

Monitoring
Corinna Harmening

Übung 2

Sommersemester 23

Erstellt am: 18.05.2023

Ramon Butzer, Vera Kozel

Inhaltsverzeichnis

1	Aufgabe 1	2
2	Aufgabe 2	4
3	Aufgabe 3	7

1 Aufgabe 1

Im Bereich des ingenieurgeodätischen Monitorings werden Deformations- oder Bewegungsprozesse oftmals überwacht, indem eine *Zeitreihe* von Messwerten erfasst wird. Diese Übung beschäftigt sich mit der Zeitreihenanalyse im Zeitbereich. Dabei wird ein Beispieldatensatz eines Pumpspeicherkraftwerk in einem Tidegebiet verwendet. Dieser besteht aus Neigungsmessungen in x - und y -Richtung sowie der Temperatur, welche mit einer Abtastrate von 120 s aufgenommen wurden. Zunächst wird die Zeitreihe bereinigt, da diese Inkonsistenzen und Lücken in der Abtastung aufweist. Dazu werden alle Messwerte mit korrekter Zeitinformation, bei denen also die Zeit ein Vielfaches von $\Delta t = 120$ s beträgt, übernommen. Durch einen Vergleich zwischen dem Originaldatensatz und der übernommenen Messwerte, werden die Zeitpunkte festgestellt, für welche kein Messwert vorliegt, oder dieser zeitlich verschoben ist. Die Lücken in der Zeitreihe werden anschließend gefüllt, indem zwischen den beiden benachbarten Messwerten linear interpoliert wird. Für die Zeitpunkte, bei denen ein zeitlich verschobener Originalwert existiert, wird ebenfalls linear interpoliert, sodass abschließend eine Zeitreihe mit konsistenter Abtastrate von Δt vorliegt. Im nächsten Schritt wird eine Trendbereinigung der Zeitreihe durchgeführt. Dabei wird in Matlab die Funktion `polyfit` verwendet.

Eine Darstellung der Zeitreihe findet sich in Abbildungen 1 und 2. Darin sind die interpolierten Werte an den ursprünglichen Datenlücken und zeitlich verschobenen Punkten, sowie der Trend gekennzeichnet.

