UFZ - Umweltforschungszentrum Leipzig-Halle GmbH

UFZ-Bericht

UFZ-Bericht • UFZ-Bericht • UFZ-Bericht • UFZ-Bericht

Nr. 4/1996 Dissertation

Immissionsprognose

Univariate Modellierung und Kürzestfristvorhersage von Wintersmogsituationen

Uwe Schlink

Sektion Expositionsforschung
und Epidemiologie
UFZ-Umweltforschungszentrum
Leipzig-Halle GmbH

ISSN 0948-9452

Immissionsprognose

Univariate Modellierung und Kürzestfristvorhersage von Wintersmogsituationen

Dissertation

zur Erlangung des akademischen Grades

Dr. rer. nat.

der Fakultät für Physik und Geowissenschaften
der Universität Leipzig

eingereicht von

Dipl.-Phys. Uwe Schlink, geb. 18.03.59 in Quedlinburg

angefertigt am

Umweltforschungszentrum Leipzig-Halle GmbH

betreut von

Professor Dr. O. Herbarth, Universität Leipzig

Beschluß über die Verleihung des Doktorgrades vom 8. Januar 1996

Bibliographische Beschreibung:

Schlink, Uwe
Immissionsprognose, Univariate Modellierung und Kürzestfristvorhersage von Wintersmogsituationen
Universität Leipzig, Diss.,
170 S., 67 Lit., 66 Abb., 5 Tab.

Referat:

Ist es möglich, ein SO_2-Warnsystem für Wintersmogsituationen zu konstruieren, welches seine Schlüsse allein aus dem Verlauf und der Variabilität der Halbstundenmittelwerte der Schwefeldioxidkonzentrationsreihe zieht? Wie könnte ein solcher Warnalgorithmus aussehen und wo liegen seine Grenzen?

Zur Klärung dieser Fragestellung wird die an der Station Leipzig-Mitte von 1980 bis 1993 gemessene SO_2-Reihe untersucht. Durch eine nichtlineare Transformation läßt sich die Variabilität der Meßreihe stabilisieren. Mittels Autokorrelations- und Spektralanalyse werden die Erhaltungsneigung und typische Zyklen des Immissionsverlaufs bestimmt. So ergibt sich ein Persistenzintervall von 8 Stunden. Zur Charakterisierung der Meßdaten sind unterschiedliche Skalenbereiche von Bedeutung: Die Vorgänge mit einer Periodizität von über einem Tag, die bei entsprechender Schadstoffemission Smogepisoden widerspiegeln, die 24- und die 12-Stunden-Schwingung, die durch den Rhythmus der Sonneneinstrahlung bedingt sind sowie die restlichen Periodizitäten mit Perioden kleiner als 24 Stunden, die an Windturbulenzen gleicher Frequenz gekoppelt sind.

Auf der Basis dieser Eigenschaften werden Zeitreihenmodelle konstruiert, mit denen eine Extrapolation des Reihenverlaufs möglich ist. Die damit berechneten Prognosen sind von unterschiedlichem Wert. Gewinnbringend ist die Vorhersage mit einem Komponentenmodell. Zur Verbesserung der Trendvorhersage wird anhand des Phasenraumbildes eine Methode der lokalen harmonischen Approximation entwickelt und getestet. Dabei kommt es zu einer fortlaufend aktualisierten Frequenzschätzung, um dem zeitabhängigen Spektrum von Immissionsreihen gerecht zu werden.

Für die Auslösung einer Warnmeldung wird die Immissionsprognose darauf bezogen, wie die Konzentrationsreihe nacheinander zwei Schwellwerte durchläuft. Dementsprechend werden Pseudo- und Überschreitungsfälle unterschieden. Eine Wahrscheinlichkeitsaussage über das Eintreten eines Überschreitungsfalls läßt sich aus Prediktoren ableiten. Dazu werden faktische Prediktoren und ein mit Hilfe der Zeitreihenmodellierung prognostizierter Prediktor herangezogen. Die Treffsicherheit der Vorhersage wird diskutiert und kann mit der von Niederschlagsprognosen verglichen werden.

Immissionsprognose

Univariate Modellierung und Kürzestfristvorhersage von Wintersmogsituationen

Inhalt

	Seite
Inhalt	2
Einleitung	7
Überblick	12
1. Teil: Die Meßdaten	**15**
1. Meßdatengewinnung	16
2. Box-Cox-Transformation	23
3. Spektraluntersuchungen	27
4. Autokorrelationsuntersuchungen	41
2. Teil: Die Zeitreihenmodellierung	**45**
1. Stochastische Modelle	46
2. ARIMA-Modelle	48
3. Modellierung der Tagesverläufe	52
4. Komponentenmodell	56
5. Zustandsraummodelle	60
3. Teil: Die Trendextrapolation	**63**
1. Das Inter- und Extrapolationsverhalten beim Kalman-Filter	64
2. Die Trendfunktion	68
3. Das Phasenraumbild	70
4. Lokale harmonische Approximation	74
5. Parameterschätzung	75
6. Modifizierte Trendextrapolation	79

Seite

4. Teil: Das Prediktionsproblem **81**

1. Definition und Lösungsansatz 82
2. Faktische Prediktoren 87
3. Prognostizierter Prediktor 94
4. Mit lokaler harmonischer Trendextrapolation prognostizierter Prediktor 98

5. Teil: Diskussion **101**

1. Fehlerquellen und Grenzen der univariaten Trendextrapolation 102
2. Ziel der Immissionsprognose: Ein Warnsystem 107

Schluß 109

Literaturverzeichnis 115

Anhänge **123**

Anhang A: Zeitreihen und ihre statistischen Kenngrößen 124

Anhang B: Die Box-Jenkins-Methode 129

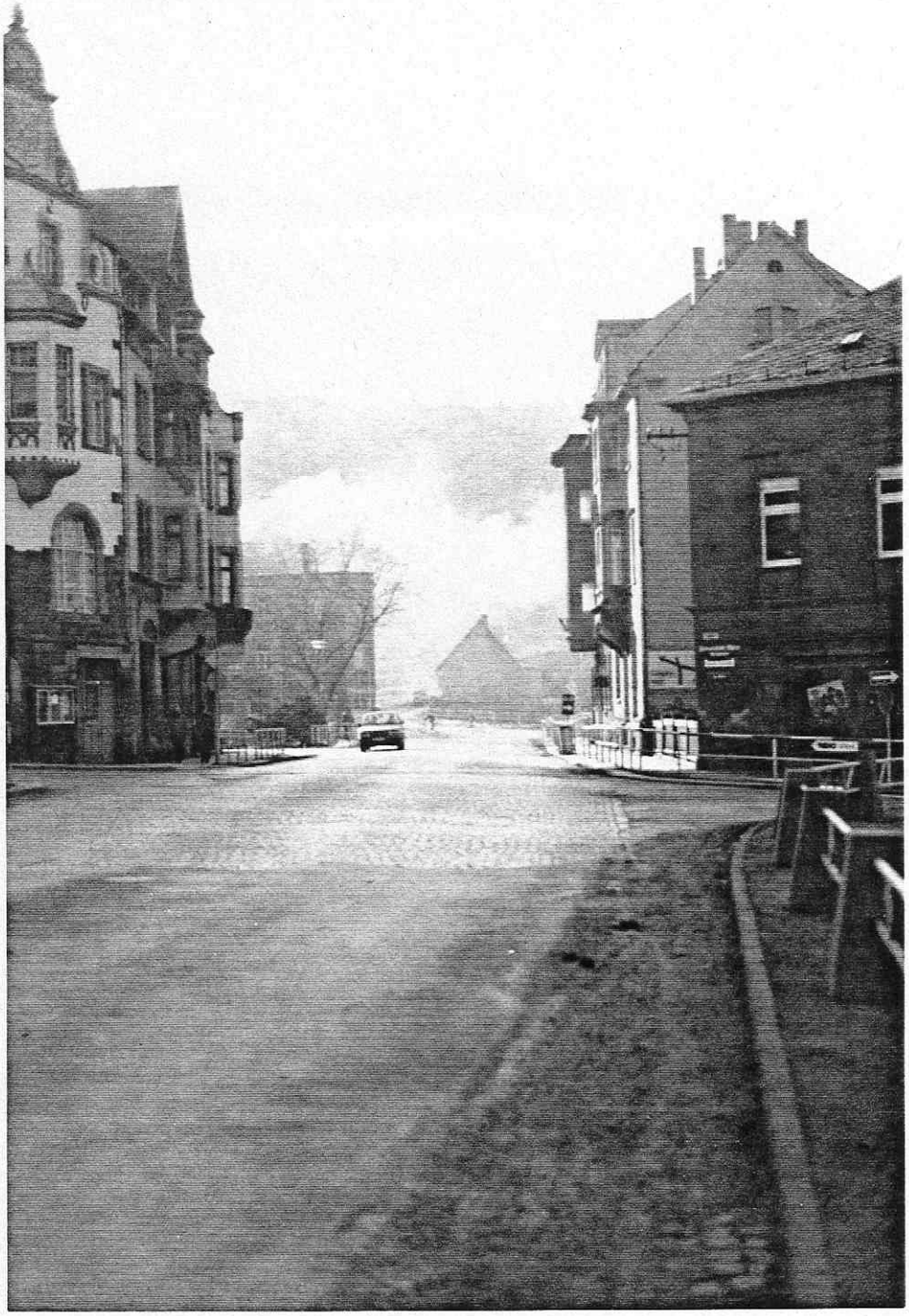
Anhang C: Zeitreihenfilterung 135

Anhang D: Der Kalman-Filter 139

Anhang E: Software 145

Anhang F: 1) Die Meßreihe CITY 151
2) Prognosesituationen an der Station Leipzig-Mitte 1980-1993 155

Anhang G: Verzeichnisse der Abbildungen, Tabellen und Symbole 161



„.... und der müde Wanderer, viele Meilen entfernt,
riecht die Stadt, zu der er sich begibt,
eher als er sie sieht."

JOHN EVELYN, Fumifugium (1661)

Einleitung

Die zunehmende Belastung der Umwelt durch Schadstoffe stellt eines der drängendsten Probleme der Gegenwart und der Zukunft dar.

In der vorliegenden Arbeit soll es um den Wintersmog gehen. Als Beispiel sei die Smogepisode des Jahres 1952 aufgeführt. Während einer Inversionswetterlage stieg in London der Schwefeldioxidgehalt der Luft um rund das Dreifache und der Schwebstaubgehalt um rund das Fünffache des Durchschnitts. Einige Tage später erhöhte sich die Sterblichkeit der Londoner Bevölkerung von rund 250 auf mehr als 800 Todesfälle pro Tag. In krasser Form wurde das Ausmaß der Luftverunreinigung und deren Auswirkungen auf Morbidität bzw. Mortalität den Behörden und weiten Kreisen der Bevölkerung deutlich.

Um die von Luftschadstoffen ausgehenden Gefahren und Risiken für Mensch und Umwelt zu minimieren, verfolgt man grundsätzlich zwei Strategien:

Die an der **Emission** orientierte Strategie besteht darin, mittels Rechtsverordnungen die Freisetzung von Schadstoffen zu kontrollieren und zu begrenzen. Rückblickend kann man als eines der ersten Umweltschutzgesetze des Abendlandes das Edikt "De conservatione aeris" des Stauferkaisers Friedrich II aus dem Jahre 1240 betrachten. Dort heißt es (Fischer (1933)):

> *"Wir sind bestrebt, die uns durch Gott geschenkte Gesundheit der Luft durch unsere Vorsorge, soweit das möglich ist, rein zu halten. Wir verfügen deshalb, daß es niemandem gestattet ist, in Gewässern, die weniger als eine Meile von einer Ansiedlung entfernt liegen, Flachs oder Hanf zu wässern, weil dadurch die Beschaffenheit der Luft ungünstig verändert wird."*

Heute gibt es eine Vielzahl von Regelungen, von denen insbesondere das Bundes-Immissionsschutzgesetz einschließlich der dazugehörigen Verordnungen und darauf aufbauend die Technische Anleitung zur Reinhaltung der Luft (TA-Luft) sowie eine Reihe von VDI-Vorschriften hervorzuheben sind. Durch die Umweltbehörden der Bundesländer erfolgt die praktische Überwachung der Einhaltung von maximalen Emissions-Konzentrationen (MEK) in der von der TA-Luft vorgeschriebenen Form.

Die Smogverordnungen der einzelnen Bundesländer, die sich an eine Muster-Smogverordnung anlehnen, verfolgen dagegen eine mehr an der **Immission** orientierte Strategie:

Die Immission ist die aus der Emission entstandene Verunreinigung der Luft. In den bodennahen Luftschichten wirkt sie unmittelbar auf Menschen, Tiere, Pflanzen oder Sachgüter ein. Zwischen Emission und Immission finden in der Atmosphäre unter dem Einfluß meteorologischer, physikalischer oder chemischer Prozesse Veränderungen der Schadstoffe statt.

Mit den in der Smog-Verordnung (1991) gegebenen Schwellen für die Immissionskonzentrationen wird nun die vorliegende Immissionssituation beurteilt. Bei der Überschreitung dieser Grenzwerte (die später noch konkret angegeben werden) sind sowohl emissionssenkende, als auch verkehrseinschränkende Maßnahmen vorgesehen. Darüberhinaus werden Verhaltenshinweise für die Bevölkerung gegeben.

Auch die TA-Luft (1993) gibt Grenzwerte vor, die die Behörden bei der Genehmigung von Anlagen binden. Das arithmetische Jahresmittel aus Halbstundenwerten darf dabei beispielsweise für SO_2 0,14 mg SO_2 / m³ und das 98%-Perzentil aller Halbstundenwerte eines Jahres 0,40 mg SO_2 / m³ nicht überschreiten.

Die VDI-Richtlinie 2310 gibt maximale Immissionskonzentrationen (MIK) für den Schutz des Menschen vor toxischen Einwirkungen mit 0,30 mg/m³ für das 24-Stunden-Mittel der Schwefeldioxidkonzentration und 1,00 mg/m³ für das Halbstundenmittel bei einmaliger Exposition an.

Dies gibt jedem die Möglichkeit, Umwelteinwirkungen zu bewerten und somit auch zu meiden. Man kann also Primärprävention betreiben und dem Entstehen pathologischer Veränderungen vorbeugen.

Mit Hilfe der in und von den Ländern der Bundesrepublik installierten Immissions-Überwachungssysteme werden Smogsituationen jedoch in der Regel erst bemerkt, wenn der kritische Zustand erreicht oder bereits vorüber ist.

Eine präventive Vermeidungsstrategie setzt demzufolge die Kenntnis der Wahrscheinlichkeit voraus, daß die Immissionskonzentration innerhalb einer bestimmten Zeit ab einem festgelegten Prognosezeitpunkt einen medizinisch relevanten Schwellenwert überschreitet (vgl. Herbarth (1993)). Methoden zur Beantwortung dieser Fragestellung zu finden, ist das Ziel der vorliegenden Arbeit.

Natürlich ist ein solches Prognosesystem auch für die Behörde, die die Auslösung von Smogalarm zu verantworten hat, eine Entscheidungshilfe. Hinter den hier durchgeführten Untersuchungen zu Vorhersagemethoden steht also im Grundmotiv die Umwelthygiene.

Der Wintersmog bildet sich besonders typisch an kalten Herbst- und Wintertagen, schwerpunktmäßig im Dezember, Januar und Februar. Die Wetterlage ist geprägt durch windschwache Verhältnisse mit Kaltluft in Bodennähe und darüberliegender Warmluft, was als Inversion bezeichnet wird. Diese stabile Situation verhindert den Luftaustausch mit der höheren Atmosphäre und führt in Bodennähe (bis ca. 200 m Höhe) zu einer starken Anreicherung von Luftverunreinigungen, speziell von Schwefeldioxid und Ruß- bzw. Schwebstaub, was dann ein Indikator für eine Smogerscheinung ist. Neben diesen bisher typischen Wintersmogschadstoffen ist künftig mit einer wachsenden Bedeutung von Schadstoffen der Verkehrsemission, wie z.B. NO_X zu rechnen.

Beispielsweise wird die Smog-Vorwarnstufe gemäß Smog-Verordnung bekanntgegeben, wenn eine austauscharme Wetterlage besteht und die über drei Stunden gemittelte Konzentration von SO_2 die Schwelle von 0,6 mg/m³ überschreitet. Dabei wird die Wettersituation vom Deutschen Wetterdienst beurteilt und muß folgende Kriterien erfüllen: Die Temperatur nimmt in der unteren Luftschicht mit der Höhe zu (Temperaturumkehr), die Windgeschwindigkeit ist sehr klein (<3m/s) und es ist nicht auszuschließen, daß die Wetterlage länger als 24 Stunden anhalten wird.

Je nach regionalen Verhältnissen erscheint der Wintersmog in verschiedenen Modifikationen: Für die in London auftretende Form ist der Nebel ein wesentlicher Bestandteil, was schließlich auch zu der Wortschöpfung "Smog" (=smoke+fog) geführt hat.

Die Untersuchungen dieser Arbeit beziehen sich auf die in der Leipziger Industrieregion des mitteldeutschen Raums auftretenden Wintersmogsituationen. Kommt es hier zur Smogwarnung, so liegt das in den meisten Fällen an einer Überschreitung des Kriteriums für Schwefeldioxid. Das SO_2 ist somit die Leitkomponente für den Wintersmog der betrachteten Region. Aus diesem Grunde ist es auch zum Gegenstand der vorliegenden Arbeit geworden.

Bemerkung: Schwefeldioxid ist ein farbloses, stechend riechendes, zum Husten reizendes Gas, welches sich leicht in Wasser löst und dabei teilweise schweflige Säure bildet. Der Hauptteil wird jedoch photochemisch zu SO_3 oxidiert, welches wiederum mit den Wassermolekülen der Luft Sulfat-Aerosole bildet, die in Abhängigkeit von den meteorologischen Bedingungen 10 bis 100.000 km weit transportiert werden können.

Die Sulfat-Aerosole werden sedimentiert oder gemeinsam mit nicht umgesetztem SO_2 durch den Niederschlag ausgewaschen, der dann einen pH-Wert von 3,5 ... 5,0 erreicht und als "saurer Regen" bezeichnet wird. Die mittlere Verweilzeit des SO_2 in der Atmosphäre ist abhängig von den meteorologischen Bedingungen und wird mit 0,5 Tagen im Jahresdurchschnitt angegeben. Sie schwankt zwischen 23 Tagen bei trockenem, sonnigem Wetter und 0,1 Tagen bei Regen. 70 ... 85 % des emittierten Schwefels werden über kontinentalem Gebiet abgeschieden.

Nach groben Schätzungen ist das Schwefeldioxid für etwa 50% der durch Luftverunreinigung entstandenen Schäden verantwortlich, so z.B. erhöhte Korrosion von Metallen, Verringerung der Haltbarkeit von Farben, Lacken und Holz, Erosion von Plasten, Zerstörung von Bauwerken. Steine zerfallen bei einem pH-Wert von 4 in saurer Atmosphäre 50 ... 100mal schneller, als im neutralen Millieu.

Aufgrund der Assimilation reagieren Pflanzen auf Luftverunreinigungen empfindlicher als Mensch und Tiere. Zentraler Angriffspunkt des SO_2 ist die Konkurrenz mit CO_2 um die CO_2 metabolisierenden Enzyme. Das führt zu einer Beeinträchtigung der Photosynthese.

Beim Menschen wird SO_2 auf der feuchten Oberfläche der Schleimhäute des Atemtrakts als schweflige Säure gelöst, teilweise unter Oxydation zu Schwefelsäure. Es ist mäßig toxisch und hat akute und chronische Wirkungen. Dazu zählen Reizung und Entzündung von Haut und Schleimhaut, der Augen und der Atemwege. Im oberen Atemtrakt kommt es zu einer Zunahme des Atemwiderstandes und einer erhöhten Infektanfälligkeit sowie zu Tracheitis, Bronchitis und Pneumonie. Hohe Schwefeldioxidkonzentrationen bewirken Atemnot, Zyanose und Herz-Kreislauf-Versagen (Horn (1989), Beck & Schmidt (1992)).

Kontinuierliche Einwirkungen von SO_2 haben bei gleicher Konzentration eine größere Schadwirkung als intermittierende. Im Langzeitversuch wirkt SO_2 kanzerogen. Risikogruppen sind vor allem Personen mit Krankheiten der Atemwege und des Herz-Kreislauf-Systems sowie alte Menschen und Kinder.□

Schwefeldioxid gelangt aus natürlichen und aus anthropogenen Quellen in die Atmosphäre. Wichtige natürliche Emittenten sind neben den Vulkanen die Bodenbakterien und Pflanzen. Der Mensch trägt durch die Verbrennung fossiler Brennstoffe wesentlich zum Schwefeleintrag in die Atmosphäre bei. 93% des globalen anthropogenen SO_2 werden auf der Nordhalbkugel emittiert, was dort natürlich auch zu einer stark erhöhten Exposition führt.

So wurde beispielsweise vor 1989 in den genannten Regionen Mitteldeutschlands stark schwefelhaltige Braunkohle für Heizzwecke verwendet. Die Folge waren extrem hohe SO_2-Konzentrationen in der Außenluft, die beispielsweise im Halbstundenmittel 4 mg/m³ durchaus erreichen und sogar noch überschreiten konnten (Leipzig-Mitte, Februar und März 1984).

Wie sich die SO_2-Konzentration im Laufe eines Jahres verhält, ist in Abb.1 am Beispiel der Tagesmittelwerte der Meßstation Leipzig-Mitte für 1989 dargestellt. Dabei ist zu beachten, daß es sich um Tagesmittel und nicht um eine gleitende Mittelung handelt. Der MIK-Wert der VDI 2310 ist zum Vergleich durch die Linie bei 0,3 mg/m³ markiert worden.

Während der Heizperiode, die die Monate Januar, Februar, März sowie Oktober, November, Dezember umfaßt, ist die Schwefeldioxidkonzentration gegenüber den Monaten April bis September deutlich erhöht. Überschreitungen des MIK-Wertes treten vorwiegend in der Heizperiode auf.

Die fortschreitende Umstellung der Brennstoffbasis findet natürlich auch in den Immissionsdaten ihren Niederschlag. Wird Wintersmog dann auch künftig noch von Bedeutung sein?

Besonders in Innenstadtgebieten, die von einer Vielzahl von Emittenten betroffen sind, ist auch in den nächsten Jahren noch mit kritischen Situationen zu rechnen. Neben der Emissionsstruktur beeinflussen vor allem die meteorologischen und die orographischen Verhältnisse die Smogneigung einer Region. Deshalb werden nicht nur industrielle Ballungszentren (wie beispielsweise der Raum Halle-Leipzig), sondern auch am Gebirgsrand oder in Kerbtälern gelegene Orte (wie z.B. Wernigerode oder Aue) unter ungünstigen Wetterbedingungen in der Heizperiode smoggefährdet sein.

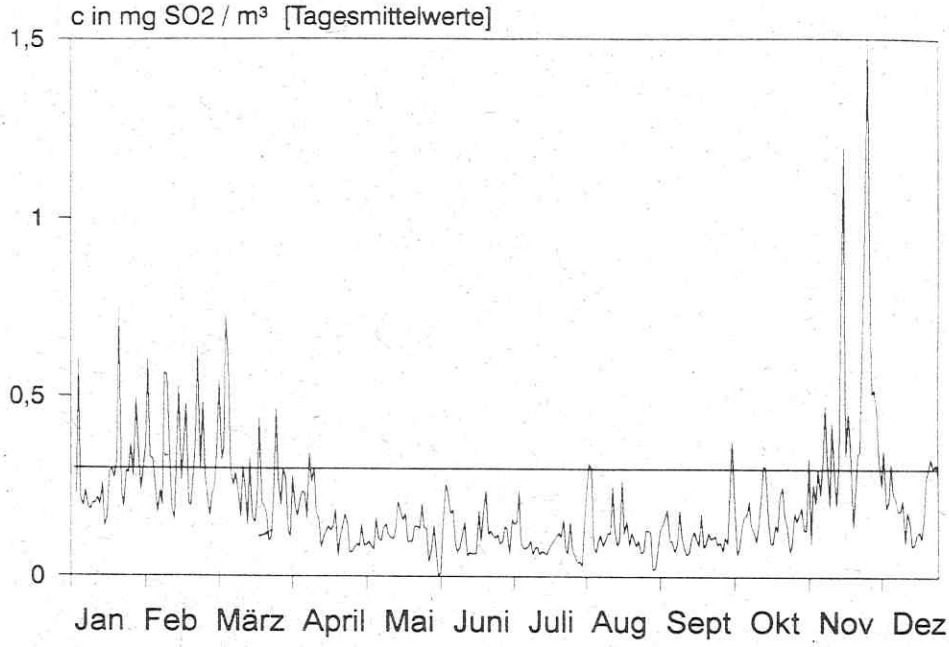


Abb.1: Typischer Jahresgang der SO_2-Konzentration; Tagesmittelwerte für Leipzig-Mitte 1989; MIK-Wert bei 0,3 mg/m³.

Die bei der Untersuchung des Wintersmogs gefundenen Methoden der Zeitreihenanalyse und -Modellierung sind darüber hinaus ein interessanter Ausgangspunkt für weiterführende Forschungen zu Problemen des Sommersmogs, die zunehmend an Aktualität gewinnen.

Wegen der gegenüber SO_2 deutlich höheren Reaktivität des Ozons werden beim Sommersmog die Vorgänge aber komplexer sein. Letzterer läuft im Rahmen eines komplizierteren chemischen Reaktionsgeflechts ab. Dessen Reaktionsdynamik muß wegen der gegenüber der Wintersmogchemie erhöhten Geschwindigkeitskonstanten in Modellierung und Prognose mit einbezogen werden.

Überblick

Ausgangspunkt dieser Arbeit war die Idee, die Persistenzeigenschaft von Zustandsgrößen der Atmosphäre für prognostische Zwecke, speziell für die Kürzestfristprognose über einige Stunden, auszunutzen und dabei deren Möglichkeiten und Grenzen abzustecken. Wie ist das zu verstehen?

Wie bereits erwähnt, verweilt das Schwefeldioxid im Mittel etwa einen halben Tag in der Atmosphäre. Dies ist also die charakteristische Zeit für die Veränderung der SO_2-Konzentration durch physikalische und chemische Vorgänge. Dazu zählen insbesondere Transportvorgänge infolge des Strömungsverhaltens der Luft sowie Diffusion, Koagulation, Kondensation, Adsorption und chemische Vorgänge, die zur Stoffumwandlung führen.

Wird von der Wettervorhersage nicht der Zustrom neuer Luftmassen angekündigt, so ist auch bei den mesoklimatischen Bedingungen keine plötzliche Veränderung (z.B. innerhalb einer Stunde) zu erwarten. Inversionswetterlagen mit einer stabilen Schichtung der Luft stellen die Grundbedingung für den Regelfall einer Smogsituation dar.

Man kann also vermuten, daß bei der Entstehung von Smogepisoden die gestiegene Erhaltungsneigung der SO_2-Konzentration eine Rolle spielt. Man bezeichnet dies auch als Persistenz. Mit Hilfe der Autokorrelationsfunktion kann man sich ein Bild von der Stärke der Erhaltungsneigung einer Serie zeitlich äquidistanter Meßwerte, die als Zeitreihe bezeichnet wird, machen. Im Teil 1 der vorliegenden Arbeit wird dies bei der Beschreibung der Meßdatenbasis erläutert.

Auf der Grundlage dieser Vorüberlegungen wird versucht, die Persistenz für eine Prognose zu nutzen. Im Teil 2 sind die Experimente dazu beschrieben. Es zeigt sich aber, daß die Verwendung der Erhaltungsneigung allein wenig gewinnbringend ist. Es müssen darüber hinaus auch noch Erkenntnisse über das spektrale Verhalten der Meßreihe (Teil 1, Abschnitt 3) einfließen, um zu einem nutzbaren Zeitreihenmodell zu kommen. Dieses besteht aus Komponenten mit unterschiedlichen und ganz typischen Eigenschaften, mit denen eine Extrapolation des Zeitreihenverlaufs möglich ist. Damit wird nicht nur die Persistenz der Zeitreihe genutzt. Der Komponentenansatz zielt darauf ab, möglichst viele ihrer phänomenologischen Eigenschaften in das Vorhersagemodell einzubeziehen.

Aber wie soll man eine für die Immissionszeitreihe adäquate Extrapolation durchführen? Im Teil 3 wird deshalb diskutiert, wie für die Trendkomponente der Meßreihe auf der Basis des Kalman-Filters bisher übliche Extrapolationen berechnet werden und ein neues Verfahren für die Fortsetzung des Zeitreihentrends angegeben. Dieses Verfahren berücksichtigt neben dem Anstieg der Trendreihe auch noch ihre Krümmung. Allgemein gesprochen modelliert das hier entwickelte Extrapolationsverfahren einen Prozeß mit zeitlich veränderlichem Spektrum. Den Zugang zu der Extrapolationsmethode liefert die Phasenraumdarstellung der Trendkomponente. Das Ziel besteht darin, die Prognosegüte zu verbessern.

Inwieweit dieses Ziel erreicht wird, dokumentiert Teil 4. Dafür ist es zunächst erforderlich, das Prediktionsproblem der Lufthygiene zu definieren. Die Prediktion dient letztendlich der Abschätzung eines gesundheitlichen Risikos bei der zu erwartenden Exposition. Jede Risikoangabe ist eine Wahrscheinlichkeitsangabe. Bei dem Prediktionsproblem der Lufthygiene muß es also um die Bestimmung einer Wahrscheinlichkeit in Bezug auf die Überschreitung von medizinisch relevanten Schwellwerten gehen. Für die Gewinnung dieser Wahrscheinlichkeit werden Indikatorgrößen herangezogen, die Prediktoren.

Smogepisoden sind durch eine oder mehrere der folgenden Eigenschaften charakterisiert:

- Es liegt eine stationäre antizyklonale Wetterlage vor.
- Es liegt eine vertikale Temperaturinversion vor.
- Die Windgeschwindigkeit ist niedrig (< 3m/s).
- Die Schadstoffkonzentrationen in der Luft steigen auf kritische Werte an.
- Die Dauer solcher Smogepisoden liegt meist zwischen 2 und 7 Tagen.

Daraus geht klar hervor, daß die Kenntnis des Wetterablaufs die conditio sine qua non für eine exakte Immissionsprognose ist. Warum beschränkt sich die vorliegende Untersuchung dann lediglich auf die univariate Betrachtung der phänomenologischen Eigenschaften der Immissionsreihe?

Generell favorisiert die Forderung nach Automatisierbarkeit des Prognosealgorithmus die Verwendung eines statistischen Ansatzes gegenüber subjektiv-synoptischen und aufwendigen dynamischen Verfahren. Statistische Modelle, die die wesentlichen kausalen Einflußgrößen (Meteorologie, Emission) auf die SO_2-Konzentration einbeziehen, arbeiten mit der regressiven Adaptation und dem Einsatz von Cluster- und Diskriminanzanalyse (Balzer (1986), Enke (1986)). Das Modell spiegelt dann kausale Zusammenhänge wider, andererseits stellt die Bereitstellung der erforderlichen Meßdaten einen zusätzlichen Aufwand dar.

Will man die Problemstellung deshalb möglichst weit vereinfachen, so bietet sich die folgende Überlegung an: Die Ergebnisse aller Wechselwirkungen der Kausalkette Emission-Transmission-Immission gehen letztendlich synergisch in den Immissionsverlauf ein - alle diesen Prozeß beeinflussenden Größen finden ihren (gewichteten) Niederschlag. Damit ergibt sich die interessante Fragestellung:

Inwieweit lassen sich aus der Phänomenologie allein dieser Immissionsreihe Informationen für eine Kürzestfristprognose gewinnen?

Eine solche Prognose antropogen beeinflußter Prozesse allein aus deren Variabilität ist dabei eine völlig neue Herangehensweise. In wenigen phänomenologischen Erscheinungen kann eine Vielzahl von kausalen Einflüssen und Wechselwirkungen enthalten sein. Dadurch wird eine vereinfachte Sicht der Smogerscheinung erreicht, die zu weniger aufwendigen Prognosealgorithmen führen soll. Diese könnten direkt an den Immissionsmeßstationen ablaufen und dort quasi im real-time-Betrieb Vorhersagen liefern, die einer handlungsorientierten Politik den entsprechenden Spielraum zur Auslösung von Präventionsmaßnahmen geben könnten.

Die damit erzielten Vorhersagen werden natürlich nur innerhalb eines sehr begrenzten Zeithorizonts gültig sein, so daß auch nach den Grenzen dieser Herangehensweise gefragt werden muß. Der Teil 5 enthält kritische Betrachtungen dazu.

Bemerkung: Man hat also bei der Modellierung grundsätzlich zwischen der prozeßorientierten und der phänomenologischen Methode zu unterscheiden. Eine sehr moderne phänomenologische Modellierungsmethode ist zum Beispiel die Anwendung der Fuzzy Logik (vgl. z.B. Bandemer (1989)).□

Bemerkung: Eine ganz analoge Situation gibt es beispielsweise auch bei den Methoden der Aktienkursbewertung. Der Verlauf (Chart) des Aktienkurses, auf den sehr viele Parameter einen Einfluß haben, läßt sich in Form einer Zeitreihe ausdrücken. Die Analysenmethoden dieses Kursverlaufs teilt man dann in zwei Bereiche ein, in die fundamentale und in die technische Analyse. Bei der Fundamentalanalyse befaßt man sich mit den konkreten Einflußfaktoren, wie Branchensituation, allgemeine Konjunkturlage, Ertrags- und Bilanzzahlen einer Gesellschaft, etc.. Im Gegensatz dazu orientiert sich die technische bzw. Chartanalyse ausschließlich an den Kursbewegungen der Aktie. Der Chartanalyst geht bei der Chartanalyse grundsätzlich davon aus, daß sich die Summe der mehr oder weniger gut abschätzbaren Bestimmungsfaktoren über die Beeinflussung der Angebots- und Nachfragekonstellation schließlich im Kurs niederschlägt (vgl. z.B. Vereinsbank (1989)).□

1983 wurde von Herbarth (1983) begonnen, den autokorrelativen Zusammenhang der Halbstundenmittelwerte von Schwefeldioxid zu untersuchen und für eine Modellierung mit ARIMA-Modellen nach Box & Jenkins (1970) zu nutzen. Auch die Studie von Wiegand et al. (1987) wandte dieses Verfahren an. Als Basis dienten dabei jedoch ausschließlich Tages- und Monatsmittelwert-Reihen. Für das Verhalten der Serie der Halbstundenmittel wurde von den Autoren ein noch offener Forschungsbedarf festgestellt.

Aus diesem Grunde finden hier, an die Arbeiten von Herbarth (1987) anschließend, prinzipiell bei allen Untersuchungen Serien von Halbstundenmittelwerten Anwendung. Diese Meßreihen wurden mit Hilfe automatisch arbeitender Stationen im Rahmen des Immissionsüberwachungssystems des Freistaates Sachsen gewonnen (vgl. Anh. F).

Reihen von zeitlich äquidistant aufeinanderfolgenden Meßwerten werden als Zeitreihen bezeichnet (vgl. Anhang A). Zu ihrer Behandlung hat sich die Zeitreihenanalyse als eigenständige Disziplin entwickelt. In diesem wissenschaftlichen Umfeld ist auch die vorliegende Arbeit zu sehen.

Bei den hier beschriebenen Untersuchungen wird einerseits mit bekannten Methoden der Zeitreihenanalyse experimentiert, um ihre Verwendbarkeit für Immissionsprobleme zu testen. Die dabei verwendeten bekannten theoretischen Ansätze sind kurz in den Anhängen skizziert.

Basierend auf den daraus gewonnenen Erkenntnissen wird eine neue Vorhersagemethode entwickelt. Der Ablauf der Datenanalysen, die Modellierungsversuche sowie die Ausarbeitung der neuen Herangehensweise und deren Beurteilung sind in den Teilen 1 bis 5 in logischer Reihenfolge erläutert.

1. Teil: Die Meßdaten

1. Meßdatengewinnung

Für die Stadt Leipzig liegen kontinuierliche Messungen der SO_2-Konzentration seit 1978 vor. Diese Serien von Halbstundenmittelwerten wurden durchgehend an der Meßstation in der Beethovenstraße 25 registriert. Die Station wird vom Sächsischen Landesamt für Umwelt und Geologie betrieben.

Am Lageplan (Abb. 1-1) wird deutlich, daß sich die Station in der Nähe des Stadtkerns von Leipzig befindet. Sie soll deshalb als Station "Leipzig-Mitte" bezeichnet werden. Das Gebiet ist durch dichte und hohe Bebauung mit Grünanteilen charakterisiert. In den Straßenschluchten zwischen den mehrstöckigen Häusern kann es zu einer starken Schadstoffanreicherung und böenartigen Windverwirbelungen kommen. Die Entnahmestelle für die Luftprobe ist auf der Photographie Abb. 1-2 mit einem Kreis markiert worden.

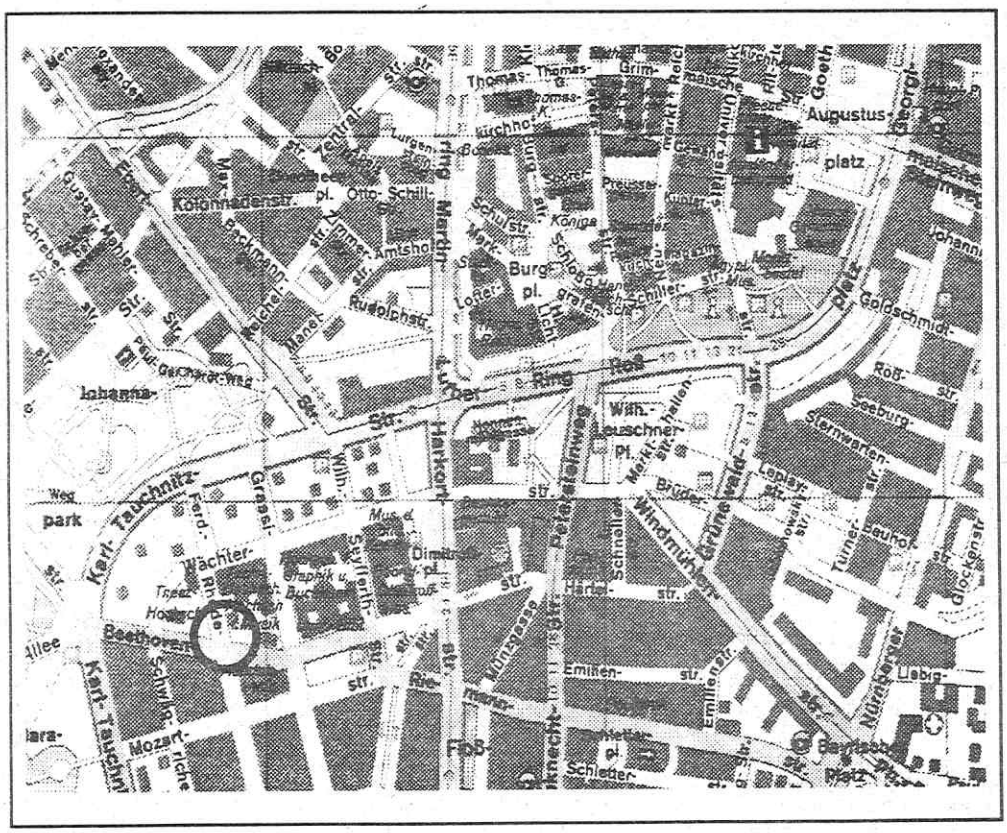


Abb. 1-1: Die Lage der Meßstation in der Beethovenstraße im Zentrum Leipzigs

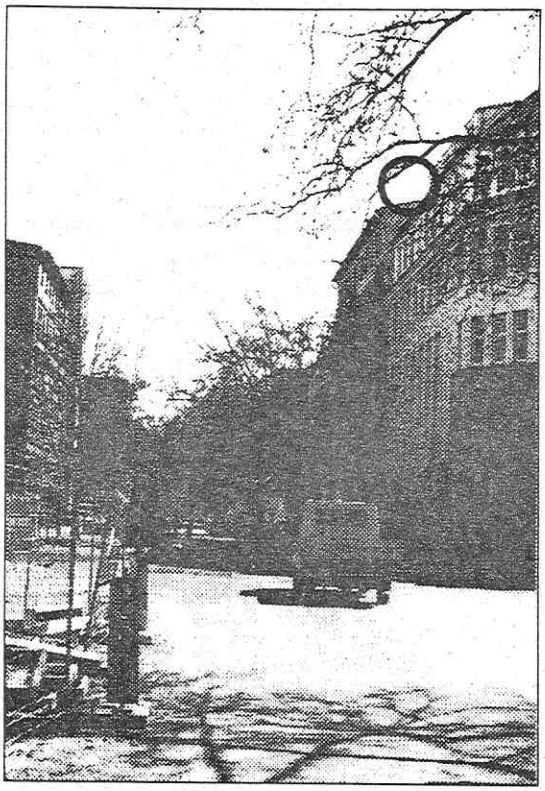


Abb.1-2: Die automatische Probenahme am Fenster des Hauses Beethovenstr.25

Dominante Emissionsquelle für SO_2 war im Innenstadtgebiet bis Anfang der 90er Jahre der "Hausbrand". Die Umstellung der Heizmaterialien von oftmals sehr schwefelhaltiger Braunkohle auf Erdgas oder -öl führte ab 1990 zu einer deutlichen Senkung der Belastung. Diese Tendenz läßt sich an Tab. 1-1 ablesen: Die Anzahl der Situationen, in denen das gleitende 3-Stunden-Mittel die Schwelle 0,6 mg/m³ (Kriterium der Smogverordnung) erreicht, nimmt von Jahr zu Jahr ab. Allerdings sind 1991 noch 22 kritische Situationen registriert worden, was auf die bei ungünstiger Witterung immer noch vorliegende Wintersmogproblematik hinweist.

Tab. 1-1: Häufigkeit der Überschreitung der Schwelle 0,6 mg/m³ durch das gleitende 3-Stunden-Mittel der Meßdaten von Leipzig-Mitte.

Jahr	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Verfügbarkeit der Daten	73%	85%	94%	79%	86%	95%	70%	90%	92%	92%	94%	89%	84%	90%
Überschreitungshäufigkeit	63	64	66	37	78	86	51	71	49	55	13	22	6	2

Wie repräsentativ ist der Meßstandort für das gesamte Stadtgebiet? In Abb. 1-3 ist das Stadtgebiet in Flächen eingeteilt worden, die sich nach ihrer Nutzungsart unterscheiden. Die Nutzungsart ist eng mit der Schadstoffemission verbunden.

Im Hinblick auf die hier interessierende Frage der Belastung wurde jedoch für jede Fläche anhand der Immission ein Langzeitluftbelastungsindex (LBI) bestimmt (Müller (1993)). Dieser ergibt sich aus der Summe der Verhältnisse von I1-Kenngröße zum Langzeitgrenzwert (IW1) der TA-Luft (Jost (1992)) für die Schadstoffe SO_2, NO_x und CO. Durch eine Klassierung des LBI erhält man Belastungsstufen. Wie dem Bericht von Müller (1993) zu entnehmen ist, ist in jedem Fall die SO_2-Konzentration der für die Wahl der Belastungsstufe ausschlaggebende Parameter. Die Flächen sind entsprechend der vorliegenden Belastungsstufe eingefärbt worden.

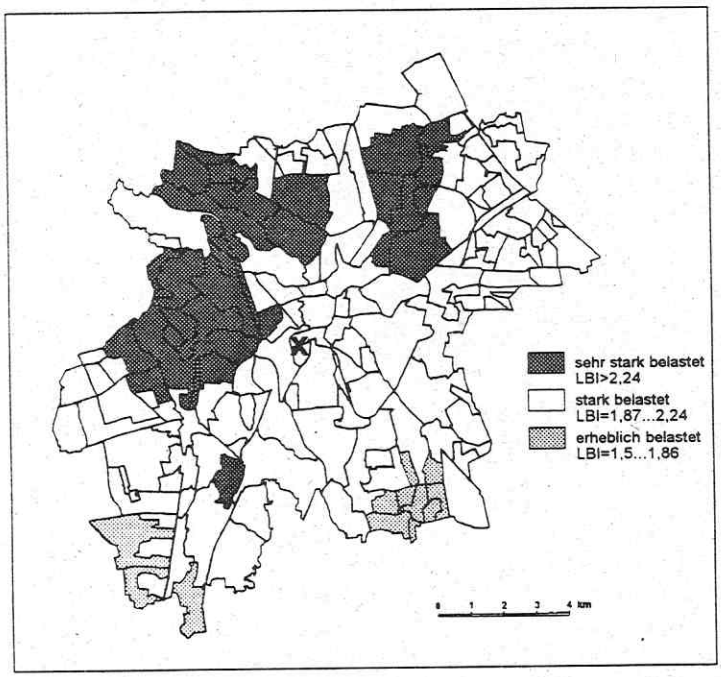


Abb. 1-3: Luftbelastung in der Stadt Leipzig 1990/91 (aus Müller (1993), LBI = Langzeitluftbelastungsindex); die Lage der Station Leipzig-Mitte ist durch ein Kreuz markiert worden.

Die durch ein Kreuz markierte Meßstation befindet sich also in einem stark belasteten Gebiet. Aufgrund der in gleicher Weise belasteten Umgebung wird es bei einsetzenden Winden, die entsprechend der langjährig bestimmten Verteilung der Windrichtungen (Abb. 1-4) vorzugsweise aus südlichen und westlichen Richtungen wehen, nicht zu einem Verdünnungseffekt kommen. Es ist also plausibel, die Meßstelle Leipzig-Mitte für das Territorium der Stadt als repräsentativ anzusehen. Untersuchungen, die diese Repräsentanz detaillierter belegen, wurden von Bredel & Herbarth (1987) durchgeführt.

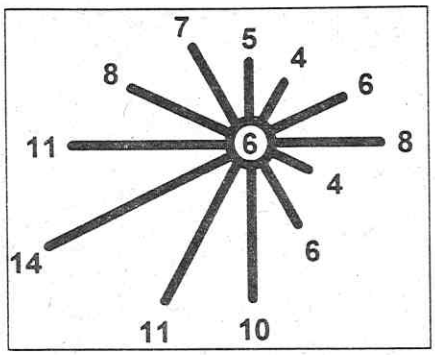


Abb. 1-4: Mittlere jährliche Verteilung der Windrichtungen in der Stadt Leipzig von 1961 bis 1980 (Messung in Prager Str. 169 in ca. 15 m Höhe; Angaben in Prozent); entnommen aus dem Umweltbericht Leipzig (1991)

Die Datenreihe der Immissionsmeßstelle Leipzig-Mitte ist im Zeitraum von 1980 bis 1993 zu 87% verfügbar. Sie stellt damit die für Leipzig umfassendste repräsentative SO_2-Meßreihe dar. Aus diesem Grunde wurde sie für diese Arbeit als Basis herangezogen.

Bemerkung (1-1): Die Messungen erfolgten seit 1990 mit einem automatisch arbeitenden Meßgerät nach dem Ultraviolett-Fluoreszenz-Meßverfahren. Das kontinuierlich angesaugte Meßgas wird mit einer UV-Lampe bei ca. 215 nm bestrahlt. Dadurch werden die im Meßgas vorhandenen SO_2-Moleküle zur Fluoreszenz angeregt, d.h. diese angeregten Moleküle geben UV-Licht mit einer Wellenlänge von 240 ... 420 nm (Maximum bei 320 nm) wieder ab. Der Grad der Fluoreszenz ist eine Funktion des SO_2-Gehalts der Probe und der Stärke des Erregerlichts. Durch eine elektronische Auswerteeinheit wird die SO_2-Konzentration aus dem Ausgangssignal bestimmt. Da beim UV-Fluoreszenzverfahren außer Schwefeldioxid auch bestimmte aromatische Kohlenwasserstoffe in der Luft ansprechen (z.B. m-Dimethylbenzol, Naphtalin usw.) ist ein 2-stufiger Kohlenwasserstoffabscheider im Gerät integriert.

Das verwendete Gerät (HORIBA) hat eine Nachweisgrenze von 0,001 ppm. Dabei gilt für SO_2 die Umrechnung: 1 ppm = 2,618 mg/m³. Die Ansprechzeit (T90) beträgt 180 Sekunden.

Vor 1990 erfolgten die Messungen mit einem automatisch nach der coulometrischen Meßmethode arbeitenden Gerät. Durch Eichungen wurde stets die Vergleichbarkeit der Meßwerte sichergestellt.

Datenausfälle sind vorwiegend durch Gerätedefekte entstanden. Aber auch bei den regelmäßigen Kalibrierungen des Meßgeräts kommt es zum Ausfall von 2 bis 3 Meßwerten. Die Reihe enthält dann an diesen Stellen Fehlwerte (missing values).□

Eine erhöhte Belastungssituation lag vom 1. bis 10.3.1989 vor. Da die in diesem Zeitabschnitt gemessene Konzentrationsreihe im weiteren sehr häufig zur Anwendung kommt, ist sie in Abb. 1-5 dargestellt und im Anhang F protokolliert worden. Zur leichteren Bezugnahme soll sie als CITY-Reihe bezeichnet werden.

Eine rein visuelle Einschätzung des Konzentrationsverlaufs läßt bereits seine wesentlichen Eigenschaften erkennen:

- Die SO_2-Konzentration variiert in irregulärer Weise, was auf den Einfluß von turbulenten Luftbewegungen hindeutet. Dieses quasi zufällige Verhalten unterscheidet den Immissionsverlauf von streng deterministischen Vorgängen und macht die Anwendung stochastischer Modellierungsverfahren erforderlich.

- Man kann optisch einen mittleren Konzentrationsverlauf ausmachen. Dieser ist offenbar am 6. und 7.3. höher, als an den anderen Tagen. Die Möglichkeit, einen solchen statistisch stabilen Mittelwert zu finden, beweist, daß die Immissionsreihe jedoch nicht absolut irregulär und zufällig ist.

- Darüber hinaus variiert die Konzentration nicht beliebig, sondern nur über einen bestimmten Wertebereich. Greift man aus der Meßreihe Werte heraus, die zeitlich hinreichend weit auseinanderliegen, so daß keine Korrelation mehr zwischen ihnen besteht, dann kann man dieses Kollektiv unabhängiger Werte auf seine Häufigkeitsverteilung untersuchen und bekommt damit eine Schätzung für die Wahrscheinlichkeitsverteilung der Meßwerte (vgl. Schlink (1990), (1993)). Es zeigt sich, daß Null eine natürliche und feste Grenze ist; negative Werte existieren nicht. Eine obere Schranke gibt es für die SO_2-Konzentrationen ebenfalls: Sie können zwar beliebig, aber nicht unendlich groß werden. Der Variations- oder Streubereich der Meßwerte verändert sich dabei im Laufe der Zeit. Beispielsweise sind der 6. und 7.3. auch durch eine vergrößerte Streubreite gekennzeichnet.

