

Korrektur dynamischer Fehler für direkte Spektralschätzung von ungleichmäßig abgetasteten Daten inklusive korrelierter Abtastintervalle

Bias Correction for Direct Spectral Estimation from Irregularly Sampled Data Including Sampling Schemes with Correlation

H. Nobach¹, N. Damaschke² und V. Kühn³

¹ Max-Planck-Institut für Dynamik und Selbstorganisation, Am Faßberg 17, 37077 Göttingen

² Institut für Allgemeine Elektrotechnik, Universität Rostock, Fakultät für Informatik und Elektrotechnik, Albert-Einstein-Straße 2, 18059 Rostock

³ Institut für Nachrichtentechnik, Universität Rostock, Fakultät für Informatik und Elektrotechnik, Richard-Wagner-Str. 31, 18119 Rostock-Warnemünde

ungleichmäßige Abtastung, direkte Spektralschätzung, Korrelation, Leistungsdichte
non-regular sampling, direct spectral estimation, autocorrelation, spectral density

Zusammenfassung

Die Prädiktion und Korrektur systematischer Fehler bei direkter Spektralschätzung aus ungleichmäßig abgetasteten Daten wird untersucht. Es werden verschiedene Abtastprozesse untersucht, die zu einer ungleichmäßigen Abtastung des untersuchten Prozesses führen. Es werden sowohl zeitkontinuierliche Abtastprozesse betrachtet, bei denen einzelne Abtastzeitpunkte mit ungleichmäßigen Abtastintervallen und einer stetigen Verteilungsfunktion entstehen als auch nominell äquidistante Abtastfunktionen mit Fehlstellen und einer entsprechend diskreten zeitlichen Verteilung. Für beide zeitlichen Verteilungen der Abtastintervalle werden jeweils verschiedene Abtastmuster untersucht. Es werden einerseits rein zufällige und voneinander unabhängige Abtastzeitpunkte untersucht, andererseits werden miteinander korrelierte und somit voneinander abhängige Abtastschemen betrachtet.

1 Einführung

Die ungleichmäßige Abtastung eines stochastischen Zufallsprozesses bewirkt, dass sich das aus den Daten gewonnene Leistungsdichtespektrum von dem charakteristischen Spektrum des betrachteten Prozesses unterscheidet. Das ungleichmäßig abgetastete, aufgenommene Signal entsteht durch die Multiplikation der Abtastfunktion mit dem ursprünglichen Signal des betrachteten Prozesses. Somit entspricht das Spektrum des aufgenommenen Signals der Faltung des Fourier-Spektrums des Prozesses mit dem Fourier-Spektrum der zeitlichen Abtastfunktion.

Die spektralen Eigenschaften der Abtastfunktion haben somit unmittelbaren Einfluss auf das Spektrum des aufgenommenen Signals. Eine möglichst korrekte Bestimmung des Spektrums des aufgenommenen Signals ergibt somit noch keine korrekte Informationsgewinnung über den betrachteten Prozess.

Aus den Untersuchungen zu Verarbeitungsmethoden für Laser-Doppler-Daten liegen bereits sehr detaillierte Kenntnisse über eine konsistente und effektive Bestimmung von Spektren aus unregelmäßig abgetasteten Daten vor [3]. Die direkte Schätzung der Spektren aus den Datenwerten ist dabei von systematischen Fehlern gekennzeichnet, die einerseits durch die zufällige Abtastung selbst entstehen [4] als auch durch das Entfernen von überlappenden Signalen aufeinanderfolgender Teilchen aus dem Datenstrom, das zunächst auf komplexe Weise einen Mindestabstand von Einzelsignalen (als Prozessortzeit bekannt) und schließlich eine Korrelation der Abtastintervalle erzeugt [2].

Übertragbare Zusammenhänge ergeben sich auch für nominell äquidistante Daten mit beobachteten Fehlstellen. Auch hier kann es ein rein zufälliges Auftreten von Ausreißern geben, die für die statistische Auswertung der Daten nicht mehr zur Verfügung stehen, oder es kann zu größeren Datenlücken kommen, bei denen die Abtastzeitpunkte korreliert sind und somit nicht als unabhängig voneinander betrachtete werden können.

Abgeleitet aus bekannten Verfahren zur erwartungstreuen Schätzung von Leitungsdichtespektren aus Laser-Doppler-Daten wird ein gemeinsamer Lösungsansatz für zeitkontinuierliche und zeitdiskrete Abtastfunktionen mit beliebigem spektralen Charakter entwickelt. Das Verfahren bewirkt eine Korrektur unterschiedlicher Wahrscheinlichkeiten für das Auftreten von Selbstprodukten einzelner Messwerte und Kreuzprodukten verschiedener Messwerte. Bei regelmäßiger, äquidistanter Abtastung (ohne Lücken) sind die Wahrscheinlichkeiten für das Auftreten von Selbst- und Kreuzprodukten identisch. Bei ungleichmäßiger Abtastung (fehlende Datenwerte in einem äquidistanten Datenstrom oder nicht äquidistante Abtastung) sind sie dagegen unterschiedlich und abhängig von den spektralen Eigenschaften der Abtastfunktion. Die statistischen Eigenschaften des Abtastprozesses können für die Korrektur sowohl aus theoretischen Betrachtungen des Abtastschemas gewonnen werden oder empirisch aus einer konkreten Realisierung der Abtastfunktion eines bereits aufgenommenen Datensatzes. Anhand von Simulationsrechnungen werden die Erwartungstreue und die Konsistenz dieser Lösungsansätze nachgewiesen.

