UNIVERSITÄT SALZBURG

Master Thesis

im Rahmen des Universitätslehrganges

„Geographical Information Science & Systems“ (UNIGIS MSc)

am Zentrum für GeoInformatik (Z_GIS)

der Paris-Lodron-Universität Salzburg

zum Thema

Ein Vergleich räumlicher Interpolationsverfahren anhand von Schwefeldaten des steirischen Bioindikatornetzes

vorgelegt von

Dipl.-Ing. Wolfgang Christian Böheim

u1047, UNIGIS MSc Jahrgang 2003

zur Erlangung des Grades

„Master of Science (Geographical Information Science & Systems) – MSc(GIS)“

Gutachter:

Ao. Univ. Prof. Dr. Josef Strobl

Salzburg, 24. April 2006

„Erst wenn der letzte Baum gefällt,
der letzte Fluss vergiftet und
der letzte Fisch gefangen ist,
werdet ihr merken,
dass man Geld nicht essen kann.“

(Prophezeiung der Cree-Indianer)

Eidesstattliche Erklärung

Ich erkläre, dass ich die vorliegende Master Thesis selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Ich versichere, dass ich das Thema dieser Arbeit bisher weder im Inland noch im Ausland in irgendeiner Form vorgelegt habe.

Salzburg, 24. April 2006

Danksagung

An dieser Stelle möchte ich mich bei all jenen bedanken, dir mir bei der Erstellung dieser Arbeit durch ihre Unterstützung zur Seite gestanden sind.

Allen voran Ao. Univ. Prof. Dr. Josef Strobl der mir die nötige Freiheit bei der Themenwahl, sowie bei der Vertiefung in einzelne Teilgebiete gab und stets ein offenes Ohr auch für nicht fachliche Fragen hatte. Ein weiterer Dank geht an die Lehrgangsleitung für ihre Geduld mit mir und der Ausräumung so mancher organisatorischer Hürde.

Ein besonderer Dank geht an die Fachabteilung für das Forstwesen des Landes Steiermark, dem Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft sowie der Landesbaudirektion-GIS Steiermark für die Bereitstellung der Schwefeldaten sowie Geobasisdaten. Ohne diese Daten wäre die Durchführung der vorliegenden Arbeit nicht möglich gewesen.

Für die Durchsicht der Arbeit, sowie ihren kritischen Anmerkungen danke ich Dipl.-Ing. Josef Ringert und Mag. Christian Debelak.

Nicht vergessen möchte ich in diesem Zusammenhang meine Eltern, die mir immer ein wichtiger Rückhalt waren und mich in allen Lebenslagen unterstützten – herzlichen Dank.

Zu guter Letzt noch ein Dank an all jene, die spätestens alle 24 Stunden gefragt haben, ob die Arbeit nicht schon fertig sei – nun ist sie es!

Salzburg, 24. April 2006 Wolfgang Christian Böheim

Kurzfassung

Ziel der vorliegenden Arbeit war es festzustellen, welches von drei ausgewählten Interpolationsverfahren sich am Besten zur Schätzung der räumlichen Verteilung der Schwefelbelastung in der Steiermark eignet, und inwieweit sich auf Basis des Verfahrens mit der höchsten Qualität Aussagen über potentiell gefährdete Regionen für die Schwefelbelastung treffen lassen. Bei den zu analysierenden Interpolationsverfahren handelte es sich um die Methode der Inversen Distanzgewichtung, dem Ordinary Kriging und dem Ordinary Cokriging.

Als Datengrundlage für die Durchführung der räumlichen Interpolation dienten Schwefeldaten des steirischen und österreichischen Bioindikatornetzes, welche vom Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft sowie der Fachabteilung für das Forstwesen des Landes Steiermark zur Verfügung gestellt wurden. Die Schwefeldaten lagen sowohl für den ersten als auch zweiten Nadeljahrgang der Bioindikatornetzbäume vor. Während die Daten des ersten Nadeljahrgangs bei allen Interpolationsverfahren zur Schätzung der Schwefelwerte herangezogen wurden, gingen die des zweiten Nadeljahrgangs als Zusatzinformation in das Ordinary Cokriging ein.

Um die Qualität einer räumlichen Interpolation beurteilen zu können ist eine intersubjektiv nachvollziehbare Prüfung unverzichtbar. Die Beurteilung der Qualität erfolgte mit Hilfe der Kreuzvalidierung. Die Grundidee dieser Methode besteht darin, aus einer vorhandenen Stichprobe eines Parameters einen gemessenen Wert herauszunehmen und diesen mittels der anderen Werte der Stichprobe unter Verwendung des jeweiligen Interpolationsverfahrens zu schätzen. Dadurch liegt für jeden Stützpunkt der Interpolation sowohl ein gemessener als auch geschätzter Wert vor, deren Differenz den lokalen Schätzfehler (Residuum) bildet. Diese Schätzfehler bilden wiederum die Ausgangsbasis für die Auswertung der Kreuzvalidierung anhand von statistischen Kennzahlen wie z.B. Schiefekoeffizient, Rangkorrelationskoeffizient, Mean Square Error (MSE) oder Mean Absolute Error (MAE).

Im Zuge der statistischen Auswertung der Kreuzvalidierungen für die drei Interpolationsverfahren konnte festgestellt werden, dass das Ordinary Cokriging die besten Schätzergebnisse für die Schwefelbelastung liefert.

Abstract

The aim of the present thesis was the evaluation of three selected interpolation methods for an estimation of the spatial distribution of sulphur exposure in Styria. The best method should be identified to provide a solid basis for the provision of possibilities for qualitative and reliable predications regarding potential vulnerable regions. The methods which are analysed within this thesis are the method of Inverse Distance Weighting, the Ordinary Kriging, and the Ordinary Cokriging.

Sulphur data out of the Styrian and Austrian bioindicator grid provided by the Federal Research and Training Centre for Forests, Natural Hazards and Landscape as well as the Department of Forestry of Styria are used as basis for the development of the spatial interpolation,. The sulphur data were available for the first and also the second needle class of the bioindicator grid trees. The data of the first needle class were used for all three interpolation methods for the estimation of the sulphur values. The sulphur data of the second needle class were used as additional information for Ordinary Cokriging only.

For the evaluation of the quality and reliability of a spatial interpolation, an inter-subjective traceable testing is indispensable. Thus, the evaluation of quality was carried out by the use of a so called “cross-validation”. The basic concept of this method is based on the extraction of a measured value out of an existing sample of a parameter and the estimation of the extracted value by the use of the whole set of values of the sample applying the three different interpolation methods. Thus, for every interpolation point, a measured and also estimated value is available. The difference between the estimated and the measured value results in the local estimation error or residual. The residuals provide the basis for the analysis of the cross-validation using statistical classification numbers like skewness coefficient, rank correlation coefficient, mean square error (MSE), or mean absolute error (MAE).

In the course of the statistical analysis of the cross-validation for the three different interpolation methods it could be ascertained that Ordinary Cokriging provides the best estimation results for the sulphur exposure.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Abkürzungsverzeichnis

A

ASI	Agenzia Spaziale Italiana

B

BFW	Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft
BGBl	Bundesgesetzblatt
BIN	Bioindikatornetz
BLUE	Best Linear Unbiased Estimator
BLUP	Best Linear Unbiased Predictor
BN	Bundesnetz

C

CIAT	Centro Internacional de Agricultura Tropical
CORINE	Coordination of Information on the Environment

D

dBASE	DataBase
DGPF	Deutsche Gesellschaft für Photogrammetrie und Fernerkundung
DLR	Deutsches Zentrum für Luft- und Raumfahrt

E

EDV	Elektronische Datenverarbeitung
ESRI	Environmental Systems Research Institute, Inc.
EU-WBS	EU-Punkt Waldschadensbeobachtungssystem

F

FAFW	Fachabteilung Forstwesen (Forstdirektion Steiermark)

G

GIS	Geographisches Informationssystem

I

ICP	International Cooperative Programme
IDW	Inverse Distance Weighting (Inverse Distanzgewichtung)
IR	Infrarot

L

LN	Lokalnetz

M

MA	Moving Average
MAE	Mean Absolute Error (Mittlerer absoluter Schätzfehler)
MSE	Mean Square Error (Mittlerer quadrierter Schätzfehler)

N

NASA	National Aeronautics and Space Administration
NIMA	National Imaging and Mapping Agency
NO_x	Stickoxide

O

O_3	Ozon
OK	Ordinary Kriging
OCK	Ordinary Cokriging

P

PE	Polyethylen

Q

QQ	Quantil-Quantil

S

SF	Schätzfehler
SO_2	Schwefeldioxid
SRTM	Shuttle Radar Topography Mission

V

VN	Verdichtungsnetz

W

WBS	Waldschadensbeobachtungssystem

1 Einleitung

Das Absterben von ganzen Wäldern ist kein Problem des 21.Jahrhunderts. Schon im 19.Jahrhundert warnten Wissenschaftler vor Schäden durch eingetragene Luftschadstoffe und der schleichenden Vergiftung des Bodens. In den siebziger und achtziger Jahren des 20.Jahrhunderts nahmen die Schäden in den Mittelgebirgslagen und Alpen so dramatisch zu, dass Forscher eine flächendeckende Entwaldung befürchteten. Dieses Schreckensszenario ist zwar nicht eingetroffen, allerdings hat sich seit damals der Zustand der Wälder nicht wesentlich verbessert, weshalb von einer Entwarnung nach wie vor keine Rede sein kann.

Der Anteil geschädigter Wälder hat sich zwar in den letzten Jahren stabilisiert, allerdings auf einem hohen Niveau. Nach einer Studie der EUROPÄISCHEN KOMMISSION (2000) waren im Jahr 1999 nahezu ein Viertel (22%) aller erhobenen Bäume in ganz Europa mittel oder schwer geschädigt. Fast 1% der Bäume war tot, 41% wiesen eine geringfügige Verlichtung auf, und über ein Drittel (36%) wurden als gesund eingestuft.

Die Ursachen für die Waldschäden sind sehr komplex, da regional sehr unterschiedliche Voraussetzungen bzw. biotisch – abiotisch bedingte Stressmuster herrschen (SMIDT 2004, MAYER 1992). Als erwiesen gilt aber, dass Immissionen das System Baum-Boden-Luft nachhaltig stören können. Zu den wichtigsten Luftschadstoffen in diesem Zusammenhang zählen Schwefeldioxid (SO_2) und Stickoxide (NO_x), die bei Verbrennungs- bzw. Hochtemperaturprozessen freigesetzt werden (MAYER 1985, SCHÜTT et. al 1984). In Verbindung mit Sauerstoff werden SO_2 und NO_x zu Schwefelsäure und Salpetersäure bzw. salpetrige Säure (DUTZ et al. 2002). Erst diese Säuren wirken als schwere Gifte auf die Bäume bzw. Pflanzen. Unter Einfluss von Sonnenlicht bildet sich zusätzlich aus den Stickoxiden Ozon (O_3), welches ebenfalls ein Pflanzengift darstellt. Säuren und Ozon zerstören die schützende Wachsschicht auf Nadeln und Blättern (TRIMBACHER 1997, SAUTER et al. 1987, KARHU & HUTTUNEN 1986). Weiters verhindern sie, dass sich die Spaltöffnungen an den Nadeln schließen wodurch der Baum den Wasserhaushalt nicht mehr regulieren kann und somit lebensnotwendige Feuchtigkeit verloren geht (TRIMBACHER & DITRICH 1989). Zugleich kommt es durch die Einwirkung von Säuren und Schwermetallen zu einer Versauerung des Bodens welche zu einer Schädigung der Wurzeln führt. Hinzukommt, dass durch die Auswaschung von Metallsalzen im Boden giftige Metallverbindungen entstehen (SCHWEIKLE 1997, DUTZ et al. 2002). Diese schädigen zusätzlich die Bodenorganismen, welche für die

Nährstoffaufnahme der Baumwurzeln unentbehrlich sind. Der ober- und unterirdisch von Giften angegriffene Baum kann also nicht genug Nährstoffe und Wasser aufnehmen, verdunstet aber, aufgrund der oben angeführten Gründe, mehr als im gesunden Zustand. Ein so geschwächter Baum widersteht einem besonders trockenen Sommer kaum noch. Abbildung 1-1 zeigt zusammenfassend die Ursachen für das Waldsterben.

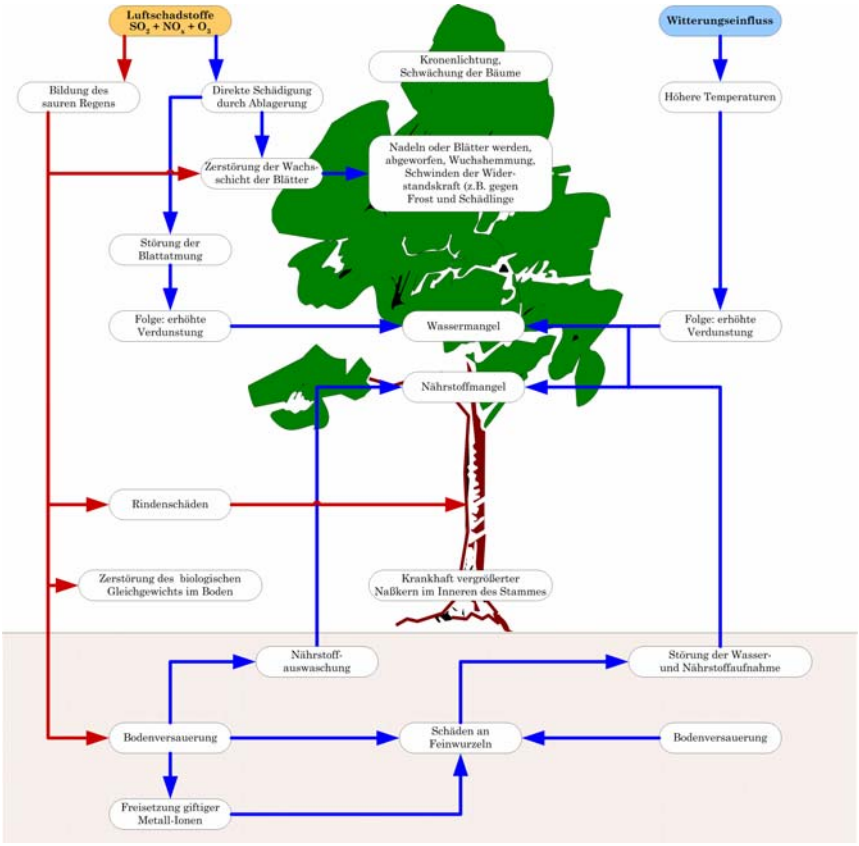


***Abbildung 1-1**: Kausalkette für das Waldsterben (MERGLER 2004)*

1.1 Problemstellung

Um die Wirkung von Schadstoffen z.B. von Schwefel auf Pflanzen (Bäume) sicher einschätzen zu können reichen physikalisch-chemische Messungen der Umweltmedien Wasser, Boden und Luft nicht aus. Ergänzend sind deshalb Verfahren des Biomonitorings und der Bioindikation notwendig. Unter dem Begriff Biomonitoring werden alle aktiven und passiven Monitoringsysteme zusammengefasst, die den Zustand und Veränderung der Umwelt anhand der resultierenden Auswirkungen auf einen lebenden Organismus bewerten (DÄSSLER 1991). Um diesen lebenden Organismus nicht zu schädigen oder gar zu töten, werden Aussagen über einen Parameter meist über eine indirekte

Bewertung (Bioindikator) abgeleitet. Unter einem Bioindikator versteht man in diesem Zusammenhang Organismen, die auf einen äußeren Reiz und im Speziellen auf eine Schadstoffbelastung mit Stoffwechseländerungen reagieren, wobei auch eine Schadstoffanreicherung, die über das normale Maß hinausgeht als Reaktion zu betrachten ist (SMIDT 2004). Am Besten als Bioindikatoren eigenen sich Pflanzen, weil sie standortsgebunden sind und meist in großer Individuenzahl natürlich vorkommen. Diese Art des Biomonitorings wird als passiv bezeichnet, da die Feststellung von Schäden im natürlichen Lebensraum bzw. am Wuchsort erfolgt.

In vielen, eigentlich in nahezu allen Fällen können räumliche Phänomene wie z.B. die Schwefelbelastung aus Kosten und aus prinzipiellen Gründen nicht vollflächig räumlich messend erfasst werden, weshalb man punktuelle Messungen in Form von Messstellen, Probepunkten oder Stichprobenrastern durchführt (BLASCHKE & STROBL 2003). Das Aussageziel der Erhebung bezieht sich jedoch meist auf das gesamte Untersuchungsgebiet in seiner flächigen Erstreckung und nicht nur auf die Standorte der jeweiligen Messeinrichtungen. Man steht deshalb vor dem verbreiteten und grundlegenden Problem, punktuelle Aussagen mit bestmöglicher Qualität auf den zwei- oder dreidimensionalen Raum zu übertragen.

Räumliche Interpolationsverfahren bieten die Möglichkeit, auf Basis von punktuell gemessenen Daten Aussagen über ihre flächenhafte Verteilung zu treffen. Unter Interpolation versteht man die Zuweisung eines berechneten Wertes zu nicht direkt gemessenen Punkten unter Annahme des Verlaufs zwischen den bekannten Punkten (BILL 1996). Je nach Charakter des gegenständlichen Phänomens wird man sich geeigneter Interpolationstechniken bedienen müssen, wobei für die Auswahl des Verfahrens das Datenniveau, der räumliche Variationsgrad, die Genauigkeit der Punktmessung sowie das Aussageziel ausschlaggebend ist.

In der Fachliteratur ist eine Vielfalt an räumlichen Interpolationsmethoden zu finden, die sich in ihren Ansätzen unterscheiden. So kann man z.B. das nichtstatistische Inverse Distance Weighting-Verfahren (IDW) und das geostatistische Interpolationsverfahren Kriging unterscheiden. Im Gegensatz zu dem nichtstatistischen Verfahren, bei dem der räumliche Zusammenhang zwischen den Daten rein intuitiv angenommen wird, liegt dem geostatistischen Verfahren ein geostatistisches Modell zugrunde. Dieses ermöglicht es, den räumlichen Zusammenhang zwischen den Werten in Abhängigkeit zur betreffenden Beobachtungsvariablen festzulegen und die im Gelände wirkenden Einflussfaktoren bei der Interpolation zu berücksichtigen. Dieser Ansatz verspricht genauere Schätzergebnisse (Hinterding 1998). Im Zusammenhang mit speziellen Fragestellungen ergibt sich aus der Fülle vorhandener Ansätze und verschiedener Interpolationsverfahren die Frage nach der Wahl eines bestimmten Verfahrens. Je nach fachlichem Hintergrund ist zu entscheiden, welches der verschiedenen Verfahren angewendet werden soll.

1.2 Zielsetzung und Gliederung der Arbeit

In der vorliegenden Arbeit wird der Frage nachgegangen, welches von drei ausgewählten Interpolationsverfahren sich am Besten zur Schätzung der räumlichen Verteilung der Schwefelbelastung in der Steiermark eignet und inwieweit sich auf Basis des Verfahrens mit der höchsten Qualität (Güte) Aussagen über potentiell gefährdete Regionen für die Schwefelbelastung treffen lassen. Ziel ist es, die Qualität der verschiedenen Verfahren zur Schätzung der Schwefelbelastung anhand der Schätzfehler vergleichend zu ermitteln. Auf Basis des Interpolationsverfahrens, welches im Vergleich die höchste Qualität aufweist, sollen in weiterer Folge Schwefelbelastungszonen ausgeschieden und in Form einer Schwefelbelastungskarte dargestellt werden. Die entsprechenden Analysen werden anhand der Schwefeldaten des steirischen Bioindikatornetzes durchgeführt.

Die vorliegende Arbeit gliedert sich in folgende Kapitel:

Kapitel 2 widmet sich zunächst kurz der Erhebungs- und Auswertemethodik innerhalb des Österreichischen Bioindikatornetzes um danach einen detaillierten Überblick über das zur Verfügung stehende Datenmaterial zu geben.

Im Anschluss daran wird in **Kapitel 3** auf die Methodik der verwendeten räumlichen Interpolationsverfahren (Inverse Distanzverfahren, Ordinary Kriging, Ordinary Cokriging) eingegangen. Um die Qualität der Schätzung durch die einzelnen Interpolationsverfahren bewerten zu können wird die Methode der Kreuzvalidierung angewandt. Die Grundlagen dieses Verfahrens, sowie die zur Auswertung der Kreuzvalidierung herangezogenen statistischen Kennwerte sind ebenfalls Gegenstand dieses Kapitels.

In **Kapitel 4** werden die Schätzergebnisse der einzelnen Interpolationsverfahren vorgestellt, interpretiert, miteinander verglichen und in Form einer sogenannten Schwefelbelastungskarte dargestellt.

Die Ergebnisse der Qualitätseinschätzung der Interpolationsverfahren werden in **Kapitel 5** diskutiert.

Kapitel 6 beinhaltet die Zusammenfassung der Arbeit und gibt einen Ausblick, wie das Interpolationsergebnis verbessert werden kann und welche der gewonnenen Erkenntnisse eventuell für andere Anwendungsgebiete relevant sein könnten.

2 Bioindikatornetz und Datengrundlage

In diesem Kapitel wird zunächst die Erhebungs- und Auswertemethodik innerhalb des Österreichischen Bioindikatornetzes vorgestellt und anschließend eine detaillierte Übersicht über das für die räumliche Interpolation zur Verfügung stehende Datenmaterial gegeben.

2.1 Bioindikatornetz

Waldschädigende Luftverunreinigungen und ihre negativen Auswirkungen auf den Wald werden seit mehr als 150 Jahre erforscht (MAYER 1992). Schon damals wiesen Forstwissenschafter und Chemiker auf „Rauchschäden" an Nadelbäumen in der Umgebung von Industriebetrieben hin. Als eine der ersten wissenschaftlichen Untersuchungen im deutschsprachigen Raum zu diesem Thema gilt die Arbeit von SCHROEDER & REUSS (1883) über die *„Auswirkungen der Emission von Schwefelsäuredämpfen auf die Vegetation des Oberharzer durch die lokalen Hüttenbetriebe*". Mit der industriellen Revolution erreichte die Rauchschadensforschung ihren ersten Höhepunkt. In Österreich wurde die chemische Pflanzenanalyse erstmals um die Jahrhundertwende des 20. Jahrhunderts zum Nachweis von Schwefel-Immissionseinwirkungen eingesetzt (PORTERLE 1891, RUSNOV 1910, 1917).

Die zunehmende Industrialisierung und das Bestreben, die Nahwirkung der Emissionen durch den Bau von höheren Schornsteinen zu vermindern, führten insbesondere seit Mitte der 60er Jahre des vorigen Jahrhunderts zu überregionalen Schäden (SMIDT et al. 2004). Besonders durch die „Politik der hohen Schornsteine" wurden Schadstoffe in weit entlegene Gebiete verfrachtet. Die geschädigte Waldfläche stieg dramatisch an. Als anthropogene Quellen für die Luftschadstoffe wurden vor allem Industrie, Verkehr und Tierhaltung erkannt.

In den Jahren 1955 bis 1980 erfolgte die pflanzenanalytische Untersuchung der Nadel- und Blattproben bereits auf 7% der österreichischen Waldfläche. Der Schwerpunkt der Erhebungen beschränkte sich dabei allerdings auf den Nahbereich von potentiellen Emissionsquellen, weshalb flächendeckende Aussagen zur Schwefelbelastung der Wälder nicht möglich waren. Die chemische Pflanzenanalyse wurde 1975 im Rahmen des österreichischen Forstgesetzes (FORSTGESETZ 1975) als Mittel zum Nachweis von Immissionseinwirkungen

gesetzlich verankert. Die Festlegung entsprechender Grenzwerte für die Beurteilung der Immissionseinwirkungen erfolgte 1984 durch die Verordnung „Forstschädliche Luftverunreinigungen" (BUNDESGESETZBLATT 1984).

Durch den 1983 erschienen „Spiegel"-Artikel „*Säureregen – da liegt was in der Luft; schwefelhaltige Niederschläge vergiften Wälder, Atemluft und Nahrung*" wurde die Öffentlichkeit und die Politik aufgerüttelt. In der Folge wurden europaweite Forschungskooperationen aufgebaut. In Österreich erfolgte die Errichtung von bundesweiten Netzen zur Überwachung des Gesundheitszustands der Wälder und deren Entwicklung (Waldzustandsinventur, Waldschadensbeobachtungssystem mit Waldbodenzustandsinventur, ICP Forests Programm Level I) sowie zur Überwachung der Schadstoffeinträge und Immissionseinwirkungen (Bioindikatornetz, ICP Forests Programm Level II).

Das Österreichische Bioindikatornetz (BIN) wurde 1983 als bundesweites, flächendeckendes Monitoringnetz eingerichtet (FÜRST 2004, 2001a). Als Bioindikatoren (passive Akkumulationsindikatoren) wurden die Fichte bzw. im trockenen Ostteil Österreichs, mangels geeigneter Fichtenflächen, auch die Baumarten Weißkiefer, Schwarzkiefer und Buche festgelegt (FÜRST 2001b). Da Nadelbäume ihre Blätter über mehrere Vegetationsperioden hinweg behalten, eignen sich diese vor allem für Akkumulationsuntersuchungen mit langen Zeitreihen. So repräsentieren die analysierten Schadstoffgehalte einen längeren Zeitraum und integrieren damit saisonale Schwankungen von Immissionen zu einem mittleren Belastungsniveau. Um flächenbezogene Aussagen zur Immissionsbelastung, sowie zur Nährstoffversorgung machen zu können, wurde ein systematisches Grundnetz mit einem Raster von 16x16km (2002: 293 Aufnahmepunkte) eingerichtet, welches direkt an das Bayrische Bioindikatornetz angekoppelt ist (SMIDT 2004).

1983 wurde das Bioindikatornetz im Auftrag des Bundesministeriums für Land- und Forstwirtschaft (dem heutigen Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft) in Zusammenarbeit mit den Landesforstbehörden erstmals beprobt (FÜRST 2001b). Zur detaillierteren Darstellung der Ergebnisse, zur Zonierung und zur Feststellung von Entwicklungen auf Bundesländerebene bzw. auf der Ebene der Bezirksforstinspektionen sowie zur Beurteilung kleinräumiger Veränderungen wurde das Grundnetz 1985 im Flachland systematisch verdichtet (2002: 483 Verdichtungspunkte) und im Gebirge den topographischen Verhältnissen angepasst (FÜRST 2004, LICK et al. 1998). Im Jahr 2002 wurden in diesem Netz 776 Punkte beprobt, davon 161 in der Steiermark.

In der Steiermark werden zusätzlich zu diesem Bundesnetz (BN) ein Verdichtungsnetz (VN) mit 785 Probepunkten sowie in der Nähe von potentiellen Emittenten ca. 40 sogenannte Lokalnetze (LN) mit insgesamt 746 Probepunkten betrieben. Abbildung 2-1 zeigt die räumliche Verteilung der Grundnetzpunkte des

Österreichischen Bioindikatornetzes am Beispiel des Jahres 2003. Auf jedem dieser Punkte sind zwei herrschende oder vorherrschende Probebäume eingerichtet.

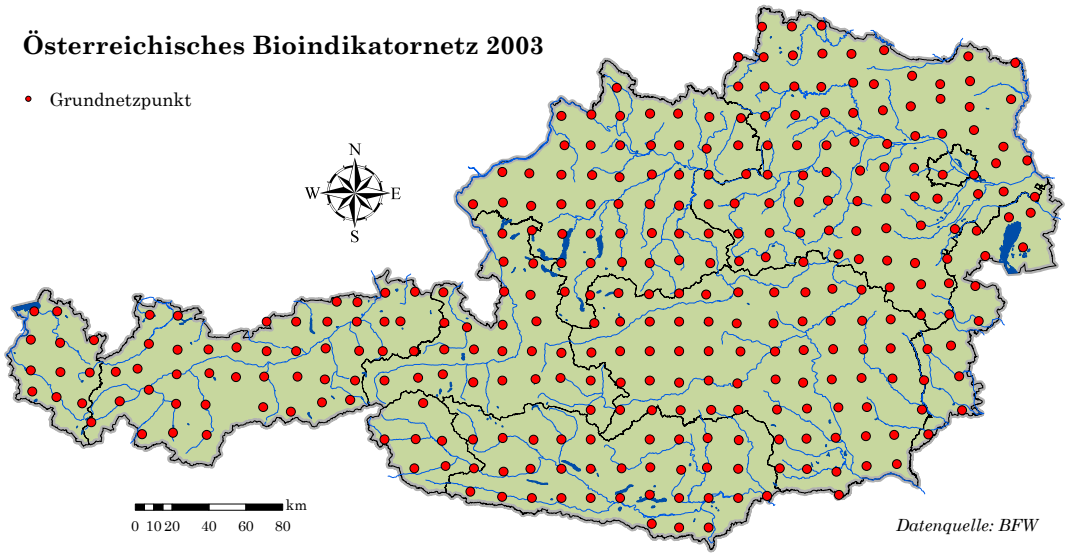


Abbildung 2-1: *Grundnetz des Österreichischen Bioindikatornetzes 2003*

Die Beprobung des Bioindikatornetzes wird vom September bis November des jeweiligen Jahres, gemäß den Bestimmungen der „*Zweiten Verordnung gegen Forstschädliche Luftverunreinigungen*“ durch die Landesforstdienste durchgeführt. Die Unterlagen für die Probenahme werden vom Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft (BFW) erstellt und den Ländern für die Probenahme zur Verfügung gestellt (FÜRST 2001b). Die Probenahme erfolgt bei den Nadelbäumen im obersten Kronendrittel, d.h. im Bereich des 6. bis 7. Astquirls (vgl. Abbildung 2-2).

Abbildung 2-2: *Probennahme an Nadelbäumen (Foto: Dr. Stefan Smidt, BFW)*

Bei den Laubbäumen wird die Probenahme in Form einer Mischprobe unter Einschluss der am Vegetationsbeginn gebildeten Blätter aus dem oberen Kronenbereich durchgeführt (SMIDT 2004). Die Äste der Nadel- und Laubbäume werden vor Ort in die Nadeljahrgänge 1 (heuriger Trieb) und 2 (Austrieb des Vorjahres) aufgetrennt, in ein Polyethylen(PE)-Säckchen verpackt und dem BFW zur weiteren Bearbeitung übermittelt.

Nach der Eingangskontrolle beim BFW werden die Proben an der Abteilung für Pflanzenanalyse zur Konservierung bei -15°C tiefgefroren, anschließend bei ca. 80°C im Umlufttrockenschrank getrocknet, von den Holzteilen befreit und vermahlen. Das Probenpulver wird in ein PE-Fläschchen aufbewahrt. Unmittelbar vor der Analyse erfolgt die Trocknung eines Probenaliqoutes bei 105°C. Die Bestimmung des Schwefelgehalts erfolgt mit einem Schwefelanalysator welcher nach dem Messprinzip der nicht dispersiven Infrarot-Detektion arbeitet.

Bei der Probenauswertung werden 280 bis 320 mg ±1mg Probenmaterial in ein Keramikschiffchen eingewogen, mit Quarzsand (ca. 0,5g) überschichtet und mit Sauerstoff bei 1400°C verbrannt. Die Verbrennungsprodukte werden getrocknet und das Schwefeldioxid in einer Infrarot(IR)-Messzelle detektiert. Danach wird aus dem Messergebnis und der Einwaage der Schwefelgehalt errechnet.

Um die Vergleichbarkeit der Daten aller Untersuchungsjahre zu gewährleisten und die Rückführbarkeit der Ergebnisse auf internationale Normale herzustellen, erfolgt eine kontinuierliche Methodenevaluierung: Anfang der 80er Jahre des 20.Jahrhunderts mit den Standardreferenzmaterialien *Citrus Leavus* sowie *Pine Needles* des amerikanischen NATIONAL BUREAU OF STANDARDS (1976, 1982) und am Ende der 80er Jahre des 20.Jahrhunderts mit dem vom europäischen *Institute of Reference Materials and Measurements* hergestellten zwei Standardreferenzmaterialien *Beech Leaves* und *Spruce Needles* (MAIER et al. 1989)

Derzeit werden in Österreich über 3000 Nadelproben pro Jahr untersucht. Dabei erfolgt nicht nur die Bestimmung der Schwefelgehalte, sondern auch eine Analyse hinsichtlich der Nährstoffe Stickstoff, Phosphor, Kalium, Calcium, Magnesium, Eisen, Mangan und Zink. In der Nähe von Emittenten werden zusätzlich die Elemente Fluor, Chlor, Kupfer, Blei und Cadmium analysiert. Sämtliche Proben werden seit 1983 in einer Datenbank archiviert deren Inhalt online abgefragt werden kann (FÜRST 2005b).

2.2 Datengrundlage

Um vorab eine Vorstellung vom Untersuchungsgebiet zu erhalten, wird dieses zunächst kurz vorgestellt und danach auf das verwendete Datenmaterial eingegangen.

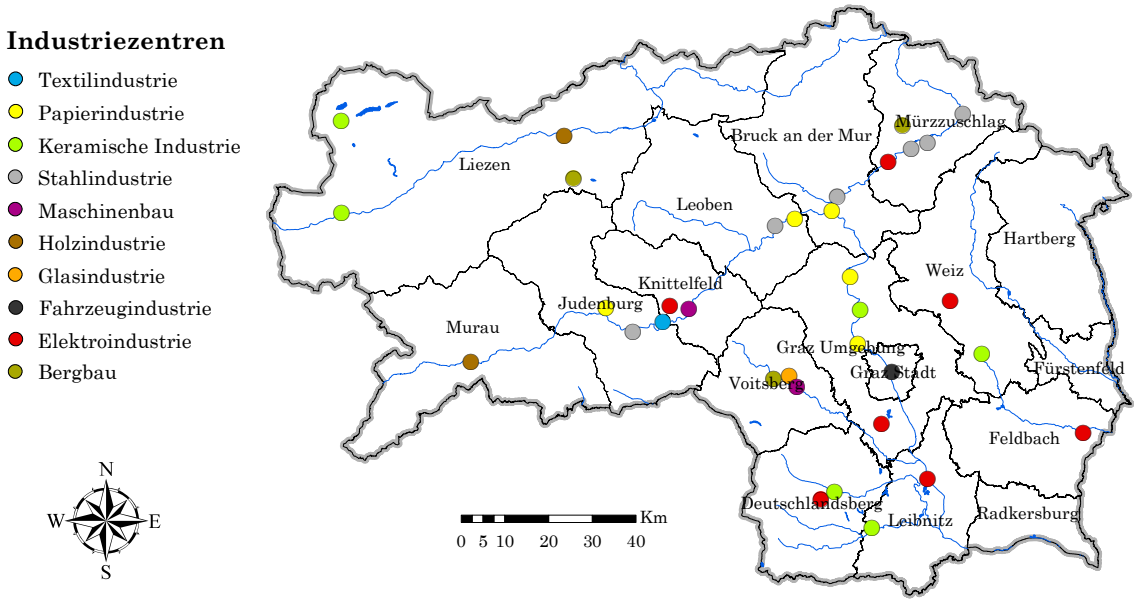
2.2.1 Untersuchungsgebiet

Das Bundesland Steiermark umfasst eine Fläche von etwa 16.400km² mit einer Nord-Süd-Erstreckung von ca. 130km und einer maximalen West-Ost-Erstreckung von 190km. Die Steiermark grenzt im Osten an das Burgenland, im Norden an Nieder- und Oberösterreich, im Westen an Salzburg und Kärnten und im Süden an Slowenien. Die Obersteiermark wird im Norden von den Kalkalpen begrenzt. Hier liegt an der Grenze zu Oberösterreich auch der höchste Berg des Landes, der Hohe Dachstein (2996m). Als Rückgrat der Steiermark kann man die Kette der Niederen Tauern bezeichnen, die sich aus den Schladminger, Wölzer, Rottenmanner und Seckauer Tauern zusammensetzen. Im Westen des Grazer Beckens bilden die Pack-, Stub- und Gleinalpe den Übergang zur Mittelsteiermark.

Die Steiermark besteht zu einem Großteil aus Berg- und Hügelland und ist mit einem Waldanteil von über 61% (1 Mio. ha) das waldreichste Bundesland Österreichs. Gemessen an der Ertragswaldfläche von rund 870.000ha sind die Baumarten Fichte (61,9%), Lärche (6,1%), Kiefer (3,9%) und Buche (6,5%) die wichtigsten Baumarten.

Klimatisch gehört die Steiermark zum europäischen Übergangsklima mit alpinen Einflüssen im Norden und Nordwesten, sowie pannonischen Einflüssen im Süden und Südosten. Diese Einflüsse spiegeln sich auch in der Verteilung der Jahresniederschlagsmenge und der mittleren Jahrestemperatur wieder. Während in der Obersteiermark teilweise mit bis zu 2000mm (Altaussee) Jahresniederschlag zu rechnen ist, muss die Vegetation in der Südsteiermark mit 900 mm oder weniger auskommen. Die kältesten Gebiete in der Steiermark sind das obere Murtal und die Umgebung um Admont. Im Gegensatz dazu gehört die Südoststeiermark zu den wärmsten Gebieten Österreichs.

Die Steiermark ist ein traditionelles Industrieland wobei sich die Industriezentren vor allem in den Regionen Mur-Mürz, Graz und Voitsberg-Köflach befinden. In diesen Regionen findet man die wichtigsten Produktionsstandorte für die Eisen- und Stahlindustrie, dem Maschinenbau, der Elektro- und Elektronikindustrie, der Papier-Zellstoff- und Kartonproduktion, der Glasindustrie, der Stein- und keramischen Industrie sowie dem Bergbau. Abbildung 2-3 gibt eine Übersicht über die räumliche Verteilung der wichtigsten Industriezentren in der Steiermark.



