

PC-Programme zur Zeitreihenanalyse: Datenmanagement, Grafik und univariate Analyseverfahren (SPSS, SYSTAT, STATISTICA, MicroTSP, MESOSAUR)

Rahlf, Thomas

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Rahlf, T. (1994). PC-Programme zur Zeitreihenanalyse: Datenmanagement, Grafik und univariate Analyseverfahren (SPSS, SYSTAT, STATISTICA, MicroTSP, MESOSAUR). *Historical Social Research*, 19(3), 78-123. <https://doi.org/10.12759/hsr.19.1994.3.78-123>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

PC-Programme zur Zeitreihenanalyse: Datenmanagement, Grafik und univariate Analyseverfahren (SPSS, SYSTAT, STATISTICA, MicroTSP, MESOSAUR)

Thomas Rahlf *

Abstract: This article compares five software-packages for time series analysis: three general statistical packages (SPSS, SYSTAT, STATISTICA), one econometric application (MicroTSP) and one program specifically designed for time series analysis (MESOSAUR). In addition to relevant statistical procedures possibilities for data transformation, data-handling and graphical facilities will also be dealt with. For SPSS and SYSTAT differences between DOS- and Windows-versions are noted. The statistical procedures for time series analysis of all three general packages are fairly similar: techniques for smoothing by moving averages, for fitting trend functions, exponential smoothing, spectral analysis and ARIMA-model building. The procedures related to the latter will be discussed in detail.

MicroTSP, which is clearly an econometric program, lacks some "classical" methods like spectral analysis and symmetric moving averages. On the other hand some newer (econometric) developments like unit root tests are implemented. Finally the program MESOSAUR will be described, which has unique data-management facilities. The various algorithms of this program even contain frequency filtering. On the other hand one can not recommend ARIMA-procedures as well as the built-in expert system. As a result some specific requirements concerning historical research will be formulated. Finally some further references to other time series analysis software and statistical programming languages will be given.

Address all Communications to Thomas Rahlf, Institut für Soziologie, Universität Halle-Wittenberg, 06099 Halle (Saale), e-mail: rahlf@soziologie.uni-halle.d400.de, Fax: (0345) 3883130. Für wertvolle Hinweise danke ich R. Metz und J. Sensch, Zentrum für Historische Sozialforschung (Köln), B. Schäfer, STATCON (Witzenhausen) und K. Christof, SPSS GmbH Software (München). Für verbleibende Fehler ist der Autor verantwortlich.

Die beiden bislang in dieser Zeitschrift erschienenen Einführungsartikel in die Zeitreihenanalyse¹ zeigen bereits eine ganze Reihe von univariaten statistischen Verfahren auf. In diesem Beitrag soll die EDV-technische Umsetzung zeitreihenanalytischer Verfahren näher behandelt werden. Dabei orientiert sich der vorliegende Artikel nicht an den dort beschriebenen Verfahren, sondern an den einzelnen Programmen bzw. den dort implementierten Prozeduren. Darüber hinaus werden auch allgemeiner Programmaufbau, Datenmanagement und Grafikmöglichkeiten beschrieben, weil sich diese Bestandteile eines Programms in der täglichen Arbeit als ebenso wichtig erweisen wie die angebotenen statistischen Verfahren.

Aufgrund der Entwicklungsgeschwindigkeit im EDV-Bereich haben Artikel wie dieser naturgemäß eine sehr geringe Halbwertszeit. Daher sollte bei den folgenden Ausführungen bedacht werden, daß manches Element, das hier beschrieben wird, zum Zeitpunkt der Lektüre bereits veraltet ist.²

Wie in nahezu jedem Spezialgebiet der Statistik gibt es auch im Bereich der Zeitreihenanalyse eine Vielzahl von Programmen, die mehr oder weniger stark spezialisiert sind. Eine Beschreibung von Software zur (univariaten) Zeitreihenanalyse muß sich daher auf eine Auswahl beschränken. Der vorliegende Beitrag beschreibt und vergleicht 5 Statistikprogramme, mit denen sich Zeitreihenanalyse betreiben läßt. Dabei ist zwischen "Allround"-Statistikpaketen, die ein möglichst breites Spektrum statistischer Analyseverfahren abdecken wollen, auf der einen Seite und Spezialprogrammen zur Zeitreihenanalyse auf der anderen Seite zu unterscheiden. Wir werden die entsprechenden Module dreier solcher allgemeinen Programme vorstellen (*SPSS*, *SYSTAT* und *STATISTICAL* sowie mit *MicroTSP* ein Spezialprogramm aus dem Bereich der Ökonometrie und schließlich mit *MESOSA UR* ein, wenn man so will, "Allround"-Zeitreihenanalyse-Programm.

Bei der Beurteilung der Zeitreihenmodule von "Allround"-Paketen, ergibt

¹ THOME, H., Beschreibende Zeitreihenanalyse: Komponentenerlegung mit Hilfe gleitender Mittelwerte, in: *Historical Social Research / Historische Sozialforschung* 17/3 (1994), S. 63-105; THOME, H., Univariate Box/Jenkins-Modelle in der Zeitreihenanalyse, in: *Historical Social Research / Historische Sozialforschung* 19/3 (1994).

² Wie kurzlebig gerade der Bereich der Software zur Zeitreihenanalyse ist, zeigt etwa der Artikel von BEAUMONT, C. / MAHMOUD, E. / McGEE, V. E., Microcomputer forecasting software: a survey, in: *Journal of Forecasting* 4 (1985), S. 305-312, der bereits völlig veraltet ist.

³ Vgl. allgemein zur Statistiksoftware SENSCH, J., *Statistische Modelle in der Historischen Sozialforschung: Anwendung formaler Modelle - Einsatz von Statistikprogramm-Paketen - Auswahlbibliographie zu statistischen Verfahren*, Köln 1994 (= *Historical Social Research / Historische Sozialforschung*, Supplement 7); EHLING, M., Anwendungsnahe Statistiksoftware - Ein Überblick über Programmpakete für den PC, in: *Allgemeines Statistisches Archiv* 78 (1994), S. 262-269.

sich das Problem, daß eine Reihe von Datenmanagement-, Grafik- und Analysemöglichkeiten zwar nicht unmittelbar in dem entsprechenden Zeitreihenanalyse-Modul enthalten ist, dafür aber in einem Basismodul oder in einem anderen Analysemodul. Dazu ein kurzes Beispiel: Trendpolynome lassen sich bei den drei hier vorgestellten allgemeinen Programmpaketen bis zu Polynomgraden ≤ 3 realisieren. Es ist aber in allen drei Paketen möglich, über ein Datenmanagement-Modul die entsprechenden Potenzen einer Zeit-Variable (oder der Fallnummer) zu erzeugen und die Parameter anschließend mit der entsprechenden Prozedur einer multiplen Regressionsanalyse zu schätzen.⁴ Ebenso ist ein direkter Befehl zur Transformation einer Zeitreihe in deren Wachstumsraten nicht erforderlich, wenn die Möglichkeit besteht, Variablen fallweise zu verschieben sowie Werte zu subtrahieren und zu dividieren⁵. Eine weitere Komplikation ergibt sich, wenn die Software nicht nur über eine Menüführung oder eine Kommandosprache, sondern auch über eine Programmiersprache verfügt. Dadurch entsteht natürlich eine nahezu unbegrenzte Flexibilität, auf die hier nicht näher eingegangen werden kann. Um diesen Vergleich nicht ins Uferlose ausarten zu lassen, werden daher nur solche Features berücksichtigt, die unmittelbar als Prozedur enthalten sind und nicht erst programmiert werden müssen. Da die meisten Anwender sicher vor dem Aufwand des Programmierens zurückschrecken, ist hiermit sicher keine große Einschränkung gegeben.

Schließlich muß noch erwähnt werden, daß mittlerweile alle großen Statistik-Software-Pakete für den PC sowohl in einer DOS als auch in einer Windows-Version angeboten werden. Während die statistischen Analyseverfahren meist identisch sind, ergeben sich in Hinsicht auf das Datenmanagement, die Grafikmöglichkeiten sowie die Speicherkapazitäten zum Teil beachtliche Unterschiede zwischen der jeweiligen DOS- und Windows-Version. Im folgenden werden daher zuerst die entsprechenden DOS-Versionen besprochen und anschließend die Unterschiede, die sich in der Windows-Version ergeben, kurz erläutert.⁶

Um verschiedene Programme hinsichtlich ihrer Qualität beurteilen zu können, läßt sich eine Reihe von Kriterien angeben. Die folgende Tabelle

⁴ Allerdings können hier, insbesondere für höhere Polynomgrade, numerische Probleme auftreten, da die X-Variablen nicht unabhängig sind und die Algorithmen zur Lösung von Gleichungssystemen innerhalb einer Multiplen Regression diese spezielle Struktur nicht berücksichtigen.

⁵ Eine Wachstumsrate ist bekanntlich definiert als

$$w_t = \frac{y_t - y_{t-1}}{y_{t-1}}$$

⁶ Bei *STATISTICA* beschränken wir uns auf die Windows-Version, da uns die DOS-Version nicht vorlag.

nach **HÖRMANN**⁷ et al. bzw. nach **EHLING**⁸ stellt einige Punkte zusammen, an denen wir uns, wenn auch nicht in dieser strikten Reihenfolge und in diesem Umfang, orientieren wollen:

Tabelle I.-

Allgemeine Auswahlkriterien zur Beurteilung von Statistiksoftware nach Ehling (gekürzt)

Nutzerfreundlichkeit

- einfache Bedienung und Nutzersprache
- gute Dokumentation der Programme und eingesetzten Algorithmen
- Qualität der Handbücher
- Einarbeitung, Aus- und Fortbildung
- einfaches Fehler-Handling

Datenverwaltung und Dateimanagement

- Datenbank (relational oder hierarchisch)
- Datendokumentation
- Veränderungen, Auswahl und Einfügen von Daten
- Dateneingabe und -löschung
- Behandlung fehlender Daten
- Import und Export von Daten

Grafische Darstellung und grafische Datenverarbeitung

- Umfang und Anwendungsbreite der verfügbaren Verfahren
- Qualität der Ergebnisse und der Grafiken

Statistische Methoden

- Umfang und Breite der statistischen Analysemethoden
 - anschauliche Präsentation der Ergebnisse der verschiedenen Methoden
 - Genauigkeit der Ergebnisse
-

Weiterhin erwähnt **EHLING** noch die Punkte "Technische Anforderungen" und "Kosten", auf die wir hier nicht weiter eingehen. Speziell für die Zeitreihenanalyse ist in bezug auf das Datenmanagement noch von Bedeutung, ob ein spezielles Format zur Verwaltung von Zeitreihen existiert.

Im folgenden wird jeweils zuerst die Dokumentation beschrieben, dann der Programmaufbau, das Datenmanagement (inklusive Datendokumentation), Datentransformations- und Grafikmöglichkeiten und schließlich die univariaten Analyseverfahren. Auf die implementierten Prozeduren zu den im einleitenden Artikel dieser Ausgabe behandelten ARIMA-Modellen wird etwas näher eingegangen. Zusätzlich wird noch das in *MESOSAUR* im-

⁷ **HÖRMANN, A., et al.**, *Comparing statistical analysis systems - a proposal for a list of comparison criteria with descriptions of BMDP, P-Stat, SAS, and SPSS*, in: **Statistical Software Newsletter** 16 (1990). S. 90-127.

⁸ **EHLING, S.** 268f.

plementierte Frequenzfilterverfahren genauer betrachtet. Da *SPSS* das am weitesten verbreitete Statistik-Paket sein dürfte, werden die einzelnen Elemente zuerst anhand dieses Paketes relativ ausführlich besprochen und bei den anderen Programmen des öfteren darauf Bezug genommen.

SPSS/PC+ Trends

Früher unter dem Namen *Statistical Package for the Social Sciences* vertrieben, heute unter "Statistical Product and Service Solutions" zählt *SPSS* zu den ältesten und am weitestverbreiteten Statistik-Programmpaketen. Zumindest in Deutschland dürfte es mittlerweile an fast jedem Universitäts-rechenzentrum vorhanden sein und bildet für die meisten Statistik-Kurse die EDV-technische Grundlage.

Wie die anderen hier zu besprechenden Pakete ist *SPSS/PC+* modular aufgebaut. Leider sind die einzelnen Teile keine Runtime-Module. Das bedeutet für den Anwender/die Anwenderin, der/die nur an zeitreihenanalytischen Verfahren interessiert ist, daß er/sie sich sowohl das Zeitreihenanalysemodul *Trends* als auch zumindest das Basismodul zulegen muß. Programmtechnisch sind die Module in *SPSS/PC+* allerdings so integriert, daß die Umgebung des Basispaketes erhalten bleibt und mit dem Kauf eines Zusatzmoduls lediglich weitere Prozeduren über Befehle zur Verfügung gestellt werden.

Die Dokumentation ist mit gut 900 DIN B5-Seiten für das Basismodul sowie nochmals 350 S. DIN A4 für das *Trends-Modul* sehr ausführlich. Neben der Befehlssyntax werden im Handbuch des Zeitreihenanalyse-Moduls auch die einzelnen Analyseverfahren in 13 Kapiteln jeweils anhand eines Beispiels verdeutlicht. Die Algorithmen des gesamten Systems werden in einem Extraband beschrieben.⁹

Beim Start des Programms erscheinen auf dem Bildschirm drei, sich teilweise überdeckende Fenster (vgl. Abb. 1). In die untere Hälfte können Befehle der *SPSS*-Kommandosprache eingegeben werden. Als Alternative können die entsprechenden Befehle von einem Menü, das sich als Voreinstellung in der oberen Bildschirmhälfte befindet und als Kommando-Generator dient, in das untere Fenster durch Auswahl übernommen werden. Dadurch wird ein Batch-Job generiert, der anschließend über eine Funktionstaste gestartet werden kann. In diesem Menü werden die Befehle ausführlich auf deutsch erläutert. Eine On-line Hilfe im üblichen Sinne gibt es nicht.¹⁰

⁹ SPSS INC., *SPSS statistical algorithms*, o.O. 1991.

¹⁰ Dafür enthält das Programm ein sog. Glossar, in dem alle Befehle sowie einige statistische Begriffe kurz beschrieben werden.

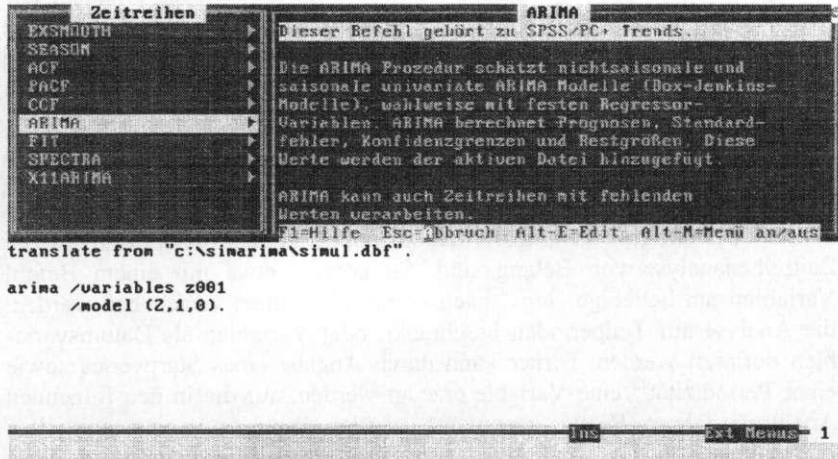


Abbildung 1: Bildschirmaufbau des Programms SPSS/PC+

Ersatzweise kann entweder ein sog. Review-Editor benutzt werden, von dem aus die Befehle direkt ausgeführt werden. Schließlich besteht die Möglichkeit, mit einem externen Editor eine Datei zu erzeugen, die die entsprechenden Befehle enthält und im Batch-Betrieb gestartet werden kann. Alle Mitteilungen des Programms werden sowohl auf dem Bildschirm angezeigt, als auch in einer ASCII-Datei protokolliert. Analyseergebnisse - und Fehlermeldungen - sind in der Regel sehr ausführlich kommentiert.

Das Datenmanagement von *SPSS/PC+* ist im allgemeinen sehr gut. Neben einem internen Format können problemlos ASCII-, dBASE, Lotus-Dateien eingelesen werden. Mehrere Dateien lassen sich auf drei Arten verknüpfen: fallweise, variablenweise und variablen weise nach Übereinstimmung in einem Schlüsselfeld. Leider können Zeitreihen nur dateiweise in den Arbeitsbereich geladen werden. Außerdem müssen alle Variablen die gleiche Fallstruktur aufweisen. So ist es z.B. nicht möglich, in einer Datei sowohl Jahres- als auch Monatswerte zu speichern. Ein spezielles Zeitreihenformat, in dem die einzelnen Werte der Variablen mit einer festen Zeitzuordnung verwaltet werden, kennt *SPSS/PC+* nicht. Die Beschränkung der Fall- und Variablenanzahl ist dabei von der *SPSS/PC+* Version abhängig." Die Datendokumentationsmöglichkeit beschränkt sich auf die Vergabe von Titeln für Datendateien. Werden die Daten im internen SPSS-Format abgespeichert,

¹¹ Seit der Version 5 kann man wahlweise eine Version erwerben, die das Extended Memory nutzt, oder sich auf die 640K. Grundspeicher beschränkt.

kann jede Variable eine bis zu 60 Zeichen lange Erläuterung erhalten. Die Möglichkeit, einzelne Werte mit erläuternden Etiketten zu versehen, dürfte bei Zeitreihen, die i.d.R. metrisch sind, kaum von Interesse sein.¹² Eine komplette Datendatei in SPSS-Format ("Systemdatei") kann zusätzlich mit einem beschreibenden Titel versehen werden.