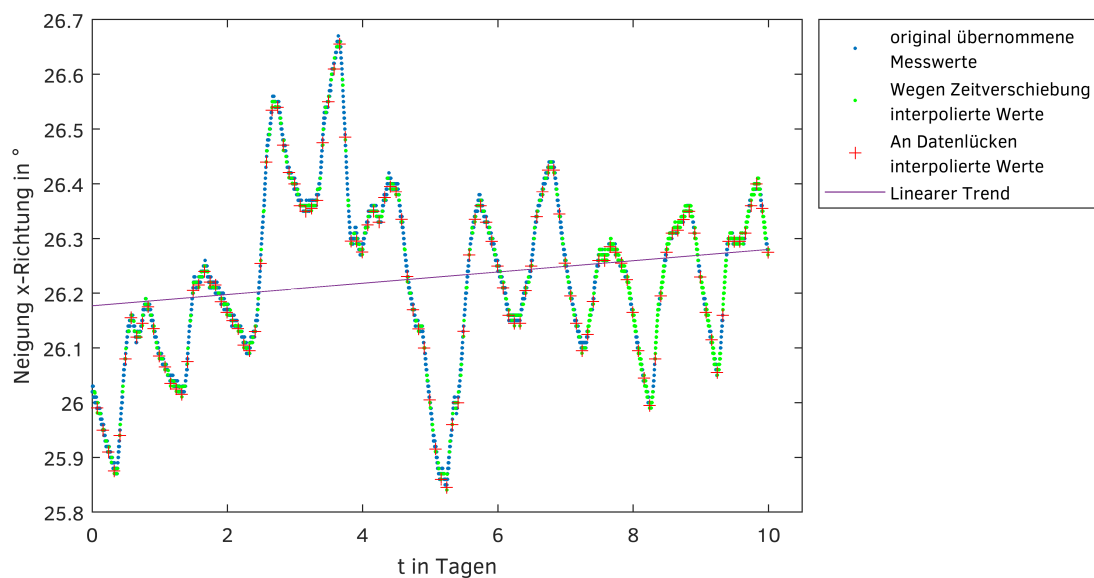


Abbildung 1: Temperaturverlauf der Zeitreihe

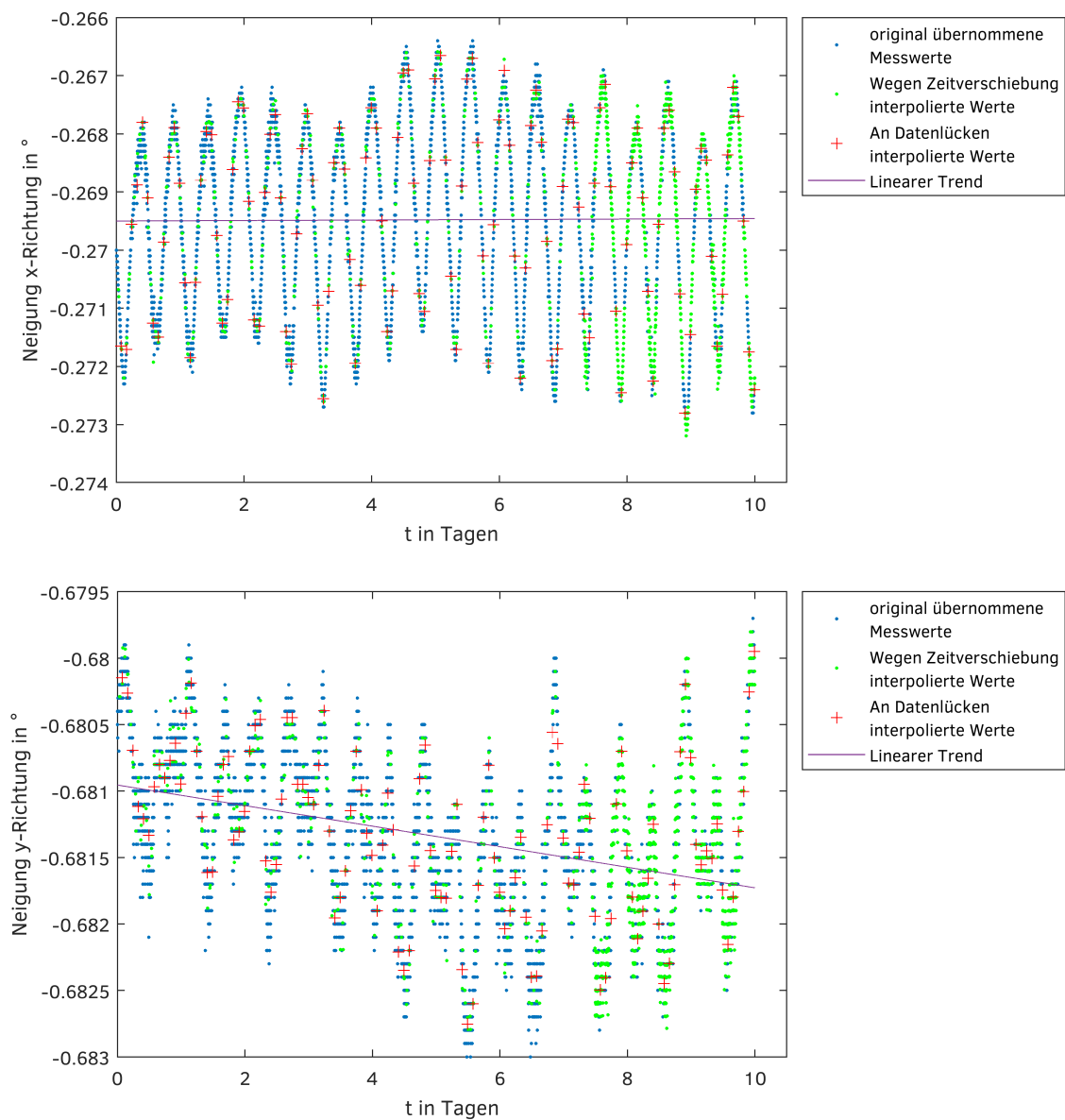


Abbildung 2: Neigungskomponenten der Zeitreihe

2 Aufgabe 2

Das Ziel dieser Aufgabe ist die Berechnung der Autokovarianz und der Autokorrelation, der stationären Zeitreihe mit n erfassten Messwerten. Die Autokovarianzfunktion beschreibt, wie sehr ein Messwert x_j mit einem anderen Messwert im Abstand k übereinstimmt und spiegelt damit den Zusammenhang zwischen zeitlich verschobenen Werte der Zeitreihe wieder. Dadurch können z.B. zyklische Signale bzw. deren Perioden in der Zeitreihe erkannt werden, da diese als lokale Maxima in der Autokovarianzfunktion hervortreten.

