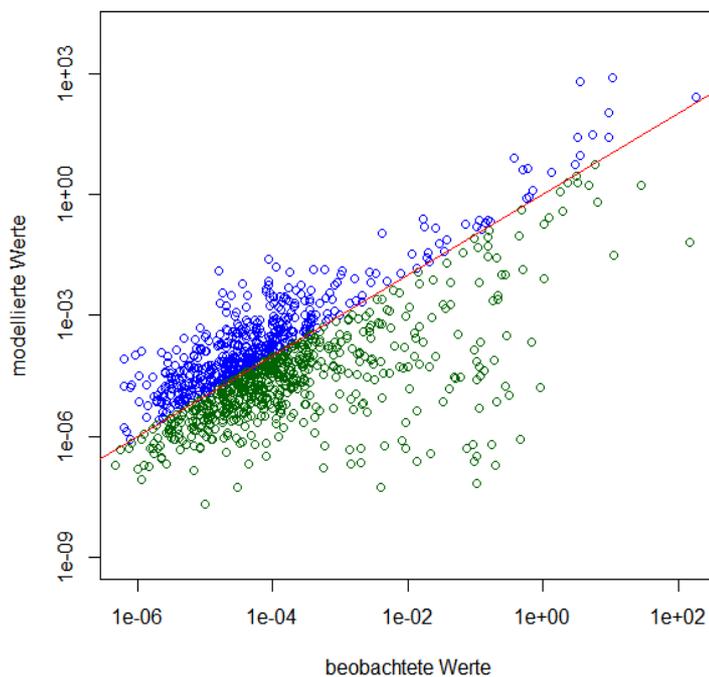


Validierung von atmosphärischen Dispersionsmodellen

Beurteilung der Qualität von meteorologischen Schadstoffausbreitungsmodellen durch statistische Parameter am Beispiel der Anwendung des Lagrange'schen Teilchensimulationsmodells FLEXPART auf die ^{137}Cs -Emissionen bei den Katastrophenereignissen in Fukushima Dai-ichi

Matthias Schlögl



Bachelorarbeit

Institut für Meteorologie (BOKU-Met)
Department Wasser – Atmosphäre – Umwelt
Universität für Bodenkultur

Jänner 2013

ISSN 1994-4179 (Print)
ISSN 1994-4187 (Online)

BOKU-Met Report 22

Validierung von atmosphärischen Dispersionsmodellen

Beurteilung der Qualität von meteorologischen
Schadstoffausbreitungsmodellen durch statistische
Parameter am Beispiel der Anwendung des
Lagrange'schen Teilchensimulationsmodells
FLEXPART auf die ¹³⁷Cs-Emissionen bei den
Katastrophenereignissen in Fukushima Dai-ichi

Matthias Schlögl

Bachelorarbeit
Bachelorstudium Umwelt- und Bioressourcenmanagement

Institut für Meteorologie (BOKU-Met)
Department Wasser – Atmosphäre – Umwelt
Universität für Bodenkultur, Wien

BOKU-Met Report 22

Jänner 2013

Diese Publikation sollte folgendermaßen zitiert werden:

Schlögl, M. (2013): Validierung von atmosphärischen Dispersionsmodellen. Beurteilung der Qualität von meteorologischen Schadstoffausbreitungsmodellen durch statistische Parameter am Beispiel der Anwendung des Lagrange'schen Teilchensimulationsmodells FLEXPART auf die ¹³⁷Cs-Emissionen bei den Katastrophenereignissen in Fukushima Dai-ichi. Bachelorarbeit für das Studium Umwelt- und Bioressourcenmanagement. *BOKU-Met Report 22*, ISSN 1994-4179 (Print), ISSN 1994-4187 (Online). Available at http://www.boku.ac.at/met/report/BOKU-Met_Report_22_online.pdf.

Impressum:

Medieninhaber und Herausgeber:

Universität für Bodenkultur, Department für Wasser – Atmosphäre – Umwelt
Institut für Meteorologie, Peter Jordan-Straße 82, 1190 Wien, Österreich

<http://www.wau.boku.ac.at/met.html>

Inhaltsverzeichnis

Abstract	3
1 Einleitung	4
2 Material und Methoden	6
3 Methodik der Modellierung	7
3.1 Inverse Modellierung.....	7
3.2 Datenassimilation.....	7
3.3 Ausbreitungsmodelle	8
3.3.1 Das Lagrange-Partikelmodell	8
3.4 Modellierung der Ausbreitung von ¹³⁷ Cs aus dem KKW Fukushima.....	9
3.4.1 Methodik	9
3.4.2 Sensitivität der Messstationen	10
3.4.3 Ergebnisse	10
4 Quantitative Evaluierung und Validierung von Modellen	11
4.1 Fehler.....	11
4.1.1 Fehler im Modell.....	11
4.1.2 Fehler in der Beobachtung	11
4.1.3 Umgang mit Fehlern	12
4.1.4 Der Zusammenhang zwischen Modell und Realität	12
4.2 Statistische Parameter und Fehlermaße.....	12
4.2.1 Mittelwerte.....	13
4.2.2 Standardabweichung.....	13
4.2.3 Bias	14
4.2.4 Mittlere quadratische Abweichung	15
4.2.5 Mean Absolute Error	16
4.2.6 Pearsons Korrelationskoeffizient	17
4.2.7 Rangkorrelationskoeffizienten	17
4.2.8 Bestimmtheitsmaß.....	18
4.2.9 Figure of Merit	19
4.2.10 Peak Estimation Accuracy.....	19
4.2.11 Scatterplot, FOEX und FAα	19
4.2.12 Tests	20

5	Analyse der Datenpaare.....	21
5.1	Ausgewählte Fehlermaße.....	21
5.1.1	Erste Analyse der Originaldaten	21
5.1.2	Aufbereitung der Daten	21
5.2	Graphische Darstellung	22
5.2.1	Scatterplot	22
5.2.2	Zeitreihen	24
5.2.3	Boxplots	25
5.2.4	Residuenplot	25
5.2.5	Statistische Tests	26
5.3	Vertiefende Analyse	26
6	Diskussion	28
7	Danksagung.....	30
8	Verzeichnisse	31
8.1	Literaturverzeichnis	31
8.2	Abbildungsverzeichnis.....	35
8.3	Tabellenverzeichnis.....	35
8.4	Abkürzungsverzeichnis.....	36
8.5	Elementsymbole	37
Anhang	38
Anhang I:	Das Dispersionsmodell FLEXPART	38
Anhang II:	Figure of Merit	39
Anhang III:	Peak Estimation Accuracy.....	40
Anhang IV:	Weitere Skill Scores	41
Anhang V:	Diagnostic Plots	42
Anhang VI:	Taylor-Diagramm	43

Kurzfassung

In dieser Arbeit wird anhand der ^{137}Cs -Emissionen bei der Unfallserie in Fukushima Dai-ichi im März 2011 der Frage nachgegangen, welche statistischen Parameter zur Beurteilung der Qualität, Zuverlässigkeit und Genauigkeit der inversen Modellierung von Quelltermen sowie von Ausbreitungsmodellen herangezogen werden können. Dazu werden nach einer kurzen Analyse der generellen Methodik der Modellierung von atmosphärischen Phänomenen verschiedene Fehlermaße auf ihre Tauglichkeit und Aussagekraft untersucht. Die gängigen Fehlermaße, welche bei der Validierung von derartigen Modellen Verwendung finden, werden vorgestellt und hinsichtlich ihrer Verwendbarkeit für die Evaluierung atmosphärischer Transportmodelle analysiert. Dabei zeigt sich, dass besonders der Einfluss der Modellperformance auf die Ausprägung der (normalisierten) Parameter zu bedenken ist. In weiterer Folge werden reale Datenpaare (bestehend aus einem Schätzer und einem Messwert) mithilfe dieser Fehlermaße analysiert. Bei einer Anwendung auf die FLEXPART-Simulation von Stohl et al. (2012) zeigt sich, dass einige Ausreißer zu relativ schlechten Gütewerten führen. Bei Verwendung robuster Methoden ergeben sich jedoch befriedigende Werte. So steigt etwa bei der Betrachtung logarithmierter Werte der Korrelationskoeffizient von 0,24 auf 0,76. Eine Eliminierung der Extremwerte zeigt bereits bei einem um 1% gestutzten Datensatz eine Erhöhung der erklärten Varianz von 6% auf 97%.

Abstract

Based on the emissions of ^{137}Cs released during the Fukushima Dai-ichi nuclear disaster in March 2011 into the atmosphere, this paper analyses different statistical parameters for model evaluation and validation; in particular, inverse modeling of source terms and dispersion modeling are investigated. Following an introductory analysis of available methods for modeling atmospheric phenomena, various statistical parameters are scrutinized regarding their suitability and significance for air pollution modeling. As it turns out, the influence of the model performance on the (normalized) parameters has to be taken into account. Subsequently, common parameters used for model validation are evaluated critically. In further consequence, selected data pairs (consisting of an estimator and a measured value) are analyzed with using selected statistical parameters. Applied to the FLEXPART simulation of Stohl et al. (2012), it turns out that a few outliers lead to relatively low performance parameters. By using more robust methods satisfactory values are obtained. For example, the correlation coefficient increases from 0.24 to 0.76 if raw data are replaced by their logarithms, and trimming 1% of the data as outliers increases the explained variance from 6% to 97%.

1 Einleitung

Modelle spielen in allen Umweltwissenschaften eine große Rolle. Besonders im Bereich der Beurteilung der Belastung von Schutzgütern mit Schadstoffen kommt ihnen dabei eine weitreichende Bedeutung zu. Da Schadstoffbelastungen eine Funktion der Zeit sind und zudem von vielen weiteren Faktoren (wie etwa Seehöhe, geographische Breite, Exposition, etc.) beeinflusst werden, ist eine rein empirische Analyse dieser komplexen Schadstoffbelastungs-Situation wenig praktikabel. Stattdessen wurden und werden verschiedenste Arten von Modellen entwickelt, welche dazu eingesetzt werden können, die tatsächliche Situation anhand weniger Messdaten möglichst realitätsgetreu zu simulieren.

Um eine profunde Aussage über die Qualität dieser Modelle zu treffen, d.h. um beurteilen zu können, mit welcher Genauigkeit und Zuverlässigkeit die Wirklichkeit durch diese Simulationen repräsentiert wird, werden verschiedene Verfahren eingesetzt. Grundsätzlich ist der Vergleich der Messdaten mit den modellierten Werten, welcher zur Validierung von Modellen herangezogen wird, eine statistische Fragestellung. Eine der prominentesten Methoden mit der die Übereinstimmung von Modell und Wirklichkeit beurteilt werden kann, ist, neben der Zeitreihenanalyse, die Verwendung von Fehlermaßen und statistischen Parametern respektive Kennzahlen.

In der vorliegenden Arbeit soll dies nun mit Hilfe eines konkreten Beispiels dargelegt und erläutert werden. Anhand des Caesium-137-Ausstoßes¹, welcher bei der Unfallserie im Kernkraftwerk Fukushima Dai-ichi im März 2011 stattgefunden hat, soll in dieser Arbeit der Frage nachgegangen werden, welche statistischen Parameter zur Beurteilung der Qualität, Zuverlässigkeit und Genauigkeit der inversen Modellierung² von Quelltermen³ sowie von Ausbreitungsmodellen herangezogen werden können. Dazu werden zuerst verschiedene Fehlermaße auf ihre Tauglichkeit und Aussagekraft hin untersucht und die gängigen Fehlermaße, welche bei der Validierung von derartigen Modellen Verwendung finden, werden vorgestellt und hinsichtlich ihrer Verwendbarkeit für die Evaluierung atmosphärischer Transportmodelle analysiert. In weiterer Folge werden reale Datenpaare (bestehend aus einem Schätzer und einem Messwert) mithilfe dieser Fehlermaße analysiert.

Dieses Beispiel aus der Meteorologie ist vom Autor aufgrund der Tatsache ausgewählt worden, dass Ausbreitung, Transport und Deposition von Schadstoffen in der Atmosphäre, über bzw. durch die Atmosphäre eines der Haupteinsatzgebiete für (inverse) Modellierung darstellt. Dieser Teilbereich der Meteorologie beschäftigt sich dabei unter anderem mit der Rekonstruktion von Quelltermen diverser Schadstoffe durch inverse Modellierung. Der Quell-

¹ Caesium (Elementsymbol Cs, Ordnungszahl 55) ist ein chemisches Element, welches zur Gruppe der Alkalimetalle gehört. Unter Normalbedingungen liegt es im festen Aggregatzustand vor. Im reinen Zustand ist es silberweiß, durch Verunreinigungen erscheint es goldgelb. Caesium hat einen für Metalle sehr niedrigen Schmelzpunkt, welcher bei 301,7 K liegt. Der Siedepunkt liegt bei 944 K. Die Masse von Caesium beträgt 132,91 u, die spezifische Dichte beträgt unter Normalbedingungen 1,873 g/cm³. Mit einer Mohs-Härte von 0,2 ist Caesium ein sehr weiches Metall. (Hoppe A. 2012, s.p.). Das künstliche radioaktive Isotop ¹³⁷Cs, welches in größeren Mengen bei diversen Kernwaffenversuchen sowie bei den Reaktorunfällen von Tschernobyl (u.a. Garger et al., 2012) und Fukushima (Yasunari et al. 2011) freigesetzt worden ist, entsteht als Spaltprodukt bei der Kernspaltung von Uran. ¹³⁷Cs hat eine physikalische Halbwertszeit von 30,2 Jahren, die effektive Halbwertszeit beim Menschen liegt bei 70 Tagen. Aufgrund der guten Löslichkeit des in Form von Salzen vorliegenden ¹³⁷Cs können sich die Ionen gut im Körper verteilen. Insbesondere findet eine Anreicherung im Muskelgewebe, in Leber, Milz und Nieren statt (Statistisches Bundesamt 2012, s.p.; Müller 2012, s.p.).

² Bei der inversen Methode werden Konzentrationsmessungen der zu untersuchenden Substanz sowie Meteorologische Felder als Eingangsdaten für ein invertiertes, mikroskaliges, physikalisches numerisches Dispersionsmodell verwendet, um die gesuchte Emissionsrate zu bestimmen (Stockhouse 2000, s.p.). Eine ausführlichere Beschreibung der Funktionsweise der inversen Modellierung bietet Stockhouse (2000).

³ Als Quellterm wird die die Art und Menge der freigesetzten Radionuklide bezeichnet.

term wird dann wiederum als Ausgangspunkt für das Ausbreitungsmodell herangezogen. Diese Vorgangsweise wird etwa bei der Modellierung der Ausbreitung radioaktiver Stoffe in der Atmosphäre angewendet. Der Ausstoß von radioaktiven Substanzen aus dem japanischen Kernkraftwerk Fukushima Dai-ichi, welcher im März 2011 durch eine Verkettung von Naturkatastrophen und Unfällen, die zu einem totalen Ausfall des Anschlusses an das externe Stromnetz (*station blackout*) geführt haben, verursacht worden ist, ist ein aktuelles Beispiel für die Anwendung derartiger Dispersionsmodelle.

Von allen radioaktiven Substanzen, welche von dem havarierten Atomkraftwerk ausgestoßen worden sind⁴, ist das Isotop ^{137}Cs deshalb für die Analyse mit den Fehlermaßen ausgewählt worden, weil es aufgrund seiner Halbwertszeit von dreißig Jahren auch noch nach langen Transportwegen überall auf der Erde gut nachweisbar ist. Mit einer Menge von über 36 PBq (Stohl et. al 2012, 2327) wurden zudem größere Mengen an ^{137}Cs freigesetzt. Aufgrund der guten Detektierbarkeit, welche auch auf das durch die Relevanz von ^{137}Cs bei Nuklearversuchen bedingte gute Messnetz⁵ der CTBTO⁶ zurückzuführen ist, eignet sich ^{137}Cs also gut für den Einsatz in der atmosphärischen Transportmodellierung. Ein weiterer Aspekt der Freisetzung des künstlichen Isotops ^{137}Cs ist der Jahrzehnte andauernde Verbleib in der Umwelt, weshalb eine Modellierung des aufgrund seiner radiologischen Wirkung für den Menschen und seine Umwelt gefährlichen Isotops als relevant anzusehen ist.

⁴ Bei den in Fukushima Dai-ichi freigesetzten Radionukliden handelt es sich um zwei Stoffgruppen. Den Großteil machen die direkt bei der Kernspaltung entstehenden Spaltprodukte, nämlich ^{131}I , ^{133}Xe , ^{134}Cs (^{133}Cs), ^{136}Cs , ^{137}Cs , $^{110\text{m}}\text{Ag}$, (^{109}Ag), $^{129\text{m}}\text{Te}$, ^{132}Te , ^{132}I , ^{140}Ba , ^{140}La , ^{91}Sr , ^{91}Y , ^{95}Zr , ^{95}Nb und ^{99}Mo aus. Zusätzlich wurde ^{239}Pu und ^{240}Pu durch die zeitweise brennenden Brennstäbe im Abklingbecken emittiert (KIT 2011b, 1). Daneben existieren aber auch ursprünglich nicht radioaktive Substanzen, welche durch Anlagerung von Spalt-Neutronen in Radionuklide umgewandelt werden. Zu diesen Aktivierungsprodukten zählen ^{239}Np und ^{59}Fe (Shozugawa et al. 2012; Tagami et al. 2012; Stohl et al. 2012).

⁵ Zum Zeitpunkt des Reaktorunfalls in Fukushima waren 63 Messstationen des Messnetzes für aerosolgebundene Gammastrahler (International Monitoring System) einsatzbereit (CTBTO 2011, s.p.). Eine Übersicht über die einzelnen Messstationen des International Monitoring System (IMS) bietet die CTBTO (2012b, s.p.).

⁶ Die *Comprehensive Nuclear-Test-Ban Treaty Organization* (Organisation des Vertrags über das umfassende Verbot von Nuklearversuchen) ist eine internationale Organisation, welche die Einhaltung des Kernwaffenteststopp-Vertrag nach seiner Unterzeichnung überwachen soll (CTBTO 2012a, s.p.).

2 Material und Methoden

Den Grundstock der Arbeit bilden eine ausführliche Literaturrecherche sowie eine statistische Auswertung von Datenpaaren, welche sich aus beobachteten Messwerten sowie aus dem *a posteriori*-Modell⁷ abgeleiteten Werten zusammensetzen.

Die Literaturrecherche hat sich primär auf das Internet erstreckt, wo mit Hilfe der Suchmaschinen Scopus, Scirus, ScienceDirect, Springerlink, ingentaconnect sowie der Metasuchmaschinen WorldWideScience und ScientificCommons nach relevanten wissenschaftlichen Publikationen gesucht worden ist. Darüber hinaus sind die Websites verschiedener fach einschlägiger Institutionen (American Institute of Physics, Society of Industrial and Applied Mathematics) nach Informationen durchsucht worden. Des Weiteren hat die Universitätsbibliothek der Universität für Bodenkultur Wien als Quelle für diverse Dissertationen, Fachbücher und weitere Publikationen gedient, welche für die Arbeit verwendet worden sind.

Um das Thema der Modellevaluierung mittels Fehlermaßen zu veranschaulichen, bezieht sich die vorliegende Arbeit dabei auf Methodik, Daten und Modelle, welche Stohl et al. (2012) bei der Untersuchung der Freisetzung von Xenon-133 und Caesium-137 aus dem Kernkraftwerk Fukushima Dai-ichi angewendet haben.

Die Daten für die statistische Analyse der Modellqualität stammen von Messstellen der CTBTO, dem europäischen Netzwerk Ring of Five (Ro5) sowie von diversen weiteren Messstationen⁸ und sind von Frau Petra Seibert – einer Co-Autorin der Studie von Stohl et al. (2012) – bezogen worden. Bei dem Datensatz handelt es sich konkret um 1494 Wertetripel, welche sich aus einem Messwert der atmosphärischen Konzentration⁹ von ¹³⁷Cs in Bq/m³, einem ausgehend vom First Guess¹⁰ (*a priori*) modellierten Wert der atmosphärischen Konzentration von ¹³⁷Cs in Bq/m³ und einem ausgehend vom finalen *a posteriori* Modell modellierten Wert der atmosphärischen Konzentration von ¹³⁷Cs in Bq/m³ zusammensetzen¹¹. Zudem sind auch noch der Zeitpunkt, in dem die Werte gemessen worden sind, sowie die Messunsicherheit bekannt. Die verschiedenen statistischen Analysen des Datensatzes mit Hilfe diverser Fehlermaße sind mit dem Statistikprogramm R¹² unter Zuhilfenahme diverser Spezialpakete durchgeführt worden.

