INHALT

1 EINLEITUNG 8

1.1 ZEITREIHENANALYSE UND SYSTEMTHEORIE 9
1.2 NEURONALE NETZE UND MATHEMATISCHE MODELLE 10
1.3 FRAGESTELLUNGEN 11
1.4 AUFBAU DER ARBEIT 12

2 DATEN 13

2.1 DATEN DER BIOLOGIE 13
2.2 DATEN DES DEUTSCHEN WETTERDIENSTES 15
2.2.1 Meßstation Kalkhügel 15
2.2.2 Zeitraum 1992-1993 16
2.2.3 Zeitraum 1994-1996 16
2.3 DATEN DER FACHHOCHSCHULE OSNABRÜCK 17
2.4 WEITERE DATEN 18
2.5 DATENVORVERARBEITUNG 18
2.5.1 Allgemeines 18
2.5.2 Die Windrichtungs-Zeitreihe 19
2.6 ÜBERSICHT 20

3 METHODEN DER DATENANALYSE UND -PROGNOSE 22

3.1 MATHEMATISCHE VERFAHREN 22
3.1.1 Multiple Regressionsanalyse 22
3.1.2 Klassisches Komponentenmodell 23
3.1.3 AR-Modellierung und ARIMA 25
3.2 NEURONALE NETZE 28
3.2.1 Einführung 28
3.2.2 Training neuronaler Netze 32
3.2.3 Arbeitsumgebung für neuronale Fragestellungen 36
3.2.4 Neuronale Sensitivitätsanalyse 37
3.2.5 Netzoptimierung: Pruning 37
3.3 MAßE FÜR DIE ANALYSE- UND PROGNOSEGÜTE 39
3.3.1 Ex-ante- und Ex-post-Beurteilung 40
3.3.2 Nichtnormierte Gütemaße 40
3.3.3 Normierte Gütemaße 40
3.4 VERFAHREN ZUR AUSWAHL EXOGENER ZEITREIHEN 41
3.4.1 Auswahl nach Korrelationskoeffizient 42
3.4.2 Auswahl nach Sensitivität 43

4 ERGEBNISSE 44

4.1 ANALYSE DER DATENREIHEN 44

4.1.1 Die CO_2-Datenreihe ... 44
4.1.2 Andere Datenreihen ... 47
4.2 VERGLEICH DER METHODEN ... 49
4.2.1 Vorgehen ... 49
4.2.2 Erwartungen ... 49
4.2.3 Der Analyse- bzw. Prognosezeitraum ... 50
4.2.4 Anzahl exogener Zeitreihen ... 52
4.2.5 Wahl der Netzstruktur ... 53
4.3 VERGLEICH DER VERFAHREN ... 55
4.3.1 Analyse ... 56
4.3.2 Prognose ... 60
4.3.3 ARIMA-Modellierung ... 63
4.4 VERFAHREN ZUR AUSWAHL EXOGENER ZEITREIHEN FÜR NEURONALE NETZE ... 67
4.4.1 Pruning ... 67
4.4.2 Auswahl nach Sensitivität ... 69
4.4.3 Vergleich der Verfahren ... 71

5 DISKUSSION ... 72

5.1 ANALYSE DER DATENREIHEN ... 72
5.2 VERGLEICH DER METHODEN ... 73
5.3 VERGLEICH DER VERFAHREN ... 74
5.3.1 Regression, Komponentenmodell und Neuronale Netze ... 74
5.3.2 ARIMA-Modellierung ... 74
5.4 VERFAHREN ZUR AUSWAHL EXOGENER ZEITREIHEN FÜR NEURONALE NETZE ... 75
5.4.1 Pruning ... 75
5.4.2 Auswahl nach Sensitivität ... 75

6 SCHLUßFOLGERUNGEN UND AUSBLICK ... 77

7 ANHANG ... 79

7.1 TABELLEN UND DIAGRAMME ... 79
7.2 GÜNSTIGE NETZARCHITEKTUREN FÜR DIE ANALYSE ... 79
7.2.1 Günstige Netzarchitekturen für die Prognose ... 82
7.3 INHALT DER CD-ROM ... 83
7.3.1 Übersicht über die Verzeichnisstruktur ... 83
7.3.2 Daten ... 83
7.3.3 Diplomarbeit ... 84
7.3.4 Ergebnisse ... 85
7.3.5 Skripte ... 86
7.4 LITERATURVERZEICHNIS ... 88
7.5 DANKSAGUNG ... 91
7.6 VERPFLICHTUNGSERKLÄRUNG ... 92

ABBILDUNGSVERZEICHNIS

Abbildung 1.1: Darstellung eines Systems mit unbekannter Systemstruktur (Black Box) 9

Abbildung 2.1: Monatlich gemittelte CO_2-Werte in Osnabrück (Westerberg) 14

Abbildung 2.2: Histogramm der CO_2-Werte für den betrachteten Zeitraum 14

Abbildung 2.3: Windrichtungen und –stärken für einen Tag mit den daraus resultierenden Tageswerten nach verschiedenen Verfahren 20

Abbildung 3.1: Die vier Phasen der Anpassung eines ARIMA[p,d,q]-Modells an eine gegebene Zeitreihe 27

Abbildung 3.2: Aufbau eines künstlichen Neurons 29

Abbildung 3.3: Struktur eines fünfschichtigen Feed-Forward-Netzes 30

Abbildung 3.4: Feed-Forward-Netz mit Vorschicht 31

Abbildung 3.5: Feed-Forward-Netz mit Shortcut-Verbindungen zwischen Hidden-Schicht #2 und Output-Schicht 31

Abbildung 3.6: Sigmoide Aktivierungsfunktion 33

Abbildung 3.7: Schematischer Ablauf des Netztrainings 34

Abbildung 3.8: Aufteilung der Wertepaare in Test-, Trainings- und Validierungsmenge (Beispiel) 35

Abbildung 3.9: Abbruchkriterien und Verlauf von Trainings- und Testfehler während des Netztrainings 36

Abbildung 3.10: Erweiterung des SNNS-Trainings um mehrfaches Pruning 39

Abbildung 4.1: Monatliche CO_2-Konzentration (Durchschnitt der Jahre 1992 - 1996) 44

Abbildung 4.2: Linearer Trend der CO_2-Konzentration über die Jahre 1992 bis 1996 45

Abbildung 4.3: Korrelogramm der CO_2-Zeitreihe 46

Abbildung 4.4: Neuronale Sensitivitäten der einzelnen exogenen Parameter 49

Abbildung 4.5: Repräsentierung der linearen Regression als Neuronales Netz 50

Abbildung 4.6: Vergleich der Analysegüte der drei Verfahren (Zeitraum März / April 1996) 56

Abbildung 4.7: (Partielle) Autokorrelationen zur undifferenzierten CO_2-Zeitreihe (19.11.94 bis 10.12.95) 64

Abbildung 4.8: Vergleich der Verfahren im dem Prognosezeitraum 1 (19.11.94 bis 10.12.95) 65

Abbildung 4.9: (Partielle) Autokorrelationen zur undifferenzierten CO_2-Zeitreihe (24.05.94 bis 17.11.94) 66

Abbildung 4.10: Vergleich der Verfahren in Prognosezeitraum 2 (24.05.94 bis 17.11.94) 67

Abbildung 4.11: Anzahl der Neuronen in Hidden- und Input-Schicht im Verlauf des Verfahrens 68

Abbildung 4.12: Fehlerverlauf des Sensitivitätsmodells zur Auswahl geeigneter Inputparameter..... 70

Abbildung 7.13: Übersicht über die Verzeichnisstruktur der CD-ROM 83

TABELLENVERZEICHNIS

Tabelle 2.1: Übersicht über fehlende Daten in der CO_2-Zeitreihe 15

Tabelle 2.2: Übersicht über die verwendeten Datenreihen..... 21

Tabelle 3.1: Übersicht über die verwendeten nichtnormierten Fehler- bzw. Gütemaße .. 40

Tabelle 3.2: Übersicht über die verwendeten normierten Fehler- bzw. Gütemaße 41

Tabelle 4.1: Korrelationen zur CO_2-Zeitreihe, nach absolutem Wert geordnet 47

Tabelle 4.2: Korrelationsmatrix (Teil 1)..... 48

Tabelle 4.3: Korrelationsmatrix (Teil 2)..... 48

Tabelle 4.4: Übersicht über die untersuchten Zeiträume..... 51

Tabelle 4.5: Übersicht über den Einfluß exogener Zeitreihen auf die Regression 53

Tabelle 4.6: Grundkonfiguration zur Bestimmung der günstigsten Netzparameter..... 54

Tabelle 4.7: Die 5 günstigsten Topologien für die Analyse (nach RMSE) 55

Tabelle 4.8: Die 5 günstigsten Topologien für die Prognose (nach RMSE) 55

Tabelle 4.9: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Januar/Februar 1996) 56

Tabelle 4.10: Die fünf günstigsten Netzarchitekturen für die Analyse (Januar/Februar 1996)..... 57

Tabelle 4.11: Gegenüberstellung der Modellergebnisse (Januar/Februar 1996)..... 57

Tabelle 4.12: Günstigste Kombinationen exogener Zeitreihen für die Analyse (März/April 1996)..... 57

Tabelle 4.13: Die fünf günstigsten Netzarchitekturen für die Analyse (März/April 1996) 57

Tabelle 4.14: Gegenüberstellung der Modellergebnisse (März/April 1996)..... 57

Tabelle 4.15: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Mai/Juni 1996)..... 57

Tabelle 4.16: Die fünf günstigsten Netzarchitekturen für die Analyse (Mai/Juni 1996) ... 58

Tabelle 4.17: Gegenüberstellung der Modellergebnisse (Mai/Juni 1996)..... 58

Tabelle 4.18: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Juli/August 1996)..... 58

Tabelle 4.19: Die fünf günstigsten Netzarchitekturen für die Analyse (Juli/August 1996) 58

Tabelle 4.20: Gegenüberstellung der Modellergebnisse (Juli/August 1996) 58

Tabelle 4.21: Günstigste Kombinationen exogener Zeitreihen für die Analyse (September/Oktober 1996) 58

Tabelle 4.22: Die fünf günstigsten Netzarchitekturen für die Analyse (September/Oktober 1996)..... 59

Tabelle 4.23: Gegenüberstellung der Modellergebnisse (September/Oktober 1996)..... 59

Tabelle 4.24: Günstigste Kombinationen exogener Zeitreihen für die Analyse (November/Dezember 1996)..... 59

Tabelle 4.25: Die fünf günstigsten Netzarchitekturen für die Analyse (November/Dezember 1996)..... 59

Tabelle 4.26: Gegenüberstellung der Modellergebnisse (November/Dezember 1996) 59

Tabelle 4.27: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Januar/Februar 1996)..... 60

Tabelle 4.28: Die fünf günstigsten Netzarchitekturen für die Prognose (Januar/Februar 1996)..... 60

Tabelle 4.29: Gegenüberstellung der Modellergebnisse (Januar/Februar 1996)..... 61

Tabelle 4.30: Günstigste Kombinationen exogener Zeitreihen für die Prognose (März/April 1996)..... 61

Tabelle 4.31: Die fünf günstigsten Netzarchitekturen für die Prognose (März/April 1996) 61

Tabelle 4.32: Gegenüberstellung der Modellergebnisse (März/April 1996)..... 61

Tabelle 4.33: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Mai/Juni 1996)..... 61

Tabelle 4.34: Die fünf günstigsten Netzarchitekturen für die Prognose (Mai/Juni 1996).. 61

Tabelle 4.35: Gegenüberstellung der Modellergebnisse (Mai/Juni 1996)..... 62

Tabelle 4.36: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Juli/August 1996)..... 62

Tabelle 4.37: Die fünf günstigsten Netzarchitekturen für die Prognose (Juli/August 1996) 62

Tabelle 4.38: Gegenüberstellung der Modellergebnisse (Juli/August 1996) 62

Tabelle 4.39: Günstigste Kombinationen exogener Zeitreihen für die Prognose (September/Oktober 1996)..... 62

Tabelle 4.40: Die fünf günstigsten Netzarchitekturen für die Prognose (September/Oktober 1996)..... 62

Tabelle 4.41: Gegenüberstellung der Modellergebnisse (September/Oktober 1996)..... 63

Tabelle 4.42: Günstigste Kombinationen exogener Zeitreihen für die Prognose (November/Dezember 1996)..... 63

Tabelle 4.43: Die fünf günstigsten Netzarchitekturen für die Prognose (November/Dezember 1996) 63

Tabelle 4.44: Gegenüberstellung der Modellergebnisse (November/Dezember 1996) 63

Tabelle 4.45: Vergleich der Prognosefehler der einzelnen Verfahren 65

Tabelle 4.46: Vergleich der Prognosefehler der einzelnen Verfahren 66

Tabelle 4.47: Ablauf des Input-Prunings 69

Tabelle 4.48: Übersicht über den Ablauf des Sensitivitätsmodells 70

Tabelle 4.49: Die günstigsten Inputparameter (nach Sensitivitätsmodell) mit Sensitivitäten 71

Tabelle 4.50: Vergleich der Verfahren zur Verbesserung der Netztopologie 71

Tabelle 7.51: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 1) 79

Tabelle 7.52: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 2) 80

Tabelle 7.53: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 2) 81

Tabelle 7.54: Netzarchitekturen und Prognosefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert 82

1 Einleitung

CO_2 zählt zu den Treibhausgasen, die sich in der Atmosphäre anreichern und einfallende Sonnenstrahlen zwar zur Erde durchlassen, diese aber bei der Wärmerückstrahlung ins All wieder zurück zum Boden reflektieren. Auf diese Weise wird die Erde – ähnlich einem Treibhaus – über das natürliche Maß hinaus erwärmt. Neben natürlichen CO_2-Quellen wie Feuern oder der Exhalation von Tier und Mensch gibt es noch eine Reihe von anthropogenen, also vom Menschen geschaffene CO_2-Quellen. Hierzu zählen alle vom Menschen initiierten Verbrennungsreaktionen, wie sie etwa beim Heizen, im Verkehr oder bei der Stromgewinnung über fossile Brennstoffe auftreten.

Im natürlichen Gleichgewicht wird das in der Atmosphäre vorhandene Kohlendioxid von Pflanzen zu Sauerstoff veratmet[1]. Der anthropogene Anteil hat jedoch insbesondere mit der Industrialisierung so stark zugenommen, daß dieser Mechanismus nicht mehr ausreicht, um das natürliche Gleichgewicht wieder herzustellen.

> „Kohlendioxid (CO_2) ist im Unterschied etwa zu Pestiziden oder Schwermetallen kein klassisches Gift, welches am Ort der Entstehung oder im direkten Kontakt mit Leben toxisch wirkt. Egal wo CO_2 ausgestoßen wird, es wirkt in jedem Fall global auf unser Klima.“[2]

Aufgrund der großen Bedeutung der Klimaproblematik und des starken Einflusses von CO_2 wird deutlich, daß es notwendig ist, auf den CO_2-Gehalt der Luft wirkende Einflußgrößen zu identifizieren, um die zukünftige CO_2-Entwicklung besser prognostizieren zu können und somit langfristig Aussagen über die Entwicklung der globalen CO_2-Konzentration und den damit verbundenen Treibhauseffekt machen zu können.

Die besondere Lage der CO_2-Meßstation am Westerberg in Osnabrück, deren Daten im Rahmen dieser Arbeit verwendet wurden, zeichnet sich zum einen durch ein urban geprägtes Umfeld aus und zum anderen durch ihre Nähe zu größeren Vegetationsflächen. Somit sind die Meßwerte sowohl von lokalen Emissionen geprägt als auch von der Pufferwirkung der umliegenden Vegetation[3], was die verwendete Zeitreihe weltweit einzigartig macht.

---

[1] Wie aus [For 93] 175 ff. hervorgeht, übt die Vegetation einen puffernden Effekt auf den CO_2-Gehalt der Luft aus.

[2] Aus [Her 98], 4.

[3] Vgl. [For 93], 175 ff.

1.1 Zeitreihenanalyse und Systemtheorie

Für die Betrachtung des Systems, welches den CO_2-Gehalt der Luft reguliert (kurz: CO_2-System) genügt es nicht, den Einfluß einer einzelnen Zeitreihe zu betrachten. Vielmehr ist die CO_2-Konzentration der Luft (in ppm[4]) das Ergebnis einer Vielzahl von Einzeleinflüssen. Demnach ist das CO_2-System ein dynamisches System nach der klassischen Systemtheorie[5], wobei bekannte Einflüsse wie etwa die Photosynthese der Vegetation oder lokale CO_2-Emissionen ebenso eine Rolle spielen wie unbekannte bzw. nicht offensichtliche. Ist die Beschaffenheit der Systemstruktur, d.h. der Einflüsse und Wechselwirkungen nicht bekannt, spricht man von einer *Black Box*. Vereinfacht ist dieser Zusammenhang in Abbildung 1.1 dargestellt.

Abbildung 1.1: Darstellung eines Systems mit unbekannter Systemstruktur (Black Box)

Schematisch werden Eingabegrößen oder Input von dem natürlichen CO_2-System transformiert. Unter diesen Größen werden vor allem exogene[6] Zeitreihen verstanden. Alle Wirkzusammenhänge ergeben zusammen eine entsprechende Ausgabegröße. Diese Ausgabegröße oder Output (endogenen Zeitreihe) stellt in unserem Fall der CO_2-Gehalt der Luft dar.

Da der Mechanismus des Systems unbekannt ist, kann dieser nur in Form von Hypothesen vorliegen. Bis ein hypothetische Modell den Input hinreichend gut auf den Output abbildet, muß es ggf. öfter angepaßt (verbessert) werden

Das CO_2-System ist durch ein solches Modell darstellbar. Die exogenen Zeitreihen (Input) als Indikatoren für verschiedenartige Umweltparameter und –einflüsse werden in Form statistischer Zeitreihen erfaßt, die Black Box vertritt die komplexen Wechselwirkungen der Natur, vor allem chemische und biochemische Prozesse. Um das System möglichst korrekt abzubilden, strebt man an, alle Zeitreihen in die Modellierung mit einfließen zu lassen, die einen Einfluß auf den Output des Systems haben. Dies ist jedoch i.a. nicht möglich, da die Anzahl der Zeitreihen mit Einfluß sehr groß sein kann, viele dieser möglichen Zeitreihen gar nicht bekannt sind oder Zeitreihen nicht in die

---

[4] Die Einheit ppm (engl.: *parts per million*) bezieht sich hier auf Volumenanteile in der Luft. In Anlehnung an die Fachliteratur wird in dieser Arbeit durchgängig ppm statt der genaueren Bezeichnung ppm/V verwendet, da die Bezugsgröße (Volumen) immer eindeutig ist.

[5] Entsprechende Definitionen finden sich in den systemwissenschaftlichen Standardwerken, wie bei [Kli 91] 5ff., 9ff.; [Kul 79] 34 f.; [Vog 83] 50 f. und anderen.

[6] Unter exogenen Zeitreihen werden die Daten verstanden, die mögliche äußere Einflüsse auf die CO_2-Zeitreihe erfassen, wie beispielsweise Temperatur, Sonnenscheindauer, etc. Die CO_2-Datenreihe als die zu modellierende Zeitreihe wird endogene Zeitreihe genannt.

Betrachtung einbezogen werden, da von ihnen kein Einfluß vermutet wird. Es ist deswegen nötig, sich bei der Modellierung auf einige aussagekräftige Zeitreihen zu beschränken.

Im Laufe der Entwicklung der Zeitreihenanalyse (ZRA) ist es möglich geworden, die Methodik immer gründlicher anzuwenden und die entsprechenden Systeme immer genauer abzubilden. Konnte man früher nur wenige Zeitreihen gleichzeitig betrachten, stehen heute mit Computern Mittel zu Verfügung, die es erlauben, die Anzahl der Zeitreihen sowie den Beobachtungszeitraum entscheidend zu vergrößern. Somit können Wechselwirkungen und eventuelle Regelmäßigkeiten besser erfaßt werden.

1.2 Neuronale Netze und Mathematische Modelle

Zur Identifizierung der Zusammenhänge zwischen der CO_2–Konzentration und anderen Klimavariablen werden in dieser Arbeit mathematische Modelle sowie Neuronale Netze verwandt. Unter anderem sollen dabei auch die Verfahren auf ihre Analyse- und Prognosegüte im Bezug auf die CO_2-Zeitreihe verglichen werden.

Während mathematische Modelle in der Zeitreihenanalyse und –prognose bereits bewährte Verfahren darstellen, war der Anwendungsbereich Neuronaler Netze lange Zeit auf reine Mustererkennungs- und Klassifizierungsaufgaben beschränkt. Im Bereich der Zeitreihenanalyse und –prognose ist die Theorie der Neuronalen Netze eine vergleichsweise neue und innovative Methode. So gibt es beispielsweise noch keine sicheren Erkenntnisse, welche Netzarchitekturen für welche Probleme am besten geeignet sind, sondern lediglich Erfahrungswerte.

Es stehen sich also die Neuronalen Netze als innovativer und noch relativ neuer Ansatz und die mathematischen Verfahren der ZRA (im folgenden als *mathematische Modelle* bezeichnet) als bewährte Mittel mit einem umfangreichen Werkzeugkasten gegenüber. In dieser Arbeit werden beide häufig als konkurrierende Paradigmen angesehene Methoden auf eine spezielle „echte“ Zeitreihe aus dem Umweltbereich angewandt und verglichen. Somit ergibt sich für diese Arbeit, daß sie zum einen experimentell und zum anderen methodisch ist.

Experimentell, da für diese konkrete Problemstellung wenig Erfahrungen bezüglich geeigneter Architekturen Neuronaler Netze vorliegen und viele Erkenntnisse erst durch geschicktes und systematisches Experimentieren mit den vorhanden Parametern ermöglicht werden.

Methodisch, da hier in erster Linie neue Erkenntnisse beim Vergleich der beiden Methoden der Zeitreihenanalyse und –prognose (Mathematische Modelle und Neuronale Netze) gemacht werden sollen. Eine inhaltliche Bearbeitung der Ergebnisse (z.B. Identifizierung und systemwissenschaftliche Begründung von Wirkzusammenhängen) soll sich aufbauend auf den Ergebnissen anschließen.

1.3 Fragestellungen

Folgende Fragestellungen sollen im Rahmen dieser Diplomarbeit erarbeitet werden:

- *Sind Neuronale Netze zur Analyse bzw. Prognose von Umweltdaten überhaupt geeignet?*

  Während mathematische Modelle schon seit einigen Jahren auch im naturwissenschaftlichen Bereich mit Erfolg eingesetzt werden, liegt das Hauptanwendungsgebiet der Neuronalen Netze nach wie vor im Bereich der Mustererkennung und Prognose ökonometrischer Zeitreihen. In den Fällen, in denen Neuronale Netze tatsächlich zur Prognose herangezogen werden, sind die zu Grunde liegenden Zeitreihen außerdem oft vollständig. Diese Arbeit soll einen entsprechenden Vergleich für die CO_2-Zeitreihe „as is“ durchführen, d.h. auch mit Fehlwerten, die bei naturwissenschaftlichen Zeitreihen durchaus nicht ungewöhnlich sind.

- *Welche Netzarchitekturen sind am besten geeignet?*

  Wie schon eingangs erwähnt, existieren keinesfalls gesicherte Erkenntnisse über die optimale Netzkonfiguration für ein konkretes Problem. Aufgrund der vielen Parameter ist es auch nicht möglich, alle denkbaren Kombinationen zu vergleichen. Hier sollen nun durch systematisches Ausprobieren Netzwerkkonfigurationen gefunden werden, die sich für die Analyse und Prognose der CO_2-Zeitreihe besonders gut eignen.

- *Welche Verfahren eignen sich besser: Neuronale Netze oder Mathematische Modelle?*

  Als zentraler Punkt dieser Arbeit sollen Neuronale Netze und mathematische Modelle in Bezug auf ihre Prognose- und Analysequalität verglichen werden. Aufgrund der Vielzahl von mathematischen Modellen werden hier die Standardverfahren der ZRA verwandt, um sie mit möglichst günstigen Netzarchitekturen zu vergleichen.

- *Wie lassen sich Modellparameter insbesondere für die neuronale Modellierung finden?*

  Neuronale Netze lassen dem Anwender erhebliche Freiheiten bei der Wahl der Netzarchitektur bzw. Topologie. Die einzige gesicherte Erkenntnis bei der Auswahl einer Topologie ist die, daß unterschiedliche Probleme von unterschiedlichen Architekturen am besten gelöst werden. Eine Übertragbarkeit von Erkenntnissen auf ein neues Problem ist also nur sehr begrenzt gegeben. Hier sollen Verfahren getestet werden, die eine geeignete Architektur selbständig finden.

- *Wie groß ist der Einfluß der einzelnen ZR und wie läßt er sich begründen?*

  Aus systemwissenschaftlicher Sicht ist es interessant, mögliche Einflüsse auf den CO_2-Gehalt der Luft zu identifizieren. Dabei wird rein methodisch vorge-

gangen: Zunächst wird die Sensitivität der einzelnen exogenen Zeitreihen nach den verschiedenen Verfahren bestimmt und für die Zeitreihen mit großem Einfluß ihr möglicher Einfluß auf den CO_2-Gehalt der Luft gesucht.

- *Welche Informationen kann man aus den Neuronalen Netzen gewinnen?*

  Da die Formeln mathematischer Modelle Aussagen über die Modellstruktur erlauben, können aus einem guten Modell ggf. Rückschlüsse über das System gezogen werden. Dies gestaltet sich bei den Neuronalen Netzen anders: Es existiert keine Modellgleichung im eigentlichen Sinn und die Information ist über das gesamte Netz verteilt. Folglich ist auch der Hauptkritikpunkt bei der Anwendung Neuronaler Netze im Rahmen der ZRA ihr Charakter als „Black Box“: Die Gewichte eines fertig trainierten Netzes sind zwar bekannt, über die nachgebildete Struktur des modellierten Systems kann man anhand dieser jedoch nur wenig Informationen gewinnen. Hier soll nun versucht werden, mittels der neuronalen Sensitivitätsanalyse Einflüsse und Wirkzusammenhänge zu identifizieren.

1.4 Aufbau der Arbeit

In Kapitel zwei werden die der Arbeit zu Grunde liegenden Daten und Datenquellen vorgestellt und die nötigen Schritte in der Datenvorverarbeitung erläutert.

Das dritte Kapitel enthält einen systematischen Überblick über die verwendeten mathematischen Modelle und Neuronalen Netze, auf deren historischen Hintergrund ebenfalls eingegangen wird. Weiterhin werden hier Kriterien für die vergleichende Analyse (Fehlermaße) vorgestellt.

Kapitel vier beinhaltet den empirischen Teil der Arbeit, der sich der Beantwortung der in Kapitel 1.3 gestellten Fragen widmet.

Im fünften Kapitel werden die Ergebnisse aus Kapitel vier diskutiert und die aus den Untersuchungen gewonnenen Erkenntnisse besprochen.

Das sechste Kapitel faßt die wichtigsten Ergebnisse und die zentralen gewonnenen Erkenntnisse kurz zusammen.

Kapitel sieben stellt ein Fazit zu den Untersuchungen dar. Hier wird auch ein Ausblick gegeben und angeregt, welche Teilbereiche näher untersucht werden sollten als es im Rahmen dieser Arbeit möglich ist.

2 Daten

Für die Zeitreihenanalyse und –prognose ist das Vorhandensein geeigneter Datenreihen von entscheidender Bedeutung. Alle Daten, die im Rahmen dieser Arbeit verwendet wurden, stammen aus vier verschiedenen Quellen: Die CO_2-Daten, die im Mittelpunkt der Untersuchung stehen, wurden vom Fachbereich Biologie der Universität Osnabrück zur Verfügung gestellt, außerdem wurden verschiedene Wetterdaten als Zeitreihen von dem Fachbereich Gartenbau der Fachhochschule Osnabrück sowie vom Deutschen Wetterdienst verwendet. Eigene Datenreihen, die zur saisonalen Modellierung dienen, runden das Spektrum ab. Sie umfassen u.a. jahreszeitliche Schwankungen.

In diesem Kapitel werden nun die einzelnen Zeitreihen und deren Quellen detailliert vorgestellt sowie die Methoden der Datenvorverarbeitung erläutert. Tabelle 2.2 am Ende dieses Kapitels zeigt eine Übersicht über die verwendeten Daten.

2.1 Daten der Biologie

Der Fachbereich Biologie der Universität Osnabrück betreibt am Westerberg (95m über NN; 58° 18‘ N, 8° 2‘ O) in der Arbeitsgruppe Ökologie seit 1981 regelmäßige Messungen des CO_2-Gehalts der Luft[7]. Ursprünglich im Rahmen einer Dissertation [For 93] verwendet, wurde die Anlage 1994 durch die heutige, modernere Meßstation ersetzt. Die Messungen erstreckten sich damals auf drei Klimazellen, in denen der Einfluß des CO_2-Gehalts der Luft auf ökologische Systeme herausgefunden werden sollte[8]. Seit 1986/1987 finden Außenmessungen der Luft statt.

Die Messungen nimmt alle 30 Sekunden ein Ultra-Rot-Absorptions-Spektroskop (URAS) in 6 Meter Höhe über dem Erdboden vor. Der Meßbereich beträgt 0-1000 ppm, die Meßgenauigkeit ist mit Abweichungen von unter einem ppm recht hoch. Die Anlagen werden regelmäßig unter Einsatz von Eichgasen gewartet. Insgesamt sind drei Spektroskope im Einsatz, von denen zwei die Luft in den Klimazellen messen, während das dritte die Außenluft mißt. Die Meßwerte des letztgenannten finden in dieser Arbeit Verwendung.

Abbildung 2.1 zeigt die Monatsmittelwerte der CO_2-Konzentration in der Luft für den betrachteten Zeitraum, Abbildung 2.2 ein Histogramm der validen CO_2-Werte für 1992-1996.

---

[7] Vgl. hierzu die Arbeiten von [For 93] 1 ff. und [Til 93] 5.

[8] Eine ausführliche Darstellung findet sich bei [For 93] 7 ff.

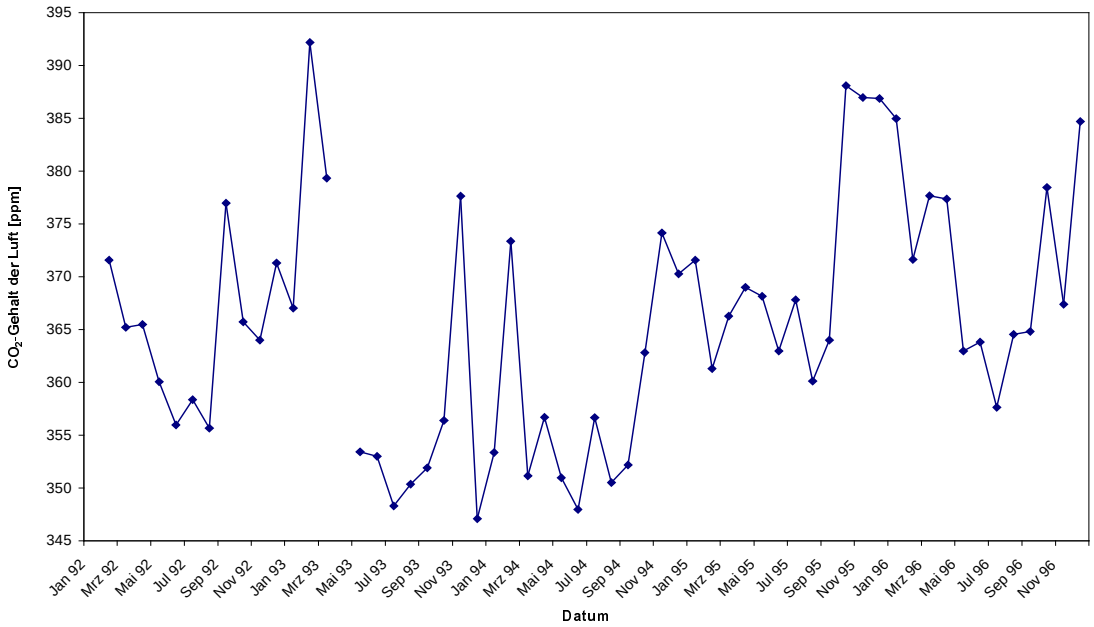

Abbildung 2.1: Monatlich gemittelte CO_2-Werte in Osnabrück (Westerberg)

Der Übersicht über die Monatsmittelwerte kann man entnehmen, daß die CO_2-Konzentrationen während der Phasen stärkerer Vegetationsaktivität (Sommer) geringer ist als während der Heizperiode im Winter, wenn größere CO_2-Emissionen und mangelnde Vegetationsaktivität zusammen kommen. Einzig die Werte für Dezember 1993 und Januar 1994 fallen aus diesem Schema.