Alle verwendeten Programme sind als Python-Code unter [1] zu finden zusammen mit einigen Herleitungen in einem Anhang.

2 Primäre direkte Spektralschätzung

Es wird ein mittelwertfreier, stochastischer Prozess $x(t)$ betrachtet. Der Prozess wird an N aufsteigenden Stellen t_i mit $i = 1 \dots N$ im Zeitraum T abgetastet. Die Abtastung erfolgt entweder an zufälligen Stellen t_i des kontinuierlichen Prozesses oder zu quantisierten Zeitpunkten mit einem festen Abtastintervall Δt , jedoch mit fehlenden Datenpunkten. Von dem aufgenommenen Datensatz wird ein erwartungstreu Spektrum bestimmt. Der betrachtete Datensatz besteht im Fall der stetigen Abtastverteilung nur aus den unregelmäßig aufgenommenen Datenpunkten, im Fall der quantisierten Abtastverteilung mit fehlenden Werten nur aus den gültigen unter den aufgenommenen Werten. Unter Vernachlässigung möglicher Fehler durch eine angenommene periodische Fortsetzung des Signals, kann das Leistungsdichtespektrum z. B. mit der Fourier-Transformation

$$S(f) = \frac{T}{N^2} \left| \sum_{i=1}^N x_i e^{-2\pi i f t_i} \right|^2 \quad (1)$$

mit der imaginären Einheit i erwartungstreu, und zwar direkt aus den Daten bestimmt werden. Alternativ kommt hier die Lomb-Scargle-Methode [6, 7] zur Anwendung. Die generalisierte Lomb-Scargle-Methode [8] wird nicht verwendet, da hier nur mittelwertfreie Prozesse untersucht werden.

Das so bestimmte Spektrum ist zwar eine erwartungstreu Schätzung der Leistungsdichte des Signals. Diese weicht aber von der Leistungsdichte des betrachteten Prozesses ab, weil die ungleichmäßige Abtastung des ursprünglich kontinuierlichen Prozesses die Leistungsdichte verfälscht. Die Multiplikation des kontinuierlichen Signals des Prozesses mit der Abtastfunktion bewirkt im Frequenzbereich eine Faltung der beiden Spektren.

Die Leistungsdichte des Signals $S(f)$ und die entsprechende Korrelationsfunktion $R(\tau)$ korrespondieren über das Wiener-Chintschin-Theorem [5]. Eine rechentechnische Umsetzung mittels der diskreten Fourier-Transformation setzt die regelmäßige Abtastung beider Funktionen in Zeitschritten $\Delta\tau$ bzw. Frequenzschritten Δf voraus. Bei einer stetigen Abtastverteilung ist der fundamentale Zeitschritt $\Delta\tau$ frei wählbar. Bei nominell äquidistanter Abtastung sind die Zeitschritte der Korrelationsfunktion identisch mit dem Zeitschritt Δt des Signals. Der Frequenzschritt ist in beiden Fällen $\Delta f = 1/T$ wobei die Messzeit T ein natürliches Vielfaches des Zeitschrittes $\Delta\tau$ ist mit $T = M\Delta\tau$. Die Korrespondenz der beiden Funktionen ist durch

$$S(f_j) = \Delta\tau \cdot \text{DFT} \{R(\tau_k)\} = \Delta\tau \cdot \sum_{k=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} R(\tau_k) e^{-2\pi i f_j \tau_k} \quad (2)$$

$$R(\tau_k) = \frac{1}{\Delta\tau} \cdot \text{IDFT} \{S(f_j)\} = \frac{1}{T} \cdot \sum_{j=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} S(f_j) e^{2\pi i f_j \tau_k} \quad (3)$$

gegeben mit $f_j = j\Delta f$, $j = -\lfloor M/2 \rfloor \dots \lfloor (M-1)/2 \rfloor$ und $\tau_k = k\Delta\tau$, $k = -\lfloor M/2 \rfloor \dots \lfloor (M-1)/2 \rfloor$, wobei $\lfloor x \rfloor$ die größte ganze Zahl kleiner oder gleich x ist.

3 Korrektur systematischer Fehler

Die hier betrachteten Abtastschemen ergeben entweder eine stetige Verteilung der Abtastintervalle (für rein zufällige Abtastung eine Exponentialverteilung, mit Korrelationen davon abweichend) oder eine diskrete Verteilung der Abtastintervalle (für eine äquidistante Abtastung ohne Fehlstellen mit der Wahrscheinlichkeit eins für das Abtastintervall $\Delta\tau$, bei auftretenden Fehlstellen kommen auch weitere natürliche Vielfache von $\Delta\tau$ vor). Für eine eindeutige Unterscheidung dieser beiden Abtastprozesse werden im Folgenden die **stetige** und die **diskrete** Verteilung der Abtastintervalle unterschieden, unabhängig davon wie zufällig bzw. unabhängig die Abtastpunkte voneinander sind.

