Saisonbereinigungsverfahren in der Preisstatistik

5

Preise

Bernd Schips
Winfried Stier

Das CENSUS-X-11-Verfahren

Darstellung, Kritik, Alternativen

Bundesamt für Statistik
Office fédéral de la statistique
Ufficio federale di statistica
Uffizi federal da statistica

BFS OFS UST

Bern, 1993

Statistik der Schweiz

Die vom Bundesamt für Statistik (BFS) herausgegebene Reihe «Statistik der Schweiz» gliedert sich in folgende Fachbereiche:

0 Bereichsübergreifende Themen
1 Bevölkerung
2 Raum, Landschaft und Umwelt
3 Erwerbsleben
4 Volkswirtschaftliche Gesamtrechnungen
5 Preise
6 Produktion, Handel und Verbrauch
7 Land- und Forstwirtschaft
8 Energie
9 Bau- und Wohnungswesen
10 Tourismus
11 Verkehr
12 Geldmenge, Finanzmärkte und Banken
13 Versicherungen
14 Gesundheit
15 Bildung und Wissenschaft
16 Kultur, Lebensbedingungen und Sport
17 Politik
18 Öffentliche Finanzen
19 Rechtspflege

Statistique de la Suisse

La série «Statistique de la Suisse» publiée par l'Office fédéral de la statistique (OFS) couvre les domaines suivants:

0 Thèmes généraux
1 Population
2 Espace, paysage et environnement
3 Emploi et vie active
4 Comptes nationaux
5 Prix
6 Production, commerce et consommation
7 Agriculture et sylviculture
8 Energie
9 Construction et logement
10 Tourisme
11 Transports et communications
12 Masse monétaire, marchés financiers et banques
13 Assurances
14 Santé
15 Education et science
16 Culture, conditions de vie et sport
17 Politique
18 Finances publiques
19 Droit et justice

Saisonbereinigungsverfahren in der Preisstatistik

Bernd Schips
Winfried Stier

Das CENSUS-X-11-Verfahren

Darstellung, Kritik, Alternativen

BFS OFS UST

Bundesamt für Statistik
Office fédéral de la statistique
Ufficio federale di statistica
Uffizi federal da statistica

Bern, 1993

Herausgeber: Bundesamt für Statistik (BFS)

Auskunft: Dieter Koch, BFS, Tel. 031 61 28 43 *(ab 25. 9. 1993: 031 322 28 43)*
Thomas Gross, BFS, Tel. 031 61 29 41 *(ab 25. 9. 1993: 031 322 29 41)*

Bearbeitung: Bernd Schips, Winfried Stier
Forschungsinstitut für empirische Wirtschaftsforschung der Hochschule St. Gallen
im Auftrag des BFS (Auftrag vom 26. Mai 1992)

Begleitung des Auftrags: Dieter Koch, Thomas Gross, BFS

Vertrieb: Bundesamt für Statistik
CH-3003 Bern
Tel. 031 323 60 60

Bestellnummer: 047-0

Preis: Fr. 11.–

Reihe: Statistik der Schweiz

Fachbereich: 5 Preise

Grafik/Layout: Autoren

Copyright: BFS, Bern 1993
Abdruck – ausser für kommerzielle Nutzung –
unter Angabe der Quelle gestattet.

ISBN: 3-303-05158-5

Inhalt

0. Erläuterung der Aufgabenstellung und der Zielsetzung dieser Arbeit 5
Literaturhinweise 5

I. Wie stellt man Saisonalität fest? 5

1. Einleitung 5
2. Stochastische Prozesse 5
2.1 Grundlagen 5
2.2 Stationäre stochastische Prozesse 6
3. Grundzüge der Spektralanalyse 8
3.1 Vorbemerkungen 8
3.2 Spektren stationärer Prozesse 9
3.3 Schätzung eines Spektrums 12
3.4 Spektralanalyse zur Aufdeckung saisonaler Bewegungen 13
4. Analyse der Zeitreihen für die Preisnotierungen 20
4.1 Beschreibung der untersuchten Zeitreihen 20
4.2 Empirische Befunde 20
Literaturhinweise 22

II. Zur Konstruktion von Saisonbereinigungsverfahren 23

1. Zielsetzungen bei einer Saisonbereinigung 23
2. Übersicht über die Entwicklung von Saisonbereinigungsverfahren 24
3. Kriterien für Saisonbereinigungsverfahren 31
4. Zum Design von Saisonfiltern (Transfer- und Phasenfunktionen) 32
5. Eigenschaften einiger gebräuchlicher Filter 35
5.1 Vorbemerkungen 35
5.2 Erste Differenzen 35
5.3 Zwölfte Differenzen 37
5.4 Wachstumsraten 39
6. Probleme bei der Konstruktion von Saisonfiltern 41
7. Zur Konstruktion von Tiefpass-Filtern 46
7.1 Filter im Zeitbereich: FIR-Filter und IIR-Filter 46
7.2 Filter im Frequenzbereich 47
7.3 Globale und lokale Filter 49
Literaturhinweise 49

III. Eigenschaften des CENSUS-X-11-Verfahrens im Frequenzbereich 52

1. Einleitung 52
2. Theoretische Überlegungen zur Bestimmung von Amplituden- und Phasenfunktion des CENSUS-X-11-Verfahrens 52
3. Experimentelles Design 53
4. Resultate für die Amplitudenfunktionen 54
5. Resultate für die Phasenfunktionen ... 56
6. Interpretationen der Amplitudenfunktionen 61
7. Interpretationen der Phasenfunktionen . 61
8. Schätzung der glatten Komponente simulierter Reihen mit dem CENSUS-X-11-Verfahren 63
Literaturhinweise 72

IV. Anwendung des CENSUS-X-11-Verfahrens auf die Zeitreihen mit Preisnotierungen und Vorschläge für das künftige Vorgehen ... 73

1. Beurteilung von CENSUS-X-11-bereinigten Zeitreihen 73
2. Anwendungen lokaler Tiefpass-Filter .. 76
3. Randeigenschaften des globalen und lokalen Tiefpass-Filters sowie der glatten Komponente des CENSUS-X-11-Verfahrens 77
4. Fazit 79
Literaturhinweis 80

0. Erläuterung der Aufgabenstellung und der Zielsetzung dieser Arbeit

In vielen Ländern erfreut sich seit einigen Jahren das Saisonbereinigungsverfahren CENSUS X 11 unter Anwendern aus dem Bereich der amtlichen Statistik, aber auch bei Vertretern von Wirtschaftsforschungsinstituten grosser Beliebtheit. Eine derartig grosse Verbreitung könnte auf unbestreitbaren methodischen Vorzügen dieses Verfahrens beruhen oder aber auch nur Ausdruck eines geschickten Marketing der betreffenden Verfahrenskonstrukteure sein.

Bezüglich der methodischen Stärken und Schwächen des Verfahrens gibt es in der einschlägigen Literatur schon seit langem eine nur anfänglich kontroverse Diskussion mit einer zunehmenden Tendenz zu einer aus methodischer Sicht äusserst kritischen Einstellung gegenüber dem CENSUS-X-11-Verfahren. In der BRD z.B. wurde deshalb schon früh nach methodischen Alternativen gesucht. Das Statistische Bundesamt z.B. verwendet heute das sogenannte BERLINER VERFAHREN [1)], die Arbeitsgemeinschaft der Wirtschaftsforschungsinstitute bemühte sich lange um die Entwicklung eines eigenständigen Verfahrens [2)], die Deutsche Bundesbank hält aber nach wie vor am CENSUS-X-11-Verfahren fest, obwohl von wissenschaftlicher Seite mehrfach überzeugende Nachweise bestimmter methodischer Unzulänglichkeiten vorgelegt wurden [3)]. Solche Unzulänglichkeiten, insbesondere am aktuellen Rand, haben dazu geführt, dass z.B. in Kanada eine spezielle Version des CENSUS-X-11-Verfahrens verwendet wird [4)] und neuerdings auch am CENSUS-Bureau in den USA eine modifizierte Version X 12 in Arbeit ist [5)].

Da die methodischen Bedenken weitgehend bekannt und akzeptiert sind, hat die Verbreitung des CENSUS-X-11-Verfahrens unter den Anwendern sicher auch etwas mit dem Marketing der Verfahrenskonstrukteure zu tun. Das CENSUS-X-11-Verfahren war eines der ersten verfügbaren, weitgehend portablen Programme zur Saisonbereinigung. Das Programm stiess auf eine Marktlücke und fand deshalb raschen Anklang. Es entstand auch der Eindruck, dass mit dem Einsatz ein und desselben Programmes eine direkte Vergleichbarkeit der saisonbereinigten Daten gegeben sei. Diese Vorstellung hält sich teilweise noch heute, obwohl die zahlreichen Parameter zur Steuerung des Verfahrens und die praktisch überall vorgenommenen Modifikationen zwangsläufig zu teilweise erheblichen und methodisch bedingten Unterschieden in den Ergebnissen führen.

In dieser Arbeit sollen daher die grundsätzlichen methodischen Einwände gegenüber dem CENSUS-X-11-Verfahren dargestellt und deren Auswirkungen erläutert werden. Als Beispiel dienen die Zeitreihen für Preisnotierungen in der Schweiz (Grosshandels- und Konsumentenpreisindizes). Abschliessend sollen Vorschläge sowohl für aus methodischer Sicht befriedigende Bereinigungsverfahren als auch für die Verarbeitung bereinigter Daten gemacht werden.

Literaturhinweise

1) Vgl. z.B. NULLAU B., 'Probleme der praktischen Anwendung des BERLINER VERFAHRENS', in: W. Wetzel (Hrsgb.), Neuere Entwicklungen auf dem Gebiet der Zeitreihenanalyse, Göttingen 1970, S. 95-129

2) Vgl. z.B. GOLDRIAN G., 'Eine neue Version des ASA-II-Verfahrens zur Saisonbereinigung von wirtschaftlichen Zeitreihen', in: Wirtschaftskonjunktur 4 (1973), S. 26-32

3) Vgl. SCHÄFFER K.A., 'Beurteilung einiger herkömmlicher Methoden zur Analyse von ökonomischen Zeitreihen', in: W. Wetzel (Hrsgb.), Neuere Entwicklungen auf dem Gebiet der Zeitreihenanalyse, Göttingen 1970, S. 131-164

4) Vgl. DAGUM E.B., The X-11 ARIMA Seasonal Adjustment Method, Statistics Canada, Catalogue 12-546 E, 1980

5) Vgl. FINDLEY D.F. et al., 'Towards X-12 ARIMA', in: Proceedings of the Bureau of Census Annual Research Conference 4 (1988), S. 591-622

I. Wie stellt man Saisonalität fest?

1. Einleitung

Dass sich in vielen unterjährigen ökonomischen Zeitreihen, also z.B. in Monats-, Quartals-, Halbsjahresreihen, jahreszeitliche Einflüsse widerspiegeln ist allgemein bekannt. Weniger bekannt ist dagegen, mit welchen Methoden man *generell* feststellen kann, ob überhaupt Saisonalität in einer konkreten Zeitreihe vorliegt, wenn man einmal von solchen Reihen absieht, bei denen schon eine rein optische Inspektion eindeutig zeigt, dass eine saisonale Komponente vorhanden ist. Nicht selten sind jedoch Reihen, bei denen nicht direkt erkennbar ist, ob sie saisonal beeinflusst sind oder nicht. Zu diesen Reihen gehören auch die meisten der hier besonders interessierenden Preisreihen. Für derartige Reihen lässt sich nur mit Hilfe spezieller Methoden feststellen, ob eine saisonale Komponente vorhanden ist.

Im folgenden sollen die praktisch wichtigsten Aspekte dieser Analysemethoden dargestellt werden, wobei ihre formal-mathematische Seite nur insoweit berücksichtigt wird, als unbedingt notwendig erscheint. Es wird insbesondere versucht, die grundlegenden Zusammenhänge graphisch darzustellen und zu verdeutlichen [1)].

2. Stochastische Prozesse

2.1 Grundlagen

Ein Charakteristikum der heutigen Zeitreihenanalyse ist darin zu sehen, dass als theoretischer Bezugsrahmen für Zeitreihenmodelle die *Theorie der stochastischen Prozesse*

verwendet wird. Zeitreihenmodelle sind somit *stochastische* Modelle, was impliziert, dass konkrete Zeitreihen als *Realisationen* stochastischer Prozesse interpretiert werden.

Unter einem *stochastischen Prozess* ist eine Menge von Zufallsvariablen $\{X_t \mid t \in T\}$ zu verstehen, die von einer Indexmenge T abhängt, welche als *Parameterraum* des Prozesses bezeichnet wird. In der Zeitreihenanalyse bedeutet T immer eine Menge von Zeitpunkten. Dabei müssen grundsätzlich zwei verschiedene Fälle unterschieden werden. Im ersten Fall ist T ein Intervall in $\mathbb{R}$, und T wird als *kontinuierlicher Parameterraum* bezeichnet. Im zweiten Fall ist T entweder eine endliche Menge von reellen Zahlen oder T enthält zumindest in jedem beschränkten Intervall nur endlich viele Zahlen. Beispiele dafür sind T = {0,1,2, ...}: = $\mathbb{N}$ oder T = {0, ±1, ±2, ...}: = $\mathbb{Z}$. T heisst dann *diskreter Parameterraum*. Ist insbesondere T endlich, dann wird $\{X_t \mid t \in T\}$ als *endlicher* stochastischer Prozess bezeichnet. Hier wird davon ausgegangen, dass T durchwegs diskret ist. Dieser Fall ist besonders wichtig, da Daten in der Praxis meistens in diskreter Form vorliegen. Bei ökonomischen Daten ist dies ausschliesslich der Fall.

Wie gewöhnliche Zufallsvariablen können auch stochastische Prozesse durch ihre Momente charakterisiert werden. Es ist jedoch zweckmässig, hier allgemeiner von *Momentfunktionen* zu sprechen, da die Momente stochastischer Prozesse grundsätzlich vom Parameter $t \in T$ abhängig sind.

Der Ausdruck

$$\mu_t := E[X_t]$$

heisst *Erwartungswertfunktion* (oder *Mittelwertfunktion*) und

$$\sigma_t^2 := Var[X_t]$$

Varianzfunktion des stochastischen Prozesses X_t. Die *Kovarianzfunktion* ist definiert durch

$$Cov(X_s, X_t) = \gamma(s,t) := E[(X_s - \mu_s)(X_t - \mu_t)],\ s, t \in T$$

und seine *Korrelationsfunktion* durch

$$Cor(X_s, X_t) = \rho(s,t) := \frac{Cov(X_s, X_t)}{[Var(X_s)Var(X_t)]^{1/2}} .$$

2.2 Stationäre stochastische Prozesse

Eine praktisch wichtige Klasse von stochastischen Prozessen bilden die *stationären Prozesse* und innerhalb dieser wiederum die *schwach stationären Prozesse*. Ein stochastischer Prozess heisst *streng stationär*, falls für seine Verteilungsfunktion gilt

$$F(x_1,x_2,..,x_n;\ t_1,t_2,..,t_n) = F(x_1,x_2,...,x_n;\ t_1+\tau,t_2+\tau,...,t_n+\tau)$$

für alle $n \in \mathbb{N}$, alle n-tupel $(t_1,t_2,...,t_n) \in T^n$ und für jedes τ, für welches $(t_1+\tau, t_2+\tau,...,t_n+\tau)$ wieder in T^n liegt.

Dies bedeutet, dass die Verteilung der n Zufallsvariablen $X_1, X_2,...,X_n$ in den Zeitpunkten $t_1, t_2,...,t_n$ und $t_1+\tau, t_2+\tau,.., t_n+\tau$ gleich ist. Streng stationäre Prozesse sind somit invariant gegenüber Verschiebungen des Beobachtungszeitpunktes, d.h. sie sind *zeittranslationsinvariant*. Für einen solchen Prozess folgt daraus, dass sowohl die Erwartungswertfunktion als auch die Varianzfunktion zeitunabhängig, also konstant sind. Jedoch gilt

$$F(x_1,x_2;\ t_1,t_2) = F(x_1,x_2;\ t_2-t_1) ,$$

d.h. die Verteilungsfunktion *2.Ordnung* (d.h.also für n = 2) hängt (nur) von der *Zeitdifferenz* t_2-t_1 ab. Dies impliziert, dass auch die Kovarianz- und die Korrelationsfunktion eines streng stationären Prozesses (nur) von der Zeitdifferenz abhängen.

Für die praktische Zeitreihenanalyse ist die Postulierung der strengen Stationarität in der Regel zu einschränkend. Ausreichend ist die *schwache Stationarität*, für die gilt

$$\mu_t = const.,$$

$$\sigma_t^2 = const.,$$

$$\gamma(s,t) = \gamma(t-s)$$

und

$$\rho(s,t) = \rho(t-s) .$$

Für schwach stationäre Prozesse sind auch die Bezeichnungen *stationär im weiteren Sinne, Kovarianz-stationär* oder einfach *stationär* gebräuchlich. Hier soll konsequent unter *stationär* immer *schwach stationär* verstanden werden. Ohne Beschränkung der Allgemeinheit kann angenommen werden, dass grundsätzlich $\mu_t = 0$ ist, da dies durch Subtraktion einer Konstanten vom Prozess immer erreicht werden kann. Dadurch werden viele formale Ausdrücke wesentlich einfacher.

Bei schwach stationären Prozessen interessiert man sich sozusagen nur für die Momentfunktionen bis zur 2. Ordnung, alle höheren Momente bleiben ausser Betracht. Streng stationäre Prozesse sind immer auch schwach stationär, die Umkehrung gilt nur in speziellen Fällen.

Der einfachste stationäre Prozess ist dadurch charakterisiert, dass die Zufallsvariablen ε_t paarweise unkorreliert sind, den Erwartungswert Null und die Varianz σ^2 haben. Es gilt also

$$\mu_t = 0,$$
$$\sigma_t^2 = \sigma^2$$

und

$$\gamma_\varepsilon(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

bzw.

$$\rho_\varepsilon(h) = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 . \end{cases}$$

Ein solcher Prozess wird als *weisses Rauschen* (*white noise*) bezeichnet.

Ein komplizierterer Prozess ist z.B. gegeben durch

$$(X_t - \mu) = \phi(X_{t-1} - \mu) + \varepsilon_t ,$$

wobei ε_t ein weisses Rauschen bezeichnet. Dieser Prozess wird *autoregressiver Prozess 1.Ordnung* genannt. Diese Bezeichnung erklärt sich dadurch, dass der Prozess in jedem Zeitpunkt t sozusagen auf sich selber "zurückgreift", wobei der time-lag eine Zeiteinheit beträgt. μ bezeichnet den Erwartungwert von X_t und ϕ den *Prozessparameter*. Dieser Prozess ist nicht für beliebige Prozessparameter ϕ stationär, sondern nur für solche, die betragsmässig kleiner als Eins sind.

Die Momentfunktionen eines AR(1)-Prozesses lauten:

$$\mu_t = 0, \quad \sigma_t^2 = \frac{\sigma_\varepsilon^2}{1-\phi^2} ,$$

$$\gamma(h) = \left[\frac{\sigma_\varepsilon^2}{1-\phi^2}\right] \phi^h, \quad h = 0,1,2,... ,$$

und

$$\rho(h) = \phi^h, \quad h = 0,1,2,...$$

Allgemein ist ein *autoregressiver Prozess p-ter Ordnung oder kurz AR(p)-Prozess*, definiert durch

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t$$

mit $\phi_p \neq 0$, wobei der Einfachheit halber (aber ohne Beschränkung der Allgemeinheit) vorausgesetzt wird, dass der Erwartungswert von X_t gleich Null ist. Auch hier können die Prozessparameter nicht beliebig gewählt werden, wenn der AR(p) stationär sein soll. Da die Stationaritätsbedingungen relativ kompliziert sind, kann hier darauf nicht näher eingegangen werden. Ebenso soll auf eine Betrachtung der Momentfunktionen verzichtet werden.

Der Prozess

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} ,$$

wobei ε_t weisses Rauschen bezeichnet, heisst (endlicher) Moving-Average-Prozess der Ordnung q - oder kurz MA(q)-Prozess - mit den *Prozessparametern* $\theta_0 = 1$, $\theta_1,...,\theta_q$ mit $\theta_q \neq 0$. Üblich ist auch die Bezeichnung *Gleitender-Durchschnitts-Prozess*. MA-Prozesse sind nichts anderes als gewogene Mittel aus unkorrelierten Zufallsvariablen, wobei jedoch die Summe der Gewichte im allgemeinen nicht Eins ist. Offensichtlich sind MA-Prozesse für beliebige Parameter stationär. Für einen MA(1)-Prozess

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

ergibt sich z.B. für die Erwartungswert- und Varianzfunktion

$$\begin{aligned} E(X_t) &= E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = 0 \\ Var(X_t) &= Var(\varepsilon_t) + \theta^2 Var(\varepsilon_{t-1}) \\ &= \sigma_\varepsilon^2(1 + \theta^2) \end{aligned}$$

und für die Autokovarianz- bzw. Autokorrelationsfunktion

$$\begin{aligned} \gamma(1) &= E\,[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\ &= \theta E(\varepsilon_{t-1}^2) \\ &= \theta\sigma_\varepsilon^2 \\ \gamma(h) &= 0 \text{ für } h \geq 2 \end{aligned}$$

und

$$\rho(1) = \frac{\theta}{1+\theta^2}$$

$$\rho(h) = 0 \text{ für } h \geq 2 .$$

Die beiden Prozesstypen können auch kombiniert werden. Man spricht dann von einem ARMA-Prozess der Ordnung (p,q) - kurz ARMA(p,q)-Prozess -, der definiert ist durch

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} .$$

Auch für diese Prozessklasse sei auf die Darstellung von Stationaritätsbedingungen und Momentfunktionen verzichtet.

3. Grundzüge der Spektralanalyse

3.1 Vorbemerkungen

Die bisherigen Ausführungen können dahingehend charakterisiert werden, dass sie ausschliesslich im *Zeitbereich* formuliert wurden. Beispielsweise wurden die Momentfunktionen allgemein als Funktionen der Zeit definiert. Neben dieser Betrachtungsform existiert jedoch noch eine andere, die wesentliche Eigenschaften eines stationären Prozesses in sehr anschaulicher Weise darzustellen erlaubt. Dabei handelt es sich um eine Betrachtung im sogenannten *Frequenzbereich*. Überlegungen und Schlussfolgerungen in diesem Bereich mögen zwar zunächst vielleicht als wesentlich komplizierter erscheinen als im gewohnten Zeitbereich. Es wird aber anhand der im folgenden präsentierten Beispiele und Graphiken leicht einzusehen sein, dass Analysen in diesem Bereich und gegebenenfalls eine Kombination von solchen in beiden Bereichen einer Analyse ausschliesslich im Zeitbereich vorzuziehen sind bzw. eine solche sinnvoll ergänzen können.

Zeitbereich und Frequenzbereich stehen einander nicht isoliert gegenüber. Vielmehr lassen sich beide Bereiche durch eine spezielle mathematische Transformation ineinander überführen. Beide Bereiche sind somit zueinander dual. Diese Transformation soll jedoch hier nicht dargestellt werden. Vielmehr soll diese Dualität anhand eines einfachen Beispiels illustriert werden.

Dazu wird von einer (diskreten) Sinusschwingung mit einer Periode von T = 12 Zeiteinheiten (ZE), z.B. 12 Monaten, ausgegangen, d.h. die zu betrachtende Zeitreihe X_t sei folgendermassen definiert:

$$X_t = \sin\lambda t \ , \ t = 1,2,\ldots ,$$

mit

$$\lambda = 2\pi/12 .$$

Im Zeitbereich ergibt sich dann die folgende Darstellung (vgl. Abb. 1):

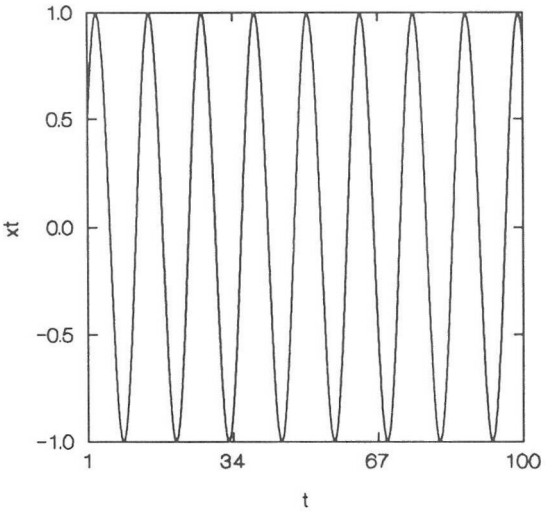


Abb. 1

Dabei ist λ die Frequenz dieser Schwingung in *radian* (rad), die hier 0.5236 beträgt. Ihre Frequenz in *Anzahl Zyklen/ZE* ist f = 1/12. Allgemein besteht zwischen diesen beiden Frequenzbegriffen die Beziehung

$$\lambda = 2\pi f ,$$

wobei $0 \leq f \leq 0.5$ ist. Dass f maximal den Wert 0.5 annehmen kann hängt damit zusammen, dass bei einer diskreten Reihe keine kleineren Perioden als 2 ZE beobachtet werden können. Beispielsweise kann bei einer Monatsreihe keine Schwingung mit z.B. einer Periode von 3 Wochen beobachtet werden.

Spricht man davon, dass diese Sinussschwingung eine Periode von 12 ZE besitze, dann beschreibt man sie im Zeitbereich. Ebenso gut kann man aber auch sagen, sie weise eine Frequenz von f = 1/12 auf, d.h. 1/12 des Zyklus "verstreicht" gewissermassen pro ZE. Diese Charakterisierung im Frequenzbereich kann natürlich auch über die Frequenz in radian, also mit λ erfolgen, was aber im allgemeinen weniger anschaulich ist. Deswegen wird im folgenden in der Regel die Frequenz als Anzahl Zyklen/ZE verstanden.

Man kann sich bei diesem einfachen Beispiel natürlich fragen, welche praktischen Vorteile eine Charakterisierung obiger Reihe im Frequenzbereich gegenüber einer solchen im gewohnten Zeitbereich bringen. Ein solcher ist offensichtlich hier nicht zu erkennen. Die Situation ändert sich aber sofort, wenn kompliziertere Reihen vorliegen. Analysen im Frequenzbereich bieten dann mindestens zwei Vorteile. Einmal geben sie Auskunft über den "Informa-

tionsgehalt" einer Reihe und zum anderen lässt sich angeben, wie sich Transformationen einer Reihe (z.B. Glättungen, Differenzenbildungen oder allgemein Filterungen) auf diesen Informationsgehalt auswirken. Dabei wird dieser im Frequenzbereich vor allem mit Hilfe des Spektrums dargestellt.

3.2 *Spektren stationärer Prozesse*

Ausgangspunkt sei ein stationärer Prozess X_t mit der Kovarianzfunktion $\gamma(h)$. Das *Spektrum* (auch "Powerspektrum" genannt) von X_t ist definiert durch

$$S(f) = \sum_{h=-\infty}^{\infty} \gamma_h e^{i\,2\pi fh}$$
$$= \gamma_0 + 2 \sum_{h=1}^{\infty} \gamma_h \cos 2\pi fh \ .$$

Das Spektrum ist somit als *Fourier-Transformierte* der Kovarianzfunktion definiert. Verwendet man in dieser Transformation anstelle der Kovarianzfunktion die Korrelationsfunktion, dann spricht man von der *Spektraldichte*. Sie ergibt sich aus S(f) durch Division mit der Varianz des Prozesses X_t.

Sowohl Spektrum als auch Spektraldichte sind gerade, stetige, periodische und nichtnegative Funktionen der Frequenz f. Das Integral über S(f), also die Fläche unter dem Spektrum, ist gleich der Varianz des Prozesses, während das Integral über der Spektraldichtefunktion gleich Eins ist. Die letztere Funktion ähnelt somit formal der Dichtefunktion einer Zufallsvariablen, was die Bezeichnung "Spektraldichtefunktion" erklärt.

Das Spektrum eines Prozesses zeigt die Aufteilung der Prozessvarianz auf einzelnen Frequenzen bzw. Frequenzbänder: Eine einem bestimmten Frequenzband zuzuordnende Prozesskomponente ist umso bedeutsamer, je grösser die "spektrale Masse" in diesem Band ist. Allgemein gibt für ein Frequenzband $[f_1, f_2]$ das Integral

$$\int_{f_1}^{f_2} S(f)\,df$$

an, welcher Anteil der Prozessvarianz auf Schwingungskomponenten mit Frequenzen $f \in [f_1, f_2]$ entfällt.

Zur Illustration seien nun Spektren einiger einfacher stochastischer Prozesse betrachtet. Für das oben eingeführte weisse Rauschen ergibt sich z.B.

$$S(f) = \sigma^2 ,$$

d.h. das Spektrum eines white-noise-Prozesses ist konstant, keine Frequenz dominiert, alle Frequenzen liefern den gleichen Beitrag zur Prozessvarianz (vgl. Abb. 2 für $\sigma^2 = 1$):

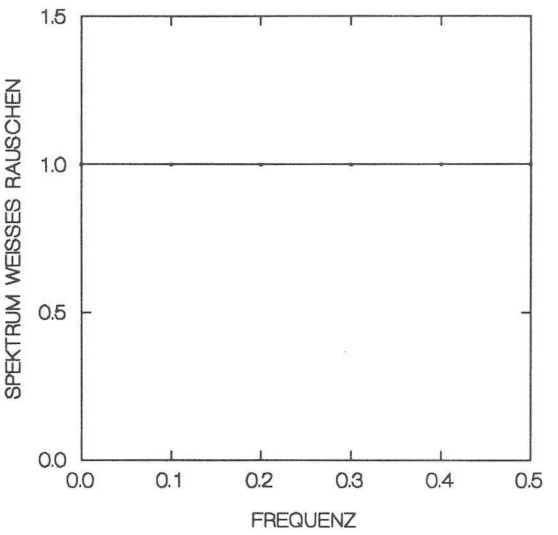


Abb. 2

Ein AR(1)-Prozess hat das Spektrum

$$S(f) = \frac{\sigma^2}{1+\phi^2-2\phi\cos 2\pi f} \ .$$

Der Verlauf dieses Spektrums hängt also vom Prozessparameter ϕ ab. Für positive ϕ ist die spektrale Masse im "Niederfrequenzbereich" (d.h. vor allem in der Nähe der Frequenz f = 0) konzentriert, während für negative ϕ eine Konzentration im "Hochfrequenzbereich" (d.h. hauptsächlich in der Nähe von f = 0.5) zu beobachten ist. Die Abb. 3 und 4 zeigen die Spektren von AR(1)-Prozessen für $\phi = 0.7$ bzw. für $\phi = -0.7$ und jeweils $\sigma^2 = 0.1$.

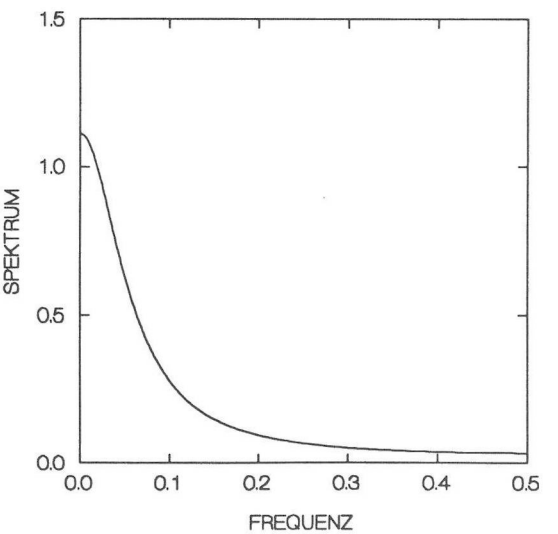


Abb. 3

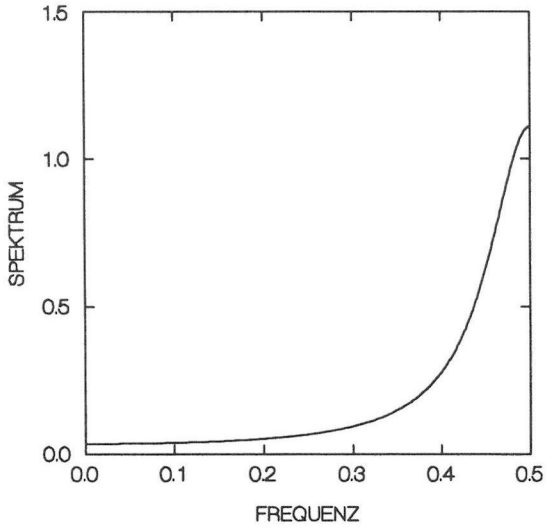


Abb. 4

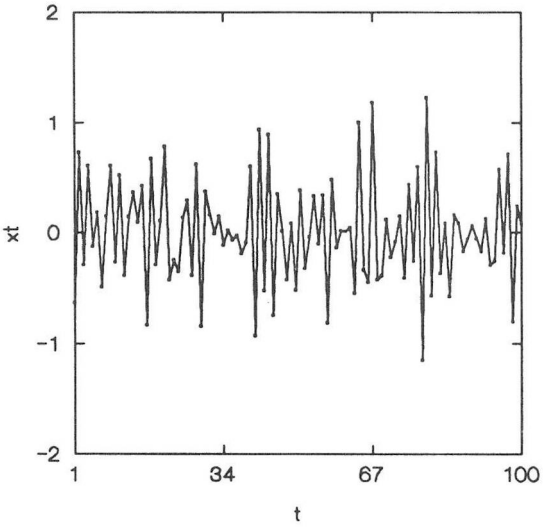


Abb. 6

Die nächsten beiden Abb. 5 und 6 zeigen je eine Realisation mit jeweils 100 Beobachtungswerten dieser beiden Prozesse.

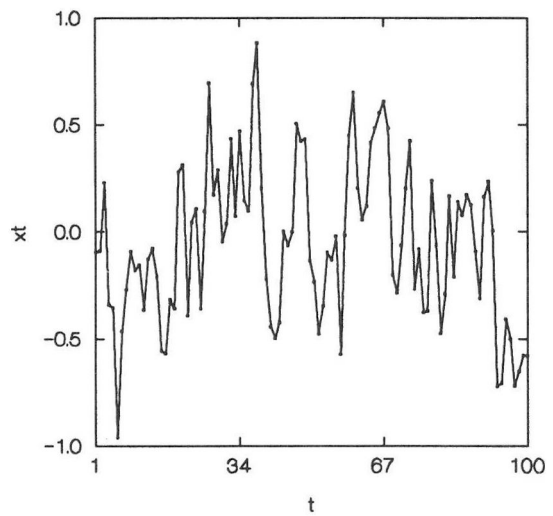


Abb. 5

Während die Zeitreihe in Abb. 5 relativ ruhig verläuft (bitte Skalen beachten), zeigt die Zeitreihe in Abb. 6 ziemlich hektische Auf-Ab-Bewegungen. Dass diese unterschiedlichen Verlaufsformen nicht zufällig, sondern stets zu erwarten sind, ergibt sich unmittelbar aus den beiden Spektren. Der bei positivem ϕ dominierende Niederfrequenzbereich führt zu relativ langsamen Bewegungsabläufen, während ein den Hochfrequenzbereich (insbesondere die Frequenz f = 0.5) akzentuierendes negatives ϕ Oszillationen induziert. Somit enthält ein Spektrum Informationen über die zeitliche Verlaufsform von Prozessrealisationen, d.h. also von Zeitreihen. Dies ist im nächsten Beispiel, dem Spektrum eines AR(2)-Prozesses, noch deutlicher.

Die Prozessgleichung für einen AR(2)-Prozess lautet

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t ,$$

und für das zugehörige Spektrum ergibt sich

$$S(f) = \frac{\sigma^2}{1 + \phi_1^2 + \phi_2^2 + 2(\phi_1\phi_2 - \phi_1)\cos 2\pi f - 2\phi_2 \cos 4\pi f} .$$

Mit den Parametern $\phi_1 = 1.6$, $\phi_2 = -0.8$ und $\sigma^2 = 0.1$ zeigt dieses Spektrum den folgenden Verlauf (vgl. Abb. 7):

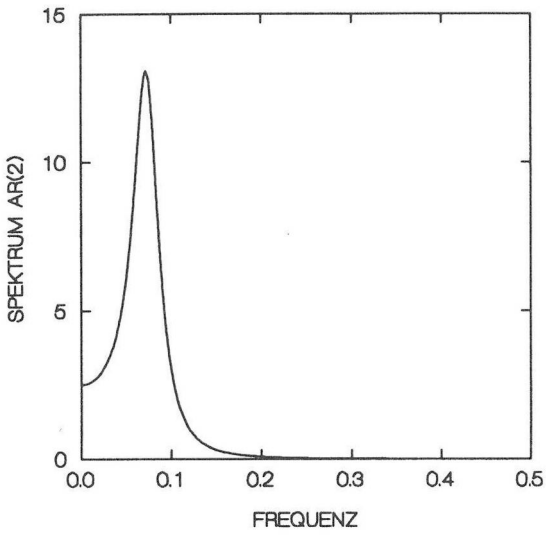


Abb. 7

Dieses Spektrum ist dadurch charakterisiert, dass es eine ausgeprägte Spitze bei der Frequenz f = 0.074 aufweist, d.h. die spektrale Masse ist stark "um" diese Frequenz konzentriert, alle anderen Frequenzbereiche sind von relativ geringer Bedeutung.

Wählt man als Parameter $\phi_1 = 1.6$, $\phi_2 = -0.99$ und $\sigma^2 = 0.1$, dann erhält man folgendes Spektrum (vgl. Abb. 8):

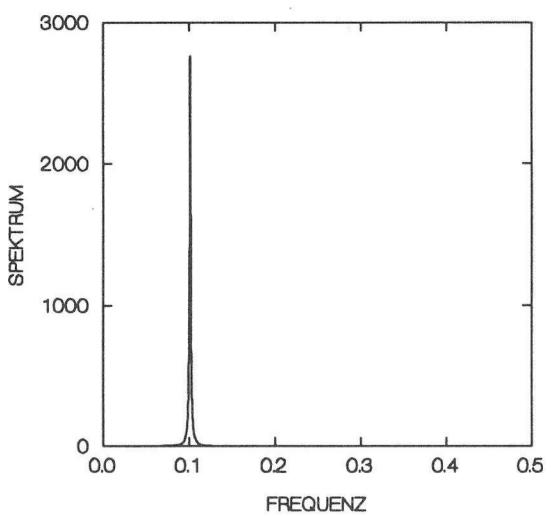


Abb. 8

Dieses Spektrum unterscheidet sich vom vorhergehenden dadurch, dass praktisch die gesamte spektrale Masse bei der Frequenz f = 0.10 nahezu vollständig konzentriert ist (man beachte die unterschiedliche Skalierung in Abb. 7 und 8). Etwas vereinfacht ausgedrückt könnte man sagen, dass das Spektrum in allen Frequenzen gleich Null ist, mit Ausnahme der Frequenz f = 0.10, in der es einen sehr grossen Wert (ca. 2500) annimmt.

Die Abb. 9 und 10 zeigen nun je eine Realisation dieser beiden AR(2)-Prozesse.

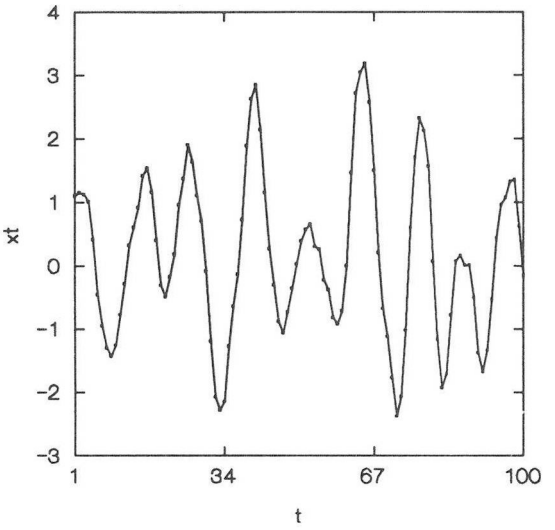


Abb. 9

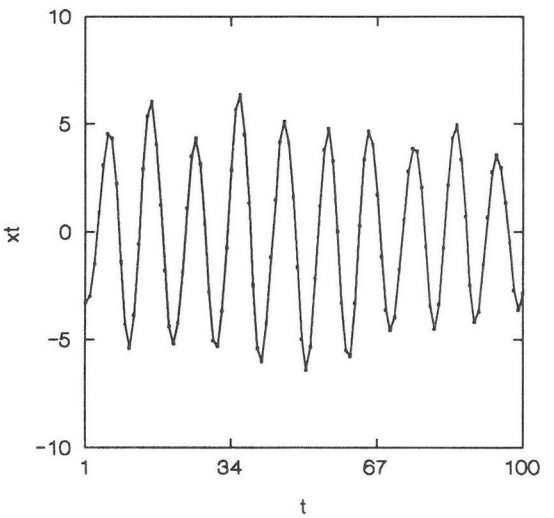


Abb. 10

Die Realisation des ersten Prozesses zeigt einen quasi-periodischen Verlauf mit einer Periode von ca. 13 ZE, diejenige des zweiten hingegen einen nahezu exakt-periodischen Verlauf mit einer Periode von 10 ZE.

Diese Beispiele zeigen, dass ausgeprägte Spitzen in einem Spektrum quasi-periodische Bewegungen in einer Zeitreihe induzieren. Je "enger" und stärker ausgeprägt solche Spitzen im Vergleich zu den sonstigen Werten eines Spektrums sind, umso regelmässiger verlaufen derartige Bewegungen.

Die eben dargelegten Eigenschaften von Spektren sind für die saisonale Komponente von Zeitreihen von Bedeutung, weil es sich bei saisonalen Bewegungen um quasi-periodische Vorgänge handelt. Somit ist zu erwarten, dass eine saisonale Komponente einer Reihe in ihrem Spektrum Spitzen "um" gewisse Frequenzen erzeugt. Diese Frequenzen werden als *saisonale Frequenzen* bezeichnet. Bei Monatsreihen sind dies die Frequenzen 1/12, 2/12, 3/12, 4/12, 5/12 und 6/12, wobei die erste Frequenz in der Regel dominiert. Die mit den Frequenzen 2/12 - 6/12 korrespon-

dierenden Schwingungen werden als *Oberwellen* bezeichnet. Je nach Ausgestaltung einer Saisonfigur sind diese mehr oder weniger ausgeprägt, können aber auch ganz fehlen. Bei Quartalsreihen sind 1/4 und 1/2 die entsprechenden saisonalen Frequenzen.

3.3 *Schätzung eines Spektrums*

Bisher wurde davon ausgegangen, dass der erzeugende Prozess einer Zeitreihe bekannt ist. Wie die Beispiele im letzten Abschnitt gezeigt haben, lässt sich das Spektrum eines bekannten stationären Prozesses *berechnen*. In der Praxis befindet man sich jedoch in einer wesentlich ungünstigeren Situation. Der erzeugende Prozess ist unbekannt. Bekannt ist lediglich *eine* Prozessrealisation, eben die vorliegende Zeitreihe, die zudem immer endlich, nicht selten sogar ziemlich kurz ist. Die Zeitreihenwerte $x_1, x_2,\ldots, x_T$ werden dann als Realisationen der Zufallsvariablen $X_1, X_2, \ldots, X_T$ aufgefasst.

Es stellt sich deshalb das Problem, mit Hilfe dieser Zeitreihe das Spekrum zu *schätzen*. Eine naheliegende Möglichkeit der Schätzung ergibt sich aus der in Abschnitt 3.2 eingeführten Definition eines Spektrums, wenn man dort die unbekannte Kovarianzfunktion durch eine geeignet *geschätzte* Kovarianzfunktion ersetzt. Ein solcher Schätzer ist gegeben durch

$$\hat{\gamma}_h = \frac{1}{T}\sum_{t=1}^{T-h}(X_t-\bar{X}_T)(X_{t+h}-\bar{X}_T), \quad h = 0,1,2,\ldots ,$$

wobei

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T}X_t$$

ein Schätzer für den Erwartungswert μ und

$$\hat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T}(X_t-\bar{X}_T)^2$$

ein Schätzer für die Varianz σ^2 des Prozesses ist.

Da die unbekannte Kovarianzfunktion nicht für unendlich viele lags h geschätzt werden kann, muss obige unendliche Summe durch eine endliche ersetzt werden, d.h. man bildet sie bis zu einem maximalen lag von M, der noch speziell festzulegen ist. Somit liegt folgender Schätzer für das unbekannte Spektrum nahe

$$\hat{S}(f) = \hat{\gamma}_0 + 2\sum_{h=1}^{M}\hat{\gamma}_h\cos 2\pi f h \quad .$$

Zu beachten ist dabei, dass das theoretische Spektrum eine kontinuierliche Funktion der Frequenz f ist, aber eine Schätzung natürlich nur an einer endlichen Anzahl von Frequenzpunkten durchgeführt werden kann (z.B. an den 101 Stellen f = 0, 0.005, 0.01, 0.015, ... ,0.495, 0.5).

Diese Schätzfunktion, die auch als *Periodogramm* bekannt ist, hat allerdings einen gravierenden Nachteil; sie ist zwar (asymptotisch) erwartungstreu, aber *nicht konsistent*. Letzteres bedeutet, dass die Varianz des Periodogramms nicht gegen Null konvergiert, wenn der Stichprobenumfang unendlich gross wird. Damit erweist sich das Periodogramm prinzipiell als ungeeigneter Schätzer für das Spektrum mit der Konsequenz, dass es zur Auffindung von "Periodizitäten" in Zeitreihen nicht geeignet ist. Allerdings lässt sich durch eine gewisse Modifikation des Periodogramms ein konsistenter Schätzer für das Spektrum ermitteln.

Es kann gezeigt werden, dass die Multiplikation der geschätzten Kovarianzen im Periodogramm mit einer Folge von Gewichten w_h zu einer konsistenten Schätzfunktion des Spektrums führt. Eine derartige Gewichtsfolge ist dadurch ausgezeichnet, dass sie ab einem bestimmten lag h = M (dem *truncation point*) identisch gleich Null wird. Sie wird üblicherweise als *lag-Fenster* bezeichnet. Lag-Fenster können auf verschiedene Weise gebildet werden. Ein einfaches lag-Fenster ist z.B. das BARTLETT-Fenster mit den Gewichten

$$w_h = \begin{cases} 1 - \dfrac{h}{M} & \text{für } h \le M \\ 0 & \text{für } h > M \end{cases} .$$

In der Praxis der Spektralschätzung hat sich weitgehend das PARZEN-Fenster durchgesetzt, das auch den hier durchgeführten Spektralschätzungen zugrundeliegt. Die Gewichte sind dabei

$$w_h = \begin{cases} 1 - 6(\dfrac{h}{M})^2 + 6(\dfrac{h}{M})^3 & \text{für } h \le \dfrac{M}{2} \\ 2(1 - \dfrac{h}{M})^3 & \text{für } \dfrac{M}{2} < h \le M \\ 0 & \text{für } h > M \end{cases} .$$

Die verschiedenen Fenster haben spezifische Vor- und Nachteile, die hier aber nicht diskutiert werden können.

Der hier verwendete Schätzer für das Spektrum lautet somit

$$\hat{S}(f) = \hat{\gamma}_0 + 2\sum_{h=1}^{M}w_h\hat{\gamma}_h\cos 2\pi f h ,$$

wobei w_h das PARZEN-Fenster bezeichnet. Dieser Schätzer ist asymptotisch erwartungstreu und konsistent.

Schätzer, die ein lag-Fenster für die Kovarianzfunktion benützen, werden als *indirekte* Spektralschätzer bezeichnet. Betrachtet man ihre Eigenschaften im Frequenzbereich, so kann man zeigen, dass die Schätzung in einem bestimmten Frequenzpunkt $f = f_0$ im Prinzip nichts anderes darstellt als ein gewogenes Mittel über den Spektralwerten von f_0 und den Spektralwerten "benachbarter" Frequenzen von f_0. Diese Mittelungs- oder Glättungseigenschaft von Spektralschätzern kann zu "leakage"-Problemen führen, d.h. die Schätzung in f_0 kann durch sehr grosse (und sehr kleine) benachbarte Spektralwerte "verfälscht" werden. Die spektrale Masse benachbarter Frequenzpunkte "sickert" gewissermassen durch und führt u.U. zu nicht-informativen Spektren. In solchen Fällen muss dafür gesorgt werden, dass diese spektrale Masse vor der Spektralanalyse "ausgefiltert" wird. Dieses Problem spielt beim Nachweis von Saisonalität in ökonomischen Reihen eine wichtige Rolle (vgl. dazu die entsprechenden Ausführungen im nächsten Abschnitt 3.4).

Ein spezielles Problem ist in der Praxis die Festlegung des *truncation point M*. Es kann gezeigt werden, dass das *Auflösungsvermögen* des Spektralschätzers umso besser ist, je grösser M ist, d.h. dicht beieinanderliegende Spitzen im theoretischen Spektrum werden auch im geschätzten Spektrum getrennt und nicht als einzige Spitze ausgewiesen. Anders ausgedrückt, der bias einer Schätzung ist umso kleiner, je grösser M ist. (Dieser bei endlicher Stichprobe stets vorhandene bias führt dazu, dass "Gipfel" im Spektrum tendenziell unterschätzt, "Täler" aber tendenziell überschätzt werden). Allerdings ist aber die Varianz des Schätzers umso grösser, je grösser der *truncation point* gewählt wird, d.h. die Schätzung ist umso instabiler, je grösser M ist. Somit besteht eine Antinomie zwischen den wünschenswerten Schätzeigenschaften *Biasreduktion* und *Varianzreduktion*. Die Erfahrung zeigt, dass bei einer Wahl von M etwa im Bereich von T/5 bis T/3 ein relativ guter Kompromiss zwischen diesen beiden divergierenden Eigenschaften erzielt werden kann.

In Abb. 11 ist das geschätzte Spektrum der Zeiteihe aus Abb. 10 wiedergegeben, wobei ein PARZEN-Fenster mit M = 20 gewählt wurde.

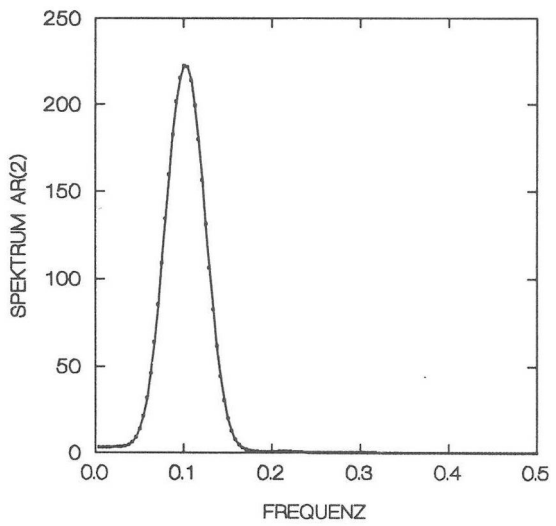


Abb. 11

3.4 Spektralanalyse zur Aufdeckung saisonaler Bewegungen

Nachdem die stochastischen und schätztheoretischen Grundlagen der Spektralanalyse dargelegt sind, kann der Einsatz dieses Werkzeuges zur Feststellung der Saisonalität von Zeitreihen demonstriert werden. Dabei soll aus speziellen Gründen, die im Verlauf der Darstellung einsichtig werden, zunächst grundsätzlich mit simulierten Reihen gearbeitet werden, also mit Reihen, deren "Informationsgehalt" sowohl im Zeit- als auch im Frequenzbereich bekannt ist.