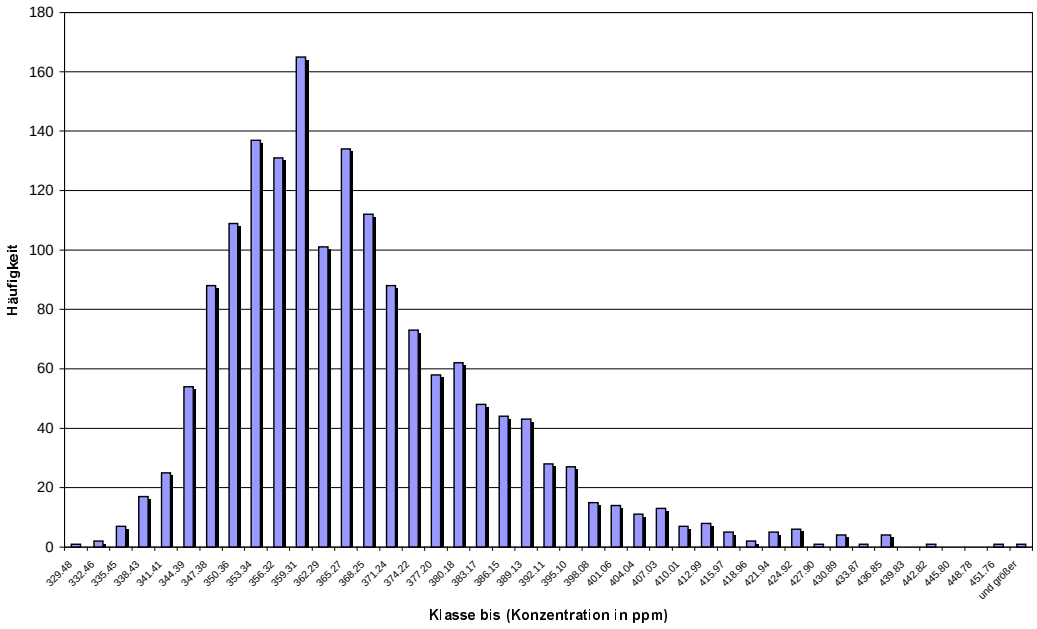

Abbildung 2.2: Histogramm der CO_2-Werte für den betrachteten Zeitraum

Das Histogramm zeigt deutlich eine schräge Verteilung. Die Daten selbst weisen zum Teil erhebliche Lücken auf, insbesondere die Daten von 1992 und 1993. Eine Übersicht

über die Fehlwerte der CO_2-Zeitreihe liefert Tabelle 2.1. Es wird deutlich, daß für die Zeitreihenanalyse nur der Zeitraum vom 06.02.1992 bis zum 31.12.1996 in Frage kommt, da für die ersten 36 Tage im Jahr 1992 keine Meßwerte vorliegen. Problematisch ist auch eine Betrachtung des Aprils 1993, für den gar keine Werte vorhanden sind, sowie für September 1992 mit nur 5 validen Werten.

Tabelle 2.1: Übersicht über fehlende Daten in der CO_2-Zeitreihe

Invalide Zeiträume (Tage)											
Von	01.01.92	27.08.92	29.12.92	16.03.93	24.03.93	10.06.93	13.11.93	20.05.94	12.03.96	16.07.96	24.12.96
Bis	05.02.92	25.09.92	31.12.92	18.03.93	06.05.93	13.06.93	14.11.93	23.05.94	15.03.96	13.08.96	27.12.96
Anzahl	36	30	3	2	44	4	2	4	5	29	4

Invalide Einzelwerte											
Tag	08.06.92	26.06.92	28.07.92	15.10.92	10.12.92	30.12.93	18.11.94	11.12.95	26.03.96	04.09.96	01.12.96

Auf die Probleme, die sich aus den unvollständigen Daten insbesondere der CO_2-Zeitreihe ergeben, wird in Kapitel 3 im Rahmen der Darstellung von Verfahren zur Zeitreihenprognose und –analyse genauer eingegangen.

2.2 Daten des Deutschen Wetterdienstes

2.2.1 Meßstation Kalkhügel

Seit 1953 betreibt der Deutsche Wetterdienst eine meteorologische Meßstation in Osnabrück, im Stadtteil Kalkhügel (52° 15‘ 26“ N, 8° 3‘ 16“ O), deren Messung für diese Arbeit verwendet wurden. Sie liegt 95m über NN. Ihre geographische Lage zeichnet sich durch eine besondere Nähe (ca. 3,8 km Luftlinie) zur CO_2-Meßstation aus. Mit dem Jahr 1992 begann die elektronische Aufzeichnung der in dieser Arbeit verwendeten Wetterdaten. Dies verspricht eine größere Genauigkeit insbesondere beim zeitlichen Abstand der Einzelmessungen.

Unsicherheiten bzw. Meßungenauigkeiten sind vor allem bei dem Bedeckungsgrad gegeben, der ein reiner Schätzwert ist. Weiterhin ist der offizielle (elektronisch gemessene) Niederschlag etwas höher als der tatsächliche Wert, und Lufttemperatur, Windgeschwindigkeit und Windrichtung können minimal gegenüber dem Westerberg variieren. Diese Abweichungen werden bei der Temperatur durch unterschiedliche Sonneneinstrahlung verursacht, bei den Windmessungen vor allem durch lokal unterschiedliche topographische Beschaffenheit.

2.2.2 Zeitraum 1992-1993

Für den Zeitraum von 1992-1993 stellte der Deutsche Wetterdienst folgende Daten als Stundenmittelwerte zur Verfügung:

- Den Bedeckungsgrad (in Achteln: 0 = wolkenlos, 8 = vollständig bewölkt),
- die Windrichtung (in 10°-Segmenten: 0 = windstill, 9 = Ost, 18 = Süd, 27 = West, 36 = Nord),
- die Windgeschwindigkeit (18 m über dem Boden, in $^{1}/_{10}\ ^{m}/_{s}$),
- die Lufttemperatur (2 m über dem Boden, in $^{1}/_{10}$ °C).

Weiterhin lagen folgende Daten als Tageswerte vor:

- Mittlere Tagestemperatur (2 m über dem Boden, in $^{1}/_{10}$ °C),
- Niederschlagshöhe (in $^{1}/_{10}$ mm).

Die Stundenwerte wurden zu Tageswerten aggregiert (siehe Abschnitt 2.5).

2.2.3 Zeitraum 1994-1996

Die Wetterdaten für den Zeitraum von 1994-1996 wurden komplett als Stundenmittelwerte zur Verfügung gestellt:

- Bedeckungsgrad (in Achteln, s.o.),
- Windrichtung (in 10°-Segmenten, s.o.),
- Windgeschwindigkeit (18 m über dem Boden, in $^{1}/_{10}\ ^{m}/_{s}$),
- Lufttemperatur (2 m über dem Boden, in $^{1}/_{10}$ °C),
- Niederschlagshöhe (in $^{1}/_{10}$ mm).

Die Einheiten entsprechen denen des Zeitraums von 1992 bis 1993. Fehlwerte werden generell durch Werte ausgedrückt, die nicht im möglichen Wertebereich liegen (-99 für Windgeschwindigkeit, -999 für Temperatur etc.).

Die Fehlwerte in den Daten beschränken sich auf den Zeitraum von 1992-1993, die der Arbeit von Tilmanns [Til 93] entstammen. Besonders fallen hier folgende Zeiträume auf: Der Zeitraum vom 01.01.1992 bis 05.02.1992 mit 36 Fehlwerten und die 44 Tage ab dem 24.03.1993, welche beide auch in den CO_2-Daten fehlen. Außerdem nicht valide sind die 65 (!) Tage ab dem 27.07.1992 und der Zeitraum vom 25.11.1992 bis zum Ende des Jahres 1992 (37 Werte).

2.3 Daten der Fachhochschule Osnabrück

Der Fachbereich Gartenbau der Fachhochschule Osnabrück betreibt seit 1954 eine agrar-meterologische Meßstation (68m über NN). Dort werden täglich folgende Wetterdaten aufgezeichnet[9]:

- Die Summe des Niederschlags (im mm),
- die Sonnenscheindauer (in Stunden),
- die Temperatur um 7^{30}, 14^{30} und 21^{30} Uhr in 2 m Höhe über dem Boden sowie deren Mittelwert (in °C),
- die relative Luftfeuchte um 7^{30}, 14^{30} und 21^{30} Uhr und deren Mittelwert (in Prozent)
- und die tägliche potentielle Verdunstung E_{pot} in mm pro Tag, die nach der Formel von Haude berechnet wird:

  $$E_{pot} = f_H \cdot (e_s - e_a).$$

  Dabei bezeichnet e_s den Sättigungsdampfdruck [mbar] und e_a den aktuellen Dampfdruck [mbar], für den gilt $e_a = e_s \cdot H_R / 100$, wobei H_R die relative Luftfeuchtigkeit [%] ist, gemessen um 14^{30} Uhr. f_H ist der sogenannte Haude-Proportionalitätsfaktor [$^{mm}/_{mbar \cdot d}$], der mit der mittleren Monatstemperatur korreliert ist und tabelliert in der Fachliteratur zu finden ist. Sein Wertebereich beträgt für Deutschland 0,26–0,39. Es ergibt sich also folgende Formel:

  $$E_{pot} = f_H \cdot e_s \cdot \left(\frac{100 - H_R}{100} \right)$$

  Diese Zeitreihe wird mit HAUDE bezeichnet.
- Außerdem steht das Sättigungsdefizit SDEF [torr] zur Verfügung. Dieses bezeichnet die Differenz zwischen dem Sättigungsdampfdruck e_s und dem tatsächlich vorhandenen Dampfdruck der Luft e_a und fand schon in der Berechnung der potentiellen Verdunstung (HAUDE) Verwendung:

  $$SDEF = e_s - e_a$$

Die Zeitreihen Niederschlag und Temperatur liegen also sowohl vom Deutschen Wetterdienst als auch von der Fachhochschule vor. Sie werden im folgenden als NS-DWD und NS-FH bzw. TEMP-DWD und TEMP-FH bezeichnet.

[9] Es werden nur die für diese Arbeit relevanten Meßdaten genannt.

2.4 Weitere Daten

Weiterhin wurden zur Modellierung der Zeitabhängigkeit bzw. der Jahresperiode folgende Datenreihen generiert:

- der Tag seit dem 01.01.1992 (Beginn der Datenreihen), von 1 bis 1827 (TG),
- der Tag im Jahr, von 1 bis 365 bzw. 366 (Schaltjahre) (JTG),
- eine Sinus- und Cosinus-Transformation des Tags im Jahr nach folgenden Formeln:

$$JTG_SIN = \sin\left(\frac{360 \cdot (Tag\ im\ Jahr)}{365\ bzw.\ 366}\right)$$

$$JTG_COS = \cos\left(\frac{360 \cdot (Tag\ im\ Jahr)}{365\ bzw.\ 366}\right)$$

  Diese Formeln bilden die unstetige JTG-Zeitreihe stetig auf das Intervall [-1..+1] ab.

Aufgrund der artifiziellen Natur dieser Datenreihen gibt es hier keine Fehlwerte.

2.5 Datenvorverarbeitung

2.5.1 Allgemeines

Für die Verwendung in der Analyse und Prognose wurden die Windgeschwindigkeit auf [$^{m}/_{s}$], die Lufttemperatur auf [°C] und die Niederschlagshöhe auf [mm] skaliert. Außerdem konnten alle Datenreihen, von denen keine Tageswerte vorlagen (die CO_2-Meßwerte sowie einige Daten des DWD), auf Tageswerte aggregiert werden. Zu diesem Zweck wurden für die meisten Zeitreihen die Mittelwerte der Stundenwerte gebildet, die pro Tag vorlagen. Ausnahmen bildeten die Niederschlagshöhe, die über den Tag aufsummiert wurde, und die Windrichtung, die aufgrund ihrer unstetigen Notation eine gesonderte Behandlung benötigte[10].

Die Transformation bzw. Aggregation der Daten im Rahmen der Datenvorverarbeitung (auch der unter 2.5.2 beschriebenen Windrichtungs-Zeitreihe) wurde durch verschiedene Programme realisiert, die in der interpretierten Programmiersprache Perl 5.0[11] geschrieben wurden. Sie finden sich auf der CD-ROM (siehe Anhang 7.3).

[10] Siehe hierzu Kapitel 2.5.2

[11] Siehe [Hal 98], [Sch 97], [Sri 97] und [Wal 96].

Ferner ist für die Verwendung der Datenreihen im Rahmen der neuronalen Modellierung mit SNNS[12] eine Skalierung der Zeitreihen notwendig, bei der ein entsprechender Wertebereich eingestellt wird. Die Beschreibung dieser Skalierungen findet sich im Kapitel über Neuronale Netze (3.2).

2.5.2 Die Windrichtungs-Zeitreihe

Wie bereits in Kapitel 2.3 erwähnt, liegen die Meßwerte der Windrichtung in 10°-Segmenten vor, wobei folgende Zuordnung getroffen wird: Norden = 36, Osten = 9, Süden = 18, Westen = 27. Für Windstille wird Null als Meßwert notiert. Mit dieser Darstellung ergibt sich eine Unstetigkeitsstelle bei Windrichtungen aus Norden. Würde man die WR-Meßwerte nun einfach durch Mittelwertbildung der Stundenwerte aggregieren, würde das Ergebnis (die vorherrschende Windrichtung) verfälscht, sobald die Unstetigkeitsstelle im Bereich der gemessenen Windrichtungen liegt, d.h., wenn der Wind überwiegend aus Norden kommt (siehe Abbildung 2.3). Um dies zu vermeiden, wird zunächst das 10°-Segment bestimmt, das von den gemessenen Windrichtungen am weitesten entfernt ist und dieses als neue Unstetigkeitsstelle verwandt. Ausgehend von diesem Segment wird der Kreis transformiert, was garantiert, daß bei der anschließenden Mittelwertbildung die Unstetigkeitsstelle keine Rolle spielt. Ist die vorherrschende Windrichtung gefunden, wird diese zurücktransformiert. Die so erhaltene Zeitreihe wird mit WR bezeichnet.

Um der Windstärke zu den jeweiligen Meßzeiten Rechnung zu tragen, wurde noch eine weitere Zeitreihe erstellt, deren Tageswerte analog zu den WR-Zeitreihe aggregiert wurden. Die einzelnen Tageswerte wurden jedoch vorher mit der entsprechenden Windrichtung gewichtet, d.h. multipliziert. So haben bei der Berechnung der vorherrschenden Windrichtung eines Tages die Meßwerte ein größeres Gewicht, zu denen eine größere Windstärke gehört. Die so erzeugte Zeitreihe wird mit WRG bezeichnet. Abbildung 2.3 zeigt die validen Meßwerte vom 22.11.1992, wobei in der linken Abbildung die zu den gemessenen Windrichtungen gehörenden Windstärken aufgetragen sind. Im rechten Teil werden die Ergebnisse der beiden vorgestellten Verfahren mit der einfachen Mittelwertbildung verglichen.

[12] Siehe [Zel 95].

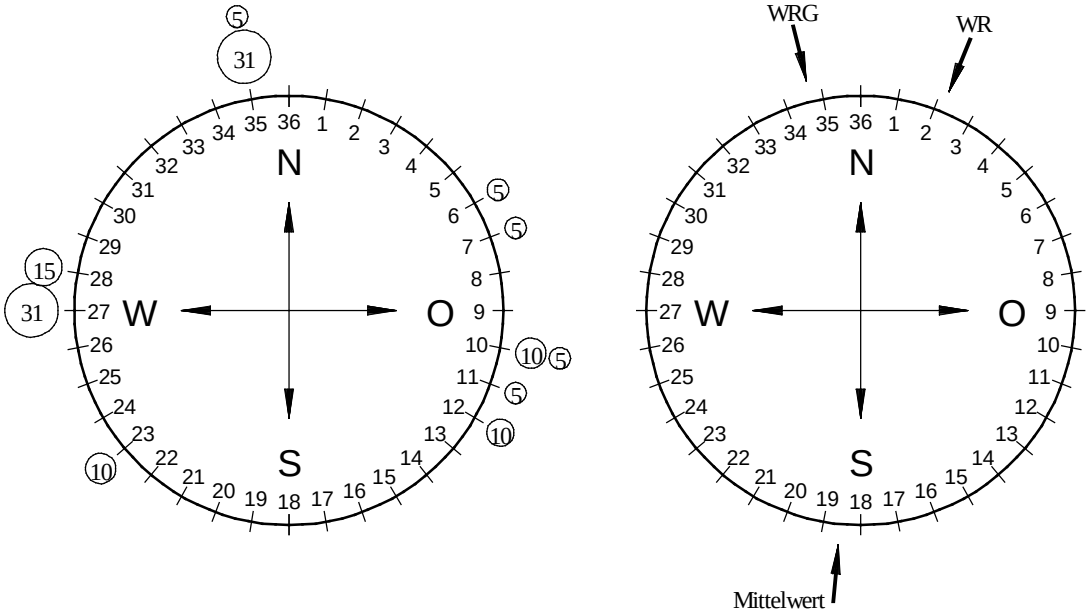

Abbildung 2.3: Windrichtungen und –stärken für einen Tag mit den daraus resultierenden Tageswerten nach verschiedenen Verfahren

Neben den Schwierigkeiten bei der Aggregation ist zu vermuten, daß die Unstetigkeit der Zeitreihe auch für die spätere Modellierung problematisch ist. Aus diesem Grund wurden die aggregierten Zeitreihen stetig auf ihre Polarkoordinaten abgebildet. Es ergeben sich jeweils zwei Zeitreihen, die den Sinus bzw. den Cosinus der WR- und WRG-Zeitreihe beinhalten. Diese werden im folgenden mit WRSIN und WRCOS bzw. WRGSIN und WRGCOS bezeichnet.

2.6 Übersicht

Tabelle 2.2 bietet eine Übersicht über die verwendeten Datenreihen, deren Herkunft, Wertebereich und Anzahl valider Werte:

Tabelle 2.2: Übersicht über die verwendeten Datenreihen

Zeitreihe	Kürzel	Quelle	Einheit	Gültig	Ungültig	Minimum	Mittelwert	Maximum
Bedeckungsgrad	BG	DWD	[%]	1576	251	0.738	70.298	100.000
CO_2-Gehalt der Luft	CO2	BIO	[ppm]	1653	174	329.481	365.181	454.747
Potentielle Verdunstung	HAUDE	FH	[mm]	1827	0	0.000	1.144	11.600
Tag im Jahr	JTAG	ART	[-]	1827	0	1.000	183.200	366.000
Cosinus des Jahrestags	JTG_COS	ART	[-]	1827	0	-1.000	0.003	1.000
Sinus des Jahrestags	JTG_SIN	ART	[-]	1827	0	-1.000	0.000	1.000
Niederschlag (DWD)	NS_DWD	DWD	[mm]	1827	0	0.000	2.458	56.200
Niederschlag (FH)	NS_FH	FH	[mm]	1827	0	0.000	2.334	38.200
Relative Luftfeuchtigkeit	RLF	FH	[%]	1827	0	34.330	76.257	100.000
Sättigungsdefizit	SDEF	FH	[torr]	1827	0	0.000	3.363	33.150
Sonnenscheindauer	SSD	FH	[h/d]	1827	0	0.000	4.032	14.700
Tag seit dem 1.1.92	TAG	ART	[-]	1827	0	1.000	914.000	1827.000
Lufttemperatur (DWD)	TMP_DWD	DWD	[°C]	1601	226	-14.713	9.468	27.609
Lufttemperatur (FH)	TMP_FH	FH	[°C]	1827	0	-13.300	9.330	27.800
Tageshöchsttemperatur (DWD)	THT_DWD	DWD	[°C]	1601	226	-12.100	12.889	34.200
Tageshöchsttemperatur (FH)	THT_FH	FH	[°C]	1827	0	-13.000	13.306	35.000
Windgeschwindigkeit	WG	DWD	[m/s]	1601	226	0.692	3.394	11.185
Windrichtung	WR	DWD	[°]	1601	226	0.870	191.217	359.583
Cosinus der Windrichtung	WRCOS	DWD	[-]	1601	226	-1.000	-0.180	1.000
Sinus der Windrichtung	WRSIN	DWD	[-]	1601	226	-1.000	-0.140	1.000
Gewichtete Windrichtung	WRG	DWD	[°]	1601	226	1.043	193.161	360.000
Cosinus der gew. Windrichtung	WRGCOS	DWD	[-]	1601	226	-1.000	-0.169	1.000
Sinus der gew. Windrichtung	WRGSIN	DWD	[-]	1601	226	-1.000	-0.160	1.000

3 Methoden der Datenanalyse und -prognose

3.1 Mathematische Verfahren

3.1.1 Multiple Regressionsanalyse

Die multiple Regressionsanalyse ist ein einfaches und weit verbreitetes Verfahren, um die Abhängigkeiten von exogenen und endogenen Variablen zu modellieren. Dabei ist man nicht zwangsläufig auf die lineare Regression beschränkt, in der Praxis spielt diese jedoch die wichtigste Rolle. Im folgenden wird das Verfahren kurz dargestellt[13].

3.1.1.1 Funktionsweise

Will man den Zusammenhang einer abhängigen (endogenen) Variable y von k unabhängigen (exogenen) Variablen x_i (i = 1, 2, ..., k) modellieren, nimmt man an, y lasse sich als gewichtete Summe der einzelnen exogenen Variablen ausdrücken. Für diese x_i existieren N Realisierungen x_{ti} (t = 1, 2, ..., N). Im linearen Fall erhält man so:

$$y_t = a_0 + a_1 \cdot x_{t1} + a_2 \cdot x_{t2} + \ldots + a_k \cdot x_{tk}$$

Man beachte, daß die x_{ti} auch zeitlich geshiftete exogene Zeitreihen darstellen können. Die Gewichte a_i (i = 0, 1, ..., k) werden in aller Regel nach der Methode der kleinsten Abweichungsquadrate bestimmt. Mit dieser Methode lassen sich die Gewichte mathematisch exakt bestimmen. Für die praktische Anwendung wurden SAS[14] (insbesondere die Statistik-Prozedur „Reg“) sowie GAUSS[15] und MS Excel 97[16] verwandt, die jedoch alle auf numerische Verfahren zurückgreifen.

Die Koeffizienten der Regressionsgleichung können allerdings nur dann zuverlässig bestimmt werden, wenn die betrachtete Zeitreihe *schwach stationär* ist. Man spricht von der schwachen Stationarität einer Zeitreihe, wenn die Mittelwerte und Varianzen zweier ausreichend großer Teilmengen $y_t \ldots y_T$ und $y_{t+k} \ldots y_{T+k}$ für alle k gleich, also zeitlich invariant sind. Ist die schwache Stationarität nicht gegeben, so wird der *Differenzenfilter 1. Ordnung* der Form

$$\Delta y_t = y_t - y_{t-1}, \quad t = 2, 3, \ldots, N$$

so oft angewandt, bis die Stationarität vorliegt.

---

[13] Hier wird die multiple lineare Regression vorgestellt. Die Erstellung nichtlinearer Regressionsmodelle erfolgt entsprechend.

[14] Siehe [Duf 92], [Gra 95], [Pfe 97] und [Ver 97].

[15] Siehe [Apt 96a], [Apt 96b], [Apt 96c] und [Apt 96d].

[16] Siehe [Bal 93], [Erb 97] und [Win 97].

3.1.1.2 Probleme

Bei der (multiplen) Regressionsanalyse treten verschiedene Probleme auf, die bei der Bewertung der Ergebnisse beachtet werden müssen:

- Um die Regressionsanalyse optimal einzusetzen, muß bereits im Vorfeld klar sein, welcher Art die Zusammenhänge zwischen exogenen und endogenen Variablen sind. Dabei ist Linearität oftmals nicht gegeben und lineare Verfahren eignen sich in diesen Fällen dementsprechend schlecht zur Modellierung.
- Ausreißer werden bei Bestimmung der Gewichte a_i zu stark bewertet, da hier die quadratische Abweichung minimiert wird. So haben 100 Werte mit Abweichungen d den gleichen Einfluß wie ein Wert mit Abweichung $10 \cdot d$. Gegebenenfalls ist hier also eine Vorverarbeitung angebracht.

3.1.2 Klassisches Komponentenmodell

Das klassische Komponentenmodell ist als eine Erweiterung der Regressionsanalyse zu verstehen, die diese um weitere, i.a. nichtlineare Terme erweitert.

3.1.2.1 Funktionsweise

Man nimmt an, daß eine Zeitreihe aus verschiedenen Komponenten zusammengesetzt ist, die den tatsächlichen Verlauf der Zeitreihe ergeben. Das klassische Komponentenmodell unterscheidet drei Komponenten[17]:

1) Der *Trend*, der eine langfristige systematische zeitliche Entwicklung des mittleren Niveaus darstellt. Er wird meist durch eine Gerade dargestellt.
2) Die *zyklische Komponente*, die eine (jahres-) zeitlich relativ unverändert auftretende Schwankung der Werte modelliert. Eine typische zyklische Schwankung ist die Saisonkomponente, die in manchen Darstellungen des Komponentenmodells als gesonderte Komponente behandelt wird[18].
3) Die *Zufalls-* oder *Restkomponente*, die nicht erklärbare Abweichungen beinhaltet. Da die beiden anderen Komponenten durch mathematische Funktionen dargestellt werden und die betrachtete Zeitreihe als Realisierung eines statistischen Prozesses noch zufallsgestört ist, muß angenommen werden, daß die Kombination aus Trend und zyklischer Komponente die Zeitreihe nicht vollständig erklären.

Es lassen sich nun aus diesen Annahmen unterschiedliche Modelle erstellen, die sich in der Art der Verknüpfung der Komponenten unterscheiden. Die zwei bedeutendsten sind das *additive Modell* und das *multiplikative Modell*. Meist wird angenommen, daß sich

---

[17] Es gibt auch die Darstellung mit vier Komponenten, bei der die zyklische Komponente noch in Konjunktur- und Saisonkomponente unterschieden wird. Vgl. hierzu [Sch 91] 9.

[18] Vgl. [Sch 91] 9.

die Komponenten additiv überlagern. So erhält man eine formale Beschreibung des additiven Modells:

$$x_t = \underbrace{m_t}_{Trend} + \underbrace{z_t}_{zyklische\ Komponente} + \underbrace{r_t}_{Rest}$$

Vielfach tritt auch der Fall ein, daß mit dem mittleren Niveau auch die Streuung um dieses Niveau zunimmt. In diesem Fall findet das multiplikative Modell Verwendung:

$$x_t = m_t \cdot z_t \cdot r_t$$

Im folgenden sollen nun gängige Verfahren vorgestellt werden, diese Komponenten zu bestimmen.

3.1.2.2 Trendkomponente

Die Trendkomponente modelliert die allgemeine zeitliche Entwicklung der Zeitreihe. Dabei werden drei Unterscheidungen vorgenommen: Entweder ist der Trend steigend (zunehmend), fallend (abnehmend) oder horizontal (gleichbleibend). Es existieren verschiedene Verfahren zur Bestimmung der Trendkomponente. Da in der Praxis vielfach von einem linearen Trend ausgegangen wird, kann hier die lineare Regression Anwendung finden, die bereits in Kapitel 3.1.1 vorgestellt wurde. Der lineare Trend ist eine Funktion der Zeit *t* nach der allgemeinen Geradengleichung

$$y = a + b \cdot t\ ,$$

wobei der Steigungskoeffizient aussagt, ob eine steigende, fallende oder horizontale Entwicklung vorliegt, je nachdem, ob er positiv, negativ oder Null ist.

Wird der Zeitreihe ein nichtlinearer Trend unterstellt, werden die entsprechenden Parameter ebenfalls mit der Methode der kleinsten Quadrate bestimmt. Fast die gesamte aktuelle Statistik-Software verfügt über numerische Verfahren zur Koeffizientenbestimmung nach dieser Methode[19].

3.1.2.3 Zyklische Komponente

Für die Bestimmung der zyklischen Komponente werden entsprechend der zeitlichen Auflösung der vorhandenen Werte (hier: Tage) über mehrere Jahre hinweg Durchschnittswerte gebildet, durch die mittels der Methode der kleinsten Quadrate eine Sinusschwingung gelegt wird. Dabei wird als Periode ein Jahr angenommen, wie es für saisonale Schwankungen üblich ist[20].

---

[19] Dies gilt insbesondere für die im Rahmen dieser Arbeit eingesetzte Statistiksoftware SAS, GAUSS und MS Excel 97.

[20] Vgl. [Bil 81], 52 ff.

3.1.2.4 Restkomponente

Die Restkomponente wird nicht wirklich modelliert. Wenn man Aussagen über sie machen könnte, so würden diese Gesetzmäßigkeiten mit in die Trend- oder zyklische Komponente aufgenommen. Wie bereits oben beschrieben, umfaßt die Restkomponente die Abweichungen der Wirklichkeit vom Modell, die bei der Realisierung eines statistischen Prozesses nicht ausbleiben. Sie entspricht damit also dem Fehler des Modells, der zu minimieren versucht wird.

Im Rahmen dieser Arbeit stehen jedoch exogene Zeitreihen zur Verfügung, so daß das univariate klassische Komponentenmodell hier in ein multivariates Verfahren erweitert werden kann. Dazu wird davon ausgegangen, daß sich der Fehler des Modells aus einem erklärbaren und einem unerklärbaren Teil zusammensetzt. Der erklärbare Teil wird als gewichtete Summe exogener Variablen ausgedrückt, wie es auch bei der Regression der Fall ist.

$$r_t = a_1 \cdot x_{t1} + a_2 \cdot x_{t2} + \ldots + a_k \cdot x_{tk} + (\text{unerklärbarer Rest})$$

3.1.2.5 Probleme

Auch bei dem klassischen Komponenten-Ansatz treten verschiedene Probleme auf, die bei der Bewertung der Ergebnisse betrachtet werden müssen:

- Durch die Festlegung der Modellstruktur (additives oder multiplikatives Modell) werden Annahmen über das System getroffen, die i.a. höchstens annähernd der Wirklichkeit entsprechen. Auch wenn dieses Modell mehr Freiheiten bietet als die Regressionsanalyse allein, ist eigentlich ein genaues Wissen um die Systemstruktur nötig, um das Modell optimal anzuwenden.

3.1.3 AR-Modellierung und ARIMA

Im Rahmen der Regressionsanalyse (Kapitel 3.1.1) wurde bereits der Ansatz der multiplen linearen Regression erläutert. Ganz ähnlich ist das Vorgehen bei einem autoregressiven Prozeß (AR-Prozeß), nur das hier statt exogener Indikatoren die gewichtete Summe zeitlich zurückliegender Realisierungen der zu prognostizierenden Variablen verwendet wird. Bei einem „autoregressive-integrated-moving-average“-Prozeß (ARIMA Prozeß, nach Box-Jenkins[21]) handelt es sich um einem stochastischen Prozeß, der autoregressive (AR) und moving-average Prozesse (MA) umfaßt[22]. Als Kurzform verwendet man die Schreibweise ARIMA[p,d,q]-Prozeß. Das ARIMA-Modell vereinigt also folgende Prozesse:

- Einen autoregressiven Prozeß der Ordnung p (AR[p]-Prozeß),

[21] Vgl. [Box 76].

[22] Vgl. [Rip 93], 53 ff.

- einen Integrationsprozeß der Ordnung d (I[d]-Prozeß), der nur dann nötig ist, wenn die Zeitreihe zur Erreichung der Stationarität vorher differenziert wurde und
- einen moving-average Prozeß der Ordung q (MA[q]-Prozeß).

3.1.3.1 Funktionsweise

Bei einem **AR[p]-Prozeß** wird davon ausgegangen, daß sich die Werte der Zeitreihe als gewichtete Summe der zeitlich zurückliegenden Realisierungen bestimmen lassen. Die Ordung $p \in \{0, 1, 2, ...\}$ des Prozesses bezeichnet dabei den größten auftretenden „lag", also die zeitliche Differenz zwischen dem zu prognostizierenden Wert und dem am weitesten zurückliegenden zur Verfügung stehenden Wert. Die allgemeine Form eines AR[p]-Prozesses lautet

$$y_t = a_0 + a_1 \cdot x_{t-1} + a_2 \cdot x_{t-2} + \ldots + a_p \cdot x_{t-p} + \varepsilon_t$$

Dabei stellt ε_t einen Fehlerterm dar, der die Abweichung vom wahren Wert beschreibt. Für die Summe der Koeffizienten a_i (i= 0, 1, ..., p) gilt, daß sie kleiner als Eins sein muß:

$$\sum_{i=0}^{p} a_i < 1$$

Der **I[d]-Prozeß** gleicht ggf. einen Differenzenfilter d-ter Ordnung

$$\Delta^d x_t = \Delta^{d-1} x_t - \Delta^{d-1} x_{t-1}, \quad t = d+1, ..., N$$

wieder aus.

Ein MA[q]-Prozeß ist ein stochastischer Prozeß, der sich in folgender Form schreiben läßt:

$$y_t = \varepsilon_t - b_1 \cdot \varepsilon_{t-1} - ... - b_q \cdot \varepsilon_{t-q}$$

Dabei ist ε_t ein White-Noise-Prozeß[23]. Diese Gleichung besagt, daß bei der Prognose die Fehler mit einem bestimmten Gewicht eingehen. Für die Gewichte b_i der Gleichung gilt dieselbe Restriktion wie für die a_i des AR-Prozesses:

$$\sum_{i=0}^{p} b_i < 1$$

Die Erstellung eines ARIMA[p,d,q]-Modells für eine bestimmte Zeitreihe gliedert sich in vier Phasen:

[23] Eine detailliertere Darstellung mit Herleitung findet sich bei [Sch 91], 71 ff. und 95 ff.

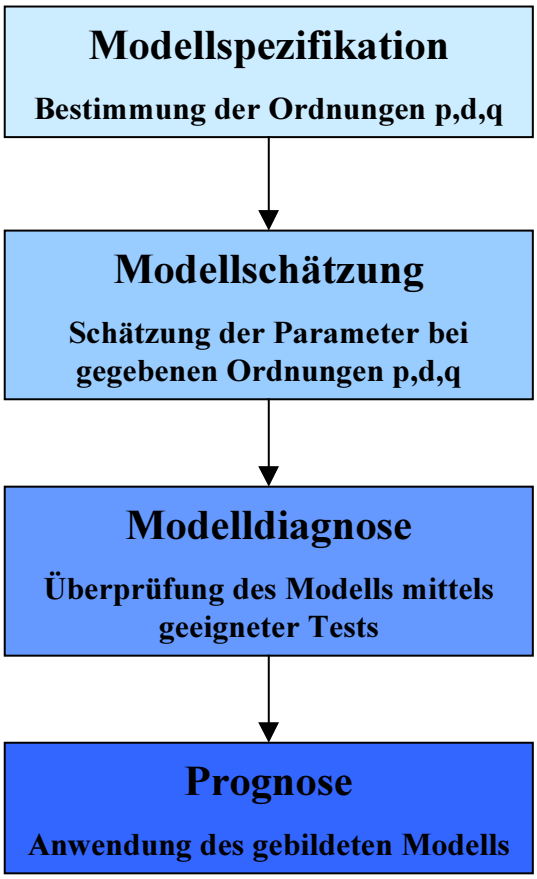

Abbildung 3.1: Die vier Phasen der Anpassung eines ARIMA[p,d,q]-Modells an eine gegebene Zeitreihe

Die ARIMA[p,d,q]-Modellierung wurde mit GAUSS durchgeführt. Alle vier Schritte in der Modellbildung werden von der Statistiksoftware unterstützt. Verfahren zur Bestimmung der Ordnungen p,d,q über (partielle) Autokorrelationskoeffizienten und zur Findung der freien Parameter finden sich in der Fachliteratur[24].

