

UNIVERSITÄT DUISBURG-ESSEN
FAKULTÄT PHYSIK
WINTERSEMESTER 18/19

Versuch : 1.12 Statistische Zeitreihenanalyse

FORTGESCHRITTENES PRAKTIKUM PHYSIK
BETREUER: ANDREAS MÜHLBACHER
GRUPPE N

JAN VERLAGE (3029893)
EDGAR JUNGBLUT (3031651)
HENDRIK MANNEL (3023047)

19.10.18
jan.verlage@stud.uni-duisburg-essen.de

Inhaltsverzeichnis

1 Grundlagen	3
1.1 Wahrscheinlichkeiten und Verteilungen	3
1.2 Stochastische Momente	4
1.3 Gaußverteilung und zentraler Grenzwertsatz	4
1.4 Statistische Abhangigkeit	5
1.4.1 Pearson-Korrelationskoeffizient	5
1.4.2 Autokorrelationsfunktion	6
1.4.3 Copula	6
1.5 Die untersuchten Systeme	7
1.5.1 Brownsche Bewegung	8
1.5.2 Aktienkurse	8
1.5.3 Temperaturen	8
2 Brownsche Bewegung	9
3 Finanzdaten	15
4 Temperaturdaten	21

1 Grundlagen

In diesem Versuch wird eine statistische Analyse von Zeitreihen durchgeführt. Dabei werden eine Teilchenbewegung (Brownsche Bewegung), Aktienkursen und Temperaturdaten betrachtet. Die Beschreibung dieser Systeme mittels stochastischer Prozesse und die Methoden zur Analyse werden in diesem Kapitel vorgestellt.

1.1 Wahrscheinlichkeiten und Verteilungen

Eine Zufallsvariable X beschreibt eine Größe, deren Werte aus einer Menge M zufällig ausgewählt werden. Diesem X liegt eine Wahrscheinlichkeitsdichte $f_X(x)$ (PDF) und eine kumulierte Wahrscheinlichkeit $F_X(x)$ zugrunde. Diese beschreibt die Wahrscheinlichkeit, dass der angenommene Wert $X \leq x$ ist. Es gilt

$$F_X(x) = \int_{-\infty}^x f_X(x') dx' . \quad (1.1)$$

Zusätzlich gelten

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \text{ und } \lim_{x \rightarrow \infty} F_X(x) = 1 , \quad (1.2)$$

außerdem folgt aus der Tatsache der steigenden Monotonie der CDF

$$F_X(x) \geq 0 \text{ und } f_X(x) \geq 0 . \quad (1.3)$$

Weiterhin ist

$$f_X(x) = \frac{\partial F_X(x)}{\partial x} . \quad (1.4)$$

Zur Approximation der Wahrscheinlichkeitsdichte kann ein Histogramm verwendet werden. Dazu werden die möglichen Werte in sog. Klassen K_i mit einer Breite b_i unterteilt. Üblicherweise wird eine konstante Breite gewählt, sodass sich in der graphischen Darstellung Balken mit Breite b und Höhe h_i ergeben. Für diese Höhe gilt

$$h_i = \sum_{j=1}^N \chi_i(x_j) , \quad (1.5)$$

wobei $\chi_i(x_j)$ die charakteristische Funktion beschreibt. Diese ist Eins, falls $x_j \in K_i$ ist und sonst Null. Mit dieser Definition werden absolute Häufigkeiten aufgetragen. Bei gleicher Breite der Balken kann durch $h_i \rightarrow h_i/N$ die relative Häufigkeiten zur Approximation der PDF verwendet werden.

1.2 Stochastische Momente

Die Momente m_k der Ordnung k einer PDF sind definiert als

$$m_k = \langle x^k \rangle = \int_{-\infty}^{\infty} x^k f_X(x) dx . \quad (1.6)$$

m_1 beschreibt den Erwartungswert, mit welchem die zentralen Momente der Ordnung k definiert werden

$$\overline{m_k} = \langle (x - m_1)^k \rangle = \int_{-\infty}^{\infty} (x - m_1)^k f_X(x) dx . \quad (1.7)$$

Es gilt definitionsgemäß $\overline{m_1} = 0$.

Das zweite zentrale Moment $\overline{m_2}$ beschreibt die Varianz. Sie ist die mittlere quadratische Abweichung und ist damit ein Maß für die Stärke der Streuung der einzelnen Werte um den Erwartungswert an. Ihre Wurzel ist die Standardabweichung σ .

Mit dem dritten zentralen Moment wird die Schiefe $\tilde{m}_3 = \overline{m_3}/\sigma^3$ definiert, sie beschreibt die Symmetrie der Verteilung. Eine symmetrische Verteilung besitzt eine Schiefe von Null. Für die Wölbung gilt $\tilde{m}_4 = \overline{m_4}/\sigma^4$. Eine große Wölbung bedeutet, dass die Verteilung um den Erwartungswert spitz ist, jedoch nur schwach abfallende Flügel besitzt.

1.3 Gaußverteilung und zentraler Grenzwertsatz

Die PDF der Gaußverteilung lautet

$$\mathcal{G}_{m_1, \overline{m_2}}(x) = \frac{1}{\sqrt{2\pi\overline{m_2}}} \exp\left(-\frac{(x - m_1)^2}{2\overline{m_2}}\right) , \quad (1.8)$$

wobei sie die Momente $\tilde{m}_3 = 0$ und $\tilde{m}_4 = 3$ besitzt. $\mathcal{G}_{0,1}$ wird Standardnormalverteilung genannt und ist zusammen mit anderen Gaußverteilungen beispielhaft in Abb. 1.1 zu sehen.

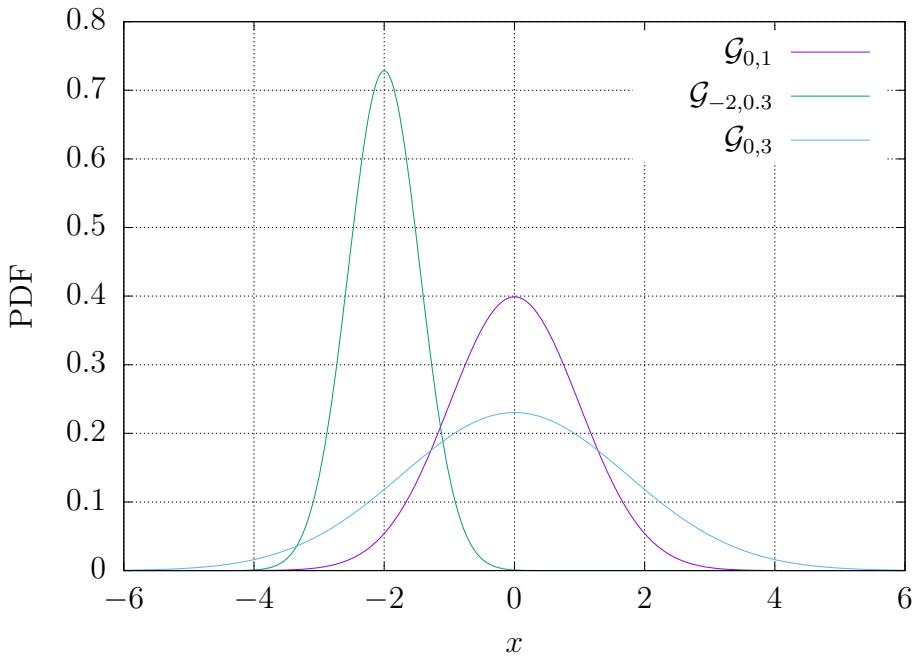


Abbildung 1.1: Wahrscheinlichkeitsdichten verschiedener Gaußverteilungen.

Zentraler Grenzwertsatz

Die Summe S_n unabhängiger Zufallszahlen mit identischer PDF, bei denen Erwartungswert und Varianz existieren, strebt gegen eine Gaußverteilung. Die Summe $S_n = \sum_i^n X_i$ definiert eine neue gaußverteilte Zufallsvariable Z_n gemäß [1]

$$Z_n = \frac{S_n - nm_1}{\sqrt{n\bar{m}_2}} . \quad (1.9)$$

1.4 Statistische Abhängigkeit

Zwei Zufallsvariablen können miteinander zusammenhängen, sie sind dann korreliert. Methoden zur Darstellung der Korrelation werden im Folgenden vorgestellt.

1.4.1 Pearson-Korrelationskoeffizient

Die Kovarianz zweier Zufallsvariablen A und B ist gegeben durch

$$\text{Kov}(A,B) = \langle (A - \langle A \rangle) \cdot (B - \langle B \rangle) \rangle , \quad (1.10)$$

woraus sich der Pearson-Korrelationskoeffizient durch Normierung mit den Standardabweichungen ergibt

$$\rho(A,B) = \text{Kov}(A,B)/\sigma_A\sigma_B . \quad (1.11)$$

Für diesen gilt $\rho(A,B) \in [-1,1]$. Allerdings kann der Pearson-Korrelationskoeffizient nur lineare Korrelation messen.

Bei $\rho(A,B) = 1$ besteht ein vollständig linearer positiver Zusammenhang zwischen A und B . Bei $\rho(A,B) = -1$ ist der Zusammenhang vollständig negativ. Für $\rho(A,B) = 0$ besteht kein linearer Zusammenhang zwischen den Zufallsvariablen. [2]

1.4.2 Autokorrelationsfunktion

Die Korrelation einer Zeitreihe X_t mit sich selbst bei einer Zeitverschiebung τ wird Autokorrelation genannt, sie ist der Korrelationskoeffizient $\rho(X_t, X_{t-\tau})$.

1.4.3 Copula

Die Copula wird verwendet um nichtlineare Korrelationen zwischen Zufallsvariablen zu beschreiben [3]. Es wird eine gemeinsame CDF für A und B aufgestellt

$$F_{A,B}(a,b) = \int_{-\infty}^a \int_{-\infty}^b f_{A,B}(a',b') da' db' , \quad (1.12)$$

wobei $f_{A,B}$ die gemeinsame PDF ist. Die PDF der einzelnen Variablen ergibt sich daraus gemäß

$$f_A(a) = \int_{-\infty}^{\infty} f_{A,B}(a,b) db \text{ bzw. } f_B(b) = \int_{-\infty}^{\infty} f_{A,B}(a,b) da . \quad (1.13)$$

Nun seien

$$u = F_A(a) \in [-1,1] \text{ und } v = F_B(b) \in [-1,1] , \quad (1.14)$$

wodurch sich die Quantils-Funktionen

$$a = F_A^{-1}(u) \text{ und } b = F_B^{-1}(v) \quad (1.15)$$

ergeben. Diese geben an, welcher Anteil an Werten kleiner ist als a bzw. b .

Die Copula ist dann definiert als

$$\text{Cop}_{A,B}(u,v) = F_{A,B}(F_A^{-1}(u), F_B^{-1}(v)) . \quad (1.16)$$

Sind die CDFs der Verteilungen stetig, so ist die Copula eindeutig. Zusätzlich gibt es eine Copuladichte $\text{cop}_{A,B}(u,v)$, welche gegeben ist durch

$$\text{cop}_{A,B}(u,v) = \frac{\partial^2}{\partial u \partial v} \text{Cop}_{A,B}(u,v) . \quad (1.17)$$

Eine Möglichkeit zur Darstellung dieser ist ein Qrank-Plot. In diesem werden die Quantils-Funktionen gegeneinander aufgetragen. Betrachtet man beispielsweise eine normalverteilte Verteilung X und eine Verteilung $Y = X^2 + \text{Rauschen}$, so ergibt sich als Korrelationskoeffizient $\rho(X,Y) \approx 0,01$.

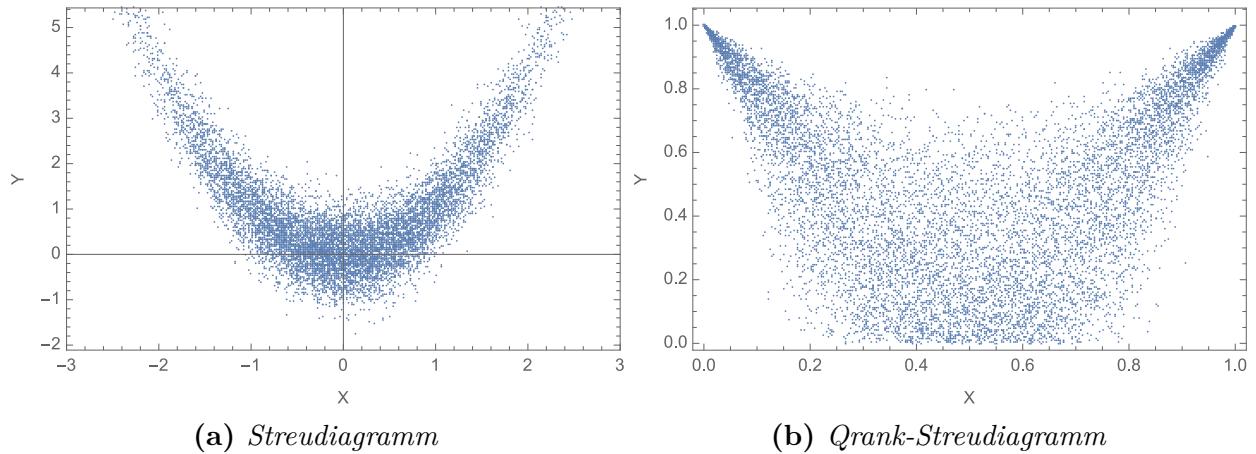


Abbildung 1.2: Darstellung des Zusammenhangs zweier Zufallsgrößen X und Y , die quadratisch zusammenhängen.

