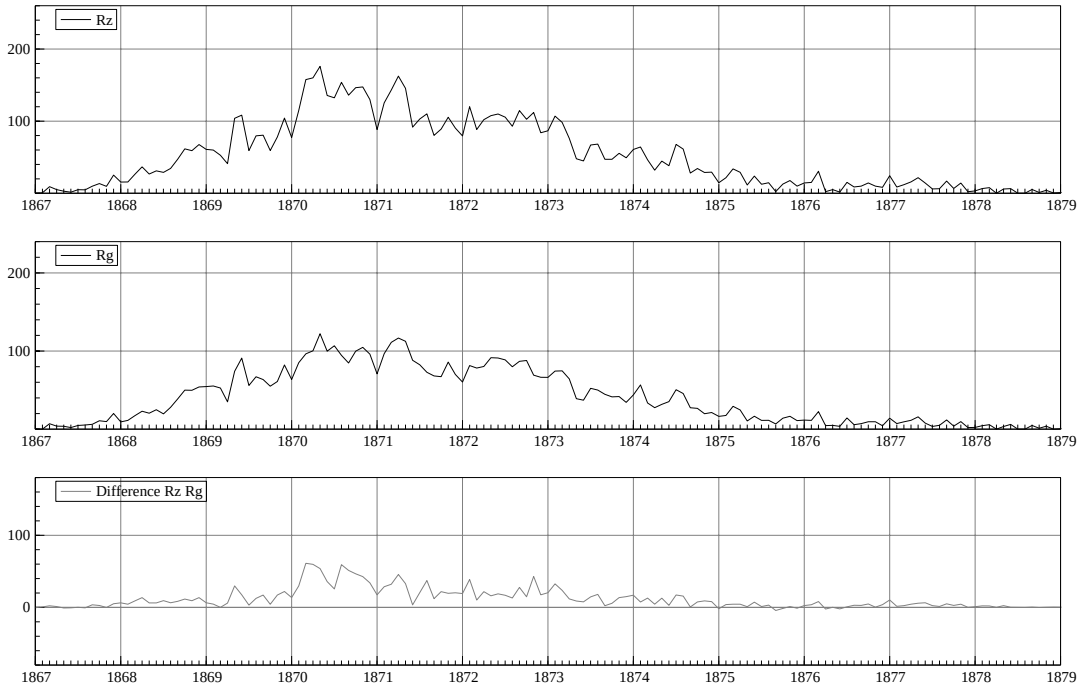


Abbildung 7.14: Verlauf der beobachteten Differenzen der Monatsmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der gruppierten Sonnenfleckenrelativzahl R_g von 1867 bis 1878.

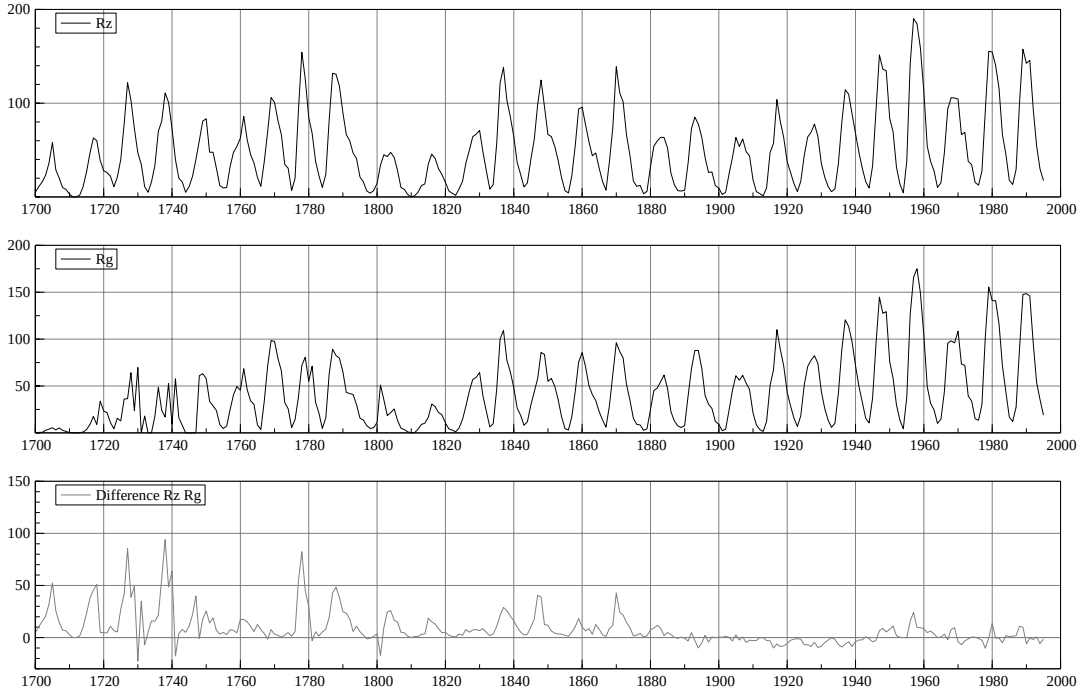


Abbildung 7.15: Verlauf der beobachteten Differenzen der Jahresmittel der Zürcher Sonnenfleckenrelativzahl R_Z und der gruppierten Sonnenfleckenrelativzahl R_I von 1700 bis 1995.

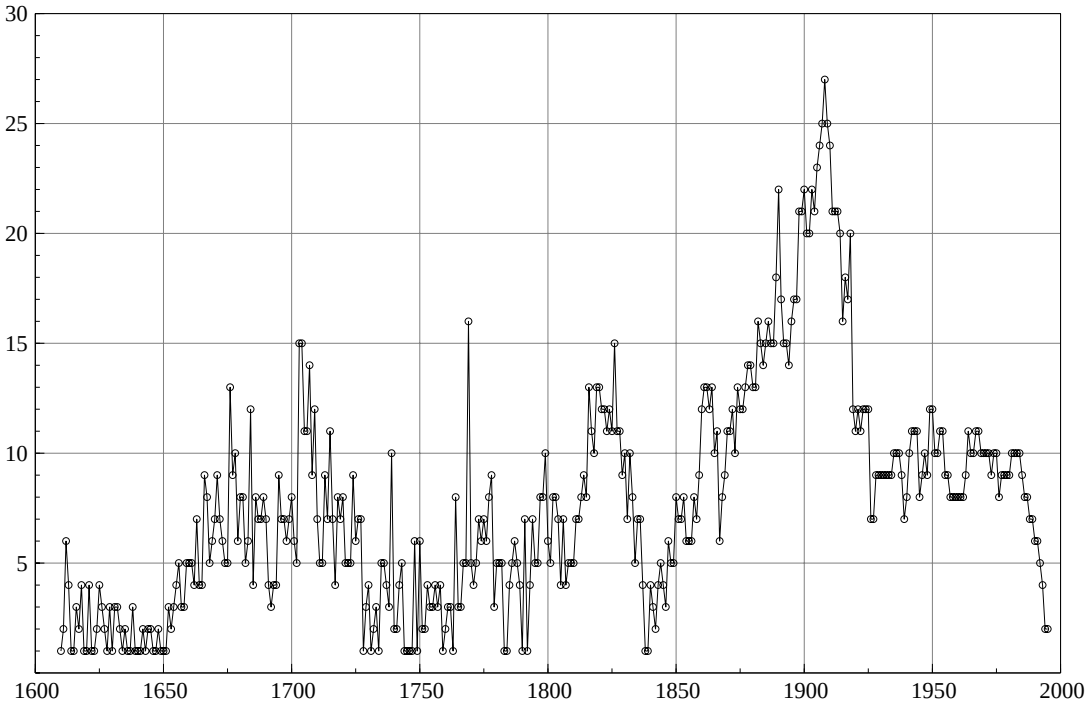


Abbildung 7.16: Anzahl der Beobachter welche zur Rekonstruktion der gruppierten Sonnenfleckenrelativzahl R_g verwendet wurden (Hoyt und Schatten 1998).

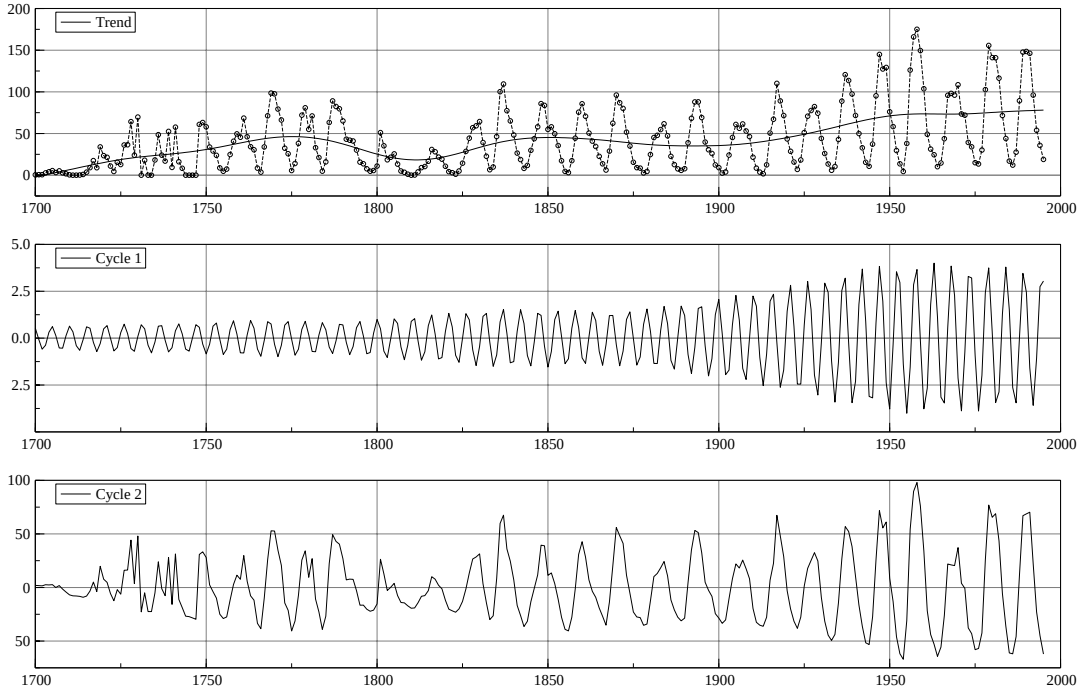


Abbildung 7.17: Dekomposition der Jahresmittel der gruppierten Sonnenfleckenrelativzahl R_g in einen stochastischen Trend und zwei stochastische Zyklen.

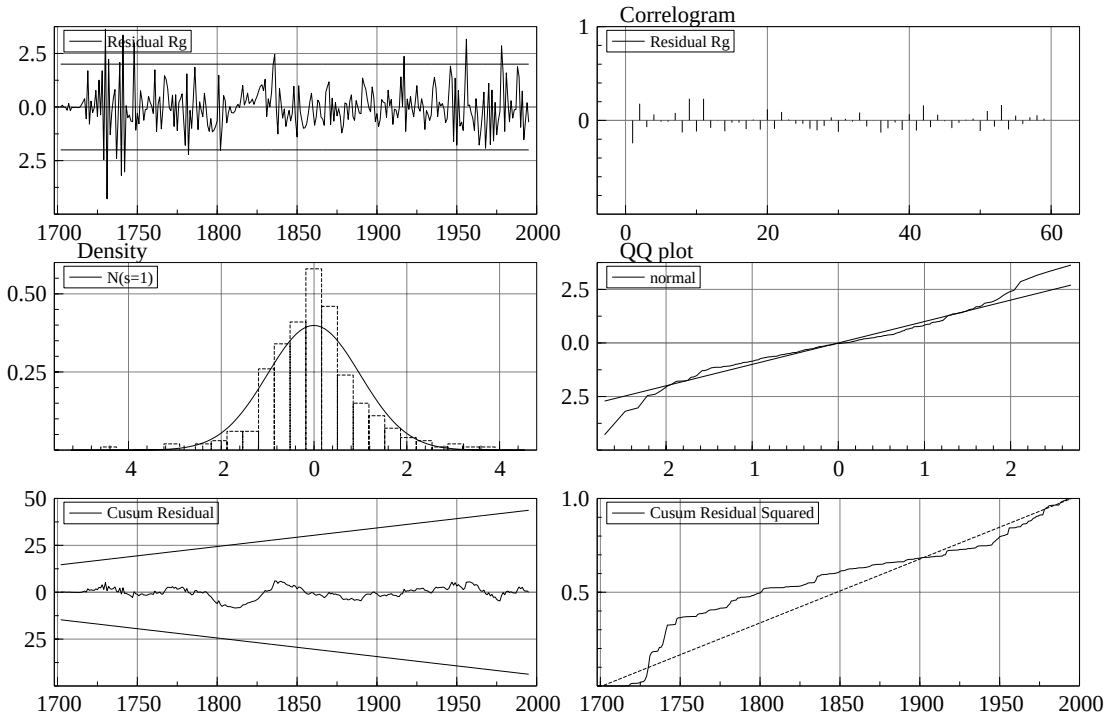


Abbildung 7.18: Diagnostik zu den Innovationen des Modells in Abbildung 7.17. Zu sehen sind die standardisierten Innovationen, deren Autokorrelationsfunktion, Histogramm, QQ-Plot und kumulierte Summe sowie die kumulierte Summe der Innovationsquadrate.

8. Schlussfolgerungen und Ausblick

Hauptgegenstand der vorliegenden Arbeit war die Homogenitätsprüfung ausgewählter Teile der Wolfschen Reihe der Zürcher Sonnenfleckenrelativzahlen R_Z, insbesondere während der Periode zwischen 1945 und 2003 mit dem auf den 1. Januar 1981 erfolgten Wechsel des Standardbeobachters. Hierzu wurde ein lineares strukturelles Zustandsraummodell entwickelt, welches für die Darstellung und die Homogenitätsprüfung der Wolfschen Reihe geeignet ist (Kapitel 3). Ausgehend von einigen in der Klimatologie zur Homogenisierung langer Beobachtungs- und Messreihen verwendeten statistischen Prüfverfahren wurden verschiedene auf dem Likelihood Ratio Prinzip beruhenden oder ihm nahestehenden statistische Tests validiert und verbessert (Kapitel 4 und 5). Schliesslich konnte gezeigt werden, dass mit Hilfe des bei der Parameterschätzung eines linearen strukturellen Zustandsraummmodells bestimmten Glättungsfehlers die Teststatistiken der Homogenitätstests berechnet werden können (Kapitel 6). Hiermit wurden die Verläufe der Tages-, der Monats- und der Jahresmittel der Wolfschen Reihe auf ihre Homogenität hin geprüft (Kapitel 7).

"Do we have the correct reconstruction of solar activity?"

Diese Frage stellten Hoyt und Schatten (1994) anlässlich des Jubiläums zum 100. Todestages von Rudolf Wolf am 6. Dezember 1993 – und beantworteten sie negativ, insbesondere was den Verlauf der Sonnenaktivität im 17., 18. und 19. Jahrhundert betrifft. Aus den Resultaten unseres 7. Kapitels wird jedoch deutlich, dass die von Hoyt und Schatten (1998) vorgelegte Rekonstruktion der Sonnenaktivität eher schlechter ist, als diejenige von Rudolf Wolf. Darüberhinaus gibt es jetzt auch Hinweise, dass sogar in den bisher als homogen erachteten Daten des 20. Jahrhunderts Inhomogenitäten aufgetreten sind.

Nachgewiesen werden konnte insbesondere im Vergleich mit der Reihe der amerikanischen Sonnenfleckenrelativzahl R_A eine sprunghafte Mittelwertsverschiebung am 1. Mai 1980 in den Reihen der Internationalen Sonnenfleckenrelativzahl R_I und der Zürcher Sonnenfleckenrelativzahl R_Z. Dieser Bruch wird möglicherweise durch den Wechsel der Standardbeobachter verursacht.

Weitere mögliche Inhomogenitätsstellen konnten im Vergleich der Zürcher Sonnenfleckenrelativzahlen R_Z mit den gruppierten Sonnenfleckenrelativzahlen R_g zu Beginn des 18. Sonnenfleckenzyklus um 1947 identifiziert werden. Andere Inhomogenitätsstellen liegen um 1847, 1861, 1894 und 1926 im Umkreis der Wechsel der jeweiligen langjährigen Standardbeobachter. Eine davon unabhängige Inhomogenitätsstelle unbekannter Ursache befindet sich um 1870.

Eine eingehendere Homogenitätsprüfung der Wolfschen Reihe kann momentan nicht vorgenommen werden, da die gefundenen Differenzen offensichtlich weniger auf Inhomogenitäten beruhen, denn auf Mängeln der diversen Rekonstruktionsverfahren und es daher offensichtlich an einer validen Auswertungsmethode fehlt, welche es erlaubt, die mehr oder weniger kurzen Serien von Einzelbeobachtungen zu einer gemeinsamen, langfristig homogenen Reihe von täglichen Sonnenfleckenrelativzahlen zusammenzufassen.

Ein hierzu geeignetes Modell stellt möglicherweise ein verallgemeinertes dynamisches lineares gemischtes Modell (DGLMM) in Zustandsraumform dar, wie es beispielsweise von Shumway und Verosub (2002) für die Rekonstruktion der zusammengesetzten Reihe der jährlichen Temperaturanomalien der Nordhemisphäre vorgeschlagen wurde. Eingehendere Erfahrungen mit einem solchen Modell sowie geeignete Software sind bisher allerdings noch nicht vorhanden.

Auch wenn bereits ein geeignetes Reduktionsverfahren vorliegen würde, müssten zuerst alle Einzelbestimmungen der Gruppen- und Fleckenzahlen der einzelnen Sonnenbeobachter zusammengetragen und elektronisch erfasst werden. Dies erscheint insofern schwierig, als die zur Konstruktion der Wolfschen Reihe berücksichtigten Originalbeobachtungen nur bis 1926 vorhanden sind und die nachfolgenden als verschollen gelten (Friedli 2001).

A. Zufallszahlengeneratoren

Die Validität von Monte-Carlo Experimenten hängt nach Knuth (1981) entscheidend von der Wahl eines geeigneten Zufallszahlengenerators ab.

Nachstehend werden drei neuere Zufallszahlengeneratoren von Park und Miller (1988), von Marsaglia (1997) und von L'Ecuyer (1997) miteinander verglichen, welche in Doornik (1998) beschrieben sind. Den Empfehlungen von Knuth (1981) folgend, wurden zudem die von den drei Generatoren erzeugten Zufallszahlenfolgen mittels einem in Press et al. (1990) explizite beschriebenen Algorithmus jeweils zusätzlich gemischt. Ein diagnostischer Vergleich dieser sechs Generatoren anhand kurzer Sequenzen von simulierten standardnormalverteilten Zufallszahlen ist in den Abbildungen A.1 bis A.3 zusammengestellt.

Es zeigt sich, dass mit allen sechs Generatoren brauchbare Zufallszahlensequenzen simuliert werden können. Die zusätzlich gemischten Zufallszahlengeneratoren sind den originalen jedoch vorzuziehen, *"since the additional shuffling wont't make a sequence any less random, and since it probably enhances the randomness substantially with very little extra cost, it can be recommended for use in combination with any other random number generator"* (Knuth, 1981; Kapitel 3.2.2).

Für die umfangreichen Monte-Carlo Experimente dieser Arbeit wurde der zusätzlich gemischte Generator von Park und Miller verwendet, da dieser der schnellste der modifizierten Generatoren ist. Der Generator von Park und Miller ist ein linearer Kongruenzgenerator mit Modulus $2^{31} - 1$ und erfüllt den Spektraltest. Einem Lemma von Knuth (1981; Kapitel 3.3.3) zufolge, liefert der originale Generator von Park und Miller somit Sequenzen von Zufallszahlen, welche weitgehend frei von Autokorrelationen sind, was durch unsere Simulation in Abbildung A.1 zumindest nicht widerlegt wird.

A.1 Ox-Programmcode

Das nachfolgende Programm liefert Sequenzen von unabhängigen, identisch standardnormal verteilten Zufallsvariablen und ist in Ox 2.20 geschrieben, einer von J. A. Doornik vom Nuffield College in Oxford entwickelten objekt-orientierten Programmiersprache mit einer umfassenden mathematischen und statistischen Funktionsbibliothek (Doornik, 1998). Die Syntax von Ox ist derjenigen von C, C++ und Java sehr ähnlich, insbesondere was den Gebrauch von Schleifen, Funktionen, Feldern und Klassen betrifft. Matrizen können wie gewöhnliche Variablen multipliziert oder invertiert werden.

```
#include <oxstd.h>                    // include the Ox standard library headers
#include <oxdraw.h>
#include <arma.h>

/*************************************************************************/
randn0(const iN, const G)
/*************************************************************************
Purpose:              This function calculates NID(0,1) distributed random num-
                      bers, based on the machine provided random number genera-
                      tor G and the refinement method ran0 of the 'Numerical
                      Recipes in Pascal' p. 216.

Input:                iN:    Number of simulated random numbers
                      G:     Machine provided random number generator

Output:               meps:  Sequence of standard normally distributed random
                             numbers

Created:              2000-12-31
Last modified:        2001-01-01

Version:              1.0

Author:               Thomas K. Friedli, IMSV, University of Berne

*************************************************************************/
{
decl dum;                                              // Dummy random numbers
decl idum;                                            // Initialisation switch
decl i;                                                      // Index over iN
decl j,k;                                          // Indices of shuffling vectors
decl RanOV=zeros(97,1);                                     // Shuffling vector
decl RanOY;                                  // Random number and shuffling index
decl meps=zeros(iN,1);                                 // Output random sequence
decl RanON=zeros(2,1);                     // Variables for normal transformation
decl iset,r,fac,eps;                       // Variables for normal transformation

ranseed(G);                                 // Selection of random number generator
idum=-1;                               // Forces initialisation of shuffling vector
iset=0;                                   // Switch for use of second random number

for (i = 0; i < iN; ++i)
{
 if (iset==0)
 {
  if (idum<0)
  {
     idum=1;
     for (j=0; j < 97; ++j)
     {
      dum=ranu(1,1);
     }
     for (j=0; j < 97; ++j)
     {
      RanOV[j]=ranu(1,1);
     }
```

```
    RanOY=ranu(1,1);
  }
  do
  {
   for (k=0; k < 2; ++k)
   {
   j=trunc(97.0*RanOY);
   RanOY=RanOV[j];
   RanON[k]=2*RanOY-1;
   RanOV[j]=ranu(1,1);
   }
   r=sqr(RanON[0])+sqr(RanON[1]);
  }
  while (r <= 0 || r >= 1);

  fac=sqrt(-2.0*log(r)/r);
  eps=RanON[0]*fac;
  meps[i]=RanON[1]*fac;
  iset=1;
 }
 else
 {
 meps[i]=eps;
 iset=0;
 }
}
return meps;
}
/*****************************************************************************/

/*****************************************************************************/
main()
/*****************************************************************************
Purpose:              This program simulates normally and identically distribu-
                      ted random numbers, based on three machine provided ran-
                      dom number generators "PM", "GM" and "LE", and the refi-
                      nement method ran0 as recommended in the 'Numerical Reci-
                      pes in Pascal' of Press et al. (1990).

Output:               meps:   Sequence of simulated random numbers

Created:              2000-12-31
Last modified:        2001-01-01

Version:              1.0

Author:               Thomas K. Friedli, IMSV, University of Berne

*****************************************************************************/
{
decl iM;                                      // Number of simulated time series
decl i;                                        // Index of simulated time series
decl iN;                                           // Number of simulated points
decl ip;                                        // Show simulation progress plot
decl G;                                               // Random number generator
decl meps1,meps2;                                      // Simulated disturbances
decl macf1,macf2,mpacf1,mpacf2;                     // Autocorrelation variables
```

```
decl j,sum,msum1,msumq1,msum2,msumq2;              // Cumulative sum variables
decl smat;                                                 // Output data matrix

iM=1;                                          // Number of simulated time series
iN=100000;                                          // Number of simulated points
G="PM";                                                // Random number generator
ip=0;                                            // Show simulation progress plot

ranseed(G);

for (i = 0; i < iM; ++i)
{
 meps1 = rann(iN,1);
 meps2 = randn0(iN,G);
 if (ip==1)
 {
  Draw(0, meps1', 0, 1);
  DrawTitle(0, G);

  Draw(1, meps2', 0, 1);
  DrawTitle(1, G~" Shuffled");

  SetDrawWindow("Simulation Progress");
  ShowDrawWindow();
 }
}
Draw(0, meps1', 0, 1);
DrawTitle(0, G);
Draw(1, meps2', 0, 1);
DrawTitle(1, G~" Shuffled");
DrawDensity(2, meps1', G, TRUE, TRUE, TRUE);
DrawAdjust(ADJ_MINMAX, 0, 0.6);
DrawTitle(2, " ");
DrawDensity(3, meps2', G~" Shuffled", TRUE, TRUE, TRUE);
DrawAdjust(ADJ_MINMAX, 0, 0.6);
DrawTitle(3, " ");
DrawQQ(4, meps1', G, QQ_N, 0, 0);
DrawTitle(4, " ");
DrawQQ(5, meps2', G~" Shuffled", QQ_N, 0, 0);
DrawTitle(5, " ");
SetDrawWindow("Distribution of Simulated Random Numbers");
ShowDrawWindow();

SaveDrawWindow("Distribution.eps");

Draw(0, meps1', 0, 1);
DrawTitle(0, G);
Draw(1, meps2', 0, 1);
DrawTitle(1, G~" Shuffled");
macf1=acf(meps1, min(iN/2,1000));
DrawMatrix(2, macf1[1:][]', "ACF "~G, 1, 1);
DrawAdjust(ADJ_INDEX,1);
DrawAdjust(ADJ_MINMAX, -1.96/sqrt(iN)-0.01, -1.96/sqrt(iN)+0.03);
DrawLine(2, 0, 1.96/sqrt(iN), min(iN/2,1000), 1.96/sqrt(iN), 14);
DrawLine(2, 0, -1.96/sqrt(iN), min(iN/2,1000), -1.96/sqrt(iN), 14);
macf2=acf(meps2, min(iN/2,1000));
DrawMatrix(3, macf2[1:][]', "ACF "~G~" Shuffled", 1, 1);
```

```
DrawAdjust(ADJ_INDEX,1);
DrawAdjust(ADJ_MINMAX, -1.96/sqrt(iN)-0.01, -1.96/sqrt(iN)+0.03);
DrawLine(3, 0, 1.96/sqrt(iN), min(iN/2,1000), 1.96/sqrt(iN), 14);
DrawLine(3, 0, -1.96/sqrt(iN), min(iN/2,1000), -1.96/sqrt(iN), 14);
mpacf1=pacf(acf(meps1, min(iN/2,1000)));
DrawMatrix(4, mpacf1[1:][]', "PACF "~G, 1, 1);
DrawAdjust(ADJ_INDEX,1);
DrawAdjust(ADJ_MINMAX, -1.96/sqrt(iN)-0.01, -1.96/sqrt(iN)+0.03);
DrawLine(4, 0, 1.96/sqrt(iN), min(iN/2,1000), 1.96/sqrt(iN), 14);
DrawLine(4, 0, -1.96/sqrt(iN), min(iN/2,1000), -1.96/sqrt(iN), 14);
mpacf2=pacf(acf(meps2, min(iN/2,1000)));
DrawMatrix(5, mpacf2[1:][]', "PACF "~G~" Shuffled", 1, 1);
DrawAdjust(ADJ_INDEX,1);
DrawAdjust(ADJ_MINMAX, -1.96/sqrt(iN)-0.01, -1.96/sqrt(iN)+0.03);
DrawLine(5, 0, 1.96/sqrt(iN), min(iN/2,1000), 1.96/sqrt(iN), 14);
DrawLine(5, 0, -1.96/sqrt(iN), min(iN/2,1000), -1.96/sqrt(iN), 14);
SetDrawWindow("Autocorrelations of Simulated Random Numbers");
ShowDrawWindow();

SaveDrawWindow("Autocorrelations.eps");

sum=0;
msum1=zeros(iN,1);
for (j=0; j < iN; ++j)
{
msum1[j]=sum+meps1[j];
sum=msum1[j];
}

sum=0;
msumq1=zeros(iN,1);
for (j=0; j < iN; ++j)
{
msumq1[j]=sum+meps1[j]*meps1[j];
sum=msumq1[j];
}

sum=0;
msum2=zeros(iN,1);
for (j=0; j < iN; ++j)
{
msum2[j]=sum+meps2[j];
sum=msum2[j];
}

sum=0;
msumq2=zeros(iN,1);
for (j=0; j < iN; ++j)
{
msumq2[j]=sum+meps2[j]*meps2[j];
sum=msumq2[j];
}

Draw(0, meps1', 0, 1);
DrawTitle(0, G);
Draw(1, meps2', 0, 1);
DrawTitle(1, G~" Shuffled");
```

```
DrawMatrix(2, msum1', G~" Cumulative Sum", 1, 1);
DrawAdjust(ADJ_MINMAX, 0, max(msum1,msum2)+5);
DrawMatrix(3, msum2', G~" Shuffled Cumulative Sum", 1, 1);
DrawAdjust(ADJ_MINMAX, 0, max(msum1,msum2)+5);
DrawMatrix(4, msumq1', G~" Cumulative Sum of Squares", 1, 1);
DrawAdjust(ADJ_MINMAX, 0, iN+50);
DrawMatrix(5, msumq2', G~" Shuffled Cumulative Sum of Squares", 1, 1);
DrawAdjust(ADJ_MINMAX, 0, iN+50);

SetDrawWindow("Cumulative Sums of Simulated Random Numbers");
ShowDrawWindow();

SaveDrawWindow("CumulativeSums.eps");

smat=meps1~meps2;
savemat("smat.in7",smat);
}
/*****************************************************************************/
```

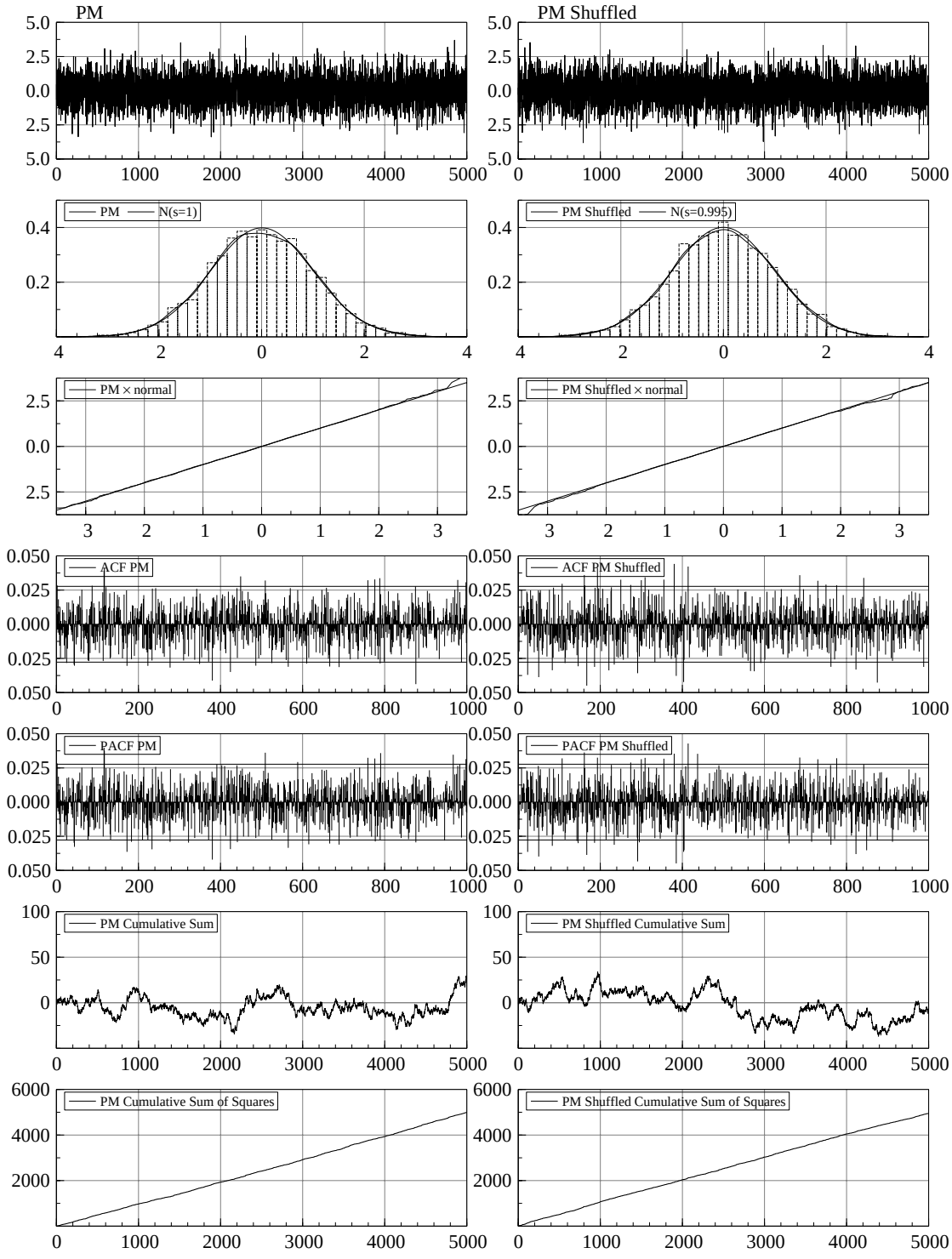


Abbildung A.1: Diagnostik zum Generator von Park und Miller (1988).

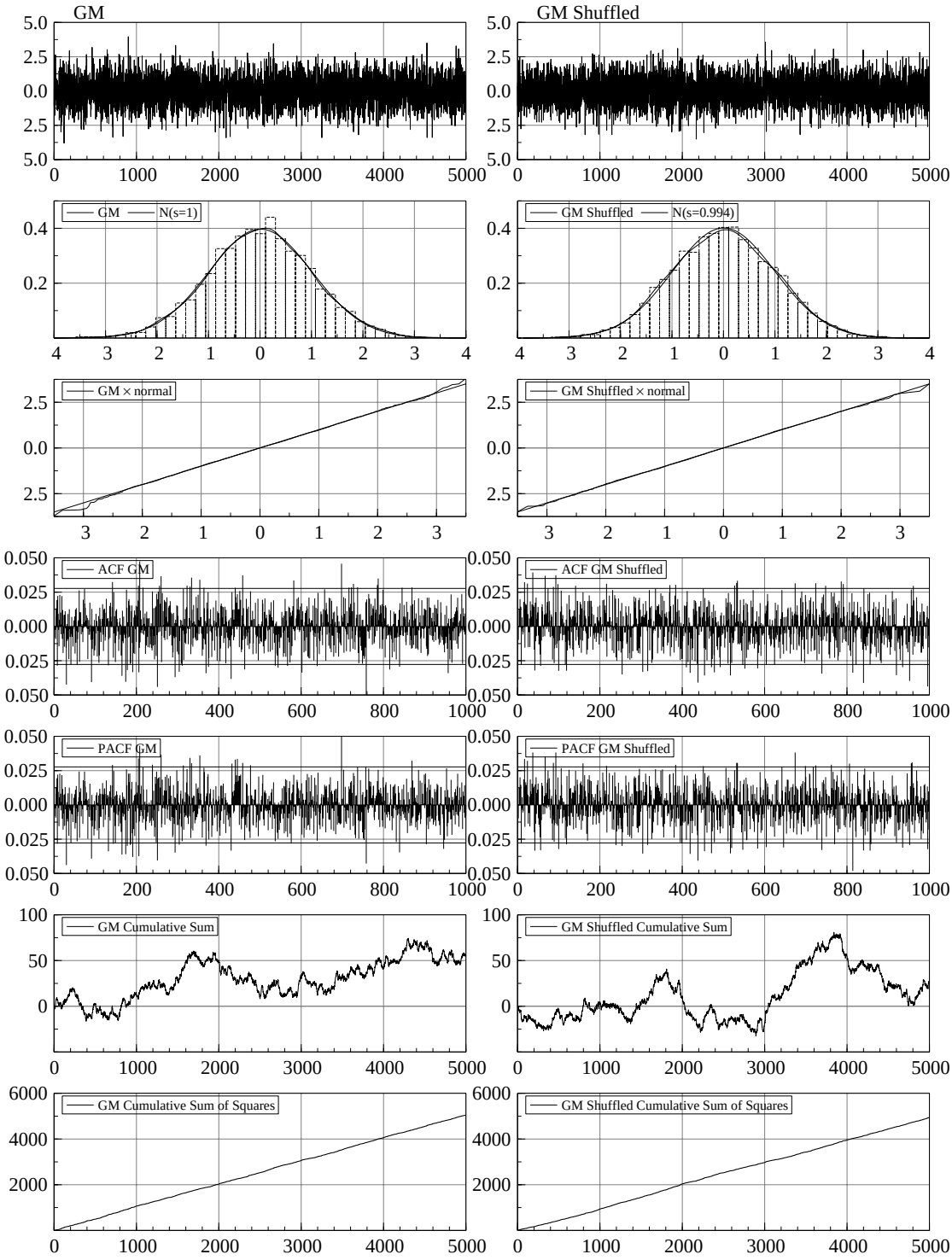


Abbildung A.2: Diagnostik zum Generator von Marsaglia (1997).

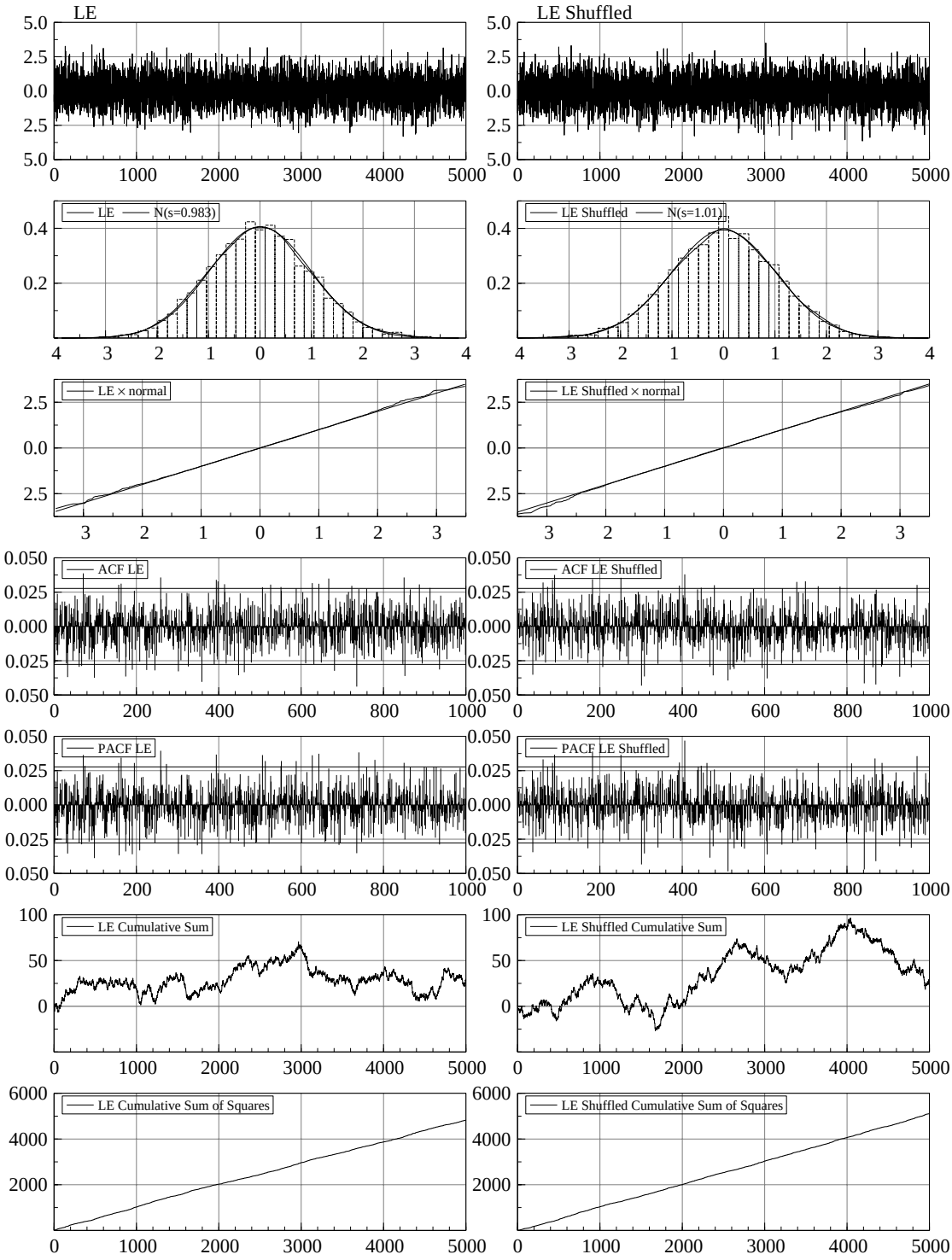


Abbildung A.3: Diagnostik zum Generator von L'Ecuyer (1997).

B. Testreihe Z80

1	0.09662657	-0.02830058	1.27715561	-0.9069845	-0.41274013
6	0.11203670	0.76984218	-1.76156687	0.06532592	-2.01159853
11	0.92871607	-0.05837132	0.35834206	0.29794732	-1.32552432
16	0.42069070	-1.55127330	0.12907378	-0.25440604	-0.88599338
21	-0.73798472	-0.36639623	0.50945495	-0.82556322	1.97456790
26	1.66090448	-0.11026234	-0.65498769	2.26249830	-2.07575876
31	0.68228835	-0.32894305	-2.08588357	-0.82152638	-1.01652472
36	-0.48102051	-0.46434802	-0.81626120	0.27992786	1.37996076
41	-0.11656543	0.89393769	0.70166888	2.46326422	0.80630575
46	1.16195934	-1.65182125	0.03206184	-1.5124517	-0.82357569
51	-0.03566778	-0.11964957	-0.35309437	-0.53539196	1.48979781
56	-0.33191027	0.01129300	1.19282610	0.96416407	1.44076680
61	-1.37869509	0.87464908	-1.11017515	-0.46821867	-0.46705127
66	1.10149617	0.82261227	-1.93622092	1.07484414	-1.25226379
71	-0.13731166	-1.07693555	1.25944428	1.07130183	-1.34867458
76	-0.13623151	2.48278914	-0.67180233	0.25655448	-0.02259524

Tabelle B.1: Simulierte standardnormalverteilte Testreihe Z80

C. Monte - Carlo Experimente

Abb.	LR1	LR2	LR3	LR4	LR5	LR6	LR7	A	B	LR8	LR9	LR10
1	LR1	LR2	LR3	LR4	LR5	LR6	LR7	A	B	LR8	LR9	LR10
2	LR1	LR2	LR3	LR4	LR5	LR6	LR7	A	B	LR8	LR9	LR10
3	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
4	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
5	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
6	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
7	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
8	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
9	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
10	LR1	LR1	LR4	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
11	LR1	LR2	LR3	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
12	LR1	LR2	LR3	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
13	LR1	LR2	LR3	LR4	LR4	LR4	LR4	LR4	B	LR8	LR9	LR9
14	LR1	-	-	-	-	-	-	-	-	-	-	-
15	LR1	-	-	-	-	-	-	-	-	-	-	-
16	LR1	-	-	-	-	-	-	-	-	-	-	-
17	LR1	-	-	-	-	-	-	-	-	-	-	-
18	LR1	-	-	-	-	-	-	-	-	-	-	-
19	LR1	-	-	-	-	-	-	-	-	-	-	-
20	LR1	-	-	-	-	-	-	-	-	-	-	-
22	LR1	-	-	-	-	-	-	-	-	-	-	-

Tabelle C.1: Konkordanztabelle zu den Abbildungen der Monte-Carlo Experimente für die Likelihood Ratio Tests (LR) sowie den Alexandersson Test (A) und den Buishand Test (B). Für jeden Test wurde derselbe Satz von Abbildungen generiert. Vollständig ist er jedoch nur für den LR1 Test wiedergegeben. Einige der Abbildungen erwiesen sich als identisch mit den korrespondierenden Abbildungen eines anderen Tests. So ist beispielsweise die Abbildung 3 des LR2 Tests, welche die Verteilung der Epochen der Maxima der LR2 Testgrösse mit $\phi = 0$ zeigen sollte, identisch mit der Abbildung 3 des LR1 Tests. Im Anhang wurde daher die Abbildung 3 des LR2 Tests nicht wiedergegeben. Ein - bedeutet, dass diejenige Abbildung weggelassen wurde obwohl sie zu keiner der vorhandenen identisch ist.

C.1 Likelihood Ratio Test 1

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.5178	2.5201	2.5226	2.7910	2.7939	2.7971	3.3372	3.3429	3.3495
100	0	2.6003	2.6027	2.6051	2.8693	2.8725	2.8755	3.4048	3.4106	3.4164
250	0	2.6906	2.6929	2.6953	2.9553	2.9584	2.9614	3.4864	3.4921	3.4988
1000	0	2.8018	2.8041	2.8064	3.0620	3.0652	3.0683	3.5725	3.5787	3.5839
50	0.1	2.7267	2.7294	2.7322	3.0302	3.0336	3.0372	3.6284	3.6360	3.6428
100	0.1	2.8280	2.8307	2.8336	3.1275	3.1309	3.1344	3.7168	3.7235	3.7309
250	0.1	2.9396	2.9425	2.9450	3.2335	3.2367	3.2398	3.8134	3.8194	3.8262
1000	0.1	3.0721	3.0747	3.0774	3.3597	3.3631	3.3664	3.9366	3.9430	3.9492
50	0.5	4.4409	4.4457	4.4504	4.9919	4.9982	5.0046	6.0777	6.0898	6.1020
100	0.5	4.7082	4.7130	4.7181	5.2551	5.2614	5.2680	6.3329	6.3441	6.3571
250	0.5	4.9850	4.9895	4.9944	5.5258	5.5321	5.5391	6.5913	6.6032	6.6147
1000	0.5	5.3034	5.3085	5.3134	5.8330	5.8391	5.8450	6.8910	6.9020	6.9128
50	0.9	16.376	16.398	16.421	18.941	18.969	18.997	24.010	24.064	24.120
100	0.9	18.955	18.979	19.002	21.642	21.673	21.704	26.952	27.012	27.077
250	0.9	21.496	21.520	21.545	24.213	24.246	24.280	29.601	29.654	29.713
1000	0.9	24.142	24.166	24.190	26.863	26.894	26.924	32.228	32.286	32.350

Tabelle C.2: Kritische Grenzen des LR1 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	1.9595	2.1864	2.2899	2.3686	2.4069	2.4567	2.4835	2.5146	2.5314
10	2.5552	2.5792	2.5900	2.6059	2.6129	2.6309	2.6308	2.6476	2.6542	2.6629
20	2.6822	2.6686	2.6786	2.6826	2.6894	2.6972	2.7083	2.7259	2.7270	2.7194
30	2.7298	2.7233	2.7398	2.7405	2.7478	2.7539	2.7484	2.7631	2.7657	2.7644
40	2.7725	2.7723	2.7716	2.7798	2.7764	2.7797	2.7972	2.7923	2.7913	2.7915
50	2.7926	2.7958	2.8035	2.8003	2.8053	2.8101	2.8095	2.8114	2.8199	2.8160
60	2.8163	2.8161	2.8241	2.8239	2.8300	2.8217	2.8329	2.8288	2.8254	2.8435
70	2.8311	2.8363	2.8357	2.8424	2.8523	2.8441	2.8454	2.8445	2.8448	2.8596
80	2.8546	2.8427	2.8641	2.8452	2.8592	2.8597	2.8601	2.8565	2.8670	2.8665
90	2.8593	2.8640	2.8699	2.8636	2.8669	2.8713	2.8834	2.8708	2.8592	2.8774
100	2.8676	2.8796	2.8738	2.8698	2.8771	2.8784	2.8768	2.8738	2.8843	2.8865
110	2.8916	2.8824	2.8864	2.8841	2.8822	2.8918	2.8868	2.8815	2.8886	2.8954
120	2.8965	2.8957	2.8876	2.8910	2.9018	2.8897	2.8984	2.9057	2.9056	2.8959
130	2.8953	2.9026	2.9045	2.9029	2.9163	2.9155	2.9043	2.9058	2.9165	2.9052
140	2.9090	2.9108	2.9095	2.9091	2.9049	2.9153	2.9082	2.9109	2.9137	2.9224
150	2.9042	2.9112	2.9046	2.9064	2.9150	2.9210	2.9199	2.9248	2.9227	2.9170
160	2.9148	2.9200	2.9237	2.9264	2.9249	2.9179	2.9241	2.9156	2.9252	2.9235
170	2.9237	2.9257	2.9299	2.9323	2.9319	2.9326	2.9301	2.9280	2.9303	2.9328
180	2.9319	2.9389	2.9320	2.9314	2.9301	2.9369	2.9338	2.9305	2.9373	2.9315
190	2.9382	2.9327	2.9416	2.9466	2.9408	2.9295	2.9445	2.9358	2.9362	2.9456

Tabelle C.3: Kritische Grenzen T_{95} des LR1 Tests für 100’000 Replikationen.