Begonnen werden soll mit einer sehr einfachen Reihe, nämlich einer Sinusschwingung mit einer Periode von 12 ZE der Länge T = 100, wie sie in Abb. 1 wiedergegeben ist.

Diese Reihe stellt gewissermassen eine Modell-Saisonfigur dar. Sie ist exakt periodisch, vollständig deterministisch und enthält keine weiteren Komponenten. Für eine solche Reihe *unendlicher Länge* existiert allerdings aus gewissen mathematischen Gründen kein Spektrum im bisher betrachteten Sinne. Durch Einführung der sogenannten *Delta-Funktion* - das ist eine "Funktion", die überall gleich Null ist und in einem einzigen Punkt ihres Definitionsbereiches unendlich gross wird - kann man jedoch wiederum ein Spektrum definieren. Dieses stellt nämlich für eine unendlich lange Sinusschwingung der Periode 12 eine Delta-Funktion dar, d.h. es ist gleich Null in allen Frequenzen mit Ausnahme der Frequenz f = 1/12, wo es einen unendlichen grossen Wert annimmt. Die ganze spektrale Masse ist sozusagen in dieser Frequenz punktuell konzentriert. Einzelheiten können und sollen hier nicht dargelegt werden. Für eine *endliche* Sinusschwingung hingegen ist die spektrale Masse zwar auch in einem Frequenzpunkt konzentriert, aber nicht ausschliesslich, und man kann das Spektrum einer solchen Schwingung mit den oben diskutierten Methoden schätzen. Zu erwarten ist dabei ein geschätztes Spektrum, dessen spektrale Masse sich symmetrisch um die Frequenz 1/12 verteilt. Dies zeigt sich auch in Abb. 12.

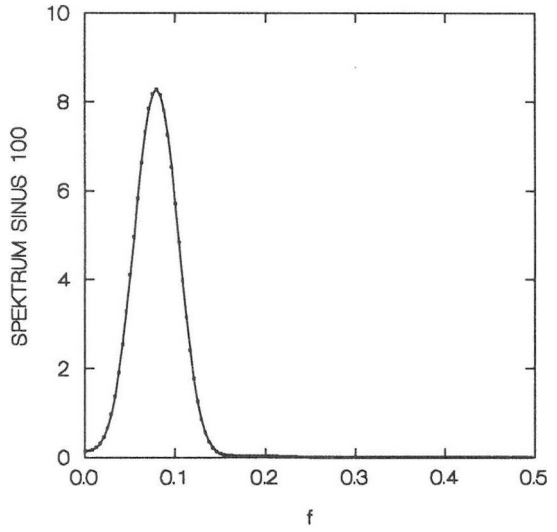


Abb. 12

Betrachtet man dieses Spektrum ausschnittsweise um die Frequenz 1/12, dann zeigt sich deutlich, dass der "peak" bei 12 ZE liegt (vgl. Abb. 13):

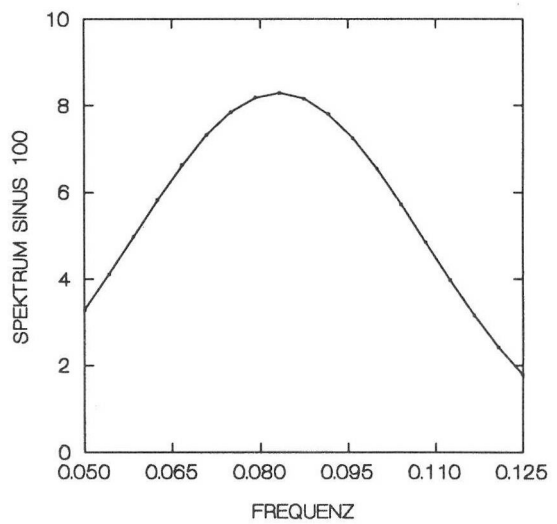


Abb. 13

Verlängert man dieses Sinus-Signal sukzessive und schätzt jeweils das Spektrum, dann ist zu erwarten, dass die Ordinate des Spektrums bei f = 1/12 immer grösser wird und sich die spektrale Masse immer mehr um diese Frequenz konzentriert. Die Abb. 14 zeigt die geschätzten Spektren für T = 100, 200, 500, 1000, 5000.

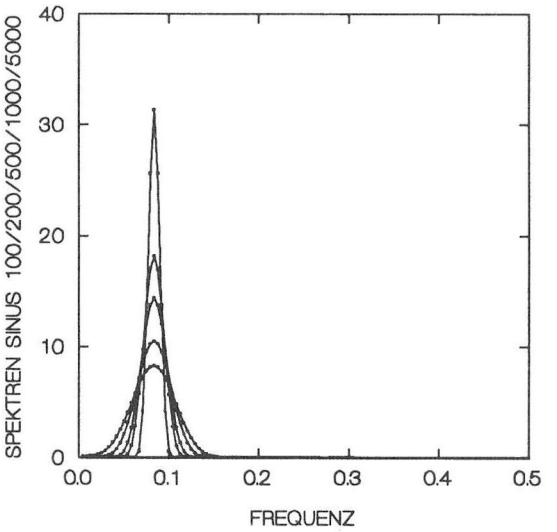


Abb. 14

Offensichtlich strebt dieses Spektrum gegen die oben erwähnte Delta-Funktion.

Als nächstes soll eine zusätzliche Komponente in die bisher betrachtete einfache Zeitreihe eingeführt werden. Dabei handelt es sich um weisses Rauschen, wobei drei verschiedene Fälle betrachtet werden, die sich durch unterschiedlich grosse Varianzen $\sigma^2 = 0.1$, 1.0 und 2.0 unterscheiden. Die drei noise-Komponenten sind in den Abb. 15-17 wiedergegeben.

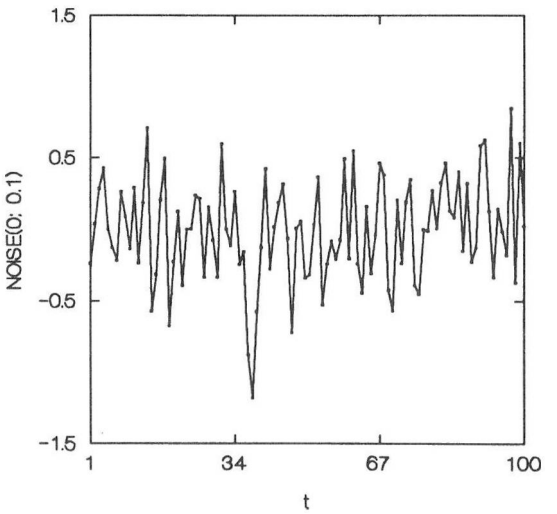


Abb. 15

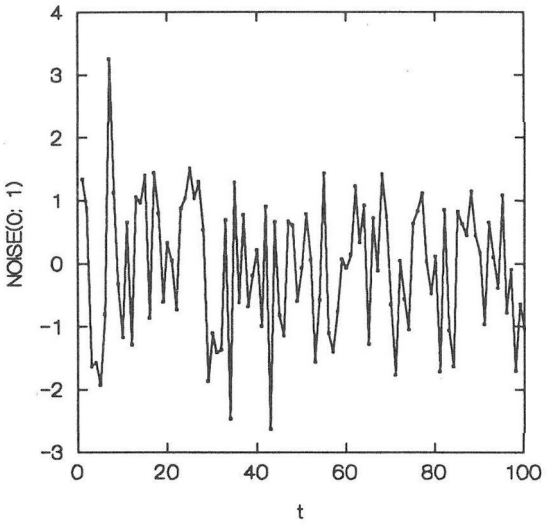


Abb. 16

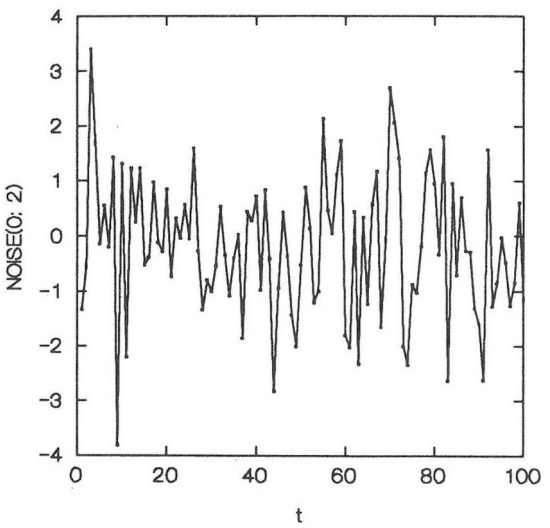


Abb. 17

Addiert man diese noise-Komponenten zum obigen Sinus-Signal, dann ergeben sich folgende Reihen (vgl. Abb. 18, 19 und 20):

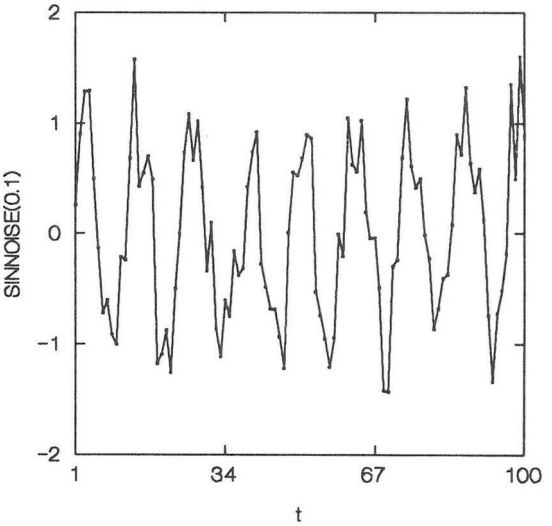


Abb. 18

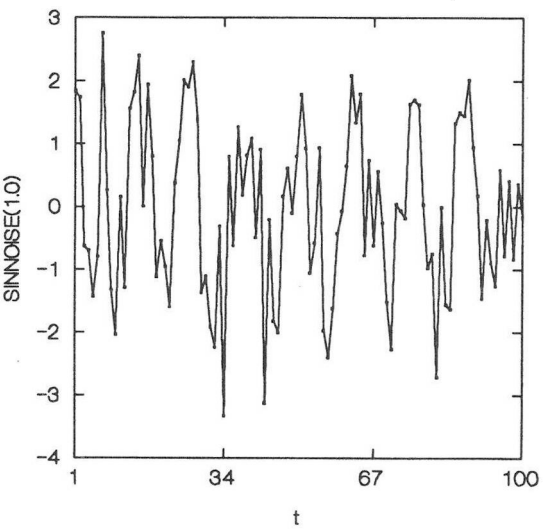


Abb. 19

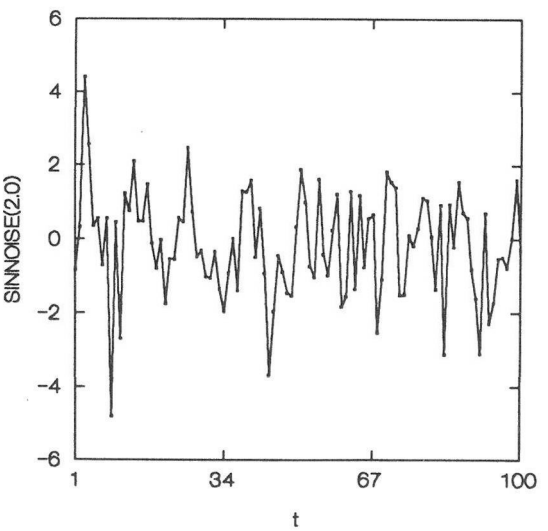


Abb. 20

Während in Abb. 18 die sinusoidale Saisonkomponente trotz der Störung noch klar erkennbar ist, verwischt die grössere Varianz der noise-Komponente in Abb. 19 die regelmässige Saisonfigur in erheblichem Ausmass. In Abb. 20 schliesslich ist diese optisch gar nicht mehr zu erkennen.

Schätzt man die Spektren dieser drei Reihen, so ergeben sich die folgenden Resultate (vgl. Abb. 21-23):

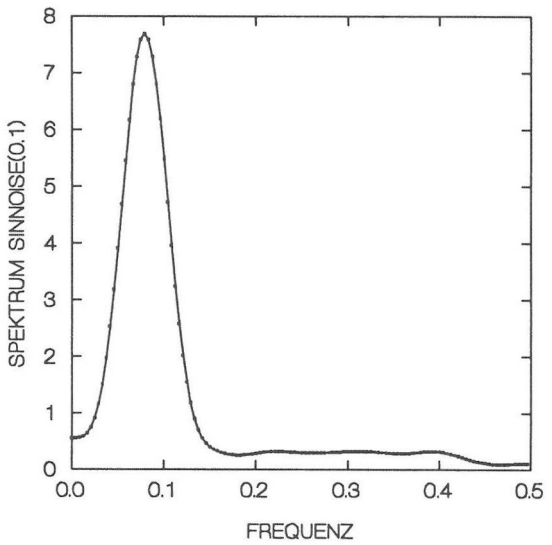


Abb. 21

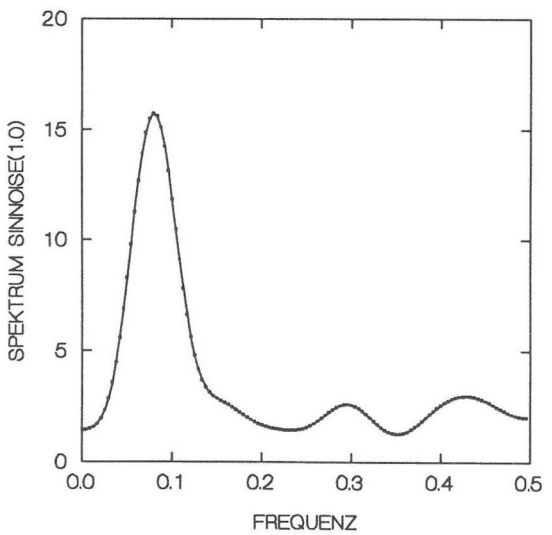


Abb. 22

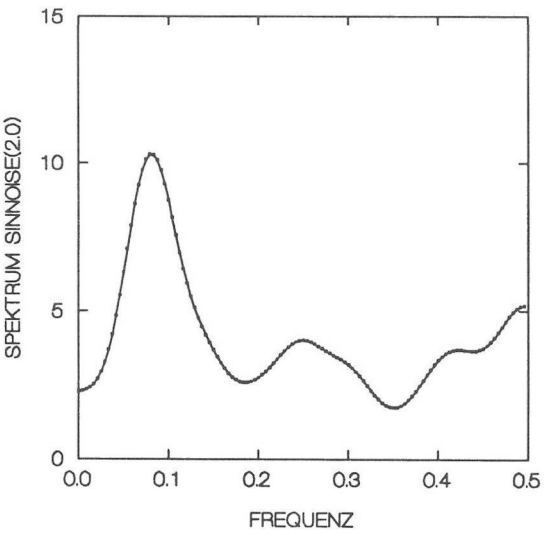


Abb. 23

In allen drei Fällen zeigt das jeweils geschätzte Spektrum eindeutig, dass eine saisonale Komponente vorhanden ist. Es zeigt sich aber auch spektrale Masse im Hochfrequenzbereich, und zwar umso mehr, je grösser die noise-Varianz ist. Aber auch im Niederfrequenzbereich und bei der saisonalen Frequenz 1/12 wird die spektrale Masse grösser. Letzteres kommt daher, dass sich das Spektrum der nunmehr aus zwei Komponenten bestehenden Reihe additiv aus den beiden Teilspektren der Sinuswelle und der noise-Komponenten zusammensetzt. Die Abb. 24, welche die drei Spektren zusammen zeigt, macht diesen Sachverhalt deutlich.

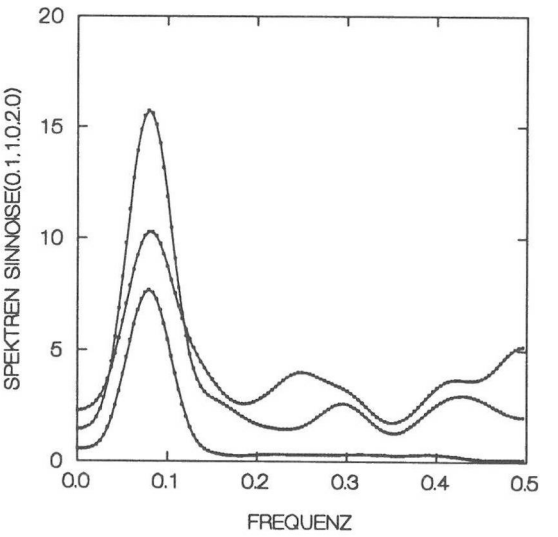


Abb. 24

Bekanntlich weisen ökonomische Zeitreihen häufig einen (meist positiven) Trend auf. Als nächstes bleibt deshalb zu untersuchen, welche Informationen eine Spektralanalyse liefert, wenn neben den bisherigen Komponenten ein Trend miteinbezogen wird. Zunächst sei der Einfachheit halber angenommen, dass die Reihe nur aus zwei Komponenten

bestehe, der bisherigen sinusoidalen Saisonfigur und einem linearen Trend

$$X_t = 2 + 0.9t + \sin(2\pi/12)t \ , \ t = 1,...,100 \ .$$

Wie aus der Abb. 25 hervorgeht, kann man trotz der das Reihenbild dominierenden Trendentwicklung die saisonale Komponente noch sehr gut erkennen.

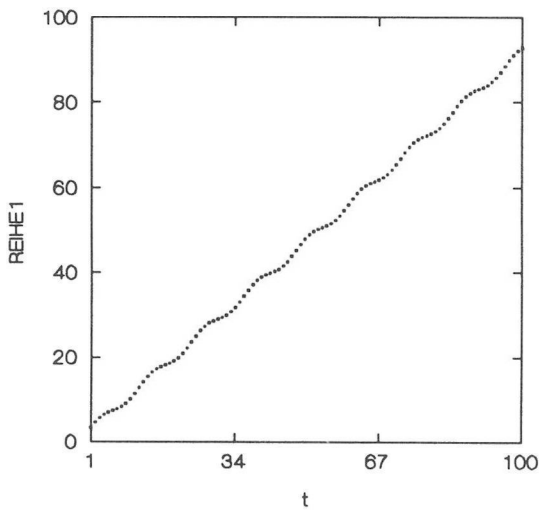


Abb. 25

Umso erstaunlicher ist es, dass das geschätzte Spektrum dieser Reihe folgenden Verlauf zeigt (vgl. Abb. 26):

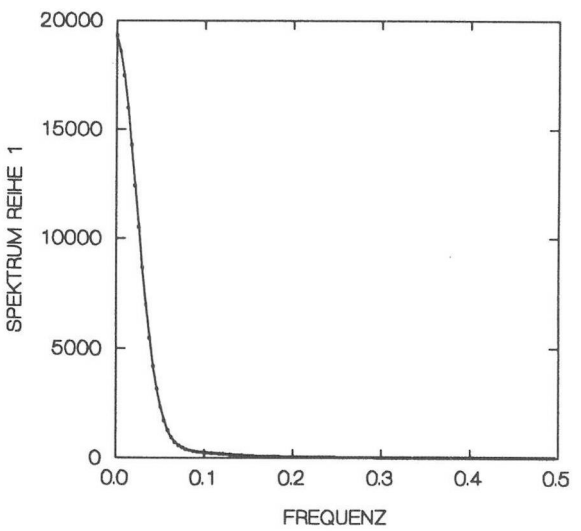


Abb. 26

Dieses Spektrum gibt keinerlei Hinweis auf die Existenz einer Saisonkomponente. Dies ändert sich auch nicht, wenn man die (hier nicht wiedergegebene) logarithmierte Form des Spektrums betrachtet, die gelegentlich dem Original-Spektrum vorzuziehen ist, weil dann feinere Details bei dieser Darstellung evidenter werden. Offensichtlich ist die gesamte spektrale Masse im Niederfrequenzbereich konzentriert, vor allem um die Frequenz f = 0. Wie ist dies zu erklären? Die Ursache dafür liegt darin, dass die Reihe einen ausgeprägten Trend aufweist. Ein Trend kann aber als Schwingung mit einer unendlich langen Periode ausgefasst werden, oder anders ausgedrückt, seine spektrale Masse ist der Frequenz f = 0 zuzuordnen. Dies hat erhebliche Konsequenzen für die Spektralschätzung, die schon in Abschnitt 3.3. unter der Bezeichnung "leakage"-Problem angedeutet wurden. Die bei f = 0 vom Trend verursachte Konzentration spektraler Masse "sickert" bei der Spektralschätzung benachbarter Frequenzen durch, so dass die (eigentlich vorhandene) Spitze bei f = 1/12 nicht mehr zum Vorschein kommt.

Man kann das Problem aber auch noch von einer anderen Seite aus betrachten. Gemäss den theoretischen Grundlagen der Spektralanalyse ist diese nur auf Realisationen stationärer Prozesse anwendbar. Solche Prozesse sind aber u.a. dadurch gekennzeichnet, dass sie einen konstanten Erwartungswert haben. Demzufolge schwanken auch die Realisationen dieser Prozesse um ein konstantes Niveau. Reihen mit Trend können daher nicht mehr als Realisationen stationärer Prozesse angesehen werden. Somit kann die Spektralanalyse auf solche Reihen nicht angewendet werden, es sei denn, es werde zuvor die Ursache der Nicht-Stationarität beseitigt. Konkret heisst das hier, dass der Trend vor der Spektralanalyse eliminiert werden muss. Damit stellt sich aber ein neues Problem. Auf welche Weise sollen Trends eliminiert werden?

Man kann sich leicht überlegen, dass es dafür keine einfachen ad-hoc-Lösungen gibt. Die naheliegende Idee etwa, für den Trend ein Polynom der Zeit anzusetzen und dessen unbekannte Koeffizienten mit Hilfe der Methode der kleinsten Quadrate zu schätzen, ist aus mehreren Gründen problematisch. Bei praktischen Reihen ist a priori nicht sicher, ob der Trend überhaupt einem Polynom folgt, und falls er das tut, ist der Polynomgrad in der Regel unbekannt, er muss also irgendwie mehr oder weniger willkürlich festgelegt werden. Ausserdem sind Kleinst-Quadrate-Schätzer gerade bei Zeitreihendaten häufig problematisch, weil gewisse schätztechnischen Voraussetzungen bei diesen Daten meistens nicht erfüllt sind. Nicht selten wird in der Praxis versucht, einen Trend durch Bildung 1. Differenzen zu eliminieren. Es wird weiter im Detail zu zeigen sein, dass diese Vorgehensweise jedoch mit *gravierenden Nachteilen* verbunden ist. (Vgl. dazu die Ausführungen in Kapitel II, Abschnitt 5.2.)

Generell lässt sich folgendes Postulat formulieren. Ein Trend sollte so eliminiert werden, dass alle übrigen Komponenten einer Reihe davon nicht tangiert werden. Im Frequenzbereich kann man dieses Postulat wie folgt formulieren. Bei der Trendelimination sollten nur diejenigen Frequenzen tangiert werden, die dem Trend zuzuordnen sind (das sind die Frequenzen f = 0 und "unmittelbar" benachbarte), alle anderen sollten unverändert belassen werden. Diesem Postulat genügt nur ein "guter" Hochpass-Filter. (1. Differenzen sind zwar auch Hochpass-Filter,

allerdings sehr schlechte, wie noch zu zeigen sein wird.) Es ist hier nicht möglich, die Konstruktion solcher Filter definiert darzulegen, vielmehr soll eine Demonstration der mit ihnen erzielten Resultate genügen. (Vgl. dazu Kapitel II, Abschnitte 4 und 7.)

Von einem guten Hochpass-Filter ist zu erwarten, dass der Filter-Output der Reihe in Abb. 25 sich kaum von der sinusoidalen Saisonkomponente unterscheidet. Die Abb. 27 zeigt deshalb beide Reihen zusammen.

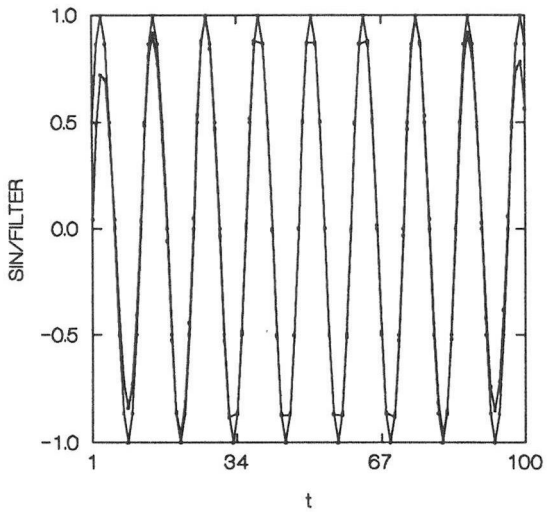


Abb. 27

Offensichtlich unterscheiden sich beide Reihen nur geringfügig am Reihenanfang bzw. Reihenende.

In Abb. 28 sind die Spektren des Filter-Outputs und der Sinusschwingung zusammen dargestellt. Die beiden Spektren sind praktisch deckungsgleich.

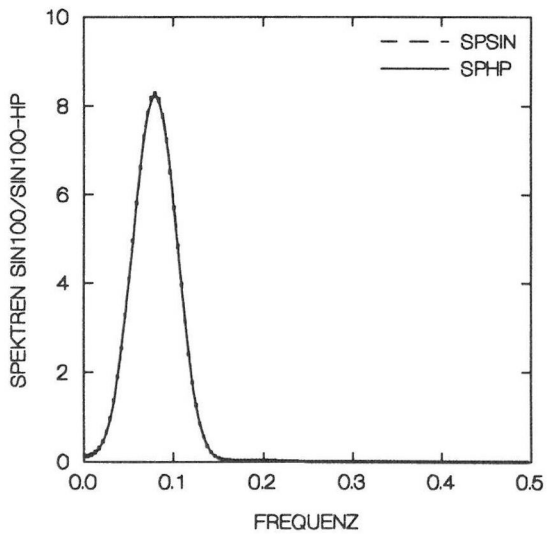


Abb. 28

Es wird später gezeigt, dass die Bildung 1. Differenzen hier zwar auch diesen linearen Trend eliminiert, aber gleichzeitig zu erheblichen Verzerrungen der Reihe führt.

Zu der bisher betrachteten trendbehafteten Reihe wird jetzt jeweils eine noise-Komponente aus den Abb. 15-17 hinzugefügt. Das Reihenmodell lautet nun

$$X_t = 2 + 0.9t + \sin(2\pi/12) + \varepsilon_t \ , \ t = 1,2,...,100 \ .$$

Wie aus den Abb. 29-31 hervorgeht, ist die saisonale Komponente jetzt kaum mehr bzw. bei den beiden grösseren noise-Varianzen gar nicht mehr erkennbar.

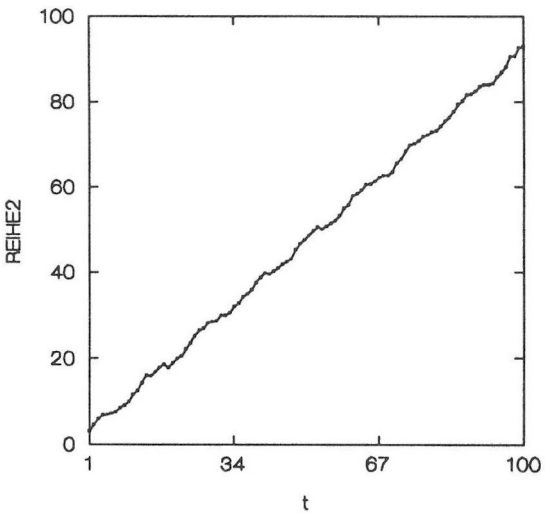


Abb. 29

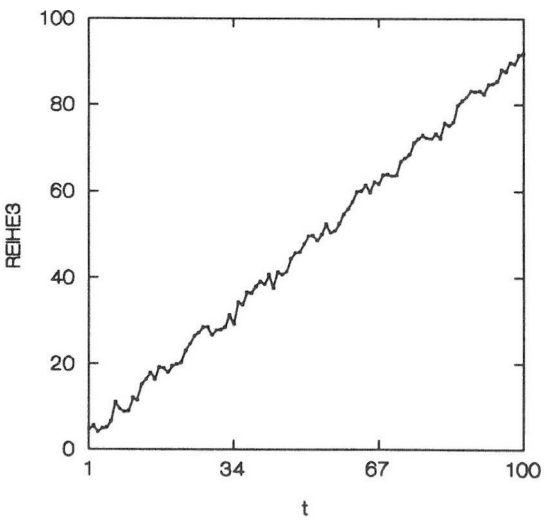


Abb. 30

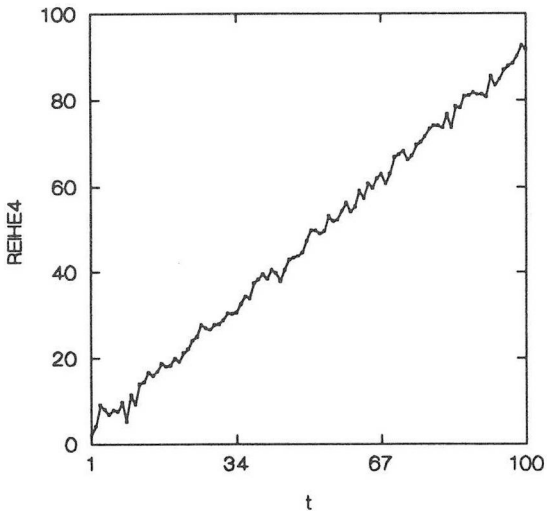


Abb. 31

Auch die (hier nicht wiedergegebenen) Spektren dieser Reihen zeigen natürlich keine saisonale Komponente. Um informative Spektren zu erhalten, muss zuerst wieder eine Hochpass-Filterung durchgeführt werden. In den Abb. 32-34 sind jeweils Filter-Output und (Saison+noise)-Komponente zusammen dargestellt. Auch hier zeigen sich wiederum nur sehr geringe Unterschiede.

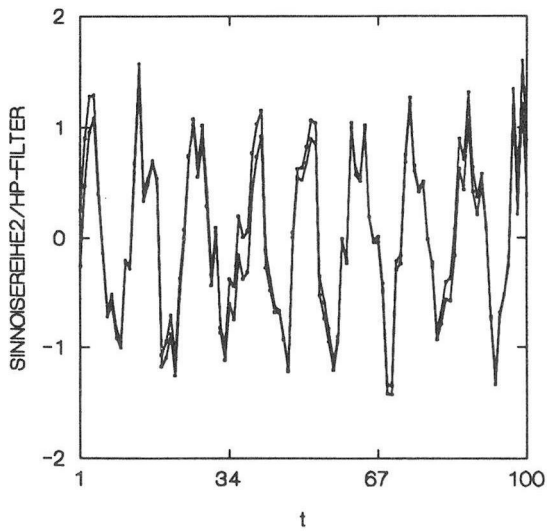


Abb. 32

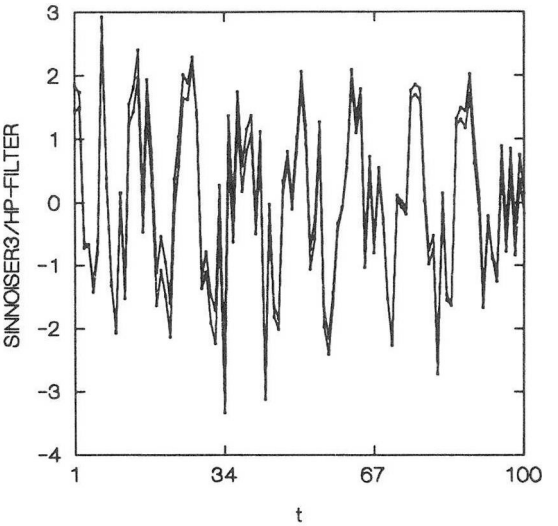


Abb. 33

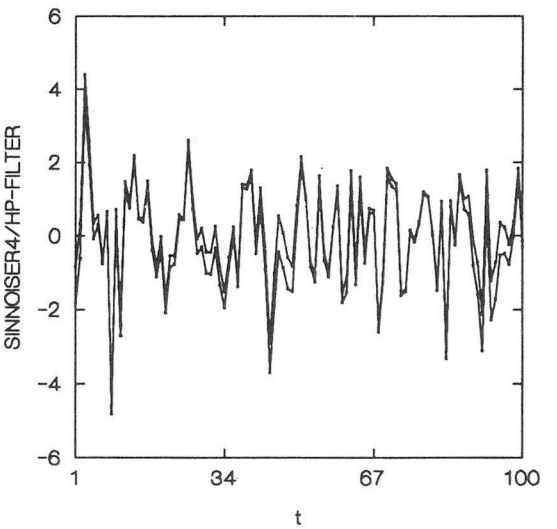


Abb. 34

Deshalb sind jeweils die Spektren von Filter-Output und (Saison+noise)-Komponente - wie bereits festgestellt - nur geringfügig voneinander verschieden. Auf ihre gemeinsame Darstellung sei deshalb verzichtet.

An den bisherigen Resultaten ändert sich nichts, wenn Reihen mit einer komplizierteren Komponentenstruktur zugrundegelegt werden, z.B. einer Saisonfigur, die auch die Oberwellen miteinbezieht, oder wenn ein autokorrelierter Störprozess unterstellt wird. Weitere Beispiele in dieser Richtung erübrigen sich daher.

Die durchgeführte Spektralanalyse simulierter Reihen lässt somit folgenden Schluss zu.

Zur Feststellung von Saisonalität in trendbehafteten ökonomischen Reihen ist in der Regel vorgängig eine Trendbereinigung erforderlich. Diese muss mit einem geeigneten Hochpass-Filter vorgenommen werden.

Aus diesem Grund basieren die im Abschnitt 4.2 präsentierten alle Spektren der zu untersuchenden Preisreihen auf hochpassgefilterten Reihen.

4. Analyse der Zeitreihen für die Preisnotierungen

4.1 Beschreibung der untersuchten Zeitreihen

Vom Bundesamt für Statistik, Abteilung Volkswirtschaft und Preise - Sektion Preise und Verbrauch, wurden die Daten des Landesindex der Konsumentenpreise für die Basen September 1966, September 1977 und Dezember 1982 sowie die Daten des Grosshandelspreisindex zur Verfügung gestellt. Die erläuterten Beispiele wurden aufgrund der Vielzahl der Reihen hauptsächlich auf den Landesindex der Konsumentenpreise eingegrenzt. Der Landesindex der Konsumentenpreise zeigt die Preisentwicklung der für die privaten Haushalte bedeutsamen Waren und Dienstleistungen und gibt an, in welchem Umfang die Lebenshaltung infolge von Preisveränderungen, möglichst unbeeinflusst durch Veränderung der Güterqualität und Änderung des Konsumverhaltens, teurer oder billiger geworden ist. Die Datenreihen ab Dezember 1982 bis Ende 1991 werden hier für die Analysen verwendet. Für spezielle Untersuchungen wurde aber vor allem der Totalindex der Konsumentenpreise bis August 1992 benutzt.

Bei den folgenden in den Darstellungen verwendeten Zeitreihen handelt es sich um die Hauptbedarfsgruppen des Landesindex der Konsumentenpreise. Bei den Hauptbedarfsgruppen der Nahrungsmittel und der Waren werden auch einige Unterbedarfsgruppen näher analysiert.

Folgende Zeitreihen werden bei den Untersuchungen vor allem betrachtet:

Abkürzung:	*Bedarfsgruppe:*	*Bezeichnung:*
PK	Totalindex	Totalind
PKA	Ausland	Ausland
PKB	Heizung und Beleuchtung	Heiz
PKD	Dienstleistungen	Dienstl
PKE	Bildung und Erholung	Bildung
PKG	Körper- und Gesundheitspflege	Pflege
PKH	Haushalteinrichtung, -unterhalt	Haush
PKI	Inland	Inland
PKK	Bekleidung	Kleid
PKM	Wohnungsmiete	Miete
PKN	Nahrungsmittel	Nahrung
PKNC	Kaffee	Kaffee
PKNE	Eier	Eier
PKNF	Fleisch und Fleischwaren	Fleisch
PKNG	Gemüse	Gemüse
PKNH	Kartoffeln	Karto
PKNM	Milch	Milch
PKNO	Früchte	Frücht
PKNP	Fische	Fisch
PKNT	Getreideprodukte	Getreid
PKT	Getränke und Tabakwaren	Getraenk
PKV	Verkehr	Verkehr
PKW	Waren	Waren
PKWD	dauerhafte	Warend
PKWN	nicht dauerhafte	Warennd

4.2 Empirische Befunde

Für die zur Verfügung stehenden Zeitreihen mit Preisnotierungen wurden Spektralanalysen zur Aufdeckung saisonaler Komponenten durchgeführt. Dazu wurden die Zeitreihen zunächst mit einem Hochpass-Filter [2] trendbereinigt und danach die Spektren geschätzt. Um etwaige Einflüsse der Hochpass-Filterung aufzudecken, wurden die Analysen auch noch mit einem alternativen Hochpass-Filter durchgeführt [3]. Die nachfolgend wiedergegebenen Ergebnisse der Spektralschätzungen bleiben davon unberührt. Aus Gründen der Übersichtlichkeit werden hier nur die geschätzten Spektren von einigen beispielhaft ausgewählten Zeitreihen wiedergegeben. Zur Illustration seien zunächst die Zeitreihe des Landesindex der Konsumentenpreise (TOTALIND), das zugehörige geschätzte Spektrum, die mit dem Hochpass-Filter bereinigte Zeitreihe und das zugehörige geschätzte Spektrum wiedergegeben (vgl. Abb. 35-38). Das Spektrum weist für diese Zeitreihe eine starke saisonale Komponente mit einer Zykluslänge von 12 Monaten aus.

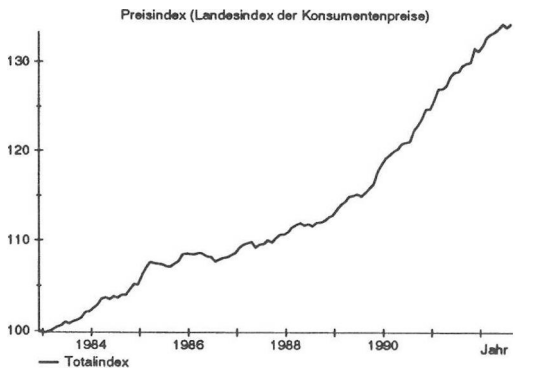


Abb. 35

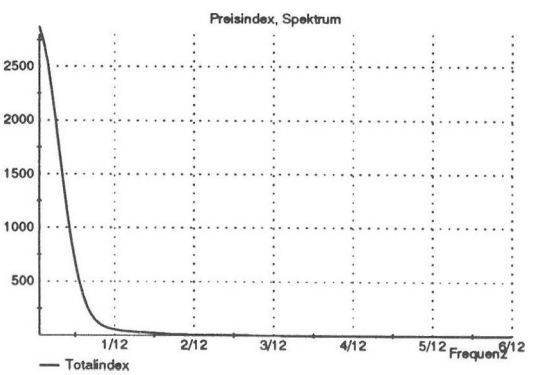


Abb. 36

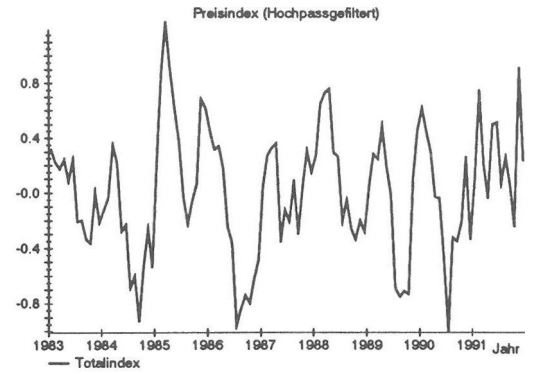


Abb. 37

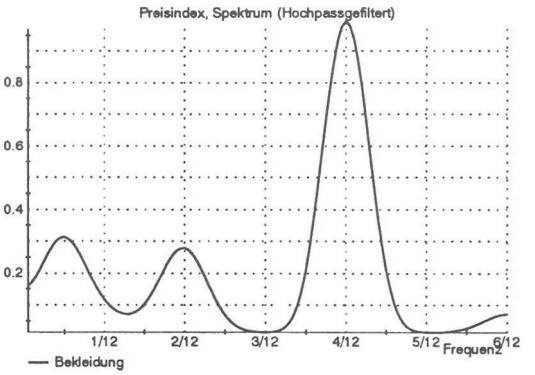


Abb. 40

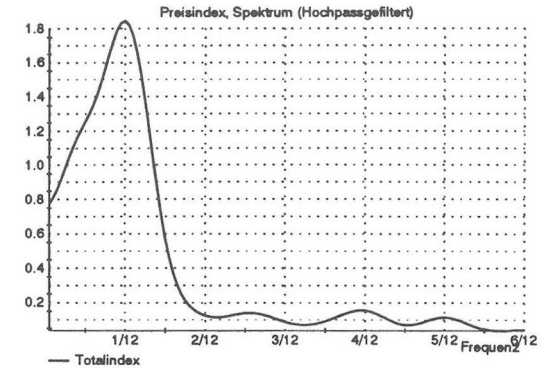


Abb. 38

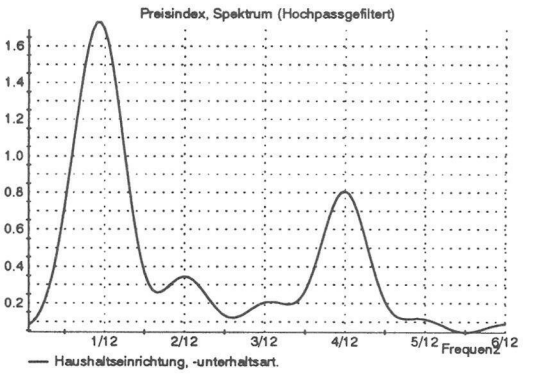


Abb. 41

Für einzelne Teilindizes gibt es - teils erhebungs-, teils aufbereitungsbedingt - durchaus davon abweichende Resultate. Der Teilindex Dienstleistungen (DIENSTL) hat neben der saisonalen Komponente mit der Zykluslänge von 12 Monaten eine starke zusätzliche Komponente mit einer Zykluslänge von 3 Monaten (vgl. Abb. 39), der Teilindex Bekleidng (KLEID) praktisch nur eine bedeutende saisonale Komponente mit einer Zykluslänge von 3 Monaten (vgl. Abb. 40), und der Teilindex für Haushalteinrichtungen und -unterhalt (HAUSH) entspricht bezüglich der saisonalen Komponente der Zeitreihe für den Teilindex Dienstleistungen (vgl. Abb. 41). Der Teilindex für die Wohnungsmiete (MIETE) hat erhebungsbedingt eine saisonale Komponente mit einer Zykluslänge von 6 Monaten (vgl. Abb. 42).

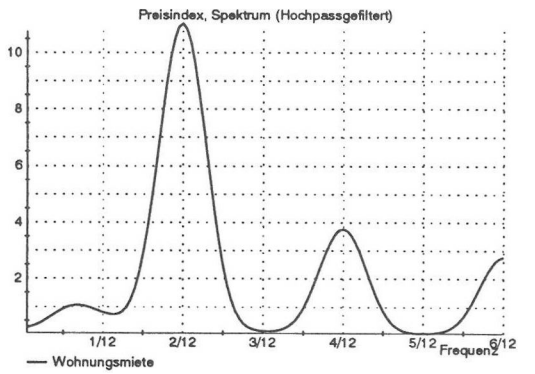


Abb. 42

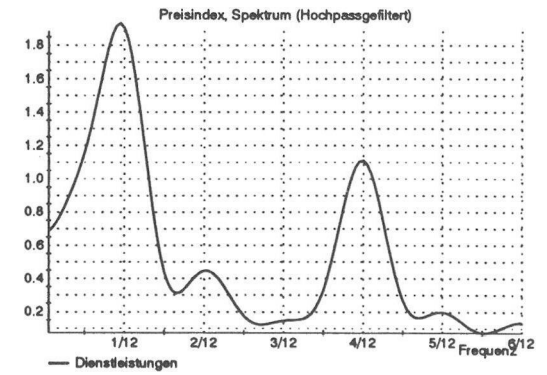


Abb. 39

In den Abb. 43-45 sind die geschätzten Spektren der hochpassgefilterten Zeitreihen für die Teilindizes "Bildung und Erholung" (BILDUNG), "dauerhafte Waren" (WAREND) und "Nahrungsmittel, Getränke und Tabakwaren" (GETRAENK) wiedergegeben. Diese Zeitreihen zeichnen sich dadurch aus, dass sie zyklische Komponenten unterschiedlicher Länge aufweisen.

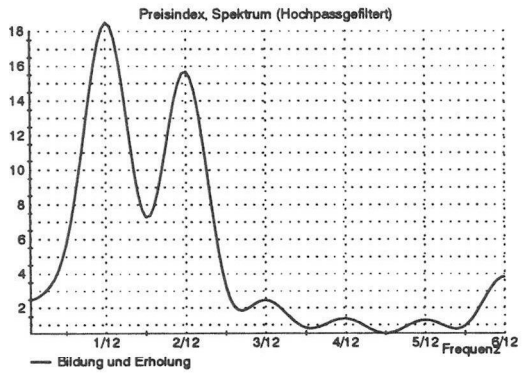


Abb. 43

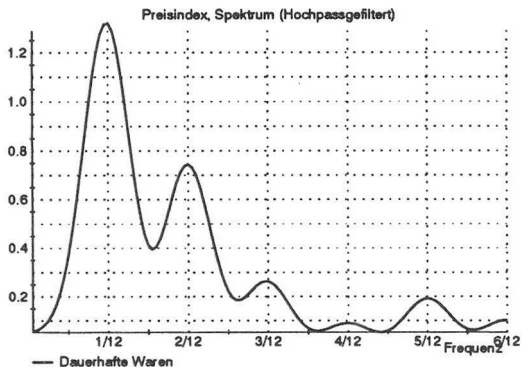


Abb. 44

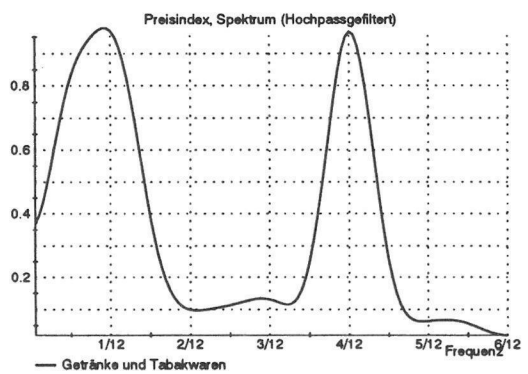


Abb. 45

Literaturhinweise

1) Vgl. dazu z.B. GRANGER C.W.J., HATANAKA M., Spectral Analysis of Economic Time Series, Princeton 1964

2) Vgl. SCHMIDT R., Konstruktion von Digitalfiltern und ihre Verwendung bei der Analyse ökonomischer Zeitreihen, Bochum 1984

3) Verwendet wurde als Alternative ein sogenannter POTTER-Filter. Vgl. OTNES R.K. et al., Applied Time Series Analysis, Vol. I, New York 1978, S. 167-171

II. Zur Konstruktion von Saisonbereinigungsverfahren

1. Zielsetzungen bei einer Saisonbereinigung

Bereits die frühen Ansätze zur Analyse von Zeitreihen ökonomischer Daten gingen von der Vorstellung aus, dass sich diese Zeitreihendaten aus unterschiedlichen Bewegungskomponenten zusammensetzen. Angeregt wurden diese Überlegungen von den Erfahrungen mit der Zerlegung von Zeitreihen in einzelne Bewegungskomponenten auf den Gebieten der Astronomie und der Meteorologie [1]. Bei ökonomischen Zeitreihen wird dabei üblicherweise von der Existenz einer Trendkomponente neben zyklischen, saisonalen und sogenannten irregulären Schwankungen ausgegangen [2]. Trotz der weiten Verbreitung dieses Komponentenansatzes gibt es keine Übereinstimmung in bezug auf die Definition der Komponenten und die Fragen ihrer gegenseitigen Verknüpfung. Dies ist auch nicht weiter erstaunlich, denn es kann über das Modell für eine ökonomische Zeitreihe auch gar keine übereinstimmende Vorstellung geben, denn die einzelnen Komponenten sind prinzipiell unbeobachtbar. Bereits die Versuche, die verschiedenen Komponenten durch eine Zuordnung von bestimmten, als exogen betrachteten Ursachengruppen zu definieren, scheitern an der Komplexität und Mehrdeutigkeit der Beziehungen. Zur Definition zyklischer Schwankungen z.B. offerieren die unterschiedlichen Konjunkturmodelle aus der ökonomischen Theorie eine breite Palette durchaus nicht immer miteinander vereinbarer Erklärungsansätze. Die sogenannte äussere Definition der Komponenten ökonomischer Zeitreihen ist daher inoperabel [3]. Für praktische Analysen ökonomischer Zeitreihen ist deshalb von einem Erklärungsansatz auszugehen, der die einzelnen Reihenkomponenten in formal eindeutiger Abhängigkeit von den vorliegenden Daten beschreibt. Diese sogenannte innere Definition der Komponenten beinhaltet das Aufstellen von ökonomisch plausiblen Hypothesen über die vermuteten Komponenten, die so konzipiert sind, dass daraus eindeutige Berechnungsvorschriften abgeleitet werden können und der mit der Zeitreihenzerlegung verbundenen Zielsetzung möglichst gut entsprochen wird.

Zielsetzungen und Verfahrenskriterien stehen damit im Zentrum der Ausgangsüberlegungen für eine Zeitreihenanalyse [4]. Bezüglich der Zielsetzungen ist zunächst zwischen retrospektiven und prospektiven Zielen zu unterscheiden. Bei der Retrospektive werden die in Form einer Zeitreihe vorliegenden Beobachtungsdaten als abgeschlossenes, historisches Abbild eines Phänomens betrachtet, das es zu untersuchen gilt, um Aufschluss über die Struktur des Prozesses zu bekommen, der die vorliegenden Daten erzeugt hat. Bei einer prospektiven Zielsetzung wird eine vorliegende Zeitreihe dagegen primär als Ausgangsbasis angesehen, um den Zustand des die Daten erzeugenden Systems im Bereich der letzten Beobachtungswerte zu diagnostizieren und um unter Umständen die künftige Entwicklung des Systems zu prognostizieren.

Lange Zeit stand vor allem eine retrospektiv-orientierte Komponentenzerlegung im Vordergrund des Interesses. Zum Beispiel sind Befunde der Art, dass Konjunkturzyklen keine streng periodischen Schwingungen, sondern Schwankungen mit wechselnder Perioden- oder Zykluslänge und variierender Amplitude sind, Ergebnisse derartig orientierter statistischer Analysen. Vor allem seit den zwanziger Jahren dieses Jahrhunderts richtet sich das Bemühen der Statistiker jedoch im engen Kontext mit der Entwicklung in der ökonomischen Theorie und den jeweiligen wirtschaftspolitischen Fragestellungen auf die kurzfristige Konjunkturdiagnose am aktuellen Rand der verfügbaren Beobachtungsreihen. Eine adäquate statistische Analyse sollte dabei Hinweise auf die momentane wirtschaftliche Situation und den weiteren Verlauf der Konjunktur geben.