Wie in Abb. 1.2a zu sehen ist, gibt es eine quadratische Abhängigkeit. Diese lässt sich durch den Pearson-Korrelationskoeffizienten nicht bestimmen. In dem Qrank-Plot in Abb. 1.2b ist diese allerdings erkennbar. Betrachtet man nun einen 3D-Plot der Copuladichte, Abb. 1.3, so erkennt man die Abhängigkeit deutlich. Große Werte von X treten mit großen Werten von Y unabhängig vom Vorzeichen auf.

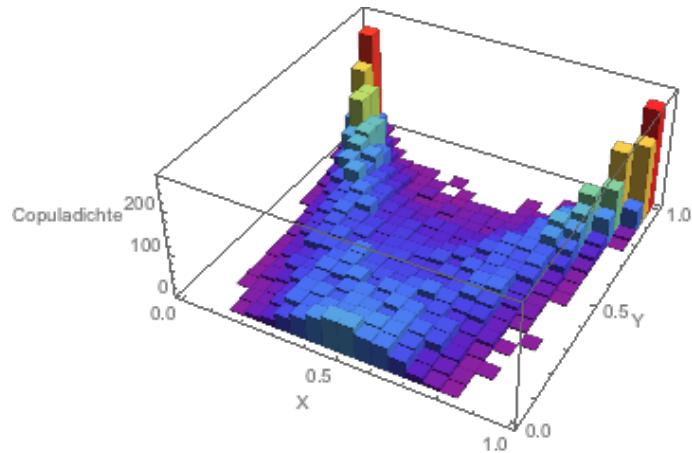


Abbildung 1.3: Copuladichte zweier quadratisch abhängiger Variablen X und Y .

Daher ist die Copula gut geeignet nichtlineare Korrelation zu bestimmen.

1.5 Die untersuchten Systeme

Nachfolgend wird auf die untersuchten Systeme eingegangen und in wie fern sie statistische Eigenschaften aufweisen.

1.5.1 Brownsche Bewegung

Die Brownsche Bewegung ist ein rein deterministischer Prozess. Jedoch wirkt die Bewegung völlig zufällig und lässt sich durch einen Zufallsprozess beschreiben. Dies liegt an der großen Anzahl an Stößen pro Sekunde (10^{21}). Jeder Stoß sorgt für eine Impulsänderung, welche unabhängig und gleichverteilt sind. Dadurch ergibt sich durch den zentralen Grenzwertsatz für alle Impulsänderungen zusammen eine Gaußverteilung.

1.5.2 Aktienkurse

Die zeitliche Entwicklung von Aktienkursen wirkt zufällig. Jedoch spiegeln Aktienkurse die zeitliche Markterwartung eines Unternehmens wider und die Akteure handeln keinesfalls zufällig. Der erste Effekt führt zu einem sog. Drift. Analog zur Brownschen Bewegung kommt der zufällige Anteil aus der großen Anzahl an Transaktionen pro Tag (10^4). Die Renditen verhalten sich jedoch nicht gemäß einer Gaußverteilung sondern besitzen ein stärkeres Gewicht in den Flanken.

1.5.3 Temperaturen

Das Wetter ist ein Paradebeispiel für ein chaotisches System. Eine Vorhersage ist sehr schwierig und nur für kleine Zeiträume möglich. Die Temperaturdaten zeigen ein zufälliges Verhalten, jedoch gibt es räumliche und zeitliche Korrelationen. Außerdem gibt es eine Abhängigkeit von der Tages- und Jahreszeit.

2 Brownsche Bewegung

Es wurde die Bewegung von Teilchen in einer Flüssigkeit mit einer Kamera in gewissen Abständen aufgenommen. Anschließend wurden die Trajektorien der Teilchen mit Hilfe von ImageJ getrackt.

Als erstes werden nun alle Trajektorien, die kürzer als 50 Zeitschritte sind verworfen. Diese sind zu kurz für eine statistische Aussage. Danach werden die Differenzen Δx_n und Δy_n gemäß

$$\Delta a_n = a_{n+1} - a_n \text{ für } n \in [1, L_i - 1] \quad (2.1)$$

berechnet, wobei L_i die Zahl der Zeitschritte der jeweiligen Trajektorie bezeichnet. Nun werden alle Trajektorien verworfen, deren Differenzen eine Kurtosis $\tilde{m}_4 > 6$ besitzen. Auf Grund des zentralen Grenzwertsatzes wird eine Gaußverteilung mit einer Kurtosis von 3 erwartet, daher wird eine Verteilung von $\tilde{m}_4 \in [0, 6]$ verwendet, um symmetrisch um den erwarteten Wert zu streuen.

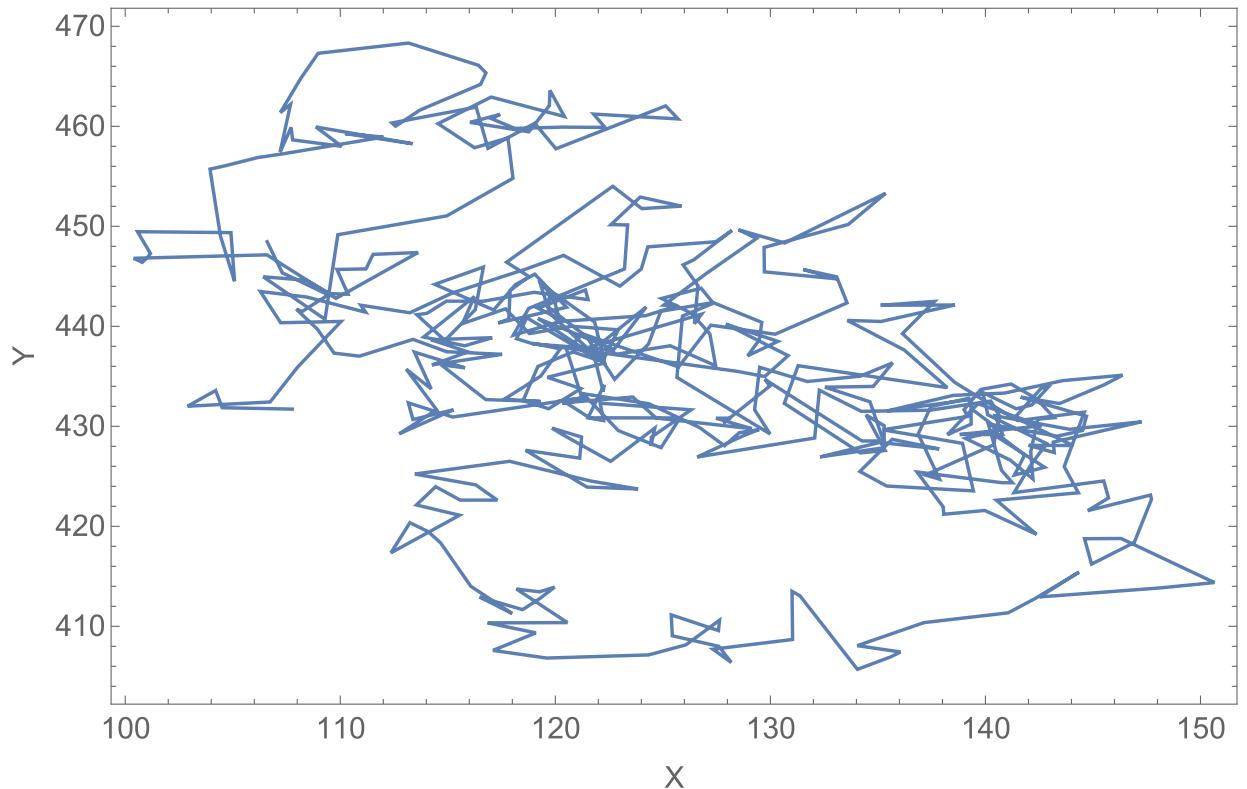


Abbildung 2.1: Darstellung der längsten Trajektorie der Brownschen Bewegung.

In Abb. 2.1 ist die längste aufgenommene Trajektorie dargestellt. Wie zu erkennen ist, gibt es eine zufällige Bewegung.

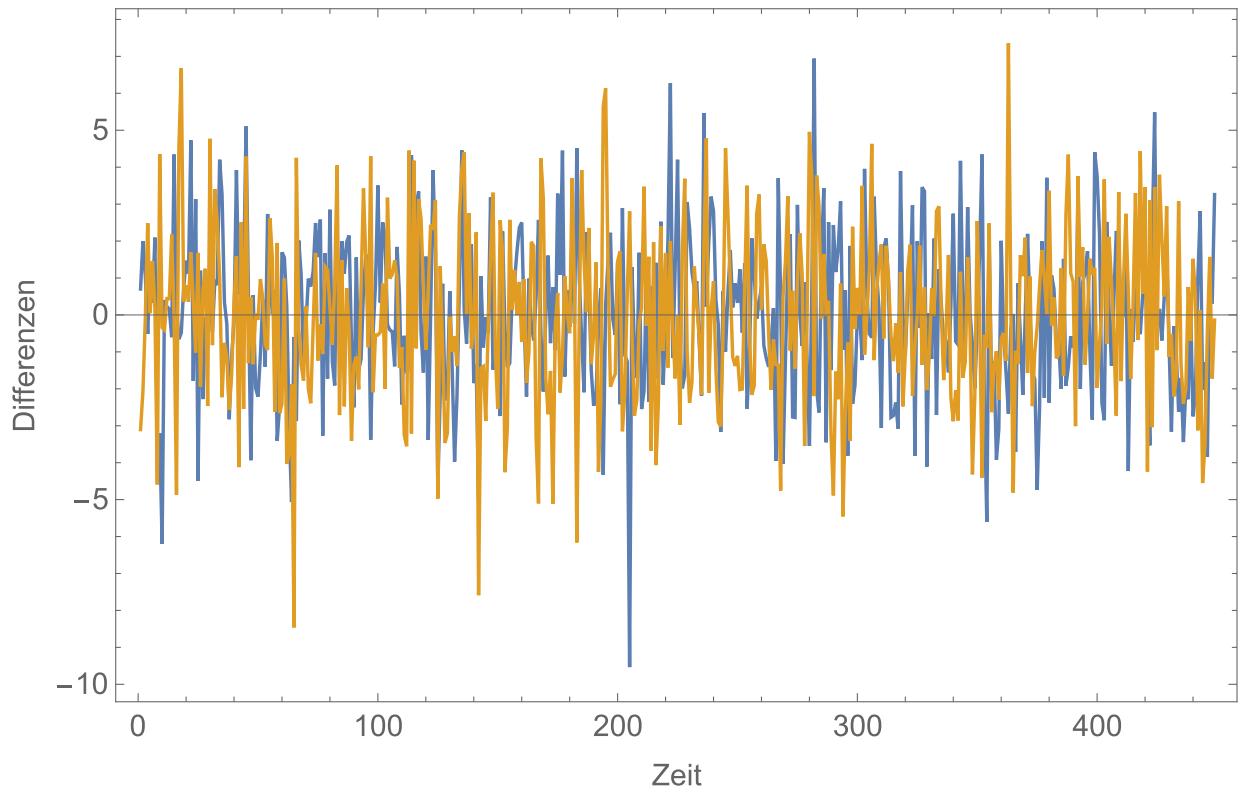


Abbildung 2.2: Zeitreihe der Differenzen Δx und Δy der längsten Trajektorie.

Diese zufällige Bewegung kann ebenfalls an den Zeitreihen der Differenzen Δx und Δy in Abb. 2.2 gesehen werden. Diese Differenzen streuen symmetrisch um 0.

Nun werden für alle Trajektorien die statistischen Momente bestimmt und der Mittelwert gebildet.

Tabelle 2.1: Erwartungswert, Varianz, Schiefe und Kurtosis der Verteilungen der Differenzen.

	Δx	Δy
m_1	-0,014	0,077
\bar{m}_2	6,787	6,993
\tilde{m}_3	-0,058	0,0153
\tilde{m}_4	3,226	3,1991

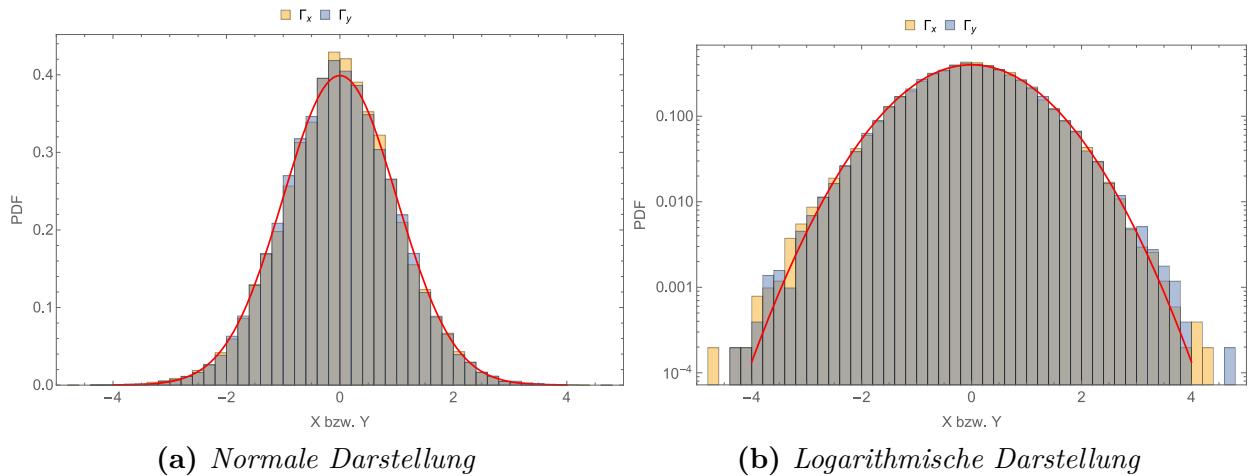
Wie in Tabelle 2.1 zu erkennen ist, stimmen der Erwartungswert, die Schiefe und Kurtosis mit der Normalverteilung überein. Dies war auf Grund des zentralen Grenzwertsatzes zu erwarten. Die Varianz ist allerdings um einiges größer als bei einer Standardnormalverteilung..