Die mittlere Abtastrate sei α und die mittlere Anzahl von Abtastungen pro Zeitintervall $\Delta\tau$ sei entsprechend $\alpha' = \alpha\Delta\tau$. Bei diskreter Verteilung der Abtastintervalle ist α' gleich der Wahrscheinlichkeit für einen gültigen Datenwert und damit kleiner oder gleich eins. Bei stetiger Verteilung der Abtastintervalle (ohne Korrelation) ist die Anzahl der Abtastungen in einem Intervall $\Delta\tau$ Poisson-verteilt mit dem Mittelwert α' , wobei α' auch größer als eins sein kann. Wenn alle Abtastungen unabhängig voneinander erfolgen, dann ist die mittlere Anzahl von Kreuzprodukten $x_i x_j$ aus zwei verschiedenen Zeitschritten α'^2 . Die mittlere Anzahl von Selbstprodukten x_i^2 in einem Zeitschritt ist α' .

Bei der Bestimmung der Korrelationsfunktion beinhaltet $R(0)$ alle Selbst- und für die stetige Verteilung der Abtastintervalle, bei der auch mehr als eine Abtastung innerhalb eines Zeitschrittes erfolgen kann, auch alle Kreuzprodukte von Abtastungen, die in jeweils einen der M Zeitschritte fallen. Alle $R(\tau_k \neq 0)$ beinhalten dagegen nur Kreuzprodukte der Abtastungen, die in jeweils zwei Zeitschritte mit dem Abstand $\tau_k = k\Delta\tau$ fallen. Es sei β_k die mittlere Anzahl von Produkten in $R(\tau_k)$ aus einem Zeitschritt. Eine stationäre Abtastung führt dann zu einer Erwartung von $M\beta_k$ Selbst- bzw. Kreuzprodukten in $R(\tau_k)$ aus dem gesamten Datensatz. Abschließend sei $\beta'_k = \beta_k/\alpha'^2$. Dann gilt (bei $\tau_k \neq 0$) für unabhängige Abtastungen $\beta'_k = 1$. Jegliche Abweichung von der unabhängigen Abtastung führt dagegen zu $\beta'_k \neq 1$. Für unabhängige Abtastung mit stetiger Verteilung der Abtastintervalle erhält man schließlich

$$\beta'_k = \begin{cases} 1 + \frac{1}{\alpha'} & \text{für } k = 0 \\ 1 & \text{sonst} \end{cases} \quad (4)$$

und für unabhängige Fehlstellen mit diskreter Verteilung der Abtastintervalle erhält man

$$\beta'_k = \begin{cases} \frac{1}{\alpha'} & \text{für } k = 0 \\ 1 & \text{sonst.} \end{cases} \quad (5)$$

Herleitungen dieser und der folgenden Zusammenhänge sind unter [1] in einem Anhang zu diesem Beitrag zu finden. Treten dagegen Korrelationen auf, dann erhält man für die stetige Verteilung der Abtastintervalle

$$\beta'_k = 1 + \frac{\gamma_k}{\alpha'}, \quad (6)$$

und für korrelierte Datenlücken bei diskreter Verteilung der Abtastintervalle

$$\beta'_k = 1 - \gamma_k + \frac{\gamma_k}{\alpha'}, \quad (7)$$

wobei γ_k die Korrelation der Abtastfunktion für den zeitlichen Abstand τ_k angibt. Für korrelierte Datenlücken mit einer diskreten Verteilung der Abtastintervalle ist γ_k die Korrelationskoeffizientenfunktion der gültigen Datenwerte. Für die stetige Verteilung der Abtastintervalle mit Korrelation ist γ_k die Korrelationsfunktion der Abtastfunktion, vermindert um α'^2 und normiert auf α' . In beiden Fällen erhält man bei unabhängiger Abtastung $\gamma_0 = 1$ und $\gamma_k = 0$ für alle $\tau_k \neq 0$. Für korrelierte Datenlücken bei einer diskreten Verteilung der Abtastintervalle bleibt $\gamma_0 = 1$ und für $\tau_k \neq 0$ wird $\gamma_k > 0$, abhängig von der bestehenden Korrelation zwischen den Abtastungen. Für die stetige Verteilung der Abtastintervalle mit Prozessortzeit wird dagegen $\gamma_0 < 1$, weil unmittelbar nach einer

erfolgten Abtastung die Wahrscheinlichkeitsdichte für eine weitere Abtastung geringer ist als für weiter voneinander entfernten Zeitpunkte.