3.1.3.2 Probleme

ARMA- und ARIMA-Prozesse setzen aufgrund ihres autoregressiven Anteils (AR) voraus, daß die Zeitreihe im betrachteten Zeitraum vollständig ist, d.h. keine Fehlwerte aufweist. Für naturwissenschaftliche Meßreihen ist dies im allgemeinen nicht gegeben, wie auch hier für die CO_2-Meßreihe. Aus diesem Grund bieten sich zwei Alternativen an: Entweder werden die fehlenden Werte durch ein geeignetes Modell (mathematisch oder neuronal) interpoliert, oder man beschränkt sich bei der Betrachtung auf Abschnitte der CO_2-Zeitreihe, die jeweils vollständig sind. Hierbei bietet sich die Zeiträume vom 19.11.94 bis zum 10.12.95 und vom 24.05.94 bis zum 17.11.94 mit 387 bzw. 178 Meßwerten an. Die Anwendung und der Modellvergleich mit den Neuronalen Netzen findet sich in Kapitel XXX (ARIMA).

[24] Zum Beispiel bei [Sch 91], [Bil 81].

3.2 Neuronale Netze

3.2.1 Einführung

Die Entwicklung Neuronaler Netze (neural networks, NN) begann 1943. Ausgelöst wurde sie durch Warren McCulloch und Walter Pitts im Rahmen ihres Aufsatzes *A logical calculus of the ideas immanent in nervous activity*. In den frühen Anfängen (1943-1955) stand die Nachbildung realer neuronaler Systeme wie dem Gehirn von Säugetieren mit künstlichen Modellen im Vordergrund. Man spricht daher auch von künstlichen Neuronalen Netzen (artificial neural networks, ANN)[25].

Erst 1986 erlangten die Neuronalen Netze durch die Entwicklung und weite Publikation des Backpropagation-Verfahrens den großen Einfluß, den sie heute haben[26]. Seitdem werden sie in vielen wissenschaftlichen Bereichen vornehmlich zur Prognose, Analyse und Bilderkennung mit Erfolg eingesetzt. Neuronale Netze bieten hier den großen Vorteil, daß sie Algorithmen und Regeln „lernen“ können, ohne daß diese vorher explizit bekannt sein müssen.

Die Anwendung Neuronaler Netze gliedert sich in zwei Phasen: Die erste ist die *Lern-* oder *Trainingsphase*, in der das Netz Zusammenhänge "lernt", die zweite ist die Recall- oder Applikationsphase, in der das Gelernte genutzt wird, um beispielsweise Prognosen durchzuführen.

Während der Trainingsphase werden dem Netz Trainingsmuster (Input- und Output-Pattern) präsentiert. Durch ein Lernverfahren werden die Variablen des Netzes (*Gewichte*, siehe unten) so angepaßt, daß der Output des Neuronalen Netzes sich den Output-Pattern möglichst gut annähert. Auf diese Weise werden Zusammenhänge zwischen Input- und Output-Pattern extrahiert, die eine Anwendung des Netzes auch auf solche Muster ermöglichen, die vorher nicht trainiert wurden.

In der Recallphase greift man auf die gelernten Zusammenhänge zurück und präsentiert dem Netz die konkrete Problemstellung, nämlich Input-Pattern, zu denen der Output nicht bekannt ist. Das Netz sollte nun in der Lage sein, den entsprechenden Output zu generieren.

3.2.1.1 Aufbau Neuronaler Netze

Neuronale Netze bestehen aus vielen miteinander vernetzten sog. *Neuronen* (Abbildung 3.2). Ein Neuron *j* besteht aus gewichteten eingehenden Informationskanälen, einer Funktion net_j zur Überlagerung der eingehenden Informationen, einer sog. *Aktivierungsfunktion* $a_j(net)$, die aus der überlagerten Information einen Output generiert und

[25] Da sich diese Arbeit ausschließlich mit künstlichen Neuronalen Netzen beschäftigt, werden diese im folgenden einfach als Neuronale Netze bezeichnet.

[26] Eine ausführliche Darstellung der Geschichte Neuronaler Netze findet sich bei [Zel 96] 28 ff.

einer Outputfunktion $o_j(a)$. Die Elemente sind in Abbildung 3.2 dargestellt[27]. Im folgenden werden die einzelnen Schritte der Arbeitsweise eines Neurons erläutert:

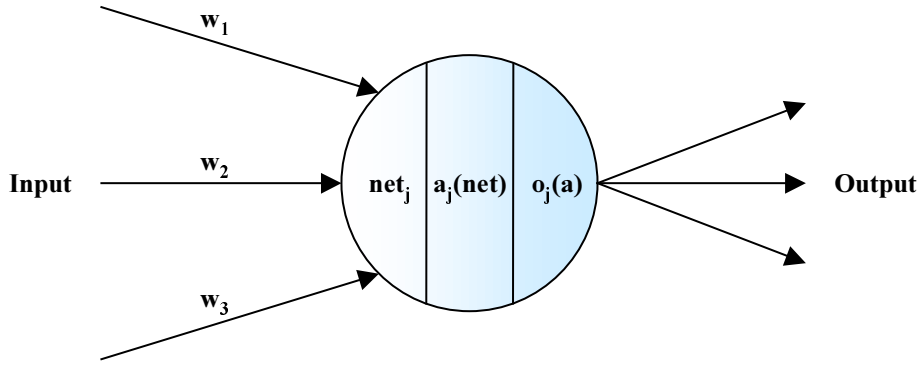

Abbildung 3.2: Aufbau eines künstlichen Neurons

1. Schritt: Zunächst ermittelt das Neuron seinen Netto-Input, der sich aus dem Output der einzelnen Neuronen der vorigen Schicht ergibt. Dazu werden die eingehenden Informationen mit den Gewichten w_i der jeweiligen Verbindungen multipliziert und durch die Eingangsfunktion *net* zu einem konkreten Wert aggregiert. Für die Modellierung der CO_2-Werte sind die eingehenden Informationen x_i im allegemeinen reelle Zahlen, es sind aber auch andere Arten von Informationen möglich, wie etwa binäre (ja / nein) Funktionen. Durch Multiplikation der x_i mit Gewichten w_i kann das Neuron eine Filterung der einzelnen Einflüsse vornehmen, zum Beispiel um unwichtige Informationen durch Nullsetzen der entsprechenden w_i auszublenden.

2. Schritt: Aus dem Nettoinput *net* wird nun der Aktivierungswert des Neurons mit Hilfe der *Aktivierungsfunktion a(net)* berechnet, die den Nettoinput auf ein festgelegtes offenes Intervall (meist]0;1[oder]-1..+1[) abbildet. Das später erläuterte Netzwerktraining fordert für Netze mit reellem Input, daß die Aktivierungsfunktion stetig differenzierbar ist. Weiterhin ist es sinnvoll, die Nichtlinearität der Funktion zu fordern, da sich andernfalls das Neuron und damit das gesamte neuronale Netz zur einfachen linearen Abbildung reduzieren.

3. Schritt: Im letzten Schritt wird aus der Aktivierung des Neurons die Ausgabe des Neurons berechnet und an die nächste Schicht weitergegeben. Meist handelt es sich bei der *Ausgabefunktion o(a)* um die Identität, wie auch bei den in dieser Arbeit verwandten Netzen. Aus diesem Grund wird sie einfach als Teil der Aktivierungsfunktion betrachtet.

3.2.1.2 Vom Neuron zum neuronalen Netzwerk

Um beliebig komplexe Zusammenhänge modellieren zu können, werden viele Neuronen zu sogenannten *Neuronalen Netzwerken* zusammengefaßt. Je nach Topologie der Vernetzung wird hier von unterschiedlichen Netz(werk)architekturen gesprochen. In

[27] Vgl. [Zim 97] 14.

dieser Arbeit wurde die sogenannte *Feed-Forward Multi-Layer-Perzeptron* (FF MLP) Architektur verwandt. Bei dieser Netztopologie sind die Neuronen in mehreren Schichten (Layers) angeordnet, bei der der Output einer Schicht der jeweils nächsten Schicht als Input dient. Der Informationsfluß ist also streng linear von Schicht zu Schicht (Feed-Forward). Die einzelnen Schichten sind dabei voll vernetzt, d.h., es existiert eine Verbindung von jedem Neuron einer Schicht zu jedem Neuron der folgenden Schicht. In Abbildung 3.3 ist dies schematisch dargestellt:

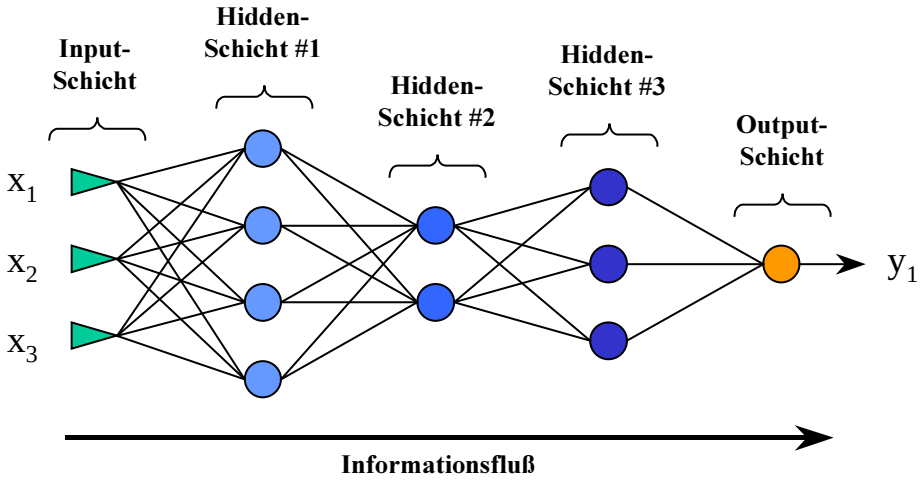

Abbildung 3.3: Struktur eines fünfschichtigen Feed-Forward-Netzes

Man unterscheidet drei Arten von Neuronen abhängig von ihrer Position im Netzwerk. Die Neuronen der ersten Schicht werden *Eingabe*- oder *Input*-Neuronen genannt. Ihnen folgen endlich viele innere Schichten, die dementsprechend *Innere*- oder *Hidden*-Schichten genannt werden. Die letzte Schicht stellt den Output des neuronalen Netzes dar und wird deshalb auch *Ausgabe*- oder *Output*-Schicht genannt.

3.2.1.3 Modifikationen der Netzarchitektur

Es existieren verschiedene Erweiterungen der eben vorgestellten Netzarchitektur, die für unterschiedliche Problemklassen eine Verbesserung der Netzgenauigkeit eröffnen. Dazu zählen vor allem die Verwendung einer *Vorschicht* und sogenannte *Shortcut-Verbindungen.*

Um die Auswirkungen von extremen Trainingsdaten während der Trainingsphase einzuschränken, wird Neuronalen Netzen eine weitere Schicht von Neuronen vorgeschaltet, die mit der ursprünglichen Input-Schicht (jetzt Hidden #1) einfach vernetzt ist, wie Abbildung 3.4 dargestellt.

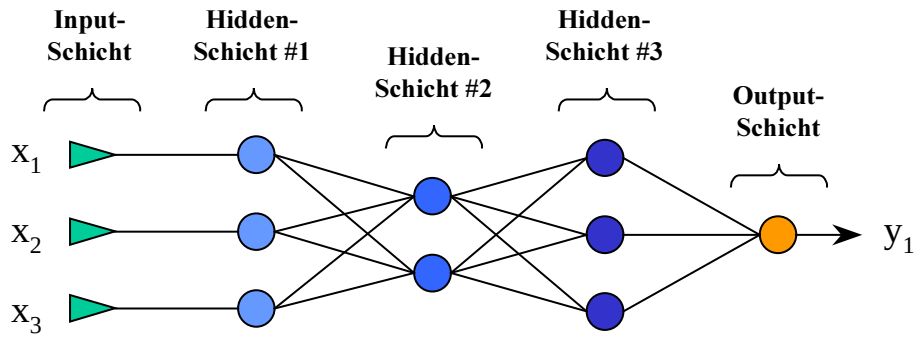

Abbildung 3.4: Feed-Forward-Netz mit Vorschicht

Dabei wird als Aktivierungsfunktion der Neuronen in der Vorschicht (Hidden #1 in Abbildung 3.4) die logistische Funktion verwandt, so daß Ausreißer auf den Sättigungsbereich der logistischen Aktivierungsfunktion abgebildet werden und der Output des Neurons seinen Charakter als Ausreißer verliert[28]. Da auch die Gewichte an den Verbindungen der Vorschicht zur ersten Hidden-Schicht beim Training verändert werden können, kann das Netz den Umgang mit Ausreißern selbst lernen.

Eine weitere Modifikation der bisher dargestellten Architektur der Feed-Forward-Netze stellen die Shortcut-Verbindungen dar. Hierbei handelt es sich um direkte Verbindungen zwischen zwei nicht benachbarten Schichten, meist der Input- und Output-Schicht. Ist in diesem Fall die Aktivierungsfunktion der Output-Schicht linear, so lassen sich mit Hilfe der Shortcut-Verbindungen lineare Einflüsse modellieren, während der Rest des Netzes (verborgene Schichten) nichtlineare Einflüsse darstellen kann. Theoretisch können Shortcut-Verbindungen zwischen beliebigen nicht benachbarten Schichten bestehen, wie in Abbildung 3.5 dargestellt.

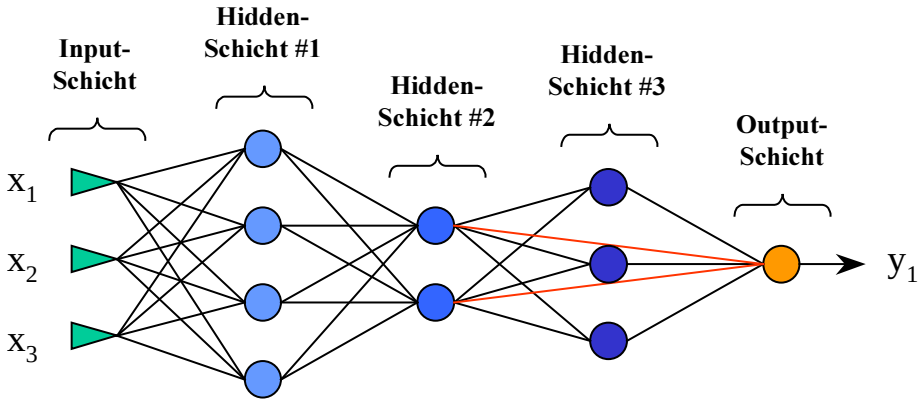

Abbildung 3.5: Feed-Forward-Netz mit Shortcut-Verbindungen zwischen Hidden-Schicht #2 und Output-Schicht

[28] Vgl. [Thi 98] 39.

3.2.2 Training neuronaler Netze

Wie schon eingangs erwähnt, gehört es zu den besonderen Stärken Neuronaler Netze, daß sie Regelmäßigkeiten erkennen können und so den Umgang mit verschiedenen Problemen *lernen*. Hier soll auf das *überwachte Lernen* eingegangen werden. Bei dem sogenannten *Netztraining* wird dabei die Gewichtsmatrix des Neuronalen Netzes in der Weise modifiziert, daß sich der Testfehler des Netzes solange verbessert, bis ein *Abbruchkriterium* erfüllt ist. Das *Lernproblem* besteht darin, die durch die *Trainingsmenge* gegebenen q n-dimensionalen Eingabevektoren x_i $\{i = 1, ..., q\}$ möglichst genau auf die ebenfalls durch die Trainingsmenge gegebenen entsprechenden q m-dimensionalen Ausgabevektoren o_i abzubilden[29]. Für das konkrete Problem der Analyse und Prognose von CO_2-Werten ist die Dimension der Ausgabevektoren $m = 1$. Als Maß für die Genauigkeit der Abbildung dient die quadratische Abweichung E_{Netz} der vorgegebenen Ausgabevektoren y_i von den durch das Netz bestimmten o_i, welche minimiert werden soll:

$$E_{Netz} = \sum_{i=1}^{q} \left(\left\| o_i - y_i \right\| \right)^2$$

Um die Minimierung des quadratischen Fehlers des Netzes zu erreichen, gibt es verschiedene Lernverfahren. Im folgenden soll nun der wohl populärste Lernalgorithmus vorgestellt werden.

3.2.2.1 Der Backpropagation-Lernalgorithmus

Es soll hier lediglich ein Überblick über den Lernalgorithmus gegeben werden. Eine genauere Darstellung mit Herleitung findet sich in der entsprechenden Fachliteratur[30]. Bereits 1974 wurde der später unter dem Namen *Backpropagation* bekannte Algorithmus in der Dissertation von Paul Werbos entwickelt, seine große Bedeutung erlangte er allerdings erst zwölf Jahre später durch die Arbeiten von Rumelhart und McCulloch (1986)[31].

Der Backpropagation-Algorithmus ist ein sogenanntes Gradientenabstiegsverfahren, d.h., es berechnet den Gradienten der Zielfunktion, hier der Fehlerfunktion, und steigt in Gegenrichtung dazu nach unten, bis ein Minimum erreicht ist. In unserem Fall wird die Minimierung des Fehlers durch Änderung aller Gewichte um einen Bruchteil des negativen Gradienten der Fehlerfunktion realisiert. Die Anwendung des Backpropagation-Verfahrens setzt ein FF-MLP und eine stetig differenzierbare Aktivierungsfunktion, die sogenannte Sigmoide $f : \Re \to]0;1[$ voraus. Für die Aktivierung a eines einzelnen Neurons i ist sie definiert durch:

---

[29] Die übereinstimmende Indizierung soll illustrieren, daß Eingabe- und Ausgabevektoren mit gleichem Index i zueinander gehören.

[30] Vgl z.B. [Zel 96], Seite 30 u. 105 ff.; sowie [Zim 97], Seite 32 ff; und [Thi 98], Seite 11 ff.

[31] Vgl. [Zel 96], Seite 30

$$a_i = f(net_i) = \frac{1}{1+e^{-net_i}}$$

Der Verlauf der sigmoiden Aktivierungsfunktion ist in Abbildung 3.6 graphisch dargestellt.

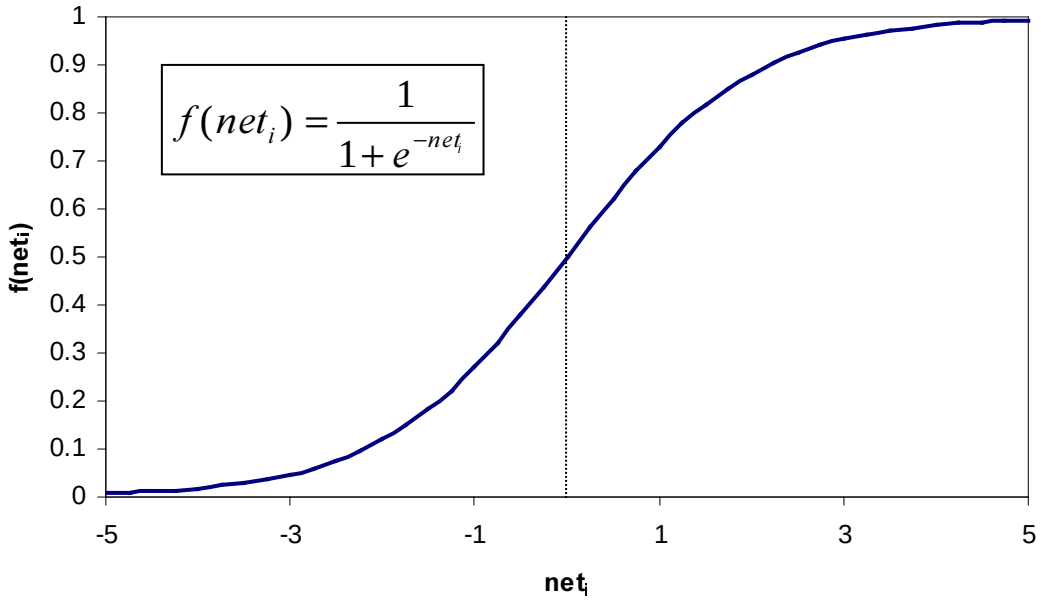

Abbildung 3.6: Sigmoide Aktivierungsfunktion

Es läßt sich leicht zeigen, daß auch die Ableitung f' auf dem gesamten Definitionsbereich definiert und stetig ist:

$$f'(net_i) = \frac{\partial f(net_i)}{\partial net_i} = \frac{e^{-net_i}}{\left(1+e^{-net_i}\right)^2} = f(net_i)\cdot(1-f(net_i))$$

Der Ablauf des Trainings läßt sich folgendermaßen beschreiben: Das Netz wird mit zufälligen Gewichten initialisiert. Ausgehend von einem durch diese Verbindungsgewichte gegebenen Punkt im Fehlergebirge erfolgt eine Modifizierung der Gewichte in der Weise, daß bei jedem Lernschritt ein tieferer Punkt auf der Oberfläche des Fehlergebirges erreicht wird. Dabei wird in Richtung der stärksten Reduzierung des quadratischen Netzfehlers vorgegangen (Gradientenabstieg), womit sich das Verfahren zur Minimumsbestimmung einer nichtlinearen *n*-dimensionalen Funktion und damit auf ein Optimierungsproblem reduziert. Auf die formale (recht umfangreiche) Darstellung der Herleitung und Lösung wird hier verzichtet (siehe oben).

Das Diagramm aus Abbildung 3.7 zeigt schematisch den Ablauf des Netztrainings, wie es im Rahmen dieser Arbeit eingesetzt wurde[32]:

[32] Vgl. [Thi 98], Seite 135

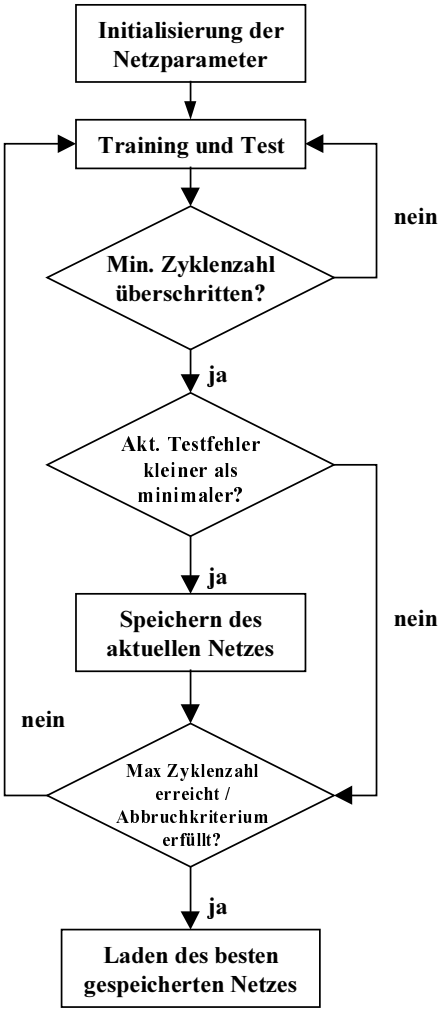

Abbildung 3.7: Schematischer Ablauf des Netztrainings

3.2.2.2 Overfitting und Generalisierung

Aufgrund der Natur des Backpropagation-Algorithmus wird sich der Fehler auf der Lernmenge asymptotisch einem Minimum nähern. Dabei steigt die Generalisierungsfähigkeit des Netzes zunächst, fällt aber später wieder ab. Um dies zu beobachten, werden die Trainingsdaten in zwei Mengen aufgeteilt: Die Test- und die Lerndaten. Nur die Lerndaten stehen für die Angleichung der Gewichte im Rahmen des Trainings zur Verfügung, während anhand der Testdaten die Generalisierungsfähigkeit des Netzes überprüft wird. Die Aufteilung zeigt Abbildung 3.8.

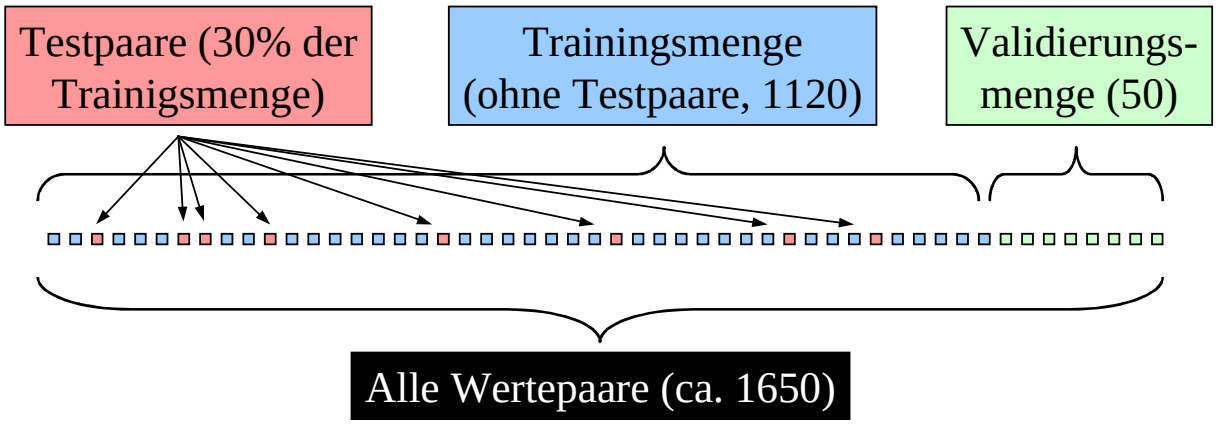

Abbildung 3.8: Aufteilung der Wertepaare in Test-, Trainings- und Validierungsmenge (Beispiel)

Nach jedem Lernschritt wird der Netzfehler auf beiden Mengen ausgewertet. Zu einem gewissen Zeitpunkt ist die optimale Generalisierungsleistung erreicht, danach wird sie schlechter. Dies kann an dem Anstieg des Fehlers für die Testmenge beobachtet werden. Ab diesem Punkt lernt das Netz die Lerndaten „auswendig“, was aber der Zielsetzung des Lernens, nämlich dem Erkennen von Wirkungszusammenhängen abträglich ist. Aus diesem Grund wird das Training abgebrochen, sobald der Fehler auf den Testdaten über einen bestimmten Prozentsatz (z.B. 20%) des bisherigen Minimums ansteigt.

Es gibt also zwei Abbruchkriterien für das Training: Entweder wird die vorgegebene maximale Zyklenzahl überschritten, oder der Fehler der Testmenge überschreitet sein bisheriges Minimum plus einen frei wählbaren Prozentsatz. Da der zweite Fall in den ersten Zyklen aufgrund der zufälligen Initialisierung der Gewichte vorzeitig eintreten kann, wird weiterhin eine minimale Zyklenzahl vorgeschrieben, vor der das Training nicht beendet wird. Die verschiedenen Abbruchkriterien werden in Abbildung Abbildung 3.9 dargestellt.

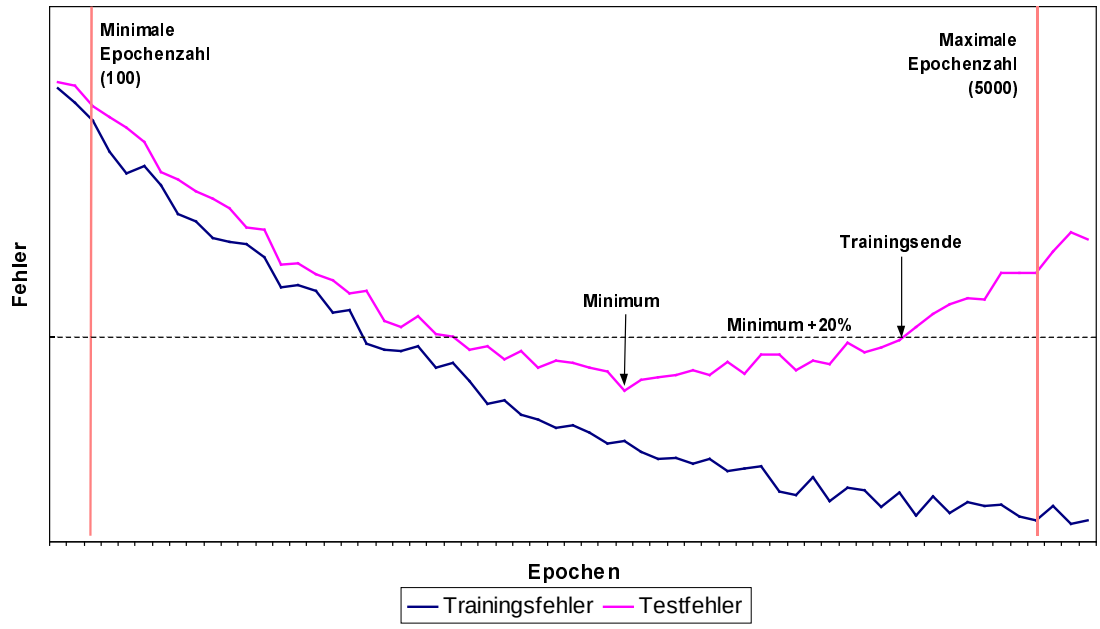

Abbildung 3.9: Abbruchkriterien und Verlauf von Trainings- und Testfehler während des Netztrainings

3.2.3 Arbeitsumgebung für neuronale Fragestellungen

Der Einsatz Neuronaler Netze im Rahmen dieser Arbeit wurde durch verschiedene Tools ermöglicht, die von der Universität Stuttgart (SNNS) und der Arbeitsgruppe Praktische Informatik der Universität Osnabrück entwickelt wurden (Arbeitsumgebung und Oberfläche zur Zeitreihenanalyse und –prognose).

3.2.3.1 Der Stuttgarter Neuronale Netze Simulator (SNNS)

„SNNS (Stuttgart Neural Network Simulator) is a simulator for neural networks developed at the Institute for Parallel and Distributed High Performance Systems (Institut für Parallele und Verteilte Höchstleistungsrechner, IPVR) at the University of Stuttgart since 1989.“[33]

Der in der Arbeitsgruppe von Andreas Zell entwickelte Simulator deckt die Bereiche Netzwerkgenerierung, Training, Test und Visualisierung der Neuronalen Netze ab[34], umfaßt jedoch keine Unterstützung bei der Verwaltung von Zeitreihen. SNNS besteht aus vier Hauptkomponenten: Dem Simulatorkern, einer graphischen Benutzeroberfläche, einem batch-Simulator und einem Netzwerkcompiler. Für die Untersuchungen im Rahmen dieser Arbeit wurde im wesentlichen mit der batch-Version von SNNS gear-

[33] Aus: [Zel 95], 1.

[34] Vgl.: [Zel 96].

beitet, die sich über ein in der Arbeitsgruppe Praktische Informatik der Universität Osnabrück entwickeltes Interface und einfache Skripte (Perl 5.0) ansteuern läßt[35].

Die Möglichkeit, SNNS-Aufrufe in Skripte einzubeziehen, erleichterte die Benutzung des Simulators und ermöglichte ein effizientes Arbeiten. Weiterhin wurden verschiedene Skripte verwendet, die Fehlerbetrachtungen zulassen.

3.2.3.2 Die Arbeitsumgebung

Unter der Arbeitsumgebung wird das Interface verstanden, mit dessen Hilfe SNNS über Skripte (Perl 5.0) angesprochen werden kann. Sie umfaßt beispielsweise Skripte zur Erstellung der von SNNS verwendeten ASCII-Netzbeschreibungsdatei aus Angaben über Topologie, Anzahl der Neuronen pro Schicht und Modifikationen der Topologie. Aber auch verschiedene Werkzeuge zur Analyse und Evaluierung von Modellzeitreihen nach verschiedenen Fehlermaßen wurden implementiert.

3.2.3.3 Die Oberfläche

Zu der Arbeitsumgebung wurde 1997 im Rahmen einer Diplomarbeit eine graphische Oberfläche entwickelt, die in erster Linie der Visualisierung der Zeitreihen dient[36]. Dabei können mehrere Zeitreihen gleichzeitig dargestellt werden, auch wenn sie in unterschiedlichen Einheiten und Zeitschritten gegeben sind. Darüber hinaus besteht die Möglichkeit, interaktiv Prognoseläufe vorzunehmen.

Für die vorliegende Arbeit zeigte sie sich besonders nützlich, wenn es darum ging, verschiedene Zeitreihen komfortabel auf einem bestimmten Zeitraum zu vergleichen.

3.2.4 Neuronale Sensitivitätsanalyse

3.2.5 Netzoptimierung: Pruning

Die vielen Freiheiten bei der Festlegung einer Netzwerkarchitektur lassen es sinnvoll erscheinen, es dem Neuronalen Netz selbst zu überlassen, eine geeignete Struktur zu finden. Dabei steht insbesondere eine Minimierung der Architektur im Vordergrund. Aufgrund der vielen freien Parameter, wie Auswahl der Eingabe- und Anzahl der Hidden-Neuronen, sowie der einzelnen Gewichte und Schwellenwerte, kann die Zahl der Trainingspaare nicht ausreichend sein, um sie alle optimal einzustellen. Ist dies der Fall, kann das Netzwerk zwar gut trainiert werden, generalisiert aber schlecht für unbekannte Eingabedaten[37]. Es sind also Verfahren notwendig, die die Komplexität[38] der Neuronalen Netze reduzieren und eine geeignete Netzwerkarchitektur finden. Verfahren dieser

[35] Vgl.: [Thi 98].

[36] Vgl. [Kro 97].

[37] Vgl. [Zel 96], 319.

[38] Die Komplexität entspricht der Anzahl freier Variablen.