Für eine derartige Diagnose des Systemzustands sind vor allem die Erfassung des Trends und der konjunkturellen Bewegungen von Bedeutung. Eine isolierte Erfassung dieser beiden Komponenten ist nicht nur technisch schwierig, sondern aus methodischen Überlegungen praktisch auch nicht durchführbar. Diese beiden Komponenten werden daher in der Regel zu einer Trend-Zyklus-Komponente, überwiegend als glatte Komponente bezeichnet, zusammengefasst. Dieses Konglomerat ist aber auch die bei einer Diagnose des Konjunkturverlaufes interessierende, sachlich relevante Information. Die möglichst gute Herausarbeitung der glatten Komponente unter Ausschaltung der saisonalen und irregulären Schwankungen wird deshalb heute oft als Hauptaufgabe einer ökonomischen Zeitreihenanalyse angesehen. Die Bestimmung einer saisonbereinigten Zeitreihe ist dabei meist nur eine verfahrensbedingte Vorstufe bei der Verfolgung dieser Aufgabenstellung. Aber auch die Ermittlung einer saisonalen Bewegungskomponente kann durchaus eine eigenständige Aufgabenstellung sein, um beispielsweise die Produktion saisonabhängiger Produkte planen zu können.

Die Verfolgung einer prospektiven Zielsetzung setzt in jedem Fall eine sorgfältige retrospektive Untersuchung der Struktur des datenerzeugenden Prozesses voraus. Neben einer Vorstellung über das der Analyse zugrundezulegende Zeitreihenmodell, insbesondere über die stochastische Natur der einzelnen Komponenten, sind dazu Kriterien zur Beurteilung der Analyseverfahren erforderlich.

Auch wenn davon ausgegangen werden darf, dass für eine Mehrzahl der Anwender eine prospektive Zielsetzung, d.h. eine Diagnose des Verlaufs der glatten Komponente am aktuellen Rand und insbesondere deren Wendepunkte in Form lokaler Maxima und Minima, allein von Interesse ist, sollen davon abweichende Positionen hier nicht unerwähnt bleiben. Bei einer prospektiven Zielsetzung ist eine Saisonbereinigung streng genommen nur ein Hilfsmittel zur Bestimmung der glatten Komponente am aktuellen Rand. Weil es in der Regel einfacher ist, die mit einer gewissen Regelmässigkeit sich wiederholenden saisonalen Schwankungen zu eliminieren, als die glatte Komponente an dieser Stelle zu schätzen, werden saisonbereinigte Werte zu

Diagnosezwecke mitherangezogen. Diese Sichtweise ist jedoch nicht unumstritten. Einzelne Anwender argumentieren dahingehend, dass die verschiedenen Komponenten, so insbesondere auch die sogenannte irreguläre Komponente, wesentliche ökonomische Informationen enthielten, die verlorengingen, wenn sich die Betrachtung auf eine Analyse der glatten Komponente am aktuellen Rand beschränken würde. Akzeptiert man eine derartige Vorstellung, dann muss man sich bewusst sein, dass man sich bei der Anwendung von Saisonbereinigungsverfahren, in denen die irreguläre Komponente als weisses Rauschen modelliert ist, aus methodischer Sicht inkonsistent verhält. Weisses Rauschen ist doch äquivalent mit der Vorstellung, dass in einem Prozess keinerlei Informationen mehr enthalten sind. Eine Interpretation des Verlaufs einer bearbeiteten Zeitreihe darf daher nicht ohne Berücksichtigung des unterstellten Zeitreihenmodells erfolgen.

2. Übersicht über die Entwicklung von Saisonbereinigungsverfahren

Die Zerlegung von Zeitreihen basiert also auf der Grundvorstellung, dass eine betrachtete Reihe sich aus mehreren Komponenten zusammensetzt. Die einzelnen Komponenten werden dabei verschiedenen Ursachenkomplexen zugeordnet. Bei Zeitreihen mit unterjährigen Daten, z.B. Monatswerten, werden häufig vier systematische und eine sogenannte irreguläre Komponente unterschieden. Die systematischen Komponenten werden dabei mit Trend, Konjunktur, Saison und Kalendereinfluss bezeichnet. Die irreguläre Komponente steht in der Regel für eine Restkomponente, die nicht spezifischen Ursachen zugeordnet werden kann, und gelegentlich auch für auf Sondereinflüsse zurückzuführende Ausreisser in den Beobachtungsdaten, die zwar häufig zeitlich genau bestimmt werden können, deren quantitatives Ausmass aber höchstens geschätzt werden kann. Das inhärente Problem einer Zerlegung von Zeitreihen ökonomischer Daten besteht nun darin, dass sich aus der ökonomischen Theorie keine direkt operable Definition für die erwähnten Komponenten, insbesondere für die systematischen Komponenten ableiten lässt. Die Zeitreihenanalyse ist deshalb darauf angewiesen, das Zusammenwirken der gedanklich anvisierten Ursachenkomplexe durch formale Modelle zu beschreiben, die dann die Grundlage für eine Schätzung der an und für sich nicht beobachtbaren einzelnen Komponenten bilden. Eine Zerlegung von Zeitreihen in einzelne Komponenten ist daher immer eine statistische Konstruktion und deshalb immer nur eine mehr oder weniger geeignete Approximation der verbal umschriebenen Ursachenkomplexe. Oder anders ausgedrückt, die aufgrund der jeweiligen Modellvorstellung ermittelten Komponenten werden durch die Etikettierung mit den Bezeichnungen für die einzelnen Ursachenkomplexe lediglich semantisch koloriert. Die Frage, ob und inwieweit das für die Zerlegung verwendete Modell für eine betrachtete Zeitreihe adäquat ist, wird mit der Etikettierung der ermittelten Komponenten nicht gelöst. Eine aus methodologischer Sicht befriedigende Beurteilung eines bestimmten Zeitreihenanalyseverfahrens und der sich damit ergebenden Befunde setzt deshalb zumindest eine detaillierte Kenntnis des zugrundegelegten Modells voraus [5].

Ein erstes Problem bei der Formulierung eines Modells als Ausgangspunkt für eine Reihenzerlegung ist die Frage nach der Zahl und dem Typ der zu erfassenden Komponenten und deren arithmetischen Verknüpfung - additiv, multiplikativ oder gemischt. Die Art der Verknüpfung ist dabei nicht aus ökonomischen Überlegungen deduzierbar, sondern immer nur ein integraler Bestandteil des gewählten formalen Modellansatzes.

Historisch gesehen sind die Verfahren zur Analyse von Zeitreihen häufig ohne explizite Definition der Modellstruktur konzipiert worden. Das heisst aber keinesfalls, dass solche Verfahren ohne eine Modellvorstellung über die Art und die Verknüpfung der einzelnen Komponenten auskommen. Durch die einzelnen Arbeitsschritte eines bestimmten Analyseverfahrens wird stets ein Modell impliziert, auch wenn es oft nicht möglich ist, das durch den Ablauf der Analyseschritte resultierende Komponentenmodell explizit zu bestimmen [6]. Das wohl bekannteste Beispiel für ein Verfahren, bei dem das zugrundeliegende Modell nur durch die Folge der Arbeitsschritte und damit nur implizit bestimmt wird, ist das CENSUS-X-11-Verfahren. Eine nur implizite Definition erschwert eine methodologische Beurteilung und führt unter Umständen zu Inkonsistenzen beim Einsatz der einzelnen Arbeitsschritte [7]. Insbesondere ist es im Falle des CENSUS-X-11-Verfahrens bisher nur unter der Annahme einer additiven Verknüpfung der Komponenten gelungen, einen stochastischen Prozess für die Erzeugung der Daten zu finden, bei dem dieses Verfahren zu optimalen Ergebnissen führt [8]. Obwohl in der Praxis vorzugsweise von einer multiplikativen Verknüpfung der Komponenten ausgegangen wird, konnte gerade für diese Annahme noch kein explizites Zeitreihenmodell gefunden werden, das damit kompatibel ist und als Basis für eine methodische Beurteilung des Verfahrens dienen könnte.

Bei den Verfahren mit explizit formulierten Komponentenmodellen ist zwischen lokalen und globalen Modellen zu unterscheiden. Lokale Modelle beschreiben die Eigenschaften der einzelnen unterstellten Komponenten jeweils nur für Teilbereiche der Beobachtungsperiode [9]. Ein Beispiel für einen lokalen Ansatz ist das BERLINER VERFAHREN, wie es vom Statistischen Bundesamt in der BRD eingesetzt wird [10].

Ohne auf eine detaillierte Darstellung der historischen Entwicklung der Zeitreihenanalyse einzugehen - dazu sei auf die spezielle Literatur verwiesen [11] - sollen nachfolgend einige interessante Entwicklungsschritte gezeigt und einige bekannte Verfahren in den Grundzügen skizziert werden. Dabei soll hier natürlich das Problem einer Saisonbereinigung im Vordergrund stehen.

Viele grundlegenden Probleme der Saisonbereinigung können bereits anhand sehr einfacher Zeitreihenmodelle

dargestellt werden [12]. Dies scheint aus zwei Gründen als zweckmässig. Einmal werden in der Praxis häufig Verfahren verwendet, die in ihren Grundzügen auf diesen Modellen aufbauen, und zum andern werden dabei Probleme transparent, die in gleicher oder ähnlicher Wiese auch bei komplizierteren Verfahren in Erscheinung treten bzw. treten können.

Die beiden nachfolgend skizzierten Modelle unterscheiden sich im Verlaufsmuster ihrer Saisonfiguren. Während das eine von einer *konstanten* Saisonfigur ausgeht, erlaubt das andere eine gewisse, noch näher zu definierende *Variabilität* in der Saisonfigur. Beide Ansätze basieren auf einem *additiven* Komponentenmodell.

Für jede der Zeitperioden $t = 1,2,...,T$ (Monate, Quartale oder Halbjahre) sollen Beobachtungswerte X_t eines metrischen Merkmales vorliegen (z.B. Umsätze eines Artikels, Anzahl der Arbeitslosen usw.). Grundlegend ist nun die Vorstellung, dass sich die beobachteten Zeitreihenwerte additiv aus verschiedenen *Bewegungskomponenten* zusammensetzen. Für dieses Komponentenmodell lässt sich dann schreiben

$$X_t = T_t + Z_t + S_t + U_t \ , \quad t = 1,2,...,T \ .$$

T_t ist die *Trendkomponente*, deren Verlauf als durch langfristig wirkende Ursachen bedingt angesehen wird. Häufig wird unterstellt, dass sie monoton wachsend oder monoton fallend verläuft. Z_t ist die *zyklische Komponente*, deren Verlauf als durch den Konjunkturzyklus bedingt angesehen wird.

Häufig werden T_t und Z_t nicht getrennt betrachtet. Die Summe dieser beiden Komponenten wird üblicherweise als *glatte Komponente* G_t bezeichnet. S_t ist die *Saisonkomponente*, deren Verlauf auf jahreszeitliche und institutionelle Ursachen zurückgeführt wird.

Saisonbereinigung bedeutet nun, dass S_t identifiziert und eliminiert werden kann. Es ist unmittelbar einleuchtend, dass saisonbereinigte Reihen für die Wirtschaftspraxis von grosser Bedeutung sind.

U_t ist die *irreguläre Komponente*. Sie wird als von Ursachen bedingt angesehen, die nicht zu den bisher aufgeführten Ursachenkomplexen zu zählen sind. Es wird angenommen, dass die U_t-Werte relativ (d.h. im Vergleich zu den G_t- uns S_t-Werten) klein sind und *regellos* um den Wert Null schwanken. Die U_t sind Residualschwankungen, die den Charakter von *Zufallsschwankungen* aufweisen sollen. Deshalb werden sie in den Zeitreihenmodellen, die den jeweiligen Bereinigungsverfahren zugrundeliegen, in der Regel als weisses Rauschen interpretiert.

Nimmt man nun an, dass die Saisonfigur konstant ist, d.h. dass gilt

$$S_t = S_{t+12} \ , \ t = 1,2,...$$

(strenge Periodizität der Saisonfigur), dann lässt sich unter Hinzunahme von zwei weiteren Voraussetzungen ein denkbar einfaches Saisonbereinigungsverfahren konstruieren [13]. Diese Voraussetzungen sind:

a) die glatte Komponente G_t kann innerhalb eines Zeitraumes von 13 Monaten hinreichend genau durch eine lineare Funktion der Zeit approximiert werden

b) die Summe oder ein gewogenes Mittel der Werte der irregulären Komponente kann mit hinreichender Genauigkeit durch die Zahl Null ersetzt werden. Es ist zweckmässig, die Saisonfigur, d.h. das 12-Tupel $(S_1, S_2,...,S_{12})$ auf die Summe Null zu normieren, d.h. $S_1 + S_2 + ... + S_{12} = 0$. Diese Normierung bedeutet keine Einschränkung der Allgemeinheit des Modellansatzes, da von den S_t-Werten eine Konstante subtrahiert und der glatten Komponente zugeschlagen werden kann.

Unter der Voraussetzung a) kann die glatte Komponente mit Hilfe eines gleitenden symmetrischen 12-Monats-Durchschnitts geschätzt werden

$$\hat{G}_t = \frac{1}{12}\left[\frac{1}{2}X_{t-6} + \frac{1}{2}X_{t+6} + \sum_{i=t-5}^{t+5} X_i\right] \ ,$$

denn die saisonale und die Restkomponente können wegen der Voraussetzung b) approximativ als Null angenommen werden.

$\hat{G}_t$ ist offensichtlich ein symmetrischer Filter. Nachteilig ist allerdings, dass sowohl am Anfang als auch am Ende einer Reihe Werte verlorengehen (je sechs), d.h. die glatte Komponente kann nicht bis zum *aktuellen Rand* geschätzt werden. Da sich das praktische Interesse aber häufig gerade auf diesen Randbereich konzentriert, ist dieser Nachteil besonders gravierend. Obiger Filter könnte natürlich auch asymmetrisch implementiert werden, womit man formal bis zum aktuellen Rand käme. Allerdings wäre das eine Scheinlösung, da ein solcher Filter eine Phasenverschiebung impliziert, die je nach Filterkonstruktion praktisch mehr oder weniger bedeutsam ist. Hier zeigt sich schon deutlich die sogenannte *Randproblematik* von Saisonbereinigungsverfahren, die auch bei wesentlich komplizierteren Verfahren praktische und theoretische Schwierigkeiten aufwirft.

Da $\hat{G}_t$ eine Schätzung für G_t ist, kann geschrieben werden

$$X_t - \hat{G}_t \approx S_t + U_t \ .$$

Die S_t- und U_t-Werte werden nun über gleichnamige

Monate gemittelt. X_{ij} sei der Reihenwert des j-ten Monats des i-ten Jahres und $S_{ij} = S_j$ aufgrund der angenommenen konstanten Saisonfigur. Dann ist

$$\tilde{S}_j := \frac{1}{n_j} \sum_i (X_{ij} - G_{ij}) \approx S_j + \frac{1}{n_j} \sum_i U_{ij} \ , \ j = 1,2,...,12 \ .$$

n_j bedeutet die Anzahl der Jahre, für die $\tilde{S}_j$ berechenbar ist. Die Normierung der 12 Werte $\tilde{S}_1, \tilde{S}_2, ..., \tilde{S}_{12}$ erfolgt durch eine Subtraktion des arithmetischen Mittels

$$\hat{S}_j := \tilde{S}_j - \frac{1}{12} \sum_{j=1}^{12} \tilde{S}_j \ , \ j = 1,2,...,12 \ .$$

Die saisonbereinigte Reihe ergibt sich schliesslich aus den Differenzen $X_{ij} - \hat{S}_j$.

Dieses einfache Verfahren wurde z.B. in den dreissiger Jahren in den USA vom "Federal Reserve Board" und vom "Londoner Wirtschaftsdienst" verwendet [14].

Der einfachste Fall einer variablen Saisonfigur liegt vor, wenn die Saisonkomponente für den Monat j als proportional zur glatten Komponente angenommen wird [15]

$$S_{ij} = \lambda_j G_{ij} \ .$$

War beim vorherigen Verfahren die absolute Grösse der Saisonausschläge konstant für gleichnamige Monate, so ist es hier die relative Grössenordnung (in bezug auf die glatte Komponente). Das additive Zeitreihenmodell lautet in diesem Kontext

$$\begin{aligned} X_{ij} &= G_{ij} + \lambda_j G_{ij} + U_{ij} \\ &= I_j G_{ij} + U_{ij} \ , \ I_j = 1 + \lambda_j \ . \end{aligned}$$

Der Faktor I_j wird als *Saisonindexziffer für den j-ten Monat* bezeichnet. Mit denselben Voraussetzungen wie oben ergibt sich

$$\frac{X_{ij}}{G_{ij}} = I_j + \frac{U_{ij}}{G_{ij}}$$

bzw.

$$\frac{X_{ij}}{\hat{G}_{ij}} \approx I_j + \frac{U_{ij}}{\hat{G}_{ij}} \ .$$

Mittelt man diese Quotienten über alle jeweils verfügbaren Jahre für gleichnamige Monate, ergibt sich eine Schätzung für die Saisonindexziffern

$$\tilde{I}_j = \frac{1}{n_j} \sum_i \frac{X_{ij}}{\hat{G}_{ij}} \ , \ j = 1,2,...,12 \ .$$

Die I_j werden auf die Summe 12 normiert. Die geschätzten Saisonindexziffern sind dann

$$\hat{I}_j = \tilde{I}_j \frac{12}{\sum_{j=1}^{12} \tilde{I}_j} \ , \ j = 1,2,...,12 \ .$$

Die Quotienten $X_{ij} / \hat{I}_j$ bilden schliesslich die saisonbereinigte Reihe. Die *Randprobleme* sind die gleichen wie im Modell mit einer konstanten Saisonfigur.

Eine konstante Saisonfigur bzw. eine variable der beschriebenen Art dürfte bei praktischen Zeitreihen - wenn überhaupt - nur selten anzutreffen sein. Die beiden dargestellten Verfahren machen aber deutlich, dass schon bei den einfachsten Saisonfiguren eine Reihe von Annahmen zu treffen sind, ohne die eine Identifikation und Elimination der saisonalen Komponente nicht möglich ist.

Es ist auch unmittelbar einzusehen. dass ohne spezielle Postulate eine Komponentenzerlegung überhaupt nicht möglich sein kann. Für jedes t ist nur ein Reihenwert vorhanden, und wenn z.B. in vier Komponenten zerlegt werden soll, dann ergeben sich 4T Unbekannte bei nur T Gleichungen, d.h. die Anzahl der Unbekannten ist stets grösser als die Anzahl der Gleichungen.

Kompliziertere Bereinigungsverfahren, die von mehr oder weniger unregelmässigen Saisonfiguren ausgehen, treffen andere bzw. diffizilere Annahmen über die Verlaufsformen der einzelnen Komponenten. Beispielsweise wird postuliert, dass die glatte Komponente durch einen Polynomansatz mindestens 3. Ordnung (damit konjunkturelle Wendepunkte erfassbar sind) modelliert werden kann, dessen Koeffizienten durch einen Regressionsansatz zu schätzen sind. Damit sind aber neue Probleme verbunden. Beim *updating*, d.h. wenn neue Beobachtungswerte hinzukommen, ändern sich die geschätzten Regressionskoeffizienten möglicherweise erheblich, so dass sich die bisher bereinigten Werte - vor allem am aktuellen Rand - eventuell nicht unerheblich verändern. Das wiederum kann zur Folge haben, dass die letzte(n) Diagnose(n) revidiert werden muss (müssen) -, ein Sachverhalt, der in der Praxis besonders unerfreulich sein kann. Damit ist wiederum das Problem der *Randstabilität* angesprochen, das bei den praktizierten Verfahren vielfach einige Mühe bereitet.

Das schon mehrfach angesprochene Verfahren des Bureau of the Census, das CENSUS-X-11-Verfahren, lässt sich folgendermassen charakterisieren. Voraus sei angemerkt, dass es das CENSUS-X-11-Verfahren als solches nicht gibt. Vielmehr wurde vom Bureau of the Census in den USA seit den 50er Jahren eine Folge von Verfahren entwickelt, die ihrerseits auf Ansätze zurückgehen, die in den 20er und 30er Jahren von Federal Reserve Board und dem National Bureau of Economic Research vertreten wurden. Im wesentlichen benutzen diese Verfahren *ratio-to-moving-average*-Methoden, d.h. es wird ein multiplikatives Zeitreihenmodell unterstellt, wobei die glatte Komponente durch einen gleitenden 12-Monatsdurchschnitt geschätzt wird und die Quotienten

$$G_t \cdot S_t \cdot U_t / \hat{G}_t \approx S_t \cdot U_t$$

für gleichnamige Monate geglättet werden. Dabei wird angenommen, dass dadurch die irreguläre Komponente *ausgebügelt* und deshalb eine Schätzung von S_t möglich wird. Eine Division von X_t durch die geschätzte Saisonkomponente $\hat{S}_t$ ergibt die saisonbereinigte Reihe. Würde man statt des multiplikativen einen additiven Ansatz wählen, so könnte man analog vorgehen, wenn man die Quotienten durch Differenzen ersetzt. Offensichtlich ist das skizzierte Vorgehen im Kern ganz ähnlich wie bei den oben dargestellten einfachen Prozeduren.

Bis 1954 wurde die Saisonbereinigung im Bureau of the Census manuell durchgeführt. In diesem Jahr wurde zum ersten Mal ein einschlägiges Computer-Programm eingesetzt (CENSUS I genannt), das 1955 modifiziert und als CENSUS II bekannt wurde. Seither wurden verschiedene Varianten dieses Programms entwickelt bis hin zur Variante X 11, die noch heute im Gebrauch ist [16]. Das Symbol *X* soll auf *experiment* hinweisen. Damit ist auch gleich angedeutet, dass die Verfahrensvarianten nicht auf der Basis von theoretischen Zeitreihenmodellen entwickelt wurden, sondern auf experimentelle Weise. Dieser Sachverhalt stellt dann auch einen der Kritikpunkte dar, die gegen X 11 erhoben werden. CENSUS X 11 ist aber zweifellos das heute weltweit am häufigsten praktizierte Saisonbereinigungsverfahren.

Da bei CENSUS X 11 kein ausformuliertes, d.h. kein explizites Zeitreihenmodell zugrundeliegt, ist es nicht möglich, eine geschlossene Darstellung des Verfahrens zu geben. Es kann lediglich eine Folge von Arbeitsgängen angegeben werden, die das Verfahren definieren. Grundsätzlich sind zwei Phasen zu unterscheiden, die sowohl bezüglich ihres Aufbaues als auch ihrer Zielsetzung nach verschieden sind. In Phase 1 wird eine sogenannte *vorläufige* Komponentenschätzung vorgenommen, die zur Identifikation und Elimination von *Extremwerten* dienen soll, und in Phase 2 wird die *endgültige* Komponentenschätzung durchgeführt, die schliesslich zur *eigentlichen* Saisonbereinigung führt. Neben der Extremwertbereinigung (d.h. Identifikation und Korrektur von Ausreissern) kennt das Verfahren eine sogenannte *arbeitstägliche* Bereinigung (d.h. Korrektur von Kalenderunregelmässigkeiten) sowie eine Bereinigung von *singulären* Effekten (z.B. verursacht durch Streiks), die hier nicht im einzelnen dargestellt werden können. Sieht man einmal von diesen *Spezialbereinigungen* ab, dann kann man die wesentlichen Arbeitsgänge von CENSUS X 11 wie folgt darstellen [17]:

a) compute the differences between the original series and a centered 12-term moving average as a first estimate of the seasonal and irregular components. Six values at each end of the series are lost

b) apply a weighted 5-term moving average to each month separately, to obtain an estimate of the seasonal component. For the next to last 12 values use an asymmetric 4-term moving average

c) adjust these seasonal components to sum to zero (approximately) over any 12-month period by subtracting a centered 12-term moving average from them. To obtain the six missing values at the end of this average, repeat the last available moving average value six times

d) subtract the adjusted seasonal component from the original series, to give a preliminary seasonally adjusted series. The seasonal component for the last six values is missing as a result of step a): for these 6 months use the estimated seasonal component for the corresponding month of the preceeding year

e) apply a 9-, 13- or 23-term Henderson moving average to the seasonally adjusted values, and subtract the resulting trend-cycle series from the original series to give a second estimate of the seasonal and irregular components. At the end of the series use *equivalent* asymmetric moving averages

f) apply a weighted 7-term moving average to each month separately, to obtain a second estimate of the seasonal component. At the end of the series use *equivalent* asymmetric moving averages

g) repeat step c)

h) subtract these final estimates of the seasonal component from the original series, giving the seasonally adjusted series.

Diese Darstellung geht von der additiven Version des Verfahrens aus. Die multiplikative Version, die in der Praxis meistens verwendet wird, ergibt sich aus relativ geringfügigen Modifikationen der obigen Darstellung (statt Subtraktion Division, Adjustierung auf Zwölf statt auf Null usw.).

Zusätzlich zur eigentlichen Saisonbereinigung bietet

CENSUS X 11 eine Anzahl von Tests, die dem Benutzer Güte- und Beurteilungskriterien bezüglich der durchgeführten Bereinigung liefern sollen.

Obwohl das CENSUS-X-11-Verfahren weitverbreitet ist und immer wieder als *erfolgreich* charakterisiert wird, hat es dennoch heftige Kritik ausgelöst. Die wichtigsten Kritikpunkte sind etwa:

- eine mangelnde analytische Transparenz, da kein ausformuliertes Zeitreihenmodell zugrundeliegt
- die Verwendung von symmetrischen zeitinvarianten Filtern in der *Reihenmitte* und von asymmetrischen zeitabhängigen Filtern an den Rändern, vor allem am aktuellen Rand, was Revisionen bei neuen Beobachtungswerten notwendig macht und zu weniger zuverlässigen Schätzungen am Rande als in der *Mitte* führt. Ausserdem implizieren solche Filter unter Umständen Phasenverschiebungen, deren Ausmass aber nicht bekannt ist. (Vgl. aber dazu die Untersuchungen in Kapitel III)
- in die verschiedenen *Spezialbereinigungen* (Extremwert-, arbeitstägliche Bereinigung usw.) geht eine Vielzahl von subjektiven Vorstellungen ein, welche das Verfahren noch undurchsichtiger machen
- insgesamt besitzt das Verfahren zu viele *Schrauben*, an denen ein Benutzer *drehen* kann. Hinzu kommt, dass manche offiziellen Benutzer eine *hauseigene Einstellung* präferieren (z.B. die Deutsche Bundesbank), so dass sich unter CENSUS X 11 durchaus verschiedene Verfahren verstehen lassen, was dazu führen kann, dass CENSUS-X-11-bereinigte Reihen nicht oder jedenfalls nicht ohne weiteres vergleichbar sind.

Die Kritik am CENSUS-X-11-Verfahren hat einerseits Untersuchungen initiiert, diesem Verfahren einen modelltheoretischen *Unterbau* zu verschaffen, andererseits aber auch zu seiner Weiterentwicklung geführt.

Das klassische additive Komponentenmodell kann z.B.auch in folgender Form dargestellt werden

$$X_t = S_t + N_t \ ,$$

wobei N_t zusammenfassend als nicht-saisonale Komponente angesehen werden kann. *Saisonbereinigung* ist in diesem Modell gleichbedeutend mit einer Schätzung von S_t, d.h. mit der Lösung eines *Signal-Extraktions*-Problems. Der Schätzfehler ist dabei

$$\delta_t = S_t - \hat{S}_t = \hat{N}_t - N_t \ .$$

Ein Bereinigungsverfahren kann als *optimal* bezeichnet werden, falls bei einer gegebenen Verlustfunktion der zu erwartende Verlust minimal ist. Für eine quadratische Verlustfunktion und eine in den Beobachtungswerten X_t lineare Schätzfunktion gilt, dass

$$\hat{S}_t = \upsilon(B)X_t = \sum_{j=-\infty}^{\infty} \upsilon_j X_{t-j}$$

den mittleren quadratischen Schätzfehler minimiert, falls X_t ein unendlich langer stationärer stochastischer Prozess ist. Der Operator $\upsilon(B)$ ist gegeben durch die inverse Fourier-Transformation der Quotienten der Spektren von S_t und X_t [18].

Da $\hat{S}_t$ einen gleitenden Durchschnitt darstellt, erhebt sich die Frage, ob es hier Zusammenhänge mit den gleitenden Durchschnitten des CENSUS-X-11-Verfahrens gibt. Da aber ökonomische Zeitreihen selten als Realisationen von stationären Prozessen interpretiert werden können, kann dieses Resultat nicht unmittelbar zur Abklärung dieser Frage herangezogen werden. Jedoch ist es häufig möglich, solche Reihen mit Hilfe des ARIMA-Ansatzes zu modellieren, was impliziert, dass eine Reihe als Realisation eines (homogen) nicht-stationären Prozesses aufgefasst wird, der durch Differenzenbildung in einem stationären Prozess transformiert werden kann.

CLEVELAND und CLEVELAND-TIAO haben gezeigt, dass obiges Optimierungsproblem auch für ARIMA-Prozesse gelöst werden kann, und dass die Lösung von derselben Form wie oben angegeben ist [19]. Das kann als Indiz dafür angesehen werden, dass für additive Zeitreihenmodelle die Verwendung von gleitenden Durchschnitten zur Saisonbereinigung angezeigt ist. Daran knüpfen auch Erweiterungen des CENSUS-X-11-Verfahrens mit ARIMA-Modellen an.

Das Charakteristische an CENSUS-X-11-ARIMA ist in einer Kombination des CENSUS-X-11-Verfahrens mit Prognosen auf der Basis von ARIMA-Modellen zu sehen. Die zugrundeliegende Idee ist einfach. Um am Reihenrand ebenfalls symmetrische Filter einsetzen zu können, liegt es nahe, die dazu erforderlichen zukünftigen Reihenwerte zu prognostizieren, z.B. mit Hilfe von ARIMA-Modellen. CENSUS-X-11-ARIMA besteht im wesentlichen aus folgenden Schritten [20]:

a) der zu bereinigenden Reihe wird ein ARIMA-Modell angepasst

b) mit Hilfe dieses Modells werden die Reihenwerte für die nächsten 12-24 Monate prognostiziert und die Ausgangsreihe entsprechend verlängert

c) die so verlängerte Reihe wird mit dem CENSUS-X-11-Verfahren saisonbereinigt.

CENSUS-X-11-ARIMA ist heute z.B. das offizielle Bereinigungsverfahren von Statistics Canada.

Da für einen praktischen Einsatz ein Verfahren weitgehend automatisiert sein muss, stellt sich hier das Problem einer möglichst *ökonomischen* Identifikation und Schätzung des jeweils in Frage kommenden ARIMA-Modells.

Gegenüber CENSUS X 11 wird zugunsten CENSUS-X-11-ARIMA geltend gemacht, dass beträchtlich weniger Revisionen am aktuellen Rand erforderlich sind. Dazu liegen ausführliche Studien vor [21].

Die bereits seit längerer Zeit angekündigte ARIMA-Erweiterung des X-11-Verfahrens durch das CENSUS Bureau der USA, die sogenannte Version X-12-ARIMA, war bei Abschluss dieses Berichtes für detaillierte Analysen noch nicht verfügbar [22]. Offensichtlich macht die Erstellung einer PC-kompatiblen Version Schwierigkeiten [23].

Das SABL-Verfahren weist starke Ähnlichkeiten zum CENSUS-X-11-Verfahren auf [24]. Wie dieses ist es iterativ aufgebaut, verwendet eine Vielzahl von Glättungsprozeduren, die nicht im Hinblick auf wohldefinierte Zielsetzungen im Rahmen eines expliziten Zeitreihenmodells entworfen sind, sondern generellen Glattheitskriterien entsprechen. Es kennt eine additive sowie - neben anderen - eine multiplikative Komponentenverknüpfung. Eine geschlossene Darstellung ist auch hier nicht möglich.

Der wesentliche Unterschied zum CENSUS-X-11-Verfahren liegt darin, dass bei SABL *resistente* Glättungsverfahren eingesetzt werden. Solche Verfahren haben die Eigenschaft der Unempfindlichkeit gegenüber Ausreissern, also gegenüber Reihenwerten, die grössenmässig stark von der Mehrzahl der Daten abweichen. So ist z.B. ein gleitender Durchschnitt keine resistente Glättungsprozedur, während z.B. ein gleitender Median als resistent zu bezeichnen ist. Ein weiteres Charakteristikum von SABL ist darin zu sehen, dass zur Beurteilung der Güte der Bereinigung vorwiegend graphische Verfahren herangezogen werden.

Das BERLINER VERFAHREN [25], von dem bis jetzt vier Versionen existieren, ist das offizielle Bereinigungsverfahren des Statistischen Bundesamtes in der BRD.

Ausgangspunkt ist auch hier das klassische Komponentenmodell, wobei angenommen wird, dass seine Komponenten durch stochastische Prozesse modelliert werden können, für deren Realisastionen gelten soll

$$G_t = \sum_{j=0}^{p} a_j t^j ,$$

$$S_t = \sum_{k=1}^{6} (b_k \cos \lambda_k t + c_k \sin \lambda_k t) ,$$

d.h. für die glatte Komponente G_t wird ein Polynom p-ten Grades der Zeit und für die Saisonkomponente ein trigonometrisches Polynom der Zeit mit den saisonalen Frequenzen $\lambda_k = 2\pi k/12,\ k = 1,...,6$ postuliert. Für die Restkomponente ε_t wird ein schwach stationärer Prozess unterstellt mit $E(\varepsilon_t) = 0$. Damit kann geschrieben werden

$$X_t = \sum_{j=0}^{p} a_j t^j + \sum_{k=1}^{6} (b_k \cos \lambda_k t + c_k \sin \lambda_k t) + \varepsilon_t .$$

Ein Charakteristikum des Verfahrens besteht nun darin, dass die Gültigkeit dieses Modells nicht für alle Zeitpunkte unterstellt wird, sondern nur für eine Teilmenge derselben, die *gleitend* das Intervall *[1,T]* überdeckt. Dabei wird zugelassen, dass obige Darstellung für G_t *und* S_t i.a. verschieden lange Teilintervalle N_1 *bzw.* N_2 erfordert.

Falls es für einen Zeitpunkt mehr als einen Stützbereich gibt, muss ein Auswahlkriterium herangezogen werden. Als solches wird prinzipiell *Varianzminimalität* gewählt. Allerdings führt dieses Kriterium nicht immer zu einer eindeutigen Lösung oder erscheint aus anderen Gründen als nicht vorteilhaft.

Die glatte und die saisonale Komponente sollten simultan geschätzt werden. Praktisch ist dies jedoch nicht der Fall, da die Stützbereichslängen N_1 *und* N_2 für ökonomische Reihen als stark unterschiedlich anzunehmen sind ($N_1 < N_2$).

Deshalb wird beim praktischen Einsatz zuerst die glatte Komponente geschätzt (Stützbereichslänge N_1) und von der Originalreihe subtrahiert. Anschliessend wird die Saisonkomponente aus der um die geschätzte glatte Komponente bereinigten Originalreihe geschätzt (Stützbereichslänge N_2). Die Subtraktion der geschätzten Saisonkomponente von der Originalreihe ergibt die saisonbereinigte Reihe.

Wie immer bei Regressionsansätzen sind auch hier - soll das Verfahren praktizierbar sein - Spezifikationsprobleme zu lösen. Diese sind hier im wesentlichen die Festlegung des Polynomgrades p, der Stützbereichslängen N_1 *und* N_2 sowie der Stützstellen τ_1 *und* τ_2 innerhalb dieser Stützbereiche. Als Gütekriterium für eine konkrete Spezifikation werden in erster Linie die Eigenschaften der daraus resultierenden Filter im Frequenzbereich angesehen, insbesondere ihre Amplitudenfunktionen (dem oben erwähnten Gütekriterium *Varianzminimalität* kommt somit nur sekundäre Bedeutung zu). Diese werden beurteilt auf dem Hintergrund einer *Referenz-Amplitudenfunktion*, die keine unerwünschten Absorptions- bzw. Verstärkungseffekte aufweist und die deshalb von den Konstrukteuren des BERLINER VERFAHRENS als *ideal* bezeichnet wird. Sie nimmt im Niederfrequenzbereich *[0, 1/12)* genau den Wert Eins - und in den saisonalen Frequenzen *2k/12, k = 1,...,6* den Wert Null an und verläuft dort in Form von fünf *Kerben*. Zwischen den saisonalen Frequenzen nimmt sie jeweils den

Wert Eins an (vgl. dazu Abschnitt 4 in diesem Kapitel).

In der Version I des BERLINER VERFAHRENS wurden mehr als 200 Filter bezüglich ihrer Amplitudenfunktion untersucht. Immer ergaben sich aber unerwünschte Verstärkungs- bzw. Absorptionseffekte. Als *günstig* für *mittlere* Reihenwerte wurde schliesslich folgende Spezifikation angesehen: $p = 3, N_1 = 23, \tau_1 = 13, N_2 = 45, \tau_2 = 24$. Für die beiden Ränder wurden asymmetrische Filter gewählt.

Die der Originalversion folgenden Versionen wurden vom Statistischen Bundesamt in der BRD entwickelt, das heute die Version 4 benützt. Alle Versionen können als Versuche angesehen werden, die *ideale* Amplitudenfunktion immer besser zu approximieren. Dazu wurde in Version 2 zusätzlich zum Polynom 3. Ordnung ein trigonometrisches Polynom mit einer Periode von 36 Monaten miteinbezogen, um konjunkturellen Bewegungen *besser* als mit einem reinen Polynomansatz Rechnung tragen zu können. Darüber hinaus wurden Stützbereiche und Stützstellen bei der Schätzung der glatten und der Saisonkomponente gegenüber der Originalversion verändert. Version 3 ist insofern bemerkenswert, als das Konzept eines Standard-Bereinigungsverfahrens aufgegeben wurde [26]. Da die bis dahin verwendeten Filter nicht verzerrungsfrei waren, musste man in Kauf nehmen, dass sich diese Verzerrungen je nach Reihe ganz unterschiedlich auswirken können. Ausserdem wurde die Randstabilität der bisherigen Versionen als unzureichend beurteilt. Die Filterauswahl erfolgt deshalb bei dieser Version in Abhängigkeit von der jeweils konkret zu bereinigenden Reihe (*reihenspezifisch*), wobei diejenige Filterfolge als *optimal* angesehen wird, die zu einer saisonbereinigten Reihe führt, deren (geschätztes) Spektrum die kleinste Distanz (im KQ-Sinne) zum (geschätzten) Spektrum der Ausgangsreihe in allen Frequenzen (mit Ausnahme der saisonalen) aufweist. Dies ist gleichbedeutend damit, dass diejenige Filterfolge ausgewählt wird, deren Amplitudenfunktion der idealen - im angegebenen Sinn - am nächsten kommt. In der Version 4 kehrte das Bundesamt allerdings wieder zu einem Standardverfahren zurück. Im Prinzip werden dieselben Ansätze wie in den bisherigen Versionen verwendet, allerdings mit zwei Neuerungen. Bei der Schätzung der glatten Komponente wird ein gewogener KQ-Ansatz verwendet, wobei als Gewichtsfunktion eine symmetrische (für mittlere Werte) bzw. eine asymmetrische (für Randwerte) Dreiecksfunktion benützt wird. Hier zeigen sich offensichtlich Parallelen zum SABL-Verfahren. Die zweite Neuerung besteht darin, dass am rechten Rand einer Reihe das Polynom 3. Ordnung für die glatte Komponente sukzessive durch eine Kombination aus zwei Polynomen 3. bzw. 1. Ordnung ersetzt wird, wobei das Gewicht des linearen Ansatzes gegen Reihenende zunimmt. Durch diese Massnahme wird die Randstabilität des Verfahrens erhöht.

Bei den neueren Entwicklungen sind Verfahren zu nennen, die auf AKAIKE und KITAGAWA/GERSCH, PAULY und SCHLICHT zurückgehen. Grundlegend ist eine Idee, die von WHITTAKER stammt und als Lösung des folgenden Problems vorgeschlagen wurde [27].

Das Zeitreihenmodell sei

$$X_t = f_t + u_t \quad , \quad t = 1,...,T \,,$$

wobei f_t eine unbekannte glatte Funktion der Zeit ist und die u_t identische und unabhängige Zufallsvariablen mit Erwartungswert Null und konstanter Varianz σ^2. Die Funktion f_t stellt die *systematische* Komponente der Zeitreihe dar und soll unter Beachtung der beiden folgenden Forderungen geschätzt werden:

- die Schätzungen von f_t sollen möglichst wenig von den gegebenen Reihenwerten abweichen
- die Funktion f_t soll möglichst glatt sein.

Da sich diese beiden Forderungen widersprechen, können sie nicht beide gleichzeitig realisiert werden. Der von WHITTAKER [28] vorgeschlagene Kompromiss besteht in der Minimierung der Zielfunktion

$$\sum_t [\, X_t - f_t \,]^2 + \lambda^2 \sum_t [\Delta_k - f_t \,]^2 \,,$$

wobei k die Ordnung des Differenzenoperators Δ bezeichnet und λ^2 einen Steuerungsparameter. Der erste Summand bringt die Abweichungen zwischen den Reihenwerten und den Werten der systematischen Komponente f_t zum Ausdruck, der zweite stellt ein Mass für die Krümmung von f_t dar. Der zu wählende Parameter λ^2 beschreibt den einzugehenden Kompromiss zwischen Anpassung und Glattheit.

AKAIKE und SCHLICHT haben diesen Ansatz auf die Analyse von Zeitreihen mit einer saisonalen Komponente erweitert. Danach setzt sich die systematische Komponente additiv aus glatter Komponente G_t und saisonaler Komponente S_t zusammen. Das von SCHLICHT formulierte *Prinzip der Saisonbereinigung* umfasst die drei Postulate:

a) die Summe aus glatter und saisonaler Komponente sollte die gegebene Reihe möglichst gut approximieren, d.h. die irreguläre Komponente $u = x - g - s$ sollte möglichst klein sein *(x, g, s, u sind $T \times 1$-Vektoren, z.B. $x' = (X_1, X_2,...,X_T)$*

   Sei $f_3 : \mathbb{R}^T \rightarrow \mathbb{R}$ ein Mass für die Grösse von u (z.B. $f_3(u) = u'u$), dann ist die Minimierung von $f_3(x - g - s)$ eine Operationalisierung des Postulats a).

b) g sollte eine möglichst glatte Kurve sein. Sei $f_1 : \mathbb{R}^T \rightarrow \mathbb{R}$ ein (beliebiges) Krümmungsmass. Postulat b) erfordert eine Minimierung von $f_1(g)$

c) die saisonale Komponente sollte ein möglichst stabiles Verlaufsmuster aufweisen und zusätzlich über einen Saisonzyklus hinweg im Durchschnitt Null ergeben.

Sei $f_2 : \mathbb{R}^T \rightarrow \mathbb{R}$ ein Mass für die Instabilität des saisonalen Musters und der Abweichung vom Durchschnitt Null. Postulat c) erfordert eine Minimierung von $f_2(s)$.

Falls die Funktionen f_1, f_2, f_3 gegeben sind, lassen sich glatte und saisonale Komponente als Lösung des folgenden Minimierungsproblems bestimmen

$$V(g, s \mid x) := f_1(g) + f_2(s) + f_3(u) = \min_{g,s}$$

SCHLICHT [28] schlägt folgende Funktionen vor

$$f_1(g) := \alpha \sum_{t=3}^{T} [\, (G_t - G_{t-1}) - (G_{t-1} - G_{t-2}) \,]^2 , \quad \alpha > 0$$

$$f_2(s) := \beta \sum_{t=s+1}^{T} (S_t - S_{t-s})^2 + \gamma \sum_{t=s}^{T} \Big(\sum_{\tau=0}^{s-1} S_{t-\tau} \Big)^2 ,$$

$$\beta \geq 0 , \gamma > 0 \qquad (s = 12 \text{ für Monatsdaten}).$$

$$f_3 (x - g - s) := \sum_{t=s+1}^{T} (X_t - G_t - S_t)^2 .$$

In $f_2(s)$ misst der erste Summand die Stabilität der Saisonkomponente und der zweite die Summe der Abweichungen dieser Komponente je Jahr vom Mittelwert Null.

Obiges Minimierungsproblem hat für die gewählte Spezifikation eine eindeutige Lösung. Freiheitsgrade besitzt das Verfahren bei der numerischen Konkretisierung der Gewichte α, β *und* γ. Je grösser α und β, desto glatter ist G_t bzw. desto stabiler ist S_t und desto grösser ist natürlich die irreguläre Komponente. Statt einer ad-hoc-Festlegung dieser Parameter kann auch versucht werden, sie so zu konkretisieren, dass die ex-post-Prognosefehler für s und g minimal werden. Letztere sind dadurch definiert, dass die Analyseresultate (für einen gegebenen Satz von Koeffizienten) bis zur Periode $t < T$ mit denjenigen für die gesamte verfügbare Reihe verglichen und die entsprechenden Differenzen gebildet werden. Als "optimal" wird dann derjenige Koeffizientensatz angesehen, der diese Prognosefehler minimiert.

Die theoretischen Überlegungen zum CENSUS-X-11-Verfahren und zu den verschiedenen Versionen des BERLINER VERFAHRENS, die als "Referenzgrösse" die Amplitudenfunktion *direkt* analysieren, geben Anstösse zu einer anders gelagerten Konstruktion von Saisonbereinigungsverfahren. Mit direkt ist gemeint, dass der Filter im Frequenzbereich konstruiert wird, ohne dass vorgängig Komponenten und Schätzmethoden im Zeitbereich spezifiziert werden. Solche - rein filtertheoretisch konzipierten - Saisonbereinigungsverfahren beruhen deshalb nicht mehr auf dem klassischen Komponentenmodell. Statt den dort üblichen traditionellen vier Komponenten sind hier a priori unendlich viele Komponenten zugelassen, die dem Kontinuum von Frequenzen *[0, 1/2]* zugeordnet sind. Zur Interpretation und Fomulierung von Zielsetzungen ist es allerdings häufig zweckmässig, Frequenzintervalle zu bilden und für diese die üblichen Termini zu verwenden. Zum Beispiel ist es naheliegend, das Intervall *[0, 1/12)* mit der glatten Komponente (im Sinne des traditionellen Komponentenmodelles) zu identifizieren. Auf einen solchen Ansatz wird in Kapitel IV noch näher eingegangen.

3. Kriterien für Saisonbereinigungsverfahren

Es gibt eine Vielzahl von Güte- oder Vergleichskriterien für Saisonbereinigungsverfahren. Quasi parallel mit der methodischen Entwicklung der Verfahren entstanden auch Konzepte für deren Beurteilung. Die gebräuchlichen Kriterien unterscheiden sich jedoch nicht nur durch ihre unterschiedliche Komplexität, sondern auch hinsichtlich ihrer theoretischen Argumentationsebene, so dass sie nicht alle und vor allem nicht ohne weiteres für einen direkten Verfahrensvergleich geeignet sind.

Ein sehr einfaches und pragmatisches Kriterium ist z.B. die Forderung nach der Invarianz der Jahressumme, oft auch als Kriterium der Summenerhaltung bezeichnet. Dabei wird gefordert, dass die Summe der saisonbereinigten Werte in einem Jahreszeitraum gleich der Summe der Beobachtungswerte sein soll. Dieses praktisch sinnvolle und unmittelbar einleuchtende Kriterium eignet sich für einen Verfahrensvergleich jedoch überhaupt nicht. Bei manchen Verfahren ist eine Normierung der Saisonfigur integraler Bestandteil, so dass dieses Kriterium automatisch erfüllt ist. Ausserdem lässt sich jedes Verfahren so modifizieren, dass dieses Kriterium erfüllt wird.

Es stellt sich daher die Frage, welche Kriterien für einen Verfahrensvergleich überhaupt in Frage kommen. Die bekannten Kriterien lassen sich zunächst danach unterscheiden, ob sie im Zeit- oder im Frequenzbereich definiert sind. Von den Kriterien im Zeitbereich seien hier beispielhaft die Orthogonalitätsbedingung und das Kriterium der Idempotenz erwähnt [29]. Die Orthogonalitätsbedingung verlangt, dass eine saisonbereinigte Reihe keine Information mehr enthalten darf, die zur Bestimmung einer Saisonkomponente genutzt werden kann. Eine Verletzung dieses Kriteriums wäre ein Indiz dafür, dass eine saisonbereinigte Zeitreihe noch saisonale Bewegungen enthält. Das Idempotenzkriterium verlangt, dass eine bereinigte Reihe sich nicht verändern darf, wenn diese mit demselben Verfahren noch einmal bereinigt wird.

Die einschlägigen Erfahrungen haben aber gezeigt, dass mit diesen und ähnlichen Kriterien praktisch nicht zwischen den

verschiedenen gebräuchlichen Saisonbereinigungsverfahren diskriminiert werden kann. Häufig erfüllen die verschiedenen Verfahren solche Kriterien in etwa gleichem Ausmass. Diese Feststellung muss auch für eine Reihe anderer Kriterien, die auf die sich ergebende Restkomponente, als Schätzung der irregulären Komponente, abstellen, gemacht werden. Zu dieser Gruppe von Beurteilungskriterien zählen etwa Run-Tests oder andere Tests auf reine Zufälligkeit wie z.B. Periodogrammtests usw. Hinzu kommt, dass alle Kriterien, die auf eine Analyse der Restkomponente abstellen, nur auf Verfahren angewendet werden können, die vom traditionellen Komponentenmodell ausgehen.

Die bisher erwähnten Kriterien gehen implizit von einer retrospektiven Zielsetzung aus. Unter dem Aspekt einer prospektiven Zielsetzung, d.h. zur Beurteilung der Fähigkeit eines Verfahrens, die konjunkturelle Grundtendenz am aktuellen Rand richtig zu diagnostizieren, wurden ebenfalls spezielle Kriterien entwickelt [30]. Dabei wird von einer Diagnose des Systemzustandes ausgegangen, wie er sich am aktuellen Rand einer Zeitreihe aufgrund des verwendeten Analyseverfahrens ergibt. Dieser Befund wird dann mit dem sich aufgrund einer retrospektiven Analyse mit dem gleichen Verfahren später erzielten Systemzustand und mit Hilfe geeigneter Masszahlen verglichen.

Neben solchen in der Zeitdimension konzipierten Kriterien gibt es aber auch Kriterien zur Beurteilung von Saisonbereinigungsverfahren im Frequenzbereich. Dazu gehört z.B. die Forderung, dass das Spektrum einer bereinigten Reihe in den saisonalen Frequenzen weder "Spitzen" noch "Einbrüche" aufweisen und im Niederfrequenzbereich mit dem Spektrum der beobachteten Zeitreihe übereinstimmen sollte. "Spitzen" in den saisonalen Frequenzen werden dabei als unzulängliche Bereinigung, "Einbrüche" als Überbereinigung interpretiert. Ein verändertes Spektrum im Niederfrequenzbereich wird als Verfälschung der Trend-Zyklus-Komponente durch die vorgenommene Saisonbereinigung aufgefasst. Ausserdem wird gefordert, dass die Kohärenzfunktion zwischen der beobachteten Zeitreihe und der bereinigten Reihe in den nicht-saisonalen Frequenzbändern möglichst nahe bei Eins liegen muss und in den saisonalen Frequenzen praktisch Null sein sollte. Diese und ähnliche spektrale Kriterien blieben jedoch ebenfalls nicht von Kritik verschont. Es stellte sich vor allem heraus, dass auch diese Kriterien nur bedingt geeignet sind, einen objektiven Vergleich zwischen Saisonbereinigungsverfahren durchzuführen [31].

Festzuhalten ist, dass jedes der hier bisher genannten Kriterien immer nur Teilaspekte eines Verfahrens erfasst. Dies legt es zunächst nahe, nach einem Katalog von Kriterien zu suchen, der dann Verfahrensvergleiche ermöglicht. Auswahl und Gewichtung der einzelnen Kriterien sind aber unter den Verfahrenskonstrukteuren umstritten, und die Festlegung einer Rangordnung für die einzelnen Kriterien ist auf objektiver Basis auch nicht möglich. Es ist daher nicht verwunderlich, dass aufgrund solcher Kriterienlisten - vielleicht zur Enttäuschung vieler Anwender - noch kein "bestes" Verfahren gefunden werden konnte.