Für unkorrelierte Abtastungen sind alle $\beta'_k = 1$ außer β'_0 . Damit hat die Korrelationsfunktion nur an der Stelle null einen systematischen Fehler, der sich durch den Vorfaktor $1/\beta'_0$, nämlich durch

$$\hat{R}(0) = \frac{1}{\beta'_0} R(0)$$

bzw. durch Subtraktion von

$$\left(1 - \frac{1}{\beta'_0}\right) R(0)$$

korrigieren lässt. Ein systematischer Fehler der Korrelationsfunktion nur an der Stelle null entspricht einer konstanten Verschiebung des gesamten Leistungsdichtespektrums. Für das Leistungsdichtespektrum $S(f)$ gilt

$$R(0) = \int_{-F/2}^{F/2} S(f) df$$

mit $F = 1/\Delta\tau$. Demnach ist das Leistungsdichtespektrum für eine Korrektur des systematischen Fehlers um

$$\frac{1}{F} \left(1 - \frac{1}{\beta'_0}\right) R(0) = \Delta\tau \left(1 - \frac{1}{\beta'_0}\right) R(0)$$

zu verschieben.

Um bei der direkten Korrektur des Spektrums die Bestimmung von Werten der Korrelationsfunktion zu vermeiden, wird stattdessen die empirische Varianz des (mittelwertfreien) Signals

$$s^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 \quad (8)$$

verwendet, wobei $\hat{R}(0) = R(0)/\beta'_0 \approx s^2$ gilt. Die Abweichung ergibt sich dadurch, dass in $R(0)$ bzw. $\hat{R}(0)$ zusätzlich zu den Selbstprodukten auch noch Kreuzprodukte von Abtastungen innerhalb einzelner Abtastintervalle eingehen, während in s^2 nur die Selbstprodukte gezählt werden. Es folgt $R(0) \approx \beta'_0 s^2$ und die Korrektur des Spektrums wird zu

$$\hat{S}(f) = S(f) - \Delta\tau (\beta'_0 - 1) s^2. \quad (9)$$

Für rein zufällige Abtastung mit stetiger Verteilung der Abtastintervalle reduziert sich die Korrektur zu

$$\hat{S}(f) = S(f) - \frac{s^2}{\alpha} \quad (10)$$

und für unabhängige Fehlstellen in einer diskreten Verteilung der Abtastintervalle zu

$$\hat{S}(f) = S(f) - s^2 \left(\frac{1}{\alpha} - \Delta\tau\right). \quad (11)$$

Bestehen dagegen Korrelationen zwischen den Abtastungen, dann müssen weitere β'_k in der Korrektur berücksichtigt werden. Die direkte Korrektur des Spektrums wird dadurch kompliziert. In diesem Fall empfiehlt sich die Korrektur der primär geschätzten Korrelationsfunktion durch

$$\hat{R}(\tau_k) = \frac{1}{\beta'_k} R(\tau_k), \quad (12)$$

mit anschließender Fourier-Transformation, um das korrigierte Leistungsdichtespektrum zu erhalten.

