

Globale Bestimmung von Gesamtozonsäulen und Ozonprofilen aus GOME Daten mit Hilfe neuronaler Netze

Dissertation
zur Erlangung des Grades
Dr. rer. nat.
der Universität Bremen

vorgelegt von

Martin D. Müller
(geboren am 11.4.1972 in Bamberg)

am

Institut für Fernerkundung / Institut für Umweltphysik
Fachbereich 1, Physik und Elektrotechnik
Universität Bremen

Tag der Einreichung: 4. Oktober 2002

Erstgutachter: Prof. Dr. J. P. Burrows
Zweitgutachter: Prof. Dr. J. Bleck-Neuhaus

für meine Eltern

Inhaltsverzeichnis

Abstract	1
Motivation und Ziele der Arbeit	2
I Grundlagen	5
1 Ozon in der Atmosphäre	7
1.1 Überblick	7
1.2 Chemismus	8
1.3 Transportvorgänge und Dynamik	10
1.4 Bedeutung für Mensch und Umwelt	11
2 Ozonmessungen	13
2.1 Bodengebundene Messungen	13
2.1.1 Ballonmessungen	13
2.1.2 Spektrophotometer-Messungen	16
2.1.3 Weitere Optische Messverfahren	17
2.2 Satelliteninstrumente	18
2.2.1 TOMS	19
2.2.2 SAGE II	19
2.2.3 HALOE	21
2.2.4 POAM II/III	22
3 Das GOME Instrument	24
3.1 Historische Einordnung	24
3.2 Eigenschaften	25
3.3 Backscatter-UV Theorie	26
3.4 Klassische GOME-Ozonbestimmung	29
3.4.1 Optimal Estimation	30
3.4.2 DOAS	34
4 Neuronale Netze	36
4.1 Grundlagen	36
4.1.1 Biologisches Vorbild	36
4.1.2 Das künstliche Neuron	40
4.1.3 Netzwerkarchitekturen	42
4.2 Belernungsalgorithmen	43
4.2.1 Veranschaulichung und Motivation	43
4.2.2 Backpropagation	44
4.2.3 Resilient Propagation	47
4.3 Einordnung in die klassische Inversionstheorie	49

4.4	Bisheriger Einsatz in der Atmosphärenforschung	52
II	Die NNORSY Methode	55
5	Softwareüberblick	57
5.1	Vorbereitung der Spektraldaten	57
5.1.1	Extraktion und Kalibration	57
5.1.2	Spektraldatenformat	59
5.1.3	Behandlung von Kanal 1A	59
5.2	Kollokation mit Ozonmessungen	60
5.3	Ergänzende Eingangsdaten	61
5.4	Datenmischung	62
5.5	Training neuronaler Netze	63
6	Gesamtozonbestimmung	65
6.1	Gesamtozon-Kollokation	65
6.2	Homogenisierung der Trainingsdaten	67
6.3	Eingangsparameter und Belernung	68
6.4	Ergebnisse und Validation des NNORSY-Gesamtozons	70
6.4.1	Teststationen-Statistik	71
6.4.2	Zeitreihen	73
6.4.3	Abhängigkeit vom Ozonwert	76
6.4.4	Zonale Mittelwerte	78
7	Bestimmung von Ozonprofilen	80
7.1	Ozonprofil-Trainingsdaten	80
7.1.1	Interpolation der Profildaten	80
7.1.2	Geografische und zeitliche Gleichverteilung	81
7.1.3	Homogenisierung der Trainingsdaten	82
7.2	Parziale Belernung neuronaler Netze	84
7.2.1	Modifikation von RPROP	84
7.2.2	Ergebnisse der Tests	85
7.2.3	Interpretation	86
7.3	Auswahl der Eingangsdaten	88
7.4	Ergebnisse der Ozonprofilbestimmung	89
7.4.1	Fehlerstatistik für Trainings- und Testdaten	89
7.4.2	Sensitivitäten der Fehler	92
7.4.3	Abschätzung der Vertikalauflösung	96
	Zusammenfassung von Teil II	100
III	Sensitivitäten und Fehler	103
8	Variation der Netzwerkkonfiguration	105
8.1	Variation der GOME-Spektraldaten	107
8.1.1	Keine Spektraldaten (NNORSY Klimatologie)	107
8.1.2	Spektralbereichs-Sensitivität	109
8.1.3	Überlegungen zur Lokalisierbarkeit von Höheninformation	111
8.2	Effekt anderer Netzwerkparameter	113
8.3	Normierung von Ein- und Ausgabe	115
8.3.1	Eingangsnormierung	115

8.3.2	Ausgangsnormierung	115
9	Abschätzung lokaler Fehler	118
9.1	Methodenbeschreibung	118
9.1.1	Retrieval mit verrauschten Eingangsdaten	118
9.1.2	Quadratische Fehlerfortpflanzung	119
9.1.3	Belernung eines Fehlernetzwerkes	120
9.2	Fallbeispiele und Interpretation	120
9.2.1	Einzelprofile	120
9.2.2	Vertikale Fehlerkorrelation	122
9.2.3	Rauschfehler-Sensitivitätsanalyse	124
9.3	Diskussion der Jacobi-Matrix	125
10	Globales Fehlerbudget	129
10.1	Übersicht und Strategie	129
10.2	Trainingsprofil-Fehler	130
10.2.1	Zufällige Fehler	130
10.2.2	Systematische Fehler	133
10.3	Andere Fehlerquellen	135
10.3.1	Eingangsdatenfehler	135
10.3.2	Kollokationsfehler	135
10.3.3	Glättungsfehler	137
10.3.4	Modellfehler	137
10.4	Zusammenstellung des Fehlerbudgets	138
	Zusammenfassung von Teil III	142
IV	Fallstudien und Anwendung	143
11	Zeitreihe ausgewählter Bodenstationen	145
11.1	Ozonesonden Hohenpeißenberg	145
11.2	Ozonesonden Syowa	148
11.3	<i>GOME Ozone Profile Working Group</i>	149
12	Integrierte Ozonsäulen	152
13	Aspekte der operationellen Anwendung	156
13.1	Unterschiede zur Offline-Version	156
13.2	Implementation für den GOME NRT-Service	156
13.3	Zeitliche Extrapolation und Nachtrainieren	157
14	Zusammenfassung und Ausblick	159
	Anhang	164
A	Liste der verwendeten Bodenstationen	164
B	TOMS-Korrekturfunktion	169
C	Behandlung der Profilkollokationen	171
D	Lebenslauf und Publikationen	173

E Abkürzungsverzeichnis	176
Literaturverzeichnis	180

Abbildungsverzeichnis

1.1	Monatsmittel des Gesamt ozons	10
2.1	Verteilung der im WOUDC registrierten Ozonsonden-Stationen	15
2.2	Absorptionsspektrum des Ozons	17
3.1	GOME Abtastmuster im nominellen Scanmodus ($\pm 31^\circ$ Schwenk)	25
3.2	Normierte Ozon-Gewichtsfunktionen	29
3.3	Glättungsfunktionen (<i>averaging Kernels</i>) des GOME OE Retrievals	33
4.1	Neuronentypen und vereinfachtes Verbindungsschema	37
4.2	Schema des Informationsflusses im menschlichen Gehirn	39
4.3	Schema eines simulierten Neurons	41
4.4	Verbindungsschema für <i>feed-forward</i> MLP-Netzwerke	42
4.5	Mögliche Lernkurve eines Menschen	43
4.6	Gradientenabstieg zum Minimum der Fehlerfunktion mit Hilfe von Backpropagation.	44
4.7	Annäherung an ein Minimum während der Belernung eines neuronalen Netzwerkes.	47
5.1	Offline-Prozesskette der NNORSY-Software	58
6.1	Bodenstationen für die Gesamt ozonbelernungen	66
6.2	2D-Histogramm der Gesamt ozonwerte im Trainingsdatensatz	67
6.3	Histogramm der Gesamt ozonwerte nach Anzahlreduktion	68
6.4	Globales NNORSY-Ozonfeld	71
6.5	Variabilität des Ozonfeldes	72
6.6	Gesamt ozon-Zeitreihen Boulder, Hohenpeißenberg und Bangkok	74
6.7	Gesamt ozon-Zeitreihen Lauder und virtuelle Stationen	75
6.8	Streudiagramm Retrieval \leftrightarrow Bodenmessungen	77
6.9	Zonale Mittelwerte Gesamt ozon	79
7.1	Typische Beispiele für drei Arten von Ozonprofildaten	80
7.2	Verteilung der Ozonprofil-Kollokationen	81
7.3	Beispiel für die Homogenisierung der Profilkollokationen	82
7.4	Mittelwert und Standardabweichung der Profilkollokationen	83
7.5	RMSE bei variabler Anzahl fehlender Trainingsdaten	86
7.6	Eigenwerte des Testdatensatzes für Parziale Belernung	87
7.7	Globale Fehler des Trainingsdatensatzes	91
7.8	Globale Fehler des Testdatensatzes	91
7.9	Testdatensatz-Fehler nach Latitudenband	93
7.10	Zieldatenverteilung für die 9 km und die 30 km Höhengschicht	94
7.11	Testdatensatz-Fehler nach Sensorentyp	95

7.12	Testdatensatz-Fehler nach Pixel <i>line-of-sight</i>	95
7.13	Abschätzung der vertikalen Profilauflösung	97
7.14	NNORSY Ozonprofil im Vergleich zu geglättetem Sondenprofil	98
8.1	Variation zwischen Belernungen	106
8.2	Belernung ohne GOME-Daten	107
8.3	Teilsäulenvergleich Klimatologienetze	108
8.4	Sensitivität gegenüber Eingangswellenlängenauswahl	110
8.5	Belernung für $\lambda \geq 300$ nm	112
8.6	Auswirkung der Ausgabe-Normierung auf das Ozonprofilretrieval	116
9.1	Relative lokale Fehler für die NNORSY-Profile	121
9.2	Kombinierte Fehlerkorrelationsmatrizen	123
9.3	Korrelationslänge der Profilfehler	123
9.4	Abhängigkeit der Rauschfehler von den Eingangsdaten	124
9.5	Mittlere Jacobi-Matrix des neuronalen Netzes	126
10.1	Zunahme des Kollokationsfehlers mit der Kollokationsentfernung	136
11.1	Zeitreihenvergleich von 10 km-Ozonteilsäulen	146
11.2	Streudiagramme von 5 km-Teilsäulen der Ozonsonden gegenüber NNORSY	147
11.3	Vergleich von Ozonprofilen im Rahmen der GOME Ozone Profile Working Group	150
12.1	Vergleich der Ozon-Monatsmittel für Apr. und Okt. 1999	153
B.1	Korrekturfunktion für EP-TOMS-Daten	170

Tabellenverzeichnis

2.1	Wichtige orbitale Ozonsensoren	18
2.2	Übersicht Okkultations-Instrumente	20
3.1	GOME Spektralkanäle	26
4.1	Verwendung neuronaler Netze in der Atmosphärenphysik	54
6.1	Neuronales Netz für Gesamtozonretrieval	69
6.2	Eckdaten der Gesamtozon-Teststationen	71
6.3	Vergleich der Statistik verschiedener Gesamtozonmessungen	72
7.1	Neuronales Netz für simuliertes Ozonprofilretrieval	85
7.2	Neuronales Netz für GOME-Ozonprofilretrieval	89
8.1	Veränderung des RMSE bei Variationen der Netzwerkkonfiguration	114
10.1	Präzision der kollokierten Ozonprofildaten	131
10.2	Profildatenfehler des Kollokationsdatensatzes	134
10.3	Rauschfehler der kollokierten Ozonprofildaten	135
10.4	Kollokationsfehler der kollokierten Ozonprofildaten	137
10.5	Glättungsfehler der NNORSY-Ozonretrievals	138
10.6	Modellfehler der NNORSY-Ozonretrievals	138
10.7	Referenz-Fehlerbudget	139
10.8	Globales NNORSY-Fehlerbudget	140
11.1	Vergleichs-Datensatz der <i>GOME Ozone Profile Working Group</i>	149

Abstract

In order to retrieve the atmospheric ozone distribution from the UV-visible satellite spectrometer GOME, inverse radiative transfer was modelled directly by means of a multilayer perceptron (MLP) type neural network. This MLP was trained exclusively on a data set of measured GOME radiances with collocated ozone measurements from ground-based instruments and TOMS in the case of total ozone retrieval, and ozonesondes plus SAGE II, HALOE and POAM III data in the case of ozone profile retrieval. Network input consists of a combination of spectral, geolocation and climatological information (latitude and time), the latter making the use of external *a priori* ozone profiles unnecessary. In addition, UKMO temperatures were used in ozone profile retrieval. This method is termed *Neural Network Ozone Retrieval System* (NNORSY).

Total ozone thus retrieved yielded RMS errors – relative to a six station test dataset – of about 12 DU, which compares well to TOMS v7 data, and is less than the operational GDP v2.7 product. The ozone profile retrieval network globally reduces standard deviation for randomly selected test collocations with respect to an ozone climatology by around 40%. Tropospheric ozone can also be retrieved in many cases, with a corresponding reduction of 10–30%. The results were further investigated and validated by a number of sensitivity studies involving statistics on regional datasets, modification of network input layout and training, estimation of local errors by creation of noisy retrievals and training of an error network. A link to classical retrieval theory was established, and discussed by means of the neural Jacobian. Vertical resolution of the profiles was assessed empirically by comparison with the collocations, and by examining profile error correlations. It seems to be in the order of 4–6 km, but gets worse above 40 km. The neural network was found capable of implicitly correcting for most instrument degradation and scan angle effects, based on the input data provided.

In conjunction with a global error budget created from ozone profile errors found in the corresponding literature – plus empiric assessments of other contributions – the above studies revealed that the neural network method compares well to classical retrieval results in most areas, but seems to thoroughly and non-linearly mix statistical and physical information. This prevents the use of most classical analysis tools, in particular concerning linear radiative theory approximations.

The above results were confirmed in a number of case studies involving the ozonesonde stations Hohenpeißenberg, Syowa, and the classical retrieval method FURM. Integrated NNORSY profiles generally agree within $\pm 5\%$ with the monthly TOMS v7 total ozone field, comparable to GDP total ozone.

Some issues remain at high solar zenith angles and very low ozone values, as well as in the tropopause region and in the oceanic and Antarctic troposphere. To resolve these, a number of measures are proposed which reflect the unique nature of the approach and should be easier to implement than the exact radiometric calibration required by classical methods. Use of MLPs for retrieving other trace gases and atmospheric parameters is limited by the necessity of a suitable set of training measurements, or an accurate simulation of sensor characteristics. However, NNORSY-GOME has been demonstrated to offer great speed advantages in an operational, prototype near-real-time system, and the relatively simple adjustment to sensors other than GOME make it a good candidate for upcoming orbital ozone instruments.

Motivation und Ziele der Arbeit

Die zunehmende Beeinflussung unserer natürlichen Umwelt durch den Menschen ist in den letzten Jahrzehnten immer stärker in das Bewußtsein der Öffentlichkeit gerückt. Die Veränderungen beziehen sich auf alle Bereiche der Umwelt, sind aber insbesondere im System Atmosphäre-Sonne am ehesten der Alltagswahrnehmung zugänglich, da es dem Menschen aus tagtäglichem, langjähriger Erfahrung intuitiv vertraut ist. Schlagworte wie Klimawandel, Sommersmog, Treibhauseffekt und Ozonloch sind daher in aller Munde, während subtile Veränderungen der Flora und Fauna, sowie der Boden- und Wasserqualität nur bestimmten Berufsgruppen bewußt werden.

Um so wichtiger ist es, die allgemein erfahrbaren Probleme des Atmosphärensystems möglichst schnell zu verstehen und deren anthropogene Ursachen publik zu machen, damit sich Entscheidungsträger aus Politik und Wirtschaft dem Druck der öffentlichen Meinung ausgesetzt sehen. Klare Lösungsvorschläge aus der Wissenschaft können dann am ehesten zur verbindlichen Umsetzung und zu einer Verbesserung der Situation führen.

Diese Erkenntnis hat in den letzten Jahren dazu geführt, dass ein zunehmendes Arsenal immer ausgefeilterer Instrumente zur Erdbeobachtung von Satelliten aus eingesetzt wird, da dies die einzige Möglichkeit darstellt, eine globale Abtastung der Atmosphäre zu realisieren. Dabei bestimmt bis heute meist ein gewisser *trade-off* zwischen räumlicher und zeitlicher Auflösung und Präzision der Messungen die Unterschiede zwischen den verschiedenen Satelliteninstrumenten. Vom Erdboden aus durchgeführte Messungen reihen sich in der Regel als Extremfall in dieses Schema ein, da sie zwar häufig sehr präzise, aber ihrer Natur gemäß auf räumlich kleine, lokale Luftmassen beschränkt sind. Sie sind aber für die Validierung der Satellitenmessungen unumgänglich.

Ein nicht zu unterschätzendes Problem bei dieser Entwicklung besteht in der Datenmenge der modernen Instrumente, welche in der letzten Zeit so schnell angestiegen ist, dass die für ihre Auswertung benötigte Rechenleistung kaum Schritt halten konnte: Während das im Vorfeld dieser Arbeit vom Autor untersuchte TOVS-Instrumentenpaket seit dem Beginn seines Einsatzes im Jahre 1978 auf 11 verschiedenen Satelliten flog und insgesamt etwa 2 TB an Level 1 Daten lieferte, entsteht diese Datenmenge beim 1995 gestarteten GOME-Instrument bereits innerhalb von gut einer Woche. Nicht zu unterschätzen ist außerdem der Aufwand, der für die Entwicklung und Validierung von Prozessierungssoftware betrieben werden muss, z. B. um systematische Fehler bei den immer empfindlicheren und komplexeren Messungen zu minimieren.

In der vorliegenden Arbeit soll ein Ansatz untersucht werden, der das Potenzial hat, mehrere der angesprochenen Problematiken im Bezug auf die Auswertung von Satellitendaten mit einem Schlag zu lösen. Das Werkzeug hierzu ist die aus dem biologischen Vorbild hervorgegangene, jedoch stark vereinfachte Simulation neuronaler Netze, deren Grundlage verschiedene statistische und informationstheoretische Theorien und Algorithmen bilden. Mit einem dem Problem angepassten geophysikalischen Überbau versehen besitzt ein solches System u. a. folgende Vorteile:

- Qualitativ hochwertige lokale Messungen atmosphärischer Parameter werden nicht nur für die Validierung verwendet, sondern bilden die direkte Grundlage des Inversionsverfahrens.
- Obwohl für die Parametrisierung („Belernung“) des neuronalen Netzes einmalig relativ viel Rechenzeit benötigt wird, ist die operationelle Anwendung des Verfahrens wesentlich schneller als klassische Algorithmen, und daher für große Datenmengen und Echtzeit-Einsatz besser geeignet.
- Die Kalibrierung der Satellitendaten gestaltet sich als weit weniger aufwändig, da das Verfahren bei geschickter Wahl der Eingangsgrößen entsprechende Korrekturen selber erlernen kann.
- Auf die gleiche Weise sinkt die Abhängigkeit der Ergebnisse von instrumentenspezifischen Parametern, was die mögliche Anpassung an zukünftige Satelliteninstrumente erleichtert.

Da die Entwicklung eines allgemeinen Verfahrens dieser Art den Rahmen der vorliegenden Arbeit sprengen würde, wurde die Umsetzung dieses vielversprechenden Ansatzes im Hinblick auf atmosphärisches Ozon konkretisiert. Aufgrund seiner besonderen Eigenschaften ist Ozon das wohl wichtigste und interessanteste atmosphärische Spurengas: Seine Verteilung ist höchst inhomogen, es bildet die vor UV-Strahlung schützende stratosphärische Ozonschicht, die überhaupt erst Leben auf der Erde ermöglichte. Es ist für die stratosphärische Temperaturverteilung maßgeblich, ebenso wie für die oxidative Kapazität der Troposphäre, womit Sommersmog, Lungenprobleme und Schäden an Pflanzen und Gebäuden einhergehen. In allen Höhenbereichen ist sein natürliches Gleichgewicht bereits nachhaltig durch anthropogene Emissionen beeinträchtigt – am Boden durch Stickoxide und Kohlenwasserstoffe, in der Stratosphäre durch Chlor- und Bromverbindungen. Obwohl seit dem Beginn systematischer Ozonmessungen in den 20er Jahren durch den Pionier Dobson bereits viele Erkenntnisse über die Chemie und Dynamik des Gases erlangt wurden, ist die Ozonforschung weiterhin aktueller Forschungsgegenstand einer wachsenden Wissenschaftlergemeinschaft, und die im Rahmen der Validierung und Anwendung des zu entwickelnden Verfahrens erhaltenen Ergebnisse können direkt in der aktuellen Forschung verwendet werden.

Die Arbeit wird auf zwei verschiedene Methoden auf Basis neuronaler Netze eingehen, die ich zusammenfassend mit *Neural Network Ozone Retrieval System (NNORSY)* bezeichnen werde. Zum einen wird ein Verfahren zur Bestimmung von **Gesamtozonsäulen** aus GOME-Daten entwickelt. Dies stellt ein verhältnismäßig einfaches Problem dar, trotzdem variieren die bislang verfügbaren operationellen Ozonprodukte für verschiedene orbitale Sensoren zum Teil erheblich. Es soll daher gezeigt werden, ob und gegebenenfalls wie dieses Problem mit neuronalen Netzen zufriedenstellend gelöst werden kann, und wo die spezifischen Unterschiede zu den bisherigen Lösungen und Ergebnissen bestehen.

Zum anderen sind die Möglichkeiten des relativ hochauflösenden GOME-Spektrometers damit aber noch nicht ausgeschöpft, da dieses Gerät erstmals mit guter flächenmäßiger Abdeckung auch die Bestimmung von **Ozonprofilen** ermöglicht. Allerdings ist dieses Inversionsproblem wesentlich schwieriger, da die Höheninformation unter anderem in kleinen, temperaturbedingten Variationen der Ozonabsorption verborgen ist und daher ein viel schlechteres Signal/Rauschverhältnis aufweist. Außerdem sind zusätzliche Informationen und Annahmen nötig, damit die Lösung eindeutig wird, und sie kann in klassischen Verfahren nur iterativ berechnet werden. Gerade bei derart komplexen Problemen kommen die Geschwindigkeitsvorteile neuronaler Netze besonders zur Geltung – sie bewegen sich durchaus im Bereich mehrerer Größenordnungen.

Das Hauptziel der vorliegenden Arbeit ist somit, diese beiden Ozonbestimmungsverfahren für den GOME-Sensor zu entwickeln und zu validieren. Dabei sollen auch Defizite, die die neuronalen Netze im Bereich der Fehlerabschätzung besitzen, behandelt werden, so dass die gewonnenen Ergebnisse gut charakterisiert werden können. Da trotz internationaler Anstrengung mehrerer Arbeitsgruppen noch kein offizielles GOME-Ozonprofilbestimmungsverfahren in Betrieb ist, versteht sich diese Arbeit auch als wissenschaftliche Grundlage für Entwicklung einer operationellen Version von NNORSY-GOME. Die hier gesammelten Erkenntnisse zur Modellierung des inversen Strahlungstransports mit Hilfe neuronaler Netze dienen jedoch gleichzeitig der Vorbereitung auf zukünftige Generationen von spektral und räumlich hochauflösenden Satelliteninstrumenten.

Anmerkungen zum Aufbau

Der hier behandelte Stoff wurde zur besseren Übersicht in vier Teile gegliedert:

Teil I beschreibt einige zum Verständnis der verwendeten Konzepte hilfreichen Grundlagen zu den Themen atmosphärisches Ozon, Satelliten-Fernerkundung und neuronale Netze.

Teil II beschreibt die entwickelten Gesamtozon- und Ozonprofilbestimmungsmethoden und ihre grundsätzliche Validierung.

Teil III konzentriert sich auf eine genauere Charakterisierung der NNORSY Ozonprofile und ihrer Fehler.

Teil IV wendet die beiden Methoden auf einige konkrete Fälle an, und geht auf den operationellen Betrieb und mögliche Verbesserungen der Verfahren ein.

Jeder Teil der Arbeit beginnt mit einer Übersicht über die in seinen Kapiteln behandelten Themen, und schließt mit einer Zusammenfassung. Davon abgesehen findet sich am Ende einiger ausgewählter Abschnitte zusätzlich ein kurzes Fazit, welches die wichtigsten getroffenen Aussagen zusammenstellt.

Ich habe mir außerdem erlaubt, einige in der englischsprachigen Fachliteratur gebräuchlichen, wissenschaftlichen und technischen Fachtermini direkt zu übernehmen, um Mißdeutungen zu vermeiden. Da z. B. der in der Satellitenfernerkundung häufig verwendete Begriff *retrieval* nur unzureichend mit *Bestimmung* übersetzt werden kann, wird er in dieser Arbeit wie ein Fremdwort verwendet, das sich immer auf die Inversion von Satellitendaten bezieht.

Teil I

Grundlagen

In diesem Teil werden die Grundlagen für das Verständnis der Arbeit gelegt.

Kapitel 1 beschreibt die wichtige Rolle des Ozons in unserer Umwelt, seine Quellen und Senken sowie die atmosphärische Dynamik, die seine globale Verteilung stark beeinflusst.

In **Kapitel 2** werden die Methoden aufgezeigt und erläutert, die im Laufe der letzten Jahrzehnte für die Messung von atmosphärischem Ozon entstanden sind. In diesem Zusammenhang werden auch die in der vorliegenden Arbeit für Training und Validierung der neuronalen Netze verwendeten Ozon-Datensätze eingeführt.

Eine Ausnahme bildet das Satellitenexperiment GOME, dem das **Kapitel 3** gewidmet ist. Hierbei wird kurz auf die technischen Aspekte des Instruments eingegangen, bevor ein Überblick über die derzeit verwendeten Auswertemethoden DOAS und Optimal Estimation gegeben wird, die der Bestimmung von atmosphärischem Ozon aus GOME-Spektren dienen.

Kapitel 4 schließt den ersten Teil der Arbeit mit einer Einführung in die Methode der neuronalen Netze ab. Zunächst wird die zugrundeliegende Idee anhand des biologischen Vorbilds vorgestellt, bevor die Diskussion auf den hier verwendeten Netzwerktyp und seine Simulation im Computer eingeschränkt wird. Es folgt ein kurzer Überblick über die Einordnung und bisherige Verwendung simulierter neuronaler Netze im Rahmen der Atmosphärenphysik.

Kapitel 1

Ozon in der Atmosphäre

1.1 Überblick

Ozon kommt in nennenswertem Maße nur in den untersten beiden Schichten der Atmosphäre vor, der Troposphäre und der Stratosphäre. Diese Trennung der beiden Schichten wird durch eine Temperaturinversion hervorgerufen: Während die Temperatur in der Troposphäre kontinuierlich abnimmt, bedingt die Absorption von UV-Strahlung durch Ozon ab einer Höhe von etwa 8 km (an den Polen) bis 18 km (in Äquatornähe) einen erneuten Anstieg der Temperatur. Die durch das Minimum der Temperatur definierte Tropopause bildet eine dynamische Barriere und läßt die meisten Prozesse in beiden Schichten getrennt verlaufen.

Als Einheit für die Ozonsäulenhöhe wird traditionell die nach dem Ozonpionier benannte *Dobson Unit* (DU) verwendet. 1000 DU entsprechen einer Ozonsäule von 1 cm unter Normalbedingungen. Die durchschnittliche Ozon-Gesamtsäule beträgt global 300–350 DU, wobei der troposphärische Anteil nur etwa 10% der Säule ausmacht. Für die Angabe vertikaler Ozonprofile gibt es je nach Anwendung verschiedene Einheitensysteme. In der vorliegenden Arbeit wurde durchgehend die Ozonmoleküldichte in $[10^{18}/\text{m}^3]$ über der geopotentiellen Höhe (GPH) in [km] aufgetragen. Erstere geht direkt in die Strahlungstransfergleichung ein und spiegelt somit den Strahlungstransport am besten wider, während letztere die Temperaturschichtung bis zu einem gewissen Grad berücksichtigt, aber nicht allzu stark von der geometrischen Höhe abweicht.

Ozon wird auf natürliche Weise hauptsächlich in der tropischen Stratosphäre über die Photodissoziation von zweiatomigem Sauerstoff gebildet. Anschließend wird es durch die Brewer-Dobson-Zirkulation [Brewer, 1949; Dobson, 1956] polwärts verfrachtet, wo es sich ansammelt, so dass die absolut größten Ozonsäulen nicht in der eigentlichen Entstehungsregion, sondern in den mittleren bis hohen Breiten zu finden sind [Fabian, 1992]. Durch den Transport entsteht so eine globale Ozonschicht in der Stratosphäre.

Troposphärisches Ozon entsteht einerseits durch photochemische Bildung [Sonnemann, 1992], andererseits können durch Austauschprozesse ozonreiche, stratosphärische Luftmassen in die Troposphäre gelangen [Levy *et al.*, 1985]. Der prozentuale Beitrag der beiden Quellen ist noch umstritten [Lelieveld & Dentener, 2000, und Referenzen darin] und variiert vermutlich stark mit den örtlichen Gegebenheiten.

Als Senken für Ozon stehen in der Stratosphäre verschiedene natürliche Abbaumechanismen zur Verfügung, zum einen die Photodissoziation mit anschließender Rekombination zu Sauerstoffmolekülen, zum anderen Abbaureaktionen mit verschiedenen Katalysatoren, von denen NO_x den größten Beitrag stellt. Aus der Troposphäre wird Ozon hauptsächlich durch trockene Deposition wieder entfernt.

1.2 Chemismus

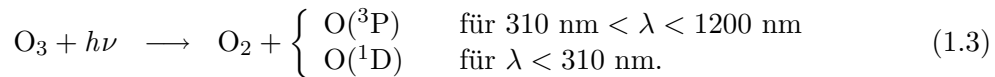
Wie bereits erwähnt, finden die chemischen Prozesse, die Entstehung und Abbau von Ozon bewirken, auf unterschiedliche Weise in Troposphäre und Stratosphäre statt. Die Ozonchemie ist äußerst komplex und auch heute noch nicht vollständig verstanden. In aktuellen Chemie-Transport-Modellen der Erdatmosphäre [z. B. *Douglass et al.*, 1996; *Khattatov et al.*, 2000] werden im Zusammenhang mit Ozon meist etwa 100–200 verschiedene Reaktionsgleichungen berücksichtigt. Dieses Kapitel kann daher nur einen kurzen Einblick in die bedeutendsten Reaktionen liefern, der aber bereits für das Verständnis der wesentlichen Effekte genügt.

Historisch gesehen wurde für Ozon zunächst die **reine Sauerstoffchemie** formuliert, bei der das Ozon durch die Photodissoziation von zweiatomigem Sauerstoff entsteht [*Chapman*, 1930]:



Diese Gleichungen machen bereits deutlich, warum das Ozon vornehmlich in der mittleren Atmosphäre zu finden ist: Die zur Aufspaltung von O_2 erforderliche harte UV-Strahlung reicht nicht in tiefere Schichten der Atmosphäre hinunter, sie wird durch das Sauerstoffmolekül absorbiert, und zwar im Schumann-Runge-Kontinuum ($\lambda < 170$ nm), den Schumann-Runge-Banden ($170 \text{ nm} < \lambda < 220$ nm), sowie dem Herzberg-Kontinuum ($\lambda > 220$ nm), zu dem auch die Gl. 1.1 gehört. In sehr großen Höhen hingegen ist die Dichte der Luft und somit die Menge an O_2 sehr gering, außerdem machen die großen Weglängen das Auffinden des für die Impulserhaltung bei der Ozonentstehung benötigten Stoßpartners M (meist Stickstoff, N_2) unwahrscheinlicher.

Der Abbau von Ozon erfolgt ebenfalls photolytisch, gemäß



Je nach Energie der beteiligten Photonen befindet sich das entstehende Sauerstoffatom im Grundzustand (Triplet-Zustand) $\text{O}({}^3\text{P})$ oder im angeregten Singulett-Zustand $\text{O}({}^1\text{D})$, der gleichzeitig Energie für verschiedene Folgereaktionen bereitstellt. Die photolytische Spaltung von Ozon bei kurzen Wellenlängen ist viel effizienter, was sich in den hohen Absorptionsquerschnitten (bis 10^{-17} cm^2) der Hartley- und Huggins-banden unterhalb von etwa 300 nm widerspiegelt, und der Grund für die hervorragende UV-Schutzwirkung der Ozonschicht ist. Ein einzelnes Ozonmolekül in 30 km Höhe hat daher eine Lebensdauer von weniger als einer Stunde [*Roedel*, 2000]. Aufgrund der Konzentrationsverhältnisse in der mittleren Stratosphäre und darunter rekombinieren die photolytisch erzeugten Sauerstoffatome aber dort meist sofort wieder mit O_2 zu Ozon, wodurch langfristige Transportprozesse und globale Verteilung des Ozons überhaupt erst ermöglicht werden [*Spichtinger-Radowsky*, 2001].

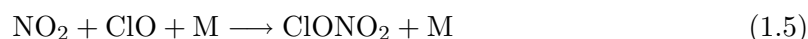
Dementsprechend liegen auch bei der Simulation der reinen Sauerstoffchemie des Ozons die Werte um etwa 30% über den gemessenen Konzentrationen. Als Grund dafür stellen sich eine Reihe von **katalytischen Abbaureaktionen** heraus, die alle nach dem Schema



ablaufen. Die wichtigsten Katalysatoren X sind hierbei

- NO.** Stickstoffoxid stellt den wichtigsten natürlichen Katalysator dar [Fabian, 1992] und ist alleine für 25% der obigen 30% Differenz verantwortlich.
- OH.** Hydroxylradikale entstehen in der Stratosphäre durch Reaktion von $O(^1D)$ mit dem allgegenwärtigen Wasserdampf.
- Cl.** Chlor gelangt auf natürliche Weise in der Form von biogenem Methylchlorid $CClH_3$ in die Stratosphäre. Seine Konzentration ist normalerweise sehr gering, aber die Katalyse mit Cl ist sehr effizient, weswegen es trotzdem beim Ozonabbau eine wichtige Rolle spielt. In jüngerer Zeit hat die Emission von Fluor-Chlor-Kohlenwasserstoffen (FCKW) zu einem starken Anstieg der stratosphärischen Chlorkonzentration geführt, da sie in der Troposphäre inert sind und erst in größeren Höhen durch UV-Photolyse ihr Chlor in der Form von Radikalen freigeben.

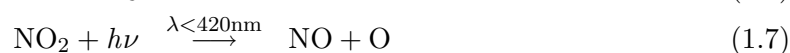
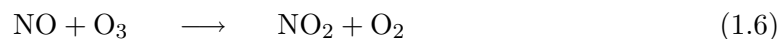
Diese Darstellung ist selbstverständlich stark vereinfacht; Gäbe es keine Wechselwirkungen zwischen den einzelnen Reaktionszyklen wie z. B. die Bildung des chemisch inaktiven Chlornitrats,



stünde ansonsten einem ungehemmten Abbau des Ozons nichts im Wege.

Ende der 70er Jahre glaubte man, die Ozonchemie im Wesentlichen beschrieben zu haben. Das regelmäßige frühjährliche Auftreten des sogenannten Ozonlochs über der Antarktis ab Beginn der 80er Jahre [Chubachi, 1984; Farman *et al.*, 1985] konnte jedoch nicht plausibel erklärt werden. Dies gelang erst, nachdem die heterogene Chemie an stratosphärischen Aerosolpartikeln berücksichtigt wurde [Solomon & R. R. Garcia, 1986; Crutzen & Arnold, 1986; McElroy *et al.*, 1986], insbesondere an den nur bei Temperaturen unter $-78^\circ C$ auftretenden Stratosphärenwolken (*polar stratospheric clouds*, PSC). Die an den Wolkentröpfchen ablaufenden Prozesse verschieben die Gewichtung der obigen Katalysereaktionen zugunsten von Chlor, indem sie einerseits die Reaktivierung des Chlors aus Depotsubstanzen wie $ClONO_2$ (Gl. 1.5) ermöglichen, andererseits deren erneute Bildung hemmen, indem das NO_2 als Salpetersäure HNO_3 in den Tröpfchen gebunden wird. Im Laufe der Polarnacht wachsen die Tröpfchen und setzen sich langsam ab. Die NO_x -armen (= denitrifizierten) Luftmassen enthalten dann viel aktives Chlor in Form von Cl_2 und $HOCl$, das zu Beginn des Frühjahrs photolytisch in Radikale zerlegt wird und sofort sein zerstörerisches Werk gemäß Gl. 1.4 beginnt.

Die Chemie des **troposphärischen Ozons** ist im Grundzustand hauptsächlich vom NO_x -Zyklus bestimmt, bei dem sich ein dynamisches Gleichgewicht zwischen NO und NO_2 einstellt [Kraus *et al.*, 1996; Harwood & Jones, 1994].



Dieses dynamische Gleichgewicht weist aufgrund seiner Strahlungsabhängigkeit einen ausgeprägten Tagesgang auf. Der atomare Sauerstoff aus Gleichung 1.6 reagiert generell über einen Dreierstoß (Gl. 1.2) zu Ozon weiter. An sonnigen Tagen stellt sich über die Reaktionen 1.6, 1.7 und 1.2 der sogenannte photostationäre Zustand ein.

Im Beisein von (zum Teil antropogen emittierten) Kohlenwasserstoffen RH (*volatile organic compounds*, VOC) ergeben sich Radikal-Kettenreaktionen, die vermehrt das NO zu

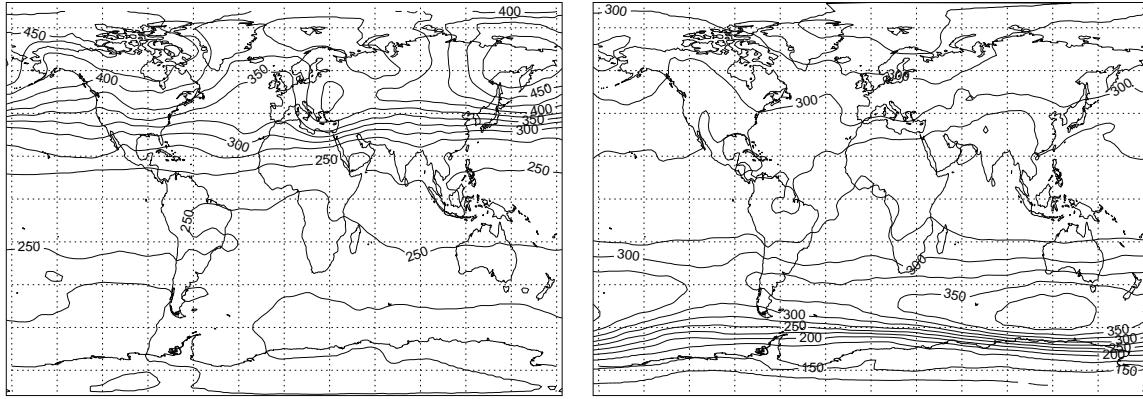
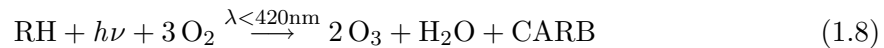
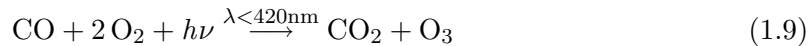


Abbildung 1.1: Monatsmittel des Gesamtozons im März (links) und September 1998. Die höchsten Werte stellen sich auf der jeweiligen Hemisphäre im Frühjahr ein. Eine Ausnahme bildet das Antarktische Ozonloch, zur Erläuterung siehe Text. (Quelle: TOMS)

NO_2 oxidieren, dessen Photolyse dann wiederum die Ozonwerte in die Höhe treibt [Skov *et al.*, 1997]. Die Summenreaktion läuft dabei nach dem Schema



ab. CARB bezeichnet hier Carbonylverbindungen wie z. B. Formaldehyd, die wiederum über Photolyse Kohlenmonoxid CO freisetzen. Häufig werden in Industrienationen zusammen mit VOC auch bedeutende Mengen Stickoxide emittiert, die ab einem Verhältnis NO zu O_3 von größer als 1/5000 als Katalysator für die (Summen-)Reaktion



wirken [Crutzen & Gidel, 1983] und damit weiteres Ozon produzieren. Wie zu erkennen, bildet sich auf diese Weise bei ausreichend Sonnenlichtzufuhr in verschmutzter Luft eine beträchtliche Menge Ozon, das zusammen mit dem durch die Radikalreaktionen entstehenden Cocktail aus teilweise toxischen organischen Verbindungen für den berüchtigten Sommersmog in Großstädten verantwortlich ist.

1.3 Transportvorgänge und Dynamik

Für die globale und lokale Verteilung von Ozon sind aufgrund seiner langen Lebensdauer meteorologische Transportprozesse von entscheidender Bedeutung. Der großräumigste davon ist die **Brewer-Dobson-Zirkulation**. Sie entsteht nach heutiger Vorstellung, indem das die winterliche Westwinddrift in mittleren Breiten überlagernde Muster planetarer und synoptischer Wellen (Rossby-Wellen, [Rossby, 1938]) gestört wird [Holton & Haynes, 1995]. Dies führt im Mittel zu einem polwärtigen Ozontransport mit gleichzeitiger leichter Absenkung. Dieser Mechanismus wirkt in der Nordhemisphäre (NH) etwa dreimal stärker als in der Südhemisphäre (SH), weil in der NH die geographiebedingten Störungen ausgeprägter sind [Rosenlof, 1995]. Hier finden sich daher auch im Mittel höhere Ozonsäulendicken als in gleichen Breiten der SH. Im Sommer kommt die Brewer-Dobson-Zirkulation aufgrund der geänderten Windverhältnisse praktisch zum Erliegen. Dies erklärt den Jahresgang der Gesamtozonsäulen in mittleren bis hohen Breiten, bei dem die höchsten Werte im Frühjahr gemessen werden, wie exemplarisch in Abb. 1.1 zu erkennen.

Die Nord-Süd Symmetrie der Abbildung wird im September jedoch empfindlich durch die sehr niedrigen Ozonwerte über der Antarktis gestört, dem oben erwähnten **Ozonloch**.

Die Entstehung dieses Lochs wird auf die niedrigen Temperaturen der in den zirkumpolaren Westwinden (= polarer Vortex) eingeschlossenen Luftmassen zurückgeführt. Durch die starke Abstrahlung des Antarktischen Eisschildes in der Polarnacht entsteht ein hoher Temperaturgradient zu den relativ warmen subpolaren Ozeanen, und durch die einheitliche Geographie können sich die resultierenden thermischen Winde ohne größere Störung den ganzen Winter über halten. Bei Temperaturen von unter -78°C in der Stratosphäre entstehen dann PSCs, und der im vorherigen Abschnitt besprochene Ozonzerstörungs-Mechanismus setzt ein.

Seit Mitte der 90er Jahre werden stark verminderte Ozonwerte auch im nordpolaren Spätwinter beobachtet [Wirth & Renger, 1996]. Hier sind die Bedingungen für die Bildung des Vortex viel schlechter: Einerseits entsteht kein so starker Temperaturgradient, da der Eisschild fehlt und die Kontinente um den Pol auch relativ stark abkühlen. Andererseits setzen sich die von der Geographie stammenden Störungen des Windfeldes bis in die Stratosphäre fort [Roedel, 2000]. Der Austausch von Luftmassen in meridionaler Richtung kommt deshalb auch im Winter nie ganz zum Erliegen. In der Tat findet gerade zu Anfang des Nord-Winters im Bereich Europas und des Nordatlantiks häufig meridionaler Austausch mit den Tropen statt, bei dem ozonarme, tropische Troposphärenluft angehoben und polwärts transportiert wird, wo sie in Zusammenspiel mit ausgeprägten Hochdruckgebieten kurze und örtlich begrenzte Episoden mit sehr niedrigen Gesamtozonsäulen („Minilöcher“) hervorruft [Hood et al., 2001].

Der Luftmassenaustausch über die Tropopausengrenze hinweg (*stratosphere-troposphere exchange*) findet häufig auch in umgekehrter Richtung statt [Appenzeller et al., 1996]. Da die Konzentration des troposphärischen Ozons von unten nach oben hin zunimmt, nahm man sogar zunächst an, es sei fast ausschließlich durch Austauschprozesse mit der ozonreichen Stratosphäre in die Troposphäre gelangt. Ozon besitzt jedoch am Boden eine Senke, die trockene Deposition, während es in der freien Troposphäre nur sehr langsam abgebaut wird. Mittlerweile wird der Beitrag der im vorangehenden Abschnitt 1.2 beschriebenen, chemischen Entstehungsprozesse global auf grob die Hälfte geschätzt, wobei die stark unterschiedliche Menge anthropogener Emissionen auf Nord- und Südhemisphäre wiederum eine starke Asymmetrie bewirkt [Lelieveld & Dentener, 2000].

Es wird bereits klar, dass sich in dem hier stark vereinfachten Wechselspiel verschiedener dynamischer Prozesse mindestens so viele ungelöste Fragestellungen finden lassen wie in der Ozonchemie. Gerade in diesem Bereich können Satellitenmessungen viel zur Wahrheitsfindung beitragen, da einerseits ihre immer besser werdende Auflösung die direkte Studie von lokalen Einzelphänomenen ermöglicht, auf die bislang nur indirekt geschlossen werden konnte. Andererseits können durch die (mehr oder weniger) lückenlose globale Abtastung auch großräumige Effekte immer besser untersucht werden.

1.4 Bedeutung für Mensch und Umwelt

Anthropogene Emissionen von Schadstoffen bewirken grob gesprochen eine Zunahme der Konzentration von Ozon in der Troposphäre, wo es hauptsächlich schädliche Wirkungen entfaltet. Die Fluor-Chlor-Kohlenwasserstoffe vor allem führen jedoch gleichzeitig zu vermehrtem Abbau der stratosphärischen Ozonschicht, die als Filter gegen harte Strahlung unerlässlich ist. Eine Zusammenfassung der verschiedenen bekannten Auswirkungen einer Änderung der Ozonkonzentrationen auf das Ökosystem findet sich z. B. bei Kondratyev & Varotsos [1996], die unter anderem auf folgende Phänomene hinweisen:

- Die durch anthropogene Emissionen von FCKW und anderen ozonzerstörenden Substanzen bedingte Abnahme des Ozons bewirkt eine erhöhte Dosis von **UV-Strahlung** an der Erdoberfläche. Daraus folgt mit einiger Sicherheit eine Schädigung aquatischer

Ökosysteme, insbesondere der Primärproduzenten (Algen). Der Einfluss auf höhere Pflanzen ist noch nicht hinreichend erforscht, da bei diesen offenbar die Sensitivität extrem schwankt und stark mit anderen Umweltfaktoren (Temperatur, CO₂, Gifte, Trockenheit, Parasiten) verknüpft ist. Beim Menschen bewirkt eine Verringerung der Ozonschichtdicke um 1% eine Zunahme der Hautkrebsfälle um 2.5% [*de Gruijl & van der Leun*, 1993].

- Ozon wirkt (wie z. B. CO₂) als **Treibhausgas**, dessen Verluste in der Stratosphäre zwischen 1980 und 1990 durch die – mengenmäßig geringere – Zunahme des troposphärischen Ozons kompensiert wurden, da es dort effizienter wirkt. Auch nach 1990 Jahrzehnt ist dieser Trend ungebrochen, so dass eine Überkompensation und damit einhergehende globale Erwärmung möglich ist.
- Das **bodennahe Ozon** bewirkt als aggressives Gas in höheren Dosen beim Menschen eine Reizung der Atemwege und der Augen. Auch höhere Pflanzen werden von Ozon angegriffen, was sich u.a. in Wachstumshemmung und verminderter Widerstandsfähigkeit äußert [z. B. *Loibl & Smidt*, 1996].

Kapitel 2

Ozonmessungen

Wie die meisten atmosphärischen Spurengase kann Ozon auf zwei verschiedene Weisen quantitativ nachgewiesen werden. Zum einen besteht die Möglichkeit, über chemische Reaktionen – von denen es wie bereits angedeutet eine große Auswahl gibt – die Ozonkonzentration *in situ* nachzuweisen. Im Zusammenhang dieser Arbeit spielt jener Nachweis nur bei Ozonsonden eine Rolle. Es sei daher für seine nähere Beschreibung auf den Abschnitt 2.1.1 verwiesen, der sich speziell mit Ozonsonden befasst.

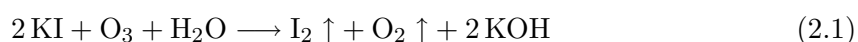
Zum anderen besitzt das Ozonmolekül ausgeprägte Absorptionsstrukturen im infraroten, sichtbaren und ultravioletten Spektralbereich, über die es mit Fernerkundungsmethoden vom Boden oder vom Orbit aus nachgewiesen werden kann. Dies eröffnet eine Vielzahl von Detektionsmöglichkeiten, von denen die Wichtigsten in den restlichen Abschnitten dieses Kapitels behandelt werden.

2.1 Bodengebundene Messungen

Die Arbeiten an der NNORSY-Methode zur Ozonbestimmung stützen sich maßgeblich auf das Vorhandensein qualitativ hochwertiger, bodengebundener Vergleichsmessungen, da diese zur Belernung der verwendeten neuronalen Netze und für die anschließende Validation unumgänglich sind. Mit „bodengebunden“ sind in diesem Zusammenhang Messungen gemeint, die vom Boden aus durchgeführt oder (im Falle von Ballons) gestartet werden, sich aber auf einen größeren Höhenbereich beziehen. Diejenigen Verfahren, die nur das bodennahe Ozon lokal messen, um z. B. Warnungen für die Bevölkerung auszugeben, werden hier nicht diskutiert – Ozonwerte aus den untersten Metern der Atmosphäre sind mit der heutigen Satellitengeneration noch nicht nachweisbar.

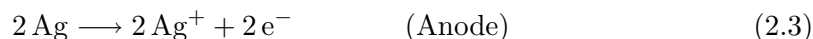
2.1.1 Ballonmessungen

Eine Ozonsonde ist im Prinzip eine fliegende chemische Reaktionskammer, die es im Zusammenspiel mit einer Radiosonde (zur Druck- und Temperaturmessung) und einem Heliumballon erlaubt, die vertikale Schichtung von Ozon *in situ* zu messen. Die verschiedenen Typen von Sonden basieren alle auf derselben Redox-Reaktion mit Kaliumjodid:



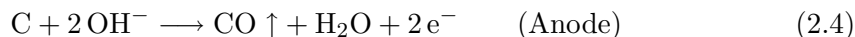
In der Praxis wird bei allen Sondentypen über eine Pumpe Umgebungsluft durch eine Kaliumjodidlösung geleitet. Die Unterschiede zwischen den wichtigsten Sondentypen liegen in der Mechanik und im Nachweis des bei der Reaktion freiwerdenden Jods [SPARC, 1998]:

Brewer-Mast-Sonden [Claude *et al.*, 1987] besitzen eine einzelne elektrochemische Zelle mit alkalischer KI-Lösung, in die ein Silberdraht als Anode und eine Platin-Netzkathode eingetaucht sind. Um die Differenz der elektrolytischen Potenziale auszugleichen, wird eine Gegenspannung von 0.42 V an die Elektroden angelegt. Ein Strom fließt erst, wenn sich die Potenziale durch freigesetztes I₂ (Gl. 2.1) verschieben. Dann finden die Reaktionen



statt. Pro Ozonmolekül fließen also im Idealfall zwei Elektronen durch den externen Stromkreis. Damit ist die nachgewiesene Ozonmenge pro Zeiteinheit direkt dem gemessenen Stromfluss proportional. Eine Variation dieses Sondentyps stellt die sogenannte **Indische Sonde** dar, bei der eine verbesserte, teflonbeschichtete Pumpe zum Einsatz kommt [Attmannspacher & Dütsch, 1970].

Carbon-Iodid („Japanische“) Ozonsonden [JMA, 1991] sind eine Weiterentwicklung des in [Komhyr, 1969] beschriebenen Typs. Hier wird die Silberanode durch Aktivkohle ersetzt und eine gemischte, neutrale KI/KBr-Lösung verwendet. Während die Kathodenreaktion dieselbe ist wie bei den Brewer-Mast-Sonden (Gl. 2.2)¹, tritt an der Anode die Reaktion



auf. Wiederum können zwei Elektronen pro Ozonmolekül nachgewiesen werden. Technisch zeichnet sich dieser Sondentyp durch eine im Vergleich zu den anderen fast doppelt so hohe Pumpleistung aus, zusätzlich wird die Sonde mit bereits integrierten Temperatur- und Drucksensoren ausgeliefert.

Elektrochemical Concentration Cell (ECC) Sonden [Komhyr *et al.*, 1995] besitzen zwei Platin-Netzkathoden, die jeweils in eine eigene Kammer mit KI-Lösung getaucht sind. Die Lösungen sind unterschiedlich konzentriert (1% KI *vs.* gesättigt) und über eine Ionenbrücke verbunden, die die Kaliumionen passieren läßt. Wird nun durch die verdünnte KI-Lösung Ozon gepumpt, so wird das freiwerdende I₂ gemäß Gl. 2.2 ionisiert, während gleichzeitig zwei K⁺-Ionen aufgrund des osmotischen Drucks zwischen den Lösungen von der gesättigten Lösung aus durch die Barriere geschoben werden. Als Ausgleich findet an der Anode in der gesättigten Lösung in der Summe die Reaktion



statt, so dass wiederum zwei Elektronen durch den externen Stromkreis wandern.

Für die Entwicklung des NNORSY-Verfahrens kamen alle drei Sondentypen zum Einsatz. Die Sondenmessdaten der meisten Stationen des weltumspannenden *Global Atmospheric Watch* (GAW) Programmes der *World Meteorological Organization* (WMO) werden an das **World Ozone and Ultraviolet Data Center (WOUDC)** in Ontario, Kanada übermittelt und dort nach eingehender Qualitätskontrolle archiviert [WOUDC, 1990; Wardle *et al.*, 1998; Hare & Fioletov, 1998]. Sie sind für die wissenschaftliche Gemeinde

¹ggf. steht statt Jod auch Brom

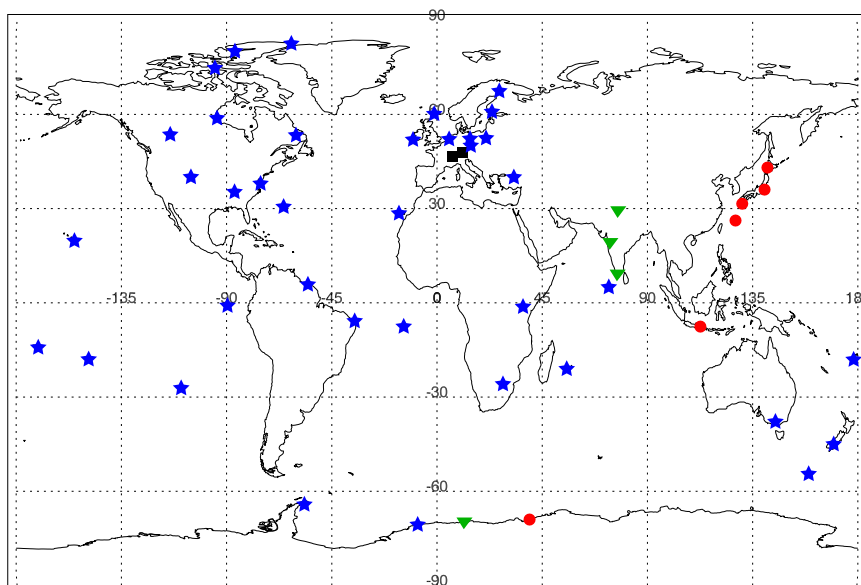


Abbildung 2.1: Verteilung der im WOUDC registrierten Ozonsonden-Stationen. Die ECC-Stationen zwischen $30^{\circ}S$ und $10^{\circ}N$ gehören der SHADOZ-Kampagne an. Symbole bezeichnen die Sondentypen: Stern = ECC, Quadrat = Brewer-Mast, Kreis = Carbon-Iodid, Dreieck = Indische Sonde

frei verfügbar². Anhang A listet die im WOUDC erfassten Stationen, für die im hier untersuchten Zeitraum (1996 bis Mitte 2001) Sondendaten in nennenswerter Anzahl vorhanden sind.

Die Verteilung der Sondenstationen auf der Erdoberfläche ist in Abb. 2.1 dargestellt. Aus geopolitischen und logistischen Gründen konzentrieren sich die Messungen in der Nordhemisphäre. Zwar existierten in den Tropen und Subtropen der Südhemisphäre einige Messstationen, ein direkter Vergleich der Daten wird jedoch zumeist durch unterschiedliche Messhäufigkeiten und verschiedene Meldeverfahren erschwert.

Um diesen Mangel auszugleichen, wurde 1998 die *Southern Hemisphere Additional Ozonesondes (SHADOZ)* Kampagne ins Leben gerufen, die neben einer Koordination der Sondenstarts in einigen Fällen zusätzliche Daten zur Verfügung stellt und ein zentrales Archiv für Sondendaten bietet. Diese Stationen sind eine ideale Datenquelle für die Belerung neuronaler Netze mit Ozonprofilen, da sie eine Lücke in der geografischen Abdeckung weitgehend schließen und ihre Qualität so gut ist, dass praktisch keine Fehlerüberprüfungen mehr durchgeführt werden müssen [Thompson *et al.*, 2001]. Die SHADOZ-Daten werden daher bei allen Belernungen, die den o. g. Zeitraum umfassen, verwendet.

Eine Zusammenstellung quantitativer Fehlerschätzungen für die Ozonsonden-Messungen findet sich in Abschn. 10.2.1. Generell werden Messfehler im Bereich 5–20% berichtet, diese schwanken aber je nach Sondentyp, Höhe und Literaturquelle beträchtlich. Die Gründe für das unterschiedliche Fehlerverhalten der Sonden sind vielfältig und noch nicht in allen Einzelheiten erforscht; Im Gegensatz zu Satelliteninstrumenten können Sonden relativ leicht ausgetauscht und durch andere Typen ersetzt werden, die zudem eine ständige Weiterentwicklung zeigen und deswegen oft nicht sofort genau genug charakterisiert werden können.

Als eine wichtige Fehlerquelle in Höhen ab etwa 25 km sowie in der Troposphäre hat sich mittlerweile die Pumpeffizienz-Korrektur erwiesen. Sie ist nötig, weil die Pumpe beim

²siehe <http://www.msc-smc.ec.gc.ca/woudc/>

Durchleiten des Ozons durch die Reaktionskammer gegen einen Widerstand von ca. 2 mbar arbeitet [Kerr *et al.*, 1994]. Die Korrektur wurde in der Vergangenheit für verschiedene Sondentypen tabelliert [z. B. Komhyr *et al.*, 1995], kann sich jedoch bei Änderung des Herstellungs- und Kalibrationsverfahrens der Sonden mitunter beträchtlich verschieben [De Backer *et al.*, 1998] und auf diese Weise Fehler im Prozentbereich hervorrufen.

Einen starken Einfluss hat weiterhin der Gesamtozon-Korrekturfaktor, der standardmäßig bei den meisten Sonden angewandt wird, nur bei ECC-Sonden meist nicht [SPARC, 1998]. Hierbei wird das Sondenprofil numerisch integriert und auf eine kollokierten Gesamtozonmessung – meist Dobson- oder Brewer-Spektrometer, u. U. auch TOMS – skaliert. Dies kann zu einer Verzerrung des Profils führen, da durch die Skalierung Absolutwert-Fehler von einer Höhenschicht zur anderen übertragen werden. Hohe Korrekturfaktoren weisen außerdem auf Instrumentenprobleme hin und können als grobes Qualitätskriterium verwendet werden [Logan, 1994].

Insgesamt lässt sich sagen, dass zur Durchführung einer guten Ozonsondenmessung trotz der Einführung verschiedener Standardinstrumente und -verfahrensweisen immer noch ein wenig „Magie“ gehört, die sich in teilweise von Station zu Station unterschiedlichen heuristischen Vorbehandlungen und Korrekturen äußert. Angesichts der für die Belernung von neuronalen Netzen benötigten enormen Datenmengen wurde darauf verzichtet, diese Prozeduren im einzelnen nachzuvollziehen und zu vergleichen. Unter der Annahme, dass die Mess- und Archivierungsteams die Qualität ihrer Sondenmessungen Messungen selber am besten einschätzen bzw. verbessern können, wurde daher die Korrektur der Sondenprofile im Rahmen dieser Arbeit auf ein Minimum beschränkt (vgl. Abschnitt 7.1).

2.1.2 Spektrophotometer-Messungen

Pionierarbeit auf dem Gebiet des optischen Ozonnachweises leistete bereits in den 20er Jahren G. M. B. Dobson [1930], der erstmals spektralphotometrische Methoden einsetzte. Die in Abschnitt 1.2 angesprochene starke Absorption im Hartley-Kontinuum besitzt ein Maximum bei 250 nm Wellenlänge, und fällt im an der Erdoberfläche nachweisbaren UV-B³ über mehrere Größenordnungen steil ab (Abb. 2.2). Diese Flanke ist sehr charakteristisch für Ozon; Dobson nutzte sie aus, indem er zwei verschiedene Wellenlängen im UV-B von einem relativ breitbandigen Doppel-Monochromator messen ließ, um anschließend durch Bildung des Verhältnisses der beiden Werte auf die Gesamtozonsäule schließen zu können [Komhyr, 1980]. Dabei können sowohl die direkte Sonnenstrahlung als auch die Himmelsstrahlung am Zenit verwendet werden, wobei Messungen der Zenitstrahlung größere Unsicherheiten bei der Ozonbestimmung ergeben. Das **Dobson-Spektrophotometer** hat sich als Standardinstrument etabliert und wird heute von weltweit etwa 200 Messstationen zum Zwecke der kontinuierlichen Gesamtozon-Überwachung eingesetzt [Wardle *et al.*, 1998].

Trotz seines Erfolges weist das Dobson-Instrument jedoch einige Schwächen auf, vor allem

- den großen Arbeits- und Zeitaufwand für die häufigen Kalibrationen, weiterhin
- die Notwendigkeit, auch für Routinemessungen ausgebildetes wissenschaftliches Personal zu benötigen und schließlich
- Probleme bei der vollständigen Automatisierung.

A. W. Brewer verbesserte das Dobson-Instrument zum sogenannten **Brewer-Spektrophotometer**, indem er das bis dahin verwendete Prisma durch ein optisches Gitter ersetzte

³hier definiert als Wellenlängenbereich 280–320 nm

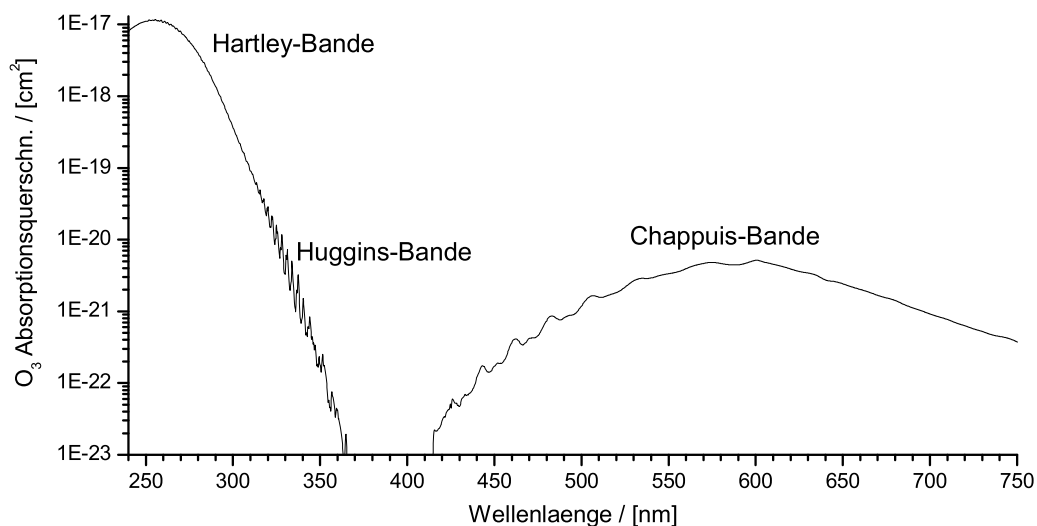


Abbildung 2.2: Absorptionsspektrum des Ozons im sichtbaren und ultravioletten Spektralbereich.

und für die Ozonbestimmung drei weitere Wellenlängen hinzunahm [Brewer & Kerr, 1973]. Mit diesem Gerät wurden die obigen Schwächen beseitigt.

Die Daten des Ozon-Bodenmessnetzwerkes der WMO werden wiederum zentral im WOUDC archiviert, sind allerdings derzeit nur als Tagesmittelwerte verfügbar. An einer genaueren Archivierung der Einzelmessungen wird gearbeitet [E. Hare, WOUDC, persönliche Mitteilung, 2000]. Der mittlerweile reiche Datenbestand an Dobson- und Brewer-Messungen⁴ gab Anlass zu einer Reihe von Vergleichsstudien [Kerr et al., 1988; Basher, 1995; Bojkov et al., 1995, und Referenzen darin], die wiederum Grundlage für eine umfassende Reevaluierung des Datenbestandes in den 90'er Jahren bildeten [WMO, 1993; Fioletov et al., 1999]. Der Fehler einzelner Messungen für beide Instrumententypen beträgt etwa 1% bei Sonnen- und 3% bei Zenitmessung [Basher, 1982; Kerr et al., 1988; Fioletov et al., 1999]. Komhyr et al. [1997] gibt die Genauigkeit von Monatsmittelwerten von 16 Dobson-Instrumenten der NOAA mit $\pm 1.5\%$ an. Die Langzeitstabilität der Geräte wird für Jahresmittel auf $\pm 0.5\%$ geschätzt [WMO, 1980],

2.1.3 Weitere Optische Messverfahren

Außer mit den relativ aufwändigen Doppel-Monochromatoren läßt sich das Ozon vom Boden aus auch aus **Filterphotometer**-Messungen bestimmen, die vor allem in der früheren Sowjetunion häufig durchgeführt wurden und zum Teil in das WOUDC-Archiv übernommen wurden. Diese Geräte besitzen laut Fioletov et al. [1999] eine Genauigkeit von etwa 3% (Sonne) bzw. 5% (Zenit).

Eine wichtige aus dem Dobson/Brewer Verfahren hervorgegangene Methode ist die **Differenzielle Optische Absorptions-Spektroskopie (DOAS)**, die erstmals von Perner & Platt [1979] angewandt wurde. Sie bildet auch die Grundlage für die operationelle Gesamt ozonbestimmung aus GOME-Daten und wird deshalb in Abschnitt 3.4.2 eingehender besprochen.

⁴Die Namen der Erfinder werden i. F. wie üblich synonym für die entsprechenden Instrumente verwendet.

2.2 Satelliteninstrumente

Aufgrund der Möglichkeit, innerhalb von Stunden oder Tagen eine mehr oder weniger globale Abdeckung zu erhalten, wurden schon relativ bald die ersten Ozonmessungen von Satelliten aus durchgeführt. Dieser Abschnitt wird die in der vorliegenden Arbeit verwendeten Satellitenmessungen kurz charakterisieren. Einzelheiten zur Zusammenstellung von Trainings- und Testdatensätzen aus diesen Daten werden dann in Teil II der Arbeit diskutiert, während genauere Angaben zur Präzision der Satelliten-Ozonprofile in Kapitel 10 zu finden sind.

Für satellitenbasierte Ozonmessungen kommen zunächst im Wesentlichen drei verschiedene Beobachtungsgeometrien in Betracht: Bei sogenannten **Nadir-Messungen** wird die aus Richtung des Erdbodens kommende Strahlung ausgewertet, wobei gleichzeitig noch mittels eines Schwenkspiegels senkrecht zur Flugrichtung andere Bodenabschnitte abgetastet werden können. Nadir-Sounder besitzen daher gewöhnlich eine viel höhere Auflösung in der Horizontalen, können aber die vertikale Schichtung des Ozons nicht so gut erfassen.

Bei der **Limb-Geometrie** betrachtet das Instrument hingegen den Rand der Erdscheibe, die Blickrichtung ist also tangential zur Erdoberfläche. Durch Veränderung der Höhe des Tangentialpunktes kann so die Strahlung aus verschiedenen Atmosphärenschichten mit hoher vertikaler Auflösung erfasst werden. Allerdings können sie in der Horizontalen nur relativ große, in Blickrichtung sehr langgestreckte Bereiche betrachten, die horizontale Auflösung ist also weniger gut. Einen Spezialfall der Limb-Geometrie bilden die **Okkultationsmessungen**, bei denen durch den Rand der Atmosphäre eine Lichtquelle beobachtet wird, meist die Sonne. Sie weisen meist ein besseres S/N-Verhältnis auf, jedoch sind die Messungen nur bei Sonnenauf- bzw. Untergang möglich, was die Anzahl der Messungen drastisch herabsetzt.

Tabelle 2.1: *Wichtige orbitale Ozonsensoren nach Beobachtungsgeometrie und Spektralbereich. Kursiv gedruckte Instrumente werden weiter unten beschrieben. MLS ist der Microwave Limb Sounder [Barath et al., 1993]. Das SCIAMACHY Instrument [Burrows et al., 1995] kann zwischen den Beobachtungsgeometrien wechseln [Bovesmann et al., 1999]. Es komplementiert die ebenfalls auf dem kürzlich gestarteten Satelliten ENVISAT installierten Sensoren GOMOS (stellare Okkultation) und MIPAS (ein Interferometer).*

Spektralbereich	Nadir	Limb	Okkultation
UV/Vis	<i>TOMS, SBUV, GOME, SCIAMACHY</i>	SCIAMACHY	<i>SAGE, POAM, GOMOS, SCIAMACHY</i>
therm. IR	TOVS	MIPAS	<i>HALOE</i>
Mikrowellen	–	MLS	–

Weiterhin gibt es mehrere Möglichkeiten für die Wahl des Spektralbereiches, da Ozon zusätzlich zu den bereits angesprochenen Dissoziations-Kontinua im **UV/sichtbaren** auch Rotations-Schwingungsbanden im **thermischen Infrarot** bei 9.1, 9.6 und 14 μm , sowie diverse starke Rotationsbanden im **Mikrowellenbereich** um 200 GHz aufweist. Auch dies führt zu spezifischen Vor- und Nachteilen: Während der Informationsgehalt der UV/sichtbaren Banden höher ist, benötigen sie die Sonne als Lichtquelle. Mikrowellen- und thermische Infrarotinstrumente messen direkt die vom Erde/Atmosphäre-System emittierte Strahlung. Sie sind nicht direkt auf die Sonne angewiesen und können daher auch bei (Polar-)

Nacht messen. Tabelle 2.1 zeigt eine Übersicht derzeitiger orbitaler Ozonsensoren nach Kategorien, ohne Anspruch auf Vollständigkeit.

2.2.1 TOMS

Das *Total Ozone Mapping Spectrometer* (TOMS) ist heute das bekannteste und am besten validierte Instrument, das dediziert für die Ozonüberwachung im Einsatz ist. Es misst die von der Atmosphäre zurückgestreute UV-Strahlung (*Backscatter-UV* (BUV) Prinzip) im Nadir. Auf dem selben Prinzip baut auch das GOME-Instrument auf, daher wird das Verfahren im Abschnitt 3.3 in Zusammenhang mit GOME näher erläutert.

Gestartet wurde das erste Gerät der TOMS-Serie auf dem Satelliten Nimbus-7 im Jahre 1978, es lieferte bis 1994 brauchbare Daten. Bereits 1991 folgte ein zweites Instrument auf Meteor-3 [Herman *et al.*, 1996], der Satellit fiel jedoch im Dezember 1994 aus. Daraus resultierte eine Datenlücke von 18 Monaten, bis im Juni 1996 das nächste TOMS-Instrument auf dem Satelliten Earth Probe in Betrieb ging (EP/TOMS, i. F. mit EPT abgekürzt) [McPeters *et al.*, 1998]. Ein weiteres Gerät, nur wenige Monate später gestartet, fiel schon Ende 1997 dem Versagen seines Trägersatelliten ADEOS zum Opfer. Die im September 2001 gestartete QuikTOMS-Mission, die die Fortsetzung der Zeitreihe garantieren sollte, stand ebenfalls unter keinem guten Stern. Durch ein Raketenversagen beim Start erreichte der Trägersatellit keinen stabilen Orbit.⁵

Die Präzision der EPT-Daten wird mit etwa 2% angegeben [McPeters *et al.*, 1998]. Die Ozonwerte wurden in der Vergangenheit regelmäßig mit Bodenmessungen des Gesamt Ozons verglichen [Fioletov *et al.*, 1999; Lambert *et al.*, 1999; Brinkma *et al.*, 2000]. Dabei fällt auf, dass es systematische Abweichungen gibt, die hauptsächlich von Jahreszeit und geografischer Breite abhängen. Bodeker *et al.* [2001] parametrisierten diese Abweichungen aufgrund statistischer Analysen und entwickelten eine analytische Korrekturfunktion für TOMS. Die Form dieser Funktion sowie ihre Koeffizienten sind dem Anhang B zu entnehmen. Die Abweichungen erreichen im südpolaren Sommer bis zu 20 DU. Sie sind vermutlich größtenteils in der Klimatologie begründet, die für das Retrieval von TOMS verwendet wird und symmetrisch zum Äquator ist [Wellemeyer *et al.*, 1997; Lambert *et al.*, 1999]. Weitere Gründe sind die Schwierigkeiten, bei sehr hellen Szenen Wolken von Eis zu unterscheiden, aber möglicherweise auch systematische Fehler in den Dobson-Messungen [Bodeker *et al.*, 2001]. Eine deutliche Verringerung der Offsets wird mit der neuen Version 8 der TOMS-Daten erwartet [P. K. Bhartia, *persönliche Mitteilung*, 2002], die für die Belernungen jedoch noch nicht zur Verfügung stand.

Für die vorliegende Arbeit wurden die TOMS-Daten korrigiert, um die Homogenität des Datensatzes zu erhöhen (siehe dazu Abschn. 6.1). Wenn wie erwartet die Übereinstimmung der V8-Daten mit Bodenmessungen besser ist als die der Version 7, wird sich diese Korrektur vermutlich in der nächsten NNORSY-Version erübrigen.

2.2.2 SAGE II

Die SAGE-Instrumente (*Stratospheric Aerosol and Gas Experiment*) messen bei acht Wellenlängen zwischen UV und nahem IR (Tab. 2.2) in solarer Okkultation. Hierbei verfolgt der Sensor, wie der Tangentialpunkt zwischen Sonne und Satellit bei jedem Sonnenauf- und -untergang durch die verschiedenen Atmosphärenschichten wandert. Bedingt durch die Orbitgeometrie kann SAGE II jeweils etwa 15 solche Ereignisse pro Tag beobachten. Ein Vorteil dieser Methode ist ihre weitgehende Selbstkalibration, da alle Werte relativ zur Tangentialpunkthöhe 200 km, d. h. außerhalb messbarer Gaskonzentrationen, normiert werden und sich etwaige Degradationseffekte somit weitgehend aufheben.

⁵Quelle: TOMS Webpage <http://toms.gsfc.nasa.gov/news/news.html>

Tabelle 2.2: Charakteristika der im Rahmen der Arbeit verwendeten Okkultations-Instrumente.

	HALOE	SAGE II	POAM III
Latitude	70°S–70°N	70°S–70°N	54°N–71°N 63°S–88°S
Kanäle [μm]	2.45, 2.80, 3.40, 3.46, 5.26, 6.25, 10.04	.385, .448, .452, .525, .600, .935, 1.02	.354, .440, .442, .603, .761, .779, .922, .936, 1.018
Orbithöhe	603 km	610 km	833 km
Aktuelle Version	19	6.10	3.0
Höhenreferenz	0.3 km	0.25 km	1.0 km
Vertikale Auflösung	2.4 km	1.2 km	0.8 km

Die Ozonbestimmung erfolgt im Bereich der Chappuis-Bande bei etwa $0.6 \mu\text{m}$. Allerdings müssen zu diesem Zweck zunächst die Beiträge anderer Extinktionsquellen herausgerechnet werden. Die ersten derartigen Algorithmen für SAGE [Chu *et al.*, 1989] lösten hierzu näherungsweise das Gleichungssystem

$$\tau(\lambda_i, z_k) = \tau_{\text{aer}}(\lambda_i, z_k) + a_i \tau_{\text{O}_3}(\lambda_i, z_k) + b_i \tau_{\text{NO}_2}(\lambda_i, z_k) \quad (2.6)$$

für alle sieben gemessenen Spektralwerte λ_i und 70 Höhenschichten z_k . Dabei ist τ die optische Dichte der entlang des Strahlungspfades von der Sonne zum Sensor liegenden Atmosphäre, die Tiefstellungen aer , O_3 und NO_2 bezeichnen die Beiträge von Aerosolen, Ozon und NO_2 , respektive. Die Konstanten a_i und b_i sind tabelliert als die Verhältnisse von Ozon- und NO_2 -Absorptionsquerschnitt an der Stelle λ_i zu deren Maxima. Der Verteilung der Aerosol-Partikelgrößen werden vier Freiheitsgrade zugestanden, die jedoch unter verschiedenen Annahmen so eingeschränkt werden, dass das resultierende Aerosol eine langsam mit der Wellenlänge variierende Absorption besitzt [Cunnold *et al.*, 1989].

Nachdem die schrägen Absorptionspfade berechnet wurden, wird anschließend das Ozonprofil mit einem iterativen, nichtlinearen Inversionsalgorithmus ermittelt, bei dem die Absorberkonzentration in den Atmosphärenschichten von außen nach innen bestimmt wird (*onion peeling* Verfahren). Die Referenzhöhe des Profils liegt bei 30 km und wird geometrisch aus den Satelliten-Ephemeriden und der Scanspiegel-Ausrichtung errechnet.

Obwohl SAGE II bei seinem Start im Jahr 1984 nur für eine Betriebsdauer von zwei Jahren ausgelegt war, liefert das Instrument auch heute noch qualitativ hochwertige Daten. Diese lange Zeitreihe macht es zusammen mit seiner Fähigkeit zur Autokalibration insbesondere für Ozontrendanalysen sehr nützlich [z. B. Wang *et al.*, 1996; SPARC, 1998; Cunnold *et al.*, 2000b,a], weswegen ein hoher Aufwand mit der Verbesserung der Produktqualität betrieben wurde. Am problematischsten ist beim oben beschriebenen Retrieval-Algorithmus der Beitrag des Aerosols, da dessen Wellenlängenabhängigkeit weder wie die Rayleigh-Streuung physikalisch (fast) exakt berechnet, noch wie bei den Spurengasen im Labor durch Spektroskopie bestimmt werden kann.

In neuerer Zeit zeigten *Steele & Turco* [1997], dass die in den sieben SAGE-Spektralkanälen enthaltene Information prinzipiell nicht ausreicht, um den Aerosolbeitrag für die Extinktion des 0.6 μm -Kanals eindeutig zu bestimmen. Ihre Empfehlung, die Aerosolparameter auf historisch gemessene Werte zu beschränken, ging in die Produktversion v5.96 ein [SPARC, 1998], was die absolute Genauigkeit der Ozonprofile unterhalb von etwa 20 km Höhe verbesserte [Rusch *et al.*, 1997]. Gemäß *Cunnold et al.* [2000b] überschätzt v5.96 die Ozonkonzentration zwischen 15–20 km im Vergleich zu Sondenmessungen und HALOE aber immer noch systematisch um 5–10% in den mittleren Breiten und 10–30% in den Tropen, was nicht allein durch die fehlerhafte Aerosolinformation erklärt werden kann. Beim Übergang zu v6.00 konnte ein Teil der Probleme unterhalb von 20 km durch Verbesserungen in den geometrischen Berechnungen behoben werden, so dass die Abweichungen um etwa ein Drittel schrumpften [Cunnold *et al.*, 2000b]. Weiterhin liegen SAGE Ozonkonzentrationen oberhalb von etwa 48 km für den Zeitraum 1994–1996 systematisch um 5–10% über den von HALOE gemessenen, aber da es in dieser Höhenregion keine Sondenmessungen mehr gibt, ist unklar, welche Daten näher an der wahren Konzentration liegen. *Newchurch et al.* [2000] fanden bei der Analyse von Ozontrends bis 50 km Höhe eine gute Übereinstimmung der SAGE (v5.96) Daten mit Umkehr- und SBUV/2-Messungen, sowohl was die Ozonabnahme als auch die jährlichen Variationen angeht.

Die seit Anfang 2002 ausschließlich zur Verfügung stehenden v6.10 Ozonprofile haben laut Dokumentation auf der SAGE II Homepage⁶ weitere Verbesserungen erfahren, über deren Umfang jedoch noch keine Untersuchungen zu finden waren. Der aktuelle Algorithmus verwendet maximal 3 Aerosolparameter, sowie maximal 6 der 7 SAGE-Kanäle, da Kanal 7 (386 nm) systematische Messfehler aufweist. Für jede Höhenschicht werden zusätzlich alle Kanäle mit einer optischen Dicke größer 7.5 ausgeschlossen, so dass das Gleichungssystem 2.6 nunmehr je nach Atmosphärenzustand unterschiedlich aussieht. Die für die Berechnungen der Rayleigh-Streuung notwendigen Temperatur- und Druckprofile entstammen den NCEP⁷-Analysen.

2.2.3 HALOE

Das *Halogen Occultation Experiment* (HALOE) wurde an Bord des NASA *Upper Atmosphere Research Satellite* (UARS) in den Orbit gebracht und liefert seit Mitte Oktober 1991 Daten [Russell *et al.*, 1993]. Die Beobachtungsgeometrie von HALOE entspricht weitgehend der von SAGE, HALOE misst jedoch nicht im UV/sichtbaren Spektralbereich, sondern bei sieben Wellenlängen im mittleren Infrarot (Tab. 2.2).

Das Verfahren zur Bestimmung des Ozonprofils beruht ebenfalls auf demselben Prinzip wie bei SAGE, nur dass die Ozonabsorption in der Schwingungsbande bei 10.04 μm bestimmt wird. Die HALOE-Ozonprofile besitzen eine vertikale Auflösung von etwa 2.4 km. In der Horizontalen ergeben sich bedingt durch die Scangeometrie Pixel von 5 km Breite und 200–400 km Länge. In den HALOE-Daten enthalten sind außerdem unter anderem Temperatur- und Druckprofile, die jedoch unterhalb von etwa 35 km Höhe aus NCEP-Analysen stammen.

Wie bei SAGE verschiebt sich die Latitudenregion, in der Sonnenauf- und -untergänge beobachtet werden im saisonalen Rhythmus: Geografische Breiten zwischen 55°S und 55°N werden über das Jahr hinweg weitgehend gleichmäßig abgetastet, während die Regionen bis $\pm 70^\circ$ nur zu bestimmten Jahreszeiten abgetastet werden. Dies muss bei Trendanalysen berücksichtigt werden [Remsberg *et al.*, 2001], spielt aber angesichts der anderen Fehlerquellen und der Datenquellen-Mischung bei der Erstellung von Trainingsdatensätzen für neuronale Netze eine untergeordnete Rolle.

⁶<http://www-sage2.larc.noaa.gov>

⁷National Centers for Environmental Prediction

Für die Genauigkeit der Ozonsäulen ist bei HALOE die Modellierung der Aerosolextinktion – im mittleren Infrarot hauptsächlich Absorption – im Wellenlängenbereich der 10.04 μm -Ozonbande maßgeblich. *Hervig et al.* [1995] fanden, dass die Aerosolkorrektur bei dieser Wellenlänge am besten alleine durch eine Extrapolation des Wertes bei 5.26 μm geschieht, da der Algorithmus dann die geringste Sensitivität gegenüber Partikelgrößenverteilung und chemischer Zusammensetzung des Aerosols aufweist. Offenbar sind die Aerosolextinktionen der HALOE-Kanäle so stark linear korreliert, dass die Hinzunahme der anderen Kanäle die Extrapolationsfehler zwar für spezielle Konstellationen senken kann, aber dann bei weitem größere Fehler auf dem Datensatz aller möglichen Atmosphärenzustände aufweist [*Steele & Turco*, 1997].

Der operationelle Algorithmus für die aktuelle Produktversion V19 funktioniert demnach für die meisten Fälle sehr gut, neigt aber bei großer optischer Dicke des Aerosols zur Unterschätzung der Aerosolextinktion um bis zu 20%, weshalb unabhängig voneinander verschiedene Algorithmen für das Screening der Ozonprofile auf der Basis dieser Extinktion entwickelt wurden [*Bhatt et al.*, 1999; *Hervig & McHugh*, 1999].

Eine detaillierte Fehlerabschätzung der HALOE V17 Ozonprofile findet sich bei *Bruehl et al.* [1996]. Für etwa 60–30 km Höhe (0.1–10 hPa) geben die Autoren den RMS-Gesamtfehler mit 8–12% an, darunter steigen die Fehler auf bis zu 30% bei 16 km (100 hPa). Beim Übergang zur Datenversion V18 wurden u. a. algorithmische Verbesserungen der Altitudenregistrierung eingeführt, wodurch die HALOE-Ozonwerte systematisch größer wurden und die Abweichungen von Ozonsonden und SAGE II auf maximal 20% (im Bereich 15–20 km) schrumpften [*Lu et al.*, 1997]. *Bhatt et al.* [1999] verglichen die V18 Daten nach der Anwendung ihrer Screening-Methode auf der Basis von 2.5 km dicken vertikalen Schichten mit Ozonsonden. Sie fanden Abweichungen von unter 10% bis hinunter auf 16 km in den Tropen und Subtropen. Außerhalb dieser Region betragen die Fehler unter 10% bis hinunter auf 11 km (200 hPa) und unter 20% bis 9 km (300 hPa), wobei nur wenige Profile nach dem Screening so weit hinunter reichen.

2.2.4 POAM II/III

Die weniger bekannten POAM-Instrumente (*Polar Ozone and Aerosol Measurement*) messen in solarer Okkultation bei neun Filterkanälen zwischen 350 und etwas über 1000 nm Wellenlänge (Tab. 2.2). POAM II [*Glaccum et al.*, 1996] flog auf dem französischen SPOT 3 Satelliten und lieferte Daten von September 1993 bis November 1996. POAM III [*Lucke et al.*, 1999] ist auf dem Satelliten SPOT 4 installiert und seit April 1998 in operationellem Betrieb. Die beiden Instrumente unterscheiden sich zwar in der technischen Ausführung, sind aber was die wissenschaftlichen Daten angeht fast identisch [*Lucke et al.*, 1999].

Aufgrund der anderen Orbitgeometrie der SPOT-Satelliten beobachten die POAM-Instrumente pro Tag durchschnittlich 14.2 Sonnenaufgänge in einem eng begrenzten Latitudenband, das im Jahreslauf zwischen 54°N und 71°N wandert. Sonnenuntergänge werden mit gleicher Häufigkeit zwischen 63°S und 88°S beobachtet. Die POAM-Ozonprofile dringen somit um bis zu 18° weiter in die Südpolregion vor als HALOE oder SAGE, was der Hauptgrund für die Aufnahme dieses Datensatzes in die Kollokationen ist. Die einzelnen Messungen bedecken eine Fläche von etwa 6 km Breite und 200 km Länge.

Die POAM-Ozonprofile werden auf ähnliche Weise wie SAGE auf der Basis von Gl. 2.6 bestimmt, allerdings erfolgt die näherungsweise Lösung des nichtlinearen Gleichungssystems für die optischen Dicken mit Hilfe des Optimal Estimation (OE) Algorithmus' [*Rodgers*, 1976]. Für die Bestimmung der vertikalen Profile nach dem Abschälverfahren (*onion peeling*) genügt aufgrund der Linearität des Problems eine einzelne, weitere OE-Iteration.

Die Profil-Referenzhöhe beträgt 40 km und wird über den Fit einer Strahlungstransfer-

berechnung für Kanal 1 (352.3 nm) an den tatsächlich gemessenen Wert bestimmt. Hierbei dienen wiederum die Temperatur- und Druckprofile aus der NCEP-Analyse als Modell für den Atmosphärenzustand. Die Genauigkeit in der Höhenbestimmung liegt bei etwa 1 km, während die Vertikalaufösung für Ozonprofile mit 0.8 km angegeben wird [Glaccum *et al.*, 1996]. Ein von Deniel *et al.* [1997] durchgeführter Vergleich der Ozonprofile mit Ozonsonden in der Nordhemisphäre ergab zwischen 20 und 25 km Abweichungen von 5–8%, darüber und darunter lagen die Werte um 2–5%, wobei POAM systematisch niedrigere Werte misst als die Sonden. Eine andere Untersuchung [Rusch *et al.*, 1997] fand im Vergleich zu SAGE II (v5.96) und MLS-Daten systematische Abweichungen von bis zu 20% unterhalb von 20 km Höhe. Da SAGE wie in Abschnitt 2.2.2 beschrieben in diesem Höhenbereich generell zu hohe Werte mißt, spricht dies für die Qualität der POAM-Daten.

Für POAM III stellen Lucke *et al.* [1999] einige Vergleiche mit ECC-Sonden und HALOE für 1998 an. Als vorläufige Ergebnisse finden sie eine Übereinstimmung von besser als 10% bis hinunter auf 15 km Höhe, wobei HALOE-Vergleichsdaten nur in der Breitenregion um 60°N vorlagen. Unterhalb von 15 km mißt POAM III tendenziell zu hoch, um bis zu 20% im Vergleich zu den Sondendaten.