Über das Basismodul lassen sich eine Fülle von Datentransformationen, Rekodierungen, bedingten Fallauswahlen usw. realisieren, die hier nicht bis ins einzelne beschrieben werden können. Das Modul *Trends* verfügt darüber hinaus über einige Datentransformationsmöglichkeiten, die speziell für die Zeitreihenanalyse von Belang sind. So können etwa mit einem Befehl Variablen um beliebige "lags" nach vorne oder hinten verschoben werden, die Analyse auf Teilperioden beschränkt, oder Variablen als Datumsvariablen definiert werden. Ferner kann durch Angabe eines Startwertes sowie einer Periodizität¹³ eine Variable erzeugt werden, auf die in den folgenden Analyseverfahren Bezug genommen werden kann (etwa zur zeitlichen Zuordnung von identifizierten Ausreißern etc.). Allerdings wird diese Datumsvariable als separate Variable verwaltet und nicht als Bestandteil der Zeitreihe. Zur Ersetzung fehlender Werte bietet das Programm eine lineare Interpolation, den Mittelwert oder Median der $k = 2, 4, \dots$ umliegenden Werte, den Mittelwert der gesamten Reihe oder den Wert eines linearen Trends über die gesamte Zeitreihe an der entsprechenden Stelle an.¹⁴ Sehr nützlich ist, daß in den meisten Prozeduren eine Reihe von Einstellungen als "Modell" abgespeichert werden können, das anschließend mit einem Befehl `apply` auf andere Daten übertragen werden kann. Außerdem lassen sich Analysen auf Unterperioden beschränken. Eine in der Zeitreihenanalyse sehr häufig vorkommende Situation ist die Aggregation von Werten über verschiedene Zeitintervalle. Dazu besteht die Möglichkeit lediglich über einen Befehl des Basismoduls. Will man also z.B. Monatswerte zu Jahreswerten aggregieren, muß eine Variable, die die Jahreszahlen enthält, explizit vorhanden sein. Die dann durchzuführende Aggregation bezieht sich auf den gesamten Datensatz: die Arbeitsdatei wird durch die aggregierte Datei ersetzt.

¹² Es sei denn, man möchte bei einer Datums- oder Zeitvariable zu einzelnen Zeitpunkten z.B. besondere Ereignisse notieren.

¹³ Möglich sind Sekunden, Minuten, Stunden, Tage, Wochen, Monate, Quartale, Jahre.

¹⁴ Für die Zeitreihenanalyse in den Geschichtswissenschaften wäre es natürlich wünschenswert, auf die Möglichkeit der expliziten Behandlung fehlender Werte genauer einzugehen, da gerade historische Zeitreihen oftmals sehr lückenhaft sind. Hierbei ist zu bedenken, daß fehlende Werte auch über (bedingte) Datentransformationsmöglichkeiten eines Programmes ersetzt werden können. Wenn im folgenden lediglich die ausdrückliche für fehlende Werte vorgesehenen Möglichkeiten erwähnt werden, muß diese Möglichkeit strenggenommen ebenfalls beachtet werden.

Sehr nachteilig ist, daß *SPSS/PC+* bislang noch nicht über eine Möglichkeit verfügt, Daten interaktiv, etwa in der Form eines Fall- bzw. Variablenorientierten Spreadsheets zu bearbeiten. Hier besteht als Alternative innerhalb von *SPSS/PC+* nur der (zusätzliche) Erwerb des (altertümlischen) interaktiven Datenbearbeitungsmoduls *Data Entry II*.¹⁵

In bezug auf die Grafikmöglichkeiten war *SPSS/PC+* bislang inakzeptabel, da es über kein eigenes hochauflösendes Grafikmodul verfügt und man auf Schnittstellen zu Präsentationsgrafik-Programmen angewiesen war, die natürlich nicht zur Erstellung spezieller statistischer oder wissenschaftlicher Grafiken konzipiert wurden. In der aktuellen Version 5.0.1 existiert nun eine Schnittstelle zu dem Grafikprogramm *AXUM*, das auch über einige wissenschaftliche Grafikmöglichkeiten verfügt.¹⁶ Prinzipiell ist eine solche Lösung jedoch nur als Notlösung zu bezeichnen.¹⁷

Die univariaten zeitreihenanalytischen Verfahren lassen sich in zwei Gruppen einteilen: zum einen verfügt *SPSS/PC+* über eine Reihe von Programmen zur Analyse im Zeitbereich, zum anderen gibt es auch einige Algorithmen für Untersuchungen im Frequenzbereich.¹⁸ Bei der Analyse im Zeitbereich bietet *Trends* mehrere Verfahren. Zur deterministischen Trendbestimmung kann eine Reihe von Funktionen angepaßt werden.¹⁹

Zur Glättung von Zeitreihen können einfache gleitende Mittelwerte verwendet werden; leider besteht nicht die Möglichkeit, Gewichte anzugeben. Obwohl es sich hierbei um eine sehr einfache Berechnung handelt, ergeben sich im Vergleich zu ungewichteten gleitenden Mittelwerten eine Reihe von

¹⁵ Eine abgespeckte Version ist als "Quick Editor" seit Version 5 im Basismodul enthalten.

¹⁶ Vgl. allgemein zu den Anforderungen an wissenschaftliche Grafiken GEBLER, J., *Statistische Graphik*, Basel / Boston / Berlin 1993.

¹⁷ Als Hauptkritikpunkte sind zu nennen: Zusätzlicher Lernaufwand für ein grundlegend neues Programm, höhere Anschaffungskosten, nicht immer gewährleistete vollständige Kompatibilität. Immerhin gibt es mittlerweile eine eigene Dokumentation über den Transfer zu Harvard Graphics. Vgl. SPSS INC., *Using SPSS/PC+ Graphics with Harvard Graphics*, o.O. 1993.

¹⁸ Analyseverfahren im Frequenzbereich wurden in den beiden einführenden Aufsätzen dieser Zeitschrift zwar nicht besprochen, der Vollständigkeit halber hier dennoch erwähnt. Vgl. zur Einführung in diesen Themenkomplex etwa GOTTMAN, J. M., *Time-series analysis. A comprehensive introduction for social scientists*, Cambridge usw. 1981, S. 179ff.; SCHLITGEN, R. / STREITBERG, B. H. J., *Zeitreihenanalyse*, München / Wien 1984, Kap. 6 oder STIER, W., *Studie zur Verbesserung der schweizerischen Aussenhandelsindizes. Im Auftrag der Eidg. Zollverwaltung*, o.O., hg. von der Eidg. Oberzolldirektion, Bern, Juni 1993, S. 34ff. Letzteres kann über Herrn Curwerra, Oberzolldirektion, Monbijoustraße 40, CH-3003 Bern kostenfrei bezogen werden.

¹⁹ Polynome 1., 2. und 3. Grades, logarithmische, inverse, Power-, Wachstums-, exponentielle und logistische Funktionen.

Vorteilen.²⁰ Darüber hinaus bietet das Programm die Option zur Berechnung von gleitenden Medianen und des T4253H-Glätters.²¹

Weiterhin können sowohl additive als auch multiplikative Saisonfaktoren berechnet und Zeitreihen damit adjustiert werden.²² Zur Untersuchung der internen Abhängigkeitsstruktur einer (stationären) Zeitreihe stehen Prozeduren zur Berechnung und (ASCII-) Darstellung der Autokorrelations- und partiellen Autokorrelationsfunktion zur Verfügung. Dabei können die Standardfehler nach zwei verschiedenen Verfahren gewählt werden: unter der Annahme, daß die Zeitreihe "white noise" ist oder eine MA-Struktur aufweist. Unter der Annahme, daß die Zeitreihe eine MA(q)-Struktur aufweist, werden die Standardfehler mit zunehmendem "lag" k größer.²³ Wird allerdings die Hypothese zugrundegelegt, daß die Zeitreihe "white noise" ist, so ist der Standardfehler für alle Autokorrelationskoeffizienten mit $1/\sqrt{n}$ identisch.²⁴ SPSS/PC+ berechnet den Standardfehler unter der Annahme von "white noise" dagegen nach der Formel²⁵

$$s(r_k) = \sqrt{\frac{(n - k)}{n(n + 2)}}$$

Mit zunehmendem "lag" und somit abnehmender Fallzahl werden die Standardfehler also fälschlicherweise kleiner!²⁶

²⁰ Vgl. dazu RAHLF, T., **Ein einfaches Programm zur Berechnung von gewogenen gleitenden Mittelwerten** in dBASE, in: *Historical Social Research / Historische Sozialforschung* 18/1 (1993), S. 123ff.

²¹ Bei der Berechnung des **T4253-Glätters** werden mehrere gleitende Mediane hintereinander angewendet.

²² Auf das sowohl in **SPSS** als auch in **STATISTICA** enthaltene Modul zur Saisonadjustierung nach dem Census bzw. X11-Verfahren, das in beiden Programmen umfangreiche Optionen bietet, gehen wir hier nicht näher ein. Bei diesem Verfahren handelt es sich, ähnlich wie bei dem vom Statistischen Bundesamt angewandten Berliner Verfahren zur Saisonbereinigung, um eine Reihe von heuristischen Prozeduren, die im Laufe der letzten Jahre immer weiter verfeinert und den praktischen Bedürfnissen angepaßt wurden. Vgl. als Einführung zu diesen Verfahren LEINER, B., *Einführung in die Zeitreihenanalyse*, München / Wien 1986, S. 54ff.

²³ Vgl. THOME, *Univariate Box/Jenkins-Modelle*, Formel (69).

²⁴ Vgl. ebda., Formel (70).

²⁵ Vgl. SPSS INC., *SPSS Statistical algorithms*, o.O. 1985, S. 26.

²⁶ Die Konfidenzgrenzen unter der Annahme von "white noise" werden üblicherweise zur Überprüfung von Residualreihen geschätzter Prozesse verwendet. Dabei ist allerdings zu bedenken, daß die Modellresiduen *geschätzte* Größen sind und somit die asymptotische Verteilung der empirischen Autokorrelationskoeffizienten wiederum nur asymptotisch für die auf den geschätzten Modellresiduen basierenden Koeffizienten gilt. Vgl. THOME, *Univariate Box/Jenkins-Modelle*, Abschnitt 9.

Neben den (partiellen) Stichprobenautokorrelationen und deren Standardabweichungen werden die Box-Ljung-Statistiken sowie die empirischen Wahrscheinlichkeiten ausgegeben. Der Befehl `fit` erstellt eine Reihe von (Residual-) Statistiken einer Zeitreihe. Im Rahmen einer Prozedur zur exponentiellen Glättung können insgesamt 12 verschiedene Modelle definiert,²⁷ α -, γ -, δ - und ϕ -Parameter sowie Anfangsschätzer für saisonale Faktoren vorgegeben werden. Die Parameter lassen sich lediglich über eine *grid-search-Prozedur* schätzen.²⁸

Bezüglich der ARIMA-Modellbildung bietet *SPSS/PC+* die Standardverfahren. Zur Überprüfung der Voraussetzung der Varianzhomogenität kann man vor der eigentlichen Untersuchung mit der im Basis-Modul enthaltenen Prozedur `examine` ein Mittelwert-Spannweite-Diagramm erstellen. Im Rahmen der `arima`-Prozedur können einfache sowie multiplikative saisonale Modelle vorgegeben werden, Initialwerte für die einzelnen Schätzer sowie die maximale Anzahl der Iterationen und vier verschiedene Konvergenzkriterien festgelegt werden. Bestimmte *lags* lassen sich explizit auf Null setzen. Per Voreinstellung wird zusätzlich noch eine Konstante geschätzt, wenn man dies nicht explizit unterdrückt.

Auf einen prinzipiell sehr wichtigen Punkt können wir hier nur bedingt eingehen: auf die Genauigkeit der Berechnungen. Zur Überprüfung der verwendeten Algorithmen werden in der Regel Testdatensätze verwendet, die in bezug auf die zu schätzenden Parameter Extremsituationen darstellen; das bekannteste dieser Beispiele dürfte der sog. Longley-Datensatz sein, der zur Überprüfung des Lösungsalgorithmus für das Gleichungssystem in der multiplen Regressionsanalyse dient.²⁹ Für die Zeitreihenanalyse sind uns

²⁷ Keinen Trend; linearer, exponentieller, gedämpfter Trend, jeweils zu verbinden mit: kein Saisonmodell; additives und multiplikatives Saisonmodell.

²⁸ Dabei wird von einem Startwert 0 ausgehend jeweils für 001, 002, ..., 0,9, 1 ein Modell berechnet und bezüglich der Residuenquadratsumme verglichen. Hierbei ist Vorsicht angebracht, da *SPSS/PC+ Trends* für jede Parameterkombination und jeden Wert der *grid-search-Prozedur* ein Modell generiert. Werden alle vier Parameter spezifiziert, berechnet das Programm $11^4 = 14.641$ Modelle! Außerdem speichert *SPSS/PC+ Trends* als Voreinstellung für jedes Modell sowohl die geschätzten Werte als auch die Residuen als neue Variable ab, wodurch selbst bei wenigen Parameterkombinationen die Arbeitsdatei sehr schnell aufgebläht wird. Das umfassende Gebiet der exponentiellen Glättung kann hier nicht näher behandelt werden. Vgl. allgemein den Überblick von GARDNER, E. S. Jr., Exponential smoothing: the state of the art, in: *Journal of Forecasting* 4 (1985), S. 1-28.

²⁹ Der Datensatz enthält hochredundante Prediktor-Variablen. Vgl. LONGLEY, J. W., An appraisal of least squares programs for the electronic computer from the point of view of the user, in: *Journal of the American Statistical Association* 62 (1967), S. 819-831. Allgemeine Test-Datensätze finden sich z.B. in ELLIOTT, A. C. / REISCH, J. S. / CAMPBELL, N. P., Benchmark data sets for evaluating microcomputer statistical

solche Datensätze nicht bekannt. Bei der ARIMA-Modellierung werden prinzipiell drei verschiedene Schätzverfahren angewendet: ein exaktes Maximum-Likelihood-Verfahren, ein Conditional-Least-Squares- sowie ein Unconditional-Least-Squares-Verfahren.³⁰ *SPSS/PC+* bietet nicht die Möglichkeit, eines dieser Schätzverfahren auszuwählen, sondern lediglich das Maximum-Likelihood-Verfahren.³¹ Nähere Angaben werden von den Handbüchern leider nicht geliefert.³² Bemerkenswert ist, daß die Zeitreihen auch fehlende Werte enthalten können; hierfür sieht *SPSS/PC+* einen Schätzalgorithmus vor, der auf dem Kalman-Filter basiert. Leider ist der Kalman-Filter nicht allgemein anwendbar.³³ Der Ergebnisausdruck ist recht ausführlich: angegeben werden per Voreinstellung die vom Programm oder Benutzer vorgegebenen Anfangsschätzer, die einzelnen Iterationen,³⁴ die Log-Likelihood-Statistik, zwei Informationskriterien (AIC, SBC), die endgültigen Parameter samt Standardfehler und empirischer Wahrscheinlichkeit sowie die Kovarianz und Korrelationsmatrix der Parameter. Für jede Schätzung werden die gefitteten Werte, die Residuen, obere und untere Konfidenzgrenze sowie der jeweilige Standardfehler als neue Variable abgespeichert. Leider fehlen sowohl Basisstatistiken, als auch ein Bestimmtheitsmaß.

Für die univariaten Prozeduren der Trendanpassung, der ARIMA-Modelle und der exponentiellen Glättung besteht die Möglichkeit, über einen Befehl `predict` die Arbeitsdatei um prognostizierte Fälle zu verlängern.³⁵

Hinsichtlich univariater Verfahren im Frequenzbereich lassen sich in *SPSS/PC+* über die Fourier-Transformation sowohl Fourierfrequenzen,

programs, in: *Collegiate Microcomputer II* (1989), S. 289-299 oder STATSOFT INC., *Validation benchmarks for statistical algorithms (version 1B)*, Tulsa, OK 1992.

³⁰ Vgl. dazu THOME, *Univariate Box/Jenkins-Modelle*, Abschnitt 8.

³¹ Der Algorithmus geht auf MELARD zurück. Vgl. MELARD, G., A fast algorithm for the exact likelihood of autoregressive-moving average models, in: *Applied Statistics* 33 (1984), S. 104-114.

³² Im Handbuch der statistischen Algorithmen fehlt ausgerechnet der Abschnitt über die ARIMA-Parameterschätzung. Vgl. SPSS INC., *SPSS statistical algorithms* (2.A.). Nach der ersten Auflage des Algorithmen-Handbuches wurde in der Box/Jenkins-Prozedur noch der approximative Newton-Raphson Algorithmus verwendet, vgl. SPSS INC., *SPSS Statistical algorithms*, S. 27.

³³ Der Kalman-Filter ist ein universelles Instrument zur Schätzung allgemeiner Zeitreihenmodelle, die in Zustandsraumform geschrieben werden. Vgl. dazu MILLS, T. C. *Time series techniques for economists*, Cambridge usw. 1990, S. 214ff. mit weiterführender Literatur.