Datenquelle: GIS Steiermark

Abbildung 2-3: *Industriezentren der Steiermark*

2.2.2 Geodaten

Für die Entwicklung der Interpolationsverfahren bzw. für die graphische Darstellung der Interpolationsergebnisse, in Form von Schwefelbelastungskarten, stehen neben den Schwefeldaten folgende Geodaten zur Verfügung:

- **Corine Land Cover Daten**

  Die Landnutzungskategorien der Corine Land Cover Daten bildeten die Grundlage für die Ausscheidung der bewaldeten Flächen in der Steiermark, mittels derer eine Waldmaske erstellt wurde. Auf Basis dieser Waldmaske werden am Ende der Arbeit die Interpolationsergebnisse graphisch dargestellt.

- **SRTM3 Daten**

  Aus den frei verfügbaren SRTM3 Daten wurde für das Untersuchungsgebiet bzw. für die Nachbarbezirke ein digitales Landschaftsmodell (Oberflächenmodell) mit einer Rasterauflösung von 3 Bogensekunden erzeugt, welches als Zusatzinformation für das Ordinary Cokriging dient.

- **Sonstige**

  Zu den sonstigen Geodaten gehören z.B. Bundesländergrenzen, Bezirksgrenzen, Flüsse, Seen oder Industriezentren. Diese Daten wurden teilweise selbst erstellt bzw. von GIS-Steiermark zur Verfügung gestellt.

Im Folgenden wird kurz auf die oben angeführten Daten eingegangen.

a) CORINE Land Cover Daten

CORINE, Synonym für eine „koordinierte Erfassung von Informationen über die Umwelt“ (Coordination of Information on the Environment), ist ein von der Europäischen Union im Jahr 1985 gegründetes Projekt zur Erfassung von Umweltdaten. Die Entwicklung wurde mit dem Fehlen von vergleichbaren, umweltrelevanten Daten innerhalb der Europäischen Union begründet. Das Corine Land Cover Programm umfasst den gesamten europäischen Raum und einige nordafrikanische Staaten, wobei mit der Bearbeitung verschiedene Stellen in den einzelnen Ländern beauftragt wurden.

Die Bodenbedeckung bzw. Landnutzung wurde entsprechend den Vorgaben der EUROPEAN ENVIRONMENT AGENCY (1993) im Arbeitsmaßstab 1:100.000 auf der Basis von Landsat 5 Bildern durch computerunterstützte visuelle Fotointerpretation erhoben. Die Satellitenbilder wurden in Österreich erstmals zwischen 1985 und 1989 jeweils im Juli und August aufgenommen, und in den Folgejahren vom Umweltbundesamt ausgewertet. Die Corine Land Cover stehen für nicht kommerzielle Zwecke kostenlos zur Verfügung. Eine Aktualisierung der CORINE Daten auf Basis von im Jahr 2000 aufgenommenen Satellitenbildern wurde mit Ende 2004 abgeschlossen. Für die Interpretation ist als kleinste Flächeneinheit 25ha und als kleinste Längeneinheit 100m festgelegt. Bei der Zuweisung der 44 Bodenbedeckungskategorien wurden verschiedene Hilfsmittel wie z.B. topographische Karten 1:50.000, Luftbilder und thematische Karten verwendet.

Folgende Landnutzungskategorien wurden in der vorliegenden Arbeit für die Erstellung der Waldmaske herangezogen:

- **Nadelwald:** Flächen mit überwiegendem Baumbewuchs, die aber auch mit Büschen und Sträuchern durchsetzt sein können; Nadelbaumarten überwiegen.
- **Laubwald:** Flächen mit überwiegendem Baumbewuchs, die aber auch mit Büschen und Sträuchern durchsetzt sein können; Laubbaumarten überwiegen.
- **Mischwald:** Flächen mit überwiegendem Baumbewuchs, die aber auch mit Büschen und Sträuchern durchsetzt sein können; weder Nadel- noch Laubbaumarten überwiegen.
- **Strauchvegetation:** Niedrige und dichte Vegetation. Überwiegend Büsche, Sträucher und Kräuter (Heidekraut, Dorngestrüpp, Besenginster, Stechginster, Goldregen usw.)

Abbildung 2-4 gibt einen Überblick über die räumliche Verteilung der oben angeführten Landnutzungskategorien.

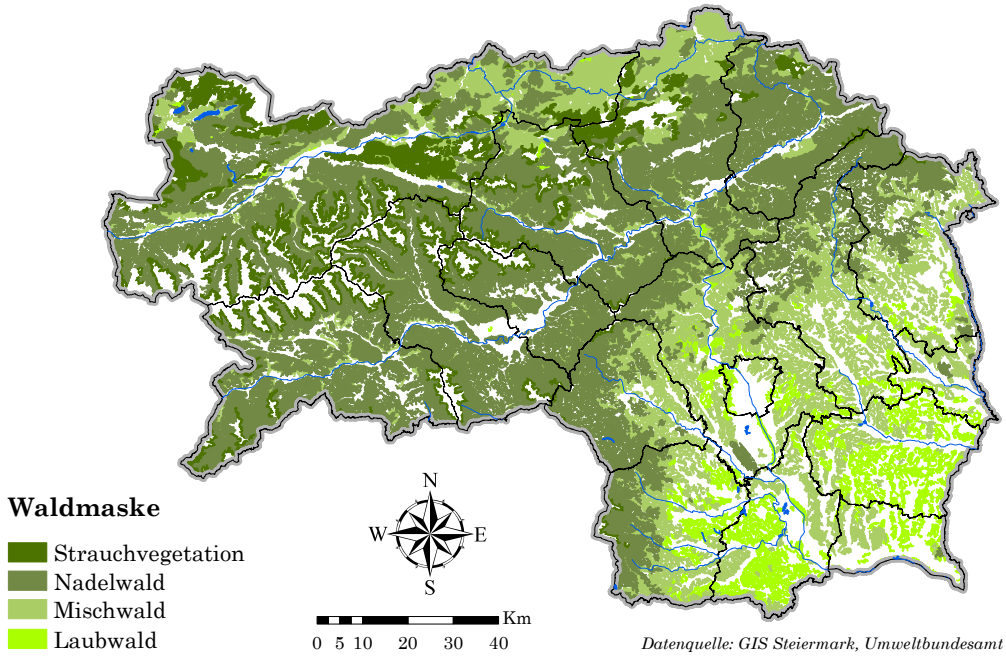


Abbildung 2-4: *Landnutzungskategorien für die Erstellung der Waldmaske*

b) SRTM3 Daten

Die SRTM-Daten sind Fernerkundungsdaten der Erdoberfläche, die bei der Shuttle Radar Topography Mission (SRTM) im Jahr 2000 aus dem Weltraum aufgezeichnet wurden. Die SRTM ist ein Gemeinschaftsprojekt der amerikanischen *National Aeronautics and Space Administration (NASA)*, der *National Imaging and Mapping Agency (NIMA)*, dem *Deutschen Zentrum für Luft- und Raumfahrt (DLR)* und der italienischen *Agenzia Spaziale Italiana (ASI)*. Ziel der Mission war es, hochauflösende, einheitlich erhobene und flächendeckende digitale Höhendaten der Erde im Bereich zwischen dem 60. Grad nördlicher und dem 56. Grad südlicher Breite zu erheben (CZEGKA et al. 2004). Die Beschränkung der Abdeckung auf diese Fläche liegt an der maximal möglichen Inklination bei Shuttle-Flügen (SCHNEIDER 2003). Die abgedeckte Fläche entspricht ca. 80% der Landmasse der Erde und ist Lebensraum für etwa 95% der Erdbevölkerung (BAMLER 1999, MARRA et al. 1998).

Das in der SRTM verwendete Verfahren zur Datenerfassung ermöglichte die Ableitung eines sehr genauen globalen Höhenmodells mit einer vertikalen Genauigkeit von 6m und einem horizontalen Pixelabstand von 30m = 1 Bogensekunde (SRTM1). Aus diesen SRTM1-Daten wurden durch eine Nachbarschaftsanalyse (Mittelwertbildung) Höhendaten mit einer Auflösung von 3

Bogensekunden (SRTM3) abgeleitet. Die Auflösung in y-Richtung entspricht demnach ca. 92,6m, in x-Richtung nimmt sie mit der geographischen Breite zu und liegt z.B. bei 47° Nord (oder 47° Süd) bei etwa 63m. Im Gegensatz zu den SRTM1-Daten sind diese SRTM3-Höhendaten frei verfügbar. Bei den aus den SRTM-Daten abgeleiteten Oberflächenmodellen handelt es sich um Landschaftsmodelle, da durch die Aufnahmemethodik jedes Landschaftselement (z.B. Vegetation, Bebauung) abgebildet wird. Die Höhengenauigkeit von mittels SRTM3-Daten erstellten Oberflächenmodellen liegt bei rund ±16m, die Genauigkeit in der horizontalen Lage bei ca. ±20m. Die SRTM3-Daten liegen als 5°x5°-Kacheln vor, die aus jeweils 6000x6000 Rasterzellen, entspricht 36 Mio. Höhenpunkten, bestehen (JARVIS et al. 2004).

Das Untersuchungsgebiet und die angrenzenden Nachbarbezirke werden von zwei Kacheln überdeckt. Um eine Erleichterung bei der Prozessierung der Daten zu gewährleisten, wurde das endgültige Oberflächenmodell aus diesem zweier-Mosaik ausgeschnitten. Insgesamt weist das in der Folge verwendete Oberflächenmodell etwa 6,1 Mio. Höhenpunkte auf, deren Werte zwischen 180m und 3600m schwanken. Die mittlere Höhe liegt im Bereich von 986,0m, die Standardabweichung bei 543,4m.

2.2.3 Schwefeldaten

In der vorliegenden Arbeit wurden keine Schwefeldaten erhoben, sondern ausschließlich bereits vorliegende Schwefeldaten ausgewertet. Diese entstammen einerseits aus Datenbeständen des **Bundesforschungs- und Ausbildungszentrums für Wald, Naturgefahren und Landschaft** sowie andererseits aus Datenbeständen der **Fachabteilung für das Forstwesen des Landes Steiermark**.

Das zur Verfügung stehende Datenmaterial beschränkt sich nicht nur auf das Untersuchungsgebiet Steiermark, sondern beinhaltet auch Schwefeldaten von benachbarten Bezirken der umliegenden Bundesländer. Durch dieses zusätzliche Datenmaterial soll auch an den Randbereichen des Untersuchungsgebiets eine stabile Interpolation gewährleistet sein. Einzige Ausnahme dabei ist der Grenzbereich zu Slowenien, wo aufgrund fehlender Schwefelwerte mit instabilen Werten der Interpolation zu rechnen ist.

Wie bereits oben erwähnt unterscheidet man bei der Probenahme im BIN zwischen dem ersten und zweiten Nadeljahrgang. Der Vergleich der unterschiedlichen Interpolationsverfahren erfolgt primär anhand der Schwefeldaten des ersten Nadeljahrgangs. Die Daten des zweiten Nadeljahrgangs fließen hingegen nur als Zusatzinformation, neben den Geländehöhendaten, in das Ordinary Cokriging ein. Anzumerken ist weiters, dass das Datenmaterial nicht ohne Vorbehandlung übernommen werden konnte. Ein Grund hierfür ist die Art der Speicherung der

Daten innerhalb der Datenbank der Fachabteilung für das Forstwesen. In der Regel werden pro BIN-Punkt zwei BIN-Bäume beprobt denen allerdings nur eine Position zugeordnet wird. Um pro BIN-Punkt nur einen Schwefelwert zu besitzen, wurden die Werte dieser beiden BIN-Bäume gemittelt.

Ein weiterer Grund ist das Vorkommen von statistischen Ausreißern welche die räumlichen Interpolationsergebnisse maßgeblich beeinflussen können. Solche Werte stellen entweder Mess- oder Übertragungsfehler oder aber kleinräumige Singularitäten dar. Bei einer Interpolation, die flächenhaft repräsentative Ergebnisse bringen soll, sind solche Daten nicht unbedingt erwünscht. Um ein valides Interpolationsergebnis zu erhalten, wurde deshalb das Datenmaterial auf solche Ausreißer überprüft. Dazu wurden die Punktdaten lagetreu visualisiert. Werte, die signifikant vom jeweiligen Mittelwert der Datengesamtheit abweichen, wurden kenntlich gemacht und hinsichtlich des Wertebereiches der nächsten Nachbarn beurteilt. Als Ausreißer erkannte Punkte wurden daraufhin aus dem Datenmaterial entfernt.

Als Ausreißertest wurde der nichtparametrische Ausreißertest **Median-3-Interquartil-Test** herangezogen, welcher keine bestimmte Verteilung der Daten voraussetzt. Die Teststatistik lautet (HINTERDING et al 2003):

- x^* ist ein Ausreißer, wenn x^* > Median + 3*(75.Perzentil -25.Perzentil) oder
- x^* < Median - 3*(75.Perzentil -25.Perzentil)

Die Verwendung derartiger Ausreißertests ist allerdings nicht unproblematisch, da die Beobachtung, die aus der Reihe fällt, keineswegs eine Fehlmessung sein muss. Als warnendes Beispiel hierfür kann die Entdeckung des „Ozon-Lochs“ dienen: Über einen Zeitraum von 10 Jahren wurden von der NASA zu bestimmten Zeiten niedrige Ozon-Gehalte in der Atmosphäre gemessen, die aber auf Grund eines automatischen Ausreißertests aus dem Datensatz eliminiert wurden (MERKEL & PLANER-FRIEDRICH 2003). Ausreißertests sind deshalb nur zu verwenden, um Hinweise auf mögliche Daten- und Erhebungsfehler zu geben. Erst eine individuelle, inhaltliche Plausibilitätsprüfung darf dazu führen, dass der betreffende Wert aus der weiteren statistischen Auswertung ausgeschlossen wird.

Die Schwefeldaten liegen für einen Zeitraum von 1983 bis 2003 vor, wobei im Mittel pro Jahr etwa 1900 BIN-Bäume beprobt wurden. Ein relativ hoher Anteil dieser BIN-Bäume wird im Rahmen der Lokalnetze betrieben, die vor allem in der Umgebung von potentiellen Emittenten eingerichtet wurden. Abbildung 2-5 zeigt am Beispiel des Erhebungsjahres 2003 die Verteilung der BIN-Punkte innerhalb des Untersuchungsgebietes und der Nachbarbezirke. Bei den dargestellten BIN-Punkten handelt es sicher um Bundesnetzpunkte, Lokalnetze sowie Punkten des Waldschadensbeobachtungsnetzes. Die räumliche Verteilung der BIN-Punkte weißt bedingt durch die Lokalnetze vor allem im Bereich der Industriezentren clusterförmige Strukturen auf.

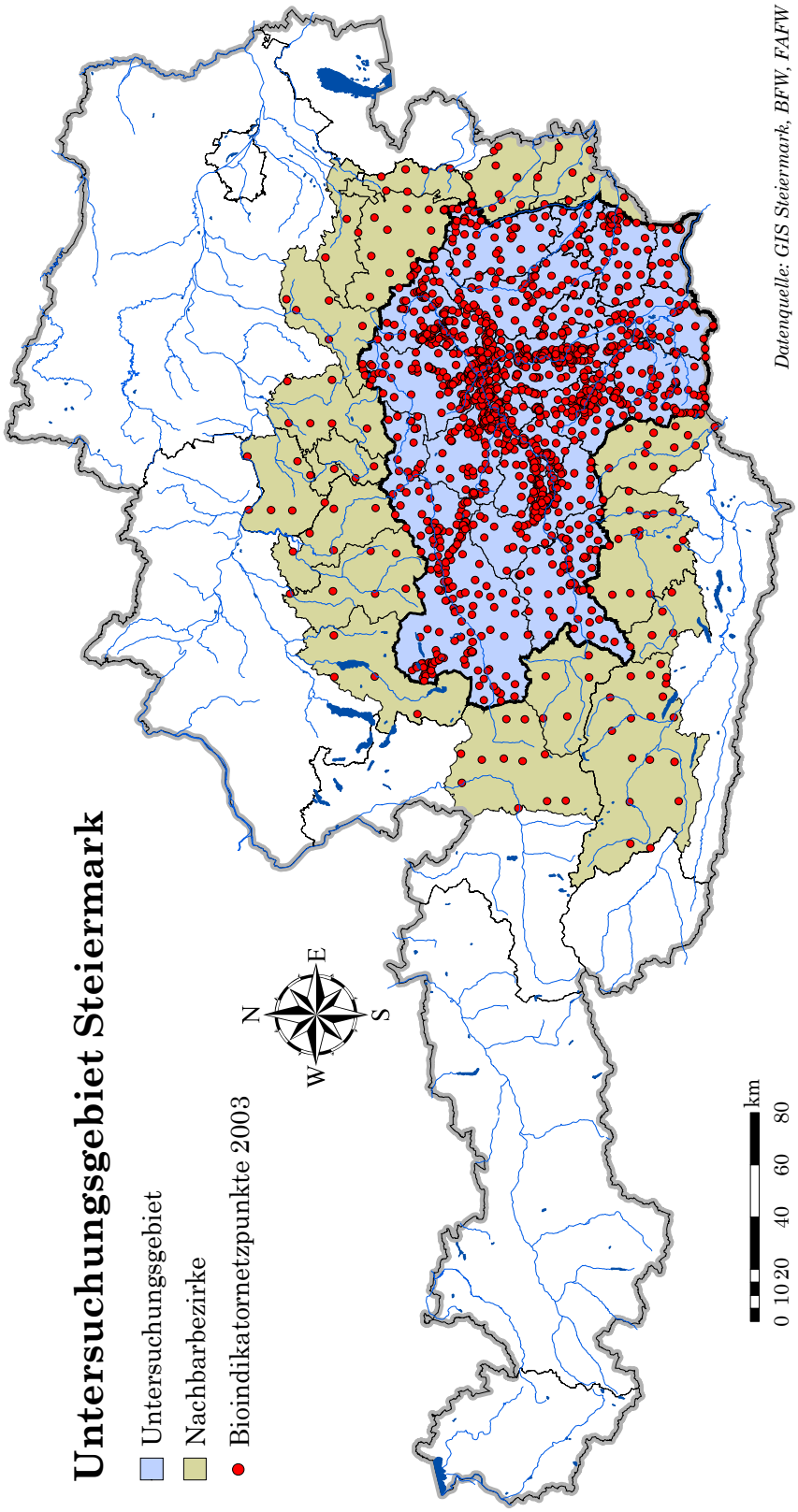


Abbildung 2-5: *Untersuchungsgebiet Steiermark*

Abbildung 2-6 gibt einen Überblick über die Anzahl der BIN-Punkte entsprechend des jeweiligen Erhebungsjahres, sowie deren Aufteilung auf die unterschiedlichen Netze.

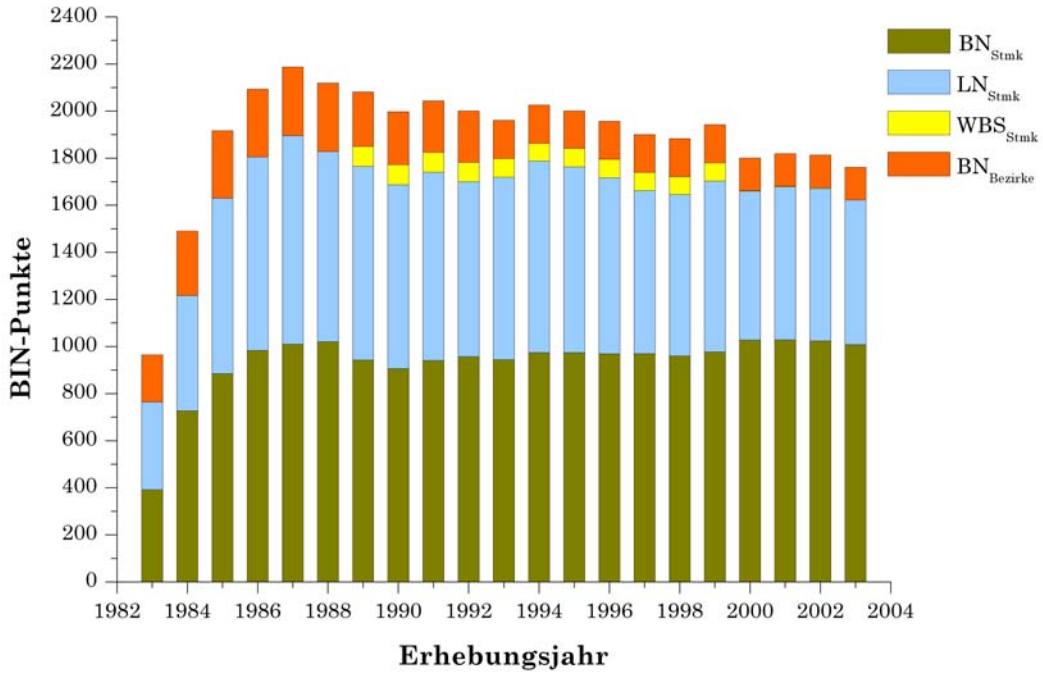


Abbildung 2-6: *Anzahl der BIN-Punkte je Erhebungsjahr bzw. Netz*

Die Anzahl der BIN-Punkte in Abbildung 2-6 setzt sich aus Bundesnetzpunkten der Steiermark (BN_{Stmk}), Bundesnetzpunkten der Nachbarbezirke ($BN_{Bezirke}$), Lokalnetzpunkte der Steiermark (LN_{Stmk}) sowie aus Punkten des Waldschadensbeobachtungssystemnetzes (WBS_{Stmk}) zusammen. An dieser Stelle sei darauf hingewiesen, dass die Fachabteilung für Forstwesen des Landes Steiermark zwar eine eigene Klasse „Verdichtungsnetzpunkte" in der Datenbank ausweist, diese aber zur Vereinfachung mit der Klasse „Bundesnetzpunkte" zusammengefasst wurde.

Die erhobenen BIN-Punkte verteilen sich nicht nur unterschiedlich im Raum sondern auch hinsichtlich der Seehöhe. Abbildung 2-7 zeigt die Höhenverteilung der BIN-Punkte am Beispiel des Erhebungsjahres 2003. Die größte Anzahl an BIN-Punkten findet man in einer Seehöhe zwischen 600m und 1000m. Mit zunehmender Seehöhe nimmt die Anzahl der Punkte allerdings immer stärker ab. Oberhalb von 2000m Seehöhe sind bedingt durch die natürliche Waldgrenze überhaupt keine BIN-Punkte mehr vorhanden. Die relativ geringe Anzahl an BIN-Punkten zwischen 200m und 600m lässt sich zum Großteil mit der intensiven landwirtschaftlichen Nutzung und den urbanen Gebieten in dieser Seehöhe begründen, womit ein geringer Waldanteil verbunden ist.

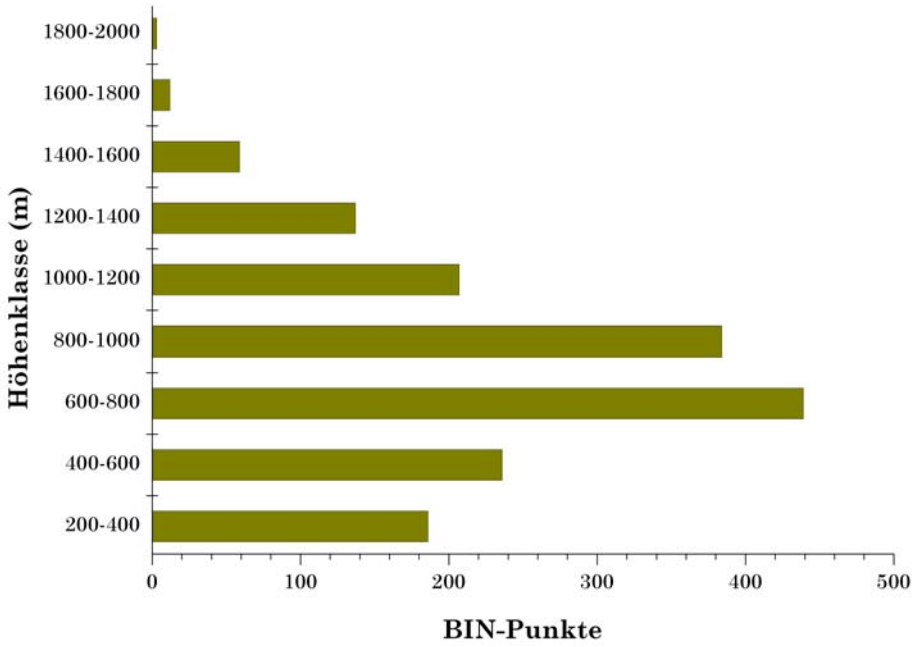


Abbildung 2-7: *Höhenverteilung der BIN-Punkte für das Jahr 2003*

Die Häufigkeitsverteilung der BIN-Punkte entsprechend der Exposition ist in Abbildung 2-8 dargestellt. Als Datenbasis diente wiederum das Erhebungsjahr 2003. Die größte Häufigkeit an BIN-Punkten tritt bei der Nord-Exposition gefolgt von der Süd-Exposition auf. Ein möglicher Grund hierfür ist die große Anzahl von BIN-Punkten im Bereich der Mur-Mürz-Furche, welche eine West-Ost-Erstreckung aufweist.

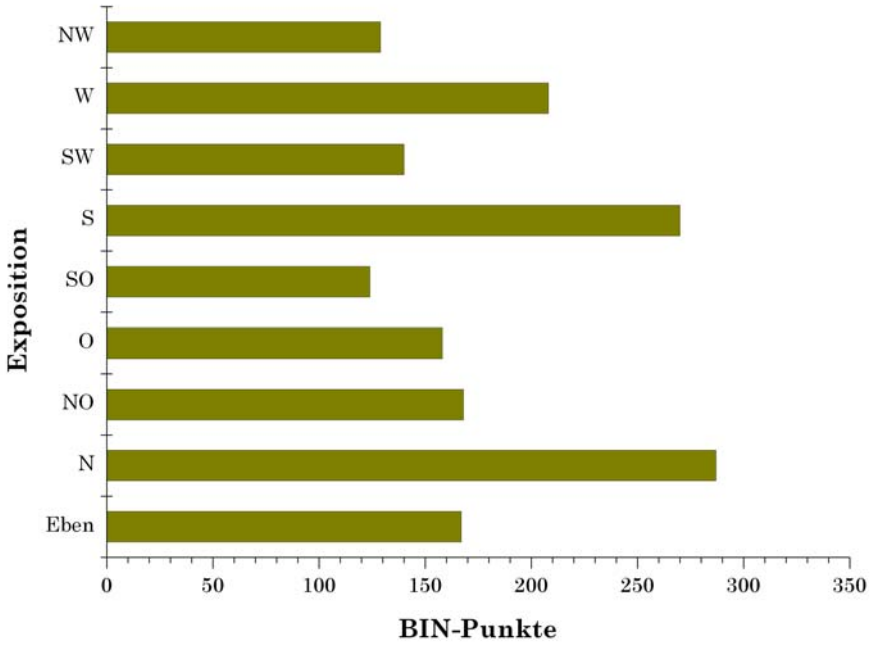


Abbildung 2-8*: Expositionsverteilung der BIN-Punkte für das Jahr 2003*

Um einen Überblick über die Struktur der Schwefeldaten für den ersten und zweiten Nadeljahrgang der einzelnen Erhebungsjahre zu bekommen, wurden im Rahmen einer explorativen Datenanalyse die wichtigsten statistischen Kennzahlen ermittelt. Die Grundlagen dieser explorativen Datenanalyse basieren auf den Arbeiten des amerikanischen Statistikers TUKEY (1977).

Die Ergebnisse der Analysen sind in den Tabelle 2-1 und 2-2 zusammengefasst. Bei den statistischen Kennzahlen handelt es sich um die Lagemaße *Mittelwert*, *Median*, *25.Perzentile*, *75.Perzentile*, dem Streuungsmaß *Standardabweichung,* sowie der *Schiefe* und der *Kurtosis*. Zusätzlich wurden die minimalen und maximalen Schwefelwerte je Erhebungsjahr angeführt.

Aus den statistischen Kennzahlen der Tabelle 2-1 und 2-2 lassen sich sowohl für die Daten des ersten, als auch zweiten Nadeljahrgang folgende Aussagen ableiten:

- Im Zeitraum 1983 bis 2003 nehmen die maximalen Schwefelwerte ab, während die minimalen Schwefelwerte sowie der Mittelwert und der Median annähernd konstant bleiben (vgl. Abbildung 2-9).
- Der Mittelwert der Schwefelwerte ist in allen Jahren größer als der Median, d.h. die Verteilung der Schwefelwerte ist rechtsschief.
- Die Kurtosis ist in jedem Erhebungsjahr größer als 3, d.h. die Verteilung der Schwefelwerte ist hochgipfelig.

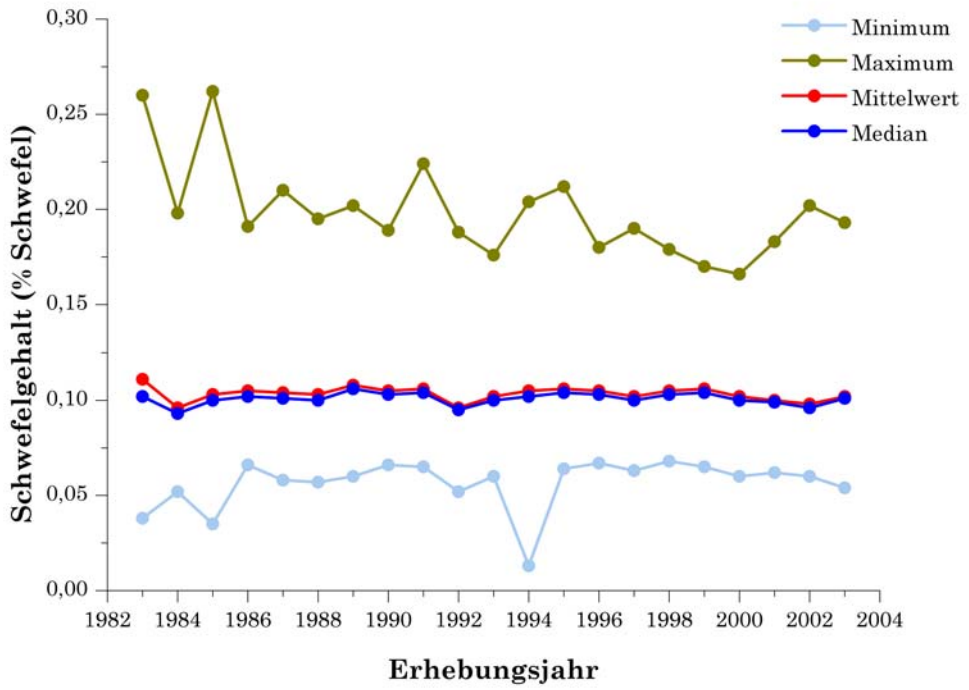


Abbildung 2-9: *Vergleich der Lagemaße für den Zeitraum von 1983 bis 2003*

Tabelle 2-1: *Kennzahlen der Schwefelwerte für den ersten Nadeljahrgang*

		Schwefelwerte (% Schwefel)								
Jahr	**BIN-Bäume**	**Minimum**	**Maximum**	**Mittelwert**	**Std.abw.**	**Schiefe**	**Kurtosis**	**25.Quartil**	**Median**	**75.Quartil**
1983	965 (765*)	0,038	0,260	0,111	0,0323	0,8671	3,5240	0,086	0,102	0,130
1984	1490 (1216)	0,052	0,198	0,096	0,0210	0,8032	4,0892	0,080	0,093	0,109
1985	1917 (1630)	0,035	0,262	0,103	0,0197	1,0283	6,2117	0,090	0,100	0,113
1986	2093 (1805)	0,066	0,191	0,105	0,0189	0,8211	4,2017	0,091	0,102	0,116
1987	2187 (1895)	0,058	0,210	0,104	0,0179	0,7635	4,4602	0,091	0,101	0,114
1988	2119 (1828)	0,057	0,195	0,103	0,0172	0,7607	4,2929	0,090	0,100	0,112
1989	2081 (1850)	0,060	0,202	0,108	0,0180	0,7307	4,2868	0,095	0,106	0,118
1990	1997 (1772)	0,066	0,189	0,105	0,0176	0,8312	4,3476	0,092	0,103	0,115
1991	2043 (1824)	0,065	0,224	0,106	0,0177	1,0116	5,6328	0,094	0,104	0,116
1992	2001 (1782)	0,052	0,188	0,096	0,0169	0,9002	4,7026	0,085	0,095	0,105
1993	1961 (1798)	0,060	0,176	0,102	0,0163	0,8473	4,3242	0,091	0,100	0,111
1994	2029 (1866)	0,013	0,204	0,105	0,0177	0,9067	5,5545	0,092	0,102	0,114
1995	2001 (1841)	0,064	0,212	0,106	0,0167	1,0172	5,1829	0,094	0,104	0,115
1996	1957 (1795)	0,067	0,180	0,105	0,0165	0,7855	3,7790	0,093	0,103	0,115
1997	1901 (1739)	0,063	0,190	0,102	0,0169	1,0677	5,0067	0,090	0,100	0,110
1998	1883 (1721)	0,068	0,179	0,105	0,0158	0,7282	4,0061	0,094	0,103	0,115
1999	1942 (1780)	0,065	0,170	0,106	0,0156	0,5757	3,4783	0,095	0,104	0,115
2000	1801 (1662)	0,060	0,166	0,102	0,0169	0,5441	3,2043	0,090	0,100	0,113
2001	1819 (1681)	0,062	0,183	0,100	0,0148	0,6328	3,9536	0,090	0,099	0,108
2002	1814 (1675)	0,060	0,202	0,098	0,0169	0,7773	4,6209	0,086	0,096	0,108
2003	1762 (1624)	0,054	0,193	0,102	0,0175	0,7164	4,1347	0,090	0,101	0,113

** Anzahl der BIN-Bäume in der Steiermark Std.abw. = Standardabweichung*

Tabelle 2-2: *Kennzahlen der Schwefelwerte für den zweiten Nadeljahrgang*

		Schwefelwerte (% Schwefel)								
Jahr	**BIN-Bäume**	**Minimum**	**Maximum**	**Mittelwert**	**Std.abw.**	**Schiefe**	**Kurtosis**	**25.Quartil**	**Median**	**75.Quartil**
1983	965 (765*)	0,054	0,390	0,129	0,049	1,272	5,079	0,091	0,117	0,156
1984	1490 (1216)	0,020	0,294	0,109	0,031	1,076	4,944	0,085	0,104	0,126
1985	1917 (1630)	0,034	0,353	0,114	0,029	1,299	7,212	0,093	0,110	0,129
1986	2093 (1805)	0,060	0,332	0,108	0,026	1,338	7,331	0,090	0,105	0,122
1987	2187 (1895)	0,010	0,277	0,113	0,025	1,149	6,133	0,096	0,110	0,126
1988	2119 (1828)	0,060	0,278	0,111	0,024	1,079	5,583	0,093	0,108	0,123
1989	2081 (1850)	0,065	0,250	0,112	0,022	1,077	5,526	0,097	0,110	0,125
1990	1997 (1772)	0,069	0,244	0,110	0,022	1,109	5,153	0,094	0,108	0,122
1991	2043 (1824)	0,068	0,314	0,109	0,022	1,751	9,516	0,093	0,105	0,120
1992	2001 (1782)	0,058	0,226	0,100	0,021	1,186	5,442	0,086	0,097	0,110
1993	1961 (1798)	0,062	0,225	0,105	0,023	1,335	5,851	0,090	0,100	0,116
1994	2029 (1866)	0,063	0,252	0,108	0,029	1,554	7,301	0,092	0,103	0,118
1995	2001 (1841)	0,010	0,236	0,108	0,022	1,373	6,531	0,093	0,103	0,118
1996	1957 (1795)	0,067	0,180	0,105	0,017	0,785	3,779	0,093	0,103	0,115
1997	1901 (1739)	0,063	0,235	0,107	0,022	1,357	6,327	0,092	0,104	0,118
1998	1883 (1721)	0,020	0,282	0,101	0,020	1,475	6,721	0,090	0,099	0,110
1999	1942 (1780)	0,069	0,218	0,107	0,017	1,075	6,275	0,095	0,105	0,117
2000	1801 (1662)	0,062	0,215	0,103	0,019	1,015	4,792	0,090	0,100	0,113
2001	1819 (1681)	0,058	0,220	0,100	0,018	1,091	5,794	0,088	0,098	0,110
2002	1814 (1675)	0,055	0,251	0,096	0,018	1,312	8,170	0,084	0,094	0,106
2003	1762 (1624)	0,058	0,200	0,097	0,017	1,093	6,228	0,085	0,095	0,106

** Anzahl der BIN-Bäume in der Steiermark Std.abw. = Standardabweichung*

Um eine Vorstellung über die Häufigkeitsverteilung der Schwefelwerte zu bekommen, wurden für die einzelnen Erhebungsjahre Histogramme erstellt. Hierbei werden die absoluten bzw. relativen Häufigkeiten, mit denen eine Merkmalsausprägung beobachtet wurde, als Rechtecke abgebildet und in Klassen zusammengefasst. Das Histogramm ist somit eine einfache Möglichkeit einen Datensatz graphisch darzustellen um aus diesen Wertebereich, die Form der Verteilung (symmetrisch, linksschief, rechtsschief, ein- oder mehrgipfelig) sowie die Bereiche größerer oder kleinerer Datenhäufigkeit abzulesen.

Abbildung 2-10 zeigt am Beispiel des Erhebungsjahres 2003 die Verteilung der Schwefeldaten für den ersten Nadeljahrgang. Dabei handelt es sich um eine eingipfelige, symmetrische und leicht rechtsschiefe Verteilung der Schwefeldaten.

