

Umweltforschungsplan
des Bundesministeriums für Umwelt,
Naturschutz und Reaktorsicherheit

Förderkennzeichen 297 42 848

"Analyse historischer Datenreihen und Entwicklung einer Methode zur quasi-
wetterbereinigten Trendanalyse von bodennahem Ozon"

Teilbericht 2
Validierung von Ozonmessreihen

von
Dr. Wolfgang Enke

I.D.E.A:
Immissionsdaten Erfassung und Analyse
Karin Seewöster
Wandelsweg 89
4594 Gelsenkirchen

IM AUFTRAG
DES UMWELTBUNDESAMTES

Mai 2000

Inhaltsverzeichnis

| | Seite |
|--|------------|
| Einleitung | 3 |
| 1 Visuelle Überprüfung | 4 |
| 2 Räumliche statistische Überprüfung | 18 |
| 3 Zeitreihenprüfung | 73 |
| 3.1 Überlegungen zur Vorgehensweise | 73 |
| 3.1.1 Räumliche Fehlwertersetzung | 75 |
| 3.1.2 Trends und Stufen | 77 |
| 3.1.2.1 Summenkurve I | 77 |
| 3.1.2.2 Summenkurve II | 79 |
| 3.1.2.3 Gleitender t-Wert | 80 |
| 3.2 Strukturbrüche in der Varianz | 82 |
| 3.3 Sensitivität der Varianz | 82 |
| 3.4 Gleitender F-Wert | 84 |
| 3.5 Kombination von Strukturbrüchen | 85 |
| 4 Anwendung auf Ozonzeitreihen | 89 |
| 5 Graphische Präsentation | 95 |
| 6 Subjektive und objektive Bewertung | 96 |
| 7 Automatische Korrektur von Interventionen | 100 |
| Terminologie | 112 |
| Anhang | |

Einleitung.

Schwankungen in den Messreihen der bodennahen Ozonkonzentration und selbstverständlich auch aller anderen Luftschadstoffe haben vielfältige und selbstverständlich auch sehr unterschiedliche Ursachen. Ein Versuch, die Ursachen dieser Schwankungen zu kategorisieren, wird mit der Einteilung in natürliche und technische Ursachen beginnen. Zu den natürlichen Ursachen können alle diejenigen gezählt werden, die die Höhe der Ozonkonzentration an einem bestimmten Ort und zu einem bestimmten Zeitpunkt bewirken oder bewirkt haben. Hierzu zählen dann sowohl biogene als auch anthropogene Einflussfaktoren, die insgesamt eine Variabilität der Ozonkonzentration in Raum (Fläche) und Zeit hervorrufen. Die Aufgabe der Messtechnik besteht darin, diese Variabilität zu erfassen und zu beschreiben. Unter die Rubrik "Natürliche Ursachen" ist auch der möglicherweise unterschiedliche Charakter benachbarter Messstationen einzuordnen, der je nach der Zugehörigkeit einer Messstation in die Kategorie Berg, Land, Stadt oder Verkehr eine räumliche Variabilität zwischen den Messreihen der Ozonkonzentration benachbarter Messstationen hervorruft.

Überlagert und auch verfälscht wird die auf natürlichen Ursachen beruhende Variabilität der Ozonkonzentration durch zusätzliche Schwankungen, die ihre Ursachen in der Messtechnik, wozu auch die Probenahme zählt, haben. Diese zusätzliche Variabilität wirkt sich überwiegend im zeitlichen Bereich aus, d.h., messtechnische Änderungen bewirken in der Mehrzahl der Fälle Schwankungen in der Messreihe einer einzelnen Messstation. Lediglich in den eher seltenen Fällen, bei denen Änderungen in der Probenahme und Messtechnik einheitlich in einem gesamten Messnetz (z.B. in Deutschland in einem Bundesland) oder - noch seltener - in allen Messnetzen Deutschlands vorgenommen werden, wirken sich diese „Schwankungen auch im räumlichen Bereich aus. Prinzipiell ist es möglich, die gesamte Variabilität in den Messreihen in die Teile wenigstens annähernd aufzuteilen, deren Ursache in natürlichen und technischen Bedingungen liegt. Das Verfahren hierfür und auch der Umfang (schließlich muss eine solche Analyse für jede einzelne Messstation vorgenommen werden) ist sehr aufwendig (siehe unten), und wird in Ansätzen in Stufe II und III

dargestellt. Diese Vorgehensweise ist insofern zwingend notwendig, da gehäuft auftretende unplausible Messwerte die Aussage, die sich allein auf diejenigen Messwerte, die mit natürlichen Schwankungen behaftet sind, beschränken soll, verfälschen werden.

Die Vernachlässigung dieser Vorstufe bei der Analyse der Messdaten muss zu Fehlinterpretationen führen, wie es sich in aller Deutlichkeit bei der 'Interpretation der Zeitreihen der Ozonkonzentration von etwa 1955 bis 1990 in der ehemaligen DDR zeigt, bei denen die Aussage "zunehmender Trend" ausschließlich auf stufenförmige Interventionen, also auf messtechnisch ' verursachte Schwankungen zurückgeführt werden kann (siehe Zeitreihe " Arkona" weiter unten). Mithin wird in diesem Vorhaben ein großes Gewicht auf die Validierung der verfügbaren Ozonreihen gelegt - ein Schritt, der eigentlich als Bestandteil einer jeden Analyse von Immissionsdaten der Analyse selbst vorgeschaltet sein sollte. Diese Validierung besteht aus 3 Stufen:

- Visuelle Überprüfung .
- Räumlich statistische Überprüfung
- Zeitreihen-Überprüfung

Ein weiterer Abschnitt dieses Berichtes behandelt schließlich das Thema:

- Fehlwertersetzung

Im letzten Abschnitt wird noch kurz auf die Ergebnisse des Validierungsprozesses eingegangen.

1 Visuelle Überprüfung

Der erste Schritt der Validierung der Messdaten besteht in der Visualisierung der Messreihen am Computerbildschirm. Bei den Messreihen, die am Bildschirm dargestellt werden, handelt es sich um die Einstundenwerte eines Jahres jeweils einer Messstation. Wahlweise können auch parallel zu der zu bewertenden Ozonmessreihe Ozondaten weiterer (z.B. benachbarter) Messstationen oder auch andere Messkomponenten der gleichen oder benachbarter Mess-Stationen eingeblendet werden, so dass - eine entsprechende Erfahrung des Betrachters vorausgesetzt - potentiell vorhandene grobe Datenfehler festgestellt und eliminiert werden können. Hiermit erfolgt ein erster, wenn

auch unzureichender räumlicher Vergleich der zu bewertenden Daten mit denjenigen benachbarter Messstationen. Die Konsistenz des zu bewertenden Abschnittes der Zeitreihe (1 Messjahr) kann allerdings erst über einen längeren Zeitraum beurteilt werden; hierfür ist dieser erste Schritt der Plausibilitätsprüfung nicht geeignet. Zudem sei deutlich herausgestellt, dass sich diese erste Validierungsstufe subjektiver Kriterien bedient und somit je nach Bearbeiter zu unterschiedlichen Ergebnissen führt.

Bei der visuellen Plausibilitätsprüfung können grobe Datenfehler erkannt werden. Das Verfahren ist subjektiv.

Der vorliegende Bericht zeigt anhand einiger Beispiele häufig auftretende Fehler in den Zeitreihen der Ozonkonzentration.

Die Ergebnisse der ersten Validierungsstufe können in diesem Bericht nur anhand von Beispielen dargestellt werden. Dass es sich hierbei um symptomatische und auch um besonders gravierende Beispiele unplausibler Messwerte handelt, sei uns vom Messnetzbetreiber, zu dessen Messnetz die genannte Station gehört, verziehen. Es sei angemerkt, dass alle Messnetze, die wir im Rahmen dieser Untersuchung betrachtet haben, gleiche oder ähnliche Unplausibilitäten aufweisen und dass wir diese oder ähnliche messtechnisch bedingten Fehler insbesondere dann für nahezu unumgänglich halten, wenn es sich bei dem zu untersuchenden Schadstoff um ein derart reaktives Gas handelt, wie es das Ozon nun einmal darstellt.

Beispiele offensichtlich nicht plausibler Messdaten lassen sich in den vorhandenen Datensätzen der Ozonkonzentration aus dem Zeitraum 1974 bis 1997 (bei Messstationen in der ehemaligen DDR auch ab etwa 1955) häufig finden. Hierbei handelt es sich in den meisten Fällen um nicht erkannte oder auch um vernachlässigte, durch das Messgerät verursachte Fehler. Oft sind auch Fehler zu finden, die durch oder bei der elektronischen Datenverarbeitung entstanden sind. Dies trifft z. B. auf die Daten der Messstation NÜRNBERG (BY053) zu, bei der die dem Umweltbundesamt übergebenen Datensätze Kalibrierwerte enthalten, wie es das Beispiel aus dem Jahr 1982 in Abb.01 zeigt. Es ist leicht einsichtig, dass solche Kalibrierwerte - wenn sie nicht aus den Kollektiven entfernt werden - Auswertungen dieses Datenkollektives

beeinträchtigen. Es ist jedoch kaum vorstellbar, dass die in der Abbildung dargestellten Fehler bei Auswertungen unbemerkt bleiben (oder etwa doch?).

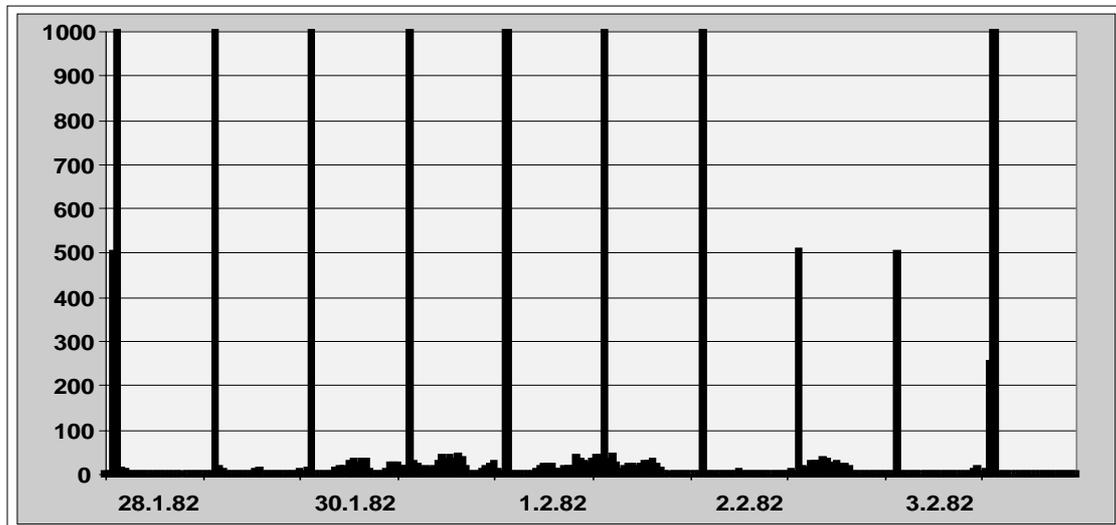
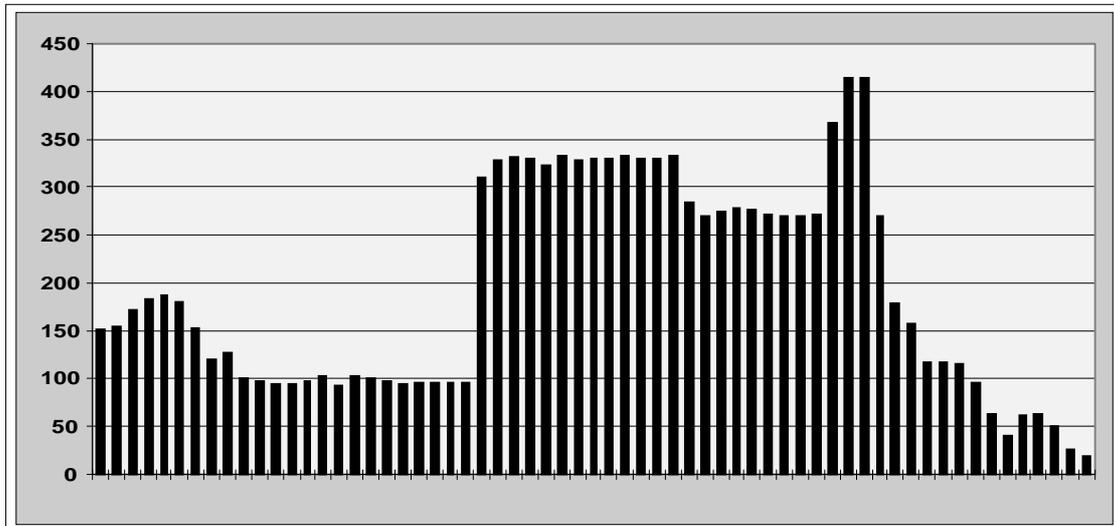


Abb. 01) Das Datenkollektiv der Station Nürnberg (BY053) enthält Kalibrierwerte

Was aber ist von den höchsten Messwerten zu halten, die aus deutschen Messstationen berichtet und in vielen Veröffentlichungen weit verbreitet wurden? Diese Messwerte, die über $500\mu\text{g}/\text{m}^3$ (nicht auf UV-Kalibrierung und 20°C bezogen) betragen, entstammen dem Datenkollektiv einer Messstation in **MANNHEIM** (BW005) aus dem Jahr 1978, aus dem einer der fragliche Zeitabschnitt in Abb.02 dargestellt ist. Es ist zu beachten, dass die hier dargestellten Werte auf UV-Kalibrierung und 20°C bezogen sind und mithin niedriger ausfallen, als die in der Literatur berichteten Messwerte.



**Abb. 02) Nicht plausible Messwerte im Datenkollektiv der Station
MANNHEIM-NORD (BW005)**

Die Abbildung stellt einen Zeitraum von 3 Tagen dar, in denen die bei der Ozonkonzentration nahezu immer vorhandene und typische zyklische Schwankung „Tagesgang“ nur am Anfang und Ende der Reihe andeutungsweise erkennbar ist, während der Zeitraum der „höchsten Konzentrationen“ einen über mehrere Stunden gleichmäßigen und durch Stufen getrennten Verlauf der „Messwerte“ anzeigt. Als Fazit ist festzuhalten, dass diese Werte und damit der höchste in Deutschland registrierte Ozonwert nicht plausibel sind.

Ebenfalls nicht plausibel sind mehrere Messwerte der Station **EGGENSTEIN** (BW004) bei Karlsruhe - immerhin eine Station, die mit häufig hohen, aber durchaus plausiblen Messwerten (Hürth-Effekt) in Erscheinung tritt - aus dem Jahr 1980, die in Abb.03 dargestellt sind. Bei dieser Station treten in der Nacht vom 30.10. auf den 1.11. 5 gleichbleibende Messwerte von $351\mu\text{g}/\text{m}^3$ auf, während kurz vor und auch nach diesem Zeitraum Ausfallwerte verzeichnet sind. Es ist sicher keine Fehlentscheidung, diese Werte als nicht plausibel einzuschätzen und aus dem Datenkollektiv zu entfernen.

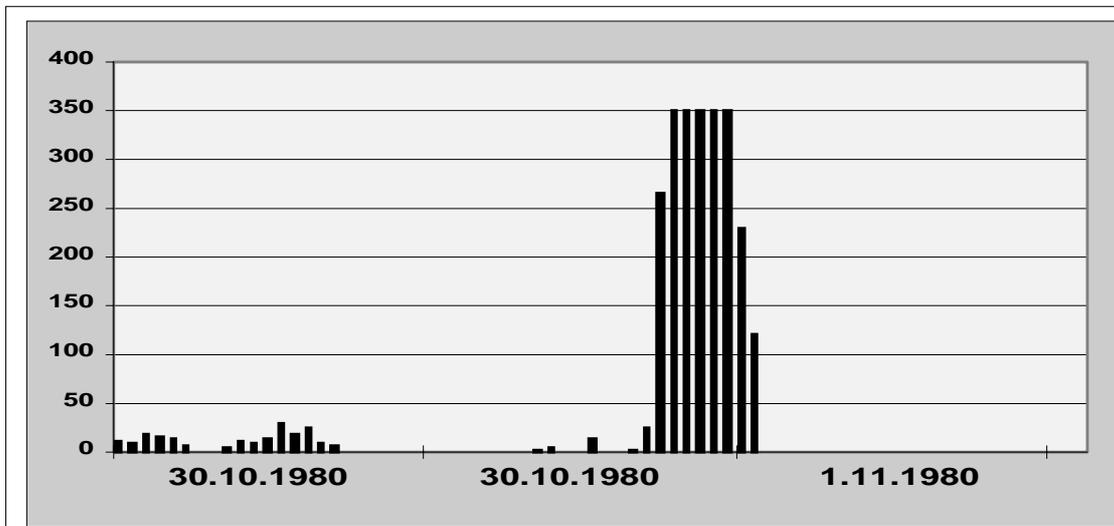


Abb. 03) Nicht plausible Messwerte, eingebettet in Ausfallperioden, Station EGGENSTEIN (BW004)

Die gleiche Entscheidung ist bei der Station **KARLSRUHE-MITTE** (BW001)- ebenfalls im Jahr 1980 - vertretbar. Hier sind gleich 2 Zeiträume mit nicht plausiblen hohen „Messwerten“ zu kritisieren (Abb.04). Der erste der beiden Zeiträume liegt im Juni in den Nachtstunden vom 12.06.1980 auf den 13.06.1980. In diesen Stunden treten „Messwerte“ bis $250\mu\text{g}/\text{m}^3$ auf. Der höchste „Messwert“ an diesem und an den folgenden Tagen beträgt sogar $318\mu\text{g}/\text{m}^3$ - und dann folgen Ausfallwerte. An den Tagen vor diesen obskuren „Messwerten“ erreicht die Ozonkonzentration bei einem üblichen Tagesgang maximal $100\mu\text{g}/\text{m}^3$. Im Wintermonat Oktober ist der nächste unplausible „Messwert“ zu erkennen, der aus einem niedrigen Niveau von unter $50\mu\text{g}/\text{m}^3$ mit einem Einstundenwert von $378\mu\text{g}/\text{m}^3$ herausragt.

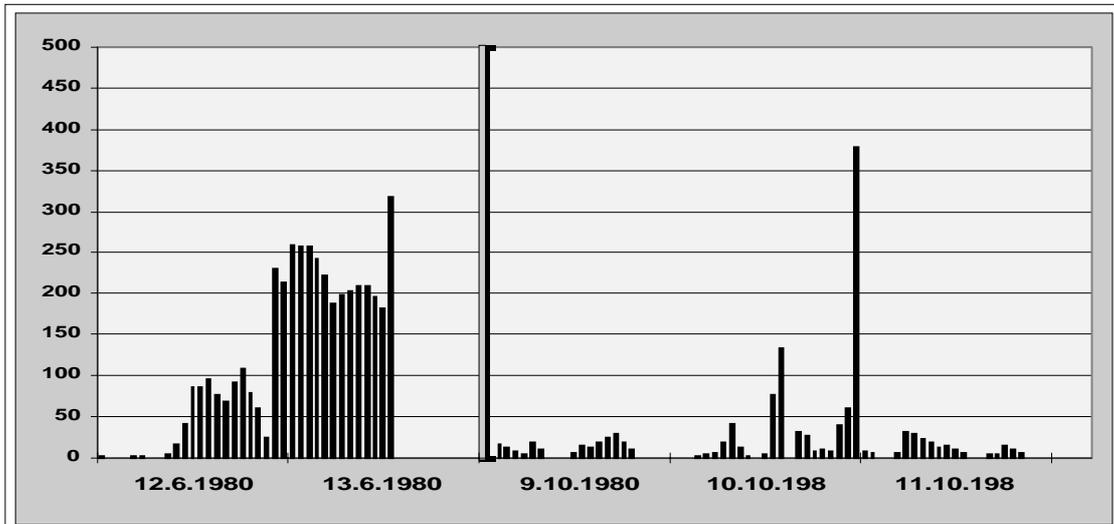
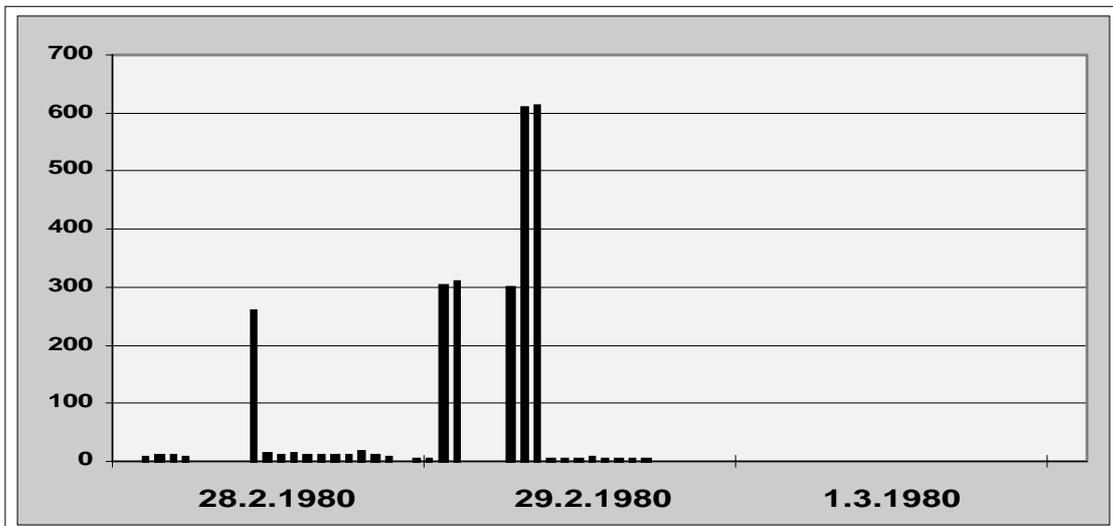


Abb. 04) Wie oben, zusätzlich eine einzelne nicht plausible Spitze im Oktober, KARLSRUHE-MITTE (BW001)

Ebenfalls im Jahr 1980 zeigt auch die Station **HEILBRONN** einige bemerkenswerte „Messwerte“, die in der Abb.05 dargestellt sind. Der Wintermonat Februar zeigte in den Tagen vor dem 28. für diese Jahreszeit üblichen Messwerte von etwa $50\mu\text{g}/\text{m}^3$ bis $70\mu\text{g}/\text{m}^3$. Am 28.02.1980 tritt jedoch ein erster, nicht plausibler Wert von etwa $260\mu\text{g}/\text{m}^3$ auf. Es folgen dann einige Messwerte bis maximal $20\mu\text{g}/\text{m}^3$ und in den Morgen- bzw. Mittagsstunden weitere nicht plausible Werte um $300\mu\text{g}/\text{m}^3$. Die Spitzenwerte der Ozonbelastung finden sich ebenfalls in den Mittagsstunden des 29.02.1980 mit den bemerkenswerten Konzentrationen von $610\mu\text{g}/\text{m}^3$ bzw. $611\mu\text{g}/\text{m}^3$. Hierauf folgen einige Messwerte mit $1\mu\text{g}/\text{m}^3$, bis das Messgerät ein Einsehen hat und nur noch Ausfallwerte meldet.



**Abb. 05) Kalibrierwerte? Nicht plausible Messwerte an der Station
HEILBRONN (BW015)**

Dass bis hierher überwiegend Messwerte aus Stationen des Landes Baden-Württemberg kritisiert wurden, ist einfach darin begründet, dass die Ozonbelastung in diesem Messnetz bereits seit Mitte der 70er Jahre registriert wurde, während die Mehrzahl der Ozon-Messstationen in den übrigen Bundesländern und im Messnetz des Umweltbundesamtes erst zu Beginn der 80er Jahre eingerichtet wurden. Selbstverständlich werden auch aus diesen Messnetzen "Ozonbelastungen" berichtet, die erstaunlich sind. So berichtet die Station **DEUSELBACH** (UB002) des Umweltbundesamtes im Jahr 1980 einige Messwerte, die zwar nicht die Höhe der weiter oben geschilderten „Konzentrationen" erreichen, aber ebenfalls als nicht plausibel verworfen werden müssen (Abb.06). Nach einem mehrere Tage andauernden Ausfall liegen die ersten wieder verfügbaren Messwerte um $230\mu\text{g}/\text{m}^3$ und auch in den nachfolgenden Stunden bis zum 05.09. um die Mittagszeit noch um $150\mu\text{g}/\text{m}^3$. Erst dann liegen die Messwerte nach einem mehrere Stunden anhaltenden Ausfall (Kalibrierung? Reparatur?) auf einem für den Monat September plausiblen Niveau von etwa $50\mu\text{g}/\text{m}^3$ bis $100\mu\text{g}/\text{m}^3$.

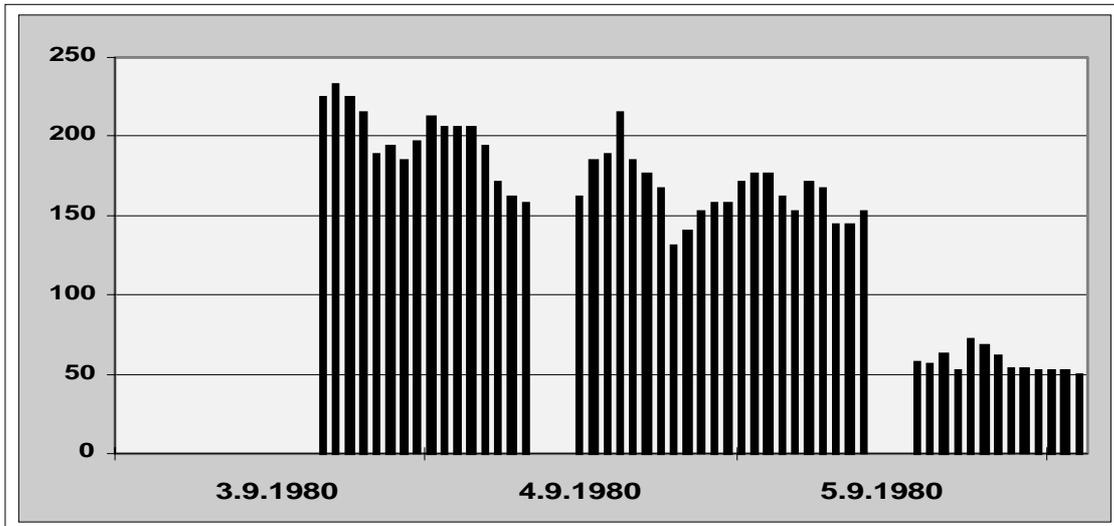


Abb. 06) Systematischer Messfehler, eingebettet in Ausfallperioden an der Station DEUSELBACH (UB002)

Ein weiteres Beispiel für nicht plausible hohe Messwerte soll die Station **WESTEIFEL** aus dem Jahr 1985 liefern. Diese Messstation des Messnetzes in Rheinland Pfalz beobachtete im Juni des Jahres 1985 insgesamt durchaus plausible Messwerte um $100\mu\text{g}/\text{m}^3$. Am 04.06. fiel das Messgerät für einige Stunden aus, meldete aber ab den Mittagsstunden erneut Messwerte, die $250\mu\text{g}/\text{m}^3$ erreichten, um dann für einen längeren Zeitraum erneut auszufallen (Abb.07). Mit hinreichender Sicherheit sind die „Messwerte“ als unplausibel aus dem Kollektiv der Messdaten dieser Station zu entfernen, zumal die benachbarten Messstationen ähnlich hohe Werte nicht aufweisen.

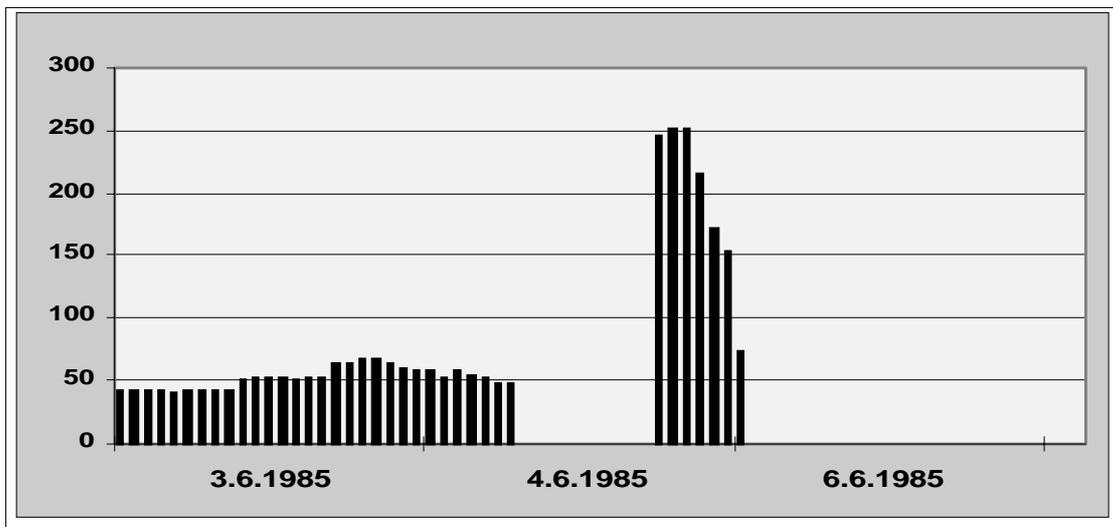


Abb. 07) Nicht plausible Messwerte, eingebettet in Ausfallperioden an der Station WESTEIFEL (RP015)

Ein vergleichbarer Fehler kann im Datenkollektiv der Messstation **LIS-ESSEN** in Nordrhein-Westfalen beobachtet werden (Abb.08), wo im April des Jahres 1982 ebenfalls über einen längeren Zeitraum Ausfallwerte gemeldet wurden. Am 13.04 und 15.04. treten jedoch „Messwerte“ auf, die für den April recht ungewöhnliche „Konzentrationen“ um $200\mu\text{g}/\text{m}^3$ erreichen. Die Gleichförmigkeit dieser „Messwerte“ lässt vermuten, dass es sich hierbei um Kalibrierwerte handelt, die bei dem Versuch, dass Messgerät wieder in Betrieb zu nehmen, hervorgerufen wurden. Auch diese Daten können mit Sicherheit aus dem Kollektiv der Station entfernt werden.

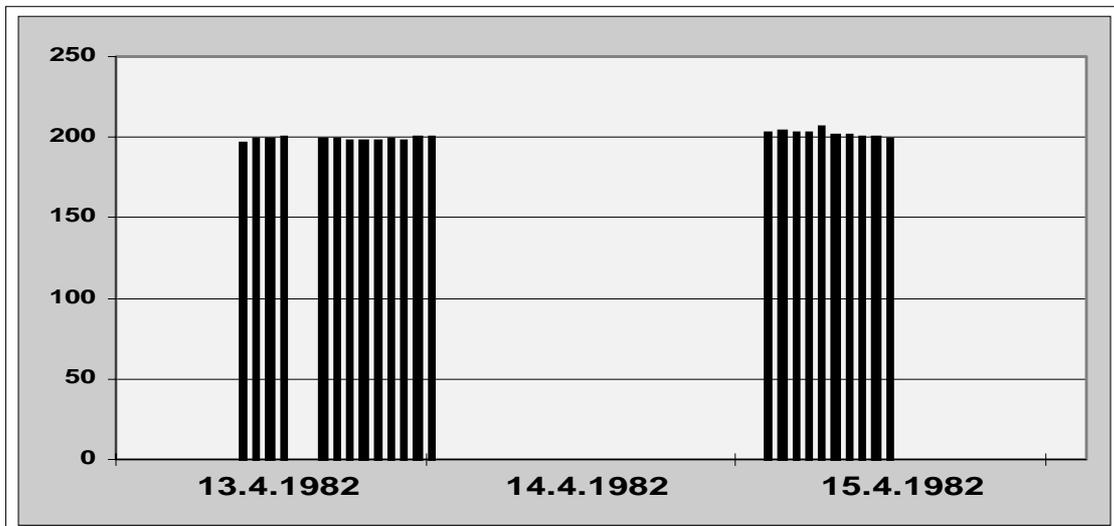


Abb. 08) Kalibrierwerte (?) im Datenkollektiv der Station LIS-ESSEN (NW028)

Es sind jedoch - wie bereits oben gesagt - nicht nur Fehler zu beobachten, die aller Wahrscheinlichkeit nach durch fehlerhafte Messungen hervorgerufen werden. In einigen Fällen, die zugegebenermaßen sehr selten sind, treten unplausible Werte auf, wenn vorhandene Ausfallwerte durch Schätzwerte ersetzt werden. Es ist an dieser Stelle nicht nachvollziehbar, wie die „Messwerte“ in Abb.09 in das Datenkollektiv der Messstation **FICHTELBERG** hineingeraten sind. Der August des Jahres 1983 brachte mit hochsommerlichen Tagen an dieser Station in Sachsen (damals Messnetz des Meteorologischen Dienstes der DDR) auch abenteuerlich erhöhte Ozonbelastungen mit sich. Allerdings trifft die Schuld an diesen „Messwerten“ nicht das Messgerät oder die Bearbeiter im Meteorologischen Dienst, die in diesem Zeitraum Ausfallwerte gemeldet haben. Entstanden sind diese unplausiblen Werte, von denen einer den bemerkenswerten Spitzenwert von nahezu $800\mu\text{g}/\text{m}^3$ erreicht, bei der nachträglichen Bearbeitung der Datenkollektive (nach der Wende). Wie auch immer diese Werte in das Datenkollektiv der Messstation **FICHTELBERG** (SN053) hineingeraten sind - bei einer nachträglichen Kontrolle hätte dieser „Missgriff“ bemerkt werden müssen!

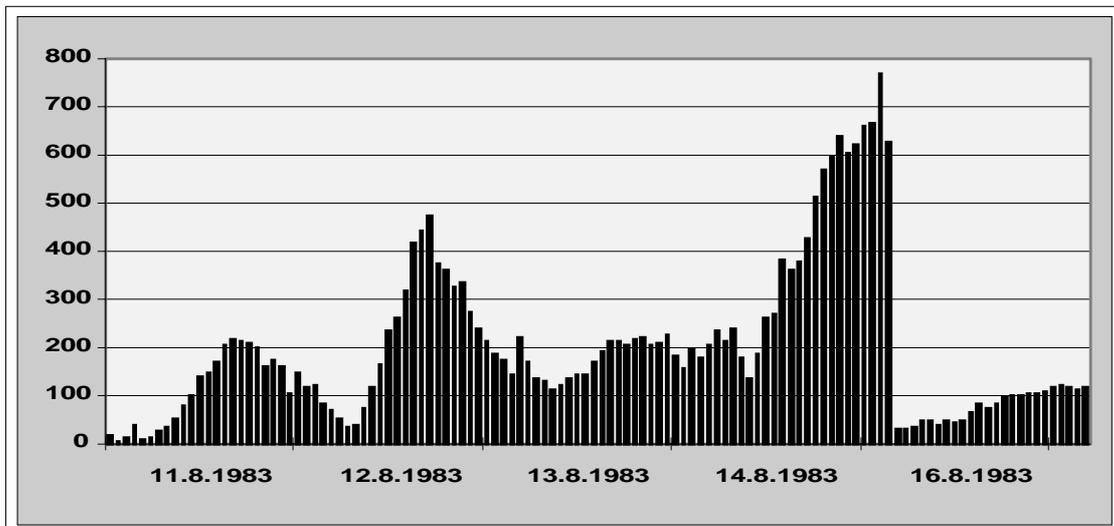
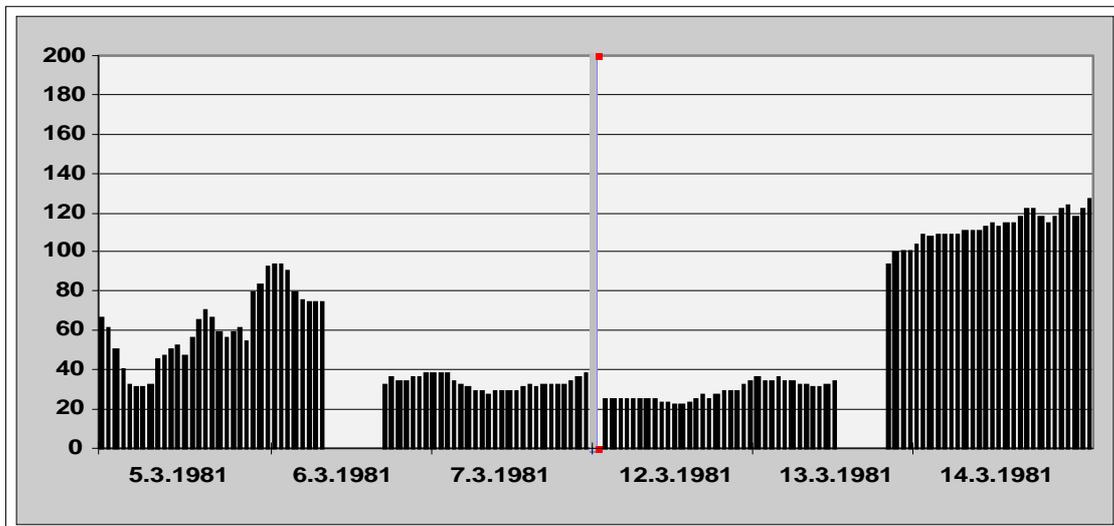


Abb. 09) Fehlerhaft berechnete Schätzwerte im Datenkollektiv der Station FICHTELBERG (SN053, Met. Dienst DDR)

Es sind jedoch nicht nur die hohen Werte, die einer Plausibilitätsprüfung oft nicht standhalten. Auch niedrige Messwerte sind häufig als nicht plausibel zu bewerten, wie dies die Abb.10 am Beispiel der UBA-Station **SCHAUINSLAND** (UB004) aus dem Jahr 1981 zeigt. Im März dieses Jahres werden die für Bergstationen üblichen relativ hohen Werte um $100\mu\text{g}/\text{m}^3$ gemessen. Am 06.03. wird jedoch eine offensichtlich missglückte Kalibrierung durchgeführt, die zur Folge hat, dass die Messwerte auf ein Niveau von $30\mu\text{g}/\text{m}^3$ abgesenkt werden. Dieser Fehler wird erst am 13.03. bei einer erneuten Kalibrierung korrigiert, ohne dass jedoch die fehlerhaften Werte der vergangenen Tage ebenfalls korrigiert worden sind.



**Abb. 10) Systematischer Messfehler im Datenkollektiv der Station
SCHAUINSLAND (UB004)**

**Messtechnisch bedingte Schwankungen sind in längeren Zeitreihen von
Schadstoffkonzentrationen nahezu unvermeidlich**

Die vorangehenden Graphiken zeigen ausgesuchte Beispiele nicht plausibler Messwerte; dass diese Beispiele in dieser eher informativen Zusammenstellung besonders krasse Fälle offensichtlich fehlerhafter Daten zeigen, ist bereits erwähnt worden. Ebenfalls erwähnt wurde bereits weiter oben, dass sich in den Datenkollektiven nahezu aller Messstationen in Deutschland (und sicher auch überall sonst) ähnliche Fehler finden lassen. Es ist weiter oben auch bereits bemerkt worden, dass wir die hier gezeigten Fehler bei der Flut von anfallenden Messdaten und insbesondere bei dem schwer messbaren, weil hoch reaktiven Schadstoff Ozon für nahezu unvermeidlich halten. Zusätzlich sollte beachtet werden, dass die Messmethodik und das Verhalten der Ozonkonzentration bezüglich der zeitlichen Schwankungen in den ersten Messjahren gegen Mitte der 70er Jahre nur wenig bekannt war.

Der Schadstoff Ozon wurde an den Messstationen des ehemaligen meteorologischen Dienstes der DDR bereits seit Mitte der 50er Jahre nach einem nasschemischen Verfahren gemessen. Bekannt wurden diese Messungen durch die in der Literatur vielfach zitierten zunehmenden Trends der Ozonkonzentration an den Stationen

ARKONA auf Rügen (MV013) und **WAHNSDORF** bei Dresden (SN051), die in Abb.11 dargestellt sind. Diese Trends beruhen jedoch auf einem krassen Strukturbruch gegen Mitte bis Ende der 70er Jahre (weitere Strukturbrüche in der Form stufenförmiger Interventionen sind in anderen Jahren erkennbar, sollen hier jedoch nicht diskutiert werden), der in Zusammenarbeit mit den Mitarbeitern des Umweltbundesamtes und Sächsischen Landesamtes für Umwelt und Geologie als Änderung im Messverfahren verifiziert werden konnte.

Diese Änderung bestand darin, dass die Menge der eingesetzten Reagenzien beim angewandten Analysenverfahren (Cauer-Methode, ein chemisches Verfahren) im Zeitraum von 1950 bis Mitte 1970 nicht ausreichend war und höhere Konzentrationen als etwa $120\mu\text{g}/\text{m}^3$ nicht erfasst werden konnten. Erst gegen Mitte der 70er Jahre wurde die Konzentration des Reagenzes zunächst an der Station **WAHNSDORF** (diese Station hatte damals die Funktion einer „Pilotstation“) und einige Zeit später auch an der Station **ARKONA** (sowie an den anderen Stationen, die zu dieser Zeit in der DDR in Betrieb waren) verdoppelt, was sich dann auch in einer Erweiterung des Messbereiches und leider auch in der genannten Stufe bemerkbar machte. Weitere Interventionen, die auf Änderungen des Messverfahrens beruhen, bewirken in den Datenkollektiven dieser Stationen messtechnisch verursachte Schwankungen, die in der Vergangenheit ebenso gerne wie falsch als zunehmender Trend der Ozonkonzentration interpretiert wurden. Trotz der Bemühungen um eine Korrekturmöglichkeit ist es uns bis heute nicht gelungen, dieses an sich wertvolle Datenmaterial soweit zu rekonstruieren, dass es für Auswertungen verwendet werden kann. So bleibt uns zunächst nur die bedauerliche Feststellung, dass die vorliegenden Daten für den Zeitraum Mitte der 50er bis Mitte der 70er Jahre für eine Trendbewertung nicht valide ist.