Art nennt man *Pruning-Verfahren* (*engl*.: pruning = Beschneiden, Ausdünnen), da sie das Netz sukzessive um unwichtige Gewichte und Neuronen reduzieren. Am Ende des Prunings wird jedoch immer das bisher beste Netz verwandt. Es werden zwei Ansätze des Prunings unterschieden: Das Pruning der Gewichte (*weight pruning*) und das Pruning von Neuronen (*Skeletonization* oder *unit pruning*), welches sich weiter aufgliedert in ein Pruning der inneren Neuronen (*hidden unit pruning*) und ein Pruning der Eingabeneuronen (*input unit pruning*). Im folgenden werden die Ansätze vorgestellt und ihre Anwendungsmöglichkeiten im Rahmen der Netzwerkoptimierung skizziert.

3.2.5.1 Weight pruning

Eine Möglichkeit, die Komplexität eines Neuronalen Netzes einzuschränken, besteht darin, die Gewichtsmatrix auszudünnen. Dabei werden einige unwichtige Verbindungen durchtrennt[39]. Das so entstandene Netz hat zwar die gleiche Struktur, ist aber in seinen Verbindungen filigraner geworden. Das hier verwendete Verfahren Optimal Brain Damage wurde 1990 von Le Chun/Denker/Solla in ihrem Artikel *Optimal Brain Damage*[40] als Ausdünnungsmethode vorgeschlagen.

Die Vorgehensweise ist die folgende: „Optimal Brain Damage [...] versucht, analytisch den Effekt einer Änderung des Parametervektors (Gewichtsvektors) auf die Fehlerfunktion [...] vorherzusagen und diejenige Änderung des Parametervektors durchzuführen, die beim Nullsetzen eines Parameters (Gewichts) die geringste Änderung der Fehlerfunktion bewirkt.“[41] Dazu wird die Fehlerfunktion durch eine Taylorreihe um ein lokales Minimum herum angenähert. Eine genauere Darstellung und Herleitung findet sich bei [Zel 96], S.320 ff. und bei [Reh 94], S. 67 ff.

3.2.5.2 Hidden unit pruning

Dieser Spezialfall der 1989 von Mozer und Smolensky[42] entwickelten Netzwerk-Skelettierung (*skeletonization*) ist als Fortsetzungsverfahren des Gewichts-Prunings zu verstehen. Ein Neuron der verborgenen Schicht wird aus dem Netz entfernt, indem alle zu und von ihm weisenden Gewichte auf Null gesetzt werden. Dies kann natürlich bereits beim Gewichts-Pruning geschehen, allerdings ist es oftmals sinnvoll, weitere Neuronen zu eliminieren[43]. Dazu werden die Signalflüsse durch die inneren Neuronen verglichen und eine Kreuzkorrelationsanalyse durchgeführt. Eine hohe Korrelation zweier Hidden-Neuronen deutet auf eine Redundanz im Informationsfluß hin und darauf, daß die innere Schicht zu groß ist.

---

[39] D.h. die entsprechenden Gewichte werden gleich Null gesetzt, womit die Verbindung effektiv gekappt ist.

[40] Ausführliche Literaturangabe in: [Zel 96], 592.

[41] Aus: [Zel 96], 320.

[42] Ausführliche Literaturangabe in: [Zel 96], 595.

[43] Vgl. [Reh 94], 75.

3.2.5.3 Input unit pruning

Das Pruning von Input-Neuronen läuft nach demselben Schema ab wie das zuvor erklärte *hidden unit pruning*, allerdings werden hier die Korrelationen der Input-Neuronen analysiert. Für die Bewertung der Inputneuronen gibt es allerdings ein alternatives Verfahren, einen Rangordnungstest. Hierzu wird als Testgröße für jedes Inputneuron der Wert der Zielfunktion berechnet, um den sie sich ändert, wenn man den Input durch seinen Mittelwert ersetzt. Wäre bereits ein lokales Minimum gefunden, so würde die Zielfunktion unter Nutzung aller Inputs kleiner gleich der Zielfunktion bei Nutzung des verkleinerten Inputvektors sein. Die Testgröße wäre somit größer gleich Null und ihr Wert würde den Erklärungsbeitrag dieses speziellen Neurons beschreiben[44]. Das Neuron mit dem kleinsten Erklärungsbeitrag wird entfernt.

3.2.5.4 Anwendung mehrerer Pruning-Verfahren

Mehrere Pruning-Verfahren können hintereinander ausgeführt werden, um eine Optimierung in verschiedenen Aspekten zu gewährleisten. Allerdings „müssen die drei oben beschriebenen Kompressionsschritte in der angegebenen Reihenfolge angewandt werden.“[45]

Im Rahmen dieser Arbeit wurde SNNS um die Funktionalität des mehrfachen Prunings erweitert. Es ist nunmehr möglich, bis zu drei Pruning-Schritte hintereinander durchzuführen. Abbildung 3.10 zeigt den Ablauf des modifizierten Netztrainings.

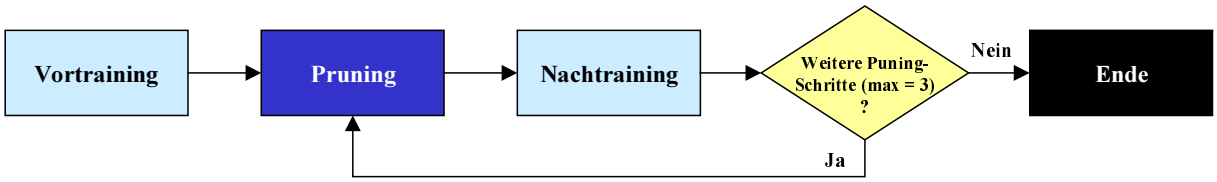

Abbildung 3.10: Erweiterung des SNNS-Trainings um mehrfaches Pruning

Somit können alle drei Pruning-Verfahren in das Netztraining eingebunden werden und eine Optimierung der Netzarchitektur in den drei vorgestellten Bereichen durchgeführt werden. Die Anwendung und der Vergleich der Verfahren zur Netzwerkoptimierung finden sich in Kapitel XXX.

3.3 Maße für die Analyse- und Prognosegüte

Für die Bewertung von mathematischen oder neuronalen Modellen ist es wichtig, vergleichbare Gütemaße für die Analyse und Prognose zu definieren. Diese werden generell in normierte und nichtnormierte Maße unterschieden. Während nichtnormierte Maße nur dann vergleichbar sind, wenn Analyse- bzw. Prognosezeitraum gleich sind,

[44] Vgl. [Reh 94], 76.

[45] Aus: [Reh 94], 65.

zeichnen sich normierte Maße durch einen festen Wertebereich (meist [0;1]) und eine generelle Vergleichbarkeit aus.

3.3.1 Ex-ante- und Ex-post-Beurteilung

Um überhaupt den Fehler einer Prognose bestimmen zu können, müssen offenbar die tatsächlichen Werte vorliegen. Bei der *Ex-ante*-Beurteilung wird die Prognose überprüft, bevor die Werte des Prognosezeitraums vorliegen. Da jedoch eine Beurteilung erst mit den tatsächlichen Werten des Prognosezeitraums stattfinden kann, beschränkt sich die *Ex-ante*-Beurteilung auf die Modellstruktur.

Im nachhinein, also *ex post*, kann eine Beurteilung stattfinden, die generell auf der Abweichung des Modellwertes $\hat{x}$ von dem wahren Wert x des Prognosezeitraums {t = 1, 2, ..., N} beruht. Diese Differenz e wird auch als Residuum bezeichnet:

$$e_t = x_t - \hat{x}_t$$

Die hier genannten Maße für Prognosefehler lassen sich auch auf den Fall der Analyse übertragen, deswegen wird im folgenden allgemein von Fehlermaßen gesprochen.

3.3.2 Nichtnormierte Gütemaße

Die in dieser Arbeit verwendeten nichtnormierten Fehler- bzw. Gütemaße können Tabelle 3.1 entnommen werden.

Tabelle 3.1: Übersicht über die verwendeten nichtnormierten Fehler- bzw. Gütemaße

Fehler- bzw. Gütemaß	Formel
Mittlerer relativer Fehler	$MRE = \frac{1}{N} \cdot \sum_{t=1}^{N} \frac{\lvert e_t \rvert}{x_t}$
Mittlerer absoluter prozentualer Fehler	$MAPE = \frac{100}{N} \cdot \sum_{t=1}^{N} \left\lvert \frac{e_t}{x_t} \right\rvert$
Summe der quadratischen Fehler (Residuen)	$SSE = \sum_{t=1}^{N} e_t^2$
Wurzel des mittleren quadratischen Prognosefehlers	$RMSE = \sqrt{\frac{1}{N} \sum_{1}^{N} e_t^2}$

3.3.3 Normierte Gütemaße

Im Rahmen dieser Arbeit wurden zwei normierte Gütemaße verwandt, die eine Klassifizierung von Modellen ermöglichen. Eine Übersicht über die verschiedenen Maße findet sich in Tabelle 3.2.

Der *Ungleichheitskoeffizient U nach Theil* vergleicht die Modellergebnisse mit der naiven Prognose, bei der für jeden Prognosewert der Wert des vorigen Zeitpunktes genommen wird. Für eine perfekte Prognose liefert er den Wert Null, wie auch alle anderen Fehlermaße. Generell spricht ein kleinerer Wert für ein besseres Modell. Als Vergleichswert dient die naive Prognose, bei der jeweils der Wert des Vortages verwandt wird. Die naive Prognose erzielt einen Ungleichheitskoeffizienten von eins. Ist U für betrachtete Modelle größer als Eins, so bringt das Modell keine Verbesserung gegenüber dem naiven Verfahren.

Der *Determinationskoeffizient* r^2 entspricht dem Quadrat des Korrelationskoeffizienten r zwischen der prognostizierten Zeitreihe und den tatsächlichen Werten. Um r zu berechnen, wird das Produkt der Abweichungen vom Mittelwert der Zeitreihen für die einzelnen Zeitpunkte t aufsummiert und durch beide Standardabweichungen geteilt. Der Wertebereich von r^2 ist [0;1], wobei Null keine und Eins eine perfekte Korrelation bedeutet.

Tabelle 3.2: Übersicht über die verwendeten normierten Fehler- bzw. Gütemaße

Fehler- bzw. Gütemaß	Formel
Ungleichheitskoeffizient von Theil	$U = \sqrt{\frac{\sum_{t=1}^{N} e_t^2}{\sum_{t=1}^{N} (x_t - x_{t-1})^2}}$
Determinationskoeffizient	$r^2 = \frac{\left(\sum_{t=1}^{N} (x_t - \bar{x}) \cdot (\hat{x}_t - \bar{\hat{x}}) \right)^2}{\sum_{t=1}^{N} (x_t - \bar{x})^2 \cdot \sum_{t=1}^{N} (\hat{x}_t - \bar{\hat{x}})^2}$

3.4 Verfahren zur Auswahl exogener Zeitreihen

Die Auswahl geeigneter exogener Zeitreihen ist ausschlaggebend für die Qualität der erstellten Modelle. Dabei möchte man einerseits möglichst viele exogene Datenreihen in Modelle integrieren, um mehr Informationen gewinnen zu können, also das Bestimmtheitsmaß möglichst hoch zu halten. Andererseits ist man bemüht, die Modelle so einfach wie möglich zu halten, sich also auf die wesentlichen Einflußfaktoren zu beschränken. Dies bedeutet aber eine Einschränkung der verwendeten exogenen Zeitreihen. Somit besteht ein Zielkonflikt: Das erstellte Modell soll so einfach wie möglich, aber so komplex wie nötig sein, um eine gute Analyse- bzw. Prognosequalität zu garantieren. Bei der multivariaten Zeitreihenanalyse stellt sich das Problem, aus den möglichen exogenen Datenreihen diejenigen auszuwählen, für die das erstellte Modell den geringsten Fehler liefert. Offenbar sind in einem realen System die Einflüsse der einzelnen Faktoren unterschiedlich groß. In vielen Fällen – wie auch in dieser Arbeit - ist es

nicht möglich, alle möglichen Kombinationen exogener Zeitreihen zu überprüfen, da der Rechenaufwand für größere Probleme die vorhandenen Möglichkeiten übersteigt: Die Anzahl $P_{n,k}$ der Möglichkeiten, aus n vorhandenen exogenen Zeitreihen k für ein Modell auszuwählen, beträgt

$$P_{n,k} = \binom{n}{k}.$$

In dieser Arbeit stehen für die Analyse 22 und für die Prognose 23 exogene Zeitreihen zur Verfügung[46]. Werden nun für die Prognose vier bzw. fünf exogene Zeitreihen ausgewählt, so erhält man

$$P_{23,4} = \binom{23}{4} = 8855 \quad \textit{bzw.} \quad P_{23,5} = \binom{23}{5} = 33649$$

verschiedene Möglichkeiten. Bei einem Zeitaufwand[47] von 50 Sekunden pro Regressionslauf (GAUSS) bzw. 21 Minuten[48] pro Netztraining (SNNS) ergeben sich für fünf exogene Zeitreihen ein zeitlicher Gesamtaufwand von 19,5 Tagen für die Regression und von 490,7 Tagen für das Netztraining. Würde man die Anzahl der möglichen Zeitreihen auf 14 verringern und sich auf vier davon beschränken, ließe sich der Aufwand auf 0,6 bzw. 14,6 Tage reduzieren.

Es ist also insbesondere für neuronale Modelle wünschenswert, im Vorfeld die beste Kombination exogener Zeitreihen für die Modellierung zu bestimmen, ohne alle Möglichkeiten zu testen. Leider kann nur dann sicher behauptet werden, die optimalen Einflußzeitreihen gefunden zu haben, wenn alle anderen Kombinationen getestet wurden. Es ist allerdings möglich, durch geeignete Algorithmen eine „gute“, d.h. vielversprechende Kombination zu bestimmen. Im folgenden werden zwei Verfahren[49] vorgestellt und diskutiert, die eine systematische Auswahl ermöglichen.

3.4.1 Auswahl nach Korrelationskoeffizient

Eine Möglichkeit, exogene Zeitreihen für ein Modell auszuwählen, besteht darin, alle verwendeten Zeitreihen nach ihrem Korrelationskoeffizienten zur endogenen Zeitreihe (hier CO_2) zu sortieren und ein Modell aus den besten k zu erstellen. Jetzt wird Schritt für Schritt die Zeitreihe mit dem schlechtesten Korrelationskoeffizienten entfernt und durch den nächsten Kandidaten auf der (geordneten) Liste der Kandidaten ersetzt. Verbessert sich das Modell, bleibt er im Modell, ansonsten wird zum bisher besten Modell zurückgekehrt. Wird eine Zeitreihe aus dem Modell entfernt, kommt sie von neuem in

---

[46] Für die Prognose kommen die zeitlich zurückliegenden Werte der CO_2-Zeitreihe hinzu.

[47] Die Werte für den Zeitaufwand beziehen sich auf eine *Sun Sparc 20 Doppelprozessor* (GAUSS) bzw. *eine HP Apollo 9000/730* (SNNS), die im Rahmen dieser Arbeit verwendet wurden.

[48] Die verwendeten Werte sind Durchschnittswerte.

[49] Für neuronale Modelle kommt das bereits in Kapitel 3.2.5 vorgestellte Pruning hinzu.

die Liste. Wird sie dem Modell hinzugefügt, kann es aber nicht verbessern, wird sie verworfen. Das Verfahren wird solange durchgeführt, bis die Liste leer ist[50].

Für dieses Verfahren spricht vor allem, daß es sowohl für multivariate mathematische als auch für neuronale Modelle eingesetzt werden kann und bei diesen die Systemstruktur im Verlauf des Verfahrens nicht ändert. Leider kann auch dieses Verfahren nicht die optimale Lösung garantieren: Es ist denkbar, daß eine Zeitreihe verworfen wird, die in Kombination mit anderen Zeitreihen eine bessere Lösung bringen würde Da die Entwicklung dieses Verfahrens noch nicht abgeschlossen ist, kann es noch nicht zu vergleichenden Tests herangezogen werden. Es ist daher nur zur Illustration einer anderen Herangehensweise gedacht.

3.4.2 Auswahl nach Sensitivität

Ein anderes Verfahren bedient sich der (neuronalen) Sensitivitätsanalyse, um geeignete Einflußzeitreihen zu bestimmen. Dazu werden zunächst alle Zeitreihen für ein Modell verwendet. Nach jedem (vollständig abgeschlossenem) Netztraining wird die neuronale Sensitivitätsanalyse durchgeführt und das Inputneuron (die exogene Zeitreihe) mit der geringsten Sensitivität „geprunt“, d.h. entfernt. Dies wird sukzessive wiederholt, bis die gewünschte Anzahl exogener Zeitreihen erreicht ist.

Der Vorteil dieses Verfahrens liegt vor allem in seiner kurzen konstanten Laufzeit, es ist jedoch möglich, daß exogene Zeitreihen aufgrund ihres geringen Einflusses zu Anfang entfernt werden, später (bei weniger Inputneuronen) allerdings das Netz verbessern könnten.

Ein weiteres Problem dieses Verfahrens ist seine Strukturinkonsistenz. Während des Ablaufs ändert sich die Netzstruktur (hier: die Anzahl der Inputneuronen) ständig. Aus diesem Grund ist eine Variation der Sensitivität einzelner Inputgrößen möglich.

[50] Siehe [Ana 98].

4 Ergebnisse

4.1 Analyse der Datenreihen

4.1.1 Die CO_2-Datenreihe

4.1.1.1 Saisonale Periodizität

Eine Aggregation der Monatswerte der CO_2-Konzentration über die Jahre 1992 bis 1996 ergab eine deutliche Periode, wie auch Abbildung 4.1 zu entnehmen ist.

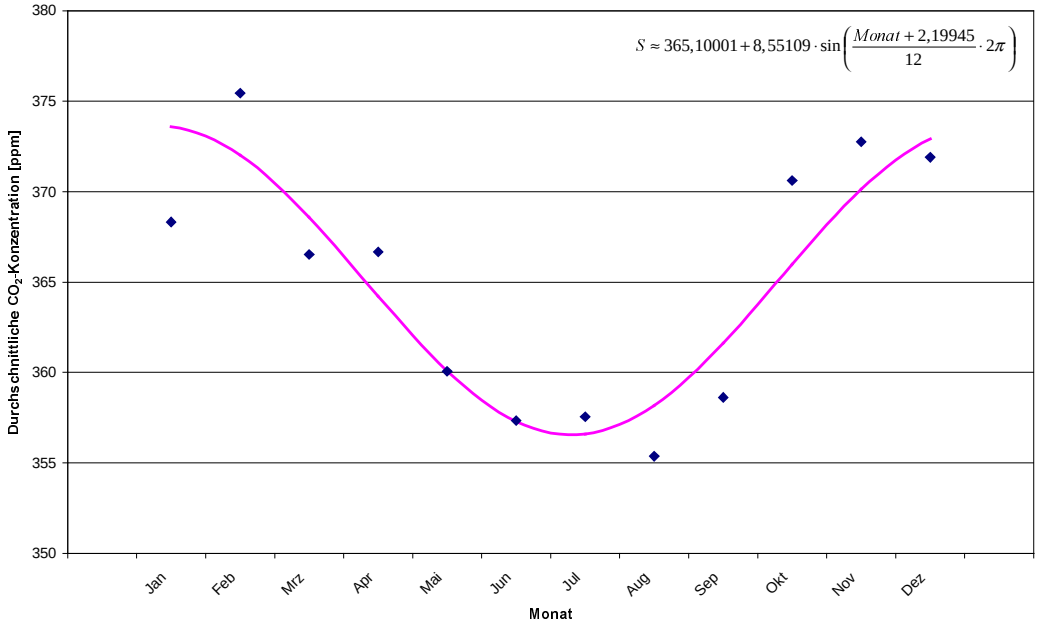

Abbildung 4.1: Monatliche CO_2-Konzentration (Durchschnitt der Jahre 1992 - 1996)

Bei der Bestimmung der Saisonkomponente wurden die Monatsmittelwerte über alle Jahre verwandt und eine Periode von zwölf Monaten vorausgesetzt. Die Modellgleichung lautet:

$$S \approx A + B \cdot \sin\left(\frac{Monat + C}{12} \cdot 2\pi\right)$$

Dabei entspricht A dem Mittelwert, um den die Werte schwanken, B einem Faktor, der die Amplitude der Schwingung darstellt und C der zeitlichen Verschiebung der Amplitude. Die Methode der kleinsten Abweichungsquadrate ergab folgende Werte für die Konstanten:

$$A = 365{,}10001 \qquad B = 8{,}55109 \qquad C = 2{,}19945$$

Somit erhält man

$$S \approx 365{,}10001 + 8{,}55109 \cdot \sin\left(\frac{Monat + 2{,}19945}{12} \cdot 2\pi\right).$$

4.1.1.2 Trend

Abbildung 4.2 zeigt den linearen Trend der CO_2-Datenreihe über die Jahre 1992 bis 1996.

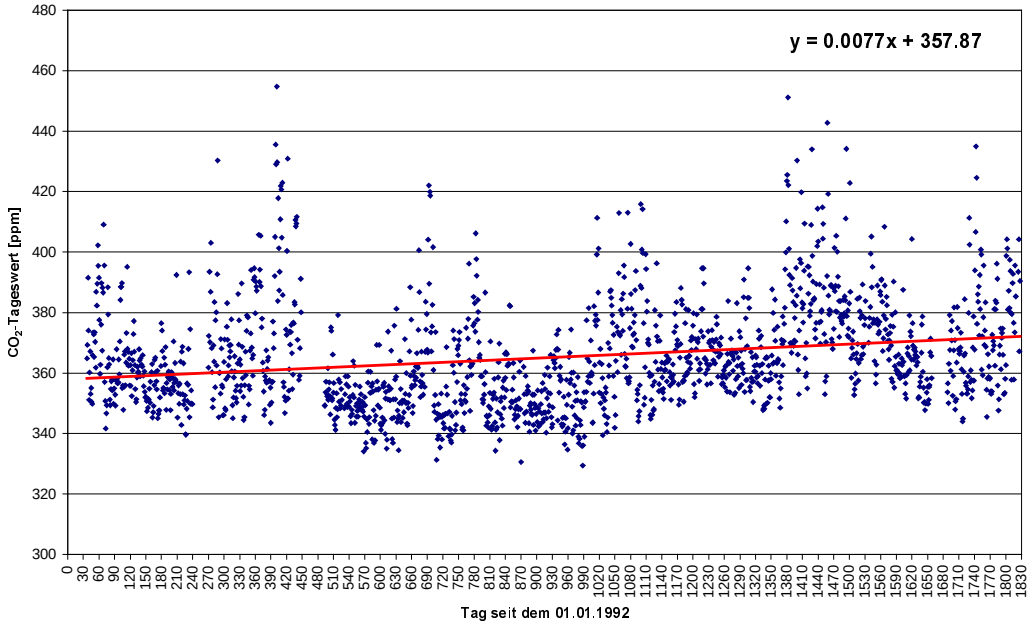

Abbildung 4.2: Linearer Trend der CO_2-Konzentration über die Jahre 1992 bis 1996

Diese Trendbetrachtung zeigt zudem einen Anstieg der durchschnittlichen jährlichen CO_2-Konzentration von ca. 2,8124 ppm.

4.1.1.3 Autokorrelation

Viele natürliche Zeitreihen sind autokorreliert, d.h., es besteht ein Zusammenhang zwischen zeitlich aufeinander folgenden Meßwerten. Die Autokorrelation entspricht dem Korrelationskoeffizienten zwischen der Zeitreihe und der zeitlich verschobenen bzw. geshifteten Zeitreihe. Eine hohe Autokorrelation spricht für die Verwendung der geshifteten Zeitreihe im (linearen) Modell, also für Autoregressive Ansätze, wie sie definitionsgemäß bei ARIMA-Modellen zu finden sind. Trägt man die Autokorrelationen gegen die Shifts auf, erhält man ein sogenanntes Korrelogramm. Mit Hilfe des Korrelogramms können Periodizitäten - falls vorhanden - identifiziert werden[51]. Abbildung 4.3 zeigt das Korrelogramm der CO_2-Datenreihe für die Shifts 0 bis 10.

[51] Dies gilt insbesondere für eine kleine Anzahl Shifts. Eine geeignetere Form, Periodizitäten in Zeitreihen zu identifizieren, stellt die Spektralanalyse dar, die ebenfalls auf der Autokorrelationsfunktion aufbaut.

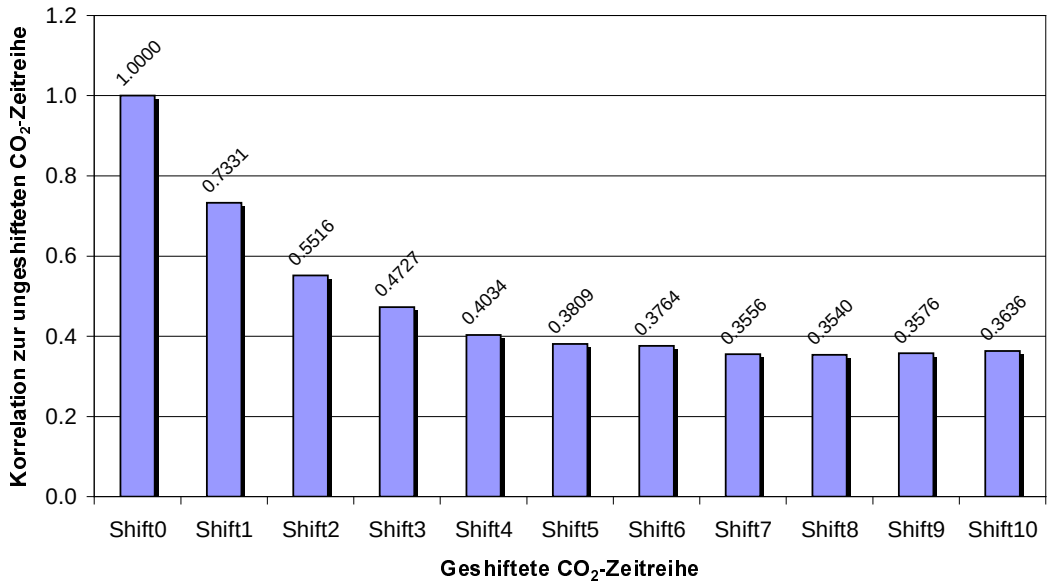

Abbildung 4.3: Korrelogramm der CO_2-Zeitreihe

Die hohe Autokorrelation für den Meßwert vom Vortag (Shift 1) und dem Tag davor (Shift 2) im Vergleich zur Korrelation zu den anderen Zeitreihen (siehe Kapitel 4.1.2), lassen eine Verwendung der geshifteten Zeitreihen bei der Modellierung sinnvoll erscheinen.

4.1.1.4 Korrelationen zu anderen Zeitreihen

Anhand des Korrelationskoeffizienten nach Pearson-Spearman[52] lassen sich im Vorfeld günstige exogene Zeitreihen für die (lineare) Modellierung identifizieren[53]. Dabei ist der Absolutbetrag des Korrelationskoeffizienten entscheidend. Tabelle 4.1 zeigt die Korrelationskoeffizienten der CO_2-Zeitreihe mit den anderen zur Verfügung stehenden Datenreihen.

[52] Vgl. [Duf 92], 70; [Sch 91], 4 und 76 ff.; [Bil 81], 62.

[53] Das beste lineare Modell setzt sich allerdings nicht immer aus den Zeitreihen mit der höchsten Korrelation zusammen. Siehe dazu Kapitel 4.1.2.

Tabelle 4.1: Korrelationen zur CO_2-Zeitreihe, nach absolutem Wert geordnet

Zeitreihe	Korrelation
CO2-Vortageswert	0.7284
WG	-0.4315
TMP_DWD	-0.3639
TMP_FH	-0.3618
JTGCOS	0.3533
WRSIN	0.3290
WRGSIN	0.3123
THT_FH	-0.3097
THT_DWD	-0.3044
WR	-0.2704
NS_DWD	-0.2651
WRG	-0.2574
ND_FH	-0.2427
TAG	0.2248
BG	-0.1680
RLF	0.1444
WRGCOS	-0.0810
WRCOS	-0.0764
SSD	-0.0641
JTGSIN	0.0290
HAUDE	-0.0226
SDEF	-0.0153
JTG	0.0073

4.1.2 Andere Datenreihen

In Kapitel 4.1.1.4 wurden die Korrelationen der CO_2-Zeitreihe mit anderen Zeitreihen als ein Weg beschrieben, eine mögliche Vorauswahl der Daten für ein Modell zu treffen. Es ist allerdings zu beachten, daß zwischen Zeitreihen, die zu den CO_2-Daten hoch korreliert sind, auch untereinander große Korrelationen bestehen können, die eine Verwendung im Modell u.U. redundant machen. Man spricht in diesem Fall von verborgenen Korrelationen. Ein natürliches Beispiel dafür sind die Bedeckungsgrad-Zeitreihe (BG) und die Sonnenscheindauer-Zeitreihe (SSD).

Um einen Überblick über die Korrelationen der Zeitreihen untereinander zu erhalten, werden oftmals sogenannte Korrelationsmatrizen aufgestellt, die die Korrelation jeder Zeitreihe mit allen anderen zeigen. Eine ausführliche Korrelationsmatrix ist in Tabelle 4.2 (Teil 1) und Tabelle 4.3 (Teil 2) gegeben.

Tabelle 4.2: Korrelationsmatrix (Teil 1)

	CO2	BG	HAUDE	JTAG	JTAGSIN	JTAGCOS	NS-DWD	NS-FH	RLF	SDEF	SSD	TAG	TMP-DWD
CO2	+1.0000	- 0.1680	- 0.0226	+0.0073	+0.0290	+0.3533	- 0.2651	- 0.2427	+0.1444	- 0.0153	- 0.0641	+0.2248	- 0.3639
BG	- 0.1680	+1.0000	- 0.3914	+0.0169	+0.0127	+0.1981	+0.2344	+0.3288	+0.5754	- 0.4004	- 0.7869	- 0.0073	- 0.1820
HAUDE	- 0.0226	- 0.3914	+1.0000	+0.0039	- 0.0370	- 0.4366	- 0.1294	- 0.1538	- 0.5334	+0.9956	+0.5273	+0.4067	+0.5552
JTAG	+0.0073	+0.0169	+0.0039	+1.0000	- 0.7765	- 0.0000	+0.0323	+0.0300	+0.1550	+0.0192	- 0.0504	+0.2000	+0.1483
JTAGSIN	+0.0290	+0.0127	- 0.0370	- 0.7765	+1.0000	- 0.0000	- 0.0529	- 0.0545	- 0.1920	- 0.0736	+0.0445	- 0.1553	- 0.2194
JTAGCOS	+0.3533	+0.1981	- 0.4366	- 0.0000	- 0.0000	+1.0000	+0.0353	+0.0291	+0.4966	- 0.4119	- 0.5393	+0.0000	- 0.8126
NS-DWD	- 0.2651	+0.2344	- 0.1294	+0.0323	- 0.0529	+0.0353	+1.0000	+0.5617	+0.2117	- 0.1280	- 0.1914	- 0.0485	+0.0039
NS-FH	- 0.2427	+0.3288	- 0.1538	+0.0300	- 0.0545	+0.0291	+0.5617	+1.0000	+0.2927	- 0.1559	- 0.2639	- 0.0543	+0.0177
RLF	+0.1444	+0.5754	- 0.5334	+0.1550	- 0.1920	+0.4966	+0.2117	+0.2927	+1.0000	- 0.5236	- 0.7410	+0.0033	- 0.4668
SDEF	- 0.0153	- 0.4004	+0.9956	+0.0192	- 0.0736	- 0.4119	- 0.1280	- 0.1559	- 0.5236	+1.0000	+0.5214	+0.4210	+0.5502
SSD	- 0.0641	- 0.7869	+0.5273	- 0.0504	+0.0445	- 0.5393	- 0.1914	- 0.2639	- 0.7410	+0.5214	+1.0000	+0.0193	+0.4917
TAG	+0.2248	- 0.0073	+0.4067	+0.2000	- 0.1553	+0.0000	- 0.0485	- 0.0543	+0.0033	+0.4210	+0.0193	+1.0000	- 0.0643
TMP-DWD	- 0.3639	- 0.1820	+0.5552	+0.1483	- 0.2194	- 0.8126	+0.0039	+0.0177	- 0.4668	+0.5502	+0.4917	- 0.0643	+1.0000
TMP-FH	- 0.3618	- 0.1838	+0.5236	+0.1444	- 0.2232	- 0.8043	+0.0045	+0.0145	- 0.5083	+0.5168	+0.5036	- 0.0426	+0.9900
THT-DWD	- 0.3044	- 0.2805	+0.6121	+0.1238	- 0.1898	- 0.8218	- 0.0405	- 0.0273	- 0.5259	+0.6078	+0.5754	- 0.0336	+0.9819
THT-FH	- 0.3097	- 0.3020	+0.5450	+0.1256	- 0.2036	- 0.8180	- 0.0306	- 0.0226	- 0.5602	+0.5397	+0.6026	- 0.0451	+0.9768
WG	- 0.4315	+0.2634	- 0.2242	- 0.1516	+0.1351	+0.3060	+0.3009	+0.2583	+0.0577	- 0.2169	- 0.2673	- 0.0765	- 0.2028
WR	- 0.2704	+0.2989	- 0.0798	- 0.0519	+0.0388	- 0.0864	+0.2030	+0.1350	+0.1671	- 0.0846	- 0.1742	- 0.0586	+0.0877
WRSIN	+0.3290	- 0.3891	+0.1502	+0.0455	- 0.0354	+0.0590	- 0.2565	- 0.1941	- 0.2407	+0.1551	+0.2708	+0.1003	- 0.0951
WRCOS	- 0.0764	- 0.0965	+0.0208	- 0.0444	+0.0916	- 0.1611	- 0.1033	- 0.1637	- 0.0480	+0.0091	+0.1344	+0.0397	- 0.0579
WRG	- 0.2574	+0.3180	- 0.0983	- 0.0511	+0.0511	- 0.0859	+0.2083	+0.1380	+0.1742	- 0.1039	- 0.1961	- 0.0628	+0.0841
WRGSIN	+0.3123	- 0.3913	+0.1469	+0.0405	- 0.0304	+0.0631	- 0.2610	- 0.1999	- 0.2395	+0.1514	+0.2695	+0.1032	- 0.1080
WRGCOS	- 0.0810	- 0.0929	+0.0322	- 0.0467	+0.0988	- 0.1735	- 0.1026	- 0.1702	- 0.0507	+0.0192	+0.1394	+0.0394	- 0.0502
CO_2-Shift1	+0.7284	- 0.0348	- 0.0248	+0.0096	+0.0352	+0.3519	- 0.2344	- 0.1961	+0.1317	- 0.0180	- 0.1172	+0.2236	- 0.3120
CO_2-Shift2	+0.5426	+0.0080	- 0.0364	+0.0095	+0.0413	+0.3512	- 0.1536	- 0.1316	+0.1410	- 0.0294	- 0.1503	+0.2237	- 0.2940