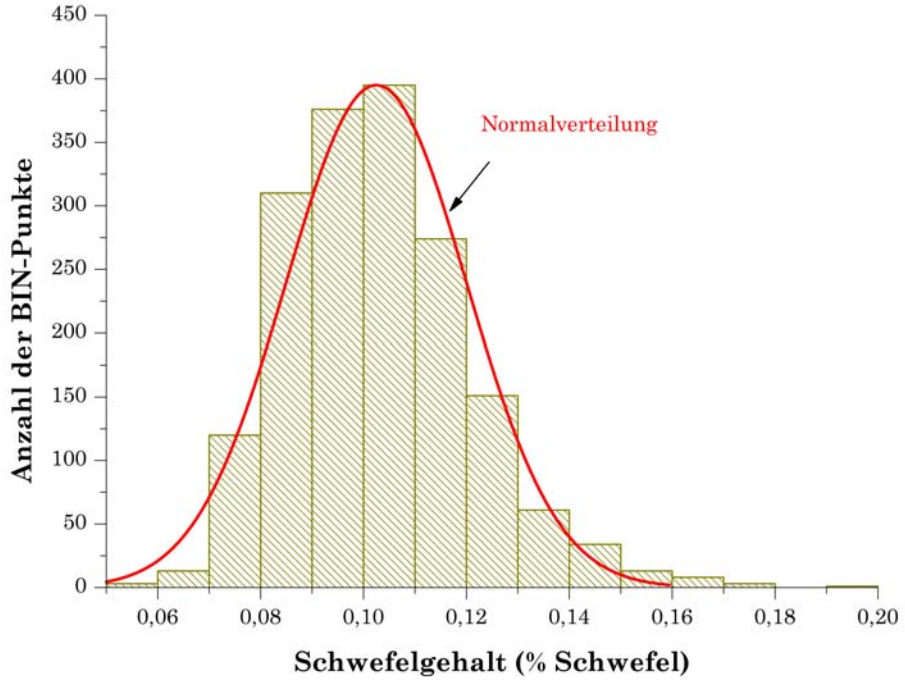


Abbildung 2-10: *Histogramm der Schwefelwerte für das Erhebungsjahr 2003*

Ein zu interpolierender Prozess sollte in der Regel normalverteilt sein. Eine Möglichkeit um einen Datensatz auf Normalverteilung zu überprüfen ist ein *Normal QQ-Plot*. Ein QQ-Plot (Quantil-Quantil-Plot) erleichtert den Vergleich einer empirischen Verteilung mit einer theoretischen Verteilung. Es vergleicht geordnete Werte einer Merkmalsausprägung mit Quantilen einer bestimmten theoretischen Verteilung (z.B. Normal, Lognormal, Exponentiell, Weibull), d.h. bei einem Normal QQ-Plot wird die theoretische Quantile der Normalverteilung mit den empirischen Quantilen der Daten verglichen. Liegt eine Normalverteilung der Daten vor, so liegen die Punkte des QQ-Plots näherungsweise auf einer Geraden.

In Abbildung 2-11 sind die beobachteten Schwefelwerte vom Erhebungsjahr 2003 (Nadeljahrgang 1) gegen die theoretischen Werte einer Normalverteilung aufgetragen. Die theoretische Normalverteilung wird durch die Gerade dargestellt. Sind die empirischen Werte normalverteilt, so müssen die einzelnen Punkte annähernd dem Verlauf der Geraden folgen. Dies ist bei den Schwefelwerten für das Jahr 2003 (Nadeljahrgang 1) offenkundig aber nicht der Fall. Im mittleren Bereich liegen die Werte zwar in der Nähe der Geraden, am oberen und unteren Ende weichen sie jedoch erheblich davon ab. Noch gravierender als die Stärke der Abweichung ist deren Form. Die einzelnen Punkte in der Grafik streuen nicht zufällig um die Gerade, sondern weisen ein klares Muster auf. Während die Werte im mittleren Bereich unterhalb der Geraden liegen, befinden sie sich im Bereich der hohen sowie der niedrigen Werte deutlich oberhalb der Geraden. Ein solches Muster lässt den Schluss zu, dass die beobachteten Werte systematisch von der Normalverteilung abweichen.

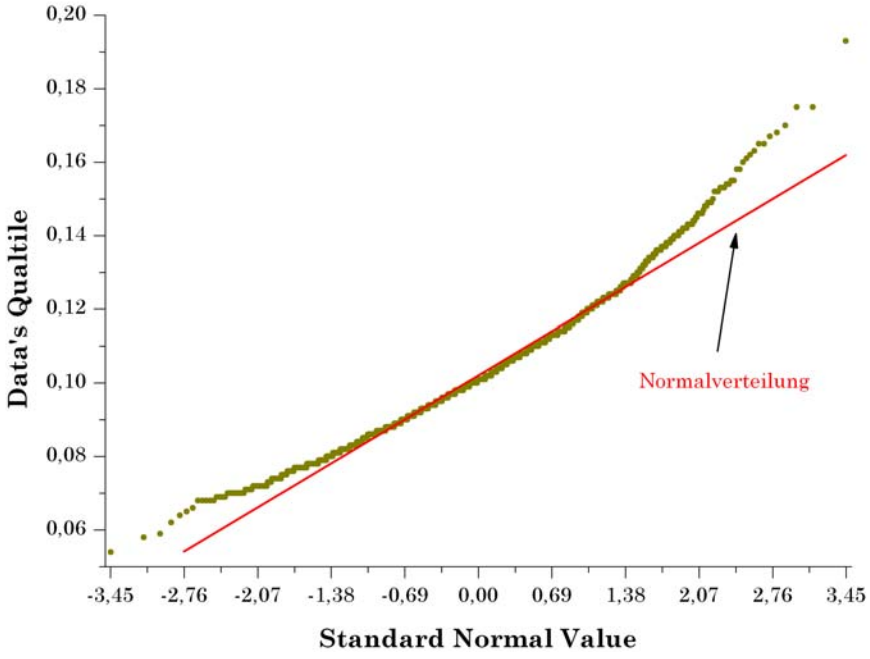


Abbildung 2-11: *Normal QQ-Plot für das Erhebungsjahr 2003*

Auf die Darstellung von weiteren Häufigkeitsverteilungen bzw. QQ-Plots wird an dieser Stelle verzichtet, da die weiteren Erhebungsjahre bzw. Nadeljahrgänge eine ähnliche Charakteristik aufweisen.

Zusammenfassend sei nochmals festgehalten, dass für die Schätzung anhand der drei Interpolationsverfahren die Schwefeldaten des ersten Nadeljahrgangs verwendet werden. Eine Ausnahme bildet das Ordinary Cokriging Verfahren, wo neben den Schwefeldaten des ersten Nadeljahrgangs, auch Geländehöhendaten sowie Schwefeldaten des zweiten Nadeljahrgangs herangezogen werden.

3 Methodik

In diesem Kapitel werden die drei in dieser Arbeit zur Anwendung kommenden räumlichen Interpolationsverfahren vorgestellt. Das Ziel der räumlichen Interpolation ist die Schätzung der kontinuierlichen räumlichen Verteilung einer Beobachtungsvariablen aus an verschiedenen Messpunkten in einem Raum gemessenen Werten. Im Anschluss an die Einführung in die methodischen Grundlagen der Interpolationsverfahren werden jene statistischen Kennzahlen vorgestellt anhand derer die Qualität dieser Verfahren beurteilt werden soll.

3.1 Inverse Distanzgewichtung

Die Inverse Distanzgewichtung (IDW, engl.: *Inverse Distance Weighting*) ist eine nichtstatistische, exakte und lokale Interpolationsmethode (JOHNSTON et. al 2001, BULLINGER 2000). Den nichtstatistischen Verfahren ist gemeinsam, dass die unbekannten Werte nur mit Hilfe der geometrischen Distanzen geschätzt werden. Lokal bedeutet, dass in die Interpolation nur Stützpunkte der Umgebung („Nachbarschaft“) einfließen, während Stützpunkte außerhalb dieser Nachbarschaft nicht berücksichtigt werden. Beim IDW-Verfahren erfolgt die Annahme, dass die punktuell gemessenen raumbezogenen Daten in Abhängigkeit von der Distanz im Raum gewisse Ähnlichkeiten in ihren Werten aufweisen (GERLACH 2001). Das bedeutet, dass Messpunkte die dem unbekannten Schätzwert näher liegen ähnlicher sind als weiter entfernte. Diese Methode basiert somit auf der auch sonst in der geographischen Analytik häufig getroffenen Grundannahme nach TOBLER (1970): „*Everything is related to everything, but near things are more related than distance things*“.

Beim IDW-Verfahren wird ein Wert der Beobachtungsvariablen an einem unbeprobten Ort durch ein gewichtetes Mittel der benachbarten gemessenen Werte geschätzt. Die Gewichte des dabei verwendeten linearen Schätzers sind proportional zu den Inversen des Abstands zwischen dem unbeprobten Ort u_0 und der Stichprobe (GERLACH 2001). An dem unbeprobten Ort wird der Wert der Beobachtungsvariablen beispielsweise nach der einfachen invers distanzgewichteten Interpolation wie folgt geschätzt.

$$z(u_0) = \frac{\sum_{i=1}^{n} \frac{1}{d_i} z(u_i)}{\sum_{i=1}^{n} \frac{1}{d_i}} \qquad \textbf{(3-1)}$$

mit:

$z(u_0)$	*Wert der Beobachtungsvariable am Ort u_0*
$z(u_i)$	*Wert der Beobachtungsvariable am Ort u_i*
d_i	*Abstand zwischen den Orten u_0 und u_i*
n	*Anzahl der Datenpunkte (Stützpunkte)*

Vereinfachend kann gesagt werden, dass nach der Auswahl der „Stützpunkte“ in der Umgebung eines gesuchten Punktes zuerst die Distanzen zu diesen berechnet werden. Anschließend wird der Wert des gesuchten Punktes als Mittelwert der Umgebungspunkte berechnet, und zwar gewichtet mit dem Kehrwert der Distanzen. Dabei wird über alle beprobten Orte summiert. Die Annahme über die Art des Zusammenhangs zwischen den Beobachtungsvariablen ist dabei intuitiv und unabhängig von der Beobachtungsvariablen (HINTERDING 1998). Die Gleichung 3-1 wird eher selten verwendet, wesentlich häufiger findet man folgende Variante:

$$z_j = \frac{\sum_{i=1}^{n} \left(\frac{1}{d_{ij}^{\ r}} z_i \right)}{\sum_{i=1}^{n} \frac{1}{d_{ij}^{\ r}}} \qquad \textbf{(3-2)}$$

mit:

z_j	*unbekannter zu ermittelnder Wert z an der Position j*
z_i	*bekannter Wert z an der Position i*
d_{ij}	*Distanz zwischen Datenpunkt an Position i und Position j*
r	*Distanzexponent*
n	*Anzahl der Datenpunkte (Stützpunkte)*

Abbildung 3-1 zeigt vereinfachend die in die obige Formel eingehenden Eingangsgrößen. Über den Distanzexponenten *r* wird das Gewicht der Stützpunkte gesteuert. Je höher der Exponent angegeben wird, umso „unruhiger“ wird die Oberfläche, da nur das Gewicht der unmittelbarsten Nachbarn in die Kalkulation eingeht (BLASCKE & STROBL 2003). Umso niedriger der Exponent gewählt wird, umso gleichförmiger gehen alle Nachbarn ungeachtet ihrer Distanz in die Berechnung ein, und desto „glatter“ wird die Schätzoberfläche. Üblicherweise arbeitet man mit einem Exponenten von 2. Dieser Wert hat allerdings den Nachteil, dass der berechnete Wert in unmittelbarer Nachbarschaft der

Stützpunkte schnell auf ein niedriges Niveau absinkt und um die Stützpunkte konzentrische Kreise, sogenannte „Bull Eyes“ entstehen (KIENZLE 2002). Ebenso werden langgestreckte Rücken und Verbindungen zwischen Stützpunkten gleicher Ausprägung verhindert (RASE 1996).

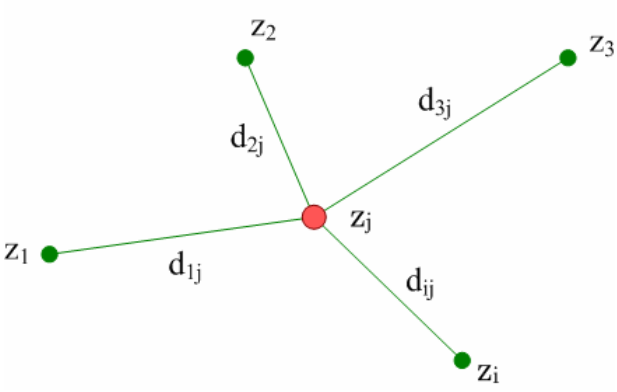


Abbildung 3-1: *Schätzung mittels IDW-Verfahren*

Das Ergebnis der IDW-Interpolation hängt, wie bereits erwähnt, nicht nur von der Größe des verwendeten Distanzexponenten ab, sondern auch von anderen Faktoren, welche die „Umgebung“ der verwendeten Stützpunkte festlegen. Dazu zählt z.B. die Auswahl einer geeigneten Anzahl an Stützpunkten die mittels unterschiedlicher Kriterien (Thiessen-Nachbarschaft, Suchradius, minimale Punkteanzahl) erfolgen kann. Häufig ergibt sich bei der lokalen Interpolation die Situation, dass bei alleiniger Berücksichtigung von Bedingungen hinsichtlich minimaler Stützpunkteanzahl oder eines Suchradius keine zufrieden stellende Konfiguration der Stützpunkte zustande kommt. Dieser Fall tritt z.B. ein, wenn die Mehrzahl oder gar alle Stützpunkte auf einer digitalisierten Isoplethe liegen. Abhilfe bietet in derartigen Fällen die Einführung einer zusätzlichen Richtungsbedingung bei der Auswahl der Interpolationsstützpunkte.

Von SHEPARD (1968) stammt eine modifiziere Form der IDW-Interpolation, bei welcher der „Bull Eyes“ Effekt weniger stark hervortritt. Bei dieser Form wird vor der eigentlichen Interpolation um jeden zu schätzenden Punkt eine lokale Trendoberfläche entsprechend der angegebenen Nachbarschaftsparameter eingepasst. Für die Gewichtsberechnung werden die Werte dieser lokalen Trendoberfläche verwendet. Einer der Nachteile der modifizierten IDW-Interpolation ist die Neigung des Algorithmus, eine Oberfläche zu erzeugen, die in der unmittelbaren Nachbarschaft der Stützpunkte flach und parallel zur Bezugsebene verläuft (RASE 1996). Diese Plateaus sind oft annähernd kreisförmig. Der Algorithmus verhindert auch die Ausbildung lang gestreckter Rücken. Diese Eigenschaften, so urteilen GORDON & WIXOM (1978), machen den Algorithmus ungeeignet als allgemein verwendbares Interpolationsverfahren.

3.2 Kriging

Kriging hat seinen Ursprung im Bergbau Südafrikas wo der Bergbauingenieur KRIGE (1951) in den frühen 50er Jahren des 20.Jahrhunderts zur Berechnung der Vorräte der Goldlagerstätte Witwatersrand den Probenwerten Einflusszonen zuordnete (GRAMS 2000). Die Ansätze von Krige wurden Ende der 50er Jahre des 20.Jahrhunderts vom Franzosen Matheron aufgegriffen und zu der Theorie der ortsabhängigen (regionalisierten) Variablen weiterentwickelt (MATHERON 1963, 1970). Zur gleichen Zeit entwickelte GANDIN (1963) unabhängig von Matheron in der Sowjetunion dasselbe Verfahren um es auf den Bereich der Meteorologie anzuwenden. In der Meteorologie heißt dieses Verfahren „optimale Interpolation" und gelegentlich findet man in der Literatur auch die Bezeichnung von der „Maximum-Entropie-Methode" (STOYAN et al. 1997)

Kriging ist ein Oberbegriff für eine Reihe von geostatistischen Schätzverfahren wie z.B. *ordinary kriging, universal kriging, disjunctive kriging, Bayesian kriging, indicator kriging, probability kriging* (GERLACH 2001). Gemeinsam ist diesen Schätzverfahren, dass es Regressionstechniken sind, die eine Minimierung der Schätzvarianz zum Ziel haben (HEINRICH 1992).

Im Gegensatz zu den meisten anderen Interpolationsverfahren (wie z.B. Inverse Distanzgewichtung) ist Kriging ein interaktiver Arbeitsprozess bei dem zuerst die der Verteilung zugrunde liegenden räumlichen Prozesse untersucht werden, ehe mittels einer als geeignet erkannten Schätzmethode die Interpolation einer Oberfläche durchgeführt wird (BLASCHKE & STROBL 2003). Kriging zählt zu den exakten Interpolationsmethoden, d.h. die Schätzung an den Messpunkten entspricht genau dem gemessenen Wert an dieser Stelle, was z.B. für Analysen mit Trendflächen oder Moving Averages nicht gilt (WALDOW 1998).

Im Kriging wird angenommen, dass der zu schätzende Prozess aus einer deterministischen und einer stochastischen Komponente zusammengesetzt ist (Heinrich 1992). Durch die unterschiedliche Definition der deterministischen Komponente werden verschiedene Kriging-Varianten festgelegt. Die stochastische Komponente wird immer als intrinsisch stationär oder stationär zweiter Ordnung angenommen (HINTERDING et al. 2003). Der Prozess, welcher der Verteilung von Werten im Raum zugrunde liegt, ist in der Regel nicht oder nur ungenügend bekannt. Deshalb wird beim Kriging-Verfahren ein geostatistisches Modell angenommen, mit dem nicht bekannte Werte aus bekannten geschätzt werden können. Mit Hilfe dieses Modells wird der räumliche Zusammenhang der Daten, welcher auch als Autokorrelation bezeichnet wird, ermittelt. Die Gewichte werden im Modell so optimiert, dass der Schätzer im Mittel den wahren Wert schätzt und keinen systematischen Fehler macht (ARMSTRONG 1998).

Da die Methode Kriging sehr komplex ist, wird in weiterer Folge die Theorie der in der vorliegenden Arbeit verwendeten Schätzverfahren nur soweit vorgestellt, wie es zum Verständnis der Vorgehensweise notwendig erscheint. Zur Herleitung der einzelnen Gleichungssysteme wird deshalb auf die Arbeiten von WACKERNAGEL (2000), ARMSTRONG (1998), GOOVAERTS (1997), ISAAKS & SHRIVASTAVA (1989), AKIN and SIEMES (1988), DUTTER (1985) sowie JOURNEL & HUIJBREGTS (1978) verwiesen.

3.2.1 Semivariogramm

Als Maß für den Zusammenhang von räumlich verteilten Werten dient das Semivariogramm. Es ist durch die folgende Gleichung definiert (ARMSTRONG 1998):

$$\gamma(h) = \frac{1}{2} Var\left[Z(u+h) - Z(u)\right] \qquad \textbf{(3-3)}$$

mit:

γ(h)	*Semivarianz*
Z(u)	*Regionalisierte Variable Z(u) am Messort u*
Z(u+h)	*Regionalisierte Variable Z(u+h) am Messort u+h*
Var	*Varianz*

Die Semivarianz *γ(h)* gibt dabei die mittlere Streuung der Differenzen der Werte von Messpunkten an, die durch den Abstand *h* getrennt sind. Es wird also ermittelt, um welchen Wert sich eine Variable in einem bestimmten Abstand verändert. Bei einer hohen mittleren Streuung ist der räumliche Zusammenhang im Abstand *h* gering. Die Modellierung der räumlichen Abhängigkeit erfolgt in zwei Schritten. Zuerst muss das empirische (bzw. experimentelle) Semivariogramm aus der vorliegenden Datengesamtheit berechnet werden, anschließend wird an dieses eine theoretische Semivariogrammfunktion angepasst (HAUSSMANN 2000).

Zur Berechnung des **empirischen Semivariogramms** werden zunächst die Abstände zwischen allen gemessenen Werten sowie die Wertedifferenz der jeweiligen Punktepaare bestimmt. Da in der Praxis oft nur wenige Abstandsvektoren vorkommen, welche dieselbe Länge haben, nach einer Faustregel aber mindestens 30 bis 50 Tupel benötigt werden, um die mittlere Varianz für einen bestimmten Abstand verlässlich abzuschätzen werden Abstandsklassen (h_k) gebildet (LORUP 2004, KRYWKOW 2000, JOURNEL & HUIJBREGTS 1978). Diese werden auch als „**lags**“ bezeichnet. Je größer die Abstandsklassen gewählt werden, umso mehr werden die Variogrammwerte geglättet. In der Praxis wird die maximale Abstandklasse nicht die größte Distanz

zwischen zwei Punkten im Untersuchungsgebiet übersteigen, da ansonsten keine Punktepaare mehr gefunden werden können. Anzumerken ist weiters, dass im Regefall kein systematisches, sondern ein mehr oder weniger unregelmäßiges Punktraster vorliegt. Bei Verwendung einer konstanten Abstandklasse würde man deshalb nur selten Punktepaare finden, die exakt diese Distanz zueinander aufweisen. Um dies zu vermeiden wird eine so genannte „**lag-Toleranz**" eingeführt, als deren Größe man üblicherweise in der Geostatistik die Hälfte der Abstandsklasse von lag 1 wählt. Die mittlere Semivarianz $\gamma(h)$ innerhalb eines jeden lags wird über die Gleichung des empirischen Semivariogramms bestimmt (GOOVAERTS 1997, ISAAKS & SRIVASTAVA 1989):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i,j=1}^{N(h)} (v_i - v_j)^2 \qquad \textbf{(3-4)}$$

mit:

N(h)	*Anzahl der Wertepaare an der Distanz h*
h	*Lag-Distanz*
v_i, v_j	*Werte der Punktepaare (head- und tail-Werte)*

Diese Werte werden in das Semivariogramm eingetragen. Dort sind an der Abszisse die mittleren Semivarianzen $\gamma(h)$ und an der Ordinate die Entfernungsklassen h_k aufgetragen (vgl. Abbildung 3-2).

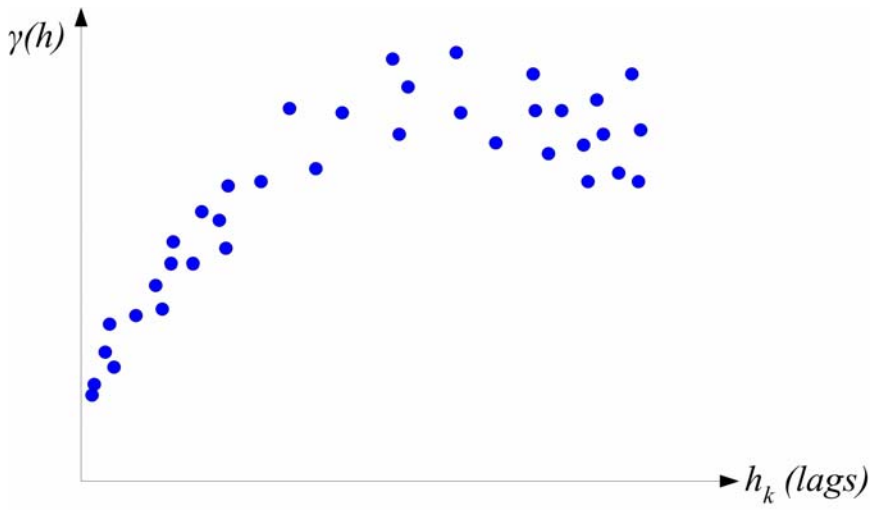


Abbildung 3-2: *Beispiel eines empirischen Semivariogramms*

Das empirische Semivariogramm zeigt also an, wie sehr Punktepaare, die derselben Abstandsklasse angehören, in ihren Werten durchschnittlich differieren. Mit Hilfe des empirischen Semivariogramms kann lediglich bestimmt werden, wie sich die räumliche Abhängigkeit in bestimmten Abständen verhält. Als Grundlage

für die Interpolation müssen jedoch Aussagen für beliebige Abstände gemacht werden können. Dazu wird angenommen, dass das aus den vorhandenen Messwerten generierte Semivariogramm den groben Verlauf des räumlichen Zusammenhangs im gesamten Untersuchungsgebiet wiederspiegelt, also repräsentativ ist. Um die fehlenden Abstände schätzen zu können, wird an das empirische Semivariogramm eine Funktion angepasst, welche den Verlauf möglichst gut wiedergibt (HAUSSMANN 2000, WALDOW 1998). Diese Funktion wird auch als **theoretisches Semivariogramm** bezeichnet. Das theoretische Semivariogramm ist in der Regel eine monoton wachsende Funktion, wenn man annimmt, dass nahe nebeneinander liegende Punkte in ihren Werten ähnlicher sind als weiter voneinander entfernte (HAUSSMANN 2000). Sie muss konditional bedingt negativ semidefinit sein, was für viele Funktionen nicht nachweisbar ist (ISAAKS & SRIVASTAVA 1989). Deshalb beschränkt man sich in der Praxis auf Familien von Funktionen (Modelle), für welche die Definitheit nachgewiesen ist. Durch die lineare Kombination dieser Modelle (verschachtelte Strukturen) kann der Umfang der möglichen Modelle beträchtlich erweitert werden (ISAAKS & SRIVASTAVA 1989). Die am häufigsten verwendeten Variogramm-Modelle sind nach GOOVAERTS (1997):

- **Das Nugget-Modell**

  Das Nugget-Modell charakterisiert Variablen deren Variabilität kleinräumiger als die geringsten Probenabstände ist. Der sogenannte Nuggeteffekt C_0 kann aber auch durch Fehler im Datensatz, entstanden bei der Probenahme oder der Analyse, verursacht werden (GERLACH 2001).

$$\gamma(h) = \begin{cases} 0 & wenn \quad h = 0 \\ C_0 & wenn \quad h \neq 0 \end{cases} \qquad \textbf{(3-5)}$$

- **Das sphärische Modell**

  Das sphärische Modell ist nach BROOKER (1986) und nach STOYAN et al. (1997) eines der am häufigsten genutzten Variogramm-Modelle. Das Modell ist gekennzeichnet durch ein lineares Verhalten bei geringen Abständen und einen allmählichen Übergang zum Plateau bei Erreichen des Schwellwertes C (vgl. Abbildung 3-4). Dabei ist a die Aussageweite, d.h. der Abstand h bei dem das Variogramm-Modell den Schwellenwert C erreicht.

$$\gamma(h) = Sph\left(\frac{h}{a}\right) = \begin{cases} 1{,}5\frac{|h|}{a} - 0{,}5\left(\frac{|h|}{a}\right)^3 & wenn \quad h \leq a \\ C & wenn \quad h > a \end{cases} \qquad \textbf{(3-6)}$$

- **Das exponentielle Modell**

  Dieses Modell hat bereits in der Nähe des Ursprungs einen gekrümmten Verlauf und erreicht den Schwellenwert nur asymptotisch. Als Aussageweite a (vgl. Abbildung 3-4) wird der Abstand gewählt, bei dem 95% des Schwellenwertes erreicht werden (PANNATIER 1996).

$$\gamma(h) = C \cdot \left(1 - e^{-\frac{3h}{a}}\right) \qquad \textbf{(3-7)}$$

- **Das Gauß´sche Modell**

  Das Gauß´sche Modell hat einen parabolischen Verlauf und erreicht den Schwellenwert nur asymptotisch. Wie beim exponentiellen Modell wird als Aussageweite der Abstand gewählt, bei dem 95% des Schwellenwertes erreicht werden. Das Gauß´sche Modell beschreibt Zufallsfunktionen von Variablen mit extrem hohen räumlichen Zusammenhang (GRAMS 2000).

$$\gamma(h) = C \cdot \left(1 - e^{-\frac{3h^2}{a^2}}\right) \qquad \textbf{(3-8)}$$

  Das Gauß´sche Modell findet Verwendung bei extrem kontinuierlichen Prozessen, z.B. bei der Modellierung glazialer Oberflächen im antarktischen Bereich oder der submarinen Morphologie (HERZFELD 1989). Aus numerischen Gründen sollte es nach ARMSTRONG (1998) mit einem Nugget-Modell kombiniert werden.

- **Das Potenzmodell (Powermodell)**

  Das Potenzmodell beschreibt einen selbstähnlichen, fraktalen Prozess und unterscheidet sich von den obigen Modellen dadurch, dass kein Schwellenwert erreicht wird (GRAMS 2000). Das Abbild dieses Prozesses ist in jeder Maßstabsebene gleich (KITANIDIS 1997).

$$\gamma(h) = c \cdot h^{\omega} \qquad \textbf{(3-9)}$$

  Der Exponent ω kann Werte zwischen $0 < \omega < 2$ annehmen. Für $\omega = 1$ ergibt sich das lineare Modell als ein Sonderfall des Potenzmodells:

$$\gamma(h) = c \cdot h \qquad \textbf{(3-10)}$$

  Die Konstante c gibt die Steigung der Geraden an. Dieses Modell eignet sich nur für Zufallsfunktionen, die keinen Schwellenwert erreichen. Für Zufallsfunktionen mit einer endlichen Varianz sollte dieser Modelltyp allerdings nicht verwendet werden (PANNATIER 1996).

Abbildung 3-3 zeigt jene Funktionsverläufe die am häufigsten für die Variogramm-Modelle verwendet werden.

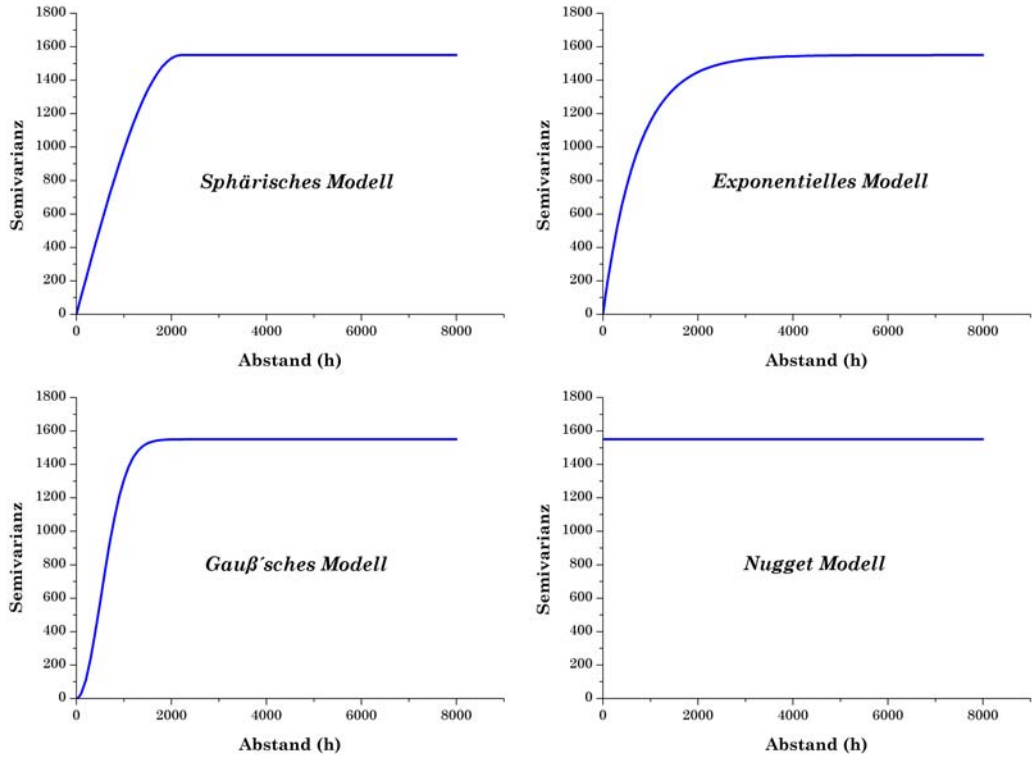


Abbildung 3-3: *Funktionsverläufe für Variogramm-Modelle*

Neben den oben angeführten Funktionen werden noch weitere Modelle (z.B. logarithmisch, periodisch) zur Anpassung an das empirische Semivariogramm verwendet, deren Funktionsverläufe z.B. in WALDOW (1998) beschrieben sind.

Der Abstand, vgl. mit Abbildung 3-4, bei dem das Semivariogramm asymptotisch den Wert der maximalen Varianz (**sill**) erreicht und dann parallel zur Abszisse verläuft stellt die Grenze des räumlichen Zusammenhangs dar und wird Aussageweite (**range**) genannt (GOOVAERTS 1997). Messwerte, die in größeren Abständen als der Aussageweite zueinander stehen, werden als räumlich unabhängig voneinander angesehen. Zur Schätzung von unbekannten Werten sollten nur Werte innerhalb dieser Aussageweite verwendet werden.

Jedes Semivariogramm beginnt bei Null, da die mittlere Varianz beim Abstand Null stets Null ergibt. Oft kommt es jedoch vor, dass schon in geringen Entfernungen ein Wertesprung sichtbar ist. Die Extrapolation der Semivariogrammfunktion auf die Ordinate ergibt dabei einen Schnittpunkt oberhalb des Ursprungs. Dieser Abstand wird „**Nugget-Effect**“ genannt und besteht aus zwei Komponenten (PLEYDELL et. al 2004, GOOVAERTS 1997):

$$Nugget\ Effect = Error\ Variance + Micro\ Variance \qquad \textbf{(3-11)}$$

Unter "**Error Variance**" wird die Varianz der Messfehler verstanden, der wiederum als Ausdruck für die Wiederholbarkeit oder Ersetzbarkeit der Messung verstanden werden kann. Unter „**Micro Variance**" wird eine zu erwartende kleinmaßstäbige Varianz der Daten verstanden.

Wenn sich die Error Variance Null nähert, hat ein Nugget-Effect zwar einen glättenden Einfluss, aber die resultierende Oberfläche nähert sich an den Stellen der Stützpunkte deren ursprünglichen Werten. Ein höherer Wert der Error Variance erlaubt ein immer stärkeres Abweichen an der Position der Stützpunkte. OLIVER & WEBSTER (1990) beschreiben den Nugget-Effect bzw. die Nugget-Varianz graphisch als die positive Abweichung, an der die Funktion des experimentellen Semivariogramms die Ordinate schneidet.

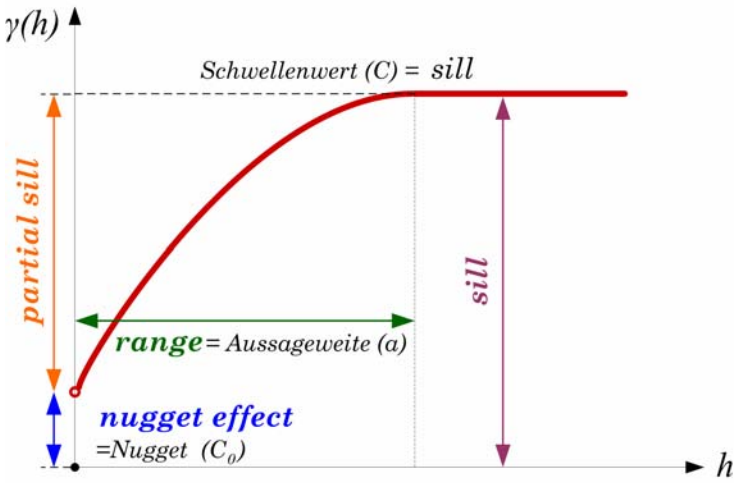


Abbildung 3-4: *Beispiel einer theoretischen Semivariogrammfunktion*

Mit der Einführung eines Nugget Effects geht man von der Prämisse ab, dass die Stützpunkte von der interpolierten Oberfläche genau getroffen werden müssen. Kriging ist dann kein exakter Interpolator mehr (Blaschke & Strobl 2003). Weiters macht die Angabe eines Nugget-Effects Kriging generell mehr zu einer glättenden Interpolationstechnik, indem Abweichungen der resultierenden Oberfläche auch an den zu interpolierenden Stützpunkten zugelassen werden, oder anders ausgedrückt: Der generelle Trend wird stärker berücksichtigt. Die Schätzungen gleichen bei erhöhtem Nugget Effect zunehmend mehr einem simplen lokalen Mittelwert (MYERS 1997). Aufgrund dieser Eigenschaften gibt der Nugget-Effect wichtige Hinweise, wie gut sich die Daten zur Interpolation eignen. Ein sehr hoher Nugget-Effect bedeutet, dass nur eine geringe räumliche Korrelation in den Stichproben zu beobachten ist. Weist ein Semivariogramm keine Steigung auf, so spricht man von einem reinen Nugget-Effekt. In diesem Fall eignen sich die Daten nicht zur Interpolation. Verbesserungen können dabei eventuell durch eine Vergrößerung der Stichprobe oder eine verbesserte Probenahme erreicht werden

(HEINRICH 1992). Gerade kleine Abstandsklassen sollten zur besseren Bestimmung kleinräumiger Variabilitäten ausreichend beprobt werden, um statistisch abgesicherte Werte zu erhalten (HEINRICH 1992).

Häufig zeigen experimentelle Semivariogramme Strukturen, die eine Modellanpassung erschweren, wie z.B. der sogenannte „**Locheffekt**“ (vgl. Abbildung 3-5 links). Beim Vorliegen eines Locheffekts fällt das experimentelle Semivariogramm nach Erreichen eines Maximums mit zunehmender Schrittweite (lag-Distanz) wieder deutlich ab, anstatt bei diesem Maximalwert zu bleiben. Der Maximalwert ist höher als die a priori Varianz (AKIN & SIEMES 1988). Ursache von Locheffekten ist das Nebeneinander von Bereichen mit sehr hohen und sehr niedrigen Werten z.B. das lokal begrenzte Vorkommen eines Schadstoffes in ansonsten unbelastetem Grundwasser (GRAMS 2000). AKIN & SIEMES (1988) schlagen vor, ein experimentelles Semivariogramm mit Locheffekt durch ein sphärisches Modell zu beschreiben, indem der Maximalwert von $\gamma(h)$ gleich dem Schwellenwert gesetzt wird.