Die diskrete Autokovarianz wird über die Formel 1 geschätzt. Dabei stellt \bar{x} den Mittelwert der Zeitreihe dar. Die Laufvariable j geht vom ersten Wert der Zeitreihe, bis zu einem Wert, der üblicherweise ca. ein Zehntel der Länge der Zeitreihe beträgt. Grund dafür ist die statistische Sicherheit, da unter Einbezug längerer Perioden, also größer gewähltem maximalen k , immer weniger Messwerte der Zeitreihe verwendet werden können

$$\hat{C}(k) = \frac{1}{n-k-1} \sum_{j=1}^{n-k} (x_j - \bar{x})(x_{j+k} - \bar{x}), \quad k = 0, 1, \dots, m = n/10 \quad . \quad (1)$$

Der Wert $C(\tau = 0)$ beinhaltet die Kovarianz jedes Messwertes mit dem im Abstand 0 liegenden Wert, also dem selben Messwert. Das bedeutet, dass $C(0)$ die empirische Varianz widerspiegelt. Der Wert $C(\tau = 1)$ enthält die Kovarianzen der Werte mit ihren jeweiligen benachbarten Wert der Zeitreihe. Deren zeitlicher Abstand ist dabei von der Abtastrate Δt abhängig. Werte $C(\tau > 1)$ beinhalten die Kovarianzen zwischen Werten im zeitlichen Abstand τ . Ein hoher Wert der Autokovarianz für ein gewisses τ deutet darauf hin, dass die Kurven der Messwerte in diesem Abstand ähnlich verlaufen.

In Formel 2 wird ein Schätzwert für die Autokorrelation K berechnet. Sie geht aus einer Normierung der Autokovarianzfunktion mit der empirischen Varianz hervor und nimmt somit nur Werte im Intervall $[-1; 1]$ an, hat ansonsten aber den selben Verlauf wie die Autokovarianzfunktion. Es gilt

$$\hat{K}(k) = \frac{\hat{C}(k)}{\hat{C}(0)} \quad . \quad (2)$$

Die Ergebnisse der für die Beispielzeitreihe gebildeten Autokorrelationsfunktionen sind in den Abbildungen 3 bis 5 dargestellt. Dabei wird die x -Achse durch den zeitlichen Abstand τ ab Beginn der Zeitreihe gebildet, welcher sich durch Multiplikation des Index k der Messwerte mit der Abtastrate $\Delta t = 120$ s ergibt.

Alle Abbildungen weisen einen periodischen Verlauf mit dem erwarteten Maximum $K(\tau = 0)$ auf, da ein Wert mit sich selbst stets maximal korreliert ist. Im weiteren zeitlichen Verlauf finden sich Minima und Nebenmaxima, welche je nach Messgröße unterschiedlich ausgeprägt sind.

Die Autokorrelationsfunktion der Temperatur (vgl. Abbildung 3) nimmt nach dem Maximum kontinuierlich ab, bis ein Minimum bei ca. 13.5 h angenommen wird. Für $\tau = 6$ h beträgt der Korrelationskoeffizient noch ca. $K = 0.5$. Dieser Umstand weist darauf hin, dass sich die Temperatur im Tagesverlauf i.d.R. nicht schlagartig ändert, sondern Prozesse der Aufheizung und Abkühlung mit einer gewissen Trägheit verlaufen. Das Minimum nach ca. einem halben Tag und das Nebenmaximum bei ca. 24 h können durch den Tagesgang der Temperatur erklärt werden. Denn nach einem halben Tag wird die Temperaturdifferenz aufgrund des Unterschied im Sonnenstand (Tag \rightarrow Nacht) maximal, nach einem ganzen Tag besteht aus dem selben Grund ein relativ großer Zusammenhang. Trotzdem ist der Wert des Nebenmaximums kleiner als der Wert nach ca. 6 h. Dies kann anhand der Temperaturkurve in Abbildung 1 nachvollzogen werden. Denn die Tageshöchsttemperatur schwankt im Messzeitraum um bis zu 0.6°C , während der Temperaturunterschied an einem Tag i.d.R. kleiner als 0.2°C ist.