Kapitel 3

Das GOME Instrument

3.1 Historische Einordnung

Bereits in den 50er Jahren des 20. Jahrhunderts stellten *Singer & Wentworth* [1957] erste Überlegungen an, wie man die vertikale Verteilung des atmosphärischen Ozons von einer Orbitalplattform aus ergründen könnte. Die Idee dahinter war, das von der Atmosphäre zurückgestreute Sonnenlicht spektral aufgelöst im UV-Bereich zu messen. Die Sonnenstrahlen erreichen oberhalb von etwa 340 nm noch größtenteils den Erdboden, während kürzerwellige Strahlung bis etwa 250 nm immer weniger weit in die Atmosphäre eindringt. Gründe hierfür sind der um vier Größenordnungen ansteigende Ozon-Absorptionsquerschnitt (Abb. 2.2) und die gleichzeitig mit λ^{-4} zunehmende Rayleigh-Streuung durch die Luftmoleküle. Dass die Dichte der Luft und damit die Anzahl der Streuzentren in der Höhe exponentiell abnimmt, führt zu einer weiteren Verstärkung des Effektes. Für eindringende Photonen einer bestimmten Wellenlänge ist die Atmosphäre also zunächst fast transparent, bis sie in einer eng begrenzten Höhenschicht auf stark zunehmende Streuung treffen, von der sie fast wie von einem Spiegel zurückgeworfen werden.

Die Information, die die Photonen aus dieser Schicht mitbringen, machte sich zuerst das *Backscatter UltraViolet* (BUV) Instrument zu Nutze, das im Jahr 1970 an Bord des Forschungssatelliten *Nimbus 7* in den Orbit geschossen wurde [*Heath et al.*, 1973]. Die Serie wurde nach erfolversprechendem Start im Jahr 1978 mit dem *Solar Backscatter UltraViolet* (SBUV) Instrument [*Heath & Park*, 1978], und ab 1984 mit den operationellen SBUV/2-Instrumenten fortgesetzt [*Frederick et al.*, 1986], die bis heute auf den Satelliten der NOAA-Serie im Einsatz sind.

Da die SBUV/2-Geräte nur 12 relativ breite Spektralkanäle besitzen und ihr S/N-Verhältnis nicht ausreicht, um nennenswerte Information aus der Troposphäre zu gewinnen [*SPARC*, 1998], wurde im Jahr 1988 von *Burrows et al.* [1995] ein ausgefeilteres, hochauflösendes orbitales Instrument vorgeschlagen, das *Scanning Imaging Absorption spectroMeter for Atmospheric CHartography* (SCIAMACHY). Es ist zur kombinierten Messung in Nadir- und Limbgeometrie vorgesehen, und kann aufgrund seines breiten Spektralbereiches von 240 nm bis über 2 μm eine große Vielzahl von atmosphärischen Spurengasen nachweisen [*Rozanov et al.*, 1993; *Bovesmann et al.*, 1999]. Die Entwicklung dieses Instruments nahm aufgrund seiner Komplexität und seines neuartigen Konzeptes eine relativ lange Zeit in Anspruch, es wurde erst 2002 als Teil des europäischen Satelliten ENVISAT gestartet.

Um bereits vorher Erfahrungen mit hochauflösenden UV-Nadir-Scannern sammeln zu können, wurde eine vereinfachte Version („SCIA-mini“) des Instruments entwickelt und unter dem Namen *Global Ozone Monitoring Experiment* (GOME) 1995 an Bord des Satelliten ERS-2¹ in den Orbit gebracht [*Burrows et al.*, 1999]. Abgesehen von den ingenieurmäßigen

¹European Remote Sensing Satellite

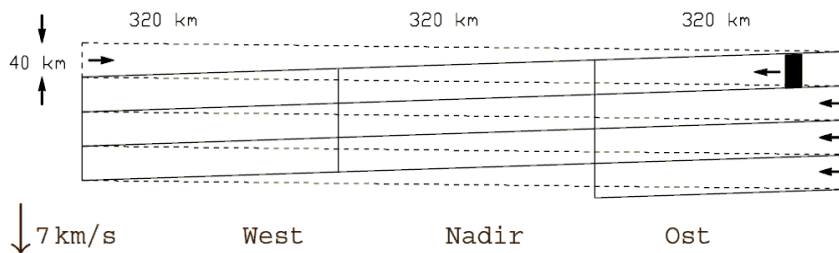


Abbildung 3.1: GOME Abtastmuster im nominellen Scanmodus ($\pm 31^\circ$ Schwenk). Abbildung aus [Bramstedt, 2001].

Aspekten hat das GOME-Instrument auch von der wissenschaftlichen Seite her bereits zu einem enormen Erkenntnisgewinn geführt, und ist bis heute nicht vollständig ausgereizt. Einzelheiten zum Instrument finden sich im Handbuch [ESA, 1995], wobei ein kurzer Überblick an dieser Stelle wiedergegeben werden soll.

3.2 Eigenschaften

GOME bewegt sich auf ERS-2 in einer **sonnensynchronen Umlaufbahn**. Diese Bahn hat eine Neigung von 98.5° gegen die Ekliptik, führt also knapp an den Polen vorbei. GOME überfliegt den Äquator immer um 10:30 Uhr Ortszeit² und braucht für eine Umrundung der Erde etwa 100 min. Genauer berechnet ergeben sich pro Tag exakt $14\frac{11}{35}$ Orbits, so dass das Instrument jeweils nach 35 Tagen (501 Orbits) wieder exakt dieselbe Bahn auf der Erdoberfläche verfolgt.

Der Eintrittsspalt des GOME-Spektrometers sieht zu jedem Zeitpunkt ein *Instantaneous Field of View* von $40 \times 2 \text{ km}^2$ auf der Erdoberfläche. Die Aufnahme eines Spektrums dauert 1.5 s. Das Blickfeld ist gekoppelt mit einem beweglichen Scanspiegel, der innerhalb von 4.5 s quer zur Flugrichtung einen Winkel von 62° überstreicht, um dann innerhalb von weiteren 1.5 s zurückgeschwenkt zu werden. Kombiniert mit der Umlaufhöhe von im Mittel 800 km und einer Nadir-Geschwindigkeit von knapp 7 km/s ergibt sich daraus das in Abb. 3.1 dargestellte **Abtastmuster** (*scan pattern*). Es besteht aus drei *Forward*-Pixeln (Ost, Nadir, West) von etwa $320 \times 40 \text{ km}^2$ und einem *Backscan*-Pixel von $960 \times 40 \text{ km}^2$ Größe. Im Rahmen dieser Arbeit werden nur die *Forward*-Pixel verwendet, da der *Backscan*-Pixel wie zu erkennen eine Art Mittelwert über die gesamte Abtastbreite (*swath*) des Instruments liefert und daher zur Verbesserung der horizontalen Auflösung nicht beiträgt. Abgesehen von der Standardkonfiguration kann GOME auch im sogenannten *polar view mode* messen, der den Scanbereich auf eine Seite verlagert und so eine Abtastung des Pols ermöglicht. Weiterhin gibt es einen *narrow swath* und einen *static scanning* Modus, bei denen die Swathbreite reduziert ist. Diese Modi werden jedoch nur recht selten angesteuert und im Rahmen der NNORSY-Ozonbestimmung nicht gesondert behandelt, da sie durch die Eingangsparameter des Modells ausreichend charakterisiert sind. Einzelheiten zu den Modi finden sich bei [Bramstedt, 2001] oder im GOME-Handbuch.

Die durch den Eintrittsspalt einfallende Strahlung wird in der GOME-Optik mittels Prismen in vier Strahlen zerlegt, die wiederum durch optische Beugungsgitter auf je eine eigene Zeile mit 1024 aktiv gekühlten Silizium-Fotodioden projiziert werden. Daraus ergeben sich vier **Spektralkanäle**, wobei der kurzwelligste, Kanal 1, in die Bänder 1A und 1B

²Geopolitisch bedingte Abweichungen der Zeitzonen von den Meridianen einmal außer acht gelassen.

Kanal	Wellenängen [nm]	Diodenanzahl	$\delta\lambda$ [nm]	t_I [s]	SNR
1A	238–283 (307)	400 (625)	0.2	12	10–434
1B	(307) 283–314	295 (70)	0.2	1.5	434–841
2	311–404	841	0.2	1.5	849–5100
3	394–611	1024	0.4	1.5	3500–4200
4	578–790	1024	0.4	1.5	3239–4214

Tabelle 3.1: *Spektralkanäle von GOME, $\delta\lambda$ bezeichnet die spektrale Auflösung, t_I die Integrationszeit, SNR das Signal-/Rauschverhältnis (nach Eichmann [2001]). Werte in Klammern gelten vor dem 7. Juni 1998.*

unterteilt ist³ (Tabelle 3.1). Band 1A benötigt aufgrund des schwachen Signals im kurzwelligen UV eine längere Integrationszeit, daher werden seine Messwerte nur alle 8 Pixel (12 s) ausgelesen. In der Tabelle sind zum Vergleich auch die Signal-/Rauschverhältnisse angegeben [Eichmann, 2001]. Am 7. Juni 1998 wurde die Grenze zwischen Band 1A und 1B von vorher 307 nm auf 283 nm verlegt, um die horizontale Auflösung dieses Bereichs zu erhöhen. Allerdings nahm man damit auch eine Zunahme des Rauschens in Kauf.

Da die Strahlungssensitivität des GOME-Spektrometers von der Polarisierung des Lichts abhängt, wird ein Teil der einfallenden Strahlung bereits vor der Aufspaltung für die vier Kanäle abgetrennt und in drei **Polarization Monitoring Devices (PMDs)** geleitet. Diese breitbandigen Detektoren decken den gesamten Spektralbereich von GOME ab und ermöglichen die Modellierung einer Polarisationskorrektur. Aufgrund ihrer Breitbandigkeit können die PMDs 16 mal schneller ausgelesen werden als die Diodenarrays. Da ihre Spektralbereiche (300–400 nm, 400–580 nm und 580–750 nm) im Wesentlichen den Grundfarben rot, grün und blau entsprechen, erzeugen die Detektoren außerdem (relativ) hoch aufgelöste Echtfarbbilder der Erde, die z. B. für die Erkennung des Wolkenbedeckungsgrades verwendet werden können.

Der GOME-Monochromator kann aufgrund der Verwendung von Diodenzeilen immer ein ganzes Spektrum auf einmal messen, und ist deswegen im Vergleich zu einfacheren Geräten sehr wartungsfreundlich, da er keine beweglichen Teile enthält. Trotzdem ist die regelmäßige **Wellenlängeneichung** im Orbit für die Langzeitstabilität unerlässlich. Der Scanspiegel kann daher alternativ zur Messung von Erdszenen eine Pt/Cr/Ne-Gasentladungslampe anvisieren, die über definierte Spektrallinien die softwaremäßige Korrektur von Wellenlängendrifts erlaubt.

Eine weitere Scanspiegelposition mißt über eine Diffusorplatte die solare Strahlungsflussdichte und ermöglicht so die Aufnahme von **Sonnenspektren**, welche im Normalbetrieb einmal täglich erfolgt. Für die geophysikalische Auswertung der GOME-Messungen wird meist das Verhältnis von rückgestreuter Erdstrahlung (*radiance*, in Einheiten von [Photonen/(nm cm² s sr)]) zur Strahlungsflussdichte der Sonne (*irradiance*, [Photonen/(nm cm² s)]) gebildet, da sich dann sowohl Veränderungen im Sonnenspektrum als auch viele instrumentelle Effekte herauskürzen.

3.3 Backscatter-UV Theorie

Wie bereits in Abschnitt 3.1 erwähnt, besteht die Idee hinter dem GOME-Instrument in der Ausnutzung der Information über Spurengase, die im rückgestreuten UV- und sichtbaren

³Strenggenommen ist Kanal 2 auch zweigeteilt, aber Band 2A umfaßt nur 9 Dioden und überlappt sich mit Band 1B, so dass es hier nicht verwendet wird.

Licht enthalten ist. Die steile Hartley-/Huggins-Absorptionsflanke des Ozons – im Gegensatz zu vielen anderen Spurengasen – bietet hier den großen Vorteil, dass sie eine Menge Höheninformation enthält, die mit vertretbarem instrumentellen Aufwand extrahiert werden kann [Rozanov *et al.*, 1993].

Um zu verstehen, wo die Ozoninformation im einzelnen verborgen ist, muss das **Strahlungsfeld** der Erdatmosphäre bekannt sein. Sei dazu ϑ der Winkel, unter dem die Strahlung in die Erdatmosphäre einfällt, und ϕ der (Azimut-)Winkel in der Tangentialebene. Die Änderung der am Satelliten bei Wellenlänge λ gemessenen, sonnennormierten Strahlungsflussdichte, $dI(s, \vartheta, \phi)$, bei Durchlaufen einer Wegstrecke ds in der Atmosphäre setzt sich zusammen aus verschiedenen Komponenten:

Extinktionsterm $-(\sigma_a(s) + \sigma_s(s))I(s, \vartheta, \phi)$: Dieser Term umfasst sowohl die Absorption durch Ozon und andere Spurenstoffe, als auch die Rayleigh-Streuung durch Luftmoleküle und die (Mie-)Streuung durch atmosphärische Staubpartikel (Aerosole) und Wolkentröpfchen. σ_a und σ_s werden als Volumen-Absorptions- und Volumen-Streuquerschnitt bezeichnet⁴ und besitzen die Dimension einer inversen Länge, [km⁻¹]. $\sigma_e = \sigma_a + \sigma_s$ ist der Volumen-Extinktionsquerschnitt.

Emissionsterm $+\sigma_a B_\lambda(T)$: Die einfallende Strahlung gewinnt durch die thermische Emission der Luft an Intensität, wobei $B_\lambda(T)$ die Planck-Funktion für die Strahlungsleistung eines schwarzen Körpers ist. Dieser Term kann allerdings im UV und im sichtbaren Spektralbereich vernachlässigt werden.

Diffuse Einstreuung $+\sigma_s \frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} S(s, \vartheta, \phi, \vartheta', \phi') I^*(s, \vartheta', \phi') d\phi' \sin \vartheta' d\vartheta'$: Bei der Berechnung des Strahlungsfeldes liegt die eigentliche Komplikation in diesem Term, da hier die gesamte Strahlung aus der Umgebung des Lichtweges berücksichtigt werden muss. $S(s, \vartheta, \phi, \vartheta', \phi')$ ist die Redistributions- oder Streuphasenfunktion, die beschreibt, welcher Anteil (pro [sr]) der von außerhalb des Lichtwegs aus der Richtung (ϑ', ϕ') einfallenden Strahlung $I^*(s, \vartheta', \phi')$ gerade in die Richtung (ϑ, ϕ) weitergestreut wird. S ist dabei auf 4π normiert.

Diese Darstellung ist bereits sehr vereinfacht; Eine bessere Beschreibung liefert der vektorielle Strahlungstransport, bei dem außer der Strahlungsintensität auch die Polarisation des Lichtes in Form von drei zusätzlichen Komponenten berücksichtigt wird [siehe z. B. Liou, 1980]. Im Zusammenhang mit der Ozonbestimmung ist die vektorielle Berechnung jedoch nicht unbedingt notwendig, wenn die Polarisierungseffekte offline berechnet und im Rahmen der Datenprozessierung geeignet kompensiert werden [Balzer & Loyola, 1996; Tanzi *et al.*, 1999].

In der Retrievalpraxis wird das Strahlungsfeld der Atmosphäre mit Hilfe eines **Strahlungstransportmodells** (*radiative transfer model*, RTM) berechnet. Für GOME wurde zu diesem Zweck am IUP ein eigenes RTM mit dem Namen GOMETRAN entwickelt [Rozanov *et al.*, 1997, 1998]. Es löst die skalare Strahlungstransportgleichung für eine planparallele, horizontal inhomogene Atmosphäre

$$\mu \frac{dI(z, \mu, \phi)}{dz} = -\sigma_e(z)I(z, \mu, \phi) + \sigma_s \frac{1}{4\pi} \int_{-1}^1 \int_0^{2\pi} S(z, \mu, \phi, \mu', \phi') I^*(z, \mu', \phi') d\phi' d\mu', \quad (3.1)$$

wobei z die Höhenkoordinate und μ der Sinus des Sonnenzenitwinkels ϑ ist. Die Gleichung wird dabei mit Hilfe finiter Elemente diskretisiert, so dass im Endeffekt die den Satellit erreichende Strahlung bei diskreten Wellenlängen geliefert wird. Entsprechend entnimmt

⁴Statt Volumen-[\cdot\cdot\cdot]querschnitt ist auch die Bezeichnung [\cdot\cdot\cdot]koeffizient üblich.

das RTM die Streukoeffizienten und -phasenfunktionen einem an diskreten Punkten gegebenen Atmosphärenzustandsvektor (Druck, Temperatur, Spurengaskonzentrationen etc.). Näheres zum Strahlungstransport, sowie speziell zum GOMETRAN-Modell und seiner Weiterentwicklung finden sich in [Buchwitz, 2000; Rozanov, 2001].

Für die folgenden Betrachtungen reicht es festzuhalten, dass durch Umformung der atmosphärischen Strahlungstransportgleichung 3.1 die vom Weltraum aus beobachtete Strahlung ausgedrückt werden kann als [Goody & Yung, 1989]

$$y(\lambda) = \int_{z=0}^{\infty} K(z, \lambda) x(z) dz. \quad (3.2)$$

Dabei muss die Messgröße $y(\lambda)$ nicht zwangsläufig identisch mit dem oben diskutierten Strahlungsfluss I sein, sie kann zum Beispiel auch durch lineare oder logarithmische Skalierung von ihm abhängen. Die gemessene Strahlung entsteht somit aus einem gewichteten Integral über den Atmosphärenzustand x , wobei die Gewichtungsfunktionen $K(z, \lambda)$ auch als *Kerne* des Beobachtungsprozesses bezeichnet werden.

In der Praxis sind die Messungen diskret, der Hilbert-Raum auf dem Gl. 3.2 wirkt, wird dann in einen Zustandsraum \mathbb{R}^n und einen Messraum \mathbb{R}^m übergeführt, so dass

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n \quad \text{und} \quad \mathbf{y} = (y_1, y_2, \dots, y_m)^T \in \mathbb{R}^m. \quad (3.3)$$

Dann bildet ein **Vorwärtsmodell** (*forward model*) \mathbf{F} wie z. B. GOMETRAN den wahren Atmosphärenzustand \mathbf{x} auf die Messung ab \mathbf{y} :

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon} \quad (3.4)$$

Im Vektor \mathbf{b} sind der Einfachheit halber alle zusätzlichen Parameter des Vorwärtsmodells zusammengefasst, während die vektorielle Größe $\boldsymbol{\epsilon}$ den Messfehler beschreibt, der hier als gaußverteilt angenommen wird. Im Allgemeinen ist das Vorwärtsmodell nicht linear, aber da es auf kontinuierlichen physikalischen Zusammenhängen basiert, kann es um einen bestimmten Zustand \mathbf{x}_a herum linearisiert werden, so dass vereinfacht gilt

$$\mathbf{y} - \mathbf{F}(\mathbf{x}_a, \mathbf{b}) = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right|_{\mathbf{x}_a} (\mathbf{x} - \mathbf{x}_a) + \mathcal{O}((\mathbf{x} - \mathbf{x}_a)^2) \approx \mathbf{K}(\mathbf{x} - \mathbf{x}_a), \quad (3.5)$$

wobei

$$\mathbf{K} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \quad (3.6)$$

die sogenannte **Gewichtsfunktionsmatrix** (*weighting function* oder *kernel matrix*) darstellt, die die Auswirkung einer Änderung des Atmosphärenzustandes auf die Messgröße beschreibt.⁵ In Abb. 3.2a sind einige der mit dieser Methode erhaltenen Gewichtsfunktionen dargestellt, d. h. die Zeilenvektoren der Matrix \mathbf{K} . Wie zu erwarten, verschiebt sich das Maximum dieser Funktionen mit zunehmender Wellenlänge in tiefere Schichten der Atmosphäre. Die Funktionen sind negativ, da eine Zunahme der Ozonkonzentration mit einer Abnahme der Messgröße (Strahlung) einhergeht. Der Absolutwert ergibt sich aus einer Kombination von O₃-Absorptionsquerschnitt und -moleküldichte und variiert daher um mehrere Größenordnungen, wie in Abb. 3.2b ersichtlich. Die „Stauung“ der Gewichtsfunktions-Maxima bei etwa 24 km fällt mit der Lage der Ozonschicht in dem betrachteten Szenario zusammen, also der höchsten Molekülanzahldichte des Ozons (hier etwa $5 \cdot 10^{-12} \text{ molec/cm}^3$).

⁵Die Ableitung nach einem Vektor bezieht sich hier auf die Bildung sogenannter Frechét-Ableitungen, d. h. es resultiert eine Matrix mit den Elementen $K_{ij} = \partial F_i / \partial x_j$.

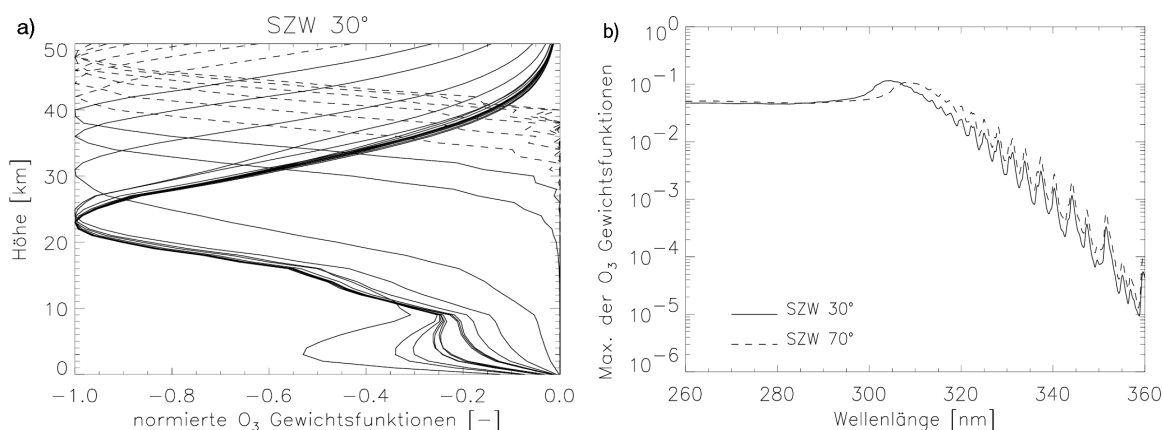


Abbildung 3.2: a) Normierte Ozon-Gewichtsfunktionen für ein typisches Juli-Szenario in 5 nm-Schritten. Von oben nach unten: gestrichelt 260–285 nm, durchgezogen 290–360 nm. b) Betrag des Maximalwerts der unnormierten Gewichtsfunktionen aus a), sowie gestrichelt für das gleiche Szenario bei 70° Sonnenzenitwinkel. Die Absolutwerte beziehen sich auf die logarithmierte, sonnennormierte Radianz als Messgröße (beide Abb. aus [Hoogen, 1998]).

Bei größerem Sonnenzenitwinkel wird durch den längeren Lichtweg die Absorption stärker. Das Gesamtbild ändert sich hierdurch kaum, aber die den Kurven zugeordneten Wellenlängen verschieben sich um einige Nanometer zum Langwelligen hin, und ihre Absolutwerte nehmen zu. Außerdem steigt der troposphärische Beitrag, der durch Mehrfachstreuung hervorgerufen wird, etwas an.

Es ist offensichtlich, dass der Höhenbereich über dem Ozonmaximum durch die Verwendung einiger weniger Wellenlängen hervorragend abgetastet werden kann, wie dies bereits relativ früh mit Hilfe der (S)BUV-Instrumente geschah (Abschn. 3.1). Die Genauigkeit dieser Instrumente reicht aber nicht aus, um genaue Ozonwerte unterhalb des Ozonpeaks, insbesondere in der Troposphäre zu liefern [SPARC, 1998] – wie aus Abb. 3.2b zu entnehmen, ist der Absolutwert der Gewichtsfunktionen, die nennenswerte **troposphärische Beiträge** liefern (ab etwa 330 nm), um zwei bis vier Größenordnungen geringer als bei 310 nm. Auch GOME stößt hier trotz seines guten S/N-Verhältnisses (Tab. 3.1) an die Grenzen seiner Messempfindlichkeit [de Beek, 1998]. Es besitzt jedoch gegenüber den SBUV-Instrumenten zwei entscheidende Vorteile:

1. Die hohe spektrale Auflösung macht es möglich, die Temperaturabhängigkeit der Ozonabsorption auszunutzen, die in den Huggins-Banden die ausgeprägten Absorptionsminima durch Druckverbreiterung auffüllt [McCartney, 1983]. Durch diesen Effekt wachsen die troposphärischen Sekundärmaxima der Gewichtsfunktionen insbesondere bei warmer Troposphäre deutlich an [Chance et al., 1997; Hoogen et al., 1999].
2. GOME mißt auch im Spektralbereich der Chappuis-Banden des Ozons, die zwar betragsmäßig viel schwächer sind als die UV-Beiträge (Abb. 2.2), aber eine unabhängige Informationsquelle über das Gesamt ozon darstellen. Unter bestimmten Umständen (großer SZW, Zustand der Atmosphäre bekannt) kann dies zur Verbesserung der troposphärischen Empfindlichkeit führen [de Beek, 1998].

3.4 Klassische GOME-Ozonbestimmung

Wie erhält man nun aus den oben beschriebenen Eigenschaften der Gewichtsfunktionen und des Vorwärtsmodells eine Lösung \hat{x} für das tatsächliche Ozonprofil? Nehmen wir zunächst

an, das Vorwärtsmodell sei linear und fehlerfrei, dann wird Gl. 3.5 zu

$$\mathbf{y} = \mathbf{K}\mathbf{x}. \quad (3.7)$$

Die Gleichung 3.7 zu invertieren, scheidet als allgemein gültige Lösung aus, da normalerweise $m \neq n$ ist, und somit $\mathbf{K} \in \mathbb{R}^n \times \mathbb{R}^m$ nicht quadratisch. Formal kann dieses Problem durch die Bildung der sogenannten Pseudo-Inversen $\mathbf{K}^\dagger = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T$ [Golub & Kahan, 1965] umgangen werden,

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{y}, \quad (3.8)$$

aber dies führt in der Praxis zu sehr instabilen Lösungen [Rodgers, 1976], d. h. Fehler in den Messungen \mathbf{y} werden extrem verstärkt. Der Grund dafür ist, dass die rechte Seite der dem Problem zugrundeliegenden Gleichung 3.2 zur Klasse der Fredholm-Integrale gehört: Diese besitzen vom mathematischen Standpunkt aus notorisch schlechte Eigenschaften, bei denen die Retrievalprobleme keine Ausnahme sind. Im Allgemeinen ist bei Satellitenmessungen die Anzahl der Messwerte (viel) größer als die der zu ermittelnden Atmosphärenparameter, also $m > n$. Damit wäre \mathbf{x} aus Gleichung 3.7 überbestimmt, was an sich noch keine Schwierigkeiten macht. Gleichzeitig ist jedoch der Rang r von \mathbf{K} meist kleiner als die Anzahl n der zu ermittelnden Komponenten x_i , wodurch \mathbf{x} unterbestimmt wäre – Rodgers [2000] spricht in diesem Zusammenhang von einem *gemischt-bestimmten* Problem.

3.4.1 Optimal Estimation

Mit anderen Worten, mit einem gegebenen Messvektor \mathbf{y} lassen sich beliebig viele Ozonprofile \mathbf{x} in Einklang bringen. Eine exakte Lösung ist somit nicht mehr möglich, das Problem verlagert sich statt dessen darauf, aus dieser Mannigfaltigkeit möglicher Lösungen eine bestimmte auszuwählen. Rodgers [1976] schlug hierzu ein Verfahren vor, das unter der Bezeichnung *Optimal Estimation* (OE) bekannt ist [s. a. Eyre, 1989; Rodgers, 1990, 2000] und in verschiedenen Variationen auch bei GOME Anwendung findet.

Statistische Grundlagen

Ausgangspunkt für diese Methode ist das *maximum likelihood* Prinzip: Die wahrscheinlichste Lösung $\hat{\mathbf{x}}$ ergibt durch Maximieren der **Likelihood-Funktion**

$$\mathcal{L} = p(\mathbf{x}|\mathbf{y}), \quad (3.9)$$

wobei $p(\mathbf{x}|\mathbf{y})$ die Wahrscheinlichkeit für Zustand \mathbf{x} ist, unter der Bedingung, dass \mathbf{y} gemessen wurde. Die Bayessche Formel für bedingte Wahrscheinlichkeiten sagt andererseits aus, dass

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}. \quad (3.10)$$

Sei nun weiterhin angenommen, dass \mathbf{x} um den gewählten Linearisierungspunkt \mathbf{x}_a aus Gl. 3.5 herum gaußverteilt ist, dass also gilt

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{S}_a|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1}(\mathbf{x} - \mathbf{x}_a)\right). \quad (3.11)$$

\mathbf{x}_a wird in diesem Zusammenhang als *a priori* Ozonprofil oder kurz einfach als *a priori* bezeichnet, weil es das Wissen über \mathbf{x} vor Beginn der Messung widerspiegelt. Die Annah-

me der Gaußverteilung geht somit davon aus, dass als Linearisierungspunkt \mathbf{x}_a eine Art mittleres Profil gewählt wird.⁶ Die Matrix

$$\mathbf{S}_a = \langle (\mathbf{x} - \mathbf{x}_a)^T (\mathbf{x} - \mathbf{x}_a) \rangle \quad (3.12)$$

ist offensichtlich symmetrisch und wird als *a priori Kovarianzmatrix* bezeichnet. Sie spielt die Rolle einer auf n Dimensionen verallgemeinerten Varianz der Verteilung. Der Ausdruck

$$\mathcal{D}_n^2 = (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) \quad (3.13)$$

heißt *Mahalanobis-Entfernung* [Crone & Crosby, 1995] und definiert im \mathbb{R}^n Hyperflächen gleicher Wahrscheinlichkeit von \mathbf{x} .

In ähnlicher Weise wird auch die Abweichung der mittels Vorwärtsmodell berechneten von der gemessenen Strahlung als gaußverteilt angenommen:

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{m/2} |\mathbf{S}_y|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{y} - \mathbf{K}\mathbf{x})^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x})\right). \quad (3.14)$$

Zur hier auftretenden *Messfehlerkovarianzmatrix* \mathbf{S}_y tragen sowohl die Messfehler ϵ aus Gl. 3.4 bei, als auch Fehler im Vorwärtsmodell, sowie Linearisierungsfehler.

Definition der Fehlerfunktion

Die Maximierung der Likelihood-Funktion 3.9 wäre mathematisch unnötig aufwändig und läßt sich schneller bewerkstelligen, indem statt dessen eine **Fehlerfunktion**⁷ $E \propto -\ln \mathcal{L}$ minimiert wird. Die Verteilung $p(\mathbf{y})$ hängt nicht explizit von \mathbf{x} ab und ist daher für die Minimierung uninteressant. Ebenfalls weggelassen werden die als additive Konstanten auftretenden Normierungsfaktoren der Gaußfunktionen. Zusammen mit den Gleichungen 3.10, 3.11 und 3.14 ergibt sich dann

$$E = (\mathbf{y} - \mathbf{K}\mathbf{x})^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x}) + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a). \quad (3.15)$$

Der zu minimierende Fehler definiert sich also als die Summe der Mahalanobis-Entfernungen von gemessenem zu berechnetem Spektrum und vom *a priori* zum wahren Profil. Diese Entfernungen werden wiederum kontrolliert von den inversen Kovarianzmatrizen. Grob gesagt trägt somit ein Term umso weniger zur Fehlerfunktion bei, je größer die mit ihm verbundene Unsicherheit ist.

Durch die Annahme von Gaußverteilungen wird in Gl. 3.10 impliziert, dass $p(\mathbf{x}|\mathbf{y})$ ebenfalls einer Gaußverteilung entspricht, die die Abweichungen zwischen Retrievals $\hat{\mathbf{x}}$ und dem wahren Wert \mathbf{x} beschreibt. Man kann also ansetzen

$$\tilde{E} = (\mathbf{x} - \hat{\mathbf{x}})^T \hat{\mathbf{S}}^{-1} (\mathbf{x} - \hat{\mathbf{x}}), \quad (3.16)$$

wobei $\hat{\mathbf{S}}$ die Retrieval-Kovarianzmatrix ist. E und \tilde{E} unterscheiden sich nur durch Normierungskonstanten der Gaußfunktionen, die immer geeignet gewählt werden können. Durch Gleichsetzen von 3.15 und 3.16 und den Vergleich gleicher Potenzen von x erhält man als OE-Lösung den besten Schätzwert

$$\hat{\mathbf{x}} = \mathbf{x}_a + (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x}_a), \quad (3.17)$$

⁶Wenn es mehrere \mathbf{x}_a gibt, handelt es sich bei $p(\mathbf{x})$ streng genommen um die gemeinsame Wahrscheinlichkeit $p(\mathbf{x}_a, \mathbf{x})$. Diese läßt sich durch die Bayessche Formel wieder in $p(\mathbf{x}_a|\mathbf{x})p(\mathbf{x})/p(\mathbf{x}_a)$ zerlegen. Da über die tatsächliche, physikalische Verteilung $p(\mathbf{x})$ meist nichts bekannt ist und $p(\mathbf{x}_a)$ nicht von x abhängt, spielen diese Terme in der obigen Betrachtung keine Rolle und wurden daher unterschlagen.

⁷Die Bezeichnung *Kostenfunktion* ist ebenfalls üblich.

und weiterhin

$$\hat{\mathbf{S}} = \mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{S}_a^{-1}. \quad (3.18)$$

Inverses Modell und Auflösungsmatrix

Wir haben nun über Umwege die Gleichung 3.7 für das lineare Vorwärtsmodell invertiert und dabei impliziert, dass es ein *inverses Modell* \mathbf{R} gibt, welches den Zusammenhang

$$\hat{\mathbf{x}} = \tilde{\mathbf{R}}(\mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon}, \mathbf{b}') = \mathbf{R}(\mathbf{y}, \mathbf{b}, \mathbf{b}', \boldsymbol{\epsilon}) \quad (3.19)$$

beschreibt. Der Vektor \mathbf{b}' faßt hierbei Parameter des inversen Modells zusammen, die nicht bereits in den Vorwärtsmodellparametern \mathbf{b} enthalten sind. Dazu zählt unter anderem das *a priori* Profil \mathbf{x}_a . Im linearen, fehlerfreien Fall ist \mathbf{R} durch die Gl. 3.17 beschrieben, so dass für seine Ableitung gilt

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}} = (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1} =: \mathbf{D}_y. \quad (3.20)$$

Die Matrix \mathbf{D}_y wird als **Beitragsfunktions-Matrix** (*contribution function matrix*) bezeichnet. Ihre Zeilen beschreiben die Reaktion des Retrievalergebnisses auf Änderungen der Messgröße. Das Einsetzen dieser Definition in Gl. 3.17 führt unter Verwendung von Gl. 3.7 – wiederum ohne Messfehler – zu

$$\hat{\mathbf{x}} - \mathbf{x}_a = \mathbf{D}_y(\mathbf{y} - \mathbf{K}\mathbf{x}_a) = \mathbf{D}_y(\mathbf{K}\mathbf{x} - \mathbf{K}\mathbf{x}_a) =: \mathbf{A}(\mathbf{x} - \mathbf{x}_a). \quad (3.21)$$

Hierdurch wird die in der Retrievaltheorie häufig verwendete **Auflösungsmatrix** (*averaging kernel matrix*) $\mathbf{A} = \mathbf{D}_y \mathbf{K}$ definiert [Backus & Gilbert, 1970]. Da wir die direkten Auswirkungen aller anderen Fehler bislang vernachlässigt haben, beschreibt Gl. 3.21 den dem Modell inhärenten Glättungsfehler, der sich durch die Tatsache erklärt, dass sowohl Profil als auch Messgröße durch gaußverteilte Fehler „verschmiert“ sind. Somit gibt \mathbf{A} einen Anhaltspunkt für die Vertikalauflösung des Retrievalverfahrens. Wären im Idealfall hingegen Profil und Messgröße exakt bekannt, dann wären alle Matrizen regulär, \mathbf{A} gleich der Einheitsmatrix und $\hat{\mathbf{x}} = \mathbf{x}$. Als Beispiel für den realen Fall zeigt Abb. 3.3 die Zeilenvektoren von \mathbf{A} (Glättungsfunktionen) eines Retrievals auf dem in Abb. 3.2 dargestellten Szenario [Hoogen, 1998]. Die Glättungsfunktionen sind offensichtlich nicht symmetrisch, können negativ werden, und ihr Extremum weicht höhenmäßig von seiner nominellen Höhe ab, vor allem unterhalb des Ozonpeaks. Dies kann dahingehend interpretiert werden, dass das Retrievalverfahren versucht, das Fehlen von physikalischer Information aus den gemessenen Spektren in Höhenbereichen, in denen die Gewichtsfunktionen sehr klein sind (Abb. 3.2), dadurch auszugleichen, dass es gewissermaßen Information aus besser bekannten Höhenschichten in diese Bereiche verschiebt. Für die 5 km Kurve aus Abb. 3.3 kommt z. B. etwa die Hälfte der Information aus Höhen über 8 km, wobei einige Höhenschichten beträchtliche negative Beiträge liefern.

Nichtlinearer Fall

Die bisherigen Betrachtungen bezogen sich auf den Fall eines linearen Vorwärtsmodells. Für den wesentlich realistischeren Fall nichtlinearer Zusammenhänge zwischen \mathbf{x} und \mathbf{y} muss die Lösung von Gl. 3.17, also das Minimum der Fehlerfunktion E , iterativ gefunden werden, indem bei jedem Iterationsschritt i um einen neuen Zustandsvektor \mathbf{x}_i herum lokal linearisiert wird. Hierzu bieten sich numerisch verschiedene Möglichkeiten an, z. B. Gradientenabstieg, das Newton-Verfahren oder eine Kombination aus beiden, das Marquardt-Levenberg Verfahren. Auf die spezifischen Unterschiede einzugehen würde an dieser Stelle zu weit führen,

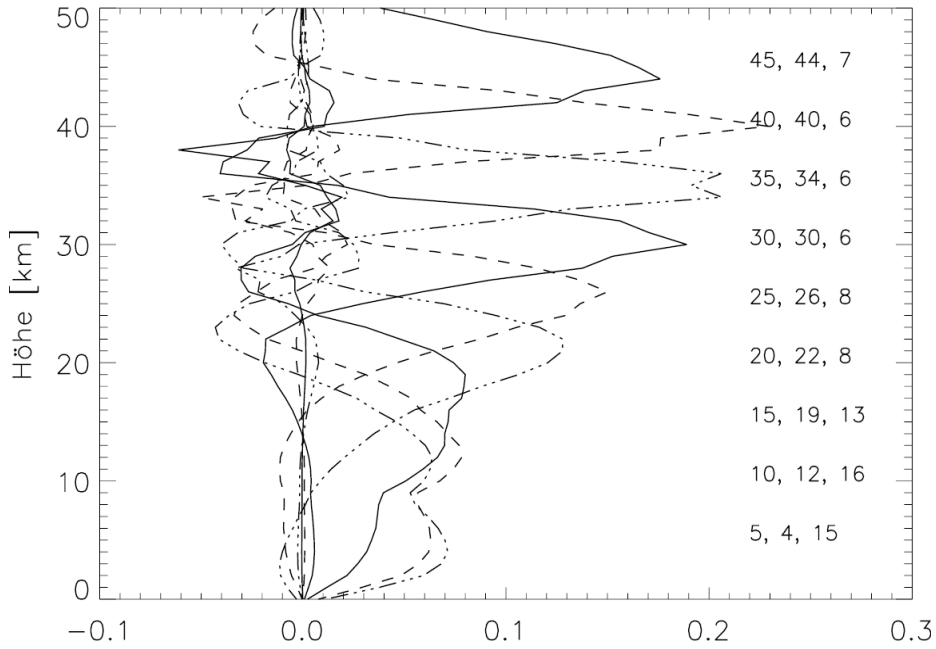


Abbildung 3.3: Glättungsfunktionen (averaging kernels) des GOME OE Retrievals für das Szenario aus Abb. 3.2. Rechts sind in [km] die nominelle und tatsächliche Altitude der Funktion, sowie die volle Halbwertsbreite (FWHM) des Hauptmaximums angegeben.

es sei daher auf die einschlägige Literatur verwiesen [z. B. *Press et al.*, 1986]. Für die am IUP entwickelte Full Retrieval Method (FURM) [*Hoogen et al.*, 1997; *de Beek*, 1998; *Hoogen et al.*, 1999], die das Newton-Verfahren verwendet, ergibt sich als Iterationsschritt

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_y^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1})^{-1} \mathbf{K}_i^T \mathbf{S}_y^{-1} [(\mathbf{y} - \mathbf{y}_i) + \mathbf{K}_i(\mathbf{x}_i - \mathbf{x}_a)]. \quad (3.22)$$

Als Startpunkt der Iteration wird meist $\mathbf{x}_0 := \mathbf{x}_a$ gesetzt, es können jedoch auch bereits erfolgte Retrievals in räumlicher Nachbarschaft oder – wie bei der Datenassimilation üblich – die vorhergesagten Profile eines (Chemie-)Transportmodelles verwendet werden [*Eyre*, 1989].

Fazit

Zusammenfassend läßt sich sagen, dass die Inversion der Strahlungstransportgleichung für Satellitenspektren ein numerisch schwieriges Problem ist, welches sich aber unter diversen Voraussetzungen iterativ lösen läßt, wie anhand des Optimal Estimation Verfahrens demonstriert. Letzteres resultiert in der informationstheoretisch besten Schätzung für den wahren Atmosphärenzustand, hier das Ozonprofil, wenn folgende Annahmen zutreffen:

- Alle Messfehler und Atmosphärenzustände sind gaußverteilt.
- Das Vorwärtsmodell ist biasfrei.
- Die Beziehung zwischen Messspektrum und Ozonprofil ist linear genug, d. h. lokale Linearisierung ist erlaubt.
- Die Kovarianzen der Messfehler, des Vorwärtsmodell-Fehlers und des *a priori* Ozonprofils sind bekannt.

- Die Kovarianzmatrizen sind nicht untereinander korreliert.
- Das Iterationsverfahren findet das globale Minimum der Fehlerfunktion E .

3.4.2 DOAS

Die *Differential Optical Absorption Spectroscopy* (DOAS) beruht – wie der Name schon sagt – nicht auf den Absolutwerten der Strahlungsintensität, sondern auf differenziellen Strukturen im Absorptionsspektrum, d. h. den Anomalien bezüglich eines gleitenden, breitbandigen Mittelwertes. Dieses Prinzip wurde in einfacher Form erstmals von *Noxon* [1975] zur Detektion von NO_2 angewandt, und später von *Platt et al.* [1979] zu der flexibleren DOAS-Methode erweitert. Grundsätzlich wird hierzu ein Referenzspektrum I_0 benötigt, in dem die Strahlung durch keine oder nur durch bekannte Mengen von Absorbern geschwächt wurde, sowie ein Messspektrum I und die Absorptionsquerschnitte α_i aller relevanten Absorber. In einem oder mehreren festen Wellenlängenfenstern wird dann zunächst ein breitbandiger Mittelwert von beiden Spektren abgezogen, z. B. in Form eines Polynomfits niedrigen Grades. Dadurch wird die Methode von der Absoluteichung des Messinstrumentes weitgehend unabhängig, und spektral langsam variierende Effekte wie Rayleigh-Streuung (Luftmoleküle) und Mie-Streuung (Aerosole, Wolken) fallen weg. Für die resultierenden, differenziellen Spektren I' und I'_0 gilt aber immer noch das Lambert-Beer-Bouguer-Absorptionsgesetz,

$$I'(\lambda) = I'_0 e^{-\alpha(\lambda)\rho s} = I'_0 e^{-\sigma(\lambda)s}, \quad (3.23)$$

wobei s die Weglänge des Lichtes, ρ die Dichte und σ den Absorptionskoeffizienten des Absorbers darstellt (vgl. Abschn. 3.3). Daraus ergibt sich durch Logarithmieren

$$\ln I'(\lambda) = \ln I'_0(\lambda) - \sum_i \alpha_i(\lambda)\rho_i s. \quad (3.24)$$

Unter der Voraussetzung, dass diese Gleichung für mehr diskrete Spektralwerte λ_j aufgestellt wird, als es Absorber gibt, erhält man ein überbestimmtes, lineares Gleichungssystem, aus dem die Spurengaskonzentrationen ρ_i durch übliche *least-squares*-Techniken gewonnen werden können. In dieser stark vereinfachten Darstellung wurden jedoch nichtlineare Effekte wie der durch inelastische Raman-Streuung verursachte Ring-Effekt [*Grainger & Ring*, 1962; *Vountas*, 1998] oder eine etwaige ungleiche Wellenlängenkalibrierung der Spektren nicht berücksichtigt. Diese müssen bei der Anwendung des DOAS-Algorithmus' entsprechend beachtet und korrigiert werden.

Da die meisten atmosphärischen Spurengase differenzielle Absorptionsstrukturen aufweisen, eignet sich DOAS sehr gut zu ihrer Detektion. Ursprünglich entwickelt wurde das Verfahren für die Bestimmung der Spurengasdichte entlang wohldefinierter Lichtpfade in Bodennähe [*Platt et al.*, 1979; *Platt & Perner*, 1983; *Platt*, 1994], um z. B. verkehrsbedingte Emissionen nachzuweisen. Später wurde die Methode auch für die Bestimmung der vertikalen atmosphärischen Spurengassäule mit Hilfe von Direktstrahlungs- oder Zenitmessungen eingesetzt [*Solomon et al.*, 1987; *Richter*, 1997; *Eisinger et al.*, 1997; *Müller*, 1997]. Da die DOAS-Methode immer die Absorption entlang des gesamten Lichtwegs detektiert, wird hierbei ein *airmass*-Faktor (AMF) benötigt [*Perliski & Solomon*, 1993; *Marquard et al.*, 2000], der das Verhältnis von schräger Säule S_{sl} zur senkrechten Säule S_{\perp} bezeichnet:

$$\mathcal{A} = \frac{S_{sl}}{S_{\perp}} \quad (3.25)$$

Noch eine Stufe komplizierter werden die Strahlungsverhältnisse bei der Anwendung von DOAS auf Backscatter-UV Spektren, da hier das Licht auf Hin- und Rückweg unter verschiedenen Winkeln durch die Erdatmosphäre dringt. Zudem dient als Hintergrund die

Erdoberfläche, deren Albedo stark variieren kann. Für die GOME-Gesamtozonbestimmung erwies sich DOAS aber auch unter diesen Verhältnissen als geeignet [Burrows *et al.*, 1994]. Der im operationellen *GOME Data Processor* (GDP) implementierte DOAS-Algorithmus behandelt die o. g. Probleme mit unterschiedlichen Methoden [Spurr, 1996], deren Beschreibung hier zu weit führen würde. Als Wellenlängenfenster wurde das Intervall 325–335 nm ausgewählt; Die im Prinzip auch mögliche Bestimmung von Gesamtozon aus den Chappuis-Banden erwies sich hingegen in der Praxis als zu störanfällig [Eisinger *et al.*, 1996; Barga *et al.*, 1999].

Kapitel 4

Neuronale Netze

Neuronale Netze, so wie sie hier Verwendung finden, können selbstverständlich einfach als ein weiteres nichtlineares Regressionsverfahren der Statistik verstanden werden. Dies wird aber den Ideen, die historischermaßen hinter ihrer Entwicklung standen, nicht gerecht. Der Anstoß für die Beschäftigung mit dieser Technik kam eindeutig aus der Biologie, daher möchte ich in Abschnitt 4.1 zunächst auf die Rolle neuronaler Netze in der Natur eingehen. Anschließend wird erläutert, wie die Umsetzung der durch die Evolution optimierten Lösung für das Problem „Informationsverarbeitung“ in stark vereinfachter Form auf einem vergleichsweise extrem primitiven digitalen Rechner erfolgt.

Der Abschnitt 4.2 beschäftigt sich damit, wie dieses digitale Modell sinnvoll eingesetzt werden kann, indem man den bei allen höheren Lebewesen ausgeprägten Prozess des Lernens simuliert. Hinter den Belernungsalgorithmen verbergen sich statistische Konzepte, die auf ähnlichen Annahmen beruhen wie OE und im Abschnitt 4.3 eingeführt werden, bevor Abschnitt 4.4 mit einigen Anwendungsbeispielen aus der Atmosphärenforschung den Bogen schließt.

4.1 Grundlagen

4.1.1 Biologisches Vorbild

Das Neuron

Die Basis für die Signalverarbeitung bildet bei allen höheren Lebewesen die Zellklasse der *Neuronen*. Das menschliche Gehirn besteht aus ca. 10 Milliarden Neuronen, während die Ganglien der neurobiologisch gut untersuchten Meeresschnecke *Aplysia* nur 20000 Neuronen enthalten [Wehner & Gehringer, 1995]. In neuerer Zeit sind sogar Amphibienarten entdeckt worden, die nur etwa 3000 Neuronen besitzen, aber dennoch zu erstaunlichen Leistungen fähig sind.

Obwohl es hunderte verschiedener Arten von Neuronen gibt, lässt sich ein gemeinsamer Grundaufbau feststellen (Abbildung 4.1). Von einem den Zellkern (Nucleus) enthaltenden Zellkörper (Soma) gehen eine Reihe von mehr oder weniger verzweigten Dendriten aus, die der Signalaufnahme dienen. Weiterhin besitzt das Neuron einen Fortsatz, das Axon, der im Vergleich zum mikroskopisch kleinen Soma außerordentlich lang werden kann – beim Menschen einige Zentimeter, bei Giraffen sogar mehrere Meter. Das Axon kann verschiedene Seitenäste aufweisen, die sich alle an ihrem Ende verzweigen und eine Reihe synaptischer Endknöpfchen ausbilden. Es dient im Wesentlichen als eine Art Leitungskabel, welches die vom Neuron ausgesandten Signale in andere Bereiche des Gehirns bzw. Körpers überträgt und dort an andere Zellen weitergibt. Neuronen sind immer in eine Matrix aus Gliazellen eingebettet, die unter anderem Stütz- und Versorgungsfunktion haben. Bei Wirbeltieren

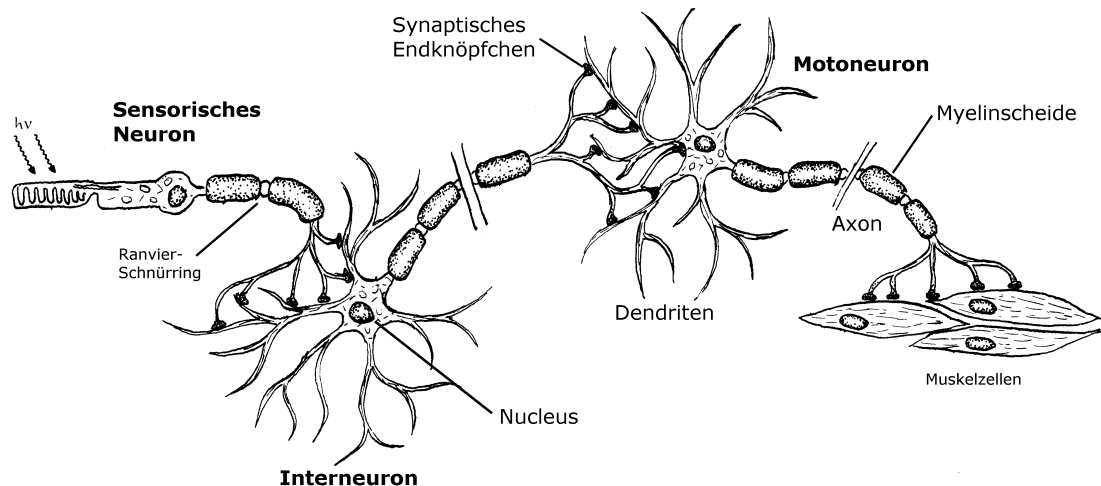


Abbildung 4.1: Neuronentypen und vereinfachtes Verbindungsschema. Umweltreize (z. B. Licht, $h\nu$) werden von sensorischen Neuronen aufgefangen, in einem Interneuron weiterverarbeitet und ggf. ein Motoneuron angewiesen, auf den Reiz durch Muskelkontraktion zu reagieren.

lagern sich spezialisierte Gliazellen an die Axone an und schließen sie in Myelinscheiden ein, die in regelmäßigen Abständen von sogenannten Ranvier-Schnürringen unterbrochen werden.

Funktionell wird zwischen drei Neuronentypen unterschieden (Abb. 4.1):

Sensorische Neuronen (Rezeptoren) nehmen Sinnesreize verschiedener Art auf und wandeln sie in elektrische Signale um.

Interneuronen sind für die Verarbeitung und Weiterleitung von Signalen zwischen Neuronen zuständig und stellen als solche mit Abstand die Mehrzahl aller Neuronen.

Motorische Neuronen (Motoneuronen) nehmen Signale von anderen Neuronen auf und leiten sie als Kommandos an Muskelzellen weiter.

Natürlich ist die Verschaltung der Neuronen selten so einfach wie in der Abbildung: Viele Neuronen sind mit mehreren hundert ihrer Artgenossen verbunden, und der Weg der Signale führt gewöhnlich erst über mehrere Interneuronen zu den Muskeln, wenn überhaupt. Die Einzigartigkeit des Zelltyps der Neuronen besteht aber in der Fähigkeit, zielgerichtet und effizient Erregungssignale empfangen und senden zu können.

Erregungsleitung

Dazu arbeiten die neuronalen Schaltkreise wie ihre elektronischen Äquivalente mit elektrischem Strom als Energieüberträger, jedoch erfolgt die Signalerzeugung und -leitung nicht durch Elektronen, sondern durch Ionen. Diese sind in der Bindegewebsflüssigkeit zwischen den Neuronen (dem Interstitium) und in den Neuronen hauptsächlich in Form von Na^+ , K^+ und Cl^- vorhanden. In der Zellmembran eines Neurons sitzen eine Reihe von Ionenpumpen, die langsam aber kontinuierlich Na^+ nach außen und K^+ nach innen befördern, so dass im Inneren des Neurons Natriummangel und Kaliumüberschuss herrschen. Da die Membran für K^+ -Ionen viel durchlässiger ist als für Na^+ , kann das Kalium dem Konzentrationsgefälle entgegen wirken (osmotischer Druck) indem es nach außen strömt. Es erzeugt dadurch

allerdings lokal eine negative Ladung an der Zellinnenwand, die dem Ausstrom entgegenwirkt, bis sich ein Gleichgewicht zwischen der beim Ausstrom eines Mols zu verrichtenden elektrischen Arbeit W_{el} und der gewonnenen osmotischen Energie E_{osm} einstellt:

$$W_{\text{el}} = E_{\text{osm}} \quad (4.1)$$

$$n \cdot F \cdot \Phi_{\text{K}^+} = R \cdot T \cdot \ln \frac{c_a(\text{K}^+)}{c_i(\text{K}^+)} \quad (4.2)$$

Dabei bedeutet	n	Ionenwertigkeit: +1 für Kaliumionen
	F	Faraday-Konstante: $96500 \frac{\text{C}}{\text{mol}}$
	Φ	elektrisches Potenzial
	R	allgemeine Gaskonstante: $8.3 \frac{\text{J}}{\text{K} \cdot \text{mol}}$
	T	Temperatur
	c_i, c_a	Innen- bzw. Außenkonzentration der Ionen

Für das Gleichgewichtspotenzial Φ_{K^+} ergibt sich bei einer Körpertemperatur von $T = 309\text{K}$ und einem K^+ -Konzentrationsverhältnis von typischerweise 20:1 (innen:außen)

$$\Phi_{\text{K}^+} \approx -80\text{mV}.$$

Dieser Wert liegt sehr nahe an den tatsächlich gemessenen neuronalen Ruhepotenzial $\Phi_0 = -75 \text{ mV}$ [Wehner & Gehrting, 1995].

Ein Signal wird im Nervensystem durch kurzzeitige Modifikation des Ruhepotenzials vermittelt. Ein Neuron „feuert“, d. h. es sendet ein Signal aus, wenn aus irgendwelchen Gründen die Potenzialdifferenz auf -40 mV absinkt¹. Dann nämlich öffnen sich in der Zellwand schlagartig Kanäle, die die Diffusion von Na^+ -Ionen erlauben. Da Na^+ in der Zelle weit weniger konzentriert vorliegt als außerhalb, besitzt es ein positives Gleichgewichtspotenzial, strömt also in die Zelle und depolarisiert die Membran lokal. Nach etwa 1 ms beginnen die Na^+ -Kanäle, sich wieder zu schließen, während die Depolarisation eine Öffnung von zusätzlichen K^+ -Kanälen bewirkt, so dass K^+ entlang seines eigenen Konzentrationsgefälles ausströmt und die Membran repolarisiert.

Währenddessen hat jedoch das von der depolarisierten Membran ausgehende elektrische Feld über Ionenbewegungen zum Absinken des Membranpotenzials am folgenden Ranvier-Schnürring geführt², so dass dort derselbe Zyklus ausgelöst wird, das Signal sich also fortpflanzt.³

Erreicht das Spannungssignal ein synaptisches Endknöpfchen, so bewirkt es eine Ausschüttung sogenannter Neurotransmitter (v. a. Acetylcholin), die über den synaptischen Spalt zur nächsten Zelle wandern und dort entweder die Potenzialdifferenz absenken (erregende Synapsen) oder erhöhen (hemmende Synapsen). Lässt die Nettoänderung durch zum etwa gleichen Zeitpunkt ankommende Signale diese Differenz unter den Schwellwert von -40 mV fallen, wird dadurch wiederum ein Signal ausgelöst, und so fort.

¹Die Verben „sinken“ und „erhöhen“ beziehen sich hier im Zusammenhang mit Potenzialdifferenzen immer auf den Absolutwert.

²Die Myelin-Scheide dazwischen wirkt als elektrischer Isolator und verhindert Ionendiffusion durch die Membran. Dies erhöht die Leitungseffizienz beträchtlich.

³Ein Zurückwandern des Signals wird dadurch verhindert, dass die Na^+ -Kanäle nach dem Schließen für 1 ms inaktiv bleiben.

Struktur von Nervensystem und Gehirn

Trotz dessen intrinsischer Komplexität kann die Betrachtung eines einzelnen Neurons keine Erklärung für die großartigen Leistungen liefern, zu denen das Gehirn höherer Lebewesen einschließlich des Menschen fähig ist. Erst im Verbund, der in Analogie zur Elektronik auch als eine Art „Schaltkreis“ angesehen wird, zeigen die Neuronen ihre eigentliche Stärke.

Der einfachste neuronale Schaltkreis, der bei Lebewesen vorkommt, ist die direkte Verbindung eines sensorischen mit einem motorischen Neuron, ein sogenannter monosynaptischer Reflexbogen: Sobald der Rezeptor einen Umweltreiz wahrnimmt, wird vom nachgeschalteten Motoneuron eine Muskelkontraktion bewirkt. Diese Art von Reflexen wird bereits bei sehr einfachen Organismen beobachtet.

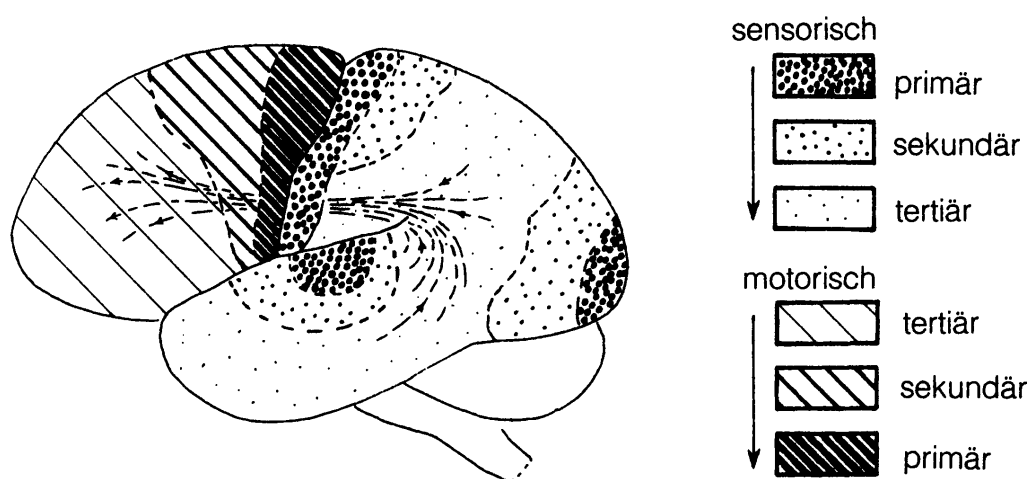


Abbildung 4.2: *Schema des Informationsflusses im menschlichen Gehirn. Im primären sensorischen Bereich ankommende Sinneswahrnehmungen wandern über die sekundären und tertiären Bereiche, in denen sie immer weiter abstrahiert und modifiziert werden, und können bei Reaktionsbedarf im primären motorischen Bereich Muskelbewegungen auslösen [nach Penrose, 1991].*

Wie bereits in Abb. 4.1 angedeutet, werden mit zunehmender Komplexität der neuronalen Steuerungsvorgänge zwischen Rezeptor und Motoneuron Interneuronen eingesetzt, die die eingehenden Sinnesreize zunächst verarbeiten und auf diese Weise differenzierteres Reagieren und/oder eine bewußte Modifikation der Muskelbewegungen erlauben. Diese Interneuronen sind üblicherweise in Schichten angeordnet: In der Netzhaut des Auges liegt hinter den Sehzellen eine Schicht Interneuronen, während in der Großhirnrinde der Säugetiere zum Beispiel sechs Neuronenschichten gezählt werden [Robert, 1995]. Allerdings zeigen eingehende Untersuchungen beim Menschen, dass der tatsächliche Informationsfluss im Gehirn noch komplizierter ist, wie Abbildung 4.2 grob andeutet. Vor der potenziellen Aktivierung von Muskelzellen läuft ein Sinnesreiz vom primären zum tertiären sensorischen Bereich, wobei die Information in zunehmendem Maße abstrahiert und mit in Form von Erinnerungen „gespeicherter“ Information assoziiert und verglichen wird. Trifft der Besitzer des Gehirns dann eine (bewußte) Handlungsentscheidung, so bewirkt sein Wille⁴ eine Folge von Signalen im tertiären motorischen Areal, die bei der Wanderung in das primäre motorische Areal von generellen Bewegungsanweisungen in gezielte Kommandos für einzelne Muskeln umgesetzt werden [Penrose, 1991].

⁴Auf welche Weise das Bewußtsein mit den physikalisch-chemischen Vorgängen im Gehirn wechselwirkt, ist derzeit allerdings noch nicht geklärt.

Neuronale Speicherung von Information

Nach der Diskussion von Signalausbreitungs-Mechanismen und Struktur des Nervengewebes bleibt anzumerken, dass die Speicherung der Information, die durch Erinnerung oder „automatische“ Reaktionen und Bewegungsabläufe abgerufen wird, nach heutigem Kenntnisstand fast ausschließlich in den Synapsen geschieht. Die Signale, die ein Neuron aussendet, sind zwar im Wesentlichen digitaler Natur („Impuls“ oder „kein Impuls“, vgl. dazu die Diskussion in [Penrose, 1991]), die Menge der auf ein Signal hin ausgeschütteten Neurotransmitter kann jedoch für jedes synaptische Endknöpfchen unterschiedlich sein und sich durch Lernprozesse verändern. Somit variiert auch die Beeinflussung des Potenzials der postsynaptischen Membran und damit die Wahrscheinlichkeit, dass die nachfolgende Nervenzelle feuert.

Es ist vielleicht nicht ganz offensichtlich, wie sich auf solche Weise zusammenhängende Information speichern lässt, da das im Gehirn verwendete Konzept grundverschieden von der aus dem Alltag vertrauten, räumlich lokalisierbaren Informations-Speicherung in Schriftzeichen, Bildern oder dem sequenziellen Speicher von Computern ist.

Am ehesten lässt sich die in Nervenansammlungen vorliegende, dezentrale Art der Datenspeicherung vielleicht mit einem Hologramm vergleichen: Ein Reflexionshologramm [Eichler & Ackermann, 1993] speichert Interferenzmuster von Lichtwellen in einer (normalerweise) flächigen Matrix. Bestrahlt man es mit kohärentem Licht, so bewirken Interferenzeffekte, dass es im Idealfall genau dieselbe Lichtverteilung zurückstretet wie das auf ihm abgebildete Objekt. Letzteres gilt natürlich auch für eine realistische bildliche Darstellung des Objektes — vom dreidimensionalen Eindruck, den ein Hologramm im Gegensatz zum Bild erzeugt, einmal abgesehen. Die Dezentralität der Datenspeicherung kommt ins Spiel, wenn eine kleine Fläche des Hologramms bzw. Bildes zerstört wird: An der betroffenen Stelle fehlt auf dem Bild dann einfach die Information über das Objekt vollständig, während sie für die restlichen Objektteile unbeeinträchtigt ist. Auf dem Hologramm ist hingegen immer noch das *gesamte* Objekt zu sehen – aber es verliert an Schärfe und Kontrast, weil die erzeugte Interferenz aufgrund der fehlenden Hologrammstelle nicht mehr perfekt ist.

Analog erzeugt der Ausfall eines Speicherchips in einem Computer eine Datenlücke, die zur völligen Unbrauchbarkeit von Programmen oder Datenstrukturen führen kann. Im menschlichen Gehirn hingegen sterben kontinuierlich Neuronen durch programmierten Zelltod ab [sog. Apoptose, Kerr *et al.*, 1972]. Erst der alters- oder schädigungsbedingte Verlust eines beträchtlichen Prozentsatzes von Neuronen – je nach Hirnareal etwa 5–30% [Selkoe, 1992] – schränkt die Gehirnfunktionen merklich ein.

Der Nachteil der neuronalen Art der Informationsspeicherung liegt jedoch auch in ihrer Dezentralität: Wie aus dem Alltag bekannt, ist die Information im Gehirn oft unstrukturiert gespeichert, und kann nicht immer auf Verlangen sofort und vollständig abgerufen werden. Weiterhin ist es sehr schwierig, außerhalb des Netzwerks den Informationsgehalt der synaptischen Verbindungsgewichte zu interpretieren, weil sie nur im Verbund wirksam werden. Dies stellt auch ein Problem der in dieser Arbeit verwendeten simulierten Neuronennetze dar, wie in späteren Kapiteln noch klar werden wird.

4.1.2 Das künstliche Neuron

Um die Signalverarbeitung eines neuronalen Netzwerks im Computer simulieren zu können, müssen zunächst einige Vereinfachungen und Idealisierungen vorgenommen werden. Abbildung 4.3 zeigt das Prinzip eines simulierten Neurons. Bei den simulierten Dendriten ankommende Signale (Eingangsmuster) y_i werden jeweils mit einem Gewicht w_i multipliziert, welches die Übertragungseffizienz der synaptischen Verbindung zur Signalquelle simuliert. Die gewichteten Signale werden aufsummiert und ergeben die *Netzeingabe* oder *Aktivierung*

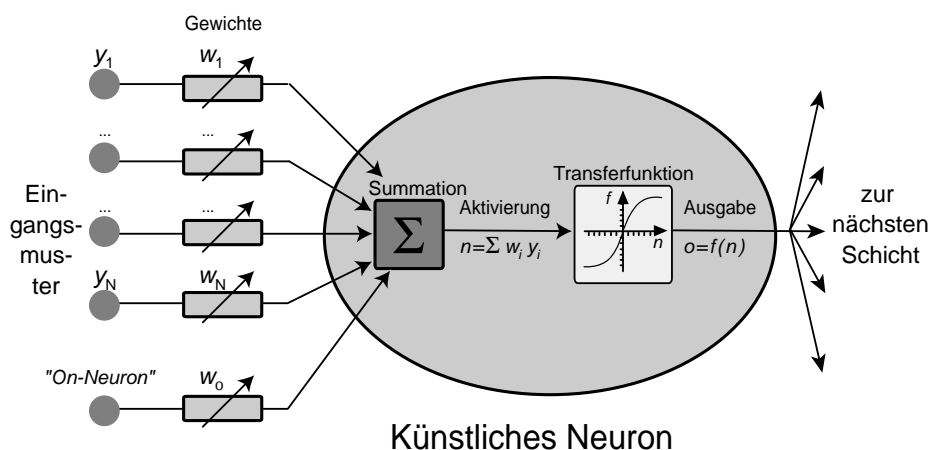


Abbildung 4.3: Schema eines simulierten Neurons.

n des Neurons.⁵ Ein Dendrit ist dabei mit dem sogenannten „On-Neuron“ verknüpft, welches dem Biasausgleich dient und immer den Wert $y_0 \equiv +1$ liefert.

In Analogie zu Abschnitt 4.1.1 entspricht n einer Verringerung der Potenzialdifferenz im Neuron. Im einfachsten Fall feuert das Neuron, sobald n eine gewisse *Aktivierungsschwelle* (*threshold*) überschreitet. Tatsächlich können durch Verknüpfung derartiger Neuronen bereits funktionsfähige neuronale Netze konstruiert werden, die alle grundlegenden Typen logischer Schaltungen nachbilden [Penrose, 1991].

Trotzdem erweist es sich in der Praxis als sinnvoller, die im Gehirn verwendeten binären Signalzustände <feuern>/<nicht feuern> durch kontinuierliche Signale zu ersetzen. Das Gehirn macht den Mangel von unterschiedlicher Signalintensität durch die Verkettung von Signalen zu komplexen Impulsfolgen, sowie durch die schiere Menge an Verknüpfungen zwischen den Neuronen wieder wett. Da viele neurobiologische Fragen insbesondere des Timings von Impulsen und Impulsfolgen noch nicht geklärt sind, erscheint es genauso berechtigt, simulierte Neuronen für die Ausgabe verschiedener Signalpegel auszustatten und damit aus den Not eine Tugend zu machen.

Anstatt durch einen einfachen Schwellwert wird der Zusammenhang zwischen Aktivierung n des Neurons und seiner Ausgabe o dann durch eine *Transferfunktion* oder *Ausgabefunktion* $f(\cdot)$ repräsentiert, die auf einem begrenzten Wertebereich kontinuierliche Werte liefert. Es zeigt sich, dass mit linearen Transferfunktionen auch nur lineare Probleme zu lösen sind [Zell, 1997], wodurch die besonderen Vorteile der neuronalen Netze nicht zur Geltung kommen. Weitere Einschränkungen entstehen dadurch, dass viele Lernverfahren (s. Kapitel 4.2) auf eine differenzierbare Transferfunktion angewiesen sind. Am häufigsten sind deshalb in der Literatur Neuronen mit sigmoider („S-förmiger“) Transferfunktion wie etwa Hyperbeltangens oder logistischer Funktion anzutreffen. Eine andere, hier nicht diskutierte Klasse von Transferfunktionen sind die *radial basis functions*, die die Form einer Glockenkurve aufweisen und z. B. für Interpolationsaufgaben eingesetzt werden [Poggio & Girosi, 1989].

⁵In manchen Modellen wird zwischen Netzeingabe n und Aktivierung a unterschieden und diese mittels einer sog. *Aktivierungsfunktion* f_{akt} verknüpft, so dass $a = f_{\text{akt}}(n)$. Jene Unterscheidung ist im Rahmen dieser Arbeit nicht notwendig, weshalb f_{akt} durchgängig der Identität gleichgesetzt wird und somit $a \equiv n$ gilt.

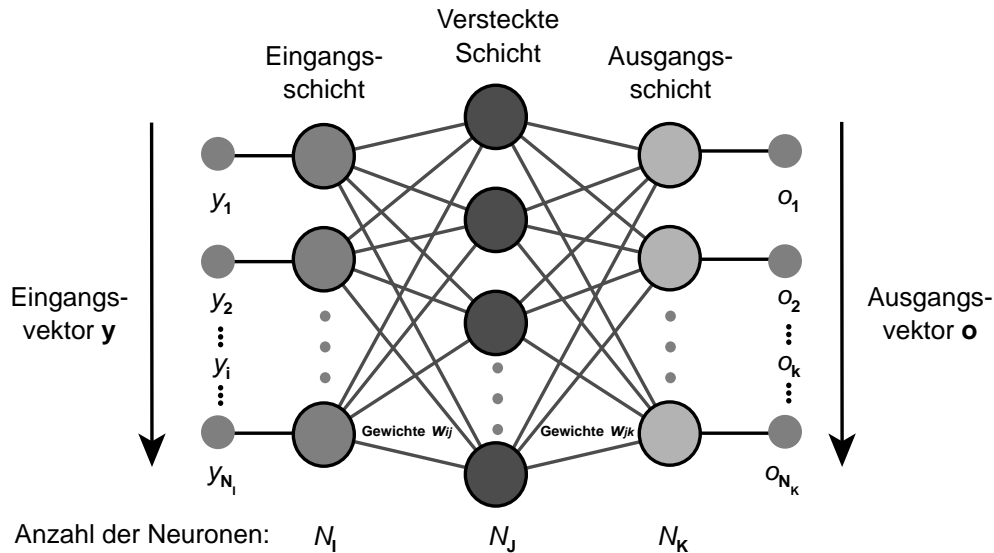


Abbildung 4.4: Verbindungsschema für feed-forward MLP-Netzwerke: Es gibt mindestens drei Neuronenschichten, die vollständig untereinander verbunden sind, aber keine Verbindungen innerhalb der Schicht aufweisen. Die Signalausbreitung erfolgt von links nach rechts, die versteckten und Ausgabeschichten können nichtlineare Transferfunktionen aufweisen.

4.1.3 Netzwerkarchitekturen

Analog zum biologischen Vorbild werden künstliche Neuronen ebenfalls gewöhnlich in Schichten angeordnet. Ausgangspunkt war hierfür das Konzept des *Perzeptrons* [Rosenblatt, 1958; Minsky & Papert, 1969], bei dem i. W. eine ein- oder zweidimensionale Schicht von Rezeptoren mit einer Ausgabeschicht verbunden sind. Diese Netze können eine einfache Mustererkennung (*pattern recognition*) durchführen, bei der meist jeweils ein Ausgangsneuron für eine bestimmte Musterklasse zuständig ist. Während zunächst nur binäre oder lineare Aktivierungsfunktionen eingesetzt wurden, entwickelte sich das Konzept im Laufe der Zeit zum nichtlinearen *multi-layer perceptron* (MLP) weiter.

Das MLP besteht aus einer Eingangsschicht, einer Anzahl von versteckten Zwischenschichten und einer Ausgangsschicht (Abb. 4.4). Die einzelnen Neuronen sind jeweils mit allen Neuronen der beiden benachbarten Schichten verbunden. Wird nun an der Eingangsschicht ein Eingangsvektor angelegt, so propagieren die Signale durch das gesamte Netzwerk zur Ausgangsschicht und werden dabei durch die Synapsengewichte und Transferfunktionen modifiziert. Da die Signale nur vorwärts laufen, wird dieser Netzwerktypus auch als *feed-forward* Netzwerk bezeichnet. Der Ausgangsvektor stellt das Klassifikations- bzw. Approximationsergebnis dar.

Es gibt viele Variationen des MLP, bei denen von der reinen, ebenenweise verbundenen *feed-forward* Topologie abgewichen wird. Manche Netze verwenden sogenannte *shortcut-connections* um Eingangsneuronen direkt mit den Ausgängen zu verbinden [Lang & Witbrock, 1989]. Ebenfalls üblich ist der Einsatz von *feedback*-Verbindungen, die gerne in Zusammenhang mit zeitlich veränderlichen Prozessen verwendet werden, weil sie dem Netz eine gewisse „Erinnerung“ an frühere Eingaben erlauben. Eine umfassende Beschreibung dieser und anderer Netzwerktopologien findet sich bei [Zell, 1997]. Für die hier gestellte Retrievalaufgabe genügt jedoch die Grundkonfiguration aus Abb. 4.4, die mathematisch auch am besten verstanden ist.

4.2 Belernungsalgorithmen

Obwohl die von einem neuronalen Netz durchgeführten mathematischen Operationen (Addition, Multiplikation, Auswertung stetiger Funktionen) für sich alleine genommen keinerlei Verständnisschwierigkeiten bereiten, ist die Mathematik der neuronalen Netze als Ganzes noch nicht vollständig verstanden. Es lässt sich beweisen, dass z. B. ein neuronales Feedforward-Netz mit $N - 1$ versteckten Neuronen in einer Zwischenschicht N Relationen zwischen Ein- und Ausgaberaum exakt nachbilden kann [Huang & Huang, 1991]. Darüberhinaus kann bereits ein dreilagiges neuronales Netz mit einer ausreichenden Anzahl von Neuronen in der Zwischenschicht eine beliebige Anzahl von Relationen mit beliebiger Genauigkeit modellieren [Stinchcombe & White, 1989]. Hornik et al. [1989] bezeichnen neuronale Netze dementsprechend auch als universale Approximatoren. Es existiert jedoch bis heute kein konstruktives Verfahren, um die zur Abbildung dieser Relationen benötigten synaptischen Verbindungsgewichte exakt zu bestimmen. Selbst wenn es eines gäbe, wäre es in der Praxis zumindest für umfangreichere Probleme höchstwahrscheinlich kaum einsetzbar: Das Finden des richtigen Gewichtssatzes ist ein sogenanntes NP-vollständiges Problem, d. h. es lässt sich nicht in höchstens polynomial mit der Anzahl der Gewichte anwachsender Zeit lösen [Blum & Rivest, 1992].

4.2.1 Veranschaulichung und Motivation

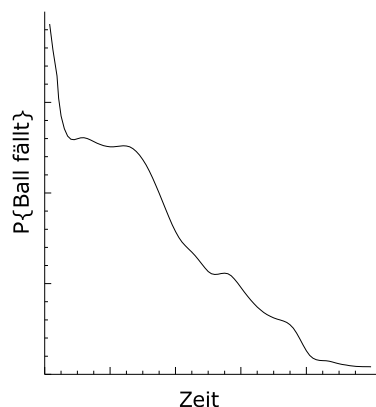


Abbildung 4.5: Mögliche Lernkurve, die ein Mensch für die Fertigkeit „Jonglieren“ durchläuft: Die Wahrscheinlichkeit, dass er einen Ball beim Auffangen fallen lässt, nimmt anfangs sehr schnell, später dann immer langsamer ab, erreicht aber nie den Wert Null. Gelegentlich treten Lernplateaus auf, bei denen trotz fortgesetzten Trainings zeitweilig keine sichtbaren Fortschritte erzielt werden.

In vielen Bereichen, in denen heute neuronale Netze eingesetzt werden, sind zudem die zu approximierenden Funktionen überhaupt nicht oder nur teilweise bekannt. Hier bietet sich wiederum die Analogie zum biologischen Vorbild an: Das Gehirn eines Jongleurs weiss z. B. auch zunächst nicht, mit welcher Kraft und zu welchen Zeiten welche Arm- und Fingermuskeln aktiviert werden müssen, um ein paar Bälle ständig in der Luft zu halten. Stattdessen sind eine ganze Reihe von Versuchen notwendig, in deren Verlauf die Hand-Auge-Koordination langsam besser wird und die Bälle länger in der Luft bleiben. Dabei spielen die Fehler eine große Rolle: jedesmal wenn ein Ball fällt, wird der Jongleur – bewusst oder unbewusst – kurz darüber nachdenken, welche Bewegung den Fehler verursacht hat, und versuchen, diese beim nächsten Mal zu unterlassen, also den Impuls an den verantwortlichen Muskel zu hemmen. Die zugehörige sogenannte *Lernkurve* ist in Abb. 4.5 dargestellt. Typisch für diese Art des Lernens ist, dass

1. die Fehlerwahrscheinlichkeit nicht gleichmäßig abnimmt, sondern sich stattdessen Abschnitte schnellen Lernens mit sogenannten *Lernplateaus* abwechseln, also Abschnitten, in denen der Lernvorgang nur schleppend vorankommt,
2. die mittlere *Lernrate*, also Verbesserung/Zeit, mit der Zeit tendenziell abnimmt und

3. kein perfektes Lernen möglich ist, d. h. die Fehlerwahrscheinlichkeit nie exakt den Wert Null erreicht. Auch nach jahrelanger Übung wird dem Jongleur gelegentlich ein Ball herunterfallen, wenn auch sehr selten.

Ein künstliches neuronales Netz verhält sich auch in dieser Hinsicht wie ein natürliches: Es muss anhand von Beispielen lernen, seine Fehlerrate zu verringern. Dies ist auch in polynomialer Zeit möglich, allerdings mit der Einschränkung, dass eventuell nicht die optimale Gewichtsconfiguration gefunden wird. Dafür bietet das Training künstlicher neuronaler Netze gleichzeitig die Chance, Zusammenhänge zwischen Eingabe- und Zielwerten zu entdecken, die gar nicht von vorneherein bekannt waren. Andererseits ist es wiederum in Ermangelung eines Verfahrens nicht möglich, die neu entdeckten Relationen aus den Gewichten des neuronalen Netzes zu rekonstruieren und in mathematisch geschlossener Form darzustellen. Wir werden auf diese Problematiken im Teil III der Arbeit zurückkommen.

Für das Training eines neuronalen Netzes wird außer einem Datensatz aus Eingabe- und Zielwerten ein sogenannter *Belernungsalgorithmus* benötigt, der mit einem iterativen Verfahren versucht, eine optimale Konfiguration von Gewichten zu finden. Den Anstoß für die Entwicklung einer ganzen Reihe solcher Algorithmen gab Mitte der 80'er Jahre die Beschreibung des *Backpropagation-Algorithmus* [Rumelhart et al., 1986], der im nächsten Unterabschnitt dieses Kapitels beschrieben wird. Ausgehend davon wird in Abschnitt 4.2.3 das für die Ozonbestimmung mit GOME bisher am erfolgreichste Belernungsverfahren vorgestellt, *Resilient Propagation* [Riedmiller & Braun, 1993].

4.2.2 Backpropagation

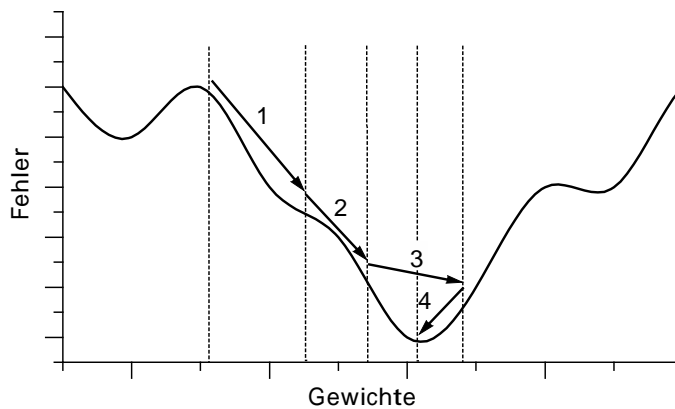


Abbildung 4.6: Gradientenabstieg zum Minimum der Fehlerfunktion $E(\mathbf{w})$ für ein (hypothetisches) Feedforward-Netz mit Hilfe von Backpropagation.

Die Idee des Backpropagation-Verfahrens ist, wie der Name bereits andeutet, Fehler von der Ausgabeschicht zur Eingabeschicht zurückzupropagieren, so dass sich alle notwendigen Gewichtsänderungen allein aus den Fehlern der Ausgabeneuronen herleiten lassen. Ziel der Gewichtsänderungen $\Delta \mathbf{w}$ ist es, ein Minimum der Gesamtfehler-Funktion $E(\mathbf{w})$ zu finden, man führt also nach dem Prinzip

$$\Delta \mathbf{w} \propto -\frac{d}{d\mathbf{w}}E(\mathbf{w}) \quad (4.3)$$

ein Gradientenabstiegsverfahren durch (Abb. 4.6). Im Vektor \mathbf{w} sind alle variablen Parameter des neuronalen Netzes zusammengefasst, also alle Gewichte w . Die folgende Konstruk-

tion des Backpropagation-Algorithmus orientiert sich an der historischen Herleitung von Rumelhart *et al.* [1986].

Gegeben sei ein mehrschichtiges Feedforward-Netz und ein Datensatz mit P Paaren von Eingangs- und Zielvektoren (auch Muster, *patterns* genannt). Weiterhin bedeutet

- t_i^p Zielwert der Belernung für Muster p und Neuron i aus Schicht I
- o_i^p Ausgabewert des neuronalen Netzes für Muster p , Neuron i aus Schicht I
- n_i^p Netzeingabe für Muster p und Neuron i aus Schicht I
- w_{ij} Gewicht der Verbindung von Schicht I , Neuron i zur nachfolgenden Schicht J , Neuron j
- f Transferfunktion des Netzwerks.
- N_I Anzahl der Neuronen in Schicht I .

Die Iteration findet in Form von **Epochen** statt. Innerhalb einer Epoche gibt es P Zyklen, in denen dem Netz alle P Trainingsmuster präsentiert werden. Danach erfolgt die Anpassung der Gewichte und die nächste Epoche beginnt. Als Maß für den Gesamtfehler einer Epoche soll hier die Fehlerfunktion

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{j=1}^{N_K} (t_j^p - o_j^p)^2 = \sum_{p=1}^P E_p \quad \text{dienen, wobei} \quad (4.4)$$

$$E_p = \frac{1}{2} \sum_{j=1}^{N_K} (t_j^p - o_j^p)^2 \quad (4.5)$$

den Fehler für Zyklus p darstellt. Gemäß dem Ansatz 4.3 sollen sich die Gewichte mit dem negativen Gradienten des Fehlers verändern. Dazu wird die Veränderung eines einzelnen Gewichts w_{ij} festgelegt als

$$\Delta w_{ij} = -\lambda \frac{\partial E}{\partial w_{ij}} = -\lambda \sum_p \frac{\partial E_p}{\partial w_{ij}}. \quad (4.6)$$

λ ist die Lernrate des Verfahrens, die die Schrittweite auf der Fehlerfläche steuert. Sie kann je nach Problem variiert werden, teilweise auch während des Belernungsprozesses. Wie bereits in Kapitel 4.1.2 erläutert, gilt weiterhin

$$o_j^p = f(n_j^p) \quad \text{und} \quad (4.7)$$

$$n_j^p = \sum_{i=0}^{N_I} o_i^p w_{ij}. \quad (4.8)$$

Der Faktor $\frac{\partial E_p}{\partial w_{ij}}$ aus Gl. 4.6 wird nun mittels Kettenregel zerlegt:

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial n_j^p} \frac{\partial n_j^p}{\partial w_{ij}} = \frac{\partial E_p}{\partial n_j^p} o_i^p. \quad (4.9)$$

Der übrig gebliebene Differentialquotient kann nun als eine Art Fehlersignal interpretiert werden, denn er beschreibt letztlich die Abhängigkeit des Fehlers E_p von der Netzeingabe n^p eines bestimmten Neurons j . Ist dieser Faktor klein, so hängt der Fehler nur schwach von der Aktivierung dieses Neurons ab, und es besteht keine Veranlassung, die Gewichte w_{ij} zu den Vorgängerneuronen schnell zu ändern. Durch die Definition

$$\delta_j^p = -\frac{\partial E_p}{\partial n_j^p} \quad (4.10)$$

ergibt sich zusammen mit Gleichung 4.6

$$\Delta w_{ij} = \lambda \sum_p o_i^p \delta_j^p. \quad (4.11)$$

Um den Faktor δ_j^p zu berechnen, ist nunmehr eine Fallunterscheidung unumgänglich:

1. Fall: Bei Neuron j handelt es sich um ein *Ausgabeneuron*. Dann ergibt sich mit den Gleichungen 4.5 und 4.7

$$\delta_j^p = \frac{\partial}{\partial n_j^p} \frac{1}{2} \sum_{j'=1}^{N_K} (t_{j'}^p - o_{j'}^p)^2 = (t_j^p - o_j^p) \frac{\partial o_j^p}{\partial n_j^p} = (t_j^p - o_j^p) f'(n_j^p) \quad (4.12)$$

2. Fall: Bei Neuron j handelt es sich um ein *Zwischeneuron*. Hier kann die Ableitung nicht direkt berechnet werden. Es bietet sich an, das Problem mittels Kettenregel und den Gleichungen 4.8 und 4.10 an die nachfolgende Neuronenschicht k weiterzureichen:

$$\delta_j^p = \frac{\partial E_p}{\partial o_j^p} \cdot \underbrace{\frac{\partial o_j^p}{\partial n_j^p}}_{f'(n_j^p)} = \frac{\partial E_p}{\partial n_k^p} \cdot \underbrace{\frac{\partial n_k^p}{\partial o_j^p}}_{\delta_k^p} \cdot \underbrace{f'(n_j^p)}_{w_{jk}} = \delta_k^p w_{jk} f'(n_j^p) \quad (4.13)$$

Auf diese Weise ist es möglich, die Korrekturen für alle Gewichte des Netzwerks zu ermitteln, indem der Faktor δ^p zunächst für alle Neuronen der Ausgabeschicht und dann rekursiv für die davor liegenden Schichten berechnet wird.

Bei der bisher besprochenen Version des Backpropagation-Verfahrens handelt es sich um die sogenannte *offline-Variante*: Gemäß der verwendeten Formel 4.11 werden zunächst die Fehlerbeiträge aller Zyklen p einer Epoche aufsummiert, bevor ein Gewichtsupdate durchgeführt wird. Eine andere Möglichkeit stellt die *online-Variante* des Algorithmus' dar, bei der die Gewichte nach jedem Zyklus gemäß

$$\Delta^p w_{ij} = \lambda o_i^p \delta_j^p \quad (4.14)$$

verändert werden. Beide Varianten haben je nach Problemstellung und vorhandenem Datenmaterial gewisse Vor- und Nachteile. Die offline-Version des Verfahrens konvergiert unter Umständen gleichmäßiger, während die online-Version die Chance auf eine wesentliche schnellere Konvergenz bietet [Zell, 1997]. Allerdings sollte die Reihenfolge der Trainingsmuster mit jeder Epoche variiert werden, um systematische Fehler zu vermeiden.