³⁴ Falls vorher eine Option `/print detailed` angegeben wurde.

³⁵ Bei Schätzung eines ARIMA(2,1,0)-Modells, $N=200$, $\langle \Delta \rangle = 1$, $\langle \epsilon \rangle = -0.75$ ergaben sich bei einer Prognose merkwürdige Werte, die immer kleiner wurden, statt sich einer Konstante anzunähern.

-perioden, sinus- und cosinus-Terme der Fourierzersetzung als auch die Werte des Periodogramms und der Spektraldichte berechnen. Zur Glättung des Periodogramms stehen das Hamming-, Tukey, Parzen-, Bartlett- und Daniell-Fenster in frei angegebener Spannweite zur Verfügung. Alternativ können eigene Gewichte definiert werden.³⁶

SPSS für Windows

Grundsätzlich bietet die Windows-Version eines Programmes gegenüber der DOS-Version natürlich mehrere Vorteile: der augenscheinlichste ist die größere Bedienungsfreundlichkeit durch die grafische Benutzeroberfläche. Diese Eigenschaft ist zwar angenehm, erweist sich aber als vergleichsweise unbedeutend im Hinblick auf andere Vorteile: Im Gegensatz zu DOS ist Windows nicht auf 640K Arbeitsspeicher begrenzt, sondern kann (zumindest theoretisch) den gesamten Arbeitsspeicher einer oder mehreren Anwendungen zur Verfügung stellen.³⁷ Prinzipiell wird dadurch auch ein sog. Multi-Tasking-Betrieb möglich, in dem zeitgleich mehrere Programme laufen können. Weiterhin kann unter Windows eine Auslagerungsdatei erstellt werden, durch die sich der Arbeitsspeicher beliebig vergrößern läßt. Außerdem können Windows-Applikationen viel "enger" zusammenarbeiten: alle Programme orientieren sich an einem einheitlichen Standard, benutzen Systemkomponenten gemeinsam (wie z.B. Druckertreiber, Hilffsystem) und können über OLE ("*Object Linking and Embedding*") sowie DDE ("*Dynamic Data Exchange*") auf einfache Weise Daten austauschen. Die Übertragung eines Statistik-Programmpaketes auf die Windows-Plattform ist unseres Erachtens erst dann als gelungen zu bezeichnen, wenn diese Eigenschaften voll unterstützt werden.

Wenden wir uns nun dem Programm *SPSS für Windows* zu. Die Dokumentation des Programmes ist seit der vor kurzem erschienenen Version 6.0.1 für das Basismodul auch auf deutsch erhältlich³⁸ und umfaßt gut 700 S.

Der Programmaufbau ist durch die Windows-Oberfläche noch bedienerfreundlicher als derjenige der DOS-Version (siehe Abb. 2). Durch die Nutzung des Windows-Hilffsystems werden die einzelnen Befehle ausführlicher erläutert, als in der DOS-Version.

³⁶ Da die Werte des Periodogramms als neue Variable in die Arbeitsdatei übernommen werden können, wäre es unseres Erachtens sinnvoller gewesen, anstelle der Spektralfenster im Rahmen der Datentransformationen gewogene gleitende Mittelwerte zu implementieren. Auf diese Weise hätten sie allgemeiner genutzt werden können.

³⁷ Da Windows ein Betriebssystem-Aufsatz auf DOS ist und kein eigenständiges Betriebssystem, ergeben sich dennoch einige Beschränkungen. Vgl. generell zum wissenschaftlichen Arbeiten unter Windows MANHART, K., *Wissenschaftliches Arbeiten unter einer grafischen Oberfläche*, in: *ZA-Information* 33, November 1993, S. 94-114.

³⁸ Als Übersetzung des Originalhandbuchs.

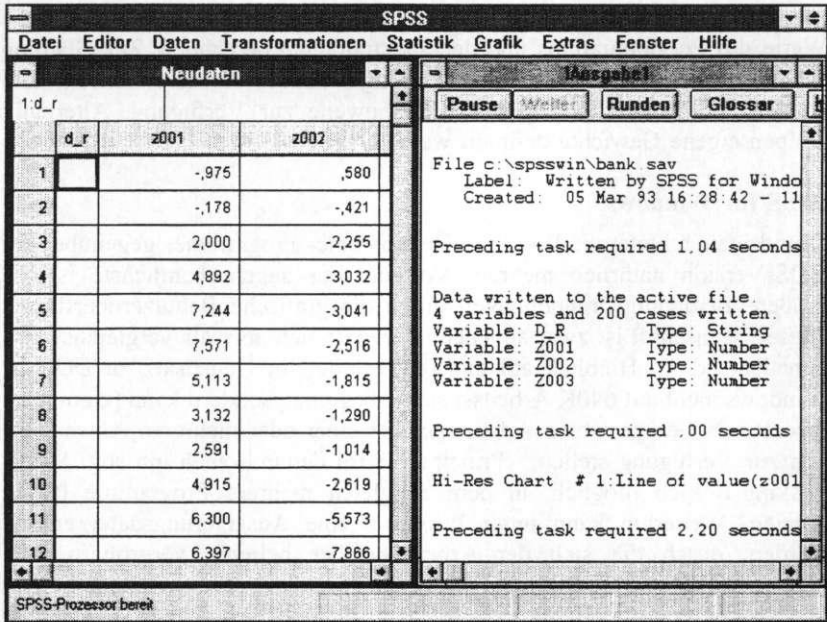


Abbildung 2: Bildschirmaufbau von SPSS für Windows

Seit kurzem ist die Menüführung deutschsprachig. Neben Dateien im eigenen Format kann *SPSS für Windows* Datendateien von Oracle- sowie des SQL-Servers direkt bearbeiten: diese müssen also nicht erst in SPSS-Format umgewandelt werden. Im Gegensatz zu *SPSS/PC+* basiert *SPSS für Windows* auf der Großrechnerversion *SPSS-X*. Insofern fallen einige Beschränkungen der DOS-Version fort. Außerdem ist der Sprachumfang der *X*-Version voll übernommen worden.³⁹ So kann etwa über einen Befehl `document` ein längerer Text zur Dokumentation eines Datensatzes eingebunden werden. Die Datentransformationsmöglichkeiten unterscheiden sich hinsichtlich der für die Zeitreihenanalyse relevanten Transformationen nicht.⁴⁰ Allerdings lassen sich nun mit der Matrixsprache auch Auswertungen realisieren, die nicht unmittelbar als Prozedur vorhanden sind. Erfahrungen hierüber liegen uns jedoch nicht vor.⁴¹ In der Windows-Version verfügt

³⁹ So z.B. die Möglichkeit der Schleifenbildung oder das Matrix-Modul.

⁴⁰ Gegenüber der DOS-Version sind vor allem die Funktionen zur Zeichenkettenmanipulation und variablenübergreifenden Verarbeitung erweitert.

⁴¹ Auf der deutschen *SPSS*-Anwendertagung 1992 in Offenburg wurde das Matrix-Modul von einigen Teilnehmern als sehr instabil bezeichnet.

SPSS nun endlich auch über einen Spreadsheet-Editor.

Die Kapazität ist durch die Windows-Plattform prinzipiell nur noch durch die Hardware beschränkt.⁴²

Erheblich besser als in der DOS-Version sind die Grafikmöglichkeiten. In *SPSS für Windows* wurde ein eigenständiges Grafikmodul implementiert, mit dem sich zu allen Prozeduren hochauflösende Grafiken erstellen lassen. Hierfür besitzt das Programm ein sog. Grafik-Karussell, in dem verschiedene Grafiken verwaltet werden können. Innerhalb dieser Grafiken lassen sich einzelne Elemente objektorientiert editieren. Die statistischen Verfahren im Bereich der univariaten Zeitreihenanalyse unterscheiden sich nicht von der DOS-Version.

SYSTAT Series

SYSTAT ist ebenso wie *SPSS* ein Allround-Statistikpaket, das für die Plattformen DOS, VAX/VMS, MAC und seit der neuesten Version 5 auch für Windows angeboten wird. Im Gegensatz zu *SPSS* sind die einzelnen Module von *SYSTAT* Runtime-Versionen, die alle im Basispaket enthalten sind. Man kann diese zwar nicht einzeln erwerben, aber immerhin sind sie alleine lauffähig, so daß man etwa nur solche installieren muß, die man wirklich benötigt. Für die Zeitreihenanalyse ist dies das Modul *Series*. Empfehlenswert sind zusätzlich die Module *Data* (zur Datenmodifikation), *Edit* (zur interaktiven Dateneditierung), *Stats* (für einfache Statistiken) sowie *SYGRAPH* (zur Erstellung hochwertiger Grafiken). Die Dokumentation besteht aus einem 255 S. (B5 Format) starken Handbuch "Getting Started", das in die grundlegenden Prinzipien des Programmes einführt, einem 248 S. umfassenden Band "Data", in dem Daten-Handling und Transformationsmöglichkeiten beschrieben werden, einem 547 S. starken Handbuch über das Grafikmodul *SYGRAPH* sowie einem 677 S. starken Handbuch "Statistics", in dem auf 67 Seiten das Modul *Series* behandelt wird. In diesem Kapitel werden anhand konkreter Beispiel-Daten, die auf Diskette mitgeliefert werden, die einzelnen Zeitreihenanalyse-Prozeduren erläutert. Erfreulich ist, daß die einzelnen Kapitel jeweils mit einem Abschnitt abschließen, in dem die Algorithmen erläutert werden.

SYSTAT verfügt wie *SPSS/PC+* über ein Menüsystem, von dem aus Befehle einer Kommandosprache in eine Kommandodatei durch Auswahl übertragen werden können (Abb. 3). Alternativ dazu läßt sich das Menüsystem auch abschalten und man kann die Befehle direkt eingeben oder vorher mit einem

⁴² Für größere Datensätze bzw. speicherintensive Berechnungen muß zuvor in einem Dialogfeld "Einstellungen" der für *SPSS für Windows* verfügbare Arbeitsspeicher heraufgesetzt werden.

Editor in einer Kommandodatei zusammenfassen und im Batch-Betrieb starten.

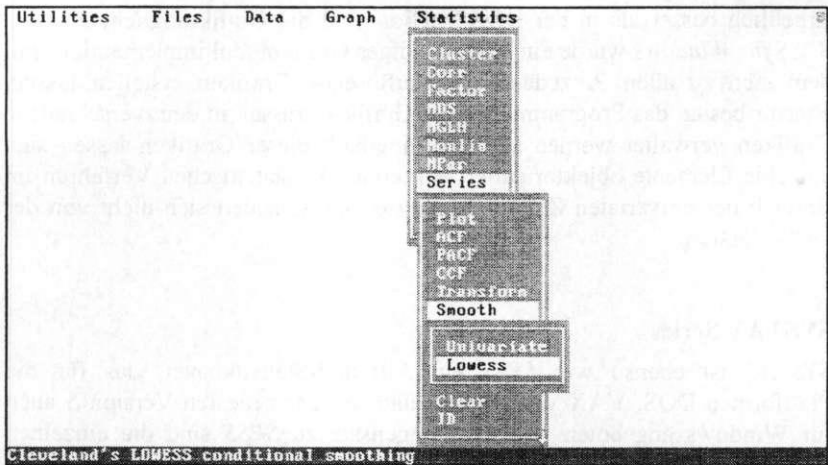


Abbildung 3: Bildschirmaufbau von SYSTAT

Das Datenmanagement, Im- und Export sowie Datentransformationsmöglichkeiten entsprechen weitgehend denen von *SPSS/PC+*. Interaktiv können Daten in einem Spreadsheetmodul *Edit* bearbeitet werden. Die Speicherkapazität ist bei einigen Verfahren auf den Arbeitsspeicher beschränkt.⁴³ Darüber hinaus bietet *SYSTAT* allerdings noch die Möglichkeit, komplexere Probleme anhand zweier "Programmiersprachen" zu behandeln: dies ist zum einen das sog. *Data Basic*, eine um Basic-Elemente erweiterte Kommandosprache sowie zum anderen *MACRO*, mit der sich sogar kleinere Anwendungen schreiben lassen. Hinsichtlich der Grafikmöglichkeiten gehört *SYSTAT* durch das Modul *SYGRAPH* sicher zum besten, was derzeit auf dem Markt ist.⁴⁴ Durch den flexiblen Aufbau des Grafik-Moduls lassen sich prinzipiell beliebig komplexe Abbildungen erstellen.⁴⁵ Obwohl beispielsweise keine

⁴³ Reicht der Arbeitsspeicher nicht aus, erscheint die lapidare Meldung: "*The problem is too large for this version of SYSTAT*".

⁴⁴ In der angekündigten Version 6.0 von *SYSTAT* ist der Grafikprogrammteil vollständig überarbeitet und ist dann unter anderem objektorientiert aufgebaut. Zum Zeitpunkt der Besprechung lag diese Version allerdings noch nicht vor.

⁴⁵ Vgl. dazu demnächst RAHLF, T., Statistische Graphik: ausgewählte Beispiele, in: STEENWEG, Helge, u.a. (Hgg.), *Geschichte und EDV: Probleme und Fortschritte, "Probleme mit dem Fortschritt?"*, Ergebnisse der 1. Tagung der Arbeitsgemeinschaft Geschichte und EDV, Göttingen, 23.-26.3.1994.

speziellen Grafiktypen zur saisonalen Analyse vorgegeben sind, lassen sich (mit einigem Aufwand) sogar sog. *Seasonal-Subseries-Plots* erstellen:⁴⁶

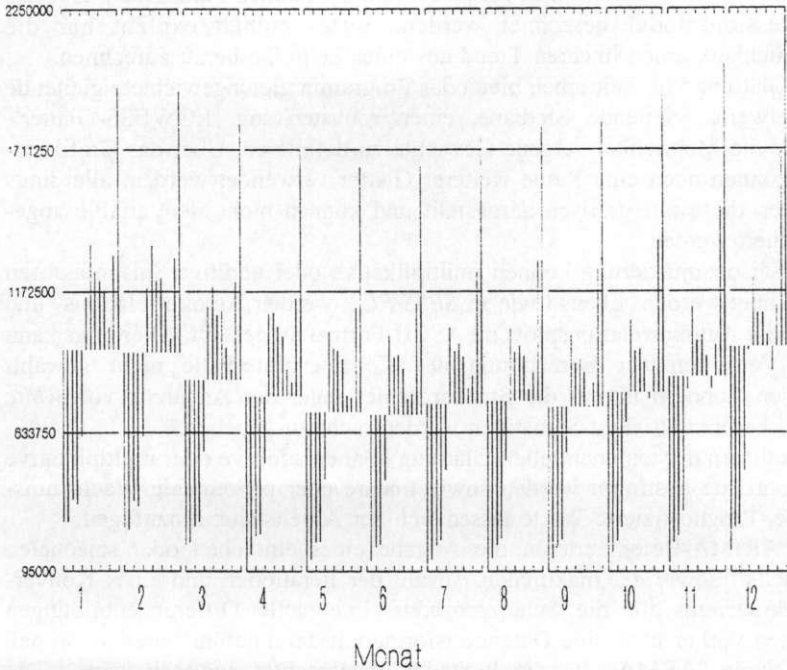


Abbildung 4: Beispiel für einen mit SYGRAPH erzeugten *Seasonal-Subseries-Plot*

Etwas umständlich ist, daß sich Befehle zu den einzelnen statistischen Analysen und zur grafischen Darstellung nicht gleichzeitig ausführen lassen, da jedesmal das Modul gewechselt werden muß. Hinsichtlich der univariaten zeitreihenanalytischen Verfahren bietet *SYSTAT Series* wenig mehr als die grundlegenden Verfahren. Zusätzlich zu den Datentransformationsmöglichkeiten des Data-Moduls kann beispielsweise noch eine Transformation in Wachstumsraten vorgenommen werden oder eine Indizierung mit einem

⁴⁶ Vgl. dazu CLEVELAND, W. S. / TERPENING, I. J., *Graphical methods for seasonal adjustment*, in: *Journal of the American Statistical Association* 11 (1982), S. 52-62.

vorgegebenen Referenzwert. Solche Transformationen hätten sich natürlich auch problemlos mit den im *Data-Modul* vorhandenen Funktionen realisieren lassen. Eine Gelegenheit, Zeitreihen zu aggregieren, besteht nicht. Fehlende Werte lassen sich durch ein "distance weighted least Squares"-Verfahren interpolieren. Zur Trendbestimmung können Funktionen über das Regressionsmodul berechnet werden. *Series* enthält explizit nur die Möglichkeit, einen linearen Trend aus einer Zeitreihe herauszurechnen. Zur Glättung von Zeitreihen bietet das Programm gleichgewichtete gleitende Mittelwerte, gleitende Mediane, einen robusten sog. LOWESS-Glätter⁴⁷ sowie die Möglichkeit, eigene Gewichte zu definieren. Über das Grafikmodul können noch eine Reihe weiterer Glätter verwendet werden; allerdings werden diese nur grafisch dargestellt und können nicht als Variable abgespeichert werden.

Zur Saisonadjustierung können multiplikative oder additive Saisonfaktoren berechnet werden. Ebenso wie in *SPSS/PC+* werden Autokorrelations- und partielle Autokorrelationsplots im ASCII-Format dargestellt. Allerdings kann das Verfahren zur Berechnung der Konfidenzintervalle nicht gewählt werden, sondern ist auf die Standardfehler unter der Annahme von *white noise* beschränkt. Test-Statistiken werden nicht ausgegeben.