Der Verlauf der Autokorrelationsfunktion der Temperatur legt außerdem nahe, dass es sich bei

diesem Prozess um farbiges Rauschen handelt. Ein Gauß-Markov-Prozess würde sich als exponentieller Abfall der Autokorrelationsfunktion äußern. Durch Überlagerung des Gauß-Markov-Prozesses mit zyklischen Schwankungen ergibt sich das hier auftretende, dem farbigem Rauschen typische, Bild. Weißes Rauschen existiert in der Realität nicht. Bei weißem Rauschen gäbe es keine Korrelation zwischen den Werten.

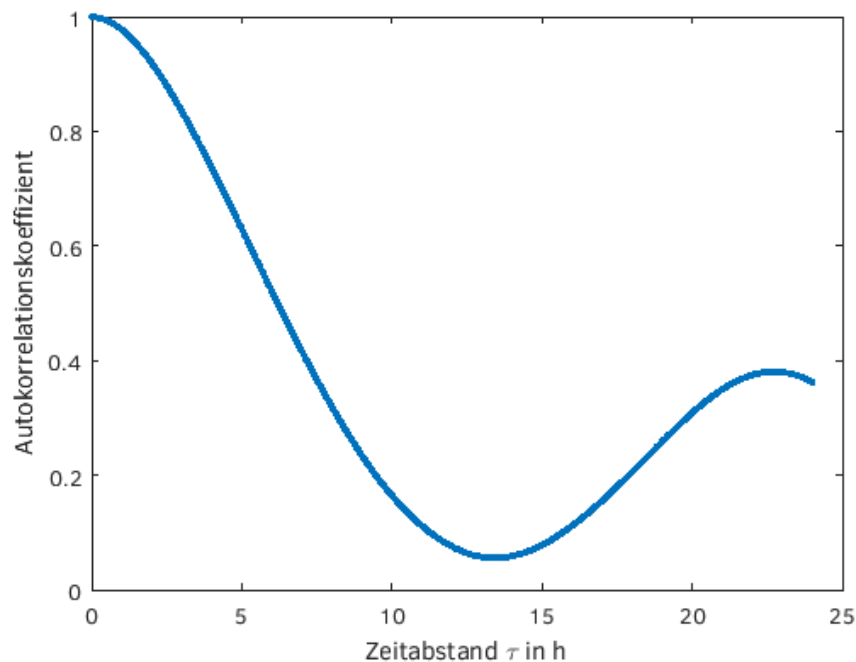


Abbildung 3: Autokorrelationsfunktion der Temperatur

Die Autokorrelationsfunktion der Neigung in x -Richtung (vgl. Abbildung 4) weist im Tagesverlauf ein Nebenmaximum bei ca. 12 h und eines bei ca. 24 h auf. Auffällig ist hierbei, dass bei diesen Perioden der Betrag des Korrelationskoeffizienten mit $K > 0.9$ deutlich größer ist, als beim Nebenmaximum der Temperaturkurve. Dieser Umstand bedeutet, dass die Neigung in x -Richtung nach einem halben bzw. ganzen Tag einen sehr ähnlichen Verlauf annimmt, was auch in Abbildung 2 anhand der gut erkennbaren periodischen Anteils verifiziert werden kann. Nach ca. 6 h und 18 h liegen Minima in der Autokorrelationsfunktion vor. Dieser Verlauf kann durch die Einwirkung der Gezeitenkräfte auf die Staumauer des Kraftwerks erklärt werden. Diese haben eine Halbtagesperiode ($T = 12$ h) und sind über kurze Zeiträume sehr stabil, weshalb der im Signal enthaltene Gauß-Markov-Prozess nur langsam abfällt, also eine große Erhaltensneigung aufweist. Der zyklische Anteil ist dagegen sehr prominent ausgeprägt.

Würde eine längere Zeitreihe betrachtet werden, wäre vermutlich der Einfluss von längeren Perioden, die zur Spring- und Nipptide führen, an einem größeren Gauß-Markov-Anteil erkennbar. Auffällig ist an der Autokorrelationsfunktion für die Neigung in x -Richtung noch das Vorzeichen. Während die Korrelation an den Maxima gegen $+1$ geht, geht sie an den Minima gegen -1 . Dieser Umstand wird durch die entgegengesetzte Steigung der Neigungskurve in 6 bzw. 18 h-Intervallen hervorgerufen. Anders gesagt ist in diesen Intervallen zwar der Neigungswert sehr ähnlich, aber die Kurve geht in die entgegengesetzte Richtung, ist also negativ korreliert.