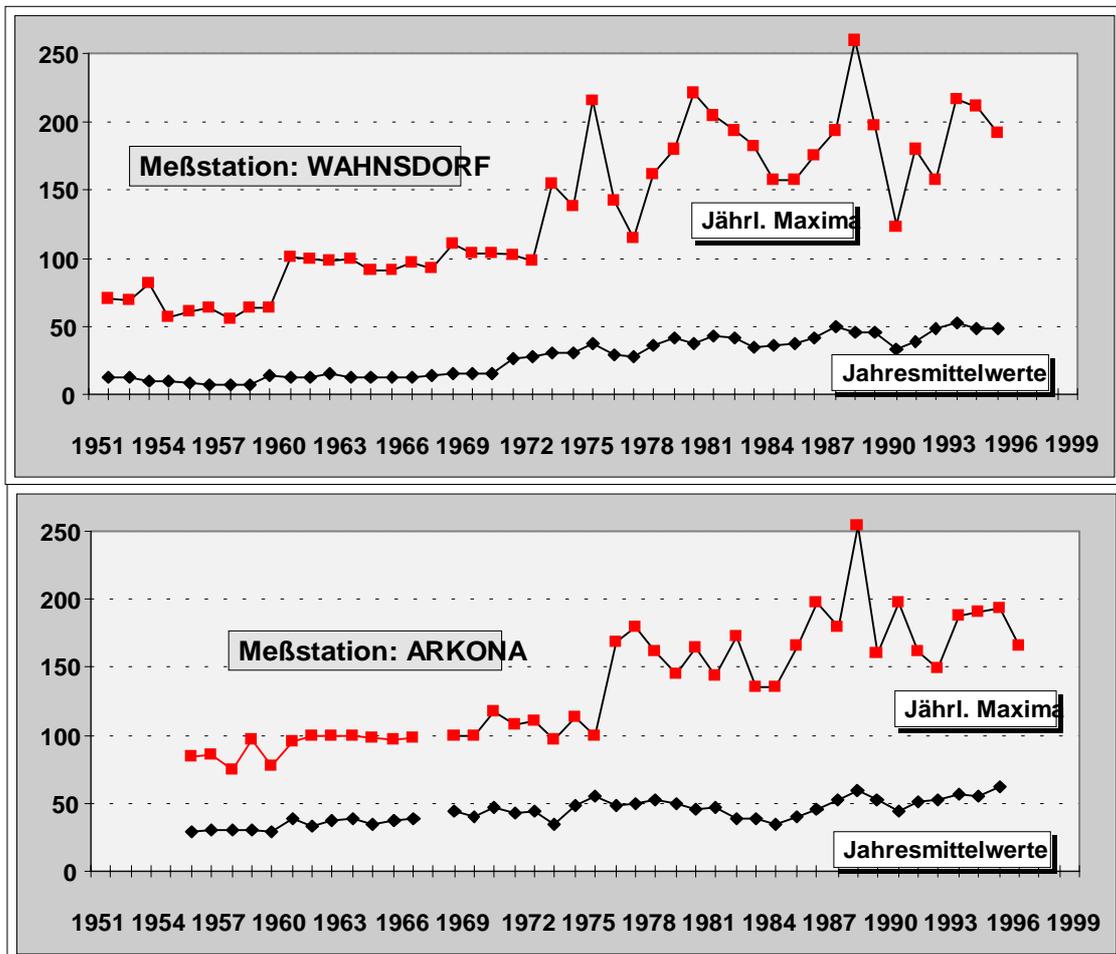


Abb. 11) Änderungen in der Messtechnik bewirken stufenförmige Interventionen in den Messreihen von ARKONA und WAHNSDORF (MV013 und SN051, Met. Dienst der DDR)

Ergebnis: Der “zunehmende Trend” der Ozonkonzentration der Messstellen ARKONA und WAHNSDORF beruht auf messtechnischen Änderungen bei der Analyse (stufenförmige Interventionen)

Die Tabelle 1 im Anhang zu diesem Bericht zeigt die Bewertungen der Ozondatakollektive aus den Jahren 1974 bis 1997. Die Mehrzahl der als nicht plausibel gekennzeichneten Datensätze wurden nach Rücksprache und Diskussion mit den Mitarbeitern des Umweltbundesamtes auf „Ausfall“ gesetzt und in einem späteren Arbeitsschritt mit Schätzwerten aufgefüllt.

2 Räumlich statistische Überprüfung

Im vorhergehenden Abschnitt wird eine visuelle Prüfung der Datenkollektive beschrieben, die am Computerbildschirm vorgenommen wird und wobei im wesentlichen nur die Messergebnisse einer einzelnen Messstation betrachtet werden. Ein Vergleich mit zeitgleichen Messwerten anderer Stationen oder anderer Messobjekte wird nur in Ausnahmefällen vorgenommen - wenn die Entscheidung „plausibel / unplausibel“ zweifelhaft ist. Dieses Verfahren ist subjektiv und lässt lediglich grobe Messfehler erkennen.

Die räumlich-statistische Prüfung auf nicht plausible Messwerte in den Datenkollektiven liefert zunächst objektive Bewertungs-Maßzahlen, die im zweiten Schritt jedoch ebenfalls subjektiv interpretiert werden

Ein objektives Verfahren liefert im ersten Schritt die in diesem Abschnitt beschriebene räumlich statistische Überprüfung der Datenkollektive. Der zweite Schritt, nämlich die Entscheidung, ob die zweifelhaften Messwerte, die das Verfahren ausgibt, tatsächlich als unplausibel zu betrachten und aus dem Datenkollektiv zu entfernen sind, bleibt erneut dem Bearbeiter überlassen und ist mithin ebenfalls subjektiv. Die Methode ist auf das Maximum gleichzeitiger Einstundenwerte in den Datensätzen aller Messstationen begrenzt und kann somit als Prüfung auf Ausreißer angesehen werden.

Voraussetzung bei der räumlich-statistischen Prüfung: Die Ozonbelastung ist ein überregionales Phänomen

Die räumlich statistische Prüfung geht von der Voraussetzung aus, dass die Belastung durch den Schadstoff Ozon eine großräumige - regionale und auch überregionale - Angelegenheit ist und dass sich hohe Messwerte nicht nur an einer Messstation, sondern auch an benachbarten Stationen finden lassen. Allerdings ist diese Annahme bei einem bestimmten Stationstyp nicht zutreffend. Hierbei handelt es sich um Messstationen in der Umgebung petrochemischer Anlagen wie sie zum Beispiel durch die Station **HÜRTH** bei Köln oder **EGGENSTEIN** bei Karlsruhe repräsentiert sind.

Als fragwürdig werden Messwerte bezeichnet, die die 3-fache Standardabweichung aller gleichzeitigen Messwerte überschreiten und um mehr als $100\mu\text{g}/\text{m}^3$ über den 3 höchsten Werten der übrigen Stationen liegen

Die Überprüfung erfolgt auf der Basis der Einstundenwerte einer Messstation, die mit den zeitgleichen Einstundenwerten aller übrigen Stationen mit Ausnahme der Bergstationen verglichen werden. Zunächst wird der Mittelwert und die Standardabweichung (die Annahme einer Normalverteilung der zeitgleichen Messwerte ist gewagt) aus den Messwerten aller Messstationen zum zu bewertenden Zeitpunkt berechnet. Als obere Bewertungsgrenze wird die dreifache Standardabweichung verwendet, die zum Mittelwert addiert wird. Ausgegeben und als potentiell unglaublich wird ein Maximum jedoch erst dann vorgeschlagen, wenn dieses Maximum um $100\mu\text{g}/\text{m}^3$ über dem Mittelwert der drei höchsten Maxima, die an den übrigen Messstationen gemessen wurden, liegt. Wird dieser Wert von $100\mu\text{g}/\text{m}^3$ verringert (etwa auf $75\mu\text{g}/\text{m}^3$ oder gar $50\mu\text{g}/\text{m}^3$), so vervielfacht sich die Zahl der angezweifelte Werte und macht eine Beurteilung der Ergebnisse nahezu unmöglich. Hier treten die häufig vorhandenen Fehler erster und zweiter Art auf, indem die Auswahl einer Schwelle je nach ihrer Höhe entweder zu viele „richtige“ Werte als nicht plausibel ausgibt oder zu viele „falsche“ Werte als plausibel durchgehen lässt.

Nach der Beseitigung der groben Unplausibilitäten im oben beschriebenen ersten Schritt der Validierung erfolgte diese zweite Stufe der Überprüfung ab dem Jahr 1980, da in den Vorjahren diese Überprüfung wegen der zu geringen Anzahl der vorhandenen Messstationen nicht durchgeführt werden kann.

Messjahr 1980

Auffällig ist die Messstation **KARLSRUHE-MITTE** (BW001) im Monat August, wo am 12.08. um die Mittagszeit zwei Messwerte mit $196\mu\text{g}/\text{m}^3$ (12 Uhr) und $197\mu\text{g}/\text{m}^3$ (13 Uhr) gemessen wurde. Die höchsten Maxima der übrigen Messstationen liegen zu diesen Zeitpunkten zwischen etwa $70\mu\text{g}/\text{m}^3$ und $80\mu\text{g}/\text{m}^3$. Diese zwei Werte sind, wie die Darstellung in Abb.12 erkennen lässt, offensichtlich als nicht plausible Ausreißer anzusehen und aus dem Datenkollektiv zu entfernen. Dies ist auch erkennbar an den

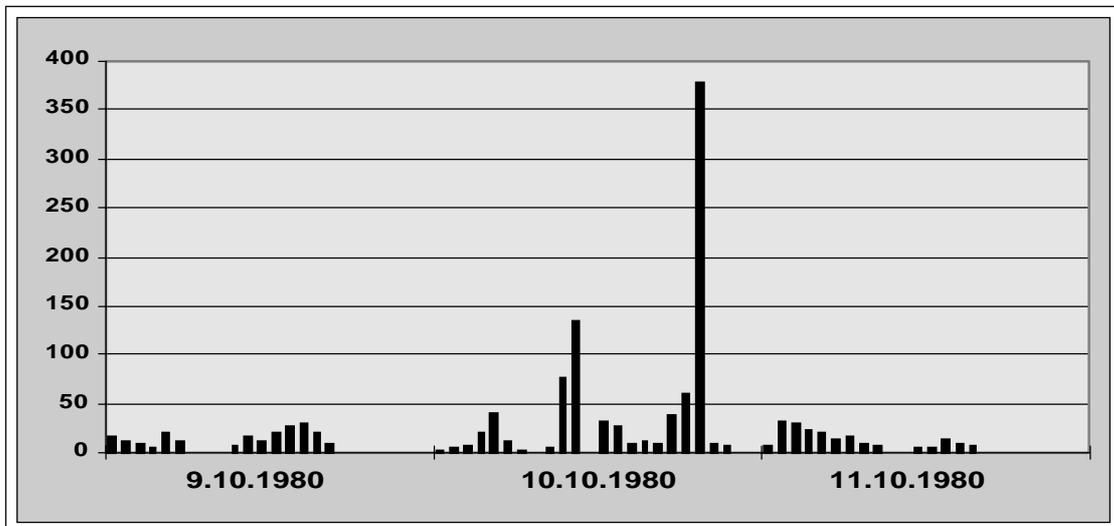


Abb. 13) Ausreißer im Datenkollektiv der Station KARLSRUHE-MITTE (BW001)

Auch die Messstation **MANNHEIM-SÜD** (BW007) ist in diesem Jahr mit einem auffälligen Messwert vertreten. Dieser Wert wird um 8 Uhr morgens am 30.07.1980 mit $169\mu\text{g}/\text{m}^3$ angegeben. Er liegt um etwas mehr als $100\mu\text{g}/\text{m}^3$ über den drei höchsten Maxima der übrigen Messstationen, die zu diesem Zeitpunkt etwas über $60\mu\text{g}/\text{m}^3$ melden. Die Darstellung in Abb.14 zeigt die Messwerte im fraglichen Zeitraum. Auf den vom Programm beanstandeten Messwert von $169\mu\text{g}/\text{m}^3$ folgen zwei weitere, nicht beanstandete Werte von $171\mu\text{g}/\text{m}^3$ und anschließend für einen längeren Zeitraum Ausfallwerte. Am Vortag wurden Messwerte über $200\mu\text{g}/\text{m}^3$ mit einem Spitzenwert von $272\mu\text{g}/\text{m}^3$ gemessen, die ebenfalls nicht beanstandet wurden. Der Wert von $169\mu\text{g}/\text{m}^3$ ist als plausibel zu betrachten.

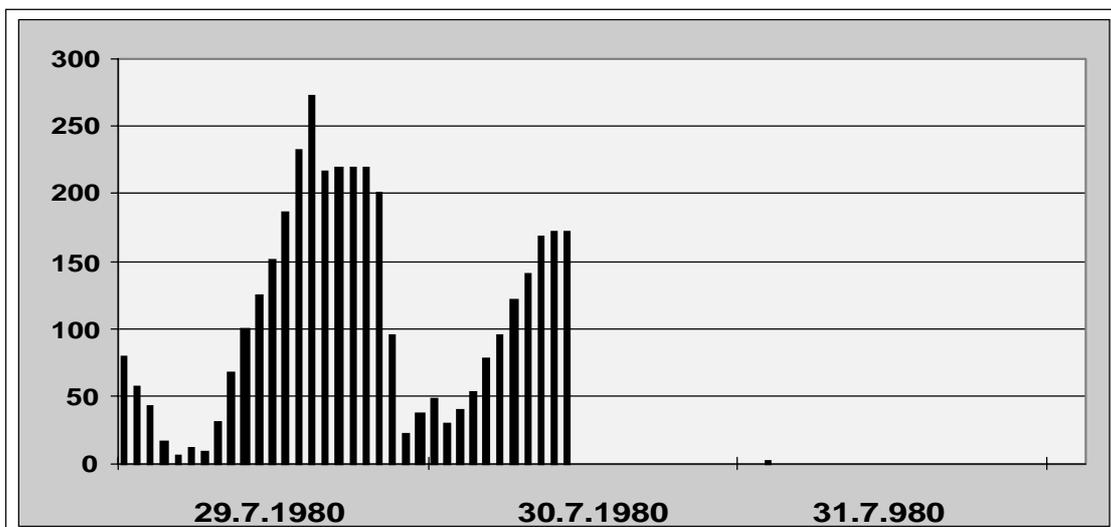


Abb. 14) Das Verfahren bezeichnet auch Werte als fragwürdig, die subjektiv jedoch als plausibel einzustufen sind, MANNHEIM-SÜD (BW007)

Die UBA-Messstation **DEUSELBACH** (UB002) meldet am 24.07 und 25.07. sowie am 05.09. mehrere fragwürdige Messwerte jeweils in den Morgenstunden (ohne Abbildung). Die Messwerte im Juli wurden während einer Episode erhöhter Ozonbelastungen ermittelt und sind plausibel. Die am 05.09. beanstandeten Messwerte wurden bereits - zusammen mit anderen Werten am vorhergehenden und gleichen Tag, die hier jedoch nicht beanstandet werden - als nicht plausibel angesehen und sind bereits aus den Datensätzen eliminiert.

Ohne Abbildung: Das Datenkollektiv der Station DEUSELBACH (UB002) enthält im Juni fragwürdige und im September nicht plausible Messwerte

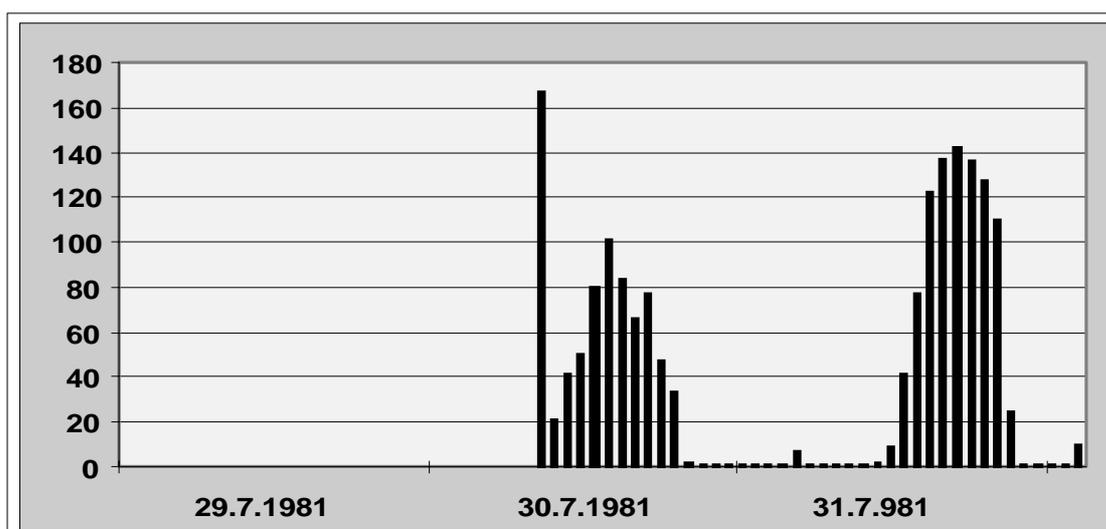
Am 18.02. ist ein Messwert von $502\mu\text{g}/\text{m}^3$ im Datenkollektiv der Messstation **NÜRNBERG** (BY053, ohne Abbildung) zu finden, der in der 1. Stufe der Validierung bereits als Kalibrierwert erkannt wurde. Dieser Wert ist selbstverständlich nicht plausibel und muss entfernt werden. Zwei weitere Messwerte von $253\mu\text{g}/\text{m}^3$, auf die diese Aussage zutrifft, werden noch am 02.03. und am 10.10. im Datenkollektiv gefunden.

Ohne Abbildung: Das Datenkollektiv der Station NÜRNBERG (BY053) enthält immer noch einige Kalibrierwerte (siehe oben)

Messjahr 1981

Im Messjahr 1981 ergeben sich in der zweiten Validierungsstufe nur wenige Messstationen, bei denen einige Messwerte bemängelt werden. Hiervon sind 3 Stationen in Baden-Württemberg und 1 Station in Bayern auffällig

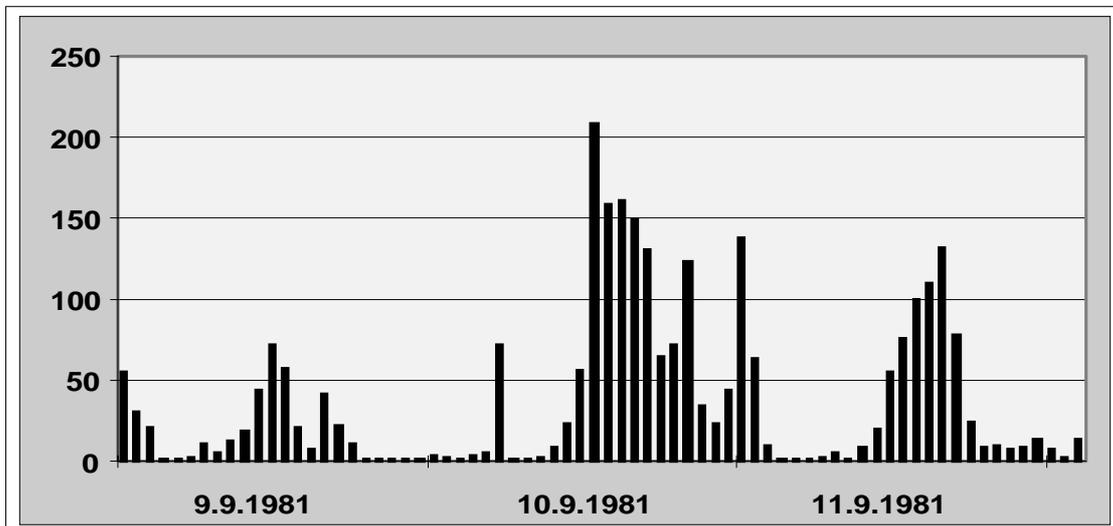
Die Station **MANNHEIM-SÜD** (BW007) berichtet am 31.08.1981 einen Messwert von $167\mu\text{g}/\text{m}^3$ um 8 Uhr morgens. Dieser Wert liegt über den höchsten Maxima der übrigen Stationen, die zwischen etwa $30\mu\text{g}/\text{m}^3$ und $60\mu\text{g}/\text{m}^3$ betragen. Bereits bei der visuellen Überprüfung wurden die Werte dieser Station in diesem Zeitraum insgesamt als fragwürdig bezeichnet. Der hier kritisierte Wert ragt in Abb.15 als nicht plausibler Ausreißer nach einer längeren Ausfallzeit aus nachfolgenden plausiblen Messwerten heraus und sollte aus dem Datenkollektiv entfernt werden.



**Abb. 15) Ausreißer oder Kalibrierwert im Datenkollektiv der Station
MANNHEIM SÜD (BW007)**

Aus Baden-Württemberg wird auch von der Station **BAD CANNSTADT** (BW013) ein fragwürdiger Messwert am 10.09. gemeldet. Dieser liegt mit $208\mu\text{g}/\text{m}^3$ weit über den Maxima anderer Messstationen, von denen Werte zwischen etwa $70\mu\text{g}/\text{m}^3$ und $90\mu\text{g}/\text{m}^3$ berichtet werden. Die Abb.16 zeigt den Wert als tatsächlich fragwürdigen Ausreißer in einem insgesamt recht plausiblen zeitlichen Umfeld. Da auch in den nachfolgenden Stunden ähnliche Spitzen beobachtet werden, ist ein Vorgang, der dem Hürth-Effekt

vergleichbar ist, nicht auszuschließen. Daher ist der Messwert mit Einschränkungen noch als plausibel anzusehen und sollte im Datensatz verbleiben.



**Abb. 16) Ausreißer oder HÜRTH-EFFEKT an der Station
BAD CANNSTADT (BW013)**

Die Messwerte der Station **HEILBRONN** (BW015) aus dem Jahr 1981 wurden in der ersten Stufe der Validierung insgesamt als fragwürdig angesehen. Die zweite Validierungsstufe meldet von dieser Station einen Messwert am 28.06. als möglicherweise nicht plausibel. Mit $269\mu\text{g}/\text{m}^3$ übersteigt dieser Wert deutlich diejenigen Maxima der übrigen Messstationen, die etwa zwischen $70\mu\text{g}/\text{m}^3$ und $90\mu\text{g}/\text{m}^3$ an diesem Tag und zu dieser Stunde melden. Wie die Abb.17 zeigt, handelt es sich bei diesem Wert um einen einzelnen „Messwert“, der in einem zeitlichen Umfeld von Aus-fallwerten eingebettet ist. Hierbei handelt es sich fraglos um einen nicht plausiblen Wert, der aus dem Kollektiv entfernt werden sollte.

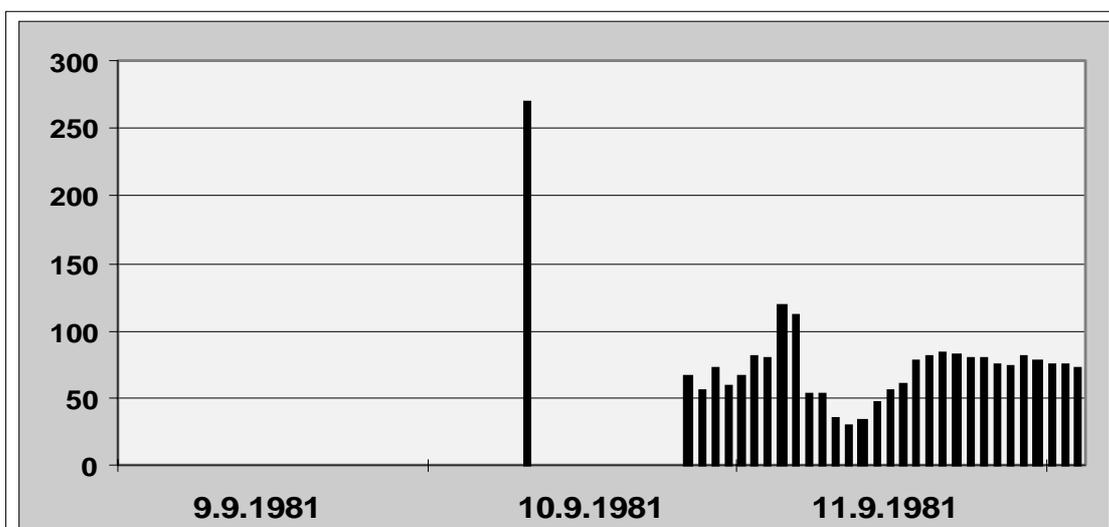


Abb. 17) Nicht plausibler Einzelwert in einer Ausfallperiode an der Station HEILBRONN (BW015)

Eine Messstation aus dem Messnetz in Bayern, nämlich **MÜNCHEN-EFFNERPLATZ (BY038)**, ist im Jahr 1981 ebenfalls mit einem Messwert in der Liste fragwürdiger Daten vertreten. Hierbei handelt es sich um ein relativ niedriges Maximum von “nur” $174\mu\text{g}/\text{m}^3$,

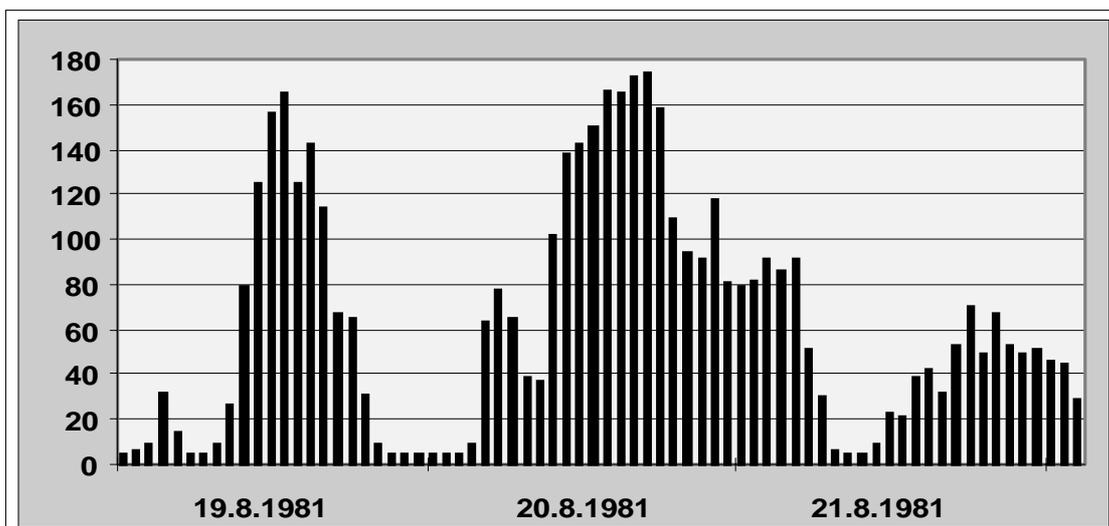


Abb. 18) Als fragwürdig bezeichnete Werte sind offensichtlich plausibel, Station MÜNCHEN-EFFNERPLATZ (BY038)

das am 20.08. um 16 Uhr ermittelt wurde und um gerade $100\mu\text{g}/\text{m}^3$ über den Maxima der übrigen Messstationen mit den drei höchsten Maxima liegt. Die Abb.18 zeigt jedoch im Zeitraum vor und nach diesem kritisierten Messwert keine ungewöhnlichen Konzentrationen, so dass der fragliche Wert als plausibel anzusehen ist und im Datenkollektiv verbleiben sollte.

Messjahr 1982

Das Messjahr 1982 ergibt eine Vielzahl von Messstationen, von denen Daten in dieser Validierungsstufe als fragwürdig bezeichnet werden. Hierbei handelt es sich um Stationen aus Baden-Württemberg, Bayern, Nordrhein-Westfalen und aus dem UBA-Messnetz. Im Mai, Juni und Juli ist die Station **EGGENSTEIN** (BW004) auffällig. Am 18. und am 31. Mai meldet diese Station Maxima von $225\mu\text{g}/\text{m}^3$ bzw. $312\mu\text{g}/\text{m}^3$. Am erstgenannten Tag liegen die Maxima der übrigen Stationen zwischen etwa $100\mu\text{g}/\text{m}^3$ und $150\mu\text{g}/\text{m}^3$ und am 31.05. zwischen etwa $190\mu\text{g}/\text{m}^3$ und $200\mu\text{g}/\text{m}^3$. Die Darstellung in der Abb.19 zeigt am 18.05. zwei Messwerte, die als Ausreißer aus einem relativ plausiblen Tagesgang der Ozonkonzentration herausragen. Diese Ausreißer sind offensichtlich auf den **HÜRTH-EFFEKT** zurückzuführen, der bei dieser Station in mehreren Jahren häufig zu beobachten ist. Daher sollten die Werte als plausibel angesehen werden. Auch der 31.05. zeigt (ohne Abbildung) keinen Konzentrationsverlauf, bei dem einzelne Werte a priori als unplausibel betrachtet werden müssen. So ist auch das kritisierte Maximum als plausibel anzusehen. Im folgenden Monat werden bei der Station **EGGENSTEIN** erneut 2 Werte zu aufeinanderfolgenden Zeitpunkten als fragwürdig ausgegeben.

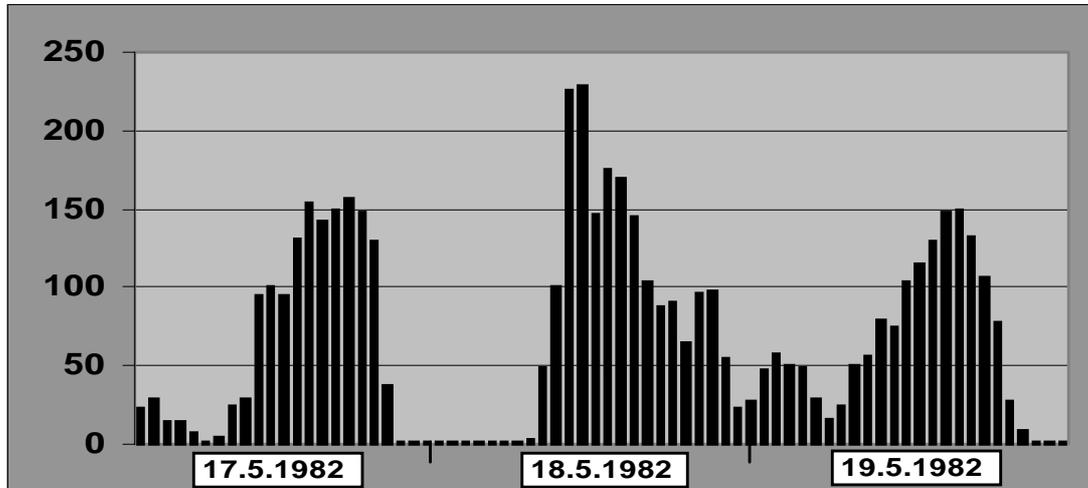


Abb. 19) Ausreißer oder HÜRTH-EFFEKT an der Station EGGENSTEIN im Mai 1982 (BW004)

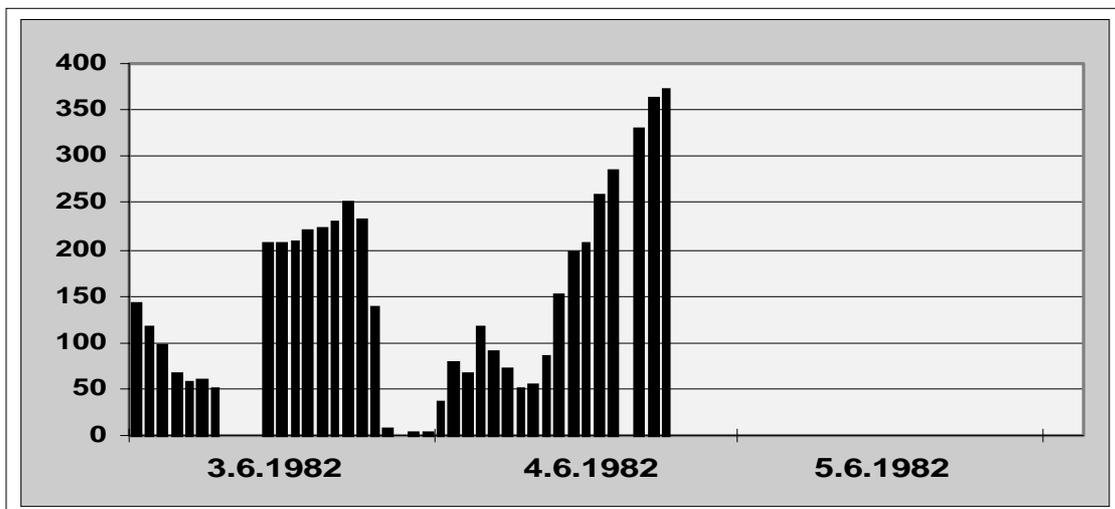


Abb. 20) Fragwürdige Messwerte an der Station EGGENSTEIN im Juni 1982 (BW004)

Es handelt sich hierbei um den 04.06. um 16 Uhr bis 17 Uhr am Nachmittag (Abb.20). Zeitlich liegen die Messwerte vor einer länger andauernden Ausfallzeit. Auch am vorhergehenden Tag sind neben vielen Ausfallwerten einige Messwerte zu verzeichnen, bei denen der übliche Tagesgang der Ozonkonzentration nur sehr undeutlich zu

erkennen ist. Darüber hinaus gibt es jedoch keine Argumente, die es rechtfertigen, diese kritisierten Maxima zu verwerfen; sie sind daher als plausibel zu betrachten. Mit $252\mu\text{g}/\text{m}^3$ meldet die Messstation **EGGENSTEIN** am 20.07.1982 ebenfalls ein Maximum, das um etwas über $100\mu\text{g}/\text{m}^3$ über den Maxima der übrigen in diesem Jahr eingerichteten Stationen liegt.

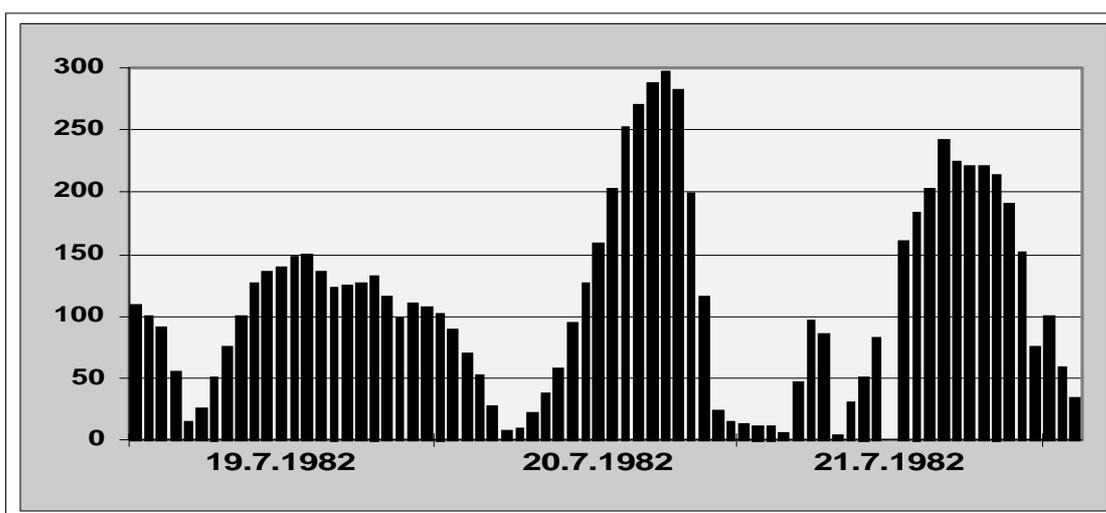


Abb. 21) Fragwürdige, aber plausible Werte an der Station EGGENSTEIN im Juli 1982 (BW004)

Bei diesem Wert handelt es sich jedoch um einen Messwert im ansteigenden Verlauf eines üblichen Tagesganges der Ozonkonzentration, der von einigen weiteren Messwerten noch überschritten wird (Abb.21). So liegt das Maximum dieses Tages bei einem nicht kritisierten Wert von $297\mu\text{g}/\text{m}^3$. Der fragliche Messwert ist mithin als plausibel anzusehen.

Eine weitere Station in Baden-Württemberg meldet Messwerte, von denen am 07.06. ein Maximum von $251\mu\text{g}/\text{m}^3$ bzw. $245\mu\text{g}/\text{m}^3$ (8 Uhr und 9 Uhr) als fragwürdig bezeichnet wird (Abb.22). Es handelt sich um die Station **MANNHEIM-MITTE** (BW 006), bei der dieser Wert um etwas über $100\mu\text{g}/\text{m}^3$ über den Maxima der übrigen Stationen (etwa $100\mu\text{g}/\text{m}^3$ bis $140\mu\text{g}/\text{m}^3$) liegt. Nach einer längeren Ausfallzeit können diese beiden Werte offensichtlich als Messfehler (oder auch Kalibrierwerte) betrachtet

werden. Beide Werte sind mithin nicht plausibel.

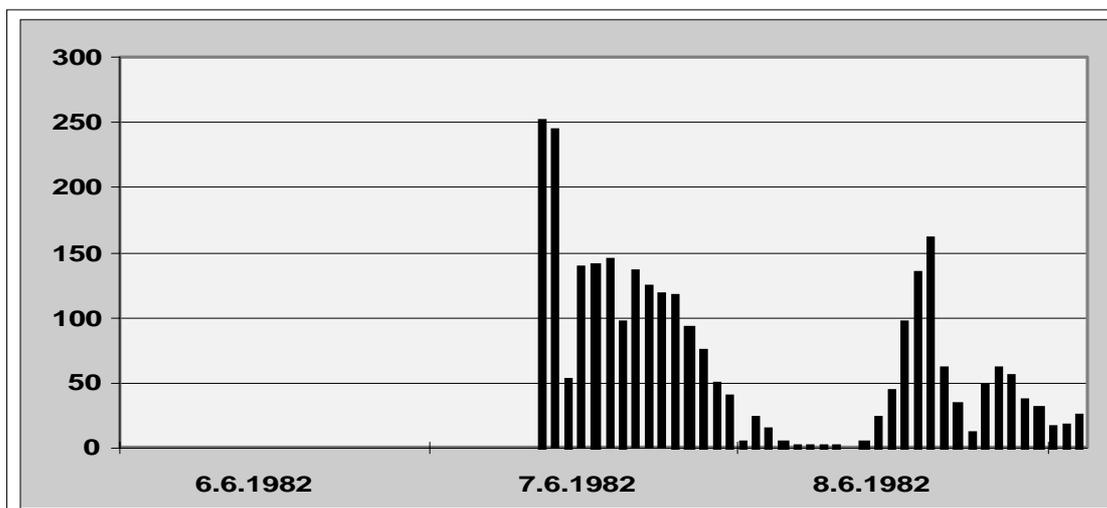


Abb. 22) Ausreißer oder Kalibrierwerte an der Station MANNHEIM-MITTE (BW006)

Die Messstation **MANNHEIM-SÜD** (BW007) ist im Wintermonat März mit 3 auffälligen Messwerten vertreten. Diese Station meldet am 02.03.1982 in den späten Nachmittagstunden Messwerte zwischen $200\mu\text{g}/\text{m}^3$ und $300\mu\text{g}/\text{m}^3$, die als fragwürdig dargestellt werden (ohne Abbildung). Allerdings sind die Messwerte aus dem Jahr 1982 insgesamt in der ersten Validierungsstufe als nicht plausibel verworfen worden, so dass hier eine zusätzliche Entscheidung nicht erforderlich ist

Ohne Abbildung: Das Datenkollektiv der Station MANNHEIM-SÜD (BW007) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Die Station **ZUFFENHAUSEN** (Stuttgart, BW011) meldet am 09.06. und 10.06. Maxima, die deutlich über den Maxima der übrigen Stationen liegen und am 09.06. in den Abendstunden bis etwa 22 Uhr mit Messwerten über $300\mu\text{g}/\text{m}^3$ recht hoch ausfallen. An diesem Tage melden die übrigen Messstationen Maxima zwischen $90\mu\text{g}/\text{m}^3$ und $130\mu\text{g}/\text{m}^3$. Der am 10.06. kritisierte Wert liegt bei 10 Uhr am Vormittag und mit $223\mu\text{g}/\text{m}^3$ erheblich niedriger als die Werte des Vortages. Bereits bei der 1. Stufe der Validitätsprüfung sind die Messwerte dieser Station insgesamt als nicht plausibel verworfen worden. Diese Einschätzung wird durch den Ausschnitt der Werte,

Ohne Abbildung: Das Datenkollektiv der Station LIS-ESSEN (NW028) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Ohne Abbildung: Das Datenkollektiv der Station WALSUM (NW034) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Ohne Abbildung: Die Monate AUG und SEP im Datenkollektiv der Station DORMAGEN (NW050) wurden bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

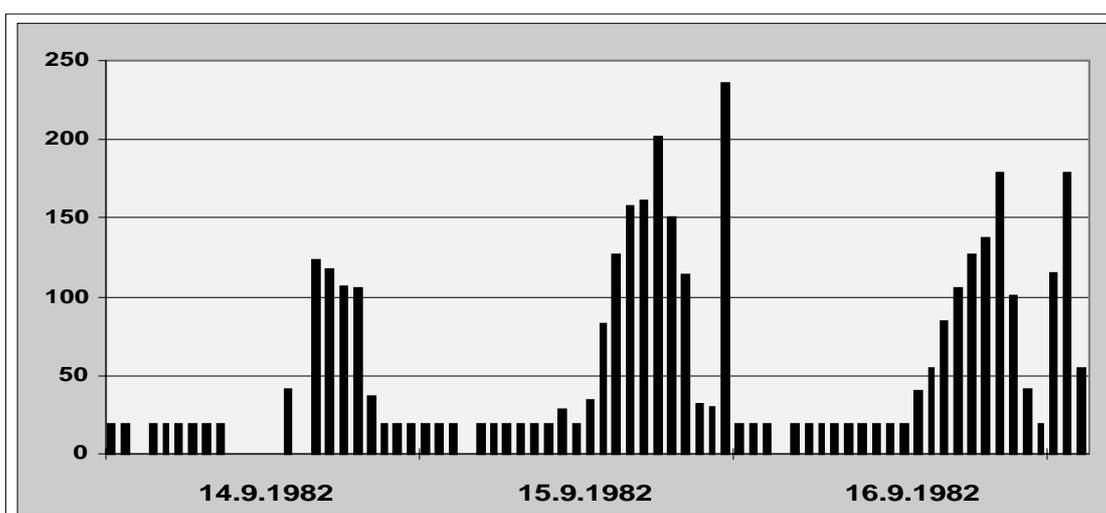


Abb. 24) Nicht plausible Messwerte und hohe "Sockelwerte" an der Station HÜRTH (NW058)

Die Messstation **SCHMÜCKE** (UB029), die heute dem Messnetz des Umweltbundesamtes angehört und im Jahr 1982 vom Meteorologischen Dienst der DDR betrieben wurde, meldet in diesem Jahr mehrere Messwerte, die von der 2. Stufe der Validierung als fragwürdig angesehen werden. Bereits in der 1. Stufe wurden die Messwerte für den Monat November (ohne Abbildung) als wenig plausibel bezeichnet. Diese Werte werden jedoch in der 2. Stufe nicht bemängelt, obwohl an den ersten Tagen des Novembers sehr hohe Werte von über $200\mu\text{g}/\text{m}^3$ berichtet wurden. Als erster fragwürdiger Messwert wird ein Wert von $235\mu\text{g}/\text{m}^3$ bemängelt, der um 21 Uhr am 27.05.1982 beobachtet wurde. Wenn dieser Wert jedoch als nicht plausibel verworfen werden muss, so sind auch weitere Messwerte an diesem und an den vorhergehenden

Tagen als nicht plausibel einzustufen. Dies kann jedoch anhand der Darstellung in Abb.25 nicht entschieden werden, so dass die Messwerte zunächst noch als plausibel zu betrachten sind. Am 06.06. werden vier Messwerte bemängelt, die in den frühen Morgenstunden mit über $200\mu\text{g}/\text{m}^3$ deutlich höher als die Maxima der übrigen Stationen liegen. Diese Werte sollten zunächst noch als plausibel betrachtet werden, auch wenn hier eine Entscheidung, ob sie und die benachbarten Werte plausibel sind oder nicht, nicht getroffen werden kann (Abb.26). Der Messwert von $194\mu\text{g}/\text{m}^3$ ragt als Spitzenwert am 05.09.1982 aus einem geringfügig niedrigeren Niveau der Messwerte der Station **SCHMÜCKE** heraus (Abb.27). Eine Entscheidung über die Plausibilität dieses Wertes erscheint als nicht erforderlich; der Wert bleibt mithin als plausibel im Datenkollektiv. Eine Serie weiterer fragwürdiger Messwerte zwischen $150\mu\text{g}/\text{m}^3$ und $250\mu\text{g}/\text{m}^3$ werden von dieser Station am 2. November zwischen 12 Uhr und 16 Uhr berichtet (ohne Abbildung). Diese Werte sind ebenso wie die Messwerte aus den 2 vorhergehenden Stunden als Ausreißer anzusehen und nicht plausibel.