Als nächstes werden die einzelnen Zeitreihen der Differenzen Δx und Δy auf Erwartungswert Null und Standardabweichung Eins normiert. Diese normierten Zeitreihen werden nun jeweils zu neuen Zeitreihen Γ_x und Γ_y zusammengefasst. Für diese ergeben sich für die Momente die in Tabelle 2.2 dargestellten Werte.

Tabelle 2.2: Erwartungswert, Varianz, Schiefe und Kurtosis der neuen Zeitreihen Γ_x und Γ_y .

	Γ_x	Γ_y
m_1	$7,5 \cdot 10^{-18}$	$-8,37 \cdot 10^{-18}$
\bar{m}_2	0,988	0,988
\tilde{m}_3	-0,048	0,012
\tilde{m}_4	3,256	3,189

Diese Momente entsprechen der Standardnormalverteilung, wie in Abb. 2.3 zu sehen ist.


Abbildung 2.3: Histogramm der neuen Zeitreihen zusammen mit einer Standardnormalverteilung.

Trägt man nun die neuen Zeitreihen Γ_x und Γ_y in einem Streudiagramm gegeneinander auf, so zeigt sich, dass die beiden Richtungen unkorreliert sind. Es ergibt sich näherungsweise ein Kreis (siehe Abb. 2.4a).

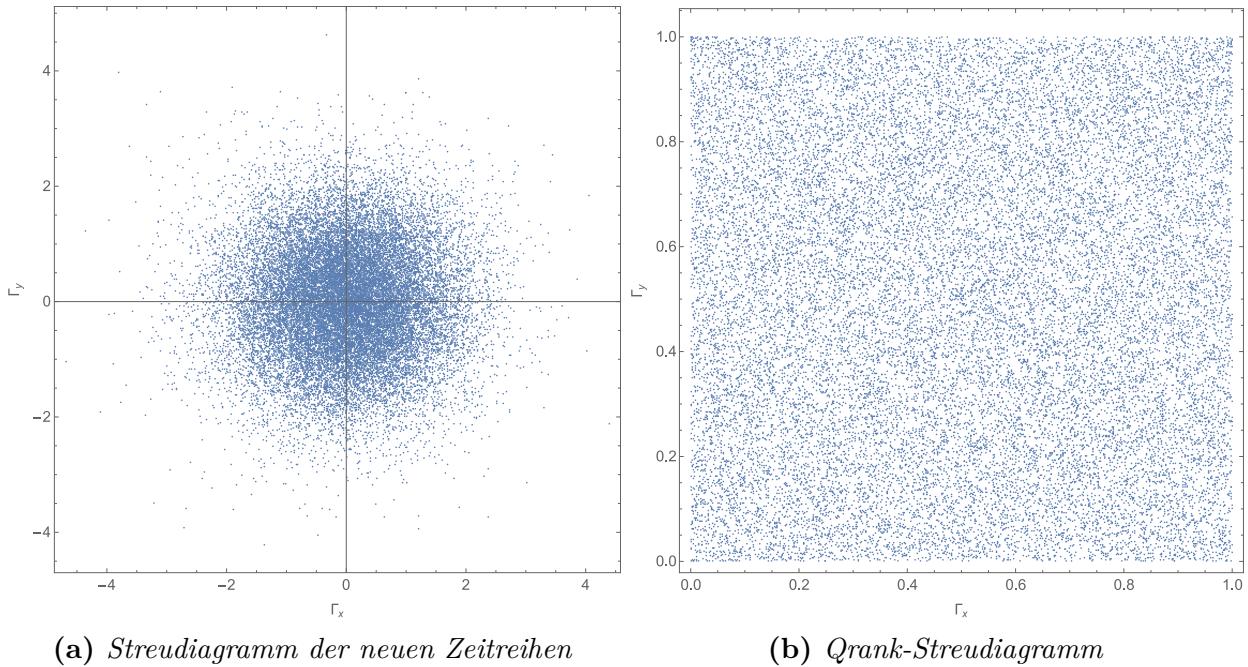


Abbildung 2.4: Streudiagramme für die Zeitreihen Γ_x und Γ_y .

Mit Hilfe einer Qrank-Funktion wird nun die Copula untersucht. Das zugehörige Streudiagramm ist in Abb. 2.4b zu sehen. Es ergibt sich ein Quadrat, d.h. es tritt jeder Wert von Γ_x mit jedem Wert von Γ_y auf. Dies zeigt sich ebenfalls an einem 3D-Histogramm der Copuladichte. Außerdem ist dieser Zusammenhang in Abb. 2.5 dargestellt. Es gibt keine Wertepaare, die deutlich häufiger miteinander auftreten als andere.

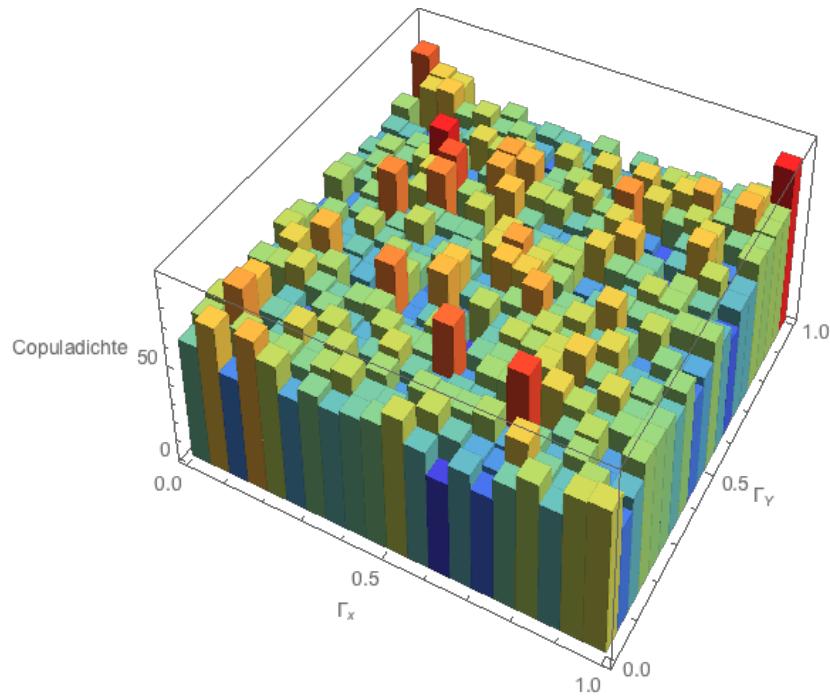
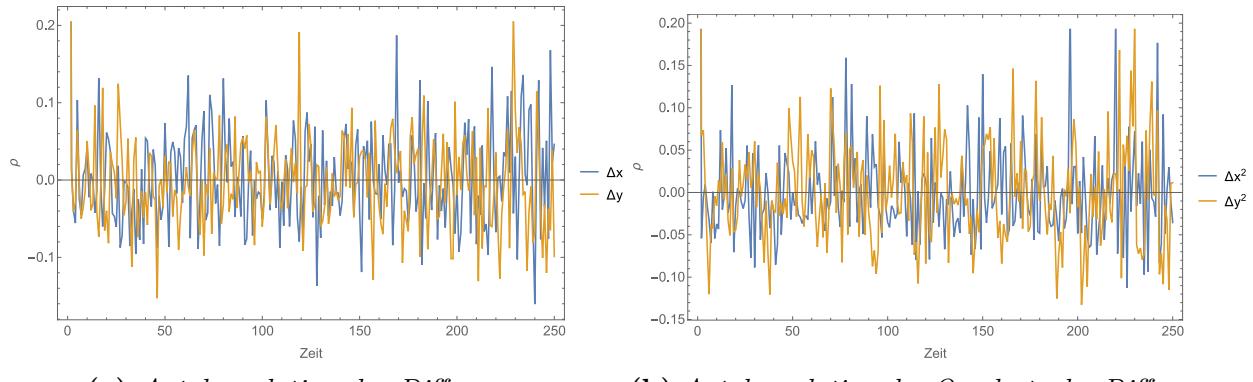


Abbildung 2.5: Copuladichte der Zeitreihen Γ_x und Γ_y .

Als nächstes werden die Autokorrelationen der Differenzen Δx und Δy und der Quadrate der Differenzen der längsten Trajektorie untersucht.



(a) Autokorrelation der Differenzen

(b) Autokorrelation der Quadrate der Differenzen

Abbildung 2.6: Autokorrelationen für die Differenzen Δx und Δy der längsten Trajektorie.

Es zeigt sich in beiden Fällen keine Autokorrelation, also kein linearer Zusammenhang. Weiterhin zeigt sich anhand eines Streudiagramms der Zeitreihe Δx und Δy jeweils mit sich selbst um einen Zeitschritt verschoben, dass es auch keine nichtlineare Korrelation gibt. In Abb. 2.7 zeigt sich für beide Zeitreihen näherungsweise ein Kreis.

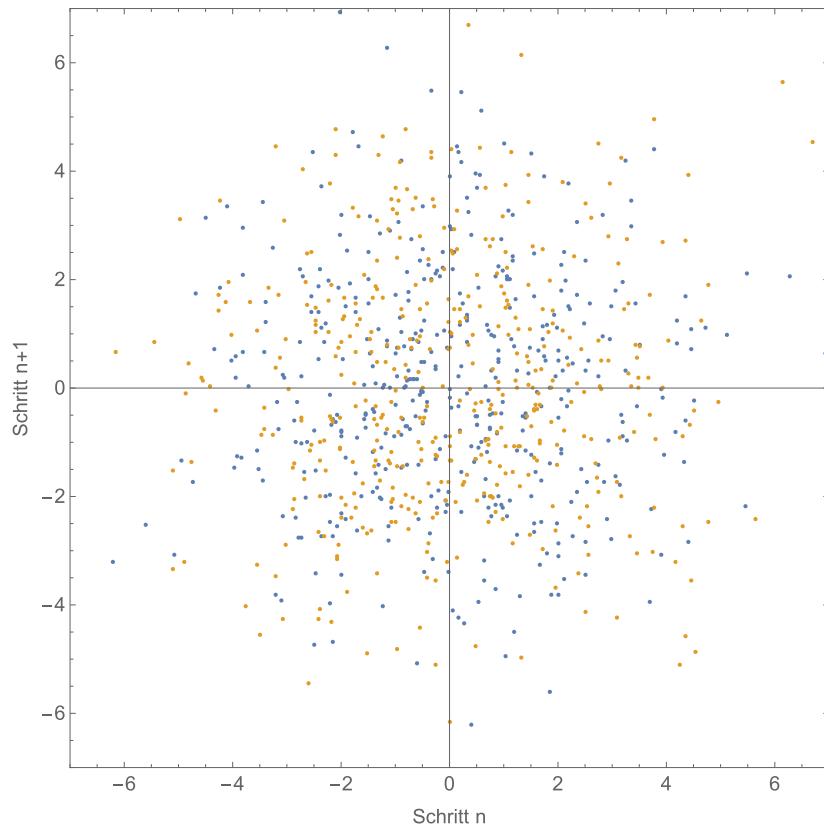


Abbildung 2.7: Streudiagramm für Δx (orange) und Δy (blau).

An den Copuladichten, Abb. 2.8 zeigt sich das gleiche. Es gibt keine Wertepaare, die deutlich häufiger miteinander auftreten als andere.