Ein erster Schritt zur Lösung dieser Anwenderprobleme besteht darin, nicht mehr die einzelnen Aspekte eines Verfahrens partiell zu betrachten und zu analysieren, sondern zu versuchen, die Eigenschaften eines Verfahrens als Ganzes zu erfassen. Da die verschiedenen Verfahren aber methodisch sehr unterschiedlich konzipiert sind, ist zunächst nach einer gemeinsamen Basis für alle Verfahren zu suchen. Da ungeachtet aller methodischen Unterschiede jedes Bereinigungsverfahren letzten Endes als Filter betrachtet werden kann, der eine beobachtete Zeitreihe in eine bereinigte Reihe transformiert, bietet sich ein Verfahrensvergleich im Frequenzbereich an. Dabei sind die Filtereigenschaften, d.h. Amplituden- und Phasenfunktion, zunächst nur bei den Verfahren explizit bekannt, die auch mit Hilfe der Filtertheorie konstruiert wurden. Bei allen anderen Verfahren muss zuerst die Gesamt-Filterwirkung ermittelt werden. Dies kann im Einzelfall auf analytischem Weg recht schwierig, in einigen Fällen, insbesondere wenn ein Verfahren aus einer nicht a priori festgelegten Abfolge vieler Arbeitsschritte besteht, z.B. beim CENSUS-X-11-Verfahren, nur approximativ möglich sein. Erschwert wird eine analytische Herleitung der Filtereigenschaften auch durch etwaige nicht-lineare Operationen im Verfahren. Es ist jedoch möglich, die benötigten Informationen über die Filtereigenschaften auf empirischem Weg zu ermitteln. Dabei wird von der Überlegung ausgegangen, dass das Verhalten eines linearen Filters, sprich eines bestimmten Saisonbereinigungsverfahrens, bekannt ist, wenn sein Verhalten für bestimmte Testsignale bekannt ist. Bei diesem Vorgehen lässt sich auch erkennen, ob die Filtereigenschaften eines Verfahrens in der Reihenmitte wesentlich anders sind als am Reihenende, d.h. am aktuellen Rand. Für Verfahren, die nicht-lineare Beziehungen verarbeiten, ist die Ableitung der Eigenschaften im Frequenzbereich mit Hilfe der Theorie linearer Filter aber nur eine Approximation. Es kann jedoch gezeigt werden, dass eine derartige Approximation, z.B. im Falle des CENSUS-X-11-Verfahrens, brauchbare Informationen liefert [32].

4. Zum Design von Saisonfiltern (Transfer- und Phasenfunktionen)

Unter einem "Filter" soll ein Algorithmus zur Transformation einer Zeitreihe verstanden werden. Formal kann der Zusammenhang zwischen einer zu filternden Zeitreihe x_t, dem sog. "Filter-Input", und der gefilterten Reihe y_t, dem sog. "Filter-Output", folgendermassen dargestellt werden:

$$y_t = F(x_t) \ .$$

Beispielsweise stellen gleitende Durchschnitte oder erste Differenzen einfache (und seit langem gebräuchliche) Filter dar.

Die Wirkungsweise eines Filters auf den Filter-Input lässt sich am besten im Frequenzbereich beurteilen, und zwar anhand seiner sogenannten "Transferfunktion" $T(f)$. Diese ist im allgemeinen eine komplexwertige Funktion, für die geschrieben werden kann

$$T(f) = |T(f)|e^{i\phi(f)} , \; 0 \le f \le 0.5 .$$

Dabei ist $|T(f)|$ der Absolutbetrag von $T(f)$ - der üblicherweise als *Amplitudenfunktion* des Filters bezeichnet wird - (im folgenden kurz $A(f)$ genannt), während $\phi(f)$ seine *Phasenfunktion* ist. i bezeichnet die imaginäre Einheit.

Mit Hilfe der Amplitudenfunktion lassen sich die *Ausblendeigenschaften* eines Filters ermitteln. Frequenzkomponenten, für die $A(f)$ kleiner als Eins ist, werden durch die Filterung abgeschwächt, während Frequenzkomponenten verstärkt werden, wenn $A(f)$ grösser als Eins ist. Für $A(f) = 1$ passieren die entsprechenden Frequenzkomponenten den Filter unverändert. Nimmt $A(f)$ den Wert Null an für gewisse Frequenzen bzw. Frequenzintervalle, dann werden die zugehörigen Komponenten völlig eliminiert.

Die Phasenfunktion $\phi(f)$ zeigt an, welche *zeitliche Verzögerung* für welche Frequenzkomponenten zwischen Filter-Input und Filter-Output besteht. $\phi(f)$ misst die Phase in radian, was allerdings wenig anschaulich ist. Deswegen empfiehlt sich eine Umrechnung in Zeiteinheiten, die sich aus einer Division von $\phi(f)$ durch $f \in (0, 0.5]$ ergibt.

In der vorliegenden Untersuchung wird vor allem von zwei Filtertypen intensiv Gebrauch gemacht, nämlich von Tief- und Hochpass-Filtern.

Ein *idealer* Tiefpass-Filter besitzt folgende Amplitudenfunktion (vgl. Abb. 46):

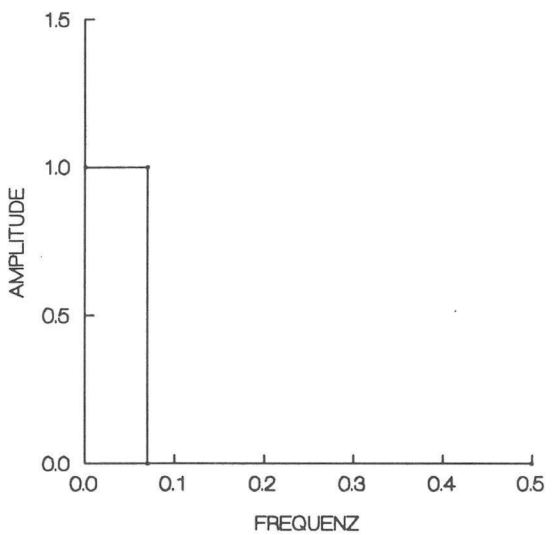


Abb. 46

Dieser Filter lässt alle Schwingungen mit einer Frequenz $f \in [0, f_A]$ unverändert passieren und absorbiert oder eliminiert alle Schwingungen mit $f \in (f_A, 0.5]$. Dabei ist f_A die sog. "Abschneidefrequenz", die in obiger Abbildung gleich 0.07 gesetzt wurde. Frequenzbereiche, die von einem Filter übertragen werden, nennt man *Passbänder*, diejenigen, die eliminiert werden, heissen *Stoppbänder*. Ein Tiefpass-Filter hat demnach das Passband $[0, f_A]$ und das Stoppband $(f_A, 0.5]$. In der Abbildung sind das die Frequenzbereiche $(0, 0.07]$ bzw. $(0.07, 0.5]$. Ein *idealer* Tiefpass-Filter ist somit dadurch ausgezeichnet, dass die Amplitude im Passband *exakt* den Wert Eins und im Stoppband *exakt* den Wert Null annimmt.

Für einen *idealen* Hochpass-Filter verläuft die Amplitudenfunktion folgendermassen (vgl. Abb. 47):

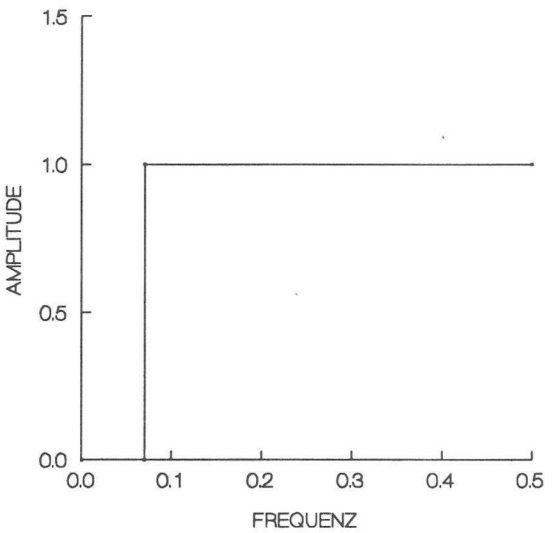


Abb. 47

Dieser Filter lässt alle Schwingungen mit den Frequenzen $f \in [f_A, 0.5]$ unverändert durch und eliminiert alle anderen. In der Abbildung wurde wieder $f_A = 0.07$ gewählt. Im Stopp- bzw. Passband nimmt die Amplitude dieses Filters wie oben *exakt* die Werte Null bzw. Eins an. Offensichtlich sind die Amplitudenfunktionen von Tief- und Hochpass-Filtern zueinander *komplementär*. Es ist

$$A_{HP}(f) = 1 - A_{TP}(f) .$$

Die Abschneidefrequenz f_A ist bei beiden Filtern jeweils von der *Zielsetzung* der Analyse her festzulegen. Will man beispielsweise bei einer monatlichen Zeitreihe alle Schwingungen mit einer Periodenlänge von 12 Monaten und darunter ausfiltern, so ist ein Tiefpass-Filter zu verwenden, bei dem $f_A < 1/12$ ist. Der resultierende Filter-Output entspricht dem, was üblicherweise als *glatte Komponente* bezeichnet wird. Sollen dagegen alle langfristigen Bewegungen einer Zeitreihe eliminiert werden, dann kann das mit Hilfe eines Hochpass-Filters erfolgen, etwa mit der gleichen Abschneidefrequenz wie beim dazu komplementären

Tiefpass-Filter. Der Filter-Output entspricht dann dem, was traditionellerweise als *trendbereinigte* Reihe bezeichnet wird.

Es sei hier angemerkt, dass mit der Festlegung der Abschneidefrequenz auch gleichzeitig die der Analyse im konkreten Fall zugrundegelegte (intendierte) Komponentendefinition offengelegt wird.

Wenn die obigen Filter als *ideal* bezeichnet wurden, dann legt das die Vermutung nahe, dass *reale* Filter die dargestellten exakten Amplitudenfunktionen nicht aufweisen. Ohne auf methodische Details einzugehen, sei hier beispielsweise nur darauf hingewiesen, dass die Amplitudenfunktionen realer Filter nicht die obige "Rechteck-Gestalt" aufweisen können. Vielmehr ist aus gewissen Gründen zwischen Pass- und Stoppband immer ein sog. "Übergangsband" vorzusehen, so dass die anzustrebenden Amplitudenfunktionen für einen Tief- bzw. Hochpass-Filter "Trapez-Gestalt" aufweisen, wie das aus den beiden folgenden Abb. 48 und 49 hervorgeht, bei denen die Länge des Übergangsbandes jeweils 0.01 beträgt.

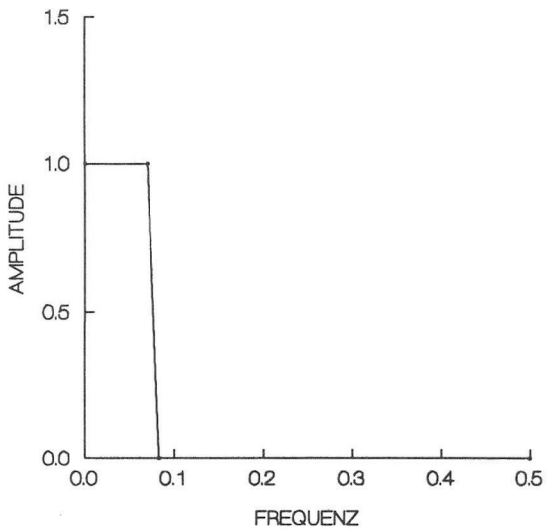


Abb. 48

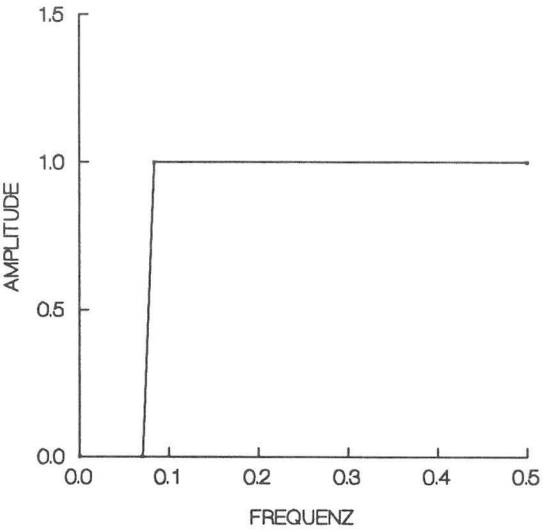


Abb. 49

Eine Zielsetzung des *filter-designs* besteht darin, Filter zu konstruieren, die eine vorgegebene Amplitudenfunktion "möglichst gut" approximieren. Auf spezielle Methoden des filter-designs kann hier nicht eingegangen werden. Es sei nur noch angemerkt, dass bei einer Filterkonstruktion auch jeweils die sich ergebende Phase beachtet werden muss. Für die Praxis der Zeitreihenanalyse sind natürlich nullphasige Filter ideal. Solche lassen sich auch konstruieren, allerdings verbunden mit dem Nachteil, dass sie über keine perfekte Randstabilität verfügen.

Ein idealer Filter zur *Saisonbereinigung* monatlicher Zeitreihen weist folgenden Verlauf auf (vgl. Abb. 50):

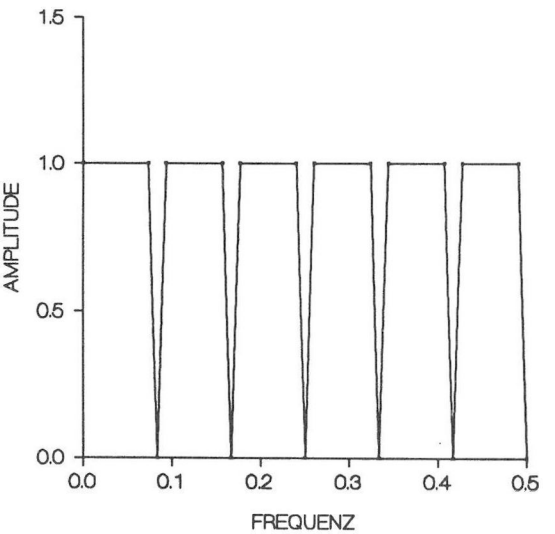


Abb. 50

Diese Amplitude hat die Eigenschaft, dass sie den Wert Null in den saisonalen Frequenzen 1/12, 2/12, 3/12, 4/12, 5/12 und 6/12 annimmt und sonst gleich Eins ist. Der zugehörige Filter - der als (multipler) Kerbenfilter bekannt ist - tangiert also nur die saisonale Komponente einer

Zeitreihe und lässt alle anderen Komponenten unverändert. (Für die Bereinigiung vierteljährlicher Zeitreihen enthält die entsprechende Amplitudenfunktion nur zwei Kerben, und zwar bei den Frequenzen 0.25 und 0.5).

Es gibt verschiedene Möglichkeiten, Filter zu entwerfen, die diese Amplitudenfunktion approximieren. Beispielsweise kann ein sogenannter "IIR-Filter" konstruiert werden, dessen Amplitude sich nur marginal von der idealen unterscheidet [33].

Bei einer Saisonbereinigung mit Hilfe eines derartigen Filters wird das sog. "klassische Komponentenmodell", das üblicherweise vier Komponenten enthält (Trend-, Zyklus- , Saison- und Restkomponente), verlassen. Statt dessen geht man a priori von *unendlich vielen* Komponenten aus, die dem *Kontinuum* von Frequenzen *[0, 0.5]* zuzuordnen sind. Es lässt sich allerdings zeigen, dass auch Saisonbereinigungsverfahren auf der Basis des klassischen Komponentenmodelles die traditionellen Analysemethoden verwenden (meistens gleitende Durchschnitte und/oder lokale Regressionsmodelle), letzten Endes die obige ideale Amplitudenfunktion realisieren wollen, wenngleich auf einem "indirekten" Weg. Dies trifft insbesondere auf das CENSUS-X-11-Verfahren zu, wie aus den Untersuchungen in Kapitel III deutlich hervorgeht. Deshalb eignet sich diese ideale Amplitudenfunktion sehr gut als *standard of reference*, wenn etwa die Filtereigenschaften verschiedener Saisonbereinigungsverfahren untersucht und verglichen werden sollen.

5. Eigenschaften einiger gebräuchlicher Filter

5.1 Vorbemerkungen

In Kapitel I, Abschnitt 3.4, wurde bei der Erörterung der Trendelimination erwähnt, dass dazu in der Praxis dafür häufig 1. Differenzen verwendet werden. Dabei wurde schon negativ vermerkt, dass diese gängige Praxis zu gravierenden Verzerrungen führen kann. Dies ist nun näher auszuführen und anhand einiger Beispiele zu demonstrieren. Gleichzeitig soll aber auf eine weitere in der Praxis häufig anzutreffende Filterung, nämlich auf die Bildung 12. Differenzen, kritisch eingegangen werden. Bei der folgenden Darstellung werden die wichtigsten Zusammenhänge wiederum durch Graphiken illustriert, wobei auf simulierte Reihen aus Kapitel I zurückgegriffen wird.

5.2 Erste Differenzen

Gegeben sei eine Zeitreihe x_t , $t = 1, 2, \ldots, T$. Unter der Bildung 1. Differenzen versteht man die Transformation

$$y_t := x_t - x_{t-1} .$$

Obwohl die eingangs Kapitel I, Abschnitt 3.1, betrachtete reine Sinusschwingung keinen Trend aufweist und deshalb keinerlei Notwendigkeit für eine Differenzenbildung besteht, soll zunächst die Wirkung einer Differenzenbildung auf diese Zeitreihe untersucht werden. Die Abb. 51 zeigt zusammen die Original- und die gefilterte Reihe.

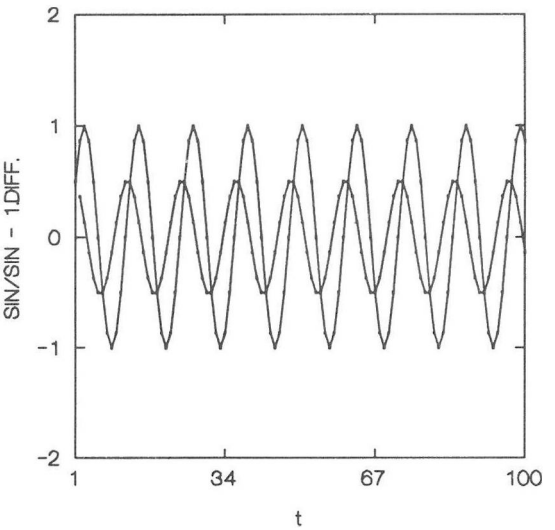


Abb. 51

Daraus lassen sich schon zwei Konsequenzen dieses Differenzen-Filters ablesen. Zum einen dämpft er die Amplitude der Sinusssschwingung und zum anderen weist der Filter-Output eine Phasenverschiebung gegenüber dem Filter-Input, d.h. also gegenüber der Originalreihe auf. Diese beiden Effekte zeigen sich bei beliebigen Input-Signalen, wie gleich noch darzulegen sein wird.

Die Amplituden-Dämpfung zeigt sich deutlich im Spektrum der gefilterten Reihe. Die Abb. 52 enthält die Spektren beider Reihen, die "peak-Ordinate" ist wesentlich kleiner bei der differenzierten Reihe.

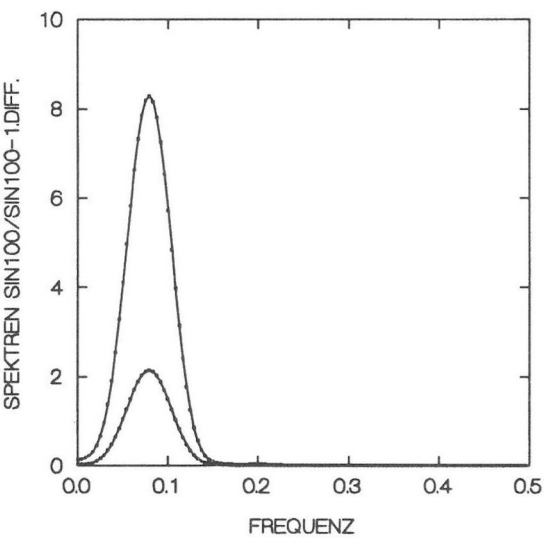


Abb. 52

Bildet man 1. Differenzen der Reihe aus Abb. 25 in Kapitel I, Abschnitt 3.4. (Trend + Sinusschwingung) und vergleicht diese mit der Sinusschwingung, dann zeigt sich folgendes Bild:

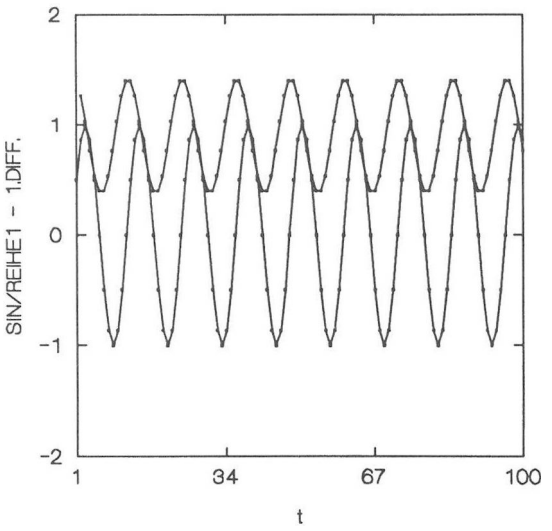


Abb. 53

Neben der Dämpfung der Sinusschwingung und einer Phasenverschiebung ist eine Parallelverschiebung des Filter-Outputs zu beobachten, was leicht verständlich wird, wenn man den Filter-Output berechnet, was hier leicht möglich ist. Man erhält

$$y_t = 2 + 0.9t + \sin(2\pi/12t) - \{2 + 0.9(t-1) + \sin(2\pi/12(t-1)\}$$
$$= 0.9 + 0.134\sin(2\pi/12t) - 0.5\cos(2\pi/12t) ,$$

d.h. vom linearen Trend bleibt bei der Differenzenbildung eine positive Konstante übrig, die im wesentlichen diese Verschiebung bewirkt.

Erste Differenzen implizieren aber nicht nur Dämpfungs-, sondern auch Verstärkungs(oder *Aufrauhungs*-)effekte. Dies zeigt deutlich folgendes Beispiel, bei dem 1. Differenzen der Reihe 4 aus Abb. 31 gebildet wurden (vgl. Abb. 54):

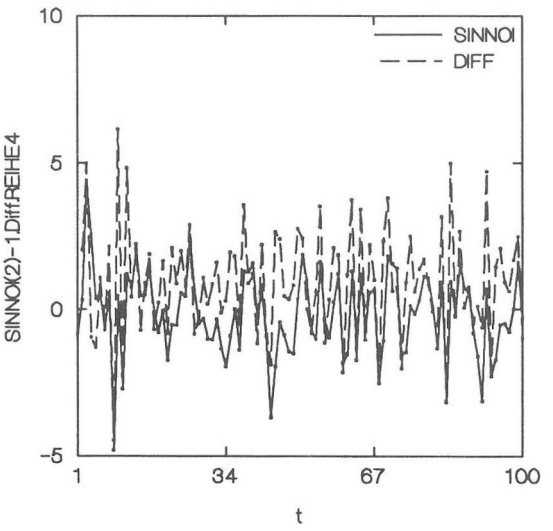


Abb. 54

Offensichtlich ist die Varianz der differenzierten Reihe (gestrichelt) beträchtlich grösser als die Varianz des "Signals"

$$\sin(2\pi/12t) + \varepsilon_t ,$$

das bei einer exakten Trendelimination zu erwarten wäre (3.5 gegenüber 2.1). Die Differenzenbildung bewirkt also eine Varianzerhöhung um 40%.

Die Abb. 55 zeigt die geschätzten Spektren des Signals, der hochpass- sowie der differenzengefilterten Reihe:

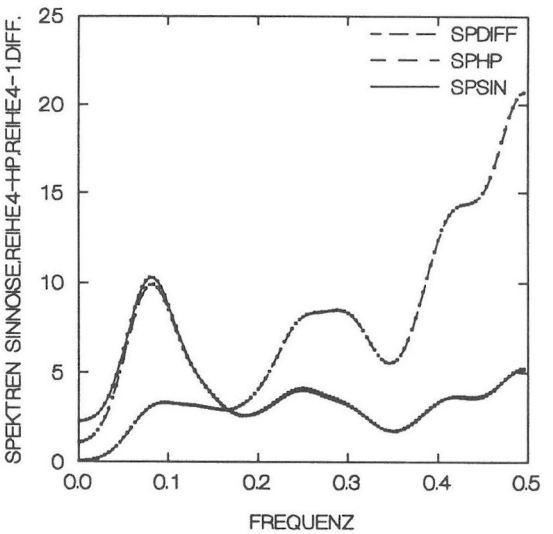


Abb. 55

Klar ersichtlich ist die starke Verschiebung spektraler Masse durch den Differenzenfilter in den Hochfrequenzbereich, was den erratischen Verlauf der 1. Differenzen erklärt. Ausserdem wird die Sinusschwingung so stark gedämpft, dass sie im Spektrum nicht mehr ausgewiesen wird.

Die an Beispielen demonstrierten Effekte einer Differenzenbildung lassen sich generell an der entsprechenden Amplitudenfunktion nachweisen. Diese lautet:

$$A(f) = [2(1 - \cos 2\pi f)]^{1/2} \ , \ f \in [0, 0.5] \ .$$

Abb. 56 zeigt den Verlauf dieser Funktion:

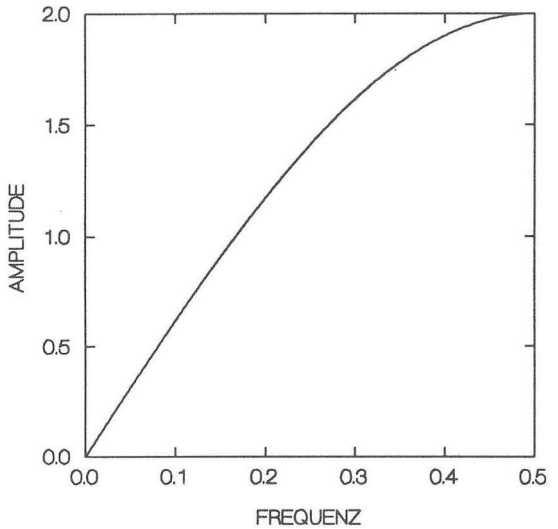


Abb. 56

Bis etwa $f = 0.2$ ist diese Amplitude kleiner als Eins (Schwingungen im Frequenzbereich $0 - 0.2$ werden also mehr oder weniger gedämpft), ab dieser Frequenz steigt sie monoton bis auf den Wert Zwei, was die Verschiebung spektraler Masse in den Hochfrequenzbereich bewirkt mit den (unerwünschten) Konsequenzen einer Varianzaufblähung und Aufrauhung. Da die Amplitude bei $f = 0$ den Wert Null annimmt, werden lineare Trends eliminiert.

Zusammenfassend lässt sich sagen:

Die in der Praxis häufig anzutreffende Trendelimination durch Bildung 1. Differenzen ist eine Operation, die mit gravierenden Nachteilen (unerwünschte Dämpfungs- bzw. Aufrauhungseffekte sowie Phasenverschiebungen) verbunden ist. Für Trendeliminationen sollten deshalb Verfahren benützt werden, welche diese vermeiden. Dazu eignen sich Filter, die unter Beachtung vorgegebener Kriterien entworfen worden sind.

5.3 Zwölfte Differenzen

Zwölfte Differenzen werden in der Praxis häufig zur Saisonbereinigung herangezogen. Die Amplitudenfunktion dieses Differenzenfilters lautet

$$A(f) = [2(1 - \cos 24\pi f)]^{1/2} \ , \ f \in [0, 0.5]$$

und zeigt folgenden Verlauf (vgl. Abb. 57):

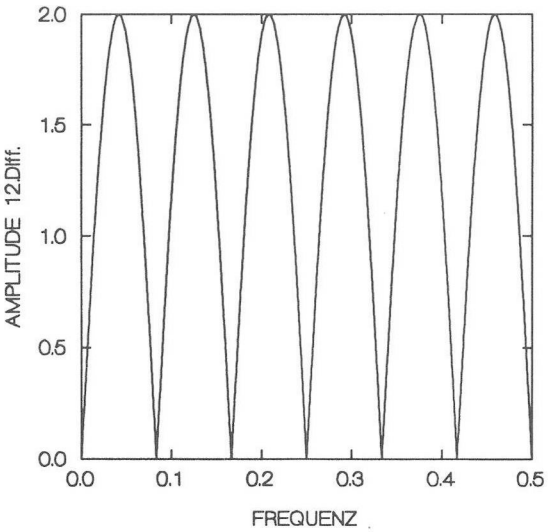


Abb. 57

Sie ist gleich Null in der Frequenz $f = 0$ sowie in *allen* saisonalen Frequenzen $f = 1/12, 2/12, 3/12, 4/12, 5/12, 6/12$. Demzufolge scheint dieser Filter sowohl für eine Trend- als auch eine Saisonbereinigung geeignet zu sein. Allerdings ist leicht nachzuweisen, dass 12. Differenzen einen ziemlich groben Saisonfilter darstellen. Wendet man ihn z.B. auf die folgende Reihe (vgl. Abb. 58) an,

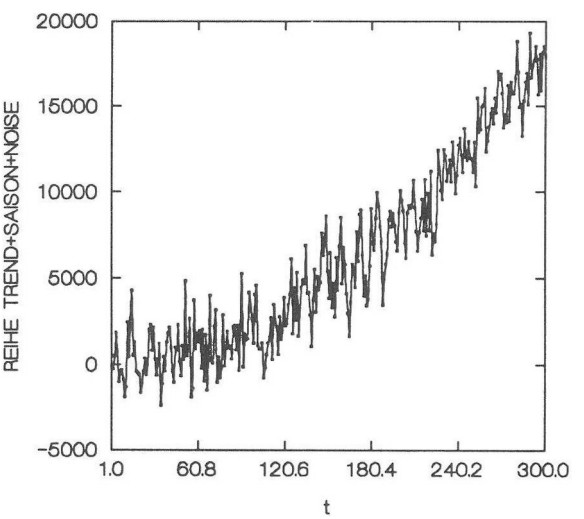


Abb. 58

der das Komponentenmodell zugrundeliegt

$$X_t = 2 + 0.9t + 0.2t^2 + 1000 \ (\sin 2\pi/12t + \varepsilon_t) \ ,$$

$$t = 1,...,300, \ mit \ \varepsilon_t \sim N \ (0, 2) \ ,$$

so ergibt sich folgender Filter-Output (vgl. Abb. 59):

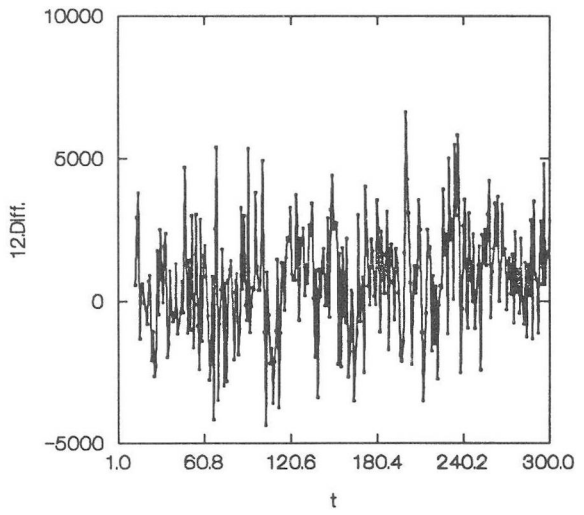


Abb. 59

Dass die saisonale Komponente nicht ganz ausgefiltert ist, sieht man an der geschätzten Autokorrelationsfunktion, die einen signifikant von Null verschiedenen Wert beim lag 12 aufweist (vgl. Abb. 60):

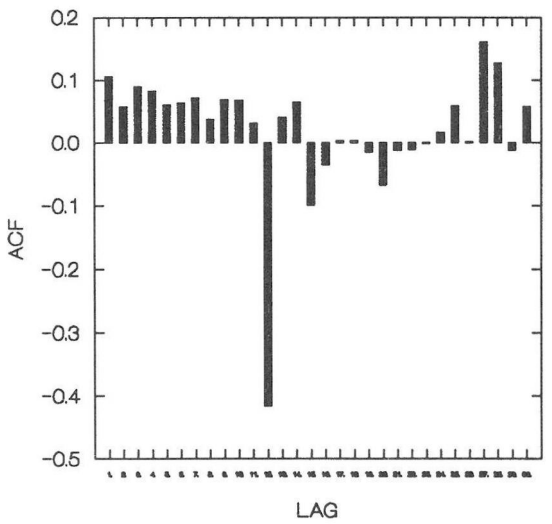


Abb. 60

Aber auch die Trendbereinigung der obigen Beispielreihe ist als unvollständig zu charakterisieren. Wie man aus dem zugrundeliegenden Komponentenmodell leicht nachrechnet, bleibt bei der Differenzenbildung nebst einer Konstanten der lineare Term 4.8t übrig. Schätzt man für den Filter-Output eine lineare Regression mit der Zeit als unabhängiger Variabler, dann ergibt sich für den Regressionskoeffizienten in der Tat ein signifikant von Null verschiedener Schätzwert von 4.43.

Wie man sich leicht überlegt, sind sowohl erste als auch zwölfte Differenzen nur in der Lage, *lineare* Trends auszufiltern.

Genau wie erste Differenzen verstärken auch zwölfte Differenzen den noise einer Reihe, d.h. es tritt der schon beschriebene Aufrauhungseffekt ein. Während die Varianz der Komponente

$$1000(\sin 2\pi/12t + \varepsilon_t)\ ,\ t = 1,...,300$$

2'348'160 beträgt, ergibt sich für die Varianz der 12. Differenzen der Wert 3'974'740, also eine Varianzaufblähung um ca. 69%.

Nicht selten werden in der Praxis auf eine Reihe sowohl erste als auch zwölfte Differenzen angewendet. Erstere zur Trend- und letztere zur Saisonbereinigung. Zwar lässt sich damit ein quadratischer Trend ausfiltern, jedoch sind die mit einer solchen Prozedur verbundenen Nachteile noch gravierender, als bisher schon dargelegt wurde. Die Amplitudenfunktion einer kombinierten 1. und 12. Differenz lautet

$$A(f) = [2(1 - \cos 2\pi f)]^{1/2}[2(1 - \cos 24\pi f)]^{1/2},\ f \in [0,0.5]$$

und zeigt folgenden Verlauf (vgl. Abb. 61):

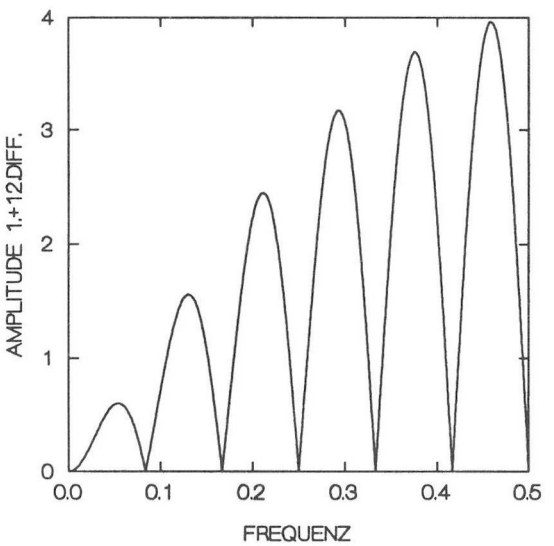


Abb. 61

Im Niederfrequenzbereich sind Dämpfungseffekte festzustellen, die von der 1. Differenz herrühren, und im Hochfrequenzbereich Aufrauhungseffekte, die aber jetzt stärker sind als bei jeweils einem Filter allein.

Bildet man erste und zwölfte Differenzen für obige Reihe, dann ergibt sich folgendes Resultat (vgl. Abb. 62):

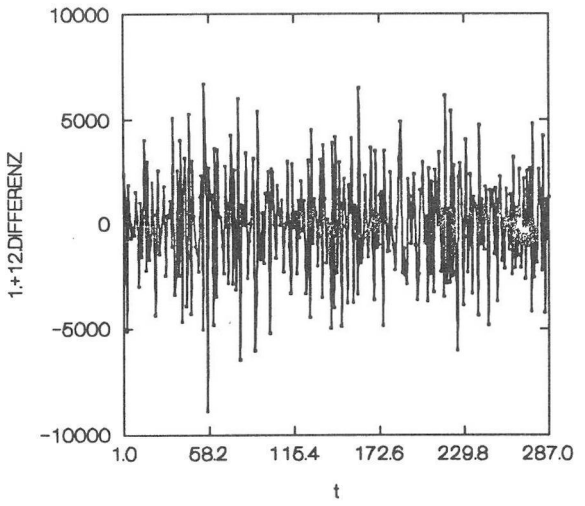


Abb. 62

Die Varianz dieses Outputs beträgt 7'115'820, was gegenüber der Varianz des "Signals" eine Aufblähung um 203% (!) und gegenüber der reinen 12. Differenz eine Aufblähung um 79% bedeutet.

Die Amplitudenfunktionen sowie die betrachteten Beispiele zeigen deutlich, dass Differenzenbildungen alles andere als harmlose Operationen sind. Insbesondere muss festgehalten werden, dass sie im allgemeinen weder für die Trend- noch für die Saisonbereinigung adäquate Werkzeuge sind.

5.4 Wachstumsraten

Abschliessend sei noch auf einen speziellen, aber in der Praxis sehr häufig angewandten Zeitreihen-Filter eingegangen, der eng mit der Differenzenbildung zusammenhängt. Es handelt sich um die Bildung von *Wachstumsraten*, wobei als Bezugsperiode meistens die Vorperiode gewählt wird

$$w_t = \frac{x_t - x_{t-1}}{x_t} .$$

Mit der Berechnung von Wachstumsraten hofft man, Informationen über die aktuelle Entwicklungstendenz einer Reihe (und damit über einen Wirtschaftsprozess) zu erhalten. Deshalb stellt sich die Frage, ob es generell überhaupt möglich ist, durch Vergleich *nur zweier* Beobachtungswerte diese Tendenz zu bestimmen. Die Antwort darauf ist ein *eindeutiges Nein*, und zwar aus folgendem Grund.

Zwar ist es richtig, dass Wachstumsraten zum Vormonat die interessierende aktuelle Entwicklungstendenz enthalten, aber diese Information wird verdeckt von den Wachstumsraten der übrigen Reihenkomponenten, insbesondere der noise-Komponente und der saisonalen Komponente. Eventuell auftretende Ausreisser akzentuieren diesen Tatbestand. Deshalb sagen Wachstumsraten gegenüber der Vorperiode nichts Relevantes über die aktuelle Entwicklunsgtendenz einer Reihe aus.

Dies ist nun im Detail zu begründen und darzulegen.

Ausgangspunkt der Überlegungen sei die simulierte Reihe aus Abb. 63, der folgendes Komponentenmodell zugrundeliegt:

$$X_t = 2 + 0.9t + \varepsilon_t \quad , \quad t = 1,2,...,100, \; mit \; \varepsilon_t \sim N(0,2)$$

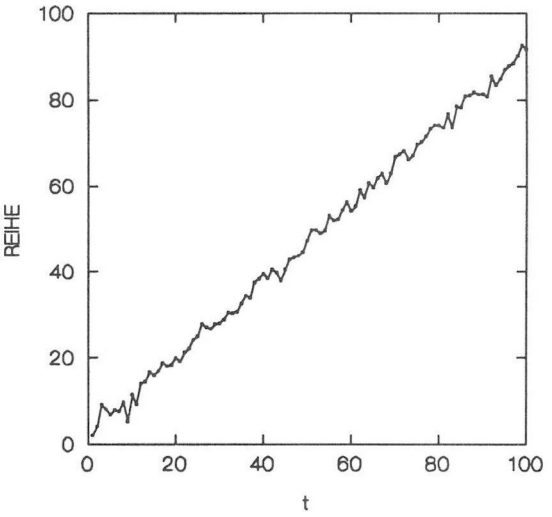


Abb. 63

Diese als völlig "harmlos" zu bezeichnende Reihe weist eine eindeutige Entwicklungstendenz auf, die natürlich durch den linearen Trend bestimmt wird. Kommt diese Entwicklungstendenz in den Wachstumsraten der Reihe, d.h. im Vergleich je zweier Reihenwerte zum Ausdruck?

Die Abb. 64 zeigt die Wachstumsraten dieser Reihe.

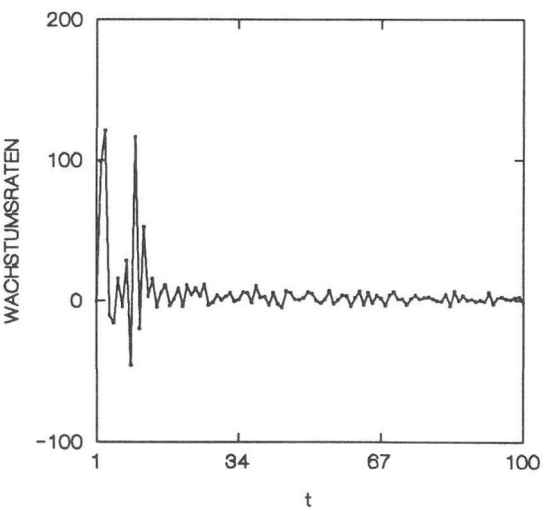


Abb. 64

Offensichtlich liefern die Wachstumsraten weder Informationen über die generelle noch über die aktuelle Entwicklungstendenz dieser Reihe, denn sie weisen eine Spannweite auf von -45.765 % bis 120.96% mit einer Varianz von 440.7! Sie sind deshalb als *völlig irreführend* zu bezeichnen, denn negative Wachstumsraten weist der lineare Trend nun mal keine auf. Diese lauten nämlich

$$WR_{Trend} = \frac{0.9}{1.1 + 0.9t}$$

und zeigen folgenden monoton fallenden (aber stets positiven) Verlauf (vgl. Abb. 65):

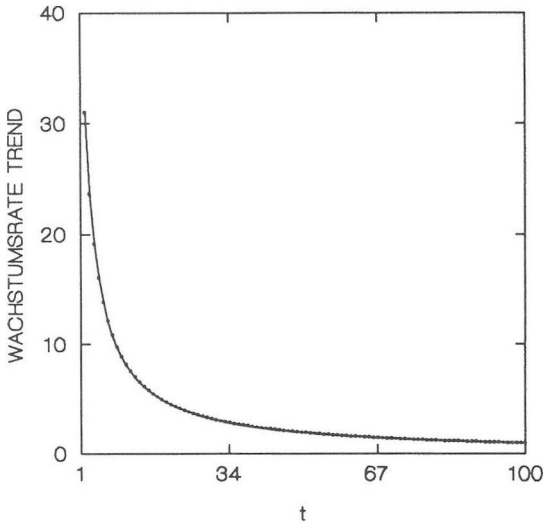


Abb. 65

Diese (tatsächlichen) Wachstumsraten liegen in einer Spanne von 0.99% und 31% mit einer Varianz von 22.4. Hier liegen demnach vollkommen andere Grössenordnungen vor als bei den oben aus der Reihe berechneten Wachstumsraten.

Die Ursache für das zu beobachtende erratische Verhalten von Wachstumsraten liegt darin, dass Wachstumsraten gegenüber der Vorperiode nicht-lineare Filter sind, die hinsichtlich ihrer Filtereigenschaften approximativ den Filtereigenschaften von 1. Differenzen gleichen und bereits dargestellt wurden.

Um zuverlässige Informationen über die Entwicklungstendenz einer Reihe zu erhalten, müssen mehr als jeweils nur zwei Reihenwerte herangezogen werden. Im betrachteten Beispiel etwa muss der noise ausgefiltert werden, was mit einem Tiefpass-Filter durchgeführt werden kann, wobei alle Beobachtungswerte verwendet werden. Ein exakter Tiefpass-Filter reproduziert praktisch den linearen Trend als korrekte Entwicklungsrichtung. Die daraus berechneten Wachstumsraten unterscheiden sich von den tatsächlichen nur geringfügig.

Die Abb. 66 zeigt zusammen die tatsächlichen Wachstumsraten (d.h. des linearen Trends) und die Wachstumsraten der Reihe sowie die auf der Basis der Tiefpass-Filterung berechneten Wachtumsraten.

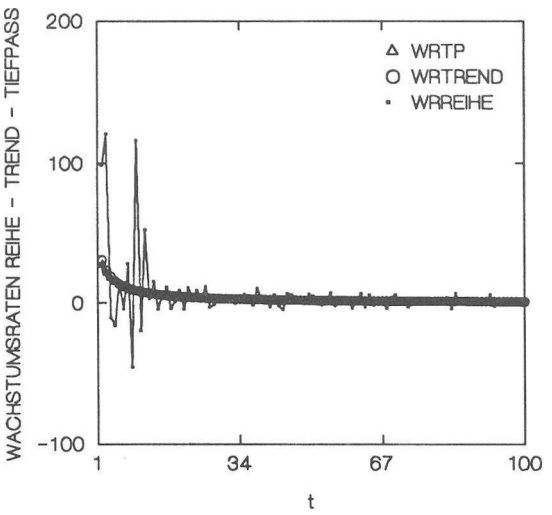


Abb. 66

Abschliessend werde noch ein besonders eindrückliches Beispiel für irreführende Informationen hinsichtlich der Entwicklungstendenz einer Reihe auf der Basis von Wachstumsraten betrachtet. Die Reihe in der Abb. 67 wird von folgendem Komponentenmodell erzeugt:

$$X_t = 0.5e^{0.05t} + \varepsilon_t \quad , \quad t = 1,2,...,100, \; mit \; \varepsilon_t \sim N(0,2)$$

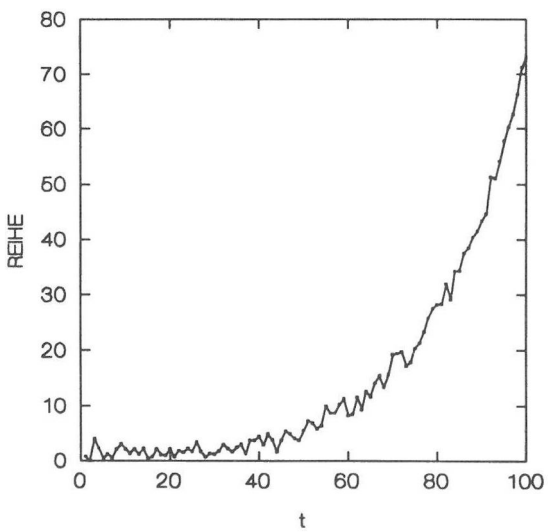


Abb. 67

Die Entwicklungsrichtung dieser Reihe wird von einer Exponentialfunktion bestimmt. Diese wächst über den gesamten Zeitbereich mit einer *konstanten* Wachstumsrate von 5.13%. Die Wachstumsraten, die aus der Reihe

berechnet werden, variieren aber im Bereich von -99.2% bis 58'441% mit einer Varianz von 34'137'500! Dies ist ein praktisch völlig irrelevantes, ja geradezu absurdes Resultat.

Als Fazit bleibt festzuhalten, dass die weitverbreitete Idee, durch Vergleich von benachbarten Beobachtungswerten Informationen über die Entwicklungsrichtung einer Reihe zu gewinnen, nicht als sinnvoll bezeichnet werden kann.

Häufig findet man in der Praxis neben den hier betrachteten Wachstumsraten auch solche, die auf einem Vorjahresvergleich basieren, also z.B. der Oktoberwert 1992 im Vergleich zum Oktoberwert 1991. Es lässt sich unschwer zeigen, dass auch solche Vorjahresvergleiche genau so negativ zu beurteilen sind wie die hier betrachteten Wachstumsraten.

Betrachten wir dazu eine Reihe, die sich additiv aus der letzten Reihe und der Saisonfigur aus Abb. 1 zusammensetzt. Die Saisonfigur ist grössenordnungsmässig gegenüber dem ausgeprägten Trend von völlig untergeordneter Bedeutung. Deshalb unterscheidet sich diese saisonale Reihe (Mittelwert: 15.04, Varianz: 343.3) kaum von der letzten (nicht-saisonalen) Reihe (Mittelwert: 15.06, Varianz: 347.7), wie das auch aus der Abb. 68, die beide Reihen zusammen darstellt, hervorgeht.

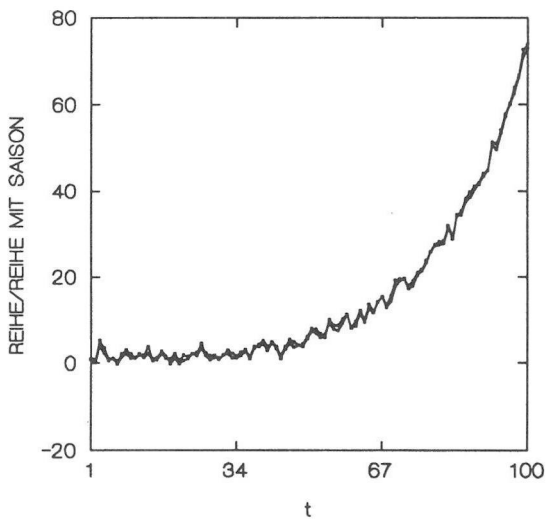


Abb. 68

Für die Wachstumsraten im Vorjahresvergleich ergibt sich (vgl. Abb. 69):

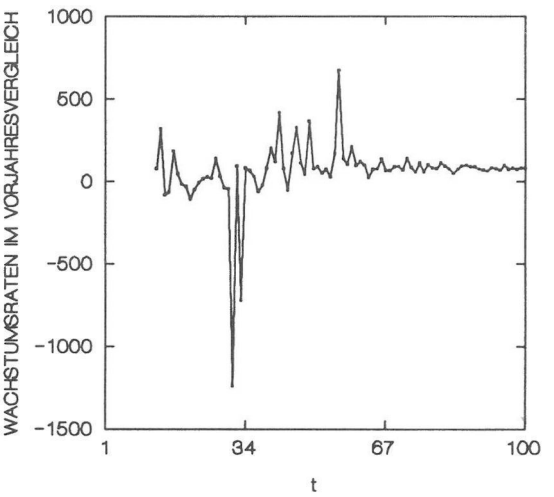


Abb. 69

Diese Wachstumsraten weisen eine Spannweite auf von -1'234.7% bis 672.4% mit einer Varianz von 37'839.5. Dieses Resultat ist nicht weniger absurd als das vorige. Die Wachstumsraten im Vorjahresvergleich liefern somit ebenso wenig Informationen über die Entwicklungstendenz dieser Reihe wie diejenigen zur Vorperiode.

Man mag gegenüber diesen Beispielen einwenden, dass die ganz extremen Ausschläge der Wachstumsraten vor allem durch die kleinen Reihenwerte am Reihenanfang bedingt sind und diese Situation nicht gerade als typisch für reale ökonomische Zeitreihen halten. Insbesondere dominiert bei diesen Werten der noise. Dieser Einwand ist sicher bis zu einem gewissen Grad berechtigt; er stellt die obigen Aussagen aber nicht grundsätzlich in Frage, schwächt sie jedoch für konkrete Zeitreihen mehr oder weniger stark ab. Ist insbesondere eine Reihe wenig noise-behaftet und weist nur geringe saisonale Ausschläge auf, bzw. liegt eine ziemlich starre Saisonfigur vor, dann führt eine Berechnung von Wachstumsraten nicht zu den demonstrierten extremen Resultaten. Im allgemeinen sind diese Bedingungen bei ökonomischen Reihen aber höchstens tendenziell erfüllt, so dass meistens bei ihren Wachstumsraten mit nicht zu vernachlässigenden Verzerrungen zu rechnen ist, die eine sichere Diagnose der Reihenentwicklung mindestens in Frage stellen.