Darüber hinaus sind ExpertInnenmeinungen von SpezialistInnen auf den Gebieten der Meteorologie (Petra Seibert vom Institut für Meteorologie, BOKU) und der Umweltstatistik (Gregor Laaha vom Institut für Angewandte Statistik und EDV, BOKU) eingeholt worden.

⁷ Der Terminus *a posteriori* bezeichnet ein genaues Modell, bei dem Wissen über einen zuvor unbekanntem Umweltzustand durch die empirische Beobachtung eines von diesem Umweltzustand abhängigen Parameters gesammelt werden konnte. Im Gegensatz dazu bezeichnet *a priori* ein näherungsweise Schätzverfahren, welches auf Vorwissen beruht.

⁸ Weitere Ausführungen zu den Messstationen finden sich in Kapitel 3.4.1 bzw. bei Stohl et al. (2012, 2318f).

⁹ Die atmosphärische Konzentration wird in Becquerel pro Kubikmeter [Bq/m³] gemessen. Becquerel ist SI-Einheit der Aktivität eines radioaktiven Stoffes und gibt die mittlere Anzahl der pro Sekunde radioaktiv zerfallenden Atomkerne an.

¹⁰ Der *First Guess* (auch *Erste Näherung*) ist eine Kurzfristvorhersage, die durch Extrapolation eines bekannten Zustandes getroffen wird und den wahrscheinlichsten Zustand zu einem anderen Zeitpunkt abschätzt (Wergen 2002, 144).

¹¹ Diese Arbeit beschränkt sich dabei auf den Vergleich von beobachteten und modellierten Werten; die Schätzwerte aus dem First Guess werden nur in Einzelfällen herangezogen.

¹² Weitere Informationen zu dem Statistikprogramm und der Programmiersprache R sowie zum R Project finden sich unter R 2012.

3 Methodik der Modellierung

Gesetzt den Fall, dass die Informationen (z.B. Menge, Ort, Zeit, Dauer, etc.) über Schadstoffquelle und -emissionen unzureichend oder gänzlich unbekannt sind, läuft die Modellierung der Ausbreitung von Luftbeimengungen in der Atmosphäre prinzipiell in zwei Schritten ab. Zuerst wird ausgehend von Messwerten der Quellterm invers modelliert. Anschließend wird in einem zweiten Schritt der so rekonstruierte Quellterm als Ausgangspunkt für ein Ausbreitungsmodell herangezogen.

Grundsätzlich ist im Bereich der Modellierung von atmosphärischen Verschmutzungen festzuhalten, dass die besondere Herausforderung in diesem komplexen Anwendungsbereich darin besteht, Trade-offs zwischen der Nachfrage nach immer mehr und immer besseren bzw. komplexeren Modellen und dem Rechenaufwand respektive der Datenunsicherheit einzugehen (Sportisse 2002, V).

Ziel dieses Kapitels ist es, einen kurzen Überblick über die inverse Methode sowie über Technik und Funktionsprinzip der Lagrange'schen Partikel-Ausbreitungsmodelle zu geben. Des Weiteren wird das Prinzip der Datenassimilation, welches zur Verbesserung der Modellqualität dient, kurz erklärt.

Diese Einführung soll als theoretischer Background zum Prinzip der atmosphärischen Ausbreitungsmodellierung dienen, welches in weiterer Folge durch Fehlerstatistik analysiert wird.

3.1 Inverse Modellierung

Im Gegensatz zu den klassischen Simulationsproblemen oder Vorwärtsproblemen handelt es sich bei der inversen Modellierung um ein sogenanntes inverses Problem. Während bei der erstgenannten Problemstellung versucht wird, von der Ursache auf eine Wirkung zu schließen, ist beim inversen Problem genau das Gegenteil der Fall. Das heißt, dass das zentrale Merkmal des inversen Problems in der Vorgangsweise begründet liegt, dass versucht wird, ausgehend von Messergebnissen Rückschlüsse auf die Werte jener Parameter zu ziehen, die ein System charakterisieren (Tarantola 2005, XI). Anders ausgedrückt bedeutet dies, dass gegebene (i.e. gemessene) Werte dazu verwendet werden, die Größe von Parametern zu schätzen, die direkt oder indirekt mit der gemessenen Größe zusammenhängen (Palmer 2012, s.p.). Bezogen auf den Anwendungsbereich inverser Modellierung in der Meteorologie impliziert dies, dass ausgehend von den Messungen der Spurenstoffkonzentrationen und meteorologischer Parameter sowie der Bestimmung von Rückwärtstrajektorien¹³ die Quellen der Luftbeimengungen und die Emissionsraten bestimmt bzw. abgeleitet werden (Seibert 2000, 4).

Neben den gegebenen Daten respektive den Messungen ist auch eine bestimmte Invertierungstechnik Grundvoraussetzung für die Anwendung der inversen Methode (Stockhause 2000, 5). Es gibt dabei mehrere Invertierungsmethoden, die auf verschiedene Modelltypen anwendbar sind. Dazu gehören einerseits direkte Verfahren, die nur auf spezielle Modelltypen angewendet werden können, wie etwa Backward-Lagrange Modelle, invertierte Gauß-Modelle oder Neuronale Netzwerkmodelle (Stockhause 2000, 9ff). Daneben gibt es noch adaptive Methoden (wie beispielsweise die Single Value Decomposition, die Kalman Filter Technik oder die Variationsanalyse), welche für jeden Modelltyp geeignet sind, da sie die simulierten Konzentrationen an die gemessenen anpassen (Stockhause 2000, 12ff).

3.2 Datenassimilation

Wie in Kapitel 3.1 soeben ausgeführt worden ist, werden zur Erstellung der inversen Modelle Daten verwendet, welche in der Regel über Beobachtung respektive Messung gewonnen werden. In aller Regel reichen die Beobachtungen alleine jedoch nicht aus, um den Anfangs-

¹³ *Trajektorie* bezeichnet allgemein eine Bahnkurve; hier ist jedoch konkret die Bahn eines Luftpartikels gemeint (Seibert 2000, 1).

zustand für ein Ausbreitungsmodell (also z.B. den Quellterm) durch inverse Modellierung zu bestimmen. Besonders für unterbestimmte Teile des Modells wird daher versucht, zusätzliche Informationen zu nutzen, welche dann in das Modell einfließen (Wergen 2002, 144). Es empfiehlt sich also, neben den Messungen und dem eigentlichen Ausbreitungsmodell auch noch einen so genannten First Guess und eine Methode zur Verknüpfung der Information aus beobachteten und modellierten Werten¹⁴ zu implementieren (Wergen 2002, 144f). Diese Verknüpfung wird als Datenassimilation¹⁵ bezeichnet.

Wird bei der Validierung des Modells festgestellt, dass es fehlerhafte Resultate liefert, so kann dies verschiedene Gründe haben. Wergen (2002, 142) nennt dabei Fehler im Anfangszustand (i.e. die Beobachtungen waren ungenau oder unvollständig) und Fehler im Vorhersagemodell als die beiden häufigsten Fehlerquellen. Um diese Fehler zu beseitigen, wird mit Hilfe der Methode der Datenassimilation¹⁶ der Modelllauf an die tatsächliche atmosphärische Situation, welche durch die gegebenen Messungen beschrieben wird, angepasst¹⁷ (Wergen 2002, 142). Dazu wird die atmosphärische Situation zu einem bestimmten Zeitpunkt bestmöglich durch die prognostischen Gleichungen des Modells beschrieben. Unter Einbeziehung der (möglicherweise ungenauen und unvollständigen) Beobachtungen wird so versucht, den wahrscheinlichsten Zustand der Atmosphäre zu dem gewählten Zeitpunkt darzustellen und zu untersuchen. So können Fehler in der Analyse identifiziert und im Modell korrigiert werden (Deutscher Wetterdienst 2012, s.p.).

Durch den Assimilationsprozess wird das Modell also für einen bestimmten Zeitpunkt an die Wirklichkeit angepasst. Bei Anwendung einer iterativen Vorgangsweise spiegelt das Modell somit nicht nur die aktuelle Situation, sondern auch zeitlich weiter zurückliegende, bereits assimilierte Beobachtungen wieder (Wergen 2002, 142). Bei einer erfolgreichen Datenassimilation werden also die wirklichen Werte durch das Modell sukzessive approximiert.

3.3 Ausbreitungsmodelle

Auch bei den Ausbreitungsmodellen gibt es verschiedene Modelltypen. Die wichtigsten Grundtypen der Ausbreitungsrechnung sind Gauß-Modelle, Euler-Ausbreitungsmodelle und Lagrange-Partikelmodelle (Stockhause 2000, 26ff).

Im Folgenden wird nur auf Lagrange-Partikelmodelle eingegangen, da das für die Modellierung der durch die Unfälle in Fukushima Dai-ichi verursachten ¹³⁷Cs-Konzentrationen verwendete Transportmodell FLEXPART ein Lagrange-Ausbreitungsmodell ist (Stohl et al. 2005).

3.3.1 Das Lagrange-Partikelmodell

Lagrange-Ausbreitungsmodelle beruhen auf dem Prinzip, dass die sich ausbreitende Schadstofffahne durch eine Gruppe masseloser Teilchen beschrieben werden kann. Für diese Teilchen wird ein (gegebenenfalls auch zeit- und raumabhängiges) Wind- und Turbulenzfeld – also Grundströmung¹⁸ und Windfluktuationen – vorgegeben, in welchem sich die Teilchen bewegen und unabhängig voneinander ausbreiten können (wissenschaft-online 2012, s.p.).

¹⁴ Dazu werden Methoden der *Schätztheorie* sowie der mathematischen Optimierung mittels *Ausgleichsrechnung* angewendet.

¹⁵ *Datenassimilation* bezeichnet also Techniken, welche dynamische Modelle mit Messdaten verbinden, um den Output zu verbessern. Bei dem verbesserten Output kann es sich beispielsweise um eine Verbesserung eines Modellparameters, um die Verbesserung der Modellergebnisse für ein konkretes Datenfeld oder um eine bessere Prognose handeln (Müller und von Storch 2004, 62).

¹⁶ Es handelt sich dabei um eine *vierdimensionale* Datenassimilation, welche sich aus drei Raumdimensionen und einer Zeitdimension zusammensetzt (Deutscher Wetterdienst 2012, s.p.).

¹⁷ Grundvoraussetzungen für Datenassimilation sind ein dynamisches Modell, ein Observationsdatensatz und ein konkretes Assimilationsschema (Müller und von Storch 2004, 62).

¹⁸ Als Grundströmung wird die vorherrschende großräumige Luftbewegung in der Atmosphäre bezeichnet, welche durch Störungen, die die Luftmassen aus der Richtung der Grundströmung ablenken, beeinflusst wird (Wetterstation Emden 2009, s.p.)

So werden die Trajektorien einzelner Partikel rekonstruiert. Schließlich wird die Konzentrationsverteilung aus der Aufsummierung der Aufenthaltsdauer der Teilchen in den Zellen eines Raums von entsprechend diskretisierter Topologie bestimmt (Mayer 2006, 18).

Mit Hilfe eines Lagrange'schen Partikelmodells kann also die Ausbreitung von verschiedenen Substanzen unter Verwendung einer großen Anzahl repräsentativer Teilchentrajektorien simuliert werden (Mayer 2006, 36).

3.4 Modellierung der Ausbreitung von ^{137}Cs aus dem KKW Fukushima

Im Folgenden wird die Studie von Stohl et al. (2012) über die Ausbreitung von Radionukliden aus dem KKW Fukushima Dai-ichi vorgestellt und hinsichtlich ihrer Methodik beschrieben¹⁹.

3.4.1 Methodik

Um die ^{137}Cs -Emissionen bestimmen und modellieren zu können, sind von Stohl et al. (2012, 2318ff) sowohl Daten über die Konzentration der Radionuklide in der Atmosphäre, als auch Depositionsdaten erhoben und für die Ausbreitungsrechnung verwendet worden. Diese Messdaten der atmosphärischen Konzentration von ^{137}Cs , welche in der Studie von Stohl et al. (2012) verwendet werden, stammen großteils von der CTBTO und von dem europäischen Netzwerk *Ring of Five* (Ro5). Daneben sind Daten des japanischen *National Institute for Environmental Studies* (NIES), des *Japan Chemical Analysis Center* (JCAC), der *Japan Atomic Energy Agency* (JAEA) und des japanischen Forschungsinstituts RIKEN verwendet worden. Zusätzlich sind noch Daten einer Studie von Hsu et al. (2012) sowie Messwerte von einzelnen Orten, welche durch direkten, persönlichen Kontakt gewonnen werden konnten, herangezogen worden (Stohl et al. 2012, 2340). Die Depositionsdaten von ^{137}Cs stammen von MEXT²⁰.

Ausgehend vom gesamten Nuklidinventar der jeweiligen Kraftwerksblöcke und der Abklingbecken ist unter Berücksichtigung des Unfallhergangs zuerst die Gesamtmenge der freigesetzten Radionuklide geschätzt worden. Um diese Daten für eine a priori Analyse (i.e. für die Inversion des Modells) verwenden zu können, sind sie durch Disaggregation²¹ in eine Zeitreihe gebracht worden. Für den First Guess sind zudem noch weitere ingenieurwissenschaftliche, technische, physikalische, chemische und materialwissenschaftliche Informationen sowie diverse Berichte und Beobachtungen bzw. Messungen herangezogen worden (Stohl et al. 2012, 2316ff).

Für die Berechnung des Quellterms bzw. die Modellierung der Ausbreitung der radioaktiven ^{137}Cs -Isotope ist das Lagrange'sche Ausbreitungsmodell FLEXPART²² verwendet worden (Stohl et al. 2012, 2321). Bei FLEXPART handelt es sich um ein Lagrange'sches Dispersionsmodell, welches ursprünglich zur Berechnung der weiträumigen, mesoskaligen Ausbreitung von Luftschadstoffen entwickelt worden ist und nunmehr nach einigen Weiterentwicklungen auf jede Form atmosphärischer Ausbreitung (Modellierung von Transport, Diffusion und Deposition) angewendet werden kann (Stohl et al. 2005, 2461). Die Ausbreitung von radioaktiv zerfallenden Tracern wird dabei mit Hilfe von Trajektorien bestimmt, welche für eine Vielzahl an Partikeln berechnet werden.

¹⁹ Ähnliche Analysen sind in kleinerem Maßstab u.a. von Schöppner et al. (2011), Leelössy et al. (2011), Yasunari et al. (2011), der BGR (Ceranna und Roß 2011) sowie der ZAMG (2011) und der CTBTO (2012c) durchgeführt worden.

²⁰ MEXT ist das japanische Ministerium für Bildung, Kultur, Sport, Wissenschaft und Technologie.

²¹ Als Disaggregation wird die Aufschlüsselung statistischer Daten einer Gesamtgröße auf unterschiedliche Einzelgrößen bezeichnet. Diese Aufschlüsselung erfolgt nach bestimmten Merkmalen des Datensatzes (VNL 2012, s.p.).

²² Eine genauere Beschreibung des Lagrange'schen Ausbreitungsmodells FLEXPART, welches für die Ausbreitungsrechnungen der Radionuklide aus dem KKW Fukushima Dai-ichi verwendet worden ist, ist in Anhang I zu finden. Weiterführende detaillierte Informationen finden sich in diversen Publikationen über FLEXPART, etwa bei Stohl et al. (1998), Stohl und Thomson (1999) oder Stohl et al. (2005).

Im Rahmen der Modellierung ist FLEXPART dazu eingesetzt worden, zuerst ausgehend von Messwerten durch inverse Modellierung ein Quellterm bestimmt, der dann wiederum in einer Vorwärtssimulation für die Ausbreitungsrechnung verwendet wird. Für die realitätsnahe Modellierung dieser Transportprozesse wird die Ausbreitung von Luftschadstoffen mit Hilfe von entsprechend aufbereiteten meteorologischen Daten für den gewünschten Zeitpunkt simuliert.

Für die Modellierung der ^{137}Cs -Emissionen von Fukushima Dai-ichi sind dabei neben den bereits erläuterten ^{137}Cs -Messwerten meteorologische Daten des *Global Forecast System* (GFS) herangezogen worden. Unter Zuhilfenahme dieser Informationen über die atmosphärische Situation ist die Ausbreitung der Isotope für den Zeitraum von 10. März 2011, 10:00 UTC bis 20. April 2011, 00:00 UTC in Dreistunden-Intervallen für drei verschiedene Höhenschichten (0-50 m, 50-300 m und 300-1000 m) simuliert worden (Stohl et al. 2012, 2321f).

3.4.2 Sensitivität der Messstationen

Die Sensitivität²³ der Messstationen²⁴ gegenüber den ^{137}Cs -Emissionen von Fukushima Dai-ichi nimmt grundsätzlich mit zunehmender Entfernung ab. So sind die Daten, welche direkt in Japan gemessen worden sind, durch hohe Sensitivitätswerte charakterisiert, während etwa Daten von Messstationen jenseits des Pazifischen Ozeans – u.a. bedingt durch Auswaschungen, etc. – durch niedrigere Sensitivität gekennzeichnet sind (Stohl et al. 2012, 2324).

3.4.3 Ergebnisse

Die a posteriori ermittelte Gesamtemission an ^{137}Cs wird von Stohl et al. (2012, 2327) mit 36,6 Pbq angegeben. Im First Guess ist die Gesamtemission a priori mit 21,5 PBq angenommen worden (Stohl et al. 2012, 2319). Die genauen Ergebnisse der Untersuchung können unter Stohl et al. (2012, 2323) nachgelesen werden.

²³ Die Messstationen verfügen über unterschiedliche Detektionsschwellen. Die Messschwelle für ^{137}Cs in der Atmosphäre liegt bei $1\mu\text{Bq}/\text{m}^3$. Bei einer Sensitivität von $10^{-15}\text{ Bq}/\text{m}^3$ pro Bq/s sind somit Emissionswerte über einem GBq/s detektierbar (Stohl et al. 2012, 2324). Die Detektionsschwelle für ^{137}Cs im Boden liegt bei $2\text{ Bq}/\text{m}^2$ (Stohl et al. 2012, 2324). Bei den Depositionswerten variiert die Sensitivität der Messungen dabei zwischen $10^{-12}\text{ Bq}/\text{m}^2$ pro Bq/s und $10^{-6}\text{ Bq}/\text{m}^2$ pro Bq/s. Somit sind Emissionswerte größer als $5 \times 10^6\text{ Bq}/\text{m}^2$ pro Bq/s respektive $5 \times 10^{12}\text{ Bq}/\text{m}^2$ detektierbar (Stohl et al. 2012, 2324).

²⁴ Nähere Informationen zur Messung radioaktiver Strahlung finden sich in Bieringer et al. 2000.

4 Quantitative Evaluierung und Validierung von Modellen

Um Qualität, Zuverlässigkeit und Genauigkeit von einem Modell beurteilen zu können – i.e., um festzustellen, wie „realistisch“ es ist – wird es validiert²⁵. Diese Kontrolle erfolgt grundsätzlich durch einen Vergleich des Modells mit der Wirklichkeit, also durch einen Vergleich von modellierten und beobachteten Werten.

Die Evaluierung von Modellen, welche sich mit dem atmosphärischen Transport von Luftschadstoffen beschäftigen, kann qualitativ und quantitativ erfolgen. Die qualitative Analyse erfolgt beispielsweise durch einen Vergleich zweier Kartenausschnitte desselben Gebiets, wobei auf einer Karte die modellierten und auf der anderen die beobachteten Emissionswerte dargestellt sind (Stohl et al. 1998, 4248). Für die quantitative Analyse werden in der Regel statistische Parameter herangezogen. Auf die Validierung mittels quantitativer Analyse wird im Folgenden näher eingegangen.

4.1 Fehler

Bei einer quantitativen Analyse ist zu bedenken, dass sowohl das Modell als auch die Messungen bzw. die beobachteten Werte Fehler enthalten können (Müller und von Storch 2004, 59). Im Sinne einer aussagekräftigen Validierung eines Modells sollten also neben den Fehlern zwischen Modell und Wirklichkeit auch die Fehler innerhalb des Modells und die Fehler in der Beobachtung bedacht werden.

4.1.1 Fehler im Modell

Der *Modellvektor*²⁶ ψ ist ein Schätzer für den tatsächlich korrekten Modellvektor ψ_t . Die Dimension des Modellvektors ψ ergibt sich dabei aus der Multiplikation der Dimension des *Zustandsvektors* ψ_i mit der Anzahl der *Orte* respektive der *Zeitschritte* $i = 1, 2, 3, \dots, K$. Die Differenz zwischen dem angenommenen Modellvektor ψ und dem tatsächlich korrekten Modellvektor ψ_t wird durch den Parameter γ , welcher alle Formen von *Modellfehlern* zusammenfasst, beschrieben (Formel 4.1).

$$\psi = \psi_t + \gamma \quad (4.1)$$

In einem perfekten Modell, welches die Realität exakt beschreibt, gibt es keine Fehler. Somit wäre $\gamma = 0$, und ergo gälte $\psi = \psi_t$. Wie oben jedoch bereits ausgeführt worden ist, gibt es keine derartigen perfekten (i.e. fehlerfreien) Modelle, weshalb immer eine bestimmte Unsicherheit in Form eines Fehlers besteht. Dabei ist zudem zu bedenken, dass modellierte Variablen als Durchschnittswerte interpretiert werden müssen (Müller und von Storch 2004, 59).