Tabelle 4.3: Korrelationsmatrix (Teil 2)

	TMP-FH	THT-DWD	THT-FH	WG	WR	WRSIN	WRCOS	WRG	WRGSIN	WRGCOS	CO_2-Shift1	CO_2-Shift2
CO2	-0.3618	-0.3044	-0.3097	-0.4315	-0.2704	+0.3290	-0.0764	-0.2574	+0.3123	-0.0810	+0.7284	+0.5426
BG	-0.1838	-0.2805	-0.3020	+0.2634	+0.2989	-0.3891	-0.0965	+0.3180	-0.3913	-0.0929	-0.0348	+0.0080
HAUDE	+0.5236	+0.6121	+0.5450	-0.2242	-0.0798	+0.1502	+0.0208	-0.0983	+0.1469	+0.0322	-0.0248	-0.0364
JTAG	+0.1444	+0.1238	+0.1256	-0.1516	-0.0519	+0.0455	-0.0444	-0.0511	+0.0405	-0.0467	+0.0096	+0.0095
JTAGSIN	-0.2232	-0.1898	-0.2036	+0.1351	+0.0388	-0.0354	+0.0916	+0.0511	-0.0304	+0.0988	+0.0352	+0.0413
JTAGCOS	-0.8043	-0.8218	-0.8180	+0.3060	-0.0864	+0.0590	-0.1611	-0.0859	+0.0631	-0.1735	+0.3519	+0.3512
NS-DWD	+0.0045	-0.0405	-0.0306	+0.3009	+0.2030	-0.2565	-0.1033	+0.2083	-0.2610	-0.1026	-0.2344	-0.1536
NS-FH	+0.0145	-0.0273	-0.0226	+0.2583	+0.1350	-0.1941	-0.1637	+0.1380	-0.1999	-0.1702	-0.1961	-0.1316
RLF	-0.5083	-0.5259	-0.5602	+0.0577	+0.1671	-0.2407	-0.0480	+0.1742	-0.2395	-0.0507	+0.1317	+0.1410
SDEF	+0.5168	+0.6078	+0.5397	-0.2169	-0.0846	+0.1551	+0.0091	-0.1039	+0.1514	+0.0192	-0.0180	-0.0294
SSD	+0.5036	+0.5754	+0.6026	-0.2673	-0.1742	+0.2708	+0.1344	-0.1961	+0.2695	+0.1394	-0.1172	-0.1503
TAG	-0.0426	-0.0336	-0.0451	-0.0765	-0.0586	+0.1003	+0.0397	-0.0628	+0.1032	+0.0394	+0.2236	+0.2237
TMP-DWD	+0.9900	+0.9819	+0.9768	-0.2028	+0.0877	-0.0951	-0.0579	+0.0841	-0.1080	-0.0502	-0.3120	-0.2940
TMP-FH	+1.0000	+0.9704	+0.9743	-0.1959	+0.0686	-0.0732	-0.0750	+0.0636	-0.0860	-0.0681	-0.3204	-0.3068
THT-DWD	+0.9704	+1.0000	+0.9850	-0.2576	+0.0453	-0.0305	-0.0478	+0.0406	-0.0450	-0.0391	-0.2817	-0.2794
THT-FH	+0.9743	+0.9850	+1.0000	-0.2583	+0.0332	-0.0178	-0.0572	+0.0279	-0.0324	-0.0488	-0.2925	-0.2936
WG	-0.1959	-0.2576	-0.2583	+1.0000	+0.2051	-0.3198	-0.1149	+0.2036	-0.3079	-0.1184	-0.2313	-0.1075
WR	+0.0686	+0.0453	+0.0332	+0.2051	+1.0000	-0.8499	-0.0237	+0.9257	-0.8403	-0.0069	-0.2234	-0.1783
WRSIN	-0.0732	-0.0305	-0.0178	-0.3198	-0.8499	+1.0000	+0.1184	-0.8413	+0.9908	+0.0988	+0.2640	+0.1965
WRCOS	-0.0750	-0.0478	-0.0572	-0.1149	-0.0237	+0.1184	+1.0000	-0.0391	+0.1503	+0.9801	-0.0526	-0.0481
WRG	+0.0636	+0.0406	+0.0279	+0.2036	+0.9257	-0.8413	-0.0391	+1.0000	-0.8481	-0.0197	-0.2187	-0.1776
WRGSIN	-0.0860	-0.0450	-0.0324	-0.3079	-0.8403	+0.9908	+0.1503	-0.8481	+1.0000	+0.1307	+0.2577	+0.1943
WRGCOS	-0.0681	-0.0391	-0.0488	-0.1184	-0.0069	+0.0988	+0.9801	-0.0197	+0.1307	+1.0000	-0.0630	-0.0570
CO_2-Shift1	-0.3204	-0.2817	-0.2925	-0.2313	-0.2234	+0.2640	-0.0526	-0.2187	+0.2577	-0.0630	+1.0000	+0.7288
CO_2-Shift2	-0.3068	-0.2794	-0.2936	-0.1075	-0.1783	+0.1965	-0.0481	-0.1776	+0.1943	-0.0570	+0.7288	+1.0000

Neben der Bestimmung von linearen Einflüssen kann die neuronale Sensitivitätsanalyse dazu herangezogen werden, nichtlineare Abhängigkeiten zu erkennen. Dazu wird ein Neuronales Netz mit allen möglichen exogenen Variablen trainiert und die Sensitivität der einzelnen Input-Neuronen bestimmt. Theoretisch kann aus der Gewichtung der Einfluß der entsprechenden Neuronen auf das neuronale Modell abgelesen werden. Abbildung 4.4 zeigt eine solche Sensitivitätsanalyse. Größere Spannweiten deuten auf eine größere Bedeutung der einzelnen Neuronen hin.

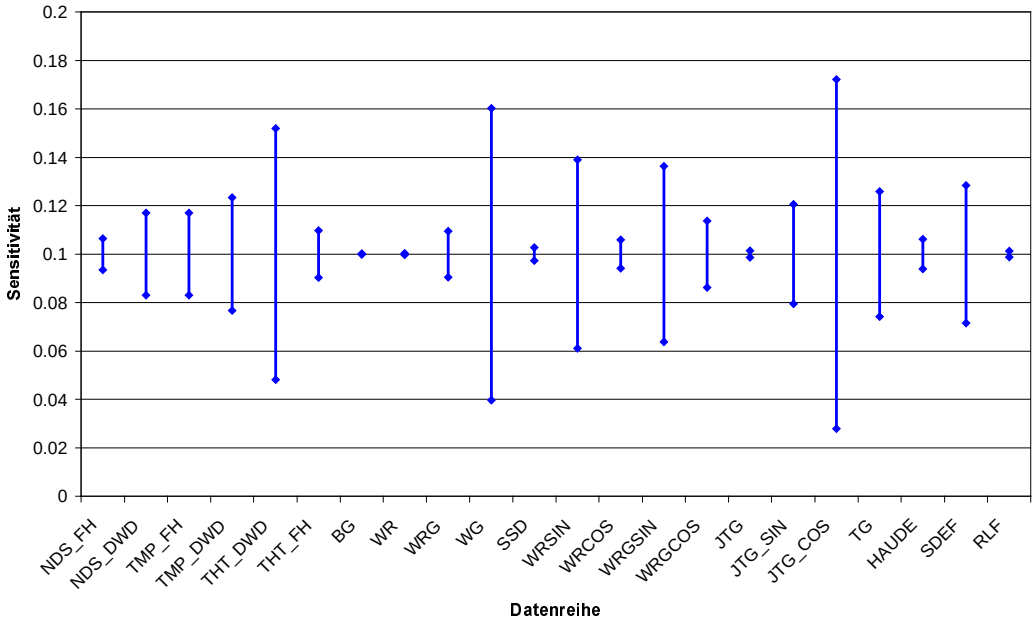

Abbildung 4.4: Neuronale Sensitivitäten der einzelnen exogenen Parameter

4.2 Vergleich der Methoden

4.2.1 Vorgehen

Für die vergleichende Analyse der Neuronalen Netze und mathematischen Modelle wurden verschiedene *Prognose*- bzw. *Analysezeiträume* gewählt, um ein möglichst umfassendes Bild über die Analyse- und Prognosefähigkeiten der einzelnen Verfahren zu gewinnen.

4.2.2 Erwartungen

Aufgrund ihres modularen Aufbaus sind Neuronale Netze trotz ihrer relativ einfach strukturierten Elemente in der Lage, komplexe Wirkzusammenhänge zu lernen und wiederzugeben. Einfache mathematische Modelle, wie etwa die multiple lineare Regression (siehe Abschnitt 3.1.1), können durch bestimmte Netzstrukturen exakt abgebildet werden. Dazu wird auf verborgene Schichten verzichtet und die Input-Schicht direkt mit dem einzigen Neuron der Output-Schicht verbunden. Als Aktivierungs- und Ausgabefunktion wird die Identität *id*() verwandt. Werden die exogenen Zeitreihen unskaliert verwendet, so entspricht die Gewichtsmatrix (w_{ij}, hier jedoch nur ein Vektor w_i) den Faktoren a_i der gelösten Regressionsgleichung. Soll eine Regression mit Konstante (a_0) nachgebildet werden, so muß ein weiteres Inputneuron mit dem konstanten Wert 1 hinzugefügt werden. Das Gewicht w_0 entspricht dann der Konstante a_0 der Regressions-

gleichung. Abbildung 4.5 zeigt ein solches Netz[54], wobei das einzige Neuron im eigentlichen Sinn[55] das der Ausgabeschicht ist. Eine äquivalente Netztopologie ist auch mit einem einzelnen Hidden-Neuron möglich.

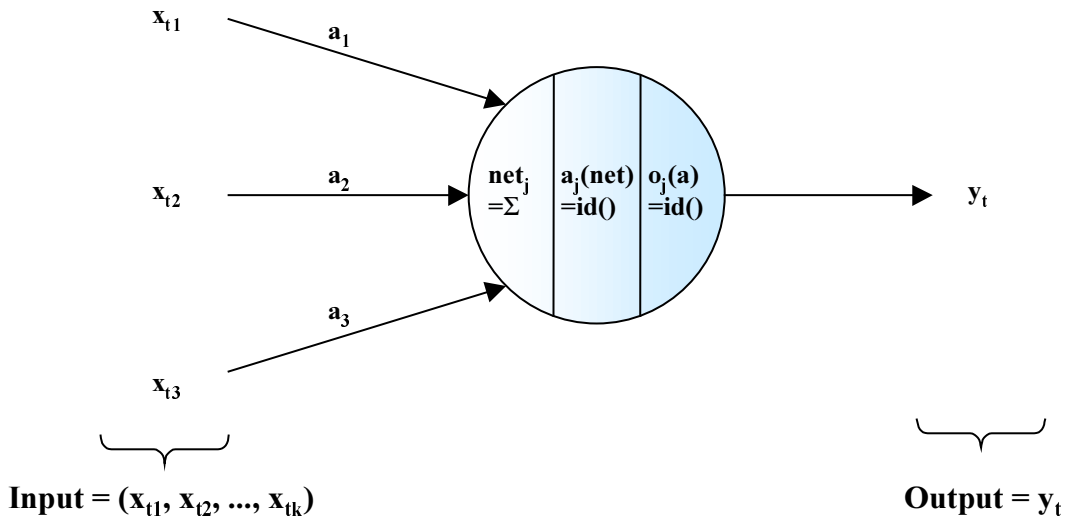

Abbildung 4.5: Repräsentierung der linearen Regression als Neuronales Netz

Aufgrund dieser Überlegungen wird klar, daß Neuronale Netze potentiell komplexere Zusammenhänge darstellen können als es die Regression kann, zumal bereits die Hinzunahme von Shortcut-Verbindungen und einem konstanten Eingabeneuron ein beliebiges Netz um die Funktionalität der multiplen linearen Regression erweitert. Folglich muß erwartet werden, daß Neuronale Netze in der Analyse und Prognose bessere Ergebnisse erzielen als die multiple lineare Regression.

4.2.3 Der Analyse- bzw. Prognosezeitraum

Für eine repräsentative Aussage bezüglich des Analyse- und Prognosezeitraums wurden sechs verschiedene Zeiträume aus dem Jahr 1996[56] gewählt, die unterschiedliche saisonale Gegebenheiten repräsentieren. Das Hauptaugenmerk liegt dabei auf der Vegetation, die einen puffernden Effekt auf die CO_2-Konzentration ausübt[57], und auf der Temperatur. Bei niedrigeren Temperaturen ist mit erhöhten Kohlendioxid-Emissionen aufgrund vermehrter Heizleistung zu rechnen, während bei extrem hohen Temperaturen durch Kühlung entsprechende Emissionen hervorgerufen werden können. Insgesamt werden

[54] Das hier dargestellte Netz verfügt jedoch nicht über das beschriebene konstante Inputneuron und somit nicht über das Gewicht a_0.

[55] Da die Neuronen der Eingabeschicht weder über Aktivierungs- noch Ausgabefunktion verfügen, gelten sie nur schematisch als Neuronen. Siehe dazu auch Kapitel 3.2.

[56] Das Jahr 1996 wurde gewählt, weil zeitlich zurückliegend bei diesem die meisten Daten existieren und so eine bessere Analyse bzw. Prognose ermöglicht wird. Abgesehen davon erscheint es sinnvoll, insbesondere bei Prognosen den jüngsten Zeitraum zu wählen

[57] Siehe Kapitel 1.

sechs Zeiträume von 30 validen Tagen[58] verwendet, die in jeweils zwei benachbarten Monaten des letzten Jahres des Datenbestandes (1996) liegen (Januar/Februar, März/April, etc.). Auf diese Weise soll den unterschiedlichen saisonalen Gegebenheiten Rechnung getragen werden.

Der untersuchte Zeitraum besteht aus den letzten 15 validen Tagen des ersten und aus den ersten 15 validen Tagen des zweiten Monats[59]. Eine Übersicht über die Zeiträume findet sich in Tabelle 4.4. Es wird jeweils eine Analyse und eine Prognose durchgeführt. Für eine *Analyse* stehen die Werte der exogenen Zeitreihen zur Zeit des zu prognostizierenden Datums zur Verfügung. Vorherige Werte der CO_2-Zeitreihe fließen dabei nicht ein. Bei einer *Prognose* wird davon ausgegangen, daß die Werte der exogenen Zeitreihe zum Prognosezeitpunkt nicht zur Verfügung stehen. Es wird daher auf die entsprechenden Vortageswerte zurückgegriffen. Hier fließt allerdings auch die CO_2-Konzentration vom Vortag mit in die Modelle ein.

Tabelle 4.4: Übersicht über die untersuchten Zeiträume

Monate	untersuchter Zeitraum	Fehlwerte
Januar/Februar	17.01.96-15.02.96	keine
März/April	16.03.96-15.04.96	26.03.96
Mai/Juni	17.05.96-15.06.96	keine
Juli/August	01.07.96-28.08.96	16.07.96-13.08.96
September/Oktober	16.09.96-15.10.96	keine
November/Dezember	16.11.96-16.12.96	01.12.96

Folgende jahreszeitliche Einflüsse werden erwartet:

Im *Sommer* entfallen die CO_2-Emissionen durch das Heizen fast völlig, werden aber evtl. in geringem Maße durch Kühlungsemissionen ausgeglichen. Die Vegetation ist voll entwickelt und durch regelmäßige Sonneneinstrahlung zu erhöhter Photosynthese-Leistung fähig. Es kann also von vergleichsweise niedrigen CO_2-Konzentrationen mit geringen Schwankungen ausgegangen werden.

Der *Winter* zeigt das genaue Gegenteil: Vermehrte CO_2-Emissionen während der Heizperiode und eine fast gänzliche Inaktivität der noch vorhandenen Vegetation, die somit auch ihre puffernde Funktion nicht mehr wahrnehmen kann. In dieser Zeit treten vermehrt hohe CO_2-Konzentrationen in der Luft auf, und lokale Emissionen spielen eine wichtigere Rolle.

[58] Valide bedeutet in diesem Zusammenhang, daß ein CO_2-Referenzwert existiert, an dem die Modellergebnisse gemessen werden könne

[59] Da die Datengrundlage nicht immer vollständig ist, mußten die Zeitabschnitte ggf. angeglichen werden. Für jeden Abschnitt existieren aber exakt 30 CO_2-Meßwerte, anhand derer die Güte der jeweiligen Modelle überprüft werden kann.

Frühling und *Herbst* sind als Perioden zunehmender bzw. abnehmender Vegetation zu betrachten, die noch nicht bzw. nicht mehr ihre volle Photosynthese-Leistung bringen. Lokale Emissionen durch Heizen sind höher als im Sommer, aber niedriger als im Winter zu erwarten.

4.2.4 Anzahl exogener Zeitreihen

Wie bereits in Kapitel 3.4 erwähnt, zeichnet sich ein Modell neben einer guten Analyse- bzw. Prognoseleistung auch durch eine geringe Zahl exogener Variablen, also Zeitreihen aus. Es wird versucht, eine Beschränkung auf die wichtigen vorzunehmen. In dem hier durchgeführten Vergleich der Verfahren wurden Modelle mit vier exogenen Zeitreihen eingesetzt. Dabei wurden alle Zeitreihen gleichwertig behandelt[60].

Um eine Vorauswahl für die multiple lineare Regression und das klassische Komponentenmodell zu treffen, wurden lineare Regressionen aller denkbaren Kombinationen von vier aus den 22 bzw. 23 möglichen exogenen Zeitreihen[61] exemplarisch für den Zeitraum vom 12.11.96 bis zum 31.12.96 durchgeführt. Die verschiedenen Läufe wurden anschließend nach RMSE sortiert. Für jede exogene Zeitreihe konnte nun die Position des ersten Auftretens in den sortierten Listen bzw. die durchschnittliche Position (Index) bestimmt werden. Auf diese Weise wurden die 14 einflußreichsten Zeitreihen als exogene Zeitreihen für den Vergleich ausgewählt. Die möglichen Kombinationen wurden so von 7315 auf 1001 reduziert. Tabelle 4.5 zeigt eine Übersicht über die Ergebnisse des Verfahrens.

[60] Dies ist nicht unbedingt selbstverständlich. Es ist auch denkbar, eine Sinus- *und* die entsprechende Cosinus-Transformation gemeinsam als *eine* Zeitreihe betrachten.

[61] Für die Prognose wurden die zeitlich zurückliegenden Werte der CO_2-Zeitreihe hinzugenommen.

Tabelle 4.5: Übersicht über den Einfluß exogener Zeitreihen auf die Regression

	Nach RMSE				Nach MAPE				Summe
Zeitreihe	1. Auftreten	Rang	Index	Rang	1. Auftreten	Rang	Index	Rang	**der Ränge**
WG	1	1	347.07	1	1	1	309.03	1	4
NDS_DWD	1	1	565.85	7	1	1	583.07	8	17
THT_FH	1	1	565.45	6	6	9	576.1	6	22
WR	1	1	646.87	11	1	1	639.31	11	24
TMP_DWD	15	14	568.64	8	1	1	545.22	2	25
WRSIN	3	6	613.53	9	4	7	560.15	4	26
THT_DWD	4	7	559.12	5	13	13	564.1	5	30
WRGSIN	5	8	645.74	10	3	6	580.32	7	31
WRG	2	5	667.03	12	2	5	652.88	12	34
TMP_FH	32	16	548.79	4	11	12	545.97	3	35
RLF	9	9	701.32	13	5	8	712.33	13	43
JTG_COS	15	13	538.05	3	147	19	613.06	10	45
HAUDE	11	10	739.16	14	8	11	745.59	15	50
TG	96	18	515.48	2	599	22	590.09	9	51
SDEF	13	12	743.46	15	7	10	748.99	17	54
JTG_SIN	12	11	782.34	18	15	14	797.53	18	61
SSD	16	15	751.22	17	19	15	746.5	16	63
NDS_FH	111	19	746.54	16	119	18	714.08	14	67
JTG	77	17	798.42	19	31	16	817.63	19	71
WRCOS	123	20	826.67	20	103	17	843.54	20	77
WRGCOS	194	22	848.31	21	202	20	862.74	21	84
BG	172	21	912.95	22	251	21	883.78	22	86

Unter Berücksichtigung der langen Rechenzeiten für ein Netztraining war es im Rahmen dieser Arbeit nicht möglich, ein ähnlich aufwendiges Verfahren zur Identifizierung günstiger Kombinationen exogener Zeitreihen zu finden. Aufgrund dieser Tatsache und aus der Überlegung der prinzipiellen Überlegenheit Neuronaler Netze und des Komponentenmodells gegenüber der Regression[62] wurde für diese Verfahren die jeweils beste Kombination der Regression genutzt.

Für die Neuronalen Netze beschränkt sich die Optimierung also auf die Netzarchitektur. Variationsmöglichkeiten bestanden insbesondere bei der Wahl der Lernfunktion (BP_momentum, BPM_ln_cosh[63] und Rprop), bei der Modifikation der Architektur (keine, Vorschicht, Shortcuts oder Vorschicht und Shortcuts) sowie bei der Skalierung der Zeitreihen (LIN, SIG4, SIG4CUT, SIG4CUT0).

4.2.5 Wahl der Netzstruktur

Neben der Auswahl geeigneter Einflußzeitreihen hat bei den Neuronalen Netzen auch die Auswahl einer günstigen Netztopologie einen wesentlichen Einfluß auf die Prognose- und Analysegüte. Ausgehend von einer Grundkonfiguration werden verschiedene Komponenten variiert und die möglichen Kombinationen auf ihre Güte getestet. Die Einstellungen, die für alle Läufe gleich waren (Grundkonfiguration), werden in Tabelle 4.6 aufgeführt. Ziel ist es, aufgrund der exemplarischen Ergebnisse mit allen vernünftig

[62] Vgl. Kapitel 3.1.2 und Kapitel 4.2.2.

[63] Vgl. [Thi 98], 33 ff.

erscheinenden Parametern eine Einschränkung der freien Variablen für Strukturfindung bei der tatsächlichen Analyse bzw. Prognose vorzunehmen.

Tabelle 4.6: Grundkonfiguration zur Bestimmung der günstigsten Netzparameter

Parameter	Konfiguration
Zeitraum	12.11.1996 - 31.12.1996
Maximale Epochenzahl	1000
Minimale Epochenzahl	100
Abbruch bei Abweichung vom Minimum	20%
Reihenfolge der Pattern beim Lernen	zufällig
Updatefunktion	topologisch
Initialisierung der Gewichte	zufällig [-0,5 .. 0,5]
Netzarchitektur	Feed Forward
Anzahl Neuronen in der Ausgabeschicht	1
Testpaare	30%
Auswahl der Testpaare	zufällig
Validierungspaare	50
Exogene Zeitreihe 1	Windgeschwindigkeit
Exogene Zeitreihe 2	Windrichtung
Exogene Zeitreihe 3	Tageshöchsttemperatur (FH)
Exogene Zeitreihe 4	Niederschlag (DWD)

Folgende vier Komponenten wurden variiert:

- **Lernfunktion**

  Die Standardlernfunktion *BackpropMomentum*[64] wurde mit *Resilent Propagation* (Rprop)[65] und *BackpropMomentum mit ln(cosh())* als Fehlerfunktion (BPM_ln_cosh)[66] verglichen.

- **Skalierungsfunktion**

  Für die Verwendung einer Zeitreihe im Rahmen eines neuronalen Modells ist die Skalierung auf den verwendeten Wertebereich (hier:]-1..+1[) sinnvoll. Zum Einsatz kamen die *lineare Skalierung* (LIN), sowie die μ-σ-*Skalierungen* MY_SIGMA4, MY_SIGMA4CUT und MY_SIGMA4CUT0[67].

[64] Eine ausführliche Darstellung der Lernfunktion findet sich bei [Zel 94], 115 und bei [Reh 94], 43 ff.

[65] Vgl. [Zel 94], 392.

[66] Die Funktion wurde im Rahmen der Arbeit von F.M. Thiesing in das System integriert. Entsprechende Erläuterungen finden sich bei [Thi 98], 33 ff.

[67] Eine Erklärung der einzelnen Skalierungsfunktionen ist bei [Thi 98], 113 ff. zu finden.

- **Modifikationen der Topologie**

  Neben der *Standardtopologie* (Feed Forward Multi Layer Perzeptron) wurden *Shortcut-Verbindungen* von der Input- zur Output-Schicht (shortC) verwendet sowie die *Vorschicht* (vorS) und eine *Kombination aus Vorschicht und SC-Verbindungen* (von der Hidden-Schicht zur Output-Schicht). Eine Erläuterung der Modifikationen findet sich in Abschnitt 3.2.1.3.

- **Anzahl der Neuronen in der verborgenen Schicht**

  Für die Hidden-Schicht wurden 3 bis 10 Neuronen getestet.

Die fünf günstigsten Parametereinstellungen werden in Tabelle 4.7 (Analyse) und Tabelle 4.8 (Prognose) aufgeführt. Eine Gesamtübersicht über die Läufe findet sich im Anhang (Kapitel 7.1 und 7.2.1).

Tabelle 4.7: Die 5 günstigsten Topologien für die Analyse (nach RMSE)

Run	mape	mre	sres	rmse	theil	Learnfunction	Scalefunction	Topology_Modification
25	1.8411	0.0184	3335.4	8.6093	0.5704	BP_momentum	linear	vorS+shortC
27	1.8095	0.0182	3369.8	8.6536	0.5745	BP_momentum	linear	vorS
5	1.7398	0.0175	3449.5	8.7553	0.5820	Rprop	linear	vorS+shortC
46	1.7793	0.0178	3472.0	8.7838	0.5825	BPM_ln_cosh	linear	none
48	1.7793	0.0178	3472.0	8.7838	0.5825	BPM_ln_cosh	linear	shortC

Tabelle 4.8: Die 5 günstigsten Topologien für die Prognose (nach RMSE)

Run	mape	mre	sres	rmse	theil	Learnfunction	Scalefunction	Topology_Modification
38	2.8852	0.0289	7648.8	13.3370	0.8673	BP_momentum	my_sigma	none
40	2.8852	0.0289	7648.8	13.3370	0.8673	BP_momentum	my_sigma	shortC
34	2.8415	0.0284	7766.7	13.4400	0.8740	BP_momentum	my_sigma0	none
36	2.8415	0.0284	7766.7	13.4400	0.8740	BP_momentum	my_sigma0	shortC
58	2.8849	0.0288	7786.5	13.4570	0.8751	BPM_ln_cosh	my_sigma	none

Aufgrund der Ergebnisse wurden Lernfunktionen auf *BP_momentum* und *BPM_ln_cosh* beschränkt und die Skalierungsfunktionen auf *linear* und μ-σ-*Skalierung*. Für die Anzahl der Neuronen in der Hidden-Schicht wurden 3, 5, 7 und 10 ausgewählt und als Topologie-Modifikationen blieben *keine*, *Vorschicht* und *Vorschicht mit Shortcuts.* Somit waren für jede Analyse / Prognose nur noch 48 Netz-Trainingsläufe mit einem geschätzten Zeitaufwand von jeweils ca. 16 Stunden nötig.

4.3 Vergleich der Verfahren

Bei den betrachteten sechs Zeiträumen und der Unterscheidung von Analyse und Prognose werden insgesamt 12 Vergleiche durchgeführt. Jeder Vergleich umfaßt einen Vergleich der Fehler (nach RMSE, in Kapitel 3.3 erläutert). Eine Sonderstellung hat hier das ARIMA-Modell, welches nur für den Fall eines zusammenhängenden Prognosezeitraums mit ausreichend Werten davor eingesetzt werden kann.

Für die Regressionsmodelle werden die bei 4 aus 15 exogenen Zeitreihen möglichen 1001 Kombinationen getestet und die beste für den Vergleich verwandt. Die so gefunden Kombinationen werden auch von den anderen Verfahren (SNNS, Komponentenmodell) verwandt.

4.3.1 Analyse

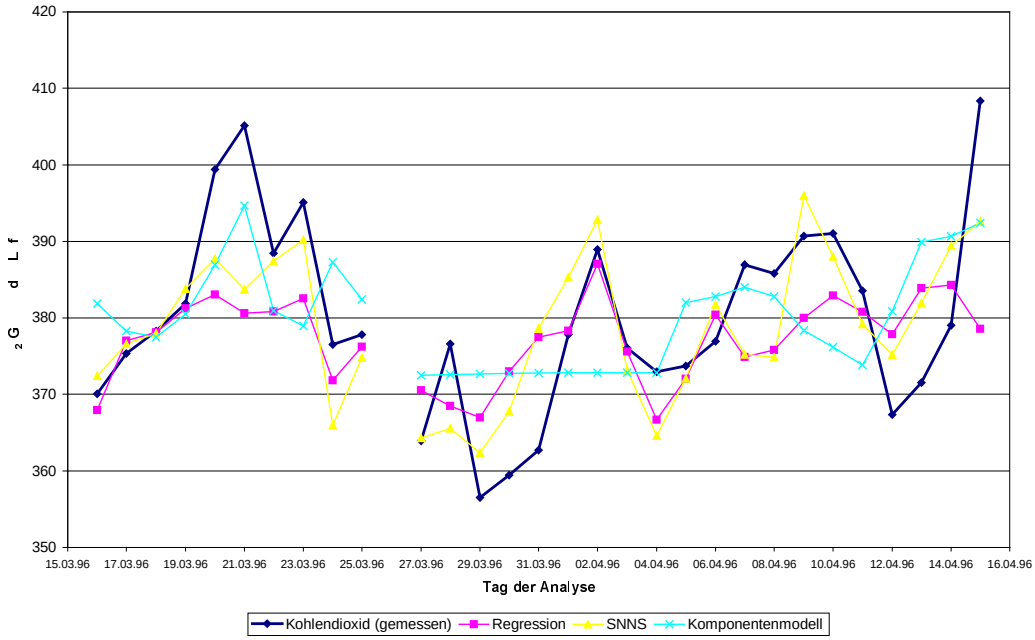

Abbildung 4.6: Vergleich der Analysegüte der drei Verfahren (Zeitraum März / April 1996)

4.3.1.1 Januar / Februar

Die fünf günstigsten Kombinationen exogener Zeitreihen für den Analysezeitraum[68] können Tabelle 4.9 entnommen werden.

Tabelle 4.9: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Januar/Februar 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
902	2.0443	0.0206	3437.2	10.704	0.6335	TMP_DWD THT_DWD WG WR
810	2.0002	0.0200	3440.7	10.709	0.6338	TG TMP_DWD WG WR
903	1.9694	0.0199	3471.9	10.758	0.6367	TMP_DWD THT_DWD WG WRSIN
221	2.0107	0.0203	3474.7	10.762	0.6369	HAUDE TMP_DWD WG WR
811	1.9359	0.0194	3475.7	10.764	0.6370	TG TMP_DWD WG WRSIN

[68] Die günstigsten Kombinationen beziehen sich auf die Regression.

