

5 Kovarianzmodellierung der Immissionsfelder

Im Kapitel 4.5 wurde die statistische Analyse der Immission auf Basis der Kovarianzmatrizen von Background und Beobachtungen beschrieben, ohne Hinweise über deren Gewinnung zu geben. Die korrekte Schätzung dieser statistischen Momente, d. h. \mathbf{HBH} , \mathbf{HB} und \mathbf{R} und der Erwartungswerte, ist das Hauptproblem bei der Analyse. Sie werden mit Hilfe von Kovarianzmodellen geschätzt und ausgedrückt. Das Kapitel beinhaltet einen Überblick über die Kovarianzmodellierung und stellt die für diese Arbeit entwickelten Kovarianzmodelle vor. Die Varianz des Beobachtungsfehlers \mathbf{R} wird häufig durch subjektive Abschätzung festgelegt. Die Beobachtungsmethode (siehe Kapitel 5.2.6) stellt eine Möglichkeit dar, den Beobachtungsfehler mit Hilfe eines Kovarianzmodells des Backgrounds und der Varianz der Beobachtungsinkremente zu bestimmen. Im Kapitel 5.3 werden die Ergebnisse vorgestellt und diskutiert.

5.1 Kovarianzmodelle für räumliche Prozesse

Da der betrachtete räumliche Zufallsprozess X bzw. seine Diskretisierung \mathbf{x} nur unzureichend bekannt ist, müssen seine Momente aus Stichproben geschätzt werden. Das Ergebnis der Schätzung ist ein Kovarianzmodell, das alle benötigten Kovarianzen für die Analyse liefert. Dieses Kovarianzmodell beruht neben der Stichprobeninformation auf vereinfachenden Annahmen, um die Information der Stichprobe z. B. auf Orte ohne Messinformation zu erweitern. Die Unterschiede dieser Annahmen sind entscheidend für die verschiedenen Typen von Kovarianzmodellen. In einem weiteren Sinne wird in dieser Arbeit auch die Ableitung der Beobachtungsfehler mit Kovarianzmodellen als Kovarianzmodellierung bezeichnet.

Die Kovarianzmatrix \mathbf{B} ist für große Zustandsvektoren nicht handhabbar¹²¹. Die zusammenfassende Beschreibung von \mathbf{B} stellt weiterhin eine wichtige Aufgabe des Kovarianzmodells dar, auch wenn \mathbf{B} vollständig bekannt wäre.

Von entscheidender Bedeutung für die Kovarianzmodellierung und die gesamte Analyse ist die Bestimmung des Erwartungswertes. Im Falle der Datenassimilation fungiert der Background, d. h. das biasfreie Modellfeld, als Erwartungswert. Die Bestimmung des Bias ist damit genauso wie die Schätzung des Beobachtungsfehler Bestandteil der Kovarianzmodellierung.

Es gibt unterschiedlichste Varianten der Kovarianzmodelle, die sich durch ihre Komplexität, ihren Aufwand und ihren Erfolg bei der Analyse unterscheiden. Im Folgenden wird versucht, eine Klassifizierung der Einteilungsprinzipien von Kovarianzmodellen zu geben. Kovarianzmodelle können unterteilt werden nach der Art

¹²¹ das Quadrat der Anzahl der Elemente des Zustandsvektors

- ihrer Gewinnung (empirisch – dynamisch)
- des wahrscheinlichkeitstheoretischen Ansatzes (bayessch)
- ihrer Darstellung (parametrisch – nichtparametrisch)
- der zugrundeliegenden Annahmen hinsichtlich des Prozesses (homogen, isotrop, normalverteilt etc.)
- der räumlichen oder zeitlichen Aggregation des Stichprobenmaterials (terminbezogen oder klimatisch¹²²)

Empirische oder dynamische Kovarianzmodelle

Empirische Kovarianzmodelle werden mit Hilfe von empirisch geschätzten Kovarianzwerten gewonnen. Dynamische Kovarianzmodelle sind aus Kenntnissen der zugrundeliegenden physikalischen Vorgänge abgeleitet (Balgovind et al., 1983). Beim Kalman-Filter wird die Kovarianzmatrix dynamisch mit einem deterministischen Modell bestimmt.

Nichtbayessche (klassische) oder bayessche Kovarianzmodelle

Die klassischen Methoden beruhen auf der Annahme, dass die Kovarianzfunktion unbekannt, aber feststehend ist und durch empirische Daten geschätzt werden kann. Bayesianische Ansätze erlauben es, eine Wahrscheinlichkeitsfunktion für die Parameter der Kovarianzfunktion zu berücksichtigen.

Parametrische oder nichtparametrische Kovarianzmodelle

Parametrische Kovarianzmodelle gehen von einer bestimmten analytischen Funktion der Kovarianzfunktion aus. Deren Parameter werden durch Approximation (Fit) der Funktion mit empirischen Kovarianzwerten gewonnen. Dabei ist ein Kompromiss zwischen der Anzahl der Parameter und der Güte ihrer Schätzung zu finden.

Nichtparametrische Kovarianzmodelle verzichten auf die Approximation mit einer analytischen Funktion. Sie verwenden Kernel-Ansätze (Altman, 1990) oder eine Eigenvektoraufspaltung (SVD) der empirischen Kovarianzmatrix \mathbf{HBH} an den Beobachtungsorten und deren Interpolation für \mathbf{HB} (siehe Kapitel 5.2.5.2).

Homogene, isotrope oder normalverteilte Kovarianzmodelle

Um ein Ensemble mit gleichen statistischen Eigenschaften für die Schätzung der empirischen Kovarianzwerte zu bilden, sind vereinfachende Annahmen notwendig. Der Standardfall für die Kovarianzmodellierung ist dabei die Annahme von Homogenität und Isotropie, so dass das Kovarianzmodell nur abstandsabhängig ist. Ist der betrachtete Prozess normalverteilt, so kann der Erwartungswert mit dem arithmetischen Mittelwert geschätzt werden.

Terminbezogene oder klimatische Kovarianzmodelle

Terminbezogene Kovarianzmodelle werden empirisch aus rein räumlichen Daten für einen Zeitpunkt geschätzt. Klimatische Kovarianzmodelle verwenden die Daten zu mehreren Terminen, um die räumliche Kovarianz zu schätzen (siehe Kapitel 4.4.2.1). Es wird dabei davon ausgegangen, dass die Felder zu verschiedenen Zeitpunkten unab-

¹²² Diese Bezeichnung wurde vom Verfasser eingeführt, um beide Methoden begrifflich zu unterscheiden. Sehr häufig wird der Unterschied zwischen beiden Ansätzen nicht herausgearbeitet. Es kann jedoch zu Fehlern der geschätzten räumlichen Kovarianz führen, falls die zeitliche Korrelation nicht beachtet wird.

hängige Realisierungen eines räumlichen Zufallsprozesses sind. Um diese Annahme zu rechtfertigen, müssen die Zeitreihen gefiltert werden (siehe Kapitel 4.4.2. und 5.2.2). Die Abbildung 5.1 vergleicht schematisch die beiden Methoden. Im klimatischen Fall kann für jedes Paar von Beobachtungen ein Kovarianzwert angegeben werden. Damit ist eine differenzierte räumliche Struktur erfassbar. Die terminbezogenen¹²³ Kovarianzmodelle benötigen Toleranzgebiete zur Ensemblebildung und verlangen aufgrund des beschränkten Datenangebots meist die Annahme von Homogenität¹²⁴ und Isotropie. Durch die Zusammenstellung von terminbezogenen Kovarianzmodellen zu verschiedenen Zeitpunkten kann jedoch die zeitliche Variabilität der Kovarianzfunktion wiedergegeben werden. Problematisch ist bei den klimatologischen Kovarianzmodellen die Erfassung des Erwartungsfeldes an den Punkten ohne Messinformation.

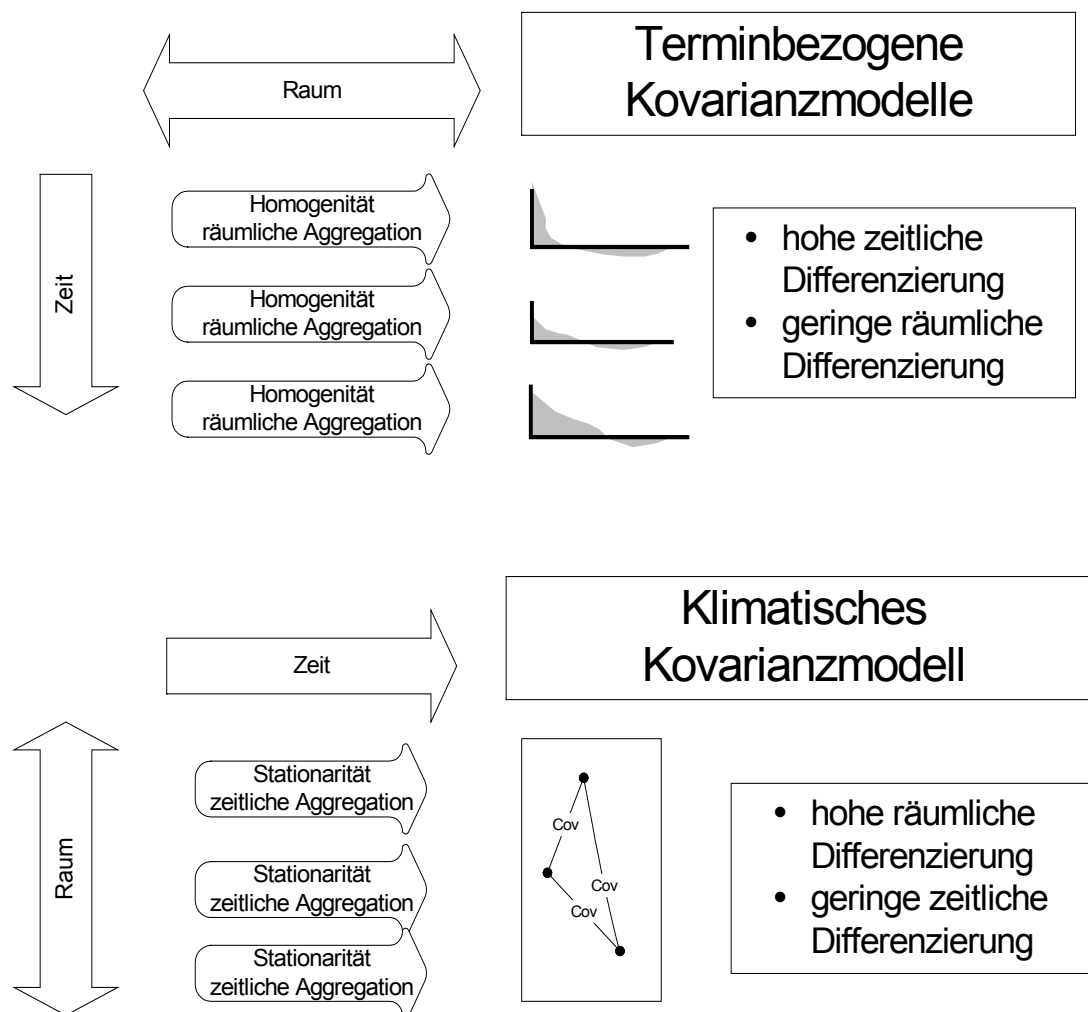


Abbildung 5.1 Schematische Übersicht der terminbezogenen und klimatischen Kovarianzmodelle

Der häufigste und einfachste Ansatz für die Kovarianzmodellierung ist das empirische, nichtbayessche, Gaußsche, homogene, isotrope, parametrische Modell. Das Kovarianzmodell ist dann eine eindimensionale analytische Funktion, die die Abstandsabhängigkeit der räumlichen Kovarianz beschreibt.

¹²³ Terminbezogene Kovarianzmodelle beruhen allein auf dem Apparat der räumlichen Statistik.

¹²⁴ Die Toleranzgebiete vereinfachen sich dadurch zu Klassen gleichen räumlichen Abstands.

Für großskalige meteorologische Felder kann die Annahme von Homogenität und Isotropie häufig gerechtfertigt werden. Die Felder der Immissionswerte sind jedoch inhomogen (siehe Kapitel 4.1), so dass diese Annahmen fragwürdig erscheinen. Es gibt in der Praxis jedoch viele Beispiele für homogene Ansätze zur Interpolation von Messwerten. Seit Anfang der neunziger Jahre gibt es verstärkt Bemühungen, empirische inhomogene Kovarianzmodelle zu entwickeln (Sampson et al., 2001), die für die Analyse von Immissionsdaten Verwendung finden (siehe Kapitel 4.3). Um die Inhomogenität zu erfassen, muss ausreichendes Datenmaterial vorhanden sein. Aus diesem Grund ist der klimatische Ansatz, d. h. die Schätzung von Kovarianzmodellen unter Verwendung von Raum-Zeit-Daten, die Basis für die Beschreibung inhomogener räumlicher Kovarianzen. Er kann jedoch die zeitliche Variabilität der Kovarianzfunktion nur ungenügend wiedergeben.

Die meteorologischen Anwendungen der statistischen Analyse beruhen vorrangig auf klimatischen Kovarianzmodellen¹²⁵ (z. B. Hollingsworth und Lönnberg, 1986 und Gandin, 1965). Im Kriging-Ansatz der Geostatistik wird das Kovarianzmodell bzw. das Variogramm (siehe Kapitel 4.4) aus den Daten des Termins¹²⁶ geschätzt¹²⁷. Reimer (1986) verwendet z. B. ein terminbezogenes Kovarianzmodell für die Analyse meteorologischer Felder auf den zeitlich variablen isentropen Koordinaten.

5.1.1 Besonderheiten der Kovarianzmodellierung von Immissionsdaten

Will man die statistischen Eigenschaften aus einer Stichprobe schätzen, so sollten die Daten aus der gleichen statistischen Grundgesamtheit stammen. Eine wesentliche Besonderheit der Immissionsmessung ist die Existenz von unterschiedlichen Regimes. Die Kenntnis des Regimes der Beobachtungen ist demzufolge eine Voraussetzung für eine sachgerechte Interpolation der Daten bzw. die Analyse. Durch den Einbezug der Modellrechnung in die Analyse wird eine weitere wichtige Informationsquelle, insbesondere an Orten ohne Beobachtung, nutzbar gemacht. Genauso wie bei den Beobachtungen kann die flächendeckende Kenntnis der Regimes der Modellfelder die statistische Analyse verbessern (siehe Kapitel 2).

Die mangelnde Berücksichtigung der Besonderheiten der Immissionsdaten (siehe Kapitel 4.1) kann die empirische Kovarianzmodellierung, insbesondere bei homogenen Ansätzen, beeinträchtigen. Im Folgenden werden die Auswirkungen mangelnder Beachtung der Besonderheiten und Ansätze zur Behandlung diskutiert. Die Lösungen beruhen auf einer Verwendung der Klassifikation der Regimes (siehe Kapitel 2) und der Modellrechnung einschließlich der aus ihr abgeleiteten statistischen Information. In Kapitel 5.1.3 werden drei konkrete, für die Analyse verwendete Kovarianzmodelle vorgestellt.

Inhomogenität der Immissionsfelder

Die Inhomogenität der Immissionsfelder bezieht sich sowohl auf die Erwartungswerte als auch auf die Kovarianzfunktion. Durch den Ansatz der Datenassimilation, eine

¹²⁵ ohne die Stationarität und Ergodizität der Zeitreihen zu überprüfen

¹²⁶ In geologischen Anwendungen sind die Daten meist überhaupt nicht zeitlich variabel.

¹²⁷ Die mangelnde Zeitabhängigkeit des Kovariogramms ist ein Unterschied zwischen Optimaler Interpolation und einfachem Kriging.

Modellrechnung als Grundlage für die Schätzung des Erwartungswertes zu verwenden, kann bereits ein hohes Maß an Inhomogenität beschrieben werden. Dies gilt z. B. auch für die Grenzfläche von Land und Meer. Aufgrund der Unzulänglichkeiten der Rechenmodelle (siehe Kapitel 5.3.2) ist jedoch davon auszugehen, dass das Feld der Fehler weder einen homogenen Erwartungswert Null noch eine homogene und isotrope Kovarianzstruktur hat.

Im Gegensatz zu anderen inhomogenen atmosphärischen Feldern wie z. B. dem Niederschlag, sind die Inhomogenitäten der Immissionsfelder in hohem Maße ortsfest. Dieser Umstand wird genutzt, indem die aus langfristigen Modellrechnungen abgeleitete räumliche Verteilung der Immissionsregimes als Klassifikationsmerkmal bei der Kovarianzmodellierung verwendet wird. Weiterhin gestattet dieser Umstand die Anwendung von klimatischen inhomogenen Kovarianzmodellen.

Geometrische Anisotropie

Das homogene Zufallsfeld ist anisotrop, wenn die Kovarianz von der Richtung abhängt. Diese Eigenschaft heißt geometrische Anisotropie, weil sie durch Skalierung und Drehung der Raumkoordinaten aufgehoben werden kann. Geometrische Anisotropie wird in der Atmosphäre z. B. bei der Form der Tiefdruckgebiete beobachtet (Daley, 1991). Die Immissionsfelder können gleichfalls eine solche Anisotropie z. B. aufgrund der vorherrschenden westlichen Windrichtung besitzen.

Die Richtungsabhängigkeit kann bei einer großen Stichprobe festgestellt werden, indem homogene parametrische Kovarianzmodelle für verschiedene Richtungen ermittelt werden. In dieser Arbeit werden unterschiedliche Kovarianzmodelle für vier Sektoren mit einer Breite von 45 ° geschätzt. Der gesuchte Kovarianzwert wird dann aus einer gewichteten linearen Kombination der Kovarianzmodelle der benachbarten Sektoren gebildet. Auf den Nachweis der statistischen Verschiedenartigkeit der Kovarianzmodelle für verschiedene Sektoren wird verzichtet. Die festgestellte Anisotropie wird in Kapitel 5.3.4.1 diskutiert.

Abweichung von der Normalverteilung

Immissionsdaten sind meist nicht normalverteilt. Das ist für die Schätzung der empirischen Kovarianzwerte mit der Momentenmethode von Bedeutung. Immissionsdaten können durch Transformation mit dem Logarithmus (Hogrefe und Rao, 2001) oder der Quadratwurzel (Meiring et al., 1998) auf eine annähernd normalverteilte Form gebracht werden. Die Kovarianzmodellierung und die Analyse erfolgt in der transformierten Größe und anschließend wird in die eigentliche Größe zurücktransformiert. Diesem formal richtigen Vorgehen steht entgegen, dass die logarithmische Transformation die Unterschiede in den höheren Wertebereichen unterbewertet, was den Intensionen der Luftreinhalte widerspricht.

Führt man die Analyse eines Feldes Z mit der log-transformierten Größe $Y=\log(Z)$ durch, so muss bei der Rücktransformation des interpolierten Werts $Y_A(\mathbf{r}_i)$ ein Bias¹²⁸ korrigiert werden. Er entspricht der Hälfte der Varianz des Analysefehlers (Cressie et al., 1999, S. 135) σ_{AY}^2 :

¹²⁸ Der Bias entsteht durch den Umstand, dass der Erwartungswert der normalverteilten log-transformierten Größe nicht dem log-transformierten Erwartungswert der nichttransformierten Größe entspricht.

$$Y = \log(Z) \quad Y = N_{\text{Gauß}}(\mu_Y, \sigma_Y^2) \quad \{5.1\}$$

$$Z_A(r_I) = \exp\left(Y_A(r_I) + \frac{\sigma_{A,Y}^2}{2}\right)$$

Neben der Transformation besteht die Möglichkeit, die empirische Kovarianz mit Hilfe verteilungsunabhängiger und robuster Methoden zu schätzen. Die auf den Perzentilen basierenden Schätzformeln weisen eine geringe Anfälligkeit für Ausreißer auf, was besonders bei NO- und Staubmessungen (PM10) wichtig ist. Hiervon wird in dieser Arbeit Gebrauch gemacht (siehe Kapitel 5.2.1 und 5.2.2).

Die Mehrzahl der Beobachtungen stammt aus urbanen und belasteten Regimes und die Stationsdichte ist heterogen

Die Mehrzahl der städtischen und verkehrsbeeinflussten Messungen verleiht den homogenen Kovarianzmodellen einen „urbanen“ Charakter, der sich in der Analyse niederschlägt. Es ist deshalb empfehlenswert, das Ensemble zur Schätzung nach dem Regimetyyp zu unterscheiden.

Die stärkere räumliche Konzentration der Beobachtungen in Ballungsräumen wird vom Analyseverfahren berücksichtigt (siehe Kapitel 4.5.5). Die heterogene Stationsdichte verfälscht jedoch die Kovarianzmodellierung, da Messungen mit geringen Abständen fast ausschließlich aus belasteten Gebieten stammen. Die Werte des Kovarianzmodells für kurze Abstände sind jedoch von besonderer Wichtigkeit: 1) Aufgrund der eingeführten Stationsauswahl aus der Umgebung des Analysepunktes (siehe Kapitel 4.5.5) und der hohen Beobachtungsdichte werden vorrangig nahe Stationen berücksichtigt. 2) Die Abschätzung der Beobachtungsfehler beruht auf der mit dem Kovarianzmodell für den Abstand Null extrapolierten Varianz.

Systematische Modell- und Beobachtungsfehler

Die Analyse {4.31} verlangt eine Biasfreiheit zwischen Modell (\mathbf{x}_B) und Beobachtung (Hy) sowie einen unabhängigen und räumlich unkorrelierten Beobachtungsfehler (siehe Kapitel 4.5.4) mit dem Erwartungswert Null. Die mangelhafte Erfüllung dieser Forderungen beruht auf der Fehlerhaftigkeit des Modells und einem systematischen Repräsentativitätsfehler der Beobachtung in den Gebieten mit hoher Immission.

Die Fehlerhaftigkeit der Eulerschen Ausbreitungsmodellierung ist vorrangig durch die ungenaue Kenntnis der Emission und im geringeren Maße von mangelhaften meteorologischen Daten bestimmt. Man kann jedoch annehmen, dass trotz fehlerhafter Mengenangaben die Lokalisierung der Emissionsschwerpunkte vertrauenswürdig ist.

Korrektur des Bias

Die sachkundige Quantifizierung und Korrektur des Bias zwischen Modell und Messung, d. h. die Abweichung der Erwartungswerte, ist problematisch. Wird der Erwartungswert der Beobachtungen als homogener Mittelwert über alle vorliegenden Beobachtungen ohne Berücksichtigung der Stationsdichte geschätzt, so bestimmen wiederum vorrangig die belasteten Beobachtungen den Wert. Das Problem wird noch verschärft, da häufig die hochbelasteten Beobachtungen eine vom Modell nicht auflösbare Skala wiedergeben. Versucht man einen räumlich variablen Bias im Kontext der klimatischen Kovarianzmodelle zu ermitteln, so steht man wiederum vor einem Inter-

polationsproblem. Es stellt sich die Frage, wie aus den nur an den Beobachtungsorten bekannten Beobachtungsinkrementen eine Korrektur für das gesamte Feld zu finden ist. Der Kompromiss ist die Annahme eines homogenen Bias, der nur durch Stationen mit einem großen Repräsentationsbereich geschätzt wird. Aus diesem Grund wird nur der Bias hinsichtlich der Land-Stationen für die Korrektur verwendet.