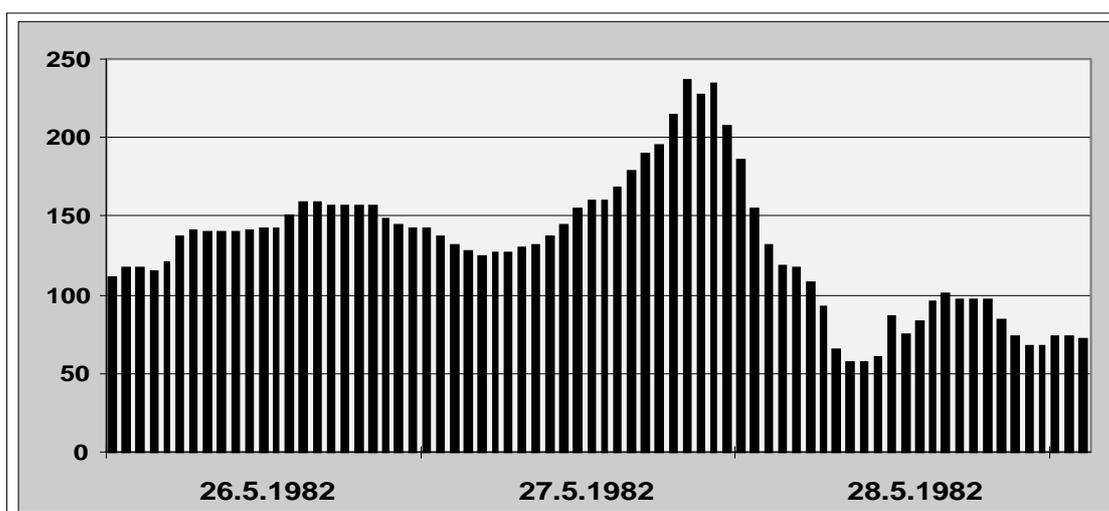


Abb. 25) Fragwürdige, aber nicht unplausible Werte an der Station SCHMÜCKE (UB029) im Mai 1982

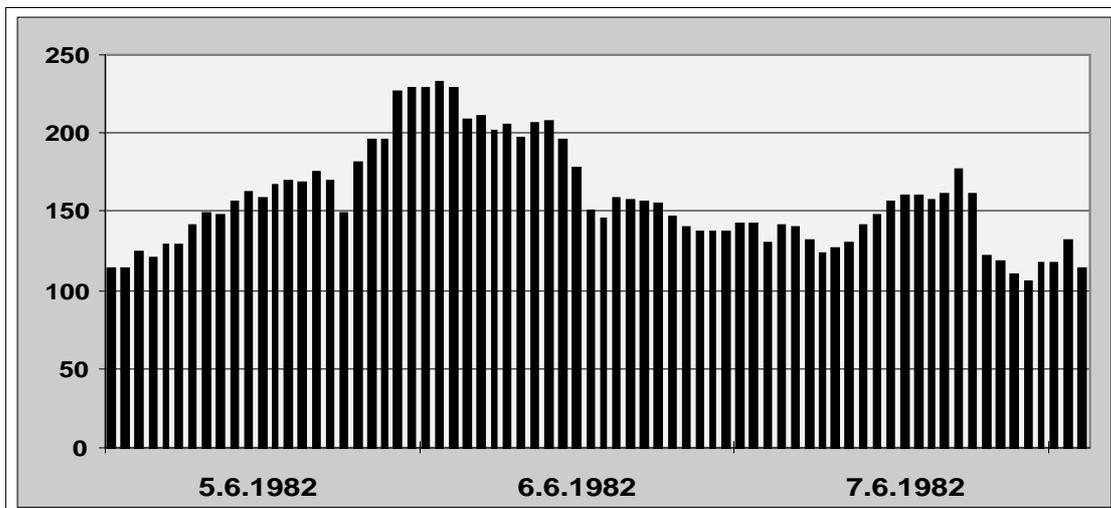


Abb. 26) Fragwürdige, aber nicht unplausible Werte an der Station SCHMÜCKE (UB029) im Juni 1982

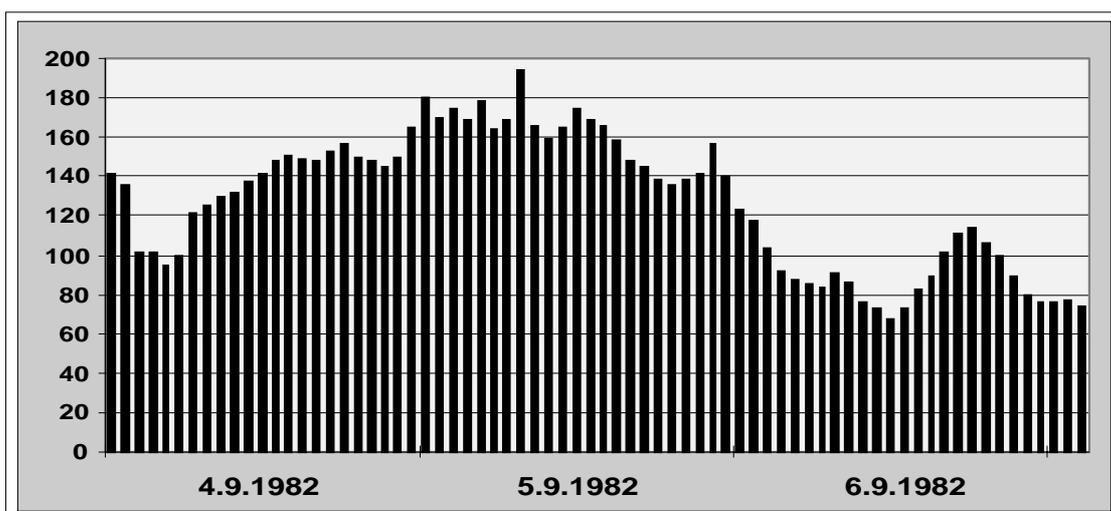


Abb. 27) Fragwürdige, aber nicht unplausible Werte an der Station SCHMÜCKE (UB029) im September 1982

Einige weitere Messwerte sind an der Station SCHMÜCKE am 2.11.1982 als nicht plausibel einzustufen

Messjahr 1983

Auch im Messjahr 1983 wird in der 2. Stufe der Validierung eine große Anzahl fragwürdiger Messwerte festgestellt. Betroffen ist hiervon hauptsächlich das Land Baden-Württemberg, was sich schon allein daraus ergibt, dass in diesem Land die meisten Ozonmessstationen eingerichtet sind. Auch das Messnetz des Umweltbundesamtes ist mit einer Station, die fragwürdige Messwerte liefert, vertreten.

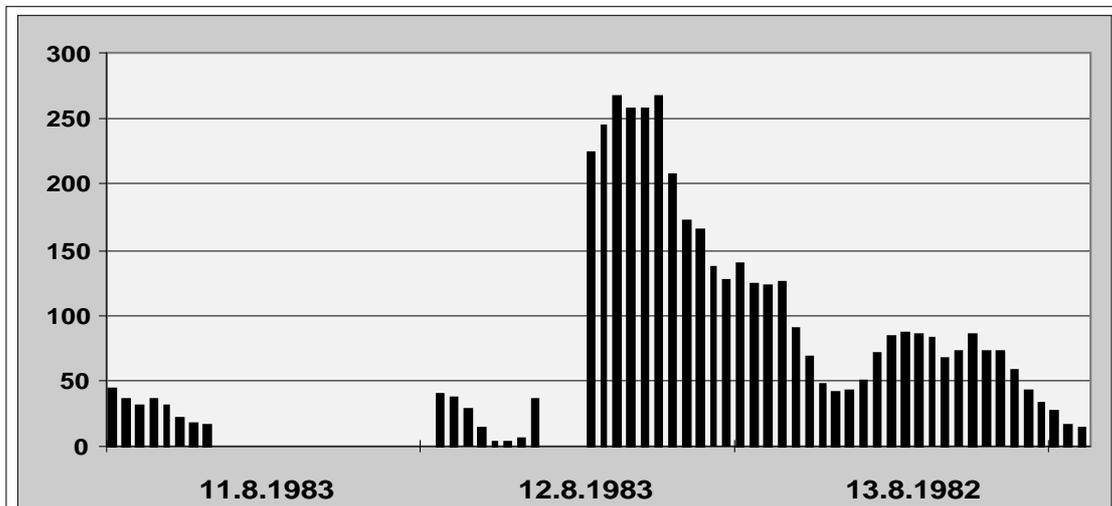
Im Juli des Jahres 1983 meldet die Messstation **KARLSRUHE-WEST** (BW002) mehrere fragwürdige Messwerte um $300\mu\text{g}/\text{m}^3$ (ohne Abbildung). Diese - und das Datenkollektiv insgesamt - sind jedoch bereits in der 1. Stufe der Validierung als nicht plausibel bezeichnet worden. Auch die Messwerte der Station **EGGENSTEIN** (BW004) sind in der ersten Validierungsstufe insgesamt und insbesondere in den Monaten Juli und August als nicht plausibel verworfen worden. Die 2. Stufe bemängelt demnach auch zahlreiche Messwerte, die im Juli und August mit Werten zwischen $300\mu\text{g}/\text{m}^3$ und $400\mu\text{g}/\text{m}^3$ gemessen wurden (ohne Abbildung).

Ohne Abbildung: Das Datenkollektiv der Station KARLSRUHE-WEST(BW002) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Ohne Abbildung: Das Datenkollektiv der Station EGGENSTEIN (BW004) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Das Datenkollektiv der Messstation **MANNHEIM-NORD** (BW005) wurde in der 1. Stufe zwar nicht verworfen, es erhielt jedoch insgesamt das Prädikat „Werte sind fragwürdig“. Die 2. Validierungsstufe meldet nur einen Messwert, der mit $224\mu\text{g}/\text{m}^3$ mehr als $100\mu\text{g}/\text{m}^3$ über den Werten der übrigen Stationen liegt. Dieser Wert wurde am 12.08.1983 ermittelt (Abb.27). Nach mehreren offensichtlich nicht plausiblen Werten und Ausfällen in den vorangehenden Stunden und Tagen wird dieser hier diskutierte Wert unmittelbar nach Ausfallwerten gemessen. Es folgen jedoch noch weitere und sogar noch höhere Messwerte (bis etwa $270\mu\text{g}/\text{m}^3$), die in dieser 2. Stufe nicht bemängelt wurden. Der nachfolgende Tag bringt allerdings nur eine Ozonbelastung von unter $100\mu\text{g}/\text{m}^3$. Wenn also der fragliche Wert als nicht plausibel bezeichnet werden soll, so müssen auch die nachfolgenden Werte ebenso bezeichnet werden. Das aber

kann hier nicht entschieden werden; damit bleiben die Werte als plausibel im Datenkollektiv.



**Abb. 27) Fragwürdige, aber nicht unplausible Werte an der Station
MANNHEIM NORD (BW005)**

Die Station **MANNHEIM-SÜD** (BW007) liefert für das Jahr 1983 ebenfalls ein Datenkollektiv, dessen Werte in der 1. Stufe das Prädikat „Werte sind fragwürdig“ erhielten. Gleichwohl wurde das Kollektiv in dieser ersten Stufe nicht verworfen. Am 29.07. und 31.07. meldet die 2. Stufe mehrere Werte um und über $200\mu\text{g}/\text{m}^3$, die zum Teil weit über den Messwerten der übrigen Stationen liegen. Dass die Einschätzung in der 1. Validierungsstufe gerechtfertigt ist, zeigt die Darstellung der Messwerte in Abb.28. Die in der zweiten Stufe kritisierten Werte sowie die übrigen Werte in deren zeitlicher Nachbarschaft erscheinen insgesamt als nicht plausibel.

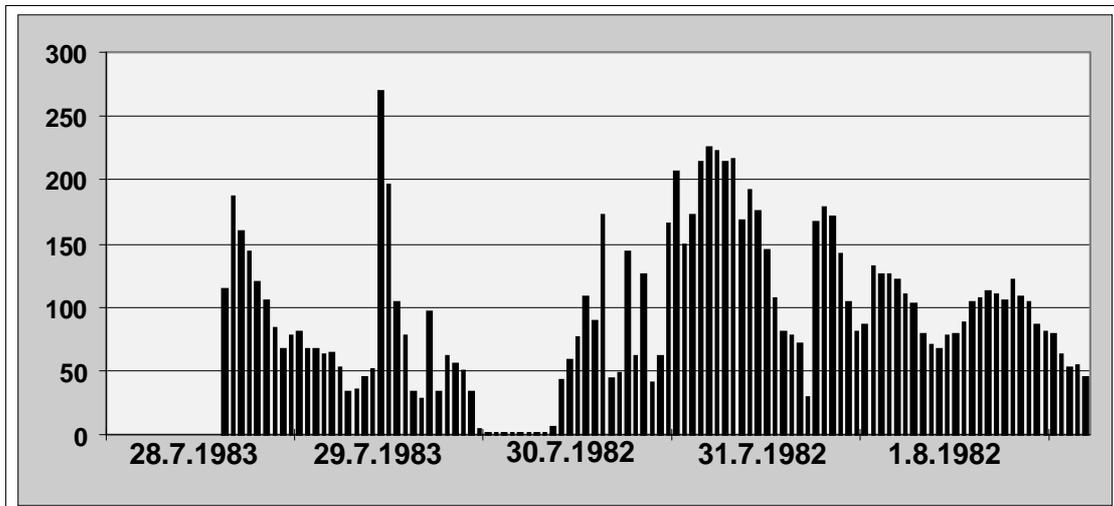


Abb. 28) Nicht plausible Messwerte an der Station MANNHEIM-SÜD (BW007)

Mit mehreren zum Teil sehr hohen, aber fragwürdigen Messwerten ist die Station **STUTTGART-HAFEN** (BW014) in der 2. Stufe der Validierung auffällig. Obwohl das gesamte Ozondatenkollektiv dieser Station in der 1. Stufe bereits als nicht plausibel verworfen wurde, sollen die in der 2. Stufe kritisierten und am 23.07. und 28.07. ermittelten Messwerte in Abb.29 zur Illustration der Bewertung der 1. Stufe dargestellt werden. In Anbetracht der zahlreichen Ausfälle und der großen Schwankungen der Messwerte in diesem Zeitraum, aber insbesondere auch im Zusammenhang mit den übrigen Messwerten dieses Jahres, die hier nicht dargestellt werden können, ist die Bewertung der 1. Stufe offensichtlich gerechtfertigt.

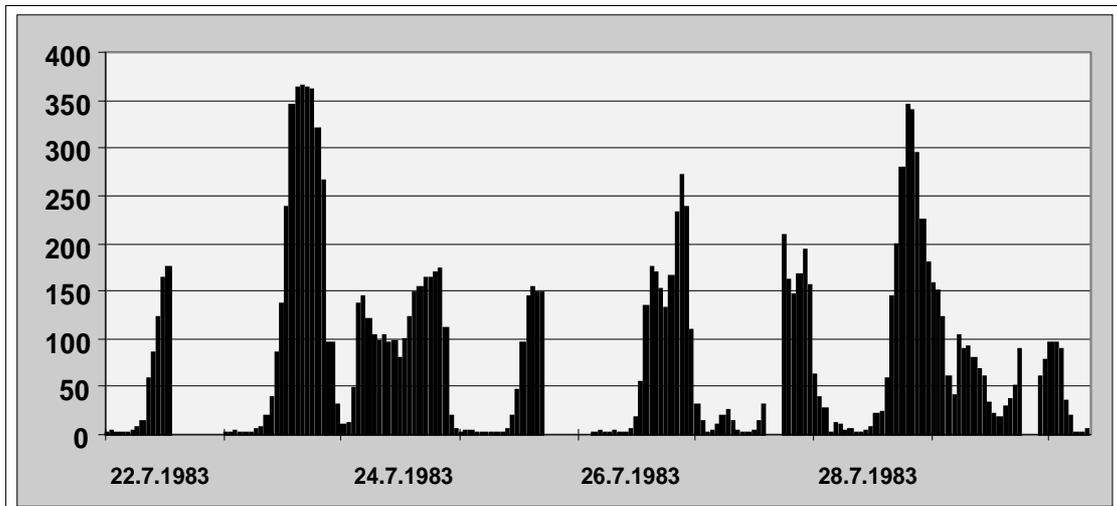


Abb. 29) Nicht plausible Messwerte in einem insgesamt nicht plausiblen Datenkollektiv der Station STUTTGART-HAFEN (BW014)

In Sachsen meldet die im Jahr 1983 vom Meteorologischen Dienst der DDR betriebene Station **WAHNSDORF** (SN051) am 09.09. (Abb.30) einen fragwürdigen Wert von $170\mu\text{g}/\text{m}^3$. Dieser und auch die zeitlich benachbarten Werte scheinen tatsächlich nicht plausibel zu sein, wie der Darstellung in Abb.2 zu entnehmen ist. Es wäre vermutlich angebracht, diese Werte insgesamt auf „Ausfall“ zu setzen, wenn das Konzentrationsniveau mit Werten zwischen etwa $120\mu\text{g}/\text{m}^3$ bis $170\mu\text{g}/\text{m}^3$ insgesamt nicht so „niedrig“ wäre, dass ein „Nicht-Eingreifen“ ebenfalls gerechtfertigt ist. Es wird daher vorgeschlagen, die Werte als plausibel zu betrachten.

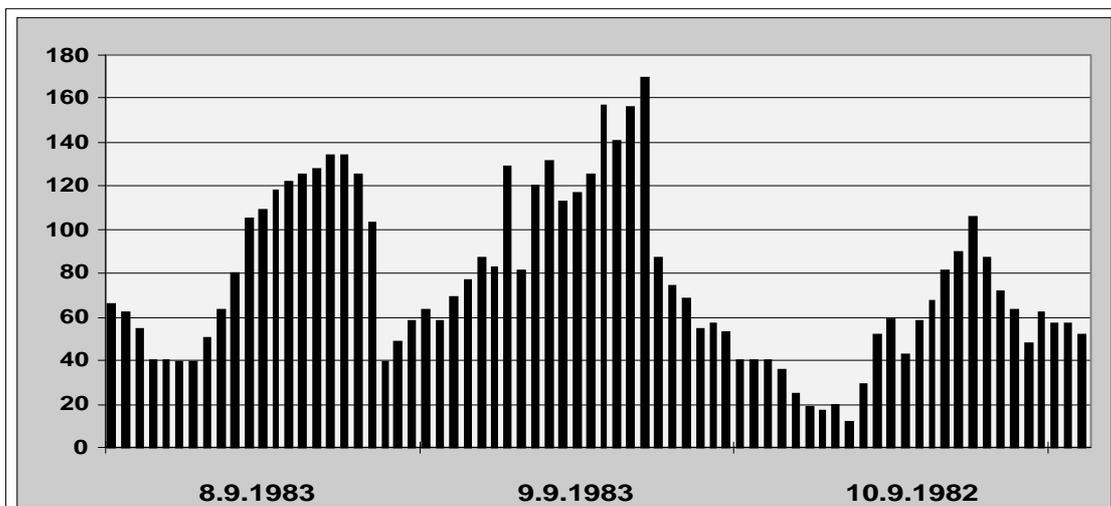


Abb. 30) Nicht plausible Messwerte an der Station WAHNSDORF (SN051), die Werte bleiben jedoch im Kollektiv

Ebenso wie im vorhergehenden Messjahr meldet auch im Jahr 1983 die Station **SCHMÜCKE** (UB029) mehrere fragwürdige Messwerte. Bereits im März dieses Jahres wird eine Serie von Messwerten als fragwürdig bezeichnet, die am 03.03.1983 in der Zeit zwischen 15 Uhr und 21 Uhr um $200\mu\text{g}/\text{m}^3$ schwanken. Die Graphik in Abb.31 zeigt die kritisierten Messwerte in einem Komplex mehrerer ebenfalls wenig plausibler Werte am 3.3.1983, die das für Wintermonate glaubhafte Niveau von etwa $50\mu\text{g}/\text{m}^3$ an den übrigen Tagen weit überragen. Es kann vermutet werden, dass es sich bei dieser Gruppe von "Werten" um fehlerhafte Schätzwerte handelt, mit denen eine Datenlücke aufgefüllt wurde. Diese sind aus dem Datenkollektiv der Station zu entfernen.

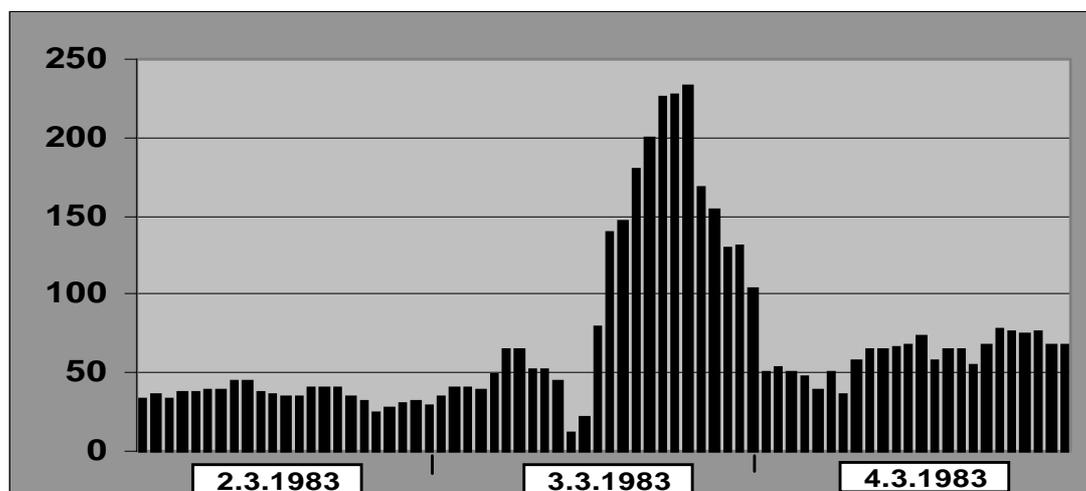


Abb. 31) Nicht plausible Werte (Schätzwerte) an der Station SCHMÜCKE (UB029) im März 1983

Die ersten Werte in den Sommermonaten werden am 14. Mai um 2 Uhr bzw. 3 Uhr morgens mit lediglich etwa $140\mu\text{g}/\text{m}^3$ festgestellt (ohne Abbildung). Diese Messwerte erscheinen deshalb fragwürdig, weil der Mittelwert der höchsten Maxima der übrigen Messstationen nur geringfügig über $30\mu\text{g}/\text{m}^3$ liegt. Das Niveau der Ozonbelastung liegt in diesen Tagen an dieser Bergstation mit Werten zwischen $100\mu\text{g}/\text{m}^3$ und $120\mu\text{g}/\text{m}^3$ so hoch, dass die fragwürdigen Werte nicht zwingend als nicht plausibel eingestuft werden müssen. Sie können daher im Datenkollektiv verbleiben. Am 25.06. wird erneut ein fragwürdiger Messwert mit $193\mu\text{g}/\text{m}^3$ festgestellt. In der Abb.32 erscheint nicht nur dieser Wert fragwürdig, sondern auch die Werte bis zum keinesfalls plausiblen Wert von nur $1\mu\text{g}/\text{m}^3$ am 24.06., da ab diesem Tag ein stetiger Anstieg bis zum Maximum in den Morgenstunden des 25.06. erfolgt, der nicht dem üblichen Tagesgang der Ozonkonzentration entspricht. Es scheint empfehlenswert zu sein, alle Messwerte vom Wert $1\mu\text{g}/\text{m}^3$ bis zum Maximum als nicht plausibel anzusehen und auf Ausfall zu setzen. Weitere fragwürdige Werte werden an den Tagen 09.07. und 10.07.1983 gemeldet. Diese Werte liegen zwischen $200\mu\text{g}/\text{m}^3$ und $240\mu\text{g}/\text{m}^3$ und stellen mithin recht hohe Konzentrationen dar, die jedoch bemerkenswerterweise im Zeitraum zwischen 22 Uhr am 09.07. und 2 Uhr am 10.07. beobachtet werden. Die Darstellung in Abb.33 zeigt - wie bereits oben angemerkt wurde - erneut einen stetigen Anstieg der

Messwerte von einem niedrigen Wert am Mittag des 09.07. bis Mitternacht und erst einen Rückgang der Konzentration in den frühen Morgenstunden des 10.07. Dieser über mehrere Stunden „verschobene“ Tagesgang der Ozonkonzentration ist auch für Bergstationen ungewöhnlich. Daher sollten die genannten fragwürdigen Werte und diejenigen in ihrer zeitlichen Umgebung als nicht plausibel eingestuft werden.

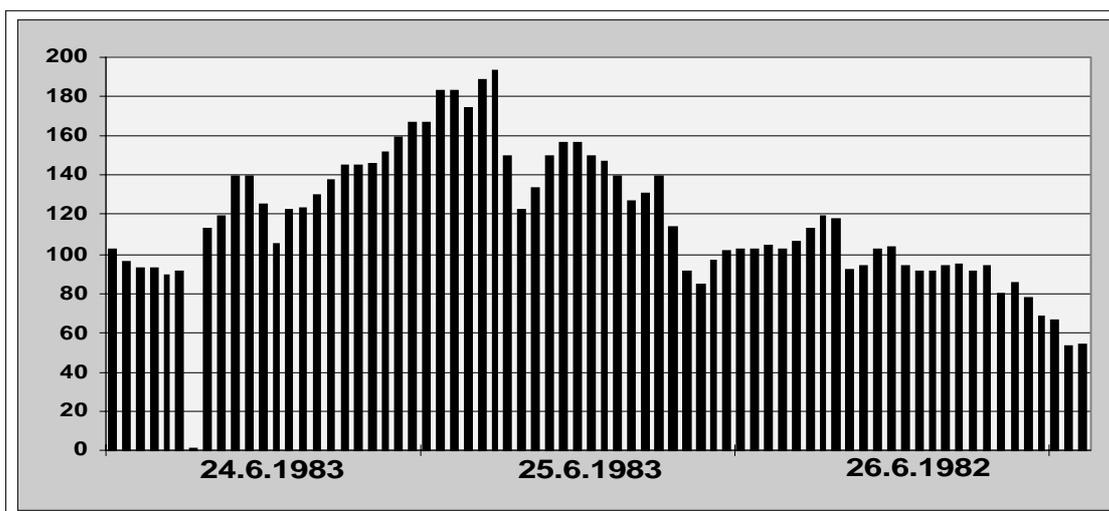


Abb. 32) Fragwürdige Messwerte an der Station SCHMÜCKE (UB029) im Monat Juni 1983

Auch die vom ehemaligen Meteorologischen Dienst der DDR in diesem Jahr betriebene Station **NEUGLOBSOW** (UB030) meldet im Jahr 1983 als fragwürdig bezeichnete Messwerte im Monat Oktober. Am 08.10.1983 werden in den frühen Morgenstunden Werte um $170\mu\text{g}/\text{m}^3$ mitgeteilt. Diese Werte sind in der Abb.34 in ihrem zeitlichen Umfeld dargestellt. Ähnlich wie bereits weiter oben scheint es sich bei diesen Werten um fehlerhaft geschätzte Erwartungswerte zu handeln, die zum Auffüllen einer Datenlücke verwendet wurden. Da auch weitere Messwerte an diesem Tag fragwürdig sind - hierbei handelt es sich um die trendähnlich linear ansteigenden Werte in den Nachmittagsstunden - wird empfohlen, alle Werte dieses Tages insgesamt als unplausibel einzustufen und aus dem Datenkollektiv der Messstation **NEUGLOBSOW** zu entfernen

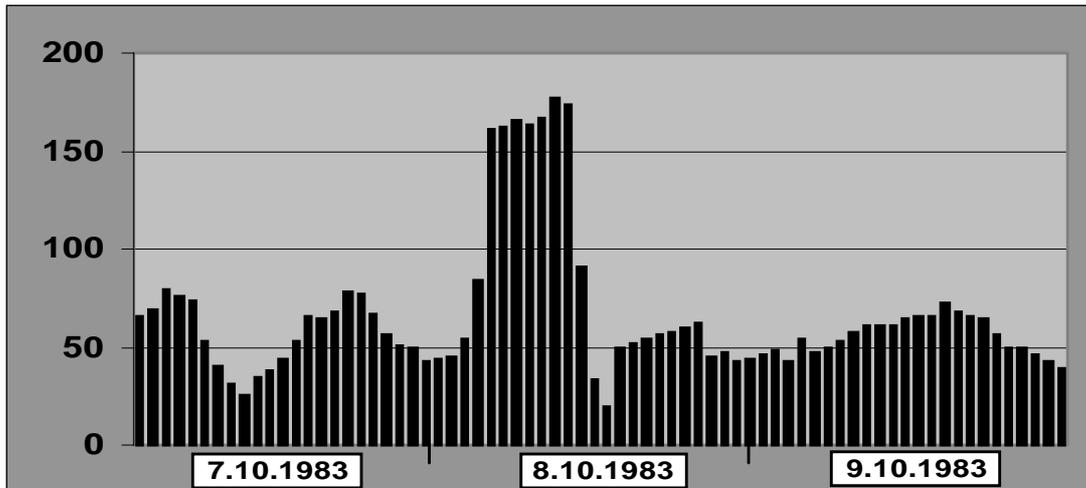


Abb. 34) Fehlerhaft geschätzte Messwerte an der Station NEUGLOBSOW (UB030)

Im Wintermonat Oktober wird bei der hessischen Station **KASSEL-NORD** (HE014) ein Messwert von $162\mu\text{g}/\text{m}^3$ als fragwürdig bezeichnet. Hierbei handelt es sich um einen von 2 Werten zwischen $150\mu\text{g}/\text{m}^3$ und $160\mu\text{g}/\text{m}^3$, die nach einem längeren Ausfall aus einer Gruppe von weiteren Messwerten um $100\mu\text{g}/\text{m}^3$ herausragen (Abb.35), die ebenfalls als fragwürdig bezeichnet werden muss, zumal das Messgerät an dieser Station wenige Tage später für einen längeren Zeitraum ausfällt. Es wird vorgeschlagen, die gesamte Gruppe der Werte bis zu den niedrigen Werten am 13.10. als nicht plausibel einzustufen und aus dem Datenkollektiv zu entfernen.

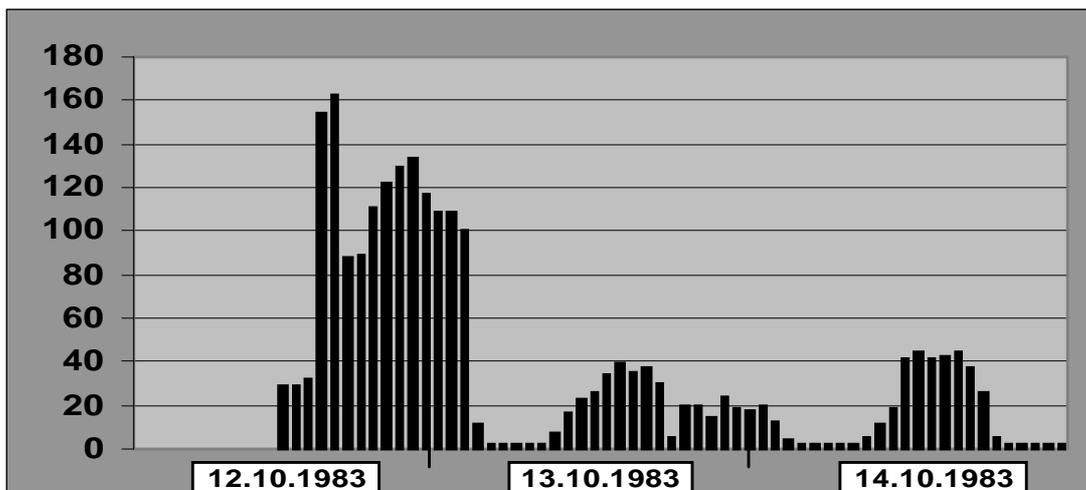


Abb. 35) Nicht plausible Messwerte an der Station KASSEL-NORD (HE014)

Auch das Land Schleswig-Holstein ist im Jahr 1983 - dem ersten Messjahr, in dem der Schadstoff Ozon an den Messstationen dieses Messnetzes registriert wurde - mit einer Messstation mit auffälligen Werten vertreten. So meldet die Messstation **GEESTHACHT** (SH003) am 16.12.1983 einen ($214\mu\text{g}/\text{m}^3$) und am 17.12.1983 drei weitere Messwerte um $150\mu\text{g}/\text{m}^3$ (Abb.36), die die Maxima der übrigen Messstationen um mehr als $100\mu\text{g}/\text{m}^3$ überragen. Es bedarf wohl kaum einer Diskussion, die beanstandeten Werte sowie einen weiteren Wert am 16.12.1983 als nicht plausibel zu bezeichnen und aus dem Datenkollektiv zu entfernen.

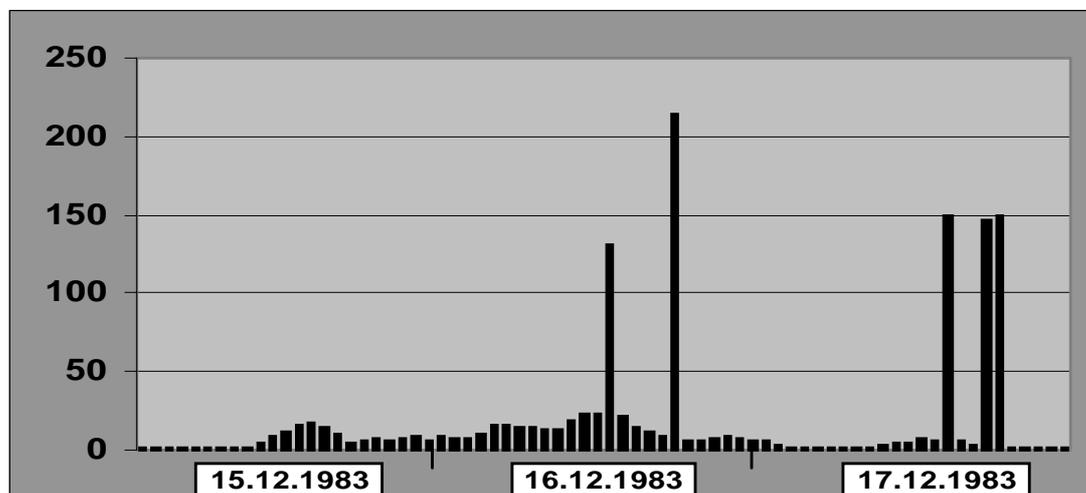


Abb. 36) Nicht plausible Messwerte an der Station GEESTHACHT (SH003)

Messjahr 1984

Im Messjahr 1984 melden überwiegend Stationen aus Baden-Württemberg fragwürdige Messwerte. Daneben werden auch ein Messwert an einer Station in Bayern sowie einige Messwerte an einer Station des Umweltbundesamtes als fragwürdig gekennzeichnet. Zunächst wird die Station **KARLSRUHE-WEST** (BW002) genannt. Das Datenkollektiv dieser Station wurde bereits bei der ersten Stufe der Validierung als nicht plausibel verworfen (ohne Abb.).

Ohne Abbildung: Das Datenkollektiv der Station KARLSRUHE-WEST (BW002) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Von der Station **EGGENSTEIN** (BW004) wird um 9 Uhr des 21.06. ein Messwert von $365\mu\text{g}/\text{m}^3$ gemeldet, der bei der 2. Stufe der Validierung auffällig ist. Dieser Wert liegt um mehr als $200\mu\text{g}/\text{m}^3$ über den 3 höchsten Maxima der übrigen Messstationen. Das gesamte Datenkollektiv auch dieser Station ist in der visuellen Prüfung als nicht plausibel bewertet worden. Es muss nicht ausführlich darüber diskutiert werden, dass dieser Wert, der in Abb.37 zu sehen ist und unmittelbar nach einer Reihe von Ausfallwerten folgt, tatsächlich als nicht plausibel einzustufen ist.

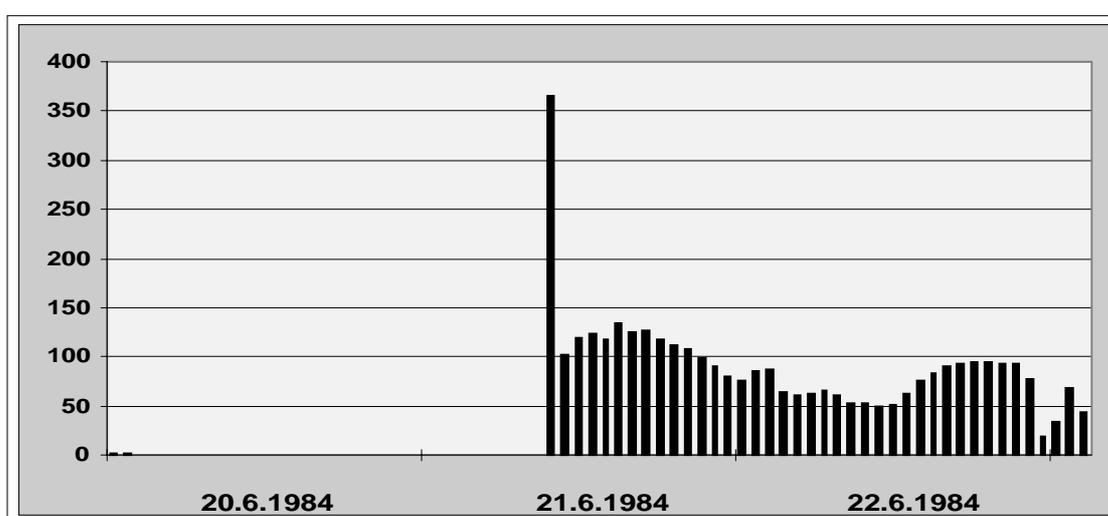
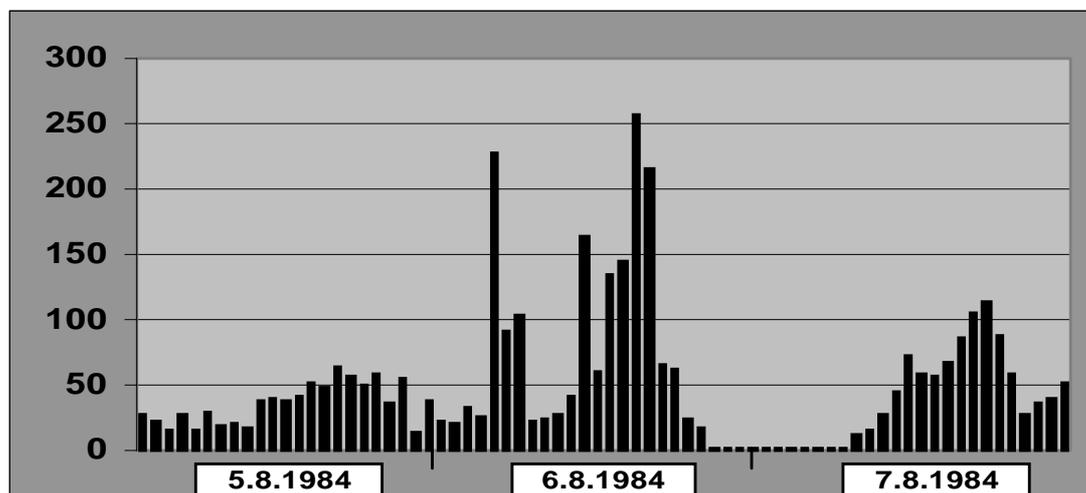


Abb. 37) Ein einzelner nicht plausibler Messwert (Kalibrierwert?) an der Station EGGENSTEIN (BW004)

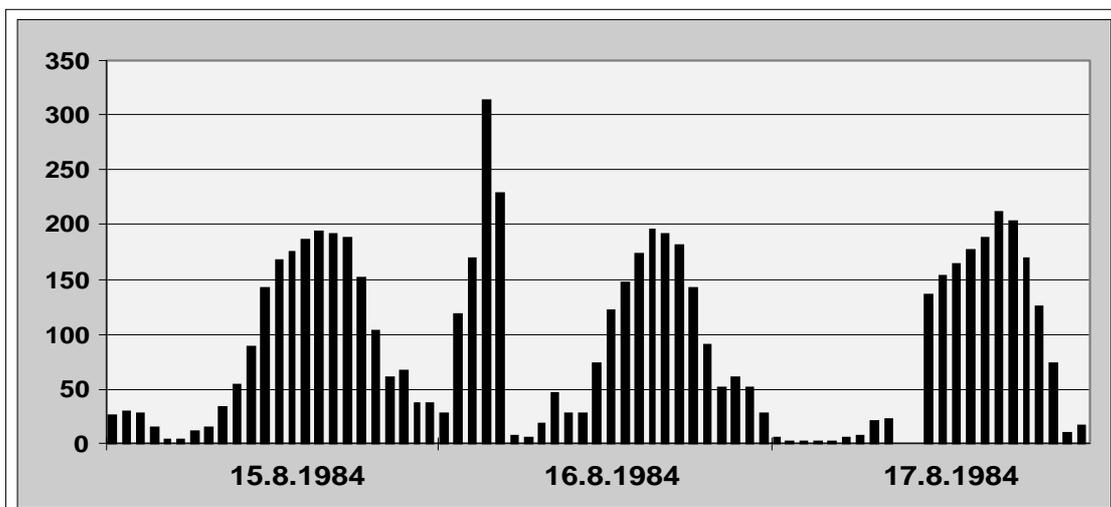
Mehrere zweifelhafte Messwerte werden von der Station **MANNHEIM-NORD** (BW005) im Monat August berichtet. Das Datenkollektiv dieser Station wurde bereits bei der ersten Stufe der Validierung als fragwürdig charakterisiert. Am 06.08. wird ein Wert von $227\mu\text{g}/\text{m}^3$ in den Morgenstunden und zwei Werte von $257\mu\text{g}/\text{m}^3$ und $216\mu\text{g}/\text{m}^3$ in den Nachmittagsstunden beanstandet. Die Maxima der übrigen Stationen liegen bei etwa $100\mu\text{g}/\text{m}^3$. Die Abb.38 zeigt die beanstandeten Messwerte, die tatsächlich als Ausreißer aus einem relativ niedrigen Niveaus herausragen und als nicht plausibel einzustufen sind. Mit Recht wird von der 2. Stufe der Validierung auch der Messwert vom 08.08. angezweifelt, der mit $311\mu\text{g}/\text{m}^3$ ebenfalls als Ausreißer aus einem Umfeld mit niedrigem Niveau herausragt (ohne Abbildung). Auch dieser Wert ist als nicht plausibel anzusehen. Das gleiche gilt für die beanstandeten Messwerte am 09.08.,

aber auch für die von der 2. Stufe nicht beanstandeten Werte an diesem und in den Morgenstunden des folgenden Tages,

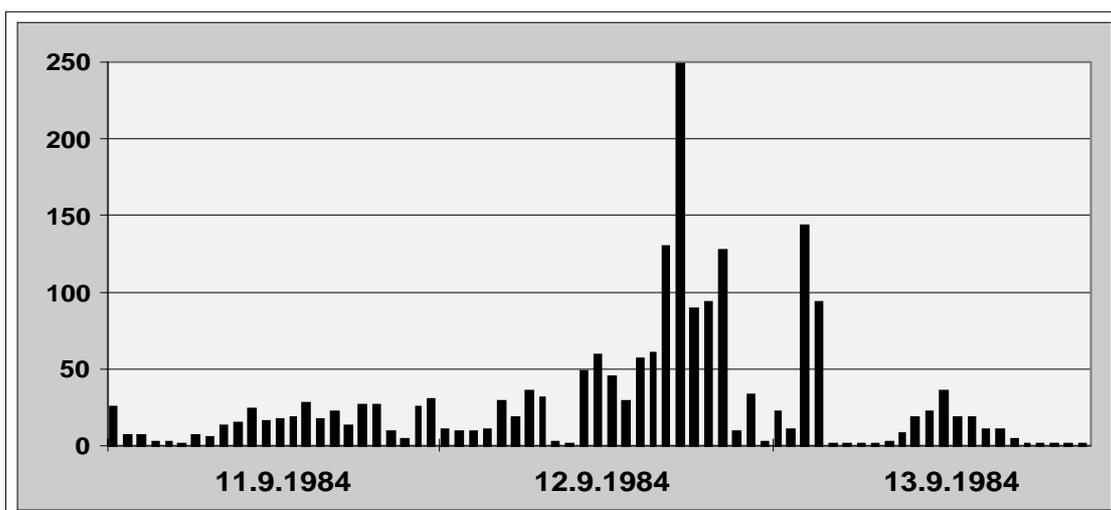


**Abb. 38) Nicht plausible Messwerte im August 1984 an der Station
MANNHEIM NORD (BW005)**

die jedoch „nur“ zwischen $150\mu\text{g}/\text{m}^3$ und $200\mu\text{g}/\text{m}^3$ liegen. Diese Werte sind als nicht plausibel einzustufen. Diese Einstufung trifft auch auf die Werte zu, die in den Morgenstunden des 16. August 1984 mit $313\mu\text{g}/\text{m}^3$ und $228\mu\text{g}/\text{m}^3$ beanstandet werden (Abb.39). Diese beiden Werte sind ebenso wie die zwei nicht beanstandeten Werte der zwei vorhergehenden Stunden sowie die ebenfalls wenig plausiblen Daten am 12. und 13.09.1984 (Abb.40) aus dem Datenkollektiv zu entfernen, wenn man sich nicht der Einschätzung der ersten Validierungsstufe anschließen und das gesamte Datenkollektiv des Jahres 1984 dieser Station als unplausibel verwerfen will.