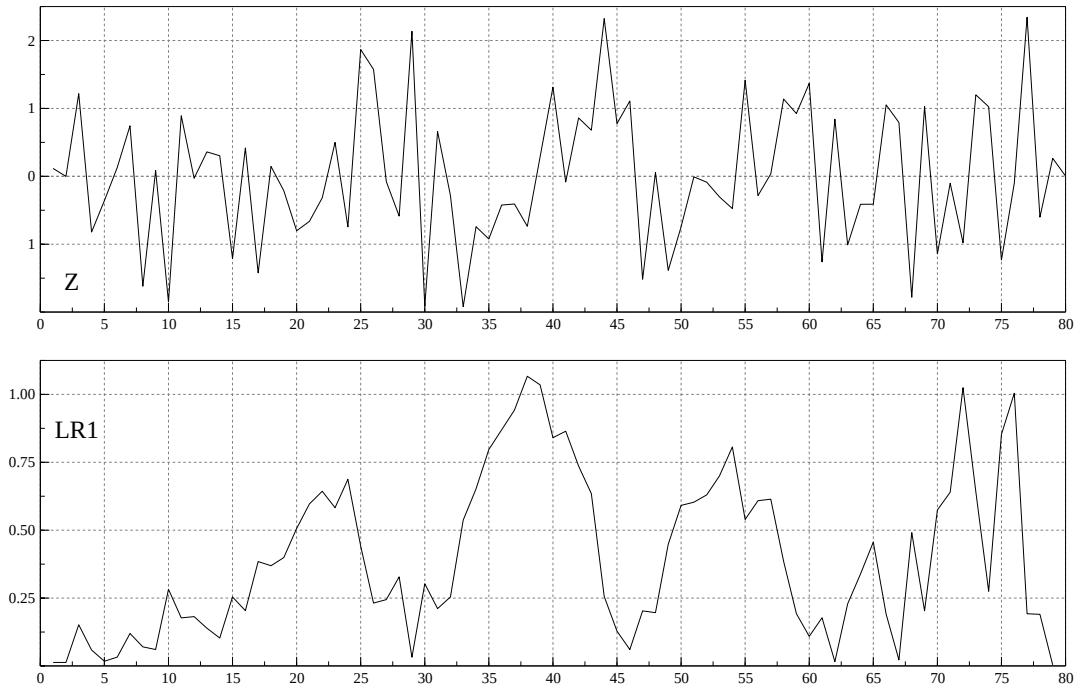


Abbildung C.1: Verlauf der LR1 Testgrösse für die Testreihe Z80

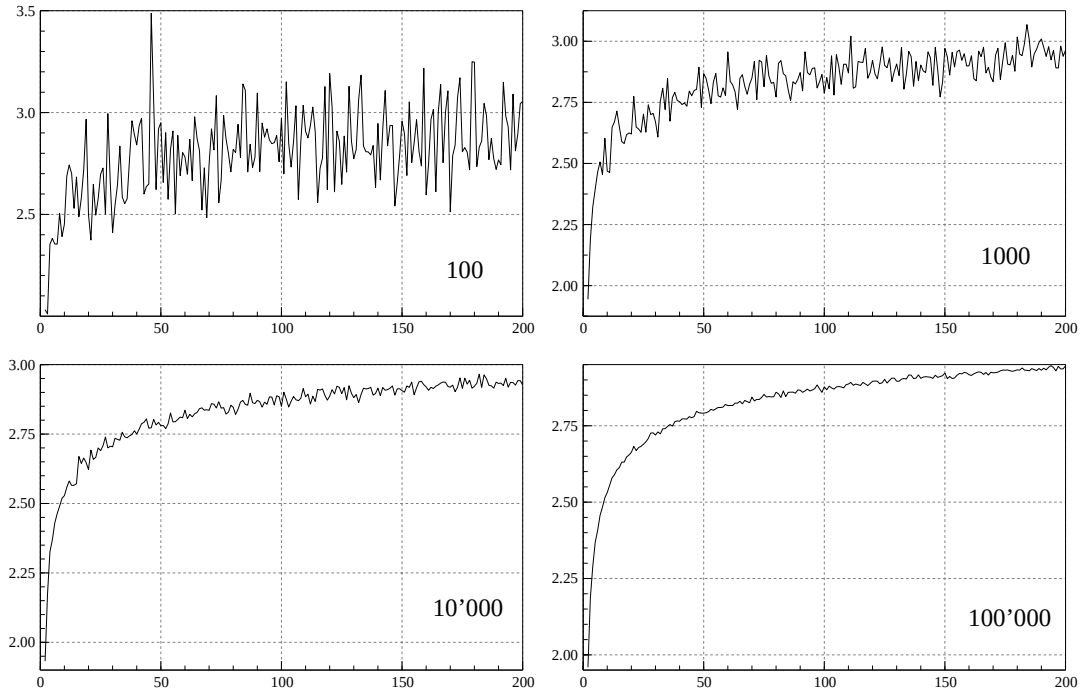


Abbildung C.2: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

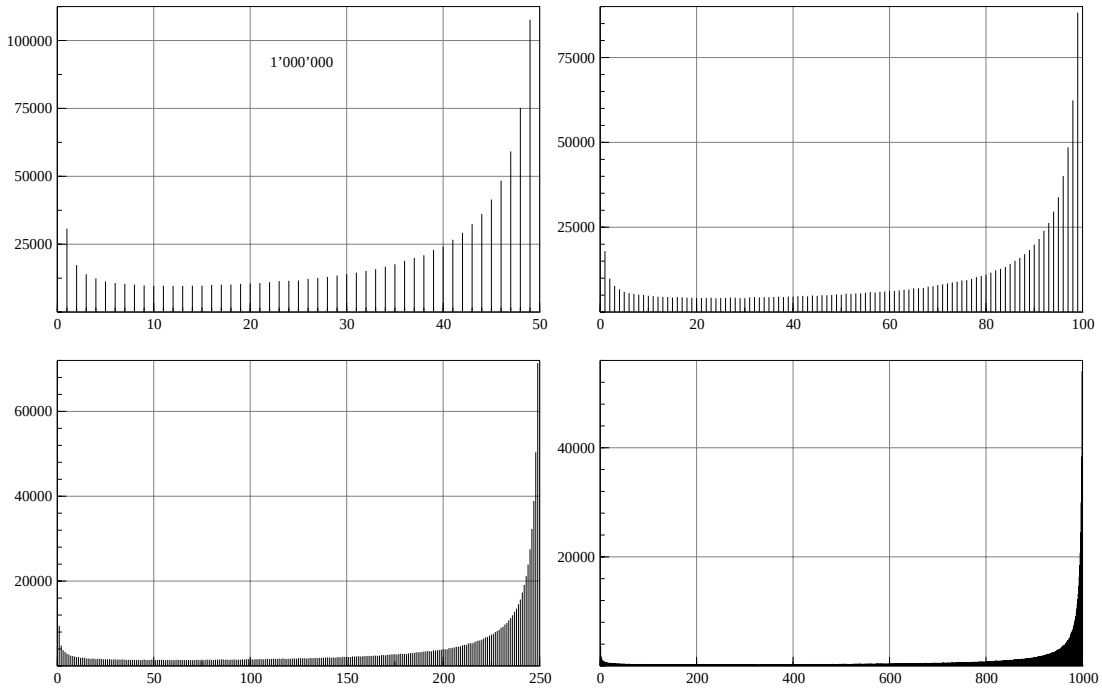


Abbildung C.3: Verteilung der Epochen der Maxima der LR1 Testgrösse mit $\phi = 0$

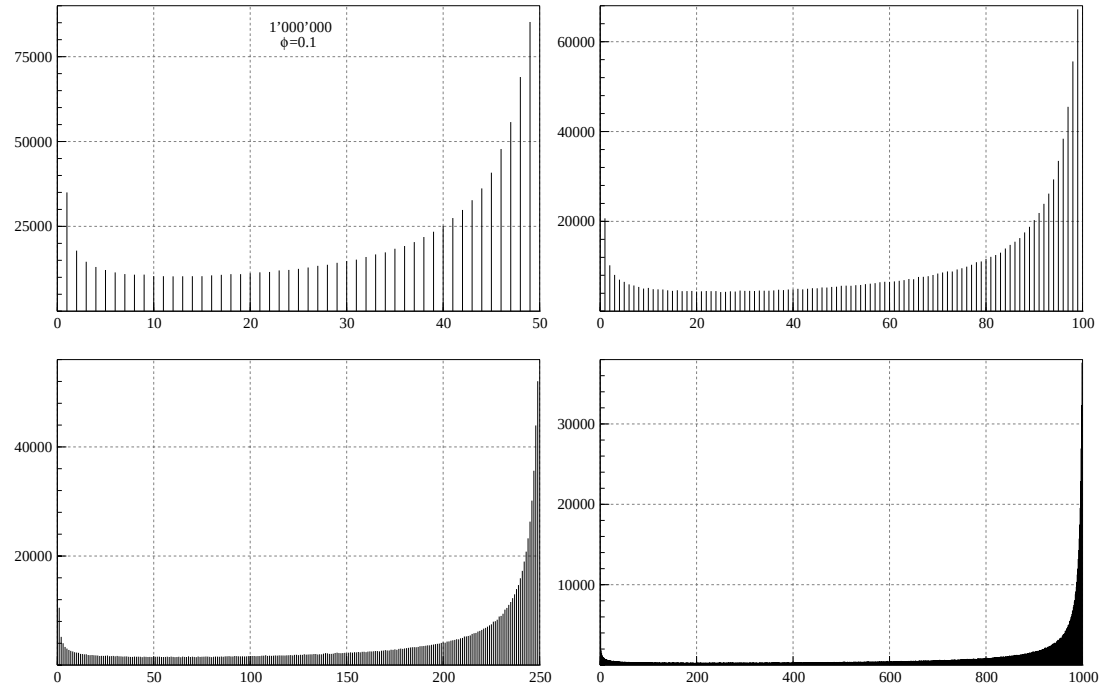


Abbildung C.4: Verteilung der Epochen der Maxima der LR1 Testgrösse mit $\phi = 0.1$

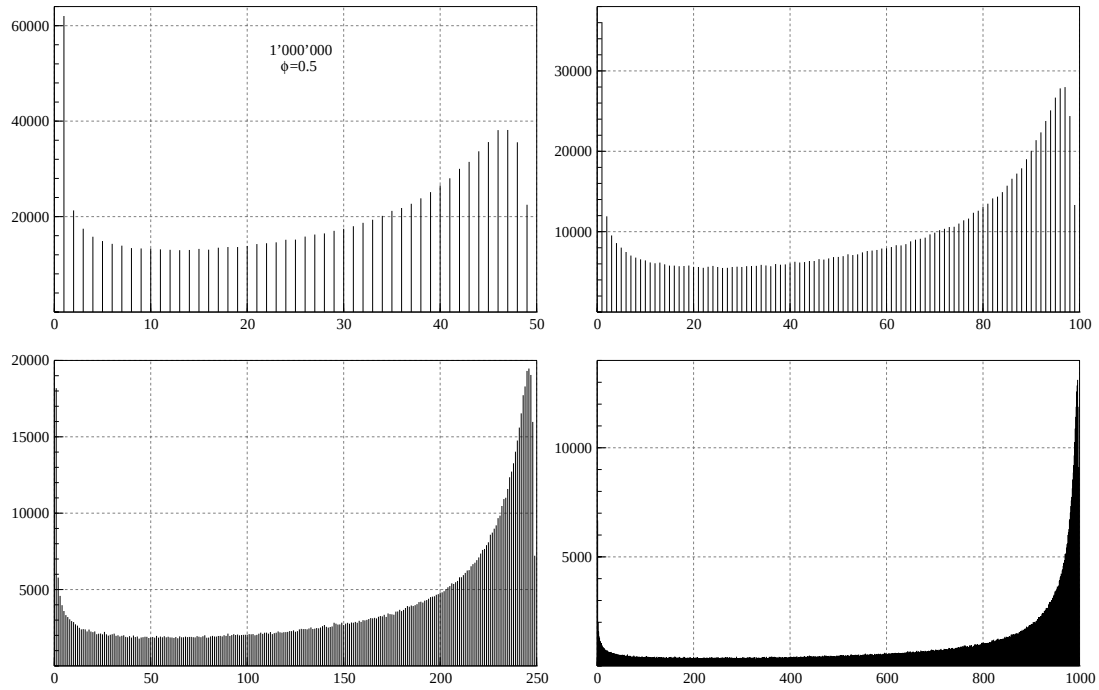


Abbildung C.5: Verteilung der Epochen der Maxima der LR1 Testgrösse mit $\phi = 0.5$

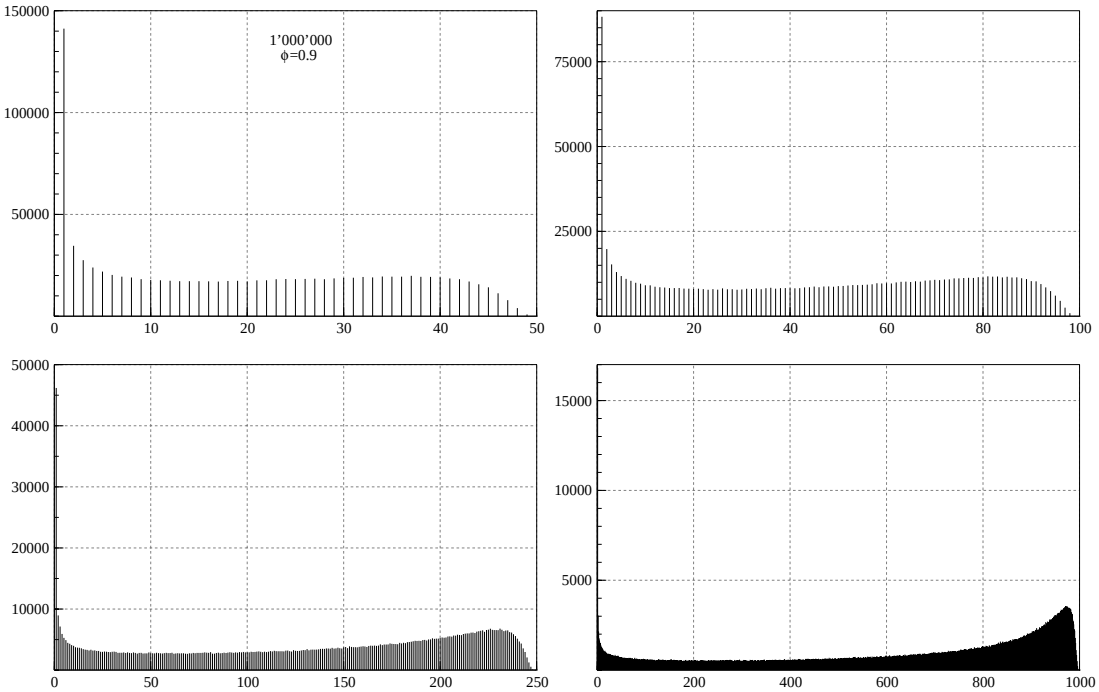


Abbildung C.6: Verteilung der Epochen der Maxima der LR1 Testgrösse mit $\phi = 0.9$

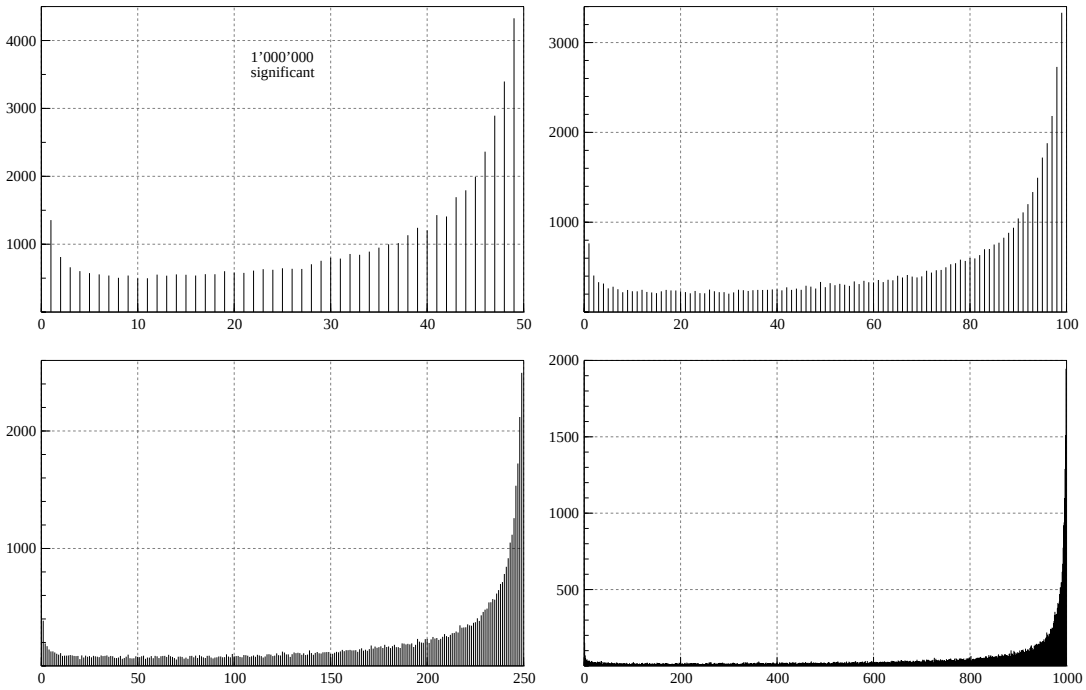


Abbildung C.7: Verteilung der Bruchepochen der LR1 Testgrösse mit $\phi = 0$

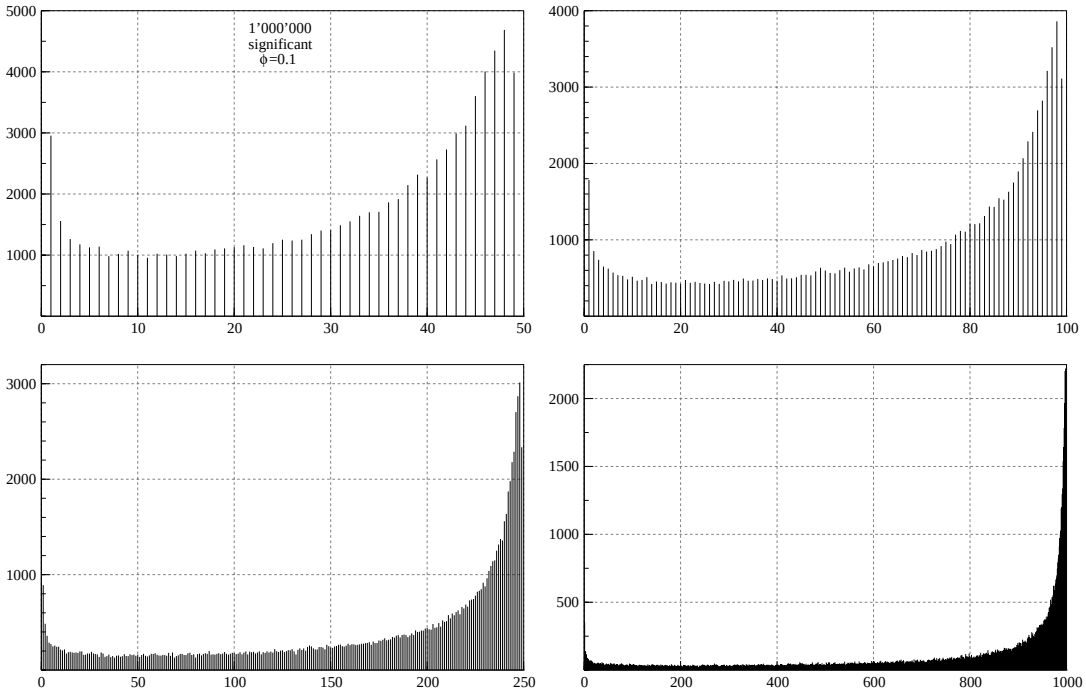


Abbildung C.8: Verteilung der Bruchepochen der LR1 Testgrösse mit $\phi = 0.1$

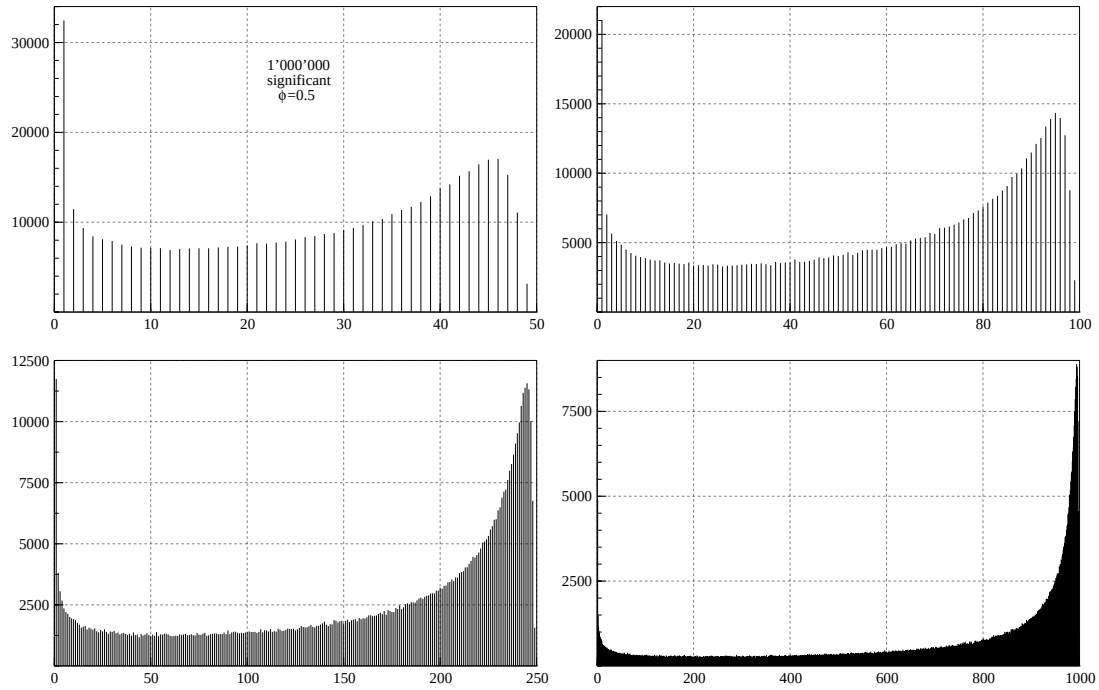


Abbildung C.9: Verteilung der Bruchepochen der LR1 Testgrösse mit $\phi = 0.5$

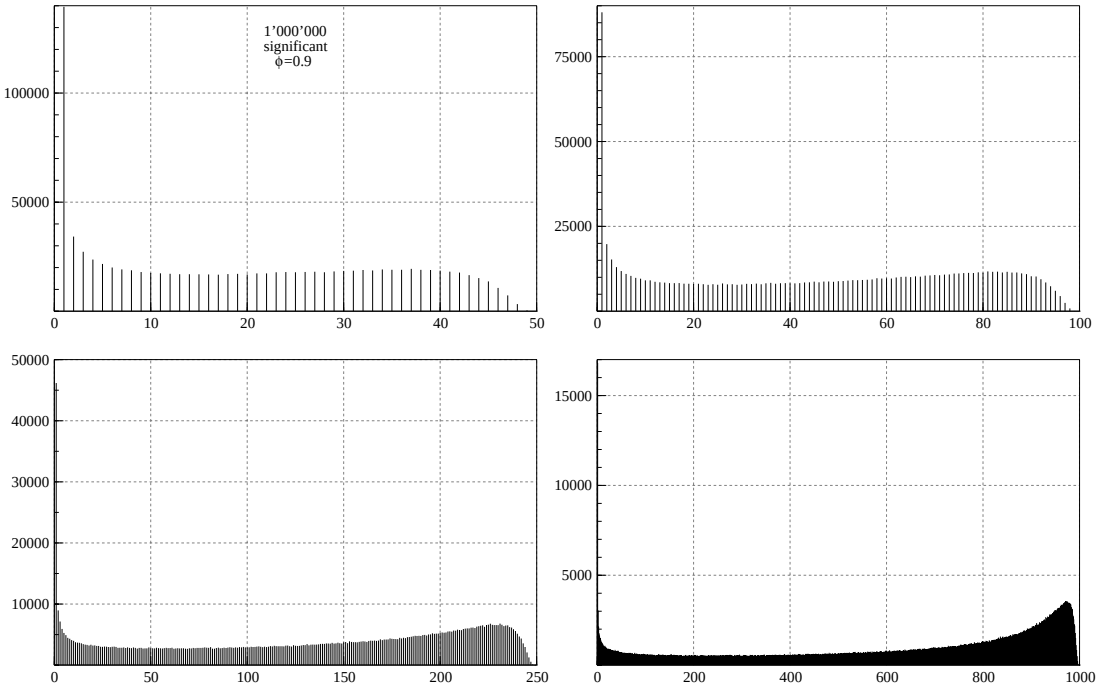


Abbildung C.10: Verteilung der Bruchepochen der LR1 Testgrösse mit $\phi = 0.9$

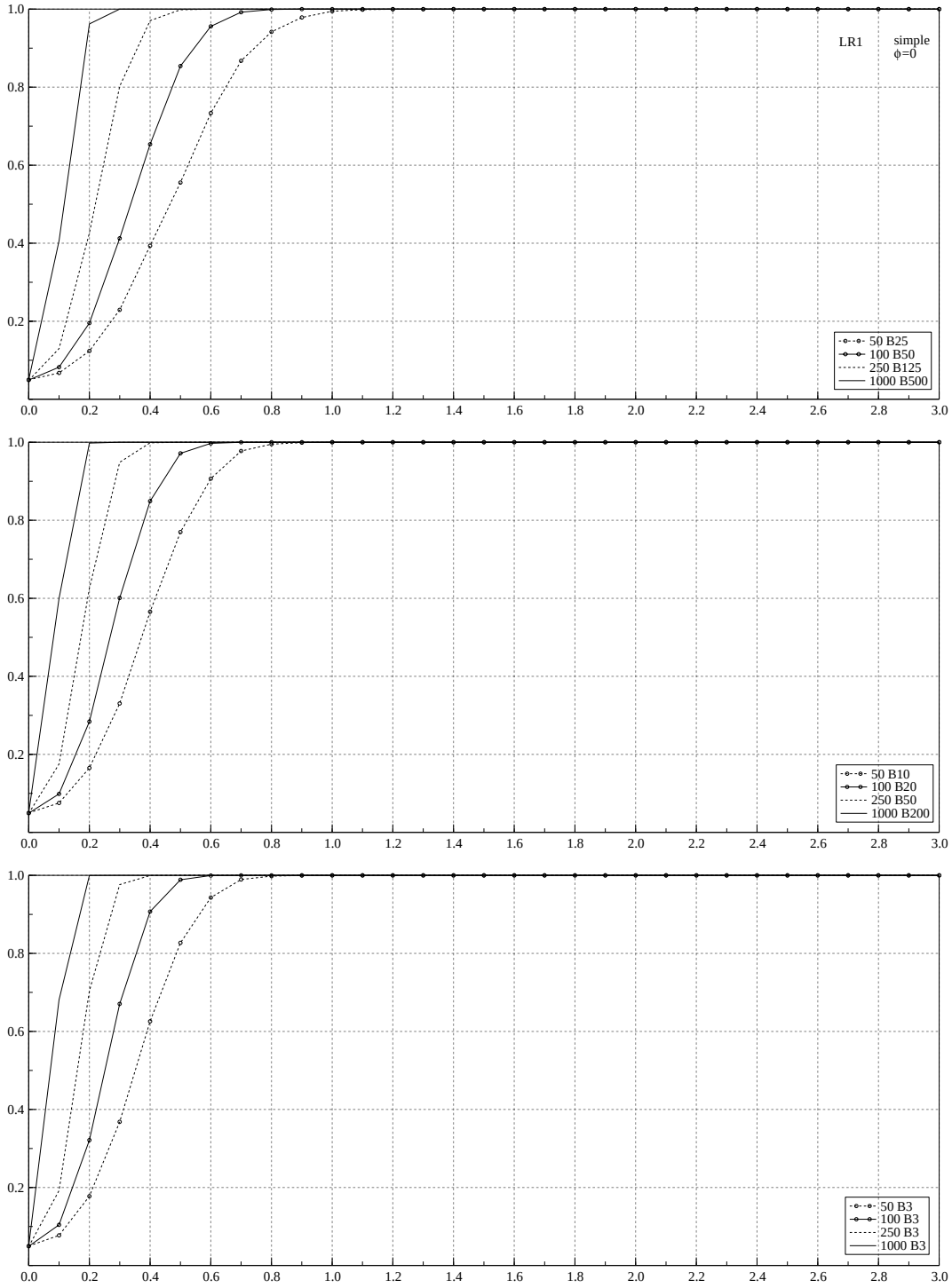


Abbildung C.11: Güte des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

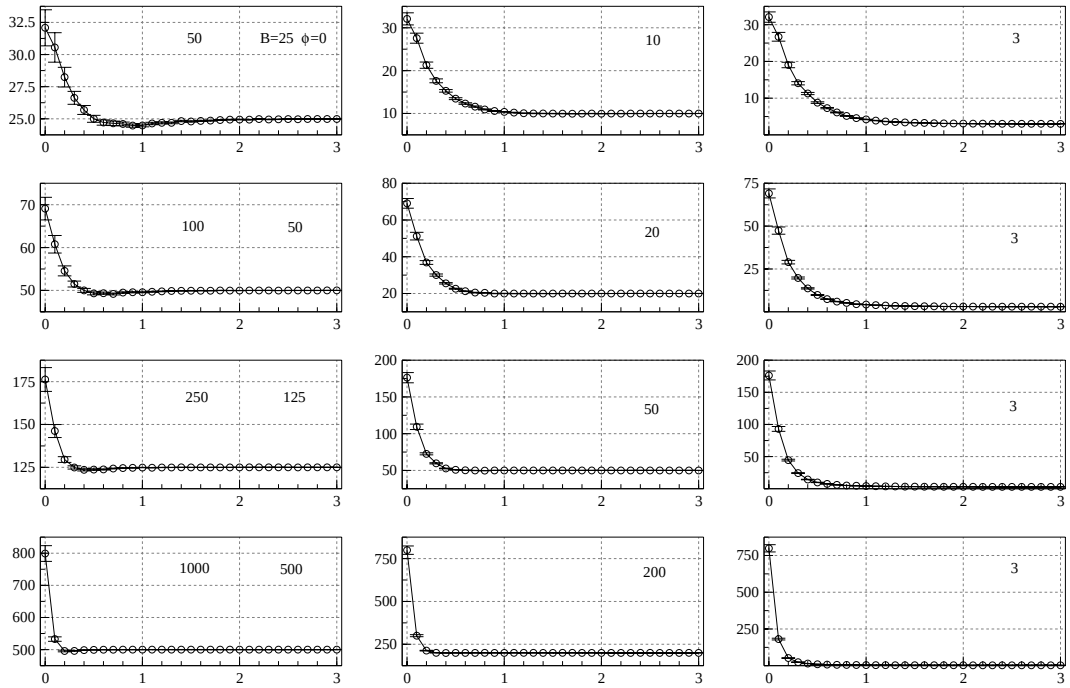


Abbildung C.12: Epochentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

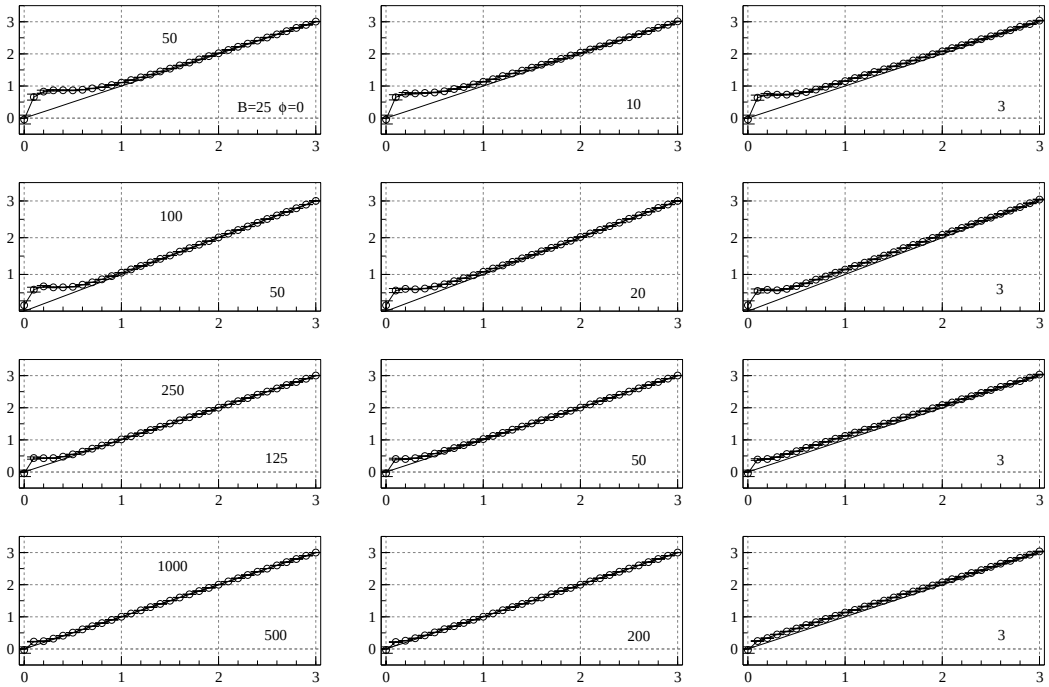


Abbildung C.13: Sprunghöhentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

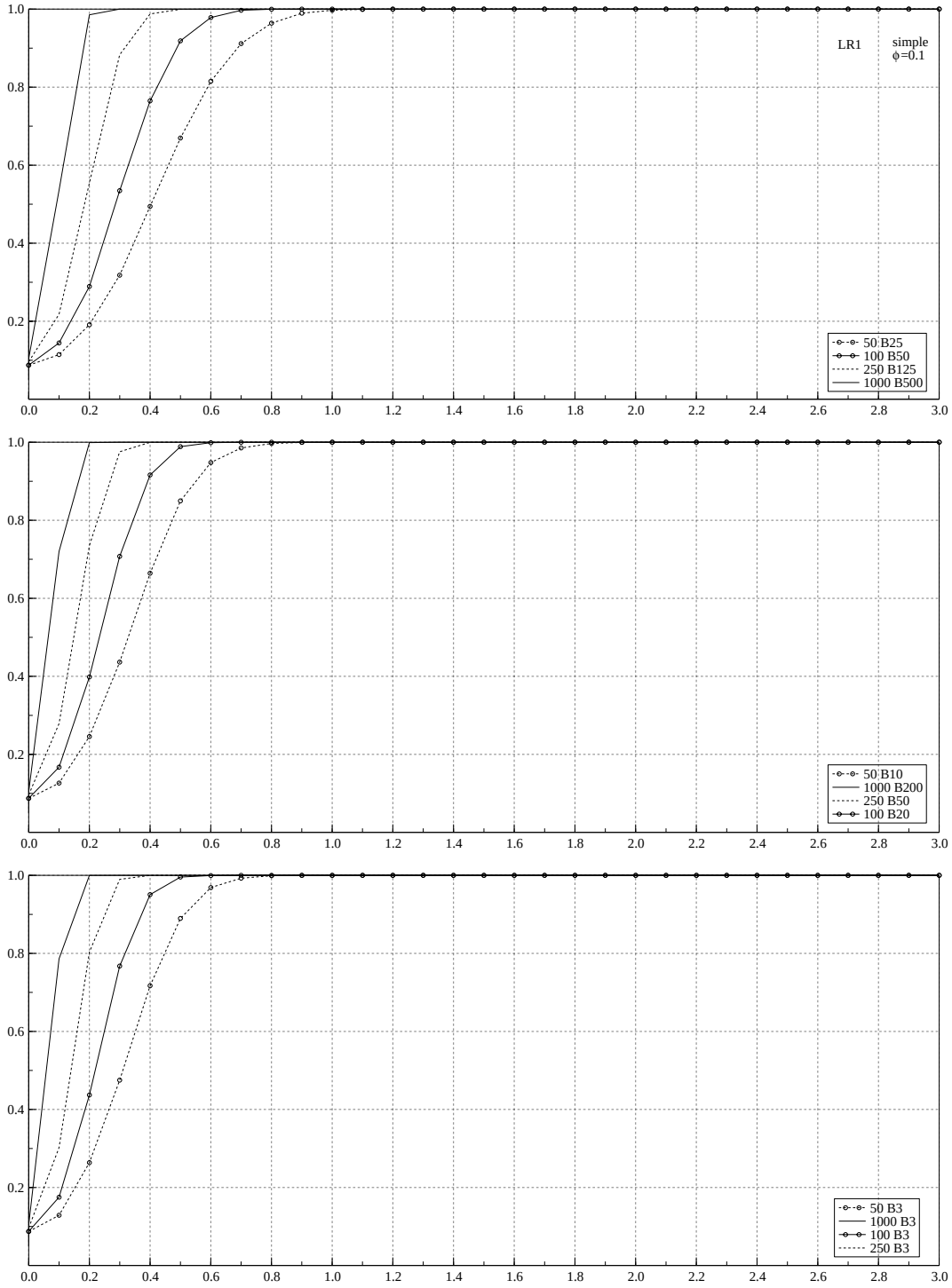


Abbildung C.14: Güte des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.1$ und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

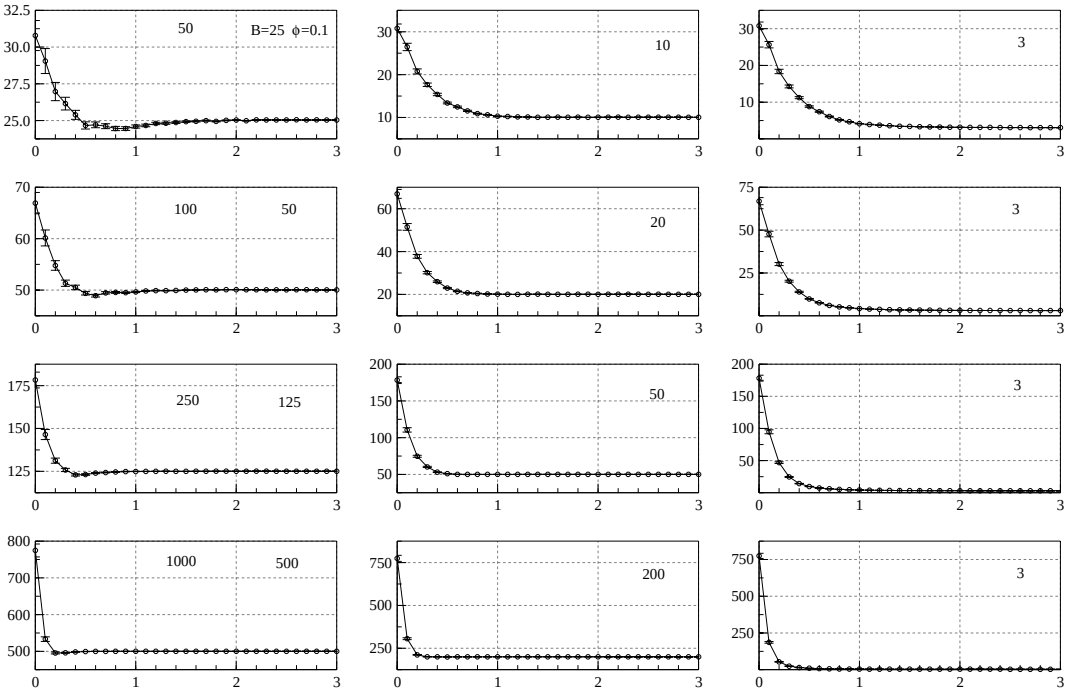


Abbildung C.15: Epochentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.1$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

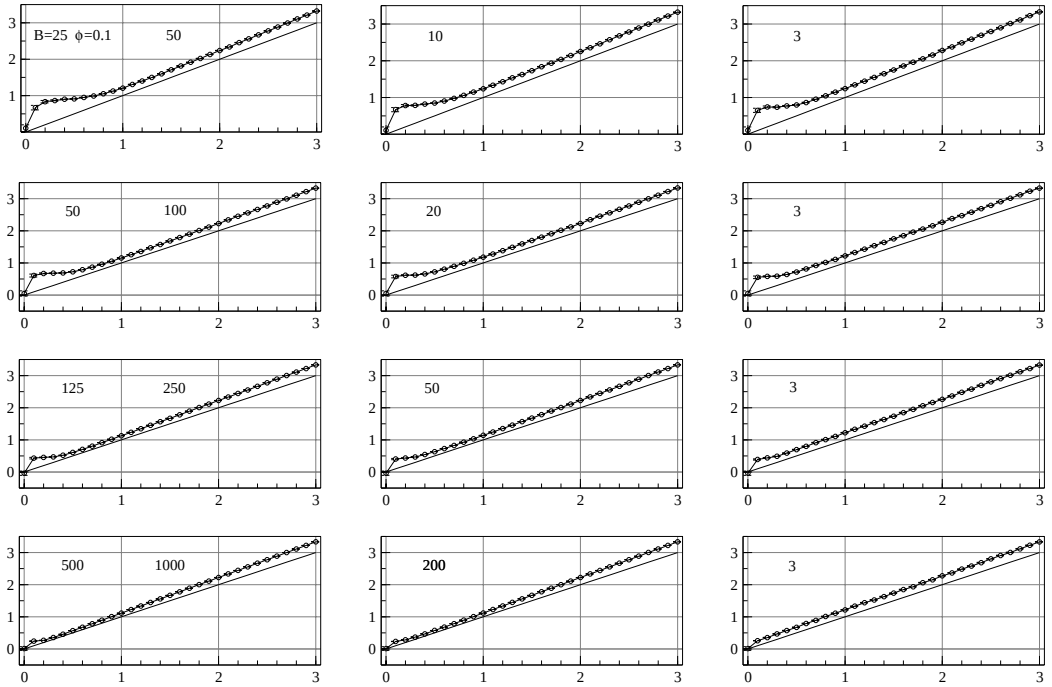


Abbildung C.16: Sprunghöhentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.1$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

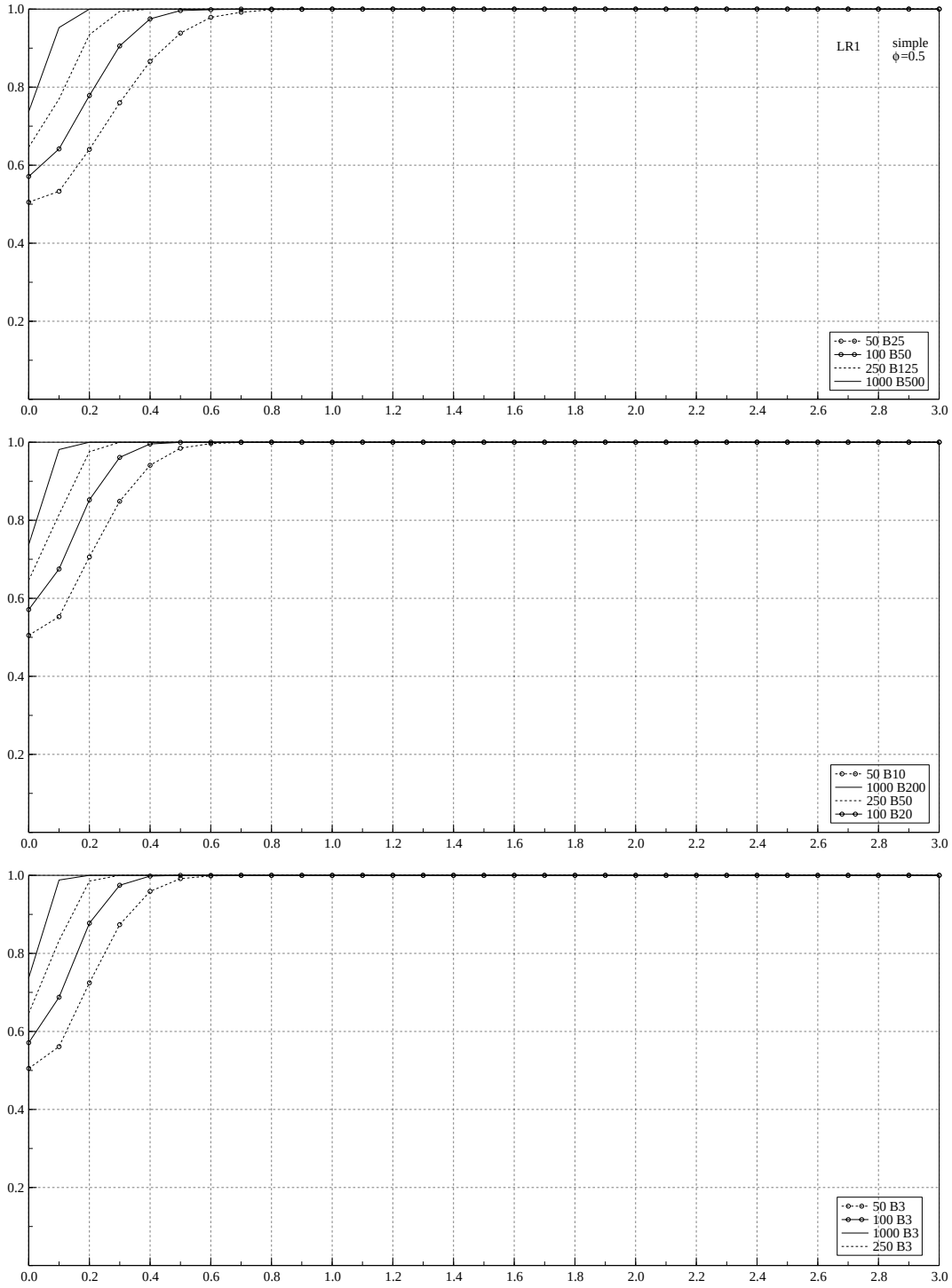


Abbildung C.17: Güte des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.5$ und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

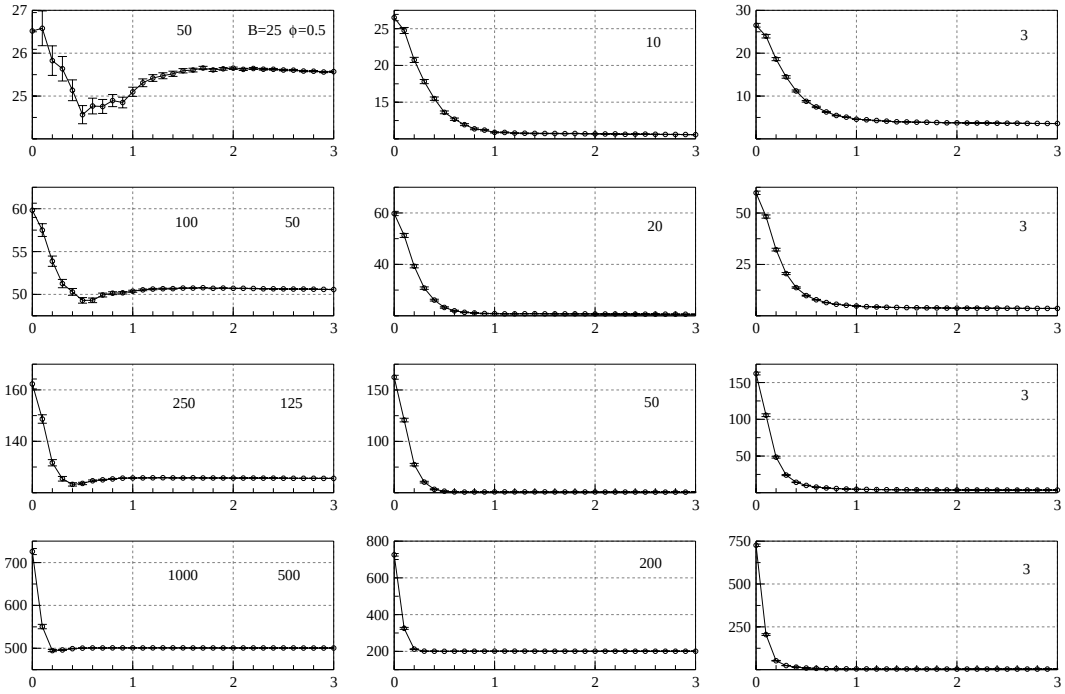


Abbildung C.18: Epochentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.5$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