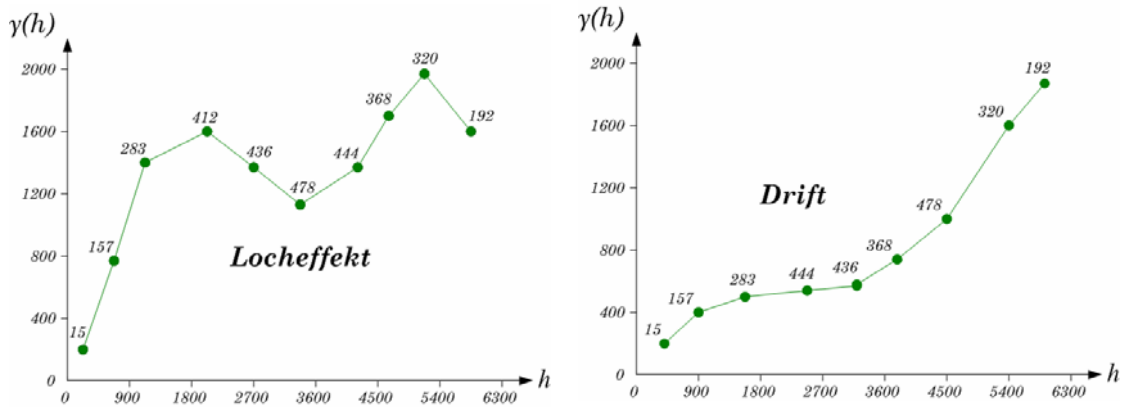


Abbildung 3-5: *Empirisches Semivariogramm mit Locheffekt und Drift*

Wenn experimentelle Semivariogramme nach Erreichen des Schwellenwertes mit zunehmendem Abstand h parabolisch ansteigen, kann dies ein Hinweis auf das Vorliegen eines **Drift**s sein (vgl. Abbildung 3-5 rechts). Das bedeutet, dass die ortsabhängige Variable bzw. die ihr zugrunde liegende Zufallsfunktion bei großen Abständen nicht mehr stationär ist (AKIN & SIEMENS 1988). Ein Drift (Trend) liegt dann vor, wenn eine richtungsgebundene, systematische Zu- oder Abnahme der Probenwerte erkennbar ist (GOOVAERTS 1997). In diesem Fall besitzt weder die Hypothese der Stationarität 2.Ordnung noch die intrinsische Hypothese Gültigkeit, so dass die klassischen Verfahren der Geostatistik nicht mehr eingesetzt werden dürften (AKIN & SIEMES 1998). Allerdings wird unter Anwendung der Hypothese der Quasistationarität trotz des Vorliegens einer Drift häufig das Normalkrigeverfahren angewandt. Bei der Hypothese der Quasistationarität wird angenommen, dass die intrinsische Hypothese lediglich

innerhalb einer größenmäßig definierten Zone, z.B. ein Kreis mir Radius r, gültig ist (GOOVAERTS 1997). Unter Annahme einer Quasistationarität können Variogrammwerte für Abstände $h \leq r$ zur Berechnung der Schätzwert herangezogen werden (ATKIN & SIEMES 1988).

Bisher wurde davon ausgegangen, dass die Zufallsvariablen sich isotrop verhalten, d.h. die räumliche Abhängigkeit in allen Richtungen gleich ist. Dies ist aber selten der Fall, weshalb eine Anisotropie bei der Berechnung des Semivariogramms und der späteren Modellanpassung berücksichtigt werden muss (GOOVAERTS 1997). In der Geostatistik werden zwei Formen der Anisotropie unterschieden: Eine geometrische Anisotropie liegt vor, wenn die Variogramme der verschiedenen Richtungen den gleichen Schwellenwert, aber verschiedene Aussageweiten aufweisen. Bei einer zonalen Anisotropie haben die Variogramme für verschiedene Richtungen unterschiedliche Schwellenwerte und unterschiedliche Aussageweiten (AKIN & SIEMES 1988).

Zur hinreichenden Bestimmung einer solchen Anisotropie benötigt man direktionale, also richtungsabhängige Semivariogramme in mindestens vier Richtungen (HAUSSMANN 2000). Aus diesen kann anschließend die Richtung der maximalen Anisotropie ermittelt werden. Direktionale Semivariogramme erhält man indem man zusätzlich zu den Abstandsklassen den Raum weiter diskretisiert. Dazu werden disjunkte Winkelklassen definiert. Auch hier gilt, dass minimal 30-50 Wertpaare in eine solche Winkelklasse fallen sollten, um statistisch abgesicherte Wert zu erhalten. Damit diese Winkelklassen mit zunehmender Entfernung nicht zu große Flächen einnehmen können maximale Bandbreiten definiert werden.

Abbildung 3-6 fasst nach DEUTSCH & JOURNEL (1998) die wichtigsten der oben vorgestellten Begriffe zusammen.

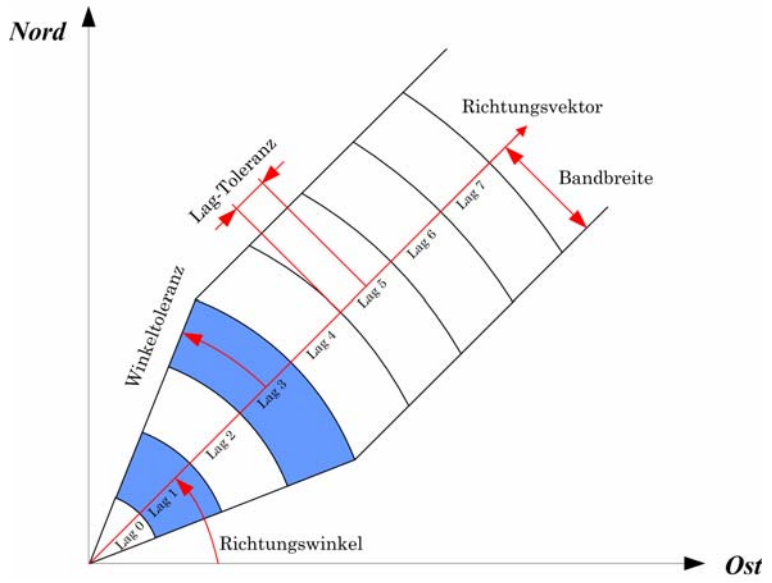


Abbildung 3-6: *Eingangsgrößen der variographischen Analyse*

3.2.2 Nicht normalverteilte Prozesse

In der Regel sollte ein zu interpolierender Prozess normalverteilt sein (STOYAN et. al 1997). Ist das nicht der Fall, wird häufig eine Transformation des Prozesses anhand der Beobachtungswerte durchgeführt (GOOVAERTS 1997). Für die Transformation der einer Interpolation zugrundeliegenden Werte sprechen zwei Gründe. Ein Grund besteht darin, dass bei einer positiven Schiefe (Verteilung ist rechtsschief), wie sie bei den Schwefeldaten auftritt, nur sehr wenige hohe Werte existieren und damit eine starke Unterschätzung in den Bereichen hoher Werte auftritt (HINTERDING et al. 2003). Den zweiten Grund stellen auftretende Probleme bei der Variogrammermittlung von nicht-normalverteilten Prozessen dar. In diesem Fall liefert der klassische Variogrammschätzer schlechte Ergebnisse.

Je nachdem wie die Beobachtungswerte verteilt sind, können unterschiedliche Transformationen herangezogen werden. Eine Möglichkeit ist die der lognormalen Transformation, die innerhalb des Lognormal Kriging Verwendung findet. Eine weitere Transformation stellt die Normal-Score Transformation dar, die im Multi-Gaußian Kriging verwendet wird.

Im Folgenden werden die Funktionsweisen dieser beiden Methoden kurz vorgestellt.

a) Lognormale Transformation

Mit dem Lognormal Kriging werden Schätzungen von Zufallsvariablen *Z(u)* vorgenommen, die im betrachteten Raum log-normalverteilt vorliegen, also eine positive Schiefe aufweisen.

Die bekannten Zufallsvariablen *Z(u)* werden zunächst durch Logarithmieren in die normalverteilten Zufallsvariablen *Y(u)* transformiert.

$$Y(u) = \log Z(u) \qquad \textbf{(3-12)}$$

Bei der Transformation der Werte ist besonders darauf zu achten, dass die logarithmierten Werte tatsächlich eine Normalverteilung aufweisen bzw. die Abweichung von der Normalverteilung vernachlässigbar ist (CHILES & DELFINER 1999). Anschließend wird ein Ordinary Kriging auf Basis der logarithmierten Werte vorgenommen, wobei die Zufallsvariable $Y^*(u)$ folgendermaßen geschätzt wird (CRESSIE 1993):

$$Y^*(u) = \sum_{i=1}^{n} w_i \log Z(u_i) = \sum_{i=1}^{n} w_i Y(u_i) \qquad \textbf{(3-13)}$$

Nach dem Ordinary Kriging werden die geschätzten Zufallsvariablen $Y^*(u)$ wieder in die ursprüngliche Verteilung zurück transformiert:

$$Z^*(u) = \exp\left\{Y^*(u) + \frac{1}{2}\sigma^2_{Y,k}(u) - \lambda\right\} \qquad \textbf{(3-14)}$$

Dabei ist $\sigma^2_{Y,k}(u)$ die Kriging-Varianz und λ der Lagrange-Multiplikator aus dem Ordinary Kriging der logarithmierten Werte (CRESSIE 1993). Durch die angegebene Formel der Rücktransformation wird eine erwartungstreue Schätzung sichergestellt, wobei jedoch Fragen über die Bedingung eines optimalen Schätzers offen bleiben (CILES & DELFINER 1999).

Um Probleme mit Nullwerten zu vermeiden, sollte diese Werte minimal erhöht werden (ARMSTRONG 1998). Eine Möglichkeit der Behandlung von Messwerten unter der Nachweisgrenze zeigt z.B. das Ministerium für Umwelt und Forsten in Deutschland auf. Danach werden zur Berechnung von Zahlenwerten die Messwerte unter der Nachweisgrenze auf die halbe Nachweisgrenze gesetzt (HINTERDING et. al 2003).

b) Normal-Score-Transformation

Die Normal-Score-Transformation ist eine graphische Transformation, welche die Normalisierung beliebiger Verteilungen unabhängig von ihrer Form zulässt. Die Normal-Score-Transformation erfolgt in drei Schritten (GOOVAERTS 1997):

1. Die wahren Wert $z(u_i)$ werden in eine aufsteigende Rangfolge geordnet:

   $$\left[z(u_i)\right]^{(1)} \le \ldots \le \left[z(u_i)\right]^{(k)} \le \ldots \le \left[z(u_i)\right]^{(n)} \qquad \textbf{(3-15)}$$

   wobei k den Rand des Wertes $z(u_i)$ unter allen Werten (n = Anzahl) angibt.

2. Die kumulative Häufigkeit des Werte $z(u_i)$ mit dem Rang k wird dann nach

   $$p^*_k = \frac{k}{n} - \frac{0{,}5}{n} \qquad \textbf{(3-16)}$$

   berechnet.

3. Der Wert $z(u_i)$ mit dem Rang k wird über das p^*_k -Quantil an die kumulative Standardnormalverteilung angepasst:

   $$y(u_i) = G^{-1}\left[F\left(z(u_i)\right)\right] = G^{-1}\left(p^*_k\right) \qquad \textbf{(3-17)}$$

   dabei stellt F die Verteilungsfunktion der Daten und G die Verteilungsfunktion der Standardnormalverteilung dar.

Danach wird ein Semivariogramm der transformierten Werte ermittelt und die transformierten Werte *y(u)* beispielsweise mittels Ordinary Kriging interpoliert, wobei sich wiederum das Problem der Rücktransformation ergibt. JONSTON et al. (2001) oder DEUTSCH & JORNEL (1998) transformieren z.B. die interpolierten Werte $y^*(u)$ mittels der Inversen der Normal-Score-Transformation direkt zurück. Dies führt aber dazu, dass die Schätzung nicht mehr erwartungstreu ist (GOOVAERTS & SAITO 2000). GOOVAERTS & SAITO (2000) schlagen deshalb alternativ einen anderen Weg vor, bei dem sie allerdings ein anderes Optimalitätsprinzip verfolgen als jenes der Erwartungstreue.

a) Zuerst wird die Gauß'sche Wahrscheinlichkeitsverteilung der Zufallsvariable *Y* am Ort *u* über den Mittelwert und die Varianz ermittelt. Der Mittelwert ist der mittels Ordinary Kriging geschätzte Wert $y^*(u)$ und die Varianz ist die über das Simple Kriging ermittelte Kriging-Varianz $\sigma^2(u)$.

b) Diese Wahrscheinlichkeitsverteilung wird diskretisiert, indem 100 Quantile $y_p(u)$ gebildet werden, die den Wahrscheinlichkeiten

$$p = \frac{k}{100} - \frac{0{,}5}{100} \qquad \textit{(mit: } k = 1,2, \ldots, 100\textit{)} \qquad \textbf{(3-18)}$$

entsprechen.

c) Die entsprechenden Quantile der lokalen Wahrscheinlichkeitsverteilung der Zufallsvariable *Z* werden ermittelt $(F(u;z|(n)) = Prob\{Z(u) \leq z|(n)\}$ und folgende Rücktransformation von $y_p(u)$ genutzt:

$$z_p(u) = F^{-1}\left[G\left(y_p(u)\right)\right] \qquad \textbf{(3-19)}$$

d) Die Kriging-Schätzung $z^*(u)$ wird aus dem Mittelwert der Verteilungsfunktion $F(u;z|(n))$ gebildet, in diesem Fall wie folgt:

$$z^*(u) = \frac{1}{100}\sum_{k=1}^{n} z_p(u) \quad \textit{(mit: } p = k/100 - 0{,}5/100\textit{)} \qquad \textbf{(3-20)}$$

Eine Annahme des Multi-Gaußian Kriging ist, dass die n-dimensionale Verteilung der Zufallsfunktion *Z(u)* normalverteilt ist (HINTERDING et. al 2003). Diese Annahme lässt sich in der Praxis nicht überprüfen, so dass nur die zweidimensionale Verteilung auf Bi-Normalverteilung überprüft wird (GOOVAERTS & SAITO 2000)

3.2.3 Ordinary Kriging

Wie bereits oben erwähnt wurde, wird im Kriging angenommen, dass sich der zu schätzende Prozess aus einer deterministischen und einer stochastischen Komponente zusammensetzt (HEINRICH 1992, CRESSIE 1993). Die Zufallsvariable *Z(u)* wird dabei als Summe der beiden Komponenten angesehen:

$$Z(u) = m(u) + R(u) \qquad \text{mit } u \in U \text{ (Untersuchungsgebiet)} \qquad \textbf{(3-21)}$$

mit:

Z(u)	*Zufallsvariable*
m(u)	*deterministische Komponente (großskaliger Anteil des Prozesses)*
R(u)	*stochastische Komponente (kleinskaliger Anteil des Prozesses)*

DEUTSCH & JOURNAL (1998) bezeichnen *m(u)* und *R(u)* auch als den großskaligen bzw. kleinskaligen Anteil des Prozesses. Im Ordinary Kriging (OK) wird die zu interpolierende Beobachtungsvariable als räumlicher stochastischer Prozess $Z = \{Z(u), u \text{ in Untersuchungsgebiet}\}$ mit einem unbekannten konstanten Mittelwert $m(u) = \text{konst.}$ Modelliert (GOOVAERTS 1997). Es seien $z(u_1),\ldots,z(u_n)$ Werte der Beobachtungsvariablen an den beprobten Orten $u_1,\ldots,u_n$. Diese werden als Realisationen der Zufallsvariablen $Z(u_1),\ldots,Z(u_n)$ angesehen. Das bedeutet jedoch nicht, dass die Gehalte völlig zufällig verteilt sind, sondern dass sie einer gewissen Wahrscheinlichkeitsstruktur folgen (HEINRICH 1992). Auch für die unbeprobten Orte werden Zufallsvariablen angenommen, deren Realisationen in einer räumlichen Interpolation aus den bekannten Realisationen geschätzt werden (WACKERNAGEL 1998, HINDERDING et al. 2000). Der stochastische Prozess wird auch als Zufallsfunktion bezeichnet.

Abbildung 3-7 zeigt vereinfacht das Ablaufschema eines räumlich stochastischen Prozesses nach HEINRICH (1992). Die Ausgangssituation des Ablaufschemas ist ein stochastischer Prozess, dessen Realisierung geschätzt werden soll. Die Zufallsvariablen an den beprobten Orten können anhand eines Messnetzes und den dort durchgeführten Messungen ermittelt werden. In Folge wird basierend auf der Schätzung des räumlichen Zusammenhangs der Zufallsvariablen des zu untersuchenden Prozesses eine bestmögliche Schätzung für jene Gebiete durchgeführt, für welche keine Beobachtungen vorliegen (HEINRICH 1992). Mit Hilfe der vorliegenden Daten der Beobachtungsvariablen wird eine Annahme über den räumlichen Zusammenhang der Werte gemacht, welcher auch als Autokorrelation bezeichnet wird. Mit dem Krige-Schätzer wird die Prozessrealisierung geschätzt und z.B. als Rastergraphik abgebildet (vgl. Abbildung 3-7).

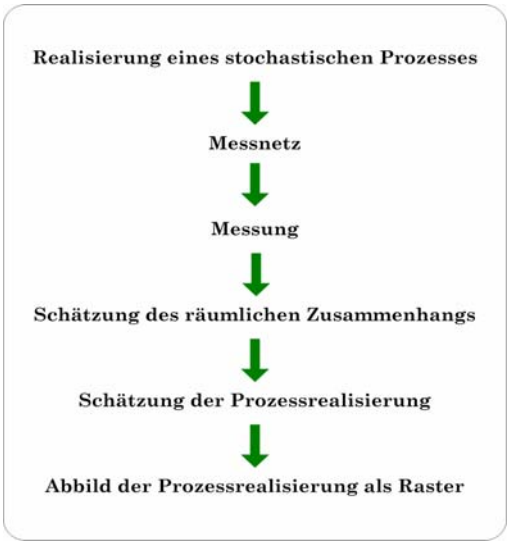


Abbildung 3-7: *Ablaufschema zur Schätzung eines stochastischen Prozesses*

Die räumliche Abhängigkeit der Beobachtungsvariablen darf lediglich von der Entfernung der Punktepaare zueinander abhängen, nicht jedoch von deren absoluten Lage im Raum (HAUSSMANN 2000). D.h. sie muss räumlich stationär und unempfindlich gegenüber Translationen sein. Da in der Praxis meist nur beschränkte Abstände zwischen den Punkten zur Berechnung der Parameter für die Interpolation benötigt werden, genügt es, wenn diese Stationarität innerhalb kleinerer Abstände gegeben ist (GOOVEARTS 1997). Die Annahme einer solchen Quasistationarität wird auch intrinsische Hypothese genannt (SCHERELIS & BLÜMEL 1988, JOURNAL & HUIJBREGTS 1978). Da die räumliche Abhängigkeit der Messpunkte als Grundlage für die Interpolation benötigt wird, muss diese im ersten Schritt bestimmt werden. Der stochastische Prozess Z wird somit als intrinsisch stationär angesehen. Das bedeutet zum einen, dass der Erwartungswert E der Zufallsvariablen dieses Prozesses konstant ist,

$$E[Z(u)] = const. \qquad \textbf{(3-22)}$$

für alle Orte u im Untersuchungsgebiet. Zum anderen hängt der räumliche Zusammenhang zwischen zwei Zufallsvariablen dieser Zufallsfunktion nicht von deren absoluter Lage, sondern nur von deren Abstandsvektor h ab (HINTERDING 1998). Unter der Annahme der intrinsischen Stationarität kann der räumliche Zusammenhang zwischen zwei Zufallsvariablen $Z(u_1)$ *und* $Z(u_2)$ des stochastischen Prozesses an den Orten u_1 und u_2 mit u_1-u_2=h, durch das so genannte Semivariogramm (vgl. Gleichung 3-4) beschrieben werden (GOOVAERTS 1997).

Das Semivariogramm ist eine Funktion des Abstandsvektors h. Es variiert mit dem Abstand und der Richtung. Ist das Semivariogramm unabhängig von der Richtung, gilt also $\gamma(h) = \gamma(|h|)$, wird es als isotrop bezeichnet. Das Semivario-

gramm wird aus den Stichprobendaten geschätzt, um den räumlichen Zusammenhang der vorliegenden Daten zu bestimmen (HINTERDING 1998). Die Variogrammabschätzung wurde bereits in Kapitel 3.2.1 „*Semivariogramm*" näher erläutert.

Um einen Wert $Z(u_0)$ an dem unbeprobten Ort u_0 aus den Probenwerten $Z(u_i)$ an den Orten $u_1,\ldots,u_n$ der Beobachtungsvariable zu schätzen, wird im Ordinary Kriging der folgende Krige-Schätzer verwendet (STOYAN et. al 1997):

$$Z^*(u_0) = \sum_{i=1}^{n} w_i Z(u_i) \qquad \textbf{(3-23)}$$

mit:

$Z^*(u_0)$ *Krige-Schätzer*

w_i *Gewichte des Krige-Schätzers*

$Z(u_i)$ *Probewerte an den Orten* u_i

Die Gewichte w_i des Krige-Schätzers für einen Punkt u_0 werden so bestimmt, dass

1. der Schätzfehler $F(u_0) = Z^*(u_0) - Z(u_0) = \sum w_i Z(u_i) - Z(u_0)$ im Mittel gleich 0 ist: $E[F(u_0)] = 0$ und **(3-24)**
2. die Varianz des Schätzfehlers minimal ist:
   $Var[F(u_0)] = min \quad \{F(u_0), w_1, \ldots, w_n \ reelle\ Zahlen\}$ **(3-25)**

$Var[F(u_0)]$ heißt in diesem Zusammenhang die Ordinary-Kriging-Varianz. Der Krige-Schätzer wird deshalb auch als sogenannter „*best linear unbiased predictor* ***(BLUP)***" bezeichnet (WERTHER 2003, HOFFMANN 2002, JOHNSTON et al. 2001, SCHAFMEISTER 1999). Krigingverfahren minimieren also den Schätzfehler (best) und liefern unter allen linearen (linear) eine erwartungstreue (unbiased) Vorhersage (predictor). In der Literatur werden die Begriffe Vorhersage und Schätzung oft synonym verwendet und anstatt von einem Prädiktor von einem Schätzer (estimator) gesprochen. Von einem „*best linear unbiased estimator* ***(BLUE)***" spricht man allerdings bei Schätzproblemen, wenn es darum geht, unbekannte aber feste Parameter zu schätzen (WERTHER 2003).

Um die Gewichte des Krige-Schätzers für einen Punkt u_0 zu bestimmen, wird die Extremwertaufgabe

$$Var[F(u_0)] = \min \quad \{F(u_0), w_1, \ldots, w_n \ reelle\ Zahlen\} \qquad \textbf{(3-26)}$$

unter der Nebenbedingung

$$E[F(u_0)] = 0 \qquad \textbf{(3-27)}$$

gelöst.

Die Nebenbedingung kann vereinfacht werden zu (HINTERDING 1998):

$$0 = E[F(u_0)] = E[Z(u_0) - \sum_{i=1}^{n} w_i Z(u_i)] = E[Z(u_0)] - (\sum_{i=1}^{n} w_i E[Z(u_i)]) =$$

$$= E[Z(u_0)](1 - \sum_{i=1}^{n} w_i) \qquad \textbf{(3-28)}$$

Daraus folgt:

$$1 = \sum_{i=1}^{n} w_i \qquad \textbf{(3-29)}$$

Dabei wurden die Stationarität des stochastischen Prozesses und die Linearität des Erwartungswertes ausgenutzt. Ohne auf die Einzelheiten der Berechnung einzugehen, gilt für die Varianz des Schätzfehlers $Var[F(u_0)]$:

$$Var[F(u_0)] = 2\sum_{i=1}^{n} w_i \gamma(u_i - u_0) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma(u_i - u_j) - \gamma(0) \qquad \textbf{(3-30)}$$

mit:

$Z^*(u_0)$	*Krige-Schätzer*
w_i	*Gewichte des Krige-Schätzers*
$Z(u_i)$	*Probewerte an den Orten* u_i

Die Extremwertaufgabe mit Nebenbedingungen wird mit der Methode der Lagrange-Multiplikatioren gelöst (JOHNSTON et. al 2001, GOOVAERTS 1997). Diese soll hier nicht weiter erläutert werden. Die Extremwertaufgabe führt zu einem linearen Gleichungssystem, dessen eindeutige Lösung die Gewichte des Krige-Schätzers sind (HERBST 2001, GOOVAERTS 2001):

$$1) \ \sum w_i \gamma(u_i - u_j) + \lambda = \gamma(u_i - u_0) \qquad \textbf{(3-31)}$$

$$2) \ \sum_{i=1}^{n} w_i = 1 \qquad \textbf{(3-32)}$$

Dabei ist λ eine Hilfsvariable, die aus der Methode der Langrange-Multiplikatoren stammt und für den Schätzer keine Bedeutung hat. In der Matrix-Schreibweise sieht das Gleichungssystem wie folgt aus (ISAAKS & SRIVASTAVA 1989):

$$\begin{bmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} * \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix} \qquad \textbf{(3-33)}$$

$$= C \qquad = w \qquad = D$$

Die Matrix *D* enthält die Semivariogrammwerte für die jeweiligen Abstände zwischen dem zu schätzenden Punkt u_0 und den zur Schätzung verwendeten Nachbarpunkten $x_1,\ldots,x_n$ (HAUSSMANN 2000). Diese Abstände entsprechen damit statistischen Abständen. Die Matrix *C* nimmt die statistischen Abstände zwischen den einbezogenen Nachbarn $x_1,\ldots,x_n$ des zu schätzenden Wertes auf. Damit werden räumliche Cluster erfasst. Liegen Nachbarpunkte statistisch nah nebeneinander, liefern sie für das Schätzen redundante Informationen. In diesem Fall werden die Gewichte gesenkt und den Gewichten der anderen Werte zugeschlagen.

Dabei ist $\gamma_{ij} = \gamma(u_i - u_j)$. Die Matrix ist aufgrund der Definitheit von γ invertierbar. Die Lösung des Gleichungssystems lautet daher:

$$w = C^{-1}D \qquad \textbf{(3-34)}$$

Die Gewichte des Krige-Schätzers sind damit für den Ort u_0 bestimmt.

3.2.4 Cokriging – Ordinary Cokriging

Für viele Anwendungen ist es sinnvoll, wenn an den Messpunkten mehrere Aufnahmegrößen zugleich erhoben werden, wie z.B. bei Luftgütemessungen neben den O_3-Gehalt auch der SO_2-Gehalt oder bei geologischen Untersuchungen die Teufen mehrerer übereinander liegender Horizonte (STOYAN et. al. 1997). In diesen Fällen interessieren dann auch die Korrelationen zwischen diesen Werten und das räumliche Verhalten der Korrelation.

Ein der multivariaten Situation angepasstes Kriging verfahren ist das Cokriging. Beim Cokriging werden zur Vorhersage einer Beobachtungsvariable $Z_1(u)$ in einem Punkt nicht nur die Messwerte dieser Beobachtungsvariable berücksichtigt, sondern auch die Messwerte einer (oder mehrer) Zusatzinformation(en), indem man deren Korrelation ausnutzt (GOOVAERTS 1997).

Die Zufallsvariable $Z_1^*(u)$ wird aus den beiden Variablen $Z_1(u)$ und $Z_2(u)$ wie folgt geschätzt:

$$Z_1^*(u) = \sum_{i=1}^{n} w_i Z_1(u_i) + \sum_{j=1}^{m} w_j Z_2(u_j) \qquad \textbf{(3-35)}$$

mit:

$Z_1^*(u)$	*Zufallsvariable*
$Z_1(u_i)$	*Beobachtungsvariable*
$Z_2(uj)$	*Sekundäre Variable*
w_i, w_j	*Gewichte*
n,m	*Anzahl der Nachbarn*

Dabei stellen n und m die Anzahl der Nachbarn dar, die in die Schätzung der Beobachtungsvariable $Z_1(u)$ einbezogen werden. $Z_1(u_i)$ und $Z_2(u_j)$ sind die gemessenen Werte der jeweiligen Nachbarschaft.

Die Schätzung ist eine gewichtete Summe der Nachbarwerte beider Variablen $Z_1(u)$ und $Z_2(u)$. Die Gewichte w_i und w_j werden unter den Bedingungen bestimmt, dass die Schätzung erwartungstreu ist und die Varianz des Schätzfehlers minimiert wird. Um das zugehörige Gleichungssystem zu lösen, muss der räumliche Zusammenhang zwischen den Variablen Z_1 und Z_2 mit Hilfe eines Kreuz-Variogramms modelliert werden. Dabei stellen u_1 und u_2 Orte des stochastischen Prozesses mit dem Abstand u_1-u_2 = h dar. Das Kreuz-Variogramm wird wie folgt beschrieben (DEUTSCH & JOURNAL 1998, GOOVAERTS 1997, JOURNEL & HUIJBREGTS 1978):

$$\gamma_{Z_1Z_2}(h) = E\left[(Z_1(u_1) - Z_1(u_2))(Z_2(u_1) - Z_2(u_2))\right] \qquad \textbf{(3-36)}$$

mit:

$\gamma_{Z_1Z_2}$	*Kreuz-Variogramm*
E	*Erwartungswert*
$Z_1(u_i)$	*Beobachtungsvariable*
$Z_2(uj)$	*Sekundäre Variable*

Cokriging kann ähnlich wie der Begriff Kriging als Oberbegriff für eine Reihe von geostatistischen Schätzverfahren wie z.B. *ordinary cokriging, simple cokriging, universal cokriging, indicator cokriging, probabilty* cokriging oder *disjunctive cokriging* angesehen werden.

In der vorliegenden Arbeit wird für den Vergleich der unterschiedlichen Interpolationsverfahren das **Ordinary Cokriging** (OCK) herangezogen. Die methodischen Grundlagen dieses Verfahrens decken sich zum Teil mit jenen des Ordinary Kriging Verfahrens, weshalb an dieser Stelle nicht näher darauf eingegangen wird. Zur Vertiefung in das Ordinary Cokriging Verfahren, bzw. allgemein in die Cokriging-Verfahren, sei auf die Arbeiten von JOHNSTON et. al (2001), CHILES & DELFINER (1999), GOOVEARTS (1998, 1997), CRESSIE (1993), ISAAKS & SRIVASTAVA (1989) und JOURNEL & HUIJBREGTS (1978) verwiesen, welche eine detaillierte Beschreibung der entsprechenden Methoden beinhalten.

Anzumerken ist, dass bei allen Interpolationsverfahren, welche Zusatzinformationen verwenden, die geeignete Auswahl der Zusatzvariablen entscheidend ist. Während geeignete Variablen das Interpolationsergebnis verbessern, können ungeeignete Variablen sogar zu einer Verschlechterung des Ergebnisses führen (BLÖSCHL & GRAYSON 2000).

3.3 Qualitätsermittlung von Interpolationsverfahren

Die kartographische Darstellung der interpolierten Werteoberfläche ist mit Hilfe von Geoinformationssystemen wie z.B. ArcView oder Geomedia relativ einfach möglich. Die graphische Darstellung gibt allerdings keine Antwort auf die Frage nach der Qualität der verwendeten Interpolationsverfahren, womit aber auch keine Entscheidung für oder gegen ein bestimmtes Verfahren getroffen werden kann.

Für die erfolgreiche räumliche Interpolation einer Werteoberfläche eines bestimmten Parameters ist deshalb eine intersubjektiv nachvollziehbare Prüfung der Qualität der entsprechenden Interpolation bzw. Methode unverzichtbar (DUMFARTH & LORUP 2000). Gebräuchlich ist hierfür in der Geostatistik unter anderem das Verfahren der Kreuzvalidierung (COLLINS & BOLSTAD 2003, GOOVAERTS 2001, GOOVAERTS 1997, CRESSIE 1993, ISAAKS & SRIVASTAVA 1989, DAVIS 1987).

Im Folgenden wird kurz auf die Methodik der Kreuzvalidierung und die zur Auswertung ihrer Ergebnisse herangezogenen statistischen Kennzahlen eingegangen.

3.3.1 Die Methode der Kreuzvalidierung

Die Methode der Kreuzvalidierung, engl. „Cross Validation“, ist eine in der Literatur anerkannte einfache Methode um die Qualität von räumlichen Interpolationsverfahren zu ermitteln. Die Methode der Kreuzvalidierung erlaubt einen Vergleich geschätzter und wahrer Werte allein auf der Basis der vorhandenen Stichprobe und ist hilfreich bei der Auswahl aus verschiedenen Interpolationsverfahren (GOOVAERTS 2001, ISAAKS & SRIVASTAVA 1989, Davis 1987).

Bei der Anwendung des Grundprinzips der Kreuzvalidierung wird ein Datensatz in zwei Teile aufgeteilt um so den Prognosefehler abschätzen zu können (HENNIG 2004). Der Nachteil dieser Vorgehensweise ist, dass zur Anpassung des Modells nur mehr die Hälfte der Daten zur Verfügung steht, was insbesondere bei kleinen Datensätzen die Qualität der Schätzung stark beeinträchtigen kann (HENNIG 2004). Darüber hinaus ist das Ergebnis vom Zufall abhängig, nämlich von der konkreten Aufteilung der Daten. Die älteste Verfeinerung der Kreuzvalidierung stammt aus den sechziger Jahren des vorigen Jahrhunderts und behebt beide Probleme (HENNIG 2004).

Die Idee dieser Verfeinerung besteht darin, aus einer vorhandenen Stichprobe eines Parameters einen gemessenen Wert zu entfernen und diesen mittels der anderen Werte der Stichprobe unter Verwendung der Interpolationsmethode zu schätzen (HENNIG 2004, SODOUDI 2004, GOOVAERTS 1997, CRESSIE 1993). Man teilt den Datensatz also *n* mal in zwei Teile mit *n-1* und einer Beobachtung auf. Wenn der Wert geschätzt ist, kann er mit dem wahren und zuvor entfernten Wert an der entsprechenden Messstelle verglichen werden. Dieser Vorgang wiederholt sich für alle verfügbaren Stichprobenwerte. In der für die Interpolation zur Verfügung stehenden Stichprobe fehlt somit immer nur jener Wert, der geschätzt werden soll. Es ist also jedem an einem bestimmten Ort u_i befindlichen wahren Wert $W(u_i)$ ein an gleicher Position befindlicher geschätzter Wert $W^*(u_i)$ zuzuordnen (GERLACH 2001).

Die Differenz zwischen wahren und geschätzten Wert bildet denn lokalen Fehler der Schätzung *SF* am Ort u_i (ISAAKS & SRIVASTAVA 1989):

$$SF(u_i) = S(u_i) - S^*(u_i) \qquad \textbf{(3-37)}$$

mit:

$SF(u_i)$ *Fehler der Schätzung am Ort* u_i

$S(u_i)$ *Wahrer Schwefelwert am Ort* u_i

$S^*(u_i)$ *Geschätzter Schwefelwert am Ort* u_i

3.3.2 Statistische Kennzahlen zur Auswertung der Kreuzvalidierung

Die Ergebnisse der Kreuzvalidierung können anhand von verschiedenen statistischen Kennwerten und empirischen Verteilungen analysiert werden. Diese Analyse gibt in Folge Aufschluss über die Qualität eines Interpolationsverfahrens (ISAAKS & SRIVASTAVA 1989). In diesem Kapitel werden jene statistischen Kennzahlen vorgestellt, die in der vorliegenden Arbeit zur Auswertung der Kreuzvalidierung herangezogen werden.

Bei einem positiven Schätzfehler *SF* aus Gleichung 3-37 ist der geschätzte Wert niedriger als der auf der gleichen Position befindliche wahre Wert. Die Oberfläche bleibt lokal hinter der durch den wahren Wert definierten tatsächlichen Lage zurück. Somit handelt es sich um eine lokale Unterschätzung (DUMFARTH & LORUP 2000). Ist der Schätzfehler *SF* hingegen negativ, ist der Schätzwert zu hoch und die Oberfläche übersteigt lokal die tatsächliche Lage. Hierbei handelt es sich um eine lokale Unterschätzung (DUMFARTH & LORUP 2000). Ist der mittlere Schätzfehler *SF* einer Stichprobe annährend Null,

$$\frac{1}{n}\sum_{i=1}^{n} S(u_i) - S^*(u_i) \approx 0 \qquad \textbf{(3-38)}$$

mit:

$S(u_i)$	*Wahrer Schwefelwert am Ort* u_i
$S^*(u_i)$	*Geschätzter Schwefelwert am Ort* u_i
n	*Stichprobenumfang*

wird zunächst davon ausgegangen, dass keine Tendenz und somit eine erwartungstreue (unbiased) Schätzung vorliegt. Hingegen kann ein signifikanter negativer (oder positiver) mittlerer Schätzfehler eine systematische Überschätzung (bzw. Unterschätzung) repräsentieren (WACKERNAGEL 2000, DUMFARTH & LORUP 2000).

Die Beschreibung der univariaten Verteilung der Schätzfehler *SF* einer Stichprobe in Form eines Histogramms ermöglicht die Ermittlung einer allfälligen systematischen Über- oder Unterschätzung. Die Schätzfehler *SF* werden auch als Residuenwerte bezeichnet. Zur Abschätzung der Art der Verzerrung ist die Erhebung des arithmetischen Mittels $\bar{x}$ der Residuen notwendig. Im Falle einer Überschätzung ist das arithmetische Mittel kleiner Null und bei einer Unterschätzung größer als Null (ISAAKS & SRIVASTAVA 1989).