4.1.2 Fehler in der Beobachtung

Analog zum Modellfehler werden alle Beobachtungen zu dem *Observationsvektor* ω zusammengefasst. Dieser angenommene Observationsvektor ω gilt wiederum als Schätzer für den tatsächlich korrekten Observationsvektor ω_t . Auch hier wird ein Parameter λ definiert, welcher sämtliche Formen von *Mess-* respektive *Beobachtungsfehlern* subsumiert und die Differenz zwischen ω und ω_t beschreibt (Formel 4.2).

$$\omega = \omega_t + \lambda \quad (4.2)$$

²⁵ Müller und von Storch (2004, 56) weisen darauf hin, dass Modelle – ebenso wie andere kognitive Konstrukte – niemals verifiziert werden können. Selbst wenn sie unter bestimmten Umständen die Wirklichkeit genau beschrieben oder vorhersagten, ist nicht sicher, dass dies unter sämtlichen potentiell möglichen Umständen der Fall ist. Ergo können Modelle nur falsifiziert werden. Dabei stellt sich nicht die Frage, ob ein Modell richtig oder falsch ist, sondern ob es die Wirklichkeit aufgrund der richtigen Gründe gut reproduziert.

²⁶ Ein Vektor bezeichnet in diesem Fall somit eine geordnete Menge von Daten des gleichen Typs. Es handelt sich dabei also um eine Matrix mit nur einer Spalte oder Zeile (Holland 2012, s.p.).

Auch hier gilt analog zu den Modellfehlern, dass niemals ausgeschlossen werden kann, dass bei der Messung Fehler passieren, weshalb auch bei der Beobachtung eine gewisse Unsicherheit in Form eines Fehlers bestehen bleibt. Neben klassischen Fehlern, welche etwa in Fehlfunktionen der Messgeräte oder Fehler in der Dokumentation der Daten begründet sind, ergibt sich die Unsicherheit auch aus der Tatsache, dass Messwerte nicht beliebig verfügbar sind, sondern zeitlich und räumlich an Messstationen gebunden sind. Außerdem ist zu bedenken, dass Beobachtungen nicht notwendigerweise immer reine Zustandsvariablen beschreiben, sondern möglicherweise von Größen beeinflusst sind, welche mit den Zustandsvariablen zusammenhängen (Müller und von Storch 2004, 59).

4.1.3 Umgang mit Fehlern

Da Fehler nicht gänzlich ausgeschlossen werden können, muss es also Ziel sein, die Parameter γ und λ zu minimieren. Diese potentiellen Fehler werden in der Regel mit statistischen Parametern (wie beispielsweise der Standardabweichung) beschrieben, welche die Unsicherheiten im Modell und bei den Messungen angeben.

4.1.4 Der Zusammenhang zwischen Modell und Realität

Unter Berücksichtigung der oben ausgeführten Gegebenheiten und Unsicherheiten kann nun folgende Observationsgleichung (Formel 4.3) aufgestellt werden, die mit Hilfe eines *Operators* C^{27} eine Verbindung von Modell und Realität herstellt:

$$\omega_t = C[\psi_t] \quad (4.3)$$

Setzt man nun den Modellvektor ψ in diese Observationsgleichung ein, erhält man eine *Modellschätzung* ω^m für die Beobachtungen (Formel 4.4).

$$\omega^m = C[\psi] \quad (4.4)$$

Diese modellierten Werte müssen nun im Rahmen der Validierung des Modells mit dem Observationsvektor ω verglichen werden (Müller und von Storch 2004, 59f). Wie bereits eingangs in Kapitel 4 erwähnt worden ist, kann dieser Vergleich qualitativ und quantitativ erfolgen. Auf die statistischen Parameter, welche für die quantitative Analyse verwendet werden, wird im Folgenden näher eingegangen.

4.2 Statistische Parameter und Fehlermaße

Im Bereich der Modellierung der Ausbreitung von Luftschadstoffen haben sich einige Parameter durchgesetzt, welche für die Evaluierung und Validierung von derartigen Ausbreitungsmodellen verwendet werden. Daneben gibt es noch einige weitere Fehlermaße, welche bislang zwar eher unbeachtet bleiben, sich aber dennoch gut für Modellvalidierungen eignen.

Die Tauglichkeit dieser statistischen Parameter für die Modellvalidierung wird in der Regel durch stochastische Verfahren überprüft. Eine der dazu am häufigsten eingesetzten Methoden ist die Monte Carlo Simulation. Dabei handelt es sich um ein stochastisches Verfahren, das auf dem Gesetz der großen Zahlen²⁸ – und somit auf vielfach durchgeführten Zufallsexperimenten – beruht. Sie wird unter anderem dazu eingesetzt, Probleme, die analytisch nicht

²⁷ Der Operator C kann dabei in unterschiedlichen Komplexitätsniveaus auftreten. Im einfachsten Fall ist C dabei eine Projektion, also ein Operator, der aus den Werten aller Gitterzellen die Werte für jene Gitterzellen selektiert, die für die Messungen herangezogen werden. In den meisten anderen Fällen muss C jedoch näher spezifiziert werden (Müller und von Storch 2004, 59).

²⁸ Das Gesetz der großen Zahlen ist eine grundlegende Gesetzmäßigkeit der Wahrscheinlichkeitstheorie. Es besagt, dass eine Folge von Zufallsvariablen bei wachsendem Stichprobenumfang gegen den Erwartungswert konvergiert (Albrecht 2011, 264). Anders ausgedrückt bedeutet dies, dass die relative Häufigkeit, mit der ein Zufallsereignis eintritt, sich der rechnerischen Wahrscheinlichkeit schrittweise immer weiter annähert, je öfter das Zufallsexperiment – unter denselben Voraussetzungen – durchgeführt wird (Statista 2012, s.p.). Mit Hilfe des Gesetzes der großen Zahlen kann das Grundmuster einer Massenerscheinung durch Betrachtung einer ausreichend großen Zahl an Einheiten der statistischen Masse beobachtet werden (Leiner 2004, 192).

oder nur mit einem unverhältnismäßigen Aufwand zu lösen sind, unter Zuhilfenahme der Wahrscheinlichkeitstheorie numerisch zu lösen. In der konkreten Anwendung zur Überprüfung der Aussagekraft der Parameter dient die Monte Carlo Simulation dazu, die originalen Messwerte durch normalverteilte randomisierte Fehler (mit gegebener Standardabweichung) zu manipulieren (Stohl et al. 1998, 4248f). Danach werden die Parameter sowohl für den Originaldatensatz, als auch für die manipulierten Daten berechnet. Aus den Ergebnissen kann man die Robustheit der Fehlermaße respektive die Sensitivität der Parameter gegenüber Messfehlern ableiten.

Im Folgenden werden nun einige verbreitete und einige weniger gebräuchliche statistische Parameter zur Modellevaluierung vorgestellt, beschrieben, und hinsichtlich ihrer Verwendbarkeit für die Evaluierung atmosphärischer Dispersionsmodelle analysiert.

Dabei bezeichnet $(x_i^m, x_i^o), i \in [1, N]$ ein Datenpaar aus modellierten (x_i^m) und beobachteten (x_i^o) Werten²⁹. N steht für die Größe der Stichprobe bzw. für die Anzahl der Datenpaare.

4.2.1 Mittelwerte

Mittelwerte sind Kenngrößen, welche zur Lagebeschreibung eines kontinuierlichen Merkmals eines Datensatzes herangezogen werden können. Mittelwerte sind somit so genannte Lagemaßzahlen (Strelec et al. 2008, 15). Man unterscheidet verschiedene Arten von Mittelwerten. Die zwei wichtigsten Lageparameter werden im Folgenden kurz vorgestellt.

Der *arithmetische Mittelwert* \bar{x} (der Durchschnitt) ist definiert als Quotient aus der Summe aller Werte und der Anzahl der Werte (Formel 4.5).

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i \quad (4.5)$$

Da der arithmetische Mittelwert – obzwar er ein geeigneter Schätzwert für den Erwartungswert jener Verteilung, aus der die Stichprobe stammt, ist – empfindlich gegenüber Ausreißern ist, kann der *Median* \tilde{x} verwendet werden. Der Median ist der Zentralwert eines Datensatzes, also jener Wert, der in einer nach Größe sortierten Zahlenreihe in der Mitte steht.

Formel 4.6 zeigt die Berechnung des Medians für eine ungerade Anzahl an Beobachtungen (n ungerade), während Formel 4.7 die Berechnung des Medians für eine gerade Anzahl an Messwerten (n gerade) definiert.

$$\tilde{x} = x_{\left(\frac{n+1}{2}\right)} \quad (4.6)$$

$$\tilde{x} = \frac{1}{2} \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right) \quad (4.7)$$

4.2.2 Standardabweichung

Die *Standardabweichung* s ist eine Streuungsmaßzahl (Strelec et al. 2008, 16), i.e. sie gibt die Streuung der Werte einer Zufallsvariablen um ihren Mittelwert an (Formel 4.8).

Das Quadrat der Standardabweichung wird als *Varianz* σ^2 bezeichnet.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.8)$$

²⁹ Diese Bezeichnungen sind auf das Anwendungsbeispiel der Validierung von atmosphärischen Transportmodellen hin gewählt und konstant beibehalten worden. In der Stochastik wird der Schätzer üblicherweise mit $\hat{\theta}$ bezeichnet, während θ (statt des beobachteten Wertes) für den geschätzten Parameter steht (Lebanon 2010).

Mittelwerte und Standardabweichung/Varianz sind einfach zu berechnende Kenngrößen. Obwohl sie zweckdienliche Grundinformationen über einen Datensatz enthalten und auch bei der Modellvalidierung verwendet werden, indem diese Parameter für die beobachteten und modellierten Werte gegenübergestellt werden (c.f. EPA 1984, 19ff; Slørdal 2002, 93f), kann man lediglich basierend auf diesen beiden Parametern keine detaillierten Schlüsse über die Modellqualität ziehen. Zwar kann festgestellt werden, dass es eine Abweichung zwischen Modell und Wirklichkeit gibt – woher diese rührt, wie sie sich äußert und welche Teile des Modells betroffen sind, ist allerdings nicht beurteilbar. Da die Ermittlung dieser Parameter jedoch einen geringen Aufwand darstellt, empfiehlt es sich, diese statistischen Kenngrößen zur statistischen Beschreibung des Datensätze mit anzugeben.

4.2.3 Bias

Als *Bias* wird die mittlere Differenz zwischen den Einzelwerten der Datenpaare – also zwischen beobachteten und gemessenen Werten – bezeichnet (Mosca et al. 2008a, 4310). Bias bezeichnet somit das Ausmaß der Verzerrung – also die systematische Differenz – zwischen zwei kompletten Datensätzen (Formel 4.9).

$$B = \frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o) = \overline{x^m} - \overline{x^o} \quad (4.9)$$

Bei einem *erwartungstreu* (also unverzerrten) Parameterschätzer x_i^m ist der *Erwartungswert* der Schätzfunktion $E(x_i^m)$ gleich dem wahren Wert des wirklichen Parameters x_i^o . Bei einem derart perfekten Modell, ist der Bias gleich 0. Man spricht von einem nicht verzerrten Modell, der Schätzer ist also *unbiased*. Dabei muss nicht zwangsweise $\psi = \omega$ gelten. Einzelne Werte des Datensatzes können sehr wohl Differenzen aufweisen, da der Bias mit der Differenz der *Mittelwerte* berechnet wird.

Für $E(x_i^m) \neq x_i^o$ gilt $E(x_i^m) - x_i^o \neq 0$ und somit $\overline{x^m} - \overline{x^o} \neq 0$. Der Schätzer beschreibt den zu schätzenden Parameter infolgedessen nicht exakt; er ist nicht erwartungstreu. In diesem Fall ist der Parameter und somit das Modell *biased*, also verzerrt. Ist der Erwartungswert größer als der zu schätzende Parameter, gilt also $E(x_i^m) > x_i^o$ bzw. $\overline{x^m} > \overline{x^o}$, so wird der Parameter systematisch überschätzt. Man erhält einen positiven Bias. Bei einem negativen Bias gilt $E(x_i^m) < x_i^o$ bzw. $\overline{x^m} < \overline{x^o}$, der zu schätzende Parameter wird somit tendenziell unterschätzt, und ergo unterschätzt das Modell die tatsächliche Situation.

Für die Modellvalidierung hat sich eine normierte Variante etabliert (c.f. Mosca et al. 1998a; Stohl et. al 1998, Seibert 2004), welche als *Fractional Bias* (FB) oder *Relative Bias* bezeichnet wird (Formel 4.10).

$$FB = \frac{2B}{\overline{x^m} + \overline{x^o}} = 2 \times \frac{\overline{x^m} - \overline{x^o}}{\overline{x^m} + \overline{x^o}} \quad (4.10)$$

Somit wird der FB durch die Normierung von den arithmetischen Mittelwerten der modellierten ($\overline{x^m}$) und der beobachteten ($\overline{x^o}$) Daten beeinflusst. Daraus ergibt sich eine Problematik, die auch in Kapitel 4.2.4 noch einmal aufgegriffen wird. Dadurch, dass die modellierten Werte in den Nenner des Bruchs eingehen, liegt der Nachteil dieses so definierten *Relative Bias* in der Eigenschaft dieser Kennzahl, Modelle, die überschätzen, systematisch besser zu bewerten (i.e., der FB ist kleiner) als Modelle, die unterschätzen.

Sofern also keine relativen Werte benötigt werden, gibt es demnach keinen Grund, den Bias zu normieren. Sollten verschiedene Größen, die miteinander verglichen werden sollen, eine Normierung notwendig machen, empfiehlt es sich, die modellierten Werte für die Normierung zu vermeiden und stattdessen nur mit den beobachteten Werten oder mit einem eigens zu definierenden, von den modellierten Werten unabhängigen, Parameter zu normieren (siehe dazu auch die Überlegungen zur Normierung der mittleren quadratischen Abweichung in Kapitel 4.2.4).

Der *Bias* ist eine wichtige Kennzahl für die Evaluierung atmosphärischer Transportmodelle, weil mit Hilfe dieser Kennzahl festgestellt werden kann, ob ein Modell systematisch verzerrt. Das Wissen um die Tatsache, dass ein Modell tendenziell überschätzt oder unterschätzt, ist eine hilfreiche Information, die bei der Interpretation von Ergebnissen, welche mit Hilfe des Modells generiert worden sind, berücksichtigt werden kann. Sollte die Verzerrung jedoch zu groß sein, müssen Korrekturen im Modell vorgenommen werden. Aus Sicht des Autors ist dabei im Rahmen der Modellvalidierung der *Bias* dem *Fractional Bias* vorzuziehen, weil der Wert des *Fractional Bias* sehr stark vom Modell beeinflusst wird.

4.2.4 Mittlere quadratische Abweichung

Die *mittlere quadratische Abweichung* (auch mittlerer quadratischer Fehler), in ihrer englischen Bezeichnung *Mean Squared Error* (MSE), gibt die mittlere Abweichung eines Schätzers vom zu schätzenden Wert – also zwischen zwei Variablen – an (Formel 4.11).

$$MSE = \frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o)^2 \quad (4.11)$$

Je kleiner die mittlere quadratische Abweichung ist, desto besser stimmen Modell und Wirklichkeit im Mittel überein. Der mittlere quadratische Fehler ist das zweite zentrale Moment des Fehlers und beinhaltet sowohl Varianz als auch Bias (Formel 4.12). Ein kleiner MSE-Wert bedeutet somit, dass der Schätzer den tatsächlichen Wert gut trifft und die Streuung dabei gering ist, i.e. dass der Schätzer in der Nähe des Erwartungswertes liegt (Lebanon 2010, 1).

$$MSE(x_i) = E[(x_i^m - x_i^o)^2] = Var(x_i) + [Bias(x_i)]^2 \quad (4.12)$$

Bei der Evaluierung von atmosphärischen Dispersionsmodellen hat es sich eingebürgert, eine abgewandelte, (i.e. normierte) Form des MSE zu Validierungszwecken zu verwenden (c.f. Mosca et al. 1998a; Stohl et. al 1998, Seibert 2004). Die gebräuchlichste Form dieses *Normalized Mean Squared Error* (NMSE) ist in Formel 4.13 dargestellt.

$$NMSE = \frac{\frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o)^2}{\overline{x^m x^o}} \quad (4.13)$$

Somit hängt der NMSE einerseits von den quadratischen Differenzen der Datenpunkte $(x_i^m - x_i^o)^2$ ab. Andererseits wird der NMSE von den arithmetischen Mittelwerten der modellierten $(\overline{x^m})$ und der beobachteten $(\overline{x^o})$ Daten beeinflusst. Wie sowohl Seibert (2004, 2) als auch Stohl et al. (1998, 4249) zu Recht anmerken, liegt der Nachteil dieses so definierten NMSE somit in der Eigenschaft dieser Kennzahl, Modelle, die überschätzen, systematisch besser zu bewerten (i.e., der NMSE ist kleiner) als Modelle, die unterschätzen.

Sofern also keine relativen Werte benötigt werden, gibt es demnach keinen Grund, den MSE zu normieren. Eine praktikable Möglichkeit wäre somit die Verwendung des *Root Mean Squared Error* (RMSE), also der Wurzel des MSE (Formel 4.14).

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o)^2} \quad (4.14)$$

Sollten verschiedene Größen, die miteinander verglichen werden sollen, eine Normierung notwendig machen, empfiehlt es sich, die modellierten Werte für die Normierung zu vermeiden. Seibert (2004, 2) schlägt hierfür vor, nur die beobachteten Werte (Formel 4.15) respektive eine – von den Datenpaaren unabhängige – charakteristische Größe (Formel 4.16) für die Normierung zu verwenden.

$$NMSE = \frac{\frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o)^2}{\overline{x^o{}^2}} \quad (4.15)$$

$$NMSE = \frac{1}{N} \sum_{i=1}^n (x_i^m - x_i^o)^2 \quad (4.16)$$

Seibert (2004, 3) weist zudem darauf hin, dass eine Normierung mit Parametern, die auf Mittelwerten beruhen, nur dann sinnvoll ist, wenn diese Mittelwerte der Größenordnung der Varianz entsprechen. Die Verwendung von Varianzen (Formel 4.17) oder von Kovarianzmatrizen zur Normierung erscheint somit zweckmäßig.

$$x_*^2 = \sigma_x^2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (4.17)$$

Da der Bias mit dem Quadrat in die Berechnung des MSE eingeht, hat eine Verzerrung des Modells signifikante Auswirkungen auf den MSE. Wenn man nur die mittlere quadratische Abweichung betrachtet, impliziert ein hoher Bias also einen hohen MSE-Wert, was bei der Interpretation auch auf eine hohe Varianz zurückgeführt werden kann. Seibert (2004, 4) merkt dazu an, dass eine Verschiebung der Daten gegeneinander um einen gewissen Offset einen großen Bias und somit einen großen RMSE ergibt, obwohl die Varianz klein sein kann, da die Daten abgesehen von der Verschiebung möglicherweise gut übereinstimmen. Um diesen Fehler und somit Fehlinterpretationen zu vermeiden, kann der *Bias-Corrected Root Mean Squared Error* ($RMSE_{BC}$) herangezogen werden (Formel 4.18).

$$RMSE_{BC} = \sqrt{\frac{1}{N} \sum_{i=1}^n [(x_i^m - \bar{x}^m) - (x_i^o - \bar{x}^o)]^2} \quad (4.18)$$

Aufgrund der Sensitivität in Bezug auf Unterschiede zwischen beobachteten und modellierten Werten und aufgrund der Aussagen, die der Parameter über die Qualität des Modells – sowohl in räumlicher als auch in zeitlicher Hinsicht – erlaubt (Mosca et al. 1998a, 4312), hat sich der NMSE in der Validierung von atmosphärischen Schadstoffausbreitungsmodellen etabliert. Es ist jedoch – wie oben ausgeführt – grundsätzlich zu bedenken, dass die mittlere quadratische Abweichung insofern mit Vorsicht zu genießen ist, als sowohl Bias als auch Varianz in diesem Parameter enthalten sind. Ein auf den ersten Blick niedrig anmutender NMSE muss also nicht zwingend bedeuten, dass das ganze Modell schlecht ist. Insbesondere Abweichungen bei den Maximalwerten haben besonders starke Auswirkungen auf den NMSE (Mosca et al. 1998a, 4312). Nichtsdestotrotz ist die mittlere quadratische Abweichung bei entsprechend überlegtem Umgang mit den Ergebnissen eine potente Kennzahl. Stohl et al. (1998, 4249) erklären zudem, dass der NMSE aufgrund seiner Verbreitung in der Modellvalidierung zur Evaluierung herangezogen wird, obwohl die Tatsache, dass der NMSE aufgrund der Abhängigkeit von der Verteilung der Messungen und der Bevorzugung von überschätzenden Modellen für derartige Evaluierungen suboptimal ist, bekannt ist.