6. Probleme bei der Konstruktion von Saisonfiltern

Im Vordergrund stehen heute in erster Linie die Eigenschaften der Transferfunktionen der durch die einzelnen Saisonbereinigungsverfahren definierten Filter. Es besteht in der Literatur Einigkeit darüber, welche Transferfunktionen als "ideal" für die Saisonbereinigung anzusehen sind. Die ideale Transferfunktion zeichnet sich dadurch aus, dass nur die saisonalen Frequenzen (z.B. gemessen in Zyklen/Zeiteinheit oder im Bogenmass) vom Filter tangiert werden, während alle anderen Frequenzen bzw. Frequenzbereiche, vor allem aber der Frequenzbereich unterhalb der ersten saisonalen

Frequenz, also z.B. bei Monatswerten der Bereich *[0, 1/12)*, unberührt bleiben sollen. Das heisst, mit dieser Forderung an eine Transferfunktion wird verlangt, dass der korrespondierende Saisonfilter die Eigenschaft hat, nur diejenige Variation einer Zeitreihe zu eliminieren, die auch als Saison interpretierbar ist. Diese Konstruktionsaufgabe für einen Saisonfilter hat zwei wichtige Nebenbedingungen zu beachten. Die gefilterte Reihe soll am aktuellen Rand möglichst stabil sein (Randstabilität), und zwischen der bereinigten Reihe und der beobachteten Zeitreihe sollen keine Phasenverschiebungen auftreten, oder diese sollen vernachlässigbar klein sein, d.h. grössenordnungsmässig unterhalb der zugrundeliegenden Grundzeitperiode liegen (Nullphase). Für die Annäherung zwischen der realisierten und der gewünschten idealen Transferfunktion sind verschiedene Kriterien zur Beurteilung der Güte einer Approximation denkbar. Die Lösung der Konstruktionsaufgabe ist dabei ein recht diffiziles und keinesfalls triviales Problem. Die Annäherung an die ideale Transferfunktion und die Einhaltung der Nebenbedingungen ergeben jedoch brauchbare und operable Kriterien zur Beurteilung der Qualität der alternativen Saisonbereinigungsverfahren.

Überraschenderweise ist jedoch bei der ganzen Diskussion der Beurteilungskriterien für Saisonbereinigungsverfahren bisher selten ausdrücklich auf das Problem eingegangen worden, dass Saisonbereinigungsverfahren überwiegend keine finalen Filter verwenden. Die Unterschiede zwischen der Bereinigung am aktuellen Rand und einer ex-post-Bereinigung sind - wie gesagt, abgesehen von der Frage der Randstabilität - bisher praktisch kaum diskutiert worden. Die Beurteilungskriterien bezogen sich auf eine ex-post-Bereinigung, die Eigenschaften der Randfilter blieben bis auf das erwähnte Stabilitätsproblem im allgemeinen ausserhalb einer Betrachtung. Nun wird aber Saisonbereinigung in der Regel zur Analyse aktueller Situationen betrieben und nicht in erster Linie für wirtschaftshistorische Zwecke. Unter diesem Aspekt sollten daher vor allem die Eigenschaften der Randfilter interessieren.

Dass sich exakte Amplitudenfunktion, exakte Nullphase und perfekte Randstabilität nicht gleichzeitig realisieren lassen, kann leicht eingesehen werden. Perfekte Randstabilität lässt sich nur mit einem Filter erzielen, der kausal (oder rekursiv) ist. Für einen solchen Filter der Ordnung (N,M) lautet die Filtergleichung:

$$y_t = \sum_{k=1}^{N} a_k y_{t-k} + \sum_{j=0}^{M} b_j x_{t-j}, \quad t = 1,...,T$$

mit

x_t : = *Filter-Input (d.h. die vorliegende Zeitreihe)*,

y_t : = *Filter-Output*,

a_k, b_j : = *nach Filtertyp festzulegende Koeffizienten*.

Es ist aus der Filtergleichung unmittelbar einsichtig, dass die einmal gefilterten Werte sich beim Hinzukommen von neuen Reihenwerten nicht mehr verändern, d.h. also, es liegt perfekte Randstabilität vor. Berechnet man die Frequenz-Antwortfunktion eines solchen Filters - was relativ einfach über eine z-Transformation der obigen Filtergleichung (für Werte auf dem Einheitskreis) erfolgen kann - dann zeigt sich, dass diese Funktion stets eine komplexe, rationale Funktion im Frequenzintervall *[0, 1/2]* ist. Der Betrag dieser Funktion ist die Amplitudenfunktion des Filters, und der Arcus-Tangens des Quotienten aus Imaginär- und Realteil dieser Funktion ist die zugehörige Phasenfunktion. Diese Phasenfunktion ist grundsätzlich nichtlinear. Somit kann es für einen perfekt randstabilen Filter keine exakte Nullphase geben.

Es lässt sich weiterhin zeigen, dass generell zwischen Amplitudenfunktion und Phasenfunktion eine gewisse Antinomie besteht. Je schärfer die Amplitudenfunktion ist (im Idealfall: exakt gleich Null in Stoppbändern und exakt gleich Eins in Durchlassbändern), umso grössere Werte nimmt die Phasenfunktion an.

In der folgenden Tabelle sind zunächst verschiedene Filtertypen aufgeführt, wobei danach unterschieden wird, ob die Filterung selbst im Zeitbereich oder im Frequenzbereich durchgeführt wird. Bei den Filtertypen, die dem ersten Fall zuzuordnen sind, werden ad-hoc-Filter, FIR-Filter und IIR-Filter unterschieden [34].

		Mit den Filtertypen realisierbare Eigenschaften					
Filterung erfolgt im	**Filtertyp**	**Exakte Amplituden-funktion**	**Exakte Null-phase**	**Perfekte Randstabi-lität**	**Den jeweiligen Filter-typen zu-zuordnende bekannte Verfahren**	**Finalität der verwendeten Filter**	**Verwen-dung eines Randfilters**
Zeitbereich	ad-hoc-Kon-struktionen	nein	teilweise, oft auch schwer kontrollierbar	nein	CENSUS X 11, BERLI-NER VER-FAHREN, ASA II, SABL	nicht gegeben	ja
Zeitbereich	FIR	nein	ja**)	-***)	keine		
Zeitbereich	IIR - ein-fache	nein	nein	ja	keine		
Zeitbereich	IIR - op-timale	nein	nein	ja	keine		
Zeitbereich	IIR - direkte	nein*)	nein*)	ja	STIER	gegeben	nein
Zeitbereich	IIR - spe-zielle	nein	nein*)	ja	SCHIPS/ STIER	gegeben	nein
Frequenz-bereich	nullphasig	ja	ja	nein*)	SCHMIDT/ STIER	nicht gegeben	ja

*) Streng theoretisch zutreffend, jedoch für die Praxis mit hinreichender Genauigkeit erreichbar.
**) Nur bei symmetrischer Implementation.
***) Hat bei diesem Filtertyp keine Bedeutung, da Filterung bis zum Rand nicht möglich.

In der 6. Spalte der Tabelle sind Verfahren aufgeführt, die den jeweiligen Filtertypen zugeordnet werden können, wobei aber kein Anspruch auf Vollständigkeit erhoben wird [35].

In der 7. Spalte wird das Problem angesprochen, ob die in einem Verfahren verwendeten Filter durchgehend bis zum letzten verfügbaren Zeitreihenwert eingesetzt werden können. Filter, bei denen das möglich ist, werden - wie bereits erwähnt - "final" genannt. Ist ein verwendeter Filter nicht final, dann muss er am aktuellen Rand einer Reihe durch einen "Randfilter" ergänzt werden, sonst ist eine Analyse bis zum letzten Reihenwert nicht möglich. Ab welchem Reihenwert ein Randfilter eingesetzt werden muss, ist natürlich von Verfahren zu Verfahren verschieden. Häufig unterscheiden sich Versionen ein und desselben Verfahrens unter anderem wesentlich darin, ab welchem Reihenwert der Randfilter eingesetzt wird, bzw. wie der jeweilige Randfilter im einzelnen konzipiert ist. Beim BERLINER VERFAHREN z.B. ist in der Version 1 ab dem Wert n-9 (n: letzter Reihenwert) ein Randfilter zur Schätzung der glatten Komponente einzusetzen, in der Version 2 ist dies der Wert n-11. Ausserdem sind die entsprechenden Randausgleichsfilter verschieden konzipiert [36]. Verfahren, die finale Filter einsetzen, weisen auch perfekte Randstabilität auf.

Neben den angeführten filtertheoretischen Prinzipien sind für die Konstruktion und Evaluation von Saisonbereinigungsverfahren auch methodologische Postulate zu berücksichtigen [37]. Solche sind ansonsten in der empirischen Forschung längst selbstverständlich. Leider stossen wissenschaftstheoretisch orientierte Überlegungen in diesem Zusammenhang vielfach immer noch auf Unverständnis. Je länger, je weniger können jedoch solche Aspekte von der Diskussion ausgeklammert bleiben. Es dürfte keine Übertreibung sein, wenn hier die Meinung vertreten wird, dass die oftmals geradezu chaotisch anmutenden Diskussionen

über Saisonbereinigungsprobleme zu einem wesentlichen Teil darauf zurückzuführen sind, dass methodologische Aspekte ignoriert werden. Dabei handelt es sich keinesfalls um "esoterische Dimensionen", die zu berücksichtigen wären. Vielmehr geht es lediglich um solche Aspekte, deren Beachtung eigentlich selbstverständlich sein sollte.

Ausgangspunkt sei die Feststellung, dass sich "Forschung" im allgemeinen als eine Folge von Entscheidungsprozessen darstellen lässt. Jedes Forschungsvorhaben, also auch die Konstruktion eines Saisonbereinigungsverfahrens, impliziert eine Fülle von Entscheidungen, die erhebliche Auswirkungen auf das Forschungsresultat haben können. Diese sind zur Sicherung der intersubjektiven Nachprüfbarkeit sorgfältig zu dokumentieren, eine Forderung, die ja auch z.B. vom "kritischen Rationalismus" erhoben wird. Diese Forderung zwingt zu einer Auseinandersetzung mit der Problematik von Begriffen, Definitionen sowie von Operationalisierungsstrategien, und dies ganz unabhängig von konkreten Bereinigungsverfahren.

Mit Begriffen lässt sich in der Forschung nur arbeiten, wenn sie definiert sind. Die Forderung nach intersubjektiver Vergleichbarkeit von Saisonbereinigungsverfahren zwingt zu einer Auseinandersetzung mit den Definitionsproblemen von Begriffen. Im wesentlichen unterscheidet man allgemein "Nominal"- und "Real"-Definitionen. Nominaldefinitionen sind nichts anderes als tautologische Umformungen. Das Definiendum ist bedeutungsgleich mit dem Definiens. Solche Definitionen haben keinen empirischen Informationsgehalt. Sie erleichtern bzw. ermöglichen erst die intersubjektive Kontrolle des Forschungsprozesses, denn nur was präzise definiert ist, kann auch eindeutig nachgeprüft werden. Im Gegensatz dazu sollen Realdefinitionen das "Wesen" einer Sache auf der Ebene der Sprache abbilden. Sie sind deshalb Behauptungen über die Beschaffenheit eines Phänomens und haben deshalb den gleichen Status wie empirische Hypothesen, können also richtig oder falsch sein. Realdefinitionen können aber in der Regel auch nicht vollständig sein, denn es ist wohl selten möglich, sämtliche Eigenschaften, die ein Sachverhalt aufweist, in die Definition aufzunehmen. Anspruch und Ziel von Realdefinitionen ist es ja gerade, nur das "Wesentliche" hervorzuheben. Aber was wirklich wesentlich ist, nicht was als wesentlich erachtet werden soll, kann nicht mit letzter Sicherheit bestimmt werden. Die Frage, welche der beiden Definitionsarten zu bevorzugen ist, hat in der Vergangenheit heftige wissenschaftstheoretische Kontroversen ausgelöst. Für die Praxis des Forschungsalltags sind allerdings solche Auseinandersetzungen wenig fruchtbar, da meistens gar nichts anderes übrigbleibt, als einen Mittelweg zwischen diesen beiden Ansätzen zu wählen. Man wird also versuchen, das "Wesentliche" eines Phänomens herauszufinden und dieses für alle weiteren Überlegungen im Sinne einer Nominaldefinition festzuschreiben.

Überblickt man die Definitionen von "Saison" bei den einzelnen Verfahren, so findet man meistens ansatzweise solche "nominalisierten" Realdefinitionen oder realdefinitorisch "angereicherten" Nominaldefinitionen. Nur "ansatzweise" deswegen, weil häufig die offengelegten Definitionen ziemlich vage sind, also im strengen Sinn des Wortes gar keine Definitionen vorliegen. Meistens handelt es sich nur um verbale Charakterisierungen des zu untersuchenden Phänomens "Saison". Auch wenn ein Verfahren ohne eine klare, präzise und damit vergleichbare Definition von "Saison" arbeitet, heisst das natürlich nicht, dass ein solches Verfahren ohne Definition auskäme. Vielmehr ist eine solche durch die einzelnen Arbeitsschritte "implizit" gegeben, aber möglicherweise nicht eruierbar.

Damit lässt sich ein erstes Gütekriterium formulieren:

- bei Saisonbereinigungsverfahren muss eine präzise, eindeutige Definition von "Saison" explizit formuliert sein.

Präzise Definitionen sind jedoch nur eine notwendige, keinesfalls aber eine hinreichende Bedingung für methodologisch sauberes Arbeiten. Da Definitionen der Ebene der Sprache angehören, kann mit ihrer Hilfe die notwendige Verknüpfung von Begriffen und realen Sachverhalten nicht vollzogen werden. Dies kann erst die Operationalisierung leisten. Darunter ist die Angabe oder Festlegung derjenigen Vorgehensweisen (oder Forschungsoperationen) zu verstehen, durch die entscheidbar wird, ob und in welchem Ausmass der mit dem Begriff bezeichnete Sachverhalt in der Realität vorliegt. Nachdem also eine präzise Definition von "Saison" vorliegt, müssen diejenigen technischen Operationen explizit dargelegt werden, die es erlauben, darüber zu entscheiden, ob und in welchem Ausmass "Saison" (im definierten Sinn) vorliegt. Konkret heisst das z.B., konstruiere einen Filter mit den und den Eigenschaften oder optimiere eine Zielfunktion mit genau spezifizierten Eigenschaften usw. Dabei ist stets zu prüfen, ob die gewählte Operationalisierung logisch gültig ist, d.h. ob Definition und Operationalisierung den gleichen Bedeutungsgehalt haben. Neben der logischen ist die empirische Gültigkeit einer Operationalisierung zu überprüfen. Dabei geht es um den Nachweis, dass eine Operationalisierung auch tatsächlich in der Realität das "erfassen" kann, was theoretisch beabsichtigt ist. Logische und empirische Validität brauchen nicht gleichzeitig voll gegeben zu sein. So kann z.B. die theoretische Spezifikation eines Filters logisch 100% valide sein, der tatsächlich realisierte Filter kann aber aus mannigfachen Gründen von seiner theoretischen Idealform abweichen und ist deshalb nicht mehr 100% empirisch valide. Diese Überlegungen führen zu drei weiteren Gütekriterien:

- bei Saisonbereinigungsverfahren muss die Operationalisierung des Begriffes "Saison" eindeutig erkennbar sein

- bei Saisonbereinigungsverfahren muss entscheidbar sein, ob und inwieweit die Operationalisierung des Begriffes "Saison" logisch gültig ist

- bei Saisonbereinigungsverfahren muss entscheidbar sein, ob und inwieweit die Operationalisierung des Begriffes "Saison" empirisch gültig ist.

Mit dem Kriterium der logischen Validität direkt verknüpft ist ein Tatbestand, der für manche Verfahren konstitutiv ist. Gemeint ist der Sachverhalt, dass diese Verfahren für den Benutzer Optionen bereitstellen, z.B. nachträgliche Glättung der bereinigten Reihe, wobei das Ausmass der Glättung durch einen vom Benutzer innerhalb gewisser Grenzen vorzugebenden Parameterwert bestimmt wird. Dabei ist die Frage zu stellen, ob das Drehen an gewissen "Schrauben" noch verträglich ist mit der ursprünglich gewählten Definition von "Saison". Das führt zum nächsten Kriterium:

- bei Saisonbereinigungsverfahren, die für den Benutzer Optionen bereitstellen, muss nachprüfbar sein, ob die wählbaren Optionsalternativen kompatibel sind mit der ursprünglich gewählten Definition von "Saison".

Selbstverständlich sollte das folgende Kriterium sein:

- bei Saisonbereinigungsverfahren ist eine Offenlegung der beabsichtigten Zielsetzung erforderlich. Insbesondere ist nachzuweisen, dass die gewählte Operationalisierung damit kompatibel ist.

Manche Bereinigungsverfahren sind iterativ aufgebaut, d.h. gewisse Phasen werden mehrfach durchlaufen. Das führt leicht zu "pragmatischen Zirkeln". Soll etwa die glatte Komponente bestimmt werden, dann muss dazu die Saison- und Restkomponente ermittelt werden. Um aber die Saisonkomponente zu bestimmen, muss zuerst eine - wenn auch vorläufige - Bestimmung der glatten Komponente durchgeführt werden. Entscheidend dabei ist, dass die vorläufigen "Phasen" die endgültige Schätzung in einer kaum zu kontrollierenden Weise präjudizieren. Deshalb ist als weiteres Kriterium zu fordern:

- bei der Konstruktion von Saisonbereinigungsverfahren sind pragmatische Zirkel möglichst zu vermeiden.

Dies scheinen die wichtigsten Kriterien zu sein, die sich für den vorliegenden Kontext aus der Beachtung methodologischer Prinzipien empfehlen. Insbesondere sind die Kriterien, die mit Definition, Operationalisierung und Zielsetzung zusammenhängen, als unverzichtbar zu bezeichnen. Weitere mögen hinzukommen, es wird hier kein Anspruch auf Vollständigkeit erhoben. Es sei hier nur noch ein Kriterium genannt, das besonders bei Verfahrensvergleichen zu berücksichtigen ist. Es ergibt sich daraus, dass bei den einzelnen Verfahren eine Reihe von Hypothesen postuliert werden, die in der Regel nicht empirisch überprüfbar sind. Im Vergleich wird man deswegen ein Verfahren, das mehr nicht-testbare Hypothesen enthält als ein anderes - unter sonst gleichen Voraussetzungen -, als inferior einstufen. Deshalb soll als letztes Kriterium formuliert werden:

- von zwei Saisonbereinigungsverfahren ist - ceteris paribus - vom methodologischen Standpunkt dasjenige als inferior anzusehen, das die grössere Anzahl von nicht-testbaren Hypothesen enthält.

Eine Auswahl wichtiger methodologischer Postulate ist in der nachfolgenden Tabelle zu finden.

Verfahren	**Evidenz der Zielsetzung des Verfahrens**	**Präzise Definition von "Saison" liegt vor**	**Erkennbare eindeutige Operationalisierung von "Saison"**	**Vorliegen von pragmatischen Zirkeln**
CENSUS X 11	nein	nein	nein	ja
BERLINER VERFAHREN	?	?	?	ja
ASA	nein	nein	nein	ja
SABL	nein	nein	nein	ja
STIER	ja	ja	ja	nein
SCHIPS / STIER	ja	ja	ja	nein
SCHMIDT / STIER	ja	ja	ja	nein

7. Zur Konstruktion von Tiefpass-Filtern

7.1 Filter im Zeitbereich: FIR-Filter und IIR-Filter

Bei vielen Problemen der Zeitreihenanalyse spielen solche Filter eine wichtige Rolle, für welche die Beziehung zwischen Filter-Input und Filter-Output in Form einer linearen Differenzengleichung

$$y_t + a_1 y_{t-1} + \ldots + a_M y_{t-M} = b_0 x_t + b_1 x_{t-1} + \ldots + b_N x_{t-N}$$

darstellbar ist. Offensichtlich handelt es sich dabei um kausale (oder nicht-antizipatorische) Filter, denn der Filter-Output hängt ausser von früheren Outputwerten nur vom gegenwärtigen Input und von früheren Inputs ab. Es lässt sich leicht zeigen, dass die Transferfunktion eines solchen Filters eine rationale Funktion ist, die ausschliesslich von den "Filter-Koeffizienten" $a_1, a_2, \ldots, a_M, b_1, b_2, \ldots, b_N$ bestimmt wird.

Anhand der Transferfunktion eines solchen Filters bzw. seiner Filtergleichung lassen sich nun zwei grundlegende Filtertypen unterscheiden. Ist $a_1 = a_2 = \ldots = a_M = 0$, dann besitzt die Transferfunktion nur Nullstellen, aber keine Pole. Der Filter-Output hängt nicht vom Output früherer Zeitpunkte ab. Solche Filter werden als *FIR* (=*F*inite-*I*mpulse-*R*esponse)-Filter bezeichnet. Synonyme sind *Transversalfilter* und *nicht-rekursive* Filter.

Die Filtergleichung für FIR-Filter lautet somit

$$Y_t = \sum_{r=0}^{N} b_r x_{t-r} \ .$$

Im Gegensatz dazu stehen Filter, die neben Nullstellen auch Pole besitzen, d.h. die Koeffizienten $a_1, a_2, \ldots, a_M$ sind nicht alle gleich Null. Solche Filter werden als *IIR* = (*I*nfinite-*I*mpulse-*R*esponse)-Filter bezeichnet. Im Hinblick auf ihre Filtergleichung ist auch die Bezeichnung *rekursive* Filter gebräuchlich.

Ein FIR-Filter ist dann bestimmt, wenn die Filter-Koeffizienten b_0 , $b_1, \ldots, b_N$ festliegen. Zum Beispiel können einfache FIR-Filter dadurch bestimmt werden, indem die Filterkoeffizienten beliebig gewählt werden. Eine naheliegende Möglichkeit ist etwa

$$b_0 = b_1 = b_2 = \ldots = b_N = \frac{1}{N+1} \ ,$$

d.h. dieser Filter ist ein einfacher *gleitender Durchschnitt der Länge (oder Ordnung) N + 1* mit

$$\sum_{r=0}^{N} b_r = 1 \ .$$

Somit stellen die bekannten einfachen gleitenden Durchschnitte (etwa der Länge N = 12) FIR-Filter dar. Es lässt sich zeigen, dass diese Filter Tiefpass-Filter sind.

Von *Filter-Design* spricht man dann, wenn die Filterkoeffizienten nicht ad hoc festgelegt (wie etwa bei den einfachen gleitenden Durchschnitten), sondern unter Beachtung gewisser Konstruktionsprinzipien abgeleitet werden. Letztere werden im allgemeinen so spezifiziert, dass eine vorgegebene Amplituden- und/oder Phasenfunktion "bestmöglich" approximiert wird. Methoden des Filter-Designs für FIR-Tiefpass-Filter können hier nicht dargestellt werden [38].

FIR-Filter und damit auch FIR-Tiefpass-Filter haben die günstige Eigenschaft, dass sie eine lineare Phase aufweisen. Falls sie symmetrisch implementiert werden, sind sie *nullphasig*. Allerdings ist damit keine Analyse bis zum aktuellen Rand möglich. Nachteilig ist jedoch, dass mit FIR-Filtern generell "scharfe" Amplitudenfunktionen nur bei relativ hoher *Filterordnung* N möglich sind, was bei kurzen Zeitreihen, wie sie in der Praxis häufig anzutreffen sind, rasch zu nicht realisierbaren Filtern führt.

Für IIR-Filter lautet die Filtergleichung

$$\sum_{k=0}^{M} a_k y_{t-k} = \sum_{r=0}^{N} b_r x_{t-r}, \quad a_0 = 1 \ .$$

Da die Transferfunktionen von IIR-Filtern im Gegensatz zu denjenigen von FIR-Filtern neben Nullstellen auch Pole besitzen, ergeben sich bei der Filterkonstruktion zusätzliche *Freiheitsgrade*. Der Verlauf der Frequenz-Antwortfunktion bzw. Amplitudenfunktion kann nicht nur über die Plazierung der Nullstellen (die nach *substanzwissenschaftlichen* Erfordernissen zu erfolgen hat), sondern auch über die Plazierung der Pole in der komplexen Ebene *gesteuert* werden. Diese Möglichkeit erhöht die Flexibilität beim Filter-Design ganz beträchtlich. Insbesondere kann man bei geschickter Plazierung dieser Pole schon bei kleiner Filterordnung (d.h. kleinem M und N) sehr scharfe Amplitudenfunktionen erzielen, insbesondere auch scharfe Tiefpass-Filter.

Beispielsweise erhält man für M = N = 1 die Filtergleichung

$$y_t = b_0 x_t - a_1 y_{t-1}$$

bzw.

$$y_t = \beta x_t + \alpha y_{t-1}$$

mit

$$\beta := b_0 \ , \ \alpha := -a_1 \ .$$

Es kann gezeigt werden, dass dieser IIR-Filter ein Tiefpass-Filter ist, falls der Parameter α im Intervall (0,1) gewählt wird. Insbesondere erhält man für Parameterwerte, die dicht bei Eins liegen, Tiefpass-Filter, deren Amplitudenfunktionen schärfer verlaufen als diejenigen der obigen einfachen FIR-Tiefpass-Filter.

Der Vorteil grösserer Flexibilität beim Filter-Design ist jedoch auch mit einem Nachteil verbunden. IIR-Filter weisen immer eine *nicht-lineare* Phase auf. Damit ist die Konstruktion nullphasiger IIR-Filter prinzipiell nicht möglich. Allerdings zeigt es sich, dass die Phasenverschiebung zwischen Filter-Input und Filter-Output häufig nur wenige (z.B. zwei oder drei) Zeiteinheiten beträgt, was für manche Anwendungen unproblematisch ist. Erachtet man jedoch eine exakte Nullphase für unabdingbar, dann können IIR-Filter nicht verwendet werden. Design-Methoden für IIR-Filter und speziell IIR-Tiefpass-Filter können hier nicht besprochen werden. Dazu sei wieder auf die oben angeführte Literatur verwiesen.

7.2 Filter im Frequenzbereich

Bei den bisher betrachteten Filtern wird die jeweilige Filter-Design-Aufgabe im *Frequenzbereich* formuliert und auch in diesem Bereich gelöst. Die Filterung selbst wird jedoch im *Zeitbereich* vollzogen (via Filtergleichung). An den beiden Filtertypen FIR und IIR zeigt sich klar die Antinomie zwischen Amplitude und Phase. Lineare (und damit auch Null-)Phase ist nur bei FIR-Filtern möglich, scharfe Amplituden erfordern jedoch IIR-Filter. Somit lassen sich beide wünschenswerte Eigenschaften nicht gleichzeitig realisieren, jedenfalls nicht mit den bisher diskutierten Design-Ansätzen.

Dies ist jedoch möglich, wenn auch die Filterung einer Zeitreihe (oder eines Signals) im Frequenzbereich vorgenommen wird. Da Signale immer im Zeitbereich gegeben sind, ist eine Filterung im Frequenzbereich nur dann möglich, wenn diese in den Frequenzbereich transformiert werden können. Dies erlaubt die Technik der sogenannten "Diskreten Fourier-Transformation" (DFT).

Sei x_t , $t = 0,1,...,T\text{-}1$ ein gegebenes Signal der Länge T, dann ist die DFT dieses Signals durch

$$DFT(x_t) = X(e^{i\lambda}) := \sum_{t=0}^{T-1} x_t e^{-i\lambda t}, \quad \lambda \in [0,2\pi]$$

gegeben.

$X(e^{i\lambda})$ repräsentiert den vollständigen Frequenzgehalt des Signals x_t. Da $X(e^{i\lambda})$ für ein Kontinuum von Frequenzen definiert ist, scheint eine direkte Berechnung nicht möglich zu sein. Ausserdem scheint eine Reproduktion des ursprünglichen Signals aus $X(e^{i\lambda})$ mit Hilfe der inversen Transformation ebenfalls nicht möglich zu sein. Man kann jedoch leicht zeigen, dass eine Kenntnis von $X(e^{i\lambda})$ für nur T Frequenzen genügt, um das originale Signal mittels einer sogenannten "Inversen Diskreten Fourier-Transformation" (IDFT) wiederzugewinnen.

Sei

$$X_m = \sum_{k=0}^{T-1} x_k e^{-i(2\pi/T)km}, \quad m = 0,1,...,T-1$$

die DFT von x_t für die Frequenzen $0, 2\pi/T, 4\pi/T,..., (T\text{-}1)2\pi/T$. Die komplexe Folge X_0 , $X_1,...,X_{T-1}$ stellt die Transformation des Signals x_0 ,...,x_{T-1} in den Frequenzbereich dar. Die IDFT von X_m ist

$$IDFT(x_m) := \frac{1}{T} \sum_{m=0}^{T-1} X_m e^{i(2\pi/T)mt}, \quad t = 0,1,...,T-1.$$

Praktisch wird häufig eine "Verfeinerung" der DFT zweckmässig sein, vor allem für kleines T. Für solche Fälle mag sich nämlich die Abtastung des Frequenzbereiches allzu grob erweisen.

Definiert man nun ein Signal durch

$$\tilde{x}_t := \begin{cases} x_t, & t = 0,1,...,T-1 \\ 0, & T \le t \le L-1 \end{cases},$$

d.h. erweitert man das gegebene Signal durch Hinzufügen von L-T Nullen, dann ist die DFT von $\tilde{x}_t$ gegeben durch

$$\tilde{X}_m := \sum_{k=0}^{L-1} \tilde{x}_k e^{-i(2\pi/L)km}, \quad m = 0,1,...,L-1 ,$$

d.h. $\tilde{X}_m$ ist für die Frequenzen $\lambda_m = (2\pi/L)m$, $m = 0,1,...,L\text{-}1$ definiert. Da aber $\tilde{x}_t = 0$ ist für $t \ge T$, kann geschrieben werden

$$\tilde{X}_m := \sum_{k=0}^{T-1} \tilde{x}_k e^{-i(2\pi/L)km}, \quad m = 0,1,...,L-1 .$$

Diese Technik der Verlängerung eines endlichen Signals durch Hinzufügen von "Nullen" erlaubt eine beliebig feine Abtastung des Frequenzintervalles bei der Berechnung der DFT.

Da die gesamte Frequenzinformation eines Signals in seiner DFT enthalten ist, liegt es nahe, eine gewünschte Amplitudenfunktion *direkt* vorzugeben und die DFT damit zu multiplizieren. Eine Rücktransformation mit Hilfe der IDFT ergibt dann die gefilterte Reihe. Soll z.B. eine Reihe tiefpassgefiltert werden, dann ordnet man den DFT-Werten im Passband den Wert Eins und im Stoppband den Wert

Null zu. Zwischen Pass- und Stoppband muss jedoch ein Übergangsband vorgesehen werden, da sonst hochfrequente Oszillationen im Filter-Output auftreten (Gibb'sches Phänomen). Diese rühren von Unstetigkeiten in der Amplitudenfunktion her, wie sie bei einer idealen rechteckigen Tiefpass-Amplitude auftreten. Es empfiehlt sich daher - auch aus einem weiteren noch darzulegenden Grund - ein "möglichst glattes" Übergangsband vorzusehen. Da schon sehr kurze Übergangsbänder ausreichend sind, um das Gibb'sche Phänomen zu vermeiden, lassen sich auf diese Weise praktisch ideale Filterpässe realisieren. Dies trifft natürlich auch auf andere Filtertypen zu wie z.B. Hoch- oder Bandpass-Filter oder Filter mit Nicht-Standard-Amplitudencharakteristiken. Von besonders praktischer Bedeutung ist aber, dass die Phasenfunktionen bei Filterung im Frequenzbereich stets identisch Null im gesamten Frequenzbereich sind. Exakte Amplitudenfunktionen und Nullphasen lassen sich also gleichzeitig realisieren.

Allerdings zeigt ein so konstruierter und implementierter Filter einen kaum interpretierbaren Verlauf an den beiden Reihenrändern. Das *Randproblem* lässt sich durch die *Endlichkeit* praktischer Zeitreihen erklären. Durch Vergleich zweier Filter-Outputs, die das Resultat eines unendlich bzw. endlich langen Filter-Input-Signals sind, kann ein "Filter-Fehler" definiert werden, dessen zeitlicher Verlauf sich anschaulich interpretieren lässt.

Zwischen Filter-Input und Filter-Output bestehen für die beiden Signale die Beziehungen

$$y_t = \sum_{k=-\infty}^{\infty} h_k x_{t-k}$$

und

$$\hat{y}_t = \sum_{k=-\infty}^{\infty} h_k I_{[1,T]}(t-k) x_{t-k} ,$$

wobei $I_{[1,T]}(t)$ ein *Daten-Fenster* ist mit der Eigenschaft

$$I_{[1,T]}(t) = \begin{cases} 1 & \textit{für } 1 < t < T \\ 0 & \textit{sonst} . \end{cases}$$

$\hat{y}_t$ ist der Filter-Output für das *endliche* Signal $x_t, t = 1,...,T$, das ein Segment aus dem unendlich langen Signal x_t , $...,1,0,1,...$ darstellt. $\hat{y}_t$ ist der "wahre" Filter-Output, da er einem *unendlich* langen Input-Signal entspricht. Ein Vergleich von y_t und $\hat{y}$ *für* $t = 1,...,T$ zeigt einen Einfluss der Empfindlichkeit eines Signals auf das Filterresultat. Für die Differenz der beiden Outputs ergibt sich

$$y_t - \hat{y}_t = \sum_{k=-\infty}^{t-T-1} h_k x_{t-k} + \sum_{k=t}^{\infty} h_k x_{t-k}$$

$$= \sum_k h_k x_{t-k}, \quad \begin{cases} k \le t-T-1 \\ k \ge t . \end{cases}$$

Definiert man einen quadratischen Filterfehler durch

$$FF_{t,T} := |y_t - \hat{y}_t|^2, \quad t = 1,...,T,$$

dann ergibt sich

$$FF_{t,T} := |y_t - \hat{y}_t|^2 = \sum_k |h_k|^2 \; |x_{t-k}|^2 .$$

Nimmt man an, dass die Fourier-Transformierte des unendlich langen Signals x_t existiert, dann gilt

$$X(e^{i\lambda}) = \sum_{t=-\infty}^{\infty} x_t e^{-i\lambda t}$$

bzw.

$$x_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\lambda}) e^{i\lambda t} d\lambda .$$

Damit ergibt sich

$$FF_{t,T} \le \sum_k |h_k|^2 \; |\frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\lambda}) e^{i\lambda(t-k)} d\lambda|^2 .$$

Mit

$$\left\| X(e^{i\lambda}) \right\|_{\infty} := \sup_{\lambda} X(e^{i\lambda})$$

folgt

$$FF_{t,T} \le \left\| X(e^{i\lambda}) \right\|_{\infty} \sum_k |h_k|^2 .$$

Die zeitvariable obere Fehlerschranke ist sowohl vom Signal (via Fourier-Transformierte bzw. Spektraldichte) als auch vom Filter selbst bestimmt. Je schneller die $|h_k|$ gegen Null konvergieren, umso kleiner ist $FF_{t,T}$. Die Konvergenzgeschwindigkeit hängt von der Gestalt der Amplitudenfunktion des Filters ab. Amplitudenfunktionen mit Unstetigkeitsstellen führen zu geringeren Konvergenzgeschwindigkeiten als Amplitudenfunktionen, die "glatt" verlaufen. Schmale Übergangsbänder wirken sich negativ auf die Konvergenzgeschwindigkeit aus. Enge Passbänder wirken in derselben Richtung. Aus dem filterspezifischen Fehleranteil

$$\delta_{t,T} = \sum_k |h_k|^2$$

$$= \sum_{k=-\infty}^{t-T-1} |h_k|^2 + \sum_{k=t}^{\infty} |h_k|^2$$

folgt nun, dass für kausale IIR-Filter $h_k = 0$ ist für $k < 0$. Deshalb ist $\delta_{t,T}$ *minimal für* $t = T$, d.h. für den rechten Rand. Für nicht-kausale Filter - dazu gehören nullphasige - ist $\delta_{t,T}$ *maximal für* $t = 1$ *und* $t = T$, d.h. an den beiden Rändern verläuft $\delta_{t,T}$ für solche Filter U-förmig. Das erklärt das zu beobachtende Verhalten des Filter-Outputs an beiden Reihenrändern.

Betrachtet man den Filterfehler $FFT_{t,T}$, dann gibt es grundsätzlich zwei Möglichkeiten, diesen zu reduzieren. Die erste besteht darin, nur solche Filter zu verwenden, deren Filtergewichte im Zeitbereich h_k sehr schnell abklingen. Das impliziert allerdings, dass man auf exakte Amplitudenfunktionen verzichtet. Kommt dies nicht in Frage, dann verbleibt als einzige Möglichkeit, das Supremum des Fourier-Spektrums, d.h. $||X(e^{i\lambda})||_\infty$, möglichst klein zu machen. Selbstverständlich darf dazu das gegebene Signal x_t nur in einer Weise manipuliert werden, die bezüglich des definitiven Filter-Outputs neutral ist. Das kann in einer relativ einfachen Weise geschehen. Soll z.B. ein Signal tiefpassgefiltert werden, dann wird - nach Spezifikation der Amplitudenfunktion - das jeweilige Fourier-Spektrum aus $x_t, t = 1,...,T$ geschätzt und diejenige Frequenz bestimmt, bei der das Supremum des Spektrums liegt. Liegt diese im Passband, dann wird eine geeignet bestimmte harmonische Komponente dieser Frequenz aus dem Signal eliminiert und das so modifizierte Signal - dessen Fourier-Spektrum nun ein kleineres Supremum besitzt als das Ausgangssignal und damit auch einen kleineren Filterfehler - in der oben beschriebenen Weise gefiltert. Dem Filter-Output wird die anfänglich eliminierte Komponente wieder hinzugefügt. Das erübrigt sich natürlich, wenn die Supremums-Frequenz im Stoppband liegt. Dieser hier nur in den Grundzügen und vereinfacht dargestellte Prozess kann *iterativ* durchgeführt werden. Auf diese Weise kann der Filterfehler iterativ verkleinert - wenn auch nicht vollständig beseitigt - werden. Die Filter-Outputs sind dann in der Regel auch am Rand interpretierbar.

7.3 Globale und lokale Filter

Die im Abschnitt 7.2 betrachteten Tiefpass-Filter können als *globale* Filter bezeichnet werden. Global deswegen, weil sie den Frequenzgehalt der *gesamten* zu filternden Reihe verarbeiten. Globale Filter weisen jedoch eine Eigenschaft auf, die bei manchen Anwendungen unter Umständen störend in Erscheinung tritt. Diese Eigenschaft besteht darin, dass beim Hinzukommen neuer Zeitreihenwerte auch die bisher gefilterten Werte sich nicht nur am Rand verändern können. Dies ist darauf zurückzuführen, dass sich der Frequenzgehalt einer Zeitreihe durch eine Aktualisierung möglicherweise merklich verändern kann.

Sieht man einmal vom aktuellen Rand ab, so sind die weiter zurückliegenden Veränderungen in aller Regel sehr klein. Jedoch können bei gewissen Anwendungen auch kleine Veränderungen zu Interpretationsproblemen führen. In solchen Fällen empfiehlt es sich, einen *lokalen* Filter zu verwenden.

Einen *lokalen* Tiefpass-Filter erhält man dadurch, dass als Filter-Input nicht die gesamte Zeitreihe verwendet wird. Vielmehr wird die Filterung über einen *gleitenden Stützbereich* ungerader Länge vorgenommen, wobei die gleiche Amplitudenfunktion wie beim globalen Tiefpass-Filter verwendet wird. Dadurch wird nur der jeweilige (zeitabhängige) *lokale Frequenzgehalt* durch den Filter verarbeitet. Der jeweils *gefilterte Wert* ist derjenige, welcher der *Mitte des Stützbereiches* zuzuordnen ist. Untersuchungen haben gezeigt, dass ein Stützbereich der Länge L = 13 verwendet werden kann. Am aktuellen Reihenrand weist ein derartiger lokaler Filter nur *sechs* Werte auf, die sich bei einer Aktualisierung noch verändern können. Alle anderen sind *absolut stabil*.

Anwendungsbeispiele für lokale Tiefpass-Filter sind in Kapitel IV zu finden. Insbesondere werden dort auch Vergleiche mit globalen Filtern sowie mit der glatten Komponente nach dem CENSUS-Verfahren vorgenommen.

Sowohl für globale als auch für lokale Filter ist PC-kompatible Software vorhanden (vorzugsweise für schnelle DOS-Rechner) [39].

Literaturhinweise

1) Vgl. NERLOVE M., GRETHER D.M. und CAVALHO J.L., Analysis of Economic Time Series. A Synthesis, New York 1979

2) Vgl. dazu PERSONS W.M., 'Indices of Business Conditions', in: The Review of Economic Statistics 1 (1919), S. 5-107

3) Vgl. WALD A., Berechnung und Ausschaltung von Saisonschwankungen, Wien 1936

4) Vgl. z.B. SCHÄFFER K.A., 'Vergleich der Effizienz von Verfahren zur Saisonbereinigung einer Zeitreihe', in: K.A. Schäffer (Hrsgb.), Beiträge zur Zeitreihenanalyse, Göttingen 1976, S. 83-104

5) Vgl. zu diesen Ausführungen SchÄFFER K.A, 'Methoden zur Zerlegung von Zeitreihen - Ein Überblick', in: K.A. Schäffer, P. Miebach, K. Edel (Hrsgb.), Probleme von Saisonbereinigungsverfahren. Eine vergleichende Studie, im Druck

6) Vgl. ebenda

7) Auf diese Problematik hat bereits WALD hingewiesen. Vgl. WALD A., Berechnung und Ausschaltung von Saisonschwankungen, Wien 1936

8) Vgl. CLEVELAND W.P., TIAO G.C., 'Decomposition of Seasonal Time Series: A Model for the Census X-11 Program', in: Journal of the American Statistical Association 71 (1976), S.531-567

9) Vgl. zu diesen Ausführungen SCHÄFFER K.A., 'Methoden zur Zerlegung von Zeitreihen - Ein Überblick', a.a.O.

10) Vgl. HEILER S., 'Überlegungen zu einem statistischen Modell einer wirtschaftlichen Zeitreihe und einem daraus resultierenden Verfahren', in: Das Berliner Verfahren. Ein Beitrag zur Zeitreihenanalyse, DIW-Beiträge zur Strukturforschung 7 (1976), S. 19-43

11) Vgl. die Literaturhinweise bei FÖRSTER W., 'Zur Historie der Saisonbereinigung und der Komponentenzerlegung bei ökonomischen Zeitreihen', in: K.A. Schäffer, P. Miebach, K. Edel (Hrsgb.), Probleme von Saisonbereinigungsverfahren. Eine vergleichende Studie, im Druck

12) Vgl. dazu STIER W., 'Saisonbereinigungsverfahren. Ein Überblick', in: OR-Spektrum 7 (1985), S. 63-79

13) Vgl. etwa BAMBERG G., BAUR F., Statistik, 5. Auflage, München-Wien 1987

14) Vgl. dazu etwa FLASKÄMPER P., Allgemeine Statistik, Hamburg 1953

15) Vgl. z.B. FALKNER H.D., 'The Measurement of Seasonal Variation', in: Journal of the American Statistical Association 19 (1924), S. 167-179

16) Vgl. SHISKIN J. et al. The X-11 Variant of the Census Method II Seasonal Adjustment Program, Bureau of Census, Technical Paper No. 15, 1967

17) Vgl. etwa WALLIS K.F., 'Seasonal Adjustment and Revision of Current Dates', in: Journal of the Royal Statistical Society 145 (1982), S. 75-86, und zu den Diagnosemöglichkeiten vgl. u.a. FINDLEY D.F. et al., 'Sliding-Spans Diagnostics for Seasonal and Related Adjustments', in: Journal of the American Statistical Association 85 (1990), S. 345-355

18) Vgl. dazu WIENER N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, New York 1949

19) Vgl. CLEVELAND W.P., Analysis and Forecasting of Seasonal Time Series, Ph. D. Thesis, Department of Statistics, University of Wisconsis 1972, und CLEVELAND W.P., TIAO G.C., 'Decomposition of Seasonal Time Series: A Model for the Census X-11 Program', in: Journal of the American Statistical Association 71 (1976), S. 581-587

20) Vgl. DAGUM E.B., 'Modelling, Forecasting and Seasonally Adjusting Economic Time Series with the X-11 ARIMA Method', in: The Statistician 27 (1978), S. 203-216

21) Vgl. etwa DAGUM E.B., 'The Effects of Asymmetric Filters on Seasonal Factor Revisions', in: Journal of the American Statistical Association 77 (1982), S. 732-738

22) Vgl. FINDLEY D. et al., 'Toward X-12-ARIMA', in: Proceedings of the Bureau of the Census Annual Research Conference 4 (1988), S. 591-622

23) Vgl. dazu SCOTT S., 'An Extended Review of the X-11-ARIMA Seasonal Adjustment Package', in: International Journal of Forecasting 8 (1992), S. 627-633

24) Vgl. CELVELAND W.P. et al., The SABL Seasonal and Calendar Adjustment Package. Bell Laboratories Murray Hill (1982)

25) Vgl. HEILER S., 'Theoretische Grundlagen des BERLINER VERFAHRENS', in: W. Wetzel (Hrsgb.), Neuere Entwicklungen auf dem Gebiet der Zeitreihenanalyse, Göttingen 1970, S. 67-94

26) Für Einzelheiten vgl. NOURNEY M., SÖLL H., 'Analyse von Zeitreihen nach dem BERLINER VERFAHREN, Version 3', in: K.A. Schäffer (Hrsgb.), Beiträge zur Zeitreihenanalyse, Göttingen 1976, S. 131-164

27) Vgl. AKAIKE H., 'Seasonal Adjustment by a Bayesian Modelling', in: Journal of Time Series Analysis 1 (1980), S. 1-13, und KITAGAWA G., GERSCH W., 'A Smoothness Priors State-Space Modelling of Time Series with Trend and Seasonality', in: Journal of the American Statistical Association 79 (1984), S. 378-389, und PAULY R., 'Zerlegung und Analyse ökonomischer Zeitreihen', in: Statistische Hefte 66, S. 291, sowie SCHLICHT E., A Seasonal Adjustment Principle and a Seasonal Adjustment Method Derwed from this Method', in: Journal of the American Statistical Association 76 (1981), S. 374-378

28) Vgl. WHITTAKER E.T., 'On a New Method of Graduation', in: Proceedings of the Edinburgh Mathematical Society 41 (1923), S. 63-75

29) Vgl. LOVELL M.C., 'Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis', in: Journal of the American Statistical Association 58 (1963), S. 993-1010

30) Vgl. insbesondere SCHÄFFER K.A., 'Vergleich der Effizienz von Verfahren zur Saisonbereinigung einer Zeitreihe', in: K.A. Schäffer (Hrsgb.), Beiträge zur Zeitreihenanalyse, Göttingen 1976, S. 83-104, und SCHÄFFER K.A., 'Analyse saisonaler Zeitreihen und Bewertung der Effizienz von Diagnosemethoden', in: K.A. Schäffer, P. Miebach, K. Edel (Hrsgb.), Probleme von Saisonbereinigungsverfahren, Eine vergleichende Studie, im Druck

31) Vgl. dazu GRETHER D.M., NERLOVE M., 'Some Properties of "Optimal" Seasonal Adjustment', in: Econometrica 38 (1970), S. 682-703, und SCHIPS B., STIER W., 'Zum Problem der Saisonbereinigung ökonomischer Zeitreihen', in: Metrika 21 (1974), S. 65-81

32) Vgl. WALLIS K.F., 'Seasonal Adjustment and Revision of Current Dates', in: Journal of the Royal Statistical Society, Ser. A 145 (1982), S. 75-86

33) Vgl. STIER W., Konstruktion und Einsatz von Digital-Filtern zur Analyse und Prognose ökonomischer Zeitreihen, Opladen 1978, und Kapitel IV in dieser Arbeit

34) Unter einem FIR-Filter versteht man einen Filter mit endlicher Impulsantwortfunktion. IIR-Filter weisen eine unendliche Impulsantwortfunktion auf. Unter der Impulsantwortfunktion versteht man dabei den Filter-Output, der sich ergibt, wenn als Filter-Input eine Impulsfunktion verwendet wird. Die (diskrete) Impulsfunktion ist eine Funktion, die für einen Zeitpunkt t_0 den Wert Eins annimmt und für alle anderen Zeitpunkte $t \neq t_0$ den Wert Null. Vgl. auch Kapitel IV in dieser Arbeit.

35) Versionen dieser Verfahren werden nicht gesondert aufgeführt, da diese für die hier interessierenden Problemkreise keine neuen Aspekte beisteuern.

36) Für Einzelheiten vgl. NULLAU B., 'Probleme der praktischen Anwendungen des BERLINER VERFAHRENS', in: W. Wetzel (Hrsgb.), Neuere Entwicklungen auf dem Gebiet der Zeitreihenanalyse, Göttingen 1970, S. 95-129, und NOURNEY M., SÖLL H., 'Analyse von Zeitreihen nach dem BERLINER VERFAHREN, Version 3', in: K.A. Schäffer (Hrsgb.), Beiträge zur Zeitreihenanalyse, Göttingen 1976, S. 131-164

37) Vgl. STIER W., 'Methodologische Überlegungen zur Güte und zum Vergleich von Saisonbereinigungsverfahren', in: K.A. Schäffer, P. Miebach, K. Edel (Hrsgb.), Probleme von Saisonbereinigungsverfahren, Eine vergleichende Studie, im Druck

38) Vgl. dazu etwa RABINER L.R., GOLD B., Theory and Application of Digital Signal Processing, Englewood Cliffs 1975

39) Nähere Angaben sind bei STIER W., FEW, Hochschule St. Gallen zu erhalten.

III. Eigenschaften des CENSUS-X-11-Verfahrens im Frequenzbereich

1. Einleitung

Saisonbereinigungsverfahren können unter sehr verschiedenen Aspekten beurteilt und verglichen werden. Hier sollen zwei Eigenschaften des CENSUS-X-11-Verfahrens betrachtet werden, denen eine grosse praktische Bedeutung zukommt. Dabei handelt es sich einerseits um die *Phaseneigenschaften* und andererseits um die *Amplitudeneigenschaften*
dieses Saisonbereinigungsverfahrens.

Beim erstgenannten Aspekt geht es um die Frage, ob und gegebenenfalls um wieviele Zeiteinheiten Phasenverschiebungen zwischen Original- und bereinigter Reihe auftreten. Unter *bereinigt* kann dabei sowohl die saisonbereinigte Reihe als auch die *glatte Komponente* verstanden werden. Diese Frage ist insbesondere für den *aktuellen Rand* einer Zeitreihe von Bedeutung, da die bereinigten Werte an dieser Stelle in der Praxis üblicherweise für Diagnosen des Systemzustands herangezogen werden. Sollte etwa ein Verfahren dort eine nennenswerte Phasenverschiebung aufweisen, so wäre dies bei der Diagnose zu berücksichtigen. Die bereinigten Werte müssten um die entsprechenden Zeiteinheiten (meistens Monate oder Quartale) *rückdatiert* werden, und es wäre zu überlegen, ob sie sich dann überhaupt noch als Diagnosebasis eigneten. Ideal ist natürlich eine Nullphase oder eine quasi-Nullphase, d.h. eine Phase, die kleiner ist als die zugrundeliegende Zeiteinheit (Monat, Quartal usw.).