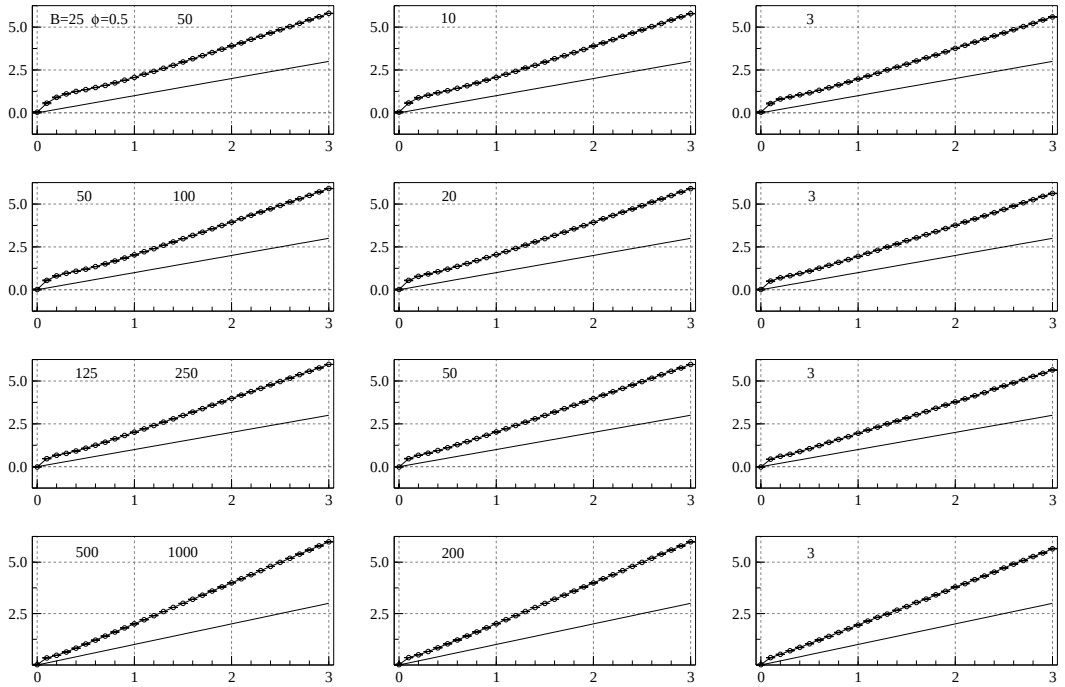


Abbildung C.19: Sprunghöhentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.5$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

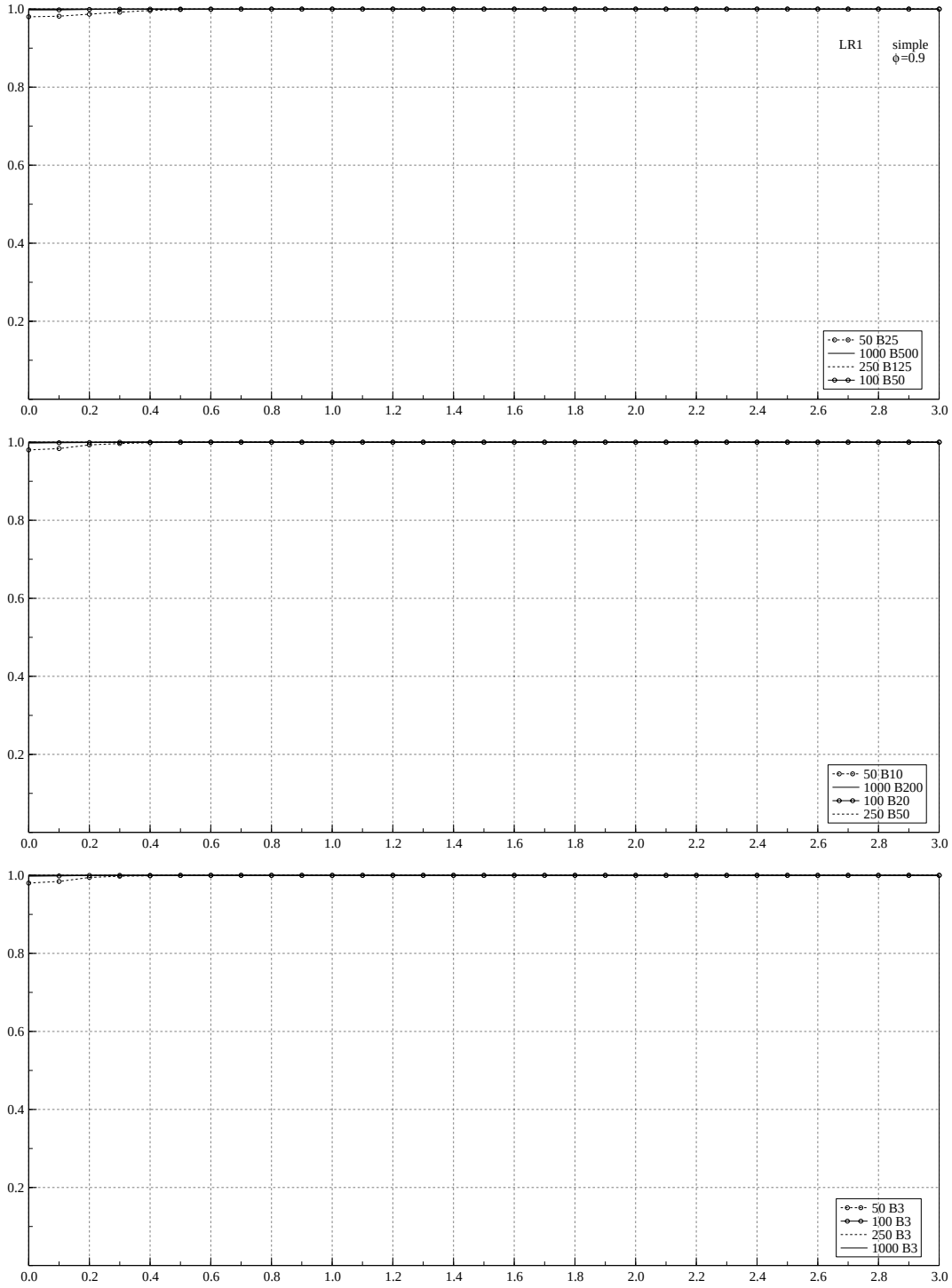


Abbildung C.20: Güte des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.9$ und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

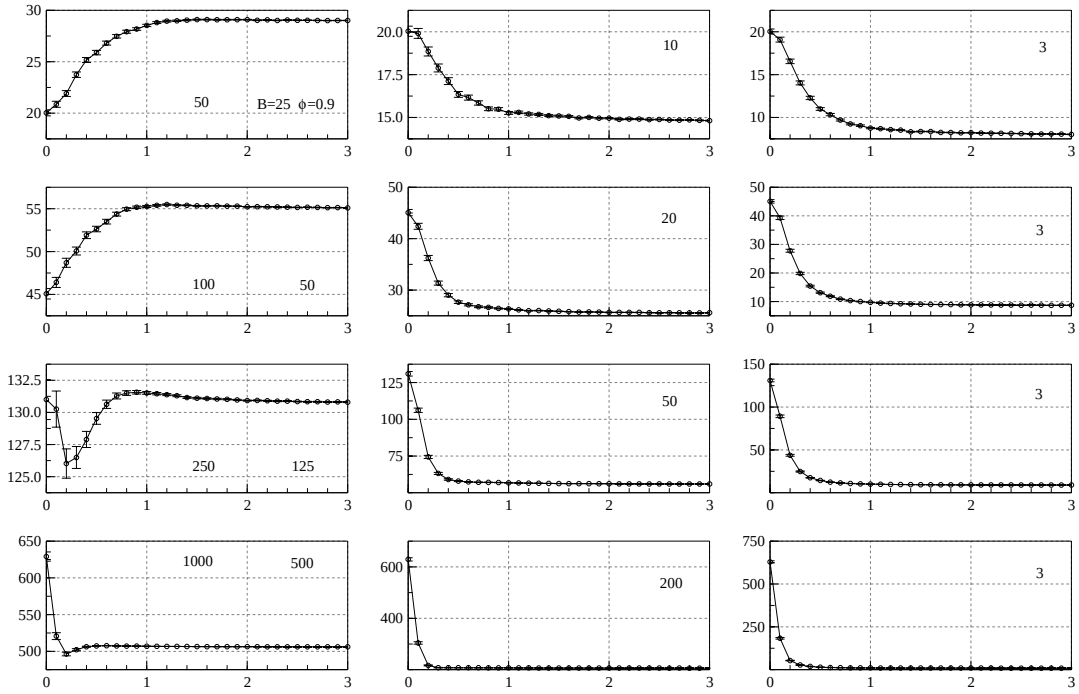


Abbildung C.21: Epochentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.9$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

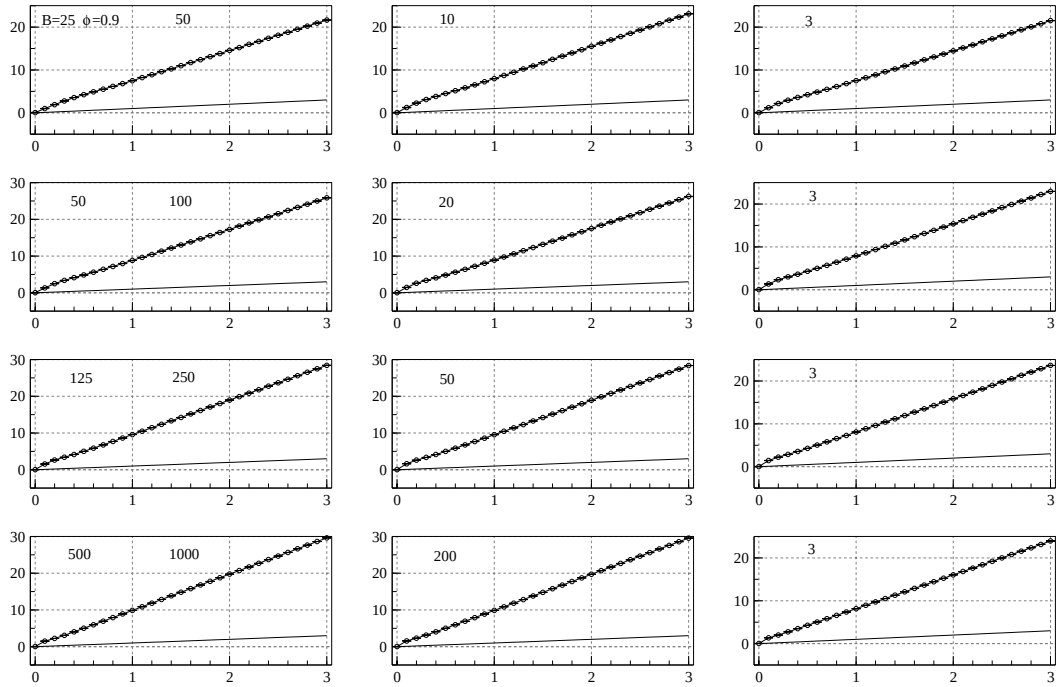


Abbildung C.22: Sprunghöhentreue des LR1 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 mit $\phi = 0.9$ und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.2 Likelihood Ratio Test 2

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.6384	2.6412	2.6439	2.9476	2.9513	2.9549	3.5786	3.5858	3.5920
100	0	2.6640	2.6666	2.6692	2.9489	2.9523	2.9558	3.5296	3.5359	3.5430
250	0	2.7194	2.7220	2.7246	2.9908	2.9937	2.9967	3.5383	3.5437	3.5496
1000	0	2.8095	2.8116	2.8142	3.0717	3.0747	3.0775	3.5875	3.5929	3.5987
50	0.1	2.8565	2.8596	2.8626	3.2001	3.2041	3.2084	3.9027	3.9105	3.9180
100	0.1	2.8903	2.8931	2.8957	3.2090	3.2130	3.2170	3.8537	3.8609	3.8686
250	0.1	2.9590	2.9615	2.9641	3.2600	3.2637	3.2672	3.8597	3.8657	3.8732
1000	0.1	3.0664	3.0689	3.0714	3.3548	3.3581	3.3613	3.9344	3.9408	3.9479
50	0.5	4.2582	4.2634	4.2686	4.8407	4.8472	4.8542	6.0489	6.0631	6.0768
100	0.5	4.3073	4.3123	4.3169	4.8367	4.8429	4.8489	5.9133	5.9264	5.9378
250	0.5	4.4223	4.4267	4.4313	4.9135	4.9190	4.9250	5.8916	5.9029	5.9141
1000	0.5	4.6240	4.6277	4.6319	5.0883	5.0937	5.0991	6.0185	6.0290	6.0407
50	0.9	11.459	11.477	11.496	13.476	13.501	13.526	17.679	17.726	17.773
100	0.9	11.011	11.026	11.044	12.834	12.855	12.877	16.630	16.671	16.719
250	0.9	10.708	10.721	10.735	12.197	12.215	12.234	15.305	15.338	15.372
1000	0.9	10.941	10.952	10.963	12.213	12.227	12.242	14.768	14.796	14.823

Tabelle C.4: Kritische Grenzen des LR2 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	12.5471	6.2066	4.8429	4.2156	3.8501	3.6516	3.4826	3.4127	3.3060
10	3.2682	3.2236	3.1700	3.1464	3.1139	3.1032	3.0771	3.0574	3.0540	3.0452
20	3.0353	3.0125	3.0048	2.9998	3.0007	2.9861	2.9920	2.9860	2.9815	2.9684
30	2.9759	2.9644	2.9730	2.9680	2.9725	2.9786	2.9635	2.9628	2.9565	2.9556
40	2.9496	2.9517	2.9520	2.9484	2.9491	2.9427	2.9564	2.9567	2.9543	2.9505
50	2.9443	2.9532	2.9477	2.9471	2.9494	2.9509	2.9502	2.9557	2.9548	2.9436
60	2.9485	2.9468	2.9546	2.9478	2.9602	2.9425	2.9542	2.9414	2.9383	2.9585
70	2.9434	2.9468	2.9458	2.9452	2.9613	2.9486	2.9420	2.9521	2.9461	2.9566
80	2.9554	2.9413	2.9671	2.9357	2.9559	2.9510	2.9478	2.9461	2.9606	2.9478
90	2.9515	2.9508	2.9601	2.9502	2.9478	2.9608	2.9757	2.9578	2.9355	2.9618
100	2.9473	2.9638	2.9539	2.9482	2.9555	2.9567	2.9542	2.9552	2.9582	2.9612
110	2.9710	2.9495	2.9614	2.9527	2.9601	2.9617	2.9560	2.9508	2.9516	2.9668
120	2.9677	2.9676	2.9545	2.9542	2.9665	2.9628	2.9658	2.9763	2.9706	2.9643
130	2.9670	2.9693	2.9643	2.9716	2.9737	2.9709	2.9653	2.9626	2.9754	2.9694
140	2.9684	2.9636	2.9685	2.9679	2.9655	2.9707	2.9711	2.9676	2.9636	2.9771
150	2.9617	2.9627	2.9615	2.9614	2.9716	2.9737	2.9750	2.9795	2.9722	2.9722
160	2.9683	2.9734	2.9761	2.9705	2.9708	2.9655	2.9759	2.9666	2.9758	2.9774
170	2.9675	2.9777	2.9756	2.9794	2.9792	2.9834	2.9820	2.9779	2.9795	2.9774
180	2.9751	2.9845	2.9786	2.9803	2.9810	2.9850	2.9764	2.9773	2.9753	2.9782
190	2.9837	2.9802	2.9853	2.9876	2.9888	2.9741	2.9848	2.9825	2.9819	2.9866

Tabelle C.5: Kritische Grenzen T_{95} des LR2 Tests für 100’000 Replikationen.

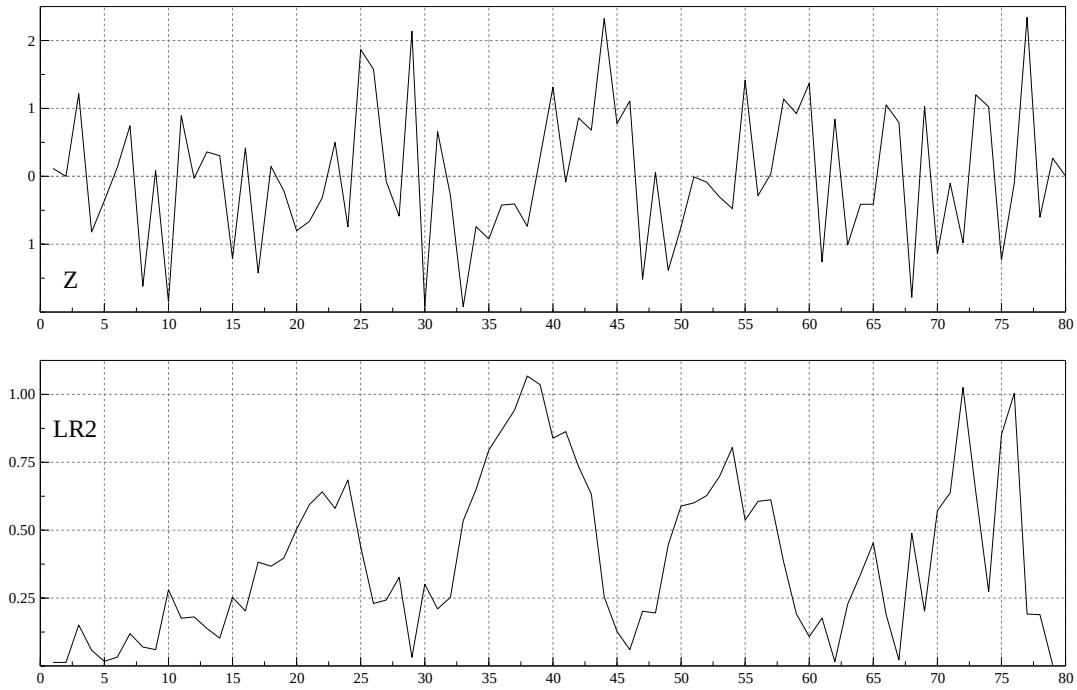


Abbildung C.23: Verlauf der LR2 Testgrösse für die Testreihe Z80

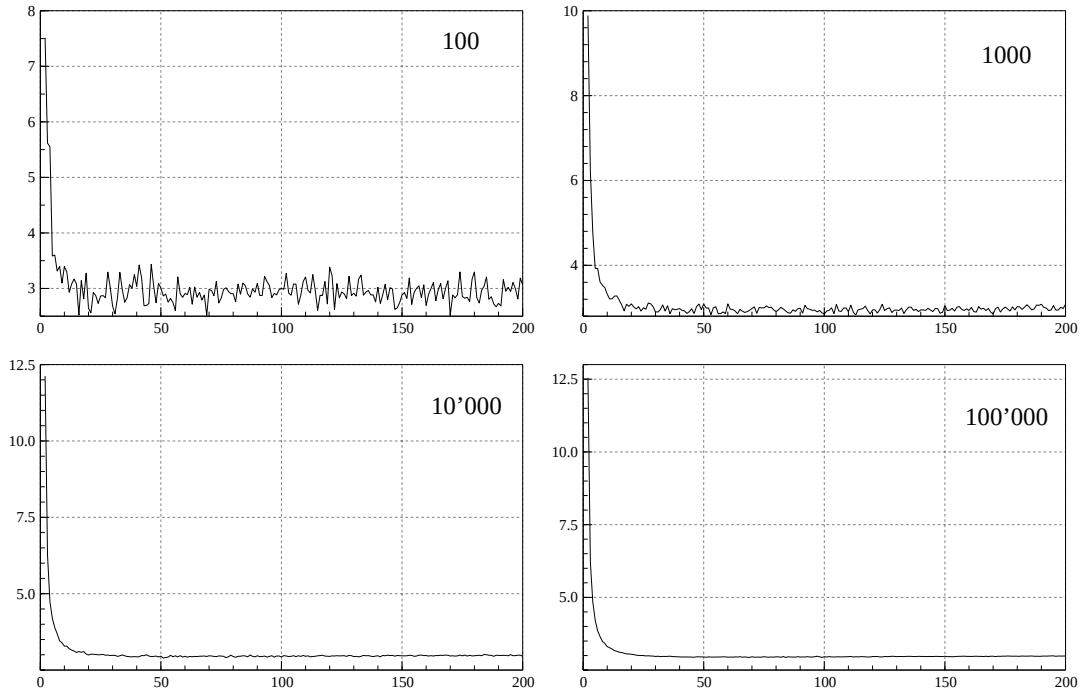


Abbildung C.24: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

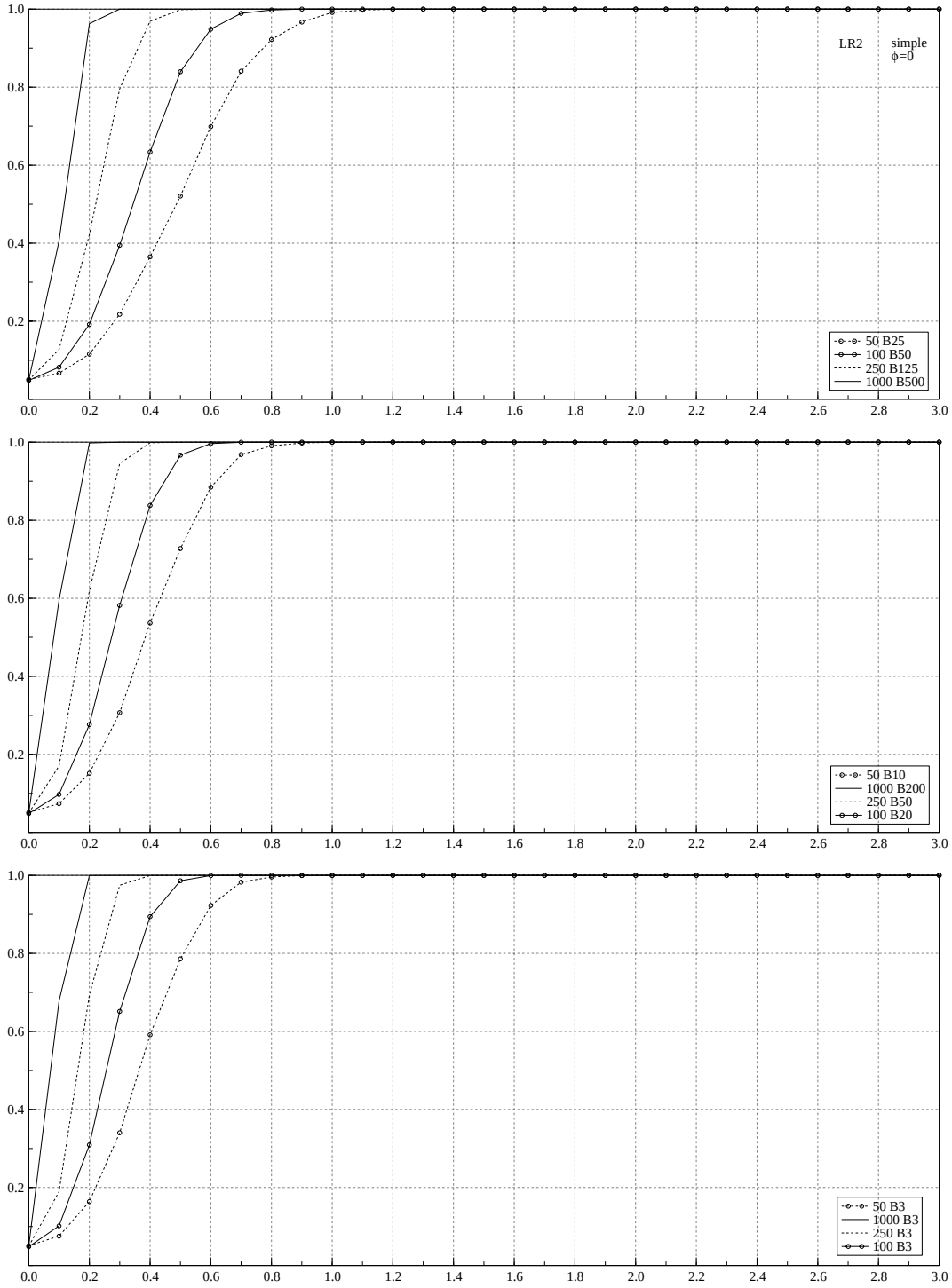


Abbildung C.25: Güte des LR2 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

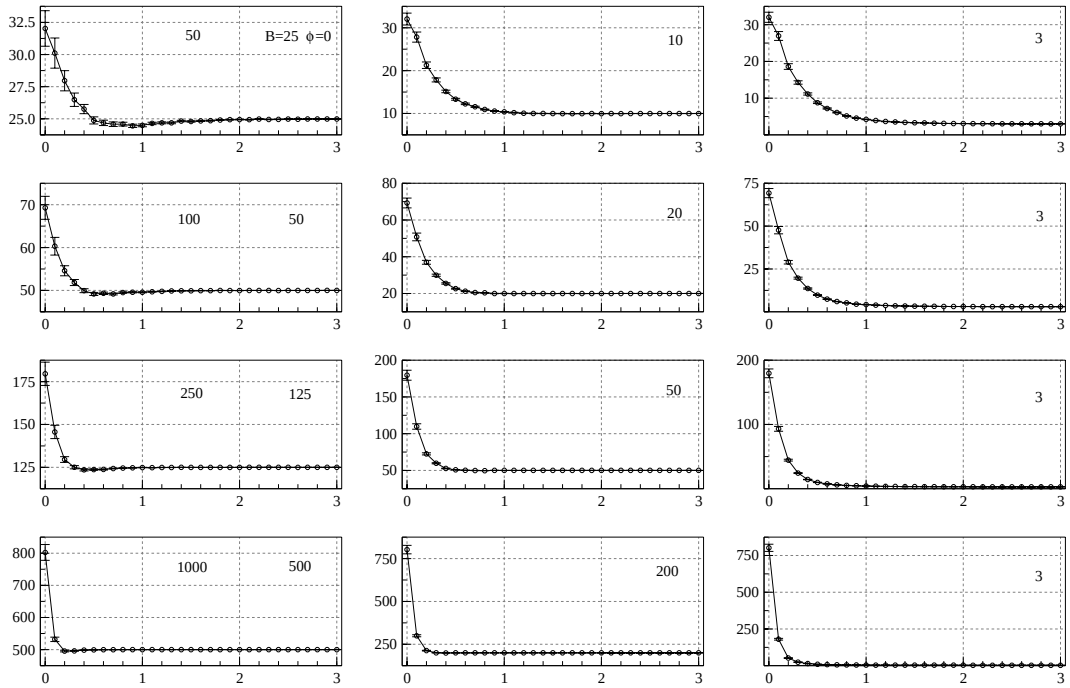


Abbildung C.26: Epochentreue des LR2 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

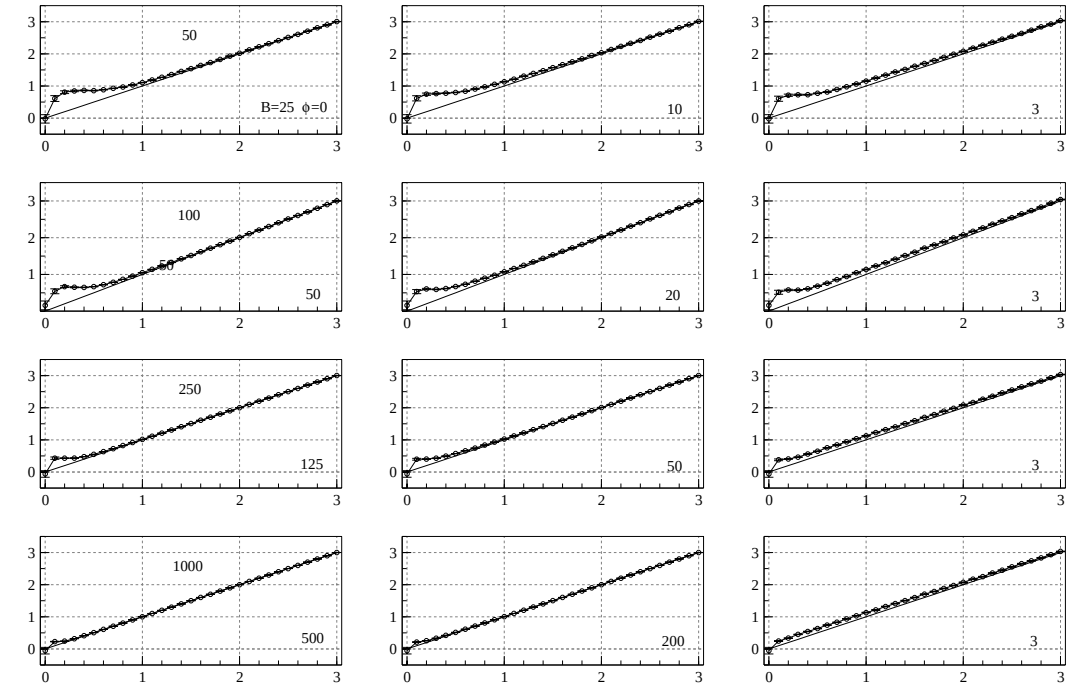


Abbildung C.27: Sprunghöhentreue des LR2 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.3 Likelihood Ratio Test 3

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.7103	2.7128	2.7150	2.9655	2.9684	2.9716	3.4805	3.4857	3.4911
100	0	2.8079	2.8101	2.8125	3.0625	3.0656	3.0687	3.5709	3.5773	3.5829
250	0	2.9082	2.9103	2.9127	3.1565	3.1591	3.1618	3.6557	3.6610	3.6662
1000	0	3.0287	3.0309	3.0330	3.2721	3.2748	3.2775	3.7549	3.7599	3.7647
50	0.1	2.9199	2.9226	2.9249	3.2033	3.2068	3.2101	3.7677	3.7737	3.7804
100	0.1	3.0451	3.0476	3.0501	3.3266	3.3297	3.3330	3.8874	3.8939	3.9001
250	0.1	3.1766	3.1791	3.1818	3.4533	3.4562	3.4593	4.0028	4.0091	4.0155
1000	0.1	3.3215	3.3238	3.3262	3.5897	3.5928	3.5957	4.1362	4.1423	4.1485
50	0.5	4.6099	4.6145	4.6191	5.1184	5.1245	5.1303	6.1295	6.1401	6.1518
100	0.5	4.9821	4.9866	4.9910	5.4960	5.5020	5.5078	6.5133	6.5252	6.5353
250	0.5	5.3509	5.3549	5.3595	5.8588	5.8646	5.8703	6.8666	6.8774	6.8895
1000	0.5	5.7263	5.7306	5.7352	6.2208	6.2270	6.2331	7.2198	7.2325	7.2428
50	0.9	13.106	13.122	13.137	14.963	14.984	15.005	18.705	18.742	18.782
100	0.9	17.431	17.453	17.472	19.731	19.759	19.786	24.283	24.329	24.379
250	0.9	21.707	21.728	21.750	24.202	24.233	24.263	29.259	29.310	29.365
1000	0.9	25.591	25.614	25.638	28.118	28.146	28.174	33.200	33.262	33.320

Tabelle C.6: Kritische Grenzen des LR3 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	1.9586	2.2214	2.3406	2.4332	2.5004	2.5544	2.5858	2.6163	2.6452
10	2.6817	2.7002	2.7079	2.7332	2.7455	2.7645	2.7793	2.7893	2.8038	2.8140
20	2.8296	2.8207	2.8437	2.8424	2.8536	2.8604	2.8690	2.8828	2.8892	2.8893
30	2.8993	2.8937	2.9085	2.9095	2.9169	2.9187	2.9190	2.9323	2.9373	2.9396
40	2.9388	2.9503	2.9499	2.9471	2.9573	2.9572	2.9618	2.9678	2.9713	2.9606
50	2.9640	2.9753	2.9818	2.9676	2.9816	2.9879	2.9895	3.0007	2.9918	2.9944
60	2.9990	3.0030	3.0078	3.0060	3.0027	3.0110	3.0100	3.0158	3.0135	3.0264
70	3.0163	3.0100	3.0194	3.0380	3.0325	3.0314	3.0272	3.0312	3.0354	3.0423
80	3.0438	3.0381	3.0527	3.0442	3.0378	3.0468	3.0485	3.0498	3.0570	3.0528
90	3.0547	3.0536	3.0565	3.0579	3.0508	3.0618	3.0661	3.0598	3.0542	3.0661
100	3.0638	3.0672	3.0646	3.0613	3.0718	3.0716	3.0709	3.0719	3.0810	3.0700
110	3.0749	3.0809	3.0750	3.0832	3.0772	3.0791	3.0822	3.0814	3.0778	3.0908
120	3.0919	3.0888	3.0840	3.0888	3.0907	3.0900	3.0918	3.0914	3.1017	3.0913
130	3.0925	3.1024	3.1036	3.0985	3.1041	3.0995	3.1009	3.1058	3.1013	3.1000
140	3.1053	3.1059	3.1004	3.1029	3.1085	3.1089	3.1043	3.1047	3.1084	3.1125
150	3.1063	3.1067	3.1046	3.1068	3.1191	3.1180	3.1188	3.1159	3.1226	3.1200
160	3.1180	3.1231	3.1283	3.1181	3.1218	3.1209	3.1168	3.1169	3.1284	3.1286
170	3.1254	3.1246	3.1248	3.1309	3.1296	3.1293	3.1228	3.1228	3.1317	3.1235
180	3.1318	3.1383	3.1367	3.1254	3.1387	3.1232	3.1346	3.1382	3.1375	3.1241
190	3.1335	3.1327	3.1373	3.1348	3.1272	3.1335	3.1437	3.1294	3.1422	3.1360

Tabelle C.7: Kritische Grenzen T_{95} des LR3 Tests für 100’000 Replikationen.

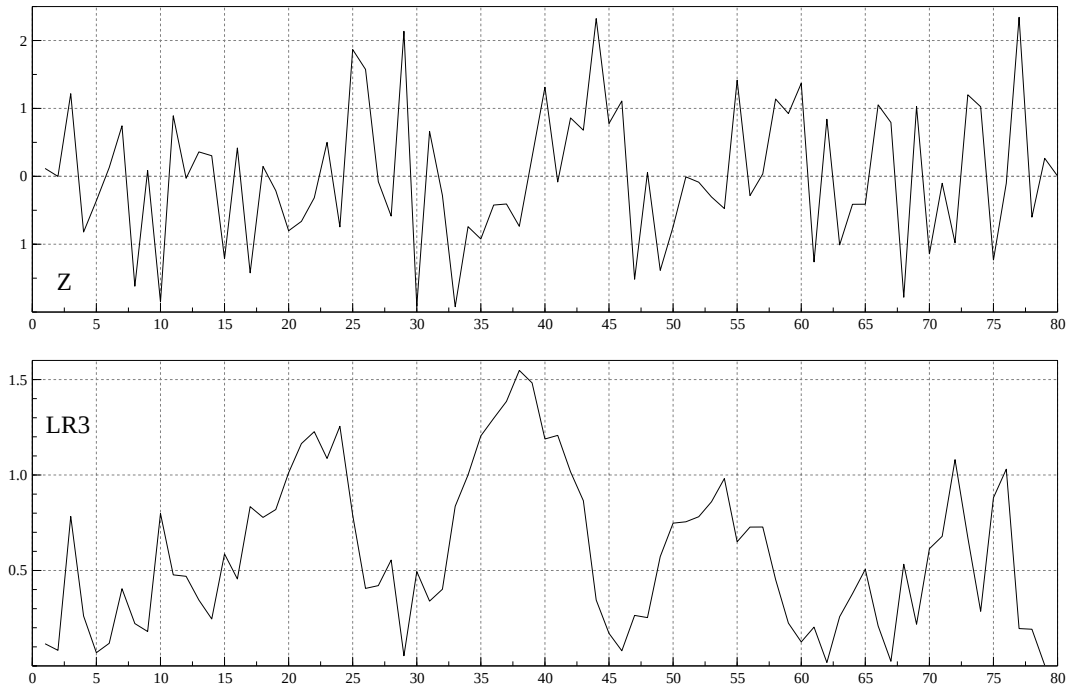


Abbildung C.28: Verlauf der LR3 Testgrösse für die Testreihe Z80

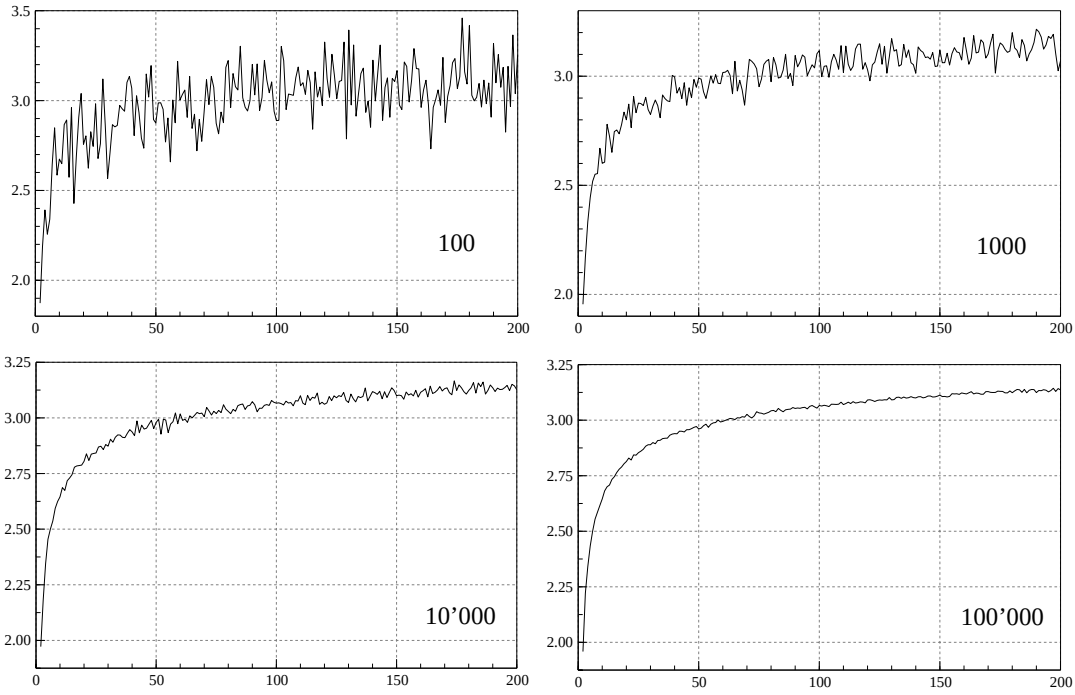


Abbildung C.29: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

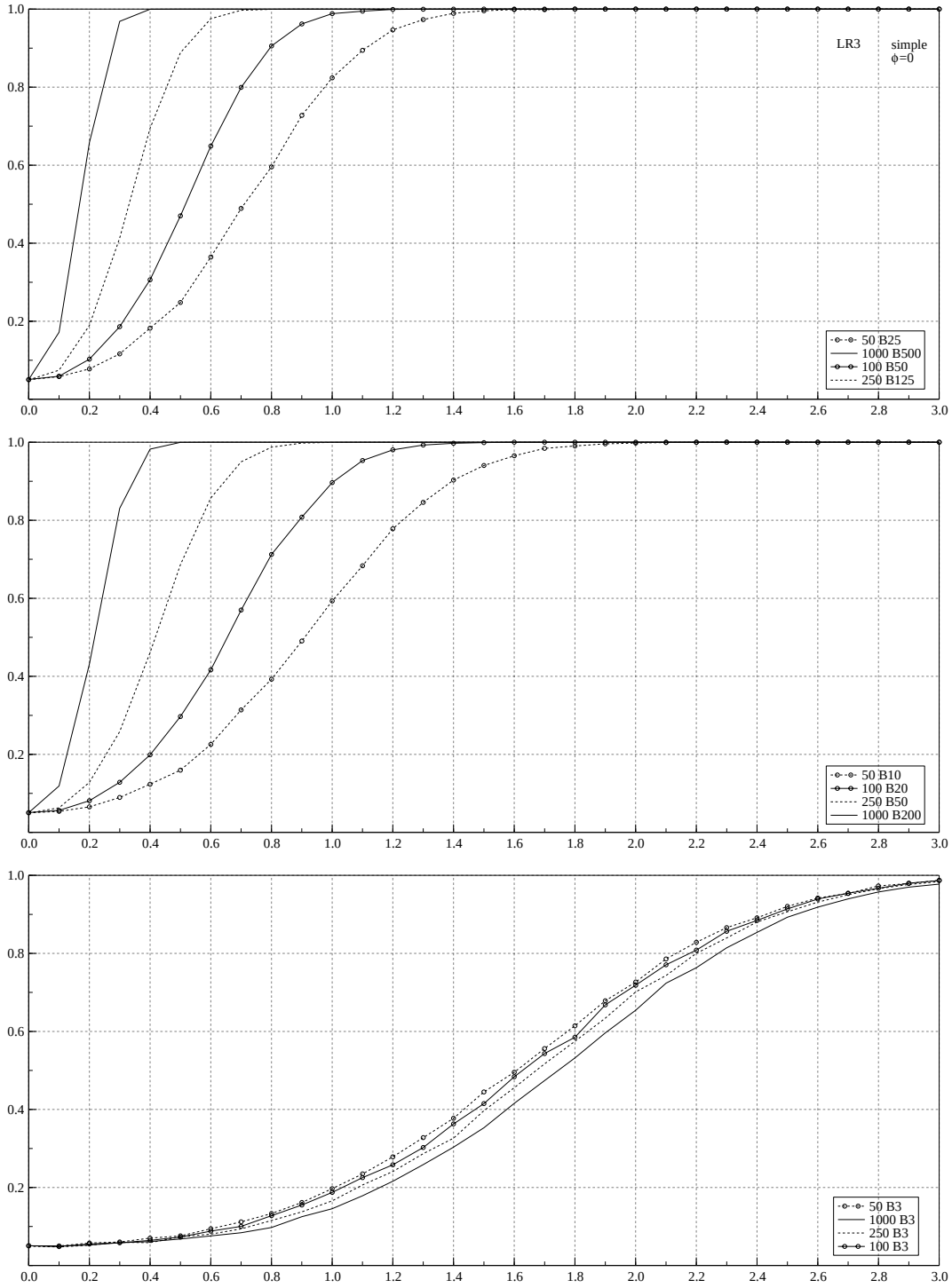


Abbildung C.30: Güte des LR3 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

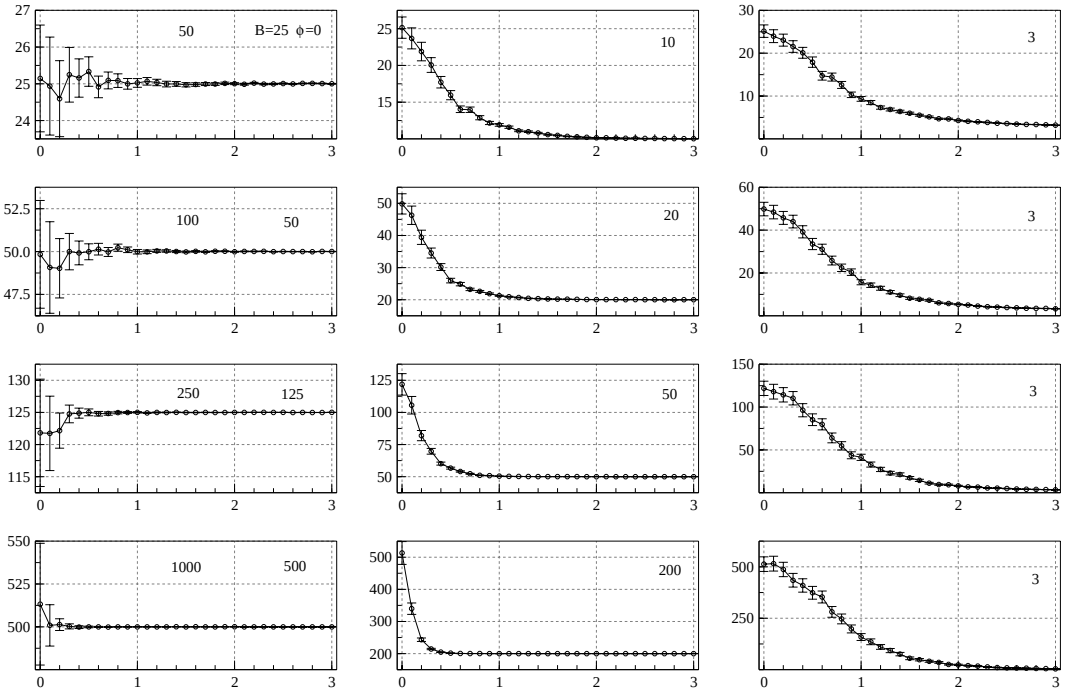


Abbildung C.31: Epochentreue des LR3 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

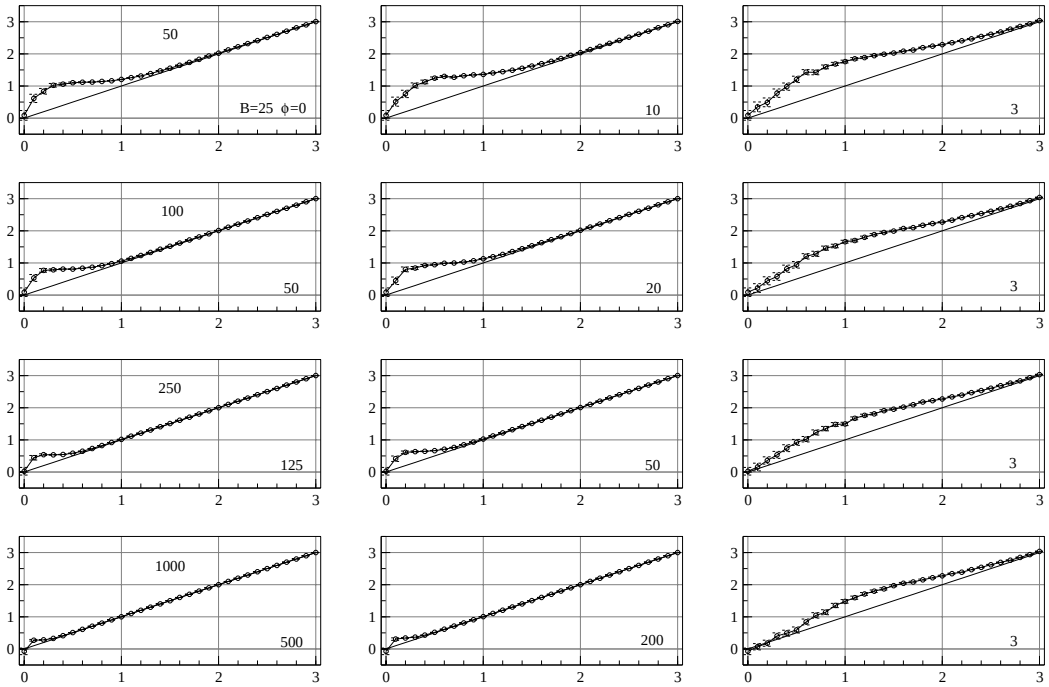


Abbildung C.32: Sprunghöhentreue des LR3 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.4 Likelihood Ratio Test 4

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.8611	2.8635	2.8662	3.1547	3.1580	3.1611	3.7584	3.7652	3.7723
100	0	2.8878	2.8903	2.8928	3.1611	3.1641	3.1671	3.7126	3.7189	3.7254
250	0	2.9436	2.9458	2.9481	3.1996	3.2027	3.2060	3.7140	3.7193	3.7250
1000	0	3.0387	3.0408	3.0430	3.2830	3.2858	3.2887	3.7714	3.7769	3.7824
50	0.1	3.0935	3.0964	3.0994	3.4194	3.4233	3.4272	4.0884	4.0967	4.1045
100	0.1	3.1298	3.1326	3.1350	3.4296	3.4333	3.4369	4.0425	4.0484	4.0559
250	0.1	3.2051	3.2075	3.2099	3.4875	3.4905	3.4940	4.0578	4.0634	4.0700
1000	0.1	3.3168	3.3193	3.3218	3.5870	3.5902	3.5935	4.1382	4.1449	4.1519
50	0.5	4.5618	4.5665	4.5713	5.1126	5.1197	5.1259	6.2516	6.2651	6.2805
100	0.5	4.6369	4.6411	4.6458	5.1401	5.1460	5.1520	6.1564	6.1681	6.1801
250	0.5	4.7812	4.7855	4.7898	5.2418	5.2470	5.2526	6.1818	6.1926	6.2030
1000	0.5	5.0024	5.0065	5.0105	5.4361	5.4409	5.4462	6.3206	6.3302	6.3408
50	0.9	9.9661	9.9780	9.9901	11.375	11.391	11.407	14.159	14.188	14.218
100	0.9	10.863	10.877	10.891	12.374	12.391	12.408	15.443	15.481	15.516
250	0.9	11.256	11.269	11.282	12.660	12.677	12.695	15.539	15.573	15.608
1000	0.9	11.745	11.755	11.765	12.945	12.959	12.973	15.370	15.398	15.423

Tabelle C.8: Kritische Grenzen des LR4 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	-	25.1099	7.5864	5.3782	4.6064	4.2310	3.9346	3.7837	3.6597
10	3.5766	3.5163	3.4481	3.4145	3.3831	3.3505	3.3192	3.3002	3.2910	3.2868
20	3.2601	3.2470	3.2426	3.2406	3.2316	3.2039	3.2094	3.2110	3.2130	3.1992
30	3.1905	3.1944	3.1869	3.1796	3.1836	3.1814	3.1730	3.1739	3.1672	3.1631
40	3.1628	3.1664	3.1725	3.1634	3.1622	3.1604	3.1698	3.1579	3.1658	3.1612
50	3.1460	3.1577	3.1594	3.1456	3.1529	3.1572	3.1619	3.1651	3.1500	3.1487
60	3.1560	3.1584	3.1574	3.1532	3.1492	3.1517	3.1519	3.1595	3.1538	3.1646
70	3.1557	3.1494	3.1504	3.1618	3.1645	3.1570	3.1534	3.1597	3.1594	3.1634
80	3.1671	3.1615	3.1667	3.1527	3.1520	3.1561	3.1578	3.1652	3.1590	3.1583
90	3.1664	3.1613	3.1596	3.1594	3.1558	3.1659	3.1653	3.1615	3.1544	3.1571
100	3.1539	3.1646	3.1608	3.1510	3.1678	3.1672	3.1650	3.1657	3.1712	3.1632
110	3.1632	3.1716	3.1562	3.1756	3.1651	3.1622	3.1689	3.1600	3.1616	3.1785
120	3.1777	3.1739	3.1642	3.1702	3.1699	3.1744	3.1705	3.1708	3.1750	3.1672
130	3.1684	3.1752	3.1781	3.1750	3.1796	3.1779	3.1795	3.1805	3.1774	3.1739
140	3.1775	3.1802	3.1682	3.1741	3.1760	3.1808	3.1767	3.1735	3.1759	3.1791
150	3.1748	3.1720	3.1768	3.1750	3.1894	3.1839	3.1819	3.1743	3.1861	3.1842
160	3.1868	3.1835	3.1904	3.1829	3.1848	3.1798	3.1746	3.1779	3.1847	3.1913
170	3.1813	3.1802	3.1794	3.1874	3.1864	3.1872	3.1820	3.1776	3.1868	3.1838
180	3.1896	3.1897	3.1924	3.1812	3.1929	3.1833	3.1896	3.1940	3.1889	3.1795
190	3.1903	3.1875	3.1878	3.1872	3.1900	3.1900	3.1975	3.1787	3.1906	3.1825

Tabelle C.9: Kritische Grenzen T_{95} des LR4 Tests für 100’000 Replikationen.