Obwohl der NMSE die gebräuchlichste MSE-Kennzahl für die Validierung von atmosphärischen Transportmodellen ist, wird in dieser Arbeit aufgrund der oben dargelegten Problematik der Normierung der RMSE verwendet. Der $RMSE_{BC}$ wird sowohl aufgrund der geringen Verbreitung³⁰ als auch aufgrund der Tatsache, dass man nicht im Vorhinein per se von einem hohen Bias und einer geringen Varianz ausgehen darf lediglich als unterstützender Beurteilungsparameter bzw. als Ergänzung zum RMSE herangezogen.

4.2.5 Mean Absolute Error

Der *Mean Absolute Error* MAE erlaubt Rückschlüsse auf systematische Fehler in einem Datensatz bzw. auf systematische Abweichungen zwischen modellierten und beobachteten

³⁰ Die einzige gefundene Referenz ist Zinevich et al. 2008, 1476, welche den $RMSE_{BC}$ – allerdings in einer normierten Form – im Rahmen der Evaluierung einer Niederschlagsmodellierung verwenden.

Werten. Er errechnet sich aus den gemittelten Absolutwerten der Abweichungen zwischen beobachteten und modellierten Werten und gibt somit die Größe der mittleren absoluten Abweichung zwischen Modell und Realität an (Formel 4.19).

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i^m - x_i^o| \quad (4.19)$$

Der Wert für den MAE ist bei einer perfekten Übereinstimmung zwischen Modell und Realität $MAE = 0$ und ist theoretisch nach oben nicht begrenzt. Je näher der Wert bei 0 ist, desto besser stimmen Modell und Wirklichkeit überein.

Yu et al. (2006, 27) erläutern, dass der MAE – ähnlich wie die Standardabweichung als Ergänzung zur mittleren Abweichung – oft zusätzlich zum Bias angegeben wird, um die Bandbreite der Abweichung zu charakterisieren. Im Prinzip ist der Informationsgehalt des MAE vergleichbar mit dem des RMSE. Der RMSE reagiert jedoch stärker auf große Abweichungen.

Auch bei diesem Parameter gibt es wiederum normalisierte Derivate, für deren Verwendung es – gesetzt den Fall, dass keine relativen Werte benötigt werden – aus den oben bei den Parametern Bias (Kapitel 4.2.3) und NMSE (Kapitel 4.2.4) bereits dargelegten Gründen keine Notwendigkeit gibt³¹.

4.2.6 Pearsons Korrelationskoeffizient

Der *Korrelationskoeffizient* r (auch Produkt-Moment-Korrelation oder *Pearsons* r) ist ein häufig verwendetes dimensionsloses Korrelationsmaß, welches dazu dient, den Grad des linearen Zusammenhangs zwischen zwei metrischen Variablen zu beschreiben. Pearsons Korrelationskoeffizient gibt dabei den Anteil an erklärter Variabilität an der Gesamtvariabilität an. Er wird berechnet, indem die Kovarianz zweier Variablen durch das Produkt der Standardabweichungen der beiden Variablen dividiert wird (Formel 4.20).

$$r = \frac{\sum_{i=1}^n (x_i^m - \bar{x}^m)(x_i^o - \bar{x}^o)}{\sqrt{\sum_{i=1}^n (x_i^m - \bar{x}^m)^2 \sum_{i=1}^n (x_i^o - \bar{x}^o)^2}} \quad (4.20)$$

Der Korrelationskoeffizient kann dabei einen Wert zwischen -1 und $+1$ annehmen, wobei die Korrelation immer signifikanter wird, je weiter der Wert von 0 entfernt ist. Ein Wert von 0 indiziert daher das Fehlen jedweder Form eines linearen Zusammenhangs (Ludwig-Mayerhofer 2012, s.p.)

4.2.7 Rangkorrelationskoeffizienten

Rangkorrelationskoeffizienten sind verteilungsfreie Korrelationsmaße, welche dazu verwendet werden, um den Grad des Zusammenhangs zwischen zwei Variablen zu beschreiben. Dadurch, dass es sich um non-parametrische Maßzahlen handelt, haben sie gegenüber Pearsons Korrelationskoeffizienten den Vorteil, dass die Annahme eines linearen Zusammenhangs keine Voraussetzung für die Anwendung von Rangkorrelationskoeffizienten ist.

Es gibt zwei gebräuchliche Rangkorrelationskoeffizienten, nämlich *Spearman's Rho* ρ und *Kendalls Tau* τ . Diese beiden werden im Folgenden kurz präsentiert. Weitere Rangkorrelationsmaße, wie etwa *Goodman und Kruskals Gamma* γ werden aufgrund der geringen Relevanz für die Validierung von atmosphärischen Dispersionsmodellen ausgeklammert.

Spearman's Rho beruht auf dem Pearsonschen Produkt-Moment-Korrelationskoeffizienten, in welchen Rangplätze eingesetzt werden; i.e. die Daten werden vor der Berechnung des Koeffizienten in Ränge konvertiert. Der Unterschied zu *Pearsons* r besteht also im Prinzip nur darin, dass nicht die Werte der Variablen, sondern deren Ränge betrachtet werden (Brosius

³¹ Diese normalisierten Parameter des Gross Error umfassen etwa den *overall normalized gross error* NGE, den *location-specific gross error* NGE_x und den *time-specific gross error* NGE_t (cf. Jacobsen 1999, 550) sowie den *mean absolute normalized gross error* MANGE (cf. Doty et al. 2002, 108).

1998, 510). Analog zu *Pearsons r* berechnet sich also *Spearman's Rho*, indem man die Kovarianz der Ränge zweier Variablen durch die Wurzel des Produkts der Standardabweichungen der Ränge der beiden Variablen dividiert (Formel 4.21).

$$\rho = \frac{\sum_{i=1}^n (r(x_i^m) - \bar{r}_{x^m})(r(x_i^o) - \bar{r}_{x^m})}{\sqrt{\sum_{i=1}^n (r(x_i^m) - \bar{r}_{x^m})^2 \sum_{i=1}^n (r(x_i^o) - \bar{r}_{x^m})^2}} \quad (4.21)$$

Kendalls Tau analysiert die Abhängigkeit bzw. den Zusammenhang zwischen zwei Datensätzen, indem die Fehlordnung der Datenpaare untereinander betrachtet wird. *Kendalls Tau* ergibt sich dabei aus dem Quotienten von der Differenz der Anzahl der konkordanten (übereinstimmenden) Datenpaare n_k und der Anzahl der diskordanten Datenpaare n_d und der Gesamtzahl der Datenpaare (Formel 4.22)³².

$$\tau = \frac{n_k - n_d}{\frac{N(N-1)}{2}} \quad (4.22)$$

Während *Spearman's Rho* ebenso wie *Pearsons r* den Anteil an erklärter Varianz an der Gesamtvariabilität in einem Korrelationsmodell angibt, drückt *Kendalls Tau* dagegen die Wahrscheinlichkeit des Vorliegens der Daten für die beiden Variablen (x_i^m, x_i^o) in derselben Rangfolge aus (StatSoft 2012, s.p.).

Ebenso wie r sind ρ und τ Korrelationskoeffizienten, also Maße für die Beurteilung des Grades einfacher linearer Zusammenhänge. Sämtliche nicht-lineare Zusammenhänge und Abhängigkeiten zwischen Datensätzen werden von diesen Korrelationsmaßen ignoriert. Da im Rahmen der Validierung atmosphärischer Transportmodelle im Idealfall ein linearer Zusammenhang zwischen beobachteten und modellierten Werten vorliegt, können diese Korrelationsmaße prinzipiell für die Validierung verwendet werden. Vom Autor wird jedoch vorzugsweise die Verwendung des *Bestimmtheitsmaßes* R^2 empfohlen, welches zudem auch den Bias mitberücksichtigt. R^2 ist bei nicht verzerrten Regressionsmodellen zwar formal gleich dem Quadrat des *multiplen Korrelationskoeffizienten* r^2 (Hackl 2008, 86f), es besitzt jedoch teilweise divergierende wahrscheinlichkeitstheoretische Eigenschaften (Assenmacher 2002, 125).

4.2.8 Bestimmtheitsmaß

Das *Bestimmtheitsmaß* R^2 (manchmal auch Determinationskoeffizient genannt) ist ein Maß für die Anpassungsgüte (*goodness of fit*) von Regressionsmodellen, welches den Anteil der Varianz, welche durch das Modell erklärt wird (i.e. die Varianz der Schätzwerte \hat{x}_i), an der Gesamtvarianz der Beobachtungen x_i angibt (Hackl 2008, 73). Das Bestimmtheitsmaß ist also der Quotient aus erklärter³³ Varianz und Gesamtvarianz, und kann ergo mit Hilfe des Verhältnisses von erklärter Quadratsumme zu gesamter Quadratsumme berechnet werden (Assenmacher 2002, 120). Formel 4.23 stellt dieses Verhältnis mathematisch dar.

$$R^2 = \frac{\sum_{i=1}^n (\hat{x}_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (4.23)$$

³² Bei dieser Formel handelt es sich um die Berechnung für Kendalls Tau-a. Dies ist die simpelste Form von Kendalls Tau, welche Ties (also verbundene Ränge oder Bindungen) nicht berücksichtigt. Daneben gibt es noch komplexere Berechnungen für Kendalls Tau-b und Kendall-Stuarts Tau-c, welche Anpassungen im Nenner zur Berücksichtigung von Ties in quadratischen (Tau-b) bzw. rechteckigen Tabellen (Tau-c) enthalten.

³³ Obwohl der Terminus „erklärte Varianz“ ein in der Stochastik gebräuchlicher Begriff ist, sei an dieser Stelle angemerkt, dass Müller und von Storch (2004, 264) darauf hinweisen, dass es semantisch korrekt wäre, hier anstatt von *erklärter* Varianz von *beschriebener* Varianz zu sprechen.

In der (englischsprachigen) Literatur ist für Gesamtvariation, erklärte Variation bzw. residuale Variation auch folgende Notation üblich (Hackl 2008, 73):

- TSS: Gesamtvariation von x_i (*total sum of squares*)
- ESS: erklärte Variation der Regresswerte (*explained* oder *estimated sum of squares*)
- RSS: nicht erklärte Variation der Residuen (*residual sum of squares*)

Das Bestimmtheitsmaß kann somit auch folgendermaßen dargestellt werden (Formel 4.24):

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (4.24)$$

R^2 kann Werte zwischen 0 und 1 annehmen. Ein Wert von 0 bedeutet, dass die Wirklichkeit durch das Modell überhaupt nicht erklärt wird. $R^2 = 1$ hingegen impliziert, dass das Regressionsmodell an die Wirklichkeit ideal angepasst ist und somit eine perfekte Erklärung für die tatsächliche Situation liefert.

Obwohl R^2 zur Evaluierung von atmosphärischen Transportmodellen nicht so häufig eingesetzt wird wie andere Parameter, finden sich doch Publikationen, welche dem Determinationskoeffizienten eine wichtige Rolle einräumen (z.B. Fournier et al. 2005). Auch im Rahmen dieser Arbeit wird das Bestimmtheitsmaß zur Beurteilung der Modellgüte herangezogen, weil damit die Qualität des Zusammenhangs von beobachteten und modellierten Daten gut beschrieben werden kann.

4.2.9 Figure of Merit

Im Rahmen der statistischen Analysen bezeichnet *Figure of Merit* einen Koeffizienten, der dazu dient, ein System oder Modell in Bezug auf seine Alternativen zu charakterisieren. Im Rahmen der Modellvalidierung gibt es zwei gebräuchliche Figures of Merit, nämlich *Figure of Merit in Space* (FMS) und *Figure of Merit in Time* (FMT). Die *Figure of Merit in Space* wird anhand der räumlichen Verteilung der Daten bestimmt, während die *Figure of Merit in Time* anhand der zeitlichen Verteilung der Daten errechnet wird.

Aufgrund der durch die Vertraulichkeit der Daten bedingten fehlenden geographischen Informationen zu den Messstellen werden FMS und FMT für die Analyse nicht weiter berücksichtigt. Nähere Informationen zu diesem Parameter finden sich in Anhang II.

4.2.10 Peak Estimation Accuracy

Die *Peak Estimation Accuracy* gibt Auskunft darüber, wie gut ein Modell einen Maximalwert (*peak value*) in räumlicher und/oder zeitlicher Hinsicht simulieren kann. Die *Paired Peak Estimation Accuracy* PPA gibt an, wie gut das Modell den Maximalwert hinsichtlich *Zeit* und *Ort* des Maximalwertes vorhersagen kann.

Aufgrund der durch die Vertraulichkeit der Daten bedingten fehlenden geographischen Informationen zu den Messstellen wird die PPA für die Analyse nicht weiter berücksichtigt. Nähere Informationen zu diesem Parameter finden sich in Anhang III.

4.2.11 Scatterplot, FOEX und FA α

Die einfachste graphische Darstellung der Datenpaare bzw. die Visualisierung ihrer Verteilung erfolgt in der Regel durch ein Streudiagramm (*scatterplot*). Dabei werden die Messwerte auf der Abszissenachse und die modellierten Werte auf der Ordinatenachse aufgetragen. Im Idealfall, also bei perfekter Übereinstimmung von Messwerten und modellierten Werten, ergibt sich eine Gerade $y = x$ respektive $x^m = x^o$, welche im Koordinatenursprung beginnt und eine Steigung von 45° aufweist. Punkte unterhalb dieser Gerade kennzeichnen unterschätzte Werte, während Punkte, welche oberhalb der Gerade liegen, Werte darstellen, die überschätzt worden sind.

Ausgehend von dieser Darstellung im Scatterplot erlaubt der *Factor of Exceedance* (FOEX) Rückschlüsse auf den Bias des Modells (Formel 4.25). O kennzeichnet dabei jene Werte, die

oberhalb der Gerade liegen und steht somit für die Anzahl der Überschätzungen – also für jene Anzahl an Datenpaaren, für die $x^m > x^o$ gilt.

$$FOEX = \left[\frac{O}{N} - 0,5 \right] \times 100 \quad (4.25)$$

Der FOEX kann somit Werte im Intervall von -50% bis $+50\%$ annehmen. Ein FOEX von -50% muss so interpretiert werden, dass alle Werte unterschätzt werden, während ein FOEX von $+50\%$ bedeutet, dass alle Werte überschätzt werden. Somit bedeutet ein FOEX von 0% , dass die Hälfte der Werte über der Geraden und die andere Hälfte unter der Geraden liegt. In diesem Fall liegt keine Verzerrung vor. Der FOEX berücksichtigt dabei allerdings nur die Anzahl der Werte die überschätzt werden; das Maß der Überschätzung bleibt unberücksichtigt (Mosca et al. 1998a, 4311).

Eine weitere Kennzahl, welche sich auf das Streudiagramm bezieht, ist $FA\alpha$. Das $FA\alpha$ -Band bezeichnet den Raum, der zwischen den beiden Linien von Formel 4.26 aufgespannt wird. x_0^m und x_0^o bezeichnen dabei die Koordinaten des Koordinatenursprungs.

$$x^m - x_0^m = (x^o - x_0^o) \pm \ln(\alpha) \quad (4.26)$$

Die gängigsten Werte für α sind 2 und 5. $FA2$ respektive $FA5$ bezeichnen somit den Anteil der prognostizierten Werte, welche innerhalb eines Faktors 2 bzw. 5 der Messwerte liegen (Mosca et al. 1998a, 4311).

Ein Scatterplot mit FOEX und verschiedenen $FA\alpha$ -Bändern kann mit jeder Form von gepaarten Daten relativ leicht erstellt werden und ist im Rahmen der Visualisierung der Datensätze bzw. einer ersten groben qualitativen Analyse sicherlich hilfreich. Die Aussagekraft dieser Parameter ist abgesehen von der graphischen Aufbereitung allerdings nicht besonders hoch. Der FOEX ist angesichts der Berechnung des Bias nur mehr bei Datensätzen mit vielen Extremwerten, welche großen Einfluss auf den Bias ausüben, interessant, und die $FA\alpha$ -Werte reagieren sehr sensitiv auf Messfehler, wie von Stohl et al. (1998, 4248f) erläutert worden ist.

4.2.12 Tests

Neben den statistischen Parametern zur Modellvalidierung können verschiedene statistische Testverfahren angewendet werden, mit Hilfe derer die Abweichungen zwischen modellierten und beobachteten Werten analysiert werden können. Dazu zählen – je nach Form der vorliegenden Daten und unter Berücksichtigung der Grundvoraussetzungen für die Anwendung dieser Testverfahren – etwa Student's t-Test, der Welch-Test, der Wilcoxon-Vorzeichen-Rang-Test, der Kolmogorov-Smirnov-Test oder die Varianzanalyse.

Diese statistischen Tests dienen dabei der Überprüfung der Gleichheit bestimmter Eigenschaften zweier Stichproben aus Sicht der statistischen Mathematik. Mit Hilfe dieser Tests kann so mit einer bestimmten Wahrscheinlichkeit (i.e. anhand eines definierten Signifikanzniveaus) die Gültigkeit einer Hypothese überprüft werden.

5 Analyse der Datenpaare

Aufbauend auf dem theoretischen Hintergrund zu den statistischen Parametern zur Modellvalidierung und dem Background zur atmosphärischen Transportmodellierung wird in diesem Kapitel nun eine Bewertung des mit FLEXPART modellierten Caesium-137-Ausstoßes aus dem havarierten Kernkraftwerk Fukushima Dai-ichi im Zeitraum zwischen 11. März 2011 und 19. April 2011 präsentiert. Wie schon in Kapitel 2 ausgeführt worden ist, werden dafür die von Stohl et al. (2012) für ihr Dispersionsmodell erhobenen und verwendeten ^{137}Cs -Konzentrationswerte als Datengrundlage für die Evaluierung des Dispersionsmodells herangezogen. Dabei handelt es sich um 1494 Datenpaare³⁴, welche aus einem modellierten und einem gemessenen Konzentrationswert bestehen. Beide Werte beziehen sich auf denselben Ort und dieselbe Zeit.

5.1 Ausgewählte Fehlermaße

Für die Analyse bzw. den Vergleich der gemessenen und beobachteten Konzentrationswerte sind die oben beschriebenen statistischen Parameter herangezogen worden. Da in dieser Arbeit keine Inter-Modell-Vergleiche angestellt und somit keine relativen Werte benötigt werden, werden die normalisierten Derivate der statistischen Parameter nicht verwendet.

5.1.1 Erste Analyse der Originaldaten

Wie aus Tab. 1 hervorgeht, sind die Unterschiede zwischen modellierten und gemessenen Werten beträchtlich. R^2 ist extrem gering, der $RMSE$ ist hoch und der Bias impliziert, dass das Modell verzerrt ist und die Wirklichkeit überschätzt. Der negative $FOEX$ hingegen deutet auf Unterschätzung der Realität hin. Dieser scheinbare Widerspruch zum Bias ist damit zu erklären, dass einige sehr hohe Werte den Bias – der aus den Mittelwerten berechnet wird – stark beeinflussen, während jedoch gleichzeitig mehr als die Hälfte der Daten unterhalb der Gerade $x^m = x^o$ liegt. r ist ebenso gering und τ konstatiert einen mittelmäßigen Zusammenhang. ρ erzielt jedoch einen verhältnismäßig guten Wert. Die nur mäßige Übereinstimmung von Modell und Wirklichkeit relativiert sich jedoch etwas, wenn man den Hintergrund der Daten sowie die Datensätze an sich näher betrachtet und einige Optimierungen vornimmt.

5.1.2 Aufbereitung der Daten

Die ^{137}Cs -Konzentrationen, welche aus geographisch weiter entfernten Stationen stammen, sind großteils sehr niedrig. Daraus ergibt sich das Problem, dass die Konzentrationen dieses Radionuklids an einigen Messstellen (besonders in Europa) teilweise so gering gewesen sind, dass sie unter den Messschwellen der Stationen gelegen sind, weshalb diese Stationen keine Messergebnisse liefern konnten. Von den 1494 verfügbaren Datenpaaren ist bei 367 Messwerten keine ^{137}Cs -Konzentrationen detektiert worden, weshalb bei diesen 367 Datenpaaren der Messwert 0 ist.