Modifikationen von Backpropagation

Wie bereits in der Einleitung dieses Kapitels erwähnt, begann mit der Entwicklung von Backprop eine sprunghafte Evolution verschiedener Belernungsalgorithmen für Feedforward-Netze. Viele davon entstanden durch Modifikation einzelner Eigenschaften von Backprop. So verwendet zum Beispiel *Quickprop* [Fahlman, 1989] anstatt des linearen Gradientenabstiegs ein Verfahren zweiter Ordnung, welches unter der Annahme lokaler Parabelform der Fehlerfläche die Konvergenz beschleunigen soll. Allerdings wird diese bessere Approximation der Fehlerfläche durch einen erhöhten Rechenaufwand erkauft.

Ein anderer Ansatz ist die Einführung eines sogenannten *Momentum-Terms*, bei der die Gleichung 4.6 durch

$$\Delta w_{ij}(\tau + 1) = \lambda \frac{\partial E}{\partial w_{ij}}(\tau + 1) + a \Delta w_{ij}(\tau) \quad (4.15)$$

ersetzt wird. a ist hier eine Kopplungskonstante von der Größenordnung 0.1, die eine Anhängigkeit von der Gewichtsänderung des vorhergehenden Zeitschritts τ generiert. Dadurch wird eine gewisse Trägheit der Gewichtsanzpassung erreicht, die verhindern soll, dass bei sehr kleinen $\frac{\partial E}{\partial w_{ij}}$ – also auf Plateaus der Fehlerfläche – die Gewichtsanzpassung stagniert. Zum anderen hilft diese Modifikation, auf dem Weg zum globalen Minimum über kleinere Unebenheiten der Fehlerfläche hinwegzuspringen [Rumelhart et al., 1986].

Eine dritte Verbesserungsmöglichkeit besteht in der dynamischen Anpassung der Lernrate an lokale Eigenschaften der Fehlerfläche [Jacobs, 1988; Tollenaere, 1990]. Aus verschiedenen Parametern wird z. B. auf Neigung und Zerklüftungsgrad der Fehlerfläche geschlossen und die Lernrate entsprechend angepasst. Auch so kann etwa eine flache Stelle schneller durchschritten werden als das bei Standard-Backpropagation möglich ist.

4.2.3 Resilient Propagation

Tatsächlich kann die Stabilität und Geschwindigkeit der Belernung durch derartige Modifikationen von Backprop erheblich gesteigert werden [Schiffmann et al., 1992]. Bei großen Netzen und Datensätzen, sowie entsprechend zerklüfteter Fehlerfläche erweisen sich diese Maßnahmen trotz alledem häufig als unzureichend, weil sie weiterhin eine inhärente Abhängigkeit von der lokalen Fehlerflächensteigung $\frac{\partial E}{\partial w_{ij}}$ besitzen. Letztlich kompensieren die erwähnten Verbesserungen zwar bis zu einem gewissen Grad diese Abhängigkeit, trotzdem kann Backprop z. B. bei zu steilen Steigungen ins Oszillieren geraten.

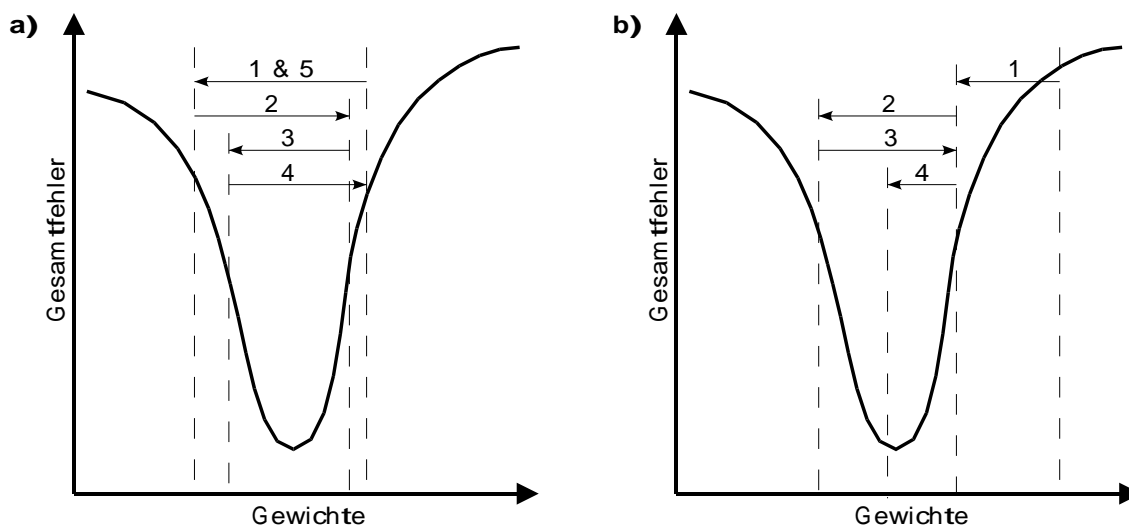


Abbildung 4.7: Annäherung an ein Minimum der Fehlerfläche während der Belernung eines neuronalen Netzwerkes. (a) Backpropagation beginnt bei steiler Fehlerfunktion zu oszillieren (hier: Periode 4 Schritte). (b) RPROP dagegen verkleinert beim Überspringen des Minimums jedesmal die Schrittweite.

Abbildung 4.7a stellt diesen Fall vereinfacht in einer Dimension dar. Backpropagation erreicht ein tiefes Minimum u. U. nie, weil der hohe Fehlergradient in dessen Umgebung eine so große Schrittweite bewirkt, dass das Minimum zyklisch übersprungen wird.

Ein weiteres Problem besteht darin, dass bei der Zurückpropagierung der Fehler gemäß den Gleichungen 4.12, 4.13 und 4.11 für jede Neuronenschicht ein weiterer Faktor δ_j^p hinzukommt. In der Praxis sind diese Faktoren meist (viel) kleiner als 1, was dazu führt, dass bei tieferen Netzen – also Netzen mit mehr als einer Zwischenschicht – die Gewichte der ersten Schichten nur noch sehr langsam verändert werden.

Der von *Riedmiller & Braun* [1993] eingeführte **RPROP-Algorithmus** (*resilient propagation*) umgeht diese Schwierigkeiten dadurch, dass die Gewichtsänderung nicht mehr direkt vom Betrag des Fehlergradienten abhängt, sondern nur noch von dessen Vorzeichen. Bei RPROP wird statt dessen zusätzlich der vorhergehende Gewichts Korrektorschritt betrachtet, wodurch eine qualitative Information über die Topologie der Fehlerfunktion gewonnen werden kann. Jedes Gewicht w_{ij} erhält seine eigene Schrittweite Δ_{ij} , die je nach Konvergenzverhalten adaptiert wird. Mit den Definitionen

$$\Gamma_{ij}(\tau) = \frac{\partial E(\tau)}{\partial w_{ij}} \cdot \frac{\partial E(\tau-1)}{\partial w_{ij}} \quad \text{sowie} \quad \text{sgn}(x) = \begin{cases} +1 & : x > 0 \\ 0 & : x = 0 \\ -1 & : x < 0, \end{cases} \quad (4.16)$$

wobei τ die Epochennummer bezeichnet, tritt an Stelle der Backpropagation-Gleichung 4.6 die Vorschrift

$$\Delta w_{ij}(\tau) = \begin{cases} \text{sgn}\left(\frac{\partial E(\tau)}{\partial w_{ij}}\right) \cdot \Delta_{ij}(\tau) & : \Gamma_{ij}(\tau) \geq 0 \quad (\text{Fall 1}) \\ -\Delta w_{ij}(\tau-1) & : \Gamma_{ij}(\tau) < 0 \quad (\text{Fall 2}) \end{cases} \quad (4.17)$$

Diese auf den ersten Blick etwas seltsam anmutende Konstruktion bewirkt, dass die Gewichte solange um Δ_{ij} vergrößert bzw. verkleinert werden, solange das Gefälle der Fehlerfläche sich nicht verändert. Bei Überspringen eines Minimums hingegen landet der Algorithmus bildlich gesprochen auf dem gegenüberliegenden Hang, der eine entgegengesetzte Steigung aufweist, so dass Γ_{ij} zwangsläufig kleiner als Null wird. In diesem Fall wird der zuletzt durchgeführte Gewichts-Schritt rückgängig gemacht.

Damit das Verfahren an dieser Stelle nicht zu oszillieren beginnt sind zwei Maßnahmen erforderlich: Zum einen muss sichergestellt werden, dass der Algorithmus bei der nächsten Iteration ($\tau + 1$) wieder den Fall 1 vorfindet, was bei der Programmierung üblicherweise dadurch realisiert wird, dass man $\frac{\partial E(\tau)}{\partial w_{ij}}$ gleich Null setzt, weil dann $\Gamma_{ij}(\tau + 1)$ ebenfalls verschwindet. Zum anderen muss die Schrittweite Δ_{ij} verringert werden, damit der folgende Schritt nicht wieder über das Ziel hinausschießt. *Riedmiller & Braun* [1993] erreichen dies durch die Vorschrift

$$\Delta_{ij}(\tau) = \begin{cases} \Delta_{ij}(\tau-1) \cdot \eta^+ & : \Gamma_{ij}(\tau) > 0 \\ \Delta_{ij}(\tau-1) \cdot \eta^- & : \Gamma_{ij}(\tau) < 0 \\ \Delta_{ij}(\tau-1) & : \Gamma_{ij}(\tau) = 0 \end{cases} \quad (4.18)$$

Die Geschwindigkeit des Wachsens bzw. Schrumpfens der Schrittweite wird dabei durch die Konstanten $\eta^+ > 1$ und $\eta^- < 1$ geregelt. Im weitesten Sinne besitzen die Konstanten η die Funktion einer Lernrate für RPROP, da sie die Konvergenzgeschwindigkeit indirekt steuern. Aufgrund ihres Zusammenspiels ist ihr Zusammenhang mit der Lerngeschwindigkeit jedoch nicht so einfach zu erfassen wie bei Backprop. Im Laufe der Voruntersuchungen stellte sich der Algorithmus als sehr stabil gegenüber der Wahl der beiden Parameter heraus. So wurde hier durchgängig $\eta^+ = 1.2$ und $\eta^- = 0.5$ gesetzt, was auch den Empfehlungen in [*Riedmiller, 1995*] entspricht.

Abbildung 4.7b zeigt schematisch das Verhalten von RPROP bei Annäherung an ein Minimum. Dieses wird zunächst übersprungen, der Algorithmus macht diesen Schritt rückgängig, verkleinert die Schrittweite und versucht so, durch Intervallschachtelung das Minimum

zu treffen. Natürlich besteht auch hier die Gefahr, anstatt des absoluten Minimums nur ein lokales Minimum zu finden. Daher wird RPROP ausschließlich im Offline-Modus betrieben: Bei der Aufsummierung der Fehler über eine ganze Epoche wird effektiv die Fehlerfläche geglättet, so dass viele lokale Minima aufgefüllt werden.

Weiterhin werden für die Gewichtsschrittweite noch ein Startwert Δ_0 sowie die minimale bzw. maximale Schrittweite $\Delta_{\min}, \Delta_{\max}$ benötigt. Auch gegenüber Variation dieser Parameter erweist sich das RPROP-Verfahren als sehr gutmütig, da eine zu kleine oder große Anfangsschrittweite durch die dynamische Adaption innerhalb weniger Epochen nach Beginn der Belernung wieder den optimalen Wert erreicht [Riedmiller, 1995]. Hier wurden nach Voruntersuchungen $\Delta_0 = 0.1$, $\Delta_{\min} = 10^{-8}$ und $\Delta_{\max} = 10$ verwendet.

4.3 Einordnung in die klassische Inversionstheorie

Die bisherigen Ausführungen sollen an dieser Stelle für eine überblicksmäßige Charakterisierung neuronaler Netze von den biologischen Ursprüngen her genügen. Es stellt sich nun die Frage, welche Aufgabe sie bei der Bestimmung von atmosphärischem Ozon erfüllen können, und wie sie dabei im Zusammenhang mit der klassischen Inversionstheorie für Satellitendaten stehen.

Gegenüberstellung mit OE

Da ein Strahlungstransfermodell im Prinzip auch eine stetige Abbildung vom Zustandsraum in den Messraum darstellt, liegt es zunächst nahe, dieses Modell $\mathbf{F}(\mathbf{x}, \mathbf{b})$ aus Gl. 3.4 mit einem neuronalen Netz nachzubilden. Dies wurde auch verschiedentlich durchgeführt [Aires et al., 2001; Chevallier et al., 2000; Schwander et al., 2001] und bringt einen Geschwindigkeitsgewinn mit sich, da die Vorwärtspropagation der Eingangsdaten (i. F. auch „Vorwärtsrechnung“) viel schneller ist als die physikalische Berechnung des Strahlungstransports. Der Rest des Retrievals kann dann mit den üblichen Methoden durchgeführt werden (Abschnitt 3.4). Dies stellt jedoch kein unabhängiges Verfahren dar, da keine neue Information in die Ozonbestimmung einfließt, und die Genauigkeit mit der Qualität des Vorwärtsmodells steigt und fällt.

In Abschnitt 3.4.1 wurde bereits gezeigt, dass die Aufgabe des Retrievalalgorithmus' in der Modellierung der bedingten Wahrscheinlichkeit $p(\mathbf{x}|\mathbf{y})$ besteht, aus der dann mit Hilfe des *maximum likelihood* Prinzips die wahrscheinlichste Lösung $\hat{\mathbf{x}}$ ausgewählt werden kann.

Zu diesem Zweck „invertieren“ OE und verwandte Verfahren quasi zunächst die Wahrscheinlichkeitsverteilung mit Hilfe des Bayesschen Satzes. Anschließend wird für den gesuchten Atmosphärenzustand \mathbf{x} eine Fehlerfunktion E minimiert, woraus sich der beste Schätzwert $\hat{\mathbf{x}}$ ergibt. Sei nun eine Menge \mathcal{T} mit P Messungen \mathbf{y}^p , sowie zugehörigen wahren Ozonprofilen \mathbf{x}^p und Vorwärtsmodell-Parametern \mathbf{b}^p gegeben. Der OE-Schätzwert $\hat{\mathbf{x}}^p$ ist dann – unter gewissen Voraussetzungen – für jeden einzelnen wahren Atmosphärenzustand \mathbf{x}^p in dem Sinne optimal, dass er die bei gegebenem Vorwärtsmodell berechneten Strahlungswerte am besten mit den beobachteten Werten und dem *a priori* Wissen in Übereinstimmung bringt. Dies wird auch als *lokales Inversionsverfahren* bezeichnet, da keinerlei Informationen über die $P - 1$ anderen Messungen in das Retrieval einfließen.

Im Gegensatz dazu stehen *globale Inversionsverfahren*, die versuchen, eine analytische Näherung für das inverse Transfermodell \mathbf{R} aus Gl. 3.19 zu finden. In diese Klasse gehören statistische Approximationsverfahren wie multiple lineare Regression [z. B. Spänkuch & Schulz, 1998; Planet, 1988], aber auch neuronale Netze, die im weitesten Sinne eine multiple nichtlineare Regression durchführen. Allerdings kann mathematisch gezeigt werden,

dass für klassische Approximationsverfahren mit festen Basisfunktionen – meist Polynome, trigonometrische Funktionen für Fourierreihenentwicklung, o. ä. – der Approximationsfehler exponentiell mit der Dimension der zu approximierenden Funktion zunimmt, während er bei neuronalen Netzen konstant bleibt [Tresp, 1995; Baron, 1993, 1994]. Gerade für das Ozonprofilretrieval ist dies ein wichtiges Ergebnis, da die Dimensionalität von Eingangs- und Ausgansraum mit 10^1 – 10^3 relativ hoch ist.

Modellierung des inversen Strahlungstransfers

Der hier verfolgte Ansatz modelliert \mathbf{R} bis auf einen Modellfehler ϵ_m direkt mit Hilfe eines MLP:

$$\hat{\mathbf{x}} = \mathbf{R}_{\text{MLP}}(\mathbf{y}, \mathbf{w}) + \epsilon_m \quad (4.19)$$

Die Aufteilung der Parameter von \mathbf{R}_{MLP} unterscheidet sich etwas von der in Gl. 3.19 vorgenommenen, da es sich hier wie gesagt um ein globales Modell handelt. Insofern bezeichnet \mathbf{w} den Vektor der synaptischen Verbindungsgewichte (vgl. Abschn. 4.2.2), die während der Belernung angepasst, aber anschließend festgehalten werden. Mit \mathbf{y} ist der Eingangsvektor des MLP gemeint, der sich aus Spektraldaten, aber auch aus denjenigen Parametern des klassischen Retrievalsystems zusammensetzt, die mit jedem einzelnen Retrieval variieren, wie z. B. Sonnenzenitwinkel, Temperaturprofil, etc. Die bei Gl. 3.19 betriebene Unterscheidung in Vorwärtsmodellparameter \mathbf{b} und zusätzliche Parameter \mathbf{b}' ist in diesem Zusammenhang irrelevant, da kein explizites Vorwärtsmodell existiert. Ferner ist die Verwendung von *lookup*-Tabellen für Komponenten von \mathbf{y} wenig sinnvoll, da hierdurch keine neue Information in das System eingebracht wird: Die Selektion eines passenden *a priori* Ozonprofils \mathbf{x}_a aus einer Ozonklimatologie impliziert z. B. einen Zusammenhang zwischen Latitude, Jahreszeit und Ozonprofil. Da das neuronale Netz theoretisch jeden solchen Zusammenhang modellieren kann, ist es aussichtsreicher, anstatt des *a priori* Profils lieber dessen Auswahlgrößen Latitude und Jahreszeit in den Vektor \mathbf{y} aufzunehmen, und so auf den Umweg über eine Tabelle – und die damit verbundene zusätzliche Diskretisierung – zu verzichten. Dies wird später noch genauer untersucht.

Wenn nun die oben eingeführte Menge \mathcal{T} groß genug und sowohl im Zustands- als auch im Messraum einigermaßen gleichmäßig verteilt ist, dann kann die bedingte Wahrscheinlichkeit $p(\mathbf{x}|\mathbf{y})$ durch die in \mathcal{T} beobachteten relativen Häufigkeiten angenähert werden. Die Likelihood-Funktion ergibt sich in diesem Fall einfach aus dem Produkt der Einzelwahrscheinlichkeiten

$$\mathcal{L} = \prod_{p=1}^P p(\mathbf{x}^p|\mathbf{y}^p), \quad (4.20)$$

und die Fehlerfunktion 3.16 wird – wiederum unter der Annahme von gaußverteilten Fehlern – zu

$$\tilde{E} = \sum_{p=1}^P (\mathbf{x}^p - \hat{\mathbf{x}}^p)^T \hat{\mathbf{S}}_p^{-1} (\mathbf{x}^p - \hat{\mathbf{x}}^p). \quad (4.21)$$

In dieser Gleichung können nun die Schätzwerte $\hat{\mathbf{x}}$ durch das inverse Modell aus Gl. 4.19 ersetzt werden:

$$\tilde{E} = (\mathbf{x}^p - \mathbf{R}_{\text{MLP}}(\mathbf{y}^p, \mathbf{w}))^T \hat{\mathbf{S}}_p^{-1} (\mathbf{x}^p - \mathbf{R}_{\text{MLP}}(\mathbf{y}^p, \mathbf{w})) \quad (4.22)$$

Zu beachten ist hier, dass die Funktion \mathbf{R}_{MLP} nicht von \mathbf{x} abhängt – sie soll ja schließlich für alle Atmosphärenzustände gültig sein. Die Minimierung von \tilde{E} erfolgt demnach bezüglich

des einzigen noch freien Parameters \mathbf{w} . Es wird also ausschließlich im Zustandsraum die Summe der Mahalanobisdistanzen zwischen der Netzausgabe

$$\mathbf{o}_p = \mathbf{o}(\mathbf{y}^p) = \mathbf{R}_{\text{MLP}}(\mathbf{y}^p, \mathbf{w}) \quad (4.23)$$

und den wahren Zuständen \mathbf{x}^p in Abhängigkeit von den Synapsengewichten des MLPs minimiert.

Übertragung in die Praxis

Hierbei stellt sich die triviale Schwierigkeit, dass der wahre Atmosphärenzustand \mathbf{x}^p natürlich unbekannt ist, andernfalls wären ja die Messungen nicht notwendig. In der Praxis ist das Verfahren daher nur anwendbar, wenn entweder mit einem Vorwärtsmodell gemäß Gl. 3.4 ein angenäherter Messvektor $\hat{\mathbf{y}} = \mathbf{y} + \boldsymbol{\epsilon}$ aus einer repräsentativen Auswahl von Atmosphärenzuständen berechnet wird, oder – alternativ dazu – unabhängige Messungen eine Schätzung

$$\mathbf{t}_p = \mathbf{x}^p + \boldsymbol{\epsilon}_t^p, \quad (4.24)$$

für jedes verwendete Ozonprofil liefern, wobei $\boldsymbol{\epsilon}_t^p$ die Messunsicherheit ist. Beides hat wiederum Vor- und Nachteile: Bei der Simulation von Messungen müssen die Instrumenteneigenschaften bereits sehr gut bekannt sein, damit die erzeugten Spektren keine systematischen Abweichungen von den tatsächlich gemessenen zeigen. Die andere Möglichkeit macht das System vom Vorhandensein geeigneter Ozonmessungen abhängig, und kann deshalb nicht für andere Spurengase verwendet werden, wenn deren Messung zu sporadisch und/oder zu ungenau erfolgt. Die hier entwickelte Methode für die Ozonbestimmung basiert auf gemessenen Ozonwerten, wobei allerdings für Voruntersuchungen auch teilweise mit berechneten Spektren gearbeitet wurde, siehe Abschnitt 7.2.

Für die praktische Anwendung des Verfahrens wird als weitere Vereinfachung für die Kovarianzmatrizen $\hat{\mathbf{S}}_p$ aus Gl. 4.22 angenommen, dass

$$\hat{\mathbf{S}}_p \equiv \sigma^2 \mathbf{I}_n \quad \forall p = 1, \dots, P \quad (4.25)$$

gilt, wobei \mathbf{I}_n die Einheitsmatrix im Zustandsraum ist. Natürlich ist für reale Retrievalprobleme die Angabe einer einzelnen Standardabweichung zur Charakterisierung des Ergebnisses meist nicht ausreichend, aber zum Einen kann der recht unwahrscheinlichen Annahme gleicher Varianzen auf der Diagonalen von $\hat{\mathbf{S}}$ durch geeignete Normierung von \mathbf{o} und \mathbf{t} entgegengewirkt werden (siehe dazu Abschnitt 8.3). Zum Anderen existieren verschiedene Möglichkeiten, die auftretenden Retrievalfehler differenzierter zu approximieren — darauf wird in den Teilen II und III der Arbeit ausführlich eingegangen. Durch Einsetzen von 4.23, 4.24 und 4.25 in Gl. 4.21 erhält man

$$\tilde{E} = \frac{1}{\sigma^2} \sum_{p=1}^P (\mathbf{o}_p - \mathbf{t}_p)^2. \quad (4.26)$$

Dieser Ausdruck für die Fehlerfunktion ist bis auf den (konstanten) Normierungsfaktor $1/\sigma^2$ identisch mit der quadratischen Fehlerfunktion E des Backpropagation-Algorithmus' (Gl. 4.4). Der Normierungsfaktor spielt für den Minimierungsprozess keine Rolle, und wird deswegen meist wie in Gl. 4.4 gleich $\frac{1}{2}$ gesetzt, so dass er sich bei Differenzieren von E herauskürzt. Wie man sich leicht klarmacht, liegt das Minimum von \tilde{E} immer beim Mittelwert der Zieldaten, also bei $\langle \mathbf{o} \rangle = \langle \mathbf{t} \rangle$. Im Grenzfall $P \rightarrow \infty$ kann man durch eine Funktionalableitung von \tilde{E} nach $\mathbf{o}(\mathbf{y})$ außerdem feststellen [Bishop, 1995a], dass

$$\mathbf{o}(\mathbf{y}) = \langle \mathbf{t} | \mathbf{y} \rangle \stackrel{4.24}{=} \langle \mathbf{x} + \boldsymbol{\epsilon}_t | \mathbf{y} \rangle = \mathbf{x}(\mathbf{y}). \quad (4.27)$$

Die Netzwerkausgabe mittelt also über biasfreies Zielwertrauschen hinweg, und folgt dabei der Form von $\mathbf{x}(\mathbf{y})$, unabhängig von deren Aussehen. Die Annahme von Gaußverteilungen bei der Konstruktion von \bar{E} bewirkt lediglich, dass der Belernungsalgorithmus die mit Gaußfehlern verrauschte Verteilung von \mathbf{t} aus Gl. 4.24 nicht von einer Verteilung unterscheiden kann, die denselben \mathbf{y} -abhängigen Zentralwert $\mathbf{x}(\mathbf{y})$ besitzt, und deren (ggf. nicht-gaußsches) Rauschen dieselbe Varianz wie $\boldsymbol{\epsilon}_t$ hat [Bishop, 1995a].

Nach dem Abschluss des Trainings stellt \mathbf{R}_{MLP} eine Approximation des (exakten) inversen Strahlungstransfermodells \mathbf{R} dar. Die erreichbare Genauigkeit dieser Approximation hängt vor allem von der Zusammenstellung des Trainingsdatensatzes \mathcal{T} ab. Wichtig sind außerdem die geeignete Wahl der Eingangsparameter, sowie die Fähigkeit des verwendeten Lernalgorithmus, das Minimum der Kostenfunktion E zuverlässig zu finden. Die Generalisierungseigenschaften neuronaler Netze ermöglichen dann, mit Hilfe des Modells \mathbf{R}_{MLP} für alle Messungen des Satelliteninstrumentes Ozonwerte zu berechnen, nicht nur für die P Messungen, die der Trainingsdatensatz \mathcal{T} enthält. Es ist wichtig, sich diesen fundamentalen Unterschied zu lokalen Inversionsverfahren klar zu machen: Neuronale Belernungsalgorithmen optimieren ein *Modell* im Hinblick auf *alle* im Trainingsdatensatz enthaltenen Fälle, während lokale Verfahren ein *Profil* optimieren, und zwar für *jeden einzelnen* Fall separat. Während die Optimierung des Netzwerks durch Training sehr rechenzeitintensiv ist, muss sie nur *einmal* durchgeführt werden, so dass bei der späteren Prozessierung großer Datenmengen ein Geschwindigkeitsvorteil gegenüber lokalen Verfahren erreicht wird.

Obwohl dieser Abschnitt nur einen knappen Einblick in die Technik der neuronalen Netze geben konnte, soll dies als Übersicht zunächst genügen. Weitere Einzelheiten und Überlegungen zur Methode finden sich bei der Diskussion des hier entwickelten Verfahrens in Teil II der Arbeit.

Fazit

Ausgehend vom Prinzip des *maximum likelihood* erweisen sich neuronale Netze vom Typ MLP unter bestimmten Voraussetzungen als geeignet, den inversen Strahlungstransfer direkt zu modellieren, indem sie eine für den Trainingsdatensatz optimale Abbildung finden. Die hier getroffenen, wesentlichen Annahmen waren

- Der Trainingsdatensatz \mathcal{T} stellt in Umfang und Verteilung eine geeignete Abtastung der auftretenden Wahrscheinlichkeitsverteilungen dar.
- Das neuronale Netz ist ausreichend komplex.
- Die Ausgangsgrößen besitzen gaußverteilte Fehler.
- Die Kovarianzmatrix der Ausgangsgrößen ist diagonal.
- Der Belernungsalgorithmus findet das globale Minimum der Fehlerfunktion E .

4.4 Bisheriger Einsatz in der Atmosphärenforschung

Neuronale Netze können im Prinzip überall dort eingesetzt werden, wo bislang klassische statistische Regressionsschemata verwendet wurden. Darüberhinaus sind sie durch ihre nicht-

linearen Eigenschaften für Probleme geeignet, an denen lineare (oder polynomiale) Regressionsschemata oft scheitern [Tresp, 1995].⁶

In der Atmosphärenforschung kann der Einsatz neuronaler Netze grob in zwei Bereiche untergliedert werden. Der erste Bereich entstand durch die Anwendung neuronaler Verfahren auf die bislang auf klassische Weise betriebene Vorhersage von Zeitreihen meteorologischer Parameter wie z. B. Trübungsgrad [Pasini et al., 2001] oder bodennahes Ozon [Ruíz-Suárez et al., 1995; Comrie, 1997]. Dieses Thema soll hier nicht weiter vertieft werden; Eine gute – wenn auch nicht mehr ganz aktuelle – Übersicht findet sich in [Hsieh & Tang, 1998].

Der zweite Bereich, zu dem auch das vorgestellte Verfahren zählt, ist das Retrieval meteorologischer Parameter aus Fernerkundungsdaten, also im Wesentlichen die Modellierung des Strahlungstransports in der Atmosphäre. Tabelle 4.1 gibt hierzu einen Überblick, ohne Anspruch auf Vollständigkeit.

Aus Platzgründen können hier nicht alle in der Tabelle aufgeführten Arbeiten ausführlich kommentiert werden. Soweit sie für die NNORSY-Methode wichtige Erkenntnisse enthalten, wird weiter unten in den entsprechenden Abschnitten darauf Bezug genommen. An dieser Stelle sollen jedoch zwei Tatsachen hervorgehoben werden: Einerseits fällt auf, dass die meisten Experimente anhand simulierter Spektraldaten vorgenommen wurden, nur wenige Studien greifen auf kollokierte Messdaten zurück.

Die andere Tatsache, die jedoch nicht direkt aus der Tabelle ersichtlich ist, besteht in der relativ geringen Dimension der meisten in der Literatur behandelten Probleme. Die neuronalen Netze enthalten meist 10–100 Neuronen (NNORSY: max. 276), die Trainingsdaten 10^2 – 10^4 Belernungsmuster (NNORSY: 10^4 – 10^6). Eine Ausnahme bildet die TOVS-Ozonbestimmung unter Mitwirkung des Autors, bei der ebenfalls sehr viele Kollokationen zum Einsatz kamen. Die Skalierung eines Schätzverfahrens um mehrere Größenordnungen kann dessen Eigenschaften jedoch grundlegend verändern. Insofern betritt die vorliegende Arbeit in zweierlei Hinsicht Neuland, da das Verhalten eines Retrievalsystems der Größe von NNORSY auf einem Datensatz aus realen Messungen – soweit dem Autor bekannt – noch nie untersucht worden ist.

⁶Die Einsatzgebiete neuronale Netze umfassen unter anderem die Vorhersage von Aktienkursen, die Erkennung von Abnormalitäten aus medizinischen Mikroskopaufnahmen, die Prüfung von Elektromotoren aufgrund der Laufgeräusche, u. v. m. Einige konkrete Beispiele finden sich in [Zell, 1997; Bishop, 1995a].

Tabelle 4.1: *Literaturüberblick zur Verwendung neuronaler Netze in der Atmosphärenphysik. Die verwendeten Akronyme finden sich im Abkürzungsverzeichnis. In der Spalte „Modell“ ist angegeben, ob das neuronale Netze Vorwärtsrechnung oder inversen Strahlungstransfer modelliert. Die Spalte „Daten“ gibt an, ob das Netz mit RTM-simulierten Daten des Sensors, oder mit Kollokationen realer Messungen der Zielwerte belernt wurde.*

Sensor	Zielparameter	Modell	Daten	Literatur
(nur Theorie)	O ₃	invers	simul.	<i>Kamgar-Parsi & Gualteri</i> [1990]
TOVS, AIRS	T	invers	simul.	<i>Escobar-Munõz et al.</i> [1993]
Bodenradiometer	T	invers	simul.	<i>Churnside et al.</i> [1994]
AIRS	T, WV	invers	simul.	<i>Motteler et al.</i> [1995]
SSM/T	T	invers	kollok.	<i>Butler et al.</i> [1996]
SSM/T	T	invers	simul.	<i>Ciotti et al.</i> [1997]
Bodenradiometer	T, WV	invers	simul.	<i>Del Frate & Schiavon</i> [1998a,b, 1999]
Bodenradiometer	T, WV, CLW	invers	simul.	<i>Solheim et al.</i> [1998]
IMG	CO	invers	simul.	<i>Hadji-Lazaro et al.</i> [1999]
AVHRR	Wolkenparameter	invers	manuell [†]	<i>Berendes et al.</i> [1999]
TOVS	O ₃	invers	kollok.	<i>Müller & Kaifel</i> [1999]; <i>Kaifel & Müller</i> [1999, 2000, 2001]
ERS-1 Scatterom.	Wind	invers	simul.	<i>Richaume et al.</i> [2000]
GOME	O ₃	invers	invert. [‡]	<i>Del Frate et al.</i> [2000]
Odin-SMR	diverse Spurengase	invers	simul.	<i>Jiménez</i> [2000]; <i>Jiménez & Eriksson</i> [2001]
Bodenradar	Regenmenge	invers	kollok.	<i>Orlandini & Morlini</i> [2000]
SSM/I	WV, CLW, T _{Boden} , Abstrahlung	invers	simul.	<i>Aires et al.</i> [2001]
SPRAFIMO	O ₃ , UV-Spektrum [§]	invers	simul.	<i>Schwander et al.</i> [2001, 2002]
TOVS u.a.	Strahlungsbudget	vorw.	simul.	<i>Chevallier et al.</i> [1998, 2000]; <i>Chevallier & Mahfouf</i> [2000]
TOVS	T	vorw.	simul.	<i>Aires et al.</i> [1999]
GOME	AMF	vorw.	invert. [‡]	<i>Loyola</i> [1999]
SSM/I	Wind, WV, SST	beide	kollok.	<i>Krasnopolsky</i> [1997]

[†] Wolken von Analysten visuell identifiziert

[‡] Klassisches Inversionsverfahren liefert Trainingsdaten für die Belernung

[§] Hochaufgel. Spektrum am Boden wird aus 5 Filtermessungen rekonstruiert

Teil II

Die NNORSY Methode

Im zweiten Teil der Arbeit wird das *Neural Network Ozone Retrieval System* (NNORSY) und seine Anwendung auf GOME-Daten beschrieben.

In **Kapitel 5** wird die Anwendung neuronaler Netze auf die Bestimmung von Ozon mit dem GOME-Instrument konkretisiert. Dabei stehen die grundlegenden Bausteine des NNORSY-Systems im Vordergrund.

Kapitel 6 beschreibt den genauen Ablauf der Gesamtozonbestimmung. Hierbei wird auf die benötigten Eingangsparameter, die Eigenschaften von Trainings- und Testdatensatz und die Ergebnisse des Retrievals eingegangen.

Für die Bestimmung von Ozonprofilen muss die in Kapitel 6 beschriebene Methodik um einige Aspekte erweitert werden, insbesondere der partiellen Belernung mit unvollständigen Profildaten und der Verwendung von zusätzlichen Eingangsparametern. Diesen widmet sich **Kapitel 7**, in dem auch die Ergebnisse des Ozonprofilretrievals vorgestellt werden.

Kapitel 5

Softwareüberblick

Die grundlegende Vorgehensweise bei der Bestimmung von Gesamtozonsäulen und höhenaufgelösten Ozonprofilen mit NNORSY ist dieselbe, daher werden in diesem Kapitel zunächst die gemeinsamen Aspekte besprochen, bevor einzeln auf die beiden Verfahren eingegangen wird. Der Ablauf des kompletten Verfahrens zur Belernung und Anwendung der neuronalen Netze ist grob in Abb. 5.1 dargestellt, und wird im Folgenden blockweise erläutert.

5.1 Vorbereitung der Spektraldaten

5.1.1 Extraktion und Kalibration

Die Daten des GOME Instruments standen für diese Arbeit in der Form binärer Level 1 Orbitdateien zur Verfügung, die aus den Level 0 Rohdaten mit Hilfe des GOME Data Processors (GDP), Version 2.7 [Balzer & Loyola, 1996] gewonnen worden waren. Diese Dateien enthalten alle relevanten Messwerte und Zusatzinformationen (Orbitdaten, Geolokation, Kalibration etc.) in einem sehr kompakten Format. Sie werden daher zunächst extrahiert, so dass orbitale Parameter sowie Wellenlängen, Radianzen und Fehler für die einzelnen GOME-Kanäle direkt zugänglich sind. Bei der Extraktion werden hier alle vom Extraktionsprogramm standardmäßig durchführbaren Korrekturen angewandt, mit Ausnahme der Degradationskorrektur.¹ Details zu diesen Korrekturen finden sich im Handbuch [Slijkhuis & Loyola, 1999].

Das Wellenlängengitter der einzelnen Kanäle, also die jedem CCD-Pixel zugeordnete Zentralwellenlänge, wird anhand der im Mittel einmal pro Monat durchgeführten Lampenkalibration und der Temperatur des Instruments zum Zeitpunkt der Messung bestimmt. Es kann daher theoretisch für jedes Spektrum eines Orbits (einschließlich des Sonnenspektrums) leicht unterschiedlich sein.

Bei der Messung des solaren Spektrums fliegt der Satellit ERS-2 im Wesentlichen auf die Sonne zu, was eine Doppler-Verschiebung

$$\Delta\lambda = \lambda \frac{v}{c} \cos\phi \sin\vartheta \quad (5.1)$$

bewirkt, wobei hier λ für die Wellenlänge steht, $\frac{v}{c}$ für das Verhältnis Satelliten- zu Lichtgeschwindigkeit, ϕ für den Azimuth- und ϑ für den Zenitwinkel der Sonne relativ zum Satelliten. Dieser Effekt wird von der GDP-Software nicht berücksichtigt, kann aber für DOAS-ähnliche Retrievalverfahren meist ignoriert werden, da er im Rahmen des spektralen „shift-and-squeeze“ [Pfeilsticker & Platt, 1994; Slaper, 1995] automatisch korrigiert wird. Bei

¹Die Befehlszeile lautete `gdp_ex -c 'LAFSNPI' <level1file> -`

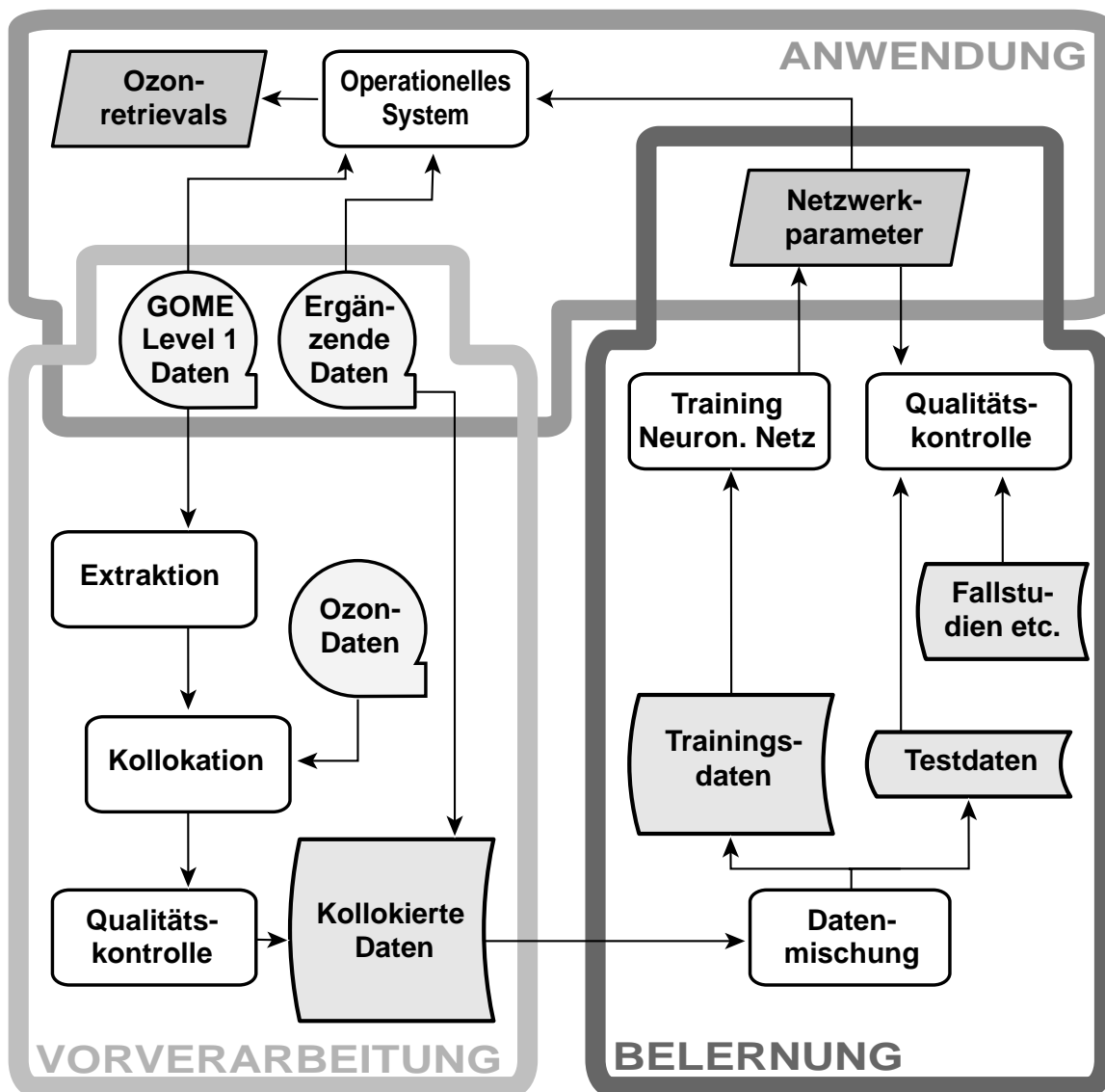


Abbildung 5.1: *Offline-Prozesskette der NNORSY-Software. Die einzelnen Prozessschritte werden in den folgenden Abschnitten dieses Kapitels erläutert.*

NNORSY hingegen wird kein *shift-and-squeeze* durchgeführt, so dass der Doppler-Shift im Rahmen der Vorverarbeitung berechnet und das Gitter des Sonnenspektrums entsprechend entzerrt wird. Die für die Anwendung von Gl. 5.1 erforderlichen Winkel zum Zeitpunkt der Messung des solaren Spektrums sind in den extrahierten Level 1 Daten nicht direkt enthalten, können aber durch Extrapolation der entsprechenden Winkel für die ersten Rückstreuungsmessungen gewonnen werden.

5.1.2 Spektraldatenformat

Weiterhin stellt sich die Frage, auf welche Weise die spektrale Information dem neuronalen Netz zur Verfügung gestellt werden soll:

- Wird nur das Rückstreu-Spektrum verwendet, ist die Ozonbestimmung sehr stark von der absoluten radiometrischen Kalibration des GOME-Sensors, sowie von eventuellen Variationen der solaren Irradianz abhängig.
- Um diese Probleme zu umgehen, kann das Rückstreu-Spektrum vorher diodenweise durch das solare Spektrum dividiert werden. Dabei können jedoch Artefakte entstehen, wenn die Wellenlängengitter nicht exakt übereinstimmen.
- Die Interpolation der Erdspektren auf das Gitter des Sonnenspektrums erzeugt hochfrequentes Rauschen.
- Das MLP könnte alternativ auch mit solarem Spektrum und Erdspektrum gleichzeitig gefüttert werden und somit die Verhältnisbildung erlernen. Allerdings wäre dann zusätzlich die Angabe der Wellenlängengitter notwendig, was die Anzahl der notwendigen Eingangsneuronen über alle Maßen erhöht.
- Die Verwendung von Wellenlängen als zusätzliche Eingangsdaten ist an sich wenig effizient, weil die Erzeugung der Wellenlängengitter durch die Kalibrationssoftware des DFD auf der Basis von Polynomen 4. Grades erfolgt [Slijkhuis & Loyola, 1999], und somit im Prinzip durch 5 Koeffizienten beschrieben werden kann.

Da neuronale Netze aufgrund ihrer statistischen Natur mit leicht verrauschten Daten recht gut zurechtkommen (vgl. Abschn. 5.4), wurde angesichts dieser Überlegungen ein Standard-Wellenlängengitter definiert, auf das sowohl Sonnen- als auch Erdspektrum unter Inkaufnahme des erzeugten Rauschens interpoliert werden. Dieses Gitter wurde durch Mittelung aus stichprobenartig ausgewählten Orbitdateien gewonnen. Dabei zeigte sich, dass lediglich Kanal 4 am langwelligen Ende eine geringe ($\sim 1\%$) Streuung der Wellenlängenschrittweite aufweist, die auf das weitgehende Fehlen von Kalibrationslinien im Lampenspektrum dieses Bereiches zurückzuführen ist [W. Thomas, pers. Mitteilung, 2000]. Die Interpolation erfolgte mit dem Akima-Algorithmus [Akima, 1970].

5.1.3 Behandlung von Kanal 1A

Einen Spezialfall für die Extraktion der GOME-Spektren stellt der Kanal 1A dar (vgl. Tab. 3.1). Einerseits ist seine Integrationszeit aufgrund des geringen S/N-Verhältnisses achtmal so lang wie die der anderen Kanäle, andererseits wurde seine Kanalgrenze am 7. Juni 1998 von 307 nm auf 283 nm verlegt. In vielen Retrieval-Algorithmen werden jeweils die 8 Pixel, die in der Integrationszeit des Kanals 1A liegen, koaddiert, so dass effektiv bei allen Wellenlängen ein Mittelwert über einen „Metapixel“ von etwa $960 \times 100 \text{ km}^2$ vorliegt [de Beek, 1998]. Dadurch wird auch Szenenhomogenität erreicht, d. h. das Instrument betrachtet für alle Spektralbereiche die gleichen Luftmassen. Allerdings verringert sich gleichzeitig

die im Vergleich zu anderen Nadir-Soundern wie TOMS oder TOVS bereits relativ schlechte horizontale Auflösung nochmals beträchtlich.

Für NNORSY wurde daher ein anderer Weg eingeschlagen, bei dem jedem der erwähnten 8 Pixel dasselbe Kanal-1A-Spektrum zugewiesen wird. Dies kombiniert die höhere horizontale Auflösung der langwelligen Kanäle mit dem großen Kanal-1A-Metapixel. Aus der Strahlungstransfertheorie ergibt sich für die Kanalgrenze je nach Sonnenzenitwinkel ein maximaler Einfluss des Ozons bei 40–42 km Höhe (bzw. 30–35 km vor Juni 1998), der sich zum Kurzwelligen hin weiter nach oben verschiebt [Hoogen, 1998]. Die Ozonverteilung in diesen Höhen der Atmosphäre variiert aber bereits so schwach (vgl. Abschn. 10.3.2), dass innerhalb des Metapixels keine großen Veränderungen zu erwarten sind und die Zuweisung desselben 1A-Spektrums an alle Einzelpixel gerechtfertigt scheint.

5.2 Kollokation mit Ozonmessungen

Für die Kollokation wird der Abstand der Mittelpunkte von GOME-Pixel und Ozonmessung nach der Formel für sphärische Dreiecke

$$\Delta = R_{\text{Erde}} \cdot \arccos(\sin(\theta_1) \sin(\theta_2) + \cos(\theta_1) \cos(\theta_2) \cos(\phi_1 - \phi_2)) \quad (5.2)$$

ermittelt, wobei θ geografische Breite und ϕ Longitude der beiden Mittelpunkte bezeichnen. $R_{\text{Erde}} = 6378.17 \text{ km}$ ist der mittlere Erdradius. Die leichte Verzerrung durch das asphärische Geoid wird somit vernachlässigt, spielt aber angesichts der anderen Kollokationsfehler keine nennenswerte Rolle.

An dieser Stelle soll nochmals auf die Philosophie hinter NNORSY hingewiesen werden, die explizit die Verwendung von *Messdaten* als Referenz vorsieht, was keineswegs selbstverständlich ist. Wie in Abschnitt 4.4 angeklungen, wird das Problem der geeigneten Zusammenstellung von Trainingsdaten in der Atmosphärenphysik gelegentlich dadurch umgangen, dass die Retrievalergebnisse eines für den betreffenden Sensor bereits vorhandenen Verfahrens als Trainingsdaten verwendet werden. Dann ist eine Kollokation mit externen Daten unnötig, aber das neuronale Netz übernimmt i. W. die systematischen und statistischen Fehler des „Lehrer“-Verfahrens, so dass es keine neue Information generieren kann, sondern nur zur Beschleunigung des Retrievals dient. Für die Überführung eines im Offline-Betrieb entwickelten Verfahrens in die operationelle Anwendung erscheint dies sinnvoll. Im Falle des GOME-Sensors existiert jedoch noch kein Konsens über das geeignetste Retrievalverfahren für Ozonprofile², was insbesondere durch Schwierigkeiten bei der absoluten Kalibration der GOME-Spektren begründet ist [Aben et al., 2000]. Die offiziellen Gesamtzondaten (Version 2.7) haben hingegen noch nicht die Qualität erreicht, die von physikalischer Seite her aufgrund des Informationsgehaltes der GOME-Spektren zu erwarten wäre [Corlett & Monks, 2001], da die bei der AMF-basierten Umwandlung schräger DOAS-Ozonsäulen in vertikale Säulen getroffene Grundannahme schwacher Absorber in einigen Bereichen zusammenbricht [Marquard et al., 2000].

Die in der Literatur (vgl. Tab. 4.1) häufig angewandte Alternative zum Lehrerverfahren besteht in der Erstellung eines Satzes von simulierten Spektren, die mit Hilfe eines RTM auf der Basis einer hinreichend großen Anzahl von verschiedenen Atmosphärenzuständen berechnet werden. Diese Methodik wurde für Voruntersuchungen auch in diesem Projekt verwendet, um die prinzipielle Durchführbarkeit des Neuronale-Netze-Verfahrens zu demonstrieren und einige Fragen der Belernung zu klären (Abschnitt 7.2). Generell eignet sich diese Möglichkeit beim momentanen Stand der Sensorhardware und der RTMs nicht für GOME, da die Auswirkungen von Instrumenteneffekten wie z. B. Degradation entweder mit

²Deswegen wurde von ESA-ESRIN die *GOME Ozone Profile Working Group* ins Leben gerufen.

simuliert oder im Rahmen der Vorverarbeitung der gemessenen Spektren herauskorrigiert werden müssten, was aufgrund mangelnden Verständnisses bisher nicht biasfrei möglich ist [Aben et al., 2000]. Für zukünftige Satelliteninstrumente ist dies jedoch durchaus ein gangbarer Weg, insbesondere zu Beginn ihrer Lebenszeit, wenn die Kalibration aus Labormessungen noch relativ frisch ist und noch nicht genügend Kollokationen für eine Belernung vorliegen.

5.3 Ergänzende Eingangsdaten

Abgesehen von den in den GOME Level 1 Daten enthaltenen Angaben wie etwa Spektralwerten, diversen Winkeln und der geografischen Position, sind für die Belernung der neuronalen Netze ggf. noch zusätzliche Eingangswerte nützlich, die den Kollokationen im Rahmen der Datenmischung hinzugefügt werden.

Die **Zeitangabe** des GOME-Pixels wurde in mehrere Eingangswerte aufgespalten: In eine lineare Zeitkoordinate (Instrumentenalter), eine sinusförmige saisonale Komponente und ein neu/alt-Flag, welches den Übergang von alter zu neuer Grenze zwischen den GOME-Spektralkanälen 1A und 1B markiert (7. Juni 1998). Die Bezeichnung *Flag* bezieht sich in diesem Zusammenhang auf einen diskreten Eingabeparameter, der entweder den Wert 1 oder -1 annimmt. Damit können Sprünge in den anderen Eingangswerten modelliert werden, die ansonsten nur schwer durch das neuronale Netz zu erfassen sind, da es sie – zumindest bei der gewählten Architektur mit kontinuierlichen Transferfunktionen – mit Hilfe kontinuierlicher Funktionen modellieren muss [Bishop, 1995a].

Als drei Flags realisiert wurde die **Blickrichtung** des GOME-Instruments (*line-of-sight*, LOS), und zwar zusätzlich zum kontinuierlichen Scanwinkel des Sensors. Je nachdem, ob ein Ost-, Nadir- und Westpixel vorliegt (Backscan-Pixel wurden nicht verwendet), wird jeweils das entsprechende Flag gleich 1, die anderen beiden gleich -1 gesetzt. Auch dies dient zum Ausgleich eventueller Diskontinuitäten zwischen den Pixeltypen, die durch die unterschiedliche Beobachtungsgeometrie entstehen und nicht durch kontinuierliche Winkelangaben modelliert werden können. Die Verwendung dreier Flags – anstatt von einem einzigen mit drei Zuständen – läßt die Modellierung einer unabhängigen Korrektur für jeden einzelnen Pixeltyp zu.

Die Kopplung von Temperatur- und Ozonprofil über die Absorption von Strahlung in der Stratosphäre [Spichtinger-Radowsky, 2001] gab Anlass zu der Überlegung, ersteres in die Kollokationen mit ein zu beziehen. Dies geschah in Form von **Temperaturdaten** der UK Met Office (UKMO) Analysen, die alle 24 h global auf 22 Druckschichten zwischen 1000 und 0.32 hPa, mit einer horizontalen Auflösung von 2.5° Breite und 3.75° Länge vorliegen [Swinbank & O'Neill, 1994]. Diese Daten wurden auf die Koordinaten jedes GOME-Pixels interpoliert und orbitweise abgespeichert, so dass sie dem Kollokationsdatensatz beigelegt werden konnten.

Zusätzlich wurde in Erwägung gezogen, die quantitative Behandlung von Wolkeneffekten durch die Einbeziehung von Wolkenparametern zu ermöglichen. Allerdings sind die derzeit im Rahmen der GOME Level 2 Daten verfügbaren Wolkenbedeckungsgrade auf der Basis des ICFA-Algorithmus [Kuze & Chance, 1994; Spurr, 1996] offenbar von fragwürdiger Qualität [Koelemeijer & Stammes, 1999; Twinder et al., 2002]. Die anderen beiden zur Charakterisierung von Wolken benötigten Parameter – Wolkenhöhe und optische Dichte – sind hingegen nicht oder nur als Klimatologie verfügbar, was für die Belernung neuronaler Netze ungeeignet ist.

Obwohl es derzeit immer noch sehr schwierig ist, aus Satellitendaten alle drei Parameter unabhängig voneinander zu bestimmen, hat das zunehmend bessere Verständnis der Wolkenphysik sowie die Weiterentwicklung der Wolkenbestimmung aus GOME-Daten da-

zu geführt, dass für die Ende 2002 geplante GDP 3.0 Datenversion einige sehr aussichtsreiche Algorithuskandidaten bereit stehen [Tuinder *et al.*, 2002, und Referenzen darin], die unter anderem auch die Zusatzinformation aus PMD-Messungen nutzen dürften [Guzzi *et al.*, 1998; von Barga *et al.*, 2000]. Angesichts dieser Tatsache und der zu erwartenden Komplexität der Wolkenbehandlung mit neuronalen Netzen wurde auf eine quantitative Wolkendiskussion im Rahmen dieser Arbeit verzichtet.

5.4 Datenmischung

Nach der Kollokation erfolgt die Aufteilung des Datenbestandes in Trainings- und Testdatensatz, sowie ggf. das Aussortieren von Kollokationen aufgrund statistischer Kriterien, die für Gesamt- und Ozonprofile unterschiedlich sind und daher weiter unten in den entsprechenden Kapiteln diskutiert werden. Die Kollokationen des Trainingsdatensatzes werden direkt für die Belernung verwendet, während der Testdatensatz dazu dient, die Konvergenz der Belernung zu überprüfen und ein Übertrainieren zu verhindern [Kohavi, 1995; Bishop, 1995a]. Zudem wird meistens zusätzlich ein Evaluierungs-Datensatz erzeugt, der sicherstellen soll, dass die Testdaten nicht ebenfalls übertrainiert werden [Plutovski, 1996]. Die Belernungen mit den hier verwendeten Ozondatenbeständen zeigten jedoch praktisch identische Ergebnisse von Test- und Evaluierungsdaten, wenn diese auf ähnliche Weise zufällig aus den Kollokationen generiert wurden. Folglich wird hier meist auf diesen dritten Datensatz verzichtet. Stattdessen erfolgt die Evaluierung anhand einzelner Bodenstationen, die explizit weder dem Test- noch den Trainingsdaten zugeteilt werden, sowie durch unabhängige Satellitenmessungen.

Ein erfolgreiches Training neuronaler Netze hängt in großem Maße von dem Aufbau eines geeigneten Trainingsdatensatzes ab. Bei dessen Zusammenstellung sind dementsprechend folgende Punkte zu beachten:

- Die **Verteilung** der Zielwerte bestimmt, ob das Retrieval später bei häufig auftretenden Atmosphärensituationen oder bei Extremereignissen geringere Fehler aufweist. Dies wird am Beispiel der Gesamt- und Ozondaten in Abschn. 6.2 ausführlich erläutert.
- Je genauer die Zuordnung von Eingangswerten und Ozonmessung möglich ist, desto besser kann der Belernungsalgorithmus die durch die Trainingsdaten repräsentierte Abbildung approximieren. Andererseits besteht dann gleichzeitig eine erhöhte Gefahr, dass das Netz übertrainiert wird und auf unabhängigen Daten starke Performance-Einbußen hinnehmen muss. Es lernt also letztlich nur die Trainingsmuster auswendig, und nicht die darunter liegenden physikalischen und statistischen Zusammenhänge. Diesem Phänomen kann durch **Regularisierung** entgegengewirkt werden. Unter diesem Begriff lassen sich eine Reihe von Maßnahmen zusammenfassen, die die Generalisierungsleistung des neuronalen Netzen auf Kosten des Fehlers beim Training verbessern. In Zusammenhang mit den Trainingsdaten spielen dabei eine Rolle

1. *Statistisches Rauschen* (Messfehler) bei den Eingangs- und Zielwerten. Diese sind in der physikalischen Anwendung immer präsent und verhindern prinzipiell eine eindeutige Zuordnung eines Eingangsvektors zu genau einem Ausgangsvektor. Im bayesianischen Sinne stellen beide Vektoren dann nur Mittelwerte der durch die Messfehler bestimmten Wahrscheinlichkeitsverteilungen im Eingabe- und Zielwertraum dar. Unter bestimmten Voraussetzungen kann mathematisch bewiesen werden, dass das Training neuronaler Netze mit verrauschten Daten der Tikhonov-Regularisierung [Tikhonov & Arsenin, 1977] äquivalent ist [Bishop, 1995b].

2. *Kollokationsfehler* entstehen durch die zeitlichen und räumlichen Gradienten des Ozonfeldes. Jede Ozonkollokation assoziiert die vom Sensor abgetasteten Luftmassen mit einem Mittelwert des atmosphärischen Ozons an unterschiedlichem Ort, zu unterschiedlicher Zeit und mit unterschiedlichem Volumen – letzteres ist insbesondere bei Ozonsonden relevant, die in jeder Höhengschicht im Wesentlichen eine Punktmessung darstellen. Diese Fehler können in guter Näherung als weißes Rauschen interpretiert werden [*Fioletov et al., 1999*]³.
 3. *Mehrfachkollokationen* kommen bei GOME häufig vor, da sich meistens innerhalb des festgesetzten Kollokationsradius³ um eine Ozonmessung mehrere Satellitenpixel befinden. In den Polarregionen, wo die Satellitenorbits stärker überlappen, können dies auch Pixel verschiedener Orbits sein. Das MLP muss dann versuchen, trotz der verschiedenen Wolkenbedeckung, SZA, LOS, etc. in diesen GOME-Pixeln dasselbe Ozon(profil) zu erkennen.
 4. *Unterschiedliche Datenquellen* für die Ozondaten bedingen ebenfalls eine größere Standardabweichung beim Training, da sie nie vollständig übereinstimmen und das neuronale Netz somit Kompromisse eingehen muss. Andererseits können sich bei geschickter Kombination verschiedener Quellen deren systematische Fehler gegenseitig kompensieren, so dass das fertig trainierte neuronale Netz auf einem unbekanntem Datensatz einen geringeren Bias besitzt als die Datenquellen einzeln betrachtet. Zum Beispiel ist es recht unwahrscheinlich, dass alle systematischen Abweichungen bei SAGE II und HALOE denselben Ursprung besitzen, da das SAGE II Instrument im UV/Vis und HALOE im mittleren Infrarot arbeitet (s. dazu auch Abschn. 10.2.2).
- Für den Trainings- und Testdatensatz muss **statistische Unabhängigkeit** sichergestellt werden, andernfalls kann der Testdatensatz seinen Zweck nicht erfüllen. Bei der Datenmischung werden daher entweder *alle* Kollokationen mit bestimmten Bodenstationen in den Testdatensatz übernommen, oder es wird (bei Profildaten) sichergestellt, dass keine Mehrfachkollokationen auf Trainings- und Testdaten verteilt sind, also dasselbe Profil nicht mit jeweils unterschiedlichen GOME-Pixeln kollokiert in beiden Datensätzen vorkommt. Natürlich ist die auf diese Weise erreichte Unabhängigkeit nie perfekt, da durch die Dynamik der Atmosphäre immer eine gewisse Kopplung auch zwischen örtlich und zeitlich verschiedenen Kollokationen vorherrscht. Ein wesentliches Ziel der Validation von NNORSY ist somit, die Generalisierungsfähigkeit des neuronalen Netzes auch außerhalb des Testdatensatzes zu zeigen.

5.5 Training neuronaler Netze

Bei allen für NNORSY verwendeten neuronalen Netzen handelt es sich um ebenenweise vollständig verbundene Feedforward-Netze, also MLPs (Abschn. 4.1.3). Es gibt eine nichtlineare Zwischenschicht mit Hyperbeltangens-Transferfunktionen, sowie unter Umständen nichtlineare Ausgangsneuronen. Wie bereits angesprochen (Abschn. 4.2) ermöglicht diese Architektur bereits die Modellierung beliebiger Abbildungen vom Eingangs- zum Ausgaberaum [*Hornik et al., 1989; Castro et al., 2000*], allerdings nur unter der zusätzlichen Bedingung dass das Modell komplex genug ist [*Bishop, 1995a*]. Die Komplexität eines MLPs wird durch die Gesamtanzahl seiner veränderlichen Gewichte (Freiheitsgrade) kontrolliert, die wiederum – falls Ein- und Ausgabe feststehen – über die Anzahl der Neuronen in der

³Wenn der Kollokationsradius jedoch so groß wird, dass die Breitenkreise darin nicht mehr als Geraden angenähert werden können, besteht die Gefahr eines meridionalen Biases, weil dann die Flächen für niedrige und hohe Latituden nicht mehr gleich groß sind. Dieser Fall kommt aber in der Praxis kaum vor.

Zwischenschicht variiert werden kann. Die Verwendung von mehr als einer Zwischenschicht hat sich bei NNORSY nicht als gewinnbringend erwiesen.

Wie groß die Zwischenschicht für eine gegebene Problemstellung minimal sein muss, lässt sich bislang nicht mathematisch herleiten [*Tamura & Tateishi, 1997*] sondern nur empirisch bestimmen. Anhaltspunkte hierfür sind

Größe der Ein- und Ausgabeschicht: *Tamura & Tateishi* [1997] geben als Faustregel die durchschnittliche Größe der anderen beiden Schichten an, aber dieser Wert kann den Autoren zufolge je nach anderen Faktoren sehr stark schwanken.

Komplexität des Problems: Diese bestimmt sich einerseits aus der Anzahl der verfügbaren Trainingsmuster, andererseits aus den Eigenschaften des zu modellierenden Zusammenhangs zwischen Eingabe und Zielvektoren. Generell werden stark redundante, lineare oder sehr verrauschte Daten weniger Information enthalten und daher auch weniger große Zwischenschichten benötigen.

Verwendeter Lernalgorithmus: Je nach Neigung zum Übertrainieren können verschiedenen Algorithmen auch verschieden komplexe Netze zugemutet werden [*Zell, 1997*].

Soweit nicht anders erwähnt, wurden alle Belernungen mit dem in Abschnitt 4.2.3 erläuterten RPROP-Algorithmus durchgeführt. Die Verbindungsgewichte wurden zu Beginn des Trainings wie bei [*Bishop, 1995a*] beschrieben mit Zufallswerten initialisiert. Ein neuronales Netz wird als fertig belernt klassifiziert, sobald der RMSE auf dem Testdatensatz nicht mehr abnimmt, oder sich innerhalb von 10^3 Epochen nur noch in der 5. geltenden Ziffer ändert. Die Rechenzeit für eine Belernung betrug je nach Umfang des Trainingsdatensatzes und Größe des Netzes in der Regel einige Tage bis zwei Wochen auf einem 400 MHz Sun UltraSPARC II Prozessor.

Kapitel 6

Gesamtozonbestimmung

6.1 Gesamtozon-Kollokation

Ausgangspunkt für das Training eines Gesamtozon-Netzwerkes sind die Kollokationen mit bodengestützten Dobson-, Brewer- und Filterspektrometermessungen. Das WOUDC enthält zur Zeit Daten von etwa 450 Bodenstationen, von denen etwa 200 im für GOME relevanten Zeitraum Messwerte lieferten.

Trotz der am WOUDC durchgeführten Qualitätskontrolle schwankt die Datenqualität der einzelnen Stationen beträchtlich [Fioletov *et al.*, 1999]. Die Arbeitsgruppe von P. K. Bhartia am *NASA Goddard Space Flight Center* (GSFC) verwendet deshalb nur 51 ausgewählte Stationen für die Validierung der TOMS Daten [P. K. Bhartia, *persönliche Mitteilung*, 2000]. Auswahlkriterium hierfür sind zum einen die Verfügbarkeit von genügend langen Messreihen und zum anderen die Qualität der Messdaten. Für die Belernung von neuronalen Netzen ist diese Stationsauswahl jedoch nicht geeignet, da die Stationen nicht alle Breitengrade abdecken und sich stark in der Nordhemisphäre konzentrieren.

Die komplette Liste der tatsächlich für die Gesamtozon-Belernungen verwendeten Stationen findet sich in Anhang A. Diese Auswahl entstand in mehreren Schritten: Zu Beginn des NNORSY-Vorgängerprojektes zur Ozonbestimmung aus TOVS-Daten [Müller & Kaifel, 1999; Kaifel & Müller, 1999, 2000, 2001] fiel nach den ersten TOVS Gesamtozon-Belernungen auf, dass die Abweichungen der Bodenmessungen von den TOVS-Ozonretrievals zwischen den einzelnen WOUDC-Stationen stark schwankten. Dies galt auch, wenn die Stationen in klimatologisch ähnlichen Regionen und/oder nahe beieinander lagen [Kaifel & Müller, 1999]. Eine qualitative Überprüfung dieses Effektes anhand von operationellen TOMS- und GOME-Ozondaten ergab, dass sehr häufig Stationen, deren Ozonwerte gegenüber TOVS große systematische Offsets oder Streuungen aufweisen, beim Vergleich mit Ozondaten von TOMS und GOME ein ähnliches Verhalten zeigten. Daraufhin wurde nach einer Möglichkeit gesucht, die Qualität der Bodenstationen automatisch zu bewerten, da die Recherche von Hintergrundinformationen zu 200 Stationen viel zu aufwändig und in vielen Fällen von zweifelhaftem Erfolg gewesen wäre.

Der Autor entwickelte daher ein semi-empirisches Bewertungssystem, welches auf den Abweichungen zwischen Bodenmessungen und kollokierten TOMS bzw. GOME Level 3 Ozondaten beruht. Für jede WOUDC-Station wurden der mittlere absolute Fehler (MAE), die mittlere Abweichung für Tages- und Monatsmittelwerte und die Standardabweichung bestimmt. Aus diesen vier Größen wurden separat Stations-Ranglisten für GOME und TOMS erstellt und mit heuristischen Gewichtungsfaktoren bewertet. Unabhängig davon geht die Anzahl der kollokierten Ozonwerte ebenfalls in die Bewertung ein, weil sie die Zuverlässigkeit der statistischen Aussagen widerspiegelt. Einzelheiten zu den Bewertungen finden sich in [Kaifel & Müller, 2000]. Von der resultierenden Gesamt-Rangliste der etwa 200 WOUDC-Stationen wurden die Kollokationen mit den 100 besten Stationen übernom-

men, sowie zusätzlich einige weitere Stationen auf der Südhemisphäre, von denen jedoch keine einen Rang schlechter als 150 innehatte. Dieses Vorgehen dient zur Verbesserung der geografischen Gleichverteilung der Stationen.¹ Das Ranglisten-Schema wurde bei den Belernungen mit TOVS-Daten intensiv getestet und hat sich dabei als erfolgreich erwiesen [Kaifel & Müller, 2001].

Die in jüngerer Zeit von Bodeker *et al.* [2001] vorgeschlagene klimatologische Korrektur, sowie neue Erkenntnisse über systematische Abweichungen zwischen TOMS und GOME Ozonwerten [Corlett & Monks, 2001] führten zu einer Revision der Liste. Später wurden noch einige neu hinzugekommene Stationen des WOUDC in die Liste aufgenommen und die geografische Verteilung verbessert.

Da es trotzdem kaum möglich ist, eine vollständige und gleichmäßige globale Abdeckung allein mit den Bodenstationen zu erreichen, wurde der Kollokationsdatensatz durch die Generierung von „virtuellen“ Stationen ergänzt. Diese entstehen, indem tägliche TOMS Level 3 Ozonkarten für bestimmte geografische Orte als Zeitreihe abgespeichert werden. Um den Datensatz möglichst homogen zu halten wurde dabei die Korrektur von Bodeker *et al.* [2001] auf die TOMS-Ozonwerte angewandt (Anhang B).

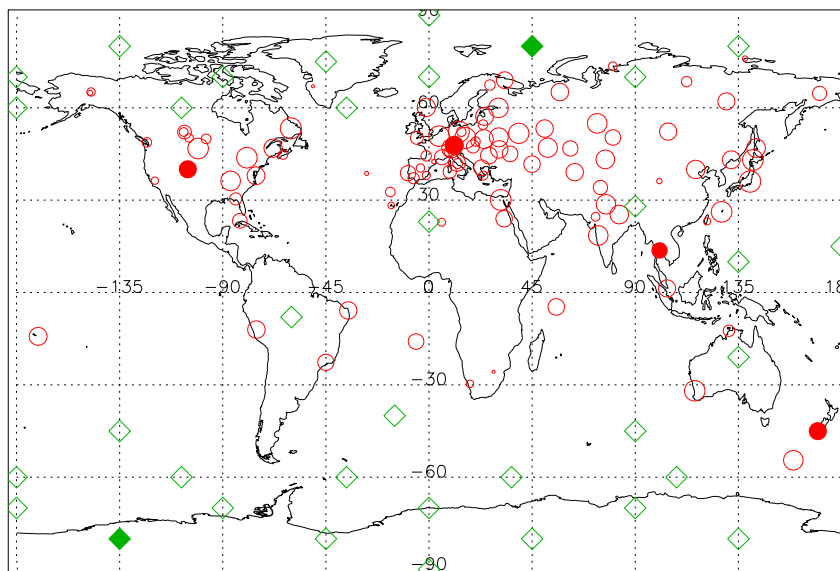


Abbildung 6.1: Verteilung der für die Gesamtozon-Belernungen verwendeten Standorte. Gefüllte Symbole bezeichnen für den Testdatensatz verwendete Stationen. Kreise: WOUDC-Stationen. Die Kreisfläche skaliert mit der Anzahl der im für GOME relevanten Zeitraum vorhandenen Messungen. Rauten: „virtuelle“ Stationen aus korrigierten TOMS Daten (s. Text).

Zunächst waren die virtuellen Stationen nur in hohen geografischen Breiten über 60° verwendet worden, um eine bessere Kontrolle über eventuelle verbliebene Differenzen zwischen TOMS und den Bodenstationen zu haben. Diese Differenzen würden sich entsprechend der hauptsächlich meridional gegliederten Verteilung des Gesamtozons in der Fehlerstatistik zonaler Mittelwerte zeigen. Müller *et al.* [2002a] konnten jedoch keine derartigen Effekte

¹Wie weiter unten noch klar werden wird, sind GOME- und TOMS-Ozonwerte gerade aufgrund des Ungleichgewichtes der Stationsanzahl von NH und SH auf die NH optimiert, weswegen die SH-Stationen in der Rangliste tendenziell schlechter eingestuft werden. Daher geht die manuelle Auswahl einiger rangniedriger SH-Stationen nicht unbedingt mit einer Qualitätsverschlechterung einher.

nachweisen, weshalb nach einer neuerlichen Revision die virtuellen Stationen bei Bedarf auch in denselben Latitudenbändern verwendet werden wie Bodenstationen. Abb. 6.1 zeigt die aktuelle Verteilung der Stationen.

6.2 Homogenisierung der Trainingsdaten

Die Standardbelernung für Gesamtozondaten umfaßt den Zeitraum Januar 1996 bis einschließlich Juli 2001. Für die verwendeten Bodenstationen zusammen mit den virtuellen TOMS-Stationen ergaben sich daraus bei einem Basis-Kollokationsradius von 120 km 338832 Rohdatensätze. Davon entfielen 17752 auf die in Abb. 6.1 gezeigten Teststationen. Diese Kollokationen wurden ohne weitere Behandlung in den Testdatensatz übernommen.

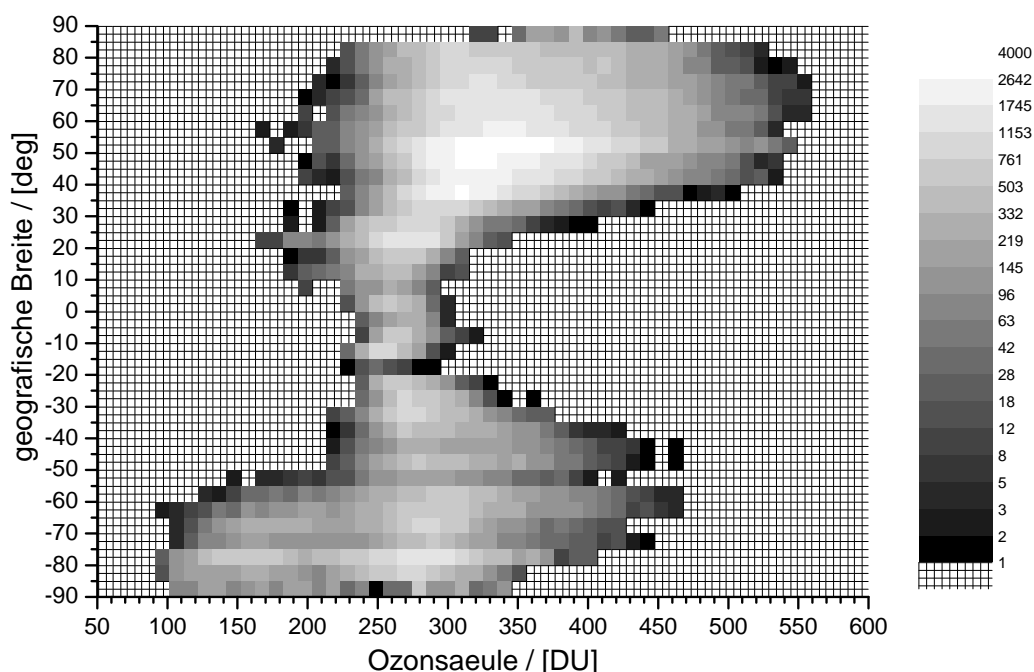


Abbildung 6.2: 2D-Histogramm der Gesamtozonwerte im Trainingsdatensatz. Die Skala gibt die Anzahl der Werte in jedem Rasterintervall (5° Lat. \times 10 DU Ozonsäule) an.

Die Verteilung der verbleibenden 321080 Kollokationen ist in Abb. 6.2 in Form eines zweidimensionalen Histogramms aufgetragen. Jeder Histogrammpunkt zählt die Anzahl der Ozonwerte, die in ein Intervall von 5° geografischer Breite mal 10 DU Ozonsäule fallen. Es ist deutlich zu sehen, dass die Gesamtozonwerte in der Nordhemisphäre im Mittel höher liegen als im Süden (vgl. Abschnitt 1.3), und dass die Dynamik des Ozonfeldes in mittleren und hohen Breiten größer ist als in den Tropen. Weiterhin fällt die bimodale Struktur der Verteilung polwärts von etwa 65° S auf, die den Unterschied zwischen Ozonloch- und Normalbedingungen kennzeichnet. Davon abgesehen konzentrieren sich die Ozonwerte in jedem Latitudenband um ihr Häufigkeitsmaximum, und die Verteilungen fallen in Form von Glockenkurven zu höheren und niedrigeren Werten hin ab.

In Abschn. 5.4 wurde schon erwähnt, dass die Zusammenstellung der Trainingsdaten immer einen gewissen Trade-Off zwischen optimaler Genauigkeit in Standard- und Extremsituationen bedingt [s. a. Diskussion in *Kaifel & Müller, 2000; Aires et al., 2001*]: Wird die Verteilung der Ozonwerte so wie sie ist akzeptiert und für die Netzwerkbelernung verwendet, dann erhält der Belernungsalgorithmus logischerweise die meiste (Fehler-)Information

aus Bereichen, in denen sich die Ozonwerte häufen, also um deren Mittelwert herum. Das neuronale Netz nimmt dadurch höhere Retrievalfehler bei extremen Ozonwerten in Kauf, weil diese viel seltener sind und bei der Bildung des Gesamt-RMSE einer Epoche (Gl. 4.4) demnach weniger ins Gewicht fallen.

Enthielte hingegen ein (hypothetischer) Trainingsdatensatz gleich viele Ozonwerte für jedes Intervall, so ergäbe sich auch eine Gleichverteilung der absoluten Retrievalfehler über alle Gesamtozonwerte. Für die Anwendung von NNORSY auf größere reale Datenmengen führte dies aber zu einem insgesamt erhöhten Fehler, weil die vielen mittelmäßigen Ozonwerte weniger genau bestimmt würden als im ersten Fall.

Da eine gleichmäßige Verteilung der Trainingsdaten-Ozonwerte in Ermangelung entsprechender Messungen ohnehin unmöglich ist, läuft die Zusammenstellung des Datensatzes auf einen Kompromiss zwischen beiden Möglichkeiten hinaus. Dieser Kompromiss wurde erzielt, indem die Maximalanzahl der Kollokationen in jeder Histogrammsäule aus Abb. 6.2 auf 500 begrenzt wurde. Die zu entfernenden, überzähligen Kollokationen innerhalb eines Histogrammbalkens wurden anhand ihrer Kollokationsdistanz aussortiert, beginnend mit den am weitesten vom GOME-Pixel entfernten Ozonmessungen. Damit werden gleichzeitig die Kollokationsfehler reduziert, und der Trainingsdatensatz schrumpft auf 208352 Werte. Abb. 6.3 zeigt die Verteilung aller kollokierten Ozonwerte vor und nach der Anzahlreduktion. Der Erfolg dieser Maßnahme wird in Abschnitt 6.4.3 diskutiert.

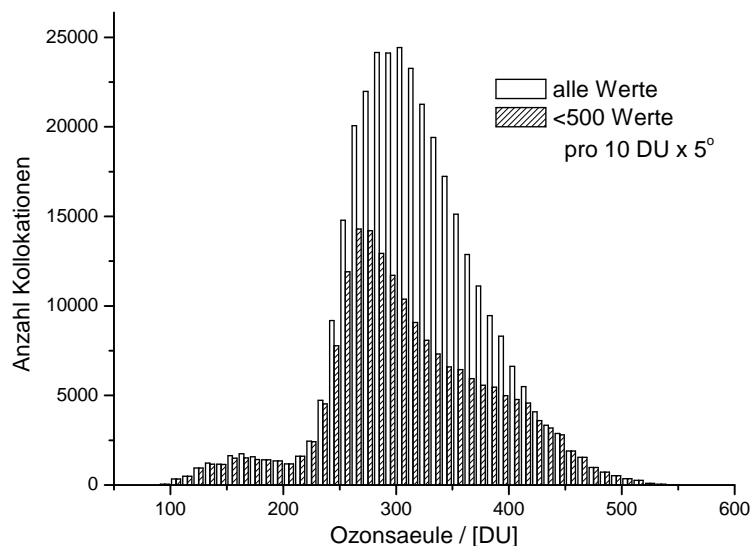


Abbildung 6.3: *Histogramm der Gesamtozon-Trainingskollokationen vor und nach der im Text beschriebenen Reduzierungsmaßnahme. Die Intervallbreite beträgt 10 DU. In der Legende wird auf das Histogramm in Abb. 6.2 Bezug genommen.*

6.3 Eingangsparmeter und Belernung

Für die Bestimmung von Gesamtozon genügt im Prinzip, wie im Abschn. 3.4.2 ausgeführt, ein relativ kleiner Spektralbereich in der Huggins-Bande des Ozons. Daher wurde zunächst das für die operationelle Gesamtozonbestimmung mit dem GDP eingesetzte Wellenlängenintervall 325–335 nm für die Belernung verwendet [Spurr, 1996; Burrows et al., 1999]. Vorstudien ergaben jedoch, dass die Ergebnisse sich durch die Verbreiterung des Intervalls auf 320–340 nm leicht verbessern lassen [Müller et al., 2001]. Um Redundanz und Messrauschen zu verringern wurden dabei jeweils zwei Spektralwerte koaddiert.

Im Gegensatz zu DOAS, welches nur auf kontinuierlichen Wellenlängenfenstern sinnvoll angewandt werden kann, spielt jedoch für NNORSY der Zusammenhang der Wellenlängen keine Rolle. Daher wurde auch eine alternative Konfiguration untersucht, bei der mehrere getrennte Spektralbereiche geringerer Auflösung zum Einsatz kamen (Tabelle 6.1). Die

Ozoninformation entstammt hier dem Huggins-Fenster 313–340 nm, allerdings wurde die spektrale Auflösung auf ca. 0.6 nm reduziert. Zusätzlich werden einige Spektralwerte im atmosphärischen Fenster zwischen 380 und 385 nm benutzt, so dass im Falle der Abwesenheit von Wolken ein klares Signal von der Erdoberfläche vorliegt, welches i. W. nur durch die NO₂-Absorption gestört wird. Die Verwendung einiger Werte aus den Chappuis-Banden um 600 nm ist für die Bestimmung des troposphärischen Anteils der Gesamtozonsäule, sowie zusammen mit dem vorherigen Fenster evtl. für die Korrektur des NO₂ Einflusses nützlich, welches in diesem Wellenlängenbereich etwas schwächer absorbiert [*de Beek*, 1998].

Tabelle 6.1: *Eingangsparameter des neuronalen Netzes für Gesamtozonretrieval. Die angegebenen Wellenlängenbereiche beziehen sich auf sonnennormierte GOME-Erdradianzen. SZA, SAA = Sonnenzenit- und -azimutwinkel bezogen auf Nordrichtung; ScanZA, ScanAA = dto., aber für Sensorblickrichtung bezogen auf Satelliten.*

Eingangsparameter	Coadding	Anzahl Werte	Zweck
313–340 nm	6	40	O ₃ Hartley/Huggins band
380–385 nm	8	7	Atmosphärisches Fenster / NO ₂
598–603 nm	8	3	O ₃ Chappuis-Bande / NO ₂
758–772 nm	8	9	O ₂ A-Bande: Wolkendetektion
SZA, SAA	-	2	Lichtwegskorrektur
ScanZA, ScanAA	-	2	Lichtwegskorrektur
line-of-sight Flags	-	3	Lichtwegskorrektur (AMF)
Latitude & Saison	-	2	Klimatologischer Background
Sensorenalter	-	1	Degradationskorr., Ozonfelldrift

Das langwellige Fenster 758–772 nm deckt die stärkste Sauerstoff-Absorption, die O₂ A-Bande ab [*Greenblatt et al.*, 1990]. Der Sauerstoff ist in der Atmosphäre gleichmäßig verteilt, daher ist die Absorption des Sauerstoffs ein Maß für die Höhe der atmosphärischen Säule, die das Licht auf seinem Weg von der Sonne bis zum Satelliten durchwandert hat. Bei Vorhandensein von Wolken beschränkt sich die durchwanderte Luftsäule auf den Bereich über der Wolkenschicht, und das Wegfallen der – aufgrund des höheren Luftdrucks besonders ausgeprägten – troposphärischen Sauerstoff-Absorption zeichnet sich in den GOME-Spektren deutlich ab [*Hoogen*, 1998]. Allerdings liefert die O₂-Absorption immer ein Mischsignal aus Wolkenbedeckungsgrad, Wolkenhöhe und -dichte. Der GDP bestimmt daher nur den Wolkenbedeckungsgrad aus dem Spektrum, und entnimmt Wolkenhöhe und -dichte einer Klimatologie [*Balzer & Loyola*, 1996]. Unabhängig davon können allerdings die räumlich besser aufgelösten PMD-Messungen zur Wolkenbestimmung beitragen [*Burrows et al.*, 1999]. Werden diese Messungen als zusätzliche Eingangsparameter für das neuronale Netz verwendet, bewirken sie jedoch keine signifikante Verbesserung des Retrievals (s. a. Kapitel 8). Vermutlich ist die durch die PMDs vermittelte, räumliche Information zu komplex und/oder ihre Auswirkung zu gering, um sie anhand der Trainingskollokationen zu erlernen. Eventuell könnte das Retrieval durch die Verwendung separat bestimmter, qualitativ hochwertiger Wolkenparameter als Eingangsdaten verbessert werden.

Andererseits ist die Ozoninformation in den Spektren ebenfalls ein Mischsignal – der Unterschied zwischen einem Pixel mit 50% Bedeckungsgrad auf 10 km Höhe und einem auf 5 km Höhe vollständig bedeckten Pixel entsteht grob betrachtet nur aus den O₃-Dichteunterschieden zwischen 0–5 km und 5–10 km, auf der Hälfte der Pixelfläche. Daher sollte der Belernungsalgorithmus bereits einen relativ guten (nichtlinearen) Zusammenhang zwischen den O₂(A)- und O₃-Mischsignalen finden können. Wie bereits in Abschnitt 5.3 angedeutet, wird daher von einer weiteren Untersuchung des Themas Wolken an dieser Stelle abgesehen.

Alles in allem ließ sich durch den Einsatz der vier verschiedenen Wellenlängenfenster der gesamte RMSE im Vergleich zur DOAS-ähnlichen Konfiguration um etwa 2% (0.25 DU) reduzieren, was im Rahmen der stochastischen Schwankungen zwischen einzelnen Netzwerkelernungen gerade noch als signifikant gewertet werden kann. Das Verfahren ist somit kaum sensibel gegenüber den Eingangswellenlängen, wie auch in verschiedenen Belernungsexperimenten mit variierenden Spektralbereichen, und -auflösungen, sowie unterschiedlicher Eingangsnormierung festgestellt wurde. Offenbar enthalten die GOME-Spektren genügend Redundanz, um dem Belernungsalgorithmus zu erlauben, flexibel auf sehr unterschiedliche Variationen der Eingangs-Konfiguration zu reagieren. Diesem Thema ist im Zusammenhang mit dem GOME-Ozonprofilretrieval der Abschnitt 8.1 gewidmet; Die dort gefundenen Ergebnisse bestätigen die hier für das Gesamtozon-Netzwerk getroffenen Aussagen im Kern. Alle folgenden Ergebnisse beziehen sich daher auf das neuronale Netz mit der in Tabelle 6.1 aufgeführten Eingangskonfiguration, welches mit den niedrigsten RMS-Fehlern aus den Testläufen hervorgegangen ist. Die Belernung nahm hier 43000 Epochen in Anspruch.

Bevor wir zum Ergebnisteil übergehen, muss jedoch noch kurz ein Punkt angesprochen werden, der bei klassischen Verfahren häufig Schwierigkeiten verursacht: Aus physikalischen Gründen läßt sich anhand der sonnennormierten Radianz nur die schräge Ozonsäule bestimmen, die beim klassischen DOAS-Retrieval anschließend mit Hilfe des **airmass-Faktors (AMF)** in eine senkrechte Säule umgewandelt wird (Abschn. 3.4.2). Zur Ermittlung dieses Faktors müssen gewisse klimatologische Annahmen gemacht werden, da eine exakte Berechnung mit Hilfe eines Strahlungstransfermodells voraussetzt, dass der Zustand der Atmosphäre inklusive des Ozonprofils bereits bekannt ist, was natürlich die Messung *ad absurdum* führt [Marquard *et al.*, 2000].

Beim NNORSY-GOME Gesamtozonretrieval entfällt die explizite Umwandlung in senkrechte Säulen, denn sie wird bereits während des Belernungsprozesses vom neuronalen Netz selber durchgeführt. Voraussetzung hierfür ist lediglich, dass das Netz mit Gesamtozonwerten trainiert, und die zur approximativen Bestimmung von AMFs notwendigen Parameter als zusätzliche Prädiktoren verwendet werden.² Diese Prädiktoren sind hier die LOS-Flags, der Scanwinkel, SZA und SAA, sowie für die klimatologische Information Latitude und Saison.