**Abb. 39) Nicht plausible Ausreißer an der Station MANNHEIM-NORD (BW005)
im Monat August 1984**



**Abb. 40) Nicht plausible Ausreißer an der Station MANNHEIM-NORD (BW005)
im Monat September 1984**

Im Wintermonat März 1984 wird ein Messwert von $179\mu\text{g}/\text{m}^3$ an der bayerischen Station **GAP-KREUZECKBAHN** (BY081) kritisiert. Dieser Wert wurde am 16.03. zu einem Zeitpunkt beobachtet, zu dem die Maxima der übrigen Messstationen ebenfalls bei recht hohen Werten um $100\mu\text{g}/\text{m}^3$ liegen. Die Abb.41 zeigt diesen Messwert in seiner zeitlichen Umgebung. Obwohl der übliche Tagesgang am 16.03. und

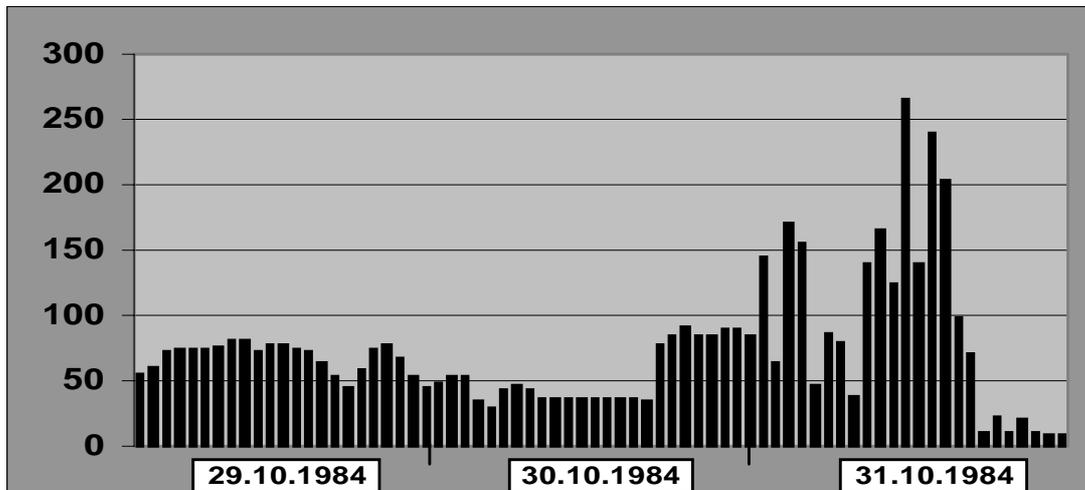


Abb. 42) Nicht plausible Werte (Schätzwerte) an der Station SCHMÜCKE (UB029)

Messjahr 1985

Auch in diesem Jahr fallen erneut einige Messstationen aus Baden-Württemberg mit fragwürdigen Messwerten auf. Bemängelt wird auch eine Station aus Nordrhein-Westfalen sowie eine Station aus Rheinland-Pfalz.

Bei der 1. Validierungsstufe wurden im Datenkollektiv der Station **KARLSRUHE-WEST** (BW002) einige nicht plausible Messwerte festgestellt. Das Datenkollektiv insgesamt wird jedoch als verwendbar eingestuft. Einen dieser Werte bemängelt auch die zweite Validierungsstufe am 25. September. Dieser Wert liegt mit $217\mu\text{g}/\text{m}^3$ um $100\mu\text{g}/\text{m}^3$ höher als die Maxima der übrigen Messstationen. Es bedarf wohl kaum einer Diskussion, diesen Ausreißer (Abb. 43) tatsächlich als nicht plausibel zu bewerten und aus dem Datenkollektiv zu entfernen.

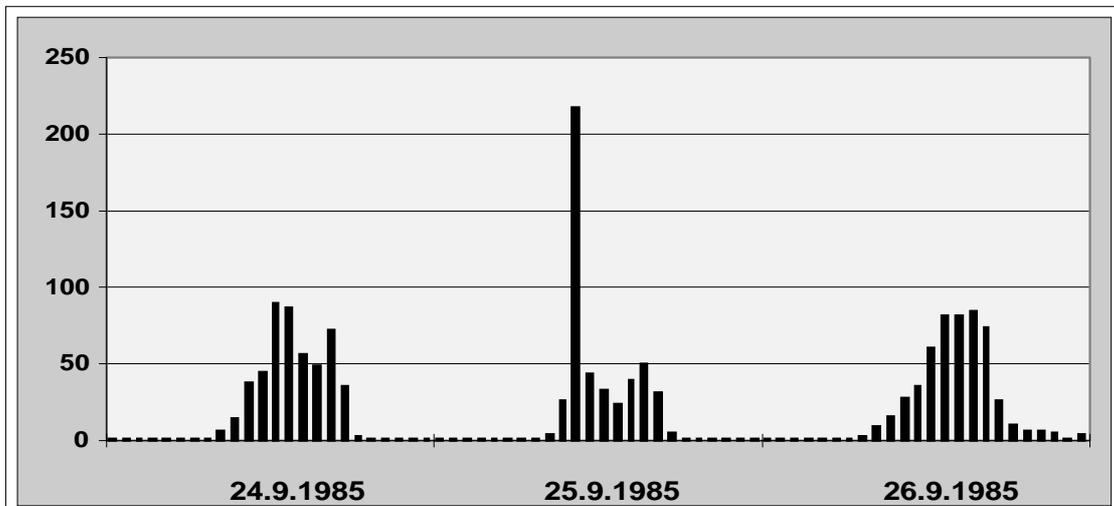


Abb. 43) Nicht plausible Werte an der Station KARLSRUHE-WEST (BW002)

Auch die Messstation **EGGENSTEIN** (BW004) ist in diesem Jahr erneut mit auffälligen Werten vertreten. In der visuellen Überprüfung wurden die Messwerte der Monate Juli und August als nicht plausibel bewertet. Einen dieser Werte bemängelt auch die 2. Stufe. Dieser Messwert wird am 26.07. in den Nachmittagsstunden mit $390\mu\text{g}/\text{m}^3$ zu einer Zeit festgestellt, in der auch von den übrigen Messstationen hohe Konzentrationen zwischen etwa $250\mu\text{g}/\text{m}^3$ und $300\mu\text{g}/\text{m}^3$ berichtet werden. Der beanstandete Messwert zeigt sich zwar als Ausreißer (Abb.44), der sich jedoch nur geringfügig aus einem Niveau hoher Werte von etwa $350\mu\text{g}/\text{m}^3$ heraushebt. Wenn auch der Übergang von den hohen Werten auf die niedrigen Werte um $50\mu\text{g}/\text{m}^3$ in den späten Nachmittagsstunden wenig plausibel ist, könnte der beanstandete Wert noch als plausibel durchgelassen werden. Allerdings muss bei dieser Bewertung das Ergebnis der 1. Validierungsstufe berücksichtigt werden.

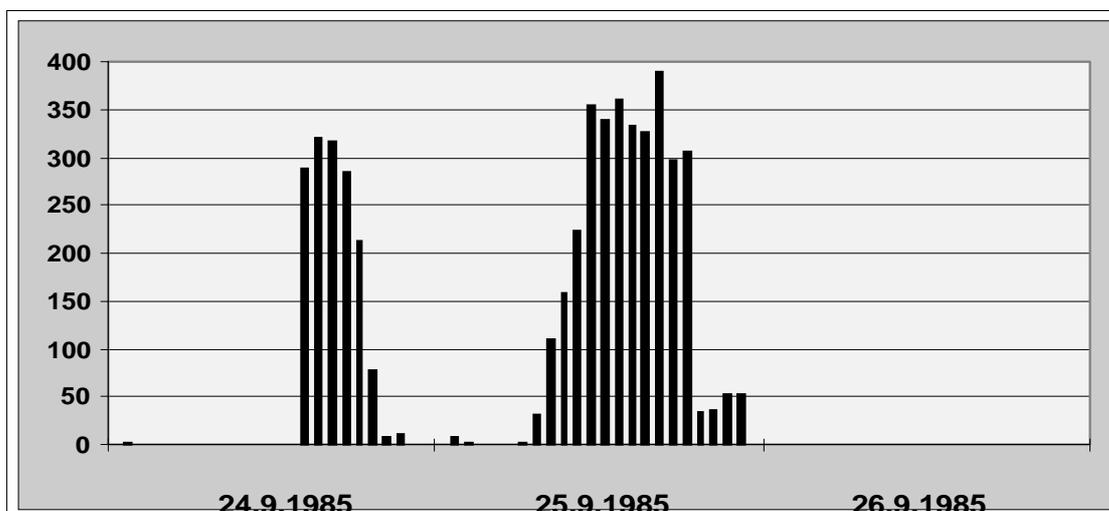


Abb. 44) Nicht plausible Werte an der Station EGGENSTEIN (BW004), bereits in der 1. Stufe als fragwürdig genannt

Mehrere fragwürdige Messwerte werden bei der Station **STUTTGART-HAFEN** (BW014) festgestellt. Das Datenkollektiv dieser Station ist jedoch bereits in der 1. Stufe der Validierung als nicht plausibel bewertet worden (ohne Abbildung). Das gleiche gilt für die Station **REUTLINGEN** (BW027), bei der die 2. Stufe der Validierung einen Wert am 26. Juli bemängelt (ohne Abbildung).

Ohne Abbildung: Das Datenkollektiv der Station STUTTGART-HAFEN (BW014) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

Ohne Abbildung: Das Datenkollektiv der Station REUTLINGEN (BW027) wurde bereits in der 1. Stufe als insgesamt nicht plausibel eingestuft

In Nordrhein-Westfalen meldet die Station **SICKINGMÜHLE** (NW015) am 07. August einen Messwert, der als fragwürdig angesehen wird (ohne Abbildung). Die Messwerte dieser und der übrigen Stationen aus dem Messnetz von Nordrhein-Westfalen wurden bei der visuellen Überprüfung jedoch nach einem Ausfall im Monat Juni insgesamt als fragwürdig - weil zu niedrig - eingestuft, so dass dieser eine Messwert, der in der 2. Stufe mit „nur“ $199\mu\text{g}/\text{m}^3$ als zu hoch (!) bewertet wird, keiner zusätzlichen Bewertung mehr bedarf.

Ohne Abbildung: Das Datenkollektiv der Station SICKINGMÜHLE (NW015) wurde bereits in der 1. Stufe als nicht plausibel eingestuft

In Rheinland-Pfalz meldet die Station **WESTEIFEL** (RP015) am 06.07. einen Messwert von $220\mu\text{g}/\text{m}^3$, der als fragwürdig bezeichnet wird (Abb.45). Nach den relativ hohen Werten des Vortages ragt eine Gruppe von 4 Messwerten in den Morgenstunden des 06.07. aus einem Konzentrationsniveau von etwa $150\mu\text{g}/\text{m}^3$ heraus, was nicht unbedingt als plausibel zu bewerten ist. Insgesamt sind die Messwerte jedoch nicht so bemerkenswert, so dass sie noch als plausibel eingestuft werden können.

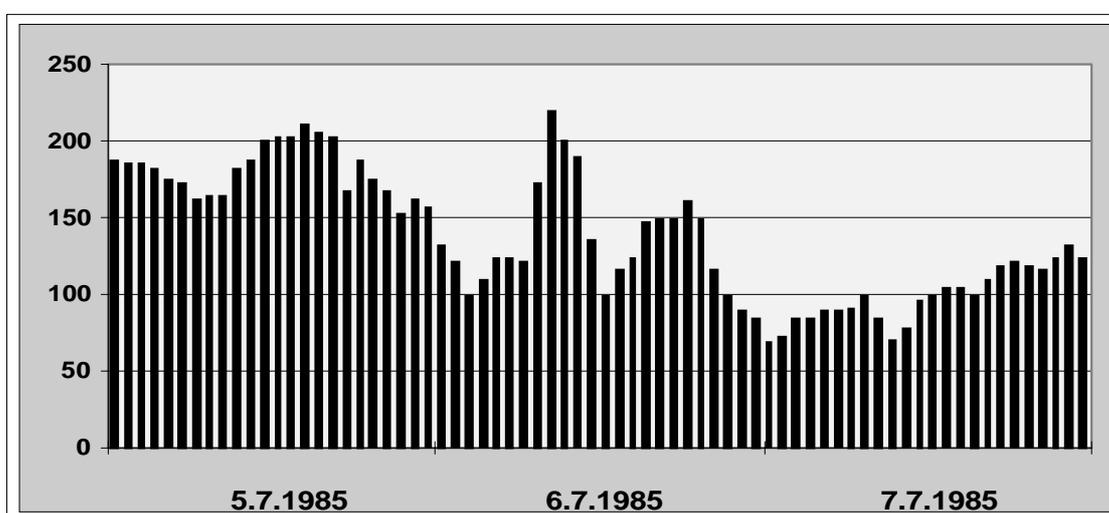


Abb. 45) Noch plausible Werte an der Station WESTEIFEL (RP015)

Messjahr 1986

In diesem wie auch in den folgenden Jahren werden nur noch wenige Messstationen aus Baden-Württemberg mit fraglichen Werten herausgefiltert. Es scheint - jedenfalls als Ergebnis der 1. und 2. Validierungsstufe - fast so, als hätten Änderungen im Analysenverfahren (z.B. andere Messgeräte) für eine „Verbesserung“ der Messergebnisse gesorgt.

So meldet die Station **EGGENSTEIN** (BW004) lediglich am 03.07.1986 einen Messwert (Abb.46), der in der 2. Stufe der Validierung auffällig wurde, während das gesamte Datenkollektiv dieser Station in der 1. Stufe als plausibel betrachtet wurde.

die Mittagszeit des folgenden Tages nicht plausibel ist, muss wohl nicht diskutiert werden.

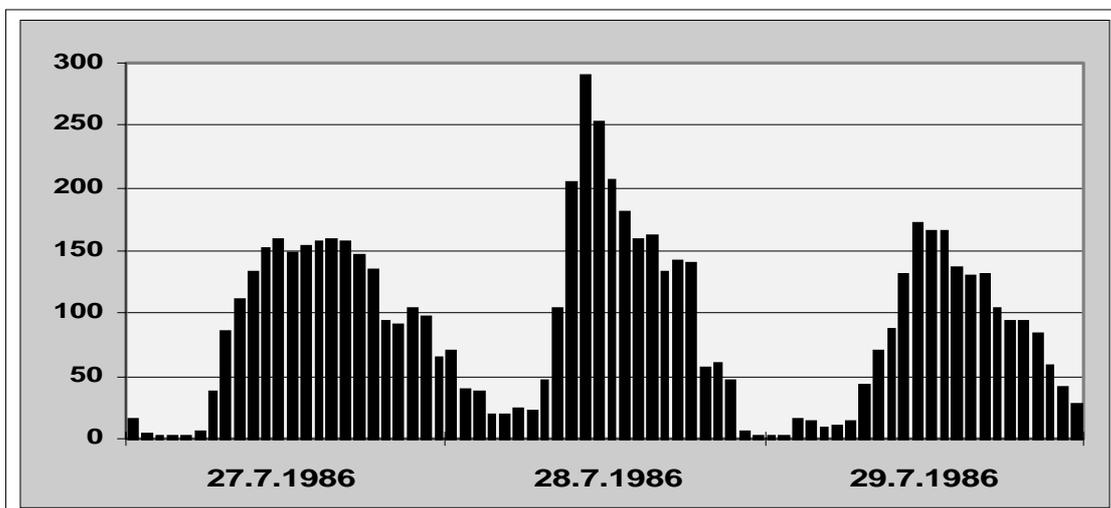


Abb. 47) Ausreißer oder HÜRTH-EFFEKT? Fragwürdige, aber noch plausible Werte in KEHL-HAFEN (BW022)

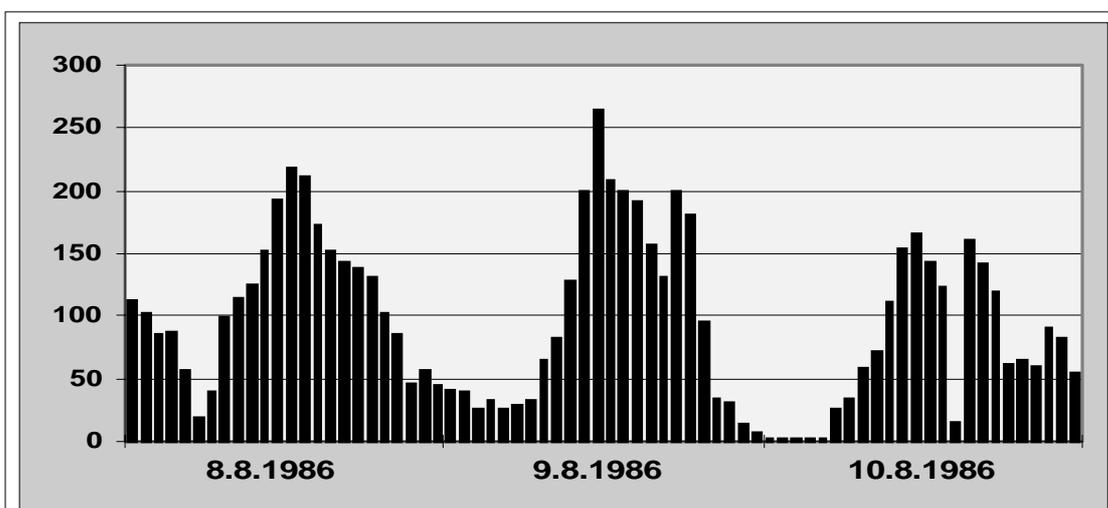


Abb. 48) Ausreißer oder HÜRTH-EFFEKT? Fragwürdige, aber noch plausible Werte in KEHL-HAFEN (BW022)

Zwei Messstationen in Bayern melden im Wintermonat Februar einige Messwerte, an deren Plausibilität gezweifelt werden kann. Die Station ARZBERG (BY002) verzeichnet am 02.02., 04.02. und 15.02. jeweils einen Messwert, der $200\mu\text{g}/\text{m}^3$ deutlich

übersteigt. Zu diesen Zeitpunkten betragen die Maxima der übrigen Messstationen um etwa $100\mu\text{g}/\text{m}^3$. Die Abb.49 zeigt die ersten beiden in der zweiten Validierungsstufe kritisierten Tage. Es ist wohl offensichtlich, dass sowohl die beanstandeten als auch weitere Messwerte in deren zeitlicher Umgebung als Ausreißer (vielleicht Kalibrierwerte) anzusehen und nicht plausibel sind. Ebenfalls nicht plausibel ist der in der zweiten Validierungsstufe als fragwürdig bezeichnete Messwert am 15.02.1986 an der Station **ARZBERG**, den die Abb.50 zeigt. Auch bei dem beanstandeten Wert könnte es sich infolge der Höhe des Messwertes von $255\mu\text{g}/\text{m}^3$ ebenfalls um einen Kalibrierwert handeln. In diesem Fall hätte sich die Kalibrierung des Gerätes auch auf die zeitliche Umgebung des Kalibrierintervalles ausgewirkt, in der auch nicht plausible Messwerte beobachtet werden.

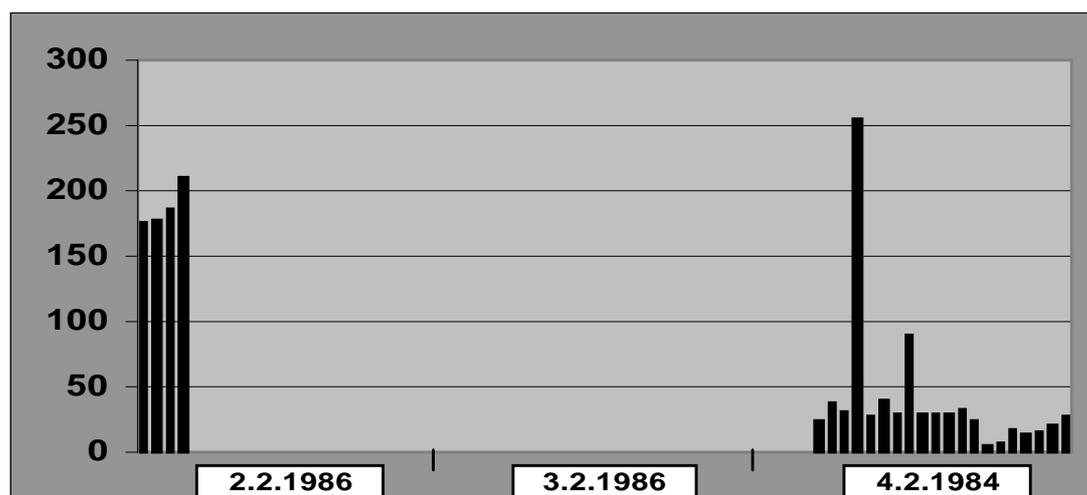


Abb. 49) Ausreißer oder Kalibrierwerte? Nicht plausible Werte an der Station ARZBERG (BY002), 2. und 4.2.1986

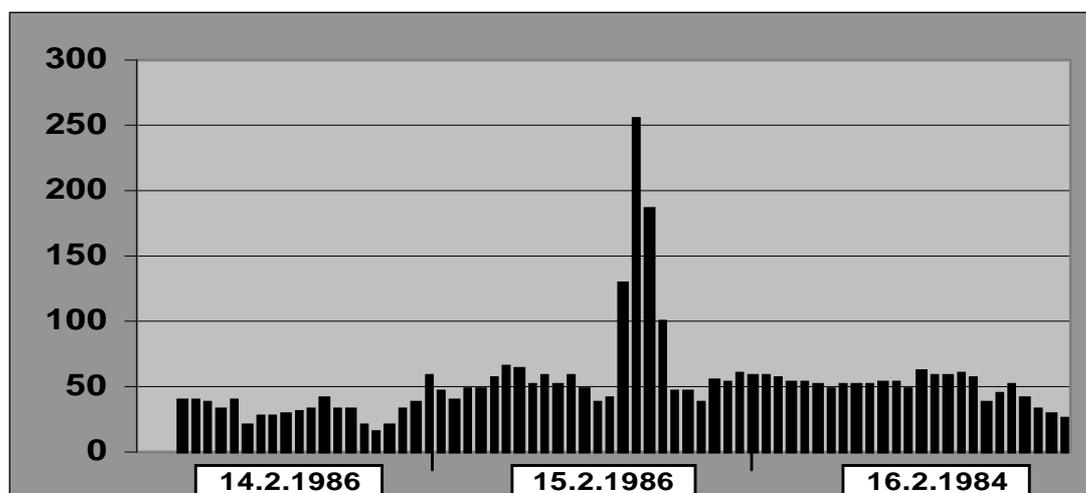


Abb. 50) Ausreißer oder Kalibrierwerte? Nicht plausible Werte an der Station ARZBERG (BY002), 15.2.1986

An der bayerischen Messstation **HAUNSTETTEN** (Augsburg, BY008) werden ebenfalls im Februar (01.02.1986) zwei fragwürdige Messwerte von etwa $190\mu\text{g}/\text{m}^3$ festgestellt (Abb.51). In Anbetracht der Jahreszeit ist es naheliegend, die Messwerte etwa ab den Mittagsstunden des 31.01. bis zu den Mittagsstunden des 01.02. als zu hoch einzustufen, obwohl anhand der Daten eine Ursache für die trendähnliche Zunahme der Messwerte in den Nachtstunden nicht erkennbar ist. Es wird empfohlen, die fraglichen Werte als unplausibel anzusehen und aus dem Datenkollektiv zu entfernen.

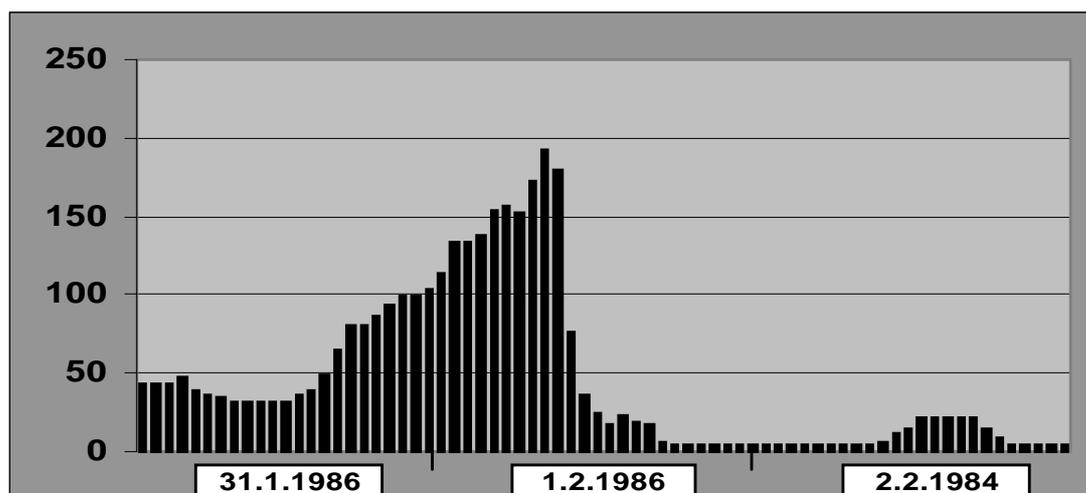


Abb. 51) Nicht plausible Werte an der Station HAUNSTETTEN (BY008) im Wintermonat Februar

Die Messstation **ZUGSPITZE** (BY083) registriert am 19.07.1986 um 1 Uhr nachts einen relativ niedrigen Messwert von $164\mu\text{g}/\text{m}^3$ (ohne Abbildung), der die Maxima der übrigen Messstationen jedoch deutlich überschreitet. Da es sich bei dieser Station jedoch um eine Bergstation mit einem nur schwach ausgeprägten Tagesgang handelt, ist dieser Messwert als plausibel einzustufen.

Ohne Abbildung: Ein gerade noch als plausibel eingestufter Messwert an der Station ZUGSPITZE (BY083)

Mehrere fragliche Messwerte werden in diesem Jahr von der Messstation **TRIER** (RP020) in Rheinland-Pfalz in den Monaten April, Mai und Juni gemeldet. Diese Messwerte liegen zum Teil weit über $200\mu\text{g}/\text{m}^3$; die Maxima der übrigen Messstationen erreichen dagegen nur Werte zwischen $100\mu\text{g}/\text{m}^3$ und $150\mu\text{g}/\text{m}^3$. Die ersten beiden fragwürdigen Messwerte werden an dieser Station mit $223\mu\text{g}/\text{m}^3$ am 04.04. und 05.04. beobachtet (Abb.52). An diesen Tagen liegen die Maxima der übrigen Messstationen bei etwa $100\mu\text{g}/\text{m}^3$. Am 04.04. ragen zwei Messwerte über $200\mu\text{g}/\text{m}^3$ aus dem abendlichen Rückgang der Ozonbelastung als Ausreißer heraus. Das gleiche gilt für vier Messwerte, die am 05.04. aus dem morgendlichen Anstieg der Ozonkonzentration als Spitze herausragen.

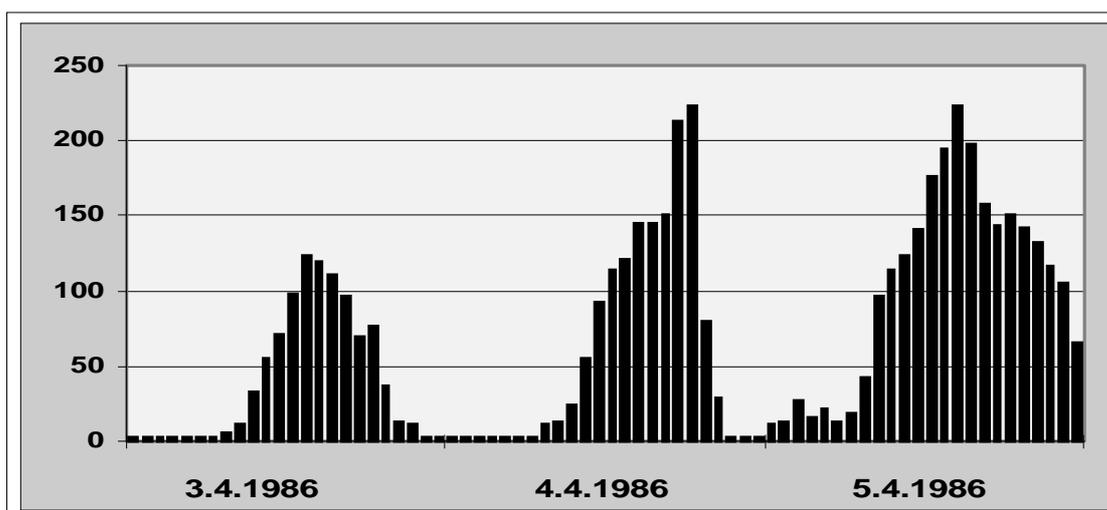


Abb. 52) HÜRTH-EFFEKT an der Station TRIER (RP020). Die Werte werden noch als plausibel angesehen

Beide Fälle könnten als **HÜRTH-EFFEKT** betrachtet werden. Es ist allerdings nicht bekannt, dass dieser Effekt auch an dieser Station beobachtet wird. Wir würden vorschlagen, diese Werte noch als plausibel anzusehen. Am 30.06. werden in den Nachmittagsstunden drei Messwerte mitgeteilt, die Konzentrationen zwischen $190\mu\text{g}/\text{m}^3$ und $220\mu\text{g}/\text{m}^3$ erreichen (ohne Abbildung). Dagegen melden die übrigen Messstationen Maxima der Ozonkonzentration, die nur selten $100\mu\text{g}/\text{m}^3$ überschreiten. Der gleichmäßige Verlauf der Ozonkonzentration an diesem Tag lässt die fraglichen Werte jedoch als plausibel erscheinen. Am 06. Juni werden zwischen 14 Uhr und 19 Uhr mehrere Messwerte gemeldet, von denen ein Wert bis nahezu an $300\mu\text{g}/\text{m}^3$ heranreicht. Dieser Messwert stellt sich als Ausreißer aus einem recht plausiblen Verlauf der Messwerte an den in der Abb.53 dargestellten Tagen dar. Es könnte sich hierbei erneut um den **HÜRTH-EFFEKT** handeln, obwohl es fraglich ist, ob dieser Effekt tatsächlich an der Messstation in Trier beobachtet werden kann. Die genannten Messwerte sollten trotz einiger Bedenken noch als plausibel angesehen werden. Wenige Tage später, nämlich am 09.06., werden an der Station **TRIER** erneut zwei Werte beobachtet, die die Maxima der übrigen Stationen um $100\mu\text{g}/\text{m}^3$ überschreiten. Diese Messwerte werden in den Abendstunden um 19 Uhr und 20 Uhr gemessen (Abb.54). Beide Messwerte sowie ein vorhergehender dritter Wert von $247\mu\text{g}/\text{m}^3$, der jedoch nicht als fragwürdig angesehen wird, ragen erneut als Spitze aus einem plausiblen Verlauf der Ozonkonzentration heraus. Erneut fällt die Entscheidung, ob es sich hierbei um den bereits häufiger genannten **HÜRTH-EFFEKT** handelt oder nicht, sehr schwer. Wir würden diese Werte zunächst als plausibel betrachten. Eventuell könnten sich in der 3. Validierungsstufe zusätzliche Hinweise ergeben, die eine Entscheidung erleichtern.

Ohne Abbildung: Ein gerade noch als plausibel eingestufte Messwert im Juni an der Station TRIER (RP020)

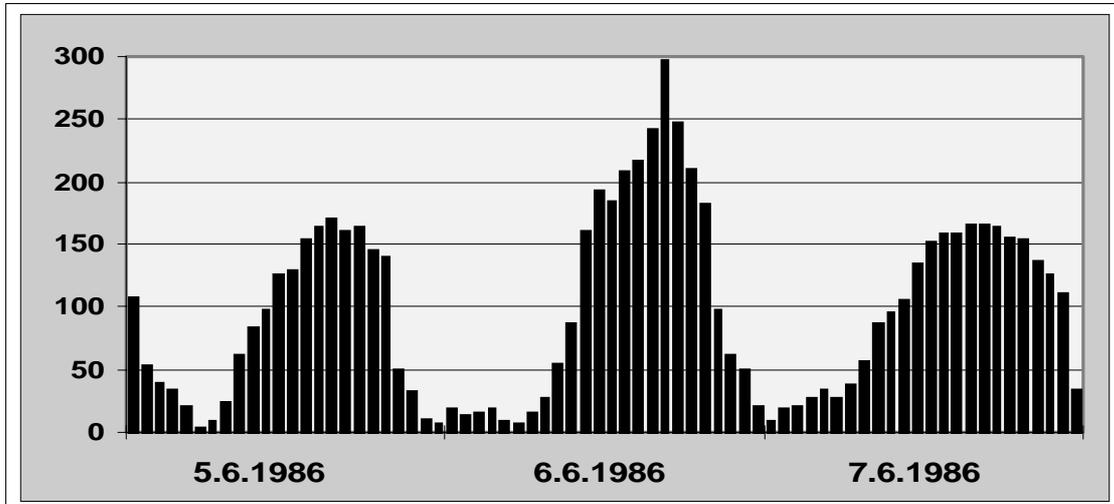


Abb. 53) HÜRTH-EFFEKT an der Station TRIER (RP020). Der Wert wird noch als plausibel angesehen

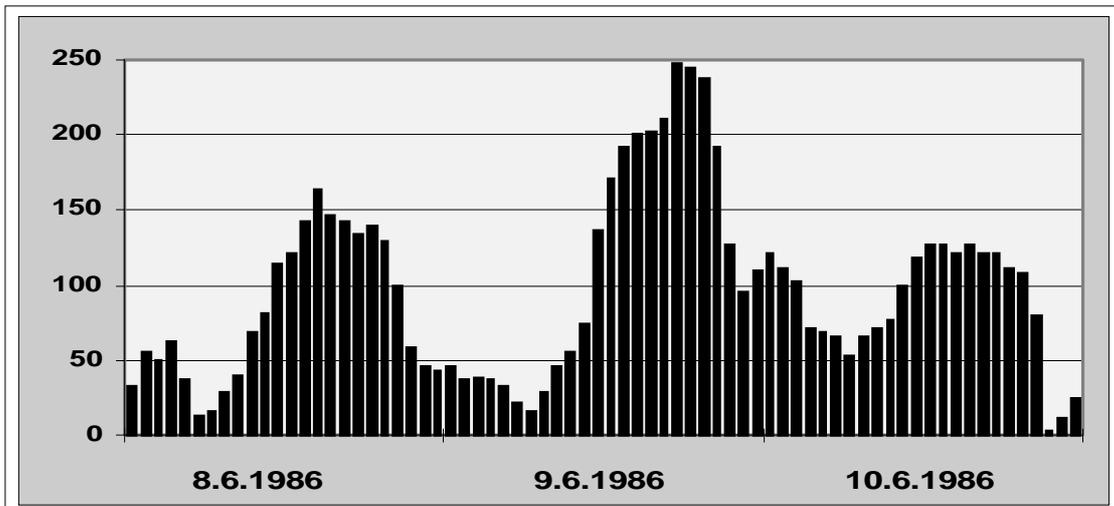


Abb. 54) HÜRTH-EFFEKT an der Station TRIER (RP020). Die Werte werden noch als plausibel angesehen

Validierung weitere Informationen hierüber bereitstellt, die eine Entscheidung bezüglich der Plausibilität dieser Daten ermöglichen.

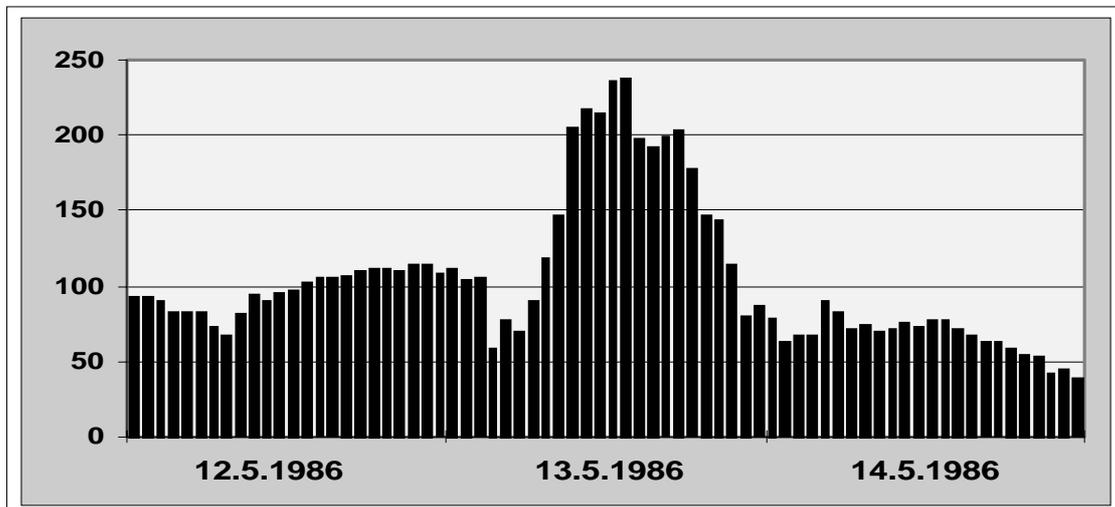


Abb. 56) Mehrere unplausible Werte an der Station SCHMÜCKE (UB029) im Mai 1986

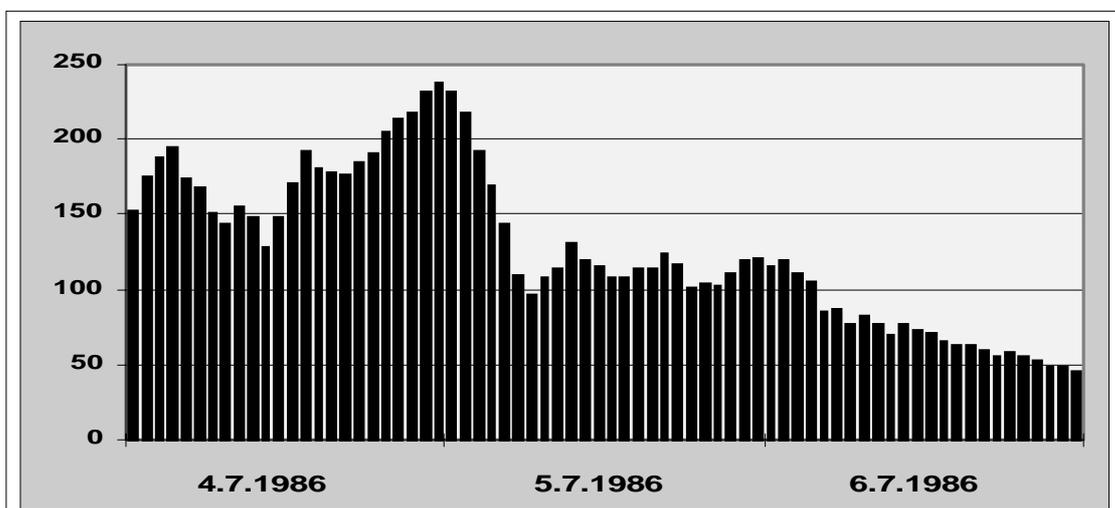


Abb. 57) Weitere unplausible Werte an der Station SCHMÜCKE (UB029)

Messjahr 1987

Nur eine Messstation aus Baden-Württemberg meldet im Jahr 1987 einige Messwerte, die in der 2. Stufe der Validierung als auffällig bezeichnet werden. Hierbei handelt es sich um die Messstation **KARLSRUHE-NORDWEST** (BW081), die am 29.06. um 7 Uhr morgens einen Messwert von $245\mu\text{g}/\text{m}^3$ meldet, der deutlich höher ist als die Maxima der übrigen Messstationen (ohne Abbildung). Dieser Messwert ist der erste Messwert nach mehreren Ausfallwerten; anschließend folgen für diese Tageszeit „normale“ Werte von etwa $50\mu\text{g}/\text{m}^3$. Der fragliche Wert ist offensichtlich nicht plausibel. Um 10 Uhr und 12 Uhr werden am 11.07.1987 zwei Werte von $273\mu\text{g}/\text{m}^3$ bzw. $308\mu\text{g}/\text{m}^3$ gemeldet, die ebenfalls deutlich über den Maxima der übrigen Messstationen liegen (Abb.58). Bei diesen Messwerten könnte es sich erneut um den HÜRTH-EFFEKT handeln, so dass sie als plausibel angenommen werden sollten.

Ohne Abbildung: Nicht plausibler Messwert an der Station KARLSRUHE-NORDWEST (BW081) im Juni 1987

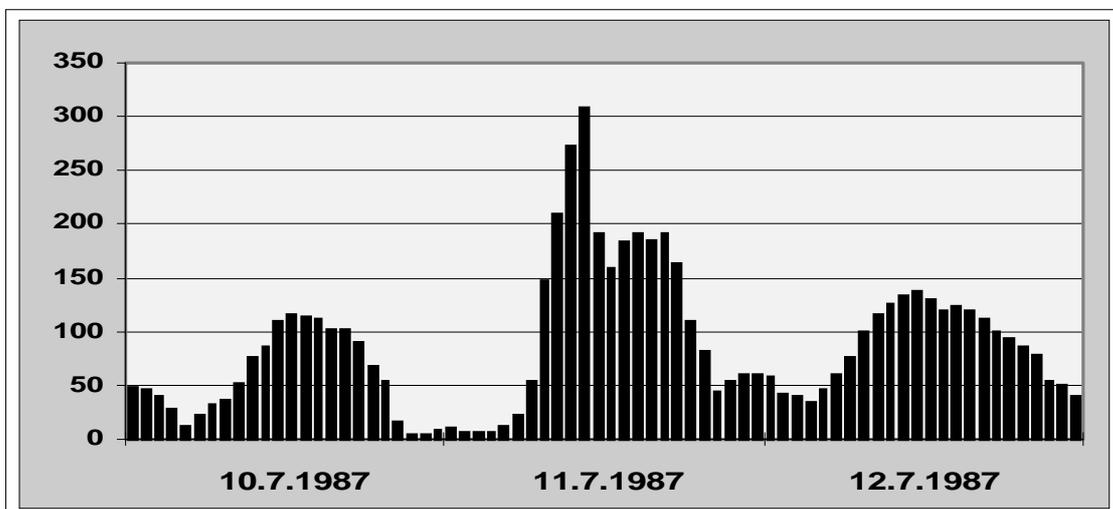


Abb. 58) HÜRTH-EFFEKT an der Station KARLSRUHE-NORDWEST (BW081)
Die Werte gelten als plausibel

Messjahr 1988

Nur ein Messwert wird von der 2. Stufe der Validierung im Sommer 1988 als fragwürdig bezeichnet. Hierbei handelt es sich um einen Wert von $318\mu\text{g}/\text{m}^3$, der am 11. August um 2 Uhr morgens an der Station **PFÄLZERWALD** (RP017) in Rheinland Pfalz beobachtet wurde. In der 1. Stufe der Validierung wurden die Messwerte dieser Station als „sehr hoch“ und damit als fragwürdig bezeichnet. Die Darstellung dieser Werte in Abb.59 zeigt, dass sie tatsächlich als nicht plausibel bezeichnet werden müssen. Dies trifft nicht allein auf die Werte des 11.08. (darunter auch der bemängelte Ausreißer in den Nachtstunden) zu, sondern auch auf die vorhergehenden Tage. Möglicherweise wird auch hier die dritte Validierungsstufe weitere Informationen bezüglich der Plausibilität dieser Daten liefern. Bis dahin sollten die Werte noch als plausibel angesehen werden, auch wenn die in der ersten und zweiten Stufe beanstandeten Daten tatsächlich wenig glaubwürdig sind.