Daher stellt sich nun die Frage, wie mit Ergebnissen verfahren werden soll, welche unterhalb der Detektionsschwelle liegen. Reimann et al. (2008, 23) weisen darauf hin, dass es wenig praktikabel ist, nicht detektierbare Werte einfach durch 0 zu ersetzen. Zum einen besteht die Möglichkeit, dass 0 ein wahrer Messwert ist, und zum anderen wird die Wirklichkeit durch die simple Ersetzung des Wertes unter dem Detektionslimit durch 0 inkorrekt wiedergegeben, da 0 eine zu gering angesetzte Schätzung für den wahren Wert ist. Darüber hinaus werden durch diese Vorgangsweise logarithmische Transformationen unmöglich, da $\log(0) = -\infty$. In Anlehnung an die Ausführungen von Reimann et al. (2008, 23f) zu diesem Thema wird somit für alle Werte unter der Detektionsschwelle ein Wert angenommen, welcher der Hälfte des Wertes des Detektionslimits entspricht. Wie in Kapitel 3.4.2 bereits erwähnt, liegt die Detekti-

³⁴ Die im Rahmen des First Guess ermittelten Werte werden somit in diesem Kapitel nicht näher betrachtet, da es sich dabei lediglich um eine Kurzfristvorhersage und nicht um das finale (a posteriori) Modell handelt.

onsschwelle der Messstationen für die atmosphärische ^{137}Cs -Konzentration bei $1\mu\text{Bq}/\text{m}^3$. Somit werden in weiterer Folge alle Werte unter der Detektionsschwelle mit einer gemessenen ^{137}Cs -Konzentration von $5 \times 10^{-7} \text{ Bq}$ angenommen.

Tab. 1 zeigt neben den ausgewählten Skill Scores der Originaldaten auch die Abweichung jener Parameter, welche unter Berücksichtigung der korrigierten Messdaten erneut berechnet worden sind, von den ursprünglichen Parametern der Originaldaten. Dabei ist ersichtlich, dass die Veränderungen respektive Verbesserungen im Vergleich zu den Skill Scores der Originaldaten nur minimal sind. Nichtsdestotrotz ist die Korrektur der Nullwerte, welche aus einer technisch bedingten Restriktion resultieren, ein wichtiger Schritt.

Tab. 1: Übersicht über statistische Kennwerte der nicht korrigierten sowie der korrigierten beobachteten und modellierten Werte des ^{137}Cs -Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).

	Parameter der Originaldaten [Bq/m ³]		Abweichung der Parameter der korrigierten Originaldaten gegenüber den Originaldaten [Bq/m ³]	
	Beobachtet	Modelliert	Beobachtet	Modelliert
μ	0,3140	1,3036	-1,23e-07	±0
σ	6,1765	27,4246	+6,25e-09	±0
Maximalwert	183,2	794,5	±0	±0
	Parameter der Originaldaten [1]		Abweichung der Parameter der korrigierten Originaldaten gegenüber den Originaldaten [1]	
Bias		0,9896		+1,23e-07
RMSE		26,6466		+2,03e-13
RMSE _{BC}		26,6283		-4,56e-09
MAE		1,2735		+4,98e-08
r		0,2383		+7,05e-10
ρ		0,7310		+2,13e-04
τ		0,5661		-3,39e-04
R^2		0,0568		+3,36e-10
FOEX		8,5007%		-13,32

5.2 Graphische Darstellung

5.2.1 Scatterplot

Ein Blick auf den Scatterplot der Datenpaare zeigt deren graphische Verteilung (Abb. 1).

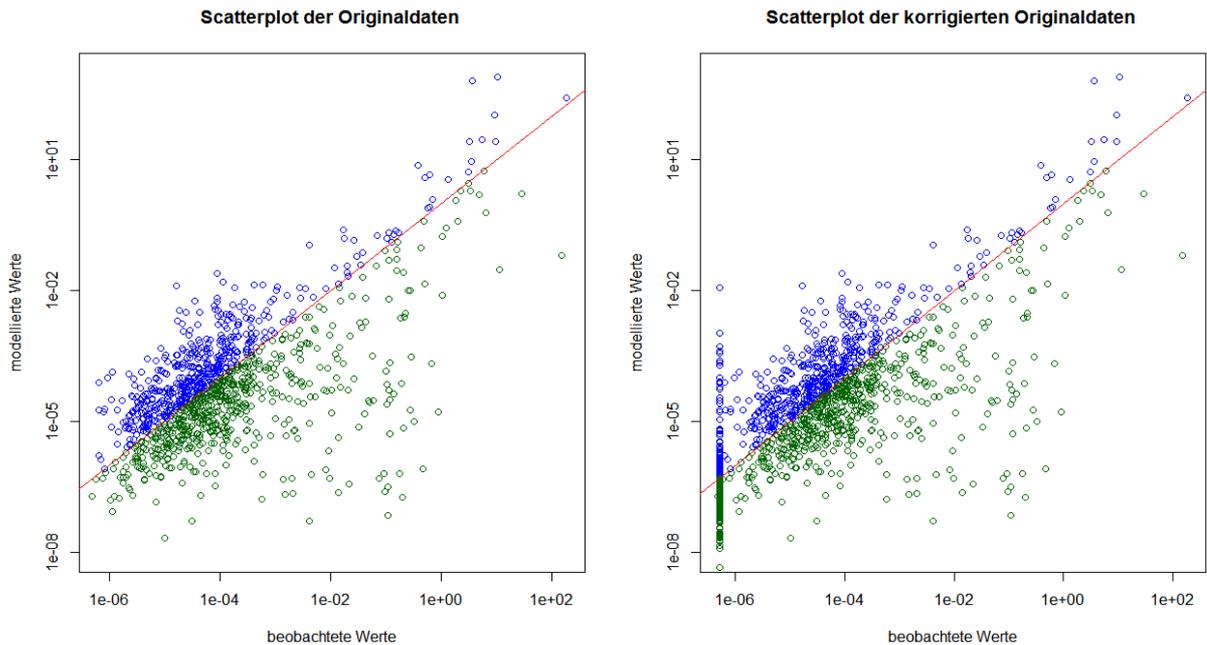


Abb. 1: Scatterplot der nicht korrigierten Originaldaten im Vergleich zum Scatterplot der korrigierten Originaldaten (eigene Darstellung; Datenquelle: Stohl et al. 2012).

Für die Interpretation des Streudiagramms ist die logarithmische Skalierung der Achsen zu beachten, welche aufgrund der enormen Unterschiede in der Höhe der ¹³⁷Cs-Konzentration für eine übersichtliche graphische Darstellung notwendig ist. Die Verteilung der Datenpaare ist somit gut erkennbar. Die rote Linie kennzeichnet die ideale Übereinstimmung zwischen modellierten und beobachteten Werten; i.e. auf der roten Gerade gilt $x^m = x^o$. Aus der Graphik geht hervor, dass zwar grundsätzlich eine Anordnung der Datenpaare um die Gerade herum feststellbar ist, die Datenpaare streuen jedoch nicht unerheblich entlang der Linie.

Tab. 2: Übersicht über statistische Kennwerte der korrigierten logarithmierten beobachteten und modellierten Werte des ¹³⁷Cs-Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).

	Parameter der korrigierten Originaldaten [Bq/m ³]		Parameter der korrigierten logarithmierten Daten [1]	
	Beobachtet	Modelliert	Beobachtet	Modelliert
μ	0,3140	1,3036	-4,5148	-4,6957
σ	6,1765	27,4246	1,5719	1,5494
Maximalwert	183,2	794,5	2,26	2,90
	Parameter der korrigierten Originaldaten [1]		Parameter der korrigierten logarithmierten Daten[1]	
Bias	0,9896		-0,1809	
RMSE	26,6467		1,1064	
RMSE _{BC}	26,6283		1,0915	
MAE	1,2735		0,7365	
r	0,2383		0,7553	
R^2	0,0568		0,5705	

Abb. 1 impliziert zudem, dass die Daten einer logarithmischen Normalverteilung unterliegen³⁵. Diese Hypothese wird dadurch untermauert, dass die korrigierten logarithmierten Daten (Tab. 2) eine im Vergleich zu den Originaldaten sehr gute Performance des Modells – und somit eine hohe Übereinstimmung von Modell und Wirklichkeit – konstatieren. Diese Werte können allerdings nicht für die Evaluierung des Modells herangezogen werden, da die logarithmische Transformation die Ausgangsdaten in ihren Lagebeziehungen zu sehr modifiziert.

5.2.2 Zeitreihen

Zeitreihen ermöglichen einen guten Vergleich zwischen modellierten und beobachteten Werten. Jacobsen (1999, 551) betont die Wichtigkeit von Zeitreihen-Plots und merkt an, dass die Performance eines Modells mit Hilfe von Zeitreihen-Diagrammen oft besser beurteilt werden kann als mit manchen statistischen Parametern.

Abb. 2 verdeutlicht den Unterschied im Zeitverlauf zwischen den durch den First Guess a priori rekonstruierten ¹³⁷Cs-Emissionen (rote Linie) und den a posteriori durch den Inversionsalgorithmus erhaltenen ¹³⁷Cs-Emissionen (blaue Linie).

Mit den vorliegenden Daten kann vom Autor keine aussagekräftige Zeitreihe der gemessenen und der a posteriori modellierten ¹³⁷Cs-Konzentrationen für den Zeitraum zwischen 11. März 2011 und 19. April 2011 erstellt werden, da aufgrund der fehlenden geographischen Informationen zu den Messstellen die Werte nicht zu Messstationen zugeordnet werden können. Ohne derartige Metadaten kann zwar prinzipiell eine Zeitreihe erstellt werden, diese kann aber in Hinblick auf ihren Aussagegehalt über tatsächliche ¹³⁷Cs-Konzentrationen nicht fundiert interpretiert werden.

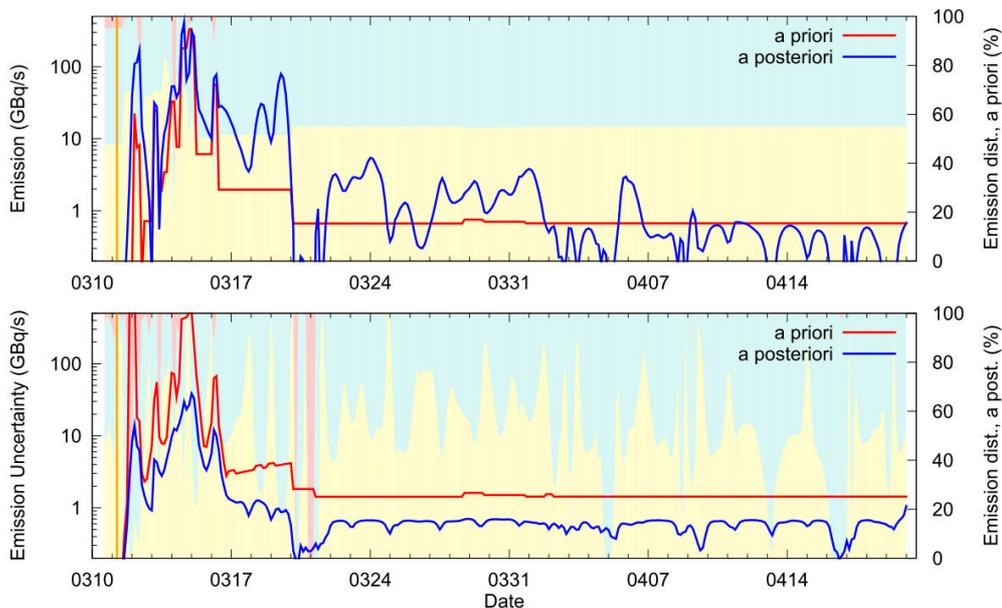


Abb. 2: Zeitreihe (oben) und assoziierte Unsicherheit (unten) der ¹³⁷Cs-Emissionen. Die rote Linie kennzeichnet die durch den First Guess a priori rekonstruierten Emissionen, die blaue Linie bezeichnet die a posteriori durch den Inversionsalgorithmus erhaltenen Emissionen. Die Hintergrundfarben gelb (0-50m), türkis (50-300m) und hellrot (300-1000m) kennzeichnen die vertikale Verteilung der Emissionen in den drei Layers (für die a priori Emissionen oben, für die a posteriori Emissionen unten). Die orange vertikale Linie kennzeichnet den Zeitpunkt des Erdbebens (Stohl et al. 2012, 2326).

³⁵ Die Tests auf Standardnormalverteilung, welche in Kapitel 5.2.5 erläutert werden, schließen eine Standard-Gauß-Verteilung aus.

5.2.3 Boxplots

Die Boxplots (Abb. 3) zeigen, dass die Mittelwerte sowie oberes und unteres Quartil bei beiden Datensätzen beinahe identisch sind. Der untere Whisker ist jedoch bei den modellierten Werten viel ausgeprägter. Es gibt zudem zahlreiche Ausreißer in beiden Datensätzen, die Ausreißer im Datensatz der modellierten Werte sind jedoch etwas größer.

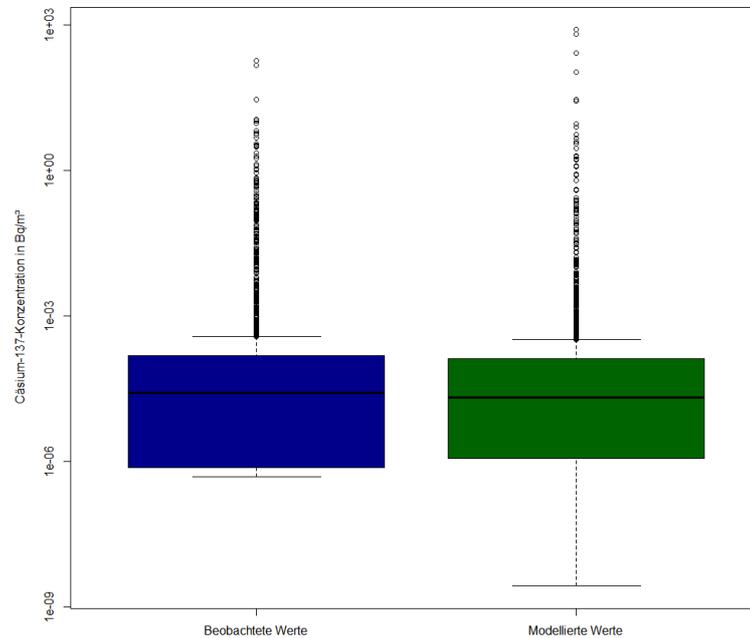


Abb. 3: Boxplots der beobachteten und modellierten ^{137}Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).

5.2.4 Residuenplot

Als Residuen werden die Schätzfehler bezeichnet, die in einem Modell auftreten. Residuen errechnen sich aus der Differenz von beobachtetem Wert und Schätzwert (Formel 5.1).

$$e_i = x_i^o - x_i^m \quad (5.1)$$

Abb. 4 zeigt die kumulierte Häufigkeit der Residuen des Modells. Die Verteilung der Fehler zeigt eindeutig, dass das Modell verzerrt ist und die Wirklichkeit überschätzt.

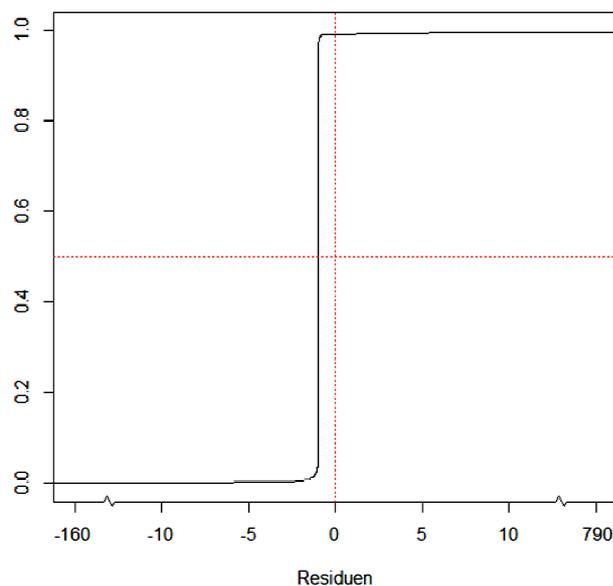


Abb. 4: Plot der kumulierten Häufigkeit der geordneten Residuen (eigene Darstellung; Datenquelle: Stohl et al. 2012).

5.2.5 Statistische Tests

Für die Wertepaare aus beobachteten und modellierten Werten sind einige grundlegende statistische Tests durchgeführt worden, mit welchen die Übereinstimmung der beiden Datensätze überprüft werden kann.

Der *Kolmogorov-Smirnov Test* überprüft die Nullhypothese gleicher Wahrscheinlichkeitsverteilungen ($H_0: F(x_m) = F(x_o)$) in zwei Datensätzen. Da der p-Wert (p-value = $1,11 \times 10^{-15}$) kleiner als das gewählte Risiko erster Art ($\alpha = 0,05$) ist, ist die Nullhypothese abzulehnen, die Wahrscheinlichkeitsverteilung ist also nicht gleich. Die beiden Datensätze sind also auch nicht normalverteilt. Dasselbe eindeutige Ergebnis liefern der *Shapiro-Wilk-Test* und der *Lilliefors-Test* (beide mit einem p-value von $2,2 \times 10^{-16}$).

Der *Welch-Test* überprüft die Nullhypothese gleicher Mittelwerte ($H_0: \mu_m = \mu_o$) zweier unabhängiger Datensätze. Diese Nullhypothese ist mit einem p-Wert von 0,1738 bei einem Risiko erster Art von $\alpha = 0,05$ anzunehmen, die Mittelwerte sind also einander gleich. Der *Zweistichproben T-Test* liefert mit einem p-Wert von 0,1512 für gepaarte Stichproben und mit einem p-Wert von 0,1737 für unabhängige Stichproben dasselbe Ergebnis.

Der *Wilcoxon-Vorzeichen-Rang-Test* ist ein nicht-parametrischer Test, welcher zur Überprüfung der Lage symmetrischer Merkmale verwendet wird (Strelec et al. 2008, 148). Mit Hilfe des Tests wird überprüft, ob zwei Stichproben ähnlich große Werte aufweisen. Die Nullhypothese der gleichen zentralen Tendenzen muss mit einem p-Wert von 0,0147 verworfen werden, die Verteilung ist also nicht gleich.

Eine *Varianzanalyse (einfache ANOVA)* kann nicht durchgeführt werden, da die Voraussetzungen der Homoskedastizität und der Normalverteilung nicht gegeben sind.

5.3 Vertiefende Analyse

Um den Einfluss der High-Leverage-Points³⁶, also jener Punkte mit einer großen Hebelwirkung auf die Regressionsgerade und somit auf das Modell, auszuschalten, können so genannte robuste Verfahren, welche unempfindlicher gegenüber Ausreißern sind, angewendet werden. Zu den robusten Verfahren zählen etwa *LMS (least median squares)*, *LTS (least trimmed squares)*, die Anwendung des *Tukey outlier filter* oder *bootstrapping*.

Exemplarisch für die robusten Verfahren wird im Folgenden das gestutzte bzw. getrimmte Mittel³⁷ (engl. *truncated* oder *trimmed mean*) verwendet. Hierfür sind die wichtigsten Parameter mit den um 1% gestutzten Werten nochmals berechnet worden (Tab. 3).

Dabei zeigt dabei ein verblüffendes Ergebnis. Bereits eine Stutzung des Datensatzes um 1% ergibt eine signifikant bessere Modellqualität. Das Bestimmtheitsmaß konstatiert mit knapp 97% eine fast perfekte Übereinstimmung, der Wert für den Bias ist sehr klein und impliziert eine nur geringe Verzerrung des Modells. Die Fehlerwerte *RMSE* sowie *MAE* zeugen von lediglich minimalen Abweichungen zwischen Modell und Wirklichkeit. Daraus ist zu schließen, dass einige wenige High-Leverage-Points eine extrem große Hebelwirkung auf das Modell ausgeübt haben.

Bei näherer Betrachtung der eliminierten Werte ist festzustellen, dass bei Betrachtung des Gesamtdatensatzes die größten Abweichungen (i.e. die größten absoluten Residuen) bei hohen Konzentrationswerten auftreten.

³⁶ Leverage ist ein Ausdruck für die Hebelwirkung von einzelnen Werten auf das Modell und stellt somit ein Maß für die Extremität (Ausreißer) der Fälle hinsichtlich der unabhängigen Variablen dar. Siehe dazu auch Anhang V respektive Abb. 5.