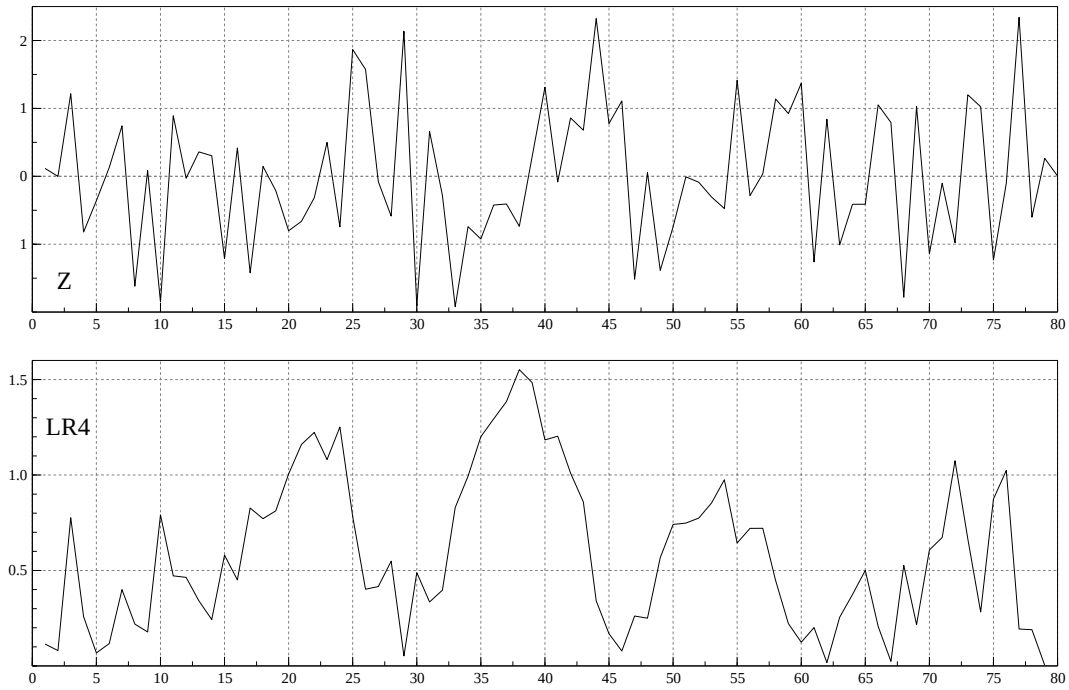


Abbildung C.33: Verlauf der LR4 Testgrösse für die Testreihe Z80

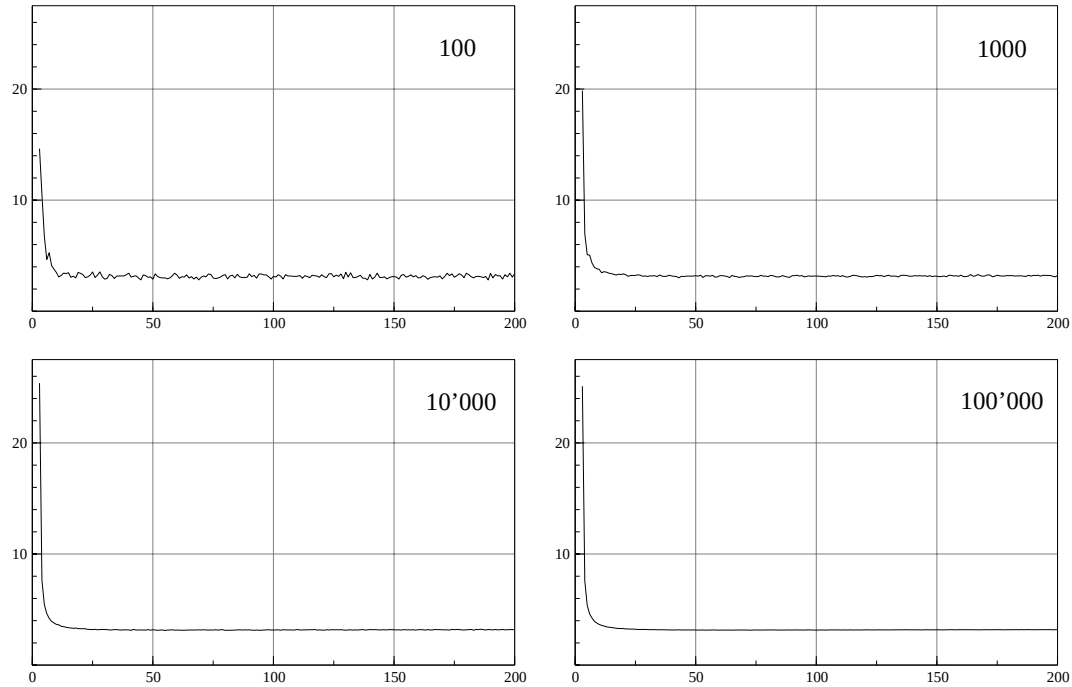


Abbildung C.34: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

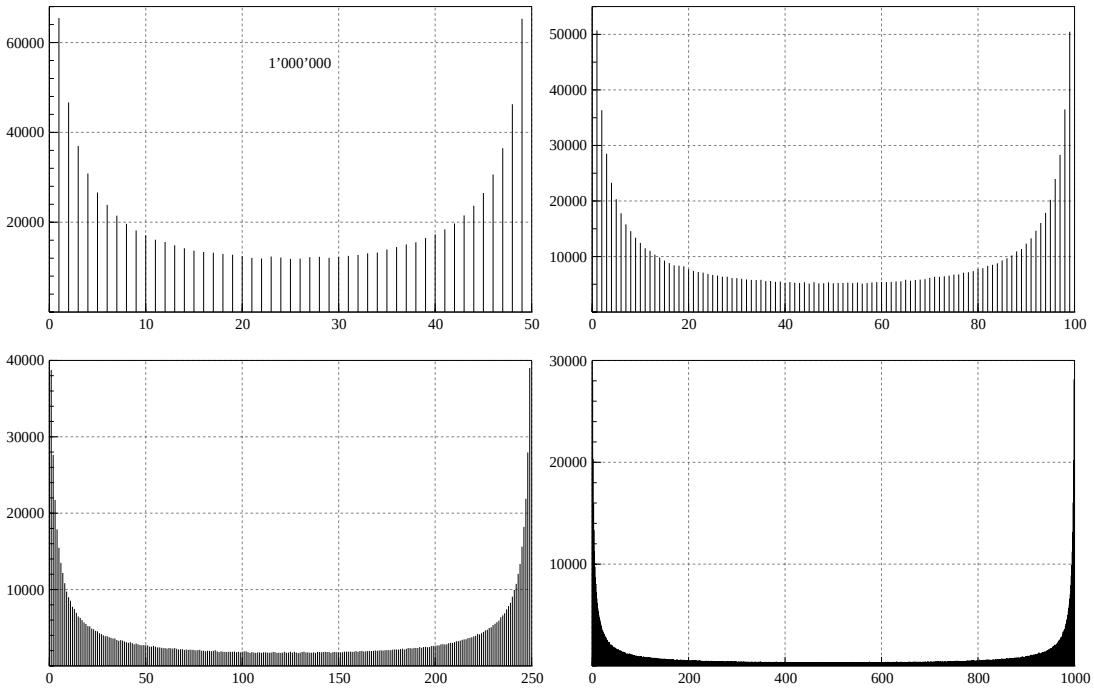


Abbildung C.35: Verteilung der Epochen der Maxima der LR4 Testgrösse mit $\phi = 0$

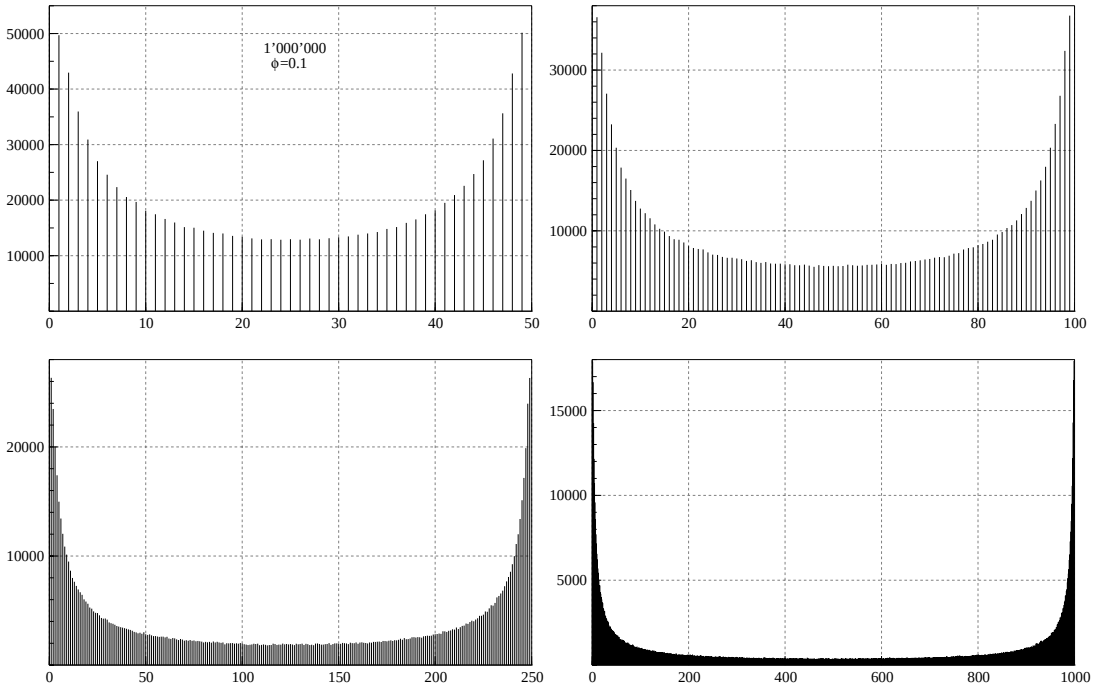


Abbildung C.36: Verteilung der Epochen der Maxima der LR4 Testgrösse mit $\phi = 0.1$

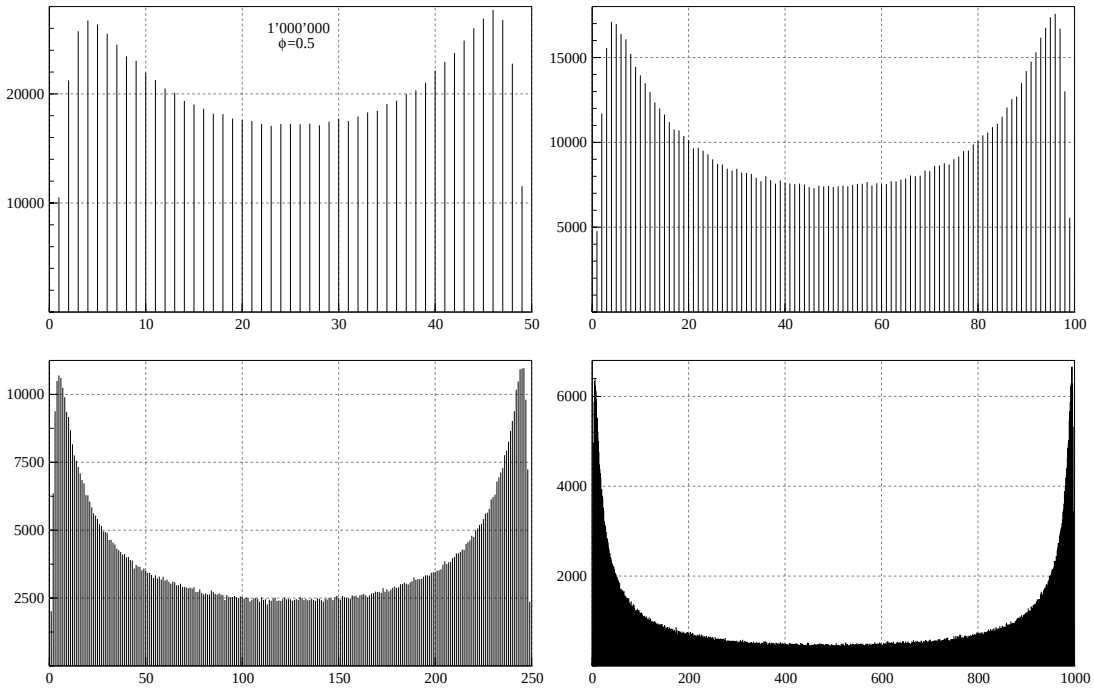


Abbildung C.37: Verteilung der Epochen der Maxima der LR4 Testgrösse mit $\phi = 0.5$

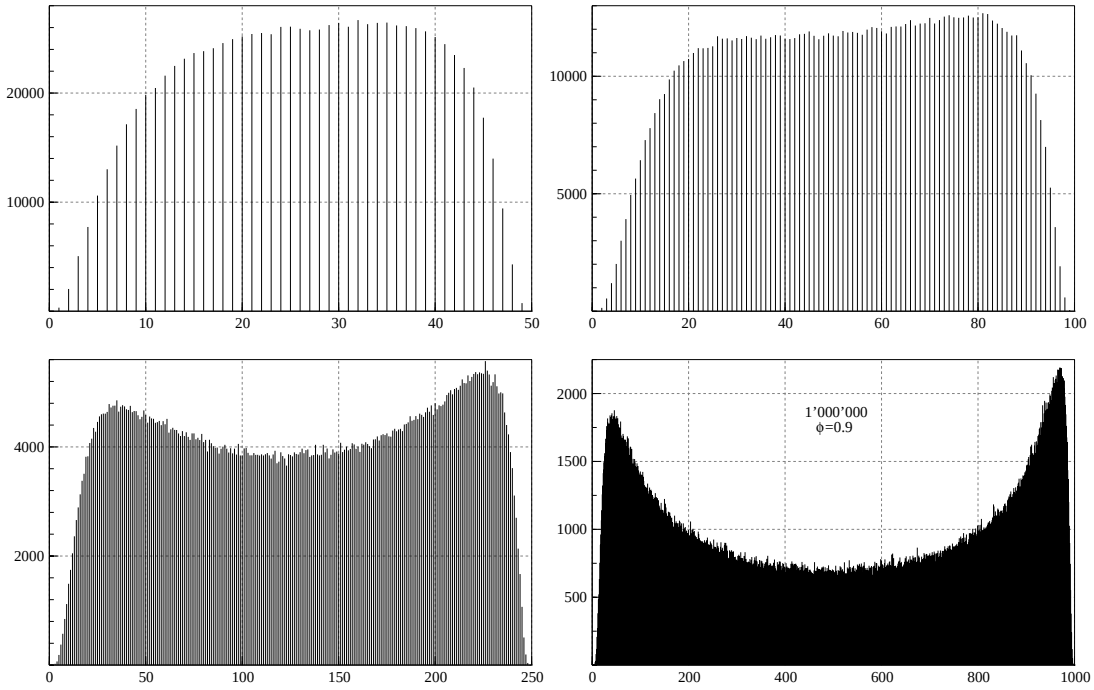


Abbildung C.38: Verteilung der Epochen der Maxima der LR4 Testgrösse mit $\phi = 0.9$

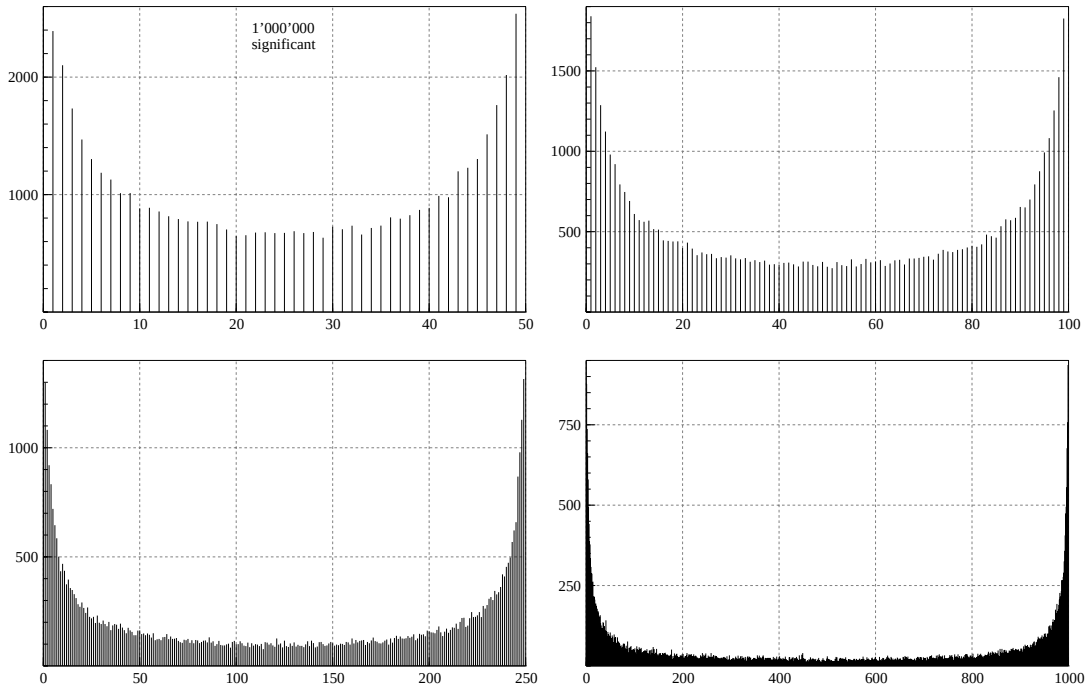


Abbildung C.39: Verteilung der Bruchepochen der LR4 Testgrösse mit $\phi = 0$

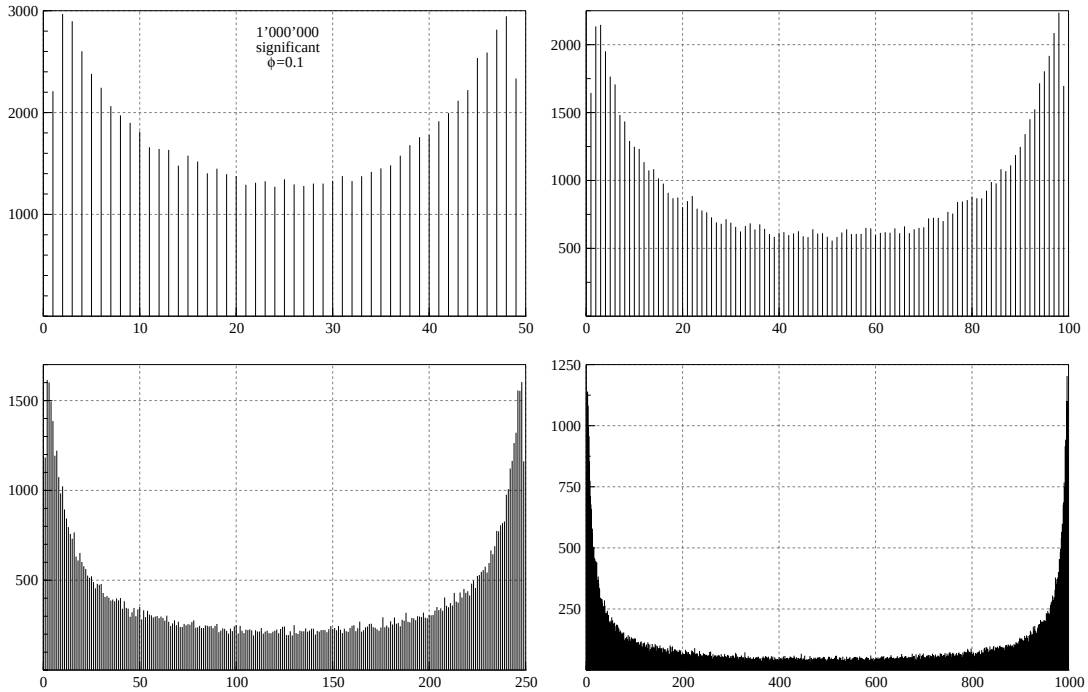


Abbildung C.40: Verteilung der Bruchepochen der LR4 Testgrösse mit $\phi = 0.1$

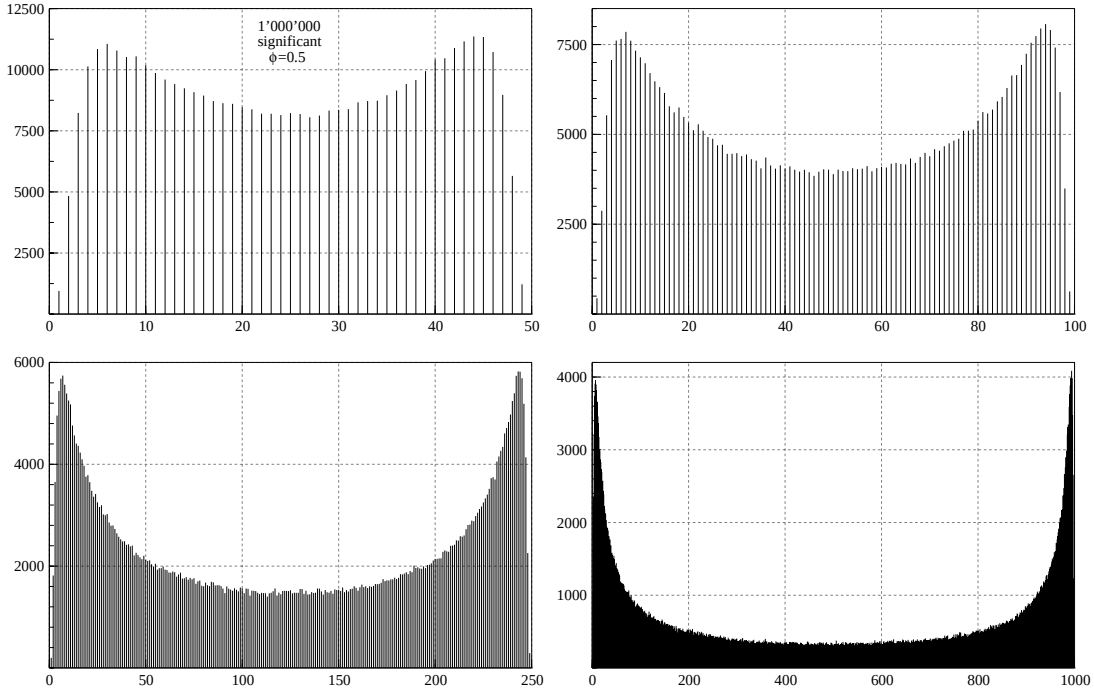


Abbildung C.41: Verteilung der Bruchepochen der LR4 Testgrösse mit $\phi = 0.5$

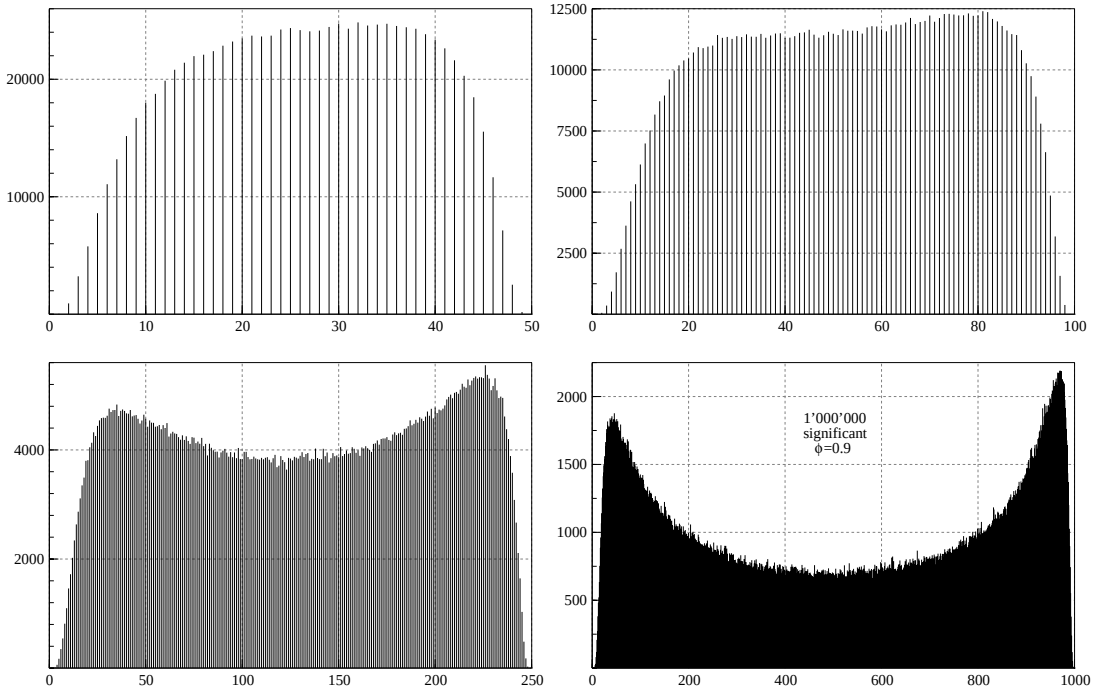


Abbildung C.42: Verteilung der Bruchepochen der LR4 Testgrösse mit $\phi = 0.9$

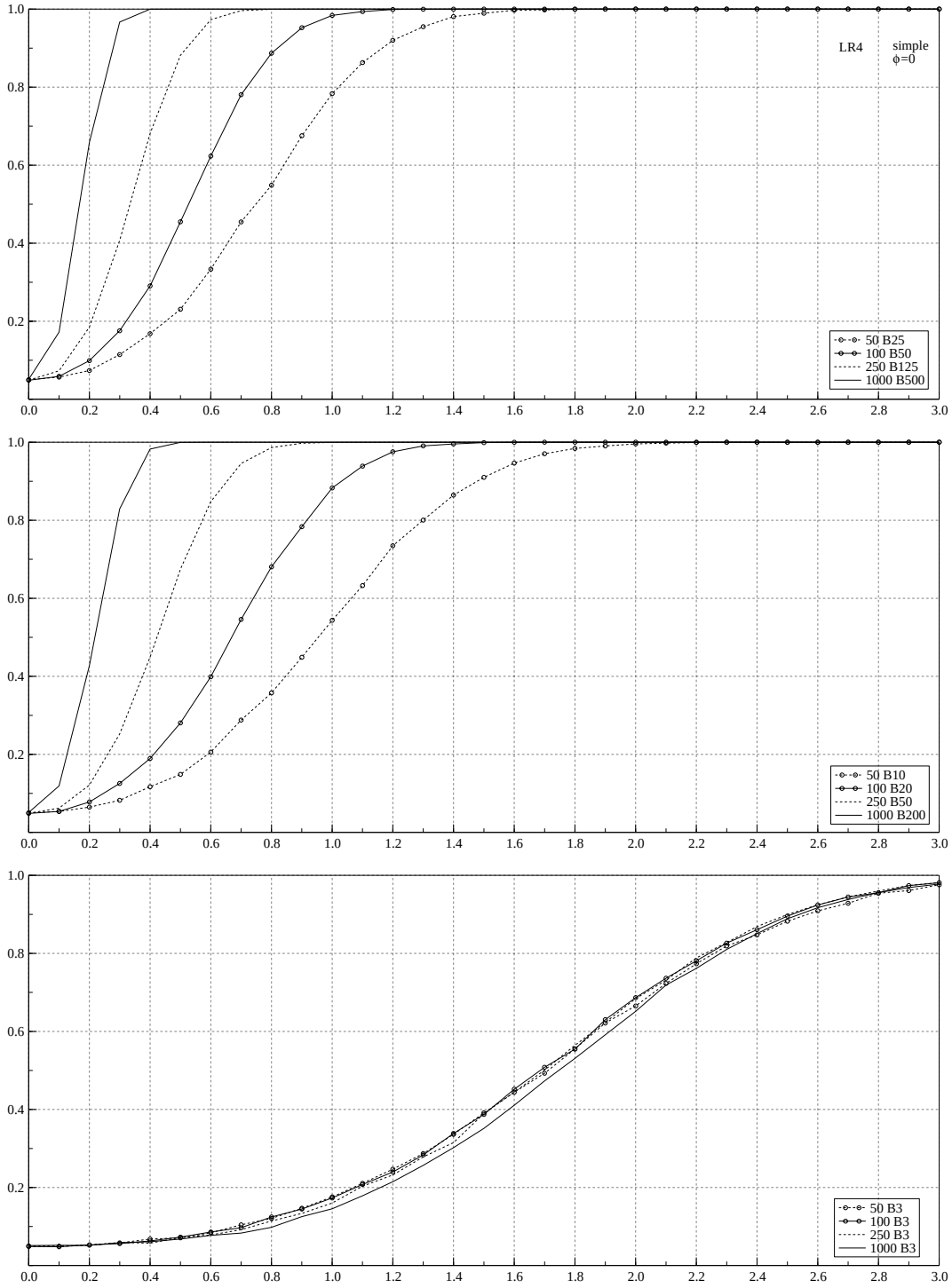


Abbildung C.43: Güte des LR4 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

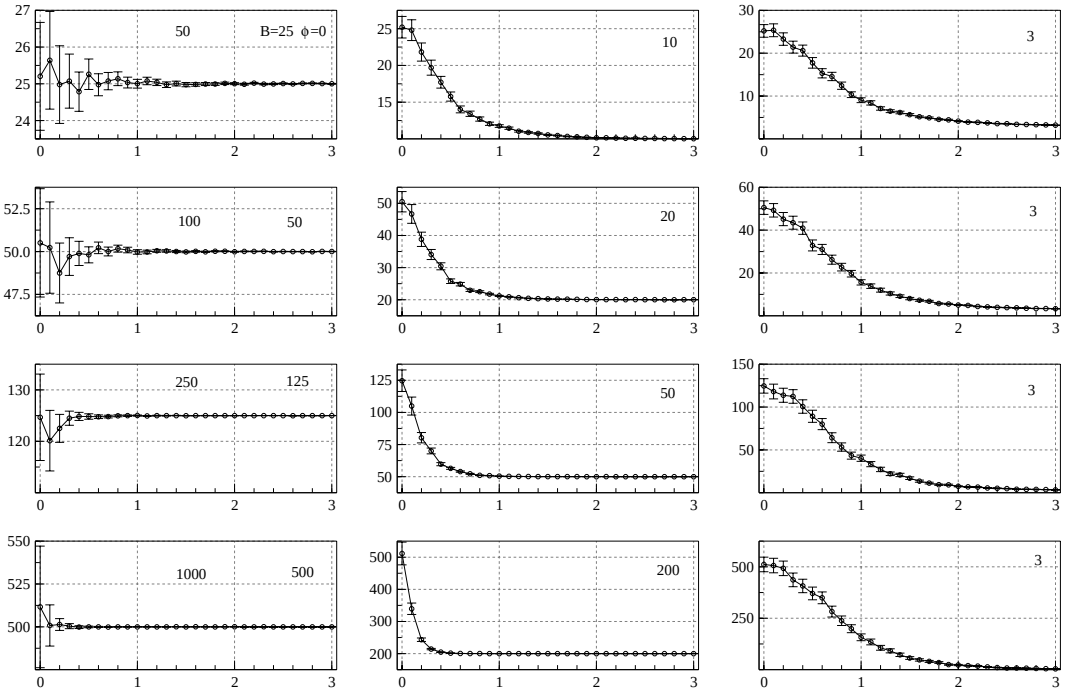


Abbildung C.44: Epochentreue des LR4 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

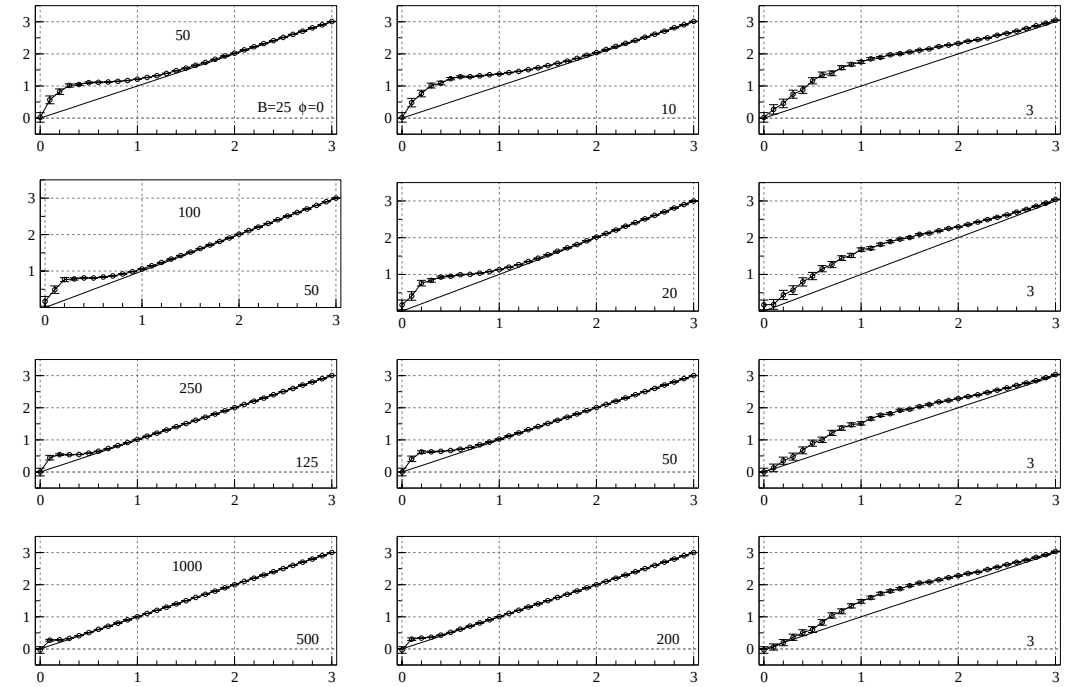


Abbildung C.45: Sprunghöhentreue des LR4 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.5 Likelihood Ratio Test 5

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.9201	2.9226	2.9253	3.2197	3.2232	3.2262	3.8359	3.8428	3.8501
100	0	2.9171	2.9196	2.9221	3.1932	3.1963	3.1993	3.7503	3.7567	3.7632
250	0	2.9555	2.9576	2.9599	3.2125	3.2155	3.2189	3.7289	3.7342	3.7400
1000	0	3.0418	3.0439	3.0460	3.2863	3.2891	3.2920	3.7752	3.7807	3.7862
50	0.1	3.1573	3.1603	3.1633	3.4899	3.4939	3.4978	4.1727	4.1812	4.1892
100	0.1	3.1616	3.1644	3.1668	3.4644	3.4681	3.4718	4.0835	4.0895	4.0971
250	0.1	3.2180	3.2204	3.2228	3.5016	3.5046	3.5081	4.0741	4.0798	4.0864
1000	0.1	3.3201	3.3226	3.3251	3.5906	3.5938	3.5971	4.1424	4.1491	4.1561
50	0.5	4.6558	4.6607	4.6655	5.2181	5.2252	5.2316	6.3805	6.3943	6.4100
100	0.5	4.6840	4.6882	4.6930	5.1923	5.1983	5.2043	6.2189	6.2307	6.2428
250	0.5	4.8005	4.8048	4.8091	5.2629	5.2681	5.2737	6.2066	6.2175	6.2279
1000	0.5	5.0074	5.0116	5.0155	5.4415	5.4463	5.4517	6.3269	6.3365	6.3472
50	0.9	10.172	10.184	10.196	11.610	11.626	11.642	14.451	14.480	14.511
100	0.9	10.973	10.988	11.002	12.499	12.517	12.534	15.600	15.639	15.674
250	0.9	11.302	11.315	11.327	12.711	12.728	12.746	15.602	15.636	15.671
1000	0.9	11.756	11.767	11.777	12.958	12.972	12.986	15.386	15.413	15.439

Tabelle C.10: Kritische Grenzen des LR5 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	-	43.4917	10.7288	6.9432	5.6417	5.0062	4.5433	4.2903	4.0917
10	3.9541	3.8519	3.7484	3.6881	3.6340	3.5818	3.5335	3.5004	3.4792	3.4646
20	3.4274	3.4055	3.3935	3.3847	3.3692	3.3348	3.3353	3.3322	3.3299	3.3115
30	3.2987	3.2992	3.2881	3.2775	3.2787	3.2736	3.2624	3.2608	3.2517	3.2453
40	3.2428	3.2446	3.2490	3.2378	3.2349	3.2315	3.2395	3.2258	3.2324	3.2264
50	3.2096	3.2202	3.2208	3.2056	3.2119	3.2151	3.2188	3.2211	3.2048	3.2025
60	3.2091	3.2106	3.2088	3.2037	3.1988	3.2005	3.2001	3.2070	3.2005	3.2108
70	3.2011	3.1940	3.1945	3.2054	3.2076	3.1993	3.1952	3.2010	3.2002	3.2037
80	3.2069	3.2008	3.2056	3.1909	3.1897	3.1935	3.1948	3.2018	3.1951	3.1940
90	3.2017	3.1962	3.1941	3.1935	3.1895	3.1994	3.1985	3.1943	3.1867	3.1891
100	3.1856	3.1961	3.1920	3.1818	3.1984	3.1975	3.1950	3.1955	3.2007	3.1923
110	3.1920	3.2003	3.1845	3.2038	3.1930	3.1898	3.1963	3.1871	3.1885	3.2053
120	3.2043	3.2002	3.1902	3.1961	3.1956	3.1999	3.1957	3.1959	3.1999	3.1919
130	3.1928	3.1995	3.2023	3.1990	3.2034	3.2015	3.2029	3.2038	3.2005	3.1968
140	3.2003	3.2028	3.1906	3.1964	3.1981	3.2028	3.1986	3.1952	3.1974	3.2006
150	3.1961	3.1930	3.1977	3.1958	3.2102	3.2045	3.2023	3.1946	3.2063	3.2043
160	3.2068	3.2033	3.2101	3.2025	3.2043	3.1992	3.1938	3.1970	3.2037	3.2103
170	3.2000	3.1988	3.1979	3.2058	3.2048	3.2055	3.2002	3.1956	3.2048	3.2017
180	3.2073	3.2074	3.2100	3.1986	3.2103	3.2006	3.2068	3.2111	3.2059	3.1963
190	3.2071	3.2043	3.2044	3.2038	3.2065	3.2064	3.2138	3.1948	3.2068	3.1986

Tabelle C.11: Kritische Grenzen T_{95} des LR5 Tests für 100’000 Replikationen.

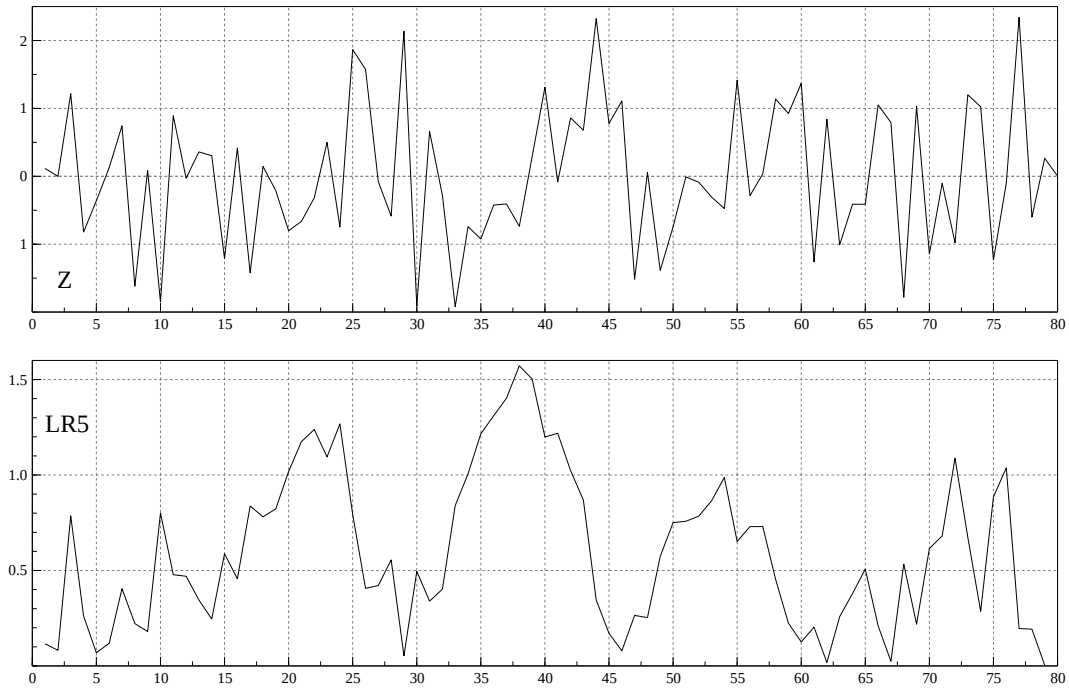


Abbildung C.46: Verlauf der LR5 Testgrösse für die Testreihe Z80

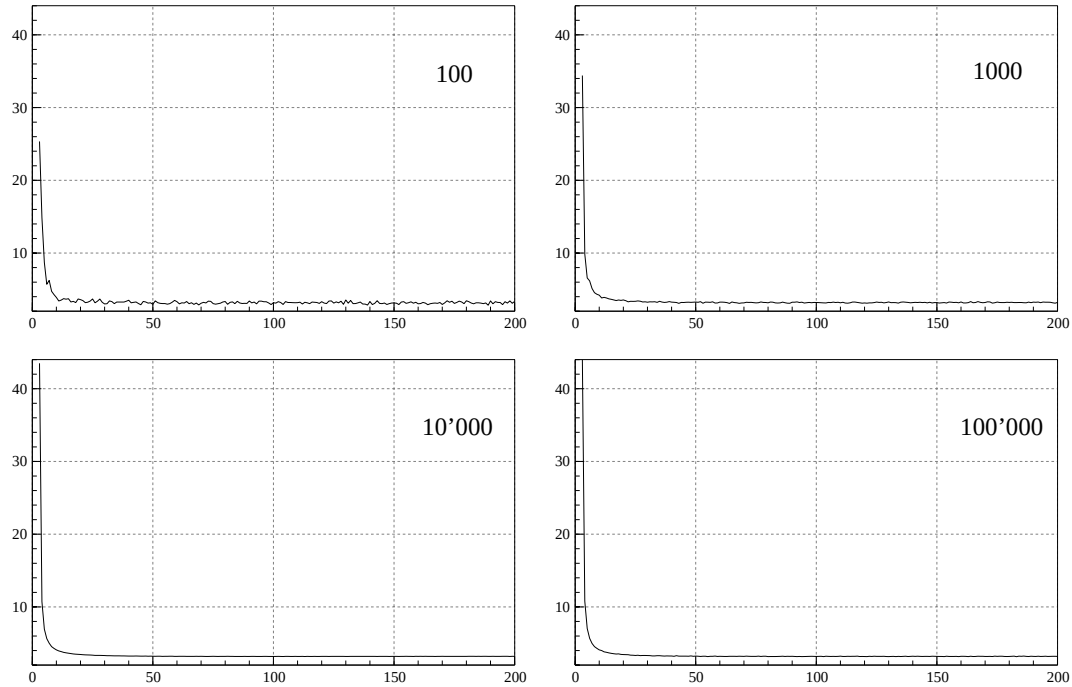


Abbildung C.47: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

C.6 Likelihood Ratio Test 6

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.6990	2.7010	2.7031	2.9302	2.9328	2.9352	3.3717	3.3764	3.3814
100	0	2.8004	2.8026	2.8048	3.0419	3.0445	3.0471	3.5115	3.5167	3.5221
250	0	2.9051	2.9072	2.9094	3.1482	3.1510	3.1542	3.6294	3.6343	3.6396
1000	0	3.0278	3.0299	3.0320	3.2687	3.2714	3.2743	3.7485	3.7540	3.7594
50	0.1	2.8830	2.8852	2.8875	3.1295	3.1324	3.1352	3.5937	3.5991	3.6042
100	0.1	3.0145	3.0169	3.0190	3.2735	3.2767	3.2798	3.7805	3.7852	3.7912
250	0.1	3.1533	3.1556	3.1579	3.4187	3.4215	3.4248	3.9453	3.9504	3.9564
1000	0.1	3.3019	3.3044	3.3069	3.5677	3.5708	3.5741	4.1073	4.1138	4.1206
50	0.5	3.8886	3.8914	3.8942	4.1986	4.2023	4.2057	4.7371	4.7427	4.7491
100	0.5	4.2417	4.2449	4.2484	4.6082	4.6123	4.6165	5.2810	5.2882	5.2956
250	0.5	4.5934	4.5972	4.6009	4.9935	4.9980	5.0028	5.7775	5.7862	5.7946
1000	0.5	4.9458	4.9498	4.9536	5.3627	5.3673	5.3724	6.2040	6.2130	6.2230
50	0.9	5.8060	5.8082	5.8105	6.0391	6.0414	6.0436	6.3514	6.3539	6.3566
100	0.9	7.3913	7.3956	7.4000	7.8085	7.8128	7.8170	8.4188	8.4248	8.4303
250	0.9	9.1944	9.2013	9.2081	9.9065	9.9145	9.9231	11.105	11.118	11.130
1000	0.9	11.019	11.028	11.036	11.990	12.002	12.012	13.835	13.855	13.874

Tabelle C.12: Kritische Grenzen des LR6 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	1.4142	1.7307	1.9661	2.1284	2.2469	2.3392	2.4011	2.4586	2.5021
10	2.5411	2.5757	2.5986	2.6266	2.6501	2.6684	2.6830	2.7000	2.7192	2.7389
20	2.7447	2.7558	2.7702	2.7847	2.7941	2.7909	2.8068	2.8197	2.8322	2.8338
30	2.8380	2.8499	2.8537	2.8571	2.8677	2.8737	2.8750	2.8824	2.8842	2.8874
40	2.8930	2.9013	2.9113	2.9097	2.9138	2.9173	2.9289	2.9244	2.9347	2.9353
50	2.9275	2.9403	2.9454	2.9382	2.9473	2.9540	2.9610	2.9667	2.9577	2.9595
60	2.9683	2.9729	2.9749	2.9741	2.9733	2.9778	2.9804	2.9889	2.9865	2.9976
70	2.9924	2.9893	2.9922	3.0036	3.0079	3.0035	3.0024	3.0094	3.0109	3.0161
80	3.0209	3.0178	3.0239	3.0135	3.0144	3.0194	3.0224	3.0302	3.0263	3.0270
90	3.0353	3.0323	3.0321	3.0332	3.0313	3.0414	3.0421	3.0399	3.0349	3.0383
100	3.0367	3.0472	3.0449	3.0374	3.0531	3.0536	3.0527	3.0543	3.0602	3.0540
110	3.0549	3.0633	3.0505	3.0686	3.0602	3.0585	3.0653	3.0582	3.0605	3.0763
120	3.0764	3.0738	3.0659	3.0720	3.0725	3.0773	3.0745	3.0755	3.0800	3.0737
130	3.0754	3.0822	3.0856	3.0834	3.0882	3.0873	3.0893	3.0909	3.0887	3.0862
140	3.0901	3.0931	3.0828	3.0887	3.0909	3.0959	3.0927	3.0904	3.0931	3.0966
150	3.0931	3.0911	3.0960	3.0948	3.1085	3.1039	3.1026	3.0962	3.1075	3.1062
160	3.1090	3.1065	3.1132	3.1068	3.1090	3.1049	3.1005	3.1040	3.1107	3.1172
170	3.1083	3.1077	3.1074	3.1152	3.1147	3.1158	3.1114	3.1077	3.1166	3.1142
180	3.1199	3.1204	3.1232	3.1132	3.1244	3.1159	3.1221	3.1265	3.1221	3.1137
190	3.1241	3.1219	3.1224	3.1222	3.1252	3.1255	3.1328	3.1156	3.1270	3.1198

Tabelle C.13: Kritische Grenzen T_{95} des LR6 Tests für 100’000 Replikationen.