Im Rahmen der exponentiellen Glättung können additive oder multiplikative Saisonterme bestimmt werden sowie lineare oder prozentuale Wachstumsterme. Prognostizierte Werte lassen sich der Arbeitsdatei hinzufügen.

Der ARIMA-Befehl erlaubt die Angabe eines einfachen oder saisonalen Modells⁴⁸ sowie der maximalen Anzahl der Iterationen und eines Konvergenzkriteriums für die Schätzprozedur. Eventuelle Differenzenbildungen müssen vorher über eine Datentransformation durchgeführt werden, so daß der Name "ARMA" für die Prozedur angebracht gewesen wäre.⁴⁹ Die Parameterschätzer werden von *SYSTAT Series* approximativ nach **McLEOD** u. **SALES** ermittelt (CLS-Verfahren).⁵⁰ Es besteht zwar die Möglichkeit, über

⁴⁷ LOWESS bedeutet *locally weighted regressing and smoothing scatterplots*.

⁴⁸ Allerdings können keine lags explizit ausgeschlossen werden. So ist es z.B. nicht möglich, ein AR(3)-Modell zu schätzen, in dem die ersten beiden Parameter gleich Null gesetzt sind.

⁴⁹ Das ist insbesondere dann nachteilig, wenn aufgrund von ARIMA-Modellen Prognosen erstellt werden sollen. Um die Originalmetrik wiederzuerlangen, muß zu den prognostizierten Werten zuerst der Wert des letzten Originalzeitpunktes addiert werden; anschließend sind noch die entsprechenden Prognosewerte aufzusummieren. Da *Series* hierfür keine Funktion vorsieht, ist solch eine Berechnung nur sehr umständlich zu realisieren.

⁵⁰ MCLEOD, A. I. / SALES, P. R. H., An algorithm for approximate likelihood calculation of ARMA and seasonal ARM A models. Algorithm AS 191, in: *Applied Statistics* 32 (1983), S. 211-223. Vgl. WILKINSON, L., *SYSTAT: The system for statistics*, Evanston,

eine **backcast**-Option Startwerte zu schätzen (ULS-Verfahren), aber man kann selbst keine Anfangswerte festlegen. Ein **forecast**-Parameter erzeugt zwar Prognosewerte, fügt diese aber nicht zur aktiven Datei hinzu. Das Listing der AR(1)MA-Prozedur liefert die Fehlerquadratsummen der einzelnen Iterationsschritte, die endgültigen Parameter sowie deren asymptotische Standardfehler und Konfidenzintervalle, die asymptotische Korrelationsmatrix der Parameter und eine Liste der prognostizierten Werte. Informationskriterien oder Bestimmtheitsmaße werden nicht ausgegeben. Die frequenzanalytischen Verfahren entsprechen weitgehend denen von SPSS.

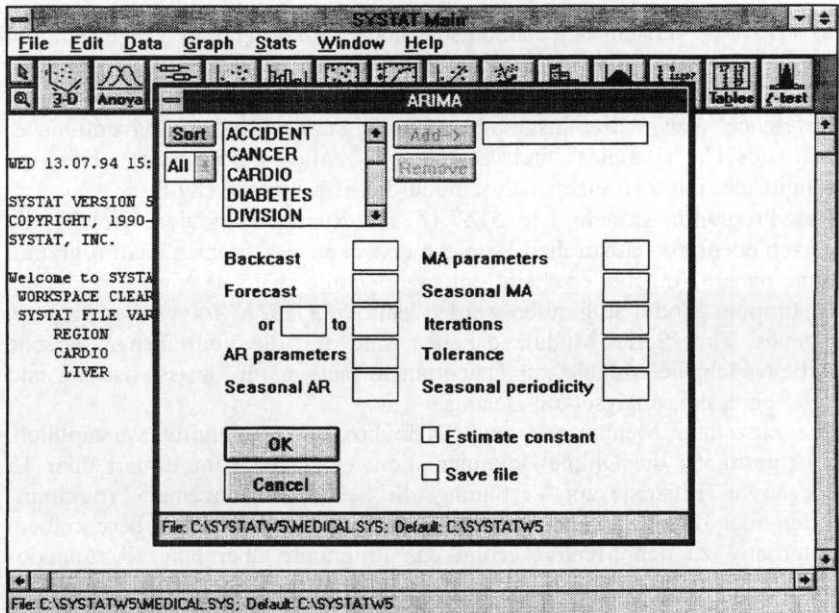


Abbildung 5: Bildschirmaufbau von SYSTAT for Windows

SYSTAT for Windows

Ebenso wie die Windows-Version von SPSS ist SYSTAT for Windows zunächst einmal bedienerfreundlicher und noch "interaktiver" (siehe Abb. 5). Hinsichtlich der univariaten Zeitreihenanalyse unterscheidet es sich nicht von der DOS-Version. In mancher Hinsicht werden die Windows-Möglichkeiten jedoch nicht voll ausgenutzt: so werden etwa die Autokorrelationsfunktionen nur als ASCII-Grafik dargestellt.⁵¹ Auch der Speicher scheint nicht voll genutzt zu werden.⁵²

STATISTICA/w

STATISTICA versteht sich als Statistik- und Grafikprogramm. Das Programm besteht wie SYSTAT aus Runtime-Modulen, die separat installiert werden können.⁵³ Als Dokumentation wird ein 150-seitiges Handbuch "Quick Reference" sowie drei insgesamt 2506 S. starke Bände "Conventions & Statistics I", "Graphics" und "Statistics II" (alle in B5-Format). Der Abschnitt, der das Zeitreihenanalysemodul betrifft, umfaßt 189 S.

Das Programm besteht wie SYSTAT aus Runtime-Modulen, welche die verschiedensten statistischen Bereiche abdecken. Startet man das Programm ohne nähere Angabe, erscheint ein sog. Module-switcher, von dem aus ein bestimmtes Modul aufgerufen werden kann. STATISTICA/w verfügt über ein eigenes Time-Series-Modul; daneben sind für die zeitreihenanalytische Arbeit noch die Module zur Datentransformation, für Basisstatistiken und zur Regressionsanalyse von Belang.

Die einzelnen Menüpunkte und Dialogboxen sind intuitiv verständlich. Einzigartig ist die Online-Dokumentation: es stehen komprimiert über 15 Megabyte Hilfstexte zur Verfügung, die neben den einzelnen Programmteilen auch die entsprechenden Verfahren zum Teil ausführlich beschreiben. Alternativ zu den Menüs verfügt das Programm über eine Kommandosprache SCL (STATISTICA Command Language), die in der Logik im großen und ganzen den Kommandosprachen von SPSS oder SYSTAT entspricht. In bezug auf den Austausch mit anderen Programmen zeigt sich STATISTICA/w recht offen: es kann Excel-, Lotus-, QuattroPro-, Paradox-, SPSS-Portable-, ASCII-, DIF- sowie Dateien im sog. *Megafile-Manager*-Format (siehe dazu unten) importieren. Daneben gibt es verschiedene interne

⁵¹ In den in Kürze erscheinenden Versionen 6.0 (DOS und WINDOWS) werden die entsprechenden Grafiken hochauflösend dargestellt.

⁵² Es gibt nach wie vor Kapazitätsprobleme, auch wenn physikalisch ausreichend Speicher vorhanden ist. In Version 6.0 soll dieses Problem weitgehend behoben worden sein.

⁵³ Da uns lediglich die Windows-Version vorlag, müssen wir auf eine Besprechung der DOS-Version verzichten.

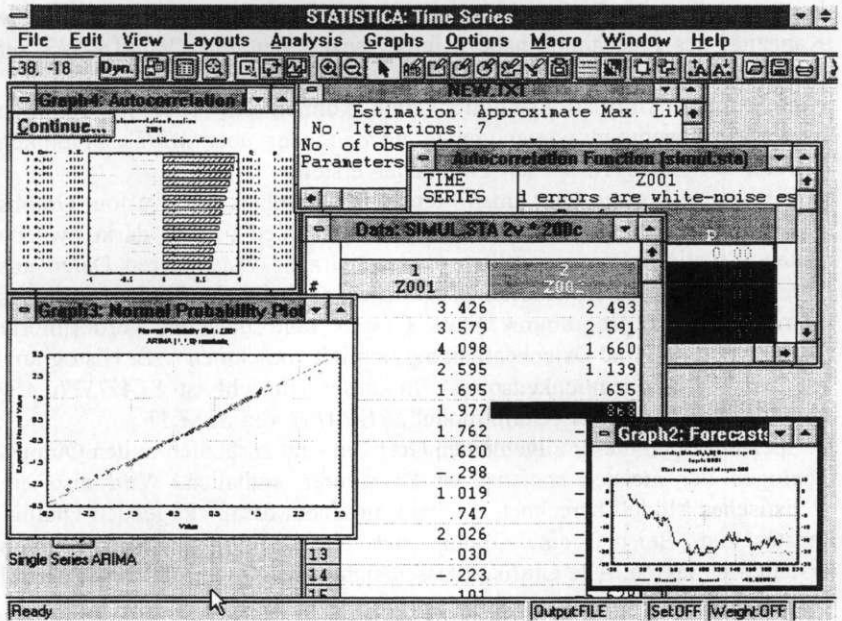


Abbildung 6: Bildschirmaufbau von STATISTICA/w

Datei-Formate, die jeweils in unterschiedlichen Fenstern dargestellt werden können: *Spreadsheet-Dateien* enthalten die Daten, *Scrollsheet-Dateien* enthalten Ergebnisse der statistischen Analysen⁵⁴, Grafiken werden in einem dritten Format gespeichert und schließlich kann man die Ergebnisse einer kompletten Sitzung oder eines Teils noch in einer Textdatei mitprotokollieren lassen.

STATISTICA/w verfügt darüber hinaus über einen sog. *Megafile Manager*, mit dem extrem große Dateien verarbeitet werden können: jede solcher Dateien kann bis zu 32.000 Spalten, jede Spalte bis zu 255 Zeichen enthalten, insgesamt also jede Zeile bis zu 8 Megabyte. Die Anzahl der Datensätze ist nur durch die Speicherkapazität der Hardware (i.d.R. Festplatte) beschränkt. Mit dem *Megafile Manager* können Daten aus unterschiedlichen Anwendungsprogrammen über Relationen oder Hierarchien zusammengesetzt werden. Dazu existiert eine eigene Programmiersprache *MML (Megafile Manager Language)*, die hier aufgrund ihrer Komplexität nicht näher

⁵⁴ Im Gegensatz zu SPSS oder SYSTAT sind diese Dateien ebenfalls wie Datendateien aufgebaut, so daß sie wiederum als Input für weitere Analyseverfahren dienen können. Eine ähnliche Lösung ist in SAS implementiert.

besprochen werden kann.⁵⁵ Auch außerhalb des *Megafile Managers* ist die Kapazität des Programms nahezu unbegrenzt:⁵⁶ der unter WINDOWS zur Verfügung gestellte Speicher wird konsequent genutzt.⁵⁷

Die Datentransformationmöglichkeiten können entweder im Rahmen einzelner Prozeduren vorgenommen werden oder über *MML*. Graphiken werden in *STA TISTICA/w* auf zwei Arten erstellt:

1. Zum einen können über einen Menüpunkt Graphics sog. Custom-Graphs oder Stats-Graphs erstellt werden. Der Unterschied besteht darin, welche Daten visualisiert werden sollen. Custom Graphs visualisieren Daten aus Spreadsheets und Scrollsheets, entsprechen also den Möglichkeiten eines normalen Grafikprogramms. Stats Graphs sind dagegen vordefinierte Grafiktypen, die eine Datenbearbeitung implizit vornehmen (z.B. Histogramme oder Wahrscheinlichkeitsplots). In dieser Hinsicht ist *STATISTICA/w* ähnlich mächtig wie das Grafikmodul *SYGRAPH* von *SYSTAT*

2. Spezialformen dieser allgemeinen Grafiken sind zusätzlich in den Output-Dialogen der meisten statistischen Prozeduren enthalten. Wird also ein statistisches Modell berechnet, erscheint in einem darauf folgenden Dialogfenster zum Beispiel eine Option, sich ein Histogramm der Residuen erstellen zu lassen. Alle Grafiken lassen sich nahezu bis in das letzte Detail editieren.

Die statistischen Verfahren zur univariaten Zeitreihenanalyse sind bei *STATISTICA/w* in Gruppen zur ARIMA-Modellierung, exponentiellen Glättung, das Census 1 und Census 2-Verfahren⁵⁸ sowie zur Spektralanalyse eingeteilt. Fehlende Werte kann man allgemein durch lineare Interpolation, den Mittelwert/Median aller oder der *n* umliegenden Punkte oder Werte einer linearen Regression ersetzen. Trendfunktionen können wiederum über ein Regressionsmodul berechnet werden.⁵⁹ Eine Reihe von Glättungen (inklusive "*distance weighted least squares*" und exponentiell gewichtete gleitende Mittelwerte) kann entweder über das Graphikmodul oder als interaktive Transformation im Rahmen der einzelnen Module vorgenommen werden. Die Prozedur zur exponentiellen Glättung orientiert sich explizit an

⁵⁵ In der Syntax ist die Sprache an Pascal angelehnt, enthält aber zusätzlich eine Reihe von Funktionen aus dem Datenmanagementbereich.

⁵⁶ So können etwa Korrelationsmatrizen bis zur Größe von 32.000*32.000 verarbeitet werden.

⁵⁷ Bei der Arbeit mit einigen im Rahmen einer Monte-Carlo-Simulationsstudie erzeugten Datensätzen ergaben sich keinerlei Speicherprobleme, selbst bei einem Umfang von N = 100.000.

⁵⁸ Obwohl die Wahlmöglichkeiten innerhalb dieser Verfahren beträchtlich sind, gehen wir (wie bei den anderen Programmen) auf deren Implementierung nicht näher ein.

⁵⁹ Über das Modul zur nichtlinearen Regressionsanalyse kann auch ein unbekannter Bruchpunkt in einer linearen Regression geschätzt werden.

der Übersicht von **GARDNER**; wie bei *SPSS* können 12 Modelle bestimmt werden, zusätzlich zur Möglichkeit des grid-search kann noch über den quasi-Newton-Algorithmus entweder der mittlere quadratische Fehler, der mittlere absolute Fehler oder der mittlere absolute prozentuale Fehler minimiert werden. Die ARIMA-Prozedur bietet die üblichen Optionen (Angabe der Ordnung für einfache oder saisonale Modelle, einige Transformationen, Iterations- und Konvergenzkriterien). Hinsichtlich des Schätzalgorithmus kann zwischen dem approximativen ULS- oder CLS-Verfahren nach **McLEOD** und **SALES** oder dem exakten **ML**-Verfahren nach **MELARD** gewählt werden. Wählt man letzteres, so wird dennoch zuerst das approximative Verfahren angewendet, um Startwerte für das exakte **ML**-Verfahren zu erhalten, in der Regel sind dann nur noch wenige Iterationen des **ML**-Verfahrens nötig. Sollten dennoch die Iterationen nicht konvergieren, gibt der Algorithmus einen sog. "Penalty-Value" aus, der den Iterationsprozeß in der Regel in eine andere Richtung lenkt.⁶⁰

Nachdem die Parameter geschätzt wurden, gibt das Programm neben den Schätzern noch asymptotische Standardfehler, t-Werte, empirische Wahrscheinlichkeiten und Konfidenzintervalle, Parameterkovarianzen und -korrelationen aus. Zur Überprüfung und Diagnose können einige Grafiken erzeugt werden (Residuenhistogramme, -autokorrelationen⁶¹). Informationskriterien und Bestimmtheitsmaße stehen nicht zur Verfügung.⁶²

⁶⁰ Generell unterscheiden sich die Schätzverfahren nur wenn die Zeitreihen sehr kurz sind und/oder die Parameter nahe an der Stationaritäts- bzw. Invertibilitätsgrenze liegen. Vgl. dazu die Ausführungen in THOME, *Univariate Box/Jenkins-Modelle*, Abschnitt 8. HARVEY geht sogar noch weiter und schreibt, daß nach herrschender Meinung die Verfahren schon bei Zeitreihen der Länge $N=100$ kaum noch Unterschiede zeigen. Vgl. HARVEY, A. C., *Time series models*, New York usw. 1993, S. 68-73. Daß *STATISTICA* ursprünglich nicht für die Gesellschaftswissenschaften, sondern für die Naturwissenschaften konzipiert wurde, wird z.B. an der entsprechenden Passage des Hilfe-Textes, der die Methoden vergleicht, besonders deutlich. Danach sollte das approximative Verfahren als schnellstes gewählt werden, wenn die Zeitreihen mehr als 30.000 (!) Werte umfassen, oder wenn bei saisonalen Modellen das Lag mehr als 365 (!) Werte umfaßt.

⁶¹ Autokorrelationsfunktion sowie partielle AKF stehen natürlich auch für die Originalreihe zur Verfügung. Beide werden als hochauflösende Grafik erstellt und können genauso wie alle anderen Grafiken editiert werden. Neben den Werten der Autokorrelation werden deren Standardfehler, die Box-Ljung-Statistik sowie die empirischen Wahrscheinlichkeiten ausgegeben. Die Standardfehler unter der Annahme von "white noise" werden nach der gleichen Formel wie in *SPSS* berechnet.