Wie die nachstehenden Ausführungen deutlich werden lassen, sollte die Phaseneigenschaft eines Verfahrens aber nicht unabhängig von seiner Amplitudeneigenschaft - also dem zweiten oben erwähnten Aspekt - betrachtet werden (und vice versa). Dies hängt nicht nur damit zusammen, dass zwischen Phase und Amplitude eines Verfahrens direkte Beziehungen existieren, sondern auch damit, dass eine isolierte Betrachtung eines einzelnen Aspektes möglicherweise zu Fehlschlüssen hinsichtlich der *Güte* eines Verfahrens führt.

Kenntnisse über die Amplitudencharakteristik eines Verfahrens erlauben auch Rückschlüsse über die einem Verfahren zugrundeliegenden, aber leider vielfach überhaupt nicht oder nur unpräzise explizierten Komponenten-Definitionen. Mit Hilfe der Amplitudencharakteristik lässt sich insbesondere die Frage klären, ob ein Verfahren eine bestimmte Komponente, z.B. in Reihenmitte gleich definiert wie am Reihenende. Sollte das nicht der Fall sein, so impliziert dies, dass die jeweiligen bereinigten Werte strenggenommen eigentlich nicht vergleichbar sind, da sie auf unterschiedlichen Definitionen beruhen.

2. Theoretische Überlegungen zur Bestimmung von Amplituden- und Phasenfunktion des CENSUS-X-11-Verfahrens

Eine analytische Bestimmung von Phasen- und Amplitudencharakteristik ist für die meisten Verfahren, die in der Praxis der Saisonbereinigung verwendet werden, insbesondere auch für das CENSUS-X-11-Verfahren nicht möglich. Dies kann verschiedene Gründe haben. So verwenden manche Verfahren nicht-lineare Filter (wenn vielleicht auch nicht durchgehend, so doch in manchen wichtigen Verfahrensschritten), oder sie setzen Filter iterativ ein, was letztlich zu einem Gesamtfilter führen kann, dessen Eigenschaften analytisch nur noch schwer, wenn überhaupt, bestimmbar sind. Insbesondere aber werden bei den meisten Verfahren für die Bereinigung des aktuellen Randes andere Prozeduren verwendet wie in Reihenmitte. Dies läuft darauf hinaus, dass insgesamt von zeitvariablen Filtern auszugehen ist. Speziell Randbereinigungen erweisen sich einer analytischen Behandlung kaum zugänglich. Für das CENSUS-X-11-Verfahren treffen praktisch alle genannten Punkte zu.

Aus diesen Gründen wird hier ein empirisch-experimenteller Zugang zur Bestimmung der Amplituden- und Phasencharakteristik dieses Verfahrens gewählt, der auf geeigneten Test-Signalen beruht. Es sei hier betont, dass empirische Untersuchungen auf der Basis konkreter Zeitreihen, die prinzipiell auch möglich wären, bewusst nicht vorgenommen werden, weil gegen die dabei erzielten Resultate stets eingewendet werden könnte, dass sie reihenspezifisch seien.

Die Bestimmung geeigneter Test-Signale ergibt sich nun aus den nachstehenden Überlegungen [1].

Sei X_t das gesuchte Test- (oder Input-) Signal und Y_t das Output-Signal des Saisonbereinigungsverfahrens, dann kann der Zusammenhang zwischen diesen beiden Signalen folgendermassen formuliert werden

$$F : X_t \Rightarrow Y_t , \quad t = 1,...,T .$$

Dabei bezeichnet F einen Filter, der das Input-Signal in das Output-Signal transformiert, d.h. das ganze Bereinigungsverfahren wird bei dieser Darstellung quasi als ein einziger Filter interpretiert.

Am Reihenrand ist diese Darstellung dahingehend zu modifizieren, dass dort der Filter üblicherweise zeitvariabel ist. (Dies gilt allerdings auch für den Reihenanfang, was aber praktisch von geringer Bedeutung ist, da die bereinigten Werte dort höchstens für eine retrospektive Zielsetzung interessant sind.) Hier gilt also

$$F_t : = X_t \Rightarrow Y_t ,$$

Der Beginn dieses Bereiches ist bei den verschiedenen Verfahren natürlich im allgemeinen unterschiedlich. Nicht selten lässt auch er sich nur empirisch ermitteln.

Schickt man das Test-Signal

$$X_t = \sin(\lambda t + \delta)$$

durch den Filter F bzw. F_t, dann ergibt sich folgender Filter-Output

$$F_t : X_t \Rightarrow Y_t = A_t(\lambda)\sin[\lambda t + \delta + \Phi_t(\lambda)] ,$$

wobei $\Phi_t(\lambda)$ die Phasenverschiebung zwischen Input und Output ist, die vom Filter F bzw. F_t, d.h. dem Bereinigungsverfahren, verursacht wird. Die Phasenfunktion $\Phi_t(\lambda)$ ist - wie bereits angemerkt - insbesonders am aktuellen Rand interessant. Hinzuzufügen ist hier, dass sie in Reihenmitte bei den meisten Verfahren praktisch gleich Null ist, da dort im Gegensatz zum Rand symmetrische Filter verwendet werden können.

$A_t(\lambda)$ bezeichnet die Amplitudenfunktion des Filters F bzw. F_t. Sie zeigt, in welchem Ausmass die verschiedenen Frequenzen bzw. Frequenzbänder zugeordnete "*power*" vom Verfahren eliminiert bzw. abgeschwächt oder auch verstärkt wird. Ein *gutes* Saisonbereinigungsverfahren sollte nur die *saisonale "power"* eliminieren, aber keine anderen Frequenzbänder tangieren.

Wie aus der obigen Input-Output-Beziehung zu entnehmen ist, beruht die hier vorgestellte empirische Analyse auf der Theorie der linearen Filter. Das impliziert, dass die erzielten Resultate für Verfahren, die nicht-lineare Filter verwenden, lediglich als Approximationen zu verstehen sind.

Da sowohl $A_t(\lambda)$ als auch $\Phi_t(\lambda)$ für alle Zeitpunkte t = 1,2, ...,T und alle Frequenzen λ bestimmt werden sollen, sind zwei Filter-Outputs Y_t^1 und Y_t^2 notwendig. Man erhält sie durch eine geeignete Wahl des Parameters δ.

Mit $\delta = \delta_1$ und $\delta = \delta_2$ ergeben sich die folgenden Outputs

$$\begin{aligned} Y_t^{1,2} &= A_t(\lambda)\sin[\lambda t + \Phi_t(\lambda) + \delta^{1,2}] \\ &= A_t(\lambda)\{\sin[\lambda t + \Phi_t(\lambda)]\cos\delta^{1,2} + \cos[\lambda t + \Phi_t(\lambda)]\sin\delta^{1,2}\} \\ &= \begin{cases} A_t(\lambda)\sin[\lambda t + \Phi_t(\lambda)] , & \delta^1 = 0 \\ A_t(\lambda)\cos[\lambda t + \Phi_t(\lambda)] , & \delta^2 = \pi/2 , \end{cases} \end{aligned}$$

woraus sich ergibt

$$\begin{aligned} Y_t^1 &= A_t(\lambda)\sin[\lambda t + \Phi_t(\lambda)] \\ Y_t^2 &= A_t(\lambda)\cos[\lambda t + \Phi_t(\lambda)] , \end{aligned}$$

d.h. die beiden gesuchten Test-Signale enthalten jetzt Zeitpunkte a

$$\begin{aligned} Y_t^1 &= \sin\lambda t \\ Y_t^2 &= \cos\lambda t . \end{aligned}$$

Aus den obigen Beziehungen lässt sich nun die empirische Amplituden- bzw. Phasenfunktion ableiten. Es ist

$$A_t(\lambda) = \sqrt{[Y_t^1]^2 + [Y_t^2]^2} .$$

Dividiert man die beiden Filter-Outputs Y_t^1 und Y_t^2, ergibt sich

$$\frac{Y_t^1}{Y_t^2} = \tan[\lambda t + \Phi_t(\lambda)] .$$

Die empirische Phasenfunktion ist deshalb

$$\Phi_t(\lambda) = \arctan \frac{Y_t^1}{Y_t^2} - \lambda t .$$

Für die praktische Bestimmung der Phasenfunktion ist allerdings die normale arctan-Funktion nicht geeignet, da sie nur im Intervall $[-\pi/2 , \pi/2]$ definiert ist. Deshalb wird dazu die im Intervall $[-\pi, \pi]$ definierte ATAN2-Funktion verwendet.

Zur leichteren Interpretation der ermittelten Phasenfunktionen ist es ratsam, diese in den Zeitbereich zu transformieren, da Phasen in radian nicht unbedingt sehr informativ sind.

3. Experimentelles Design

Zur empirischen Bestimmung von Amplitude und Phase des CENSUS-X-11-Verfahrens wurden jeweils 60 Test-Signale $X_t^1 = \sin\lambda t$ und $X_t^2 = \cos\lambda t$ der Länge T = 200 erzeugt, wobei das Frequenzintervall $[0 , \pi]$ an den Stellen $\pi k/60$, $k = 1,2,...,60$ abgetastet wurde. Amplitude und Phase wurden sowohl für die Reihenmitte als auch für den Reihenrand bestimmt, wobei für die erstere die Zeitpunkte t = 100 bis t = 170 und für den letzteren die Zeitpunkte t = 171 bis t = 200 gewählt wurden.

Dabei wurde davon ausgegangen, dass überwiegend Reihen auf der Basis von Monatswerten bereinigt werden. Eine analoge Untersuchung könnte natürlich auch z.B. für Quartalsdaten durchgeführt werden.

Die Amplitudenfunktion eines *idealen* Saisonbereinigungsverfahrens sollte die Verlaufsform eines *idealen* multiplen Kerbenfilters aufweisen, da eine derartige Funktion garantiert, dass nur die "*power*" saisonaler Frequenzen ausgefiltert wird. Eine solche Amplitudenfunktion lässt sich z.B.

mit einem rekursiven Filter realisieren [2], wie sie in nachstehender Abb. 70 wiedergegeben ist. Sie unterscheidet sich nur marginal von der idealen Amplitudenfunktion.

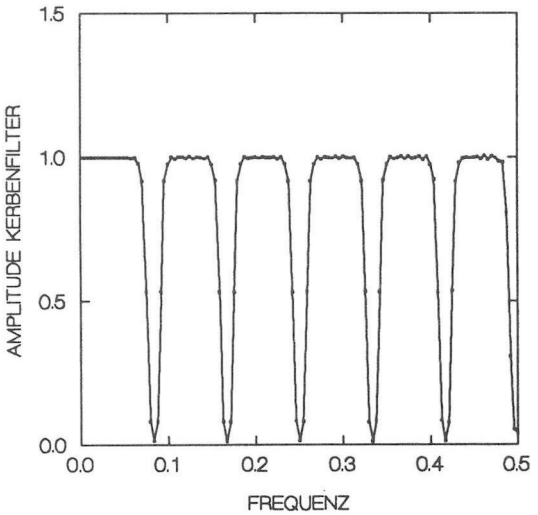


Abb. 70

Wie die meisten Verfahren liefert auch das CENSUS-X-11-Verfahren nebst saisonbereinigten Werten eine *glatte Komponente*. Diese sollte weder saisonale "*power*" noch *noise* enthalten. Idealerweise sollte eine glatte Komponente das Resultat eines exakten Tiefpass-Filters sein. Die Amplitudenfunktion eines solchen Filters nimmt den Wert Eins an im Frequenzintervall *[0, 1/12)* und ist sonst Null. Diese soll für die glatte Komponente als *standard of reference* dienen.

4. Resultate für die Amplitudenfunktionen

Die erzielten Resultate werden im folgenden graphisch dargestellt. Die Abb. 71-74 zeigen die empirisch ermittelten Amplitudenfunktionen des CENSUS-X-11-Verfahrens für die Reihenmitte und den Reihenrand für die *Saisonbereinigung*, wobei diese Funktionen sowohl für die *additive* als auch für die *multiplikative* Version ermittelt wurden.

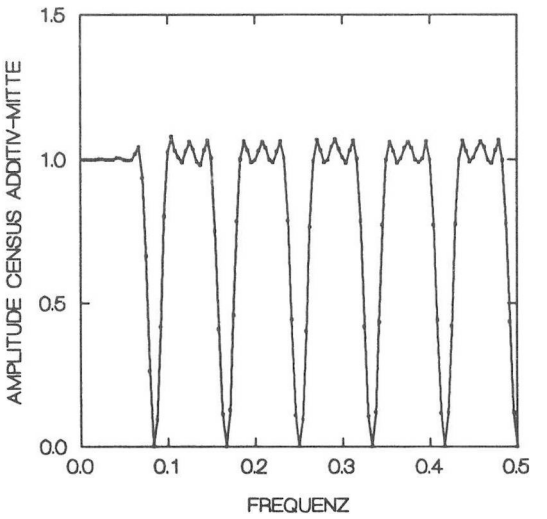


Abb. 71

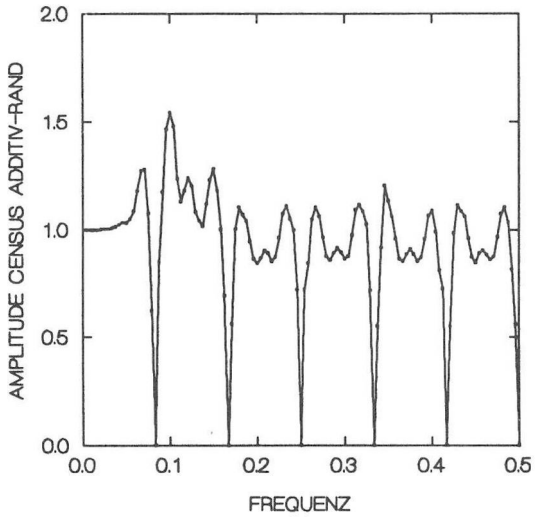


Abb. 72

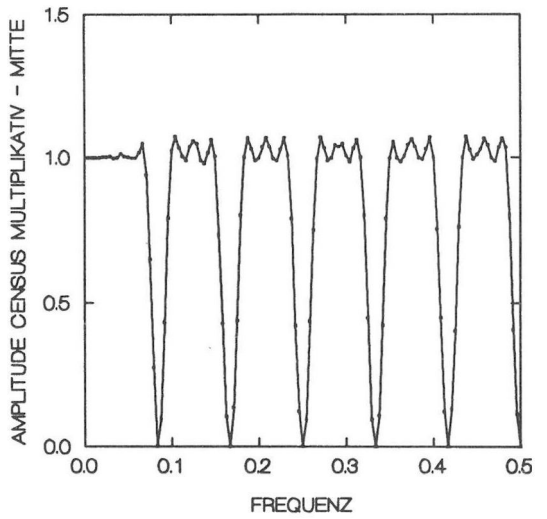


Abb. 73

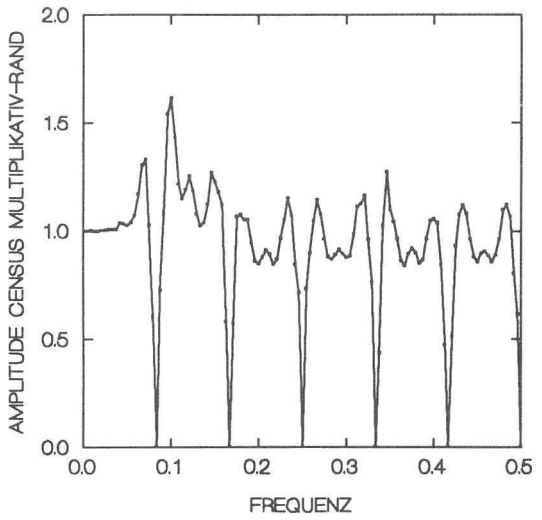


Abb. 74

In den Abb. 75 und 76 werden die Amplituden für Reihenmitte und -rand jeweils zusammen dargestellt.

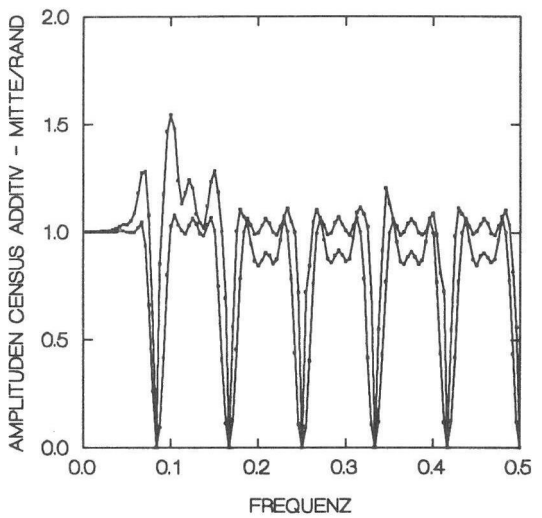


Abb. 75

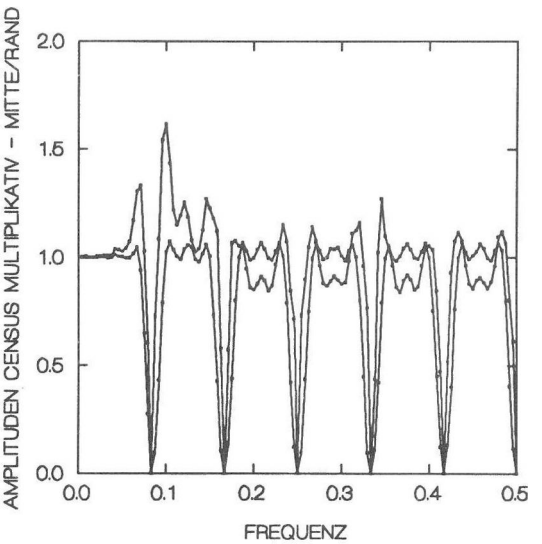


Abb. 76

Die Abb. 77-80 zeigen die Amplitudenfunktionen der *glatten* Komponente in Reihenmitte und am Reihenrand für die additive bzw. multiplikative Version des CENSUS-X-11-Verfahrens.

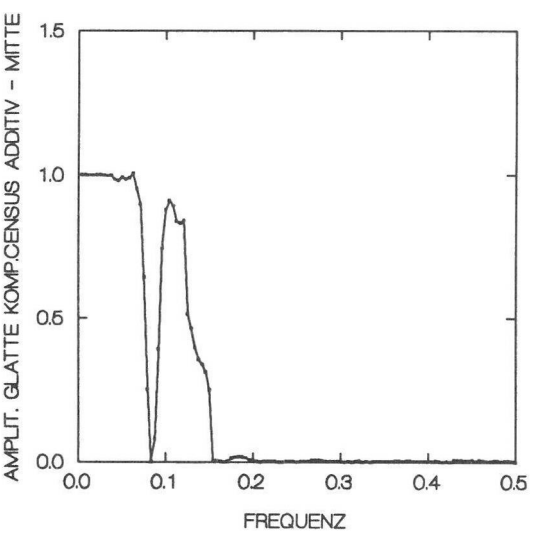


Abb. 77

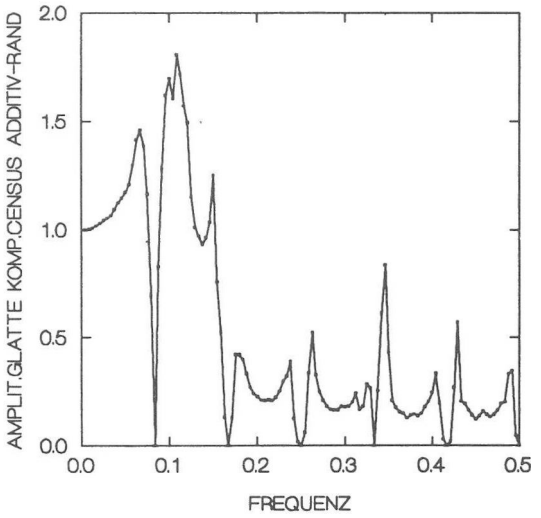


Abb. 78

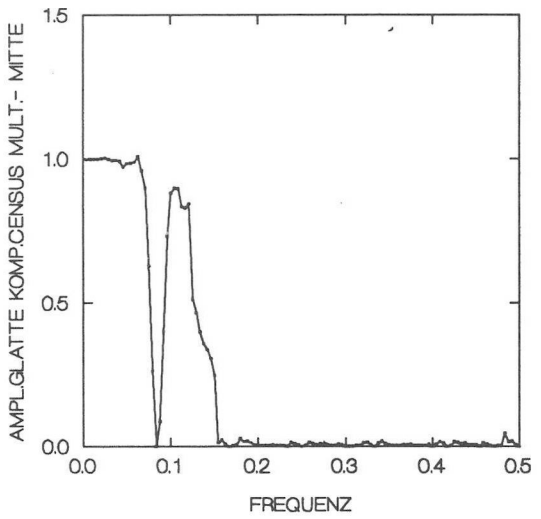


Abb. 79

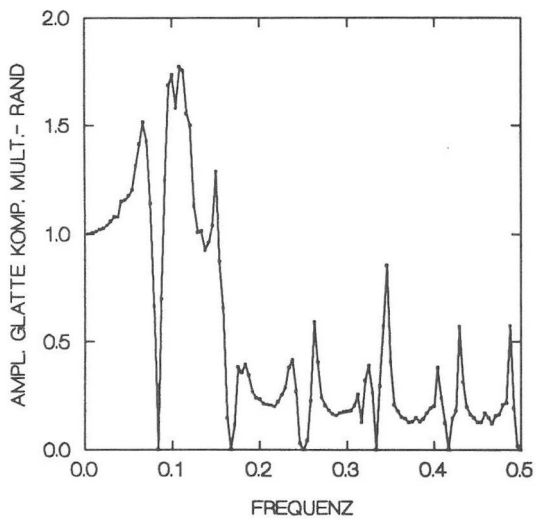


Abb. 80

In den Abb. 81 und 82 sind die Amplituden für Reihenmitte und Reihenrand zusammen dargestellt.

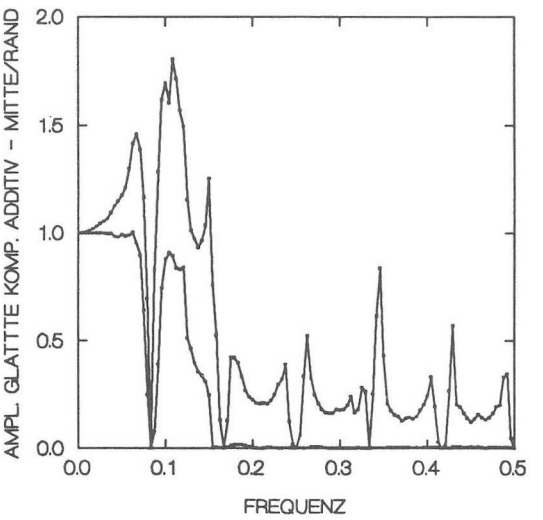


Abb. 81

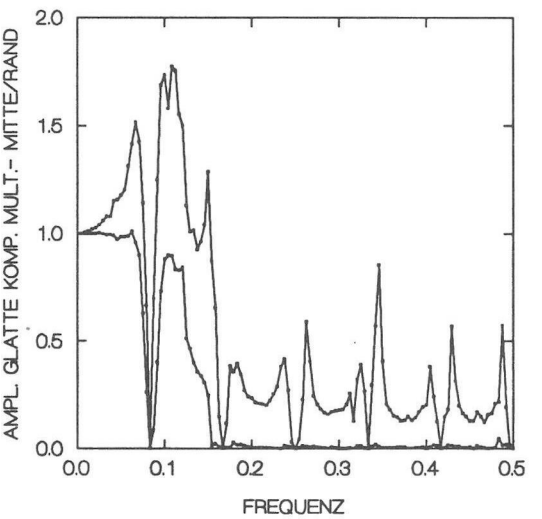


Abb. 82

5. Resultate für die Phasenfunktionen

Die Ergebnisse für die Phasenfunktionen werden analog zu denjenigen der Amplitudenfunktionen sowohl für die *additive* als auch für die *multiplikative* Version dargestellt, wobei zusätzlich zwischen Phasen in *radian* und Phasen in *Zeiteinheiten* unterschieden wird. Für die *Saisonbereinigung* der *additiven* Version sind diese Phasenfunktionen in den Abb. 83-86 zu finden. Abb. 87 und 88 stellen sie zusammen für Reihenmitte und Reihenrand dar.

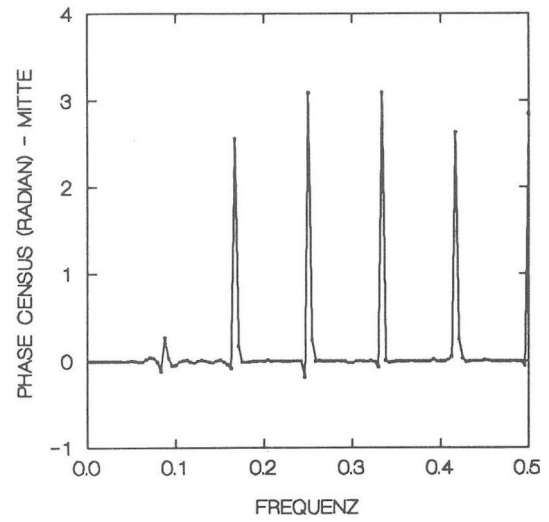


Abb. 83

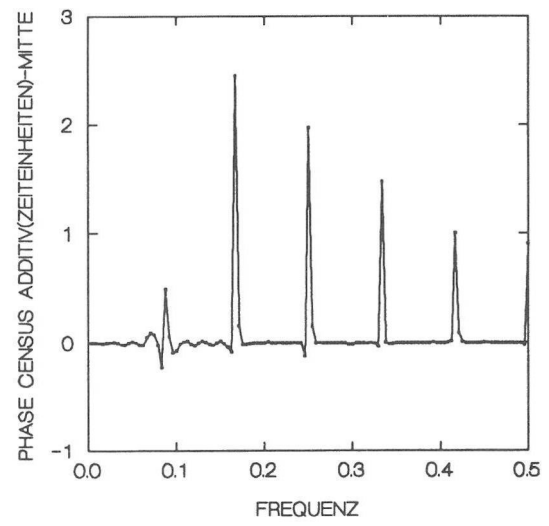


Abb. 84

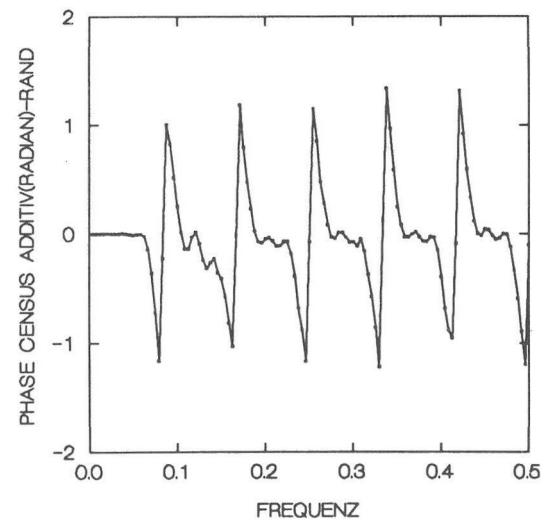


Abb. 85

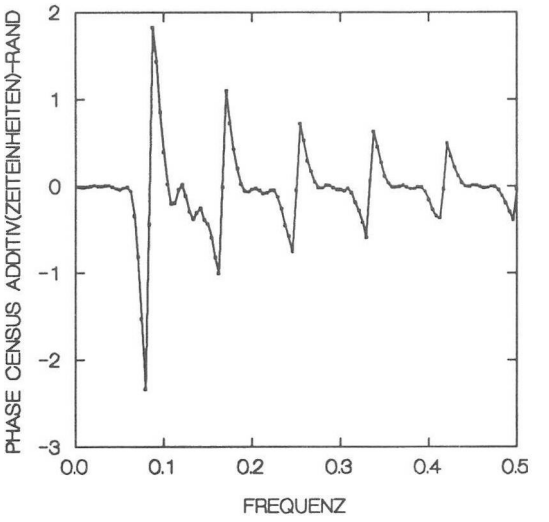


Abb. 86

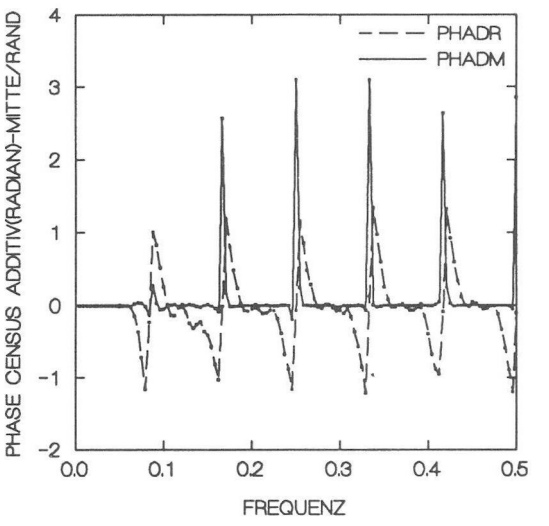


Abb. 87

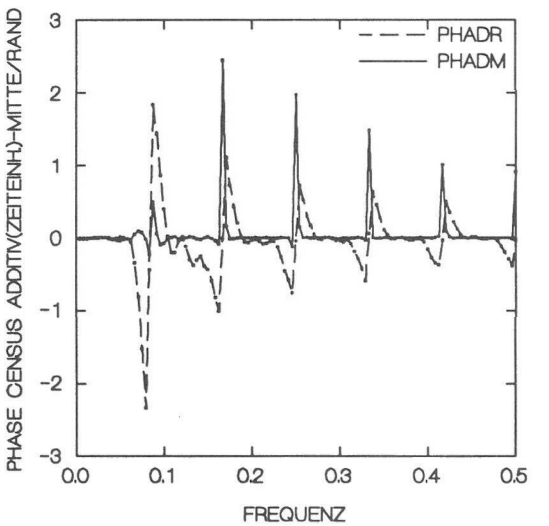


Abb. 88

Die Abb. 89-92 zeigen die Phasenfunktionen der *multiplikativen* Version, wobei die Abb. 93 und 94 "Mitte" und "Rand" zusammen präsentieren.

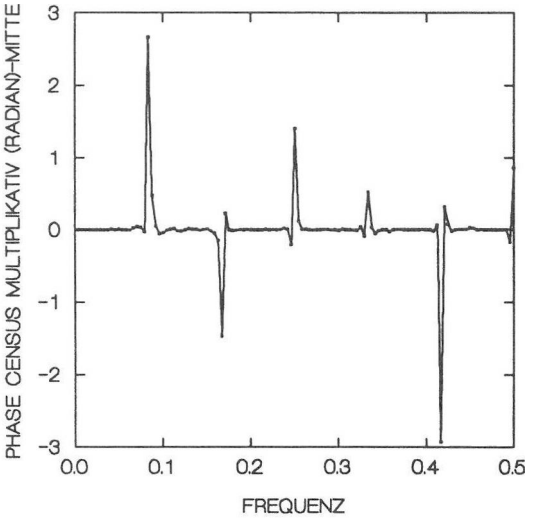


Abb. 89

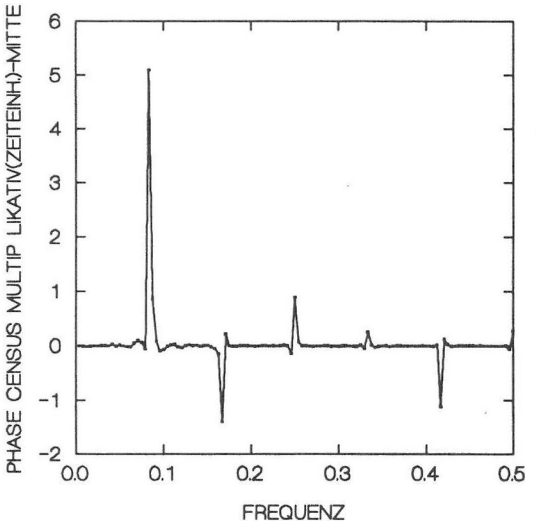


Abb. 90

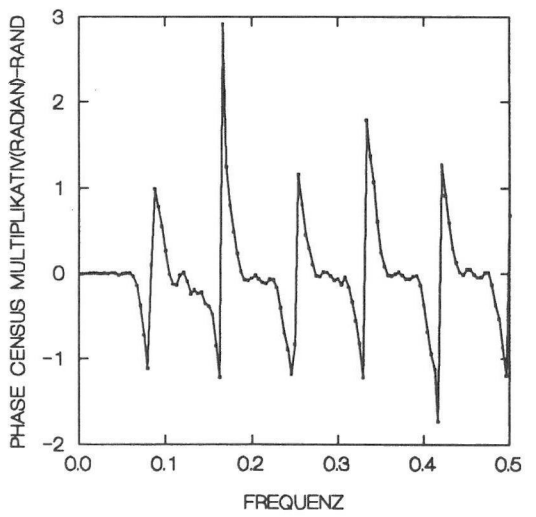


Abb. 91

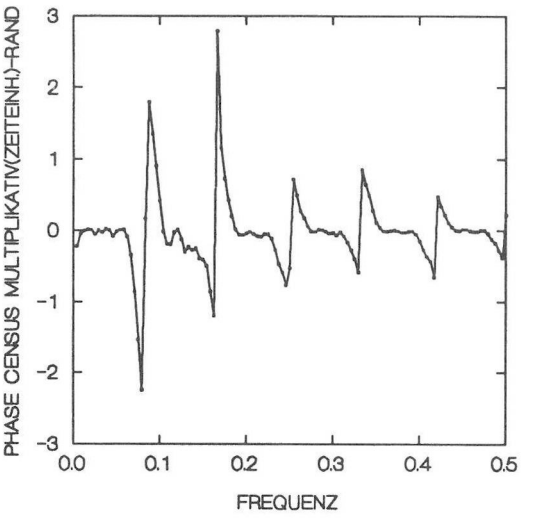


Abb. 92

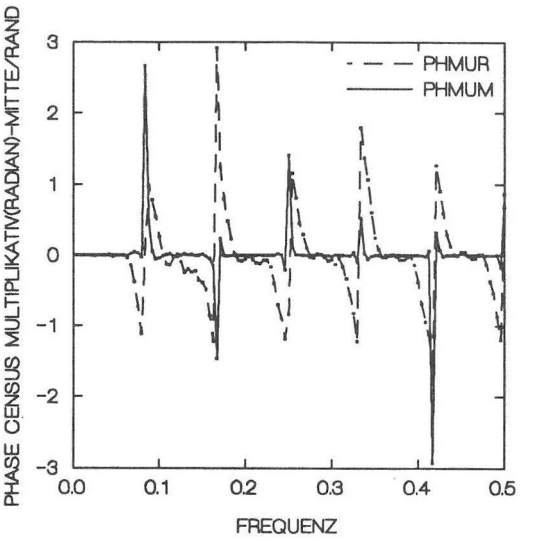


Abb. 93

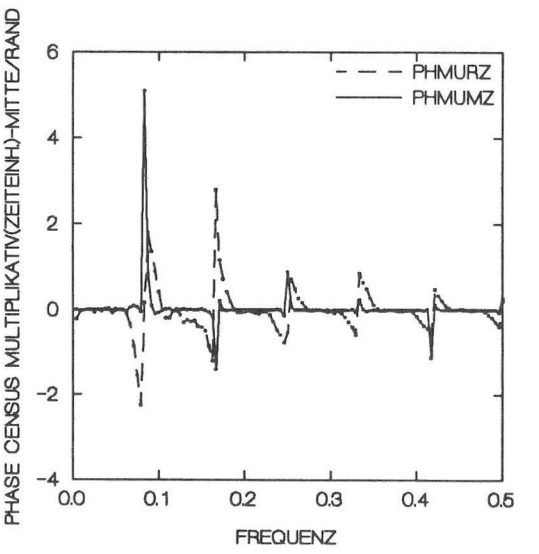


Abb. 94

Die Phasenfunktionen *der glatten Komponente* der *additiven* Version sind in den Abb. 95-98 wiedergegeben. Die Abb. 99 und 100 stellen wiederum "Mitte" und "Rand" gemeinsam dar.

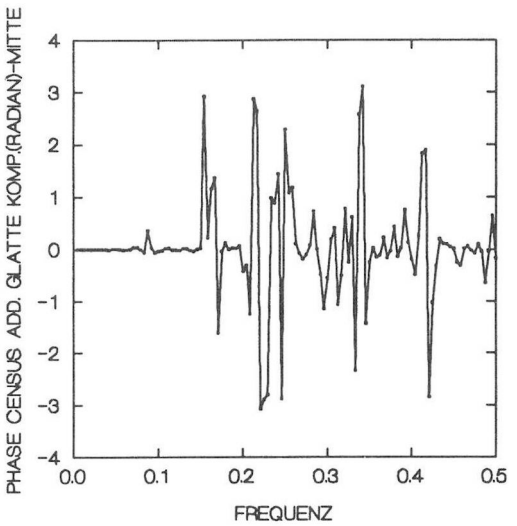


Abb. 95

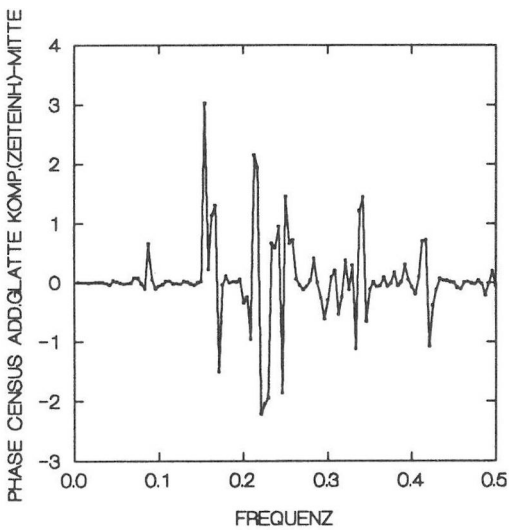


Abb. 96

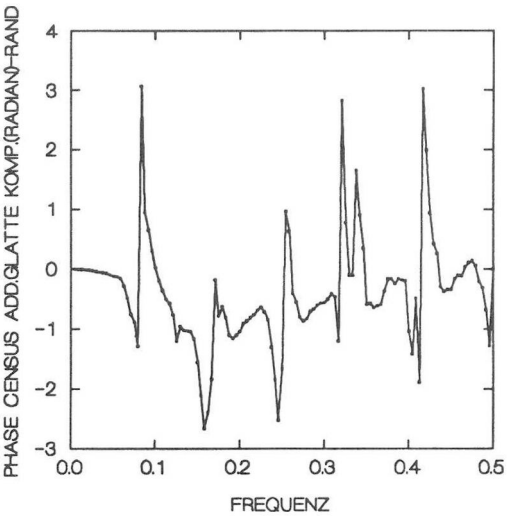


Abb. 97

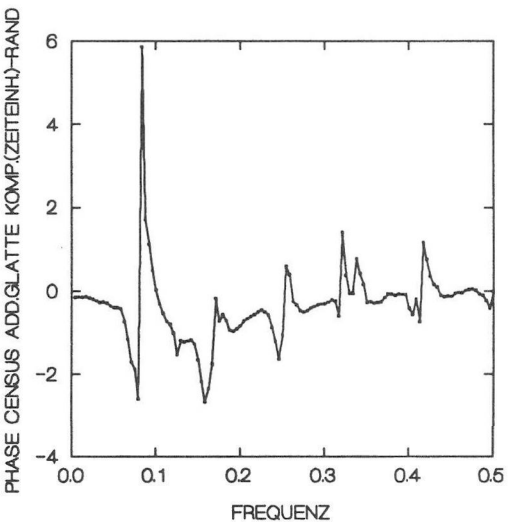


Abb. 98

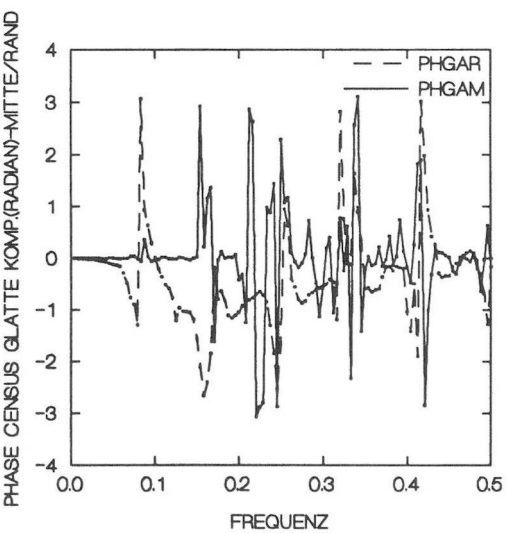


Abb. 99

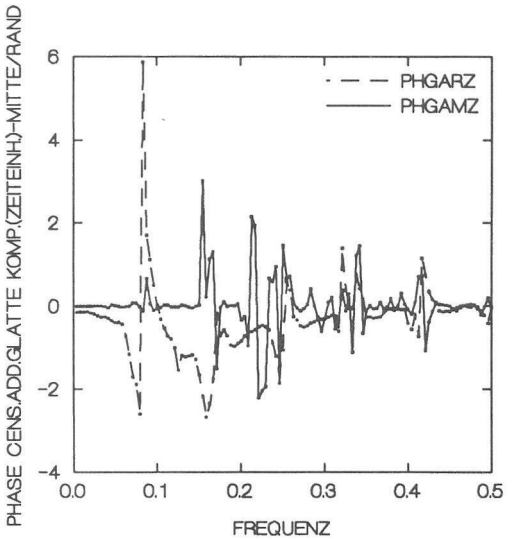


Abb. 100

Schliesslich ergeben sich für die Phasenfunktionen der *glatten* Komponente der *multiplikativen* Version die Verläufe der Abb. 101-104, die in den Abb. 105 und 106 gemeinsam für "Mitte" und "Rand" dargestellt sind.

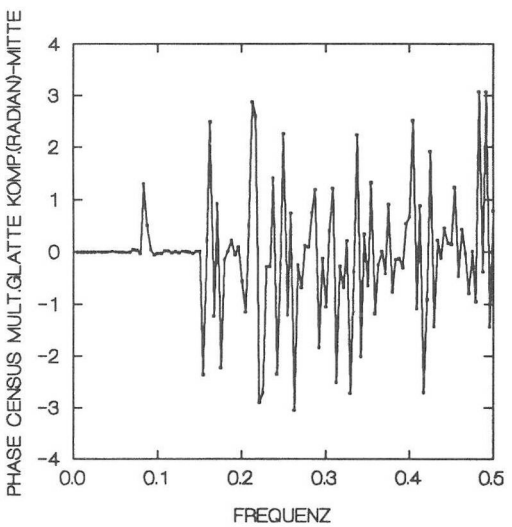


Abb. 101

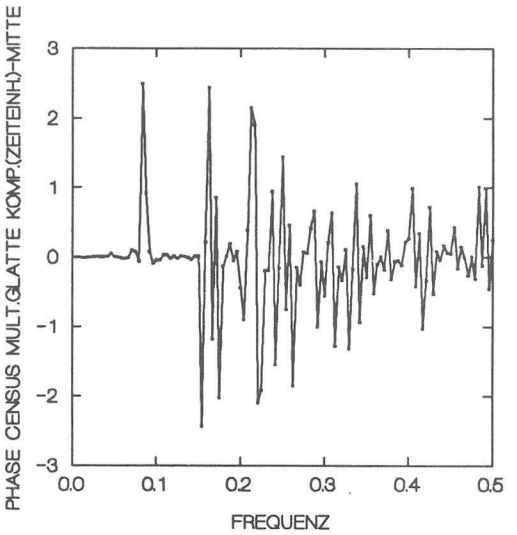


Abb. 102

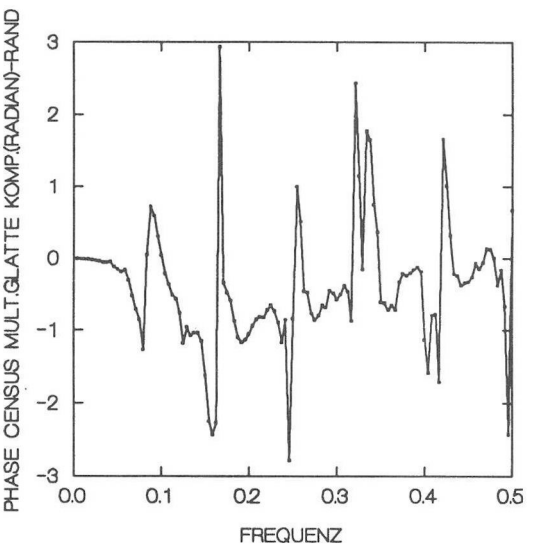


Abb. 103

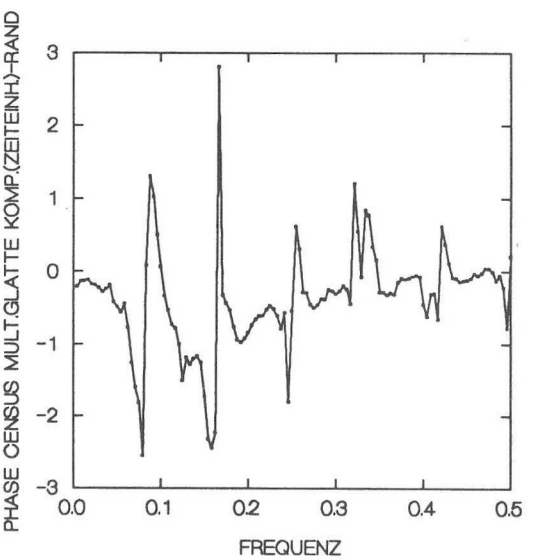


Abb. 104

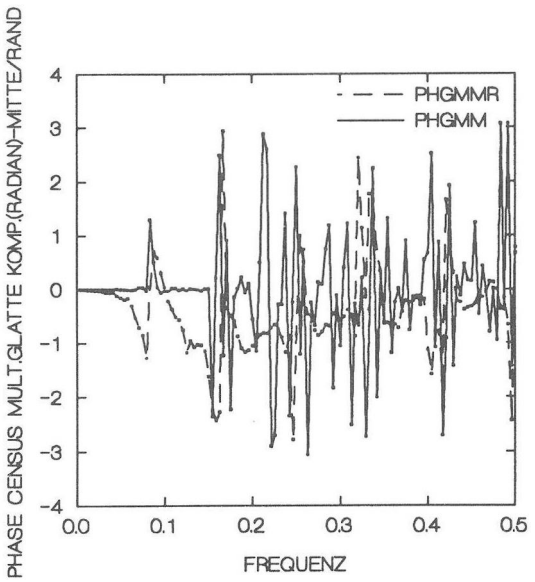


Abb. 105

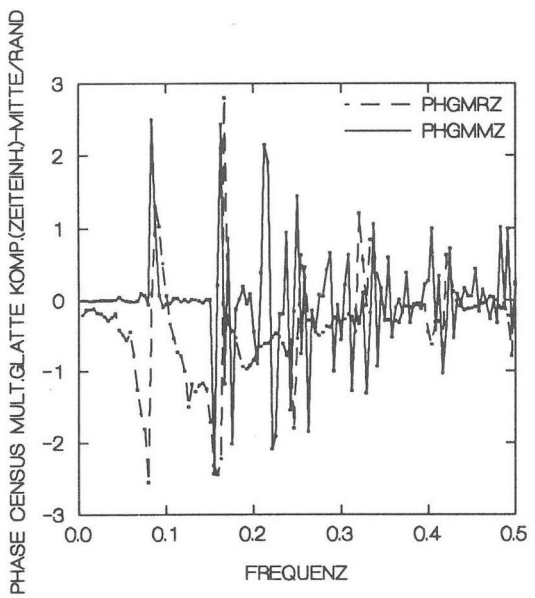


Abb. 106

6. Interpretationen der Amplitudenfunktionen

Vergleicht man die *Amplitudenfunktionen* für die *Saisonbereinigung* der beiden CENSUS-X-11-Versionen, so kann man zwischen diesen Versionen kaum nennenswerte Differenzen feststellen. Dies gilt sowohl für die Reihenmitte als auch für den Reihenrand (vgl. Abb. 71 und 72 bzw. Abb. 73 und 74). Für die Reihenmitte ist bei beiden Versionen eine relativ gute Annäherung an die Amplitude des Kerbenfilters festzustellen (vgl. dazu Abb. 70).

Beide CENSUS-X-11-Versionen zeigen am Reihenrand eine *verzerrte Amplitude*, wobei der Grad der Verzerrung in beiden Fällen als etwa gleich einzustufen ist (vgl. Abb. 72 und 74).

Die Amplituden der glatten Komponente zeigen, dass das CENSUS-X-11-Verfahren in beiden Versionen ausgesprochen *schlechte Tiefpass-Filtereigenschaften* aufweist und dies sogar in *Reihenmitte*, was eigentlich nicht zu erwarten ist. Wie man aus Abb. 77 und Abb. 79 entnehmen kann, weisen beide Amplituden für die Reihenmitte zwischen der ersten und zweiten saisonalen Frequenz einen beträchtlichen *peak* auf - die Amplituden nehmen dort einen Wert von ca. 0.9 an - obwohl sie in diesem Frequenzbereich Null oder mindestens annähernd Null sein müssten. Am Reihenrand tritt dieser peak verstärkt auf (vgl. Abb. 78 und 80).

Dieser Sachverhalt ist keinesfalls von rein akademischem Interesse, da er schwerwiegende praktische Konsequenzen hat, insofern nämlich, als er für beobachtbare Oszillationen in der glatten Komponente verantwortlich ist, d.h. beide CENSUS-X-11-Versionen weisen offensichtlich eine Art Slutzky-Effekt auf.

Dieser Slutzky-Effekt kann an konkreten Reihen demonstriert werden. In nachfolgender Abb. 107 ist beispielhaft eine Monatsreihe (Umsätze einer Firma) zusammen mit ihrer glatten CENSUS-X-11-Komponente (multiplikativ) wiedergegeben.

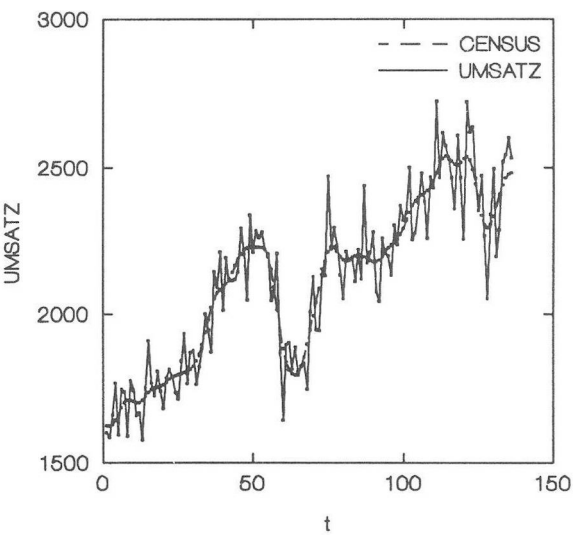


Abb. 107

Dass diese glatte Komponente einen Slutzy-Effekt aufweist, sieht man deutlich, wenn man sie zusammen mit der glatten Komponente eines exakten Hochpass-Filters vergleicht (vgl. Abb. 108).