Systematischer Beobachtungsfehler

Das übliche statistische Modell für den Beobachtungsfehler, d. h. von Repräsentativitäts- und Messfehler, geht von einem räumlich unkorrelierten Beobachtungsfehler (siehe Kapitel 4.5.4) mit Erwartungswert Null aus. Diesen Ansatz nützt die Beobachtungsmethode (siehe Kapitel 5.2.6). Im klimatischen Ansatz besteht die Möglichkeit, die Beobachtungsfehler stationsbezogen zu bestimmen. Dabei muss sinnvoll zwischen dem unkorrelierten und dem systematischen Beobachtungsfehler unterschieden werden. Letzteren könnte man mit dem Beobachtungsoperator korrigieren.

Einfluss eines nichtkorrigierten Bias

Bei der Kovarianzschätzung führt ein Bias zu einer starken Überschätzung, da es sich um einen quadratischen Term handelt. Die Linearität der Analysegleichung {4.31} bewirkt einen linearen Einfluss des Bias. Geht man von einem Bias des Modellzustandsvektors \mathbf{x}_B und geringen Beobachtungsfehlern aus, so wird der Bias der Analyse gleichwertig oder im Einflussgebiet der Beobachtungen geringer sein. Ist eine Beobachtung aufgrund subskaliger Einflüsse biasbehaftet, so wird für sie eine große Beobachtungsfehlervarianz (siehe 5.2.6) bestimmt. Damit wird der Einfluss der Beobachtung auf die Analyse reduziert (siehe Kapitel 4.5.5). Alternativ könnte solch ein Bias durch eine Korrektur mit dem Beobachtungsoperator H berücksichtigt werden.

Zeitliche Periodizität

Die Periodizität, bzw. die damit verbundene zeitliche Korrelation, ist ein Problem für die Schätzung der empirischen Kovarianzwerte der klimatischen Kovarianzmodelle (siehe Kapitel 4.4.2). Um die Zeitreihendaten als unabhängige Beobachtungen eines räumlichen Prozesses ansehen zu können, ist die Identifizierung der zeitlich korrelierten Komponente („pre-whitening“) notwendig, damit die zeitliche Kovarianz nicht die Schätzung der räumlichen verfälscht. Hierfür sind die Klassenbildung und zeitliche Filterung¹²⁹ geeignet.

In Kapitel 3.9 wurde die typische zeitliche Variabilität der hier betrachteten Stoffe diskutiert. Der Tagesgang ist meist die am stärksten ausgeprägte zeitliche Struktur. Auch die räumliche Kovarianz kann starke Tagesschwankungen aufweisen. Elbern et al. (1998) zeigen dies z.B für die Beobachtungsinkremente der Ozonimmission. Der Tagesgang in der Raum-Zeit-Kovarianz lässt sich demzufolge nur unzureichend separieren (siehe Kapitel 4.4.2). Aus diesem Grund ist es sinnvoll, den Tageszyklus nicht durch Filterung zu entfernen, sondern für jede Tagesstunde ein eigenes räumliches Kovarianzmodell zu bilden.

Weitere wichtige Zyklen sind Jahresgang und Wochengang. Diese werden durch robuste Filterung entfernt. Damit wird von einer Separation der mittleren zeitlichen und räumlichen Komponente ausgegangen.

Meiring et al. (1998) bilden gleichfalls tagesstundenbezogene Kovarianzmodelle und verwenden ein autoregressives Modell AR(2), um den restlichen Teil der zeitlichen

¹²⁹ Die technische Durchführung der Filterung kann unter Artefakten des Verfahrens leiden.

Korrelation zu beseitigen. Host und Follestad (1999) entfernen nur den Tages- und Jahresgang durch Filterung.

5.1.2 Datenbasis für die empirischen Kovarianzmodelle

Die Kovarianzmodelle können für verschiedene Zufallsfelder bestimmt werden. Die Wahl der Datenbasis hängt von der Aufgabenstellung ab.

Für die Bestimmung der empirischen Kovarianzmodelle der Immissionsfelder stehen mehrjährige Zeitreihen von stündlichen Beobachtungen und die stündlichen Jahresrechnungen für die Gitterpunkte des Modells REM/Calgrid zur Verfügung (siehe Kapitel 1.4.2). Die Schätzung der empirischen Kovarianzwerte kann auf folgender Datenbasis beruhen und den angegebenen Aufgabenstellungen dienen:

1. Beobachtungsinckremente, d. h. die Abweichung zwischen Beobachtung und Modell \Rightarrow Analyse und Datenassimilation, Beobachtungsfehlerbestimmung
2. Beobachtungsdaten \Rightarrow Messwertinterpolation, Beobachtungsfehlerbestimmung, Beobachtungsklimatologie, Modellevaluierung
3. Modellrechnung \Rightarrow Modellklimatologie, Modellevaluierung

1. Bei der Datenassimilation ist die Kovarianzstruktur des biasfreien Modellfehlerfeldes und der Modellbias gesucht (siehe Kapitel 4.5.4). Die Kovarianzmodellierung erfolgt für die Beobachtungsinckremente. Gleichzeitig kann mit der Beobachtungsmethode aus Kapitel 5.2.6 der auflösungsabhängige Beobachtungsfehler mit Hilfe des Kovarianzmodells abgeschätzt werden.

2. Von modellunabhängiger und damit allgemeinerer Aussagekraft ist die Kovarianzstruktur des Immissionsfeldes, wie es durch die Beobachtungen registriert wird. Problematisch ist dabei die Ermittlung des kontinuierlichen Erwartungswertfeldes $\mu(\mathbf{r})$ und der Kovarianz aufgrund der beschränkten Beobachtungsdichte. Der Beobachtungsfehler lässt sich in gleicher Weise wie bei den Beobachtungsinckrementen abschätzen. Betrachtet man das Modellfeld als eine biasfreie Schätzung des Erwartungswertes $\mu(\mathbf{r})$, so entspricht dies dem Vorgehen bei der Datenassimilation.

3. Die Vorteile bei der Schätzung von Kovarianzmodellen aus Modelldaten sind, dass für alle Raumpunkte Informationen vorliegen und dass diese nicht mit dem Beobachtungsfehler kontaminiert sind. Im Idealfall einer Übereinstimmung der statistischen Eigenschaften (Erwartungswert und Kovarianzfunktion) des Modells mit denen des „wahren“ Feldes \mathbf{x}_{true} können die Modellwerte als empirische Datenbasis für die Kovarianzmodellierung verwendet werden. Die Kovarianzstruktur des Modells kann auch direkt mit der der reinen Beobachtungen in einer anspruchsvollen Modellevaluierung verglichen werden (siehe Kapitel 5.3.3.1).

Die Kovarianzmodellierung für Modelldaten hat Tradition in der numerischen Wettervorhersage. Die „NMC¹³⁰-Methode“ vergleicht Modellprognosen für einen Termin aus einer 24- bzw. 48-stündigen Vorhersage (Parrish und Derber, 1992). Die Analysis-Ensemble-Methode verwendet ein Ensemble von Prognosen mit leicht variierten Anfangsbedingungen (Fisher und Courtier, 1995). Diese Methoden gehen schließlich im Kalman-Filter auf, bei dem die Kovarianz aus gestörten Modellzuständen gebildet wird (siehe Kapitel 4.5.7).

¹³⁰ National Meteorological Center

5.1.3 Drei Kovarianzmodelle (A, B und C) für Immissionsdaten

Im Rahmen der Arbeit sind drei empirische Kovarianzmodelle A, B und C für die Analyse von stündlichen Immissionsdaten entwickelt und verglichen worden. Es handelt sich dabei um ein homogenes, terminbezogenes Kovarianzmodell A und zwei klimatische, inhomogene Kovarianzmodelle B und C, die jeweils für Ozon, NO₂, NO₂ und PM₁₀ abgeleitet werden.

Das terminbezogene homogene Modell A liefert einen räumlich homogenen Wert der gesuchten Größen für jeden Termin. Die klimatischen Kovarianzmodelle geben stationsbezogene bzw. räumlich inhomogene Werte für jeweils eine Stunde des Tages an, die für einen langen Zeitraum Gültigkeit haben (siehe Abbildung 5.1).

Das klimatische Kovarianzmodell B beruht auf Kombination von mehreren parametrischen, homogenen Kovarianzmodellen, die für jeweils eine Kombination von Immissionsregimes aufgestellt werden. Das klimatische Kovarianzmodell C ist nichtparametrisch und wird durch die Eigenvektoraufspaltung der empirischen Kovarianzmatrix **HBH** gebildet. Die klimatischen Kovarianzmodelle B und C sind für jede Tagesstunde separat formuliert und berücksichtigen so die Variabilität der räumlichen Kovarianz im Tagesgang.

Alle drei Kovarianzmodelle werden hinsichtlich der erzielten Analysequalität in Kapitel 6.2.2 untersucht. Der Vergleich unterschiedlicher Kovarianzmodelle als Ansatz einer Modell- und Beobachtungsklimatologie erfolgt in Kapitel 5.3.

Terminbezogenes homogenes parametrisches Kovarianzmodell A

Es stellt den klassischen Ansatz der Optimalen Interpolation für homogene Felder dar. Der Ausgangspunkt der Schätzung ist eine Stichprobe der Beobachtungsinkremente zu dem gegebenen Termin. Aus ihnen werden empirische Kovarianzwerte pro Abstandsklasse und für 4 verschiedene Richtungssektoren geschätzt und mit Hilfe einer analytischen Funktion approximiert (siehe 5.2.3). Es erfolgt eine Korrektur mit dem homogenen Bias der Land-Stationen. Die durch die Beobachtungsmethode bestimmte Varianz des Beobachtungsfehlers gilt für alle Stationen.

Der Vorteil dieses Kovarianzmodells ist seine einfache Schätzung aus vorliegenden Daten. Mit diesem Kovarianzmodell kann auch die Analyse von Jahresmittelwerten der Beobachtung direkt erfolgen. Im Rahmen der aktiven Datenassimilation mit OI wird dieses Kovarianzmodell verwendet, um während der Modellrechnung Beobachtungen zu assimilieren. Auch die Eigenvektoren für Kovarianzmodell C werden mit diesem Modell A ohne Biaskorrektur interpoliert.

Klimatisches inhomogenes parametrisches Kovarianzmodell B auf Basis der Immissionsregimes der Modellrechnung

Der Ansatz, die Inhomogenität zu beschreiben, beruht auf einer Superposition verschiedener homogener Kovarianzmodelle. Es wird dabei angenommen, dass innerhalb eines Regimes eine homogene Kovarianzstruktur zu finden ist. Die Verteilung der Regimes wird aus der Modellrechnung gewonnen.

Die Basis für die Modellbildung sind die empirischen klimatischen stationspaarbezogenen Kovarianzwerte. Sie werden aus den Zeitreihen der Werte zu einer Tagesstunde für das ganze Jahr und getrennt für Sommer und Winter gebildet. Somit liegt für jede Tagesstunde und Saison ein homogenes Kovarianzmodell für jede Kombination von Regimes vor. Die Auswahl des anzuwendenden Kovarianzmodells für die Analyse-

punkte, d. h. für **HB**, wird durch das Regime der Gitterbox aus der Modellrechnung und dem der verwendeten Beobachtung festgelegt.

Die Varianzen des Beobachtungsfehlers werden, aufbauend auf diesem Kovarianzmodell, für jede Kombination des Regimes des Backgrounds am Stationsort und dem Regime der Stationsmessung bestimmt.

Klimatisches inhomogenes Kovarianzmodell C auf Basis der EOF der empirischen Kovarianzmatrix

Anders als beim klimatischen Kovarianzmodell B wird bei diesem Kovarianzmodell die Struktur der räumlichen Inhomogenität aus der empirischen Kovarianzmatrix abgeleitet. Mit Hilfe einer Eigenvektoraufspaltung der empirischen Kovarianzmatrix lassen sich die wichtigen Anteile der Kovarianz, d. h. die Eigenvektoren, an den Stationsorten bestimmen und in einem weiteren Schritt auf die Analysepunkte ohne Messinformation interpolieren. Aus den interpolierten Eigenvektoren wird dann die vollständige Kovarianzmatrix für alle geforderten Paare von Orten rekombiniert. Eine genauere Beschreibung dieses Kovarianzmodells erfolgt in Kapitel 5.2.5.2.

Das hier verwendete Kovarianzmodell auf Basis der EOF unterscheidet sich von anderen Anwendungen (Holland et al., 1999, Obled und Creutin, 1986, und van Egmond und Onderdelinden, 1981) durch die vorherige Separation des Beobachtungsfehlers und die Anwendung der Optimalen Interpolation (Kovarianzmodell A) zur Interpolation der Eigenvektoren. Die anderen Autoren verwenden hierfür mathematische formale Verfahren ohne statistische Komponente.

5.2 Die Schätzung der empirischen Kovarianzmodelle

5.2.1 Schätzung homogener terminbezogener Kovarianzwerte

Das Ensemble zur Schätzung der empirischen Kovarianzwerte $\hat{C}(r)$ wird für den homogenen Prozess aus der Menge $N(r)$ von Beobachtungen $y(\mathbf{r}_i)$ am Ort \mathbf{r}_i im gleichen Abstandsintervall $r+\Delta r$ gewonnen:

$$\hat{C}(r) = \frac{1}{|N(r)|} \sum_{i,j \in N(r)} \left(y(\mathbf{r}_j) - \hat{\mu}(\mathbf{r}_j) \right) \left(y(\mathbf{r}_i) - \hat{\mu}(\mathbf{r}_i) \right) \quad \{5.2\}$$

$$N(r) = \left\{ (\mathbf{r}_i, \mathbf{r}_j) : r \leq |\mathbf{r}_i - \mathbf{r}_j| < r + \Delta r, \quad i, j = 1, \dots, n \right\}$$

$$\hat{\mu}(r) = \hat{\mu} = \bar{y} = \frac{1}{N} \sum_i y(\mathbf{r}_i) \quad \{5.3\}$$

Die Breite der Abstandsintervalle wird von der Datenlage und der räumlichen Auflösung bestimmt. Cressie (1993) gibt an, dass mindesten 30 Paare von Daten für die Bildung eines empirischen Kovarianzwertes vorhanden sein sollten.

Liegen keine anderen Informationen vor, so wird der homogene Erwartungswert aus dem Mittelwert der Beobachtungen geschätzt. In diesem Fall kann jedoch auf Bildung des Erwartungswertes verzichtet werden, wenn auf das geschätzte empirische Variogramm $2\hat{\gamma}(r)$ zurückgegriffen wird:

$$2\hat{\gamma}(r) = \frac{1}{|N(r)|} \sum_{i,j \in N(r)} (y(\mathbf{r}_j) - y(\mathbf{r}_i))^2 \quad \{5.4\}$$

Aus ihm kann mit Kenntnis der Varianz $\hat{C}(0)$ der Kovarianzwert mit Hilfe von Formel {4.10} bestimmt werden. Die Verwendung von Variogrammen hat den Vorteil, dass sie robuster gegenüber einem unbeseitigten Trend in den Daten, d. h. gegenüber einer Abweichung von der Annahme der Homogenität, sind (Cressie, 1993, S. 72).

Neben den klassischen¹³¹ Schätzungen gibt es robuste Ansätze, die weniger fehlerhaft auf Ausreißer und Abweichungen von der Normalverteilung reagieren. Robuste Schätzungen der Kovarianz sind mit der folgenden Gleichung möglich:

$$\hat{C}(r) = \frac{1}{4} \text{MAD} \{y(\mathbf{r}_j) + y(\mathbf{r}_i)\} - \text{MAD} \{y(\mathbf{r}_j) - y(\mathbf{r}_i)\} \quad \{5.5\}$$

Sie beruhen auf dem Zusammenhang zwischen Kovarianz und Varianz und deren robuster Schätzung mit der absoluten Medianabweichung (MAD):

$$\begin{aligned} \text{Cov}(a, b) &= \{ \text{Var}(a + b) - \text{Var}(a - b) \} / 4 \\ \overline{\text{Var}(a)} &= 1.486 \text{MAD}(a) \\ \text{MAD} &\equiv \text{med} \{ |x_i - \text{med} \{x_i, i = 1, n\}|; i = 1, n \} \end{aligned} \quad \{5.6\}$$

Variogrammwerte können in gleicher Weise mit Hilfe des Medians robust geschätzt werden:

$$2\hat{\gamma}(R) = \frac{1}{0.457} \text{Median} \left\{ \left(y(\mathbf{r}_j) - y(\mathbf{r}_i) \right)^2 \right\} \quad \{5.7\}$$

Der Vorfaktor ist die Korrektur des Bias, der durch das Ersetzen des Mittelwertes durch den Median entsteht¹³² (Smith, 2001). Die folgende, häufig verwendete robuste Schätzung ist auf weiteren theoretischen Argumenten aufgebaut (Cressie, 1993):

$$2\hat{\gamma}(R) = \frac{1}{.457 + .495/|N(R)|} \left\{ \frac{1}{|N(r)|} \sum_{i,j \in N(R)} \left| y(\mathbf{r}_j) - y(\mathbf{r}_i) \right|^{\frac{1}{2}} \right\}^4 \quad \{5.8\}$$

¹³¹ auch Momentenmethode

¹³² Der Bias ist damit zu erklären, dass für eine annähernd symmetrisch-verteilte Größe x die Größe x^2 nicht symmetrisch ist. Der Median von x^2 bedarf damit einer Korrektur, wenn er als Schätzung für den Erwartungswert herangezogen wird.

5.2.2 Schätzung klimatischer Kovarianzwerte

5.2.2.1 Filterung der zeitlichen Periodizität und Persistenz

Die klimatischen Kovarianzmodelle werden separat für jede Tagesstunde gewonnen. Um die notwendige zeitliche Unabhängigkeit des Ensembles zu erreichen, muß nun noch die Filterung von Jahres- und Wochengang und „synoptischer Persistenz“ erfolgen (siehe Kapitel 4.4.2).

Der Jahresgang wird mit einem Tiefpass geschätzt, der durch einen gleitenden Medianfilter (Wilks, 1995) von 21 Tagen realisiert wird. Die Abweichung des Jahresganges vom Median der gesamten Zeitreihe ist die Korrekturgröße der Zeitreihe $(y_d^h)^y$ für jede Tagesstunde h , die dann keinen Jahresgang mehr aufweist:

$$(y_d^h)^y = y_d^h - \left(\text{med}_{d'=d-10, d+10} \{y_{d'}^h\} \right) + \left(\text{med}_{d'=1, 365} \{y_{d'}^h\} \right) \quad \{5.9\}$$

Der Wochengang als eine typische Eigenschaft von Immissionsdaten stellt eine periodische aber keine harmonische Schwingung dar. Die üblichen, auf der spektralen Betrachtungsweise aufgebauten Filter sind demzufolge nicht ohne weiteres anwendbar. Die Beseitigung des Wochengangs erfolgt jedoch prinzipiell nach dem gleichen Vorgehen wie beim Jahresgang. Der Wochengang wird ermittelt, in dem für jeden Wochentag der Median der jahresganggefilterten Zeitreihe der Tagesstundenwerte bestimmt wird. Die Differenz zum Median der gesamten Zeitreihe bildet die Korrekturgröße:

$$(y_d^h)^{wy} = (y_d^h)^y - \left(\text{med}_{d'=\text{mod}(d+1,7), 365, 7} \{ (y_{d'}^h)^y \} \right) + \left(\text{med}_{d'=1, 365} \{ (y_{d'}^h)^y \} \right) \quad \{5.10\}$$

Neben Jahres- und Wochengang hat die Dauer der Wettersituationen eine zeitliche Korrelation zur Folge. Im Gegensatz zu den durch externe Faktoren bestimmten systematischen Schwankungen ist die Ausprägung dieser Erhaltungsneigung variabel. Um diesen Anteil aus den Daten zu entfernen, wird auf ein autoregressives Modell zurückgegriffen. Autoregressive Modelle bilden die Zeitreihe aus einer Regression zurückliegender Werte. Aufbauend auf die Erfahrung von Meiring et al. (1998) wurde ein AR-Modell zweiter Ordnung gewählt. Die mit diesem AR(2) Modell konstruierte Zeitreihe wird biaskorrigiert von der ursprünglichen, d. h. der jahres- und wochenganggefilterten Reihe abgezogen. Somit realisiert sich der Hochpass wiederum durch die Subtraktion einer tiefpassgefilterten Zeitreihe.

$$(y_d^h)^* = (y_d^h)^{wy} - \left(\text{AR2} \left((y_d^h)^{wy} \right) \right) + \left(\text{med}_{d'=1, 365} \left\{ \text{AR2} \left((y_{d'}^h)^{wy} \right) \right\} \right) \quad \{5.11\}$$

mit

$$\text{AR2}(y_t) = \mu_x + \alpha_1 (y_{t-1}) + \alpha_2 (y_{t-2})$$

5.2.2.2 Diskussion der Filterung am Beispiel

Die Abbildung 5.2 zeigt die tägliche Zeitreihe der NO₂-Konzentration um 7 Uhr für die Vorstadt-Station Burg (BB001, U1) und die zugehörige Autokorrelationsfunktion. Deutlich sichtbar ist die hohe Persistenz und eine 7-tägige Schwingung, die durch den Wochengang bedingt ist.

Abbildung 5.3, Abbildung 5.4 und Abbildung 5.5 zeigen die Anwendung des Jahresgangs-, Wochengangs- und AR(2)-Filters auf die Messreihe und die entsprechenden zeitliche Autokorrelationsfunktionen. Die nacheinanderfolgende Anwendung der drei Filter führt zu der Zeitreihe in Abbildung 5.6. Die entsprechend dargestellte Autokorrelationsfunktion weist fast keine signifikanten zeitlichen Autokorrelationen mehr auf.

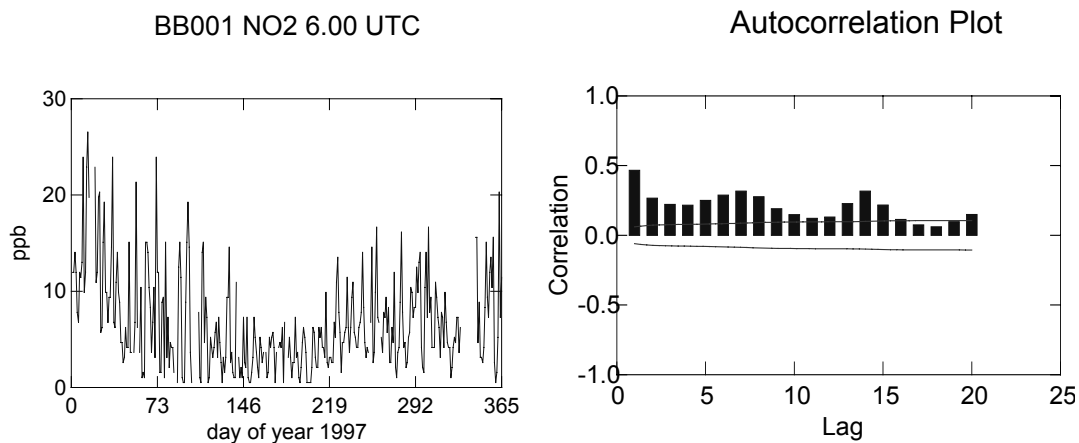


Abbildung 5.2 Tägliche Zeitreihe der NO₂-Konzentration um 7 Uhr (MEZ) an der Station Burg (BB001) und Autokorrelationsfunktion (Lag = Anzahl der verschobenen Tage).