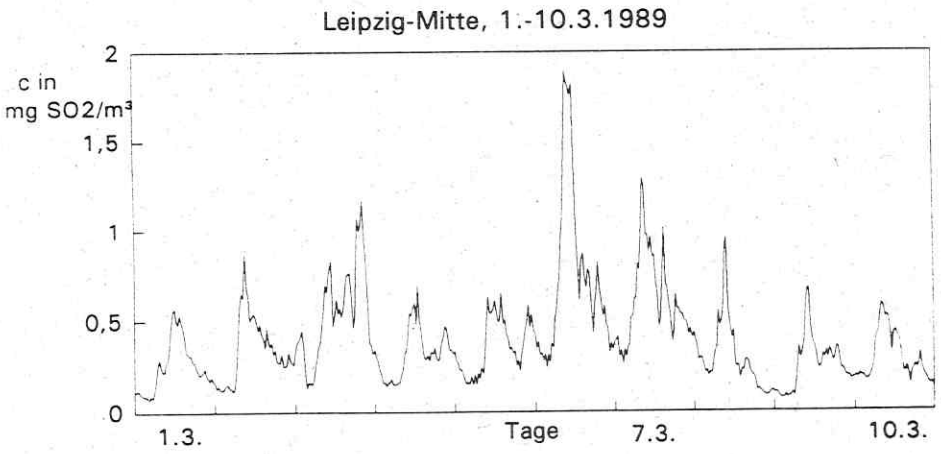


Abb. 1-5: Meßreihe CITY für Leipzig-Mitte während einer erhöhten Belastungssituation

- Die Variation der Immissionsreihe vollzieht sich in einem deutlich erkennbaren Rhythmus. Wenn man die Abstände zwischen den markanten Anstiegsflanken betrachtet, so läßt sich dieser als Tageszyklus identifizieren. Darüberhinaus könnte man vermuten, daß eine wöchentliche Periodizität vorliegt. Die genaue Untersuchung längerer Zeitabschnitte bestätigt diese Vermutung jedoch nicht: Im allgemeinen liegt keine Wochenperiode vor. Auf die in den Immissionsreihen enthaltenen periodischen Strukturen wird im Abschnitt 3 noch detaillierter eingegangen.

Typische Konzentrationsverläufe über einen Tag sind in Abb. 1-6 gesondert aufgezeichnet. Charakteristisch für ein Tagesprofil in der Heizperiode sind der morgendliche Anstieg ab ca. 8 Uhr sowie die oft auftretende zweite Konzentrationserhöhung am Abend. Beide Immissionserhöhungen haben ihre Ursachen in verstärkten Emissionen, die vorrangig während der "Anheizphasen" am Morgen und am Abend zustandekommen. Die Senkung des Immissionswertes am Nachmittag ist eine Folge der Erwärmung durch die Sonneneinstrahlung, die zu einem vertikalen Transport führt (vgl. Höschele (1966)).

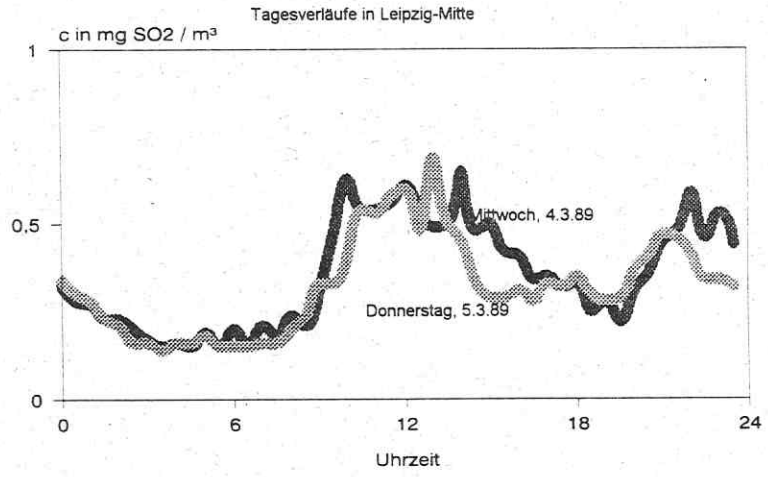


Abb. 1-6: Typische Tagesverläufe der SO_2-Konzentration

Die in Abb. 1-6 dargestellten Tagesverläufe kann man auch für mehrere aufeinanderfolgende Tage räumlich hintereinander aufzeichnen, so daß ein dreidimensionales Gebirge entsteht. Die Höhenlinien eines solchen Gebirges, die Isoplethen, sind in Abb. 1-7 widergegeben. Dabei wurde die Datenreihe CITY aus der Abb. 1-5 verwendet.

In dieser Darstellung sind durch die direkte Vergleichsmöglichkeit der Konzentrationsniveaus verschiedener Tage zur gleichen Tageszeit die Veränderungen des Tagesprofils besonders gut erkennbar.

Der alles überragende Peak am 6.3. wird ebenso deutlich, wie die zeitliche Verschiebung der Isoplethe 0,5 mg/m³ (durch "J" gekennzeichnet) von Tag zu Tag. Letzteres Phänomen bedeutet eine Verschiebung der Phasenlage der Anstiegsflanke. Es könnte aber auch durch eine zur Tagesrhythmik hinzutretende langsame Veränderung eines "Grundlevels" oder "Grundniveaus" entstehen. Im Abschnitt 4 des 2. Teils wird diese Eigenschaft im Komponentenmodell abgebildet.

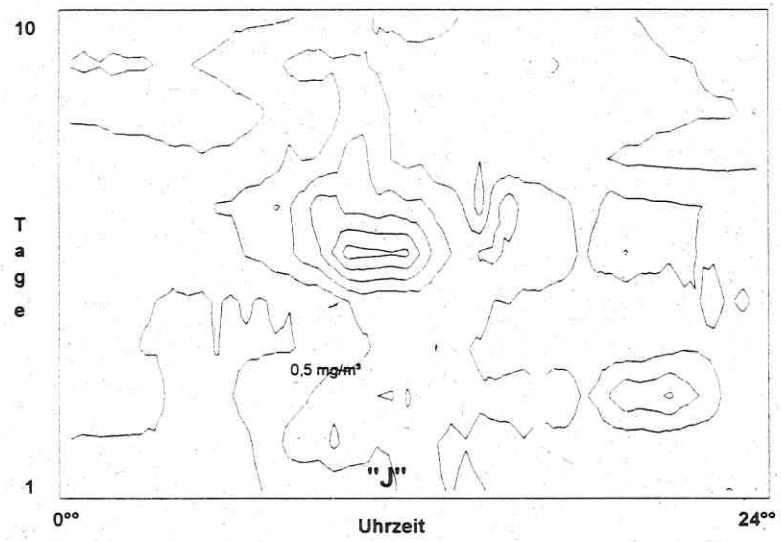


Abb. 1-7: Isolinien der SO_2-Konzentration zur Darstellung des täglichen Ganges und des Ganges über 10 Tage; 1.-10.3.89; Leipzig-Mitte;

Bisher wurde ein visueller Eindruck von den Wesenszügen der Immissionsreihen gewonnen. In den folgenden Abschnitten soll nun mit Hilfe des Periodogramms und der Autokorrelationsfunktion eine exaktere Charakterisierung der Meßdaten erfolgen.

Zuvor muß jedoch noch eine für die Zeitreihenanalyse wichtige Voraussetzung geschaffen werden: Die Stationarität der Meßreihe.

Bei der statistischen Bearbeitung von Zeitreihen wird nämlich stets davon ausgegangen, daß es sich um Realisierungen stochastischer Prozesse handelt. Die für die Bearbeitung notwendigen theoretischen Beziehungen lassen sich im allgemeinen nur unter der Voraussetzung stationärer stochastischer Prozesse ableiten, so daß die Berechnung entsprechender statistischer Kenngrößen ausschließlich für stationäre Zeitreihen sinnvoll ist. Was man unter Stationarität versteht, ist im Anhang A (speziell Definition (A-5)) skizziert. Bei der Zeitreihenmodellierung ist es in Spezialfällen möglich, auch instationäres Verhalten in das Modell einzubeziehen. Darauf soll jedoch später an den geeigneten Stellen eingegangen werden.

2. Box - Cox -Transformation

Im vorhergehenden Abschnitt ist auf S. 20 + 21 das Verhalten einer Immissionsreihe verbal charakterisiert worden. Dabei konnten ein mittlerer Konzentrationsverlauf und eine Streubreite zur Beschreibung verwendet werden. Da für beide Charakteristika eine Zeitabhängigkeit zu erkennen war, handelt es sich also um eine Mittelwert**funktion** m_t der Zeit. Auch die Standardabweichung s_t ist eine **Zeitfunktion**, deren Argumente die diskreten Zeitpunkte t = ..., -1, 0, 1, 2, ... sind.

Bemerkung: Über diese heuristische Definition soll hier nicht hinausgegangen werden. Eine exakte Formulierung müßte auf einem zugrundeliegenden stochastischen Prozeß aufbauen und viele Realisierungen ν dieses Prozesses zu einem Ensemble zusammenfassen. m_t wäre dann das zu jedem Zeitpunkt gebildete Ensemblemittel

$$m_t = \frac{1}{\nu} \sum_{\substack{\text{alle } \nu \text{ Reali-} \\ \text{sierungen} \\ \text{des Ensembles}}} z_t$$

und s_t die dazugehörige Standardabweichung.

$$s_t = \sqrt{\frac{1}{\nu-1} \sum_{\substack{\text{alle } \nu \text{ Reali-} \\ \text{sierungen} \\ \text{des Ensembles}}} (z_t - m_t)^2}$$

m_t und s_t sind Schätzungen für die Erwartungswertfunktion μ_t und die Varianzfunktion σ_t des stochastischen Prozesses. Das Ensemblemittel ist sehr genau von der Mittelung über die Zeit bei der Betrachtung einer einzigen Realisierung (zeitliches Mittel) zu unterscheiden.□

Für die statistische Auswertung von Meßreihen muß im allgemeinen die Stationarität des zugrundeliegenden stochastischen Prozesses vorausgesetzt werden (zur Stationarität vgl. Anhang A). In den für die Praxis relevanten Fällen kann man sich jedoch mit schwacher Stationarität begnügen, was bedeutet, daß der Erwartungswert und die Varianz des stochastischen Prozesses zeit**un**abhängig sind, die Meßreihe also einen zeitlich konstanten Mittelwert m_t = const. und eine zeitlich konstante Streuung s_t = const. aufweist.

Bei den hier vorliegenden Daten variiert der Meßwert umso mehr, je größer er ist. Dies läßt sich bereits am Verlaufsbild der Immissionsreihe CITY (Abb. 1-8) visuell erkennen. Die Maxima und Minima der Kurve sind unterschiedlich ausgeprägt: Ein hoher Immissionswert ist auch mit stärkeren Fluktuationen verbunden, als ein niedriger.

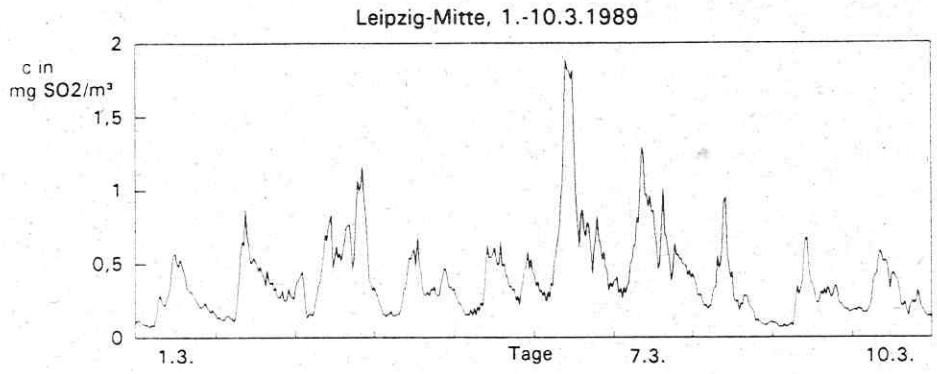


Abb. 1-8: CITY-Meßreihe

Zur Beseitigung dieser zeitlichen Veränderung der Streuung geht man von folgender Überlegung aus: Wenn allgemein zwischen der Varianz der Meßwerte $var(z_t) = \sigma_t^2$ und deren Erwartungswert μ_t der Zusammenhang

$$var(z_t) = \mu_t^c, \qquad (1\text{-}1)$$

mit einer beliebigen Konstanten c, besteht, dann läßt sich durch eine nichtlineare Transformation der z_t eine Stabilisierung der Varianz erreichen. Die verbreitetste ist die logarithmische Transformation. Eine umfassendere Klasse von Transformationen für positive Meßwerte sind die Box-Cox-Transformationen (Box & Cox (1964)) mit dem Transformationsparameter θ:

$$u_t = F(z_t) = \begin{cases} (z_t^{\theta} - 1) / \theta & \text{für} \quad \theta \neq 0 \\ \ln(z_t) & \text{für} \quad \theta = 0 \end{cases} \qquad (1\text{-}2)$$

Die Festlegung von $u_t = \ln(z_t)$ für $\theta = 0$ ist wegen $\lim ((z^{\theta} - 1) / \theta) \rightarrow \ln(z)$ für $\theta \rightarrow 0$ auch sinnvoll. Eine Methode zur Schätzung von θ wird von Box & Cox (1964) angegeben.

Man kann θ aber auch näherungsweise aus der Bedingung $var(u_t) = const.$ bestimmen. Die Annahme (1-1) lautet für die Schätzwerte s_t und m_t: $s_t^2 = var(z_t) = m_t^c$ (Schlittgen & Streitberg (1991)). Für die Box-Cox-Transformation $u = F(z)$ führt dies auf die Beziehung:

$$\begin{aligned} var(u) &= var(F(z)) \\ &= (du/dz)^2 \,|_m \, var(z) \\ &= (m_t^{\theta-1})^2 \, m_t^c \\ &= const. = k^2 \end{aligned} \qquad (1\text{-}3)$$

Wegen $m_t^c = s_t^2$ folgt somit

$$s_t = k\ m_t^{1-\theta}\ , \qquad (1\text{-}4)$$

was sich schreiben läßt als

$$\ln s_t = \ln k + (1-\theta) \ln m_t. \qquad (1\text{-}5)$$

Die Näherung besteht nun darin, daß man die Ensemblemittel m_t und s_t durch Zeitmittelung schätzt. Man bildet also aus gleichlangen Abschnitten der gemessenen Reihe jeweils Mittelwert und Standardabweichung. Für diese muß sich bei Auftragung in doppelt-logarithmischem Maßstab der Parameter θ aus dem Anstieg bestimmen lassen. Fehlwerte dürfen natürlich nicht mit in die Rechnung eingehen. Allerdings ist mit dieser Methode lediglich eine grobe Abschätzung möglich. Man kann damit nur entscheiden, ob θ = -0.5, 0, 0.5, 1 oder 2 gewählt werden sollte. Schließlich ist auf die Prüfung der Voraussetzung für die benutzte Näherung, nämlich die Ersetzbarkeit der Ensemblemittelung durch die Zeitmittelung, die als Ergodizität der Zeitreihe bezeichnet wird, verzichtet worden.

Für eine solche Abschätzung wurde die Meßreihe eines Jahres (Leipzig-Mitte 1989) in Abschnitte von jeweils zwei Tagen zerlegt und für diese Abschnitte die Logarithmen von Mittelwert m und Standardabweichung s bestimmt. Eine lineare Regression führte auf einen Anstieg von $\approx 0{,}45 = 1 - \theta$ (vgl. Abb. 1-9).

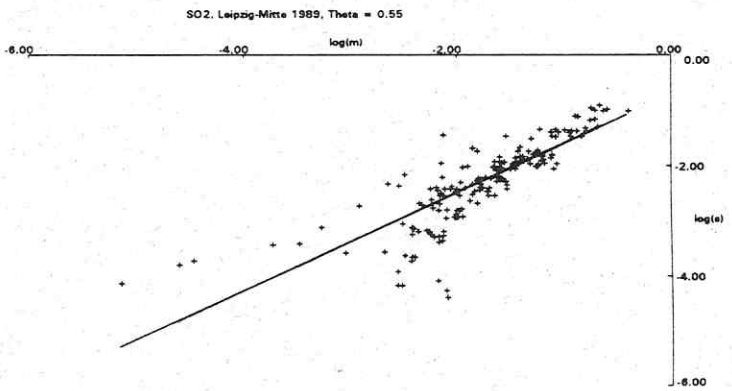


Abb. 1-9: Logarithmus der Standardabweichungen s und Logarithmus der Mittelwerte m aus Abschnitten von jeweils 96 Meßwerten der Meßreihe von Leipzig-Mitte 1989 und deren Regressionsgerade.

Der Transformationsparameter liegt damit bei $\theta \approx 0{,}55$. Für die weitere Verarbeitung wurde deshalb in ausreichender Näherung $\theta = 1/2$ gesetzt.

Die Anwendung der Box-Cox-Transformation auf die Datenreihe CITY in Abb. 1-8 ergibt den in Abb. 1-10 gezeigten Verlauf. Ein visueller Vergleich der Konturen der Maxima und Minima läßt die Veränderungen im Verhalten der Varianz erkennen.

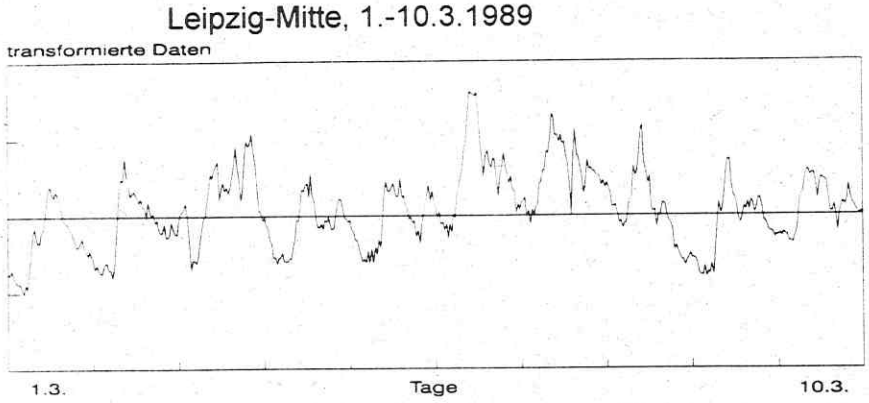


Abb. 1-10: Transformierte CITY Daten der Abb. 1-8

Die Überprüfung der Transformationsparameter bei Meßreihen anderer Jahre ergab ähnliche Resultate, die Tab. 1-2 zu entnehmen sind.

Tabelle 1-2: Transformationsparameter θ am Beispiel verschiedener SO_2-Jahresreihen der Station Leipzig-Mitte

Jahr	1985	1987	1989	1991	1993
θ	0,59	0,60	0,55	0,67	0,55

Für die weiteren Analysen ist außer einer stabilen Varianz auch noch ein stationärer, d.h., zeitlich konstanter Erwartungswert erforderlich. Um dies zu erreichen, wurde im Teil 1 stets von einem linearen Trendanteil ausgegangen und dieser eliminiert. In den folgenden Abschnitten fand dies neben der Box-Cox-Transformation stets Berücksichtigung, auch wenn es nicht an allen Stellen explizit erwähnt ist.

Im Ergebnis der Untersuchungen des 1. Teils wird die Erkenntnis stehen, daß gerade die niederfrequenten Anteile der Reihe die Smogepisoden charakterisieren. Bei der Modellierung im Teil 2 wird der Begriff "Trend" deshalb weiter gefaßt: Er beinhaltet nicht nur lineare sondern auch niederfrequente Anteile. Dieser Trendbegriff, der über die Linearität hinausgeht, ist dann die Basis für die Modellierung der Immissionsreihen und die Prognose von Smogepisoden. Mit der Nichtlinearität sind Lösungsschwierigkeiten verbunden, die im Teil 3 besprochen werden.

Im Laufe der hier geführten Untersuchungen kristallisiert sich somit gerade die zeitliche Veränderung des Erwartungswertes, also der Trend der Meßreihe als wesentlicher Ausgangspunkt für eine Immissionsprognose heraus.

3. Spektraluntersuchungen

Die Meßdaten sollen zunächst mit Hilfe der Fourieranalyse systematisch auf das Vorhandensein periodischer Strukturen untersucht werden. Ausgangspunkt ist wieder die in Abb. 1-11 gezeigte CITY-Reihe. Sie enthält 480 Halbstundenmittelwerte.

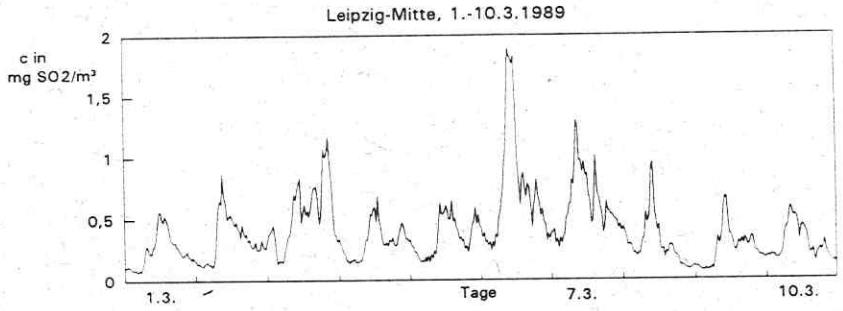


Abb. 1-11: SO_2-Konzentrationsreihe CITY für Leipzig-Mitte

Das Frequenzspektrum einer Reihe wird durch ihr Periodogramm geschätzt. Die Grundlagen und Formeln dazu sind im Anhang A aufgeschrieben. Abb. 1-12 zeigt das Periodogramm der Daten von Abb. 1-11, wobei zuvor lineare Trendanteile entfernt wurden.

Der Beobachtungsabstand der Meßwerte beträgt 0,5 h, so daß die Frequenzen in $1/(0{,}5h) = 2 \cdot h^{-1}$ angegeben sind.

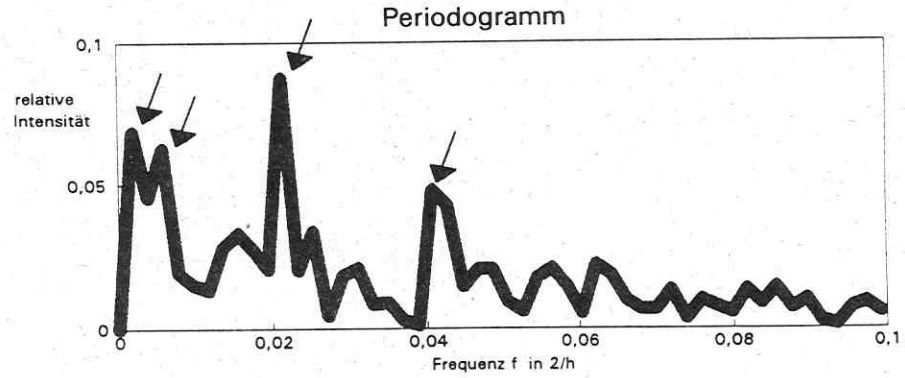


Abb. 1-12: Periodogramm der CITY-Daten von Abb. 1-11. Periode=1/f. Markante Peaks liegen bei f ≈ 0.0019, 0.0056, 0.02, 0.04 vor (durch Pfeile markiert). Dies entspricht Perioden von 263, 89, 24 und 12 Stunden.

Die Peaks im Periodogramm lassen sich als Intensitäten (oder Energien) interpretieren, mit denen die einzelnen Frequenzen besetzt sind. Da hier **Halbstunden**mittelwerte verwendet werden, charakterisiert z.B. eine Periodizität von 48 Meßwerten den 24-Stunden-Rhythmus. Je nachdem, ob hohe oder tiefe Frequenzen vorherrschen, ist die Reihe rauh oder glatt.

Bemerkung (1-2): Durch die bei der Periodogrammberechnung verwendete Fouriertransformation wird die Meßreihe eineindeutig vom Zeit- in den Frequenzbereich transformiert. Das Periodogramm trifft dann eine Aussage über die Glattheit einer Meßreihe. Analog dazu wird diese Aussage im Zeitbereich durch die Autokorrelationsfunktion (vgl. Abschnitt 4) gemacht. Hier ist die Glattheit mit der Durbin-Watson-Statistik meßbar (Box & Jenkins (1976)). Das Periodogramm kann man auch als Fourier-Transformierte der Autokorrelationsfunktion darstellen (Schlittgen & Streitberg (1991)).□

Das Periodogramm läßt sich aber auch bereits mit einer einfachen visuellen Methode auf die Existenz periodischer Strukturen untersuchen: Man verwendet dazu das kumulative Periodogramm, welches eine Schätzung der Spektralverteilungsfunktion ist (Abb. 1-13). Da die theoretische Spektraldichte für einen unstrukturierten Rauschprozeß konstant ist, ist ihre Verteilung und dieser entsprechend das kumulative Periodogramm durch einen konstanten linearen Anstieg bestimmt. Abweichungen von dieser Linie zeigen Abweichungen vom weißen Rauschen deutlicher als das Periodogramm selbst an. Darauf baut ein Test von Durbin (1969) über das Vorliegen von weißem Rauschen auf.

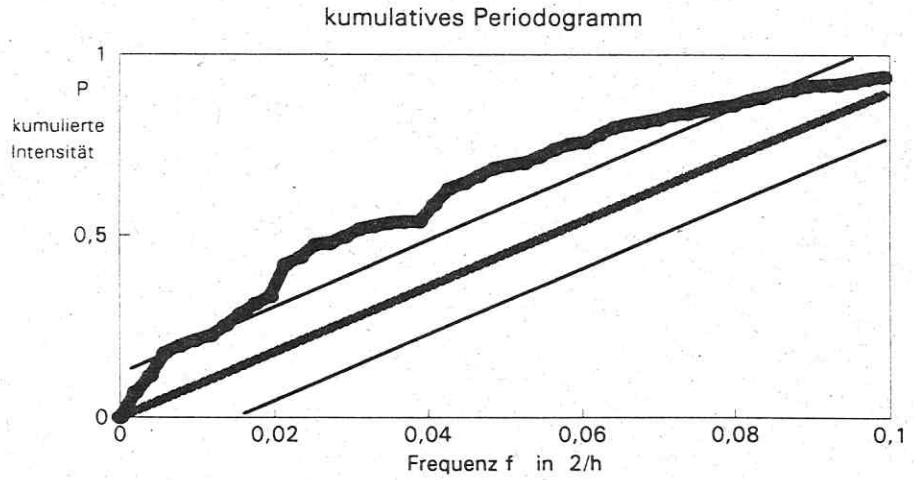


Abb. 1-13: Kumulative Darstellung des Periodogramms aus Abb. 1-12. Die Gerade wäre der Verlauf des kumulativen Periodogramms bei weißem Rauschen. Die Parallelen kennzeichnen das 95%-Konfidenzintervall.

Das kumulative Periodogramm (Abb. 1-13) weist darauf hin, daß es sich bei den Daten der Abb. 1-11 nicht um weißes Rauschen handelt, sondern daß sie periodische Strukturen enthalten.

Der Vertrauensbereich c der für das weiße Rauschen gültigen Geraden P = const. · f ± c wird nach der von Stephens (1970) angegebenen (empirisch gefundenen) Formel

$$c = \frac{\sqrt{-\frac{1}{2}\ln\frac{\alpha}{2}}}{\sqrt{M-1}+0.2+\frac{0.68}{\sqrt{M-1}}} - \frac{0.4}{M-1}$$

mit M = int(N/2) und der statistischen Sicherheit von (1-α) berechnet.

Für die Fourieranalyse sind sinusförmige Verläufe am geeignetsten. Um diese zu bekommen, kann man die Meßdaten transformieren. Die dafür verwendete Box-Cox-Transformation wurde bereits im Abschnitt 2 diskutiert.

Prinzipiell sollte vor einer Spektralanalyse auch der (lineare) Trend entfernt werden, damit andere langperiodische Effekte nicht verdeckt werden.

Das von den transformierten Daten ermittelte Periodogramm zeigt Abb. 1-14.

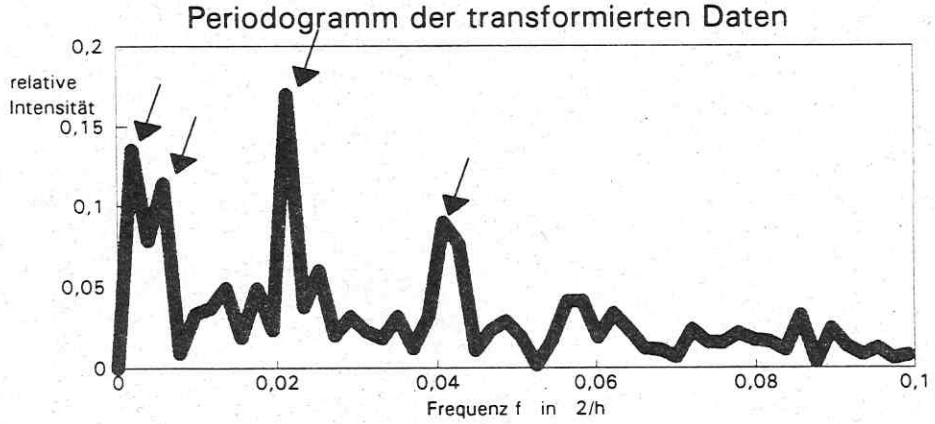


Abb. 1-14: Periodogramm der transformierten und trendbereinigten CITY-Reihe; Erläuterungen siehe Abb. 1-12.

Das Periodogramm in Abb. 1-14 läßt im Vergleich zu Abb. 1-12 die dominanten Perioden deutlicher hervortreten. Ein prinzipiell anderes Verhalten zeigt sich jedoch nicht.

Bei einer Spektralanalyse sind zwei weitere Tatsachen zu beachten:

1) Da nur Perioden eingehen, die ganz in die Gesamtlänge N der Zeitreihe hineinpassen, gibt es eine Diskretisierung der Frequenzachse: $f_i=i/N$. So kann es vorkommen, daß eine wesentliche Frequenz der Reihe mit keiner der Fourierfrequenzen f_i übereinstimmt. Sie würde sich dann in den nächstgelegenen Fourierfrequenzen bemerkbar machen. Das vom Periodogramm gezeigte Bild enthält dann einen Peak endlicher Breite. Dieser durch die Endlichkeit der untersuchten Reihe bewirkte Effekt wird als "Durchsickern" (Leakage) bezeichnet.

2) Die Fourieranalyse schätzt aus N Datenpunkten N Koeffizienten. Jeder Wert des Periodogramms baut auf nur 2 Koeffizienten auf. Zwar ist er von jedem anderen Punkt des Periodogramms unabhängig, dafür besitzt er jedoch eine sehr große Varianz. Einzelne Punkte im Periodogramm enthalten also einen hohen Rauschanteil und sind somit sehr unsicher (Schlittgen & Streitberg (1991)).

Deshalb wird das Periodogramm in der Regel geglättet. Dies verkleinert die Varianz der Werte (auf Kosten von Auflösung und Bias). Das geglättete Periodogramm ist ein Schätzer für die Spektraldichte (vgl. Anhang A).

Die Glättungsprozeduren bei Periodogrammen heißen "Fenster" (engl.: windows). Sie mitteln symmetrisch die in einem Fenster festgelegter Breite (Faustregel: ca. 10...20% der Datenbreite) enthaltenen Werte mit festgelegten Gewichten. Es ist also für eine Glättung ein Fenster**typ** sowie die (ungerade) **Breite** zu wählen.

Als Beispiel wurde das Periodogramm aus Abb. 1-14 mit einem Tukey-Hamming(5)-Fenster geglättet und in Abb. 1-15 dargestellt. Die Gewichte berechnen sich beim Tukey-Hamming-Fenster zu $w_j=0.54+0.46*\cos(\pi j/p)$, wobei $j=0,1,2,...,p$ und $p=int(span/2)$, hier mit der Fensterbreite $span=5$ ist (SPSS Manual (1987)). Anschließend werden die Gewichte auf die Summe 1 normiert.

Das geglättete Periodogramm (Abb. 1-15) zeigt jetzt nur noch drei wesentliche Perioden an: 135 Stunden, 24 Stunden und 12 Stunden, wobei letztere offenbar eine Oberwelle der Tagesperiode ist. Der in Abb. 1-12 ersichtliche Doppelpeak ist also kein Charakteristikum der CITY-Reihe.

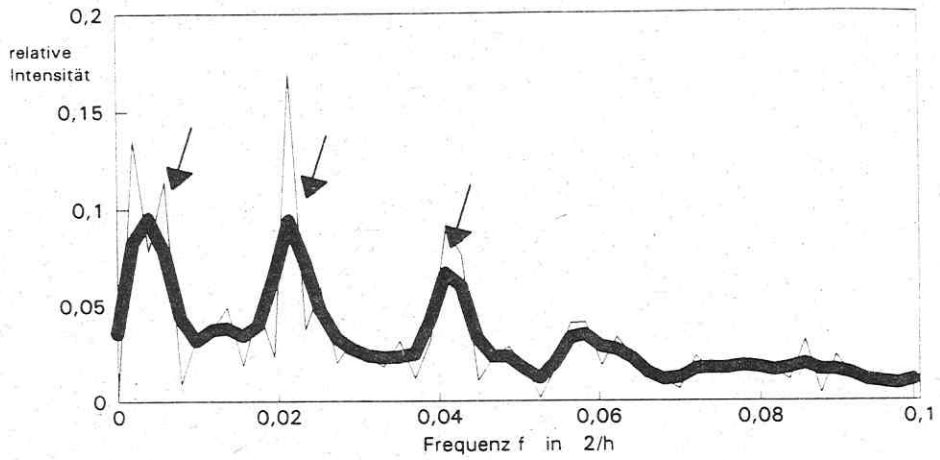


Abb. 1-15: Periodogramm aus Abb. 1-14 (dünne Linie) mit einem Tukey-Hamming(5)-Fenster geglättet (dicke Linie).

Enthält eine Zeitreihe einen sehr dominanten Zyklus, so ist es sinnvoll, durch Differenzbildung eine Saisonbereinigung vorzunehmen. Dadurch können im Periodogramm Verdeckungserscheinungen durch den dominanten Peak vermieden werden und andere Strukturen deutlicher hervortreten.

Das Vorliegen eines 24-Stunden-Rhythmus (Periode 48) ist bei der SO_2-Immission trivial, da sowohl die Emission, als auch die meteorologischen Einflüsse dem Tageszyklus unterworfen sind. Deshalb soll im folgenden mittels Differenzbildung über 48 Meßwerte eine Saisonbereinigung vorgenommen werden. Die Differenzenreihe ist in Abb. 1-16 widergegeben.

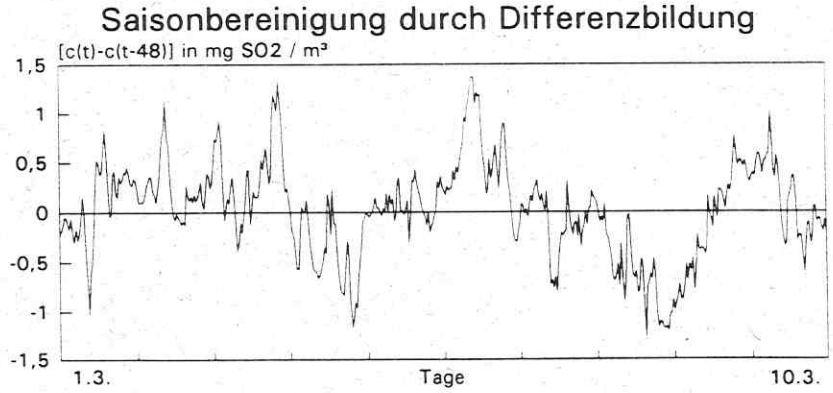


Abb. 1-16: Durch Differenzbildung über 48 Werte saisonbereinigte CITY-Daten

Daraus erhält man das in Abb. 1-17 gezeigte Periodogramm.

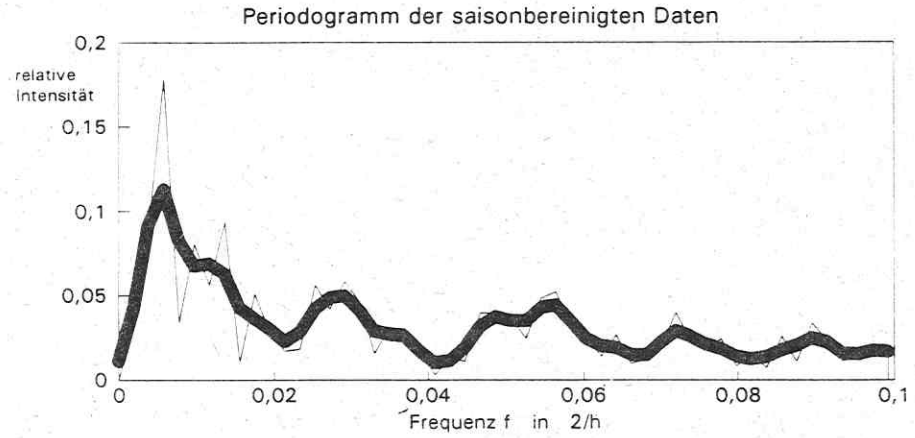


Abb. 1-17: Periodogramm (dünne Linie) der Daten von Abb. 1-16 und mit Tukey-Hamming(5)-Fenster geglättetes Periodogramm (dicke Linie).

Die 24stündige und die 12stündige Periodizität sind verschwunden. Jetzt hebt sich lediglich die Periode von 135 Stunden (f=0.0037) deutlich heraus. Weitere Zyklen lassen sich nicht entdecken.

Bisher wurden die Periodogramme nur im Bereich $f < 0.1$, d.h. Periode $> 1/f =$ 10 Meßwerte = 5 Stunden, also für die langwelligen Zyklen aufgezeichnet.

Betrachtet man das vollständige und das kumulative (normierte) Periodogramm in Abb. 1-18, so zeigt sich aber, daß in der Differenzenreihe (Abb. 1-16) verstärkt langwellige Schwingungen enthalten sind ($f<0.1$, d.h. Periode > 5 Stunden), die bereits in Abb. 1-17 genauer bestimmt wurden.

In diesem Frequenzbereich von 0 ... 0.1 treten also offenbar die wesentlichsten Unterschiede zum weißen Rauschen auf. Somit ist es berechtigt, die Spektraluntersuchungen auf diesen Bereich zu konzentrieren.

Für die hier betrachtete CITY-Reihe wurde neben der Tagesschwingung ein dominanter Zyklus von 135 Std. $\approx$ 5,6 Tage gefunden. Er beschreibt offenbar den Ablauf der hier untersuchten Episode. Ist dies ein charakteristischer Zyklus für Smogepisoden?

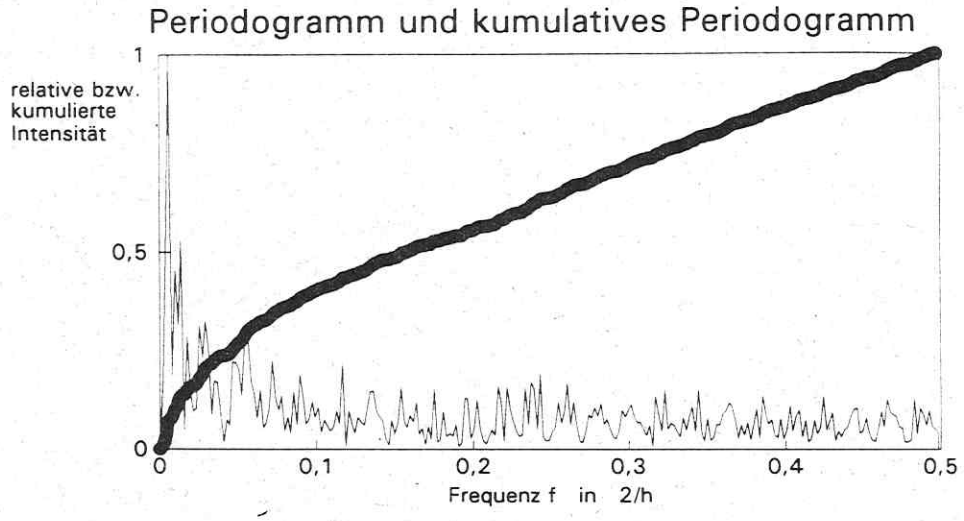


Abb. 1-18: Periodogramm (dünne Linie) und kumulatives Periodogramm (dicke Linie) der Daten von Abb. 1-16

Zur Aufklärung dieser Frage wurde eine Vielzahl von Abschnitten der Datenreihe von Leipzig-Mitte (1980 bis 1993) untersucht, was zu folgendem Ergebnis führte:

Die dominanten Periodizitäten verändern sich im Laufe der Zeit. Zur Beschreibung des Verlaufs von Immissionsdaten ist ein festes Frequenzspektrum nicht ausreichend. Vielmehr muß man von einer Zeitabhängigkeit der einzelnen Fourierkoeffizienten und damit von einem zeitlich veränderlichen Spektrum ausgehen.

Um dieses Verhalten anschaulich zu machen, wird zu einem Zeitabschnitt fester Breite (10 Tage) das Periodogramm sowie das saisonbereinigte Periodogramm in Abb. 1-19 widergegeben. Durch kontinuierliches Weiterrücken dieses Ausschnitts um jeweils einen Tag lassen sich die Veränderungen im Periodogramm dann gut beobachten.

Beispielsweise kann man aus Abb. 1-19 erkennen, wie sich im Verlaufe des Weiterrückens des Beobachtungsfensters vom 26.2. bis zum 1.3. die Peaks im Periodogramm verändern:

Die dominante langwellige Periode (α) von ca. 64 Stunden (2.7 Tage) wird intensiver. Dabei verkleinert sich ihre Frequenz, so daß sich die Periode auf 135 Stunden (5.6 Tage) verlängert (β). Als triviale Periode (γ) ist die 24-h-Schwingung ($f \approx 0.02$) in unterschiedlich starkem Maße stets enthalten. Die 12-h-Periode (δ) zeigt sich bei $f \approx 0.041$.

Durch die saisonbereinigende Differenzbildung wurden die 24stündige Periode und ihre Harmonische entfernt. Dies ist aus dem Vergleich von mittlerem und oberem Periodogramm gut ablesbar. Vermutlich machen sich auch die Harmonischen der Perioden > 24h bei Perioden < 24h bemerkbar. Die Schwingungen mit Periodenlängen über 24 Stunden sind deshalb bei den weiteren Untersuchungen von speziellem Interesse.

Mit der Länge des Beobachtungsfensters ist natürlich auch die Länge der größten meßbaren Periode sowie der Stützstellenabstand im Periodogramm festgelegt. Man muß also mit einer für den Anwendungszweck geeigneten Fensterbreite arbeiten. Da Smogepisoden in der Regel 2 bis 7 Tage andauern, wurde hier für die Beobachtung jeweils eine Dekade herangezogen.

Das so für einen Zeitabschnitt von 10 Tagen berechnete Periodogramm läßt die langfristigen Veränderungen der spektralen Anteile besonders gut erkennen, wenn man es für aufeinanderfolgende Dekaden räumlich hintereinander aufzeichnet. Über der so gebildeten Frequenz - Zeit - Ebene erhebt sich dann eine Gebirgsfläche, deren Höhe die Intensität einer bestimmten Frequenz während einer bestimmten Dekade repräsentiert.

Die Höhenliniendarstellung eines solchen Spektrogramms ist in Abb. 1-20 aufgezeichnet.

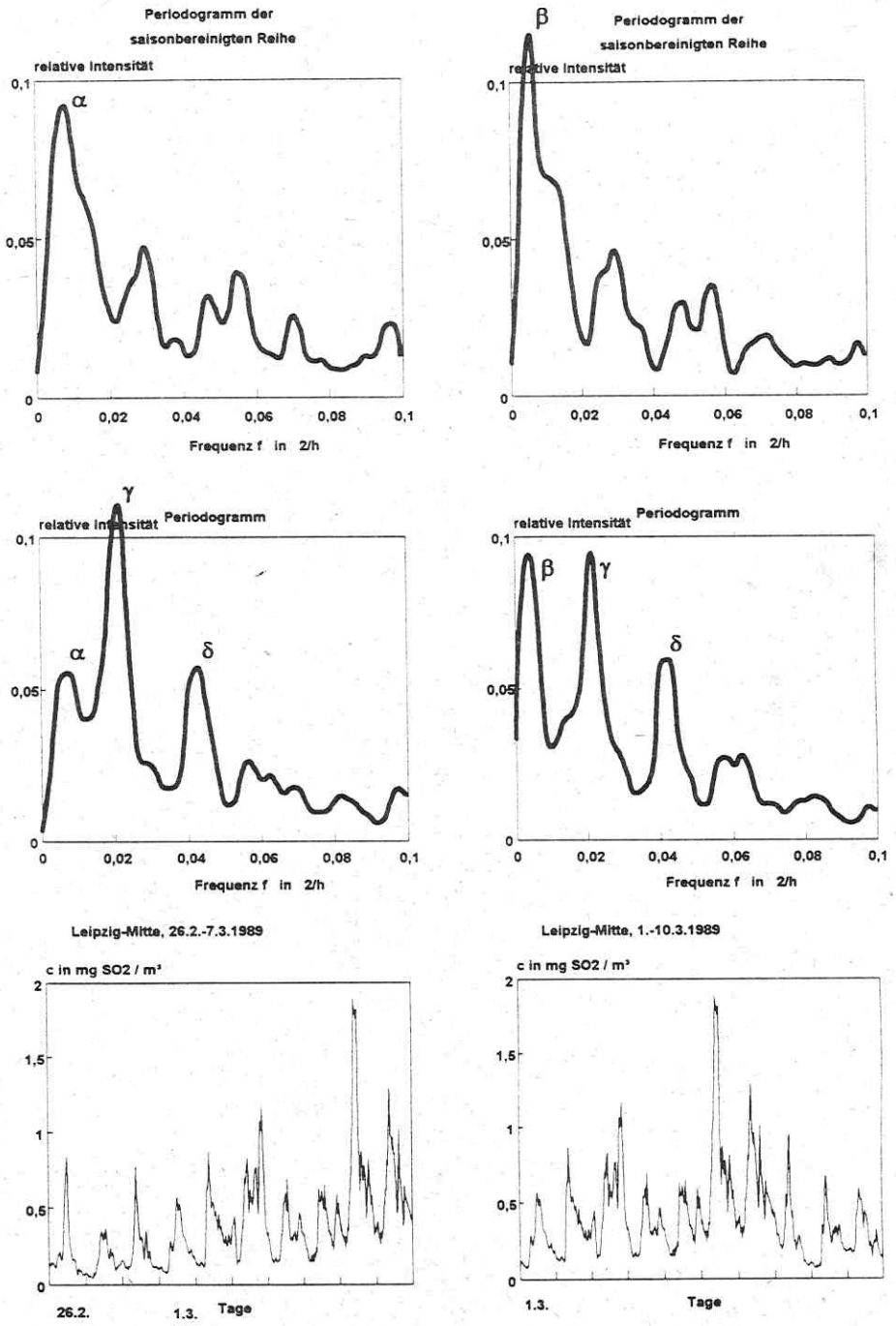


Abb. 1-19: Verschiedene Abschnitte der SO_2-Immissionsreihe von Leipzig-Mitte, 1989 mit dazugehörigem Periodogramm (unten: Zeitreihe z_t; mitte: aus der unten dargestellten Dekade berechnetes Periodogramm I(f); oben: Periodogramm I(f) der saisonbereinigten Reihe)

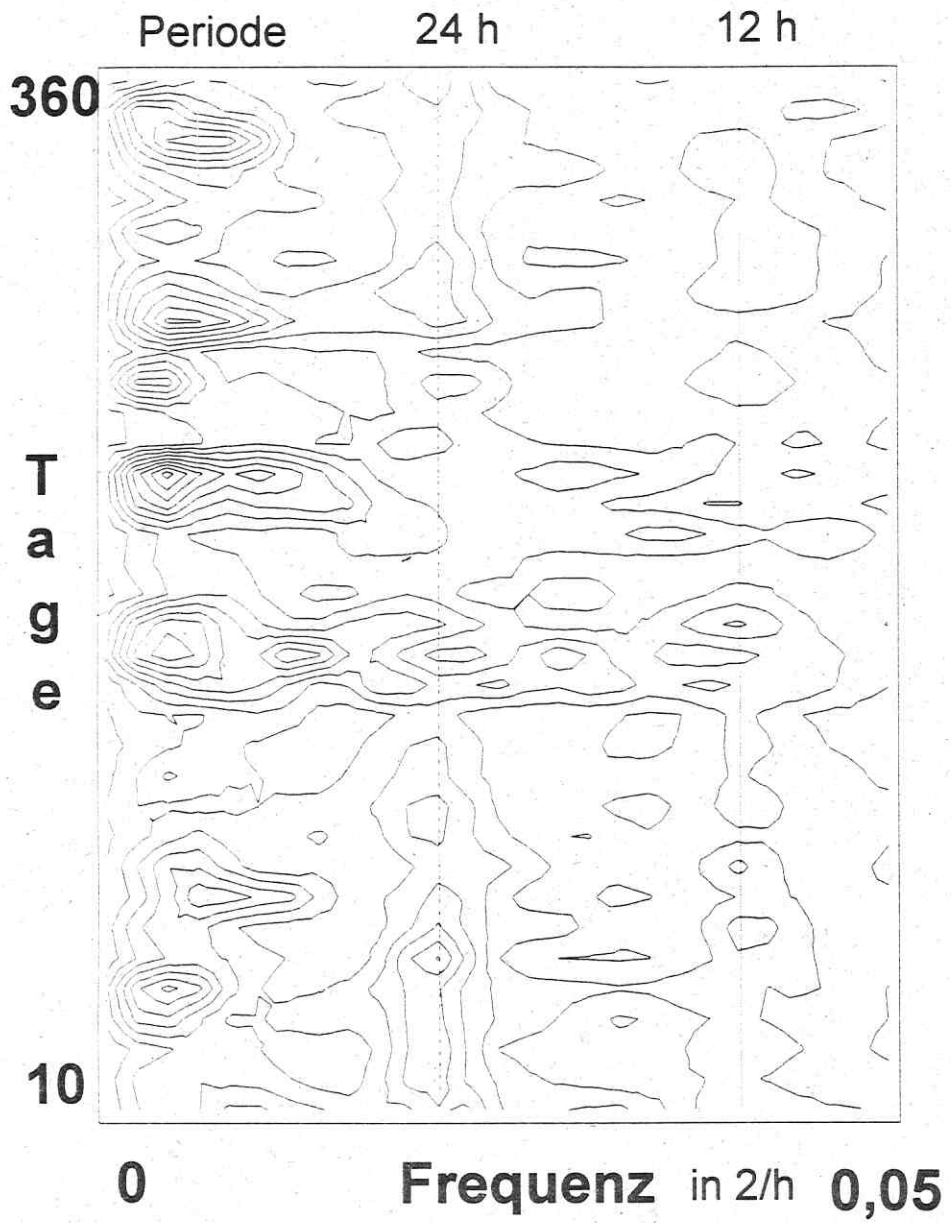


Abb. 1-20: Höhenlinien des Spektrogramms für die Daten von Leipzig-Mitte für das gesamte Jahr 1989; die einzelnen Periodogramme wurden jeweils aus Meßreihenabschnitten von 10 Tagen nach Box-Cox-Transformation und Trendelimination berechnet und anschließend mit einem Tukey-Hamming(5)-Fenster geglättet.

Die Frequenz 0,05 entspricht einer Periode von 10 Stunden. Somit repräsentieren die Gipfel des Spektrogramms unmittelbar links von dieser Frequenz die Periode von 12 Stunden (gepunktete Linie). Die durch eine Strich-Punkt-Linie gekennzeichneten Gipfel entsprechen der 24 Stunden Periode.

Im Gegensatz zu diesen Periodizitäten, die zwar in ihrer Intensität einen Jahresgang aufweisen, in ihrer Frequenz jedoch stabil bleiben, gibt es auch lokale Maxima bei sehr niedrigen Frequenzen unterschiedlicher Größe.

