

TEXTE

30/2025

Abschlussbericht

Aufarbeitung und Analyse von Luftqualitätsdaten der Deutschen Demokratischen Republik

von:

Jörg Schmidtke, Paul Schmidt, Kerstin Schmidt
BioMath GmbH, Hamburg/ Rostock

Herausgeber:

Umweltbundesamt

TEXTE 30/2025

REFOPLAN des Bundesministeriums Umwelt,
Naturschutz, nukleare Sicherheit und Verbraucherschutz

Forschungskennzahl 3722 52 203 0

FB001554

Abschlussbericht

Aufarbeitung und Analyse von Luftqualitätsdaten der Deutschen Demokratischen Republik

von

Jörg Schmidtke, Paul Schmidt, Kerstin Schmidt
BioMath GmbH, Hamburg/ Rostock

Im Auftrag des Umweltbundesamtes

Impressum

Herausgeber

Umweltbundesamt
Wörlitzer Platz 1
06844 Dessau-Roßlau
Tel: +49 340-2103-0
Fax: +49 340-2103-2285
buergerservice@uba.de
Internet: www.umweltbundesamt.de

Durchführung der Studie:

BioMath GmbH
Eppendorfer Weg 255
20251 Hamburg

Abschlussdatum:

August 2024

Redaktion:

Fachgebiet II 4.2 Beurteilung der Luftqualität
Andrea Minkos, Susan Kessinger

Publikationen als pdf:

<http://www.umweltbundesamt.de/publikationen>

ISSN 1862-4804

Dessau-Roßlau, Februar 2025

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

Danksagung

Wir danken allen Personen, die über ihre Arbeit im Meteorologischen Dienst der DDR oder in den Bezirkshygieneinstituten an der Erhebung der Messwerte und ihrer Qualitätssicherung beteiligt waren. Unser Dank gilt außerdem denjenigen, die in der Nachwendezeit geistesgegenwärtig und eigenverantwortlich die Daten aus dem Datenspeicher auslasen und somit dem UBA zur Verfügung stellen konnten, sowie allen Mitarbeiterinnen und Mitarbeitern in den Immissionsmessnetzen der Länder, die mit Informationen zur Verarbeitung der Daten beigetragen haben.

Kurzbeschreibung: Aufarbeitung und Analyse von Luftqualitätsdaten der Deutschen Demokratischen Republik

Messwerte der Luftqualitätsparameter Blei, Chlor, Fluor, Kohlenmonoxid, Ozon, Schwebstaub, Schwefeldioxid, Schwefelkohlenstoff, Schwefelwasserstoff, Stickoxide, Stickstoffdioxid, Stickstoffmonoxid und Sulfat aus der Deutschen Demokratischen Republik (DDR) für den Zeitraum von 1969 bis 1990 wurden aufgearbeitet. Die Aufarbeitung umfasste die Extraktion der Messwerte aus den Rohdaten mit anschließender Klassifikation und Plausibilitätskontrolle. Das Ergebnis sind qualitätsgesicherte Daten der Luftqualitätsmessungen der DDR.

Abstract: Processing and analysing air quality data of the German Democratic Republic

Measured values of the air quality parameters carbon disulphide, carbon monoxide, chlorine, fluorine, hydrogen sulphide, lead, nitrogen dioxide, nitrogen monoxide, nitrogen oxides, ozone, sulphate, sulphur dioxide and suspended particulates from the German Democratic Republic (GDR) for the period from 1969 to 1990 were processed. The processing included the extraction of the measured values from the raw data with subsequent classification and plausibility checks. The result is quality-assured data on air quality measurements in the GDR.

Inhaltsverzeichnis

Abbildungsverzeichnis.....	9
Tabellenverzeichnis.....	9
Abkürzungsverzeichnis.....	11
Zusammenfassung.....	12
Summary.....	16
1 Hintergrund und Problemstellung.....	20
2 Zielstellung.....	21
3 Literaturrecherche.....	22
4 Datenmanagement.....	24
4.1 Datensichtung und Strukturierung.....	24
4.1.1 Übermittlung der Rohdaten.....	24
4.1.2 Strukturierung der Rohdaten.....	26
4.2 Standardisierung der Rohdaten.....	27
4.2.1 ASCII Typ-A.....	27
4.2.2 ASCII Typ-B.....	31
4.2.3 ASCII Typ-C.....	36
4.2.4 ASCII Typ-D.....	38
4.2.5 ASCII Typ-E.....	39
4.2.6 ASCII Typ-F.....	41
4.2.7 ASCII Typ-G.....	42
4.2.8 IBM Lotus 1-2-3 oder Borland Quattro Pro.....	44
4.2.9 Microsoft Excel.....	47
4.2.10 Metadaten der Stationen.....	49
4.2.11 Ergebnisse der Standardisierung.....	50
4.3 Harmonisierung der standardisierten Rohdaten von SO ₂	51
4.3.1 SO ₂ -Jahreswerte.....	53
4.3.2 SO ₂ -Monatswerte.....	55
4.3.3 SO ₂ -Tageswerte.....	57
4.3.4 SO ₂ -Messwerte zu Zeitpunkten.....	59
4.3.5 Zeitliche Auflösungen im Zusammenhang.....	62
4.3.6 Ergebnis der Harmonisierung.....	63
4.4 Plausibilitätsprüfung.....	64
4.4.1 Datenbasis.....	64

4.4.2	Plausibilität.....	65
A	Anhang	69
A.1	Stationen.....	69
A.2	Anlagedateien	81
5	Quellenverzeichnis	82

Abbildungsverzeichnis

Abbildung 1:	Anzahl Dateien je Jahr der letzten Änderung.....	13
Abbildung 2:	Anzahl Stationen je Parameter und Jahr	15
Figure 3:	Number of files per year of last change	17
Figure 4:	Number of stations per parameter and year	19
Abbildung 5:	Inhalt ASCII Typ-A	29
Abbildung 6:	Inhalt ASCII Typ-B	33
Abbildung 7:	Inhalt ASCII Typ-C	36
Abbildung 8:	Inhalt ASCII Typ-D	38
Abbildung 9:	Inhalt ASCII Typ-E.....	40
Abbildung 10:	Inhalt ASCII Typ-F.....	42
Abbildung 11:	Inhalt ASCII Typ-G	43
Abbildung 12:	Inhalt Datei vom Typ IBM Lotus 1-2-3	46
Abbildung 13:	Zusammenhang der SO ₂ -Jahreswerte bezüglich der Datenquellen	54
Abbildung 14:	Zusammenhang der SO ₂ -Monatswerte bezüglich der Datenquellen	56
Abbildung 15:	Zusammenhang der SO ₂ -Tageswerte bezüglich der Datenquellen	58
Abbildung 16:	Zusammenhang der SO ₂ -Messwerte zu Zeitpunkten bezüglich der Datenquellen.....	61
Abbildung 17:	Zusammenhang der SO ₂ -Tagesmittelwerte bezüglich der schaltjahrrelevanten Datenquellen	62
Abbildung 18:	Datenquellen der SO ₂ -Messwerte.....	63
Abbildung 19:	Zusammenhang der SO ₂ -Messwerte bezüglich der Datenquellen	63

Tabellenverzeichnis

Tabelle 1:	Metadaten.....	24
Tabelle 2:	Anzahl Rohdatendateien der Messwerte in der Stammverzeichnisstruktur.....	25
Tabelle 3:	Klassifizierung der Rohdatendateien der Messwerte	26
Tabelle 4:	Verzeichnisstruktur der Rohdatendateien vom ASCII Typ-A....	28
Tabelle 5:	Extrahierte Variablen der Rohdatendateien vom ASCII Typ-A.	29
Tabelle 6:	Verzeichnisstruktur der Rohdatendateien vom ASCII Typ-B....	31
Tabelle 7:	Aufbau der Rohdatendateien vom ASCII Typ-B.....	33
Tabelle 8:	Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-B.....	34
Tabelle 9:	Stationsnummern in den Rohdatendateien vom ASCII Typ-B, die nicht in den Metadaten vorhanden sind	35
Tabelle 10:	Aufbau der Rohdatendateien vom ASCII Typ-C.....	36

Tabelle 11:	Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-C.....	37
Tabelle 12:	Aufbau der Rohdatendateien vom ASCII Typ-D	38
Tabelle 13:	Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-D	39
Tabelle 14:	Extrahierte Variablen der Rohdatendateien vom ASCII Typ-E .	40
Tabelle 15:	Aufbau des Datenblocks der Rohdatendateien vom ASCII Typ-G	43
Tabelle 16:	Extrahierte Variablen der Rohdatendateien vom ASCII Typ-G.	44
Tabelle 17:	Verzeichnisstruktur der Rohdatendateien vom Typ IBM Lotus 1-2-3 oder Borland Quattro Pro	44
Tabelle 18:	Inhalt und Bedeutung der ersten sechs Byte von Dateien mit der Extension wk1 im Dateinamen.....	45
Tabelle 19:	Aufbau der Dateien des Typs IBM Lotus 1-2-3	46
Tabelle 20:	Extrahierte und zusätzlich ermittelte Variablen der Rohdatendateien vom Typ IBM Lotus 1-2-3.....	47
Tabelle 21:	Klassifizierung der Microsoft Exceldateien.....	48
Tabelle 22:	Extrahierte und zusätzlich ermittelte Variablen der Rohdatendateien der Typen A bis G der Microsoft Excel 97-2003-Arbeitsmappen.....	48
Tabelle 23:	Datei mit Stationen aus den Metadaten	50
Tabelle 24:	Datei mit den Abständen zwischen den Stationen	50
Tabelle 25:	Dateitypen mit SO ₂ -Werten unterschiedlicher zeitlicher Auflösungen.....	51
Tabelle 26:	Datei mit SO ₂ -Messwerten	52
Tabelle 27:	Datenquellen der SO ₂ -Jahreswerte	53
Tabelle 28:	Anzahl gleicher und ungleicher SO ₂ -Jahreswerte bezüglich der Datenquellen	53
Tabelle 29:	Datenquellen der SO ₂ -Monatswerte	55
Tabelle 30:	Anzahl gleicher und ungleicher SO ₂ -Monatswerte bezüglich der Datenquellen	55
Tabelle 31:	Datenquellen der SO ₂ -Tageswerte	57
Tabelle 32:	Anzahl gleicher und ungleicher SO ₂ -Tageswerte bezüglich der Datenquellen	57
Tabelle 33:	Datenquellen der SO ₂ -Messwerte zu Zeitpunkten.....	59
Tabelle 34:	Anzahl gleicher und ungleicher SO ₂ -Messwerte zu Zeitpunkten bezüglich der Datenquellen.....	59
Tabelle 35:	Schadstoffe und zeitliche Auflösung für die Plausibilitätsprüfung.....	64
Tabelle 36:	Q _{CutOff} je Schadstoff und zeitlicher Auflösung.....	66
Tabelle 37:	Finale Ergebnisdateien	67

Abkürzungsverzeichnis

BHI	Bezirkshygieneinstitut
CL	Chlor
CO	Kohlenmonoxid
CS₂	Schwefelkohlenstoff
DDR	Deutsche Demokratische Republik
DWD	Deutscher Wetterdienst
FL	Fluor
GIS	Geoinformationssystem
H₂S	Schwefelwasserstoff
MD	Meteorologischer Dienst
NO	Stickstoffmonoxid
NO₂	Stickstoffdioxid
NO_x	Stickoxide
O₃	Ozon
PB	Blei
RE	Registrierung eines gemessenen Schadstoffs
SO₂	Schwefeldioxid
SO₄	Sulfat
SP	Tages-Spitzenwert der Registrierung eines gemessenen Schadstoffs
SPSS	Software von IBM für die statistische Datenanalyse
TE	Terminwert eines gemessenen Schadstoffs
TG	Tageswert eines gemessenen Schadstoffs
TSP oder SWS	Schwebstaub (Total suspended particles)
UBA	Umweltbundesamt

Zusammenfassung

Ziel des Projektes war es, Messwerte der Luftqualitätsparameter Blei, Chlor, Fluor, Kohlenmonoxid, Ozon, Schwebstaub, Schwefeldioxid, Schwefelkohlenstoff, Schwefelwasserstoff, Stickoxide, Stickstoffdioxid, Stickstoffmonoxid und Sulfat aus der Deutschen Demokratischen Republik (DDR) für den Zeitraum von 1969 bis 1990 aufzuarbeiten und als qualitätsgeprüfte Messwerte bereitzustellen.

Zunächst wurden mit einer Literaturrecherche Informationen zu angewandten Messmethoden und zu Plausibilitätsprüfungen nach den damals geltenden technischen Regeln zusammengestellt.

Die Aufarbeitung der Messwerte der Luftqualitätsparameter und deren Plausibilitätsprüfung erfolgte in den vier Schritten

- ▶ Strukturierung der Rohdaten
- ▶ Standardisierung der Datenformate
- ▶ Prüfung und Harmonisierung der Messwerte verschiedener Datenquellen und
- ▶ Prüfung der Plausibilität.

Strukturierung der Rohdaten

Das UBA stellte 1.943 Dateien mit Messwerten der Luftqualitätsparameter und 2 Dateien mit Informationen zu den Messwerten und den Stationen bereit.

Die Dateien der Messwerte wurden in einer Liste zusammengestellt und um relevante Dateiattribute (z. B. Dateierweiterung, Datum der letzten Änderung, usw.) ergänzt. Danach wurde jede Datei nach ihrer Art in *Textdatei* oder *Binärdatei* eingeteilt. Die binären Dateien wurden anschließend weiter in *Lotusdatei*¹, *Exceldatei* oder *sonstige Datei* klassifiziert (siehe Abbildung 1). Drei der sonstigen Binärdateien wurden Microsoft Write, eine zu einem CD-Brennprogramm zugeordnet.

Grundlage für die Klassifizierung waren einerseits die Namen der Dateierweiterungen und andererseits die ersten 6 Bytes der Dateien. Mit Hilfe der ersten Bytes wurde festgestellt, zu welchen Anwendungsprogrammen die Dateien gehören. Die Binärdateien gehörten zu den Programmen

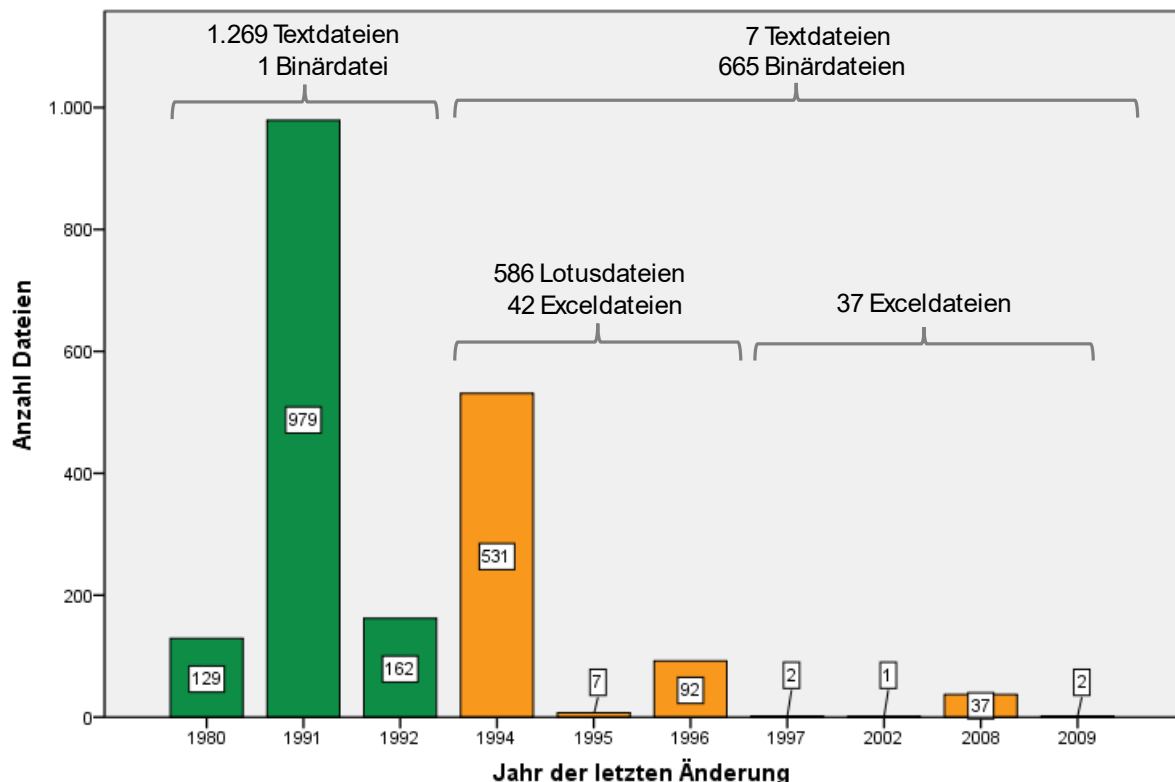
- ▶ IBM Lotus 1-2-3 Version 2.01
- ▶ Microsoft Excel 4.0
- ▶ Microsoft Excel 97 und
- ▶ Microsoft Excel 2000-2003.

Es wurde festgestellt, dass es zwei zeitliche Cluster bezüglich der letzten Änderung der Dateien gab. Im Zeitraum bis 1992 wurden Textdateien erstellt, im Zeitraum von 1994 bis 1996 wurde mit Lotusdateien und einigen wenigen Exceldateien gearbeitet. Aus dieser Betrachtung wurde vermutet, dass es sich im Zeitraum bis 1992 um Dateien handelt, die aus dem Datenspeicher der DDR erstellt wurden und die Dateien von 1994 bis 1996 zum Vorhaben Dresden (Fleischer &

¹ Das Programm Lotus 1-2-3 ist eine Tabellenkalkulationssoftware aus der Zeit von 1983 bis 2003.

Pohl, 1995) gehören. Auch die Namen der Verzeichnisstruktur für die Dateien wiesen auf dieses Vorhaben hin.

Abbildung 1: Anzahl Dateien je Jahr der letzten Änderung



Dateien mit dem Zeitstempel 1980 wurden auf PCs bearbeitet, deren Systemuhr nicht korrekt arbeitete.

Quelle: eigene Darstellung, BioMath

Standardisierung der Datenformate

Nach der Dateiklassifizierung erfolgte die Einteilung der Dateien nach inhaltlichen Gesichtspunkten in die folgenden Typen

- ▶ 7 Typen Textdateien (ASCII Typ-A bis ASCII Typ-G)
- ▶ 1 Typ IBM Lotus 1-2-3 und
- ▶ 7 Typen Exceldateien (Excel Typ-A bis Excel Typ-G).

Dazu wurden von allen Textdateien die ersten 10 Zeilen und von allen Binärdateien die ersten 10 Byteblöcke mit den abschließenden Steuerzeichen $OD_{Hex}0A_{Hex}^2$ sowie deren Längen ermittelt. Mit diesen Informationen und manuellen Sichtungen der Dateiinhalte wurden dann die Typisierungen durchgeführt.

Anschließend wurden alle Dateien eines Typs in einem Standardformat zusammengefasst. Während dieser Zusammenfassung wurden diverse physische und auch inhaltliche Fehler beseitigt. Es wurden z. B. Steuerzeichen aus Textdateien entfernt, unvollständige Textzeilen entfernt oder ggf. mit Fehlwerten aufgefüllt, vertauschte Ziffern in den Stationsnummern korrigiert, usw. Hilfreich bei der Fehlerbeseitigung waren auch die Attribute der Dateien und die Bezeichnungen der Komponenten der Dateipfade. Aus Ihnen konnte oftmals auf den

² $OD_{Hex}0A_{Hex}$ sind die beiden Steuerzeichen, mit denen in Textdateien der Zeilenumbruch gekennzeichnet ist.

Luftschadstoff sowie das Jahr und den Zeitraum der Messungen geschlossen werden. Weiterhin wurden über Häufigkeits- und Kreuztabellen die Anzahlen der Messwerte auf Korrektheit überprüft und - wenn sachlogisch möglich - Korrekturen vorgenommen.

Im Ergebnis der Standardisierung lagen, abhängig vom Luftqualitätsparameter, die Messwerte als

- ▶ Stunden- oder Halbstundenwerte
- ▶ Tageswerte (Messwerte, die über einen gesamten Tag ermittelt wurden)
- ▶ Terminwerte (Messwerte die zu bestimmten Zeitpunkten ermittelt wurden)
- ▶ Tagesmittelwerte
- ▶ Monatsmittelwerte und
- ▶ Jahresmittelwerte

vor.

Prüfung und Harmonisierung der Messwerte verschiedener Datenquellen

Der überwiegende Teil der Messwerte bezog sich auf den Luftschadstoff SO₂ und lag in 13 verschiedenen Dateitypen vor. Gleiche und ähnliche Messwerte sowie berechnete Mittelwerte von SO₂ traten in diesen Dateien mehrfach auf. Die Messwerte aller anderen Luftschadstoffe waren über 3 verschiedene Dateitypen verteilt, ohne mehrfach aufzutreten.

Die SO₂-Messreihen mussten so harmonisiert werden, dass je Station und Zeitpunkt nur ein eindeutiger Messwert vorhanden war.