⁶² Nicht ganz verständlich ist eine Fehlermeldung, die bei der Schätzung einer simulierten MA(2)-Reihe mit den Parametern $\theta_1=1.5$, $\theta_2=-0.8$: "*Some parameter estimates are very close to or greater than 1.0; model may be non-stationary or non-invertible; check model-adequacy.*"

Im Frequenzbereich entsprechen die Möglichkeiten größtenteils *SPSS* oder *SYST AT*. Periodogramme und Spektren werden darüber hinaus - ebenso wie bei den anderen Prozeduren - als hochauflösende, editierbare Grafik dargestellt und zusätzlich als Scrollsheet tabelliert.

MicroTSP

Das Programm *MicroTSP* stammt aus dem Bereich der Ökonometrie und wurde speziell zur (ökonometrischen) Analyse von Zeitreihen für das Betriebssystem DOS entwickelt.⁶³

Als Dokumentation wird ein kurze, aber völlig ausreichende Installationsanleitung sowie ein etwa 400 Seiten starkes spiralgebundenes DIN A5-Handbuch mitgeliefert. Nach drei einführenden Kapiteln über grundlegende Funktionen des Programms werden einzelne Analyseverfahren anhand von Fallstudien erläutert. Die Beispieldateien werden auf Diskette mitgeliefert. In einem dritten Teil werden schließlich alle Prozeduren und Befehle systematisch beschrieben.

Das Programm verfügt sowohl über eine Kommandosprache als auch über ein Menüsystem, das über Funktionstasten aufzurufen ist. Will man die Befehle nicht direkt eingeben, können sie aus diesen Menüs in die Kommandozeile geschrieben werden, von der aus die Befehle direkt ausgeführt werden. Falls ein Befehl nicht vollständig ist, fragt *MicroTSP* den Anwender nach den fehlenden Angaben. Alternativ dazu kann man sich auch Kommandodateien erstellen, die im Batch-Betrieb abgearbeitet werden können. Innerhalb solcher Befehlsdateien bietet *MicroTSP* auch die Möglichkeit, über FOR..NEXT-Schleifen unter Angabe von Variablen komplexere Abfragen zu schreiben.⁶⁴ Wahlweise kann das Ergebnis einzelner Analysen oder der gesamten Sitzung in einer ASCII-Datei mitprotokolliert werden.

Das Datenmanagement von *MicroTSP* ist etwas eigenwillig. Falls während einer Arbeitssitzung neue Daten eingegeben werden sollen, kann mit dem Befehl `create` eine Zeitindizierung durch Angabe des Startzeitpunktes und der Periode erzeugt werden. Möglich sind undatierte Werte, Jahre, Quartale oder Monate. Anschließend können Daten mit einem (rudimentären) Editor eingegeben werden. Im Gegensatz zu den bislang besprochenen Programmen verfügt *MicroTSP* über ein spezifisches Zeitreihenformat. Alle Zeitreihenwerte sind also unmittelbar mit einem Zeitpunkt verknüpft.

Die Daten lassen sich auf zweierlei Weise permanent sichern. Zum einen besteht die Möglichkeit, einzelne Zeitreihen in einer separaten Datei im

⁶³ Vgl. zum folgenden auch HIGGINS, M. L. / LTAIFA, N., *MicroTSP Version 7.0: a review*, in: *Journal of Applied Econometrics* 7 (1992), S. 423-429.

⁶⁴ Matrizen bzw. Arrays können allerdings nicht definiert werden.

ASCII-Format abzuspeichern. Diese Datei kann in den ersten Zeilen erläuternde Kommentare enthalten. Werden diese Dateien bei einer neuen Sitzung wieder geladen,⁶⁵ muß vorher ein entsprechend großer Teil des Arbeitsspeichers mit dem Befehl `create` und der Angabe einer Zeitspanne für die Daten reserviert werden.

```

range 1788 - 1988 | series: current=2 maximum=198 | output LPT1:
-----
URR1      URR2
current SMP1 1788 - 1988 | path D:\TSP\DATA\      print PUFF
-----
>load beispiel

-----
Single Equation Estimation
-----
(1) Ordinary L.S. & ARMA      LS
(2) OLS (White Covariance)   LS(H)
(3) Weighted Least Squares   LS(W)
(4) Two-Stage L.S. & ARMA    TSLS
(5) Weighted TSLS            TSLS(W)
(6) Nonlinear Least Squares  NLS
(7) Weighted NLS             NLS(W)
(8) Starting values for NLS  PARAM
(9) Logit -binary dependent  LOGIT
(A) Probit -binary dependent  PROBIT
-----
System Estimation
-----
(B) System (file) estimation  SYS
(C) Vector Autoregression     VARREST
F1 Break (F3-F10 menus)

-----
1-BREAK 2-LAST| 3-FILES 4-DATA 5-GRAPH 6-STATS 7-EQUATIONS 8-ESTIMATE 9-CONTROL

```

Abbildung 7: Bildschirmaufbau von MicroTSP

Der freibleibende Speicher wird vom Programm genutzt.⁶⁶ Als Alternative kann man mit einem Befehl `save` alle sich in der Arbeitsdatei befindenden Reihen in einer Binärdatei abspeichern. Die Datendefinition dieser Reihen wird in einer separaten ASCII-Datei abgespeichert. Importmöglichkeiten bestehen für Lotus-, Symphony- und DIF- (*Data Interchange Format*) Dateien.

Zur Datentransformation steht eine Reihe von Funktionen zur Verfügung, die im großen und ganzen denen der oben besprochenen Programme entspricht. Sehr vorteilhaft ist, daß für eine Reihe von Datentransformationen wie z.B. der Bildung von *lags* oder Differenzen von Zeitreihen diese nicht explizit als neue Variablen erzeugt werden müssen, sondern sie einfach durch Angabe des *lags* in Klammern hinter einem Variablennamen bzw. durch Angabe der Funktion in eine Analyse miteinbezogen werden können. Wenn ein Modell geschätzt wird, bleibt dieses solange im Arbeitsspeicher erhalten, bis es durch ein neues ersetzt wird. Alle folgenden Prozeduren

⁶⁵ Man kann mehrere Dateien (1 Datei = 1 Zeitreihe) in die Arbeitsumgebung laden.

⁶⁶ Das bedeutet, daß die Größe der zu verarbeitenden Dateien durch den Arbeitsspeicher begrenzt ist. Extended memory wird von *MicroTSP* nicht unterstützt.

können dann auf dieses Modell Bezug nehmen.⁶⁷ Ungewöhnlich, aber durchaus nützlich, ist die Möglichkeit, aggregierte Werte auf eine niedrigere Aggregationsebene auszuweiten. So kann z.B. aus einer Zeitreihe mit 20 Jahreswerten eine Reihe mit $12 \cdot 20 = 240$ Monatswerten erzeugt werden, wobei für die einzelnen Monatswerte entweder die Jahreswerte übernommen oder diese vorher durch 12 geteilt werden.⁶⁸ In einer Arbeitssitzung können allerdings nur Zeitreihen desselben Aggregationsniveaus geladen werden. Erwähnenswert ist schließlich die Möglichkeit, mit einem Befehl `sample` eine bestimmte Zeitperiode auszuwählen, die so lange verwendet wird, bis mit einem weiteren `sample`-Befehl eine neue (Unter-) Periode definiert wird.

Die Grafikmöglichkeiten sind recht umfangreich. Über Kommandos können hochauflösend Linien-, Balken-, Kreis- und Streudiagramme erzeugt werden, in denen sich - ebenfalls über Kommandos - Markierungssymbole, Skalierungen, Beschriftungen, Linienstärken und Farben verändern lassen.⁶⁹ Die erzeugte Grafiken können auch separat als Datei abgespeichert werden, allerdings nur in einem internen Format oder - über Umwege - als HPGL-Datei.

Hinsichtlich der statistischen Verfahren ist die Herkunft des Programmes aus der Ökonometrie bis hin zur Syntax klar zu erkennen: so gibt es etwa keinen expliziten Befehl für die ARIMA-Modellierung, sondern diese wird über einen allgemeinen Regressionsbefehl `ls` (*least-squares*) ausgeführt, in dem der Name der Zeitreihe als abhängige Variable und die AR- und MA-Terme als "Prädiktoren" formuliert werden.⁷⁰ "AR" und "MA" sind dabei vom Programm reservierte Variablennamen. Ebenso kann eine Konstante einbezogen werden, indem man ein "c" in die Schätzgleichung mit aufnimmt. Die Anforderung einer ARIMA(2,0,1)-Schätzung mit der Modellgleichung

⁶⁷ So müssen z.B. die Residuen eines Modells nicht explizit abgespeichert werden, sondern sind als spezifische Systemvariable mit dem Namen "resid" vorhanden. Wird ein neues Modell geschätzt, werden die Werte dieser Variablen mit den neuen Residuen überschrieben. Analog nehmen die verschiedenen Testprozeduren immer auf die Parameter des zuletzt geschätzten Modells Bezug.

⁶⁸ Das ist z.B. ganz nützlich, wenn man in einer Grafik sowohl Monats- als auch Jahreswerte darstellen möchte.

⁶⁹ Histogramme von Originalreihen oder Residuen werden nur mit ASCII-Zeichen im Textmodus dargestellt.

⁷⁰ Da das Programm vorrangig zur Schätzung von einer Vielzahl verschiedener Regressionsmodelle konzipiert wurde, scheint die Aufnahme der Möglichkeit einer ARIMA-Modellierung im Programm wohl von den Autoren auch nicht als eigenständiges Verfahren geplant worden zu sein, sondern wohl eher, um im Rahmen des Regressionsmodells eine kompliziertere Fehlerstruktur berücksichtigen zu können.

$$Y_t = \mu + \phi_1 Y_{t-1} - \phi_2 Y_{t-2} + \epsilon_t + \theta \epsilon_{t-1}, \epsilon \sim N(0, \sigma_\epsilon^2)$$

lautet in *MicroTSP* z.B. für eine Variable Y mit dem Namen "Var1":⁷¹

```
ls Var1 c ar(1) ar(2) ma(1)
```

und für eine saisonale Zeitreihe "Var2" z.B.

```
ls Var2 c ar(1) ar(2) sar(4)
```

Sind in der Gleichung MA-Terme enthalten, verwendet das Programm das nichtlineare Kleinst-Quadrate-Schätzverfahren nach **MARQUARDT**, bei dem wahlweise Startwerte vorgegeben oder vom Programm bestimmt werden können, ebenso wie Angaben zu Iterationen, Konvergenz und Schrittabständen. Eine backcasting-Option gibt es nicht. Für ein berechnetes Modell werden neben den Werten der Koeffizienten die Standardfehler, t-Statistiken und die empirischen Wahrscheinlichkeiten ausgegeben. Der Sinn einer *generellen* Angabe solcher auf der t-Verteilung basierenden Wahrscheinlichkeiten muß allerdings bezweifelt werden.⁷² Bei *MicroTSP* wird dieses Problem im Zusammenhang mit einem implementierten Testverfahren besonders deutlich, wie das folgende Beispiel zeigt: Das Programm bietet neben zahlreichen anderen Tests auch einen *unit root* Test. Dieser Test untersucht in seiner einfachsten Form, ob in einer Zeitreihe, die sich als AR(1)-Prozeß

$$Y_t = m + \rho Y_{t-1} + e_t$$

schreiben läßt, ob der Koeffizient $\rho = 1$ ist. Zur Schätzung des Koeffizienten wird zuvor auf beiden Seiten der Gleichung Y_{t-1} abgezogen und in der Gleichung

$$\Delta Y_t = m + \gamma Y_{t-1} + e_t$$

⁷¹ Wie man an der Gleichung sieht, gehen die MA-Koeffizienten **positiv in die Gleichung ein, so daß das Vorzeichen bei einem Vergleich mit anderen Programmen umgedreht werden muß.**

⁷² Das gilt natürlich für die anderen hier besprochenen Programme ebenso.

getestet, ob $y = p-1 = 0$ ist. Der Schätzer dieses Koeffizienten hat bekanntlich keine der herkömmlichen Stichprobenverteilungen. Ausgehend von einer Arbeit DICKEYS⁷³ wurden seither in einer unübersehbaren Fülle von Beiträgen für die verschiedensten Situationen (d.h. Erweiterungen des einfachen AR(1)-Prozesses) Verteilungen - meist empirisch auf der Basis von Monte-Carlo-Simulationen - ermittelt.⁷⁴

Im folgenden haben wir einen einfachen Random Walk simuliert (N=200) und anschließend mit *MicroTSP* auf die Existenz einer Einheitswurzel $H_0: p = 1$ d.h. $H_0: y = 0$ getestet. Darauf liefert das Programm folgendes Ergebnis,

```

Augmented Dickey-Fuller: UROOT(N,1) X
-----
Dickey Fuller t-statistic      0.0449
MacKinnon critical values: 1%  -2.5753
                             5%  -1.9413
                             10% -1.6165
-----
    
```

Abbildung 8: Unit root-Test mit *MicroTSP*

nach dem erwartungsgemäß die Nullhypothese einer Einheitswurzel nicht zurückgewiesen werden kann, da der Wert der Dickey-Fuller-f-Statistik noch größer ist als der kritische 10%-Wert, der hier in einer Version von MACKINNON angegeben wird.⁷⁵

Nach Ausgabe dieser Testgrößen fragt *MicroTSP*, ob die entsprechende Testgleichung ebenfalls ausgegeben werden soll. Beantwortet man mit "y", erhält man folgende Ausgabe,⁷⁶

⁷³ DICKEY, D. A., *Estimation and hypothesis testing for nonstationary time series*, Ph. D. dissertation, Iowa State University 1976.

⁷⁴ Vgl. allgemein etwa HASSLER, U., Einheitswurzeltests - Ein Überblick, in: *Allgemeines Statistisches Archiv* 78 (1994), S. 207-228.

⁷⁵ MACKINNONs kritische Werte basieren auf umfangreicheren Simulationsstudien, die für beliebige Stichprobengrößen und die Alternativen reiner AR(p)-Prozeß, AR(p)-Prozeß mit Konstante sowie AR(p)-Prozeß mit Konstante und linearem Trend kritische Werte bereitstellen. Vgl. HALL / JOHNSTON / LILIE, S. 16-6.

⁷⁶ Es fällt auf, daß das Bestimmtheitsmaß R^2 negativ ist. Obwohl in der Schätzgleichung keine Konstante enthalten ist, wird von *MicroTSP* anscheinend zwar die Summe der Residuenquadrate anhand einer Regression ohne Konstante schätzt, die Varianz von X dagegen als Summe der quadrierten Abweichungen vom empirischen Mittelwert und nicht als einfache Summe der quadrierten X -Werte. Leider fehlt in der Dokumentation ein Hinweis auf die Schätzprozedur. Vgl. zum Problem aber etwa MADDALA, G. S., *Introduction to econometrics*, New York 1992, S. 83.

```

LS // Dependent Variable is D(X)
Augmented Dickey-Fuller: UROOT(N,1) X
=====
VARIABLE      COEFFICIENT   STD. ERROR   T-STAT.    2-TAIL SIG.
-----
D(X(-1))      0.0611150    0.0715016   0.8547354   0.3937
X(-1)         0.0092217    0.0049360   0.0449172   0.9642
=====
R-squared      -0.006072    Mean of dependent var   -0.098848
Adjusted R squared  -0.011205    S.D. of dependent var   0.996702
S.E. of regression  1.002271    Sum of squared resid    196.8913
Log likelihood    -280.3939    Durbin-Watson stat      1.993050
=====
    
```

Abbildung 9: Ausgabe der unit root Testgleichung von *MicroTSP*

Abbildung 9: Ausgabe der unit root Testgleichung von *MicroTSP*

in der die t -Statistik des gelagten Zeitreihenwertes $X(-1)$ mit 0.0449... der t -Statistik des Ausdrucks des unit root Tests entspricht. Offensichtlich macht hier jedoch die Angabe der empirischen zweiseitigen Wahrscheinlichkeit (letzte Spalte) keinen Sinn, da der Schätzer eben nicht f -verteilt ist.

Die anderen univariaten zeitreihenanalytischen Verfahren entsprechen im großen und ganzen denen der bisher besprochenen Programme.⁷⁷ Zur Identifikation der internen Struktur einer Zeitreihe bietet *MicroTSP* die Autokorrelations- und partielle Autokorrelationsfunktion an, den Standardfehler der geschätzten Autokorrelationen allerdings nur unter der Annahme des "white noise" und die Box-Pierce sowie Ljung-Box-Statistik nur für das letzte angegebene "lag".

Herkömmliche deskriptive gleitende Mittelwerte kennt *MicroTSP* überhaupt nicht, lediglich asymmetrische, in die nur vergangene Werte einbezogen werden können. Trendfunktionen müssen über die Multiple Regression modelliert werden. Zur saisonalen Adjustierung können Saisonfaktoren berechnet werden. Im Rahmen der exponentiellen Glättung schätzt *MicroTSP* fünf verschiedene Modelle über eine Fehlerquadratminimierung.⁷⁸ Fehlende Werte sind in allen hier besprochenen Schätzverfahren nicht zugelassen, sondern müssen vorher über eine bedingte Datentransformation ersetzt werden. Über Analyseverfahren im Frequenzbereich verfügt *MicroTSP* nicht.