6.4 Ergebnisse und Validation des NNORSY-Gesamtozons

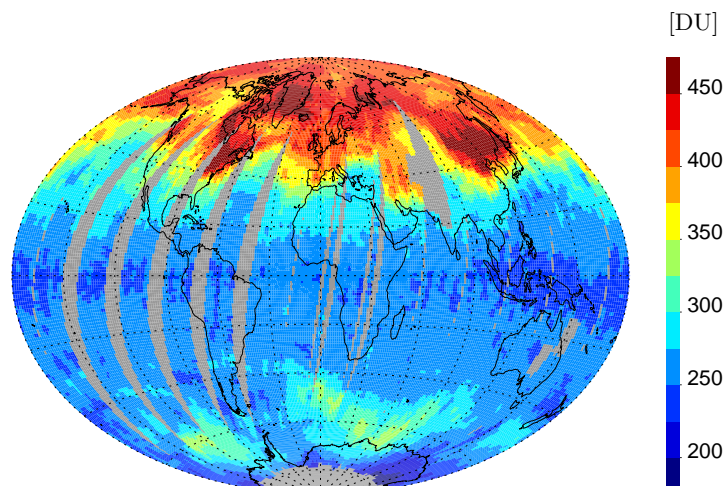
Abb. 6.4 gibt zunächst exemplarisch einen visuellen Eindruck der Gesamtozonwerte, die das aus der obigen Belernung resultierende neuronale Netz liefert. Der räumliche Zusammenhang des Ozonfeldes ist hier offensichtlich gewährleistet, denn es sind weder großflächige Artefakte, noch übermäßiges Rauschen der Einzelwerte oder starke systematische Abweichungen zwischen Ost-, Nadir-, und West-Pixel erkennbar. Ähnliches zeigt sich auch bei stichprobenartiger Überprüfung anderer Tage des Retrievalzeitraumes.

Es kann somit davon ausgegangen werden, dass die vorgeschlagene Methode für einzelne Pixel vernünftige Werte liefert. Die folgenden Abschnitte werden nun die statistischen Eigenschaften der NNORSY-Gesamtozonwerte – aufgeschlüsselt nach den Dimensionen Latitude, Zeit und Ozonwert – näher durchleuchten. Hierzu wurde der gesamte GOME-Datenbestand von 1996 bis Ende Juli 2001 prozessiert.

Zunächst sollen die Retrievalfehler im Bezug auf die Messungen der sechs global verteilten Teststationen untersucht werden (Abschn. 6.4.1). Dabei dienen operationelle TOMS

²In der Tat ist es schwierig, ein neuronales Netz mit dieser Methode auf schräge Ozonsäulen zu trainieren, denn diese müssten zunächst aus den bodengestützten Gesamtozon-Messungen für die jeweilige Position des Satelliten errechnet werden, was wiederum die approximative Bestimmung von AMFs voraussetzt.

Abbildung 6.4: Globales NNORSY-Ozonfeld gemittelt über 26. bis 28. März 2001, ohne Glättung und ohne Interpolation fehlender Werte. Die Rasterung erfolgte entsprechend dem TOMS-Standard auf ein $1.25^\circ \times 1^\circ$ Gitter (geogr. Länge \times Breite). Fehlende Werte erscheinen in grau. Die schmalen Streifen in der Mitte entstammen dem narrow swath Scanmodus des Sensors.



v7 und nicht interpolierte GDP v2.7 Ozonwerte als Referenz. Ein feineres Bild des Fehlerverhaltens im zeitlichem Regime geben die in Abschn. 6.4.2 präsentierten Zeitreihen. Abschnitt 6.4.3 wird kurz den Dynamikbereich der Retrievalausgabe im Vergleich zu den Kollokationen betrachten, bevor die Veränderung der Fehler mit der Latitude in Abschn. 6.4.4 anhand zonaler Mittelwerte und unter Einbeziehung aller verfügbarer Bodenstationen charakterisiert wird.

6.4.1 Teststationen-Statistik

Die wichtigsten Daten der in Abb. 6.1 markierten Teststationen sind in Tabelle 6.2 zusammengefasst.

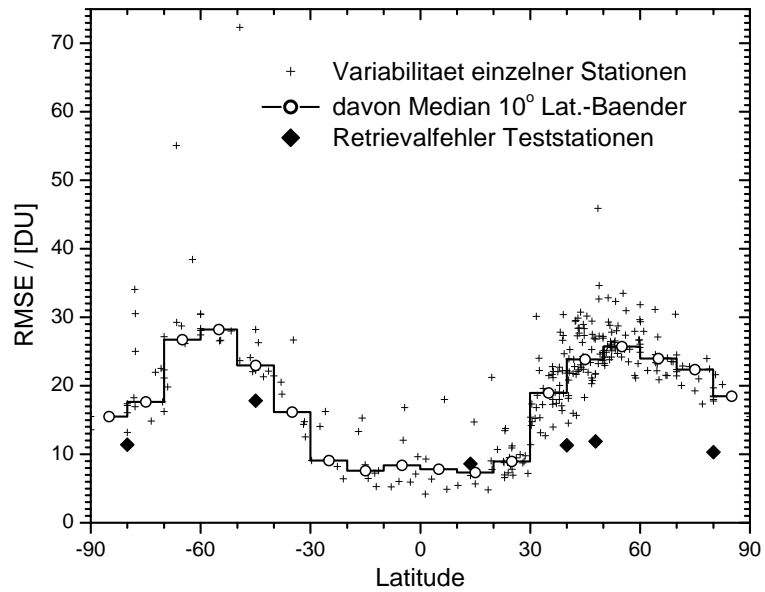
Tabelle 6.2: Daten der für die Kontrolle der Gesamtzonbelernungen verwendeten Teststationen. Die Anzahl der Werte bezieht sich auf Kollokationen mit gültigen NNORSY-Retrievals innerhalb 160 km, für den Zeitraum Jan. 1996 bis Juli 2001. TOMS Daten für die Generierung der virtuellen Stationen wurden wie in Abschn. 2.2.1 beschrieben modifiziert.

Stationsname	Instrument	Lat [$^\circ$ N]	Lon [$^\circ$ O]	WMO #	# Werte
Boulder	Dobson	40.0	-105.3	67	390
Hohenpeissenberg	Brewer (Mk II)	47.8	11.0	99	556
Bangkok	Brewer (Mk IV)	13.7	100.6	216	492
Lauder	Dobson	-45.0	169.7	256	450
Virtuelle Station 2	mod. TOMS	-80.0	-135.0	-	702
Virtuelle Station 18	mod. TOMS	80.0	45.0	-	885

Einige statistische Eigenschaften der Kollokationen von NNORSY mit den Teststationen sind der Tabelle 6.3 zu entnehmen. Die RMS-Retrievalfehler liegen generell um die 11 DU, für die tropische Station Bangkok etwas niedriger, dafür in Lauder deutlich höher.

Diese Werte sind am besten in Zusammenhang mit der Variabilität des Ozonfeldes zu interpretieren. In Abb. 6.5 ist zu diesem Zweck eine einfache Analyse der Differenzen von Gesamtzonmessungen jeweils zweier konsekutiver Tage dargestellt, also im Mittel die Variabilität innerhalb von 24 h. Den RMSE-Werten entsprechen relative Änderungen von

Abbildung 6.5: Variabilität des Ozonfeldes anhand der Differenz von Messwerten direkt aufeinanderfolgender Tage. Alle WOUDC-Bodenstationen und virtuelle TOMS-Stationen wurden verwendet, sofern für den Zeitraum 1979–2001 mehr als 100 Differenzwerte gebildet werden konnten. Zum Vergleich ist der NNORSY-Retrievalfehler für die Teststationen dargestellt.



etwa 3–4% innerhalb von 30 Breitengraden um den Äquator, etwa 7–8% in den nördlichen mittleren Breiten, 5% an den Polen und etwa 10% in dem Bereich, der im südpolaren Frühjahr regelmäßig vom Rand des antarktischen Ozonlochs überstrichen wird.

Man muss an dieser Stelle zwei Fehlerquellen klar differenzieren, die sich beide auf den Retrieval-RMSE auswirken: Zum einen wird das neuronale Netz auf Basis der Kollokationen von GOME-Pixeln mit Ozon-Tagesmitteln³ belernt, die also durchaus ein um mehrere Stunden τ verschobenes Ozonfeld sehen können und somit – bei angenommener linearer Zunahme des Fehlers mit der Zeit, vgl. [Fioletov *et al.*, 1999] – bereits einen entsprechenden Bruchteil ($\tau / 24$ h) der obigen Fehler beinhalten. Diese Fehlerquelle kann der Belernungsalgorithmus zum Teil kompensieren (s. dazu die Diskussion in Kapitel 10). Zum anderen gehen dieselben Fehler aber auch direkt in die Berechnung des Testdatenretrieval-RMSE

³Aus den WOUDC-Daten ist nicht ersichtlich, ob es sich bei den Tageswerten um Mittelwerte oder tatsächlich um eine einzelne Messung irgendwann im Verlauf des Tages handelt – beide Fälle sind möglich [E. Hare, persönliche Mitteilung, 1999].

Tabelle 6.3: Statistik der Gesamtozonwerte verschiedener Sensoren im Vergleich zu den gewählten Teststationen, basierend auf täglichen Level 3 Ozonkarten. Der relative Bias entspricht dem Wert (Sensor-Station)/Station. TOMS-Daten standen ab Juli 1996 zur Verfügung, nicht interpolierte GDP-Daten ab Aug. 1997. Die virtuellen Stationen enthalten nur TOMS-Daten, daher ist ein Vergleich mit TOMS nicht sinnvoll.

Stationsname	NNORSY			TOMS v7			GDP v2.7		
	RMSE [DU]	Bias [%]	σ [%]	RMSE [DU]	Bias [%]	σ [%]	RMSE [DU]	Bias [%]	σ [%]
Boulder	11.3	1.8	3.2	8.9	-0.7	2.7	15.4	1.1	4.8
Hohenpeissenberg	11.9	2.1	2.9	7.7	0.3	2.3	13.4	0.1	4.1
Bangkok	8.6	-1.5	3.0	12.9	2.7	4.3	9.5	-1.0	3.6
Lauder	17.8	0.2	5.8	14.4	2.8	3.9	26.1	-4.2	6.8
Virtuelle Station 2	11.4	-0.7	4.8	–	–	–	19.5	-2.2	7.7
Virtuelle Station 18	10.3	0.4	2.9	–	–	–	16.6	-0.5	4.9

ein.

Ein beträchtlicher Teil des Retrieval-RMSE kommt somit durch Kollokationsfehler zu Stande. Die RMSE-Werte sind ebenfalls in Abb. 6.5 dargestellt. Damit kann die Abbildung auch als Antwort auf eine andere, insbesondere für spätere Assimilations-Anwendungen interessante Fragestellung interpretiert werden: Um wie viel verbessert das NNORSY-Retrieval die Vorhersage des Ozonwertes an einem bestimmten Ort, wenn der Wert des Vortages bekannt ist, und als Baseline angenommen wird, er verändere sich innerhalb eines Tages nicht?

Wie zu erkennen, liegt die tägliche Variabilität des Ozonfeldes in Bangkok bereits im Bereich der Messunsicherheit. NNORSY kann hier im Vergleich zum Vortageswert keine Verbesserung erreichen. Bei den weiter nördlich gelegenen Stationen ist hingegen eine RMSE-Verbesserung um den Faktor 2 möglich, während sie in Lauder und über der Antarktis deutlich geringer ausfällt. Dabei ist zusätzlich zu bedenken, dass der Belernungsalgorithmus wie weiter oben erläutert diejenigen Bereiche des Messwertraumes am besten optimiert, in denen die Datendichte am größten ist [Bishop, 1995a]. Dies erklärt die Differenzen zwischen Nord- und Südhemisphäre, da die Kollokationsanzahl in der NH überwiegt, wie in Abschn. 7.1.3 erwähnt. Andererseits spricht es für die Flexibilität des NNORSY-Ansatzes, dass auch in der SH noch deutliche Verbesserungen erreicht werden.

In Tabelle 6.3 sind zusätzlich die statistischen Größen für GDP- und unkorrigierte TOMS-Daten angegeben. Die Streuung der NNORSY-Daten liegt in allen Fällen unter der des GDP, aber mit Ausnahme der Station Bangkok generell über der von TOMS. Die Werte stimmen mit den Analysen in [Bodeker *et al.*, 2001; Lambert *et al.*, 1999; McPeters *et al.*, 1998; Corlett & Monks, 2001] gut überein.

Im übrigen würde die vollständige Analyse aller bei NNORSY Gesamt Ozon auftretenden Fehler den Rahmen dieser Arbeit bei weitem sprengen, zumal der Schwerpunkt auf der Ozonprofilbestimmung liegt. Zweck der hier vorgestellten, schlaglichtartigen Untersuchungen ist es vielmehr, die qualitätsmäßige Einordnung der NNORSY-Daten relativ zu GDP v2.7 und TOMS v7 mit ausreichender Sicherheit zu ermöglichen.

6.4.2 Zeitreihen

Wichtig im Zusammenhang mit Trendanalysen ist die zeitliche Stabilität der Ozonretrievals. In Abb. 6.6 und 6.7 sind Zeitreihen der Differenzen zwischen gemessenen Ozonsäulen an den Teststationen und den kollokierten Satellitendaten dargestellt. Die Monatsmittelwerte wurden nur verwendet, wenn an mehr als fünf Tagen Differenzwerte vorlagen. Der Vollständigkeit halber sind hier bereits die integrierten Ozonsäulen aus der NNORSY-Ozonprofilbestimmung mit einbezogen, die erst in Abschn. 12 ausführlich diskutiert werden.

Einer der an diesen Graphen zu beobachtenden Effekte sind die saisonalen Schwankungen der Differenzen, die mitunter die Größenordnung der statistischen Schwankungen erreichen oder übertreffen [Fioletov *et al.*, 1999] und sich auf eine Reihe verschiedener Ursachen zurückführen lassen. Darunter sind die Variation des SZA mit der Jahreszeit, die Abhängigkeit des Fehlers von der Magnitude des Gesamt Ozonwertes, sowie das Verhalten der verwendeten Monatsklimatologien für TOMS, GDP und die Dobson bzw. Brewer-Instrumente [Lambert *et al.*, 1999; Corlett & Monks, 2001]. Die saisonalen Schwankungen sind offenbar bei den GDP-Monatsmitteln am größten, während sie bei NNORSY in der Größenordnung derer des TOMS-Sensors liegen.

NNORSY Gesamt Ozon weist in der ersten Jahreshälfte 1996 einen deutlich positiven Bias gegenüber den Stationen Boulder und Hohenpeißenberg auf. Da dieser Effekt bei den integrierten NNORSY-Ozonprofilen – die ähnliche Prädiktoren aber einen völlig unterschiedlichen Trainingsdatensatz verwenden – in ähnlicher Form auftritt, ist der Grund hierfür

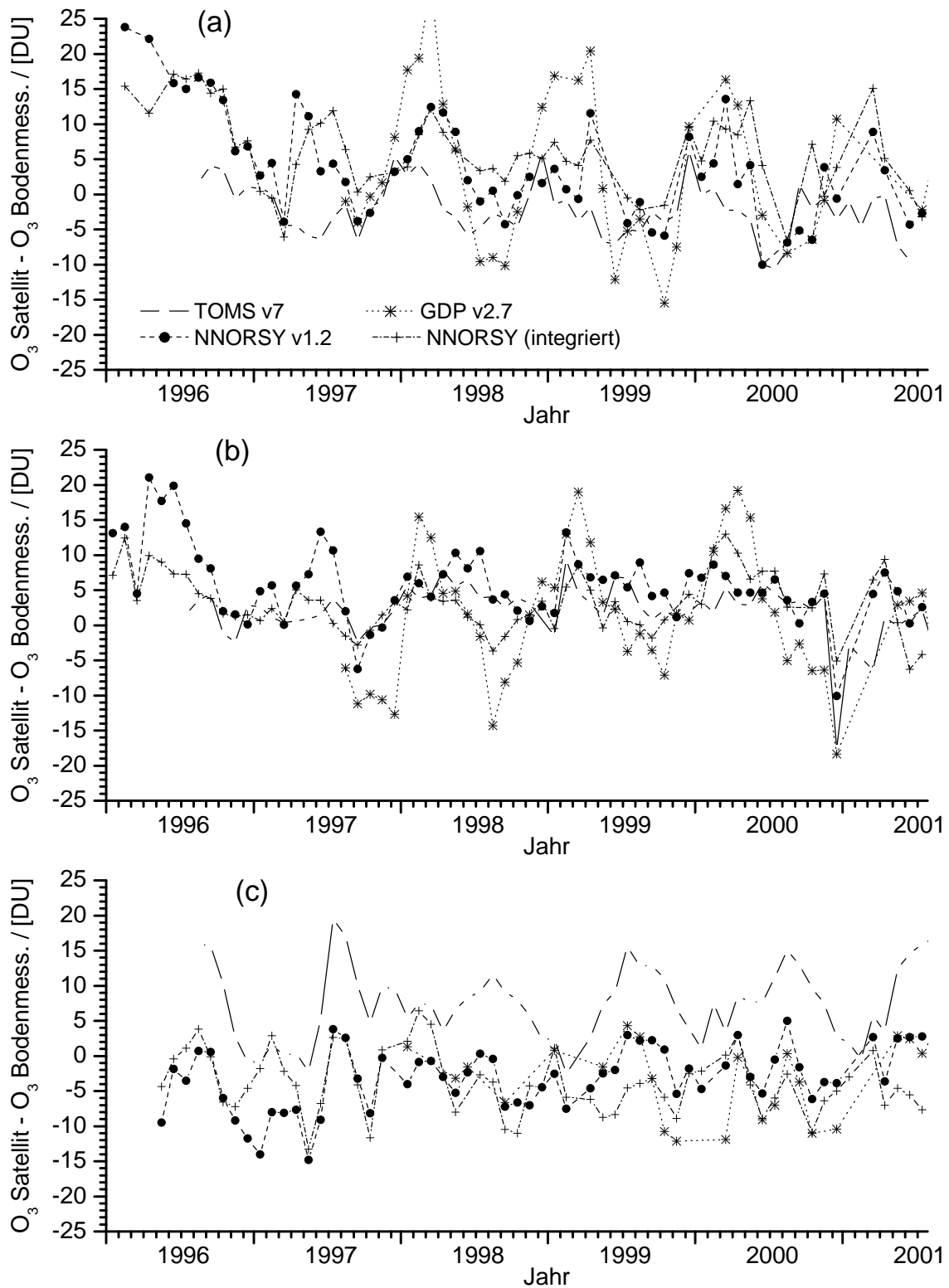


Abbildung 6.6: Zeitreihen der Differenz zwischen Satellitenmessungen und den verwendeten Teststationen (a) Boulder, (b) Hohenpeißenberg und (c) Bangkok.

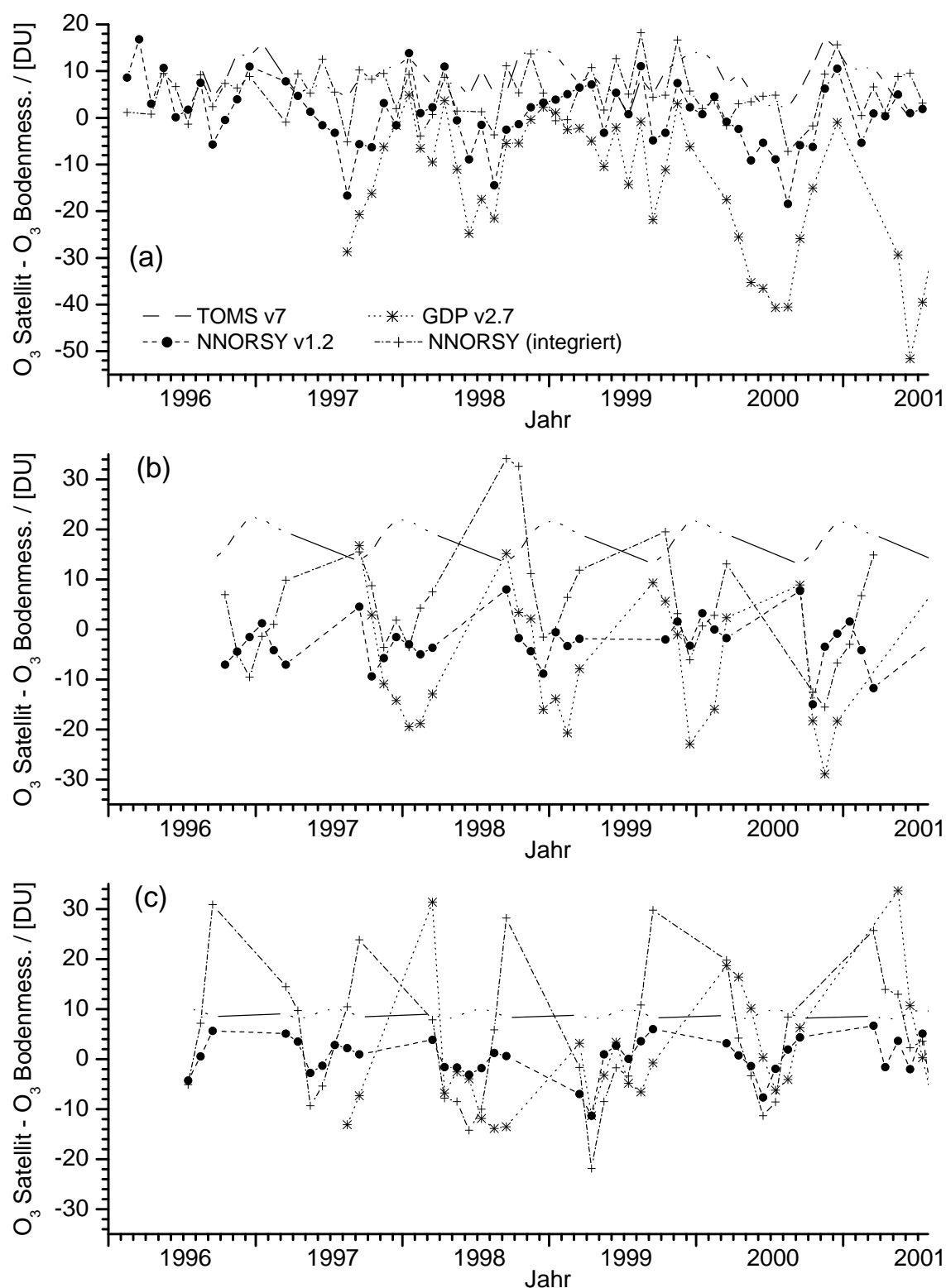


Abbildung 6.7: Wie Abb. 6.6, aber für (a) Lauder sowie für die virtuellen Teststationen bei $80^\circ S$ (b) und $80^\circ N$ (c). Bei (b) und (c) spiegelt die TOMS-Kurve lediglich die klimatologische Korrekturfunktion [Bodeker et al., 2001] wider.

vermutlich im GOME-Spektrum zu suchen: Vor dem 27. März 1996 wurde aufgrund der bei hellen Szenen möglichen Übersättigung der Detektor-Photodioden die Integrationszeit während eines Pixelscans auf ein viertel (0.375 s) begrenzt [Callies *et al.*, 1996]. Lediglich Band 1A war von diesem Vorgehen ausgenommen. Das Problem wurde später durch eine Softwarelösung („*coadding-patch*“) behoben. Es ist nicht auszuschließen, dass eine Veränderung der statistischen Eigenschaften zwischen vor und nach diesem Datum gemessenen Spektren zum Teil für die beobachteten Biases verantwortlich ist. Davon abgesehen sind jedoch keine weiteren Drifts der NNORSY-Datenreihen festzustellen.

Wie sich bereits in Tab. 6.3 andeutete, liegen die TOMS-Werte bei den Stationen Bangkok und Lauder weiter von den Bodenmessungen entfernt als NNORSY, während die Schwankungsbreite eine vergleichbare Größe aufweist. Diese systematische Abweichung spiegelt sich wie oben erwähnt in der von Bodeker *et al.* [2001] konstruierten Korrekturfunktion für TOMS wider, deren Auswirkung an den Polen in Reinform in den Diagrammen 6.7b und c zu erkennen ist. Während Corlett & Monks [2001] bereits berichteten, dass GDP-Retrievals im globalen Latitudenbandmittel für die Jahre 1996–1999 zwischen 50–60°S die absolut größten negativen Differenzen zu TOMS aufweisen, findet sich dieser Trend auch in den Daten aus Lauder, und scheint sich sogar nach 1999 noch deutlich zu verstärken. Beim Vergleich von Monatsmittelwert-Karten der verschiedenen Sensoren (Kapitel 12) fällt beim GDP zudem eine gewisse Ost-West-Asymmetrie auf. Diese Probleme wurden bei den NNORSY-Daten nicht beobachtet.

An den Polen stimmen die NNORSY-Gesamtozonwerte sehr gut mit den aus korrigierten TOMS-Daten bestehenden virtuellen Stationen überein. Allerdings streuen die TOMS-Werte insbesondere über der Antarktis auch beträchtlich um die Messungen von Bodenstationen [McPeters & Labow, 1996; Fioletov *et al.*, 1999; Bodeker *et al.*, 2001], welche wiederum aufgrund der widrigen äußeren Umstände höhere Fehler als üblich aufweisen können. Insofern ist es sehr schwierig, hier eine absolute Wahrheit im Bezug auf das Ozonfeld zu finden. Unter der hier getroffenen Annahme, dass das korrigierte TOMS-Ozonfeld der Wahrheit am nächsten kommt, zeigen sich sowohl bei GDP als auch bei den integrierten NNORSY-Daten relativ große Schwankungen der Differenzen. Beim GDP wurde dieser Effekt u. a. auf die fehlende Modellierung von Ozonlochsituationen in der verwendeten Klimatologie und auf generelle Probleme bei hohem SZA zurückgeführt [Lambert *et al.*, 1999; Marquard *et al.*, 2000]. Auf die integrierten NNORSY-Profile wird im Rahmen der Profilbestimmungskapitel näher eingegangen.

6.4.3 Abhängigkeit vom Ozonwert

Bei der Diskussion der Trainingsdaten-Zusammenstellung in Abschn. 6.2 wurde auf den Kompromiß hingewiesen, der zwischen Genauigkeit in häufigen und Genauigkeit in extremen Situationen eingegangen werden muss. Abb. 6.8a stellt das Ergebnis dieses Kompromisses für den vorliegenden Fall anhand eines Streudiagrammes dar. Dabei kamen Monatsmittelwerte des Gesamtozons aller derzeit⁴ im WOUDC-Archiv vorhandenen Gesamtozonsäulenmessungen zum Einsatz, um eine möglichst breite Streuung der Werte zu erreichen. Dass dabei auch die für die Netzwerkbelernung verwendeten Trainingsdaten in die Untersuchung eingehen, ist in diesem Fall unbedenklich, da einerseits die Vergleiche mit den Teststationen dem System eine gute Generalisierungsfähigkeit attestiert haben, und sich andererseits eventuelle Probleme mit der Ausgabedynamik auch in den Trainingsretrievals zeigen sollten.

In der Tat wird in der Abbildung eine leichte Reduktion des Dynamikbereiches von NNORSY gegenüber den Kollokationen beobachtet, die sich jedoch nur in den äußeren 40 DU des gesamten Wertebereiches durch eine leichte Tendenz weg von der Diagonalen und hin zu mittleren Werten bemerkbar macht. Der dadurch entstehende relative Fehler be-

⁴Stand: 14. Juni 2002

trägt am oberen Ende der Skala etwa 4%. Bei sehr niedrigen Ozonwerten kann er hingegen durchaus 10–20% erreichen. Diese Fälle treten jedoch nur äußerst selten auf: Von den über 8000 Messpunkten im Diagramm liegen lediglich 36 bei unter 140 DU, das sind weniger als 0.45%. Zudem ist zu bedenken, dass derart niedrige Ozonsäulen (bisher) ausschließlich über der Antarktis gemessen werden, wo wie oben ausgeführt die exakte Bestimmung von Ozonwerten aufgrund anderer Faktoren ohnehin relativ schwierig ist. Vor diesem Hintergrund kann der gefundene Kompromiss als sehr gut brauchbar gewertet werden. Er stellt auch eine deutliche Verbesserung gegenüber früheren Belernungen mit TOVS-Daten dar, bei denen die Kippung der Regressionsgeraden weg von der Diagonalen durchweg ausgeprägter war [Kaifel & Müller, 2000]. Der Vergleich mit entsprechenden Streudiagrammen für GDP- und TOMS-Daten (Abb. 6.8b und c) zeigt, dass auch diese Daten nicht völlig frei von Ozonwertabhängigkeit sind und im Bereich niedriger Ozonwerte eine Tendenz zur Überschätzung aufweisen.

Davon abgesehen weisen bei NNORSY auch Korrelationsgrad und die statistischen Größen für diesen Datensatz hervorragende Werte auf, obwohl angesichts der großen Anzahl der Kollokationen (103505 insgesamt) keine rigorose Qualitätskontrolle der Messungen vorgenommen wurde.

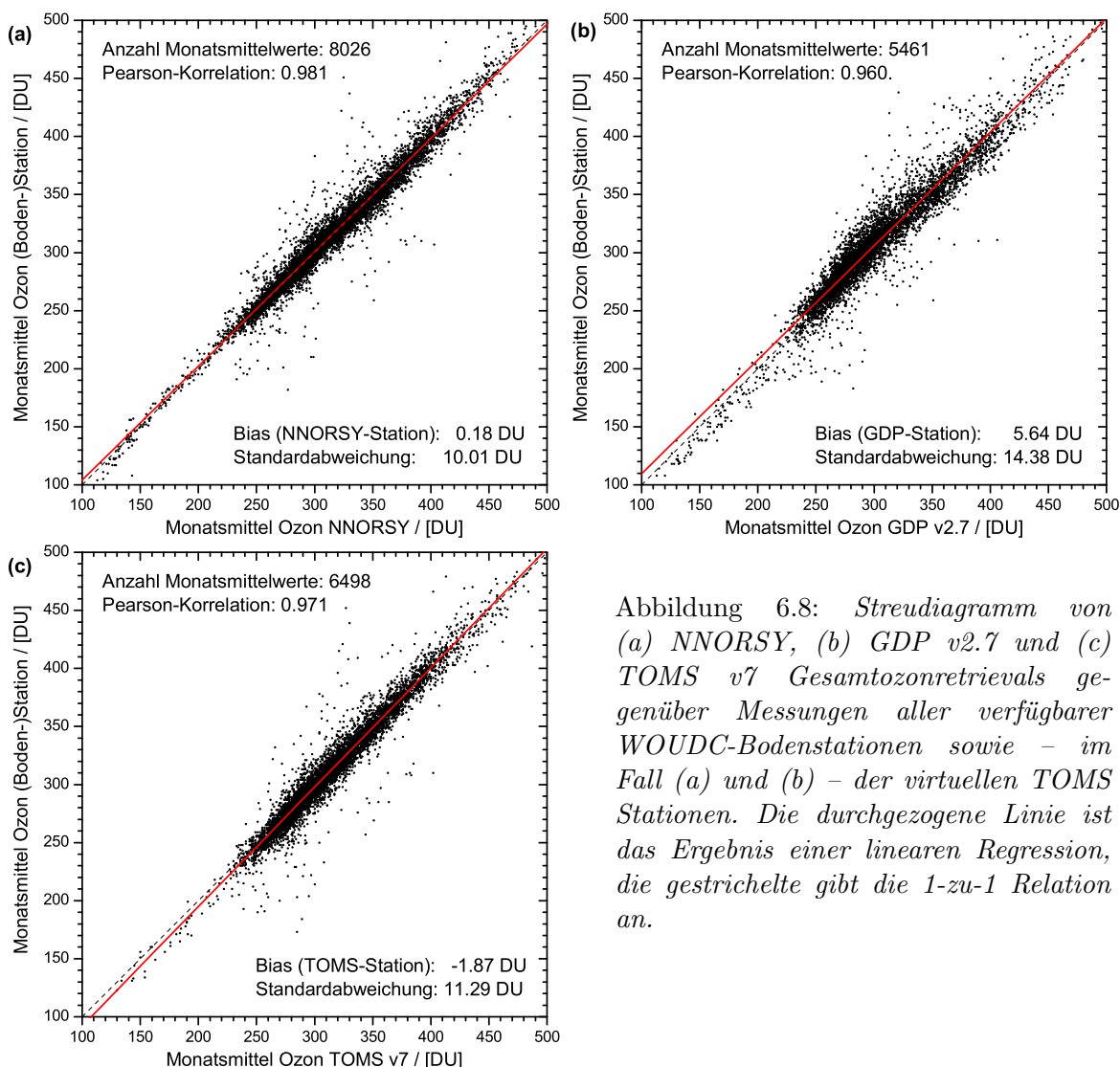


Abbildung 6.8: Streudiagramm von (a) NNORSY, (b) GDP v2.7 und (c) TOMS v7 Gesamtozonretrievals gegenüber Messungen aller verfügbarer WUODC-Bodenstationen sowie – im Fall (a) und (b) – der virtuellen TOMS Stationen. Die durchgezogene Linie ist das Ergebnis einer linearen Regression, die gestrichelte gibt die 1-zu-1 Relation an.

6.4.4 Zonale Mittelwerte

Die bisherigen Untersuchungen waren auf den räumlich eng begrenzten Bereich um die Bodenstationen herum beschränkt. Aber wie variiert die Qualität des Retrievals genau mit der Latitude?

Zur Beantwortung dieser Frage ist in Abb. 6.9a der gesamte NNORSY Ozondatenbestand als zonale Mittelwerte dargestellt. Zunächst läßt sich feststellen, dass bekannte Anomalien des Ozonfeldes wie etwa die extrem niedrigen Werte in der Umgebung von Neuseeland (etwa 45°S) im August 1997 [Brinkma et al., 1998] oder die ausgeprägte nordpolare Ozonabnahme im März 1997 [Müller et al., 1997; Eichmann et al., 1999] gut sichtbar sind, also nicht etwa durch die teilweise statistische Natur des NNORSY-Verfahrens verschmiert werden.

Aussagekräftiger ist jedoch der Vergleich mit dem korrigierten TOMS-Ozonfeld, welches hier wiederum als derzeit beste Annäherung an das wahre Ozonfeld verstanden wird (Abb. 6.9b). In den meisten Fällen liegt die Differenz zwischen den beiden Ozonfeldern unter 5 DU, nennenswerte Abweichungen treten nur am Rand der polaren Tag-Nacht-Grenze (dem *Terminator*) auf, wo die Sonnenzenitwinkel nahe bei 90° liegen. Interessanterweise ist in Abb. 6.9b von der im vorherigen Abschnitt diskutierten, leicht reduzierten Extremwertdynamik nicht viel zu bemerken; lediglich ein sehr eng begrenzter Latitudenbereich um den Südpol weist im September eine Überschätzung der im Zentrum des Ozonlochs sehr niedrigen Ozonsäulen auf. Hingegen werden die relativ hohen Ozonsäulen im nördlichen Winter offenbar systematisch überschätzt. Dieser Effekt ist auch in den Differenzen-Jahresgängen der Stationen Boulder und Hohenpeißenberg (Abb. 6.6a und b) ansatzweise erkennbar. Die Gründe hierfür sind momentan nicht ganz klar, aber da der Effekt ebenfalls bei GDP und bei den integrierten NNORSY-Ozonprofilen auftritt, hängt er vermutlich nicht mit den Trainingsdaten zusammen, sondern eher mit den GOME-Spektren oder mit dem hohen SZA. Die resultierenden relativen Abweichungen bewegen sich jedoch mit 4–7% in einem tolerablen Rahmen, daher wurde das Phänomen noch nicht in aller Ausführlichkeit ergründet. Einige weitere Erkenntnisse hierzu sind in den Fallstudien in Abschn. 12 dokumentiert. Es bleibt anzumerken, dass die zunehmende Degradation des TOMS Sensors ab Mitte 2000 zu einer merklichen Verschlechterung der Datenqualität führt⁵ so dass die Vergleiche gegen Ende der Zeitreihe mit Vorsicht zu behandeln sind.

Die in Abb. 6.9c dargestellten zonalen Differenzen zwischen GDP und TOMS bestätigen das bereits von Corlett & Monks [2001] gefundene Ergebnis, dass die GDP-Werte im globalen Mittel deutlich niedriger als TOMS liegen. Die Korrekturfunktion von Bodeker et al. [2001], die die TOMS-Ozonsäulen systematisch um 0–20 DU verringert, kann diesen Offset auch nur zum Teil ausgleichen. Allerdings treten nahe des südlichen polaren Terminators systematisch zu hohe Werte auf, die auf die in Abschnitt 6.4.2 erwähnten Probleme hindeuten. Insgesamt ergeben die Differenzen zwischen GDP und TOMS somit ein wesentlich unruhigeres Bild als diejenigen zwischen NNORSY und TOMS. Es darf dabei jedoch nicht unerwähnt bleiben, dass eine gewisse Glättung dieser Differenzen bei NNORSY durch die Verwendung von virtuellen TOMS-Stationen für das Netzwerktraining erreicht wird. Für die geophysikalische Qualität der Ozondaten, an der der Benutzer letztendlich interessiert ist, spielt jedoch deren algorithmische Herkunft keine große Rolle. Indes wird durch die geringen Abweichungen von NNORSY sowohl gegenüber den Bodenstationen als auch gegenüber dem korrigierten TOMS-Ozonfeld wiederum demonstriert, dass sich diese beiden Quellen ausgezeichnet zu einem konsistenten Bild des Ozonfeldes vereinigen lassen, was die Resultate in [Bodeker et al., 2001] und [Müller et al., 2002a] bestätigt.

⁵Quelle: <http://jwocky.gsfc.nasa.gov/news/news.html>

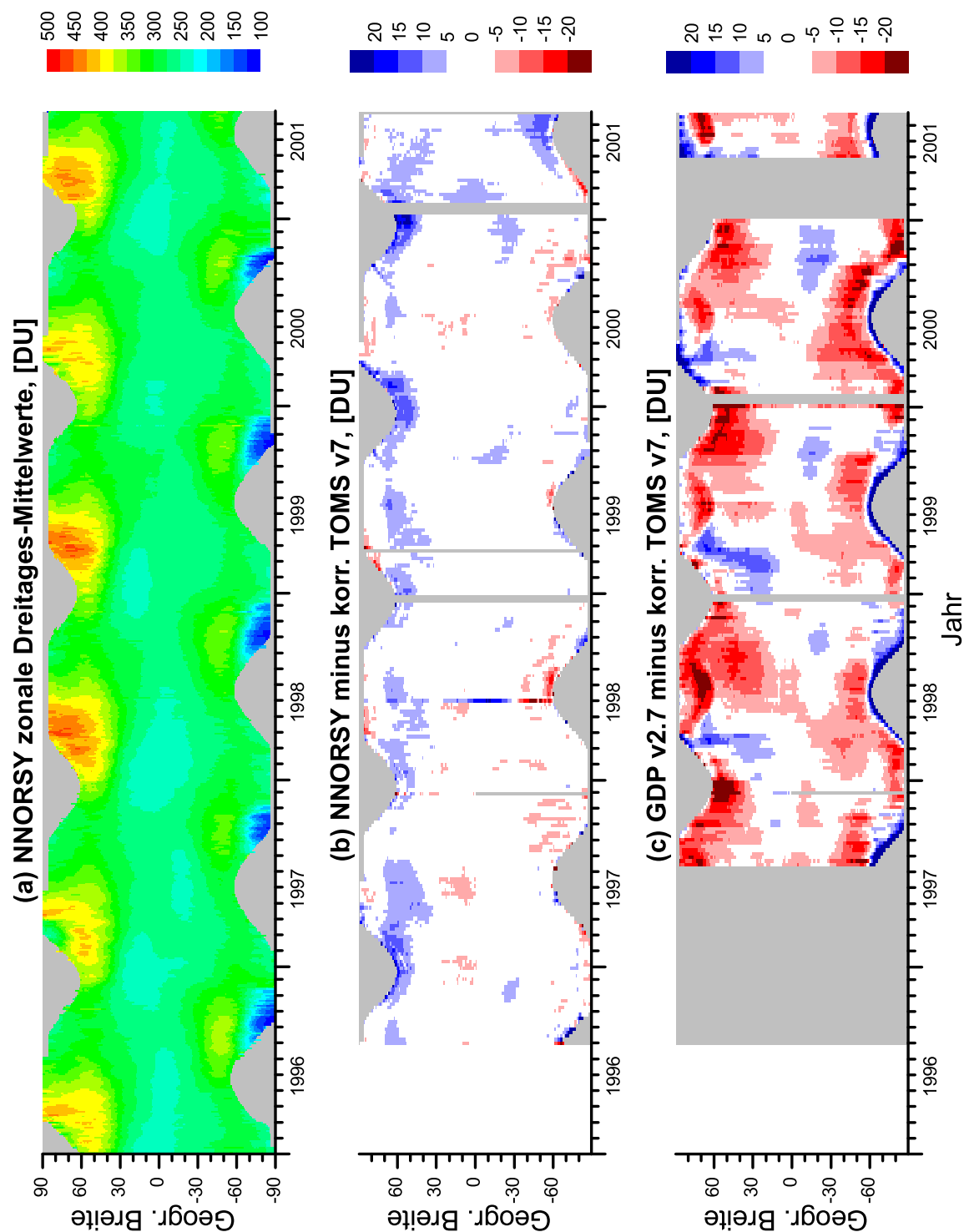


Abbildung 6.9: Zonale Mittelwerte (180 Breitenkreise) von NNORSY Gesamtozon und Differenzen zu korrigierten TOMS-Daten. Für die Differenzbildung wurden Wochenmittel der Level 3 Ozonkarten verwendet [vgl. Corlett & Monks, 2001].

Kapitel 7

Bestimmung von Ozonprofilen

7.1 Ozonprofil-Trainingsdaten

Der Datensatz zum Training von Profil-Netzwerken setzt sich wie wie bereits erwähnt aus Kollokationen mit verschiedenen Profilmessungen zusammen. In diesem Unterkapitel wird die Aufbereitung der Rohkollokationen in eine für neuronale Netze geeignete Form beschrieben. Zunächst werden alle Profile auf ein einheitliches Gitter interpoliert (Abschn. 7.1.1). Anschließend kann die geografische und zeitliche Verteilung der Daten optimiert werden (Abschn. 7.1.2). Im letzten Abschnitt 7.1.3 dieser Sektion werden dann die angewandten Methoden zur Reduzierung der verbleibenden systematischen und zufälligen, höhenabhängigen Fehler im Trainingsdatensatz aufgezeigt.

7.1.1 Interpolation der Profildaten

Die aus verschiedenen Datenquellen stammenden Ozonprofile zeichnen sich durch unterschiedliche Höhenauflösung und durch unregelmäßige vertikale Abstände zwischen den Messpunkten aus (Abb. 7.1).

Die Verarbeitung durch ein neuronales Netz setzt ein festes Höhengitter voraus, welches nicht unbedingt regelmäßig, aber auf jeden Fall für jedes verwendete Ozonprofil gleich sein muss. Für die vorliegende Arbeit wurden alle Profildaten durch numerische Integration auf ein Gitter mit Schichten von 1 km geopotenzieller Höhe (GPH) abgebildet, wobei die Schichtmitten bei 1, 2, ..., 60 km liegen.

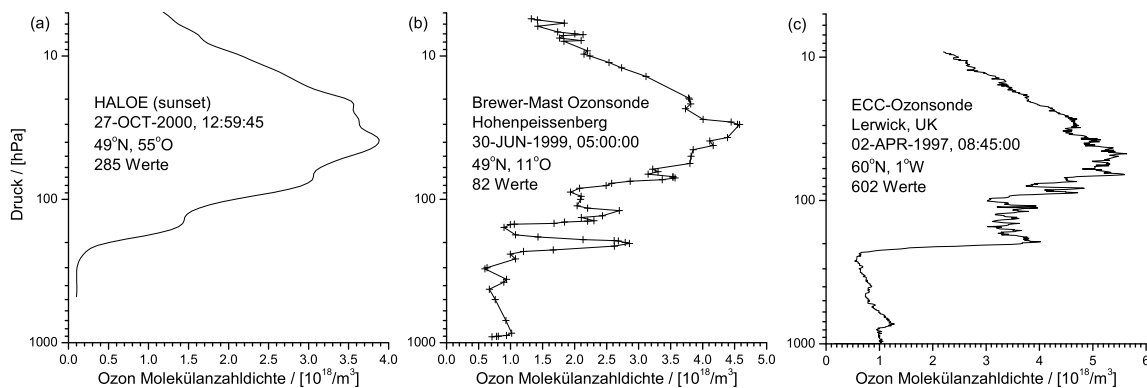


Abbildung 7.1: *Typische Beispiele für drei Arten von Ozonprofildaten. (a) Durch optische Fernerkundung gewonnenenes Okkultations-Profil. (b) Niedrig aufgelöstes Ozonsondenprofil. (c) Hochaufgelöstes Ozonsondenprofil*

Im Rahmen von Vorversuchen waren auch Belernungen auf der Basis von Druckniveaus oder geometrischer Höhe durchgeführt worden. Die Unterschiede zu den GPH-Schichten sind sehr gering, tendenziell sind die Ergebnisse jedoch minimal schlechter, da in der GPH mehr Information über den vertikalen Aufbau der Atmosphäre enthalten ist. Diese kann dem Verfahren z. B. bei der Unterscheidung von Troposphäre und Stratosphäre möglicherweise von Nutzen sein [Roedel, 2000].

7.1.2 Geografische und zeitliche Gleichverteilung

Für die Belernung neuronaler Netze ist es wünschenswert, dass die Verteilung der Kollokationen über Latitude relativ gleichförmig ist, damit Situationen, die nur in bestimmten Breitengraden auftreten, nicht überbewertet werden. Allerdings ist zu berücksichtigen, dass die Variabilität des Ozonfeldes ebenfalls stark breitenabhängig ist: Die strahlungsdominierte Ozonverteilung in der tropischen Stratosphäre neigt z. B. zu weit geringeren Schwankungen, als diese in den mittleren Breiten üblich sind. Um die Anzahl der unterschiedlichen atmosphärischen Situationen im Trainingsdatensatz zu maximieren, ist es also durchaus berechtigt, mehr Kollokationen aus den variableren Bereichen zu verwenden [Bishop, 1994]. Alle hier beschriebenen Maßnahmen wurden auf den gesamten Kollokationsdatensatz angewandt, d. h. vor dessen Aufteilung in Test- und Trainingsdaten. Bei einem Kollokationsradius von maximal 250 km im Zeitraum Januar 1996 bis Juli 2001 ergaben sich 203131 Rohkollokationen.

Die Verteilung der Kollokationen kann nun durch breiten- oder stationsabhängige, gezielte Verringerung des maximalen Kollokationsradius' optimiert werden. Diese Aufgabe läßt sich schlecht automatisieren, daher wurden die durchzuführenden Maßnahmen nach gründlicher Durchmusterung der Kollokationsstatistiken empirisch festgelegt. Einen gewissen Sonderfall stellt POAM III dar, da dieses Instrument nur in eng begrenzten Latitudenbändern um 70°S und 60°N Messungen liefert (Abschn. 2.2.4). Hier wurden nur Kollokationen der Südhemisphäre, sowie nördlich von 65°N verwendet – in den nördlichen mittleren Breiten herrscht durch die hohe Bodenstationsdichte ohnehin ein solches Überangebot an Kollokationen, dass die dortigen POAM-Messungen für spätere Validierungszwecke zurückgehalten wurden. Die Beschreibung der Maßnahmen im einzelnen ist dem Anhang C zu entnehmen.

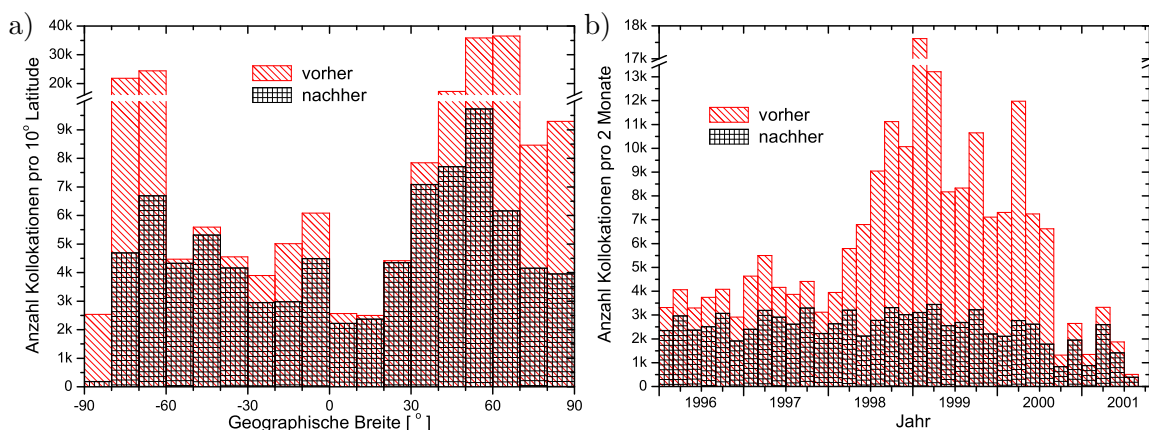


Abbildung 7.2: Verteilung der Ozonprofil-Kollokationen für den Zeitraum 1996 bis Juli 2001 über Latitude und Zeit, vor und nach der durchgeführten Datensatz-Homogenisierung.

Abb. 7.2a zeigt die ursprüngliche und die reduzierte Verteilung der Kollokationen über die geografische Breite.

Die zeitliche Verteilung der Kollokationen über die Jahre 1996–2001 ist ebenfalls sehr ungleichmäßig (Abb. 7.2b), insbesondere da POAM- und SHADOZ-Daten erst ab Mitte

1998 zur Verfügung stehen. Hier bietet sich als Ziel auf jeden Fall eine Gleichverteilung an, damit etwaige Trends im globalen Ozonfeld sowie die zeitabhängige Degradation des GOME-Sensors mit über die Zeit unveränderter Gewichtung in die Belernungen eingehen. Nach dahingehender Optimierung (Anhang C) wurden die verbleibenden Kollokationen zufällig in Trainings- und Testdaten aufgeteilt, wobei Mehrfachkollokationen nicht getrennt wurden (vgl. Abschn. 5.4). Der Trainingsdatensatz enthielt 70048 Kollokationen, der Testdatensatz 12281.

7.1.3 Homogenisierung der Trainingsdaten

Wie bereits im Abschnitt 2.2 angesprochen, liefern die verschiedenen Ozonmessverfahren und -sensoren mitunter recht unterschiedliche Werte [Lu et al., 1997; Steele & Turco, 1997; Deniel et al., 1997; SPARC, 1998]. Abgesehen von der Schwierigkeit, falsche Messwerte aufzuspüren und auszusondern, zeigt sich, dass systematische Differenzen meist nur behoben werden können, indem ein großer Teil der Messungen entfernt und damit die natürliche Varianz des Ozonfeldes verringert wird. Im Zusammenhang mit der Belernung neuronaler Netze ist es wichtig, zwischen Repräsentativität und notwendiger Menge der Trainingsdaten auf der einen Seite, und deren Homogenität auf der anderen Seite einen guten Kompromiß zu finden. Ein kleiner und homogener Datensatz deckt möglicherweise nicht alle auftretenden atmosphärischen Situationen ab und verringert damit die Flexibilität des neuronalen Netzes, während ein allumfassender, mit systematischen Biases und großen Fehlern kontaminierter Datensatz zu einem sehr flexiblen, aber ungenauem Netz führt. Bishop [1995a] spricht in diesem Zusammenhang von einem *bias-variance* Dilemma.

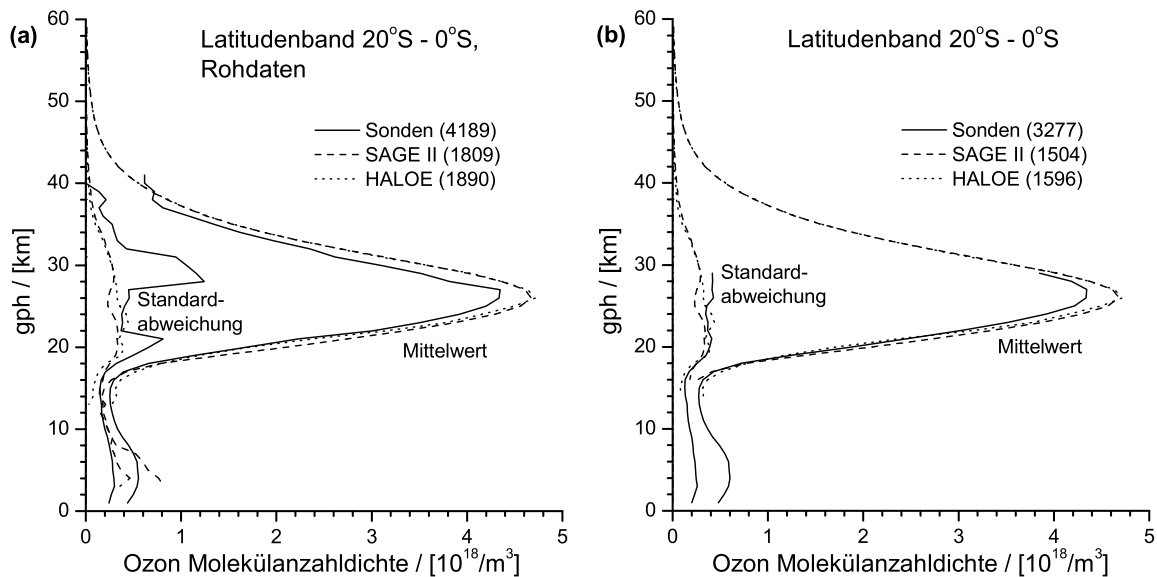


Abbildung 7.3: Mittelwert und Standardabweichung der Profilkollokationen (a) vor und (b) nach der Homogenisierung des Datensatzes. Die jeweilige Anzahl von Kollokationen ist in Klammern angegeben. HALOE und SAGE-Profile sind in diesem Latitudenband gleichverteilt, die Sondenmessungen liegen im Mittel näher am Äquator.

Die Qualitätssicherung und Auswahl geeigneter Profil-Belernungsdaten erfolgt in mehreren Schritten. Bereits bei der Kollokation und Interpolation der Profildaten werden Profile mit offensichtlichen Fehlern wie negativen Werten oder Ozonmoleküldichten von $> 15 \cdot 10^{18} m^{-3}$ komplett ausgesondert. Für die Okkultationsmessungen liegen zudem höhenabhängige Präzisionsfehlerschätzungen vor. Ihnen gemäß werden alle einzelnen Ozonwerte

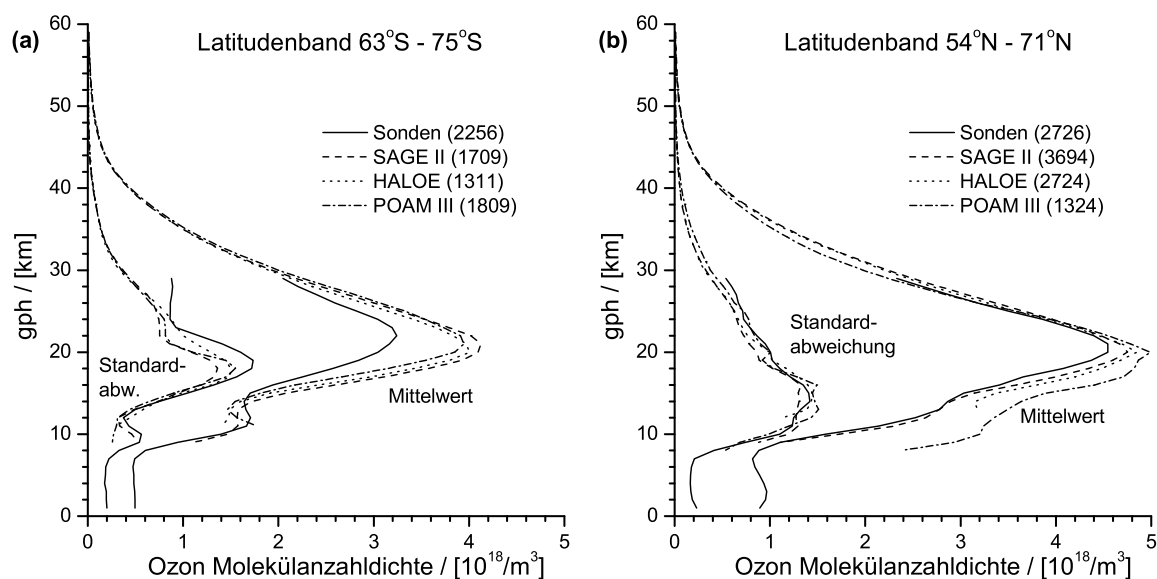


Abbildung 7.4: Wie Abb. 7.3, aber für zwei Bänder nach der Homogenisierung. In Diagramm (a) gruppieren sich alle Sonden um 71°S herum, in (b) liegen POAM-Daten nur nördlich von 65° vor.

mit $>15\%$ relativem Fehler entfernt. Leider verbleibt trotzdem eine zu große Zahl von Ausreißern und systematischen Fehlern im Datensatz, die die subsequente Anwendung einer Reihe zusätzlicher heuristischer Maßnahmen erforderlich macht. Zur statistischen Untersuchung der Datensatz-Diskrepanzen wurde der Rohdatensatz von 80°S bis 80°N auf solche Weise in $10\text{--}20^{\circ}$ breite Latitudenbänder aufgeteilt, dass die Kollokationen in den Streifen einigermaßen gleichverteilt sind. Abb. 7.3a zeigt hierzu als Beispiel die Rohdaten des Bandes $0\text{--}20^{\circ}\text{S}$. Ausgehend von diesen Statistiken und Angaben aus der einschlägigen Literatur (s. Kapitel 2) wurden weitere Homogenisierungsschritte auf den Rohdatenbestand angewandt, deren Details sich ebenfalls in Anhang C finden.

In Abb. 7.3b ist das Latitudenband $0\text{--}20^{\circ}\text{S}$ nach den Homogenisierungsmaßnahmen dargestellt. Die in Abb. 7.3a erkennbaren Probleme mit der Standardabweichung der Sonden in großen Höhen, der verschwindenden Variabilität von HALOE bei 13 km und den zu niedrigen SAGE-Werten unterhalb von 16 km wurden eliminiert. Abb. 7.4a und b stellen den homogenisierten Trainingsdatensatz in zwei anderen Bändern der Süd- und Nordhemisphäre dar.

Fazit

Aufgrund der zusätzlichen Höhendimension und der Verwendung von nunmehr fünf unterschiedlichen Datenquellen (WOUDC, SAGE, HALOE, POAM, SHADOZ), erweist sich die Homogenisierung der Ozonprofil-Trainingsdaten als deutlich komplexer als die der Gesamt-ozondaten. Trotzdem konnte durch den Einsatz verschiedener mathematischer, physikalischer und empirischer Hilfsmittel ein für die Belernung neuronaler Netze geeigneter Datensatz erzeugt werden. Der hierzu betriebene nicht unerhebliche Aufwand läßt sich durch die Wichtigkeit solider Trainingsdaten für den Erfolg des NNORSY-Verfahrens rechtfertigen. In wie weit die zwangsläufig im Datensatz verbleibenden Inhomogenitäten das Retrievalergebnis beeinflussen, wird weiter unten an verschiedenen Stellen diskutiert werden.

7.2 Parziale Belernung neuronaler Netze

Die meisten in der Literatur erwähnten Ansätze von MLP-Retrievals in der Atmosphärenphysik (Abschn. 4.4) verwenden simulierte Spektren oder die Ergebnisse eines klassischen Retrievalverfahrens zur Generierung von Trainings- und Testdatensatz. Damit wird ein Problem vermieden, welches bei der Verwendung von Kollokationen unweigerlich auftritt: Sowohl vertikale Rasterweite als auch gültiger Höhenbereich der Zielprofile können sehr unterschiedlich ausfallen. Während das Problem der unterschiedlichen Rasterung durch geeignete Interpolation gelöst werden kann (Abschn. 7.1.1), stellt sich die Frage, wie ein neuronales Netz mit Ozonprofilen belernt werden kann, von denen diejenigen aus Sondenmessungen nur bis 25–30 km Höhe hinauf reichen, während diejenigen aus Satellitenmessungen selten tiefer als bis 10–15 km Höhe in die Atmosphäre hinunter reichen. Verschiedene mögliche Lösungsansätze für dieses Dilemma wären:

- Die Profile könnten mit Hilfe einer **Klimatologie** ergänzt werden. Dies führt aber zu Sprüngen an der Nahtstelle zwischen individuellem Profil und Klimatologie, die nicht durch Skalierung der Klimatologie behoben werden können, ohne Artefakte einzuführen. Außerdem wird die Methode so von einer weiteren Datenquelle zweifelhafter Genauigkeit abhängig.
- Ideal wären **Tripelkollokationen** zwischen Sonde, Satellitenprofil und GOME-Pixel, aber davon gibt es viel zu wenige, um ein sinnvolles Training zu ermöglichen.
- Man könnte zwei **verschiedene Netze** für Sonden- und Satellitendaten trainieren, doch würde dies schon aufgrund der ungleichmäßigen Verteilung der Sondenstationen sicher zu systematischen Abweichungen zwischen den beiden Retrievals führen, die untersucht und geeignet ausgeglichen werden müssten.
- Der hier realisierte Ansatz modifiziert stattdessen den **Belernungsalgorithmus**, so dass auch unvollständige Profile für die Belernung verwendet werden können.

7.2.1 Modifikation von RPROP

Ausgehend von einem unvollständigen Kollokationsdatensatz wird zunächst für jeden Zielwert t_j^p ein Fehlerflag m_j^p bestimmt, welches bei Vorhandensein des Wertes in einer Höhengschicht j gleich eins ist, bei dessen Fehlen hingegen gleich null. Anschließend kann der Ausdruck für den Gesamtfehler der Epoche (Gl. 4.5) wie folgt modifiziert werden:

$$E_p = \frac{1}{2} \left(\sum_{j=1}^{N_K} m_j^p \right)^{-1} \sum_{j=1}^{N_K} (t_j^p - o_j^p)^2 \cdot m_j^p. \quad (7.1)$$

Um die Eigenschaften des neuen Algorithmus – der i. F. als **Parziale Belernung** bezeichnet wird – zu untersuchen, wurde er zunächst an einem simulierten GOME-Datensatz getestet. Dies hat den Vorteil, dass keine unkontrollierbaren Messungenauigkeiten das Ergebnis beeinflussen können. Die Simulationen basieren auf zwei Klimatologien, von denen eine dem MPI 2D Chemie-Transportmodell [*Christoph Brühl, private Mitteilungen, 1999*] entstammt, die andere wurde von *Fortuin & Kelder [1998]* aus SAGE, HALOE und Ozonsondendaten konstruiert. Mark Weber (IUP, Universität Bremen) skalierte die in ihnen enthaltenen Ozon- und Temperaturprofile in physikalisch sinnvollem Rahmen und verwendete sie zusammen mit unterschiedlichen Beobachtungsgeometrien und Oberflächenbedingungen als Eingabe für das pseudo-sphärische Strahlungstransportmodell GOMETRAN [*Rozanov et al., 1997*]. Auf diese Weise wurde ein Trainingsdatensatz von 17294 simulierten GOME-Spektren und zugehörigen Ozonprofilen auf 34 Höhengschichten erzeugt. Der Testdatensatz

enthielt 1925 derartige Kollokationen, bei denen andere Ozonprofile, sowie Zwischenwerte für die skalaren Parameter Verwendung fanden.

Das zu belernende MLP besaß 34 Ausgabe-, 35 versteckte und 98 Eingabeneuronen, die in Tabelle 7.1 näher spezifiziert sind. Da es in diesem Abschnitt primär um die Validierung des neuen Algorithmus' geht, sollen die Eingabeparameter hier nicht weiter diskutiert werden, der Leser sei hierfür auf die Abschnitte 6.3 und 7.3 verwiesen. Die Auswahl des Testproblems ist auch nicht besonders kritisch; Sehr ähnliche Ergebnisse wurden zum Beispiel mit dem in der Literatur über neuronale Netze gerne als Referenz benutzten Encoder-Decoder-Problem erzielt [Müller *et al.*, 2002b], welches kaum Gemeinsamkeiten mit der Ozonbestimmungs-Problematik besitzt.

Tabelle 7.1: *Eingangsparameter neuronaler Netze für simuliertes Ozonprofilretrieval. Der angegebene Wellenlängenbereich bezieht sich auf sonnennormierte Erdradianzen, die mit die GOMETRAN Strahlungstransfermodell [Rozanov *et al.*, 1997] berechnet wurden.*

Eingangsparameter	Anzahl Werte	Zweck
290–310 nm	61	O ₃ Hartley/Huggins-Banden
Temperaturprofil	34	Starke Korrelation mit O ₃
Sonnenzenitwinkel	1	Lichtwegskorrektur
Bodenalbedo	1	Bodenkorrektur
Bodenhöhe über N.N.	1	Troposph. O ₃ Korrektur

Für verschiedene Belernungen wurden im Trainingsdatensatz 0, 10, 50, 80, 90, 95 und 98 Prozent der Zieldaten zufällig ausgewählt und als fehlend markiert, während der Testdatensatz unverändert blieb. Allerdings ist die zufällige Auswahl der Daten nicht besonders realistisch, da in der Praxis die Ozonwerte üblicherweise zusammenhängend vom oberen Ende des Profils her fehlen (Ozonsonden), oder von unten her (Satellitenmessungen). Um eine etwas realistischere Verteilung fehlender Daten zu simulieren, wurden deshalb zwei weitere Datensätze mit 10 und 50% an fehlenden Werten konstruiert. Hier erfolgte zunächst eine zufällige Zuweisung der Anzahl fehlender Werte an jedes Zielprofil, unter Beachtung des Gesamt-Prozentsatzes. Anschließend wurde die zugewiesene Anzahl Werte bei jedem Profil mit gleicher Wahrscheinlichkeit entweder von oben oder von unten her als fehlend markiert. Offensichtlich ist es kaum möglich, auf diese Weise Trainingsdatensätze mit wesentlich weniger als 50% Daten zu erzeugen, da dies unweigerlich zu Lücken in mittleren Höhen der Profile führt, in denen nur sehr wenige oder gar keine gültigen Ozondaten mehr vorhanden sind.

7.2.2 Ergebnisse der Tests

Einige Testbelernungen bis Epoche 30000 zeigten, dass die Abnahme des RMSE während der Belernung sehr gleichmäßig geschieht und ab einigen 1000 Epochen fast linear verläuft [Müller *et al.*, 2002b]. Daher wurden in den folgenden Experimenten für jeden der erzeugten unvollständigen Trainingsdatensätze jeweils zehn neuronale Netze mit partieller Belernung bis Epoche 3000 trainiert. Die Ergebnisse der zehn Trainingsläufe streuen aufgrund der Stabilität von RPROP untereinander nur schwach (Abb. 7.5) und lassen insgesamt folgende Schlüsse zu:

- Die simulierten Trainingsdaten sind stark redundant, eine merkliche Verschlechterung der Ergebnisse des Testdatensatzes tritt erst auf, wenn zwischen 50 und 80% der Werte fehlen. Gründe für die Redundanz sind vermutlich die Absenz von Messfehlern in den simulierten Spektren, sowie die – durch das Monatsmittel bedingte – recht

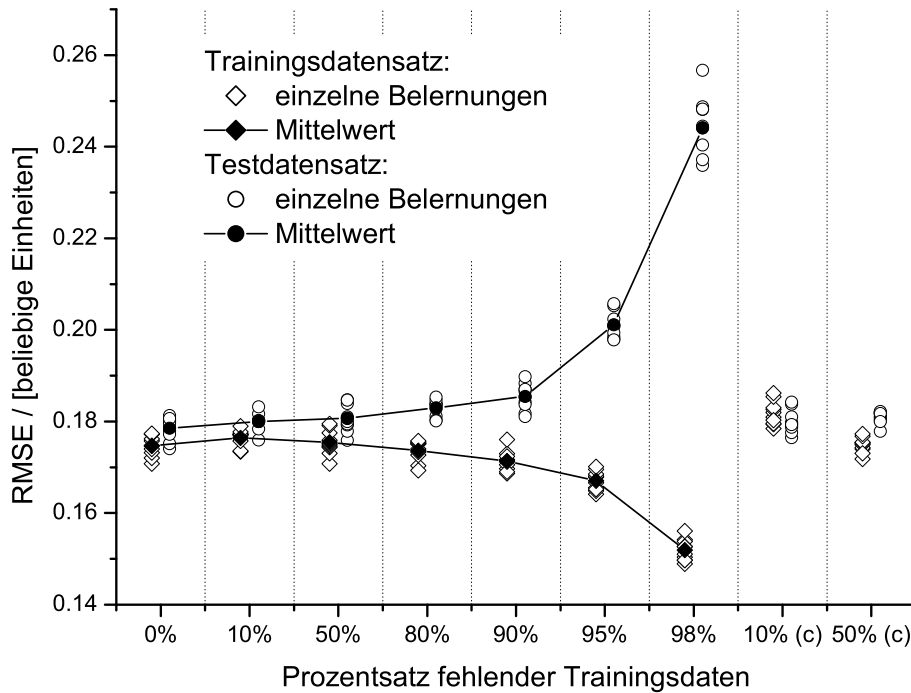


Abbildung 7.5: Minimaler RMSE bei Trainingsende (Epochenzahl 3000) für verschiedene Prozentanteile fehlender Daten. In jeder Kategorie wurden zehn Netze bis Epoche 3000 belernt. Bei den mit (c) markierten Kategorien erfolgte das Wegschneiden der Daten immer entweder vom oberen oder unteren Ende des Ozonprofils aus (siehe Text).

glatte Form der Atmosphärenprofile in den zugrundeliegenden Klimatologien. Letzeres bedingt eine starke Korrelation der 34 verwendeten Ozonkonzentrations-Schichten, die das neuronale Netz implizit zur Rekonstruktion fehlender Werte verwenden kann.

- Der Trainings-RMSE hat bis etwa 90% fehlender Daten nur eine minimale Tendenz zur Abnahme. Zwischen 90 und 98% beginnt der Fehler schneller zu fallen. Dies kann auf die verringerte Variabilität des Datensatzes zurückgeführt werden, nachdem die meisten seiner Werte entfernt wurden.
- Im Gegensatz dazu nimmt der Test-RMSE bis zur 90%-Schwelle leicht zu und zeigt darüber ebenfalls eine Verstärkung dieser Tendenz. Die Generalisierungsfähigkeit des Netzes wird mit abnehmender Varianz der Trainingsdaten offenbar schlechter.
- Die beiden Belernungskategorien mit realistischerer Auswahl fehlender Daten verhalten sich nicht grundlegend anders als die Kategorien mit rein zufälliger Auswahl.

7.2.3 Interpretation

Gemäß den obigen Ergebnissen bildet für den verwendeten Trainingsdatensatz offenbar ein Bruchteil von 90% fehlender Daten die kritische Schwelle für beschleunigte Degradation der Generalisierungsleistung. Es liegt nahe, dass dies in direkter Beziehung zum Informationsgehalt der Zielwerte steht. Zur Überprüfung wurde eine **Eigenvektorzerlegung** (*principal component analysis*, PCA) der $P=1925$ Profile des Testdatensatzes berechnet (Abb. 7.6). Ebenfalls dargestellt ist der kumulative Bruchteil η der Varianz, welcher durch die Verwendung von M Eigenvektoren beschrieben wird [Wilks, 1995, Gl. 9.40],

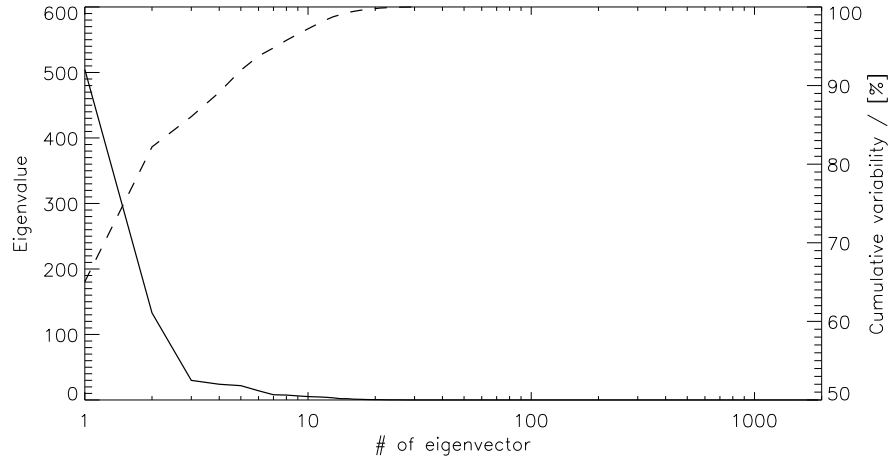


Abbildung 7.6: *Eigenwerte (durchgezogen, linke Ordinate) für die Eigenvektorzerlegung des aus 1925 modifizierten Klimatologie-Ozonprofilen bestehenden Testdatensatzes, und Informationsgehalt η gemäß Gl. 7.2 (gestrichelt).*

$$\eta_M = \frac{\sum_{m=1}^M \lambda_m}{\sum_{m=1}^P \lambda_m}, \quad (7.2)$$

wobei λ_m der m -te Eigenvektor ist. Auf der anderen Seite können die N Ausgabeneuronen für die N Höhengschichten des Profils nicht mehr als N unabhängige Informationen enthalten. Bei informationstheoretisch optimaler Zerlegung der Profile ist somit zu erwarten, dass die ersten N Eigenvektoren ausreichen, um die gesamte Profildynamik des Datensatzes zu beschreiben, d. h.

$$1 - \eta_{M \geq N} \simeq 0. \quad (7.3)$$

Tatsächlich gilt für den verwendeten Datensatz (Abb. 7.6) mit $N=34$, dass $(1 - \eta_{34}) \simeq 5.7 \cdot 10^{-6}$, das Kriterium ist also angenähert erfüllt. Anstelle der N Ozonwerte für die einzelnen Höhengschichten läßt sich also jedes Profil auch – fast exakt – als Linearkombination der N ersten Eigenvektoren des Datensatzes ausdrücken.

Sei nun k der Bruchteil fehlender Werte im Trainingsdatensatz, dann sollten von den N unabhängigen Informationen pro Profil im Mittel

$$\kappa = N \cdot (1 - k) \quad (7.4)$$

verbleiben. Der steile Abfall der Generalisierungsleistung beginnt wie oben festgestellt zwischen $k=0.9$ und $k=0.95$, was einem κ von 3.4 bzw. 1.7 entspricht. Da andererseits das Kriterium 7.2 näherungsweise erfüllt ist, entspricht dies auch der Anzahl von Eigenvektoren, die zur Beschreibung der Varianz des unvollständigen Datensatzes benötigt werden. Aus Abb. 7.6 ist zu ersehen, dass die Beträge der Eigenwerte bereits ab dem 3. stark abnehmen. Die ersten zwei bis drei Eigenvektoren beschreiben demnach einen großen Teil der Varianz. Dies fällt mit der beobachteten Schwelle für κ zusammen. Mit anderen Worten, sobald so viele Daten fehlen, dass die theoretisch verbleibende Anzahl unabhängiger Eigenvektoren pro Profil unter die zur Beschreibung der Varianz des kompletten Datensatzes benötigte Anzahl fällt, beginnt die Generalisierungsleistung des Netzes abzunehmen.

Fazit

Die partielle Belernung mit dem RPROP-Algorithmus ermöglicht die Verwendung unvollständiger Zielwerte (Ozonprofile) im Trainingsdatensatz. Der neue Algorithmus verhält sich stabil und scheint sich von der Lage der fehlenden Werte nicht wesentlich beeinflussen zu lassen. Zudem konnte anhand eines simulierten Ozonretrievals mit redundanten Trainingsdaten gezeigt werden, dass die Generalisierungsleistung der erzeugten neuronalen Netze erst dann merklich abnimmt, wenn die in den unvollständigen Trainingsdaten verbleibende Information auch theoretisch nicht mehr ausreicht, um die Varianz der Testdaten zu reproduzieren. Eine partielle Belernung scheint somit den Informationsgehalt unvollständiger Trainingsdaten optimal auszuschöpfen.

7.3 Auswahl der Eingangsdaten

Für die Bestimmung von Ozonprofilen muss die Eingangsdatenauswahl im Vergleich zum Gesamt Ozonnetzwerk etwas modifiziert werden (Tabelle 7.2), da die Redundanz nun viel geringer ist. Die Auswahl geeigneter Wellenlängenbereiche basiert auf den in Abschn. 3.3 besprochenen Gewichtsfunktionen. Im Prinzip wäre es möglich, die gesamte vom GOME-Spektrum abgedeckte Hartley- und Hugginsbande zu verwenden (vgl. Abb. 2.2), aber die Anzahl der Eingangsparameter muss wie bereits erwähnt der Anzahl der Trainingskollationen Rechnung tragen, um Übertrainieren zu vermeiden und die Belernungsdauer im Rahmen zu halten.

Zunächst wurde der Wellenlängenbereich bis auf 270 nm als Untergrenze erweitert, um genug spektrale Information aus größeren Höhen anzubieten. Bei der Entwicklung des FURM-Algorithmus hatte sich gezeigt, dass GOME unterhalb von 290 nm bislang unverstandene instrumentale Effekte aufweist, die die klassische Auswertung erschweren, da sie nicht auf einfache Weise entfernt werden können, wie z. B. durch Korrekturpolynome [Bramstedt, 2001]. Ob NNORSY in der Lage ist, aus diesem kurzwelligen Teil des GOME-Spektrums Information zu gewinnen, wird im III. Teil der Arbeit näher untersucht werden. Die spektrale Auflösung wurde im Band 1A durch Koaddition auf etwa 1.5 nm reduziert, in Band 1B auf etwa 1 nm und in Band 2 auf 0.5 nm, wobei in den Huggins-Banden darauf geachtet wurde, dass die Mittelung die Flanken des Ozonspektrums nicht verschmiert. NNORSY hat sich allerdings auch bei der Ozonprofilbestimmung als sehr robust gegenüber der Spektralwertauswahl und -vorbehandlung erwiesen, wie in Abschn. 8.1 ausgeführt.

Am langwelligen Ende der Huggins-Banden zeigte sich, dass die Verwendung der Werte bis 325 nm etwas bessere Ergebnisse erzielt, als wenn der gesamte Bereich bis 340 nm genutzt wird – hier scheint bereits zuviel Redundanz vorzuliegen, oder die statistische Information über die Verteilung des Ozons in den diesem Wellenlängenbereich entsprechenden, niedrigen Höhenschichten ist verlässlicher als die spektrale Information, und es ergeben sich Widersprüche. Dieser Effekt wird ebenfalls in Teil III der Arbeit nochmals aufgegriffen.

Eine wichtige zusätzliche Informationsquelle für das Profilretrieval-Netzwerk stellen die UKMO-Temperaturdaten dar, da aus den GOME-Spektren selber i. W. nur ein einziger Temperaturparameter unabhängig bestimmt werden kann [V. Rozanow, *persönliche Mitteilung*, 2000]. Die Verwendung von Temperaturprofilen ist jedoch ein zweiseitiges Schwert: Zwar lassen sich die Retrievalergebnisse hierdurch beträchtlich verbessern, aber gleichzeitig wird das Verfahren von einer externen Datenquelle abhängig, die möglicherweise nicht in Echtzeit zur Verfügung steht. Wenn die Temperaturdaten „zu genau“ sind, koppelt der Belernungsalgorithmus das Ozonprofil so stark damit, dass die GOME-Spektralwerte mehr oder weniger ignoriert werden und das Retrievalergebnis fast nur noch von der Genauigkeit der Temperaturinformation abhängt. Dies konnte in Vorstudien nachgewiesen werden, bei denen die von den Ozonsonden ebenfalls gemessenen Temperaturprofile in hoher Auflösung

Tabelle 7.2: *Eingangsparameter neuronaler Netze für das Retrieval von Ozonprofilen. Zur näheren Erläuterung der Größen siehe Tabelle 6.1.*

Eingangsparameter	Coadding	Anzahl Werte	Zweck
270–325 nm	4–12	74	O ₃ Hartley/Huggins Banden
380–385 nm	4	13	Atmosphärisches Fenster / NO ₂
598–603 nm	4	6	O ₃ Chappuis-Bande / NO ₂
758–772 nm	8	9	O ₂ (A) Bande – Wolkendetektion
SZA, SAA	-	2	Lichtwegskorrektur
ScanZA, ScanAA	-	2	Lichtwegskorrektur
line-of-sight Flags	-	3	Lichtwegskorrektur
Latitude & Saison	-	2	Klimatologischer Background
Sensorenalter	-	1	Degradationskorr., Ozonfelddrift
UKMO T-Profil	4 [†]	10	Meteorologischer Background

[†]4 km Höhenschichten

als Eingangsdaten verwendet wurden.¹ Die optimalen Ergebnisse ließen sich mit der in Tabelle 7.2 angegebenen Konfiguration erzielen, bei der die UKMO-Daten als Mittelwerte von 4 km dicken Schichten auf 4–40 km Höhe in die Belernung eingingen. Die Verwendung von 1 km dicken Schichten erweist sich demgegenüber nicht als vorteilhaft, weil sie in keinem Verhältnis zur relativ schlechten horizontalen und zeitlichen Auflösung der UKMO-Daten steht (vgl. Abschn. 5.3).

Die restlichen Eingangswerte entsprechen denen des Gesamtzonnetzwerks und wurden in Abschn. 6.3 bereits erläutert. Die spektrale Auflösung ist für das Profilretrieval jedoch etwas höher. Insgesamt ergeben sich so 122 Eingangsneuronen. Für die Zwischenschicht erwiesen sich bei der gegebenen Trainingsdatensatzgröße 45 Neuronen als geeigneter Wert. Die Ausgabe erfolgte auf 60 Höhenschichten, zentriert um die geopotenziellen Höhen 1–60 km. Die Belernung des Netzwerkes für Kollokationen von 1996 bis Mitte 2001 war nach 16000 Epochen abgeschlossen.

7.4 Ergebnisse der Ozonprofilbestimmung

Die grundlegende Validation der Belernungsergebnisse erfolgt auf der Basis von Trainings- und Testdatensatz-Statistiken. Hierzu werden Bias μ_R und Standardabweichung σ_R der Retrievalprofile relativ zu den kollokierten Ozonprofilen bestimmt. Um den Informationsgewinn der Methode besser einschätzen zu können, wird jedem Retrieval außerdem ein nach Latitude und Monat ausgewähltes Ozonprofil der Klimatologie von *Fortuin & Kelder* [1998] entgegengestellt und dafür die entsprechenden Statistiken μ_C bzw. σ_C berechnet. Das Verhältnis σ_R/σ_C gibt dann Aufschluss über die Verbesserung des Wissens über die atmosphärische Ozonverteilung durch Einsatz der NNORSY-Methode.

7.4.1 Fehlerstatistik für Trainings- und Testdaten

Abb. 7.7 zeigt die globale Statistik für den **Trainingsdatensatz**. Die relative Standardabweichung in Höhenschicht i ist hier definiert als

¹Nebenbei bemerkt zeigt dies einen wichtigen Aspekt der NNORSY-Methode: Man erweist dem Netz keinen Gefallen, wenn für die Belernung bessere Daten benutzt werden, als später im operationellen Betrieb zur Verfügung stehen. Die Kovarianz der Eingangsdaten – und somit deren Gewichtung – wird im Verlauf der Belernung bereits fixiert und optimal berücksichtigt. Spätere Kovarianzänderungen stören das dabei austarierte Gleichgewicht und werden so im Regelfall die Ausgabefehler vergrößern.

$$\delta_{R,i} = \frac{\sum_{p=1}^{P_i} \sqrt{\frac{1}{P_i-1} (\hat{x}_i^p - t_i^p - \langle \hat{x}_i - t_i \rangle)^2}}{\frac{1}{P_i} \sum_{p=1}^{P_i} t_i^p} = \frac{\sigma_{R,i}}{\langle t_i \rangle}, \quad (7.5)$$

also relativ zum globalen Mittelwert, nicht zum jeweiligen Einzelprofil — andernfalls ergeben sich bedingt durch die Glättungsfehler unrealistisch hohe Werte, da die Berechnung der Statistiken auf 1 km-Höhenschichten erfolgt. Auf die Glättungsfehler wird im Kapitel 10 ausführlich eingegangen. $\hat{\mathbf{x}}$ bezeichnet das Retrievalprofil, \mathbf{t} das kollokierte Zielprofil. Man beachte, dass aufgrund fehlender Zielwerte die Anzahl der Werte P_i in jeder Höhengschicht unterschiedlich sein kann: Mit dem Fehlerflag m aus Gl. 7.1 gilt nämlich $P_i = \sum_{p=1}^P m_i^p$. Die mittlere systematische Abweichung (Bias) ist definiert als

$$\mu_{R,i} = \frac{1}{P_i} \sum_{p=1}^{P_i} (\hat{x}_i^p - t_i^p) = \langle \hat{x}_i - t_i \rangle. \quad (7.6)$$

Das Belernungsverfahren konnte den Bias zwischen Netzausgabe und Trainingskollokationen praktisch auf Null reduzieren, wie dies auch von der Theorie her zu erwarten ist [Bishop, 1995a]. Die Differenz der mittleren Ozonprofile bei 9–16 km im Bild 7.7a ist dadurch zu erklären, dass die Trainingsdaten für tropische Breiten mehr Satelliten- und weniger Sondenprofile enthalten, wobei erstere meist nicht in diese Höhen hinabreichen. Die Retrievals $\hat{\mathbf{x}}$ sind hingegen überall definiert. Bei der Bildung des Mittelwerts über alle Profile findet sich daher in den Retrievalprofilen eine größere Anzahl niedriger Ozonwerte aus der tropischen oberen Troposphäre, als in den Trainingskollokationen.

Die größten relativen Fehler (Abb. 7.7b) treten ebenfalls in diesem Höhenbereich und ein wenig darunter auf. Sie sind bezeichnend für die starke Dynamik des dortigen Ozonfeldes in mittleren bis hohen Breiten, die sich durch vertikale, filamentartige Strukturen [Hood et al., 2001], Variationen der Tropopausenhöhe [Roedel, 2000] und starke horizontale Konzentrationsgradienten auszeichnet. In den Tropen bedingen die generell niedrigen Ozonwerte in dieser Höhenregion bereits bei kleinen Abweichungen große relative Fehler. Die Reduktion der Standardabweichung relativ zur Klimatologie liegt oberhalb von 9 km weiterhin bei etwa 40%, wie in den meisten anderen Höhenregionen auch, da die Dynamik durch Monats- und Breitenkreismittelwerte ebenfalls im Vergleich mit anderen Höhen schlechter beschrieben werden kann. Allerdings ist NNORSY unterhalb von 9 km Höhe nur etwa 20–30% besser als die Klimatologie. Hier wird die Gewinnung spektraler Information schwierig (vgl. Abschn. 8.1.1), und die relativ schlechte räumlich-zeitliche Auflösung der UKMO-Temperaturprofile kann den Änderungen des Ozonfeldes nur noch unzureichend folgen.

Beim Übergang zum **Testdatensatz** werden die Fehler generell etwas größer (Abb. 7.8). Das neuronale Netz wird hier mit Eingangsdaten und Ozonprofilen konfrontiert, die nicht in den Trainingsdaten enthalten waren. Die resultierende Fehlerstatistik kann somit einen Eindruck von der Generalisierungsleistung des Netzes geben. Die globalen Mittelwerte und Variationsbereiche der beiden Profildatensätze stimmen allerdings gut überein, damit sind die ermittelten Relativfehler direkt vergleichbar. Die große Ähnlichkeit der Fehlerstatistiken für Trainings- und Testdatensatz zeigt deutlich, dass das neuronale Netz mit der Generalisierung in dem Bereichen, die von beiden Datensätzen abgedeckt werden, keine Schwierigkeiten hat.

Oberhalb von etwa 25 km Höhe liegen die relativen Fehler auch bei den Testkollokationen im Bereich zwischen 5–10%, und die Reduktion der Standardabweichung gegenüber der Klimatologie schwankt um die 40%. Darunter steigt der relative Fehler an, während

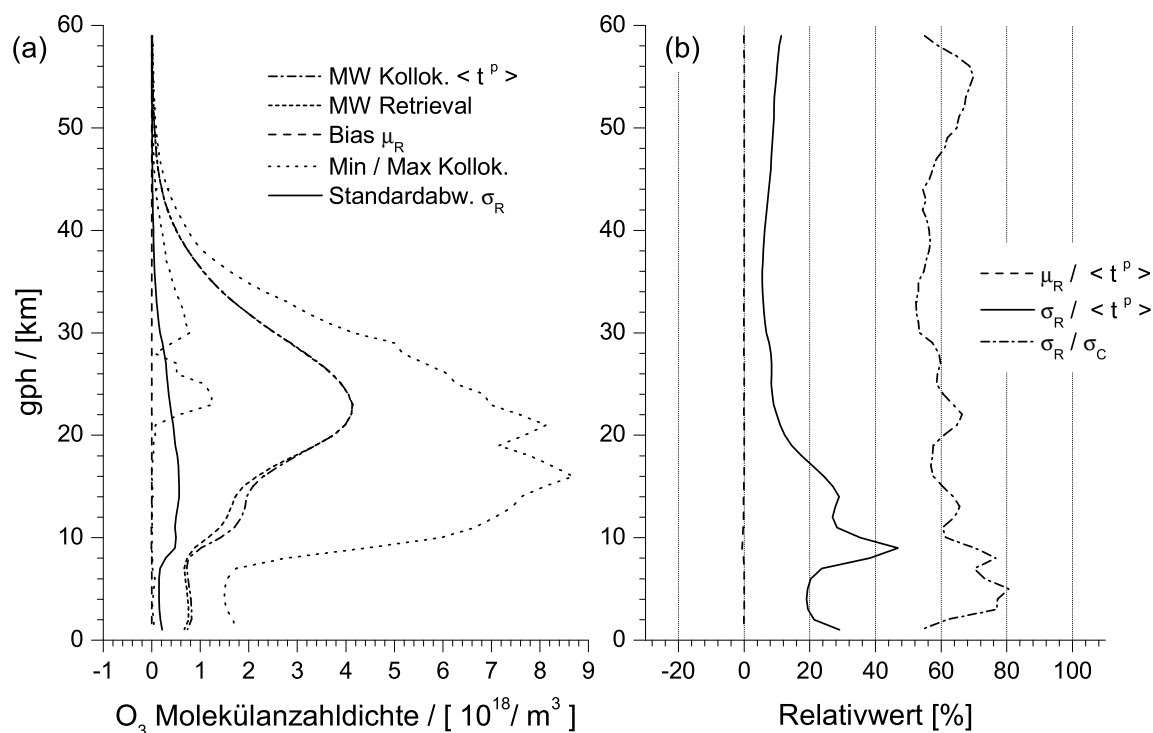


Abbildung 7.7: Globale Statistik für den Trainingsdatensatz nach vollständiger Belernung des neuronalen Netzes. (a) Absolutwerte. Minimum- und Maximumwerte beziehen sich auf die Kollokationen. (b) Werte relativ zum Mittelwert ($\langle t^p \rangle$) des gesamten Datensatzes, sowie Standardabweichung σ_R relativ zur analog berechneten Standardabweichung der F&K-Klimatologie, σ_C .