Dazu wurden die Messwerte von SO₂ bezüglich des Auftretens in den verschiedenen Dateitypen geprüft. Innerhalb dieser Prüfung wurde festgestellt, dass ein großer Teil der Stunden- und Halbstundenmesswerte in Sprüngen von 5 oder 10 vorlagen, und dass Mittelwerte ohne Nachkommastellen immer abgerundet wurden. Weiterhin wurden Tagesmittelwerte in den Schaltjahren ab dem Monat März den falschen Tagen zugeordnet. Die Prüfung umfasste nicht nur den Vergleich der Messwerte über die verschiedenen Dateitypen, sondern auch die Beurteilung des Datums der letzten Änderungen der Dateien der Messwerte.

Das Ergebnis der Prüfung war, dass sich alle SO₂-Messwerte auf die Werte in den Dateien vom ASCII Typ-A, ASCII Typ-B und ASCII Typ-D aus dem Zeitraum bis 1990 zurückführen lassen. In der anschließenden Harmonisierung wurden dann standardisierte Dateien für

- ▶ Stunden- oder Halbstundenwerte
- ▶ Tageswerte und
- ▶ Terminwerte

von SO₂ erstellt.

Prüfung der Plausibilität

Ein Teil der Plausibilitätsprüfung erfolgte bereits während der Standardisierung und Harmonisierung.

Für alle Luftqualitätsparameter wurde die untere Grenze der Messwerte mit Null festgelegt. Negative Messwerte traten allerdings nur als Codierung von Fehlwerten auf. Eine obere Grenze

der Messwerte konnte nicht definiert werden, auch nicht für SO₂. Für die Plausibilitätsprüfung von Nullwerten lagen keine belastbaren Informationen vor, sodass Nullwerte grundsätzlich als plausibel angesehen wurden.

Für die Detektion von Werten, die so extrem sind, dass sie sehr wahrscheinlich nicht in die Messreihe gehören, wurde ein Verfahren auf Basis von 95 %- oder 99 %-Perzentilen entwickelt. In diesem Verfahren wurden die Messwerte ins Verhältnis zu den monatlichen Perzentilen gesetzt und ein Toleranzwert bestimmt, ab dem ein Messwert als nichtplausibel eingestuft wird. Für jeden Luftschadstoff und jede zeitliche Auflösung wurden sowohl die Perzentile als auch die Toleranzwerte definiert. Diese Definitionen erfolgten konservativ, d. h. es wurde sich eher für die Plausibilität eines Messwertes entscheiden als dagegen.

Es wurden keine Messwerte verändert. Unplausible Messwerte wurden nachvollziehbar als Fehlwerte gekennzeichnet.

Im Ergebnis wurden je Schadstoff validierte Messreihen als Stunden- bzw. Halbstundenwerte, Tageswerte, Terminwerte sowie Tagesmittelwerte in Textdateien bereitgestellt (siehe Abbildung 2).

Abbildung 2: Anzahl Stationen je Parameter und Jahr

Parameter	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
CL-Tag										4	5	5	5									
CO-Zeit												1	2	1	1		2	3			3	4
CS ₂ -Tag										2	2	2	2									
FL-Tag										7	8	8	8									
H ₂ S-Tag										1	1		2									
H ₂ S-Zeit												2	2	1								
NO-Tag																						17
NO-Zeit																						5
NO ₂ -Tag										1	1	1	4	4	4	5	5	15	2	24	26	
NO ₂ -Termin														1	1	1		1		1	1	1
NO ₂ -Zeit																						5
NO _x -Tag																			1	1	3	6
NO _x -Zeit										1	1	1	1	1		1	1	1	1	1	1	1
O ₃ -Termin	6	6	5	5	5	5	5	5	5	6	5	5	3									
O ₃ -Zeit					1	2	2	2	2	3	3	3	6	8	7	9	9	9	9	9	9	1
PB-Tag										1	1	1	1									
SO ₂ -Tag					3	3	3	34	46	43	39	43	46	49	57	66	67	68	7	71	59	
SO ₂ -Tagesmittel	1	1	2	2	2	3	3	9	39	44	51	41	5	54	62	65	64	55	79	126	128	
SO ₂ -Termin	8	12	13	13	13	13	14	15	2	2	2	2	3	3		4		6	7	6	4	
SO ₂ -Zeit	1	1	2	2	2	3	3	9	39	44	51	41	5	54	62	65	64	55	79	126	112	
SO ₄ -Tag											4	4										
SWS-Tag	16	17	18	18	18	18	2	3	32	31		24	26	31	35	42	42	45	5	55	57	
SWS-Termin																						2
SWS-Zeit																				1	5	5

Zeit: Stunden- bzw. Halbstundenwerte, Tag: Tageswert, Termin: Terminwerte und Tagesmittel: Tagesmittelwert

Quelle: eigene Darstellung, BioMath

Summary

Aim of the project was to process measured values of the air quality parameters carbon disulphide, carbon monoxide, chlorine, fluorine, hydrogen sulphide, lead, nitrogen dioxide, nitrogen monoxide, nitrogen oxides, ozone, sulphate, sulphur dioxide and suspended particulates from the German Democratic Republic (GDR) for the period from 1969 to 1990, and to provide them as quality-checked measured values.

First, a literature search was carried out to compile information on the measurement methods used and on plausibility checks in accordance with the technical regulations in force at the time.

The processing of the measured values of the air quality parameters and their plausibility check was carried out in four steps

- ▶ Structuring of the raw data
- ▶ Standardization of the data formats
- ▶ Checking and harmonizing the measured values from different data sources and
- ▶ Checking the plausibility.

Structuring the raw data

The UBA provided 1,943 files with measured values of the air quality parameters and 2 files with information on the measured values and the stations.

The files containing the measured values were compiled in a list and supplemented with relevant file attributes (e.g. file extension, date of last change, etc.). Each file was categorized according to its type as a *text file* or *binary file*. The binary files further were classified into *Lotus file*³, *Excel file* or *other file* (see Figure 3). Three of the other binary files were assigned to Microsoft Write and one to a CD burning program.

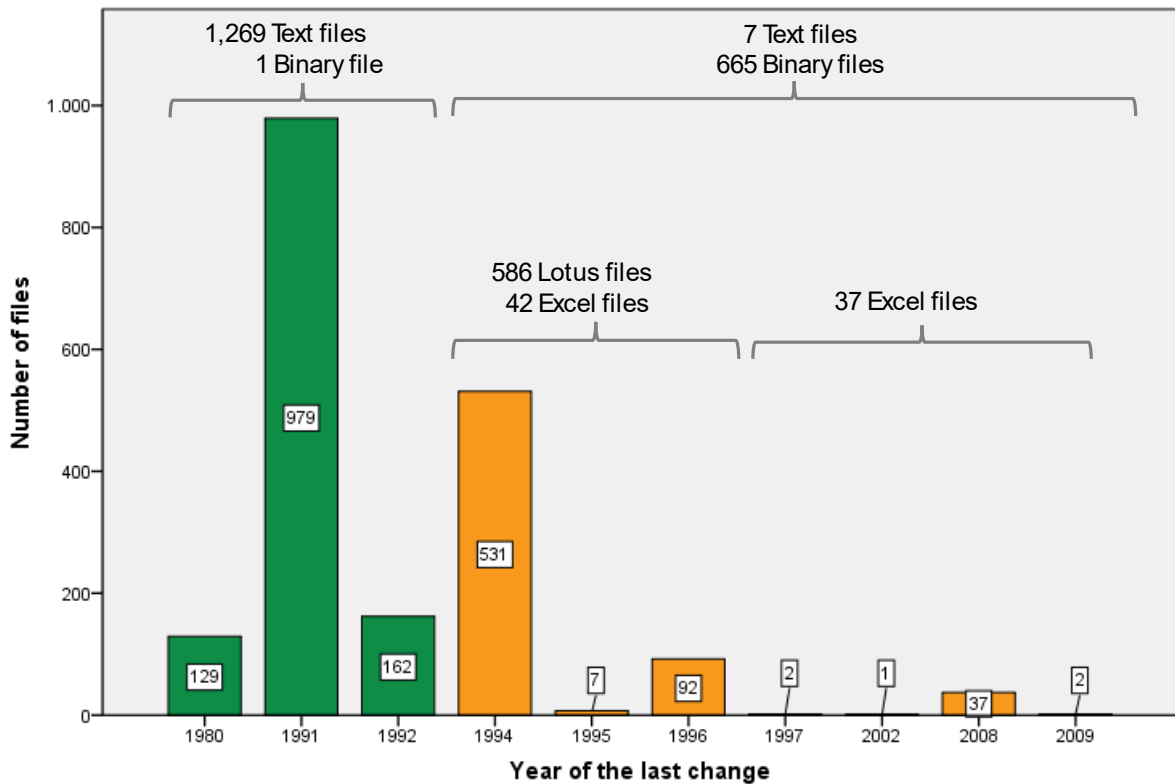
The classification was based on the names of the file extensions and the first 6 bytes of the files. The first bytes were used to determine which application programs the files belonged to. The binary files belonged to the programs

- ▶ IBM Lotus 1-2-3 Version 2.01
- ▶ Microsoft Excel 4.0
- ▶ Microsoft Excel 97 and
- ▶ Microsoft Excel 2000-2003.

It was determined that there were two temporal clusters with regard to the last change of the files. In the period up to 1992, text files were created, while in the period from 1994 to 1996, Lotus files and a few Excel files were used. From this observation it was assumed that in the period up to 1992 the files were created from the data storage of the GDR and that the files from 1994 to 1996 belonged to the Dresden project (Fleischer & Pohl, 1995). The names of the directory structure for the files also pointed to this project.

³ The Lotus 1-2-3 program is a spreadsheet software from the period 1983 to 2003.

Figure 3: Number of files per year of last change



Files with the timestamp 1980 were processed on PCs whose system clock was not working correctly.

Source: own illustration, BioMath

Standardization of the data formats

After file classification, the files were divided into the following types according to their content

- ▶ 7 types of text files (ASCII type-A to ASCII type-G)
- ▶ 1 type IBM Lotus 1-2-3 and
- ▶ 7 types of Excel files (Excel type-A to Excel type-G).

For this purpose, the first 10 lines of all text files and the first 10 bytes blocks of all binary files with the final control characters $OD_{Hex}0A_{Hex}^4$ and their lengths were determined. The typing was carried out using this information and manual inspection of the file contents.

All files of one type were combined into a standard format. During this consolidation, various physical and content-related errors were eliminated. For example, control characters were removed from text files, incomplete text lines were removed or filled with missing values, transposed digits in the station numbers were corrected, etc. Also, the attributes of the files and the names of the components of the file paths were helpful in eliminating errors since it was often possible to deduce the air pollutant as well as the year and period of the measurements. Furthermore, the numbers of measured values were checked for correctness using frequency and cross tables, and corrections were made where logically possible.

As a result of the standardization, depending on the air quality parameter, the measured values were available as

⁴ $OD_{Hex}0A_{Hex}$ are the two control characters used to indicate line breaks in text files.

- ▶ Hourly or half-hourly values
- ▶ Daily values (measured values that were determined over an entire day)
- ▶ Date values (measured values that were determined at specific times)
- ▶ Daily mean values
- ▶ Monthly mean values and
- ▶ Annual mean values

Examination and harmonization of measured values from different data sources

The majority of the measured values related to the air pollutant SO₂ and were available in 13 different file types. Identical and similar measured values as well as calculated mean values of SO₂ occurred several times in these files. The measured values of all other air pollutants were distributed across 3 different file types without occurring more than once.

The SO₂ measurement series had to be harmonised so that only one unique measured value was available for each station and time. For this purpose, the measured values of SO₂ were checked for duplicates in the different file types. This check revealed that a large proportion of the hourly and half-hourly measured values were available in jumps of 5 or 10, and that mean values without decimal places were always rounded down. Furthermore, in leap years daily averages from the month of March onwards were assigned to the wrong days. The check included not only the comparison of the measured values across the different file types, but also the assessment of the date of the last changes to the measured value files.

The result of the test was that all SO₂ measurement values can be traced back to the values in the ASCII Type-A, ASCII Type-B and ASCII Type-D files from the period up to 1990. In the subsequent harmonization, standardized files were created for

- ▶ Hourly or half-hourly values
- ▶ Daily values and
- ▶ Term values

of SO₂.

Plausibility check

Part of the plausibility check was already carried out during standardization and harmonization.

For all air quality parameters, the lower limit of the measured values was set at zero. However, negative measured values only occurred as coding of missing values. An upper limit of the measured values could not be defined, not even for SO₂. No reliable information was available for the plausibility check of zero values, so that zero values were generally regarded as plausible.

A method based on 95 % or 99 % percentiles was developed for the detection of values that are so extreme that they very probably do not belong in the measurement series. In this procedure, the measured values were set in relation to the monthly percentiles and a tolerance value was determined above which a measured value is classified as missing. Both the percentiles and the tolerance values were defined for each air pollutant and each temporal resolution. These definitions were conservative, i.e. a decision was made in favour of the plausibility of a measured value rather than against it.

No measured values were changed. Implausible measured values were clearly marked as missing values.

As a result, validated measurement series were provided in text files for each pollutant as hourly or half-hourly values, daily values, date values and daily mean values (see Figure 4).

Figure 4: Number of stations per parameter and year

Parameters	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	
Cl-Day										4	5	5	5										
CO-Time												1	2	1	1		2	3		3	4		
CS2-Day										2	2	2	2										
Fl-Day										7	8	8	8										
H2S-Day										1	1		2										
H2S-Time												2	2	1									
NO-Day																							17
NO-Time																							5
NO2-Day										1	1	1	4	4	4	5	5	15	2	24	26		
NO2-Date														1	1	1		1		1	1	1	
NO2-Time																							5
NOx-Day																			1	1	3	6	
NOx-Time										1	1	1	1	1		1	1	1	1	1	1	1	1
O3-Date	6	6	5	5	5	5	5	5	5	6	5	5	3										
O3-Time						1	2	2	2	3	3	3	6	8	7	9	9	9	9	9	9	9	1
PB-Day										1	1	1	1										
SO2-Day						3	3	3	34	46	43	39	43	46	49	57	66	67	68	7	71	59	
SO2-Daily average	1	1	2	2	2	2	3	3	9	39	44	51	41	5	54	62	65	64	55	79	126	128	
SO2-Date	8	12	13	13	13	13	13	14	15	2	2	2	2	3	3		4		6	7	6	4	
SO2-Time	1	1	2	2	2	2	3	3	9	39	44	51	41	5	54	62	65	64	55	79	126	112	
SO4-Day												4	4										
SWS-Day	16	17	18	18	18	18	18	2	3	32	31		24	26	31	35	42	42	45	5	55	57	
SWS-Date																							2
SWS-Time																					1	5	5

Time: hourly or half-hourly values, Day: daily value, Date: date values and Daily average: daily mean value

Source: own illustration, BioMath

1 Hintergrund und Problemstellung

In der DDR war Umweltverschmutzung im Allgemeinen und die Belastung der Luft im Besonderen allgegenwärtig. Die Verbrennung stark schwefelhaltiger Braunkohle und unzureichende Abgasminderungssysteme waren die Hauptursache für die hohen Schadstoffgehalte der Luft, die sich nicht nur negativ auf die Gesundheit der Menschen auswirkte, sondern auch in dem durch sauren Regen verursachtes Waldsterben in den Mittelgebirgen widerspiegelte. Die Ionenkonzentrationen im Niederschlag (wet-only-Sammler) wie z. B. Sulfationen war in den 1980er-Jahren überdurchschnittlich hoch im Vergleich zu Westdeutschland oder zu den 1990er-Jahren (Marquardt, Bruggemann & Ihle, 1996). Enorme Luftverschmutzung gab es damals nicht nur in der DDR, sondern auch in den angrenzenden Staaten Polen und Tschechoslowakei (Hůnová, 2020).

Die Luftqualität wurde von staatlicher Seite überwacht. Die 15 Bezirkshygieneinstitute waren zuständig für die Messung von Luftschadstoffen in den belasteten, quellnahen Gebieten der Ballungsräume und Städte. Die Luftqualität im quellfernen Raum wurde vom Meteorologischen Dienst der DDR gemessen. Die Beobachtungen waren hauptsächlich auf die durch Braunkohleverbrennung verursachten Schadstoffe Schwefeldioxid, Schwebstaub, Stickstoffoxide und Schwefelwasserstoff ausgerichtet und zweitrangig auf Kohlenmonoxid, Ozon und halogenierte Verbindungen (Werner & Meckel, 1989). Die gemessenen Daten wurden jedoch der Bevölkerung nicht bekannt gegeben.

Auch nach dem Ende der DDR wurden die Luftqualitätsdaten aus der DDR nicht veröffentlicht. Die Rohdaten sowohl der Bezirkshygieneinstitute als auch des Meteorologischen Dienstes zu den Konzentrationen mehrerer Schadstoffe wie SO_2 , TSP, NO_x und O_3 lagen am UBA vor. Diese Daten umfassten den Zeitraum von 1969 bis 1990 in verschiedenen zeitlichen Auflösungen (z. B. Stunden- oder Tageswerte). Stichprobenhafte Prüfungen ergaben, dass die Daten nicht validiert waren. Aufgrund unzureichender Kenntnisse über die damals angewandten Messmethoden und den jeweiligen Prüfroutinen konnten die Daten bisher nicht endgültig validiert werden.

Die Nachfrage seitens der Öffentlichkeit nach diesen Daten bzw. der daraus abzuleitenden Einschätzung der Luftqualität in der DDR und deren Änderung in den 1990er-Jahren bestand weiterhin. Das UBA wollte daher alle vorliegenden Rohdaten soweit möglich einer gründlichen Plausibilitätsprüfung unterziehen, um die Luftqualitätsdaten zu archivieren und zu veröffentlichen.

2 Zielstellung

Ziel des Projektes war die Bereitstellung einer Datenmenge mit qualitätsgesicherten Daten der Luftqualitätsmessungen der DDR, die in die Immissionsdatenbank am UBA eingelesen werden konnte. Dabei sollte eine gute Balance zwischen möglichst hoher Datenqualität und großem Datenumfang hergestellt werden.

Das Schließen von Datenlücken mittels statistischer Methoden war nicht Gegenstand des Projektes.

3 Literaturrecherche

Für die Zusammenstellung von Informationen, wie angewandte Messmethoden oder Plausibilitätsprüfungen und aktive Messstationen, wurde nach Literatur zu den damals geltenden technischen Regeln recherchiert. Das UBA hatte zum Auftakt des Projektes einige Publikationen bereitgestellt/ empfohlen.

Zum Verstehen und Aufarbeiten der Daten lag die folgende Literatur vor:

- ▶ Feister und Warmbt (1987). Long-term measurements of surface ozone in the German Democratic Republic. *Journal of Atmospheric Chemistry*, 5(1).
<https://doi.org/10.1007/BF00192500>
- ▶ Fleischer und Pohl (1995). Aufbereitung der Immissionsmeßreihen des ehemaligen Meteorologischen Dienstes der DDR für wissenschaftliche Auswertungen.
- ▶ Lahmann und Werner (1992). Immissionsüberwachung in der DDR : vorveröffentlichter Teilbericht zu "Feststellung und Bewertung von Immissionen - Leitfaden zur Immissionsüberwachung in Deutschland": Forschungsbericht 10402267.
- ▶ Landesumweltamt Brandenburg (1995). Luftqualität 1975 bis 1990 (Studien und Tagungsberichte).
- ▶ Ratzlaff (1990). Ermittlung von Wahrscheinlichkeitsdichteverteilungen für Konzentrationswerte in Luft- und Niederschlagsproben, insbesondere zur Risikoabschätzung bei Belastungen im Extrembereich. TU Dresden.
- ▶ Teichert und Warmbt (1955). Ozonuntersuchungen am Meteorologischen Observatorium Wahnsdorf (Abhandlungen des Meteorologischen und Hydrologischen Dienstes der Deutschen Demokratischen Republik). Berlin.
- ▶ Warmbt (1966). Surface ozone and artificial β - activity in Dresden-Wahnsdorf. *Tellus*, 18(2-3), 441–450. <https://doi.org/10.1111/j.2153-3490.1966.tb00256.x>
- ▶ Werner (1992a). Wintersmog in der ehemaligen DDR. Teil I: Ursachen, Kriterien und Datengewinnung. *Staub - Reinh. D. Luft.* (52), 119–126.
- ▶ Werner (1992b). Wintersmog in der ehemaligen DDR. Teil II: Datenanalyse, Diskussion und Ausblick. *Staub - Reinh. D. Luft*, 52, 239–244.

Weiterhin lagen einige Betriebsanleitungen damals eingesetzter Messgeräte (z. B. SO₂-Immissionsmessgerät CM-5, Ozonanalysator Typ Ozonograph II) oder Anleitungen sowie TGL-Standards vor.