Seit Frühjahr 1994 wird *MicroTSP* unter dem Namen *Econometric Views* mit einigen Erweiterungen auch für Windows und Mac angeboten.

⁷⁷ Dabei ist allerdings immer zu bedenken, daß *MicroTSP* vorrangig für die ökonometrische, d.h. multivariate Analyse im Rahmen des Regressionsmodells konzipiert ist.

⁷⁸ Einfaches und doppeltes (= Trendkomponente) exponentielles Glätten sowie das Holt-Winters Verfahren ohne, mit additiver oder multiplikativer Saisonkomponente.

MESOSAUR

MESOSAUR gehört zu einer Reihe von Programmen, die von *SYSTAT* als externe Zusatzprogramme für spezielle statistische Analyseverfahren angeboten werden und die völlig unabhängig von *SYSTAT* sind. Die Produktpolitik ist dabei nicht ganz klar, da *SYSTAT* selbst über ein eigenes Modul zur Zeitreihenanalyse verfügt.

Die Dokumentation besteht aus einem 248 S. umfassenden Handbuch, das im ersten Teil die einzelnen Menüpunkte beschreibt und im zweiten Teil anhand dreier "Fallstudien" einige statistische Verfahren erläutert.

Als einziges der hier vorgestellten Programme verfügt *MESOSAUR* über keine eigene Kommandosprache, sondern ist vollständig menügeführt. Das Menü orientiert sich zwar nicht am SAA-Standard, ist aber nach kurzer Einarbeitung leicht handhabbar. Die Online-Hilfefunktion beschränkt sich auf ein hypertext-verknüpftes Glossar, in dem ca. 30 statistische Begriffe erläutert werden. Die kontextsensitive Hilfe zu den einzelnen Befehlen beschränkt sich auf wenige Hinweise.

Importieren lassen sich Dateien im ASCII-, dBASE III+- und SYSTAT-Format. Darüber hinaus verfügt *MESOSAUR* über ein einzigartiges Datenmanagement. Jede Variable muß als Zeitreihe definiert werden und wird als zeitindizierter Vektor behandelt. Als Zeitintervalle sind neben ganzzahligen Nummern Minuten, Stunden, Tage, Wochen, Monate, Quartale oder Jahre möglich. Bis auf Minutenebene lassen sich Zeitreihen vom Jahre 0 bis zum Jahr 3000 indizieren. Eine Zeitreihe kann bis zu 16.000 Werte enthalten und mit bis zu 80 Zeichen Kommentar versehen werden.⁷⁹ Durch die Verwaltung einzelner Zeitreihen als Vektoren können in *MESOSAUR* im Gegensatz zu allen anderen hier vorgestellten Programmen Zeitreihen beliebiger Zeitintervalle in einer einzigen Systemdatei gespeichert werden.⁸⁰ Dabei belegt eine Zeitreihe, die 16.000 Beobachtungen umfaßt lediglich 70 Kilobyte. In einer Datei können bis zu 123 Zeitreihen abgespeichert werden.

Zur Datenverwaltung während einer Arbeitssitzung dient ein sog. Codebook.

⁷⁹ Lange Zeitreihen sind für die Wirtschaftsgeschichte durchaus nicht ungewöhnlich. So umfassen etwa die erhaltenen wöchentlichen Getreidepreis- und Umsatznotierungen der Reichsstadt Köln den Zeitraum von 1531 bis 1798 annähernd 14.000 Werte oder die erhaltenen täglichen Aktienkursnotierungen der Londoner Börse, beginnend 1698 weit über 40.000 Werte, um nur zwei Beispiele zu nennen. Vgl. EBELING, D. / IRSIGLER, F. (Bearbb.), *Getreideumsatz, Getreide- und Brotpreise in Köln 1368-1797. Erster Teil: Getreideumsatz und Getreidepreise: Wochen-, Monats- und Jahrestabelle*, Köln / Wien 1976 (= *Mitteilungen aus dem Stadtarchiv von Köln* 65); NEAL, L., *The rise of financial capitalism. International capital markets in the age of reason*, Cambridge usw. 1990.

⁸⁰ So kann in einer Datei z.B. eine Monatszeitreihe von 1700 bis 1750, eine Jahreszeitreihe von 1820 bis 1990 und ein Börsenkurs im Minutenabstand vom 5.5.1994, 10.15 Uhr bis 12.00 Uhr problemlos in derselben Datei abgespeichert werden.

In dieses Codebook können aus verschiedenen Dateien insgesamt bis zu 256 Zeitreihen eingelesen werden (siehe Abb. 10). Da alle Zeitreihen als Vektoren behandelt werden, können sie auch einzeln in eine schon bestehende Datei "hineingespeichert" bzw. aus verschiedenen Dateien in das Codebook geladen werden.⁸¹ In der alltäglichen Arbeit erweist sich dieses Datenmanagement als überaus vorteilhaft.

Die meisten Befehle lassen sich zusätzlich zur Menüführung über eine Taste bzw. eine Tastenkombination unmittelbar ausführen. So kann z.B. eine Zeitreihe grafisch dargestellt werden, indem man lediglich die Taste "p" drückt. Die Grafikmöglichkeiten sind zwar recht eingeschränkt, aber völlig ausreichend, um einen visuellen Eindruck einer Zeitreihe, eines Streudiagramms, Periodogramms, Spektrums oder Histogramms zu bekommen.⁸² Sehr nützlich für längere Zeitreihen ist, daß Bildschirmausschnitte der Grafik über eine Zoom-Funktion vergrößert werden können.

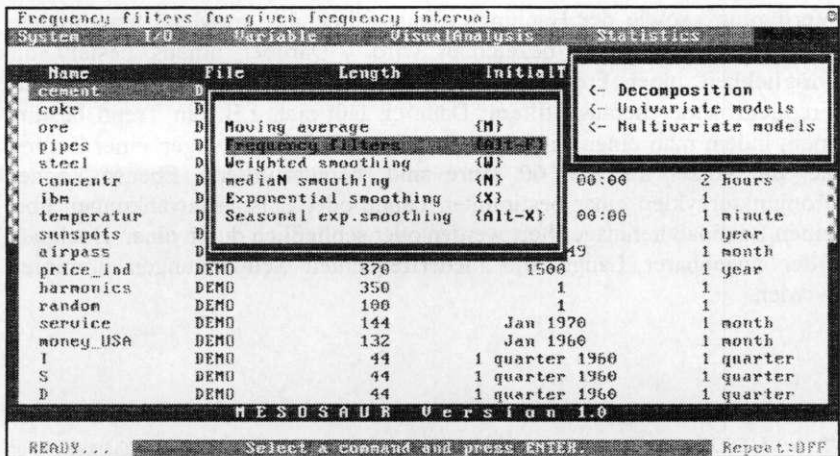


Abbildung 10: Bildschirmaufbau von MESOSAUR

Die Datentransformationsmöglichkeiten sind in MESOSAUR in drei Rubriken unterteilt: arithmetische, mathematische und "spezielle" Transformationen, die den üblichen Rahmen abdecken. Unüblich ist die Möglichkeit, eine Zeitreihe zu invertieren.⁸³ Das erweist sich z.B. als hilfreich, um

⁸¹ Es ist sogar möglich, Zeitreihen mit dem selben Namen in einer Datei mehrfach abzuspeichern.

⁸² Die Grafikfunktionen basieren auf der HALO-Bibliothek.

⁸³ Aus einer Reihe mit den Werten 1,2,3,4,5,6 wird eine Reihe mit den Werten 6,5,4,3,2,1.

im Rahmen der ARIMA-Modellierung mit dem sog. backcasting Anfangswerte zu schätzen.⁸⁴

Die univariaten statistischen Analyseverfahren beinhalten Basisstatistiken (Mittelwert etc.), Autokorrelations-, partielle Autokorrelationsfunktion, Periodogramm, integriertes Periodogramm⁸⁵ sowie das Spektrum, das als mit dem Parzen-Fenster geglättetes Periodogramm berechnet wird.⁸⁶ Letztere können als Variable abgespeichert werden. Außerdem bietet das Programm noch drei verschiedene Tests: einen Turning Point Test sowie den Kendall-schen und Spearmanschen Rangkorrelationstest. Alle Verfahren beziehen sich entweder auf den gesamten Zeitraum oder auf eine Teilperiode, die sich vorher angeben läßt.⁸⁷

Zur Glättung können gleitende Mittelwerte mit Gewichten, die denen lokaler Trendpolynome entsprechen, gewählt werden.⁸⁸ Als Alternative kann man auch exponentiell abnehmende Gewichte bestimmen. Zur robusten Glättung stehen gleitende Mediane mit einer Spannweite von 3 bis 21 zur Verfügung⁸⁹ sowie der Hanning-cosinus Filter, der im Menü irrtümlicherweise als "Hamming" bezeichnet wird.⁹⁰ Darüber hinaus besteht die Möglichkeit, über Frequenzfilter aus einer Zeitreihe genau angebbare Frequenzbänder herauszufiltern. Dadurch läßt sich z.B. ein Trend bestimmen, indem man einen Tiefpaß wählt, der alle Schwingungen einer Zeitreihe, die kürzer als z.B. 60 Jahre sind, herauszufiltern. Ebenso können Konjunkturzyklen einer bestimmten Länge oder Saisonschwankungen über einen Bandpaß herausgefiltert werden oder schließlich durch einen Hochpaßfilter angebbarer Länge alle niederfrequenten Schwingungen eliminiert werden.

⁸⁴ In *MESOSAUR* selbst können allerdings keine Startwerte für die Schätzung von ARIMA-Modellen angegeben werden.

⁸⁵ Das integrierte Periodogramm dient zum Test einer Zeitreihe auf "white noise". Zu diesem Zweck sind in die Darstellung zusätzlich Konfidenzgrenzen eingezeichnet. SCHLITZGEN / STREITBERG, S. 279, betonen, daß dieser Test lediglich auf Originalreihen, nicht auf geschätzte Residualreihen angewendet werden kann.

⁸⁶ Die Korrelationsfunktionen und Periodogramme sind auch grafisch hochauflösend darstellbar.

⁸⁷ Die Auswahl einer Teilperiode ist insbesondere bei langen Zeitreihen ein wenig umständlich, da die Anfangs- und Endwerte Schritt für Schritt in einem Menü über die Plus- und Minustasten erhöht bzw. verringert werden müßten.

⁸⁸ Die Gewichte des linearen lokalen Trendpolynoms zu einem Zeitpunkt / entsprechen gerade dem einfachen gleitenden Mittelwert.

⁸⁹ Davon können auch mehrere gleichzeitig gewählt werden.

⁹⁰ Im Menü "Visual Analysis" besteht die Möglichkeit der interaktiven Bestimmung gleitender Mediane.

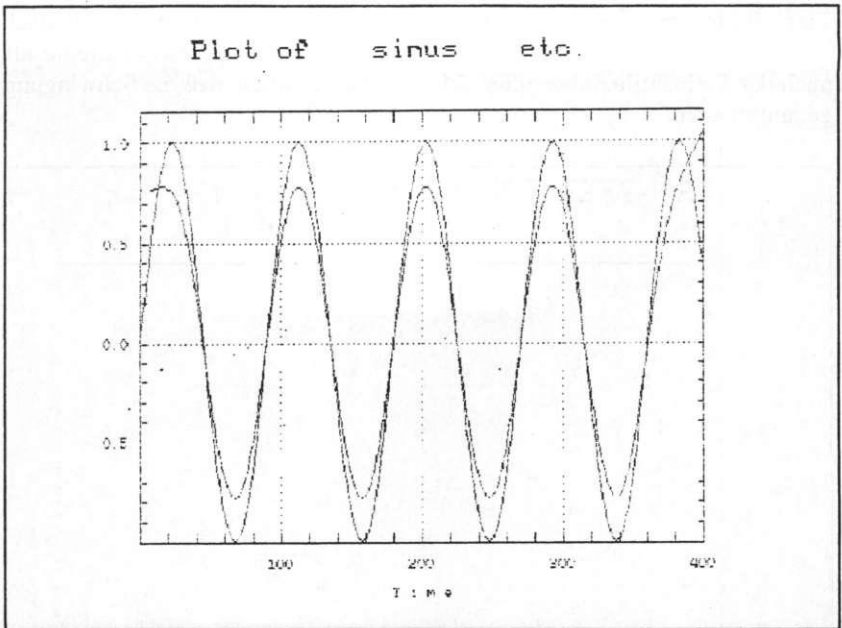


Abbildung 11: Sinusschwingung und von MESOSAUR berechneter Tiefpaßfilter ($\lambda=1/65$)

Dieses Verfahren, das insbesondere für die Konjunkturgeschichte von herausragender Bedeutung ist,⁹¹ wird in der Dokumentation leider nicht näher erläutert. Die Theorie, die hinter dem Verfahren steht sowie die Berechnung sind außerordentlich kompliziert, so daß die Implementation in einem kommerziellen PC-Programm sehr überraschend ist. Bislang ist uns nur ein nichtkommerzielles Programm bekannt, das diese Berechnung durchführt.⁹² Da die deskriptive Trendbestimmung in vielen historischen Zusammenhängen von besonderem Interesse ist, bringen wir ein Beispiel zur Untersuchung der Trennschärfe des Frequenzfilters im niederfrequenten Bereich. Zu diesem Zweck wurde eine deterministische Sinusschwingung mit einer Periode von exakt 90 Zeiteinheiten erzeugt, deren Periodogramm also die gesamte Dichte bei $\lambda=1/90$ konzentriert. Für diese Reihe wurde ein

⁹¹ Vgl. dazu etwa METZ, R., *Ansätze, Begriffe und Verfahren der Analyse ökonomischer Zeitreihen*, in: *Historical Social Research / Historische Sozialforschung* 13/3 (1988), S. 21-103.

⁹² METZ, R. / STIER, W., *NULFIL. Ein Programm zum Filtern im Frequenzbereich mit Spektralminimierung*, in: Spree, R., *Lange Wellen wirtschaftlicher Entwicklung in der Neuzeit. Historische Befunde, Erklärungen und Untersuchungsmethoden*, Köln 1991 (= *Historical Social Research / Historische Sozialforschung, Supplement 4*), S. 139-144.

Tiefpaßfilter mit $\lambda=1/65$ berechnet, der diese Schwingung theoretisch vollständig übertragen müßte. In Abb. 11 ist sowohl die Originalreihe als auch der Tiefpaßfilter abgebildet. Man erkennt deutlich, daß die Schwingung gedämpft wird.⁹³

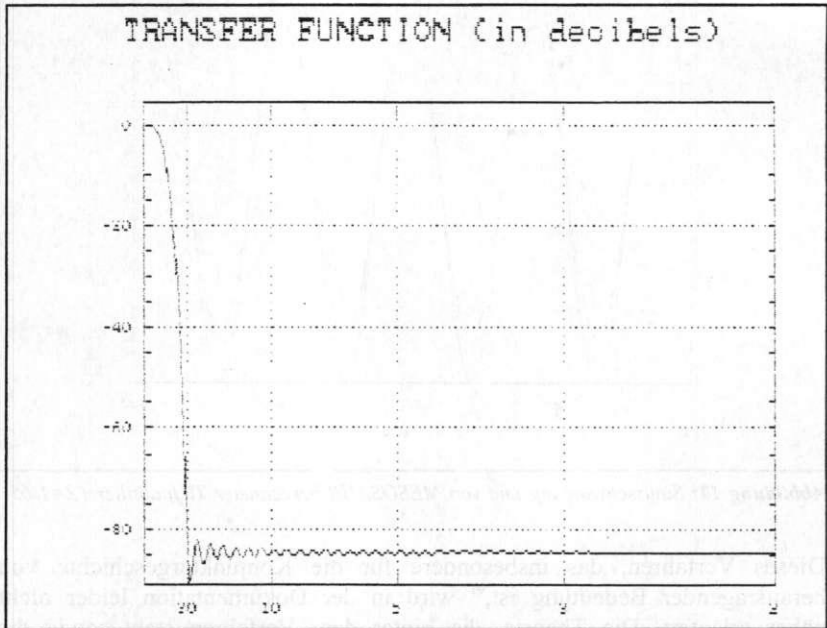


Abbildung 12: Von MESOSAUR verwendete Transferfunktion eines Tiefpaßfilters ($\lambda=1/65$)

Neben der Berechnung eines Frequenzfilters bietet MESOSAUR auch eine grafische Darstellung der Transferfunktion des gewählten Filters. Die Transferfunktion des von uns angewendeten Tiefpaßfilters ist in Abb. 12 dargestellt.⁹⁴ Man sieht, daß die Transferfunktion nicht "rechtwinklig" verläuft, sondern Frequenzen, die höher als $1/65$ sind, nicht vollständig eliminiert werden und niedrigere Frequenzen nur teilweise übertragen werden.

⁹³ Außerdem fällt auf, daß der **verwendete** Filter nicht "randstabil" ist. Auf diese Problematik gehen wir hier aber nicht näher ein. Vgl. zum gesamten Komplex METZ, R. / STIER, W., Filter design in the frequency domain, in: KLEINKNECHT, A. / MANDEL, E. / WALLERSTEIN, I., *New findings in long-wave research*, London 1992, S. 45-79.

⁹⁴ Abb. 11 und 12 lassen im übrigen erkennen, daß die Grafikausgabe von MESOSAUR eher mäßig ist.