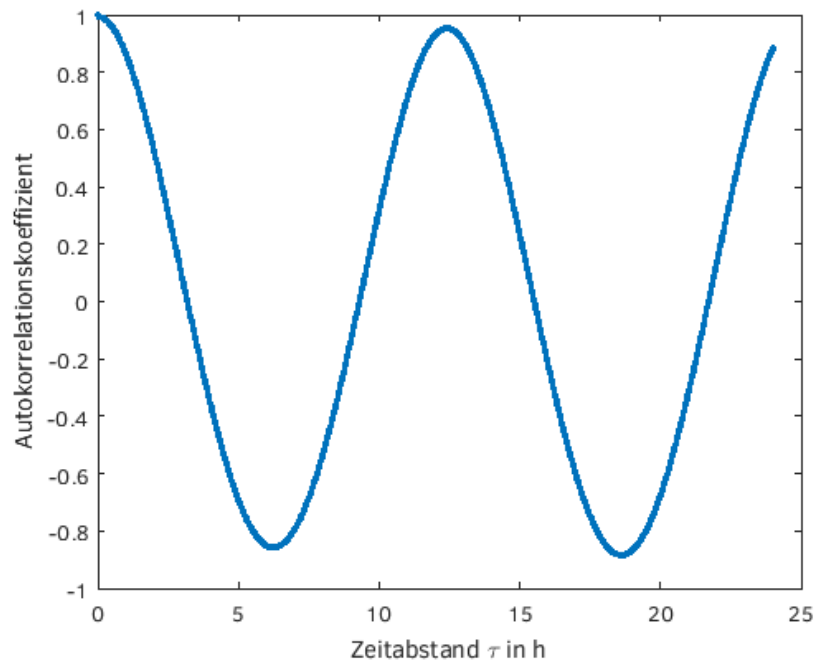


Abbildung 4: Autokorrelationsfunktion der Neigung in x-Richtung

Die Autokorrelationsfunktion der Neigung in y -Richtung (vgl. Abbildung 5) sieht ähnlich aus wie die der Neigung in x -Richtung. Allerdings ist das Nebenmaximum nach ca. 12 h nur halb so stark und das Nebenmaximum nach ca. 24 h um ca. 10 % weniger ausgeprägt wie das Nullmaximum. Dies weist darauf hin, dass der Zusammenhang zwischen Werten im Tagesgang signifikanter ausgeprägt ist, als der Zusammenhang im Halbtagesgang. Auch betragen die Werte der Minima bei 6 h bzw. 18 h nur ca. -0.2 , sind also nicht so stark negativ korreliert, wie bei der Neigung in x -Richtung.

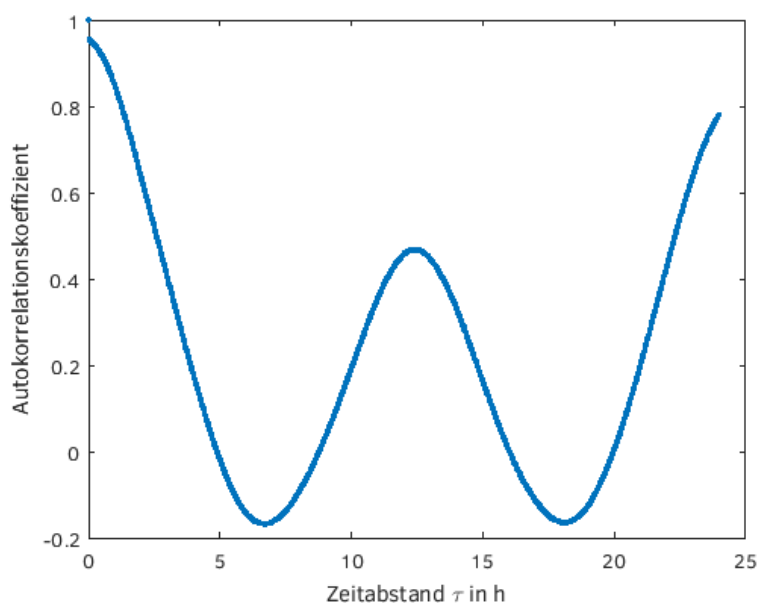


Abbildung 5: Autokorrelationsfunktion der Neigung in y-Richtung

3 Aufgabe 3

In diesem Aufgabenteil werden die Zusammenhänge zwischen Datenpunkten unterschiedlicher Zeitreihen untersucht. Dies kann mithilfe der Kreuzkovarianz geschehen, welche in Formel 3 berechnet wird