Die Berechnung des **arithmetischen Mittels** $\bar{x}$ geht aus Gleichung 3-39 hervor.

$$\bar{x} = \frac{1}{n} SF \qquad \textbf{(3-39)}$$

mit:

$\bar{x}$	*arithmetisches Mittel der n-Residuenwerte*
SF	*i-ter Residuenwert*
n	*Anzahl der Residuenwerte*

Diese Form der Verzerrungsanalyse ist für eine einwandfreie Aussage noch nicht ausreichend. So kann z.B. eine Verteilung auch verzerrungsfrei sein, wenn das arithmetische Mittel wegen der Kombination vieler kleiner Unterschätzungen mit einigen hohen Überschätzungen oder umgekehrt gleich Null ist. Eine daraus resultierende schiefe Verteilung würde ebenfalls eine Verzerrung der geschätzten Oberfläche implizieren (ISAAKS & SRIVASTAVA 1989). Daher sollte auch eine annähernd symmetrische Verteilung ein Auswahlkriterium für eine geeignete Interpolation sein.

Ein sehr einfaches Schiefemaß für metrisch-skalierte Variablen ist die Differenz zwischen **Median** x_m und arithmetischem Mittel $\bar{x}$ (GERLACH 2001). Umso größer die Differenz dieser beiden Kennwerte, umso stärker asymmetrisch ist die

Residuenverteilung. Der Median ist wie das arithmetische Mittel ein Lagemaß und jener Wert, der die der Größe nach aufsteigend geordneten Residuenwerte in zwei gleich große Intervalle teilt (STOYAN et. al 1997, BARTSCH 1991). Der Median ist somit ein Lagemaß für ordinal-skalierte Variablen, welches auch auf metrisch-skalierte Variablen angewandt werden kann (STREIT 2001).

Ein anderes, in der Statistik gebräuchliches Maß für die Schiefe ist der **Schiefekoeffizient *g*** nach Pearson (GERLACH 1991). Dieser wird durch das Potenzmoment 3.Ordnung und die Standardabweichung (vgl. Gleichung 3-41) folgendermaßen definiert (HARTUNG et al. 2002):

$$g = \frac{\frac{1}{n}\sum_{i=1}^{n}(SF - \bar{x})^3}{\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(SF - \bar{x})^2\right)^3}} \qquad \textbf{(3-40)}$$

mit:

g	*Schiefekoeffizient*
SF	*i-ter Residuenwert*
$\bar{x}$	*Arithmetisches Mittel der n-Residuenwerte*
n	*Anzahl der Residuenwerte*

Ist $g = 0$, bedeutet dies, dass die Häufigkeitsverteilung symmetrisch ist (STOYAN et. al 1997). Je stärker negativ g ist, desto linksschiefer bzw. negativ schiefer ist die Verteilung. Die Verteilung ist umso rechtsschiefer bzw. positiv schief, je stärker positiv der Wert g ist. Die Ermittlung der Schiefe ist nur sinnvoll, wenn die Häufigkeitsverteilung eingipfelig ist (HARTUNG et al. 2002).

Eine weitere Eigenschaft, die man sich von der Residuenverteilung erhofft ist eine geringe Streuung der Residuenwerte um ihren Mittelwert. Die **Standard-abweichung *s*** ist ein geeignetes Maß für die Schätzung der Streuung der Werte (HARTUNG et al. 2002):

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(SF - \bar{x})^2} \qquad \textbf{(3-41)}$$

mit:

s	*Standardabweichung*
SF	*i-ter Residuenwert*
$\bar{x}$	*Arithmetischem Mittel der n-Residuenwerte*
n	*Anzahl der Residuenwerte*

Je größer der Wert der Standardabweichung ist, umso weiter streuen die Residuenwerte um ihren Mittelwert (ISAAKS & SRIVASTAVA 1989). Laut ISAAKS & SRIVASTAVA (1989) ist eine Schätzung mit einer geringen Abweichung von der Erwartungstreue und einer geringen Standardabweichung der Residuenwerte gegenüber einer erwartungstreuen Schätzung mit einer großen Standardabweichung vorzuziehen.

Die Darstellung von Häufigkeitsverteilungen erfordert eine Klasseneinteilung der darzustellenden Werte. Eine Klasseneinteilung ist ein relativ subjektives Verfahren und sollte daher fachlich sinnvoll vorgenommen werden (STREIT 2001).

Die Größe des **mittleren quadrierten Schätzfehlers** (engl. *Mean Square Error*, **MSE**) ist ein weiterer, hilfreicher statistischer Kennwert für den Vergleich verschiedener Schätzungen (WACKERNAGEL 2000, BATTAGLIA 1996):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(S(u_i) - S^*(u_i))^2 \qquad \textbf{(3-42)}$$

mit:

MSE	*Mittlerer quadrierter Schätzfehler*
$S(u_i)$	*Wahrer Schwefelwert am Ort* u_i
$S^*(u_i)$	*Geschätzter Schwefelwert am Ort* u_i
n	*Stichprobenumfang*

Je kleiner der Wert des *MSE* ist, umso besser ist auch die entsprechende Schätzmethode (HINTERDING et. al 2003). Neben dem mittleren quadrierten Schätzfehler *MSE* ist auch der **mittlere absolute Schätzfehler** (engl. *Mean Absolute Error*, **MAE**) ein wichtiger Kennwert zur Abschätzung der Qualität einer Methode (HINTERDING et. al 2003, BRINKMANN 2002, ISAAKS & SRIVASTAVA 1989):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|S(u_i) - S^*(u_i)\right| \qquad \textbf{(3-43)}$$

mit:

MAE	*Mittlerer absoluter Schätzfehler*
$S(u_i)$	*Wahrer Schwefelwert am Ort* u_i
$S^*(u_i)$	*Geschätzter Schwefelwert am Ort* u_i
n	*Stichprobenumfang*

Auch hier gilt, je kleiner der Wert des *MAE* ist, umso besser ist die entsprechende Schätzmethode (HINTERDING et. al 2003). Die beiden angeführten Kennwerte der Schätzfehler unterscheiden sich dadurch, dass bei der Berechnung des *MSE* durch das Quadrieren höhere Residuenwerte stärker gewichtet werden. Außerdem

werden Residuenwerte, die einen Wert kleiner als eins annehmen, in ihrer Gewichtung stark minimiert. In die Berechung des *MAE* gehen die verschiedenen Residuenwerte gleichgewichtet ein.

Optisch gute Einblicke in die Qualität der vorgenommen Schätzung vermitteln auch so genannte **Korrelationsdiagramme**. In diese werden die wahren gegen die geschätzten Schwefelwerte aufgetragen. Bei einer optimalen Schätzung liegen die wahren und geschätzten Werte exakt an einer in einem Winkel von 45° stetig ansteigenden Geraden. In der Realität treten aber bei jeder Interpolation Fehlschätzungen auf. Daher ist bei jedem Korrelationsdiagramm dieser Art die Hauptdiagonale von einer mehr oder minder gut an sie angepassten Punktewolke umgeben. Je genauer die Schätzungen der gewählten Methode sind, d.h. umso ähnlicher sich die wahren und geschätzten Werte sind, umso besser geeignet ist die Methode für die Interpolation (ISAAKS & SRIVASTAVA 1989).

Die Qualität der betreffenden Methode kann in diesem Zusammenhang über die Ermittlung beispielsweise des **Rangkorrelationskoeffizienten r_s** nach Spearman zwischen den wahren und geschätzten Werten quantifiziert werden (GERLACH 2001). Der Rangkorrelationskoeffizient ist ein Maß für die Korrelation zwischen zwei ordinal skalierten Merkmalen, die nicht aus einer normalverteilten Grundgesamtheit stammen müssen (LÜTZKENDORF 2001). Er gibt Aufschluss über die Stärke des Zusammenhangs zwischen den beiden Variablen und leistet gute Dienste als Maßzahl für die Nähe der Punkte zur Diagonalen „*Wahrer Wert = Geschätzter Wert*“. Eine Interpolation mit einem hohen Koeffizienten wir im Allgemeinen anderen, durch niedrigere Werte gekennzeichneten Schätzungen, vorzuziehen sein. Der Wert des Koeffizienten liegt stets im Intervall [-1, +1]. Ist r = 1, liegen die Punkte exakt an der aufsteigenden Geraden „*Wahrer Wert = Geschätzter Wert*“. Ist r = -1, liegen sie exakt an der absteigenden Geraden „*Wahrer Wert ≠ Geschätzter Wert*“. Je näher r an +1 liegt, umso besser ist die Punktewolke an die Hauptdiagonalen des Korrelationsdiagramms angepasst. Je näher sich r der 0 annähert, umso diffuser erscheint die Punktewolke (DUMFARTH & LORUP 2000).

Der Rangkorrelationskoeffizient nach Spearman (monotoner Zusammenhang) wird wie folgt berechnet(HARTUNG et al. 2002):

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n(n^2-1)} \qquad \textbf{(3-44)}$$

mit:

r_s *Rangkorrelationskoeffizient*

d_i *Differenz zwischen den Rangzahlen zweier Variablen der i-ten Stichprobe*

n *Stichprobenumfang*

Die in diesem Kapitel vorgestellten Methoden werden in der Literatur häufig, wie beispielsweise von ISAAKS & SRIVASTAVA (1989), zur statistischen Analyse der Ergebnisse einer Kreuzvalidierung herangezogen. Vergleiche der Qualität verschiedener Interpolationsverfahren zur Modellierung eines Parameters finden sich z.B. bei PALOMINO & MARTIN (1994), sowie bei COLLINS & BOLSTAD (2003). PALOMINO & MARTIN (1994) verwenden die statistischen Kennwerte der Standardabweichung und des Korrelationskoeffizienten. COLLINS & BOLSTAD (2003), die in ihrem Artikel verschiedene Interpolationstechniken zur Schätzung der Lufttemperatur vergleichen, erheben die statistischen Kennwerte des arithmetischen Mittels $\bar{x}$ der Residuen, des *MAE* und des *MSE* zur Auswertung der Kreuzvalidierung. Die Ermittlung der statistischen Kennzahlen erfolgte in der vorliegenden Arbeit mit Hilfe der Statistiksoftware Origin 7.0.

3.4 Vorgehensweise und verwendete Software

In diesem Abschnitt wird auf die praktische Vorgehensweise bei der Durchführung der räumlichen Interpolation der Schwefeldaten eingegangen, deren wichtigsten Arbeitsschritte in Abbildung 3-8 dargestellt sind. Als Software für die Interpolation wurde der **ArcGIS™ Geostatistical Analyst** verwendet.

Abbildung 3-8: *Arbeitsschritte der räumlichen Interpolation*

a) Datenaufbereitung

Die Aufbereitung der Datensätze ist einer der wichtigsten Arbeitsschritte vor dem Beginn der eigentlichen Interpolation. Da die Schwefeldaten, wie bereits in Kapitel 2.2.3 „*Schwefeldaten*“ erwähnt wurde, von unterschiedlichen Institutionen stammen, lagen diese einerseits in Form einer Access-Datenbank und andererseits in Form eines Excel-Files vor. In einem ersten Schritt musste deshalb eine

einheitliche Datenbasis ausgewählt werden, wofür sich auf den ersten Blick die Access-Datenbank anbot. Die Vorteile einer derartigen Lösung, wie z.B. einer direkten Anbindung an ArcGIS™, sind unbestritten. Bei der Verwendung der Datenbank traten allerdings Zugriffsprobleme auf, deren Ursachen vor allem in der komplexen Struktur der Datenbank lagen. Um diese Probleme zu vermeiden wurde von der Wahl der Access-Datenbank Abstand genommen und die Aufbereitung der Datensätze mittels Excel durchgeführt.

Das Vorkommen von statistischen Ausreißern kann die Ergebnisse einer räumlichen Interpolation maßgeblich beeinflussen. Um diesen Einfluss zu vermeiden, wurde die Datengesamtheit auf derartige Werte überprüft und als Ausreißer erkannte Punkte aus dem Datenmaterial entfernt. Im Zuge der Aufbereitung des Datenmaterials wurden außerdem zahlreiche BIN-Punkte mit gleicher Bezeichnung festgestellt, welche ebenfalls aus dem Datensatz entfernt wurden.

In der Regel werden auf jedem BIN-Punkt zwei Bäume beprobt, d.h. pro Erhebungsjahr liegen für den ersten und zweiten Nadeljahrgang jeweils zwei Schwefelwerte vor. Da eine Interpolation mit redundanten Daten nicht möglich ist, wurden die Schwefelwerte der beiden Bäume summiert und arithmetisch gemittelt. In die Interpolation der vorliegenden Arbeit gehen somit nur die gemittelten Schwefelwerte der jeweiligen Nadeljahrgänge ein. Die räumliche Verteilung der BIN-Punkte lag nur in Form eines „Layers“ vor, wobei die zugehörige Attributtabelle keine Informationen über die xy-Koordinaten enthielt. Um von den einzelnen BIN-Punkten dennoch Koordinaten zu erhalten, wurden diese mit Hilfe eines vb-Scripts aus diesem Layer bestimmt. Den Abschluss der Datenaufbereitung bildete die Zusammenstellung der Schwefeldatensätze für die Erhebungsjahre. Anhand dieser Datensätze wurden in der Folge mittels ArcGIS™ 8.2 Layer erzeugt und einer Geodatabase zugeordnet.

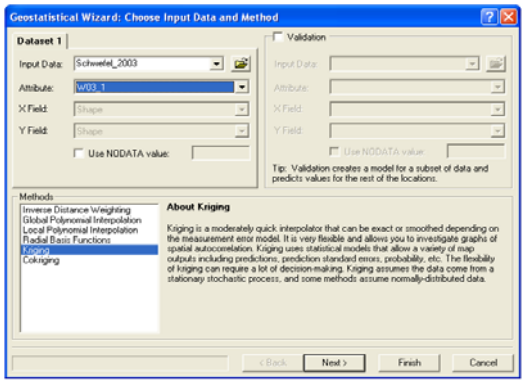
b) Explorative Datenanalyse - Trendanalyse

Vor der eigentlichen Modellanpassung bietet der Geostatistical Analyst die Möglichkeit eine explorative Datenanalyse sowie Trendanalyse durchzuführen. Die explorative Datenanalyse ist, wie bereits oben erwähnt wurde, ein Teilgebiet der Statistik, wobei die Suche nach Strukturen und auffälligen Besonderheiten im Vordergrund steht. Die Instrumente der explorativen Datenanalyse sind zum einen Werkzeuge der beschreibenden univariaten Statistik (z.B. Streudiagramme, Histogramme, Normal QQ-Plot), darüber hinaus aber auch statistische und dynamische Verfahren der multivariaten Analyse und interaktiver graphische Prozeduren (z.B. Trendanalyse, Voronoi Karten). Mit Hilfe der Trendanalyse kann zudem festgestellt werden, ob eine systematische Zu- oder Abnahme der Messwerte im Raum vorliegt.

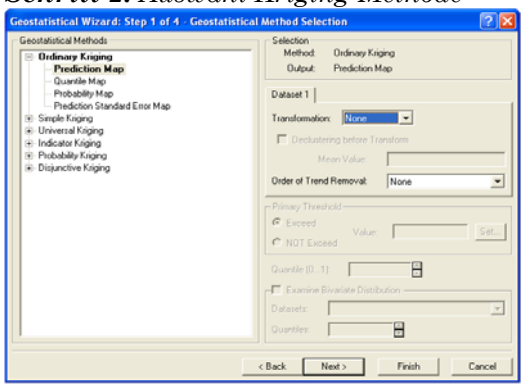
c) Modellanpassung

Die Modellanpassung der ausgewählten Interpolationsmethoden erfolgte im Anschluss an die explorative Datenanalyse bzw. Trendanalyse mit Hilfe des im Geostatistical Analyst vorhandenen „**Wizard**“. Dieser führt den Anwender Schritt für Schritt durch die Interpolation, wobei je nach gewählter Methode die Anzahl der notwendigen Arbeitsschritte unterschiedlich hoch sein kann. Abbildung 3-9 zeigt beispielhaft die Arbeitsabfolge bei der Durchführung eines Schätzvorgangs mittels Ordinary Kriging, wofür maximal sechs Arbeitsschritte notwendig sind:

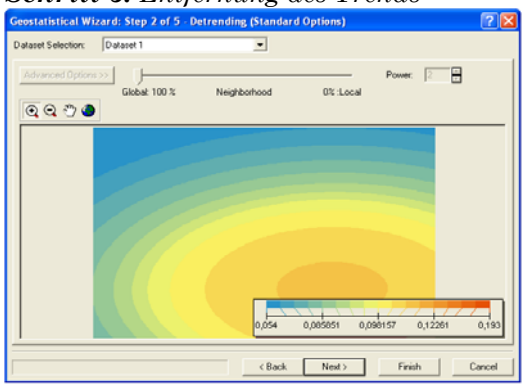
Schritt 1: *Daten- und Methodenauswahl*



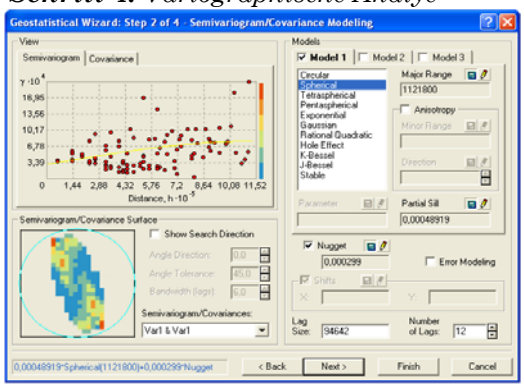
Schritt 2: *Auswahl Kriging-Methode*



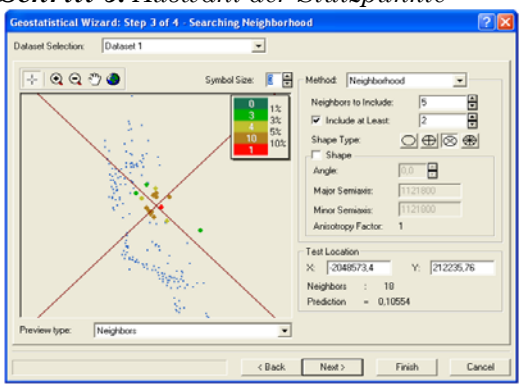
Schritt 3: *Entfernung des Trends*



Schritt 4: *Variographische Analye*



Schritt 5: *Auswahl der Stützpunkte*



Schritt 6: *Kreuzvalidierung*

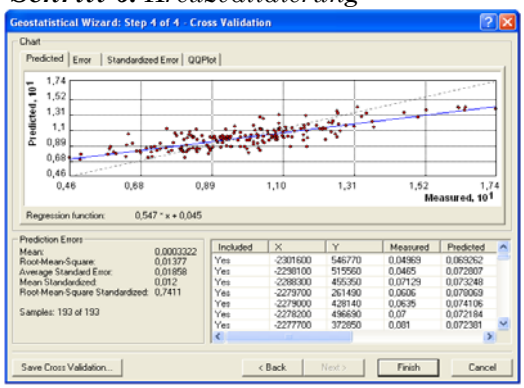


Abbildung 3-9: *Arbeitsschritte für den Schätzvorgang*

In den **Schritten 1** und **2** erfolgt zunächst die Auswahl des entsprechenden Datensatzes und der gewünschten Interpolationsmethode. Ist ein Trend im ausgewählten Datensatz vorhanden, so kann dieser bei Bedarf mit Hilfe von **Schritt 3** entfernt werden. Die Anpassung des ausgewählten Modells an das empirische Semivariogram erfolgt interaktiv in **Schritt 4**. Interaktiv bedeutet, dass der „Wizard" die Möglichkeit bietet Modellparameter zu ändern und deren Auswirkungen in einem eigenen Fenster, welches das empirische Semivariogramm bzw. die Kovarianzfunktion und gleichzeitig das Modell anzeigt, zu beobachten. Für die Modellanpassung können bis zu drei Variogramm-Modelle miteinander verschachtelt werden, wofür insgesamt elf verschiedene Funktionen (Modelle) zur Verfügung stehen. **Schritt 5** dient zur Festlegung der Suchumgebung anhand derer die Stützpunkte für die Interpolation ausgewählt werden. Dieser Vorgang erlaubt wiederum eine Interaktion mit dem „Wizard". Zu diesem Zweck wird die räumliche Verteilung der Stützpunkte in einem Anzeigefenster dargestellt. Durch Änderung der Parameter sowie Klicken an einen Ort wird die Suchumgebung visualisiert und die gefundenen Stützpunkte markiert. Zum Abschluss der Interpolation wird in **Schritt 6** die Kreuzvalidierung durchgeführt, deren Ergebnisse graphisch und tabellarisch dargestellt werden. Entsprechen die Ergebnisse der Kreuzvalidierung nicht den Erwartungen, so kann man im Wizard einige Schritte zurückgehen und entsprechenden Änderungen durchführen. Verläuft die Kreuzvalidierung in diesem iterativen Vorgang hingegen zufrieden stellend so wird nach dem Beenden des Wizard eine Zusammenfassung der Parameter angezeigt und ein Layer mit den Interpolationsergebnissen erstellt.

d) Ermittlung der Qualität von den Interpolationsverfahren

Die Validierung der Interpolationsergebnisse erfolgte wie bereits erwähnt wurde mit Hilfe der Kreuzvalidierung. Um die Ergebnisse der Kreuzvalidierung anhand von statistischen Kennzahlen und Diagrammen beurteilen zu können, wurden diese in einer dBASE-Datei abgespeichert und in die Statistiksoftware **ORIGIN 7.0** importiert. Anschließend erfolgte die Berechnung der statistischen Kennzahlen und die Erstellung der entsprechenden Diagramme in ORIGIN

e) Vergleich der Interpolationsverfahren

Um Festzustellen welches der ausgewählten Interpolationsverfahren sich am besten zur Schätzung der räumlichen Verteilung der Schwefelbelastung in der Steiermark eignet, wurde ein Vergleich der Qualität dieser Verfahren durchgeführt. Der Vergleich erfolgte dabei anhand der statistischen Kennzahlen zur Beurteilung der Qualität von Interpolationsverfahren.

f) Schwefelbelastungskarte

Die in der Verordnung gegen forstschädliche Luftverunreinigungen (BUNDESGESETZBLATT 1984) festgelegten Grenzwerte von 0,11% Schwefelgehalt im Nadeljahrgang 1 und von 0,14% Schwefelgehalt im Nadeljahrgang 2 ermöglichen erst den Nachweis von Immissionsbelastungen in den Wäldern.

Basierend auf diesen Grenzwerten erfolgte, in Anlehnung an die Vorgehensweise der Fachabteilung für das Forstwesen des Landes Steiermark, eine Klassifizierung der Schwefelgehalte für den ersten du zweiten Nadeljahrgang (vgl. Tabelle 3-1). Diese Klasseneinteilung bildete die Grundlage bei der Erstellung der Schwefelbelastungskarten für die Steiermark, anhand derer die Schwerpunktgebiete der Immissionsbelastung festgestellt werden können. Für die Ausweisung der jeweiligen Schwefelbelastungszonen wurden nur die Grenzwerte des ersten Nadeljahrgangs herangezogen.

Tabelle 3-1: *Klassifizierung der Schwefelgehalte von Nadelproben*

	% Schwefel im Nadeljahrgang	
Klasse	**1**	**2**
unbelastet	< 0,081	<0,111
Grenzbereich	0,081 – 0,110	0,111 – 0,140
leicht belastet	0,111 – 0,140	0,141 – 0,170
belastet	0,141 – 0,170	0,171 – 0,200
stark belastet	>0,170	>0,200

4 Ergebnisse

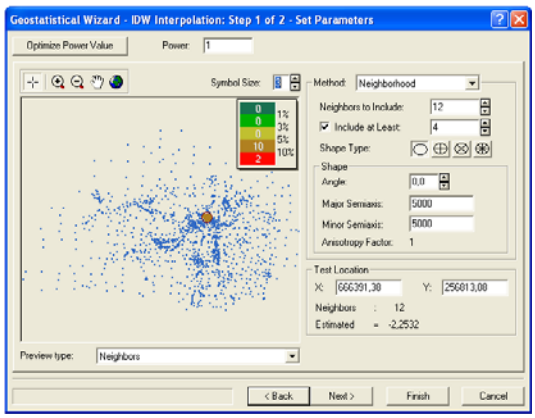
Im vorangegangenen Kapitel wurden die methodischen Grundlagen jener Interpolationsverfahren vorgestellt, die zur Regionalisierung der gemessenen Schwefeldaten herangezogen werden. Dieses Kapitel beschäftigt sich nun mit den Ergebnissen der räumlichen Interpolation anhand dieser Verfahren. An dieser Stelle ist anzumerken, dass die im Folgenden vorgestellten Ergebnisse nur anhand von Daten des Erhebungsjahres 2003 ermittelt wurden.

4.1 Schätzergebnis Inverse Distanzgewichtung

a) Durchführung

Die IDW-Interpolation im Geostatistical Analyst ist ein iterativer Vorgang der in zwei Arbeitsschritten abläuft (vgl. Abbildung 4-1). In Schritt 1 erfolgen zunächst die Einstellungen für die Interpolation, wie z.B. Anzahl der Stützpunkte, anhand derer die Schätzung durchgeführt wird. Die Ergebnisse dieser Schätzung werden anschließen in Schritt 2 validiert. Die verwendeten Schwefeldaten lagen nicht in Normalverteilung vor, weshalb vor dem Beginn der Interpolation eine lognormale Transformation durchgeführt werde musste. Im Laufe des iterativen Vorgangs zeigte sich, dass mit einem Distanzexponenten von 1 und einer minimalen bzw. maximalen Stützpunkteanzahl von 4 bzw. 12 das beste Schätzergebnis zu erreichen ist. Die Auswahl der Stützpunkte erfolgte dabei anhand eines Suchkreises mit einem Radius von 5000m.

Schritt 1: *IDW-Parameter*



Schritt 2: *Kreuzvalidierung*

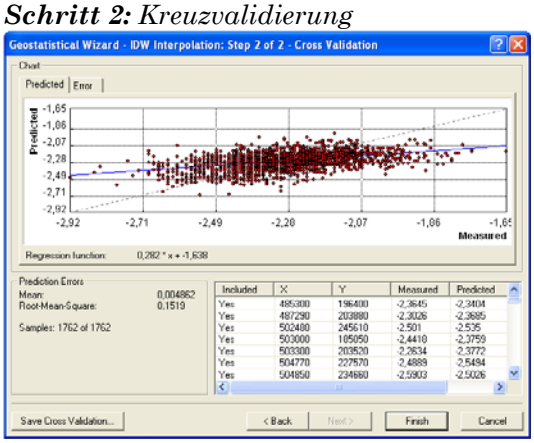


Abbildung 4-1*: Arbeitsschritte für die Schätzung mittels IDW-Verfahren*

b) Statistische Kennzahlen der Kreuzvalidierung

In der Tabelle 4-1 sind die Ergebnisse der statistischen Auswertung der Kreuzvalidierung für das IDW-Verfahren dargestellt. Die erhobenen statistischen Kennzahlen wurden bereits in Kapitel 3.3 „*Qualitätsermittlung von Interpolationsverfahren*" erläutert, weshalb an dieser Stelle nicht mehr näher auf diese eingegangen wird.

Tabelle 4-1: *Ergebnis der Kreuzvalidierung für das IDW-Verfahren*

Statistische Kennzahlen	IDW
Stichprobenanzahl n	1762
Arithmetisches Mittel $\bar{x}$	0,00037
Median x_m	-0,00070
Standardabweichung s	0,01599
Schiefekoeffizient g	0,57399
MAE	0,01237
MSE	0,00026
Rangkorrelationskoeff. r_s	0,458**

** Die Korrelation ist auf dem Niveau von 0,01 signifikant (zweiseitig)

Das arithmetische Mittel der Residuen ist beim IDW-Verfahren größer als Null, d.h. für die Schwefelwerte liegt im Mittel eine systematische Unterschätzung vor (vgl. Tabelle 4-1). Ein sehr einfaches Maß um die Verzerrung einer geschätzten Werteoberfläche festzustellen ist die Differenz aus Median und arithmetischem Mittel. Betrachtet man die Werte dieser beiden Kennzahlen, so lässt sich eine leicht asymmetrische Verteilung der Residuen feststellen. Da der Wert für das arithmetische Mittel größer ist als jener für den Median besitzt die asymmetrische Verteilung der Residuen eine positive Schiefe. Diese Aussage wird durch den Schiefekoeffizienten bestätigt, dessen Wert ebenfalls größer als Null ist und somit auf eine positive Schiefe hinweist. Die leicht rechtsschiefe Verteilung der Residuen ist auch in Abbildung 4-2 zu erkennen. Abbildung 4-2 zeigt die Verteilung der Residuen in Form eines Histogramms, dessen Balken die absolute Häufigkeit der Residuen eines Werteintervalls von 0,01% Schwefel darstellen. Die Interpretation des Wertes für die Standardabweichung, als Maß für die Streuung der Residuenwerte um den entsprechenden Mittelwert, ist ohne Vergleich mit einem anderen Interpolationsverfahrens nur schwer möglich. Eine ähnliche Aussage kann für die beiden statistischen Kennzahlen MAE und MSE getroffen werden, die ohne Vergleichswerte ebenfalls schwer zu beurteilen sind. Um die Werte dieser Kennzahlen nicht falsch zu interpretieren wird an dieser Stelle darauf verzichtet und auf das Kapitel 4.4 „*Vergleich der Schätzergebnisse*" verwiesen.

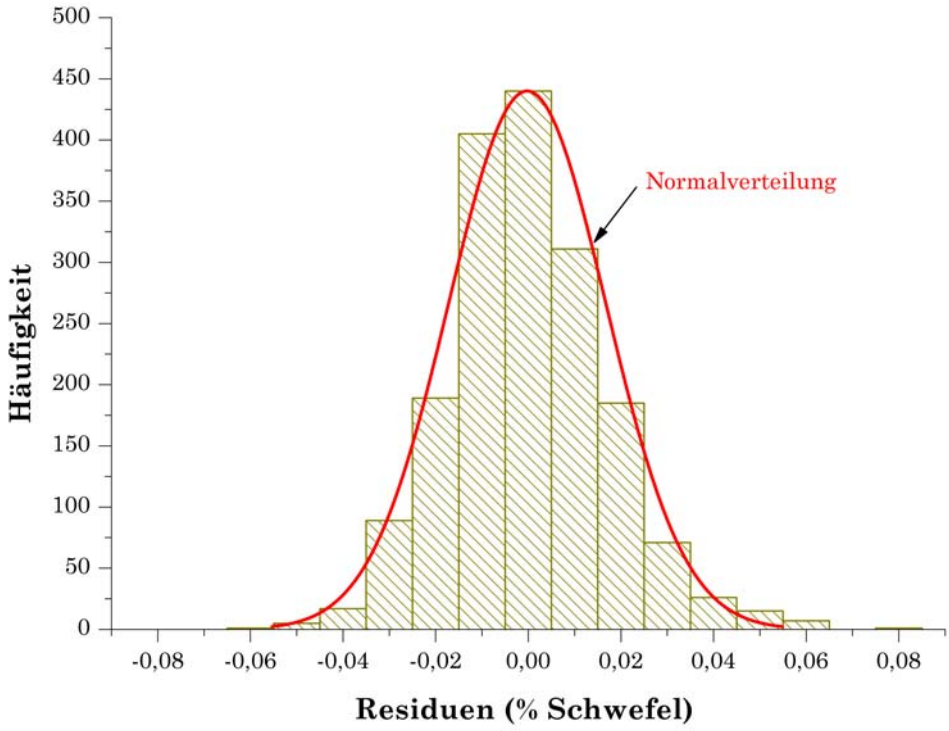


Abbildung 4-2: *Histogramm der Residuen für das IDW-Verfahren*

Abbildung 4-3 zeigt das Korrelationsdiagramm der wahren und geschätzten Schwefelwerte für das IDW-Verfahren.

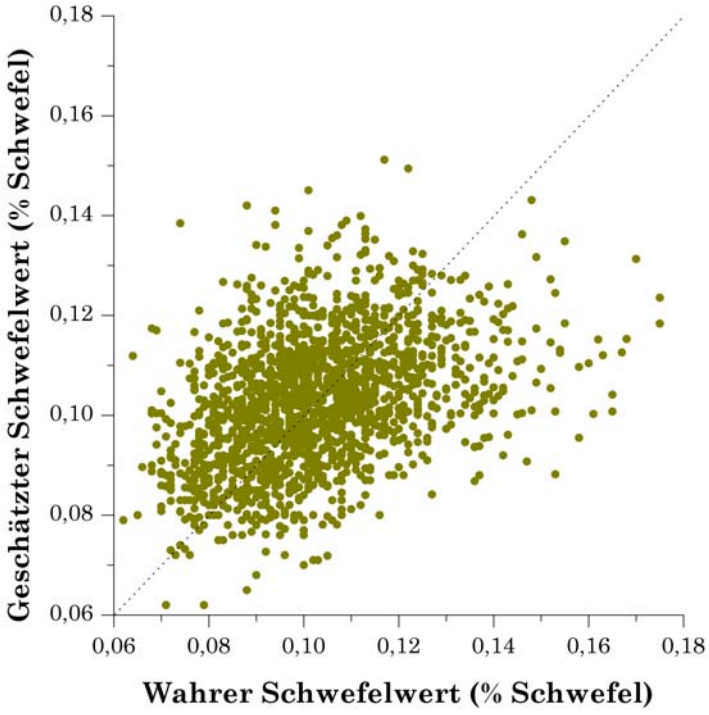


Abbildung 4-3: *Korrelationsdiagramm für das IDW Verfahren*

Im Korrelationsdiagramm (vgl. Abbildung 4-3) ist eine mittelstarke positive Korrelation zwischen den Schwefelwerten zu erkennen, die durch den Rangkorrelationskoeffizient aus Tabelle 4-1 bestätigt wird. Die Korrelation ist auf dem Niveau von 0,01 signifikant. Anhand des Korrelationsdiagramms ist festzustellen, dass die hohen wahren Schwefelwerte unterschätzt und die niedrigen wahren Schwefelwerte überschätzt werden. Die Unterschätzung der hohen Schwefelwerte ist dabei wesentlich stärker ausgeprägt als die Überschätzung der niedrigen Schwefelwerte. Dieser Umstand wirkt sich deutlich auf den Wertebereich der geschätzten Werte aus, welcher im Vergleich zu den wahren Schwefelwerten eingeschränkt ist. Deutlicher zum Ausdruck kommen diese Über- und Unterschätzungen der Schwefelwerte in Abbildung 4-4, welche die Streuung der Residuen bezogen auf die wahren Schwefelwerte zeigt.

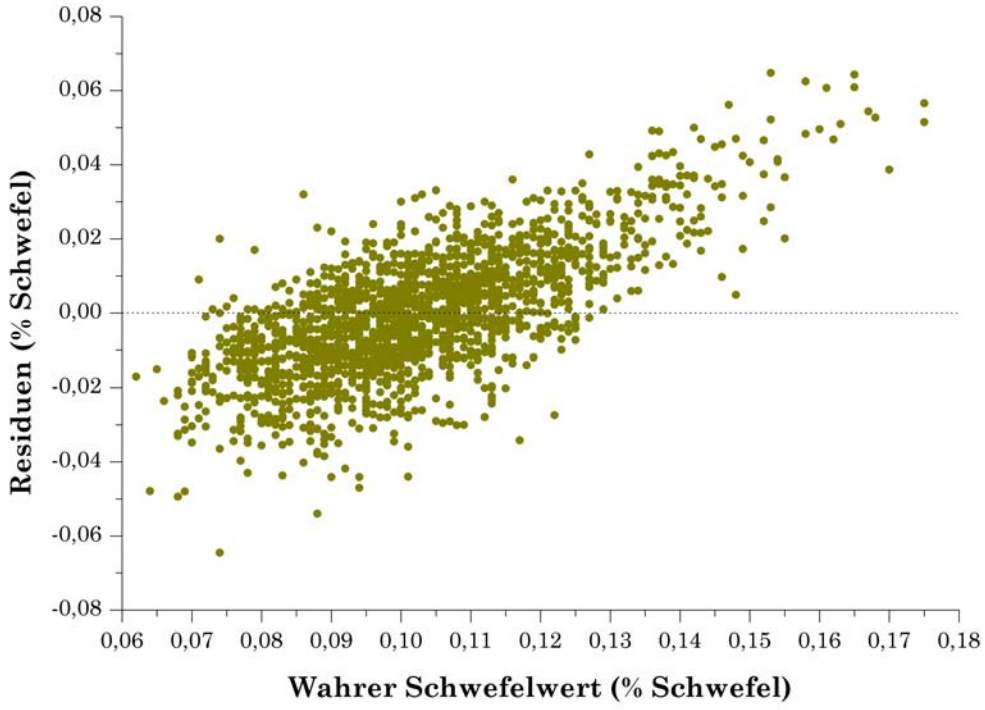


Abbildung 4-4: *Streuungsdiagramm der Residuen für das IDW-Verfahren*

d) Schwefelbelastungskarte

In Abbildung 4-5 sind die geschätzten Schwefelwerte ohne Berücksichtigung der tatsächlichen Waldausbreitung in Form einer Karte dargestellt. Die Klassifikation der Schwefelwerte erfolgte anhand der Quantile-Methode, d.h. in jeder Schwefelklasse ist die gleiche Anzahl an Werten vorhanden. Die Schwefelbelastung zeigt regional deutliche Unterschiede, wobei die höchsten Werte in der Mur-Mürz Furche, dem Grazer Becken, der Osteiermark, dem Ennstal und im Gebiet Eisenerz-Hieflau vorhanden sind (vgl. Abbildung 4-5). Zudem ist teilweise der für das IDW-Verfahren typische „*bulls-eye-Effekt*“ zu erkennen.