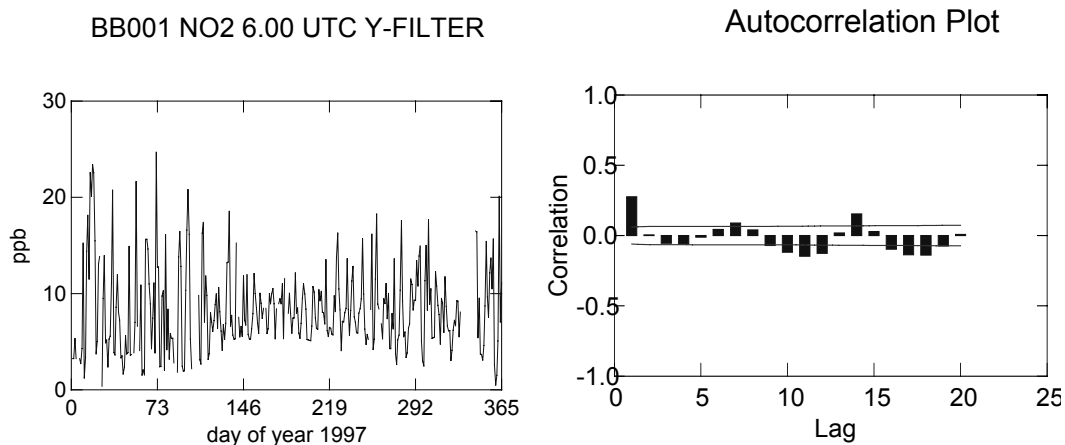


Abbildung 5.3 Tägliche Zeitreihe der NO₂-Konzentration um 7 Uhr (MEZ) an der Station Burg (BB001), jahresganggefiltert, und deren Autokorrelationsfunktion (Lag = Anzahl der verschobenen Tage).

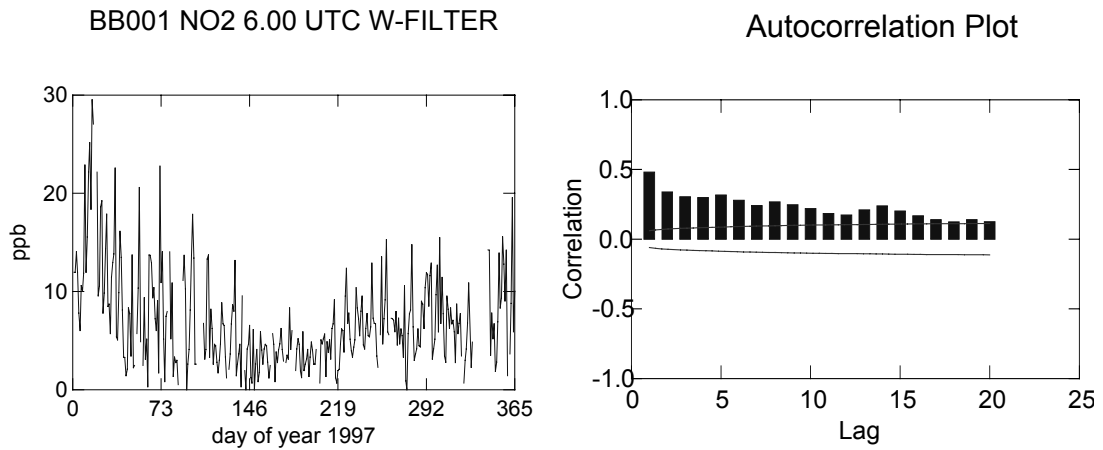


Abbildung 5.4 Tägliche Zeitreihe der NO₂-Konzentration 7 Uhr (MEZ) an der Station Burg (BB001), wochengangefiltert, und deren Autokorrelationsfunktion (Lag = Anzahl der verschobenen Tage).

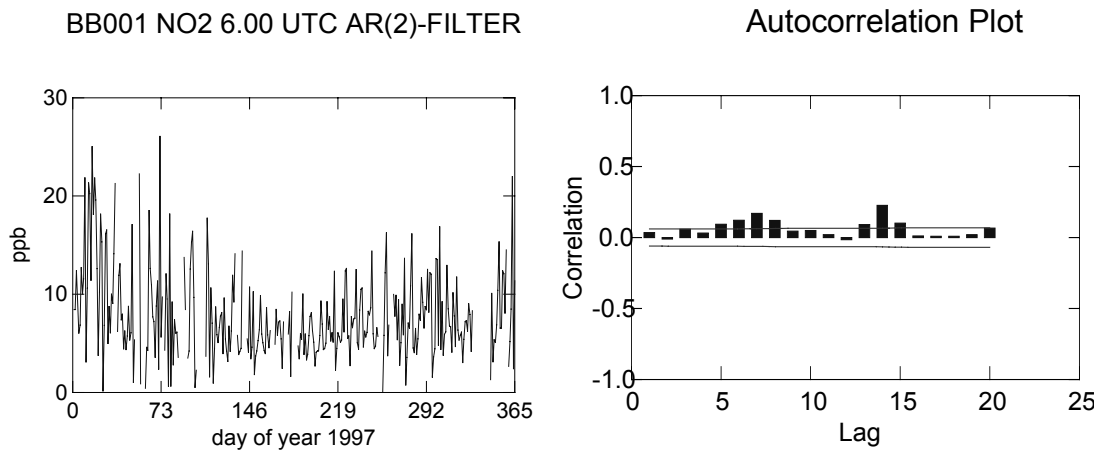


Abbildung 5.5 Tägliche Zeitreihe der NO₂-Konzentration um 7 Uhr (MEZ) an der Station Burg (BB001), AR(2)-gefiltert, und deren Autokorrelationsfunktion (Lag = Anzahl der verschobenen Tage).

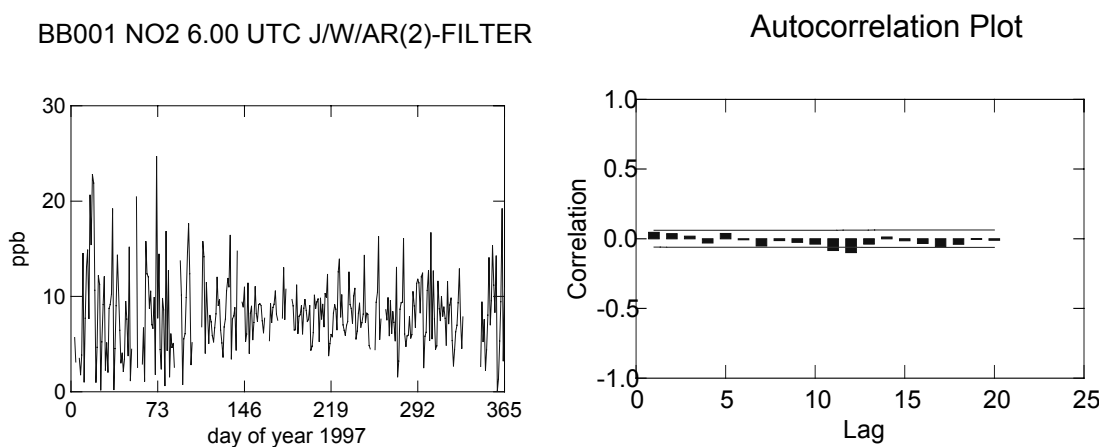


Abbildung 5.6 Tägliche Zeitreihe der NO₂-Konzentration um 7 Uhr (MEZ) an der Station Burg (BB001), Jahresgang-, Wochengang- und AR(2)-gefiltert, und deren Autokorrelation (Lag = Anzahl der verschobenen Tage).

5.2.2.3 Schätzung stationspaarbezogener Kovarianzwerte

Die gefilterten Zeitreihen $(x_d^h)^*$ bilden das Ensemble, um für jede Tagesstunde $h = 1 \dots 24$ und jedes Stationspaar $(\mathbf{r}_i, \mathbf{r}_j)$ einen empirischen Kovarianzwert $\hat{C}^h(\mathbf{r}_i, \mathbf{r}_j)$ zu schätzen:

$$\hat{C}^h(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{N} \sum_{d=1}^N \left((x_d^h)^*(\mathbf{r}_i) - \overline{(x_d^h)^*(\mathbf{r}_i)} \right) \left((x_d^h)^*(\mathbf{r}_j) - \overline{(x_d^h)^*(\mathbf{r}_j)} \right)$$

$h \dots$ Stunde des Tages
 $\mathbf{r}_i, \mathbf{r}_j \dots$ Beobachtungsorte
 $d = 1 \dots N \dots$ Tage pro Jahr/Saison
 $()^* \dots$ jahres- und wochengangsbereinigte Zeitreihe
 $\overline{(x)} = \frac{1}{N} \sum_{d=1}^N x_d$

{5.12}

Analog zur Schätzformel {5.5} kann eine robuste Schätzung $\hat{C}^h(\mathbf{r}_i, \mathbf{r}_j)$ des stationsbezogenen Kovarianzwerts erfolgen. Die Bildung der empirischen Variogrammwerte 2γ ist nur sinnvoll unter der Annahme intrinsischer Homogenität (siehe Gleichung {4.9}). Aus diesen Werten kann ein Variogrammmodell und daraus mit Gleichung {4.10} das Kovarianzmodell gebildet werden.

5.2.3 Analytische Form des parametrischen Kovarianzmodells

Die Aussage der empirischen Kovarianzwerte $\hat{C}(R)$ bzw. $\hat{C}(\mathbf{r}_i, \mathbf{r}_j)$ lässt sich durch die Approximation mit einer analytischen Funktion erweitern. Die Parameter einer vorgegebenen Funktion werden dabei aufgrund der vorliegenden Datenpunkte bestimmt. Die analytische Form der Funktion stellt sicher, dass das Kovarianzmodell den theoretischen Eigenschaften (siehe Kapitel 4.4), d. h. positive Definität und Konvergenz gegen Null für große Abstände, genügt. Weiterhin sind durch die Wahl der Funktion die spektralen Eigenschaften des Feldes bestimmt, falls es sich um einen homogenen Prozess 2. Ordnung handelt¹³³.

Für eine umfassende Zusammenstellung von Funktionsklassen, die als Kovarianzmodell für räumliche Zufallsfunktionen dienen können, sei auf Cressie Cressie (1993) verwiesen. Verschiedene Kovarianzmodelle für meteorologische Anwendungen und deren Abhängigkeit von der Art des Backgrounds (Modell oder klimatisches Feld aus Beobachtungen) werden in Thiebaut (1985) und Daley (1991) vorgestellt. Die vorge-

¹³³ Mit Hilfe des Wiener-Chinchin-Theorems lässt sich das Energiespektrum eines stationären Prozesses zweiter Ordnung aus der Fouriertransformierten der Kovarianzfunktion gewinnen. Für die Anwendung der spektralen Betrachtungsweise wird aus Gründen der Vereinfachung Homogenität und Isotropie für den Erwartungswert und den stochastischen Anteil vorausgesetzt.

schlagenen Kovarianzmodelle bestehen meist aus der Kombination einer linear¹³⁴ oder quadratisch abfallenden Exponentialfunktion und Kosinustermen¹³⁵. Der Exponentialterm impliziert ein abfallendes Energiespektrum („rotes Rauschen“), das aufgrund des Kosinusterns einen hervortretenden Wellenlängenbereich („farbiges Rauschen“) besitzt (Schweschnikow, 1965). In den Anwendungen zur Interpolation von Immissionsbeobachtungen ist meist kein Nulldurchgang der empirischen Kurve und somit auch kein hervortretender Frequenzbereich feststellbar. In vielen Kriging-Anwendungen werden neben exponentiellen und Gaußschen Funktionen, Kombinationen von Besselfunktionen oder sphärische Variogramme verwendet (Smith, 2001).

Für die Auswahl einer theoretisch möglichen Funktion ist die Güte der Approximation der empirischen Werte das entscheidende Merkmal. Im Fall der Immissionsdaten gilt dies besonders für geringe Abstände, da meist keine weit entfernten Beobachtungen für die Analyse verwendet werden¹³⁶. Von besonderem Interesse ist das Verhalten der Funktion beim Abstand Null für die Robustheit der Schätzung der Varianz des Beobachtungsfehlers (Beobachtungsmethode, siehe Kapitel 5.2.6). Denn für die Beobachtungsmethode wird die Varianz des Backgroundfehlers durch die Extrapolation mit Hilfe des Kovarianzmodells für den Abstand Null gewonnen.

Getestet wurden das exponentielle, Gaußsche und sphärische Kovarianzmodell ({5.13}, {5.14} und {5.15}), da sie augenscheinlich gut der Abstandsabhängigkeit der ermittelten empirischen Kovarianzwerte des Modellfehlers (Background) entsprechen.

Die genannten Funktionen hängen von zwei Parametern ab; es sind dies die geschätzte Varianz σ_B^2 und der räumliche Skalenparameter (Range) L (Abbildung 5.7). Letzterer ist ein Maß für die Größe des Gebiets, in dem eine räumliche Beziehung in Form einer Kovarianz besteht:

$$f_{\text{exp}}(r) = \sigma_B^2 \exp\left(-\frac{|r|}{L}\right) \quad \{5.13\}$$

$$f_{\text{gau}}(r) = \sigma_B^2 \exp\left(-\frac{r^2}{L^2}\right) \quad \{5.14\}$$

$$f_{\text{sph}}(r) = \begin{cases} \sigma_B^2 \left(1 + \frac{1}{2}\left(\frac{r}{L}\right)^3 - \frac{3}{2}\left(\frac{r}{L}\right)\right) & r \leq L \\ 0 & r > L \end{cases} \quad \{5.15\}$$

¹³⁴ Der Funktion f_1 ist aufgrund des linearen Exponentialterms an der Stelle $r = 0$ nicht stetig differenzierbar. In der Praxis liegen nur diskrete Beobachtungen des Prozesses vor, so dass die Annahme mangelnder Differenzierbarkeit kein Hindernis für die Anwendung dieser Kovarianzfunktion ist. Eine detaillierte Diskussion ist bei Taubenheim (1969) zu finden.

¹³⁵ Die Lage der Nullstelle der Kovarianzfunktion, bedingt durch den Kosinustern, korrespondiert mit der Wellenlänge mit maximaler Spektralenergie.

¹³⁶ In meteorologischen Anwendungen gilt das nicht, wenn z. B. über dem Meer nur wenige Beobachtungen vorliegen.

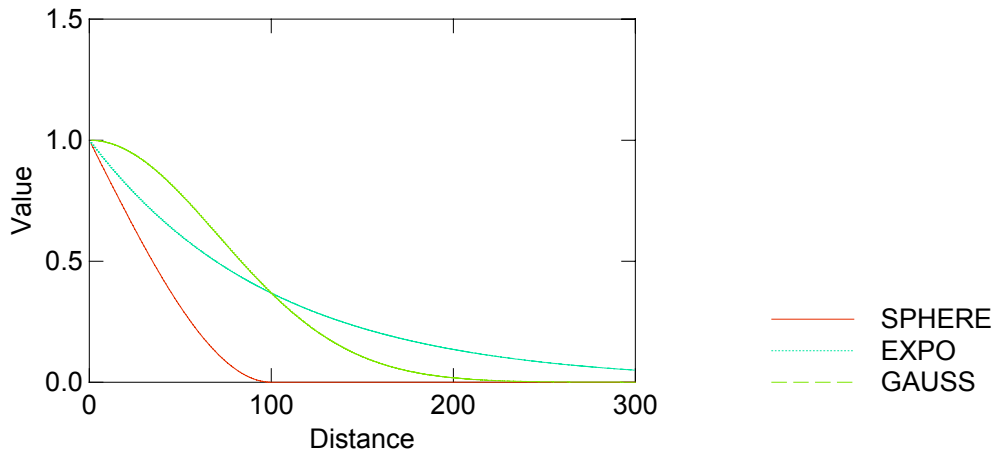


Abbildung 5.7 Exponentielles, Gaußsches und sphärisches Kovarianzmodell (Gleichung {5.13}, {5.14} und {5.15}), $\sigma_B^2=1.0$, $L=100$

5.2.4 Approximation des parametrischen Kovarianzmodells

Bei der Approximation des terminbezogenen homogenen Kovarianzmodells existiert für jede Abstandsklasse ein Wert. Im Fall der klimatischen Kovarianzmodelle ist die Punktwolke der paarweisen Kovarianzwerte zu approximieren. Das parametrische Kovarianzmodell stellt eine Kurve durch die Punktwolke dar. Für jede Abstandsklasse existieren damit mehrere Werte, deren Streuung die Bandbreite möglicher Kovarianzen angibt und auf die Unzulänglichkeit der Annahme von Homogenität und Isotropie verweist. Für die größere Robustheit und um die Anzahl der Normalgleichungen zu verringern, wird in diesem Fall der Median der Kovarianzwerte als Maß für die Kovarianz in dieser Abstandsklasse gebildet.

Nach der Methode der gewichteten kleinsten Quadrate (w.l.s.¹³⁷) werden die nichtlinearen Parameter θ_i ($i=1, k$) der Funktion f_c durch die Minimierung der folgenden Kostenfunktion gebildet:

$$\sum_{i=1}^n w_i \left(\widehat{C}(R_i) - f_c(R_i, \boldsymbol{\theta}) \right)^2 \rightarrow \min, \quad \boldsymbol{\theta} = \{\theta_{1...k}\} \quad \{5.16\}$$

Die Gewichte w_i bestimmen den Einfluss, den der einzelne empirische Kovarianzwert bei der Approximation erhält. Eine statistisch motivierte Wahl der w_i ist ein Maß für die Sicherheit von $\widehat{C}(R)$. Bei den klimatischen Kovarianzmodellen wird hierfür der Quartilabstand IQR^{138} der Kovarianzwerte der Abstandsklasse herangezogen. Die Streuung der terminbezogenen Kovarianzwerte wird mit einem Ansatz für Streuung von Korrelations- bzw. Kovarianzwerten (Taubenheim, 1969) gewonnen. Die geforderte Unabhängigkeit der Daten wird vernachlässigt. Diese Vertrauensbereiche verringern sich mit zunehmender Stichprobenanzahl und zunehmender relativer Größe der Kovarianzwerte. Von besonderer praktischer Bedeutung ist das Kovarianzmodell für kurze

¹³⁷ weighted least squares

¹³⁸ Differenz zwischen 75%-Perzentil und 25%-Perzentil

Abstände r (siehe Kapitel 5.2.3). Eine bessere Übereinstimmung in diesem Bereich wird mit einer höheren Wichtung erzwungen:

$$w_i^* = \left(\frac{r_{Max} - r_i}{r_{Max}} \right)^2 \sqrt{N(R)} \begin{cases} \frac{1}{IQR} & \dots \text{ klimatisches KM} \\ \frac{\widehat{C(r_i)}}{\overline{C(r_i)}} & \text{terminbezogene KM} \end{cases} \quad \{5.17\}$$

Im Falle einer linearen Abhängigkeit der Parameter θ_i führt Gleichung {5.16} zu den Gaußschen Normalgleichungen. Für die nichtlineare Abhängigkeit der verwendeten Kovarianzmodelle {5.13}, {5.14} und {5.15} wird eine Taylorentwicklung der Funktion f_c hinsichtlich θ_i , d. h. σ_B^2 und L , gebildet, die eine lineare Funktion für $d\theta_i$ darstellt:

$$f_c(R, \boldsymbol{\theta} + d\boldsymbol{\theta}) = f_c(R, \boldsymbol{\theta}) + \sum_{i=1}^K \frac{\partial f(R, \boldsymbol{\theta})}{\partial \theta_i} d\theta_i \quad \{5.18\}$$

Mit dem iterativen Verfahren nach Levenberg-Marquard (Press, 1992) wird das Minimum des Ausdrucks {5.16} gesucht. Für die praktische Durchführung der Approximation müssen initiale Werte der Parameter σ_B^2 und L vorgegeben werden. Die Größe der beiden Parameter wird durch die Schnittpunkte mit der y - und x -Achse einer linearen Regression der Daten gewonnen.

5.2.5 Nichtparametrisches inhomogenes Kovarianzmodell aus EOFs

Neben der Methode, die empirischen stationsbezogenen klimatischen Kovarianzwerte durch eine analytische Funktion zu approximieren, können auch Ansätze aus dem Umfeld der Hauptkomponentenanalyse (PCA¹³⁹) verwendet werden, um die Kovarianzfunktion zu modellieren. Die PCA beruht auf einer reduzierten Darstellung eines Datensatzes von Zeitreihen an verschiedenen Orten, und zwar mit Hilfe einer Auswahl von Eigenvektoren (EOF) der empirischen Kovarianzmatrix und zeitlich variierenden Koeffizienten (Hauptkomponenten, Zwiers und von Storch, 2000).

Der Gedanke, durch eine Eigenwertaufspaltung die bedeutsamen Anteile der empirischen Kovarianzmatrix zu separieren, findet auch bei der räumlichen Kovarianzmodellierung Anwendung (siehe Kapitel 5.2.5.2). Ausgangspunkt ist der klimatische Ansatz, der auf die stationsbezogene Kovarianzmatrix zurückgreifen kann. Sie wird durch die Eigenvektorzerlegung komprimiert dargestellt. Neben der Erfassung von inhomogenen Strukturen sind numerische Aspekte eine wichtige Motivation für die reduzierende Eigenvektorzerlegung (SVD¹⁴⁰) der empirischen Kovarianzmatrix (siehe Kapitel 5.2.5.3).

¹³⁹ Principal Component Analysis

¹⁴⁰ Singular Value Decomposition

5.2.5.1 Matrixapproximation durch SVD

Durch eine Hauptachsentransformation mit der orthogonalen Matrix \mathbf{V} kann eine beliebige symmetrische Matrix \mathbf{A} , wie es alle Kovarianzmatrizen sind, in eine Diagonalmatrix $\mathbf{\Lambda}$ transformiert werden, deren Diagonalelemente die Eigenwerte λ_i von \mathbf{A} sind:

$$\mathbf{A} = \mathbf{V}^T \mathbf{A} \mathbf{V} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{mit } \mathbf{A} \mathbf{x}_{ei} = \lambda_i \mathbf{x}_{ei} \quad \text{und} \quad \mathbf{V} \mathbf{V}^T = \mathbf{I} \quad \{5.19\}$$

Die Spalten der Matrix \mathbf{V} sind dabei die zugehörigen Eigenvektoren \mathbf{x}_{ei} . Gilt für gewisse Eigenwerte $\lambda_i=0$, so ist \mathbf{A} singulär. Die Aufteilung der singulären Anteile einer Matrix mit Hilfe der Eigenwerte wird als singular value decomposition (SVD) bezeichnet. In umgekehrter Reihenfolge kann \mathbf{A} aus \mathbf{V} und $\mathbf{\Lambda}$ gebildet werden. Die SVD kann man benutzen, um die Matrix \mathbf{A} zu approximieren¹⁴¹ (Press, 1992). Dazu werden kleine Eigenwerte in $\mathbf{\Lambda}$ gleich Null gesetzt ($\mathbf{\Lambda}'$) und die entsprechenden Spalten, d. h. die zugehörigen Eigenvektoren, in der Matrix \mathbf{V} entfernt (\mathbf{V}'). Die approximierte Matrix \mathbf{A}' ergibt sich dann aus:

$$\mathbf{A}' = \mathbf{V}' \mathbf{\Lambda}' \mathbf{V}'^T \quad \{5.20\}$$

Der Rang der Matrix \mathbf{A}' ist geringer als der von \mathbf{A} , da dies für $\mathbf{\Lambda}'$ und \mathbf{V}' gilt. Basiert man die rechentechnische Speicherung von \mathbf{A}' auf \mathbf{V}' und $\mathbf{\Lambda}'$, so verringert sich der Aufwand zur Speicherung der originalen Matrix \mathbf{A} . Dies wird bei der RRSQR-Approximation der Kovarianzmatrix ausgenutzt (siehe Kapitel 4.5.8). Hierbei wird die räumliche Kovarianzmatrix \mathbf{P} durch ihre Wurzel dargestellt: $\mathbf{P} = \mathbf{S} \mathbf{S}^T$. Die Matrix \mathbf{S} wird durch $\mathbf{S} = \mathbf{V}' \mathbf{\Lambda}'^{1/2}$ approximiert.