$$\hat{C}_{xy}(k) = \frac{1}{n-k-1} \sum_{j=1}^{n-k} (x_j - \bar{x})(y_{j+k} - \bar{y}), \quad k = 0, 1, \dots, m = n/10 \quad . \quad (3)$$

Dabei sind mit \bar{x} und \bar{y} die Mittelwerte der jeweiligen Messgrößen gegeben. Auch hierbei wird der Index k üblicherweise durch $\frac{1}{10}$ der Anzahl an Messpunkten begrenzt.

Durch Normierung der Kreuzkovarianz mit dem geometrischen Mittel der x - und y -Varianzen ergibt sich die Kreuzkorrelationsfunktion

$$\hat{K}_{xy}(k) = \frac{\hat{C}_{xy}(k)}{\sqrt{\hat{C}_x(0) \cdot \hat{C}_y(0)}} \quad . \quad (4)$$

Diese Normierung sorgt wie bei der Autokorrelationsfunktion dafür, dass der Wertebereich maximal im Intervall $[-1; 1]$ liegt. Allerdings muss das globale Maximum einer Kreuzkorrelationsfunktion im Gegensatz zur Autokorrelationsfunktion nicht bei $k = 0$ liegen, da der größtmögliche Zusammenhang zweier unterschiedlicher Kurven ggf. einem Zeitversatz unterliegt. Die Ergebnisse der für die Beispielzeitreihe gebildeten Kreuzkorrelationsfunktionen sind in Abbildungen 6 bis 8 dargestellt.

Die Kreuzkorrelationsfunktion der Neigungen in x - und y -Richtungen (vgl. Abbildung 6) weist im Zeitraum von 24 h zwei Maxima an den bereits bekannten Stellen 6 h und 18 h auf. In diesen zeitlichen Abständen besteht zwischen den Neigungskomponenten also der größte Zusammenhang. Minima bestehen ca. an den Stellen 0 h, 12 h und 24 h. Erklärt werden kann das Aussehen dieser Funktion dadurch, dass in den Messwerten beider Neigungskomponenten eine Halbtagesperiode prävalent ist, diese jedoch um 6 h verschoben ist. Ein Maximum in den Daten der x -Komponente fällt mit einem Minimum in den Daten der y -Komponente zusammen und alle 12 h wiederholt sich dieses Phänomen. Der Betrag der Kreuzkorrelationsfunktion liegt in diesem Fall zwischen -0.7 und 0.6 . Dieser Umstand zeigt, dass an den angegebenen Zeitabständen τ die Neigungskomponenten relativ stark positiv oder negativ korreliert sind.

Für den vorliegenden Sachverhalt ergibt sich aus der Form dieser Kreuzkorrelationsfunktion die Erkenntnis, dass die Staumauer des Pumpspeicherkraftwerks sich sowohl in x - als auch in y -Richtung im 12 h-Rhythmus neigt, der Abstand zwischen dem maximalen Betrag der Neigungskomponenten jedoch 6 h beträgt.

Die Kreuzkorrelationsfunktionen zwischen den Neigungskomponenten und der Temperatur wird stets so berechnet, dass die Temperatur der Funktion als erster Parameter und die Neigungskomponente als zweiter Parameter übergeben wird. Grund dafür ist, dass die Temperatur die potentielle Einflussgröße ist, welche auf die Neigungskomponenten einwirkt. Dies kann jedoch nur in einem zeitlichen Versatz geschehen. Kurz gesagt werden im folgenden die Zusammenhänge zwischen der Temperatur eines Zeitpunktes und den Neigungskomponenten in einem gewissen zeitlichen Abstand τ zum Zeitpunkt betrachtet.