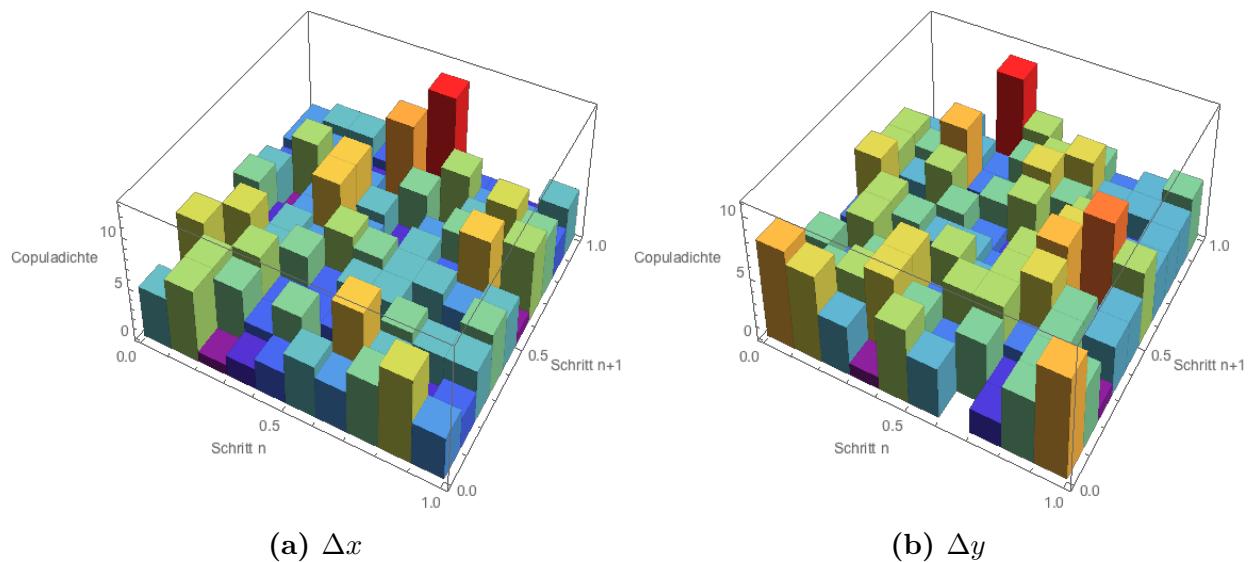


Abbildung 2.8: Copuladichten für die Zeitreihen Δx und Δy mit sich selbst um einen Zeitschritt verschoben.

3 Finanzdaten

In diesem Versuchsteil werden die statistischen Eigenschaften von Aktienpreiszeitreihen untersucht. Dazu werden mit Hilfe des Programmes „Mathematica“ die Tagesabschlusspreise $S(t)$, im Folgenden immer als Preise der Aktien bezeichnet, von neun verschiedenen Unternehmen aus dem S&P-500-Index heruntergeladen. Die verwendeten Unternehmen sind Apple (AAPL), IBM (IBM), 3M (MMM), Microsoft (MSFT), General Electric (GE), British Petroleum (BP), Coca Cola (KO), McDonalds (MCD) und Disney (DIS). Eine Preiszeitreihe ist beispielsweise für die Apple-Aktie in Abb. 3.1 dargestellt.

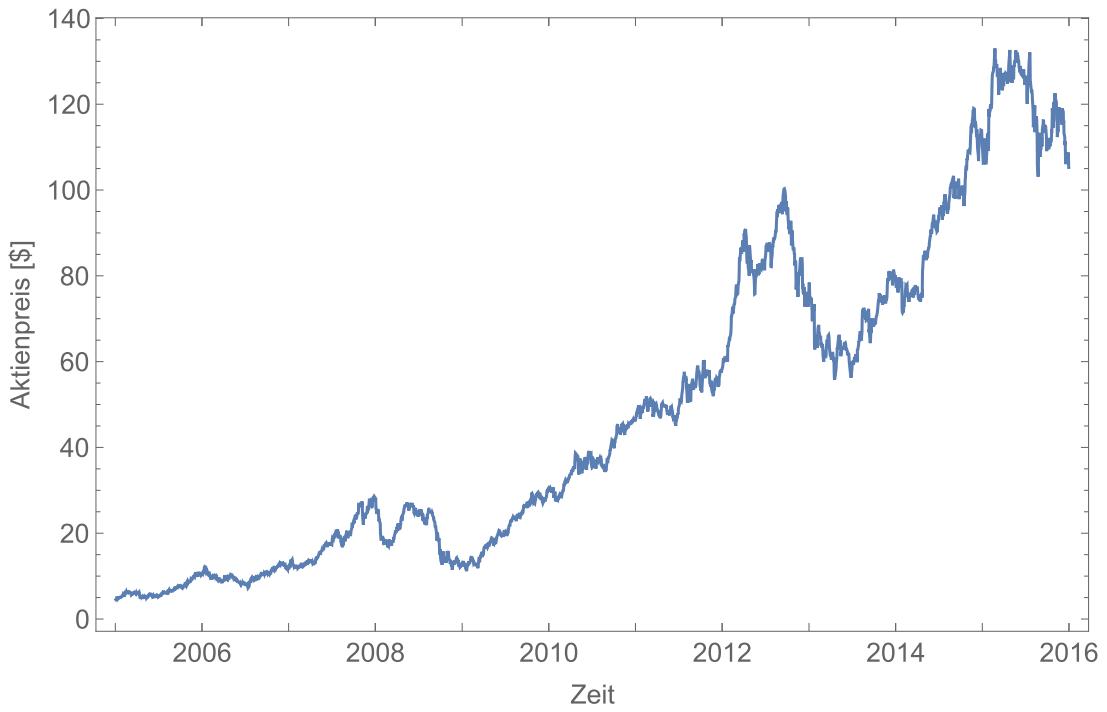


Abbildung 3.1: Preiszeitreihe der Apple-Aktie.

Über den Zeitraum von zehn Jahren ist neben Schwankungen ein exponentieller Anstieg des Preises festzustellen.

Die relativen Preisänderungen

$$R(t) = \frac{S(t + \Delta t) - S(t)}{S(t)} , \quad (3.1)$$

werden als Returns (Renditen) bezeichnet. Eine Betrachtung der täglichen Returns, $\Delta t = 1 \text{ d}$, zeigt, dass deren Verteilung von einer Gaußverteilung abweicht (Abb. 3.2). Es gibt sehr große Returns,

welche bei einer Gaußverteilung nicht auftreten würden. Große Preisbewegungen in beide Richtungen sind dadurch möglich.

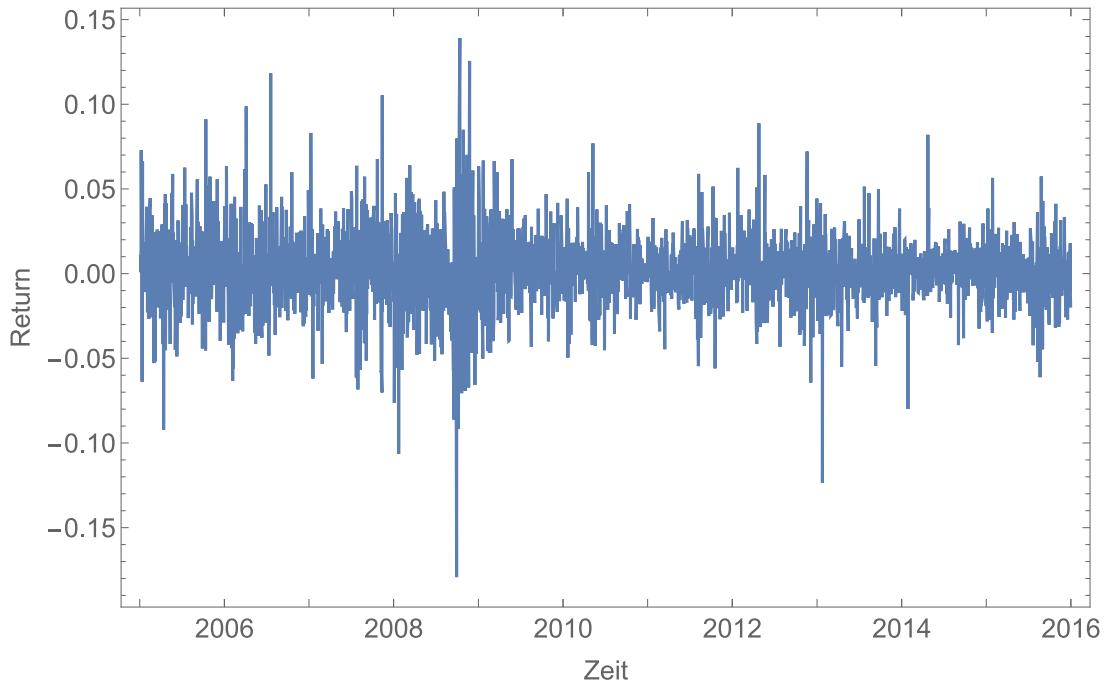


Abbildung 3.2: Returnzeitreihe der Apple-Aktie.

Die statistischen Momente für alle ausgewählten Aktien sind in Tabelle 3.1 dargestellt.

Tabelle 3.1: Erwartungswert, Varianz, Schiefe und Kurtosis der ausgesuchten Aktien.

	AAPL	IBM	MMM	MSFT	GE	BP	KO	MCD	DIS
m_1	0.0014	0.0002	0.0003	0.0004	0.0001	0.0000	0.0003	0.0006	0.0006
\overline{m}_2	0.0005	0.0002	0.0002	0.0003	0.0004	0.0004	0.0001	0.0002	0.0003
\tilde{m}_3	-0.0323	-0.0510	-0.1780	0.4421	0.4397	-0.09475	0.7808	0.3305	0.4435
\tilde{m}_4	7.8605	9.1701	9.3147	14.8476	16.2576	14.6145	17.7966	8.1446	12.1559

Der Mittelwert liegt für alle Aktien sehr nah an der Null, ist jedoch immer positiv, es wurde allerdings auf vier Nachkommastellen gerundet, weshalb bei dem Erwartungswert von BP eine Null steht. Dies entspricht einer positiven Entwicklung der Aktienkurse. Die Standardabweichung ist um eine Größenordnung größer als der Mittelwert, die täglichen Schwankungen sind also vergleichsweise groß gegenüber dem täglichen Trend. Die Verteilungen besitzen entweder eine kleine Schiefe in negative Richtung oder eine deutlich größere in positive Richtung. Große Returns treten also verstärkt in positiver Richtung auf. Alle Verteilungen besitzen eine deutlich größere Kurtosis als eine Gaußverteilung, d.h. sie besitzen viel ausgeprägtere Flügel, große Returns sind also nicht unwahrscheinlich. Ein Investor muss auf Grund der kleinen Schiefe und großen Kurtosis dementsprechend jederzeit mit Kursbewegungen in beide Richtungen rechnen.

Um eine bessere Vergleichbarkeit der Aktien untereinander und zur Gaußverteilung zu erhalten,

werden nun alle Returnverteilungen normiert, indem der Mittelwert abgezogen und durch die Standardabweichung geteilt wird. In Abb. 3.3 ist die normierte Verteilung der Returns der Apple-Aktie zusammen mit einer Standardnormalverteilung dargestellt. Es ist deutlich zu erkennen, dass die Returnverteilung viel breiter ist als die Standardnormalverteilung.

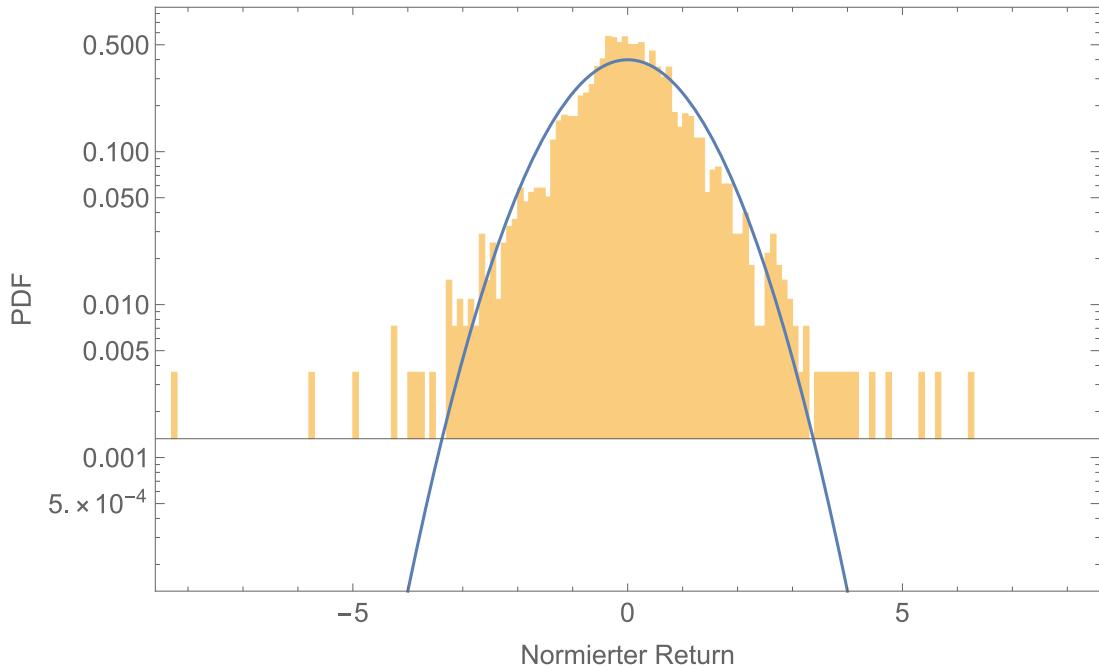


Abbildung 3.3: Histogramm der normierten Returns der Apple-Aktie zusammen mit einer Standardnormalverteilung.

Um nun die Returns zweier Aktien untereinander zu vergleichen, wird ein Streudiagramm aus den normierten Returns gebildet. Für die Apple- und IBM-Aktie, zeigt dieses einen leichten linearen Trend. Ein Streudiagramm der Quantilsfunktionen der Returns zeigt kleine Häufungen in der unteren linken und oberen rechten Ecke, also ebenfalls einen leicht linearen Trend. Diese beiden Darstellungen sind in Abb. 3.4 zu sehen.