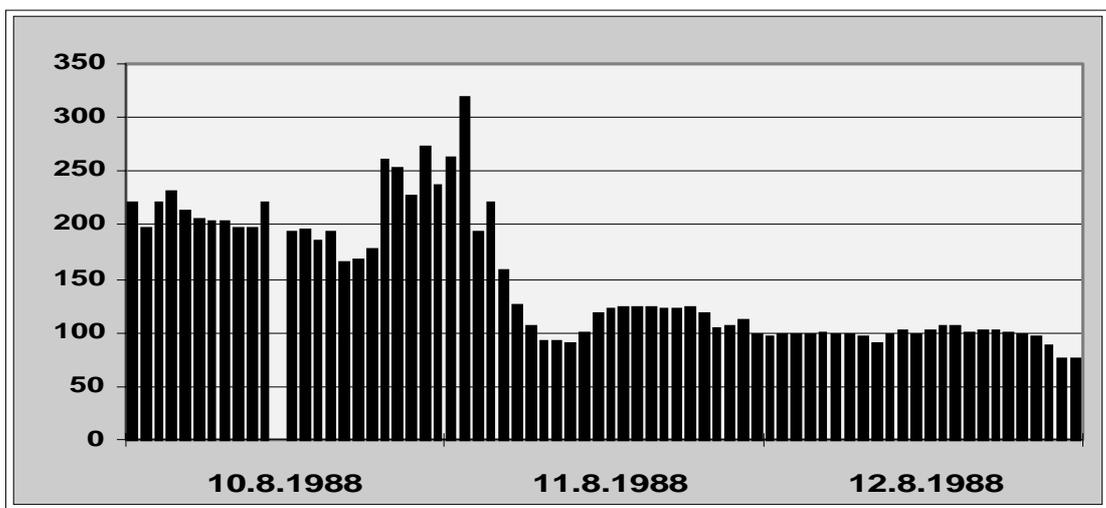


Abb. 59) Plausibel oder nicht plausibel? Fragwürdige Werte an der Station PFÄLZERWALD (RP017)

Im Wintermonat Februar werden von der bayerischen Bergstation **GAP-ZUGSPITZE** (BY083) zwei Messwerte berichtet, die mit etwa $190\mu\text{g}/\text{m}^3$ in den beiden Mitternachtsstunden vom 12.02. und 13.02. um knapp $100\mu\text{g}/\text{m}^3$ über den Maxima der übrigen Stationen liegen. Diese beiden Werte sowie der vorhergehende und nachfolgende Wert ragen deutlich aus einem recht hohen Niveau der Ozonkonzentration

als Ausreißer heraus und erscheinen in der Abb.60 tatsächlich als wenig plausibel. Diese vier Messwerte sollten aus dem Datenkollektiv entfernt werden.

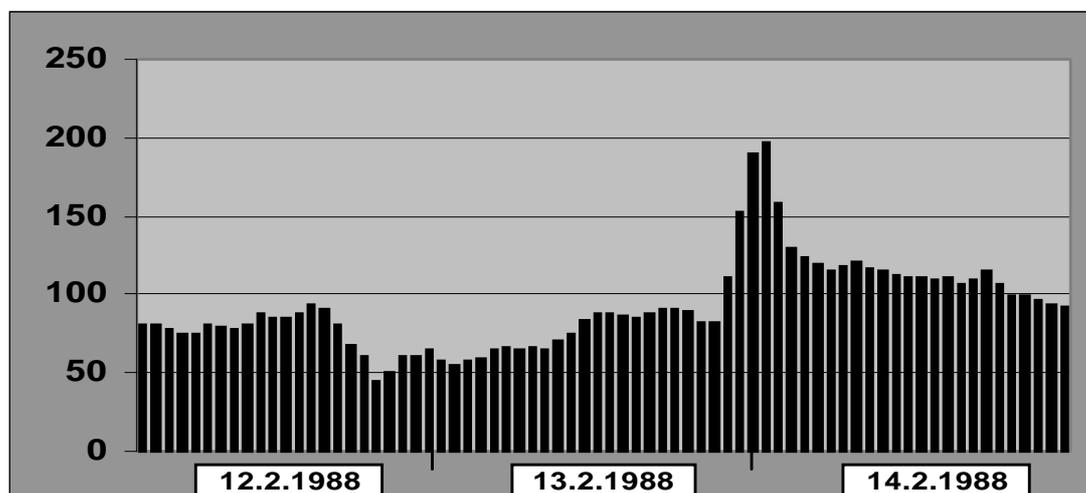


Abb. 60) Nicht plausible Messwerte an der Station ZUGSPITZE (BY083) im Wintermonat Februar

Messjahr 1989

Im Jahr 1989 werden Messwerte aus mehreren Messnetzen als fragwürdig eingestuft. Hierbei ist das Messnetz in Rheinland-Pfalz mit 3 und das Messnetz des Umweltbundesamtes mit 4 Messstationen (3 dieser Messstationen wurden in diesem Jahr noch vom Meteorologischen Dienst der DDR betrieben) vertreten.

Im Wintermonat Februar ist (zum ersten Mal) eine Messstation in Niedersachsen als auffällig herausgefiltert worden, die um 22 Uhr am 17.02.1989 einen Messwert von $235\mu\text{g}/\text{m}^3$ meldet. Dieser Messwert, der an der Messstation **NORDENHAM** (NI036) beobachtet wurde, stellt sich in Abb.61 als krasser Ausreißer dar und muss aus dem Datenkollektiv entfernt werden. Zu entscheiden ist, ob die diesem Wert benachbarten Daten ebenfalls als unplausibel zu verwerfen sind, auch wenn sie von der Höhe der Messwerte von etwas unter $100\mu\text{g}/\text{m}^3$ kaum in der Lage sind, spätere Auswertungen ernsthaft zu beeinflussen.

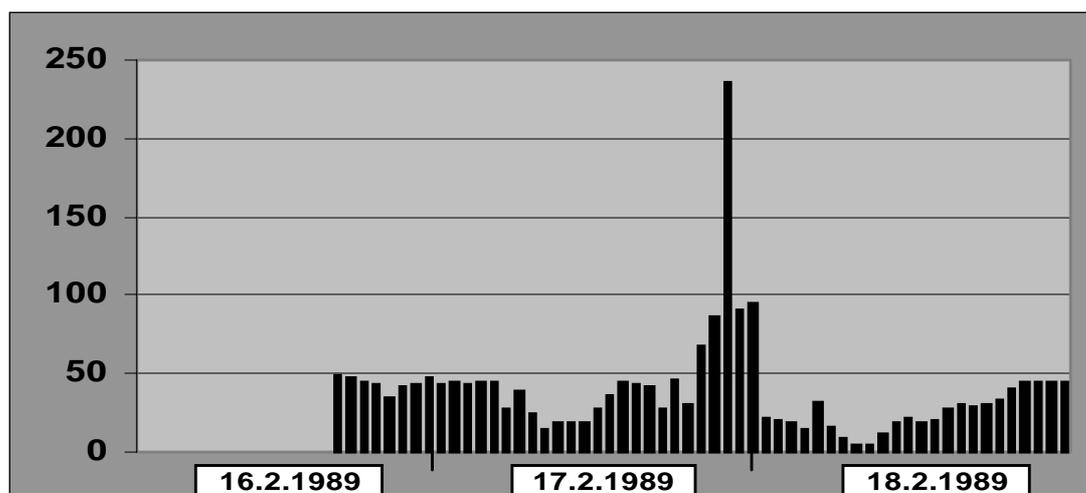


Abb. 61) Nicht plausibler Messwert an der Station NORDENHAM (NI036) im Wintermonat Februar

Erneut ist die Messstation **SCHMÜCKE** (UB029) mit auffälligen Werten herausgefiltert, die in dem Datenkollektiv dieser Station im Wintermonat Dezember enthalten sind. Diese Daten (01.12. und 02.12.1989) sind in der Graphik in Abb.62 in ihrem zeitlichen Umfeld dargestellt. Offensichtlich handelt es sich erneut um (nach der Wende) fehlerhaft geschätzte Ersatzwerte, die eine ursprünglich vorhandene Datenlücke schließen sollen. Es ist schon erstaunlich, dass solche Schätzwerte aufgrund einer fehlerhaften Berechnung (was ja durchaus einmal vorkommen kann) nicht erkannt werden und im Datenkollektiv bleiben. Immerhin erreicht der höchste dieser Schätzwerte eine bemerkenswerte "Konzentration" von $531\mu\text{g}/\text{m}^3$. Es bedarf keiner Diskussion, dass die Daten als nicht plausibel einzustufen und auf "Ausfall" zu setzen sind.

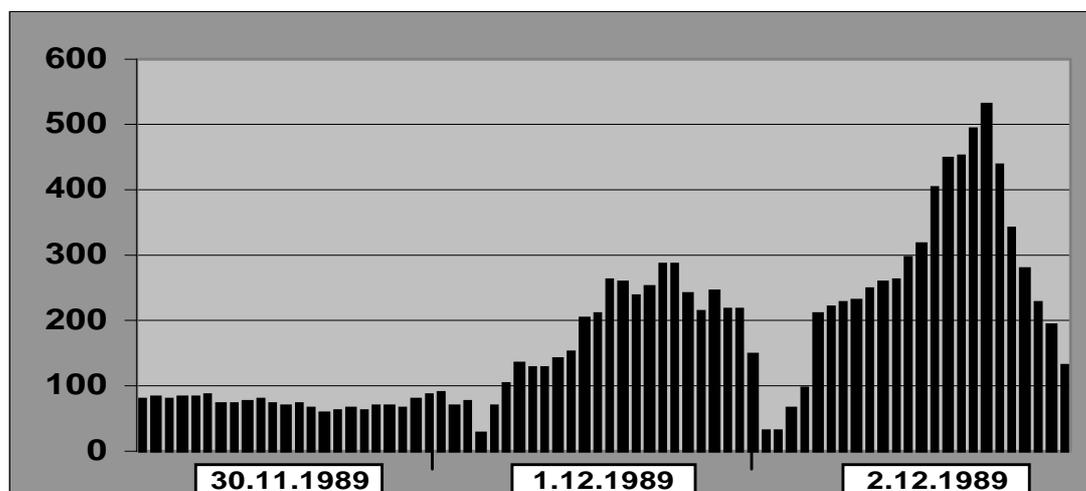


Abb. 62) Fehlerhaft berechnete Schätzwerte an der Station SCHMÜCKE (UB029)

Eine weitere Messstation (NEUGLOBSOW, UB030) des ehemaligen Meteorologischen Dienstes der DDR, die heute dem Messnetz des Umweltbundesamtes angehört, zeigt im Wintermonat November offensichtlich den gleichen Fehler, der bereits bei der Messstation **SCHMÜCKE** bemängelt wurde. Die Messwerte ab den Mittagsstunden des 26.11. bis zur Stufe in den frühen Morgenstunden des 27.11. sind offensichtlich fehlerhaft geschätzte Erwartungswerte und nicht plausibel und müssen aus dem Datenkollektiv der Station entfernt werden.

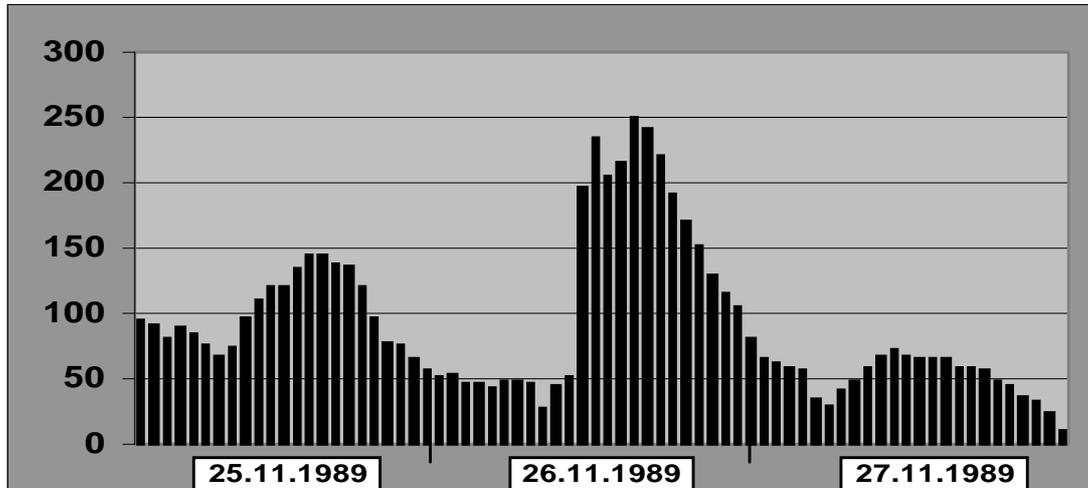


Abb. 63) Fehlerhaft berechnete Schätzwerte an der Station NEUGLOBSOW (UB030)

Die Station **LUDWIGSHAFEN-MUNDENHEIM (RP003)** meldet am 02.07. einen Messwert von $208\mu\text{g}/\text{m}^3$, der die Maxima der übrigen Messstationen um mehr als $100\mu\text{g}/\text{m}^3$ überschreitet. Die Abb.64 zeigt jedoch während dieser Tage im wesentlichen plausible und ausreißerfreie Tagesgänge der Ozonkonzentration, so dass der kritisierte Messwert dieses Tages nicht verworfen und als plausibel angesehen werden sollte.

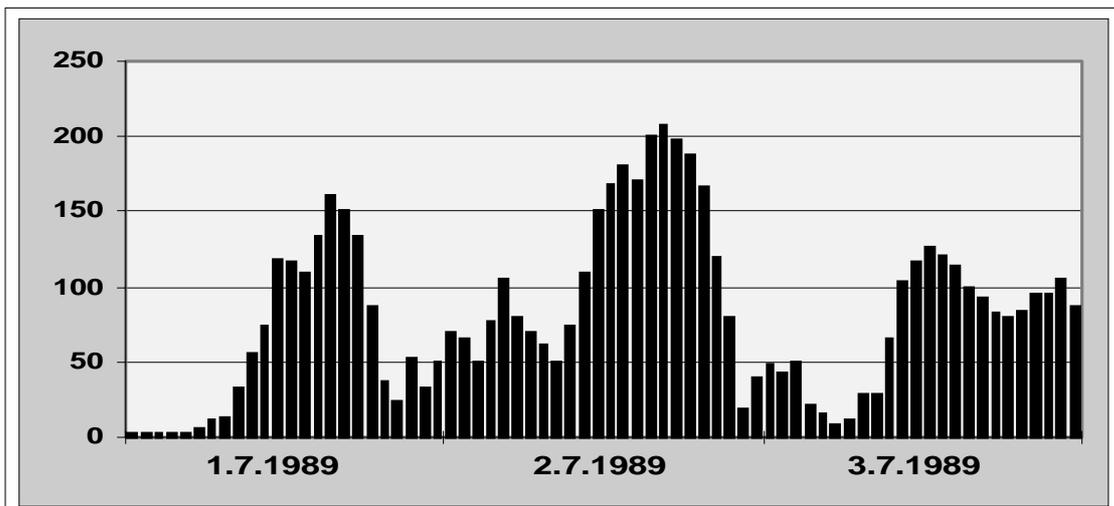


Abb. 64) Station LH-MUNDENHEIM (RP003): Der kritisierte Messwert ist als plausibel einzustufen

Die Messstation **WESTERWALD-HERDORF** (RP016) meldet am 22.07. einen Messwert von $273\mu\text{g}/\text{m}^3$, der als fragwürdig dargestellt wird. In der ersten Stufe der Validierung wurde das Datenkollektiv dieser Station als plausibel angesehen. Auch die Abb.65 zeigt keine Auffälligkeiten, die an der Plausibilität der Messwerte zweifeln lässt. Der Messwert ist mithin als plausibel anzusehen und sollte im Datenkollektiv dieser Messstation verbleiben.

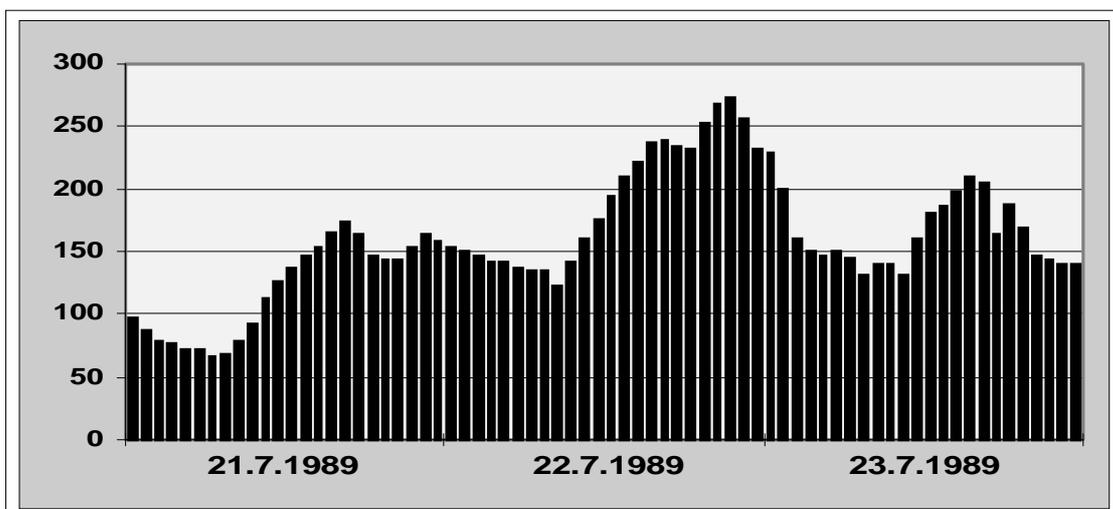


Abb. 65) Station WESTERWALD (RP016): Der kritisierte Messwert ist als plausibel einzustufen

Dagegen stellen sich die Messwerte um $200\mu\text{g}/\text{m}^3$, die am 18.04.1989 von der Station **KAISERSLAUTERN** (RP019) um die Mittagszeit gemeldet werden, deutlich als Ausreißer dar, die bereits in der ersten Stufe der Validierung als nicht plausibel angesehen wurden (ohne Abbildung). Die 2. Stufe bestätigt die Einschätzung, dass beide Werte nicht plausibel sind und aus dem Datenkollektiv der Station entfernt werden müssen.

Ohne Abbildung: Nicht plausibler Messwert an der Station KAISERSLAUTERN (RP019) im April 1989

Die UBA-Messstation **HOHENWESTEDT** (UB006) in Schleswig-Holstein meldet am 20.08. einen Spitzenwert von $351\mu\text{g}/\text{m}^3$, der die Maxima der übrigen Stationen um mehr als $100\mu\text{g}/\text{m}^3$ übersteigt (Abb.66). Dieser Messwert wurde bereits in der ersten Stufe der

Validierung als fragwürdig gekennzeichnet. Es ist jedoch nicht nur dieser eine Messwert und auch nicht nur dieser Tag, - der 20.08. -, der als fragwürdig anzusehen ist. Auch am Vortag werden einige Werte gemeldet, an deren Plausibilität Zweifel berechtigt sind. Darüber hinaus zeigt auch der folgende Tag, - der 21.08. -, einen unüblichen Tagesgang der Ozonkonzentration. Auch die Messwerte an diesem Tag müssen als nicht plausibel angesehen werden.

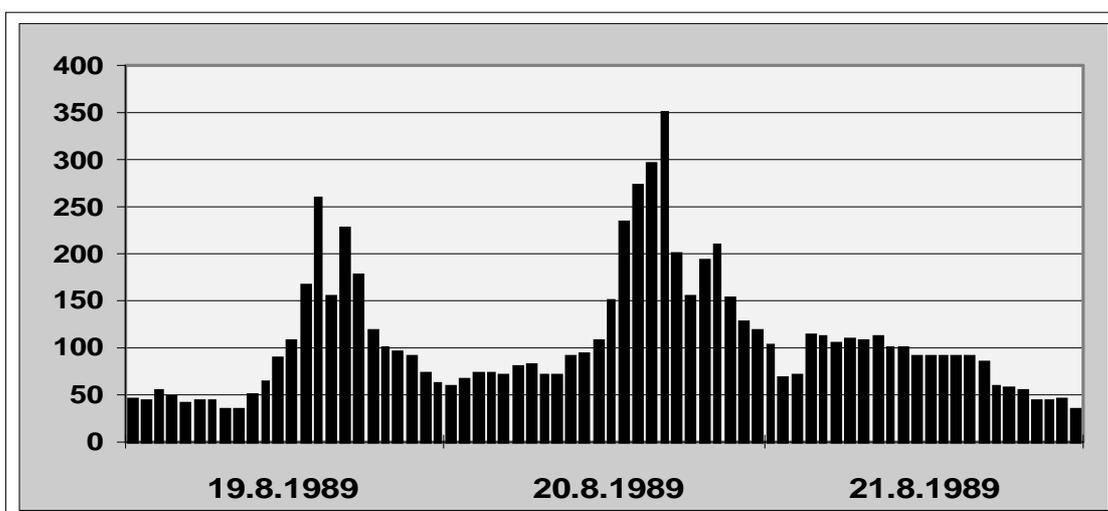


Abb. 66) Nicht plausible Messwerte an der Station HOHENWESTEDT (UB006) im August 1989

Ebenfalls im August meldet die UBA-Station **LINDENBERG** (UB025) mehrere Messwerte die sowohl in der ersten als auch in der zweiten Stufe als zweifelhaft angesehen werden müssen. Zwischen 13 Uhr und 19 Uhr werden an dieser Station Messwerte ermittelt, die weit über $200\mu\text{g}/\text{m}^3$ liegen und die die Maxima der übrigen Stationen um mehr als $100\mu\text{g}/\text{m}^3$ übersteigen. Die Abb.67 zeigt die bemängelten Messwerte, die zunächst gar nicht so unplausibel zu sein scheinen. Erst wenn man feststellt, dass 4 Messwerte zwischen 13 Uhr und 17 Uhr den gleichen Messwert, nämlich $286\mu\text{g}/\text{m}^3$ betragen, können Zweifel an der Richtigkeit der Werte aufkommen. Trotz dieser Zweifel können diese Messwerte zunächst noch als plausibel betrachtet werden - in der Hoffnung - dass eine weitere Überprüfung (z.B. die 3. Stufe der Validierung) hier die erforderliche Klarheit verschaffen kann.

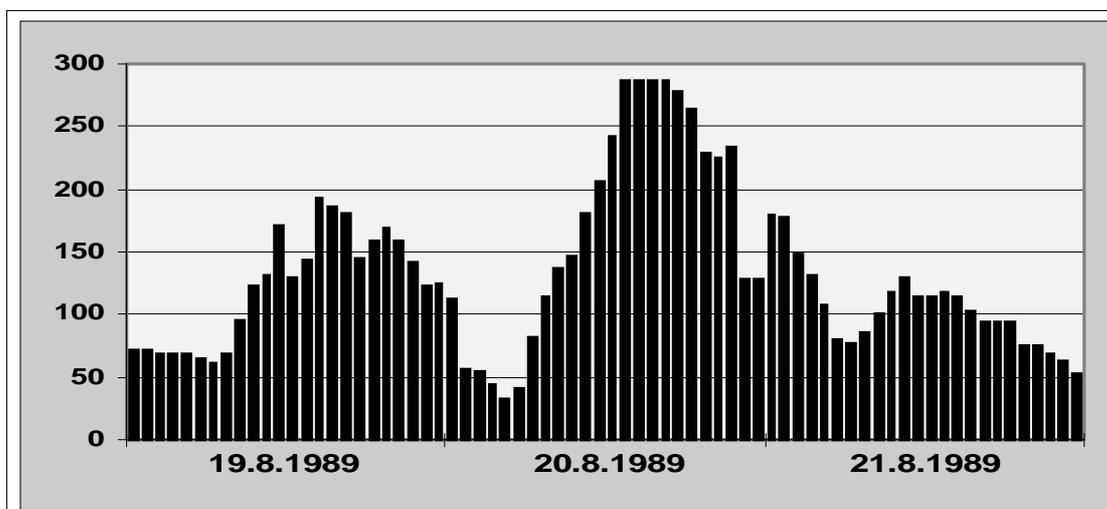


Abb. 67) Plausibel oder nicht plausibel? Fragwürdige Werte an der Station LINDENBERG (UB025)

Messjahr 1990

Im Messjahr 1990 werden - ebenso wie in den folgenden beiden Messjahren 1991 und 1992 - nur noch von einer Station in Nordrhein-Westfalen Messwerte gemeldet, die die zweite Stufe des Validierungsverfahrens als fragwürdig darstellt. Hierbei handelt es sich um die Messstation **HÜRTH** (NW058) bei Köln, bei der auch in der Vergangenheit häufig Messwerte beobachtet werden, die als Ausreißer aus dem üblichen Tagesgang der Ozonkonzentration zum Teil sehr weit herausragen. Diesen Ausreißern verdankt der **HÜRTH-EFFEKT** seinen Namen. Hierbei werden bei Messstationen in der Nachbarschaft von petrochemischen Anlagen hohe Ozonkonzentrationen beobachtet, die auf eine (vorübergehende) hohe VOC-Emission zurückzuführen sind. Offensichtlich handelt es sich auch bei den Messwerten am 17.06.1990 zwischen 10 Uhr und 13 Uhr um Ozonkonzentrationen ($323\mu\text{g}/\text{m}^3$, $267\mu\text{g}/\text{m}^3$ und $268\mu\text{g}/\text{m}^3$), die durch diesen Effekt hervorgerufen werden (Abb.68). In sehr viel schwächerer Ausführung ist dieser Effekt auch am 16.06.1990 zu beobachten, wie dies die Abb.68 zeigt. Übrigens wurden diese Werte auch in der ersten Stufe als auffällig markiert und als plausibel gekennzeichnet. Auch am 13.07.1990 ist dieser Effekt zu beobachten. An diesem Tag meldet die 2. Stufe der Validitätsprüfung einen Messwert von $255\mu\text{g}/\text{m}^3$, der um mehr als $100\mu\text{g}/\text{m}^3$ über

den Maxima der benachbarten Messstationen liegt. Die Abb.69 zeigt diesen Messwert, der mit dem darauffolgenden Wert offensichtlich erneut vom **HÜRTH-EFFEKT** geprägt ist. Mithin sollten auch diese beiden Werte als plausibel eingestuft werden. Zum drittenmal in diesem Jahr wird bei dieser Station dieser Effekt am 31.07. beobachtet (Abb.70). Hier bemängelt die 2. Stufe der Validierung einen um 13 Uhr ermittelten Wert von $308\mu\text{g}/\text{m}^3$, der deutlich über den Maxima der übrigen Messstationen liegt. Erneut handelt es sich um mehrere aufeinanderfolgende Werte in den Mittagsstunden, bei denen der Einfluss des **HÜRTH-EFFEKTES** sehr deutlich zu erkennen ist.

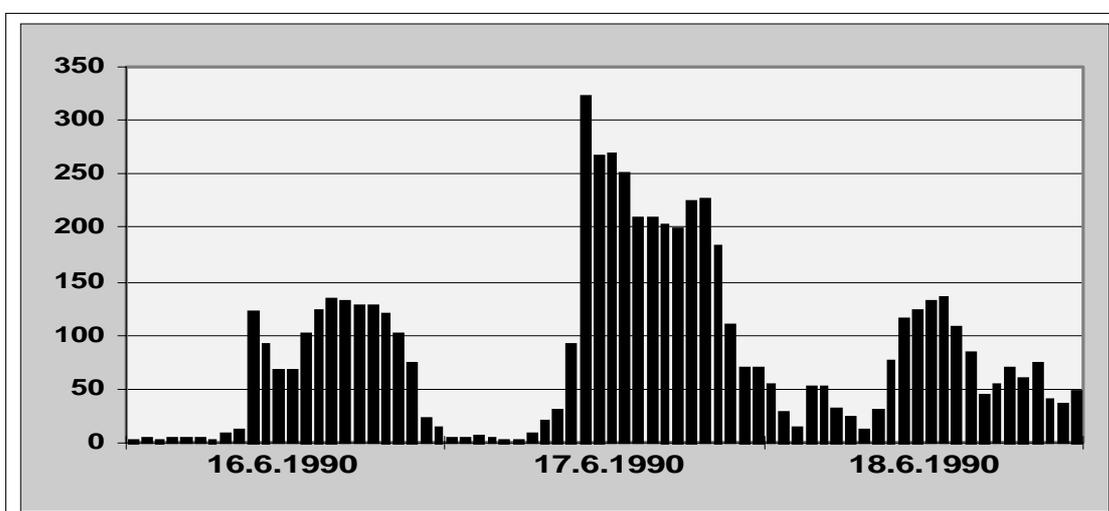


Abb. 68) Ein typisches Beispiel für den HÜRTH-EFFEKT an der Messstation HÜRTH (NW058) im Juni 1990

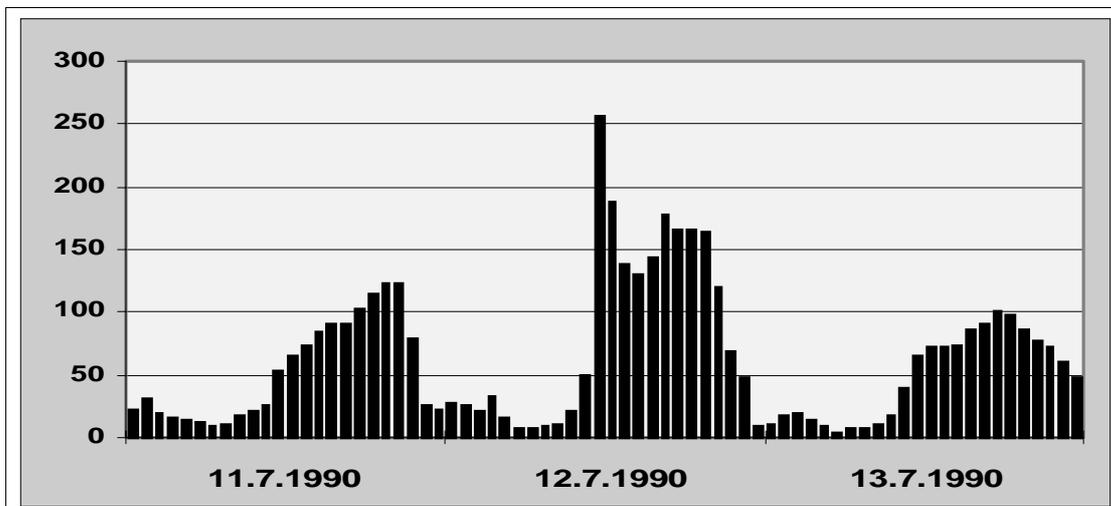


Abb. 69) Ein weiteres Beispiel für den HÜRTH-EFFEKT an der Messstation HÜRTH (NW058) im Juli 1990

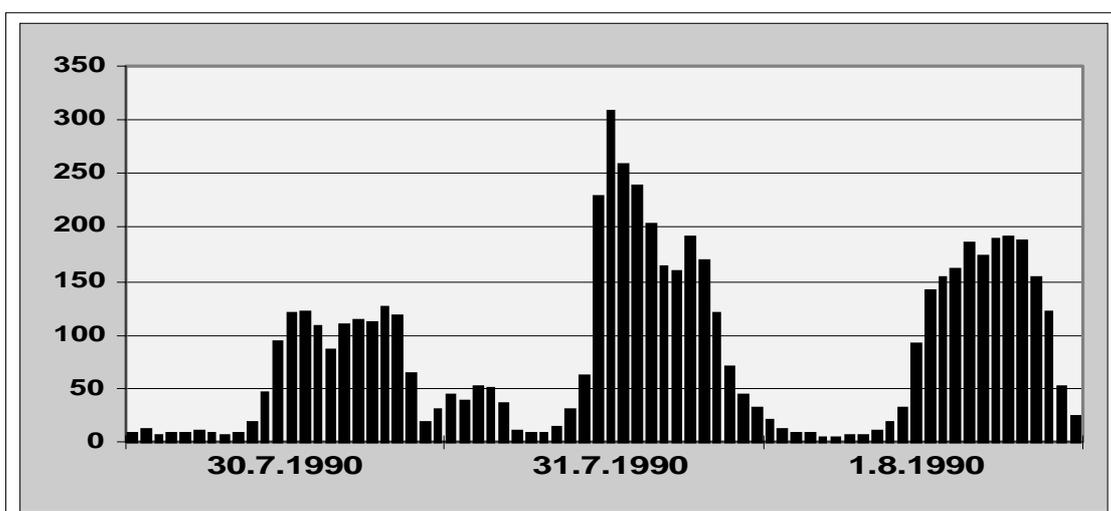


Abb. 70) Noch ein Beispiel für den HÜRTH-EFFEKT an der Messstation HÜRTH (NW058) ebenfalls im Juli 1990

Messjahr 1991

Im Jahr 1991 wird dieser Effekt erneut an der Station **HÜRTH** (NW058) im September beobachtet. Am 02.09. verzeichnet die zweite Validierungstufe einen recht hohen

Messwert von $282\mu\text{g}/\text{m}^3$, der - wie die Abb.71 zeigt - zusammen mit dem nachfolgenden Messwert, der mit etwas unter $200\mu\text{g}/\text{m}^3$ bereits deutlich niedriger ausfällt, dem **HÜRTH-EFFEKT** zuzurechnen und somit plausibel ist. Weitere Messwerte oder Messstationen sind in diesem Jahr weder in der ersten noch in der zweiten Validierungsstufe als auffällig gekennzeichnet. Offensichtlich sind die Messungen ab dem Jahr 1990 oder 1991 deutlich besser geworden; zumindest treten die krassen Ausreißer, die in diesen beiden Stufen der Plausibilitätsprüfung erkannt werden können, nicht mehr auf. Ob sich diese Aussage auch in der 3. Stufe bestätigen lässt, bleibt abzuwarten.

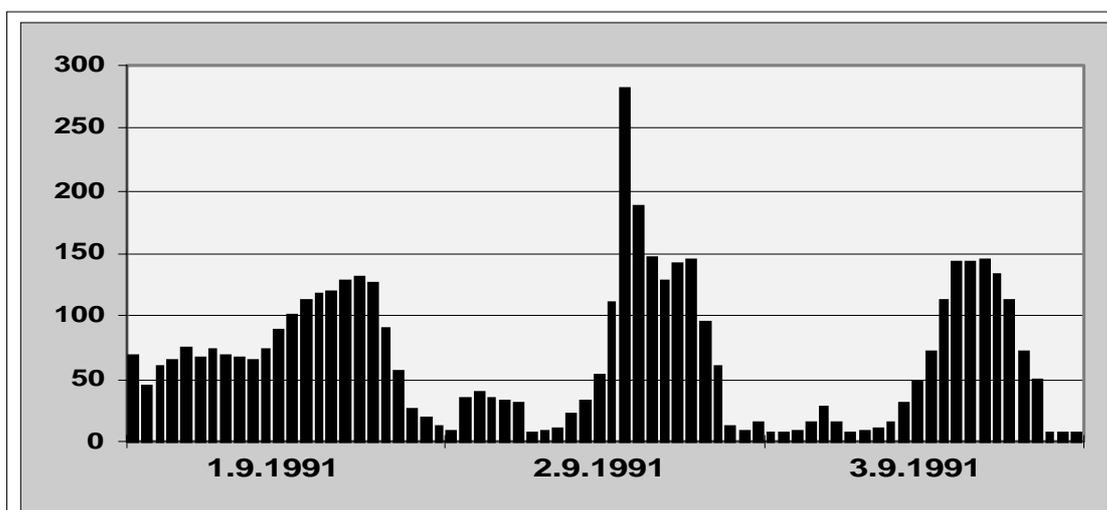
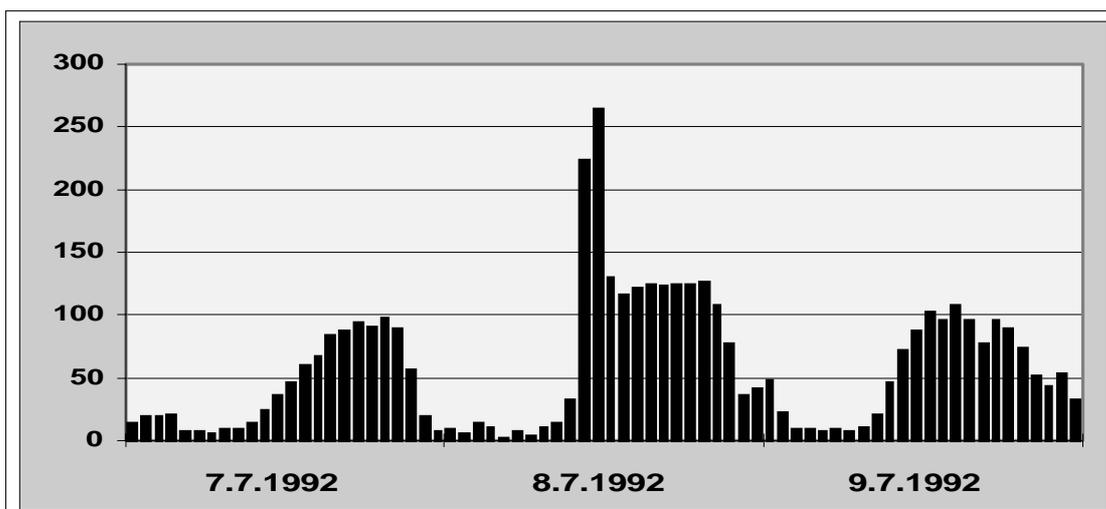


Abb. 71) Der HÜRTH-EFFEKT im September 1991 an der Messstation HÜRTH bei Köln (NW058)

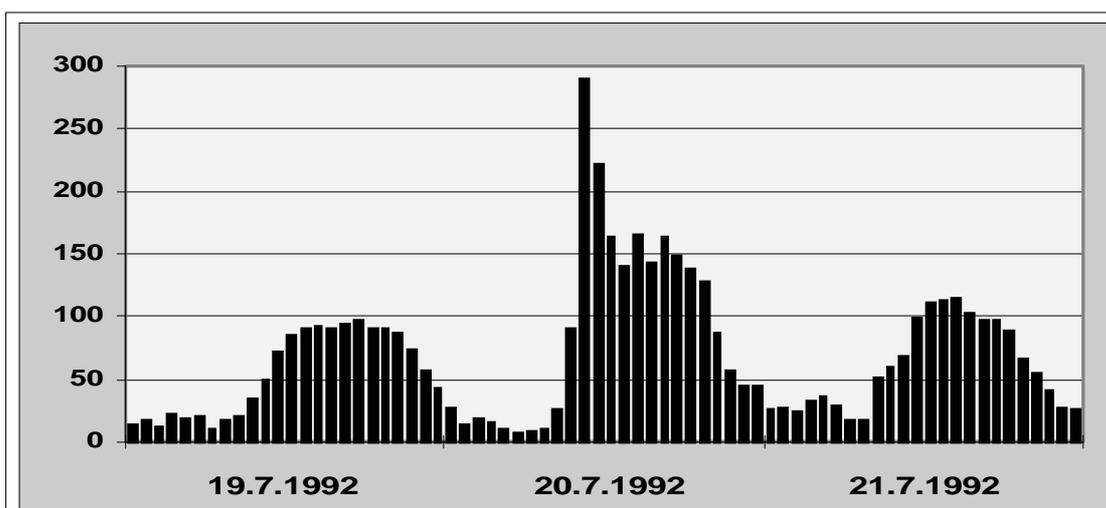
Messjahr 1992

Als einzige Messstation ist im Jahr 1992 erneut die Station **HÜRTH** (NW058) mit zwei auffälligen Messwerten vertreten, die beide im Monat Juli ermittelt wurden. Der erste Messwert wird am 08.07. um 12 Uhr mit $264\mu\text{g}/\text{m}^3$ beobachtet (Abb. 72) und ist wie die Messwerte der beiden Vorjahre dem **HÜRTH-EFFEKT** zuzuschreiben. Dieser Effekt wirkt sich offensichtlich auch auf die nachfolgenden Stunden aus, in denen der übliche Tagesgang der Ozonkonzentration deutlich abgeflacht erscheint. Auch diese Messwerte sind zusammen mit den vorhergehenden Werten als plausibel einzustufen. Das gleiche

gilt für den Messwert von $289\mu\text{g}/\text{m}^3$, der am 20.07. bei dieser Station um 11 Uhr am Vormittag festgestellt wird (Abb.73). Das Erscheinungsbild gleicht einschließlich der nachfolgenden Abflachung des üblichen Tagesganges der Ozonkonzentration erstaunlich gut dem Bild des erstgenannten Falles in diesem Monat. Auch an diesem Tag sind die bemängelten Messwerte als plausibel einzustufen



**Abb. 72) Der HÜRTH-EFFEKT im Juli 1992 an der Messstation
HÜRTH bei Köln (NW058)**



**Abb. 73) Der HÜRTH-EFFEKT im Juli 1992 an der Messstation
HÜRTH bei Köln (NW058)**

Messjahre 1993 bis 1997

Die zweite Stufe der Validierung bemängelt in diesen Jahren keinen Messwert mehr als fragwürdig. Offensichtlich tritt auch der **HÜRTH-EFFEKT** nicht mehr in der Deutlichkeit auf, dass die Methodik des Validierungsverfahrens anspricht.

3 Zeitreihenprüfung

Die Zeitreihenprüfung stellt die 3. Stufe des Validierungsprozesses dar. Sie ist sowohl hinsichtlich des Verfahrens als auch hinsichtlich der Ergebnisse am aufwendigsten. In diesem Bericht wird lediglich das Verfahren detailliert beschrieben. Da die Ergebnisse in Form von jeweils einer Abbildung pro Messstation und Untersuchungszeitraum (jeweils 10 Jahre) vorliegen, werden an dieser Stelle lediglich einige wenige Beispiele diskutiert. Ebenso beschränken sich die Darstellungen auf den Zeitraum 1990 bis 1997, da die Validierungsverfahren zunächst auf diesem Zeitraum aufgebaut waren. Mittlerweile sind die Ozondatakollektive für den gesamten Zeitraum, für die Daten verfügbar sind - nämlich 1980 bis 1997 -, validiert.

3.1 Überlegungen zur Vorgehensweise

Auch wenn ein Datenkollektiv der Ozonkonzentration von systematischen Fehlern, nicht plausiblen Werten und Ausreißern nach den ersten beiden hier dargestellten Validierungsstufen bereinigt ist, kann es in seinem zeitlichen Verlauf trotzdem Trends und Strukturbrüche aufweisen, die sowohl reale Ursachen (meteorologische Bedingungen, Änderungen der Emission der Vorläuferstoffe u.a.m.) haben als auch artifiziell (z.B. Änderungen der Probenahme- und Messtechnik) sein können. Deshalb ist es notwendig, die meteorologisch oder emissionsseitig bedingten Einflüsse nach der Beseitigung zyklischer Schwankungen (Tages- und Jahresgang) aus der Messreihe zu eliminieren. Eine Möglichkeit hierzu bietet die Ersetzung der zu untersuchenden Zeitreihe mit Erwartungswerten, die aus anderen Ozonmessreihen unter Einbeziehung meteorologischer Einflussgrößen berechnet werden. Um nicht nur die Situation eines konkreten Tages für die Interpolation zu verwenden, wie dies bei rein räumlichen Interpolationsverfahren (Shepard-Verfahren, Thin Plate Spline, Triangulation oder IDW-Verfahren) der Fall ist, sondern auch die statistischen Beziehungen zwischen

verschiedenen Beobachtungsreihen und der zu ersetzenden Reihe, wird ein stufenweises screening-Regressionsverfahren eingesetzt. Zur Auswertung wird die Differenz zwischen der insgesamt fehlwertersetzten Messreihe (Erwartungswerte) und den Originalwerten einer Station verwendet (in den Abbildungen die blaue Kurve), die Zeitreihe der Erwartungswerte selbst ist dagegen nicht dargestellt. In dieser Differenzenreihe treten durch die weitgehende Eliminierung meteorologischer und emissionsseitig bedingter Einflüsse mögliche Strukturbrüche, Trends oder Messfehler deutlicher hervor, als dies bei den Originalreihen der Fall ist.

Um die Wirksamkeit und das Verhalten des entwickelten Prüfverfahrens zu demonstrieren, wird eine normal verteilte Zufallsfolge von 1464 Werten erzeugt (Simulation von 8 Sommerhalbjahren, in der Abszisse der Abb.74 mit 1990 bis 1998 beziffert, gelbe Kurve in Abb.74). Diese Zufallsfolge wurde so skaliert, dass ihr Wertebereich einer von zyklischen und langfristigen Schwankungen bereinigten typischen Ozonreihe nahe kommt.