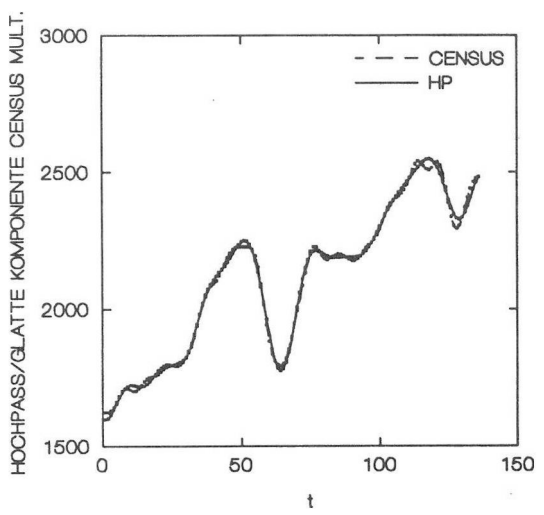


Abb. 108

Offensichtlich oszilliert die CENSUS-X-11-Komponente um die Hochpass-Komponente mit einer Periode von ca. 9 Monaten, was in etwa der oben erwähnten *peak-Frequenz* entspricht. Optisch ganz deutlich werden diese Oszillationen, wenn man Abb. 108 nur ausschnittsweise in Abb. 109 betrachtet.

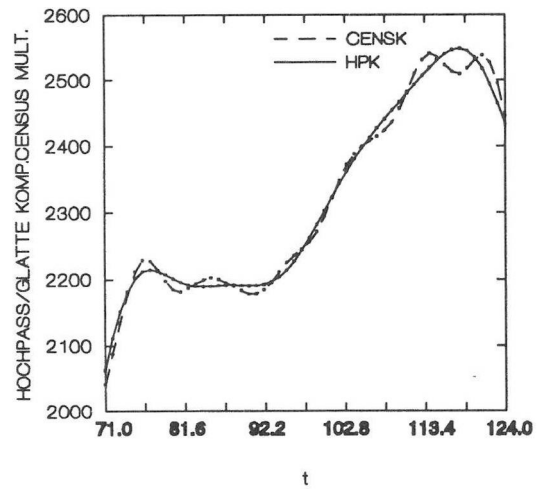


Abb. 109

Die Erkenntnis, dass die glatte Komponente nach dem CENSUS-X-11-Verfahren einen Slutzky-Effekt enthält, ist allerdings nicht ganz neu. Schon vor gut 20 Jahren kam SCHÄFFER auf einem anderen Weg, nämlich über relativ umfangreiche und komplizierte Modellrechnungen zu dem Schluss, dass die glatte CENSUS-X-11-Komponente einen Slutzky-Effekt aufweisen dürfte [3]. Diesem Resultat scheint aber nicht die gebührende Aufmerksamkeit geschenkt worden zu sein.

Noch schlechter als in Reihenmitte sind die CENSUS-X-11-Amplituden für die glatte Komponente am Reihenrand (vgl. die Abb. 78 und 80). Diese haben mit der Amplitude eines Tiefpass-Filters offenbar überhaupt nichts mehr gemein. Abgesehen davon, dass der erwähnte *peak* etwa doppelt so hoch ist wie in Reihenmitte, ist festzustellen, dass der Hochfrequenzanteil einer Reihe nur in völlig unzureichender Weise ausgefiltert wird (die Amplituden müssten ab der ersten saisonalen Frequenz Null oder fast gleich Null sein). Dies führt dazu, dass die Randwerte der glatten Komponente *hochgradig volatil* sind, ein Phänomen, das den intensiven Benutzern des CENSUS-X-11-Verfahrens (z.B. der DEUTSCHEN BUNDESBANK) seit Jahren wohlbekannt ist, für das aber bislang offensichtlich eine theoretische Erklärung fehlte.

7. Interpretationen der Phasenfunktionen

Für beide CENSUS-X-11-Versionen sind die Phasen in der *Reihenmitte* (sowohl in radian als auch in Zeiteinheiten) *praktisch gleich Null* im ersten (und entscheidenden) Frequenzband, das sich bis zur ersten saisonalen Frequenz erstreckt (vgl. dazu Abb. 83 und 84 sowie Abb. 89 und 90). Dies ist allerdings nicht weiter erstaunlich, werden doch in Reihenmitte ausschliesslich *symmetrische Filter* verwendet.

Erstaunlich ist dagegen das Phasenverhalten am Reihenrand. Dort zeigt sich praktisch der *gleiche Phasenverlauf wie in der Mitte*, obwohl dort *asymmetrische Filter* verwendet werden (vgl. Abb. 85 und 86 bzw. Abb. 91 und 92).

Genau der gleiche Sachverhalt ist für die glatte Komponente festzustellen (vgl. Abb. 95 und 96 sowie Abb. 97 und 98), was noch um einiges erstaunlicher ist (besonders für den Rand), werden doch bei der Schätzung dieser Komponente am Rand stark *asymmetrische Filter* eingesetzt.

Die unerwarteten guten Phaseneigenschaften vom CENSUS-X-11-Verfahren am Reihenrand lassen sich leicht erklären, wenn man Amplituden- und Phasenfunktionen *zusammen* betrachtet.

Bei asymmetrischen Filtern besteht generell eine *Antinomie* zwischen Amplitude und Phase, d.h. je schärfer die Amplitude, umso schlechter (d.h. umso grösser) ist die Phase. Mit anderen Worten, das CENSUS-X-11-Verfahren erkauft sich die günstigen Phaseneigenschaften am Rand mit einer *Zerstörung der Amplitudenfunktion*. Die überraschenden quasi-Nullphasen am Rand resultieren somit lediglich aus miserablen Rand-Amplitudenfunktionen.

Es sei hier noch hinzugefügt, dass eine methodologische Konsequenz dieses Sachverhaltes darin zu sehen ist, dass die Randwerte (die saisonbereinigten, aber in noch stärkerem Masse diejenigen der glatten Komponente) mit den entsprechenden Werten in Reihenmitte eigentlich *nicht vergleichbar* sind, denn an den (realisierten) Amplitudenfunktionen lassen sich im Frequenzbereich die einem Verfahren zugrundeliegenden Komponentendefinitionen, oder genauer ausgedrückt, die *operationalisierten Komponen-*

tendefinitionen ablesen. Stark divergierende Amplitudenfunktionen wie hier zwischen Reihenmitte und -rand sind danach gleichbedeutend mit *unterschiedlichen Komponententendefinitionen.*

8. Schätzung der glatten Komponente simulierter Reihen mit dem CENSUS-X-11-Verfahren

Wie in Abschnitt 1 dieses Kapitels dargelegt wurde, sind insbesondere die Amplitudenfunktionen des CENSUS-X-11-Verfahrens für die glatte Komponente als problematisch zu bezeichnen. Um mögliche praktische Auswirkungen dieser Problematik zu veranschaulichen, soll zunächst für die vier simulierten Zeitreihen aus Kapitel I, Abschnitt 4 (vgl. die Abb. 25, 29, 30 und 31) die glatte Komponente mit dem CENSUS-X-11-Verfahren geschätzt werden. Diesen Schätzungen sei jeweils eine Bestimmung der glatten Komponente mit Hilfe eines Tiefpass-Filters gegenübergestellt.

Wie aus Kapitel I, Abschnitt 4 ersichtlich ist, liegt diesen Reihen das einfache Komponentenmodell

$$X_t = 2 + 0.9t + \sin(2\pi/12) + \varepsilon_t \quad , t = 1,2,...,100$$

zugrunde. Die glatte Komponente aller vier Reihen ist somit durch die lineare Funktion 2 + 0.9t gegeben. Die Reihen unterscheiden sich nur dadurch, dass der noise-Term ε_t jeweils eine unterschiedliche Varianz aufweist. Bei Reihe 1 (vgl. Abb. 25) ist sie gleich Null (d.h. diese Reihe enthält keinen noise), bei Reihe 2 (vgl. Abb. 29) gleich 0.1, bei Reihe 3 (vgl. Abb. 30) gleich 1.0 und bei Reihe 4 (vgl. Abb. 31) schliesslich gleich 2.0.

Die Abb. 110-113 zeigen nun jeweils die mit dem CENSUS-X-11-Verfahren geschätzte glatte Komponente dieser Reihen.

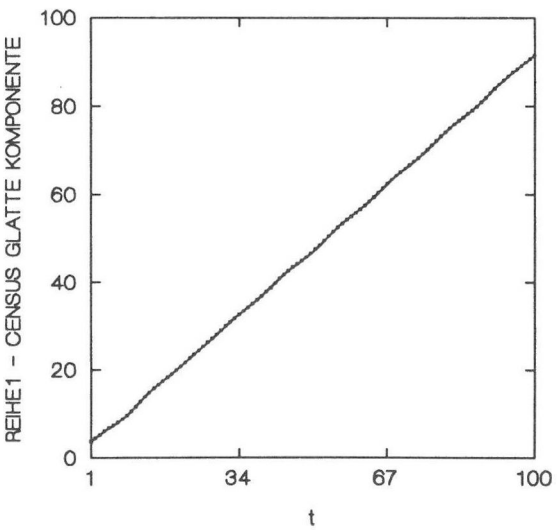


Abb. 110

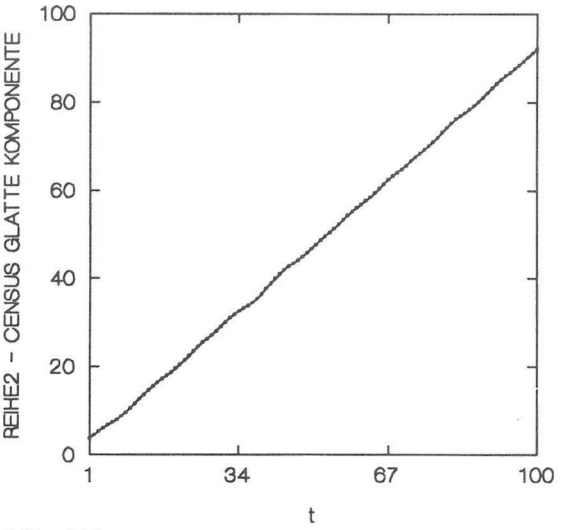


Abb. 111

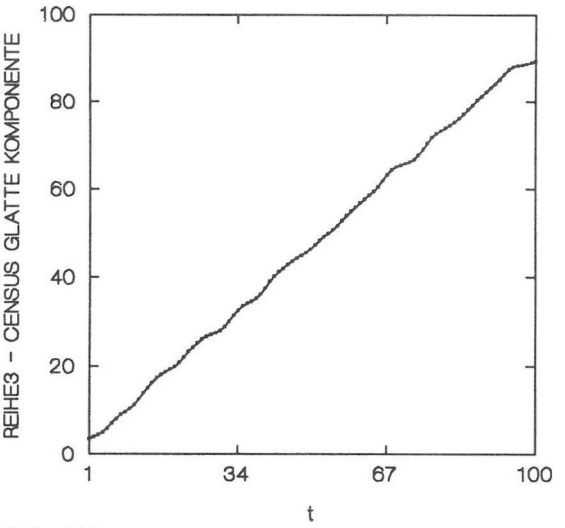


Abb. 112

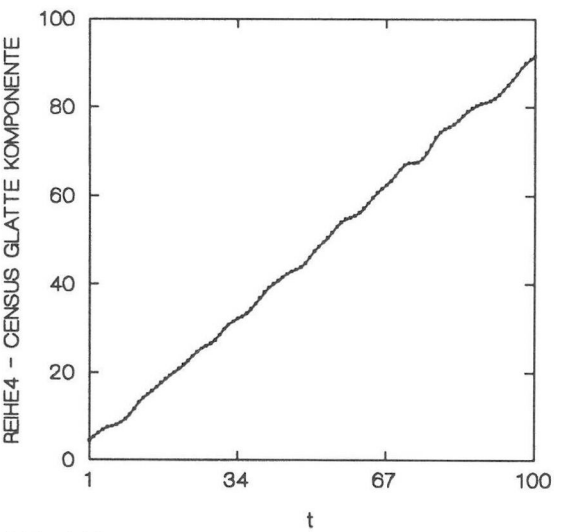


Abb. 113

Offensichtlich weist die glatte Komponente des CENSUS-X-11-Verfahrens Oszillationen auf, die insbesondere bei den Reihen 3 und 4 (vgl. die Abb. 112 und 113) ins Auge springen. Aber sie sind auch, wenngleich weniger stark

ausgeprägt, schon bei den Reihen 1 und 2 (vgl. die Abb. 110 und 111) zu erkennen.

Zum Vergleich sei der Filter-Output eines Tiefpass-Filters in der Abb. 114 für Reihe 4 dargestellt.

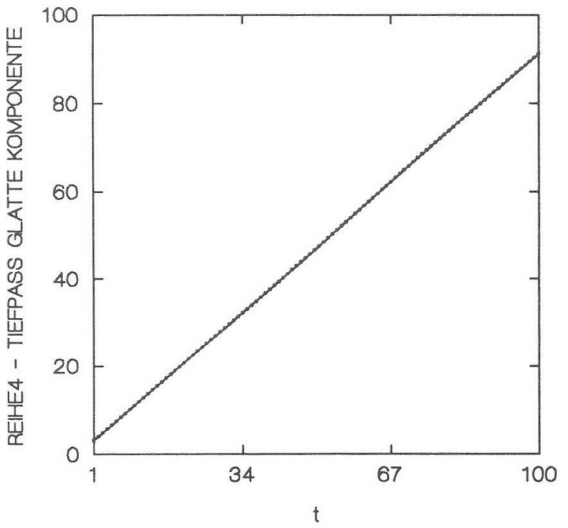


Abb. 114

Diese glatte Komponente ist optisch praktisch nicht von einer Geraden zu unterscheiden, es treten also keine Oszillationen in der bereinigten Reihe auf. Für die übrigen drei Reihen ergibt sich jeweils fast das gleiche Resultat, so dass sich Wiederholungen hier erübrigen.

Dass der Tiefpass-Filter durchwegs bessere Resultate liefert, zeigt sich auch, wenn man mit Hilfe eines *Gütemasses* die Differenzen zwischen der theoretischen glatten Komponente, also der Funktion 2 + 0.9t und der mit dem CENSUS-X-11-Verfahren geschätzten sowie der mit einem Tiefpass-Filter erzielten Komponente vergleicht. Verwendet man etwa als Gütemass den *mittleren quadratischen Fehler*

$$MQF_C = \sqrt{\frac{1}{100}\sum_{t=1}^{100}(GK_t - GKC_t)^2}$$

bzw.

$$MQF_{TP} = \sqrt{\frac{1}{100}\sum_{t=1}^{100}(GK_t - GKTP_t)^2}$$

mit

$$GK_t = 2 + 0.9t$$

und

GKC_t : *Glatte Komponente, CENSUS*
$GKTP_t$: *Glatte Komponente, Tiefpass* ,

so ergeben sich für das CENSUS-X-11-Verfahren und für den Tiefpass-Filter folgende Werte:

	CENSUS-X-11	Tiefpass-Filter
Reihe 1 :	0.2275	0.0371
Reihe 2 :	0.2913	0.0965
Reihe 3 :	0.6333	0.1930
Reihe 4 :	0.7522	0.2983

Zusätzlich zu diesem Gütemass kann man z.B. noch die Standardabweichungen der *prozentualen relativen* Differenzen berechnen, d.h. die Standardabweichungen der Differenzen bezogen auf die jeweiligen Werte der theoretischen glatten Komponente. Mit dem CENSUS-X-11-Verfahren und dem Tiefpass-Filter ergibt sich folgendes Bild:

	CENSUS-X-11	Tiefpass-Filter
Reihe 1 :	3.26 %	0.55 %
Reihe 2 :	3.69 %	0.86 %
Reihe 3 :	2.85 %	2.38 %
Reihe 4 :	9.11 %	1.71 %

Auch dieses Streuungsmass zeigt durchwegs bessere Werte für den Tiefpass-Filter.

Auf der Basis dieser Differenzen lassen sich aber nicht nur solche oder ähnliche Masszahlen berechnen. Informativer sind möglicherweise Überlegungen, welche die Differenzen selbst als Zeitreihen betrachten. Damit ist eine Zielsetzung verbunden, die sich etwa so charakterisieren lässt: Enthalten die Differenzen noch Informationen, die sie "eigentlich" nicht mehr enthalten dürften?

Die Differenzen für die mit dem CENSUS-X-11-Verfahren bereinigten Reihen sind in den nächsten 4 Abbildungen dargestellt (vgl. Abb. 115-118):

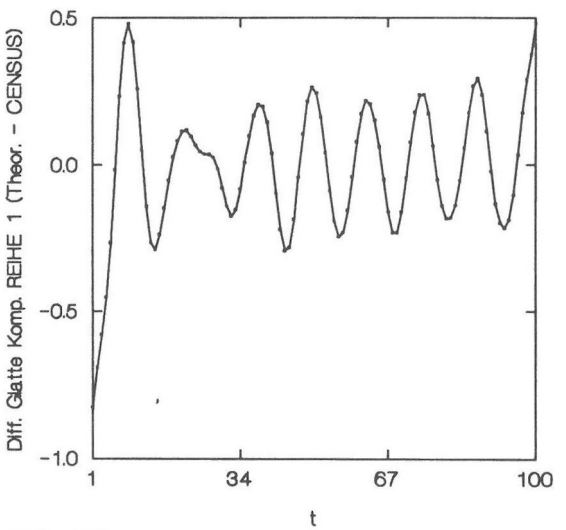


Abb. 115

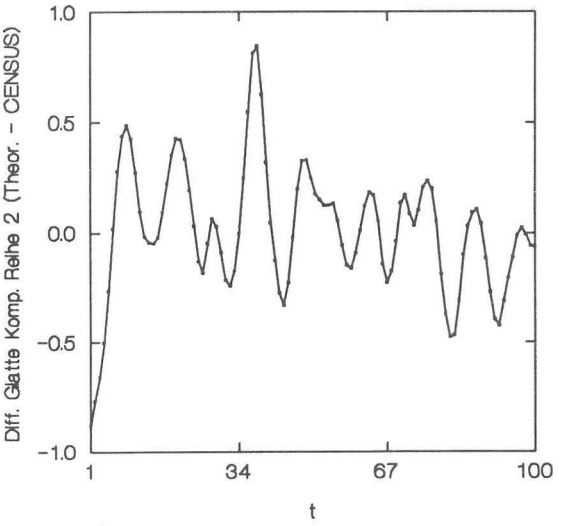


Abb. 116

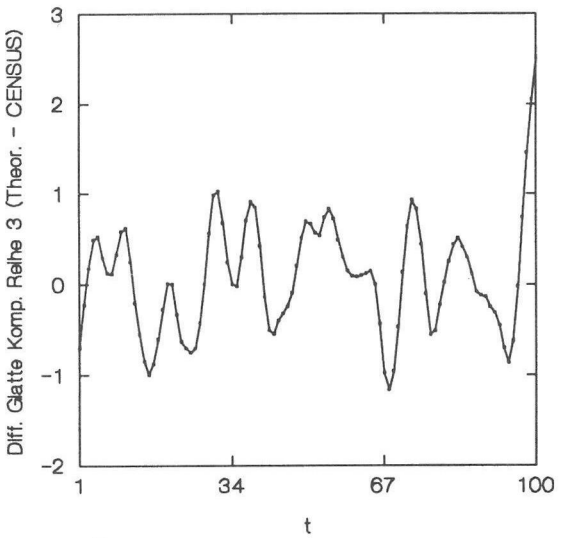


Abb. 117

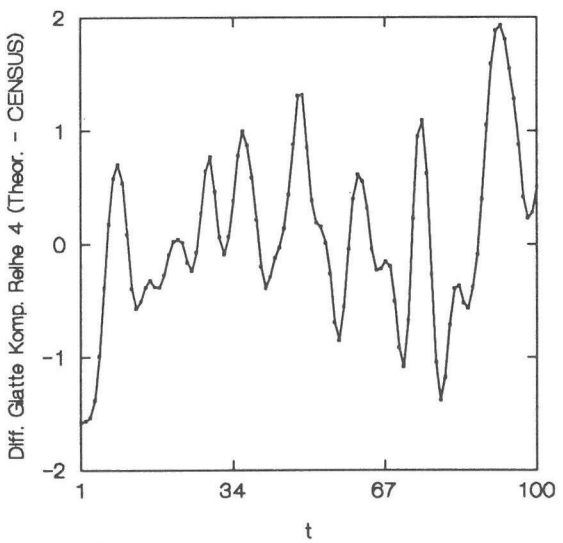


Abb. 118

Auffallend ist, dass alle vier Differenzenreihen ziemlich regelmässig verlaufende kurzfristige Oszillationen aufweisen. Diese zeigen sich auch klar in den geschätzten Spektren dieser Differenzen, die nachstehend wiedergegeben sind (vgl. Abb. 119-122):

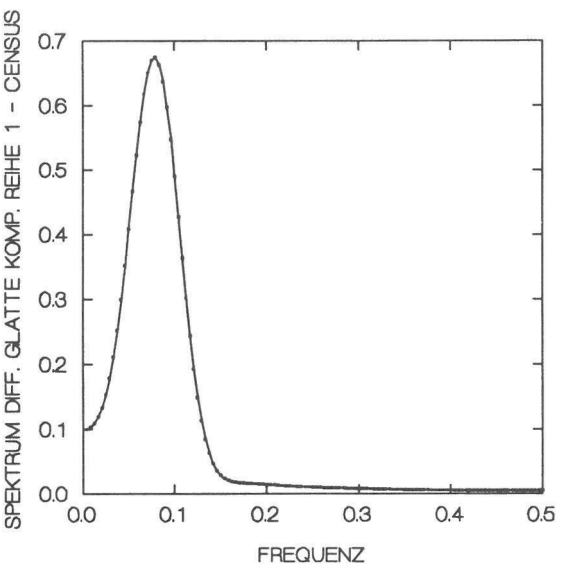


Abb. 119

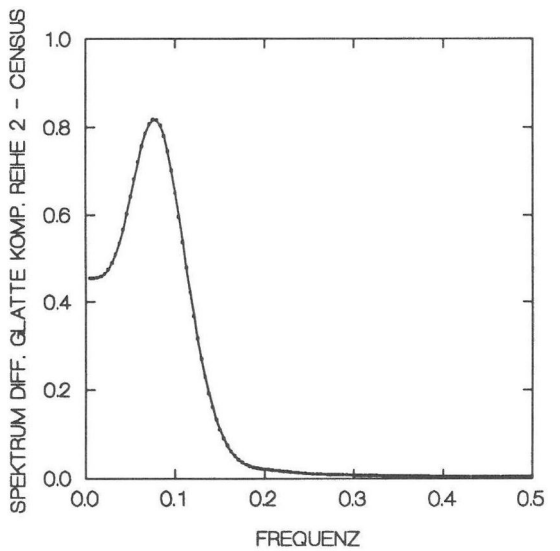


Abb. 120

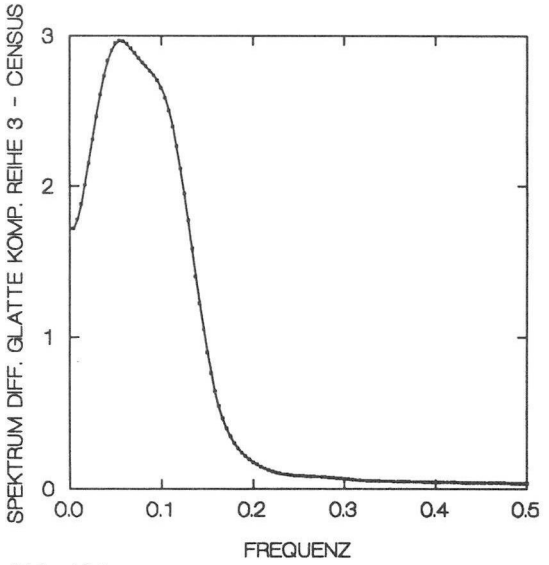


Abb. 121

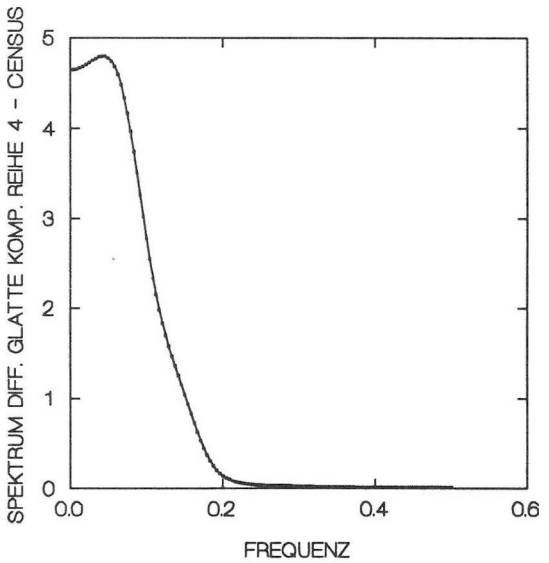


Abb. 122

Für die ersten drei Differenzenreihen liegt der jeweilige *peak* im Spektrum bei ca. 12, 13 und 18 Monaten. Für die vierte Differenzenreihe liefert das Spektrum keine klare Information, weil diese einen leichten Trend aufweist. Filtert man jedoch diesen mit einem Hochpass aus, dann zeigt das Spektrum folgenden Verlauf (vgl. Abb. 123):

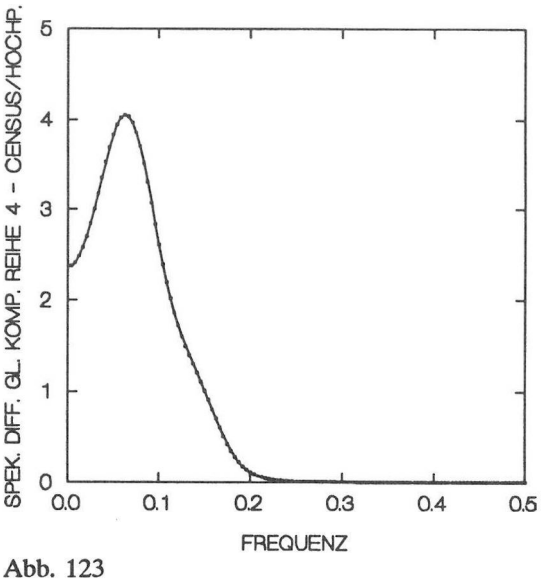


Abb. 123

Das Maximum dieses Spektrums liegt bei ca. 16 Monaten.

Somit bestätigen die Spektren den optischen Eindruck ausgeprägter, aber relativ kurzfristiger, zyklischer Bewegungen obiger Differenzenreihen.

Die Differenzenreihen des Tiefpass-Filters zeigen für alle vier Reihen einen ziemlich gleichartigen Verlauf, so dass sie zusammen in Abb. 124 darstellbar sind.

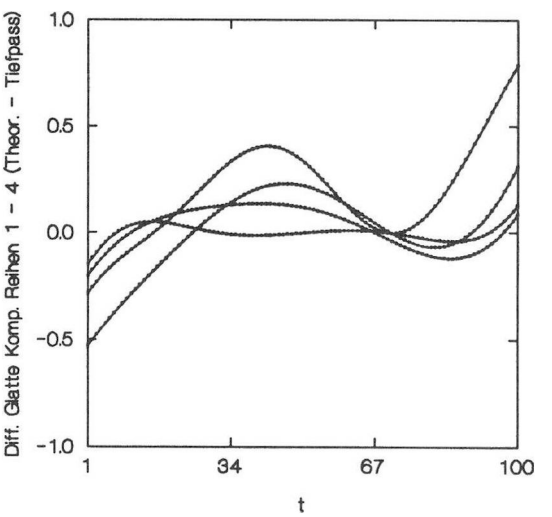


Abb. 124

Im Gegensatz zu den CENSUS-X-11-Differenzen verlaufen diese sehr ruhig und ohne erkennbare zyklische Bewegungen. Es ist deshalb zu erwarten, dass die Spektren dieser Differenzenreihen keine peaks aufweisen. In der Tat zeigen sie alle den gleichen Verlauf; eine Konzentration spektraler Masse um die Frequenz Null (was ihren ruhigen Verlauf reflektiert), aber keinerlei Hinweis auf zyklische Bewegungen wie bei den Differenzen, die sich mit dem CENSUS-

X-11-Verfahren ergeben. Daran ändert sich auch nichts, wenn man die Tiefpass-Differenzen hochpassfiltert. Stellvertretend sei deshalb nur ein Spektrum wiedergegeben (vgl. Abb. 125):

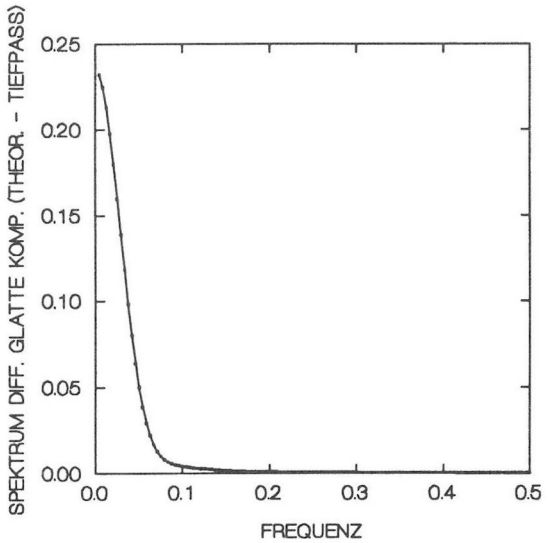


Abb. 125

Die zu beobachtenden zyklischen Bewegungen in den CENSUS-X-11-Differenzen zeigen sich aber auch in den geschätzten glatten Komponenten der vier Reihen, wenn man diese durch einen Hochpass-Filter schickt, wie das aus den Abb. 126-129 hervorgeht:

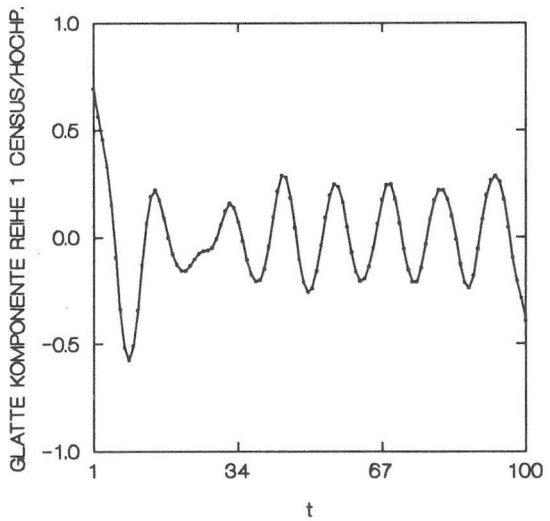


Abb. 126

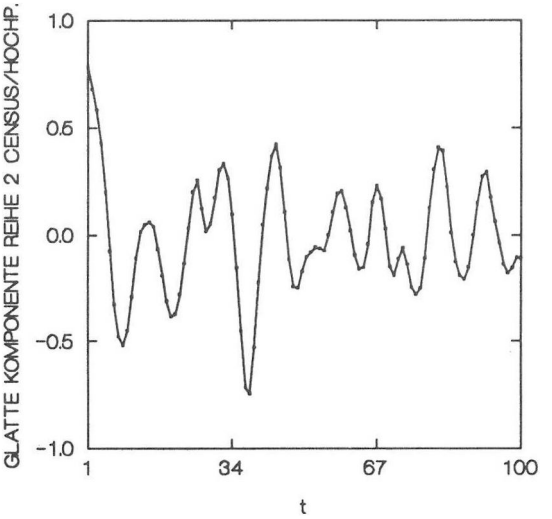


Abb. 127

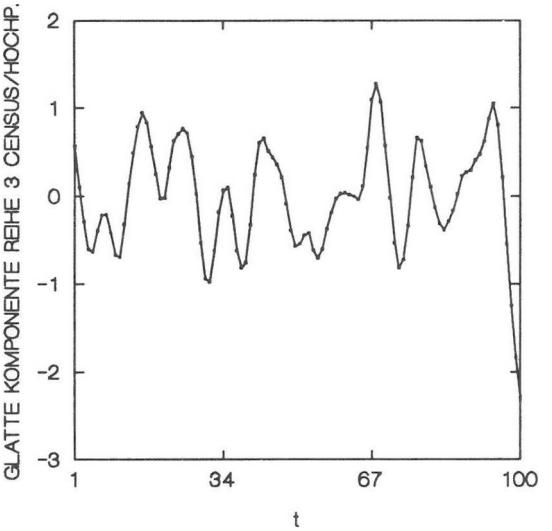


Abb. 128

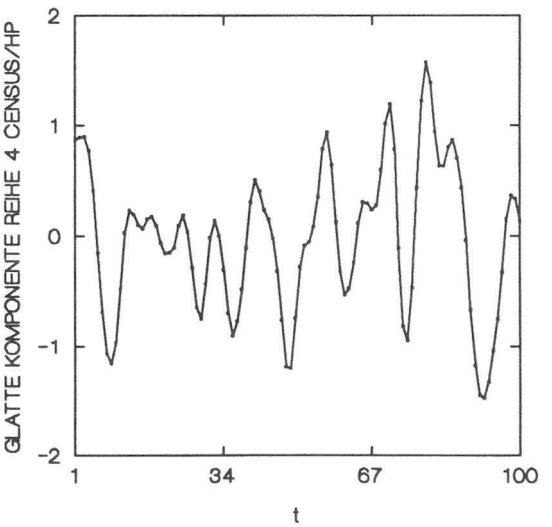


Abb. 129

Vergleicht man nun Abb. 126 mit Abb. 115, Abb. 127 mit Abb. 116, Abb. 128 mit Abb. 117 und schliesslich Abb. 129 mit Abb. 118, so stellt man unschwer fest, dass die jeweiligen Reihenpaare nahezu perfekt spiegelbildlich zur Null-Linie verlaufen. Zur Illustration sei das Reihenpaar der Abb. 126 und 115 zusammen dargestellt (vgl. Abb. 130):

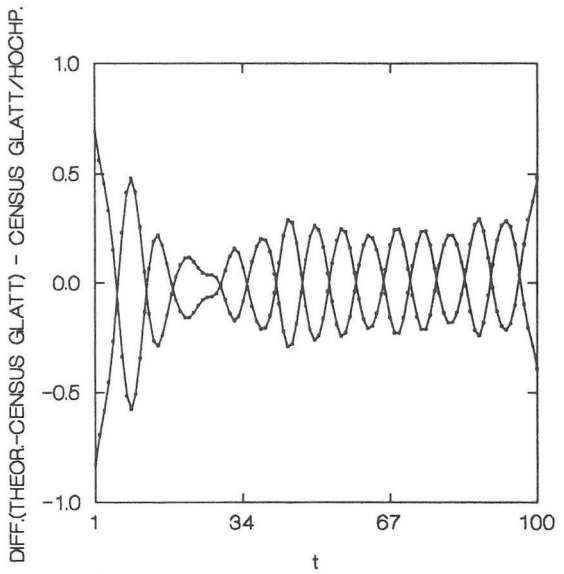


Abb. 130

Dass diese Spiegelbildlichkeit hier kein Zufall ist, sondern generell gilt, folgt aus elementaren filtertheoretischen Zusammenhängen.

Die Spektren der hochpassgefilterten glatten Komponente nach dem CENSUS-X-11-Verfahren enthalten somit genau die gleichen Informationen wie die Spektren der Differenzenreihen, d.h. sie zeigen dieselben oben erwähnten Zyklen an.

Somit lässt sich folgender Schluss ziehen:

Die glatte Komponente, die mit dem CENSUS-X-11-Verfahren ermittelt wird, enthält kurzfristige zyklische Bewegungen, die vom Verfahren induziert werden. Das CENSUS-X-11-Verfahren weist einen Slutzky-Effekt auf.

Für die Ursache dieses Effektes sei auf die Ausführungen zur Amplitudenfunktion vom CENSUS-X-11-Verfahren in Abschnitt 6 in diesem Kapitel verwiesen.

Die zyklischen Bewegungen der glatten Komponente nach dem CENSUS-X-11-Verfahren sind nicht nur als Schönheitsfehler des Verfahrens zu betrachten. Berechnet man nämlich auf ihrer Basis *Wachstumsraten gegenüber der Vorperiode, hier also dem Vormonat*, wie dies in der Praxis häufig geschieht, dann zeigen sich diese Bewegungen auch in diesen Wachstumsraten, was mit der grossen Gefahr irriger Schlussfolgerungen über die generelle Entwicklungsrichtung einer Reihe verbunden ist. Die Abb. 131-134 zeigen den Verlauf der Wachstumsraten, die aus der glatten Komponente nach dem CENSUS-X-11-Verfahren der vier Reihen gebildet wurden.

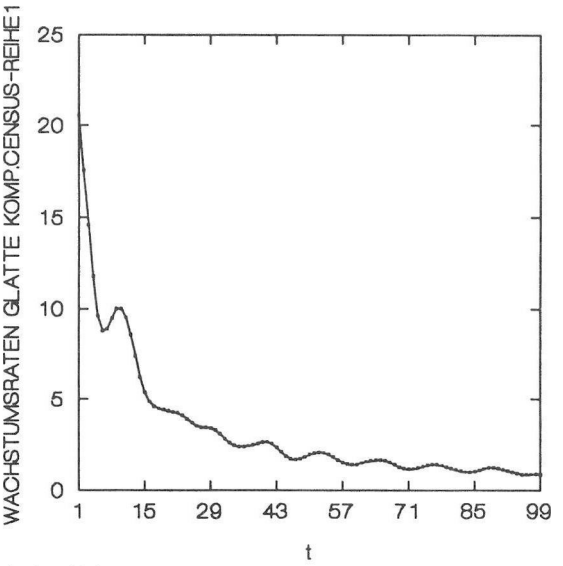


Abb. 131

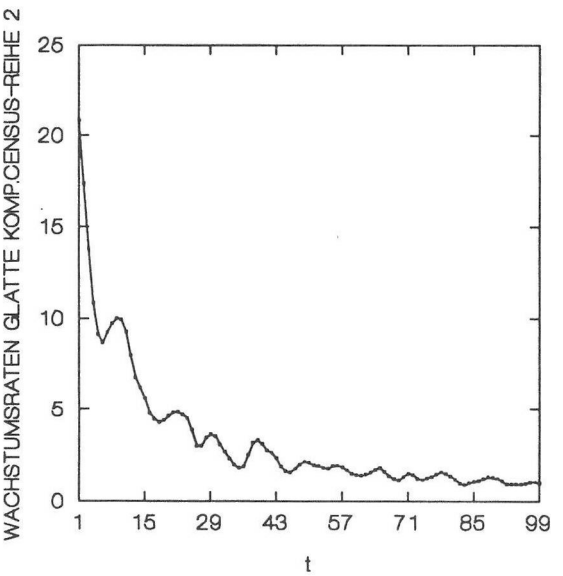


Abb. 132

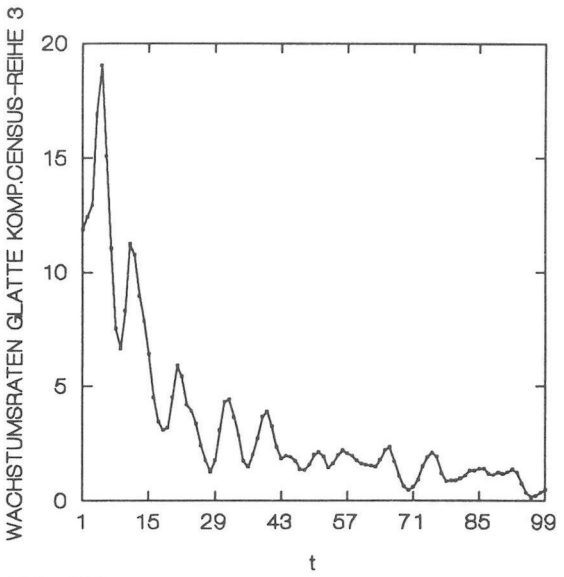


Abb. 133

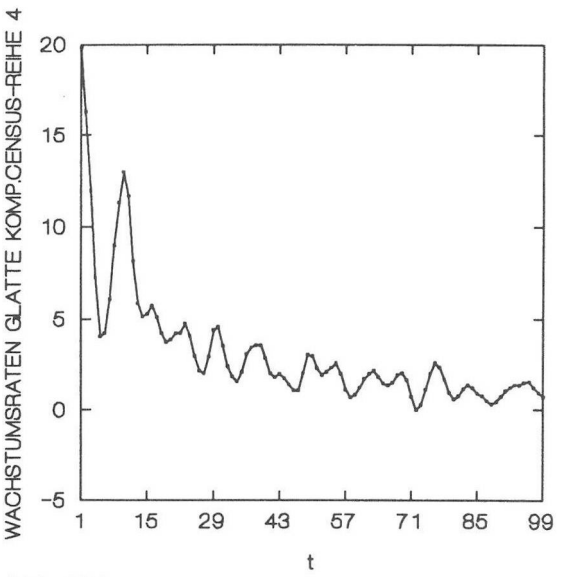


Abb. 134

Diese Wachstumsraten oszillieren um die wahren Raten, die bei einem linearen Trend einen monoton fallenden Verlauf aufweisen. Die Abb. 135 zeigt diese Oszillationen beispielhaft für Reihe 4.

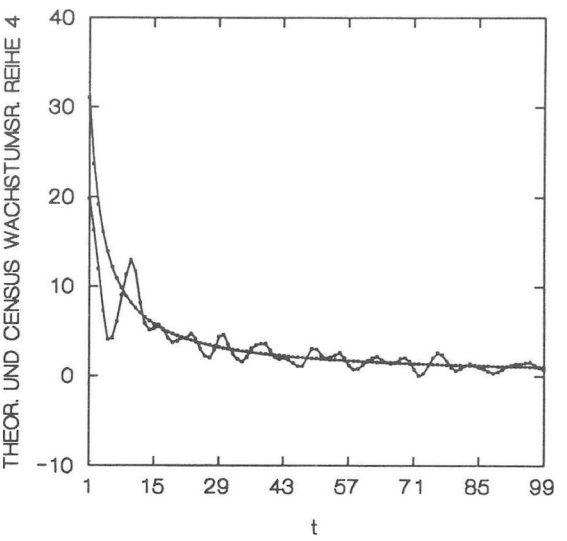


Abb. 135

Im Gegensatz dazu zeigen die Wachstumsraten der mit einem Tiefpass-Filter erzielten glatten Komponente monoton fallende Verläufe, wie sie für alle vier Reihen zusammen mit den wahren Wachstumsraten in Abb. 136 dargestellt sind.

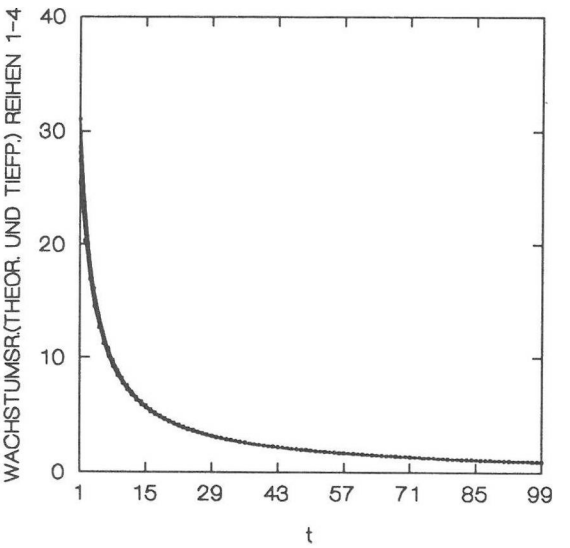


Abb. 136

Offensichtlich sind diese Raten optisch kaum voneinander und von den wahren Raten zu unterscheiden.

Die Spektren der hochpassgefilterten Wachstumsraten der mit dem CENSUS-X-11-Verfahren bereinigten Reihe zeigen in Abb. 137 folgenden Verlauf:

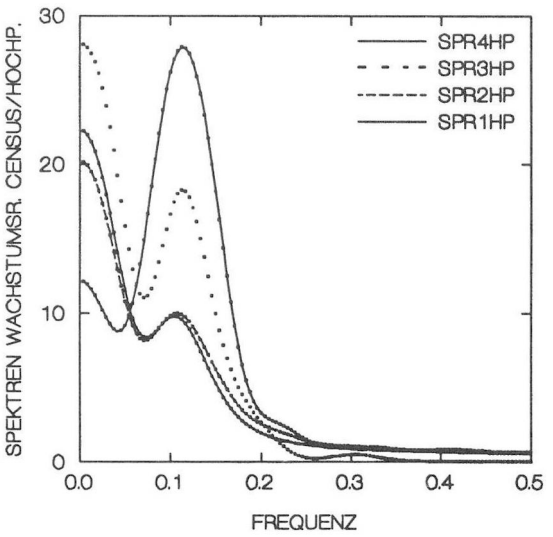


Abb. 137

Alle vier Spektren bestätigen und präzisieren den obigen optischen Eindruck. Das Verfahren indiziert einen Zyklus von 8-9 Monaten, wobei dieser bei den Reihen 3 und 4 besonders ausgeprägt ist.

Die kurzfristigen Oszillationen der glatten Komponente nach dem CENSUS-X-11-Verfahren zeigen sich nicht nur in den Wachstumsraten gegenüber der Vorperiode, die auf der Basis der glatten Komponente berechnet werden, sie können sich auch in den entsprechenden Vorjahres-Wachstumsraten bemerkbar machen. Beispielhaft zeigt die Abb. 138 die wahren Vorjahres-Wachstumsraten des linearen Trends sowie die entsprechenden Raten für die glatte Komponente nach dem CENSUS-X-11-Verfahren und dem Tiefpass-Filter.

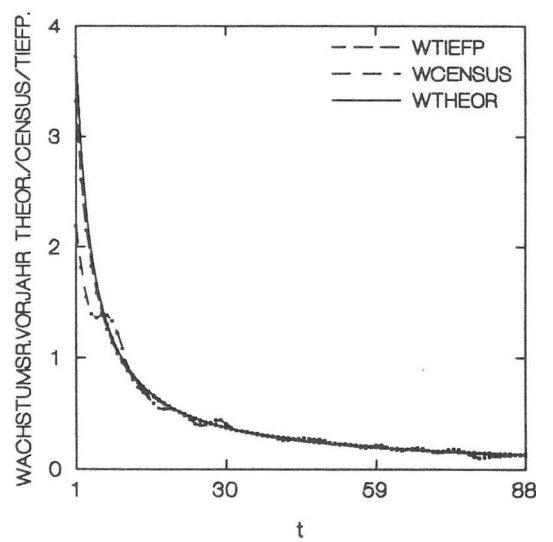


Abb. 138

Während die Wachstumsraten des Tiefpass-Filters sich kaum von den tatsächlichen Raten der Simulationsreihe unterscheiden, zeigen die Wachstumsraten, die sich aus der mit dem CENSUS-X-11-Verfahren ermittelten glatten Komponente ergeben, wiederum einen oszillativen Verlauf, wenn auch weniger ausgeprägt als bei den Vormonatsraten, was daran liegt, dass im Simulationsbeispiel ein linearer Trend zugrundeliegt.

Abschliessend sei die CENSUS-X-11-Komponente der in der Abb. 139 dargestellten Reihe betrachtet. Es handelt sich dabei um eine Reihe, die im Prinzip schon in Kapitel II, Abschnitt 5.3 untersucht wurde. Da das verfügbare CENSUS-X-11-Programm maximal 240 Werte analysieren kann, wurde diese Reihe von vorne um 60 Werte gekürzt. Das zugrundeliegende Komponentenmodell sei hier wiederholt:

$$X_t = 3000 + 0.9t + 0.2t^2 + 1000\ (\sin 2\pi/12 + \varepsilon_t)\ ,$$

mit

$$\varepsilon_t \sim N(0,2)$$

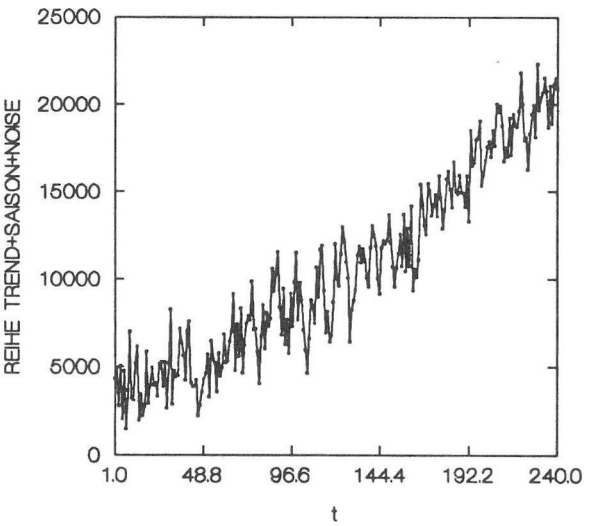


Abb. 139

Schätzt man mit dem CENSUS-X-11-Verfahren die glatte Komponente dieser Reihe, die ja hier einem Polynom zweiter Ordnung folgt, dann erhält man eine stark um die wahre glatte Komponente oszillierende Reihe, wie aus der folgenden Abb. 140 ersichtlich ist:

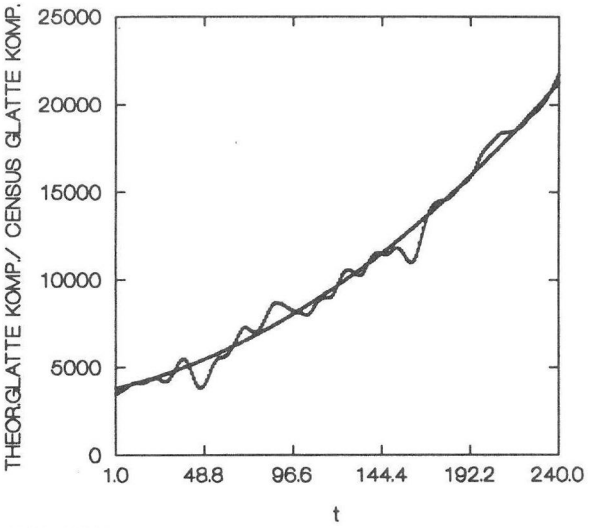


Abb. 140

Diese "glatte" Komponente dürfte für diagnostische Zwecke wohl völlig unbrauchbar sein. Dies trifft auch für ihre stark zyklisch verlaufenden Wachstumsraten zu (vgl. Abb. 141):

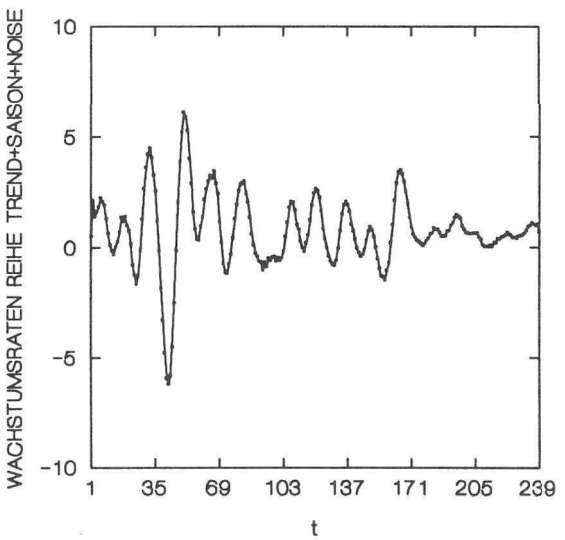


Abb. 141

Aus dem Spektrum dieser Wachstumsraten ergibt sich, dass diese eine Periodizität von ca. 18 Monaten aufweisen (vgl. Abb. 142):

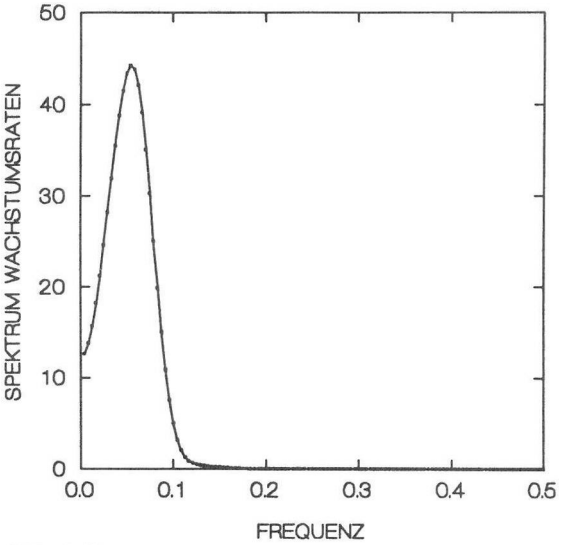


Abb. 142

Im Gegensatz zum CENSUS-Verfahren ergibt sich mit einem Tiefpass-Filter ein akzeptables Resultat. Die tatsächliche glatte Komponente und die mit einem Tiefpass-Filter ermittelte Komponente unterscheiden sich nur wenig (vgl. Abb. 143):

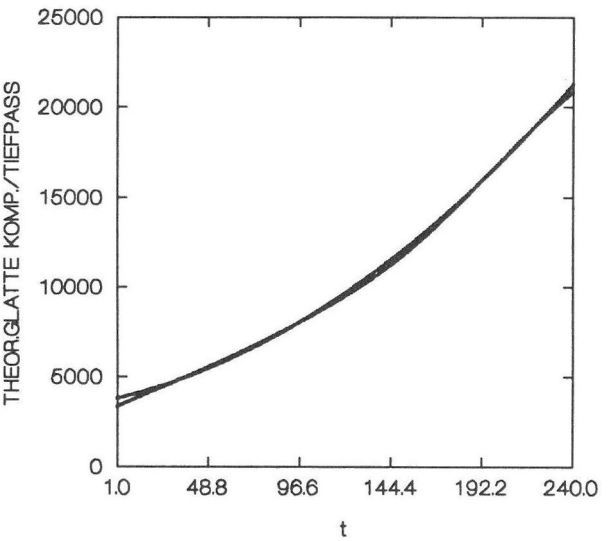


Abb. 143

Das unbefriedigende Resultat des CENSUS-X-11-Verfahrens lässt sich leicht erklären. Bei der bereinigten Reihe handelt es sich um eine sehr stark "verschmutzte" Reihe, die bewusst so konstruiert wurde, dass im Niederfrequenzbereich aussergewöhnlich viel spektrale Masse konzentriert ist, die *nicht* von der glatten Komponente, sondern von der Sinusschwingung (also der Saison) und vom noise herrührt. Dies wird deutlich, wenn man die (geschätzten) Spektren der drei Reihenkomponenten betrachtet, die nachstehend auszugsweise bis zur zweiten saisonalen Frequenz (d.h. bis zur Frequenz 2/12 = 0.1667) wiedergegeben sind. Dabei bedeuten:

SPGK : Spektrum des Trendpolynoms $3000+0.9t+0.2t^2$
SPNOISE : Spektrum von noise 1000 ε_t
SPSINUS : Spektrum von 1000 $\sin(2\pi/12)$

FREQUENZ	SPGK	SPNOISE	SPSINUS
0.004166667	1086793000	6526690	6546031
0.008333334	1008096000	6562007	6579868
0.0125	888678400	6595290	6610358
0.01666667	743644800	6595407	66605499
0.02083333	589780500	6531433	6533556
0.025	442551700	6381506	6374255
0.02916667	313724600	6139848	6129918
0.03333334	210074200	5820196	5833685
0.0375	133321100	5454762	5551754
0.04166667	81098960	5088893	5377269
0.04583333	48525940	4772599	5415260
0.05	29874610	4550844	5760439
0.05416667	19917740	4454729	6472215
0.05833333	14710750	4495451	7553007
0.0625	11787440	4662242	8936130
0.06666667	9915929	4924546	10487880
0.07083333	8644649	5237796	12025150
0.075	7856216	5551391	13345860
0.07916667	7468170	5817185	14265850
0.08333334	7318606	5996814	14653820
0.0875	7192830	6066666	14455860
0.09166667	6909065	6019878	13704040
0.09583333	6388157	5865449	12507160
0.1	5667964	5625042	11026570
0.1041667	4864193	5328386	9443570
0.1083333	4106664	5008245	7926344
0.1125	3485385	4695778	6603966
0.1166667	3028006	4416885	5552296
0.1208333	2710255	4189857	4792982
0.125	2485152	4024375	4303341
0.1291667	2311784	3921766	4032556
0.1333333	2169970	3876247	3918972
0.1375	2057952	3876826	3904148
0.1416667	1979788	3909528	3941324
0.1458333	1932764	3959573	3998082
0.15	1902578	4013244	4054658
0.1541667	1867831	4059265	4100217
0.1583333	1809713	4089638	4129118
0.1625	1720489	4099991	4138528
0.1666667	1606035	4089561	4127628

Wie man unschwer erkennt, ist in gewissen Frequenzbereichen die spektrale Masse, die der Saison und dem noise zuzurechnen sind, wesentlich grösser als die entsprechende spektrale Masse, welche auf die glatte Komponente entfällt.