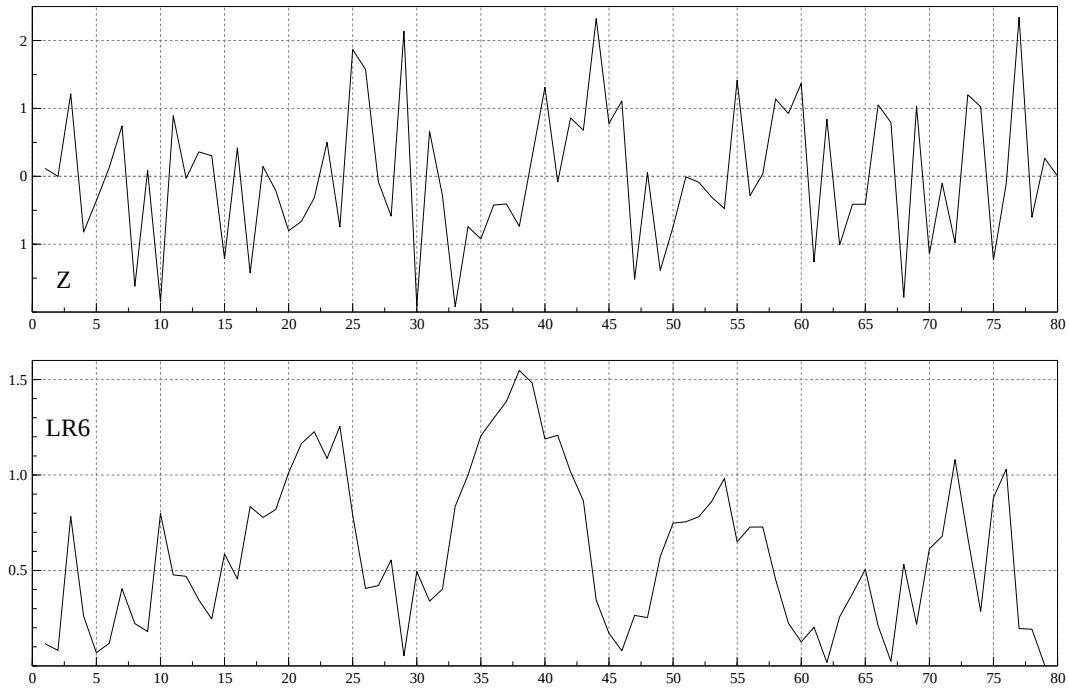


Abbildung C.48: Verlauf der LR6 Testgrösse für die Testreihe Z80

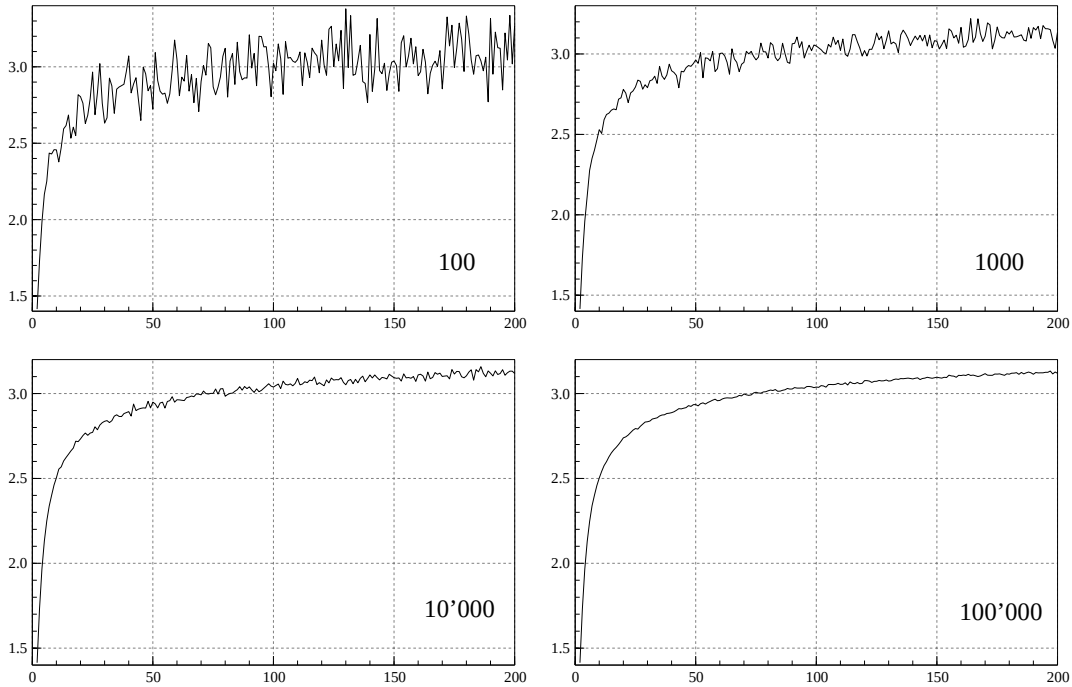


Abbildung C.49: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

C.7 Likelihood Ratio Test 7

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	2.6719	2.6738	2.6759	2.9008	2.9034	2.9057	3.3378	3.3425	3.3474
100	0	2.7864	2.7886	2.7908	3.0266	3.0293	3.0318	3.4939	3.4991	3.5044
250	0	2.8993	2.9014	2.9036	3.1419	3.1447	3.1479	3.6221	3.6270	3.6323
1000	0	3.0263	3.0283	3.0305	3.2670	2.2698	3.2726	3.7467	3.7521	3.7575
50	0.1	2.8540	2.8562	2.8585	3.0980	3.1009	3.1037	3.5575	3.5629	3.5679
100	0.1	2.9994	3.0018	3.0039	3.2571	3.2602	3.2633	3.7615	3.7663	3.7722
250	0.1	3.1470	3.1493	3.1516	3.4119	3.4147	3.4179	3.9374	3.9425	3.9484
1000	0.1	3.3003	3.3028	3.3052	3.5659	3.5690	3.5723	4.1052	4.1117	4.1186
50	0.5	3.8495	3.8523	3.8551	4.1564	4.1601	4.1634	4.6895	4.6951	4.7014
100	0.5	4.2205	4.2236	4.2271	4.5851	4.5892	4.5934	5.2545	5.2617	5.2691
250	0.5	4.5842	4.5880	4.5917	4.9835	4.9880	4.9928	5.7659	5.7746	5.7830
1000	0.5	4.9433	4.9473	4.9511	5.3600	5.3646	5.3698	6.2009	6.2099	6.2199
50	0.9	5.7476	5.7498	5.7521	5.9784	5.9807	5.9829	6.2876	6.2901	6.2927
100	0.9	7.3542	7.3585	7.3629	7.7694	7.7737	7.7779	8.3766	8.3826	8.3880
250	0.9	9.1760	9.1829	9.1897	9.8867	9.8947	9.9032	11.083	11.096	11.108
1000	0.9	11.014	11.023	11.031	11.984	11.996	12.006	13.828	13.848	13.867

Tabelle C.14: Kritische Grenzen des LR7 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	1.0000	1.4131	1.7027	1.9037	2.0511	2.1656	2.2461	2.3180	2.3737
10	2.4228	2.4661	2.4966	2.5311	2.5602	2.5836	2.6029	2.6240	2.6467	2.6695
20	2.6785	2.6924	2.7093	2.7261	2.7376	2.7367	2.7544	2.7689	2.7829	2.7862
30	2.7919	2.8050	2.8101	2.8148	2.8265	2.8335	2.8359	2.8442	2.8469	2.8511
40	2.8575	2.8666	2.8772	2.8765	2.8812	2.8854	2.8976	2.8938	2.9046	2.9058
50	2.8986	2.9119	2.9175	2.9108	2.9204	2.9275	2.9349	2.9410	2.9325	2.9348
60	2.9439	2.9489	2.9512	2.9507	2.9503	2.9551	2.9581	2.9669	2.9648	2.9762
70	2.9713	2.9684	2.9716	2.9833	2.9878	2.9836	2.9828	2.9900	2.9918	2.9972
80	3.0022	2.9993	3.0056	2.9955	2.9966	3.0018	3.0050	3.0129	3.0092	3.0101
90	3.0186	3.0158	3.0158	3.0170	3.0153	3.0255	3.0264	3.0244	3.0195	3.0231
100	3.0216	3.0322	3.0301	3.0227	3.0385	3.0392	3.0384	3.0402	3.0461	3.0401
110	3.0411	3.0496	3.0370	3.0551	3.0469	3.0453	3.0522	3.0452	3.0476	3.0635
120	3.0637	3.0612	3.0534	3.0596	3.0602	3.0650	3.0624	3.0635	3.0680	3.0619
130	3.0636	3.0705	3.0739	3.0719	3.0768	3.0759	3.0780	3.0797	3.0776	3.0751
140	3.0791	3.0821	3.0720	3.0780	3.0803	3.0853	3.0822	3.0799	3.0827	3.0862
150	3.0829	3.0809	3.0858	3.0848	3.0985	3.0940	3.0927	3.0864	3.0977	3.0965
160	3.0994	3.0969	3.1037	3.0973	3.0996	3.0955	3.0912	3.0947	3.1014	3.1080
170	3.0992	3.0987	3.0984	3.1062	3.1058	3.1070	3.1026	3.0990	3.1079	3.1055
180	3.1113	3.1118	3.1147	3.1048	3.1160	3.1075	3.1138	3.1182	3.1139	3.1055
190	3.1159	3.1137	3.1143	3.1142	3.1171	3.1175	3.1248	3.1077	3.1191	3.1120

Tabelle C.15: Kritische Grenzen T_{95} des LR7 Tests für 100’000 Replikationen.

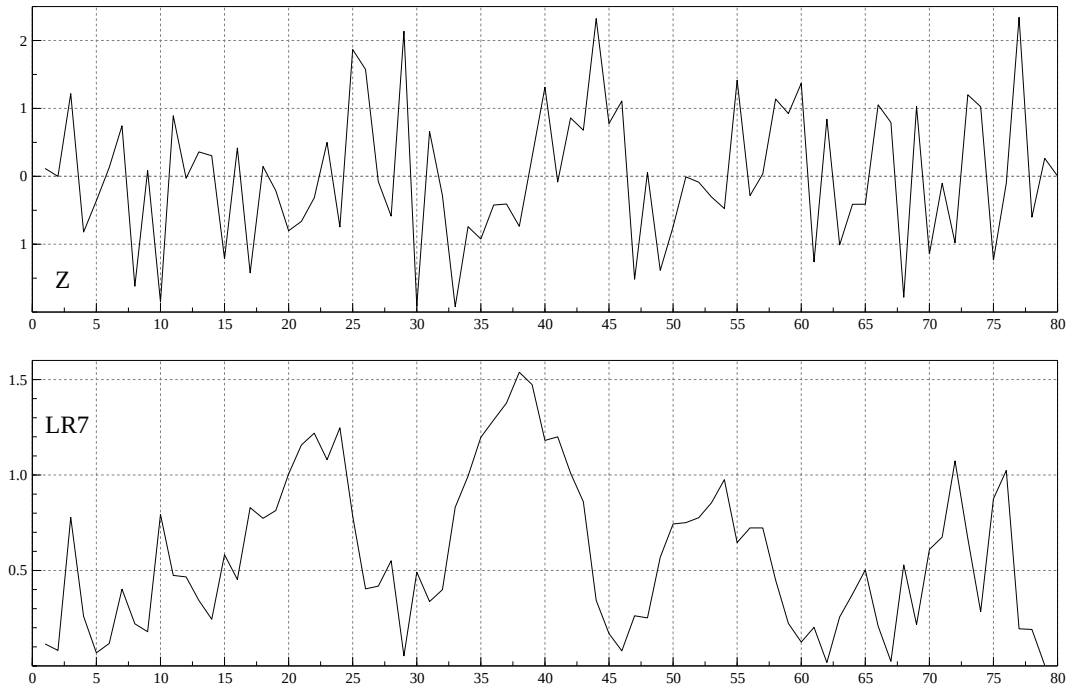


Abbildung C.50: Verlauf der LR7 Testgrösse für die Testreihe Z80

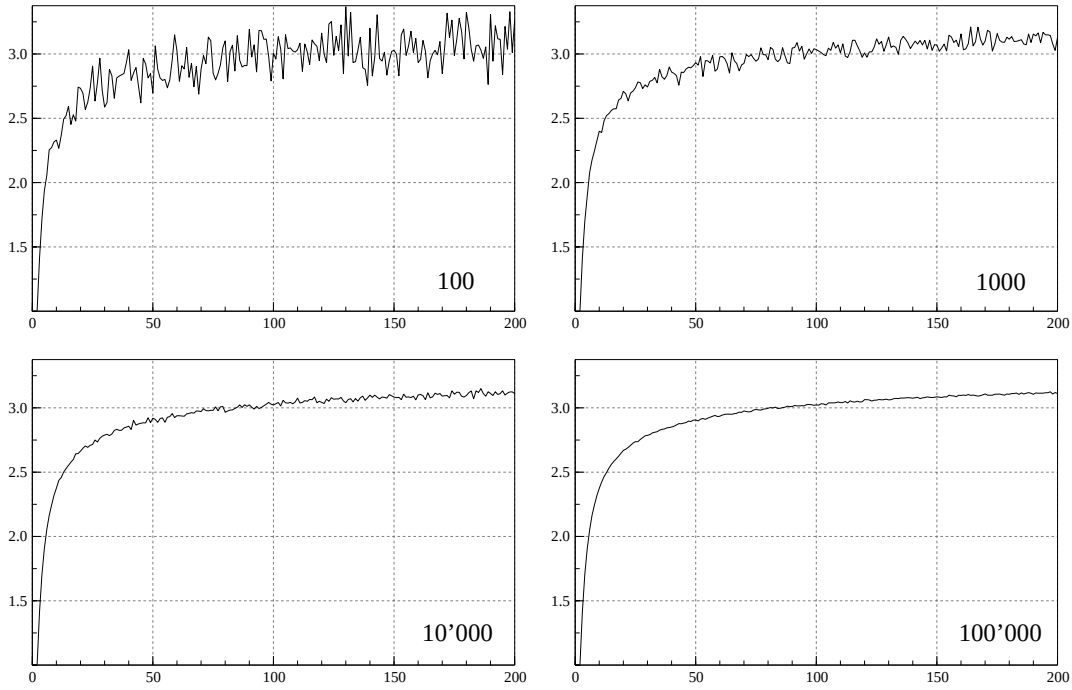


Abbildung C.51: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

C.8 Alexandersson Test

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	7.2845	7.2953	7.3068	8.5864	8.6015	8.6152	11.368	11.400	11.434
100	0	7.8423	7.8546	7.8671	9.2529	9.2692	9.2849	12.331	12.367	12.405
250	0	8.4399	8.4519	8.4646	9.9111	9.9290	9.9488	13.172	13.208	13.247
1000	0	9.1675	9.1801	9.1931	10.684	10.702	10.721	14.052	14.092	14.133
50	0.1	8.3116	8.3246	8.3379	9.7936	9.8120	9.8296	12.914	12.953	12.990
100	0.1	9.0873	9.1018	9.1146	10.716	10.737	10.757	14.292	14.328	14.373
250	0.1	9.9435	9.9578	9.9724	11.688	11.707	11.729	15.565	15.605	15.653
1000	0.1	10.903	10.919	10.935	12.728	12.751	12.774	16.870	16.923	16.980
50	0.5	15.121	15.143	15.165	17.628	17.660	17.688	22.440	22.493	22.554
100	0.5	17.992	18.019	18.049	21.235	21.274	21.312	27.889	27.965	28.044
250	0.5	21.099	21.134	21.169	24.935	24.980	25.028	33.379	33.480	33.577
1000	0.5	24.461	24.500	24.538	28.759	28.808	28.863	38.489	38.602	38.726
50	0.9	33.709	33.735	33.762	36.471	36.499	36.526	40.341	40.373	40.406
100	0.9	54.631	54.695	54.760	60.973	61.040	61.106	70.876	70.978	71.070
250	0.9	84.538	84.664	84.789	98.139	98.298	98.467	123.33	123.61	123.88
1000	0.9	121.43	121.62	121.80	143.77	144.04	144.30	191.41	191.96	192.48

Tabelle C.16: Kritische Grenzen des Alexandersson Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	2.0000	2.9952	3.8657	4.5301	5.0483	5.4717	5.7655	6.0445	6.2606
10	6.4570	6.6343	6.7525	6.8991	7.0229	7.1202	7.1987	7.2901	7.3942	7.5014
20	7.5332	7.5942	7.6738	7.7547	7.8069	7.7891	7.8784	7.9507	8.0213	8.0304
30	8.0542	8.1219	8.1436	8.1629	8.2239	8.2581	8.2658	8.3082	8.3183	8.3370
40	8.3694	8.4175	8.4754	8.4664	8.4902	8.5105	8.5787	8.5520	8.6122	8.6158
50	8.5702	8.6455	8.6754	8.6328	8.6868	8.7262	8.7673	8.8013	8.7480	8.7588
60	8.8107	8.8384	8.8498	8.8450	8.8405	8.8671	8.8828	8.9337	8.9191	8.9859
70	8.9547	8.9358	8.9531	9.0218	9.0474	9.0208	9.0142	9.0565	9.0658	9.0967
80	9.1258	9.1071	9.1438	9.0811	9.0866	9.1170	9.1349	9.1819	9.1583	9.1628
90	9.2133	9.1947	9.1939	9.2004	9.1890	9.2500	9.2543	9.2413	9.2104	9.2315
100	9.2215	9.2852	9.2714	9.2257	9.3214	9.3246	9.3189	9.3290	9.3645	9.3268
110	9.3325	9.3840	9.3057	9.4165	9.3650	9.3542	9.3960	9.3527	9.3664	9.4638
120	9.4643	9.4484	9.3996	9.4375	9.4404	9.4696	9.4525	9.4590	9.4862	9.4478
130	9.4582	9.5001	9.5207	9.5075	9.5371	9.5313	9.5440	9.5538	9.5402	9.5245
140	9.5485	9.5670	9.5034	9.5402	9.5540	9.5844	9.5651	9.5503	9.5671	9.5887
150	9.5675	9.5546	9.5849	9.5780	9.6629	9.6344	9.6262	9.5863	9.6563	9.6484
160	9.6660	9.6501	9.6922	9.6521	9.6660	9.6404	9.6132	9.6347	9.6762	9.7169
170	9.6617	9.6579	9.6558	9.7042	9.7013	9.7083	9.6809	9.6581	9.7132	9.6983
180	9.7338	9.7368	9.7547	9.6922	9.7620	9.7089	9.7476	9.7753	9.7477	9.6952
190	9.7602	9.7461	9.7497	9.7483	9.7666	9.7688	9.8142	9.7067	9.7780	9.7330

Tabelle C.17: Kritische Grenzen T_{95} des Alexandersson Tests für 100’000 Replikationen.

N	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
12	5.762	5.769	5.777	6.613	6.623	6.632	8.116	8.131	8.144
18	5.265	6.274	6.284	7.289	7.301	7.312	9.279	9.303	9.324
24	6.571	6.582	6.591	7.694	7.707	7.721	9.990	10.02	10.04
30	6.812	6.823	6.834	8.005	8.019	8.033	10.47	10.49	10.52
36	6.992	7.002	7.014	8.233	8.247	8.262	10.82	10.85	10.88
42	7.131	7.142	7.153	8.397	8.413	8.427	11.09	11.12	11.15
48	7.257	7.269	7.280	8.563	8.580	8.595	11.32	11.35	11.39
50	7.285	7.295	7.307	8.586	8.602	8.615	11.37	11.40	11.43
54	7.357	7.370	7.380	8.669	8.685	8.701	11.49	11.52	11.55
60	7.450	7.462	7.474	8.791	8.806	8.823	11.67	11.70	11.74
66	7.513	7.525	7.536	8.858	8.874	8.890	11.77	11.80	11.84
72	7.605	7.617	7.628	8.961	8.978	8.994	11.92	11.95	11.99
78	7.656	7.667	7.678	9.023	9.039	9.056	12.00	12.04	12.07
84	7.711	7.722	7.734	9.108	9.108	9.125	12.11	12.15	12.19
90	7.765	7.776	7.789	9.156	9.172	9.189	12.20	12.24	12.27
96	7.820	7.832	7.844	9.203	9.220	9.237	12.26	12.31	12.34
100	7.842	7.855	7.867	9.253	9.269	9.285	12.33	12.37	12.41
102	7.864	7.875	7.887	9.272	9.289	9.307	12.34	12.38	12.42
108	7.901	7.913	7.925	9.307	9.323	9.341	12.38	12.42	12.46
114	7.933	7.945	7.956	9.348	9.365	9.382	12.47	12.50	12.54
120	7.965	7.977	7.990	9.384	9.400	9.417	12.52	12.56	12.59
180	8.253	8.265	8.278	9.706	9.724	9.742	12.92	12.96	12.99
240	8.431	8.443	8.456	9.901	9.919	9.935	13.19	13.23	13.27
250	8.440	8.452	8.465	9.911	9.929	9.949	13.17	13.21	13.25
300	8.559	8.572	8.586	10.04	10.06	10.08	13.32	13.37	13.41
360	8.656	8.669	8.681	10.14	10.16	10.18	13.45	13.49	13.54
420	8.736	8.749	8.762	10.23	10.25	10.27	13.55	13.59	13.63
480	8.819	8.831	8.845	10.31	10.33	10.35	13.65	13.69	13.73
540	8.885	8.898	8.910	10.38	10.40	10.42	13.73	13.76	13.80
600	8.920	8.932	8.945	10.43	10.45	10.46	13.79	13.83	13.86
660	8.965	8.979	8.993	10.47	10.49	10.51	13.84	13.88	13.92
720	9.011	9.024	9.038	10.52	10.54	10.56	13.91	13.94	13.98
780	9.039	9.053	9.066	10.55	10.57	10.59	13.95	13.99	14.03
840	9.095	9.108	9.120	10.61	10.63	10.64	14.01	14.05	14.09
900	9.115	9.128	9.140	10.64	10.66	10.67	14.01	14.06	14.10
960	9.135	9.148	9.162	10.66	10.68	10.70	14.07	14.11	14.15
1000	9.168	9.180	9.193	10.68	10.70	10.72	14.05	14.09	14.13
1020	9.170	9.183	9.195	10.70	10.72	10.74	14.11	14.15	14.19
1080	9.196	9.208	9.222	10.71	10.73	10.75	14.10	14.13	14.18
1140	9.229	9.243	9.256	10.75	10.77	10.79	14.16	14.20	14.25
1200	9.238	9.251	9.264	10.75	10.77	10.79	14.18	14.23	14.27
6000	9.866	9.880	9.894	11.41	11.43	11.45	14.83	14.87	14.90

Tabelle C.18: Kritische Grenzen des Alexandersson Tests für 1’000’000 Replikationen.

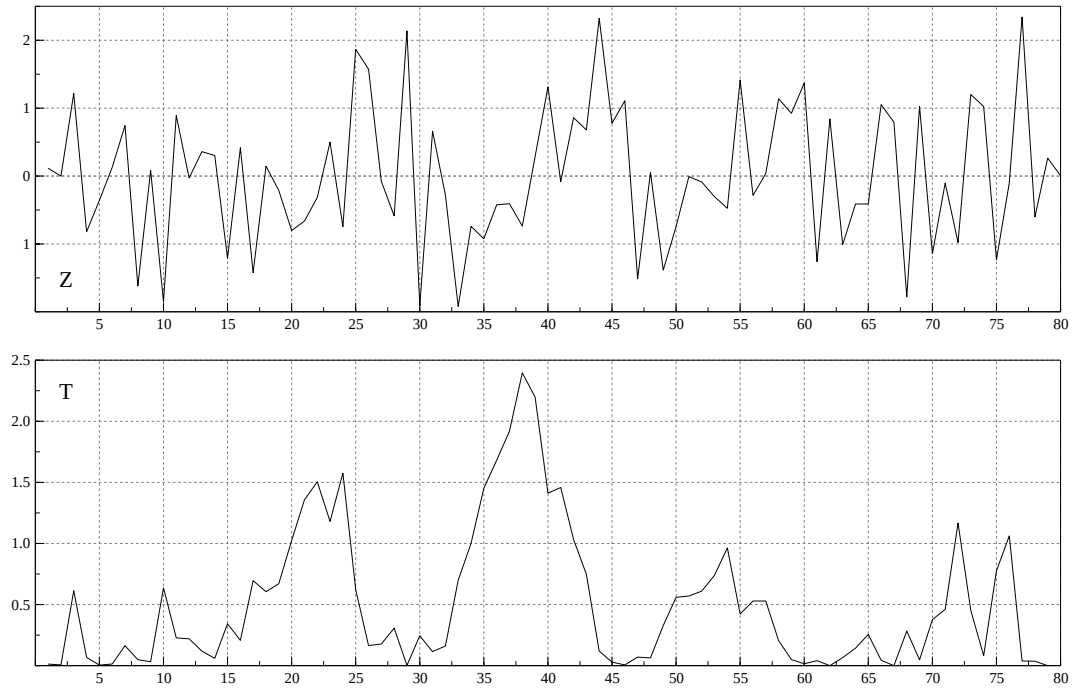


Abbildung C.52: Verlauf der Alexandersson Testgrösse für die Testreihe Z80

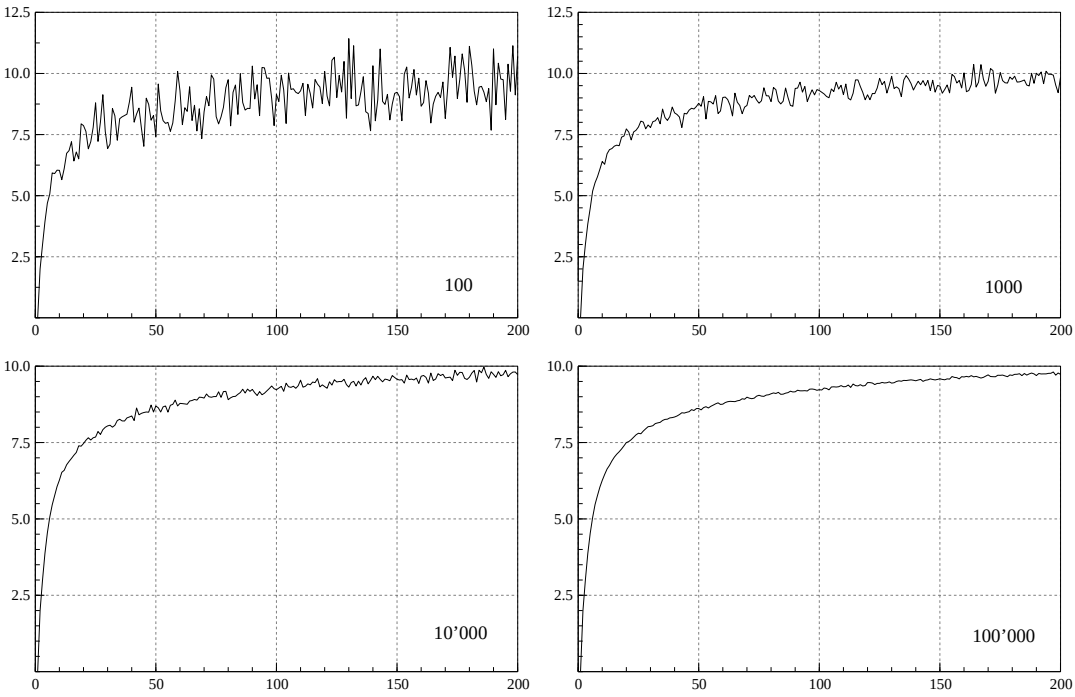


Abbildung C.53: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

C.9 Buishand Test

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
50	0	8.0804	8.0888	8.0970	8.9680	8.9784	8.9885	10.654	10.674	10.694
100	0	11.652	11.664	11.676	12.949	12.966	12.981	15.530	15.560	15.589
250	0	18.739	18.757	18.775	20.848	20.873	20.898	24.997	25.040	25.087
1000	0	38.076	38.112	38.152	42.286	42.337	42.385	50.646	50.744	50.839
50	0.1	8.7617	8.7707	8.7798	9.7127	9.7238	9.7346	11.520	11.542	11.562
100	0.1	12.721	12.734	12.746	14.135	14.152	14.168	16.894	16.922	16.951
250	0.1	20.596	20.616	20.635	22.898	22.923	22.952	27.434	27.486	27.538
1000	0.1	41.929	41.971	42.015	46.577	46.629	46.688	55.817	55.914	56.021
50	0.5	12.395	12.406	12.418	13.609	13.622	13.636	15.749	15.772	15.793
100	0.5	18.675	18.693	18.713	20.678	20.700	20.723	24.393	24.429	24.470
250	0.5	31.109	31.140	31.170	34.551	34.591	34.631	41.242	41.321	41.395
1000	0.5	64.635	64.700	64.766	71.844	71.934	72.023	86.061	86.211	86.382
50	0.9	19.662	19.672	19.683	20.670	20.680	20.690	22.011	22.023	22.034
100	0.9	34.870	34.895	34.919	37.310	37.334	37.361	40.905	40.938	40.972
250	0.9	66.594	66.655	66.716	72.833	72.903	72.983	83.526	83.629	83.745
1000	0.9	152.14	152.29	152.45	168.76	168.95	169.15	200.31	200.67	201.02

Tabelle C.19: Kritische Grenzen des Buishand Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	1.0000	1.4131	1.9010	2.2624	2.5747	2.8593	3.1198	3.3706	3.5989
10	3.8209	4.0341	4.2069	4.4060	4.5992	4.7708	4.9445	5.0904	5.2544	5.4304
20	5.5834	5.7183	5.8686	6.0136	6.1464	6.2699	6.4177	6.5511	6.6862	6.7784
30	6.9321	7.0448	7.1529	7.2564	7.4095	7.5118	7.6151	7.7451	7.8329	7.9538
40	8.0542	8.1557	8.2703	8.3734	8.4624	8.5846	8.6927	8.7617	8.8672	8.9587
50	9.0389	9.1559	9.2546	9.3328	9.4557	9.5413	9.6046	9.7164	9.7898	9.8791
60	9.9838	10.0664	10.1475	10.2234	10.3145	10.4126	10.4686	10.5690	10.6673	10.7928
70	10.8407	10.9007	10.9664	11.0830	11.1482	11.2351	11.2762	11.3735	11.4839	11.5333
80	11.6290	11.6997	11.7687	11.8582	11.8913	11.9571	12.0394	12.0976	12.1919	12.2409
90	12.3296	12.3966	12.4797	12.5378	12.6137	12.6642	12.7652	12.7903	12.8537	12.9640
100	13.0037	13.0941	13.1269	13.1985	13.2913	13.3804	13.4086	13.5311	13.5829	13.6234
110	13.6613	13.7376	13.8059	13.9350	13.9172	14.0380	14.1042	14.0987	14.2076	14.2777
120	14.3198	14.3985	14.4913	14.5146	14.5647	14.6351	14.6814	14.7200	14.8068	14.8291
130	14.8856	15.0009	15.0640	15.0978	15.2113	15.2537	15.3043	15.3182	15.4045	15.4692
140	15.5684	15.5569	15.6026	15.6960	15.7548	15.8057	15.8796	15.9084	15.9695	16.0637
150	16.0250	16.1591	16.1842	16.2135	16.3247	16.3645	16.3650	16.4788	16.5002	16.5847
160	16.6284	16.7227	16.7260	16.7722	16.8428	16.8669	16.9529	16.9652	17.0920	17.1403
170	17.1848	17.2532	17.2374	17.2690	17.3788	17.4183	17.4660	17.4631	17.4805	17.5708
180	17.6851	17.7727	17.7898	17.8080	17.8780	17.8916	17.9307	18.0173	18.0537	18.0541
190	18.1476	18.1635	18.2517	18.3462	18.3461	18.3905	18.5120	18.4230	18.5393	18.6048

Tabelle C.20: Kritische Grenzen T_{95} des Buishand Tests für 100’000 Replikationen.

N	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
12	3.677	3.680	3.683	4.018	4.022	4.025	4.600	4.604	4.610
18	4.623	4.627	4.631	5.094	5.099	5.105	5.936	5.945	5.953
24	5.425	5.430	5.435	5.995	6.001	6.008	7.048	7.060	7.073
30	6.132	6.137	6.143	6.786	6.794	6.802	8.028	8.041	8.055
36	6.767	6.774	6.780	7.502	7.511	7.519	8.895	8.911	8.925
42	7.355	7.362	7.370	8.154	8.163	8.172	9.697	9.711	9.729
48	7.901	7.908	7.917	8.766	8.775	8.785	10.44	10.45	10.47
50	8.080	8.089	8.097	8.968	8.978	8.989	10.65	10.67	10.69
54	8.412	8.420	8.429	9.343	9.354	9.366	11.14	11.16	11.18
60	8.897	8.906	8.915	9.882	9.893	9.904	11.81	11.83	11.85
66	9.351	9.360	9.369	10.39	10.40	10.41	12.40	12.42	12.45
72	9.804	9.814	9.824	10.90	10.91	10.92	13.02	13.05	13.07
78	10.23	10.24	10.25	11.35	11.36	11.38	13.56	13.58	13.61
84	10.63	10.65	10.66	11.82	11.83	11.85	14.13	14.15	14.18
90	11.02	11.03	11.04	12.25	12.27	12.28	14.64	14.67	14.70
96	11.40	11.42	11.43	12.68	12.69	12.71	15.17	15.21	12.23
100	11.65	11.66	11.68	12.95	12.97	12.98	15.53	15.56	15.59
102	11.77	11.78	11.80	13.09	13.10	13.12	15.67	15.69	15.72
108	12.14	12.15	12.16	13.48	13.49	13.51	16.14	16.16	16.19
114	12.47	12.48	12.49	13.85	13.87	13.88	16.59	16.62	16.65
120	12.80	12.81	12.83	14.23	14.25	14.26	17.04	17.07	17.10
180	15.83	15.84	15.86	17.60	17.62	17.64	21.09	21.13	21.17
240	18.36	18.38	18.40	20.41	20.43	20.45	24.50	24.55	24.59
250	18.74	18.76	18.78	20.85	20.87	20.90	25.00	25.04	25.09
300	20.60	20.62	20.65	22.91	22.94	22.97	27.48	27.53	27.58
360	22.62	22.64	22.67	25.13	25.16	25.18	30.16	30.22	30.28
420	24.49	24.51	24.53	27.21	27.24	27.27	32.61	32.67	32.73
480	26.21	26.24	26.27	29.13	29.17	29.20	34.90	34.96	35.04
540	27.83	27.86	27.89	30.93	30.97	31.00	37.11	37.18	37.26
600	29.36	29.39	29.42	32.64	32.67	32.71	39.20	39.27	39.34
660	30.82	30.85	30.89	34.23	34.27	34.32	41.13	41.21	41.29
720	32.22	32.26	32.29	35.79	35.83	35.87	42.94	43.02	43.10
780	33.58	33.62	33.65	37.30	37.34	37.38	44.75	44.83	44.91
840	34.88	34.91	34.95	38.74	38.78	38.84	46.52	46.60	46.69
900	36.08	36.12	36.15	40.06	40.11	40.16	48.09	48.19	48.28
960	37.29	37.33	37.37	41.43	41.48	41.52	49.76	49.87	49.95
1000	38.08	38.11	38.15	42.29	42.34	42.39	50.65	50.74	50.84
1020	38.50	38.54	38.58	42.75	42.80	42.85	51.26	51.36	51.45
1080	39.61	39.65	39.69	44.01	44.06	44.11	52.69	52.78	52.89
1140	40.72	40.76	40.80	45.24	45.29	45.35	54.20	54.30	54.41
1200	41.77	41.81	41.85	46.41	46.46	46.52	55.62	55.73	55.84
6000	94.03	94.12	94.21	104.5	104.6	104.7	125.2	125.4	125.6

Tabelle C.21: Kritische Grenzen des Buishand Tests für 1’000’000 Replikationen.

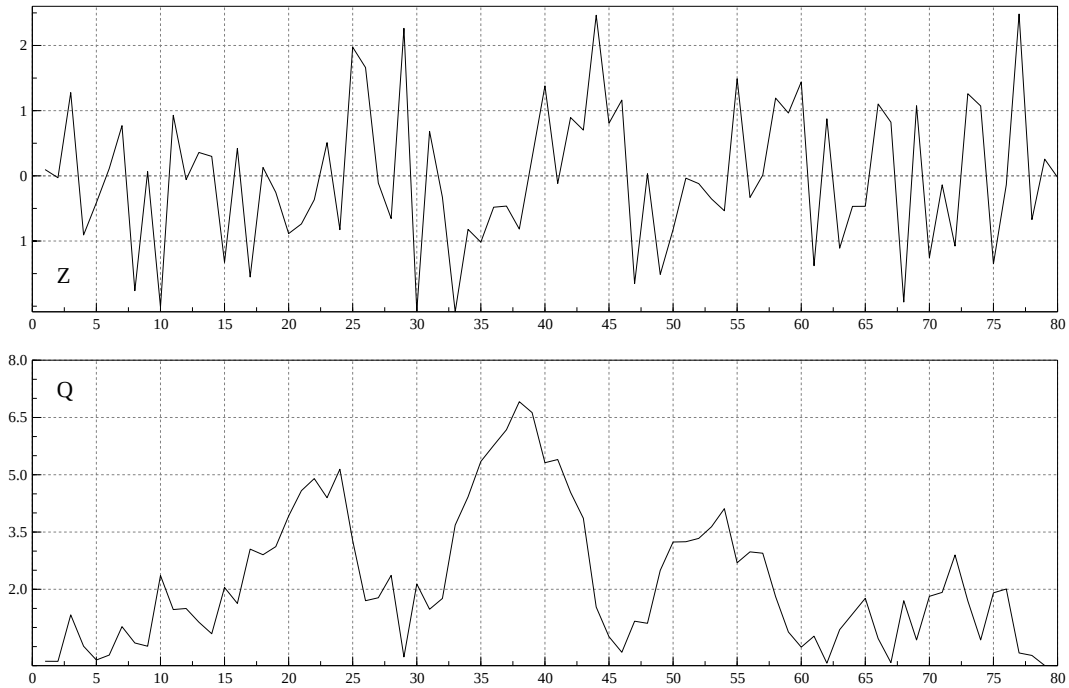


Abbildung C.54: Verlauf der Buishand Testgrösse für die Testreihe Z80

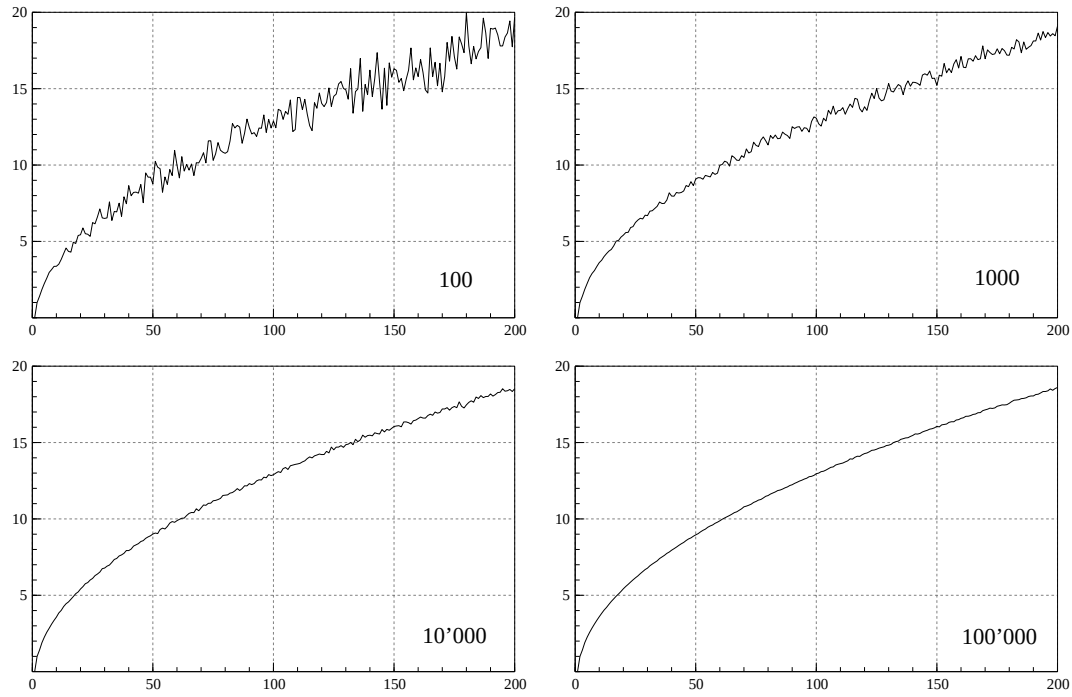


Abbildung C.55: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

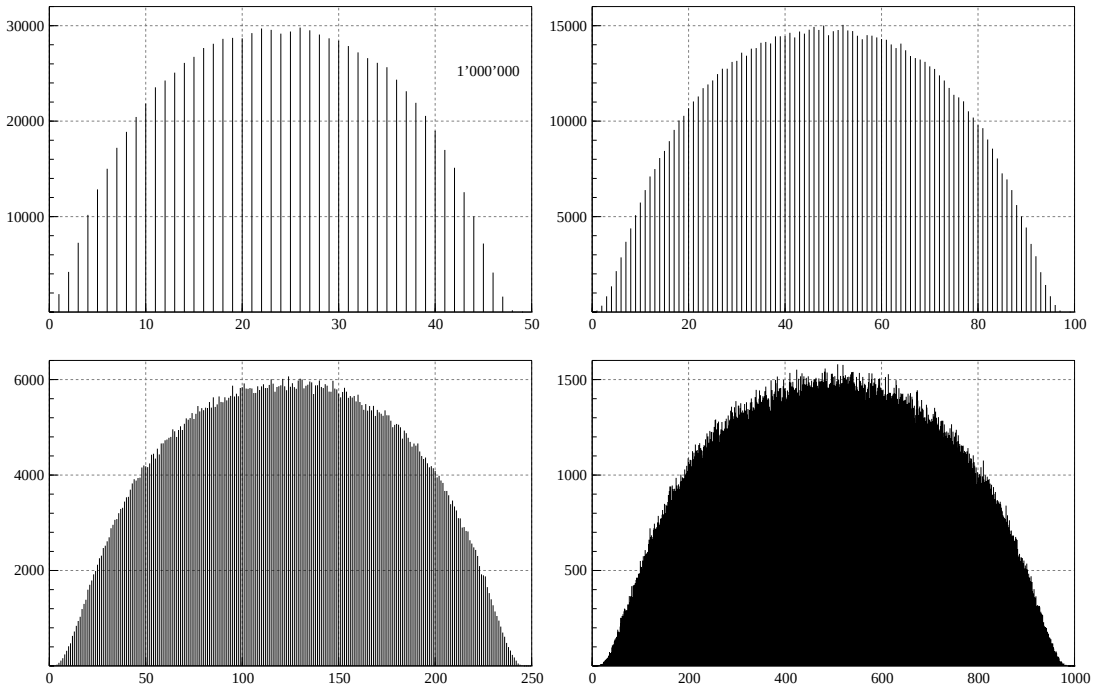


Abbildung C.56: Verteilung der Epochen der Maxima der Buishand Testgrösse mit $\phi = 0$

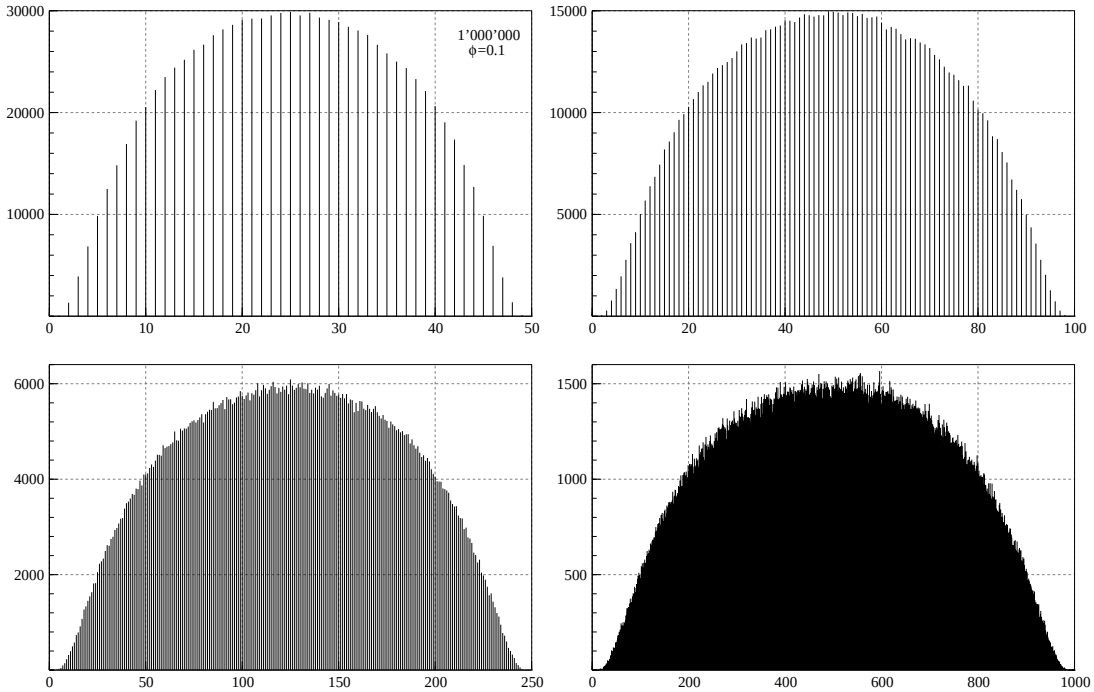


Abbildung C.57: Verteilung der Epochen der Maxima der Buishand Testgrösse mit $\phi = 0.1$

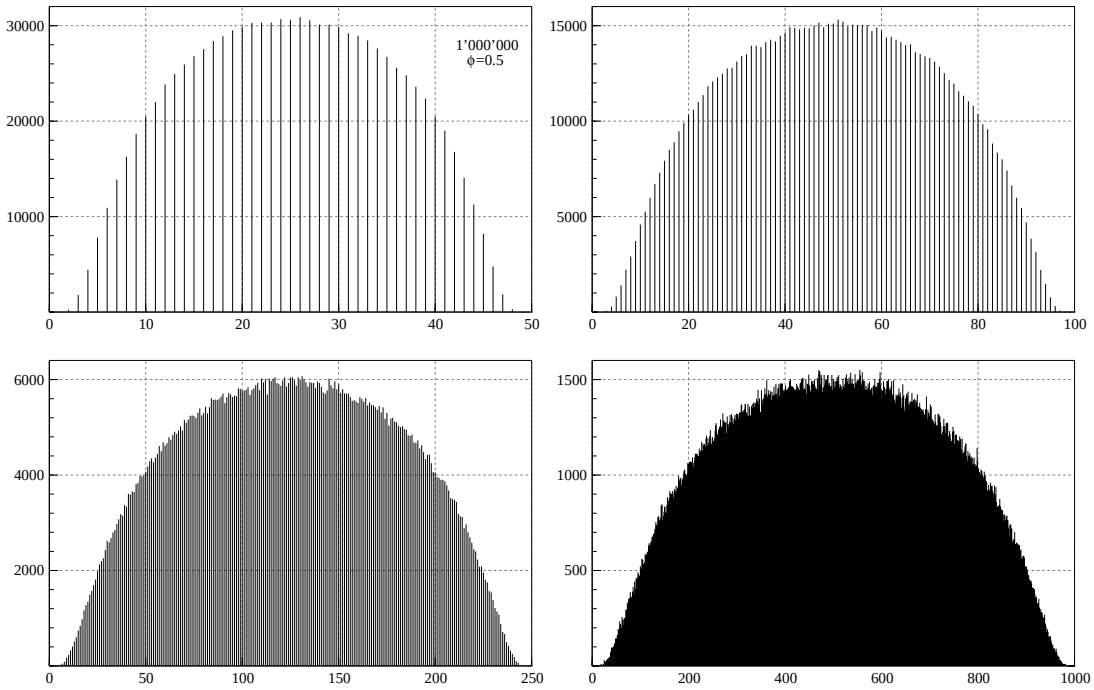


Abbildung C.58: Verteilung der Epochen der Maxima der Buishand Testgrösse mit $\phi = 0.5$