Diese zeigen sich im Spektrogramm (Abb. 1-20) als stark lokalisierte Gebilde, die einander ablösen. Sie sind durch Periodenlängen von über einem Tag charakterisiert. Wie lassen sich diese Vorgänge interpretieren? Wodurch könnten sie kausal bedingt sein und was für ein phänomenologisches Verhalten ist zu erwarten?

Beispielsweise dauern Smogepisoden etwa 2 bis 7 Tage an (vgl. Einleitung). Somit könnte man die niederfrequenten Vorgänge in der Meßreihe mit den Episoden identifizieren. Eine solche Interpretation stößt jedoch auf Widerspruch, denn die niederfrequenten Erscheinungen treten auch im Sommer auf (vgl. Abb. 1-20), Smogepisoden des Schwefeldioxids jedoch nicht.

Letztere sind an eine bestimmte Höhe des SO_2-Niveaus, wie sie nur in der Heizperiode infolge verstärkter Emissionen zustande kommen kann, gebunden. Bei der Periodogrammberechnung geht die Information über dieses Niveau jedoch durch die nichtlineare varianzstabilisierende Transformation und die Trendelimination verloren. Das Spektrogramm bildet also lediglich die Dynamik der Immissionsreihe ab. Setzt man eine definierte Emissionsstärke voraus, so wird die Dynamik der Smogepisoden jedoch in erster Linie durch das Wettergeschehen bestimmt.

Die durch das Zusammenspiel verschiedener Luftmassen bestimmten Wetterabläufe weisen im allgemeinen keine feste Periodizität auf. Sie sind offenbar für die niederfrequenten Vorgänge mit sich ändernder Frequenz in der Immissionsreihe verantwortlich.

Der erwähnte Einfluß der Transformation der Daten auf das Spektrogramm wird im Vergleich von Abb. 1-20 mit Abb. 1-21 deutlich. Zur Erstellung von Abb. 1-21 wurden die Meßdaten nicht mit Trendelimination und Transformation vorbehandelt. In der Nichtheizperiode ist die Amplitude besonders der niederfrequenten Schwingungen der SO_2-Konzentration dann so gering, daß sie im Spektrogramm nicht mehr hervortreten (Abb. 1-21). Erst durch eine nichtlineare Box-Cox-Transformation werden sie registrierbar (Abb. 1-20).

Bemerkung (1-3): Da Änderungen von Schadstoffkonzentrationen wesentlich durch den Transport von Luft und damit durch den Wind hervorgerufen werden, ist ein Vergleich mit dem Frequenzspektrum der Windgeschwindigkeit (Stull (1989)) interessant. Das Windspektrum enthält drei typische Bereiche (vgl. Abb. 1-22):

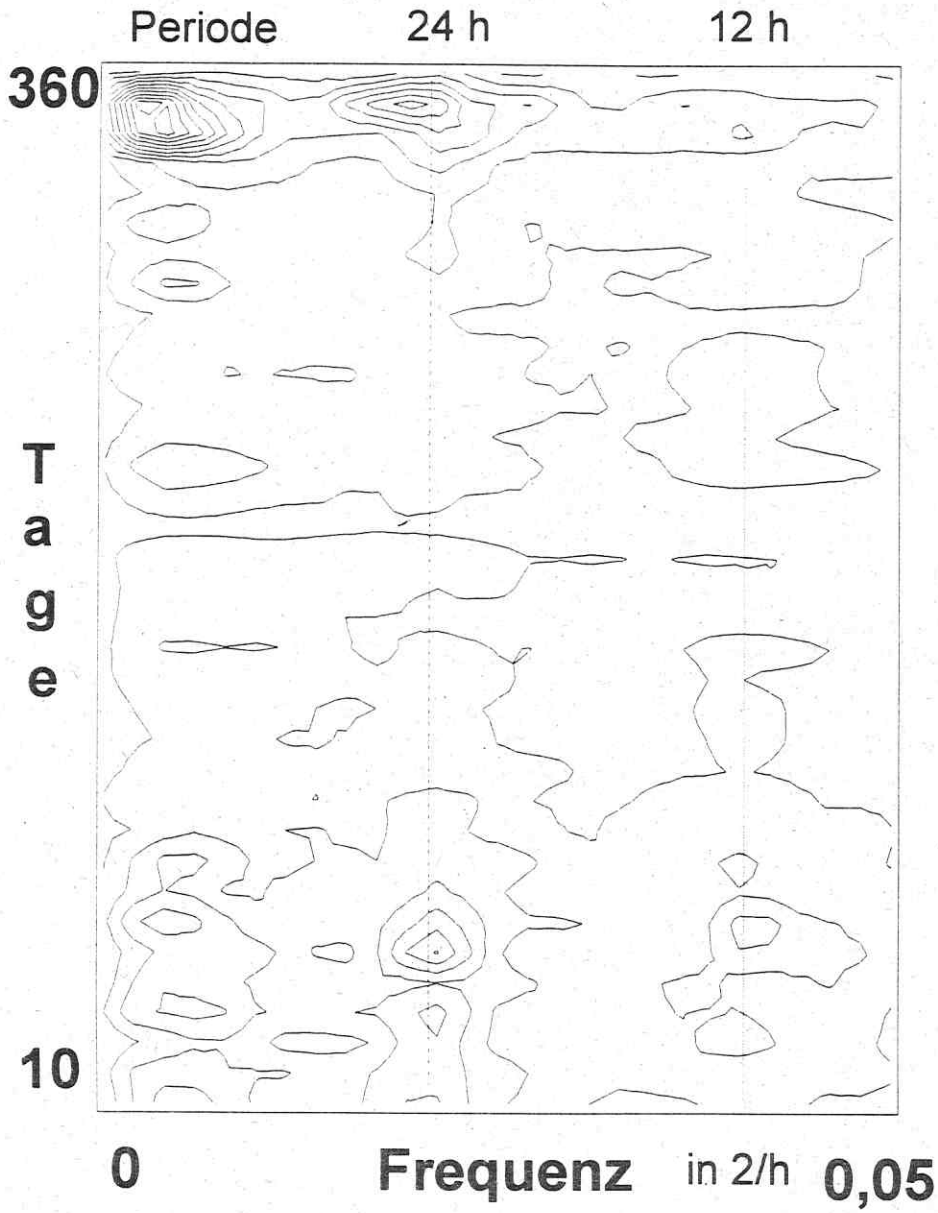


Abb. 1-21: Höhenlinien des Spektrogramms für die Daten von Leipzig-Mitte 1989; die einzelnen Periodogramme wurden jeweils aus Meßreihenabschnitten von 10 Tagen berechnet (keine Transformation) und anschließend mit einem Tukey-Hamming(5)-Fenster geglättet. In der Jahresmitte (Sommer) sind keine Episoden erkennbar.

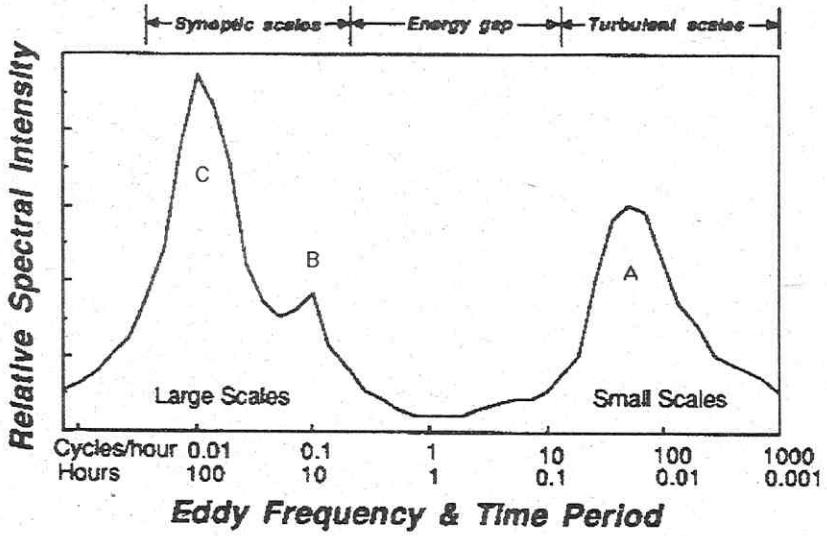


Abb. 1-22: Schematisches Spektrum der Windgeschwindigkeit in Bodennähe; aus einer Studie von Van der Hoven (1957)

- Mikroskalige Erscheinungen (A) bei Periodendauern von 10 s bis 10 min, die vorwiegend Turbulenzen darstellen.
- Einen Peak (B) bei der 24-Stunden-Periode, der den Windanstieg am Tag und das Abflauen während der Nacht repräsentiert.
- Veränderungen mit einer Periodendauer in der Umgebung von 100 Stunden, die mit dem Durchgang von Fronten und Wettersystemen verbunden sind, also den Zyklus der Rossby-Wellen zeigen. Dieser Bereich (C) des Spektrums wird synoptischer Skalenbereich genannt.

Die Korrespondenz zwischen den Bereichen des Windspektrums und des Konzentrationsspektrums war zu erwarten. Sie erlaubt die Identifikation der im Zeitablauf auftretenden Phänomene der Immissionsreihe.

Darüber hinaus ist sie auch für die gesamte Zielstellung dieser Arbeit interessant. Schließlich bietet sie an, die Prognosealgorithmen für die Windvorhersage auf die Immissionsprognose zu übertragen. Für den mikroskaligen Bereich wird dies auch in Form einer stochastischen Zeitreihenmodellierung realisiert.

Die Phänomene der synoptischen Skala stellen jedoch großräumige (≈ 1000 km) Wirbelerscheinungen dar, die von der Synoptik explizit behandelt und deterministisch vorhergesagt werden können. Die Grundlagen dazu wurden von Bjerknes (1919) und Bergeron (1928) gelegt.

Aus bereits an anderen Stellen dieser Arbeit genannten Gründen soll aber hier der univariate Ansatz verfolgt werden, d.h. allein Eigenschaften der Immissionsreihe werden für eine Vorhersage herangezogen. Für die Prognose ist es dann extrem wichtig, die niederfrequenten Vorgänge des synoptischen Skalenbereichs mit ihrer veränderlichen Frequenz modellieren zu können. Dieser Aufgabe widmet sich der 3. Teil der vorliegenden Arbeit.□

Die bei den Periodogrammuntersuchungen der SO_2-Konzentrationsreihe gefundenen Eigenschaften sollen abschließend zusammengefaßt werden:

1) Die unterschiedlich stark ausgeprägte Tagesperiode ist permanent vorhanden und kann durch Differenzbildung aus der Reihe entfernt werden.

2) Die Zyklen mit kurzer Periodendauer (unter 24 Stunden) kann man als vorwiegend stochastische Effekte auffassen, die lediglich für die sehr kurzfristige Entwicklung wichtig sind. Diese Zyklen sind äußerst instabil.

3) Für die Untersuchung von Smogepisoden erscheinen Zeitverläufe mit Perioden größer als die Tagesperiode interessant. Betrachtet man die Entwicklung des Periodogramms bei weiterrückendem Beobachtungsfenster, so lassen sich aber auch in diesem niederfrequenten Bereich keine stabilen Perioden erkennen. Vielmehr werden beispielsweise im Laufe der Zeit Dreitageszyklen durch Fünftageszyklen abgelöst, es gibt also "wandernde Peaks", d.h., das Spektrum ist zeitabhängig.

Bemerkung: Bei allen hier durchgeführten Untersuchungen wurden konventionelle Techniken der Spektralanalyse angewendet. Welchen Einfluß dabei die Abtastlänge, also die Zeit zwischen zwei aufeinanderfolgenden Meßwerten sowie die Gesamtlänge der untersuchten Zeitreihe auf die Schätzung des Spektrums des realen Immissionsverlaufs haben, ist an den entsprechenden Stellen diskutiert worden.

In den letzten Jahren sind Techniken entwickelt worden, die durch die Verwendung eines parametrischen Modells den Einfluß der genannten Meßbedingungen auf die Schätzung verringern. Hier ist besonders die von Burg (1967) entwickelte Maximum Entropie Spektral Analyse zu erwähnen (vgl. auch Olberg & Rakoczi (1984)). Eine präzise Untersuchung von Immissionsreihen mit dieser Analysentechnik könnte bisher noch verborgene Rhythmen entdecken. Diese spezielle Problematik soll hier jedoch nicht weiter verfolgt werden.□

Die hier gefundenen Resultate bilden die Grundlage für den 2. Teil der Arbeit und werden dort im Abschnitt 4 zur Konstruktion eines Komponentenmodells genutzt. Im nächsten Abschnitt soll jedoch erst noch ein weiteres Charakteristikum der Immissionszeitreihen diskutiert werden: Ihre Autokorreliertheit.

4. Autokorrelationsuntersuchungen

Der Ausgangspunkt für die Beschäftigung mit der Kürzestfristprognose von Immissionszeitreihen war deren Autokorreliertheit (Herbarth (1984)).

Der autokorrelative Zusammenhang der Meßdaten soll hier am Beispiel der Reihe CITY demonstriert werden. In den Anhängen A und B sind die zur Berechnung der Autokorrelation erforderlichen Beziehungen aufgeführt.

Nach Box-Cox-Transformation und Elimination linearer Trendanteile ergeben sich das in Abb. 1-23 dargestellte Korrelogramm und die in Abb. 1-24 widergegebene partielle Autokorrelationsfunktion.

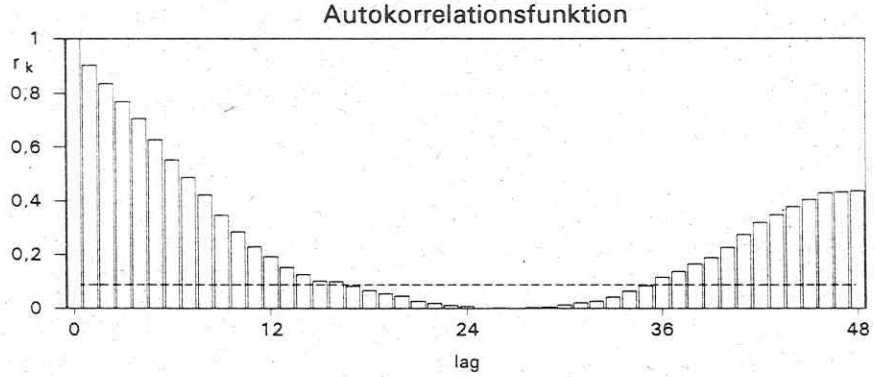


Abb. 1-23: Korrelogramm der CITY Datenreihe von Abb. 1-11; die gestrichelte Linie markiert das Konfidenzintervall für verschwindende Autokorrelation

Bei der Festlegung der Konfidenzgrenzen für das Korrelogramm geht man von der Annahme einer Reihe aus aufeinanderfolgenden unabhängigen Zufallszahlen (weißes Rauschen) aus. Für große N sind deren empirische Autokorrelationen r_1, r_2, ..., r_k näherungsweise unabhängige normalverteilte Zufallsvariable mit demselben Erwartungswert (-1/N) und derselben Varianz (1/N). In das Korrelogramm kann dementsprechend ein Konfidenzband eingezeichnet werden, in dem für jedes k>0 die Konfidenzgrenzen $-1/N \pm 1{,}96/\sqrt{N}$ zum 95% Konfidenzniveau betragen (Schlittgen & Streitberg (1991)). Alle Abweichungen von diesem Konfidenzband deuten darauf hin, daß sich die betrachtete Reihe vom weißen Rauschen unterscheidet.

Das Korrelogramm enthält wesentliche Informationen über Abhängigkeiten zwischen den Meßwerten zu verschiedenen Zeitpunkten innerhalb der Reihe. Die bereits bei der Fourieranalyse gefundene Tagesperiode zeigt sich auch hier in einer erhöhten Abhängigkeit von Zeitpunkten im Abstand von 48 Halbstundenschritten (lag=48).

Der erste Wert der ACF zum lag 0 ist aufgrund der Normierung eins. Der Wert zum lag 1 beträgt ≈ 0.9. Diese Korrelation von 0.9 zwischen zwei direkt aufeinanderfolgenden Meßpunkten wäre im Sinne einer Korrelationsanalyse dahingehend zu interpretieren, daß das Bestimmtheitsmaß $B = r^2 \cdot 100\%$ hier 81% beträgt. Das Bestimmtheitsmaß B drückt aus, welcher Anteil der Varianz der Meßreihe mit dem Regressionsmodell (hier aus der um einen Zeitschritt verschobenen Meßreihe) erklärt werden kann (vgl. Schönwiese (1992)).

Die Korrelationsanalyse setzt aber u.a. voraus, daß die untersuchten Stichproben voneinander unabhängig sind. Dies ist natürlich nicht der Fall. Eine solche Interpretation kann somit bei Autokorrelationsuntersuchungen nicht angewendet werden. Durch die Autokorrelationskoeffizienten $r_k > 0$ (k=1, 2, ...) wird ja gerade die Abhängigkeit zwischen den Meßwerten ausgedrückt.

Diese Abhängigkeit führt zu einer Erhaltungsneigung, auch Persistenz genannt. Erst bei lag = 16 gerät r_k in den Konfidenzbereich verschwindender Autokorrelation (vgl. Abb. 1-23), ist also vom weißen Rauschen nicht mehr zu unterscheiden. Das Persistenzintervall für die CITY-Reihe beträgt somit 8 Stunden und drückt aus, daß ein Meßwert die Information der letzten 8 Stunden bis zum Meßzeitpunkt beinhaltet.

Die hier feststellbare starke Autokorreliertheit und damit starke Persistenz von Immissionsreihen weckt Hoffnungen auf ihre Verwendbarkeit für Vorhersagen. Dazu ist es erforderlich, die von Box & Jenkins (1971) entwickelten Modelle anzupassen. Bei der Identifikation eines geeigneten ARIMA(p,d,q)-Modells verwendet man neben der ACF auch noch die in Abb. 1-24 gezeigte partielle Autokorrelationsfunktion (PACF).

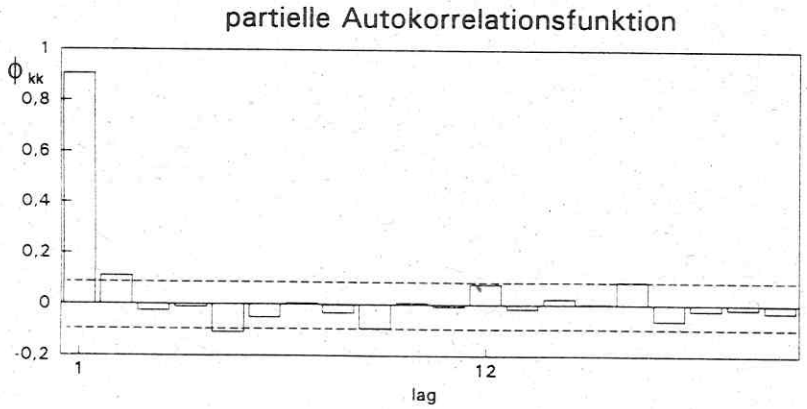


Abb. 1-24: Partielle Autokorrelationsfunktion der CITY Datenreihe von Abb. 1-11; die gestrichelte Linie markiert das Konfidenzintervall für verschwindende partielle Autokorrelation

Da die ACF allmählich abklingt, die PACF jedoch nach lag 1 abbricht, läßt sich nach dem von Box & Jenkins angegebenen Klassifikationsschema (s. Anhang B) ein AR(1)-Modell anpassen (vgl. auch Teil 2, Abschnitt 2).

Die Beurteilung des Verlaufs von ACF und PACF erfolgt unter Berücksichtigung der eingezeichneten Konfidenzgrenzen.

Der Vertrauensbereich in Abb. 1-24 ist bereits unter der Annahme eines AR(p)-Prozesses eingezeichnet worden. Dafür gilt nämlich (Schlittgen & Streitberg (1991)), daß die partiellen Autokorrelationen der Ordnung > p asymptotisch unabhängig normalverteilt sind, mit dem Erwartungswert 0 und der Varianz 1/N. Also hat ein 95%-Konfidenzintervall die Grenzen $\pm 2\sqrt{1/N}$.

Im 2. Teil wird nun die Autokorreliertheit der SO_2-Reihen zur Anpassung von AR(p)-Modellen verwendet.

2. Teil: Die Zeitreihenmodellierung

1. Stochastische Modelle

Um Immissionsmodelle aufzustellen, muß man prinzipiell von dem Gesamtsystem "Atmosphäre" ausgehen. Die gemessenen Konzentrationen eines Luftschadstoffs stellen dabei Realisierungen eines Merkmals dieses Systems dar.

Das zeitliche Verhalten von Systemen (z.B. physikalischen, ökologischen, ökonomischen, technischen, medizinischen, etc.) kann mit Hilfe der Systemtheorie nach einheitlichen mathematischen Methoden beschrieben werden. Nach DIN 66201 wird unter einem System eine abgegrenzte Anordnung von aufeinander einwirkenden Gebilden verstanden. Mit Prozeß bezeichnet man die Umformung und/oder den Transport von Materie, Energie und/oder Information. Das Verhalten eines Systems ergibt sich aus dem Verhalten seiner Prozesse.

Die Gewinnung der mathematischen Modelle von Prozessen und die Darstellung des zeitlichen Verhaltens aufgrund gemessener Signale bezeichnet man als Prozeßanalyse bzw. Prozeßidentifikation (Isermann (1974)).

Bei der Untersuchung von Prozessen und Systemen sind solche mit örtlich konzentrierten Parametern und solche mit verteilten Parametern zu unterscheiden. Letztere sind sowohl von der Zeit als auch vom Ort abhängig und werden deshalb durch partielle Differentialgleichungen beschrieben. Im Zusammenhang mit Immissionsdaten gibt es dazu Arbeiten von Omatu & Seinfeld (1981, 1982, 1983, 1987).

Einfacher zu untersuchen sind Systeme mit konzentrierten Parametern, bei denen man zur mathematischen Behandlung die Zustandsgrößen als in einem Punkt konzentriert betrachten darf. Sie werden durch gewöhnliche Differentialgleichungen beschrieben.

Das hier betrachtete System Atmosphäre liefert die Immissionsdaten also als Output. Input sind Klima- und Emissionsdaten. Für die angestrebte Modellierung des Immissionsverlaufs sollte jedoch ein minimaler Aufwand betrieben werden. Da die Registrierung aller Emissionsdaten äußerst aufwendig wäre, beschränkt sich die vorliegende Untersuchung darauf, nur den Verlauf des Ausgangssignals, also der gemessenen Immissionswerte zu modellieren. Es ist also lediglich ein Signalprozeß zu untersuchen.

Der als System-Output gemessene Immissionswert wird hier außerdem als repräsentativ für ein bestimmtes räumliches Gebiet betrachtet, so daß von einer Ortsabhängigkeit der Zustandsgrößen abgesehen werden kann. Alle innerhalb des Systems enthaltenen funktionalen Zusammenhänge, also auch die meteorologischen, stecken zwar letztlich summarisch im Output, gehen aber im Detail nicht in die phänomenologischen Modellgleichungen ein.

Bei dem zu analysierenden Signal handelt es sich um ein stochastisches Signal, welches vom Zufall abhängt. Es kann deshalb nicht exakt beschrieben werden und führt nicht zu einem deterministischen Modell. Meßbare stochastische Signale sind jedoch nicht völlig regellos, sondern besitzen innere Zusammenhänge, die dann in dem identifizierten stochastischen Modell zum Ausdruck kommen.

Wegen des zufälligen Charakters lassen sich ähnliche Signalverläufe, die aus statistisch identischen Signalquellen entstehen, zu einem Ensemble (Familie) zusammenfassen. Dieses Ensemble wird durch einen stochastischen Prozeß repräsentiert. Ein einziger Signalverlauf ist eine Realisierung des stochastischen Prozesses.

Der Signalverlauf wird bei dem hier diskutierten Immissionsproblem in Form der aufeinanderfolgenden zeitdiskreten Halbstundenmittelwerte der SO_2-Konzentration angegeben. Es handelt sich also um eine Zeitreihe.

Derartige Immissionszeitreihen wurden bereits von (Herbarth (1982)) im Hinblick auf eine ARIMA-Modellierung (Box & Jenkins (1971)) untersucht. Dabei erwies sich eine Kombination von harmonischem und autoregressivem Modell als optimal. Die hier realisierte Modellkonstruktion setzt diesen Weg konsequent fort.

Zu diesem Zweck beginnt der nächste Abschnitt mit den einfachsten Modellierungsversuchen nach der ARIMA-Methode. Später wird dann zum komplexeren Komponentenmodell übergegangen.

2. ARIMA-Modelle

Das Ziel dieses Abschnitts ist es, zu untersuchen, welche Prognosemöglichkeiten man mit den ARIMA-Modellen hat, die auf der im Teil 1, Abschnitt 4 festgestellten Autokorreliertheit der Immissionsreihe aufbauen.

Das ARIMA-Modell nach Box & Jenkins (1970) vereinigt das Modell der gleitenden Mittel (MA, engl.: moving average) mit dem autoregressiven Modell (AR). Mit dem Backshiftoperator B ($Bz_t := z_{t-1}$) und den Polynomen $\Phi_p(B)$ und $\Theta_q(B)$ besitzt dieses Modell die Darstellung (Grundlagen s. Anhang B)

$$\Phi_p(B)\ \nabla^d z_t = \Theta_q(B)\ \alpha_t. \qquad (2\text{-}1)$$

Die Zeitreihe z_t geht also aus einem gefilterten weißen Rauschen α_t hervor.

Bevor man sich über eine Prognose Gedanken macht, ist es zunächst erforderlich, ein Modell an die Meßreihe anzupassen. Die Identifikation eines geeigneten Modells, also das Festlegen der Ordnungen der Polynome $\Phi_p(B)$ und $\Theta_q(B)$ und das Schätzen der Parameter erfolgt unter Nutzung der Autokorrelationsfunktion (ACF) und der partiellen Autokorrelationsfunktion (PACF) (vgl. Box & Jenkins (1970)).

Beispielhaft soll dies an den ersten beiden Tagen der CITY Reihe (1. und 2.3.89, Leipzig-Mitte) demonstriert werden. In den Abbn. 2-1..4 sind die Originaldaten, sowie (nach varianzstabilisierender Box-Cox-Transformation) die ACF und PACF widergegeben. Am Verhalten von ACF und PACF ist erkennbar, daß sich ein AR(2)-Modell anpassen läßt.

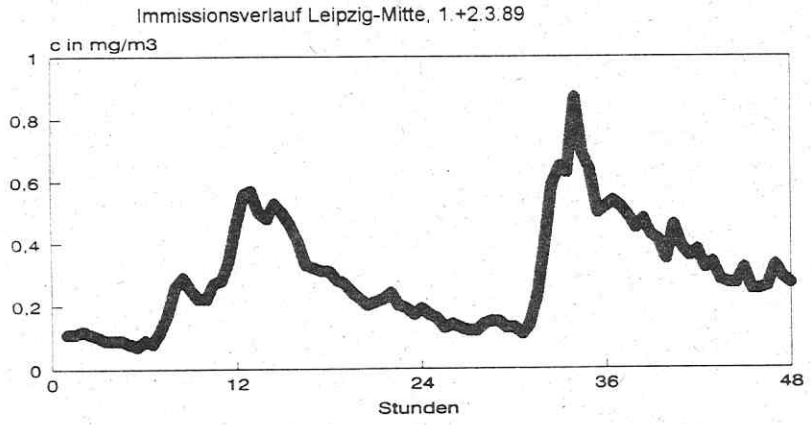


Abb. 2-1: Die ersten beiden Tage der CITY-Reihe

Bemerkung (2-1): Eine korrekte Identifikation erfordert noch den Vergleich mit anderen denkbaren Modellen anhand eines Kriteriums (Akaike Information Criterion (Akaike (1970)): AIC := ln σ_α^2 + 2p/N; σ_α^2=Varianz der Residuenreihe α_t, p=Ordnung des AR-Modell-Polynoms Φ, N=Anzahl der Meßwerte der Originalreihe), welches minimal sein muß. Auch dieser Vergleich favorisiert das AR(2)-Modell (AIC=-3,45) dicht gefolgt vom AR(1)-Modell (AIC=-3,40).□

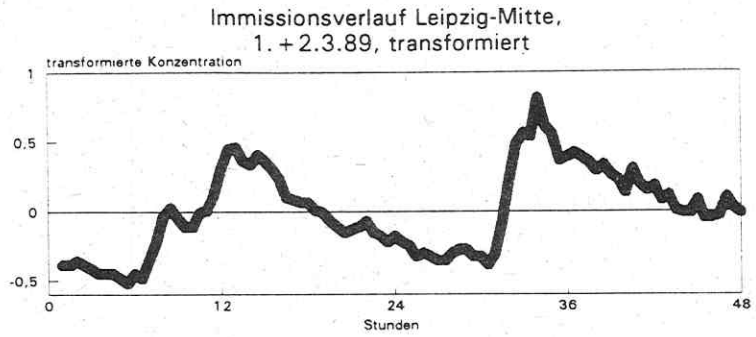


Abb. 2-2: Daten aus Abb. 2-1 nach Box-Cox-Transformation und Elimination des Mittelwertes.

Die Schätzung der Parameter ergibt:

$$\phi_1 = 1.18; \quad \phi_2 = -0.25.$$

Bemerkung (2-2): Nach den Schätzformeln von (Box & Jenkins (1970)) berechnen sich beim AR(2)-Modell ϕ_1 und ϕ_2 entsprechend den Yule-Walker-Gleichungen aus den Autokorrelationskoeffizienten r_1 und r_2 zu

$$\phi_1 = r_1(1-r_2)/(1-r_1^2) \quad \text{und} \quad \phi_2 = (r_2-r_1^2)/(1-r_1^2).\square$$

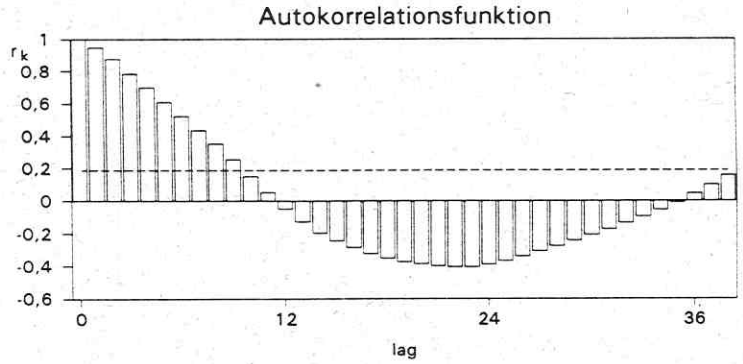


Abb. 2-3: Autokorrelationsfunktion der varianzstabilisierten Daten von Abb. 2-2

Somit kann die in Abb. 2-1 gezeigte Meßreihe nach Elimination ($z_t \rightarrow z_t$ - m) ihres Mittelwertes m mit dem AR(2)-Modell

$$(1 - 1.18\,B + 0.25\,B^2)\, z_t = \alpha_t \qquad (2\text{-}2)$$

modelliert werden. Gemäß Anhang C berechnet sich dann die Prognose $^*z_t(l)$ ab dem Zeitpunkt t für einen l Zeitschritte voraus liegenden Wert aus:

$$
\begin{aligned}
^*z_t(1) &= 1.18\, z_t - 0.25\, z_{t-1} \\
^*z_t(2) &= 1.18\, ^*z_t(1) - 0.25\, z_t \\
&\ldots \\
^*z_t(l) &= 1.18\, ^*z_t(l\text{-}1) - 0.25\, ^*z_t(l\text{-}2).
\end{aligned}
\qquad (2\text{-}3)
$$

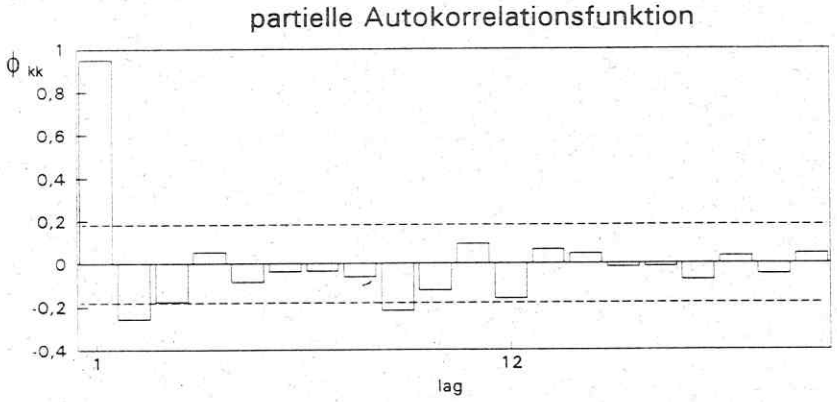


Abb. 2-4: Partielle Autokorrelationsfunktion der varianzstabilisierten Daten von Abb. 2-2

In Abb. 2-5 sind zwei solche Prognosen zu sehen, die zu verschiedenen Zeitpunkten erstellt wurden. Bei günstiger Wahl des Prognosezeitpunktes (hier die Prognose nach 35 Stunden) wird der künftige Verlauf der Meßkurve im wesentlichen vorhergesagt, was Abb. 2-5 deutlich erkennen läßt. Andererseits liegt die zweite Prognose (nach 32 Std.) völlig daneben.

Beide Prognosen spiegeln letztendlich lediglich das exponentielle Abklingverhalten in Richtung auf den Mittelwert wider, welches durch die Prognosegleichungen rekursiv ausgedrückt wird. Wird nun die Prognose zu einem Zeitpunkt gestartet, an dem in der Meßreihe ein solches Abklingen einsetzt, so kommt es zu einer adäquaten Vorhersage. Eine derartig günstige Situation wird sich aber nur zufällig ergeben. Im allgemeinen kann die AR(2)-Prognose das Verhalten der Immissionsreihe nicht vorhersagen.

Vor dem Prognosepunkt ist in Abb. 2-5 der Wert der Prognosekurve aus den letzten beiden Meßwerten stets neu berechnet worden (vgl. erste Prognoseformel). Dieses rekursive Verfahren legt nahe, sich auf eine Prognose für den nächsten Wert zu beschränken.

Dann kann man unter Nutzung aller bisher erfolgten Messungen (bzw. der letzten 48) ständig fortlaufend neu die Parameter ϕ_1 und ϕ_2 schätzen und einen Prognosewert angeben.

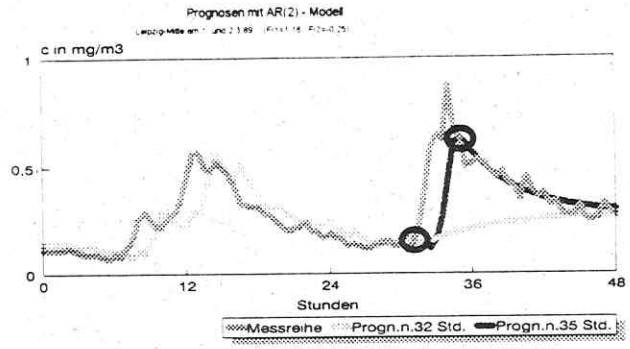


Abb. 2-5: Prognosen mit einem angepaßten AR(2)-Modell, die zu zwei verschiedenen Zeitpunkten gestartet wurden.

Eine derartige Prognose ist in Abb. 2-6 gezeigt. Auch sie liefert wenig Informationen, da sie der gedämpften und zeitverschobenen Meßkurve entspricht.

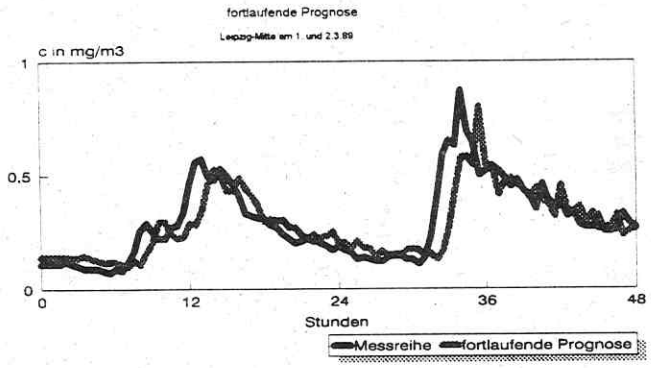


Abb. 2-6: Fortlaufend neu erstellte Prognose

Abschließend läßt sich feststellen, daß ein reines ARIMA-Modell die gegebenen Immissionszeitreihen nicht so zufriedenstellend abbildet, daß damit eine adäquate Vorhersage des Reihenverlaufs möglich scheint. Da eine periodische Struktur erkennbar ist, bieten sich Versuche mit saisonalen Modellen oder eine vorherige Abtrennung harmonischer Komponenten (Fourieranalyse) an. Auf einer solchen Zerlegung im spektralen Bereich beruht das im 4. Abschnitt konstruierte Komponentenmodell.

Zuvor soll jedoch noch ein Versuch beschrieben werden, die einzelnen Tagesverläufe mit autoregressiven Modellen abzubilden und dadurch jedem Tag einen Punkt im Parameterraum der Modelle zuzuordnen. Anlaß für diese Untersuchung war die Hoffnung, daß die Punkte im Parameterraum Strukturen bilden, die sich für eine Prognose ausnutzen lassen.

3. Modellierung der Tagesverläufe

Natürlich kann man auch für die Meßreihenabschnitte der einzelnen Tage eine Modellierung finden. Mit Hilfe dieser tageweisen Modellbildung ist es dann möglich, jeden Tag durch die Angabe weniger Modellparameter zu charakterisieren. Welches Bild ergibt sich daraus und welche Schlußfolgerungen lassen sich für eine Prognose ziehen?

Bei Betrachtung einzelner Tagesverläufe (Reihen mit 48 Meßwerten) wurde in fast allen Fällen das AR(1)- oder das AR(2)-Modell identifiziert. Deshalb konnte AR(2) (mit dem Spezialfall AR(1) für ϕ_2=0) als Standardmodell gewählt und für jeden Tagesgang die Parameter ϕ_1 und ϕ_2 geschätzt werden. Das AR(2)-Modell lautet:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \alpha_t \qquad (2\text{-}4)$$

In der ϕ_1-ϕ_2-Ebene läßt sich somit jeder Tag durch einen Punkt markieren, dessen Lage den Typ des Tagesganges widergibt (siehe dazu die Abbn. 2-7,8,9; hier Verwendung der Box-Jenkins-Schätzer aus Bemerkung (2-2)).

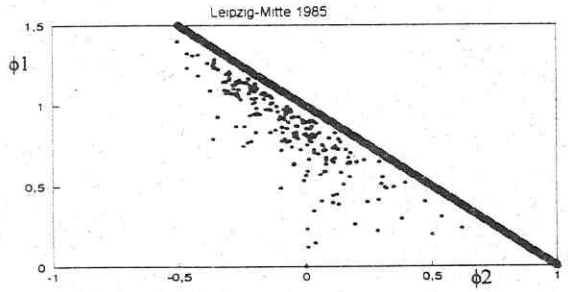


Abb. 2-7: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1985.

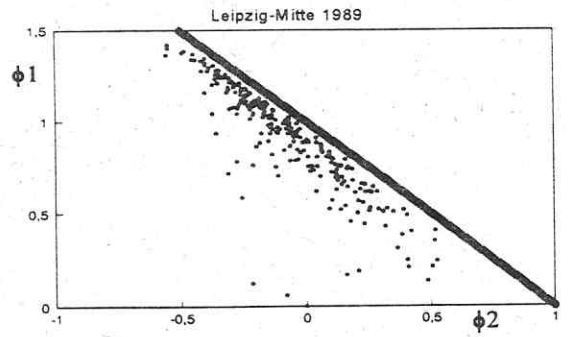


Abb. 2-8: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1989.

In Abb. 2-10 sind für typische Lagen in der ϕ_1-ϕ_2-Ebene die Verläufe der dazugehörigen Zeitreihe illustriert worden. Die verschiedenen Verlaufstypen unterscheiden sich in ihrem spektralen Verhalten, welches mit den Parametern des AR(2)-Modells gemäß (Kendall & Ord (1990))

$$I(f) \propto \frac{1}{(1+\phi_1^2+\phi_2^2-2\phi_1(1-\phi_2)\cos 2\pi f - 2\phi_2 \cos 4\pi f)} \qquad (2\text{-}5)$$

zusammenhängt (I(f) = relative Intensität des Periodogramms bei der Frequenz f). Für

$$\cos 2\pi f = -\frac{\phi_1(1-\phi_2)}{4\phi_2} \qquad (2\text{-}6)$$

besitzt das Periodogramm I(f) einen Extremwert, der bei $\phi_2 < 0$ ein Maximum und bei $\phi_2 > 0$ ein Minimum ist.

Bei den hier geschätzten Parameterwerten gilt stets $\phi_1 > 0$. Ist darüber hinaus $\phi_2 < 0$, so gibt es im Periodogramm einen Niederfrequenz-Peak. Die Existenz von niederfrequenten Zyklen in Zeitreihen ist also allein eine Folge des autoregressiven Zusammenhangs und bedarf nicht unbedingt einer periodischen Modellierung. Diese Erkenntnis ist Yule zu verdanken, nach dem auch der AR(2) als Yule Prozeß bezeichnet wird.

Die Abbn. 2-7,8,9 zeigen, daß es eine für Prognosezwecke ausnutzbare Clusterung bei bestimmten Verlaufstypen innerhalb der ϕ_1-ϕ_2-Ebene nicht gibt. Unter meteorologischen Gesichtspunkten, die ein wesentlicher Aspekt des Immissionsgeschehens sind, sind die erkennbaren fließenden Übergänge auch zu erwarten. Eine gewisse Strukturierung als Folge des Emissionsverhaltens wäre dennoch denkbar gewesen.

Die tageweise Modellierung ist also im Hinblick auf eine Prognose von Smogsituationen nicht gewinnbringend. Dafür gibt es noch einen weiteren Grund:

Der Verlaufstyp steht zur Stärke der Immissionsbelastung in keinem Zusammenhang. Vielmehr zeigt sich, daß die Höhe der Immission in der geschätzten Streuung σ_α^2 der α_t zum Ausdruck kommt.

Bemerkung (2-3): Die in den Abbn. 2-7,8,9 eingezeichnete Linie repräsentiert die Gleichung $\phi_1 + \phi_2 = 1$. Erfüllen die geschätzten Parameter diese Bedingung, so handelt es sich um eine instationäre Reihe. Man kann dies anhand folgender Umformung erkennen. Beim Yule-Prozeß AR(2) gilt:

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 \\ &= 1 - B(\phi_1 + \phi_2 B) \end{aligned} \qquad (2\text{-}7)$$

und durch Einsetzen der Relation $\phi_1 + \phi_2 = 1$ ergibt sich:

$$\begin{aligned} \Phi(B) &= 1 - B(1 - \phi_2 + \phi_2 B) \\ &= (1 + \phi_2 B)(1 - B). \end{aligned} \qquad (2\text{-}8)$$

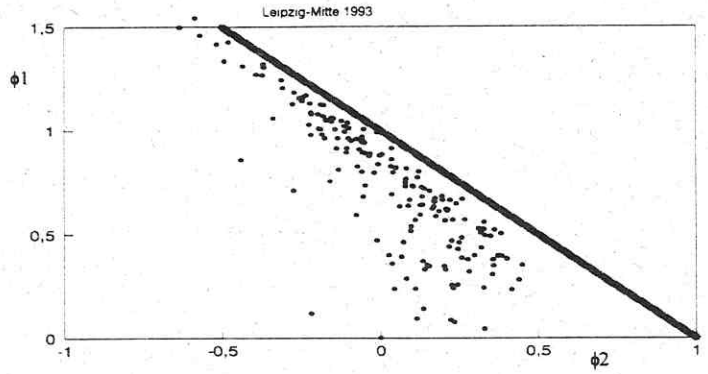


Abb. 2-9: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1993.

Das Polynom enthält also eine Nullstelle für B = 1 und stellt somit ein Modell für eine instationäre Reihe dar (vgl. Anhang C). Durch die Differenzbildung $(1 - B)\ z_t = z_t - z_{t-1}$ wird eine stationäre Reihe gebildet, die dann durch $(1 + \phi_2 B)$ als AR(1) modelliert wird.

Da in den Abbn. 2-7,8,9 die Parameterwerte für die untersuchten Tage in der Regel zu dieser eingezeichneten Linie Abstand halten, sind die Tagesabschnitte der Immissionsreihe also stationäre Zeitreihen.□

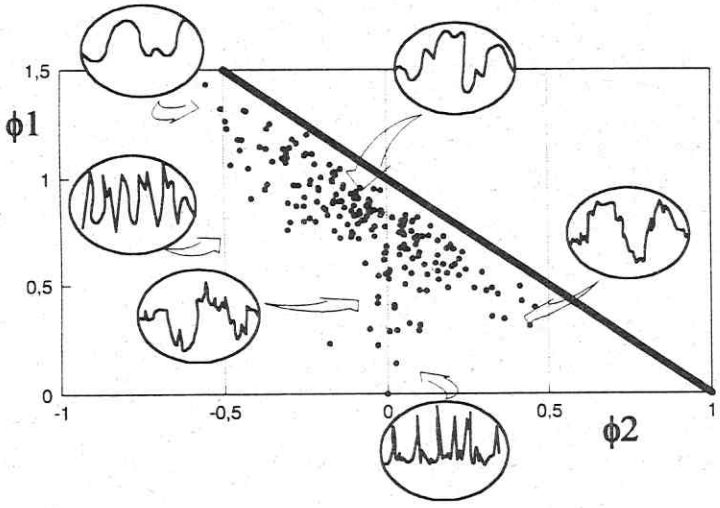


Abb. 2-10: Typischer Verlauf der AR(2)-Reihen an verschiedenen Punkten der ϕ_1-ϕ_2-Ebene.

4. Komponentenmodell

Bisher wurden die Möglichkeiten der Modellierung und Prognose von SO_2-Immissionsdaten mit ARIMA-Modellen untersucht. Die Grundlage war dabei der autokorrelative Zusammenhang der Meßreihe. Weitere Eigenschaften, wie z.B. das Vorhandensein periodischer Strukturen, waren bestenfalls implizit im Modell enthalten. Auch die Zeitabhängigkeit des Erwartungswertes der Reihe, also ihr instationäres Verhalten ist bisher nicht explizit einbezogen worden.

An dieser Stelle soll nun ein Modelltyp eingeführt werden, der es erlaubt, die in der Immissionsreihe enthaltenen Phänomene explizit zu behandeln. Es zeigte sich im Ergebnis der Untersuchungen des spektralen Verhaltens von SO_2-Serien $\{z_t\}$ mit Hilfe des Periodogramms im Teil 1, Abschnitt 3, daß es günstig ist, $\{z_t\}$ durch Zerlegung in das folgende **Komponentenmodell** aufzuspalten:

$$z_t = F^{-1} (T_t + S1_t + S2_t + R_t) \qquad (2\text{-}15)$$

Mit F^{-1} soll die varianzstabilisierende Transformation (siehe Teil 1, Abschnitt 2) symbolisiert werden. Die Komponenten weisen zwei wesentliche Eigenschaften auf:

Zum einen ist die Komponentenzerlegung eine Aufspaltung im spektralen Bereich. Das bedeutet aufgrund der Orthogonalität von Fouriertermen, daß die resultierenden Komponenten voneinander unabhängig sind. Man kann also jede Komponente für sich betrachten. Wechselwirkungen gibt es nicht. Wie gut die Zerlegung ist, wird natürlich von der Durchlaßcharakteristik des verwendeten Filters bestimmt (vgl. Anhang C und D). Dementsprechend fällt dann auch die Unabhängigkeit der Teilreihen aus.

Zum anderen bildet jede Komponente einen anderen in der Meßreihe enthaltenen phänomenologischen Effekt ab. Die Modellbestandteile verhalten sich also ganz verschieden und müssen deshalb selbst auch wieder mit Modellen unterschiedlicher Struktur beschrieben werden:

Die aufgeführten Komponenten lassen sich durch Transformation und anschließende Filterung aus der Meßreihe extrahieren. Eine Erläuterung des dazu verwendeten Filteralgorithmus folgt im nächsten Abschnitt. In Abb. 2-11 wird das Vorgehen an der CITY-Reihe verdeutlicht.

Die Trendkomponente $\{T_t\}$ beschreibt die niederfrequentesten Veränderungen der Immissionskonzentration (Frequenz < 1/24h). Die Abschneidefrequenz des Glättungsfilters (vgl. Anh. D) wurde gerade so gewählt, daß die Trendkomponente die Tagesperiode noch nicht enthält. Die Güte der Glättung läßt sich durch Betrachten der Ableitung der Trendfunktion beurteilen, die eine leichte Wellenbewegung des Trends deutlicher als dieser selbst sichtbar macht und die vom Kalman-Algorithmus automatisch mitgeliefert wird.

Der Trend von SO_2-Daten ist eine nichtperiodische Funktion, die das instationäre Verhalten des Erwartungswertes beschreibt. Er wird im Teil 3, Abschnitt 2 noch genauer analysiert. Da sich Smogepisoden sowie das Niveau der Immission ausschließlich in der Trendfunktion zeigen, ist diese von besonderem Interesse. Wegen der Unabhängigkeit von den anderen Komponenten kann man sich also bei weiteren Untersuchungen von Smogepisoden allein auf den Trend konzentrieren, was im Teil 3 dieser Arbeit erfolgt.

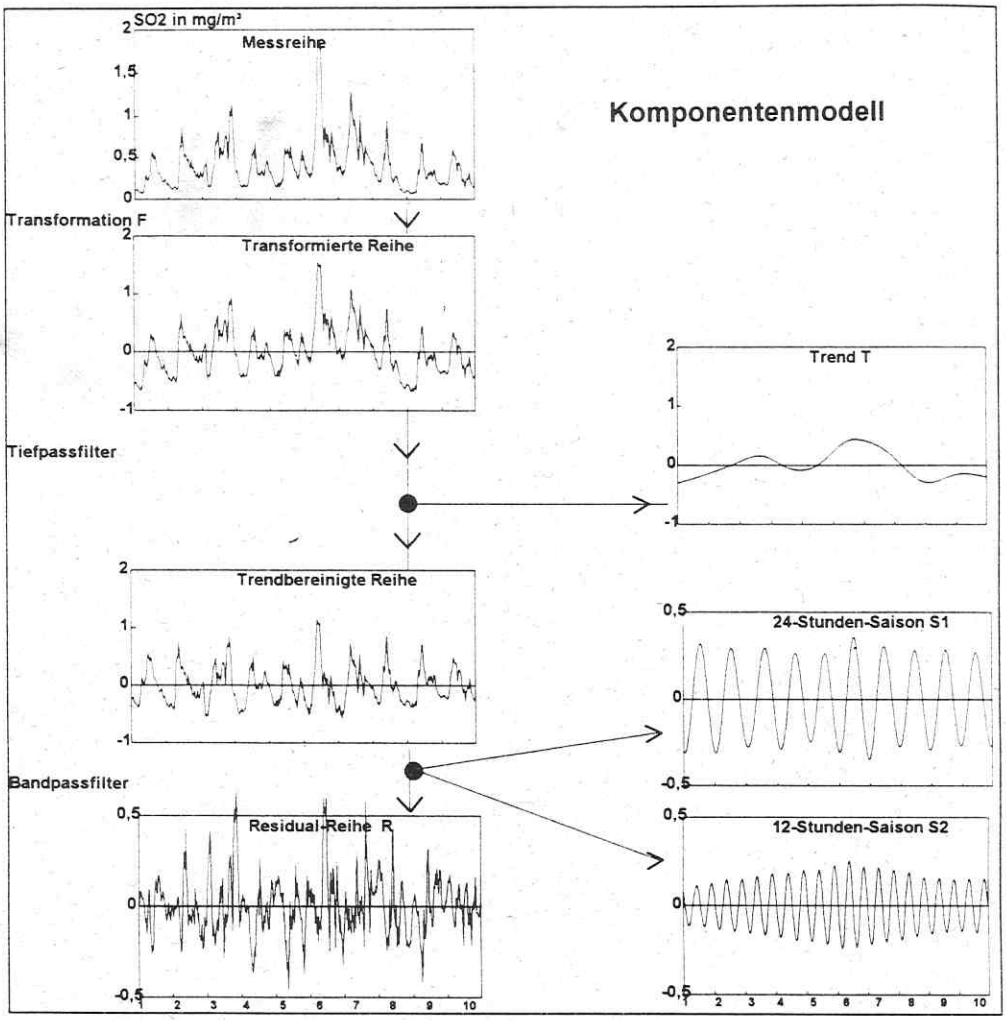


Abb. 2-11: Darstellung der aus der CITY-Reihe mit dem Kalman-Filter extrahierten Komponenten (NVR(Trend) = 0.00001, NVR(S1) = NVR(S2) = 0.001)

Die Saisonkomponenten $\{S1_t\}$ und $\{S2_t\}$ geben den Tagesrhythmus wieder. Da dieser letztendlich astronomisch festgelegt ist, handelt es sich um eine stabile Periodizität, die mit einer harmonischen Regression modelliert werden kann.