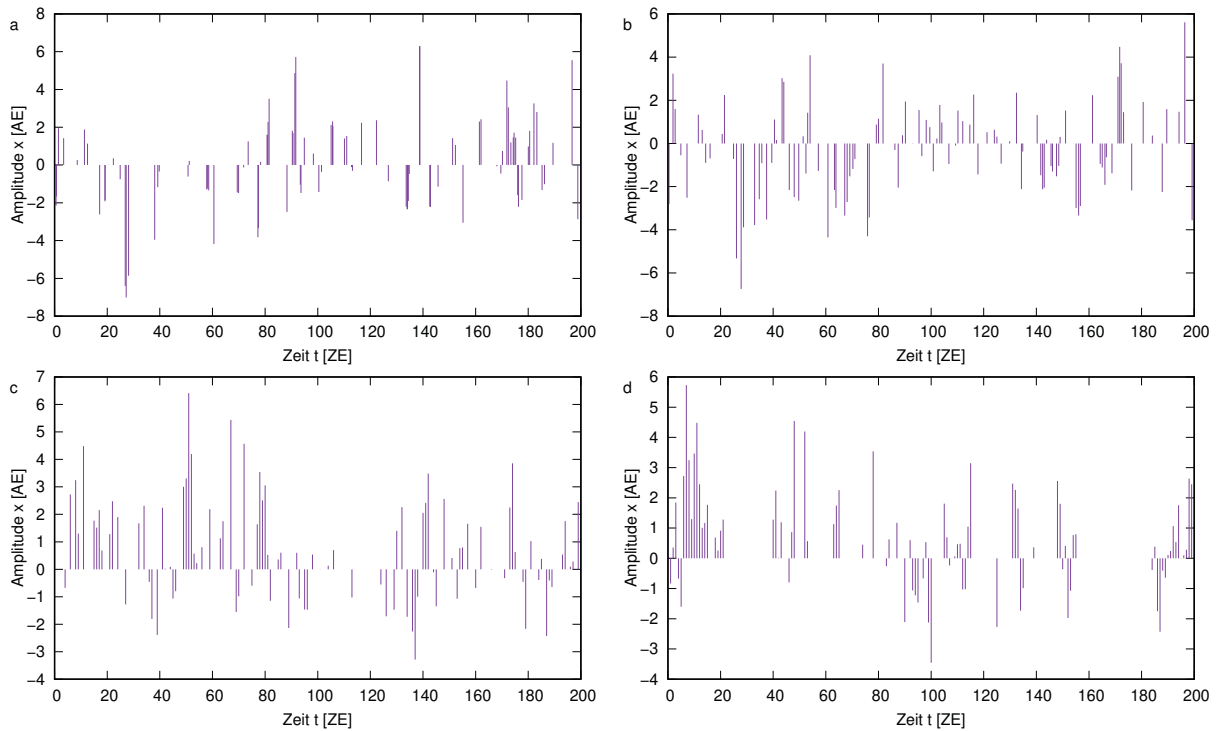


Abbildung 1: Einzelrealisierungen der Signale (a) ideal zufällige Abtastung (b) zufällige Abtastung mit Mindestabstand, (c) unabhängige Fehlstellen und (d) korrelierte Datenlücken (AE - Amplitudeneinheit, ZE - Zeiteinheit)

4 Empirische Bestimmung der Korrekturkoeffizienten

Stehen keine theoretischen Kenntnisse über die Korrelation der Abtastpunkte zur Verfügung, dann können die β'_k auch empirisch aus den bereits aufgenommenen Daten bestimmt werden. Dafür wird die direkte Spektralschätzung mittels der Fourier-Transformation aus Gleichung (1) noch einmal auf die Daten und ihre Abtastzeitpunkte angewendet, allerdings jetzt mit der Einheitsamplitude. Die Lomb-Scargle-Methode kommt hier aufgrund von Schwierigkeiten mit Zeitreihen mit konstanter Amplitude nicht zum Einsatz. Es ergibt sich

$$\beta'_k = \text{IDFT} \left\{ \frac{M}{N^2} \left| \sum_{i=1}^N e^{-2\pi i f_j t_i} \right|^2 \right\} = \frac{1}{N^2} \cdot \sum_{j=-\lfloor M/2 \rfloor}^{\lfloor (M-1)/2 \rfloor} \left| \sum_{i=1}^N e^{-2\pi i f_j t_i} \right|^2 \cdot e^{2\pi i f_j \tau_k} \quad (13)$$

wieder mit $f_j = j\Delta f$, $j = -\lfloor M/2 \rfloor \dots \lfloor (M-1)/2 \rfloor$ und $\tau_k = k\Delta\tau$, $k = -\lfloor M/2 \rfloor \dots \lfloor (M-1)/2 \rfloor$.

5 Simulation

Es wurde ein Moving-Average-Prozess mit einem vorgegebenen Spektrum mit exponentiell abnehmender Leistungsdichte und einer markanten Delle simuliert. Für die diskrete Verteilung der Abtastintervalle wurde eine Prozessordnung von 200 gewählt. Für die stetige Verteilung der Abtastintervalle wurde dieser Prozess 10-fach überabgetastet. Die Prozessordnung war entsprechend 2000. Jede Realisierung generiert jeweils ein Signal der Dauer $T = 200$ ZE (Zeiteinheiten). Die Simulation erfolgt so, dass das Signal an beiden Enden zusammenpasst, um zusätzliche Fehler durch die implizite Annahme einer periodischen Fortsetzung des Signals zu verhindern. Die Signale sind mittelwertfrei und haben eine Standardabweichung von 2 AE (Amplitudeneinheiten). Auf diese Zeitreihen wurden vier verschiedene Abtastverfahren angewendet:

- zufällige und untereinander unabhängige Abtastung der überabgetasteten Zeitreihe mit einer mittleren Abtastrate von $\alpha = 0.5 \text{ ZE}^{-1}$,
- zufällige Abtastung der überabgetasteten Zeitreihe mit einer mittleren Abtastrate von $\alpha =$