³⁷ Um einen getrimmten Datensatz zu erhalten, wird ein definierter Prozentsatz der größten und kleinsten Werte des geordneten Datensatzes abgeschnitten. Mit den restlichen Werten werden die Berechnungen durchgeführt.

Tab. 3: Übersicht über getrimmte statistische Kennwerte der korrigierten beobachteten und modellierten Werte des ^{137}Cs -Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).

	Parameter der korrigierten Originaldaten [Bq/m^3]		Parameter der korrigierten Originaldaten mit 1% Trim [Bq/m^3]	
	Beobachtete Werte	Modellierte Werte	Beobachtete Werte	Modellierte Werte
μ	0,3140	1,3036	0,0237	0,0134
	Parameter der korrigierten Originaldaten [1]		Parameter der korrigierten Originaldaten mit 1% Trim [1]	
Bias	0,9896		-0,0103	
RMSE	26,6467		0,1730	
RMSE_{BC}	26,6283		0,1727	
MAE	1,2735		0,0205	
R²	0,0568		0,9686	

6 Diskussion

In dieser Arbeit ist gezeigt worden, dass es eine Vielzahl von statistischen Kenngrößen, Parametern, Skill Scores und Fehlermaßen gibt, welche für die Beurteilung der Qualität von atmosphärischen Transportmodellen herangezogen werden können. Die aussagekräftigsten Grundparameter sind dabei Bias, Bestimmtheitsmaß und *RMSE*. Die weiterführenden Fehlermaße wie *MAE*, Korrelations- und Rangkorrelationskoeffizienten sowie die diversen Derivate wie Fractional Bias oder *NMSE* können jedoch mit einem – gemessen an ihrem Aussagegehalt – vertretbaren Mehraufwand berechnet und individuell – je nach Aufgabenstellung oder Fokus des Modells – zusätzlich zu Interpretationszwecken verwendet werden. Bei der Verwendung dieser Parameter für Analysezwecke ist dabei jedoch mit Vorsicht vorzugehen. Dabei ist der Einfluss der Modellperformance auf die Ausprägung der (normalisierten) Parameter zu bedenken.

Obwohl es mit den in dieser Arbeit vorgestellten Parametern methodische Grundlagen für eine standardisierte Herangehensweise an die Evaluierung atmosphärischer Dispersionsmodelle gibt, sollte die Interpretation dieser Skill Scores also situationsabhängig und individuell für jedes Modell erfolgen. Jeder Fall ist also – zumindest bis zu einem bestimmten Grad – gesondert zu betrachten.

Obzwar die statistischen Parameter im untersuchten Fall bei Betrachtung des Originaldatensatzes auf Basis der ausgewählten Parameter eine eher schlechte Übereinstimmung von Modell und Wirklichkeit konstatieren, relativiert sich dieser Umstand jedoch etwas, wenn man den Hintergrund der Daten näher betrachtet. Die Berücksichtigung der Komplexität atmosphärischer Transportmodellierung im Allgemeinen und der suboptimalen Datengrundlage in diesem Fall im Speziellen ist in die Bewertung des Modells mit einzubeziehen. Im konkreten Fall der Evaluierung des Dispersionsmodells für die ^{137}Cs -Emissionen von Stohl et al. (2012) sind dabei die die Messdaten betreffenden Einschränkungen zu berücksichtigen. Die teilweise unter der Detektionsschwelle liegenden Messwerte der Messstationen, welche sich in großer räumlicher Entfernung zu Japan befinden, sowie die hohe Sensitivität der modellierten Werte erhöhen die Modellunsicherheit. Hinzu kommt im vorliegenden Fall der Umstand, dass die Informationspolitik aufgrund der heiklen Materie sowohl seitens der Betreiberfirma des KKW Fukushima Dai-ichi, TEPCO, als auch seitens der japanischen Behörden sowie der IAEO suboptimal angelegt gewesen ist, wodurch auch die Beschaffung von Konzentrationswerten der Ortsdosisleistung in Japan erschwert worden ist. Darüber hinaus ist auch die durch Erdbeben und Tsunami bedingte kapitale Zerstörung der Infrastruktur in Japan zu bedenken, wovon auch Messstationen und Kommunikationsnetzwerke betroffen gewesen sind.

Die teilweise erheblichen Abweichungen zwischen gemessenen und modellierten Werten führen Stohl et al. (2012, 2331) auf die Resuspension (i.e. auf Aufwirbelungen) von ^{137}Cs von kontaminierten Oberflächen, welche in dem Modell nicht berücksichtigt wird, sowie auf die Kontamination der Detektoren zurück. Daneben stellt das Ersetzen der Werte unter der Detektionsschwelle auch einen gewissen Unsicherheitsfaktor dar.

Auch unerwartete lokale Wetterbedingungen können zudem einen Einfluss auf die Modellperformance haben. Derartige atmosphärische Transportmodelle sind generell durch den hohen Grad an Dynamik geprägt, welcher der Meteorologie eigen ist. Daher geht eine Vielzahl an Aspekten, Einflussfaktoren, Kenngrößen und zu berücksichtigenden Parametern in das Formelwerk zur Modellierung ein, um ein möglichst realitätsnahes Ergebnis zu liefern. Trotz aller Bemühungen kann eine Modellierung daher nicht perfekt sein, da jede atmosphärische Situation – bedingt durch die verschiedenen Faktoren, welche Auswirkungen auf die Atmosphäre haben und diese beeinflussen – einzigartig ist.

Durch die Anwendung robuster Verfahren zeigt sich im Rahmen der Evaluierung allerdings eine signifikant bessere Modellperformance. Dies lässt Rückschlüsse darauf zu, dass einige High-Leverage-Points großen Einfluss auf die Modellgüte ausüben. Nach der Eliminierung dieser Ausreißer steigt die Modellgüte unerwartet stark an, was den enormen Einfluss dieser Extremwerte unterstreicht. Obwohl diese Ausreißer einzelne Ausnahmefälle darstellen, sind

sie doch Teil des zu evaluierenden Datensatzes und auch als solche zu interpretieren. Dabei ist festzuhalten, dass solche High-Leverage-Punkte aufgrund verschiedener Aspekte zu Stande kommen können. So kann es sich etwa dabei um Messfehler handeln, welche durch die oben bereits angesprochene Kontamination der Detektoren oder durch Resuspension bedingt sein können. Daneben können diese extremen Abweichungen möglicherweise auf Fehler in den Input-Daten für das Modell zurückgeführt werden.

Allerdings ist auch festzuhalten, dass diese Skill Scores nicht notwendigerweise endgültig valide Parameter darstellen. Die Variabilität im Datensatz – sowohl bezüglich Ort als auch Zeit – stellt nämlich einen großen Unsicherheitsfaktor für die Ausbreitungsrechnung dar. Bereits eine kleine Abweichung im Rahmen der Modellierung, welche eine Schadstoffwolke nicht zur exakt richtigen Zeit am exakt richtigen Ort errechnet, kann mitunter eine erhebliche Verschlechterung der Parameter mit sich bringen.

Für eine Evaluierung des untersuchten Schadstoffausbreitungsmodells hinsichtlich der Verwendbarkeit der Ergebnisse für Katastrophenmanagement (z.B. Evakuierungen) können die durchgeführten Analysen aus Sicht des Autors nicht herangezogen werden. Für eine derartige Interpretation müsste die Modellperformance für große Konzentrationswerte untersucht werden. Im vorliegenden Fall ist dies allerdings nicht möglich, da die hohen Konzentrationswerte die größten Residuen aufweisen und somit oftmals Ausreißer darstellen. Zudem enthält der Datensatz der beobachteten Werte ohnehin lediglich fünf Werte mit einer Konzentration von über 10 Bq/m^3 . Die Werte des Datensatzes liegen damit im Rahmen der Aktivitätskonzentrationen natürlicher radioaktiver Stoffe in der Luft (KIT 2011a, 2). Da somit keine repräsentativen Daten über hohe Konzentrationswerte vorliegen, kann dieses Modell nicht adäquat hinsichtlich seiner Verwendbarkeit im Bereich des Zivilschutzes beurteilt werden. Darüber hinaus ist festzuhalten, dass das IMS der CTBTO prinzipiell zur Überwachung der Einhaltung des Kernwaffenteststopp-Vertrages dienen sollen. In dieser Funktion sind die Messstationen darauf ausgelegt, kleinste Spuren künstlicher Radionuklide nachzuweisen, weshalb die Messdaten aufgrund der notwendigen zeitintensiven Messverfahren mit einer zeitlichen Latenz von etwa drei Tagen nach Ende des Probenahmezeitraums vorliegen (BFS 2012, s.p.). Da diese Messstationen somit nicht auf einen raschen Nachweis von Radionukliden in der Umwelt ausgerichtet sind, erscheint eine Anwendung eines derartigen Transportmodells für den Ernstfall aufgrund der Zeitverzögerung von mehreren Tagen nur bedingt praktikabel. Zudem weisen Koohkan et al. (2012) in ihrer Studie darauf hin, dass die Verwendung der CTBTO-Daten für inverse Modellierung – und somit auch für die Anwendung von Dispersionsmodellen wie FLEXPART – aufgrund der unterschiedlichen Verteilung der IMS-Messstationen für unterschiedliche Regionen der Welt eine unterschiedlich gute Performance der Modelle liefert.

Die Methodik der Dispersionsmodellierung ist allerdings überaus vielfältig. In der vorliegenden Studie ist versucht worden, ein globales Optimum zu erreichen, wobei eine teilweise schlechtere Anpassungsgüte für einzelne Regionen zu Gunsten einer global optimalen Modellierung in Kauf genommen worden sind. Es besteht nun die Möglichkeit, die Ausbreitungsrechnungen im Detail an kleinräumigere Strukturen anzupassen, wodurch sich bei überbestimmten Regionen eine weitaus bessere Anpassungsgüte erreichen lässt.

Zusammenfassend ist festzuhalten, dass die Performance des FLEXPART-Modells im untersuchten Fall – trotz einer Überschätzung der Wirklichkeit, eines geringen Bestimmtheitsmaßes und eines großen *RMSE* – als durchaus gut zu bewerten ist, da die schlechte Modellgüte des Ausgangsdatsatzes auf einige wenige Extremwerte zurückzuführen ist. Der grundsätzliche Trend des Konzentrationsverlaufes wird allerdings gut wiedergegeben. Bedenkt man zudem die oben erwähnte Dynamik und Komplexität der Materie sowie die suboptimale Datengrundlage, ist das Modell ein durchaus potentes Werkzeug für die Simulation der Ausbreitung von Schadstoffen in der Atmosphäre.

7 Danksagung

Besonderer Dank gebührt Frau Univ.Doz. Dr.phil. Petra Seibert für die große Kooperationsbereitschaft, für die Zurverfügungstellung des Datenmaterials sowie für die fachliche Unterstützung. Ebenso großer Dank gebührt Herrn Assoc.Prof. Dipl.-Ing. Dr.techn. Gregor Laaha für den fachlichen Input und die freundliche Unterstützung beim Erstellen der Arbeit. Bedanken möchte ich mich auch bei Frau Dr.nat.techn. Kerstin Scherr, welche diese Arbeit begleitet und betreut hat.

8 Verzeichnisse

8.1 Literaturverzeichnis

- Albrecht, P., 2011. Stichwort: Gesetz der großen Zahlen. In: Wagner, F., Hrsg. Gabler Versicherungslexikon. Wiesbaden: Gabler Verlag. 264.
- Assenmacher, W., 2002. Einführung in die Ökonometrie. München; Wien: Oldenbourg Wissenschaftsverlag.
- de Bellevue, J. L., 2006. The Lagrangian Particle Dispersion Model FLEXPART. Presentation to UKZN Training Session, 09 November 2006.
- Bieringer, J., Weiss, W., Alberts, W., Burgkhardt, B., Hecker, O., Matzke, M., Nolte, G., Rimpler, A., Rühle, H., Salfeld, C., Scheib, H. und Waltermann, R., 2000. Grundlagen und Hinweise zur Messung von Ortsdosis und Ortsdosisleistung. Messanleitungen für die Überwachung radioaktiver Stoffe in der Umwelt und externer Strahlung. Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU). München – Jena: Urban & Fischer.
- BFS (Bundesamt für Strahlenschutz), 2012. Spurenanalyse Weltweit. Salzgitter. Verfügbar in: http://www.bfs.de/de/ion/imis/spurenmessungen_weltweit.html/ [Abfrage am 13. Juni 2012].
- Ceranna, L. und Roß, O., 2011. Ausbreitungssimulationen von Radionukliden emittiert durch den Reaktorunfall in Fukushima, Japan. Hannover: Bundesanstalt für Geowissenschaften und Rohstoffe. Verfügbar in: http://www.bgr.bund.de/DE/Themen/Erdbeben-Gefaehrdungsanalysen/Seismologie/Kernwaffenteststopp/Verifikation/Atmosphaer-Transport/Besondere%20Ereignisse/atm_fukushima_inhalt.html [Abfrage am 14. Juni 2012].
- Brosius, F., 1998. SPSS 8. Professionelle Statistik unter Windows. Hamburg: International Thomson Publishing.
- Bureau of Meteorology, 2012. Seasonal Streamflow Forecasts – Glossary. Stichwort: Skill Score. Melbourne. Verfügbar in: <http://www.bom.gov.au/water/ssf/glossary.shtml#item-skill-score> [Abfrage am 13. April 2012].
- CTBTO (Comprehensive Nuclear-Test-Ban-Treaty Organization), 2011. Fukushima-related Measurements by the CTBTO. Wien. Verfügbar in: <http://www.ctbto.org/?id=3234> [Abfrage am 13. Juni 2012]
- CTBTO (Comprehensive Nuclear-Test-Ban-Treaty Organization), 2012a. Who We Are. Wien. Verfügbar in: <http://www.ctbto.org/specials/who-we-are/> [Abfrage am 2. April 2012].
- CTBTO (Comprehensive Nuclear-Test-Ban-Treaty Organization), 2012b. International Monitoring System. Wien. Verfügbar in: <http://www.ctbto.org/map/#ims> [Abfrage am 2. April 2012].
- CTBTO (Comprehensive Nuclear-Test-Ban-Treaty Organization), 2012c. CTBTO Tracks Fukushima's Radioactive Release. Wien. Verfügbar in: <http://www.ctbto.org/verification-regime/the-11-march-japan-disaster/?Fsize=kuyzyweqhcib1%20or%201%3D%40%40version--#play/9b7PwKraek> [Abfrage am 13. Juni 2012].
- Deutscher Wetterdienst, 2012. Datenassimilation. Offenbach. Verfügbar in: http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_aufgabenspektrum_vorhersagedienst&_state=maximized&_windowLabel=T20408456401154670743272 [Abfrage am 1. April 2012].
- Doty, K., Tesche, T. W., McNally, D. E., Timin, B. und Mueller, S. F., 2002. Meteorological Modeling for the Southern Appalachian Mountains Initiative (SAMI). Final Report. Muscle Shoals.
- Eckhardt, S., Forster, C., James, P., Spichtinger-Rakowsky, N. und Stohl, A., 2005. Atmosphärischer Ferntransport und seine Auswirkungen auf die Spurengaskonzentrationen in der freien Troposphäre über Mitteleuropa (ATMOFAST) – Das Teilvorhaben „Transportmodellierung“. Endbericht. Technische Universität München, Lehrstuhl für Ökologiklimatologie. Freising-Weihenstephan.
- EPA, 1984. Interim Procedures for Evaluating Air Quality Models. United States Environmental Protection Agency. Monitoring and Data Analysis Division. Office of Air Quality Planning and Standards. Washington, D.C.
- Forster, C., Wandering, U., Wotawa, G., James, P., Mattis, I., Althausen, D., Simmonds, P., O'Doherty, S., Kleefeld, C., Jennings, S. G., Schneider, J., Trickl, T., Kreipl, S., Jäger, H. und Stohl,

- A., 2001. Transport of Boreal Forest Fire Emissions from Canada to Europe. *Journal of Geophysical Research*, 106, 22887–22906.
- Fournier, N., Weston, K. J., Dore, A. J. und Sutton M. A., 2005. Modelling the Wet Deposition of Reduced Nitrogen over the British Isles Using a Lagrangian Multi-Layer Atmospheric Transport Model. *Quarterly Journal of the Royal Meteorological Society*, 131, 703–722.
- Garger, E.K., Kuzmenko, Y., Sickinger, S., Tschiersch, J., 2012. Prediction of the ¹³⁷Cs activity concentration in the atmospheric surface layer of the Chernobyl exclusion zone. *Journal of Environmental Radioactivity*, 110, 53–58.
- Hackl, P., 2008. Einführung in die Ökonometrie. München: Pearson Studium.
- Holland, H., 2012. Stichwort: Vektor. Wiesbaden. Verfügbar in: <http://wirtschaftslexikon.gabler.de/Definition/vektor.html> [Abfrage am 13. Juni 2012].
- Hoppe, A., 2012. Das interaktive Periodensystem der Elemente (PSE). Sichtwort: Caesium. Verfügbar in: <http://www.periodensystem.info/elemente/caesium/> [Abfrage am 08. Juni 2012].
- Hsu, S.-C., Huh, C.-A., Chan, C.-Y., Lin, S.-H., Lin, F.-J. und Liu, S. C., 2012. Hemispheric Dispersion of Radioactive Plume Laced with Fission Nuclides from the Fukushima Nuclear Event. *Geophysical Research Letters*, 39, L00G22, doi:10.1029/2011GL049986.
- Jacobson, M. Z., 1999. *Fundamentals of Atmospheric Modeling*. Cambridge: Cambridge University Press.
- KIT (Karlsruher Institut für Technologie), 2011a. Hintergrundinformationen zu ausgewählten Themen zum nuklearen Störfall in Japan. Nr. 18: Welche Einheiten und Größen sind im Strahlenschutz wichtig, und was ist deren Bedeutung? Karlsruhe.
- KIT (Karlsruher Institut für Technologie), 2011b. Hintergrundinformationen zu ausgewählten Themen zum nuklearen Störfall in Japan. Nr. 31: Plutoniumkontamination beim japanischen Kernkraftwerk Fukushima I. Karlsruhe.
- Koohkan, M. R., Bocquet, M., Wu, L. und Krysta, M., 2012. Potential of the International Monitoring System Radionuclide Network for Inverse Modelling. *Atmospheric Environment* 54, 557–567.
- Leelössy, Á., Mészáros, R. und Lagzi, I., 2011. Short and Long Term Dispersion Patterns of Radionuclides in the Atmosphere around the Fukushima Nuclear Power Plant. *Journal of Environmental Radioactivity*, 102, 1117–1121.
- Lebanon, G., 2010. *Bias, Variance, and MSE of Estimators*. Georgia Institute of Technology. Atlanta.
- Leiner, B., 2004. Einführung in die Statistik. München: Oldenbourg Wissenschaftsverlag GmbH.
- Ludwig-Mayerhofer, W., 2012. ILMES – Internet-Lexikon der Methoden der empirischen Sozialforschung. Stichwort: Produkt-Moment-Korrelation. Siegen. Verfügbar in: http://www.lrz.de/~wlm/ilm_p4.htm [Abfrage am 11. April 2012].
- Mayer, D., 2006. Vergleichende Anwendung von Gauß- und Lagrange'schen Partikelmodell zur Ausbreitung von Viren. Diplomarbeit. Institut für Meteorologie und Geophysik, Universität Wien.
- Mosca, S., Graziani, G., Klug, W., Bellasio, R. und Bianconi, R., 1998a. A Statistical Methodology for the Evaluation of Long-range Atmospheric Dispersion Models: An Application to the ETEX Exercise. *Atmospheric Environment*, 32, 4307–4327.
- Mosca, S., Graziani, G., Klug, W., Bellasio, R. and Bianconi, R., 1998b. ATMES-II – Evaluation of Long-Range Dispersion models Using Data of the 1st ETEX Release. Appendix B – Models' Description. Office of Official Publications of the European Communities.
- Müller, P. und von Storch, H., 2004. *Computer Modelling in Atmospheric and Oceanic Sciences*. Berlin/Heidelberg: Springer Verlag.
- Müller, R., 2012. Radioaktivität und Strahlung. Die gefährlichen Isotope. Cäsium-137, Iod-131, Strontium-90. Liestal. Verfügbar in: <http://www.radioaktive-strahlung.org/radioaktivitaet/isotope.htm> [Abfrage am 08. Juni 2012].
- Murphy, A. H., 1988. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Monthly Weather Review*, 116, 2417–2424.
- Murphy, A. H. und Epstein, E. S., 1989. Skill Scores and Correlation Coefficients in Model Verification. *Monthly Weather Review*, 117, 572–582.