Die hier geschilderten Probleme treten allerdings nur im niederfrequenten Bereich auf. Zur Ermittlung kürzerer Zyklen oder saisonaler Schwankungen liefert der Filter durchaus zufriedenstellende Resultate.

Schließlich gehen wir noch kurz auf die Robustheit des Frequenzfilters ein. Die folgende Abbildung zeigt für die Reihe der Hauskäufe in Lübeck von 1284 bis 1700⁹⁵ einen Tiefpaßfilter mit einer Frequenz von 1/65, d.h. alle Schwingungen mit einer Dauer von weniger als 65 Jahren werden eliminiert. Im Vergleich dazu wurde ein von *SYSTAT* Series erzeugter LOWESS-Glätter in die Grafik aufgenommen. Man erkennt eine sehr hohe Übereinstimmung in weiten Bereichen der Reihe. Eine auffällige Abweichung zeigt sich um 1440: hier sind die Werte des Frequenzfilters höher.

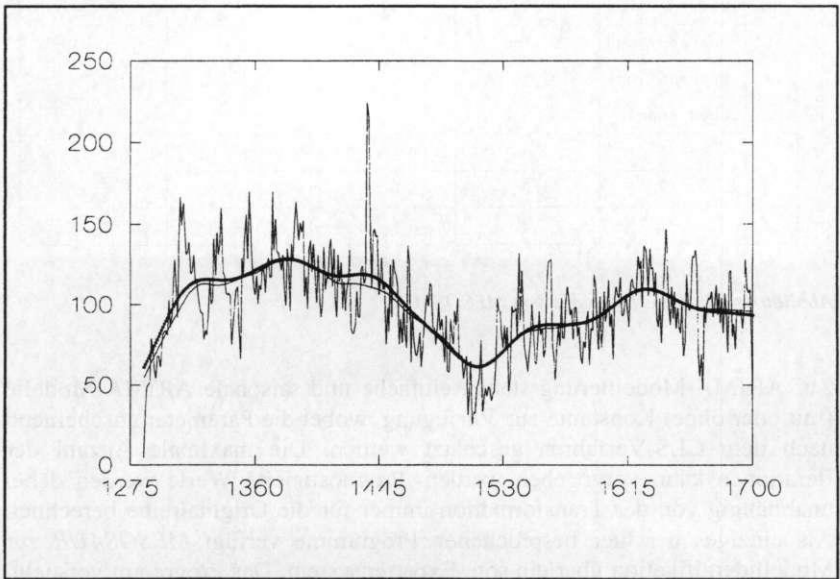


Abbildung 13: Vergleich eines Mesosaur-Tiefpaßfilters ($\lambda=1/65$) mit einem LOWESS-Glätter von *SYSTAT* Series (Tension-Parameter=0.9)

⁹⁵ Vgl. allgemein zum Hintergrund HAMMEL, R., Häusermarkt und wirtschaftliche Wechsellagen in Lübeck von 1284 bis 1700, in: *Hansische Geschichtsblätter* 106 (1988), S. 41-107 sowie zu verschiedenen deskriptiven Auswertungsverfahren RAHLF, T., Ausgewählte Methoden der konjunkturstatistischen Untersuchung am Beispiel des Lübecker Häusermarktes von 1284 bis 1700, in: HAMMEL-KIESOW, Rolf (Hg.), *Wege zur Erforschung städtischer Häuser und Höfe. Beiträge zur fächerübergreifenden Zusammenarbeit am Beispiel Lübecks im Spätmittelalter und in der frühen Neuzeit*, Neumünster 1993 (= Häuser und Höfe in Lübeck 1), S. 374-380.

Die Erklärung dafür liefern die "Extremwerte" 1437-1439. Hierdurch ergeben sich für den Tiefpaßfilter höhere Werte als für den LOWESS-Glätter.⁶⁶ Da der LOWESS-Glätter über den Median ermittelt wird, beeinflussen Extremwerte den Verlauf nicht.

ARIMA MODEL FITTING FOR VARIABLE *cesent*
Moving Average order

M o v i n g A v e r a g e o r d e r		0	1	2	3	4
0		Unrecognized	Unrecognized	Overdefined		
1	Underdefined	Failed	Failed			
2	Underdefined	Failed	Failed			
3	Underdefined	Failed				
4	Underdefined					
5	Good model					
6						
7						
8						

Press INS for EXPERT menu

Abbildung 14: Expertenmodus von MESOSAUR

Zur ARIMA-Modellierung stehen einfache und saisonale ARIMA-Modelle (mit oder ohne) Konstante zur Verfügung, wobei die Parameter anscheinend nach dem CLS-Verfahren geschätzt werden. Die maximale Anzahl der Iterationen kann vorgegeben werden. Prognostizierte Werte werden dabei unabhängig von der Transformation immer für die Originalreihe berechnet. Als einziges der hier besprochenen Programme verfügt *MESOSAUR* zur Modellidentifikation über ein sog. Expertensystem. Das Programm versucht, für eine Originalreihe zuerst den optimalen Grad der Differenzenbildung zu ermitteln. Dann werden nacheinander ARMA-Modelle der Ordnung (0,0) bis höchstens zur Ordnung (8,4) getestet und hinsichtlich verschiedener Kriterien (t-Werte, Residuenautokorrelationen, Akaike-Kriterium) beurteilt (siehe Abb. 14). Wird ein Modell als gut befunden, können anschließend die Parameter für dieses oder jedes andere Modell geschätzt werden. Zusätzlich

⁶⁶ Vgl. zur Berechnungsweise CLEVELAND, W. S., Robust locally weighted regression and smoothing scatterplots, in: *Journal of the American Statistical Association* 74 (1979), S. 829-836 sowie RAHLF, Statistische Graphik. Der Tension-Parameter gibt an, wieviel Prozent der Daten in die Berechnung eines Wertes einbezogen werden. 0.9 bedeutet, daß jeweils 90% einbezogen werden.

wird für jedes Modell kurz erläutert, warum *MESOSAUR* zu einer Ablehnung bzw. Befürwortung eines Modells rät. Wird kein Modell als adäquat beurteilt, erscheint eine Meldung: "*Complicated data structure. Call for a knowledgeable specialist*". Nach welchen Kriterien genau *MESOSAUR* von Modell zu Modell vorgeht und welche Kombination von Kriterien letztendlich zur Ablehnung oder Annahme eines Modells führt, ist leider unklar. Allgemein zeigt unsere Erfahrung, daß selbst bei Simulationen von Prozessen einfacher Modelle, insbesondere von MA(q)-Modellen, die meisten Modelle nicht erkannt werden.⁹⁷

Fazit

Die hier vorgestellten Programme können natürlich keine repräsentative Auswahl der gesamten zur Zeit angebotenen Software zur univariaten Zeitreihenanalyse sein.

Hinsichtlich der univariaten Verfahren weisen die hier beschriebenen Programme einige Gemeinsamkeiten auf. Alle Programme bieten die Standardanalyseverfahren der ARIMA-Modellierung sowie "traditionelle" Methoden, wie die Bestimmung von Trendfunktionen, gleitenden Mittelwerten, Saisonadjustierung und exponentieller Glättung mit mehr oder weniger umfangreichen Möglichkeiten. Überblickt man die neueren Lehrbücher zur Zeitreihenanalyse, so fällt auf, daß mit Ausnahme der ARIMA-Modellierung diese Verfahren kaum noch behandelt werden, letztere dafür aber meist in einem Umfang, der in keinem der hier besprochenen Programme implementiert ist. Diese Diskrepanz hat unseres Erachtens zwei Gründe: zum einen benötigt die Umsetzung statistischer Verfahren in ein kommerzielles Programmsystem, das stabil und benutzerfreundlich sein soll, sicher einige Zeit. Zum anderen ist zwar die ARIMA-Modellbildung mit Methoden, die über die ursprünglichen von Box/Jenkins entwickelten hinausgehen, in der Forschung mittlerweile zwar recht etabliert; die kommerziellen Programmanbieter richten sich aber natürlich nicht nur nach den Interessen von - meist nicht sehr zahlungskräftigen - Forschern, sondern nach den Erfordernissen der "Praktiker", d.h. z.B. Wirtschaft und Industrie.

In Hinblick auf die ARIMA-Modellierung wäre die Implementation einiger "neuerer" Verfahren wünschenswert; hierbei ist vor allem an Identifikations-

⁹⁷ Werden dennoch MA-Modelle im ARIMA-Menü angegeben und geschätzt, so ergeben sich oft geschätzte Standardfehler, die erheblich von den Angaben der anderen hier besprochenen Programme abweichen. Die t-Werte sind in der Regel vom Betrag erheblich kleiner.

hilfen zur Bestimmung der Modellordnung zu denken,⁹⁸ Methoden für die Erkennung und Schätzung von Ausreißern,⁹⁹ robuste Verfahren zur Parameterschätzung von ARIMA-Modellen,¹⁰⁰ oder sog. "robuste" Glättungsverfahren.¹⁰¹ Wie man bei *MicroTSP* sieht, werden einige neuere Verfahren, die von der Ökonometrie ausgehen, bereits umgesetzt. Die eigenwillige Datenverwaltung des Programms sowie das völlige Fehlen frequenzanalytischer Verfahren zwingt jedoch dazu, zusätzlich noch andere Programme zu benutzen.

Ein insbesondere für Historiker bedeutsames Manko ist bei allen Programmen die recht rudimentäre Möglichkeit der Datendokumentation. Aber auch hierbei gilt es zu bedenken, daß - zumindest die hier besprochenen Softwarepakete - nicht in Hinblick auf die Erfordernisse von Historikern entwickelt worden sind. Konzepte wie der unmittelbare Zugriff auf Oracle- oder SQL-Datenbanken von *SPSS für Windows* oder die Megafilos von *STA TISTICA/w* sind aber sehr vielversprechend und sollten in der historischen Forschung mehr beachtet werden. Ist man primär an der Analyse einzelner Zeitreihen interessiert und arbeitet nicht mit komplexen Datenbanken, stellt die Datenverwaltung von *MESOSAUR* eine nahezu optimale Lösung dar. Auch die einfache Bedienung, das relativ breite Methodenspektrum sowie die Implementierung von Frequenzfiltern dürfte dieses Programm für Historiker interessant machen. Für Standardanalysen, Datenverwaltung von wenigen, unterschiedlichen Zeitreihen sowie schnelle, explorative Analysen ist dieses Programm unbedingt zu empfehlen, wenn auch die Prozedur zur ARIMA-Modellierung einige Schwächen aufweist.

Die drei hier besprochenen allgemeinen Statistik-Pakete bieten unter anderem den Vorteil, daß eine große Breite von Datentransformationsmöglichkeiten und Analysen vorgenommen werden kann, die nicht unmittelbar für die Zeitreihenanalyse vorgesehen ist. Auf der anderen Seite beschränken sich natürlich die einzelnen zeitreihenanalytischen Verfahren auf einen sehr allgemeinen Rahmen. Erstaunlich ist, daß die jeweiligen Windows-Versionen von *SPSS* und *SYSTA T* nicht die Möglichkeiten nutzen, die dieser Betriebs-

⁹⁸ Hierbei wäre vor allem an die inverse und extendierte Autokorrelationsfunktion, die Corner-Methode oder die Vektorkorrelationen zu denken. Vgl. dazu SCHMITZ, B., *Einführung in die Zeitreihenanalyse. Modelle, Softwarebeschreibung, Anwendungen*, Bern / Stuttgart / Toronto 1989, S. 77ff.

⁹⁹ Vgl. etwa THOME, H., Ausreißer und fehlende Werte in der Zeitreihenanalyse: Ihre Modellierung im Rahmen des Box/Jenkins-Ansatzes, in: *ZA-Information* 31 (1992), S. 37-69.

¹⁰⁰ Vgl. dazu etwa SCHLITZGEN, R., Ein moderner Zugang zur Analyse von Zeitreihen mit Ausreißern, in: *Allgemeines Statistisches Archiv* 75 (1991), S. 75-102.

¹⁰¹ Vgl. allgemein zu Glättungen SCHLITZGEN, R., Robuste Glättung von Zeitreihen, in: *Allgemeines Statistisches Archiv* 74 (1990), S. 223-250.

System-Aufsatz bietet. Lediglich bei *STATISTICA/w* scheint dies konsequent gelungen zu sein.

Wie bereits erwähnt, kann der hier dargelegte Überblick über die fünf beschriebenen Programme nur einen kleinen Einblick in die große Zahl der zur Zeit verfügbaren Software zur Zeitreihenanalyse geben. Er soll daher nicht abschließen, ohne noch kurz einige andere Programme zu erwähnen, die zu den verschiedenen Verfahren angeboten werden (natürlich ohne Anspruch auf Vollständigkeit).

In Hinblick auf die univariaten Zeitreihenanalyseverfahren ist *BMDP* sicher mit *SPSS* oder *SYSTAT* vergleichbar. Das Programm stammt ursprünglich aus dem medizinischen Bereich, findet aber auch in den Gesellschaftswissenschaften Verwendung.

Das umfangreichste allgemeine Statistik-Programmpaket ist sicher *SAS*, das auch über mehrere Module zur Zeitreihenanalyse verfügt. Im Gegensatz zu allen anderen uns bekannten Statistikprogrammen kann *SAS* nicht käuflich erworben, sondern nur "gemietet" werden. Beide Programme wurden ursprünglich für Großrechner-Betriebssysteme entwickelt, werden seit einigen Jahren aber auch in PC-Versionen angeboten.

Für ökonometrische Verfahren, die auch limitierte abhängige Variablen explizit berücksichtigen hat das Programm *LIMDEP* in letzter Zeit starke Verbreitung gefunden.¹⁰² Im Bereich der Ökonometrie werden neben *MicroTSP* *MICROFIT*,¹⁰³ *SHAZAM*,¹⁰⁴ *PC-GIVE**¹⁰⁵ und *RATS*¹⁰⁶ verwendet. Für die von A. C. **HARVEY** beschriebenen strukturellen Zeitreihenmodelle¹⁰⁷

¹⁰² LIMDEP bietet zwar für speziell für die Zeitreihenanalyse vergleichsweise wenig Prozeduren, eröffnet aber durch die sehr mächtige Programmiersprache eine Reihe von Möglichkeiten. So ist etwa unter den mitgelieferten Beispielprogrammen ein Programm zur Schätzung von ARCH-Modellen enthalten. Vgl. allgemein zu den Komponenten des Programms PAARSCH, H. J., LIMDEP, version 6.0: a review, in: *Journal of Applied Econometrics* 9 (1994), S. 91-98; ROSENQVIST, G., Software review: LIMDEP version 6.0, in: *Applied Statistics* 43 (1994), S. 290-293; SENSCH, J., Statistical computing software reviews: LIMDEP 6.0 (1992), in: *Historical Social Research / Historische Sozialforschung* 19/2 (1994), S. 111-121.

¹⁰³ MCKENZIE, C. R., MICROFIT 3.0: a review, in: *Journal of Applied Econometrics* 8 (1993), S. 413-419.

¹⁰⁴ VEALL, M. R., SHAZAM 6.2: a review, in: *Journal of Applied Econometrics* 6 (1991), S. 317-320.

¹⁰⁵ ERICSSON, N. R. / LYSS, H., An update to PC-GIVE: version 6.01, in: *Journal of Applied Econometrics* 6 (1991), S. 321-325.

¹⁰⁶ LARSEN, H., A review of RATS 3.0: software for econometric modelling, forecasting, and simulation, in: *Journal of Applied Econometrics* 5 (1990), S. 87-91.

¹⁰⁷ Vgl. dazu HARVEY, S. 253ff.

wurde das Programm *STAMP* entwickelt.¹⁰⁸ Speziell für die Modellierung von ARIMA-Modellen gibt es das Programm *Autobox* der Firma Automatic Forecasting Systems.¹⁰⁹ Hierbei handelt es sich um ein Expertensystem zur (weitgehend) automatischen Identifikation und Schätzung von ARIMA-Modellen, einschließlich verschiedenster Transformationsmöglichkeiten, der automatischen Erkennung und Schätzung von Interventionen usw. Für die Klasse der ARIMA-Modelle gehört das Programm sicher zu den elaboriertesten. Leider besteht die Politik der Firma darin, so wenig wie möglich offenzulegen, so daß die implementierten Verfahren zum großen Teil nicht nachzuvollziehen sind. Ein spezielles Zeitreihenanalyseprogramm mit einem Schwerpunkt auf der ARIMA-Modellierung ist *SCA*, daß in einer der nächsten Ausgaben dieser Zeitschrift näher besprochen werden wird.

Hinzuweisen ist schließlich noch auf spezielle statistische Programmiersprachen. Im Gegensatz zu Anwendungsprogrammen bieten diese natürlich eine viel größere Flexibilität, die freilich auch mit einem höheren Einarbeitungsaufwand verbunden ist. In den letzten Jahren sind eine Reihe von Programmiersprachen entwickelt worden, die speziell für mathematisch / statistische Probleme konzipiert sind, wie etwa *GAUSS*¹¹⁰ *SC*,¹¹¹ *S-Plus*¹¹² oder *MATHLAB*.¹¹³ Für diese Programmiersprachen existieren bereits zahlreiche Funktionsbibliotheken zu den unterschiedlichsten Spezialproble-

¹ 14

men.

¹⁰⁸ DIEBOLD, F., Structural time series and modelling package: a review, in: *Journal of Applied Econometrics* 4 (1988), S. 195-204.