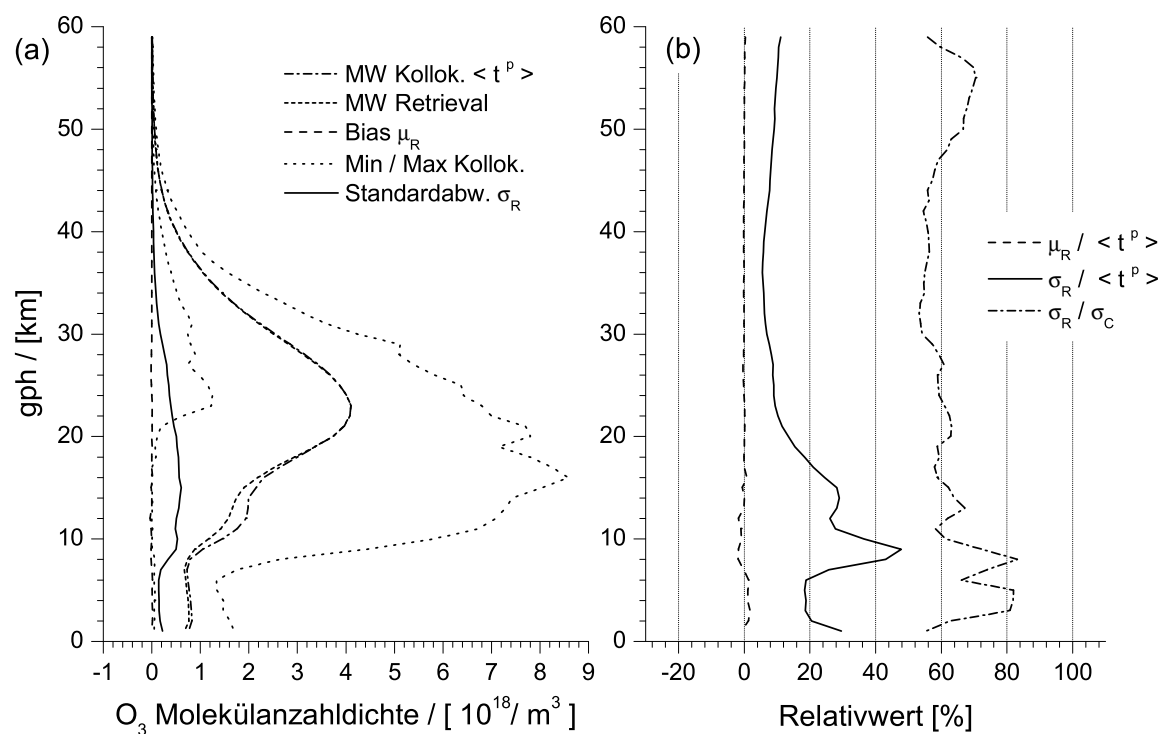


Abbildung 7.8: Wie Abb. 7.7, aber für den statistisch unabhängigen Testdatensatz.

σ_R/σ_C zunächst konstant bleibt, d. h. der Fehleranstieg spiegelt die zunehmende Variabilität des Ozonprofils wider, die in gleichem Maße auch die Klimatologiefehler steigen läßt. Erst unterhalb von 10 km Höhe büßt NNORSY etwas von seiner Überlegenheit gegenüber der Klimatologie ein. Dies ist als Zeichen dafür zu interpretieren, dass das S/N-Verhältnis für die Ozonbestimmung aus den GOME-Spektren schlechter wird, denn das neuronale Netz verläßt sich hier offenbar mehr auf statistische Information, die ja auch der Klimatologie zugrunde liegt und somit gegenüber dieser weniger Vorteile bringt. Hier zeigt sich bereits das Wechselspiel zwischen physikalischer und statistischer Information, welches sehr charakteristisch für das vorgestellte Verfahren ist und in Teil III der Arbeit ausführlich untersucht wird. Zunächst soll jedoch die Fehlerstatistik genauer durchleuchtet werden.

7.4.2 Sensitivitäten der Fehler

Da bei der Qualitätskontrolle einer Netzwerkbelerung die Generalisierungsleistung wichtiger ist als das Reproduzieren der Trainingskollokationen, werden sich die folgenden Ausführungen auf die Diskussion des Testdatensatzes beschränken.

Geografische Breite

Die globale Verteilung des Ozons ist in erster Linie breitenabhängig, daher liegt es nahe, die Testkollokationen in Latitudenbänder einzuordnen und die entsprechenden Fehler auf diesen Bändern zu bestimmen (Abb. 7.9). Die Werte schwanken auf diesen Bändern teilweise beträchtlich.

In mittleren bis hohen Breiten kann das System oberhalb des Ozonpeaks sehr gut generalisieren, die Fehler liegen in ähnlichen Bereichen wie die des Trainingsdatensatzes. In südpolaren Breiten ist die Standardabweichung in mittleren Höhen teilweise um bis zu 60% geringer als die der Klimatologie, was auf quantitativ gutes Erfassen von Ozonloch-situationen schließen läßt, die in der Klimatologie nicht berücksichtigt werden können. Im nordpolaren Raum ist die Standardabweichungs-Reduktion ebenfalls recht hoch, allerdings besteht hier das Problem, dass polwärts von 75°N keine Okkultations-Profile mehr zur Verfügung stehen (s. Tab. 2.2), so dass über die Genauigkeit in Polnähe keine direkte Aussage möglich ist.

Am problematischsten sind offenbar die Tropen. Hier bilden sich troposphärische Biases von bis zu 20% aus, und die Verbesserung gegenüber der Klimatologie ist in größeren Höhen relativ gering. Mehrere Faktoren spielen dafür eine Rolle:

1. Die Ozonverteilung in größeren Höhen ist in den Tropen hauptsächlich einstrahlungsbedingt, und variiert daher von Tag zu Tag sehr wenig, d. h. die Monatsklimatologie kann das Ozonfeld bereits gut repräsentieren.
2. Aufgrund des Auftretens vieler hoher und dichter Wolken ist es schwierig, physikalische Information über die generell niedrigen troposphärischen Ozonwerte in den Spektraldaten auszumachen.
3. Die Verteilung der Ozonwerte im Höhenbereich um 9–13 km sieht in den Tropen ganz anders aus als außerhalb davon (Abb. 7.10), da der Bereich in den Tropen immer in der Troposphäre liegt (niedrige Ozonwerte), jenseits des Tropopausenbruches jedoch als Übergangsregion zwischen Tropo- und Stratosphäre eine breite Streuung von Ozonwerten aufweist. Im Prinzip ist dies kein Problem, da das MLP jede Verteilung modellieren kann. Es muss aber aufgrund seiner quadratischen Fehlerfunktion einen globalen Kompromiss für die Standardabweichung suchen (s. Abschn. 4.3). Da der (eingangsdatenabhängige) Zielwert in der betrachteten Region nahe bei Null liegt,

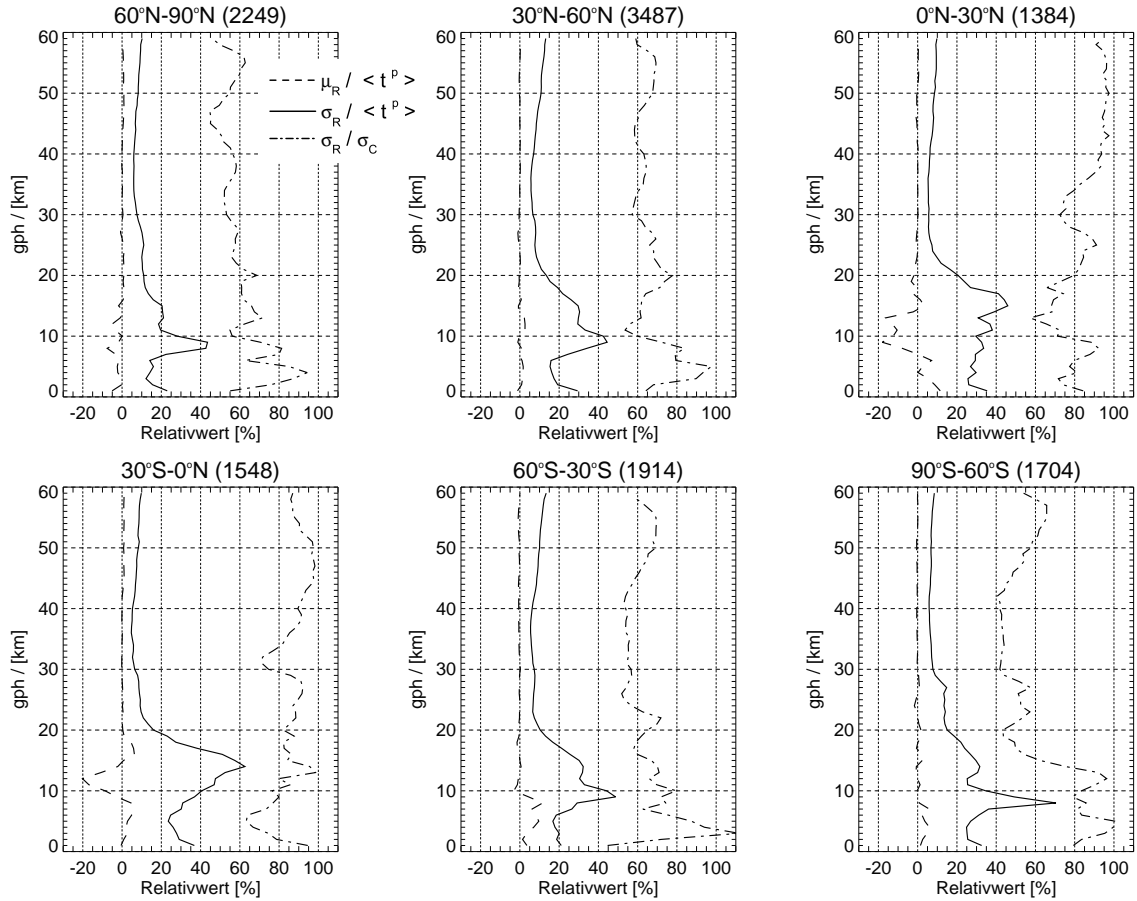


Abbildung 7.9: Wie Abb. 7.7b, aber Testkollokationen nach Latitudenband geordnet. Die Werte in Klammern geben die Anzahl der Kollokationen an.

und daher nach unten hin begrenzt, die (globale) Standardabweichung aber relativ groß ist, tendiert das MLP hier zu etwas niedrigen Werten. In anderen Höhenregionen gelingt das Finden eines guten Kompromisses offenbar sehr gut.

4. Da aufgrund des größeren horizontalen Abstandes der GOME-Orbits in niedrigen Breiten nicht so viele Kollokationen zustande kommen, ist trotz der Verwendung von SHADOZ-Daten die dortige Troposphäre relativ schwach im Trainingsdatensatz vertreten. Der Belernungsalgorithmus optimiert daher diesen Bereich nicht so gut, was den im vorherigen Punkt beschriebenen Effekt verstärkt.

Angesichts der generell niedrigen troposphärischen Ozonwerte in den Tropen entsprechen die Biases jedoch nur einer absoluten Änderung von $\simeq 0.1 \cdot 10^{18} \text{ m}^{-3}$. Da die durchschnittliche außertropische Ozondichte auf 12 km Höhe bei $\simeq 2 \cdot 10^{18} \text{ m}^{-3}$ liegt, und wie erwähnt die Anzahl der tropischen Kollokationen relativ gering ist, verschwindet der negative Bias bei globaler Betrachtung, ohne einen merklichen, positiven Bias in anderen Breitenregionen zu hinterlassen. Würde das NNORSY-Höhengitter z. B. relativ zur Tropopausenhöhe definiert, könnte möglicherweise ein großer Teil der Biases eliminiert werden. Dies wurde jedoch noch nicht näher untersucht, da die hierzu notwendigen, hochaufgelösten Tropopausenhöhen bisher nicht zur Verfügung standen.

Die Schwierigkeiten bei der Bestimmung von Ozonwerten unterhalb des Ozonpeaks zeigen sich in allen Latitudenbändern, wobei die Höhenvariation mit der geografischen Breite

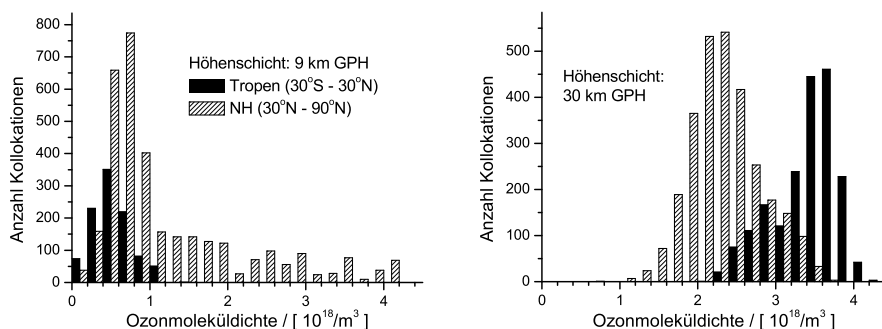


Abbildung 7.10: Zieltatenverteilung für die 9 km und die 30 km Höhengschicht (Intervall $0.2 \cdot 10^{18} \text{ m}^3$). Um 9 km herum variieren Standardabweichung und Mittelwertlage in und außerhalb der Tropen sehr stark, bei 30 km (und in den meisten anderen Schichten) hingegen nur mäßig.

gut beobachtet werden kann: Die größten relativen Fehler treten in der Nähe der Tropopause auf, in den Tropen um 14 bis 17 km Höhe, polwärts des Tropopausenbruchs bei 8 bis 10 km Höhe. In den südlichsten beiden Bändern ergibt sich außerdem in der Troposphäre eine minimale Verschlechterung gegenüber der Klimatologie. Da hier allerdings nur relativ wenige Bodenstationsmessungen vorhanden sind, wird dieser Effekt im Rahmen der Fallstudien in den Kapiteln 11 und 12 nochmals näher betrachtet.

Sensorentyp

Um zu überprüfen, wie sich die verschiedenen für die Kollokation verwendeten Sensorentypen verhalten, wurden in Abb. 7.11 die Fehlerstatistiken getrennt für Ozonsonden, SAGE, HALOE und POAM-Profile dargestellt. Die Sonden bieten das gewohnte Bild mit relativ großen Fehlern im Tropopausenbereich und gleichmäßigem σ_R/σ_C . SAGE II und HALOE verhalten sich in größeren Höhen fast gleich, lediglich am unterem Geltungsbereich der Profile liegen die Okkultationswerte im Mittel etwas niedriger als das Retrieval, und der Dynamikbereich von HALOE nimmt unrealistisch stark ab. Aufgrund der großen Datenmenge ist es wie bereits erwähnt nicht möglich, sämtliche Profilfehler auszuschließen, andererseits ist die Anzahl der durch Aerosole und Wolken kontaminierten Zielwerte bereits so gering, dass der Belernungsalgorithmus darüber hinweg mitteln kann, ohne global nennenswerte Biases zu generieren. Die Vergleiche mit POAM III Daten weisen in der Höhe eine Reduzierung der klimatologischen Standardabweichung von über teilweise 60% auf, was einerseits auf eine sehr gute Datenqualität hinweist, andererseits aber zum Teil durch die breitengradmäßig eingeschränkte Verwendung der Daten bedingt ist. Insbesondere sind die tropischen Breiten ausgespart, für die weiter oben die schlechtesten σ_R/σ_C -Werte gefunden wurden.

Pixel line-of-sight

Im Gegensatz zu den meisten anderen GOME Ozonprofil-Retrievalverfahren verwendet NNORSY die volle horizontale Auflösung des Instruments, ohne das S/N-Verhältnis einzelner Spektren durch Koaddition zu verbessern, und ohne eine explizite Scanwinkel-Korrektur vorzunehmen. Eine solche Korrektur wird dem neuronalen Netz ermöglicht, indem sowohl der Scanwinkel als auch drei LOS-Flags als Eingangsdaten verwendet werden (Abschn. 5.3). Daher ist es wichtig zu überprüfen, inwieweit die statistischen Fehler der drei verwendeten Pixeltypen Ost, Nadir und West differieren.

Anhand von Abb. 7.12 ist zu erkennen, dass die Variation der relativen Fehler nur sehr gering ist. Oberhalb von 20 km stimmen sowohl Bias als auch relative Standardabweichung

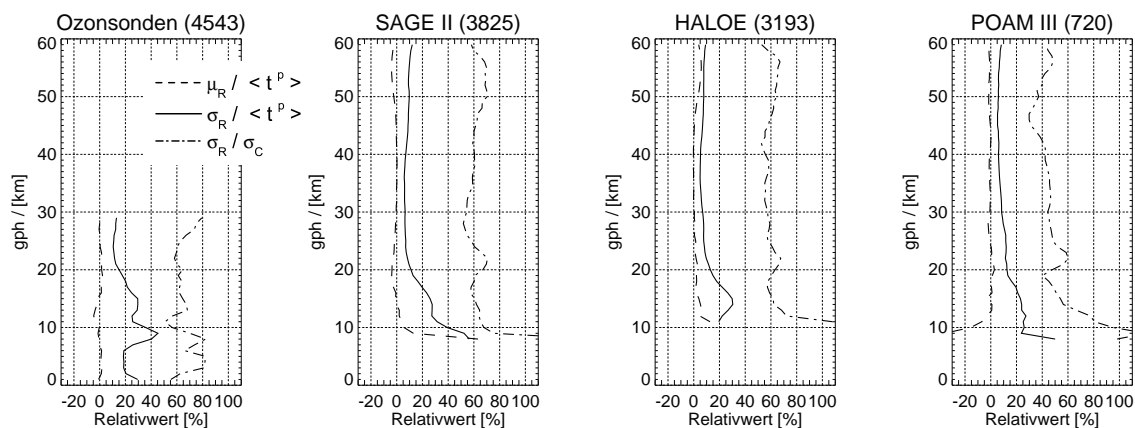


Abbildung 7.11: Wie Abb. 7.7b, aber Testkollokationen nach Sensorentyp geordnet. Die Werte in Klammern geben die Anzahl der Kollokationen an.

sehr gut überein, darunter nehmen die Schwankungen leicht zu, bleiben aber generell im Bereich von $\pm 5\%$. Die Verbesserung gegenüber der Klimatologie ist für die Nadir-Pixel in fast allen Höhenbereichen deutlicher als für die anderen beiden Typen. Allerdings liegen diese Änderungen teilweise im Bereich der Streuung verschiedener Netzwerkbelernungen (vgl. Abb. 8.1), und sind demnach nicht unbedingt statistisch signifikant. Die Scanwinkel-Korrektur, die das neuronale Netz aus den Trainingsdaten erlernt, scheint also alles in allem sehr gut zu arbeiten. Die Daten der Ost- und West-Pixel liefern nur wenig schlechtere Profile als die Nadir-Pixel.

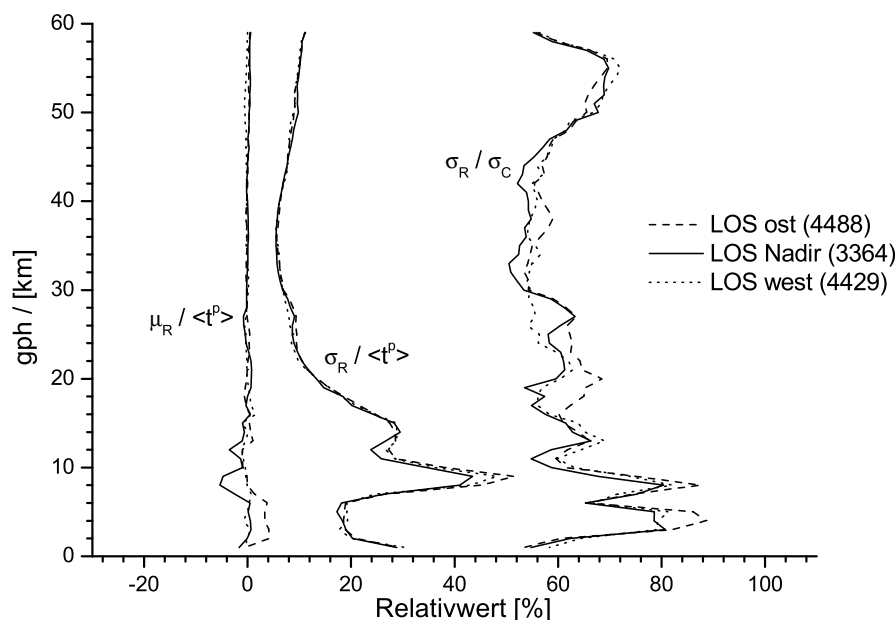


Abbildung 7.12: Wie Abb. 7.7b, aber Testkollokationen nach Pixel LOS geordnet. Die Werte in Klammern geben die Anzahl der Kollokationen an.

7.4.3 Abschätzung der Vertikalauflösung

Nachdem das Verhalten der NNORSY-Retrievals auf Basis des Testdatensatzes ausführlich charakterisiert wurde, bleibt noch zu klären, wie hoch deren Auflösung in vertikaler Richtung ist. Bei klassischen OE-Retrievals wird zu diesem Zweck die in Abschn. 3.4.1 eingeführte *averaging kernel* Matrix

$$\mathbf{A} = \mathbf{D}_y \mathbf{K} \quad (7.7)$$

analysiert, deren Zeilenvektoren (AKs) als Grundlage für die Abschätzung der Vertikalaufklärung dienen können. Die Beitragsfunktionsmatrix $\mathbf{D}_y = \frac{\partial \mathbf{x}}{\partial \mathbf{y}}$ läßt sich bei einem inversen, neuronalen RTM als neuronale Jacobi-Matrix direkt durch Differenzieren der vom neuronalen Netz vermittelten Abbildung berechnen (Abschn. 9.4). \mathbf{K} müsste dann gemäß Gl. 3.5 aus einem separaten Vorwärtsmodell gewonnen werden, da beim NNORSY-Retrieval keine explizite Vorwärtsrechnung durchgeführt wird.

Jiménez [2000] wandte dieses Verfahren auf simulierte Mikrowellen-Limb-Messungen an und fand ähnliche AKs wie bei OE, wenn auch etwas verrauschter. Der Autor weist jedoch darauf hin, dass dieses Retrievalproblem recht linear ist. Bei NNORSY erweisen sich jedoch mehrere Faktoren im Hinblick auf die AK-Berechnung als sehr ungünstig: Das System ist hochgradig nichtlinear und liefert fast defekte Jacobi-Matrizen, die in Kap. 9 diskutiert werden. Weiterhin mischt NNORSY zur Erkennung von sehr kleinen vertikalen Strukturen statistische und physikalische Information (Kap. 8), indem manche Vorwärtsmodell-Parameter (z. B. SZA, Scanwinkel) und klimatologische Koordinaten (Latitude, Zeit) direkt als Prädiktoren in das Retrieval eingehen. Wird die Vertikalaufklärung hingegen nur über die – physikalisch berechnete – AK-Matrix abgeschätzt, so ergibt sich ein verzerrtes Bild der in der Praxis beobachteten Auflösung.

Wodurch kann die Vertikalaufklärung überhaupt definiert werden? Es ist nicht immer möglich, ein optimales Kriterium dafür anzugeben, siehe dazu die ausführliche Diskussion in *Rodgers* [2000]. Bei FURM wird die volle Halbwertsbreite (FWHM) der AKs als Maß verwendet, was zwar sehr anschaulich ist, jedoch interpretationsmäßige Schwierigkeiten verursacht, wenn die AKs stark ausgeprägte negative Beiträge besitzen oder stark asymmetrisch sind. Generell gilt aber, dass für technische Validationen eines Retrievalalgorithmus’ – im Gegensatz zu geophysikalischen Validationen – oder für Vergleiche zwischen verschiedenen Algorithmen die kollokierten Ozonprofile i. A. mit den AKs gefaltet werden, um die Glättungsfehler zu minimieren. Dies ist wichtig, da die Glättungsfehler einen beträchtlichen Beitrag zum Gesamtfehler des Retrievals stellen können (Abschn. 10.3.3). Beim visuellen Vergleich von NNORSY-Retrievals und Ozonsonden-Kollokationen auf dem 1 km Höhenraster wird dies sofort klar, da sich die NNORSY-Profile bei feineren Sondenprofilstrukturen wie ein gleitender Mittelwert verhalten, und demnach die Sondenprofile lokal relativ stark vom Retrieval abweichen können. Daher sollten sich z. B. die in den Abbildungen 7.7 bis 7.12 dargestellten Fehler verringern lassen, indem die Sondenprofile vor dem Vergleich mit NNORSY geglättet werden. Natürlich gilt dies nur bis zu einem gewissen Punkt: Wenn die Glättung zu stark ist, mittelt sie über die noch vom Retrieval erfassten vertikalen Strukturen hinweg, und der Glättungsfehler steigt wieder an. Der Glättung, bei der die beobachteten Fehler am geringsten sind, entspricht dann die Vertikalaufklärung des Retrievalverfahrens.

Die Glättung der kollokierten Profile kann durch die Multiplikation mit einer Matrix \mathbf{A}_δ geschehen, deren Zeilenvektoren identische normierte Gaußfunktionen der Halbwertsbreite δ enthalten. Diese Funktionen sind jeweils um eine andere Höhengschicht i zentriert, so dass die Matrix in etwa der AK-Matrix des OE-Verfahrens entspricht. Anders ausgedrückt, entspricht die Multiplikation eines Profils mit \mathbf{A}_δ der Faltung mit einer Gaußfunktion der

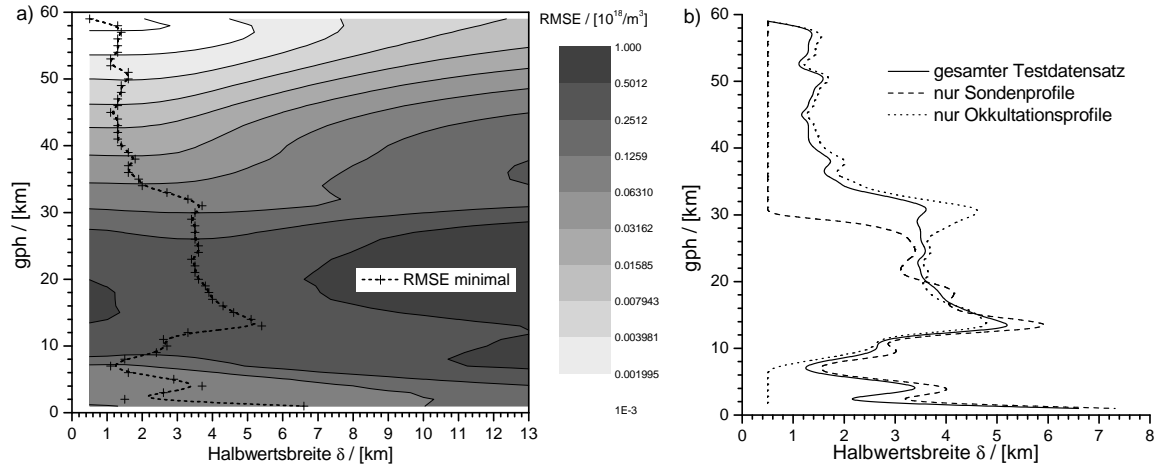


Abbildung 7.13: (a) Variation des RMSE auf dem Testdatensatz bei Faltung der Kollokationsprofile mit Gaußfunktionen unterschiedlicher FWHM. (b) FWHM, bei der ein Minimum des RMSE für die entsprechende Höhe auftritt, für verschiedene Teildatensätze.

FWHM δ . Für einen Datensatz von P Kollokationen errechnet sich der RMSE der Höhengschicht i bei Glättungs-FWHM δ zu

$$\text{RMSE}_i(\delta) = \sqrt{\frac{1}{P} \sum_{p=1}^P (\hat{\mathbf{x}}_i^p - (\mathbf{A}_\delta \mathbf{t}^p)_i)^2}. \quad (7.8)$$

Abb. 7.13a zeigt die Variation des RMSE mit i und δ für den gesamten Testdatensatz. Fehlende Profilstücke der Okkultations- und Sondenprofile \mathbf{t} wurden vor der Multiplikation mit \mathbf{A}_δ durch NNORSY-Retrievals ergänzt, da sonst zu viele unkontrollierbare Randeffekte auftreten. Daher sind die Absolutwerte der Fehler nicht interpretierbar, worauf gleich noch näher eingegangen wird. Der RMSE wurde anstatt der Standardabweichung gewählt, weil dadurch die bei der Verschmierung von steilen Ozonprofil-Flanken entstehenden Biases berücksichtigt werden. Allerdings sind die Ergebnisse unter Verwendung der Standardabweichung fast identisch.

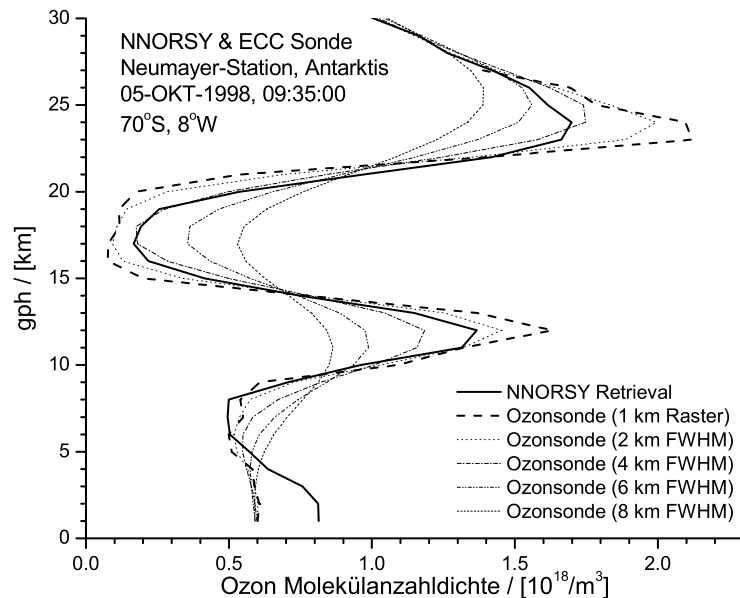
Die Konturen des RMSE variieren kontinuierlich mit der FWHM der Glättungsfunktion, und die RMSE-Minima bilden einen deutlichen Sattel aus, der im Höhenbereich 11–32 km zwischen 3 und 5 km FWHM liegt, sonst meist darunter. Dieselbe Statistik auf dem Trainingsdatensatz angewandt liefert ein fast identisches Ergebnis (nicht dargestellt). Es ist anschaulich klar, dass diese heuristische Methode nur in den Bereichen gut arbeitet, in denen das Ozonprofil erkennbare vertikale Strukturen zeigt. Da die oberhalb von 30 km ausschließlich verwendeten Okkultationsmessungen bereits sehr glatt erscheinen, bewirkt eine zusätzliche Glättung hier fast nur die Zunahme der Biases, und der Glättungsfehler zeigt kein zuverlässiges Minimum. Der starke Rückgang der Minimumskurve unterhalb von 10 km entsteht durch die Abflachung der steilen Flanke des Ozonprofils beim Übergang von der Troposphäre zur Stratosphäre (vgl. mittleres Ozonprofil in Abb. 7.8a): Schon bei relativ geringer Glättungs-FWHM wird viel Ozon aus der Flanke nach unten in die Troposphäre verschoben, und führt dort zu großen Fehlern. Die vertikale Auflösung in dieser Höhenregion abzuschätzen, ist auch bei klassischen Verfahren wie OE problematisch, da die zugehörigen AKs asymmetrisch sind, mehrere Peaks aufweisen können und ihr Hauptmaximum mitunter stark von seiner nominellen Höhe abweicht [de Beek, 1998; Hoogen et al., 1999]. Die Näherung der AKs mit Gaußfunktionen entspricht daher in dieser Region den

physikalischen Verhältnissen vermutlich nicht besonders gut. In der unteren Troposphäre machen zusätzlich Randeﬀekte die Interpretation der RMSE-Minimumskurve unsicher.

Es ist offensichtlich, dass die Minimumskurve bis zu einem gewissen Grad vom mittleren Ozonprofil abhängt, insbesondere von dessen Flankensteilheit. Bei einem rein physikalischen Verfahren müsste man diese Abhängigkeit größtenteils als Artefakt der beschriebenen Methode zur Vertikalaufhebungsbestimmung werten. Bei NNORSY hingegen resultiert die Vertikalaufhebung tatsächlich zum Teil aus der im Trainingsdatensatz repräsentierten Profillform: Eine steile Flanke in einem Trainings-Ozonprofil führt naturgemäß zu großen Fehlern, wenn das Retrieval ihre Altitude und Steilheit nicht richtig erfasst. Große Fehler weisen aber den Belernungsalgorithmus an, die entsprechende Stelle bevorzugt zu optimieren, so dass am Ende die Vertikalaufhebung des Verfahrens sicher in einer – wenn auch nicht ohne weiteres quantifizierbaren – Weise mit der Trainingsprofillform zusammenhängt.

Weiterhin ist zu untersuchen, wie sich fehlende Profildaten auf die Berechnung der RMSE-Minima auswirken. Zu diesem Zweck ist in Abb. 7.13b die Minimumskurve nochmals aufgetragen, zusammen mit analogen Kurven für reine Okkultations- und Sondenkollokationen. Die Übereinstimmung dieser beiden Kurven ist in Höhen zwischen 14 km und 25 km recht gut; Hier liegen in beiden Datensätzen nur relativ wenige Datenlücken vor, so dass beide Kurven als zuverlässig gelten können. Da fehlende Werte durch NNORSY-Retrievals ergänzt werden und der Vergleich eines Profils mit einer geglätteten Version seiner selbst ($\delta > 0$) immer Fehler produziert, ist sichergestellt, dass die Minima-Kurven in Bereichen mit vielen fehlenden Daten zu $\delta = 0$ hin tendieren. Wird nicht oder mit einer Klimatologie ergänzt, ist das Verhalten der Minima hingegen unberechenbar. Hieraus folgt, dass in jeder Höhe der größte in Abb. 7.13b dargestellte FWHM-Wert die zuverlässigste Schätzung für die Vertikalaufhebung angibt.

Abbildung 7.14: NNORSY Ozonprofil im Vergleich zu einem geglätteten Sondenprofil des Testdatensatzes. Mit FWHM ist die Halbwertsbreite δ der Gaußfunktion gemeint, mit der das Sondenprofil zur Glättung gefaltet wurde.



Es mag zunächst verwundern, dass damit die Vertikalaufhebung trotzdem noch deutlich unter den strahlungstransfertheoretischen Grenzen von 6–8 km liegt [Hoogen, 1998]. Wie Abb. 7.14 zeigt, können stratosphärische NNORSY-Retrievals jedoch in der Praxis tatsächlich mit geglätteten Sondenprofilen für $\delta=2$ km bis $\delta=4$ km verglichen werden. Weitere Beispiele und eine Gegenüberstellung mit OE-Verfahren finden sich in Abschn. 11.3. Die weiter oben demonstrierte, gute Generalisierungsfähigkeit des neuronalen Netzes sowie die im nächsten Kapitel vorgestellten Experimente mit Belernungen ohne GOME-Spektraldaten

sprechen dagegen, dass das System hier nur eine gewissen Ozonfeldstruktur „auswendig“ lernt. Wie viel Physik und wie viel Statistik den Ergebnissen hier im Einzelnen zugrunde liegt, und wie sich der scheinbare Widerspruch zur Strahlungstransfertheorie auflösen läßt, wird der anschließende, III. Teil dieser Arbeit zu erhellen versuchen.

Fazit

Mit Hilfe von Vergleichen zu vertikal hoch aufgelösten, kollokierten Ozonprofilen wurde versucht, die Vertikalaufösung des Retrievalverfahrens abzuschätzen. Im Höhenbereich 12 bis 32 km ergeben sich für eine gaußförmige Glättungsfunktion FWHM-Werte von 3 bis 6 km, darüber und darunter ist keine zuverlässige Aussage möglich. In nullter Näherung kann die Vertikalaufösung von NNORSY mit etwa 4 km angegeben werden. Die hier erhaltenen Ergebnisse werden im Zusammenhang mit der Profilfehler-Korrelationslänge in Abschn. 9.2 nochmals aufgegriffen und ergänzt.

Zusammenfassung von Teil II

Die GOME-Ozonbestimmung mit NNORSY verläuft in drei Schritten: Zunächst wird eine Kollokationsdatenbank aus GOME-Spektren, zusätzlichen Eingangsdaten aus den Level 1 Daten und externen Quellen, sowie den gewünschten Ozonmessungen (Gesamtsäulen oder Profile) aufgebaut. Dann wird diese Datenbank in geeigneter Weise in Trainings- und Testdaten aufgeteilt, von denen die Trainingsdaten zur Belernung des neuronalen Netzes, und die Testdaten zur Kontrolle und Validierung der Ergebnisse dienen. Das auf diese Weise belernte Netzwerk kann dann im dritten Schritt verwendet werden, um den gesamten GOME-Datenbestand zu prozessieren.

Als Eingangsparameter für das Retrieval wurden bei Gesamt Ozon die sonnennormierten GOME-Radianzen von 313–340 nm, sowie ein atmosphärisches Fenster (382 nm), die O₃ Chappuis-Bande (600 nm) und die O₂ A-Bande (um 765 nm) verwendet. Weiterhin gingen Lageinformationen des Sensors und klimatologische Parameter (Latitude, Zeit) in das Retrieval ein. Das Ozonprofil-Netzwerk verwendet zusätzlich den GOME Kanal 1 bis hinunter auf 270 nm, sowie aus UKMO-Analysen ermittelte Temperaturprofile.

Für Gesamt Ozon wurde ein Kollokationsdatensatz aus bodengestützten Gesamt Ozonmessungen und klimatologisch korrigierten TOMS-Daten zusammengestellt, wobei letztere zur Verbesserung der geografischen Verteilung der Messungen dienen. Der Testdatensatz bestand aus sechs Stationen, deren Gesamt Ozonmessungen im Vergleich zu NNORSY durchschnittlich RMS-Fehler von knapp 12 DU lieferten. Darüberhinaus wurden die Retrievals anhand von Zeitreihen dieser Stationen (1996 bis Mitte 2001), sowie von zonalen Mittelwerten und Streudiagrammen aller Bodenstationen analysiert und mit den GDP v2.7 und den TOMS v7 Datenprodukten verglichen. Die Resultate lassen den Schluss zu, dass NNORSY das Gesamt Ozonfeld deutlich genauer als der GDP, und nur wenig schlechter als TOMS erfasst.

Bei der Ozonprofilbestimmung erweist sich der Aufbau eines Trainingsdatensatzes aufgrund der zusätzlichen Höhendimension und der Verwendung von vier verschiedenen Profildatentypen (Ozonsonden, HALOE, SAGE II, POAM III) im Vergleich zu Gesamt Ozonretrievals als erheblich komplexer. Durch eine Reihe von Qualitätskontrollen und Homogenisierungsmaßnahmen konnte jedoch eine für die Belernung geeignete Kollokationsdatenbank aufgebaut werden. Zusätzlich erforderten die in der Höhe unvollständigen Profildaten die Entwicklung und Validierung eines partiellen Belernungsverfahrens, was durch eine Modifikation von RPROP erreicht wurde.

Der Vergleich der Retrievalergebnisse mit den unabhängigen Testdaten ergab oberhalb von 10 km Höhe global eine durchschnittliche Reduktion der Standardabweichung gegenüber der Klimatologie von etwa 40%, unterhalb lag die Reduktion bei 10–30%. Eine Sensitivitätsuntersuchung auf diesen Daten ergab in tropischen Breiten generell ein schlechteres Abschneiden des Systems als außerhalb davon. Eine Abhängigkeit der Fehler von GOME-Pixeltyp (Ost, Nadir, West) konnte hingegen nicht festgestellt werden.

Ebenfalls auf Basis der Kollokationen wurde die Vertikalaufösung der NNORSY-Ozonprofile im Höhenbereich 12–32 km empirisch auf etwa 4 km abgeschätzt. Dieser im Vergleich

zu klassischen Verfahren niedrige Wert scheint den visuellen Eindruck zu bestätigen und läßt sich vermutlich durch die Mischung statistischer und physikalischer Information im Retrievalnetzwerk begründen.

Teil III

Sensitivitäten und Fehler

Der folgende dritte Teil der Arbeit hat die bessere Charakterisierung der NNORSY-Ozonprofilbestimmung zum Ziel.

Kapitel 8 geht zunächst dem Verhalten des Systems auf den Grund, indem die Fehlerstatistiken verschiedener Konfigurationen der neuronalen Netze verglichen werden. Dabei steht die Frage nach dem Informationsgehalt der Eingangsdaten im Mittelpunkt.

In **Kapitel 9** wird untersucht, wie sich für einzelne NNORSY-Retrievals Fehler spezifizieren lassen, die von den Eingangsdaten abhängen. Dazu werden drei verschiedene Methoden verglichen und die Jacobi-Matrix des neuronalen Netzes interpretiert.

Kapitel 10 versucht schließlich, ein über die bisherigen Statistiken hinausgehendes Fehlerbudget aufzustellen, indem die wichtigsten Fehlerquellen von der Kollokation bis zum Retrieval analysiert und kombiniert werden. Dadurch kann die Effizienz der Methode und ihre Genauigkeit in der Praxis abgeschätzt werden.

Kapitel 8

Variation der Netzwerkkonfiguration

Eine gängige Möglichkeit für die Sensitivitätsanalyse bei neuronalen feed-forward Netzen besteht in der gezielten Variation der Eingangs- und Zwischenneuronen. In der Tat gibt es sogenannte *evolutionäre Algorithmen*, die die Netzwerkkonfiguration durch zufällige „Mutation“ der Eingangs- und/oder Zwischenschicht, anschließendes Training und Selektion der erfolgreichsten Varianten optimieren [Fogel et al., 1990; Branke, 1995]. Die Größe des Ozonbestimmungsproblems in Kombination mit der verfügbaren Rechenleistung verhinderten leider bei der vorliegenden Arbeit den Einsatz eines derartigen Algorithmus'. Die Methodik kann aber genutzt werden, um anhand von Vergleichen zwischen einzelnen Netzen gezielt die Wirkung ausgewählter Eingangsparameter zu studieren.

Die Vergleiche zwischen Netzen verschiedener Konfiguration können auf der Basis von Trainings- und Testdatensatz erfolgen, wenn diese Datensätze für alle Konfigurationen dieselben Kollokationen enthalten. Aufgrund der großen Anzahl von Profilen in den Datensätzen (10^3 bis 10^5) ist die statistische Aussagekraft der Vergleiche recht hoch. Um die Signifikanz von Vergleichen einzelner Netze zu bewerten, muss allerdings berücksichtigt werden, dass aufgrund der zufälligen Initialisierung der Gewichte zu Beginn des Trainings (s. Abschn. 5.5) i. A. jedes Netz ein anderes Minimum der Fehlerfunktion findet, und damit die Ergebnisse für gleiche Netzkonfiguration ebenfalls gewissen Schwankungen unterworfen sind. Abb. 8.1 zeigt die Variation für zehn identische Konfigurationen mit verschiedener Gewichtsinitialisierung.

Bei dieser Art von Darstellung – die i. F. mehrfach verwendet wird – wird zunächst die Standardabweichung gegenüber Trainings- und Testkollokationen für alle beteiligten Netzwerke separat bestimmt. Anschließend werden für die Testdaten alle Standardabweichungen auf diejenige des Referenznetzwerks normiert, für die Trainingsdaten wird genauso verfahren. Zur besseren Übersicht werden die resultierenden Kurven außerdem mit einer Gaußfunktion von 4 km voller Halbwertsbreite gefaltet, dies hat sich als realistische Abschätzung der Vertikalauflösung erwiesen (s. dazu Abschnitt 7.4.3).

Offenbar liegen die zufälligen Abweichungen im Bereich weniger Prozent. Die formale statistische Untersuchung einer größeren Anzahl identischer Trainingsläufe wäre zwar wünschenswert, ist jedoch aufgrund der langen Rechenzeiten sehr aufwändig und konnte im Rahmen dieser Arbeit nicht durchgeführt werden. Da jedoch auch andere Experimente mit dem RPROP-Trainingsalgorithmus zeigen, dass die Ergebnisse üblicherweise nur in einem sehr engen Bereich variieren (vgl. Abb. 7.5), wird für die folgenden Untersuchungen davon ausgegangen, dass Abweichungen über 5% statistische Signifikanz besitzen. Dies mag angesichts von Abb. 8.1 etwas hoch angesetzt erscheinen, aber es ist zu beachten, dass zusätzliche Unsicherheiten durch die Zusammenstellung des Trainings- und Testdatensatzes entstehen,

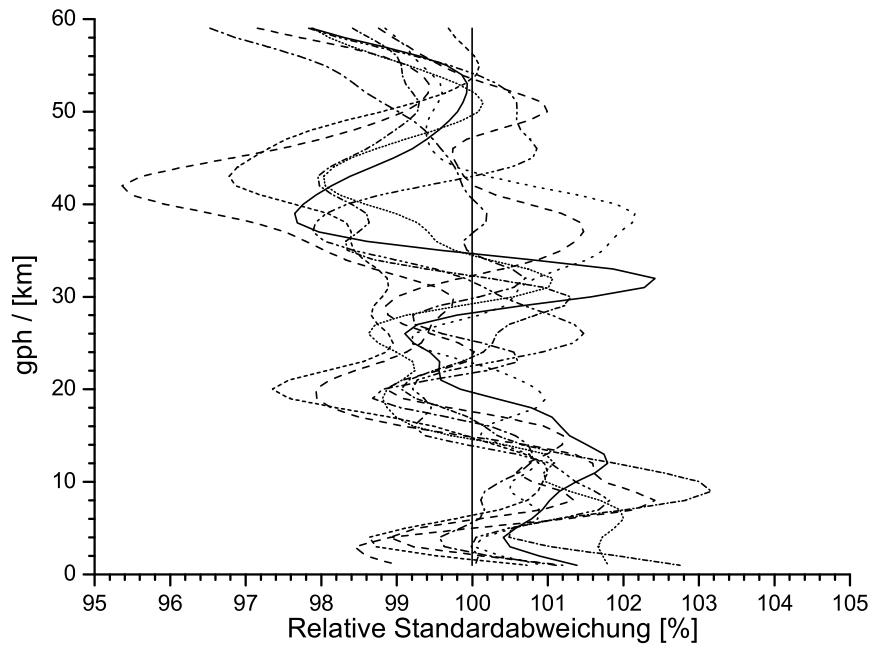


Abbildung 8.1: Vergleich der Standardabweichungen des Testdatensatzes von zehn Trainingsläufen neuronaler Netze gleichen Aufbaus, relativ zu einem zufällig ausgewählten Lauf. Die Schwankungen entstehen durch die zufällige Initialisierung der Verbindungsgewichte.

die sich für zukünftige Belernungen durch die Hinzunahme neuerer Kollokationen sicher ändern wird.

Ein weiterer Punkt der bei den folgenden Vergleichen berücksichtigt werden muss, liegt in der Anpassungsfähigkeit der neuronalen Netze. Die interne Gewichtung einzelner Eingangsneuronen variiert selbstverständlich mit deren Anzahl und Auswahl. Eine generelle Interpretation gestaltet sich deshalb als schwierig, die Argumentation verläuft aber wenigstens für den Gesamtfehler grob nach folgendem Schema: Ist das neue Ergebnis bei Training mit einem fehlenden Eingangsparameter – oder einem zusammenhängenden Satz derselben – im Vergleich zu vorher

schlechter, dann enthielt der Parameter offensichtlich nicht ersetzbare Information.

gleich gut, dann war der Parameter entweder für das Ergebnis nicht relevant oder redundant, d. h. sein Gewicht wird auf andere Eingangsneuronen mit gleichem Informationsgehalt aufgeteilt. Offensichtlich ist die Unterscheidung zwischen diesen beiden Fällen nicht immer leicht zu treffen, denn um Redundanzen auszuschließen müssten streng genommen alle möglichen Kombinationen der Eingangsparameter untersucht werden.

besser, dann steht der Parameter meist im Widerspruch zu einem (oder mehreren) anderen, so dass der Belernungsalgorithmus die Gewichtung aller widersprüchlichen Eingangsneuronen generell verringert. Bei Auflösung des Widerspruchs durch Weglassen des fraglichen Neurons kann dagegen die Information aus den verbliebenen Parametern wieder genutzt werden.

Falls die Verbesserung nur beim Testdatensatz auftritt, gibt es noch zwei weitere Möglichkeiten: Entweder wird die Anzahl der freien Parameter (Verbindungsgewichte) des Netzwerks gerade durch das Weglassen so weit gesenkt, dass eine bessere Generalisierung möglich ist. Dies tritt bei den untersuchten Konfigurationen jedoch vermutlich

nur ein, wenn eine beträchtliche Anzahl Eingangsneuronen gleichzeitig wegfällt, und diese nicht relevant sind. Oder die fraglichen Eingangswerte beschreiben den funktionalen Zusammenhang zwischen Netzein- und -ausgabe für den Trainingsdatensatz extrem gut und tragen dementsprechend viel zur Netzausgabe bei, das Ergebnis läßt sich jedoch nicht gut generalisieren.

8.1 Variation der GOME-Spektraldaten

8.1.1 Keine Spektraldaten (NNORSY Klimatologie)

Bevor die Frage nach dem Informationsgehalt der GOME-Kanäle angegangen wird, ist zu klären, wie viel Information das neuronale Netz überhaupt aus den Spektraldaten zieht, und wie viel auf die breiten- und zeitabhängige Klimatologie des Ozonfeldes zurückzuführen ist.

Zu diesem Zweck stellt Abb. 8.2 den Vergleich von Trainingsläufen dar, die lediglich mit klimatologischen Eingangswerten durchgeführt wurden. Die UKMO-Temperaturprofile, alle Spektraldaten sowie auf die Satelliten- bzw. Sensorposition bezogene Parameter wie Scanwinkel usw. wurden weggelassen, die Anzahl der Zwischenschicht-Neuronen wurde jedoch nicht verändert.

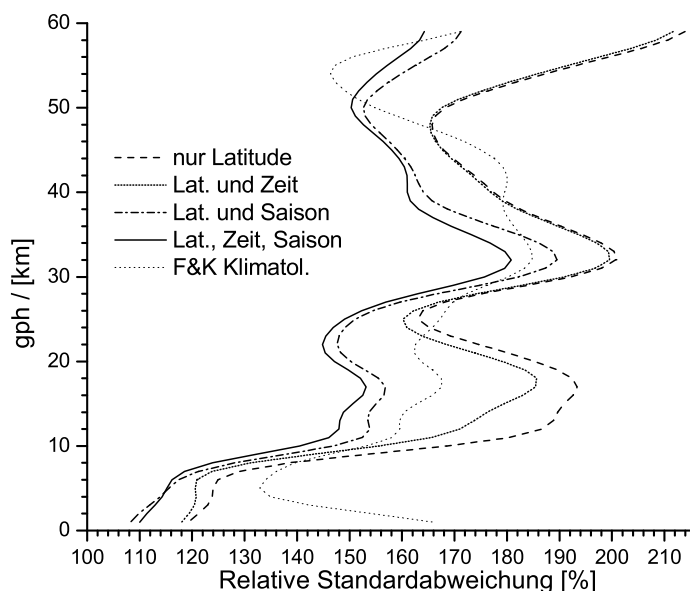


Abbildung 8.2: Testdatensatz-Standardabweichungen von klimatologischen Belernungen (ohne GOME-Daten) sowie der Klimatologie von Fortuin & Kelder [1998]. Alle Werte sind relativ zur entsprechenden Statistik der Standardbelernung (Tab. 7.2) angegeben. Die verwendeten Eingangsparameter sind die Latitude, das (lineare) Sensorenalter „Zeit“, die (zyklische) Jahreszeit „Saison“.

Wird lediglich die **Latitude** als Prädiktor für das Ozonprofil verwendet, so liegt die Standardabweichung in der Stratosphäre um 60 bis 115% über derjenigen der Referenzbelernung mit GOME- und UKMO-Daten. Auffallend ist jedoch, dass der Fehlerquotient in der unteren Troposphäre deutlich geringer ausfällt. Dies deutet bereits darauf hin, dass nur relativ wenig spektrale Information aus diesen Höhenschichten bis zum Satelliten vordringt bzw. verarbeitet werden kann.

Eine zusätzliche **lineare Zeitkoordinate** als Eingangsparameter bringt offensichtlich in der oberen Stratosphäre keinen großen Gewinn, und nur mäßige Verbesserungen unterhalb von 25 km. Das neuronale Netz kann aus einem monotonen Eingangswert keine zyklischen Schwankungen modellieren [Bishop, 1995a] und versucht daher, eine nicht-zyklische Funktion zu finden, die die ausgeprägten jahreszeitlichen Variationen des Ozonfeldes modelliert. Zur Verdeutlichung ist eine solche Funktion in Abb. 8.3 für die Teilsäule 12–16 km dargestellt, in der die Unterschiede zur reinen Latitudenbelernung am deutlichsten hervortreten. Tatsächlich kann das Netz hier zumindest ab 1999 einen Teil des Jahreszeitenzyklus erfassen,

aber die polynomähnliche Modellierungsfunktion stellt ansonsten keine gute Repräsentation des Zyklus dar. Das Ergebnis hängt zudem stark von den Anfangsbedingungen ab, jede Belernung findet eine völlig andere Funktion. Die Diskussion über lokale Minima bei diesem Problem und deren Vermeidung durch optimale Netzkonfiguration und Startwerte soll an dieser Stelle jedoch nicht vertieft werden. Es bleibt anzumerken, dass die Ergebnisse sich verbessern, wenn ein kürzerer Belernungszeitraum gewählt wird (1–2 Jahre).

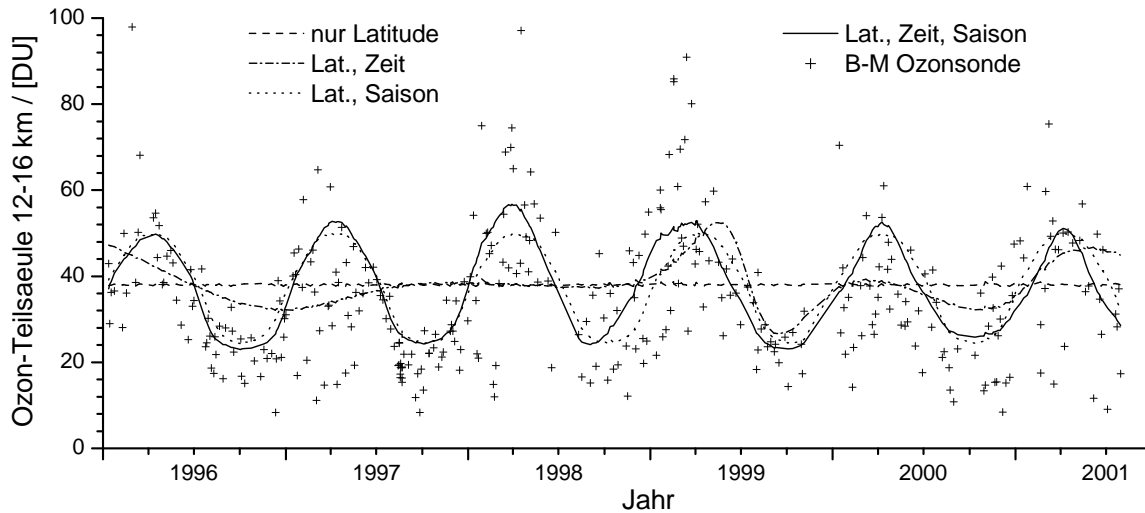


Abbildung 8.3: Vergleich der klimatologischen Belernungen aus Abb. 8.2 mit Ozonsondenmessungen der Station Hohenpeißenberg, integriert über die Höhenschicht 12–16 km. „Zeit“ bezeichnet das Sensorenalter, „Saison“ die zyklische Jahreszeitenkomponente. Die Sondenmessungen wurden explizit nicht für die Belernungen verwendet.

Ein umgekehrtes Bild ergibt sich, wenn stattdessen eine zyklische **Jahreszeitkoordinate** als zweiter Eingangsparameter verwendet wird. Diese verbessert das Resultat gegenüber der reinen Latitudenbelernung in allen Höhenschichten, Am deutlichsten jedoch oberhalb von 50 km und zwischen 10 und 24 km. Da die Moleküldichten in sehr großen Höhen äußerst gering sind, neigen die relativen Fehler dort leicht zu hohen Werten, das Ergebnis ist demnach in diesem Höhenbereich nicht überzubewerten. Der starke Einfluss der Jahreszeit macht sich insbesondere in Höhen um 15 km bemerkbar, wie auch der Vergleich mit Abb. 8.3 zeigt. Hier passt das Netz nun eine rein zyklische Funktion an die Ozonwerte des Trainingsdatensatzes an, die auch den Jahreszyklus in Hohenpeißenberg relativ gut widerspiegelt.

Wenn **beide Zeitkoordinaten** kombiniert werden, modifiziert die lineare Zeit Amplitude und Zeitpunkt der jahreszeitlichen Schwankungen, so dass klimabedingte Verschiebungen des Zyklus von Jahr zu Jahr berücksichtigt werden können. Dies resultiert in leichten Verbesserungen über den ganzen Höhenbereich (Abb. 8.2). Wiederum gilt, dass der positive Einfluss der linearen Zeit besser zum Tragen kommt, wenn der Belernungszeitraum kürzer ist. Zu beachten ist weiterhin, dass diese NNORSY-Klimatologie in den nördlichen mittleren bis hohen Breiten sowie in den Tropen keine wesentlichen Vorteile gegenüber der F&K-Klimatologie bringt (nicht dargestellt). Lediglich in der Troposphäre sinkt die Standardabweichung merklich, und der Höhenbereich über 25 km weist geringere Biases auf (s. dazu auch Abb. 11.1). Die in Abb. 8.2 zu erkennenden Verbesserungen beruhen zum großen Teil auf der genaueren Modellierung der Ozonfelddynamik um die Antarktis herum, insbesondere des Ozonlochs. Wie oben angedeutet ist die hier beschriebene NNORSY-Klimatologie jedoch nur ein Nebenprodukt des GOME-Ozonbestimmungsverfahrens, und daher in keiner Weise auf gute klimatologische Eigenschaften optimiert. Dass sich trotzdem

in einigen Bereichen ohne großen Zusatzaufwand schon deutliche Verbesserungen erzielen ließen, zeigt die prinzipiell sehr gute Eignung neuronaler Netze für die Konstruktion von Klimatologien aus inhomogenen Messungen. Mit entsprechender Optimierung von Datenauswahl, Belernungsverfahren und Netzwerkkonfiguration ließen sich die Ergebnisse sicher noch beträchtlich steigern.

Ein weiterer wichtiger Punkt, der sich anhand von Abb. 8.2 erkennen läßt, ist die im Mittel relativ schwache Verbesserung der Retrievalergebnisse in der mittleren bis unteren Troposphäre bei Hinzunahme von GOME-Spektraldaten und UKMO-Temperaturprofilen. Hieraus läßt sich folgern, dass der Beitrag spektraler Information in diesem Höhenbereich des Retrievals am geringsten ist, wie sich bereits in den Testdatensatz-Statistiken in Abschnitt 7.4 andeutete und aufgrund der Strahlungstransfertheorie zu erwarten war. Trotzdem liegen die beobachteten Verbesserungen von 10–30% weit oberhalb der mit 5% angesetzten Signifikanzgrenze. Beim Vergleich mit einer Belernung ohne T-Profil (nicht dargestellt) reduziert sich der Effekt zwar auf 10–20%, ist aber immer noch deutlich signifikant. Das neuronale Netz kann also zweifellos troposphärische Ozoninformation aus den GOME-Spektren extrahieren.

Fazit

Die Belernung von neuronalen Netzen mit lediglich klimatologischen Eingangsparametern führt zur Generierung einer NNORSY-Klimatologie, die am genauesten ist, wenn sowohl Latitude als auch eine lineare und eine saisonale Zeitkoordinate verwendet werden. Die Standardabweichung für Trainings- und Testdatensatz liegt dann oberhalb von 10 km Höhe etwa 45 bis 75% über der Referenzbelernung. Unterhalb von 10 km ist der Erkenntnisgewinn durch die Satellitendaten zwar geringer, aber immer noch deutlich vorhanden. Die NNORSY-Klimatologie erreicht auf den Testdaten im Mittel etwas geringere Biases und Standardabweichungen als die F&K-Klimatologie.

8.1.2 Spektralbereichs-Sensitivität

Gleich zu Anfang dieser Sektion muss darauf hingewiesen werden, dass sich das NNORSY-Verfahren überraschenderweise als relativ insensibel gegenüber der Auswahl spektraler Information aus den GOME-Daten erwiesen hat, wie bereits in Abschn. 7.3 erwähnt wurde. Während der Entwicklung des Verfahrens wurde eine große Anzahl verschiedener Spektralfensterkombinationen untersucht, allerdings waren die verwendeten Trainingsdaten, Netzwerkkonfigurationen, Belernungsalgorithmen und Pre-/Postprozessierungsschritte ebenfalls ständiger Weiterentwicklung unterworfen, so dass die Resultate nicht immer vergleichbar sind. Die wichtigsten Experimente wurden für die vorgestellte Version des NNORSY-Systems nochmals nachvollzogen und werden in diesem Abschnitt zusammengefaßt.

Alle hier vorgestellten Experimente wurden ohne die UKMO-Temperaturprofile als Eingangsdaten durchgeführt, damit das Netz gezwungen ist, möglichst viel Information aus den GOME-Spektren zu extrahieren. Wie sich in Abschn. 9.2 zeigen wird, bedingt die Verwendung der UKMO T-Profile eine Abnahme des Einflusses der Spektralinformation, so dass die Methode noch insensibler gegenüber deren Variation wird.

Im vorhergehenden Abschnitt wurde bereits dargestellt, dass der Fehler der NNORSY-Klimatologie im Vergleich zur Standardkonfiguration des Retrievalsystems um 10–75% höher liegt. Wie viel spektrale Information ist jedoch tatsächlich nötig, damit das neuronale Netz die hier klaffende Lücke schließen kann? Um dieser Frage nachzugehen, wurde ein Netz mit **minimaler Spektralinformation** trainiert. Außer Geolokation, Sensorenalter und Saison kamen hierbei lediglich einige auf 0.5 nm koaddierte Messpunkte bei den Wellenlängen 270 nm, 283 nm, 296 nm, 313 nm, 380 nm, 600 nm und 765 nm zum Einsatz,

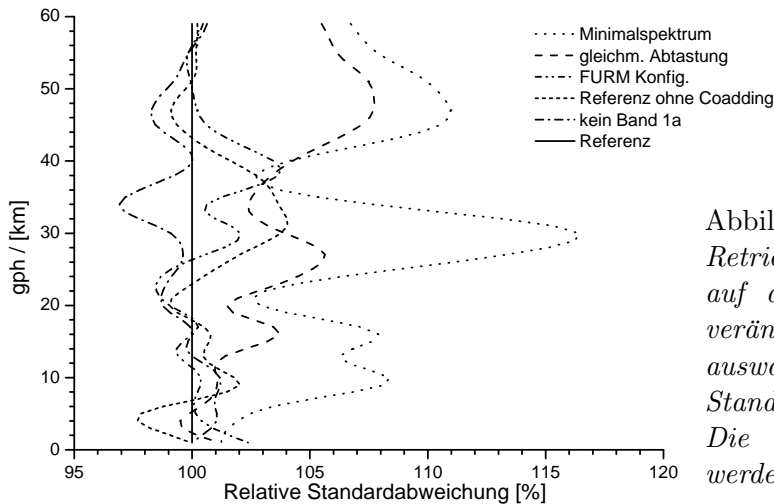


Abbildung 8.4: Sensitivität der Retrieval-Standardabweichung auf dem Testdatensatz gegenüber veränderter Eingangswellenlängenauswahl. Als Referenz dient der Standardlauf ohne UKMO T-Profil. Die einzelnen Konfigurationen werden im Text erläutert.

sowie in den Huggins-Banden die *on-* und *off-peak* Wellenlängen 325 nm und 323.8 nm, um die Temperaturabhängigkeit des Ozonabsorptionsquerschnitts zu erfassen [de Beek, 1998]. Jedes der im Standardlauf enthaltenen Wellenlängenfenster wurde also auf 1–2 Werte reduziert, wodurch sich insgesamt ein Netz mit nur 18 Eingangsneuronen ergibt (Standard ohne T-Profil: 112). Wie in Abbildung 8.4 zu sehen, reduziert bereits die in diesen neun Spektralwerten erhaltene Information den Fehler soweit, dass die Referenz in den meisten Höhen nur um 3–12% besser liegt. Lediglich im mittleren Höhenbereich, um 30 km, liegen die Fehler noch um 16% über der Referenz. Dies könnte ein Hinweis darauf sein, dass die relativ hoch aufgelösten GOME-Spektren nicht so viel mehr Ozoninformation enthalten, als die wenigen hier ausgewählten Spektralwerte — die zudem keineswegs den Anspruch erheben, eine optimale Wahl zu sein.¹ Allerdings setzt diese These voraus, dass das neuronale Netz die vorhandene Spektralinformation auch optimal nutzt, was bei nur neun Spektralwerten relativ wahrscheinlich ist [vgl. hierzu die Untersuchungen am TOVS-Sensor, Müller & Kaifel, 1999; Kaifel & Müller, 2000], aber für die standardmäßig verwendeten 102 Spektralwerte nicht garantiert werden kann (vgl. Abschn. 9.3).

Auf der anderen Seite wäre es interessant zu sehen, ob die momentane Standardkonfiguration tatsächlich alle Information in den GOME-Daten ausnutzt. Leider macht es wenig Sinn, ein neuronales Netz mit den gesamten ca. 3500 GOME-Spektralkanälen als Eingabe zu belernen – ohne eine gewaltige Menge an zusätzlichen Trainingsdaten (und Rechenzeit) könnte der Belernungsalgorithmus die vielen kleinen, rauschbedingten und/oder systematischen Widersprüche in diesen Daten nicht auflösen, und das Ergebnis wäre immer extrem suboptimal. Eine niedrig aufgelöste **gleichmäßige Abtastung** des gesamten GOME-Spektralbereiches kann jedoch testweise für die Belernung verwendet werden. Dabei wird ohne Rücksicht auf die Strahlungstransfertheorie jeder 25. Messwert aus Band 1, jeder 50. aus Band 2 und jeder 100. aus den Bändern 3 und 4 verwendet. Dies ergibt ein Netz mit 77 Eingangsneuronen. Wie ebenfalls aus Abb. 8.4 ersichtlich, schrumpft dadurch der Abstand zur Standardkonfiguration in manchen Bereichen bereits deutlich unter die mit 5% angenommene Signifikanzgrenze, wobei das Standardnetz allerdings noch in allen Höhen besser abschneidet.

Aufgrund der stochastischen Natur der neuronalen Netze ist die Optimierung der Spektralwertauswahl durch Belernung und Vergleichen ein sehr mühsamer und zeitaufwändiger Prozess. Die für die Standardkonfiguration erfolgte, physikalisch motivierte Auswahl der Wellenlängenfenster ist hier auf jeden Fall ein sehr effizienter Ansatzpunkt. Natürlich gibt

¹Die Verwendung von TOMS- oder SBUV/2-Spektralfenstern lieferte jedoch sehr ähnliche Ergebnisse.

es auch im Rahmen der physikalisch sinnvollen Spektralbereichswahl große Variationsmöglichkeiten. Aus den beiden bisherigen Experimenten ist aber bereits abzusehen, dass es sehr schwierig sein wird, die getroffene Wahl mit großem Effekt weiter zu verbessern. Um dies zu demonstrieren, wurden testweise einige Netze mit Variationen der von *de Beek* [1998] beschriebenen **FURM-Konfiguration** belernt, die von der NNORSY-Standardkonfiguration relativ stark abweicht. Sie verwendet sehr breite Koadditions-Intervalle (1–10 nm) von 250–290 nm, etwa 0.5 nm Koaddition im Bereich 290–340 nm, sowie 5 Messwerte aus den Chappuis-Banden zwischen 560 und 672 nm. Um die Information der bei FURM verwendeten Wolkenparameter in das Experiment einzubringen, wurde hier zusätzlich das standardmäßige O₂ A-Bandenfenster benutzt. Wie in Abb. 8.4 zu erkennen, ergibt sich mit diesem Netz (141 Eingangsneuronen) keine signifikante Änderung gegenüber der Referenz.

Dies wirft unter anderem die Frage auf, inwieweit sich das Koaddieren von Spektralwerten auf das Retrievalergebnis auswirkt, da die bei der FURM-Konfiguration gewählten großen Intervalle im Band 1A offenbar keine merkliche Verbesserung bewirken. Tatsächlich ergibt ein Trainingslauf mit der Standardkonfiguration, aber **ohne Koaddieren**, wenig Änderung gegenüber dem Standardlauf. Hierbei wurde anstatt jeweils n Spektralwerte zu mitteln nur jeder n . Wert verwendet. Lediglich in dem direkt über dem Ozonpeak liegenden Höhenbereich zeichnen sich leichte Verbesserungen durch Koaddition ab, die jedoch nicht ganz an die Signifikanzgrenze heranreichen. Offenbar ist das neuronale Netz in der Lage, über das Rauschen einzelner Dioden hinwegzumitteln, wenn insgesamt genügend Spektralwerte vorliegen. Wie die Ausnutzung dieser Redundanz im konkreten Fall aussieht, wird in Abschn. 9.3 anhand der Jacobi-Matrix des neuronalen Netzes näher untersucht. Es ist davon auszugehen, dass die Koaddition in der Praxis eventuell Vorteile bringt, falls sich die Eingangsdatenqualität gegenüber den Trainingsdaten verschlechtert. Die implizite Rauschkorrektur des Netzes [*Bishop*, 1995b] müsste dann im Fall einzelner, ungemittelter Messwerte größere Schwankungen ausgleichen, als wenn diese bereits durch Koaddition reduziert werden.

Nachdem die bisherigen Betrachtungen sich mit den Auswirkungen von Spektralfensterwahl und -auflösung befassten, soll im Folgenden die Frage nach der Höhenzuordnung spektraler Information kurz angeschnitten werden. So sollte z. B. eine Belernung, deren spektrale Eingangswerte weniger weit in das UV hineinreichen, im oberen Bereich des Ozonprofils höhere Fehler produzieren. Wie aus Abb. 8.4 ersichtlich, ist dies jedoch nicht der Fall, wenn **Band 1A** aus der Standardkonfiguration entfernt wird – die Änderung der Standardabweichung ist nicht signifikant. Wie Eingangs des Kapitels angedeutet kann hieraus jedoch nicht zwangsläufig geschlossen werden, dass das neuronale Netz keine Information aus diesem Band gewinnt. In der Tat verringerte sich beim Hinzufügen der standardmäßigen Band 1A Spektralwerte zur besten klimatologischen Netzkonfiguration aus Abschn. 8.1.1 der erreichte Testdaten-RMSE um 10.2%, verteilt über die ganze Profilhöhe. Jedoch läßt sich dieser Informationsgewinn offenbar auch an anderer Stelle der Eingangsdaten erzielen. Aber wie viel Spektralinformation muss tatsächlich entfernt werden, bis der Ausgleich in einigen Schichten nicht mehr gelingt?

8.1.3 Überlegungen zur Lokalisierbarkeit von Höheninformation

Abb. 8.1.3 zeigt ein Experiment, bei dem die Standardkonfiguration auf Spektralbereiche ≥ 300 nm beschnitten wurde. Um die Aussagekraft zu erhöhen, stellt die Abbildung Mittelwert und Fehler von zehn Paaren von Netzen – jeweils Standard- und reduziertes Netz – dar, die mit unterschiedlicher Gewichtsinitialisierung belernt wurden. Auffällig ist zunächst das Maximum der Fehlerzunahme um 29 km Höhe, was zumindest für kleine Sonnenzenitwinkel in etwa der Eindringtiefe der UV-Strahlung bei 300 nm entspricht [*Hoogen*, 1998].

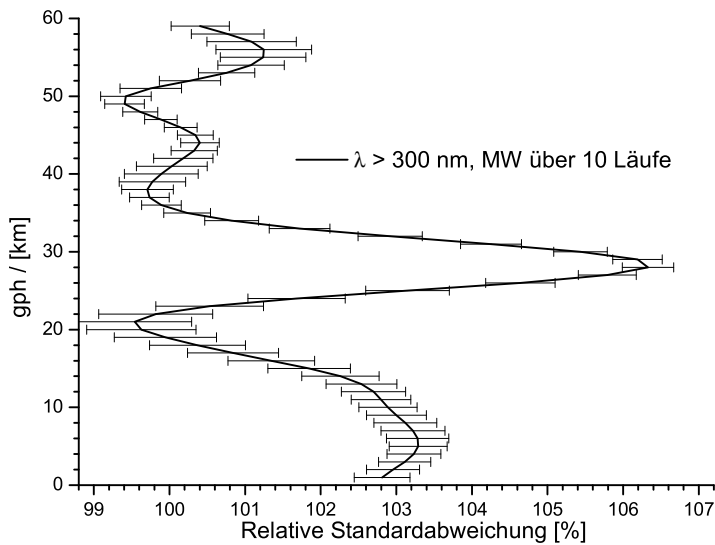


Abbildung 8.5: Wie Abb. 8.4, aber gemittelt über zehn Trainingsläufe. Die Balken geben die Standardfehler des Mittelwertes an.

Bei größeren SZAs oder kürzeren Wellenlängen dringt die Strahlung jedoch um 10–20 km weniger weit in die Atmosphäre ein (Abb. 3.2). Da ab 300 nm wieder Spektralwerte verfügbar sind, sollte diese höhere Region um 40–50 km nach klassischer Theorie eigentlich die größten Fehler aufweisen – dies ist jedoch nicht der Fall, statt dessen taucht ein sekundäres Fehlermaximum in der Troposphäre auf. Beim Entfernen anderer Spektralbereiche ergibt sich ein ähnliches Bild (nicht dargestellt): Die Standardabweichung nimmt zwar gemittelt über die Höhe um einige Prozent zu, die entstehenden Fehlerzunahme-Profile verhalten sich jedoch chaotisch und entziehen sich der strahlungstransfertheoretischen Interpretation. Die entstehenden Fehlerprofile variieren zudem relativ stark mit der Auswahl von Trainings- und Testdaten, sowie mit der verwendeten Ein- und Ausgabenormierung (Abschn. 8.3).

Diese Ergebnisse legen nahe, dass es bei NNORSY offenbar sehr schwierig ist, die spektrale Information unterschiedlichen Höhenbereichen des Ozonprofils zuzuordnen, wie dies bei klassischen Verfahren mit Hilfe der Beitragsfunktionen möglich ist. Insbesondere erscheint es auf den ersten Blick unlogisch, dass Messungen im kurzwelligen Spektralbereich die Berechnung der Ozonkonzentration in niedrigen Atmosphärenschichten beeinflussen, aus denen ganz offensichtlich nie Strahlung bis zum Sensor vordringen kann.

Um diesen Effekt zu verstehen, ist es hilfreich, ein kleines **Gedankenexperiment** für eine stark vereinfachte Konstellation durchzuführen. Nehmen wir an, wir betrachten einen Satz Ozonprofile aus einem schmalen Latitudenband, die alle denselben integrierten Gesamt Ozonwert und in etwa dieselbe Profilform besitzen, nur dass der Ozonpeak um einige Kilometer in der Höhe variiert. Zur Bestimmung der Profile seien zugehörige UV-Spektren gegeben, die alle ohne Wolken und Instrumenteneffekte, sowie beim selben Sonnenzenith- und Scanwinkel vom Satelliten gemessen wurden. Wir wollen nun ein Profil bestimmen, dessen Peak nach oben verschoben ist. Da die Minima der Gewichtsfunktionen \mathbf{K} aus der Strahlungstransportrechnung zu kurzen Wellenlängen hin nach oben wandern, wird das zugehörige Spektrum gegenüber dem mittleren Spektrum etwas weniger Strahlungsintensität im kurzwelligen, dafür etwas mehr im langwelligen Spektralbereich aufweisen.

Beim Retrieval mittels Optimal Estimation böte sich ein Ozonprofil mit mittlerer Peakhöhe als *a priori* an. Gemäß den physikalischen Beitragsfunktionen wandelt das Verfahren die fehlende Strahlungsintensität im kurzwelligen Bereich in einen positiven Beitrag zur Ozonkonzentration oberhalb des *a priori* Peaks um, während die zusätzliche langwellige Strahlung zu einer Ozonabnahme unterhalb des Peaks führt. Dadurch wandert der Peak nach oben. Die Zuordnung von Wellenlänge zum Höhenbereich ist eindeutig möglich.

Ein neuronales Netz hingegen kennt die Strahlungstransportrechnung nicht, es kann jedoch aus einer Stichprobe des Datensatzes lernen, dass ein „Kippen“ des Spektrums zum Langwelligen hin mit einer Verschiebung des Ozonpeaks nach oben einhergeht. Aber woher soll es wissen, dass das Nachlassen der Strahlung im Kurzwelligen mit Ozonzunahme oberhalb des Peaks assoziiert ist, während die Zunahme im Langwelligen von einer Ozonabnahme unterhalb des Peaks herrührt? Die Beitragsfunktionen sowohl des langwelligen als auch des kurzwelligen Bereiches werden daher Betragsmaxima unter- und oberhalb des Ozonpeaks aufweisen, die höhenmäßige Lokalisierung der Information ist nicht mehr möglich.

An diesem sehr einfachen Beispiel wird bereits deutlich, wie das neuronale Netz statistische und physikalische Information mischt. In der Retrieval-Praxis ist die Situation natürlich um ein vielfaches komplizierter, die implizite Klimatologie sowie Störungen durch Wolken, Instrumentenfehler, Rauschen, Beobachtungsgeometrie, Ozonfeld-Dynamik, etc. tragen dazu bei, den Zusammenhang zwischen Eingangswellenlänge und Ozonprofil-Änderung weiter zu verschleiern. Weiterhin tragen auch die verschiedenen Normierungsschritte zwischen Netzeingang und -ausgang zur Delokalisierung der Information bei, wie dies selbst für den mathematisch einfacheren Fall der Vorwärtsmodell-Simulation mit Hilfe neuronaler Netze bereits beobachtet wurde [Aires *et al.*, 1999; Chevallier & Mahfouf, 2000]. Trotz der geschilderten Schwierigkeiten wird in Kapitel 9 nochmals versucht werden, physikalische und statistische Information mit anderen Methoden zu trennen.

8.2 Effekt anderer Netzwerkparameter

Tabelle 8.1 gibt einen Überblick über die Sensitivität der NNORSY-Methode gegenüber anderen Eingangsparametern (vgl. Tab. 7.2), soweit sich diese aus den Belernungsexperimenten ergeben. Auf die Darstellung als Fehlerprofile wurde aufgrund der schlechten Interpretierbarkeit verzichtet; Der RMS-Gesamtfehler einer Belernung stellt bereits ein gutes Maß für den effektiven Informationsgewinn durch einen Eingangswert dar. Wie der Belernungsalgorithmus diese Information über die Höhe verteilt, variiert mit der Konfiguration des gesamten restlichen Systems und kann an dieser Stelle nicht mit allgemeiner Gültigkeit wiedergegeben werden.

Die RMSE der zweimal zehn Belernungen aus Abb. 8.1.3 weisen relative Stichproben-Standardabweichungen von 0.45% und 0.41% für die Test- und 0.53% bzw. 0.58% für die Trainingsdaten auf. In Anbetracht der zu Beginn des Kapitels angestellten Überlegungen können demnach bereits Veränderungen von etwa 1% als signifikant angesehen werden.

Die Tabelle zeigt zunächst, dass die Veränderungen bei Trainings- und Testdatensatz meist unterschiedlich sind. Ist nach Weglassen eines Eingangsparameters die Zunahme des Fehlers für den Testdatensatz nennenswert größer, dann nimmt die Generalisierungsleistung des Netzes offenbar stärker ab als seine Fähigkeit, die Trainingsprofile zu rekonstruieren. Mit anderen Worten kann der Belernungsalgorithmus durch den betreffenden Parameter allgemein gültige physikalische und/oder statistische Gesetzmäßigkeiten ableiten. Zu diesen Parametern zählen eindeutig **LOS-Flags**, **SZA** und **Latitude**.

Im umgekehrten Fall von viel größeren Trainings-Fehlerzunahmen dient ein Teil der fehlenden Information dazu, spezifische Merkmale des Trainingsdatensatzes zu erlernen, die nicht auf die Testdaten – und somit höchstwahrscheinlich auch nicht in die operationelle

Variation	Training	Test
keine LOS Flags	1.2%	3.9%
kein SZA	1.9%	5.5%
keine Latitude	2.3%	5.5%
keine Saison	1.4%	0.5%
kein Sensorenalter	3.5%	2.3%
kein UKMO T-Profil	9.1%	8.2%
mit PMD-Kanälen	-0.2%	0.9%
mit Tropopausenhöhe	-0.1%	1.1%
65 Zwischenneuronen	0.0%	0.9%
35 Zwischenneuronen	2.1%	2.3%

Tabelle 8.1: *Veränderung des RMSE auf Trainings- und Testdatensatz gegenüber der Standardbelernung bei Variationen der Netzwerkkonfiguration. Die 2σ -Signifikanzgrenze liegt bei etwa 1%.*

Anwendung hinein – übertragbar sind. Solange der Testdatenfehler bei Hinzunahme der Parameter trotzdem abnimmt, wie zum Beispiel in der Tabelle bei den beiden **Zeitkoordinaten**, gibt es keinen Grund diese Parameter nicht zu verwenden. Dabei ist jedoch größere Vorsicht geboten; Es empfiehlt sich hier die Überprüfung des Ergebnisses anhand von weiteren, unabhängigen Evaluierungsdaten. Dadurch wird sichergestellt, dass nicht die oben erwähnten Trainingsdaten-Merkmale z. B. aufgrund des Datenauswahlverfahrens zufällig auch im Testdatensatz auftreten, aber nicht in der operationellen Anwendung [Plutovski, 1996]. Den Zeitkoordinaten ist daher eine Fallstudie in Abschn. 11 gewidmet.

Die absolut größte Verbesserung der Ergebnisse bringt die Verwendung des **UKMO-Temperaturprofils**. Wie sich in Abschn. 9.2 zeigen wird, bewirkt sie in der Tat eine grundsätzliche Änderung der Eingangsdatengewichtung, bei der die Spektralinformation an Bedeutung verliert, aber das gesamte Verfahren stabilisiert wird. Hingegen führt die Hinzunahme der **PMD-Information**, wie bereits in Abschn. 7.3 andiskutiert, nicht zu merklichen Vorteilen, sondern tendenziell eher zu einer Verschlechterung der Generalisierungsfähigkeit. Wahrscheinlich ist die in den 48 PMD-Werten pro Pixel enthaltene, flächige Information auf zu komplexe Weise mit Bodenalbedo, Wolken und Ozonprofil verknüpft, um sie widerspruchsfrei auf Basis der Trainingsdaten zu modellieren.

Das Hinzufügen der **Tropopausenhöhe** (TPH) zu den Eingangswerten geht ebenfalls nicht mit Verbesserungen einher, wenn gleichzeitig die Temperaturprofile verwendet wurden, wie in der Standardbelernung der Fall. Die TPH-Werte wurden von Mark Weber (IUP) aus UKMO-Analysen berechnet. Die TPH ist dabei als die Höhe der 3.5 PVU Fläche, bzw. des 380 K potentiellen Temperaturniveaus definiert, wobei der jeweils niedrigere Wert maßgeblich war. Möglicherweise treten hier Widersprüche zu den UKMO-Temperaturprofilen auf, deren Ursache nicht ganz klar ist. Für sich alleine betrachtet tragen die Daten durchaus Informationen über das Ozonprofil bei: Die Hinzunahme der TPH zu einer T-Profil freien Referenzbelernung läßt die Fehler auf Trainings- und Testdatensatz um je 1.4% sinken. Es ist jedoch zu vermuten, dass die Information aus den Tropopausenhöhen noch besser genutzt werden könnte, indem das Retrieval-Höhengitter dynamisch und relativ zur Tropopause definiert wird [Bojkov & Fioletov, 1997]. Darauf wird später noch kurz eingegangen.

Bei den Arbeiten mit verschiedenen Trainingsdatensätzen hatte sich im Laufe der letzten Jahre empirisch eine Anzahl von etwa 45 Neuronen in der Zwischenschicht als geeignete Balance zwischen Freiheitsgraden des Netzwerkes und Informationsgehalt der Trainingsdaten erwiesen. Zu viele Zwischenneuronen führen zu Übertrainieren, also Zunahme des Testdaten-RMSE; Allerdings setzt dieser Effekt bei der großen und komplexen hier verwendeten Datenmenge erst relativ spät ein. Zu wenige Zwischenneuronen bedingen eine Zunahme von Trainings- und Test-RMSE, weil die Komplexität des Datensatzes nicht mehr vollständig modelliert werden kann. Zwischen diesen Extrema liegt ein relativ breiter Be-

reich, in dem die Neuronenanzahl ohne nennenswerte Fehleränderung variiert werden kann. Aus praktischen Gründen ist es zweckmäßig, die Zwischenschichtgröße am unteren Ende dieses Bereichs zu orientieren. Dies spart zum einen Rechenzeit bei der Belernung², zum anderen kann Übertrainieren auch bei Belernungskontrolle anhand eines Testdatensatzes nie mit absoluter Sicherheit ausgeschlossen werden [Plutovski, 1996], so dass ein gewisser Sicherheitsabstand zum oberen Ende der Skala angebracht ist. Die entsprechenden Resultate in Tab. 8.1 belegen, dass die Auswahlbedingung mit 45 Zwischenneuronen gut erfüllt ist.

8.3 Normierung von Ein- und Ausgabe

Die Normierung von Ein- und Ausgabewerten dient bei neuronalen Netzen einerseits zur Erzeugung statistisch gleichmäßigerer Verteilungen, andererseits zur Vermeidung von numerischen Fehlern durch die begrenzte Rechengenauigkeit der Computer.

8.3.1 Eingangsnormierung

Durch eine einfache lineare Skalierung der Eingangswerte wird erreicht, dass alle Eingangsneuronen Werte von der Größenordnung 1 erhalten, so dass die Verbindungsgewichte in der nachfolgenden Schicht ebenfalls alle in etwa dieselbe Größenordnung besitzen. Meist wurde hierzu jeder verwendete Wert der Rohdaten \tilde{y}_i mit seinem Mittelwert \bar{y}_i und seiner Standardabweichung σ_i gemäß

$$y_i = \frac{\tilde{y}_i - \bar{y}_i}{\sigma_i} \quad (8.1)$$

skaliert. Eine andere, vor allem für Daten mit festem Wertebereich $[\tilde{y}_{i,\max}; \tilde{y}_{i,\min}]$, wie etwa Winkelkoordinaten, verwendete Normierung ist

$$y_i = \frac{2(\tilde{y}_i - \tilde{y}_{i,\min})}{\tilde{y}_{i,\max} - \tilde{y}_{i,\min}} - 1, \quad (8.2)$$

wodurch der Wertebereich auf das Intervall $[-1;1]$ abgebildet wird.

Da die gesamte für die statistische Beschreibung des Retrievalprozesses verwendete Bayessche Theorie von gaußverteilten Größen ausgeht, ist es wünschenswert, dass sowohl Eingangs- als auch Ausgangswerte möglichst gaußähnlich verteilt sind. In der Praxis ist dies nicht immer der Fall, z. B. beim Sonnenzenitwinkel oder bei periodischen Größen wie der saisonalen Zeitkoordinate. Es hat sich jedoch gezeigt, dass die hier verwendeten neuronalen Netze recht tolerant gegenüber nicht-gaußverteilten Eingangsgrößen sind. Die Logarithmierung der Spektralwerte brachte für T-Profil freie Belernungen eine Verringerung des RMS-Gesamtfehlers um knapp 5% mit sich, da die Verteilung der Messwerte für die meisten Wellenlängen dadurch etwas gauß-ähnlicher wurde. Für Belernungen mit T-Profil waren die Verbesserungen jedoch nicht mehr signifikant.

8.3.2 Ausgangsnormierung

Die Auswirkungen unterschiedlicher Skalierung der Netzwerkausgabe, also der Ozonprofile, sollen anhand von Abb. 8.6a erläutert werden. Bei linearer, unskalierte Ausgabe wird das Netz so trainiert, dass die Ausgangsneuronen direkt die Ozonmoleküldichte in Einheiten von $[10^{18}/m^3]$ angeben. Diese Werte schwanken jedoch um etwa drei Größenordnungen, von im Mittel etwa 0.7 bei 1 km, 4 bei 23 km bis 0.004 bei 60 km GPH. Dies führt aufgrund

²Für die (operationelle) Anwendung ist die Rechenzeit hingegen völlig unkritisch, siehe Kapitel 13.