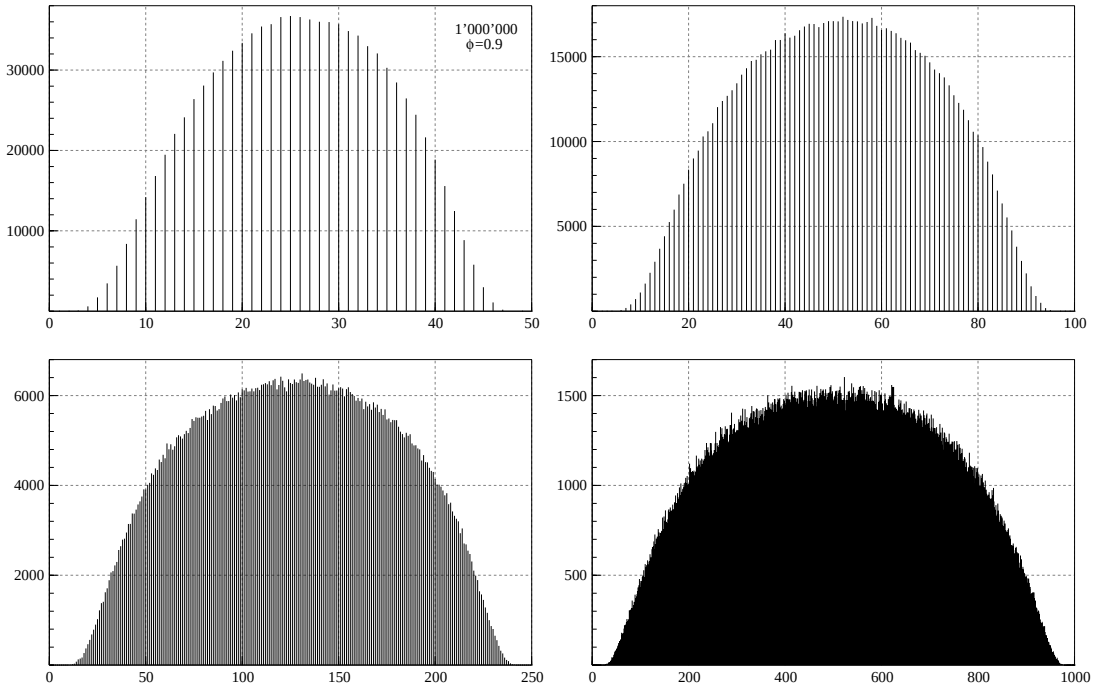


Abbildung C.59: Verteilung der Epochen der Maxima der Buishand Testgrösse mit $\phi = 0.9$

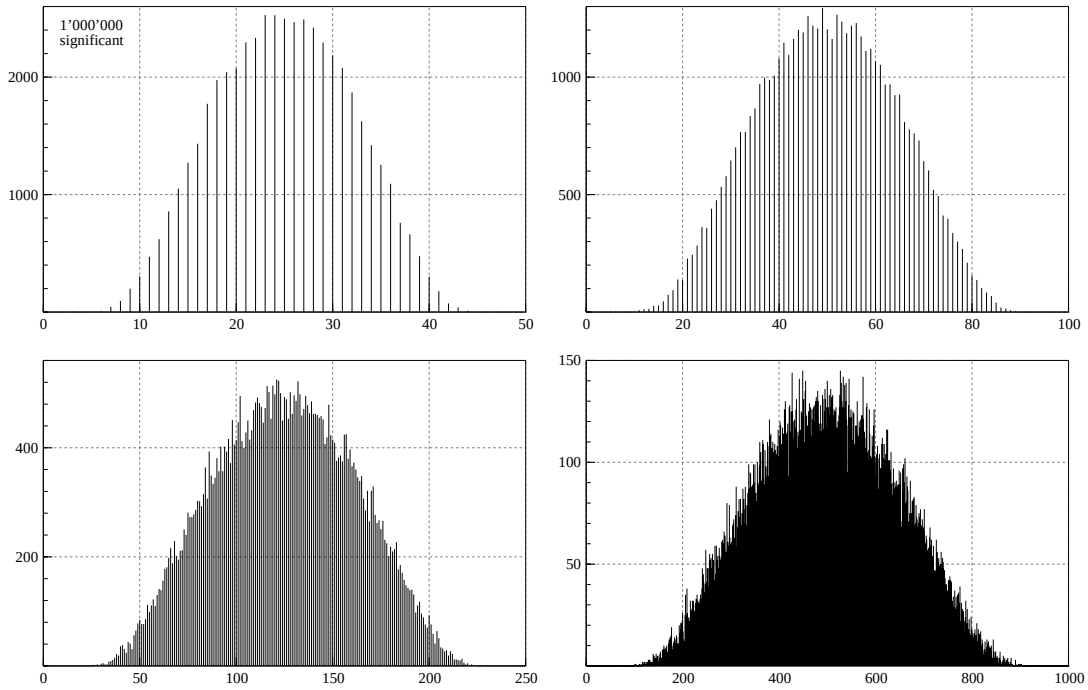


Abbildung C.60: Verteilung der Bruchepochen der Buishand Testgrösse mit $\phi = 0$

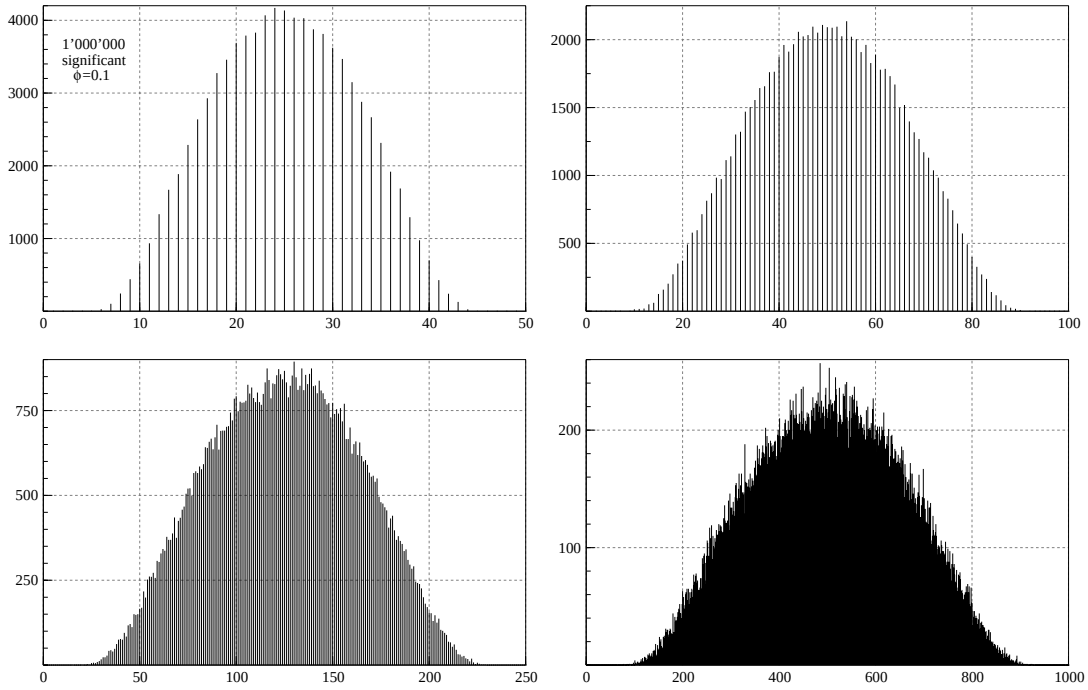


Abbildung C.61: Verteilung der Bruchepochen der Buishand Testgrösse mit $\phi = 0.1$

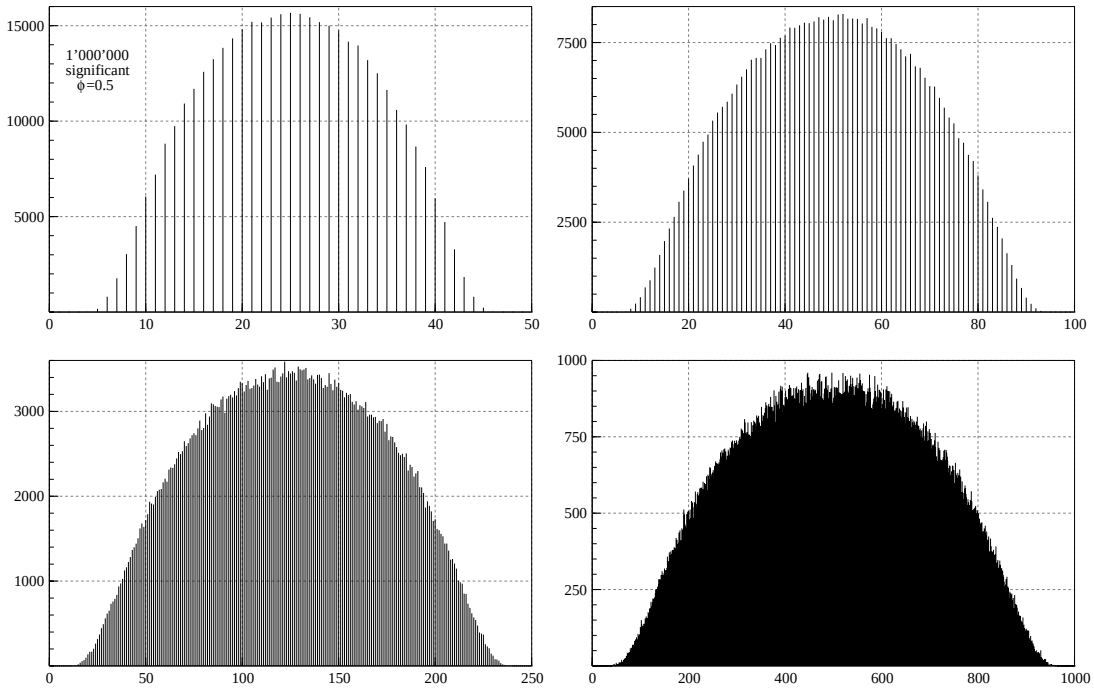


Abbildung C.62: Verteilung der Bruchepochen der Buishand Testgrösse mit $\phi = 0.5$

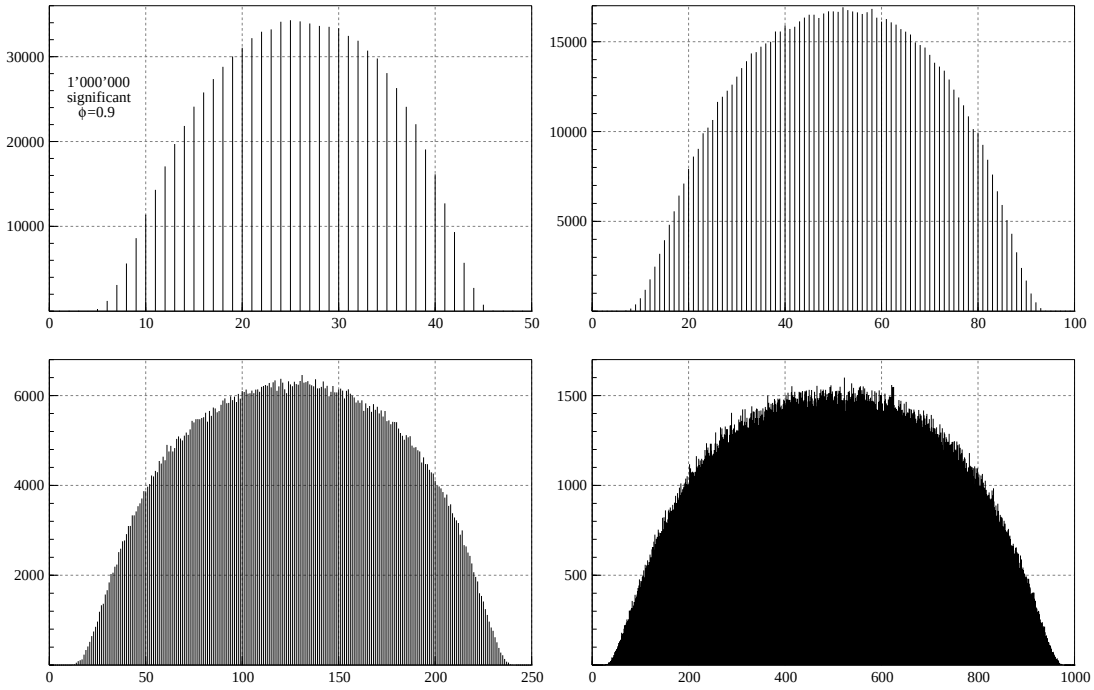


Abbildung C.63: Verteilung der Bruchepochen der Buishand Testgrösse mit $\phi = 0.9$

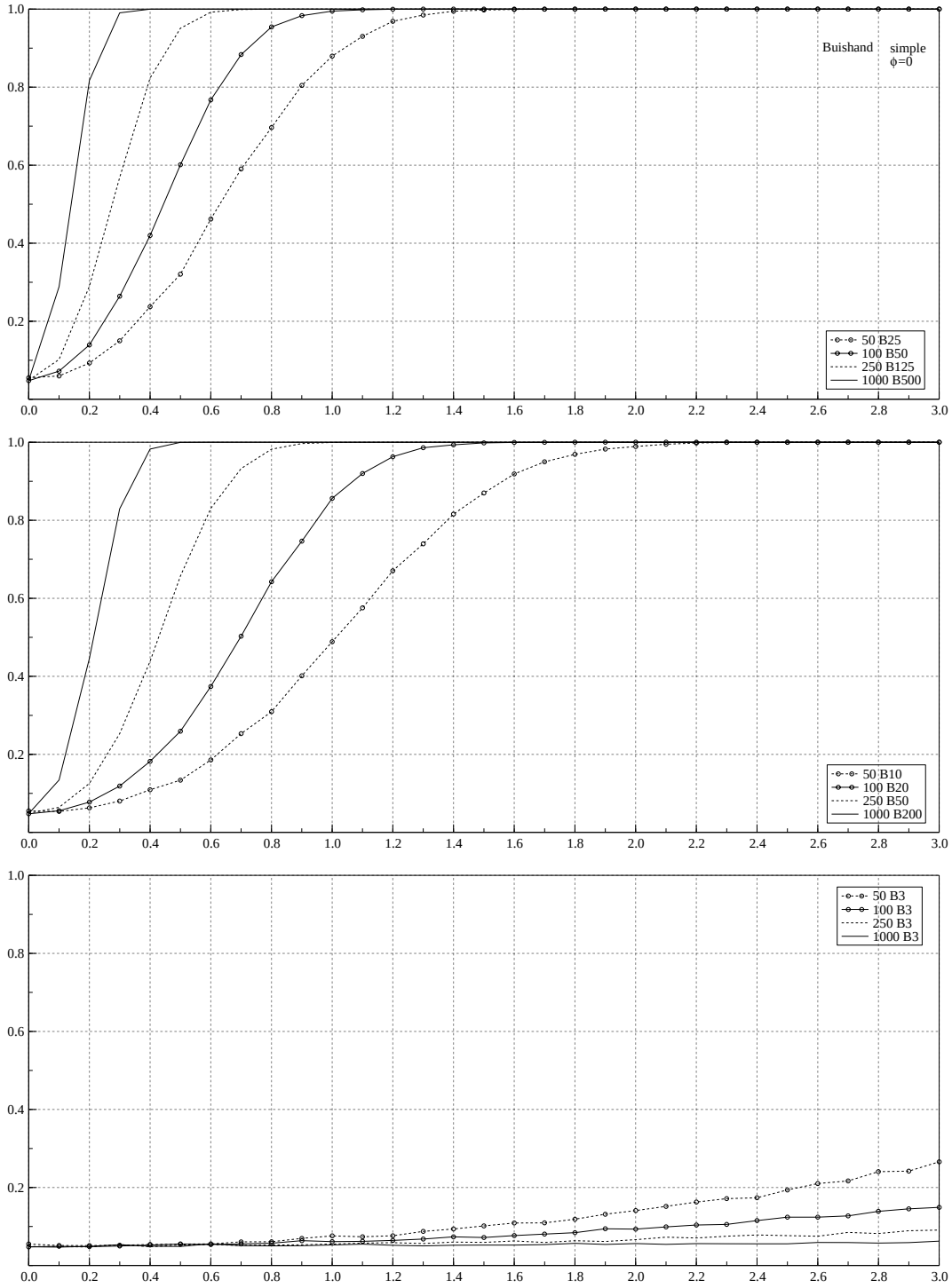


Abbildung C.64: Güte des Buishand Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

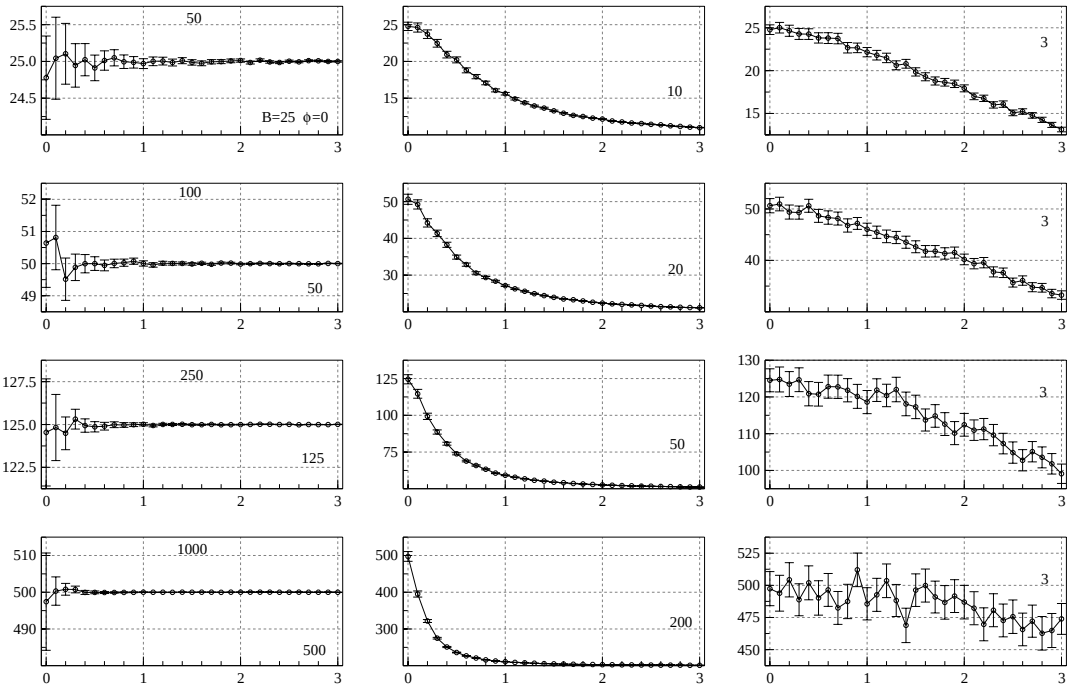


Abbildung C.65: Epochentreue des Buishand Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

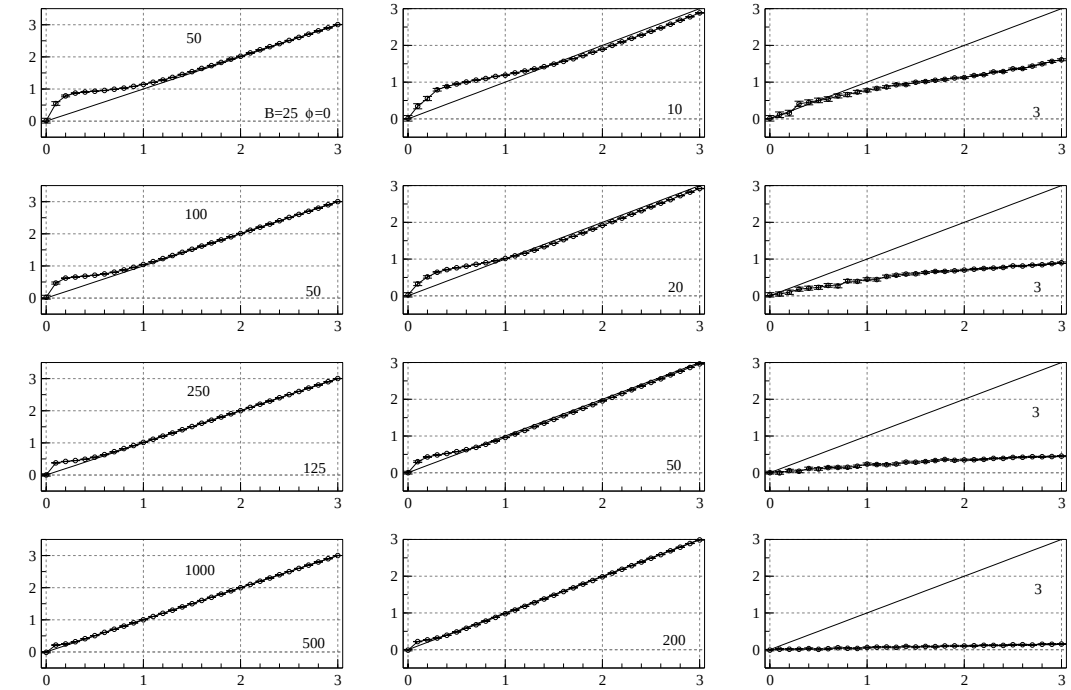


Abbildung C.66: Sprunghöhentreue des Buishand Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 250 und 1000 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.10 Likelihood Ratio Test 8

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
25	0	4.2909	4.2946	4.2983	4.6942	4.6987	4.7034	5.4470	5.4547	5.4635
50	0	4.7478	4.7518	4.7564	5.2160	5.2213	5.2274	6.1191	6.1295	6.1393
75	0	4.9900	4.9943	4.9985	5.4923	5.4988	5.5045	6.4797	6.4912	6.5033
100	0	5.1417	5.1464	5.1515	5.6624	5.6690	5.6755	6.7059	6.7174	6.7283
25	0.1	4.6955	4.6994	4.7032	5.1270	5.1314	5.1368	5.9187	5.9268	5.9371
50	0.1	5.1365	5.1408	5.1453	5.6411	5.6467	5.6522	6.6084	6.6189	6.6298
75	0.1	5.4231	5.4278	5.4330	5.9682	5.9742	5.9806	7.0312	7.0407	7.0526
100	0.1	5.6102	5.6151	5.6202	6.1747	6.1815	6.1880	7.2965	7.3084	7.3203
25	0.5	6.1516	6.1558	6.1602	6.6110	6.6161	6.6214	7.3501	7.3566	7.3635
50	0.5	7.1856	7.1917	7.1982	7.8283	7.8354	7.8429	8.9642	8.9771	8.9870
75	0.5	7.7895	7.7964	7.8036	8.5298	8.5388	8.5474	9.8973	9.9133	9.9274
100	0.5	8.1761	8.1833	8.1911	8.9792	8.9892	8.9989	10.479	10.493	10.507
25	0.9	7.9861	7.9891	7.9920	8.2600	8.2627	8.2654	8.6164	8.6195	8.6223
50	0.9	11.150	11.156	11.161	11.668	11.673	11.678	12.358	12.365	12.371
75	0.9	13.373	13.381	13.388	14.129	14.136	14.144	15.188	15.199	15.209
100	0.9	14.956	14.965	14.975	15.919	15.930	15.940	17.335	17.347	17.361

Tabelle C.22: Kritische Grenzen des LR8 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	2.2576	2.7667	3.1629	3.4568	3.6724	3.8362	3.9596	4.0727	4.1643
10	4.2400	4.3187	4.3585	4.4273	4.4830	4.5241	4.5646	4.6012	4.6387	4.6857
20	4.7119	4.7346	4.7660	4.7977	4.8233	4.8359	4.8687	4.8979	4.9242	4.9181
30	4.9578	4.9765	4.9895	5.0060	5.0321	5.0445	5.0576	5.0816	5.0928	5.1009
40	5.1189	5.1339	5.1480	5.1635	5.1730	5.1910	5.2173	5.2107	5.2275	5.2334
50	5.2348	5.2648	5.2817	5.2687	5.3056	5.3112	5.3118	5.3267	5.3286	5.3440
60	5.3512	5.3642	5.3746	5.3869	5.3951	5.4168	5.4043	5.4204	5.4372	5.4542
70	5.4567	5.4544	5.4488	5.4872	5.4890	5.4903	5.4907	5.5071	5.5195	5.5144
80	5.5372	5.5396	5.5622	5.5485	5.5351	5.5571	5.5610	5.5630	5.5828	5.5673
90	5.5883	5.5979	5.6035	5.6021	5.6101	5.6267	5.6388	5.6246	5.6232	5.6442
100	5.6412	5.6535	5.6552	5.6653	5.6640	5.6920	5.6811	5.7056	5.7096	5.6950
110	5.6952	5.7158	5.7070	5.7428	5.7116	5.7381	5.7440	5.7236	5.7395	5.7514
120	5.7621	5.7742	5.7763	5.7683	5.7764	5.7839	5.7913	5.7775	5.8003	5.7845
130	5.7902	5.8164	5.8290	5.8089	5.8312	5.8387	5.8382	5.8389	5.8401	5.8450
140	5.8576	5.8482	5.8441	5.8667	5.8741	5.8687	5.8853	5.8797	5.8838	5.8997
150	5.8779	5.8924	5.8851	5.8919	5.9075	5.9227	5.8971	5.9105	5.9118	5.9218
160	5.9191	5.9410	5.9348	5.9246	5.9541	5.9308	5.9392	5.9497	5.9602	5.9720
170	5.9781	5.9756	5.9581	5.9627	5.9864	5.9843	5.9834	5.9723	5.9632	5.9823
180	6.0009	6.0112	6.0087	6.0023	6.0123	6.0068	6.0060	6.0251	6.0209	6.0007
190	6.0153	6.0218	6.0370	6.0484	6.0396	6.0394	6.0700	6.0283	6.0564	6.0510

Tabelle C.23: Kritische Grenzen T_{95} des LR8 Tests für 100’000 Replikationen.

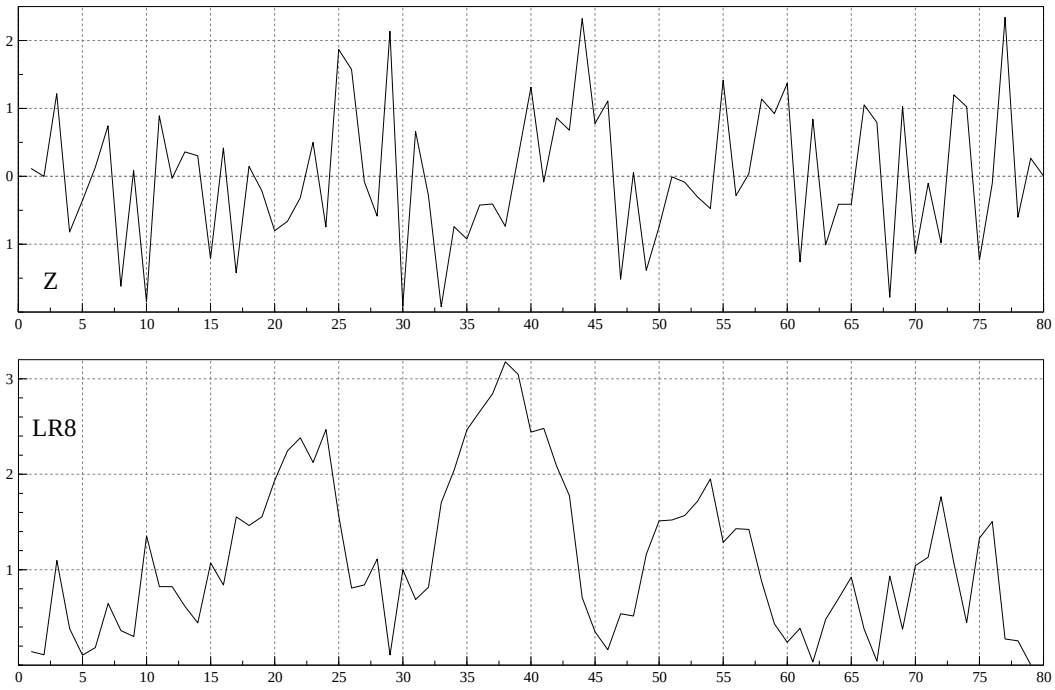


Abbildung C.67: Verlauf der LR8 Testgrösse für die Testreihe Z80

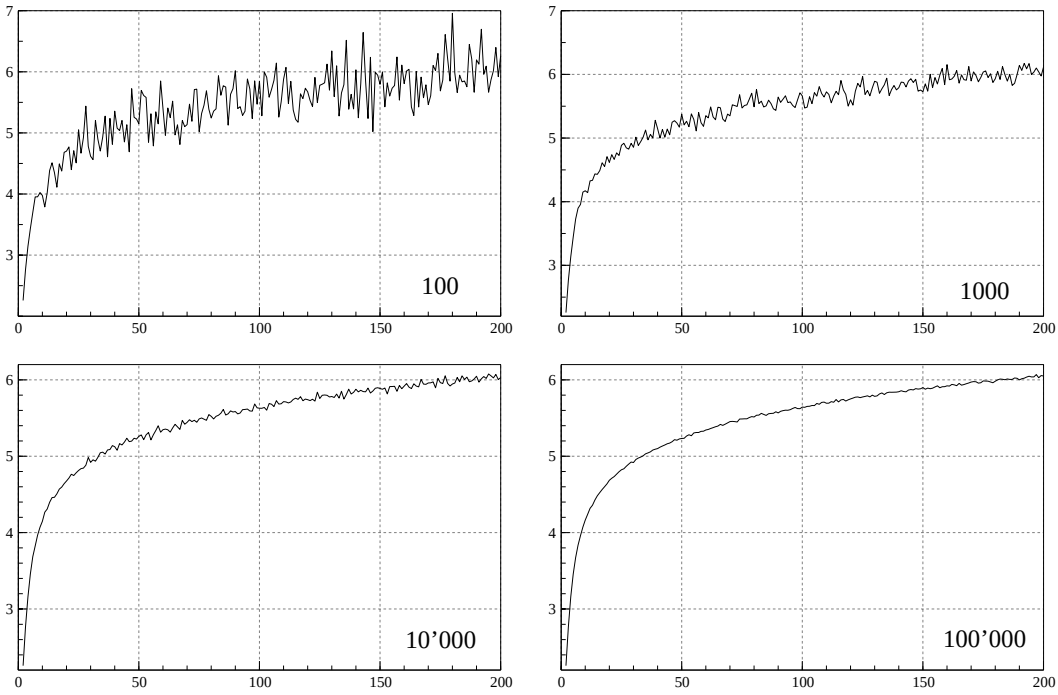


Abbildung C.68: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

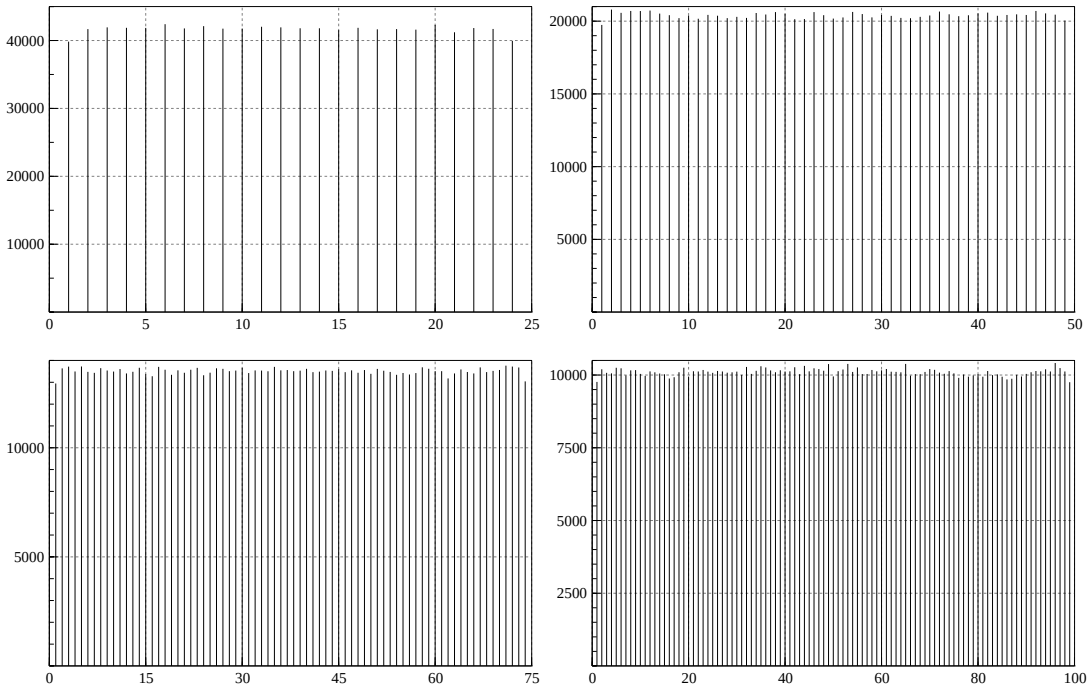


Abbildung C.69: Verteilung der Epochen der Maxima der LR8 Testgrösse mit $\phi = 0$

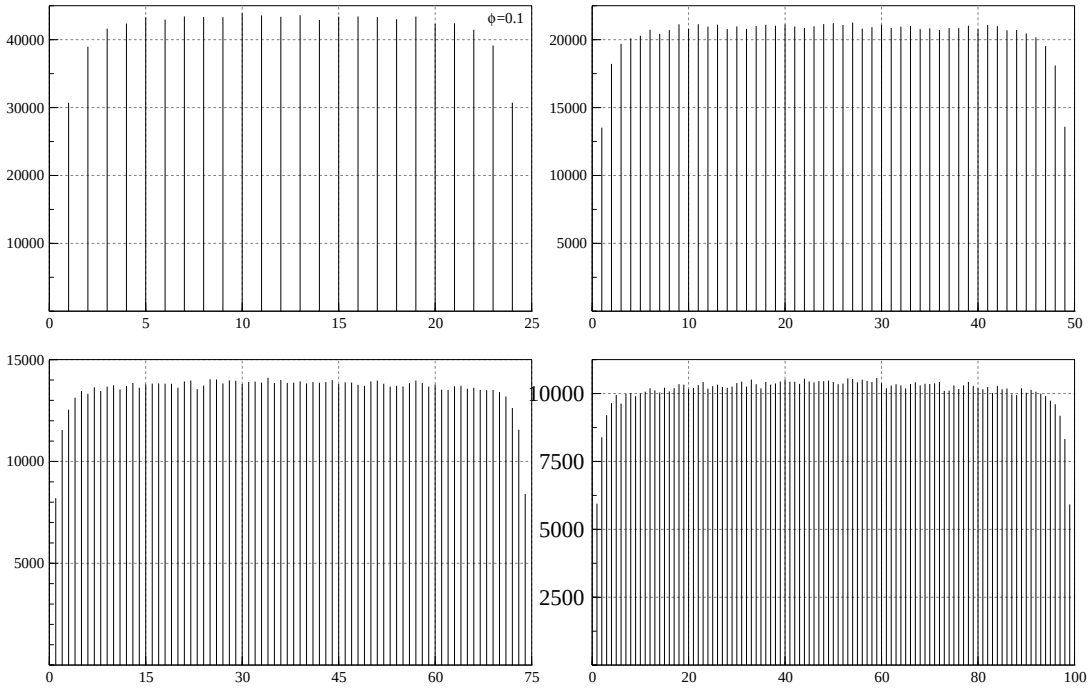


Abbildung C.70: Verteilung der Epochen der Maxima der LR8 Testgrösse mit $\phi = 0.1$

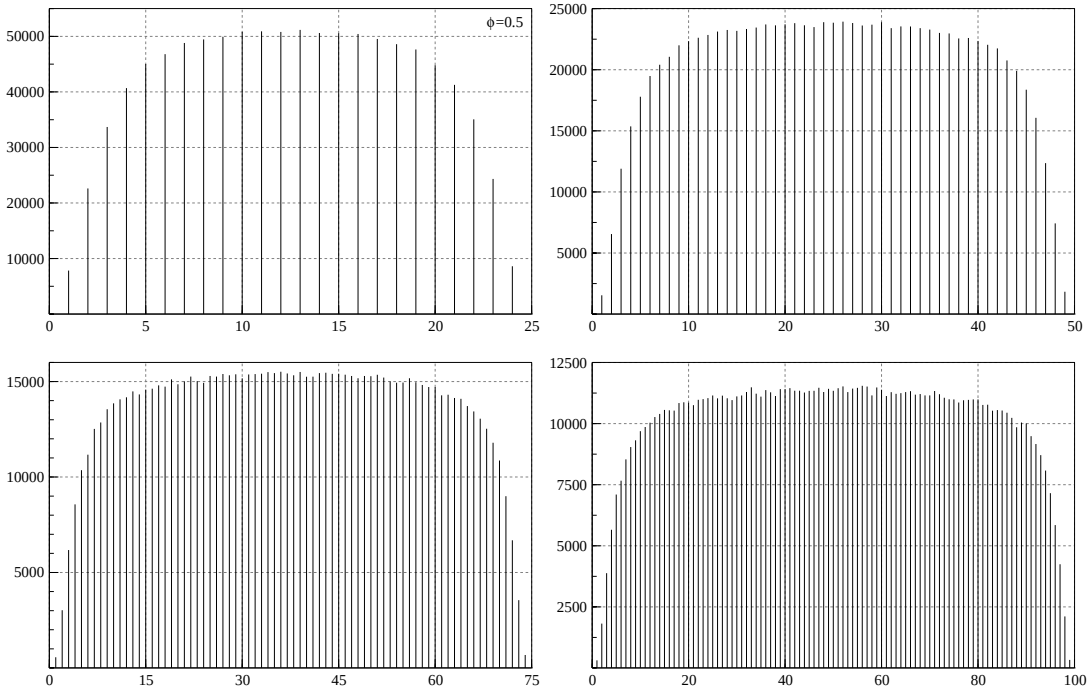


Abbildung C.71: Verteilung der Epochen der Maxima der LR8 Testgrösse mit $\phi = 0.5$

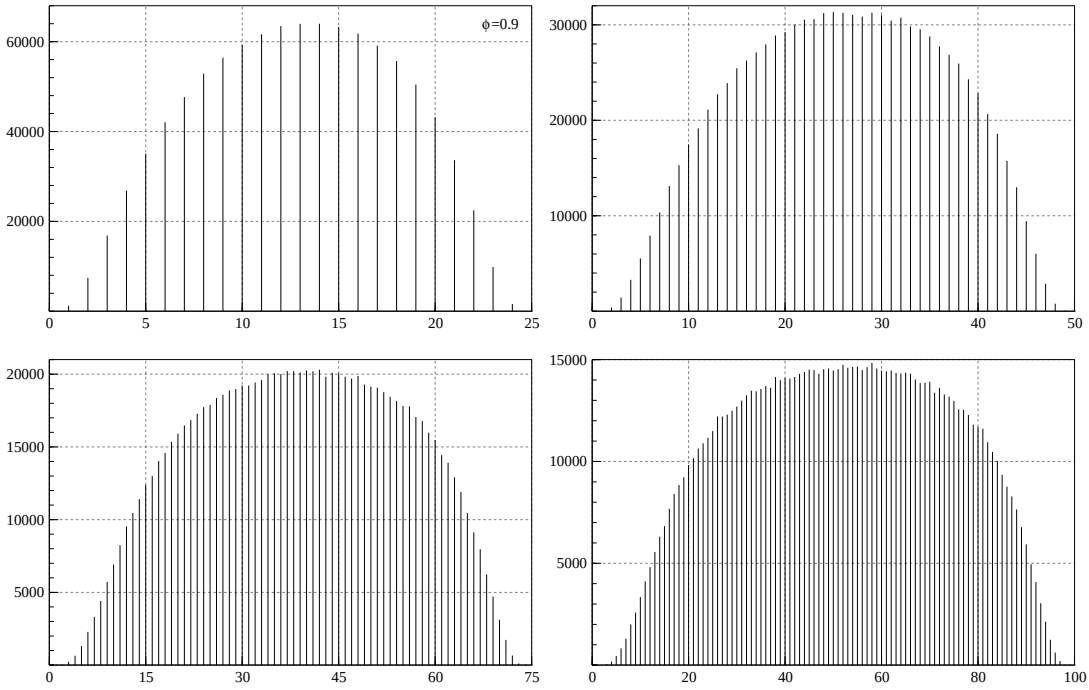


Abbildung C.72: Verteilung der Epochen der Maxima der LR8 Testgrösse mit $\phi = 0.9$

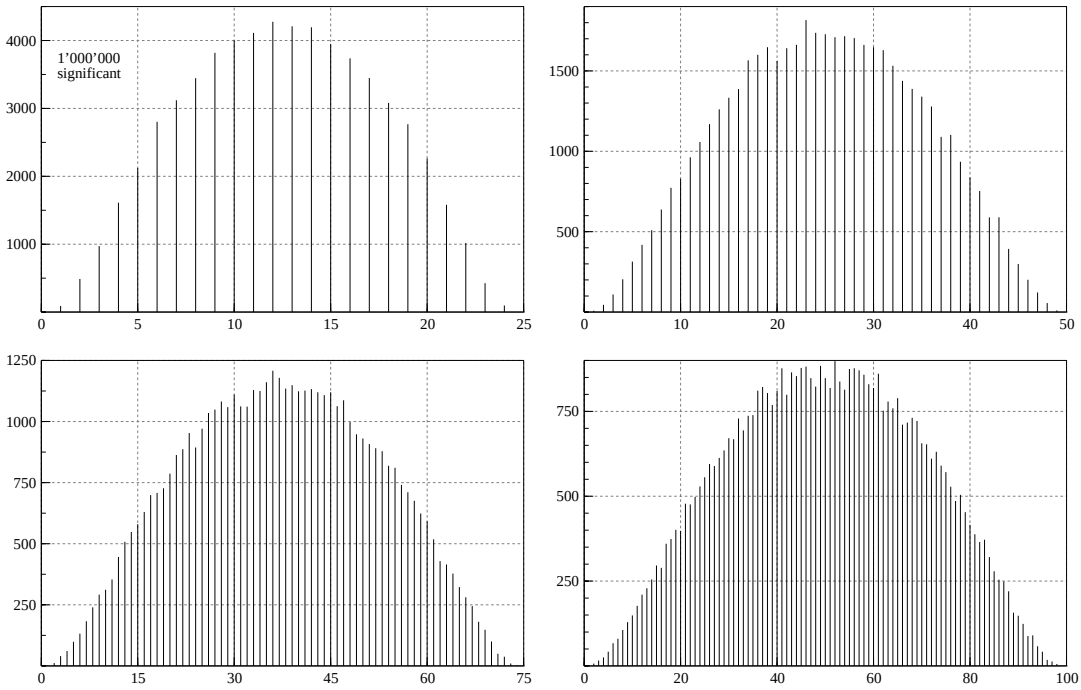


Abbildung C.73: Verteilung der Bruchepochen der LR8 Testgrösse mit $\phi = 0$

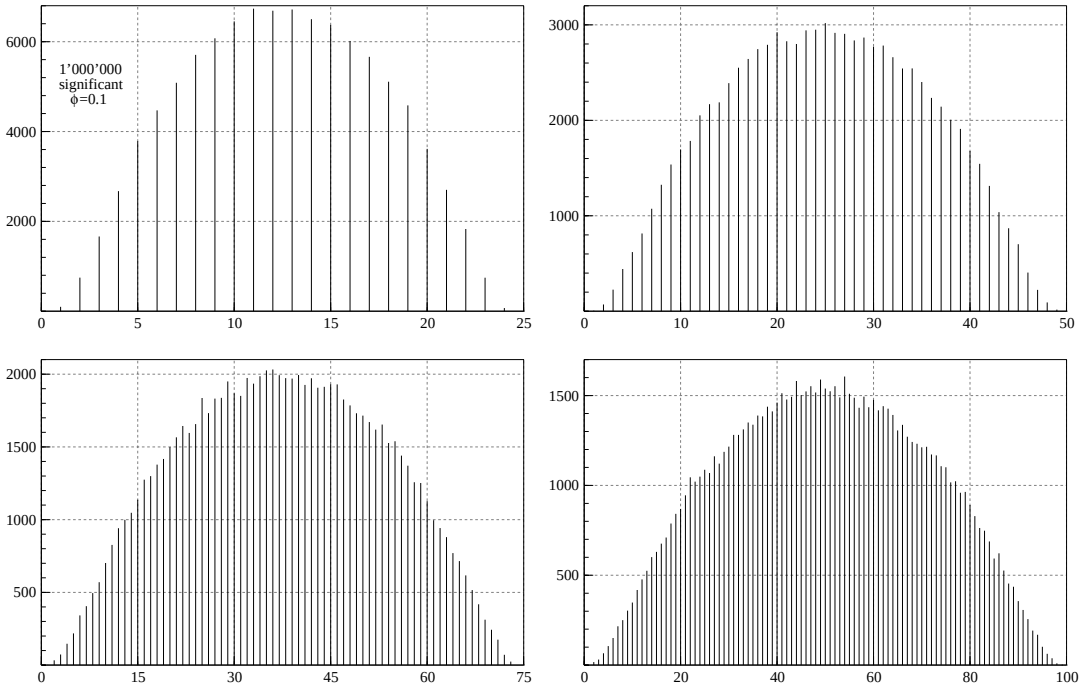


Abbildung C.74: Verteilung der Bruchepochen der LR8 Testgrösse mit $\phi = 0.1$

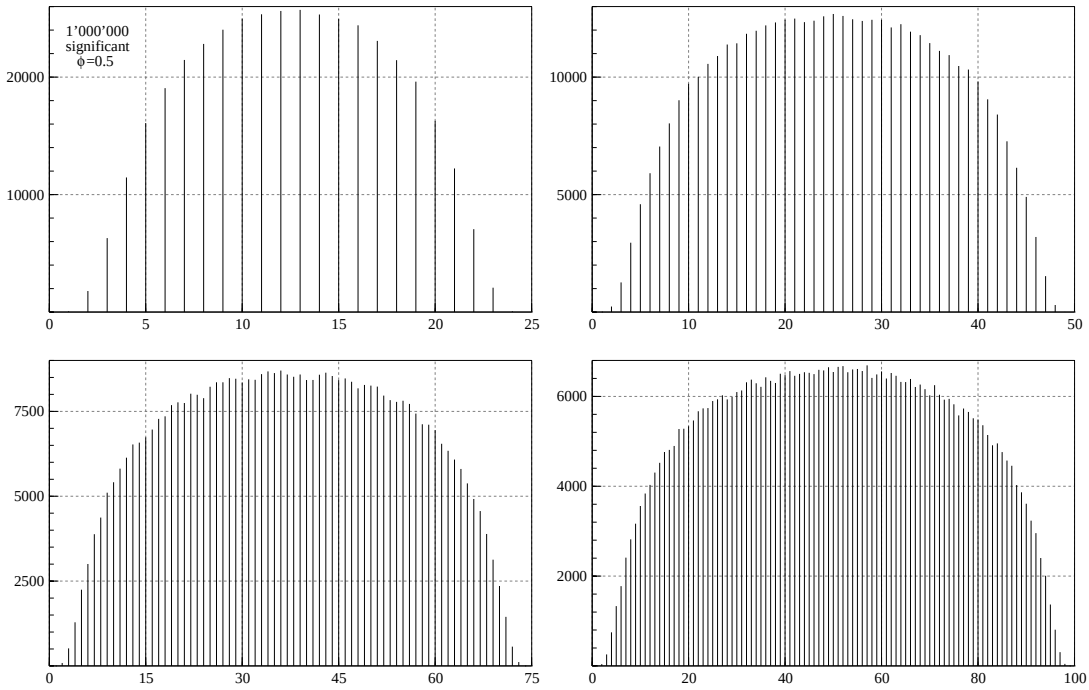


Abbildung C.75: Verteilung der Bruchepochen der LR8 Testgrösse mit $\phi = 0.5$

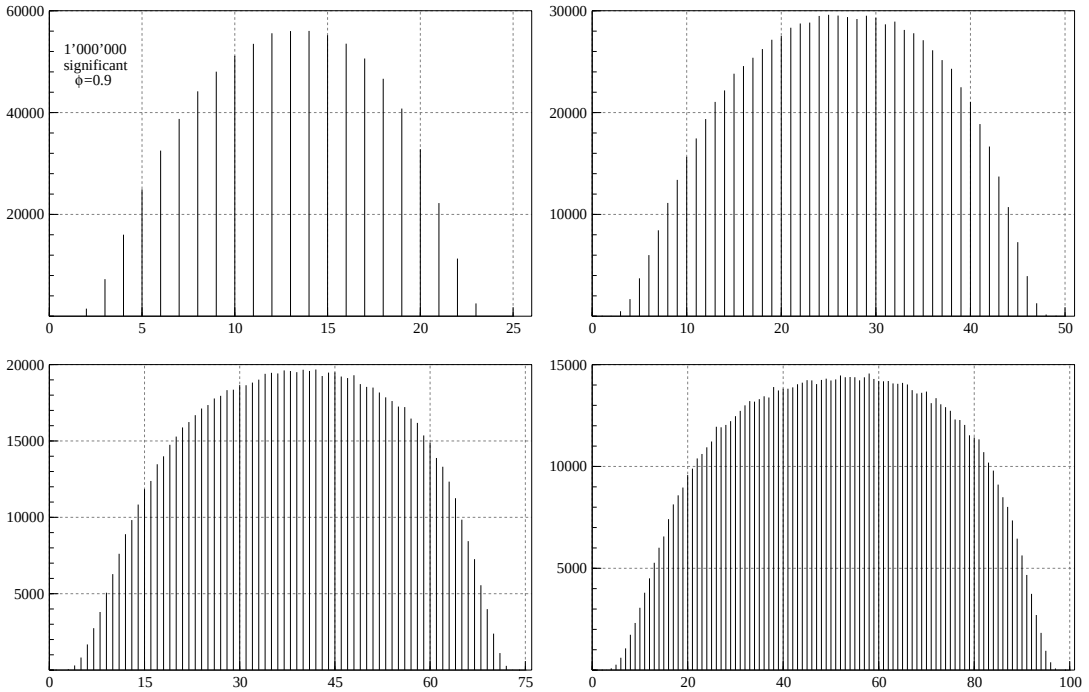


Abbildung C.76: Verteilung der Bruchepochen der LR8 Testgrösse mit $\phi = 0.9$

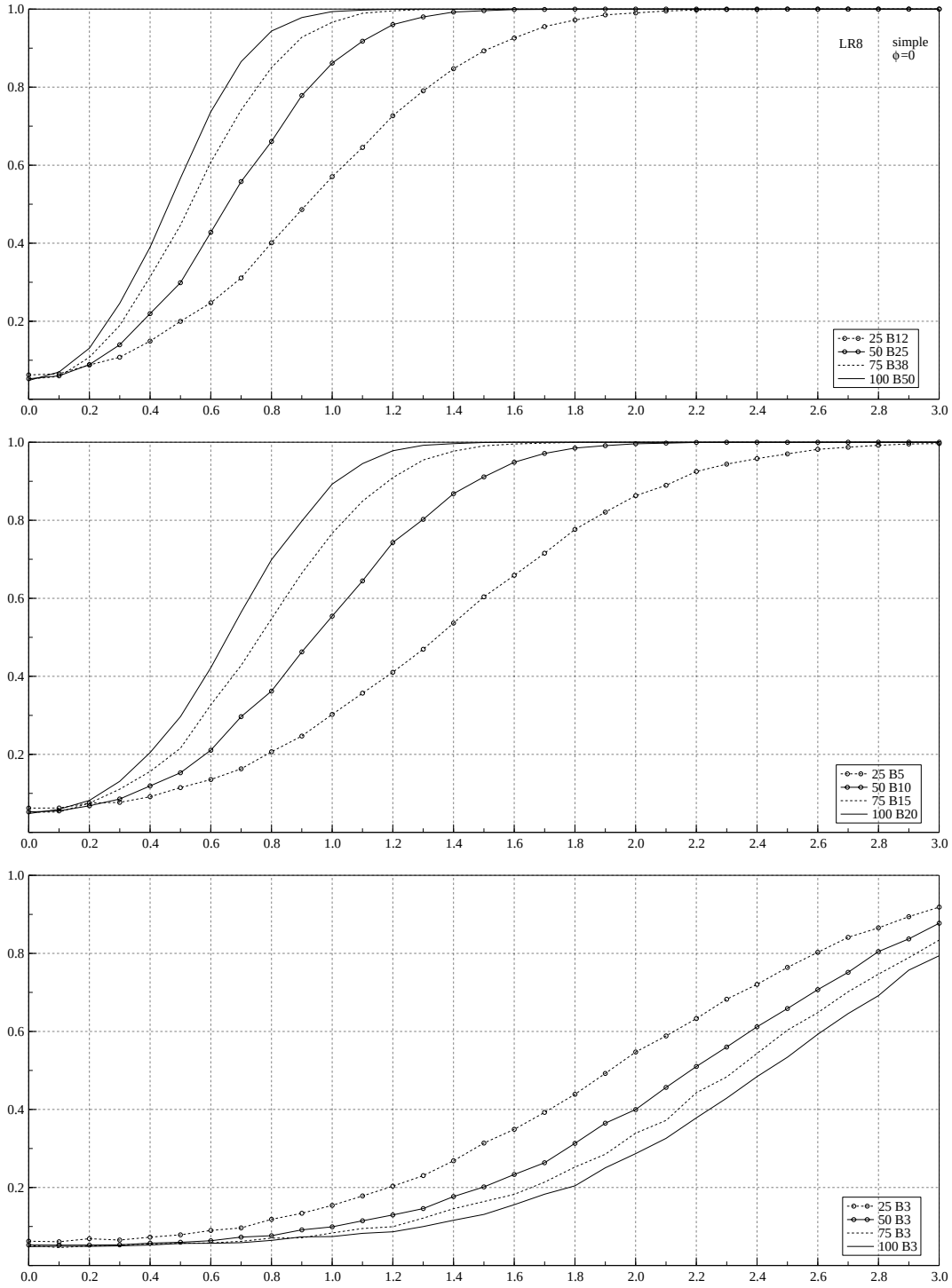


Abbildung C.77: Güte des LR8 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