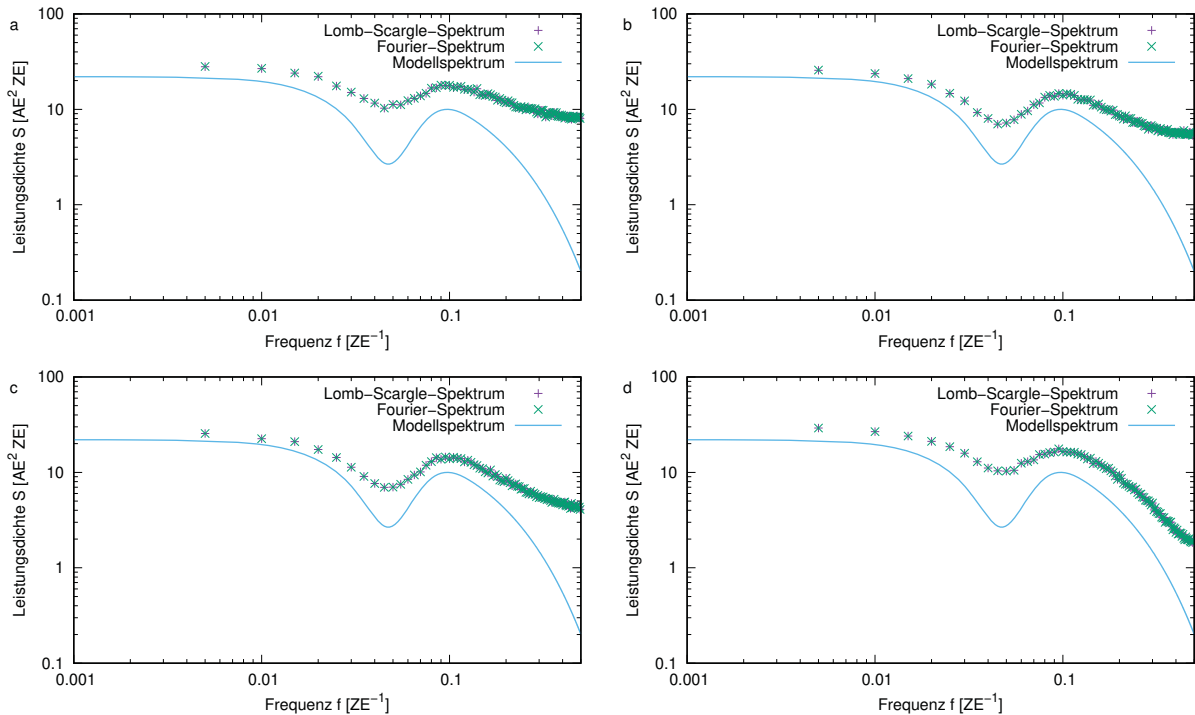


Abbildung 2: Primäre Leistungsdichteschätzung (ohne Korrektur), gemittelt über 1000 Signale für (a) ideal zufällige Abtastung (b) zufällige Abtastung mit Mindestabstand, (c) unabhängige Fehlstellen und (d) korrelierte Datenlücken (AE - Amplitudeneinheit, ZE - Zeiteinheit)

0.5 ZE^{-1} und einem zeitlichen Mindestabstand von 0.5 ZE für eine grobe Nachahmung einer Prozessortzeit,

- (c) zufällige und untereinander unabhängige Entfernung von Werten aus der äquidistanten Zeitreihe mit einer Wahrscheinlichkeit von 50 %,
- (d) Folgen von gültigen bzw. ungültigen Werten der äquidistanten Zeitreihe, wobei die Gültigkeit in jedem Zeitschritt mit einer Wahrscheinlichkeit von 20 % wechselt (ein Markov-Prozess). Das ergibt ebenfalls durchschnittlich 50 % gültige Werte, allerdings in zusammenhängenden und damit korrelierten Blöcken exponentialverteilter Dauer.

Abb. 1 zeigt individuelle Realisierungen der Signale und ihrer Abtastungschemen. Die Zeitreihen in Abb. 1a und b wurden zufällig abgetastet, wobei in Abb. 1b ein Mindestintervall eingehalten wurde, während die Abtastung in Abb. 1a rein zufällig und unabhängig voneinander erfolgte. In Abb. 1c und d wurden die Daten nominell äquidistant abgetastet und anschließend Datenpunkte verworfen. Die Auswahl in Abb. 1c erfolgte rein zufällig und unabhängig voneinander, während in Abb. 1d größere, zusammenhängende Lücken zu erkennen sind. Unabhängig vom Abtastungschemen wurde in allen vier Fällen eine mittlere Datenrate von $\alpha = 0.5 \text{ ZE}^{-1}$ realisiert.

Von diesen Signalen wurden die primären Leistungsdichtespektren mit der Fourier-Transformation und mit der Lomb-Scargle-Methode bestimmt. In Abb. 2 ist die über 1000 Realisierungen gemittelte Leistungsdichte dargestellt. Für alle vier Abtastungschemen ist ein signifikanter systematischer Fehler zu erkennen. Dabei sind diese Spektren selbst erwartungstreue Schätzungen der Leistungsdichte der abgetasteten Signale. Der systematische Fehler besteht hingegen zum beobachteten stochastischen Prozess. Die unregelmäßige Abtastung verändert den spektralen Charakter des aufgenommenen Signals gegenüber dem Prozess. Das erhaltene Signalspektrum hängt damit vom Abtastungschemen ab. Entsprechend sind die Spektren in Abb. 2 je nach Abtastungschemen unterschiedlich, während es keine signifikanten Unterschiede zwischen den primären Schätzungen des Spektrums mittels der Fourier-Transformation oder der Lomb-Scargle-Methode gibt.

In Abb. 3 wurden die primären Schätzungen entsprechend den oben eingeführten Methoden korrigiert. Für zufällige und unabhängige Abtastung mit stetiger Verteilung der Abtastintervalle und für zufällige