Die Restkomponente $\{R_t\}$ trägt der Tatsache Rechnung, daß es sich um einen stochastisch beeinflußten Prozeß handelt. Sie kann mit einem Markov-Prozeß AR(1) modelliert werden.

Mit diesem Komponentenmodell wird man den in der SO_2-Immissionsreihe enthaltenen unterschiedlichen Phänomenen gerecht. Durch ihre Isolierung werden sie einer ihrem Charakter entsprechenden weiteren Modellierung zugänglich gemacht.

Bemerkung (2-4): Bei Zeitreihen von Lufttemperaturen im Abstand von 6 Stunden findet Essenwanger (1980) ähnliche Ergebnisse. Die Reihen lassen sich in drei Komponenten zerlegen: Weißes (der rotes) Rauschen, Zyklen und Quasi-Zyklen. Die Quasi-Zyklen sind dabei durch eine feste Periode aber zeitabhängige Amplitude und Phasenlage charakterisiert. Er kann einige dieser Quasi-Zyklen mit synoptisch-skaligen Phänomenen in Verbindung bringen.□

Eine Zerlegung der Reihe in Komponenten bringt Vorteile für die Interpretation und die Vorhersage, da jede Komponente eine eigene Charakteristik und Dynamik aufweist. So kann man den Trend als Ergebnis der permanenten Einwirkung äußerer Kräfte auffassen, beispielsweise als Widerspiegelung von Emission und Großwetterlage. Die Periodizitäten können den Schwingungen des untersuchten physikalischen Systems (hier der Atmosphäre) zugeordnet werden.

Bei der Anwendung solcher Komponentenmodelle sollte jedoch nicht vergessen werden, daß die Existenz der Komponenten und ihre additive Überlagerung nicht a priori gegeben sondern lediglich eine Erfahrungstatsache ist, die sich aus der Analyse umfangreichen Datenmaterials ergibt und auch für die Zukunft immer wieder neu verifiziert werden muß.

5. Zustandsraummodelle

Für die Zerlegung einer gemessenen Immissionsreihe in die im Abschnitt 4 beschriebenen Komponenten muß man sich geeigneter Filteralgorithmen bedienen. Das bedeutet, daß die Filter in der Lage sein müssen, die bei den Untersuchungen im Teil 1, Abschnitt 3 gefundenen Spektralbereiche möglichst sauber voneinander zu trennen.

Interessanterweise ergibt sich ein für diesen Zweck sehr nützlicher Filteralgorithmus aus einer Modellierungsaufgabe der Regelungstechnik, bei der es um die Schätzung einer nicht meßbaren Zustandsgröße geht. Das Modell wird als Zustandsraummodell bezeichnet. Fundamentale Arbeiten dazu wurden von Kalman geleistet (vgl. Kalman & Bucy (1961)), obwohl die statistischen Grundlagen bereits von Plackett (1950) formuliert wurden.

"Zustandsraum" meint eine Modelldarstellung in Form einer oder mehrerer Bewegungsgleichungen der unbekannten Zustandsgröße und die zusätzliche Angabe einer Beobachtungsgleichung nach der die Zustandsgröße mit der gemessenen Beobachtungsgröße zusammenhängt. Sowohl Zustands- als auch Beobachtungsgleichung enthalten dabei in der Regel stochastische Einflüsse, z.B. Rauschterme (μ und η).

Das lineare Zustandsraummodell für den vektoriellen Zustand $\mathbf{x}_t$ lautet dann:

Beobachtungsgleichung	$z_t = \mathbf{H}(t\text{-}1)\ \mathbf{x}_t + \mu_t$	(2-16)
Zustandsgleichung	$\mathbf{x}_t = \mathbf{F}(t\text{-}1)\ \mathbf{x}_{t-1} + \mathbf{G}(t\text{-}1)\ \eta_{t-1}$	(2-17)
wobei $\mathbf{x}_t = (y_t\ x_t\ \ldots)^T$ ist.		(2-18)

Die Schätzung des unbekannten Zustandes $\mathbf{x}_t$, der hier die Komponenten des Modells aus Abschnitt 4 zusammenfaßt, aus der gemessenen Beobachtungsgröße z_t erfolgt mit einem rekursiven Verfahren.

Die multidisziplinäre Anwendung der Zustandsraummodellierung hat zu einer sehr uneinheitlichen Terminologie geführt. Während die ingenieur-technische Literatur, deren Terminologie hier benutzt werden soll, vom "Kalman-Filter" spricht, bevorzugen Statistiker oft die Bezeichnungen "Bayes'sche Vorhersage" oder "Dynamisches Lineares Modell" (Harrison & Stevens (1976), Fildes (1983)) oder "Lineare Dynamische Regression" (Duncan & Horn (1972)).

Die komplette Komponentenzerlegung des Abschnitts 4 benötigt eine Kombination verschiedener Filter. Dazu muß ein Kalman-Filter verwendet werden, daß auf einem Zustandsraummodell (2-16) + (2-17) mit folgender Struktur basiert.

Beobachtungsgleichung $z_t = T_t + S1_t + S2_t + R_t$

$$= T_t + {}_{11}\theta_t\cos(2\pi tf_1) + {}_{21}\theta_t\sin(2\pi tf_1) + {}_{12}\theta_t\cos(2\pi tf_2) + {}_{22}\theta_t\sin(2\pi tf_2) + R_t, \quad (2\text{-}19)$$

was sich auch schreiben läßt als

$$z_t = \mathbf{H}(t\text{-}1)\,\mathbf{x_t} + \mu_t \quad (2\text{-}20)$$

mit

$$\mathbf{x}_t = (\,T_t\;\; D_t\;\; {}_{11}\theta_t\;\; {}_{21}\theta_t\;\; {}_{12}\theta_t\;\; {}_{22}\theta_t\,)^T, \quad (2\text{-}21)$$

$$\mathbf{H}(t\text{-}1) = (\,1\;\; 0\;\; \cos(2\pi tf_1)\;\; \sin(2\pi tf_1)\;\; \cos(2\pi tf_2)\;\; \sin(2\pi tf_2)\,), \quad (2\text{-}22)$$

$$\mu_t = R_t, \quad (2\text{-}23)$$

und

Zustandsgleichung $\mathbf{x}_t = \mathbf{F}\,\mathbf{x}_{t-1} + \mathbf{G}\,\eta_{t-1}$ (2-24)

mit

$$\eta_t = (\,0\;\; T\eta_t\;\; {}_{11}\eta_t\;\; {}_{21}\eta_t\;\; {}_{12}\eta_t\;\; {}_{22}\eta_t\,)^T \quad (2\text{-}25)$$

$$\mathbf{F} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}; \qquad G_{ij} = \delta_{ij} = \begin{cases} 1 & \text{für } i=j \\ 0 & \text{für } i\neq j \end{cases} \quad (2\text{-}26)$$

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & NVR(T) & 0 & 0 & 0 & 0 \\ 0 & 0 & NVR(1) & 0 & 0 & 0 \\ 0 & 0 & 0 & NVR(1) & 0 & 0 \\ 0 & 0 & 0 & 0 & NVR(2) & 0 \\ 0 & 0 & 0 & 0 & 0 & NVR(2) \end{pmatrix} = \mathrm{cov}(\eta). \quad (2\text{-}27)$$

Dabei wurde als Glättungsstärke für den Trend NVR(T) = 0.00001 und einheitlich für alle periodischen Anteile NVR(1) = NVR(2) = 0.01 gewählt. Eine detaillierte Beschreibung der Zustandsraummodelle und der darauf aufbauenden Filteralgorithmen ist im Anhang D nachzulesen.

Dem Trend liegt also ein IRW-Modell und den Saisonkomponenten ein DHR-Modell zugrunde.

3. Teil: Die Trendextrapolation

1. Das Interpolations- und Extrapolationsverhalten beim Kalman-Filter

Im vorigen Abschnitt wurde die Meßreihe in unabhängige Komponenten mit verschiedenen typischen Eigenschaften zerlegt, was einer Aufspaltung im Spektralbereich entspricht. Die Auftrennung der Reihe geschah mit einem Kalman-Filter. Diese Herangehensweise verfolgt die Aufgabenstellung der Modellierung einer Reihe. Für die Modellbildung ist es dabei wichtig, ein Kalman-Filter und damit ein Zustandsraummodell (welches im Anhang D erläutert wird) mit geeignetem spektralen Filterverhalten zu wählen.

Die in der vorliegenden Arbeit untersuchte Fragestellung soll jedoch primär die der Prognose einer Zeitserie sein. Sicher ist auch gerade in diesem Fall ein adäquates Modell für den Reihenverlauf wichtig, mit dem dann eine Extrapolation vorgenommen werden kann. Das für die Komponentenzerlegung verwendete Kalman-Filter stellt bereits ein solches Modell dar, das durch eine vorgegebene Zustandsgleichung definiert ist (vgl. Anhang D).

Da jedoch der Vorhersagebereich dadurch charakterisiert ist, daß keine Meßwerte vorliegen, ist zu fragen, wie das Modell in diesem Fall reagiert. Eine analoge Problematik wie bei der Vorhersage, also der Extrapolation, liegt bei der Interpolation von Meßlücken, die durch Fehlwerte entstanden sind, vor. Dabei ist eine Interpolation gewissermaßen eine "Lückenschließung" von zwei Seiten her, eine Extrapolation dagegen ein von einer Seite ausgehendes Verfahren.

Ein gegebenes Kalman-Filter besitzt also wegen des zugrunde liegenden Zustandsraummodells neben seinem *spektralen Durchlaßverhalten*, welches beispielhaft im Anhang D erläutert wird, noch ein typisches *Interpolations- und Extrapolationsverhalten*. Letzteres wird bei innerhalb oder außerhalb der Meßreihe vorliegenden Fehlwerten wirksam.

Für die Anpassung des Modells wird natürlich ausschließlich der Bereich der vorliegenden Meßdaten und deren Frequenzspektrum genutzt. Für das Prognose-Verhalten eines Kalman-Filters sind aber seine Interpolations- und Extrapolationseigenschaften wichtiger, als seine spektralen Eigenschaften! Die Modellanpassung erfolgt auf der Basis des spektralen Verhaltens, die Prognose mit dem Modell nutzt dann aber dessen Interpolations- und Extrapolationsverhalten. Diese Erkenntnis drückt ein **Dilemma der Zeitreihenprognose** aus und verdeutlicht, warum man streng zwischen Modellierung und Prognose unterscheiden sollte.

Methoden zur Zeitreihenmodellierung sind in der Vergangenheit intensiv entwickelt worden und in einschlägigen Softwarepaketen verfügbar (z.B. SAS, SPSS, microCAPTAIN). Wie obige Diskussion zeigt, ist jedoch die Anwendung der damit gefundenen Modelle für eine Prognose zu hinterfragen.

Dieser 3. Teil der Arbeit soll deshalb die bei der Prognose mit dem Kalman-Filter auftretenden Effekte transparent machen und zu ersten Vorschlägen für eine Verbesserung des Prognosealgorithmus führen. Damit wird jetzt über die Anwendung bekannter Algorithmen hinausgegangen und eine methodische Weiterentwicklung vorgenommen. Wegen ihrer Bedeutung für die Smogepisoden sollen die auftretenden Fragen speziell an der Trendkomponente der Meßreihe diskutiert werden.

Anhand der übersichtlichen Zustandsraumdarstellung des dem Kalman-Filter zugrunde liegenden Modells lassen sich bereits theoretische Schlüsse über das Verhalten bei Fehlwerten ziehen. Dazu soll das Interpolationsverhalten an folgenden Beispielen betrachtet werden:

Beispiel (3-1): Die Zustandsraumdarstellung des verwendeten Trendmodells (integrated random walk)

Beobachtungsgleichung	$z_t = y_t + \mu_t$	μ_t=weißes Beobachtungsrauschen	(3-1)
Zustandsgleichungen	$y_t = y_{t-1} + x_{t-1}$	(IRW-Modell)	
	$x_t = x_{t-1} + \eta_t$	η_t=weißes Rauschen	

Der Kalman-Filteralgorithmus für dieses IRW-Trendmodell soll kurz mit IRWSMOOTH bezeichnet werden. Er schätzt aus der beobachteten Meßreihe $\{z_t\}$ den Verlauf des Zustandes $\{y_t\}$, also der Trendkomponente. In den Bereichen, in denen keine Information über die Beobachtungsgröße z_t vorliegt (Fehlwerte), muß man davon ausgehen, daß sich der Zustand entsprechend seiner Modellgleichung in ungestörter Weise entwickelt. D.h., die Zustandsgleichung gilt unverändert, jedoch ist für η_t dessen Erwartungswert $E[\eta_t] = 0$ zu verwenden.

Damit ergeben sich aus (3-1) die Zustandsgleichungen

$$\begin{aligned} y_t &= y_{t-1} + x_{t-1} \\ x_t &= x_{t-1} \end{aligned} \qquad (3\text{-}2)$$

Der Zustand y_t ist also in Bereichen mit Fehlwerten durch einen *konstanten Zuwachs* x_t gekennzeichnet, der Filteralgorithmus versucht die Reihe y_t *linear* zu interpolieren bzw. extrapolieren.

Für den Fall einer Interpolation ist dies in der Abb. 3-1 dargestellt. Eine unterschiedlich starke Glättung, die mit dem NVR eingestellt wird (s. Anhang D), wirkt sich bei der Interpolation in verschieden großen Anstiegen aus. Da die Filterung bei einer Interpolation zweiseitig ist, versucht der Algorithmus in beiden Richtungen die Reihe mit jeweils konstantem Anstieg fortzusetzen. Natürlich muß dann im mittleren Bereich ein Übergang gefunden werden.

Es ist erkennbar, daß das Ergebnis der Interpolation von der Wahl der Glättungsstärke (NVR) abhängt. Natürlich spielt auch die Breite des Bereiches der Fehlwerte eine Rolle.□

Beispiel (3-2): Ein alternatives Trendmodell (double integrated random walk)

Natürlich könnte man anstelle des IRW-Modells auch ein DIRW-Modell (s. Anhang D) für den Trend verwenden. Im Spektralbereich hätte das entsprechende Kalman-Filter, welches ebenfalls ein Tiefpaß ist, eine andere Abschneidefrequenz, so daß zur Abtrennung derselben Trendfunktion wie beim IRW-Glätter ein entsprechend anderer NVR-Wert gewählt werden müßte. Der Vergleich beider Modelle soll anhand von Abb. 3-2 diskutiert werden:

Wählt man für jedes der beiden Filter eine passende Glättungsstärke (z.B. NVR = 10^{-5} für IRWSMOOTH und 10^{-7} für DIRWSMOOTH), so lassen sich sehr ähnliche Filterergebnisse erzielen: Mit beiden Methoden erhält man nahezu gleich verlaufende Trendfunktionen, jedoch nur solange Meßdaten vorliegen!.

Deutliche Unterschiede kann man dagegen im **Interpolationsbereich** der Abb. 3-2 erkennen: Da hier für die Zustands- gleichungen des DIRW-Modells gilt:

$$\begin{aligned} y_t &= y_{t-1} + x_{t-1} \\ x_t &= x_{t-1} + w_{t-1} \\ w_t &= w_{t-1}, \end{aligned} \qquad (3\text{-}3)$$

versucht also der entsprechende Filteralgorithmus (DIRWSMOOTH) die zweite Differenz w_t der Zustandswerte y_t konstant zu halten. Die Reihe wird also nicht linear, wie beim IRW-Modell, sondern parabelförmig fortgesetzt. Dies führt zu sehr verschiedenem Interpolations- und natürlich analog auch verschiedenem Extrapolationsverhalten (vgl. Abb. 3-2).□

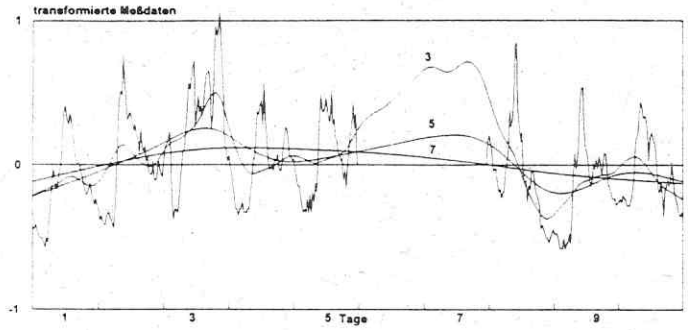


Abb. 3-1: Verlauf der transformierten CITY-Reihe mit Fehlwerten an den Tagen 6 und 7, die mit IRWSMOOTH bei verschiedenen NVR interpoliert wurden (-log(NVR)=3, 5, 7).

Beispiel (3-3): Die Zustandsraumdarstellung des Periodenmodells (dynamic harmonic regression)

Beobachtungsgleichung $z_t \triangleq a_t \cos(\omega t) + b_t \sin(\omega t) + \mu_t$ (3-4)

Zustandsgleichungen $a_t = a_{t-1} + \eta 1_t$ (RW-Modelle)
$b_t = b_{t-1} + \eta 2_t$

Die Beobachtungsgleichung ist die Gleichung einer harmonischen Regression. Durch das Zulassen der Zeitabhängigkeit der Regressionskoeffizienten a_t und b_t, die entsprechend ihrer Zustandsgleichungen als random walk (RW) modelliert wird, filtert der dazugehörige Kalman-Algorithmus (DHR) nicht nur exakt eine Frequenz (ω) aus der Meßreihe heraus (wie dies eine "normale" harmonische Regression mit konstanten Parametern a und b tun würde), sondern ein um ω gelegenes Frequenzband, dessen Breite durch die Glättungsstärke NVR = cov(ε1)/cov(μ) = cov(ε2)/cov(μ) bestimmt wird (Ng et al. (1990)).

Die Zustandsgleichungen lassen erkennen, daß bei Fehlwerten die Koeffizienten a und b konstant gehalten werden, d.h., das Filter für das Periodenmodell setzt die Periode ($1/\omega$) mit konstanter Amplitude fort.□

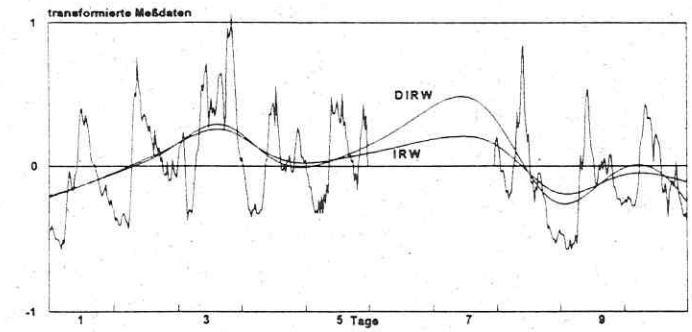


Abb. 3-2: Verlauf der transformierten CITY-Reihe mit Fehlwerten an den Tagen 6 und 7, die mit IRWSMOOTH bzw. DIRWSMOOTH interpoliert wurden (NVR = 0,00001 bzw. 0,0000001).

Die Untersuchungen zum Interpolations- und Extrapolationsverhalten des Kalman-Algorithmus zeigen an den Beispielen 3-1 bis 3-3, daß die Wahl des zugrundeliegenden Zustandsraummodells eine sehr wesentliche Rolle spielt. Es ist also bei der Verwendung des Kalman-Filters für die **Prognose** nicht nur auf das aus einem Modellansatz resultierende spektrale Filterverhalten zu achten (wie dies bei der Zeitreihen**modellierung** wesentlich ist), sondern auch auf das Verhalten bei Vorliegen von Fehlwerten.

So steckt beispielsweise in einer Prognose von Immissionsreihen mit dem Periodenmodell der dynamischen harmonischen Regression (DHR) die Annahme, daß die betrachtete Periodizität in der nahen Zukunft stabil und mit konstanter Amplitude erhalten bleibt.

Die Prognose des Trends einer Immissionsreihe mit dem integrated random walk Modell (IRW) geht hingegen davon aus, daß sich der Trend vom Prognosepunkt aus linear extrapolieren läßt. Wie der nachfolgende Abschnitt zeigen wird, ist jedoch ein solcher Verlauf für die Trendfunktion absolut untypisch. Aus diesem Grunde erscheint eine Trendprognose mit dem IRWSMOOTH-Algorithmus unzweckmäßig. Die Anwendung dieses Filters liegt offenbar mehr im Bereich der Modellierung, wo es zur Abtrennung des niederfrequenten Anteils einer Zeitreihe sehr geeignet ist.

2. Die Trendfunktion

Die Untersuchungen zum Interpolations- und Extrapolationsverhalten von Filtern speziell für die Trendkomponente werfen die Frage nach charakteristischen Eigenschaften dieser Komponente auf. Erst mit Kenntnis dieser typischen Eigenschaften ist man in der Lage, einen adäquaten Modellansatz zu machen, der dann nicht nur für eine Modellierung, sondern auch für eine Extrapolation des Trends genutzt werden kann.

Man könnte befürchten, daß der in IRWSMOOTH bereits enthaltene Modellansatz bei der Glättung in irgendeiner Weise der Trendkomponente aufgeprägt wird. Dies ist jedoch nicht der Fall. Das Filter wirkt lediglich als Tiefpaß, so daß man davon ausgehen kann, daß die Trendfunktion ein reales Charakteristikum der Meßreihe ist. Das verwendete Zustandsraummodell kommt ausschließlich in Bereichen mit Fehlwerten zum Tragen.

Eine aus einer Immissionsmeßreihe extrahierte Trendfunktion soll in Abb. 3-3 genauer betrachtet werden.

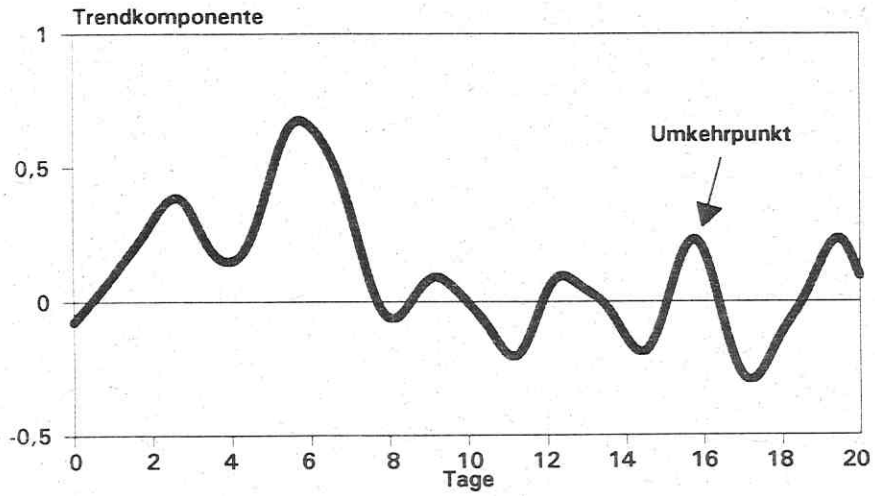


Abb. 3-3: Mit IRWSMOOTH extrahierte Trendfunktion für die Meßstelle Leipzig-Mitte vom 1.3.-20.3.89 (transformiert); die ersten 10 Tage entsprechen also der CITY-Reihe.

Die Tatsache, daß visuell keinerlei Tagesperiode erkennbar ist, läßt vermuten, daß der Tiefpaß aus den Originaldaten wirklich nur den interessierenden Frequenzbereich herausgefiltert hat. Zur Kontrolle wurde die vom Kalman-Algorithmus automatisch mit berechnete Ableitung der Trendfunktion betrachtet. Durch das Differenzieren würden kleine Schwankungen deutlicher hervortreten. Aber auch die Ableitungsfunktion zeigt visuell keine Tagesperiode, was obige Vermutung bestätigt.

Charakteristisch ist nun, daß es in dieser Trendkomponente keine festen Periodizitäten gibt. Schwingungsvorgänge, wie sie sich in Abb. 3-3 zwischen etwa dem 10. und dem 20. Tag andeuten, lösen sich später wieder auf und man stelit außerdem bei genauer Betrachtung fest, daß sich die Periodenlänge im Laufe der Zeit ändert.

Das einzig Typische am Immissions-Trend ist sein glatter Verlauf und die Tatsache, daß er über einen längeren Zeitbereich nicht instationär wird. Das alternierende Auftreten von Umkehrpunkten, an denen die Richtung des Trendverlaufs umkehrt, verhindert ein beliebig weites Entfernen vom Ausgangsniveau.

Der Abstand zwischen den Umkehrpunkten weist die in Abb. 3-4 gezeigte asymmetrische Verteilung auf. Je mehr Zeit seit dem letzten Umkehrpunkt vergangen ist, um so größer wird die Wahrscheinlichkeit für eine erneute Umkehr. Diese Eigenschaft könnte für einen Prognosealgorithmus genutzt werden, was hier jedoch nicht geschehen soll.

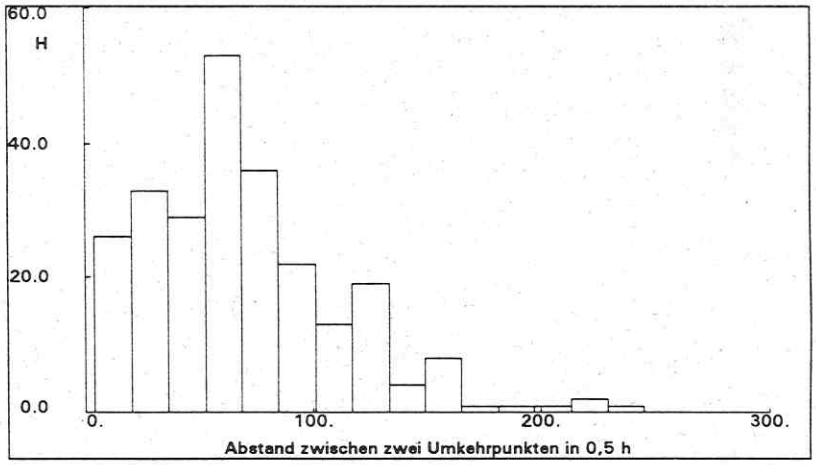


Abb. 3-4: Häufigkeitsverteilung H der Abstände zwischen den Umkehrpunkten der Trendfunktion von der Meßreihe Leipzig-Mitte 1989

Bemerkung (3-1): Im Teil 1, Abschnitt 3 wurden die Skalenbereiche des Frequenzspektrums prozeßorientiert diskutiert. Dabei konnte die Trendkomponente mit den Phänomenen des synoptischen Skalenbereiches identifiziert werden. Die Umkehrpunkte im Trend entstehen also durch den Zustrom neuer Luftmassen. Somit muß die in Abb. 3-4 gezeigte Häufigkeitsverteilung mit der Häufigkeit des Andauerns von Großwetterlagen über eine bestimmte Zeit korrespondieren.□

3. Das Phasenraumbild

Für die graphische Darstellung der Trendfunktion bietet sich aber auf Grund der Tatsache, daß das Kalman-Filter neben dem geglätteten Trend T_t auch dessen Ableitung $D_t = T_t'$ mit liefert noch eine weitere Möglichkeit an. Es läßt sich die Trendbewegung in dem aus Trend und Ableitung aufgespannten Phasenraum widergeben. Abb. 3-5 zeigt ein solches Phasenraumbild für die Trendfunktion der Daten des Jahres 1989 der Station Leipzig-Mitte.

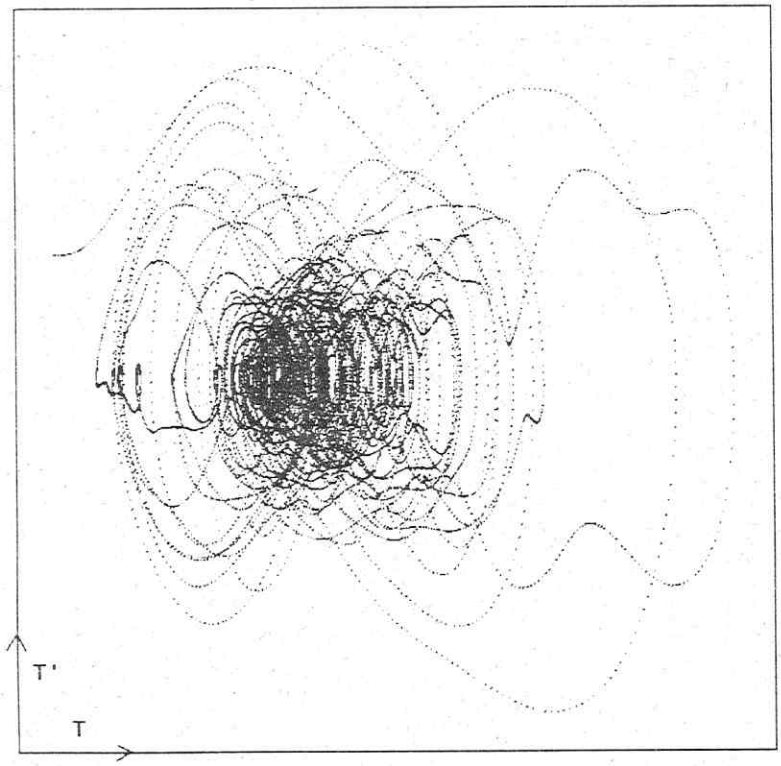


Abb. 3-5: Phasenraumbild für die Trendfunktion T und deren Ableitung T ' zu den Daten von Leipzig-Mitte 1989.

Um die Bewegung der Phasenraumtrajektorie genauer zu analysieren, ist sie in Abb. 3-6 für die 10 Tage der CITY Reihe dargestellt. Ausgangspunkt für eine Phasenraumdarstellung sind die Trendkomponente T und deren Ableitung T ' , die in ihrem zeitlichen Verlauf links oben und unten aufgezeichnet sind. Verwendet man dann T als Abszisse und T ' als Ordinate, so entsteht das rechts unten befindliche Phasenraumbild mit dem typischen Trajektorienverlauf.

Diese Bahn erinnert an eine Hyperzykloide, die ein auf dem Umfang einer Ellipse gelegener Punkt beschreibt, wenn die Ellipse gleichzeitig mit einer Winkelgeschwindigkeit Ω rotiert und ihr Mittelpunkt entlang der Abszissenachse eine Translation erfährt.

Insgesamt wird am Phasenraumbild deutlich, daß es sich beim Trend und seiner Ableitung keinesfalls um unabhängige stochastische Prozesse handelt. Denn nur durch einen funktionalen Zusammenhang beider kann eine solche stetige Bahn entstehen.

Daraus leitet sich die Grundidee für eine neue Prognosemethode ab: Man geht davon aus, daß das typische Phasenraumverhalten der Trajektorie zeitlich invariant bleibt und nutzt es für eine Extrapolation. Dazu muß der Zusammenhang zwischen T und T ' , nämlich die o.g. Hyperzykloidenbewegung mathematisch formuliert werden.

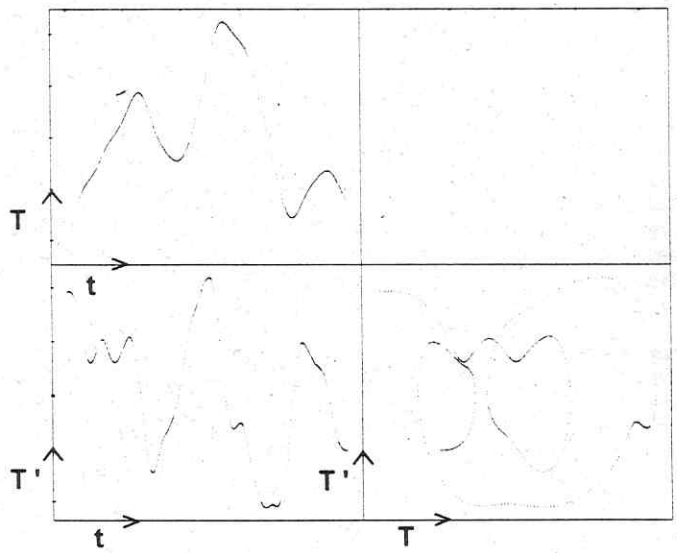


Abb. 3-6: Phasenraumtrajektorie (rechts unten) der Trendfunktion der CITY-Reihe; zur Illustration der Entstehung der Phasentrajektorie sind die Trendreihe T und ihre Ableitung T ' im Zeitverlauf (t) links oben und links unten mit angegeben.

Für die Bewegung der Trendfunktion T_t und ihrer Ableitung T_t' entlang einer Hyperzykloide im Phasenraum gilt folgende Bewegungsgleichung:

$$T_t = c(t) - a(t) \cos(\Omega(t)\, t - t_o(t)) \qquad (3\text{-}5)$$

$$T_t' = \Omega(t)\, a(t) \sin(\Omega(t)\, t - t_o(t)) \qquad (3\text{-}6)$$

Dieses Modell (3-5)+(3-6) soll als Hyperzykloidenmodell der Phasenraumtrajektorie der Trendkomponente bezeichnet werden.

In den Beziehungen (3-5)+(3-6) kommt der Zusammenhang beider Funktionen zum Ausdruck: T_t' ist wirklich die Ableitung der Trendfunktion. Die Zeitabhängigkeit der Parameter c(t), a(t), $\Omega(t)$ und $t_0(t)$ wird als sehr schwach angenommen.

$$\frac{1}{c}\frac{dc}{dt}, \frac{1}{a}\frac{da}{dt}, \frac{1}{\Omega}\frac{d\Omega}{dt}, \frac{1}{t_o}\frac{dt_o}{dt} << \Omega \qquad (3\text{-}7)$$

D.h., sie sind zwar nicht konstant und müssen für jedes Stück der Trajektorie neu bestimmt werden, im Bereich der Prognose können sie jedoch mit hinreichender Genauigkeit als konstant betrachtet werden.

Aber wie sieht denn nun beispielsweise eine IRWSMOOTH-Extrapolation bzw. -Interpolation im Phasenraum aus?

Eine solche Interpolation wurde ja bereits für die CITY-Reihe mit einem Bereich von Fehlwerten an den Tagen 6 und 7 in Abb. 3-1 gezeigt. Wie dies für die Trendkomponente und deren Ableitung im Phasenraum aussieht, verdeutlicht Abb. 3-7. Der Bereich der Interpolation ist verstärkt gezeichnet. Da dieses Filter ein Konstanthalten des Anstiegs bewirkt, zeigt die Phasenraumtrajektorie hier einen nahezu waagerechten Verlauf.

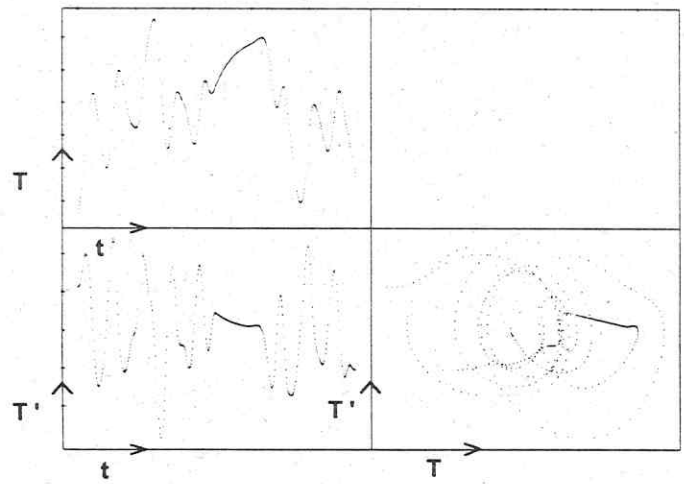


Abb. 3-7: Phasenraumtrajektorie (rechts unten) der Trendfunktion der CITY-Reihe mit Fehlwerten an den Tagen 6 und 7 (NVR=0.001); zur Illustration der Entstehung der Phasentrajektorie sind die Trendreihe und ihre Ableitung im Zeitverlauf links oben und links unten mit angegeben.

Das von der üblichen kreisenden Bewegung abweichende Verhalten der Trajektorie im Interpolationsbereich, wo sich der Ordinatenwert nur wenig ändert, ist deutlich zu erkennen. Die Interpolation ignoriert also die Verlaufseigenschaften der Phasenraumkurve vollkommen.

Ein Trendverlauf nach dem Hyperzykloidenmodell würde dann wie folgt aussehen:

$$T_t = c - a \cos(\Omega t - t_o), \qquad (3\text{-}8)$$

Verwendet man diese Funktion zur Extrapolation, so bringt dies eine Erhöhung der Genauigkeit gegenüber der linearen Trendprognose $T_t \sim t$ des IRWSMOOTH-Filters. Denn wenn man für kleine Zeiten ($t << 1/\Omega$) diese modifizierte Trendprognose entwickelt, dann ergibt sich

$$T_t \approx c - a\{1 - \Omega t + ...\} - t_o = c - a - t_o + a \Omega t + ..., \qquad (3\text{-}9)$$

was die lineare Prognose als zweiten Term der Taylorreihe enthält.

4. Lokale harmonische Approximation

Das Hyperzykloidenmodell im Phasenraum bedeutet, die Trendfunktion durch eine harmonische Oszillation zu modellieren. Die Parameter dieses Modells sind jedoch selbst wieder zeitabhängig, so daß die geschätzten Werte nur lokal gültig sind. Diese Eigenschaft entspricht dem gefundenen Spektralverhalten für die Trendkomponente: Es ereignen sich einzelne Episoden, deren Periode jedoch nicht konstant ist; das Spektrum ist zeitabhängig.

Durch eine lokale harmonische Approximation ist man in der Lage, ein solches zeitabhängiges Spektrum zu modellieren. Das gefundene Modell läßt sich für Extrapolationen jedoch nur bis zu einem bestimmten Prognosehorizont verwenden. Darüber hinaus ist mit so starken Änderungen der Parameter c(t), a(t), $\Omega(t)$ und $t_0(t)$ zu rechnen, daß der Trendverlauf nicht mehr adäquat wiedergegeben wird.

Bemerkung (3-1): Was läßt sich über die Größe dieses Prognosehorizontes aussagen?

Da die Trendfunktion gerade durch das Herausfiltern extrem niederfrequenter Anteile aus der Meßreihe entstanden ist, enthält sie auch nur Schwingungen mit Periodenlängen, die größer als ein Tag sind. Für kürzere Zeiten kann der Trend also als nahezu linear betrachtet und auch so (z.B. mit IRWSMOOTH) extrapoliert werden. Eine Trendprognose über einen Tag ist aber für den angestrebten Zweck einer Kürzestfristprognose absolut ausreichend.

Darüber hinaus wird mit der lokalen harmonischen Approximation eine lokal dominante Periode des Trends, die die gerade ablaufende Episode beschreibt, durch Ω modelliert. Somit werden Veränderungen der Parameter c, a, Ω und t_0 länger als $1/\Omega$ dauern (vgl. Formel (3-7)). Bei einer Smogepisode von beispielsweise 7 Tagen würde der Prognosehorizont nach lokaler harmonischer Approximation sogar mindestens 7 Tage betragen. Einschränkend ist jedoch zu bemerken, daß $1/\Omega$ nur ein Schätzwert für die Dauer der Episode ist und die Gültigkeit des Hyperzykloidenmodells voraussetzt.□

Ein Algorithmus für die Anwendung einer lokalen harmonischen Approximation zur Verbesserung der Trendprognose hätte also folgendermaßen auszusehen:

1) Die Meßreihe (sie enthalte keine Fehlwerte) wird nach der Transformation der Daten mit einem Kalman-Filter in ein Komponentenmodell zerlegt. Für die weiteren Schritte sind die Trendfunktion und die am Prognosepunkt vorliegenden Werte der Amplituden der Saisonkomponenten wichtig.

2) An die Trendkomponente wird fortlaufend lokal eine harmonische Schwingung angepaßt, d.h. für $T_t = c - a \cos(\Omega t - t_0)$ sind die Parameter c, a, Ω und t_0 zu schätzen.

3) Die für eine Umgebung des Prognosepunktes gültigen Parameterwerte werden verwendet, um die Trendkomponente mit der Funktion $T_t = c - a \cos(\Omega t - t_0)$ zu extrapolieren.

4) Aus dem extrapolierten Trend und den mit konstanten Amplituden fortgesetzten Saisonkomponenten wird die Prognosereihe zusammengesetzt.

5. Parameterschätzung

Um erst einmal prinzipiell die Arbeitsweise einer solchen lokalen harmonischen Approximation mit variabler Frequenz Ω zu testen und zu sehen, ob sie sich für die interessierenden Prognosezwecke eignet, wurde hier eine Schätzung der Parameter c, a, Ω und t_0 aus vier Meßpunkten vorgenommen. Dabei muß man sich jedoch immer der folgenden Probleme bewußt sein und versuchen, zwischen den sehr gegensätzlichen Anforderungen einen praktikablen Kompromiß zu finden:

- Wegen der lokalen Gültigkeit des Parametersatzes wäre es natürlich optimal, für jeden Punkt eine Schätzung zu besitzen. Dies würde auch eine Analyse des Zeitverlaufs der Parameter ermöglichen, welche weitere interessante Aspekte aufzeigen könnte.

- Zur Schätzung von vier Unbekannten benötigt man jedoch vier Gleichungen und damit vier Punkte (Stützstellen) der Meßreihe.

- Da es sich, wie schon erwähnt, um sehr langperiodische Veränderungen handelt, ist Ω entsprechend klein. Um die bei der Schätzung von Ω auftretenden numerischen Ungenauigkeiten möglichst klein zu halten, wäre es günstiger, den Abstand zwischen den Stützstellen möglichst groß zu wählen.

- Dies steht jedoch im Widerspruch zur Zeitabhängigkeit von Ω, die eine möglichst punktuelle Schätzung erfordert.

- Die Schätzgleichung für Ω (wird noch in Formel (3-15) angegeben und) ist wegen der Winkelfunktionen nichtlinear und somit schwer lösbar.

Die günstigste Lösung wäre eine rekursive Schätzung von Ω mit einem Filteralgorithmus. Da es sich aber um eine nichtlineare Schätzgleichung handelt, ist dazu die Entwicklung eines geeigneten nichtlinearen Filters erforderlich - die linearen Kalman-Filter versagen hier.

Für die verwendeten vier Stützstellen lassen sich dann mit dem Hyperzykloiden-modell (3-8) folgende vier Gleichungen aufschreiben:

$$T_i = c - a\cos(\Omega i - t_0) \qquad (3\text{-}10)$$
$$= c - a\{\cos(t_0)\cos(\Omega i) + \sin(t_0)\sin(\Omega i)\}, \quad \text{für} \quad i=0, 1, 2, 3$$

Zur Eliminierung von c und a werden zweckmäßigerweise die Verhältnisse gebildet:

$$v2 = (T_2 - T_0) / (T_1 - T_0) \qquad (3\text{-}11)$$

$$v3 = (T_3 - T_0) / (T_1 - T_0), \qquad (3\text{-}12)$$

was auf

$$v2 = [1 - \cos(2\Omega) - \tau \sin(2\Omega)] / [1 - \cos(\Omega) - \tau \sin(\Omega)] \qquad (3\text{-}13)$$

$$v3 = [1 - \cos(3\Omega) - \tau \sin(3\Omega)] / [1 - \cos(\Omega) - \tau \sin(\Omega)] \qquad (3\text{-}14)$$

mit $\tau = \tan(t_o)$ führt.

Entfernt man noch τ, so erhält man als nichtlineare Bestimmungsgleichung für Ω:

$$[1 - \cos(2\Omega) - v2\,(1 - \cos(\Omega))]\,[\sin(3\Omega) - v3 \sin(\Omega)] = [1 - \cos(3\Omega) - v3\,(1 - \cos(\Omega))]\,[\sin(2\Omega) - v2 \sin(\Omega)] \qquad (3\text{-}15)$$

Um eine Lösung für diese Gleichung in Abhängigkeit von v2 und v3 zu finden, werden die Winkelfunktionen bis zur 5. Ordnung in Ω in eine Taylorreihe entwickelt. Dieses Vorgehen ist auch begründet, da, wie bereits diskutiert, zu erwarten ist, daß Ω sehr kleine Werte annehmen wird ($\Omega << 1/(i \cdot 0.5h)$).

$$[1-(1 - 4\Omega^2/2 + 16\Omega^4/24) - v2\,(1-(1 - \Omega^2/2 + \Omega^4/24))]\,[(3\Omega - 27\Omega^3/6) - v3\,(\Omega - \Omega^3/6)] \approx$$

$$[1-(1 - 9\Omega^2/2 + 81\Omega^4/24) - v3\,(1-(1 - \Omega^2/2 + \Omega^4/24))]\,[(2\Omega - 8\Omega^3/6) - v2\,(\Omega - \Omega^3/6)]$$

Durch Zusammenfassen von Termen bis zur 5. Ordnung in Ω folgt:

$$\Omega^5 \{ (2 - v2)(81 - v3)/24 - (3 - v3)(16 - v2)/24 + (9 - v3)(8 - v2)/12 - (27 - v3)(4 - v2)/12 \} + \Omega^3 \{ (3 - v3)(4 - v2)/2 - (2 - v2)(9 - v3)/2 \} = 0 \qquad (3\text{-}16)$$

Wegen $\Omega \neq 0$ führt dies auf:

$$\Omega^2 = 4\,(v3 + 3(1 - v2)) / (v3 + 7(1 - v2)). \qquad (3\text{-}17)$$

Damit ist eine Bestimmungsgleichung für Ω gefunden.

Für ein Zahlenbeispiel werden wieder die CITY-Daten herangezogen. Die Datenreihe der Tage 4, 5 und 6 soll die Basis für eine Prognose des 7. Tages sein. Dazu wird sie mit dem oben beschriebenen Kalman-Filter in Komponenten zerlegt. Aus vier Punkten mit einem Abstand von jeweils 2 Stunden wird Ω fortlaufend aus der Trendkomponente berechnet. Die dabei entstehenden Werte sind in Tabelle 3-1 angegeben.

Da die Phasenraumtrajektorie natürlich nicht exakt einer Hyperzykloide folgt, ändert sich der Ω-Wert entlang dieser Trajektorie. Solange die Kurve in der Phasenebene konvex gekrümmt ist, erhält man stets einen positiven Wert für Ω. Ist die Trajektorie jedoch sehr schwach gekrümmt und fast eine Gerade, so wird Ω sehr klein. Verläuft andererseits die Phasenraumkurve in einem bestimmten Abschnitt konkav (vgl. Abb. 3-6), dann läßt sich eine Ellipsenbahn überhaupt nicht anpassen und es wird $\Omega^2 < 0$. Dies ist an den in Tabelle 3-1 mit einem "?" markierten Stellen der Fall.

Tabelle 3-1: Fortlaufend aus der Trendkomponente geschätzte Frequenz Ω der lokalen harmonischen Approximation für die Tage 4, 5 und 6 der CITY-Reihe; die Zahlenwerte sind auf 3 Stellen gerundet.

Nr.	4. Tag	5. Tag	6. Tag
1	?	?	0,04
2	?	?	0,039
3	?	?	0,038
4	?	?	0,037
5	?	?	0,037
6	?	?	0,036
7	?	?	0,036
8	?	?	0,036
9	?	?	0,036
10	?	?	0,037
11	?	?	0,037
12	?	?	0,037
13	?	?	0,038
14	?	?	0,038
15	?	?	0,038
16	?	?	0,038
17	?	?	0,038
18	?	?	0,038
19	?	?	0,037
20	?	?	0,036
21	?	?	0,035
22	?	?	0,033
23	0,013	?	0,03
24	0,033	?	0,024
25	0,041	0,009	0,014
26	0,045	0,017	?
27	0,046	0,021	?
28	0,045	0,024	?
29	0,041	0,027	?
30	0,037	0,029	?
31	0,03	0,031	?
32	0,023	0,033	?
33	0,017	0,034	?
34	0,015	0,034	?
35	0,021	0,034	?
36	0,028	0,033	?
37	0,033	0,032	?
38	0,036	0,031	?
39	0,036	0,031	?
40	0,031	0,031	?
41	0,022	0,031	?
42	?	0,032	?
43	?	0,033	?
44	?	0,035	?
45	?	0,037	?
46	?	0,038	?
47	?	0,04	?
48	?	0,04	?

Da sich die Trajektorie jedoch global gesehen "um ein Zentrum wickelt" (vgl. Abb. 3-5), muß ihr überwiegender Verlauf konvex sein.

Für die Extrapolation wird ein mittlerer Wert aus den letzten Ω-Daten verwendet. Im hier angegebenen Beispiel beträgt dieser $\Omega_{mittel} \approx 0.035$.

Die Parameter c, a und t_o lassen sich nun sehr komfortabel über eine lineare Regression der Trendwerte T_t gegen $\cos(\Omega t)$ und $\sin(\Omega t)$ bestimmen

$$T_t = \xi_o + \xi_1 \cos(\Omega t) + \xi_2 \sin(\Omega t), \qquad (3\text{-}18)$$

denn nach dem Hyperzykloidenmodell (3-8) gilt

$$T_t = c - a \cos(t_o) \cos(\Omega t) - a \sin(t_o) \sin(\Omega t). \qquad (3\text{-}19)$$

Damit erhält man aus den Regressionskoeffizienten ξ_o, ξ_1, ξ_2:

$$t_o = \arctan(\xi_2/\xi_1) \qquad (3\text{-}20)$$

$$a = -\xi_1/\cos(t_o) \qquad (3\text{-}21)$$

$$c = \xi_o \qquad (3\text{-}22)$$

6. Modifizierte Trendextrapolation

Wie die Extrapolation mit einer solchen harmonischen Funktion im Phasenraum für das hier betrachtete Beispiel aussieht, zeigt Abb. 3-8. Die Regression zur Schätzung der Parameter c, a und t_0 basierte dabei nur auf dem Trendverlauf der letzten 24 Stunden (dicke Linie). Für den dargestellten Fall ergab sich c = -0.6, a = 0.3 und t_0 = -1.4. Anschließend wurde der Trend entlang der gestrichelt gezeichneten Ellipsenbahn fortgesetzt. Der Schnittpunkt zwischen Trend-Trajektorie und Extrapolationsellipse wurde als Anschlußpunkt verwendet.