5.2.5.2 Auf SVD basierendes inhomogenes Kovarianzmodell C

Das Kovarianzmodell C ist durch die Implementierung des Kalman-Filters in der RRSQR-Form motiviert (Kapitel 4.5.8). Ausgangspunkt ist die Eigenvektorzerlegung der empirischen Kovarianzmatrix.

Die empirische Kovarianzmatrix \mathbf{C} der Beobachtungsinkremente entspricht der Matrix $\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}$. Durch sie ist bereits ein wichtiger Teil der für die Gewinnung der Analysegewichte \mathbf{K} notwendigen Kovarianzinformation gegeben (siehe Gleichung {4.35}). Unbekannt bleibt noch $\mathbf{H} \mathbf{B}$, d. h. die Kovarianz zwischen den Beobachtungsorten und den Analysepunkten ohne Beobachtung. Grundgedanke des Kovarianzmodells C ist es, die wichtigen Eigenvektoren der gesamten Kovarianzmatrix \mathbf{B} an den Stationsorten, d. h. für die Matrix $\mathbf{H} \mathbf{B} \mathbf{H}^T$ zu bestimmen und diese dann auf die anderen Analyseorte zu extrapolieren. Die Elemente eines Eigenvektors der Matrix $\mathbf{H} \mathbf{B} \mathbf{H}^T$ sind einem Stationort zugeordnet. So kann der Eigenvektor als ein an den Beobachtungsorten vorgegebenes skalares Feld aufgefasst werden. Mit einer räumlichen Interpolation mit dem homogenen terminbezogenen Kovarianzmodell \mathbf{A} können sie für alle Analysepunkte bestimmt werden. Damit wird die Kovarianzmodellierung für die Interpolation der

¹⁴¹ z. B. für eine effektive Speicherung von großen Matrizen beim reduced rank square root (RRSR) Kalman-Filter bzw. für die eindeutige Singularisierung bei linearen Gleichungssystemen.

Immissionsbeobachtung auf eine weitere Interpolationsaufgabe zurückgeführt. Im Gegensatz zu den originalen Größen kann jedoch bei den Elementen der bedeutsamen Eigenvektoren von einer größeren Glattheit und damit einer einfacheren Interpolation ausgegangen werden (Smith, 2001). Die Eigenvektoren, multipliziert mit den entsprechenden Eigenwerten, werden dann zu einer Kovarianzmatrix $\hat{\mathbf{B}}$ für alle Ortspaare rekombiniert:

1. empirische Kovarianzmatrix für die Stationsorte:

$$\hat{\mathbf{C}} = \widehat{\mathbf{H}\mathbf{B}\mathbf{H}^T} \quad \mathbf{C} \in R^{n \times n} \quad n \dots \text{Anzahl der Beobachtungen}$$

2. Eigenvektoraufspaltung:

$$\hat{\mathbf{C}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \quad \mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2 \dots \mathbf{v}_n] \quad \mathbf{v}_i \in R^n$$

3. Reduzierung für $\lambda_k \ll \lambda_1$:

$$\mathbf{V} \rightarrow \tilde{\mathbf{V}} = [\mathbf{v}_1, \mathbf{v}_2 \dots \mathbf{v}_k] \quad k \ll n \quad \{5.21\}$$

4. Interpolation ($R^n \rightarrow R^N$ $N \dots$ Anzahl der Gitterpunkte):

$$\mathbf{v}_i \rightarrow \mathbf{v}_i^* \quad \mathbf{v}_i^* \in R^N$$

5. Rekombination der Kovarianzmatrix für alle Feldpunkte:

$$\hat{\mathbf{C}}^* = \tilde{\mathbf{V}}^* \tilde{\mathbf{\Lambda}} \tilde{\mathbf{V}}^{*T} = \hat{\mathbf{B}} \quad \mathbf{C}^*, \hat{\mathbf{B}} \in R^{N \times N}$$

Die Bestimmung der Eigenvektoren und Eigenwerte der gesuchten Backgroundkovarianzmatrix kann nicht direkt auf der Kovarianzmatrix der Beobachtungsinkremente beruhen, da diese die Beobachtungsfehlervarianz enthält. Um $\mathbf{H}\mathbf{B}\mathbf{H}^T$ zu schätzen, ist folglich \mathbf{R} von der empirischen Kovarianzmatrix zu subtrahieren, d. h. die Hauptdiagonalelemente müssen um die Werte der Beobachtungsfehlervarianz (siehe Kapitel 5.2.6) verringert werden. Nach der Bestimmung der Eigenwerte (Press, 1992) stellt sich die Frage, ab welchem Eigenwert λ_i die zugehörigen Eigenvektoren vernachlässigbar sind. Hierzu kann der zugehörige Anteil an der Gesamtvarianz, der sich durch die Summe aller Eigenwerte ergibt, verwendet werden. Ist der Anteil größer als 5%, so wird der zugehörige Eigenwert als relevant betrachtet:

$$\frac{\lambda_i}{\sum_{j=1}^N \lambda_j} \geq 0.05 \quad \{5.22\}$$

Van Egmond und Onderdelinden (1981) verwenden Eigenwerte bis zur Größe des geschätzten Messinstrumentenfehlers. Hier wird die SVD auf die „kontaminierte“ Kovarianz, d. h. die Schätzung von $\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}$, angewandt. Die Frage, ob der Beobachtungsfehler von den Eigenvektoren mit Eigenwerten in dieser Größe erfasst und durch dieses Vorgehen entfernt wird, bleibt unbeantwortet.

5.2.5.3 SVD für numerische Stabilität

Wendet man eine Hauptachsentransformation auf die Matrix \mathbf{HBH}^T an, so werden die Kovarianzwerte an den Beobachtungsorten linear kombiniert. Diese „neuen Beobachtungsorte“ weisen nun keine Korrelation hinsichtlich der Backgroundfehler auf. Dadurch vereinfacht sich das Gleichungssystem und seine Lösung. Gleichzeitig lässt sich anhand der Eigenwerte überprüfen, ob die Kovarianzmatrix positiv definit oder schlecht konditioniert¹⁴² ist.

Ein lineares Gleichungssystem $\mathbf{Ax}=\mathbf{b}$ kann hinsichtlich der Eigenvektorbasis \mathbf{V} umgeformt werden, wobei die Matrix \mathbf{A} durch $\mathbf{\Lambda}$ ersetzt wird:

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} &\rightarrow \mathbf{\Lambda x}^* = \mathbf{b}^* \\ \mathbf{x}^* = \mathbf{V}^T \mathbf{x} &\text{ und } \mathbf{b}^* = \mathbf{V}^T \mathbf{b} \end{aligned} \quad \{5.23\}$$

Befinden sich unter den ausgewählten Stationen zwei eng benachbarte und haben sie einen geringen Beobachtungsfehler¹⁴³, so hat die Matrix des Gleichungssystems zwei ähnliche Spalten und Zeilen. Dadurch ist das Gleichungssystem schlecht konditioniert, was zu numerischen Problemen bei seiner Lösung führt. Die Konditionierung einer reellen, symmetrischen und positiv definiten Matrix lässt sich aus dem Verhältnis des größten zum kleinsten Eigenwert ableiten (Kiesewetter und Maess, 1974). In der numerischen Praxis empfiehlt es sich, die Eigenvektoren mit sehr kleinen Eigenwerten zu vernachlässigen und nur die approximierte Matrix \mathbf{A}' bzw. $\mathbf{\Lambda}'$ zu verwenden¹⁴⁴. Im Gegensatz zur SVD bei dem Kovarianzmodell \mathbf{C} steht hierbei nicht der Aspekt der Handhabbarkeit sondern jener der numerischen Stabilität im Vordergrund. Deswegen werden hierbei erst Eigenwerte, die das um 10-fache kleiner als die in {5.22} sind, vernachlässigt.

5.2.6 Schätzung der Beobachtungsfehlervarianz

Der Beobachtungsfehler wird als eine räumlich nicht korrelierte Komponente der Beobachtung mit verschwindendem Erwartungswert angesehen (siehe Kapitel 4.5.2). Die Beobachtungsmethode nach Hollingsworth und Lönnberg (1986) benutzt die Beobachtungssinkremente¹⁴⁵, d. h. die Differenz zwischen Beobachtung und Background ($\mathbf{y} - \mathbf{Hx}_B$) zur Schätzung seiner Varianz.

Bildet man die Kovarianzmatrix der Beobachtungssinkremente unter der Annahme, dass die Fehler des Backgrounds und der Beobachtung nicht korreliert sind, so ergibt sie sich aus der Kovarianzmatrix der Beobachtungen \mathbf{R} und der mit dem linearisierten Beobachtungsoperator \mathbf{H} auf die Variable der Beobachtungen transformierten Matrix des Backgroundfehlers \mathbf{B} (vergleiche Gleichung {4.23} und {4.24}):

$$E\left((\mathbf{y} - \mathbf{Hx}_B)(\mathbf{y} - \mathbf{Hx}_B)^T\right) = \mathbf{R} + \mathbf{HBH}^T \quad \{5.24\}$$

¹⁴² Die Matrix ist „fast“ singular.

¹⁴³ Die Varianz des Beobachtungsfehlers trägt zur numerischen Stabilität bei.

¹⁴⁴ d. h. die Eigenwerte und ihre reziproken Werte gleich Null zu setzen

¹⁴⁵ die durch die „nahrhafte“ Null ($\mathbf{Hx}_{\text{true}} - \mathbf{Hx}_{\text{true}}$) erweitert werden

Mit der weiteren Annahme, dass die Beobachtungsfehler nicht korreliert sind¹⁴⁶ und der Beobachtungsoperator \mathbf{H} eine Verschiebung auf den Ort der Beobachtung ist, folgt für einen Kovarianzwert C_{ij} der Beobachtungsincremente an den Orten i und j :

$$C_{ij} = E((y_i - Hx_i)(y_j - Hx_j)) = \begin{cases} \sigma_{Bi}^2 + \sigma_{Oi}^B & i = j \\ b_{ij} & i \neq j \end{cases} \quad \{5.25\}$$

Für verschiedene Orte i und j gleicht der Kovarianzwert der Beobachtungsincremente dem der Backgroundfehler, d. h. dem entsprechenden Element aus \mathbf{HBH}^T ; für gleiche Orte entspricht der Wert der Summe der Varianzen des Beobachtungsfehlers σ_{Oi}^2 und des Backgrounds σ_{Bi}^2 .

Der extrapolierte Wert eines Kovarianzmodells für den Abstand Null wird nun als Schätzung der Varianz des Backgroundfehlers verwendet. Die Güte der Schätzung wird sowohl durch die Gültigkeit der statistischen Schätzung der Kovarianzen als auch durch die Wahl des Kovarianzmodells bestimmt¹⁴⁷ (siehe Kapitel 5.2.3). Von Bedeutung ist dabei analytischen Form des parametrischen Kovarianzmodells.

Die Differenz zur Varianz der Inkremente liefert dann die gesuchte Varianz des Beobachtungsfehlers σ_{Oi}^2 . Das Prinzip der Schätzung der Varianz des Beobachtungsfehler mit Hilfe eines Kovarianzmodells ist in Abbildung 5.8 dargestellt.

Für den homogenen terminbezogenen Ansatz ergibt die Beobachtungsmethode einen Wert für die Varianz des Beobachtungsfehlers aller Stationen. Mit dem klimatischen Ansatz können stationsspezifische Fehler ermittelt werden. Für jede Station ist aus der Zeitreihenauswertung der Beobachtungsincremente die Gesamtvarianz bekannt. Die Aufgabe besteht nun in einer möglichst guten Schätzung der Varianz des Backgrounds bzw. des glatten Feldes am Stationsort. Hierzu wird ein lokales homogenes und isotropes Kovarianzmodell auf Basis der empirischen Kovarianzwerte mit der betrachteten Station in einem Radius von 100 km herangezogen.

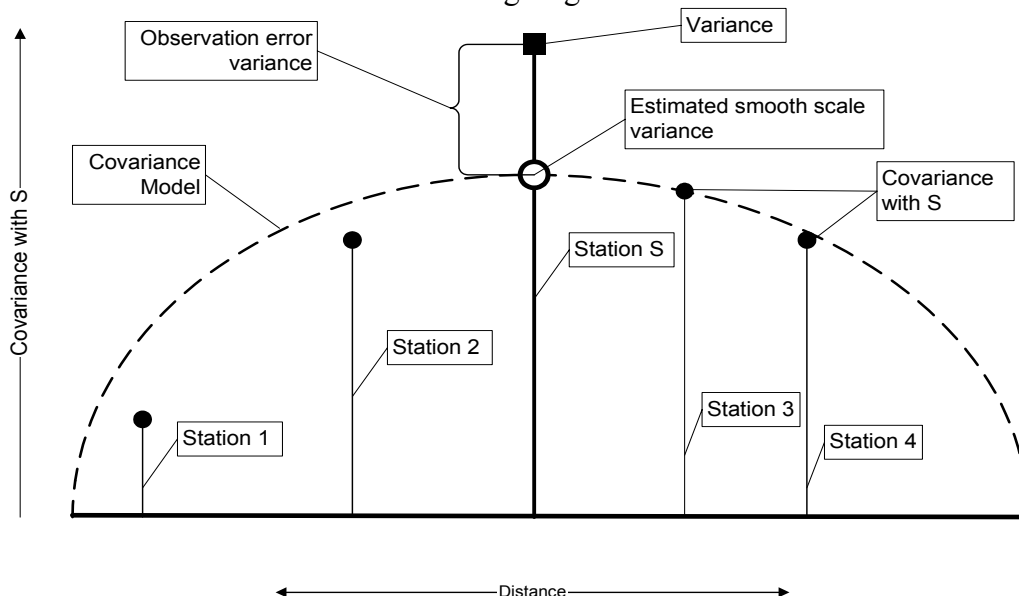


Abbildung 5.8 Schätzung der Varianz des klimatischen Beobachtungsfehlers (observation error variance) der Station S mit Hilfe eines Kovarianzmodells nach der Beobachtungsmethode.

¹⁴⁶ \mathbf{R} ist damit eine Diagonalmatrix

¹⁴⁷ Von besonderer Bedeutung ist die Steigung des Kovarianzmodells in der Nähe des Abstandes Null.

5.3 Diskussion der Ergebnisse der Kovarianzmodellierung

Dieses Kapitel beinhaltet die Darstellung der Ergebnisse der Kovarianzmodellierung nach Kapitel 5.2. Es werden im folgenden die Schätzungen der Beobachtungsfehlervarianz, des Modelbias und der räumlichen Kovarianz, einschließlich der Varianz¹⁴⁸ des „glatten“ Feldes, diskutiert.

Die räumliche Kovarianz wird für drei Datenquellen aus Kapitel 5.1.2 geschätzt, d. h. für die Beobachtungsinckremente, die Beobachtungen und die reinen Modelldaten. Die Kovarianz der Beobachtungsinckremente dient der Analyse, die Kovarianz der Beobachtung und Modelldaten wird für die klimatologische Darstellung und die Modellevaluierung benötigt. Der Erwartungswert der Beobachtungsinckremente, d. h. der Modelbias, gibt den systematischen Modellfehler an. Er muss für die Analyse beseitigt werden.

Häufig wird die Kovarianzmodellierung nur als Zwischenschritt zur Analyse betrachtet und ihren Ergebnissen keine große Aufmerksamkeit geschenkt. Hier schließt sich nun eine genauere Diskussion der Ergebnisse der Kovarianzmodellierung aus folgenden Gründen an:

- Die Größe des Beobachtungsfehlers, der hauptsächlich ein Maß für mangelnde Repräsentativität ist, soll veranschaulicht werden
- Der Vergleich von Bias und der räumlichen Kovarianz von Modellrechnung und Beobachtung ist eine Modellevaluierung, die 1. und 2. statistische Momente berücksichtigt.
- Bewertung des Vermögens der Kovarianzmodelle A, B und C aus Kapitel 5.1.3 die Struktur der räumlichen Kovarianz und deren zeitliche Veränderung widerzugeben
- Abschätzung der Stärke der Veränderung des Modellfeldes (Background) durch die Beobachtung bei der Analyse.
- Vergleich der Kovarianzmodellierung von passiven und aktiven Verfahren (Kalman-Filter, aktive Assimilation mit OI)

Die Frage des Erfolgs der Kovarianzmodelle der Beobachtungsinckremente für die Analyse ist in der Arbeit Inhalt des nächsten Kapitels 6. Die direkte Interpretation der Beobachtungsinckremente hat aufgrund des Differenzcharakters und der spezifischen Konfiguration des Chemie-Transport-Modells eine geringere Aussagekraft. Deshalb wird in diesem Kapitel die räumliche Struktur der Kovarianzen der „reinen“ Beobachtungen und der Modellwerte präsentiert und verglichen¹⁴⁹ (siehe Kapitel 5.1.2).

Die präsentierten Ergebnisse hängen von dem verwendeten Kovarianzmodell und seinen Annahmen ab. So erzeugt das homogene Modell A einen räumlich homogenen Wert der betrachteten Größen für jeden Termin. Die klimatischen Kovarianzmodelle liefern stationsbezogene bzw. räumlich inhomogene Werte für jeweils eine Stunde des Tages über einen langen Zeitraum. Sie beruhen auf der empirischen klimatischen Kovarianzmatrix der Beobachtungsorte. Die separate Bildung der klimatischen Kovarianzmodelle für jede Tagesstunde erfasst nur die tagesgangabhängige Variabilität.

¹⁴⁸ Die Varianzen werden dabei meist durch ihre Wurzel, d. h. die Standardabweichung, dargestellt.

¹⁴⁹ Für eine Modellevaluierung hinsichtlich der zweiten räumlichen Momente, d. h. der Kovarianzfunktion, ist es zunächst empfehlenswerter, die beiden abgeleiteten Kovarianzmodelle zu vergleichen und nicht das Kovarianzmodell der Abweichungen zu untersuchen.

Die betrachtete räumliche Variabilität bezieht sich innerhalb der gewählten Modellvorstellung des Zufallsfeldes (siehe Kapitel 4.4.1) auf die Abweichungen vom Erwartungswert. Hier wird zusätzlich auf den Erwartungswert der Beobachtungsinkremente eingegangen, da er der Bias zwischen Modell und Beobachtung ist.

Von besonderer praktischer Bedeutung ist der ermittelte Beobachtungsfehler, da er für alle Arbeiten zur statistischen Analyse und Datenassimilation sowie für die Modellevaluierung¹⁵⁰ von grundlegender Bedeutung ist.

Die Auswertung der terminbezogenen homogenen Kovarianzmodelle (A) für verschiedene Termine eignet sich zur Untersuchung der zeitlichen Variabilität der Kovarianz und der Stärke der Anisotropie. Die Darstellung der inhomogenen klimatischen Kovarianzmodelle (B, C) liefert einen Eindruck von der Ausprägung der Inhomogenität der zweiten Momente der Felder bzw. von der Möglichkeit, sie zu erfassen (siehe Abbildung 5.1). Die Gegenüberstellung beider Ansätze kann die Rechtfertigung der Grundannahmen zur Stichprobenbildung, d. h. Homogenität oder Stationarität, des jeweils anderen Ansatzes überprüfen.

Die Kovarianzmodellierung erfolgte empirisch unter der Verwendung von Beobachtungsdaten. Eine dynamische Form der Kovarianzmodellierung beinhaltet der Kalman-Filters (siehe Kapitel 4.5.7) für das Chemie-Transport-Modell REM/Calgrid (siehe Kapitel 1.4.2). Ein Vergleich der empirischen Kovarianzmodellierung mit den Kovarianzen des Kalman-Filters wird in Kapitel 5.3.5 gegeben.

5.3.1 Der Beobachtungsfehler

Nach der Definition in Kapitel 4.5.2 ist der Beobachtungsfehler eine räumlich nicht korrelierte Komponente der Messung. Seine Varianz wird mit der Beobachtungsmethode aus Kapitel 5.2.6 geschätzt. Der Umstand der Unkorreliertheit wird von den hier gewonnenen Ergebnissen bestätigt: Unabhängig davon, ob Beobachtungen oder Beobachtungsinkremente verwendet werden, sind die ermittelten Fehlergrößen sehr ähnlich. Weiterhin hat die Wahl der analytischen Kovarianzfunktion (siehe Kapitel 5.2.3) keinen erkennbaren Einfluss auf die geschätzten Beobachtungsfehlervarianzen.

In der Literatur wird der Beobachtungsfehler häufig ad hoc als Prozentsatz des Messwertes angenommen oder mit statistischen Methoden geschätzt (Tilmes und Zimmerman, 1998). Die in der Arbeit gewonnenen Ergebnisse sprechen insbesondere bei Ozon gegen dieses Vorgehen, da für alle Regimes ein gleich großer Beobachtungsfehler ermittelt wurde und weil der absolute Beobachtungsfehler am Nachmittag geringer als in den Nachstunden¹⁵¹ ist.

5.3.1.1 Der stationsbezogene Beobachtungsfehler

Die Gesamtheit der stationsbezogen bestimmten Standardabweichungen, d. h. die Wurzel der Varianz, der Beobachtungsfehler aller betrachteten Spurenstoffe, ist in Abbildung 5.9 zu finden. Abbildung 5.12 gibt einen Überblick über die räumliche Verteilung der Beobachtungsfehler für Ozon und PM10.

¹⁵⁰ Die „perfekte“ Modellsimulation muss demzufolge nur im Rahmen der Standardabweichung des Beobachtungsfehlers mit der Messung übereinstimmen.

¹⁵¹ Die Ozonwerte sind nachmittags höher als in der Nacht

Für NO und PM10 sind die geschätzten stationsbezogenen Beobachtungsfehler deutlich größer als für NO₂, Ozon und SO₂ an allen Tagesstunden. Dies entspricht der Erwartung, dass Messungen für das kurzlebige NO in der Nähe der Quellen von nur geringer räumlicher Repräsentativität sind. Die stündliche Variabilität von PM10 ist sehr hoch und dementsprechend werden große Fehler geschätzt. Da PM10 keinen ausgeprägten Tagesgang aber eine große Tagesschwankung besitzt (siehe Kapitel 3.9.1), ist es im Sinne der räumlichen Repräsentativität empfehlenswert, die stündlichen Werte durch das Tagesmittel zu ersetzen. Die zugehörigen Beobachtungsfehler sind dann deutlich geringer.

Der Zusammenhang zwischen Beobachtungsfehler und dem Immissionsregime, d. h. der typischen Belastung, ist in Abbildung 5.10 und Tabelle 7 dargestellt. Für die primär emittierten Spurenstoffe, d. h. mit Ausnahme von Ozon, steigt der Beobachtungsfehler mit zunehmender mittlerer Immission an. Bei PM10, SO₂ und NO bleibt der relative, d. h. auf die mittlere Belastung bezogene Fehler, nahezu konstant in allen Regimes; bei NO₂ sinkt der relative Fehler leicht mit zunehmender Belastung.