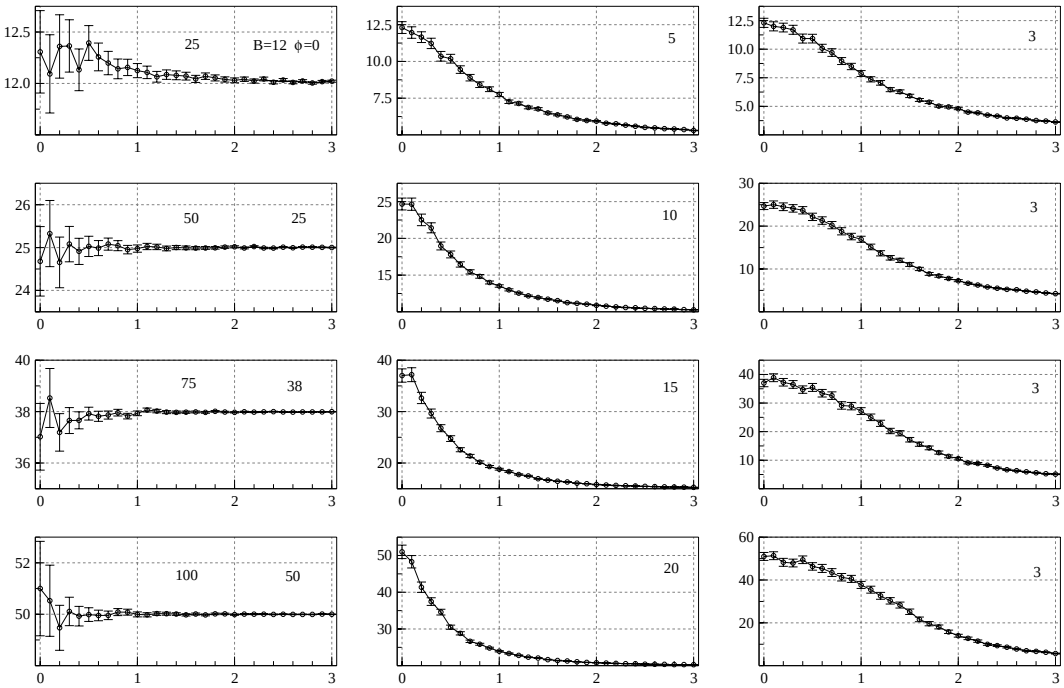


Abbildung C.78: Epochentreue des LR8 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

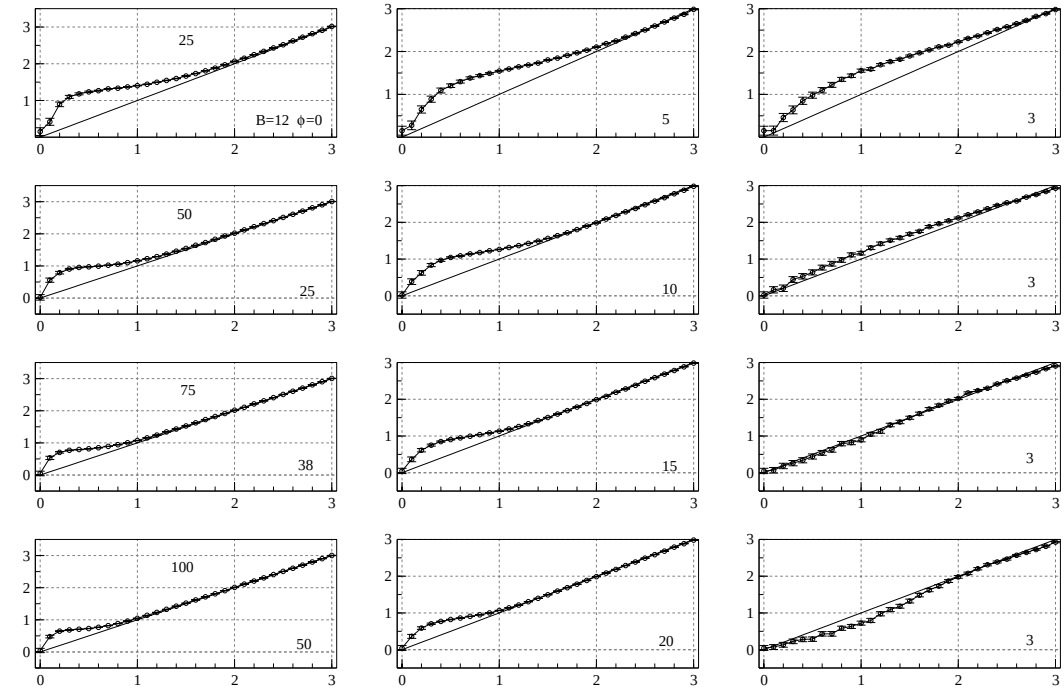


Abbildung C.79: Sprunghöhentreue des LR8 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.11 Likelihood Ratio Test 9

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
25	0	2.7068	2.7090	2.7111	2.9346	2.9372	2.9399	3.3564	3.3606	3.3651
50	0	2.9734	2.9759	2.9781	3.2349	3.2378	3.2409	3.7397	3.7450	3.7503
75	0	3.1136	3.1161	3.1186	3.3903	3.3934	3.3969	3.9369	3.9432	3.9493
100	0	3.2600	3.2624	3.2652	3.5518	3.5552	3.5585	4.1355	4.1417	4.1483
25	0.1	2.8621	2.8643	2.8664	3.0988	3.1014	3.1041	3.5285	3.5334	3.5385
50	0.1	3.1921	3.1946	3.1970	3.4727	3.4760	3.4795	4.0144	4.0196	4.0251
75	0.1	3.3783	3.3813	3.3839	3.6852	3.6887	3.6923	4.2842	4.2907	4.2967
100	0.1	3.5340	3.5367	3.5393	3.8584	3.8622	3.8663	4.5043	4.5121	4.5195
25	0.5	3.6583	3.6608	3.6633	3.9100	3.9123	3.9152	4.3074	4.3108	4.3146
50	0.5	4.3712	4.3745	4.3775	4.7336	4.7377	4.7419	5.3680	5.3740	5.3808
75	0.5	4.7700	4.7739	4.7778	5.1933	5.1976	5.2025	5.9689	5.9769	5.9847
100	0.5	5.0826	5.0871	5.0915	5.5502	5.5550	5.5604	6.4218	6.4314	6.4406
25	0.9	4.6643	4.6658	4.6674	4.8082	4.8096	4.8110	4.9945	4.9962	4.9978
50	0.9	6.6322	6.6351	6.6381	6.9147	6.9175	6.9204	7.2933	7.2964	7.2997
75	0.9	8.0181	8.0225	8.0266	8.4412	8.4454	8.4496	9.0378	9.0429	9.0479
100	0.9	9.1248	9.1302	9.1358	9.6831	9.6889	9.6946	10.501	10.509	10.517

Tabelle C.24: Kritische Grenzen des LR9 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	0.0000	1.3819	1.6947	1.9324	2.1006	2.2263	2.3267	2.3978	2.4659	2.5202
10	2.5655	2.6097	2.6427	2.6785	2.7143	2.7424	2.7631	2.7875	2.8151	2.8425
20	2.8591	2.8791	2.9016	2.9220	2.9376	2.9458	2.9678	2.9873	3.0047	3.0124
30	3.0267	3.0445	3.0574	3.0666	3.0830	3.0984	3.1032	3.1176	3.1282	3.1358
40	3.1478	3.1552	3.1734	3.1751	3.1899	3.1943	3.2147	3.2170	3.2327	3.2346
50	3.2384	3.2542	3.2644	3.2595	3.2793	3.2881	3.2978	3.3039	3.3019	3.3138
60	3.3223	3.3353	3.3394	3.3443	3.3519	3.3575	3.3583	3.3728	3.3762	3.3943
70	3.3913	3.3907	3.3966	3.4108	3.4189	3.4191	3.4220	3.4329	3.4360	3.4429
80	3.4580	3.4577	3.4694	3.4611	3.4566	3.4710	3.4700	3.4812	3.4856	3.4869
90	3.4935	3.4986	3.5041	3.5088	3.5040	3.5198	3.5248	3.5206	3.5157	3.5326
100	3.5292	3.5403	3.5390	3.5347	3.5466	3.5622	3.5562	3.5642	3.5665	3.5657
110	3.5673	3.5754	3.5668	3.5912	3.5753	3.5818	3.5976	3.5824	3.5880	3.6040
120	3.6086	3.6096	3.6066	3.6137	3.6121	3.6155	3.6183	3.6152	3.6280	3.6201
130	3.6234	3.6377	3.6416	3.6339	3.6414	3.6418	3.6440	3.6453	3.6427	3.6523
140	3.6533	3.6504	3.6494	3.6561	3.6632	3.6603	3.6642	3.6612	3.6686	3.6725
150	3.6654	3.6644	3.6651	3.6676	3.6805	3.6863	3.6809	3.6778	3.6840	3.6888
160	3.6840	3.6920	3.6894	3.6848	3.6911	3.6844	3.6886	3.6875	3.6966	3.7061
170	3.7021	3.6978	3.6948	3.7017	3.7025	3.7034	3.7044	3.6949	3.7038	3.7000
180	3.7148	3.7114	3.7131	3.7032	3.7185	3.7048	3.7102	3.7172	3.7185	3.7029
190	3.7153	3.7202	3.7170	3.7237	3.7232	3.7167	3.7297	3.7051	3.7255	3.7157

Tabelle C.25: Kritische Grenzen T_{95} des LR9 Tests für 100’000 Replikationen.

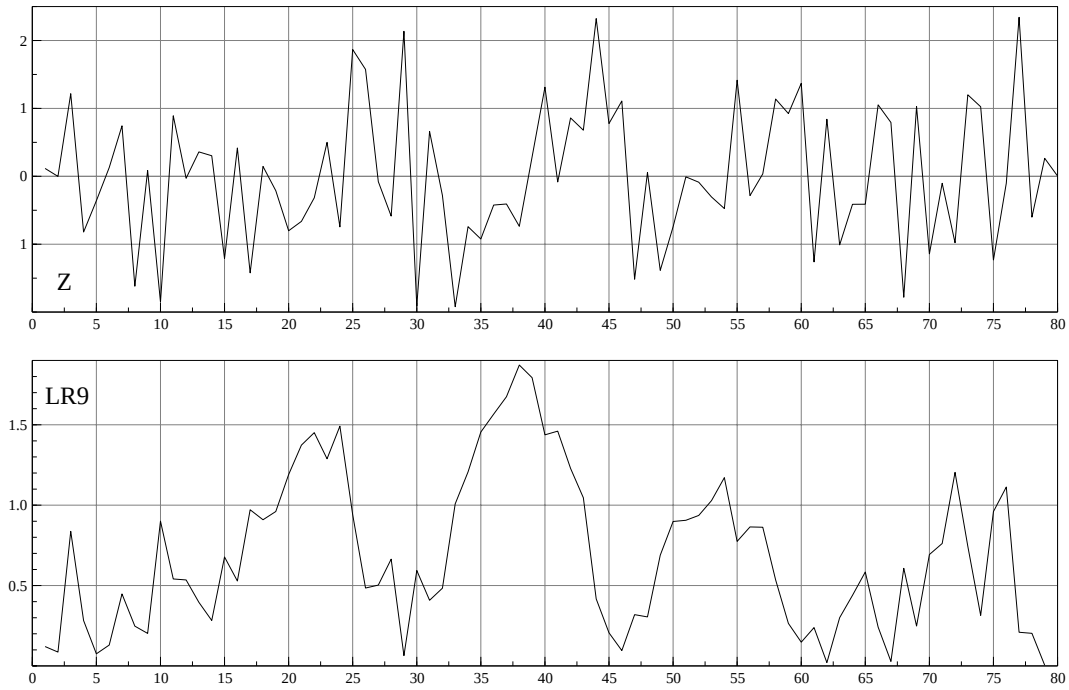


Abbildung C.80: Verlauf der LR9 Testgrösse für die Testreihe Z80

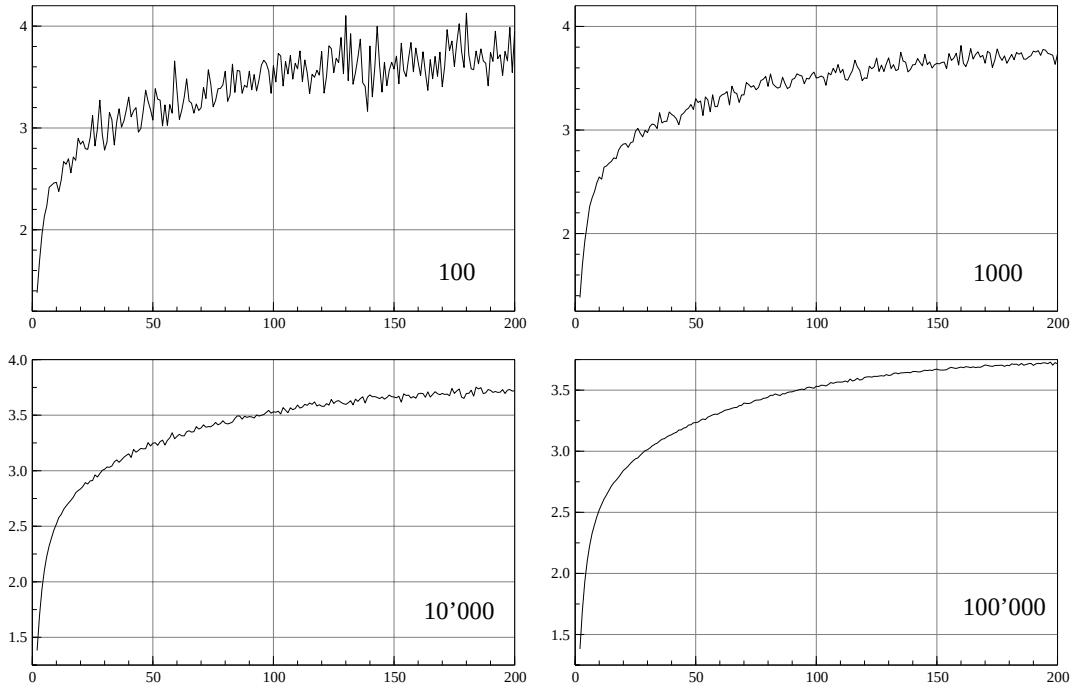


Abbildung C.81: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

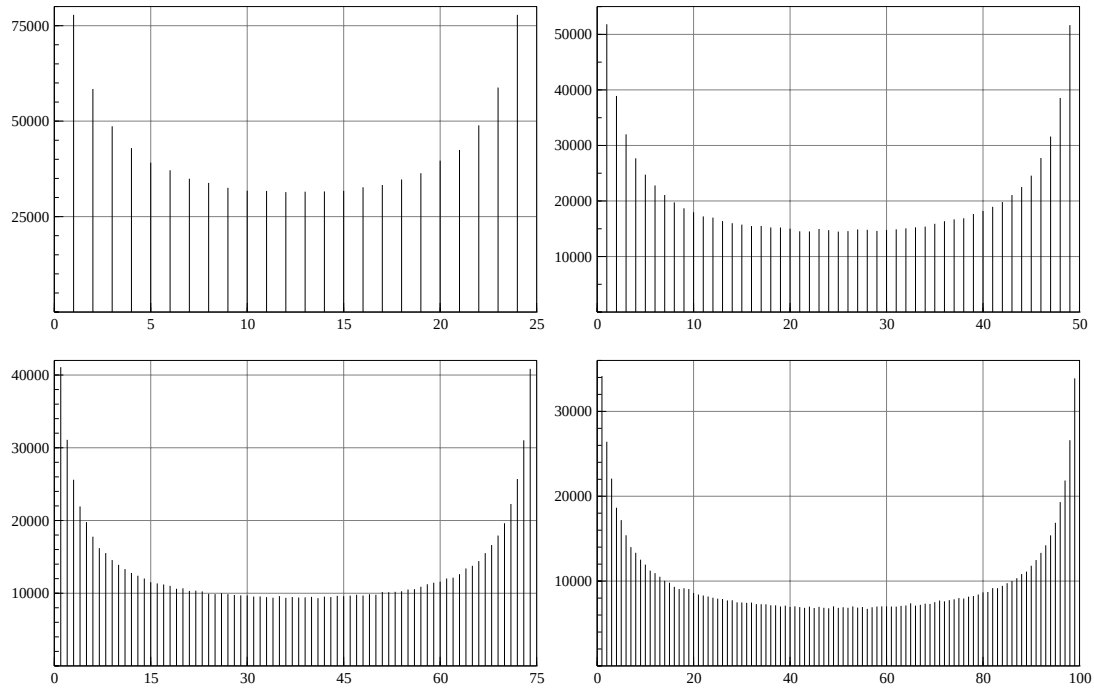


Abbildung C.82: Verteilung der Epochen der Maxima der LR9 Testgrösse mit $\phi = 0$

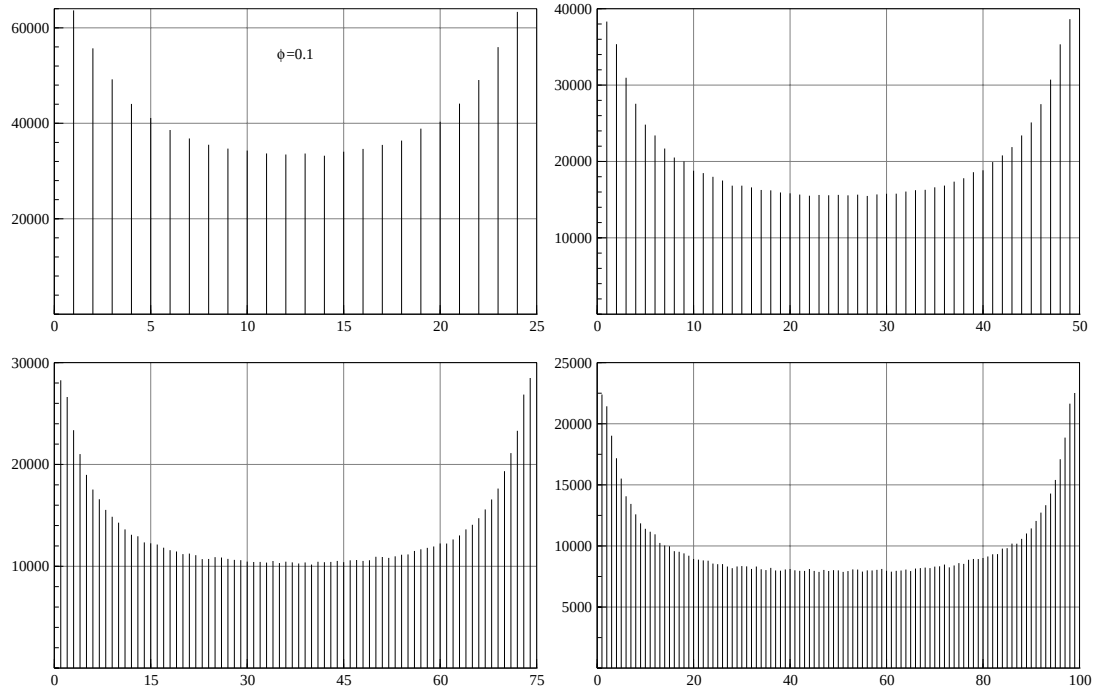


Abbildung C.83: Verteilung der Epochen der Maxima der LR9 Testgrösse mit $\phi = 0.1$

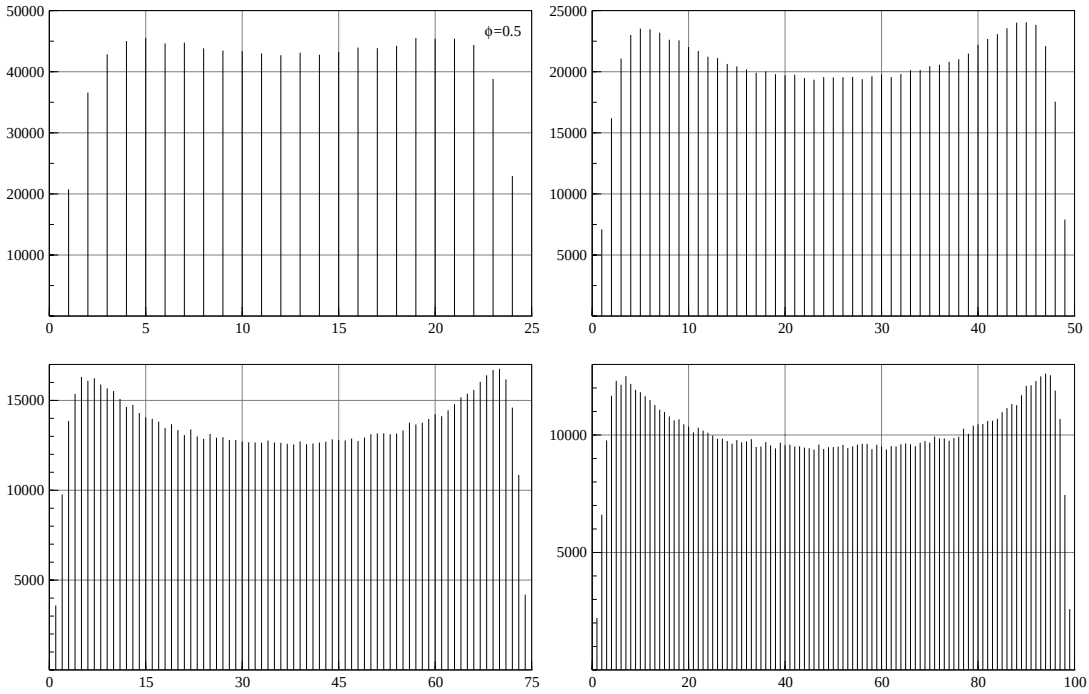


Abbildung C.84: Verteilung der Epochen der Maxima der LR9 Testgrösse mit $\phi = 0.5$

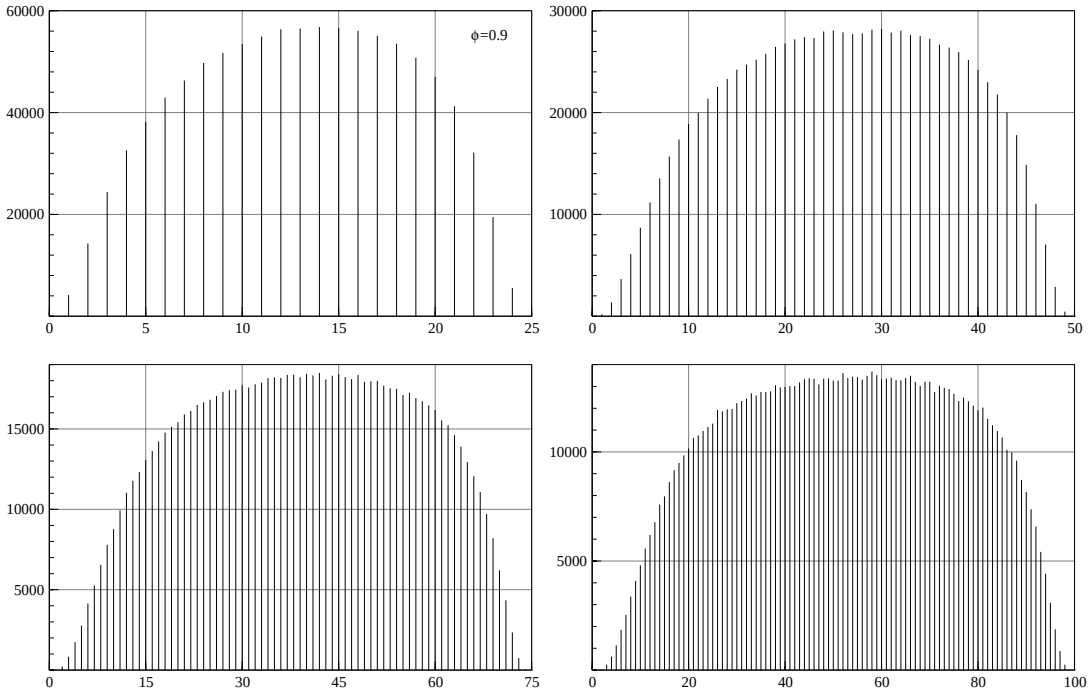


Abbildung C.85: Verteilung der Epochen der Maxima der LR9 Testgrösse mit $\phi = 0.9$

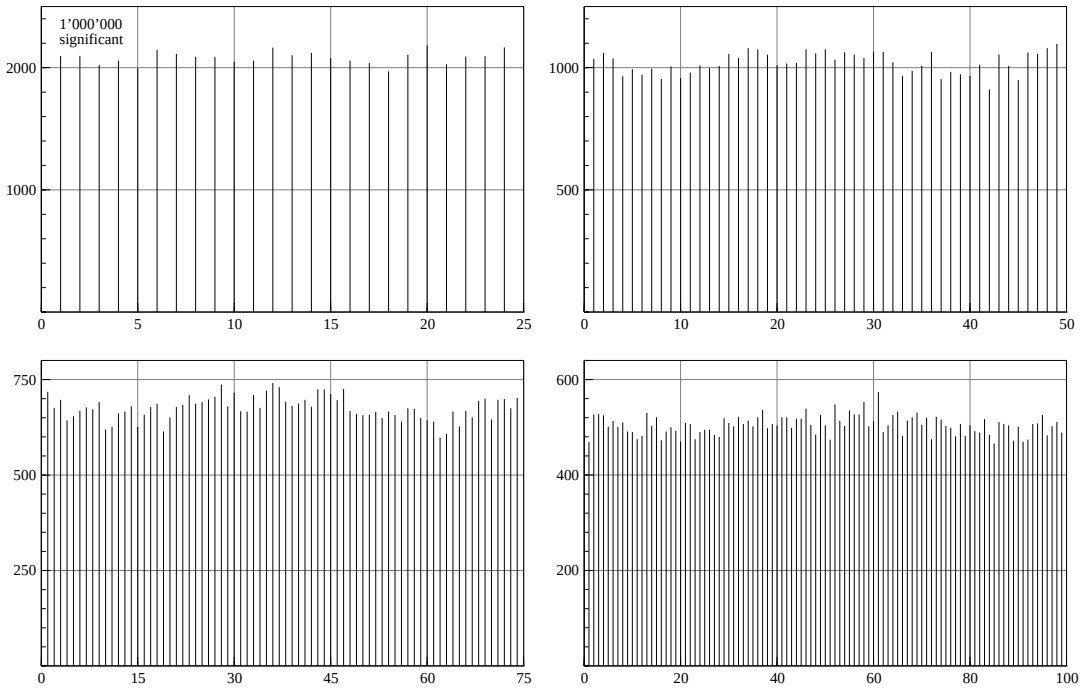


Abbildung C.86: Verteilung der Bruchepochen der LR9 Testgrösse mit $\phi = 0$

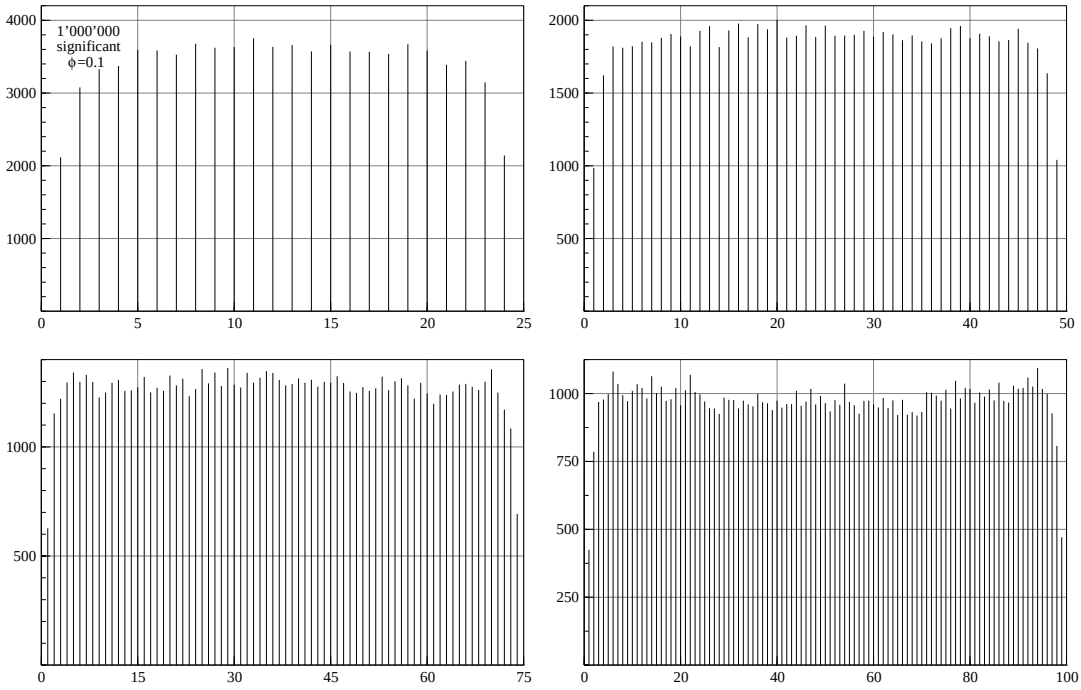


Abbildung C.87: Verteilung der Bruchepochen der LR9 Testgrösse mit $\phi = 0.1$

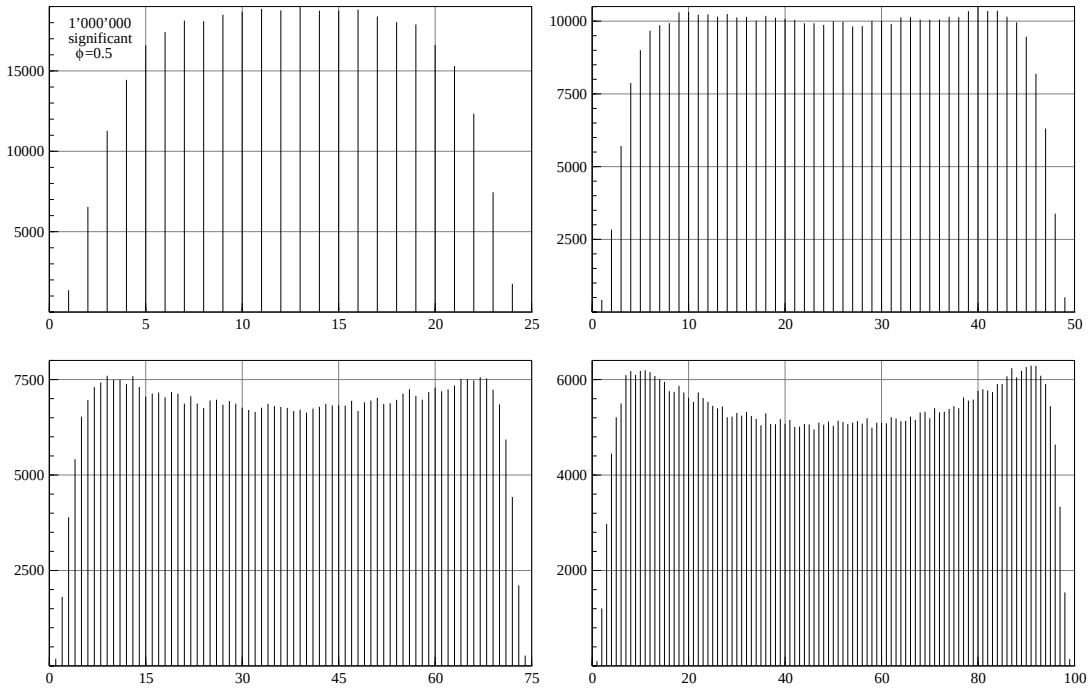


Abbildung C.88: Verteilung der Bruchepochen der LR9 Testgrösse mit $\phi = 0.5$

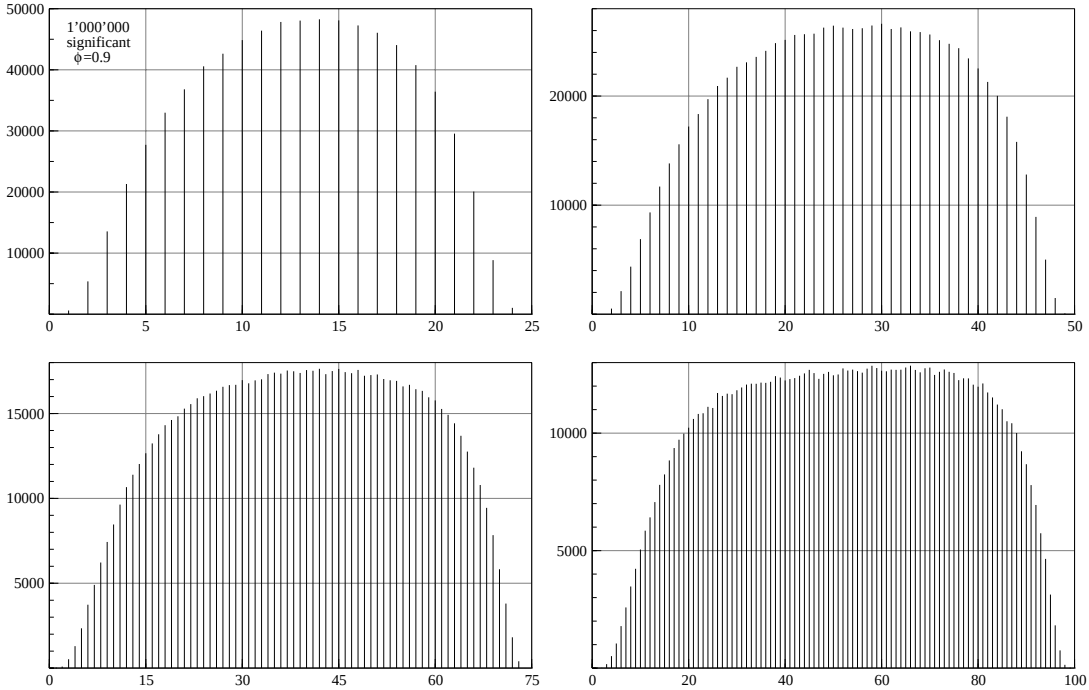


Abbildung C.89: Verteilung der Bruchepochen der LR9 Testgrösse mit $\phi = 0.9$

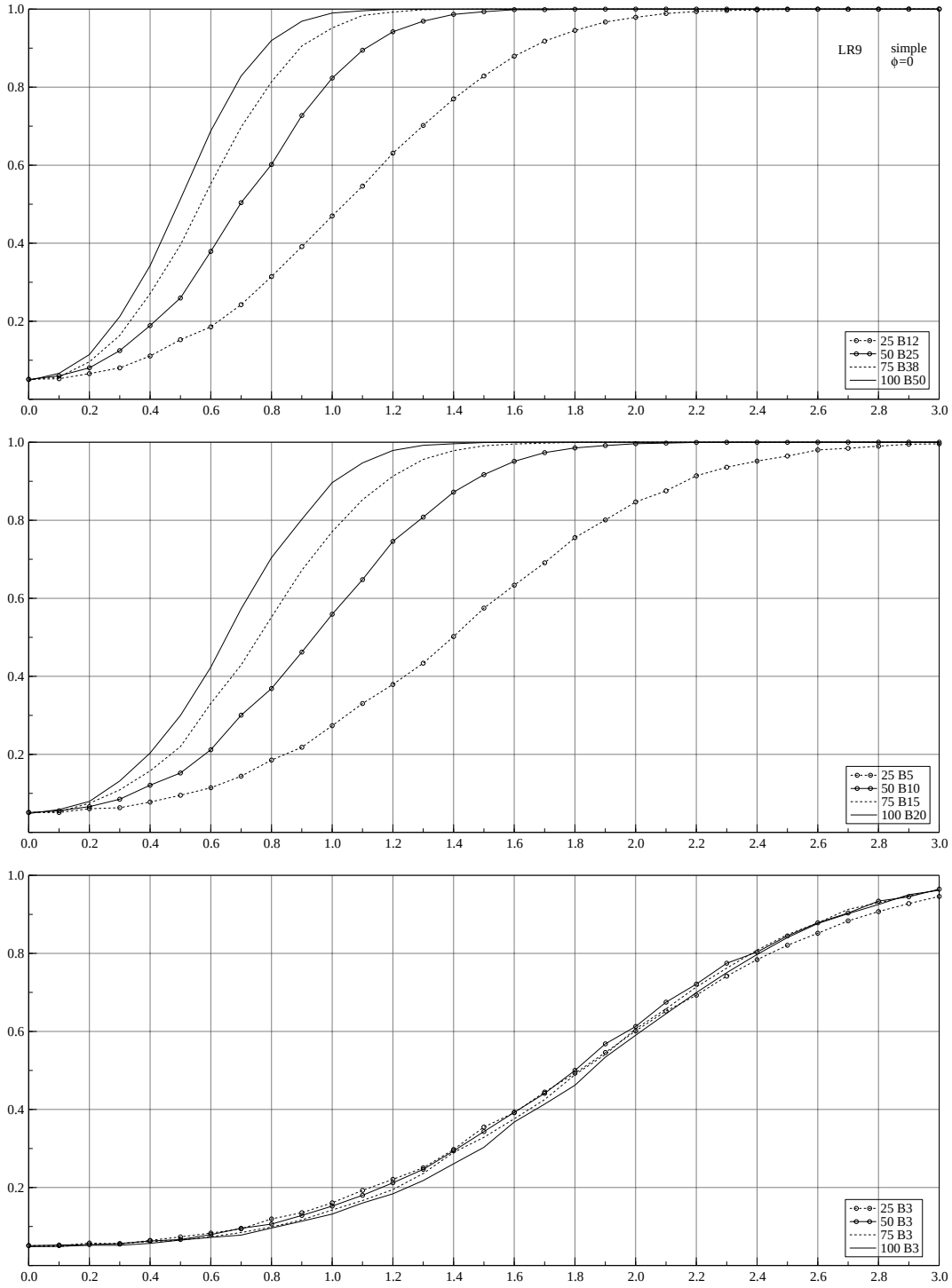


Abbildung C.90: Güte des LR9 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen. Die oberste Abbildung gibt die Gütefunktionen bei einer Versatzinhomogenität in der Mitte der Zeitreihe wieder, die mittlere Abbildung bei einer Versatzinhomogenität nach einem Fünftel der Werte und die unterste Abbildung beim 3. Wert der Zeitreihe.

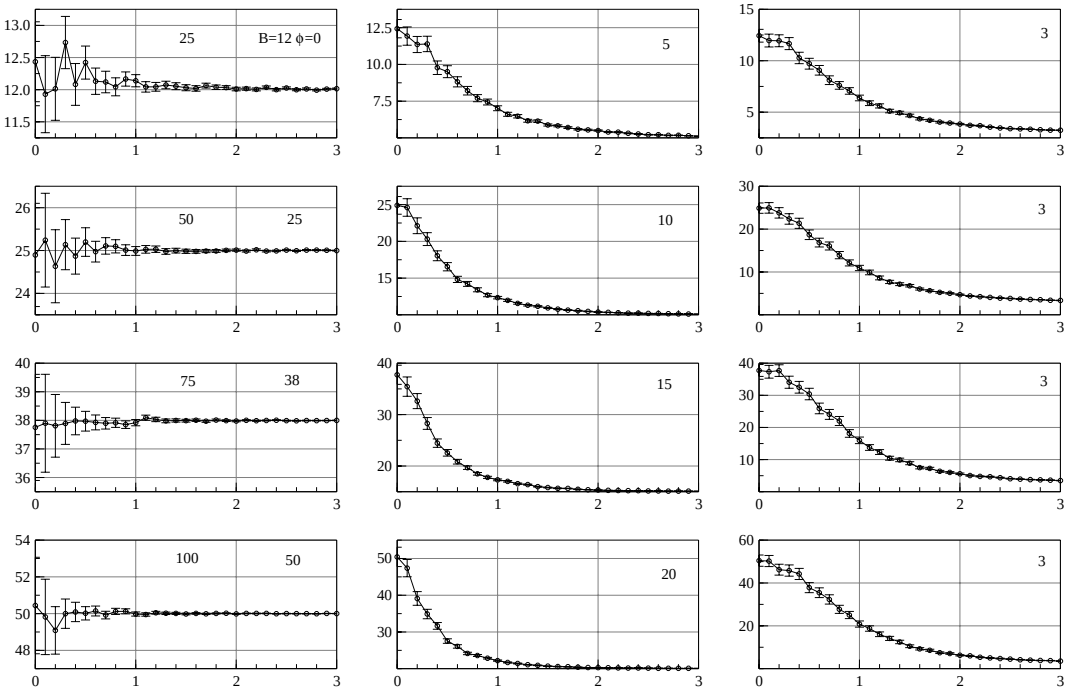


Abbildung C.91: Epochentreue des LR8 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

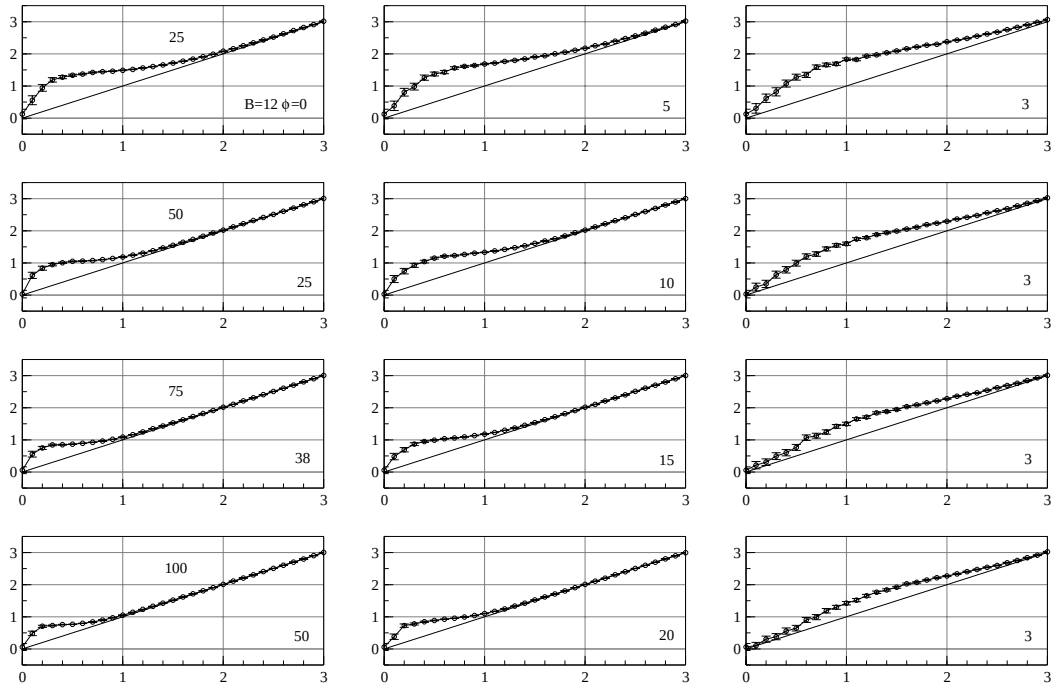


Abbildung C.92: Sprunghöhentreue des LR8 Tests unter der Alternative einer einfachen Versatzinhomogenität für Zeitreihen der Länge $N = 25$, 50, 75 und 100 und Sprunghöhen zwischen Null und 3 Standardabweichungen.