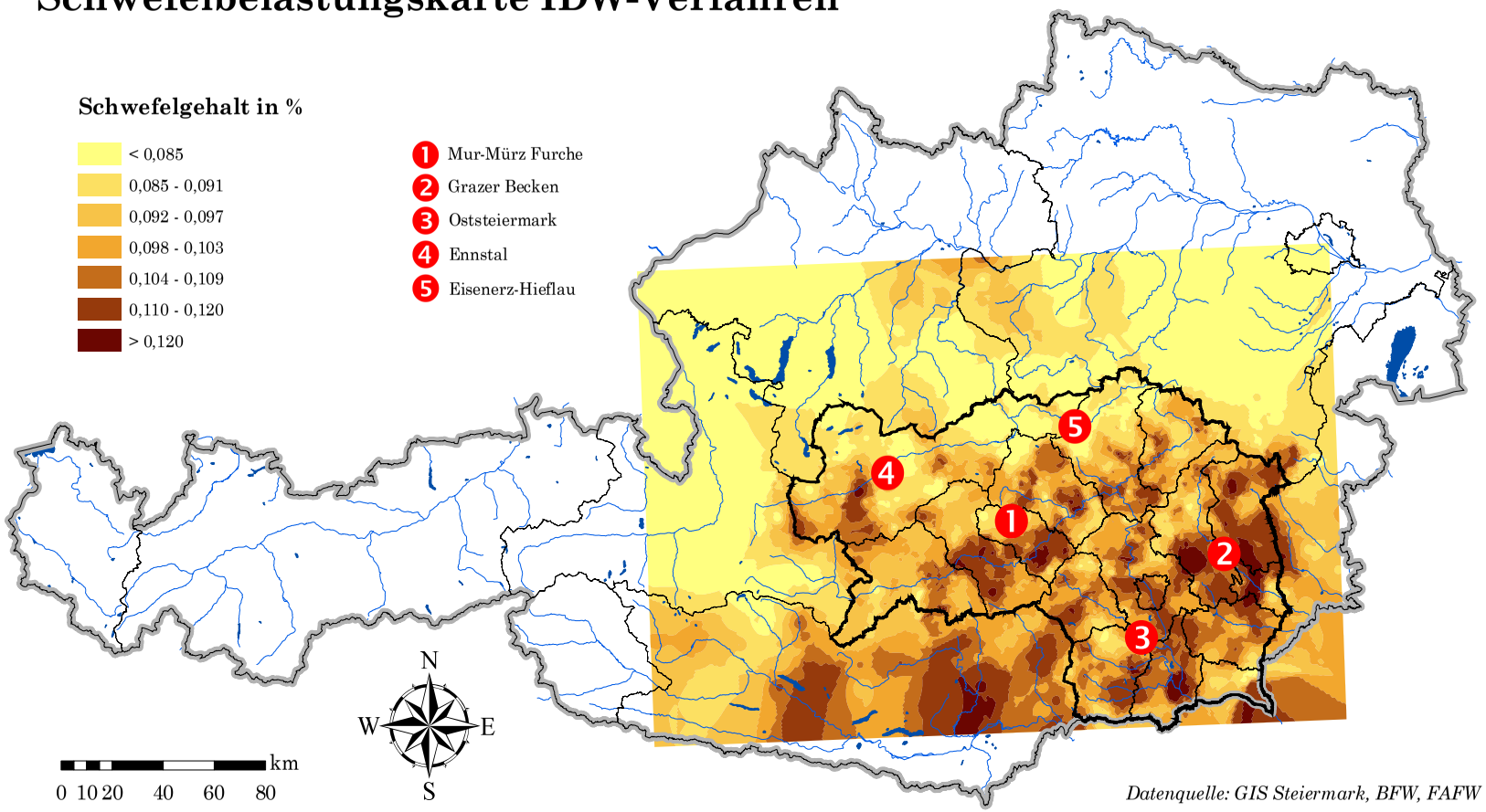


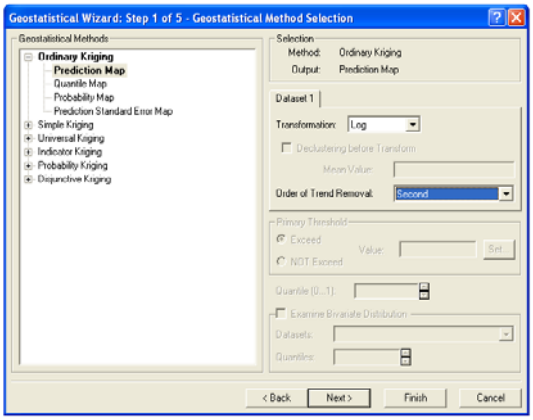
Abbildung 4-5: *Schwefelbelastungskarte mittels IDW-Verfahren*

4.2 Schätzergebnis Ordinary Kriging

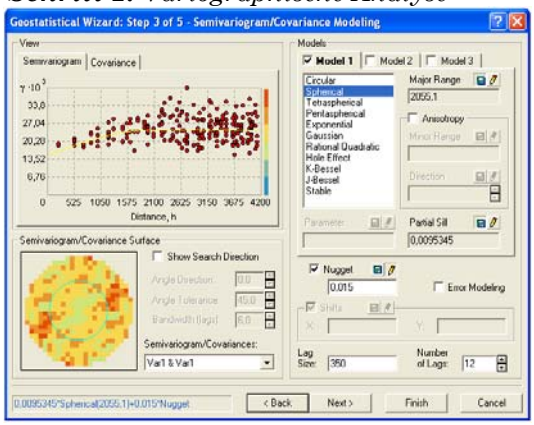
a) Durchführung

Für eine erfolgreiche Interpolation einer Werteoberfläche mittels Ordinary Kriging sind mehrere Arbeitsschritte notwendig (vgl. Abbildung 4-6), welche in Kapitel 3.4 „*Vorgehensweise und verwendete Software*“ bereits vorgestellt wurden. Auf die Nichtnormalverteilung der verwendeten Schwefeldaten wurde bereits hingewiesen. Im Rahmen der Trendanalyse konnte weiters ein „umgekehrt u-förmiger“ Trend in Richtung Nordwest-Südost festgestellt werden. Da beide Ursachen die variographische Analyse erschweren können, wurde eine lognormale Transformation der Schwefeldaten durchgeführt und der Trend mit Hilfe eines Polynoms zweiter Ordnung entfernt.

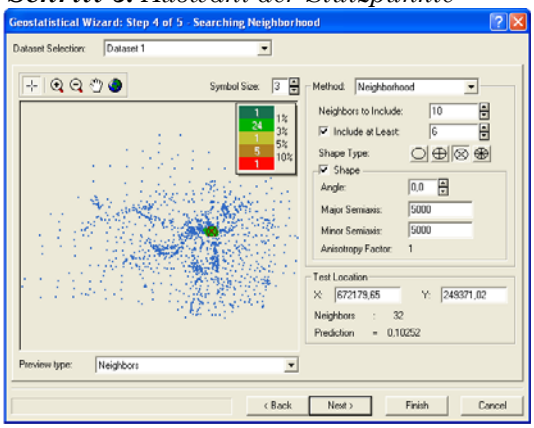
Schritt 1: *Modellauswahl*



Schritt 2: *Variographische Analyse*



Schritt 3: *Auswahl der Stützpunkte*



Schritt 4: *Kreuzvalidierung*

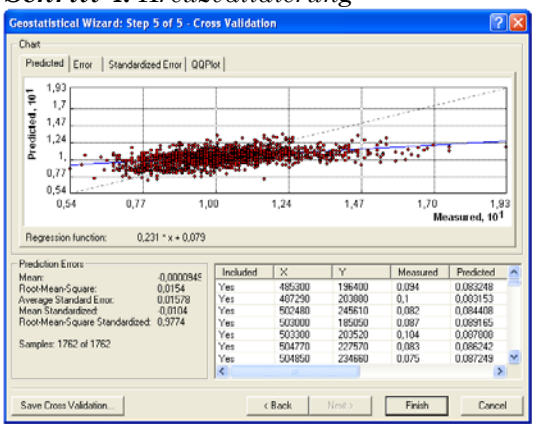


Abbildung 4-6: *Arbeitsschritte für die Schätzung mittels Ordinary Kriging*

Die variographische Analyse ist ein iterativer Vorgang, welcher vereinfacht als Wechselspiel zwischen empirischen und theoretischen Semivariogramm angesehen werden kann. Im Zuge der Analyse hat sich gezeigt, dass bei einer lag-Distanz von

350m die besten Schätzergebnisse für die Schwefelwerte zu erwarten sind. Die Anpassung an das empirische Semivariogramm erfolgte durch ein theoretisches Semivariogramm, das sich aus einem sphärischen Modell und einem Nugget-Modell zusammensetzt. Die Gleichung 4-1 zeigt das theoretische Semivariogramm, welches durch die drei Parameter Partial Sill, Range und Nugget beschrieben wird.

$$\underset{\text{Partial Sill}}{0{,}0095345} * Spherical(\underset{\text{Range}}{2055{,}1}) + \underset{\text{Nugget}}{0{,}015} * Nugget \qquad \textbf{(4-1)}$$

Die Auswahl, der in die Schätzung eingehenden Stützpunkte, erfolgte durch einen Suchkreis (4 Sektoren) mit einem Radius von 5000m. Die minimale bzw. maximale Stützpunkteanzahl wurde mit 6 bzw. 10 festgelegt. Eine vorhandene Anisotropie ging in die Interpolation nicht ein, da deren Berücksichtigung zu keiner Verbesserung des Schätzergebnisses geführt hat.

b) Statistische Kennzahlen der Kreuzvalidierung

Die Tabelle 4-2 gibt eine Übersicht über die Ergebnisse der statistischen Auswertung der Kreuzvalidierung für das Ordinary Kriging Verfahren.

Tabelle 4-2: *Ergebnis der Kreuzvalidierung für das Ordinary Kriging*

Statistische Kennzahlen	Ordinary Kriging
Stichprobenanzahl n	1762
Arithmetisches Mittel $\bar{x}$	0,00009
Median x_m	-0,00141
Standardabweichung s	0,01541
Schiefekoeffizient g	0,71278
MAE	0,01179
MSE	0,00024
Rangkorrelationskoeff. r_s	0,493**

** Die Korrelation ist auf dem Niveau von 0,01 signifikant (zweiseitig)

Das arithmetische Mittel der Residuen ist bei der Schätzung mittels Ordinary Kriging größer als Null, d.h. für die geschätzten Schwefelwerte liegt im Mittel eine systematische Unterschätzung vor. Betrachtet man das arithmetische Mittel und den Median, so lässt sich aus diesen beiden Kennzahlen eine rechtsschiefe Verteilung der Residuen ableiten. Der Wert für den Schiefekoeffizienten deutet ebenfalls auf eine positive Schiefe der Residuenverteilung hin, welche in Abbildung 4-7 zu erkennen ist. Die Abbildung 4-7 zeigt die Verteilung der Residuen in Form eines Histogramms, dessen Balken die absolute Häufigkeit eines Werteintervalls von 0,01% Schwefel darstellen.

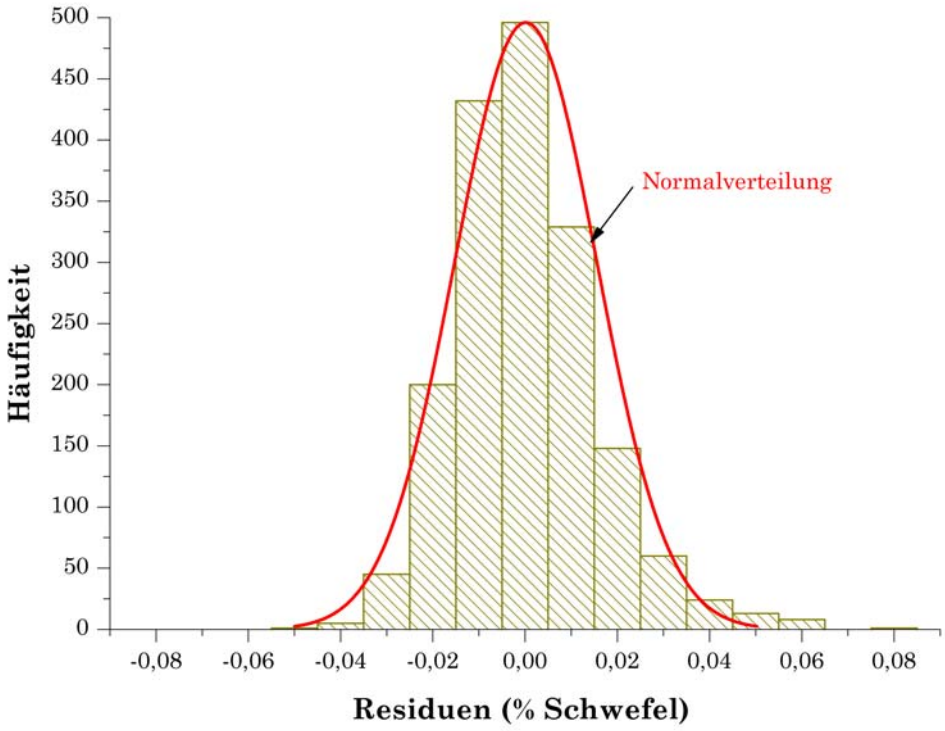


Abbildung 4-7: *Histogramm der Residuen für das Ordinary Kriging*

In Abbildung 4-8 ist das Korrelationsdiagramm der wahren und geschätzten Schwefelwerte für das Ordinary Kriging dargestellt.

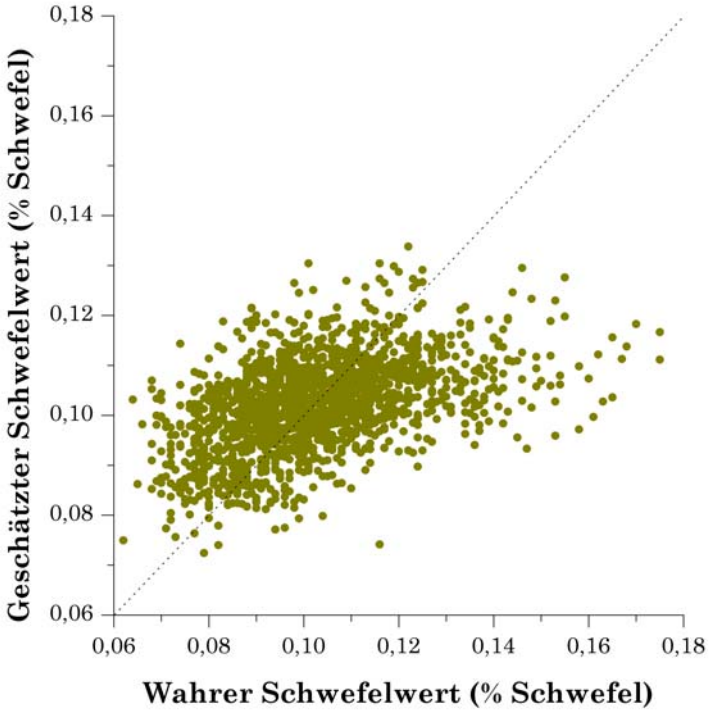


Abbildung 4-8: *Korrelationsdiagramm für das Ordinary Kriging*

Im Korrelationsdiagramm ist zwischen den Schwefelwerten eine mittelstarke positive Korrelation feststellbar, die durch den Rangkorrelationskoeffizient aus Tabelle 4-2 bestätigt wird. Die Korrelation ist auf dem Niveau von 0,01 signifikant. Das Korrelationsdiagramm zeigt weiters eine Unterschätzung der hohen wahren Schwefelwerte und eine Überschätzung der niedrigen wahren Schwefelwerte. Das Ausmaß der Unterschätzung ist dabei wesentlich stärker ausgeprägt als jenes der Überschätzung. Dieser Umstand spiegelt sich auch im Wertebereich der geschätzten Schwefelwerte wieder, welcher im Vergleich zu den wahren Schwefelwerten eingeschränkt ist. Etwas deutlicher zum Ausdruck kommen die Über- und Unterschätzungen der geschätzten Schwefelwerte in Abbildung 4-9, welche die Streuung der Residuen bezogen auf die wahren Schwefelwerte zeigt.

Abbildung 4-9: *Streuungsdiagramm der Residuen für das Ordinary Kriging*

c) Schwefelbelastungskarte

Die geschätzten Schwefelwerte sind in Abbildung 4-10 anhand einer Schwefelbelastungskarte dargestellt, welche die tatsächliche Waldausbreitung nicht berücksichtigt. Die Klassifikation der Schwefelwerte erfolgte anhand der Quantile-Methode. Die höchsten Werte für die Schwefelbelastung treten in der Mur-Mürz-Furche, dem Grazer Becken und in der Oststeiermark auf (vgl. Abbildung 4-10). Im Gegensatz zum Schätzergebnis des IDW-Verfahrens (vgl. Abbildung 4-5) ist allerdings der Flächenanteil für die Klasse „*>0,120% Schwefel*“ wesentlich geringer ausgeprägt.

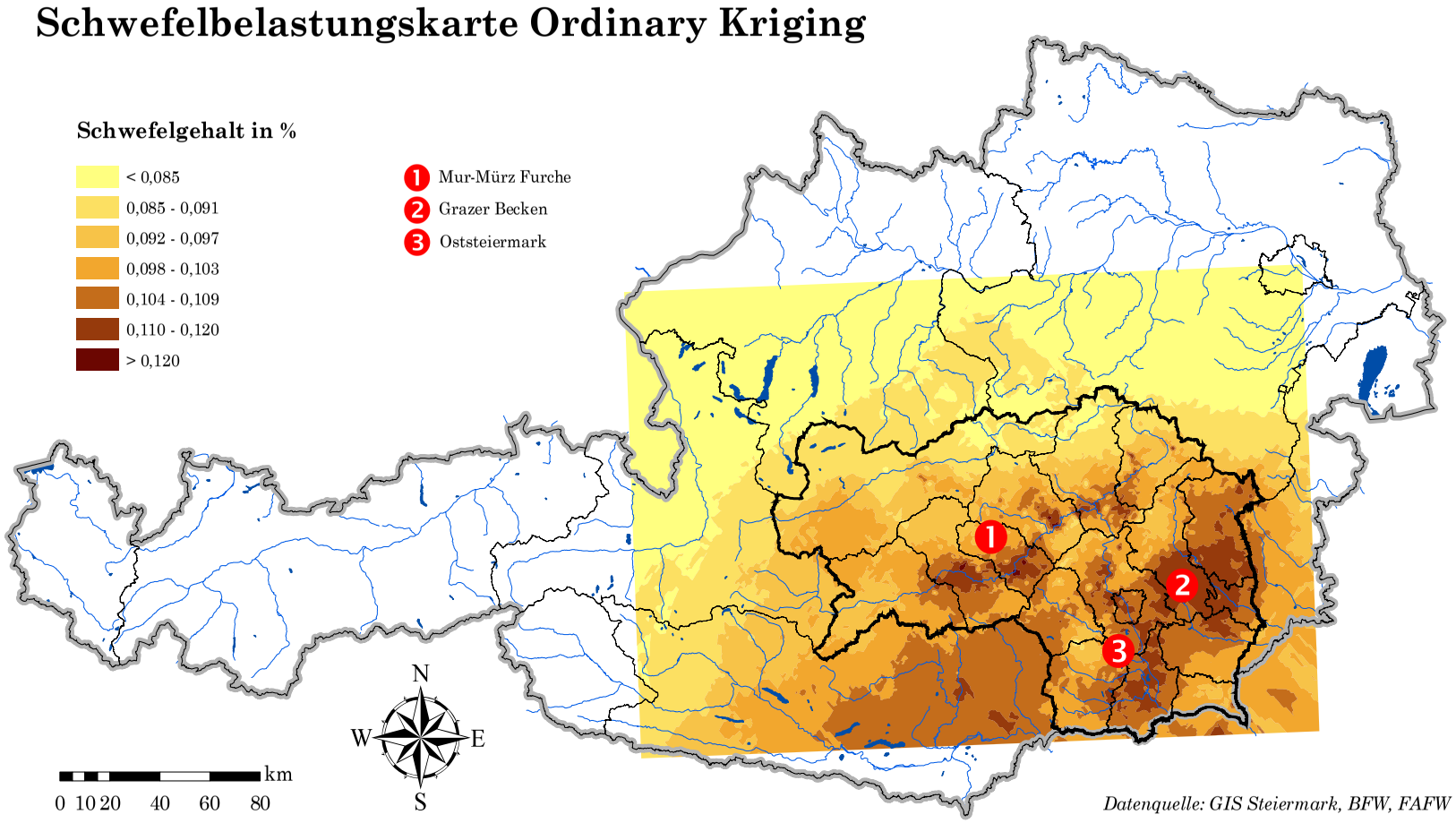


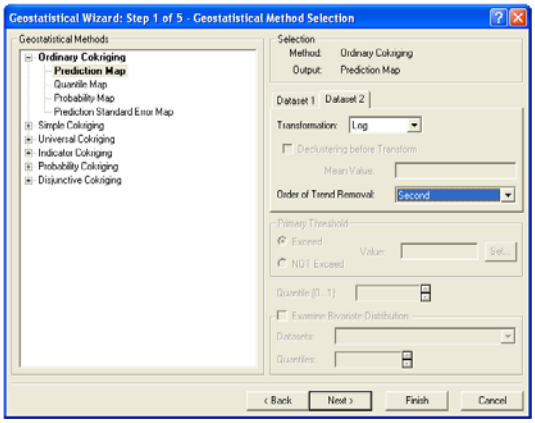
Abbildung 4-10: *Schwefelbelastungskarte mittels Ordinary Kriging*

4.3 Schätzergebnis Ordinary Cokriging

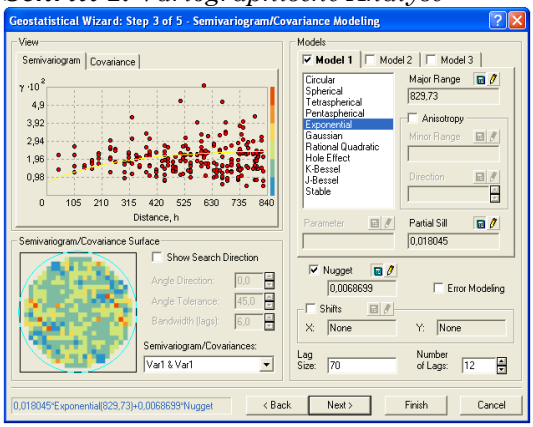
a) Durchführung

Die Durchführung der Schätzung mittels Ordinary Cokriging erfolgte anhand von mehreren Arbeitsschritten (vgl. Abbildung 4-11). Ursprünglich war angedacht, neben den Schwefeldaten des ersten Nadeljahrgangs auch Zusatzinformationen in Form von Geländehöhendaten und Schwefeldaten des zweiten Nadeljahrgangs in die Schätzung mit einzubeziehen. Im Laufe der variographischen Analyse musste allerdings festgestellt werden, dass die Berücksichtigung der Geländehöhendaten zu keiner Verbesserung des Schätzergebnisses der Schwefeldaten führt. Aus diesem Grund wurden die Geländehöhendaten für die Schätzung mittels Ordinary Cokriging nicht mehr berücksichtigt und als Zusatzinformation nur mehr die Schwefeldaten des zweiten Nadeljahrgangs verwendet.

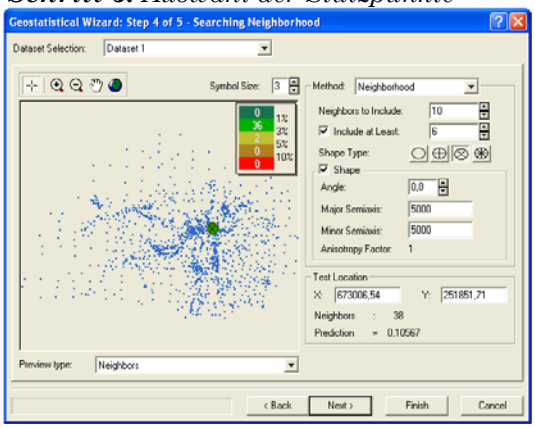
Schritt 1: *Modellauswahl*



Schritt 2: *Variographische Analyse*



Schritt 3: *Auswahl der Stützpunkte*



Schritt 4: *Kreuzvalidierung*

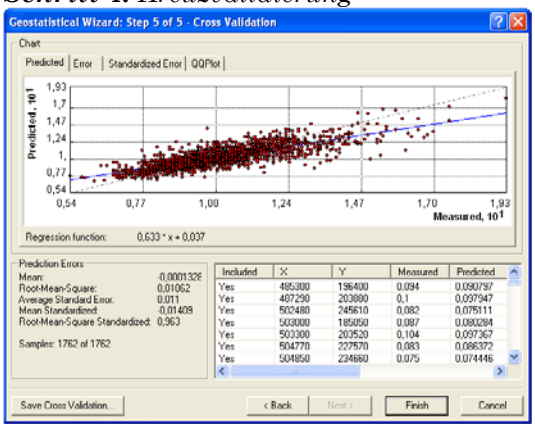


Abbildung 4-11: *Arbeitsschritte für die Schätzung mittels Ordinary Cokriging*

Wie bei den Schwefeldaten des ersten Nadeljahrgangs konnte auch bei jenen des zweiten eine Nichtnormalverteilung und ein „umgekehrt u-förmiger“ Trend in Richtung Nordwest-Südost festgestellt werden. Da, wie bereits erwähnt, die Nichtnormalverteilung und der Trend die Modellanpassung erschweren können, erfolgte sowohl für die Daten des ersten als auch zweiten Nadeljahrgangs eine lognormale Transformation und die Entfernung des Trends mit Hilfe eines Polynoms zweiter Ordnung.

Der Aufwand für die Berechnung eines Cokriging Systems ist deutlich höher als jener für ein einfaches Kriging System. Der Grund hierfür liegt darin, dass bei der variographischen Analyse je nach Anzahl an Eingangsvariabeln (Datensätzen) mehrere Variogramme notwendig sind. Betrachtet man beispielsweise ein System bestehend aus zwei Eingangsvariablen (z.B. Ozon und Stickstoff), so müssen im Rahmen der variographischen Analyse drei Variogramme erstellt werden - je eines für die beiden Variablen und ein Cross-Variogramm. Bei einem aus drei Eingangsvariablen bestehenden System sind bereits sechs Variogramme notwendig.

Im vorliegenden Fall gingen die Schwefeldaten des ersten und zweiten Nadeljahrgangs in die Schätzung mittels Ordinary Cokriging ein, d.h. bei der variographischen Analyse mussten drei Variogramme erstellt werden. Bei den Variogrammen handelte es sich einerseits um die empirischen Semivariogramme der beiden Schwefeldatensätze und andererseits um das Cross-Variogramm. Die Anpassung der theoretischen Semivariogramme an die beiden empirischen erfolgte durch die Kombination zweier Modelle, genauer gesagt, durch die Kombination eines exponentiellen Modells mit einem Nugget Modell. Im Zuge der variographischen Analyse konnte weiters festgestellt werden, dass bei einer lag-Distanz von 70m die besten Schätzergebnisse für die Schwefelwerte zu erwarten sind.

Die Gleichung 4-2 zeigt das theoretische Semivariogramm für den ersten Nadeljahrgang und die Gleichung 4-3 jenes für den zweiten Nadeljahrgang. Die Form der beiden Semivariogramme wird durch die drei Parameter Partial Sill, Range und Nugget beschrieben (vgl. Gleichung 4-2). Die Gleichung 4-4 zeigt die Funktion der Cross-Covariance.

Partial Sill *Range* *Nugget*

$$0{,}018045 * \mathit{Exponential}(829{,}73) + 0{,}0068699 * \mathit{Nugget} \qquad \textbf{(4-2)}$$

$$0{,}021826 * \mathit{Exponential}(829{,}73) + 0{,}0063267 * \mathit{Nugget} \qquad \textbf{(4-3)}$$

$$0{,}019846 * \mathit{Exponential}(829{,}73) \qquad \textbf{(4-4)}$$

Die Auswahl der in die Schätzung eingehenden Stützpunkte, erfolgte für die beiden verwendeten Schwefeldatensätze durch Suchkreise (4 Sektoren) mit einem Radius von jeweils 5000m. Die minimal notwendige Stützpunkteanzahl je Sektor wurde für den Schwefeldatensatz des ersten Nadeljahrgangs mit 6 und für den zweiten Nadeljahrgang mit 4 festgelegt. Die maximale Stützpunkteanzahl pro Sektor wurde für beide Schwefeldatensätze mit 10 angenommen. Im Gegensatz zum Ordinary Kriging konnte im Rahmen der variographischen Analyse keine Richtungsabhängigkeit der räumlichen Autokorrelation festgestellt werden.

b) Statistische Kennzahlen der Kreuzvalidierung

In der Tabelle 4-3 sind die Ergebnisse der statistischen Auswertung der Kreuzvalidierung für das Ordinary Cokriging Verfahren zusammenfassend gegenübergestellt.

Tabelle 4-3: *Ergebnis der Kreuzvalidierung für das Ordinary Cokriging*

Statistische Kennzahlen	**Ordinary Cokriging**
Stichprobenanzahl n	1762
Arithmetisches Mittel $\bar{x}$	0,00013
Median x_m	-0,00058
Standardabweichung s	0,01063
Schiefekoeffizient g	0,57863
MAE	0,00808
MSE	0,00011
Rangkorrelationskoeff. r_s	0,779**

** Die Korrelation ist auf dem Niveau von 0,01 signifikant (zweiseitig)

Das arithmetische Mittel der Residuen ist beim Ordinary Cokriging Verfahren größer als Null, d.h. für die geschätzten Schwefelwerte liegt im Mittel eine systematische Unterschätzung vor (vgl. Tabelle 4-3). Betrachtet man die Werte für das arithmetische Mittel und den Median, so lässt sich wie beim Ordinary Cokriging eine leicht asymmetrische Verteilung der Residuen ableiten, welche eine positive Schiefe aufweist. Die positive Schiefe wird weiters durch den Wert des Schiefekoeffizienten nach Pearson bestätigt (vgl. Tabelle 4-3). Diese leicht rechtsschiefe Verteilung der Residuen ist auch in Abbildung 4-12 zu erkennen. Die Abbildung 4-12 zeigt die univariate Verteilung der Residuen in Form eines Histogramms, dessen Balken die Häufigkeit der Residuen eines Werteintervalls von 0,005% Schwefel darstellen.

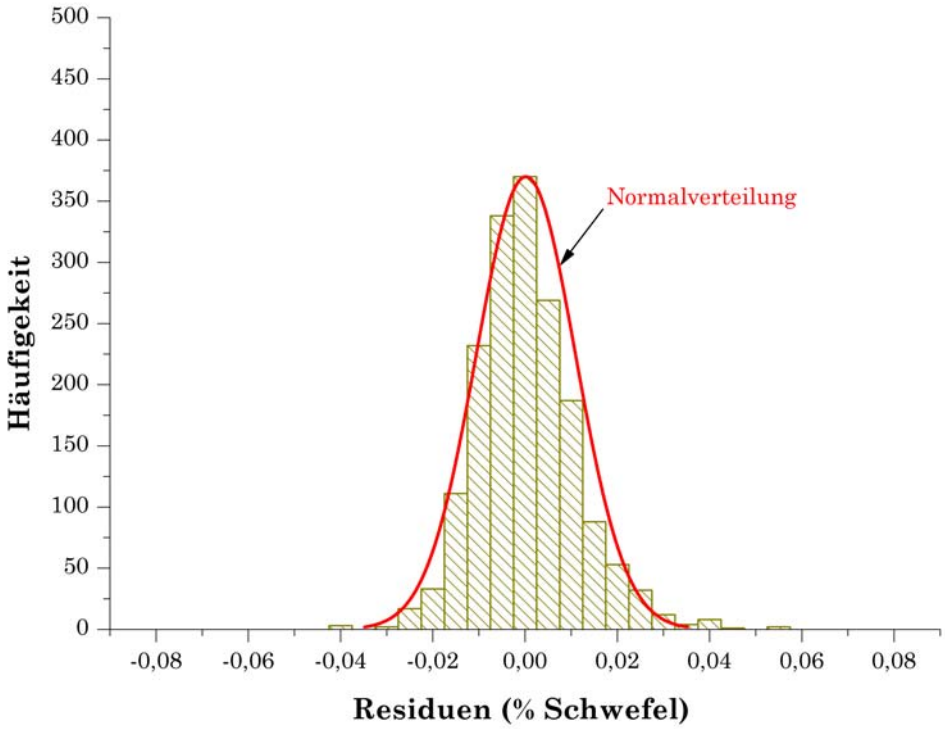


Abbildung 4-12: *Histogramm der Residuen für das Ordinary Cokriging*

Abbildung 4-13 zeigt das Korrelationsdiagramm der wahren und geschätzten Schwefelwerte für das Ordinary Cokriging.

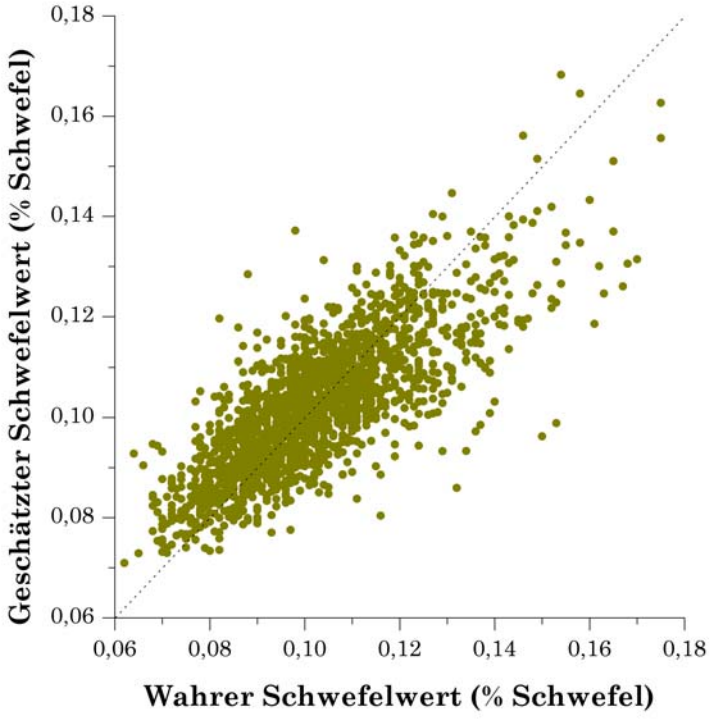


Abbildung 4-13: *Korrelationsdiagramm für das Ordinary Cokriging*

Im Korrelationsdiagramm (vgl. Abbildung 4-13) ist zwischen den wahren und geschätzten Schwefelwerten eine starke positive Korrelation zu erkennen, die durch den Rangkorrelationskoeffizient aus Tabelle 4-3 bestätigt wird. Die Korrelation ist auf dem Niveau von 0,01 signifikant. Das Korrelationsdiagramm zeigt weiters eine geringe Unterschätzung der hohen wahren Schwefelwerte und eine Überschätzung der niedrigen wahren Schwefelwerte. Im Gegensatz zu den bisher vorgestellten Interpolationsergebnissen sind die Wertebereiche der wahren und geschätzten Schwefelwerte fast identisch. Etwas deutlicher zum Ausdruck kommen diese Über- und Unterschätzungen in Abbildung 4-14, welche die Streuung der Residuen bezogen auf die wahren Schwefelwerte zeigt.

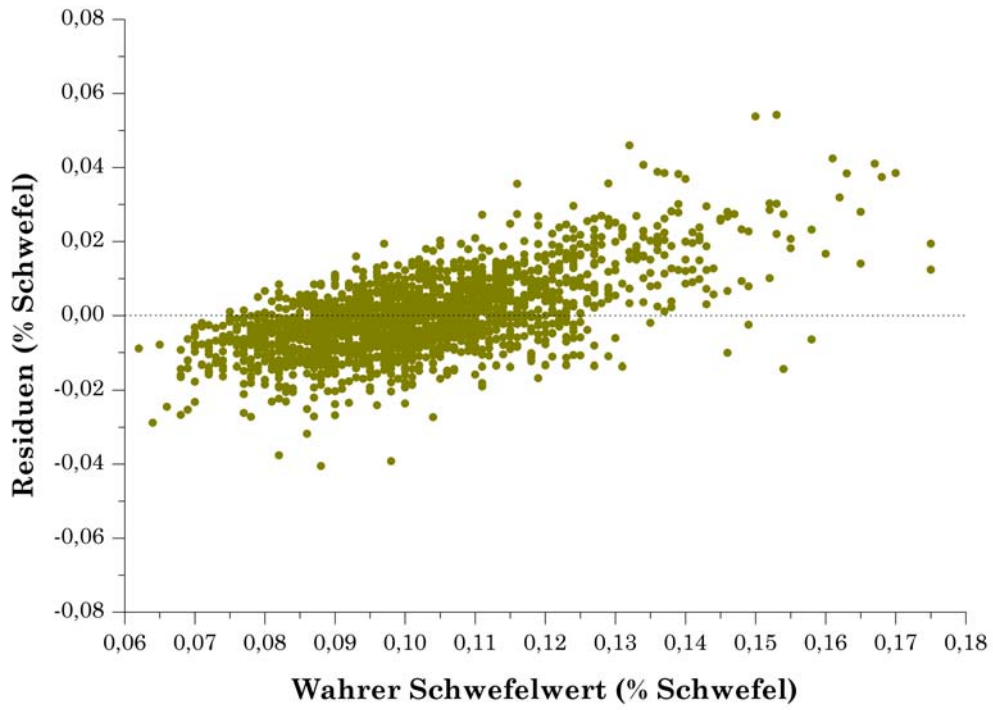


Abbildung 4-14: *Streuungsdiagramm der Residuen für das Ordinary Cokriging*

c) Schwefelbelastungskarte

Die Abbildung 4-15 zeigt das Ergebnis der Werteoberflächenschätzung für das Ordinary Cokriging Verfahren in Form einer Schwefelbelastungskarte, wobei die tatsächliche Waldausbreitung nicht berücksichtigt wurde. Die Schwefelwerte wurden dabei mittels der Quantile-Methode klassifiziert. Die höchsten Werte für die Schwefelbelastung treten wie bereits oben festgestellt wurde in der Mur-Mürz Furche, dem Grazer Becken und in der Oststeiermark auf (vgl. Abbildung 4-15). Anzumerken ist, dass die Schwefelbelastungskarten aus dem Ordinary Kriging und dem Ordinary Cokriging nur geringfügige Unterschiede hinsichtlich der räumlichen Verteilung der Schwefelwerte aufweisen.