Stoff/Regime:	B	R / 1	U1 / 2	U2 / 3	U3 / 4	S / 5	S2	Tagesgang
O ₃ (ppb)	8	6	6	6	6	6		schwach
NO ₂ (ppb)		3	4	5	6	7	8	schwach
NO (ppb)		3	7	10	15	20	30	8 Uhr-Max
PM10(µg/m ³)		8	8	10	12	15		schwach
SO ₂ (µg/m ³)		2	3	4	6	18		schwach

Tabelle 7 Geschätzte Standardabweichung des klimatischen Beobachtungsfehlers der stündlichen Messung, dargestellt als Median aller Stationen innerhalb eines Regimes (Erklärung der Regimes in Tabelle 4)

Bei Ozon ist die Standardabweichung des Beobachtungsfehlers für alle Regimes mit Ausnahme der Berg-Stationen nahezu konstant. Die scheinbar geringe räumliche Repräsentativität der Berg-Stationen erklärt sich aus dem hier gewählten Vorgehen, den Beobachtungsfehler mit Hilfe der umgebenden Stationen abzuschätzen. In der Tat sind die Berg-Stationen aufgrund der geringen Deposition in der Nacht nicht repräsentativ für die Verhältnisse in der Bodenschicht über Land. Sie entsprechen eher den Verhältnissen in der zweiten Modellschicht. Der relative Beobachtungsfehler der Berg-Stationen ist ungefähr von gleicher Größe wie der der Land-Stationen. Mit zunehmender Urbanisierung und Verkehrseinfluss steigt der relative Beobachtungsfehler.

Der Ausgangspunkt für die klimatische Kovarianzmodellierung sind separate Kovarianzmodelle für jede Tagesstunde. Betrachtet man die geschätzten Standardabweichungen des Beobachtungsfehlers für jede Tagesstunde, so stellt man nur einen geringen Tagesgang fest (siehe Abbildung 5.11). Bei NO gibt es während des morgendlichen Immissionsmaximums einen Anstieg des Fehlers. Interessant ist, dass für Ozon in der Zeit der höchsten Konzentrationen eine leichte Abnahme des absoluten und eine große des relativen klimatischen Beobachtungsfehlers festzustellen ist. Offensichtlich ist bei der nachmittäglichen Ozonbildung und guter Durchmischung die räumliche Repräsentativität größer als in den nächtlichen Stunden, wo lokale NO-Immission und Depositionsvorgänge die Ozonkonzentration steuern.

SD of climatic Obs. error 2001 all stations

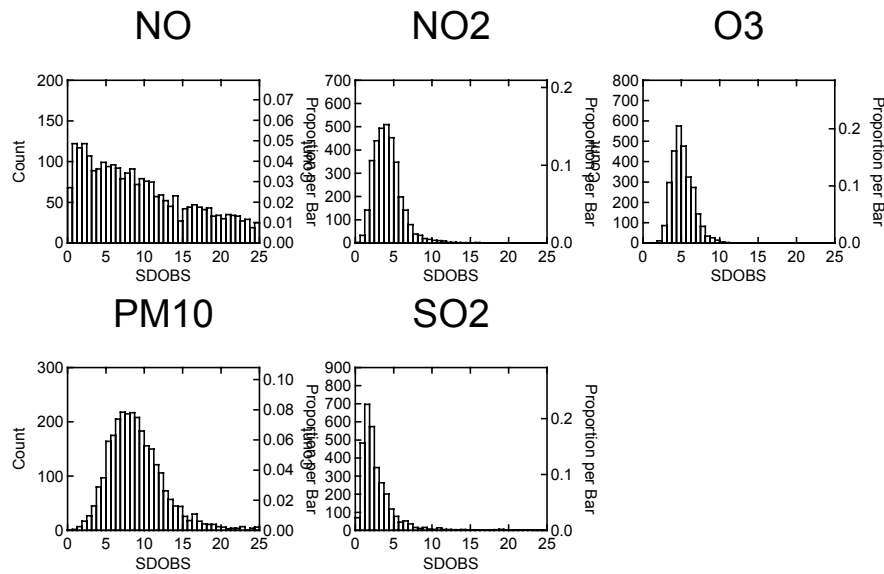


Abbildung 5.9 Histogramme der Standardabweichung des klimatischen Beobachtungsfehlers (SDOBS) für alle Stationen und alle Tagesstunden (oben in ppb, unten in $\mu\text{g}/\text{m}^3$)

SD of climatic Obs. error vs AQ regime

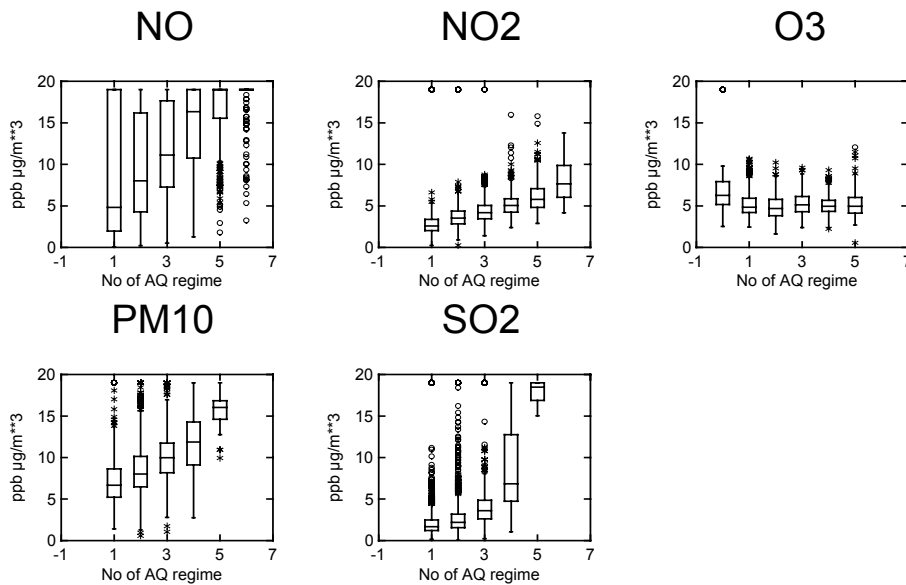


Abbildung 5.10 Standardabweichung¹⁵² des klimatischen Beobachtungsfehlers für alle Stationen und alle Tagesstunden, geordnet nach dem Immissionsregime (obere Reihe in ppb, 0=Berg/Küste, 1=ländlich, 2=Vorstadt, 3=Stadt, 4=belastete Stadt, 5=Verkehr, 6=Verkehr extrem, untere Reihe in $\mu\text{g}/\text{m}^3$, Belastungsstufen 1 – 5)

¹⁵² Die grafische Darstellung beschränkt sich auf Werte kleiner als 20 ppb bzw. $\mu\text{g}/\text{m}^3$

Daily Var. of climatic Obs. error (SD)

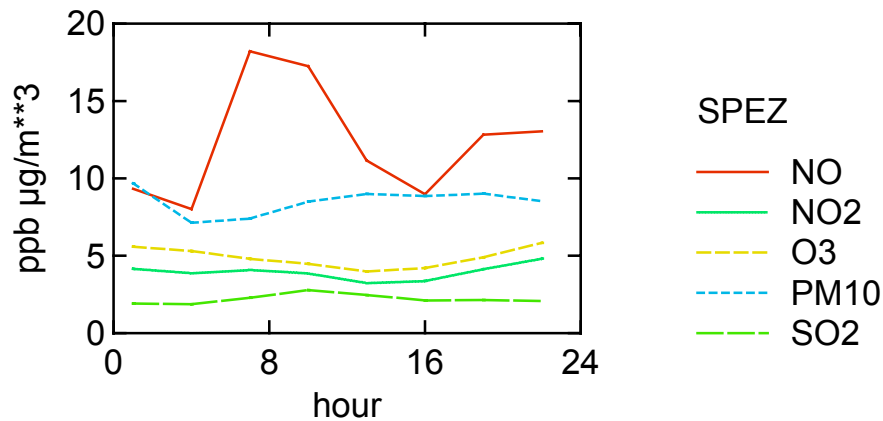


Abbildung 5.11 Tagesgang der Standardabweichung des klimatischen Beobachtungsfehlers für alle Stationen (O3, NO2 und NO in ppb, SO2 und PM10 in $\mu\text{g}/\text{m}^3$)

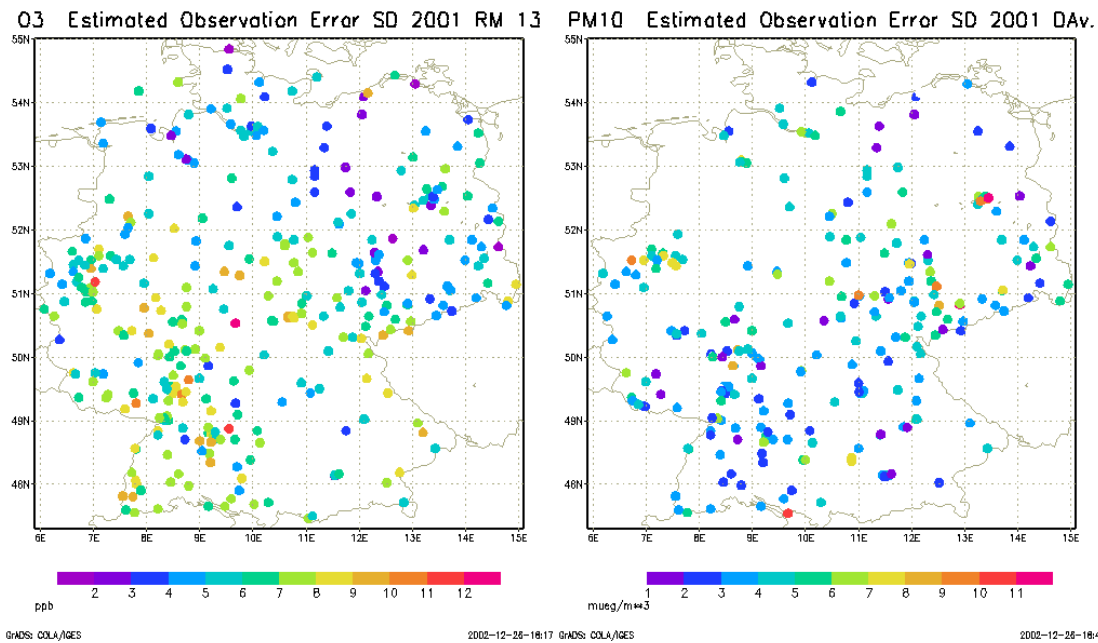


Abbildung 5.12 Karte der stationsbezogenen Standardabweichung des Beobachtungsfehlers für Ozon um 13 Uhr (links) und der Tagesmittelwerte von PM10 (rechts)

5.3.1.2 Der homogene terminbezogene Beobachtungsfehler

Im Gegensatz zum stationsbezogenen Beobachtungsfehler liegt der terminbezogene Beobachtungsfehler als Zeitreihe für alle Stunden des Jahres vor. Pro Termin fasst ein Wert den typischen Beobachtungsfehler aller Stationen zusammen, mit Ausnahme der Stationen der Verkehrs-Regimes (S, S2 bzw. 5, siehe Tabelle 4 in Kapitel 2.3.2.2). Die

Charakteristik dieses Wertes ist durch die urbanen Stationen geprägt, da sie den größten Anteil der verwendeten Stationen stellen.

Der terminbezogene Beobachtungsfehler ist auf Basis des terminbezogenen Kovarianzmodells A ermittelt worden (siehe Kapitel 5.1.3). Durch die Zusammenschau der terminbezogenen Kovarianzmodelle für alle Zeitpunkte kann die zeitliche Variabilität der Kovarianzstruktur untersucht werden.

Abbildung 5.13 zeigt analog zu Abbildung 5.9 die Häufigkeitsverteilung der Standardabweichung des homogenen Beobachtungsfehlers für alle Termine. Die Wertebereiche sind von ähnlicher Größenordnung wie die der individuellen klimatischen Fehler. Für NO und PM10 sind die individuellen Unterschiede größer als die Unterschiede zwischen den homogenen Werten zu verschiedenen Zeitpunkten. Die Darstellung der Beobachtungsfehler, geordnet nach Tagesstunden, liefert ein ähnliches Bild wie im klimatischen Fall (siehe Abbildung 5.14 und Abbildung 5.11).

Der Jahresgang, ermittelt durch die monatliche Zusammenfassung des homogenen Beobachtungsfehlers, enthält Abbildung 5.15. Für NO₂ und PM10 ergibt sich keine erkennbare Jahresgangabhängigkeit, NO und SO₂ zeigen ein Sommerminimum, während Ozon ein leichtes Maximum im Sommer aufweist.

Insgesamt kann festgestellt werden, dass der homogene Beobachtungsfehler keinen sehr ausgeprägten Jahresgang liefert. Dies rechtfertigt das Konzept des klimatischen Beobachtungsfehlers, da die individuellen Unterschiede größer erscheinen als die zeitliche Variabilität.

SD of homogeneous Obs. error all hours

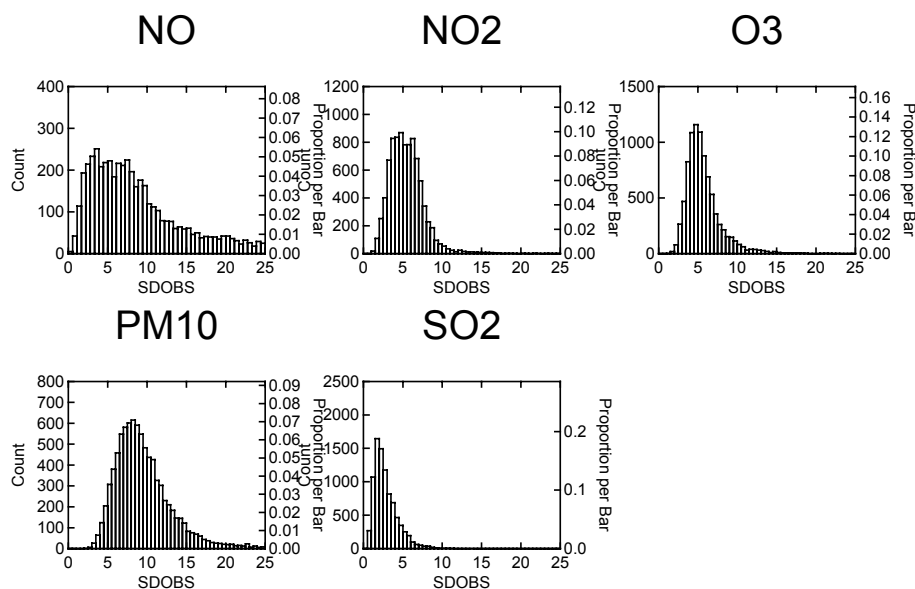


Abbildung 5.13 Histogramme der Standardabweichung des terminbezogenen homogenen Beobachtungsfehlers (SDOBS) für alle Termine im Jahr 2001 (obere Reihe in ppb, untere in $\mu\text{g}/\text{m}^3$)

Daily Var. of hom. Obs. error (SD)

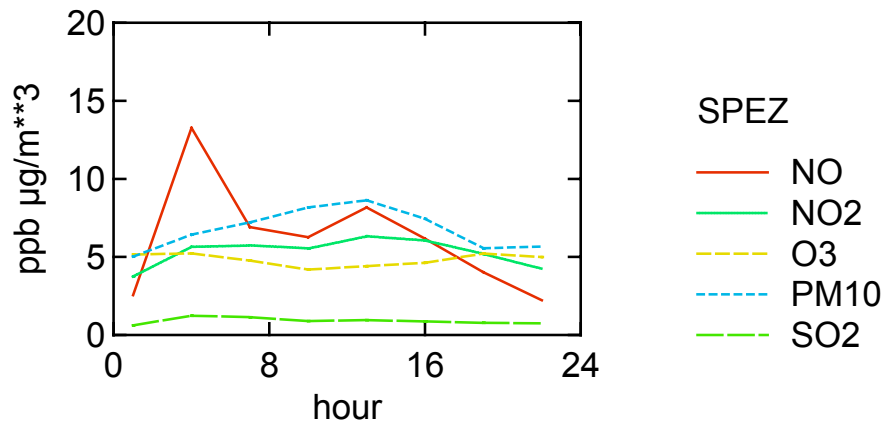


Abbildung 5.14 Tagesgang der Standardabweichung des homogenen terminbezogenen Beobachtungsfehlers für alle Stunden des Jahres (O₃, NO₂ und NO in ppb, SO₂ und PM₁₀ in µg/m³)

Annual Var. of hom. Obs. error (SD)

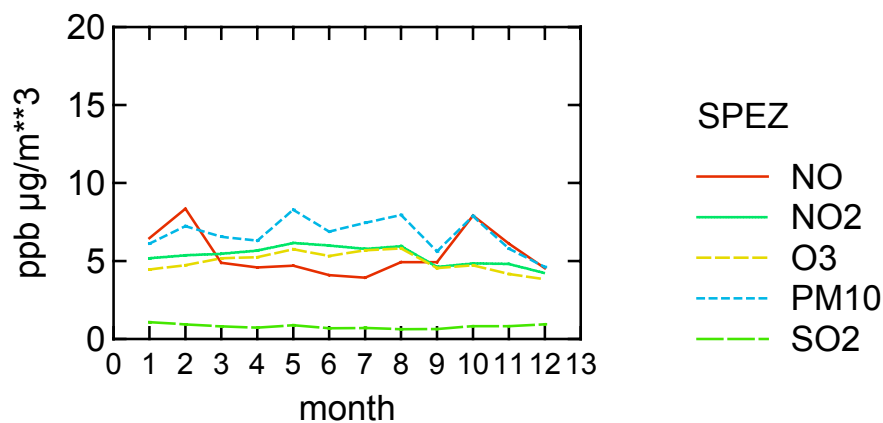


Abbildung 5.15 Jahrgang der Standardabweichung des homogenen Beobachtungsfehlers für alle Termine (für O₃, NO und NO₂ in ppb, für SO₂ und PM₁₀ in µg/m³)

5.3.2 Der Bias

Der Bias beschreibt die Abweichung der Erwartungswerte von Modell und Beobachtung und ist demzufolge von grundlegender Bedeutung für die Modellevaluierung. Ein positiver Bias ist Anzeichen für eine Unterschätzung der Beobachtung durch das Chemie-Transport-Modell. Für die Analyse stellt sich zusätzlich die Frage, inwieweit die Hypothese vom Erwartungswert Null der Beobachtungsinkremente {4.33} Gültigkeit hat bzw. durch welche Korrektur eine Biasfreiheit erreicht werden kann. Die Gegenüberstellung der zeitlichen Variabilität innerhalb der Regimes (siehe Kapitel 3.9) gibt einen Einblick in die zeitliche Variabilität des Bias. Eine Diskussion des Bias-

Tagesgangs für verschiedene chemisch gekoppelte Stoffe liefert weitere Hinweise zum Modellverständnis.

Die Histogramme des terminbezogenen und des stationsbezogenen Bias sind in Abbildung 5.16 und Abbildung 5.17 dargestellt. Die Maximalwerte des Bias liegen in beiden Fällen in der Größe der Erwartungswerte selbst. Die individuellen Unterschiede oder die Unterschiede zu verschiedenen Zeiten haben eine ähnlich große Spannweite.

Wie schon in Kapitel 5.1.1 ausgeführt, ist die Bewertung des terminbezogenen Bias zwischen Modell und Beobachtung nicht ohne eine Berücksichtigung der Zusammensetzung der Beobachtungen hinsichtlich ihrer Regimezugehörigkeit beantwortbar. Der stationsbezogene Bias im klimatischen Fall ist in Abbildung 5.18 für die verschiedenen Regimetypen der Beobachtung dargestellt. Bei NO_x weisen die ländlichen Regimes fast keinen Bias auf; mit zunehmender Belastung steigt der Bias an. Dieses Verhalten ist konsistent mit der Annahme, dass diese Regimes nicht die räumliche Ausdehnung einer Modellgitterbox haben und damit nicht vom Modell erfasst werden können¹⁵³. Trotz einer gewissen Streuung ist die Simulation von SO_2 in allen Regimes nahezu biasfrei, d. h. auch die belasteteren Regimes werden vom Modell richtig erfasst. Bei Ozon und PM10 sind die suburbanen bzw. die Stationen der Belastungsstufe 2 biasfrei. Die Werte der Land-Stationen der Ozonbeobachtung werden im Mittel unterschätzt bzw. bei PM10 vom Modell überschätzt.

Der bisher betrachtete Bias setzt sich gleichmäßig aus allen Stunden zusammen. Eine genauere Beurteilung der Ursachen des Bias kann eine Betrachtung seines Tagesganges ergeben. Der Tagesgang des Bias für alle Stationen bzw. nur für die Land-Stationen ist in Abbildung 5.19 und Abbildung 5.20 dargestellt. Es zeigt sich dabei eine deutliche Tagszeitabhängigkeit, die vom terminbezogenen und stationsspezifischen Ansatz qualitativ gleichwertig wiedergegeben wird.

Am auffälligsten ist der gegenläufige Tagesgang des Bias für Ozon und PM10. Die Ozonsimulation ist nachts deutlich zu gering und tagsüber für die Gesamtheit der Stationen zu hoch. Letzteres gilt für die Land-Stationen nur in geringem Umfang. Für PM10 gilt der umgekehrte Verlauf mit einer nächtlichen Überschätzung und einer Unterschätzung am Tage. Da beide Stoffe nicht unmittelbar chemisch miteinander gekoppelt sind, ist die mögliche Ursache dieses Verhaltens vorrangig in den vertikalen Austauschprozessen zu vermuten: Stabile Verhältnisse sind zu stabil und labile zu labil. Der zu geringe nächtliche Austausch führt zu erhöhten PM10-Werten und zu überschätztem Ozonabbau durch Deposition und Titration mit NO ohne eine Kompensation durch vertikale Ozoneinmischung in die Bodenschicht. Tagsüber führt der überschätzte Austausch zur Verringerung der PM10-Konzentration und zu Ozoneinmischung von oben bzw. zu einer sehr starken Verdünnung des NO .

Ozon, NO und NO_2 sind durch das sich in wenigen Minuten einstellende fotochemische Gleichgewicht gekoppelt. Die nachmittägliche Überschätzung von Ozon ist mit unterschätzten NO_2 -Werten verbunden, was für richtige O_x ¹⁵⁴-Verhältnisse sprechen könnte. Nachts gilt das nur bei den Land-Stationen, da für diese eine leichte Unterschätzung durch das Modell im klimatologischen Fall zu verzeichnen ist.

¹⁵³ Dieser Umstand widerspricht nicht der Tatsache, dass einige Gitterboxen z. B. im Ruhrgebiet als Gebiete höchster NO_x -Belastung simuliert werden (siehe Kapitel 3.4 und 3.5)

¹⁵⁴ O_x ist die Summe aus NO_2 und O_3 . Die Untersuchung von O_x ist sinnvoll, da dadurch die Bilanz über das zwischen beiden Stoffen ausgetauschte O-Radikal gebildet wird. Bei unveränderter Einstrahlung und konstantem NO_x wird das Verhältnis der beiden Stoffe nur von der meist sehr variablen NO -Konzentration gesteuert.

Der Jahresgang des homogenen terminbezogenen Bias für 2001 auf der Basis von Monatswerten zeigt große Schwankungen zwischen den verschiedenen Monaten (alle Stationen in Abbildung 5.21, Land-Stationen in Abbildung 5.22). Frühjahr und Winter sind Zeiten mit der stärksten Unterschätzung von Ozon und mit der stärksten Überschätzung von PM10. Insbesondere die starke Schwankung von Monat zu Monat innerhalb einer Jahreszeit unterstützt die These, dass die Wettersituation sehr wichtig für die Entstehung des Bias von Ozon und PM10 ist. Auch hier ergibt sich das beim Tagesgang gefundene gegenläufige Verhalten der jeweiligen Bias.