- NILU (Norwegian Institute for Air Research), Department of Atmospheric and Climate Research, 2010. FLEXPART. Flex* Publications – Partial listing of Publications for which FLEXTRA and/or FLEXPART were used. Kjeller. Verfügbar in: <http://transport.nilu.no/flexpart/flexpubs> [Abfrage am 1. April 2012].
- NILU (Norwegian Institute for Air Research), Department of Atmospheric and Climate Research, 2012. FLEXPART. Kjeller. Verfügbar in: <http://transport.nilu.no/flexpart> [Abfrage am 1. April 2012].
- Palmer, P., 2012. Inverse Modelling. School of GeoSciences. University of Edinburgh. Verfügbar in: <http://www.geos.ed.ac.uk/research/eochem/inv.html> [Abfrage am 29. März 2012].
- R, 2012. The R Project for Statistical Computing. Verfügbar in: <http://www.r-project.org/> [Abfrage am 2. April 2012].
- Reinmann, C., Filzmoser, P., Garrett, R. G. und Dutter, R., 2008. Statistical Data Analysis Explained. Applied Environmental Statistics with R. Chichester: John Wiley & Sons Ltd.
- Schöppner, M., Plastino, W., Povinec, P.P., Wotawa, G., Bella, F., Budano, A., de Vincenzi, M. und Ruggieri, F., 2011. Estimation of the Time-Dependent Radioactive Source-Term from the Fukushima Nuclear Power Plant Accident Using Atmospheric Transport Modeling. *Journal of Environmental Radioactivity*. doi:10.1016/j.jenvrad.2011.11.008.
- Seibert, P., 2000. Quell-Rezeptor-Beziehungen atmosphärischer Spurenbestandteile. Habilitationsschrift. Institut für Meteorologie und Physik, Universität für Bodenkultur Wien.
- Seibert, P., 2004. Remarks on Statistical Parameters for Model Evaluation. Wien.
- Shozugawa, K., Nogawab, N. und Matsuo, M., 2012. Deposition of Fission and Activation Products after the Fukushima Dai-ichi Nuclear Power Plant Accident. *Environmental Pollution*, 163, 243–247.
- Slørdal, L. H., 2002. Applying Model Calculations to Estimate Future Urban Air Quality with Respect to Requirements of the EU Directives on NO₂ and PM₁₀. In: Sportisse, B., Hrsg. *Air Pollution Modelling and Simulation*. Berlin/Heidelberg: Springer Verlag.
- Sperber, K. R., 1999. Are Revised Models Better Models? A Skill Score Assessment of Regional Interannual Variability. *Geophysical Research Letters*, 26, 1267–1270.
- Spichtinger, N., Wenig, M., James, P., Wagner, T., Platt, U., und Stohl, A., 2001. Satellite Detection of a Continental-Scale Plume of Nitrogen Oxides from Boreal Forest Fires. *Geophysical Research Letters*, 28, 4579–4582.
- Sportisse, B., 2002. *Air Pollution Modelling and Simulation*. Berlin/Heidelberg: Springer Verlag.
- Statista, 2012. Statista-Lexikon. Stichwort: Gesetz der großen Zahl. Hamburg: Statista. Verfügbar in: <http://de.statista.com/statistik/lexikon/definition/58/gesetz-der-grossen-zahl/> [Abfrage am 1. April 2012].
- Statistisches Bundesamt, 2012. Gesundheitsberichterstattung des Bundes. Glossar. Stichwort: Cäsium-137. Bonn. Verfügbar in: http://www.gbe-bund.de/glossar/Caesium_137.html [Abfrage am 08. Juni 2012]
- StatSoft, 2012. StatSoft Glossar zu STATISTICA. Stichwort: Kendalls Tau. Hamburg. Verfügbar in: <http://www.statsoft.de/glossary/K/KendallTau.htm> [Abfrage am 11. April 2012].
- Stockhause, M., 2000. Bestimmung von Emissionsraten diffuser Quellen mit Hilfe inverser Modellierung. Schriftenreihe des Fraunhofer-Instituts für Atmosphärische Umweltforschung, 65-2000. Garmisch-Partenkirchen: Shaker Verlag.
- Stohl, A., 1996. The FLEXPART Model – Performance Evaluation Using Tracer Data from the First ETEX Release. In: Mosca, S., Graziani, G., Klug, W., Bellasio, R. and Bianconi, R., 1998. *ATMES-II – Evaluation of Long-Range Dispersion models Using Data of the 1st ETEX Release*. Appendix B – Models' Description. Office of Official Publications of the European Communities.
- Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., und Seibert, P., 2002. A Replacement for Simple Back Trajectory Calculations in the Interpretation of Atmospheric Trace Substance Measurements. *Atmospheric Environment*, 36, 4635–4648.
- Stohl, A., Forster, C., Eckhardt, S., Spichtinger, N., Huntrieser, H., Heland, J., Schlager, H., Wilhelm, S., Arnold, F. und Cooper, O., 2003. A Backward Modeling Study of Intercontinental Pollution Transport Using Aircraft Measurements. *Journal of Geophysical Research*, 108, 4370.

- Stohl, A., Forster, C., Frank, A., Seibert, P. und Wotawa, G., 2005. Technical note: The Lagrangian Particle Dispersion Model FLEXPART, Version 6.2. *Atmospheric Chemistry and Physics*, 5, 2461–2474.
- Stohl, A., Hittenberger, M. und Wotawa, G., 1998. Validation of the Lagrangian Particle Dispersion Model FLEXPART against Large Scale Tracer Experiments. *Atmospheric Environment*, 32, 4245–4264.
- Stohl, A., Seibert, P., Wotawa, G., Arnold, D., Burkhardt, J. F., Eckhardt, S., Tapia, C., Vargas, A. und Yasunari, T. J., 2012. Xenon-133 and Caesium-137 Releases into the Atmosphere from the Fukushima Dai-ichi Nuclear Power Plant: Determination of the Source term, Atmospheric Dispersion, and Deposition. *Atmospheric Chemistry and Physics*, 12, 2313–2343.
- Stohl, A., und Thomson, D.J., 1999. A Density Correction for Lagrangian Particle Dispersion Models. *Boundary-Layer Meteorology*, 90, 155–167.
- Stohl, A. und Trickl, T., 1999. A textbook example of long-range transport: Simultaneous Observation of Ozone Maxima of Stratospheric and North American Origin in the Free Troposphere over Europe. *Journal of Geophysical Research*, 104, 30445–30462.
- Strelec, H., Rasch, D., Moder, K., Spangl, B. und Matiasch, L., 2008. *Statistik*. Wien: Universität für Bodenkultur – Department für Raum, Landschaft und Infrastruktur – Institut für Angewandte Statistik und EDV.
- Tagami, K., Uchida, S., Uchihori, Y., Ishii, N., Kitamura, H. und Shirakawa, Y., 2011. Specific Activity and Activity Ratios of Radionuclides in Soil Collected About 20 km from the Fukushima Daiichi Nuclear Power Plant: Radionuclide Release to the South and Southwest. *Science of the Total Environment*, 409, 4885–4888.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).
- Taylor, K. E., 2001. Summarizing Multiple Aspects of Model Performance in a Single Diagram. *Journal of Geophysical Research*, 106, 7183–7192.
- VNL, 2012. Stichwort: Disaggregation. Verein Netzwerk Logistik. Steyr. Verfügbar in: <http://www.vnl.at/Disaggregation.191.0.html> [Abfrage am 2. April 2012].
- Wergen, W., 2002. Datenassimilation – ein Überblick. *promet*, 27, 3, 142–149.
- Wetterstation Emden, 2009. *Wetterlexikon*. Stichwort: Grundströmung. Emden. Verfügbar in: <http://www.meteo-dynamics.de/wetterlexikon/atmosphaere/grundstroemung> [Abfrage am 17. Mai 2012].
- wissenschaft-online, 2012. *Lexikon der Geowissenschaften*. Stichwort: Ausbreitungsmodell. Heidelberg: Spektrum der Wissenschaft Verlagsgesellschaft mbH. Verfügbar in: <http://www.wissenschaft-online.de/abo/lexikon/geo/1185> [Abfrage am 30. März 2012].
- Yasunari, T.J., Stohl, A., Hayano, R.S., Burkhardt, J.F., Eckhardt, S. und Yasunari, T., 2011. Cesium-137 Deposition and Contamination of Japanese Soils due to the Fukushima Nuclear Accident. *Proceedings of the National Academy of Sciences*, 108, 19530–19534.
- Yu S., Eder B., Dennis R., Chu S.-H. und Schwartz S. E., 2006. New Unbiased Symmetric Metrics for Evaluation of Air Quality Models. *Atmospheric Science Letters*, 7, 26–34.
- ZAMG (Zentralanstalt für Meteorologie und Geodynamik), 2011. ZAMG Fukushima Web Page. Wien. Verfügbar in: <http://www.zamg.ac.at/wetter/fukushima/> [Abfrage am 14. Juni 2012].
- Zinevich, A., Alpert, B. und Messer, H., 2008. Estimation of rainfall fields using commercial microwave communication networks of variable density. *Advances in Water Resources*, 31, 1470–1480.

8.2 Abbildungsverzeichnis

Titelgraphik: Scatterplot der nicht korrigierten Originaldaten (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	Titelblatt
Abb. 1: Scatterplot der nicht korrigierten Originaldaten im Vergleich zum Scatterplot der korrigierten Originaldaten (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	23
Abb. 2: Zeitreihe (oben) und assoziierte Unsicherheit (unten) der ^{137}Cs -Emissionen. Die rote Linie kennzeichnet die durch den First Guess a priori rekonstruierten Emissionen, die blaue Linie bezeichnet die a posteriori durch den Inversionsalgorithmus erhaltenen Emissionen. Die Hintergrundfarben gelb (0-50m), türkis (50-300m) und hellrot (300-1000m) kennzeichnen die vertikale Verteilung der Emissionen in den drei Layers (für die a priori Emissionen oben, für die a posteriori Emissionen unten). Die orange vertikale Linie kennzeichnet den Zeitpunkt des Erdbebens (Stohl et al. 2012, 2326).....	24
Abb. 3: Boxplots der beobachteten und modellierten ^{137}Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	25
Abb. 4: Plot der kumulierten Häufigkeit der geordneten Residuen (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	25
Abb. 5: Diagnostic Plots für das Modell (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	42
Abb. 6: Geometrischer Zusammenhang zwischen Pearsons Korrelationskoeffizienten, der Wurzel aus der bias-korrigierten mittleren quadratischen Abweichung und den Standardabweichungen der modellierten und beobachteten Datensätze (verändert nach Taylor 2001, 7184).....	43
Abb. 7: Taylor-Diagramm für die ^{137}Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	43
Abb. 8: Taylor-Diagramm für die ^{137}Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	44

8.3 Tabellenverzeichnis

Tab. 1: Übersicht über statistische Kennwerte der nicht korrigierten sowie der korrigierten beobachteten und modellierten Werte des ^{137}Cs -Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	22
Tab. 2: Übersicht über statistische Kennwerte der korrigierten logarithmierten beobachteten und modellierten Werte des ^{137}Cs -Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	23
Tab. 3: Übersicht über getrimmte statistische Kennwerte der korrigierten beobachteten und modellierten Werte des ^{137}Cs -Ausstoßes in Fukushima Dai-ichi für den Zeitraum zwischen 11. März 2011 und 19. April 2011 (eigene Darstellung; Datenquelle: Stohl et al. 2012).....	27

8.4 Abkürzungsverzeichnis

ATMES II	Atmospheric Transport Model Evaluation Study II
Bq	Becquerel
BFS	Bundesamt für Strahlenschutz
BGR	Bundesanstalt für Geowissenschaften und Rohstoffe
CTBTO	Comprehensive Nuclear-Test-Ban Treaty Organization
FB	Fractional Bias
FMS	Figure of Merit in Space
FMT	Figure of Merit in Time
FOEX	Factor of Excedance
GFS	Global Forecast System (der Wetter- und Ozeanografiebehörde der USA)
IAEO	Internationale Atomenergie-Organisation
JCAC	Japan Chemical Analysis Center
JAEA	Japan Atomic Energy Agency
KKW	Kernkraftwerk
MEXT	Ministry of Education, Culture, Sports, Science & Technology in Japan.
MAE	Mean Absolute Error
MANGE	Mean Absolute Normalized Gross Error
MSE	Mean Squared Error
NIES	National Institute for Environmental Studies
NILU	Norsk Institutt for Luftforskning (Norwegian Institute for Air Research)
NGE	Normalized Gross Error
NMSE	Normalized Mean Squared Error
PPA	Paired Peak Estimation Accuracy
RMSE	Root Mean Squared Error
Ro5	Ring of Five
SPPA	Spatially Paired Peak Estimation Accuracy
TEPCO	Tokyo Electric Power Company
TPPA	Temporally Paired Peak Estimation Accuracy

8.5 Elementsymbole

Ag.....	Silber (Argentum)
Ba.....	Barium
Cs.....	Cäsium
Fe.....	Eisen
I.....	Iod
La.....	Lanthan
Mo.....	Molybdän
Nb.....	Niob
Np.....	Neptunium
Pu.....	Plutonium
Sr.....	Strontium
Te.....	Tellur
Xe.....	Xenon
Y.....	Yttrium
Zr.....	Zirconium

Anhang

Anhang I: Das Dispersionsmodell FLEXPART

Für die Modellierung der Ausbreitung der radioaktiven ^{137}Cs -Isotope ist das Lagrange'sche Partikelmodell FLEXPART verwendet worden (Stohl et al. 2012, 2321). FLEXPART ist ursprünglich zur Berechnung der weiträumigen, mesoskaligen Ausbreitung von durch eine Punktquelle emittierten Luftschadstoffen entwickelt worden und kann nunmehr nach einigen Weiterentwicklungen in de facto allen Bereichen, in denen atmosphärischer Transport eine Rolle spielt, eingesetzt werden. Der Einsatzbereich umfasst somit die Modellierung von Transport, Diffusion, Deposition (sowohl trockene Deposition als auch Nassablagerung) sowie den radioaktiven Zerfall von Tracern, welche von einer Punkt-, Linien-, Flächen- oder Volumenquelle emittiert werden (Stohl et al. 2005, 2461). Die Ausbreitung der Tracer wird mit Hilfe von Trajektorien bestimmt, welche für eine Vielzahl an Partikeln berechnet werden. Um Turbulenz und Konvektion zu simulieren, werden zusätzlich zu den von den Input-Daten beschriebenen Bewegungen auch stochastische Komponenten verwendet (Eckhardt et al. 2005, 87). Das Programm FLEXPART kann – analog zu der oben beschriebenen Vorgangsweise bei der Modellierung atmosphärischer Transportprozesse – in zwei Modi ausgeführt werden. So kann ausgehend von Messwerten durch inverse Modellierung ein Quellterm bestimmt werden, der dann wiederum in einer Vorwärtssimulation für die Ausbreitungsrechnung verwendet wird. Für die realitätsnahe Modellierung dieser Transportprozesse kann die Ausbreitung von Luftschadstoffen mit Hilfe von entsprechend aufbereiteten meteorologischen Daten zu einem gewünschten Zeitpunkt simuliert werden.

Das FLEXPART-Modell ist seit seiner Entwicklung in verschiedenen Bereichen der atmosphärischen Transportmodellierung zum Einsatz gekommen (NILU 2010). Dabei ist das Modell auch anhand diverser Anwendungsbeispiele (u.a. Stohl 1996, Stohl et al. 1998, Stohl und Trickl 1999, Forster et al. 2001, Spichtinger et al. 2001, Stohl et al. 2002, Stohl et al. 2003) evaluiert sowie weiterentwickelt, korrigiert und optimiert (u.a. Stohl und Thomson 1999, NILU 2012) worden. Aus den im Rahmen der *Atmospheric Transport Model Evaluation Study II* (ATMES II) von Mosca et al. (1998b) zwischen 49 Dispersionsmodellen durchgeführten Inter-Modell-Vergleichen geht FLEXPART als eines der besten atmosphärischen Transportmodelle hervor (Mosca et al. 1998b, 50ff; Stohl 1996, 463).

Aktuell liegt das Modell unter der GNU GPL V3.0 in Version 9 vor und wird von 35 Institutionen aus 17 Ländern eingesetzt (NILU 2012, s.p.).

Anhang II: Figure of Merit

Im Rahmen der statistischen Analysen bezeichnet *figure of merit* einen Koeffizienten, der dazu dient, ein System oder Modell in Bezug auf seine Alternativen zu charakterisieren. Im Rahmen der Modellvalidierung gibt es zwei gebräuchliche figures of merit, nämlich *Figure of Merit in Space* (FMS) und *Figure of Merit in Time* (FMT).

Die *Figure of Merit in Space* wird anhand der räumlichen Verteilung der Daten bestimmt. Sie ist definiert als der Anteil der Durchschnittsmenge der Flächen der gemessenen und modellierten Datensätze an der Vereinigungsmenge und ergibt sich aus dem Quotienten der Flächen von Durchschnittsmenge und Vereinigungsmenge (Formel I.1). Um die FMS bestimmen zu können, wird zu einer gegebenen Zeit ein Schwellenwert für die Konzentration der Partikel definiert. Alle Werte oberhalb dieses Schwellenwertes werden berücksichtigt.

$$FMS = 100 \frac{A_m \cap A_o}{A_m \cup A_o} \quad (I.1)$$

Je höher die FMS, desto besser ist die Performance des Modells. Da eine niedrige FMS jedoch auch durch die Verschiebung der Punktwolken um einen gewissen Offset begründet sein kann, empfiehlt es sich, für die Analyse auch eine (geo)graphische Darstellung der Flächen der gemessenen und der modellierten Daten heranzuziehen (Mosca et al. 1998a, 4313).

Analog zur FMT wird die *Figure of Merit in Time* anhand der zeitlichen Verteilung der Daten bestimmt. Dazu wird für einen gegebenen Ort L eine Zeitreihe (mit den Zeitpunkten t_i) der modellierten Werte und der Messwerte erstellt. Die FMT evaluiert nun für ein definiertes Zeitintervall die Schnittmenge der modellierten und gemessenen Konzentrationen an diesem bestimmten Ort. Dieser Trend wird mit einer Zeitreihe bestehend aus den Maximalkonzentrationen aus modellierten und gemessenen Werten normiert (Mosca et al. 1998a, 4313). Die mathematische Darstellung der FMT ist in Formel I.2 ersichtlich.

$$FMT = 100 \frac{\sum_{i=1}^n \min\{x^o(L, t_i), x^m(L, t_i)\}}{\sum_{i=1}^n \max\{x^o(L, t_i), x^m(L, t_i)\}} \quad (I.2)$$

Die Interpretation der FMT erfolgt analog zur FMS. Während hohe Werte auf eine hohe Modellgüte schließen lassen, empfiehlt sich bei niedriger FMT eine unterstützende qualitative Analyse mit Hilfe eines Konzentrations-Zeit-Diagramms.

Da bei FMS und FMT wiederum die Modellwerte in Form der Vereinigungsmenge in den Nenner eingehen, empfiehlt Seibert (2004, 3), für die Normierung ausschließlich die Fläche der beobachteten Werte zu verwenden, i.e. die Fläche der Vereinigungsmenge durch die Fläche der Messwerte zu ersetzen.

Anhang III: Peak Estimation Accuracy

Die *Peak Estimation Accuracy* gibt Auskunft darüber, wie gut ein Modell einen Maximalwert (*peak value*) in räumlicher und/oder zeitlicher Hinsicht simulieren kann. Die *Paired Peak Estimation Accuracy PPA* gibt nun an, wie gut das Modell den Maximalwert hinsichtlich *Zeit t* und *Ort L* des Maximalwertes vorhersagen kann (Formel II.1)

$$PPA = \frac{x_i^o(\hat{L}, \hat{t}) - x_i^m(\hat{L}, \hat{t})}{x_i^m(\hat{L}, \hat{t})} \quad (\text{II.1})$$

Das Zirkumflex (^) kennzeichnet dabei die Maximalwerte, welche zu einem *Zeitpunkt t* an einem *Ort L* auftreten. Für die Berechnung werden also jene modellierten Werte verwendet werden, welche zeitlich und räumlich mit dem beobachteten Maximalwert einhergehen (Jacobson 1999, 551).

Da der PPA-Wert sehr empfindlich auf zeitliche oder räumliche Verschiebungen reagiert, kann dieser Parameter gesondert als *Temporally Paired Peak Estimation Accuracy* TPPA bzw. als *Spatially Paired Peak Estimation Accuracy* SPPA angegeben werden (Doty et al 2002, 107).

Anhang IV: Weitere Skill Scores

Ausgehend von den eben dargestellten grundlegenden statistischen Parametern zur Modellvalidierung gibt es für die Evaluierung derartiger Modelle noch diverse weitere Fehlermaße. Derartige Parameter, welche verwendet werden können, um die Performance von Modellen zu bewerten, werden allgemein als *skill score* oder in der sinngemäßen Übersetzung als *Schätzgütemaß* bezeichnet.