Tabelle 4.10: Die fünf günstigsten Netzarchitekturen für die Analyse (Januar/Februar 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
25	1.9770	0.0196	2534.3	9.1912	0.544	BPM_ln_cosh	linear	7	none
28	1.7213	0.0172	2623.8	9.3520	0.5535	BPM_ln_cosh	linear	3	vorS+shortC
27	1.8825	0.0187	2769.8	9.6087	0.5687	BPM_ln_cosh	linear	10	vorS
5	1.8426	0.0184	2962.1	9.9366	0.5881	BP_momentum	linear	5	vorS+shortC
31	1.9492	0.0192	2980.9	9.9681	0.5899	BPM_ln_cosh	linear	7	vorS+shortC

Tabelle 4.11: Gegenüberstellung der Modellergebnisse (Januar/Februar 1996)

Modell	mape	rmse
SNNS	1.9770	9.191
Komponentenmodell	2.0102	9.350
Regression	2.0443	10.704

4.3.1.2 März / April

Tabelle 4.12: Günstigste Kombinationen exogener Zeitreihen für die Analyse (März/April 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
193	2.0911	0.0211	3403.1	10.651	0.9713	HAUDE TG WG WR
195	2.1394	0.0216	3485.9	10.780	0.9822	HAUDE TG WG WRG
194	2.1571	0.0218	3573.5	10.914	0.9806	HAUDE TG WG WRSIN
196	2.1882	0.0221	3688.2	11.088	0.9962	HAUDE TG WG WRGSIN
698	2.2067	0.0224	3887.3	11.383	1.0438	RLF TG WG WR

Tabelle 4.13: Die fünf günstigsten Netzarchitekturen für die Analyse (März/April 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
27	1.9412	0.0194	2181.4	8.6730	0.7791	BPM_ln_cosh	linear	10	vorS
28	1.7794	0.0178	2183	8.6762	0.7835	BPM_ln_cosh	linear	3	vorS+shortC
4	1.7725	0.0177	2201.7	8.7133	0.7868	BP_momentum	linear	3	vorS+shortC
13	1.9057	0.019	2226	8.7612	0.7908	BP_momentum	μ-σ	7	none
25	1.8211	0.0181	2267.6	8.8428	0.7927	BPM_ln_cosh	linear	7	none

Tabelle 4.14: Gegenüberstellung der Modellergebnisse (März/April 1996)

Modell	mape	rmse
SNNS	1.9412	8.673
Komponentenmodell	1.8870	10.264
Regression	2.0911	10.651

4.3.1.3 Mai / Juni

Tabelle 4.15: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Mai/Juni 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
666	1.4651	0.0147	1744.2	7.625	0.6099	NDS_DWD WG WRSIN WRGSIN
1000	1.5503	0.0156	1920.9	8.002	0.6400	WG WRSIN WRG WRGSIN
998	1.5525	0.0156	1926.7	8.014	0.6410	WG WR WRSIN WRGSIN
586	1.6513	0.0165	1963.2	8.089	0.6470	NDS_DWD TG WRSIN WRGSIN
665	1.5225	0.0153	1984.1	8.132	0.6504	NDS_DWD WG WRSIN WRG

Tabelle 4.16: Die fünf günstigsten Netzarchitekturen für die Analyse (Mai/Juni 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
12	1.5563	0.0156	1585.6	7.2699	0.5814	BP_momentum	linear	5	none
34	1.4839	0.0148	1604	7.3121	0.5848	BPM_ln_cosh	linear	7	vorS
9	1.5825	0.0159	1606.2	7.3171	0.5852	BP_momentum	linear	10	none
23	1.5549	0.0155	1640.3	7.3943	0.5914	BP_momentum	μ-σ	3	none
11	1.6413	0.0165	1703.5	7.5354	0.6027	BP_momentum	linear	3	none

Tabelle 4.17: Gegenüberstellung der Modellergebnisse (Mai/Juni 1996)

Modell	mape	rmse
SNNS	1.5563	7.269
Komponentenmodell	1.4662	7.561
Regression	1.4651	7.625

4.3.1.4 Juli / August

Tabelle 4.18: Günstigste Kombinationen exogener Zeitreihen für die Analyse (Juli/August 1996)

Run	mape	mre	sres	rmse	theil	Variables
413	1.4584	0.0145	1371.9	6.763	0.7542	JTG_COS TG WG WR
301	1.5033	0.0150	1377.6	6.777	0.7438	JTG_COS NDS_DWD TG WG
415	1.4670	0.0146	1393.0	6.814	0.7578	JTG_COS TG WG WRG
346	1.5039	0.0150	1403.5	6.840	0.7528	JTG_COS RLF TG WG
414	1.4873	0.0148	1419.5	6.879	0.7519	JTG_COS TG WG WRSIN

Tabelle 4.19: Die fünf günstigsten Netzarchitekturen für die Analyse (Juli/August 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
3	1.5851	0.0157	1642.4	7.3991	0.8384	BP_momentum	linear	10	vorS
36	1.7182	0.0171	1647.8	7.4113	0.8388	BPM_ln_cosh	linear	5	none
29	1.6863	0.0168	1666.5	7.4531	0.8411	BPM_ln_cosh	linear	5	vorS+shortC
28	1.7662	0.0175	1708.8	7.5471	0.8486	BPM_ln_cosh	linear	3	vorS+shortC
12	1.7503	0.0174	1712.9	7.5562	0.8573	BP_momentum	linear	5	none

Tabelle 4.20: Gegenüberstellung der Modellergebnisse (Juli/August 1996)

Modell	mape	rmse
Regression	1.4584	6.763
Komponentenmodell	1.4844	6.802
SNNS	1.5851	7.399

4.3.1.5 September / Oktober

Tabelle 4.21: Günstigste Kombinationen exogener Zeitreihen für die Analyse (September/Oktober 1996)

Run	mape	mre	sres	rmse	theil	Variables
301	2.9175	0.0296	7300.8	15.600	0.8471	JTG_COS NDS_DWD TG WG
346	2.8857	0.0293	7464.6	15.774	0.8566	JTG_COS RLF TG WG
26	2.9490	0.0299	7519.3	15.832	0.8597	HAUDE JTG_COS TG WG
413	3.1626	0.0320	7560.2	15.875	0.8620	JTG_COS TG WG WR
415	3.1040	0.0314	7560.5	15.875	0.8621	JTG_COS TG WG WRG

Tabelle 4.22: Die fünf günstigsten Netzarchitekturen für die Analyse (September/Oktober 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
26	2.3275	0.0233	4166.8	11.785	0.6400	BPM_ln_cosh	linear	10	vorS+shortC
28	2.3929	0.0240	4316.8	11.996	0.6514	BPM_ln_cosh	linear	3	vorS+shortC
23	2.8355	0.0282	4530.1	12.288	0.6673	BP_momentum	μ-σ	3	none
30	2.4488	0.0246	4613.4	12.401	0.6734	BPM_ln_cosh	linear	3	vorS
32	2.8082	0.0280	4809.2	12.661	0.6875	BPM_ln_cosh	linear	5	vorS

Tabelle 4.23: Gegenüberstellung der Modellergebnisse (September/Oktober 1996)

Modell	mape	rmse
SNNS	2.3275	11.785
Komponentenmodell	2.5517	13.414
Regression	2.9175	15.600

4.3.1.6 November / Dezember

Tabelle 4.24: Günstigste Kombinationen exogener Zeitreihen für die Analyse (November/Dezember 1996)

Run	mape	mre	sres	rmse	theil	Variables
332	1.9653	0.0196	2535.7	9.194	0.6641	JTG_COS NDS_DWD WG WR
652	1.8676	0.0187	2547.9	9.216	0.6717	NDS_DWD THT_FH WG WR
334	1.9806	0.0197	2567.3	9.251	0.6680	JTG_COS NDS_DWD WG WRG
654	1.8886	0.0189	2582.8	9.279	0.6761	NDS_DWD THT_FH WG WRG
333	1.9918	0.0199	2588.7	9.289	0.6721	JTG_COS NDS_DWD WG WRSIN

Tabelle 4.25: Die fünf günstigsten Netzarchitekturen für die Analyse (November/Dezember 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
25	1.8194	0.0181	2227.6	8.617	0.6267	BPM_ln_cosh	linear	7	none
36	1.9211	0.0191	2375.9	8.899	0.6458	BPM_ln_cosh	linear	5	none
2	1.9271	0.0192	2477.6	9.088	0.6540	BP_momentum	linear	10	vorS+shortC
33	1.9985	0.0198	2497.1	9.123	0.6605	BPM_ln_cosh	linear	10	none
35	1.9942	0.0198	2573.0	9.261	0.6698	BPM_ln_cosh	linear	3	none

Tabelle 4.26: Gegenüberstellung der Modellergebnisse (November/Dezember 1996)

Modell	mape	rmse
Komponentenmodell	1.8923	7.999
SNNS	1.8194	8.617
Regression	1.9653	9.194

4.3.2 Prognose

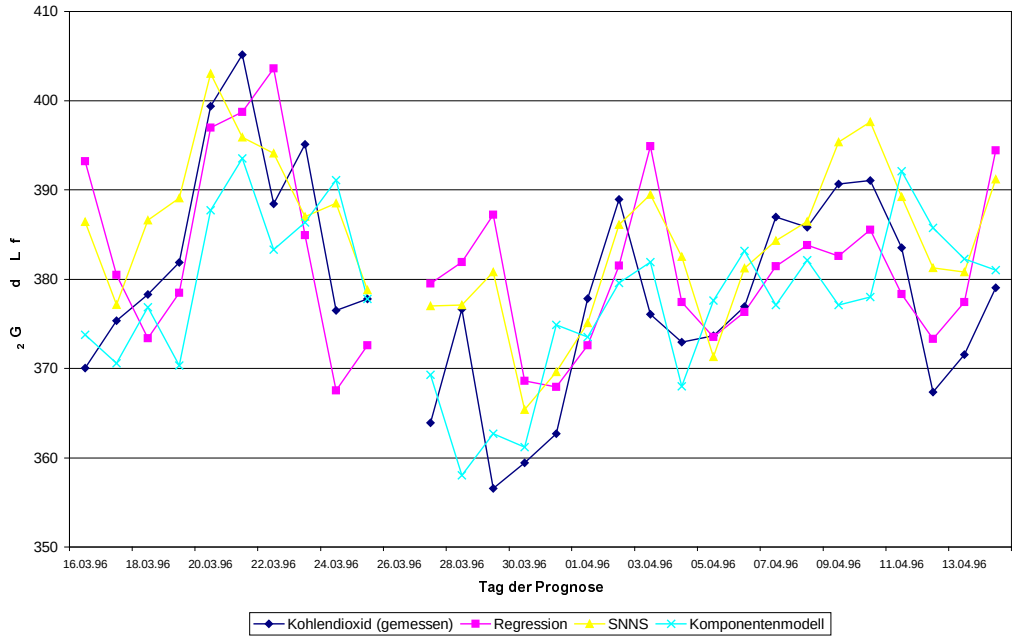

Bild 4.1: Vergleich der Prognosegüte der drei Verfahren (Zeitraum März / April 1996)

4.3.2.1 Januar / Februar

Die fünf günstigsten Kombinationen[69] exogener Zeitreihen für den Prognoseezeitraum können Tabelle 4.9 entnommen werden.

Tabelle 4.27: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Januar/Februar 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
710	2.2844	0.0228	5690.8	13.773	0.8151	RLF TMP_DWD TMP_FH WG
95	2.3152	0.0231	5722.3	13.811	0.8174	HAUDE NDS_DWD TMP_FH THT_FH
884	2.3021	0.0230	5727.3	13.817	0.8177	TMP_DWD TMP_FH THT_FH WRSIN
709	2.3223	0.0232	5735.2	13.827	0.8183	RLF TMP_DWD TMP_FH THT_FH
886	2.3112	0.0231	5735.2	13.827	0.8183	TMP_DWD TMP_FH THT_FH WRGSIN

Tabelle 4.28: Die fünf günstigsten Netzarchitekturen für die Prognose (Januar/Februar 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
7	2.3158	0.0231	5298.4	13.290	0.7865	BP_momentum	linear	7	vorS+shortC
25	2.3077	0.0231	5303.8	13.296	0.7869	BPM_ln_cosh	linear	7	none
31	2.3101	0.0231	5317.1	13.313	0.7879	BPM_ln_cosh	linear	7	vorS+shortC
29	2.2803	0.0228	5346	13.349	0.790	BPM_ln_cosh	linear	5	vorS+shortC
30	2.2198	0.0223	5378	13.389	0.7924	BPM_ln_cosh	linear	3	vorS

[69] Die günstigsten Kombinationen beziehen sich auf die Regression.

Tabelle 4.29: Gegenüberstellung der Modellergebnisse (Januar/Februar 1996)

Modell	mape	rmse
SNNS	2.3158	13.290
Regression	2.2844	13.773
Komponentenmodell	2.2842	13.782

4.3.2.2 März / April

Tabelle 4.30: Günstigste Kombinationen exogener Zeitreihen für die Prognose (März/April 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
682	2.0769	0.0209	2838.4	10.068	0.8934	RLF TG TMP_FH WG
693	2.0592	0.0207	2857.3	10.102	0.8963	RLF TG THT_FH WG
832	2.0672	0.0208	2875.8	10.134	0.8992	TG TMP_FH WG WRSIN
688	2.0829	0.0210	2886.9	10.154	0.9010	RLF TG THT_DWD WG
675	2.0751	0.0209	2887.6	10.155	0.9011	RLF TG TMP_DWD WG

Tabelle 4.31: Die fünf günstigsten Netzarchitekturen für die Prognose (März/April 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
45	1.9260	0.0192	2441.1	9.3372	0.8285	BPM_ln_cosh	μ-σ	10	none
12	1.9274	0.0192	2452.1	9.3581	0.8304	BP_momentum	linear	5	none
27	1.8731	0.0187	2479.8	9.4109	0.8350	BPM_ln_cosh	linear	10	vorS
48	1.8908	0.0190	2491.8	9.4335	0.8371	BPM_ln_cosh	μ-σ	5	none
25	1.9263	0.0192	2540.1	9.5247	0.8451	BPM_ln_cosh	linear	7	none

Tabelle 4.32: Gegenüberstellung der Modellergebnisse (März/April 1996)

Modell	mape	rmse
SNNS	1.9260	9.337
Regression	2.0769	10.068
Komponentenmodell	2.0911	10.439

4.3.2.3 Mai / Juni

Tabelle 4.33: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Mai/Juni 1996)

Run	mape	mre	sres	rmse	theil	Exogene Variablen
205	1.7549	0.0178	3200.4	10.329	0.8261	HAUDE TMP_DWD TMP_FH WG
167	1.7612	0.0178	3269.7	10.44	0.8350	HAUDE TG TMP_DWD TMP_FH
177	1.7810	0.018	3310.9	10.505	0.8402	HAUDE TG TMP_FH WG
203	1.7891	0.0181	3315.7	10.513	0.8408	HAUDE TMP_DWD TMP_FH THT_DWD
232	1.8066	0.0183	3324.8	10.527	0.8420	HAUDE TMP_FH THT_DWD WG

Tabelle 4.34: Die fünf günstigsten Netzarchitekturen für die Prognose (Mai/Juni 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
11	1.6998	0.0172	3061.3	10.102	0.8079	BP_momentum	linear	3	none
33	1.7482	0.0177	3083.9	10.139	0.8109	BPM_ln_cosh	linear	10	none
25	1.7114	0.0174	3088.1	10.146	0.8115	BPM_ln_cosh	linear	7	none
32	1.7664	0.0178	3090.7	10.150	0.8118	BPM_ln_cosh	linear	5	vorS
12	1.7423	0.0177	3128.3	10.212	0.8167	BP_momentum	linear	5	none

Tabelle 4.35: Gegenüberstellung der Modellergebnisse (Mai/Juni 1996)

Modell	mape	rmse
Komponentenmodell	1.7082	10.098
SNNS	1.6998	10.102
Regression	1.7549	10.329

4.3.2.4 Juli / August

Tabelle 4.36: Günstigste Kombinationen exogener Zeitreihen für die Prognose (Juli/August 1996)

Run	mape	mre	sres	rmse	theil	Variables
205	1.5732	0.0158	1540.3	7.2880	0.826	HAUDE TMP_DWD TMP_FH WG
794	1.6056	0.0161	1545.9	7.3012	0.8275	TG TMP_DWD TMP_FH WG
167	1.5497	0.0155	1546.3	7.3022	0.8276	HAUDE TG TMP_DWD TMP_FH
793	1.5550	0.0156	1548.2	7.3066	0.8281	TG TMP_DWD TMP_FH THT_FH
797	1.5830	0.0159	1559.2	7.3324	0.8311	TG TMP_DWD TMP_FH WRG

Tabelle 4.37: Die fünf günstigsten Netzarchitekturen für die Prognose (Juli/August 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
9	1.5587	0.0156	1460.5	7.0965	0.8043	BP_momentum	linear	10	none
33	1.5013	0.0151	1528.5	7.2600	0.8228	BPM_ln_cosh	linear	10	none
30	1.5815	0.0158	1538.5	7.2837	0.8255	BPM_ln_cosh	linear	3	vorS
35	1.5177	0.0153	1596.8	7.4203	0.8410	BPM_ln_cosh	linear	3	none
27	1.5693	0.0158	1602.1	7.4328	0.8424	BPM_ln_cosh	linear	10	vorS

Tabelle 4.38: Gegenüberstellung der Modellergebnisse (Juli/August 1996)

Modell	mape	rmse
SNNS	1.5587	7.0965
Komponentenmodell	1.5665	7.2435
Regression	1.5732	7.2880

4.3.2.5 September / Oktober

Tabelle 4.39: Günstigste Kombinationen exogener Zeitreihen für die Prognose (September/Oktober 1996)

Run	mape	mre	sres	rmse	theil	Variables
794	3.0393	0.0305	7320.1	15.621	0.8482	TG TMP_DWD TMP_FH WG
826	3.1179	0.0313	7432.3	15.74	0.8547	TG TMP_FH THT_FH WG
798	3.2475	0.0326	7455.9	15.765	0.8561	TG TMP_DWD TMP_FH WRGSIN
895	3.249	0.0328	7495.1	15.806	0.8583	TMP_DWD TMP_FH WRSIN WRGSIN
890	3.1397	0.0317	7497.3	15.809	0.8584	TMP_DWD TMP_FH WG WRGSIN

Tabelle 4.40: Die fünf günstigsten Netzarchitekturen für die Prognose (September/Oktober 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
11	2.9996	0.0298	5356.8	13.363	0.7256	BP_momentum	linear	3	none
3	3.0307	0.0303	6108.3	14.269	0.7749	BP_momentum	linear	10	vorS
1	2.9082	0.0292	6529.6	14.753	0.8011	BP_momentum	linear	7	none
9	3.0955	0.0308	6641.4	14.879	0.808	BP_momentum	linear	10	none
25	2.9767	0.0299	6684.9	14.927	0.8106	BPM_ln_cosh	linear	7	none

Tabelle 4.41: Gegenüberstellung der Modellergebnisse (September/Oktober 1996)

Modell	mape	rmse
SNNS	2.9996	13.363
Komponentenmodell	2.8973	14.012
Regression	3.0393	15.621

4.3.2.6 November / Dezember

Tabelle 4.42: Günstigste Kombinationen exogener Zeitreihen für die Prognose (November/Dezember 1996)

Run	mape	mre	sres	rmse	theil	Variables
348	2.4604	0.0245	4424.8	12.352	0.8884	JTG_COS RLF TG WRSIN
350	2.458	0.0245	4426.7	12.355	0.8886	JTG_COS RLF TG WRGSIN
386	2.4769	0.0248	4435.8	12.368	0.8895	JTG_COS RLF WRG WRGSIN
66	2.4711	0.0247	4444.0	12.379	0.8903	HAUDE JTG_COS WRG WRGSIN
422	2.4642	0.0245	4446.0	12.382	0.8905	JTG_COS TG WRG WRGSIN

Tabelle 4.43: Die fünf günstigsten Netzarchitekturen für die Prognose (November/Dezember 1996)

Run	mape	mre	sres	rmse	theil	Lernfunktion	Skal.-Fkt.	Hidden	Topologie Mod.
45	2.5803	0.0256	4565.1	12.547	0.9024	BPM_ln_cosh	μ-σ	10	none
37	2.5845	0.0257	4583.5	12.572	0.9042	BPM_ln_cosh	μ-σ	7	none
48	2.597	0.0258	4664.1	12.682	0.9121	BPM_ln_cosh	μ-σ	5	none
35	2.4982	0.0248	4672.1	12.693	0.9129	BPM_ln_cosh	linear	3	none
28	2.4989	0.0248	4680.7	12.705	0.9137	BPM_ln_cosh	linear	3	vorS+shortC

Tabelle 4.44: Gegenüberstellung der Modellergebnisse (November/Dezember 1996)

Modell	mape	rmse
Komponentenmodell	2.3903	11.986
Regression	2.4604	12.352
SNNS	2.5803	12.547

4.3.3 ARIMA-Modellierung

Aufgrund ihres autoregressiven und des Moving-Average-Prozesses kann die ARIMA-Modellierung nur für Datenreihen ohne Fehlwerte verwendet werden. Selbst einzelne „missing values“, die bei naturwissenschaftlichen Datenreihen oftmals zu finden sind, schließen eine zuverlässige und aussagekräftige ARIMA-Modellierung aus. Aus diesem Grund wird der Vergleich dieser Art von Modellen mit den Neuronalen Netzen auf den längsten, vollständigen Abschnitten der CO_2-Zeitreihe durchgeführt. Dies ist zum einen der Zeitraum vom 19.11.94 bis zum 10.12.95 mit 387 Tagen und zum anderen die 178 Tage zwischen dem 24.05.94 und dem 17.11.95. Als Vergleich wird jeweils eine Prognose durchgeführt, die die letzten 30 Tage der Zeiträume umfaßt.

Für die multivariaten Modelle SNNS und Regression müssen noch die Inputzeitreihen festgelegt werden. Das neuronale Netz erhält als Input CO2 (Vortageswert), WG, TMP_DWD, NDS_DWD und WR_SIN. Die Regression erhält ebenfalls den CO2-

Vortageswert und analog zu dem in Kapitel 4.2.1 vorgestellten Vorgehen vier von 11 möglichen exogenen Zeitreihen[70].

4.3.3.1 Zeitraum 1 (19.11.94 bis 10.12.95, 387 Tage)

Die Autokorrelationen und partiellen Autokorrelationen identifizieren die CO_2-Zeitreihe auf diesem Abschnitt als einen ARIMA[1,0,0]-Prozeß. In Abbildung 4.7 werden Autokorrelation und partielle Autokorrelation des undifferenzierten Zeitreihenabschnitts gezeigt.

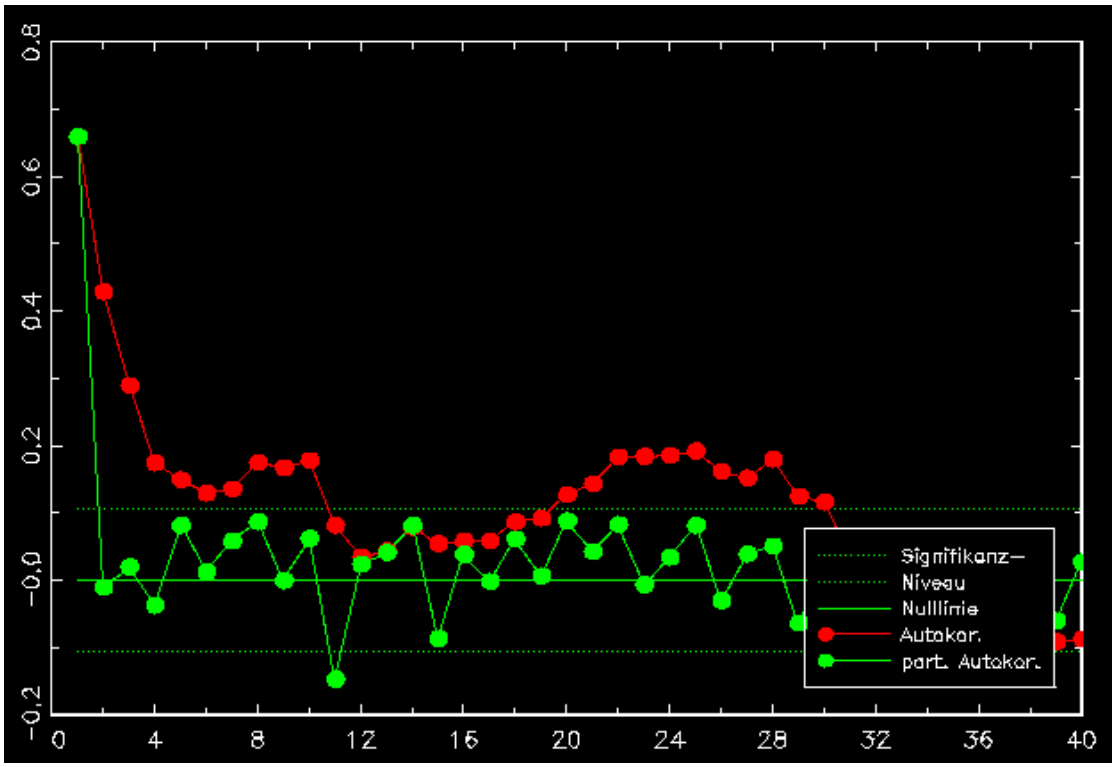

Abbildung 4.7: (Partielle) Autokorrelationen zur undifferenzierten CO_2-Zeitreihe (19.11.94 bis 10.12.95)

Als günstigste Netzarchitektur auf dem untersuchten Intervall erwies sich ein FF-MLP mit BPM_ln_cosh als Lernfunktion, linearer Skalierung der Zeitreihen, 7 Neuronen in der Hiddenschicht und ohne Vorschicht oder Shortcuts. Auf diesem Abschnitt sind die multivariaten Verfahren (SNNS, Regression) dem ARIMA-Modell deutlich überlegen. Eine Gegenüberstellung der Modellfehler findet sich in Tabelle 4.45.

[70] Um Rechenzeit zu sparen und da der Vergleich ARIMA / SNNS im Vordergrund stand, wurde die Anzahl der möglichen exogenen Inputs verringert (Siehe Kapitel 4.2.1). Die Beschränkung auf 11 Zeitreihen ergab eine Reduzierung von 1001 auf 330 Regressionsläufe.

Tabelle 4.45: Vergleich der Prognosefehler der einzelnen Verfahren

Verfahren	MAPE	RMSE
SNNS	2.6502	13.4593
Regression	2.7535	14.4789
ARIMA[1,0,0]	3.0123	15.9489

Abbildung 4.8 zeigt die CO_2-Zeitreihe und die Modellreihen im Vergleich. Es ist deutlich zu erkennen, daß die Modellzeitreihen dem wahren Wert „hinterherlaufen“.

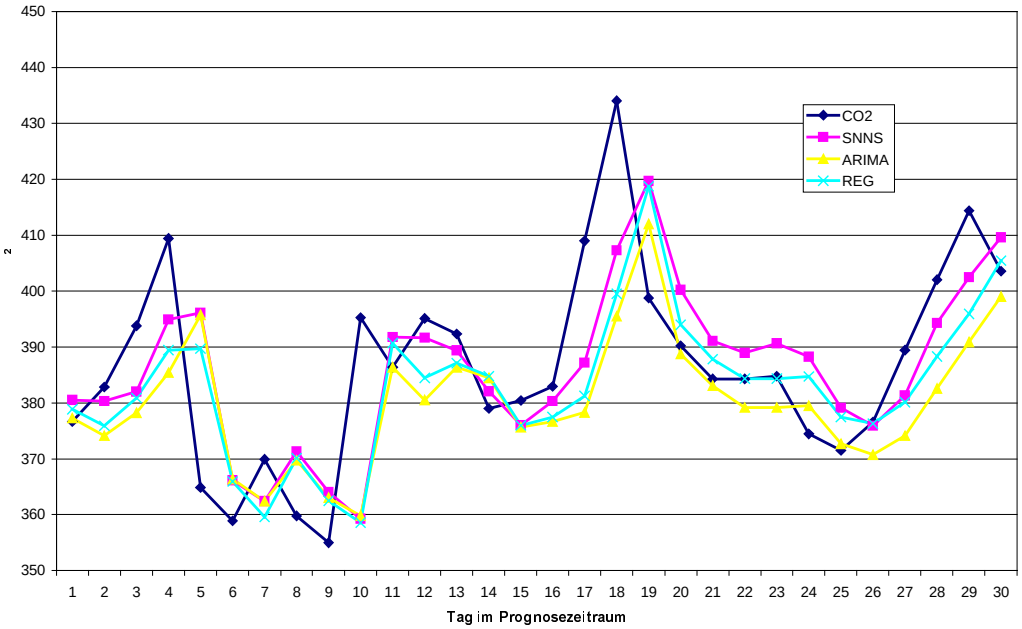

Abbildung 4.8: Vergleich der Verfahren im dem Prognosezeitraum 1 (19.11.94 bis 10.12.95)

4.3.3.2 Zeitraum 2 (24.05.94 bis 17.11.95, 178 Tage)

Auch dieser Abschnitt der CO_2-Zeitreihe wird mittels der Autokorrelationen und partiellen Autokorrelationen als ein ARIMA[1,0,0]-Prozeß identifiziert. Abbildung 4.9 zeigt die Autokorrelation und partielle Autokorrelation des undifferenzierten Zeitreihenabschnitts.

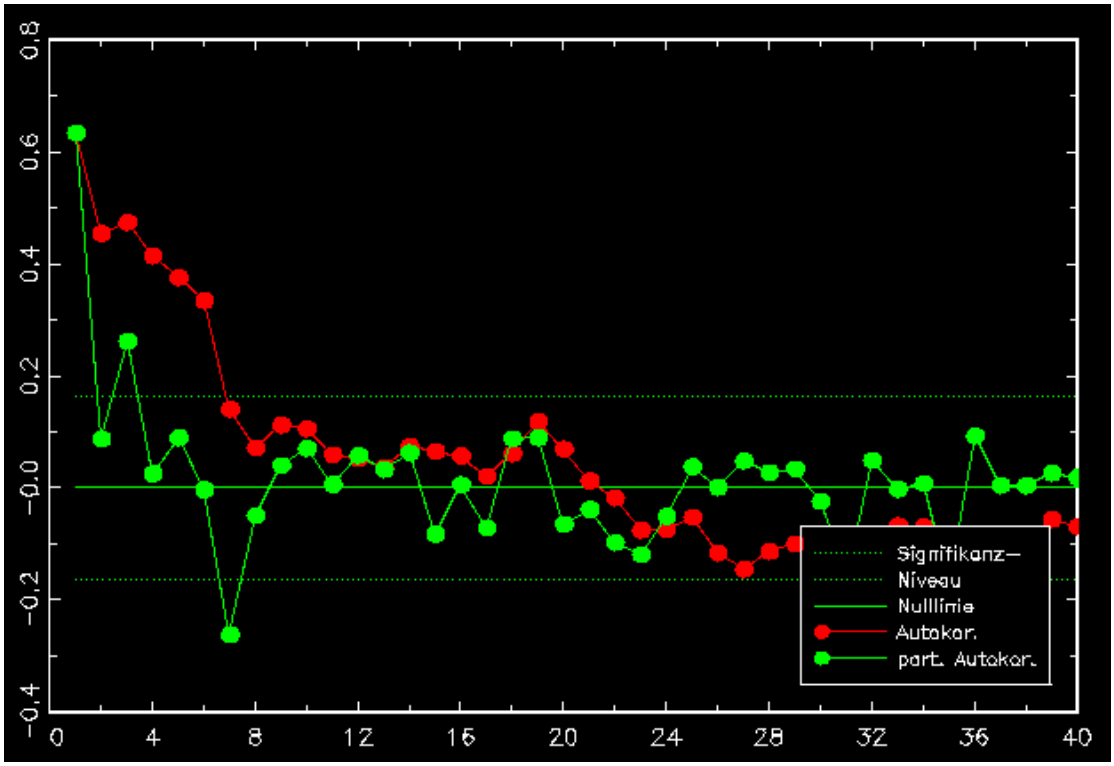

Abbildung 4.9: (Partielle) Autokorrelationen zur undifferenzierten CO_2-Zeitreihe (24.05.94 bis 17.11.94)

Wieder setzen sich das ARIMA-Modell und die SNNS-Prognose deutlich gegenüber der Regression ab. Diesmal allerdings liefert das ARIMA-Verfahren eine bessere Prognoseleistung als das beste Neuronale Netz. Eine Gegenüberstellung der Modellfehler findet sich in Tabelle 4.46.

Tabelle 4.46: Vergleich der Prognosefehler der einzelnen Verfahren

Verfahren	MAPE	RMSE
ARIMA[1,0,0]	2.3490	10.3926
SNNS	2.3632	10.6344
Regression	2.7213	12.0172

Abbildung 4.10 zeigt die CO_2-Zeitreihe und die Modellreihen im Vergleich.

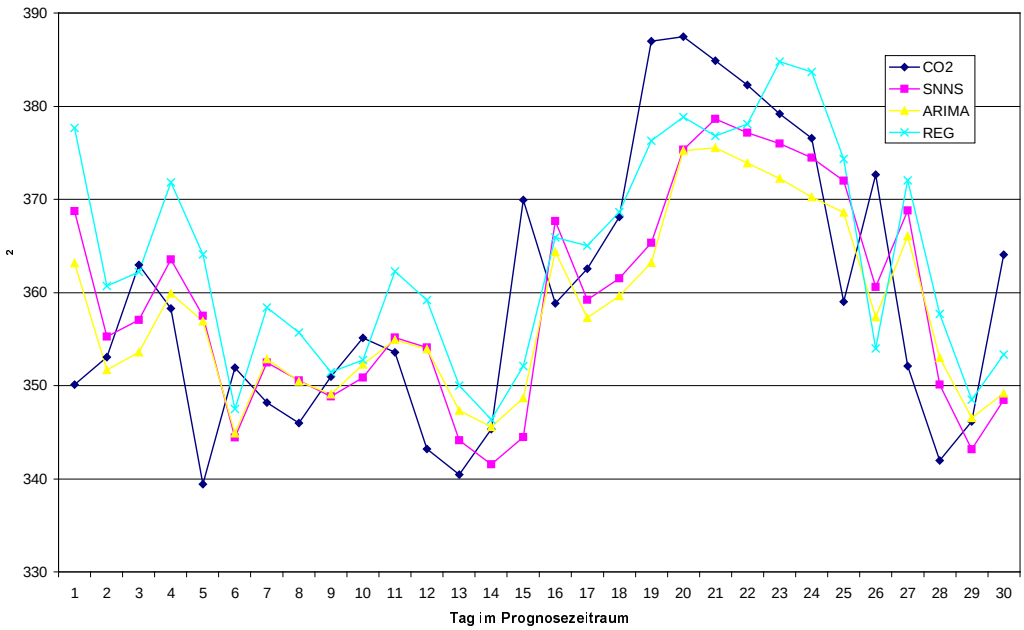

Abbildung 4.10: Vergleich der Verfahren in Prognosezeitraum 2 (24.05.94 bis 17.11.94)

4.4 Verfahren zur Auswahl exogener Zeitreihen für Neuronale Netze

4.4.1 Pruning

Das Pruning-Verfahren wurde in Kapitel 3.2.5 vorgestellt. Um eine weitgehende Optimierung der Netzarchitektur zu erreichen, wurde SNNS um die Möglichkeit erweitert, bis zu drei verschiedene Pruning-Verfahren nacheinander anzuwenden. Im folgenden werden die einzelnen Schritte beim Ablauf dargestellt. Dazu wird von einem recht umfangreichen Netz mit allen[71] 22 Inputs und 20 Neuronen in der verborgenen Schicht ausgegangen. Im ersten Schritt werden nun die Gewichte bzw. Links geprunt, um die Zahl der freien Variablen einzuschränken. Interessant wird es dann in den nächsten beiden Schritten, wenn erst die Hidden-Neuronen und dann die Inputs geprunt werden. Abbildung 4.11 zeigt die Anzahl der Neuronen in den beiden Schichten im Verlauf des Prunings.

[71] Da eine Analyse durchgeführt wird, fließt die CO_2-Zeitreihe nicht autoregressiv mit ein.