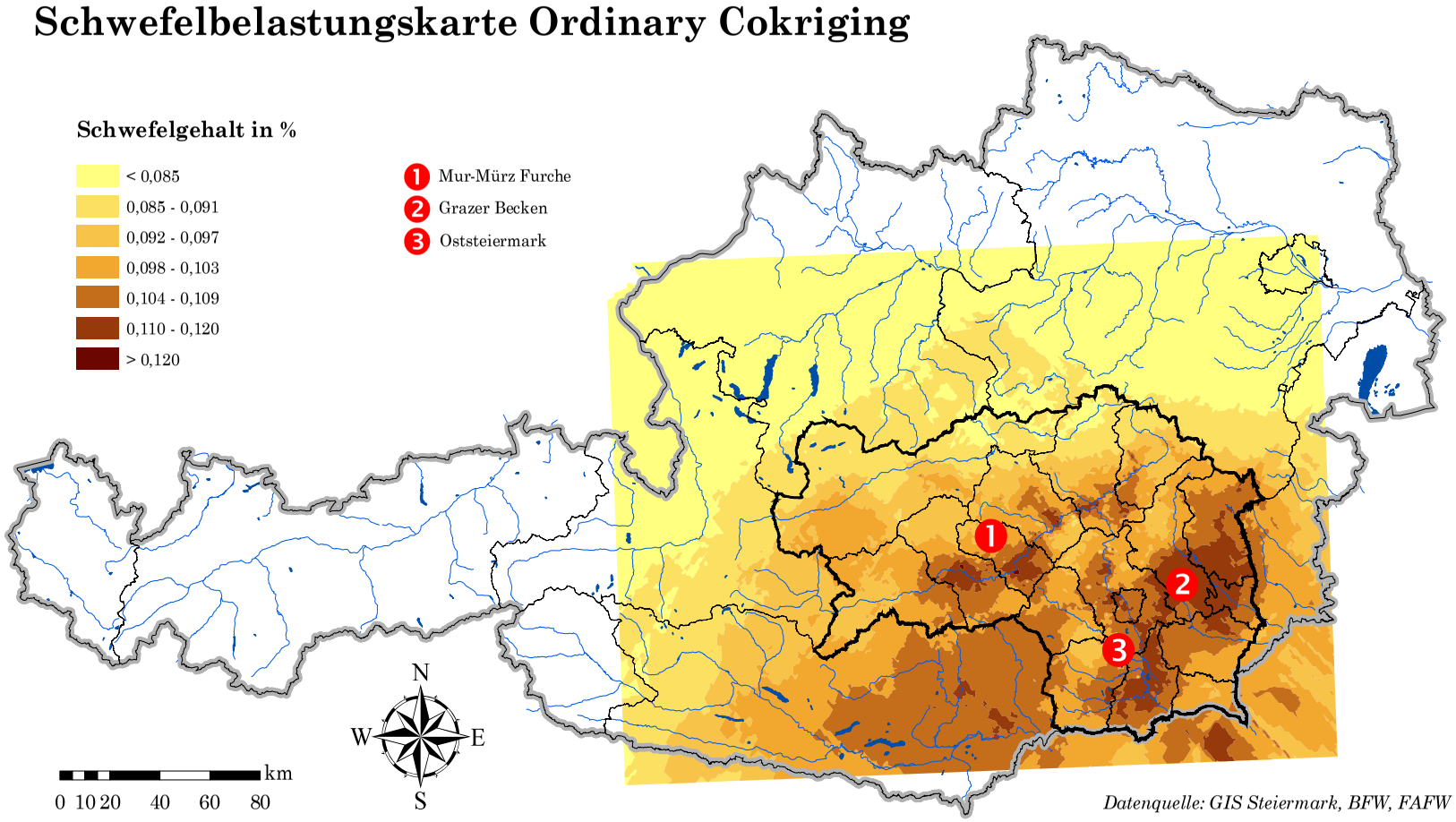


Abbildung 4-15: *Schwefelbelastungskarte mittels Ordinary Cokriging*

4.4 Vergleich der Schätzergebnisse

In der Tabelle 4-4 sind die Ergebnisse der statistischen Auswertung der Kreuzvalidierung für die drei verwendeten Interpolationsverfahren vergleichend gegenübergestellt.

Tabelle 4-4: *Vergleich der statistischen Kennzahlen der Kreuzvalidierungen*

Statistische Kennzahlen	IDW-Verfahren	Ordinary Kriging	Ordinary Cokriging
Stichprobenanzahl n	1762	1762	1762
Arithmetisches Mittel $\bar{x}$	0,00037	0,00009	0,00013
Median x_m	-0,00070	-0,00141	-0,00058
Standardabweichung s	0,01599	0,01541	0,01063
Schiefekoeffizient g	0,57399	0,71278	0,57863
MAE	0,01237	0,01179	0,00808
MSE	0,00026	0,00024	0,00011
Rangkorrelationskoeff. r_s	0,458**	0,493**	0,779**

** Die Korrelation ist auf dem Niveau von 0,01 signifikant (zweiseitig)

Das arithmetische Mittel $\bar{x}$ der Schätzfehler bzw. Residuen ist beim Ordinary Kriging und dem Ordinary Cokriging sehr ähnlich, wohingegen das arithmetische Mittel des IDW-Verfahrens wesentlich größer ist (vgl. Tabelle 4-4). Da alle drei Werte größer als Null sind, liegt für alle drei Interpolationsverfahren im Mittel eine systematische Unterschätzung der Schwefelwerte vor. Das Ausmaß dieser Unterschätzung ist beim Ordinary Kriging etwas geringer ausgeprägt als beim Ordinary Cokriging.

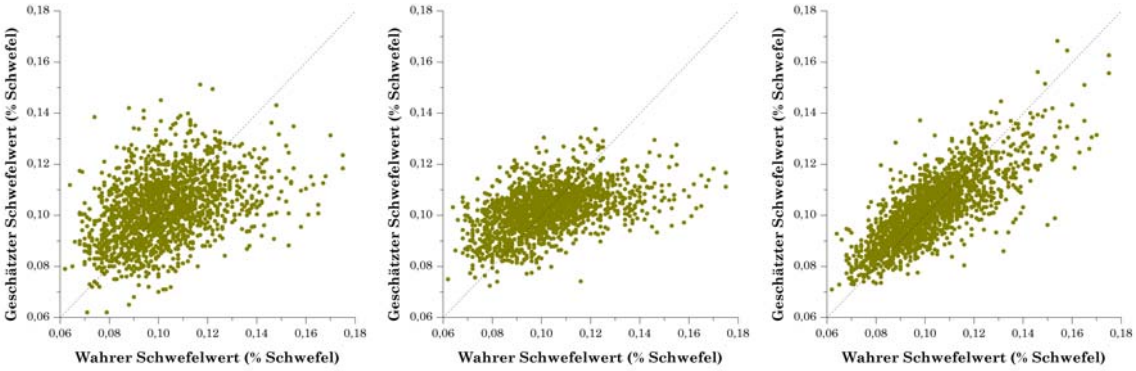
Die Werte der Standardabweichung s, als Maß für die Streuung der Residuenwerte um ihren Mittelwert, sind für die einzelnen Verfahren sehr unterschiedlich (vgl. Tabelle 4-4). Den höchsten Wert für die Standardabweichung weist das IDW-Verfahren auf, gefolgt vom Ordinary Kriging. Den kleinsten Wert und damit die geringste Streuung der Residuenwerte um den Mittelwert besitzt das Ordinary Cokriging.

Die Schiefekoeffizienten g sprechen bei den drei Verfahren ebenfalls für eine Verzerrung der geschätzten Werteoberfläche der Schwefelwerte. Der Wert g ist für alle Interpolationsverfahren größer als Null. Somit liegt für alle Residuen-verteilungen eine positive Schiefe vor, welche für das Ordinary Kriging am größten und für das Ordinary Cokriging etwas kleiner ist. Den kleinsten Wert für den

Schiefekoeffizienten nimmt das IDW-Verfahren an. Damit ist bei diesem Verfahren die positive Schiefe der Residuenverteilung etwas geringer ausgeprägt als bei den beiden anderen Verfahren.

Der MAE liegt bei allen Interpolationsverfahren im Bereich von Null, wobei aus dem IDW-Verfahren der höchste Wert resultiert (vgl. Tabelle 4-4). Für den MSE besitzt das IDW-Verfahren ebenfalls den größten Wert. Einen etwas geringeren Wert nimmt das Ordinary Kriging an, wohingegen der Wert für das Ordinary Cokriging wesentlich kleiner ist als für die beiden anderen Verfahren.

Abbildung 4-16 zeigt den Vergleich der Korrelationsdiagramme für die drei verwendeten Interpolationsverfahren.



IDW-Verfahren **Ordinary Kriging** **Ordinary Cokriging**

Abbildung 4-16: *Vergleich der Korrelationsdiagramme*

Aus den Korrelationsdiagrammen in Abbildung 4-16 geht hervor, dass die Schwefelwerte für das IDW-Verfahren und dem Ordinary Kriging ähnliche streuende Punktewolken abbilden. Eine positive Korrelation zwischen den wahren und geschätzten Schwefelwerten ist hier, wie auch beim Ordinary Cokriging zu erkennen. Die Rangkorrelationskoeffizienten r_s in der Tabelle 4-4 bestätigen, dass ein positiver signifikanter Zusammenhang zwischen den wahren und geschätzten Schwefelwerten für alle drei Verfahren besteht. Dieser Zusammenhang ist für das Ordinary Kriging sowie IDW-Verfahren mittelstark und liegt bei ungefähr 0,50. Im Vergleich dazu ist der Rangkorrelationskoeffizient für das Ordinary Cokriging mit rund 0,78 wesentlich höher. Auffällig ist zudem, dass die Punktewolke für das Ordinary Cokriging über ein größeres Werteintervall der geschätzten Schwefelwerte streut, als die Punktewolken der beiden anderen Verfahren (vgl. Abbildung 4-16). Weiters zeigt sich, dass die hohen Schwefelwerte beim IDW-Verfahren und dem Ordinary Kriging stärker unterschätzt werden als beim Ordinary Cokriging. Die Überschätzung der niederen Schwefelwerte ist beim Ordinary Cokriging ebenfalls geringer ausgeprägt als bei den beiden anderen Verfahren (vgl. Abbildung 4-16).

Die zur Abschätzung der Qualität von Interpolationsverfahren wesentlichen statistischen Kennwerte der Residuen sind neben dem Rangkorrelationskoeffizient, der MAE und der MSE. Die Werte für den MAE und den MSE sind für die drei Interpolationsverfahren sehr unterschiedlich und weisen für das Ordinary Kriging sowie IDW-Verfahren die größten Werte auf. Im Gegensatz dazu sind der MAE und der MSE für das Ordinary Cokriging wesentlich kleiner, was für eine höhere Qualität des Ordinary Cokriging Verfahrens zur Schätzung der Schwefelwerte spricht. Bei der Berechnung des MSE werden die größeren Residuenwerte durch das Quadrieren stärker gewichtet als die kleineren Residuenwerte. Das IDW-Verfahren und Ordinary Kriging verursachen also größere Residuenwerte als das Ordinary Cokriging. Außerdem weist die Residuenverteilung mit den entsprechenden statistischen Kennwerten auf eine geringere Verzerrung der geschätzten Werteoberfläche der Schwefelwerte beim Ordinary Cokriging hin.

Zusammenfassend kann festgehalten werden, dass die Qualität des Ordinary Cokriging Verfahrens zur flächenhaften Schätzung der Schwefelwerte wesentlich höher einzuschätzen ist als jene der beiden anderen Verfahren. Vergleicht man Qualität des IDW-Verfahrens mit jener des Ordinary Kriging Verfahrens so sieht man, dass das Ordinary Kriging eine etwas höhere Qualität besitzt.

4.5 Schwefelbelastungskarte der Steiermark

In Abbildung 4-17 sind die mittels Ordinary Cokriging geschätzten Schwefelwerte für das Jahr 2003 in Form einer Schwefelbelastungskarte für die Steiermark, unter Berücksichtigung der tatsächlichen Waldausbreitung, dargestellt. Die Klassifizierung der Schwefelwerte erfolgte in Anlehnung an die Vorgehensweise der Fachabteilung für das Forstwesen des Landes Steiermark anhand von fünf Klassen (vgl. Tabelle 3-1).

Die Schwefelbelastung der steirischen Wälder liegt zum Großteil unterhalb des Grenzwertes von 0,11% Schwefel für Daten des ersten Nadeljahrgangs. Belastete Waldgebiete findet man vor allem in der Umgebung von Ballungsräumen und Industriezentren, wie z.B. Leoben, Judenburg, Knittelfeld oder Bruck an der Mur (vgl. Abbildung 4-17). Diese Waldgebiete wurden mehrheitlich als *„leicht belastet*“ klassifiziert.

Vereinzelt wurden Waldflächen auch als „*belastet*“ ausgeschieden, wie etwa im Raum Knittelfeld. Die Flächen dieser Waldgebiete sind allerdings so gering, dass sie in der Karte nur schwer zu erkennen sind. „*Unbelastete*“ bzw. „*stark belastete*“ Waldgebiete konnten überhaupt nicht festgestellt werden.

Schwefelbelastungskarte der Steiermark für das Jahr 2003

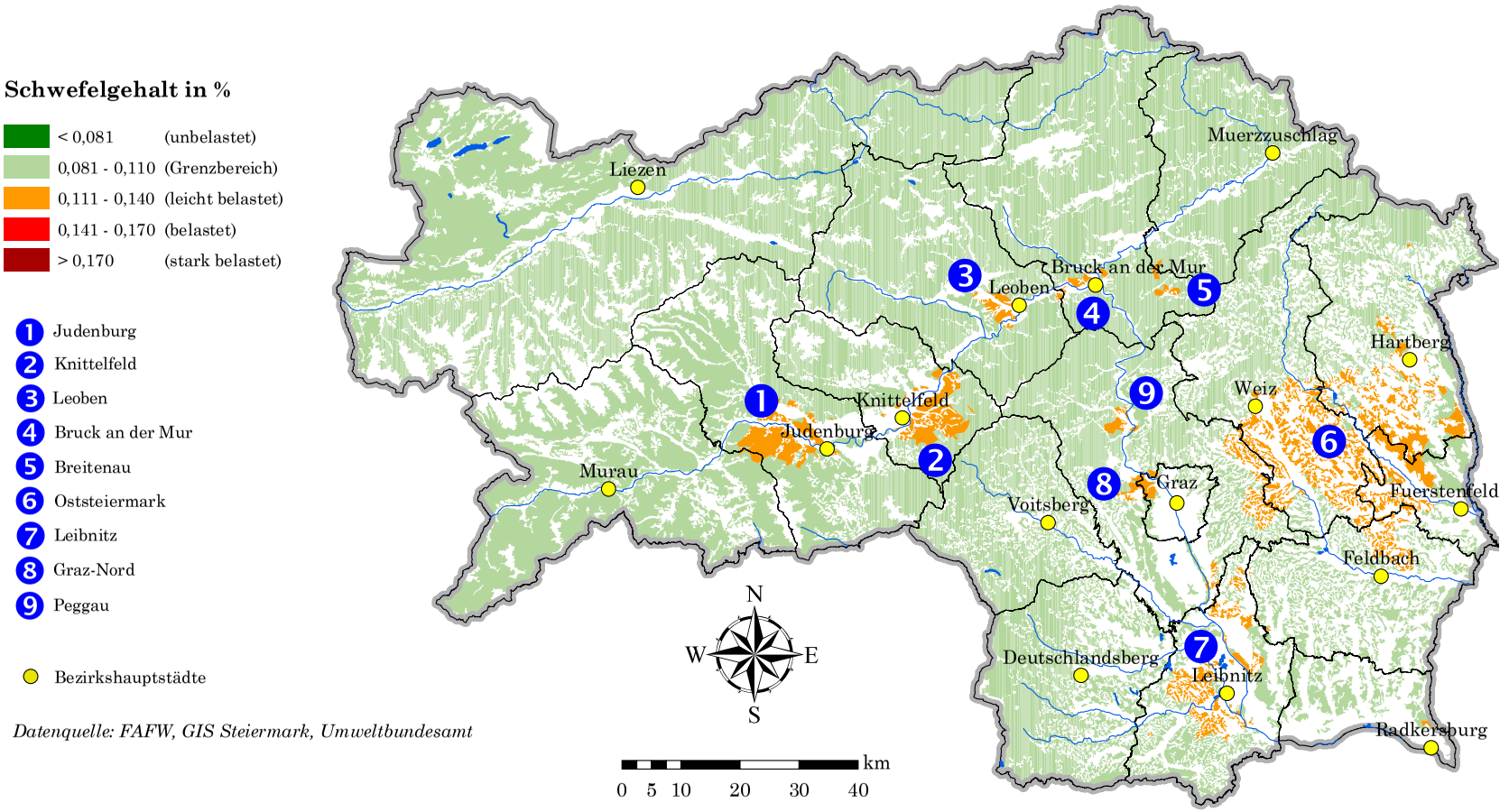


Abbildung 4-17: *Schwefelbelastungskarte der Steiermark für das Jahr 2003*

5 Diskussion der Ergebnisse

Die Ergebnisse der Kreuzvalidierung von den verwendeten Interpolationsverfahren wurden bereits im vorangegangen Kapitel vorgestellt bzw. interpretiert. In diesem Kapitel sollen nun die Ergebnisse unter dem Blickwinkel der Vor- und Nachteile der einzelnen Interpolationsverfahren betrachtet werden.

Krigingverfahren haben, wie alle lokalen Schätzmethoden, den entscheidenden Nachteil, dass die geschätzten Werteoberflächen geglättet sind und damit die Extremwerte im Allgemeinen über- bzw. unterschätzt werden (DOBLER et al 2003, JOHNSTON et. al. 2001, GOOVAERTS 1997). Die Glättung ist darüber hinaus nicht gleichförmig sondern an den Stützpunkten am geringsten. Eine Karte auf Basis optimaler lokaler Krigingschätzungen erscheint aus diesem Grund in dichter beprobten Gebieten häufig variabler bzw. detailreicher als in weniger dicht beprobten Regionen. Dabei handelt es sich aber nicht um einen Effekt des betrachteten räumlichen Prozesses, sondern um „Artefakte“ der Methode (DOBLER et. al 2003). Bei Fragestellungen, die v.a. die räumliche Verbreitung hoher bzw. extremer Werte beinhaltet, sind Krigingschätzer daher nur bedingt geeignet insbesondere zur Abschätzung der Aussagesicherheit (BLASCHKE & STROBL 2003). Auch bei sehr ungleich verteilten Stützpunkten oder Clusterbildung wird von Kriging, im Unterschied zu anderen Verfahren, abgeraten (CRESSIE 1993). Für die Schätzung anhand des IDW-Verfahrens wird von JOHNSTON et. al (2001) ebenfalls eine gleichmäßige Verteilung der Stützpunkte vorgeschlagen. Der Grund dafür ist, dass bei diesem Verfahren die Messpunkte, die innerhalb von Punktehäufungen liegen bei der Gewichtung überbewertet werden, während Werte außerhalb von Clustern mit einem zu geringen Anteil in die Schätzung einfließen (HAUSSMANN 2000). Dies spielt bei rasterförmig vorliegenden Daten keine Rolle, allerdings liegen Umweltdaten selten in dieser Form vor.

Betrachtet man die Ergebnisse der verwendeten Interpolationsverfahren in der vorliegenden Arbeit, so ist generell eine Überschätzung der niedrigen wahren Schwefelwerte und eine Unterschätzung der hohen wahren Schwefelwerte festzustellen. Diese Über- und Unterschätzungen sind besonders beim IDW-Verfahren und dem Ordinary Kriging stark ausgeprägt, während sie beim Ordinary Cokriging Verfahren wesentlich geringer ausfallen. Diese Problematik kann sich vor allem dann negativ auswirken, wenn es darum geht entsprechende Schwefelbelastungskarten zu erstellen. So kann es beispielsweise vorkommen, dass

durch die Unterschätzung der hohen Schwefelwerte stark belastete Waldgebiete in ihrer flächenmäßigen Erstreckung zu gering ausgewiesen werden bzw. im Extremfall überhaupt zur Gänze verschwinden.

Die räumliche Verteilung der in der vorliegenden Arbeit für die Schätzung herangezogenen BIN-Punkte ist relativ ungleichmäßig und weist zum Teil clusterförmige Strukturen auf. Diese Strukturen sind vor allem in der Umgebung der Industriezentren sichtbar und sind auf die dort eingerichteten Lokalnetze des Bioindikatornetzes zurückzuführen (vgl. Abbildung 2-5). Auf die Problematik der clusterförmigen Strukturen wurde bereits hingewiesen, welche allerdings in der vorliegenden Arbeit nicht berücksichtigt wurde. Stattdessen lag das Hauptaugenmerk darauf möglichst viele BIN-Punkte in die Schätzung einzubeziehen. Diese Denkweise deckt sich auch mit HEINRICH (1992) der empfiehlt, „*… je mehr Messpunkte, desto besser*". Allgemein gilt, je mehr Werte zur Schätzung eines unbekannten Punktes eingesetzt werden, desto zuverlässiger wir im Allgemeinen die Schätzung. (GERLACH 2001). Hingegen ist bei einer sehr geringen Anzahl an Messpunkten die Gefahr einer Fehlschätzung relativ groß.

Um die Problematik der clusterförmigen Strukturen bei der Lösung von zukünftigen Fragestellungen im Zusammenhang mit der Schwefelbelastung zu vermeiden ist zu überlegen ob es nicht sinnvolle wäre, je nach Auflösungsebene der Betrachtung unterschiedliche Datensätze für die Schätzung zu verwenden. So könnte man beispielsweise auf regionaler Ebene (Bundesland) nur die BIN-Punkte des Bundes- und Verdichtungsnetzes verwenden, während man für Fragestellungen auf lokaler Ebene (Bezirk, Gemeinde) auch jenen der Lokalnetze hinzunimmt. Für Gebiete mit kleinräumigen und stark wechselnden Nutzungsformen schlagen GOOVAERST (1997) und SINOWSKI & AUERSWALD (1999) vor, die Räume auf Basis von leicht zu erhebenden Flächeninformationen zu untergliedern und innerhalb dieser Teilräume eine Kriging-Interpolation durchzuführen.

Bei der Durchführung der Schätzung anhand von Interpolationsverfahren stellt sich immer wieder die Frage nach der Anzahl, der in die Schätzung einzubeziehenden Stützpunkte. Diese Anzahl ist zum Großteil von der zu untersuchenden Beobachtungsvariable und der vorgegebenen Messnetzdichte abhängig. Zu beachten ist, dass mit zunehmender Stützpunkteanzahl die Gewichtung der einzelnen Punkte immer geringer wird, da in Summe die Werte der Gewichte eins ergeben muss. Die Hinzunahme von weiteren Stützpunkten geht dabei vor allem zulasten der Gewichte jener Stützpunkte, die näher am zu schätzenden Punkt liegen. Beim IDW-Verfahren kann dieser Umstand durch die Erhöhung des Distanzexponenten kompensiert werden, wodurch die näher liegenden Stützpunkte ein stärkeres Gewicht erhalten (CLARK & HARPER 2000).

In der vorliegenden Arbeit wurden für die Durchführung der Schätzung anhand des IDW-Verfahrens zwischen 4 und 12 Stützpunkte herangezogen. Im Gegensatz

dazu erfolgte die Schätzung beim Ordinary Kriging bzw. Ordinary Cokriging mit bis zu 40 Stützpunkten (10 pro Sektor). Die Schätzung mit einer größeren Stützpunkteanzahl ist zwar statistisch stabiler, führt allerdings auch zu einer stärkeren Glättung der Schätzoberfläche. Vergleicht man die geschätzten Werteoberflächen mit Hilfe der Schwefelbelastungskarten für die einzelnen Interpolationsverfahren (vgl. Abbildung 4-5, Abbildung 4-10, und Abbildung 4-15) so erkennt man, dass die Schätzoberfläche für das IDW-Verfahren wesentlich „unruhiger" ist, als jene der beiden anderen Verfahren. Beim IDW-Verfahren kommen somit die lokalen Maxima der Schwefelwerte wesentlich stärker zur Geltung als bei den beiden anderen Verfahren. Im Gegensatz dazu werden beim Ordinary Kriging bzw. Ordinary Cokriging größere zusammenhängende Flächen mit gleicher Schwefelbelastung ausgeschieden als beim IDW-Verfahren.

Die automatische Erstellung der Schwefelbelastungskarte für die Steiermark (vgl. Abbildung 4-17) anhand des Geostatistical Analyst bringt zwar einige Vorteile mit sich, allerdings stellt sich auch die Frage, inwiefern diese Art der Regionalisierung einer manuellen Bearbeitung vorzuziehen ist. Die Schwefelbelastungskarte ist ein gutes Beispiel für eine Karte, die bei manueller Erstellung wahrscheinlich ein anderes Aussehen hätte. Bei der manuellen Erstellung der Karte kann der Bearbeiter sein zusätzliches Wissen um die lokale Situation der Schwefelausbreitung einbringen. Allerdings muss dieses zusätzliche Wissen zuerst einmal vorhanden sein, damit die manuelle Erstellung einen Vorteil gegenüber der automatisierten Regionalisierung hat.

Für den Vergleich von manuell und automatisiert erstellten Karten ist der von DAHLBERG (1975) durchgeführte Versuch von Interesse. Bei diesem Versuch wurde aus einem umfangreichen Datensatz zu geologischen Strukturdaten eine Teilmenge ausgewählt und erfahrenen Erdölgeologen zur Verfügung gestellt. Die manuell konstruierten Karten der Geologen wurden in weiterer Folge mit der automatisch generierten Karte verglichen, welche aus dem gesamten Datensatz der Strukturdaten erstellt wurde. Der Vergleich hat gezeigt, dass alle Karten in Bereichen mit hoher Stützpunktedichte ähnliche Ergebnisse lieferten, die mit den Informationen aus dem gesamten Datensatz in guter Übereinstimmung standen. In Bereichen mit geringer Stützpunkteanzahl haben sich jedoch die Karten zum Teil erheblich unterschieden. Während einige der Geologen sehr gute Schätzungen lieferten, wichen die Karten der anderen Geologen zum Teil erheblich von der Realität ab. Die mittels Computer erzeugte Karte lieferte weder eine überdurchschnittlich gute, noch schlechte Schätzung und entsprach in ihrer Qualität einer von einem „durchschnittlichen" Geologen manuell konstruierten Karte.

Die Entwicklung der Interpolationsverfahren und Computerprogramme in den letzten Jahren gibt Anlass zur Vermutung, dass sich seit dem damaligen Versuch von DAHLBERG (1975) die Programme erheblich verbessert haben. Aus diesem

Grund ist anzunehmen, dass eine heutzutage mit einem Computer erstellte Karte, die richtige Programmanwendung vorausgesetzt, einer manuell erstellten Karte gleichzustellen oder gar vorzuziehen ist.

Bei der heutigen Vielfalt an Interpolationsverfahren ist die Auswahl eines geeigneten Verfahrens und die Erfahrung des Bearbeiters entscheidend für die Qualität der Karte. Dies belegt auch ein Versuch von ENGLUND (2000) der die Varianz zwischen Geostatistikern und verschiedenen Verfahren untersuchte. Bei diesem Versuch ließ ENGLUND (2000) zwölf Wissenschafter mit geostatistischer Erfahrung unabhängig voneinander eine Regionalisierung durchführen. Die Ergebnisse variierten erheblich, was ENGLUND (2000) auf die Wahl unterschiedlicher Methoden, die abweichende Interpretation der Daten (Umgang mit Extremwerte) und zum Teil auch auf Fehler in der Bearbeitung zurückführte.

Abschließend soll an dieser Stelle noch kurz auf das vereinfachende Ablaufschema der Schätzung eines räumlichen Prozesses eingegangen werden, das in Kapitel 3.2.3 „*Ordinary Kriging*“ erläutert wurde (vgl. Abbildung 3-7). Der sequenzielle Ablauf dieses Schemas ist zum Teil etwas irreführend. Das Schema gibt zwar den von oben nach unten gerichteten Handlungsbedarf wieder, vergisst allerdings auf den dahinter stehenden Planungsbedarf, der sich in Wechselbeziehungen zwischen den dargestellten Schritten vollzieht. Das bedeutet, dass bereits die Planung eines Messnetzes auf die Bedürfnisse des zu verwendeten Schätzverfahrens zugeschnitten sein sollte (GERLACH 2001). Dies spiegelt sich auch in der Geostatistik wieder, in der sehr intensiv auf die Frage der Gestaltung von Messnetzen eingegangen wird. Ein wichtiges Problem besteht darin Messnetze zu finden, welche die vorgegebenen Genauigkeiten unter Berücksichtigung der Kostenvorgaben garantieren (STOYAN et. al 1997). Bei gegebener Messnetzdichte spielt auch die Geometrie der Messnetze eine Rolle. So haben z.B. mathematische Untersuchungen gezeigt, dass unter bestimmten Bedingungen Dreiecksgitter zweckmäßig sind, während unter anderen Bedingungen hexagonale Netze Vorteile bringen (STOYAN et. al 1997).

6 Zusammenfassung und Ausblick

6.1 Zusammenfassung

Um die Auswirkungen von Schadstoffen auf Pflanzen einschätzen zu können reichen physikalisch-chemische Messungen der Umweltmedien Wasser, Boden und Luft oft nicht aus. Ergänzend werden deshalb Verfahren des Biomonitorings eingesetzt, die den Zustand und die Veränderung der Umwelt anhand der Auswirkungen auf einen lebenden Organismus bewerten. Um diesen Organismus nicht zu schädigen oder gar zu töten, werden die Aussagen über die Auswirkungen aus einer indirekten Bewertung in Form eines Bioindikators abgeleitet. Als Bioindikatoren eignen sich besonders Pflanzen, da sie standortsgebunden sind und meist in großer Individuenanzahl natürlich vorkommen. Diese Art des Biomonitorings wird als passiv bezeichnet, da die Feststellung von Schäden im natürlichen Lebensraum bzw. am Wuchsort erfolgt. Als Beispiel für ein derartiges passives Biomonitoringsystem kann das österreichische Bioindikatornetz angesehen werden, in welchem Bäume als Bioindikatoren verwendet werden um Immissionsauswirkungen auf die Wälder festzustellen (FÜRST 2004)

In nahezu allen Fällen können räumliche Phänomene aus Kosten und prinzipiellen Gründen nicht vollflächig räumlich erfasst werden, weshalb man punktuelle Messungen in Form von z.B. Stichprobenrastern durchführt (BLASCHKE & STROBL 2003). Das Aussageziel der Erhebung bezieht sich jedoch meist nicht nur auf die Standorte der jeweiligen Messeinrichtung, sondern auf die gesamte Fläche des Untersuchungsgebiets. Man steht deshalb vor dem verbreiteten Problem, punktuelle Aussagen mit bestmöglicher Qualität auf den zwei- oder dreidimensionalen Raum zu übertragen. Räumliche Interpolationsverfahren bieten die Möglichkeit auf Basis von punktuell gemessenen Daten Aussagen über die flächenhafte Verteilung zu treffen. In der Fachliteratur ist eine Vielfalt an räumlichen Interpolationsmethoden zu finden, die sich zum Teil grundsätzlich in ihren Ansätzen unterscheiden. Im Zusammenhang mit speziellen Fragestellungen, wie z.B. die Schwefelbelastung von Wäldern, stellt sich deshalb häufig die Frage nach der Wahl der „geeignetsten“ Methode.

Ziel der vorliegenden Arbeit war es festzustellen, welches von drei ausgewählten Interpolationsverfahren sich am Besten zur Schätzung der räumlichen Verteilung der Schwefelbelastung in der Steiermark eignet, und inwieweit sich auf Basis des Verfahrens mit der höchsten Qualität Aussagen über potentiell gefährdete

Regionen für die Schwefelbelastung treffen lassen. Bei den zu analysierenden Interpolationsverfahren handelte es sich um die Methode der Inversen Distanzgewichtung, dem Ordinary Kriging und dem Ordinary Cokriging.

Die Inverse Distanzgewichtung ist eine nichtstatistische, lokale Interpolationsmethode, welche die unbekannten Werte nur mit Hilfe der geometrischen Distanzen schätzt. Im Gegensatz dazu sind die beiden anderen Verfahren, wesentlicher komplexer, da diesen ein geostatistisches Modell zugrunde liegt. Gemeinsam ist diesen Verfahren, dass es Regressionstechniken sind, die eine Minimierung der Schätzvarianz zum Ziel haben (HEINRICH 1992).

Die Durchführung der räumlichen Interpolation erfolgte anhand von Schwefeldaten des steirischen und österreichischen Bioindikatornetzes, welche von der Fachabteilung für das Forstwesen des Landes Steiermark sowie dem Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft zur Verfügung gestellt wurden. Um an den Randbereichen des Untersuchungsgebiets eine stabile Interpolation zu gewährleisten beinhaltete der Datensatz zusätzlich Schwefeldaten von den benachbarten Bundesländern. Einzige Ausnahme bildete dabei der Grenzbereich zu Slowenien, wo bedingt durch fehlende Schwefeldaten von Slowenien mit instabilen Werten der Interpolation zu rechnen ist. Die Schwefeldaten lagen sowohl für den ersten als auch zweiten Nadeljahrgang der Bioindikatornetzbäume vor. Während die Daten des ersten Nadeljahrgangs bei allen drei Interpolationsverfahren zur Schätzung der Schwefelwerte herangezogen wurden, gingen die des zweiten nur als Zusatzinformation, neben den Geländehöhendaten, in das Ordinary Cokriging ein.

Für die erfolgreiche räumliche Interpolation einer Werteoberfläche ist eine intersubjektiv nachvollziehbare Prüfung der Qualität der entsprechenden Interpolationsmethode unverzichtbar (DUMFARTH & LORUP 2000). Im Rahmen dieser Arbeit wurde dafür die in der Geostatistik häufig angewandte Methode der Kreuzvalidierung herangezogen. Die Grundidee dieser Methode besteht darin, aus einer vorhandenen Stichprobe einen gemessenen Wert zu entfernen und diesen mittels der anderen Werte der Stichprobe unter Verwendung des Interpolationsverfahrens zu schätzen (HENNIG 2004, SODOUDI 2004, GOOVAERTS 1997, CRESSIE 1993). Damit liegt für jeden Stichprobenwert sowohl ein wahrer als auch ein geschätzter Wert vor, deren Differenz den Schätzfehler (Residuum) an dem entsprechenden Ort bildet. Die Auswertung der Kreuzvalidierung erfolgte auf Basis der Schätzfehler anhand von verschiedenen statistischen Kennzahlen (z.B. Median, Schiefekoeffizient, Rangkorrelationskoeffizient, MSE, MAE), welche letztendlich Aufschluss über die Qualität eines Interpolationsverfahren gaben.

Vergleicht man die statistischen Kennzahlen der verwendeten Interpolationsverfahren so kann unter anderem festgestellt werden, dass die drei Interpolationsverfahren die wahren Schwefelwerte im Mittel unterschätzen. Das Ausmaß dieser Unterschätzung ist beim Ordinary Kriging Verfahren geringer ausgeprägt als bei

den beiden anderen Verfahren. Für eine Verzerrung der geschätzten Werteoberfläche sprechen auch die Werte der Schiefekoeffizienten, welche für alle drei Verfahren auf eine positive Schiefe hinweisen. Zwischen den wahren und geschätzten Schwefelwerten besteht bei allen drei Verfahren ein signifikanter positiver Zusammenhang (Korrelation). Diese Korrelation ist für das Ordinary Kriging sowie IDW-Verfahren mittelstark ausgeprägt und liegt bei rund 0,50. Im Vergleich dazu ist der Rangkorrelationskoeffizient für das Ordinary Cokriging Verfahren mit 0,78 wesentlich höher. Weiters zeigt sich, dass die hohen Schwefelwerte beim IDW-Verfahren und beim Ordinary Kriging stärker unterschätzt werden als beim Ordinary Cokriging. Die Unterschätzung der niedrigen Schwefelwerte ist beim Ordinary Cokriging ebenfalls geringer als bei den beiden anderen Verfahren. Die Werte für den MAE und den MSE sind für die drei Interpolationsverfahren sehr unterschiedlich und weisen für das Ordinary Kriging sowie IDW-Verfahren die größten Werte auf. Im Gegensatz dazu sind der MAE und MSE für das Ordinary Cokriging wesentlich kleiner, was wiederum für eine höhere Qualität des Ordinary Cokriging Verfahrens zur Schätzung der Schwefelwerte spricht.