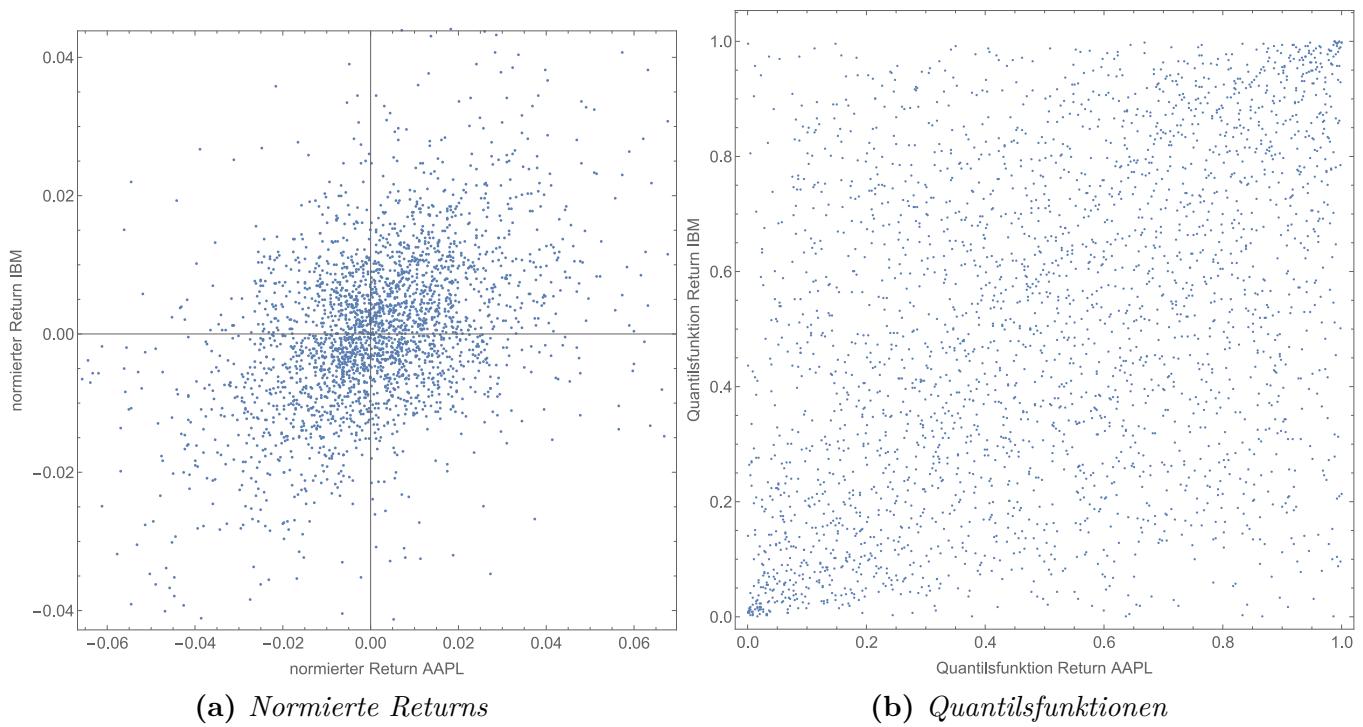


Abbildung 3.4: Streudiagramme der Returns und Quantilsfunktionen der normierten Returns der Apple- und IBM-Aktie.

Die Copuladichte dieses Zusammenhangs ist in Abb. 3.5 dargestellt. Es ist zu erkennen, dass große positive und negative Returns jeweils gemeinsam auftreten.

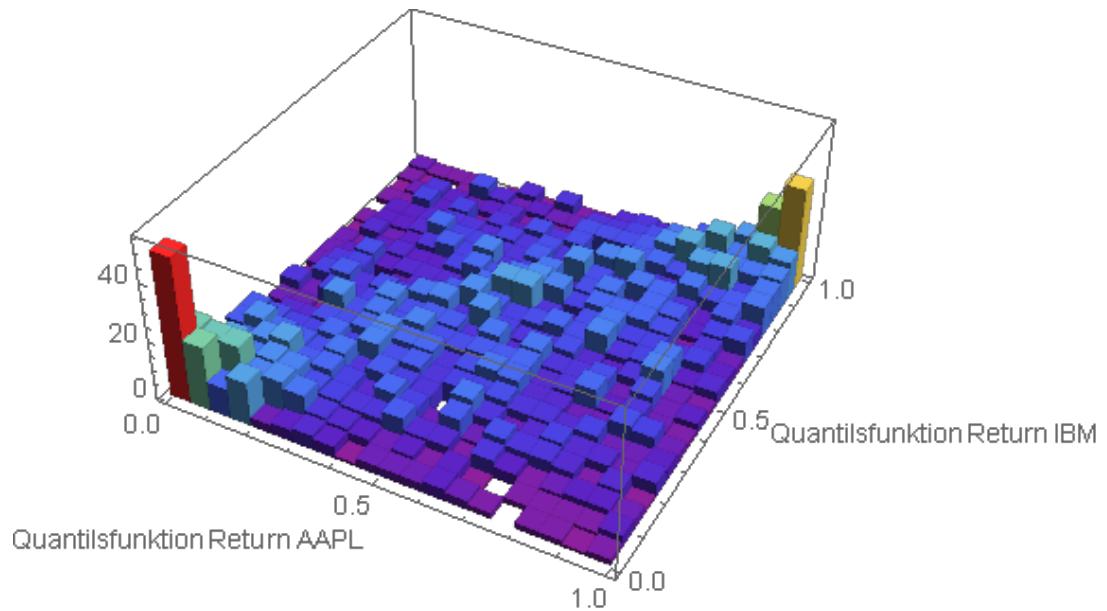


Abbildung 3.5: Copuladichte der Returns der Apple- und IBM-Aktie.

Nun soll nicht die Korrelation der Returns von Unternehmen untereinander, sondern die Autokorrelation der Returns einer Aktie untersucht werden. Am Beispiel der Apple-Aktie ist die Autokorrelation der Returns und des Quadrates der Returns in Abb. 3.6 dargestellt. Es ist zu erkennen, dass für die Returns keine Autokorrelation vorliegt, die Funktion schwankt um Werte nahe der Null. Für die Quadrate der Returns (und somit die Beträge) ist dies allerdings anders. Die Autokorrelation fällt lediglich langsam ab, d.h. auf Preisschwankungen folgen Preisschwankungen in der selben Größenordnung, sowohl in positiver als auch negativer Richtung. Dieser Effekt nennt sich Volatility-Clustering. In Abb. 3.2 ist dies bereits zu erkennen, es gibt Zeiten großer und kleiner Returns.

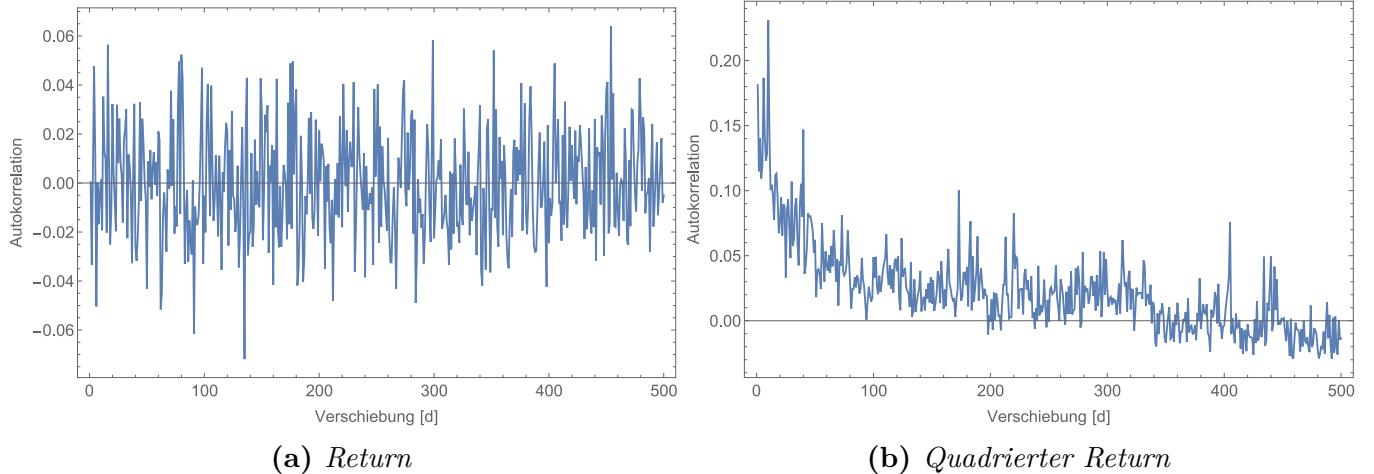


Abbildung 3.6: Autokorrelationsfunktionen der Returns und der quadrierten Returns der Apple-Aktie.

Um neben linearen Zusammenhängen auch noch andere Zusammenhänge zu berücksichtigen, werden nun Streudiagramme für die Returns der Apple-Aktie an zwei aufeinanderfolgenden Tagen erstellt. Diese sind in Abb. 3.7 dargestellt. Das normale Streudiagramm lässt keinen deutlichen Zusammenhang erkennen, es ist annähernd kreisförmig. In dem Streudiagramm der Quantilsfunktionen hingegen lässt sich jeweils eine leichte Häufung in den Ecken erkennen. Große Returns bedingen also erneut große Returns, sowohl in positiver als auch negativer Richtung.

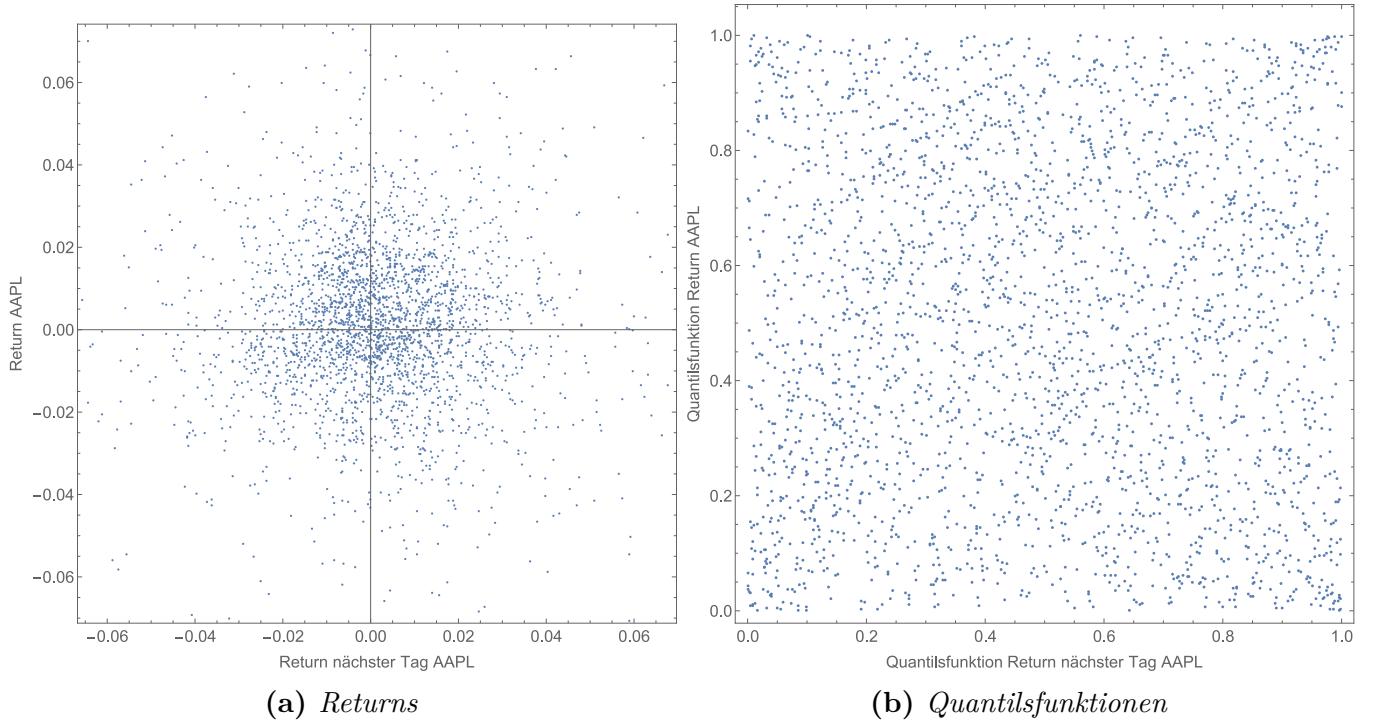


Abbildung 3.7: Streudiagramme der Returns und Quantilsfunktionen der Returns der Apple-Aktie an zwei aufeinanderfolgenden Tagen.

In einem dreidimensionalen Histogramm der Copuladichte, dargestellt in Abb. 3.8, ist dieser Zusammenhang noch deutlicher zu erkennen. Große Returns bedingen große Returns am Folgetag. In diesem Fall bedingen insbesondere große negative Returns große positive Returns am Folgetag.