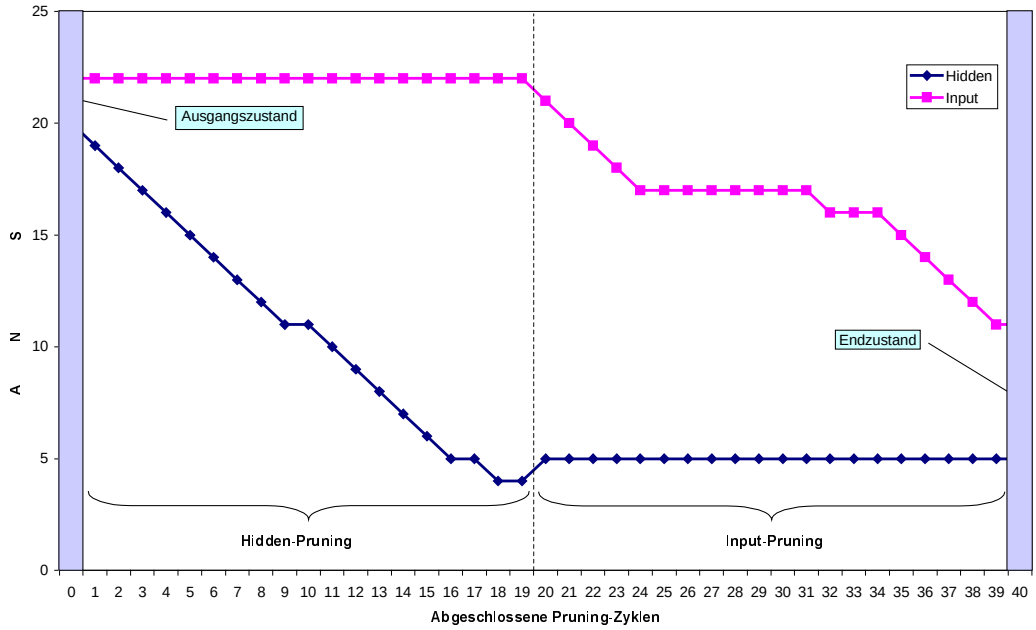

Abbildung 4.11: Anzahl der Neuronen in Hidden- und Input-Schicht im Verlauf des Verfahrens

Die Anzahl der Neuronen in der verborgenen Schicht reduziert sich zunächst von 20 auf vier reduziert, den besten Testfehler liefert jedoch das Netz mit fünf Hidden-Neuronen, so daß nach Ende des Hidden-Prunings (nach Zyklus 19) wieder mit fünf Neuronen in der verborgenen Schicht gearbeitet wird. Eine Stagnation der Neuronenanzahl, wie sie für die Zyklen 26 bis 31 für die Input-Neuronen beobachtet werden, rührt daher, daß das Entfernen eines Neurons das Netz auch mit Nachtraining nicht verbessert.

Für die Auswahl der zu entfernenden Input-Neuronen wird die „Wichtigkeit“ jedes Neurons, also der Beitrag zum Netzfehler, berechnet. Entfernt wird immer das jeweils unwichtigste Neuron. Tabelle 4.47 zeigt eine Übersicht über die Reihenfolge beim Input-Pruning.

Tabelle 4.47: Ablauf des Input-Prunings

Zyklus	BG	HAUDE	JTG	JTG_SIN	JTG_COS	NDS_DWD	NDS_FH	RLF	SDEF	SSD	TG	TMP_DWD	TMP_FH	THT_DWD	THT_FH	WG	WR	WRSIN	WRCOS	WRG	WRGSIN	WRGCOS
20	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
21	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+
22	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+
23	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	+
24	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
25	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
26	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	-
27	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
28	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
29	-	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	+
30	-	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
31	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	+
32	-	-	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
33	-	-	+	+	+	+	+	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+
34	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+	+	-	+	-	+	+	+
35	-	-	+	+	+	+	+	+	-	-	+	-	+	+	+	+	-	+	-	+	+	+
36	-	-	+	+	+	+	-	+	-	-	+	-	+	+	+	+	-	+	-	+	+	+
37	-	-	+	+	+	+	-	-	-	-	+	-	+	+	+	+	-	+	-	+	+	+
38	-	-	+	+	+	+	-	-	-	-	+	-	+	+	+	+	-	+	-	+	+	-
39	-	-	+	+	+	+	-	-	-	-	+	-	+	+	+	-	-	+	-	+	+	-
40	-	-	-	+	+	+	-	-	-	-	+	-	+	+	+	+	-	+	-	+	+	-

Die endgültige Netzkonfiguration ist ein Netz mit 11 Inputs und fünf Neuronen in der verborgenen Schicht.

4.4.2 Auswahl nach Sensitivität

Zur Bewertung des Verfahrens wurde eine Analyse auf den letzten 50 Tagen der zur Verfügung stehenden Daten durchgeführt. Ausgehend von 22 exogenen Zeitreihen wurde als Zielgröße 4 Zeitreihen festgelegt und der Fehlerverlauf des Verfahrens betrachtet. Abbildung 4.12 zeigt eine graphische Darstellung des Fehlerverlaufs. Durch die Fehlerwerte für 22 bis fünf Inputzeitreihen wurde eine Trendlinie mit Hilfe der Methode der kleinsten Quadrate gelegt. Die zugrundeliegende Topologie ist ein dreischichtiges FF-MLP mit der Lernfunktion BPM_ln_cosh, ohne Vorschicht, mit Shortcuts von der Input- zur Outputschicht und mit logistischer Aktivierungsfunktion. Alle anderen Einstellungen wurden aus Abschnitt 4.2.5 übernommen.

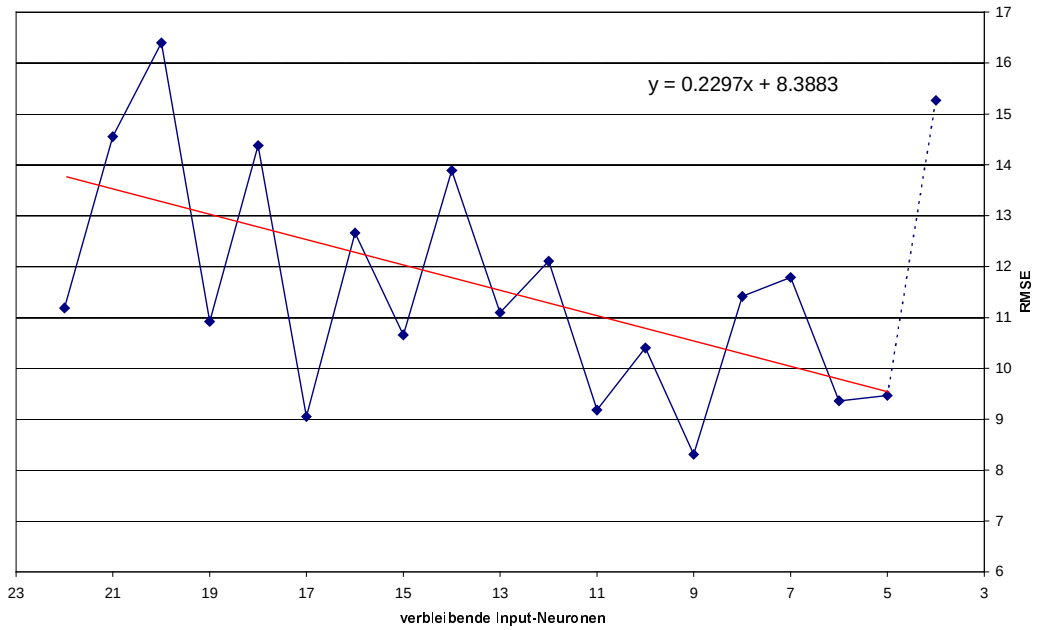

Abbildung 4.12: Fehlerverlauf des Sensitivitätsmodells zur Auswahl geeigneter Inputparameter

Die in den verschiedenen Verfahrensschritten entfernten Input-Neuronen mit ihren Sensitivitäten werden zusammen mir dem jeweiligen Netzfehler in Tabelle 4.48 aufgeführt.

Tabelle 4.48: Übersicht über den Ablauf des Sensitivitätsmodells

Verfahrens-schritt	Inputs davor	RMSE davor	Entferntes Inputneuron	Sensitivität des Neurons	Inputs danach	RMSE danach
1	22	11.185	BG	0.000521	21	14.555
2	21	14.555	WRCOS	0.000233	20	16.398
3	20	16.398	WRG	0.000339	19	10.924
4	19	10.924	JTG	0.004377	18	14.378
5	18	14.378	THT_FH	0.003276	17	9.047
6	17	9.047	TMP_FH	0.000532	16	12.662
7	16	12.662	HAUDE	0.008275	15	10.655
8	15	10.665	RLF	0.006033	14	13.889
9	14	13.889	WR	0.002944	13	11.092
10	13	11.092	JTG_SIN	0.000367	12	12.105
11	12	12.105	WRGCOS	0.001470	11	9.181
12	11	9.181	NDS_FH	0.029431	10	10.403
13	10	10.403	SDEF	0.007034	9	8.311
14	9	8.311	SSD	0.017033	8	11.412
15	8	11.412	THT_DWD	0.034345	7	11.791
16	7	11.791	TMP_DWD	0.001584	6	9.361
17	6	9.361	TG	0.022068	5	9.466
18	5	9.466	JTG_COS	0.100004	4	15.264

Die günstigste Konfiguration ergab sich nach 13 Verfahrensschritten, als noch neun Input-Neuronen übrig waren. Die Sensitivitäten dieser Topologie finden sich in Tabelle 4.49.

Tabelle 4.49: Die günstigsten Inputparameter (nach Sensitivitätsmodell) mit Sensitivitäten

Parameter	Sensitivität
NDS_DWD	0.0838
TMP_DWD	0.1163
THT_DWD	0.1205
WG	0.1192
SSD	0.0170
WRSIN	0.1193
WRGSIN	0.1297
JTG_COS	0.1052
TG	0.0306

4.4.3 Vergleich der Verfahren

Beide Verfahren wurden im Rahmen eine Analyse auf den letzten 50 Tagen des Jahres 1996 getestet und miteinander verglichen[72]. Eine Auswertung der Netzfehler der endgültigen Netze ist in Tabelle 4.50 gegeben.

Tabelle 4.50: Vergleich der Verfahren zur Verbesserung der Netztopologie

Verfahren	mape	mre	sres	rmse	theil
Sensitivitätsmodell	1.6859	0.0169	3108.3	8.311	0.5511
Pruning-Verfahren	2.8692	0.0282	7445.6	12.863	0.8483

Auf dem betrachteten Intervall zeigt die Netzarchitektur des Sensitivitätsmodells einen deutlich geringeren Netzfehler als die des Pruning-Verfahrens.

[72] Überlegungen zur prinzipiellen Vergleichbarkeit der Verfahren finden sich in Kapitel 5.

5 Diskussion

5.1 Analyse der Datenreihen

Die Betrachtung der jahreszeitlichen Periodizität zeigt eine Schwingung der CO_2-Werte mit vergleichsweise niedrigen Werten im Sommer und hohen Werten im Winter. Allein die durchschnittliche Tageskonzentration für die Januarmonate entspricht nicht diesem Verlauf.

Aus der Trendberechnung über die Jahre 1992 bis 1996 geht ein Anstieg der durchschnittlichen jährlichen CO_2-Konzentration von ca. 2,8124 ppm hervor. Dieser Wert zeigt eine sehr gute Übereinstimmung mit den Ergebnissen von Forstreuter [For 93] für den Zeitraum von 1984 bis 1991:

> „The continuous measurements of the atmospheric CO_2-concentration from 1984 to 1991 showed a significant yearly increase of 2.8 ppm.“[73]

Das Korrelogramm zeigt deutlich eine starke Autokorrelation, insbesondere zum Vortageswert. Eine absolute Korrelation von mehr als 0,4 existiert nur für die Windgeschwindigkeit (WG) und die geshiftete CO_2-Zeitreihe selbst (Shift 1 bis 4). Für folgende Zeitreihen ist die Korrelation kleiner als 0,05: Sinus des Jahrestages (JTG_SIN), Potentielle Verdunstung (HAUDE), Sättigungsdefizit (SDEF) und Tag im Jahr (JTG).

Generell müssen hier lineare und nichtlineare Einflüsse unterschieden werden. Zeitreihen, die eine hohe Korrelation zu den CO_2-Daten aufweisen, sind linear stark mit dieser verknüpft und eignen sich so als Inputparameter für lineare Modelle wie etwa die multiple lineare Regression. Der auffälligste Wert ist hier sicherlich die hohe negative Korrelation der Windgeschwindigkeit zu den CO_2-Daten. Da sich in der näheren Umgebung der Meßstation viele Emissionsquellen wie die Universitäts- und Fachhochschulgebäude, Wohnungen und Industrie befinden, läßt sich dies als Abhängigkeit von lokalen Emissionen erklären: Bei geringer Wingeschwindigkeit haben lokale Emissionen einen größeren Einfluß auf die CO_2-Konzentration in der Luft, bei starken Winden ist der Austausch der Luft um die Meßstation größer, hohe lokale Konzentrationen werden ausgeglichen. Interessant ist weiterhin, daß die Windrichtung (WR) offenbar einen deutlich geringeren Einfluß hat.

Auch die negative Korrelation der CO_2-Zeitreihe zur Temperatur (TMP) läßt sich durch lokale Emissionen erklären. Hierfür sind zum einen Heizemissionen der umliegenden Gebäude verantwortlich, zum anderen ist die Temperatur ein Indikator der Jahreszeit und damit der Vegetationstätigkeit. Im Winter (mit niedrigen Temperaturen) kommen

---

[73] Aus: [For 93], 175.

also höhere Emissionen und mangelnde Vegetationsaktivität zusammen. Diese jahreszeitlichen Einflüsse werden neben der Temperatur (und damit auch der Tageshöchstemperatur THT) in natürlicher Weise auch von der Cosinus-Transformation des Jahrestages (JTG_COS) erfaßt, allerdings ist diese Zeitreihe positiv zur CO_2-Zeitreihe korreliert, da die Transformation im Winter (Dezember/Januar) hohe Werte und im Sommer (Juni/Juli) niedrige Werte annimmt.

Die versteckten Korrelationen lassen sich am Beispiel der Sonnenscheindauer (SSD) verdeutlichen: Es ist offensichtlich, daß die Sonnenscheindauer auf die Temperatur besonders im Sommer einen starken Einfluß ausübt. Weiterhin ist sie in natürlicher Weise mit dem Bedeckungsgrad (BG) korreliert und durch die Regulation der Sonneneinstrahlung auch mit der relativen Luftfeuchtigkeit (RLF). Diese fließt wiederum in die Formel zur Bestimmung der potentiellen Verdunstung (HAUDE) ein, eine hohe Korrelation ist hier also nicht verwunderlich. Das Sättigungsdefizit (SDEF) ist bis auf den Haude-Proportionalitätsfaktor identisch mit der potentiellen Verdunstung, deswegen auch hier die extrem hohe Korrelation.

Erstaunlich ist die verhältnismäßig niedrige Korrelation der beiden Niederschlagszeitreihen (NS_DWD und NS_FH) von nur 0,56. Offenbar gibt es hier lokal deutlich unterschiedliche Niederschlagsmengen.

Die neuronale Sensitivitätsanalyse zeigt eine etwas andere Bewertung der Wichtigkeit der einzelnen Inputs. Zwar zeigt sich auch hier wieder ein starker Einfluß der Windrichtung, die größte Sensitivität hat allerdings der Cosinus des Jahrestages. Die Tageshöchsttemperatur vom Deutschen Wetterdienst (TMP_DWD) ist für das Neuronale Netz offenbar bedeutsamer als die der Fachhochschule (TMP_FH), während sich für den lineare kein Unterschied zwischen den beiden Zeitreihen zeigt. Auffällig ist auch die extrem niedrige Sensitivität der Windrichtung (WR), die für die lineare Modellierung eine wichtige Komponente darstellt[74].

5.2 Vergleich der Methoden

Der theoretische Vergleich der Abbildbarkeit der multiplen linearen Regression auf ein vergleichsweise einfaches Neuronales Netz macht deutlich, daß Neuronale Netze bei geeigneter Topologie zur Zeitreihenanalyse besser geeignet sein müssen: Neben den (linearen) Zusammenhängen die auch die Regression erfaßt, können sie weitere (vornehmlich nichtlineare) Einflüsse modellieren. Fehlen jedoch die entsprechenden Verbindungen, der konstante Input, oder werden andere Aktivierungs- und Ausgabefunktionen verwendet, ist dies nicht mehr zwangsläufig gegeben.

Die Beschränkung der Untersuchungen auf ein einzelnes Jahr (1996) ist etwas problematisch, denn möglicherweise handelt es sich um besonders untypisches Jahr. Da allerdings für alle Verfahren die gleichen Bedingungen gelten, stellt dies keine Minderung

[74] Siehe Tabelle 4.5.

der Aussagekraft der Untersuchung dar. Fehlwerte in den einzelnen Zeitabschnitten spielen ebenfalls für den Vergleich keine Rolle, da sie keinen Einfluß auf die Modellbildung haben.

In An*betracht der Rechenzeit erschien die Beschränkung der möglichen exogenen Zeitreihen auf 14 als sinnvoll. Da die Bestimmung geeigneter Inputvariablen ohnehin nur für die Regression durchgeführt wurde und da diese Verfahren das Kriterium für die Auswahl lieferten, stellt dies keine Einschränkung für die Gültigkeit der Ergebnisse dar. Auch für die Wahl der günstigsten Netzstruktur wurde die Anzahl zu testender Topologien aus Gründen der Rechenzeit eingeschränkt und somit u.U. besonders gute Netze übersehen. Die Anzahl der freien Parameter ist jedoch gerade bei Neuronalen Netzen zu groß, als daß eine umfassende Betrachtung möglich wäre.

Es zeigt sich jedoch, daß die Wahl der verwendeten Netzstruktur einen deutlichen Einfluß auf den jeweiligen Netzfehler hat.

5.3 Vergleich der Verfahren

5.3.1 Regression, Komponentenmodell und Neuronale Netze

Bei dem Vergleich der Verfahren auf den sechs zweimonatigen Zeiträumen wird deutlich, daß sich die Neuronalen Netze recht gut für die Prognose- bzw. Analyseaufgaben eignen. Zwar weist hin und wieder das Komponentenmodell den geringeren Fehler auf und in seltenen Fällen sogar die Regression, aber im Großen und Ganzen sind Neuronale Netze hier die besseren Verfahren. Dies trifft gleichermaßen für Analysen und Prognosen zu.

Im Sommer zeigen sich die mathematischen Modelle besser als im Winter, was vermutlich auf die durch die Vegetation bedingte geringere Schwankung der CO_2-Werte zurückzuführen ist.

5.3.2 ARIMA-Modellierung

Die Vergleichbarkeit von ARIMA-Modellen und Neuronalen Netzen bzw. der multiplen linearen Regression ist nur bedingt gegeben, da das ARIMA-Verfahren strikt univariat ist, die Neuronalen Netze und auch die Regression aber eine multivariate Modellierung ermöglichen. Unter der in Kapitel 1 getroffenen Annahme, daß die CO_2-Konzentration von anderen Umweltparametern beeinflußt wird, kann eine univariate Erklärung der CO_2-Werte „aus sich selbst heraus“ den Prozeß höchstens näherungsweise beschreiben. Somit muß angenommen werden, daß die ARIMA-Modellierung weniger gut geeignet ist als die multivariaten Verfahren.

Die ARIMA-Prognose auf dem gegebenen Zeitabschnitt zeigt die für AR- und MA-Prozesse typische Verschiebung gegenüber der Originalzeitreihe. In unserem speziellen Fall schnitt das ARIMA-Modell nur auf dem Zeiten Intervall besser gegenüber den Neuronalen Netzen ab.

5.4 Verfahren zur Auswahl exogener Zeitreihen für Neuronale Netze

Prinzipiell unterscheiden sich die beiden untersuchten Verfahren zur Netzwerkoptimierung schon im Ansatz: Während das Pruning-Verfahren neben Input-Neuronen auch noch Neuronen der verborgenen Schicht und Gewichte zu optimieren versucht, beschränkt sich das Sensitivitätsmodell ganz auf die Inputs. Auch der Aufwand variiert erheblich: Das Sensitivitätsmodell führt für *jeden* Verfahrensschritt ein komplettes Netztraining durch, das Pruning-Verfahren hingegen beschränkt sich auf *ein* Netztraining plus für jeden Verfahrensschritt ein Nachtraining, das aber i.d.R. aus vergleichsweise wenigen Trainingszyklen besteht.

5.4.1 Pruning

Beim Pruning des hier gewählten Neuronalen Netzes ergab sich besonders beim Neuronen-Pruning eine wesentliche Veränderung der Netzstruktur. Die Reihenfolge der geprunten Input-Neuronen läßt darauf schließen, welche Bedeutung die einzelnen Inputs für ein neuronales Modell haben. Offenbar sind hier die Einflüsse anders als bei der Regression: Während dort die Zeitreihen TG, JTG_SIN und JTG einen sehr geringen Einfluß ausüben bzw. bereits bei der Vorauswahl exogener Variablen herausgefallen sind, bleiben sie im Netz erhalten oder fallen erst sehr spät aus dem Modell heraus. Auf der anderen Seite haben Zeitreihen mit großem Einfluß bei der Regression (WR, TMP_DWD) im Netz einen geringeren Einfluß und werden gleich zu Anfang geprunt.

Da das vollständig geprunte Netz erheblich filigraner ist als das Ausgangsnetz, reduziert das Verfahren die Zahl der freien Parameter effektiv. Die Dauer eines dreischrittigen Pruning-Laufes hängt stark von den gewählten Optionen bei den einzelnen Verfahren ab. Insbesondere die Zahl der Trainingsepochen für das Nachtraining hat bei vielen zu entfernenden Elementen (z.B. Gewichte) einen starken Einfluß auf die Laufzeit. Die hier gewählte Konfiguration lief 67 Minuten[75] und ist damit erheblich schneller als die Auswahl nach Sensitivität.

5.4.2 Auswahl nach Sensitivität

Bei der Auswahl nach Sensitivität wird in jedem Verfahrensschritt ein Input-Neuron entfernt, bis eine gewünschte Anzahl von Inputs übrig bleibt. Dabei kann entweder das letzte erzeugte Netz verwendet werden, oder das mit dem bisher geringsten Netzfehler.

Aus der Fehlerentwicklung geht eine deutliche globale Verbesserung des Netzfehlers im Verlauf des Verfahrens hervor. Lediglich der Wert des letzten Verfahrensschrittes ist zu hoch, was eine Beendigung des Verfahrens an dieser Stelle nahelegt. Auch hier läßt die Reihenfolge der entfernten Neuronen Rückschlüsse über die Bedeutung der entsprechenden Inputzeitreihen im neuronalen Modell zu. Das bei Pruning beobachtete Phänomen findet sich hier wieder: Zeitreihen mit großem Einfluß bei der Regression fallen

[75] Dies ist ein Durchschnittswert aus drei Messungen auf einer *HP Apollo 9000/730*.

recht früh aus dem Modell (THT_FH), während umgekehrt Zeitreihen mit geringem Einfluß erst sehr spät entfernt werden (TG, SDEF, SSD). Die beiden für die Regression entscheidendsten Datenreihen (WG und NDS_DWD) waren offenbar auch für die Netze wichtig: Sie wurden nicht aus dem Modell entfernt.

6 Schlußfolgerungen und Ausblick

Es hat sich gezeigt, daß Neuronale Netze, die ursprünglich zur Mustererkennung konzeptioniert waren, sich auch für die Analyse und Prognose von Zeitreihen eignen. Während den in der Praxis verwandten mathematischen Modellen oft Annahmen über die Einflüsse auf die zu prognostizierende bzw. analysierende Zeitreihe vorausgehen, müssen diese für Neuronale Netze nicht getroffen werden. Somit ergeben sich größere Freiheiten bei der Modellierung. Es ist allerdings zu bedenken, daß auch bereits bekannte Wirkzusammenhänge bei der neuronalen Modellierung aufgrund der Vorgaben für die Modellstruktur i.a. nicht genutzt werden können.

Gegenüber den mathematischen Verfahren zeigen die Neuronalen Netze die Eigenschaft, auch nichtlineare Zusammenhänge zu erfassen, ohne dabei auf komplexe Berechnungsverfahren zurückgreifen zu müssen. Ferner sind Neuronale Netze in Verbindung mit der neuronalen Sensitivitätsanalyse geeignet, lineare und nichtlineare Einflüsse exogener Zeitreihen zu identifizieren und einen Anhaltspunkt über die Stärke eines Einflusses zu geben.

Die Frage nach den am besten geeigneten Netzarchitekturen läßt sich nicht befriedigend beantworten. Offenbar ist die einzige gesicherte Erkenntnis, daß die lineare Skalierungsfunktion der μ-σ-Skalierung vorzuziehen ist. Ansonsten unterscheiden sich die jeweils günstigsten Architekturen zu stark, um eine eindeutige Antwort geben zu können. Für die erfolgreiche Anwendung von Neuronalen Netzen im Umweltbereich fehlt daher noch ein zuverlässiges Verfahren zur automatischen Bestimmung einer geeigneten Netzarchitektur. Ohne ein Verfahren, welches selbst eine geeignete Topologie finden kann, setzt der Einsatz Neuronaler Netze ein Verständnis der Theorie der Neuronalen Netze voraus. Die noch in der Entwicklung befindliche Auswahl nach Korrelationskoeffizient müßte eingehend mit den anderen Verfahren verglichen werden. Die beiden besprochenen Verfahren sind bereits gute Ansätze, lassen sich aber sicher noch verbessern, beispielsweise durch Variation der Lernfunktion während des Verfahrens. Hier liegt noch ein großes Potential für die Anwendung Neuronaler Netze in der Praxis.

Es hat sich gezeigt, daß die CO_2-Konzentration der Luft vor allem von lokalen Emissionen und der Vegetation geprägt ist. Die wesentlichen Einflüsse lassen sich in jahreszeitliche Faktoren (Temperatur und damit verknüpft: Vegetation), die Durchmischung der Luft (Windgeschwindigkeit) und die CO_2-Konzentration des Vortages unterteilen.

Mit Hilfe der neuronalen Sensitivitätsanalyse ist es möglich, in die „Black Box“ eines Neuronalen Netzes zu schauen und neben linearen auch nichtlineare Einflüsse aufzuzeigen. Auch wenn sich die so identifizierten Einflußgrößen mit denen teilweise decken, die über die Korrelationskoeffizienten bestimmt wurden, können hier doch teilweise neue Erkenntnisse über das betrachtete Modell gemacht werden.

Insgesamt hat sich gezeigt, daß Neuronale Netze durchaus für den komplexen Vorgang der Analyse und Prognose von Zeitreihen geeignet sind. Neben den Notwendigkeiten einer geeigneten Auswahl von Einflußzeitreihen und deren Vorverarbeitung, die auch bei anderen Verfahren der Zeitreihenanalyse gegeben sind, ist bei den Neuronalen Netzen besonders die Wahl der Netztopologie für die Analyse- und Prognosegüte ausschlaggebend. Die im Rahmen dieser Arbeit implementierten Verfahren zur Netzwerkoptimierung bieten dabei eine wertvolle Unterstützung.

7 Anhang

7.1 Tabellen und Diagramme

7.2 Günstige Netzarchitekturen für die Analyse

Tabelle 7.51: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 1)

mape	mre	sres	rmse	theil	Lern-funktion	Skalierungs-funktion	Hidden-Neuronen	Topologie-modifikation
1.8601	0.0186	3216.8000	8.4549	0.5593	BP_momentum	linear	10	vorS+shortC
1.8611	0.0187	3243.5000	8.4899	0.5623	BPM_ln_cosh	linear	3	vorS+shortC
1.8293	0.0183	3275.5000	8.5316	0.5656	BPM_ln_cosh	linear	5	vorS
1.8290	0.0183	3301.2000	8.5650	0.5682	BP_momentum	linear	6	vorS+shortC
1.8411	0.0184	3335.4000	8.6093	0.5704	BP_momentum	linear	7	vorS+shortC
1.8926	0.0189	3341.0000	8.6165	0.5679	BPM_ln_cosh	linear	9	vorS
1.8676	0.0187	3350.0000	8.6282	0.5719	BP_momentum	linear	9	none
1.8529	0.0186	3359.5000	8.6404	0.5740	BP_momentum	linear	8	none
1.8095	0.0182	3369.8000	8.6536	0.5745	BP_momentum	linear	7	vorS
1.9499	0.0194	3371.1000	8.6552	0.5708	BP_momentum	linear	4	vorS+shortC
1.7218	0.0173	3392.2000	8.6823	0.5766	BPM_ln_cosh	linear	4	none
1.8764	0.0188	3422.6000	8.7212	0.5759	BPM_ln_cosh	linear	3	none
1.7398	0.0175	3449.5000	8.7553	0.5820	Rprop	linear	7	vorS+shortC
1.8693	0.0187	3457.1000	8.7650	0.5805	BP_momentum	linear	5	vorS+shortC
1.7793	0.0178	3472.0000	8.7838	0.5825	BPM_ln_cosh	linear	7	none
1.7793	0.0178	3472.0000	8.7838	0.5825	BPM_ln_cosh	linear	7	shortC
1.8124	0.0183	3472.2000	8.7840	0.5834	BPM_ln_cosh	linear	7	vorS
1.7680	0.0177	3490.7000	8.8074	0.5841	BPM_ln_cosh	linear	8	none
1.7930	0.0180	3498.3000	8.8170	0.5849	BPM_ln_cosh	linear	6	none
1.8708	0.0187	3504.4000	8.8248	0.5842	BP_momentum	linear	4	vorS
1.8711	0.0187	3508.0000	8.8292	0.5855	BPM_ln_cosh	linear	3	vorS
1.8780	0.0188	3561.2000	8.8959	0.5889	BPM_ln_cosh	linear	7	vorS+shortC
1.8620	0.0187	3567.3000	8.9035	0.5908	BP_momentum	linear	3	vorS
1.7502	0.0176	3579.7000	8.9190	0.5919	BPM_ln_cosh	linear	6	vorS
1.9249	0.0192	3605.9000	8.9515	0.5928	BP_momentum	linear	6	vorS
1.8637	0.0187	3637.4000	8.9906	0.5964	BPM_ln_cosh	linear	4	vorS
1.8642	0.0188	3690.9000	9.0565	0.6012	BP_momentum	linear	8	vorS+shortC
1.8447	0.0186	3749.2000	9.1278	0.6054	BPM_ln_cosh	linear	9	none
1.8213	0.0184	3777.6000	9.1622	0.6092	BP_momentum	linear	3	none
1.8038	0.0181	3786.6000	9.1731	0.6093	BP_momentum	linear	4	none
2.0475	0.0203	3793.8000	9.1818	0.6039	BP_momentum	linear	5	vorS
1.8249	0.0183	3844.6000	9.2432	0.6138	BP_momentum	linear	6	none
1.9048	0.0192	3871.7000	9.2756	0.6169	BP_momentum	linear	5	none
1.8140	0.0183	3874.0000	9.2784	0.6152	BPM_ln_cosh	linear	10	none
1.7821	0.0180	3875.7000	9.2804	0.6170	Rprop	linear	7	none

Tabelle 7.52: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 2)

mape	mre	sres	rmse	theil	Lern-funktion	Skalierungs-funktion	Hidden-Neuronen	Topologie-modifikation
1.7821	0.0180	3875.7000	9.2804	0.6170	Rprop	linear	7	shortC
1.8918	0.0189	3883.7000	9.2901	0.6061	BP_momentum	μ-σ-X	7	none
1.8918	0.0189	3883.7000	9.2901	0.6061	BP_momentum	μ-σ-X	7	shortC
1.9777	0.0198	3884.9000	9.2914	0.6000	BPM_ln_cosh	μ-σ	4	none
2.0640	0.0206	3884.9000	9.2915	0.6035	BP_momentum	μ-σ	6	vorS+shortC
1.9200	0.0192	3893.2000	9.3013	0.6054	BPM_ln_cosh	μ-σ-X	7	none
1.9200	0.0192	3893.2000	9.3013	0.6054	BPM_ln_cosh	μ-σ-X	7	shortC
2.0509	0.0205	3920.8000	9.3343	0.6029	BP_momentum	μ-σ	7	vorS+shortC
1.8291	0.0184	3938.5000	9.3553	0.6215	Rprop	linear	7	vorS
1.8847	0.0190	3963.9000	9.3855	0.6198	BPM_ln_cosh	linear	6	vorS+shortC
2.0179	0.0202	3967.0000	9.3891	0.6125	Rprop	μ-σ-X	7	vorS
1.9249	0.0194	3973.1000	9.3963	0.6171	BPM_ln_cosh	μ-σ-0	7	vorS+shortC
2.0978	0.0209	3985.6000	9.4111	0.6224	BPM_ln_cosh	linear	10	vorS+shortC
2.0159	0.0202	4007.0000	9.4363	0.6113	BPM_ln_cosh	μ-σ	10	none
1.9019	0.0192	4044.6000	9.4805	0.6195	Rprop	μ-σ-0	7	vorS
2.0615	0.0207	4057.8000	9.4959	0.6138	BP_momentum	μ-σ	4	vorS+shortC
2.0080	0.0202	4071.0000	9.5114	0.6117	BP_momentum	μ-σ	4	none
1.9638	0.0197	4118.7000	9.5669	0.6186	BP_momentum	μ-σ-X	7	vorS+shortC
1.9530	0.0198	4134.8000	9.5857	0.6313	Rprop	μ-σ-0	7	vorS+shortC
1.9896	0.0200	4136.3000	9.5874	0.6186	Rprop	μ-σ	7	vorS
1.9426	0.0196	4155.3000	9.6094	0.6382	BPM_ln_cosh	linear	5	vorS+shortC
2.0622	0.0207	4206.8000	9.6688	0.6313	BPM_ln_cosh	μ-σ	3	none
1.9526	0.0197	4215.0000	9.6781	0.6423	BP_momentum	linear	8	vorS
2.0782	0.0208	4228.6000	9.6937	0.6276	BPM_ln_cosh	μ-σ	9	none
1.9170	0.0194	4240.9000	9.7079	0.6413	BPM_ln_cosh	linear	8	vorS+shortC
2.0770	0.0209	4256.6000	9.7258	0.6279	BP_momentum	μ-σ	3	vorS+shortC
2.0527	0.0206	4287.2000	9.7606	0.6336	BP_momentum	μ-σ	8	none
2.0864	0.0210	4288.2000	9.7619	0.6292	BP_momentum	μ-σ	9	vorS+shortC
1.9229	0.0195	4303.5000	9.7792	0.6439	BPM_ln_cosh	linear	4	vorS+shortC
2.0659	0.0207	4314.6000	9.7918	0.6369	BPM_ln_cosh	μ-σ	7	none
2.0659	0.0207	4314.6000	9.7918	0.6369	BPM_ln_cosh	μ-σ	7	shortC
2.0813	0.0209	4357.7000	9.8406	0.6356	BP_momentum	μ-σ	10	none
2.1253	0.0213	4358.7000	9.8417	0.6399	BP_momentum	μ-σ	7	none
2.1253	0.0213	4358.7000	9.8417	0.6399	BP_momentum	μ-σ	7	shortC
2.0546	0.0206	4359.9000	9.8431	0.6405	Rprop	μ-σ-X	7	none
2.0546	0.0206	4374.5000	9.8595	0.6421	BPM_ln_cosh	μ-σ	8	none
2.0895	0.0210	4374.6000	9.8597	0.6378	BPM_ln_cosh	μ-σ	5	none
2.0682	0.0208	4379.6000	9.8653	0.6426	BP_momentum	μ-σ	3	none
2.0770	0.0209	4384.8000	9.8711	0.6363	BP_momentum	μ-σ	9	none
1.9243	0.0195	4385.9000	9.8724	0.6564	BPM_ln_cosh	linear	5	none
2.0661	0.0207	4436.0000	9.9286	0.6384	BP_momentum	μ-σ-X	7	vorS
2.1115	0.0212	4443.1000	9.9366	0.6412	BP_momentum	μ-σ	10	vorS+shortC
1.9235	0.0195	4452.0000	9.9466	0.6538	BPM_ln_cosh	μ-σ-0	7	vorS
2.1046	0.0211	4473.6000	9.9706	0.6456	BPM_ln_cosh	μ-σ	6	none
1.9599	0.0198	4495.1000	9.9945	0.6605	BPM_ln_cosh	linear	9	vorS+shortC
2.0896	0.0210	4497.9000	9.9977	0.6459	Rprop	μ-σ	7	none
2.0896	0.0210	4497.9000	9.9977	0.6459	Rprop	μ-σ	7	shortC
1.9835	0.0200	4543.5000	10.0480	0.6631	Rprop	μ-σ-0	7	none