- ▶ TGL 39612 (1982) - Nutzung und Schutz der Atmosphäre. REINHALTUNG DER LUFT. Begriffe.
- ▶ TGL 39615/02 (1983) - Nutzung und Schutz der Atmosphäre. Manuelle Immissionsmethoden. Allgemeine Forderungen zur Probenahme.
- ▶ TGL 39615/03 (1984) - Nutzung und Schutz der Atmosphäre. Manuelle Immissionsmethoden. Bestimmung von Schwefeldioxid in der atmosphärischen Luft.
- ▶ TGL 39614/03 (1985) - Nutzung und Schutz der Atmosphäre. Manuelle Emissionsmethoden. Bestimmung von Stickoxiden durch alkalimetrische Titration.

In einer Scopus Suche wurden lediglich Artikel identifiziert, die vor 1990 erschienen sind und die Daten einzelner Parameter sowie Stationen/ Regionen auswerten. Diese Publikationen sind nicht digital verfügbar. Eine Beschaffung – um gegebenenfalls zusätzliche Informationen zu angewandten Messmethoden oder Plausibilitätsprüfungen und aktiven Messstationen zu erlangen – versprach keinen substanziellen Mehrwert im Vergleich zu vorliegenden Informationen und Aufwand und wurde daher nicht weiterverfolgt.

4 Datenmanagement

Das UBA stellte die verfügbaren Rohdaten der Luftqualitätsmessungen der DDR zur Verfügung. Diese Rohdaten wurden in einer Sichtung strukturiert und dann für die Weiterverarbeitung standardisiert. Ziel der Standardisierung waren Datentabellen, in denen jeder Datensatz genau einen Messwert und seine Beschreibungen wie Schadstoff, Herkunft, Datum, Uhrzeit usw. beinhaltet.

Speziell für den Schadstoff SO₂ wurde dann eine Harmonisierung durchgeführt, weil gleiche Messwerte bezüglich Station und Zeiteinheit in mehreren Dateien vorhanden waren. Bei den anderen Schadstoffen war dies nicht der Fall.

Abschließend wurden die standardisierten Rohdaten bzw. harmonisierten SO₂-Rohdaten einer Plausibilitätsprüfung unterzogen und die finalen Daten erzeugt.

4.1 Datensichtung und Strukturierung

4.1.1 Übermittlung der Rohdaten

Die übermittelten Rohdaten bestanden einerseits aus Metadaten und andererseits aus Messdaten. Zusätzlich wurden GIS-Daten übermittelt. Die Metadaten umfassten die folgenden drei Dateien (siehe Tabelle 1)

Tabelle 1: Metadaten

Dateiname	Inhalt
Parameterbeschreibung_Originale.doc	Beschreibung von 29 Parametern des MD mit Datum vom 11.03.1991
Parameterliste_DDR.xls	Inhalt ist identisch mit Parameterbeschreibung_Originale.doc
Stationsbeschr_MD_DDR.xls	Beschreibung von 300 Stationen, deren Betreibern und Messreihenparametern mit Datum vom 25.02.1991

Die Parameterbeschreibung *Parameterliste_DDR.xls* beinhaltet die folgenden Attribute

- ▶ Kurzbezeichnung
- ▶ Parametername
- ▶ Maßeinheit
- ▶ Angaben zum Messbereich

Die Datei der Stationsbeschreibungen *Stationsbeschr_MD_DDR.xlsx* hat die folgenden Inhalte

- ▶ Das Tabellenblatt Stationsbeschreibung listet 320 Stationen inkl. Stationsnummer, Kurzbezeichnung, Stationsname, Betreiber, Seehöhe und geografische Koordinaten, Landschaftstyp, Bebauungstyp und Nutzungstyp auf.
- ▶ Das Tabellenblatt Betreiberbeschreibung listet 42 Betreiber inkl. Nummer, Kürzel und Betreibername auf.

- ▶ Das Tabellenblatt Parameterbeschreibung listet 29 Luftparameter inkl. Nummer, Kurzbezeichnung, Name, Maßeinheit und Grenzwerte auf.
- ▶ Die weiteren Tabellenblätter beinhalten Messreihenbeschreibungen zu den Parametern SO2RE, SO2TG, SO2TE, O3TE, O3RE, O3SP, SWSR, SWSTT, NO_ (NO2TG, NO2TE, NO2RE, NOTG, NORE, NOxTE, NOxTG, NOxRE) und Blei TG, also zu 16 von den 29 gelisteten Parametern.
- ▶ Die Messreihenbeschreibungen stellen Beginn und Ende der Messungen, die Messhäufigkeit (pro Tag), die Dauer der Probenahme, das Verfahren und die Messtermine (Uhrzeiten) des entsprechenden Parameters für jede der 320 Stationen dar.

Die Dateien der Rohdaten der Messwerte lagen in der folgenden Stammverzeichnisstruktur vor (siehe Tabelle 2).

Tabelle 2: Anzahl Rohdatendateien der Messwerte in der Stammverzeichnisstruktur

Stammverzeichnis	Anzahl Dateien
Blei	6
Chlor	4
Fluor	4
Kohlenmonoxid_Org_Potsdam	8
NO	3
NO2_REG_Org_Potsdam	1
NO2_tageswerte_Org_Potsdam	16
NO2_Terminwerte_org_Potsdam	7
NOx_REG_Org_Potsdam	16
Ozon_Org_Ptsdam	35
Schwebstaub_Org_Potsdam	38
Schwefelkohlenstoff	4
Schwefelwasserstoff	6
SO2_Reg	1.117
SO2_Regspitzen	7
SO2_Tageswerte	17
SO2_terminwerte	21
SO2_Vorh_Dresden	630
Sulfat_Org_potsddam	2
Gesamt	1.942

Aus den Bezeichnungen der Verzeichnisse ergibt sich, dass zum Parameter SO₂ 1.792 von den insgesamt 1.942 Dateien gehören. Das sind etwas mehr als 92 % der Dateien. Die restlichen 8 % der Dateien beziehen sich auf Schwebestaub, Ozon, NO, NO₂ und NO_x sowie CO, NO, Schwefelkohlenstoff, Schwefelwasserstoff, Blei, Chlor und Fluor.

Die übermittelten GIS-Daten bestehen aus

- ▶ ArcGIS-Shape Dateien von Messtationen in unterschiedlichen Koordinatensystemen
- ▶ Informationen zu Koordinatensystemen
- ▶ Geografische Karten in verschiedenen Formaten
- ▶ Excel-Importdaten verschiedener Sachverhalte (z. B. Kraftwerke, Stationen DWD)

4.1.2 Strukturierung der Rohdaten

Die Strukturierung der Rohdaten bezieht sich auf die Messdaten. Diese lagen als Dateien in der Stammverzeichnisstruktur vor (siehe Tabelle 2).

Für die Durchmusterung dieser Dateien wurde ein Tool programmiert, das die folgenden Informationen zu jeder Datei zusammenstellt (siehe Rohdaten-Strukturierungs-Excel-Mappe: *Rohdatendateien.xlsx*):

- ▶ Auflösung des Stammverzeichnisses in alle Unterverzeichnisse
- ▶ Dateiname
- ▶ Extension des Dateinamens
- ▶ Größe der Datei in Byte
- ▶ Ermittlung der ersten 10 Datenzeilen und deren Längen

Mit Hilfe dieser Informationen wurde jede Datei nach Art und Typ klassifiziert (siehe Tabelle 3).

Tabelle 3: Klassifizierung der Rohdatendateien der Messwerte

Art der Datei	Typ der Datei	Anzahl Dateien	Beschreibung
Text	ASCII Typ-A	695	SO ₂ Tagesmittelwerte, Monatsmittelwerte und Jahresmittelwert je Station und Jahr in einer Datei, Monate in Spalten, Tage in Zeilen
Text	ASCII Typ-B	562	verschiedene Parameter unterschiedlicher Stationen in unterschiedlicher zeitlicher Auflösung, geblockt in Zeilen mit fester Anzahl von Spalten
Text	ASCII Typ-C	2	SO ₂ unterschiedlicher Stationen in fester zeitlicher Auflösung für das Jahr 1990 im csv-Format
Text	ASCII Typ-D	1	SO ₂ unterschiedlicher Stationen, ähnlich dem ASCII Typ-B aber mit Tabulatoren als Trennzeichen der Spalten

Art der Datei	Typ der Datei	Anzahl Dateien	Beschreibung
Text	ASCII Typ-E	13	SO ₂ -Monatswerte unterschiedlicher Stationen, ähnlich dem ASCII Typ-A
Text	ASCII Typ-F	1	NO _x ohne zeitliche Zuordnung, ähnlich dem ASCII Typ-B (es fehlen die ersten sieben Spalten mit dem Hinweis „eventuell März“)
Text	ASCII Typ-G	1	SO ₂ -Monats- und Jahresmittelwerte verschiedener Stationen des Meteorologischen Dienstes
Text	ASCII Typ-ohne Inhalt	1	nicht relevant, Datei ist ohne Inhalt
Binär	IBM Lotus 1-2-3 oder Borland Quattro Pro	586	SO ₂ -Tagesmittelwerte ohne Angabe von Stationen
Binär	Microsoft Excel 4-Arbeitsmappe	2	SO ₂ -Tagesmittelwerte ohne Angabe von Stationen, inhaltliches Format entspricht dem Typ IBM Lotus 1-2-3
Binär	Microsoft Excel 97-2003-Arbeitsmappe	66	SO ₂ unterschiedlicher Stationen in fester zeitlicher Auflösung
Binär	Microsoft Write	3	nicht relevant, Inhalt ist unbedeutender Text
Binär	Microsoft Excel 4.0-Makro	8	nicht relevant, eine Sammlung von Makros ohne Dokumentation
Binär	Nero	1	nicht relevant, Indexdatei für das Brennen von CDs mit der Software Nero Burning ROM des Herstellers Nero AG
Gesamt		1942	

Die Art der Datei beschreibt die Möglichkeiten, wie die Weiterverarbeitung der Datei erfolgen kann. Der Typ der Datei spezifiziert bei den Text-Dateien das inhaltliche Format, bei den Binär-Dateien zunächst nur das Software-Format. Das inhaltliche Format der Binär-Dateien erschließt sich erst, wenn die Dateien mit der zugehörigen Software geöffnet werden.

4.2 Standardisierung der Rohdaten

4.2.1 ASCII Typ-A

In jeder Datei des ASCII Typ-A sind für ein Jahr die SO₂-Tages- und Monatsmittelwerte sowie der Jahresmittelwert gespeichert. Alle Dateien des ASCII Typ-A sind strukturiert in Verzeichnissen abgelegt (siehe Tabelle 4).

Tabelle 4: Verzeichnisstruktur der Rohdatendateien vom ASCII Typ-A

Rohdatenverzeichnis	Unterverzeichnis	Anzahl Dateien	Jahr der Messung ⁵
Rohdaten\SO2_Reg\79_90_TM_JM\	SO2R78	36	1978
	SO2R79	41	1979
	SO2R80	42	1980
	SO2R81	40	1981
	SO2R82	41	1982
	SO2R83	42	1982
	SO2R84	47	1984
	SO2R85	49	1985
	SO2R86	48	1986
	SO2R87	49	1987
	SO2R88	72	1988
	SO2R89	90	1989
	SO2R90	98	1990
Gesamt	13 Unterverzeichnisse	695	13 Jahre

Die Dateien vom ASCII Typ-A sind Textdateien mit der Zeichencodierung OEM 850⁶. Mit dieser Zeichencodierung werden die enthaltenen Umlaute und Sonderzeichen korrekt dargestellt (siehe Abbildung 5).

⁵ Das Jahr der Messung wurde aus den Namen der Unterverzeichnisse bestimmt.

⁶ Die Zeichencodierung OEM 850 wurde im Jahre 1987 unter MS-DOS eingeführt.

Abbildung 5: Inhalt ASCII Typ-A

Tagesmittelwerte
 " MÖlbis
 so2-Werte Januar-Dezember 1990 /181230BHI "

Tag	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
1.	58	263	261	397	146	166	216	154	-99	0	81	121
2.	37	331	233	577	0	212	320	103	-99	82	135	71
3.	44	318	217	647	471	301	0	97	5	61	61	55
4.	239	292	811	591	348	227	0	145	2	36	32	42
5.	402	416	403	523	346	168	168	131	0	65	56	95
6.	149	331	243	397	370	153	269	120	32	29	63	128
7.	170	278	811	368	406	317	52	77	57	46	69	198
8.	305	323	811	-99	415	195	715	15	32	20	86	205
9.	434	357	376	340	277	152	1163	23	20	32	67	95
10.	354	415	378	190	210	120	174	62	4	46	113	50
11.	390	413	305	110	116	81	17	130	0	24	124	77
12.	516	590	313	60	132	74	405	163	8	11	28	152
13.	516	483	321	157	127	109	31	155	0	3	42	146
14.	654	209	362	148	85	113	0	193	0	3	111	166
15.	576	190	374	190	96	108	-99	182	-99	1	72	131
16.	418	177	418	467	77	87	359	183	-99	61	163	150
17.	370	355	382	230	119	92	28	167	62	98	-99	116
18.	218	456	373	130	14	105	8	126	7	99	-99	115
19.	401	469	482	212	1	136	0	97	68	90	80	129
20.	530	310	508	178	0	164	7	130	8	60	127	82
21.	365	383	527	177	0	247	0	85	120	45	40	89
22.	383	334	597	175	19	482	0	64	4	36	26	101
23.	344	370	421	175	79	2616	0	67	2	30	24	55
24.	386	393	549	177	141	2556	0	93	12	-99	37	56
25.	392	483	421	186	118	1306	5	91	79	60	21	43
26.	369	556	413	174	118	529	14	-99	0	68	17	128
27.	401	282	366	269	113	253	15	89	0	40	61	134
28.	319	334	304	153	109	27	57	113	2	13	-99	12
29.	441		345	172	109	0	31	161	8	18	36	114
30.	271		364	260	133	35	14	165	1	41	102	-99
31.	193		356		139		120	123		148		20
MoMw	343	361	420	270	172	383	190	116	26	47	69	102
AnzW	31	28	31	29	28	29	22	30	20	29	27	30
Jahresmittelwert	= 208											

length: 3.754 lines: 43 Ln: 1 Col: 1 Pos: 1 Windows (CR LF) OEM 850 INS

Darstellung der Textdatei MMÖLB90.SO2 mit der Zeichencodierung OEM 850

Quelle: eigene Darstellung, BioMath

Für die Weiterverarbeitung der Dateien vom ASCII Typ-A wurde ein Tool in Delphi⁷ entwickelt, das die Zeichencodierung OEM 850 berücksichtigt und aus den Zeilen und Spalten dieser Textdateien die folgenden Variablen ermittelte (siehe Tabelle 5).

Tabelle 5: Extrahierte Variablen der Rohdatendateien vom ASCII Typ-A

Variable	Label
StationNameKurz	Kurzbezeichnung der Station
Jahr	Jahr
Stationnummer	Stationsnummer

⁷ Delphi ist eine integrierte Entwicklungsumgebung für die Programmiersprache Object Pascal.

Variable	Label
Betreiber	Betreiber
BetreiberKennzeichen	Kennzeichen des Betreibers
Tag	Tag als Zahl
Monat	Monat als Zahl
Messwert	Tagesmittelwert
QMWMon	Monatsmittelwert
QMWMonKz	Kennzeichen des Monatsmittelwertes
QMWMonAnz	Anzahl der Messwerte zur Berechnung des Monatsmittelwertes
QMWMonAnzKz	Kennzeichen der Anzahl der Messwerte zur Berechnung des Monatsmittelwertes
QMWJahr	Jahresmittelwert
QMWJahrKz	Kennzeichen des Jahresmittelwertes

Zur Feststellung der Richtigkeit der ermittelten Variablen wurden einerseits Prüfungen auf die Anzahl der Zeilen, der Zeilenlängen und der Sonderzeichen durchgeführt sowie deskriptive Statistiken, wie z. B. Häufigkeiten in Kreuztabellen und statistische Maßzahlen, ermittelt und bewertet. Weiterhin wurden Monats- und Jahresmittelwerte nachgerechnet, mit den gespeicherten Monats- und Jahresmittelwerten in den Dateien verglichen und ggf. Ursachen für die Abweichungen ermittelt. Ferner wurde geprüft, ob die Stationsnummern in den Metadaten vorhanden waren.

Dabei wurden folgende Besonderheiten identifiziert:

- ▶ Betreiber und Betreiberkennzeichen an falscher Position in Zeile 2
- ▶ Jahreszahlen in Zeile 2 unvollständig
- ▶ Stationsnummern nicht korrekt bezüglich der Metadaten zu den Stationen
- ▶ Tagesmittelwerte nicht in der richtigen Spalte einer Zeile
- ▶ Tagesmittelwerte mit Buchstaben vor den Werten
- ▶ Keine Tagesmittelwerte oder Fehlwerte am 29. Februar von Schaltjahren vorhanden
- ▶ Falsch berechnete Monats- und Jahresmittelwerte

Die unvollständigen Jahreszahlen wurden mit Hilfe der Dateinamen der Rohdaten und Kreuztabellen bezüglich der Jahre vervollständigt.

Die Stationsnummern wurden mit Hilfe der Stationskurznamen und der zusätzlichen Metainformation zu den Stationsnummern korrigiert. Im Ergebnis dieser Korrekturen lag nun eine Tabelle vor, mit deren Hilfe man ggf. unbekannte Stationsnummern in den anderen Rohdaten korrigieren konnte.

Alle Korrekturen wurden in der Rohdaten-Strukturierungs-Excel-Mappe (siehe Abschnitt 4.1.2) dokumentiert.

Die Ursache für falsch berechnete Monats- und Jahresmittelwerte liegt sehr wahrscheinlich darin, dass die Berechnung der Mittelwerte – insbesondere im Schritt der Summation der Einzelwerte – mit einem nicht geeigneten numerischen Zahlentyp erfolgte, denn es wurden nur die Mittelwerte falsch berechnet, deren Summe größer als 65535 ist. Dies deutet darauf hin, dass für die Summation der numerische Zahlentyp Word eingesetzt wurde. Dieser Zahlentyp besteht aus zwei Byte und kann positive ganze Zahlen nur im Bereich von 0 bis 65535 darstellen. Wird nun während der Summation der Wert von 65535 überschritten, kommt es zum arithmetischen Überlauf (Overflow) und der Summenwert nimmt eine beliebige Zahl an, d. h. sowohl der Summenwert als auch der daraus berechnete Mittelwert sind unbestimmt.

Die gespeicherten falschen Monats- und Jahresmittelwerte wurden nicht verändert. Falls diese Mittelwerte später relevant werden sollten, werden sie aus den Tagesmittelwerte neu berechnet.

Durch Nachrechnung der Mittelwerte für Monat und Jahr ergab sich das Folgende.

Die gespeicherten Monatsmittelwerte wurden wie folgt berechnet:

1. Summation der Tagesmittelwerte eines Monats unter Berücksichtigung der Fehlwerte mit der Codierung -99
2. Ermittlung der Anzahl der Tage des Monats mit gültigen Tagesmittelwerten einschließlich der Null-Werte
3. Ganzzahlige Division der Summe durch die Anzahl der Tage, d. h. die Nachkommastellen der so berechneten Monatsmittelwerte wurden abgeschnitten

Der gespeicherte Jahresmittelwert wurde wie folgt berechnet:

1. Summation der Monatsmittelwerte unter Berücksichtigung der Fehlwerte mit der Codierung -99
2. Ermittlung der Anzahl der Monate mit gültigen Monatsmittelwerten
3. Ganzzahlige Division der Summe durch die Anzahl der Monate, d. h. die Nachkommastellen des so berechneten Jahresmittelwertes wurden abgeschnitten

In einem weiteren Schritt wurden die extrahierten Variablen (siehe Tabelle 5) aufgearbeitet und um zusätzliche Variablen ergänzt, sodass final für jeden Tagesmittelwert ein Datensatz mit allen zugehörigen Informationen vorliegt.

4.2.2 ASCII Typ-B

In jeder Datei des ASCII Typ-B sind Messwerte für ein Jahr und einen Luftschadstoff für mehrere Stationen gespeichert. Alle Dateien des ASCII Typ-B sind strukturiert in Verzeichnissen abgelegt (siehe Tabelle 6).