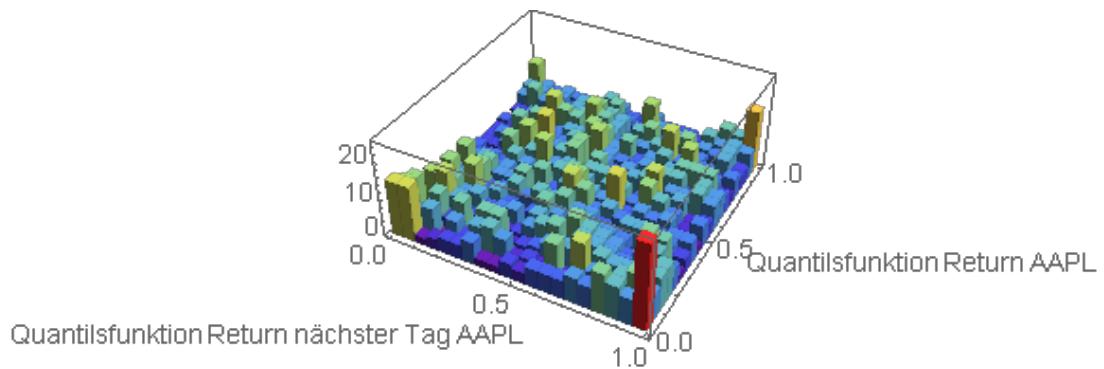


Abbildung 3.8: Copuladichte der Returns der Apple-Aktie an zwei aufeinanderfolgenden Tagen.

4 Temperaturdaten

In diesem Versuchsteil werden die statistischen Eigenschaften von Temperaturzeitreihen betrachtet. Dazu wurden die Temperaturen von neun Städten analysiert. Es wurde jeder Kontinent abgedeckt und zwei Städte aus Deutschland gewählt. Die Städte sind Duisburg, Frankfurt (Oder), Hooker (USA), Vitoria (Brasilien), Johannesburg (Südafrika), Pjöngjang (Nordkorea), Perth (Australien), Honolulu (Hawaii), Alert (Kanada). Eine Übersicht über die genauere Lage ist in Abb. 4.1 zu sehen.



Abbildung 4.1: Lage der ausgewählten Städte. Jeder Kontinent ist abgedeckt und zwei Städte liegen in Deutschland.

Für alle Städte wurde mit dem Programm „Mathematica“ die Temperatur vom 1.1.08 bis 1.1.18 heruntergeladen und analysiert. Eine Temperaturzeitreihe von Alert (Kanada) ist in Abb. 4.2 exemplarisch dargestellt. Die Jahreszeiten erzeugen klar erkennbare Temperaturschwankungen.

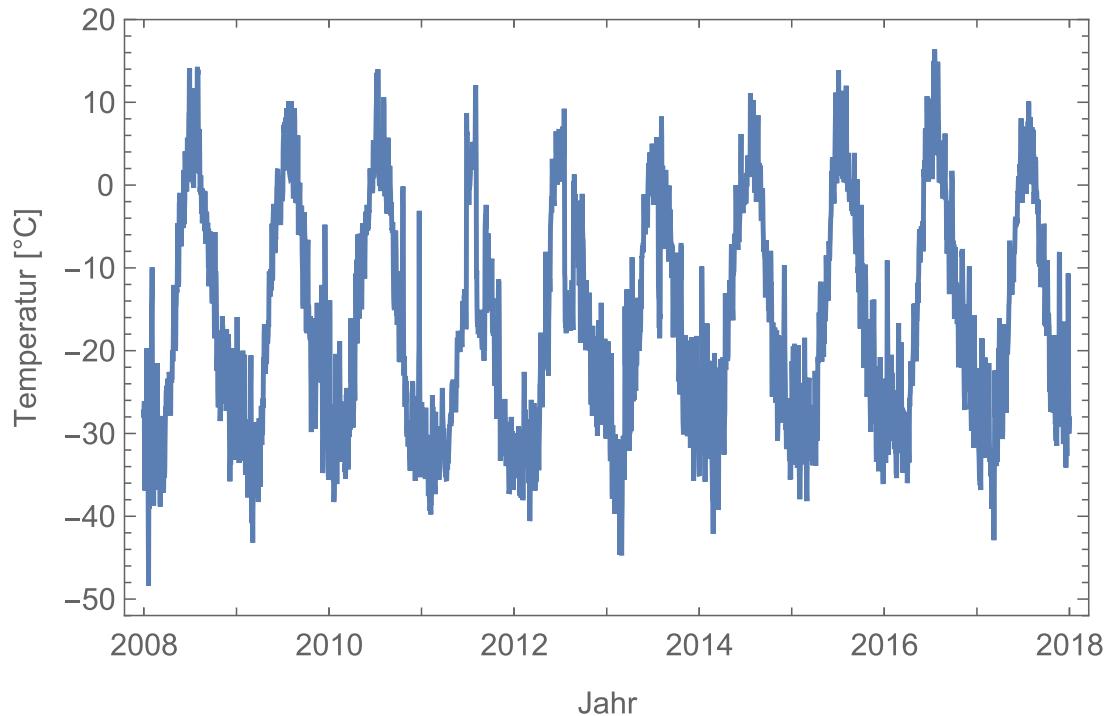


Abbildung 4.2: Temperaturzeitreihe von Alert (Kanada).

Aus diesen Verteilungen kann man die statistischen Momente berechnen. Diese sind in Tabelle 4.1 dargestellt.

Tabelle 4.1: Erwartungswert, Varianz, Schiefe und Kurtosis der Temperaturen der ausgewählten Städte.

	Du	Fr	Ho	Vi	Jo	Pj	Pe	Ho	Al
m_1	11.7200	9.6744	13.5470	24.8880	16.7410	11.1790	18.6640	25.7560	-16.8330
\bar{m}_2	6.9958	7.4198	10.8320	2.5190	4.3126	11.4560	5.2030	2.0120	13.3140
\tilde{m}_3	-0.0825	-0.0778	-0.2085	-0.1335	-0.4336	-0.2955	0.2902	-0.1054	0.2836
\tilde{m}_4	2.4522	2.4394	2.0464	2.1383	2.8263	1.8259	2.3647	2.6888	1.9255

Die Temperaturdaten werden nun normiert, indem der Mittelwert abgezogen und durch die Standardabweichung geteilt wird. In dem Histogramm der normierten Temperaturen, Abb. 4.3, sind zwei Maxima rechts und links des Mittelwertes zu erkennen. Es scheint wie zwei separate Gaußverteilungen um eine hohe und eine tiefe Temperatur. Dies kann damit zusammenhängen, dass Sommer und Winter dominant und die Übergangsjahreszeiten sehr kurz sind. Die Temperaturen streuen um einen Wert für den Winter und um einen Wert für den Sommer. Diese Aufspaltung ist extremer, je weiter die Stadt vom Äquator entfernt ist. Dies hängt damit zusammen, dass die Ausprägung der Jahreszeiten mit dem Abstand zum Äquator zunimmt und es am Äquator nur noch ein Tageszeitenklima gibt.

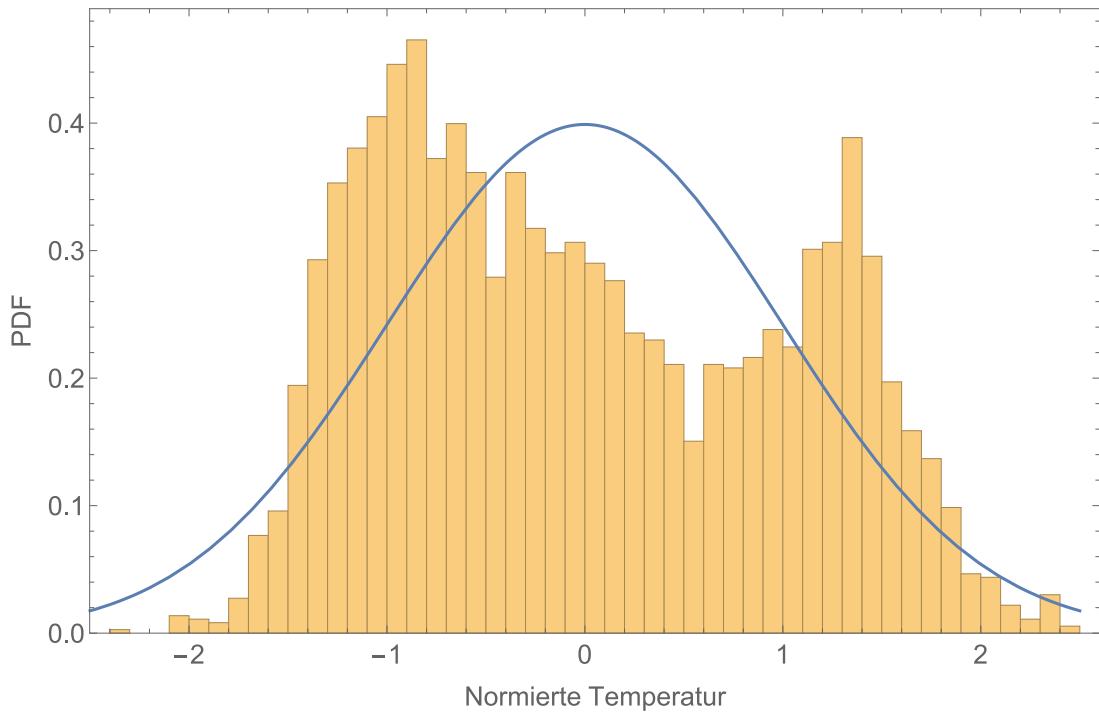


Abbildung 4.3: Histogramm der Temperatur in Alert (Kanada) und Standardnormalverteilung.

Nun wird der Zusammenhang zwischen den Temperaturen zweier Orte betrachtet. In Abb. 4.4a ist ein Streudiagramm der normierten Temperaturen von Duisburg und Frankfurt zu sehen. Es ist ein deutlicher Zusammenhang zu erkennen. Hohe und tiefe Temperaturen treten in Duisburg und Frankfurt jeweils gemeinsam auf. In Abb. 4.4b ist das Streudiagramm der Quantilsfunktionen und in Abb. 4.5 die Copuladichte der Temperaturen von Duisburg und Frankfurt dargestellt. Diese bestätigen den beschriebenen Zusammenhang.

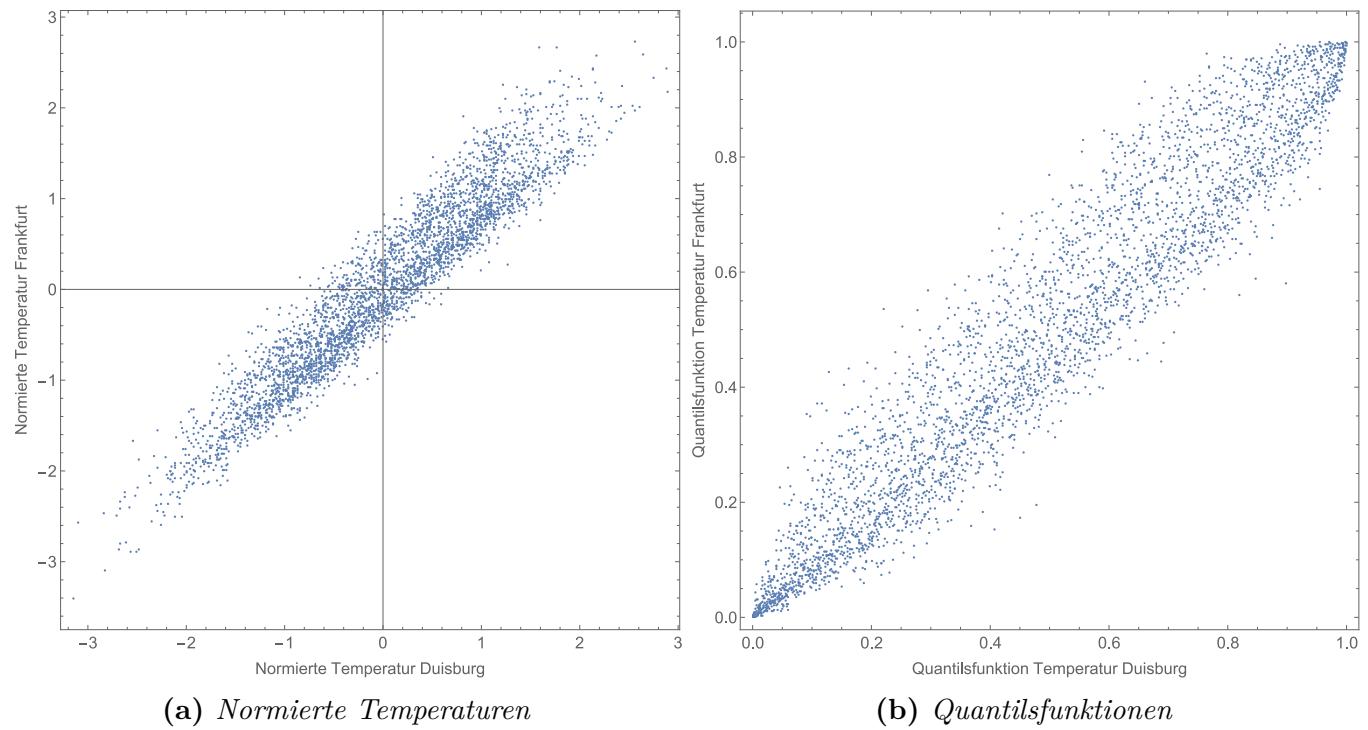


Abbildung 4.4: Streudiagramm der normierten Temperaturen und der Quantilsfunktionen der Temperaturen von Duisburg und Frankfurt.