NO und NO₂ zeigen einen geordneteren Jahresgang mit einer Unterschätzung der Gesamtheit der NO-Beobachtungen im Winter und einer Überschätzung von NO₂ im Sommer. Betrachtet man lediglich die Land-Stationen, so ist nur für NO₂ eine leichte sommerliche Unterschätzung und eine winterliche Überschätzung zu erkennen. Für das ländliche Ozon und PM10 ist die winterliche Unterschätzung bzw. Überschätzung besonders ausgeprägt.

Bias of climatic Model error 2001 all stations

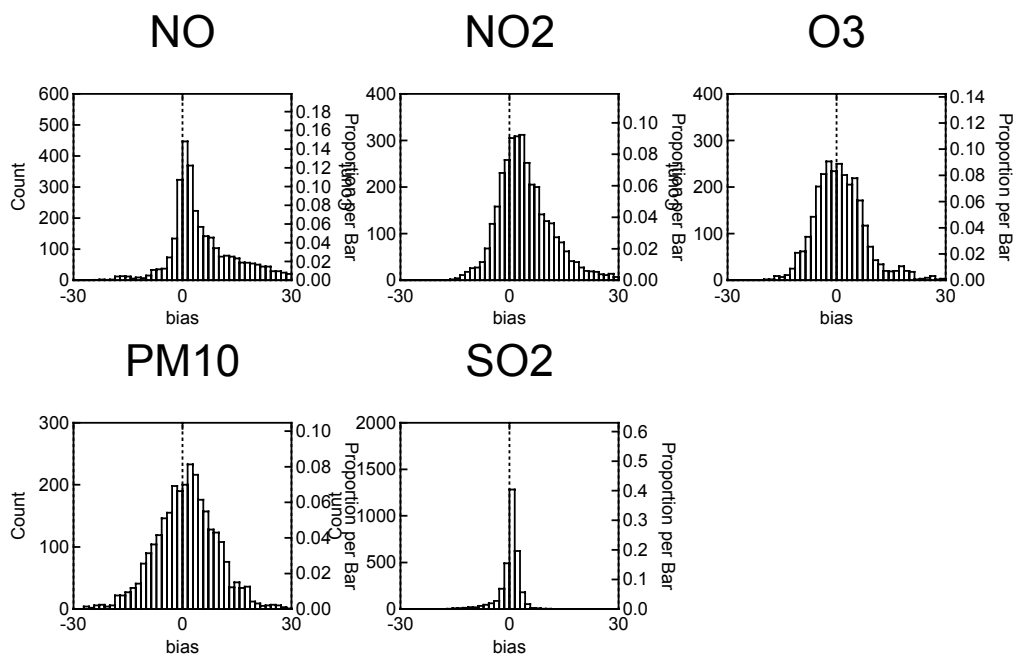


Abbildung 5.16 Histogramme des stationsbezogenen Bias zwischen Modell und Beobachtung für alle Stationen (obere Reihe in ppb, unten in $\mu\text{g}/\text{m}^3$)

Bias of homogeneous Model error 2001 all hours

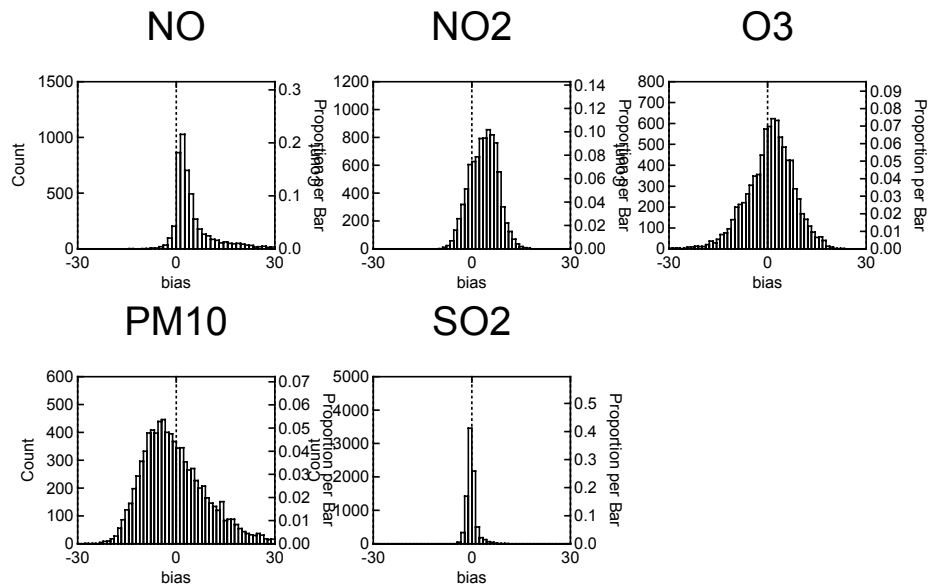


Abbildung 5.17 Histogramme des terminbezogenen homogenen Bias zwischen Modell und Beobachtung für alle Termine 2001 (obere Reihe in ppb, unten in $\mu\text{g}/\text{m}^3$).

Bias of climatic Obs. error vs AQ regime

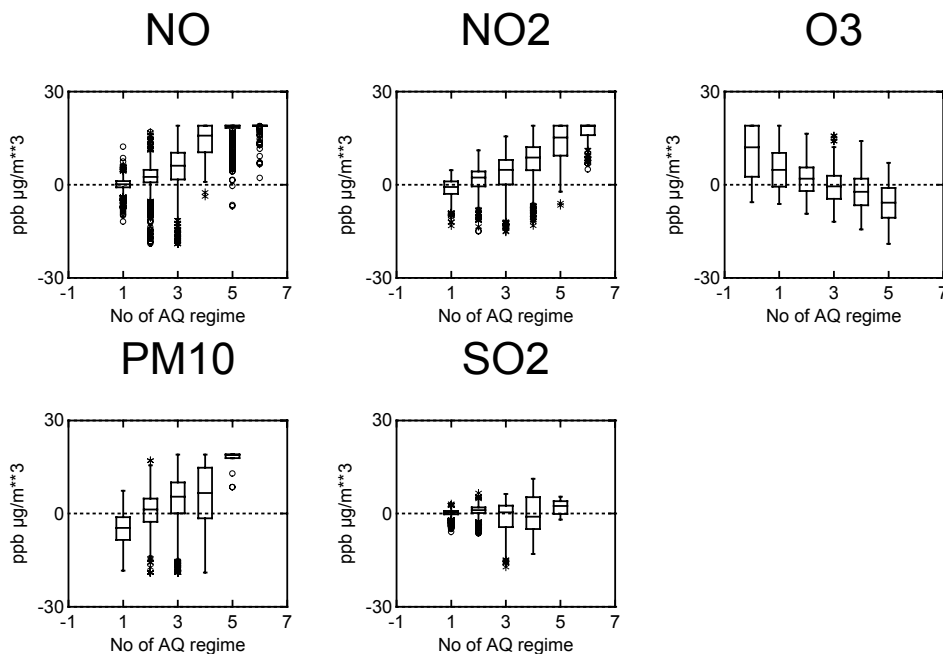


Abbildung 5.18 Bias zwischen Beobachtung und Modellrechnung für alle Stationen, geordnet nach dem Regime (AQ regime) der Beobachtung (obere Reihe in ppb, 0=Berg/Küste, 1=ländlich, 2=Vorstadt, 3=Stadt, 4=belastete Stadt, 5=Verkehr, 6=Verkehr extrem, untere Reihe in $\mu\text{g}/\text{m}^3$, Belastungsstufen 1– 5).

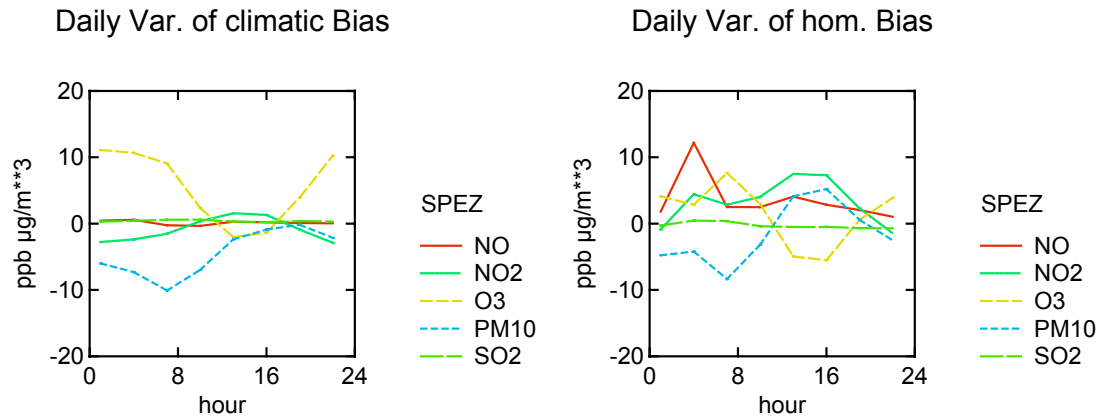


Abbildung 5.19 Tagesgang des klimatischen Bias für alle Stationen (links) und des homogenen terminbezogenen Bias für alle Stunden des Jahres 2001 (rechts) (O3, NO2 und NO in ppb, SO2 und PM10 in µg/m³)

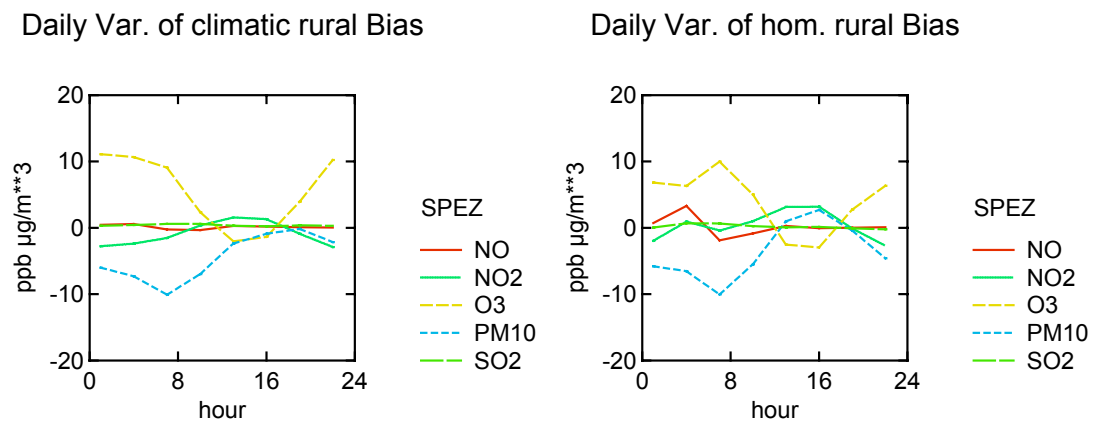


Abbildung 5.20 Tagesgang des klimatischen Bias für alle Land-Stationen (links) und des homogenen terminbezogenen Bias der Land-Stationen für alle Stunden des Jahres (rechts) (O3, NO2 und NO in ppb, SO2 und PM10 in µg/m³)

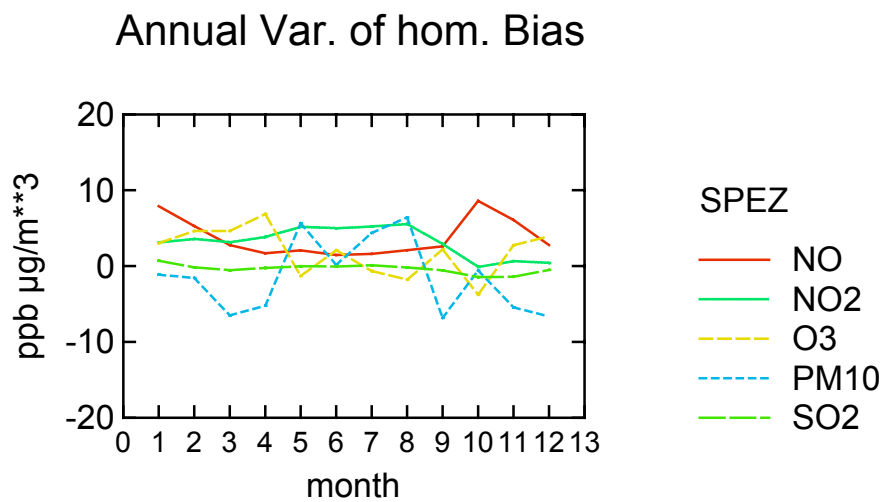


Abbildung 5.21 Jahresgang des homogenen terminbezogenen Bias aller Stationen für alle Termine 2001 (O3, NO und NO₂ in ppb, für SO₂ und PM10 in µg/m³)

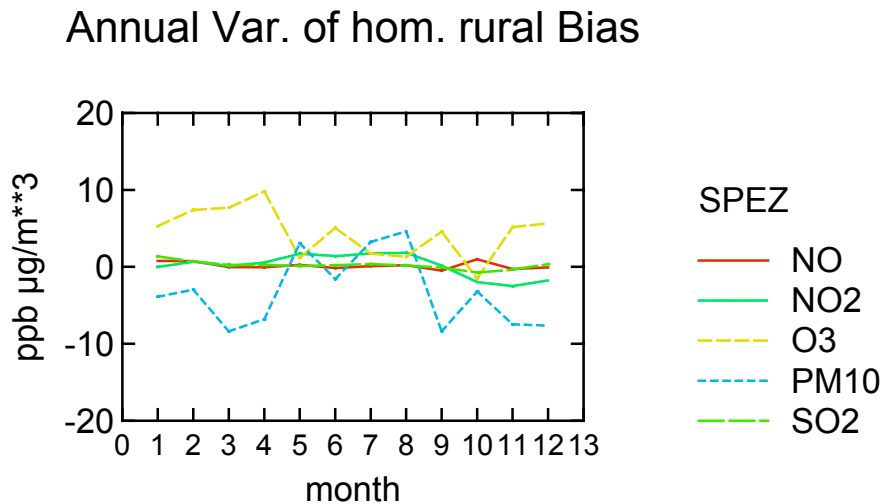


Abbildung 5.22 Jahresgang des homogenen terminbezogenen Bias der Land-Stationen für alle Termine 2001 (O₃, NO und NO₂ in ppb, für SO₂ und PM₁₀ in µg/m³)

5.3.3 Die Varianz des „glatten“ Feldes

Während der Beobachtungsfehler eine singuläre Eigenschaft der Messorte ist und der Bias die Differenz von Erwartungswerten beschreibt, werden nun die Variationen des glatten, d. h. des räumlich korrelierten Immissionsfeldes untersucht (siehe Kapitel 4.4.1.1). Die hier untersuchte Varianz ist die Kovarianz für den Abstand Null, die nicht durch den Beobachtungsfehler beeinflusst ist.

Die folgende Diskussion der Varianz soll zwei Aufgaben erfüllen:

- Vergleich der räumlichen Variabilität von Modellrechnung und Beobachtung
- Abschätzen der Stärke des Einflusses der Beobachtungsinkremente auf die Veränderung des Modellfeldes bei der Analyse

5.3.3.1 Varianz von Modell, Beobachtung und Beobachtungsinkrement

Der Vergleich der Varianz des glatten Feldes von Modellrechnung und Beobachtungen soll überprüfen, ob die Variabilität der beiden Quellen übereinstimmt. Der Vergleich der Varianz und Kovarianz gibt erste Hinweise, ob das Modell in der Lage sein könnte, die Variabilität¹⁵⁵ des gesuchten „glatten“ Immissionsfeldes zu erfassen.

Das generelle Ziel der Modellierung ist es jedoch, nur das „glatte“ Beobachtungsfeld ohne das „Rauschen“ der Beobachtungsfehler zu simulieren. Für den Vergleich der Variabilität ist es demzufolge notwendig, den Varianzanteil des Beobachtungsfehlers zu entfernen und nur den räumlich korrelierten Teil der Beobachtungen zu betrachten.

Die Schätzung der Varianz des glatten Feldes erfolgt mit einem Kovarianzmodell, da Beobachtungsdaten und Beobachtungsinkremente durch den räumlich unkorrelierten Beobachtungsfehler „kontaminiert“ sind.

¹⁵⁵ z. B. kann die abgebildete Variabilität des Modells durch eine Erhöhung der Auflösung verbessert werden.

Würde die Modellrechnung genau dem Erwartungswertfeld der Beobachtungen entsprechen (siehe Kapitel 4.4.1), dann wäre die Varianz des glatten Feldes von Beobachtungsincrementen und Beobachtungsdaten die gleiche. Eine verringerte Varianz der Beobachtungsincremente ist demzufolge ein Hinweis darauf, dass das Modell bereits einen Teil der Variabilität wiedergegeben hat. Dies ist für die meisten der hier behandelten Stoffe der Fall. Folglich ist die Varianz des glatten Feldes der Beobachtungsincremente kleiner als die der reinen Beobachtungsdaten.

Die Varianz ist ein Maß für die Variabilität des Feldes jenseits der des Erwartungswertes. Die Stärke der Varianz hängt von der Methode zur Schätzung des Erwartungswertes sowie von der Korrektur des Bias ab. Aus Gründen der Vereinfachung werden hier nur die klimatischen Varianzen an den Beobachtungsorten untersucht. Der Erwartungswert wird dabei durch den zeitlichen Mittelwert geschätzt.

Abbildung 5.23 zeigt die geschätzten Standardabweichungen, d. h. die Wurzel der Varianz, für das glatte Feld der Modellrechnung, der Beobachtung und der Beobachtungsincremente im Tagesgang. Mit Ausnahme der nachmittäglichen Ozonimmission ist die räumliche Variabilität der Modellrechnung deutlich geringer als die des unkontaminierten glatten Immissionsfeldes aus den Beobachtungen. Dieser Unterschied ist bei der NO-Immission am größten. Durch den Einbezug der Modellrechnung wird nachmittags bei Ozon, PM10 und SO₂ die Variabilität verringert, da hier die Beobachtungsincremente eine geringere Varianz als die Beobachtung aufweisen. Dieser Umstand ist Ausdruck einer sinnvollen Prognoseleistung durch das Modell.

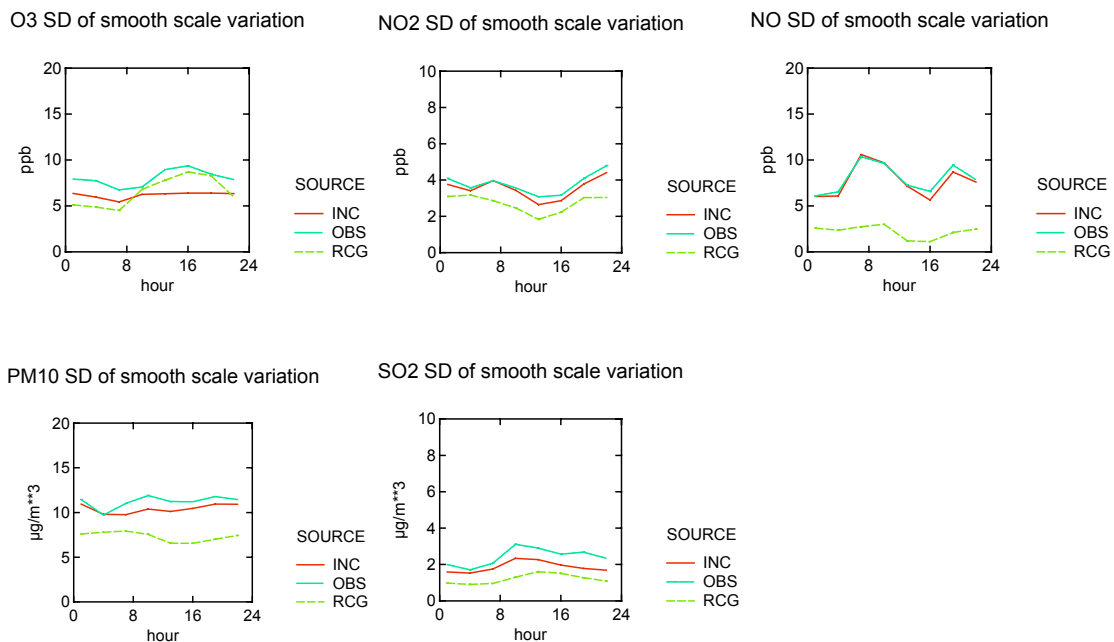


Abbildung 5.23 Tagesgang der klimatischen Varianz des „glatten“ rauschfreien Feldes für die Modelldaten (RCG), die Beobachtungsdaten (OBS) und die Beobachtungsincremente (INC), dargestellt als Standardabweichung. Der dargestellte Wert ist der Median über alle Stationen.

5.3.3.2 Rausch-Signal-Verhältnis

Die Stärke des lokalen Einflusses der Beobachtung auf die Analyse hängt von dem Verhältnis der Varianzen der zugehörigen Fehler ab. Große Beobachtungsfehlervarian-

zen σ_o^2 im Vergleich zu der des Backgrounds bzw. Modells σ_B^2 führen zu einer nur geringen Änderung des Backgrounds bzw. Modellfeldes am Stationsort nach der Gleichung {4.31}. Als Maß dafür dient das sogenannte Rausch-Signal-Verhältnis η , dass als das Verhältnis der Beobachtungsfehlervarianz zur Gesamtvarianz der Beobachtungssinkremente definiert ist (Gandin, 1965):

$$\eta = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_B^2} \quad \{5.26\}$$

Für Geopotentialfelder liegt der Wert des Rausch-Signal-Verhältnisses im Bereich von 10-15% (Daley, 1991). Nur die nachmittäglichen Ozonbeobachtungen liegen im klimatischen Fall in einem ähnlichen Größenbereich. Die Werte η für den Rest der hier betrachteten bodennahen Immissionsfelder sind deutlich größer (über 50%). Damit ist im Allgemeinen nicht zu erwarten, dass die analysierten Felder am Messort sehr genau mit der Beobachtung übereinstimmen. Der hohe Rauschanteil tritt sowohl bei den klimatischen (Abbildung 5.24) als auch mit den terminbezogenen Kovarianzmodellen auf.

In der Praxis der Analyse kann der Bias häufig nicht feldübergreifend bestimmt und korrigiert werden. Dadurch werden erhöhte Varianzen des Backgroundfeldes σ_B^2 geschätzt, woraus sich ein verringertes Rausch-Signal-Verhältnis ergibt (siehe auch Kapitel 5.1.1).

Climatic noise signal ratio

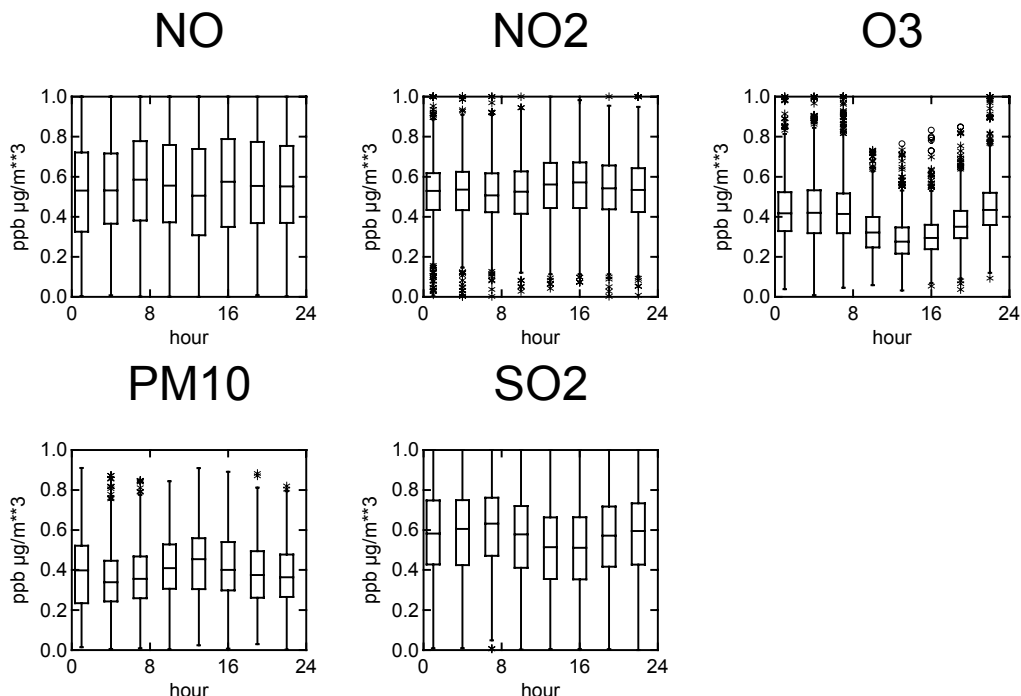


Abbildung 5.24 Verhältnis der Beobachtungsfehlervarianz zur Gesamtvarianz des klimatischen Kovarianzmodells für die Stunden des Tages (hour)

5.3.4 Räumliche Strukturen der Kovarianzmodelle A, B und C

Die Heterogenität der Immissionsfelder ist in dieser Arbeit häufig hervorgehoben worden. Sie betrifft die Erwartungswerte und die davon abweichende räumlich korrelierte stochastische Komponente.