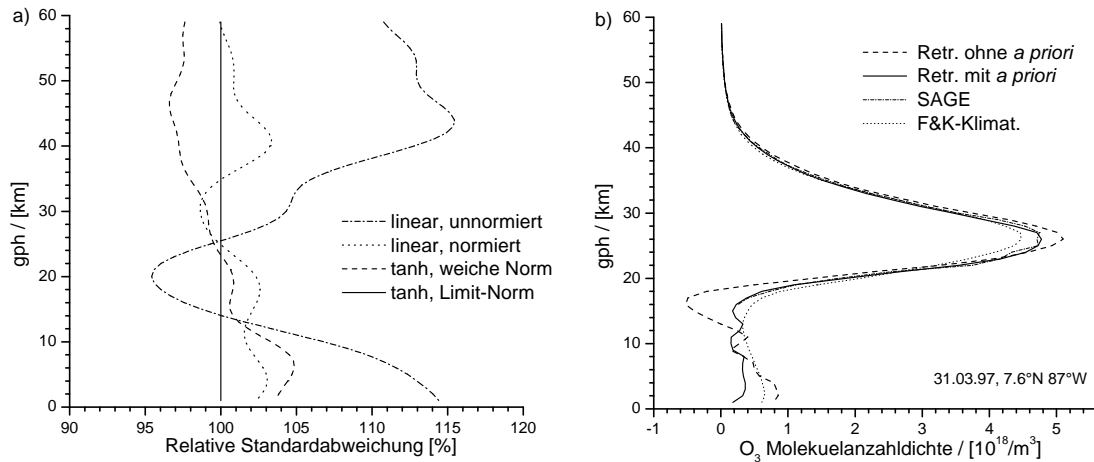


Abbildung 8.6: Auswirkung der Ausgabe-Normierung auf das NNORSY-Ozonprofilretrieval. (a) Testdatensatz-Statistik für Belernungen mit verschiedenen Ausgabe-Neuronen (linear, tanh-Transferfkt.) und Normierungen. Limit-Norm bedeutet die Fixierung auf Null als unteren Grenzwert der Ozonkonzentration. (b) Die so eingebrachte a priori-Information hilft, in schlecht abgetasteten Gebieten des Zustandsraumes – hier die äquatoriale Troposphäre – instabile Schwingungen des Systems zu vermeiden.

von Rundungsfehlern zu einer Überbewertung der mittleren Höhenlagen auf Kosten der anderen Bereiche. Dieser Effekt kann durch Normierung der Werte aller Höhenschichten auf das Intervall $[-1;1]$ vermieden werden.

Weiterhin bietet der verwendete Belernungsalgorithmus die Möglichkeit, die Ausgabe-Neuronen statt mit linearer Transferfunktion ebenfalls mit einer Sigmoiden zu versehen. Hierzu wird dieselbe *tanh*-Funktion wie bei den Neuronen der versteckten Schicht benutzt. Da das Netz jetzt nur noch Ausgabewerte im Intervall $]-1;1[$ liefern kann, muss das Ozonprofil zwangsläufig normiert werden. Dabei ist zu beachten, dass die Randwerte des Intervalls nur schwer erreicht werden, weil die Steigung des Arcustangens dort so gering ist, d. h. die Aktivierung des Ausgabeurons muß sehr hohe Werte annehmen. Aus diesem Grund wird das Ozonprofil mittels min/max-Norm (Gl. 8.2) auf das Intervall $[-0.9;0.9]$ skaliert, wobei als Maximalwerte für Ozon die um einen Sicherheitsrahmen von 10% erhöhten, im Trainingsdatensatz auftretenden maximalen Ozonmoleküldichten verwendet wurden.

Man beachte, dass dies eine Form der Einbringung von *a priori*-Wissen darstellt: Die Lösung des Retrievalproblems wird hier auf einen physikalisch „vernünftigen“ Bereich des Zustandsraumes eingeschränkt. Im Mittel ergeben sich durch diese Maßnahme keine wesentlichen Verbesserungen gegenüber der normierten, linearen Ausgabe (Abb. 8.6a), weswegen letztere auch bei den meisten Problemstellungen ausreichend ist [Bishop, 1995a] — das neuronale Netz neigt bei guten Trainingsdaten ohnehin nicht zu Ausreißern.

Da *a priori* ebenfalls bekannt ist, dass die Ozonmoleküldichte nicht negativ werden kann, wird standardmäßig das Intervall $[-1.0;0.9]$ für die Normierung verwendet und die untere Intervallgrenze mit der Ozondichte 0 assoziiert. Dies ergibt zwar wiederum im Mittel keine besseren Ergebnisse, im Einzelfall jedoch durchaus, insbesondere bei partieller Belernung in Bereichen, in denen die Trainingsdaten lückenhaft sind. Ein Beispiel davon ist in Abb. 8.6b zu sehen: Für das betreffende Netz wurden keine SHADOZ-Ozonsondendaten verwendet, weswegen die Information über die obere tropische Troposphäre fast ausschließlich aus – teilweise durch hohe Wolken kontaminierten – Okkultationsdaten stammt (s. Abschn. 7.1). Im Fall der linearen Ausgabe kommt das Retrievalverfahren hier ins Schwingen und produziert negative Werte. Die *a priori* Regularisierung unterdrückt hingegen die (unphysikalische)

Schwingung und führt insgesamt zu einem stabileren Ergebnis.

Fazit

Im diesem Kapitel wurde untersucht, wie sich die Retrievalfehler auf den Kollokationsdatensätzen bei Belernung mit unterschiedlicher Netzwerkkonfiguration verhalten. Diese Methode der Sensitivitätsanalyse liefert ein relativ klares, höhenaufgelöstes Bild von den Auswirkungen größerer Änderungen der Eingangskonfiguration sowie der Normierung. So konnte gezeigt werden, dass die NNORSY-Methode in der Tat eine Klimatologie repräsentiert, wenn keine Spektraldaten verwendet werden, und dass es im Höhenbereich oberhalb des Ozonpeaks, von etwa 25 km bis 40 km am stärksten von der Hinzunahme der GOME-Messungen profitiert. Dieses Resultat steht in guter Übereinstimmung mit der Strahlungstransfertheorie und dem GOME S/N-Spektrum.

Ein weiteres wichtiges Ergebnis ist die Tatsache, dass aufgrund der Eigenschaften neuronaler Netze eine feinere Analyse der Höhenzugehörigkeit spektraler Information mit der angewandten Methode nicht befriedigend möglich ist. Theoretische Überlegungen legen den Schluss nahe, dass dies im vorliegenden Fall auch mit anderen Ansätzen schlecht machbar ist, weil das neuronale Netz physikalische und statistische Zusammenhänge intern zu sehr vermengt. Insgesamt ist die Methode sehr tolerant gegenüber Variation von Spektralfenstern und -auflösung.

Aus demselben Grund wurde für die Sensitivitätsanalyse anderer Eingangsparameter der über die Höhe gemittelte RMSE verwendet. Die NNORSY-Methode profitiert demnach in der Anwendung stark vom UKMO-Temperaturprofil, mittelstark von SZA, Latitude und LOS-Flags und eher schwach vom Sensorenalter. Für saisonale Komponente und PMD-Messungen ließen sich keine signifikanten Änderungen nachweisen, während die aus UKMO-Daten errechnete Tropopausenhöhe nur bei T-Profil freien Belernungen leichte Verbesserungen bewirkt.

Für das neuronale Netzwerk stellen Ein- und Ausgabenormierung eine Art der Regularisierung durch das Einbringen von *a priori* Information über geophysikalische Zusammenhänge dar. Die Logarithmierung und Normierung der GOME-Spektren erlauben dem neuronalen Netz, spektrale Information leichter zu extrahieren, während die nichtlineare Beschränkung der Ozonprofilausgabe auf physikalisch sinnvolle Werte zur Stabilität der Methode beiträgt.

Kapitel 9

Abschätzung lokaler Fehler

Die bisherige Fehlerdiskussion in Kapitel 7 beschränkte sich auf Vergleiche mit Kollokationen verschiedener anderer Instrumente. Dieses Verfahren besitzt den Nachteil, dass spezielle Situationen nicht individuell berücksichtigt werden können, weil sie immer in der erforderlichen Menge von Kollokationen untergehen. Mit anderen Worten ist es denkbar, dass die Retrievalfehler lokal stark von den Breitengrad- oder saisonalen Mittelwerten abweichen.

Dieses Kapitel wird diesem Umstand Rechnung tragen, indem drei verschiedene Methoden zur Abschätzung von Fehlern einzelner Ozonprofile vorgestellt und verglichen werden. Anschließend werden die Methoden eingesetzt, um durch weitere Sensitivitätsanalysen einen besseren Einblick in die innere Struktur des Retrievalnetzwerks zu erhalten.

9.1 Methodenbeschreibung

9.1.1 Retrieval mit verrauschten Eingangsdaten

Eine sehr praxisnahe Möglichkeit, die Präzision des Retrievals abzuschätzen, besteht darin, den Retrievalprozess C Mal unter Verwendung von künstlich verrauschten Eingangswerten zu wiederholen. Sei dazu ϵ der künstlich erzeugte, gaußverteilte Fehler des unnormierten Eingangsdatenvektors $\tilde{\mathbf{y}}$, dann ist

$$\mathbf{x}_c = \mathbf{R}_{\text{MLP}}(\nu_{\text{in}}(\tilde{\mathbf{y}} + \epsilon_c), \mathbf{w}) \quad (9.1)$$

das Retrievalergebnis von Versuch c . In Ermangelung besseren Wissens wird hierbei die Eingangsfehlerkovarianz als diagonal angenommen [de Beek, 1998], d. h. die einzelnen Komponenten von ϵ sind nicht korreliert. Die Eingangsnormierung $\nu_{\text{in}}(\cdot)$ des Netzwerks wird in Abschnitt 8.3 näher erläutert. Der Bezug auf den unnormierten Eingangsvektor $\tilde{\mathbf{y}}$ anstatt auf \mathbf{y} dient hier der besseren Kontrolle der Eingangsfehler, da diese für $\tilde{\mathbf{y}}$ leichter zu spezifizieren sind. $\tilde{\mathbf{y}}$ enthält die vom GOME-Instrument gemessenen, sonnennormierten und logarithmierten Radianzen¹, die Winkelkoordinaten, Flags, das Temperaturprofil, etc. Damit läßt sich die Eingangsnormierung schreiben als

$$y_i = \nu_{\text{in}}(\tilde{y}_i) = \begin{cases} \tilde{y}_i & \text{für Flags} \\ \frac{\tilde{y}_i - \langle \tilde{y}_i \rangle}{\sigma_i} & \text{für Spektralwerte und Temperatur} \\ 2 \frac{\tilde{y}_i - \tilde{y}_{\text{min},i}}{\tilde{y}_{\text{max},i} - \tilde{y}_{\text{min},i}} - 1 & \text{für Winkel- und Zeitkoordinaten,} \end{cases} \quad (9.2)$$

wobei σ_i , $\tilde{y}_{\text{min},i}$ und $\tilde{y}_{\text{max},i}$ Standardabweichung, Minimum und Maximum von Eingangswert \tilde{y}_i bezeichnen, jeweils bezogen auf den Trainingsdatensatz.

¹Wie zu erkennen durchlaufen die Spektralwerte mehrere Normierungsschritte, von denen mit *Eingangsnormierung* i. F. nur der letzte gemeint ist.

Sei weiterhin $\hat{\mathbf{x}}$ das Retrievalergebnis bei unverraushtem Eingangsvektor. Dann kann eine **Rauschfehlerkovarianzmatrix** definiert werden mittels

$$\mathbf{S}_{\text{RF}} = \langle (\mathbf{x}_c - \hat{\mathbf{x}})(\mathbf{x}_c - \hat{\mathbf{x}}) \rangle. \quad (9.3)$$

Der Erwartungswert $\langle \cdot \rangle$ bezieht sich hierbei auf alle $c = 1, \dots, C$ Versuche. Für C lagen die Werte bei den durchgeführten Untersuchungen zwischen 10^3 und 10^4 . Damit dauert die Bestimmung von \mathbf{S}_{RF} auf einer durchschnittlichen Workstation etwa 0.5 s.

9.1.2 Quadratische Fehlerfortpflanzung

Alternativ zu der eben vorgestellten heuristischen Fehlerfortpflanzung ist die lokale Linearisierung des neuronalen Netzes zum Zweck einer klassischen, quadratischen Fehlerfortpflanzungsrechnung ebenfalls möglich. Die Aktivierung von Ausgabeneuron k des hier verwendeten Perzeptronnetzes mit einer Zwischenschicht läßt sich analytisch schreiben als (vgl. Abschn. 4.2.2)

$$n_k = \sum_{j=0}^{N_J} o_j v_{jk} = \sum_{j=1}^{N_J} f \left(\sum_{i=0}^{N_I} y_i w_{ij} \right) v_{jk} + o_0 v_{0k}. \quad (9.4)$$

Dabei bezeichnet \mathbf{y} den normierten Netzwerk-Eingabevektor, $f(\cdot)$ ist die Transferfunktion, hier der Hyperbeltangens. Die Gewichte zwischen Eingangs- und Zwischenschicht sind in der Matrix $\mathbf{W} = (w)_{ij}$, diejenigen hinter der Zwischenschicht in $\mathbf{V} = (v)_{jk}$ zusammengefasst.² Weiterhin ist $o_0 \equiv 1$ und $y_0 \equiv 1$ (On-Neuronen, vgl. Abschn. 4.1.2). Bei den hier verwendeten, nichtlinearen Ausgabeneuronen gilt dann für die Netzwerkausgabe

$$x_k = f(n_k). \quad (9.5)$$

Das Ozonprofil $\hat{\mathbf{x}}$ ergibt sich durch die inverse *Ausgabenormierung* aus den Werten der Ausgangsneuronen:

$$\hat{x}_k = \nu_{\text{out}}^{-1}(x_k) = \frac{1}{2}(x_k - 1)(\hat{x}_{\text{max},k} - \hat{x}_{\text{min},k}) + \hat{x}_{\text{min},k}. \quad (9.6)$$

Die Minimal- und Maximalwerte von $\hat{\mathbf{x}}$ können auf verschiedene Weisen festgelegt werden, es sei hierzu wiederum auf Abschnitt 8.3 verwiesen. Differenziation des Ozonprofils $\hat{\mathbf{x}}$ nach $\tilde{\mathbf{y}}$ führt unter Verwendung von Kettenregel und Gln. 9.4 bis 9.6 zu

$$\begin{aligned} G_{ki} &\equiv \frac{\partial \hat{x}_k}{\partial \tilde{y}_i} = \frac{\partial \hat{x}_k}{\partial x_k} \frac{\partial x_k}{\partial n_k} \frac{\partial n_k}{\partial y_i} \frac{\partial y_i}{\partial \tilde{y}_i} = \\ &= \frac{1}{2}(\hat{x}_{\text{max},k} - \hat{x}_{\text{min},k}) f'(n_k) \nu'_{\text{in}}(\tilde{y}_i) \sum_{j=1}^{N_J} f' \left(\sum_{i=0}^{N_I} y_i w_{ij} \right) w_{ij} v_{jk}. \end{aligned} \quad (9.7)$$

\mathbf{G} stellt die **neuronale Jacobi-Matrix** dar. Durch die Nichtlinearität des neuronalen Netzes ist sie von den Eingangsdaten abhängig. Ihr entspricht in klassischen Retrievalverfahren die Beitragsfunktionsmatrix $\mathbf{D}_{\mathbf{y}}$. Die Linearisierung des Systems um einen Linearisierungspunkt $\tilde{\mathbf{y}}_a$ erfolgt analog zu Gl. 3.5 gemäß

$$\hat{\mathbf{x}} - \mathbf{R}_{\text{MLP}}(\nu_{\text{in}}(\tilde{\mathbf{y}}_a + \boldsymbol{\epsilon}), \mathbf{w}) = \mathbf{G}\boldsymbol{\epsilon} + \mathcal{O}(\boldsymbol{\epsilon}^2) \quad (9.8)$$

²Wie weiter oben eingeführt, werden alle Gewichte zusammen auch als Vektor \mathbf{w} bezeichnet. Die Aufteilung auf zwei Matrizen dient hier lediglich der Übersichtlichkeit.

Sofern die Messfehler ϵ nicht zu weit vom Linearisierungspunkt wegführen, sollte sich mit dieser Gleichung deren Effekt auf das Retrievalprofil berechnen lassen. Wie im letzten Abschnitt wird ϵ als gaußverteilt und unkorreliert angesetzt. Mit der (diagonalen) Messfehler-Kovarianzmatrix \mathbf{S}_y ergibt sich dann für die Kovarianz des Ausgabefelders der Ausdruck

$$\mathbf{S}_G = \mathbf{G}^T \mathbf{S}_y \mathbf{G}. \quad (9.9)$$

9.1.3 Belernung eines Fehlernetzwerkes

Ein völlig anderer Ansatz zur Abschätzung lokaler Fehler ist die Belernung eines neuronalen Fehlernetzes [Satchwell, 1994; Dybowski, 1997]. Hierzu wird nach der vollständigen Belernung des Retrievalnetzes ein zweites Netz mit denselben Trainingsdaten und demselben Aufbau trainiert, allerdings mit dem Unterschied, dass als Zielvektor $\tilde{\mathbf{t}}$ der Absolutwert des Retrievalfehlers dient:

$$\tilde{\mathbf{t}}^p = |\hat{\mathbf{x}}^p - \mathbf{t}^p| \quad \forall p = 1, \dots, P \quad (9.10)$$

Bei der Mittelwertbildung über alle Trainingsmuster p ergibt sich der mittlere absolute Fehler (MAE) des Trainingsdatensatzes

$$\text{MAE} = \frac{1}{P} \sum_{p=1}^P |\hat{\mathbf{x}}^p - \mathbf{t}^p|. \quad (9.11)$$

Aufgrund der bereits beschriebenen Eigenschaft neuronaler Netze, den Mittelwert des Trainingsdatensatzes fast exakt zu approximieren, entspricht der MAE auch der mittleren Ausgabe des Fehlernetzwerkes nach der Belernung. Anders ausgedrückt verteilt das Fehlernetzwerk den MAE über alle Trainingsmuster, indem es die Eingangsdaten als Maßstab dafür nimmt, wie hoch die Abweichung für ein einzelnes Retrievalprofil sein wird.

Ursprünglich wurde für diese Methode der Zielwert $\tilde{\mathbf{t}}^p = (\hat{\mathbf{x}}^p - \mathbf{t}^p)^2$ vorgeschlagen, woraus sich bei der Mittelung die Varianz anstatt des MAE ergibt. Dieser Ansatz stellte sich allerdings hier als ungeeignet heraus, er lieferte unphysikalische Undulationen der Fehler-schätzung oberhalb von etwa 35 km Höhe, die auf numerische Probleme schließen lassen: Der bereits bestehende Wertebereich für die Ozondichten von drei Größenordnungen Umfang wird bei der Quadrierung auf sechs Größenordnungen erhöht.

9.2 Fallbeispiele und Interpretation

Für die folgenden Betrachtungen wurden zur Vereinfachung unkorrelierte gaußsche Eingangsfehler angenommen. Deren Standardabweichung betrug 2% für GOME-Kanal 3 und 4, 3% für Kanal 2, 5% für Kanal 1B und 10% für Kanal 1A [R. van Oss, persönliche Mitteilung, 2002]³, 1° für die Winkelkoordinaten und 3 K für das UKMO-Temperaturprofil [S. Tellmann, persönliche Mitteilung, 2002, und eigene Untersuchungen].

9.2.1 Einzelprofile

Um die Ergebnisse der drei unterschiedlichen Fehlerschätzungen im konkreten Fall zu vergleichen, wurden drei Kollokationen des Testdatensatzes ausgewählt, die unterschiedliche Atmosphärenzustände repräsentieren: Ein Profil im Ozonmaximum der nördlichen mittleren Breiten, eines aus dem Äquatorialbereich und ein typisches antarktisches Ozonlochprofil.

³Diese Werte enthalten auch einige systematische Kalibrationsfehler, die das neuronale Netz wahrscheinlich zum Teil kompensieren kann. Dafür kommen Fehler aus der für NNORSY verwendeten Doppler-Korrektur und der Interpolation auf das Standardgitter hinzu, die etwa in der selben Größenordnung liegen dürften.

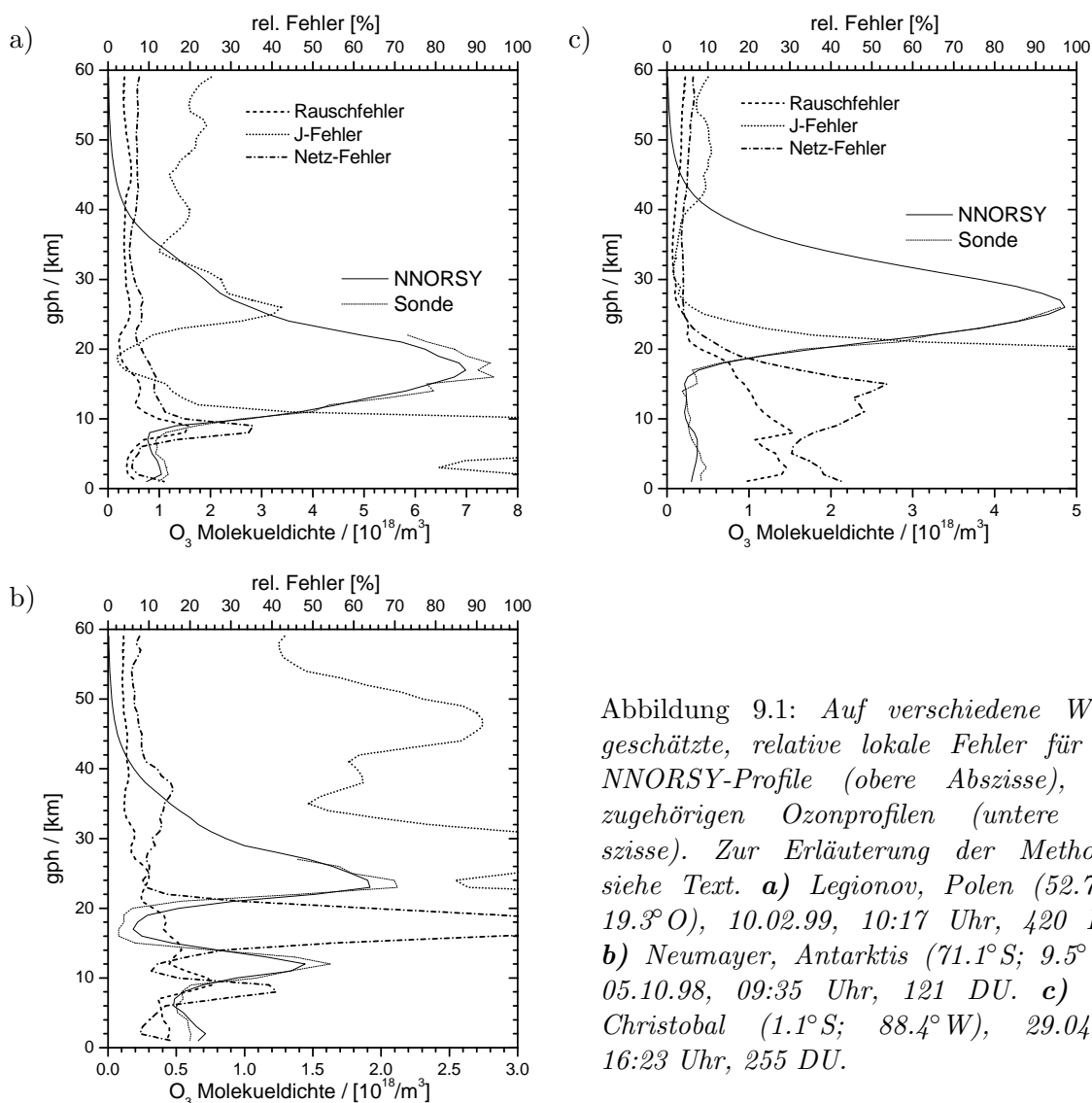


Abbildung 9.1: Auf verschiedene Weise geschätzte, relative lokale Fehler für die NNORSY-Profile (obere Abszisse), mit zugehörigen Ozonprofilen (untere Abszisse). Zur Erläuterung der Methoden siehe Text. **a)** Legionov, Polen (52.7°N ; 19.3°O), 10.02.99, 10:17 Uhr, 420 DU. **b)** Neumayer, Antarktis (71.1°S ; 9.5°W), 05.10.98, 09:35 Uhr, 121 DU. **c)** San Christobal (1.1°S ; 88.4°W), 29.04.99, 16:23 Uhr, 255 DU.

Abb. 9.1a zeigt das Ozonprofil und die relativen Fehler für die erste Situation. Zunächst fällt auf, dass die Fehlerkurven der Rauschmethode und der Belernung eines Fehlernetzwerks (i. F. NN-Methode) qualitativ sehr ähnlich sind, die NN-Fehler sind jedoch durchweg um den Faktor 1.25 bis 2 größer als die Rauschfehler. Dies ist insofern zu erwarten, als die Rauschmethode im Wesentlichen einen *precision*-Fehler liefert, da sie nur zufällige Abweichungen berücksichtigt. Die NN-Methode liefert hingegen auch Informationen über Biases, und schließt alle Fehler ein, die sich im Vergleich zu den Kollokationen ergeben, also auch Unsicherheit der Zielprofile, Kollokationsfehler, Glättungsfehler, etc. Sie entspricht demnach eher dem Fehler im Vergleich zum wahren Ozonprofil (*accuracy*-Fehler); Dieses Thema wird in Kap. 10 genauer ausgeführt. Die höchsten Fehler für beide Methoden zeigen sich wie zu erwarten am Oberrand der Troposphäre, an dem die niedrigen Ozonwerte gekoppelt mit dem steilen Ozongradienten bereits bei geringer vertikaler Verschiebung des Profils eine große relative Abweichung verursachen.

Die aus der quadratischen Fehlerfortpflanzung (J-Methode) resultierende Kurve fällt in der Abbildung stark aus dem Rahmen, die relativen Werte liegen in den meisten Höhenbereichen um ein Vielfaches über denen der anderen Methoden. Lediglich im Bereich des Ozonmaximums ergeben sich sehr geringe Werte. Die Verhältnisse ändern sich auch nicht grundlegend, wenn zu den anderen beiden atmosphärischen Situationen übergegangen wird. Hier erreichen die Fehler aus der J-Methode teilweise mehrere tausend Prozent, was in keinem Verhältnis zu den tatsächlich beobachteten Abweichungen zwischen Retrieval

und Kollokationen steht. Eine weitere Auffälligkeit ist die Diskrepanz von Rauschfehler und NN-Fehler im Bereich des Ozonlochs (Abb. 9.1c). Letzterer erreicht hier Werte von über 100%. Derart große Abweichungen werden tatsächlich beobachtet, da die begrenzte vertikale Auflösung des Retrievals eine gewisse Verschmierung der ozonarmen Region um 17 km bewirkt, wodurch die erhaltenen Ozondichten in dieser Region durchaus um den Faktor 2–3 über denen kollozierter Sonden liegen können. Der tropische Fall b) weist hingegen ein dem Fall a) analoges Verhalten auf, nur dass die relativen Fehler in der Troposphäre aufgrund der geringeren Ozondichte noch höher ausfallen als im Fall a).

Allgemein fällt auf, dass die mit der Rauschmethode erhaltenen Fehler im Vergleich zu den aus der OE-Theorie für ähnliche Fälle berechneten *precision*-Fehlern relativ klein sind [de Beek, 1998]. Selbst der mit 10% Standardabweichung recht großzügig verrauschte Kanal 1A bewirkt offenbar keine starke Profilunsicherheit – dies wird in Abschn. 9.2.3 differenzierter untersucht. Der scheinbare Widerspruch zur Strahlungstransfertheorie kann auch an dieser Stelle mit den speziellen Eigenschaften neuronaler Netze erklärt werden: Zum einen stellen die GOME-Spektren nur *eine* Informationsquelle dar, deren Einfluss der Belernungsalgorithmus mit den anderen Eingabeparametern vergleicht, und bei starkem Rauschen im Trainingsdatensatz entsprechend senkt (Abschn. 9.2.3). Zum anderen bewirkt die Redundanz in den spektralen Eingangsdaten, dass *einzelne* Spektralwerte – gerade im Kanal 1A – nur einen sehr geringen Einfluss auf das Ausgabeprofil haben, und erst in der Summe ihre Information zum Retrieval beitragen (Abschn. 9.3). Dadurch wird das Verfahren sehr robust gegenüber Rauschen mit Mittelwert Null.

9.2.2 Vertikale Fehlerkorrelation

Um der Frage etwas näher zu kommen, warum die J-Methode so unrealistisch hohe Fehler liefert, sollte zunächst untersucht werden, ob die in linearer Näherung berechnete Jacobi-Matrix das Fehlerverhalten überhaupt zutreffend beschreiben könnte. Zu diesem Zweck wurden für die Fälle a) und c) aus Abb. 9.1 Korrelationsmatrizen der Fehler ermittelt, und zwar einerseits aus den Spalten der Jacobi-Matrix – also den Beitragsfunktionen für jeden Eingangswert – und andererseits direkt aus den im Rahmen der Rauschmethode berechneten 10^4 verrauschten Retrievals (Abb. 9.2). Diese beiden Korrelationsgrößen sind zwar nicht exakt vergleichbar, sollten bei diagonaler Kovarianzmatrix des Eingangsrauschens aber zumindest ähnliche Ergebnisse produzieren: Beide Größen drücken aus, wie der Fehler, also die Abweichung vom unverrauschten Retrievalprofil \hat{x} , in den Höhenschichten korreliert ist, wenn statt Eingangsneuron A das Eingangsneuron B gestört wird.

Generell besteht in Abb. 9.2 eine weitgehende Übereinstimmung der berechneten Fehlerkorrelationen im Bereich der Diagonalen. Gleichzeitig existieren jedoch beträchtliche Korrelationen bzw. Antikorrelationen außerhalb derselben. Dies kann als Hinweis auf den in Abschn. 8.1.3 diskutierten Effekt verstanden werden, dass nämlich das System dazu neigt, bei Verschiebungen der Ozonpeakhöhe gekoppelte Änderungen ober- und unterhalb des Peaks vorzunehmen. Die extradiagonalen Korrelationen oberhalb von 25 km Höhe stimmen für beide Matrizenhälften recht gut überein, darunter gibt es im Fall a) gewisse Differenzen; Die Antikorrelation zwischen Fehlern unter- und oberhalb des Ozonpeaks scheint bei der J-Methode stärker ausgeprägt zu sein als bei der Rauschmethode. In weniger extremen atmosphärischen Situationen mit mittlerem Gesamt ozonegehalt, sowie im Mittel aller Testkollokationen (nicht dargestellt) ähnelt der Grad der Übereinstimmung eher dem Fall b). Die Jacobi-Matrix an sich scheint also die Zusammenhänge zwischen den Höhenschichten recht gut darstellen zu können. Wieso es trotzdem zu den beobachteten, hohen Fehlerschätzungen kommen kann, wird im nächsten Abschnitt (9.3) diskutiert werden.

Eine wesentliche Gemeinsamkeit der beiden Matrizenhälften besteht in den meisten Fällen in der Abnahme des Korrelationsgrades um die Diagonale herum. Diese kann als ein

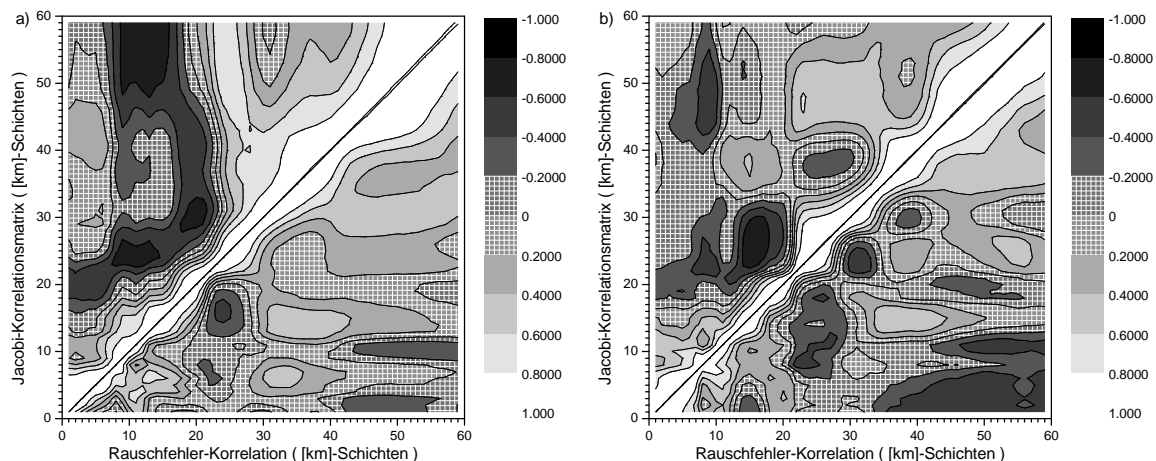


Abbildung 9.2: Kombinierte Fehlerkorrelationsmatrizen für die Fälle a) und b) aus Abb. 9.1. Die obere linke Hälfte der dargestellten Matrizen entstand aus der Jacobi-Matrix, die untere rechte Hälfte direkt aus den Rauschfehlern.

Maß für die Auflösung des Retrievals angesehen werden, da sich kleinräumige Strukturen in der Vertikalen offensichtlich nur dann unabhängig bestimmen lassen, wenn ihre Fehler nicht zu stark korreliert sind. Wie aus Abb. 9.2 zu entnehmen, ist der Bereich zwischen den Konturen für $c = 0.6$ und $c = 0.8$ i. A. sehr schmal, die Korrelation nimmt hier also steil ab. In Abb. 9.3 ist daher die über den Testdatensatz gemittelte Korrelationslänge aufgetragen, wobei die Fehler bei einem Korrelationsgrad von $c > 0.7$ als korreliert galten.

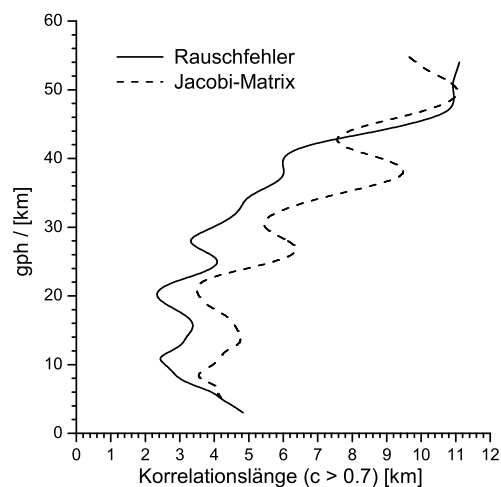


Abbildung 9.3: Korrelationslänge l der Profilverfehler für einen Korrelationsgrad von $c > 0.7$. Die Größe l ist hier als Breite der $c=0.7$ Kontur (vgl. Abb. 9.2) definiert, gemessen senkrecht zur Diagonalen und gemittelt über den gesamten Testdatensatz. Sie wurde für die Rauschfehler-Korrelationsmatrix und die aus der Jacobi-Matrix ermittelten Korrelationen separat berechnet.

Im mittleren Höhenbereich von etwa 10–30 km ergeben sich Korrelationslängen von 2.5–6.5 km, wobei die Jacobi-Matrix systematisch höhere Werte liefert als die Rauschfehlerkorrelation. Dies stimmt recht gut mit der in Abschn. 7.4.3 empirisch ermittelten Höhenauflösung überein; Die Differenzen in der Kurvenform im Vergleich zu Abb. 7.13 resultieren vermutlich zum einen aus Artefakten jener empirischen Methode (s. d.), zum anderen aus den hier vernachlässigten Korrelationen zwischen weit entfernten Höhenschichten, die sich im Zweifelsfall leicht negativ auf die Vertikalauflösung auswirken dürften. Insofern kann der in Abschn. 7.4.3 ermittelte Vertikalauflösungs-Schätzwert von etwa 4 km für die mittleren Höhenschichten als bestätigt gelten. Umgekehrt läßt damit Abb. 9.3 den Schluss zu, dass NNORSY oberhalb von 30 km stark an vertikaler Auflösung verliert, bis bei 50 km Höhe nurmehr Strukturen von 11–12 km Dicke aufgelöst werden. Allerdings kommen hier in

den kollozierten Okkultations-Profilen auch keine feineren Strukturen vor — die Korrelationslänge spiegelt also bis zu einem gewissen Grad nur die Kovarianz der Trainingsprofile wider. In der Troposphäre scheint die NNORSY-Vertikalaufösung ebenfalls etwas abzunehmen, aber hier zeigt der diagonalnahe Bereich der Korrelationsmatrizen viele und hoch variable Feinstrukturen (vgl. Abb. 9.2), so dass die Werte mit Vorsicht zu behandeln sind.

9.2.3 Rauschfehler-Sensitivitätsanalyse

Wenden wir uns nun zunächst wieder der Rauschfehlermethode zu. Um die einzelnen Fehlerbeiträge zu differenzieren, wurden die Komponenten des Eingabevektors blockweise verrauscht und die resultierenden Fehlerprofile bestimmt. Zu beachten ist hierbei, dass es wenig Sinn macht, *einzelne* Eingangsneuronen mit Rauschen zu belegen – dies würde wiederum nur die Beitragsfunktionen der Jacobi-Matrix liefern. Nur durch Verrauschen mehrerer verwandter Eingangsneuronen läßt sich die effektive Netzwerk-Antwort auf das Rauschen bestimmen.

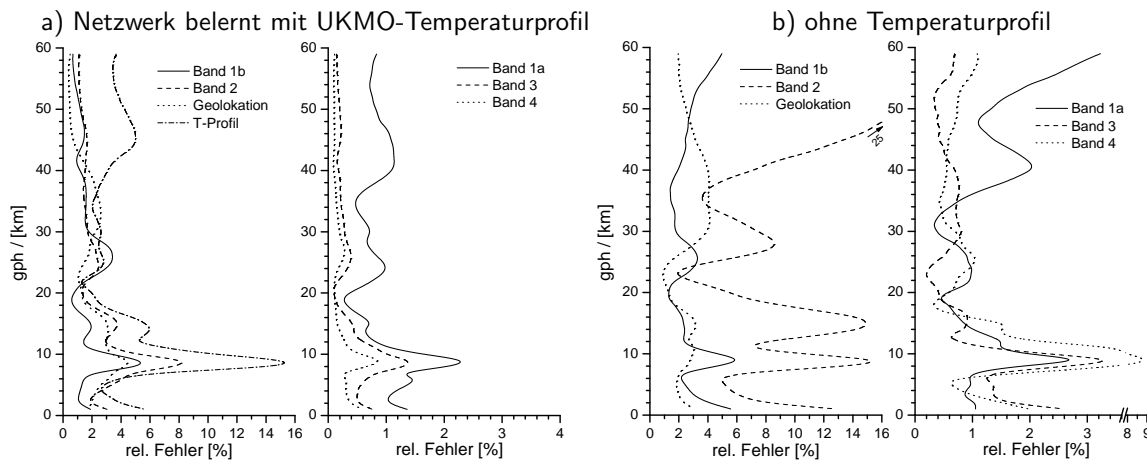


Abbildung 9.4: Abhängigkeit der Rauschfehler von den Eingangsdaten. Die innerhalb der GOME-Bänder verwendeten Spektralbereiche sind Tabelle 7.2 zu entnehmen.

Abb. 9.4a zeigt das Resultat für den Fall a) aus Abb. 9.1. Leider läßt sich offenbar auch auf diese Weise keine eindeutige Höhenzuordnung zwischen den GOME-Spektralbändern und den Ausgabefeldern ermitteln, wie schon in Abschn. 8.1 festgestellt. Die Fehlerprofile für die verschiedenen GOME-Bänder ähneln sich stark, wobei sich die Hauptbeiträge in allen Fällen um die Tropopause herum sammeln. Zumindest bei Band 1A/B scheint jedoch das Verhältnis der Fläche unter der Kurve oberhalb von 20 km zu derjenigen unterhalb von 20 km größer zu sein als bei den anderen Bändern, was auf stärkeren Einfluss des Kurzwelligen auf die Ozonverteilung in größeren Höhen hindeutet, wie theoretisch zu erwarten. Auch durch feinere Differenzierung der zu verrauschenden Eingangswerte, etwa in 5 nm Intervalle, ließ sich kein besserer Nachweis der Höhenzuordnung von spektraler Information finden (nicht dargestellt). Mögliche Gründe für dieses Verhalten werden in Abschn. 9.3 diskutiert.

Interessanterweise liefert das UKMO-Temperaturprofil den größten Beitrag zu den Fehlern. Dies deutet wiederum auf den starken Zusammenhang zwischen Temperatur- und Ozonverteilung hin. Auch hier ist die Höhenzuordnung nicht eindeutig; Obwohl nur Temperaturen bis 40 km Höhe in das Netz eingehen, ist der Fehlerbeitrag auch in größeren Höhen stärker als der aller anderen Eingangsdatengruppen. Um zu überprüfen, ob die den Fehler verursachende Temperatur-Unsicherheit mit 3°C vielleicht nur unrealistisch hoch angesetzt wurde, zeigt Abb. 9.4b dieselbe Situation für eine Belernen, bei der die UKMO-Daten

nicht als Eingabe verwendet wurden, die Konfiguration aber ansonsten gleich blieb. Dadurch wird die Form der Fehlerprofile uneinheitlicher, jedoch nehmen die Fehler generell zu. Das neuronale Netz versucht in diesem Fall, mehr Information aus den Spektren zu gewinnen, die jedoch ein viel schlechteres S/N-Verhältnis aufweisen als das Temperaturprofil. Für Band 1A nehmen die Fehler in der Höhe überproportional zu, während die O₂ Banden in Kanal 4 eine deutliche Zunahme um die Tropopausengrenze herum zeigen, was wieder auf die gemäß Strahlungstransferrechnung zu erwartende Höhenabhängigkeit der spektralen Information hindeutet. Allerdings nimmt auch der Fehler aus Kanal 2 in der Höhe stark zu, obwohl hier fast keine Strahlung aus den großen Höhen beim Sensor ankommt.

Lernt das neuronale Netz aus den Temperaturprofilen einen physikalischen oder einen statistischen Zusammenhang zwischen Temperatur und Ozon? Diese Frage läßt sich leider hier nicht eindeutig beantworten. Selbstverständlich bestehen starke physikalische Zusammenhänge dahingehend, dass einerseits die Temperaturschichtung der Stratosphäre wesentlich durch die Absorption von solarer Strahlung an Ozonmolekülen hervorgerufen wird [z. B. *Fabian*, 1992], und andererseits die massive Ozonzerstörung über der Antarktis wie in Abschn. 1.2 erläutert erst durch sehr niedrige Temperaturen ermöglicht wird [*Spichtinger-Radowsky*, 2001, und Referenzen darin]. Untersuchungen von *Spänkuch & Schulz* [1995, 1998] haben jedoch auch gezeigt, dass es schon mit multipler linearer Regression – zumindest lokal – möglich ist, gute Ozonprofilschätzungen aus der Temperaturverteilung zu gewinnen, ohne irgendeine Art von Strahlungstransportrechnung durchzuführen.

Zusammenfassend kann an dieser Stelle auf jeden Fall festgehalten werden, dass die Verwendung des UKMO-Temperaturprofils das NNORSY-Verfahren stabilisiert. Die Temperaturdaten wirken nach bisherigem Verständnis trotz ihrer in allen vier Dimensionen relativ schlechten Auflösung als zusätzliche *a priori* Information, anhand derer sich das neuronale Netz unter Einbeziehung der klimatologischen Koordinaten (vgl. Abschn. 8.1.1) eine Art *a priori* Ozonprofil konstruieren kann, welches dann durch die GOME-Spektren modifiziert wird. Ohne Temperaturdaten kann das Netz nur ein klimatologisches Ozonprofil als Ausgangspunkt verwenden, und die Korrekturen durch die Spektralwerte werden größer und unsicherer, wodurch die Fehler ansteigen. In beiden Fällen läßt sich jedoch der Zusammenhang zwischen spektraler Wellenlänge und Ozon-Höhenschicht nicht eindeutig nachweisen.

9.3 Diskussion der Jacobi-Matrix

Es konnte bisher nicht geklärt werden, wo die Höheninformation verloren geht, und warum die mittels Jacobi-Matrix propagierten Fehler so groß sind, obwohl die aus ihr ermittelte Korrelation der Ozonprofilfehler gut mit der Empirie übereinstimmt. Zur Beantwortung der letzteren Frage ist in Abb. 9.5 die über den gesamten Testdatensatz gemittelte und normierte Jacobi-Matrix $\bar{\mathbf{G}}$ dargestellt [*Aires et al.*, 2001], für die gilt

$$\bar{\mathbf{G}} = (\boldsymbol{\sigma}_x^{-1})^T \left(\frac{1}{P} \sum_{p=1}^P \mathbf{G}_p \right) \boldsymbol{\sigma}_y. \quad (9.12)$$

Die Vektoren der Standardabweichungen $\boldsymbol{\sigma}_x$ und $\boldsymbol{\sigma}_y$ von Ozonprofilen bzw. Eingangsdaten wurden hierbei aus dem Testdatensatz berechnet. Jede Spalte der Matrix $\bar{\mathbf{G}}$ ist also mit der Eingangsvariabilität gewichtet. Gleichzeitig sind die Zeilen auf die Ausgangsvariabilität normiert, damit die Werte zwischen den Höhenschichten vergleichbar werden. Somit kann $\bar{\mathbf{G}}$ als der über den Testdatensatz gemittelte Einfluss interpretiert werden, den jeder Eingangswert auf das Ausgabeprofil hat. Jede Spalte von $\bar{\mathbf{G}}$ enthält eine normierte, dimensionslose Beitragsfunktion.

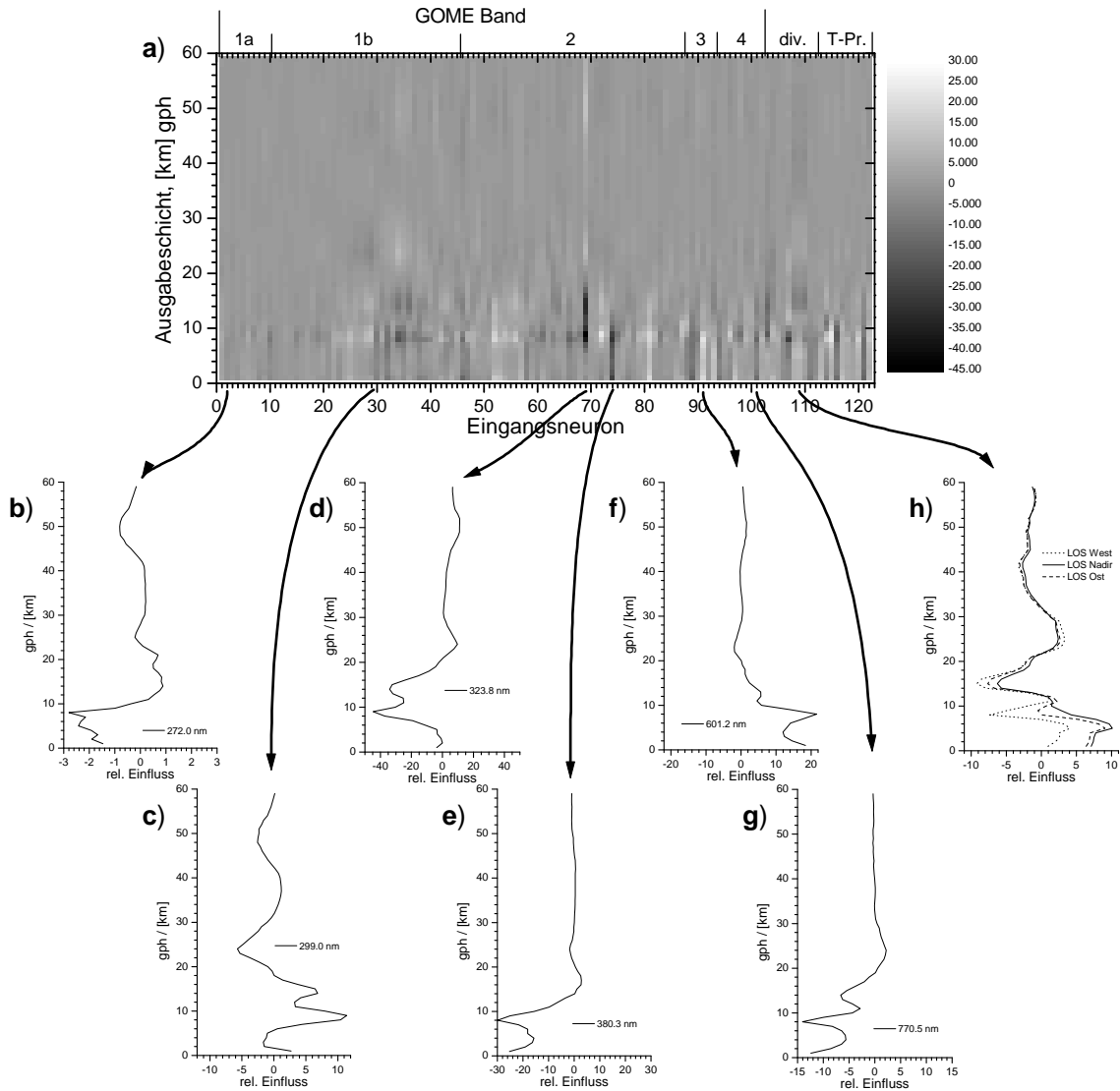


Abbildung 9.5: Normierte, über den Testdatensatz gemittelte Jacobi-Matrix des durch das neuronale Netz repräsentierten, inversen Strahlungstransfermodells. Die innerhalb der GOME-Bänder verwendeten Spektralbereiche sind Tabelle 7.2 zu entnehmen, „div.“ meint diverse Eingabeparameter, „T-Pr.“ das UKMO-Temperaturprofil. Durch die Normierung mit Ein- und Ausgabestandardabweichung sind alle Zahlenwerte dimensionslos und relativ zueinander direkt vergleichbar. Sie spiegeln den mittleren Einfluss eines Eingabeparameters auf das Ozonprofil wider. Dies erlaubt z. B. die Aussagen, dass sich fast alle Eingangsdaten einzeln betrachtet stärker auf das Ozonprofil unterhalb von 20–25 km auswirken, als oberhalb davon, oder dass der GOME-Spektralwert bei 380.3 nm (Teilbild e) für die Höhe 8 km etwa einen zehnmal größeren Einfluss hat als derjenige bei 272 nm (b). Der niedrigste in der Matrix vorkommende Wert liegt bei -44.7, der höchste bei 22.4. Wie in Teilbild h) dargestellt und im Text diskutiert, bewirkt die Ähnlichkeit der Kurven jedoch bei korrelierten Eingabeparametern Kompensations- bzw. Verstärkungseffekte.

Ein erster Blick auf Abb. 9.5a zeigt, dass die Beitragsfunktionen im Höhenbereich unter 20 km am variabelsten sind, und dass die Beiträge beginnend mit dem kurzwelligen Drittel von GOME Band 1B ($\lambda \lesssim 295$ nm) relativ schwach werden. Weiterhin fällt auf, dass in vielen Bereichen helle und dunkle Streifen direkt beieinander liegen, und sich manche senkrechten Muster an verschiedenen Stellen der Matrix in ähnlicher Form finden. Die Graphen b)–h) greifen verschiedene Beitragsfunktionen heraus. Die hier auftretenden Muster sind als typisch für die Methode anzusehen; Sie treten immer wieder in ähnlicher Form auf, auch wenn die Netzwerkconfiguration wie z. B. in Abschn. 8.1 beschrieben verändert wird. Um den Rahmen nicht zu sprengen beschränkt sich die folgende Diskussion daher auf den Standardlauf (Tab. 7.2).

Wie schon bei der Rauschfehleranalyse festgestellt (Abb. 9.4), liegt für fast alle Beitragsfunktionen ein absolutes Extremum bei etwa 9 km vor, unabhängig vom Spektralbereich. Davon abgesehen sind nur wenige Strukturen erkennbar, die eine Höhenzuordnung des Einflusses zuließen, so etwa das Extremum bei 50 km in b) und dasjenige bei 37 km in c), die in etwa einer gemittelten Höhe der theoretischen Gewichtsfunktions-Minima für diesen Spektralbereich entsprechen könnten [Hoogen, 1998; de Beek, 1998], sowie die relative Strukturlosigkeit der Kurven oberhalb von 20 km für die Fälle e)–g), die mit dem erwarteten Einfluss von atmosphärischem Fenster, Chappuis-Bande und O₂ A-Bande, respektive, recht gut übereinstimmen. Allerdings zeigen sich generell oberhalb von 30 km Höhe nur sehr breite und flache Strukturen in den Beitragsfunktionen, was auch die großen Fehlerkorrelationslängen aus Abb. 9.3 erklärt. Wie bereits erwähnt spiegelt sich hier der Aufbau des Trainingsdatensatzes bzw. die Struktur des Ozonfeldes wider, nicht eine Beschränkung der neuronalen Netze. Es gibt keinen Grund anzunehmen, dass die Methode sich nicht auch hier die theoretisch in den GOME-Spektren enthaltene Information zu Nutze machen würde, wenn entsprechend kleinräumige, vertikale Strukturen in den Trainingsdaten auftauchten.

Das Hauptproblem bei der Interpretation der Matrix $\bar{\mathbf{G}}$ ist, dass sich viele Effekte gegenseitig kompensieren: Ein großer negativer Beitrag bei einer Wellenlänge kann z. B. durch 5 kleinere, positive Beiträge bei anderen Wellenlängen ausgeglichen werden, die nicht notwendigerweise in der Nähe oder überhaupt beieinander liegen müssen, solange sie nur innerhalb des Trainingsdatensatzes entsprechend korreliert sind. Das Vorzeichen der Beitragsfunktionen ist daher ebenfalls nur schwer zu interpretieren. Dasselbe gilt für die nicht-spektralen Eingangsparameter. Da die Kompensation der Beiträge verschiedener Eingangswerte in der klassischen Fehlerfortpflanzungsrechnung (Gl. 9.9) nicht berücksichtigt wird, ist klar, dass die erhaltenen Fehlerschätzungen viel zu hoch liegen. Im Sinne der linearen Fehleranalyse von Rodgers [2000] würde das NNORSY-Verfahren daher vermutlich als instabil gelten, die Praxis zeigt jedoch, dass es durchaus stabile Ergebnisse liefert. Aires et al. [2001] weisen in diesem Zusammenhang darauf hin, dass selbst bei bekannten Eingangskovarianzmatrizen die lineare Fehleranalyse scheitern kann, weil die Beschreibung des Zusammenspiels zweier MLP-Eingangsparameter intern mitunter viel komplexer ist, als sich dies durch einen einzigen Zahlenwert (der Kovarianz) im Rahmen der linearen Theorie ausdrücken ließe.

Ein einfaches Beispiel für die beobachteten Kompensationseffekte ist in Abb. 9.5h anhand der line-of-sight Flags sehr deutlich zu sehen. Es ist klar, dass diese Flags stark korreliert sind, da immer eines den Wert 1 und die anderen beiden den Wert -1 erhalten (vgl. Abschn. 5.3). Ihr effektiver Einfluss auf das Retrieval setzt sich also immer aus zwei negativen und einer positiven Beitragsfunktion zusammen. Da sich die Funktionen oberhalb von 10 km Höhe stark ähneln, kompensiert sich hier ein negativer und ein positiver Beitrag fast vollständig, während die quadratische Fehlerfortpflanzung durch die gleiche Gewichtung aller drei einen um den Faktor $\sqrt{3}$ zu hohen Einfluss auf den Fehler annehmen würde. Wie man sich leicht klarmacht, ist die Überschätzung noch größer, wenn statt dessen ein großer Beitrag durch viele kleine kompensiert wird, wie dies bei den Spektralwerten der Fall zu sein scheint.

Fazit

Durch verschiedene Methoden der lokalen Fehlerabschätzung und Sensitivitätsanalyse wurde versucht, die inneren Zusammenhänge der verwendeten neuronalen Netze zu erhellen und ihre Ausgabe besser zu charakterisieren. Dabei bestätigte sich der in Abschn. 8.1.3 andiskutierte Effekt, dass die vorliegende Methode nur eine äußerst grobe Beschreibung des physikalischen Zusammenhangs zwischen Wellenlänge der Eingangs-Spektralwerte und Höhenverteilung des Ozons in der Ausgabe zulässt. Gleichzeitig stößt die lineare Abschätzung der Ausgangsfehler aufgrund der komplizierten Abhängigkeiten von Eingangswerten und Beitragsfunktionen schnell an ihre Grenzen. Damit ist klar, dass sich das NNORSY-Verfahren nur sehr eingeschränkt und mit relativ großem Aufwand für die Gewinnung neuer physikalischer Erkenntnisse analysieren und nutzen lässt. Die Fehlercharakterisierung mittels verrauschter Eingangsdaten sowie durch die Belernung eines Fehlernetzwerkes liefert jedoch durchaus brauchbare Werte und ist für den operationellen Betrieb gut geeignet. Für den allgemeinen Gebrauch kann die mittels NN-Methode gewonnene Fehlerabschätzung empfohlen werden, da sie nicht von Schätzungen des Eingangsrauschens abhängt und im globalen Mittel die Fehlerstatistik der Kollokationsdatensätze widerspiegelt. Zusätzlich könnte noch auf einfache Weise der Modellfehler, der in Kapitel 10 quantifiziert wird, angegeben oder hineingerechnet werden. Damit kann das gesetzte Ziel, die Schwächen der lokalen Fehlercharakterisierung bei neuronalen Netzen zu überwinden, als erreicht gelten.

Kapitel 10

Globales Fehlerbudget

10.1 Übersicht und Strategie

In Abschnitt 4.3 wurde bereits angedeutet, dass die Belernung von neuronalen Netzen mit quadratischer Fehlerfunktion auf die globale Minimierung eines einzigen Fehlerparameters hinausläuft, nämlich des RMSE der Trainingskollokationen. Welche Komponenten dieser Fehler enthält, wie er sich auf die Höhenschichten verteilt und in welchem Maße er von den Eingangsparametern abhängig ist, soll im Folgenden eruiert werden.

Bei Betrachtung des Beobachtungssystems NNORSY-GOME lassen sich verschiedene Beiträge zum Gesamtfehler identifizieren:

1. Rauschfehler der Eingangsdaten. Diese pflanzen sich naturgemäß in die Profile fort und wurden in Kapitel 9 ausführlich diskutiert.
2. Unsicherheit der kollokierten Ozonprofile. Die für die Belernung verwendeten Messungen besitzen endliche Genauigkeit (Abschn. 2.1 und 2.2) sowie ggf. systematische Abweichungen untereinander (Abschn. 7.1.3).
3. Kollokationsfehler. Dadurch dass die verschiedenen Messinstrumente i. A. unterschiedliche Luftmassen betrachten, entstehen Unsicherheiten, die von der Variabilität des Ozonfeldes abhängen.
4. Glättungsfehler. Die verschiedenen expliziten und systeminhärenten Regularisierungsmechanismen führen zu einer reduzierten Vertikalauflösung des Retrievalsystems (Abschn. 7.4.3).
5. Modellfehler. Der Belernungsalgorithmus findet meist nicht das globale Minimum der Fehlerfunktion, daher variieren die Ergebnisse zwischen einzelnen Trainingsläufen (Kapitel 8).

Bislang wurden im Wesentlichen zwei Arten von Fehlern diskutiert: Einerseits die *precision*-Fehler einzelner Retrievalprofile (Kap. 9), also der Beitrag 1. Andererseits die Fehlerstatistik auf den Kollokationsdatensätzen, welche die Beiträge 1 bis 4 umfaßt.

Leider macht es die kompakte Form des NNORSY-Retrievals sehr schwierig, die einzelnen Fehlerbeiträge zu isolieren. Anders als z. B. bei OE ist die aktive Fehlergewichtung in Form von Kovarianzmatrizen nicht Teil des Retrievals, sondern wird vom Belernungsprozess implizit selber vorgenommen. Insbesondere ist nicht immer klar, wie sich die einzelnen Fehlerbeiträge addieren: Die Diskrepanzen zwischen quadratischer Fehlerfortpflanzung und Rauschexperimenten in Kap. 9 machen dies deutlich.

Als Strategie für die Aufstellung eines gesamten Fehlerbudgets wird darum folgende Vorgehensweise verfolgt: Zunächst werden die Fehlerbeiträge 1–4 so gut es geht separat

abgeschätzt. Da anzunehmen ist, dass sich diese Fehler zum Teil gegenseitig kompensieren, wird anschließend durch quadratische Addition der Fehler ein **Referenz-Fehlerbudget** erstellt. Der Vergleich mit den bei Trainings- und Testdaten tatsächlich beobachteten Fehlern liefert dann eine Aussage darüber, wie gut der Retrievalalgorithmus die vorhandene Information ausnutzt. Liegt dessen Fehler über der Referenz, so hat das Retrieval offenbar Schwierigkeiten, die theoretisch verfügbare Information aus den Trainingsdaten zu extrahieren. Liegt er dagegen wesentlich darunter, spricht das für eine weitgehende interne Fehlerkompensation, und damit für ein gutes Funktionieren der Methode. Wenn das Retrieval generell ähnlich gut oder besser als die Referenz arbeitet, erlaubt die Kombination der Kollokations-Fehlerstatistiken mit den Abschätzungen für Beitrag 5 die Angabe eines **globalen Fehlerbudgets**.

In der hier vorgestellten Analyse nicht enthalten ist der Beitrag des Repräsentativitätsfehlers: Zwar haben die bisherigen Ergebnisse gezeigt, dass NNORSY die aus dem Trainingsdatensatz erlernten Zusammenhänge sehr gut auf den Testdatensatz übertragen kann, aber damit ist noch nicht sichergestellt, wie gut die Methode in Regionen von Messwert- und Zustandsraum arbeitet, die von keinem dieser Datensätze abgedeckt werden. Ansatzweise wird dieses Thema in den Fallstudien des Teils IV dieser Arbeit untersucht. Eine formale Analyse der Repräsentativitätsfehler ist möglich und wird bei kleineren Problemstellungen regelmäßig durchgeführt bzw. bereits bei der Belernung berücksichtigt [z. B. *Bishop*, 1995a; *Plutowski*, 1996, und Referenzen darin]. Im vorliegenden Fall würde sie jedoch entweder einen weiteren globalen, völlig unabhängigen Ozondatensatz, oder aber die Belernung eines großen Ensembles von Netzen mit jeweils unterschiedlicher Zusammenstellung von Trainings- und Testdaten erfordern. Beide Ansätze würden den Rahmen dieser Arbeit sprengen. —

Angesichts der in Abschnitt 7.4 deutlich gewordenen Variation der Genauigkeit für verschiedene Breitenregionen werden die folgenden Fehlerabschätzungen für die Tropen (30°S–30°N), mittlere bis hohe nördliche Breiten (30°N–90°N) und mittlere bis hohe südliche Breiten (30°S–90°S) separat durchgeführt. In der Höhe werden die Schichten 0–5 km, 6–14 km, 15–25 km, 26–40 km und 41–60 km betrachtet.

10.2 Trainingsprofil-Fehler

Die Abschätzung der Trainingsprofil-Fehler erfolgt in drei Schritten. Zunächst wird die Präzision der Trainingskollokationen abgeschätzt, dann die systematischen Biases zwischen den einzelnen Datenquellen betrachtet und schließlich alles unter Berücksichtigung der Anzahl der jeweiligen Messwerte zusammengeführt.

10.2.1 Zufällige Fehler

Ausgehend von den teilweise schon in Kapitel 2 zitierten Untersuchungen aus der Literatur und unter zusätzlicher Berücksichtigung der in Abschnitt 7.1.3 beschriebenen Datensatz-Homogenisierungsmaßnahmen fasst Tabelle 10.1 die geschätzte Präzision der Sensorentypen ECC-Ozonsonden, andere Ozonsonden, HALOE, SAGE II und POAM III zusammen.

Ozonsonden

Die Genauigkeit der Sondenmessungen ist höhenabhängig und je nach Sondentyp (vgl. Abschn. 2.1.1) unterschiedlich. In der Stratosphäre kann die Genauigkeit der ECC-Sonden mit etwa 5% angenommen werden [*Komhyr et al.*, 1995; *Hilsenrath et al.*, 1986], die der anderen Typen ist minimal schlechter [*Beekmann et al.*, 1994; *SPARC*, 1998].

Tabelle 10.1: *Geschätzte Präzision der kollokierten und homogenisierten Ozonprofil-*
daten des Trainingsdatensatzes in Prozent, nach Sensorentyp.

GPH / [km]	Sensor	Geografische Breite		
		30°S-90°S	30°S-30°N	30°N-90°N
41–60	nicht-ECC	–	–	–
	ECC	–	–	–
	SAGE II	10	10	10
	HALOE	8	8	8
	POAM III	5	–	5
26–40	nicht-ECC	20	20	20
	ECC	12	12	12
	SAGE II	6	6	6
	HALOE	10	10	10
	POAM III	5	–	5
15–25	nicht-ECC	7	7	7
	ECC	5	7	5
	SAGE II	10	10	10
	HALOE	15	15	15
	POAM III	10	–	10
6–14	nicht-ECC	12	14	12
	ECC	7	10	7
	SAGE II	15	15	15
	HALOE	20	–	20
	POAM III	20	–	20
0–5	nicht-ECC	10	10	10
	ECC	6	8	6
	SAGE II	–	–	–
	HALOE	–	–	–
	POAM III	–	–	–

Generell nimmt die Präzision der Ozonwerte ab einer Höhe von etwa 28 km stark ab, besonders bei nicht-ECC-Sonden [Kerr *et al.*, 1994], und die Sondentypen zeigen kein einheitliches Verhalten mehr [SPARC, 1998]. Tendenziell unterschätzen Brewer-Mast Sonden die Ozonwerte über 28 km um 5–15%, während Japanische Sonden zu einer Überschätzung um bis zu 10% neigen [Fujimoto *et al.*, 1998]. Da die nicht-ECC Sonden i. F. nicht getrennt betrachtet werden, spiegeln die hohen Fehler in der 26–40 km-Schicht ausnahmsweise diese systematischen Abweichungen wider.

In der untersten Stratosphäre und Troposphäre sind die Ozonwerte sehr niedrig, so dass die Sonden ihrer Nachweisgrenze näher kommen und die Resultate verschiedener Vergleichskampagnen zwischen 1970 und 1991 teilweise beträchtlich differieren [Logan, 1999]. Neuere Labormessungen [Smit *et al.*, 1998] ergaben, dass die systematischen Abweichungen zwischen Sondentypen innerhalb von $\pm 5\%$ liegen, während die Präzision von ECC-Sonden mit 5%, die der anderen Typen mit 10–15% angegeben wird. SPARC [1998] schätzen die Präzision respektive mit 5–10% bzw. 10–20% ab. Tendenziell sind die Fehler um so größer, je geringer die Ozonkonzentration ist, daher wurden die Fehler für die besonders ozonarme tropische Troposphäre etwas höher angesetzt als für SH und NH.

In der sogenannten planetaren Grenzschicht, dem untersten Kilometer der Troposphäre, macht sich insbesondere in urbanen Gebieten der photochemische Tagesgang des Ozons

durch die Abhängigkeit von der Sonneneinstrahlung (vgl. Abschnitt 1.2) bemerkbar, was zu stark schwankenden Sondenprofilwerten führt [Fortuin & Kelder, 1998]. Werte unterhalb von 500 m wurden daher in dieser Arbeit nicht verwendet.

SAGE II

Während systematische Abweichungen der SAGE II Ozonprofile bereits in Abschn. 2.2.2 andiskutiert wurden, sind zuverlässige Angaben über die Präzision aktueller, einzelner Profile nur schwer zu finden, da die meisten Studien sich mit Langzeittrends beschäftigen. Gemäß Cunnold *et al.* [1989] weisen die SAGE v5.93 Ozonprofile zwischen 24–48 km zufällige Fehler von etwa 5–7% auf, darüber steigen die Fehler auf über 20% (bei 60 km) an. Rusch *et al.* [1998] führten im Rahmen der Entwicklung von v5.96 eine Monte-Carlo-Fehleranalyse durch, bei der sie zufällige Fehler von etwa 10% im Bereich 15–19 km fanden, darüber sanken die Fehler schnell auf 3% bei 25 km ab. Höhen über 30 km werden jedoch nicht diskutiert.

Verbesserungen beim Übergang zu v6.00 und v6.10 beschränken sich größtenteils auf die Reduzierung systematischer Fehler des Ozonprofils unterhalb von 20 km und oberhalb von 50 km (vgl. Abschn. 2.2.2). Eine stichprobenartige Analyse der Fehlerangaben aus dem v6.10 Datenprodukt – nach den hier durchgeführten Homogenisierungsmaßnahmen (Abschn. 7.1.3) – ergab Unsicherheiten des Ozonprofils von etwa 9% (bei 6–14 km), 2,5% (15–25 km) und zwischen 1–2% oberhalb von 25 km Höhe. Diese Werte werden laut Online-Dokumentation auf der SAGE II Website aus der Standardabweichung der – mit einem 0,5 km Medianfilter geglätteten – Einzelmessungen (Sonnenscans) für jede Höhe gewonnen. Die Fehler sind aufgrund der geänderten Berechnung bemerkenswerterweise um einen Faktor von 5–20 geringer als die in v5.96 angegebenen. Sie erscheinen jedoch im Vergleich zu den Ozonsonden sowie HALOE und POAM als unrealistisch klein. Die Werte in der Tabelle sind daher als Kompromiß zwischen den verschiedenen zitierten Untersuchungen zu betrachten.

HALOE

Die in [Bruehl *et al.*, 1996] durchgeführte, umfangreiche Analyse der zufälligen Fehler der HALOE Ozonprofile V17 hat im Wesentlichen auch noch für die hier verwendeten V19 Daten Gültigkeit, lediglich unterhalb von 20 km ergaben sich durch Korrekturen der Höhenzuordnung signifikante Verbesserungen [Bhatt *et al.*, 1999]. Wie in Abschn. 2.2.3 besprochen, sind hingegen in den dem Datenprodukt beiliegenden Fehlerangaben nur Messrauschen und Aerosoleffekte enthalten. Da außerdem die hier verwendete Screeningprozedur für Ausreißer nicht perfekt ist, werden die Präzisionsfehler im Bereich 6–14 km mit 20% recht hoch angesetzt. Die anderen Höhenschichten folgen den Angaben von Bruehl *et al.* [1996], mit einem kleinen Zuschlag für nicht berücksichtigte Glättungseffekte aufgrund der nominellen HALOE Höhenauflösung, die mit 2,3 km schlechter ist als die des NNORSY-Retrievalgitters (1 km).

POAM III

Neueste Analysen der Genauigkeit von POAM III Version 3 Daten finden sich in [Lumpe *et al.*, 2002]. Die Autoren unterscheiden hier zwischen theoretischen Fehlerbetrachtungen und der Berechnung des Präzisionswertes aus statistischen Analysen der Varianz eines großen Datensatzes von Retrievals. Da nicht ganz klar ist, welcher der beiden Werte für das aufzustellende Fehlerbudget die bessere Wahl ist, spiegeln die Angaben in Tab. 10.1 einen Mittelwert der beiden Fehler wider. Ein Zuschlag für Glättungseffekte erübrigt sich hier, da die nominelle Vertikalaufklärung von POAM III besser als 1 km ist [Lumpe *et al.*, 2002].

10.2.2 Systematische Fehler

Der zweite Schritt besteht in der quantitativen Analyse der in Abschnitt 7.1.3 bereits qualitativ betrachteten Biases zwischen den einzelnen Profildatenquellen. Wie in Kapitel 2 bereits auszugsweise dargelegt, existiert eine große Anzahl von Studien, in denen (unter anderem) die hier verwendeten Ozonmessverfahren miteinander verglichen werden. Jedes Verfahren für sich versucht natürlich, dem „wahren“ Ozonfeld besonders nahe zu kommen, weicht aber in bestimmten Punkten durch systematische Fehler von ihm ab. Üblicherweise – wie auch im Rahmen dieser Arbeit – wird angenommen, dass Ozonsonden dem wahren Ozonfeld am nächsten sind, da sich hier die Instrumentenparameter am besten kontrollieren lassen. Wie oben dargelegt gibt es aber auch bei den Sonden erhebliche Biases zwischen verschiedenen Typen, Messprozeduren, etc. Sieht man auf einer Metaebene das wahre Ozonfeld stark vereinfacht als einen Messpunkt an, so würden die mit unterschiedlichen systematischen Fehler behafteten Ozonretrievalverfahren im Idealfall eine Punktwolke darum bilden. Der NNORSY-Trainingsdatensatz mit seinen verschiedenen Datentypen entspricht dann einer Teilmenge dieser Punktwolke, von der angenommen wird, dass ihr Schwerpunkt nahe am wahren Ozonfeld liegt. Ob dies tatsächlich der Fall ist, kann – wie bei anderen Verfahren auch – letztlich nur durch eine Reihe von Validationsstudien ermittelt werden. Die Auswahl der Trainingsdaten kann diesen Fall jedoch wahrscheinlicher machen: Die Mischung von Daten aus physikalisch und mathematisch sehr unterschiedlichen Verfahren verringern die Chance, dass der Trainingsdaten-Mittelwert in einer ganz bestimmten Richtung vom Zielwert abweicht. So mißt z.B. – ganz grob betrachtet – HALOE offenbar in der NH im Höhenbereich 15–20 km im Vergleich zu Ozonsonden tendenziell zu niedrig [Lu *et al.*, 1997], SAGE II zu hoch [Cunnold *et al.*, 2000b], und POAM III bei 20 km zu niedrig [Deniel *et al.*, 1997] und bei 15 km zu hoch [Lucke *et al.*, 1999]. Eine systematische Literaturstudie aller dieser Abweichungen für die verwendeten Datentypen, unter Berücksichtigung aller Höhen- und Latitudenbereiche, des Datenmischungsverhältnisses und der Versionshistorie der Daten würde helfen, um die Lage des Trainingsdaten-Mittelwertes besser einschätzen zu können, muss aber an dieser Stelle aus Platzgründen ebenfalls an zukünftige Arbeiten verwiesen werden. Einige Anhaltspunkte in dieser Richtung werden jedoch die weiter unten folgenden Fallstudien liefern.

Wenden wir uns stattdessen einer Untersuchung der durch die Sensorenbiases hervorgerufenen Streuung der „Trainingsdaten-Punktwolke“ zu, um beim obigen Bild zu bleiben. Dazu wird der Trainingsdatensatz in K Latitudenbänder unterteilt. Ein einzelner, gemessener Zielwert für die Belernung wird mit t_{ijk}^p bezeichnet, wobei $i = 1, \dots, H$ die Höhengschicht, $j = 1, \dots, J$ den Sensorentyp (vgl. Tab. 10.1) und $k = 1, \dots, K$ das Latitudenband angibt. Der Index p nummeriert alle vorhandenen Werte durch. Weiterhin existiert für jede Rasterposition ijk eine Indexmenge

$$\mathcal{I}_{ijk} = \{p_1, p_2, \dots, p_{N_{ijk}}\}, \quad (10.1)$$

in der die Indizes der dort vorhandenen Zielwerte hinterlegt sind. Dann ist

$$\langle t_{ij} \rangle_k = \frac{1}{N_{ijk}} \sum_{p \in \mathcal{I}_{ijk}} t_{ijk}^p \quad (10.2)$$

der über Latitudenband k gemittelte Zielwert für Sensor j in der Höhengschicht i . Die systematischen Abweichungen zwischen den Sensorentypen j und j^* lassen sich in dieser Nomenklatur schreiben als

$$\mu_{ijj^*k} = \langle t_{ij} \rangle_k - \langle t_{ij^*} \rangle_k. \quad (10.3)$$

Wenn diese Biases nicht alle in die gleiche Richtung zeigen und mit Höhengschicht und Latitude variieren – was wie oben erläutert für die verwendeten Daten plausibel ist – wird das neuronale Netz sie in erster Näherung als eine weitere Fehlerquelle behandeln, da es gewissermaßen durch die systematischen Widersprüche verunsichert wird und das intern erzeugte, mittlere Retrievalprofil infolgedessen eine geringere Präzision aufweist. Für die Fehlerfortpflanzung der Biases spielt zusätzlich die Häufigkeitsverteilung der vorhandenen Messungen eine Rolle. Jeder Differenz wird daher ein heuristischer Gewichtungsfaktor

$$\alpha_{ijj^*k} = \left(\frac{1}{N_{ijk}} + \frac{1}{N_{ij^*k}} \right)^{-1} \quad (10.4)$$

zugewiesen, der eine höhere Gewichtung derjenigen Differenzen bewirkt, bei denen von beiden beteiligten Sensoren ungefähr gleich viele Messwerte vorhanden sind, ein etwaiger Widerspruch also besonders signifikant ist. Damit läßt sich der mittlere quadratische Biasfehler für Höhengschicht i im Latitudenband k angeben zu

$$\langle \mu_i \rangle_k = \frac{\sqrt{\sum_{j \neq j^*} (\alpha_{ijj^*k} \mu_{ijj^*k})^2}}{\sum_{j \neq j^*} \alpha_{ijj^*k}}. \quad (10.5)$$

Diesen Formalismus direkt auf die in Tabelle 10.1 verwandten Höhengschichten und Latitudenbänder anzuwenden ist nicht sinnvoll, da die meridionale Verteilung der Sensorenmessungen innerhalb der breiten Latitudenbereiche stark schwankt, was wiederum in den Mittelwerten zu Differenzen führen kann. Das neuronale Netz nimmt diese Differenzen jedoch nicht wahr, weil es anhand der kontinuierlichen Latitude als Eingangskoordinate zwischen ihnen unterscheiden kann. Andererseits ist eine gewisse Mindestbreite der Latitudenbänder erforderlich, um genügend Trainings-Kollokationen für zuverlässige Mittelwertbildung zu haben. Als Kompromiß wurden Bänder von 10° Breite gebildet. Die Höhengauflösung der Netzausgabe von 1 km wurde zunächst beibehalten. Nach Berechnung der Statistiken wurden diese hoch aufgelösten Daten $\langle \mu_i \rangle_k$ zunächst durch den Mittelwert aller Messungen des jeweiligen Bandes geteilt, um relative Werte $\langle \delta_i \rangle_k$ zu erhalten. Anschließend wurden sie durch arithmetische Mittelung – gewichtet mit der normierten Anzahl der Werte – auf die in Abschnitt 10.1 festgelegten Bänder \tilde{k} und Schichten \tilde{i} heruntergerechnet, was die Biasfehler $\langle \delta_{\tilde{i}} \rangle_{\tilde{k}}$ in Tabelle 10.2 ergibt.

Tabelle 10.2: Aus Tab. 10.1 entnommene, für den Testdatensatz über alle Sensoren gemittelte Präzisionsfehler der kollokierten Ozonprofile, sowie durch Biases zwischen den kollokierten Ozonprofildaten hervorgerufene Eingangsfehler (zur Berechnung siehe Text).

GPH [km]	Präzisionsfehler			Biasfehler		
	30°S-90°S	30°S-30°N	30°N-90°N	30°S-90°S	30°S-30°N	30°N-90°N
41-60	5.5%	6.0%	5.7%	8.7%	5.5%	8.5%
26-40	4.4%	5.1%	5.4%	3.3%	1.1%	2.8%
15-25	5.4%	6.3%	5.6%	4.8%	5.5%	3.4%
6-14	10.5%	11.3%	9.3%	9.4%	2.9%	8.1%
0-5	7.2%	8.6%	7.2%	–	–	–

Die in Tab. 10.1 angegebenen zufälligen Fehler $\sigma_{\tilde{i}\tilde{k}}$ der Satelliten- und Sonden-Kollokationen können nun ebenfalls mit ihrer Häufigkeit im Trainingsdatensatz gewichtet und

gemittelt werden, wobei sich hier wiederum der arithmetische Mittelwert anbietet, da nicht unbedingt davon auszugehen ist, dass sich diese Fehler gegenseitig kompensieren. Es gilt somit

$$\langle \sigma_{\tilde{i}} \rangle_{\tilde{k}} = \frac{\sum_j N_{\tilde{i}\tilde{j}\tilde{k}} \sigma_{\tilde{i}\tilde{j}\tilde{k}}}{\sum_j N_{\tilde{i}\tilde{j}\tilde{k}}}. \quad (10.6)$$

Diese Werte zeigt Tabelle 10.2, wiederum relativ zum jeweiligen Mittelwert aller Messungen für die Schicht \tilde{i} . Strenggenommen müssten diese Werte für Trainings- und Testdatensatz separat ermittelt werden, aber aufgrund der großen Kollokationsanzahl und der sehr ähnlichen Verteilung der Profile können die Unterschiede zwischen beiden Datensätzen vernachlässigt werden.

10.3 Andere Fehlerquellen

10.3.1 Eingangsdatenfehler

Für die auf Unsicherheiten der Eingangsdaten zurückzuführenden Fehlerbeiträge kann eine gute Abschätzung gewonnen werden, indem die in Abschn. 9.1.1 eingeführte Rauschfehler-Methode auf den global verteilten Trainingsdatensatz angewandt wird. Das Ergebnis nach entsprechender Mittelwertbildung ist in Tabelle 10.3 dargestellt.

Tabelle 10.3: *Durch Unsicherheiten der Netzwerk-Eingangsdaten hervorgerufene Fehler, gemittelt über alle kollokierten GOME-Pixel.*

GPH [km]	Geografische Breite		
	30°S-90°S	30°S-30°N	30°N-90°N
41-60	3.6%	3.2%	3.5%
26-40	3.9%	2.3%	3.8%
15-25	8.3%	12.4%	7.4%
6-14	16.7%	23.0%	15.0%
0-5	9.9%	13.7%	7.5%

10.3.2 Kollokationsfehler

Um die horizontalen Kollokationsfehler abzuschätzen, bietet es sich an, Fehlerstatistiken für jeweils alle verfügbaren Kollokationen zu erstellen, die eine Kollokationsdistanz zwischen r und $r + \Delta r$ aufweisen. Dafür muss jedoch ein eigener Datensatz kreiert werden, da die Trainings- bzw. Testdaten aufgrund der sensoren- und stationsabhängigen Gleichverteilungs-Maßnahmen (s. Abschn. 7.1.2) zu Verzerrungen des Ergebnisses führen würden. Δr darf nicht zu klein gewählt werden, damit die Anzahl der vorhandenen Kollokationen in jedem Abstandsring statistisch aussagekräftig bleibt. Die relativen Fehlerstatistiken von Retrieval gegenüber den kollokierten Ozonprofilen werden dann wie in Abschn. 7.4 beschrieben für die gewünschten drei Latitudenbänder separat erstellt und wiederum auf die oben definierten fünf Höhenschichten heruntergerechnet. Man erhält für jedes Koordinatenpaar $[\tilde{i}; \tilde{k}]$ eine Kurve des relativen Fehlers gegenüber r . Die lineare Extrapolation dieser Kurve zum Abstand $r=0$ km ergibt dann den relativen Fehler in Abwesenheit jeglicher Kollokationsfehlerbeiträge. Für einzelne Kollokationen mit Abstand r_1 kann somit der

durchschnittliche Kollokationsfehlerbeitrag bestimmt werden, indem aus der Kurve die Differenz der relativen Fehler bei r_1 und $r=0$ km abgelesen wird. Für den Testdatensatz wird der Kollokationsfehler einfach bei der mittleren Kollokationsdistanz $r_1 = \langle r \rangle_{\text{test}}$ abgelesen.

Das Problem hierbei ist allerdings, dass aufgrund der ungleichmäßigen Abstandsverteilung der Belernungskollokationen – durch die Homogenisierungsmaßnahmen und die mit r quadratisch zunehmende Kreisfläche – Höhenbereiche, in denen viel statistische Information in das Retrieval einfließt, keine monotone Zunahme des Fehlers mit der Kollokationsentfernung zeigen. Daher wurden nur solche Höhenschichten für die Auswertung verwendet, bei denen in etwa eine lineare Zunahme beobachtet wurde. Die Fehler der anderen Schichten lassen sich daraus näherungsweise bestimmen, indem entsprechend der Variabilität des Ozonfeldes (Abb. 7.3, 7.4) skaliert wird. Für die SH herrscht eine grob lineare Zunahme des Kollokationsfehlers in der Höhenschicht 15–25 km (Abb. 10.1a), in den Tropen zwischen 26–40 km (Abb. 10.1b).

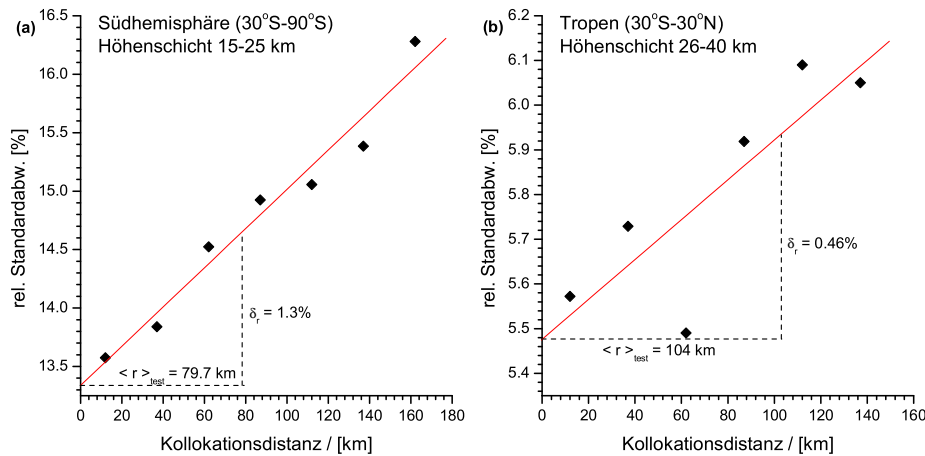


Abbildung 10.1: Zunahme des Kollokationsfehlers mit der Kollokationsentfernung r . Die Berechnung der Standardabweichung von (NNORSY minus Kollokationsprofil) erfolgte wie in Abschn. 7.4 beschrieben, jedoch auf einem räumlich nicht homogenisierten Datensatz und auf Entfernungs-„Ringene“ von $\Delta r = 25$ km Breite. $\langle r \rangle_{\text{test}}$ ist der mittlere Kollokationsradius im Testdatensatz, für die entsprechende Breitenregion.

In der NH fand sich leider in keiner Höhenschicht eine realistische lineare Zunahme des Kollokationsfehlers. Die Gründe hierfür sind nicht ganz klar; möglicherweise hängt die Verschleierung der Fehlerzunahme mit den Restinhomogenitäten des Trainingsdatensatzes zusammen. Das Ozonfeld der NH ist jedoch bereits recht gut untersucht, so dass ein Experiment von *Fioletov et al.* [1999] hier Abhilfe schaffen kann. Die Autoren bestimmten den Kollokationsfehler des SBUV-Gesamtozons im Vergleich zu 17 NH-Bodenstationen. Eine lineare Regression auf diesen Daten bis etwa 150 km Kollokationsdistanz ergibt für den durchschnittlichen Kollokationsradius $\langle r \rangle_{\text{test}} = 96$ km des Testdatensatzes einen relativen Gesamtozon-Kollokationsfehler von etwa 1.1%. Dieser Wert stimmt gut mit den Gesamtozonfehlern überein, die man durch die oben beschriebene Umrechnung der Fehler einer Höhenschicht auf das gesamte Ozonprofil erhält, nämlich 1.2% (SH) und 1.1% (Tropen). Umgekehrt kann der NH-Gesamtozonfehler unter Zuhilfenahme der Ozonfeldvariabilität (Abb. 7.3, 7.4) auf die Höhenschichten verteilt werden.

Die zeitlichen Kollokationsfehler können näherungsweise separiert werden, indem die Tagesvariabilität des Gesamtozon-Feldes aus Abb. 6.5 ebenfalls mit der Ozonprofil-Variabilität aus Abb. 7.3 und 7.4 auf die festgelegten Höhenschichten skaliert wird. Auch hier läßt sich die Variation für einen bestimmten zeitlichen Abstand (im Testdatensatz im Mittel ca. 5 h)

durch lineare Interpolation gewinnen [Fioletov et al., 1999]. Das Ergebnis für alle Breitenregionen zeigt Tabelle 10.4.

Tabelle 10.4: Mittelwerte der Kollokationsfehler für den Testdatensatz (zur Berechnung siehe Text).

GPH [km]	räumlich			zeitlich		
	30°S-90°S	30°S-30°N	30°N-90°N	30°S-90°S	30°S-30°N	30°N-90°N
41–60	0.7%	0.5%	0.8%	1.3%	0.2%	0.8%
26–40	0.8%	0.5%	0.7%	1.5%	0.2%	0.7%
15–25	1.3%	1.5%	1.0%	2.5%	0.6%	1.1%
6–14	1.6%	2.7%	2.3%	3.2%	1.2%	2.5%
0–5	0.9%	1.9%	0.7%	1.8%	0.8%	0.8%

10.3.3 Glättungsfehler

Bei der Auswertung spektraler Daten mit Hilfe von OE kann der Glättungsfehler leicht aus den AKs berechnet werden [Rodgers, 1976] und stellt normalerweise einen beträchtlichen Anteil des Gesamtfehlers [Rodgers, 2000]. In Ermangelung dieser theoretischen Möglichkeiten muss bei NNORSY wieder auf die Heuristik zurückgegriffen werden, die schon in Abschn. 7.4.3 für die Abschätzung der Vertikalaufösung gute Dienste geleistet hat. Die Annahme dabei ist, dass der Glättungsfehler näherungsweise verschwindet, wenn die hohe vertikale Auflösung der kollokierten Ozonprofile künstlich degradiert wird, bis sie der Vertikalaufösung des NNORSY-Retrievals entspricht. Im Gegensatz zu den Betrachtungen in Abschn. 7.4.3 interessiert jedoch diesmal nicht, wie stark die Profile geglättet werden müssen, sondern wie klein der durch die Glättung erreichte minimale Fehler im Vergleich zum Ausgangswert wird.