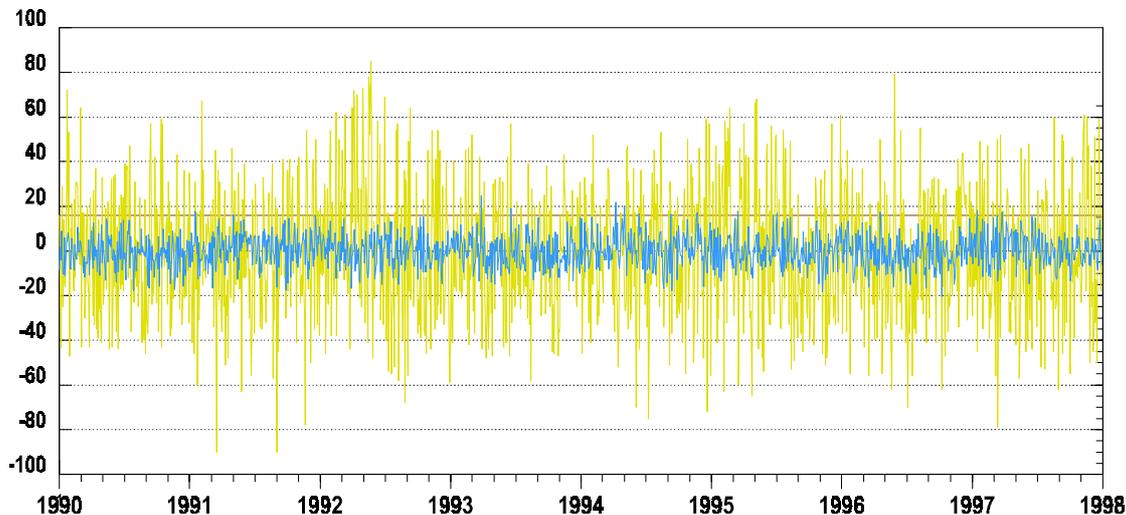


Abb. 74) Zeitreihe der Originalwerte (normal verteilte Zufallsfolge = gelbe Kurve) und der Differenz zwischen den Originalwerten und den Erwartungswerten (blaue Kurve)

Diese Kurve wird im folgenden schrittweise den Plausibilitätsprüfungen unterzogen, wobei die einzelnen Schritte jeweils kurz beschrieben und diskutiert werden. Erst anschließend werden diese Schritte an den originalen Datenkollektiven durchgeführt.

3.1.1 Räumliche Fehlwertersetzung

Eine der wesentlichsten Voraussetzungen für die Anwendung dieses und aller anderen zeitreihenanalytischen Verfahren ist das Vorliegen kompletter Datenkollektive, die mithin keine Ausfallwerte beziehungsweise kleinere oder größere Lücken aufweisen dürfen. Das gleiche Verfahren, das im folgenden zur Validierung eingesetzt wird, nämlich das stufenweise screening-Regressionsverfahren, bei dem Messwerte benachbarter Stationen und meteorologische Messgrößen als exogene Variablen zur Berechnung von Schätzwerten eingesetzt werden, wird zum "Auffüllen" der Datenlücken verwendet, so daß lückenlose Datensätze zur weiteren Bearbeitung vorliegen. Diese "horizontale Fehlwertersetzung" mit Hilfe von Messwerten benachbarter Messstationen (multivariates Verfahren) bildet somit eine Erweiterung der zeitlichen Fehlwertersetzung, bei der lediglich die Informationen aus der Zeitreihe selbst verwendet wird (univariates Verfahren)

| Modif. Zufallsfolge | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------------------|------|------|------|------|------|------|------|------|------|
| Korrelation (r^2) | 0.92 | 0.85 | 0.80 | 0.76 | 0.67 | 0.54 | 0.48 | 0.48 | 0.52 |

Tabelle 01) Bestimmtheitsmaß zwischen der Ausgangsreihe und 9 modifizierten Zufallsfolgen.

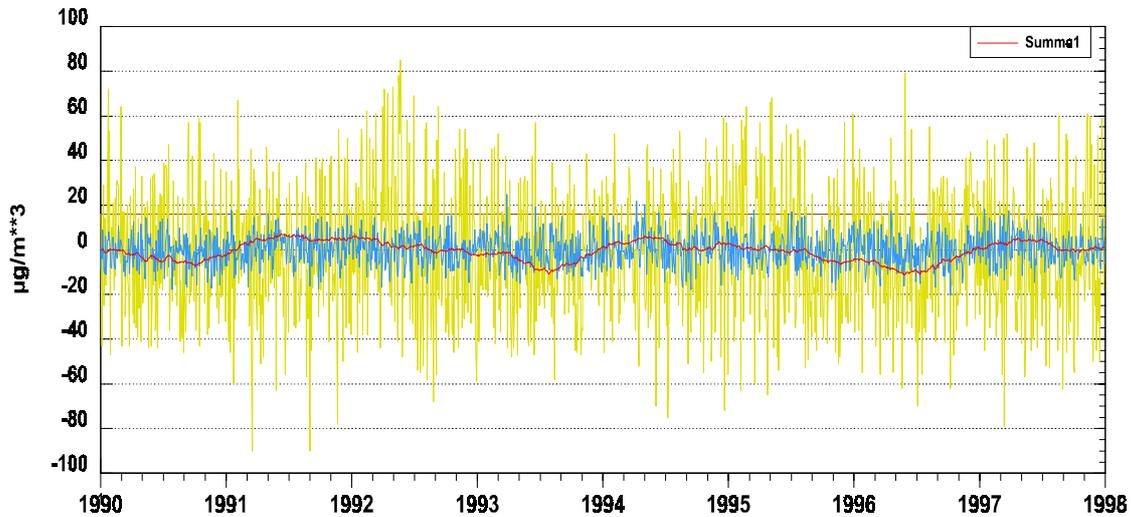
.Für die horizontale Fehlwertersetzung mit Zufallszahlenfolgen werden „Nachbarstationen" simuliert, indem auf die zu ersetzende Zufallsfolge (Originalwerte = gelbe Kurve in Abb.74, im weiteren als „**Kurve I**" bezeichnet) 9 andere Zufallsfolgen mit unterschiedlicher Amplitude aufgeprägt wurden. Die so erhaltenen 9 Zufallsfolgen sind der Ausgangsfolge per Definition ähnlich und können zur Simulation von Erwartungswerten für eine komplette Fehlwertersetzung der Ausgangsreihe herangezogen werden. Die einfache lineare Korrelation zwischen der Ausgangsreihe und den modifizierten Zufallsfolgen ist aus Tab.01 ersichtlich. Das Spektrum der Korrelationen wurde so gewählt, daß es mit Korrelationen, wie sie zwischen realen

Ozonzeitreihen beobachtet werden, näherungsweise übereinstimmt. In der Tab.02 ist die Güte der horizontalen Fehlwertersetzung bei schrittweiser Einbeziehung weiterer Prediktoren aufgeführt. Als Maß für die Güte der kompletten Fehlwertersetzung wird die Restvarianz und das Bestimmtheitsmaß ($BM = r^2$) verwendet. Wegen der statistischen Stabilität und zur Vermeidung von Überbestimmtheiten der räumlichen Ersetzung werden maximal 8 Prediktoren zugelassen. Die zu ersetzende Zufallsfolge in Abb.74 hat eine Ausgangsvarianz von $27,6\mu\text{g}/\text{m}^3$.

| Anz. verwendeter Prediktoren | Bezeichnung_der_verwendeten_Prediktoren | Faelle | Restvarianz | BM |
|------------------------------|---|--------|-------------|------|
| 2 | Zufallsprediktor | 1464 | 11.6 | 81.3 |
| 4 | Zufallsprediktor | 1464 | 10.8 | 83.9 |
| 5 | Zufallsprediktor | 1464 | 9.7 | 86.9 |
| 3 | Zufallsprediktor | 1464 | 9 | 88.8 |
| 6 | Zufallsprediktor | 1464 | 8.4 | 90.2 |
| 3 | Zufallsprediktor | 1464 | 8.1 | 90.8 |
| 8 | Zufallsprediktor | 1464 | 7.9 | 91.4 |
| 10 | Zufallsprediktor | 1464 | 7.8 | 91.6 |

Tab. 02) Kenngrößen der Zufallszahlenfolgen (“Nachbarstationen”)

Bei alleiniger Verwendung des Zufallsprediktors 2 zur räumlichen Ersetzung beträgt die verbleibende Restvarianz in der Zeitreihe der Erwartungswerte $11,4\mu\text{g}/\text{m}^3$. Sie reduziert sich bei Hinzunahme weiterer Einflußgrößen auf $7,0\mu\text{g}/\text{m}^3$, wobei sich das Bestimmtheitsmaß auf 93,6% erhöht.



**Abb. 75) Summation der Kurve II (rote Kurve = Summenkurve I)
Testgröße zur Identifikation von Strukturbrüchen**

Die **Differenz** zwischen der Reihe der Erwartungswerte und derjenigen der Originalwerte ist als blaue Kurve in Abb.74 eingezeichnet. Diese aus den berechneten Erwartungswerten bestehende Reihe wird im weiteren Verlauf mit „**Kurve II**“ bezeichnet. Sie weist eine geringere Variabilität auf als die Originalreihe und läßt weder einen Trend noch Strukturbrüche erkennen.

3.1.2 Trends und Stufen

Zur Validierung der Zeitreihe „**Differenz**“, also der nach kompletter Fehlwertersetzung verbleibenden Varianz der Zeitreihe (**Kurve II**), werden nun 5 Testgrößen eingeführt.

3.1.2.1 Summenkurve I

Diese Testgröße stellt die Summe aller positiven und negativen Differenzen der **Kurve II** bis zu einem gegebenen Zeitpunkt dar. Für eine normal verteilte Zufallsfolge schwankt diese **Summenkurve I** um den Nullpunkt (Abb.75).

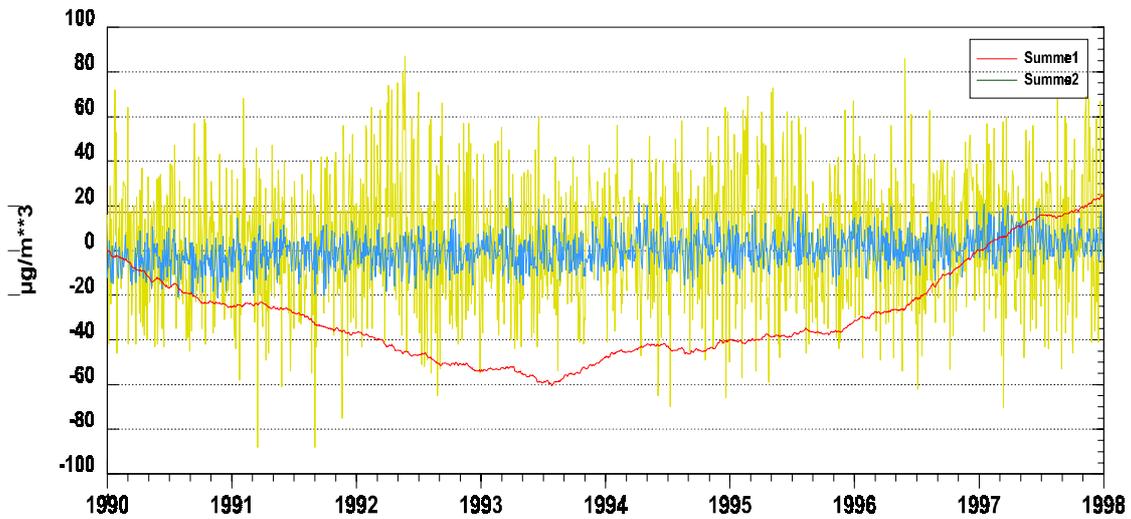


Abb. 76) Verlauf der Summenkurve I bei Anwesenheit eines Trends in der Originalreihe Trend in der Originalreihe = $2\mu\text{g}/\text{m}^3$ pro Jahr

Prägt man der Ausgangsreihe der Originalwerte (**Kurve I in Abb. 74**) einen linearen Trend von $2\mu\text{g}/\text{m}^3$ pro Jahr auf (Abb.76), so reagiert die **Summenkurve I** sehr empfindlich. Da die Ersetzung der Fehlwerte durch Erwartungswerte mittels Methode der kleinsten Quadrate (multiple Regression) erfolgt, ist die Summe der Abweichungen über den gesamten Zeitraum nahe Null, d.h. die **Summenkurve I** beginnt und endet immer in der Nähe von Null. Kleinere Abweichungen sind numerische Ungenauigkeiten und können vernachlässigt werden.

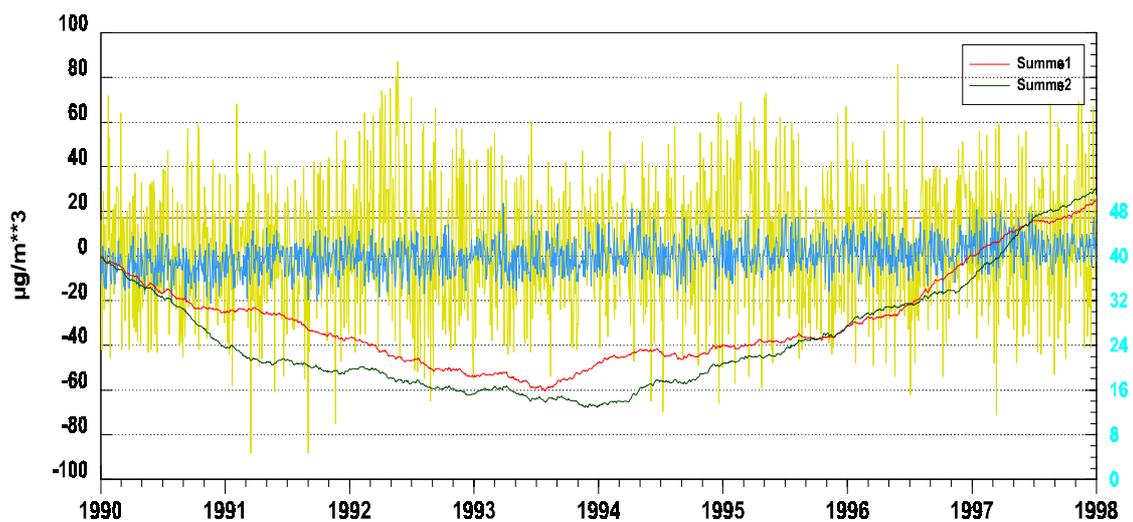


Abb. 77) Verlauf der Summenkurven I und II bei Anwesenheit eines Trends in der Originalreihe

Es besteht jedoch ein Problem:

Es ist a priori nicht zu unterscheiden, ob ein festgestellter Trend der Zeitreihe der Originalwerte zuzuordnen ist oder ob dieser durch einen Trend in einer zur Berechnung der Erwartungswerte verwendeten Zeitreihe (Prediktoren) vorgetäuscht wird. Hierzu wird eine zusätzliche Testgröße benötigt.

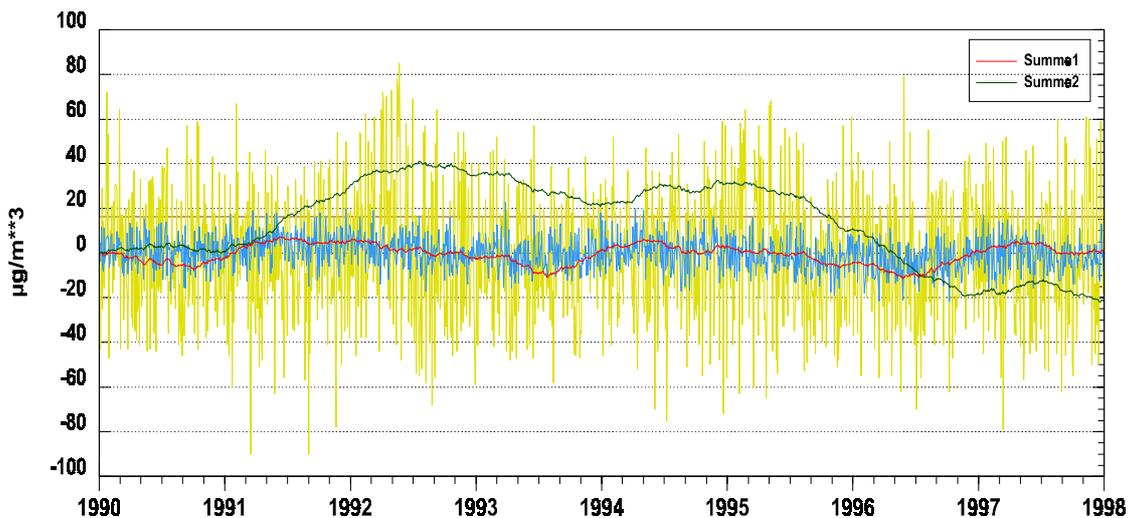


Abb.78) Verlauf der Summenkurven I und II bei Anwesenheit eines Trends im stärksten Prediktor (Originalreihe ist trendfrei)

3.1.2.2 Summenkurve II

Die Einführung einer weiteren Testgröße, nämlich der **Summenkurve II**, erlaubt eine Beantwortung der Frage, ob ein Trend in der Originalreihe oder in dem stärksten Prediktor vorliegt. Der Unterschied zur **Summenkurve I** liegt darin, daß in dem stufenweisen Screening- Verfahren zur Herleitung der multiplen Regressionsbeziehung der stärkste Prediktor, also der Prediktor, der zur Fehlwertersetzung und Summation an erster Stelle verwendet wird, zur Erzeugung der blauen Kurve nicht zur Verfügung steht. Der Vergleich der beiden Summenkurven (**Summenkurve I und II**) erlaubt Rückschlüsse, woher der gefundene Trend stammt. Verlaufen wie in Abb.77 die beiden Kurven gleichsinnig, so stammt der Trend aus der Originalreihe. In Abb.78 wurde dem zweiten Zufallsprediktor ein linearer Trend von $3\mu\text{g}/\text{m}^3$ pro Jahr aufgeprägt. Dieser

Trend hat zur Folge, daß der Verlauf der **Summenkurve II** diesen Trend anzeigt, während die **Summenkurve I** keinen systematischen Verlauf aufweist. Außerdem folgt, daß ein gleichgerichteter Trend in der Originalreihe und im stärksten Prediktor einen entgegengesetzten Verlauf der beiden Summenkurven bewirkt.

Zur Entscheidungshilfe bezüglich der Frage, ob die Kurven einen **signifikanten** Trend anzeigen, wird ein weiterer Test eingeführt

3.1.2.3 Gleitender t-Wert

Zur Prüfung der Null-Hypothese, daß zwei Mittelwerte aus einer Grundgesamtheit stammen, eignet sich der t-Test. Überschreitet der berechnete t-Wert die Signifikanzschranke einer vorgegebenen Irrtumswahrscheinlichkeit, hier 5%, so muß die Null-Hypothese abgelehnt werden, was bedeutet, daß die Mittelwerte mit einer definierten Wahrscheinlichkeit unterschiedlichen Grundgesamtheiten angehören.

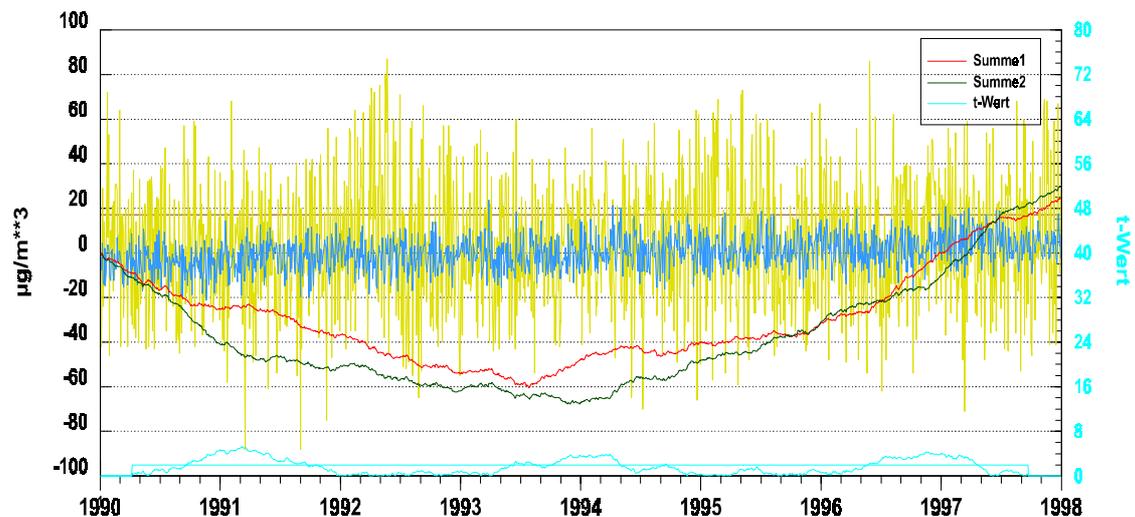


Abb. 79) Testgröße "Gleitender t-Wert (blaue Kurve) bei Anwesenheit eines Trends in der Originalreihe

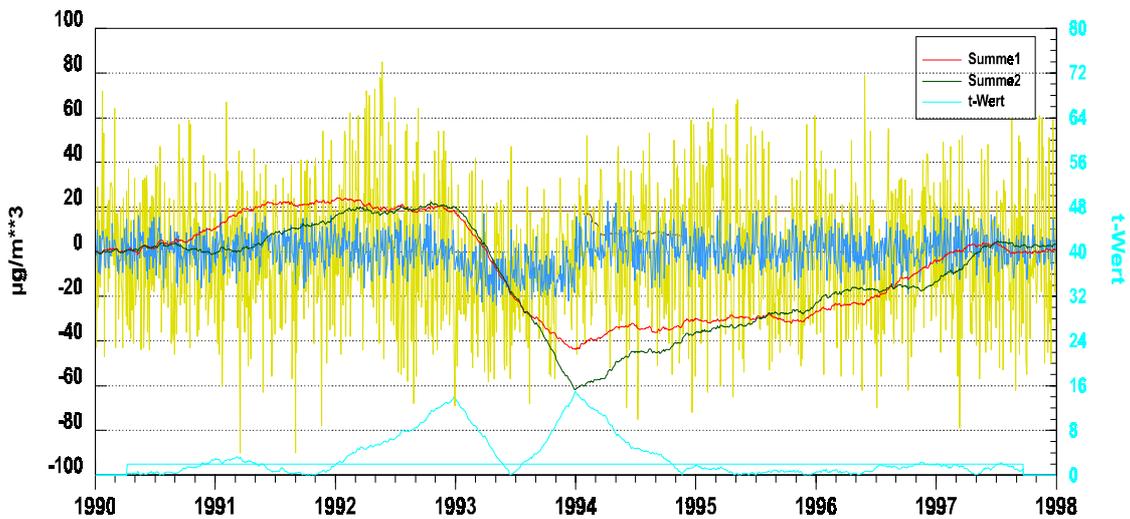


Abb. 80) Testgröße “Gleitender t-Wert” (blaue Kurve) bei Anwesenheit einer stufenförmigen Intervention in der Originalreihe

Die Zahl der effektiven Freiheitsgrade, die zur Bestimmung der Signifikanzschwelle notwendig sind, werden über die **Bartelsche Erhaltungszahl** und über die Berechnung der **Autokorrelation** der Zeitreihe bestimmt. Die blaue Linie in Abb.79 stellt den für jeden Tag berechneten t-Wert in einem Zeitfenster von plus/minus 100 Tagen um diesen Zeitpunkt dar. Dieses Fenster gleitet über die gesamte Zeitreihe. Die dünne blaue Linie markiert die Signifikanzschwelle des gleitenden t-Tests. Bei einer trendfreien und normal verteilten Zufallsfolge beobachtet man praktisch keine Überschreitung des Signifikanzniveaus. Liegt eine Intervention wie z.B. ein Trend vor (Abb.79), so spricht der gleitende t-Test jedoch an. In dieser Abbildung zeigen die Summenkurven I und II diesen Trend an, die 3 Pfeile zeigen das Ansprechen des gleitenden t-Tests auf Signifikanz dieses Trends.

Die Sensibilität sowohl der Summenkurven als auch des gleitenden t-Tests auf systematische Fehler in den Messreihen demonstriert die Abb.80. Hier wurde auf die normal verteilte Zufallsfolge in einem Jahr (1993) eine kleine systematische Abweichung von $-10\mu\text{g}/\text{m}^3$ aufgeprägt. In der Originalreihe (gelbe Kurve in Abb.80) ist diese Manipulation kaum sichtbar; in der **Kurve II** (blaue Kurve in Abb.80) läßt sich die Verschiebung gerade noch erkennen. Deutlich wird der Zeitraum durch die **Summenkurven I** und **II** markiert. Da beide Kurven gleichsinnig verlaufen, ist die

systematische Abweichung der Originalreihe zuzuordnen. Zusätzlich läßt sich aus der Differenz zwischen Beginn und Ende der Verschiebung aus der **Summenkurve II** der Betrag des systematischen Fehlers ermitteln ([Wert bei Beginn minus Wert am Ende] *20/Anzahl der Tage). Die Zahl 20 ergibt sich durch die Skalierung der Y-Achse. Deutlich markiert auch der gleitende t-Test den Beginn und das Ende der systematischen Verschiebung.

3.2 Strukturbrüche in der Varianz

Eine zweite Kategorie von Fehlern in Messreihen ist die Änderung der Empfindlichkeit von Messgeräten, welche auch eine Änderung in der Variation der Messwerte verursachen kann. Meist tritt dieser Fehler in Kombination mit einer Mittelwertverschiebung auf.

3.3 Sensitivität der Varianz

Das Sensitivitätsmaß ist definiert als gleitendes Mittel des Verhältnisses der Streuung der Originalreihe zur Streuung der komplett fehlwertersetzten Reihe über ein Zeitfenster von 30 Tagen. Für eine normal verteilte Zufallsfolge liegt das Sensitivitätsmaß nahe Null. Mit diesem Maß lassen sich zwei Arten von Defekten in den Zeitreihen erkennen:

1. Die Empfindlichkeit der Ausgangsreihe verändert sich, ihre Änderung von Termin zu Termin ist gleichsinnig mit anderen Messstationen.
2. Die Varianz der Originalreihe bleibt gleich, ihre Synchronität, d.h. gleichsinnige tägliche Änderungen zu anderen Messreihen, ist nicht mehr gewährleistet.

Im letzten Fall hat der fehlwertersetzte Zeitraum eine ähnliche Varianz wie die Originalreihe, folglich erhöht sich das Sensitivitätsmaß.

In Abb.81 wurde die Streuung der Originalreihe 1993 um 50% und in Abb.82 um 30% reduziert. Deutlich ist das Ansprechen des Sensitivitätsmaßes zu erkennen, während die Maßzahlen für systematische Fehler nicht reagieren. Als quantitativer Anhaltspunkt, dessen Überschreitung auf eine Inkonsistenz der Varianz der Originalreihe hindeutet, wurde ein empirisch optimiertes Verhältnis der Streuung der Originalreihe und der räumlichen Fehlwertersetzung verwendet. Diese Schranke berücksichtigt sowohl die

Variabilität der Ausgangsreihe als auch die Güte ihrer räumlichen Ersetzung. Je größer die Varianz der Ausgangsreihe und je geringer die Güte der räumlichen Ersetzung ist, desto höher ist diese Schranke. In Abb.81 überschreitet das Sensitivitätsmaß diese Linie deutlich, in Abb.82 nur noch leicht (diese Schranke ist in den Abbildungen als dünne braune Linie bei etwa $20\mu\text{g}/\text{m}^3$ eingezeichnet).

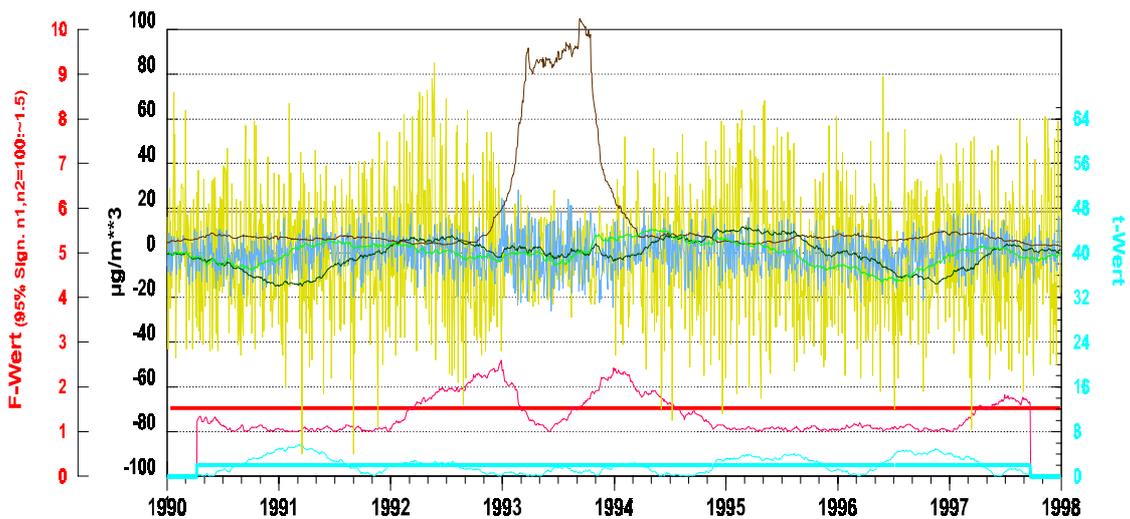


Abb. 81) Beispiel für das beschriebene “Sensitivitätsmaß der Varianz”

Im Jahr 1993 ist die Varianz der Originalreihe um 50% reduziert

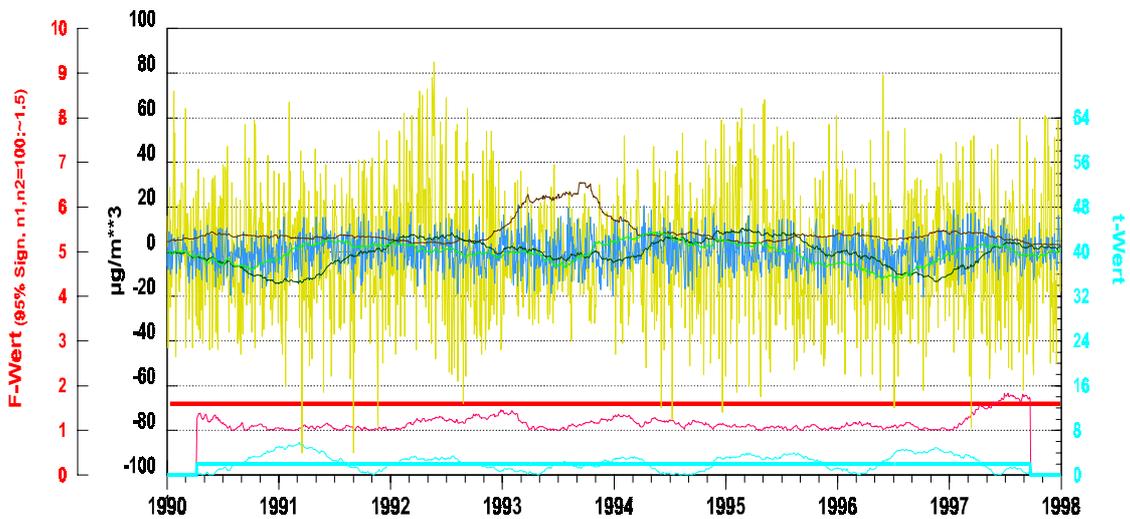


Abb. 82) Beispiel für das beschriebene “Sensitivitätsmaß der Varianz”

Im Jahr 1993 ist die Varianz der Originalreihe um 30 Prozent reduziert

3.4 Gleitender F-Wert

Zur Prüfung der Null-Hypothese, daß zwei Streuungen aus einer Grundgesamtheit stammen, eignet sich der F-Test. Überschreitet der berechnete F-Wert die Signifikanzschranke einer vorgegebenen Irrtumswahrscheinlichkeit, hier 5%, so muß die Null-Hypothese abgelehnt werden. Die Zahl der effektiven Freiheitsgrade, die zur Bestimmung der Signifikanzschwelle notwendig sind, werden über die Bartelsche Erhaltungszahl und über die Berechnung der Autokorrelation der Zeitreihe bestimmt. Die rote Linie im unteren Bereich der Graphiken in den Abb.81 und Abb.82 stellt den für jeden Tag berechneten F-Wert in einem Zeitfenster von plus/minus 100 Tagen um diesen Zeitpunkt dar. Dieses Fenster gleitet über die gesamte Zeitreihe. Die dazugehörige rote Gerade markiert die Signifikanzschwelle des gleitenden F-Tests. Bei einer normal verteilten Zufallsfolge, wie sie in Abb.74 gegeben ist, ist die Überschreitung des Signifikanzniveaus praktisch nicht vorhanden. Liegt eine deutliche Veränderung der Varianz der **Kurve I** vor, wie in Abb.81 veranschaulicht, so spricht der gleitende F-Test deutlich an. Eine nur schwache Reduktion der Varianz um 30%, wie sie in Abb.82 zu sehen ist, kann jedoch durch diesen Test statistisch nicht gesichert werden.

3.5 Kombinationen von Strukturbrüchen

Um das Verhalten der Testgrößen in den Fällen zu demonstrieren, in denen mehrere Strukturbrüche in Kombination auftreten - wie es in der Realität häufig der Fall ist - werden im folgenden einige Beispiele gezeigt.

Kombiniert man eine systematische Verschiebung der Zeitreihe ($-10\mu\text{g}/\text{m}^3$ im Jahr 1993) mit einer Reduktion der Varianz um 30% ebenfalls im Jahr 1993, wie in Abb.83 gezeigt, sprechen alle Testkurven entsprechend an. Der deutliche Abfall der Summenkurven I und II mit Beginn des Jahres 1993 zeigt den Beginn der Stufe mit $-10\mu\text{g}/\text{m}^3$ und der Umkehrpunkt dieser beiden Kurven mit dem Ende des Jahres 1993 zeigt das Ende dieser Stufe an. Der annähernd parallele Verlauf beider Kurven läßt erkennen, daß die Stufe in der zu untersuchenden Reihe selbst vorhanden ist und nicht etwa in den Reihen der Prediktoren. Die Signifikanz des Strukturbruches ergibt sich aus den beiden "Spitzen" des **Gleitenden t-Wertes** zu Beginn und Ende des Jahres 1993. Die Reduzierung der Varianz um 30% zeigt sich am deutlichsten in der Sensitivitätskurve, die im Jahr 1993 das empirisch optimierte Verhältnis zwischen der Originalreihe und der fehlwertersetzten Reihe deutlich übersteigt. Der Verlauf des gleitenden F-Wertes zeigt die Varianzreduzierung ebenfalls an und überschreitet auch das Signifikanzniveau, läßt jedoch infolge eines relativ breiten Maximums den Beginn und das Ende der Intervention kaum erkennen.

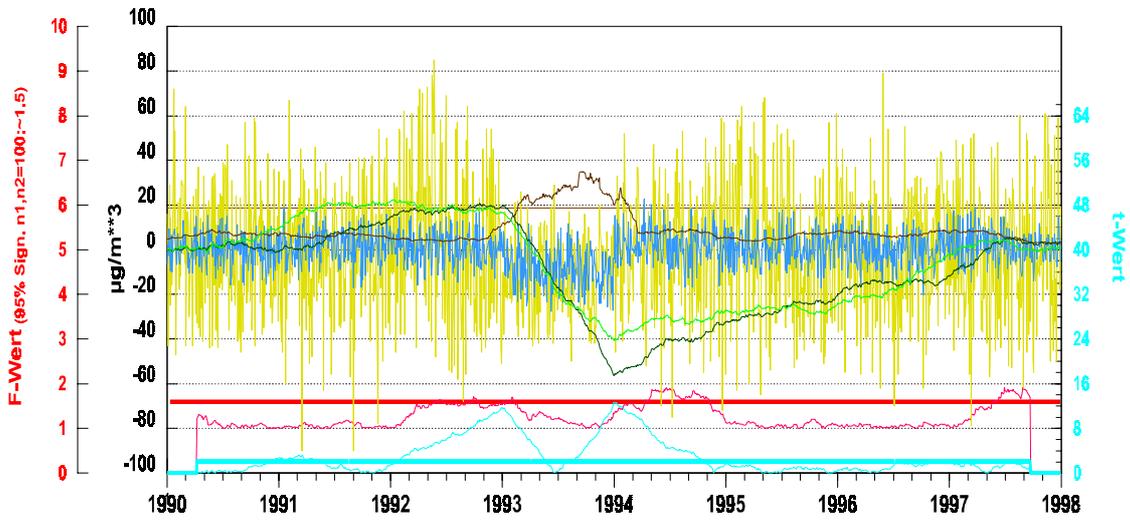


Abb. 83) Sensitivität um 30% reduziert und systematische Abweichung von $-10 \mu\text{g}/\text{m}^3$ im Jahr 1993

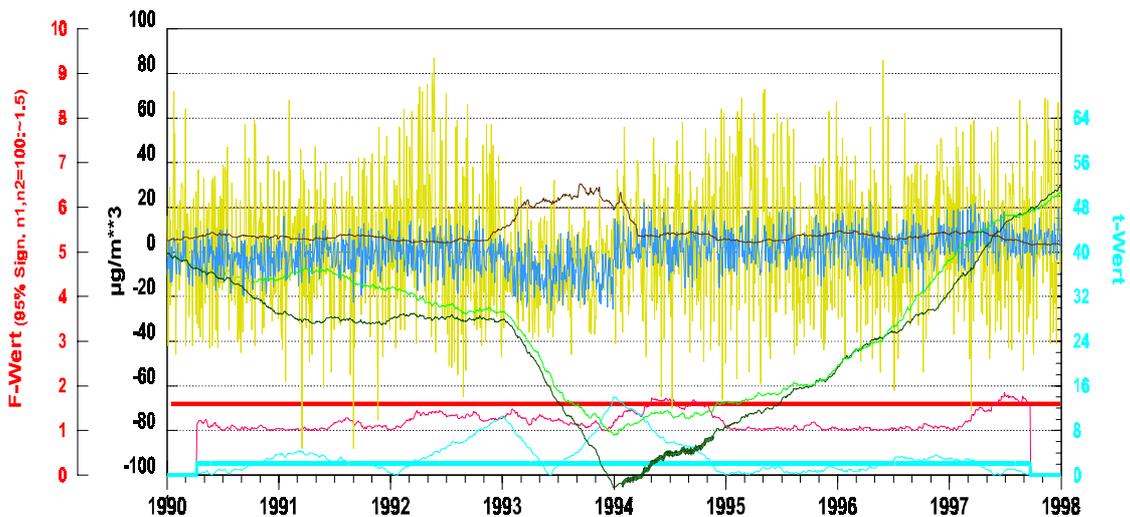


Abb. 84) Sensitivität um 30% reduziert und systematische Abweichung von $-10 \mu\text{g}/\text{m}^3$ im Jahr 1993 sowie ein Trend von $+2 \mu\text{g}/\text{m}^3$ pro Jahr

Die zusätzliche Aufprägung eines linearen Trends von $2 \mu\text{g}/\text{m}^3$ pro Jahr (insgesamt also drei Interventionen, nämlich Stufe, Varianzreduktion und Trend), wie in Abb.84 demonstriert, bewirkt lediglich eine Modifikation der **Summenkurven I und II** (Abb.84). Diese beiden Kurven weisen bereits zu Beginn der Zeitreihe eine abfallende

Tendenz auf, die auf einen Trend in der Originalreihe hinweisen (dies zeigt der annähernd parallele Verlauf beider Kurven).

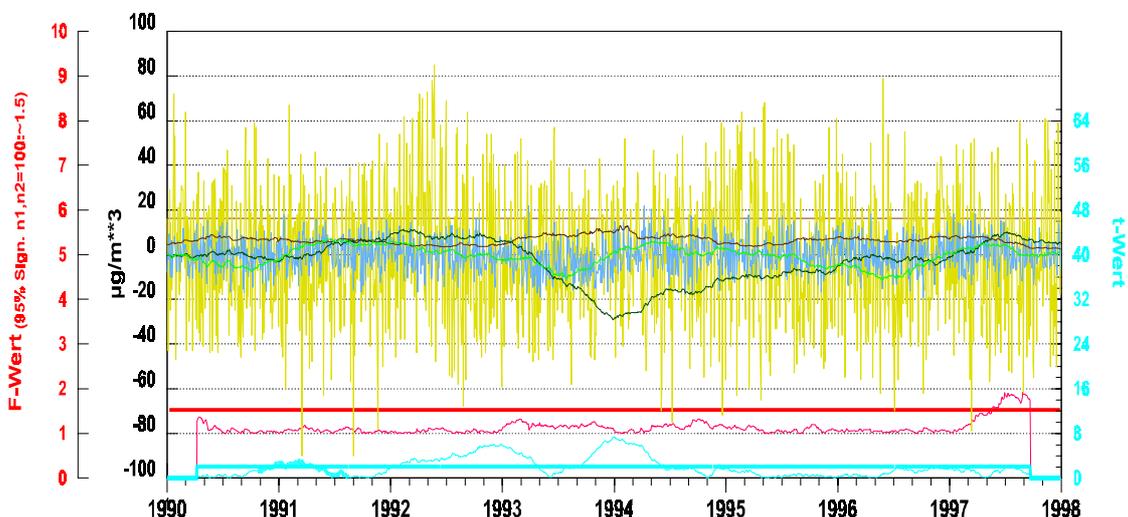


Abb. 85) Systematische Abweichung im stärksten Prediktor im Jahr 1993 (+10 $\mu\text{g}/\text{m}^3$)

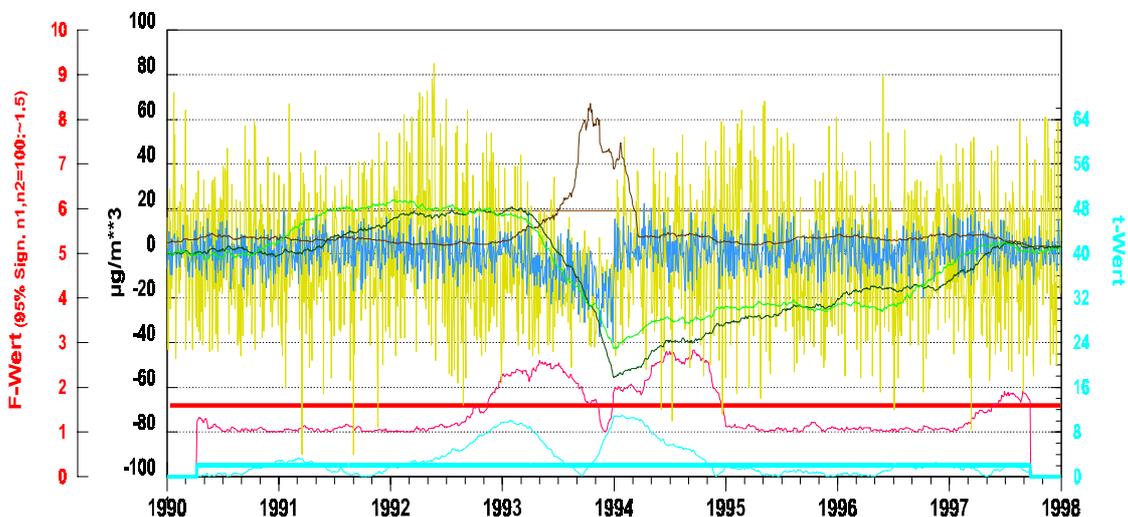


Abb. 86) Zunehmende systematische Abweichung bis - 18 $\mu\text{g}/\text{m}^3$ und Sensitivität abnehmend auf - 50%, beides im Jahr 1993

Durch die Stufe im Jahr 1993 verstärkt sich diese abfallende Tendenz deutlich. In der Summe beider systematischer Änderungen (Trend und Stufe) erreichen beide Prüfgrößen gegen Ende des Jahres 1993 den niedrigsten Wert (etwa $-50\mu\text{g}/\text{m}^3$ in

Abb.83 gegenüber $-100\mu\text{g}/\text{m}^3$ in Abb.84). Während die Prüfgröße **Gleitender t-Wert** die Signifikanz des Strukturbruches "**Stufe im Jahr 1993**" anzeigt, erreicht der **Gleitende F-Wert** die Signifikanzschwelle nicht. Trotz der tatsächlichen Reduzierung der Varianz um 30% hält die Prüfgröße diese Änderung nicht oder noch nicht für eine systematische Änderung in den Messwerten.