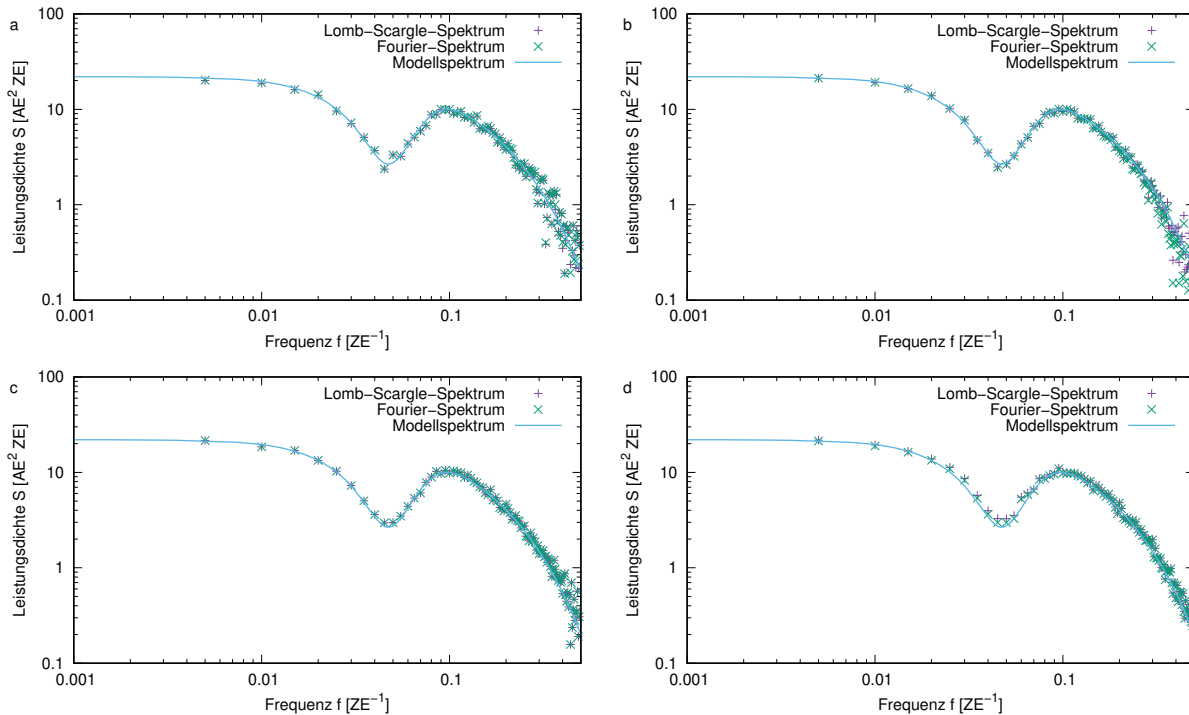


Abbildung 3: Leistungsdichteschätzung mit Korrektur, gemittelt über 1000 Signale, mit Korrekturfaktoren aus theoretischer Herleitung aus dem Abtastmodell für (a) ideal zufällige Abtastung (b) zufällige Abtastung mit Mindestabstand, (c) unabhängige Fehlstellen und (d) korrelierte Datenlücken (AE - Amplitudeneinheit, ZE - Zeiteinheit)

Auswahl der verworfenen Werte bei diskreter Verteilung der Abtastintervalle (Abb. 3a und c) wurde das Spektrum direkt nach Gleichung (10) bzw. (11) korrigiert, und zwar mit bekannter mittlerer Abtastrate von $\alpha = 0.5 \text{ ZE}$. Für korrelierte Abtastung und stetiger Verteilung der Abtastintervalle mit dem Mindestabstand und für korrelierte Datenlücken und diskreter Verteilung der Abtastintervalle (Abb. 3b und d) wurde das primäre Spektrum in die entsprechende Korrelationsfunktion transformiert, die dann nach Gleichung (12) korrigiert und schließlich in das entsprechende Spektrum zurücktransformiert wurde. Die β'_k wurden aus den Modellen der Abtastschemen theoretisch hergeleitet. Die Herleitungen dieser Zusammenhänge sind unter [1] in einem Anhang zu diesem Beitrag zu finden. Sie gelten nur für genau die hier verwendeten Abtastschemen und sind nicht allgemeingültig.

Die Ergebnisse in Abb. 3 zeigen, dass die vorgeschlagene Korrektur zu erwartungstremen (in Fall b leider nur fast erwartungstreu) Schätzungen der Leistungsdichte führt. Im Fall b bleibt ein kleiner Restfehler, der durch Aliasing entsteht. Durch die zufällige Abtastung werden höherfrequente Anteile des Spektrums über einen Frequenzbereich verschmiert, und zwar auch solche Anteile, die ursprünglich in dem betrachteten Frequenzbereich lagen. Die vorgeschlagene Korrektur ist nicht dafür gedacht, solche Aliasing-Fehler zu korrigieren. Sie sind das Resultat der insuffizienten Informationsextraktion durch die Abtastung. Dadurch fehlende Informationen können durch die hier beschriebene Korrektur nicht wieder hinzugefügt werden.

In Abb. 4 wurde die Korrektur abschließend durch die mit Gleichung (13) empirisch bestimmten Werte β'_k durchgeführt. Die Korrektur erfolgt hier für alle vier Abtastschemen anhand der Korrelationsfunktion mit anschließender Transformation in die dargestellten Spektren. Auch hier zeigen die Ergebnisse, dass die Korrektur für alle Abtastschemen funktioniert, bis auf den kleinen Aliasing-Fehler in Fall b, der so nicht zu korrigieren ist.