Die Kreuzkorrelationsfunktion zwischen der Temperatur und der Neigung in x -Richtung ist in Abbildung 7 dargestellt. Darin ist zunächst erkennbar, dass der Betrag des Kreuzkorrelationskoeffizienten im Intervall $[-0, 1; 0, 16]$ liegt und somit deutlich geringere Werte annimmt, wie bei den bisher betrachteten Funktionen. Auch in dieser Funktion lässt sich ein zyklisches Signal erkennen. Betrachtet man die Maxima und Minima der Funktion, so lässt sich die Periode des zyklischen Signals auf ca. 10 h bis 13 h grob abschätzen. Das globale Minimum tritt bei ca.

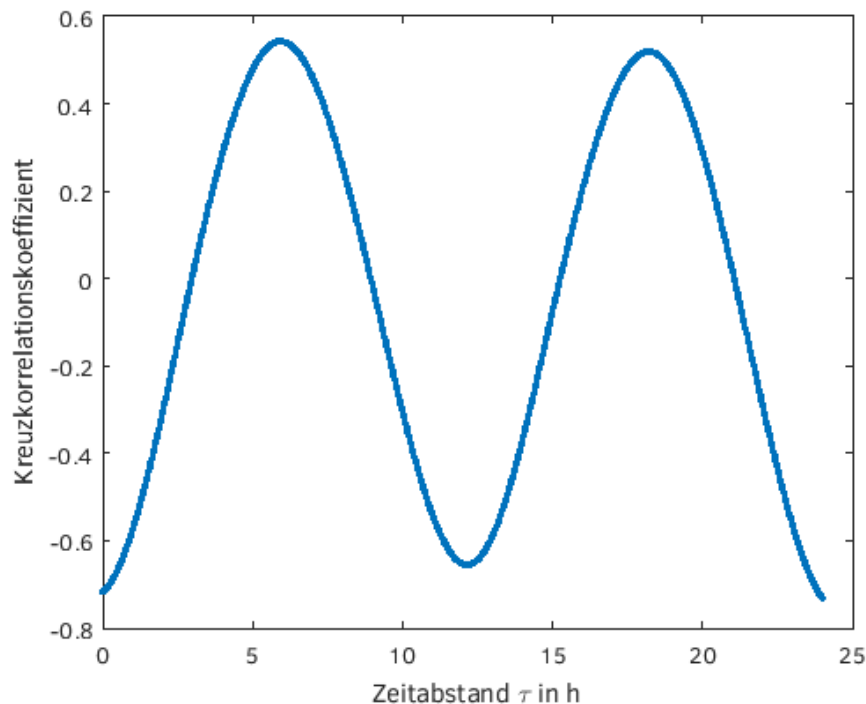


Abbildung 6: Kreuzkorrelation zwischen der x-Neigung und y-Neigung

$\tau = 2.5$ h auf, das globale Maximum bei ca. $\tau = 22$ h. Der Verlauf der Funktion ist schwieriger zu erklären. Augenscheinlich liegt ein gewisser periodischer Zusammenhang zwischen der Temperatur (24 h-Periode) und der Neigung in x -Richtung (12 h-Periode) vor, welcher im Zeitabstand $\tau = 22$ h sein größtes Ausmaß annimmt.

Unter der Annahme, dass die Gezeiten den größten Einfluss auf die Neigungskomponente in x -Richtung ausüben, kann die Vermutung angestellt werden, dass der Hochpunkt der Korrelationsfunktion im Bild zwar durch die Tagesperiode der Temperatur hervorgerufen wird, die eigentliche *Ursache* dafür aber das zweifache Durchlaufen der Halbtagesperiode der Gezeiten ist. In diesem Fall wäre dieser Sachverhalt ein schönes Beispiel für eine *Scheinkorrelation* bzw. einen *Fehlschluss* handeln. Ein hoher Wert in der Kreuzkorrelationsfunktion muss nicht bedeuten, dass eine der untersuchten Größen von der anderen abhängt.

Die Kreuzkorrelationsfunktion zwischen der Temperatur und der Neigung in y -Richtung ist in Abbildung 8 dargestellt. Auch hier tritt ein klares zyklisches Signal auf. Die Beträge des Kreuzkorrelationskoeffizienten liegen im Intervall $[-0.4, 0.2]$. Die maximale Korrelation tritt bei ca. $\tau = 6$ h auf. Dies könnte ein Hinweis darauf sein, dass die Temperatur als Einflussgröße im Sinne einer thermischen Ausdehnung auf die y -Neigungskomponente des Kraftwerkes einwirkt. Der Zeitabstand zwischen Höchsttemperatur und Maximum der Neigungskomponente könnte in diesem Fall durch die Trägheit des Materials gegenüber Ausdehnung hervorgerufen werden.