C.12 Likelihood Ratio Test 10

N	ϕ	$T_{90,l}$	T_{90}	$T_{90,u}$	$T_{95,l}$	T_{95}	$T_{95,u}$	$T_{99,l}$	T_{99}	$T_{99,u}$
25	0	2.6521	2.6543	2.6563	2.8754	2.8779	2.8805	3.2886	3.2927	3.2971
50	0	2.9435	2.9460	2.9482	3.2024	3.2053	3.2083	3.7021	3.7073	3.7126
75	0	3.0928	3.0953	3.0978	3.3677	3.3707	3.3742	3.9105	3.9168	3.9229
100	0	3.2437	3.2461	3.2488	3.5340	3.5374	3.5406	4.1148	4.1209	4.1275
25	0.1	2.8043	2.8064	2.8085	3.0362	3.0388	3.0414	3.4572	3.4620	3.4670
50	0.1	3.1601	3.1625	3.1648	3.4378	3.4411	3.4445	3.9740	3.9792	3.9846
75	0.1	3.3365	3.3389	3.3417	3.6371	3.6407	3.6440	4.2226	4.2285	4.2352
100	0.1	3.5148	3.5178	3.5208	3.8323	3.8358	3.8395	4.4597	4.4664	4.4730
25	0.5	3.5844	3.5868	3.5892	3.8310	3.8333	3.8360	4.2204	4.2237	4.2274
50	0.5	4.3272	4.3305	4.3335	4.6860	4.6901	4.6942	5.3140	5.3200	5.3268
75	0.5	4.7014	4.7052	4.7090	5.1149	5.1196	5.1246	5.8726	5.8815	5.8891
100	0.5	5.0373	5.0417	5.0457	5.4910	5.4960	5.5013	6.3345	6.3441	6.3528
25	0.9	4.5701	4.5715	4.5731	4.7111	4.7124	4.7138	4.8936	4.8952	4.8969
50	0.9	6.5656	6.5684	6.5714	6.8452	6.8480	6.8508	7.2200	7.2231	7.2264
75	0.9	7.8781	7.8822	7.8865	8.2886	8.2930	8.2972	8.8676	8.8728	8.8786
100	0.9	8.9713	8.9769	8.9822	9.5058	9.5110	9.5167	10.287	10.294	10.301

Tabelle C.26: Kritische Grenzen des LR10 Tests für 1’000’000 Replikationen.

N	1	2	3	4	5	6	7	8	9	10
0	-	0.9771	1.3838	1.6735	1.8788	2.0323	2.1541	2.2429	2.3249	2.3909
10	2.4461	2.4986	2.5390	2.5811	2.6222	2.6554	2.6806	2.7090	2.7401	2.7705
20	2.7902	2.8129	2.8379	2.8605	2.8783	2.8886	2.9123	2.9335	2.9525	2.9618
30	2.9775	2.9965	3.0108	3.0212	3.0386	3.0551	3.0609	3.0763	3.0878	3.0963
40	3.1092	3.1174	3.1363	3.1388	3.1543	3.1594	3.1803	3.1833	3.1996	3.2021
50	3.2065	3.2227	3.2334	3.2292	3.2493	3.2586	3.2687	3.2753	3.2738	3.2860
60	3.2949	3.3083	3.3128	3.3180	3.3261	3.3320	3.3331	3.3479	3.3517	3.3700
70	3.3673	3.3671	3.3732	3.3877	3.3960	3.3965	3.3997	3.4108	3.4141	3.4213
80	3.4366	3.4366	3.4484	3.4404	3.4362	3.4508	3.4500	3.4613	3.4659	3.4675
90	3.4742	3.4795	3.4852	3.4901	3.4855	3.5014	3.5066	3.5026	3.4979	3.5149
100	3.5117	3.5229	3.5218	3.5177	3.5297	3.5454	3.5395	3.5477	3.5501	3.5495
110	3.5512	3.5594	3.5510	3.5754	3.5598	3.5663	3.5822	3.5672	3.5729	3.5890
120	3.5937	3.5948	3.5919	3.5991	3.5977	3.6012	3.6040	3.6010	3.6139	3.6061
130	3.6095	3.6239	3.6278	3.6203	3.6279	3.6284	3.6307	3.6320	3.6295	3.6392
140	3.6403	3.6375	3.6366	3.6434	3.6505	3.6478	3.6517	3.6488	3.6563	3.6602
150	3.6532	3.6523	3.6531	3.6556	3.6686	3.6744	3.6692	3.6662	3.6724	3.6773
160	3.6726	3.6806	3.6781	3.6736	3.6799	3.6733	3.6775	3.6765	3.6857	3.6952
170	3.6913	3.6870	3.6841	3.6910	3.6919	3.6929	3.6939	3.6845	3.6935	3.6898
180	3.7045	3.7012	3.7029	3.6931	3.7085	3.6948	3.7003	3.7073	3.7087	3.6931
190	3.7055	3.7105	3.7074	3.7141	3.7137	3.7072	3.7202	3.6958	3.7161	3.7064

Tabelle C.27: Kritische Grenzen T_{95} des LR10 Tests für 100’000 Replikationen.

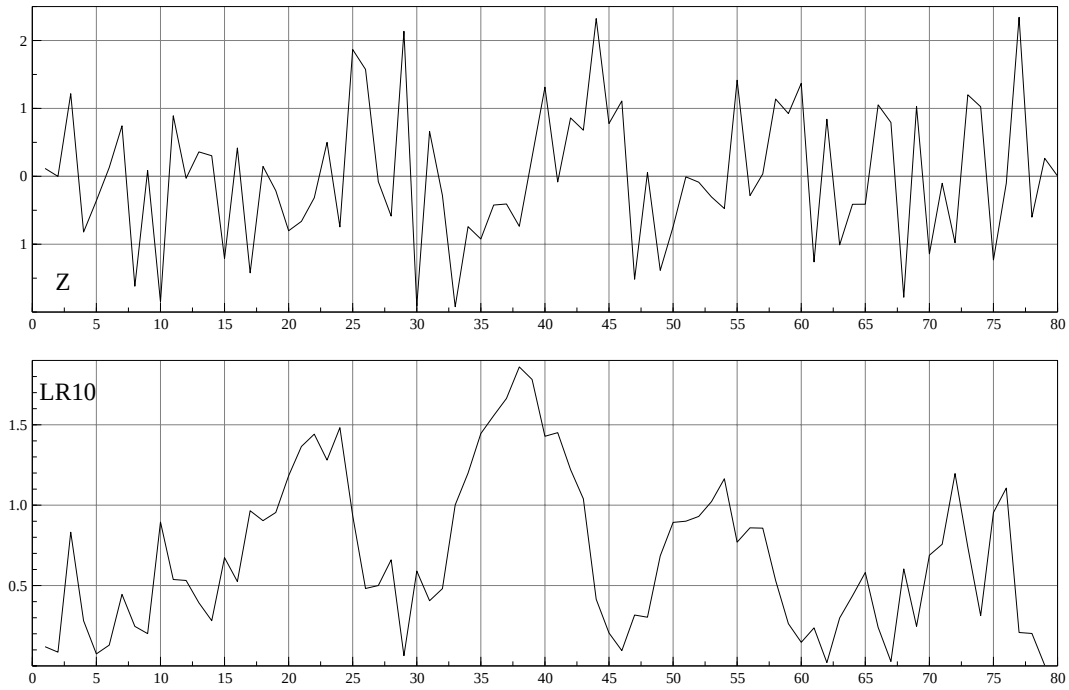


Abbildung C.93: Verlauf der LR8 Testgrösse für die Testreihe Z80

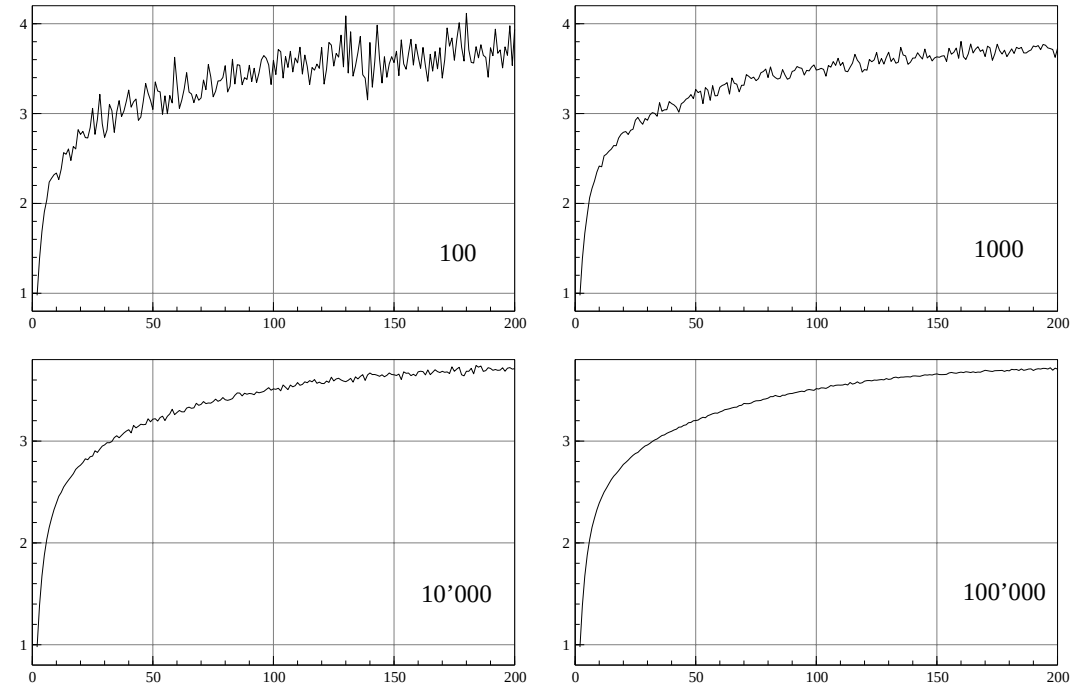


Abbildung C.94: Verlauf der kritischen Grenzen für verschiedene Anzahl Replikationen

Literaturverzeichnis

ALEXANDERSSON, H. (1984). A homogeneity test based on ratios and applied to precipitation series. *Reports Dept. of Meteorol., Uppsala, Sweden,*, **79**:1–55.

ALEXANDERSSON, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, **6**:661–675.

ANDERSON, B. D. O. AND MOORE, J. B. (1979). *Optimal Filtering*. Prentice-Hall Information and System Science Series. Prentice-Hall: Englewood Cliffs, N.J. ISBN 0-13-638122-7 hardback.

ANDREWS, D. F. AND HERZBERG, A. M. (1985). *Data - A Collection of Problems from Many Fields for the Student and Research Worker*. Springer: New York, Berlin, Heidelberg, Tokyo. ISBN 0-387-96125-9 hardback.

ANTOCH, J., HUŠKOVÁ M., AND JARUŠKOVÁ, D. (2002). Off-Line Statistical Process Control. In LAURO, C., ANTOCH, J., VINZI, V. E., AND SAPORTA, G. , Ed., *Multivariate Total Quality Control – Foundation and Recent Advances*, pp. 1–86. Physica-Verlag. Heidelberg, New York.

ANTOCH, J., HUŠKOVÁ M., AND PRÁŠKOVÁ, Z. (1997). Effect of dependence on statistics for determination of change. *Journal of Statistical Planning and Inference*, **60**:291–310.

BABU, G. J. AND FEIGELSON, E. D. (1996). *Astrostatistics*. Chapman & Hall: London, Weinheim, New York, Tokyo, Melbourne, Madras. ISBN 0-412-98391-5 hardback.

BARTLETT, M. S. (1995). Periodogram analysis and continous spectra. *Experimental Astronomy*, **6**:163–210.

BAUDENBACHER, M. (1997). *Homogenisierung langer Klimareihen, dargelegt am Beispiel der Lufttemperatur*. Veröffentlichungen der Schweizerischen Meteorologischen Anstalt: Vol. 58. ISSN 1422-1381 paperback.

BECK, R., HILBRECHT, H., REINSCH, K., AND VÖLKER, P. (1982a). *Handbuch für Sonnenbeobachter*. Fachgruppe SONNE der Vereinigung der Sternfreunde (VdS) e.V.: Berlin, Bonn. ISBN 3-923787-00-6 paperback.

BECK, R., HILBRECHT, H., REINSCH, K., AND VÖLKER, P. (1995b). *Solar Astronomy Handbook*. Willmann-Bell, Inc.: Richmond. ISBN 0-943396-47-6 hardback.

BENESTAD, R. E. (2002). *Solar Activity and Earth's Climate*. Springer-Praxis books in environmental sciences. Praxis Publishing Ltd.: Chichester, UK. ISBN 3-540-43302-3 hardback.

BERGHMANS, D., CLETTE, F., CUGNON, P., GABRYL, J.-R., HOCHEDEZ, J. F., VAN DER LINDEN, R. A. M., AND VERWICHTE, E. (2002). The solar influences data analysis centre. *Journal of Atmospheric and Solar-Terrestrial Physics*, **64**:757–761.

BOSSHARD, W. (1996). *Homogenisierung klimatologischer Zeitreihen, dargelegt am Beispiel der relativen Sonnenscheindauer*. Veröffentlichungen der Schweizerischen Meteorologischen Anstalt: Vol. 57. ISSN 1422-1381 paperback.

BOX, G. E. P., JENKINS, G. M., AND REINSEL, G. C. (1994). *Time Series Analysis: Forecasting and Control*. Prentice Hall: New Jersey. ISBN 0-13-060774-6 hardback.

BROCKWELL, P. J. AND DAVIS, R. A. (1991). *Time Series: Theory and Methods*. Springer Series in Statistics. Springer: New York, Berlin, Heidelberg. ISBN 0-387-97429-6 hardback.

BROCKWELL, P. J. AND DAVIS, R. A. (2003). *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer: New York, Berlin, Heidelberg. ISBN 0-387-95351-5 hardback.

BROWN, P. E. AND DE JONG, P. (2001). Nonparametric smoothing using state space techniques. *The Canadian Journal of Statistics*, **29, No 1**:37–50.

BROWN, R. L., D. J. E. J. M. (1975). Techniques for Testing the Constancy of Regression Relastionships over Time. *Journal of the Royal Statistical Society. Series B*, **37**:149–192.

BUISHAND, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, **58**:11–27.

CHATFIELD, C. (2004). *The Analysis of Time Series. An Introduction*. Texts in Statistical Science. Chapman & Hall: Boca Raton, London, New York, Washington. ISBN 1-58488-317-0 paperback.

CHOW, S. C. AND LIU, J. P. (1992). *Design and analysis of bioavailability and bioequivalence studies*. Statistics: textbooks and monographs; vol. 133. Dekker: New York. ISBN 0-8247-8682-3 hardback.

CONOVER, W. J. (1980). *Practical Nonparametric Statistics*. John Wiley & Sons: New York, Chichester, Brisbane, Toronto. ISBN 0-471-08457-3 paperback.

CONRAD, V. AND POLLAK, W. (1950). *Methods in Climatology*. Harvard University Press.

CORTÉS, T. R. AND SÁNCHEZ, F. (1996). *The Structure of the Sun*. Cambridge University Press: Cambridge, New York, Melbourne. ISBN 0-521-56307-0 hardback.

CSÖRGŐ, M. AND HORVÁTH, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons: Chichester, New York, Weinheim, Brisbane, Singapore, Toronto. ISBN 0-471-95522-1 hardback.

DARLING, D. A. AND ERDÖS, P. (1956). A Limit Theorem for the Maximum of Normalized Sums of Independent Random Variables. *Duke Mathematical Journal*, **23**:143–155.

DE JONG, P. AND PENZER, J. (1998). Diagnosing Shocks in Time Series. *Journal of the American Statistical Association*, **93, No 442**:796–806.

DE JONG, P. (1988). The likelihood for a state space model. *Biometrika*, **75**:165–169.

DE JONG, P. (1991). The diffuse Kalman filter. *Annals of Statistics*, **19**:1073–1083.

DOORNIK, J. A. (1998). *Ox - An object-oriented matrix programming language*. Timberlake Consultants Ltd.: Kent. ISBN 0-9533394-0-8 paperback.

DURBIN, J. AND KOOPMAN, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford Statistical Science Series 24. Oxford University Press: Oxford, New York. ISBN 0-19-852354-8 hardback.

DWIVEDI, B. N. (2003). *Dynamic Sun*. Cambridge University Press: Cambridge. ISBN 0-521-81057-4 hardback.

EDDY, J. A. (1979). *A New Sun — The Solar Results From Skylab*. NASA SP-402. NASA: Washington, D.C.

EKRUTT, J. W. (1981). *Die Sonne — Die Erforschung des kosmischen Feuers*. Geo-Bücher. Gruner + Jahr AG & Co.: Hamburg. ISBN 3-570-01720-6 hardback.

FAHRMEIR, L. AND TUTZ, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer Series in Statistics. Springer: New York, Berlin, Heidelberg, Tokio. ISBN 0-387-95187-3 hardback.

FALLOW, A., HASSIG, L., AND SOMERVILLE, R. M. S. (1990). *Die Sonne*. Reise durch das Universum. Time-Life Books B.V.: Amsterdam. ISBN 90-6182-457-5 hardback.

FEIGELSON, E. D. AND BABU, G. J. (1992). *Statistical Challanges in Astronomy*. Springer: New York, Berlin, Heidelberg. ISBN 0-387-97911-5 hardback.

FEIGELSON, E. D. AND BABU, G. J. (2003). *Statistical Challanges in Modern Astronomy*. Springer: New York, Berlin, Heidelberg. ISBN 0-387-95546-1 hardback.

FRIEDLI, T. K. AND SCHUEPBACH, E. (2001). Future Research Avenues in Homogeneity Testing of Tropospheric Ozone Records. In P. M. MIDGLEY, M. REUTHER, M. WILLIAMS , Ed., *Proceedings from the EUROTRAC-2 Symposium 2000*. Springer.

FRIEDLI, T. K., FRÖHLICH, M., MUSCHG, A., REBSAMEN, H. P., AND SCHNITTER, B. (1998). *Sempers ehemalige Eidgenössische Sternwarte in Zürich*. Schweizerische Kunstführer GSK, Nr. 631/632: Bern. ISBN 3-85782-631-2 paperback.

FRIEDLI, T. K. AND KELLER, H. U. (1993). Rudolf Wolf als Pionier der Sonnenfleckenforschung. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **138/4**:267–281.

FRIEDLI, T. K. (1984). *De maculis in Sole observatis oder die Datierung der Beobachtungen von Johannes Fabricius*. Wettbewerb Schweizer Jugend forscht 1985. Unpublished: Bern.

FRIEDLI, T. K. (1988). *Grundlagen der Solarstatistik*. Wettbewerb Schweizer Jugend forscht 1989. Unpublished: Bern. ETH BIB 754 434 HS Q.

FRIEDLI, T. K. (1989a). *Beiträge zur solarstatistischen Theorie der Sonnenfleckenrelativzahl nach Wolf – zusammengestellt zur Sonnenbeobachtertagung vom 10./11. Juni 1989 in Carona TI*. Sonnenbeobachtergruppe der SAG.: Bern.

FRIEDLI, T. K. (1989b). *Beiträge zur solarstatistischen Theorie der Sonnenfleckengruppenklassifikation nach Waldmeier & Friedli – zusammengestellt zur Sonnenbeobachtertagung vom 5. November 1989 in Zürich*. Sonnenbeobachtergruppe der SAG.: Bern.

FRIEDLI, T. K. (1997). Erste Beobachtungen am Wolfschen Normalrefraktor in Schliern b. Köniz. *Mitteilungen der Rudolf Wolf Gesellschaft*, **3 (9)**:3–46.

FRIEDLI, T. K. (1999). *Einführung in die Sonnenaktivitätsüberwachung*. Rudolf Wolf Gesellschaft.: Zürich und Bern. Ausgabe Januar 1999.

FRIEDLI, T. K. (2001). *Time Series Analysis of Sunspot Numbers*. NF Final Report. Department of mathematical Statistics, University of Berne, Switzerland: Bern.

FRIIS-CHRISTENSEN, E., FRÖHLICH, C., HAIGH, J. D., SCHÜSSLER, M., AND VON STEIGER, R. (2000). *Solar Variability and Climate*. Space Science Series of ISSI, Vol. 11. Kluwer Academic Publishers: Dordrecht, Boston, London. ISBN 0-7923-6741-3 hardback.

GINGERICH, O. (1984). *Astrophysics and twentieth-century astronomy to 1950: Part A*, Vol. 4A of *The General History of Astronomy*. Cambridge University Press: Cambridge, New York, Melbourne. ISBN 0-521-24256-8 hardback.

GIOVANELLI, R. G. (1987). *Geheimnisvolle Sonne*. VCH: Weinheim. ISBN 3-527-26501-5 hardback.

GISLER, O., BAUDENBACHER, M., AND BOSSHARD, W. (1997). *Homogenisierung schweizerischer klimatologischer Messreihen des 19. und 20. Jahrhunderts*. Schlussbericht NFP 31. vdf Hochschulverlag AG an der ETH Zürich. ISBN 3-7281-2423-0 paperback.

GLEISSBERG, W. (1952). *Die Häufigkeit der Sonnenflecken*. Akademie Verlag: Berlin.

HARGREAVES, J. K. (1992). *The solar-terrestrial environment*. Cambridge Atmospheric and Space Science. Cambridge University Press: Cambridge, New York, Melbourne. ISBN 0-521-32748-2 hardback.

HARVEY, A. AND KOOPMAN, S. J. (1992). Diagnostic Checking of Unobserved-Components Time Series. *Journal of Business & Economic Statistics*, **10, No 4**:377–389.

HARVEY, A. AND KOOPMAN, S. J. (1996). Structural time series models in medicine. *Statistical Methods in Medical Research*, **1996; 5**:23–49.

HARVEY, A. AND KOOPMAN, S. J. (2000). Signal extraction and the formulation of unobserved components models. *Econometrics Journal*, **3**:84–107.

HARVEY, A. C. AND SHEPHARD, N. (1993). Structural Time Series Models. In *Handbook of Statistics, Vol. 11*, pp. 261–302. Elsevier.

HARVEY, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge Univesity Press: Cambridge, New York, Melbourne. ISBN 0-521-40573-4 paperback.

HARVEY, A. C. (1993). *Time Series Models*. Harvester Wheatsheaf: New York, London, Toronto, Sidney, Tokyo, Singapore. ISBN 0-7450-1200-0 paperback.

HARVEY, K. L. (1992). *The Solar Cycle*. Astronomical Society of the Pacific Conferenc Series, Volume 27. Astronomical Society of the Pacific: San Fancisco. ISBN 0-937707-46-5 hardback.

HATHAWAY, D. H., WILSON, R. M., AND REICHMANN, E. J. (1999). A synthesis of solar cycle prediction techniques. *Journal of Geophysical Research*, **104**:22,375–22,388.

HAWKINS, D. M. (1977). Testing a Sequence of Observations for a Shift in Location. *Journal of the American Statistical Association*, **72, (357)**:180–186.

HAWKINS, D. M. AND OLWELL, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*. Statistics for Engeneering and Physical Science. Springer: New York, Berlin, Heidelberg. ISBN 0-387-98365-1 hardback.

HERZOG, J. AND MÜLLER-WESTERMEIER, G. (1998). *Homogenitätsprüfung und Homogenisierung klimatologischer Messreihen im Deutschen Wetterdienst*. Berichte des Deutschen Wetterdienstes. Vol. 202. Selbstverlag des Deutschen Wetterdienstes. ISBN 3-88148-336-5 paperback.

HOUGHTON, J. T., DING, Y., GRIGGS, D. J., NOGUER, M., VAN DER LINDEN, P. J., DAI, X., MASKELL, K., AND JOHNSON, C. A. (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report pf the Intergovernamental Panel on Climate Change*. Cambridge University Press: Cambridge, New York, Oakleigh, Melbourne, Madrid, Cape Town. ISBN 0-521-01495-6 paperback.

HOYT, D. V. AND SCHATTEN, K. H. (1992a). A new look at Wolf sunspot numbers in the late 1700's. *Solar Physics*, **138**:387–397.

HOYT, D. V. AND SCHATTEN, K. H. (1992b). New information on solar activity, 1779-1818, from Sir William Herschel's unpublished notebooks. *The Astrophysical Journal*, **384**:361–384.

HOYT, D. V. AND SCHATTEN, K. H. (1992c). Sir William Herschel's unpublished notebooks: abstracts of Solar observations. *The Astrophysical Journal Supplement Series*, **78**:301–340.

HOYT, D. V. AND SCHATTEN, K. H. (1994). The one hundredth year of Rudolf Wolf's death: Do we have the correct reconstruction of solar activity? *Geophysical Research Letters*, **21 (18)**:2067–2070.

HOYT, D. V. AND SCHATTEN, K. H. (1995a). A new interpretation of Christian Horrebow's sunspot observations from 1761 to 1777. *Solar Physics*, **160**:387–392.

HOYT, D. V. AND SCHATTEN, K. H. (1995b). A revised listing of the number of sunspot groups made by Pastorff, 1819 to 1833. *Solar Physics*, **160**:393–399.

HOYT, D. V. AND SCHATTEN, K. H. (1995c). Observations of sunspots by Flamsteed during the Maunder Minimum. *Solar Physics*, **160**:379–385.

HOYT, D. V. AND SCHATTEN, K. H. (1995d). Overlooked sunspot observations by Hevelius in the early Maunder Minimum, 1653-1684. *Solar Physics*, **160**:371–378.

HOYT, D. V. AND SCHATTEN, K. H. (1996). How well was the Sun observed during the Maunder Minimum? *Solar Physics*, **165**:181–192.

HOYT, D. V. AND SCHATTEN, K. H. (1997). *The Role of the Sun in Climate Change*. Oxford University Press: New York, Oxford. ISBN 0-19-509414-X paperback.

HOYT, D. V. AND SCHATTEN, K. H. (1998). Group Sunspot Numbers: a new Solar Activity Reconstruction. *Solar Physics*, **181**:491–512.

HRUSKA, J., SHEA, M. A., SMART, D. F., AND HECKMAN, G. (1994). *Solar-Terrestrial Predictions - IV. 3 Volumes*. NOAA Environmental Research Laboratories: Boulder, Colorado, USA.

HÜSLER, J. AND ZIMMERMANN, H. (2001). *Statistische Prinzipien für medizinische Projekte*. Hans Huber: Bern, Göttingen, Toronto, Seattle. ISBN 3-456-83437-3 hardback.

IZENMAN, A. J. (1985). J. R. Wolf and the Zürich Sunspot Relative Numbers. *The Mathematical Intelligencer*, **7**:27–33.

JASCHEK, C. AND MURTAGH, F. (1990). *Errors, Bias and Uncertainties in Astronomy*. Cambridge University Press: Cambridge, New York, Melbourne, Sidney. ISBN 0-521-39300-0 hardback.

JAZWINSKI, A. H. (1970). *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engeneering 64. Academic Press: New York, San Francisco, London. ISBN 0-12-381550-9 hardback.

JEVONS, W. S. (1878). The Periodicity of Commercial Crises and Its Physical Explanation. In H. S. FOXWELL , Ed., *Investigations in Currency and Finance.*, pp. 106–219. New York: Macmillan.

KALMAN, R. E. AND BUCY, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *J. Basic Eng., Trans. ASME, Series D*, **83 (3)**:95–108.

KALMAN, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng., Trans. ASME, Series D*, **82 (1)**:35–45.

KARKOSCHKA, E. (1979). Neue Relativzahl-Mittelung. *SONNE*, **3**:33–37.

KELLER, H. U. AND FRIEDLI, T. K. (1995). The Sunspot Activity in the Years 1976 - 1995. *Mitteilungen der Rudolf Wolf Gesellschaft*, **3 (7)**:3–46.

KIEPENHEUER, K. O. (1957). *Die Sonne*. Springer: Berlin, Göttingen, Heidelberg.

KITAGAWA, G. AND GERSCH, W. (1996). *Smoothness Priors Analysis of Time Series*. Lecture Notes in Statistics 116. Springer: New York, Berlin, Heidelberg, Tokyo. ISBN 0-387-94819-8 paperback.

KLEIN, J. L. (1997). *Statistical Visions in Time. A History of Time Series Analysis 1662-1938*. Cambridge University Press: Cambridge, New York, Melbourne, Sidney. ISBN 0-521-42046-6 hardback.

KNUTH, D. E. (1981). *The Art of Computer Programming. Vol. 2.: Seminumerical algorithms*. 2nd ed. Addison-Wesley: Reading, Massachusetts. ISBN 0-201-03822-6 hardback.

KOOPMAN, S. J., HARVEY, A. C., DOORNIK, J. A., AND SHEPHARD, N. (2000). *Stamp: Structural Time series Analyser, Modeller and Predictor*. Timberlake Consultants Press: London. ISBN 0-9533394-4-0 paperback.

KOOPMAN, S. J., SHEPHARD, N., AND DOORNIK, J. A. (1999a). Statistical algorithms for models in state space form using SsfPack 2.2. *Econometrics Journal*, **2**:113–166.

KOOPMAN, S. J., SHEPHARD, N., AND DOORNIK, J. A. (2002b). *Statistical algorithms for models in state space*. mimeo: Free University, Amsterdam.

KOOPMAN, S. J. (1993). Disturbance smoother for state space models. *Biometrica*, **80**:117–126.

KRÄHENBÜHL, S., KÜNZLE, T., FRIEDLI, T. K., STEINER, N., AND HÜSLER, J. R. (2004). *Vervollständigung und Homogenisierung von Schadstoff-Messreihen des NABEL*. Pilot-Projekt. Meteotest.

LADIRAY, D. AND QUENNEVILLE, B. (2001). *Seasonal Adjustment with the X-11 Method*. Lecture Notes in Statistics 158. Springer: New York, Berlin, Heidelberg, Tokyo. ISBN 0-387-95171-7 paperback.

LANG, K. R. (1996). *Die Sonne, Stern unserer Erde*. Springer: Berlin, Heidelberg. ISBN 3-540-59437-X hardback.

LANG, K. R. (2001). *The Cambridge Encyclopedia of the Sun*. Cambridge University Press: Cambridge, New York, Oakleigh, Madrid, Cape Town. ISBN 0-521-78093-4 hardback.

LAURO, C., ANTOCH, J., VINZI, V. E., AND SAPORTA, G. (2002). *Multivariate Total Quality Control – Foundation and Recent Advances*. Physica-Verlag: Heidelberg, New York.

LEADBETTER, M. R., LINDGREN, G., AND ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer: New York, Heidelberg, Berlin. ISBN 3-540-90731-9 hardback.

LETFUS, V. (1999). Daily Relative Sunspot Numbers 1749-1848: Reconstruction of missing observations. *Solar Physics*, **184**:201–211.

L' ECUYER, P. (1997). Tables of maximally-equidistributed combined LSFR generators. *Mimeo. University of Montreal, Canada*.

MARONNA, R. AND YOHAI, V. J. (1978). A Bivariate Test for the Detection of a Systematic Change in Mean. *Journal of the American Statistical Association*, **73, (363)**:640–645.

MARSAGLIA, G. (1997). A random number generator for C. *Discussion paper on sci.stat.math*.

MCCULLAGH, P. AND NELDER, J. A. (1989). *Generalized Linear Models*. Monographs on Statistics and Applied Probability 37. Chapman and Hall: London, New York, Tokyo, Melboure, Madras. ISBN 0-412-31760-5 hardback.

MCINTOSH, P. S. AND DRYER, M. (1972). *Solar Activity Observations and Prediction*. Progress in Astronautics and Aeronautics, v. 30. The MIT Press: Cambridge, Massachusetts, and London, England. ISBN 0-262-13086-6 hardback.

MCKINNON, J. A. (1987). *Sunspot Numbers: 1610-1985, based on "The Sunspot-Activity in the Years 1610-1960, revised by John A. McKinnon"*. Report UAG-95. World Data Center A for Solar-Terrestrial Physics: Boulder, Colorado. ISSN 0579-7144 paperback.

MEEUS, J. (1983). *Astronomical Tables of the Sun, Moon, and Planets*. Willmann-Bell, Inc.: Richmond. ISBN 0-943396-02-6 paperback.

MEEUS, J. (1997). *Mathematical Astronomy Morsels*. Willmann-Bell, Inc.: Richmond. ISBN 0-943396-51-4 hardback.

MEEUS, J. (2002). *More Mathematical Astronomy Morsels*. Willmann-Bell, Inc.: Richmond. ISBN 0-943396-74-3 hardback.

MEEUS, J. (1958). Une formule d'adoucissement pour l'activité solaire. *Ciel et Terre*, **74(6)**:445–449.

MITTER, M. (2004). *Lineare nicht-Gauss'sche Zustandsraummodelle*. Master's Thesis: University of Bern.

MONKS, P. S., RICKARD, A. R., DENTENER, F., JONSON, J. E., LINDSKOG, A., ROEMER, M., SCHUEPBACH, E., FRIEDLI, T. K., AND SOLBERG, S. (2003). *Tropospheric Ozone and Precursors, Trends, Budgets and Policy*. TROTREP Synthesis and Integration Report. University of Leicester.

NICOLSON, I. (1982). *Die Sonne*. Herder: Freiburg im Breisgau. ISBN 3-451-19461-9 hardback.

PAP, J. M., FRÖHLICH, C., HUDSON, H. S., AND SOLANKI, S. K. (1994). *The Sun as a Variable Star – Solar and Stellar Irradiance Variations*. IAU Colloquium 143. Cambridge University Press: Cambridge, New York, Melbourne. ISBN 0-521-420068-7 hardback.

PARK, S. AND MILLER, K. (1988). Random number generators: Good ones are hard to find. *Communications of the ACM*, **31**:1192–1201.

PEPIN, R. O., EDDY, J. A., AND MERRILL, R. B. (1980). *The Ancient Sun. Fossil Record in the Earth, Moon and Meteorites*. Geochimica et Cosmochimica Acta, Supplement 13. Pergamon Press: New York, Oxfrod, Toronto, Sydney, Frankfurt, Paris. ISBN 0-08-1026324-0 hardback.

PETERSON, T. C., EASTERLING, D. R., KARL, T. R., GROISMAN, P., NICHOLLS, N., PLUMMER, N., TOROK, S., AUER, I., BOEHM, R., GULLETT, D., VINCENT, L., HEINO, R., TUOMENVIRTA, H., MESTRE, O., SZENTIMREY, T., SALINGER, J., FORLAND, E. J., HANSSEN-BAUER, I., ALEXANDERSSON, H., JONES, P., AND PARKER, D. (1998). Homogeneity Adjustments of *in situ* Atmospheric Climate Data: A Review. *International Journal of Climatology*, **18**:1493–1517.

POTTER, K. W. (1981). Illustration of a new test for detecting a shift in mean in precipitation series. *Monthly Weather Review*, **109**:2040–2045.

POYNTING, J. H. (1877). The Drunkenness Statistics of the Large Towns in England and Wales. In *Collected Scientific Papers*, pp. 497–503. Cambridge University Press.

POYNTING, J. H. (1884). A Comparison of the Fluctuations in the Price of Wheat and in Cotton and Silk Imports into Great Britain. *Journal of the Royal Statistical Society*, **47**:43–74.

PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., AND VETTERING, W. T. (1990). *Numerical Recepes in Pascal. The Art of Scientific Computing*. Cambridge University Press: Cambridge, New York, Port Chester, Melboure, Sidney. ISBN 0-521-37516-9 hardback.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. Probability and Mathematical Statistics. Academic Press: London, San Diego, New York, Berkeley, Boston, Sidney, Tokyo. ISBN 0-12-564922-3 paperback.

QIN, Z. (1998). The transitional time scale from stochastic to chaotic behavior for solar activity. *Solar Physics*, **178**:423–431.

RAO, T. S., PRIESTLEY, M. B., AND LESSI, O. (1997). *Applications of Time Series Analysis in Astronomy and Meteorology*. Chapman & Hall: London, Weinheim, New York, Tokyo, Melbourne, Madras. ISBN 0-412-63800-2 hardback.

REINSCH, K., BECK, R., HILBRECHT, H., AND VÖLKER, P. (1999). *Die Sonne beobachten*. Sterne und Weltraum, Hüthig GmbH: Heidelberg. ISBN 3-87973-930-7 paperback.

RIND, D. (2002). The Sun's Role in Climate Variations. *Science*, **296**:673–677.

ROSENBERG, B. (1973). Random coefficients models: the analysis of a cross-section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement*, **2**:399–428.

ROTH, G. D. (1989). *Handbuch für Sternfreunde*. Springer: Berlin, Heidelberg, New York. ISBN 3-540-50763-9 Bände 1 und 2 hardback.

RUF, K. (1977). American and Zürich Sunspot Relative Numbers. *Journal of Interdisciplinary Cycle Research*, **8**:215–217.

RUTTEN, R. J. AND SCHRIJVER, C. J. (1994). *Solar Surface Magnetism*. NATO ASI Series C: Mathematical and Physical Sciences, vol. 433. Kluwer Academic Publishers: Dordrecht, Boston, London. ISBN 0-7923-2845-0 hardback.

SALAKHUTDINOVA, I. I. (1998). A fractal structure of the time series of global indices of solar activity. *Solar Physics*, **181**:221–235.

SCHEFFLER, H. AND ELSÄSSER, H. (1990). *Physik der Sterne und der Sonne*. B.I. Wissenschaftsverlag: Mannheim, Wien, Zurich. ISBN 3-411-14172-7 hardback.

SCHLITTGEN, R. AND STREITBERG, B. H. J. (1999). *Zeitreihenanalyse*. Lehr- und Handbücher der Statistik. Oldenbourg: München, Wien. ISBN 3-486-24982-7 hardback.

SCHOVE, D. J. (1983). *Sunspot Cycles*. Benchmark Papers in Geology, Volume 68. Hutchinson Ross Publishing Company: Stroudsburg, Pennsylvania. ISBN 0-87033-424-X hardback.

SCHRIJVER, C. J. AND ZWAAN, C. (2000). *Solar and Stellar Magnetic Activity*. Cambridge Astrophysics Series; 34. Cambridge University Press: Cambridge, UK. ISBN 0-521-58286-5 hardback.

SCHUEPBACH, E. AND FRIEDLI, T. K. (1998). Quality control procedures applied to the 10-year (1988-97) ozone record at Jungfraujoch in the Swiss Alps. In *Papers of the 11th World Clean Air and Environment Congress, Durban, South Africa, 13-18 September 1998*, pp. 3D–3(I–5). IUAPPA.

SCHUEPBACH, E., FRIEDLI, T. K., ZANIS, P., MONKS, P. S., AND PENKETT, S. A. (2001). State space analysis of changing seasonal ozone cycles (1988-1997) at Jungfraujoch (3580 m above sea level) in Switzerland. *Journal of Geophysical Research*, **106**:20,413–20,427.

SCHUSTER, A. (1898). On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *Terr. Mag. Atmos. Elect.*, **3**:13–41.

SCHÜTTEL, M. A. (2002). *Modellierung saisonaler Ozonzyklen mit linearen Gauss'schen Zustandsraummodellen*. Master's Thesis: University of Bern.

SCHWABE, H. S. (1843). Die Sonne. *Astronomische Nachrichten*, **20 (473)**:285–286.

SCHWABE, H. S. (1844). Sonnen-Beobachtungen im Jahre 1843. *Astronomische Nachrichten*, **21 (495)**:233–236.

SCHWEPPE, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*, **11**:61–70.

SHAPLEY, A. H. (1946). American Observations of Relative Sunspot-Numbers in 1945 for Application to Ionospheric Predictions. *Popular Astronomy*, **54**:351–358.

SHAPLEY, A. H. (1949). Reduction of Sunspot-Number Observations. *Publications of the Astronomical Society of the Pacific*, **61**:13–21.

SHUMWAY, R. H. AND VEROSUB, K. L. (2002). Dynamic Mixed Models for Merging Multiple Time Series Fragments. *Draft Prepared for the State Space and Unobserved Components Models Conference, Amsterdam*.

SHUMWAY, R. H. AND STOFFER, D. S. (2000). *Time Series Analysis and Its Applications*. Springer Texts in Statistics. Springer: New York, Berlin, Heidelberg. ISBN 0-387-98950-1 hardback.

SINGH, M. (1994). *Die Sonne — Das Gestirn in der Kulturgeschichte*. Ernst Wasmuth Verlag/UNESCO: Tübingen, Berlin. ISBN 38030 3059 5 hardback.

SONETT, C. P., GIAMPAPA, M. S., AND MATTHEWS, M. S. (1991). *The Sun in Time*. Space Science Series. The University of Arizona: Tucson. ISBN 0-8165-1297-3 hardback.

STEINER, N. (2004). *Exakte Initialisierung von linearen, Gauss'schen Zustandsraummodellen in S-Plus*. Master's Thesis: University of Bern.

STEWART, B. (1883). Meteorology: Terrestrial Magnetism. In *Encyclopedia Britannica, 9th ed.*, pp. 159–184. New York: Charles Scribner's Sons.

TAYLOR, P. O. (1985). Computation of American Relative Sunspot Numbers. *Journal of the American Association of Variable Star Observers*, **14**:28–32.

TAYLOR, P. O. (1991). *Observing the Sun*. Practical Astronomy Handbook Series, Volume 3. Cambridge University Press: Cambridge, Melbourne, Sydney. ISBN 0-521-40110-0 hardback.

THOMAS, J. H. AND WEISS, N. O. (1992). *Sunspots: Theory and Observations*. NATO ASI Series C: Mathematical and Physical Sciences, vol. 375. Kluwer Academic Publishers: Dordrecht, Boston, London. ISBN 0-7923-1852-8 hardback.

THÖNI, C., WEIGEL, S., FRIEDLI, T. K., AND HÜSLER, J. R. (2001). *Homogenisierung und Interpolation fehlender Werte von Schadstoff-Messreihen des NABEL*. Vorstudie. Meteotest.

TONG, H. (1990). *Non-linear Time Series*. Clarendon Press: Oxford. ISBN 0-19-852300-9 paperback.

TRUMPLER, R. J. AND WEAVER, H. F. (1953). *Statistical Astronomy*. University of California Press: Berkeley, Los Angeles.

WALDMEIER, M. (1935). Neue Eigenschaften der Sonnenfleckenkurve. *Astronomische Mitteilungen, Nr. 133*:1–25.

WALDMEIER, M. (1948). 100 Jahre Sonnenfleckenstatistik. *Astronomische Mitteilungen, Nr. 152*:1–8.

WALDMEIER, M. (1961). *The Sunspot-Activity in the Years 1610-1960*. Schulthess & CO AG: Zürich.

WALDMEIER, M. (1963). Die Eidgenössische Sternwarte 1863-1963. *Astronomische Mitteilungen, Nr. 250*:1–23.

WALDMEIER, M. (1968). Sonnenfleckenkurven und die Methode der Sonnenaktivitätsprognose. *Astronomische Mitteilungen, Nr. 286*:1–13.

WALDMEIER, M. (1971). *Astronomische Mitteilungen, Nr. 304*:1–10.

WALDMEIER, M. (1978). *Astronomische Mitteilungen, Nr. 358*:1–27.

WEI, W. W. S. (1990). *Time Series Analysis*. Addison-Wesley Publishing Company, Inc.: Redwood City, California, New York, Tokyo. ISBN 0-201-15911-2 hardback.

WEST, M. AND HARRISON, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer Series in Statistics. Springer: New York, Berlin, Heidelberg, Tokio. ISBN 0-387-94725-6 hardback.

WHITTLE, P. (1954). A statistical investigation of sunspot observations with special reference to H. Alven's sunspot model. *Astrophysical Journal*, **120**:251–260.

WILLIAMSON, S. P., EVANS, D. L., NEYLAND, M. A., WALDRON, R. E., DEVINE, J. F., WHATLEY, D., WEILER, E. J., LEINEN, M. S., AND BABCOCK, M. R. (2000). *National Space Weather Program. Implementation Plan. 2nd. Edition.* FCM-P31-2000: Washington, DC.

WILSON, P. R. (1994). *Solar and Stellar Activity Cycles*. Cambridge Astrophysics Series, Volume 24. Cambridge University Press: Cambridge, Melbourne, Sydney. ISBN 0-521-43081-X hardback.

WILSON, R. M. (1998). A comparison of Wolf's reconstructed record of annual sunspot number with Schwabe's observed record of 'clusters of spots' for the interval of 1826-1868. *Solar Physics*, **182**:217–230.

WOLD, A. (1940). The fitting of straight lines if both variables are subject to error. *Annals of Mathematical Statistics, Nr. 50*, **11**:284–300.

WOLFER, A. (1895). *Astronomische Mittheilungen, Nr. 96*:187–232.

WOLF, R. (1851). Sonnenflecken-Beobachtungen in der zweiten Hälfte des Jahres 1850. *Mittheilungen der Naturforschenden Gesellschaft in Bern, Nr. 207*:89–95.

WOLF, R. (1852). Neue Untersuchungen über die Periode der Sonnenflecken und ihre Bedeutung. *Mittheilungen der Naturforschenden Gesellschaft in Bern, Nr. 255-257*:249–270.

WOLF, R. (1856). *Mittheilungen über die Sonnenflecken, Nr. 1*:12.

WOLF, R. (1857). *Mittheilungen über die Sonnenflecken, Nr. 3*:27–50.

WOLF, R. (1873). *Astronomische Mittheilungen, Nr. 33*:82–85.

WOLF, R. (1875). *Astronomische Mittheilungen, Nr. 38*:393–423.

WOLF, R. (1877). *Astronomische Mittheilungen, Nr. 42*:29–60.

WOLF, R. (1878). *Beobachtungen der Sonnenflecken. Von der Zeit ihrer Entdeckung bis auf die Gegenwart. Gesammelt von Rudolf Wolf*. Wissenschaftshistorische Sammlungen der ETH Bibliothek. Handschrift Hs 386:46: Zürich.

WOLF, R. (1880). *Astronomische Mittheilungen, Nr. 50*:269–316.

WOLF, R. (1890). *Handbuch der Astronomie, ihrer Geschichte und Litteratur*. F. Schulthess: Zürich.

WORLD METEOROLOGICAL ORGANIZATION (1986). *Guidelines on the Quality Control of Surface Climatological Data*. Prepared by P.F. Abbott (U.K.) as Rapporteur in the WMO Commission for Climatology: WCP-85. WMO/TD-No. 111.

WORSLEY, K. J. (1979). On the Likelihood Ratio Test for a Shift in Location of Normal Populations. *Journal of the American Statistical Association*, **74, (366)**:365–367.

YAO, Y.-C. AND DAVIS, R. A. (1986). The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. *Sankhyā: The Indian Journal of Statistics*, **48, Series A, Pt. 3**:339–353.

YOUNG, P. (1984). *Recursive Estimation and Time-Series Analysis. An Introduction*. Communications and Control Engeneering Series. Springer: Berlin, Heidelberg, New York, Tokio. ISBN 0-387-13677-0 hardback.

ZANIS, P., SCHUEPBACH, E., SCHEEL, H. E., BAUDENBACHER, M., AND BUCHMANN, B. (1999). Inhomogeneities and trends in the surface ozone record (1988-1996) at Jungfraujoch in the Swiss Alps. *Atmospheric Environment*, **33**:3777–3786.

ZIVOT, E. AND WANG, J. (2003). *Modeling Financial Time Series with S-Plus*. Springer: New York, Berlin, Heidelberg. ISBN 0-387-95549-6 paperback.

Curriculum vitae

Ich wurde am 3. September 1967 als einziges Kind von Katja und Peter Friedli-Schneider in Bern geboren und besuchte dort 1974–1978 die Primarschule Bitzius und 1978–1982 die Sekundarschule Laubegg. 1982 trat ich in das Gymnasium Bern-Kirchenfeld ein, welches ich im Herbst 1986 mit der Matura Typus C abschloss. Nach zwei Zwischenjahren immatrikulierte ich mich 1988 an der Universität Bern in den Fächern Astronomie (Hauptfach), Mathematik und Physik. Im Dezember 1996 erwarb ich bei Prof. Dr. Gerhard Beutler das Diplom in Astronomie.

Von 1992–1993 und von 1995–1997 arbeitete ich als Hilfsassistent am Astronomischen Institut und zwischen 1992 und 1995 zusätzlich für das Bundesamt für Landestopografie (Wabern) am SLR-Observatorium in Zimmerwald. Seit 1997 bin ich als Assistent am Institut für mathematische Statistik und Versicherungslehre tätig.

In den Jahren 1998–2001 erteilte ich am Weiterbildungszentrum für Gesundheitsberufe (WEG) in Aarau zwei Semesterwochenstunden Statistik für Pflegewissenschaftler und seit 2001 unterrichte ich drei Semesterwochenstunden Statistik für Biotechnologen an der Hochschule Wädenswil.

Seit August 1996 bin ich mit der Aktuarin Maria Hostettler Friedli verheiratet und inzwischen Vater von Anna Lea (*1999) und Franziska Manuela (*2001) geworden.