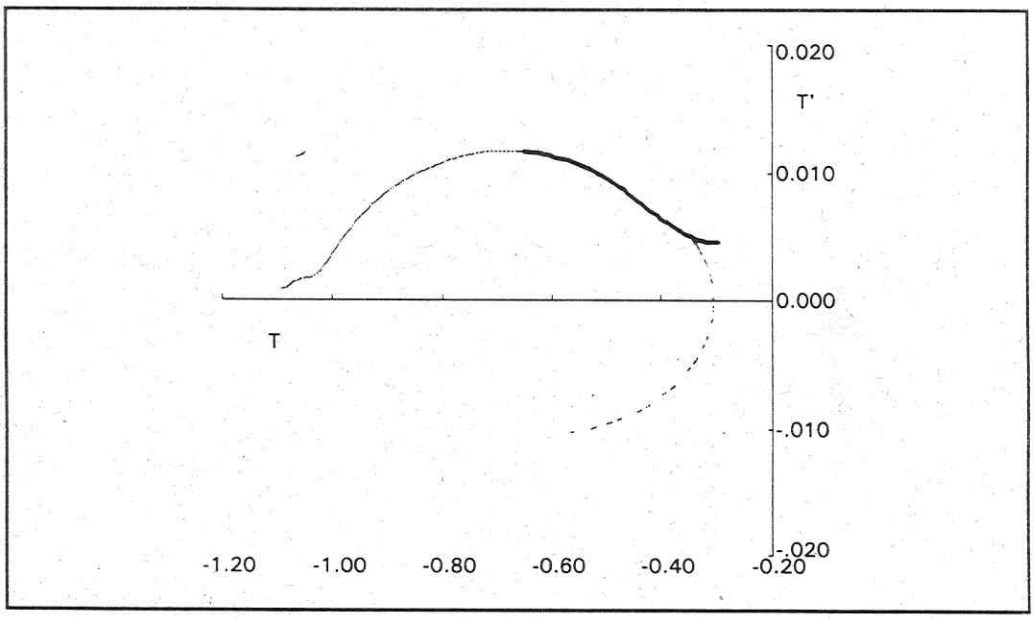


Abb. 3-8: Phasenraumbild für eine Extrapolation der Trendfunktion nach dem Hyperzykloidenmodell.

Mit dieser harmonischen Trendextrapolation ergibt sich eine von der linearen Extrapolation deutlich unterschiedliche Prognose des Reihenverlaufs für den 7. Tag, die in Abb. 3-9 gezeigt ist. Dabei ist die Gesamtprognose wieder aus den einzelnen Komponenten (Trend und Saisonalitäten) zusammengesetzt worden.

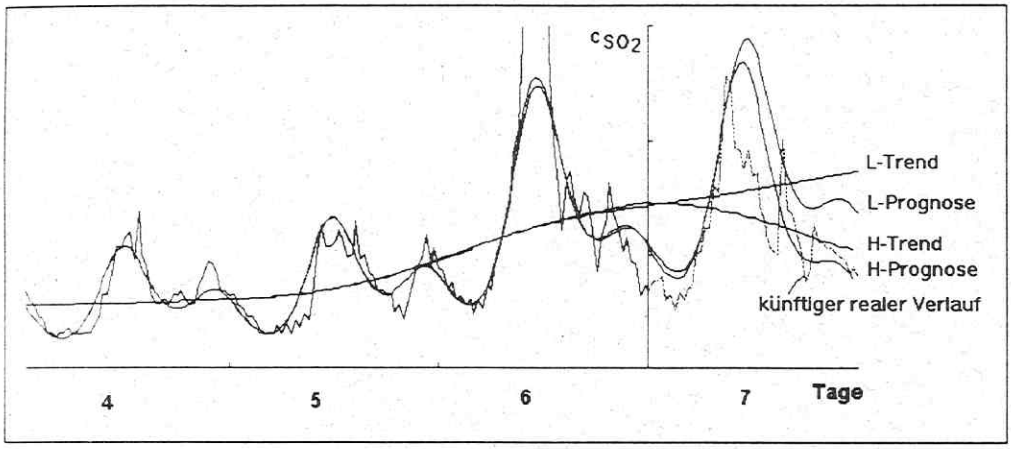


Abb. 3-9: Vergleich zwischen linearer "L" und lokaler harmonischer "H" Trendextrapolation am Beispiel der Tage 4, 5, 6 und 7 der CITY-Reihe.

Rein visuell kommt man an dem in Abb. 3-9 dargestellten Beispiel zu der Einschätzung, daß eine durch lokale harmonische Approximation modifizierte Extrapolation der Trendkomponente die im Nachhinein gemessene Datenreihe adäquater beschreibt, als bei linearer Trendextrapolation.

Wie wirkungsvoll die hier entwickelte Methode bei der Vorhersage von Smogsituationen arbeitet, muß jedoch erst noch an einer Vielzahl von Prognosefällen statistisch untersucht werden. Dies geschieht im Teil 4. Voraussetzung für eine solche Statistik ist dabei die klare Festlegung, welche Immissionssituationen für eine Vorhersage überhaupt interessant sind.

4. Teil: Das Prediktionsproblem

1. Definition und Lösungsansatz

In der sächsischen Smog-Verordnung (1991) sind eine ganze Reihe von notwendigen Bedingungen für die Auslösung der Smog-Vorwarnstufe genannt. Um die Untersuchungen überschaubar zu gestalten, soll in der vorliegenden Arbeit ein Kriterium herausgegriffen werden: Die Überschreitung der Schwelle 0,6 mg/m^3 durch das gleitende Mittel über drei Stunden (M3). Von Interesse ist nun die Frage, ob in den folgenden 24 Stunden ein Andauern dieser Situation nicht ausgeschlossen werden kann. Wenn dies der Fall ist, dann wird in dieser Zeit das gleitende Mittel über 24 Stunden (M24) die Schwelle 0,6 mg/m^3 ebenfalls erreichen. Das 24-Stunden-Mittel ist eine umweltmedizinisch relevante Kenngröße (vgl. Herbarth & Bredel (1990)) und wird deshalb als zweites Kriterium verwendet.

Mit Hilfe dieser beiden Kriterien werden folgende Begriffe festgelegt:

Definition (4-1): Ein **Nullfall** ist eine Immissionssituation, in deren Verlauf stets M3 < 0,6 mg/m^3 ist. Da diese Situationen für eine Vorwarnung nicht von Interesse sind, bleiben sie im weiteren unberücksichtigt.□

Definition (4-2): Ein **Pseudofall** ist eine Immissionssituation, in deren Verlauf M3 die Schwelle 0,6 mg/m^3 überschreitet, aber stets M24 < 0,6 mg/m^3 ist.□

Definition (4-3): Ein **Überschreitungsfall** ist eine Immissionssituation, in deren Verlauf M3 die Schwelle 0,6 mg/m^3 überschreitet und M24 innerhalb der folgenden 24 Stunden ebenfalls die Schwelle 0,6 mg/m^3 überschreitet.□

Definition (4-4): Der **Prognosepunkt** ist der Zeitpunkt t_0, an dem M3 = 0,6 mg/m^3 ist.□

Definition (4-5): Die **Prognoseaufgabe** besteht darin, für eine gegebene Immissionssituation am Prognosepunkt eine Wahrscheinlichkeitsaussage über das künftige Eintreten eines Überschreitungsfalls (bzw. Pseudofalls) zu machen.□

Die Wahrscheinlichkeitsaussage für die Prognose ist also aus dem Immissionsverlauf vor dem Prognosepunkt zu gewinnen. Um ein Bild über den gesamten Konzentrationsverlauf der verschiedenen Situationen zu bekommen, wurden für die Meßstelle Leipzig-Mitte von 1980 bis 1993 alle Pseudo- und alle Überschreitungsfälle herausgesucht und in den Abbn. 4-1 und 4-2 dargestellt. Der interessierende Zeitbereich umfaßt 48 Stunden vor und 24 Stunden nach dem Prognosepunkt. Anhang F enthält eine Liste dieser hier untersuchten Situationen.

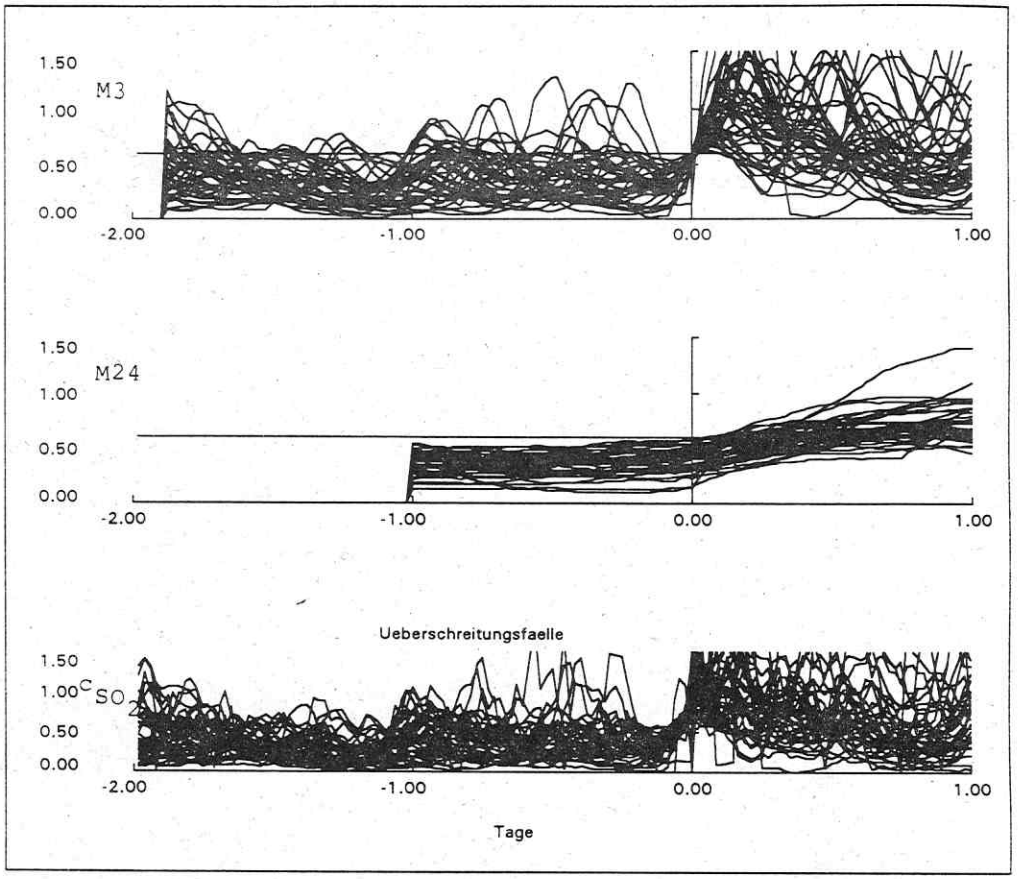


Abb. 4-1: Die 34 Überschreitungsfälle an der Meßstelle Leipzig-Mitte in der Zeit von 1980 bis 1993 (unten) und deren gleitende 24- und 3-Stunden-Mittel (mitte und oben)

Entsprechend der Prognoseaufgabe sind nun vor dem Prognosepunkt ($t<t_0=0$) gelegene Parameter zu finden, aus denen sich die Wahrscheinlichkeit dafür bestimmen läßt, daß sich die vorliegende Situation zu einem späteren Überschreitungsfall entwickelt.

Definition (4-6): Ein **Prediktor** ist ein Parameter, der aus dem Verlauf der Meßreihe bis zum Prognosepunkt gewonnen wird und mit dessen Hilfe eine Vorhersage über das Verhalten der Reihe nach dem Prognosepunkt möglich ist.□

Solche Parameter können zum einen geeignete Verlaufseigenschaften der Meßreihe sein. Es handelt sich dann um faktische Prediktoren. Andererseits ist es auch möglich, ein Zeitreihenmodell anzupassen und mit diesem Modell Parameter vorherzusagen. Diese sollen prognostizierte Prediktoren heißen.

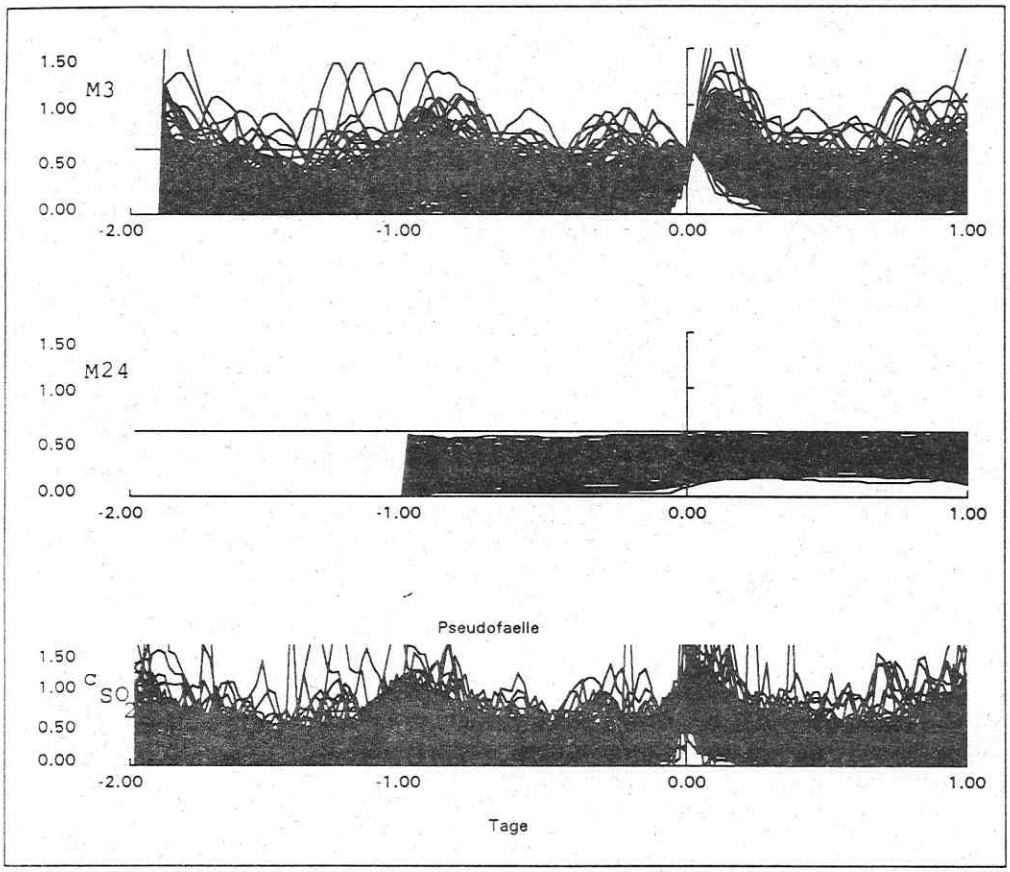


Abb. 4-2: Die 157 Pseudofälle an der Meßstelle Leipzig-Mitte in der Zeit von 1980 bis 1993 (unten) und deren gleitende 24- und 3-Stunden-Mittel (mitte und oben).

Für die Prognoseaufgabe ist lediglich die Identifizierung einer Situation als Überschreitungs- oder Pseudofall wesentlich. Deshalb kann man jeweils alle Überschreitungs- und alle Pseudofälle zu einer Menge zusammenfassen, die mit $\mathbf{U}$ und $\mathbf{P}$ bezeichnet werden.

Bemerkung (4-1): Nicht mehr einzelne Reihenverläufe, sondern diese bezüglich der Grenzwertüberschreitung zusammengefaßten Mengen sind jetzt der Untersuchungsgegenstand des 4. Teils der Arbeit. Mit statistischen Methoden wird versucht, Parameter (Kenngrößen) zu finden, durch die sich die Mengen charakterisieren und unterscheiden lassen. Da die Parameter bereits vor Kenntnis der Zuordnung einer Situation zu einer dieser Mengen bestimmt werden sollen, sind sie Prediktoren.□

Durch einen Prediktor wird nun eine Berechnungsvorschrift gegeben, die jedem einzelnen Fall eine reelle Zahl zuordnet. Eine Einteilung dieser Prediktorwerte in z.B. zwei Klassen bewirkt im allgemeinen eine Aufteilung aller Fälle und somit der Mengen **P** und **U** in der in Abb. 4-3 illustrierten Art.

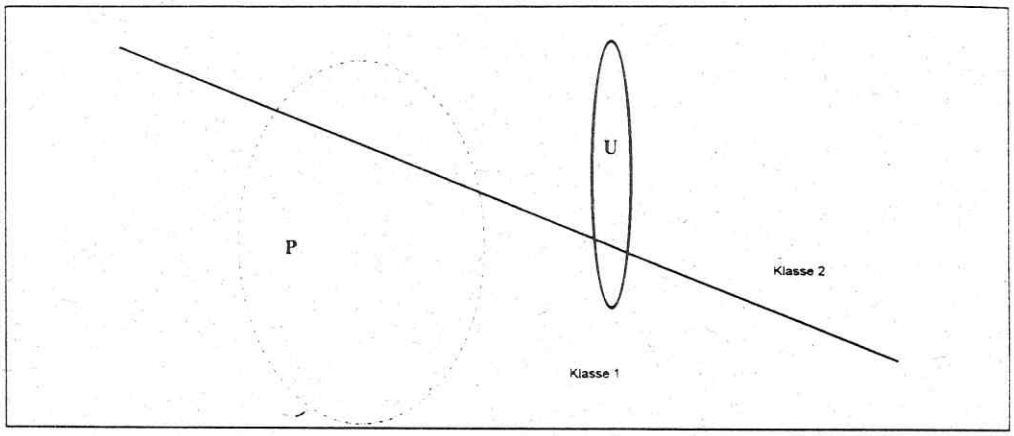


Abb. 4-3: Aufteilung der Mengen **P** und **U** durch zwei Klassen eines Prediktors

Da mit dem Prediktor der Typ eines vorliegenden Falles vorhergesagt werden soll, ist er so zu konstruieren, daß er die beiden Mengen bestmöglich unterscheidbar macht. In Abb. 4-4 ist gezeigt, wie durch einen idealen Prediktor mit zwei Kiassen die Pseudo- von den Überschreitungsfällen getrennt werden: Alle in Klasse 1 liegenden Fälle sind Pseudofälle. Befindet sich dagegen ein Fall in Klasse 2, so stellt er eine Überschreitungssituation dar.

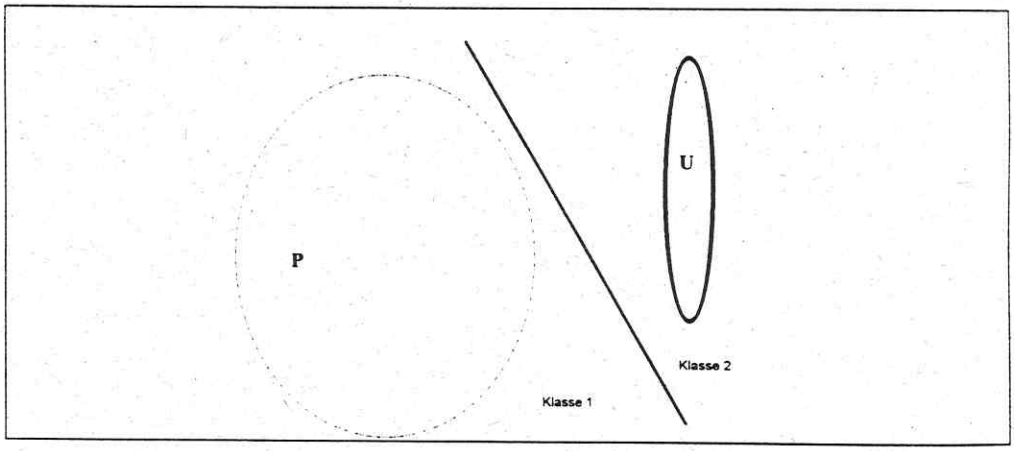


Abb. 4-4: Aufteilung der Mengen **P** und **U** durch zwei Klassen eines idealen Prediktors

Im allgemeinen wird man dieses Ideal nicht erreichen können. Durch Verwendung mehrerer Klassen und weiterer Prediktoren werden beide Mengen weiter aufgespalten. Bei zwei Prediktoren und vier Klassen könnte dies folgendermaßen aussehen (Abb. 4-5):

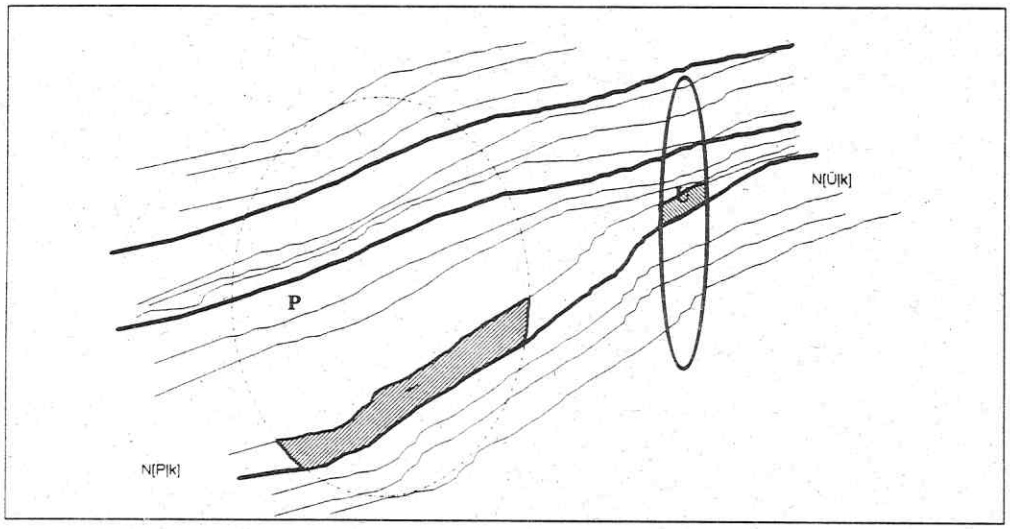


Abb. 4-5: Schematische Veranschaulichung der Aufteilung der Mengen **P** und **U** durch zwei Prediktoren zu je vier Klassen.

Prediktor π_1 bildet 4 Klassen, die durch die dicken Linien voneinander getrennt werden. Jede dieser Klassen wird durch den Prediktor π_2 (dünne Linien) wieder in 4 Klassen aufgespalten. Die Anzahl aller in einem herausgegriffenen Klassenbereich (k_1, k_2) gelegenen Fälle ist dunkel markiert und mit N[P|**k**] bzw. N[Ü|**k**] bezeichnet.

Allgemein lassen sich die p Prediktoren π_i (i=1,..,p) zu dem Vektor $\boldsymbol{\pi} = (\pi_1 \ \pi_2 \ \pi_3 \ ... \ \pi_p)^T$ zusammenfassen, ebenso die Klassen $\mathbf{k} = (k_1 \ k_2 \ k_3 \ ... \ k_p)^T$ für diese Prediktoren. Es bezeichne weiterhin N die Anzahl aller betrachteten Fälle und N[P|**k**] bzw. N[Ü|**k**] die Anzahl aller Pseudo- bzw. Überschreitungsfälle in dem Klassenbereich **k**.

Die Wahrscheinlichkeit, daß sich für einen konkreten Fall der Prediktorenvektor $\boldsymbol{\pi}$ in **k** befindet berechnet sich dann zu

$$P(\mathbf{k}) = (N[P|\mathbf{k}] + N[Ü|\mathbf{k}]) / N. \qquad (4\text{-}1)$$

2. Faktische Prediktoren

Wintersmogsituationen entstehen beim gleichzeitigen Auftreten von austauscharmer Wetterlage und ungeminderter Schadstoffemission. Sie sind also alle durch ähnliche Auslösebedingungen charakterisiert. Davon ausgehend kann man vermuten, daß es durch die übereinstimmenden Entwicklungsbedingungen bei den Überschreitungsfällen im Vorfeld zu Konzentrationsverläufen kommt, die einander sehr ähnlich sind und sich von denen der Pseudofälle unterscheiden.

Diese Vermutung läßt sich formulieren als

Arbeitshypothese: Für die Mengen $\mathbf{U}$ und $\mathbf{P}$ existieren Prediktoren, mit denen sich die beiden Mengen trennen lassen.□

Sind diese Prediktoren gefunden, so ist damit die Arbeitshypothese bestätigt und eine Prognose realisierbar. Es ist also nach für eine Vorhersage effizienten Parametern des Konzentrationsverlaufs zu suchen. Dazu wird aus allen in Abb. 4-1 dargestellten Überschreitungsfällen der Station Leipzig-Mitte im Zeitraum von 1980 bis 1993, also allen 34 in $\mathbf{U}$ enthaltenen Elementen durch arithmetische Mittelung ein mittlerer Konzentrationsverlauf gebildet. Der Streubereich läßt sich zu jedem Zeitpunkt (verteilungsfrei) durch das 0,1- bzw. 0;9-Perzentil charakterisieren. Bei den 157 Pseudofällen wird analog vorgegangen.

Das Ergebnis zeigt Abb. 4-6. In Abb. 4-7 bzw. Abb. 4-8 ist anstelle der Halbstundenmeßwerte das gleitende 24- bzw. 3-Stunden-Mittel aufgetragen.

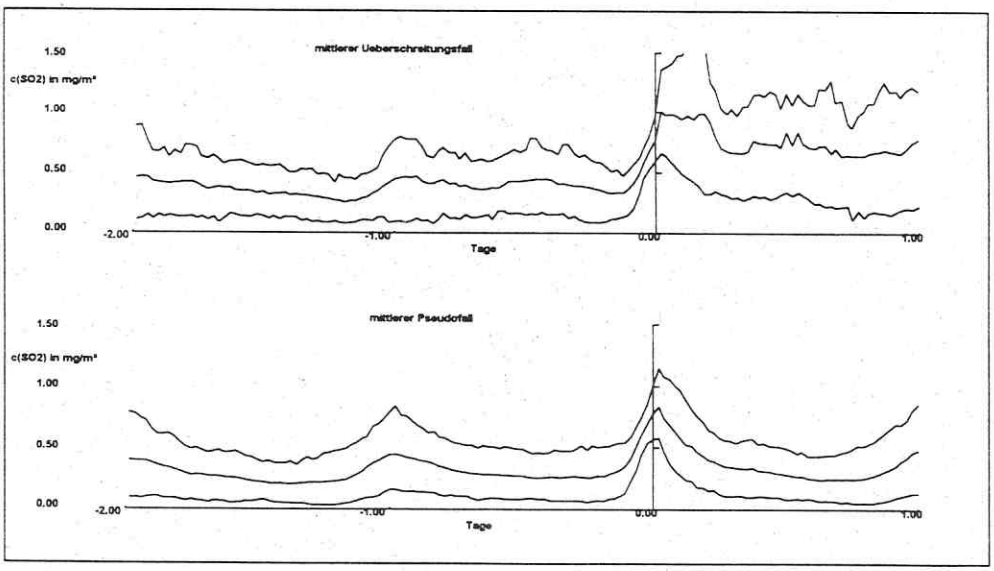


Abb. 4-6: Mittlere Realisierungen zu den in den Abbn. 4-1 und 4-2 gezeigten Fällen (0,1-Perzentil, Mittelwert, 0,9-Perzentil)

An diesen Darstellungen läßt sich ablesen, worin sich der mittlere Pseudo- vom mittleren Überschreitungsfall im Bereich vor t_0 anders verhält. Die gefundenen Unterschiede werden dann für die Aufstellung von Prediktoren genutzt:

Zum einen unterscheiden sich die mittleren Fälle in dem Wert, den das gleitende Mittel über die letzten 24 Stunden am Prognosepunkt annimmt (Parameter $M24(t_0)$). Für die Pseudofälle findet man am Prognosepunkt einen kleineren M24-Wert als für die Überschreitungsfälle (Abb. 4-7). Dieser Parameter soll deshalb als erster Prediktor verwendet werden.

Die Streubereiche für diesen Prediktor im Pseudo- und im Überschreitungsfall überschneiden sich jedoch (wie Abb. 4-7 zeigt), so daß keine eindeutige Aussage entsteht. Deshalb wird nach einem weiteren Prediktor gesucht.

Betrachtet man den Verlauf von M3 (Abb. 4-8), so zeigt sich jeweils ein typisches Verhalten in den letzten 3 Stunden vor t_0. Aus diesem Grund soll der Anstieg von M3 über die letzten 3 Stunden (mit M3A3 bezeichnet) als zweiter Prediktor verwendet werden. In Abb. 4-9 sind die beiden Prediktoren für einen Beispielfall eingezeichnet.

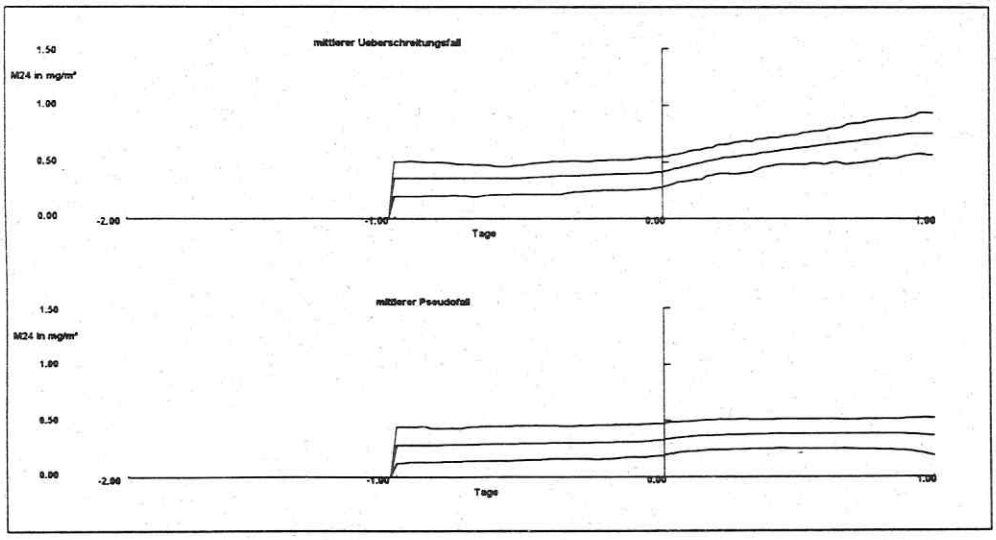


Abb. 4-7: Mittlere Realisierungen zu den M24 der in den Abbn. 4-1 und 4-2 gezeigten Fälle (0,1-Perzentil, Mittelwert, 0,9-Perzentil)

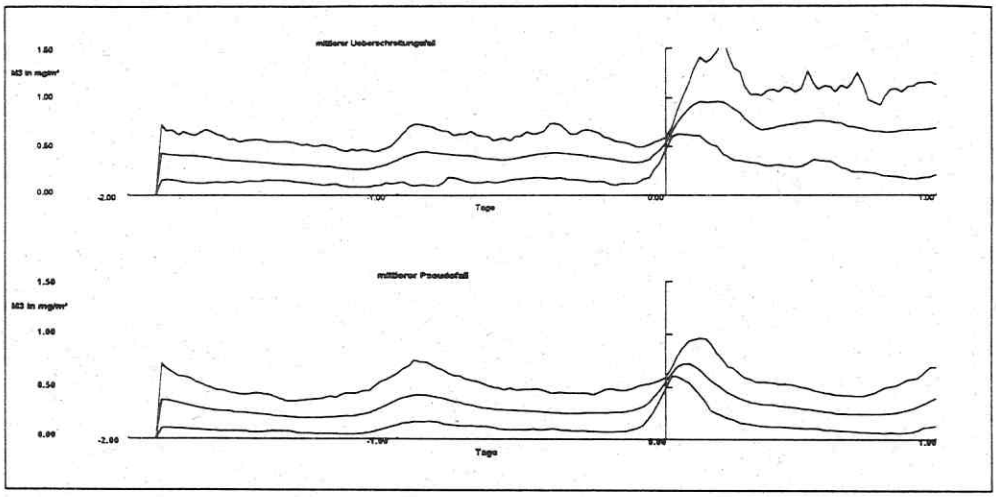


Abb. 4-8: Mittlere Realisierungen zu den M3 der in den Abbn. 4-1 und 4-2 gezeigten Fälle (0,1-Perzentil, Mittelwert, 0,9-Perzentil)

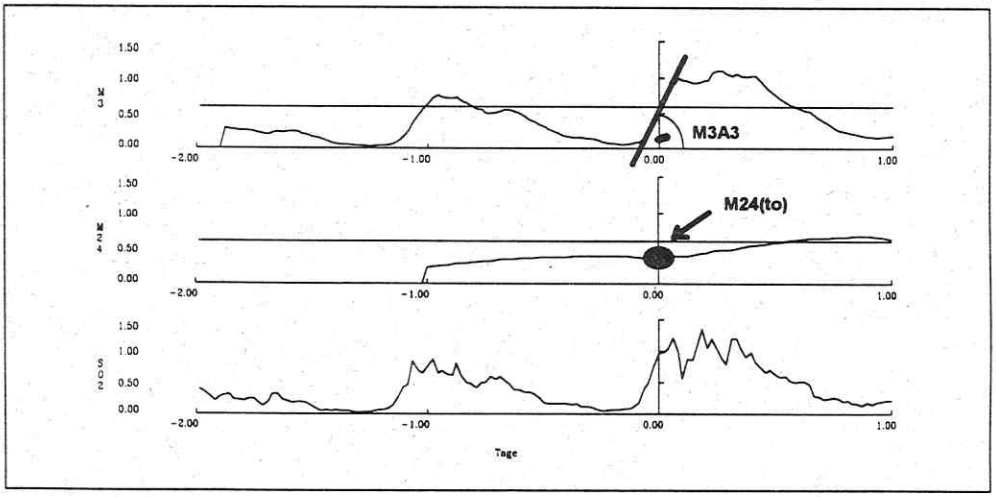


Abb. 4-9: Die verwendeten Prediktoren M24 und M3A3 anhand eines Beispiels für einen Überschreitungsfall

Damit lassen sich für jeden Fall zwei Werte berechnen, die dann einen Punkt in der von M24 und M3A3 aufgespannten Prediktorenebene beschreiben. In Abb. 4-10 wurden so alle Pseudo- und Überschreitungsfälle mit unterschiedlichen Symbolen markiert.

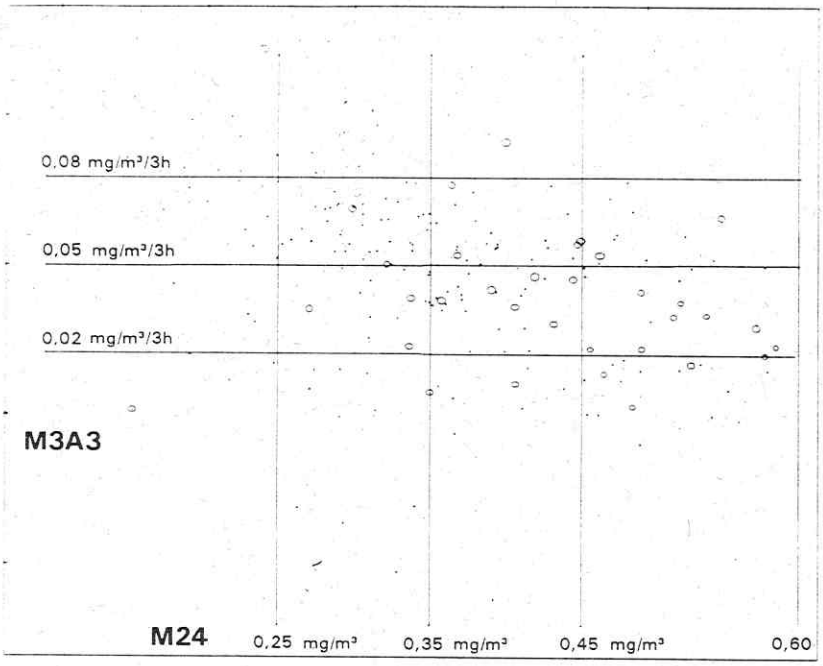


Abb. 4-10: Prediktorenebene für die Prediktoren M24(t_0) und M3A3(t_0). Die Pseudofälle sind durch Punkte, die Überschreitungsfälle durch Kreise markiert

Zur Interpretation der Verhältnisse in der Prediktorenebene soll auf die eingangs (in Abschnitt 1) gemachten Überlegungen zurückgegriffen werden. Dort wurden die Prediktorwerte in zwei Klassen eingeteilt, um damit eine Zuordnung zu den Mengen **U** und **P** zu bekommen (Abb. 4-3).

In Abb. 4-10 sind nun die Überschreitungsfälle durch Kreise und die Pseudofälle durch Punkte markiert worden. Sie bilden die Mengen **U** und **P**. Beide Mengen überlappen und durchdringen sich so stark, daß eine einfache Trennung durch Bildung von zwei Klassen nicht möglich ist. Dies hätte nur bei einer Lokalisierung beider Mengen in separaten Bereichen der Prediktorenebene, also einer Clusterbildung funktioniert.

Um eventuell doch bestehende Unterschiede in der Lage von **U** und **P** herauszuarbeiten, wird jeder Prediktorbereich in vier Klassen eingeteilt (vgl. Abb. 4-10). Die Klassengrenzen sind willkürlich festgelegt worden und orientieren sich an Mittelwert und Streubereich. Zusätzlich ist die Schranke von 0,6 mg/m^3 eingezeichnet, welche entsprechend der Definition der Prognosesituation von M24 zum Zeitpunkt t_0 noch nicht überschritten werden kann. Sie wird allerdings bei den Überschreitungsfällen innerhalb der nächsten 24 Stunden erreicht.

Für die so entstehenden Klassenbereiche **k** lassen sich folgende statistischen Angaben machen:

- die Anzahl der in **k** liegenden Überschreitungs- und Pseudofälle (N[Ü|**k**] und N[P|**k**]),

- die Wahrscheinlichkeit, einen Fall in dem Bereich **k** zu finden

$$P(\mathbf{k}) = \left(N[Ü|\mathbf{k}] + N[P|\mathbf{k}]\right) / \left(\Sigma_k \left(N[Ü|\mathbf{k}] + N[P|\mathbf{k}]\right)\right) \qquad (4\text{-}2)$$

- die Wahrscheinlichkeit, daß ein in **k** befindlicher Fall ein Überschreitungsfall ist

$$P(Ü|\mathbf{k}) = N[Ü|\mathbf{k}] / \left(\left(N[Ü|\mathbf{k}] + N[P|\mathbf{k}]\right) \right) \qquad (4\text{-}3)$$

- die Wahrscheinlichkeit, daß ein Überschreitungsfall in den Bereich **k** fällt

$$P(\mathbf{k}|Ü) = \left(P(\mathbf{k}) \cdot P(Ü|\mathbf{k})\right) / \left(\Sigma_k \left(P(\mathbf{k}) \cdot P(Ü|\mathbf{k})\right)\right) \qquad (4\text{-}4)$$

- die a priori Treffsicherheit, welche sich aus $\max(P(Ü|\mathbf{k}),\ (1-P(Ü|\mathbf{k})))$ abschätzen läßt.

Liegen nämlich für einen gegebenen Fall die Werte der Prediktoren, also der Prediktorenvektor $\pi = (\pi_1\ \pi_2\ \pi_3\ \ldots\ \pi_p)^T$, in dem Klassenbereich **k** und ist $P(Ü|\mathbf{k}) > 0.5$, so wird dieser Fall mit der Wahrscheinlichkeit von $P(Ü|\mathbf{k})$ ein Überschreitungsfall. Ist dagegen $P(Ü|\mathbf{k}) \leq 0.5$, dann wird es mit der Wahrscheinlichkeit $(1-P(Ü|\mathbf{k}))$ einen Pseudofall geben.

Diese Wahrscheinlichkeiten geben also gewissermaßen eine a priori Treffsicherheit der Prognose dadurch an, daß sie ausdrücken, welcher Anteil der früheren Fälle mit der gemachten Prognose in Einklang steht, sich also zur Erklärung der Prognose heranziehen läßt.

Die a priori Treffsicherheit ist umso größer, je näher $P(Ü|\mathbf{k})$ bei 0 oder 1 liegt, d.h. je dominanter die Pseudo- oder die Überschreitungsfälle im Klassenbereich **k** sind.

Für die Prediktorenebene der Abb. 4-10 sind die statistischen Angaben in Tab. 4-1 zusammengefaßt. Die beiden Klassenbereiche, in denen $P(Ü|\mathbf{k}) > 0.5$ gilt, also ein Überschreitungsfall vorhergesagt würde, sind durch Umrahmung der Wahrschein- lichkeitsangabe hervorgehoben worden.

Dabei fällt auf, daß sich in einem der Bereiche die Angaben auf nur einen Überschreitungsfall gründen. Die Aussage in diesem Bereich kann also nicht besonders sicher sein. Aus diesem Grunde soll im weiteren die a priori Treff- sicherheit TS mit der Belegungswahrscheinlichkeit $P(\mathbf{k})$ gewichtet werden:

$$TS(\mathbf{k}) = P(\mathbf{k}) \cdot \max(P(Ü|\mathbf{k}),\ (1-P(Ü|\mathbf{k}))). \qquad (4\text{-}5)$$

$$TS_{total} = \Sigma_k\, TS(\mathbf{k}) \qquad (4\text{-}6)$$

ist dann die totale a priori Treffsicherheit dieser Vorhersagemethode.

Tabelle 4-1:

Grenzbedingungen, Wahrscheinlichkeiten und a priori Treffsicherheiten für die Prognose mit den Prediktoren M24(to) [in mg/m³] und M3A3 [in mg/m³/3h]

M3A3 \ M24	< 0,25 N(P)(N(Ü))	0,25 ... 0,35	0,35 ... 0,45	> 0,45
> 0,08	11(1)	8(0)	0(1)	0(0)
0,05 ... 0,08	7(0)	35(2)	16(3)	10(3)
0,02 ... 0,05	3(0)	12(3)	13(6)	6(9)
< 0,02	3(1)	11(1)	7(1)	15(3)

P(k)	0,063	0,042	0,005	0
	0,037	0,2	0,1	0,068
	0,016	0,078	0,1	0,078
	0,021	0,063	0,042	0,094
P(Ü|k)	0,083	0	1	
	0	0,054	0,158	0,231
	0	0,2	0,316	0,6
	0,25	0,083	0,125	0,167
P(k|Ü)	0,03	0	0,03	
	0	0,06	0,09	0,09
	0	0,09	0,18	0,26
	0,03	0,03	0,03	0,09
a priori TS(k)	0,058	0,042	0,005	
	0,037	0,183	0,084	0,052
	0,016	0,062	0,068	0,047
	0,016	0,058	0,037	0,078

Die Diskussion der Vorhersage soll am Beispiel des Klassenbereiches k_{34} (M24 > 0.45 mg/m³ und 0.02 mg/m³/3h < M3A3 ≤ 0.05 mg/m³/3h) geführt werden:

Liegt für einen neuen Fall der Prediktorenvektor π = (M24 M3A3)^T in k_{34}, dann lautet wegen $P(Ü|k_{34})$ = 0.6 die Prognose auf "Überschreitungsfall". Diese Prognose basiert auf den 9 in k_{34} bereits realisierten Überschreitungsfällen, welche 4.7% aller Fälle sind. Bei anderen Klassenbereichen lautet die Prognose auf "Pseudofall". Alle weiteren Überlegungen gehen analog.

Betrachtet man die Prognosen für alle Klassenbereiche, dann werden sie entsprechend der totalen a priori Treffsicherheit, der Summe aller TS(**k**) durch 85% aller Fälle bestimmt.

Vor besonderem Interesse für die Smogwarnung sind natürlich die Überschreitungssituationen. Solche Situationen sind seltene Ereignisse. Beispielsweise waren unter den an der Station Leipzig-Mitte von 1980-1993 insgesamt auftretenden 191 Prognosefällen (siehe Anhang F) 18% Überschreitungsfälle. Spektakuläre Trefferraten lassen sich bei der Vorhersage solcher seltenen Ereignisse nicht erwarten. Dafür liegt diese Aufgabe dem Grenzbereich objektiv möglicher Vorhersagbarkeit zu nahe.

In solchen Fällen ist es schwer, allein mit Trefferraten z.B. die Prognose selbst nur eines seltenen Ereignisses angemessen zu würdigen; denn sowohl nach täglicher Erfahrung, als auch gemäß der Informationstheorie (z.B. Klaus (1969)) besitzt die Kenntnis vom Eintreten eines Ereignisses einen umso größeren Wert, je unwahrscheinlicher es ist (Balzer (1989)). Auch wenn die Güte der in diesem 4. Teil vorgeschlagenen Prognosealgorithmen mit Hilfe ihrer Treffsicherheit bei der Vorhersage von Überschreitungssituationen charakterisiert wird, sollte der genannte Aspekt bei der Beurteilung nicht vergessen werden.

Da es lediglich in k_{34} und k_{13} zur Vorhersage einer Überschreitung kommt, beträgt die a priori Treffsicherheit dieser Prognose mit den Prediktoren M24 und M3A3 ca. 5%. Zur Abschätzung der Güte eines Vorhersagealgorithmus wäre jedoch eine a posteriori Treffsicherheit, die das Ergebnis einer neuen Realisierung mit ihrer zuvor gestellten Prognose vergleicht, zu bestimmen. Das bedeutet, daß man grundsätzlich zwischen den Daten zur **Kalibrierung** des Modells und den Daten zur **Validierung** der gestellten Prognose unterscheiden sollte.

Bei der Validierung entsteht dann die a posteriori Treffsicherheit, die den Anteil der richtig vorhergesagten Fälle unter den Validierungsdaten beschreibt. Sie ist im allgemeinen kleiner, als die a priori Treffsicherheit.

Bei den hier getesteten faktischen Prediktoren handelt es sich um sehr einfache Kenngrößen. Welche Ergebnisse ein mit Hilfe der im Teil 2 erläuterten Zeitreihenmodellierung bestimmter prognostizierter Prediktor liefert, soll im folgenden Abschnitt untersucht werden.

3. Prognostizierter Prediktor

Neben den aus der faktischen Zeitreihe gewonnenen Prediktoren kann man zur Lösung der gestellten Aufgabe auch Parameter verwenden, die aus einem prognostizierten Verlauf der Schadstoffkonzentration abgeleitet werden. Dazu wird an die gemessenen Daten ein Zeitreihenmodell angepaßt und mit diesem Modell eine Extrapolation durchgeführt.

Als Prediktor soll hier die Zeitspanne Δt verwendet werden, die vom Prognosepunkt t=0 bis zu der mit dem Modell vorhergesagten Überschreitung vergeht (vgl. Abb. 4-11). Ein Überschreitungsfall wird dann vorhergesagt, wenn das sich aus M24(Zeitreihenprognose) ergebende $\Delta t < 24$ Std. ist. Diese Festlegung orientiert sich an der Definition (4-3) des Überschreitungsfalls, bei dem die zwischen dem Prognosepunkt (M3 = 0,6 mg/m³) und dem "Überschreitungspunkt" (M24 = 0,6 mg/m³) vergehende Zeit weniger als einen Tag betragen soll.

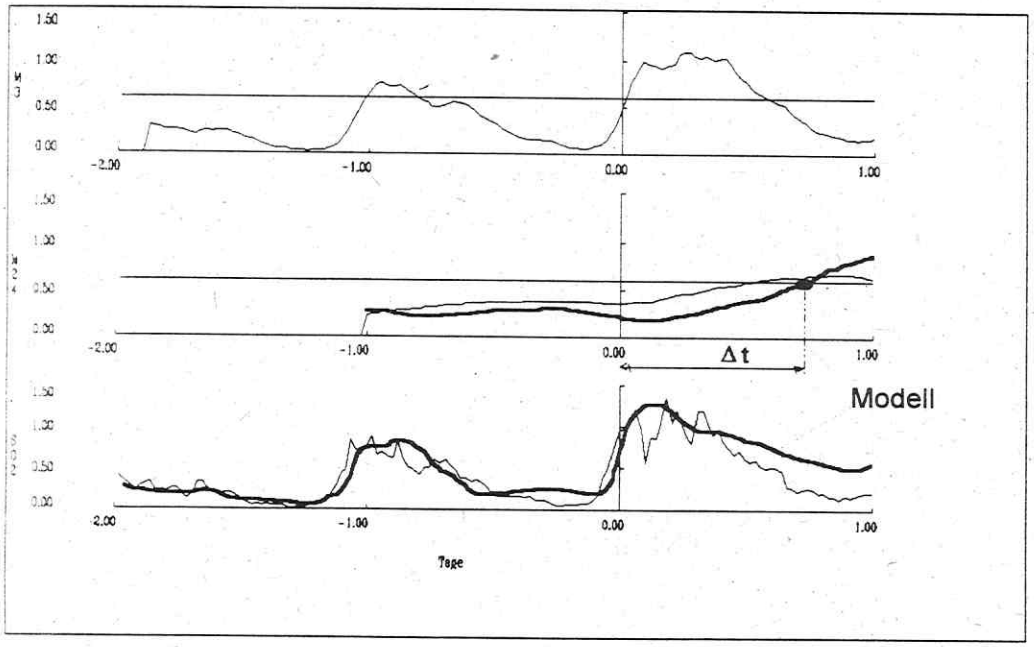


Abb. 4-11: Der prognostizierte Prediktor Δt im Falle einer Beispielsituation; die modellierte Zeitreihe ist durch eine dicke Linie markiert.

Im Teil 2, Abschnitt 5 ist für die Immissionsreihe bereits ein Komponentenmodell aufgestellt worden. Damit ist man nun in der Lage, den Verlauf der Meßreihe für eine gewisse Zeit vorherzusagen. Aus dieser Zeitreihenprognose läßt sich dann auch der voraussichtliche Zeitpunkt für eine Schwellwertüberschreitung durch M24 bestimmen, so daß der prognostizierte Prediktor Δt berechnet werden kann. Für die Untersuchung wurden die Prognosesituationen der Station Leipzig-Mitte von 1980 bis 1993 verwendet, ein Abschnitt von zwei Tagen vor t=0 zugrunde gelegt und daraus die Zeitreihenprognose für die beiden folgenden Tage berechnet.

In den Abbn. 4-12 und 4-13 wurden für die Pseudo- und die Überschreitungsfälle die Histogramme von Δt aufgezeichnet. Die Fälle, in denen die Extrapolation des Modells an den beiden auf t=0 folgenden Tagen zu keiner Überschreitung der Schwelle 0.6 mg/m³ durch M24 kommt, sind im rechten Histogrammteil gesondert dargestellt.