Im Fall der terminbezogenen homogenen Kovarianzmodelle wird die Abweichung von einem konstanten Erwartungswert betrachtet. Im klimatischen Fall wird die Beziehung der Abweichungen vom stationsbezogenen Erwartungswert betrachtet. Die räumliche Heterogenität dieser Erwartungswerte, d. h. der mittleren Verhältnisse der Beobachtungen, wird aus den Betrachtungen zur Regimeklassifikation (siehe Kapitel 3) deutlich.

In diesem Kapitel wird nun versucht, die Inhomogenität und Anisotropie der räumlichen Beziehung der Abweichungen, d. h. der Kovarianzfunktion, darzustellen. Die Darstellung der Inhomogenität der Kovarianzfunktion ist aufgrund der im Vergleich zum Feld verdoppelten Dimension keine leichte Aufgabe. Dies gilt sowohl für die geschätzte empirische Kovarianzmatrix an den Stationsorten als auch für die gesamte Kovarianzfunktion, so wie sie für die Analyse benötigt wird. Eine Möglichkeit wäre die Darstellung des Feldes der Varianzen. Eine Veranschaulichung der räumlichen Beziehungen könnte für einen festgelegten Punkt erfolgen, indem man das zugehörige Feld der Kovarianzen abbildet. Eine indirekte Veranschaulichungen des Feldes der Varianzen findet sich in Kapitel 6.2.1. Eine beispielhafte Darstellung der empirischen Kovarianzen hinsichtlich einer Station ist Abbildung 5.36.

Im Gegensatz zu diesen Darstellungsformen wird in diesem Kapitel das Vermögen der Kovarianzmodelle A, B und C (siehe Kapitel 5.1.3), räumliche Strukturen wiederzugeben, dargelegt. Für das homogene terminbezogene Kovarianzmodell A beschränkt sich die Untersuchung der räumlichen Struktur auf die Untersuchung der geometrischen Anisotropie in Kapitel 5.3.4.1. Das klimatische Kovarianzmodell B beruht auf homogenen Kovarianzmodellen für verschiedene Kombinationen von Regimes (siehe Kapitel 5.3.4.2). Eine allgemeinere Möglichkeit zur Wiedergabe der Inhomogenität liegt in der Darstellung der Eigenvektoren der Kovarianzmatrix in Kapitel 5.3.4.3, was dem Grundgedanken des Kovarianzmodells C entspricht.

5.3.4.1 Homogene Anisotropie (KM A)

In dem homogenen Ansatz ist die Anisotropie feldübergreifend. Sie äußert sich in ellipsenförmigen Strukturen der Kovarianzfunktion. Durch eine lineare Koordinatentransformation¹⁵⁶ kann die Isotropie und damit eine reine Abstandsabhängigkeit der Kovarianzfunktion hergestellt werden (siehe Kapitel 4.4.1.2) Dieses anisotrope Verhalten wird in den meteorologischen Strömungsmustern häufig festgestellt¹⁵⁷ und mit anisotropen Kovarianzmodellen erfasst (Thiebaut, 1976).

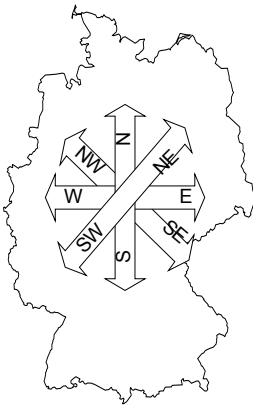
Mögliche Ursachen für eine geometrische Anisotropie der Immissionsfelder sind zum einen in geografischen Faktoren, wie der Lage der Messstationen, der Form des Untersuchungsgebiets und der räumlichen Verteilung der Ballungsgebiete in Deutschland zu suchen. Dem gegenüber stehen meteorologische Einflüsse, wie die mittlere Windrichtung und -stärke oder die Temperaturverteilung. Die mit den terminbezogenen Kovari-

¹⁵⁶In den modernen Formen der Deformationsansätze erfolgt eine nichtlineare Verzerrung des Koordinatensystems (siehe Kapitel 4.3.2.3).

¹⁵⁷Die horizontale Neigung der Achsen der Rossby-Wellen ist z. B. Voraussetzung für den meridionalen Impulstransport.

anzmodellen ermittelbare Zeitabhängigkeit der Prozesse kann den Zusammenhang zwischen der veränderlichen Wettersituation erklären und damit eine Trennung der beiden Ursachengruppen ermöglichen.

Windstärke und -richtung scheinen die aussichtsreichsten Kandidaten für die Erklärung der witterungsbedingten Anisotropie zu sein. Der zunächst naheliegenden Erklärung der Anisotropie als Wirkung des Transportes durch den Wind in eine Richtung steht jedoch der Umstand gegenüber, dass mit Zunahme der Windstärke die Verdünnung vergrößert und somit die „Störung“ abgeschwächt wird.



Die Überprüfung der Anisotropie erfolgte nur für die Beobachtungsdaten und die Modellrechnung. Dafür wurden terminbezogene Kovarianzmodelle hinsichtlich eines homogenen Erwartungswertfeldes für vier Richtungssektoren (Nord/Süd, Ost/West, Nordwest/Südost und Südwest/Nordost) mit einem Winkel von 45° gebildet¹⁵⁸. Der für die sphärische Kovarianzfunktion ermittelte räumliche Skalierungsparameter L (siehe Kapitel 5.2.3) wurde dann für verschiedene Richtungen verglichen. Als Maß für die Anisotropie wurde die relative Differenz des Rangeparameters zu dem um 90° gedrehten Sektor gewählt. Die deutlichere Anisotropie ergab sich für Modellrechnung und

Beobachtung hinsichtlich des Sektorenpaars Nord/Süd und Ost/West¹⁵⁹ am Mittag.

Abbildung 5.25 und Abbildung 5.26 zeigen Histogramme dieser relativen Differenz für die Nord/Süd und Ost/West Sektoren um 13 Uhr für Beobachtung und Modelldaten. Auffällig ist die deutliche Anisotropie der Modelldaten mit längeren Kovarianzen in Ost-West-Richtung für NO , NO_2 und SO_2 . Diese Struktur lässt sich so in den Beobachtungsdaten nicht feststellen. Für NO existiert sogar die stärkere Anisotropie in Nord-Süd-Richtung.

Nach der Betrachtung der mittleren Anisotropie stellt sich die Frage nach einem Zusammenhang zwischen Anisotropie und meteorologischer Situation. Dafür wurden die Wettersituationen mit Hilfe der Großwetterlagen nach Hess und Brezowsky H. (1977) klassifiziert. Die über 30 Wetterlagen sind weiterhin nach Gerstengarbe und Werner (1999) in den drei Grundströmungsmustern „Zonal“, „Meridional“ und „Übergang“ zusammengefasst worden. Mit dieser Einteilung konnte kein Zusammenhang zwischen der Ausprägung der Isotropie und der Wetterlage festgestellt werden. Mögliche Ursachen für die systematische Anisotropie sind demzufolge eher in der Form des Untersuchungsgebietes und in den unterschiedlichen realen und modellierten räumlichen Emissionsverteilungen zu vermuten.

¹⁵⁸ Problematisch erschien dabei der Umstand, dass die für verschiedene Sektoren geschätzte Varianz leichte Unterschiede aufwies. Die Frage, ob es sich dabei um Zufälligkeiten oder zu berücksichtigende Abweichungen handelt, konnte nicht beantwortet werden.

¹⁵⁹ Für die Beobachtungsinckremente tritt die größere Anisotropie im Sektorenpaar Nordwest/Südost und Südwest/Nordost auf, d. h. in einem um 45° gegen das geografische gedrehten Koordinatensystem.

ANISOTROPY W-O/N-S (OBS)

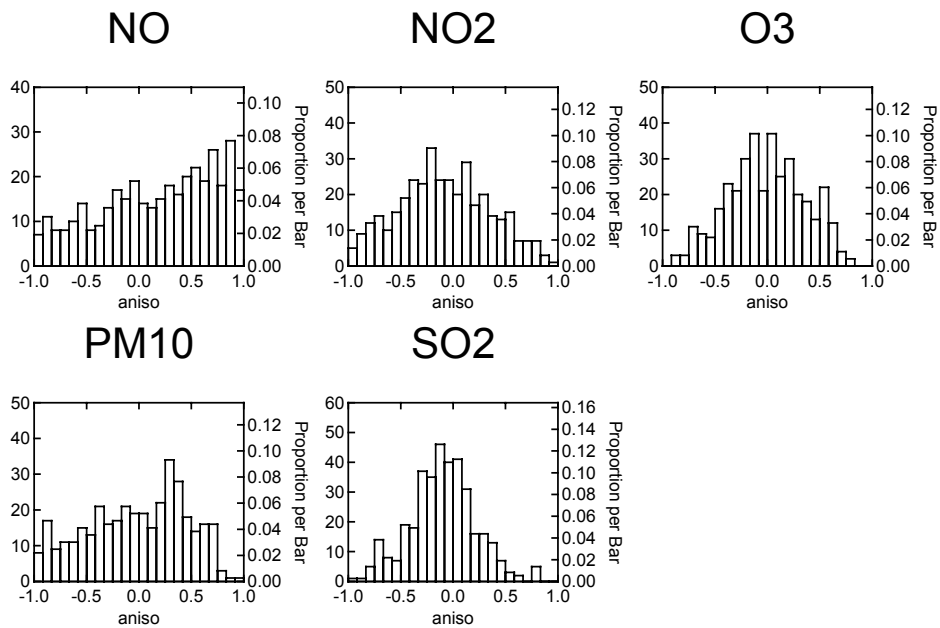


Abbildung 5.25 Relative Differenz (aniso) zwischen dem terminbezogenen homogenen Range-Parameter L in Ost-West- und in Nord-Süd-Richtung für die Beobachtungswerte um 13 Uhr. Positive Differenzen stehen für einen größeren Kovarianz-Range L in Nord-Süd Richtung

ANISOTROPY W-O/N-S (RCG)

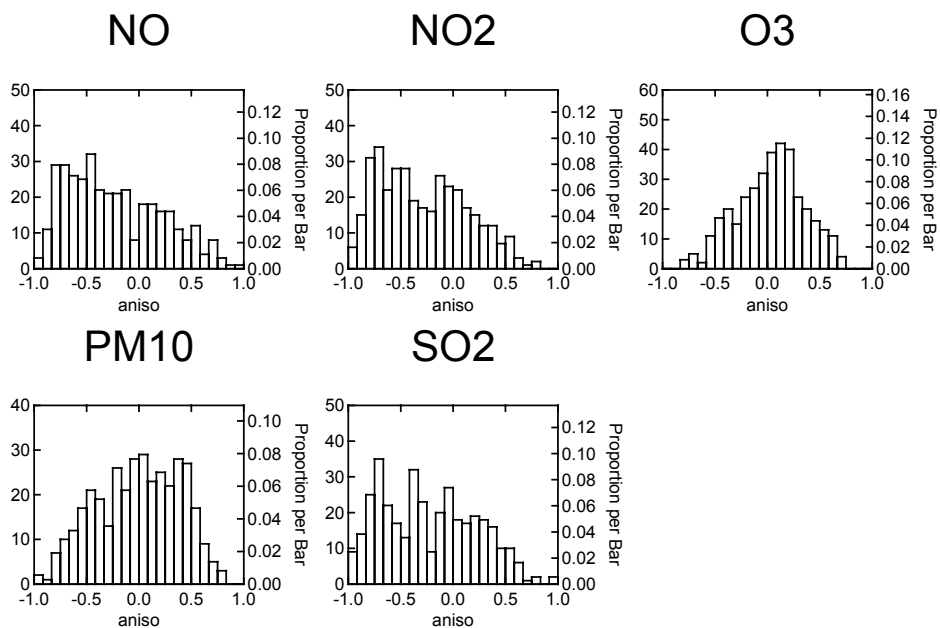


Abbildung 5.26 Relative Differenz(aniso) zwischen dem terminbezogenen homogenen Range-Parameter L in Ost-West- und in Nord-Süd-Richtung für die Modellwerte um 13 Uhr. Positive Differenzen stehen für einen größeren Kovarianz-Range L in Nord-Süd Richtung

5.3.4.2 Kovarianz für verschiedene Immissionsregimes (KM B)

Die unabhängigen Variablen für das Kovarianzmodell B sind neben dem Abstand die Regimes der beiden Orte der Kovarianz (siehe Kapitel 5.1.3). Die Regimes sind die der Stationen für \mathbf{HBH}^T bzw. die des Analysepunkts und der Station für \mathbf{HB} . Das Regime des Analysepunkts wird aus der Modellrechnung abgeleitet. Folgerichtig wird die durch das Kovarianzmodell B wiedergegebene Struktur durch die Verteilung der beobachteten Regimes und die Modellrechnung bestimmt. Die Regimeverteilung der Stationen wurde bereits in Kapitel 3 beschrieben. Da die abgeleiteten Regimes der Modellrechnung stark von der mittleren Belastung abhängen, können die in Kapitel 6.3 vorgeführten Felder der modellierten Jahresmittelwerte eine Vorstellung von der räumlichen Verteilung der modellierten Regimes geben.

Im Folgenden werden nun die Kovarianzen für verschiedene Regimes am Beispiel für die Beobachtungsincremente, Beobachtung und Modellrechnung diskutiert. Dazu wird die entsprechende abstandabhängige empirische Kovarianzfunktion innerhalb des ländlichen und des urbanen Regimes für Ozon (16 Uhr) und NO_2 (7 Uhr) präsentiert.

Auffälligstes Merkmal der Kurven ist die Abhängigkeit von der Varianz und damit gewissermaßen die Skalierung der Kurven. Für das nachmittägliche Ozon ergeben sich keine großen Unterschiede in der Kovarianzstruktur innerhalb des ländlichen und des urbanen Regimes (Abbildung 5.27). In beiden Fällen ist die Kovarianzfunktion für Modellrechnung und Beobachtung von ähnlichem Verlauf. Bei den Beobachtungsincrementen ist die Varianz deutlich kleiner (siehe Kapitel 5.3.3). Die Stärke des relativen Abfalls ist geringer, was auf einen im Vergleich zu Messung und Beobachtung größeren räumlichen Skalierungsparameter L (Range) schließen lässt.

Bei NO_2 um 7 Uhr sind die Varianzen der Abweichung vom Erwartungswert erwartungsgemäß in urbanen Regimes deutlich höher als in den ländlichen (Abbildung 5.28). Im urbanen Regime ist die Form des Abfalls für Modelldaten, Beobachtungsincremente und Beobachtung etwa gleichartig; die Varianz der Beobachtung und der Inkremente ist jedoch deutlich höher als die der Modellrechnung. Die Existenz einer beachtlichen Kovarianz innerhalb des urbanen Regimes für große Abstände erklärt sich durch den Umstand, dass durch die Regimeunterscheidung bei der Approximation gewissermaßen von einem deutschlandweiten Gebiet mit urbaner Charakteristik ausgegangen wird.

Innerhalb der ländlichen Regimes ist die Kovarianz deutlich geringer als in den urbanen; der Skalierungsparameter ist aufgrund des geringen Abfalls jedoch höher.

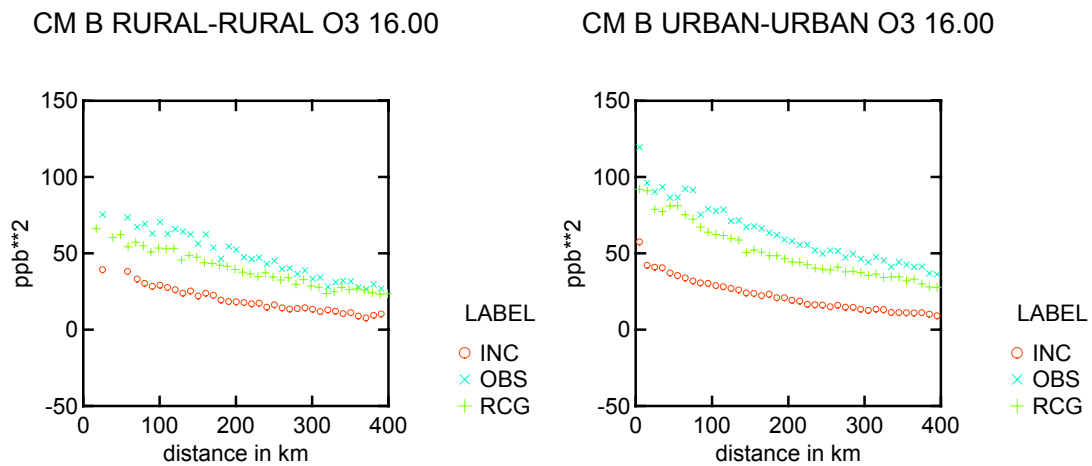


Abbildung 5.27 Klimatische Kovarianz (Kovarianzmodell B) für Ozon um 16 Uhr zwischen den ländlichen (links) und den urbanen Regimes (rechts) der Beobachtung (OBS), der Modellrechnung (RCG) und der Beobachtungsincremente (INC). Dargestellt ist der Median aller empirischen Kovarianzwerte innerhalb einer Abstandsklasse.

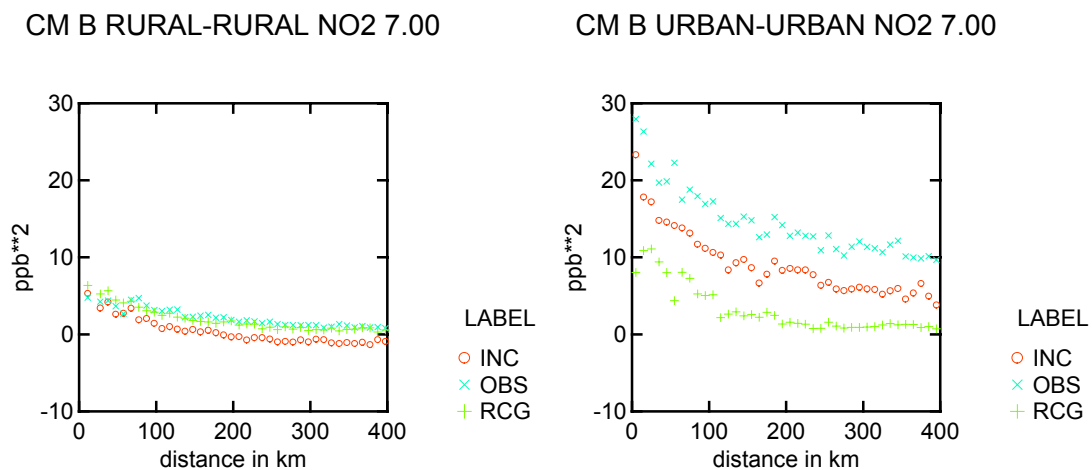


Abbildung 5.28 Klimatische Kovarianz (Kovarianzmodell B) für NO₂ um 7 Uhr zwischen den ländlichen Regimes der Beobachtung (OBS), der Modellrechnung (RCG) und der Beobachtungsincremente (INC). Dargestellt ist der Median aller empirischen Kovarianzwerte innerhalb einer Abstandsklasse.

5.3.4.3 Die Eigenvektoren der Kovarianzmatrix (KM C)

Die Eigenvektoren, multipliziert mit der Wurzel des zugehörigen Eigenwertes, können als die die Kovarianzmatrix aufbauenden Felder aufgefasst werden (siehe Kapitel 4.5.8 und 5.2.5.2). Diese Felder besitzen eine „voneinander unabhängige“¹⁶⁰ räumliche Variation. Das Quadrat der Elemente an einem Ort ergibt den Anteil an der Varianz; das Produkt zwischen zwei Orten ergibt deren Kovarianzanteil. Die separat geschätzte

¹⁶⁰ aufgrund ihrer Orthogonalität

Beobachtungsfehlervarianz wird vor der Eigenvektoraufspaltung von der empirischen Kovarianzmatrix abgezogen, wenn es die der Beobachtungen oder Beobachtungsinkremente ist.

Die prozessorientierte Interpretation der Eigenvektoren (EOF) ist ein viel und kontrovers diskutiertes Hilfsmittel in der globalen Klimaforschung (Dommengeset und Latif, 2002). Die Eigenvektoren werden dabei als wichtige Moden der Variabilität interpretiert. Die Interpretation wird jedoch durch den formalen mathematischen Charakter der Eigenvektoren eingeschränkt. Dies gilt vorrangig für die Eigenvektoren ab dem zweitgrößten Eigenwert, da diese orthogonal zu allen vorangegangenen sein müssen. Die Muster der Eigenvektoren zeigen eine Abhängigkeit von der Form des betrachteten Gebietes. Weiterhin ergeben sich statistisch motivierte Entartungen, wenn die zugehörigen Eigenwerte von ähnlicher Größe sind (Richman, 1986). Ein Lösungsansatz dafür ist die Drehung der Eigenvektoren, der hier aber nicht verfolgt wird.

Im Folgenden werden für Ozon und NO₂ beispielhaft Eigenvektoren der empirischen klimatischen Kovarianzmatrix aus den Beobachtungsdaten, der Modellrechnung¹⁶¹ und den Beobachtungsincrementen miteinander verglichen (siehe Abbildung 5.29 bis Abbildung 5.34). Die an den Stationsorten vorliegenden Eigenvektoren wurden für eine bessere Darstellung analog zum Kovarianzmodell C räumlich interpoliert. Bei der Interpretation ist aus den dargelegten Gründen Vorsicht geboten. Der Vergleich scheint jedoch möglich, da die Artefakte für alle drei Datenbasen gleichwertige Auswirkungen haben sollten¹⁶².

Für Ozon um 13 Uhr zeigen Modellrechnung und Beobachtungsdaten eine sehr ähnliche Struktur im ersten Eigenvektor (siehe Abbildung 5.29 und Abbildung 5.30). Das Zentrum des „Monopols“ ist jedoch in der Modellrechnung nach Süden verschoben. Der zugehörige Eigenwert erklärt bei der Beobachtung 44% und für die Modellrechnung 51% der Gesamtvarianz. Die nächsten beiden Eigenvektoren zeigen übereinstimmend¹⁶³ einen Dipol mit Nord-Süd und NW-SO Orientierung. Ihr Erklärungsanteil liegt bei 13 % (Modell) und 8 % (Beobachtung) bzw. bei 7 und 6 %. Der erste Eigenvektor der Beobachtungsinkremente (Abbildung 5.31) hat einen wesentlich geringeren Erklärungsanteil von nur 13 % und zeigt einen SW-NE gelagerten Dipol. Der zweite Eigenvektor der Beobachtungsinkremente entspricht in seiner Struktur dem dritten Eigenvektor von Modellrechnung und Beobachtung.