In Anbetracht der im erwähnten Absatz angesprochenen Problematik von Randeffekten, und deren Vermeidung durch Ergänzung unvollständiger Profile mit Retrievalwerten, erscheint es in diesem Zusammenhang sinnvoller, für die Statistik der untersten beiden Höhengschichten (0–5 km und 6–14 km) nur Sondenkollokationen zu verwenden, und für die obersten beiden (26–40 km und 41–60 km) nur Okkultationsprofile. Dann ist der Effekt der Ergänzungen bzw. Randeffekte in erster Näherung vernachlässigbar.

Um die Werte in Tabelle 10.5 zu erhalten, wurde die in Abschn. 7.4.3 beschriebene Prozedur für jedes der drei Latitudenbänder separat auf den Testdatensatz angewandt, und für jede Höhengschicht der relative Beitrag des Glättungsfehlers bestimmt, indem die Differenz zwischen Fehler ohne Glättung und Fehler bei optimaler Glättung durch den mittleren Ozonwert der Retrievals geteilt wurde.

10.3.4 Modellfehler

Um die Modellfehler abzuschätzen kann auf die am Anfang von Kapitel 8 beschriebene Versuchsreihe zurückgegriffen werden, bei der zehn neuronale Netze identischen Aufbaus, aber unterschiedlicher Gewichtsinitialisierung mit denselben Trainingsdaten belernt wurden. Dazu wurde für jedes Retrievalprofil des dortigen Testdatensatzes die Standardabweichung aus den zehn Trainingsläufen bestimmt, und diese Werte über die drei hier verwendeten Latitudenbereiche gemittelt. Das Ergebnis zeigt Tabelle 10.6.

Tabelle 10.5: *Glättungsfehler der NNORSY-Ozonretrievals (zur Berechnung siehe Text).*

GPH [km]	Geografische Breite		
	30°S-90°S	30°S-30°N	30°N-90°N
41–60	0.6%	1.8%	0.4%
26–40	7.8%	13.4%	11.5%
15–25	17.2%	14.3%	22.1%
6–14	14.2%	18.4%	17.3%
0–5	4.1%	21.5%	13.3%

Tabelle 10.6: *Aus der Standardabweichung von zehn zufällig initialisierten Trainingsläufen abgeschätzte Modellfehler der NNORSY-Ozonretrievals.*

GPH [km]	Geografische Breite		
	30°S-90°S	30°S-30°N	30°N-90°N
41–60	2.5%	2.1%	3.1%
26–40	2.3%	1.5%	2.5%
15–25	3.9%	5.9%	2.9%
6–14	9.5%	12.8%	5.8%
0–5	9.7%	6.9%	4.4%

10.4 Zusammenstellung des Fehlerbudgets

Nachdem nun die einzelnen Fehlerbeiträge abgeschätzt wurden, läßt sich das **Referenz-Fehlerbudget** für die Testdatensatz-Statistiken wie folgt aufstellen:

$$\delta_{\text{Ref}} = \sqrt{\underbrace{\delta_{\text{RF}}^2}_{\text{Rauschfehler}} + \underbrace{\delta_{\mu}^2}_{\text{Profilbiasf.}} + \underbrace{\delta_{\sigma}^2}_{\text{Profilpräz.}} + \underbrace{\delta_t^2 + \delta_r^2}_{\text{Kollokationsf.}} + \underbrace{\delta_G^2}_{\text{Glättungsf.}}} \quad (10.7)$$

Diese Gleichung gilt separat für jeden Tabelleneintrag – die Indizes \tilde{i} und \tilde{k} wurden zur Vereinfachung weggelassen. In Tabelle 10.7 wird dieses Referenz-Fehlerbudget den auf dem Testdatensatz tatsächlich beobachteten Standardabweichungen von (NNORSY-Retrieval minus Kollokation) gegenübergestellt (vgl. Abb. 7.8 und 7.9). Dabei muss beachtet werden, dass die Profilpräzisionsfehler δ_{σ} doppelt eingehen: Einmal im Trainingsdatensatz, wo sie die Belernung stören, und einmal im Testdatensatz bei der Berechnung der Fehlerstatistik. Anders ausgedrückt setzen sich streng genommen die in Kapitel 7.4 vorgestellten Testdatenfehler aus der quadratischen Summe von Retrievalfehler und Kollokationsprofilfehler zusammen, was bislang vernachlässigt wurde. Für die Werte in Tabelle 10.7 wurde daher der Kollokationsprofilfehler aus den Testdatensatz-Standardabweichungen herausgerechnet. Die Fehlerbeiträge δ_{RF} und δ_{μ} wirken sich hingegen nur auf die Belernung aus, während δ_t , δ_r und δ_G durch die Art ihrer Berechnung aus den Testdatensatz-Retrievals bereits Belernungs- und Teststatistikanteile enthalten.

Wie aus der Tabelle zu entnehmen, hat die NNORSY-Methode oberhalb von 14 km keinerlei Schwierigkeiten, die vorgegebenen Referenzwerte zu unterschreiten. Allerdings fällt die Unterschreitung in den Tropen bei 15–25 km geringer aus als anderswo, da in diesen Höhen bereits Teile der Troposphäre enthalten sind. Dies führt uns zum Problempunkt des

Tabelle 10.7: Gegenüberstellung von Referenz-Fehlerbudget und der beobachteten NNORSY-Testdatenstandardabweichung. NNORSY-Fehlerwerte, die über der Referenz liegen, wurden hervorgehoben.

GPH [km]		Geografische Breite		
		30°S-90°S	30°S-30°N	30°N-90°N
41–60	Referenz	11.1%	9.0%	10.9%
	NNORSY	7.0%	5.7%	7.7%
26–40	Referenz	10.5%	14.5%	13.6%
	NNORSY	5.3%	3.0%	4.2%
15–25	Referenz	20.6%	20.8%	24.3%
	NNORSY	13.1%	17.7%	11.2%
6–14	Referenz	26.3%	31.8%	26.2%
	NNORSY	31.3%	40.4%	27.3%
0–5	Referenz	13.1%	27.0%	16.9%
	NNORSY	23.0%	25.5%	14.4%

Verfahrens, der offenbar in der unterdurchschnittlich guten Erfassung der 6–14 km Schicht liegt, also der oberen Troposphäre sowie – außerhalb der Tropen – der untersten Stratosphäre. Angesichts der Ergebnisse der vorherigen Kapitel verwundert dies nicht, da dort bereits an verschiedenen Stellen auf die relativ großen Unsicherheiten in den Höhen um 9 km hingewiesen wurde. Hier besteht großes Verbesserungspotenzial, zum Beispiel durch Einführung eines tropopausenreferenzierten, dynamischen Höhengitters [Logan, 1999]. Dadurch ließe sich eine bessere Übereinstimmung von Ozonprofilen diesseits und jenseits des subtropischen Tropopausenbruchs realisieren, was die Homogenität der Trainingsdaten deutlich erhöhen würde. Laut *Bojkov & Fioletov* [1997] reduziert sich die Standardabweichung des Ozons an der Tropopause um 15–30%, wenn die Höhe anstatt in Druckkoordinaten in [km] relativ zur Tropopause gemessen wird. Eine solche Reduktion würde dem Belernungsalgorithmus erlauben, schwächere Zusammenhänge zu modellieren, die bislang durch die starken Unterschiede zwischen tropischen und extratropischen Ozonprofilen dominiert und verdeckt wurden.

Allerdings ist zu bedenken, dass dazu jederzeit eine möglichst genaue Tropopausenreferenz aus anderen Quellen bereitgestellt werden muss, um den Trainingsdatensatz – und später die Retrievals – zu generieren. Des Weiteren ist nicht ganz auszuschließen, dass ein Teil der beobachteten hohen NNORSY-Fehler im Bereich 6–14 km durch einzelne, nicht aussortierte Datenfehler in den Kollokationen hervorgerufen werden. Hierzu müsste die Qualitätskontrolle insbesondere für Okkultationsprofile nochmals überprüft und ggf. unter Einbeziehung aktueller Literatur revidiert werden. Weiterhin sollte untersucht werden, ob eine explizite Behandlung von Wolken im Retrieval die Ergebnisse verbessern kann, sofern für jedes GOME-Pixel eine zuverlässige Messung aller notwendigen Wolkenparameter vorliegt.

Die erhöhten NNORSY-Fehler in der 0–5 km Schicht der SH sind geprägt durch die Sondenkollokationen über der Antarktis, da über den Südozeanen fast keine Sondenmessungen vorliegen. Hier tritt eine ganze Reihe von Problemen auf, für deren Diskussion auf Abschnitt 11.2 und Kapitel 12 verwiesen wird. Hier, wie auch in der tropischen Troposphäre, ist aufgrund der relativ geringen Anzahl von Messungen mit einem möglicherweise beträchtlichen Beitrag des Repräsentativitätsfehlers zu rechnen, insbesondere über den Ozeanen. Einen Anhaltspunkt dafür geben auch die Modellfehler in Tab. 10.6: In Regionen,

die vom Trainingsdatensatz nur schlecht abgedeckt sind, ergeben sich erwartungsgemäß die größten Unsicherheiten.

Dementsprechend ändert auch die quadratische Summation von NNORSY-Fehlern aus Tabelle 10.7 und den Modellfehlern nicht viel am Gesamtbild des **globalen Fehlerbudgets** (Tab. 10.8). Beim Vergleich dieser Werte mit entsprechenden Literaturangaben für andere Sensoren und Verfahren ist zu beachten, dass es sich bei den angegebenen Höhenschichten *nicht* um Fehler auf Basis integrierter Teilsäulen handelt, sondern um über die Höhe gemittelte Fehler in der Retrievalauflösung von 1 km. Daher enthalten sie einen beträchtlichen Anteil an Glättungsfehlern (Tab. 10.5). Integrierte Teilsäulen, sowie ein kurzer, qualitativer Vergleich mit anderen Verfahren werden in Kapitel 11 besprochen.

Tabelle 10.8: *Globales NNORSY-Fehlerbudget. Diese Werte spiegeln eine accuracy-Schätzung auf 1 km Höhenschichten, ohne Berücksichtigung der Repräsentativitätsfehler wider.*

GPH [km]	Geografische Breite		
	30°S-90°S	30°S-30°N	30°N-90°N
41-60	7.4%	6.1%	8.3%
26-40	5.8%	3.3%	4.9%
15-25	13.7%	18.7%	11.6%
6-14	32.7%	42.4%	27.9%
0-5	24.9%	26.5%	15.0%

Die obigen Ausführungen lassen unweigerlich die Frage aufkommen, ob es theoretisch möglich ist, mit der vorgeschlagenen NNORSY-Methode eine höhere Genauigkeit zu erreichen, als die Trainingsdaten selber besitzen? Es ist anschaulich sofort klar, dass dies für die Biases gilt: Angenommen, das neuronale Netz würde mit Ozonprofilen zweier Retrievalverfahren trainiert, die jeweils für dasselbe wahre Ozonprofil systematisch zu hohe bzw. um den gleichen Betrag zu niedrige Werte liefern. Der als Netzwerkausgabe resultierende Kompromiß läge dann in allen Fällen näher am wahren Ozonwert, als die zur Erzeugung der Belernungsdaten verwendeten Verfahren für sich gestellt.

Ähnliches gilt für die Varianz der Retrievals: Besäße z. B. das globale Ozonfeld nur eine begrenzte Anzahl von diskreten, anhand der Eingangsdaten unterscheidbaren Zuständen, und bestünde gleichzeitig der Trainingsdatensatz aus 100 verrauschten Messungen für jeden dieser Zustände, so entspräche die Netzwerkausgabe für jeden Zustand in etwa dem Mittelwert jener 100 Messungen. Dadurch ließe sich das Rauschen im Idealfall um den Faktor 10 reduzieren, das neuronale Netz besäße also eine geringere Fehlervarianz als die Einzelmessungen.

In der Realität sind diese Idealbedingungen natürlich nicht erfüllt, aber das Gedankenexperiment zeigt, dass NNORSY tatsächlich als ein neues Verfahren angesehen werden kann, welches in gewissem Sinne über die systematischen und zufälligen Fehler seiner Generatoren hinwegmittelt und insofern zumindest das Potenzial besitzt, genauer als diese zu werden.

Fazit

Der Vergleich der bei NNORSY beobachteten Testdatensatzfehler mit einer Referenzabschätzung der quadratisch summierten Einzelfehlerbeiträge ergab, dass das Verfahren offensichtlich in den meisten Latitudenregionen und Höhen eine besser als quadratische Fehlerkompensation aufweist. Probleme bestehen hauptsächlich im Bereich 6–14 km und lassen

sich u. a. durch die latitudenabhängige Verschiebung der Tropopausenhöhe erklären, hinzu kommen eventuell Wolkeneinfluss und fehlerbehaftete Okkultationsprofile.

Trotzdem vergleicht sich das unter Einbeziehung von Modellfehlern berechnete, globale Fehlerbudget von NNORSY gut mit den *accuracy*-Fehlern anderer Instrumente. Insbesondere liegen die Schätzungen oberhalb von 25 km teilweise im Bereich der *precision* von Okkultations-Soundern. Im Bereich des Ozonmaximums und darunter machen die hohen Glättungsfehler einen direkten Vergleich der Genauigkeiten schwierig, daher wird dieser in den Fallstudien des Teils IV der Arbeit näher untersucht. Die Fehler in der unteren bis mittleren Troposphäre liegen für ein satellitengestütztes Verfahren relativ niedrig, bergen aber außerhalb der Nordhemisphäre die Gefahr von Unterrepräsentation im Trainingsdatensatz, und den damit verbundenen Repräsentativitätsfehlern.

Zusammenfassung von Teil III

Die der NNORSY-Methode eigene Mischung von statistischen und physikalischen Informationen führt zu einem relativ komplizierten Fehlerverhalten des Systems. Um dessen Verständnis näher zu kommen, wurde versucht, durch Belernung von neuronalen Netzen mit unterschiedlicher Konfiguration einen Einblick in die relative Bedeutsamkeit der Eingangsgrößen zu erlangen. Dabei erwiesen sich MLPs bei Belernung mit rein klimatologischen Parametern als fähig, eine Repräsentation des Ozonfeldes zu finden, deren Standardabweichung auf den Testdaten mit derjenigen der F&K-Klimatologie vergleichbar ist, und in der Stratosphäre um 45–75% über der Referenzbelernung liegt.

Bei spektralen Eingangsdaten konnte keine eindeutige **Zuordnung von Spektralbereich und Höhe** des größten Einflusses auf das Ozonprofil festgestellt werden. Zum selben Ergebnis kommen auch Sensitivitätsanalysen mit künstlich verrauschten Eingangsdaten. Die Betrachtung des minimal erreichten RMSE auf Trainings- und Testdaten läßt hingegen Rückschlüsse auf den Beitrag von einzelnen (Gruppen von) Eingangsparametern zu. Dabei stellte sich abgesehen von den Spektraldaten das **UKMO T-Profil** als wichtigste Informationsquelle im Bezug auf die Ozonprofilbestimmung heraus. Die Hinzunahme von T-Profilen bewirkt u. a. einen geringeren Einfluss der Spektraldaten auf das Retrieval, wie sich anhand abnehmender Rauschfehler feststellen läßt.

Für die Höhenkorrelation der Retrievalprofile wurden mit Hilfe der Rauschfehler und der neuronalen Jacobi-Matrix gut übereinstimmende Ergebnisse gefunden, die u. a. die empirische Vertikalauflösung des Retrievals indirekt bestätigen. Die klassische lokale Fehleranalyse mittels Jacobi-Matrix scheint hingegen aufgrund von nichtlinearen Kompensationseffekten im MLP für NNORSY ungeeignet zu sein, kann aber bei Bedarf durch die Rauschfehlermethode oder die Belernung eines Fehlernetzwerkes ersetzt werden.

Einige weitere Aussagen über das Fehlerverhalten des Systems lassen sich durch den Vergleich der Trainings- und Testdatensatz-Fehlerstatistiken mit externen Fehlerangaben gewinnen. Hierzu wurden Präzisionsfehlerangaben für die kollokierten Profildaten und GOME-Spektren aus der Literatur entnommen, und eigene Abschätzungen für Kollokationsfehler, Glättungsfehler, Modellfehler und systematische Abweichungen zwischen den Sensorentypen verwendet. Es zeigte sich, dass das System in den meisten Regionen und Höhen geringere Fehler aufweist als die Quadratsumme der Einzelbeiträge. Gewisse Probleme bestehen noch in der Tropopausenregion und der antarktischen Troposphäre. Dieses **globale Fehlerbudget** beschreibt zusammen mit den Einzelprofil-Fehlern aus verrauschten Eingangsdaten (*precision*-Fehler) und dem neuronalen Fehlernetzwerk (*accuracy*) die Fehlercharakteristika des NNORSY-Algorithmus.

Teil IV

Fallstudien und Anwendung

Der letzte Teil dieser Arbeit enthält einige Fallstudien, anhand derer einzelne Aspekte der NNORSY-Methode genauer diskutiert, Vergleiche mit anderen Verfahren angestellt und einige Schlussfolgerungen zur Anwendbarkeit der Methode gezogen werden.

In **Kapitel 11** werden Zeitreihen der Bodenstationen Hohenpeißenberg und Syowa im Hinblick auf das (jahres-)zeitliche Verhalten der Retrievalfehler und deren Höhenabhängigkeit analysiert. Ein Unterkapitel stellt außerdem einige Ozonprofilvergleiche mit anderen Verfahren vor.

Um von den bisher hauptsächlich auf den Kollokationsdatensätzen ermittelten Fehlern auf deren flächendeckende Verteilung zu schließen, bietet sich ein Vergleich von integrierten Ozonprofilen mit verschiedenen satellitenbasierten Gesamtozonmessungen an. Diesem widmet sich **Kapitel 12**.

In **Kapitel 13** wird kurz auf den Status der Echtzeitanwendung von NNORSY eingegangen und Verbesserungsmöglichkeiten für den bereits existierenden Prototyp aufgezeigt.

Kapitel 14 fasst nochmals die wichtigsten im Rahmen des Projektes gesammelten Erkenntnisse zusammen, und gibt dabei gleichzeitig Empfehlungen für zukünftige Verbesserungen der NNORSY-Methode.

Kapitel 11

Zeitreihe ausgewählter Bodenstationen

Um die unabhängige Überprüfung der zeitlichen Stabilität der NNORSY-Ozonprofile zu ermöglichen, wurden Messungen der Stationen Hohenpeißenberg und Syowa wie bereits erwähnt weder im Trainings- noch im Testdatensatz verwendet. In den Abschnitten 11.1 und 11.2 werden für diese beiden Stationen Zeitreihen auf der Basis von integrierten 10 km-Höhenschichten untersucht. Ähnliche Graphen finden sich auch in [Hoogen, 1998]. Das dabei gewonnene Bild wird mit der Betrachtung von höher aufgelösten Streudiagrammen ergänzt. Für alle Darstellungen wurden ggf. NNORSY-Mehrfachkollationen innerhalb von 250 km Radius um das Ozonsondenprofil gemittelt. Weiterhin stellt Abschnitt 11.3 einige extern durchgeführte Vergleiche mit anderen Retrievalverfahren vor. Auf die globale Verteilung der beobachteten Effekte wird dann im nächsten Kapitel eingegangen.

11.1 Ozonesonden Hohenpeißenberg

Die Zeitreihe für Hohenpeißenberg in Abb. 11.1a zeigt in allen drei Teilsäulen eine gute Übereinstimmung zwischen NNORSY und den Ozonsonden. Wie zu erwarten entstehen im Frühjahr die größten Abweichungen, da zu dieser Jahreszeit die räumlich-zeitliche Ozonfeldvariabilität in den nördlichen mittleren Breiten besonders hoch ist. Es sind keine signifikanten Drifts gegenüber den Sonden zu erkennen; die in Abb. 6.6b präsente leichte Überschätzung des Gesamt Ozons zu Beginn der Zeitreihe läßt sich hier in den Teilsäulen nicht beobachten. Auch nach dem Umschalten von ERS-2 auf den gyrolosen Steuerungsmodus im Januar 2001 entsteht keine sichtbare Verschlechterung der NNORSY-Retrievals, obwohl in den Level 1 Daten eine Reihe von Problemen auftrat. Davon abgesehen werden auch die durch Degradation des Sensors hervorgerufenen Veränderungen des GOME-Spektrums [Aben *et al.*, 2000] von NNORSY kompensiert; Obwohl noch keine formale Trendanalyse mit diesen Daten durchgeführt wurde, ist zu erwarten, dass sich die Trends von Sonden und NNORSY-Ozonprofilen recht gut entsprechen werden.

Die ebenfalls in der Abbildung eingezeichneten Teilsäulen der F&K-Klimatologie können den kurzfristigen Schwankungen des Ozonfeldes selbstverständlich nicht folgen, weisen aber zudem in der Höhenschicht 20–30 km systematisch zu hohe Werte auf. Dies mag bereits auf merkliche Verluste an durchschnittlicher Ozonkonzentration in diesem Bereich hindeuten [Reid *et al.*, 2000], da die Klimatologie aus Ozonmessungen von 1980–1991 konstruiert wurde. Die NNORSY-Retrievals weisen hingegen nur minimale Biases im Vergleich zu den Sonden-Teilsäulen auf.

Abb. 11.2a zeigt denselben Datensatz als Streudiagramme, gemittelt über 5 km dicke Höhenschichten. Die höchsten Pearson-Korrelationskoeffizienten von 0.902 und 0.935 wer-

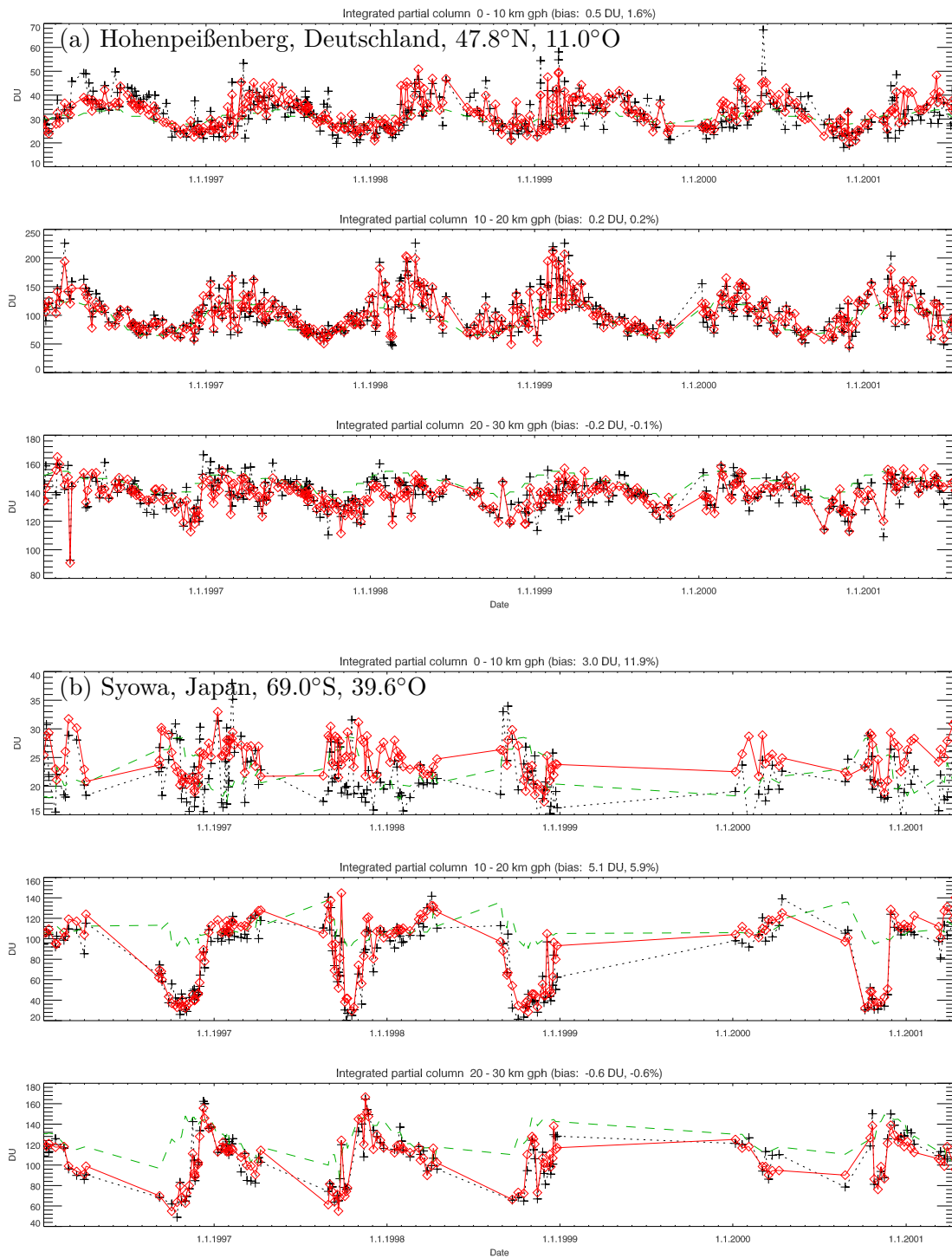


Abbildung 11.1: Zeitreihenvergleich von 10 km-Teilsäulen aus NNORSY-Ozonprofilen (\diamond) mit entsprechenden Werten (+) der Stationen Hohenpeißenberg und Syowa. Die Daten beider Stationen waren komplett von der Belerung ausgenommen worden. Die F&K-Klimatologie [Fortuin & Kelder, 1998] ist zusätzlich als gestrichelte Line eingetragen.

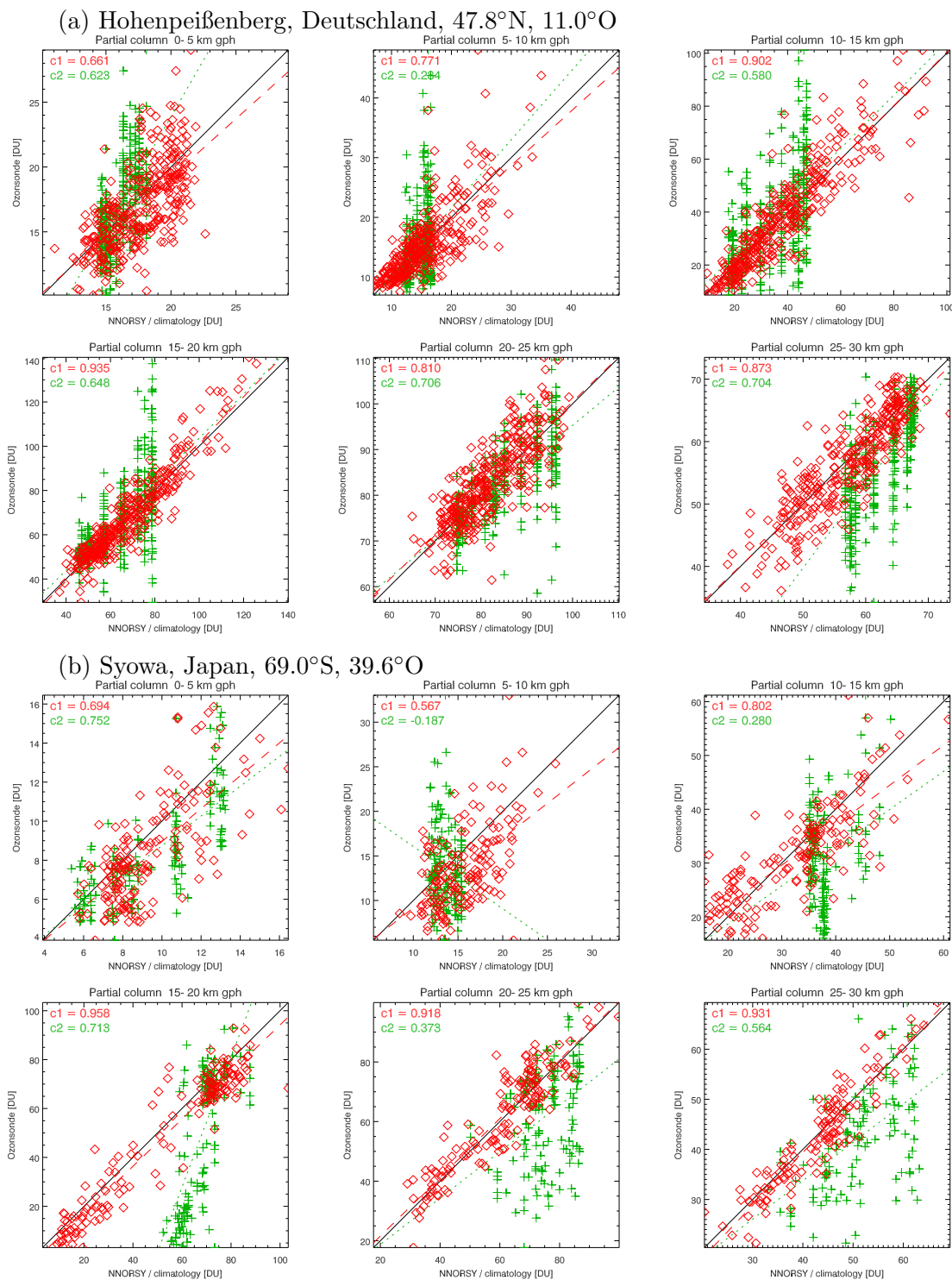


Abbildung 11.2: Wie Abb. 11.1, aber Streudiagramme von 5 km-Teilsäulen der Ozonsonden gegenüber NNORSY (\diamond) und der F&K-Klimatologie (+). Regressionsgeraden sind gestrichelt bzw. gepunktet dargestellt. c_1 und c_2 geben die entsprechenden Pearson-Korrelationskoeffizienten an.

den in den beiden Schichten zwischen 10 km und 20 km Höhe beobachtet, was den visuellen Eindruck aus Abb. 11.1a bestätigt. In der Schicht 5–10 km bildet NNORSY den ozonarmen troposphärischen Hintergrund recht genau ab, aber für größere Ozonwerte streuen die Messpunkte stärker. Diese hohen Werte treten meist als Folge von Luftmassenaustausch zwischen Stratosphäre und Troposphäre oder photochemischen Prozessen auf [*Lelieveld & Dentener, 2000*]. Die Ozonbestimmung vom Satelliten aus kann in diesen Fällen durch Wolken beträchtlich gestört werden; Zudem spielen sich die betreffenden Ereignisse meist auf kurzen Zeitskalen und räumlichen Entfernungen ab, was die Kollokationsfehler vergrößert. Dennoch verläuft die Regressionsgerade zwischen NNORSY und den Sonden-Teilsäulen fast diagonal, und das System scheint im Gegensatz zur Klimatologie das Auftreten hoher Ozonwerte meist richtig zu diagnostizieren, wenn auch die Magnitude nicht immer korrekt erfasst wird. Die Streudiagramme für die niedrigste und die beiden höchsten Schichten weisen keine besonderen Merkmale auf, allerdings zeigt sich oberhalb von 20 km wie schon in Abb. 11.1a auch hier wieder der positive Bias der F&K-Klimatologie.

11.2 Ozonesonden Syowa

Ein etwas anderes Bild ergibt sich beim Betrachten analoger Darstellungen für die japanische Station Syowa, gelegen am Rande des antarktischen Kontinents (Abb. 11.1b und Abb. 11.2b). Hier ist die Streuung der troposphärischen Ozonwerte stärker ausgeprägt als in Hohenpeißenberg, so dass die NNORSY-Ergebnisse unterhalb von 10 km Höhe fragwürdig sind. Da die GOME Messungen hier fast immer bei großem SZA ausgeführt werden, ist es denkbar, dass sich die Gewichtsfunktionen mit troposphärischem Maximum (vgl. Abb. 3.2) in den langwelligen Spektralbereich jenseits der für die Eingangsdaten gewählten 325 nm-Grenze verschieben. Wie bereits in Abschn. 7.3 diskutiert, führte die testweise Verlagerung dieser Grenze auf 335 nm oder 340 nm jedoch global gesehen zu minimal größeren Fehlern, während für Syowa keine nennenswerte Verbesserung bemerkt wurde. Es ist daher eher wahrscheinlich, dass die großen Sonnenzenitwinkel in Verbindung mit der durch die Eisbedeckung hohen Albedo der Antarktis – die sich im UV nur schwer von Wolken unterscheiden lässt [*McPeters et al., 1998; Corlett & Monks, 2001*] – der Grund für die Probleme sind [*Manney et al., 2001*]: Wenn dadurch das S/N-Verhältnis troposphärischer Information in den GOME-Spektren zu gering wird, geht diese Information im globalen Trainingsdatensatz unter, da der Belernungsalgorithmus zunächst die ebenfalls enthaltenen, deutlicheren Zusammenhänge (größere Höhen, andere Latituden, etc.) modelliert. Hier könnte eventuell das Training mit einem spezialisierten Datensatz Abhilfe schaffen, dieser Ansatz wurde aber noch nicht weiter verfolgt.

Für die SZA/Albedo-These spricht auch, dass die Retrievals wie in Abb. 11.1b zu sehen im Mittel um über 10% zu hoch liegen. Es ist durchaus wahrscheinlich, dass das neuronale Netz die hohe Albedo intern zum Teil als vermeintliche Wolkenbedeckung interpretiert, und dementsprechend eine Korrektur für die „verdeckte“ troposphärische Ozonsäule addiert.

Oberhalb der Tropopause erfasst NNORSY die Ozondynamik offenbar recht gut. Während die 10–20 km Schicht in Abb. 11.1b immer noch eine leichte Überschätzung der Ozonsäule aufweist, weisen die Korrelationskoeffizienten in Abb. 11.2b zumindest oberhalb von 15 km bereits extrem hohe Werte auf. Besonders in der 15–20 km Schicht bilden sich zwei deutlich abgesetzte Punktwolken, die die scharfe Trennung zwischen Ozonlochbedingungen und ungestörter polarer Ozonschicht unterstreichen. Wiederum liegt die F&K Monatsklimatologie im Mittel zu hoch, da sich ihre Werte auf den Bereich der ungestörten Ozonsituation beschränken und die Ozonminima nicht erfassen. Dies kann wiederum als direkter Nachweis der Vertiefung des Ozonloches seit den 80er Jahren, aus denen die Klimatologie stammt, gewertet werden [*Spichtinger-Radowsky, 2001*].

Es bleibt zu bemerken, dass auch in dieser Zeitreihe trotz der größeren Fehler in den

unteren Atmosphärenschichten keine zeitliche Variation systematischer Biases zu erkennen ist. Über eine eventuelle Zunahme der Streuung nach Beginn des Jahres 2001 kann allerdings aufgrund der wenigen Messungen noch keine verlässliche Aussage getroffen werden; Die Ozonlöcher bis einschließlich 2000 werden offenbar mit vergleichbarer Genauigkeit erfasst.

11.3 *GOME Ozone Profile Working Group*

Am 4. April 2001 fand bei ESA/ESRIN in Frascati (Italien) das erste *GOME1 Exploitation Meeting* statt, bei dem sich die *GOME Ozone Profile Working Group* (GOPWG) konstituierte. Sie hat sich zum Ziel gesetzt, alle Arbeitsgruppen, die sich mit GOME-1 Ozonprofilretrieval befassen, zusammen zu bringen. Im Rahmen der GOPWG sollen Ergebnisse der unterschiedlichen Verfahren vorgestellt und verglichen werden. Hierzu wurde ein erster Testdatensatz konstruiert, der den einzelnen Arbeitsgruppen GOME-Spektren und *a priori* Daten bestehend aus *a priori* Ozonprofilen, Temperatur- und Druckprofilen zur Verfügung stellt. Anhand dieser Daten wurden mit den unterschiedlichen Methoden Ozonprofile berechnet. Bei dem verwendeten GOME-Datensatz handelt es sich um ca. 60 Pixel aus dem Jahre 1997, die sich auf drei verschiedene Standorte verteilen (Tab. 11.1) und mit den dortigen Ozonsonden- und Lidar-Messungen kolloziert wurden.

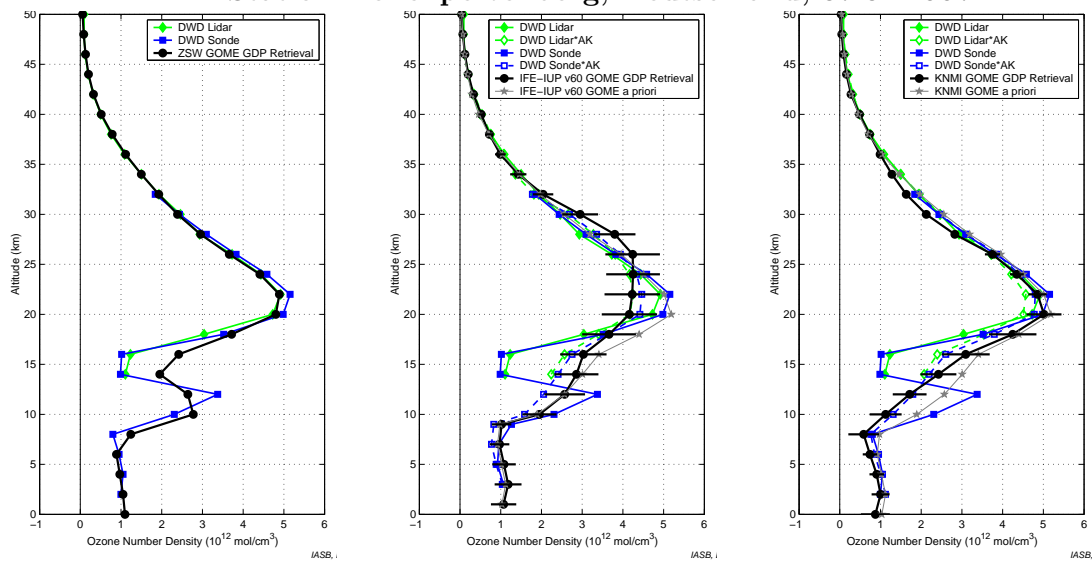
Tabelle 11.1: *Vergleichs-Datensatz der GOME Ozone Profile Working Group, Phase 1.*

Ort	geogr. Breite	geogr. Länge	Anzahl der Profile
Hohenpeißenberg	47.78°N	11.01°E	20
Mauna Loa	19.48°N	155.61°W	17
Lauder	45.04°S	169.68°E	24

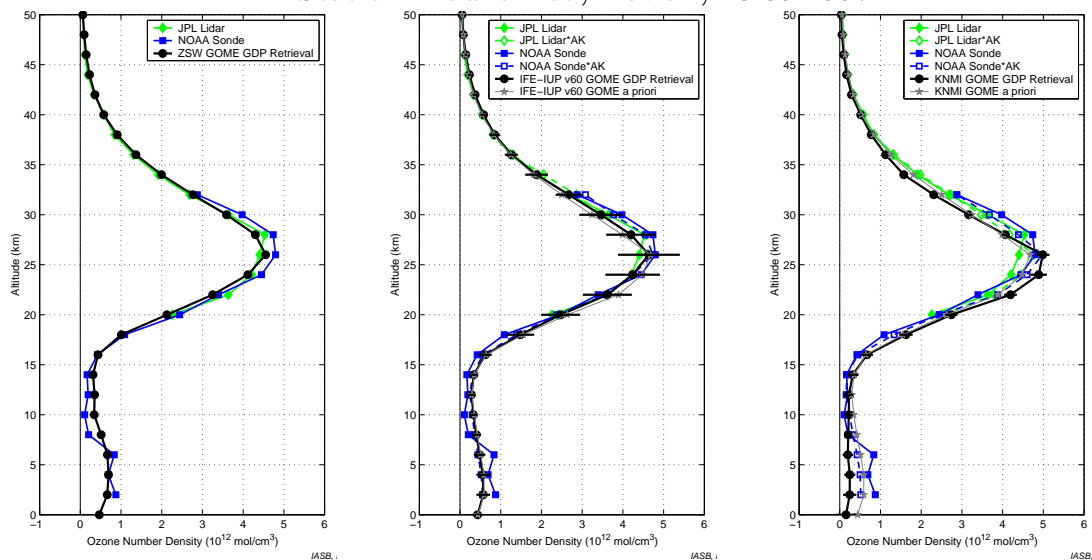
Die Daten aller Beteiligten wurden von J.-Ch. Lambert und V. Soebijanta am *Belgian Institute for Space Aeronomy* (IASB), Brüssel, ausgewertet. Ein Vergleich einiger Ozonprofile ist in Abb. 11.3 dargestellt, aus Platzmangel können hier jedoch nicht die Ergebnisse aller sieben beteiligten Gruppen gezeigt werden. Weiterhin ist anzumerken, dass die verwendeten NNORSY-Daten auf einem Belernungdatensatz beruhen, der nur das Jahr 1997 umfaßt und nicht mehr ganz aktuell ist. Allerdings fand keine der drei dargestellten Stationen bei der Belernung Verwendung. Bei FURM Version 6.0 handelt es sich ebenfalls um vorläufige Daten – die Version wird derzeit noch validiert und verbessert [S. Tellmann, *persönliche Mitteilung*, 2001]. Das KNMI verwendet ebenfalls OE für die Profilbestimmung [van der A, 1999; van der A *et al.*, 2001], jedoch mit einem anderen Vorwärtsmodell [van Oss & Spurr, 2001].

Bei den dargestellten Plots für IUP und KNMI sind gestrichelt die mit den AKs der Retrievals gefalteten Sonden- und Lidarprofile eingezeichnet. Für Hohenpeißenberg zeigt sich hierbei, dass die vertikale Auflösung dieser gefalteten Profile nicht ausreicht, um die laminare Struktur bei 12 km Höhe zu erfassen. Das NNORSY-Profil zeigt jedoch ein deutliches Sekundärmaximum in etwa dieser Höhe. Unter der Voraussetzung, dass die OE-Verfahren die physikalische GOME-Spektralinformation optimal nutzen, ist die bessere Vertikalaufklärung von NNORSY nur durch die Nutzung statistischer Information zu erklären: Das neuronale Netz „weiß“, dass derartige Laminae an entsprechender Stelle und Jahreszeit auftreten können, und „achtet“ auf dementsprechende Anzeichen in den GOME-Spektren und T-Profilen.

Station Hohenpeißenberg, Deutschland, 07.04.1997



Station Mauna Loa, Hawaii, 18.09.1997



Station Lauder, Neuseeland, 26.10.1997

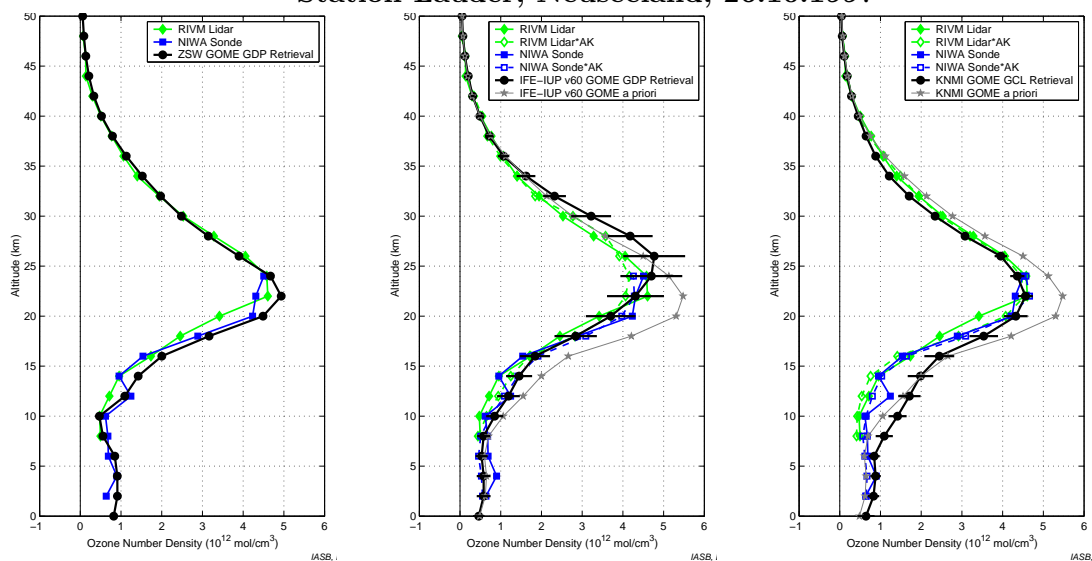


Abbildung 11.3: Vergleich von Ozonprofilen im Rahmen der GOME Ozone Profile Working Group (Auszug). Dargestellt sind die Profile für NNORSY, FURM 6.0 und das NRT-System des KNMI, Niederlande [Alle Plots: V. Soebijanta, IASB, 2002].

Die anderen beiden dargestellten Szenarien zeigen, dass NNORSY auch im äquatorialen Bereich und in der SH konkurrenzfähige Profile liefert. Diese Tendenz zeigte sich auch in den von der Gruppe durchgeführten statistischen Analysen (nicht dargestellt).

Fazit

NNORSY liefert im Vergleich mit den Bodenstationen Hohenpeißenberg und Syowa stabile Ozon-Teilsäulen, die die Ozodynamik auch bei schnellen Veränderungen des Ozonfeldes gut erfassen. Die besten Ergebnisse erzielt das System zwischen 10 und 20 km Höhe, am problematischsten ist wie zu erwarten die Troposphäre, insbesondere über der Antarktis. Die im Rahmen der GOPWG durchgeführten Vergleiche bestätigen die gute Qualität und hohe Vertikalauflösung von NNORSY im Vergleich zu anderen Retrievalverfahren.

Kapitel 12

Integrierte Ozonsäulen

Die bisherigen Untersuchungen des NNORSY-Ozonprofilretrievals beschränkten sich entweder auf einzelne Ozonsondenstationen, oder auf die mehr oder weniger global verteilten Kollokationen in Trainings- und Testdatensatz. Eine weitere Möglichkeit, das globale Verhalten des Systems zu überprüfen, besteht im Vergleich der integrierten NNORSY-Ozonprofile¹ mit verfügbaren Gesamttozonfeldern, hier von GDP und TOMS. Dies erlaubt gleichzeitig einen genaueren Blick auf das in Kapitel 6 vorgestellte NNORSY-Gesamttozonretrieval.

Da eine ausführliche Untersuchung des kompletten Datenbestandes zu umfangreich wäre, sollen an dieser Stelle nur eine typische Frühjahrs- und eine Herbstsituation am Beispiel der Monatsmittel von April und Oktober 1999 betrachtet werden. Dass diese beiden Monate keine Spezialfälle darstellen, ist anhand der Zeitreihen in den Abbildungen 6.6, 6.7, 6.9 und 11.1 erschießbar.

Abb. 12.1 zeigt die beiden Monatsmittel aus dem NNORSY-Gesamttozonretrieval. Die restlichen sechs Ozonkarten entstanden durch Differenzbildung der verschiedenen GOME-Retrievals zum korrigierten TOMS-Ozonfeld (vgl. Abschn. 2.2.1). Die Differenzen wurden für jeden Tag des Monats einzeln aus den Ozon Level 3 Tageskarten in TOMS-Auflösung (288×180 Felder) bestimmt, und die gültigen Differenzwerte für jeden Rasterpunkt über alle Tage gemittelt.

Zunächst kann festgehalten werden, dass die globalen, flächengewichteten Biases gegenüber TOMS für alle drei GOME-Methoden nur wenige DU betragen, wobei – wie bereits weiter oben erwähnt – trotz der TOMS-Korrektur die GDP Ozonwerte systematisch kleiner als die des TOMS sind. Des weiteren enthalten die Vergleiche **GDP–TOMS** zwei Artefakte, deren Ursachen nicht ganz klar sind: Zum einen ist im April eine systematische negative Abweichung in der Südhemisphäre erkennbar, deren Grenze genau am Äquator verläuft. Dies legt nahe, dass es sich um eine von den unterschiedlichen verwendeten Klimatologien hervorgerufene Erscheinung handelt.

Zum anderen tritt im Oktober 1999 eine deutliche Veränderung der GDP–TOMS Differenzen für alle GOME-Orbits auf, die den Äquator zwischen etwa 70°O und 180°O überqueren. Dieser Effekt ist bereits für Tagesmittelwerte deutlich erkennbar, und tritt systematisch auch in anderen Monaten auf. Die Grenze verläuft genau am Längengrad jener Stelle nördlich von Indien, an der aus kommunikationstechnischen Gründen keine GOME-Daten empfangen werden [*W. Thomas, persönliche Mitteilung, 2001*]. Wenn es sich hier um GDP/TOMS-Unterschiede in der Erzeugung von Level 3 Tageskarten handelt, z. B. durch die unterschiedliche zeitliche Zuordnung von Pixeln bzw. Orbits, ist nicht einzusehen, warum sich der Effekt nicht innerhalb eines Monats herausmittelt. Ein längenabhängiger Algorithmusfehler o. ä. ist jedoch schwer vorstellbar. Was auch immer die Ursache, trat dieser Effekt bisher nur bei GDP-Ozondaten auf, obwohl diese nicht anders behandelt wurden

¹Hierbei wird die unterste Profilschicht (0.5–1.5 km) linear bis zum Boden extrapoliert.

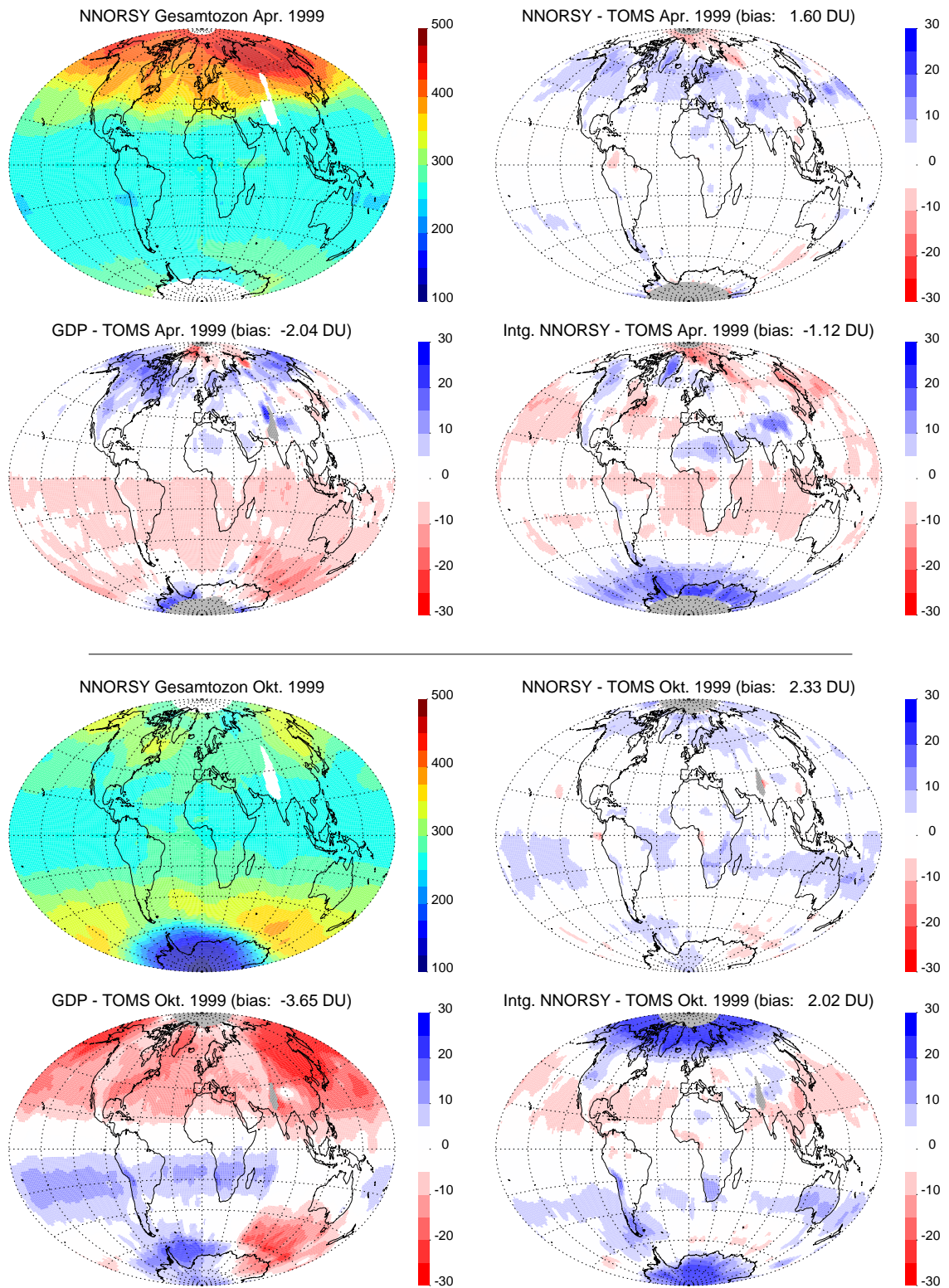


Abbildung 12.1: Vergleich der Monatsmittelwerte von NNORSY und GDP v2.7 Gesamtozonretrieval, sowie integrierten NNORSY Ozonprofilen. Als Referenz dienten hierbei korrigierte TOMS v7 Ozondaten. Die NNORSY-Ozonzfelder für April und Oktober sind zur Übersicht mit dargestellt. Alle Angaben in [DU].

als TOMS, NNORSY und – in früheren Untersuchungen – TOVS, so dass er an dieser Stelle einem Fehler des GDP-Produktes zugeschrieben werden muss. Man beachte, dass der Effekt anhand zonaler Mittelwerte wie in Abb. 6.9 oder in [Corlett & Monks, 2001] verwendet nicht zu entdecken ist, und zudem zonale Biases teilweise kompensiert.

Die Differenzen zwischen **NNORSY Gesamtozon** und TOMS zeigen hingegen eine überwiegend zonale Struktur, die lediglich in Regionen hoher Ozondynamik nennenswert längenabhängige Variationen aufweist. Der absolute Bias gegenüber TOMS erreicht hier lokal selten Werte von 15 DU, und überschreitet auch in anderen Monaten fast nie die 20 DU Grenze. Allerdings treten die größten Abweichungen dann in zeitlicher Nähe zu den Equinoxen und an der polaren Tag/Nacht Schwelle auf, wie in den zonalen Mittelwerten in Abb. 6.9 zu sehen.

Interessanter sehen die Vergleiche zwischen TOMS und **integrierten NNORSY Ozonprofilen** aus, insbesondere da hier keinerlei TOMS-Information in die Belernung einfließt. Auch hier sind die globalen Biases minimal, jedoch ist die lokale Variation stärker ausgeprägt als bei NNORSY Gesamtozon und besser verstanden als beim GDP. Drei Effekte fallen in dieser Fallstudie besonders ins Auge:

1. **Sonnenzenitwinkel:** Probleme des Ozonprofilretrievals bei hohem Sonnenzenitwinkel konnten bereits anhand der Zeitreihen in Abb. 6.7b und c vermutet werden, und zeigen sich hier in flächenhafter Ausdehnung. Diese Probleme wurden in Abschn. 11.2 bereits ausführlich diskutiert. Der dort beobachtete, positive Offset der unteren Profilhälfte scheint zum großen Teil für die Überschätzung des Gesamtozons verantwortlich zu sein. Zu beachten ist allerdings auch, dass die hier beobachteten Biases nicht anhand der Testkollokationen nachweisbar sind (Abb. 7.9). Sollten demnach sowohl diese Kollokationen als auch das korrigierte TOMS das wahre Ozonfeld zuverlässig repräsentieren, ist zu folgern, dass der Testdatensatz hier am Rande des gültigen Latitudenbereiches offensichtlich ein Repräsentationsproblem hat (vgl. Abschn. 10.1). Dies ist auch insofern wahrscheinlich, als dieser Latitudenbereich während eines Jahres teilweise in der Polarnacht versinkt, so dass trotz der vielen Kollokationen die Ozonfelddynamik nahe des polaren Terminators im Trainingsdatensatz chronisch unterrepräsentiert ist.
2. **Oberflächenalbedo:** Ein ähnliches Problem entsteht in der Troposphäre über den Ozeanen, da hier praktisch keine Sondenmessungen vorliegen, und die Okkultationsprofile sehr unzuverlässig sind. NNORSY wird hier die nur aus landgestützten Sondaufstiegen bekannte, troposphärische Ozonverteilung auf die Ozeane extrapolieren. Zusätzlich ist mit Sicherheit die über Land erlernte, implizite Wolken- bzw. Albedokorrektur für die troposphärische Ozonsäule über den Ozeanen aufgrund der niedrigen Oberflächenalbedo fehlerhaft: Es tritt der umgekehrte Effekt wie in Abb. 11.1b auf, nämlich dass das Netz die Mischalbedo von Meer und Wolken als zu niedrigen Wolkenbedeckungsgrad interpretiert, und dementsprechend eine zu geringe Troposphärenkorrektur addiert (vgl. Diskussion in Abschn. 11.2).

Wiederum zeigen sich keine nennenswerten Biases in den Testkollokationen. Da die Abdeckung der Ozeanflächen mit Okkultationsmessungen jedoch sehr gut ist, sollte hier in der Stratosphäre kein Repräsentationsproblem bestehen. Die beobachteten Biases wären demnach fast vollständig dem troposphärischen Gesamtozonbeitrag zuzuschreiben.

Für diejenigen Flächen mit extrem hoher Albedo unterschätzt das Netz aus teilweise den selben Gründen die Gesamtozonsäule, wie im April 1999 über der Sahara, dem Himalaya, Grönland und der Antarktis zu erkennen. Jedoch können hier auch der SZA und die Bodenelevation eine Rolle spielen.

3. Bodenelevation: Die drei letzten der gerade erwähnten Gebiete haben außer der hohen Albedo auch die große Höhe über dem Meeresboden gemeinsam. Diese wird bei der Integration der Profile über die geopotenzielle Höhe nicht berücksichtigt, d. h. Teile des Ozonprofils können sich „unter der Erdoberfläche“ befinden. Die resultierenden Fehler können durchaus für etwa 5–10 DU der Überschätzung verantwortlich sein; Eine Korrektur wurde bislang allerdings noch nicht implementiert. Das Verwenden der Elevation als zusätzlichen Eingangswert für die Netzwerkbelernungen macht hingegen wenig Sinn, da kaum Stationen in großen Höhen vorhanden sind und bei (stratosphärischen) Okkultationsprofilen die Bodenhöhe wenig relevant ist.

Alles in allem stimmen die integrierten NNORSY-Profile global in den meisten Fällen bis auf 0–5% mit dem korrigierten TOMS-Feld überein. Lokal werden in Situationen mit niedrigen Ozonsäulen und hohem Sonnenzenitwinkel bis zu 10–20% Abweichung beobachtet, andere lokale Maxima liegen im Bereich von 5–10%. In allen Fällen ist ein großer Fehleranteil in der Troposphäre zu suchen.

Kapitel 13

Aspekte der operationellen Anwendung

Die bisherige Diskussion des NNORSY-Verfahrens beschränkte sich auf die Offline-Version der Software: Die einzelnen Prozessschritte werden anhand von Skriptdateien und Parameterfiles konfiguriert und per Hand gestartet. Um zu demonstrieren, dass das Verfahren auch im Echtzeitbetrieb einsetzbar ist, wurde zusätzlich ein Prototyp für operationelles Retrieval entwickelt.

13.1 Unterschiede zur Offline-Version

Für die operationelle Anwendung sind nur die Teile des NNORSY-Paketes notwendig, die das Preprocessing und die NN-Vorwärtsrechnung ausführen. Kollokation und Belernung neuronaler Netze können weiterhin offline durchgeführt werden.

Dabei ist zu beachten, dass zusätzliche Eingangsdaten wie insbesondere die UKMO-Temperaturprofile ebenfalls in Echtzeit zur Verfügung stehen müssen, damit die Prozessierung nicht verzögert wird. Alternativ kann ein *first-look*-Produkt ohne die Verwendung zeitkritischer Zusatzinformation erzeugt werden, welches dann z. B. am nächsten Tag mit einem *second-day*-Produkt höherer Qualität ergänzt wird, in das alle notwendigen Zusatzinformationen einfließen.

Die Prozessierungsgeschwindigkeit ist aufgrund der extrem schnellen Retrievals mit neuronalen Netzen normalerweise kein Hemmnis für die Echtzeitanwendung. Ein Flaschenhals entsteht hier allenfalls bei der Vorverarbeitung, nicht beim Retrieval selber: Die im Zeitraum dieser Arbeit zur Verfügung stehende Version 2.0 der GDP-Extraktorsoftware kann keine Binärdaten ausgeben, sondern nur ASCII-Daten. Dabei entstehen pro Orbit etwa 280 MB an formatiertem Output, der von NNORSY wieder eingelesen und in binäre Daten umgewandelt werden müssen. Die Interpolation auf das für die Netzeingabe verwendete Standard-Wellenlängengitter erfordert ebenfalls mehr Rechenleistung als die Retrieval-Vorwärtsrechnung des neuronalen Netzes.

13.2 Implementation für den GOME NRT-Service

Das DLR initiierte im Januar 1997 in Zusammenarbeit mit dem IFE/IUP und der ESA einen GOME *near-real-time* (NRT) Service, indem ein GDP-System auf der Station Kiruna (Schweden) installiert wurde. Eine weitere NRT-Station wurde im November 2001 in Gatineau (Kanada) in Betrieb genommen, in enger Kooperation mit IFE/IUP, ESA und Natural Resources Canada (NRCan). Diese Station liefert GDP-Retrievals von GOME-Orbits, die von den kanadischen Empfangsanlagen in Gatineau and Prince Albert aufgefangen werden.

Um eine globale Abdeckung zu erreichen fehlt damit nur ein weiterer Orbit, der wird von der Station Maspalomas (Gran Canaria, Spanien) empfangen wird.

Ein NNORSY *first-look* Prozessierungssystem (ohne Verwendung von T-Profilen) wurde an das Ende der GDP-Prozesskette in Kiruna, Gatineau and Maspalomas angefügt, d. h. die Produkte werden mit globaler Abdeckung ausgeliefert. Dabei werden die Level 2 Daten mit dem restlichen Datenstrom an das DLR gesendet, und dort in grafische Darstellung überführt. Die resultierenden Bilder können dann auf der GOME NRT Seite des DLR (N-GONG) ¹ abgerufen werden. Obwohl an den drei Empfangsstationen nur jeweils ein 100 MHz-Prozessor für NNORSY zur Verfügung steht, können alle GOME-Pixel in Echtzeit abgearbeitet werden, was die Leistungsfähigkeit des Systems anschaulich demonstriert.

13.3 Zeitliche Extrapolation und Nachtrainieren

Die im Rahmen dieser Arbeit durchgeführten Untersuchungen bezogen sich generell auf den Zeitraum, aus dem auch die Trainings- und Testkollokationen stammten. Im operationellen Betrieb wird der Belernungszeitraum jedoch immer einige Wochen bis Monate zurückliegen, da die für Kollokationen benötigten Sonden-, Gesamt ozon- und Okkultationsmessungen nur mit Verzögerung bereitgestellt werden. Solange die GOME-Spektren sich nicht zu stark mit der Zeit verändern, führt dies nur zu einer mäßigen Zunahme der Retrievalfehler, die im konkreten Fall mit Hilfe aktueller Ozonsondenprofile von Stationen, zu denen direkter Kontakt besteht, quantifiziert werden könnten. Im Abstand von einigen Monaten würde dann die Kollokationsdatenbank mit allen neuen Messdaten aktualisiert und das neuronale Netz durch Nachtrainieren angepasst.

Bedingt durch den Ausfall der Gyros des Satelliten ERS-2, auf dem das GOME Instrument installiert ist, werden seit Anfang des Jahres 2001 nur noch sehr sporadisch vollständige Sonnenspektren aufgenommen: Die durch den gyrolosen Navigationsmodus bedingte Taumelbewegung des Satelliten führt zu einer periodisch unvollständigen Ausleuchtung der Diffuserplatte, und damit zu von Tag zu Tag stark schwankenden Intensitäten, so dass die meisten Sonnenspektren von der GDP-Software verworfen werden. Ein weiteres Problem ist, dass die Sonnenspektren seit etwa 1998 unterschiedlich von den Erdspektren degradieren, was auf eine Degradation des Scanspiegels zurückgeführt wird [Aben *et al.*, 2000; Richter & Wagner, 2001]. Die GDP-Extraktorsoftware kann zwar eine Degradationskorrektur durchführen, diese wird aber auf Sonnen- und Erdspektren gleichermaßen angewandt und hat deshalb bei der Bildung von sonnennormierten Spektren keinen Effekt.

Die NNORSY-Ozonbestimmung erfolgt unter der Annahme, dass die Degradation des Instruments in zeitlich kontinuierlicher Weise fortschreitet. Dann kann das neuronale Netz bei Verwendung eines linearen Zeitparameters als Eingabewert beim Training eine entsprechende zeitabhängige Funktion erlernen, die die Degradation kompensiert. Für Daten vor 2001 kann dies anhand der Konstanz der Fehler im Vergleich zu unabhängigen Kollokationen mit Ozonprofilen nachgewiesen werden (vgl. Abschn. 11).

Das Problem bei der Belernung mit GOME-Daten des Jahres 2001 und danach besteht darin, dass das neuronale Netz die zeitlich kontinuierliche Degradation nicht von den sprunghaft veränderlichen Effekten unterscheiden kann, die durch die Verwendung veralteter Sonnenspektren hervorgerufen werden. Diese Effekte sind von zweierlei Art:

- Die langsame, kontinuierliche Veränderung der Sonnenspektren aufgrund von Prozessen auf der Sonne wird durch die regelmäßigen GOME-Sonnenmessungen quantisiert. Bei täglicher Messung sind die Quantisierungseffekte vernachlässigbar, aber wenn die Messintervalle wie in 2001 mehrere Monate lang werden, ist mit Sprüngen zu rechnen.

¹http://auc.dfd.dlr.de/GOME_NRT

- *Richter & Wagner* [2001] stellten fest, dass die Diffuserplatte des Instrumentes eine spektrale Signatur induziert, die in komplexer Weise vom Einstrahlwinkel der Sonne abhängt, der im Jahresverlauf variiert. Diese Variation spielt normalerweise für GOME Gesamt ozon keine Rolle, kann aber die mittels DOAS bestimmten schrägen Säulen von NO_2 und anderen schwachen Absorbern um bis zu 100% schwanken lassen. Da Ozonprofile ebenfalls eine genaue Kalibrierung benötigen, ist auch hier mit Beeinträchtigung zu rechnen. Wie bei Punkt 1) kann NNORSY diesen zeitabhängigen Effekt kompensieren, wenn eine ausreichende, regelmäßige Abtastung der Funktion vorliegt. Dies ist ab 2001 jedoch nicht mehr der Fall.

Ein operationelles System muss diese Fehlerquellen geeignet berücksichtigen. Dies könnte z. B. geschehen, indem die schlecht ausgeleuchteten Sonnenspektren nicht verworfen werden, sondern zusammen mit einer Zeitvariablen, die die Periodizität der Ausleuchtung widerspiegelt, als Eingangsdaten Verwendung finden.

Fazit

Im Rahmen dieser Arbeit wurde die Eignung von NNORSY für den Echtzeiteinsatz untersucht, und anhand der Implementation eines Prototyps in die NRT-Prozesskette dreier GOME-Empfangsstationen erfolgreich nachgewiesen. Durch Berücksichtigung von instrumentellen Effekten neueren Datums und die Zuführung möglichst aktueller Temperaturprofile ließe sich die Datenqualität der operationellen NNORSY-Ozonprofile und -säulen noch weiter verbessern.

Kapitel 14

Zusammenfassung und Ausblick

Die Methode der neuronalen Netze, wie sie bei NNORSY angewandt wird, hat sich im Rahmen dieser Arbeit als geeignet erwiesen, um den inversen atmosphärischen **Strahlungstransport** im Bezug auf Ozon zu modellieren. Sie ermöglicht, aus einem GOME-Satellitenspektrum mit beigefügten geophysikalischen und statistischen Informationen ohne Umwege auf ein atmosphärisches Ozonprofil bzw. den Gesamtzongehalt zu schließen. Hierbei wurde deutlich, dass die aus dem Trainingsdatensatz gewonnene *a priori* Information ausreicht, um die mehrdeutige Strahlungstransfergleichung eindeutig zu lösen, und dass die durch das natürliche Messrauschen von GOME-Spektren und kollozierten Ozonprofilen und im Zusammenspiel mit einer geschickten Wahl der Trainingsdaten vermittelte Regularisierung des Systems die Lösung sehr gut stabilisiert.

Die **Genauigkeit** der erhaltenen Ozonretrievals scheint für Gesamtzoon besser zu sein als die des operationellen GDP v2.7 Ozonproduktes, und ist offenbar nur wenig schlechter als die der TOMS v7 Daten. Zum Training des entsprechenden Netzes gingen allerdings außer den Gesamtzoonmessungen von Bodenstationen auch klimatologisch korrigierte TOMS-Daten in die Belernung ein, so dass letztere Ergebnisse mit Vorsicht zu bewerten sind. Da sowohl TOMS als auch der GDP in naher Zukunft zu einer neuen, verbesserten Datenversion übergehen werden, sollte sich demnächst eine gute Möglichkeit zur unabhängigen Validierung ergeben.

Für die Ozonprofilbestimmung wurde ein Kollokationsdatensatz aus GOME-Pixeln mit Ozonsonden, sowie SAGE II, HALOE und POAM III Ozonprofilen zusammengestellt. Um eine für die Netzwerkbelernung geeignete Verteilung und Qualität dieses Datensatzes sicherzustellen, mussten eine Reihe von Homogenisierungsmaßnahmen angewandt werden. Das Retrievalergebnis scheint nach den hier durchgeführten Analysen von der Genauigkeit her mit klassischen Verfahren vergleichbar zu sein, eine genauere Untersuchung im Vergleich mit fünf anderen GOME-Ozonprofilbestimmungsmethoden wird derzeit im Rahmen der GOME Ozone Profile Working Group unternommen.

Ein großer Vorteil von NNORSY besteht darin, dass die absolute **Kalibrierung** der GOME-Spektralwerte nicht bekannt, sondern nur im Trainingsdatensatz gut repräsentiert sein muss. Dies bedeutet insbesondere, dass instrumentelle Effekte wie etwa die Degradation von Sensorkomponenten nicht aufwändig modelliert werden müssen, sondern automatisch vom System erlernt werden, sofern die sie kontrollierenden Parameter (Sensorenalter, Scanwinkel etc.) mit in die Belernung eingehen. Gleichzeitig besteht wenig Gefahr, durch falsche Kalibration Biases in das Retrieval einzuführen: Findet der Belernungsalgorithmus keinen konsistenten Zusammenhang zwischen Eingangsparemetern und Ozon, so wird der Eingang in der Folge meist einfach ignoriert. Auf den hier untersuchten Zeitreihen von 1996 bis Juli 2001 konnten dementsprechend weder für Ozon-Teilsäulen noch für Gesamtzoon instrumentenbedingte Drifts im Vergleich zu Bodenstationen bzw. Ozonsonden festgestellt

werden, obwohl für die GOME-Spektren nur die GDP-Standardkalibration (ohne Degradationsoption) verwendet wurde.

Die Verinnerlichung einer ganzen Reihe von Korrekturen seitens des Belernungsalgorithmus' wirkt sich allerdings an anderer Stelle nachteilig aus, da sich diese Korrekturen nicht mehr separieren lassen und somit die Analyse der Fehler einzelner Retrievals stark erschwert wird. Zur besseren **Fehlercharakterisierung** wurden daher einige empirische Methoden entwickelt. So läßt sich der *precision*-Fehler recht gut durch künstliches Verrauschen der Eingangsdaten abschätzen, während *accuracy*-Fehler durch Belernung eines separaten Fehlernetzwerks zugänglich werden. Interessanterweise führte die Verwendung der Jacobi-Matrix des neuronalen Netzes nicht zu sinnvollen Fehlerwerten, vermutlich weil die nichtlineare Modellierung von Eingangsdatenkorrelationen seitens des Netzes im Widerspruch zu der linearen Näherung durch Jacobi- und Kovarianzmatrizen steht.

Die neuronale Jacobi-Matrix verdeutlichte jedoch wiederum ein Phänomen, welches bereits bei Sensitivitätsanalysen des *precision*-Fehlers sowie der Architektur des Netzwerkes beobachtet wurde, nämlich die fast vollständige Verschleierung der **Höhenzuordnung** zwischen Eingangsdaten und Ozonprofilen. Die in diesem Sinne ermittelten Fehlerprofile bzw. Beitragsfunktionen von GOME-Spektralwerten, Temperaturprofil, und anderen Größen wie z. B. SZA oder Scanwinkel ließen keine oder nur sehr grobe Ähnlichkeit mit Vorhersagen aus der Strahlungstransferrechnung erkennen. Es bleibt zu folgern, dass hier die Vermischung von physikalischen und statistischen Zusammenhängen innerhalb des Netzes so gründlich ist, dass sie im Nachhinein nicht wieder separiert werden kann. Allerdings ist es durchaus möglich, anhand des RMS-Gesamtfehlers auf Trainings- und Testdatensatz Aussagen über den Informationsgehalt spezifischer Eingangsparameter zu treffen, und deren „physikalischen“ Beitrag zur Modellierung anhand der vermittelten Generalisierungsfähigkeit abzuschätzen.

In diesem Zusammenhang fiel auf, dass die größte Verringerung des Retrieval-RMSE, um etwa 8–10%, von den verwendeten UKMO **Temperaturprofilen** ausgeht. Außerdem stabilisieren die T-Profile das Retrieval, so dass die *precision*-Fehler abnehmen. Interessanterweise finden sich in der Literatur nur sehr wenige Studien, die eine direkte Verknüpfung von Temperaturprofil und Ozonprofil versuchen. Bei klassischen Retrievalverfahren wie z. B. FURM geht das T-Profil nur indirekt über die Strahlungstransportrechnung ein. Obwohl natürlich insbesondere die stratosphärische Temperaturverteilung i. W. durch die Absorption energiereicher Strahlung durch Ozon hervorgerufen wird und daher durch das RTM Berücksichtigung findet, stellt sich angesichts der hier gefundenen Zusammenhänge die Frage, ob möglicherweise eine empirische Relation zwischen T- und Ozonprofil gefunden und z. B. zur Konstruktion des *a priori* Ozonprofils genutzt werden kann, wie sie anscheinend das neuronale Netz implizit durchführt. Weiterhin deutet sich hier sowohl für NNORSY als auch für klassische Retrievalverfahren eine spürbare Qualitätsverbesserung an, die durch die Verwendung genauerer T-Profile erreicht werden könnte. Mit NNORSY könnte man sogar noch einen Schritt weiter gehen, etwa indem die Daten eines T-Sounders mit denen eines Ozoninstruments auf derselben Orbitalplattform für eine gemeinsame MLP-Belernung kombiniert werden.

Die direkte Kombination der Temperaturinformation mit klimatologischen Größen und den GOME-Spektren bewirkt bei NNORSY außerdem, dass die bei klassischen Verfahren theoretisch mögliche **Vertikalauflösung** in manchen Fällen unterschritten wird. Dies ist ebenfalls eine Folge der direkten Mischung von spektraler, geophysikalischer und statistischer Information. Die genaue Quantisierung der Höhenauflösung erweist sich hingegen aufgrund des oben angedeuteten Fehlens einer eindeutigen Höhenzuordnung als schwierig. Empirische Schätzmethode ergeben als Referenzwert etwa 4–5 km vertikale Auflösung im

mittleren Höhenbereich. Dieser Wert wird durch die beobachtete Fehlerkorrelationslänge in den Retrievalprofilen, sowie durch visuelle Vergleiche des Retrievals mit geglätteten Sondenprofilen bestätigt.

Probleme der Profilvertrievals bestehen derzeit noch um den Nordpol herum, global in der Region um die Tropopause, sowie in der Troposphäre über der Antarktis und den Ozeanen. Obwohl auch hier bei Vergleichen mit Ozonsonden Verbesserungen gegenüber der Annahme klimatologischer Ozonprofile erzielt wurden, besteht noch großes Optimierungspotenzial. Durch die geringe Anzahl von – nur innerhalb einzelner Messkampagnen durchgeführten – schiffsgestützten Sondaufstiegsmessungen besitzt das neuronale Netz zudem sehr wenig Information über die ozeanische Troposphäre, und die Validierung der Retrievalergebnisse gestaltet sich als schwierig. Insbesondere in den Tropen macht sich hier ein negativer Bias der Retrievals bemerkbar. Polwärts von 71°N liegen hingegen keine Satellitenmessungen mehr vor, so dass die Korrektur der dort auftretenden hohen SZA und anderer Effekte vom Belernungsalgorithmus aus den Daten einiger weniger Bodenstationen abgeleitet werden muss, und beim Vergleich integrierter Gesamt Ozonwerte mit TOMS zu systematischen, positiven Biases führt. Um den Südpol gleichen die POAM-Daten den Mangel an stratosphärischer Information aus, jedoch führen die speziellen antarktischen Bedingungen (SZA, Wolken, Eis, Elevation) in der Troposphäre zu positiven Biases und hohen Fehlern.

Allgemein läßt sich sagen, dass das Belernungsverfahren immer zunächst die deutlichsten Zusammenhänge zwischen Eingangs- und Ausgangsdaten modelliert. Schreitet der Belernungsprozess weiter fort, werden auch schwächere Signale wie das des troposphärischen Ozons in die Modellierung einbezogen. Dabei besteht jedoch gleichzeitig die Gefahr, dass mehr und mehr Rauschen sowie spezifische Eigenschaften des Trainingsdatensatzes abgebildet werden, was die Generalisierungsleistung des Systems verschlechtert und die Fehler auf dem unabhängigen Testdatensatz wieder steigen läßt. Gemäß den hier gewonnen Erkenntnissen könnte die **Verringerung der troposphärischen Fehler** am ehesten erreicht werden durch

Verstärkung des Ozonsignals: Wenn das Höhengitter des Retrievals mit der Tropopausenhöhe referenziert wird, verringern sich nach [Bojkov & Fioletov, 1997; Logan, 1999] die beobachteten Ozonwertschwankungen stark, und die Gradienten des Ozonprofiles werden steiler. Damit könnte sich die Belernung mehr auf das eigentliche Signal konzentrieren und müsste nicht in erster Linie die latituden-, jahreszeit- und meteorologieabhängigen Verschiebungen der Tropopausenhöhe modellieren. Dabei wird natürlich ein zuverlässiger, extern berechneter Tropopausenhöhenwert benötigt.

Schwächung anderer Signale: Das Netz könnte durch geeignete Datensatzauswahl mehr auf die kritischen Bereiche spezialisiert werden, indem z. B. nur SH-Daten, oder nur Höhenschichten unter 25 km in die Belernung eingehen. Allerdings muss dabei wiederum sehr sorgfältig ein Übertrainieren vermieden werden, was bei den bisher verwendeten, globalen Datensätzen weniger kritisch ist.

Reduzierung des Rauschens: Nachdem nun das grundsätzliche Verhalten des Systems bekannt ist, könnte sich die explizite Untersuchung von Wolkeneffekten als lohnenswert erweisen, die bislang aufgrund ihrer Komplexität vernachlässigt wurde. Dies gilt insbesondere, da die neue GDP-Version 3.0 demnächst erstmals für alle GOME-Pixel ein zuverlässiges und leicht zugängliches Wolkenprodukt bereitstellt.

Im Bereich der **Eingangsdatenauswahl** (*feature selection*) bestehen ebenfalls generell noch Verbesserungsmöglichkeiten. Bislang wurden analog zu bestehenden physikalischen

Verfahren (normierte) Spektralwerte als Eingangsparameter verwendet. Diese weisen jedoch ein hohes Maß an Redundanz auf. Gelänge es, den tatsächlichen Informationsgehalt dieser Spektren durch die Anwendung spezieller Transformationen – wie etwa Fourierzerlegung oder PCA – in eine geringere Anzahl von Parametern zu destillieren, ohne das Rauschverhalten des neuronalen Netzes zu verschlechtern, könnte der Belernungserfolg verbessert werden. Probleme können hier insbesondere dadurch entstehen, dass gerade kleine Variationen des Spektrums oft wertvolle Information enthalten, die bei der Transformation zusammen mit dem Rauschen entfernt wird [Bishop, 1995a]. In diesem Zusammenhang weisen Berendes et al. [1999] allerdings darauf hin, dass das Gebiet *feature selection* inzwischen den Umfang einer Sub-Disziplin erreicht hat [Richards, 1993], systematische Untersuchungen über den Nutzen verschiedener Aufbereitungsverfahren für Eingangsdaten aber noch weitgehend ausstehen.

Prinzipiell ließen sich mit NNORSY auch **andere Spurengase** nachweisen, allerdings stellt sich hier das Problem der Vergleichsmessungen, da für keines davon ein Äquivalent des Ozon-Bodenmessnetzes besteht. Sofern die Instrumenteneigenschaften gut genug bekannt sind, könnten neuronale Netze auf der Basis simulierter Spektren trainiert werden, wie dies in der Literatur häufig zu finden ist. Für GOME empfiehlt sich dieses Vorgehen allerdings nicht unbedingt, da insbesondere die Sensordegradation schwer zu simulieren ist, und die Spurengassignale in den Spektren naturgemäß viel schwächer sind als die des Ozons.

Umgekehrt sollte sich jedoch die Anpassung von NNORSY auf die **Ozonbestimmung mit anderen Sensoren** relativ einfach bewerkstelligen lassen, da der Arbeitsaufwand für die Erstellung eines Kollokationsdatensatzes angesichts der vielen bereits gesammelten und aufbereiteten Messungen wesentlich geringer sein dürfte, und ein exaktes Verständnis aller Instrumenteneigenschaften für die NNORSY-Methode nicht unbedingt vorausgesetzt wird, sondern vom Verfahren erlernt werden kann. Insbesondere für theoretisch schwer zugängliche Probleme wie die gleichzeitige Auswertung von Limb- und Nadirmessungen des SCIAMACHY-Sensors oder die oben erwähnte Kombination von Temperatur- und Ozonsondungs könnte sich die Technik neuronaler Netze als interessante Alternative zur klassischen Auswertung anbieten.

Anhang

Anhang A

Liste der verwendeten Bodenstationen

Die folgende Tabelle listet alle in der Arbeit verwendeten WMO-Stationen auf, deren Messwerte ausnahmslos dem WOUDC-Archiv entnommen wurden. In der vorletzten Spalte bezeichnet „x“ Gesamt Ozon-Trainingsstationen und „Test“ die Teststationen. Die letzte Spalte gibt „x“ für Ozonprofil-Belehnungsstationen an (Trainings- und Testdaten wurden zufällig ausgewählt) und „Val“ für Validationsstationen. Die am Ende der Tabelle aufgeführten virtuellen Stationen wurden wie in Absch. 6.1 beschrieben aus modifizierten TOMS-Daten erstellt.