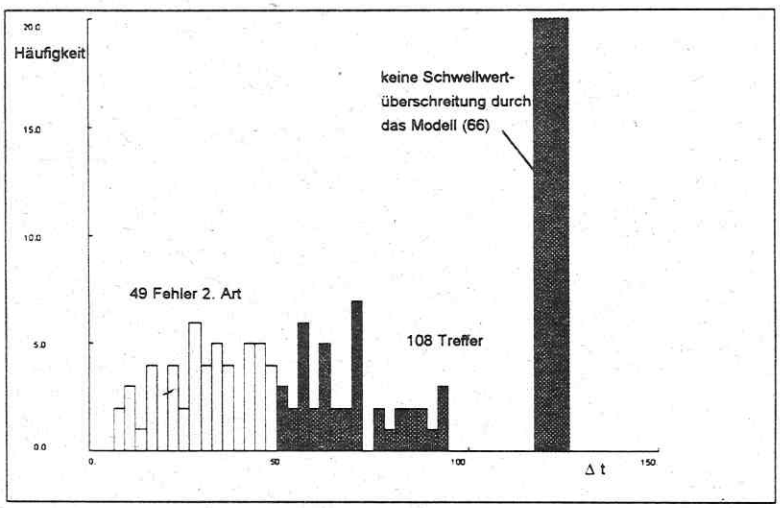


Abb. 4-12: Histogramm für den prognostizierten Prediktor Δt aller Pseudofälle (Fallzahlen in Klammern angegeben)

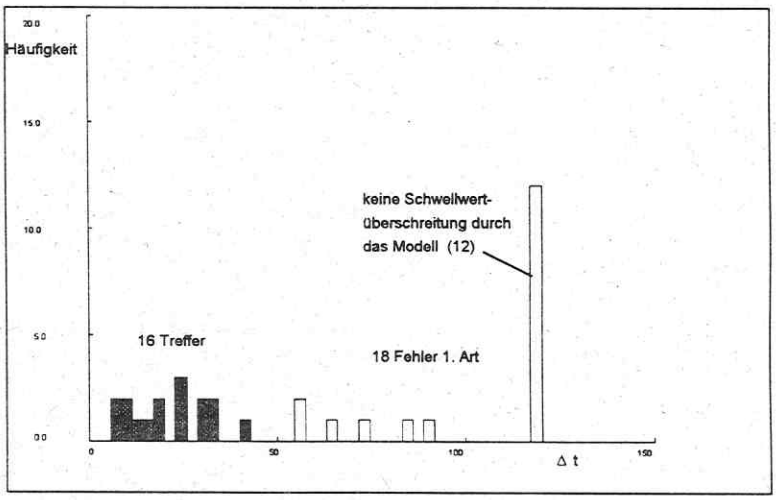


Abb. 4-13: Histogramm für den prognostizierten Prediktor Δt aller Überschreitungsfälle (Fallzahlen in Klammern angegeben)

Um aus dem prognostizierten Prediktor Δt eine Vorhersage ableiten zu können, wird wieder eine Klassenteilung des Wertebereichs von Δt eingeführt. Und zwar in Orientierung an der Definition (4-3) des Überschreitungsfalls die Klassen

$$k_1 = \{ 0<\Delta t\leq 48 \} \text{ und } k_2 = \{ \Delta t>48 \}. \qquad (4\text{-}7)$$

Für $\Delta t \in k_1$ wird dann ein Überschreitungsfall und für $\Delta t \in k_2$ ein Pseudofall vorhergesagt.

Die grau eingefärbten Balken geben jeweils die richtig prognostizierten Fälle an, also die Treffer. Ein Treffer liegt vor,

- wenn $\Delta t \leq 48$ Halbstundenschritte vorhergesagt wurde und dann auch ein Überschreitungsfall eintrat, oder
- wenn es innerhalb der auf t=0 folgenden 24 Stunden nicht zu einer Überschreitung kam, d.h., wenn das Modell $\Delta t > 48$ Halbstundenschritte vorhergesagt hatte und dann auch ein Pseudofall eintrat.

Daneben gibt es aber auch zwei Arten der Fehlprognose, die sich aufgliedern in

- Fehler 1. Art, wenn ein Pseudofall prognostiziert wurde ($\Delta t > 48$ Halbstundenschritte), jedoch real ein Überschreitungsfall eintrat, und
- Fehler 2. Art, wenn ein Überschreitungsfall vorhergesagt wurde ($\Delta t \leq 48$ Halbstundenschritte), tatsächlich aber ein Pseudofall realisiert wurde.

Die weißen Balken repräsentieren die Fehlprognosen. Ein Prediktor mit 100% Treffsicherheit hätte also jeweils lediglich die Δt-Werte in den grauen Balken geliefert.

Welche Güte hat nun die Vorhersage mit dem Prediktor Δt?

Die besonders interessierende Prognose eines Überschreitungsfalls ($\Delta t \in k_1$) wurde bei diesen insgesamt 191 Fällen 16+49 = 65 mal gestellt. Davon erwiesen sich 16 Fälle als Treffer, das bedeutet eine Treffsicherheit von 25%.

Da Δt ein prognostizierter Prediktor ist, das Modell also an den vor dem Prognosepunkt t_o gelegenen Meßdaten kalibriert und mit den auf t_o folgenden Meßdaten validiert wird, ist die berechnete Treffsicherheit bereits eine a posteriori gültige Kenngröße.

Sie soll mit der Treffsicherheit bei Niederschlagsvorhersagen verglichen werden. Da der Niederschlag (je nach geographischer Region) ebenfalls ein seltenes Ereignis sein kann, ergeben sich interessante Parallelen zwischen seiner Vorhersagegüte und der von Smogepisoden. Von Balzer (1984) wurde die in Abb. 4-14 dargestellte Abhängigkeit zwischen Vorhersagedistanz x [in Stunden], der Wahrscheinlichkeit p(E) des Niederschlagsereignisses E und der Treffsicherheit TS der Prognose angegeben. Ihr liegen 23 verschiedene Arten von Niederschlagsvorhersagen zugrunde.

Sie zeigt, daß es relativ gute Vorhersageleistungen im Mittelfristbereich (x > 80 Std.) gibt, sofern p(E) > 50% ist. Dagegen ist ein vergleichsweise rasches Nachlassen der Vorhersageleistung im Kurzfristbereich (x < 48 Stunden) zu verzeichnen, wenn p(E) < 10% ist.

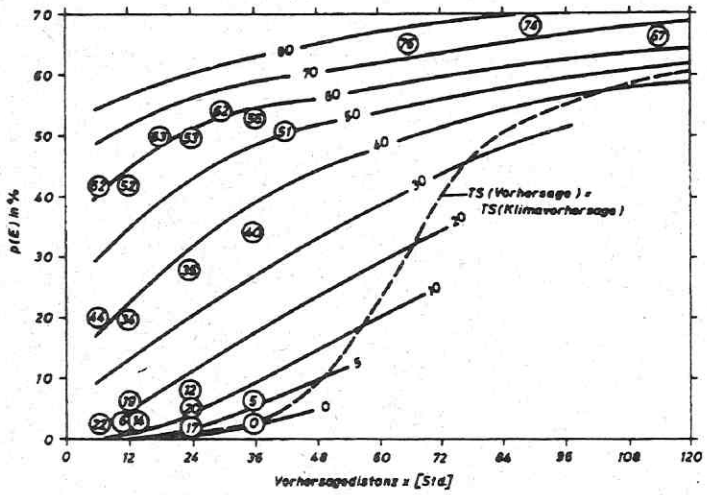


Abb. 4-14: Treffsicherheit (TS in %) von 23 verschiedenen Arten von Niederschlagsvorhersagen (durch die Kreise markiert), die sich in der Vorhersagedistanz x und der Wahrscheinlichkeit des Ereignisses p(E) unterscheiden (aus Balzer (1984)).

Entnimmt man der Abb. 4-14 die Treffsicherheit für eine Wahrscheinlichkeit von Überschreitungsereignissen von p(E) = 18% und x = 24 Stunden, so erhält man TS ≈ 27%. Mit diesem Wert ist die in der vorliegenden Untersuchung bei der univariaten Smogprognose mit dem Prediktor Δt erreichte Güte in Übereinstimmung.

Wie lassen sich aber die auftretenden Fehlprognosen verstehen?

Die Fehler 1. Art kommen durch eine falsche Extrapolation des Zeitreihenmodells zustande. Hier wird offenbar Δt viel zu groß (Δt>48) vorhergesagt, was darauf hindeutet, daß die Zeitreihenprognose im längerfristigen Bereich nicht zufriedenstellend arbeitet.

Dies könnte möglicherweise durch eine verbesserte Trendprognose optimiert werden. Ganz ähnlich ist die Situation bei den Fehlern 2. Art: Δt wird viel zu klein vorhergesagt.

Der folgende Abschnitt untersucht, ob eine modifizierte Trendprognose den Anteil der Fehlvorhersagen verringert.

4. Mit lokaler harmonischer Trendextrapolation prognostizierter Prediktor

Mit der in Teil 3, Abschnitt 4 entwickelten lokalen harmonischen Trendextrapolation verfügt man über eine alternative Methode zur Berechnung des prognostizierten Prediktors, der zur Unterscheidung von dem mit dem reinen Komponentenmodell bestimmten Wert mit $\Delta t(h)$ bezeichnet werden soll. Für den direkten Vergleich von $\Delta t(h)$ mit Δt sind die bereits für Δt diskutierten Histogramme hier ebenfalls erstellt worden (Abbn. 4-15 und 4-16).

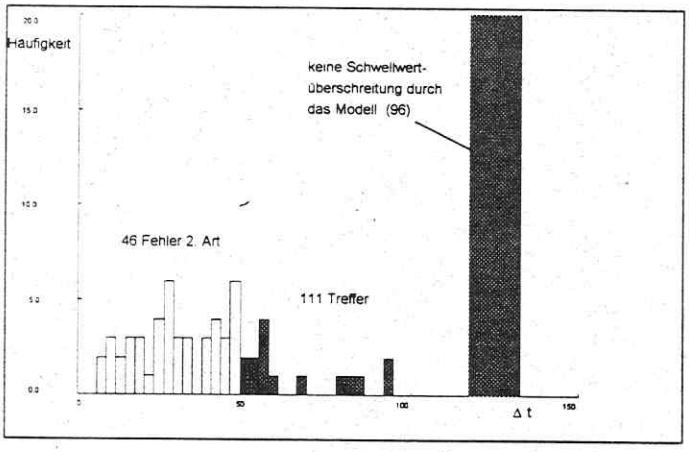


Abb. 4-15: Histogramm für den prognostizierten Prediktor $\Delta t(h)$ aller Pseudofälle (Fallzahlen in Klammern angegeben)

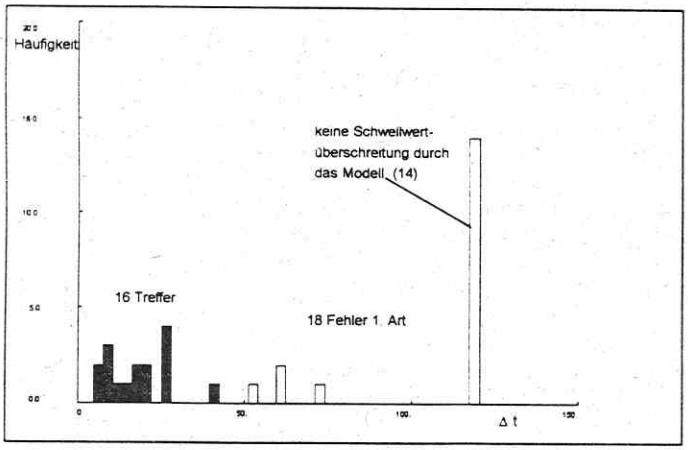


Abb. 4-16: Histogramm für den prognostizierten Prediktor $\Delta t(h)$ aller Überschreitungsfälle (Fallzahlen in Klammern angegeben)

Dabei lassen sich folgende Veränderungen erkennen:

- Die Trefferzahl bei den Überschreitungsfällen ist unverändert, ebenso die Anzahl der Fehlprognosen 1. Art.

- Der modifizierte Algorithmus hat lediglich bei den Pseudofällen die Trefferzahl um 3 erhöht und entsprechend die Zahl der Fehler 2. Art verringert. Dabei kam es aber zu einer deutlichen Umschichtung der Häufigkeiten bei den Treffern: Die Zahl der Fälle, in denen überhaupt keine Schwellwertüberschreitung durch das Modell vorhergesagt wurde, stieg um 30, was die Prognosehäufigkeit von Werten im Bereich $48 < \Delta t(h) < 96$ entsprechend reduzierte.

- Die Treffsicherheit bei der Vorhersage eines Überschreitungsfalls berechnet sich hier zu 16/(16+46) = 26%, im Vergleich zu 25% bei Prognose mit Δt.

Wie im Teil 3, Abschnitt 6 beschrieben wurde, nutzt die lokale harmonische Trendextrapolation die Krümmung des Trends für seine Fortsetzung mit aus. Damit ist sie umfassender, als die lineare Extrapolation. Warum die Prognosegüte jedoch hinter den damit verbundenen Erwartungen zurückbleibt, soll im 5. Teil analysiert werden.

5. Teil: Diskussion

1. Fehlerquellen und Grenzen der univariaten Trendextrapolation

Die bei der Zeitreihenmodellierung im Teil 2 eingeführte Abstraktion, die Prognoseaufgabe lediglich als univariate Problemstellung zu formulieren, hat natürlich ihre Konsequenzen. Der Prognosegüte sind damit prinzipielle Schranken gesetzt. Der Teil 5 soll anhand der in der vorliegenden Arbeit vorgeschlagenen Vorhersagealgorithmen verdeutlichen, an welchen Stellen und durch welche Effekte sie in ihrem Prognosehorizont begrenzt werden. Den Ausgangspunkt bildet dabei eine Analyse der Fehlprognosen bei der Trendextrapolation.

Zuvor soll jedoch eine interessante Eigenschaft der Trendkomponente diskutiert werden. Im Teil 1, Abschnitt 4 wurde die Autokorrelation der Immissionsreihe analysiert. Die Autokorreliertheit, also die Persistenzeigenschaft der SO_2-Reihe war ein Motiv für die Beschäftigung mit Algorithmen der univariaten Kürzestfristprognose. Bei der Zeitreihenmodellierung im 2. Teil wurde dann aber festgestellt, daß die Nutzung der Autokorrelation allein keine zufriedenstellenden Prognoseergebnisse bringt.

Die Persistenz beschreibt gewissermaßen die "Trägheit" der Schadstoffkonzentration des atmosphärischen Systems. Ihre Gegenspieler sind äußere Einflußgrößen, wie der Luftmassenstrom, eine verstärkte oder reduzierte Emission, die Veränderung der Luftschichtung, etc., die die Einstellung eines stabilen Gleichgewichtszustandes stören.

Es ist nun bemerkenswert, daß dieser Widerspruch bei der Komponentenzerlegung der Meßreihe primär an die Trendkomponente "vererbt" wird und sich dort drastisch verschäft. Zur Verdeutlichung dieses Tatbestandes betrachte man die in Abb. 5-1 vergleichend dargestellten Autokorrelationsfunktionen der CITY-Reihe und deren Trendkomponente.

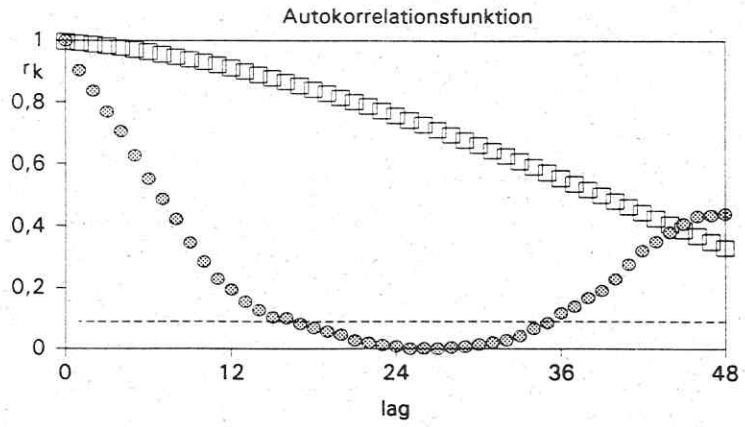


Abb. 5-1: Autokorrelationsfunktion der CITY-Reihe (Punkte) und ACF der Trendkomponente der CITY-Reihe (Kästchen)

Der Trend zeigt für Zeitabstände unter 24 Stunden eine stärkere Erhaltungsneigung, als die Meßdaten selbst. Andererseits war im Teil 3, Abschnitt 2 herausgearbeitet worden, daß gerade der Trendverlauf auf Wetteränderungen der synoptischen Scale reagiert. Da die saisonalen und die Restkomponente (Teil 2, Abschnitt 5) bis auf wenige Ausnahmen (auf die später noch kurz eingegangen wird) zeitlich relativ unverändert bleiben, muß sich eine Änderung des Immissionszustandes logischerweise vorrangig in der Trendkomponente ausdrücken. Dabei sind jedoch Zeiten von 24 Stunden und mehr charakteristisch.

Prägnanter ausgedrückt wird der Trend durch zwei widersprüchliche Phänomene bestimmt: Die Persistenz einerseits und die Wechselwirkung mit anderen atmosphärischen Vorgängen andererseits. Dies ist nur auf der Ebene von verschiedenen Zeitscales möglich!

Aufgrund dieser besonderen Bedeutung des Trends wurden zwei Methoden der Extrapolation ausprobiert: Die lineare, welche durch das Zustandsraummodell bereits automatisch vom für die Trendextraktion verwendeten Kalman-Algorithmus realisiert wird und eine harmonische mit lokal angepaßter Frequenz, welche aus dem Verhalten der Phasenraumtrajektorie des Trends entwickelt worden war.

Wie diese beiden Extrapolationen bei verschiedenen Verlaufstypen des Trends arbeiten, ist an den Beispielen in Abb. 5-2 vergleichend dargestellt.

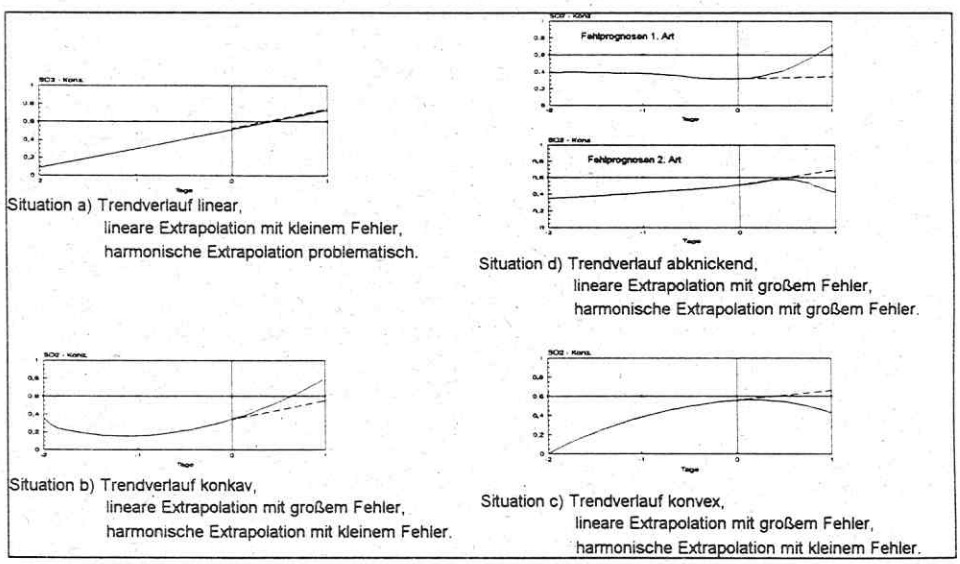


Abb. 5-2: Vergleich von linearer (gestrichelt) und lokaler harmonischer (durchgezogene Linie) Trendextrapolation (t>0) anhand ausgewählter Beispielsituationen.

Die Situation a) läßt sich mit einer linearen Trendextrapolation vorhersagen. Darüberhinaus kann der nahezu lineare Verlauf auch durch ein Stück einer Schwingung mit sehr großer Periode approximiert werden. Prinzipiell ist deshalb auch die harmonische Extrapolation anwendbar. Es entstehen lediglich numerische Probleme, da die Frequenz der anzupassenden Winkelfunktion äußerst klein wird.

Bemerkung (5-1): So waren die Schwierigkeiten bei der Schätzung sehr kleiner Frequenzen durch den in Teil 3, Abschnitt 5 vorgeschlagenen Algorithmus der lokalen harmonischen Extrapolation eine Ursache für die unvermutet geringe Verbesserung der Prognosegüte gegenüber der linearen Trendextrapolation. Im Abschnitt Parameterschätzung wurde bereits auf die Problematik der nichtlinearen Schätzgleichung hingewiesen. Ein optimales Verfahren würde die Extrapolation gleich mit in die Filterung einbeziehen, wie dies bei der linearen Extrapolation durch das Kalman-Filter geschieht.

Inwieweit hier die nichtlinearen Filteralgorithmen nach Stratonovitch (1960) und Kushner (1964) (vgl. auch Krebs (1980)) nützlich sind, ein Optimum an Information aus dem Trendverlauf verfügbar zu machen, ist eine interessante Fragestellung, die jedoch über den Rahmen dieser Arbeit hinausgeht.□

Im Vergleich zu a) ist die harmonische Extrapolation der linearen in den Beispielen b) und c) bei konkavem oder konvexem Trendverlauf mit nahzu konstanter Schwingungsdauer deutlich überlegen. Wird die atmosphärische Situation im Bsp. a) durch konstante Immissionsbedingungen (beispielsweise feste Inversionshöhe und konstanter Schadstoffausstoß) bestimmt, so liegen den Beispielen b) und c) mit einer festen Periode schwingende Änderungen der Immissionsbedingungen zugrunde.

Bei etwa der Hälfte aller Prognosesituationen bleibt jedoch die Schwingfrequenz des Trends selbst in der kurzen Zeitspanne von drei Tagen nicht konstant, was oft zu einem Umkehrpunkt der Trendkurve im zu prognostizierenden Zeitbereich führt (Bsp. d)). Diese Tatsache wird auch durch das Histogramm der Zeitabstände zwischen den Umkehrpunkten des Trends (Abb. 3-4), welches bei etwa 35 Stunden ein Maximum zeigt, untermauert.

Hier versagen natürlich univariate Techniken und so kommt es zu den in den Abschnitten 3 und 4 des 4. Teils erstellten Fehlprognosen 1. und 2. Art. An dieser Stelle läßt sich die Prognosegüte nur durch die Berücksichtigung des kausalen Zusammenspiels der atmosphärischen Zustandsgrößen verbessern. Solche Situationen sind univariat prinzipiell nicht vorhersagbar - eine multivariate Betrachtung wird unerläßlich.

Bemerkung (5-2): Die in Abb. 5-2 aufgezählten Situationen sind natürlich Idealisierungen, die zur Diskussion der Vorhersagemethode vorgenommen werden mußten. Reale Situationen sind Übergangsformen zwischen diesen. Ein optimaler Prognosealgorithmus sollte möglichst viele Verlaufsformen des Trends adäquat verarbeiten können.□

Bemerkung (5-3): Die Betrachtung konkreter Fälle von Fehlprognosen zeigt eine weitere Fehlerursache, die jedoch bereits durch Verbesserung des univariaten Modells beseitigt werden könnte: In einigen wenigen Situationen entsteht ein Überschreitungsfall durch eine Veränderung der Amplituden der Saisonkomponenten. Das im Komponentenmodel verwendete Saisonmodell (DHR) extrapoliert jedoch den periodischen Verlauf mit konstanter Amplitude. Es ließe sich also vermutlich auch durch eine Modifizierung des Prognosealgorithmus für den periodischen Anteil eine Steigerung der Treffsicherheit erreichen. Beispielsweise könnte man die RW-Modellierung der zeitabhängigen Amplituden durch ein IRW-Modell (siehe Anhang D) ersetzen. Letzteres würde die prognostizierte Amplitude zeitproportional verändern.□

Abschließend sollen noch einige Sachverhalte über die Unterschiede der lokalen harmonischen Trendextrapolation zu zwei bekannten Filteralgorithmen (der DHR- und der SIRW-Methode) angemerkt werden.

Bemerkung (5-4): Der von Young et al. (1991) angegebene Algorithmus zur dynamischen harmonischen Regression (DHR, siehe Anhang D), der die Zeitabhängigkeit der Frequenz zuläßt, wurde in der vorliegenden Arbeit zwar für die Modellierung und Extrapolation der saisonalen Anteile der Meßreihe verwendet, jedoch nicht für die Trendprognose. Die Gründe dafür sind folgende:

Ein mit dem DHR-Modell konstruiertes Kalman-Filter extrahiert alle in einem bestimmten Frequenzband gelegenen Schwingungen aus der Meßreihe. Im Gegensatz dazu wird bei der *statischen* harmonischen Regression nur genau die Schwingung einer festen Frequenz aus der Meßreihe herausgefiltert.

Dabei gehen beiden Verfahren von der Vorgabe einer festen Frequenz aus. Die Ausdehnung auf ein Frequenzband, dessen Breite vom gewählten NVR abängt, läßt dann beim DHR-Algorithmus eine gewisse Zeitabhängigkeit der Frequenz zu. Diese Zeitabhängigkeit wird jedoch nur bei der Modellierung, nicht aber bei der Prognose wirksam. Das DHR-Modell extrapoliert nämlich den Reihenverlauf in gleicher Weise, wie ein statisches harmonisches Modell: Mit einer Schwingung konstanter Amplitude und der in der Zustandsgleichung (bzw. Regressionsgleichung) festgelegten Frequenz.

Für die Trendextrapolation muß man deshalb von einer etwas anderen Sachlage ausgehen. Das Verhalten der Trendfunktion läßt sich nicht durch eine typische Periodizität, die im Laufe der Zeit stärker oder schwächer ausgeprägt ist, charakterisieren. Vielmehr ändert sich die Periode der Trendbewegung ständig und variiert über einen relativ großen Bereich. Alle dabei vorkommenden Frequenzen sind völlig gleichberechtigt. Die Dominanz einer bestimmten Schwingung ist bestenfalls lokal vorhanden. Wollte man für die Modellierung dieses Verhaltens die dynamische harmonische Regression verwenden, so hätte man eine sehr große Bandbreite zu wählen (NVR groß) und käme bei der Festlegung der charakteristischen Periode in große Schwierigkeiten. Für welchen Wert sollte man sich entscheiden?

Es liegt also nahe, die Periode lokal stets neu zu schätzen, wie dies in dem in Teil 3, Abschnitt 6 realisierten Verfahren getan wurde.□

Bemerkung (5-5): Die Isolierung der Trendfunktion $T_t = y_t$ erfolgte im Rahmen des Komponentenmodells entsprechend einem integrated random walk Ansatz in der Zustandsraumdarstellung. Dieser Ansatz ist verallgemeinerbar, indem ein zusätzlicher Parameter α eingeführt wird:

Zustandsgleichung

$$y_t = \alpha y_{t-1} + x_{t-1} \qquad \text{(SIRW-Modell)} \qquad (5\text{-}1)$$
$$x_t = x_{t-1} + \eta_t \qquad \eta_t = \text{weißes Rauschen}$$

Aus diesem smoothed integrated random walk Modell (SIRW) ergibt sich dann für $\alpha = 1$ als Spezialfall das IRW- und für $\alpha = 0$ das random walk Modell (RW). Hätte man nun ein Extrapolationsverhalten entsprechend der oben beschriebenen lokalen harmonische Approximation auch mit einem SIRW-Ansatz erzielen können?

Für die Prognose nimmt die SIRW Zustandsgleichung die Form

$$\begin{aligned} y_t &= \alpha y_{t-1} + x_{t-1} \\ x_t &= \quad x_{t-1} \end{aligned} \qquad (5\text{-}2)$$

an. x_t wird also konstant gehalten, so daß $y_t - \alpha y_{t-1} = \text{const.}$ ist. Wie man sieht, bestimmt sich die Ableitung (korrekter: die Differenzenreihe) der Trendfunktion y_t nun nicht mehr zu x_t, sondern aus

$$\begin{aligned} y_t - y_{t-1} &= \text{const.} - (\alpha\text{-}1)\, y_{t-1} = \text{const.} - (1\text{-}1/\alpha)\,(y(t) - \text{const.}) \\ &= \text{const.} - \beta\,(y(t) - \text{const.}) \end{aligned} \qquad (5\text{-}3)$$

mit $\beta = (1\text{-}1/\alpha)$, $-\infty < \beta \leq 0$.

Im Falle der lokalen harmonischen Trendextrapolation gilt dagegen lokal die Beziehung

$$y_t = c - a\cos(\Omega t - t_0) \qquad (5\text{-}4)$$

und damit

$$\begin{aligned} y_t - y_{t-1} &= a\,\{\cos(\Omega t - t_0 - \Omega) - \cos(\Omega t - t_0)\} \\ &= a\,\{\cos(\Omega t - t_0)\cos(\Omega) + \sin(\Omega t - t_0)\sin(\Omega) - \cos(\Omega t - t_0)\} \end{aligned} \qquad (5\text{-}5)$$

Wird dieser Ausdruck für kleine Ω entwickelt, so erhält man in niedrigster Ordnung

$$y_t - y_{t-1} = -(\Omega^2 a/2)\cos(\Omega t - t_0) = -\Omega^2/2\,(y_t - c). \qquad (5\text{-}6)$$

Vergleicht man diesen Ausdruck mit (5-3), so läßt sich erkennen, daß im Falle kleiner Ω die SIRW-Extrapolation ein Spezialfall der lokalen harmonischen Extrapolation ist.□

2. Ziel der Immissionsprognose: Ein Warnsystem

In hochbelasteten industriellen Ballungsgebieten sind zur Feststellung von Smogsituationen seit einigen Jahren Überwachungssysteme mit automatischen Meßgeräten installiert worden. Beispiele dafür sind die Meßnetze in Nordrhein-Westfahlen und in Sachsen. Mit einem solchen Kontrollsystem registriert man jedoch eine Smogsituation erst, wenn der kritische Zustand erreicht oder bereits vorüber ist (Herbarth (1983)).

Zur Abwendung von smogbedingten Schäden ist es erforderlich, Emissionssenkungen sowie präventive Verhaltensweisen der Bevölkerung bereits im Vorfeld zu veranlassen. Damit stellt sich die Frage nach einem Warnsystem. Der Begriff des Warnsystems meint allgemein einen Algorithmus, der die Möglichkeit der Entstehung einer kritischen Situation zu einem Zeitpunkt, zu dem die Kriterien den kritischen Bereich noch nicht erreicht haben, signalisiert.

Kritische Situationen sind nicht der Normalzustand, sie sind seltene Ereignisse. Praktisch läßt sich diese Feststellung durch die geringe Anzahl von Überschreitungsfällen (vgl. Anhang F) belegen. Dabei hatte man bereits eine Vorauswahl von prognosewürdigen Situationen mit erhöhtem Immissionsniveau getroffen und die Vielzahl der Nullfälle völlig unberücksichtigt gelassen.

Wie sich in Teil 4, Abschnitt 3 herausstellte, ist eine Treffsicherheit von 25% für die Vorhersage eines Überschreitungsfalls mit der Güte von Niederschlagsvorhersagen vergleichbar. Und dies, obwohl es sich nur um eine univariate Smogprognose handelt. Ein automatisches Vorwarnsystem, welches im on-line-Betrieb (z.B. direkt am PC, der die Meßdaten aufnimmt) aus den einlaufenden Immissionsdaten eine Warnmeldung ableitet, ist damit realisierbar. In diesem Sinne kann die im Teil 4, Abschnitt 2 gemachte Arbeitshypothese bestätigt werden.

Im Zeitraum von 1980-1993 hat das Warnsystem für die Immissionsmeßstation Leipzig-Mitte 65 Überschreitungsfälle vorhergesagt, wovon allerdings nur 16 eingetreten sind (Teil 4, Abschnitt 3). Hier machen sich die Fehler 2. Art bemerkbar. Da der Zweck eines Vorwarnsystems darin besteht, die Möglichkeit einer kritischen Situation anzukündigen, bedeuten die Fehler 2. Art nur eine unwesentliche Minderung der Vorhersagequalität.

Für die Sicherheit des Warnsystems haben jedoch die Fehler 1. Art eine größere Bedeutung, denn durch das Auftreten von Überschreitungsfällen ohne vorherige Warnung wird die Qualität der Vorhersage wesentlich gemindert. In dem hier betrachteten Prognosebeispiel traten 18 solcher Fälle auf. An dieser Stelle besteht also Bedarf der Verbesserung der Prognosegüte.

Schluß

In der vorliegenden Arbeit wurden speziell die Immissionsdaten der Meßstation Leipzig-Mitte des Zeitraumes 1980-1993 mit dem Ziel der Erstellung eines Algorithmus für die Kürzestfristprognose von Überschreitungssituationen untersucht. Die Prognosestellung sollte allein anhand der in der Vergangenheit registrierten Halbstundenwerte der SO_2-Konzentration, also univariat erfolgen.

Zur Stabilisierung der Varianz der Meßreihe wurde die Transformation

$$u_t = F(z_t) = 2(\sqrt{z_t} - 1)$$

gefunden. Zwar besitzt die Meßreihe eine Autokorreliertheit, jedoch ist diese Eigenschaft zur Konstruktion eines aussagekräftigen Vorhersagealgorithmus nicht ausreichend.

Wie die Fourieranalyse zeigt, gibt es drei wesentliche und voneinander unabhängige Spektralbereiche: Den hochfrequenten Bereich (Periode < 12 Stunden) der instabilen Irregularitäten, den saisonalen Anteil mit den Perioden von 24 und 12 Stunden und den niederfrequenten Bereich (Periode > 24 Stunden). Letzterer wird nach einer Zerlegung der Meßreihe in Komponenten als Trendkomponente (oder kurz Trend) bezeichnet. Für die Komponentenzerlegung wird ein Kalman-Filter verwendet. Es stellt sich heraus, daß Smogepisoden am deutlichsten durch die Trendkomponente beschrieben werden. Deshalb wird diese genauer untersucht. In der Phasenraumdarstellung zeigen sich dann charakteristische Trajektorienverläufe des Trends.

Die Feststellung, daß Modellierung und Extrapolation zwei unterschiedliche Aspekte des Kalman-Algorithmus sind, ist Anlaß, das Inter- und Extrapolationsverhalten des Kalman-Filters für den Trend zu untersuchen. Beim Vergleich dieses Verhaltens mit dem Verhalten der Trendtrajektorie im Phasenraum zeigt sich die Möglichkeit einer verbesserten Trendvorhersage. Diese wird als lokale harmonische Trendextrapolation bezeichnet. Die Schätzgleichung der dafür erforderlichen Parameter ist jedoch nichtlinear. Sie kann nur näherungsweise gelöst werden. Bei der Anwendung dieser Methode auf konkrete Prognosefälle bleibt die Güte dieser Vorhersagen deshalb hinter dem erwarteten Maß zurück.

Bei den 191 registrierten Prognosesituationen des Meßzeitraumes, die definierten Bedingungen entsprechen, wurden zwei sehr einfache Kenngrößen ausgewählt und faktische Prediktoren genannt. Mit ihnen ist bereits eine Wahrscheinlichkeitsaussage über eine künftige Schwellwertüberschreitung der SO_2-Konzentration möglich. Die Seltenheit von Überschreitungssituationen bedingt generell eine niedrige Treffsicherheit der Vorhersagen. Durch die Anwendung der in den Teilen 2 und 3 entwickelten Algorithmen für die Zeitreihenprognose zur Bestimmung eines prognostizierten Prediktors kann jedoch eine Treffsicherheit erreicht werden, die mit der von vergleichbaren Niederschlagsvorhersagen übereinstimmt. Wodurch Fehlprognosen zustande kommen, war an ausgewählten Situationen im Teil 5 zu diskutieren. Hier zeigen sich auch die Grenzen des univariaten Herangehens.

Die hier dokumentierten Untersuchungen haben verdeutlicht, daß man die in der Variabilität der Immissionsreihe enthaltenen Phänomene für die Konstruktion eines on-line arbeitenden Algorithmus zur Smogvorwarnung nutzen kann. Ein Vorschlag für einen solchen Algorithmus wurde in Teil 4 angegeben und Modifikationen zum Zwecke der Erhöhung der Treffsicherheit sowie seine prinzipiellen Grenzen diskutiert.

Der Prognosealgorithmus geht von einer Zeitreihenmodellierung der Immissionsreihe aus. Man kann erwarten, daß die in den Algorithmus einbezogenen Phänomene, wie Tagessaisonalität und die Widerspiegelung von Smogereignissen durch den Trendverlauf vom konkreten Meßstandort nicht abhängen. Damit ist eine gewisse Universalität des Warnsystems garantiert. Eine Anpassung an die örtlichen Gegebenheiten, wie orografische Situation oder Emissionsstruktur der Region ist somit nicht erforderlich.

Die weitere Forschungsarbeit sollte einerseits den Aspekt der multivariaten Zeitreihenmodellierung einbeziehen, d.h. die Abhängigkeiten von meteorologischen Zustandsgrößen und der Emissionsstärke zur Steigerung der Prognosegüte ausnutzen.

Andererseits haben die vorliegenden Untersuchungen auch deutlich gemacht, daß Immissionsreihen ein zeitlich veränderliches Spektrum aufweisen. Man bezeichnet dies auch als evolutionäres Spektrum. Von Priestley (1965) und (1981) und Hampe (1982) wurde eine Theorie dieser Erscheinungen erarbeitet. Dabei geht es auch um die Schätzung des jeweils **lokal** gültigen Spektrums, wie das folgende Zitat von Priestley (1981) zeigt:

> *In "the case of more general types of non-stationary processes, we are led to the notion of a continuously changing spectrum, or more precisely, a time-dependent spectrum. Clearly, in such cases we could never hope to estimate the spectrum at a particular instant of time, but if we assume that the spectrum is changing smoothly over time then, by using estimates which involve only local functions of the data, we may attempt to estimate some form of "average" spectrum of the process in the neighbourhood of any particular time instant. We therefore consider the class of non-stationary processes whose statistical characteristics are changing "smoothly" over time."*

Die Anwendung lokaler Spektralschätzer und daraus abgeleiteter Prognosemethoden auf Immissionsdaten erscheinen deshalb erfolgversprechend.

Wie diffizil bei der im Teil 3, Abschnitt 5 realisierten Schätzung der Parameter für die lokale harmonische Extrapolation die Festlegung von Stützstellenabstand, Mittelungsbereich für die Frequenz etc. ist, ist dem Autor bewußt. Deshalb wäre es auch interessant, die Arbeitsweise von nichtlinearen Filteralgorithmen, die evtl. zu einer adäquateren Trendextrapolation im Sinne des Trajektorienverlaufs im Phasenraum führen könnte, zu testen.

Akaike, H (1970): Statistical predictor identification; *Annals of the Institute of Statist. Math.* **22**, S. 203.

Balzer, K (1984): Gegenwärtige Grenzen der Vorhersagbarkeit und Gibt es einen Trend der Prognosegüte?; *Abh. Met. Dienst d. DDR*, Nr. 133, S. 37.

Balzer, K (1986): Eine selbstlernende, optimale Mensch-Maschine-Kombination von operativen Immissionsprognosen; *Z. Meteorol.* **36**, Nr. 2, S. 127-133.

Balzer, K (1989): Statistische Vorhersage seltener, extremer meteorologischer Erscheinungen; *Abh. Met. Dienst d. DDR*, Nr. 141, S. 115-118.

Bandemer, H (1989): *Einführung in die Fuzzy-Methoden: Theorie und Anwendungen unscharfer Mengen*, Akademie Verlag.

Beck, EG und Schmidt, P (1992): *Hygiene Präventivmedizin*, Ferdinand Enke Verlag Stuttgart.

Bergeron, T (1928): Über die dreidimensional verknüpfende Wetteranalyse, I. Teil; *Geof. Publ.* **5**, Nr. 6.

Bjerknes, J (1919): On the structure of moving cyclones; *Geof. Publ.* **1**, Nr. 2.

Box, GEP und Cox, DR (1964): An analysis of transformations; *J R Statist Soc B* **26**, S. 211-252.

Box, GEP und Jenkins, GM (1970): *Time series analysis - forecasting and control*, Prentice Hall, Englewood Cliffs, New Jersey, Revised Edition 1976.

Box, GEP und Pierce, DA (1970): Distribution of residual autocorrelations in autoregressive - integrated moving average time series models, Revised Edition 1976; *J. Am. Stat. Ass.* **65**, Nr. 332, S. 1509-1525.

Bredel, H und Herbarth, O (1987): Methodische Untersuchungen zur optimalen Beschreibung der Raum-Zeit-Struktur des Immissionsfeldes kommunaler Ballungsgebiete - Ermittlung der räumlichen Immissionsverteilung;
Z. gesamte Hyg. **33**, Nr. 8, S. 370-372.

Burg, JP (1967): Maximum Entropy Spectral Analysis, *Unpubl. paper presented at the 37th Meeting of Soc. Explor. Geophys., Oklahoma City*.

Doob, JL (1953): *Stochastic Processes*, Wiley, New York.

Duncan, DB und Horn, SD (1972): Linear dynamic recursive estimation from the viewpoint of regression analysis;
J Amer Statist Ass **67**, S. 815-821.

Durbin, J (1969): Tests of serial correlation in regression analysis based on the periodogram of least-squares residuals;
Biometrica **56**, S. 1-15.

Enke, W (1986): Ein modernes Verfahren zur Aufstellung statistischer Prognosealgorithmen als Einheit von Cluster-, Diskriminanz- und Regressionsanalyse; *Z. Meteorol.* **36**, Nr. 2, S. 117-122.

Essenwanger, OM (1980): On red noise and quasi-periodicity in the time series of atmospheric temperature; *Statistical Climatology - Developments in Atmospheric Sciences* **13**, S. 165.

Fildes, R (1983): An evaluation of Bayesian forecasting; *J Forecasting* **2**, S. 137.

Fischer, A (1933): *Geschichte des Deutschen Gesundheitswesens I*, Berlin.

Hamming, RW (1973): *Numerical Methods for Scientists and Engineers*, McGraw-Hill.

Hampe, JF (1982): *Zur Theorie und Anwendung der Schätzung zeitabhängiger Spektren*, Haag und Herchen, Frankfurt am Main.

Harrison, PJ und Stevens, CF (1976): Bayesian forecasting (with discussion); *J Roy Statist Soc B* **38**, S. 205-247.

Heller, WD et al. (1978): *Stochastische Systeme*, de Gruyter, Berlin.

Herbarth, O (1982): Zeitreihenanalyse und Prognose lufthygienischer Konzentrationsreihen; *Z. gesamte Hyg.* **28**, Nr. 7, S. 508-09.

Herbarth, O (1993): Smogprognosen - eine Hilfe für die Medizin; *Mitteilungen Deutsche Meteorologische Gesellschaft e.V.*, Nr. 4, S. 17-20.

Herbarth, O und Bredel, H (1983): Automatische Immissionsüberwachung und Immissionsprognose; *Z. gesamte Hyg.* **29**, Nr. 9, S. 533.

Herbarth, O und Bredel, H (1987): Methodische Untersuchungen zur optimalen Beschreibung der Raum-Zeit-Struktur des Immissionsfeldes kommunaler Ballungsgebiete - Ermittlung der zeitlichen Immissionsverteilung; *Z. gesamte Hyg.* **33**, Nr. 9, S. 442-43.

Herbarth, O und Bredel, H (1990): Methodical aspects of finding threshold values for air protection using epidemiologic studies; *Zbl. Hyg.* **190**, S. 335-343.

Ho, Yu-Chi (1962): On the stochastic approximation method and optimal filtering theory; *Journal of mathematical analysis and applications* **6**, S. 152.

Höschele, K (1966): Der zeitliche Verlauf und die örtliche Verteilung der SO_2-Konzentration in einem Stadtgebiet mit einer Analyse der Einflußgrößen; *Meteorologische Rundschau* **19**, Nr. 1, S. 14-22.

Horn, K; Pinigin, M; Dörre, W; Hammje, K und Knauer, A (1989): *Grundlagen der Lufthygiene*, VEB Verlag Volk und Gesundheit, Berlin.

Hoven, I Van der (1957): Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour; *J. Meteor.* **14**, S. 160.

Isermann, R (1974): *Prozeßidentifikation*, Springer-Verlag Berlin Heidelberg New York.

Jost, D (1993): *Die neue TA Luft*, WEKA Fachverlag für technische Führungskräfte GmbH.

Kalman, RE (1960): A new approach to linear filtering and prediction problems; *Journal of Basic Engineering* **82**, S. 35.

Kalman, RE und Bucy, RS (1961): New results in linear filtering and prediction theory; *Trans. ASME, Series D* **83**, S. 95-108.

Kendall, M und Ord, JK (1993): *Time Series*, Edward Arnold.

Klaus, G (1969): *Wörterbuch der Kybernetik*, Dietz Verlag, Berlin.

Krebs, V (1980): *Nichtlineare Filterung*, R. Oldenbourg Verlag München Wien.

Kushner, HJ (1964): On the dynamical equations of conditional probability density functions, with applications to optimal stochastic control theory; *J. Math. Anal. Appl.* 8, S. 332-344.

Kushner, HJ (1967): Approximations to optimal nonlinear filters; *IEEE Trans. on Autom. Control* **AC-12**, S. 546-556.

Kushner, HJ (1967): Dynamical equations for optimal nonlinear filtering; *J. Differential Equations* 3, S. 179-190.

Leipzig (1991): Umweltbericht der Stadt Leipzig 1991; *Rat der Stadt Leipzig, Dezernat für Umweltschutz und Sport.*

Müller, U (1993): Beziehungen zwischen Stadtklima und ausgewählten Luftinhaltsstoffen in Abhängigkeit von urbanen Flächennutzungsstrukturen am Beispiel der Stadtregion Leipzig; *Umweltforschungszentrum Leipzig-Halle GmbH, Unveröffentlichter Abschlußbericht (1.3.93).*

Ng, CN und Young, PC (1990): Recursive estimation and forecasting of non-stationary time series; *Journal of Forecasting* **9**, S. 173-204.

Norton, JP (1975): Optimal smoothing in the identification of linear time-varying systems;
Proceedings Instn. Elect. Engrs. **122**, Nr. 6, S. 663.

Olberg, M und Rakoczi, F (1984): *Informationstheorie in Meteorologie und Geophysik*, Akademie-Verlag, Berlin.

Omatu, S und Seinfeld, JH (1981): Filtering and smoothing for linear discrete-time distributed parameter systems based on Wiener-Hopf theory with application to estimation of air pollution; *IEEE Transactions* **SMC-11**, Nr. 12, S. 785-801.

Omatu, S und Seinfeld, JH (Apr. 1982): Estimation of atmospheric species concentrations from remote sensing data; *IEEE Transactions on geoscience and remote sensing* **GE-20**, Nr. 2, S. 142-53.

Omatu, S und Seinfeld, JH (1983): Optimization of sensor and actuator locations in a distributed parameter system; *Journal of the Franklin Institute* **315**, Nr. 5/6, S. 407-21.

Omatu, S und Seinfeld, JH (1987): Optimal estimators for distributed-parameter systems with time-averaged pointwise measurements; *Int. J. Control* **45**, Nr. 6, S. 1963-74.

Peil, J (1985): Verfahren der lokalen Approximation zur nichtparametrischen Schätzung unbekannter stetiger Funktionen aus Meßdaten; *Gegenbaurs morph Jahrb* **131**, Nr. 3, S. 367.

Plackett, RL (1950): Some theorems in least squares; *Biometrica* **37**, S. 149-157.

Priestley, MB (1965): Evolutionary spectra and non-stationary processes; *J. R. Statist. Soc.* **B 27**, S. 204.

Priestley, MB (1981): *Spectral Analysis and Time Series, Vol. 1 and 2,*
Academic Press, London.

Schlink, U (1990): Zur rechnergestützten Auswertung von Immissionsmeßdaten; *Z. gesamte Hyg.* **36**, Nr. 1, S. 149-157.

Schlink, U (1993): Die Definition der Immissionskenngrößen nach Bundesrecht und DDR-Recht;
Staub - Reinhaltung der Luft **53**, Nr. 4, S. 139-141.

Schlittgen, R und Streitberg, BHJ (1991): *Zeitreihenanalyse*, R. Oldenbourg Verlag GmbH, München Wien.

Schönwiese, C-D (1992): *Praktische Statistik für Meteorologen und Geowissen- schaftler*, Gebr. Borntraeger, Berlin Stuttgart.

SPSS (1987): SPSS/PC+ Trends, Dokumentation der Erweiterung des SPSS für Zeitreihenanalyse und -prognose; *SPSS Inc.*.

Verordnung der Sächsischen Staatsregierung und des Sächsischen Staats- ministeriums für Umwelt und Landesentwicklung zur Verhinderung schädlicher Umwelteinwirkungen bei austauscharmen Wetterlagen (1991): *Smog-Verordnung*, Bd. Nr. 31, Sächsisches Gesetz- und Verordnungsblatt 9. Dez. 1991, S. 390.

Stephens, MA (1970): Use of the Kolmogorov-Smirnov, Cramer-von Mises and related statistics without extensive tables; *J R Statist Soc* **B 32**, S. 115-122.

Stratonovich, RL (1966): A new form of representing stochastic integrals and equations; *J. SIAM Control* **4**, S. 362-371.

Stull, RB (1989): *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers.

VDI: *VDI 2310*, VDI-Verlag GmbH, Düsseldorf .

Vereinsbank, Bayerische (1989): *Oft verwendet - kurz erklärt*, München.

Wiegand, G; Weigelt, M und Stahl, H (1987): Methoden zur Bewertung der Luftqualität, Teil A: Zeitreihenanalyse von Immissionsmessungen; Teil B: Extremwertuntersuchungen von Schadstoffkonzentrationen; *Bundesminister für Umwelt, Naturschutz und Reaktorsicherheit, Forschungsbericht 1 04 02 440, Luftreinhaltung* .

Young, P und **Benner**, S (1991): microCAPTAIN 2; *User handbook, first edition, Centre for Research on Environmental Systems, Institute of Environmental and Biological Sciences, Lancaster University*.

Young, PC; **Ng**, C und **Armitage**, P (1989): A systems approach to recursive economic forecasting and seasonal adjustment; *Computers Math. Applic.* **18**, Nr. 6/7, S. 481-501.

Young, PC; **Ng**, CN; **Lane**, K und **Parker**, D (1991): Recursive forecasting, smoothing and seasonal adjustment of non-stationary environmental data; *Journal of Forecasting* **10**, S. 57-89.

Anhänge

Anhang A

Zeitreihen und ihre statistischen Kenngrößen

Immissionsmeßreihen sind in zeitlicher Reihenfolge geordnete Zahlen z_t (t=1..N). Solche Reihen bezeichnet man als Zeitreihen oder -serien.

Für die systematische Behandlung einer Zeitreihe nimmt man an, daß sie eine Realisierung eines dynamischen Vorgangs mit Zufallscharakter ist. Ein dynamischer Prozeß beschreibt die zeitliche Veränderung einer Größe und enthält somit die Abhängigkeiten zwischen den Werten dieser Größe zu verschiedenen Zeitpunkten. Gerade diese Abhängigkeit ist für die Untersuchung von Zeitreihen interessant.

Die Theorie der stochastischen Prozesse und der Zeitreihenanalyse ist mathematisch exakt und detailliert in Lehrbüchern abgehandelt (vgl. Heller u.a. (1978), Doob (1953), Box & Jenkins (1970), Schlittgen & Streitberg (1991)). Hier sollen als Arbeitsgrundlage nur einige Definitionen und Zusammenhänge angegeben werden:

Definition (A-1): Eine reelle **Zufallsvariable** Z ist eine (meßbare) Abbildung, die den Ergebnissen ξ eines Zufallsvorgangs jeweils reelle Zahlen $Z(\xi)$ zuordnet.□

Definition (A-2): Ein **stochastischer Prozeß** ist eine Folge $\{Z_t\}$ $(t \in T)$ von Zufallsvariablen Z_t, wobei der als Zeitparameter bezeichnete Index t Element einer höchstens abzählbaren Indexmenge T ist. In der Regel wird T als die Menge der natürlichen Zahlen gewählt.□

Einem stochastischen Prozeß liegt ein bestimmter Zufallsvorgang zugrunde. Realisiert sich ein Ereignis ξ des Zufallsvorgangs, so nimmt jede Zufallsvariable Z_t einen Wert $z_t(\xi)$ an. Man erhält somit eine Folge $\{z_t(\xi)\}$ $(t \in T)$ von Zahlen. Ein stochastischer Prozeß ist demnach als eine Abbildung zu verstehen, die jedem Ergebnis ξ eines Zufallsvorgangs eine **Zeitreihe** $\{z_t(\xi)\}$ zuordnet. Zur Vereinfachung der Schreibweise wird im weiteren diese Zeitreihe mit $\{z_t\}$ bezeichnet.

Bei der Zeitreihenmodellierung spielen zwei stochastische Prozesse eine fundamentale Rolle. Sie sollen deshalb an dieser Stelle genannt werden.