Aufgrund der ungünstigen Amplitudenfunktion der glatten Komponente des CENSUS-X-11-Verfahrens sowohl in Reihenmitte als auch am Reihenrand wird diese spektrale Masse in zyklische Bewegungen transformiert, die fälschlicherweise der glatten Komponente zugeschlagen werden. Anders ausgedrückt, das CENSUS-X-11-Verfahren ist nicht in der Lage, die Reihenkomponenten zu trennen.

Es sei hier ausdrücklich betont, dass das oben dargestellte gute Tiefpass-Resultat nur möglich ist, wenn ein Tiefpass-Filter mit *extrem kurzem Passband* verwendet wird. Nur dann kann die nicht zur glatten Komponente gehörende spektrale Masse eliminiert werden. Im vorliegenden Fall wurde deshalb als Passband der Frequenzbereich [0, 0.005] gewählt.

Die eben betrachtete Reihe ist sicherlich nicht typisch für ökonomische Reihen. Die Tiefpass-Filterung zeigt aber zum einen, dass nicht immer mit einem Standard-Passband gearbeitet werden kann bzw. sollte (für "normale" ökonomische Reihen z.B. mit dem Passband [0, 0.07]). Zum andern demonstriert sie die *Flexibilität* eines filtertheoretischen Ansatzes, eine Flexibilität, wie sie mit einem Standardverfahren wie etwa dem CENSUS-X-11-Verfahren nicht zu erzielen ist.

Literaturhinweise

1) Vgl. dazu STIER W., EDEL K., 'Eigenschaften von Saisonbereinigungsverfahren im Frequenzbereich', in: K.A. Schäffer, K. Edel, P. Miebach (Hrsgb.), Probleme von Saisonbereinigungsverfahren, eine vergleichende Studie, im Druck

2) Vgl. STIER W., Verfahren zur Analyse saisonaler Schwankungen in ökonomischen Zeitreihen, Berlin-Heidelberg-New York 1980

3) Vgl. SCHÄFFER K.A., 'Beurteilung einiger herkömmlicher Methoden zur Analyse von ökonomischen Zeitreihen', in: W. Wetzel (Hrsgb.), Neuere Entwicklungen auf dem Gebiet der Zeitreihenanalyse, Göttingen 1970, S. 131-164

IV. Anwendung des CENSUS-X-11-Verfahrens auf die Zeitreihen mit Preisnotierungen und Vorschläge für das künftige Vorgehen

1. Beurteilung von CENSUS-X-11-bereinigten Zeitreihen

Untersucht man Zeitreihen, die mit dem CENSUS-X-11-Verfahren saisonbereingt worden sind, so stellt man fest, dass nicht immer alle saisonalen Frequenzen ausgefiltert wurden. Häufig ist dies den bereinigten Reihen jedoch kaum direkt anzusehen. Erst eine detaillierte Untersuchung mit Hilfe der Spektralanalyse - vorgenommen an den geeignet hochpassgefilterten saisonbereinigten Reihen - kann solche unzulänglichen Bereinigungsergebnisse aufdecken. Beispielsweise erhält man für den Totalindex folgende Spektren (vgl. Abb. 144):

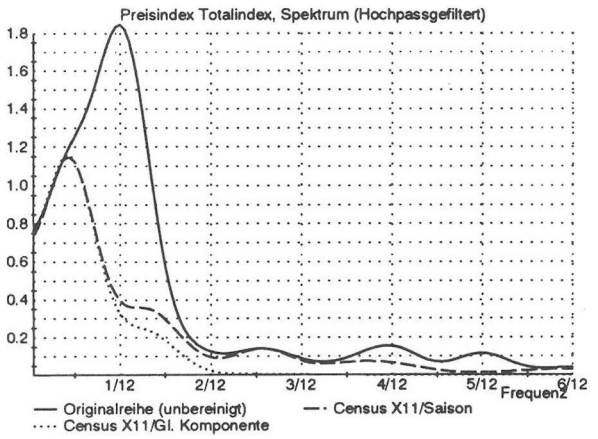


Abb. 144

Diese Abbildung zeigt, dass schon die erste saisonale Frequenz sowohl bei der Saisonbereinigung als auch bei der glatten Komonponente nur *unzureichend ausgefiltert* ist. Für die Saisonbereinigung (nicht aber für die glatte Komponente) trifft dies aber auch für die zweite, dritte, vierte und sechste saisonale Frequenz zu.

Bei anderen Reihen zeigt sich diese Eigenschaft des CENSUS-X-11-Verfahrens noch deutlicher, z.B. bei den Reihen DIENSTL und MIETE, deren Spektren (für die hochpassgefilterte Original- und saisonbereinigte Reihe sowie die hochpassgefilterte glatte Komponente) in Abb. 145 und 146 wiedergegeben sind.

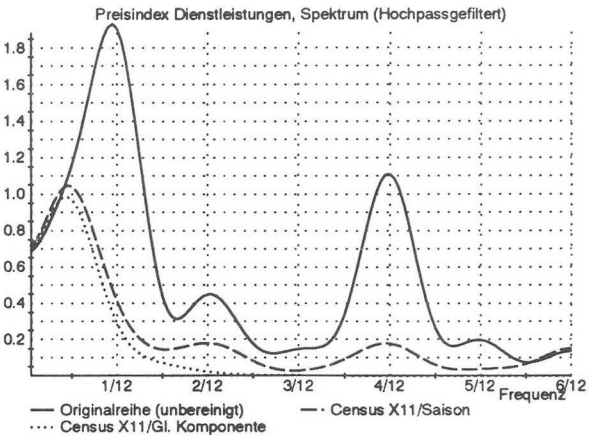


Abb. 145

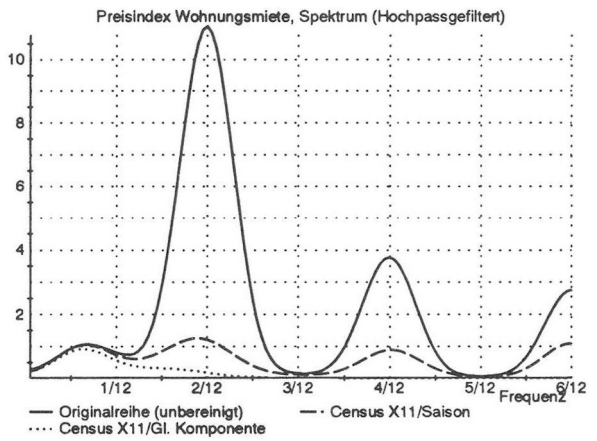


Abb. 146

Bei beiden Reihen zeigt sich insbesondere eine unzureichende Dämpfung des Quartalszyklus, aber auch des Halbjahreszyklus, der speziell die Reihe MIETE dominiert.

Die für die glatte CENSUS-X-11-Komponente unten erwähnte Randeigenschaft des "Nachzeichnens" (vgl. dazu die Ausführungen in Abschnitt 3 dieses Kapitels) ist auch für Randreihen (zu diesem Begriff sei ebenfalls auf Abschnitt 3 verwiesen), die mit diesem Verfahren saisonbereinigt werden, zu konstatieren, und zwar wiederum unabhängig davon, ob zuerst der Totalindex bereinigt wird und daraus die Inflationsraten (d.h. die Wachstumsraten gegenüber dem Vorjahr) berechnet werden oder umgekehrt. Die Resultate beider Möglichkeiten sind in der nächsten Abb. 147 dargestellt, wobei die gestrichelte senkrechte Linie den Beginn des Randes andeutet.

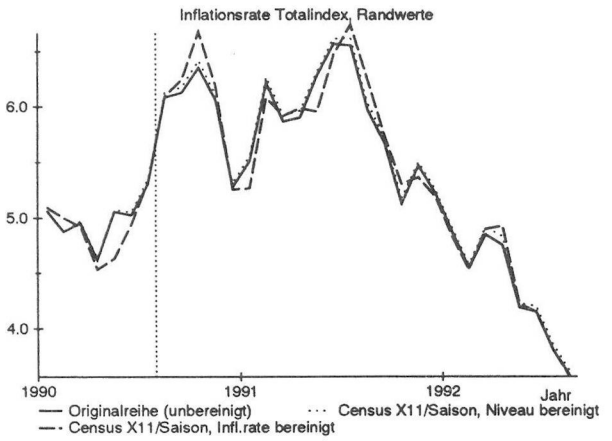


Abb. 147

Bei beiden Vorgehensweisen ist praktisch kaum eine saisonale Bereinigung am aktuellen Rand festzustellen.

In diesem Zusammenhang stellt sich die generelle Frage, ob zuerst die Niveaureihe mit dem CENSUS-X-11-Verfahren saisonbereinigt werden soll, um daraus die (saisonbereinigten) Wachstumsraten zu berechnen, oder ob eher die umgekehrte Vorgehensweise zu empfehlen ist. Betrachtet man diese Frage vom Standpunkt der Filtertheorie aus, so handelt es sich bei beiden Vorgehensweisen um eine *Reihen- (oder Serien-)* Schaltung zweier Filter, d.h. der Filter-Output des ersten Filters ist der Filter-Input des zweiten. Es kann gezeigt werden, dass der Output des Gesamtfilters von der Reihenfolge der verwendeten Filter *unabhängig* ist, falls *beide* Filter *linear* sind. Davon kann man sich leicht überzeugen, wenn man eine beliebige Zeitreihe nacheinander, z.B. durch einen gleitenden fünfgliedrigen und drei-gliedrigen Durchschnittsfilter (beides sind lineare Filter) schickt. Eine Vertauschung der Filter-Reihenfolge hat keinen Einfluss auf das Endresultat. Das ändert sich sofort, wenn nur *einer* der beiden Filter nichtlinear ist. Auch davon kann man sich leicht überzeugen, wenn man beispielsweise bei einer beliebigen Zeitreihe zuerst Wachstumsraten berechnet (dies stellt eine nichtlineare Filterung dar) und daraufhin etwa eine Glättung mit einem gleitenden Durchschnitt vornimmt. Vertauscht man hier die Reihenfolge der Filteroperationen, also glättet man zuerst und berechnet dann Wachstumsraten, so zeigt sich, dass das Endresultat von der Filter-Reihenfolge abhängt.

Da bei vorliegendem Problem *zwei* nicht-lineare Filteroperationen involviert sind, ist von vornherein klar, dass die Wahl der Vorgehensweise nicht ohne Einfluss auf das Endresultat ist. Es wird allerdings kaum möglich sein, für eine der beiden Filterfolgen eine *generelle* Überlegenheit nachzuweisen. Vielmehr scheint gemäss den bisherigen Erfahrungen eine Entscheidung zugunsten einer bestimmten Filterfolge von der Zielsetzung der Analyse, der Art und Grössenordnung der resultierenden Unterschiede, vor allem aber auch von der jeweils betrachteten Reihe abhängig zu sein. Darauf soll nun anhand konkreter Beispiele noch etwas näher eingegangen werden.

Steht bei der Bereinigung die *prospektive* Zielsetzung im Vordergrund, dann sollte die festzulegende Filterfolge nach ihren Randeigenschaften bestimmt werden. Wie oben dargelegt wurde, verstärkt der Wachstumsraten-Filter die noise-Komponente einer Zeitreihe beträchtlich. Da die Amplitudenfunktion des CENSUS-X-11-Verfahrens *ungünstige Randeigenschaften* aufweist - sowohl für die Saisonbereinigung als auch für die glatte Komponente - wird bei der Saisonbereinigung bzw. Bestimmung der glatten Komponente die aufgerauhte noise-Komponente zusätzlich noch verstärkt. In Anbetracht dieses Umstandes liegt bei Verwendung des CENSUS-X-11-Verfahrens die Empfehlung nahe, zuerst die Niveauwerte zu bereinigen und anschliessend die Wachstumsraten zu berechnen. Dass diese Empfehlung mindestens für gewisse Zeitreihen sinnvoll ist, zeigt sich gerade am Beispiel der Reihe TOTALIND. Vergleicht man den Verlauf der saisonbereinigten Randwerte dieser Reihe (vgl. Abb. 147) für beide Alternativen, so zeigt sich, dass eine Saisonbereinigung von TOTALIND mit anschliessender Berechnung der Wachstumsraten der umgekehrten Vorgehensweise vorzuziehen ist, da letztere nicht nur die Rand-Wachstumsraten nachzeichnet (vgl. dazu obige Ausführungen), sondern zusätzlich noch die Ausschläge der Wachstumsraten verstärkt. Ein Vergleich der beiden Alternativen für die glatte Komponente (vgl. Abb. 162 in diesem Kapitel) führt tendenziell zur gleichen Beurteilung.

Eine eindeutige Entscheidung scheint aber für diese Reihe nicht möglich zu sein, wenn die *retrospektive* Zielsetzung der Analyse zugrundegelegt wird. Die nächste Abb. 148 enthält die beiden alternativen glatten Komponenten der Wachstumsraten der Reihe TOTALIND.

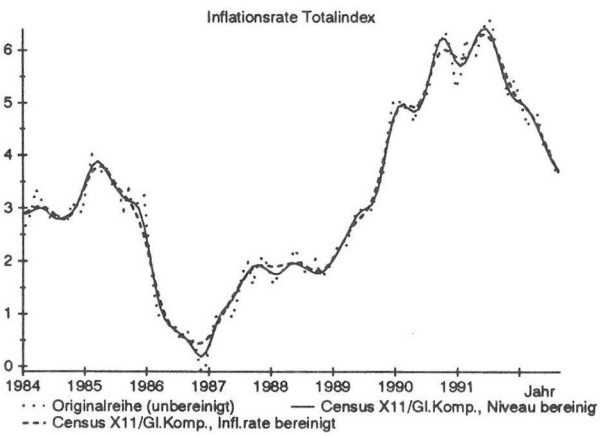


Abb. 148

Über den gesamten Zeitraum hinweg unterscheiden sich die beiden glatten Komponenten nur wenig, so dass eine Präferenz für die eine oder andere rein subjektiv wäre.

Die beiden alternativ saisonbereinigten Reihen der Wachstumsraten von TOTALIND sind zusammen mit den Inflationsraten der unbereinigten Reihe in Abb. 149 dargestellt.

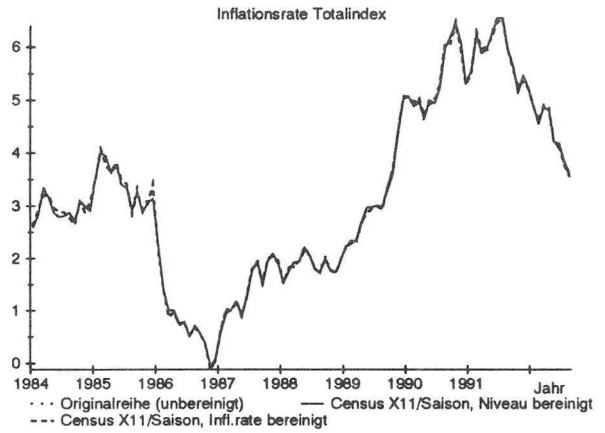


Abb. 149

Auch hier sind zwar Unterschiede vorhanden, aber durchwegs sehr kleine, so dass auch dieser Vergleich keine Entscheidungsgrundlage zugunsten einer der beiden Alternativen liefert.

Als nächstes seien die Wachstumsraten der Reihe DIENSTL betrachtet. Für die glatte Komponente erhält man folgende Resultate (vgl. Abb. 150):

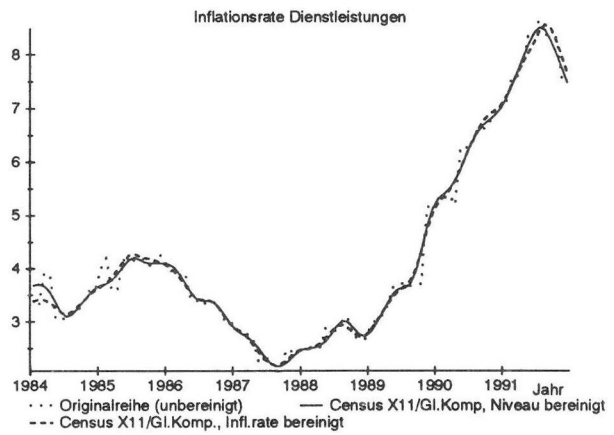


Abb. 150

Durchwegs ist eine starke (positive) Korrelation zwischen beiden glatten Komponenten mit meistens recht geringen Unterschieden festzustellen. Etwa ab dem Jahr 1991 weist die Filterfolge "Wachstumsraten/Glatte Komponente" eine Phasenverschiebung auf, so dass man wohl eher der umgekehrten Vorgehensweise den Vorzug geben dürfte. Für die saisonbereingten Wachstumsraten ergibt sich (vgl. Abb. 151):

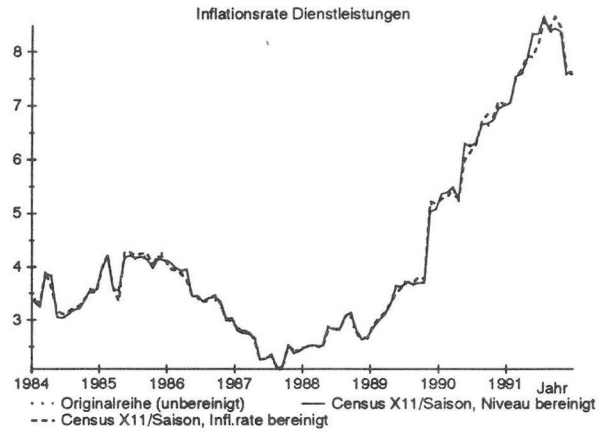


Abb. 151

Wie bei der obigen Reihe liefert auch dieser Vergleich keine Kriterien für eine Bevorzugung einer Filterfolge.

Als letztes Beispiel sei die Reihe MIETE betrachtet. Für die glatte Komponente erhält man (vgl. Abb. 152):

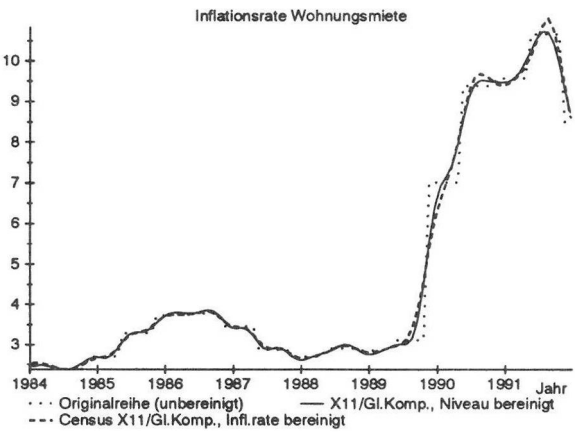


Abb. 152

Unterschiede zeigen sich etwa ab dem Jahr 1990. Die Filterfolge "Wachstumsraten/Glatte Komponente" erzeugt am Reihenende ein "oszillatives Überschiessen" der glatten Komponente, so dass man eher der alternativen Filterfolge den Vorzug geben würde. Zur gleichen Schlussfolgerung führt auch ein Vergleich der beiden saisonbereinigten Wachstumsraten, wie sich aus nachstehender Abb. 153 ergibt.

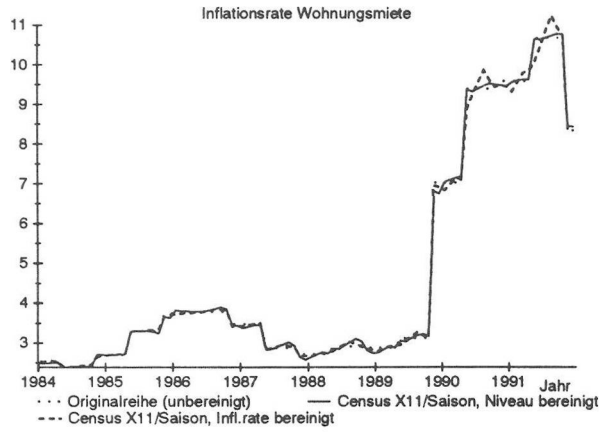


Abb. 153

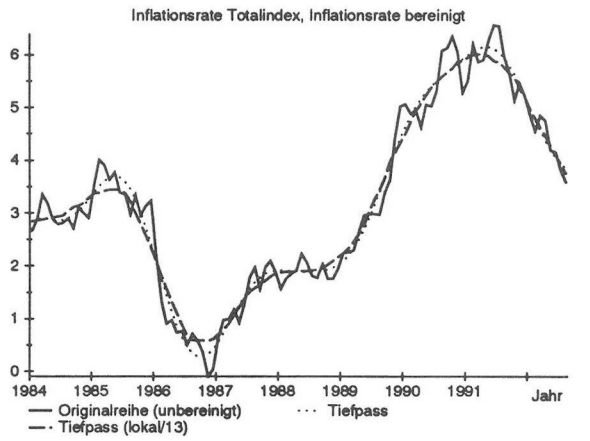


Abb. 155

Die dargestellten Beispiele zeigen, dass die Frage nach der Festlegung der Filterfolge bei der retrospektiven Zielsetzung nicht immer eindeutig entscheidbar ist, wenn die Analyse mit dem CENSUS-X-11-Verfahren durchgeführt wird. Die obigen Ergebnisse sprechen allerdings auch bei dieser Zielsetzung eher für die Filterfolge "Bereinigung der Niveauwerte/Wachstumsraten" als für die umgekehrte. Allgemein scheint eine eindeutige Entscheidung aber eher möglich zu sein, und zwar zugunsten der eben erwähnten Folge, wenn die prospektive Zielsetzung verfolgt wird.

2. Anwendungen lokaler Tiefpass-Filter

Die Abb. 154 zeigt die berechnete Inflationsrate bei Anwendung eines lokalen Tiefpass-Filters auf den Totalindex.

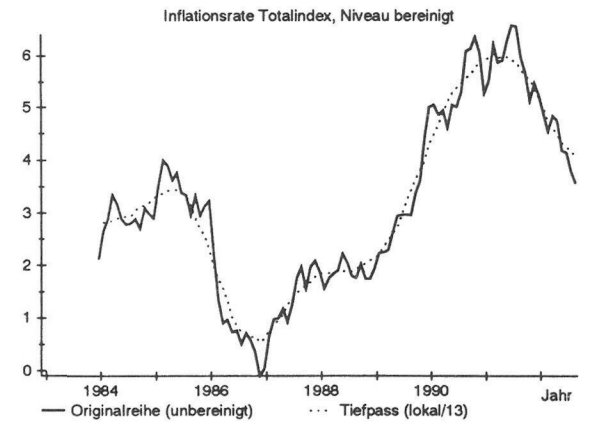


Abb. 154

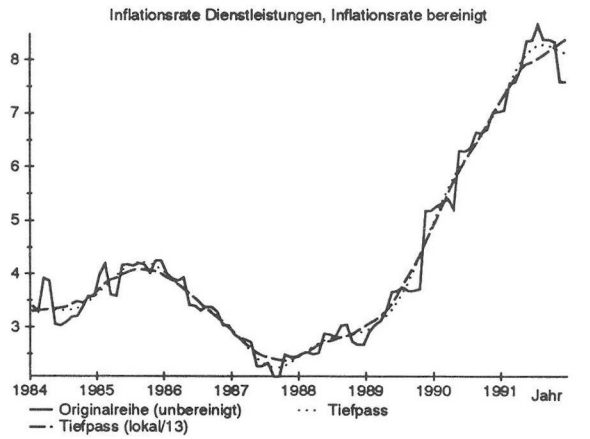


Abb. 156

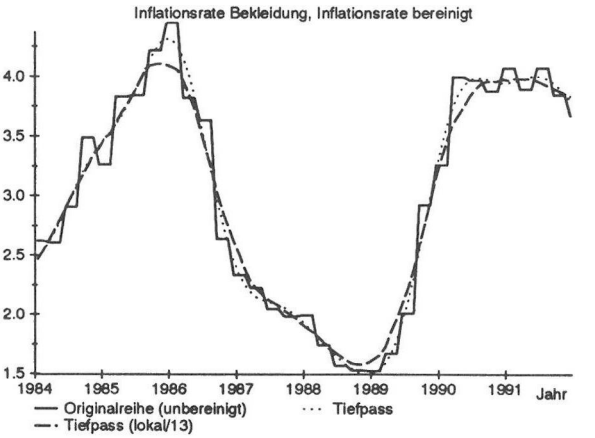


Abb. 157

In den folgenden vier Abb. 155-158 werden für die Wachstumsraten gegenüber dem Vorjahr (Inflationsraten) der Zeitreihen TOTALIND, DIENSTL, KLEID sowie MIETE jeweils globaler und lokaler Tiefpass-Filter einander gegenübergestellt.

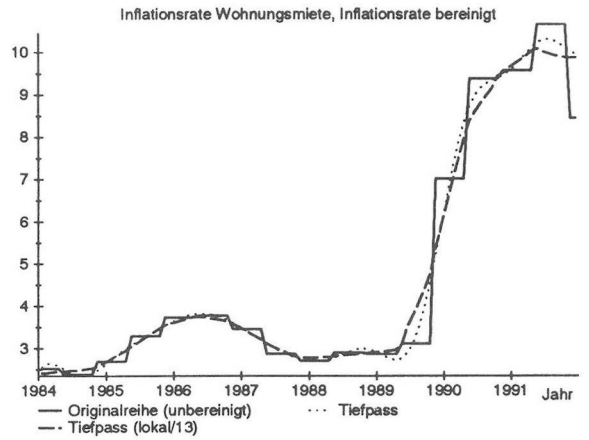


Abb. 158

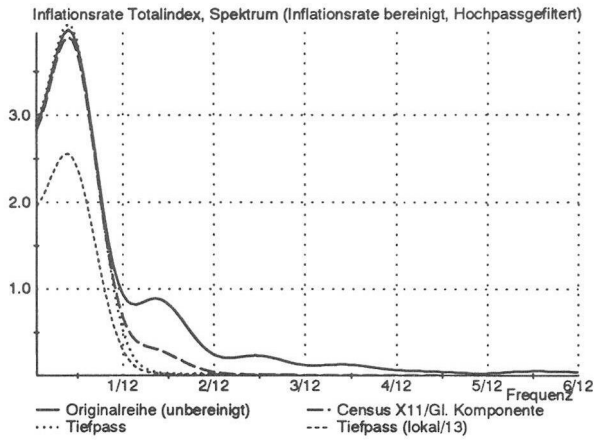


Abb. 159

Wie aus diesen Abbildungen hervorgeht, bestehen keine systematischen Unterschiede zwischen beiden Filtern. Der lokale Tiefpass verläuft jedoch etwas unruhiger und weist etwas geringere Ausschläge auf als der globale Tiefpass. Der Grund dafür ist einmal darin zu sehen, dass ein lokaler Filter eine *zeitabhängige Information* verarbeitet (die Frequenzinformationen sukzessiver Stützbereiche), während einem globalen Filter nur genau *eine* Frequenzinformation zur Verfügung steht, d.h. es liegt hier quasi *ein* Stützbereich vor. Da die Varianzen der sukzessiven Stützbereiche jeweils kleiner sind als die Varianz der Gesamtreihe, ist zum anderen zu erwarten, dass die Varianz der lokal gefilterten Reihe insgesamt kleiner ist als diejenige der global gefilterten, was die moderateren Ausschläge erklärt.

Schätzt man die Spektren der obigen global und lokal tiefpassgefilterten Zeitreihen (nach Hochpass-Filterung), dann zeigt sich, dass die Spektren der lokal gefilterten Reihen durchwegs unterhalb der Spektren der global gefilterten Reihen liegen. Beispielhaft sei dies in der nächsten Abbildung für die Zeitreihe TOTALIND gezeigt, wobei zusätzlich das (ebenfalls nach Hochpass-Filterung) geschätzte Spektrum für die glatte Komponente nach dem CENSUS-X-11-Verfahren miteinbezogen wurde (vgl. Abb. 159):

(Aus dieser Abbildung ist ausserdem zu ersehen, dass mit beiden Tiefpass-Filtern eine bessere Bereinigung im Hochfrequenzbereich zu erzielen ist als mit dem CENSUS-X-11-Verfahren.)

Damit kompatibel sind die Varianzen der (hochpassgefilterten) globalen und lokalen Tiefpass-Filter der obigen Reihen, die nachstehend für alle vier Reihen wiedergegeben sind:

Varianz der Reihe	*Globaler TP*	*Lokaler TP*
TOTALIND	0.247	0.156
DIENSTL	0.0353	0.0165
KLEID	0.0504	0.0239
MIETE	0.193	0.0959

3. Randeigenschaften des globalen und lokalen Tiefpass-Filters sowie der glatten Komponente des CENSUS-X-11-Verfahrens

Für praktische Zwecke, insbesondere für die aktuelle Diagnose, sind die *Randeigenschaften* eines Saisonbereinigungsverfahrens von besonderer Bedeutung. Man kann die Randeigenschaften eines Verfahrens dadurch ermitteln, dass man eine konkrete Reihe kürzt und sie sukzessive um je einen Wert verlängert. Daraus lässt sich eine *bereinigte Randreihe* konstruieren, indem man die bereinigte Reihe um den jeweils letzten bereinigten Wert verlängert. Dies ergibt eine Reihe, die *nur* aus Randwerten besteht und deshalb hier kurz als *Randreihe* bezeichnet werden soll.

Die Abb. 160 zeigt die beiden Randreihen der Inflationsrate für den globalen und lokalen Tiefpass-Filter, wobei die gestrichelte senkrechte Linie den Beginn des Randes anzeigt.

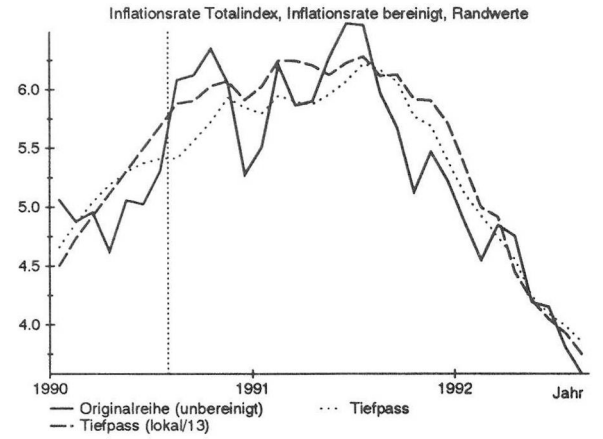


Abb. 160

Offensichtlich bereinigen beide Tiefpass-Filter die Randwerte im Gegensatz zum CENSUS-X-11-Verfahren recht gut. Die mit letzterem erzielten Randwerte der glatten Komponente weisen die Eigenschaft auf, dass sie praktisch die *Originalreihe einfach mehr oder weniger "nachzeichnen"* (mit einer Phasenverschiebung). Dies geht deutlich aus der nächsten Abbildung hervor, die für die Inflationsrate sowohl die Randwerte des lokalen Tiefpass-Filters als auch die Randwerte der glatten Komponente des CENSUS-X-11-Verfahrens enthält (vgl. Abb. 161):

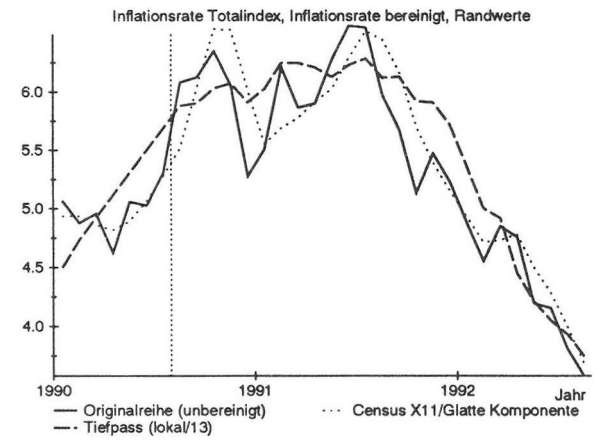


Abb. 161

Dieses Phänomen für die Randwerte zeigt sich unabhängig davon, ob man zuerst die glatte Komponente des Totalindex schätzt und daraus die Inflationsrate berechnet, oder ob man vorgängig die Inflationsrate berechnet und für diese die glatte CENSUS-Komponente schätzt. Beide Möglichkeiten sind zusammen in der Abb. 162 dargestellt.

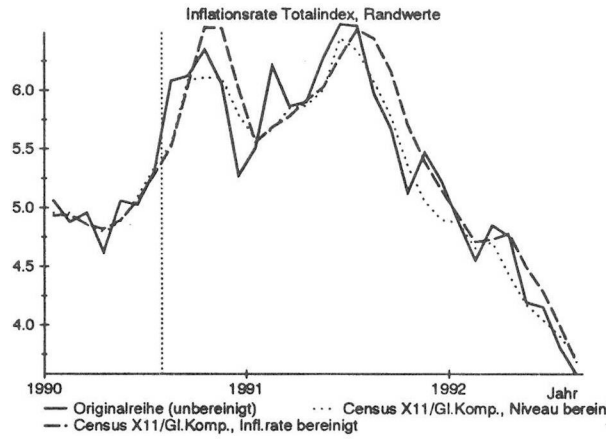


Abb. 162

Offensichtlich ist bei beiden Vorgehensweisen eine "Nachzeichnung" festzustellen. Bei der CENSUS-X-11-bereinigten Inflationsrate ist aber zusätzlich noch eine Phase von ca. einem Monat zu konstatieren.

Diese auffallend schlechten Randeigenschaften (im Zeitbereich) der glatten Komponente von CENSUS X 11 treten nicht nur bei dieser Reihe zutage. Sie sind vielmehr als ein *reihenunabhängiges Charakeristikum* dieses Verfahrens anzusehen, da sie auf die völlig unzulänglichen Eigenschaften der Rand-Amplitudenfunktion der glatten Komponente zurückzuführen sind, wie sie schon in Kapitel III ausführlich dargestellt und kommentiert wurden. Daraus lässt sich für die Praxis der folgende Schluss ziehen:

Die mit dem CENSUS-X-11-Verfahren erzielten Randwerte für die glatte Komponente sind für Diagnosezwecke als unbrauchbar zu qualifizieren.

Dass der lokale Tiefpass-Filter auch für den Rand akzeptable bereinigte Werte liefert, zeigt schliesslich eine Gegenüberstellung von lokal endgültig bereinigten mit lokalen Randwerten, wobei letztere eine leichte Phasenverschiebung aufweisen (vgl. Abb. 163):

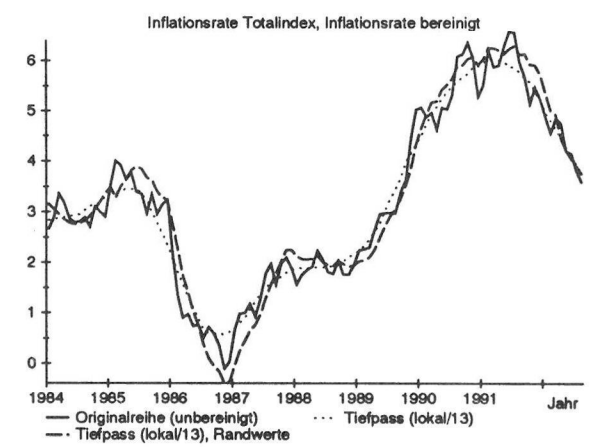


Abb. 163

4. Fazit

Die wichtigsten Resultate der vorliegenden Studie lassen sich in Kürze etwa folgendermassen zusammenfassen:

Die Untersuchungen des CENSUS-X-11-Verfahrens hinsichtlich der Eigenschaften der beiden *Rand-Amplitudenfunktionen* - für die eigentliche Saisonbereinigung und für die glatte Komponente - haben gezeigt, dass beide *besonders ungünstige Verlaufsformen* aufweisen. Dies hat *praktische Konsequenzen*. Die Randwerte einer Zeitreihe werden weder saisonbereinigt, noch wird die glatte Komponente dieser Werte bestimmt. Vielmehr ist in beiden Fällen praktisch eine *Nachzeichnung* der Bewegungen der Originalwerte, verbunden mit einer Varianzaufblähung zu beobachten. Daraus folgt, dass mit dem CENSUS-X-11-Verfahren die *prospektive* Zielsetzung einer Bereinigung *nicht zu realisieren* ist. Daran ändert auch nicht der Sachverhalt, dass die beiden Phasenfunktionen für die Randbereinigung einen günstigen Verlauf - insbesondere im Niederfrequenzbereich - aufweisen, die für die *Randreihen* nur zu einer *geringen Phasenverschiebung* führt. Wie gezeigt wurde, ist dieser lediglich die Konsequenz der beiden völlig unzulänglichen Rand-Amplitudenfunktionen.

Für aktuelle diagnostische Zwecke (prospektive Zielsetzung) ist das CENSUS-X-11-Verfahren sowohl hinsichtlich der Saisonbereinigung als auch der glatten Komponente unbrauchbar.

Als günstiger sind die Eigenschaften von CENSUS X 11 zu beurteilen, wenn eine *retrospektive* Zielsetzung verfolgt wird. Die *Amplitudenfunktion* für die *Saisonbereinigung* kommt einer *idealen sehr nahe*, und auch die *Phasenfunktion* weist praktisch eine *Nullphase* aus. Nicht ganz unproblematisch ist allerdings auch bei dieser Zielsetzung die *glatte Komponente*, da diese - wie gezeigt werden konnte - einen *Slutzky-Effekt* aufweist.

Für Untersuchungen mit retrospektiver Zielsetzung (z.B. auch für wirtschaftshistorische Analysen) kann die Saisonbereinigung des CENSUS-X-11-Verfahrens unbedenklich verwendet werden. Die glatte Komponente dieses Verfahrens kann allerdings für solche Untersuchungen nur dann ohne Vorbehalt eingesetzt werden, wenn der Slutzky-Effekt für die Analyse als unwichtig oder unbedeutend erachtet wird.

Sind neben den Niveauwerten einer Zeitreihe auch ihre Wachstumsraten (z.B. zum Vorjahr) von Interesse, so ist bei Verwendung des CENSUS-X-11-Verfahrens die Serienschaltung "Saisonbereinigung/Wachstumsraten" bzw. "Glatte Komponente/Wachstumsraten" zu empfehlen. Die umgekehrte Filterfolge würde speziell für den aktuellen Rand zu einer zusätzlich unerwünschten Aufrauhung der Reihe führen, die vom CENSUS-X-11-Verfahren aufgrund seiner ungünstigen Randeigenschaften noch verstärkt wird. Aber selbst bei dieser Filterfolge muss bedacht werden, dass sich so bereinigte Wachstumsraten am aktuellen Rand nicht für diagnostische Zwecke eignen. Die oben erwähnte Nachzeichnung der Zeitreihenwerte durch die Randwerte ist auch hier festzustellen.

Da für aktuell diagnostische Zwecke die glatte Komponente einer Zeitreihe der saisonbereinigten Zeitreihe vorzuziehen ist - letztere enthält ja noch die noise-Komponente, die eine Diagnose des aktuellen Verlaufs (geht die Entwicklung aufwärts oder abwärts oder stagniert sie?) erschwert oder unmöglich macht - stellt sich die Frage nach einer *Alternative* zu der dafür unbrauchbaren glatten Komponente des CENSUS-X-11-Verfahrens. Als Alternative wird hier die Verwendung eines *Tiefpass-Filters* vorgeschlagen, der entweder als *globaler* oder als *lokaler* Tiefpass implementiert werden kann.

Mit beiden Filtern ist eine Bestimmung der glatten Komponente am aktuellen Rand möglich, d.h. sie sind für die prospektive Zielsetzung und damit für aktuelle Diagnosen einsetzbar.

Der globale Tiefpass-Filter verläuft etwas glatter als der lokale, weist aber den *Nachteil* auf, dass sich beim *updating*, d.h. beim Hinzukommen neuer Zeitreihenwerte, auch weiter zurückliegende gefilterte Werte verändern können. Dies ist beim lokalen Tiefpass-Filter nicht der Fall. Bei einer Stützbereichslänge von 13 Monaten können sich nur die letzten sechs gefilterten Werte verändern. Die Reihe der *Randwerte* weist - ähnlich wie beim CENSUS-X-11-Verfahren - nur eine *geringe Phasenverschiebung* auf.

Bei beiden Tiefpass-Filtern ist die Filterfolge für eine Bereinigung von Wachstumsraten infolge ihrer günstigen Amplitudeneigenschaften *im Prinzip beliebig*. Da aber die Amplitudenfunktionen beider Filter keine exakten Rechteckfunktionen sein können - die Gründe dafür wurden oben dargestellt - empfiehlt sich hier die Filterfolge Wachstumsratenberechnung/Tiefpass-Filter.

Generell ist abschliessend anzumerken, dass als Alternativen zum CENSUS-X-11-Verfahren natürlich auch an Verfahren wie z.B. etwa nach KITAGAWA-GERSCH, PAULY, SCHLICHT usw. zu denken wäre, wie bereits in Kapitel II erwähnt. Allerdings sind dazu bisher noch keine Untersuchungen hinsichtlich der Eigenschaften der Amplituden- und Phasenfunktionen dieser Verfahren bekannt geworden, insbesondere keine für den praktisch besonders wichtigen aktuellen Rand. Lediglich über Randeigenschaften des Verfahrens von KITAGAWA-GERSCH liegt eine vergleichende Untersuchung (im Zeitbereich) von SCHÄFFER vor, wobei der Vergleich mit dem CENSUS-X-11-Verfahren vorgenommen wird. In Anbetracht der in dieser Studie dargestellten Randeigenschaften des CENSUS-X-11-Verfahrens dürfte es wohl kaum überraschen, dass dieser Vergleich zugunsten des Verfahrens von KITAGAWA-GERSCH ausgefallen ist [1].

Literaturhinweis

1) Vgl. SCHÄFFER K.A., 'Probleme bei der Analyse der Arbeitslosenreihe', in: L. Reyher, J. Kühl (Hrsgb.), Resonanzen, Arbeitsmarkt und Beruf - Forschung und Politik, Nürnberg 1988, S. 385-400

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat – als zentrale Statistikstelle des Bundes – die Aufgabe, statistische Informationen breiten Benutzerkreisen zur Verfügung zu stellen.

Die Verbreitung der statistischen Information geschieht gegliedert nach Fachbereichen (vgl. Umschlagseite 2) und mit verschiedenen Mitteln:

Programme des publications de l'OFS

En sa qualité de service central de statistique de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public.

L'information statistique est diffusée par domaine (cf. verso de la première page de couverture); elle emprunte diverses voies:

Diffusionsmittel	*Kontakt* *N° à composer*	*Moyen de diffusion*
Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse	031 323 60 50	Communiqués de presse: information rapide concernant les résultats les plus récents
Publikationen zur vertieften Information (zum Teil auch als Diskette)	031 323 60 60	Publications: information approfondie (certaines sont disponibles sur disquette)
Online-Datenbank	031 323 60 86	Banque de données (accessible en ligne)
Individuelle Auskünfte	031 323 60 11	Service de renseignements individuels

Nähere Angaben zu den verschiedenen Diffusionsmitteln liefert das jährlich nachgeführte **Publikationsverzeichnis.** Es kann gratis bezogen werden über 031 323 60 60.

La **Liste des publications**, qui est mise à jour chaque année, donne davantage de détails sur les divers moyens de diffusion. Pour la recevoir, il suffit de la demander au n° 031 323 60 60.

Preisstatistik

Bisher wurden die folgenden Publikationen zur Revision der Preisstatistik veröffentlicht:

- Das zukünftige preisstatistische System der Schweiz: Grundlagen, Überlegungen und Fragen, Bern 1989.
- Die Revision des Landesindexes der Konsumentenpreise, Grobkonzept, Bern 1990.
- Die Revision des Landesindexes der Konsumentenpreise, Konzept für einen neuen Landesindex, Bern 1993.
- Die Revision des schweizerischen Grosshandelspreisindexes: Grobkonzept für einen neuen Produzentenpreisindex, Bern 1990.
- Die Revision des schweizerischen Grosshandelspreisindexes: Konzept für den neuen Produzentenpreisindex, Bern 1992.
- Verbrauchserhebung 1990: Ausgaben und Einnahmen der privaten Haushalte, Bern 1992.

Für 1993 ist die Veröffentlichung der folgenden Publikationen vorgesehen:

- Grundlagen des revidierten Landesindexes der Konsumentenpreise.
- Grundlagen des revidierten Produzentenpreisindexes.

In diesen Publikationen werden die Grundlagen und Methoden der revidierten Preisstatistiken detailliert dargestellt.

- Lange Reihen des Landesindexes der Konsumentenpreise.
- Lange Reihen des Produzentenpreisindexes.

In diesen Publikationen werden die langen Zeitreihen, die sich aus der Verkettung des revidierten Landesindexes und des Produzentenpreisindexes mit den bisherigen Preisreihen ergeben, dargestellt.

Statistique des prix

Ont déjà paru sur la révision de la statistique des prix les publications suivantes:

- Aménagement futur du système statistique des prix en Suisse: Bases, considérations et questions, Berne 1989.
- La révision de l'indice suisse des prix à la consommation, Concept de base, Berne 1990.
- Révision de l'indice suisse des prix à la consommation, Conception du nouvel indice suisse des prix à la consommation, Berne 1993.
- La révision de l'indice suisse des prix de gros: Concept de base pour un nouvel indice des prix à la production, Berne 1990.
- La révision de l'indice suisse des prix de gros: Concept de détail pour un nouvel indice des prix à la production, Berne 1992.
- Enquête sur la consommation 1990: Les dépenses et les revenus des ménages privés, Berne 1992.

Pour 1993, il est prévu de publier les documents suivants, en plus de la présente publication:

- Bases de l'indice révisé des prix à la consommation.
- Bases de l'indice révisé des prix à la production.

Dans ces publications seront présentées en détail les nouvelles bases et méthodes de la statistique de prix.

- Séries de longue durée de l'indice des prix à la consommation.
- Séries de longue durée de l'indice des prix à la production.

Dans ces publications seront présentées les séries de longue durée, qui sont le résultat de l'enchaînement des séries consécutives à chacune des révisions, pour les indices des prix à la consommation et à la production.

Das Ziel dieser vom Bundesamt für Statistik (BFS) in Auftrag gegebenen Expertise besteht darin, einen Überblick über die Saisonbereinigung von wirtschaftlichen Zeitreihen zu erhalten – dies im Lichte des heutigen Forschungsstandes und im Hinblick auf die praktische Umsetzung in der amtlichen Statistik.

Neben allgemeinen Betrachtungen zur Saisonalität und zur Konstruktion von Saisonbereinigungsverfahren wird speziell das verbreitete Verfahren CENSUS-X-11 näher vorgestellt und diskutiert. Im Zentrum der Untersuchung steht dessen Verwendbarkeit für die Preiszeitreihen in der Schweiz.

Die Expertise liefert damit in theoretischer und praktischer Hinsicht einen fruchtbaren Beitrag zur Diskussion des komplexen Themas der Bereinigung von Zeitreihen in der Preisstatistik.

Bestellnummer:
047-0

Bestellungen:
☏ 031 323 60 60

Preis:
Fr. 11.–

ISBN 3-303-05158-5