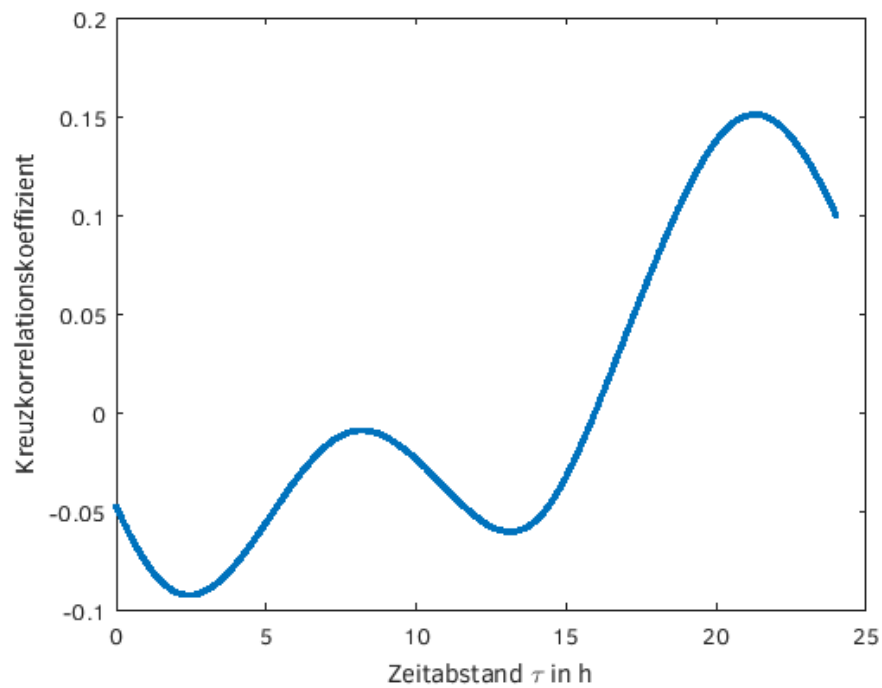


Abbildung 7: Kreuzkorrelation zwischen der x-Neigung und Temperatur

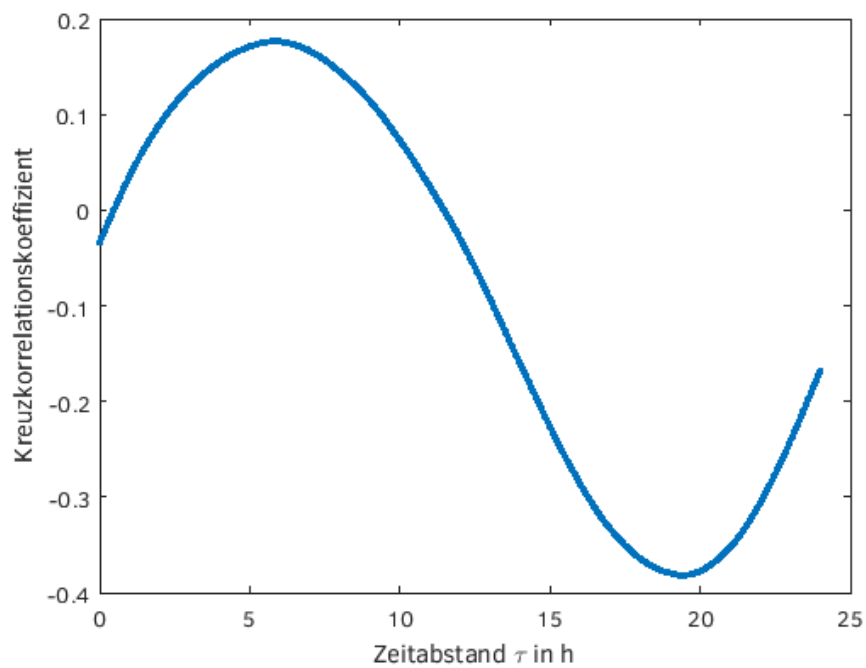


Abbildung 8: Kreuzkorrelation zwischen der y-Neigung und Temperatur