WMO	Name	Land	Latitude	Longit.	O ₃ -Säule	O ₃ -Profil
1	LEOPOLDVILLE	COG	-4.30	15.55	-	x
2	TAMANRASSET	DZA	22.80	5.52	x	x
3	ALMA-ATA	KAZ	43.23	76.93	x	x
4	BANARAS	IND	25.00	83.00	-	x
5	DIKSON ISLAND	RUS	73.50	80.23	-	x
6	HABBANIYA	IRQ	33.37	43.57	-	x
7	KAGOSHIMA	JPN	31.58	130.57	x	x
8	KODAIKANAL	IND	10.23	77.47	x	x
9	MOUNT ABU	IND	24.60	72.70	-	x
10	NEW DELHI	IND	28.65	77.22	x	x
11	QUETTA	PAK	30.11	66.57	x	-
12	SAPPORO	JPN	43.06	141.33	x	x
14	TATENO	JPN	36.06	140.10	x	x
16	VLADIVOSTOK	RUS	43.12	131.90	x	-
18	ALERT	CAN	82.50	-62.32	-	x
19	BISMARCK	USA	46.77	-100.75	x	-
20	CARIBOU	USA	46.87	-68.03	x	-
21	EDMONTON (STONY PL.)	CAN	53.55	-114.10	x	x
24	RESOLUTE	CAN	74.72	-94.98	x	x
27	BRISBANE	AUS	-27.42	153.12	x	-
30	MINAMITORISHIMA	JPN	24.30	153.97	x	-
31	MAUNA LOA	USA	19.53	-155.57	x	-
35	AROSA	CHE	46.78	9.68	x	-
36	CAMBORNE	GBR	50.22	-5.32	x	-
40	HAUTE PROVENCE	FRA	43.93	5.70	x	-
42	ST. PETERSBURG	RUS	59.97	30.30	x	-
43	LERWICK	GBR	60.13	-1.18	x	x
45	MESSINA	ITA	38.20	15.55	x	-
50	POTSDAM	DEU	52.22	13.05	x	-

Fortsetzung nächste Seite

WMO	Name	Land	Latitude	Longit.	O ₃ -Säule	O ₃ -Profil
51	REYKJAVIK	ISL	64.13	-21.90	x	-
53	UCCLE	BEL	50.80	4.35	x	-
55	VIGNA DI VALLE	ITA	42.08	12.22	x	-
65	TORONTO	CAN	43.78	-79.47	x	-
67	BOULDER	USA	40.03	-105.25	Test	x
68	BELSK	POL	51.84	20.79	x	-
73	AHMEDABAD	IND	23.03	72.65	x	-
74	VARANASI	IND	25.32	83.03	x	-
76	GOOSE BAY	CAN	53.30	-60.36	x	x
77	CHURCHILL	CAN	58.75	-94.07	x	x
79	TALLAHASSEE	USA	30.40	-84.35	x	-
82	LISBON	PRT	38.77	-9.15	x	-
84	DARWIN	AUS	-12.42	130.88	x	-
85	IRKUTSK	RUS	52.26	104.35	x	-
86	FEODOSIJA	UKR	45.03	35.38	x	-
87	KIEV	UKR	50.40	30.45	x	-
89	NY ALESUND	NOR	78.93	11.88	x	-
91	BUENOS AIRES	ARG	-34.58	-58.48	x	-
96	HRADEC KRALOVE	CZE	50.18	15.83	x	-
99	HOHENPEISSENBERG	DEU	47.80	11.02	Test	Val
100	BUDAPEST-LORINC	HUN	47.43	19.18	x	-
101	SYOWA	JPN	-69.00	39.58	x	Val
105	FAIRBANKS	USA	64.82	-147.87	x	-
106	NASHVILLE	USA	36.25	-86.57	x	-
107	WALLOPS ISLAND	USA	37.93	-75.48	x	x
109	HILO	USA	19.72	-155.07	-	x
111	AMUNDSEN-SCOTT	ATA	-89.98	-24.80	x	-
112	BOLSHAYA ELAN	RUS	46.92	142.73	x	-
113	DUSHANBE	TJK	38.58	68.78	x	-
115	SAMARA (KUIBYSHEV)	RUS	53.25	50.45	x	-
116	MOSCOW	RUS	55.75	37.57	x	-
117	MURMANSK	RUS	68.97	33.05	x	-
118	NAGAEVO	RUS	59.58	150.78	x	-
119	ODESSA	UKR	46.48	30.63	x	-
120	OMSK	RUS	54.93	73.40	x	-
121	RIGA	LVA	57.19	24.25	x	-
122	EKATERINBURG	RUS	56.80	60.63	x	-
123	YAKUTSK	RUS	62.08	129.75	x	-
128	KARARGANDA	KAZ	49.80	73.13	x	-
129	PECHORA	RUS	65.12	57.10	x	-
130	PETROPAVLOVSK/KAMCH.	RUS	52.97	158.75	x	-
132	SOFIA	BGR	42.82	23.38	x	-
142	IGARKA	RUS	67.47	86.57	x	-
143	KRASNOYARSK	RUS	56.00	92.88	x	-
144	MARKOVO	RUS	64.68	170.42	x	-
147	SEMIPALATINSK	KAZ	50.35	80.25	x	-
148	VITIM	RUS	59.45	112.58	x	-
152	CAIRO	EGY	30.08	31.28	x	-
153	VORONEZ	RUS	51.70	39.17	x	-
156	PAYERNE	CHE	46.49	6.57	-	x
158	CASABLANCA	MAR	33.57	-7.67	x	-
159	PERTH	AUS	-31.92	115.95	x	-
165	OSLO	NOR	59.91	10.72	x	-
174	LINDENBERG	DEU	52.21	14.12	x	x
175	NAIROBI	KEN	-1.27	36.80	x	x
183	ATIRAY (GUREV)	KAZ	47.03	51.85	x	-

Fortsetzung nächste Seite

WMO	Name	Land	Latitude	Longit.	O ₃ -Säule	O ₃ -Profil
184	LWOW	UKR	49.82	23.95	x	-
187	POONA	IND	18.53	73.85	x	x
190	NAHA	JPN	26.20	127.68	x	x
191	SAMOA	ASM	-14.25	-170.56	x	x
192	MEXICO CITY	MEX	19.33	-99.18	x	-
199	BARROW	USA	71.32	-156.60	x	-
200	CACHOEIRA PAULISTA	BRA	-22.68	-45.00	x	-
201	SESTOLA	ITA	44.22	10.77	x	-
205	THIVANDRUM	IND	8.48	76.95	-	x
208	XIANGHE	CHN	39.98	116.37	x	-
209	KUNMING	CHN	25.03	102.68	x	-
213	EL ARENOSILLO	ESP	37.10	-6.73	x	-
214	SINGAPORE	SGP	1.33	103.88	x	-
216	BANGKOK	THA	13.67	100.61	Test	-
218	MANILA	PHL	14.63	121.83	x	-
219	NATAL	BRA	-5.84	-35.21	x	x
221	LEGIONOWO	POL	52.40	20.97	-	x
226	BUCHAREST	ROM	44.48	26.13	x	-
232	VERNADSKY FARADAY	UKR	-65.25	-64.27	x	-
233	MARAMBIO	ATA	-64.23	-56.72	x	x
241	SASKATOON	CAN	52.11	-106.71	x	-
242	PRAHA	CZE	50.02	14.45	-	x
245	ASWAN	EGY	23.97	32.78	x	-
252	SEOUL	KOR	37.57	126.95	x	-
253	MELBOURNE	AUS	-37.80	144.97	x	-
254	LAVERTON	AUS	-37.87	144.75	-	x
256	LAUDER	NZL	-45.03	169.68	Test	x
261	THESSALONIKI	GRC	40.52	22.97	x	-
262	SODANKYLA	FIN	67.34	26.51	x	x
265	IRENE	ZAF	-25.56	28.19	x	x
267	SONDRESTROM	GRL	67.00	-50.62	x	-
268	ARRIVAL HEIGHTS	ATA	-77.83	166.67	x	-
271	ARHANGELSK	RUS	64.58	40.50	x	-
274	NIKOLAEVSK-NA-AMURE	RUS	53.15	140.70	x	-
276	TURA	RUS	64.17	100.07	x	-
277	CIMLJANSK	RUS	47.73	42.25	x	-
279	NORRKOEPING	SWE	58.58	16.15	x	-
282	KISLOVODSK	RUS	43.73	42.66	x	-
284	VINDELN	SWE	64.24	19.77	x	-
287	FUNCHAL (MADEIRA IS.)	PRT	32.64	-16.89	x	-
290	SATURNA ISLAND	CAN	48.78	-123.13	x	-
293	ATHENS	GRC	37.98	23.73	x	-
295	MT. WALIGUAN	CHN	36.17	100.53	x	-
301	J.R.C. ISPRA (VARESE)	ITA	45.80	8.63	x	-
305	ROME UNIVERSITY	ITA	41.90	12.52	x	-
308	MADRID	ESP	40.45	-3.72	x	-
311	HAVANA	CUB	23.28	-82.55	x	-
312	KAUNAS	LTU	54.52	23.54	x	-
314	BELGRANO II	ATA	-77.87	-34.63	x	-
315	EUREKA	CAN	80.00	-86.18	-	x
316	DEBILT	NLD	52.10	5.18	x	x
317	LAGOS	NGA	6.60	3.33	x	-
318	VALENTIA OBSERVATORY	IRL	51.93	-10.25	x	x
319	MONTREAL (DORVAL)	CAN	45.48	-73.75	x	-
320	WINNIPEG	CAN	49.90	-97.24	x	-
321	HALIFAX (BEDFORD)	CAN	44.70	-63.61	x	-

Fortsetzung nächste Seite

WMO	Name	Land	Latitude	Longit.	O ₃ -Säule	O ₃ -Profil
322	PETALING JAYA	MYS	3.10	101.65	x	-
323	NEUMAYER	ATA	-70.65	-8.25	-	x
325	LINAN	CHN	30.30	119.73	x	-
326	LONGFENGSHAN	CHN	44.75	127.60	x	-
327	ANGRA DO HEROISMO	PRT	38.66	-27.22	x	-
328	ASCENSION ISLAND	COG	-7.98	-14.42	-	x
330	HANOI	VNM	21.03	105.85	x	-
331	POPRAD-GANOVCE	SVK	49.03	20.32	x	-
332	POHANG	KOR	36.03	129.38	x	-
336	ISFAHAN	IRN	32.47	51.12	x	-
338	BRATTS LAKE (REGINA)	CAN	50.21	-104.71	x	-
339	USHUAIA	ARG	-54.85	-68.31	x	-
340	SPRINGBOK	ZAF	-29.67	17.90	x	-
341	HANFORD	USA	36.32	-119.63	x	-
342	COMODORO RIVADAVIA	ARG	-45.78	-67.50	x	-
343	SALTO	URY	-31.38	-57.97	x	-
345	SONGKHLA	THA	7.20	100.60	x	-
346	MURCIA	ESP	38.00	-1.17	x	-
347	ISSYK-KUL	KGZ	42.62	76.98	x	-
348	ANKARA	TUR	39.95	32.88	-	x
349	LASHA	CHN	29.40	91.03	x	-
350	TAHKUSE	EST	58.52	24.94	x	-
351	KING GEORGE ISLAND	URY	-62.18	-58.90	x	-
354	MINSK	BLR	53.83	27.47	x	-
376	MRSA MTROUH	EGY	31.33	27.22	x	-
394	BROADMEADOWS	AUS	-37.69	144.95	-	x
398	HONG KONG UNIVERSITY	HKG	22.28	114.20	x	-
399	USHUAIA II	ARG	-54.82	-68.32	x	-
400	MAITRI	ATA	-70.46	11.45	-	x
401	SANTA CRUZ	ESP	28.46	-16.26	-	x
404	JOKIOINEN	FIN	60.81	23.50	-	x
409	HURGHADA	EGY	27.28	33.75	x	-
410	AMBERD	ARM	40.38	44.25	x	-
412	DIEKIRCH	LUX	49.87	6.17	x	-
418	HUNSTVILLE	USA	35.28	-86.58	-	x
419	BORDEAUX	FRA	44.84	-0.53	x	-
426	SAN JULIAN	ARG	-49.30	-67.76	x	-
429	MARCAPOMACOCHA	PER	-11.40	-76.32	x	-
432	PAPEETE (TAHITI)	PYF	-18.00	-149.00	-	x
434	SAN CHRISTOBAL	ECU	-0.92	-89.60	-	x
436	LA REUNION ISLAND	REU	-21.08	55.48	-	x
437	WATUKOSEK (JAVA)	IDN	-7.57	112.65	-	x
438	SUVA (FIJI)	FJI	-18.13	178.32	-	x
441	EASTER ISLAND	CHL	-27.17	-109.42	-	x
442	PILAR	ARG	-31.66	-63.88	x	-
-	VIRTUAL STN 1	ANT	-90.00	0.00	x	-
-	VIRTUAL STN 2	ANT	-80.00	135.00	Test	-
-	VIRTUAL STN 3	ANT	-80.00	-45.00	x	-
-	VIRTUAL STN 4	ANT	-80.00	45.00	x	-
-	VIRTUAL STN 5	ANT	-80.00	135.00	x	-
-	VIRTUAL STN 6	VAR	-70.00	-180.00	x	-
-	VIRTUAL STN 7	VAR	-70.00	-90.00	x	-
-	VIRTUAL STN 8	VAR	-70.00	0.00	x	-
-	VIRTUAL STN 9	VAR	-70.00	90.00	x	-
-	VIRTUAL STN 10	VAR	-60.00	-180.00	x	-
-	VIRTUAL STN 11	VAR	-60.00	-108.00	x	-

Fortsetzung nächste Seite

WMO	Name	Land	Latitude	Longit.	O ₃ -Säule	O ₃ -Profil
-	VIRTUAL STN 12	VAR	-60.00	-36.00	x	-
-	VIRTUAL STN 13	VAR	-60.00	36.00	x	-
-	VIRTUAL STN 14	VAR	-60.00	108.00	x	-
-	VIRTUAL STN 15	ARC	90.00	0.00	x	-
-	VIRTUAL STN 16	ARC	80.00	-135.00	x	-
-	VIRTUAL STN 17	ARC	80.00	-45.00	x	-
-	VIRTUAL STN 18	ARC	80.00	45.00	Test	-
-	VIRTUAL STN 19	ARC	80.00	135.00	x	-
-	VIRTUAL STN 20	VAR	70.00	-180.00	x	-
-	VIRTUAL STN 21	VAR	70.00	-90.00	x	-
-	VIRTUAL STN 22	VAR	70.00	0.00	x	-
-	VIRTUAL STN 23	VAR	70.00	90.00	x	-
-	VIRTUAL STN 24	VAR	60.00	-180.00	x	-
-	VIRTUAL STN 25	VAR	60.00	-108.00	x	-
-	VIRTUAL STN 26	VAR	60.00	-36.00	x	-
-	VIRTUAL STN 27	VAR	75.00	-45.00	x	-
-	VIRTUAL STN 28	VAR	23.00	0.00	x	-
-	VIRTUAL STN 29	VAR	-21.00	135.00	x	-
-	VIRTUAL STN 30	VAR	-45.00	90.00	x	-
-	VIRTUAL STN 31	VAR	28.00	90.00	x	-
-	VIRTUAL STN 32	VAR	15.00	180.00	x	-
-	VIRTUAL STN 33	VAR	-45.00	-135.00	x	-
-	VIRTUAL STN 34	VAR	-40.00	-15.00	x	-
-	VIRTUAL STN 35	VAR	-8.00	-60.00	x	-
-	VIRTUAL STN 36	VAR	10.00	135.00	x	-

Anhang B

TOMS-Korrekturfunktion

Die in [Bodeker *et al.*, 2001] beschriebene Korrekturfunktion wurde entwickelt, um Ozontrendanalysen unter Verwendung verschiedener Satelliteninstrumente zu ermöglichen, indem deren Messungen möglichst gut an das Ozon-Bodenmessnetz angepasst werden. Sie hat die analytische Form

$$\Delta(\theta, t) = A + Bt + C \sin(2\pi t) + D \cos(2\pi t) + E \sin(4\pi t) + F \cos(4\pi t). \quad (\text{B.1})$$

Die Variable t gibt hierbei die Zeit seit 1. 1. 1998 in Jahren an, θ ist die geografische Breite in Radians. Die Koeffizienten A werden als

$$A = \sum_{i=0}^6 A_i \theta^i \quad (\text{B.2})$$

und die Koeffizienten B bis F als

$$B = B_0 + B_1 \sin \theta + B_2 \cos \theta \quad (\text{B.3})$$

expandiert. Dies ergibt insgesamt 22 freie Parameter, deren Bestimmung mit Hilfe einer *least-squares* Regression auf den Differenzen zwischen Satelliten- und Bodenmessungen erfolgte. Im Rahmen dieser Arbeit wurden nur die Koeffizienten für Earth Probe TOMS V7 Daten benötigt. Diese lauten [G. Bodeker, *persönliche Mitteilung*, 2000]:

$A_0 =$	- 6.719668	$B_0 =$	0.698368	$E_0 =$	0.957396
$A_1 =$	3.805661	$B_1 =$	- 0.003095	$E_1 =$	- 1.079085
$A_2 =$	19.768902	$B_2 =$	- 1.386023	$E_2 =$	- 1.600956
$A_3 =$	1.850527	$C_0 =$	- 2.810475	$F_0 =$	- 3.476455
$A_4 =$	-40.115544	$C_1 =$	1.160466	$F_1 =$	0.144585
$A_5 =$	- 1.289032	$C_2 =$	3.692910	$F_2 =$	4.610161
$A_6 =$	15.738706	$D_0 =$	- 1.487623		
		$D_1 =$	0.405276		
		$D_2 =$	0.732498		

In Abbildung B.1 ist die Korrekturfunktion grafisch dargestellt. Dabei ist zu beachten, dass die die Latitude θ polwärts von 70°S und 65°N festgehalten wird, um unrealistisch hohe Werte durch die Potenzierung zu vermeiden.

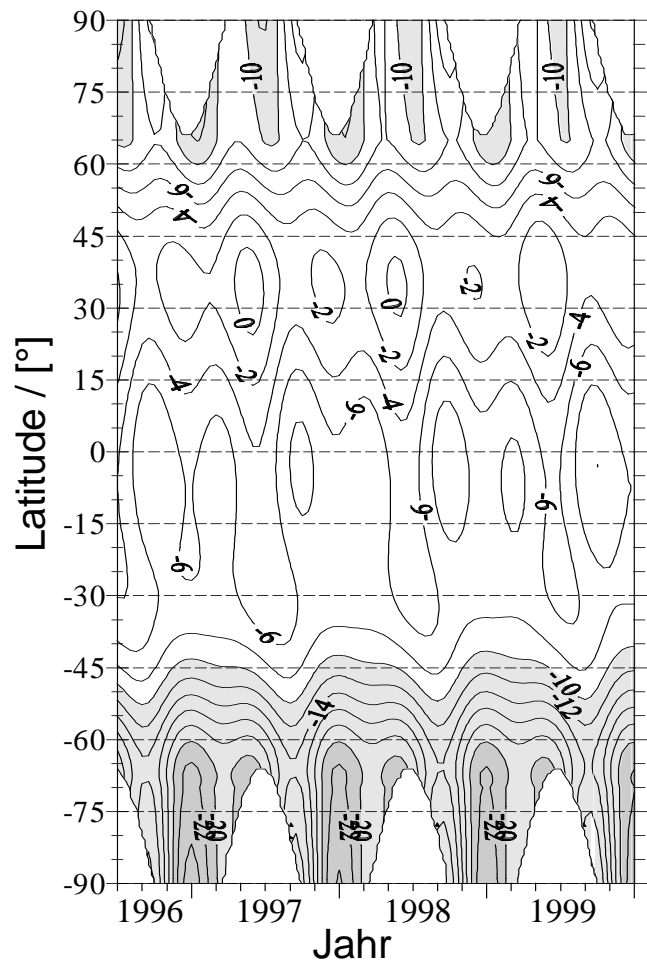


Abbildung B.1: Korrekturfunktion für EP-TOMS-Daten in [DU], aus [Bodeker et al., 2001].

Anhang C

Behandlung der Profilkollokationen

Die in Kapitel 7 beschriebene, statistische Optimierung und Homogenisierung des Ozonprofil-Kollokationsdatensatzes wird in diesem Abschnitt näher konkretisiert. Basis ist der komplette Datensatz, der den Zeitraum Jan. 1996 bis einschließlich Juli 2001 umfasst und Kollokationsdistanzen von maximal 250 km aufweist. Die nachfolgende Beschreibung bezieht sich dabei auf die Stationsnummern gemäß WMO (Anhang A):

- Die Daten der Stationen 187 und 394 wurden wegen zu häufiger Datenfehler [Logan, 1999] komplett, die der Stationen 99 und 101 zur späteren Validation aussortiert.
- Der Kollokationsradius für die Stationen 18, 21, 76, 77, 156, 174, 221, 242, 315, 316 und 323 (mittlere und hohe Breiten) wurde auf 150 km reduziert. Für die Stationen in den nördlichen mittleren Breiten geschah dies aufgrund der hohen Stationsdichte. Bei den polnäheren Stationen ergeben sich durch die Überlappung der GOME-Orbits extrem viele Kollokationen.
- Aus demselben Grund wurde der Kollokationsradius für POAM III global auf 50 km beschränkt. NH-Kollokationen südlich von 65°N wurden entfernt.

Einige der obigen Maßnahmen – insbesondere die Reduzierung der POAM-Kollokationen – verringerten bereits die zeitliche Ungleichverteilung. Zusätzlich erwies sich folgendes Vorgehen als geeignet:

- Der Kollokationsradius zwischen 45°N–55°N wurde von 1996 bis einschließlich April 2000 auf 120 km herabgesetzt.
- Der Kollokationsradius zwischen 25°S–5°N wurde von 1996 bis einschließlich Juli 2000 auf 170 km herabgesetzt.
- Der Kollokationsradius zwischen 35°N–75°N wurde von Januar bis einschließlich April 1997 auf 90 km herabgesetzt.

Im Zuge der Fehlerreduktion des Datensatzes wurden über das in Abschn. 7.1 beschriebene, grundlegende Screening die folgenden Anpassungen vorgenommen:

- Alle Ozonsondenprofile wurden zwischen 26 km und 30 km Höhe abgeschnitten, da deren Zuverlässigkeit i. A. oberhalb dieses Höhenbereichs stark abnimmt [SPARC, 1998]. Die Obergrenze wurde hier für jedes Profil zufällig festgelegt, da das Abschneiden aller Profile auf gleicher Höhe in Vorversuchen zu Artefakten geführt hat.

- Bei ersten globalen Profilbestimmungen mit NNORSY hatte sich gezeigt, dass über der sogenannten *South Atlantic Anomaly* (SAA) starkes Rauschen der Ozonwerte in der Troposphäre und über 35 km auftritt. Bei der SAA handelt es sich um eine Störung des durch das Erdmagnetfeld erzeugten Schutzschirms gegen hochenergetische, geladene Teilchen.¹ Der Einschlag dieser Partikel erzeugt im GOME-Sensor punktuell Rauschen (*shot noise*) einzelner Photodioden. Da es hierfür im GDP noch keine entsprechende Korrektur gibt [D. Loyola, *persönliche Mitteilung*, 2001], und derartige, hoch variable statistische Fehler vom Belernungsalgorithmus schlecht berücksichtigt werden können, wurden für den aktuellen Datensatz alle Kollokationen des entsprechenden Gebietes ($10^{\circ}\text{S}-36^{\circ}\text{S} \cap 0^{\circ}-70^{\circ}\text{W}$) von der Belernung ausgeschlossen.
- Die Okkultationsmessungen werden wie bereits erwähnt durch Aerosole, PSCs und Wolken gestört, außerdem nehmen unterhalb von 15–20 km die Fehler bei der Höhenregistrierung und die horizontale Ausdehnung der Pixel durch den langen Lichtweg stark zu. Da in den entsprechenden, niedrigen Höhen i. A. ausreichend Ozonsondendaten vorhanden sind, wurden die Okkultationsprofile bei einer Höhe abgeschnitten, die in etwa der Tropopausehöhe ± 5 km entspricht, und demnach mit der geografischen Breite variiert. Gemäß Literaturangaben über die Zuverlässigkeit der einzelnen Sensoren und eigenen Untersuchungen variierte die Referenzhöhe zusätzlich je nach Sensor um einige Kilometer.
- Die Anwendung subtilerer Techniken wie etwa der von *Bhatt et al.* [1999] vorgeschlagenen Screeningprozedur für Aerosolkontamination bei HALOE (vgl. 2.2.3), oder einer vom Autor entwickelten, auf der troposphärischen SAGE-Profilform beruhende Fehlerdetektion führte leider bislang nicht zum gewünschten Erfolg und wurde deshalb verworfen.

¹Die deutsche ROSAT-Satellitenmission enthielt einen eigens zur Untersuchung der SAA konzipierten Sensor, den SAA Detector (SAAD). Nähere Informationen hierzu finden sich z.B. unter <http://www.ll.mit.edu/ST/sbv/saa.html>.

Anhang D

Lebenslauf und Publikationen

Lebenslauf des Autors

Name	<i>Martin</i> Dietmar Müller	
Geburtsdatum	11. April 1972	
Geburtsort	Bamberg	
Familienstand	ledig	
Anschrift	Riemenschneiderweg 7a, D-96450 Coburg	
<hr/>		
Schulbildung	Sept. 1982 – Juli 1991	Gymnasium Ernestinum Coburg
<hr/>		
Weiterführende Ausbildung	Nov. 1991 – Feb. 1997	Studium der Physik an der Universität Bayreuth
	Okt. 1993 – März 1994	Auslandssemester an der University of St. Andrews, Schottland
Praktika	Aug. 1994 – Sept. 1994	Robert Bosch GmbH, Stuttgart Abteilung <i>Zentrale Analytik</i>
	März 1995	DLR, Göttingen Abteilung <i>Aerothermodynamik</i>
<hr/>		
Wehrdienst	März 1997 – Dez. 1997	Fernmeldeaufklärungsregiment 220, Donauwörth Programmiertätigkeit in der Gruppe <i>Technische Erfassung</i>
<hr/>		
Berufliche Tätigkeit	seit Jan. 1998	Wissenschaftlicher Mitarbeiter und Netzwerkadministrator am Zentrum für Sonnenenergie- und Wasserstoff-Forschung (ZSW), Stuttgart

Publikationsliste

Fachzeitschriftenartikel

- Müller, M.D., A.K. Kaifel, M. Weber, J.P. Burrows: Neural Network Scheme for the Retrieval of Total Ozone from Global Ozone Monitoring Experiment Data, *Applied Optics*, 41(24), 5051–5058, 2002.
- Müller, M.D., A.K. Kaifel, M. Weber, J.P. Burrows: Partial training of neural networks with incomplete target data, applied to atmospheric science, eingereicht bei *Neural Networks*
- Müller, M.D., A.K. Kaifel, M. Weber, S. Tellmann, J.P. Burrows, D. Loyola: Ozone Profile Retrieval from GOME Data using a Neural Network Approach (NNORSY), eingereicht bei *Journal of Geophysical Research*

Publizierte Konferenzbeiträge

- Kaifel, A.K. und M.D. Müller: *Results of TOVS ozone retrieval with neural networks*. In: Techn. Proc. 11th Int. ATOVS Study Conference, Budapest, September 2000, pp. 153–166, Bureau of Met. Res. Center, Melbourne, Australia, 2001.
- Müller, M.D. und A.K. Kaifel: *Efficient processing of multi-year global TOVS data using ITPP, 3I and neural networks*. In: Techn. Proc. 10th Int. ATOVS Study Conference, pp. 397–407, Boulder, CO, USA, 1999.
- Müller, M.D., A.K. Kaifel, M. Weber: *Ozone from GOME data using neural network technique*. In: Sawaya-Lacoste, H. (Ed.), Proc. Of ERS-ENVISAT Symposium, Gothenburg, Sweden, 16–20 October 2000. ESA-ESTEC, ESA Public. Div. (SP-461), Noordwijk, The Netherlands, 2001.
- Müller, M.D., A.K. Kaifel, M. Weber, S. Tellmann: *Real-time total ozone and ozone profiles retrieved from GOME data using neural networks*. In: Proc. 2001 EUMETSAT Meteorological Satellite Data User's Conference, Antalya, 1-5 October 2001. EUMETSAT, Darmstadt, Germany, 2001.
- Kaifel, A.K., M.D. Müller, M. Weber, S. Tellmann: *Neural Network Ozone Retrieval System for GOME spectra (NNORSY-GOME)*. In: Proc. EUROTRAC-2 Symposium 2002, P.M. Midgley, M. Reuther (Eds.), Margraf Verlag, Weikersheim, Germany, in Druck.

Vorträge und Präsentationen

- Seminarvortrag “Möglichkeiten der Verwendung von neuronalen Netzen für die Ozonbestimmung aus GOME-Daten” an der Universität Bremen, IUP Seminar, November 1999.
- Seminarvortrag “Real-Time Total Ozone and Ozone Profiles from TOVS and GOME Data Using Neural Networks” am Data Assimilation Office (DAO), GSFC, Greenbelt, USA, September 2001.
- Präsentation “Ozone Profiles from GOME Data Using Neural Network Technique” im Rahmen des ESA/ESRIN GOME Data Exploitation Meeting, Frascati, Italien, April 2001.

- Vortrag “Real-Time Total Ozone and Ozone Profiles from TOVS and GOME Data Using Neural Networks” am Cooperative Institute for Meteorological Satellite Studies (CIMSS), Madison, WI, USA, September 2001.
- Präsentation “Recent developments of the ZSW GOME ozone retrieval scheme” im Rahmen des GOME Ozone Profile Validation Group meeting, Univ. of Leicester, UK, November 2001.
- Seminarvortrag “The Neural Network Ozone Retrieval System (NNORSY) for GOME ozone profiles” an der Universität Bremen, IUP Seminar, Mai 2002.

Anhang E

Abkürzungsverzeichnis

Mathematische Symbole

Δ	Gewichtsschrittweite bei RPROP
$\Delta(\cdot)$	TOMS-Korrekturfunktion [DU]
δ	Halbwertsbreite der Glättungsfunktionen
δ	Standardabweichung relativ zum Mittelwert
$\delta\lambda$	spektrale Auflösung [m]
ϵ	statistischer Fehlervektor
θ	geografische Breite [$^{\circ}$ N]
ϑ	Sonnenzenitwinkel
λ	Eigenwert
λ	Lernrate des Backpropagation-Verfahrens
λ	Wellenlänge [m]
μ	Sinus von ϑ
μ	mittlere Abweichung, Bias
ν	Frequenz [Hz]
$\nu_{\text{in}}(\cdot)$	Eingangsnormierung des MLP
$\nu_{\text{out}}(\cdot)$	Ausgangsnormierung des MLP
σ	Standardabweichung
σ_a	Volumen-Absorptionsquerschnitt [m^{-1}]
σ_e	Volumen-Extinktionsquerschnitt [m^{-1}]
σ_s	Volumen-Streuquerschnitt [m^{-1}]
τ	Belernungsepochennummer [-]
τ	optische Dichte [-]
ϕ	Azimutwinkel
\mathcal{A}	<i>airmass</i> -Faktor [-]
\mathbf{A}	Auflösungsmatrix (<i>averaging kernel matrix</i>)
\mathcal{D}	Mahalanobis-Distanz
\mathbf{D}_y	Beitragsfunktionsmatrix (<i>contribution function matrix</i>)
E	Fehlerfunktion
$\mathbf{F}(\cdot)$	diskretes Strahlungstransport-Vorwärtsmodell
\mathbf{G}	neuronale Jacobi-Matrix
\mathbf{I}	Einheitsmatrix
I	Strahlungsflussdichte [Photonen/(nm cm ² s)]
\mathbf{K}	Gewichtsfunktionsmatrix (<i>kernel matrix</i>)
\mathcal{L}	Likelihood-Funktion
MAE	mittlerer absoluter Fehler (<i>mean absolute error</i>)
N	Neuronenanzahl einer MLP-Schicht

P	Anzahl der (Trainings-)Muster
$\mathbf{R}(\cdot)$	inverses Strahlungstransportmodell
$\mathbf{R}(\cdot)_{\text{MLP}}$	inverses, neuronales Strahlungstransportmodell
\mathbf{S}	Kovarianzmatrix
S	Ozonsäule [DU]
$S(\cdot)$	Streufunktion
\mathcal{T}	Kollokations-Datensatz
\mathbf{V}	Matrix der Verbindungsgewichte von MLP-Schicht 2 nach 3
\mathbf{W}	Matrix der Verbindungsgewichte von MLP-Schicht 1 nach 2
$\tilde{\mathbf{y}}$	unnormierter Messvektor
\mathbf{b}	Vorwärtsmodell-Parameter
c	Lichtgeschwindigkeit [$2.998 \cdot 10^8 \text{ m s}^{-1}$]
$f(\cdot)$	neuronale Transferfunktion (hier: tanh)
h	Plancksches Wirkungsquantum [$6.626 \cdot 10^{-34} \text{ J s}$]
m	Dimension des Messraumes
n	Dimension des Zustandsraumes
n	Netzeingabe, neuronale Aktivierung
\mathbf{o}	Netzwerkausgabe
$p(\cdot)$	Wahrscheinlichkeitsdichte
\mathbf{t}	Zielwertvektor für Netzausgabe (<i>target vector</i>)
t	Zeit [s]
v	Geschwindigkeit [m s^{-1}]
$\hat{\mathbf{x}}$	beste Schätzung für Ozonprofil
\mathbf{w}	Gewichtsvektor des neuronalen Netzes
\mathbf{x}	Zustandsvektor der Atmosphäre (meist Ozonprofil)
\mathbf{y}	normierter Messvektor (Netzeingabe)
z	geometrische Höhe über Erdboden [m]

Akronyme

ADEOS	ADvanced Earth Observing Satellite
AIRS	Atmospheric InfraRed Sounder
AK	Averaging Kernel
AMF	AirMass Factor
ASCII	American Standard Code for Information Interchange
AVHRR	Advanced Very High Resolution Radiometer
BUV	Backscattered Ultra-Violet
CLW	Cloud Liquid Water
DFD	Deutsches Fernerkundungs-Datenzentrum
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DOAS	Differential Optical Absorption Spectroscopy
DU	Dobson Unit
DWD	Deutscher WetterDienst
ECC	Elektrochemical Concentration Cell
ECMWF	European Centre for Medium-Range Weather Forecasts
ENVISAT	ENVIronmental SATellite
EPT	Earth Probe TOMS
ERS	European Remote-sensing Satellite
ESA	European Space Agency
ESRIN	European Space Research Institute
F&K	Fortuin & Kelder (Klimatologie)

FCKW	Fluor-Chlor-KohlenWasserstoff
FURM	FULL Retrieval Method
FWHM	Full Width Half Maximum
GAW	Global Atmosphere Watch
GDP	Gome Data Processor
GOME	Global Ozone Monitoring Experiment
GOMETRAN	GOME-StrahlungsTRANsportmodell
GOMOS	Global Ozone Monitoring by Occultation of Stars
GOPWG	GOME-1 Ozone Profile Working Group
GPH, gph	Geopotenzielle Höhe
GSFC	Goddard Space Flight Center
HALOE	HALogen Occultation Experiment
IMG	Interferometric Monitor for Greenhouse gases
IR	InfraRot
ISCCP	International Satellite Cloud Climatology Project
IUP	Institut für UmweltPhysik (Universität Bremen)
KNMI	Koninklijk Nederlands Meteorologisch Instituut
Lidar	LIght Detection And Ranging
LOS	Line-Of-Sight
MAE	Mean Absolute Error
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLP	Multi-Layer Perceptron
MLS	Microwave Limb Sounder
MPI	Max-Planck-Institut
NASA	National Aeronautics and Space Administration
NCEP	National Center for Environmental Protection
NH	NordHemisphäre
NN	Neuronales Netz
NNORSY	Neural Network Ozone Retrieval System
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real-Time
OE	Optimal Estimation
PCA	Principal Component Analysis
PMD	Polarisation Monitoring Device
POAM	Polar Ozone and Aerosol Measurement instrument
PSC	Polar Stratospheric Clouds
PV	Potential Vorticity
RPROP	Resilient PROPagation
RMSE	Root Mean Square Error
RTM	Radiative Transfer Model
S/N, SNR	Signal zu Rausch Verhältnis
SAA	Solar Azimuth Angle
SAA	South Atlantic Anomaly
SAGE	Stratospheric Aerosol and Gas Experiment
SBUV	Solar Backscattered Ultra-Violet instrument
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SH	SüdHemisphäre
SHADOZ	Southern Hemisphere Additional OZonesondes
SMR	Sub-Millimetre Receiver
SPARC	Stratospheric Processes And their Role in Climate

SPOT	Satellite Pour l'Observation de la Terre
SPRAFIMO	SPektral-RAdiometer auf FIIter-MOdelL Basis
SSM/I	Special Sensor Microwave/Imager
SSM/T	Special Sensor Microwave/Temperature
SST	Sea Surface Temperature
SZA	Solar Zenith Angle
TIROS	Television InfraRed Observation Satellite
TOA	Top Of Atmosphere
TOMS	Total Ozone Mapping Spektrometer
TOVS	TIROS Operational Vertical Sounder
UARS	Upper Atmosphere Research Satellite
UKMO	United Kingdom Meteorological Office
UV	UltraViolett
VOC	Volatile Organic Compounds
WMO	World Meteorological Organization
WOUDC	World Ozone and Ultraviolet radiation Data Center
WV	Water Vapor
ZSW	Zentrum für Sonnenenergie- und Wasserstoff-Forschung

Literaturverzeichnis

- Aben, I., M. Eisinger, E. Hegels, R. Snel & C. Tanzi (2000): *GOME Data Quality Improvement (GDAQI), Final Report*, Techn. Ber. TN-GDAQI-003SR/2000, Space Research Organization Netherlands (SRON), Utrecht, Netherlands.
- Aires, F., C. Prigent, W. B. Rossow & M. Rothstein (2001): “A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations”, *J. Geophys. Res.* **106**(D14), 14887–14908.
- Aires, F., M. Schmitt, A. Chedin & N. Scott (1999): “The ”weight smoothing” regularization of MLP for Jacobian stabilization”, *IEEE-NN* **10**(6), 1502.
- Akima, H. (1970): “A new method of interpolation and smooth curve fitting based on local procedures”, *J. of the ACM* **17**(4), 589–602.
- Appenzeller, C., J. R. Holton & K. H. Rosenlof (1996): “Seasonal variation of mass transport across the tropopause”, *J. Geophys. Res.* **101**, 15071–15078.
- Attmannspacher, W. & H. U. Dütsch (1970): *International Ozone Sonde Intercomparison at the Observatory Hohenpeissenberg*, Berichte des Deutschen Wetterdienstes 120, DWD.
- Backus, G. E. & J. F. Gilbert (1970): “Uniqueness in the inversion of inaccurate gross earth data”, *Phil. Trans. R. Soc. Lond.* **266**, 123–140.
- Balzer, W. & D. Loyola (1996): *Product Specification Document of the GOME Data Processor*, Techn. Ber. ER-PS-DLR-GO-0016, Deutsches Zentrum für Luft- und Raumfahrt, Deutsches Fernerkundungs-Datenzentrum, Oberpfaffenhofen, Germany.
- Barath, F., M. C. Chavez, R. E. Cofield, D. A. Flower, M. A. Frerking, M. B. Gram, W. M. Harris, J. R. Holden, R. F. Jarnot, W. G. Kloezeman, G. J. Klose, G. K. Lau, M. S. Loo, B. J. Maddison, R. J. Mattauach, R. P. McKinney, G. E. Peckham, H. M. Pickett, G. Siebes, F. S. Soltis, R. A. Suttie, J. A. Tarsala, J. W. Waters & W. J. Wilson (1993): “The Upper Atmosphere Research Satellite Microwave Limb Sounder instrument”, *J. Geophys. Res.* **98**, 10751–10762.
- Bargen, A. V., W. Thomas & D. Loyola (1999): *GOME Data Processor Update Report for GDP 0-to-1 Version 2.0 and GDP 1-to-2 Version 2.7*, Techn. Ber. ER-TN-DLR-GO-0043, Deutsches Zentrum für Luft- und Raumfahrt, Deutsches Fernerkundungs-Datenzentrum, Oberpfaffenhofen, Germany.
- Baron, A. (1993): “Universal approximation bounds for superposition of sigmoidal functions”, *IEEE Trans. Information Theory* **39**, 930–945.
- Baron, A. (1994): “Approximation and estimation bound for artificial neural networks”, *Machine Learning* **14**, 115–133.

- Basher, R. E. (1982): *Review of the Dobson Spectrophotometer and its Accuracy*, WMO Ozone Rep. 13, World Meteorol. Organ., Geneva.
- Basher, R. E. (1995): *Survey of WMO-Sponsored Dobson Spectrophotometer Intercomparisons*, Techn. Ber., World Meteorological Organization (WMO), Wellington, New Zealand.
- Beekmann, M., G. Ancellet, G. Megie, H. G. J. Smit, & D. Kley (1994): "Intercomparison campaign for vertical ozone profiles including electrochemical sondes of ECC and Brewer-Mast type and a ground based UV-differential absorption lidar", *J. Atmos. Chem.* **19**, 259–288.
- Berendes, T. A., K. S. Kuo, A. M. Logar, E. M. Corwin, R. M. Welch, B. A. Baum, A. Pretre & R. C. Weger (1999): "A comparison of paired histogram, maximum likelihood, class elimination, and neural network approaches for daylight global cloud classification using AVHRR imagery", *J. Geophys. Res.* **104**(D6), 6199–6213.
- Bhatt, P. P., E. E. Remsberg, L. L. Gordley, J. M. McInerney, V. G. Brackett & J. M. Russell III (1999): "An evaluation of the quality of Halogen Occultation Experiment ozone profiles in the lower stratosphere", *J. Geophys. Res.* **104**(D8), 9261–9275.
- Bishop, C. M. (1994): "Novelty detection and neural network validation", *IEEE Proceedings: Vision, Image and Signal Processing* **141**(4), 217–222.
- Bishop, C. M. (1995a): *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.
- Bishop, C. M. (1995b): "Training with noise is equivalent to Tikhonov regularization", *Neural Computation* **7**(1), 108–116.
- Blum, A. L. & R. L. Rivest (1992): "Training a three-node neural network is NP-complete", *Neural Networks* **5**(1), 117–127.
- Bodeker, G. E., J. C. Scott, K. Kreher & R. L. McKenzie (2001): "Global ozone trends in potential vorticity coordinates using TOMS and GOME intercompared against the dobson network: 1978-1998", *J. Geophys. Res.* **106**(D19), 23029–23042.
- Bojkov, R. D., L. Bishop & V. E. Fioletov (1995): "Total ozone trends from quality controlled ground-based data (1964–1994)", *J. Geophys. Res.* **100**, 25867–25876.
- Bojkov, R. D. & V. E. Fioletov (1997): "Changes of the lower stratospheric ozone over Europe and Canada", *Geophys. Res. Lett.* **102**, 1337–1347.
- Bovesmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noel, V. Rozanov, K. V. Chance & A. H. P. Goede (1999): "SCIAMACHY - Mission objectives and measurement modes", *J. Atmos. Sciences* **56**, 125–150.
- Bramstedt, K. (2001): *GOME Ozonprofile: Weiterentwicklung und Validierung*, Dissertation, University of Bremen.
- Branke, J. (1995): "Evolutionary algorithms for neural network design and training", in: *1st Nordic Workshop on Genetic Algorithms and its Applications, Vaasa, Finland, January 1995*.
- Brewer, A. W. (1949): "Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere", *Quart. J. Roy. Soc.* **75**, 351–363.
- Brewer, A. W. & J. B. Kerr (1973): "Total ozone measurements in cloudy weather", *Pure Appl. Geophys.* **106–108**, 928–937.

- Brinksma, E. J., J. B. Bergwerff, G. E. Bodeker, K. F. Boersma, I. S. Boyd, B. J. Connor, J. F. de Haan, W. Hogervorst, J. W. Hovenier, A. Parrish, J. J. Tsou, J. M. Zawodny & D. P. J. Swart (2000): "Validation of 3 years of ozone measurements over Network for the Detection of Stratospheric Change station Lauder, New Zealand", *J. Geophys. Res.* **105**(D13), 17291–17306.
- Brinksma, E. J., Y. J. Meijer, B. J. Connor, G. L. Manney, J. B. Bergwerff, G. E. Bodeker, I. S. Boyd, J. B. Liley, W. Hogervorst, J. W. Hovenier, N. J. Livesey & D. P. J. Swart (1998): "Analysis of record-low ozone values during the 1997 winter over Lauder, New Zealand", *Geophys. Res. Lett.* **25**(15), 2785–2788.
- Bruehl, C., S. R. Drayson, I. Russel, J. M., P. J. Crutzen, J. McInerney, P. N. Purcell, H. Claude, H. Gernand, T. McGee, I. McDermid & M. R. Gunson (1996): "HALOE ozone channel validation", *J. Geophys. Res.* **101**, 10.217–10.240.
- Buchwitz, M. (2000): *Strahlungstransport- und Inversions-Algorithmen Zur Ableitung Atmosphärischer Spurengasinformationen Aus Erdfernerkundungsmessungen in Nadirgeometrie im Ultravioletten Bis Nahinfraroten Spektralbereich Am Beispiel SCIAMACHY*, Dissertation, University of Bremen.
- Burrows, J., D. Diebel, B. Kerridge, R. Munro, D. Pemberton, U. Platt & H. Frank (1994): *A Study of Methods for Retrieval of Atmospheric Constituents*, Final Report ESA Contract 9687/91/NL/BI, European Space Agency, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands.
- Burrows, J., E. Hölzle, A. Goede, H. Visser & W. Fricke (1995): "SCIAMACHY - Scanning Imaging Absorption Spectrometer for Atmospheric Chartography", *Acta Astronautica* **35**, 445–451.
- Burrows, J., M. Weber, M. Buchwitz, V. Rozanov, A. Ladstädter-Weissenmayer, A. Richter, R. de Beek, R. Hoogen, K. Bramstedt, K.-U. Eichmann, M. Eisinger & D. Perner (1999): "The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results", *J. Atmos. Sciences* **56**, 151–175.
- Butler, C. T., R. Z. Meredith & A. P. Stogryn (1996): "Retrieving atmospheric temperature parameters from DMSP SSM/T-1 data with a neural network", *J. Geophys. Res.* **101**, 7075–7083.
- Callies, J., C. Caspar, A. Lefebvre & A. Hahne (1996): "GOME operations during the validation campaign", in: *GOME Geophysical Validation Campaign Proceedings, Final Results, ESA-WPP 108*, S. 5–19, European Space Agency, Frascati.
- Castro, J. L., C. J. Mantas & J. M. Benítez (2000): "Neural networks with a continuous squashing function in the output are universal approximators", *Neural Networks* **13**, 561–563.
- Chance, K. V., J. P. Burrows, D. Perner & W. Schneider (1997): "Satellite measurements of atmospheric ozone profiles, including tropospheric ozone, from UV/visible measurements in the nadir geometry: A potential method to retrieve tropospheric ozone", *J. Quant. Spectrosc. Rad. Transfer* **57**(4), 467–476.
- Chapman, S. (1930): "A theory of upper atmospheric ozone", *Quart. J. Roy. Soc.* **3**, 103–125.
- Chevallier, F., F. Chérury, R. Armante, C. J. Stubenrauch & N. A. Scott (2000): "Retrieving the clear sky vertical longwave radiative budget from TOVS: Comparison of a neural network-based retrieval and a method using geophysical parameters", *J. Appl. Meteor.* Accepted 1999.

- Chevallier, F., F. Chérury, N. A. Scott & A. Chédin (1998): “A neural network approach for a fast and accurate computation of a longwave radiative budget”, *J. Appl. Meteor.* **37**, 1385–1397.
- Chevallier, F. & J.-F. Mahfouf (2000): “Evaluation of the Jacobians of infrared radiation models for variational data assimilation”, *J. Appl. Meteor.* Submitted May 2000.
- Chu, W. P., E. W. McCormick, J. Lenoble, C. Brogniez & P. Pruvost (1989): “SAGE II inversion algorithm”, *J. Geophys. Res.* **94**, 8339–8351.
- Chubachi, S. (1984): “A special ozone observation at Syowa station, Antarctica”, in: C. S. Zerefos & A. Ghazi (Hg.), *Proc. Ozone Symp. 1984*, S. 243–247, D. Reidel Publ. Co., New York.
- Churnside, J. H., T. A. Stermitz & J. A. Schroeder (1994): “Temperature profiling with neural network inversion of microwave radiometer data”, *J. Atmos. Ocean. Tech.* **11**, 105–109.
- Ciotti, P. B. A. P., G. d’Auria, F. Marzano, N. Pierdicca & S. Bonafoni (1997): “The role of a priori information in designing retrieval algorithms for microwave radiometric profiling of the atmosphere”, in: *Proc. Of IEEE 1997 International Geoscience and Remote Sensing Symposium (IGARSS 1997)*, S. 2100–2102, IEEE.
- Claude, H., R. Hartmansgruber & U. Köhler (1987): *Measurement of Atmospheric Ozone Profiles Using the Brewer/Mast Sonde: Preparation, Procedure, Evaluation*, WMO Global Ozone Research and Monitoring Project Report No. 17 WMO/TD 179, World Meteorological Organization, Geneva: AREP/ENV.
- Comrie, A. C. (1997): “Comparing neural networks and regression models for ozone forecasting”, *J. Air & Waste Manage. Assoc.* **47**, 653–663.
- Corlett, G. K. & P. S. Monks (2001): “A comparison of total column ozone values derived from the Global Ozone Monitoring Experiment (GOME), the Tiros Operational Vertical Sounder (TOVS), and the Total Ozone Mapping Spectrometer (TOMS)”, *J. Atmos. Sciences* **58**(9), 1103–1116.
- Crone, L. & D. Crosby (1995): “Statistical applications of a metric on subspaces to satellite meteorology”, *Technometrics* **37**, 324–328.
- Crutzen, P. J. & F. Arnold (1986): “Nitric acid cloud formation in the cold antarctic stratosphere: A major cause for the springtime „ozone hole“”, *Nature* **324**, 651–655.
- Crutzen, P. J. & L. T. Gidel (1983): “A two-dimensional photochemical model of the atmosphere. 2: The tropospheric budgets of anthropogenic chlorocarbons CO, CH₄, CH₃Cl and the effect of various NO_x sources on tropospheric ozone”, *J. Geophys. Res.* **88**, 6641–6661.
- Cunnold, D. M., W. P. Chu, R. A. Barnes, M. P. McCormick & R. E. Veiga (1989): “Validation of SAGE II ozone measurements”, *J. Geophys. Res.* **94**(D6), 8447–8460.
- Cunnold, D. M., M. J. Newchurch, L. E. Flynn, H. J. Wang, J. M. Russell, R. McPeters, J. M. Zawodny & L. Froidevaux (2000a): “Uncertainties of upper stratospheric ozone trends from 1979 to 1996”, *J. Geophys. Res.* **105**(D4), 4427–4444.
- Cunnold, D. M., H. J. Wang, L. W. Thomason, J. M. Zawodny, J. A. Logan & I. A. Megretskaia (2000b): “SAGE (version 5.96) ozone trends in the lower stratosphere”, *J. Geophys. Res.* **105**(D4), 4445–4457.

- De Backer, H., D. De Muer, E. Schoubs & M. Allaart (1998): "A new pump correction profile for Brewer-Mast ozonesondes", in: R. Bojkov & G. Visconti (Hg.), *Proceedings of the XVIII Quadrennial Ozone Symposium, 1996*, L'Aquila, Italy.
- de Beek, R. (1998): *Bestimmung Von Ozonvertikalprofilen Aus Messungen Des Satelliten-instrumentes GOME im Ultravioletten und Sichtbaren Spektralbereich*, Dissertation, University of Bremen, berichte aus der Umweltphysik D46.
- de Gruijl, F. R. & J. C. van der Leun (1993): "Influence of ozone depletion on the incidence of skin cancer", in: A. R. Y. et al. (Hg.), *Environmental UV Photobiology*, Plenum Press.
- Del Frate, F., S. Casadio & C. Zehner (2000): "Retrieval of ozone profiles by using GOME measurements and a neural network algorithm", in: *Proc. Of IEEE 2000 International Geoscience and Remote Sensing Symposium (IGARSS 2000), Honolulu, July 2000*, IEEE.
- Del Frate, F. & G. Schiavon (1998a): "A combined natural orthogonal functions / neural network technique for radiometric estimation of atmospheric profiles", *Radio Science* **33**, 405–410.
- Del Frate, F. & G. Schiavon (1998b): "Neural networks for the retrieval of water vapor and liquid water from radiometric data", *Radio Science* **33**, 1373–1386.
- Del Frate, F. & G. Schiavon (1999): "Nonlinear principal component analysis for the radiometric inversion of atmospheric profiles by using neural networks", *IEEE Trans. Geosci. Remote Sensing* **37**, 2335–2342.
- Deniel, C., F. Dalaudier, E. Chassefiere, R. M. Bevilacqua, E. P. Shettle, J. D. Lumpe, K. W. Hoppel, J. S. Hornstein, D. W. Rusch & C. E. Randall (1997): "A comparative study of POAM II and ECC sonde ozone measurements obtained over northern europe", *J. Geophys. Res.* **102**(D19), 23629–23642.
- Dobson, G. (1930): "Observations of the amount of ozone in the earth's atmosphere and its relation to other geophysical conditions", in: *Proc. Roy. Soc. London*, S. 411–433.
- Dobson, G. M. B. (1956): "Origin and distribution of polyatomic molecules in the atmosphere", *Proc. R. Soc. London, Ser. A* **236**, 187–193.
- Douglass, A. R., C. J. Weaver, R. B. Rood & L. Coy (1996): "A three-dimensional simulation of the ozone annual cycle using winds from a data assimilation system", *J. Geophys. Res.* **101**, 1463–1483.
- Dybowski, R. (1997): *Assigning Confidence Intervals to Neural Network Predictions*, Techn. Ber., Division of Infection, UMDS (St Thomas' Hospital), London.
- Eichler, J. & G. Ackermann (1993): *Holographie*, Springer, Wien.
- Eichmann, K.-U. (2001): *Die Ozonverteilungen der Nordhemisphäre 1997–2000 Gemessen mit GOME: Einfluss Von Dynamik und Chemie in der Stratosphäre*, Dissertation, Universität Bremen, IUP.
- Eichmann, K.-U., K. Bramstedt, M. Weber, V. V. Rozanov, R. Hoogen & J. P. Burrows (1999): "O₃ profiles from GOME satellite data - II: Observations in the Arctic spring 1997 and 1998", *Physics and Chemistry of the Earth* **24**(5), 453–457.
- Eisinger, M., A. Richter, J. Burrows & A. Piters (1996): "Studies on DOAS ozone column retrieval from the UV and visible measurements of GOME", in: *GOME Geophysical Validation Campaign Proceedings- Final Results (24–26 January, Frascati, 1996)*, ESA-WPP 108, S. 161–174, European Space Agency.

- Eisinger, M., A. Richter, A. Ladstädter-Weissenmayer & J. Burrows (1997): "DOAS zenith sky observations: 1. BrO measurements over Bremen 53°N 1993–1994", *J. Atmos. Chem.* **26**, 93–108.
- ESA (1995): *GOME User's Manual*, European Space Agency.
- Escobar-Munõz, J., A. Chedin, F. Chérüy & N. Scott (1993): "Multi-layer neural networks for the retrieval of atmospheric variables from satellite-borne vertical sounding", *Atmospheric Physics-Numerical Analysis* **317**, 911–918.
- Eyre, J. R. (1989): "Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation", *Quart. J. Roy. Soc.* **115**, 1001–1037.
- Fabian, P. (1992): *Atmosphäre und Umwelt*, Springer Verlag, Berlin.
- Fahlman, S. E. (1989): "An empirical study of learning speed in back-propagation networks", in: D. Touretzky, G. Hinton & T. Sejnowski (Hg.), *Proc. Of the 1988 Connectionist Models Summer School*, Morgan Kaufmann, Carnegie Mellon University.
- Farman, J. C., B. G. Gardiner & J. D. Shanklin (1985): "Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction", *Nature* **315**, 207–210.
- Fioletov, V. E., J. B. Kerr, E. W. Hare, G. J. Labow & R. D. McPeters (1999): "An assessment of the world ground-based total ozone network performance from the comparison with satellite data", *J. Geophys. Res.* **104**(D1), 1737–1747.
- Fogel, D. B., L. J. Fogel & V. W. Porto (1990): "Evolving neural networks", *Biological Cybernetics* **63**, 487–493.
- Fortuin, J. P. F. & H. Kelder (1998): "An ozone climatology based on ozonesonde and satellite measurements", *J. Geophys. Res.* **103**(D24), 31079–31734.
- Frederick, J. E., R. P. Cebula & D. F. Heath (1986): "Instrument characterization for the detection of long-term changes in stratospheric ozone: An analysis of the SBUV/2 radiometer", *J. Atmos. Ocean. Tech.* **3**, 472–480.
- Fujimoto, T., O. Uchino & T. Nagai (1998): "Four year intercomparison of MRI ozone lidar, ozonesonde and SAGE II data", in: R. Bojkov & G. Visconti (Hg.), *Proceedings of the XVIII Quadrennial Ozone Symposium, 1996*, L'Aquila, Italy.
- Glaccum, W., R. Lucke, R. M. Bevilacqua, E. P. Shettle, J. S. Hornstein, D. T. Chen, J. D. Lumpe, S. S. Krigman, D. J. Debrestian, M. D. Fromm, F. Dalaudier, E. Chassefiere, C. Deniel, C. E. Randall, D. W. Rusch, J. J. Olivero, C. Brogniez, J. Lenoble & R. Kremer (1996): "The Polar Ozone and Aerosol Measurement (POAM II) instrument", *J. Geophys. Res.* **101**(D9), 14479–14487.
- Golub, G. & W. Kahan (1965): "Calculating the pseudo-inverse and singular values of a matrix", *SIAM Numerical Analysis*, *B* **2**, 205–224.
- Goody, R. M. & Y. L. Yung (1989): *Atmospheric Radiation*, Oxford Univ. Press, Oxford, 2. Aufl.
- Grainger, J. F. & J. Ring (1962): "Anomalous Fraunhofer profile lines", *Nature* **163**, 762–765.
- Greenblatt, G. D., J. J. Orlando, J. B. Burkholder & A. R. Ravishankara (1990): "Absorption measurements of oxygen between 330 and 1140 nm", *J. Geophys. Res.* **95**(D11), 18577–18582.

- Guzzi, R., T. Kurosu & J. Burrows (1998): *GOME Cloud and Aerosol Data Products*, Final Report ESA Contract 11572/95/NL/CN, European Space Agency, ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands.
- Hadji-Lazaro, J., C. Clerbaux & S. Thiria (1999): “An inversion algorithm using neural networks to retrieve atmospheric CO total columns from high-resolution nadir radiances”, *J. Geophys. Res.* **104**(D19), 23841–23854.
- Hare, E. W. & V. E. Fioletov (1998): “An examination of the total ozone data in the World Ozone and Ultraviolet Radiation Data Centre”, in: R. Bojkov & G. Visconti (Hg.), *Proceedings of the XVIII Quadrennial Ozone Symposium, 1996*, L’Aquila, Italy.
- Harwood, M. H. & R. L. Jones (1994): “Temperature dependent UV-visible absorption cross sections of NO₂ and N₂O₄: Low-temperature measurements of the equilibrium constant for 2NO₂ ⇌ N₂O₄”, *J. Geophys. Res.* **99**(D11), 22955–22964.
- Heath, D. F., C. L. Mateer & A. J. Krueger (1973): “The Nimbus 4 Backscatter Ultraviolet (BUV) atmospheric ozone experiment – two years of operation”, *Pure and Applied Geophys.* **106**, 1238–1253.
- Heath, D. F. & H. Park (1978): “The Solar Backscatter Ultraviolet (SBUV) and Total Ozone Mapping Spectrometer (TOMS) experiment”, in: C. R. Madrid (Hg.), *The Nimbus-7 Users’ Guide*, S. 175–211, NASA Goddard Space Flight Center, Greenbelt, Maryland.
- Herman, J. R., P. K. Bhartia, A. J. Krueger, R. D. McPeters, C. G. Wellemeyer, C. J. Seftor, G. Jaross, B. M. Schlesinger, O. Torres, G. Labow, W. Byerly, S. L. Taylor, T. Swissler, R. P. Cebula & X.-Y. Gu (1996): *Meteor-3 Total Ozone Mapping Spectrometer (TOMS) Data Products User’s Guide*, Reference publication, NASA.
- Hervig, M. E., J. M. R. III, L. L. Gordley, J. Daniels, S. R. Drayson & J. H. Park (1995): “Aerosol effects and corrections in the Halogen Occultation Experiment”, *J. Geophys. Res.* **100**, 1067–1079.
- Hervig, M. E. & M. McHugh (1999): “Cirrus detection using HALOE measurements”, *Geophys. Res. Lett.* **26**, 719–722.
- Hilsenrath, E., W. Attmannspacher, A. Bass, W. Evans, R. Hagemeyer, R. Barnes, W. Komhyr, K. Mauersberger, J. Mentall, M. Proffitt, D. Robbins, S. Taylor, A. Torres & E. Weinstock (1986): “Results from the balloon ozone intercomparison campaign (BOIC)”, *J. Geophys. Res.* **91**, 13137–13152.
- Holton, J. R. & P. H. Haynes (1995): “Stratosphere-troposphere exchange”, *Rev. Geophys.* **33**, 403–439.
- Hood, L. L., B. E. Soukharev, M. Fromm & J. P. McCormack (2001): “Origin of extreme ozone minima at middle to high northern latitudes”, *J. Geophys. Res.* **106**(D18), 20925–20940.
- Hoogen, R. (1998): *Inversion Globaler Ozonvertikalverteilungen Aus Messungen Des Satelliteninstrumentes GOME*, Dissertation, Universität Bremen, IUP.
- Hoogen, R., V. Rozanov, R. de Beek, K. Bramstedt, K.-U. Eichmann, M. Weber, M. Buchwitz, T. Kurosu & J. P. Burrows (1997): “Ozone profiles from GOME satellite data – Part I: Advances in retrieval algorithm development”, in: *Proc. Fourth European Symposium on Stratospheric Ozone Research*, ESA, Schliersee.

- Hoogen, R., V. V. Rozanov & J. P. Burrows (1999): "Ozone profiles from GOME satellite data: Algorithm description and first validation", *J. Geophys. Res.* **104**(D7), 8263–8280.
- Hornik, K., M. Sinchcombe & H. White (1989): "Multilayer feedforward networks are universal approximators", *Neural Networks* **2**, 359–366.
- Hsieh, W. W. & B. Tang (1998): "Applying neural network models to prediction and data analysis in meteorology and oceanography", *Bull. Am. Meteorolog. Soc.* **79**, 1855–1870.
- Huang, S. C. & Y. F. Huang (1991): "Bounds on number of hidden neurons in multilayer perceptrons", *IEEE Transactions on Neural Networks* **2**, 47–55.
- Jacobs, R. (1988): "Increased rates of convergence through learning rate adaptation", *Neural Networks* **1**(4).
- Jiménez, C. (2000): *Inversion of Microwave Limb Sounding Observations of the Atmosphere by a Neural Network Technique*, Techn. Ber. 364L, Dept. Radio and Space Science, Chalmers University of Technology, Göteborg, Sweden.
- Jiménez, C. & P. Eriksson (2001): "A neural network technique for inversion of atmospheric observations from microwave limb sounders", *Radio Sci.* **36**(5), 941–954.
- JMA (1991): *Annual Report on Monitoring the Ozone Layer*, Techn. Ber. 2, Japan Meteorological Agency.
- Kaifel, A. K. & M. D. Müller (1999): "Results of neural network total ozone retrieval on global NOAA-TOVS data", in: J. Le Marshall & J. D. Jasper (Hg.), *Techn. Proc. 10th Int. ATOVS Study Conference, Boulder, CO, USA, 27th January – 2nd February 1999*, S. 262–275, Bureau of Meteorology Research Centre, Melbourne, Australia.
- Kaifel, A. K. & M. D. Müller (2000): *Bestimmung der Gesamtozonsäule mit Neuronalen Netzen und 3-D Analyse der Atmosphäre mit Dem '3I'-System*, Techn. Ber., Abschlußbericht BMBF-Vorhaben 01 LO 9514/8, Zentrum für Sonnenenergie- und Wasserstoff-Forschung, Stuttgart.
- Kaifel, A. K. & M. D. Müller (2001): "Results of TOVS ozone retrieval with neural networks", in: J. Le Marshall & J. D. Jasper (Hg.), *Techn. Proc. 11th Int. ATOVS Study Conference, Budapest, Hungary, 20–26 September 2000*, S. 153–165, International ATOVS Working Group, Bureau of Meteorology Research Centre, Melbourne, Australia.
- Kamgar-Parsi, B. & J. Gualteri (1990): "Solving inversion problems with neural networks", in: *Proc. Of Int. Joint Conf. On Neural Networks*, S. 955–960.
- Kerr, J. B., I. A. Asbridge & W. F. J. Evans (1988): "Intercomparison of total ozone measured by the Brewer and Dobson spectrometers at Toronto", *J. Geophys. Res.* **93**(11), 11129–11140.
- Kerr, J. B., H. Fast, C. T. McElroy, S. J. Oltmans, J. A. Lathrop, E. Kyro, A. Paukkunen, H. Claude, U. Köhler, C. R. Sreedharan, T. Takao & Y. Tsukagoshi (1994): "The 1991 WMO international ozonesonde intercomparison at Vanscoy, Canada", *Atmosphere-Ocean* **32**(4), 685–716.
- Kerr, J. F. R., A. H. Wyllie & A. R. Currie (1972): "Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics", *British J. of Cancer* **26**, 239–257.
- Khattatov, B. V., J.-F. Lamarque, L. V. Lyjak, R. Menard, P. F. Levelt, X. X. Tie, J. C. Gille & G. P. Brasseur (2000): "Assimilation of satellite observations of long-lived chemical species in global chemistry-transport models", *J. Geophys. Res.* **105**, 29135–29154.

- Koelemeijer, R. B. A. & P. Stammes (1999): "Effects of clouds on ozone column retrieval from GOME UV measurements", *J. Geophys. Res.* **104**(D7), 8281–8294.
- Kohavi, R. (1995): "A study of cross-validation and bootstrap for accuracy estimation and model selection", in: *Proc. Int. Joint Conf. On Artificial Intelligence*.
- Komhyr, W. D. (1969): "Electrochemical concentration cell for gas analysis", *Ann. Geophys.* **25**, 203–210.
- Komhyr, W. D. (1980): *Operations Handbook: Ozone Observations with a Dobson Spectrophotometer*, WMO Ozone Rep. 6, World Meteorol. Organ., Geneva.
- Komhyr, W. D., R. A. Barnes, G. B. Brothers, J. A. Lathrop & D. A. Oppermann (1995): "Electrochemical concentration cell ozonesonde performance evaluation during STOIC 1989", *J. Geophys. Res.* **100**, 9231–9244.
- Komhyr, W. D., G. C. Reinsel, R. D. Evans, D. M. Quincy, R. D. Grass & R. K. Leonard (1997): "Total ozone trends at sixteen NOAA/CMDL and cooperative dobson spectrometer observatories during 1979–1996", *Geophys. Res. Lett.* **24**, 3225–3228.
- Kondratyev, K. Y. & C. A. Varotsos (1996): "Global ozone dynamics and environmental safety", *ESPR - Environ. Sci. and Pollut. Res.* **3**, 153–157.
- Krasnopolsky, V. M. (1997): *Neural Networks for Standard and Variational Satellite Retrievals*, Techn. Ber., Contribution 148, NOAA Environmental Modeling Center, Ocean Modeling Branch, Washington (DC).
- Kraus, A. B., F. Rohrer, E. S. Grobler & D. H. Ehhalt (1996): "The global tropospheric distribution of NO_x estimated by a three-dimensional chemical tracer model", *J. Geophys. Res.* **101**, 18 587–18 604.
- Kuze, A. & K. Chance (1994): "Analysis of cloud top height and cloud coverage from satellites using the O_2 a and b bands", *J. Geophys. Res.* **99**, 14481–14491.
- Lambert, J.-C., M. V. Roozendaal, M. D. Mazière, P. C. Simon, J.-P. Pommereau, F. Goutail, A. Sarkissian & J. F. Gleason (1999): "Investigation of pole-to-pole performances of spaceborne atmospheric chemistry sensors with the NDSC", *J. Atmos. Sciences* **56**(2), 176–193.
- Lang, K. J. & M. J. Witbrock (1989): "Learning to tell two spirals apart", in: D. Touretzky & G. E. H. T. Sejnowski (Hg.), *Proc. Of the 1988 Connectionist Models Summer School*, Morgan Kaufmann, Carnegie Mellon University.
- Lelieveld, J. & F. J. Dentener (2000): "What controls tropospheric ozone?", *J. Geophys. Res.* **105**(D3), 3531–3551.
- Levy, H., II, J. D. Mahlman, W. J. Moxim & S. C. Liu (1985): "Tropospheric ozone: The role of transport", *J. Geophys. Res.* **1997**(90), 3753–3771.
- Liou, K.-N. (1980): *An Introduction to Atmospheric Radiation*, Academic Press, San Diego.
- Logan, J. A. (1994): "Trends in the vertical distribution of ozone: An analysis of ozonesonde data", *J. Geophys. Res.* **99**(D12), 25553–25585.
- Logan, J. A. (1999): "An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models and development of a gridded climatology for tropospheric ozone", *J. Geophys. Res.* **104**(D13), 16115–16149.

- Loibl, W. & S. Smidt (1996): "Ozone exposure: Areas of potential risk for selected tree species", *Environ. Sci. & Pollut. Res.* **3**, 213–217.
- Loyola, D. (1999): "Using artificial neural networks for the calculation of air mass factors", in: *ESAMS'99 - European Symposium on Atmospheric Measurements from Space, Noordwijk, Netherlands, 18–22 January 1999*, S. 573–576, ESA, WPP-161, Noordwijk.
- Lu, J., V. A. Mohnen, G. K. Yue, R. J. Atkinson & W. A. Matthews (1997): "Intercomparison of stratospheric ozone profiles obtained by SAGE II, HALOE, and ozonesondes during 1994–1995", *J. Geophys. Res.* **102**(D13), 16137–16144.
- Lucke, R. L., D. R. Korwan, R. M. Bevilacqua, J. S. Hornstein, E. P. Shettle, D. T. Chen, M. Daehler, J. D. Lumpe, M. D. Fromm, D. Debrestian, B. Neff, M. Squire, G. König-Langlo & J. Davies (1999): "The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results", *J. Geophys. Res.* **104**, 18.785–18.799.
- Lumpe, J. D., R. M. Bevilacqua, K. W. Hoppel & C. E. Randall (2002): "POAM III retrieval algorithm and error analysis", *J. Geophys. Res.* In press.
- Manney, G. L., H. A. Michelsen, R. M. Bevilacqua, , M. R. Gunson, F. W. Irion, N. J. Livesey, J. Oberheide, M. Riese, J. M. Russell III, G. C. Toon & J. M. Zawodny (2001): "Comparison of satellite ozone observations in coincident air masses in early November 1994", *J. Geophys. Res.* **106**, 9923–9944.
- Marquard, L. C., T. Wagner & U. Platt (2000): "Improved air mass factor concepts for scattered radiation differential optical absorption spectroscopy of atmospheric species", *J. Geophys. Res.* **105**(D1), 1315–1328.
- McCartney, E. J. (1983): *Absorption and Emission by Atmospheric Gasses – The Physical Processes*, John Wiley, New York.
- McElroy, M. B., R. J. Salawitch, S. C. Wofsy & J. A. Logan (1986): "Reductions of antarctic ozone due to synergistic interactions of chlorine and bromine", *Nature* **321**, 759–762.
- McPeters, R. D., A. J. Krueger, P. K. Bhartia, J. R. Herman, C. G. Wellemeyer, C. J. Seftor, G. Jaross, O. Torres, L. Moy, G. Labow, W. Byerly, S. L. Taylor, T. Swissler & R. P. Cebula (1998): *Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide*, Technical Publication 1998-206895, NASA.
- McPeters, R. D. & G. J. Labow (1996): "An assessment of the accuracy of 14.5 years of Nimbus 7 TOMS Version 7 ozone data by comparison with the Dobson Network", *Geophys. Res. Lett.* **23**, 3695–3698.
- Minsky, M. & S. Papert (1969): *Perceptrons*, MIT Press, Cambridge, MA.
- Motteler, H. E., L. L. Strow, L. McMillin & J. A. Gualteri (1995): "Comparison of neural network and regression-based method for temperature retrievals", *Appl. Opt.* **34**, 5390–5397.
- Müller, M. D. (1997): *Ein spektroskopisches Verfahren zur bodengestützten Messung des atmosphärischen Gesamtstickstoffdioxids NO₂*, Diplomarbeit, Universität Bayreuth.
- Müller, M. D., A. Kaifel, M. Weber & J. P. Burrows (2002a): "Neural network scheme for the retrieval of total ozone from Global Ozone Monitoring Experiment data", *Appl. Opt.* **41**(24), 5051–5058.

- Müller, M. D. & A. K. Kaifel (1999): "Efficient processing of multi-year global TOVS data using ITPP, 3I and neural networks", in: J. LeMarshall & J. D. Jasper (Hg.), *Techn. Proc. 10th Int. ATOVS Study Conference, Boulder, Colorado, 27 Jan – 2 Feb 1999*, S. 397–407, Int. ATOVS Study Working Group (ITWG), Bureau of Meteorology Research Centre, Melbourne, Australia.
- Müller, M. D., A. K. Kaifel & M. Weber (2001): "Ozone from GOME data using neural network technique", in: H. Sawaya-Lacoste (Hg.), *Proc. Of ERS-ENVISAT Symposium, Gothenburg, Sweden, 16–20 October 2000*, ESA-ESTEC, ESA Public. Div. (SP-461), Noordwijk, Netherlands.
- Müller, M. D., A. K. Kaifel, M. Weber & J. P. Burrows (2002b): "Partial training of neural networks with incomplete target data, applied to atmospheric science", *Neural Networks* Submitted.
- Müller, R., J.-U. Groöß, D. McKenna, P. Crutzen, C. Brühl, J. M. Russell & A. Tuck (1997): "HALOE observations of the vertical structure of chemical ozone depletion in the Arctic vortex during winter and early spring 1996-1997", *Geophys. Res. Lett.* **24**(22), 2717–2720.
- Newchurch, M. J., L. Bishop, D. Cunnold, L. E. Flynn, S. Godin, S. H. Frith, L. Hood, A. J. Miller, S. Oltmans, W. Randel, G. Reinsel, R. Stolarski, R. Wang, E.-S. Yang & J. M. Zawodny (2000): "Upper-stratospheric ozone trends 1979-1998", *Journal of Geophysical Research* **105**(D11), 14.625–14.636.
- Noxon, J. F. (1975): "Nitrogen dioxide in the stratosphere and troposphere measured by ground-based absorption spectroscopy", *Science* **189**, 547–549.
- Orlandini, S. & I. Morlini (2000): "Artificial neural network estimation of rainfall intensity from radar observations", *J. Geophys. Res.* **105**(D20), 24849–24861.
- Pasini, A., V. Pelino & S. Potestà (2001): "A neural network model for visibility nowcasting from surface observations: Results and sensitivity to physical input variables", *J. Geophys. Res.* **106**(D14), 14951–14959.
- Penrose, R. (1991): *Computerdenken*, Spektrum Verlag, Heidelberg.
- Perliski, L. M. & S. Solomon (1993): "On the evaluation of air mass factors for atmospheric near-ultraviolet and visible absorption spectroscopy", *J. Geophys. Res.* **98**(D6), 10363–10374.
- Perner, D. & U. Platt (1979): "Detection of nitrous acid in the atmosphere by differential optical absorption", *Geophys. Res. Lett.* **6**(12), 917–920.
- Pfeilsticker, K. & U. Platt (1994): "Airborne measurements during the arctic stratospheric experiment: Observation of O_3 and NO_2 ", *Geophys. Res. Lett.* **21**(13), 1375–1378.
- Planet, W. (1988): *Data Extraction and Calibration of TIROS-N/NOAA Radiometers*, NOAA Technical Memorandum NESS 107, U.S. Department of Commerce, NOAA, Washington DC, USA.
- Platt, U. (1994): "Differential Optical Absorption Spectroscopy (DOAS)", in: M. W. Sigrist (Hg.), *Air Monitoring by Spectroscopic Techniques*, S. 27–84, John Wiley, New York.
- Platt, U. & D. Perner (1983): "Measurements of atmospheric trace gases by long path differential UV/visible absorption spectroscopy", in: D. A. Killinger & A. Mooradian (Hg.), *Optical and Laser Remote Sensing*, S. 95–105, Springer Verlag, New York.

- Platt, U., D. Perner & H. W. Paetz (1979): "Simultaneous measurement of atmospheric CH₂O, O₃, and NO₂ by differential optical absorption", *J. Geophys. Res.* **84**, 6329–6335.
- Plutovskii, M. E. P. (1996): *Survey: Cross-Validation in Theory and in Practice*, Techn. Ber., Department of Computational Science Research, David Sarnoff Research Center, Princeton, New Jersey, USA.
- Poggio, T. & F. Girosi (1989): *A Theory of Networks for Approximation and Learning*, A.I. Memo 1140, MIT.
- Press, W. H., B. P. Flannery, S. A. Teukolsky & W. T. Vetterling (1986): *Numerical Recipes*, Cambridge University Press, New York.
- Reid, S. J., A. F. Tuck & G. Kiladis (2000): "On the changing abundance of ozone minima at northern midlatitudes", *J. Geophys. Res.* **105**, 12169–12180.
- Remsberg, E. E., P. P. Bhatt & L. E. Deaver (2001): "Ozone changes in the lower stratosphere from the Halogen Occultation Experiment for 1991 through 1999", *J. Geophys. Res.* **106**(D2), 1639–1653.
- Richards, J. A. (1993): *Remote Sensing Digital Image Analysis: An Introduction*, Springer Verlag, New York, 2. Aufl.
- Richaume, P., F. Badran, M. Crepon, C. Mejía, H. Roquet & S. Thiria (2000): "Neural network wind retrieval from ERS-1 scatterometer data", *J. Geophys. Res.* **105**(C4), 8737–8752.
- Richter, A. (1997): *Absorptionsspektroskopische Messungen stratosphärischer Spurengase über Bremen, 53° N*, Dissertation, Universität Bremen, Institute of Environmental Physics.
- Richter, A. & T. Wagner (2001): *Diffuser Plate Spectral Structures and their Influence on GOME Slant Columns*, Techn. Ber., Institute für Umweltphysik: Universität Bremen und Universität Heidelberg.
- Riedmiller, M. (1995): *Untersuchungen zu Konvergenz und Generalisierungsfähigkeit überwachter Lernverfahren mit dem SNNS*, Techn. Ber., Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe.
- Riedmiller, M. & H. Braun (1993): "A direct adaptive method for faster backpropagation learning: The Rprop algorithm", in: *Proc. of the ICNN*, S. 586–591, IEEE, San Francisco.
- Robert, J.-M. (1995): *Nervenkitzel*, Spektrum Akademischer Verlag, Heidelberg.
- Rodgers, C. D. (1976): "Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation", *Rev. Geophys. Space Phys.* **14**(4), 609–624.
- Rodgers, C. D. (1990): "Characterisation and error analysis of profiles retrieved from remote sounding measurements", *J. Geophys. Res.* **95**(D5), 5587–5595.
- Rodgers, C. D. (2000): *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing, Singapore.
- Roedel, W. (2000): *Physik Unserer Umwelt – Die Atmosphäre*, Springer Verlag, Berlin, 3. Aufl.
- Rosenblatt, F. (1958): "The perceptron: A probabilistic model for information storage and organization in the brain", *Psychological Review* **65**(6), 386–408.

- Rosenlof, K. H. (1995): "Seasonal cycle of the residual meridional circulation in the stratosphere", *J. Geophys. Res.* **100**, 5173–5191.
- Rossby, C.-G. (1938): "Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacement of the semi-permanent centers of action", *J. Marine Research (Sears Foundation)* **2**, 38–55.
- Rozanov, A. (2001): *Modeling of Radiative Transfer Through a Spherical Planetary Atmosphere: Application to Atmospheric Trace Gases Retrieval from Occultation-and Limb-Measurements in UV-Vis-NIR*, Dissertation, Universität Bremen, IUP.
- Rozanov, V., D. Diebel, R. J. D. Spurr & J. P. Burrows (1997): "GOMETRAN: A radiative transfer model for the satellite project GOME, The plane-parallel version", *J. Geophys. Res.* **102**, 16683–16695.
- Rozanov, V., Y. Timofeyev, M. Biryulina, J. Burrows, R. Spurr & D. Diebel (1993): "Accuracy of atmospheric constituent retrieval from multichannel remote sensing instruments", in: S. Keevallik & O. Kärner (Hg.), *IRS '92: Current Problems in Atmospheric Radiation*, S. 394–397, Deepak Publishing, Ganotibm VAm USA.
- Rozanov, V. V., T. Kurosu & J. P. Burrows (1998): "Retrieval of atmospheric constituents in the UV/visible: A new analytical approach to calculating weighting functions", *J. Quant. Spectrosc. Rad. Transfer* **60**, 277–299.
- Ruiz-Suárez, J. C., O. A. Mayora-Ibarra, J. Torres-Jiménez & L. G. Ruiz-Suárez (1995): "Short-term ozone forecasting by artificial neural networks", *Adv. Eng. Soft.* **23**, 143–149.
- Rumelhart, D. E., G. Hinton & R. Williams (1986): "Learning internal representations by error backpropagation", in: D. Rumelhart & J. McClelland (Hg.), *Parallel Distributed Processing, Vol. 1*, Bd. 1, S. 318–362, MIT Press.
- Rusch, D. W., R. M. Bevilacqua, C. E. Randall, J. D. Lumpe, K. W. Hoppel, M. D. Fromm, D. J. Debrestian, J. J. Olivero, J. S. Hornstein, F. Guo & E. P. Shettle (1997): "Validation of POAM II ozone measurements with coincident MLS, HALOE, and SAGE II observations", *J. Geophys. Res.* **102**, 23615–23627.
- Rusch, D. W., C. E. Randall, M. T. Callan, M. Horanyi, R. T. Clancy, S. C. Solomon, S. J. Oltmans, B. J. Johnson, U. Koehler, H. Claude & D. De Muer (1998): "A new inversion for stratospheric aerosol and gas experiment ii data", *J. Geophys. Res.* **103**(D7), 8465–8476.
- Russell, L. J., L. L. Gordley, J. H. Park, S. R. Drayson, W. D. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries & J. P. Crutzen (1993): "The halogen occultation experiment", *J. Geophys. Res.* **98**, 10777–10797.
- Satchwell, C. (1994): "Neural networks for stochastic problems: More than one outcome for the input space", Presented at NCAF Conference, Aston University.
- Schiffmann, W., M. Joost & R. Werner (1992): *Synthesis and Performance Analysis of Multilayer Neural Network Architectures*, Techn. Ber. 16/1992, Univ. of Koblenz, Inst. of Physics.
- Schwander, H., A. Kaifel, A. Ruggaber & P. Koepke (2001): "Spectral radiative-transfer modeling with minimized computation time by use of a neural-network technique", *Appl. Opt.* **40**(3), 331–335.
- Schwander, H., P. Koepke, A. Kaifel & G. Seckmeyer (2002): "Modification of spectral UV irradiance by clouds", *J. Geophys. Res.* **107**(D16), AAC 7–1–7–12.

- Selkoe, D. J. (1992): "Alterndes Gehirn – alternder Geist", *Spektrum der Wissenschaft* (11), 124–132.
- Singer, S. F. & R. C. Wentworth (1957): "A method for the determination of the vertical ozone distribution from a satellite", *J. Geophys. Res.* **62**, 299–308.
- Skov, H., A. H. Egeløv & K. Granby (1997): "Relationships between ozone and other photochemical products at Ll. Valby, Denmark", *Atmospheric Environment* **31**(5), 685–691.
- Slaper, H. (1995): "Comparing ground-level spectrally resolved solar UV measurements using various instruments: A technique resolving effects of wavelength shift and slit width", *Geophys. Res. Lett.* **22**(20), 2721–2724.
- Slijkhuis, S. & D. Loyola (1999): *GOME Data Processor Extraction Software User's Manual*, Techn. Ber. ER-SUM-DLR-GO-0045, Deutsches Zentrum für Luft- und Raumfahrt, Deutsches Fernerkundungs-Datenzentrum, Oberpfaffenhofen, Germany.
- Smit, H. G. J., W. Sträter, M. Helten, D. Kley, D. Ciupa, H. Claude, U. Köhler, B. Hoegger, G. Levrat, B. Johnson, S. J. Oltmans, J. B. Kerr, D. W. Tarasick, J. Davies, M. Shitamichi, S. K. Srivastav, C. Vialle & G. Velghe (1998): "JOSIE: The 1996 WMO international intercomparison of ozonesondes under quasi flight conditions in the environmental simulation chamber at Jülich", in: R. Bojkov & G. Visconti (Hg.), *Proceedings of the XVIII Quadrennial Ozone Symposium, 1996*, L'Aquila, Italy.
- Solheim, F., J. R. Godwin, E. R. Westwater, Y. Han, S. J. Keihm, K. Marsh & R. Ware (1998): "Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods", *Radio Science* **33**, 393–404.
- Solomon, S. & D. W. R. R. Garcia, F. S. Rowland (1986): "On the depletion of antarctic ozone", *Nature* **321**, 755–758.
- Solomon, S., A. L. Schmeltekopf & R. W. Saunders (1987): "On the interpretation of zenith sky absorption measurements", *J. Geophys. Res.* **92**(D7), 8311–8319.
- Sonnemann, G. (1992): *Ozon*, Akademie Verlag GmbH, Berlin.
- Spänkuch, D. & E. Schulz (1995): "Diagnosing and forecasting total column ozone by statistical relations", *J. Geophys. Res.* **100**, 18873–18885.
- Spänkuch, D. & E. Schulz (1998): "Statistical estimate of the vertical ozone structure", *Quarterly Journal of the Hungarian Meteorological Service* **102**(4), 201–218.
- SPARC (1998): *SPARC/IO₃C/GAW Assessment of Trends in the Vertical Distribution of Ozone*, WCRP-SPARC Report WMO TD No. 935, World Meteorological Organization.
- Spichtinger-Radowsky, N. (2001): *Ozonloch und thermische Struktur der antarktischen Stratosphäre: Zeitliche Entwicklung von Ende der fünfziger Jahre bis heute*, Dissertation, Ludwig-Maximilians-Universität München, Fakultät für Geowissenschaften.
- Spurr, R. (1996): *GOME Level 1 to 2 Algorithms Description*, Techn. Ber. ER-TN-DLR-GO-0025, Deutsches Zentrum für Luft- und Raumfahrt, Deutsches Fernerkundungs-Datenzentrum, Oberpfaffenhofen, Germany.
- Steele, H. M. & R. P. Turco (1997): "Separation of aerosol and gas components in the Halogen Occultation Experiment and the Stratospheric Aerosol and Gas Experiment (SAGE) II extinction measurements: Implications for SAGE II ozone concentrations and trends", *J. Geophys. Res.* **102**(D16), 19665–19681.

- Stinchombe, M. & H. White (1989): "Universal approximation using feedforward networks with non-sigmoid hidden layer activation functions", in: *Proc. Int. Joint Conf. On Neural Networks, Vol.*, S. 613–617.
- Swinbank, R. & A. O'Neill (1994): "A stratosphere-troposphere data assimilation system", *Monthly Weather Review* **122**.
- Tamura, S. & M. Tateishi (1997): "Capabilities of a four-layered feedforward neural network: Four layers versus three", *IEEE Transactions on Neural Networks* **8**(2), 251–255.
- Tanzi, C. P., E. Hegels, I. Aben & K. Bramstedt (1999): "GOME polarization monitoring: The first three years", Proc. 25th Annual European Optical Meeting on Atmospheric studies by Optical Methods.
- Thompson, A. M., J. C. Witte, R. D. McPeters, S. J. Oltmans, F. J. Schmidlin, J. A. Logan, M. Fujiwara, V. W. J. H. Kirchhoff, F. Posny, G. J. R. Coetzee, B. Hoegger, S. Kawakami, T. Ogawa, B. J. Johnson, H. Voemel & G. Labow (2001): "The 1998-2000 SHADOZ (Southern Hemisphere Additional Ozonesondes) tropical ozone climatology: Comparisons with TOMS and ground-based measurements", *submitted to J. Geophys. Res.* .
- Tikhonov, A. N. & V. Y. Arsenin (1977): *Solutions of Ill-Posed Problems*, V. H. Winston, Washington, DC.
- Tollenaere, T. (1990): "SuperSAB: Fast adaptive backpropagation with good scaling properties", *Neural Networks* **3**(5).
- Tresp, V. (1995): "Die besonderen eigenschaften neuronaler netze bei der approximation von funktionen", *KI* **3**, 12–17.
- Tuinder, O. N. E., R. de Winter-Sorkina & P. J. H. Builtjes (2002): "Retrieval methods of effective cloud cover for the GOME instrument: An intercomparison", *Atmos. Chem. Phys. Discuss.* **2**, 623–668.
- van der A, R. (1999): "Ozone profile retrieval and its application to the polar view mode of GOME", in: *ESAMS'99 - European Symposium on Atmospheric Measurements from Space, Noordwijk, Netherlands, 18-22 January 1999*, S. 295–298, ESA, WPP-161, Noordwijk.
- van der A, R. J., A. J. M. PETERS, R. F. van Oss, P. J. M. Valks, J. H. G. M. van Geffen, H. M. Kelder & C. Zehner (2001): "Near-real time delivery of GOME ozone profiles", in: H. Sawaya-Lacoste (Hg.), *Proc. Of ERS-ENVISAT Symposium, Gothenburg, Sweden, 16-20 October 2000*, ESA-ESTEC, ESA Public. Div. (SP-461), Noordwijk, Netherlands.
- van Oss, R. F. & R. Spurr (2001): "Fast and accurate ozone profile retrieval for GOME-2 using 4-stream LIDORT", in: H. Sawaya-Lacoste (Hg.), *Proc. Of ERS-ENVISAT Symposium, Gothenburg, Sweden, 16-20 October 2000*, ESA-ESTEC, ESA Public. Div. (SP-461), Noordwijk, Netherlands.
- von Bargaen, A., T. Kurosu, K. Chance, D. Loyola, B. Aberle & R. Spurr (2000): *ERS-2, Cloud Retrieval Algorithm for GOME (CRAG)*, Final report, Deutsches Zentrum für Luft- und Raumfahrt (DLR) und Smithsonian Astrophysical Observatory (SAO), ESA/ESTEC, PO Box 299, 2200 AG Noordwijk, The Netherlands.
- Vountas, M. (1998): *Modellierung und Parametrisierung des Ring Effektes: Einfluss auf die Bestimmung von stratosphärischen Spurengasen*, Dissertation, University of Bremen.

- Wang, H. J., D. M. Cunnold & X. Bao (1996): "A critical analysis of stratospheric aerosol and gas experiment ozone trends", *J. Geophys. Res.* **101**(D7), 12,495–12,514.
- Wardle, D. I., E. W. Hare, D. V. Barton & C. T. McElroy (1998): "The World Ozone and Ultraviolet Radiation Data Centre - Content and submission", in: R. Bojkov & G. Visconti (Hg.), *Proceedings of the XVIII Quadrennial Ozone Symposium, 1996*, L'Aquila, Italy.
- Wehner, R. & W. Gehring (1995): *Zoologie*, Georg Thieme Verlag, Stuttgart, 23rd Aufl.
- Wellemeier, C. G., S. L. Taylor, C. J. Seftor, R. D. McPeters & P. K. Barthia (1997): "A correction for total ozone mapping spectrometer profile shape errors at high latitude", *J. Geophys. Res.* **102**, 9029–9038.
- Wilks, D. S. (1995): *Statistical Methods in the Atmospheric Sciences*, Academic Press.
- Wirth, M. & W. Renger (1996): "Evidence of large scale ozone depletion within the arctic polar vortex 94/95 based on airborne lidar measurement", *Geophys. Res. Lett.* **23**, 813–816.
- WMO (1980): *Report of the Meeting of Experts on Assessment of Performance Characteristics of Various Ozone Observing Systems (Boulder, Colorado, USA, 1980)*, WMO Ozone Rep. 9, World Meteorological Organization, Geneva.
- WMO (1993): *Handbook for Ozone Data Reevaluation*, WMO Ozone Rep. 29, World Meteorological Organization, Geneva.
- WOUDC (1990): *Ozone Data for the World: Catalog of Ozone Stations and Catalogue of Ozone Data for 1985 - 1989*, Techn. Ber., Atmospheric Environment Service in cooperation with WMO, Downsview, Ontario, Canada.
- Zell, A. (1997): *Simulation Neuronaler Netze*, Oldenbourg Verlag, München, 2nd Aufl.

Danksagung

Ich möchte mich an dieser Stelle bei allen Personen bedanken, die mich in den vergangenen Jahren bei meiner Promotion unterstützt haben. Mein besonderer Dank gilt hierbei

- *Prof. Dr. John P. Burrows*, der sich bereit erklärt hat, mich als externen Doktoranden anzuerkennen und mir damit die Möglichkeit gab, dieses interessante Thema in die etablierte Retrieval-Fachwelt einzubringen. Ihm sowie auch *Prof. Dr. Jörn Bleck-Neuhaus* möchte ich außerdem für Aufmunterung und Ansporn in schwierigen Phasen der Arbeit danken.
- *Anton Kaifel*, der mir als Kollege, inoffizieller Chef und Betreuer am ZSW immer als Hauptansprechpartner zur Verfügung stand. Ohne die bereitwillige Weitergabe seiner reichhaltigen Erfahrung mit Satellitendaten, Programmierung und neuronalen Netzen wäre diese Arbeit nicht möglich gewesen.
- *Mark Weber* für die wissenschaftliche Betreuung der Arbeit, interessante Diskussionen, Anregungen, Datenbereitstellung, Übernachtungsmöglichkeit in Bremen, Korrekturlesen, und vieles mehr,
- meinen Kollegen *Oliver Reutter*, *Jasmine Kaptur* und *Michael Wohlfart* für das gute Arbeitsklima, die gute Zusammenarbeit und die Abnahme lästiger Netzwerkjobs in Stoßzeiten,
- *Silvia Tellmann* für die Unterstützung bei der Vorverarbeitung und Kollokation der Profildaten,
- *Rüdiger de Beek*, *Klaus Bramstedt*, *Kai-Uwe Eichmann*, *Andreas Richter* und *Vladimir Rozanov* für die Beantwortung von vielen Fragen und die Bereitstellung von Daten,
- *Greg Bodeker* für die Übermittlung seiner TOMS-Korrekturfunktionen,
- *Ed Hare* für die Hilfe mit den WOUDC-Daten,
- den Mitgliedern der GOPWG für Motivationsschübe und richtungsweisende Diskussionen,
- *Martina Mertens* für seelischen Beistand und Verbesserungstips bei der grafischen Gestaltung der Arbeit,
- meinen Eltern für ihre Zuneigung und Unterstützung jeglicher Art, und dafür, dass sie mir auch in hektischen Zeiten immer einen ruhigen Hafen zum konzentrierten Arbeiten und zum Sammeln neuer Energie zur Verfügung stellten.

Schließlich möchte ich noch den hunderten von anonymen Personen danken, durch deren unermüdlichen Einsatz die unzähligen Messungen von Bodenstationen, Ozonsonden und Satellitenprofilen entstanden, und öffentlich verfügbar gemacht wurden. Ohne sie hätte diese Arbeit gar nicht erst begonnen werden können.

Stuttgart, im September 2002