Eine systematische Abweichung des stärksten Prediktor von $+10\mu\text{g}/\text{m}^3$ im Jahre 1993 (Stufe), der zur Fehlwertersetzung verwendet wird, wirkt sich wie in Abschnitt 3.1.2 beschrieben auf beide Summenkurven aus (Abb.85). Ein Vergleich der **Summenkurve I** und **II** zeigt jedoch, daß die Ursache nicht in der Ausgangsreihe zu suchen ist, denn beide Kurven zeigen zum fraglichen Zeitpunkt (im Jahr 1993) keinen synchronen Verlauf. Auch hier weist der **Gleitende t-Wert** auf die Signifikanz des Strukturbruches hin.

In Abb.86 wurden sowohl eine systematische Abweichung zunehmend auf $18\mu\text{g}/\text{m}^3$ als auch eine Abnahme der Sensibilität einer Messreihe auf 50% (im Jahr 1993) simuliert. Diese in ihrem zeitlichen Verlauf trendähnlich abnehmende Stufe, die gleichzeitig mit einer Reduzierung der Varianz der Messwerte verbunden ist, zeigt eine deutliche Verschiebung der - im übrigen parallel verlaufenden - **Summenkurven I** und **II** bei Beginn und Ende des Zeitraumes der Stufe. Für die Signifikanz der Stufe spricht der Gleitende t-Wert im unteren Teil der Abbildung mit hoch signifikanten Werten zu Beginn und Ende des Jahres 1993. Das gleiche Ergebnis lassen die Prüfgrößen der Varianz mit ebenfalls hoch signifikanten F-Werten im fraglichen Zeitraum erkennen.

Was in diesen aus Zufallsfolgen konstruierten Datensätzen, bei denen in einem relativ kurzen Zeitabschnitt ein oder mehrere Strukturbrüche "eingebaut" sind und die im übrigen recht gleichmäßig verlaufen, noch deutlich und auch eindeutig erkennbar ist, wird sich in den Datenkollektiven realer Ozonzeitreihen nur schwierig und mit sehr viel Erfahrung interpretieren lassen. Im folgenden werden einige Beispiele realer Ozonzeitreihen gezeigt, die diese Interpretationsschwierigkeiten demonstrieren.

4 Anwendung auf Ozonzeitreihen

Zur Untersuchung der realen Ozonzeitreihen auf Inkonsistenzen, d.h. auf Brüche, Verwerfungen, Veränderung der Sensitivität der Messgeräte usw. werden nicht die Originalreihen selbst verwendet, sondern die jahres- und tagesgangbereinigten Daten, so wie sie nach der Eliminierung grober Fehler in der Datenkontrolle Stufe 1 und der Stufe 2 zur Verfügung stehen. Aus den Abb.79 und Abb.84 ist ersichtlich, daß sich vorhandene Trends in den Reihen störend auf die Identifikation von systematischen Fehlern sowie Änderungen in der Varianz der Zeitreihen auswirken. Deshalb werden zusätzlich aus allen Datenreihen die linearen Trends beseitigt.

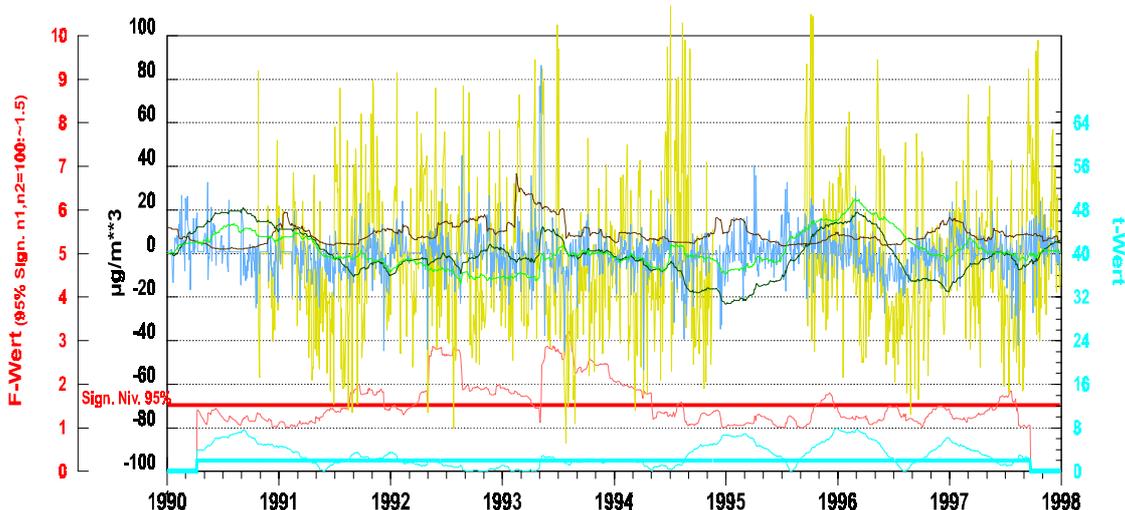
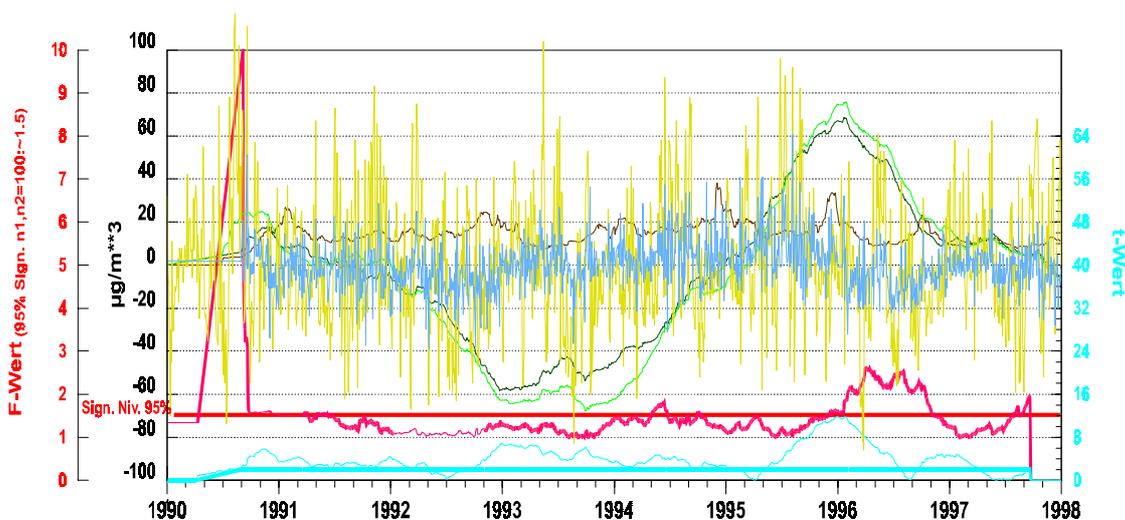


Abb. 87) Original-Ozonreihe der Station LÖRICK (NW071), Zeitraum 1990 – 1997 Messnetz Nordrhein-Westfalen (Raum Düsseldorf)

In Abb.87 sind die Ergebnisse der Validierung der Ozonzeitreihe der Station **LÖRICK** (NW071) dargestellt. Die Kurven zeigen im Vergleich zu vielen anderen Messstationen einen relativ ruhigen Verlauf. Schon die Höhe des Bestimmtheitsmaßes von 90.7% bei der räumlichen Fehlwertersetzung deutet auf eine Messreihe hin, die kaum mit Fehlern behaftet sein kann. Zur räumlichen Fehlwertersetzung wurden im Rahmen der multiplen Regression zusätzlich meteorologische Größen angeboten. In diesem Fall ist das Tagesminimum der Lufttemperatur des Vortages die zweitstärkste Einflußgröße. Die Restvarianz von $10,7\mu\text{g}/\text{m}^3$ zeigt deutlich, daß die Messreihe mit anderen Ozonreihen

harmonisiert. Der im Vergleich zu den nachfolgenden Abbildung relativ “ruhige” Verlauf der **Summenkurven I** und **II** deutet auf die Abwesenheit stufenförmiger Interventionen hin (ein eventuell vorhandener langfristiger (globaler) Trend ist zu Beginn der Untersuchungen eliminiert worden). Lediglich der F-Wert deutet im Juli 1992 und im Mai 1993 auf einen möglichen Fehler in Form einer Varianzreduzierung etwa von Mitte 1992 bis Mitte 1993 hin.



**Abb. 88) Original-Ozonreihe der Station SCHAUNSLAND (UB004),
Zeitraum 1990 - 1997 Messnetz Umweltbundesamt**

Auch die Messreihe **SCHAUNSLAND** (UB004, Abb.88) kann man noch als brauchbar bezeichnen. Zu sehen ist eine systematische Verschiebung von $3\mu\text{g}/\text{m}^3$ zwischen 1993 und 1996 gegenüber dem übrigen Zeitraum. Es handelt sich hierbei entweder um einen lokalen Trend (in diesem Fall bedeutet dieser Trend eine Umkehr des ebenfalls lokalen abnehmenden Trends vom Beginn der Zeitreihe bis zum Jahr 1993) oder sogar um eine langperiodische Schwingung. Auf jeden Fall ist dieser Strukturbruch in der Ozonreihe **SCHAUNSLAND** selbst begründet, da die Summenkurve I und II weitgehend einen klaren parallelen Verlauf aufweisen. Eine weitere vermutliche Inkonsistenz kann im Jahr 1996 identifiziert werden. Hier ist der Abfall der **Summenkurven I** und **II** am größten. Zu diesem Zeitpunkt pendelt entweder der lokale Trend zurück oder - und das ist wahrscheinlicher - es handelt sich hierbei um einen Teil der genannten langperiodischen Schwingung. Gleichzeitig mit dem

letzten genannten Strukturbruch zeigt der **Gleitende F-Wert** eine signifikante Änderung bei der Variation der Messwerte an, die etwa zu Beginn des Jahres 1996 einsetzt. Ob allerdings diese Strukturbrüche in einer systematischen Änderung der Ozonkonzentration bzw. in Änderungen bei den Emissionen der ozonbildenden Stoffe begründet ist oder ob es sich um messtechnisch bedingte artifizielle Schwankungen handelt, beantwortet diese Analyse nicht.

Deutliche systematische Fehler sind bei der Station **VILLINGEN** (BW039, Abb.89) zu beobachten. Die systematischen Fehler sind gekoppelt mit Sprüngen in der Varianz der Messreihe.

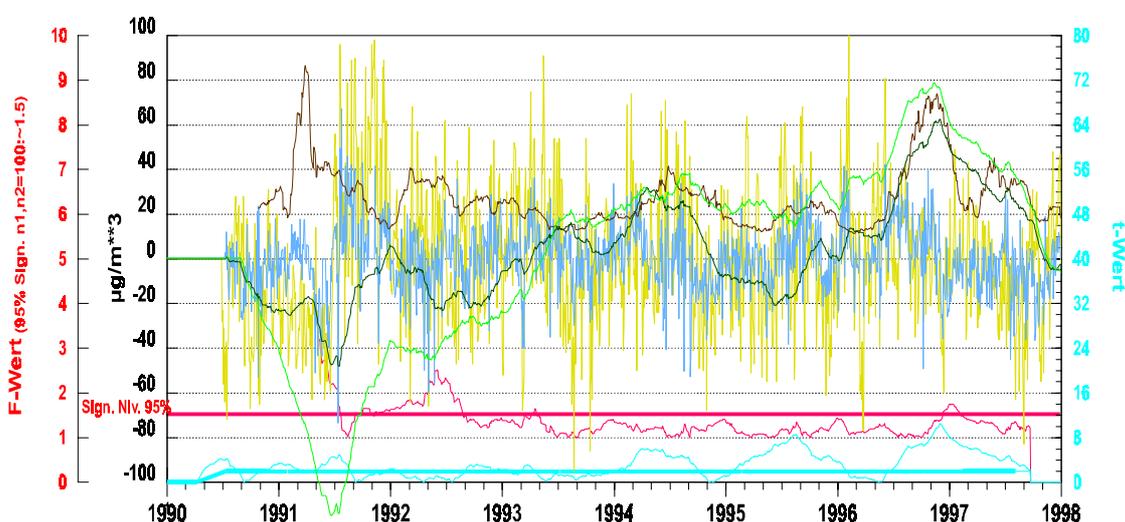


Abb. 89) Original-Ozonreihe der Station VILLINGEN (BW039), Zeitraum 1990 – 1997 Messnetz Baden-Württemberg

Sowohl die **Summenkurven I** und **II** als auch das **Sensitivitätsmaß** zeigen Strukturbrüche an. Im Jahr 1991 ist dieser Strukturbruch bereits in der Originalreihe (gelbe Kurve) zu erkennen. Aber auch die Zeitreihe der fehlwertersetzten Daten (in dieser Abbildung graue, sonst blaue Kurve) zeigt diesen Bruch, der bis in die Mitte des Jahres 1991 hinein reicht, deutlich an. Der **Gleitende t-Wert** zeigt den Beginn und das Ende dieser Intervention als signifikant an, was beim Ausmaß des Strukturbruches nicht verwunderlich ist. Auch die Sensibilitätskurve (braune Kurve in der Abbildung) übersteigt die Linie, die die optimierte Varianz anzeigt und die bei etwa $+20\mu\text{g}/\text{m}^3$ liegt,

deutlich und zeigt mithin an, daß in diesem Zeitraum eine krasse Änderung in der Variation der Messwerte auftritt. Ab etwa Mitte 1994 zeigt der **Gleitende t-Wert** mehrere Stufen. Die erste Markierung ist im Jahr 1994 erkennbar, eine zweite erfolgt gegen Mitte des Jahres 1995 und eine letzte zu Beginn des Jahres 1997. Die **Summenkurven I** und **II** lassen besonders in diesem letzten Zeitraum ein trendähnliches Verhalten (lokaler Trend) vermuten, der mit einer Veränderung der Varianz gekoppelt ist. Hierauf weist der steile Anstieg der braunen Kurve (=Sensitivitätsmaß) hin, dessen Signifikanz durch die Spitze beim **Gleitenden F-Wert** bestätigt wird. Trotz dieser “qualitativen Mängel” kann das Datenkollektiv der Station **VILLINGEN** ähnlich wie dasjenige der Station **SCHAUINSLAND** als weitgehend plausibel eingestuft werden. Dieses Qualitätsurteil ist bei einem großen Teil der nachfolgenden Stationen nicht mehr in dieser Eindeutigkeit zu vergeben.

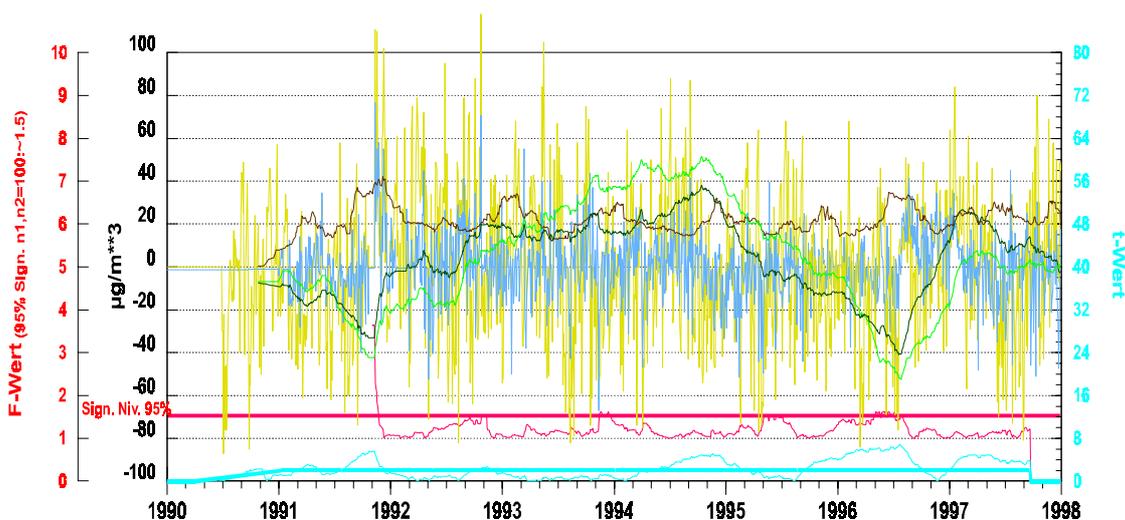


Abb. 90) Original-Ozonreihe der Station WALDSHUT BW040), Zeitraum 1991 – 1997 Messnetz Baden-Württemberg

Bei der Station **WALDSHUT** (BW040, Abb.90) ist gegen Ende des Jahres 1991 ein systematischer Fehler zu erkennen. Hier zeigen sowohl die beiden **Summenkurven** als auch der **Gleitende t-Wert** eine Stufe an, die sich auch in der Originalreihe ausmachen läßt. Daß sich in diesem Zeitraum auch die Variabilität der Messwerte ändert, zeigen übereinstimmend die Sensitivitätskurve als auch der **Gleitende F-Wert**. Nach einem sehr ruhigen Verlauf der Kurven der Prüfgrößen erfolgt gegen Ende des Jahres 1994

eine stufenförmige Intervention, die der **Gleitende t-Wert** als signifikant bezeichnet. Allerdings ist diese Stufe mit niedrigeren Messwerten gegenüber den umliegenden Zeiträumen nicht mit einer signifikanten Änderung der Variabilität verbunden, da der **Gleitende F-Wert** zwar Spitzen in diesem Zeitraum anzeigt, die jedoch die Signifikanzschwelle nicht übersteigen. Auch dieser Messstation ist noch ein insgesamt plausibles Datenkollektiv mit “kleinen Mängeln” zu bescheinigen.

Als über längere Abschnitte unbrauchbar müssen dagegen die Messreihen **AALEN** (BW029, Abb.91), **WELZHEIMERWALD** (BW030, Abb.92) und **WAHNSDORF**

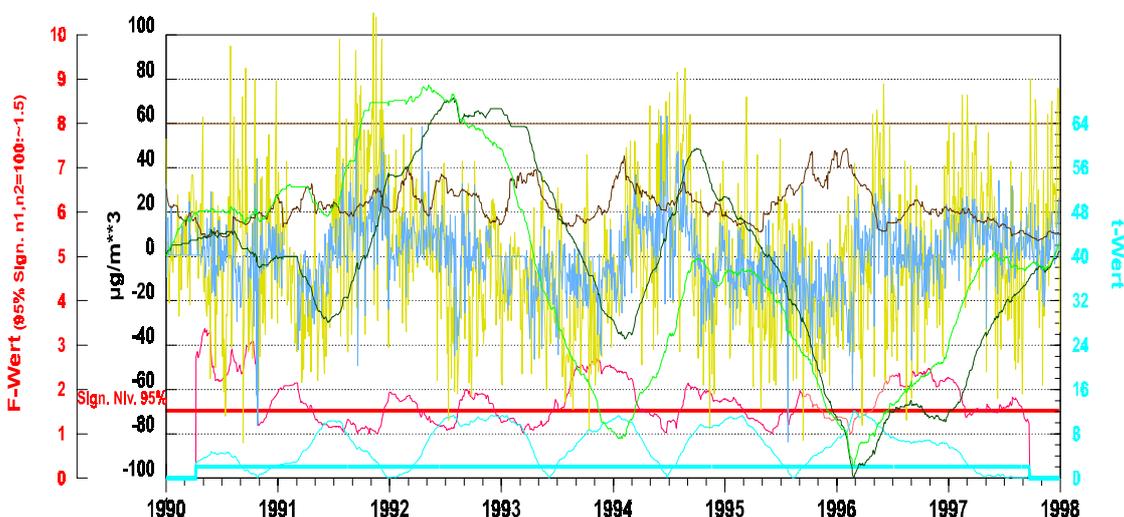


Abb. 91) Original-Ozonreihe der Station AALEN (BW029), Zeitraum 1990 – 1997 Messnetz Baden-Württemberg

(SN051, Abb.93) bezeichnet werden. So zeigt die Station **AALEN** im Messnetz des Landes Baden-Württemberg in nahezu regelmäßigen Abständen stufenförmige Interventionen, die ebenso regelmäßig mit Änderungen in der Varianz der Messwerte einher gehen. Die Summenkurven I und II zeigen diese Stufen ebenfalls, die auch als periodische Schwingungen interpretiert werden können, wofür jedoch keine hinreichenden Erklärungsgründe vorliegen.

Die Station **WELZHEIMERWALD** im Norden des Ballungsraumes Stuttgart zeigt zunächst einen nahezu vollständig parallelen Verlauf der Summenkurven I und II. Das zeigt an, daß alle Strukturbrüche in der Zeitreihe selbst enthalten sind. Gegen Ende des

Jahres 1990 ist ein deutlicher und signifikanter Bruch in der Variation der Messwerte ersichtlich, der jedoch nur von kurzer Dauer ist. Ein gravierender Fehler ist im Zeitraum 1992 und 1993 zu verzeichnen. In diesem Zeitraum ist sowohl die Varianz der Messreihe gestört als auch ein systematischer Fehler zu beobachten. Dieser Umstand wird anscheinend zu Beginn des Jahres 1993 bemerkt und das Gerät wird zu diesem Zeitpunkt abgeschaltet (Ausfallzeitraum).

Gegen Mitte des Jahres 1995 erfolgt eine stufenförmige Erhöhung der Messwerte bei einer gleichzeitigen Varianzänderung (Erhöhung der Varianz). Diese Stufe wird gegen Ende 1995 beendet. Insgesamt ist der Verlauf dieser Zeitreihe unter ursachenanalytischen Gesichtspunkten kaum zu interpretieren.

Die Ozonzeitreihe der Station **WAHNSDORF** bei Dresden leidet unter einer messtechnisch bedingten Unterbrechung im Jahr 1991. Noch im Jahr 1990 wurden die Messungen des ehemaligen Meteorologischen Dienstes der DDR weitergeführt. Mit dem Ausbau des Messnetzes des heutigen Landes Sachsen mit Beginn des Jahres 1992 erfolgte eine Änderung in der Geräteausstattung der vorhandenen Stationen, die in den meisten Fällen - so auch in diesem - mit einem Strukturbruch einher geht. Daß die neue Messtechnik offensichtlich noch nicht "im Griff" war, zeigt der lokale abfallende Trend der Messwerte im Jahr 1992. Dieser Trend kehrt sich gegen Ende dieses Jahres um und zeigt bis gegen Ende 1994 eine zunehmende Tendenz. Beide Trends sind signifikant, wie der Verlauf des **Gleitenden t-Wertes** anzeigt. Eine zusätzliche Stufe ist um die Mitte des Jahres 1996 erkennbar, die bis Mitte 1997 reicht. Alle stufenförmigen Interventionen in dieser Zeitreihe sind mit einer hochsignifikanten Änderung der Variation gekoppelt. Insgesamt gesehen kann diese Reihe allenfalls seit etwa 1994 interpretiert werden, was nicht allein dadurch verursacht ist, daß Messwerte für das Jahr 1991 nicht vorliegen und dass die Ausfallwerte vor der Anwendung zeitreihenanalytischer Verfahren durch Schätzwerte ersetzt werden müssen.

Der hier beschriebenen Prüfung wurden alle vorhandenen Ozonzeitreihen unterzogen. Es ist naheliegend, daß alle Ergebnisse der Validierungsstufe 3 hier nicht dargestellt werden können. Abschließend sei darauf hingewiesen, daß die vorliegenden Prüfverfahren Resultate erbringen, die einer zusätzlichen subjektiven Interpretation

bedürfen. Die Gefahr, bei der Überlagerung heterogener Effekte Fehlinterpretationen zu erhalten, sollte nicht übersehen werden.

5 Graphische Präsentation der Validierungsstufe 3

Die im vorangegangenen Abschnitt näher erläuterten Abbildungen wurden für folgende Zeiträume erstellt:

- April bis September 1990 bis 1997,
6-stündiges Mittel (13 Uhr bis 18 Uhr)
- April bis September 1990 bis 1997,
5-stündiges Mittel (2 Uhr bis 6 Uhr)
- April bis September 1980 bis 1989,
6-stündiges Mittel (13 Uhr bis 18 Uhr)
- Oktober bis März 1990 bis 1997,
6-stündiges Mittel (13 Uhr bis 18 Uhr)

Da wegen der Übersichtlichkeit der dargebotenen Kurven Farbabbildungen notwendig sind, kann das gesamte Material nicht in dem vorliegenden Zwischenbericht beigelegt werden. Die Abbildungen können aber auf Anfrage zur Verfügung gestellt werden. Auf die Darstellung der Winterhalbjahre und die der Nachtstunden für den Zeitraum 1980 bis 1989 sowie die der Nachtstunden für die Winterhalbjahre des Zeitraums 1990 bis 1997 wurde verzichtet, da der notwendige Aufwand unverhältnismäßig hoch wäre. Außerdem zeichnete sich ab, daß keine wesentlichen neuen Erkenntnisse zu erwarten waren.

Detaillierte Auswertungen der Abbildungen werden schwerpunktmäßig für den Zeitraum April bis September 1990 bis 1997, 6-stündiges Mittel (13 Uhr bis 18 Uhr) vorgenommen. An dieser Stelle können jedoch bereits einige generelle Aussagen gemacht werden:

- Die im Abschnitt 3 vorgestellte Methode der räumlichen Kontrolle der Messreihen lässt sich in den 90er Jahren besser als in den 80er Jahren anwenden. Dies liegt zum

einen an der deutlich höheren Stationsdichte und zum anderen in der verbesserten Qualität der Messreihen in den 90er Jahren.

- Messfehler in den Datenreihen prägen sich am deutlichsten in den Nachmittagsstunden der Sommermonate aus. In den Nachtstunden sind Fehler aufgrund geringerer Amplituden schwerer zu identifizieren.

- Die räumlichen Abhängigkeiten zwischen den Messreihen weisen im Sommer eine deutlich höhere Tagesamplitude auf als im Winter.

- Die Residuen nach kompletter Fehlwertersetzung sind im Winter deutlich kleiner als im Sommer, das Bestimmtheitsmaß hingegen weist im Sommer höhere Werte auf.

6 Subjektive und objektive Bewertung

Schon eine erste Analyse der in Phase 3 erzeugten Abbildungen macht deutlich, daß es erhebliche qualitative Unterschiede bezüglich der Güte der vorliegenden Ozonmessreihen gibt. Deshalb wurden exemplarisch für die Jahre 1990 bis 1998, die Monate April bis September als Mittel über diesen Zeitraum von 14 Uhr bis 18 Uhr sowohl subjektiv als auch objektiv bewertet. Dazu wurde eine Tabelle (Anlage, Tab. 1 "Subjektive und objektive Bewertung...") erstellt, die in Spalte 1 die Stationsbezeichnung, in Spalte 2 eine subjektive Bewertung (1 = sehr gut ... 5 = schlecht), in Spalte 3 die Restvarianz und in Spalte 4 das Bestimmtheitsmaß der komplett fehlwertersetzten Reihen enthält. Außerdem wurden mögliche Fehler in den Zeitreihen zeitlich und abhängig vom Typ näher spezifiziert und die Länge der untersuchten Reihe sowie größere Ausfälle dokumentiert.

Bei der Auswertung der Tabelle ergeben sich unerwartete Resultate:

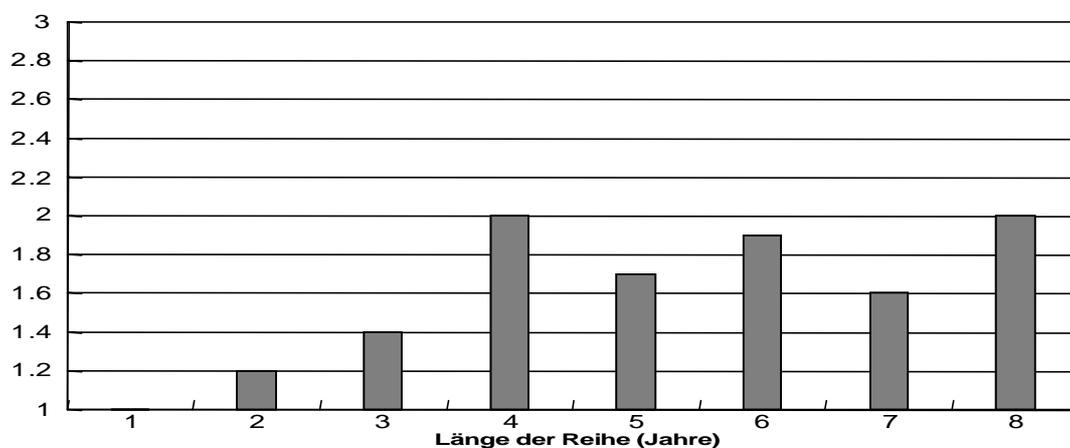
In Abb.94 sind der subjektive Bewertungsmaßstab, die Reststreuung (Streuung der Residuen nach kompletter Ersetzung der Reihen) und das Bestimmtheitsmaß zur Beurteilung der räumlichen Kopplung zwischen den Stationen in Abhängigkeit von der Länge der Messreihe dargestellt. Es zeigt sich besonders deutlich bei der Streuung der Residuen, daß die Messreihen um so besser sind, je kürzer die Reihe selbst ist. Zieht man in Betracht, daß die kurzen Zeitreihen vornehmlich ab 1995 zu finden sind, so spielt neben der verbesserten Messmethodik die größere räumliche Dichte der

Messstationen eine entscheidende Rolle. Bemerkenswert ist, daß Messreihen ≤ 3 Jahre bei der subjektiven Bewertung praktisch nicht beanstandet wurden.

Überraschend ist das Ergebnis einer länderspezifischen Validierung für den oben genannten Zeitraum (Abb.95). Nicht nur im subjektiven Bewertungsmaßstab, sondern auch bei der Reststreuung und dem Bestimmtheitsmaß zeigen sich deutliche Unterschiede zwischen den Bundesländern. Bei der Interpretation der Ergebnisse sollte man jedoch berücksichtigen, daß es sich bei den Daten mancher Bundesländer (z.B. Baden-Württemberg und den Messreihen des Umweltbundesamtes) um überwiegend lange Reihen mit den oben erläuterten Konsequenzen handelt.

Es sei an dieser Stelle jedoch betont, daß es sich um relative Unterschiede handelt, die Qualität der Reihen jedoch im Mittel als gut eingeschätzt werden kann. Mit dem beschriebenen Verfahren eröffnet sich nun die Möglichkeit, die Qualität und Zuverlässigkeit der Ozonmessreihen weiter zu verbessern. Die Vorgehensweise wird im folgenden Kapitel näher beschrieben.

subjektiver Bewertungsmaßstab



Bestimmtheitsmaß

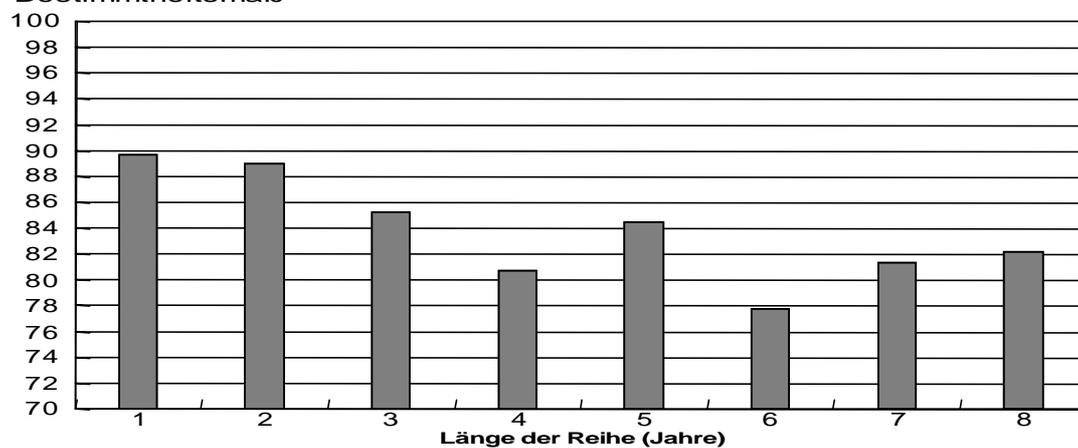
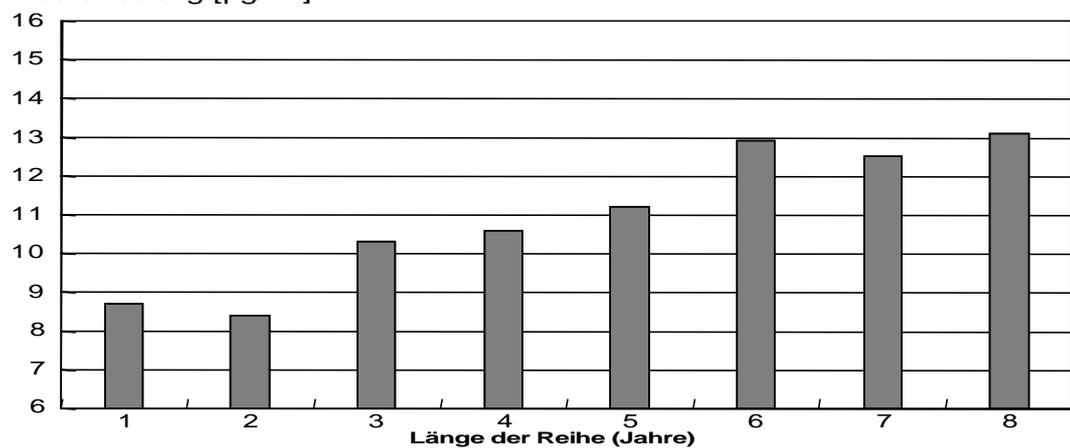
Reststreuung [$\mu\text{g}/\text{m}^3$]

Abb. 94) Validierung der Ozonmessreihen Stufe 3 in Abhängigkeit von der Länge der Messreihe, April bis September 1990 - 1997, 6-Stundenmittel (13 - 18 Uhr)

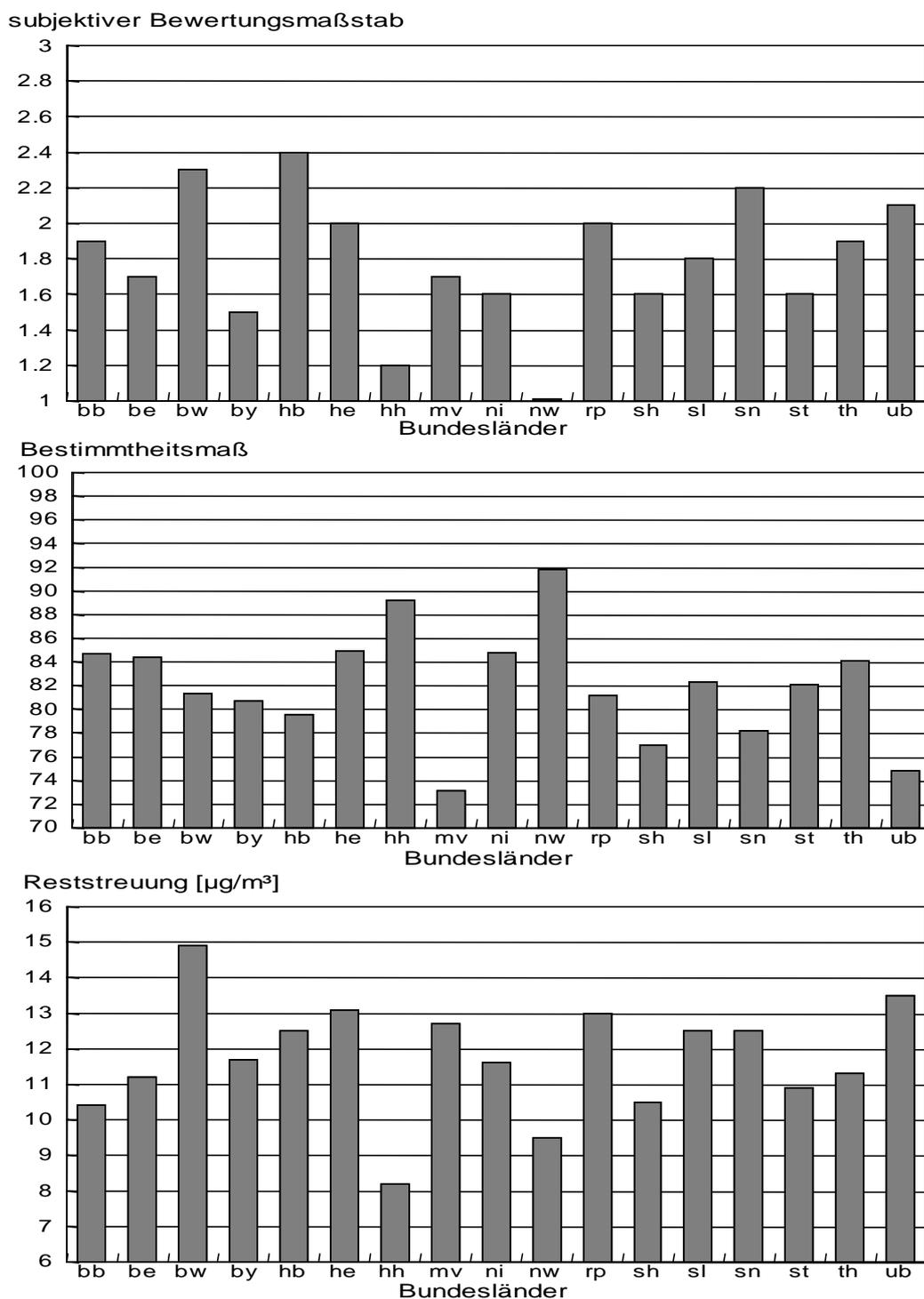


Abb. 95) Länderspezifische Validierung der Ozonmessreihen Stufe 3, April bis September 1990 - 1997, 6-Stundenmittel (13 - 18 Uhr)

7 Automatische Korrektur von Interventionen

Die Validierung der vorausgegangenen Etappen verdeutlicht die Notwendigkeit, in den Datenreihen deutlich vorhandene Brüche und Stufen vor einer Wetterbereinigung und Trendanalyse zu eliminieren. Verwendet werden hierzu die fehlwertersetzten und saisonbereinigten Datenreihen als Abweichung von den Jahresgängen, wie sie nach der Validierungsstufe 2 zur Verfügung stehen. Es wird in folgenden Schritten vorgegangen:

a) Der Gesamtdatensatz wird in die Zeiträume 01.10.1979 bis 30.09.1989 und 01.10.1989 bis 31.12.1997 jeweils in Sommer- und Winterhalbjahr (April bis September bzw. Oktober bis März) untergliedert. Dies trägt dem Umstand Rechnung, daß:

die Güte und Zahl der Ozonmessreihen nach 1990 deutlich zugenommen hat und damit räumliche Abhängigkeiten besser sind als vor 1990 und die Verhältnisse zwischen Sommer- und Winterhalbjahr sehr verschieden sind, was auf die unterschiedlichen räumlichen Abhängigkeiten zwischen den Stationen zurückzuführen ist.

b) Zur Stabilisierung der Aussagen über mögliche Brüche und Stufen in den Zeitreihen, deren Ursachen mit größter Wahrscheinlichkeit messtechnisch bedingt sind, werden jeweils 5-stündige Mittelwerte berechnet. Dadurch werden zufällige Schwankungen in den Messreihen stärker geglättet, so daß systematische Fehler besser hervortreten.

c) Alle Datensätze werden von linearen Trends bereinigt, um realistische räumliche Abhängigkeiten zwischen den Messstationen zu modellieren. Ausgewählte Reihen sind in den Abb.96 bis Abb.100 dargestellt.

d) Für die zu korrigierende Reihe wird in Analogie zu Punkt c eine komplette "Fehlwertersetzung" unter Nutzung der Ozonzeitreihen aller Stationen als potentielle Einflussgrößen vorgenommen. Die Selektion der signifikanten Einflußfaktoren wird durch eine stufenweise screening Regressionsanalyse erreicht.

e) Erneute Trendbereinigung der Differenz zwischen fehlwertersetzter und originaler Reihe (saisonbereinigte Abweichungen, im weiteren als Zeitreihe der Residuen bezeichnet). Diese Zeitreihe ist in den Abb.96 bis Abb.100 als 'Abw.-Ers.' bezeichnet und stellt die tägliche Abweichung zwischen fehlwertersetztem und beobachtetem 5-stündigen Mittel dar (schwarze Kurve in der Mitte der Y-Achse). Für eine Messreihe ohne wesentliche Brüche und Stufen, wie sie z.B. für die Stationen Spremberg-Süd (BB026, Abb.96) oder Arzberg (BY002, Abb.97) zu finden ist, variiert die Zeitreihe der Residuen naturgemäß um Null. Hingegen zeigen die Residuen der Messreihen von den Stationen Frankenberg (HE027, Abb.98) und Wahnsdorf (SN051, Abb.100) einen unruhigen Verlauf. Die Station Stuttgart-Mitte (BW012, Abb.99) weist bei einem im allgemeinen ruhigen Verlauf im Jahr 1993 einen deutlich erkennbaren Fehler auf.

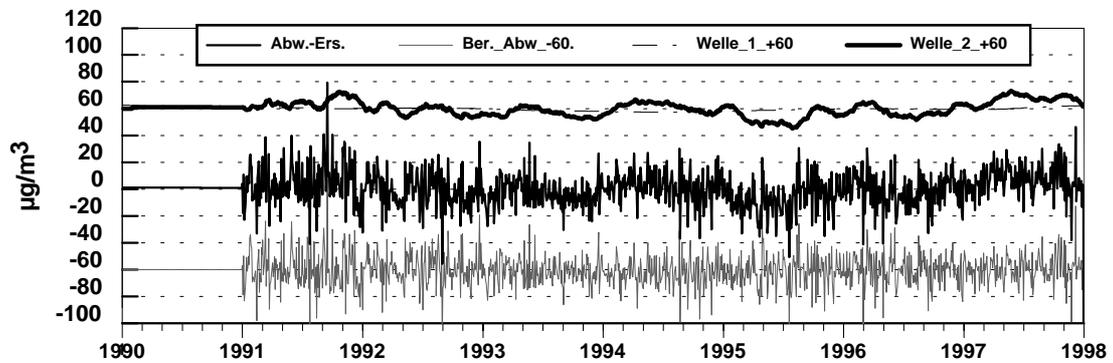
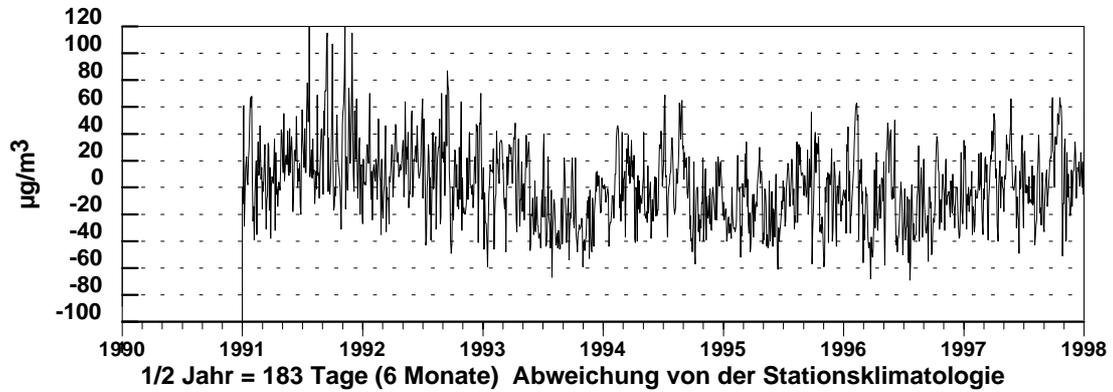
f) Zur Eliminierung langperiodischer Schwingungen aus der Zeitreihe der Residuen wird ein gleitendes Mittel der Residuen über 200 Tage berechnet und von der Originalreihe subtrahiert. Dies erweist sich als notwendig, um mögliche Trends, die sich bei mehreren Stationen in einem anfänglichen Anstieg der Ozonwerte und einer Abnahme zum Ende des betrachteten Zeitabschnittes äußern, in den Reihen zu belassen. In den Abb.96 bis Abb.100 ist diese Welle (mit "Welle 1+60" bezeichnet) jedoch nicht von Relevanz.

g) Das 30-tägige Mittel der vom linearen Trend und von langperiodischen Schwankungen bereinigten Reihe der Residuen, in den Abb.96 bis Abb.100 als "Welle 2+60" bezeichnet, wird als der systematische Messfehler betrachtet. Zur besseren Darstellung wurden diese Kurve um $60\mu\text{g}/\text{m}^3$ nach oben verschoben. Bei der Station Arzberg und Spremberg-Süd schwankt diese Zeitreihe, wie nicht anders zu erwarten, unwesentlich um die Nulllinie (fette schwarze Kurve oben). Deutliche systematische Fehler werden hingegen bei der Station Stuttgart-Mitte und Wahnsdorf signalisiert.

h) Der systematische Fehler wird nun aus den Originaldaten (nach Validierungsstufe 2) entfernt und wie unter Punkt b bis d eine erneute räumliche Fehlwertersetzung der kompletten Zeitreihe vorgenommen. Die Reihe der nun berechneten Residuen, in den Abbildungen als "Ber._Abw._-60" bezeichnet (Mitte, graue Kurve), zeigt den

erwarteten ruhigen Verlauf und führt zudem, wie aus den Tabellen in den Abb.96 bis Abb.100 ersichtlich ist, zu einer deutlichen Reduktion der Reststreuung "var2" (Streuung der Residuen nach der Beseitigung systematischer Fehler) gegenüber "var1" (Streuung der Residuen vor der Beseitigung systematischer Fehler). siehe die darunter liegende Tabelle. Für die Station Arzberg ergibt sich lediglich eine Verbesserung der Reststreuung von $9,7\mu\text{g}/\text{m}^3$ auf $8,8\mu\text{g}/\text{m}^3$ (Abb.97). Bei der mit einem deutlichen Fehler (Jahr 1993) behafteten Reihe Stuttgart-Mitte reduziert sich hingegen die Reststreuung von $13,8\mu\text{g}/\text{m}^3$ auf $9,5\mu\text{g}/\text{m}^3$ (Abb.99).