Tabelle 7.53: Netzarchitekturen und Analysefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert (Teil 2)

mape	mre	sres	rmse	theil	Lern-funktion	Skalierungs-funktion	Hidden-Neuronen	Topologie-modifikation
1.9835	0.0200	4543.5000	10.0480	0.6631	Rprop	μ-σ-0	7	shortC
2.1302	0.0214	4701.2000	10.2210	0.6603	BP_momentum	μ-σ	5	vorS+shortC
2.0533	0.0208	4701.1000	10.2210	0.6723	BPM_ln_cosh	μ-σ-0	7	none
2.0533	0.0208	4701.1000	10.2210	0.6723	BPM_ln_cosh	μ-σ-0	7	shortC
2.1553	0.0217	4745.8000	10.2700	0.6620	BP_momentum	μ-σ	6	none
2.1456	0.0216	4772.1000	10.2980	0.6627	BP_momentum	μ-σ	5	none
2.1522	0.0217	4837.6000	10.3680	0.6725	BP_momentum	μ-σ	8	vorS+shortC
2.2541	0.0226	5116.0000	10.6620	0.6974	Rprop	μ-σ	7	vorS+shortC
2.0994	0.0213	5163.5000	10.7120	0.7114	BP_momentum	linear	9	vorS
2.4432	0.0241	5214.5000	10.7650	0.7041	BP_momentum	linear	9	vorS+shortC
2.1879	0.0222	5428.5000	10.9830	0.7274	BP_momentum	linear	10	vorS
2.1910	0.0223	5803.8000	11.3570	0.7444	BP_momentum	μ-σ-0	7	vorS
2.2928	0.0233	5827.8000	11.3800	0.7527	BP_momentum	linear	3	vorS+shortC
2.5836	0.0254	5901.1000	11.4510	0.7481	BP_momentum	linear	10	none
2.3007	0.0234	6008.6000	11.5550	0.7579	BP_momentum	μ-σ-0	7	none
2.3007	0.0234	6008.6000	11.5550	0.7579	BP_momentum	μ-σ-0	7	shortC
2.6109	0.0257	6157.0000	11.6970	0.7631	BP_momentum	linear	7	none
2.6109	0.0257	6157.0000	11.6970	0.7631	BP_momentum	linear	7	shortC
2.3144	0.0236	6408.0000	11.9330	0.7777	BP_momentum	μ-σ-0	7	vorS+shortC
2.5299	0.0252	6508.2000	12.0260	0.7888	Rprop	μ-σ-X	7	vorS+shortC
2.7800	0.0273	6616.0000	12.1250	0.7930	BPM_ln_cosh	linear	8	vorS
2.8117	0.0276	6662.5000	12.1680	0.7932	BPM_ln_cosh	linear	10	vorS
4.3035	0.0442	17520.0000	19.7320	1.2769	BPM_ln_cosh	μ-σ-X	7	vorS+shortC
4.3036	0.0442	17522.0000	19.7330	1.2769	BP_momentum	μ-σ	3	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	3	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	3	vorS
4.3036	0.0442	17522.0000	19.7330	1.2769	BP_momentum	μ-σ	4	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	4	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	4	vorS
4.3036	0.0442	17522.0000	19.7330	1.2769	BP_momentum	μ-σ	5	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	5	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	5	vorS
4.3036	0.0442	17522.0000	19.7330	1.2769	BP_momentum	μ-σ	6	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	6	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	6	vorS
4.3036	0.0442	17522.0000	19.7330	1.2769	BP_momentum	μ-σ	7	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	7	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	7	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BP_momentum	μ-σ	8	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	8	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	8	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BP_momentum	μ-σ	9	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	9	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	9	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BP_momentum	μ-σ	10	vorS
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	10	vorS+shortC
4.3037	0.0442	17523.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ	10	vorS
4.3036	0.0442	17522.0000	19.7330	1.2769	BPM_ln_cosh	μ-σ-X	7	vorS

7.2.1 Günstige Netzarchitekturen für die Prognose

Tabelle 7.54: Netzarchitekturen und Prognosefehler im Zeitraum vom 12.11.96 - 31.12.96, nach RMSE sortiert

mape	mre	sres	rmse	theil	Lern-funktion	Skalierungs-funktion	Hidden-Neuronen	Topologie-modifikation
2.8852	0.0289	7648.8	13.3370	0.8673	BP_momentum	my_sigma	7	none
2.8852	0.0289	7648.8	13.3370	0.8673	BP_momentum	my_sigma	7	shortC
2.8415	0.0284	7766.7	13.4400	0.8740	BP_momentum	my_sigma0	7	none
2.8415	0.0284	7766.7	13.4400	0.8740	BP_momentum	my_sigma0	7	shortC
2.8849	0.0288	7786.5	13.4570	0.8751	BPM_ln_cosh	my_sigma	7	none
2.8849	0.0288	7786.5	13.4570	0.8751	BPM_ln_cosh	my_sigma	7	shortC
2.9170	0.0291	7906.3	13.5600	0.8818	Rprop	my_sigma	7	none
2.9170	0.0291	7906.3	13.5600	0.8818	Rprop	my_sigma	7	shortC
2.8250	0.0281	7939.9	13.5890	0.8836	BPM_ln_cosh	linear	7	vorS
2.9369	0.0294	8014.3	13.6520	0.8878	Rprop	my_sigma	7	vorS
2.8938	0.0289	8026.5	13.6620	0.8885	BPM_ln_cosh	my_sigmaX	7	none
2.8938	0.0289	8026.5	13.6620	0.8885	BPM_ln_cosh	my_sigmaX	7	shortC
2.9142	0.0292	8081.9	13.7100	0.8915	BP_momentum	linear	7	vorS
2.9030	0.0292	8084.7	13.7120	0.8917	Rprop	my_sigma0	7	vorS
2.9065	0.0292	8217.4	13.8240	0.8990	Rprop	my_sigma0	7	vorS+shortC
2.9093	0.0291	8282.5	13.8790	0.9025	BP_momentum	linear	7	vorS+shortC
2.9237	0.0291	8371.6	13.9530	0.9074	BP_momentum	my_sigmaX	7	none
2.9237	0.0291	8371.6	13.9530	0.9074	BP_momentum	my_sigmaX	7	shortC
3.0075	0.0302	8375.2	13.9560	0.9075	Rprop	my_sigma0	7	none
3.0075	0.0302	8375.2	13.9560	0.9075	Rprop	my_sigma0	7	shortC
3.0053	0.0302	8417.0	13.9910	0.9098	BPM_ln_cosh	linear	7	vorS+shortC
2.9677	0.0296	8453.0	14.0210	0.9118	Rprop	my_sigmaX	7	vorS
3.0006	0.0301	8472.5	14.0370	0.9128	BPM_ln_cosh	my_sigma0	7	none
3.0006	0.0301	8472.5	14.0370	0.9128	BPM_ln_cosh	my_sigma0	7	shortC
2.9992	0.0301	8536.4	14.0900	0.9162	BPM_ln_cosh	my_sigma0	7	vorS+shortC
3.0587	0.0306	8540.7	14.0930	0.9165	Rprop	my_sigma	7	vorS+shortC
2.9874	0.0299	8835.7	14.3350	0.9322	Rprop	linear	7	vorS+shortC
3.0680	0.0309	8847.5	14.3440	0.9328	BPM_ln_cosh	my_sigma0	7	vorS
3.0058	0.0299	8866.1	14.3590	0.9338	Rprop	my_sigmaX	7	none
3.0058	0.0299	8866.1	14.3590	0.9338	Rprop	my_sigmaX	7	shortC
3.0171	0.0303	8872.1	14.3640	0.9341	BP_momentum	my_sigma0	7	vorS+shortC
3.1153	0.0315	9110.9	14.5560	0.9466	BP_momentum	my_sigma0	7	vorS
3.0607	0.0307	9325.8	14.7270	0.9577	Rprop	linear	7	vorS
3.1720	0.0320	9436.1	14.8140	0.9633	BPM_ln_cosh	linear	7	none
3.1720	0.0320	9436.1	14.8140	0.9633	BPM_ln_cosh	linear	7	shortC
3.1092	0.0311	9553.1	14.9050	0.9693	Rprop	linear	7	none
3.1092	0.0311	9553.1	14.9050	0.9693	Rprop	linear	7	shortC
3.1368	0.0314	9669.5	14.9960	0.9752	BP_momentum	linear	7	none
3.1368	0.0314	9669.5	14.9960	0.9752	BP_momentum	linear	7	shortC
3.2304	0.0321	10169.0	15.3780	1.0000	Rprop	my_sigmaX	7	vorS+shortC
3.3333	0.0332	10862.0	15.8930	1.0335	BP_momentum	my_sigmaX	7	vorS+shortC
3.3602	0.0333	11495.0	16.3500	1.0632	BP_momentum	my_sigmaX	7	vorS
4.2565	0.0437	16560.0	19.6250	1.2762	BPM_ln_cosh	my_sigmaX	7	vorS
4.2572	0.0437	16565.0	19.6270	1.2763	BPM_ln_cosh	my_sigmaX	7	vorS+shortC
4.2589	0.0437	16579.0	19.6360	1.2769	BP_momentum	my_sigma	7	vorS+shortC

7.3 Inhalt der CD-ROM

Aufgrund der Fülle an anfallenden Daten und Ergebnissen war es nicht möglich, sämtliche Zeitreihen, Tabellen, Diagramme und Skripte in schriftlicher Form mit in diese Arbeit aufzunehmen. Um diese dennoch aufzuführen, wurden sie auf eine CD-ROM gebrannt, die dieser Arbeit beiliegt[76].

7.3.1 Übersicht über die Verzeichnisstruktur

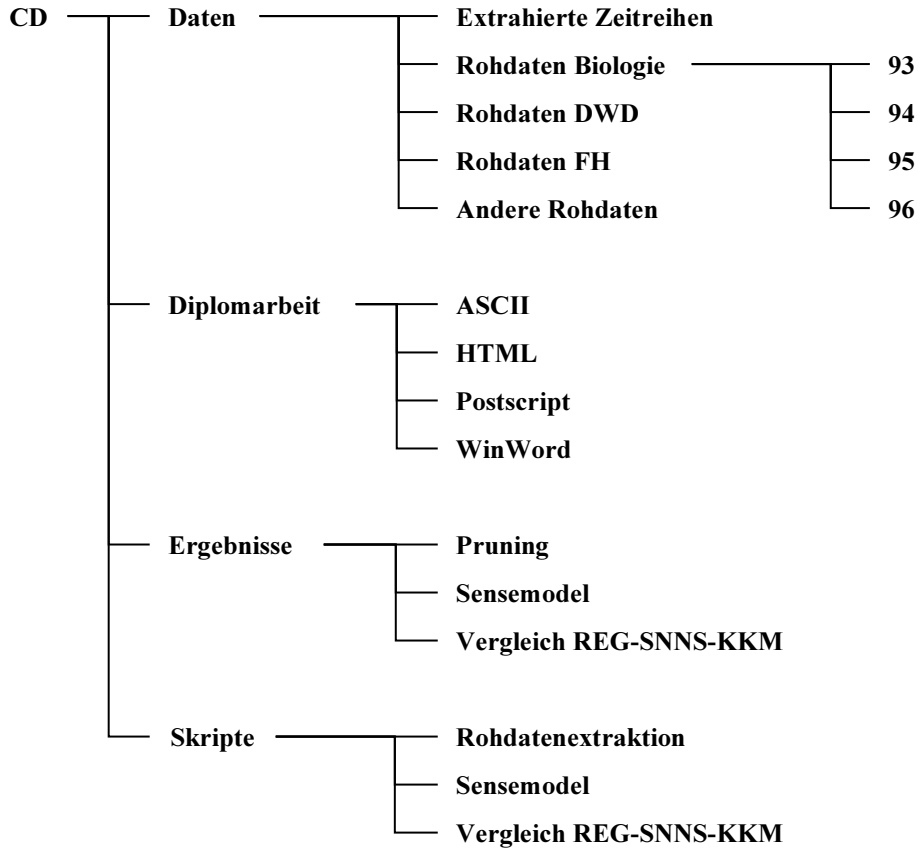

Abbildung 7.13: Übersicht über die Verzeichnisstruktur der CD-ROM

7.3.2 Daten

Das Datenverzeichnis gliedert sich in Rohdaten (verschiedene Quellen) und extrahierte Zeitreihen, die im Rahmen der Arbeit Verwendung fanden.

[76] Da die CD-ROM auch die verwendeten Daten enthält, die zum Teil nur für diese wissenschaftliche Arbeit bestimmt waren (insbesondere die Daten des DWD und der FH), enthalten nur die drei offiziellen Exemplare eine CD-ROM.

7.3.2.1 Extrahierte Zeitreihen

Enthält die verwendeten Datenreihen. Alle Datenreihen liegen im ASCII-Format vor und enthalten 1827 Zeilen, entsprechend dem betrachteten Zeitraum. Fehlwerte werden analog zu SAS als Punkt dargestellt. Der Name der Zeitreihendateien setzt sich wie folgt zusammen:

`AAA-BBB-9296.txt`

Dabei gibt AAA das Kürzel der Zeitreihe an (zwei oder mehr Buchstaben), BBB steht für die Datenquelle (DWD, FH, BIO oder ART für artifizielle selbsterstellte Zeitreihen).

7.3.2.2 Rohdaten Biologie

Enthält für die Jahre 1993-1996 die Tagesdateien der CO_2-Werte vom Fachbereich Biologie in entsprechenden Unterverzeichnissen. Die Werte für 1992 standen nicht unmittelbar zur Verfügung, sie wurden der Arbeit von Tillmanns entnommen [Til 93] und finden sich im Verzeichnis "Sonstige Rohdaten" (Siehe 7.3.2.5)in der Datei `daymean.92`.

7.3.2.3 Rohdaten DWD

Hier finden sich die Daten des Deutschen Wetterdienstes. Für die Jahre 1992 und 1993, sowie für den Zeitraum von 1994-1996 existieren ASCII-Dateien in verschiedenen Formaten. Sie heißen `sy92.out`, `sy93.out` und `sy94x96.dat`. Die Daten liegen stundenweise vor.

7.3.2.4 Rohdaten FH

Die Wetterdaten von der Fachhochschule liegen als selbsterklärende Excel-Datei (Version 97) in der Datei `Wetterdaten FH.xls` als Tageswerte vor.

7.3.2.5 Sonstige Rohdaten

Enthält die Dateien `daymean.92` und `daymean.93` aus der Arbeit von Tillmanns [Til 93], die u.a. die CO_2-Werte von 1992 und 1993 als Tageswerte enthalten.

7.3.3 Diplomarbeit

Hier findet sich der Text der Diplomarbeit, in vier verschiedenen Formaten: ASCII, HTML, Postscript Color und WinWord-Format (Version 97). Das Titelblatt befindet sich jeweils in einer gesonderten Datei mit Namen `Titelblatt.*` (* = txt, doc, etc.).

7.3.3.1 ASCII

`diplomarbeit.txt`: Die Arbeit in ASCII (MS-DOS) ohne Abbildungen und Tabellen.

7.3.3.2 HTML

`diplomarbeit.htm`: Die Arbeit in HTML mit Abbildungen und Tabellen. Diese Datei ist das Zentraldokument. Die Konvertierung erfolgte mit WinWord 97. Andere Dateien sind Filialdokumente und Abbildungen.

7.3.3.3 Postscript

`diplomarbeit.ps`: Die Arbeit in (farbigem) Postscript.

7.3.3.4 WinWord

`diplomarbeit.doc`: Die Arbeit in ihrem ursprünglichem Format: WinWord[77], Version 97. Der Aufbau erfolgte größeren wissenschaftlichen Arbeiten entsprechend mit Zentral- und Filialdokumenten. Zentraldokument ist "Diplomarbeit.doc". Andere Dateien sind Filialdokumente.

7.3.4 Ergebnisse

Dieser Ordner enthält verschiedene in der Arbeit verwendete Tabellen und Abbildungen sowie einige Tabellen, die aufgrund ihrer Größe nicht mit in die schriftliche Form der Arbeit aufgenommen werden konnten.

7.3.4.1 Sensemodel

Hier finden sich die Ergebnisse der Zeitreihenauswahl nach Sensitivität, wie sie von dem entsprechenden Skript erzeugt wurden.

`sensemodel.log`: Zeigt die Entwicklung des neuronalen Modells mit den Sensitivitäten der einzelnen Parameter im jeweiligen Verfahrensschritt.

`sense.errordevelopment`: Die Fehlerentwicklung des Verfahrens.

7.3.4.2 Pruning

Neben den Ergebnissen des Prunings dieser Arbeit finden sich hier auch die SNNS-Erweiterungen und eine Beispieldatei.

`bmanPRUNE.tmpl` und `bmanPRUNE3.tmpl`: Templates für SNNS (einfaches und mehrfaches Pruning), die bei jedem Neuronen-Prunze-Zyklus das Ablegen der Netzdatei veranlassen.

`BP.prune`: Beispieldatei für mehrfaches Pruning unter SNNS, die auch in dieser Arbeit verwendet wurde.

`prunetab.sh`: Shellskript zur Generierung einer Übersicht über die geprunten Neuronen im jeweiligen Zyklus. Aufruf: `sh prunetab.sh`.

`auswertung.txt`: Beispiel einer solchen Auswertung.

---

[77] Siehe [Nic 98].

7.3.4.3 Vergleich REG-SNNS-KKM

`?-1996.error`: Der Analysefehler für die 1001 Regressionsläufe zur Bestimmung der besten Inputvariablen für das entsprechende Monatspaar (?=1: Januar/Februar, ?=2: März/Arpil, ...).

`?p-1996.error`: Der Prognosefehler für die 1001 Regressionsläufe zur Bestimmung der besten Inputvariablen für das entsprechende Monatspaar (?=1: Januar/Februar, ?=2: März/Arpil, ...).

`SNNS_x?.error`: Der Analysefehler für die 48 verschiedenen Netztopologien in dem entsprechenden Monatspaar (?=1: Januar/Februar, ?=2: März/Arpil, ...).

`SNNS_x?p.error`: Der Prognosefehler für die 48 verschiedenen Netztopologien in dem entsprechenden Monatspaar (?=1: Januar/Februar, ?=2: März/Arpil, ...).

7.3.5 Skripte

Im Rahmen dieser Arbeit ist es nötig geworden, verschiedene Skripte und Programme[78] zu erstellen. Diese dienten vorwiegend dazu, die vorhandenen Daten in ein für den jeweiligen Verarbeitungsschritt der Zeitreihenanalyse sinnvolles Format zu transformieren. Skripte dieser Art werden in drei Gruppen unterteilt: Skripte zur Extraktion der Rohdaten, Skripte zu Datenvorverarbeitung[79] und SAS-Skripte. Außerdem kommen noch Skripte hinzu, die das Erstellen und Auswerten Neuronaler Netze bzw. mathematischer Modelle steuern (sog. Modellskripte), sowie verschiedene andere Skripte unterschiedlicher Aufgabenbereiche, die unter der Rubrik „Hilfsprogramme und –skripte“ laufen. Im folgenden sollen nun die wichtigsten erstellten Skripte in den jeweiligen Rubriken kurz vorgestellt werden.

Der Ordner "Skripte" enthält alle wichtigen im Rahmen dieser Arbeit implementierten Skripte, die alle in der interpretierten Programmiersprache Perl 5.0 geschrieben sind.

7.3.5.1 Rohdatenextraktion

Im Rahmen dieser Arbeit kamen neben den selbsterstellten Zeitreihen Daten von drei verschiedenen Institutionen zum Einsatz. Während die Daten der Fachhochschule im Excel 97 Format vorlagen und keiner nennenswerten Verarbeitung bedurften, lagen die Daten der Biologie in einem und die des deutschen Wetterdienstes in drei verschiedenen Formaten vor, die eine entsprechende Weiterverarbeitung notwendig machten. Folgende Skripte kamen dabei zum Einsatz:

---

[78] Da vorwiegend die interpretierte Skriptsprache Perl 5.0 verwandt wurde, wird im folgenden einfach von Skripten gesprochen.

[79] Vgl. Kapitel 2.5.

`CO2_extract.pl:` Extrahiert aus den bis zu 366 CO_2-Dateien pro Jahr eine Zeitreihe mit den aggregierten Tageswerten.

`DWD_extract.pl`: Extrahiert aus den Dateien vom Deutschen Wetterdienst (`SY92.OUT`, `SY93.OUT` und `sy_94x96.dat`) alle Zeitreihendaten und legt sie in entsprechenden Dateien (z.B. WR.txt, BG.txt, etc.) ab.

`NS_extract*.pl:` Extrahiert die täglichen Niederschlagwerte für die Jahre 1992-1993 bzw. 1994-1996. Da die Niederschlagswerte nicht in den Dateien des Deutschen Wetterdienstes von 1992-1993 enthalten waren, greift dieses Skript auf die Datei `KL.OUT`[80] und die DWD-Datei `sy_94x96.dat` zu.

`WR_extract*.pl:` Extrahiert die tägliche Windrichtung für die Jahre 1992-1993 bzw. 1994-1996 nach dem in Kapitel 2.5.2 vorgestellten Verfahren. Dabei wird auf die Dateien `SY92.OUT`, `SY93.OUT` und s`y_94x96.dat` zugegriffen.

`WRG_extract*.pl:` Extrahiert die gewichtete Windrichtung für die Jahre 1992-1993 bzw. 1994-1996 nach dem in Kapitel 2.5.2 vorgestellten Verfahren. Dabei wird ebenfalls auf die Dateien `SY92.OUT`, `SY93.OUT` und s`y_94x96.dat` zugegriffen.

7.3.5.2 Sensemodel

`sensemodel.pl`: Das Skript zur Zeitreihenauswahl nach Sensitivität.

`BP.sense`: Das entsprechende BP-Template.

7.3.5.3 Vergleich REG-SNNS-KKM

`make_REG1.pl`: Beispielskript zum Finden der besten Kombination exogener Zeitreihen für die Regression. Dieses Skript generiert und wertet alle 1001 Regressionen für den Zeitraum Januar/Februar aus.

`make_SNNS1.pl`: Das analoge Skript zu Bestimmung der besten Topologie für Neuronale Netze. Dieses Skript generiert und wertet 48 verschiedene Topologien für den Zeitraum Januar/Februar aus.

[80] Siehe [Til 93], Anhang, Diskette 2.

7.4 Literaturverzeichnis

[Ana 98] **Ananieva, Maria:** *Modellierung wirtschaftlicher Zeitreihen mit Neuronalen Netzen.* Osnabrück, Universität: Magisterarbeit im Studiengang Computerlinguistik und künstliche Intelligenz, voraussichtlich Oktober 1998.

[Apt 96a] **Aptec Systems, Inc.:** *GAUSS Applications: Time Series*. 1996.

[Apt 96b] **Aptec Systems, Inc.:** *GAUSS Volume II: Command and Reference Manual*. Revised version, 1996.

[Apt 96c] **Aptec Systems, Inc.:** *GAUSS for UNIX*. Revised version, 1996.

[Apt 96d] **Aptec Systems, Inc.:** *GAUSS System and Graphics Manual*. Revised version, 1996.

[Bal 93] **Baloui, Said:** *Excel 5.0 – das Kompendium*. Haar bei München: Markt-und-Technik-Verlag, 1993. – ISBN 3-87791-549-3

[Bil 81] **Billeter, Peter**; **Vlach, Vladimir**: *Zeitreihen-Analyse: Einführung in die praktische Anwendung*, Würzburg: Physica-Verlag, 1981. – ISBN 3-7908-0257-3

[Box 76] **Box, George E. P.; Jenkins, Gwilym M.**; *Time series analysis: Forecasting and control. Revised edition.* San Francisco: Holden-Day, 1976. – ISBN 0-8162-1104-3

[Duf 92] **Dufner, Julius; Jensen, Uwe; Schumacher, Erich:** *Statistik mit SAS.* Stuttgart: Teubner, 1992 (Teubner Studienbücher Mathematik). – ISBN 3-519-02088-2

[Edw 84] **Edwards, Allen L.:** *An Introduction to Linear Regression and Correlation.* New York : W.H. Freeman and Company, 1984. – ISBN 0-7167-1593-5

[Erb 97] **Erben, Wilhelm:** *Statistik mit Excel 5 oder 7: Lehr- und Übungsbuch mit zahlreichen Excel-Beispieltabellen.* 2., durchges. Aufl. München: Oldenbourg Verlag, 1986. – ISBN 3-486-24043-9

[For 93] **Forstreuter, Manfred:** *Langzeitwirkungen der atmosphärischen CO_2-Anreicherung auf den Kohlenstoff- und Wasserhaushalt von Rotklee-Wiesenschwingelgemeinschaften.* Berlin, Technische Universität, Fachbereich 14, Diss., 1993. – ISBN 3-7983-1579-5

[Gra 95] **Graf, Alexander; Ortseifen, Carina:** *Statistische und grafische Datenanalyse mit SAS.* Heidelberg: Spektrum Verlag, 1995. – ISBN 3-86025-706-4

[Hal 98] **Hall, Joseph N.; Schwartz, Randal L.:** *Effective Perl programmning: Writing better programs with Perl*. Reading, Massachusetts: Addison-Wesley-Longman, 1998. – ISBN 0-201-41975-0

[Her 98] **Herbst, Vela:** Gefahr erkannt – doch nicht gebannt: Weltklimaveränderungen. In: *Wirtschaftsspiegel,* 3/1998 (Mai / Juni 1998), S. 2-5

[Kli 91] **Klir, George J.:** *Facets of systems science*. New York: Plenum Press, 1991. – ISBN 0-306-43959-X

[Kro 97] **Krone, Stephan:** *Eine interaktive Oberfläche zur grafischen Analyse von Zeitreihen.* Osnabrück, Universität, Fachbereich Mathematik/Informatik, Diplomarbeit, 1997

[Kul 79] **Kulla, Bernhard:** *Angewandte Systemwissenschaften.* Würzburg: Physica-Verlag, 1979. – ISBN 3-7908-0206-9

[Lei 86] **Leiner, Bernd:** *Einführung in die Zeitreihenanalyse*. 2., verb. Aufl. München: Oldenbourg Verlag, 1986. – ISBN 3-486-20299-5

[Mil 90] **Miller, Alan J.:** *Subset Selection in Regression.* London: Chapman and Hall, 1990 (Monographs on Statistics and Applied Probability 40), 1990 – ISBN 0-412-35380-6

[Nic 97] **Nicol, Natascha; Albrecht, Ralf:** *Wissenschaftliche Arbeiten schreiben mit WinWord 97: Formvollendete und normgerechte Examens-, Diplom- und Doktorarbeiten.* Bonn: Addison-Wesley-Longman, 1997. – ISBN 3-89319-981-0

[Pfe 97] **Pfeifer, Andreas; Schuhmann, Marco:** *Statistik mit SAS.* München: Oldenbourg Verlag, 1997. – ISBN 3-486-23953-8

[Reh 94] **Rehkugler, Heinz; Zimmermann, Hans Georg (Hrsg*):*** *Neuronale Netze in der Ökonometrie*. München: Verlag Vahlen, 1994. – ISBN 3-8006-1871-0

[Rip 93] **Ripper, Klaus:** *Theoretische und empirische Aspekte des Einsatzes Neuronaler Netze zur Prognose von Zeitreihen.* Darmstadt, technische Hochschule, Fachbereich Rechts- und Wirtschaftswissenschaften, Diplomarbeit, 1993

[Sch 97] **Schilli, Michael:** *Effektives Programmieren mit Perl 5: Objektorientierung, Graphische Oberflächen, Internet-Anwendungen.* Bonn: Addison-Wesley-Longman, 1997. – ISBN 3-8273-1095-4

[Sch 91] **Schlittgen, Rainer; Streitberg, Bernd H. J.:** *Zeitreihenanalyse*. 3. unveränderte Auflage. München: Oldenbourg Verlag, 1991. – ISBN 3-486-21908-1

[Sch 85] **Schrödter, Harald:** *Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden.* Berlin: Springer-Verlag, 1985. – ISBN 3-540-15355-1

[Sri 97] **Srinivasan, Sriram:** *Advanced Perl programming: Foundation and techniques for Perl application developers*. Cambridge: O'Reilly, 1997. – ISBN 1-565-92220-4

[Spi 90] **Spiegel, Murray R.:** *Statistik* / Schramm, Reginald; Heitmeyer, Carola (Übers., Bearb.). 2. überarb. und erw. Auflage. Hamburg: McGraw-Hill 1990: – ISBN 3-89028-908-8

[Thi 98] **Thiesing, Frank M.:** *Analyse und Prognose von Zeitreihen mit Neuronalen Netzen*. Osnabrück, Universität, Fachbereich Mathematik/Informatik, Diss.; Aachen: Shaker Verlag 1998. – ISBN 3-8265-4242-8

[Til 96] **Tillmanns, Peter:** *Globale und lokale Einflußfaktoren auf die bodennahe CO_2-Konzentration in einem urban geprägten Meßgebiet*. Osnabrück, Universität, Fachbereich Mathematik/Informatik, Diplomarbeit, 1996

[Ver 97] **Verbeke, Geert; Molenberghs, Geert:** *Linear mixed models in practice: A SAS-oriented approach*. New York: Springer, 1997 (Lecture notes on statistics No. 126). – ISBN 0-387-98222-1

[Vog 83] **Vogt, Roland:** *Die Systemwissenschaften: Grundlagen und wissenschaftstheoretische Einordnung*. Frankfurt am Main: HAAG+HERCHEN, 1983. – ISBN 3-88129-602-6

[Wal 96] **Wall, L., Schwartz, R. L.:** *Programming Perl*. Cambridge: O'Reilly, 1996. – ISBN 1-565-92149-6

[Win 97] **Winter, Rick et al.:** *Office 97 Professional: Referenz und Anwendungen; Outlook, Word, Excel, Access, PowerPoint*. Haar bei München: Markt-und-Technik-Verlag, 1997. – ISBN 3-8272-1018-6

[Zel 96] **Zell, Andreas:** *Simulation neuronaler Netze*. 1. unveränderter Nachdruck. Bonn: Addison-Wesley-Longman, 1996. – ISBN 3-89319-554-8

[Zel 95] **Zell, Andreas**: *SNNS: Stuttgart Neural Network Simulator User Manual*. Report No. 6/95. Version 4.1. University of Stuttgart, Germany, 1995 – (ftp://ftp.informatik.uni-stuttgart.de/pub/SNNS/SNNSv4.1.Manual.ps.gz)

[Zim 97] **Zimmerer, Thomas:** *Künstliche neuronale Netze versus ökonometrische und zeitreihenanalytische Verfahren zur Prognose ökonomischer Zeitreihen*. (Universität Regensburg, Diss.) Frankfurt am Main: Europäischer Verlag der Wissenschaften, 1997. – ISBN 3-631-32465-0

7.5 Danksagung

Mein Dank gilt an dieser Stelle allen, die zum Gelingen dieser Arbeit beigetragen haben:

- Herrn Prof. Dr. OLIVER VORNBERGER für die Betreuung der Arbeit und die Unterstützung bei der Beschaffung der Daten,
- Herrn Prof. Dr. MICHAEL MATTHIES für seine Bereitschaft zur Begutachtung der Arbeit,
- Herrn Dr. FRANK M. THIESING für die konstruktive Kritik und die hilfreichen Anmerkungen,
- Herrn JAN RIMMEK stellvertretend für den Fachbereich Gartenbau der Fachhochschule Osnabrück für hilfreiche und unbürokratische Unterstützung bei der Beschaffung der Daten der Fachhochschule,
- dem Deutschen Wetterdienst für die Bereitstellung der Wetterdaten der Meßstation Osnabrück,
- dem Fachbereich Biologie der Universität Osnabrück für die Bereitstellung der CO_2-Meßdaten
- sowie meinen Eltern und Freunden für das Korrekturlesen und die Unterstützung insbesondere während der „heißen“ Phasen dieser Arbeit.

7.6 Verpflichtungserklärung

Ich versichere, die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Osnabrück, den 11. September 1998

Jan Hannemann