Definition (A-3): Ein **White-Noise-Prozeß** (weißes Rauschen) ist eine Folge $\{\varepsilon_t\}$ $(t \in T)$ von identisch verteilten und unabhängigen Zufallsvariablen ε_t.□

Definition (A-4): Sei $\{\varepsilon_t\}$ $(t \in T)$ ein White-Noise-Prozeß. Dann ist durch die Rekursion

$$X_t = \begin{cases} \varepsilon_1 & \text{für } t=1 \\ X_{t-1} + \varepsilon_t & \text{für } t=2, 3, \ldots \end{cases}$$

ein **Random-Walk-Prozeß** (RW) $\{X_t\}$ $(t \in T)$ festgelegt.□

Bemerkung (A-1): Der Zuwachs des Random-Walk-Prozesses ist also ein weißes Rauschen. Die Random-Walk-Gleichung wird beispielsweise bei der eindimensionalen Brownschen Bewegung eines Mikroteilchens realisiert.□

Für die Behandlung von Zeitreihen als Realisierungen stochastischer Prozesse ist es von Bedeutung, daß gewisse Charakteristika des Prozesses, auf die später noch eingegangen werden soll, gegen Zeitverschiebungen invariant sind. Geht man allgemein von der Forderung aus, daß sämtliche Charakteristika des Prozesses im Zeitverlauf invariant bleiben, so muß es sich um einen streng stationären Prozeß handeln:

Definition (A-5): Ein stochastischer Prozeß $\{Z_t(\xi)\}$ $(t \in T)$ heißt **streng stationär**, wenn die gemeinsame Verteilungsfunktion jedes endlichen Systems von Zufallsvariablen $(Z_{t(1)}, Z_{t(2)}, \ldots, Z_{t(n)})$ des Prozesses identisch ist mit der gemeinsamen Verteilungsfunktion des um s Zeitpunkte verschobenen Systems $(Z_{t(1+s)}, Z_{t(2+s)}, \ldots, Z_{t(n+s)})$.□

In den praktischen Anwendungen der Zeitreihenanalyse interessieren vorwiegend die Momente erster und zweiter Ordnung (Mittelwert, Varianz, Kovarianzen). Sollen diese Charakteristika gegen Zeitverschiebungen invariant sein, so spricht man von **schwacher Stationarität**, was ein Spezialfall der strengen Stationarität ist. Wenn im weiteren von Stationarität die Rede ist, soll damit stets die schwache Stationarität gemeint sein, d.h. Mittelwert, Varianz und Kovarianz sind zeitlich konstant. Bei Vorliegen von Meßreihen die nicht stationär sind, bedient man sich im allgemeinen gewisser "Kunstgriffe", wie z.B. der im Teil 1 angewandten nichtlinearen Box-Cox-Transformation zur Erzielung einer stationären Varianz.

Die genannten Charakteristika lassen sich für stationäre Zeitreihen wie folgt festlegen:

Definition (A-6): Der **Mittelwert** z_{MW} wird durch das arithmetische Mittel gebildet.

$$z_{MW} = \frac{1}{N} \sum_{t=1}^{N} z_t \,. \square$$

Er beschreibt die zentrale Lage der Werte der Zeitreihe $\{z_t\}$.

Definition (A-7): Die **Varianz** c_0 = var(z) berechnet sich zu:

$$c_o = \frac{1}{N}\sum_{t=1}^{N}(z_t - z_{MW})^2 . \square$$

Definition (A-8): Die **Autokovarianz** c_k ist:

$$c_k = \frac{1}{N-k}\sum_{t=1}^{N-k}((z_t - z_{MW})(z_{t+k} - z_{MW})).$$

Die Autokovarianz ist eine Funktion des Zeitabstandes k, der auch als **Lag** bezeichnet wird. □

Definition (A-9): Als **Autokorrelationsfunktion** r_k bezeichnet man den Quotienten

$$r_k = \frac{c_k}{c_o}.$$

Der Graph dieser Funktion heißt **Autokorrelogramm**. □

Oft ist es von Interesse, in den Zeitreihen enthaltene Periodizitäten aufzudecken. Dazu läßt sich das Periodogramm nutzen:

Definition (A-10): Das **Periodogramm** oder Stichprobenspektrum der Zeitreihe $\{z_t\}$ (t=1,...,N) ist die folgende Funktion der Frequenz f_k:

$$P(f_k) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} z_t e^{-2\pi i f_k t}\right|^2 \quad \text{mit } f_k = \frac{k}{N};\ k = 1,...,(\tfrac{N}{2}) \quad \square$$

Es gibt für jede Frequenz f_k ein Maß dafür an, mit welcher "Stärke" (oder "Intensität") harmonische Wellen dieser Frequenz in der Ausgangsreihe enthalten sind.

Durch den auf 1 festgelegten Zeitabstand zwischen zwei Messungen ist die kleinste registrierbare Periode 2. Dies begrenzt die im Periodogramm darstellbare größte Frequenz auf 0.5, welche auch als Nyquist-Frequenz bezeichnet wird. Andererseits kann bei N Meßwerten eine maximale Periode von N und damit eine kleinste Frequenz von 1/N vom Periodogramm erfaßt werden. Somit genügt es, $P(f_k)$ für Frequenzen f_k zwischen 0 und 0,5 zu betrachten.

Geht man von einem naiven Standpunkt aus, so kann man das Periodogramm als heuristisches Hilfsmittel zur Aufdeckung von Periodizitäten betrachten. Ob auffällige Peaks im Stichprobenspektrum tatsächlich real existierenden periodischen Erscheinungen entsprechen, muß dann anhand der konkreten Sachlage diskutiert werden.

Als direkter Schätzer für die Spektraldichte des zugrundeliegenden stochastischen Prozesses ist das Periodogramm ungeeignet. Selbst für Reihenlängen $N \rightarrow \infty$ nähert sich $P(f_k)$ der entsprechenden Spektraldichte nicht an (Schlittgen & Streitberg (1991)). Durch geeignete Modifikationen läßt sich das Periodogramm allerdings zu einem geeigneten Schätzer für die Spektraldichte ausbauen.

So erhält man beispielsweise durch Glättung des Periodogramms einen konsistenten Schätzer für die Spektraldichte. Eine Glättung des Stichprobenspektrums mit einem Tukey-Hamming(5)-Fenster ist im Teil1, Abschnitt 3 an den Immissionsmeßdaten durchgeführt und erläutert worden.

Anhang B

Die Box-Jenkins-Methode

Von (Box & Jenkins (1970)) wurden die sogenannten ARIMA-Modelle, die vorwiegend in der Ökonometrie, aber auch bei technischen und naturwissenschaftlichen Problemen Anwendung finden, untersucht. Im folgenden sollen ihre Definition und Eigenschaften kurz erläutert werden.

Das Auffinden eines adäquaten Modells geschieht in drei Schritten: Identifikation des Modells, Schätzung der Modellparameter, Bewertung der Anpassungsgüte des gefundenen Modells. Unter Umständen muß diese Prozedur mehrmals durchlaufen werden.

Das so gefundene Modell ermöglicht dann eine Prognose des weiteren Zeitreihenverlaufs.

Bei einem ARIMA-Modell wird die Reihe **abhängiger**, zeitbezogener Meßwerte (Zeitreihe) $z_1, z_2, \ldots, z_N$ als Realisierung eines stochastischen Prozesses $\{z_t\}_{t\in T}$, $T=\{\ldots,-1,0,1,\ldots\}$, betrachtet.

z_t besitze die Darstellung:

$$z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2}-\ldots-\phi_{p+d} z_{t-p-d} = \alpha_t-\theta_1\alpha_{t-1}-\theta_2\alpha_{t-2}-\ldots-\theta_q\alpha_{t-q} \qquad \text{(B-1)}$$

wobei $\{\alpha_t\}_{t\in T}$ eine Folge unkorrelierter Zufallsgrößen ist mit $E(\alpha_t) = 0$ (B-2)

und $D^2(\alpha_t) = \sigma_\alpha{}^2$ (α_t: "weißes Rauschen").

Diese Differenzengleichung (B-1) stellt also einen stochastischen Signalprozeß als Funktion eines diskreten weißen Rauschens $\{\alpha_t\}$ dar.

Mit den Definitionen des Backshiftoperators B und der Polynome $\Phi(B)$ und $\Theta(B)$

$$Bz_t := z_{t-1} \qquad \text{(B-3)}$$

$$\Phi_{p+d}(B) := 1-\phi_1 B-\phi_2 B^2-\ldots-\phi_{p+d} B^{p+d} \qquad \text{(B-4)}$$

$$\Theta_q(B) := 1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q \qquad \text{(B-5)}$$

lautet die Differenzengleichung in Operatorform:

$$\Phi_{p+d}(B)\, z_t = \Theta_q(B)\, \alpha_t. \qquad \text{(B-6)}$$

Weiterhin besitze $\Phi_{p+d}(B)$ die Gestalt $\Phi_{p+d}(B) = \Phi_p(B)(1-B)^d$, und der Differenzoperator ∇ sei festgelegt durch

$$\nabla z_t := z_t - z_{t-1} = (1 - B)\, z_t \qquad \text{(B-7)}$$
$$\nabla^2 z_t := \nabla(\nabla z_t) = z_t - 2z_{t-1} + z_{t-2} \quad \text{usw.},$$

so daß die allgemeine Formulierung des Modells

$$\Phi_p(B)\, \nabla^d z_t = \Theta_q(B)\, \alpha_t \qquad \text{(B-8)}$$

ist.

Ein stochastischer Prozeß (mit disktreter Zeit t), der eine Darstellung (B-8) mit geeigneten Parametern p,d,q,σ_α^2 und Koeffizienten $\phi_1,\phi_2,...,\phi_p,\theta_1,\theta_2,...,\theta_q$ gestattet, heißt ARIMA(p,d,q)-Prozeß: autoregressiver integrierter Prozeß der gleitenden Mittel der Ordnung (p,d,q) (engl.: AutoRegressive Integrated Moving Average process).

Bemerkung (B-1): Ein ARIMA-Prozeß wird als stationär und invertierbar bezeichnet, wenn d=0 ist und für alle Nullstellen von $\Phi_p(B)$ und $\Theta_q(B)$ stets $|B| > 1$ gilt. Im weiteren werden nur derartige Modelle betrachtet, so daß die Wahl der Koeffizienten $\phi_1,\phi_2,...,\phi_p,\theta_1,\theta_2,...,\theta_q$ gewissen Restriktionen unterliegt.□

Spezialfall (B-1): Der autoregressive Prozeß AR(p) der Ordnung p ist gekennzeichnet durch

$$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)\, z_t = \alpha_t \qquad (t=1,..,N). \qquad \text{(B-9)}$$

In diesem Fall hängt der Signalwert z_t vom Zufallswert α_t und von den mit den Parametern ϕ_i gewichteten vergangenen Werten z_{t-1}, z_{t-2}, ... ab. Daher die Bezeichnung "autoregressiv".□

Beispiel (B-1): Der AR(1)-Prozeß (≡ARIMA(1,0,0), Random Walk)

$$z_t = \phi_1 z_{t-1} + \alpha_t \qquad \text{(B-10)}$$

ist für $\phi_1=1$ instationär und beschreibt die eindimensionale Brownsche Bewegung eines Mikroteilchens.□

Für eine experimentell gegebene Zeitreihe $\{z_i\}_{i=1...N}$ ist nun ein geeignetes ARIMA(p,d,q)-Modell zu identifizieren und anschließend aus $\{z_i\}_{i=1,...,N}$ die Modellparameter $p,d,q,\sigma_\alpha^2,\phi_1,\phi_2,...,\phi_p, \theta_1,\theta_2,...,\theta_q$ zu schätzen.

Die Eigenschaft der Autokorreliertheit eines stochastischen Signals wird durch zwei Funktionen beschrieben: Die **Autokorrelationsfunktion (ACF)** und die **partielle Autokorrelationsfunktion (PACF)**. Box & Jenkins (1970) haben untersucht, welches Verhalten für die ACF und die PACF bei mit den ARIMA-Modellen realisierten Zeitreihen theoretisch zu erwarten ist. Dabei zeigte sich, daß es ein der Struktur eines Modells entsprechendes typisches Verhalten gibt, welches in Tabelle B-1 zusammengefaßt ist.

Die **Identifikation** eines Modells basiert nun darauf, die aus der Meßreihe geschätzten ACF ($=r_k$) und PACF ($=\phi_{kk}$) mit den theoretischen Funktionsverläufen zu vergleichen.

Bei den Schätzungen werden folgende Formeln verwendet:

$$z^{(d)}_{MW} = \frac{1}{N-d} \sum_{i=1}^{N-d} \nabla^d z_i \qquad \text{(Mittelwert)} \qquad \text{(B-11)}$$

$$c_o^{(d)} = \frac{1}{N-d} \sum_{i=1}^{N-d} (\nabla^d z_i - z^d_{MW})^2 \qquad \text{(Varianz)} \qquad \text{(B-12)}$$

$$c_k^{(d)} = \frac{1}{N-d} \sum_{i=1}^{N-d-k} (\nabla^d z_i - z^d_{MW})(\nabla^d z_{i+k} - z^d_{MW}) \qquad \text{(Autokovarianz)} \quad k=1,2,...,K \quad (K \leq (N-d)/4) \qquad \text{(B-13)}$$

$$r_k^{(d)} = c_k^{(d)} / c_o^{(d)} \qquad \text{(Autokorrelation, ACF)} \qquad \text{(B-14)}$$

und für die partielle Autokorrelation ('d' weggelassen), (PACF)

(B-15)

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & r_1 & \cdots & r_{k-2} & r_1 \\ r_1 & 1 & \cdots & r_{k-3} & r_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ r_{k-1} & r_{k-2} & \cdots & r_1 & r_k \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & \cdots & r_{k-2} & r_{k-1} \\ r_1 & 1 & \cdots & r_{k-3} & r_{k-2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ r_{k-1} & r_{k-2} & \cdots & r_1 & 1 \end{vmatrix}}; \quad \phi_{11} = r_1; \quad k = 1,2,3,...,L$$

Praktisch werden für die Modellidentifikation die ca. 20 ersten Autokorrelationen r_k der Originalzeitreihe $\{z_t\}_{t=1...N}$ und gegebenenfalls der 1. und 2. Differenzenreihe (∇z_t und $\nabla^2 z_t$) betrachtet. Fällt die Autokorrelation der Originalreihe $r_k(\{z_t\})$ sehr schnell ab, so ist d=0 zu wählen und es handelt sich um eine stationäre Originalreihe. Sinkt jedoch die Autokorrelation nur allmählich ab, dann ist $\{z_t\}$ nicht stationär und man betrachtet die Autokorrelation der 1. Differenzenreihe $r_k(\{\nabla z_t\})$. Fällt diese sehr schnell ab, so ist d=1 zu wählen und die 1. Differenzenreihe ist stationär. Andernfalls ist das Verfahren analog fortzusetzen. In der Praxis wird i.a. d=0, 1 oder 2 gewählt.

Nach dem Festsetzen des Parameters d betrachtet man nun das ARMA(p,q)-Modell, welches der d-ten Differenzenreihe $w_t := \nabla^d z_t$ anzupassen ist.

Für die Wahl der Parameter p und q werden Eigenschaften der Autokorrelationsfunktion (ACF) und der partiellen Autokorrelationsfunktion (PACF) von AR(p)≡ ARMA(p,0)-, MA(q)≡ARMA(0,q)- und ARMA(p,q)-Prozessen genutzt, die in Tab. (B-1) angegeben sind (Box & Jenkins (1970)). Für lange Meßreihen, d.h. für großes N, zeigen die Schätzungen r_k und ϕ_{kk} der Zeitreihen ähnliches Verhalten wie die ACF und die PACF der korrespondierenden stochastischen Prozesse und werden daher zur Entscheidungsfindung herangezogen. In der Praxis wählt man meist p=0, 1 oder 2 und q=0, 1 oder 2.

Anfangsschätzungen für die Modellparameter ϕ_p und θ_q gewinnt man aus den Yule-Walker-Gleichungen. Als Beispiel sei die Herleitung dieser Gleichungen für einen AR(p)-Prozeß angegeben:

Ausgangspunkt ist die Modellgleichung

$$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)\, z_t = \alpha_t \qquad (t=1,..,N). \qquad \text{(B-16)}$$

Nach Multiplikation mit z_{t-k} und Erwartungswertbildung entsteht für k=0:

$$c_o-\phi_1 c_1-\phi_2 c_2-...-\phi_p c_p=E[\alpha_t z_t]=E[\alpha_t^2]=\sigma_\alpha^2 \qquad \text{(B-17)}$$

und für k=1,2,...,p:

$$\rho - \mathbf{R}\,\Phi = 0, \qquad \text{(B-18)}$$

wobei

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \dots & \rho_{p-1} \\ \rho_1 & \rho_2 & & \dots & \dots & \rho_{p-2} \\ . & & & & & . \\ . & & \dots & \dots & . & \\ \rho_{p-1} & & \dots & \dots & 1 & \end{pmatrix}; \quad \rho = \begin{pmatrix} \rho_1 \\ \rho_2 \\ . \\ . \\ \rho_p \end{pmatrix}; \quad \Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ . \\ . \\ \phi_p \end{pmatrix} \qquad \text{(B-19)}$$

und $\rho_k = E[z_t, z_{t-k}]$ ist.

Dieses System ist nach Φ aufzulösen und für ρ_k hat man die Schätzwerte r_k zu verwenden. Aus der Beziehung für k=0 ergibt sich z.B.:

$$s_a^2 = c_0(1-\phi_1 r_1 - \phi_2 r_2 - \ldots - \phi_p r_p) \qquad \text{(B-20)}$$

Bemerkung (B-2): Im Falle eines MA(q)-Prozesses (≡ARIMA(0,0,q)) geht die Herleitung analog, führt aber auf ein nichtlineares Gleichungssystem. Durch Methoden, wie die Minimum-Quadratsummen-Schätzung lassen sich die aus den Yule-Walker-Gleichungen berechneten Anfangswerte noch verbessern (Box & Jenkins (1970)).□

Die *Prognose* erfolgt speziell bei AR(p)-Reihen (die hier von Interesse sind) über eine rekursive Anwendung der Modellgleichung, wobei $\alpha_t = E[\alpha_t] = 0$ gesetzt wird. Ein am Zeitpunkt t über l Zeitschritte vorhergesagter Wert sei mit ${}^*z_t(l)$ bezeichnet:

Beispiel AR(1): $z_t = \phi_1 z_{t-1} + \alpha_t$ (B-21)

$$\begin{array}{l} {}^*z_t(1) = \phi_1 z_t \\ {}^*z_t(l) = \phi_1 \, {}^*z_t(l-1) \\ \hline \rightarrow \quad {}^*z_t(l) = \phi_1^l z_t \end{array} \qquad \square$$

Beispiel AR(2): $z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \alpha_t$ (B-22)

$$\begin{array}{l} {}^*z_t(1) = \phi_1 z_t + \phi_2 z_{t-1} \\ {}^*z_t(2) = \phi_1 \, {}^*z_t(1) + \phi_2 z_t \\ \hline \rightarrow \quad {}^*z_t(l) = \phi_1 \, {}^*z_t(l-1) + \phi_2 \, {}^*z_t(l-2) \end{array} \qquad \square$$

Die Prognoseformeln setzen dabei $z_{MW} = 0$ voraus. Andernfalls muß vor der Prognose die Transformation $z_t \rightarrow z_t - z_{MW}$ angewendet werden werden.

Tab. B-1: Übersicht über die Eigenschaften von autoregressiven Prozessen (AR), Prozessen der gleitenden Mittel (MA) und gemischten ARMA-Prozessen (nach Box & Jenkins (1970))

	AR(p)	MA(q)	ARMA(p,q)
Modell durch z's ausgedrückt	$\Phi(B)z_t=\alpha_t$	$\Theta^{-1}(B)z_t=\alpha_t$	$\Theta^{-1}(B)\Phi(B)z_t=\alpha_t$
α's ausgedrückt	$z_t=\Phi^{-1}(B)\alpha_t$	$z_t=\Theta(B)\alpha_t$	$z_t=\Phi^{-1}(B)\Theta(B)\alpha_t$
Stationaritäts-bedingung	Wurzeln von $\Phi(B)=0$ außerhalb des Einheitskreises	immer stationär	Wurzeln von $\Phi(B)=0$ außerhalb des Einheitskreises
Invertibilitätsbedingung	immer invertierbar	Wurzeln von $\Theta(B)$ außerhalb des Einheitskreises	Wurzeln von $\Theta(B)$ außerhalb des Einheitskreises
ACF	unendlich (gedämpfte Exp.- und/od. Sinusform)	endlich	unendlich (gedämpfte Exp.- und/od.Sinusform nach den ersten q-p lags)
	klingt allmählich aus	**bricht nach lag q ab**	**klingt allmählich aus**
PACF	endlich	unendlich (gedämpfte Exp.- und/od. Sinusform dominiert)	unendlich (gedämpfte Exp.- und/od. Sinusform dominiert nach den ersten p-q lags)
	bricht nach lag p ab	**klingt allmählich aus**	**klingt allmählich aus**

Anhang C

Zeitreihenfilterung

Eine Meßreihe kann auch unter dem Aspekt der Signalverarbeitung analysiert werden. So lassen sich verschiedene Filteralgorithmen anwenden, die die Reihe $\{x_k\}$ in gezielter Weise umformen bzw. interessierende Eigenschaften extrahieren. Unter einem Filter G versteht man allgemein die Anwendung eines Algorithmus der Form

$$y_i = G(x_{k1}, x_{k2}, x_{k3}, \dots) \quad \text{mit } i = 1, \dots, N \text{ und } k1, k2, \dots = 1, \dots, N$$

auf die Meßreihe $\{x_k\}$. Somit kann ein Wert der gefilterten Reihe $\{y_i\}$ Meßwerte x_k verschiedener Zeitpunkte enthalten.

Beispielsweise gibt es Filter zum Differenzieren und zum Integrieren des Meßwerteverlaufs oder auch Filter zum Glätten (z.B. gleitende Mittelung), welche für die Zerlegung der Meßreihe in Komponenten von Interesse sind.

Man kann die bei einer Filterung erfolgenden Veränderungen des Zeitverlaufs besser verstehen, wenn man über eine Fouriertransformation $\mathcal{F}$ des Filters in den Frequenzbereich wechselt. Um dies zu demonstrieren, wird der folgende Glättungsfilter G der gewichteten gleitenden Mittelung über 5 Halbstundenwerte x_j definiert:

$$\begin{aligned} y_i &= G(x_j) \\ &= (x_{i-2}+5x_{i-1}+8x_i+5x_{i+1}+x_{i+2})/20 \\ &= [(B^{-2}+5B^{-1}+8+5B^{1}+B^{2})/20]\ x_i \end{aligned} \qquad \text{(C-1)}$$

In der Darstellung wurde vom Backshiftoperator B^k

$$B^k x_i := x_{i-k} \qquad \text{(C-2)}$$

Gebrauch gemacht. Wendet man diesen Operator B^k auf eine harmonische Schwingung $e^{i\omega t}$ mit der Kreisfrequenz $\omega=2\pi f$ an

$$\begin{aligned} B^k e^{i\omega t} &= e^{i\omega(t-k)} \\ &= e^{i\omega k} e^{i\omega t}, \end{aligned} \qquad \text{(C-3)}$$

dann erhält man seine Fourierdarstellung $B^k(\omega) = e^{-i\omega k}$, mit der man zur Fourierdarstellung des Glättungsfilters $y(\omega) = \mathcal{F}\{y_t\}$ kommt:

$$y(\omega) = \{(e^{+2i\omega} + 5e^{+i\omega} + 8 + 5e^{-i\omega} + e^{-2i\omega})/20\}\, x(\omega)$$
$$= H(\omega)\, x(\omega) \qquad \text{(C-4)}$$

Die Funktion $H(\omega)$, die sich zusammenfassen läßt zu

$$H(\omega) = (2\cos(2\omega) + 10\cos(\omega) + 8)/20, \qquad \text{(C-5)}$$

ist die Transferfunktion, welche multipliziert mit der Amplitude eines Eingangssignals bestimmter Frequenz die Amplitude des Ausgangssignals für diese Frequenz ergibt. Wegen der Symmetrie des Filters ist diese Transferfunktion reell. Bei nichtsymmetrischen Filtern enthält sie dagegen sinus-Terme mit imaginären Koeffizienten, was eine Phasenverschiebung der geglätteten Reihe zur Folge hat.

Wie man an der Transferfunktion (Abb. C-1) des Filters G erkennen kann, handelt es sich um einen Tiefpaß, der die höherfrequenten Anteile aus der Zeitreihe entfernt. Die Frequenz, bei der die Transferfunktion auf 0,5 abgesunken ist (hier bei einer Frequenz von ca. 0,2) wird als Abschneidefrequenz bezeichnet.

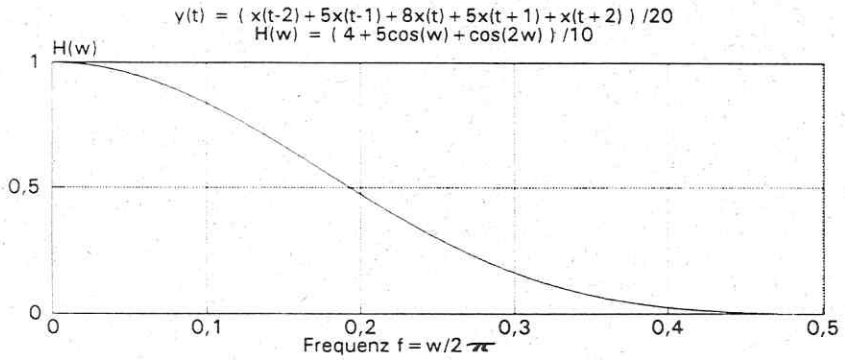


Abb. C-1: Transferfunktion einer symmetrischen gewichteten gleitenden Mittelung

Die Eigenschaften eines solchen Glättungsfilters (z.B. die Steilheit der Flanke und der Wert der Abschneidefrequenz) lassen sich natürlich durch die Wahl der Glättungsspanne, d.h. die Anzahl der in die Glättung einbezogenen Meßwerte und deren Gewichte modifizieren.

In der Praxis werden Zeitreihen häufig durch ungewichtete gleitende Mittelung geglättet. Um eine Vorstellung über die Eigenschaften dieses Filters zu bekommen, ist für den Fall der 24-Stunden-Mittelung (48 Meßwerte) die Transferfunktion in Abb. C-2 skizziert worden.

$$y(t) = (x(t-24) + \ldots + x(t-1) + x(t) + x(t+1) + \ldots x(t+24)) / 49$$
$$H(w) = (2\cos(24w) + 2\cos(23w) + \ldots + 2\cos(2w) + 2\cos(w) + 1) / 49$$

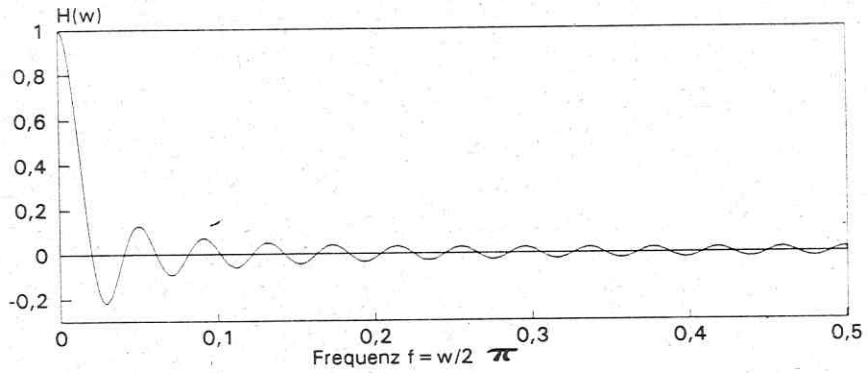


Abb. C-2: Transferfunktion einer symmetrischen ungewichteten gleitenden Mittelung über 48 Meßwerte

Diese Transferfunktion zeigt eine Abschneidefrequenz bei ca. 0,01, d.h. bei sehr langen Perioden (ca. 50 Std.). Als Tiefpaß arbeitet das ungewichtete Mittel jedoch sehr unsauber: Auch höhere Frequenzen können mit unterschiedlicher Intensität passieren. Bei bestimmten Frequenzen gibt es durch das negative Vorzeichen von H(w) sogar eine Phasenumkehr. Diese Unsauberkeiten machen sich auch visuell in einer unruhigen Kontur der gemittelten Reihe bemerkbar. Die Glättung einer Reihe mit einem ungewichteten gleitenden Mittel führt also nicht zu einer Abtrennung hochfrequenter Anteile und ist somit keine optimale Glättung.

Bemerkung: An dieser Stelle soll auch auf eine ganz andere Möglichkeit, die Meßreihe zu glätten und damit die tieffrequenten Anteile herauszufiltern hingewiesen werden. Sie besteht in der lokalen Anpassung von Polynomen (Local Polynomial Approximation, (Peil (1985))). Dabei nutzt man die Tatsache, daß sich jedes globale Modell für die Meßgröße y(t)

$y_t = g(t) + \varepsilon(t)$ $\quad t=1,2,3,\ldots;$ $\varepsilon(t)$=weißes Rauschen; $g(t)$=deterministische Funktion

in eine Taylorreihe

$$y_t = g(t_0) + \ldots + (t-t_0)^m g^{(m)}(t_0)/m! + \varepsilon(t) \qquad \text{(C-6)}$$

entwickeln läßt.

Diese Reihe ist dann ein lokales nichtparametrisches Modell, welches nur in einer durch eine Gewichtsfunktion

$$w(t;t_0,h) := \exp[\, -(t-t_0)^2/(2h^2)\,] \qquad h>0 \qquad \text{(Gaußfunktion)} \qquad \text{(C-7)}$$

definierten Umgebung von t_0 gilt. Durch ein Aneinanderfügen der lokalen approximierten Taylorpolynome erhält man dann die geglättete Reihe. Die Koeffizienten des lokal gültigen Polynoms schätzt man nach der Methode der kleinsten Quadrate aus den Daten der Meßreihe.

Wie stark die jeweiligen Meßwerte in die Schätzer eingehen, wird durch den Parameter h der Gewichtsfunktion gesteuert. Er legt die Größe der für die Schätzung genutzten Umgebung fest. Damit bestimmt er die Stärke der Glättung. Die Verwendung der LPA zur Glättung von Immissionsreihen läßt bei geeigneter Wahl von h deutlich das Tiefpaßverhalten erkennen.□

Anhang D

Der Kalman-Filter

Neben den im Anhang C genannten Filtern, die lediglich die Werte der Meßreihe direkt nutzen, gibt es noch sogenannte rekursive Filter. Bei letzteren wird der Output über Zwischengrößen berechnet.

Bei einer festgelegten Glättungsstärke bewirkt eine nichtrekursive Filterung ein begrenztes "Erinnerungsvermögen" der Outputreihe an die Vergangenheit. Rekursive Filter beziehen dagegen die gesamte Vergangenheit ein und sind deshalb gerade für Extrapolationsaufgaben besser geeignet (Hamming (1973)). Bei der Abarbeitung des rekursiven Algorithmus kann es allerdings Stabilitätsprobleme geben.

Zu den rekursiven Filtern gehört der Kalman-Filter (Kalman (1960)). Dabei wird von einer Zustandsgröße Y ausgegangen, deren zeitliche Veränderung eine Zustandsgleichung beschreibt. Diese enthält bereits stochastische Terme, ist also nicht deterministisch.

Wesentlich für den Filteralgorithmus ist nun die Annahme, daß diese Zustandsgröße zwar existiert, aber nicht direkt gemessen werden kann, sondern nur mit einem überlagerten Rauschanteil. Die Formulierung dieses Modells durch eine Beobachtungs- und eine Zustandsgleichung wird Zustandsraumdarstellung genannt. Aus der Kenntnis der Beobachtungsgröße eine Schätzung für die Zustandsgröße zu finden, ist dann die Aufgabe des Kalman-Filters.

An einem einfachen Zustandsraummodell soll dies demonstriert werden:

Beobachtungsgleichung	$z_t = y_t + \mu_t$	μ_t=weißes Beobachtungsrauschen	(D-1)
Zustandsgleichung	$y_t = y_{t-1} + x_{t-1}$	(IRW-Modell)	
	$x_t = x_{t-1} + \eta_{t-1}$	η_t=weißes Rauschen	

Die beobachtete Meßgröße z_t ist eine mit μ_t verrauschte Zustandsgröße y_t. Für die Dynamik der Zustandsgröße y_t wurde das integrated random walk Modell (IRW) angesetzt. Dieses ergibt sich durch "Aufsummieren" der Glieder einer Random-Walk-Reihe $\{x_t\}$ (vgl. Definition A-4).

Der Kalman-Filter extrahiert dann aus der Meßreihe $\{z_t\}$ einen Schätzer für y_t. Von Ng & Young (1990) wurde der Algorithmus in rekursiver prediction-correction-Form angegeben und seine Eigenschaften untersucht. Die von den Autoren verwendete Bezeichnung IRWSMOOTH soll hier ebenfalls benutzt werden.

Der Parameter für die Glättung ist das Verhältnis var(μ)/var(η), das sogenannte noise-variance-ratio (NVR). IRWSMOOTH ist ein sehr guter Tiefpaß mit einer in der Umgebung der Abschneidefrequenz steil abfallenden Transferfunktion (Abb. D-1).

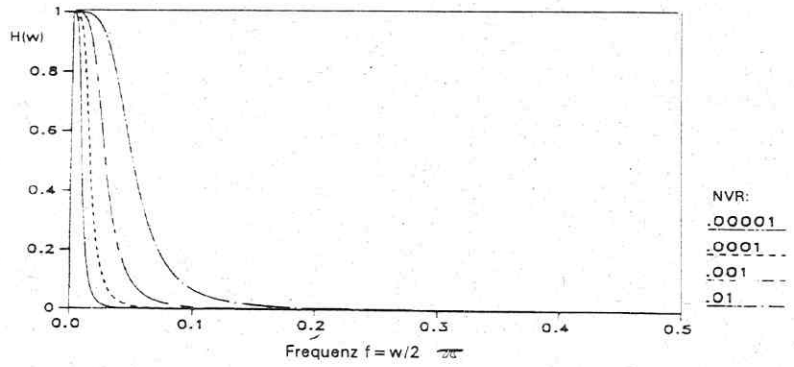


Abb. D-1: Transferfunktion des IRWSMOOTH - Glättungsfilters (aus Ng & Young (1990)).

Zwischen dem Parameter NVR und der Abschneidefrequenz f_{50} gibt es einen einfachen Zusammenhang, so daß sich letztere sehr gut einstellen läßt (Ng & Young (1990)):

$$f_{50} = 0{,}158\ NVR^{0{,}25} \qquad \text{(D-2)}$$

Durch eine temporäre Vergrößerung des NVR lassen sich mit dem rekursiven Kalman-Algorithmus bei Strukturbrüchen der Zeitreihe auch Interventionen modellieren.

Bemerkung (D-1): IRWSMOOTH gehört zu den einfachsten Kalman-Filtern. Natürlich läßt sich das IRW-Modell auch durch ein anderes ersetzen, was die Eigenschaften des Filters modifiziert. Beim IRW gibt die Größe x_t die Ableitung der Zustandsgröße y_t wieder. Sie wird bei der Anwendung des Filters automatisch mit geschätzt und wird in Bereichen, in denen das Filter zur Prognose genutzt wird, konstant gehalten (η=0). Analog dazu kann man durch ein Kalman-Filter mit double integrated random walk (DIRW) erreichen, daß im Prognosebereich die zweite Ableitung der Trendfunktion konstant gehalten wird. Die Zustandsraumdarstellung des DIRW-Modells lautet nämlich:

Beobachtungsgleichung $z_t = y_t + \mu_t$ μ_t=weißes Beobachtungsrauschen (D-3)

Zustandsgleichung (DIRW-Modell)

$$y_t = y_{t-1} + x_{t-1}$$
$$x_t = x_{t-1} + w_{t-1}$$
$$w_t = w_{t-1} + \eta_{t-1}$$

η_t=weißes Rauschen.□

Beobachtungs- und Zustandsgleichung lassen sich verallgemeinern und in eine Vektorform bringen:

$$z_t = \mathbf{H}(t)\, \mathbf{x}_t + \mu_t \tag{D-4}$$

$$\mathbf{x}_t = \mathbf{F}(t-1)\, \mathbf{x}_{t-1} + \mathbf{G}(t-1)\, \eta_{t-1} \tag{D-5}$$

$$\text{wobei } \mathbf{x}_t = (y_t \;\; x_t \;\; \dots)^T \text{ ist.} \tag{D-6}$$

Beispiel (D-1): Für das o.g. integrated random walk Modell hätten dann die Matrizen der Zustandsraumdarstellung und die Kovarianzmatrizen folgendes Aussehen:

$$\mathbf{H}(t) = \mathbf{H} = (1\ 0); \quad \mathbf{G}(t) = \mathbf{G} = \begin{bmatrix} 0 & 0 \\ & \\ 0 & 1 \end{bmatrix}; \qquad \mathbf{P}(t) = \text{cov}(\mathbf{x}_t) \tag{D-7}$$

$$\mathbf{F}(t) = \mathbf{F} = \begin{bmatrix} 1 & 1 \\ & \\ 0 & 1 \end{bmatrix}; \; \mathbf{Q}(t) = \mathbf{Q} = \begin{bmatrix} 0 & 0 \\ & \\ 0 & 1 \end{bmatrix} \cdot \text{NVR} = \text{cov}(\eta_t).\square \tag{D-8}$$

Der Kalman-Filter-Algorithmus basiert nun auf folgenden Gleichungen, die in zwei Schritten (prediction, correction) nacheinander berechnet werden (Ho (1962), Young (1991)):

prediction:

$$\mathbf{x}_{t|t-1} = \mathbf{F}(t-1)\, \mathbf{x}_{t-1} \tag{D-9}$$

$$\mathbf{P}(t|t-1) = \mathbf{F}(t-1)\, \mathbf{P}(t-1)\, \mathbf{F}^T(t-1) + \mathbf{G}(t-1)\, \mathbf{Q}(t-1)\, \mathbf{G}^T(t-1) \tag{D-10}$$

correction:

$$\mathbf{x}_t = \mathbf{x}_{t|t-1} + \mathbf{P}(t|t-1)\, \mathbf{H}^T(t) \left[\mathbf{1} + \mathbf{H}(t)\, \mathbf{P}(t|t-1)\, \mathbf{H}^T(t)\right]^{-1} \{z_t - \mathbf{H}(t)\, \mathbf{x}_{t|t-1}\} \tag{D-11}$$

$$\mathbf{P}(t) = \mathbf{P}(t|t-1) - \mathbf{P}(t|t-1)\, \mathbf{H}^T(t) \left[\mathbf{1} + \mathbf{H}(t)\, \mathbf{P}(t|t-1)\, \mathbf{H}^T(t)\right]^{-1} \mathbf{H}(t)\, \mathbf{P}(t|t-1) \tag{D-12}$$

Der Algorithmus wird rekursiv auf z_t angewendet und führt auf die bedingte Zustandsreihe $\mathbf{x}_{t|t-1}$. Für die Glättung ist noch ein rekursives absteigendes Filter erforderlich:

$$\mathbf{x}_{t|N} = \mathbf{F}^{-1}(t) \left[\mathbf{x}_{t+1|N} + \mathbf{G}(t)\, \mathbf{Q}(t)\, \mathbf{G}^T(t)\, \mathbf{L}(t)\right] \tag{D-13}$$

$$\mathbf{L}(t) = \left[\mathbf{I} - \mathbf{P}(t+1)\, \mathbf{H}^T(t+1)\, \mathbf{H}(t+1)\right]^T \left[\mathbf{F}^T(t+1)\, \mathbf{L}(t+1) - \mathbf{H}^T(t+1)\right] \{y_{t+1} - \mathbf{H}(t)\, \mathbf{F}(t)\, \mathbf{x}_t\} \tag{D-14}$$

$$\text{mit } \mathbf{L}(N) = \mathbf{0}. \tag{D-15}$$

Bemerkung (D-2): Interessanterweise gibt es eine exakte Analogie der Zustandsraumdarstellung zum Problem der Regression mit zeitabhängigen Parametern (Young, Ng, Armitage (1989)). Dabei wird die Beobachtungsgleichung als Regressionsgleichung und die Zustandsgleichung als Modellgleichung für die Zeitabhängigkeit der Regressionsparameter angesehen. Die Umwandlung der expliziten Lösung des Regressionsproblems in eine rekursive gelang Ho (1962).

Mit diesem einen Ansatz der Zustandsraumgleichungen können also zwei wesentliche Aufgaben der Zeitreihenmodellierung bewältigt werden: Das Abtrennen von Komponenten durch Filterung bzw. Glättung und die Verwendung von Regressionsmodellen mit zeitabhängigen Parametern für diese Komponenten. Die Lösungsalgorithmen sind somit für beide Aufgaben identisch. Diese Tatsache eröffnet dem Kalman-Algorithmus im Bereich der Zeitreihenanalyse äußerst breite Anwendungsmöglichkeiten.□

Beispiel (D-2): Sehr nützlich ist die harmonische Regression

$$s_t = {}_1\theta_t \cos \omega t + {}_2\theta_t \sin \omega t \qquad \text{(D-16)}$$

mit zeitabhängigen Parametern ${}_1\theta_t$ und ${}_2\theta_t$. Die Dynamik der Parameter wird durch Random Walk (s. Anhang A) modelliert:

(D-17)

$${}_1\theta_t = {}_1\theta_{t-1} + \eta_t$$
$${}_2\theta_t = {}_2\theta_{t-1} + \eta_t.$$

Dieses Modell wird als **Dynamische Harmonische Regression** (DHR) bezeichnet. Die Anwendung des Kalman-Filter-Algorithmus führt auf eine rekursive Schätzung der Parameterwerte ${}_1\theta_t$ und ${}_2\theta_t$.□

Im Anhang B war bereits eine andere Klasse von Modellen vorgestellt worden: Die ARIMA-Modelle. Welche Beziehungen bestehen zwischen ARIMA- und Zustandsraummodellen?

Typisch für die ARIMA-Modellierung ist es, daß die Zeitreihe als Ganzes aufgefaßt wird. Es ist aber auch möglich, ARIMA-Modelle in die Zustandsraumformulierung zu bringen.

Beispiel (D-3): Exemplarisch sei hier das AR(2)-Modell

$$(1 - \phi_1 B - \phi_2 B^2)\, z_t = \varepsilon_t \qquad \text{(D-18)}$$

d.h.,

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \varepsilon_t. \qquad \text{(D-19)}$$

angeführt.

Es gilt die Stationaritätsbedingung $\phi_1 + \phi_2 < 1$. (D-20)

Das Modell wird also durch die eine Gleichung (D-18) repräsentiert. Um das AR(2)-Modell in die Zustandsraumdarstellung zu bringen, werden neue Variable eingeführt:

$$y_t = z_t,$$
$$x_t = z_{t-1} \qquad \text{(D-21)}$$

wodurch die

Zustandsgleichungen

$$y_t = \phi_1 y_{t-1} + \phi_2 x_{t-1} + \varepsilon_t \qquad \text{(D-22)}$$
$$x_t = y_{t-1}$$

entstehen.

Der vektorielle Zustand $(y\ x)^T$ zum Zeitpunkt t wird aus dem Zustand zum Zeitpunkt t-1 rekursiv berechnet. Durch Einsetzen beider Gleichungen ineinander läßt sich die Äquivalenz zum AR(2)-Modell bestätigen. Sie ergibt sich aus der allgemeinen Tatsache, daß eine skalare Differenzengleichung p-ter Ordnung als (p-dimensionale) vektorielle Differenzengleichung 1. Ordnung formuliert werden kann. Um vom Zustandsvektor auf die beobachtete Größe z_t zu kommen, muß noch die Beobachtungsgleichung

$$z_t = (1\ \ 0)\ (y_t\ \ x_t)^T \qquad \text{(D-23)}$$

gelten. Anstelle von Gleichung (D-18) kann für das AR(2)-Modell also auch die Zustandsraumdarstellung (D-22)+(D-23) angegeben werden.□

Anhang E

Software

Die verwendeten Konzentrationsreihen wurden mit Hilfe eines Datenverwaltungsprogramms aus dem Datenpool der Immissionsmessungen gewonnen.

Für die statistischen Auswertungen kam PC-ISP/DGS+ zur Anwendung. Es handelt sich dabei um eine Statistik-Software, die interpretativ arbeitet und besonders für die interaktive Arbeit im Dialog geeignet ist. ISP ist als Befehlssprache implementiert, so daß es sehr flexibel und universell einsetzbar ist. Die in dieser Sprache enthaltenen Befehle sind teilweise äußerst mächtig und erlauben so eine rasche Problembearbeitung.

Im allgemeinen ist die Verarbeitungsgeschwindigkeit in ISP sehr hoch, da das System matrixorientiert arbeitet und die Ausführung von Schleifen weitestgehend intern und sehr schnell erfolgt. Spezielle, in den hier gemachten Untersuchungen genutzte Algorithmen sind nicht in ISP implementiert, so daß dafür Makros erstellt werden mußten. Diese Makros erfordern jedoch wegen der interpretativen Abarbeitung relativ viel Rechenzeit. Dies ist besonders bei dem Makro für die Kalman-Filterung der Fall. Die Entwicklung eines ausführbaren Programms zur Kalman-Filterung, welches sich dann problemlos in ISP einbinden ließe, könnte weitere Arbeiten auf diesem Gebiet wesentlich intensivieren.

Eine sehr aktuelle Software zur Zeitreihenanalyse ist das von (Young et.al. (1991)) entwickelte Programm microCAPTAIN. Es arbeitet menügesteuert die wesentlichsten Schritte der von den Autoren entwickelten Methoden zur Modellierung von Zeitserien ab. Die bei der Anwendung von microCAPTAIN auf die hier interessierenden Immissionsdaten gemachten Erfahrungen bildeten einen Teil des wissenschaftlichen Ausgangspunktes der vorliegenden Arbeit.