Im Laufe der Zeit sind verschiedene Derivate von gängigen statistischen Parametern entwickelt worden, in welche diverse Korrekturen, Anpassungen oder Änderungen implementiert worden sind oder welche etwa durch verschiedene Normierungsverfahren charakterisiert sind (c.f. Murphy 1988, Murphy und Epstein 1989, Sperber 1999, Taylor 2001, Seibert 2004).

In den Geowissenschaften sind diverse, speziell auf das jeweilige Anwendungsgebiet zugeschnittene Skill Scores weit verbreitet. In der Meteorologie finden sich so etwa Schätzgütemaße, welche in Formel III.1 in einer allgemeinen, abstrahierten Form angegeben sind. S_{mod} steht dabei für die Schätzgüte der Modellvorhersage, S_{ref} für die Schätzgüte einer Referenzvorhersage bzw. die Schätzgüte einer Standardvorhersage, und S_{perf} drückt die Schätzgüte einer perfekten Vorhersage aus (Bureau of Meteorology 2012, s.p.).

$$Skill\ Score = \frac{S_{mod} - S_{ref}}{S_{perf} - S_{ref}} \quad (III.1)$$

Ein konkretes Beispiel für ein derartiges Schätzgütemaß in der Meteorologie ist etwa der *Brier skill score* (Müller und von Storch 2004, 265).

Auch für die Evaluierung von Schadstoffausbreitungsmodellen können prinzipiell noch eine große Zahl an weiteren Schätzgütemaßen (wie etwa jene von Taylor 2001, 7190f oder Seibert 2004, 6) oder an diversen Derivaten herangezogen werden. Um den Rahmen nicht zu sprengen, werden diese adaptierten Skill Scores in dieser Arbeit jedoch nicht weiter betrachtet, zumal ohnehin zu bezweifeln ist, dass diese zunehmend komplexeren Parameter signifikant bessere Modellvalidierungen ermöglichen als die gängigen und verbreiteten Parameter.

Anhang V: Diagnostic Plots

Abb. 5 zeigt die *Diagnostic Plots* für die modellierten und gemessenen ^{137}Cs -Werte.

In der Abbildung links oben (*Residuals vs Fitted*) sind die Residuen gegen die vorhergesagten Werte aufgetragen. Aus dieser Abbildung kann man ablesen, ob Homoskedastizität (Varianzgleichheit) oder Bias vorliegen. Die ungleiche Streuung der Punkte um die Gerade impliziert jedoch Heteroskedastizität. Zudem erkennt man einen systematischen Trend.

Im *Normal Q-Q-Plot (Normal-Quantile-Quantile-Plot)* sind auf der y-Achse die standardisierten Residuen und auf der x-Achse die bei Normalverteilung erwartete Werte aufgetragen. Dies gibt Auskunft darüber, ob die standardisierten Residuen normalverteilt sind. Bei Normalverteilung sollten alle Werte möglichst auf der Geraden liegen. Während die Gerade größtenteils sehr gut getroffen wird, sprechen die doch sehr starken Abweichungen an den Enden gegen eine Normalverteilung.

Der *Scale Location-Plot* links unten zeigt auf der Ordinatenachse die Wurzel aus dem Betrag der standardisierten Residuen und auf der Abszissenachse die vorhergesagten Werte. Die rote Linie zeigt, dass die Varianz der Residuen nicht homogen ist, da sich ein auffälliger Trend nach oben abzeichnet, welcher eine Schiefe der Verteilung (*skewness*) impliziert.

Der *Residuals vs. Leverage-Plot* stellt schließlich die standardisierten Residuen (Ordinate) gegen Leverage (Abszisse) dar. Leverage ist dabei ein Ausdruck für die Hebelwirkung von einzelnen Werten auf das Modell und stellt somit ein Maß für die Extremität (Ausreißer) der Fälle hinsichtlich der unabhängigen Variablen dar. In diesem Plot ist zudem auch noch *Cook's Distanz* dargestellt, welche ebenso ein Maß für den Einfluss eines Datenpunktes auf die Regressionsgerade darstellt, indem sie Rückschlüsse darüber zulässt, wie sehr ein einzelner Datenpunkt die Koeffizienten der Regressionsgleichung beeinflusst. Alle Datenpunkte, die außerhalb der Linien für *Cook's distance* in dem Diagramm liegen, überschreiten den kritischen Wert für einen extremen Einfluss auf die Regressionsgerade.

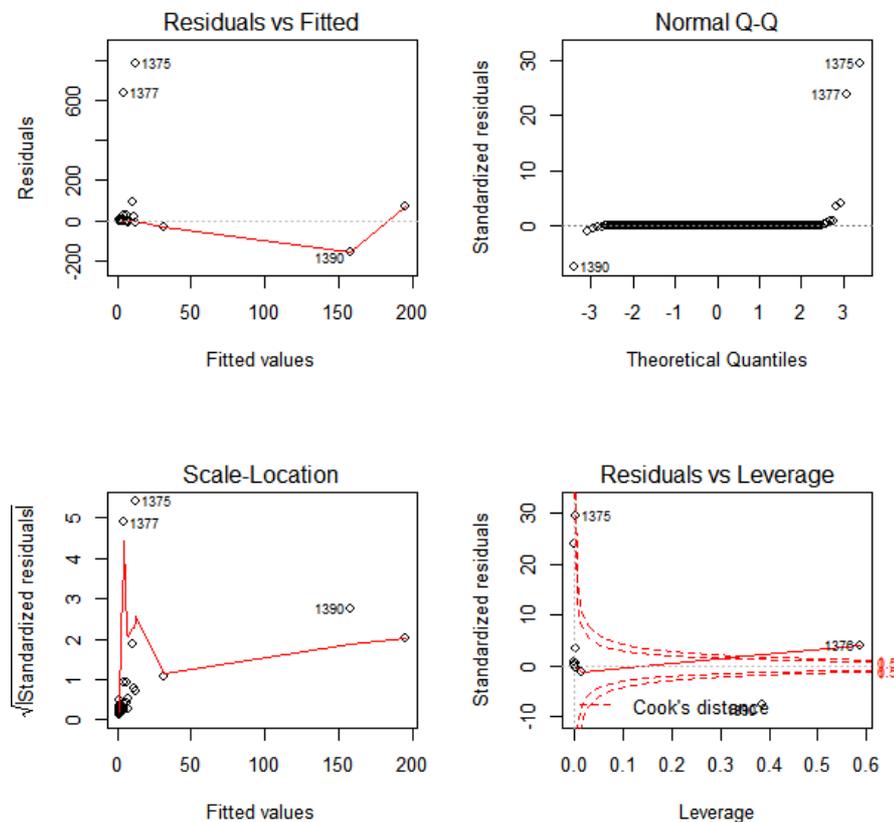


Abb. 5: Diagnostic Plots für das Modell (eigene Darstellung; Datenquelle: Stohl et al. 2012).

Anhang VI: Taylor-Diagramm

Taylor (2001, 7183ff) hat einen Zusammenhang zwischen dem Korrelationskoeffizienten r , dem $RMSE_{BC}$ und den Standardabweichungen der Messwerte σ_{x^o} und modellierten Werte σ_{x^m} aufgezeigt und beschrieben (Formel V.1).

$$RMSE_{BC}^2 = \sigma_{x^o}^2 + \sigma_{x^m}^2 - 2\sigma_{x^o}\sigma_{x^m}r \quad (V.1)$$

Der geometrische Zusammenhang zwischen diesen Werten ist in Abb. 6 dargestellt.

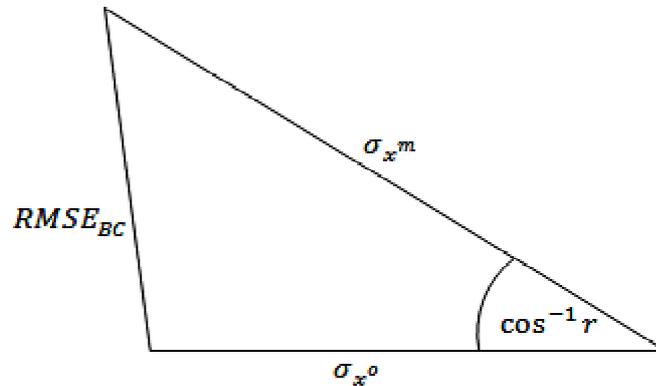


Abb. 6: Geometrischer Zusammenhang zwischen Pearsons Korrelationskoeffizienten, der Wurzel aus der bias-korrigierten mittleren quadratischen Abweichung und den Standardabweichungen der modellierten und beobachteten Datensätze (verändert nach Taylor 2001, 7184).

Das Ergebnis dieser Überlegungen ist das sogenannte Taylor-Diagramm. Dabei handelt es sich um einen Graphen, welcher diese Parameter für die Performance eines Modells zusammenfasst bzw. ausdrückt. Es handelt sich dabei um ein Polardiagramm, bei welchem die Radialkoordinate die Standardabweichung angibt und die Winkelkoordinate φ durch $r = \cos \varphi$ definiert ist. In dieses Diagramm wird nun die Beobachtung³⁸ bei $\varphi = 0$ und Radius σ_{x^o} geplottet, während das Modell für $\varphi = r$ und Radius σ_{x^m} geplottet wird. Die Distanz zwischen den beiden Punkten stellt somit den $RMSE_{BC}$ dar (Seibert 2004, 5).

Abb. 7 zeigt ein einfaches Taylor-Diagramm für die mit FLEXPART modellierten und die gemessenen Caesium-137 Werte sowie die Werte des First Guess.

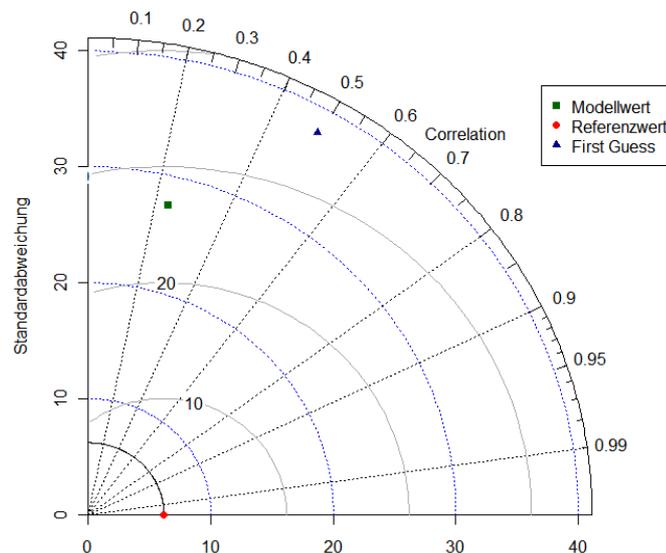


Abb. 7: Taylor-Diagramm für die ¹³⁷Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).

³⁸ Taylor (2001, 7184) bezeichnet die Beobachtung in diesem Zusammenhang allgemein als *reference*.

Wie aus Abb. 7 ersichtlich ist, ist der Korrelationskoeffizient zwischen Modellwerten und Referenzwerten $r = 0,24$. Die Standardabweichung der modellierten Werte beträgt $\sigma_{x^m} = 27,42$, während die Standardabweichung der Messwerte $\sigma_{x^o} = 6,18$ ist. Der $RMSE_{BC}$ beträgt 26,63.

Abb. 8 zeigt ebenso ein Taylor-Diagramm für die mit FLEXPART modellierten und die gemessenen Caesium-137 Werte, allerdings ist dieses Diagramm um die negativen Korrelationswerte ergänzt.

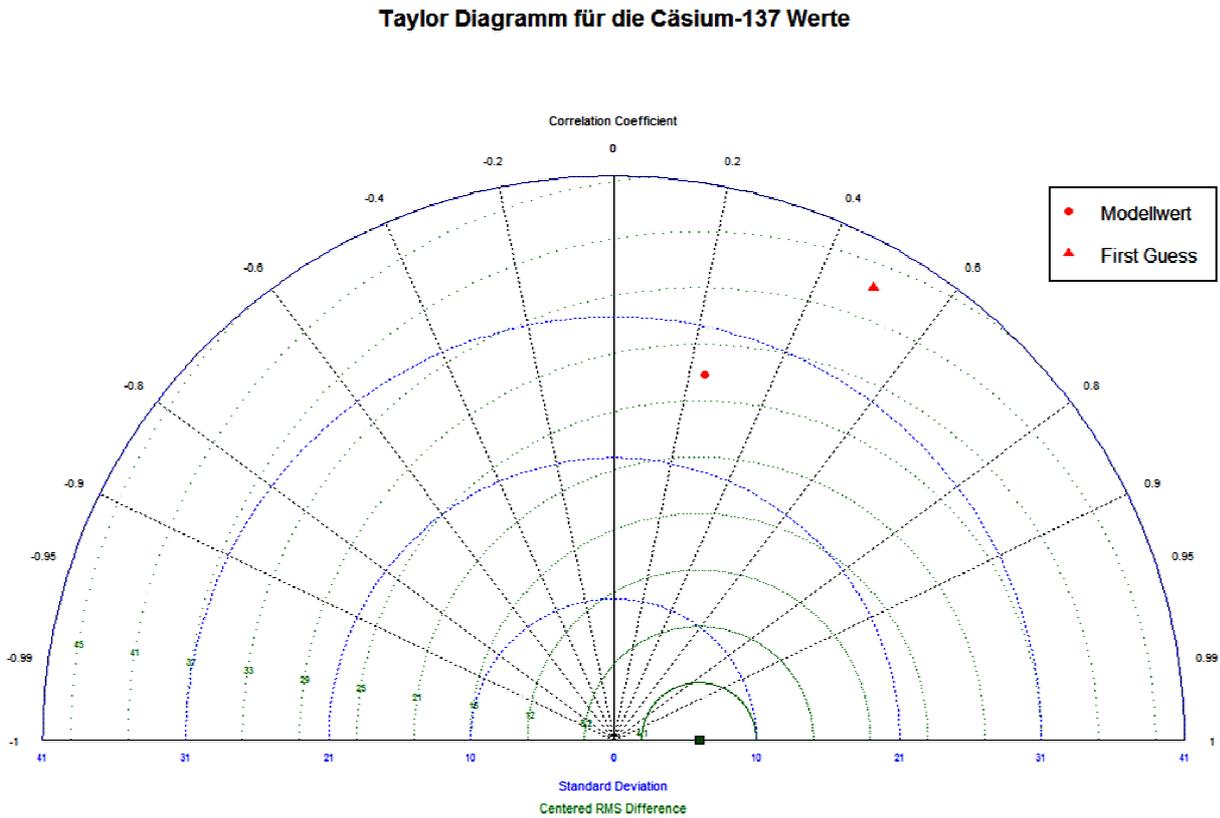


Abb. 8: Taylor-Diagramm für die ^{137}Cs Werte (eigene Darstellung; Datenquelle: Stohl et al. 2012).

Bisher erschienen in der Reihe BOKU-Met Report:

- 1 Eitzinger, J., Kubu, G., Formayer, H., Haas, P., Gerersdorfer, T., Kromp-Kolb, H. (2009): **Auswirkungen einer Klimaänderung auf den Wasserhaushalt des Neusiedlersees.** (Endbericht im Auftrag der Burgenländischen Landesregierung vom 15. Juli 2005).
- 2 Frank, A., Seibert, P. (2009): **Diagnose von Extremereignissen aus großräumigen meteorologischen Feldern.** (Endbericht StartClim.4, November 2003).
- 3 Formayer, H., Matulla, C., Haas, P., Groll, N. (2009): **Statistische Downscalingverfahren zur Ableitung von Extremereignissen in Österreich aus GCM-Feldern.** (Endbericht StartClim.5, November 2003).
- 4 Schwarzl, I., Haas, W. (2009): **Kommunikation an der Schnittstelle Wissenschaft und Bildung.** (Endbericht StartClim.11, November 2003).
- 5 Formayer, H., Haas, P., Matulla, C., Frank, A., Seibert, P. (2009): **Untersuchungen regionaler Klimaänderungsszenarien hinsichtlich Hitze- und Trockenperioden in Österreich.** (Endbericht StartClim2004.B, November 2003).
- 6 Schwarzl, I., Lang, E. (2009): **"Hängen Hitze und Leistungsfähigkeit zusammen?"** Ein Projekt an der Schnittstelle Wissenschaft und Bildung. (Endbericht StartClim2004.G, Jänner 2005).
- 7 Formayer, H., Kromp-Kolb, H. (2009): **Hochwasser und Klimawandel. Auswirkungen des Klimawandels auf Hochwasserereignisse in Österreich.** (Endbericht WWF 2006).
- 8 Gerersdorfer, T., Frank, A., Formayer, H., Haas, P., Moshhammer, H. (2009): **Untersuchung zur nächtlichen Abkühlung in einem sich ändernden Klima.** (Endbericht StartClim2005.A1b, November 2006).
- 9 Krüger, B. C., Schicker, I., Formayer, H., Moshhammer, H. (2009): **Feinstaub und Klimawandel – Gibt es Zusammenhänge in Nordostösterreich?** (Endbericht StartClim2006.A, Juli 2007).
- 10 Rössler, M., Laube, W., Weihs, P. (2009): **Avoiding bird collisions with glass surfaces. Experimental investigations of the efficacy of markings on glass panes under natural light conditions in Flight Tunnel II.** (Final report, March 2007).
- 11 Formayer, H., Hofstätter, M., Haas, P. (2009): **Untersuchung der Schneesicherheit und der potenziellen Beschneigungszeiten in Schladming und Ramsau.** (Endbericht STRATEGIE, Oktober 2007).
- 12 Kromp-Kolb, H., Formayer, H., Haas, P., Hofstätter, M., Schwarzl, I. (2009): **Beobachtete Veränderung der Hitzeperioden in Oberösterreich und Abschätzung der möglichen zukünftigen Entwicklungen.** (Endbericht Band 1 der Forschungsreihe „Auswirkungen des Klimawandels auf Oberösterreich“, Februar 2007).
- 13 Moshhammer, H., Gerersdorfer, T., Hutter, H.-P., Formayer, H., Kromp-Kolb, H., Schwarzl, I. (2009): **Abschätzung der Auswirkungen von Hitze auf die Sterblichkeit in Oberösterreich.** (Endbericht Band 3 der Forschungsreihe „Auswirkungen des Klimawandels auf Oberösterreich“, Juli 2007).
- 14 Formayer, H., Kromp-Kolb, H., Schwarzl, I. (2009): **Auswirkungen des Klimawandels auf Hochwasserereignisse in Oberösterreich.** (Endbericht Band 2 der Forschungsreihe „Auswirkungen des Klimawandels auf Oberösterreich“, Mai 2007).
- 15 Simic, S., Schmalwieser, A.W., Moshhammer, H. (2009): **Gesundheitsrisiken für die österreichische Bevölkerung durch die Abnahme des stratosphärischen Ozons.** (Endbericht StartClim2007.B, Juni 2008).
- 16 Formayer, H., Clementschitsch, L., Hofstätter, M., Kromp-Kolb, H. (2009): **Vor Sicht Klima! Klimawandel in Österreich, regional betrachtet.** (Endbericht Global 2000, Mai 2008).
- 17 Eitzinger, J., Kubu, G. (eds.) (2009): **Impact of Climate Change and Adaptation in Agriculture.** (Extended Abstracts of the International Symposium, University of Natural Resources and Applied Life Sciences (BOKU), Vienna, June 22-23 2009).
- 18 Formayer, H., Kromp-Kolb, H. (2009): **Klimawandel und Tourismus in Oberösterreich.**

- 19 Fleischhacker V., Formayer H., Seisser O., Wolf-Eberl S., Kromp-Kolb, H. (2009): **Auswirkungen des Klimawandels auf das künftige Reiseverhalten im österreichischen Tourismus.** (Forschungsbericht im Auftrag des Bundesministeriums für Wirtschaft, Familie und Jugend).
- 20 Seibert, P., Radanovics, R., Krüger, B. C. (2010): **Sources of Air Pollution Relevant for the Austrian Biosphere Reserve Wienerwald.** (Final Report for the Project "Sources of Air Pollution Relevant for Austrian Biosphere Reserves: Quantification, Trends, Scenarios", Man and Biosphere Programme [MaB], Austrian Academy of Sciences [ÖAW]).
- 21 Arnold, D., Morton, D., Schicker, I., Seibert, P., Rotach, M. W., Horvath, K., Dudhia, J., Satomura, T., Müller, M., Zängl, G., Takemi, T., Serafin, S., Schmidli, J., Schneider, S. (2012): **High Resolution Modelling in Complex Terrain.** (Report on the HiRCoT 2012 Workshop, Vienna, 21-23 February 2012).

Alle Berichte sind unter <http://www.boku.ac.at/met/report/> online verfügbar.