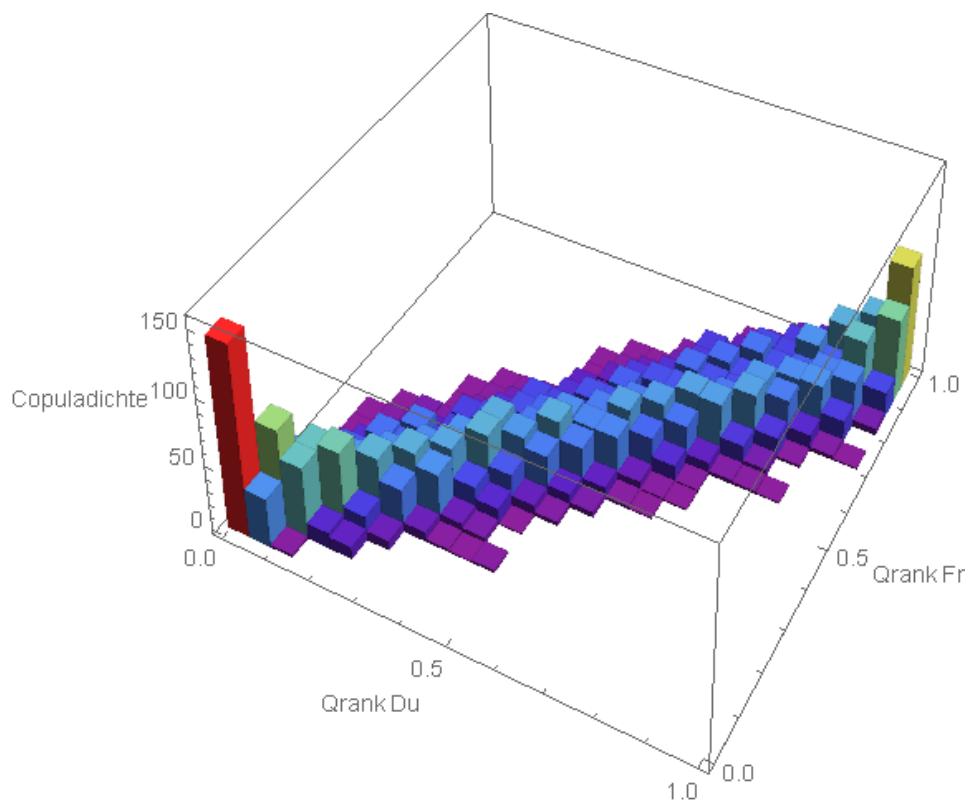


Abbildung 4.5: *Copuladichte der Temperaturen von Duisburg und Frankfurt.*

Wird die Copuladichte zwischen zwei Städten in unterschiedlichen Hemisphären betrachtet, ist erkennbar, dass sich die Temperaturen invers zueinander verhalten. In Abb. 4.6 zeigt sich, dass wenn in Duisburg ein warmer Tag ist, in Perth (Australien) wahrscheinlich ein kalter Tag ist. Dieser Zusammenhang ist allerdings nicht mehr so stark wie der zwischen Duisburg und Frankfurt, da Duisburg und Perth eine deutlich größere räumliche Distanz zueinander haben.

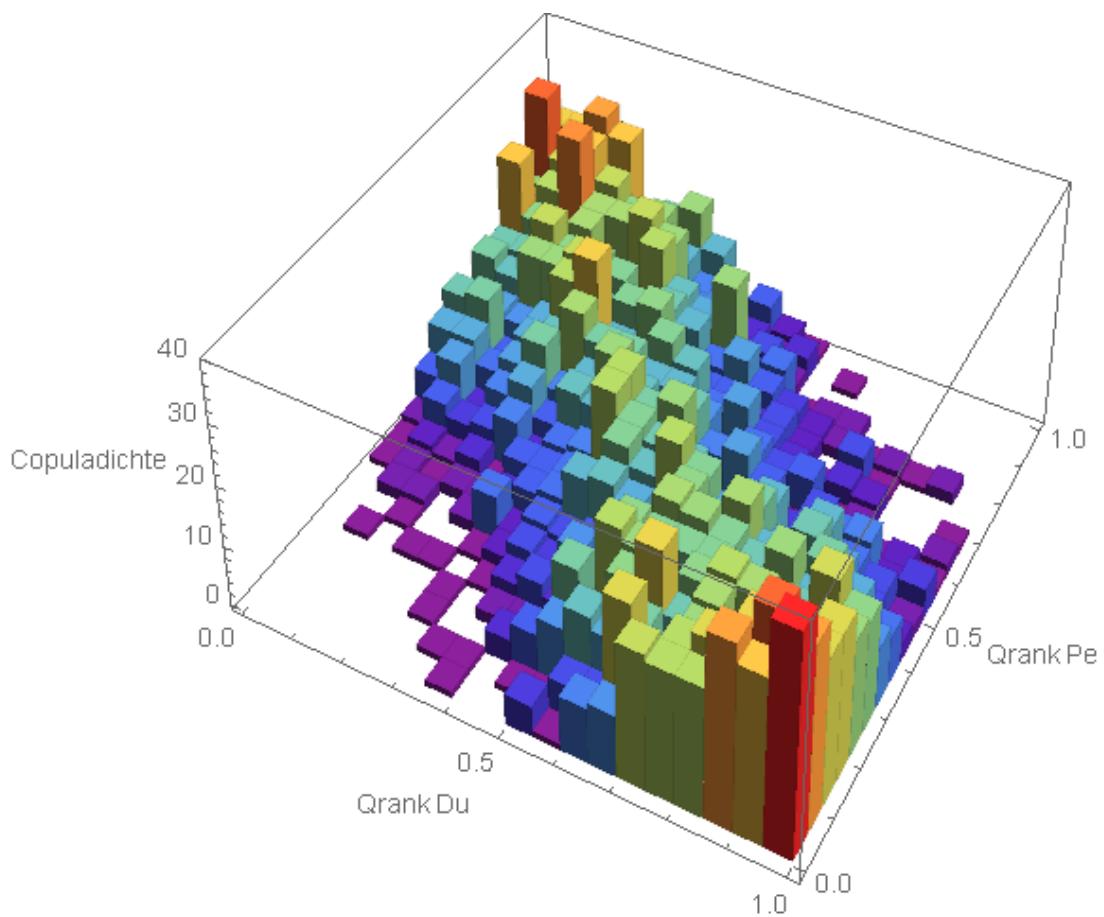


Abbildung 4.6: Copuladichte der Temperaturen von Duisburg und Perth.

Die Autokorrelation in Abb. 4.7 der Temperatur einer Stadt weist ein periodisches Verhalten auf, dies ist auf die Jahreszeiten zurückzuführen.

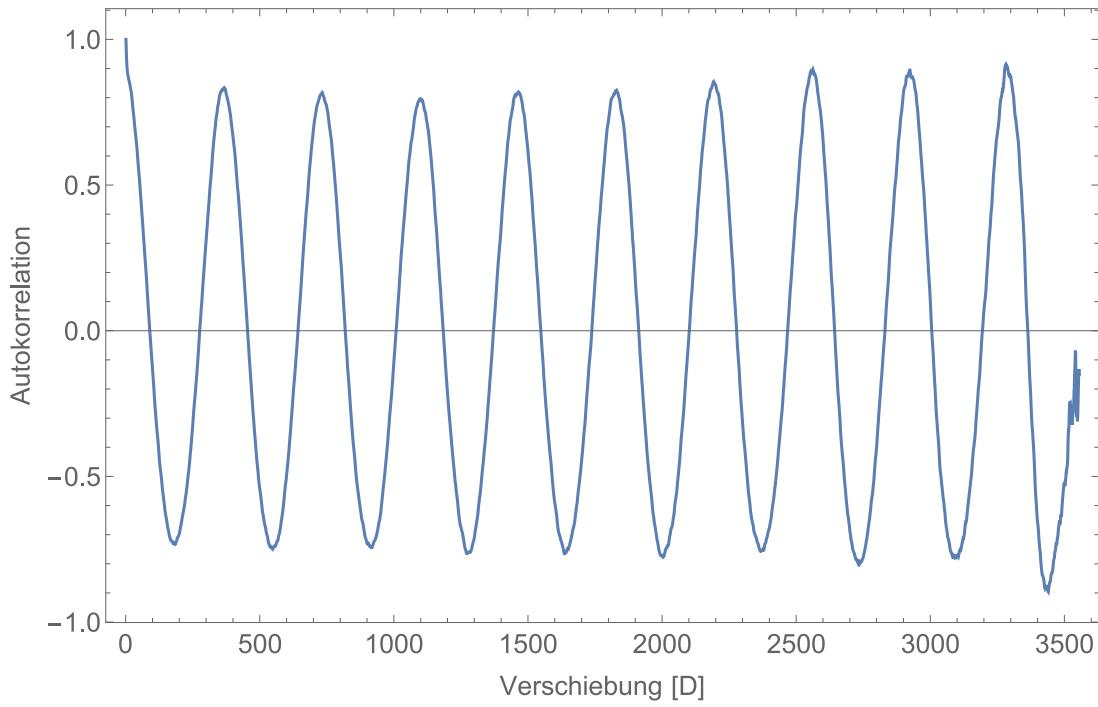


Abbildung 4.7: Autokorrelation der Temperatur in Alert (Kanada).

Aus der Temperaturzeitreihe lassen sich auch die Temperaturdifferenzen bestimmen. Diese sind exemplarisch für Alert (Kanada) in Abb. 4.8 dargestellt.

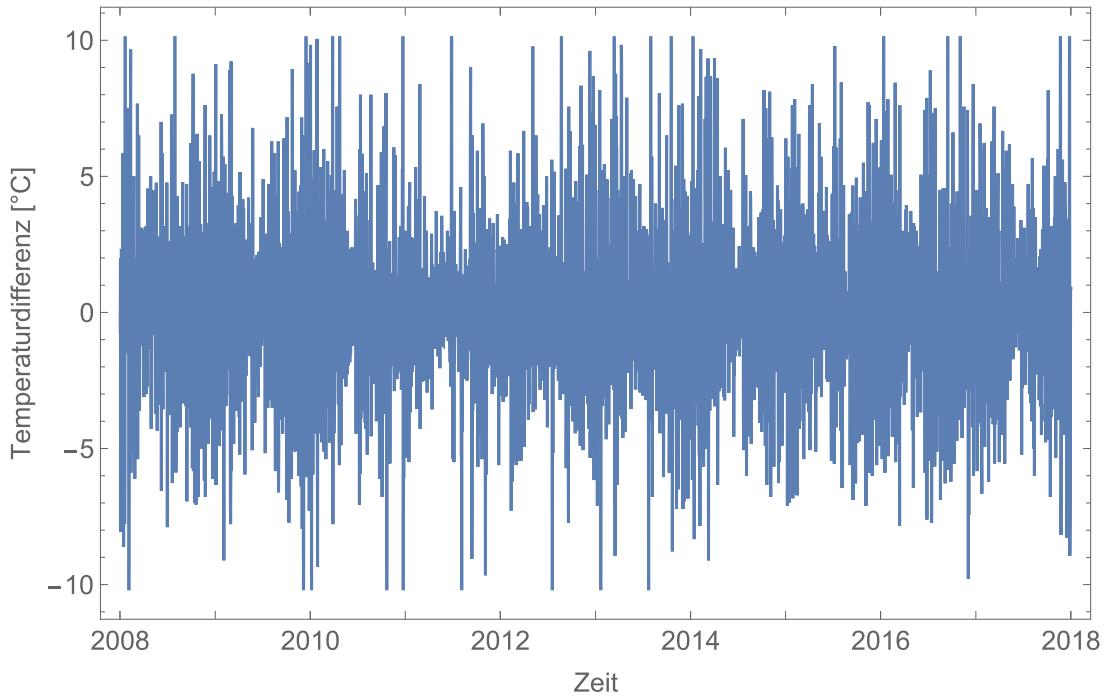


Abbildung 4.8: Temperaturdifferenzen in Alert (Kanada).

Zunächst werden wieder die statistischen Momente der Zeitreihen der Temperaturdifferenzen bestimmt. Diese sind in Tabelle 4.2 zusammengefasst.

Tabelle 4.2: Erwartungswert, Varianz, Schiefe und Kurtosis der Temperaturdifferenzen der ausgewählten Städte.

	Du	Fr	Ho	Vi	Jo	Pj	Pe	Ho	Al
m_1	0.0017	0.0023	-0.0028	-0.0003	0.0002	0.0003	-0.0007	0.0003	-0.0002
\bar{m}_2	2.3785	2.2484	4.3049	1.2675	2.4016	2.5412	2.3706	1.0088	3.1882
\tilde{m}_3	-0.2246	-0.0905	-0.4990	-0.2732	-0.4457	-0.5056	-0.2866	-0.1128	0.3624
\tilde{m}_4	3.8067	3.6633	4.6071	4.7675	4.1347	4.1830	3.8002	6.0369	5.3916

Die Temperaturdifferenzen wurden normiert und in Abb. 4.9 als Histogramm dargestellt. Es ist erkennbar, dass die Verteilung der Temperaturdifferenzen eine größere Kurtosis als eine Standardnormalverteilung besitzt, da sie spitzer ist.