¹⁰⁹ BEWLEY, R., AUTOBOX: a review, in: *Historical Social Research / Historische Sozialforschung* 16/2 (1991), S. 195-201. Mittlerweile liegt das Programm in Version 3.0 vor.

¹¹⁰ ANDERSON, R. G., The GAUSS programming system: a review, in: *Journal of Applied Econometrics* 1 (1992), S. 215-219.

¹¹¹ AUSTIN, G., SC version 1.107: a review, in: *Journal of Applied Econometrics* 9 (1994), S. 219-223.

¹¹² HALLMAN, J., Review of S-PLUS, in: *Journal of Applied Econometrics* 8 (1993), S. 213-219.

¹¹³ Einen Vergleich von *GAUSS* und *MATHLAB* bietet RUST, J., *GAUSS and MATHLAB: a comparison*, in: *Journal of Applied Econometrics* 8 (1993), S. 307-324.

¹¹⁴ So gibt es z.B. für *GAUSS* die folgenden kommerziell vertriebenen, menügeführten Anwendungsprogramme zur Zeitreihenanalyse (ohne Anspruch auf Vollständigkeit): *COINT* zur Analyse von kointegrierten Zeitreihen, *TSAGL* (Time series analysis and graphics library), eine ökonometrische Anwendung *GAUSSX*, *Time Series* speziell für Regressions- und ARIMA-Modelle oder *MULTI* zur Analyse von Vektor-ARIMA-Modellen. Alle Produkte setzen jeweils die Existenz von *GAUSS* voraus. Für *S-Plus* (und sicher auch für die anderen aufgeführten Programmiersprachen) gibt es ähnliche Anwendungen.

Literatur

- ANDERSON, R. G.**, The GAUSS programming system: a review, in: *Journal of Applied Econometrics* 7 (1992), S. 215-219.
- AUSTIN, G.**, SC version 1.107: a review, in: *Journal of Applied Econometrics* 9 (1994), S. 219-223.
- BEAUMONT, C. / MAHMOUD, E. / MCGEE, V. E.**, Microcomputer forecasting software: a survey, in: *Journal of forecasting* 4 (1985), S. 305-312.
- BEWLEY, R.**, AUTOBOX: a review, in: *Historical Social Research / Historische Sozialforschung* 16/2 (1991), S. 195-201.
- CLEVELAND, W. S.**, Robust locally weighted regression and smoothing scatterplots, in: *Journal of the American Statistical Association* 74 (1979), S. 829-836
- CLEVELAND, W. S. / TERPENNING, I. J.**, Graphical methods for seasonal adjustment, in: *Journal of the American Statistical Association* 77 (1982), S. 52-62.
- DICKEY, D. A.**, *Estimation and hypothesis testing for nonstationary time series*, Ph. D. dissertation, Iowa State University 1976.
- DIEBOLD, F.**, Structural time series and modelling package: a review, in: *Journal of Applied Econometrics* 4 (1988), S. 195-204.
- EBELING, D. / IRSIGLER, F.** (Bearbb.), *Getreideumsatz, Getreide- und Brotpreise in Köln 1368-1797. Erster Teil: Getreideumsatz und Getreidepreise: Wochen-, Monats- und Jahrestabelle*, Köln / Wien 1976 (= Mitteilungen aus dem Stadtarchiv von Köln 65)
- EHLING, M.**, Anwendungsnahe Statistiksoftware - Ein Überblick über Programmpakete für den PC, in: *Allgemeines Statistisches Archiv* 78 (1994), S. 262-269.
- ELLIOTT, A. C. / REISCH, J. S. / CAMPBELL, N. P.**, Benchmark data sets for evaluating microcomputer statistical programs, in: *Collegiate Microcomputer* 11 (1989), S. 289-299.
- ERICSSON, N. R. / LYSS, H.**, An update to PC-GIVE: version 6.01, in: *Journal of Applied Econometrics* 6 (1991), S. 321-325.
- GARDNER, E. S. Jr.**, Exponential smoothing: the state of the art, in: *Journal of Forecasting* 4 (1985), S. 1-28.
- GEBLER, J.**, *Statistische Graphik*, Basel / Boston / Berlin 1993.
- GOTTMAN, J. M.**, *Time-series analysis. A comprehensive introduction for social scientists*, Cambridge usw. 1981.
- HALL, R. E. / JOHNSTON, J. / LILIEN, D. M.**, *MicroTSP User's Manual. Version 7.0*, Irvine, California: Quantitative Micro Software 1990.

- HALLMAN, J., Review of S-PLUS, in: *Journal of Applied Econometrics* 8 (1993), S. 213-219.
- HAMMEL, R., Häusermarkt und wirtschaftliche Wechsellagen in Lübeck von 1284 bis 1700, in: *Hansische Geschichtsblätter* 106 (1988), S. 41-107.
- HÄRTUNG, J. / ELPELT, B. / KLÖSENER, K.H., *Statistik. Lehr- und Handbuch der angewandten Statistik*, München / Wien 1989.
- HARVEY, A. C., *Time series models*, New York usw. 1993.
- HASSLER, U., Einheitswurzeltests - Ein Überblick, in: *Allgemeines Statistisches Archiv* 78 (1994), S. 207-228.
- HIGGINS, M. L. / LTAIFA, N., MicroTSP Version 7.0: a review, in: *Journal of Applied Econometrics* 7 (1992), S. 423-429.
- HÖRMANN, A., et al., *Comparing Statistical analysis systems - a proposal for a list of comparison criteria with descriptions of BMDP, P-Stat, SAS, and SPSS*, in: *Statistical Software Newsletter* 16 (1990), S. 90-127.
- KIDD, M. P., A review of the Econometrics Toolkit (ET), in: *Journal of Applied Econometrics* 1 (1992), S. 323-328.
- KUZNETSOV, S. E. / KHALILEEV, A. A., *MESOSAUR: a companion to SYSTAT*, Evanston, IL: SYSTAT, Inc. 1991.
- LARSEN, H., A review of RATS 3.0: software for econometric modelling, forecasting, and simulation, in: *Journal of Applied Econometrics* 5 (1990), S. 87-91.
- LEINER, B., *Einführung in die Zeitreihenanalyse*, München / Wien 1986.
- LONGLEY, J. W., An appraisal of least squares programs for the electronic computer from the point of view of the user, in: *Journal of the American Statistical Association* 62 (1967), S. 819-831.
- MADDALA, G. S., *Introduction to econometrics*, New York 1992.
- MANHART, K., Wissenschaftliches Arbeiten unter einer grafischen Oberfläche, in: *ZA-Information* 33, November 1993, S. 94-114.
- MCKENZIE, C. R., MICROFIT 3.0: A review, in: *Journal of Applied Econometrics* 8 (1993), S. 413-419.
- MCLEOD, A. I. / SALES, P. R. H., An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. Algorithm AS 191, in: *Applied Statistics* 32 (1983), S. 211-223.
- MELARD, G., A fast algorithm for the exact likelihood of autoregressive-moving average models, in: *Applied Statistics* 33 (1984), S. 104-114.
- METZ, R., Ansätze, Begriffe und Verfahren der Analyse ökonomischer Zeitreihen, in: *Historical Social Research / Historische Sozialforschung* 13/3 (1988), S. 21-103.

- METZ, R. / STIER, W.**, NULFIL. Ein Programm zum Filtern im Frequenzbereich mit Spektralminimierung, in: **SPREE, R.**, *Lange Wellen wirtschaftlicher Entwicklung in der Neuzeit. Historische Befunde, Erklärungen und Untersuchungsmethoden*, Köln 1991 (= Historical Social Research / Historische Sozialforschung, Supplement 4), S. 139-144.
- METZ, R. / STIER, W.**, Filter design in the frequency domain, in: **KLEIN-KNECHT, A. / MANDEL, E. / WALLERSTEIN, I.**, *New findings in long-wave research*, London 1992, S. 45-79.
- MILLS, T. C.** *Time series techniques for economists*, Cambridge usw. 1990.
- NEAL, L.**, *The rise of financial capitalism. International capital markets in the age of reason*, Cambridge usw. 1990.
- PAARSCH, H. J.**, LIMDEP, version 6.0: a review, in: *Journal of Applied Econometrics* 9 (1994), S. 91-98.
- PICKLES, A.**, Software review: SPSS/PC+ version 4.0, in: *Applied Statistics* 41 (1992), S. 438-441.
- PROVAR, F.**, Software review: SPSS version 5 for Windows, in: *Applied Statistics* 42 (1993), S. 686-689.
- RAHLF, T.**, Ein einfaches Programm zur Berechnung von gewogenen gleitenden Mittelwerten in dBASE, in: *Historical Social Research / Historische Sozialforschung* 18/1 (1993), S. 122-144.
- RAHLF, T.**, Ausgewählte Methoden der konjunkturstatistischen Untersuchung am Beispiel des Lübecker Häusermarktes von 1284 bis 1700, in: **HAMMEL-KIESOW, Rolf** (Hg.), *Wege zur Erforschung städtischer Häuser und Höfe. Beiträge zur fächerübergreifenden Zusammenarbeit am Beispiel Lübecks im Spätmittelalter und in der frühen Neuzeit*, Neumünster 1993 (= Häuser und Höfe in Lübeck 1), S. 374-380.
- RAHLF, T.**, Statistische Graphik: ausgewählte Beispiele, in: **STEENWEG, Helge**, u.a. (Hgg.), *Geschichte und EDV: Probleme und Fortschritte, "Probleme mit dem Fortschritt?"*, Ergebnisse der 1. Tagung der Arbeitsgemeinschaft Geschichte und EDV, Göttingen, 23.-26.3.1994.
- ROSENQVIST, G.**, Software review: LIMDEP version 6.0, in: *Applied Statistics* 43 (1994), S. 290-293.
- RUST, J.**, GAUSS and MATHLAB: a comparison, in: *Journal of Applied Econometrics* 8 (1993), S. 307-324.
- SCHLITGEN, R.**, Robuste Glättung von Zeitreihen, in: *Allgemeines Statistisches Archiv* 74 (1990), S. 223-250.
- SCHLITGEN, R.**, Ein moderner Zugang zur Analyse von Zeitreihen mit Ausreißern, in: *Allgemeines Statistisches Archiv* 75 (1991), S. 75-102.
- SCHMITZ, B.**, *Einführung in die Zeitreihenanalyse. Modelle, Softwarebe-*

Schreibung, Anwendungen, Bern / Stuttgart / Toronto 1989

SCHLITGEN, R. / **STREITBERG, B. H. J.**, *Zeitreihenanalyse*, München / Wien 1984.

SENSCH, J., Statistical computing software reviews: LIMDEP 6.0 (1992), in: *Historical Social Research / Historische Sozialforschung* 19/2 (1994), S. 111-121.

SENSCH, J., *Statistische Modelle in der Historischen Sozialforschung: Anwendung formaler Modelle - Einsatz von Statistikprogramm-Paketen - Auswahlbibliographie zu statistischen Verfahren*, Köln 1994 (= *Historical Social Research / Historische Sozialforschung*, Supplement 7).

SPSS INC., *SPSS Statistical algorithms*, o.O. 1985.

SPSS INC., *SPSS statistical algorithms*, o.O. 1991.

SPSS INC., *Using SPSS/PC+ Graphics with Harvard Graphics*, o.O. 1993.

STATSOFT INC., *Validation benchmarks for Statistical algorithms (version 1B)*, Tulsa, OK 1992.

STIER, W., *Studie zur Verbesserung der schweizerischen Aussenhandelsindizes. Im Auftrag der Eidg. Zollverwaltung*, o.O., hg. von der Eidg. Oberzolldirektion, Bern, Juni 1993.

THOME, H., Beschreibende Zeitreihenanalyse: Komponentenzerlegung mit Hilfe gleitender Mittelwerte, in: *Historical Social Research / Historische Sozialforschung* 17/3 (1994), S. 63-105.

THOME, H., Ausreißer und fehlende Werte in der Zeitreihenanalyse: Ihre Modellierung im Rahmen des Box/Jenkins-Ansatzes, in: *ZA-Information* 31 (1992), S. 37-69.

THOME, H., Univariate Box/Jenkins-Modelle in der Zeitreihenanalyse, in: *Historical Social Research / Historische Sozialforschung* 19/3 (1994).

VEALL, M. R., SHAZAM 6.2: a review, in: *Journal of Applied Econometrics* 6 (1991), S. 317-320.

WILKINSON, L., *SYGRAPH: The system for graphics*, Evanston, IL: SYSTAT Inc. 1990.

WILKINSON, L., *SYSTAT: The system for statistics*, Evanston, IL: SYSTAT Inc. 1990.

Tabelle A1:
Übersicht über ausgewählte Programmfunktionen

●: vorhanden/möglich ○: eingeschränkt/indirekt vorh./mögl. -: nicht vorh./mögl.

	SPSS/PC+ (Trends)	SPSS für Windows	SYSTAT (Series)	SYSTAT for Windows	STATISTICA/w	MESOSAUR	MicroTSP
Im- und Exportformate							
ASCII	●	●	●	●	●	●	●
dBASE	●	●	●	●	●	●	-
Paradox	-	-	-	-	●	-	-
Oracle	-	●	-	-	-	-	-
SQL-Server	-	●	-	-	-	-	-
Lotus	●	●	●	●	●	-	●
Excel	-	●	-	●	●	-	-
SPSS	●	●	-	-	○	-	-
SYSTAT	-	-	●	●	-	●	-
STATISTICA	-	-	-	-	●	-	-
Handhabung							
Menüführung	●	●	●	●	●	●	●
Kommandosprache	●	●	●	●	●	-	●
Kommandogenerator	●	●	●	●	●	-	●
Programmiersprache	-	○	●	●	●	-	●
Hochauflösende Graphik							
Grafik der Zeitreihe	-	●	●	●	●	●	●
Histogramme	-	●	●	●	●	●	-
Scatterplots	-	●	●	●	●	●	●
Boxplots	-	●	●	●	●	-	-
Quantilplots	-	●	●	●	●	-	-
Normal Probability Plots	-	●	●	●	●	-	-
Periodogramm	-	●	○	○	●	●	-
Integriertes Periodogramm	-	○	○	○	●	●	-
Spektrum	-	●	○	○	●	●	-
Editionsmöglichkeiten	-	●	●	●	●	○	○

	SPSS/PC+ (Trends)	SPSS für Windows	SYSTAT (Series)	SYSTAT for Windows	STATISTICA/w	MESOSAUR	MicroTSP
Statistische Verfahren							
Mittelwert, Varianz	●	●	●	●	●	●	●
Median	●	●	●	●	●	●	●
Autokorrelation	●	●	●	●	●	●	●
w.n. Standardfehler	●	●	-	-	●	-	●
MA Standardfehler	●	●	●	●	●	●	-
Box-Ljung Statistik	●	●	-	-	●	-	○
Partielle Autokorrelation	●	●	●	●	●	●	●
Periodogramm	●	●	●	●	●	●	-
Integriertes Periodogramm	○	○	○	○	○	●	-
Spektrum	●	●	○	○	●	●	-
Spektralfenster:							
Hamming	●	●	-	-	●	-	-
Daniell (Rechteck)	●	●	-	-	●	-	-
Bartlett	●	●	-	-	●	-	-
Parzen	●	●	-	-	●	●	-
Tukey	●	●	-	-	●	-	-
benutzerdefiniert	●	●	○	○	●	-	-
einfache gleitende Mittelwerte	●	●	●	●	●	●	-
benutzerdefinierbare	-	-	●	●	-	-	-
gleitende Mediane	●	●	●	●	●	●	-
T4253H	●	●	-	-	-	-	-
LOWESS	-	-	●	●	-	-	-
DWLS	-	-	●	●	●	-	-
Frequenzfilter	-	-	-	-	-	●	-
Saisonadjustierung:							
Multiplikativ	●	●	●	●	●	●	●
Additiv	●	●	●	●	●	●	●

	SPSS/PC+ (Trends)	SPSS für Windows	SYSTAT (Series)	SYSTAT for Windows	STATISTICA/w	MESOSAUR	MicroTSP
Univariate ARIMA-Modelle	●	●	●	●	●	●	●
Automatische Identifikation	-	-	-	-	-	●	-
Schätzmethode für Parameter:							
ULS	-	-	●	●	●	-	-
CLS	-	-	●	●	●	●	●
Exakte ML	●	●	-	-	●	-	-
Informationskriterien							
AIC	●	-	-	-	-	●	-
SBC	●	-	-	-	-	-	-
"Neuere" Verfahren							
ARCH	-	-	-	-	-	-	●
GARCH	-	-	-	-	-	-	-
ARFIMA	-	-	-	-	-	-	-
Unit-root Tests	-	-	-	-	-	-	●

Bezugsadressen

MESOSAUR 1.0
 STATCON B. Schäfer
 Mündener Str. 1
 37213 Witzenhausen

STATISTICA/w 4.5
 Loll+Nielson
 Quickbornstr. 3
 20253 Hamburg

MicroTSP 7.0
 Quantitative Micro Software
 4521 Campus Drive, Suite 336
 Irvine, California 92715

SYSTAT 5.0.3
SYSTAT for Windows 5.0.3
 siehe MESOSAUR

SPSS/PC+ 5.0.1
SPSS für Windows 6.0
 SPSS GmbH Software
 Rosenheimer Straße 30
 81669 München