Bei der Beseitigung der systematischen Fehler aus den Ozonreihen besteht zusätzlich ein bisher noch nicht beachtetes Problem: Es ist nicht eindeutig, ob der gefundene systematische Fehler der betrachteten Reihe selbst zuzuordnen ist oder ob er über die räumliche Verknüpfung mit einer fehlerbehafteten anderen Messreihe vorgetäuscht wird. Deshalb ist es sinnvoll, das Verfahren (Punkte a bis h) erneut anzuwenden. Dabei werden systematische Fehler, die von anderen Messreihen stammen, wieder korrigiert. In den meisten Fällen ergeben sich jedoch nur noch marginale Änderungen (vgl. Abb.96 bis Abb.100, unterer Kasten).



**Güte
der
horizontalen
Ersetzung:**

| | Bezeichnung der verwendeten Prediktoren | Faelle | var1 | var2 |
|-----|---|--------|------|------|
| 28 | BE027_MC27_Marienfelde-Schichau | 1281 | 15.9 | 14.1 |
| 459 | UB031_Lueckendorf | 1281 | 15 | 13.1 |
| 458 | UB030_Neuglobsow | 1281 | 14.7 | 12.5 |
| 63 | BW029_Aalen | 1281 | 14.1 | 12.1 |
| 175 | HE003_Frankfurt-Bockenheim | 1281 | 13.6 | 12 |
| 321 | SH007_Kiel-Schuetzenwall_Verk. | 1281 | 13.3 | 11.8 |
| 213 | HH008_Sternschanze | 1281 | 13.2 | 11.7 |
| 364 | SN051_Radebeul-Wahnsdorf | 1281 | 13 | 11.6 |

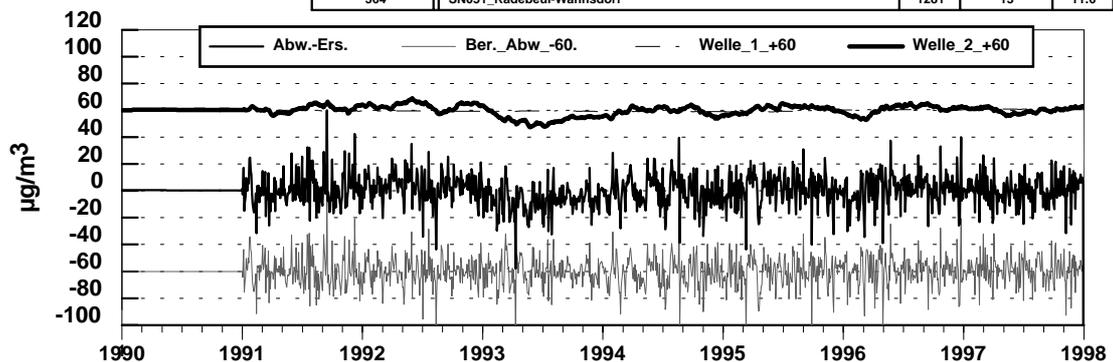
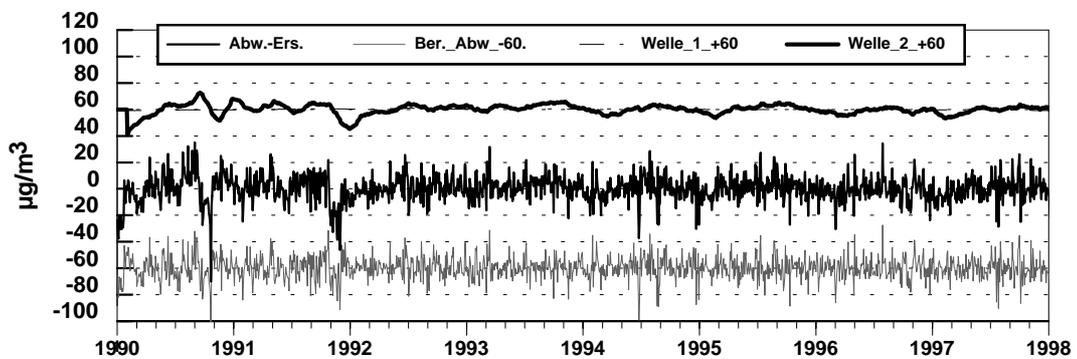
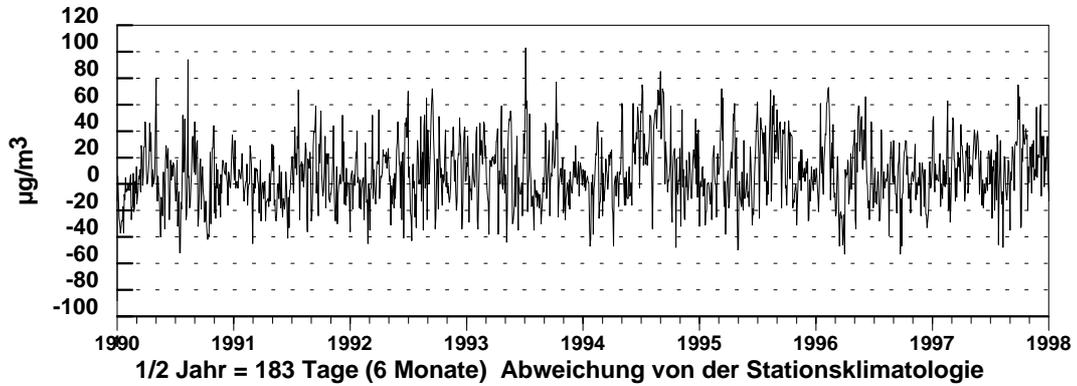


Abb. 96: Validierung der Ozon- Meßreihe BB026 - Spremberg vor- und nach der Bereinigung von Strukturbrüchen, April bis September, 5-stündiges Mittel (13 bis 17 Uhr)



Güte der horizontalen Ersetzung:

| | Bezeichnung der verwendeten Prediktoren | Faelle | var1 | var2 |
|-----|---|--------|------|------|
| 131 | BY020_Hof | 1464 | 11.2 | 10.1 |
| 150 | BY072_Tiefenbach | 1464 | 10.7 | 9.6 |
| 26 | BE014_MC14_Charlottenburg-Lersch | 1464 | 10.2 | 9.1 |
| 142 | BY053_Nuernberg-Willi-B.-P.(Mari) | 1464 | 10.1 | 9 |
| 39 | BW002_Karlsruhe-West | 1464 | 10 | 8.9 |
| 180 | HE008_Frankfurt-Ost | 1464 | 9.9 | 8.9 |
| 185 | HE014_Kassel-Nord | 1464 | 9.9 | 8.8 |
| 435 | UB004_Schauinsland | 1464 | 9.8 | 8.8 |

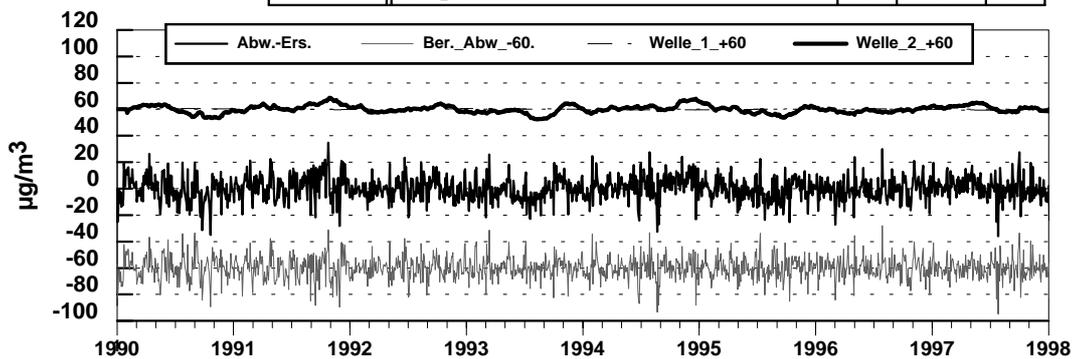
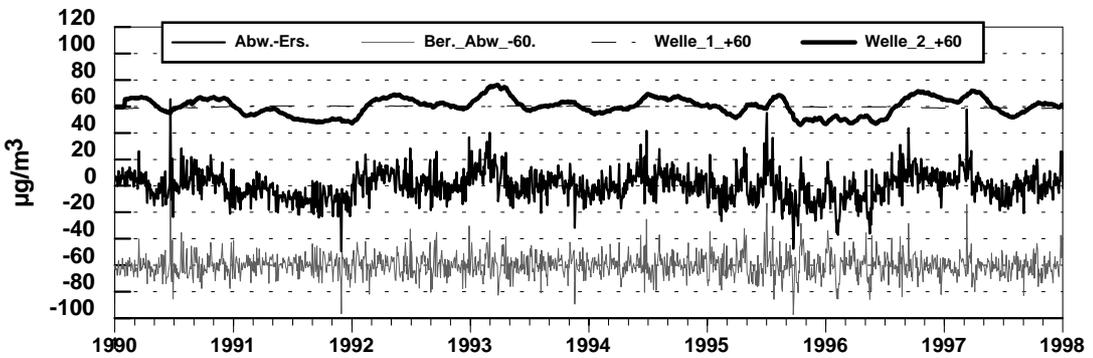
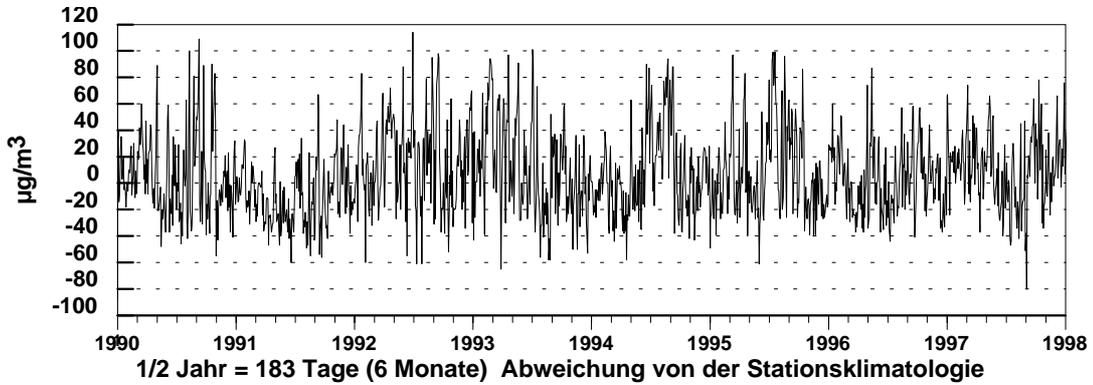


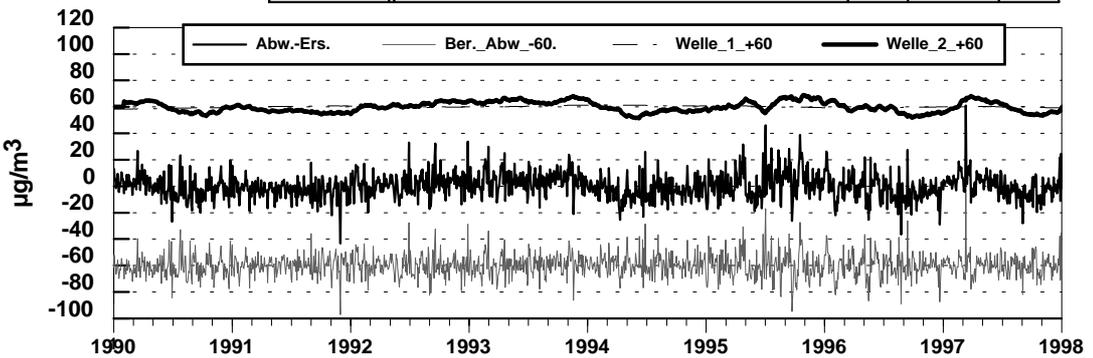
Abb. 97: Validierung der Ozon- Meßreihe BY002 - Arzberg vor- und nach der Bereinigung von Strukturbrüchen, April bis September, 5-stündiges Mittel (13 bis 17 Uhr)



/2 Jahr = 183 Tage (6 Monate) 1. Iteration

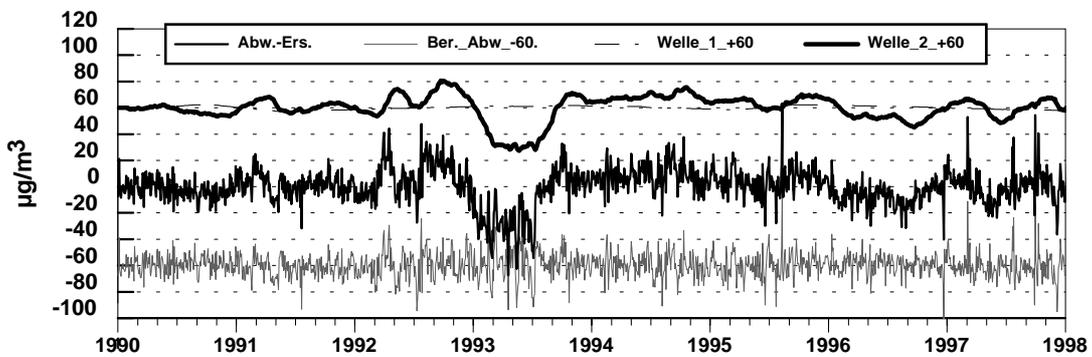
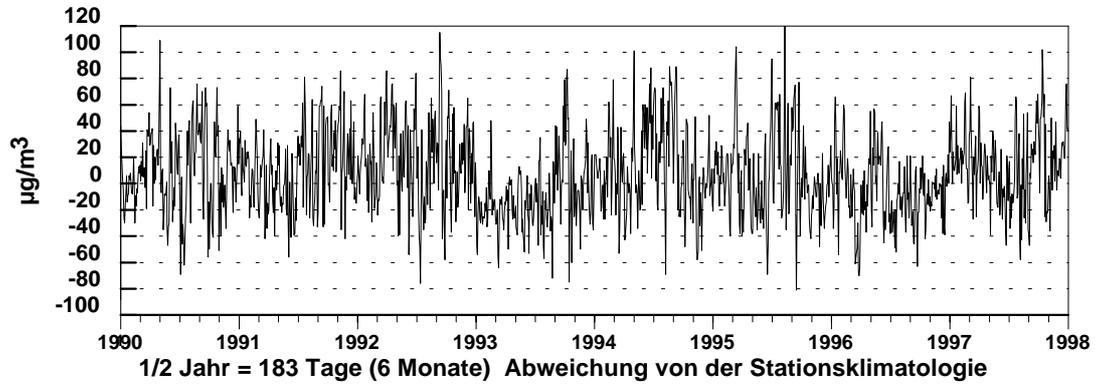
Güte
der
horizontalen
Ersetzung:

| | Bezeichnung der verwendeten Prediktoren | Faelle | var1 | var2 |
|-----|---|--------|------|------|
| 204 | HE033_Borken | 1464 | 12.7 | 9.8 |
| 248 | NI043_Lingen | 1464 | 12.4 | 9.5 |
| 26 | BE014_MC14_Charlottenburg-Lersch | 1464 | 12.1 | 9.3 |
| 196 | HE025_Koenigstein | 1464 | 11.9 | 9.2 |
| 203 | HE032_Bebra | 1464 | 11.7 | 9 |
| 284 | NW065_Rothaargebirge | 1464 | 11.5 | 8.9 |
| 76 | BW042_Bernhausen | 1464 | 11.3 | 8.8 |
| 47 | BW011_Stuttgart-Zuffenhausen | 1464 | 11.2 | 8.8 |



/2 Jahr = 183 Tage (6 Monate) 2. Iteration

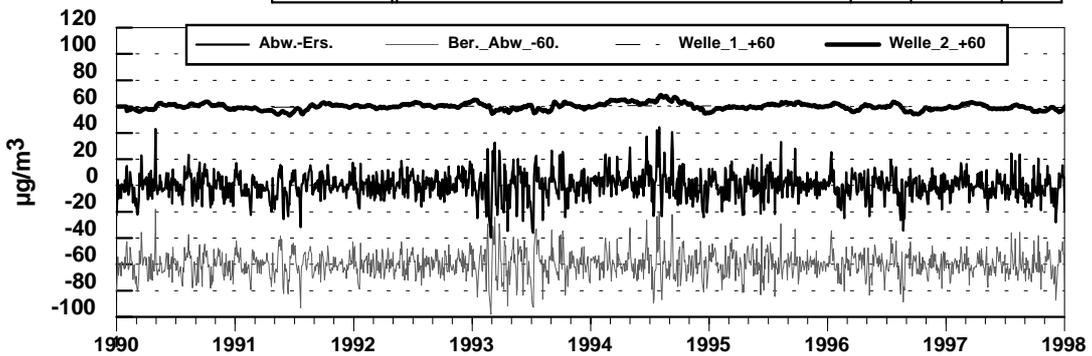
Abb. 98: Validierung der Ozon- Meßreihe HE027 - Frankenberg vor- und nach der Bereinigung von Strukturbrüchen, April bis September, 5-stündiges Mittel (13 bis 17 Uhr)



1/2 Jahr = 183 Tage (6 Monate) 1. Iteration

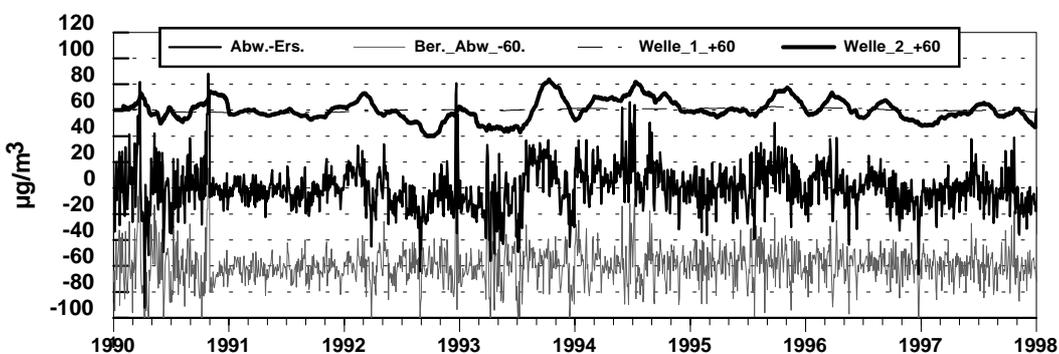
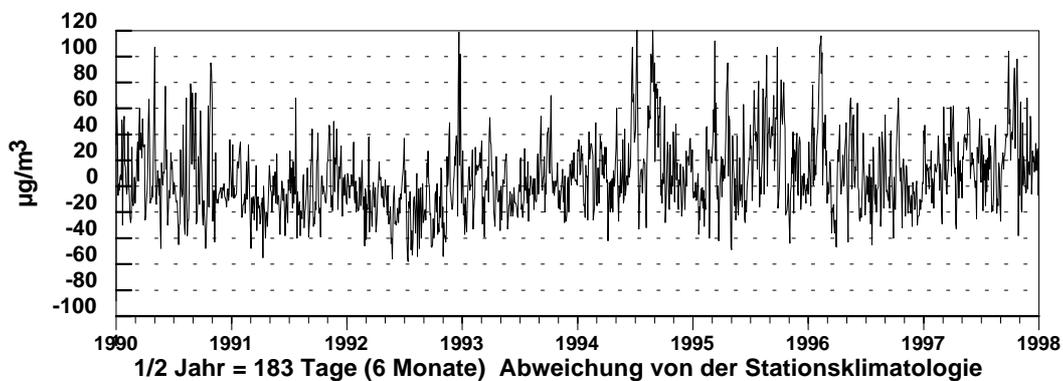
Güte
der
horizontalen
Ersetzung:

| | Bezeichnung der verwendeten Prediktoren | Faelle | var1 | var2 |
|-----|---|--------|------|------|
| 58 | BW024_Ludwigsburg | 1464 | 15.2 | 10.3 |
| 67 | BW033_Pforzheim-West | 1464 | 14.9 | 10 |
| 442 | UB013_Ansbach_(W_IX) | 1464 | 14.6 | 9.8 |
| 72 | BW038_Friedrichshafen | 1464 | 14.3 | 9.7 |
| 197 | HE026_Spessart | 1464 | 14.2 | 9.6 |
| 50 | BW014_Stuttgart-Hafen | 1464 | 14 | 9.5 |
| 52 | BW016_Neckarsulm | 1464 | 13.9 | 9.5 |
| 75 | BW041_Rheinfelden | 1464 | 13.8 | 9.4 |



1/2 Jahr = 183 Tage (6 Monate) 2. Iteration

Abb. 99: Validierung der Ozon- Meßreihe BW048 - Stuttgart vor- und nach der Bereinigung von Strukturbrüchen, April bis September, 5-stündiges Mittel (13 bis 17 Uhr)



**Güte
der
horizontalen
Ersetzung:**

| | Bezeichnung der verwendeten Prediktoren | Faelle | var1 | var2 |
|-----|---|--------|------|------|
| 366 | SN053_Fichtelberg_DWD | 1464 | 19.3 | 17 |
| 28 | BE027_MC27_Marienfelde-Schichauw | 1464 | 17.9 | 15.4 |
| 298 | RP014_Hunsr ck-Leisel_Idar | 1464 | 17.8 | 15.3 |
| 48 | BW012_Stuttgart-Mitte | 1464 | 17.5 | 15.2 |
| 56 | BW022_Kehl-Hafen_(Kehl) | 1464 | 17.3 | 15.1 |
| 434 | UB003_Brotjacklriegel | 1464 | 17.2 | 15 |
| 362 | SN049_Carlsefeld_DWD | 1464 | 17.1 | 15 |
| 232 | NI014_Buedenstedt/Reinsdorf | 1464 | 17 | 14.9 |

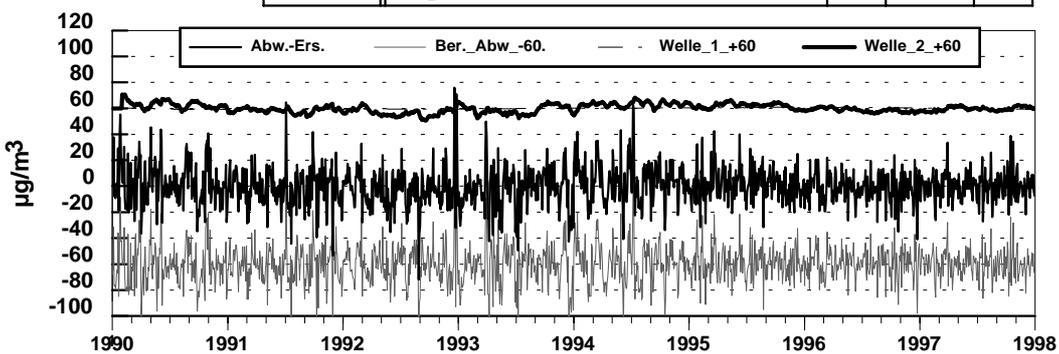


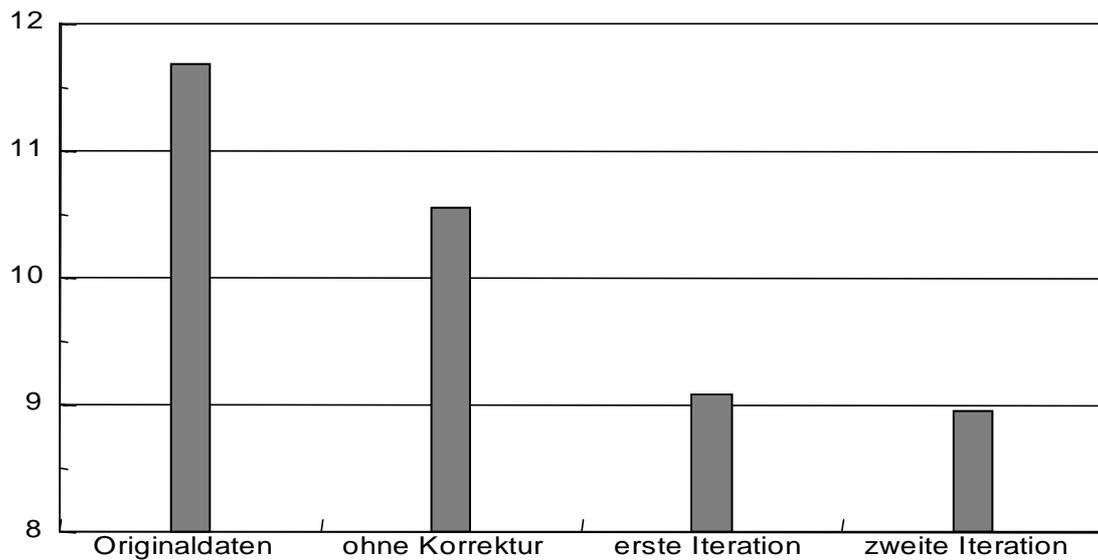
Abb.100: Validierung der Ozon- Meßreihe SN051 - Wahnsdorf vor- und nach der Bereinigung von Strukturbrüchen, April bis September, 5-stündiges Mittel (13 bis 17 Uhr)

Am Beispiel der Sommerhalbjahre 1990 bis 1997 wurde die Bereinigung der dritten Stufe für 15 Uhr, also dem Mittelwert zwischen 13 Uhr und 17 Uhr, für 470 Stationen durchgeführt. Die Ergebnisse sind in Abb.101 zusammengefaßt. Verwendet man die nicht fehlwertersetzten Datenreihen (Originaldaten), so beträgt die mittlere Streuung der Residuenreihen $11,6\mu\text{g}/\text{m}^3$. Bei Verwendung der fehlwertersetzten Reihen (ohne Korrektur) beträgt die mittlere Streuung der Residuenreihen $10,6\mu\text{g}/\text{m}^3$. Diese Abnahme der Reststreuung wird plausibel durch die Vorgehensweise der räumlichen Fehlwertersetzung - es werden die gleichen Ozonmessreihen verwendet wie später zur Validierung der Stufe 3. Die Beseitigung von Brüchen und Stufen in der ersten Iterationsstufe führt zu einer weiteren deutlichen Abnahme der mittleren Reststreuung auf $9,1\mu\text{g}/\text{m}^3$. Eine erneute Anwendung der Validierungsprozedur der Stufe 3 in einer zweiten Iteration führt hingegen nur zu einer marginalen zusätzlichen Verbesserung der mittleren Reststreuung, was für die Wirksamkeit des Verfahrens schon bei einmaliger Anwendung (1. Iteration) spricht. Es wäre ja auch denkbar, daß Brüche in Messreihen Auswirkungen der Fehler anderer Messreihen sind, die über die räumliche Verknüpfung in der untersuchten Reihe nur vorgetäuscht werden. Dies scheint jedoch nur selten der Fall zu sein. Ein weiterer Indikator dafür, daß die auftretenden Brüche und Stufen in den Messreihen keine "natürlichen" Ursachen haben, stellt das Verfahren der Wetterbereinigung dar. Man kann davon ausgehen, daß das reale Wettergeschehen keinen Einfluss auf die Qualität der Messreihen hat. Für die in Abb.96 bis Abb.100 ausgewählten Stationen wurde deshalb eine Wetterbereinigung durchgeführt (siehe erster Zwischenbericht). Das Bestimmtheitsmaß bzw. die Reduktion der Varianz gegenüber der Ausgangsreihe mit und ohne Beseitigung der Brüche und Stufen ist aus Abb.102 und Abb.103 zu entnehmen. Daraus lassen sich folgende Aussagen gewinnen:

- Eine Korrektur der Ozonmessreihen mit Stufe 3 der Validierung führt immer zu einer Verbesserung der Wetterbereinigung
- Der Qualitätsgewinn ist um so größer, je stärker die Reihe mit systematischen Fehlern behaftet war.

- Auch in der zweiten Iterationsstufe ist noch ein Qualitätsgewinn bezüglich Wetterbereinigung zu erreichen. Er fällt jedoch erwartungsgemäß deutlich geringer aus.

Reststreuung $\mu\text{g}/\text{m}^3$



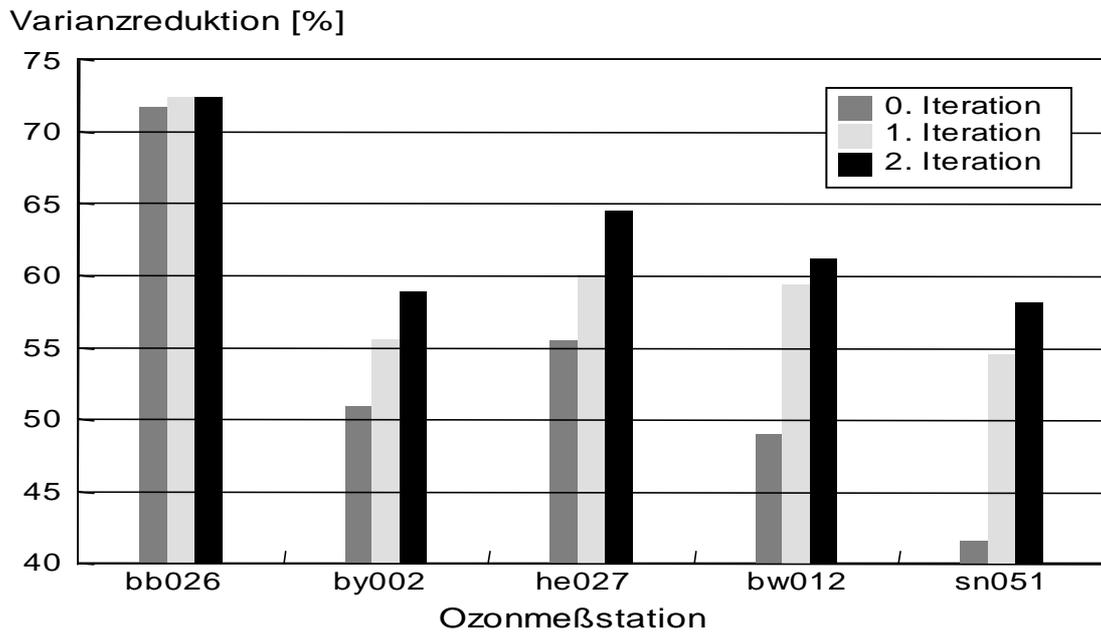
Originaldaten: Verwendung der Datensätze ohne Fehlwertersetzung

ohne Korrektur: Verwendung der fehlwertersetzten Meßreihen

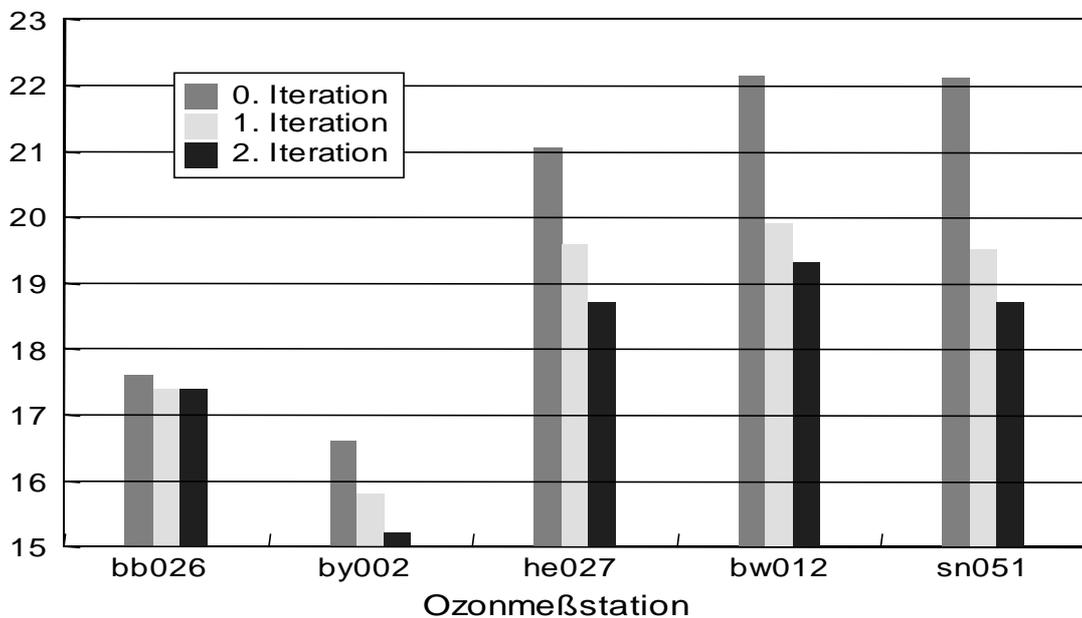
erste Iteration: Brüche und Sprünge werden in einer ersten Stufe eliminiert (Validierung Stufe 3)

zweite Iteration: Brüche und Sprünge werden in einer zweiten Stufe eliminiert

Abb. 101) Gewichtete mittlere Reststreuung bei kompletter Fehlwertersetzung als Mittel über 470 Stationen, 1990 bis 1997, April bis September, 5-stündiges Mittel (13 - 17 Uhr)



**Abb. 102) Reduktion der Varianz nach Wetterbereinigung 1990 bis 1997,
April bis September, 5-stündiges Mittel (13 - 17 Uhr)**



**Abb. 103) Restreung nach Wetterbereinigung als Mittel über 470 Stationen,
1990 bis 1997, April bis September, 5-stündiges Mittel (13 - 17 Uhr)**

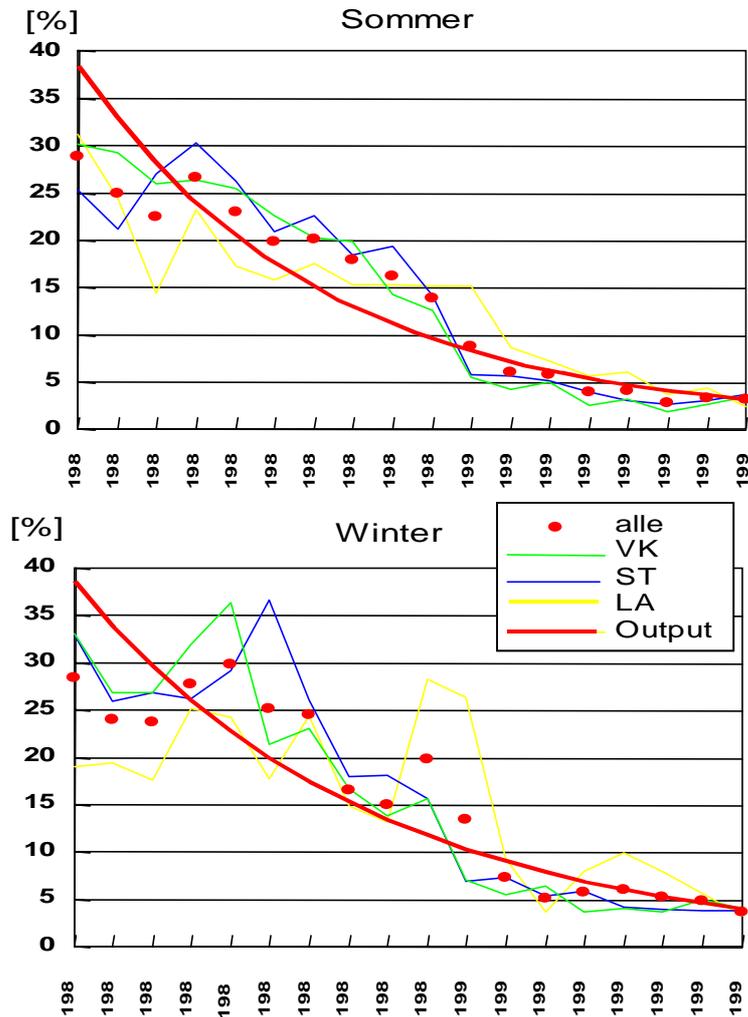


Abb. 104) Approximierter Messfehler, abgeschätzt aus dem Verhältnis der Streuung von LIMBA21 zu LIMBA26 im zeitlichen Verlauf als Mittel über alle Stunden

Nach den bisherigen Untersuchungen empfiehlt es sich, die Validierung der Stufe 3 für den gesamten Zeitraum 1980 - 1997, alle Tageszeiten und Messreihen durchzuführen, da anderenfalls die in der nachfolgenden Projektphase zu bestimmenden Trends nachhaltig verfälscht sein können. Die saisonbereinigten Ozonreihen weisen eine Variabilität auf, die zu einem Teil auf natürliche Schwankungen der Ozonkonzentration basiert und zu einem weiteren Teil durch zufällige und systematische Messfehler hervorgerufen ist. Werden diese Messfehler wie oben beschrieben eliminiert, so kann das Verhältnis saisonbereinigter Originalreihe zur fehlerbereinigten Reihe eine quantitative Approximation des Fehlers darstellen. Dieses prozentuale Verhältnis für

jedes Messjahr wird in der Abb.104 als Mittel über alle Messstationen und Stunden getrennt nach Sommer- und Winterhalbjahr dargestellt. Hierbei wird nach den Stationskategorien "Verkehr", "Stadt" und "Land" unterschieden. Die durchgehende Kurve stellt eine exponentielle Glättungsfunktion über alle Stationstypen dar. Es zeigt sich eine deutliche Steigerung der Qualität der Messreihen im Messzeitraum. Besonders auffällig ist ein Qualitätssprung im Jahr 1990, deren Ursache im Rahmen dieser Arbeit nicht geklärt werden kann. In den 90er Jahren beträgt der approximierte Messfehler etwa 5% der Variabilität der Originalreihen, während dieser Fehler zu Beginn des Untersuchungszeitraumes bei etwa 30% lag.

Diese Ergebnisse legen nahe, die Methodik in ein operationell anwendbares Verfahren umzusetzen, das eine weitgehend automatische Datenkontrolle und Korrektur am aktuellen Rand der Zeitreihe gestattet. Ein solches Verfahren bringt zusätzlich eine Optimierung der Betreuung der Messstationen im Sinne einer Qualitätssicherung.

Terminologie:

Datenkollektive:

Als Datenkollektive werden im vorliegenden Fall die Meßdaten einer Meßkomponente (z. B. Ozon) einer Meßstation bezeichnet. Hierbei handelt es sich um die Einstundenwerte eines Messjahres vom 01. Januar bis 31. Dezember, die in einer Matrix (Spalten = Uhrzeit, Zeilen = Tage) vorliegen.

Originalwerte:

Als Originalwerte werden in diesem Zusammenhang die saison- und trendbereinigten (Tagesgang, Jahresgang) Einzelwerte einer Ozonmessreihe bezeichnet. Aufgrund der Saison- und Trendbereinigung schwanken diese Werte um „Null“.

Erwartungswerte:

Erwartungswerte (saison- und trendbereinigt) werden als Ersatz für ausgefallene Originalwerte (Datenlücken) verwendet. Sie werden (im vorliegenden Fall) aus zeitgleichen Datenkollektiven exogener Variablen, die an der gleichen oder einer anderen Station gemessen werden (z. B. meteorologische Variablen) oder aus zeitgleichen Datenkollektiven der Ozonkonzentration, die an anderen Stationen gemessen werden, berechnet.

Zeitreihe:

Eine Zeitreihe besteht aus einer Reihe zeitlich aufeinanderfolgender Daten, die in äquidistanten Zeitschritten erhoben wurden.

Datenlücken:

Datenlücken sind in nahezu allen Zeitreihen eines Luftschadstoffes enthalten. Sie entstehen durch Kalibrierungen, Wartungen und Reparaturen des Meßgerätes, durch die Eliminierung unplausibler Meßwerte oder auch durch unvollständige Zeitreihen zu Beginn und gegen Ende der Messungen.

Saisonale Schwankungen:

Saisonale Schwankungen (hier: Jahresgang und Tagesgang, ggf. auch Wochengang der Ozonkonzentration) gehören zu den systematischen, zyklischen Schwankungen, die für die Ozonkonzentration charakteristisch sind. Vor der Berechnung von Regressionen und Korrelationen werden diese Schwankungen aus den Datenkollektiven entfernt. Hierdurch werden „Scheinkorrelationen“ vermieden.

Trend:

Trends gehören zu den systematischen, langfristigen Schwankungen. Zur Bewertung von Tendaussagen ist das Vorliegen hinreichend langer Zeitreihen erforderlich, um eine Fehlinterpretation zu vermeiden, die sich dann ergibt, wenn es sich bei vermuteten Trends tatsächlich um den an- oder absteigenden Teil einer zyklischen Schwankung oder gar um einen Teil einer zufälligen Schwankung handelt.

Residuen:

Als Residuen einer Zeitreihe werden die zufälligen Schwankungen bezeichnet, die nach der Eliminierung systematischer (regulärer) Anteile der Gesamtschwankung mittels geeigneter statistischer Verfahren zurückbleiben (irreguläre Schwankungen). Im Idealfall, d.h., weitere Anteile der irregulären Schwankungen können durch kein weiteres statistisches Verfahren mehr erklärt werden, bilden die Residuen das sogenannte „Weiße Rauschen“, das im gesamten und in Teilbereichen der Zeitreihe den Mittelwert = 0 und die (normierte) Standardabweichung = 1 aufweist.

Differenz:

Mit Differenz wird im vorliegenden Zusammenhang die Zeitreihe bezeichnet, die sich aus der Differenz der Originalwerte und der Erwartungswerte ergibt.

t-Wert:

Mit dem t-Test wird die (Null-)Hypothese geprüft, daß zwei Mittelwerte einer gemeinsamen Grundgesamtheit entstammen. Überschreitet der berechnete t-Wert eine vorgegebene Signifikanzschränke, so ist die Null-Hypothese zu verwerfen und die Mittelwerte entstammen mit einer bestimmten Wahrscheinlichkeit unterschiedlichen Grundgesamtheiten.

F-Wert:

Mit dem F-Test wird die (Null-)Hypothese geprüft, daß zwei Streuungen einer gemeinsamen Grundgesamtheit entstammen. Überschreitet der berechnete F-Wert eine vorgegebene Signifikanzschränke, so ist die Null-Hypothese zu verwerfen und die Streuungen entstammen mit einer bestimmten Wahrscheinlichkeit unterschiedlichen Grundgesamtheiten.