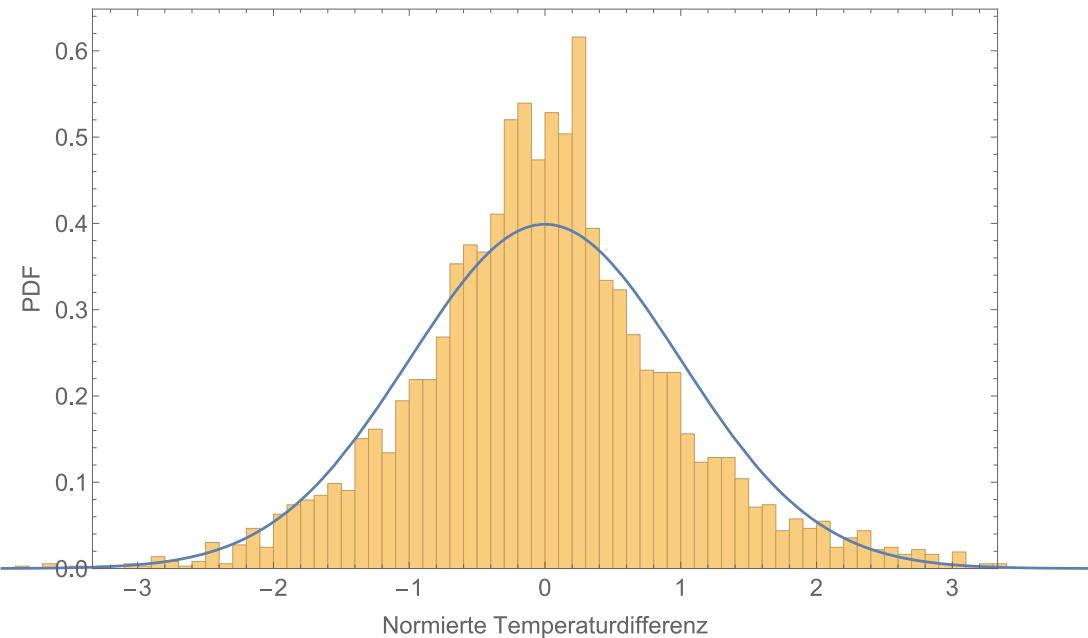


Abbildung 4.9: Histogramm der normierten Temperaturdifferenzen von Alert (Kanada) mit Standardnormalverteilung.

Im Streudiagramm in Abb. 4.10a sind die täglichen Temperaturänderungen von Duisburg und Frankfurt gegeneinander aufgetragen. Es lässt sich ein linearer Zusammenhang erkennen. Das Streudiagramm der Quantilsfunktionen der Temperaturdifferenzen von Duisburg und Frankfurt ist in Abb. 4.10b gegeben. Es lassen sich Häufungen in der rechten unteren und linken oberen Ecke, also ein linearer Trend erkennen.

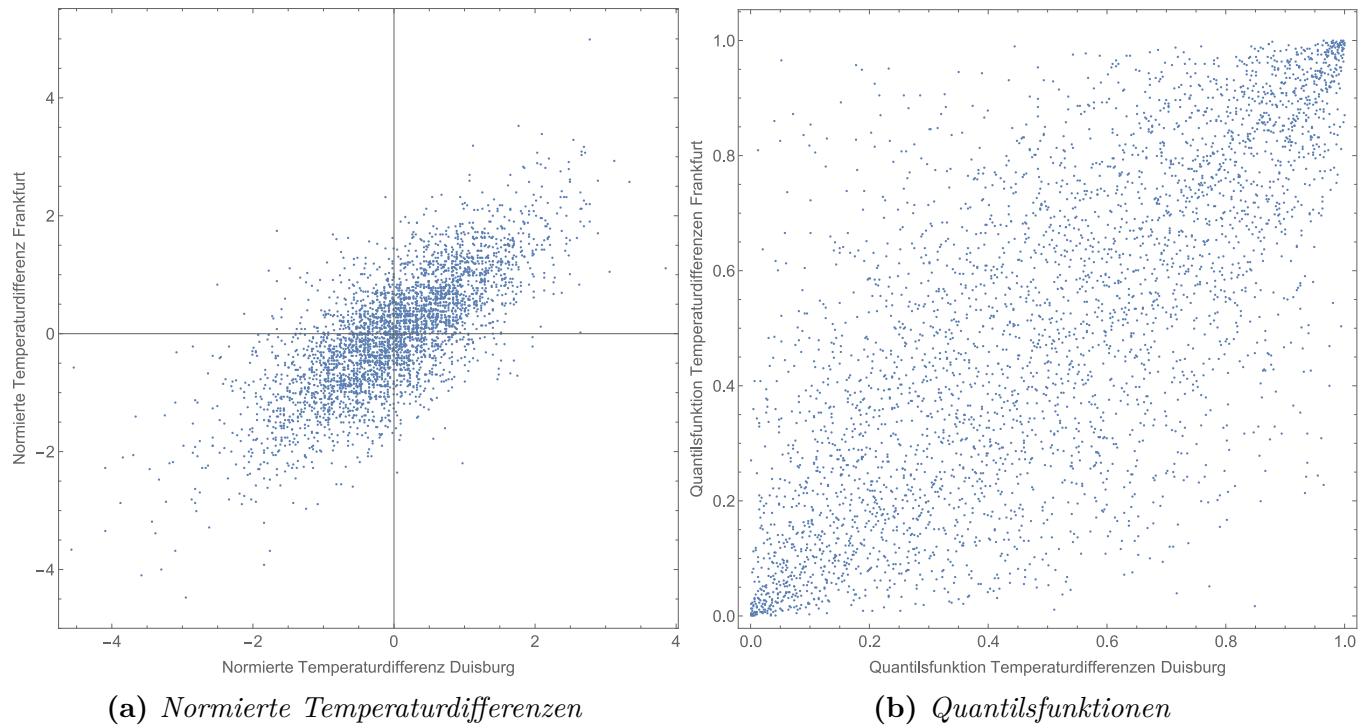


Abbildung 4.10: Streudiagramme der normierten Temperaturdifferenzen und der Quantilsfunktionen der Temperaturdifferenzen für Duisburg und Frankfurt.

In der Copuladichte in Abb. 4.11a ist wieder zu erkennen, dass die Temperaturänderungen in Duisburg und Frankfurt ähnlich verlaufen. Der Zusammenhang zwischen den Hemisphären ist jetzt jedoch nicht mehr gegeben, wie an der Copuladichte der Temperaturdifferenzen für Duisburg und Perth zu erkennen ist (Abb. 4.11b). Das bedeutet, dass die Temperaturänderungen in Duisburg unabhängig sind von den Temperaturänderungen in Perth (Australien).

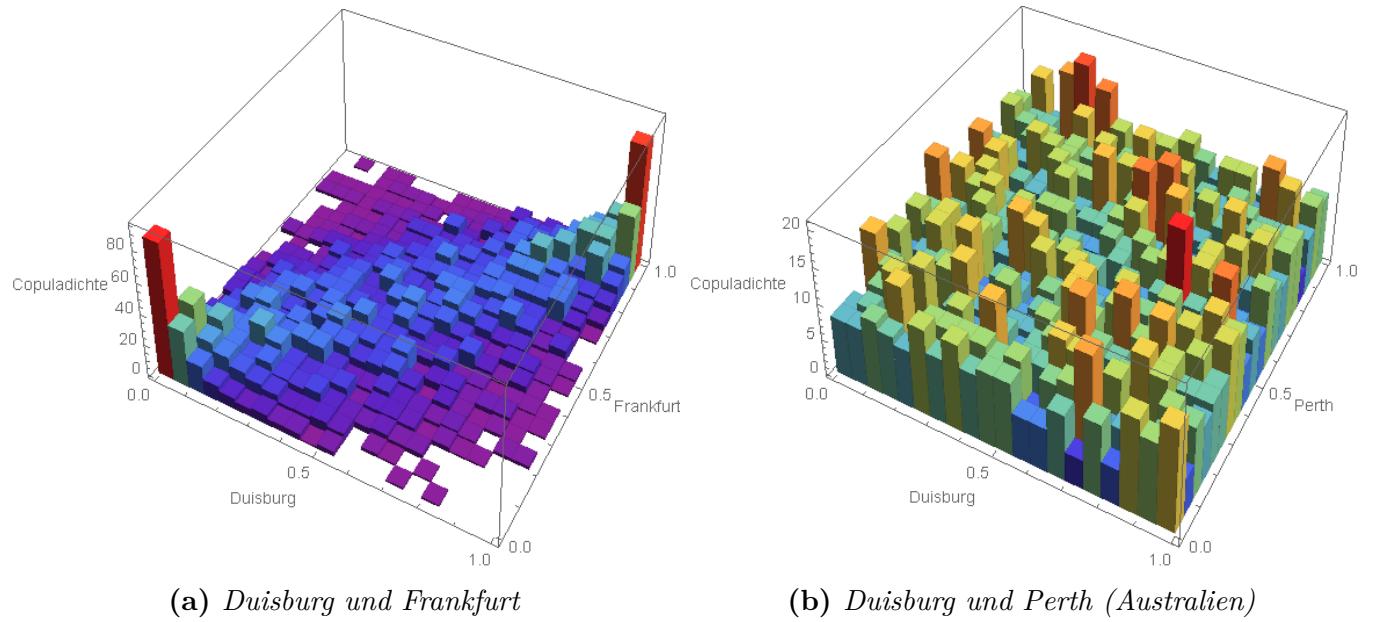


Abbildung 4.11: Copuladichte der Temperaturdifferenzen von Duisburg und Frankfurt und von Duisburg und Perth (Australien).

Die Autokorrelation der Temperaturdifferenzen in Abb. 4.12 schwankt um Werte nahe der Null. Temperaturänderungen sind also von einem auf den nächsten Tag unabhängig voneinander.

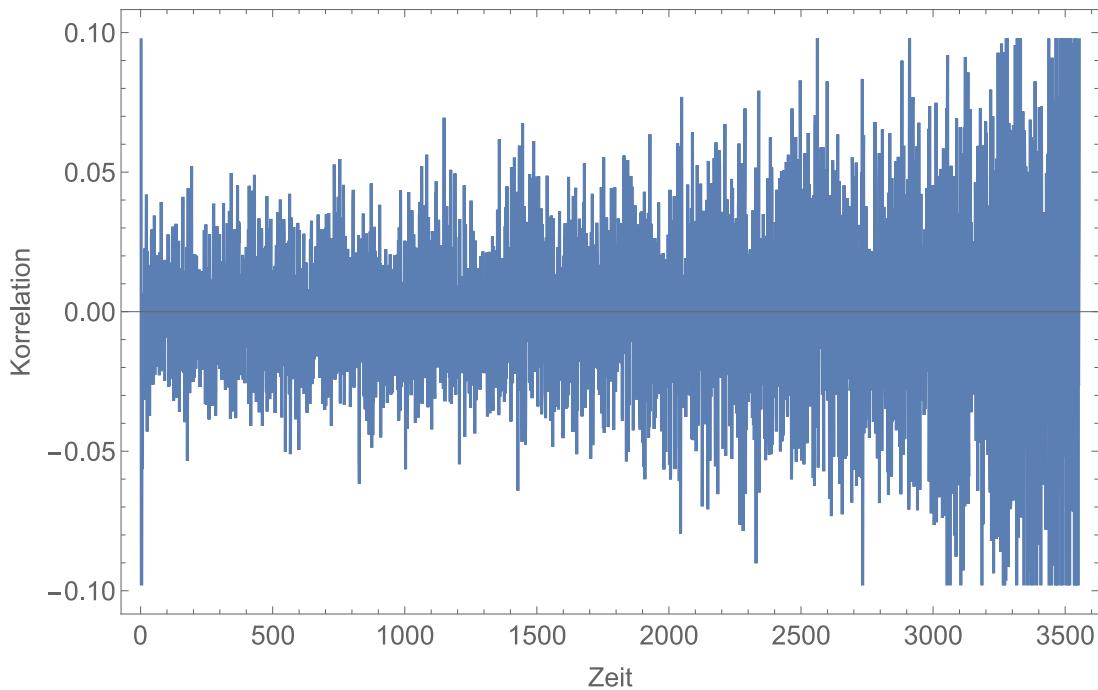


Abbildung 4.12: Autokorrelation der Temperaturdifferenzen in Alert.

Literaturverzeichnis

- [1] KRENGEL, ULRICH: *Einführung in die Wahrscheinlichkeitstheorie und Statistik: Für Studium, Berufspraxis und Lehramt*, Seite 160. vieweg studium; Aufbaukurs Mathematik. Vieweg+Teubner Verlag, 2015.
- [2] KRENGEL, ULRICH: *Einführung in die Wahrscheinlichkeitstheorie und Statistik: Für Studium, Berufspraxis und Lehramt*, Seite 53ff. vieweg studium; Aufbaukurs Mathematik. Vieweg+Teubner Verlag, 2015.
- [3] FRIEWALD, NILS: *Copula Funktionen- Eine Einführung*. <https://www.uni-muenster.de/Physik.TP/~lemm/seminarSS08/copula0506.pdf>, 2015. Zugriff: 18.10.18.