Zusammenfassend kann festgehalten werden, dass die Qualität des Ordinary Cokriging Verfahrens zur flächenhaften Schätzung der Schwefelwerte wesentlich höher einzuschätzen ist als jene der beiden anderen Verfahren. Vergleicht man die Qualität des IDW-Verfahrens mit jener des Ordinary Kriging Verfahrens so sieht man, dass das Ordinary Kriging eine etwas höhere Qualität besitzt.

6.2 Ausblick

Um die Qualität der flächenhaften Schätzung der Schwefelwerte anhand des Ordinary Cokriging Verfahrens zu verbessern, ist anzudenken ob es nicht sinnvoll wäre weitere Zusatzinformationen für die Schätzung zu verwenden. Neben Standortsinformationen wie z.B. Bodentyp oder Exposition würden sich vor allem andere Schadstoffgehalte in den Nadel(Blatt)proben wie Fluor, Chlor aber auch Nährstoffe wie Stickstoff, Phosphor, Kalium, Kalzium und Magnesium anbieten. Diese Nährstoffe bzw. Schadstoffgehalte werden je nach Bedarf für einen Teil der BIN-Punkte bei der Auswertung im Labor untersucht und liegen somit vor. Weiters ist zu überlegen, ob es in Zukunft nicht sinnvoll wäre Stützpunkte des slowenischen Bioindikatornetzes in die Schätzung einzubeziehen. Damit wäre zumindest gewährleistet, dass auch für den Süden der Steiermark stabile Interpolationsergebnisse vorliegen.

Bisher wurde nur der „stationäre“ Zustand der Schwefelbelastung betrachtet, anhand dessen allerdings keine Aussage über die zeitliche Veränderung der Schwefelbelastung getroffen werden kann. Um die Veränderung feststellen zu

können müssen wiederholt gemessene Werte über einen längeren Zeitraum in Form einer Zeitreihe (engl. time series) vorliegen. Diese Bedingung erfüllen die erhobenen Schwefeldaten, welche für einen mehr als 20 jährigen Zeitraum vorhanden sind. Bei der statistischen Analyse von Zeitreihen tritt das Problem der zeitlichen Abhängigkeit auf, da im Allgemeinen nicht davon ausgegangen werden kann, dass die Messwerte voneinander unabhängig sind (STOYAN et al. 1997). Diese zeitlichen Abhängigkeiten sind oft ziemlich kompliziert und entstehen durch Überlagerungen, Verknüpfungen von „Trends", periodischen Schwankungen sowie kurzzeitigen, zufälligen Einflüssen. Das wesentliche Ziel einer Zeitreihenanalyse ist die Entdeckung von Zusammenhängen in Zeitreihen bzw. zwischen verschiedenen Zeitreihen. Für die Entdeckung derartiger Zusammenhänge sind vor allem die explorative Datenanalyse sowie die Form der Datendarstellung wichtig (SCHLITTGEN & STREITBERG 2001). Zu viele Einzelheiten in der Darstellung würden das Wesentliche überdecken, während eine zu starke Verallgemeinerung zum Verschwinden von Details führen könnte. Gelingt es, für die Zeitreihe(n) ein mathematisches Modell zu finden, dann ist ein wesentlicher Schritt zu deren Verständnis gegeben. Die Anwendung von anspruchsvolleren Methoden der Zeitreihenanalysen, welche auf Modellannahmen beruhen wie z.B. MA(q)-Prozess, ist für Umweltdaten problematisch (STOYAN et. al 1997). Der Grund hiefür ist, dass die Zeitreihen von Umweltdaten selten stationär sind und durch Ausreißer sowie fehlende Werte in der Qualität beeinträchtigt werden. Eine gute Einführung in die Methodik der Zeitreihenanalyse findet man z.B. in CHATFIELD (2003), SCHLITTGEN & STREITBERG (2001), STOYAN et. al. (1997), STORM (1985) und SCHMITZ (1989). Die mathematischen Grundlagen der Zeitreihentheorie sind in BROCKWELL & DAVIS (1991) und BOX et. al (1994) dargestellt. In MICHELS (1992) werden nicht parametrische Methoden in der Analyse und Prognose von Zeitreihen ausführlich auf Umweltprobleme, wie z.B. Hochwasserschutz oder Gewässer- und Luftbelastung angewandt.

Gänzlich unberücksichtigt blieben bisher auch die komplexen räumlichen Strömungs- und Turbulenzvorgänge welche maßgeblich für die Ausbreitung von Luftschadstoffen in bodennahen Luftschichten verantwortlich sind. Diese sind z.B. für bodennahe Quellen neben den allgemeinen meteorologischen Bedingungen auch von den topographischen Gegebenheiten abhängig (ÖTTL et. al 2004, BERLEKAMP 2000, PRABHA 1996). In erster Linie sind dabei das thermische Verhalten des Talbereiches und seine Wechselwirkungen mit der überlagerten mesoskalen Strömung bzw. synoptischen Strömung unter verschiedenen Wetterlagen von Bedeutung (PRABHA 1996). Die nächtliche Abkühlung erzeugt im hügeligen oder bergigen Gelände eine Strömung, welche talauswärts gerichtet ist (Bergwind). Die Strömung ist gleichzeitig mit einer bodennahen Inversion verbunden, wobei es zu einer starken Reduktion der vertikalen Durchmischungsaktivität der bodennahen Atmosphäre kommt (HOFKO 1994). Am Tag kommt es dagegen auf Grund der starken Erwärmung der Hänge und der Talatmosphäre zu

einem Taleinwärtsströmen von Luftmassen (Talwind). Diese Strömung ist besonders intensiv während der Sommermonate und durch die Sonneneinstrahlung gut vertikal durchmischt (HOFKA 1994). Daraus ist zu erkennen, dass für die Ausbreitung von Schadstoffe die vertikalen Strukturen unter den für das Gebiet typischen Strömungssituationen von besonderer Bedeutung sind. Aus diesem Grund muss die Frage nach der Häufigkeit, der Mächtigkeit, Lufttemperatur und der Turbulenzaktivität dieser bodennahen Luftschichten am Standort beantwortet werden. Für die Erfassung, Analyse und Überwachung der Luftschadstoffbelastungen werden Messverfahren, Windkanal-Experimente und Modellrechnungen eingesetzt. Der Einsatz von Modellen wird sich in Zukunft weiter steigern, da die Erfassung von Messdaten und die experimentelle Ermittlung von Daten einen wesentlich höheren zeitlichen und finanziellen Aufwand erfordern. Durch den Einsatz von Modellen ist außerdem die Durchführung von Simulationen, z.B. Planungsvarianten, oder der Prognose von zukünftigen Situationen möglich. Je nach Problemstellung, Zielsetzung, räumlicher Ausdehnung und Datenverfügbarkeit werden entweder einfache statistische Modelle (z.B. Regressionsanalyse) oder physikalisch sehr anspruchsvolle Modelle (z.B. Gaußmodell) eingesetzt. Für vertiefende und weiterführende Informationen zur Ausbreitungsmodellierung sei z.B. auf HELBIG et. al (1999) und ZENGER (1998) verwiesen.

7 Literaturverzeichnis

AKIN, H. und SIEMES, H. (1988): Praktische Geostatistik – Eine Einführung für den Bergbau und die Geowissenschaften. Springer Verlag, Berlin-Heidelberg-New York, 304 S.

ARMSTRONG, M. (1998): Basic Linear Geostatistics. Springer Verlag, Berlin-Heidelberg, 1st Edition, 155 S.

BARTSCH H.-J. (1991): Taschenbuch mathematischer Formeln. Fachbuchverlag Leibzig, 14.Auflage, 605 S.

BATTAGLIA, G. J. (1996): Mean Square Error. AMP Journal of Technology. Vol. 5 (1996), pp. 31-36.

BERLEKAMP, J. FUEST, S. de LANGE, N. und LUBERICHS, M. (2000): Emissions- und Immissionskataster für das Stadtgebiet Münster als Komponente eines kommunalen Umweltinformationssystems. In: Umweltinformatik – Umweltinformation für Planung, Politik und Öffentlichkeit. Metropolis-Verlag, Band 26, 728 S.

BILL, R. (1996): Grundlagen der geographischen Informationssysteme. Band 2: Analysen, Anwendungen und neue Entwicklungen. Wichmann-Verlag, Karlsruhe-Heidelberg, 463 S.

BLÖSCHL, G. and GRAYSON, R. (2000): Spatial observations and interpolation. In: Grayson, R. and Blöschl, G, (Eds.) Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge Universtiy Press, Cambrige, pp. 17-50.

BOX, G. E. P., JENKINS, G. M. and REINSEL, G. C. (1994): Time Series Analysis: Forecasting and Control. Publisher Prentice Hall, Englewood Cliffs, 3rd Edition, 598 p.

BRINKMANN, J. (2002): Räumliche Variabilität von Böden und Bodeneigenschaften auf dem landwirtschaftlichen Versuchsgut Frankenforst im Pleiser Hügelland. Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 157 S.

BROCKWELL, P. J. and DAVIS, R. A. (1991): Time Series: Theory and Methods, Springer Verlag, Berlin-Heidelberg-New York, 2nd Edition, 592 p.

BROOKER, P.I. (1986): A Parametric Study of Robustness of Kriging Variance as a Function of Range and Relative Nugget Effect for a Spherical Semivariogram. Journal of the International Association for Mathematical Geology 18 (5), pp. 477-488.

BULLINGER, A. (2000): Vergleich von Interpolationsverfahren und digitalen Geländemodellen für Wattgebiete – GIS gestützte Untersuchungen am Beispiel der morphologischen Modellierung von Tidebecken in Schleswig-Holstein. Diplomarbeit, Universität Kiel, 120 S.

BUNDESGESETZBLATT (1984): Zweite Verordnung gegen Forstschädliche Luftverunreinigungen 1984. BGBl 199/1984. Erlassen durch den Bundesminister für Land- und Forstwirtschaft 24.April 1984.

CHATFIELD, C. (2003): The Analysis of Time Series. Publisher Chapman & Hall/CRC Press, London-Boca Raton, 6th Edition, 333 p.

CHILES, J.-P. and DELFINER, P. (1999): Geostatistics – Modeling Spatial Uncertainty. Publisher John Wiley & Sons, New York-Toronto, 1st Edition, p 720.

CLARK, I. and HARPER, W. V. (2000): Practical Geostatistics. Publisher Ecosse North Amer LlC, Spiral edition, 442 p.

CRESSIE, N. A. C. (1993): Statistics for Spatial Data. Publisher John Wiley & Sons, New York, 928 p.

CZEGKA, W., BEHRENDS, K. und BRAUNE, S. (2004): Die Qualität der SRTM-90m Höhendaten und ihre Verwendbarkeit in GIS. 24. Wissenschaftlich-Technische Tagung der DGPF, 15.-17.9.2004, Halle, 11 S.

DAHLBERG, E.C. (1975): Relative Effectiveness of Geologists and Computers in Mapping potential Hydrocarbon Exploration Targets. Journal of the International Association for Mathematical Geology, Vol. 7 (5/6), pp. 374-394.

DÄSSLER, H.-G. (1991): Einfluss von Luftverunreinigungen auf die Vegetation. Urban & Fischer Verlag (Elsevier), 4. Auflage, 266 S.

DAVIS, M. (1987): Production of conditional simulations via the LU decomposition of the covariance matrix. Journal of the International Association for Mathematical Geology, Vol. 19 (2), pp. 91-98.

DEUTSCH, C.V. and JOURNAL, A.G. (1998): Geostatistical Software Library and User's Guide. Oxford-University Press, New York-Oxford, 2nd Edition, 340 p.

DOBELER, L, HINTERDING, A., MÜLLER, A., GERLACH, N. und GABEL, F. (2003): Geostatistische und statistische Methoden du Auswerteverfahren für Geodaten mit Punkt- bzw. Flächenbezug. Abschlussbericht Teil 3: Empfehlungen für die Anwendung statistischer und geostatistischer Methoden zur flächenbezogenen Auswertung von Daten über Stoffgehalte in Böden. Institut für Geoinformatik der Westfälischen Wilhelms-Universität Münster, 55 S.

DUTTER, R. (1985): Geostatistik – Eine Einführung mit Anwendungen. Teubner Verlag, Stuttgart, 159 S.

DUTZ, M., GEHRIG, S., HARTMANN, K. und NELLES, M. (2002):Grundlagen der Umwelttechnik, Teil 3: Umweltauswirkungen durch Schadstoffe, Lärm und Strahlung. Studienunterlagen. Fachhochschule Hildesheim-Holzminden-Göttingen, 58 S.

ENGLUND, E. J. (2000): A Variance of Geostatisticians. Journal of the International Association for Mathematical Geology 22, pp 417-455.

EUROPÄISCHE KOMMISSION (2000): Der Waldzustand in Europa – Kurzbericht 2000. UN/ECE und EK, Genf und Brüssel, 33 S.

EUROPEAN ENVIRONMENT AGENCY (1993): Corine Land Cover, Technical Guide, Report EUR 12585 EN, Bruxelles Luxembourg, 144p.

FORSTGESETZ 1975: Bundesgesetzblatt 440/1975.

FÜRST, A. (2001a): Forstliches Biomonitoring in Österreich. Österreichische Forstzeitung 3/2001, 37-38.

FÜRST, A. (2001b): Österreichisches Bioindikatornetz – Schwefelimmissions-einwirkungen 2000. FBVA-Berichte, Wien (119), 84S.

FÜRST, A. (2004): Österreichisches Bioindikatornetz. Praxisinformation des Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft, Nr. 5 - 2004, 1315.

GANDIN, L.S. (1963): Objective Analysis of Meteorlogical Field. Leningrad, Hydrometeorological Publ. House. English translation: Israel Program for Scientific Translations, Jerusalem, 1965, 242 p.

GERLACH, N. (2001): Ein Vergleich räumlicher Interpolationsverfahren für Windgeschwindigkeitsdaten in komplexem Gelände. Westfälische Wilhelms-Universität Münster, Fachbereich Geowissenschaften, 108 S

GOOVAERTS, P. (1997): Geostatistics for Natural Resources Evaluation. Oxford University Press, New York-Oxford, 1st Edition, 487 p.

GOOVAERTS, P. (1998): Ordinary Cokriging Revisited. Journal of the International Association for Mathematical Geology Vol. 30 (1), pp. 21-42.

GOOVAERTS, P. (2001): Geostatistical modelling of uncertainty in soil science. Geoderma, Vol. 103 (2001), pp. 3-26.

GOOVAERTS, P. and SAITO, H. (2000): Geostatistical Interpolation of Positively Skewed and Censored Data in a Dioxin-Contaminated Site. Environmental Science and Technology 34, pp. 4228-4235.

GORDON, W.J. and WIXOM, J.A. (1978): Shepard´s method of "Metric interpolation" to bivariate and multivariate interpolation. Mathematics of Computation, Vol 32 (1978), pp. 253-264.

GRAMS, S. (2000): Einsatz geostatistischer Verfahren zur Charakterisierung der Grundwasserbeschaffenheit im Bereich der Rieselfelder südlich Berlins. Dissertation an der Technischen Universität Berlin, 146 S.

HARTUNG, J., ELPELT, B., und KLÖSENER, K.-H. (2002): Statistik: Lehr- und Handbuch der angewandten Statistik. Oldenburg-Verlag, München-Wien, 13.Auflage, 975 S.

HAUSSMANN, M. (2000): GIS – gestützte Regionalisierung und Bewertung von punkthaft gemessenen Schwermetallgehalten in Oberböden. Diplomarbeit an der Universität Stuttgart, 148 S.

HEINRICH, U. (1992): Zur Methodik der räumlichen Interpolation mit geostatistischen Verfahren – Untersuchungen zur Validität flächenhafter Schätzungen diskreter Messungen kontinuierlicher raumzeitlicher Prozesse. Wiesbaden, Deutscher Universitäts-Verlag, 1.Auflage, 124 S.

HELBIG, A., BAUMÜLLER, J. und KERSCHGENS, M. J. (1999): Stadtklima und Luftreinhaltung. Springer Verlag, Berlin-Heidelberg, 2. Auflage, 467 S.

HERBST, M. (2001): Regionalisierung von Bodeneigenschaften unter Berücksichtigung geomorphometrischer Strukturen für die Modellierung der Wasserflüsse eines mikroskaligen Einzugsgebiets. Dissertation, Rheinische Friedrichs-Wilhelms-Universität Bonn, S. 130.

HERZFELD, U.C. (1989): Variography of submarine morphology: problems for deregularization and cartographical implications. Journal of the International Association for Mathematical Geology, Vol. 21(7), pp. 693-713.

HINDERDING, A., KÜHNE, S., SLEPTSOV, B., STREIT, U. und HEINEN, T. (2000): Integration wissenbasierter Methoden und gestatischer Verfahren zur räumlichen Interpolation. Wasser und Boden 52 (3), 18-21.

HINTERDING, A., MÜLLER, A., GERLACH, N. und Gabel, F. (2003): Geostatistische und statistische Methoden und Auswertverfahren für Geodaten mit Punkt- und Flächenbezug. Abschlussbericht Teil1: Grundlagen, Institut für Geoinformatik der Westfälischen Wilhelms-Universität Münster. 145 S.

HOFFMANN, H. (2002): Stochastisch-deterministische Modelle zur Analyse und Prognose räumlicher und zeitlicher Ozonimmissionsstrukturen in Sachsen. Dissertation, Technische Universität Bergakademie Freiberg, 131 S.

HOFKO, M. (1994): Statistisch-klimatologische Bearbeitung von Windmessungen aus dem Raum Linz. Diplomarbeit, Universität Innsbruck, 90 S.

ISAAKS, E.H. and SRIVASTAVA, R.M. (1989): An introduction to Applied Geostatistics. Oxford University Press, New York-Oxford, 1st Edition, 592 p.

JARVIS, A., RUBIANO, J., NELSON, A. FARROW, A. and MULLIGAN, M. (2004): Pratical use of SRTM data in the tropics – Comparisons with digital elevation models generated from cartographic data. (Working document No. 198). Centro Internacional de Agricultura Tropical (CIAT), 35 p.

JONSTON, K., VER HOEF, J.M., KRIVORUCHKO, K. and LUCAS, N. (2001): Using ArcGIS™ Geostatistical Analyst. ESRI, Redlands, 1st Edition, 300 p.

JOURNEL, A.G. and HUIJBREGTS, C.J. (1978): Mining Geostatistics. Academic Press, London-New York-San Francisco, 600 p.

KARHU, M. and HUTTUNEN, S. (1986): Erosion effects of air pollution on needle surfaces. Water, Air & Soil Pollution 31, pp. 417-423.

KIENZLE, A. (2002): Seismische Mikrozonierung von Bukarest – eine GIS gestützte statistische Analyse der Erdbebengefährdung. Dissertation, Universität Karlsruhe, 184 S.

KITANIDIS, P.K. (1997): Introduction to Geostatistics – Applications in Hydrogeology. Cambridge University Press, Cambridge, 249 p.

KRIGE, D.G. (1951): A Statistical Approach to some basic Mine Valuation Problems on the Witwatersrand. Journal of the Chemical, Metallurgical and Mining of South Africa 52(6), pp. 119-139.

LICK, H., GUNDL, K., KLAUSBAUER, R., STADLMANN, G., ALTENBURGER, H., DORFER, A., FREITAG, B. und SCHUSTER, T. (1998): Der Zustand des steirischen Waldes. Bericht über Untersuchungen und Erhebungen der Fachabteilung für das Forstwesen, Amt der steiermärkischen Landesregierung, 107 S.

MAIER, E.A., MUNTAU, H. and GRIEPINK, B. (1989): Certified reference materials – beech leaves and spruce needles – for the quality control in monitoring damage in forests by acid deposition. Journal of Analytical Chemistry Fresenius 335, pp. 833-838.

MARRA, M., CARANADE, R.E., LEDNER, R. and LEBERL, F.W. (1998): Accuracy of DEMs from Remotely Sensed Radar Images. In: GIS/LIS 98 Proceedings, Fort Worth, Texas, 10-12. November 1998, 356 p.

MATHERON, G. (1963): Principles of Geostatistics. Economic Geology 58 (8), pp. 1246-1266.

MATHERON, G. (1970): The Theory of Regionalized Variables its Applications. Les Cahiers du Centre de Morphologie Mathématiques, Ecole des Mines de Paris, Fontainebleau, 211p.

MAYER, H. (1985): Waldverwüstende Immissionsschäden in Österreich. Herausgeber Institut für Waldbau, Universität für Bodenkultur Wien, vorläufige gekürzte Fassung, 11 S.

MAYER, H. (1992): Waldbau auf soziologisch-ökologischer Grundlage. Verlag Gustav Fischer, Stuttgart-Jena-New York, 4. Auflage, 522 S.

MERKEL, B. und PLANER-FRIEDRICH, B. (2003): Integrierte Datenauswertung Hydrogeologie. Freiberg Online Geology, Vol. 7 (2003), 1-61.

MICHELS, P. (1992): Nichtparametrische Analyse und Prognose von Zeitreihen. Physika-Verlag, München-Wien, 234 S.

MYERS, J. (1997): Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping. Publisher John Wiley & Sons, New York, 571 p.

NATIONAL BUREAU OF STANDARDS (1976): Certificate of analyses – standard reference material 1575 – pine needles, 4 p.

NATIONAL BUREAU OF STANDARDS (1982): Certificate of analyses – standard reference material 1572 – citrus leaves, 4 p.

OLIVER, M.A. and WEBSTER, R. (1990): Kriging: A method of interpolation for Geographical Information Systems. International Journal of Geographical Information Systems 4 (2), pp. 313-332.

ÖTTL, D., STURM, P. und EICHLSEDER, H. (2004): Modellierung der Luftschadstoffe durch den KFZ-Verkehr entlang der Autobahnabschnitte Hallein und Salzburg. Forschungsgesellschaft für Verbrennungskraft-maschinen und Thermodynamik mbH, Bericht Nr. FVT-44/04/Öt V&U 04/07/6300, 53 S.

PALOMINO, I and MARTIN, F. (1994): A Simple Method of Spatial Interpolation of the Wind in Complex Terrain. Journal of Applied Meteorology 34, pp. 1678-1693.

PANNATIER, Y. (1996): Variowin – Software for Spatial Data Analysis in 2D. Springer Verlag, Berlin-Heidelberg-New York, 91p.

PLEYDELL, D. R. J., RAOUL, F., TOURNEUX, F., DANSON, F.M., GRAHAM, A. J., CRAIG, P. S. and GIRAUDOUX, P. (2004): Modelling the spatial distribution of Echinococcus multilocularis infection in foxes. Acta Tropica 91 (2004), pp. 253-265.

PORTERLE, K. (1891): Über die Beschädigung von Fichtenwaldbeständen durch schwefelige Säure. Österreichisches landwirtschaftliches Centralblatt 1, 27-38.

PRABHA T. V. (1996): Modelling of Pollutant Transport and Dispersion for Operational Use over Urban Area with Complex Terrain Surrounding. Dissertation, Universität für Bodenkultur Wien, 162 p.

RASE, W.-D. (1996): Interpolation von stetigen Oberflächen aus flächenbezogenen Informationen. In: Dollinger, F. und Strobl, J. (Eds.) Angewandte Geographische Informationsverarbeitung VIII = Salzburger Geographische Materialien, Heft 24. Selbstverlag des Institutes für Geographie der Universität Salzburg. 17-26.

RUSNOV, P. (1910): Über die Feststellung von Rauchschäden im Nadelwald. Centralblatt für das gesamte Forstwesen 36, 257-268.

RUSNOV, P. (1914): Über den heutigen Stand der Rauchschadensforschung. Centralblatt für das gesamte Forstwesen 43, 335-344.

SAUTER, J.J., KAMMERBAUER, H., PAMBOR, L. and HOCK, B. (1987): Evidence fort he accelerated micromorphological degradation of epistomatal waxes in Norway spruce by motor vehicle emissions. European Journal of Forest Pathology 17, pp. 444-448.

SCHAFMEISTER, M.-T. (1999): Geostatistik für die hydrologische Praxis. Springer-Verlag, Berlin-Heidelberg-New York, 172 S.

SCHERELIS, G. und BLÜMEL, W.D. (1988): Geostatistik und ihre Anwendungsperspektiven in der Geoökologie am Beispiel des Kriging-Verfahrens. Karlsruher Manuskripte zur Mathematischen und Theoretischen Wirtschafts- und Sozialgeographie, 19 S.

SCHLITTGEN, R. und STREITBERG, B. H. J. (2001): Zeitreihenanalyse. Oldenbourg Verlag, München-Wien, 9.Auflage, 571 S.

SCHMITZ, B. (1989): Einführung in die Zeitreihenanalyse. Verlag Hans Huber, Bern, 235 S.

SCHNEIDER, P. (2003): Validierung und Skalierung von SRTM-Geländemodellen für Anwendungen in der hydrologischen Modellierung. Diplomarbeit Ludwig-Maximilians-Universität München, 147 S.

SCHROEDER, J.v. und REUSS, J. (1883): Die Beschädigung der Vegetation durch Rauch und die Oberharzer Hüttenrauchschäden. Mit 5 Tafeln und 2 Karten. Parey-Verlag, Berlin, 333 S.

SCHÜTT, P., BLASCHKE, H., HOLDENRIEDER, O. und KOCH, W. (1984): Der Wald stirbt an Stress. Bertelsmann Verlag, München, 264 S.

SCHWEIKLE, V. (1997): Bodenversauerung – Ursachen, Auswirkungen, Maßnahmen. Landesanstalt für Umweltschutz Baden-Württemberg, 28 S.

SHEPARD, D. (1968): A two-dimensional interpolation function for irregularly-spaced data. Proceedings ACM National Conference, pp. 517-524.

SINOWSKI, W. and AUERSWALD, K. (1999): Using relief parameters in a discriminant analysis to stratify geological areas with different spatial variability of soil properties. Geoderma 89, pp. 113-128.

SMIDT, S. (2004): Waldschädigende Luftverunreinigungen. Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft (Version 5.4.2004), 183 S.

SMIDT, S., FRIEDL, H., Fürst, A. Breitenbach, M. und Plattner, J. (2004): Waldschädigende Luftverunreinigungen. Praxisinformation des Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft, Nr. 5 - 2004, 3-6.

SODOUDI, S. (2004): Verifikation der EZMW-Niederschlagsprognose für den Iran. Dissertation, Freie Universität Berlin, 205 S.

STORM, R. (1995): Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle. Fachbuchverlag Leipzig, 9.Auflage, 424 S.

STOYAN, D., STOYAN, H. und JANSEN, U. (1997): Umweltstatistik – Statistische Verarbeitung und Analyse von Umweltdaten. B.G. Teubner Verlagsgesellschaft, Stuttgart-Leipzig, 348 S.

TOBLER, W. (1970): A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography, Vol. 46 (2), pp. 234-240.

TRIMBACHER, C. (1997): Fichtennadeluntersuchungen in der Steiermark – Rasterelektronenmikroskopische und chemisch-analytische Ergebnisse. Umweltbundesamt Wien, UBA-BE-082, 37 S.

TRIMBACHER, C. und DITRICH, H. (1989): Epicuticuläre Wachsstrukturen von Fichtennadeln als Indikator für Langzeitimmissionsbelastungen. Umweltbundesamt Wien, UBA-Report-89-039, 114 S.

TUKEY, J. W. (1977): Exploratory Data Analysis. Publisher Addison-Wesley, Massachusetts, 710 p.

WACKERNAGEL, H. (2000): Multivariate Geostatistics – An Introduction with Applications. Springer Verlag, Berlin-Heidelberg, 1st Edition, 312 p.

WERTHER, T. (2003): Optimal Interpolation in Semi-Hilbert Spaces. Dissertation, Universität Wien, 175 p.

ZENGER, A. (1998): Atmosphärische Ausbreitungsmodellierung - Grundlagen und Praxis. Springer Verlag, Berlin-Heidelberg, 159 S.

8 Quellenverzeichnis

BAMLER, R.: Digital Terrain Models from Radar Interferometry (1997). http://ww.ifp.uni-stuttgart.de/publications/phowo97/bamler.pdf (2005-12-29)

BLASCHKE, T. und STROBL, J.: Räumliche Interpolation (2003). http://elearn.sbg.ac.at/courses/1/03J898008/content/_56440_1/03J898008_L08_Int (2005-08-05)

COLLINS, F.C. and BOLSTAD, P.V.: A Comparison of Spatial Interpolation Techniques in Temperature Estimation (2000). http://www.uni-salzburg.at/geo/idrisi/gis_environmental_modeling/sf_papers/collins_fred/collins.html (2005-08-01)

DUMFARTH, E. und LORUP, E.J.: Geostatistik I – Theorie und Praxis (2000). http://www.geo.sbg.at/staff/lorup/lv/geostats2000 (2005-06-05)

FÜRST, A.: Österreichisches Bioindikatornetz (2005). http://bfw.ac.at/600/1004.html (2005-06-17)

HENNIG, C.: Modellwahl und Variablenselektion in der Statistik (2004). http://www.math.uni-hamburg.de/home/hennig/lehre/modelselect.html (2006-01-27)

KRYWKOW, J.: Kriging, Geostatistik – Eine Einführung (2000). http://www.geog.fu-berlin.de\~jkrywkow/gis2/sem9/kriging.html (2005-08-14)

LORUP, E.J.: Explorative räumliche Datenanalyse – Teil B (2004). https://elearn.sbg.ac.at/webapps/portal/frameset.jsp?tab=courses&url=/bin/common/course.pl?course_id=_1143_1 (2005-07-12)

LÜTZKENDORF, S.: SAS – Statistische Maßzahlen (2001). http://www.informatik.hu-berlin.de/~luetzken/sas/sas-smas.html (2005-07-23)

STREIT, U.: Einführung in die Geostatistik 2.1 (2001). http://castafiore.uni-muenster.de/vorlesungen /Geostatistik/index.html (2004-10-12)

WALDOW, H.v.: Einführung in die Geostatistik (1998). http://www.geog.fu-berlin.de/~jkrywkow/harald/geostatistik/referat.html (2005-08-04)

9 Glossar

A

Autokorrelation	Die Autokorrelations-Funktionen beschreiben die Ähnlichkeit eines Signals bzw. einer Zeitfunktion mit sich selbst.
Anisotropie	Richtungsabhängigkeit
Abiotisch	Als abiotisch werden alle Umweltfaktoren zusammengefasst an denen Lebewesen nicht erkennbar beteiligt sind. Dazu gehören z.B. der geologische Untergrund oder das Klima.

B

Bioindikator	Bioindikatoren sind Organismen, die auf einen äußeren Reiz und im Speziellen auf eine Schadstoffbelastung mit Stoffwechseländerung reagieren, wobei auch eine Schadstoffanreicherung, die über das normale Maß hinausgeht, als Reaktion zu betrachten ist. Als Bioindikatoren eigenen sich am besten Pflanzen, weil sie standortsgebunden sind und meist in großer Individuenzahl natürlich vorkommen.
Biomonitoring	Biomonitoring umfasst alle Monitoring-Systeme, die den Zustand und die Veränderung der Umwelt anhand der resultierenden Auswirkungen auf einen lebenden Organismus bewerten
Biotisch	Als biotisch werden alle Umweltfaktoren zusammengefasst, an denen Lebewesen erkennbar beteiligt sind. Sie ergeben sich aus den Wechselwirkungen zwischen einzelnen Arten innerhalb eines Ökosystems.

C

Cokriging	Das Cokriging ist ein der multivariaten Situation angepasstes Kriging-Verfahren. D.h. zur Vorhersage eines Merkmals in einem Punkt werden nicht nur die Messwerte dieses Merkmals, sondern auch Messwerte anderer Größen benutzt.

D

Data Mining	Unter „Data Mining“ versteht man das systematische (in der Regel automatisierte oder halbautomatische) Entdecken oder Extrahieren unbekannter Information aus großen Mengen von Daten.

E

Emission	Jede Art der Abgabe von Stoffen, Energien und Strahlen an die Umgebung durch eine bestimmte Quelle wird als Emission bezeichnet. Häufig handelt es sich dabei um die Abgabe von Schadstoffen bzw. Schadenergien.
Exposition	Die Exposition die die Lage eines Berghanges im Bezug auf die Himmelsrichtung an.

F

Forstschädliche Luftverunreinigungen	Forstschädliche Luftverunreinigungen im Sinne des Forstgesetzes sind Luftverunreinigungen, die messbare Schäden am Waldboden, oder Bewuchs (Gefährdung der Waldkultur) verursachen.

G

Geostatistik	Fachbereich der Statistik, der sich mit der Analyse von Daten mit räumlichem Bezug beschäftigt.

H

Herrschende Bäume	In der Regel den Hauptbestand bildende Bäume mit verhältnismäßig gut entwickelten Kronen.
Histogramm	Ein Histogramm ist ein grafisches Mittel zur Darstellung der Häufigkeitsverteilung eines Merkmals.

I

Immission	Ist die Einwirkung der emittierenden Schadstoffe auf Pflanzen, Tiere und Menschen sowie Gebäude, nachdem sie sich in der Luft, dem Wasser oder dem Boden ausgebreitet oder auch chemisch oder physikalisch umgewandelt haben.
Infrarot	Infrarot oder Wärmestrahlung gehört zur elektromagnetischen Strahlung und liegt im Spektrum zwischen sichtbaren Licht und Mikrowellenstrahlung.

Intrinsische Hypothese Eine Form von räumlicher Stationarität, die weniger restriktiv als die Stationarität zweiter Ordnung ist und bei der die Bedingungen für die Ortsfestheit auf die ersten Unterschiede und nicht die zugrunde liegende regionalisierte Variable beschränkt ist. Die intrinsische Hypothese ist nützlich zur Modellierung regionalisierter Variablen, bei denen die Form des Variogramms eine Form des Geltungsbereiches ist.

Isotrop Räumliche Abhängigkeit ist in allen Richtungen gleich

Isoplethe Linien die Punkte mit gleichen Werteausprägungen verbindet z.B. Isohypse (Höhenlinien)

K

Korrelation Die Korrelation ist eine Beziehung zwischen zwei oder mehr quantitativen statistischen Variablen. Es gibt eine positive (je mehr, desto mehr) und eine negative (je mehr, desto weniger) Korrelation.

Kreuzvalidierung Die Kreuzvalidierung ("engl. Cross Validation") ist ein Testverfahren zur Messung der Qualität von Algorithmen und Methoden z.B. aus dem Data Mining.

Kriging Nach Danie KRIGE benanntes Verfahren bzw. Gruppe von Verfahren zur Interpolation von im Raum verteilter Variablen.

Kurtosis Parameter, der die Form (Höhe) der Verteilungsfunktion (Summenfunktion) beschreibt. Sie ist definiert als das „standardisierte zentrale Moment 4. Ordnung". Verteilungen mit stark ausgeprägten Rändern werden „leptokurtic" genannt, solche mit wenig ausgeprägten Rändern heißen „platykurtic". Eine Verteilung welche dieselbe Kurtosis wie die Normalverteilung aufweist, wird „mesokurtic" genannt.

M

Median Der Median ist der mittlere Messwert einer Verteilung: die eine Hälfte der Messwerte ist größer als der Median, die andere Hälfte ist kleiner. Der Median reagiert, anders als das arithmetische Mittel, nicht auf Extremwerte

N

Neuartige Waldschäden	Schlagwort für das auch als Waldsterben bezeichnete Phänomen von großflächigen Schädigungen des Waldes durch z.B. sauren Regen, Ozon, Schwermetalle.

Q

Quartile	Die drei Trennpunkte, die eine Häufigkeitsverteilung in vier gleiche Teile zerlegen. D.h. sie sind die Quantile für p=0,25, p=0,50 und p=0,75.

R

Räumliche Interpolation	Berechnung von Schätzwerten für Punkte, die zwischen den Stützpunkten, deren Wert bekannt ist, liegen.
Regionalisierte Variable	Variable mit Raumbezug, die aufgefasst wird als Mischform einer deterministischen und einer rein zufälligen Variable.
Residuum	Differenz zwischen beobachteten und geschätzten Wert eines Parameters

S

Schwefeldioxid	Schwefeldioxid ist ein farbloses Gas, das 2,3fach schwerer ist Luft. Es weist einen stechenden Geruch auf, der ab einer Konzentration von ca. 1,3mg/m3 wahrgenommen wird.
Stickoxide	Ist eine Sammelbezeichnung für die gasförmigen Oxide des Stickstoffs. Sie werden mit NOx abgekürzt, da es auf Grund der vielen Oxidationsstufen des Stickstoffs mehrere Stickstoff-Sauerstoff-Verbindungen gibt.
Schiefe	Parameter der die Form (die Neigung) der Verteilungsfunktion (Summenfunktion) beschreibt. Die Schiefe ist definiert als das „standardisierte zentrale Moment 3.Ordnung“
Synoptische Strömung	Luftströmung aufgrund des Wettergeschehens

T

Trendoberflächen	Trendoberflächen (auch: polynominale Schätzungen) sind in mathematischen Sinne die Fortsetzung des Konzeptes der Regressionskurven. Mit dem Unterschied, dass wir nun keine Linie anpassen, sondern eine Fläche.

V

Variogramm	Quantitative, häufig auch graphische, Darstellung der Variabilität von regionalisierten Variablen in Abhängigkeit von deren räumlicher Distanz.
Variographie	Die Variographie ist Werkzeug zur explorativen Analyse der räumlichen Kontinuität.
Vorherrschende Bäume	Bäume mit kräftig entwickelten Kronen

W

Waldsterben	Unter dem Begriff Waldsterben wird üblicherweise das Absterben ganzer Wälder aufgrund einer einzigen oder komplexen Krankheitsursache verstanden.