Als Beispiel für ein ISP-Macro enthalten die folgenden Seiten das Listing des Macros für die Zeitreihenprognose mit lokaler harmonischer Trendextrapolation.

```
#                                                    Prognose mit Komponentenmodell und lokaler harmonischer Trendextrapolation
#       mit nt wird die Zahl der zu verwendenden Tage eingestellt
#
set "termin=n
input/macro > prognk
#########################################################################
#  macro fuer die laufende prognose    einer zeitreihe                  #
#  unter nutzung von komponentenmodell und kalman-filter                 #
#  mit lokaler harmonischer Trendextrapolation                           #
#########################################################################
#
cli "cls"
set pty=l
set out=n
set interact=y
set ymin=0
set ymax=1.5
set labxtix=-2,-1,0,1
set locxtix=-2,-1,0,1
pi=3.1415926
#
# einlesen der gesamten messreihe
input/dims=480/format=(f10.6) 'sample.all' > al
al= if (al<=0) then ? else al
# al=al/1000
#
# schleife fuer die darstellung und prognose
nt=4        # nt Tage
laenge=nt*48
i=iota(array(laenge))/48-(nt-1)
anf=48*3+1
weiter=6
while ((weiter==6)|(weiter==4)|(weiter==5)|(weiter==9))
  e=anf+laenge-1
  #
  # herausgreifen des darzustellenden abschnitts von nt tagen
  reihe=al(anf:e)
  #
  # herausgreifen der ersten (nt-1) tage als prognosebasis
  last=(nt-1)*48
  basis= if (index(reihe,1)<((nt-1)*48)) then reihe else ?
  zukunft= if (index(reihe,1)<((nt-1)*48)) then ? else reihe
  #
  # varianzstabilisierende trafo nach box-cox
  transf=(sqrt(basis)-1)*2
  #
  # kalman-filterung mit dem komponentenmodell fuer die prognoseerstellung
  putstr "Berechnung laeuft..."
  kalman transf > prognos trend abl p1c p1s p2c p2s
  alt=trend
  palt=prognos
  delete prognos trend abl p1c p1s p2c p2s
  transf=transf(1:last)
  kalman transf > prognos trend abl p1c p1s p2c p2s
  delete prognos
  #
  # ellipsen-schaetzung im phasenraum
  ell trend abl > tp
  #
  # fortsetzen der parameter in den prognosebereich:
  # 1. trendanteil
  tr=tp
  # 2. periodischen anteil
  p=p1c(0)
  reshape/dims=48/cycle p > p
  glue/axis=1 p1c p > p1c
  p=p1s(0)
  reshape/dims=48/cycle p > p
  glue/axis=1 p1s p > p1s
  p=p2c(0)
  reshape/dims=48/cycle p > p
  glue/axis=1 p2c p > p2c
  p=p2s(0)
  reshape/dims=48/cycle p > p
  glue/axis=1 p2s p > p2s
  j=iota(p1c)
  p=p1c*cos(2*pi*j/48)+p1s*sin(2*pi*j/48)+p2c*cos(2*pi*j/24)+p2s*sin(2*pi*j/24)
  # gesamtfunktion:
  f=tr+p
  lastp=last+1
  laste=last+48
  glue/axis=1 (trend(1:last)) (tr(lastp:laste)) > trend
  # jetzt alles darstellen !!!
  # vorher ruecktransformation
  prognos=(f/2+1)**2
  trend=(trend/2+1)**2
  alt=(alt/2+1)**2
  palt=(palt/2+1)**2
  #
  # darstellung des gemessenen und des prognostizierten verlaufs
  gscat i reihe > oreihe /interact=n/out=y/pcol=7
  gscat i zukunft > ozuk /interact=n/out=y/pcol=4
  gscat i prognos > oprogn /interact=n/out=y/pcol=
  gscat i trend > otrend /interact=n/out=y/pcol=
  gscat i alt > oalt/interact=n/out=n/pcol=<
  gscat i palt > opalt/interact=n/out=n/pcol=;
  tag=ftos(int(anf/48)+1)
  cli "cls"
gplot oreihe ozuk oprogn otrend oalt opalt /scrw=0,100,1000,800/xlab='days'/ylab=SO2/title='Prognose '/out=y
putstr ""
putstr ""
putstr ("erster Tag = $tag%")
```

```
  #
  # weitere ablaufsteuerung
    weiter=6
    anf= if (weiter==6) then (anf+48) else anf
    anf= if (weiter==4) then (anf-48) else anf
    anf= if (weiter==5) then (anf+5*48) else anf
    anf= if (weiter==9) then (anf+9*48) else anf
    if (weiter==0) then
      quit
    end if
end while
cli 'cls'
putstr "Macro-Ende"
return

#########################################################################
set *termin=y


set *termin=n
input/macro > kalman
#########################################################################
# kalman - filterung nach dem komponentenmodell                         #
#  mit IRWSMOOTH fuer trend                                             #
# kalman input > output                                                 #
#########################################################################
# bei der programmierung war darauf zu achten, dass H als zeilenvektor
# in der form trn(h) zu verarbeiten ist. ansonsten gelten die beziehungen
# in der ueblichen notation. die vektoren sind immer spaltenvektoren.
pi=3.1415926
getarg "input-reihe " > input
y=input
#
# 1. Festlegungen
nd=dims(y)
# einheitsmatrix 6x6 erzeugen
glue (1) (array(6)) > tmp
reshape/cycle=y/dims=6,6 tmp > unit6x6
delete tmp
# Periodizitaeten
per1=48
per2=24
# Matrixdefinitionen
f=unit6x6
f(1,2)=1
g=unit6x6
q=0*f
q(2,2)=0.00001          # NVR fuer trend
q(3,3)=0.001            # NVR fuer periode 1
q(4,4)=0.001            # NVR fuer periode 1
q(5,5)=0.001            # NVR fuer periode 2
q(6,6)=0.001            # NVR fuer periode 2
h=array(6)
x=0*array(6,nd)  # zustandsvariable x(k)
xb=0*array(6,nd) # bedingtes x(k|k-1)
#
# 2. Startwerte
x(1,1)= if (missing(y(1))) then 0 else (y(1))
p=0*array(6,6,nd)
p(*,*,1)=10*unit6x6
pb=0*f                          # bedingtes P(k|k-1)
pw=p(*,*,1)                     # p(k) wert zu einem festen k
#
# 3. Rekursion, filtering-pass
nm=dims(y)-1
tmp1= if (missing(y(1))) then 0 else (y(1))
do k=iota(array(nm))+1
  km=k-1
  glue (1) (0) (cos(2*pi*km/per1)) (sin(2*pi*km/per1)) (cos(2*pi*km/per2)) (sin(2*pi*km/per2)) > h
  # prediction:
  xb(*,k)=mpy(f,x(*,km))
  pb=mpy(f,mpy(pw,trn(f)))+mpy(g,mpy(q,trn(g)))
  #
  wert= if (missing(y(k))) then (tmp1) else (y(k))
  y(k)=wert                     # ersetzt fehlende messwerte durch den vorherigen bedingten
  tmp1= mpy(trn(h),xb(*,k))
  #
  # correction:
  nn=1/(1+mpy(trn(h),mpy(pb,h)))                    # nn = nenner
  x(*,k)=xb(*,k)+mpy(pb,h)*nn*(y(k)-mpy(trn(h),xb(*,k)))
  pw=pb-mpy(mpy(pb,h),mpy(trn(h),pb))*nn
  p(*,*,k)=pw
end do
# 4. smoothing-pass
ys=0*y
tr=ys
l=0*x
z=0*x
z(*,nd)=x(*,nd)
ein=unit6x6
do k=nd-(iota(array(nd))-1)
  kp=k+1
  km=k-1
  glue (1) (0) (cos(2*pi*k/per1)) (sin(2*pi*k/per1)) (cos(2*pi*k/per2)) (sin(2*pi*k/per2)) > h
  if (k<nd) then
    hwa=trn(ein-mpy(p(*,*,kp),mpy(h,trn(h))))
    hwb=mpy(trn(f),l(*,kp))-h*(y(kp)-mpy(trn(h),mpy(f,x(*,k))))
    l(*,k)=mpy(hwa,hwb)
    z(*,k)=mpy(inv(f),(z(*,kp)+mpy(mpy(g,q),mpy(trn(g),l(*,k)))))
  end if
```

```
  # geglaetteter schaetzer ys fuer die original-reihe
  tmps=mpy(trn(h),z(*,k))
  ys(k)=tmps
end do
tr=trn(z(1,*))
abl=trn(z(2,*))
p1c=trn(z(3,*))
p1s=trn(z(4,*))
p2c=trn(z(5,*))
p2s=trn(z(6,*))
output=ys
putarg "output-reihe " output
putarg "trend " tr
putarg "ableit " abl
putarg "p1c " p1c
putarg "p1s " p1s
putarg "p2c " p2c
putarg "p2s " p2s
return

set *termin=y
#########################################################################
```

```
set *termin=n
input/macro > elli4
#########################################################################
# elli   - lokale harmonische Trendextrapolation durch Ellipsenberechnung im Phasenraum      #
# elli trend ableitg > trendprog                                                              #
# trendprog ist um 48 laenger als trend                                                       #
#########################################################################
#
pi=3.1415926
getarg "Trend " > t
getarg "Ableitung " > d
ig=ftos(dims(t)+48)
te=t
de=d
omega1=0*te
omega2=omega1
i=4
nm=dims(te)-3*i
x0=0*te
x1=0*te
x2=0*te
x3=0*te
do k=iota(array(nm))+3*i
  k3=k
  k2=k-i
  k1=k-2*i
  k0=k-3*i
  xi=te(k3)
  x3(k)=xi
  xi=te(k2)
  x2(k)=xi
  xi=te(k1)
  x1(k)=xi
  xi=te(k0)
  x0(k)=xi
end do
x0=x0**1000
x1=x1**1000
x2=x2**1000
x3=x3**1000
v2=(x2-x0)/(x1-x0)
v3=(x3-x0)/(x1-x0)
omega=sqrt(4*(v3-3*(v2-1))/(v3-7*(v2-1)))
omega=omega/i
ij=iota(te)
cli 'cls'
lo=dims(omega)
lo=lo-40
la=lo-20
mo=mean(omega(la:lo))
mo= if (missing(mo)) then 0.001 else mo
mos=ftos(mo)
gscat ij omega /ymin=0/ymax=0.1/locxtix=/labxtix=/out=y/title=$mos%/scrw=0,100,1000,400
omega=mo
#
pp=dims(te)
np=48
beg=pp-(np-1)
x=te(beg:pp)
xs=de(beg:pp)
#
da=dims(x)+48
ti=iota(array(da))-1
# ti ist die zeit
# die gleich die 48 prognoseschritte mitenthaelt
# es kann aber nur im alten messbereich angepasst werden, deshalb:
ta=ti(1:np)
```

```
#
#  schaetzen der parameter t0, a, c
#
glue/axis=2 (cos(omega*ta)) (sin(omega*ta)) > bb
regress bb x > res coef/const=y
t0=atan(coef(2)/coef(1))
a=-coef(1)/cos(t0)
c=coef(3)
#
print omega t0 a c
#
# fertige ellipse im alten bereich
sx=c-a*cos(omega*ta-t0)
gscat (iota(x)) x > oxx
gscat (iota(sx)) sx > osx
cli 'cls'
gplot oxx osx /ymin=/ymax=/labxtix=/locxtix=/out=y
sy=a*omega*sin(omega*ta-t0)
# diese ellipse soll jetzt am letzten schnittpunkt mit der originalen
# phasenraumkurve an diese anschliessen und sie spaeter weiterfuehren
# der abstand zwischen ellipse und originalkurve ist:
delta=(x-trn(sx))*(x-trn(sx))/x/x+(xs-trn(sy))*(xs-trn(sy))/xs/xs
#
dm=min(trn(min(delta)))
mi1= if (delta==dm) then (index(delta,1)) else ?
mi2= if (delta==dm) then (index(delta,2)) else ?
mi1=sum(trn(sum(mi1)))
mi2=sum(trn(sum(mi2)))
# somit sind der alte und der neue abschnitt des prognosetrends
mib=beg+mi1-1
trendalt=t(1:mib)
dalt=d(1:mib)
#
mip=mi2+1
trendneu=c-a*cos(omega*ti(1:np)-t0)
dneu=a*omega*sin(omega*ti(1:np)-t0)
trendneu=c-a*cos(omega*ti(mip:0)-t0)
dneu=a*omega*sin(omega*ti(mip:0)-t0)
glue trendalt trendneu > trendprog
glue dalt dneu > dprog
reshape /dims=$lg% trendprog > trendprog
reshape /dims=$lg% dprog > dprog
#
gscat trendprog dprog > oprog /pcol=4
gscat trendneu dneu > oneu /pcol=4
gscat t d > oalt /pcol=?
gscat x xs > ox /pcol=>
cli 'cls'
gplot oalt oprog ox /ymin=/ymax=/locxtix=/labxtix=/out=y
#
putarg "trendprog " trendprog
return

set *termin=y
#########################################################################
```

Anhang F

1) Die Meßreihe CITY für Leipzig-Mitte in der Zeit vom 01.-10.03.1989, Halbstundenmittelwerte der SO_2-Konzentration in mg/m³

Uhrzeit \ Datum	01/03/89	02/03/89	03/03/89	04/03/89	05/03/89
00:00	0,11	0,16	0,33	0,34	0,33
00:30	0,11	0,13	0,38	0,3	0,28
01:00	0,12	0,14	0,39	0,28	0,27
01:30	0,11	0,13	0,41	0,23	0,23
02:00	0,1	0,12	0,45	0,21	0,23
02:30	0,09	0,12	0,36	0,16	0,2
03:00	0,09	0,14	0,21	0,16	0,17
03:30	0,09	0,15	0,13	0,14	0,15
04:00	0,08	0,15	0,16	0,16	0,16
04:30	0,07	0,13	0,16	0,16	0,15
05:00	0,09	0,13	0,15	0,18	0,19
05:30	0,08	0,11	0,17	0,15	0,15
06:00	0,12	0,14	0,24	0,15	0,2
06:30	0,17	0,23	0,29	0,15	0,15
07:00	0,26	0,4	0,36	0,16	0,21
07:30	0,29	0,59	0,37	0,16	0,18
08:00	0,25	0,65	0,45	0,2	0,24
08:30	0,22	0,63	0,56	0,24	0,21
09:00	0,22	0,87	0,7	0,33	0,32
09:30	0,27	0,69	0,66	0,33	0,48
10:00	0,28	0,64	0,74	0,42	0,63
10:30	0,34	0,5	0,8	0,54	0,54
11:00	0,46	0,52	0,83	0,53	0,54
11:30	0,56	0,54	0,47	0,58	0,56
12:00	0,57	0,52	0,54	0,6	0,61
12:30	0,5	0,49	0,62	0,48	0,55
13:00	0,48	0,45	0,54	0,69	0,49
13:30	0,53	0,48	0,57	0,5	0,5
14:00	0,5	0,43	0,52	0,45	0,65
14:30	0,46	0,41	0,57	0,33	0,48
15:00	0,41	0,35	0,66	0,29	0,5
15:30	0,33	0,46	0,75	0,29	0,42
16:00	0,32	0,39	0,76	0,31	0,41
16:30	0,31	0,36	0,77	0,28	0,34
17:00	0,31	0,38	0,67	0,33	0,35
17:30	0,28	0,32	0,46	0,32	0,32
18:00	0,27	0,34	0,53	0,35	0,33
18:30	0,24	0,28	0,85	0,3	0,25
19:00	0,22	0,27	1,07	0,28	0,28
19:30	0,2	0,27	1	0,29	0,22
20:00	0,21	0,32	1,03	0,37	0,31
20:30	0,22	0,25	1,17	0,41	0,36
21:00	0,24	0,25	0,95	0,47	0,45
21:30	0,2	0,26	0,8	0,46	0,48
22:00	0,19	0,33	0,6	0,41	0,59
22:30	0,17	0,29	0,38	0,34	0,46
23:00	0,19	0,27	0,37	0,34	0,53
23:30	0,17	0,26	0,32	0,32	0,44

Uhrzeit \ Datum	06/03/89	07/03/89	08/03/89	09/03/89	10/03/89
00:00	0,43	0,38	0,43	0,1	0,18
00:30	0,35	0,39	0,39	0,1	0,19
01:00	0,38	0,41	0,36	0,1	0,18
01:30	0,31	0,3	0,28	0,09	0,2
02:00	0,32	0,33	0,29	0,07	0,19
02:30	0,28	0,26	0,29	0,07	0,19
03:00	0,31	0,34	0,25	0,07	0,17
03:30	0,24	0,3	0,21	0,09	0,17
04:00	0,31	0,35	0,22	0,07	0,17
04:30	0,27	0,36	0,19	0,08	0,19
05:00	0,37	0,51	0,21	0,08	0,21
05:30	0,35	0,52	0,21	0,1	0,25
06:00	0,51	0,62	0,28	0,08	0,38
06:30	0,55	0,62	0,33	0,14	0,42
07:00	0,67	0,81	0,34	0,25	0,43
07:30	0,73	0,78	0,55	0,35	0,53
08:00	0,95	0,96	0,47	0,29	0,59
08:30	1,06	1,29	0,5	0,32	0,57
09:00	1,44	1,25	0,61	0,39	0,51
09:30	1,89	0,97	0,92	0,58	0,52
10:00	1,83	0,97	0,95	0,67	0,51
10:30	1,81	0,88	0,56	0,67	0,45
11:00	1,76	0,96	0,48	0,47	0,32
11:30	1,82	0,85	0,4	0,37	0,42
12:00	1,39	0,86	0,44	0,36	0,44
12:30	1	0,71	0,26	0,32	0,42
13:00	0,81	0,65	0,24	0,25	0,4
13:30	0,61	0,46	0,25	0,23	0,38
14:00	0,83	0,02	0,18	0,25	0,29
14:30	0,87	0,71	0,23	0,31	0,21
15:00	0,73	1,01	0,22	0,29	0,21
15:30	0,68	0,7	0,28	0,33	0,24
16:00	0,78	0,68	0,28	0,29	0,2
16:30	0,76	0,59	0,27	0,34	0,14
17:00	0,61	0,53	0,22	0,32	0,21
17:30	0,43	0,38	0,2	0,27	0,23
18:00	0,59	0,44	0,19	0,29	0,25
18:30	0,66	0,64	0,17	0,35	0,23
19:00	0,82	0,56	0,11	0,35	0,25
19:30	0,68	0,56	0,12	0,3	0,32
20:00	0,62	0,53	0,11	0,23	0,25
20:30	0,52	0,53	0,1	0,23	0,2
21:00	0,58	0,5	0,09	0,21	0,19
21:30	0,47	0,49	0,09	0,19	0,17
22:00	0,46	0,48	0,08	0,19	0,15
22:30	0,32	0,42	0,09	0,19	0,14
23:00	0,37	0,45	0,1	0,17	0,15
23:30	0,34	0,4	0,11	0,18	0,13

Anhang F

2) Liste der 191 Prognosesituationen an der Station Leipzig-Mitte von 1980 bis 1993

Druckprotokoll der Meßstelle STA 182022 01 Leipzig-Mitte SO_2
ausgewertetes Intervall: 01.01.80 00:00 - 31.12.93 23:30

Grenzwert: 0.600 mg/m^3 bezogen auf das 3.0 h-Mittel
Intervall: 96 Werte vor t_o, 192 Werte gesamt
Vor t[krit] ist für mindest. 24.0 h das gleitende Mittel über 24.0 h < 0.600 mg/m^3.
(*), wenn innerhalb der auf t_o folgenden 24.0 h das gleitende Mittel über 24.0 h den Wert 0.600 mg/m^3 erreicht. Die Zeit bis zur Überschreitung ist in Klammern angegeben.

lfd.Nr.	Dateiname	Intervall	Zeit bis Überschreitung
1.	S001A182.022	20.02.80 09:00 - 24.02.80 08:30	
2.	S002A182.022	20.02.80 14:00 - 24.02.80 13:30	
3.	S003A182.022	20.02.80 16:00 - 24.02.80 15:30	
4.	S004A182.022	21.02.80 09:00 - 25.02.80 08:30	
5.	S005A182.022	12.03.80 08:30 - 16.03.80 08:00	
6.	S006A182.022	12.04.80 07:30 - 16.04.80 07:00	
7.	S007A182.022	15.10.80 18:00 - 19.10.80 17:30	
8.	S008A182.022	20.10.80 18:00 - 24.10.80 17:30	
9.	S009A182.022	05.11.80 09:00 - 09.11.80 08:30	
10.	S001X182.022	06.11.80 09:30 - 10.11.80 09:00	*(12.5h)
11.	S010A182.022	09.11.80 07:30 - 13.11.80 07:00	
12.	S011A182.022	17.11.80 09:00 - 21.11.80 08:30	
13.	S012A182.022	04.01.81 08:00 - 08.01.81 07:30	
14.	S013A182.022	05.01.81 09:00 - 09.01.81 08:30	
15.	S014A182.022	17.01.81 08:00 - 21.01.81 07:30	
16.	S002X182.022	25.01.81 07:00 - 29.01.81 06:30	*(4.0h)
17.	S015A182.022	05.04.81 08:30 - 09.04.81 08:00	
18.	S016A182.022	18.04.81 10:00 - 22.04.81 09:30	
19.	S017A182.022	19.04.81 10:30 - 23.04.81 10:00	
20.	S018A182.022	25.04.81 09:30 - 29.04.81 09:00	
21.	S019A182.022	25.04.81 15:00 - 29.04.81 14:30	
22.	S020A182.022	15.10.81 10:30 - 19.10.81 10:00	
23.	S021A182.022	16.10.81 16:00 - 20.10.81 15:30	
24.	S022A182.022	24.10.81 10:00 - 28.10.81 09:30	
25.	S023A182.022	25.10.81 08:30 - 29.10.81 08:00	
26.	S024A182.022	08.11.81 11:30 - 12.11.81 11:00	
27.	S025A182.022	08.11.81 18:00 - 12.11.81 17:30	
28.	S026A182.022	15.11.81 07:00 - 19.11.81 06:30	
29.	S027A182.022	15.11.81 17:00 - 19.11.81 16:30	
30.	S028A182.022	17.11.81 18:30 - 21.11.81 18:00	
31.	S029A182.022	24.11.81 09:00 - 28.11.81 08:30	
32.	S003X182.022	25.12.81 10:30 - 29.12.81 10:00	*(5.5h)
33.	S030A182.022	21.01.82 12:00 - 25.01.82 11:30	
34.	S031A182.022	22.01.82 11:30 - 26.01.82 11:00	
35.	S032A182.022	23.01.82 10:00 - 27.01.82 09:30	
36.	S033A182.022	26.01.82 09:00 - 30.01.82 08:30	
37.	S004X182.022	02.02.82 08:30 - 06.02.82 08:00	*(13.0h)
38.	S034A182.022	14.02.82 08:30 - 18.02.82 08:00	
39.	S005X182.022	19.02.82 20:00 - 23.02.82 19:30	*(10.5h)
40.	S006X182.022	20.02.82 04:30 - 24.02.82 04:00	*(2.0h)
41.	S035A182.022	06.03.82 08:00 - 10.03.82 07:30	
42.	S036A182.022	07.03.82 08:00 - 11.03.82 07:30	
43.	S037A182.022	16.03.82 09:30 - 20.03.82 09:00	
44.	S038A182.022	18.03.82 10:00 - 22.03.82 09:30	
45.	S039A182.022	21.03.82 08:30 - 25.03.82 08:00	
46.	S040A182.022	22.03.82 09:00 - 26.03.82 08:30	
47.	S041A182.022	21.10.82 11:30 - 25.10.82 11:00	
48.	S042A182.022	23.10.82 15:00 - 27.10.82 14:30	

49.	S043A182.022	24.10.82	19:00	- 28.10.82	18:30	
50.	S044A182.022	26.10.82	10:00	- 30.10.82	09:30	
51.	S045A182.022	28.10.82	11:00	- 01.11.82	10:30	
52.	S046A182.022	06.11.82	08:00	- 10.11.82	07:30	
53.	S047A182.022	07.11.82	09:30	- 11.11.82	09:00	
54.	S048A182.022	01.12.82	11:00	- 05.12.82	10:30	
55.	S049A182.022	01.12.82	13:00	- 05.12.82	12:30	
56.	S050A182.022	02.12.82	13:00	- 06.12.82	12:30	
57.	S051A182.022	05.12.82	09:00	- 09.12.82	08:30	
58.	S052A182.022	17.12.82	12:30	- 21.12.82	12:00	
59.	S053A182.022	29.12.82	12:00	- 02.01.83	11:30	
60.	S054A182.022	01.02.83	12:30	- 05.02.83	12:00	
61.	S055A182.022	02.02.83	07:00	- 06.02.83	06:30	
62.	S007X182.022	12.02.83	00:30	- 16.02.83	00:00	*(9.5h)
63.	S008X182.022	12.02.83	07:00	- 16.02.83	06:30	*(3.0h)
64.	S056A182.022	02.04.83	09:30	- 06.04.83	09:00	
65.	S057A182.022	03.11.83	11:00	- 07.11.83	10:30	
66.	S058A182.022	03.11.83	18:30	- 07.11.83	18:00	
67.	S059A182.022	04.11.83	19:30	- 08.11.83	19:00	
68.	S060A182.022	05.11.83	10:00	- 09.11.83	09:30	
69.	S009X182.022	06.11.83	18:00	- 10.11.83	17:30	*(20.5h)
70.	S010X182.022	07.11.83	09:30	- 11.11.83	09:00	*(5.0h)
71.	S061A182.022	25.12.83	09:30	- 29.12.83	09:00	
72.	S062A182.022	10.01.84	13:00	- 14.01.84	12:30	
73.	S063A182.022	10.01.84	16:30	- 14.01.84	16:00	
74.	S064A182.022	21.01.84	18:30	- 25.01.84	18:00	
75.	S065A182.022	22.01.84	10:00	- 26.01.84	09:30	
76.	S066A182.022	23.01.84	07:30	- 27.01.84	07:00	
77.	S011X182.022	24.01.84	08:00	- 28.01.84	07:30	*(3.5h)
78.	S067A182.022	05.02.84	09:30	- 09.02.84	09:00	
79.	S012X182.022	24.02.84	11:00	- 28.02.84	10:30	*(10.0h)
80.	S068A182.022	17.03.84	08:30	- 21.03.84	08:00	
81.	S069A182.022	03.04.84	10:30	- 07.04.84	10:00	
82.	S070A182.022	04.04.84	08:30	- 08.04.84	08:00	
83.	S071A182.022	28.04.84	08:30	- 02.05.84	08:00	
84.	S072A182.022	19.10.84	10:00	- 23.10.84	09:30	
85.	S073A182.022	27.10.84	14:00	- 31.10.84	13:30	
86.	S013X182.022	02.02.85	09:30	- 06.02.85	09:00	*(7.5h)
87.	S014X182.022	18.03.85	10:30	- 22.03.85	10:00	*(21.5h)
88.	S015X182.022	18.03.85	18:30	- 22.03.85	18:00	*(13.5h)
89.	S016X182.022	19.03.85	06:00	- 23.03.85	05:30	*(2.0h)
90.	S074A182.022	03.04.85	18:30	- 07.04.85	18:00	
91.	S075A182.022	04.04.85	00:30	- 08.04.85	00:00	
92.	S076A182.022	13.04.85	07:30	- 17.04.85	07:00	
93.	S077A182.022	14.04.85	08:00	- 18.04.85	07:30	
94.	S017X182.022	15.04.85	08:30	- 19.04.85	08:00	*(5.5h)
95.	S078A182.022	27.04.85	07:30	- 01.05.85	07:00	
96.	S079A182.022	29.04.85	11:00	- 03.05.85	10:30	
97.	S080A182.022	02.06.85	09:30	- 06.06.85	09:00	
98.	S081A182.022	19.10.85	10:00	- 23.10.85	09:30	
99.	S082A182.022	27.10.85	14:00	- 31.10.85	13:30	
100.	S083A182.022	28.10.85	11:30	- 01.11.85	11:00	
101.	S084A182.022	29.10.85	10:00	- 02.11.85	09:30	
102.	S085A182.022	11.11.85	08:30	- 15.11.85	08:00	
103.	S018X182.022	13.11.85	08:30	- 17.11.85	08:00	*(13.5h)
104.	S019X182.022	18.11.85	07:30	- 22.11.85	07:00	*(3.0h)
105.	S086A182.022	14.12.85	08:30	- 18.12.85	08:00	
106.	S087A182.022	20.12.85	11:00	- 24.12.85	10:30	
107.	S088A182.022	20.12.85	21:00	- 24.12.85	20:30	
108.	S089A182.022	21.12.85	07:30	- 25.12.85	07:00	
109.	S090A182.022	21.12.85	18:00	- 25.12.85	17:30	
110.	S091A182.022	22.12.85	10:00	- 26.12.85	09:30	
111.	S020X182.022	03.02.86	06:00	- 07.02.86	05:30	*(3.5h)
112.	S092A182.022	14.03.86	12:00	- 18.03.86	11:30	
113.	S021X182.022	15.03.86	07:00	- 19.03.86	06:30	*(12.5h)
114.	S093A182.022	30.03.86	03:00	- 03.04.86	02:30	

115. S094A182.022 30.03.86 12:00 - 03.04.86 11:30
116. S095A182.022 31.03.86 09:00 - 04.04.86 08:30
117. S096A182.022 01.04.86 07:30 - 05.04.86 07:00
118. S097A182.022 11.04.86 11:30 - 15.04.86 11:00
119. S098A182.022 12.04.86 08:30 - 16.04.86 08:00
120. S099A182.022 22.04.86 14:30 - 26.04.86 14:00
121. S100A182.022 25.01.87 07:30 - 29.01.87 07:00
122. S022X182.022 07.02.87 06:30 - 11.02.87 06:00 *(5.0h)
123. S023X182.022 16.02.87 07:00 - 20.02.87 06:30 *(13.5h)
124. S024X182.022 16.02.87 18:30 - 20.02.87 18:00 *(2.0h)
125. S025X182.022 18.02.87 07:30 - 22.02.87 07:00 *(19.5h)
126. S101A182.022 01.03.87 07:30 - 05.03.87 07:00
127. S026X182.022 02.03.87 08:30 - 06.03.87 08:00 *(5.0h)
128. S102A182.022 08.04.87 09:30 - 12.04.87 09:00
129. S103A182.022 05.08.87 12:00 - 09.08.87 11:30
130. S104A182.022 02.09.87 14:00 - 06.09.87 13:30
131. S105A182.022 03.10.87 08:30 - 07.10.87 08:00
132. S106A182.022 20.10.87 10:30 - 24.10.87 10:00
133. S107A182.022 20.10.87 13:00 - 24.10.87 12:30
134. S108A182.022 24.10.87 08:30 - 28.10.87 08:00
135. S109A182.022 29.10.87 11:00 - 02.11.87 10:30
136. S027X182.022 02.12.87 18:30 - 06.12.87 18:00 *(17.5h)
137. S028X182.022 03.12.87 10:00 - 07.12.87 09:30 *(2.0h)
138. S029X182.022 11.12.87 18:00 - 15.12.87 17:30 *(10.5h)
139. S110A182.022 22.12.87 11:30 - 26.12.87 11:00
140. S111A182.022 13.02.88 09:00 - 17.02.88 08:30
141. S112A182.022 14.02.88 01:30 - 18.02.88 01:00
142. S113A182.022 14.02.88 08:30 - 18.02.88 08:00
143. S114A182.022 01.03.88 08:30 - 05.03.88 08:00
144. S115A182.022 03.04.88 09:00 - 07.04.88 08:30
145. S116A182.022 01.10.88 03:30 - 05.10.88 03:00
146. S117A182.022 09.10.88 10:00 - 13.10.88 09:30
147. S118A182.022 19.10.88 14:00 - 23.10.88 13:30
148. S119A182.022 21.10.88 12:30 - 25.10.88 12:00
149. S120A182.022 22.10.88 10:00 - 26.10.88 09:30
150. S121A182.022 30.10.88 12:30 - 03.11.88 12:00
151. S122A182.022 02.11.88 09:30 - 06.11.88 09:00
152. S123A182.022 03.11.88 11:00 - 07.11.88 10:30
153. S124A182.022 05.11.88 10:00 - 09.11.88 09:30
154. S125A182.022 07.11.88 10:30 - 11.11.88 10:00
155. S126A182.022 08.11.88 06:30 - 12.11.88 06:00
156. S127A182.022 21.01.89 12:00 - 25.01.89 11:30
157. S128A182.022 23.01.89 10:00 - 27.01.89 09:30
158. S129A182.022 25.01.89 11:00 - 29.01.89 10:30
159. S130A182.022 25.01.89 18:00 - 29.01.89 17:30
160. S131A182.022 26.01.89 11:00 - 30.01.89 10:30
161. S132A182.022 29.01.89 11:30 - 02.02.89 11:00
162. S030X182.022 30.01.89 10:00 - 03.02.89 09:30 *(13.0h)
163. S133A182.022 01.02.89 15:00 - 05.02.89 14:30
164. S134A182.022 18.02.89 15:00 - 22.02.89 14:30
165. S135A182.022 19.02.89 11:30 - 23.02.89 11:00
166. S031X182.022 20.02.89 08:00 - 24.02.89 07:30 *(3.5h)
167. S136A182.022 22.02.89 08:30 - 26.02.89 08:00
168. S137A182.022 24.02.89 12:00 - 28.02.89 11:30
169. S032X182.022 04.03.89 08:00 - 08.03.89 07:30 *(3.5h)
170. S138A182.022 18.03.89 08:00 - 22.03.89 07:30
171. S139A182.022 18.03.89 12:30 - 22.03.89 12:00
172. S140A182.022 25.03.89 08:00 - 29.03.89 07:30
173. S141A182.022 26.03.89 09:00 - 30.03.89 08:30
174. S142A182.022 10.04.89 08:30 - 14.04.89 08:00
175. S143A182.022 23.10.89 13:30 - 27.10.89 13:00
176. S033X182.022 04.02.90 09:00 - 08.02.90 08:30 *(10.0h)
177. S144A182.022 14.01.91 10:30 - 18.01.91 10:00
178. S034X182.022 15.01.91 10:00 - 19.01.91 09:30 *(9.0h)
179. S145A182.022 20.01.91 08:30 - 24.01.91 08:00
180. S146A182.022 21.01.91 15:30 - 25.01.91 15:00

```
181. S147A182.022  31.01.91 11:30 - 04.02.91 11:00
182. S148A182.022  06.02.91 10:00 - 10.02.91 09:30
183. S149A182.022  07.02.91 11:30 - 11.02.91 11:00
184. S150A182.022  08.02.91 20:00 - 12.02.91 19:30
185. S151A182.022  30.11.91 09:00 - 04.12.91 08:30
186. S152A182.022  30.11.91 12:00 - 04.12.91 11:30
187. S153A182.022  01.12.91 11:30 - 05.12.91 11:00
188. S154A182.022  08.12.91 09:30 - 12.12.91 09:00
189. S155A182.022  08.12.91 18:00 - 12.12.91 17:30
190. S156A182.022  15.12.91 10:00 - 19.12.91 09:30
191. S157A182.022  26.01.92 12:30 - 30.01.92 12:00

Zahl der *-Markierungen: 34   = 17.8%,
mittlere Zeitspanne bis zur Überschreitung: 8.7 h
```

Anhang G

Abbildungsverzeichnis

Seite

Abb. 1: Typischer Jahresgang der SO_2-Konzentration 11

Abb. 1-1: Die Lage der Meßstation in der Beethovenstraße im Zentrum Leipzigs 16

Abb. 1-2: Die automatische Probenahme am Fenster des Hauses Beethovenstraße 25 17

Abb. 1-3: Luftbelastung in der Stadt Leipzig 1990/91 18

Abb. 1-4: Mittlere jährliche Verteilung der Windrichtungen in der Stadt Leipzig von 1961 bis 1980 19

Abb. 1-5: Meßreihe CITY für Leipzig-Mitte während einer erhöhten Belastungssituation 20

Abb. 1-6: Typische Tagesverläufe der SO_2-Konzentration 21

Abb. 1-7: Isolinien der SO_2-Konzentration zur Darstellung des täglichen Ganges und des Ganges über 10 Tage 22

Abb. 1-8: CITY-Meßreihe 24

Abb. 1-9: Logarithmus der Standardabweichungen s und Logarithmus der Mittelwerte m aus Abschnitten von jeweils 96 Meßwerten der Meßreihe von Leipzig-Mitte 1989 und deren Regressionsgerade 25

Abb. 1-10: Transformierte CITY-Daten der Abb. 1-8 26

Abb. 1-11: SO_2-Konzentrationsreihe CITY für Leipzig-Mitte 27

Abb. 1-12: Periodogramm der CITY-Daten von Abb. 1-11 27

Abb. 1-13: Kumulative Darstellung des Periodogramms aus Abb. 1-12 28

Abb. 1-14: Periodogramm der transformierten und trendbereinigten CITY-Reihe 29

Abb. 1-15: Periodogramm aus Abb. 1-14 (dünne Linie) mit einem Tukey-Hamming(5)-Fenster geglättet (dicke Linie) 31

Abb. 1-16: Durch Differenzbildung über 48 Werte saisonbereinigte CITY-Daten 31

Abb. 1-17: Periodogramm (dünne Linie) der Daten von Abb. 1-16 und mit Tukey-Hamming(5)-Fenster geglättetes Periodogramm (dicke Linie) 32

Abb. 1-18: Periodogramm (dünne Linie) und kumulatives Periodogramm (dicke Linie) 33

Abb. 1-19: Verschiedene Abschnitte der SO_2-Immissionsreihe von Leipzig-Mitte 1989 mit dazugehörigem Periodogramm 35

Abb. 1-20: Höhenlinien des Spektrogramms für die Daten von Leipzig-Mitte für das gesamte Jahr 1989; die einzelnen Periodogramme wurden jeweils aus Meßreihenabschnitten von 10 Tagen nach Box-Cox-Transformation und Trendelimination berechnet und anschließend mit einem Tukey-Hamming(5)-Fenster geglättet 36

Abb. 1-21: Höhenlinien des Spektrogramms für die Daten von Leipzig-Mitte 1989; die einzelnen Periodogramme wurden jeweils aus Meßreihenabschnitten von 10 Tagen berechnet (keine Transformation) und anschließend mit einem Tukey-Hamming(5)-Fenster geglättet 38

Abb. 1-22: Schematisches Spektrum der Windgeschwindigkeit in Bodennähe 39

Abb. 1-23: Korrelogramm der CITY Datenreihe von Abb. 1-11 41

Abb. 1-24: Partielle Autokorrelationsfunktion der CITY Datenreihe von Abb. 1-11 42

Abb. 2-1: Die ersten beiden Tage der CITY-Reihe 48

Abb. 2-2: Daten aus Abb. 2-1 nach Box-Cox-Transformation und Elimination des Mittelwertes 49

Abb. 2-3: Autokorrelationsfunktion der varianzstabilisierten Daten von Abb. 2-2 49

Abb. 2-4: Partielle Autokorrelationsfunktion der varianzstabilisierten Daten von Abb. 2-2 50

Abb. 2-5: Prognosen mit einem angepaßten AR(2)-Modell, die zu zwei verschiedenen Zeiten gestartet wurden 51

Abb. 2-6: Fortlaufend neu erstellte Prognose 51

Abb. 2-7: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1985 52

Abb. 2-8: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1989 52

Abb. 2-9: Lage der an alle Tagesgänge angepaßten AR(2)-Modelle in der ϕ_1-ϕ_2-Ebene. Daten: SO_2-Halbstundenwerte von Leipzig-Mitte, 1993 54

Abb. 2-10: Typischer Verlauf der AR(2)-Reihen an verschiedenen Punkten der ϕ_1-ϕ_2-Ebene 54

Abb. 2-11: Darstellung der aus der CITY-Reihe mit dem Kalman-Filter extrahierten Komponenten 57

Abb. 3-1: Verlauf der transformierten CITY-Reihe mit Fehlwerten an den Tagen 6 und 7, die mit IRWSMOOTH bei verschiedenen NVR interpoliert wurden 66

Abb. 3-2: Verlauf der transformierten CITY-Reihe mit Fehlwerten an den Tagen 6 und 7, die mit IRWSMOOTH bzw. DIRWSMOOTH interpoliert wurden 66

Abb. 3-3: Mit IRWSMOOTH extrahierte Trendfunktion für die Meßstelle Leipzig-Mitte vom 1.3.-20.3.89 (transformiert) 68

Abb. 3-4: Häufigkeitsverteilung H der Abstände zwischen den Umkehrpunkten der Trendfunktion von der Meßreihe Leipzig-Mitte 1989 69

Abb. 3-5: Phasenraumbild für die Trendfunktion und deren Ableitung zu den Daten von Leipzig-Mitte 1989 70

Abb. 3-6: Phasenraumtrajektorie (unten rechts) der Trendfunktion der CITY-Reihe 71

Abb. 3-7: Phasenraumtrajektorie (unten rechts) der Trendfunktion der CITY-Reihe mit Fehlwerten an den Tagen 6 und 7 72

Abb. 3-8: Phasenraumbild für eine Extrapolation der Trendfunktion nach dem Hyperzykloidenmodell 79

Abb. 3-9: Vergleich zwischen linearer "L" und lokaler harmonischer "H" Trendextrapolation am Beispiel der Tage 4, 5, 6 und 7 der CITY-Reihe 80

Abb. 4-1: Die 34 Überschreitungsfälle an der Meßstelle Leipzig-Mitte in der Zeit von 1980 bis 1993 (unten) und deren gleitende 24- und 3-Stunden-Mittel (mitte und oben) 83

Abb. 4-2: Die 157 Pseudofälle an der Meßstelle Leipzig-Mitte in der Zeit von 1980 bis 1993 (unten) und deren gleitende 24- und 3-Stunden-Mittel (mitte und oben) 84

Abb. 4-3: Aufteilung der Mengen $\mathbf{P}$ und $\mathbf{U}$ durch zwei Klassen eines Prediktors 85

Abb. 4-4: Aufteilung der Mengen $\mathbf{P}$ und $\mathbf{U}$ durch zwei Klassen eines idealen Prediktors 85

Abb. 4-5: Schematische Veranschaulichung der Aufteilung der Mengen $\mathbf{P}$ und $\mathbf{U}$ durch zwei Prediktoren zu je vier Klassen 86

Abb. 4-6: Mittlere Realisierungen zu den in den Abbn. 4-1 und 4-2 gezeigten Fällen 87

Abb. 4-7: Mittlere Realisierungen zu den M24 der in den Abbn. 4-1 und 4-2 gezeigten Fälle 88

Abb. 4-8: Mittlere Realisierungen zu den M3 der in den Abbn. 4-1 und 4-2 gezeigten Fälle 89

Abb. 4-9: Die verwendeten Prediktoren M24 und M3A3 anhand eines Beispiels für einen Überschreitungsfall 89

Abb. 4-10: Prediktorenebene für die Prediktoren M24(t_o) und M3A3(t_o) 90

Abb. 4-11: Der prognostizierte Prediktor Δt im Falle eine Beispielsituation 94

Abb. 4-12: Histogramm für den prognostizierten Prediktor Δt aller Pseudofälle 95

Abb. 4-13: Histogramm für den prognostizierten Prediktor Δt aller Überschreitungsfälle 95

Abb. 4-14: Treffsicherheit (TS in %) von 23 verschiedenen Arten von Niederschlagsvorhersagen, die sich in der Vorhersagedistanz x und der Wahrscheinlichkeit des Ereignisses p(E) unterscheiden 97

Abb. 4-15: Histogramm für den prognostizierten Prediktor $\Delta t(h)$ aller Pseudofälle 98

Abb. 4-16: Histogramm für den prognostizierten Prediktor $\Delta t(h)$ aller Überschreitungsfälle 98

Abb. 5-1: Autokorrelationsfunktion der CITY-Reihe (Punkte) und ACF der Trendkomponente der CITY-Reihe (Kästchen) 102

Abb. 5-2: Vergleich von linearer und lokaler harmonischer Trendextrapolation anhand ausgewählter Beispielsituationen 103

Abb. C-1: Transferfunktion einer symmetrischen gewichteten gleitenden Mittelung 136

Abb. C-2: Transferfunktion einer symmetrischen ungewichteten gleitenden Mittelung über 48 Meßwerte 137

Abb. D-1: Transferfunktion des IRWSMOOTH-Glättungsfilters 140

Tabellenverzeichnis

Seite

Tab. 1-1: Häufigkeit der Überschreitung der Schwelle 0,6 mg/m³ durch das gleitende 3-Stunden-Mittel der Meßdaten von Leipzig-Mitte 17

Tab. 1-2: Transformationsparameter θ am Beispiel verschiedener SO_2-Jahresreihen der Station Leipzig-Mitte 26

Tab. 3-1: Fortlaufend aus der Trendkomponente geschätzte Frequenz Ω der lokalen harmonischen Approximation für die Tage 4, 5 und 6 der CITY-Reihe 77

Tab. 4-1: Grenzbedingungen, Wahrscheinlichkeiten und a priori Treffsicherheiten für die Prognose mit den Prediktoren $M24(t_0)$ [in mg/m³] und M3A3 [in mg/m³/3h] 92

Tab. B-1: Übersicht über die Eigenschaften von autoregressiven Prozessen AR(p), Prozessen der gleitenden Mittel (MA) und gemischten ARMA-Prozessen 134

Symbolverzeichnis

a(t)	Parameter im Hyperzykloidenmodell
α_t	unabhängige, identisch verteilte Zufallszahlen (weißes Rauschen)
B	Bestimmtheitsmaß
B	Backshiftoperator
c	Konzentration;
c	Vertrauensbereich im kumulativen Periodogramm
c(t)	Parameter im Hyperzykloidenmodell
∇	Differenzenoperator
$\Theta_q(B)$	Polynom q-ten Grades
θ	Box-Cox-Transformationsparameter
η_t	weißes Rauschen im Zustandsraummodell
ε_t	weißes Rauschen
F(t)	Matrix im Zustandsraummodell
F(.)	varianzstabilisierende Box-Cox-Transformation
f	Frequenz
$\Phi_p(B)$	Polynom p-ten Grades
ϕ_1, ϕ_2	AR - Parameter
G(t)	Matrix im Zustandsraummodell
H(t)	Matrix im Zustandsraummodell
I	relative Intensität einer Frequenz im Spektrum
k	Vektor eines Klassenbereichs in der Prediktorenebene
M3A3	Anstieg des gleitenden Mittels über 3 Stunden, Prediktor
M24(t_o)	gleitendes Mittel über 24 Stunden am Prognosepunkt, Prediktor
m, m_t	Mittelwert, Mittelwertfunktion
μ_t	Erwartungswertfunktion
μ_t	weißes Rauschen beim Zustandsraummodell
N	Anzahl der Werte einer Meßreihe
N[P|**k**]	Anzahl der Pseudofälle in **k**
N[Ü|**k**]	Anzahl der Überschreitungsfälle in **k**
NVR	noise variance ratio beim Kalman Filter
ω	Kreisfrequenz (= $2\pi f$)
$\Omega(t)$	Winkelgeschwindigkeit im Hyperzykloidenmodell

p	Ordnung des autoregressiven Modells AR(p)
P	kumulierte Intensität im Spektrum
$P(\mathbf{k})$	Wahrscheinlichkeit, daß sich der Prediktorenvektor $\boldsymbol{\pi}$ in $\mathbf{k}$ befindet
$\mathbf{P}$	Menge aller Pseudofälle einer Station
$\boldsymbol{\pi}$	Prediktorenvektor
q	Ordnung des Modells der gleitenden Mittel MA(q)
r_k	Autokorrelationsfunktion ACF, Autokorrelationskoeffizienten
R_t	Residualkomponente
s, s_t	Standardabweichung
span	Fensterbreite, Glättungsspanne, Anzahl der in eine Glättung einbezogenen Werte
$S1_t$, $S2_t$	Saisonale Komponenten
σ_t	Varianz
σ_α^2	Varianz der Reihe α_t,
t	Zeit (Bei den Meßreihen wird die Zeit nicht kontinuierlich sondern in äquidistanten Schritten (t= ..., -1, 0, 1, 2, ...) durchlaufen.)
$TS(\mathbf{k})$	a priori Treffsicherheit
$t_o(t)$	Parameter im Hyperzykloidenmodell
t_o	Prognosepunkt
T_t	Trendkomponente
T_t'	Ableitung des Trends
Δt	Zeitspanne vom Prognosepunkt bis zur Überschreitung, Prediktor
$\Delta t(h)$	mit lokaler harmonischer Trendextrapolation vorhergesagte Zeitspanne vom Prognosepunkt bis zur Überschreitung, Prediktor
u_t	transformierte Meßwerte
$\mathbf{U}$	Menge aller Überschreitungsfälle einer Station
$\mathbf{x}_t$	Zustandsvektor
ξ_o, ξ_1, ξ_2	Regressionskoeffizienten
z_t, $\{z_t\}$	Meßwerte, Reihe von Meßwerten
$^*z_t(l)$	ab dem Zeitpunkt t für l Schritte im voraus prognostizierter Wert

Zusammenfassung der wissenschaftlichen Ergebnisse zur Dissertation

"Immissionsprognose, Univariate Modellierung und Kürzestfristvorhersage von Wintersmogsituationen"

der Fakultät für Physik und Geowissenschaften
der Universität Leipzig

eingereicht von
Dipl.-Phys. Uwe Schlink, geb. 18.03.59 in Quedlinburg

angefertigt am
Umweltforschungszentrum Leipzig-Halle GmbH
Sektion Expositionsforschung und Epidemiologie

November 1994

Die Auslösung von Smogalarm und das Meiden hoher Luftschadstoffbelastungen bei Smog setzen eine Vorschau auf die zu erwartende meteorologische Situation voraus. Das Ziel der vorliegenden Arbeit ist deshalb die Entwicklung eines Prognosealgorithmus.

Die umwelthygienische Beurteilung der Immissionssituation erfolgt anhand der Wahrscheinlichkeitsverteilung der Meßwerte (vgl. Larsen (1969), Schlink (1990), (1993)).

Bei der Wintersmogvorhersage für die ersten Stunden einer Episode kann man davon ausgehen, daß im Konzentrationsverlauf der Leitkomponente SO_2 alle meteorologischen und Emissionseinflüsse integriert sind. Somit ist eine univariate Vorhersagemethode ausreichend.

Dazu wurde bisher (vgl. Herbarth (1982), Wiegand et al. (1987)) die Autokorreliertheit der Meßreihe ausgenutzt. Dies führte auf die Anwendung der Modelle von Box & Jenkins (1970).

Die hier dokumentierten eigenen Untersuchungen zeigen aber, daß bei den vorliegenden Daten (Halbstundenmittel) diese Modelle keine zufriedenstellende Prognose liefern.

Deshalb wurde ein neuer Ansatz gewählt. Er bezieht auch die Periodizitäten der Meßreihe ein und wird mit dem rekursiven Kalman-Filter (Young et al. (1991)) realisiert.

Die vorliegende Arbeit entwickelt auf dieser Basis das folgende Prognoseverfahren:

Ausgehend von der Spektralanalyse der Meßreihe $\{z_t\}$ wird ein Zeitreihenmodell aus Komponenten unterschiedlicher Zeitscales aufgestellt:

$$z_t = F^{-1} (T_t + S1_t + S2_t + R_t) \qquad (1)$$

Dabei stellt F eine nichtlineare Funktion dar, mit der $\{z_t\}$ transformiert wird, um eine zeitlich stabile Varianz zu erreichen. $\{S1_t\}$ und $\{S2_t\}$ geben die 24- und die 12-Stunden Rhythmik wieder. $\{R_t\}$ bezeichnet einen irregulären Restterm. Die Komponente $\{T_t\}$ beschreibt die niederfrequenten Vorgänge und wird als Trendreihe bezeichnet.

Für die Vorhersage ist von Bedeutung, daß die Smogepisoden speziell durch diese Trendkomponente $\{T_t\}$ widergespiegelt werden. Sie besitzt ein zeitlich veränderliches Frequenzspektrum und wird hier detailliert untersucht.

Zu diesem Zweck wird der Trendverlauf im Phasenraum dargestellt. Die dabei entstehende charakteristische Trajektorie in Form einer Hyperzykloide erlaubt es, eine neue Extrapolationsmethode abzuleiten. Diese approximiert lokal den Trendverlauf durch eine harmonische Funktion:

$$T_t = c - a \cos(\Omega t - t_0) \qquad (2)$$

Deren Parameter c(t), a(t), $\Omega(t)$, $t_0(t)$ sind schwach zeitabhängig und bringen die zeitliche Veränderlichkeit des Spektrums zum Ausdruck. Eine Approximation, bei der diese Parameter fortlaufend neu geschätzt werden, wird diesem zeitabhängigen Spektrum der Trendkomponente gerecht.

Die harmonische Funktion (2) kann auch für die Trendextrapolation genutzt werden. Dabei ergibt sich gegenüber der linearen Extrapolation eine Verbesserung. Dies zeigen Vergleiche mit dem linearen Interpolations- und Extrapolationsverhalten des Kalman-Filters, mit dem die Trendkomponente aus der Meßreihe berechnet wird.

Ziel der Immissionsprognose ist es, für einen gemessenen Konzentrationsverlauf eine Wahrscheinlichkeit anzugeben, mit der es künftig zu einer Schwellwertüberschreitung kommt. Die Datenanalyse zeigt, daß sich dafür drei charakteristische Größen des Konzentrationsverlaufs eignen, die Prediktoren. In die Berechnung der Prediktoren geht die oben genannte Zeitreihenprognose mit dem Komponentenmodell (1) ein.

Um die Güte der entwickelten Verfahren einzuschätzen, werden sie mit gemessenen Konzentrationsreihen getestet. Die Trefferrate einer Prognose hängt von der Häufigkeit des vorhergesagten Ereignisses und vom Vorhersagezeitraum ab. Wintersmogepisoden sind Ereignisse, die etwa 1 bis 5 mal im Jahr auftreten und ca. 2 bis 7 Tage andauern. Die Trefferrate der Prognose liegt bei 26 %. Literaturvergleiche (Balzer (1989)) zeigen, daß bei der Vorhersage von Niederschlagsereignissen vergleichbarer Häufigkeit die gleiche Trefferrate erreicht wird. Prognostiziert man dagegen die viel häufiger eintretende Situation, daß der angegebene Schwellwert nicht überschritten wird, so läßt sich eine Trefferrate von 86 % erreichen.

Der entwickelte Prognosealgorithmus erfüllt die gestellte Aufgabe, indem er Eigenschaften der SO_2-Konzentrationsreihe nutzt, die bisher für Prognosezwecke unzugänglich waren. Da über die Meßdaten stets auch eine Anpassung an den Meßort gegeben ist, ist der Algorithmus für die Vorhersage von Wintersmogereignissen unter den klimatischen Bedingungen des mitteleuropäischen Bereiches mit seinen Zyklonen universell einsetzbar.

Die kritische Diskussion der Extrapolationsalgorithmen zeigt, daß mit den hier entwickelten Verfahren die Grenzen der univariaten Prognose ausgeschöpft sind. Eine weitere Verbesserung der Prognosegüte ist nur durch die Einbeziehung anderer atmosphärischer Zustandsgrößen möglich.

Balzer, K (1989): Statistische Vorhersage seltener, extremer meteorologischer Erscheinungen; *Abh. Met. Dienst d. DDR*, Nr. 141, S. 115.

Box, GEP und Jenkins, GM (1970): *Time series analysis - forecasting and control*, Prentice Hall, Englewood Cliffs, New Jersey, Revised Edition 1976.

Herbarth, O (1982): Zeitreihenanalyse und Prognose lufthygienischer Konzentrationsreihen; *Z. gesamte Hyg.* **28**, Nr. 7, S. 508-09.

Larsen, RE (1969): A new mathematical model of air pollutant concentration averaging time and frequency; *Journal Air Pollution Control Association* **19 (1)**, S. 24.

Schlink, U (1990): Zur rechnergestützten Auswertung von Immissionsmeßdaten; *Z. gesamte Hyg.* **36**, Nr. 1, S. 149-157.

Schlink, U (1993): Die Definition der Immissionskenngrößen nach Bundesrecht und DDR-Recht; *Staub - Reinhaltung der Luft* **53**, Nr. 4, S. 139-141.

Wiegand, G; Weigelt, M und Stahl, H (1987): Methoden zur Bewertung der Luftqualität, Teil A: Zeitreihenanalyse von Immissionsmessungen; Teil B: Extremwertuntersuchungen von Schadstoffkonzentrationen; *Bundesminister für Umwelt, Naturschutz und Reaktorsicherheit, Forschungsbericht 1 04 02 440, Luftreinhaltung* .

Young, PC; Ng, CN; Lane, K und Parker, D (1991): Recursive forecasting, smoothing and seasonal adjustment of non-stationary environmental data; *Journal of Forecasting* **10**, S. 57-89.

Sektion Expositionsforschung und Epidemiologie
UFZ-Umweltforschungszentrum Leipzig-Halle GmbH
Permoserstraße 15
04318 Leipzig
Telefon 0341/235-2365
Telefax 0341/235-2288