Tabelle 6: Verzeichnisstruktur der Rohdatendateien vom ASCII Typ-B

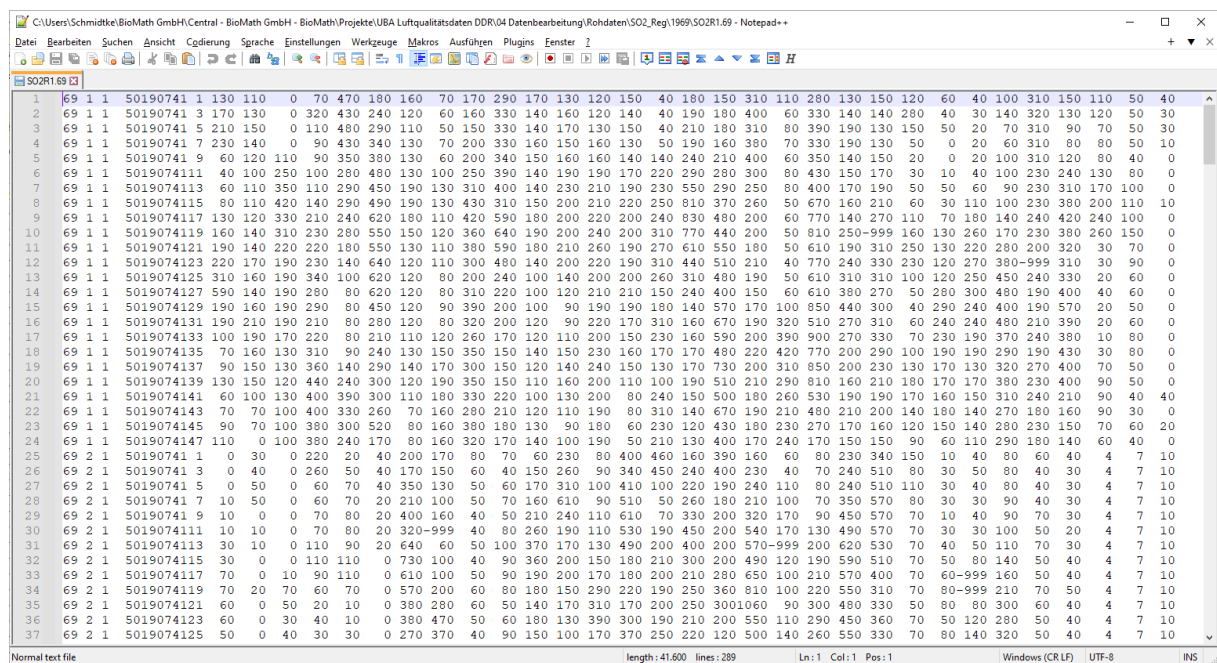
Rohdatenverzeichnis	Unterverzeichnis	Anzahl Dateien
Rohdaten\	Blei	4

Rohdatenverzeichnis	Unterverzeichnis	Anzahl Dateien
	Chlor	4
	Fluor	4
	Kohlenmonoxid_Org_Potsdam	8
	NO	3
	NO2_REG_Org_Potsdam	1
	NO2_tageswerte_Org_Potsdam	16
	NO2_Terminwerte_org_Potsdam	7
	NOx_REG_Org_Potsdam	15
	Ozon_Org_Ptsdam	35
	Schwebstaub_Org_Potsdam	38
	Schwefelkohlenstoff	4
	Schwefelwasserstoff	6
	SO2_Reg	370
	SO2_Regspitzen	7
	SO2_Tageswerte	17
	SO2_terminwerte	21
	Sulfat_Org_potsddam	2
Gesamt	18 Unterverzeichnisse	562

Die Namen der Unterverzeichnisse geben Hinweise auf die Schadstoffe und deren zeitliche Erfassung. Die Bezeichnung *Reg* weist auf registrierte Messwerte hin, d. h. Messwerte, die je Stunde oder je halbe Stunde erfasst worden sind. *Terminwerte* sind Messwerte die nur zu bestimmten Zeitpunkten erhoben worden sind. *Tageswerte* sind Messwerte mit einer Abdeckung über 24 Stunden von 7 bis 7 Uhr und nicht zu verwechseln mit Tagesmittelwerten.

Die Dateien vom ASCII Typ-B sind auch Textdateien mit der Zeichencodierung OEM 850 und bestehen aus Zahlenkolonnen (siehe Abbildung 6). Durch die Verwendung dieser Zeichencodierung können Steuerzeichen in den Textdateien korrekt identifiziert werden.

Abbildung 6: Inhalt ASCII Typ-B



Darstellung der Textdatei SO2R1.69 mit der Zeichencodierung OEM 850

Quelle: eigene Darstellung, BioMath

Zunächst wurde der inhaltliche Aufbau der Rohdatendateien analysiert und mit der Beschreibung aus dem UBA-Bericht „Immissionsmessungen von Luftschadstoffen in der DDR von 1969 bis 1990“ abgeglichen.

Die Rohdatendateien haben 142 Zeichen je Zeile und feste Spalten mit dem folgenden Aufbau (siehe Tabelle 7).

Tabelle 7: Aufbau der Rohdatendateien vom ASCII Typ-B

Spalte	Bedeutung
1-2	Jahr zweistellig
3-4	Monat als Zahl
5-6	Codierung des Luftschadstoffs als Zahl
7-10	Betreiber als Zahl 4-stellig
11-16	Stationsnummer 6-stellig
17-18	Zeitpunkt der Messung als Zahl zwischen 1 und 48
19-22	Messwert 4-stellig für Tag 1
23-26	Messwert 4-stellig für Tag 2
...	...
139-142	Messwert 4-stellig für Tag 31

Für die Weiterverarbeitung der Dateien vom ASCII Typ-B wurde ein Tool in Delphi entwickelt, das alle Rohdatendateien in eine Datei zusammenfasst und die zusammengehörigen Spalten in Variablen abbildet (siehe Tabelle 8).

Tabelle 8: Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-B

Spalte	Bedeutung
IDDatensatz	Laufende Nummer die den Datensatz eindeutig identifiziert
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
DateiName	Dateiname der Rohdatendatei mit vollständigem Pfad
DateiTyp	Typ der Rohdatendatei (immer ASCII Typ-B)
NurDateiName	Dateiname der Rohdatendatei
VerzBez9	Verzeichnisebene der Rohdatendatei
VerzBez10	Verzeichnisebene der Rohdatendatei
VerzBez11	Verzeichnisebene der Rohdatendatei
LineNr	Nummer der Zeile in der Rohdatendatei
LineLaenge	Länge der Zeile in der Rohdatendatei
Jahr	Jahr zweistellig wie in den Rohdaten
Monat	Monat als Zahl wie in den Rohdaten
Komponente	Codierung des Luftschadstoffs als Zahl wie in den Rohdaten
BetreiberNummer	Betreiber als Zahl wie in den Rohdaten
StationNummer	Stationsnummer wie in den Rohdaten
Zeitinfo	Zeitpunkt der Messung als Zahl wie in den Rohdaten
Messwert1	Messwert Tag 1 wie in den Rohdaten
Messwert2	Messwert Tag 2 wie in den Rohdaten
...	...
Messwert31	Messwert Tag 31 wie in den Rohdaten
Kommentar	Kommentar
AlleMesswerteMissing	Kennzeichen als Zahl (0-nicht alle Messwerte sind Fehlwerte, 1-alle Messwerte sind Fehlwerte)

Zur Feststellung der Richtigkeit der ermittelten Variablen wurden einerseits Prüfungen auf Steuerzeichen und Zeilenlängen durchgeführt und andererseits deskriptive Statistiken, wie z. B. einfache Häufigkeiten, Häufigkeiten in Kreuztabellen und statistische Maßzahlen, ermittelt und bewertet. Ferner wurde geprüft, ob die Stationsnummern und Betreiber in den Metadaten vorhanden waren.

Dabei wurden folgende Besonderheiten identifiziert:

- ▶ Steuerzeichen an unterschiedlichen Positionen
- ▶ zusätzliche Leerzeichen in Zeile
- ▶ weniger als 142 Zeichen in der Zeile (Datenzeile abgeschnitten)
- ▶ Jahr und Monat teilweise falsch
- ▶ Codierung der Luftschadstoffe nicht eindeutig
- ▶ Zeitpunkt der Messung fehlt
- ▶ Stationsnummern nicht in der Metadatei
- ▶ Codierung der Betreiber nicht eindeutig
- ▶ Fehlwerte werden mit -999 oder -99 codiert
- ▶ Keine Fehlwerte am 29. Februar von Schaltjahren vorhanden

Der größte Teil der Besonderheiten konnte korrigiert werden. Vorhandene Messwerte wurden dabei nicht korrigiert. Die Korrekturen erfolgten in zwei voneinander getrennten Arbeitsschritten.

1. Es wurden die physischen Besonderheiten korrigiert. Dazu gehörte die automatische Entfernung von Steuerzeichen durch das eingesetzte Tool zur Zusammenführung der Rohdatendateien und eine manuelle Anpassung der Rohdaten. Die manuelle Anpassung beinhaltete das Entfernen oder Hinzufügen von logisch erschließbaren einzelnen Zeichen in den Zeilen, deren Länge nicht 142 betrug.
2. Es wurden die inhaltlichen Besonderheiten mit SPSS korrigiert.

Es konnten alle Besonderheiten bis auf die folgenden acht Stationsnummern korrigiert werden (siehe Tabelle 9).

Tabelle 9: Stationsnummern in den Rohdatendateien vom ASCII Typ-B, die nicht in den Metadaten vorhanden sind

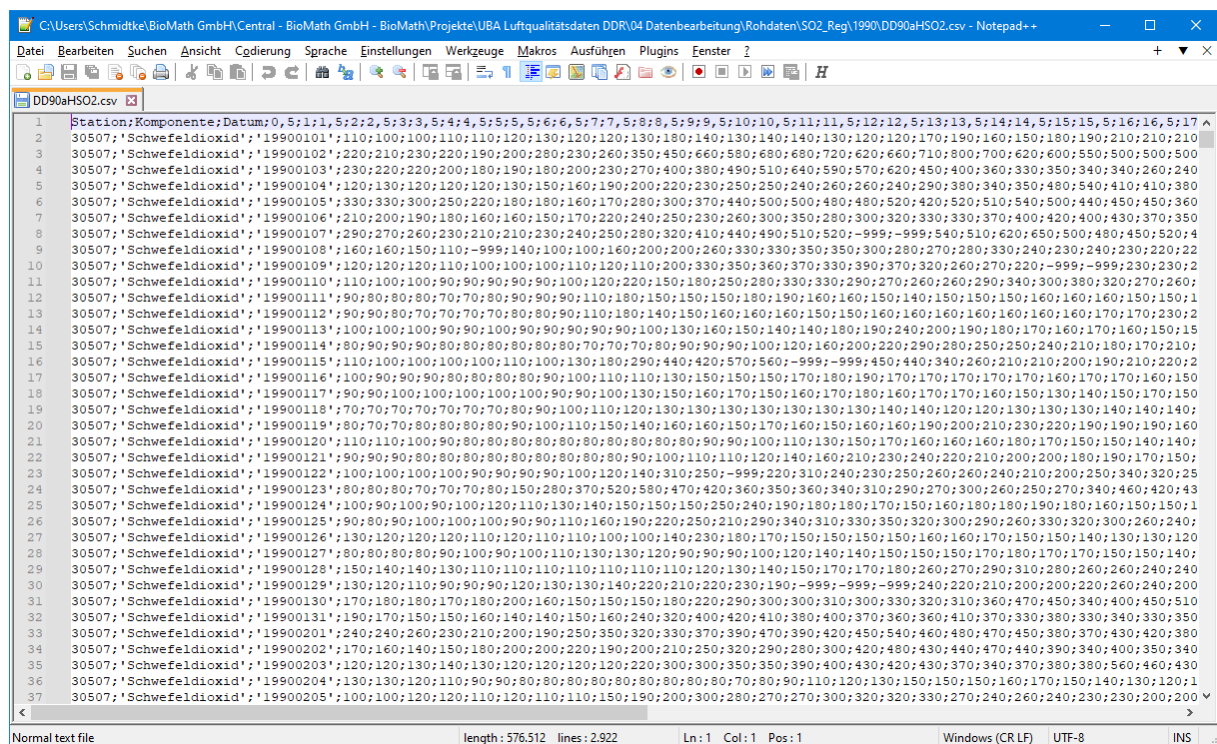
Stationsnummer	Rohdaten Dateiname	Anzahl Zeilen	Anmerkung
73816	SO2TE.86A	60	ist ein durchgehender Block in der Rohdatendatei
150651	SO2RMÜ.86	48	ist ein durchgehender Block in der Rohdatendatei
170206	SO2RMÜ.85	240	ist ein durchgehender Block in der Rohdatendatei
170206	SO2RMÜ.86	240	ist ein durchgehender Block in der Rohdatendatei
170609	SO2R2.86	576	ist ein durchgehender Block in der Rohdatendatei
183616	NO2TG.82A	3	in allen Zeilen sind alle Messwerte Fehlwerte
205142	SO2RMÜ.86	96	ist ein durchgehender Block in der Rohdatendatei
205438	SO2RMÜ.85	192	ist ein durchgehender Block in der Rohdatendatei
205438	SO2RMÜ.86	288	ist ein durchgehender Block in der Rohdatendatei
244852	SO2TG.89A	3	in allen Zeilen sind alle Messwerte Fehlwerte

Das Ergebnis der Korrekturen wurde für die Weiterverarbeitung in einer Textdatei mit der Struktur laut Tabelle 8 gespeichert.

4.2.3 ASCII Typ-C

Vom ASCII TYP-C gibt es 2 Dateien. Sie sind Textdateien im csv-Format mit dem Trennzeichen Semikolon, dem Hochkomma zu Kennzeichnung von Text und mit der Zeichencodierung UTF-8. (siehe Abbildung 7).

Abbildung 7: Inhalt ASCII Typ-C



Darstellung der Textdatei DD90aHSO2.csv

Quelle: eigene Darstellung, BioMath

In der ersten Zeile stehen die Variablenamen, in den folgenden Zeilen die Daten. Die 2 Dateien unterscheiden sich in der Anzahl der Spalten für die Messwerte (siehe Tabelle 10).

Tabelle 10: Aufbau der Rohdatendateien vom ASCII Typ-C

Spalte	Variable	Bedeutung
1	Station	Stationsnummer
2	Komponente	Bezeichnung der Komponente, ist immer Schwefeldioxid
3	Datum	Datum achtstellig im Format JJJJMMTT
4-51	numerische Zahlen	Datei DD90aHSO2.csv: 48 Messwerte mit numerischen Variablenamen in Schritten von 0,5 beginnend mit 0,5, das entspricht Schritten von 30 Minuten beginnend bei 0:30 Uhr

Spalte	Variable	Bedeutung
4-27	Wertiiii	DD90aSSO2.csv: 24 Messwerte mit den folgenden Variablennamen Wert0000, Wert0100, Wert0200, Wert0300, ... , Wert2300, das entspricht Stundenwerten beginnend bei 0:00 Uhr.

Die Buchstaben H und S in den Dateinamen und die Anzahl der jeweiligen Messwerte weisen auf halbstündliche und stündliche Messungen hin, wobei die halbstündlichen Messungen bei 0:30 Uhr und die stündlichen Messungen bei 0:00 Uhr beginnen.

Mit einem entwickelten Tool in Delphi wurden die Dateien vom ASCII Typ-C zusammenfasst und die zusammengehörigen Spalten in Variablen abgebildet (siehe Tabelle 11).

Tabelle 11: Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-C

Spalte	Bedeutung
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
Volldateiname	Dateiname der Rohdatendatei mit vollständigem Pfad
DateiName	Dateiname der Rohdatendatei
DateiTyp	Typ der Rohdatendatei (immer ASCII Typ-C)
StationNummer	Stationsnummer wie in den Rohdaten
Komponente	Bezeichnung der Komponente wie in den Rohdaten, ist immer Schwefeldioxid
Datum	Datum wie in den Rohdaten (siehe Spalte 3 der Rohdaten)
Tag	Tag ermittelt aus Datum
Monat	Monat ermittelt aus Datum
Jahr	Jahr ermittelt aus Datum
Stunde	Stunde ermittelt aus den Variablennamen der Messwerte
Minute	Minute ermittelt aus den Variablennamen der Messwerte
Messwert	Messwert wie in den Rohdaten
NumOfLine	Nummer der Zeile in den Rohdaten
NumTokenOfLine	Nummer der Spalte in den Rohdaten

Es wurden folgende Besonderheiten identifiziert und korrigiert:

- ▶ Messwerte ohne Zahlenangaben
- ▶ unterschiedliche Zählung der Stunden

Messwerte ohne Zahlenangaben wurden auf -999 gesetzt und somit als Fehlwerte gekennzeichnet.

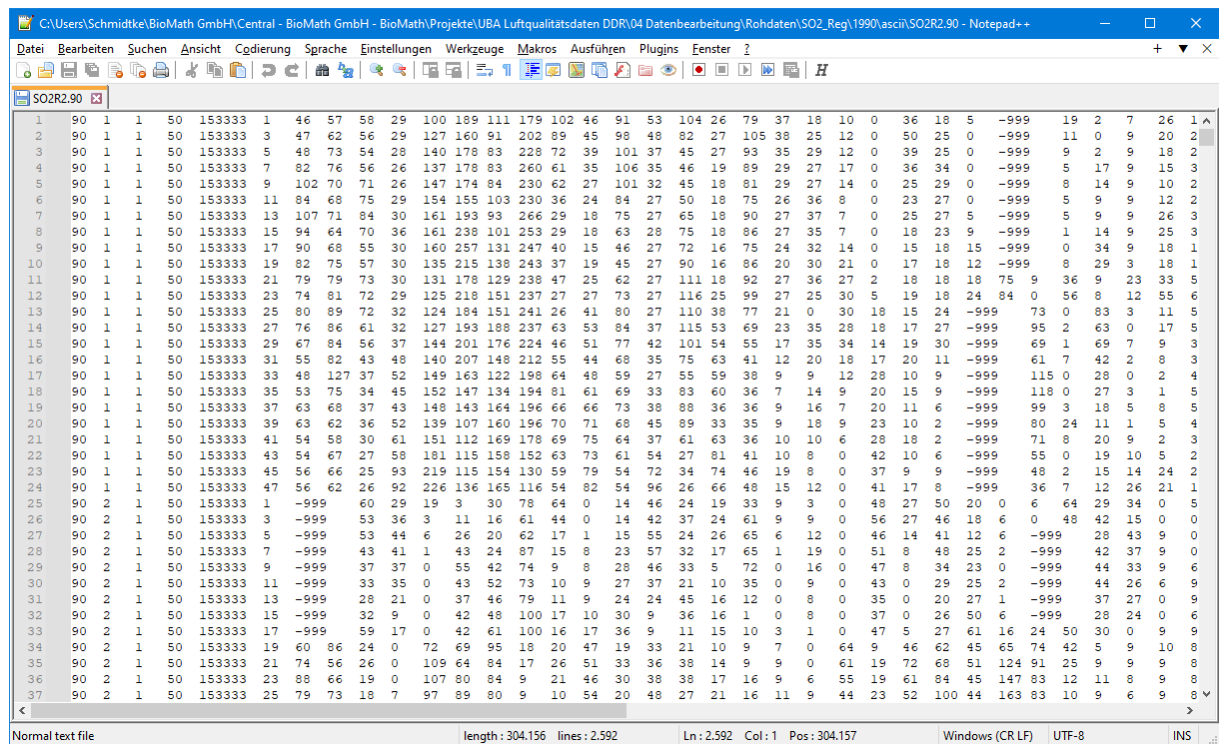
Die Messwerte, die der Uhrzeit 24:00 Uhr zugordnet waren, wurden auf den Folgetag mit der Uhrzeit 0:00 Uhr transformiert.

Alle Stationsnummern der Dateien vom ASCII Typ-C waren in den Metadaten vorhanden.

4.2.4 ASCII Typ-D

Vom ASCII TYP-D gibt es 1 Datei. Sie ist eine Textdatei mit der Zeichencodierung UTF-8 und ähnelt im Aufbau den Dateien des ASCII Typ-B. Die Variablen befinden sich allerdings nicht in festen Spalten, sondern sind mit dem Tabulator getrennt (siehe Abbildung 8).

Abbildung 8: Inhalt ASCII Typ-D



Darstellung der Textdatei SO2R2.90

Quelle: eigene Darstellung, BioMath

In der Datei gibt es keine Variablennamen. Daten beginnen unmittelbar in Zeile 1 und haben den folgenden Aufbau (siehe Tabelle 12).

Tabelle 12: Aufbau der Rohdatendateien vom ASCII Typ-D

Spalte	Bedeutung
1	Jahr zweistellig, immer 90
2	Monat als Zahl
3	Unbekannt, ist immer 1
4	Codierung des Luftschadstoffs als Zahl, ist immer 50
5	Stationsnummer
6	Zeitpunkt der Messung als ungerade Zahl zwischen 1 und 47
7 - 37	31 Messwerte für die Tage 1 bis 31

Mit einem entwickelten Tool in Delphi wurde die Datei vom ASCII Typ-D zur Weiterverarbeitung umstrukturiert und um Variablen ergänzt (siehe Tabelle 13).