Die Struktur der Eigenvektoren der NO₂-Immission um 7 Uhr ist wesentlich heterogener als bei Ozon um 13 Uhr (Abbildung 5.32 und Abbildung 5.33). Der erste Eigenvektor trägt nur 22 % (Beobachtung) bzw. 29% (Modellrechnung) der Gesamtvarianz. Der erste Eigenvektor der Beobachtungsinkremente gleicht in seiner Struktur stark dem der Beobachtung (Abbildung 5.34). In den Beobachtungen sind die Räume Berlin und München weitere wichtige Moden der Variabilität, die in den Modelldaten so nicht zu finden sind. Hier herrscht die größte Variabilität in der Nähe der Ruhrgebietes.

¹⁶¹ Die Kovarianzmatrix der Modelldaten wird nur für die Messorte gebildet, so dass die Größe und Bedeutung der Kovarianzmatrix in allen Fällen die gleiche ist.

¹⁶² Die Untersuchung der Frage, ob beim RRSQRT-Ansatz für Kalman-Filter degenerierte oder „unphysikalische“ Eigenvektoren entstehen und welche Auswirkungen sie haben, scheint lohnenswert.

¹⁶³ Die Struktur der Eigenvektoren ist vom Vorzeichen unabhängig.

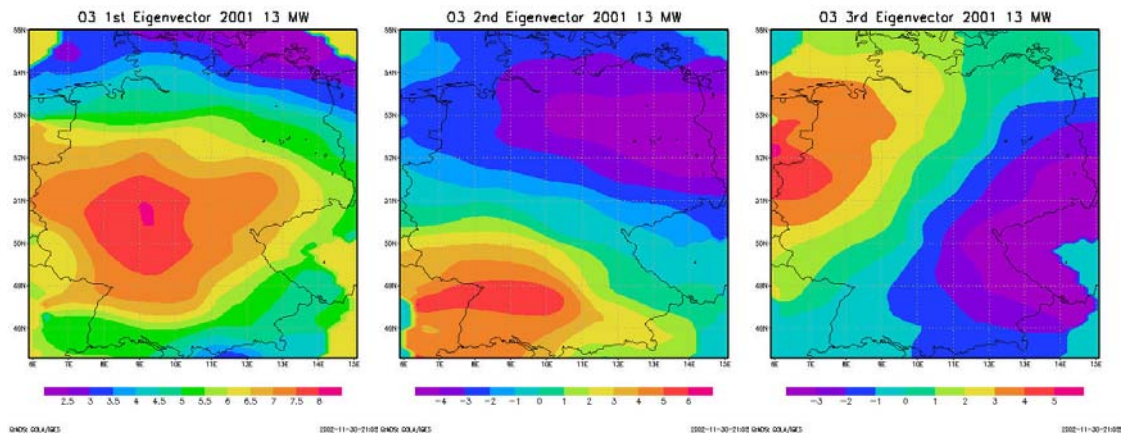


Abbildung 5.29 Die ersten drei Eigenvektoren (interpoliert) der empirischen Kovarianzmatrix der Beobachtungen für Ozon um 13 Uhr.

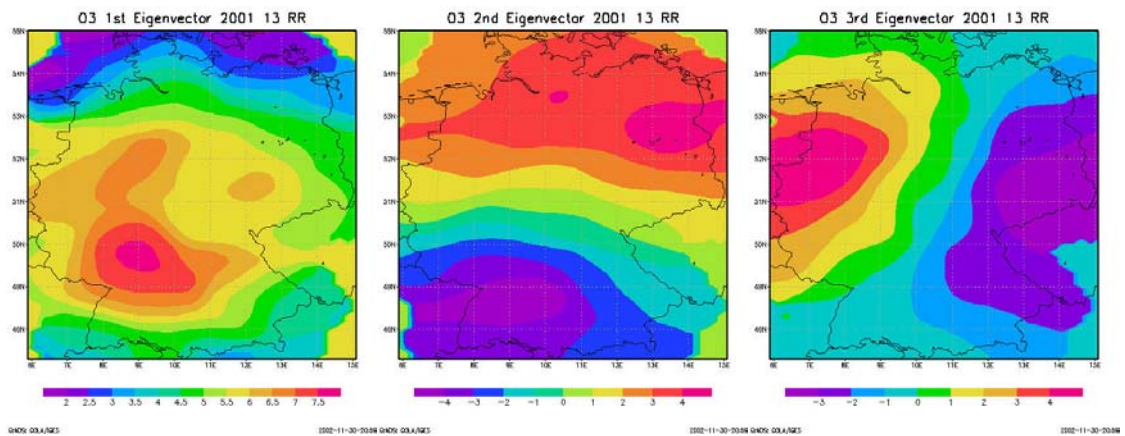


Abbildung 5.30 Die ersten drei Eigenvektoren (interpoliert) der empirischen Kovarianzmatrix der Modellrechnung für Ozon um 13 Uhr.

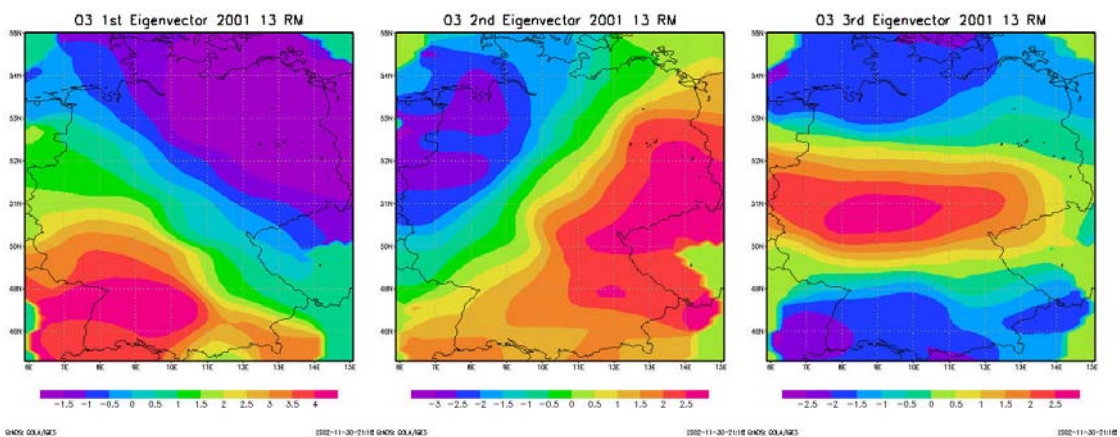


Abbildung 5.31 Die ersten drei Eigenvektoren (interpoliert) der empirischen Kovarianzmatrix der Beobachtungsinkremente (Beobachtung minus Modell) für Ozon um 13 Uhr.

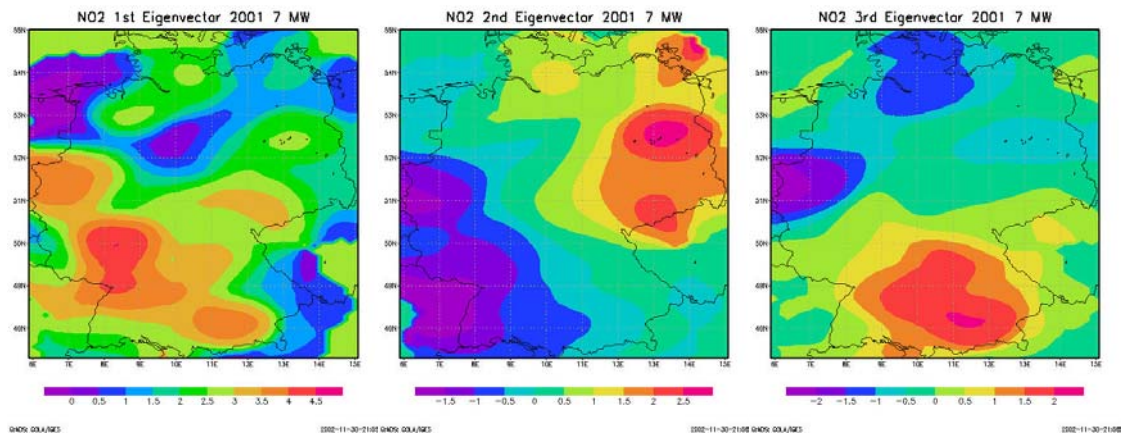


Abbildung 5.32 Die ersten drei Eigenvektoren (formal interpoliert) der empirischen Kovarianzmatrix der Beobachtungen für NO₂ 7.00 Uhr.

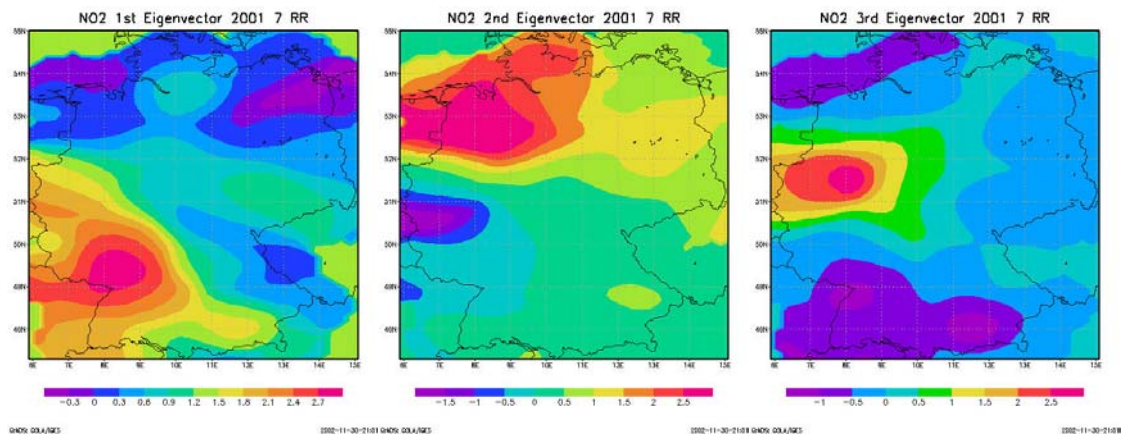


Abbildung 5.33 Die ersten drei Eigenvektoren (formal interpoliert) der empirischen Kovarianzmatrix der Modellrechnung für NO₂ 7 Uhr.

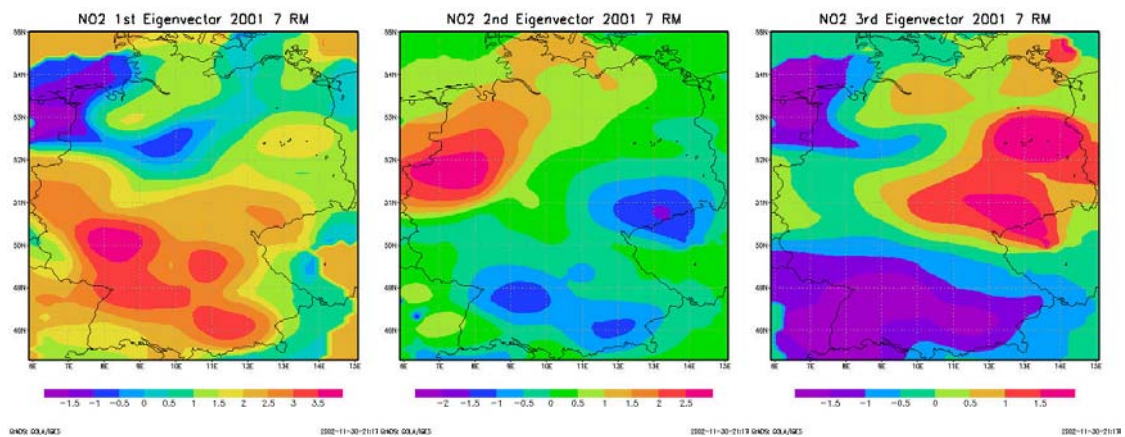


Abbildung 5.34 Die ersten drei Eigenvektoren (formal interpoliert) der empirischen Kovarianzmatrix der Beobachtungszinkremente für NO₂ 7 Uhr.

5.3.5 Alternative dynamische Kovarianzmodelle

Im Rahmen der Arbeit werden vorrangig passive Ansätze der Datenassimilation angewendet. Die geschätzten Kovarianzmodelle des Backgrounds sind demzufolge unabhängig vom vorhergehenden Verlauf der Datenassimilation. Zusätzlich wurden jedoch zwei aktive Verfahren getestet:

1. aktive OI während des Modelllaufs für das gesamte Jahr 2001 (siehe 6.1.7)
2. der Kalman-Filter Lauf für den Juli 2001.

Bei diesen Verfahren erfolgt die Kovarianzmodellierung während des Modelllaufs, denn beim aktiven Ansatz beeinflusst das Analyseergebnis zu einem Termin das Backgroundfeld für den nächsten Termin. Aus diesem Grund sind klimatische Ansätze weniger geeignet.

1. Bei der aktiven OI geschieht die Kovarianzmodellierung empirisch mit dem terminbezogenen homogenen Kovarianzmodell A der Beobachtungsinkremente.

Vergleicht man das Rausch-Signal-Verhältnis der aktiven und passiven Variante, so stellt man ein um ca. 20% höheres Verhältnis bei gleichwertiger Beobachtungsfehlervarianz fest. Dies spricht für eine geringere Ausprägung der Varianz des Background- bzw. Modellfehlers und damit für den positiven Effekt der Analyse im vorangehenden Zeitschritt. Das Endergebnis der analysierten Felder scheint jedoch nicht wesentlich verbessert, wie die Diskussion in Kapitel 6.4 zeigt.

2. Beim Kalman-Filter ist die Prognose der Kovarianzmatrix die grundlegende Eigenschaft des Verfahrens (siehe Kapitel 4.5.7). Hier wird ein zeitlich veränderliches und räumlich inhomogenes Kovarianzmodell prognostiziert. Das Kovarianzmodell des Kalman-Filters liefert nicht nur die Kovarianz zwischen verschiedenen Punkten eines Feldes, sondern zwischen allen Elementen des Modellzustandsvektors. Der RRSQRT-Ansatz des Kalman-Filters beruht auf der Eigenvektoraufspaltung der dynamisch modellierten Kovarianzmatrix \mathbf{P} (siehe Kapitel 4.5.8).

Eine Evaluierung des Kalman-Filters beschränkt sich meist auf den Nachweis seiner Funktionalität, d. h. der verbesserten Wiedergabe von Beobachtungen, die nicht an der Assimilation beteiligt waren (van Loon et al., 1999). Die Frage nach einer Verbesserung der Assimilationsleistung kann durch dieses Vorgehen nicht beantwortet werden. Der Vergleich mit den in Kapitel 5.3.4.3 diskutierten Eigenvektoren und Eigenwerten der empirischen klimatischen Kovarianzmatrix kann zur Prozessevaluierung des Kalman-Filters herangezogen werden. Dieser Vergleich ist angemessen, da die Stärke der Beeinflussung des Modellzustandsvektors durch die Beobachtungen mit dem Rausch-Signal-Verhältnis abgeschätzt werden kann. Die Beobachtungsfehlervarianz in der passiven Datenassimilation (Analyse) und der Kalman-Filter-Anwendung ist dieselbe, so dass auch die Varianzen des Backgrounds bzw. des Modellzustandsvektors vergleichbar sind.

Der hier verwendete Kalman-Filter-Lauf ist eine erste Testversion. Der durchgeführte Vergleich hat demzufolge nur methodischen Charakter. Abbildung 5.35 zeigt beispielhaft die mit dem Kovarianzmodell C an den Stationsorten geschätzte empirische Standardabweichung für 13 Uhr im Sommerhalbjahr 2001 und den Mittelwert der Standardabweichung zur selben Zeit aus dem Kalman-Filter Lauf für den Juli 2001. Die Standardabweichungen des Kalman-Filter-Laufes sind um den Faktor 10 geringer.

Demzufolge ist auch die Möglichkeit der Veränderung des Feldes durch Beobachtungen deutlich herabgesetzt¹⁶⁴.

Ein weiterer Unterschied ist die unterschiedliche Struktur der räumlichen Kovarianz. Abbildung 5.36 zeigt die Kovarianz zur Station Burg (BB001). Während bei den empirischen geschätzten Werten erwartungsgemäß eine Abnahme der Kovarianz mit wachsender Entfernung auftritt, ist dies beim Kalman-Filter nicht zu erkennen. Die Stärke der Kovarianz hängt offensichtlich nicht von der Entfernung ab.

Dieser Umstand lässt sich mit der hier angewandten dynamischen Methode zur Modellierung der Kovarianz durch die Modellprognosefehler (model noise, siehe 4.5.7) erklären. Der Modellfehler wird induziert, indem für das gesamte Modellgebiet eine konstante Änderung von NO_x - und VOC-Emission sowie vom vertikalen turbulenten Austauschkoefizienten K_z angesetzt wird. Folgerichtig sind diese Modellprognosefehler über das gesamte Gebiet räumlich hoch korreliert. Folglich empfiehlt es sich, die Variation der Modellparameter gebietsabhängig zu machen. Aufgrund des erhöhten numerischen Aufwandes ist dies nur begrenzt möglich. Ansätze zur lokalen Begrenzung der Kovarianz sind in Builtjes et al. (2000) zu finden. Die in der vorliegenden Arbeit empirisch modellierten Kovarianzen können diese Ansätze verbessern, da sie eine Spezifizierung von Einflussgebieten ermöglichen.

Weitere zukünftige Untersuchungen könnten sich dem Einfluss der formalen Eigenvektorzersetzung und Reduktion (RRSQRT, siehe 4.5.8) auf die physikalisch-chemische Konsistenz der gebildeten Modellzustandsvektoren widmen. Dabei spielt die Frage der Entartung bzw. Rotation der Eigenvektoren eine Rolle. Die Ergebnisse des Kapitels 5.3.4.3 bilden hierfür eine Arbeitsgrundlage.

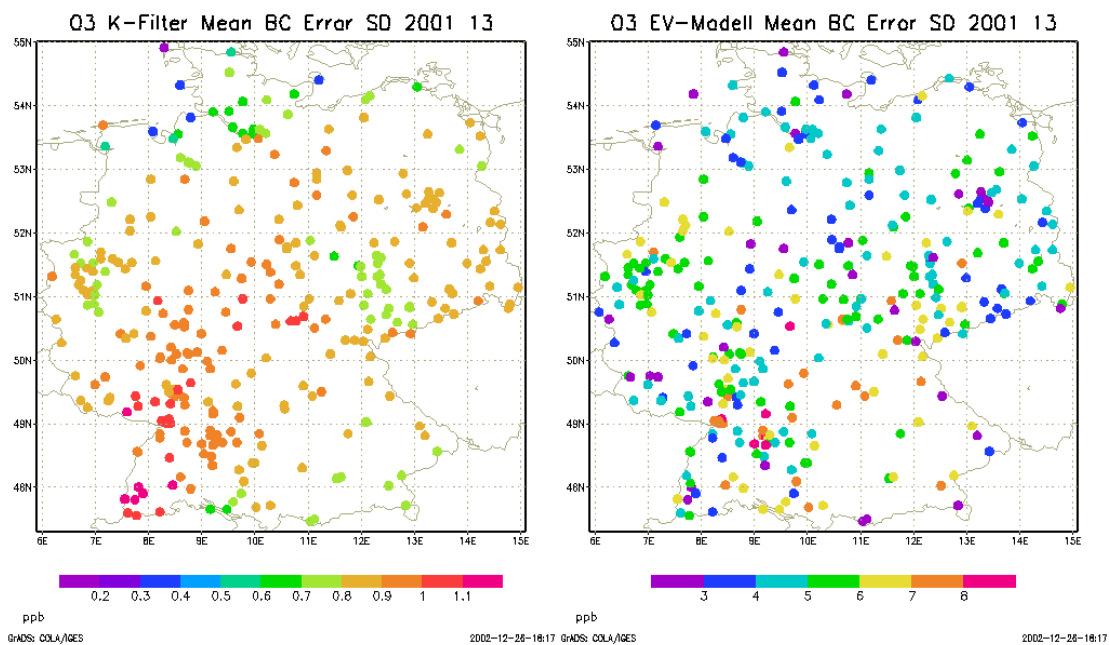


Abbildung 5.35 Standardabweichung (in ppb) des Fehlers des Backgrounds für Ozon um 13 Uhr aus dem Kalman-Filter-Lauf (links, Mittelwert im Juli 2001) und dem empirischen klimatischen Kovarianzmodell C (Sommer 2001). Die Skalierung ist für die Karten unterschiedlich.

¹⁶⁴ Dieser Umstand wird noch verstärkt, da in der gegebenen Implementierung die Analyse mit dem Feld der Stundenmittelwerte und nicht mit dem prognostizierten Modellfeld erfolgt. Letzteres wird „nur“ über die Kovarianzen zum Feld der Stundenmittelwerte mit der Analyse verändert.

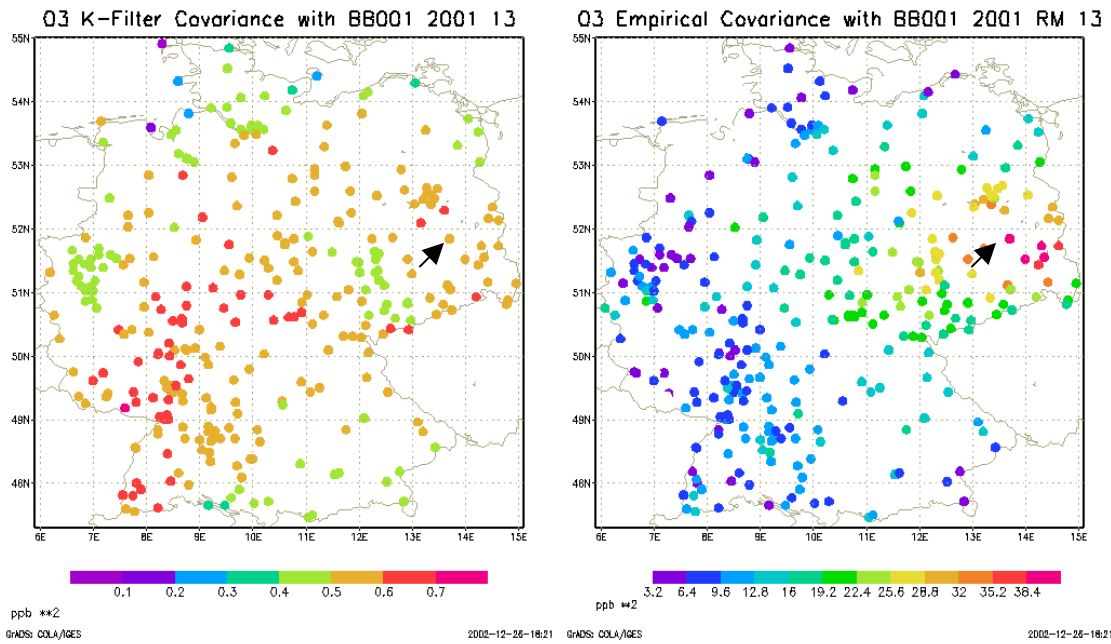


Abbildung 5.36 Kovarianz (in ppb^2) zur Station Burg (BB001) (Pfeil) des Fehlers des Backgrounds für Ozon um 13 Uhr aus der Kalman-Filter-Lauf (links, Mittelwert im Juli 2001) und dem empirischen klimatischen Kovarianzmodell C (Sommer 2001). Die Skalierung ist für die Karten unterschiedlich.