6 Zusammenfassung

Eine zufällige Abtastung von Zeitreihen bewirkt einen systematischen Fehler der Spektralschätzung gegenüber dem betrachteten Prozess. Die oft als Referenz verwendete Lomb-Scargle-Methode zur Spektralschätzung unregelmäßig abgetasteter Zeitreihen zeigt in dieser Untersuchung keine Vorteile

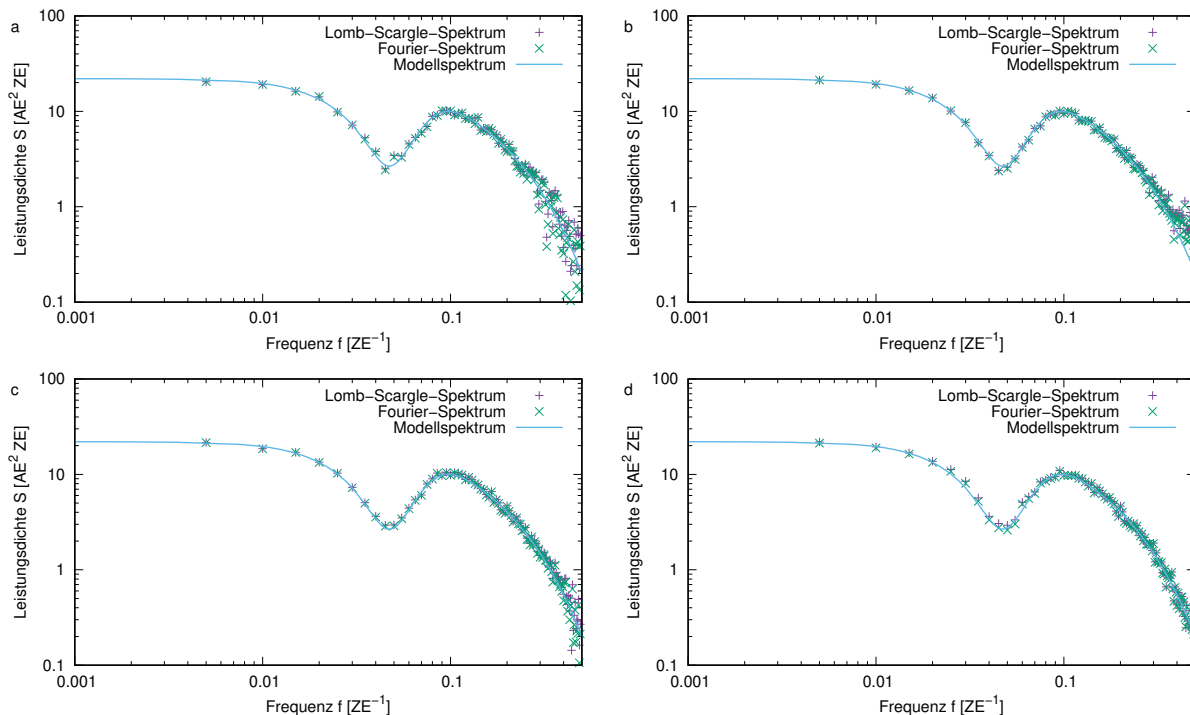


Abbildung 4: Leistungsdichteschätzung mit Korrektur, gemittelt über 1000 Signale, mit empirisch ermittelten Korrekturfaktoren für (a) ideal zufällige Abtastung (b) zufällige Abtastung mit Mindestabstand, (c) unabhängige Fehlstellen und (d) korrelierte Datenlücken (AE - Amplitudeneinheit, ZE - Zeiteinheit)

gegenüber einer direkten Spektralschätzung mittels der Fourier-Transformation. Die systematischen Fehler durch die ungleichmäßige Abtastung können mit den vorgestellten Methoden analysiert, prognostiziert und schließlich korrigiert werden. Insbesondere ist das vorgeschlagene Verfahren nicht nur auf rein zufällige Abtastenschemen anwendbar, sondern auch auf Abtastenschemen mit Korrelation zwischen den Abtastungen. Es ist auch nicht auf bestimmte Abtastmodelle beschränkt, sondern universell für beliebige spektrale Zusammensetzung der Abtastung anwendbar.

Literatur

- [1] URL <http://www.nambis.de/publications/gala19.html>.
- [2] P Buchhave and C M Velte. Reduction of noise and bias in randomly sampled power spectra. *Experiments in Fluids*, 56:79, 2015.
- [3] N Damaschke, V Kühn, and H Nobach. A fair review of non-parametric bias-free autocorrelation and spectral methods for randomly sampled data in laser Doppler velocimetry. *Digital Signal Processing*, 76:22–33, 2018.
- [4] M Gaster and J B Roberts. The spectral analysis of randomly sampled records by a direct transform. *Proc. R. Soc. Lond. A.*, 354:27–58, 1977.
- [5] A Khintchine. Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109:604–615, 1934.
- [6] N R Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39:447–462, 1976.
- [7] J D Scargle. Studies in astronomical time series analysis. II. statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853, 1982.
- [8] M Zechmeister and M Kürster. The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms. *Astronomy and Astrophysics*, 496(2):577–584, 2009.