Tabelle 13: Aufbau der zusammengeführten Rohdatendateien vom ASCII Typ-D

Spalte	Bedeutung
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
DateiName	Dateiname der Rohdatendatei mit vollständigem Pfad
DateiTyp	Typ der Rohdatendatei (immer ASCII Typ-D)
NurDateiname	Dateiname der Rohdatendatei
VerzBez9	Verzeichnisebene der Rohdatendatei
VerzBez10	Verzeichnisebene der Rohdatendatei
VerzBez11	Verzeichnisebene der Rohdatendatei
LineNr	Nummer der Zeile in der Rohdatendatei
NumTokenOfLine	Anzahl Spalten in der Rohdatendatei
Jahr	Jahr zweistellig wie in den Rohdaten
Monat	Monat wie in den Rohdaten
Unbekannt	Unbekannt wie in den Rohdaten, ist immer 1
Komponente	Codierung des Luftschadstoffs als Zahl wie in den Rohdaten, ist immer 50
StationNummer	Stationsnummer wie in den Rohdaten
Zeitinfo	Zeitpunkt der Messung als Zahl wie in den Rohdaten
Messwert1 – Messwert31	Messwerte Tag 1 bis Tag 31 wie in den Rohdaten

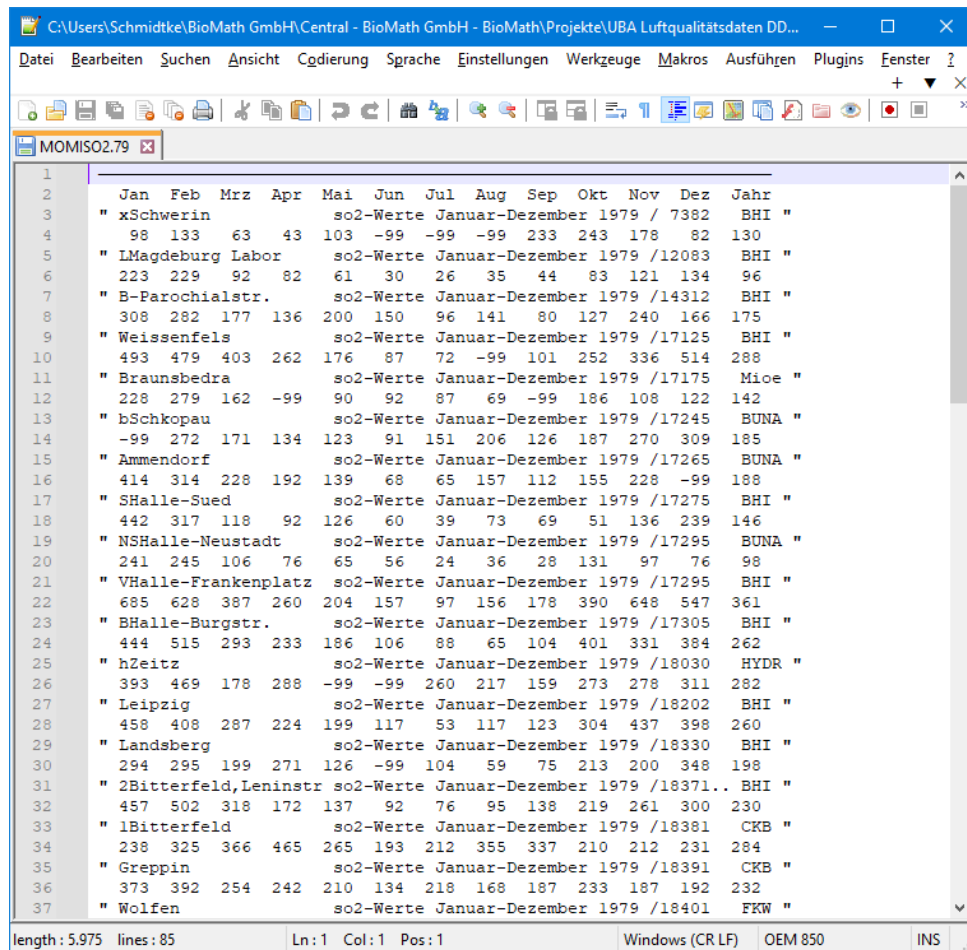
Es wurden keine Besonderheiten identifiziert.

Alle Stationsnummern der Datei vom ASCII Typ-D waren in den Metadaten vorhanden.

4.2.5 ASCII Typ-E

Vom ASCII Typ-E gibt es 13 Dateien. Es sind Textdateien mit der Zeichencodierung OEM 850 und sie beinhalten Monats- und Jahresmittelwerte von SO₂. Teilweise ähnelt der Aufbau dieser Textdateien dem ASCII Typ-B (siehe Abbildung 9).

Abbildung 9: Inhalt ASCII Typ-E



Darstellung der Textdatei MOMISO2.79 mit der Zeichencodierung OEM 850

Quelle: eigene Darstellung, BioMath

Die Zeile 1 besteht aus den dreistelligen Kürzeln der Monate und dem Jahr. Bei den folgenden Zeilen gehören immer zwei Zeilen zusammen. Die erste Zeile davon beschreibt die Station, den gemessenen Luftschadstoff, das Jahr und den Betreiber, die zweite Zeile die Mittelwerte der Monate und des Jahres.

Mit einem entwickelten Delphi-Tool wurden aus den Zeilen und Spalten der Textdateien vom ASCII Typ-E die folgenden Variablen ermittelt (siehe Tabelle 14)

Tabelle 14: Extrahierte Variablen der Rohdatendateien vom ASCII Typ-E

Variable	Label
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
VollDateiName	Dateiname der Rohdatendatei mit vollständigem Pfad
Dateityp	Typ der Rohdatendatei (immer ASCII Typ-E)
Messparameter	SO ₂ -Mittelwerte
Jahr	Jahr
StationNameKurz	Kurzbezeichnung der Station

Variable	Label
StationNummer	Stationsnummer
Betreiber	Betreiber
Mon1 - Mon12	Mittelwerte der Monate 1 bis 12
Jahr	Jahresmittelwert

Dabei wurden folgende Besonderheiten identifiziert:

- ▶ falsche Zeilenumbrüche
- ▶ negative Mittelwerte
- ▶ eine Jahreszahl ist unvollständig
- ▶ eine Stationsnummer fehlt
- ▶ falsche Stationsnummern

Fast alle Besonderheiten konnten korrigiert werden. Vorhandene Messwerte wurden dabei nicht korrigiert.

Da die Dateien eine handhabbare Größe hatten, wurden die Korrekturen manuell durchgeführt. Bei der unvollständigen Jahreszahl war nur das Jahrhundert vorhanden. Das Jahrzehnt und das Jahr konnte auf Grund der anderen Jahreszahlen in der Datei erschlossen werden. Die fehlende Stationsnummer wurde mithilfe der Variablen StationNameKurz und Betreiber aus den Metadaten eindeutig bestimmt.

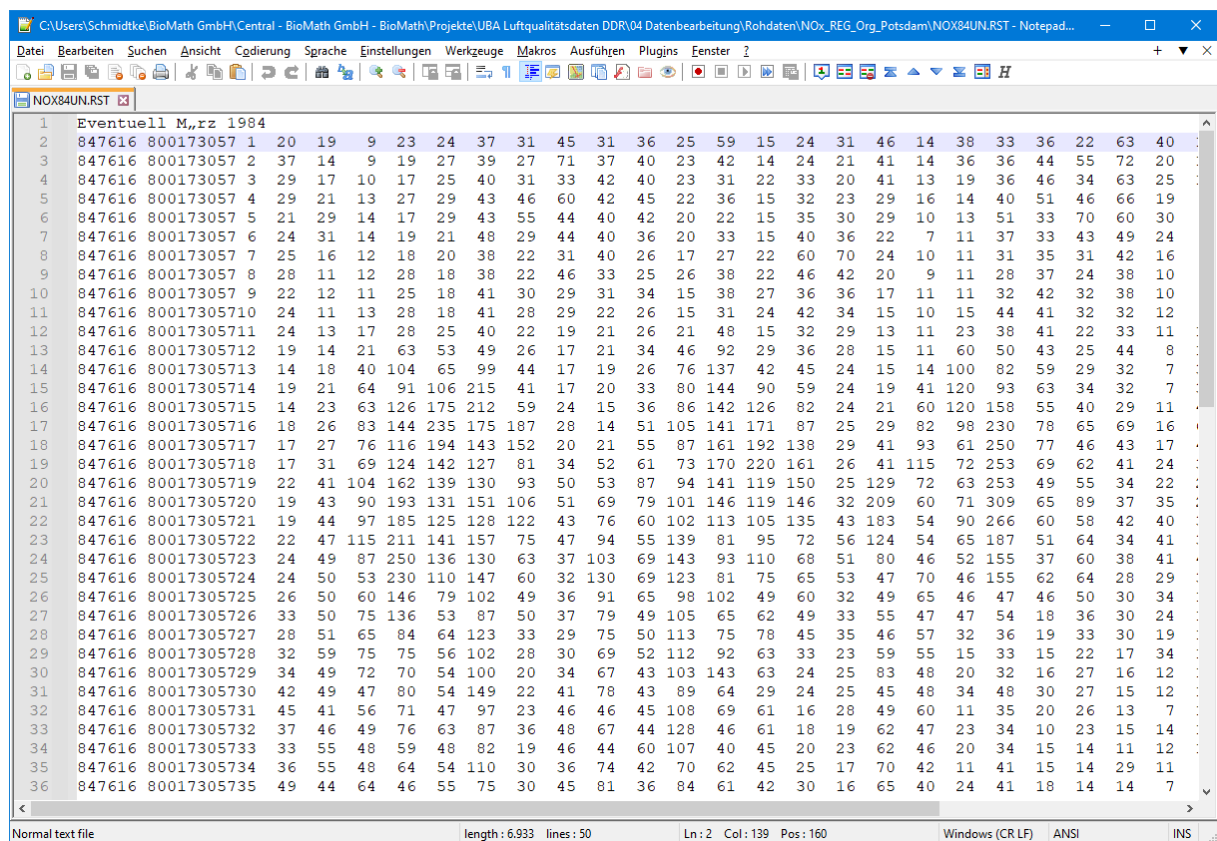
Von den insgesamt 185 verschiedenen Stationsnummern in den Rohdaten konnten 49 in den Metadaten aufgefunden werden. Die restlichen 136 konnten nicht aufgefunden werden. Eine Überprüfung ergab, dass bei diesen 136 Stationsnummern die letzte Ziffer fehlte, also die Stationsnummern am Ende abgeschnitten waren.

Unter Berücksichtigung der abgeschnittenen Stationsnummer, der Kurzbezeichnung der Station und des Betreibers konnten zu 135 abgeschnittenen Stationsnummern die vollständigen Stationsnummern eindeutig zugeordnet werden. Mit dieser Methode konnte allerdings die Station mit der abgeschnittenen Nummer 3050 (Kurzbezeichnung=Rostock, Betreiber=BHI) nicht eindeutig einer vollständigen Stationsnummer zugeordnet werden. Für die Zuordnung standen die Stationen mit den Nummern 30507 (Kurzbezeichnung=Rostock Gertr. Straße, Betreiber=BHI-R) und 30508 (Kurzbezeichnung=Rostock Wielandstraße, Betreiber=BHI-R) zur Auswahl. Über den Vergleich der Monats- und Jahresmittelwerte dieser Stationen mit den zugehörigen Mittelwerten aus den Rohdaten des ASCII Typ-A konnte dann eine eindeutige Zuordnung getroffen werden (siehe Anlagedatei *All ASCII Typ-E (unbekannte Stationen).xlsx*).

4.2.6 ASCII Typ-F

Vom ASCII Typ-F gibt es 1 Datei. Sie ist eine Textdatei mit der Zeichencodierung ANSI und der Aufbau der Datei entspricht dem ASCII Typ-B (siehe Abbildung 10).

Abbildung 10: Inhalt ASCII Typ-F



Darstellung der Textdatei NOX84UN.RST mit der Zeichencodierung OEM 850

Quelle: eigene Darstellung, BioMath

Die Datei hat 49 Zeilen und folgende Besonderheiten:

- ▶ In der ersten Zeile steht: Eventuell März 1984.
- ▶ Der Monat (Spalte 2 und 3) ist mit der Zahl 76 codiert

Alle Messwerte beziehen sich auf den Luftschadstoff NO_x (Spalte 5-6 mit dem Eintrag 16) und die Station mit der Nummer 173057 (Spalte 11-16). Für diese Station, den Luftschadstoff NO_x , den Monat März und das Jahr 1984 gib es bereits Messwerte in der Datei NOX84.RST im gleichen Verzeichnis.

Da die Inhalte dieser Datei nicht einem korrekten Monat zugeordnet werden konnten, wurde diese Datei nicht weiterverarbeitet.

4.2.7 ASCII Typ-G

Vom ASCII Typ-G gibt es 1 Datei. Sie ist eine Textdatei mit der Zeichencodierung ANSI und beinhaltet Monats- und Jahresmittelwerte von SO_2 (siehe Abbildung 11).

Abbildung 11: Inhalt ASCII Typ-G

```

C:\Users\Schmidtk\BioMath GmbH\Central - BioMath GmbH - BioMath\Projekte\UBA Luftqualitätsdaten DDR\04 Datenbearbeitung\Rohdaten\SO2_Vorh_Dresden\...
Datei Bearbeiten Suchen Ansicht Codierung Sprache Einstellungen Werkzeuge Makros Ausführen Plugins Fenster ?
MDMWSO2.ASC
1 Monats- und Jahresmittelwerte der SO2-Konzentration von Stationen des Meteorologischen Dienstes
2
3 Angaben in æg/m3
4
5 Ort Betreiber Stationscode
6 Jan Feb Mrz Apr Mai Juni Jul Aug Sep Okt Nov Dez JM
7 Arkona MD 44126
8 1988 20 7 9 6 3 2 3 2 3 8 9 4 6
9 1989 10 6 6 5 2 1 1 2 2 5 11 11 5
10 1990 10 10 3 3 1 1 2 1 1 8 5 7 4
11
12
13 Artern MD 172318
14 1988 120 81 81 71 65 34 24 30 51 69 103 65 66
15 1989 80 75 50 55 45 31 23 27 24 39 104 168 60
16 1990 80 68 37 47 28 10 15 17 59 46 83 44
17
18 Angermnde MD 90260
19 1988 48 34 29 18 12 10 14 20 22 33 22 15 23
20 1989 52 42 33 14 8 9 5 13 9 12 12 34 20
21 1990 58 57 26 11 5 5 3 6 4 15 12 22 18
22
23 Aue MD 233643
24 1988 150 128 141 118 94 40 58 74 84 109 140 108 104
25 1989 135 99 72 92 73 38 17 47 89 #DIV/0! #DIV/0! #DIV/0! 73
26 1990 123 81 77 65 54 30 27 33 24 59 33 65 55
27
28 Berlin-Buch MD 143830
29 1988 72 56 43 20 9 4 12 11 18 31 49 35 30
30 1989 92 79 58 35 13 12 8 11 12 31 49 62 38
31 1990 80 70 33 22 9 12 5 12 17 37 45 45 32
32
33 Boltenhagen MD 75914
34 1988 60 25 43 22 9 7 7 11 17 19 18 19 21
35 1989 51 44 29 16 13 7 6 5 6 14 33 48 23
36 1990 66 26 27 13 7 5 4 5 12 19 29 35 21

```

Darstellung der Textdatei MDMWSO2.ASC mit der Zeichencodierung ANSI

Quelle: eigene Darstellung, BioMath

Die Informationen in dieser Datei stehen in tabulatorgetrennten Spalten und sind in Blöcken von je 4 Zeilen organisiert (siehe Tabelle 15).

Tabelle 15: Aufbau des Datenblocks der Rohdatendateien vom ASCII Typ-G

Zeile	Spalte	Bedeutung
1	1	Ort der Station
1	2	Betreiber
1	3	Stationsnummer
2	1	Jahr
2	2-13	Monatsmittelwerte für die Monate 1 bis 12
2	14	Jahresmittelwert

Manuell wurden aus den Inhalten der Textdateien vom ASCII Typ-G die folgenden Variablen ermittelt (siehe Tabelle 16).

Tabelle 16: Extrahierte Variablen der Rohdatendateien vom ASCII Typ-G

Variable	Label
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
VollDateiName	Dateiname der Rohdatendatei mit vollständigem Pfad
Dateityp	Typ der Rohdatendatei (immer ASCII Typ-G)
Ort	Ort aus den Rohdaten
Betreiber	Betreiber aus den Rohdaten
StationNummer	Stationsnummer
Messparameter	SO ₂ -Mittelwerte
Mon1 - Mon12	Mittelwerte der Monate 1 bis 12
Jahr	Jahresmittelwert

Die Datei hatte die folgende Besonderheit:

- Messwerte mit dem Inhalt #DIV/0!

Diese Inhalte weisen darauf hin, dass der Ursprung dieser Datei eine Tabellenkalkulation ist. Alle Inhalte #DIV/0! wurden auf -999 gesetzt und somit als Fehlwerte gekennzeichnet.

Alle Stationsnummern der Datei vom ASCII Typ-G waren in den Metadaten vorhanden.

4.2.8 IBM Lotus 1-2-3 oder Borland Quattro Pro

Alle Dateien vom Typ IBM Lotus 1-2-3 oder Borland Quattro Pro haben die Extension wk1 und sind in Verzeichnissen strukturiert abgelegt (siehe Tabelle 17).

Tabelle 17: Verzeichnisstruktur der Rohdatendateien vom Typ IBM Lotus 1-2-3 oder Borland Quattro Pro

Rohdatenverzeichnis	Unterverzeichnis	Anzahl Dateien
Rohdaten\SO2_Vorh_Dresden\	BHI	219
	MD	367
Gesamt	2 Unterverzeichnisse	586

Die Dateierweiterung wk1 lässt vermuten, dass diese Dateien von IBM Lotus 1-2-3 erstellt wurden.

IBM Lotus 1-2-3 ist ein Tabellenkalkulationsprogramm, das 1983 für DOS und ab 1991 für Windows zur Verfügung stand.

Um genau festzustellen, ob es sich tatsächlich um IBM Lotus-Dateien handelt, wurde eine Prüfroutine entwickelt, die die ersten sechs Bytes dieser Dateien einliest und sie mit der binären Dateibeschreibung von IBM Lotus 1-2-3 abgleicht. Im Ergebnis konnte festgestellt werden, dass bei allen Dateien die ersten sechs Bytes die folgenden gleichen Inhalte hatten (siehe Tabelle 18).

Tabelle 18: Inhalt und Bedeutung der ersten sechs Byte von Dateien mit der Extension wk1 im Dateinamen

Byte	Inhalt hexadezimal	Bedeutung
1 und 2	0000 H	Code für BOF (Begin Of File); muss immer 0000 H sein
3 und 4	0002 H	Länge in Byte für den folgenden Eintrag des Dateiformats; muss immer 0002 H sein
5 und 6	0406 H	Code für das Dateiformat; 0406 H steht für das Dateiformat IBM Lotus 1-2-3 Version 2.01

Somit liegen alle Dateien der Klassifikation Typ IBM Lotus 1-2-3 oder Borland Quattro Pro im Format IBM Lotus 1-2-3 Version 2.01 vor.

Da die Originalsoftware von IBM nicht zur Verfügung stand, wurde nach einer Alternative gesucht. Microsoft Excel unterstützte bis zur Version 2000 (intern Version 9.0) die Bearbeitung von IBM Lotus 1-2-3 Dateien. Diese Excelversion stand auf einem alten Windows XP Rechner noch zur Verfügung, sodass alle Daten damit geöffnet und die Tabelleninhalte als Textdateien mit der Codepage UTF-8 zur Weiterverarbeitung abgespeichert werden konnten. Dabei wurde der Rohdatendateiname beibehalten, jedoch die Extension wk1 in txt geändert.

Die wesentlichen Inhalte der Dateien sind die ersten fünf Spalten, wobei jeweils in der ersten Zeile die Variablenamen stehen (siehe Abbildung 12 und Tabelle 19).

Abbildung 12: Inhalt Datei vom Typ IBM Lotus 1-2-3

	A	B	C	D	E
1	CASENAME	BA_DATEN	INFO	ORGDATEN	Mittel
2	1\1\78	50,7	PISO2		76,06799316
3	2\1\78	92		92	77,0269165
4	3\1\78	70		70	76,80369568
5	4\1\78	36		36	76,59642029
6	5\1\78	159		159	76,40344238
7	6\1\78	55		55	78,15666962
8	7\1\78	55		55	79,44194031
9	8\1\78	55		55	80,5531311
10	9\1\78	106		106	80,77879333
11	10\1\78	83		83	81,28530121
12	11\1\78	62		62	81,76286316
13	12\1\78	89		89	82,21389771
14	13\1\78	102		102	84,31622314
15	14\1\78	102		102	85,04475403
16	15\1\78	102		102	85,06925201
17	16\1\78	64		64	85,69251251
18	17\1\78	73		73	89,45612335
19	18\1\78	4		4	93,04050446
20	19\1\78	53		53	96,45817566
21	20\1\78	75		75	98,49321747
22	21\1\78	75		75	99,63781738
23	22\1\78	75		75	100,4065628
24	23\1\78	111		111	101,3128052
25	24\1\78	77		77	102,8896255
26	25\1\78	76		76	104,4020844
27	26\1\78	101		101	105,8540039
28	27\1\78	71		71	108,5599976
29	28\1\78	71		71	109,6399994
30	29\1\78	71		71	112,7399979
31	30\1\78	129		129	115,6000061
32	31\1\78	118		118	114,2800064
33	1\2\78	115		115	115,0400085

Darstellung der Textdatei BATCM78.TXT in Microsoft Excel 365

Quelle: eigene Darstellung, BioMath

Tabelle 19: Aufbau der Dateien des Typs IBM Lotus 1-2-3

Spalte	Variablenname	Bedeutung
A	CASENAME oder DATUM	Datum in der Form Tag\Monat\Jahr mit Monat als Zahl und Jahr zweistellig oder Datum der Form Tag-Monat-Jahr mit Monat als dreibuchstabile Abkürzung und Jahr zweistellig
B	Buchstabenfolgen z. B. BA_DATEN, KA_DATEN, DATEN, BR_DATEN, ...	Numerische Zahlen
C	INFO oder Mittel	Alphanumerische Zeichenkette oder Mittelwert
D	ORGDATEN oder Mittel	Numerische Zahlen oder Mittelwert
E	Mittel	Mittelwert

In einigen wenigen Dateien ist in den Spalten F bis T eine Matrix (Monate in Spalten, Tage in Zeilen) mit den Inhalten der Spalte B und den zugehörigen Berechnungen der Mittelwerte für die Monate und das Jahr.

Für die Weiterverarbeitung der Textdateien vom Typ IBM Lotus 1-2-3 wurde ein Tool in Delphi entwickelt, das alle Textdateien in eine standardisierte Datei überführte. Dabei wurde die Spalte B in zwei Variablen aufgeteilt und aus dem Datum der Tag, Monat und das Jahr in separate Variablen überführt (siehe Tabelle 20).

Tabelle 20: Extrahierte und zusätzlich ermittelte Variablen der Rohdatendateien vom Typ IBM Lotus 1-2-3

Variablenname	Bedeutung
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
VollDateiname	Dateiname der Rohdatendatei mit vollständigem Pfad
Dateiname	Dateiname der Rohdatendatei
Line	Nummer der Zeile in der Rohdatendatei
Betreiber	Betreiber (BHI oder MD)
DATUM	Datum aus Spalte A der Rohdatendatei
DATENKennung	Variablenname der Spalte B der Rohdatendatei
DATEN	Inhalt der Spalte B der Rohdatendatei
INFO	Inhalt der Variablen INFO der Rohdaten
ORGDATEN	Inhalt der Variablen ORGDATEN der Rohdaten (SO ₂ -Tageswert)
Mittel	Inhalt der Variablen INFO der Rohdaten
ATag	Tag als Zahl (extrahiert aus Variable DATUM)
AMonat	Monat als Zahl (extrahiert aus Variable DATUM)
AJahr	Jahr als vierstellige Zahl (extrahiert aus Variable DATUM und die Ziffern 19 vorangestellt)
Kommentar	Kommentar

In der Variable ORGDATEN sind SO₂-Tageswerte gespeichert, die Variable besaß keine numerische Codierung von Fehlwerten.

Die ersten beiden Buchstaben des Dateinamens und die ersten beiden Buchstaben der Variable DATENKennung sind Kürzel von Stationen, die mit Hilfe der Anhänge der Teilberichte 3, 4 und 5 des Vorhabens Dresden (Fleischer & Pohl, 1995) in Stationsnummern überführt werden konnten (siehe Anlagedatei *All IBM Lotus.xlsx*).

In einer deskriptiven Analyse wurde festgestellt, dass es in einer Rohdatendatei Datensätze mit falschen Monatsangaben gab. Diese Datensätze wurden von der weiteren Verarbeitung ausgeschlossen.

4.2.9 Microsoft Excel

Die 68 Microsoft Exceldateien wurden hinsichtlich ihrer inhaltlichen Struktur klassifiziert (siehe Tabelle 21)

Tabelle 21: Klassifizierung der Microsoft Exceldateien

Klassifizierung	Inhalt der Dateien	Anzahl Dateien
Microsoft Excel 4-Arbeitsmappe	keine Inhalte oder Daten des IBM Lotus Typ	2
Microsoft Excel 97-2003-Arbeitsmappe Typ-A	Blei-Messwerte je Tag in Spalten	2
Microsoft Excel 97-2003-Arbeitsmappe Typ-B	SO ₂ -Messwerte stündlich in Spalten	4
Microsoft Excel 97-2003-Arbeitsmappe Typ-C	SO ₂ -Messwerte halbstündlich in Spalten	2
Microsoft Excel 97-2003-Arbeitsmappe Typ-D	SO ₂ -Messwerte halbstündlich in Zeilen	26
Microsoft Excel 97-2003-Arbeitsmappe Typ-E	SO ₂ -Tagesmittelwerte über zwei Tabellenblätter	2
Microsoft Excel 97-2003-Arbeitsmappe Typ-F	SO ₂ -Monats- und Jahresmittelwerte	1
Microsoft Excel 97-2003-Arbeitsmappe Typ-G	SO ₂ -Tagesmittelwerte über mehrere Tabellenblätter	29

Jede klassifizierte Datei hat ihr eigenes inhaltliches Format, das auch manchmal innerhalb einer Klassifizierung von Datei zu Datei minimal variiert. Auf Grund dieser Vielfalt wird an dieser Stelle auf die exakte Beschreibung der Rohdatendateien verzichtet.

Sowohl manuell als auch mit Hilfe eines entwickelten Delphi-Tools wurden die Rohdaten in die folgenden Variablen überführt (siehe Tabelle 22).

Tabelle 22: Extrahierte und zusätzlich ermittelte Variablen der Rohdatendateien der Typen A bis G der Microsoft Excel 97-2003-Arbeitsmappen

Typ	Variablenname	Bedeutung
A - G	IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
A - G	VollDateiname	Dateiname der Rohdatendatei mit vollständigem Pfad
A - G	Verz9	Verzeichnisebene der Rohdatendatei
A - G	Verz10	Verzeichnisebene der Rohdatendatei
A - G	Verz11	Verzeichnisebene der Rohdatendatei
A - G	Dateiname	Dateiname der Rohdatendatei
E - G	Sheet	Bezeichnung des Tabellenblatts der Rohdatendatei
A - G	LineNr	Nummer der Zeile in der Rohdatendatei
A - C, E - G	Komponente	Komponente
A - G	StationNummer	Stationsnummer
A, D - G	Betreiber oder BetreiberName	Betreiber
D - G	Ort	Ort
B, C	Datum	Datum der Form TT.MM.JJJJ

Typ	Variablenname	Bedeutung
A, D – G	Jahr	Jahr als Zahl vierstellig
A, D, E	Monat	Monat als Zahl
A, D	Zeitinfo	Nummer der Zeit
A, E	Tag1 - Tag31	Tageswert bzw. Tagesmittelwert
B, C	Wert1 - Wert47/Wert48	Stunden- oder Halbstundenwert
D	Messwert1 - Messwert31	Tagesmittelwert
F, G	Jan - Dez	Monatsmittelwert
F	JahrMittelwert	Jahresmittelwert
D	ImmerEins	nicht identifizierbarer Inhalt
D	NummerUnbekannt	nicht identifizierbarer Inhalt

Dabei wurden folgende Besonderheiten identifiziert:

- ▶ falsche Stationsnummern (Typ D und E)
- ▶ keine Information zu den Stationen (Typ G)
- ▶ mehrere Jahre auf einem Tabellenblatt (Typ G), normalerweise ist immer ein Jahr auf einem Tabellenblatt
- ▶ Mittelwerte haben Nachkommastellen (Typ E und F)

In einer Rohdatendatei vom Typ D war bei einer Stationsnummer eine Ziffernfolge vertauscht. Über den zugehörigen Stationsnamen konnte mit Hilfe der Metadaten die korrekte Stationsnummer ermittelt werden. In den Rohdatendateien vom Typ E waren 1049 Stationsnummern nicht in den Metadaten zu finden. Diese Stationsnummern waren bezüglich zu den korrekten Stationsnummern hochgezählt, d. h. in Excel wurde sehr wahrscheinlich die automatische Ausfüllfunktion nicht korrekt verwendet.

Innerhalb der Rohdaten des Typ G gab es zu den Stationen nur manchmal eine Angabe und diese bezog sich nur auf den Ort. Allerdings weisen auch die letzten 4 bzw. 5 Buchstaben der Dateinamen auf einen Ort hin. Mit Hilfe der Stationsnamen aus den Rohdaten bzw. aus den Dateinamen konnten über die Metadaten in fast allen Fällen die Stationsnummern eindeutig ermittelt werden. Nur bei der Datei SO2PLAU.XLS konnte zunächst die Zuordnung nicht eindeutig getroffen werden. Mit Hilfe der zusätzlichen Information über den Messbeginn und das Messende wurde dann die Zuordnung eindeutig getroffen.

4.2.10 Metadaten der Stationen

Die Metadaten zu den Stationen umfassen insgesamt 320 Stationen, wobei es zu 5 Stationen keine geografischen Koordinaten gibt. Zur Prüfung und Ermittlung der korrekten Stationsnummern in den Daten wurden die folgenden Variablen extrahiert (siehe Tabelle 23).

Tabelle 23: Datei mit Stationen aus den Metadaten

Variablenname	Bedeutung
LN	Laufende Nummer
Stationsnummer	Stationsnummer ganzzahlig
Kurzbezeichnung	Kurzbezeichnung der Station
Stationsname	Name der Station
Betreiber	Abkürzung des Betreibers

Für die spätere Bewertung von Messwerten an benachbarten Stationen wurde aus den in der Metadatei hinterlegten geografischen Koordinaten die Entfernungen als einfache euklidische Distanz⁸ zwischen den Stationen in km berechnet (siehe Tabelle 24 und Anlagedatei *Stationen (Metadaten)(Entfernungen).xlsx*).

Tabelle 24: Datei mit den Abständen zwischen den Stationen

Variablenname	Bedeutung
LfdNr	Laufende Nummer in dieser Datei
LN-1 und LN-2	Laufende Nummer aus den Metadaten für Station-1 und Station-2
StatNr-1 und StatNr-2	Stationsnummer aus den Metadaten
KurzBez-1 und KurzBez-2	Kurzbezeichnung der Station aus den Metadaten
Stationsname-1 und Stationsname-2	Name der Station aus den Metadaten
Betreiber-1 und Betreiber-2	Abkürzung des Betreibers aus den Metadaten
Entfernung	Entfernung zwischen der Station-1 und Station-2 in km

4.2.11 Ergebnisse der Standardisierung

Alle Rohdaten vom ASCII Typ-A (SO₂-Tagesmittelwerte) wurden vollständig standardisiert. Somit ist jeder SO₂-Tagesmittelwert der Rohdaten eindeutig bezüglich des Tages und der Station einschließlich Fehlwertkennzeichnung definiert.

Die Rohdaten vom ASCII Typ-B konnten bis auf sehr wenige Ausnahmen vollständig standardisiert werden. Diese Ausnahmen betreffen SO₂-Messreihen von sechs Stationen, die nur in den Jahren 1985 oder 1986 gemessen wurden. Die kürzeste Messreihe ist einen Monat, die längste ein Jahr lang.

Der größte Datenbestand liegt für SO₂, gefolgt von O₃ und TSP vor. Von diesen drei Luftschadstoffen gibt es Messwerte für die Jahre 1969 bis 1990. Alle anderen Schadstoffe wurden nur über wenige Jahre gemessen (z. B. Blei von 1978 bis 1981 oder NO_x von 1978 bis 1982 und 1983 bis 1990).

⁸ Da die relativen Entfernungen zwischen den Stationen interessant sind genügt die euklidische Distanz für diese Betrachtungen.

Die Rohdaten vom Typ IBM Lotus 1-2-3 beinhalten SO₂-Messwerte, die sich jeweils auf einen Tag beziehen.

Die Rohdaten der ASCII Typen C und D sind SO₂-Werte in der Auflösung von Stunden bzw. Halbstunden.

Die Rohdaten der ASCII Typen E und G beinhalten SO₂-Mittelwerte der Monate und Jahre ohne Nachkommastellen. Hier handelt es sich um extrahierte Zeilen der Dateien vom ASCII Typ-A.

Die Rohdatendatei vom ASCII Typ F weisen NO_x-Werte aus, die keinem zeitlichen Abschnitt zugeordnet werden können. Die Datei wurde deshalb von der Weiterverarbeitung ausgeschlossen.

Zwei der Rohdatendateien vom Excel-Typ beinhalten Werte von Blei, alle anderen Dateien beziehen sich auf SO₂-Werte. Etwa die Hälfte der Rohdatendateien von SO₂-Werten besteht aus Stunden- oder Halbstundenwerten. Die andere Hälfte gehört laut Verzeichnisstruktur zum Vorhaben Dresden (Fleischer & Pohl, 1995), in ihnen sind Tages- und Jahresmittelwerte ohne Nachkommastellen abgespeichert.

Über alle Daten hinweg gibt es Unterschiede in der Zahlendarstellung von Mittelwerten. Entweder haben die Mittelwerte Nachkommastellen oder keine. In einer Nachrechnung hat sich gezeigt, dass nach der Berechnung der Mittelwerte die Nachkommastellen abgeschnitten und nicht gerundet wurden.

Weiterhin wurde festgestellt, dass die Stunden- und Halbstundenwerte von SO₂ überwiegend in 10er Schritten vorliegen. Nur in einigen wenigen Ausnahmen liegen diese Werte in 5er Schritten vor.

4.3 Harmonisierung der standardisierten Rohdaten von SO₂

Die Messwerte des Luftschadstoffs SO₂ lagen in unterschiedlichen Dateitypen vor. Für alle anderen Schadstoffe sind die Messwerte in nur einem Dateityp (ASCII Typ-B) gespeichert. Deshalb war eine Harmonisierung für den Luftschadstoff SO₂ notwendig.

Ziel der Harmonisierung waren Messreihen, die je Station und Zeitpunkt nur einen Wert haben.

Messwerte zu SO₂ lagen in den zeitlichen Auflösungen

- ▶ Jahre
- ▶ Monate
- ▶ Tage und
- ▶ Zeitpunkte (Stunden oder halbe Stunden)

vor (siehe Tabelle 25).

Tabelle 25: Dateitypen mit SO₂-Werten unterschiedlicher zeitlicher Auflösungen

Typ der Datei	Anzahl Dateien	Zeitpunkt	Tag	Monat	Jahr
ASCII Typ-A	695		X	X	X
ASCII Typ-B	562	X	X		
ASCII Typ-C	2	X			
ASCII Typ-D	1	X			

Typ der Datei	Anzahl Dateien	Zeitpunkt	Tag	Monat	Jahr
ASCII Typ-E	13			X	X
ASCII Typ-G	1			X	X
Excel Typ-B*	4	X			
Excel Typ-C*	2	X			
Excel Typ-D*	26	X			
Excel Typ-E*	2		X		
Excel Typ-F*	1			X	X
Excel Typ-G*	29		X		

* Microsoft Excel 97-2003-Arbeitsmappe

Für die Harmonisierung der SO₂-Messwerte wurden die standardisierten Dateien zunächst in eine einheitliche Struktur überführt (siehe Tabelle 26), um Prüfungen auf Gleichheit der Messwerte unterschiedlicher Dateitypen durchzuführen.

Tabelle 26: Datei mit SO₂-Messwerten

Variablenname	Bedeutung
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
DateiTyp	Typ der Rohdatendatei
StationNummer	Stationsnummer
Messparameter	Bezeichnung des Messparameters (SO ₂ -Zeitpunkt, SO ₂ -Tag, SO ₂ -Monat oder SO ₂ -Jahr)
Jahr	Jahr als Zahl vierstellig
Monat	Monat als Zahl
Tag	Tag als Zahl
Stunde	Stunde als Zahl
Minute	Minute als Zahl
Wert	Messwert ohne Nachkommastellen
TagImJahr	Tag im Jahr als Zahl
WochentagBez	Bezeichnung des Wochentages zweistellig (Mo, Di, Mi, Do, Fr, Sa oder So)
MonatBez	Bezeichnung des Monats dreistellig (Jan, Feb, Mrz, Apr, Mai, Jun, Jul, Aug, Sep, Okt, Nov oder Dez)

Je nach zeitlicher Auflösung wurden in diesen Dateien die Variablen Monat, Tag, Stunde und Minute mit Inhalten gefüllt.

Die Prüfung auf Gleichheit der Messwerte erfolgte getrennt nach den zeitlichen Auflösungen.

Um Rundungsfehler auszuschließen, wurde die Gleichheit auch bei einer Abweichung von ± 1 angenommen.

4.3.1 SO₂-Jahreswerte

Die SO₂-Jahreswerte waren in vier Datenquellen vorhanden (siehe Tabelle 27). Die Zusammenhänge der SO₂-Jahreswerte bezüglich der Datenquellen sind in Tabelle 28 aufgelistet und in Abbildung 13 dargestellt.

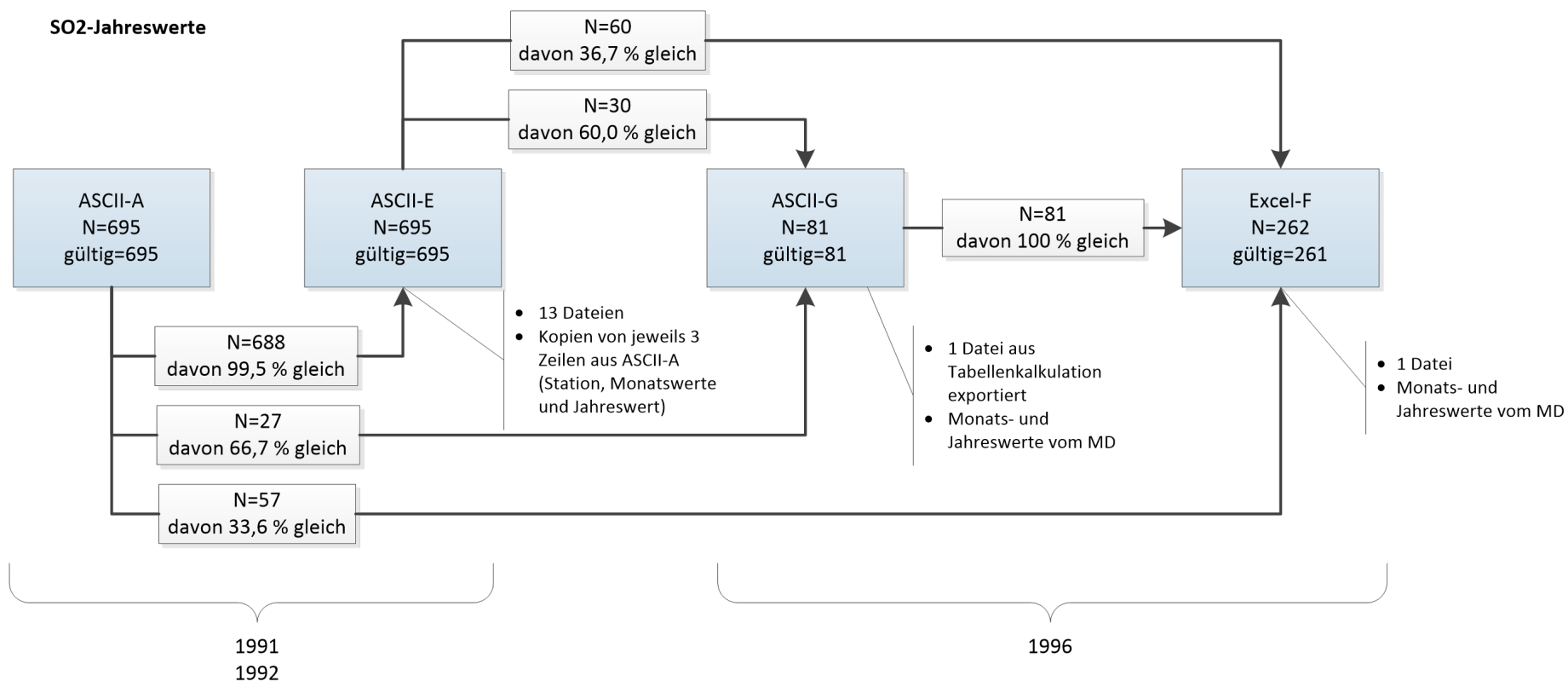
Tabelle 27: Datenquellen der SO₂-Jahreswerte

Datenquelle	Jahr der Erstellung	Anzahl Datensätze	Anzahl gültiger Werte
ASCII Typ-A	1991 oder 1992	695	695
ASCII Typ-E	1991 oder 1992	695	695
ASCII Typ-G	1996	81	81
Excel Typ-F	1996	262	261

Tabelle 28: Anzahl gleicher und ungleicher SO₂-Jahreswerte bezüglich der Datenquellen

Datenquelle-1	Datenquelle-2	Anzahl gleicher Messwerte	Anzahl ungleicher Messwerte	Anzahl Messwerte bei der nur Datenquelle-1 gültig ist	Anzahl Messwerte bei der nur Datenquelle-2 gültig ist
ASCII Typ-A	ASCII Typ-E	684	4	0	0
ASCII Typ-A	ASCII Typ-G	18	9	0	0
ASCII Typ-A	Excel Typ-F	19	38	0	0
ASCII Typ-E	ASCII Typ-G	18	12	0	0
ASCII Typ-E	Excel Typ-F	19	41	0	0
ASCII Typ-G	Excel Typ-F	81	0	0	0

Abbildung 13: Zusammenhang der SO₂-Jahreswerte bezüglich der Datenquellen



Quelle: eigene Darstellung, BioMath

In den Dateien des ASCII Typ-A sind die SO₂-Jahreswerte SO₂-Jahresmittelwerte, die aus den Monatsmittelwerten berechnet wurden. Die Übereinstimmung von 99,5 % zwischen den Werten der Dateien des ASCII Typ-A und ASCII Typ-E bestätigt die Hypothese, dass es sich bei den Dateien vom ASCII Typ-E um kopierte Zeilen aus den Dateien des ASCII Typ-A handelt.

Die SO₂-Jahreswerte in den Dateien des ASCII Typ-G und Excel Typ-F stimmen vollständig überein. Sie unterscheiden sich aber deutlich von den Werten in den Dateien des ASCII Typ-A und ASCII Typ-E. Vermutlich stammen diese Werte aus einer späteren Weiterverarbeitung.

4.3.2 SO₂-Monatswerte

Die SO₂-Monatswerte waren in vier Datenquellen vorhanden (siehe Tabelle 29). Die Zusammenhänge der SO₂-Monatswerte bezüglich der Datenquellen sind in Tabelle 30 aufgelistet und in Abbildung 14 dargestellt.

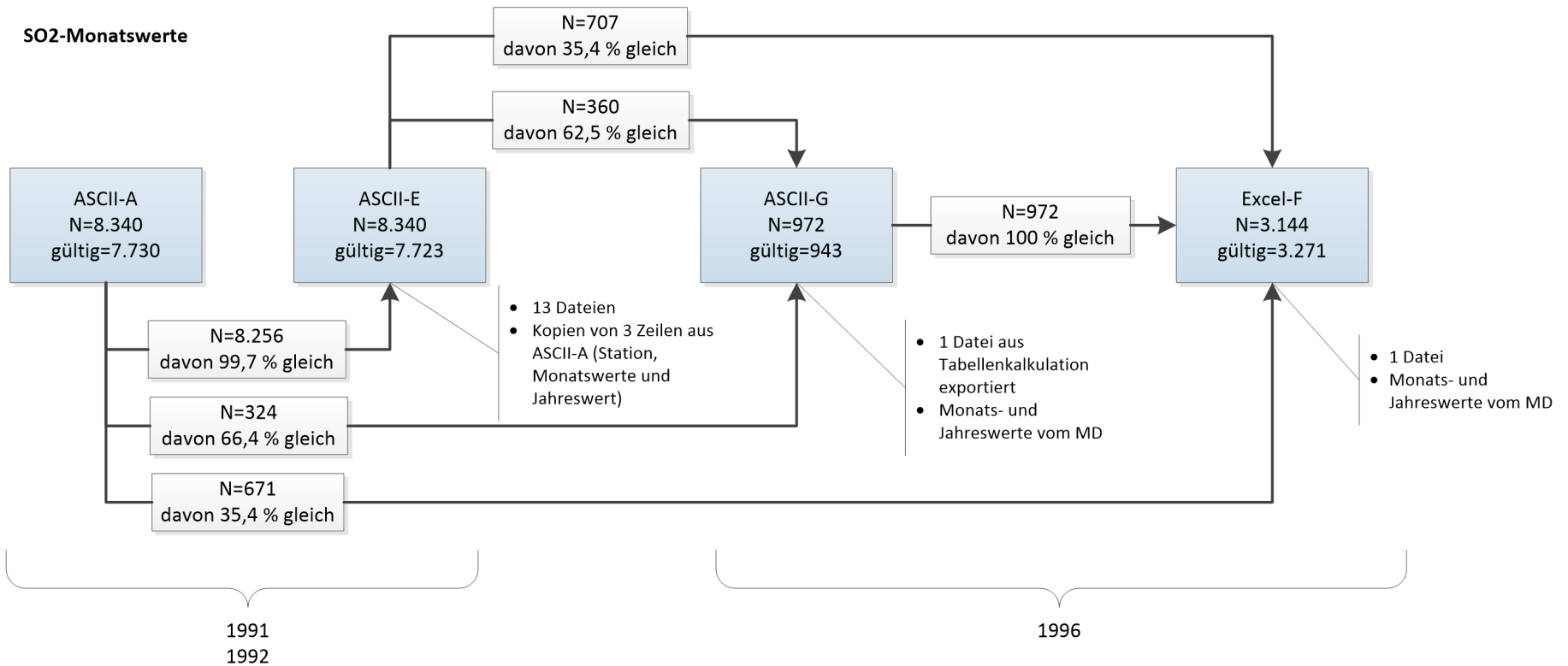
Tabelle 29: Datenquellen der SO₂-Monatswerte

Datenquelle	Jahr der Erstellung	Anzahl Datensätze	Anzahl gültiger Werte
ASCII Typ-A	1991 oder 1992	8.340	7.730
ASCII Typ-E	1991 oder 1992	8.340	7.723
ASCII Typ-G	1996	972	943
Excel Typ-F	1996	3.144	3.271

Tabelle 30: Anzahl gleicher und ungleicher SO₂-Monatswerte bezüglich der Datenquellen

Datenquelle-1	Datenquelle-2	Anzahl gleicher Messwerte	Anzahl ungleicher Messwerte	Anzahl Messwerte bei der nur Datenquelle-1 gültig ist	Anzahl Messwerte bei der nur Datenquelle-2 gültig ist
ASCII Typ-A	ASCII Typ-E	8.233	19	4	0
ASCII Typ-A	ASCII Typ-G	225	99	0	0
ASCII Typ-A	Excel Typ-F	255	416	3	10
ASCII Typ-E	ASCII Typ-G	225	135	0	0
ASCII Typ-E	Excel Typ-F	255	452	3	10
ASCII Typ-G	Excel Typ-F	972	0	0	0

Abbildung 14: Zusammenhang der SO₂-Monatswerte bezüglich der Datenquellen



Quelle: eigene Darstellung, BioMath

In den Dateien des ASCII Typ-A sind die SO₂-Monatswerte SO₂-Monatsmittelwerte, die aus den Tagesmittelwerten berechnet wurden. Die Zusammenhänge der SO₂-Monatsmittelwerte entsprechen den Zusammenhängen der SO₂-Jahresmittelwerte.

4.3.3 SO₂-Tageswerte

Die SO₂-Tageswerte waren in fünf Datenquellen vorhanden (siehe Tabelle 31). Die Zusammenhänge der SO₂-Tageswerte bezüglich der Datenquellen sind in Tabelle 32 aufgelistet und in Abbildung 15 dargestellt.

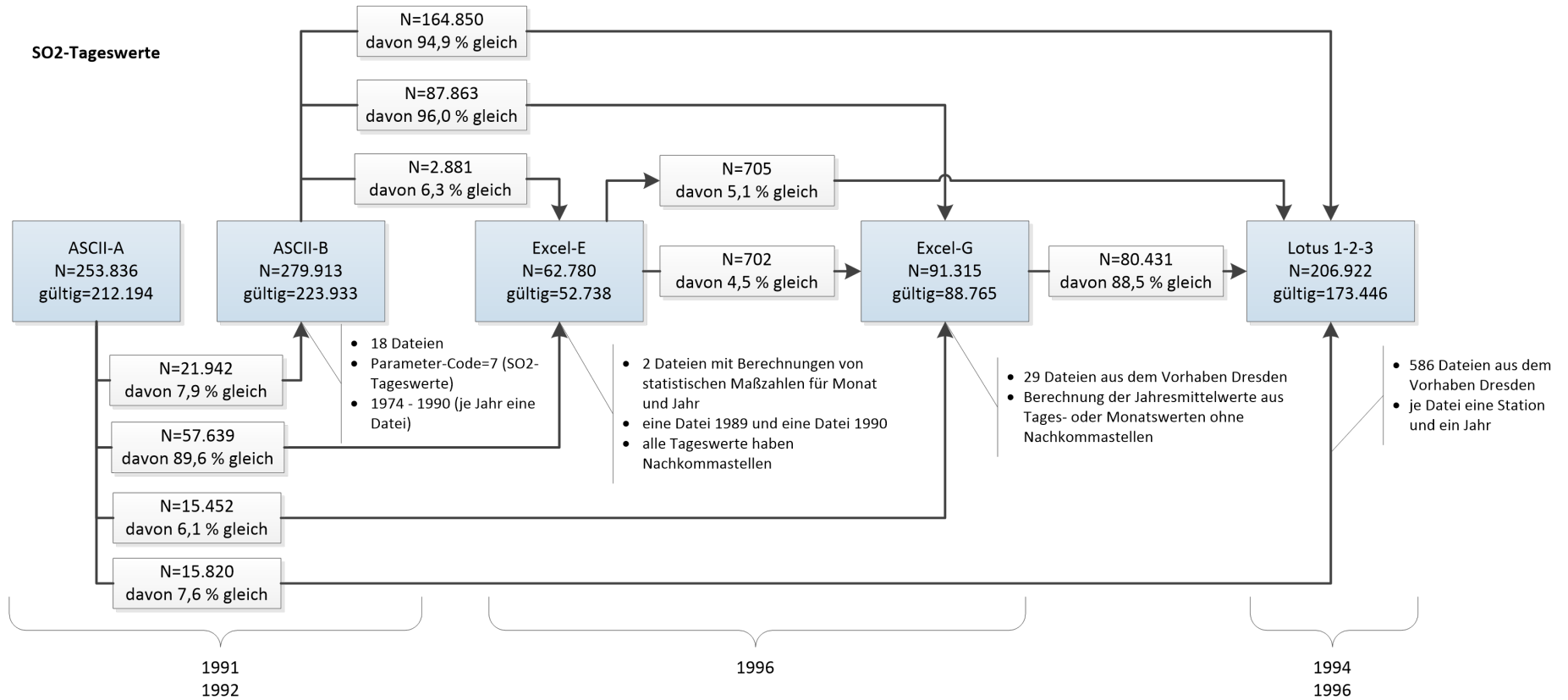
Tabelle 31: Datenquellen der SO₂-Tageswerte

Datenquelle	Jahr der Erstellung	Anzahl Datensätze	Anzahl gültiger Werte
ASCII Typ-A	1991 oder 1992	253.836	212.194
ASCII Typ-B	1991 oder 1992	279.913	223.933
Excel Typ-E	1996	62.780	52.738
Excel Typ-G	1996	91.315	88.765
Lotus 1-2-3	1994 oder 1996	206.922	173.446

Tabelle 32: Anzahl gleicher und ungleicher SO₂-Tageswerte bezüglich der Datenquellen

Datenquelle-1	Datenquelle-2	Anzahl gleicher Messwerte	Anzahl ungleicher Messwerte	Anzahl Messwerte bei der nur Datenquelle-1 gültig ist	Anzahl Messwerte bei der nur Datenquelle-2 gültig ist
ASCII Typ-A	ASCII Typ-B	2.129	19.813	3.714	1.072
ASCII Typ-A	Excel Typ-E	54.957	2.682	3.576	105
ASCII Typ-A	Excel Typ-G	1.011	14.441	124	864
ASCII Typ-A	Lotus 1-2-3	1.285	14.535	364	713
ASCII Typ-B	Excel Typ-E	276	2.605	317	1.182
ASCII Typ-B	Excel Typ-G	87.161	702	470	2.402
ASCII Typ-B	Lotus 1-2-3	162.550	2.300	5.774	577
Excel Typ-E	Excel Typ-G	33	669	0	28
Excel Typ-E	Lotus 1-2-3	37	668	1	24
Excel Typ-G	Lotus 1-2-3	77.804	2.627	7.273	169

Abbildung 15: Zusammenhang der SO₂-Tageswerte bezüglich der Datenquellen



Quelle: eigene Darstellung, BioMath

In den Dateien des ASCII Typ-A werden die SO₂-Tageswerte explizit als SO₂-Tagesmittelwerte bezeichnet. Für die anderen Dateien ist nicht klar, ob es sich um Mittelwerte oder Tageswerte, d. h. Messwerte, die über einen gesamten Tag ermittelt wurden, handelt.

Bezüglich der Gleichheit der Werte konnten zwei Cluster gebildet werden:

- ▶ SO₂-Tagesmittelwerte (ASCII Typ-A und Excel Typ-E)
- ▶ SO₂-Tageswerte (ASCII Typ-B, Excel Typ-G und Lotus 1-2-3)

Innerhalb der beiden Cluster stimmen die Werte zu über 85 % überein. Diese etwas geringere Übereinstimmung kann daran liegen, dass in Excel und Lotus gerundete Werte gespeichert wurden.

4.3.4 SO₂-Messwerte zu Zeitpunkten

Die SO₂-Messwerte zu Zeitpunkten waren in sechs Datenquellen vorhanden (siehe Tabelle 33). Die Zusammenhänge der SO₂-Messwerte zu Zeitpunkten bezüglich der Datenquellen sind in Tabelle 34 aufgelistet und in Abbildung 16 dargestellt.

Tabelle 33: Datenquellen der SO₂-Messwerte zu Zeitpunkten

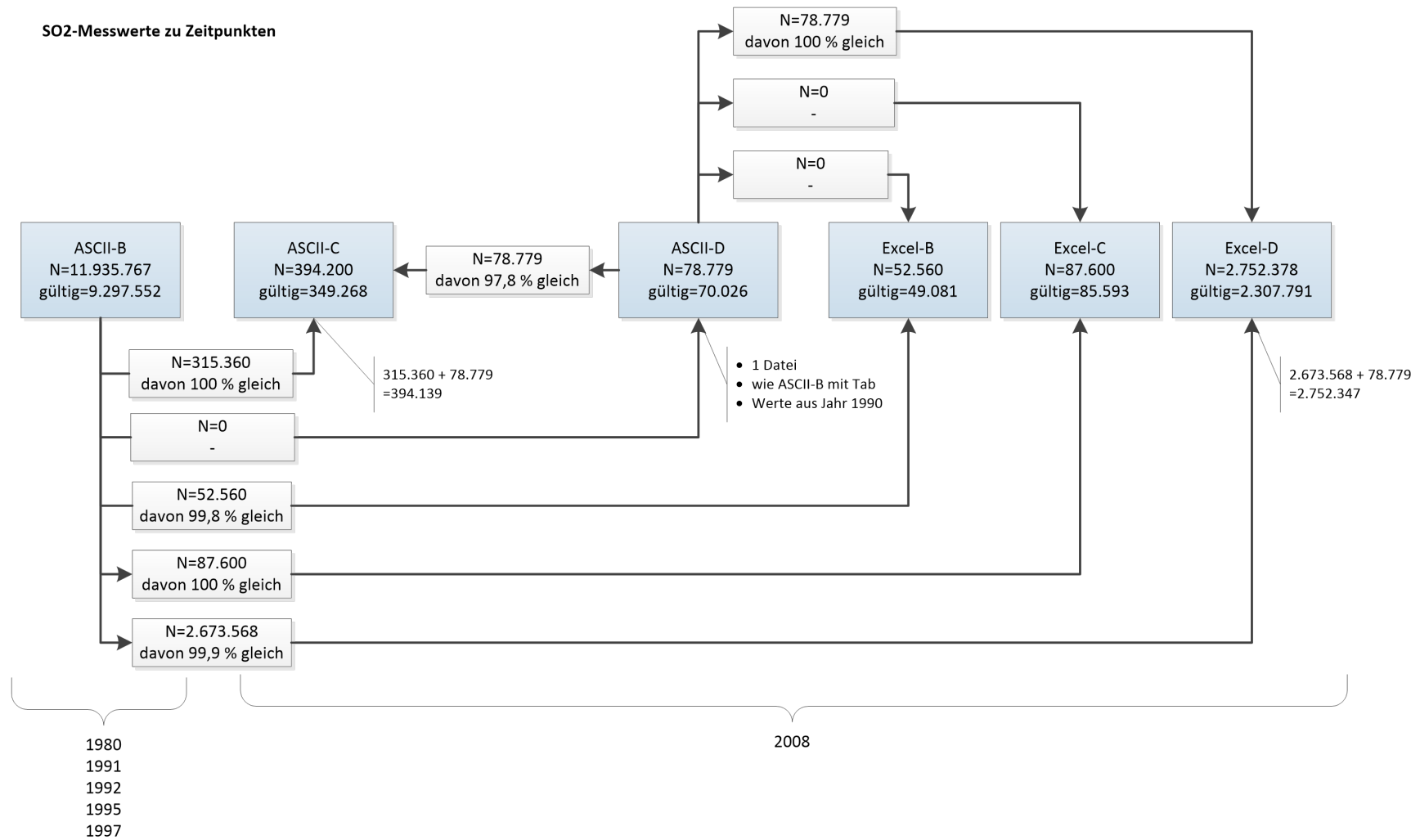
Datenquelle	Jahr der Erstellung	Anzahl Datensätze	Anzahl gültiger Werte
ASCII Typ-B	1980, 1991, 1992, 1995 oder 1997	11.935.767	9.297.552
ASCII Typ-C	2008	394.200	349.268
ASCII Typ-D	2008	78.779	70.0268
Excel Typ-B	2008	52.560	49.081
Excel Typ-C	2008	87.600	85.593
Excel Typ-D	2008	2.752.378	2.307.791

Tabelle 34: Anzahl gleicher und ungleicher SO₂-Messwerte zu Zeitpunkten bezüglich der Datenquellen

Datenquelle-1	Datenquelle-2	Anzahl gleicher Messwerte	Anzahl ungleicher Messwerte	Anzahl Messwerte bei der nur Datenquelle-1 gültig ist	Anzahl Messwerte bei der nur Datenquelle-2 gültig ist
ASCII Typ-B	ASCII Typ-C	315.360	0	0	0
ASCII Typ-B	ASCII Typ-D	0	0	0	0
ASCII Typ-B	Excel Typ-B	52.444	116	116	0
ASCII Typ-B	Excel Typ-C	87.600	0	0	0
ASCII Typ-B	Excel Typ-D	2.673.544	6	13	5
ASCII Typ-C	ASCII Typ-D	77.076	1.665	5	33
ASCII Typ-C	Excel Typ-B	0	0	0	0

Datenquelle-1	Datenquelle-2	Anzahl gleicher Messwerte	Anzahl ungleicher Messwerte	Anzahl Messwerte bei der nur Datenquelle-1 gültig ist	Anzahl Messwerte bei der nur Datenquelle-2 gültig ist
ASCII Typ-C	Excel Typ-C	0	0	0	0
ASCII Typ-C	Excel Typ-D	392.464	1.665	8	33
ASCII Typ-D	Excel Typ-B	0	0	0	0
ASCII Typ-D	Excel Typ-C	0	0	0	0
ASCII Typ-D	Excel Typ-D	78.779	0	0	0
Excel Typ-B	Excel Typ-C	0	0	0	0
Excel Typ-B	Excel Typ-D	0	0	0	0
Excel Typ-C	Excel Typ-D	0	0	0	0

Abbildung 16: Zusammenhang der SO₂-Messwerte zu Zeitpunkten bezüglich der Datenquellen



Für eine bessere Übersichtlichkeit sind nur die wichtigsten Zusammenhänge dargestellt.

Quelle: eigene Darstellung, BioMath

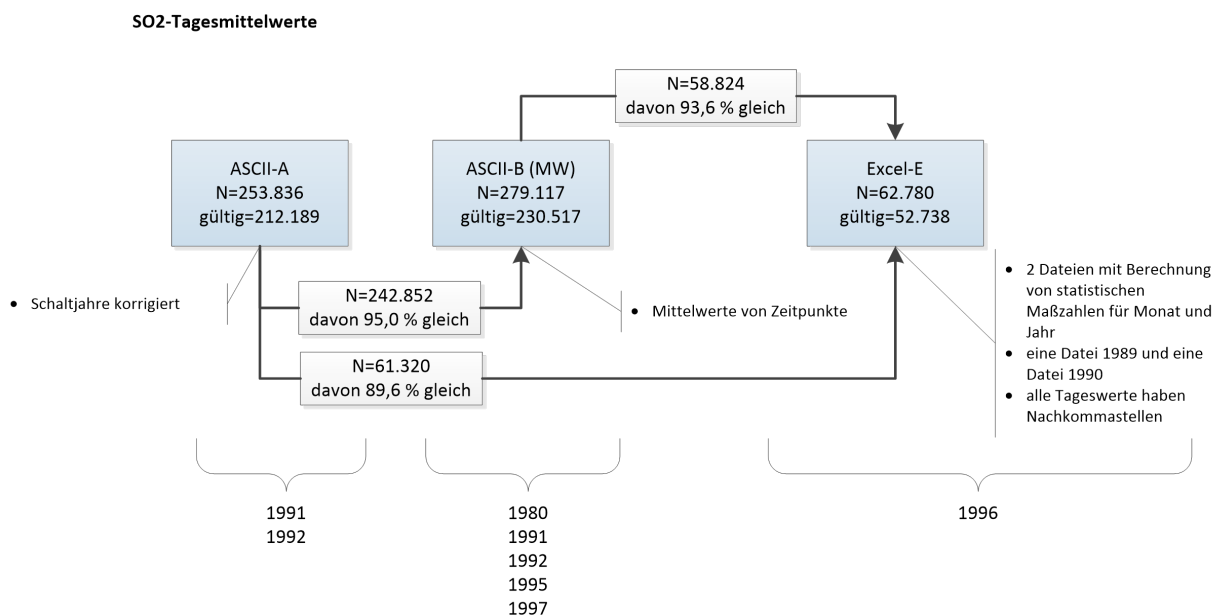
Für die SO₂-Stunden- und Halbstundenwerte gibt es sehr klare Abgrenzungen bezüglich der Datenquellen. Wenn Werte in unterschiedlichen Datenquellen vorhanden waren, stimmten sie zu fast 100 % überein. Weiterhin konnte festgestellt werden, dass sich alle SO₂-Stunden- und Halbstundenwerte auf die Dateien vom ASCII Typ-B und ASCII Typ-D zurückführen lassen.

4.3.5 Zeitliche Auflösungen im Zusammenhang

Für die SO₂-Stunden- und Halbstundenwerte aus den Dateien vom ASCII Typ-B und ASCII Typ-D wurden die SO₂-Tagesmittelwerte gebildet und diese mit denen aus den Dateien vom ASCII Typ-A verglichen. Dabei stellte sich heraus, dass es eine sehr hohe Übereinstimmung in den Jahren gab, die keine Schaltjahre waren. In einer weiteren Analyse wurden die SO₂-Tagesmittelwerte für die Schaltjahre näher untersucht. Im Ergebnis wurde festgestellt, dass in den Dateien vom ASCII Typ-A die SO₂-Tagesmittelwerte ab dem 29. Februar um einen Tag verschoben waren. D. h. die SO₂-Tagesmittelwerte vom 29. Februar waren dem 1. März, die vom 1. März dem 2. März usw. zugeordnet.

Nach der Korrektur der SO₂-Tagesmittelwerte aus den Dateien des ASCII Typ-A wurde der Zusammenhang der SO₂-Tagesmittelwerte bezüglich der drei schaltjahrrelevanten Datenquellen analysiert (siehe Abbildung 17).

Abbildung 17: Zusammenhang der SO₂-Tagesmittelwerte bezüglich der schaltjahrrelevanten Datenquellen



Quelle: eigene Darstellung, BioMath

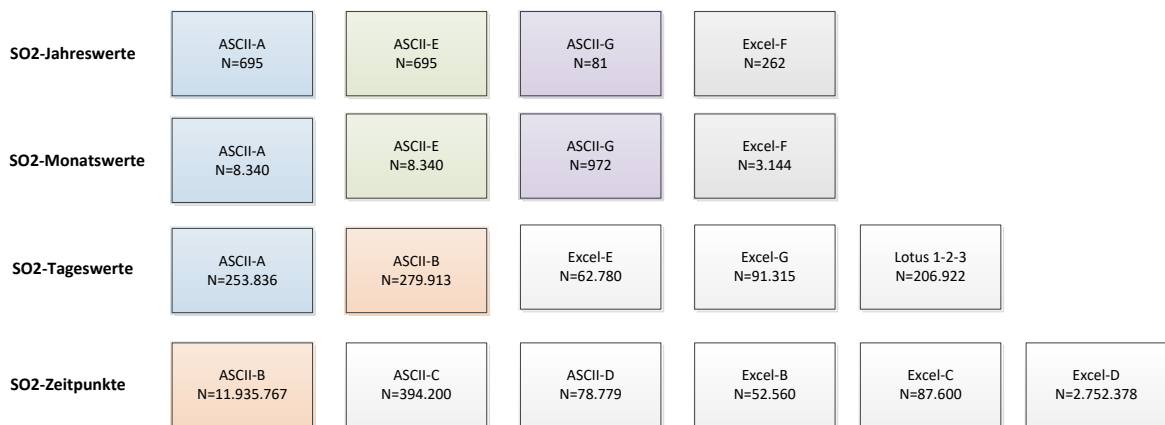
Der überwiegende Teil der SO₂-Tagesmittelwerte bezüglich der Dateien des ASCII Typ-A und Excel Typ-B kann aus den SO₂-Stunden- und Halbstundenwerten berechnet werden. Für 189 Stationen sind SO₂-Stunden- und Halbstundenwerte für mindestens 1 und maximal für 20 Jahre vorhanden. Unter Berücksichtigung der SO₂-Stunden- und Halbstundenwerte der Datei des ASCII Typ-D sind lediglich für 17 Stationen für das Jahr 1990 SO₂-Tagesmittelwerte vorhanden.

4.3.6 Ergebnis der Harmonisierung

Ausgangspunkt der Harmonisierung der SO₂-Messwerte waren die in Abbildung 18 dargestellten Datenquellen für die zeitlichen Auflösungen

- ▶ Jahre
- ▶ Monate
- ▶ Tage und
- ▶ Zeitpunkte (Stunden oder halbe Stunden).

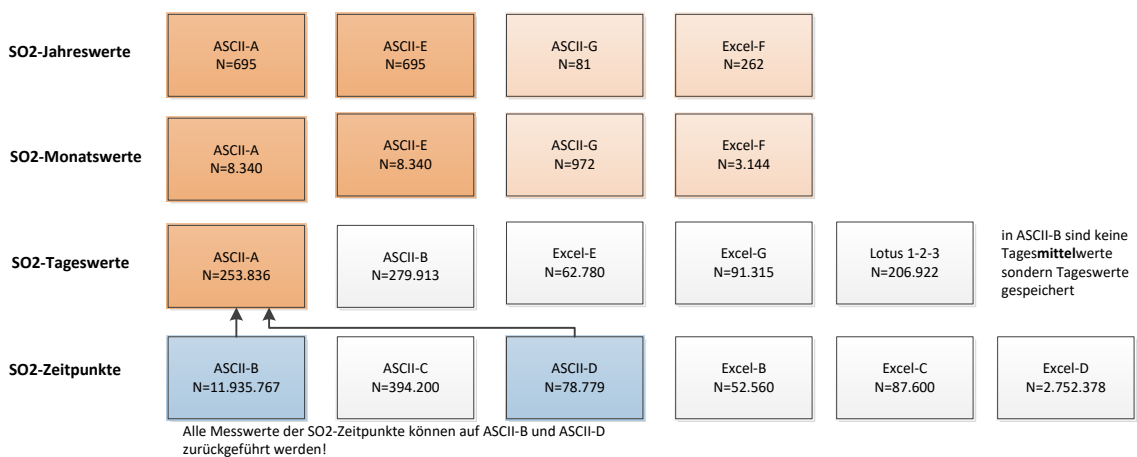
Abbildung 18: Datenquellen der SO₂-Messwerte



Quelle: eigene Darstellung, BioMath

Aus den SO₂-Messwerten zu Zeitpunkten (ASCII Typ-B und ASCII Typ-D) konnten fast alle SO₂-Tagesmittelwerte (ASCII Typ-A) ermittelt werden. Aus diesen SO₂-Tagesmittelwerten wurden dann die SO₂-Monats- und Jahresmittelwerte (ASCII Typ-A und ASCII Typ-E) berechnet (siehe Abbildung 19).

Abbildung 19: Zusammenhang der SO₂-Messwerte bezüglich der Datenquellen



Quelle: eigene Darstellung, BioMath

Eine Besonderheit stellen die SO₂-Tageswerte dar. Sie sind SO₂-Werte, die über 24 Stunden – ggf. tagesübergreifend – ermittelt wurden und keine Tagesmittelwerte sind. Das zeigte sich auch in den Betrachtungen zu den Zusammenhängen der SO₂-Tageswerte bezüglich der Datenquellen (siehe Abschnitt 4.3.3). Alle SO₂-Tageswerte lassen sich auf die Dateien des ASCII Typ-B zurückführen. Die SO₂-Tageswerte der Dateien des Excel Typ-G und Lotus 1-2-3 bilden jeweils eine Teilmenge der SO₂-Tageswerte der Dateien des ASCII Typ-B.

Die SO₂-Werte der Dateien des Excel Typ-E sind Tagesmittelwerte und bilden eine Teilmenge der SO₂-Tagesmittelwerte der Dateien des ASCII Typ-A.

Die weitere Verarbeitung von SO₂-Werten erfolgte nur aus den Datenquellen

- ▶ ASCII Typ-B
- ▶ ASCII Typ-D und ggf.
- ▶ ASCII Typ-A.

4.4 Plausibilitätsprüfung

4.4.1 Datenbasis

Die Plausibilitätsprüfung wurde für die folgenden Schadstoffe und zeitlichen Auflösungen durchgeführt (siehe Tabelle 35). Die SO₂-Tagesmittelwerte wurden nur dann berücksichtigt, wenn für diese Stationen und Jahre keine Stunden- oder Halbstundenwerte vorliegen.

Tabelle 35: Schadstoffe und zeitliche Auflösung für die Plausibilitätsprüfung

Schadstoff	Kürzel	Zeitpunkt Stunden- Halbstunden- werte	Termin	Tageswert	Tagesmittelwert
Chlor	CL			X	
Kohlenmonoxid	CO	X			
Schwefelkohlenstoff	CS2			X	
Fluor	FL			X	
Schwefelwasserstoff	H2S	X		X	
Stickstoffmonoxid	NO	X		X	
Stickstoffdioxid	NO2	X	X	X	
Stickoxide	NOx	X		X	
Ozon	O3	X	X		
Blei	PB			X	
Schwefeldioxid	SO2	X	X	X	X
Sulfat	SO4			X	
Schwebstaub	SWS	X	X	X	

Je Schadstoff und zeitlicher Auflösung wurden Textdateien im Format wie in Tabelle 26 beschrieben erzeugt, wobei der Inhalt der Variable Messparameter entsprechend dem Schadstoff belegt wurde.

4.4.2 Plausibilität

Im ersten Schritt wurden Messreihen von Stationen, deren Stationsnummern nicht in den Metadaten referenziert waren, entfernt.

Im Prozess der Standardisierung wurden die in den Daten definierten Fehlwerte und die leeren Dateninhalte in einheitliche Fehlwerte überführt.

Bei allen Schadstoffen sind negative Werte nicht erlaubt. In den Daten für die Plausibilitätsprüfung waren nur für die definierten Fehlwerte negative Werte vorhanden.

Eine obere Grenze der Werte konnte nicht festgelegt werden, auch nicht für SO₂. In den Betriebsanleitungen des Messgerätes CM5 für SO₂ gibt es zwar Angaben zu den Messbereichen (0 ... 0,45 mg SO₂/m³, 0 ... 1,5 mg SO₂/m³ oder 0 ... 4,5 mg SO₂/m³), aber einige Messreihen enthielten auch aufeinanderfolgende Werte, die größer waren als die obere Grenze der Messbereiche. Das betraf insbesondere die drei Stationen 225902-Erfurt Platz Pariser K., 182130-Mölbis und 182225-Leipzig-Nord mit Werten über 4,5 mg/m³.

Für die Plausibilitätsprüfung von Nullwerten lagen keine belastbaren Informationen vor, auch nicht für SO₂. In den Betriebsanleitungen des Messgerätes CM5 für SO₂ gibt es zwar die Angabe für die untere Nachweisgrenze von 10 µg/m³, es ist aber nicht bekannt, ob sie für alle Jahre und Stationen gleichermaßen gilt. Weiterhin gibt es SO₂-Messreihen, die Nullwerte und nur Werte in Sprüngen von 5 µg/m³ aufweisen (siehe auch 4.2.11). Es ist also nicht klar, ob Nullwerte Fehlwerte oder gültige Werte sind. Deshalb wurden die Nullwerte als gültige Werte angesehen und bei der Prüfung auf Extremwerte mitberücksichtigt.

Für die Detektion von Werten, die so extrem sind, dass sie sehr wahrscheinlich nicht in die Messreihe gehören, gibt es die Möglichkeiten der Ausreißertests oder die Bewertung der Verteilung der Werte über einen Zeitabschnitt. Bei beiden Verfahren sind Grenzen festzulegen, ab wann ein Wert so extrem ist, dass er nicht zur Messreihe gehört.

Die Betrachtung der Werte in einem Zeitabschnitt von einem Monat ist für die Detektion gut geeignet, da es dann eine Unabhängigkeit vom Jahresverlauf gibt.

Ausreißertests basieren auf Annahmen über die Verteilung der Werte. Diese Annahmen sind bei den vorliegenden Werten nicht immer erfüllt, insbesondere die Symmetrie der Verteilung liegt nicht vor. Auch ein großer Stichprobenumfang z. B. bei Stundenwerten über einen Monat, hat starken Einfluss auf die Trennschärfe solcher Tests.

Besser geeignet ist die Bewertung des 95. oder 99. Perzentils, abhängig von der Anzahl der Werte in einem Monat. Grundlage für die Bewertung ist dann der Absolutbetrag des Quotienten aus dem Wert und dem 95. oder 99. Perzentil – berechnet für den Monat. Für diese Quotienten Q₉₅ oder Q₉₉ ist dann ein Wert Q_{CutOff} festzulegen, ab dem ein Wert so extrem ist, dass er sehr wahrscheinlich nicht zur Messreihe gehört.

Die Detektion von Extremwerten wurde wie folgt durchgeführt:

1. Festlegung des Schadstoffes und dessen zeitlicher Auflösung (siehe Tabelle 35)
2. Festlegung des Perzentils (siehe Tabelle 36)
3. Berechnung der Perzentile je Station, Jahr und Monat
4. Berechnung der Q9x-Werte für jeden Messwert
5. Berechnung der Perzentile je Station, Jahr und Monat ohne Nullwerte
6. Berechnung der Q9x-Werte für jeden Messwert ohne Nullwerte
7. Berechnung des Maximums der beiden Q9x-Werte
8. Festlegung vom Q_{CutOff}
 - a. Sortieren der Messwerte absteigend nach den Q9x-Werten
 - b. Absteigende Bewertung des jeweiligen Messwertes durch Betrachtung der 10 Werte vor und nach dem Messwert mit dem Kriterium, ob es möglich ist, dass ein so extremer Wert in die Messreihe an dieser Stelle plausibel wäre
 - c. Kann der Messwert als plausibel angenommen werden, so wird der Q9x-Wert als Q_{CutOff} festgelegt
9. alle Messwerte, deren Q9x-Wert größer ist als Q_{CutOff} , werden auf Fehlwerte gesetzt

Bei der Festlegung von Q_{CutOff} wurde auch berücksichtigt, dass bei Tages- und Terminwerten durchaus hohe Sprünge im Messwertverlauf möglich sind. Die Festlegung des Q_{CutOff} -Wertes erfolgte immer konservativ, d. h. es wurde sich eher für die Plausibilität eines Messwertes entscheiden als dagegen.

Die Werte für Q_{CutOff} sind in Tabelle 36 gelistet. Werte die größer sind als Q_{CutOff} , wurden auf Fehlwerte gesetzt.

Tabelle 36: Q_{CutOff} je Schadstoff und zeitlicher Auflösung

Schadstoff	Kürzel	zeitliche Auflösung	Perzentil	Q_{CutOff}	Anzahl gültige Werte	Anzahl Werte größer Q_{CutOff}
Chlor	CL	Tageswert	95	2,5	3.556	4
Kohlenmonoxid	CO	Zeitpunkt	95	2,5	17.3504	12
Schwefelkohlenstoff	CS2	Tageswert	95	2,0	1.458	0
Fluor	FL	Tageswert	95	3,0	6.950	5
Schwefelwasserstoff	H2S	Zeitpunkt	99	11,0	63.369	0
Schwefelwasserstoff	H2S	Tageswert	95	2,0	831	0
Stickstoffmonoxid	NO	Zeitpunkt	99	3,0	24.019	0
Stickstoffmonoxid	NO	Tageswert	95	2,5	4.703	7
Stickstoffdioxid	NO2	Zeitpunkt	99	4,0	23.913	0
Stickstoffdioxid	NO2	Termin	95	3,0	4.887	4
Stickstoffdioxid	NO2	Tageswert	95	2,0	29.820	65
Stickoxide	NOx	Zeitpunkt	99	2,2	150.778	4

Schadstoff	Kürzel	zeitliche Auflösung	Perzentil	Q _{CutOff}	Anzahl gültige Werte	Anzahl Werte größer Q _{CutOff}
Stickoxide	NO _x	Tageswert	95	2,1	2.724	11
Ozon	O ₃	Zeitpunkt	99	2,0	760.519	0
Ozon	O ₃	Termin	95	3,5	86.976	2
Blei	PB	Tageswert	95	1,9	997	2
Schwefeldioxid	SO ₂	Zeitpunkt	99	6,3	9.326.163	89
Schwefeldioxid	SO ₂	Termin	95	7,5	206.252	13
Schwefeldioxid	SO ₂	Tageswert	95	4,4	223.933	35
Schwefeldioxid	SO ₂	Tagesmittelwert	95	3,0	238.911	0
Sulfat	SO ₄	Tageswert	95	3,5	2.793	0
Schwebestaub	SWS	Zeitpunkt	99	2,7	124.873	84
Schwebestaub	SWS	Termin	95	2,0	2.273	0
Schwebestaub	SWS	Tageswert	95	2,9	206.344	24

Im Ergebnis der Plausibilitätsprüfungen wurden Textdateien für jeden Schadstoff und jede zeitliche Auflösung erstellt (siehe Tabelle 37).

Tabelle 37: Finale Ergebnisdateien

Variablenname	Bedeutung
IDDatei	Eindeutige Nummer zur Identifikation der Rohdatendatei
DateiTyp	Typ der Rohdatendatei
StationNummer	Stationsnummer
Messparameter	Bezeichnung des Messparameters (z. B. SO ₂ -Zeitpunkt, PB-Tag, NO ₂ -Termin)
Jahr	Jahr als Zahl vierstellig
Monat	Monat als Zahl
Tag	Tag als Zahl
Stunde	Stunde als Zahl
Minute	Minute als Zahl
Wert	Messwert ohne Nachkommastellen
TagImJahr	Tag im Jahr als Zahl
WochentagBez	Bezeichnung des Wochentages zweistellig (Mo, Di, Mi, Do, Fr, Sa oder So)
MonatBez	Bezeichnung des Monats dreistellig (Jan, Feb, Mrz, Apr, Mai, Jun, Jul, Aug, Sep, Okt, Nov oder Dez)

Variablenname	Bedeutung
Q95 oder Q99	Absolutbetrag des Quotienten aus dem Wert und dem 95. oder 99. Perzentil bezüglich des Monats
QCutOff	Festgelegter CutOff-Wert bezüglich Q95 oder Q99 zur Fehlwertsetzung
WertFinal	Finaler Wert
FehlwertFlag	Kennzeichen des Fehlwertes für WertFinal 1 = Wert wurde in den Rohdaten als Fehlwert gekennzeichnet 2 = Q95 oder Q99 ist größer als der CutOff-Wert 3 = Wert wurde manuell als Fehlwert gekennzeichnet

Der Separator ist das Semikolon und das Dezimaltrennzeichen ist der Punkt.

Eine deskriptive Übersicht der finalen Ergebnisdateien sind in der Datei *Übersicht validierte finale Daten.xlsx* zu finden.

Abschließend wurden die relevanten Ergebnisdaten so aufbereitet, dass sie in die Immissionsdatenbank am UBA eingelesen werden konnten. Alle Daten werden vom UBA zur Verfügung gestellt (immission@uba.de).

A Anhang

A.1 Stationen

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
1	30507	R-ROSG	Rostock Gertr. Straße	BHI-R	NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
2	30508	R-ROSW	Rostock Wielandstraße	BHI-R	SO ₂ -Tag
3	30608	R-ROSS	Rostock Schr.-Platz	BHI-R	SWS-Tag
4	31105	WR-R	Warnemünde	MD	SWS-Tag
5	32643	ZN-R	Zingst	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
6	40623	R-GREI	Greifswald	BHI-R	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
7	40624	GR-R	Greifswald MS	MD	SO ₂ -Tag, SWS-Tag
8	44126	AR-R	Arkona	MD	NO-Tag, NO ₂ -Tag, O ₃ -Termin, O ₃ -Zeit, SO ₂ -Tag, SO ₂ -Termin, SO ₄ -Tag, SWS-Tag
9	62441	BI-S	Boizenburg	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
10	70046	S-ZZWW	Wittenberge Zellw	BHI-S	SO ₂ -Termin
11	70148	WI-S	Weisen/Wittenberge	MD	SO ₂ -Termin, SWS-Tag
12	73824	S-KHIS	Schwerin Pieckstraße	BHI-S	SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
13	73825	S-SCHW	Schwerin	BHI-S	SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
14	73826	S-BORN	Schwerin Bornhoev	BHI-S	SO ₂ -Termin
15	73923	SW-S	Schwerin MS	MD	SO ₂ -Tag, SWS-Tag
16	75914	BO-R	Boltenhagen	MD	SO ₂ -Tag, SO ₂ -Termin, SWS-Tag
17	84637	TE-N	Teterow	MD	SWS-Tag
18	84711	S-GEPP	Güstrow Parr.Platz	BHI-S	SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
19	84811	S-GPK	Güstrow Grabenstr.	BHI-S	SO ₂ -Termin
20	90260	AN-F	Angermünde	MD	SO ₂ -Tag, SWS-Tag
21	90902	NG-P	Neuglobsow	MD	NO-Tag, NO ₂ -Tag, O ₃ -Zeit, SO ₂ -Tag, SO ₂ -Termin, SO ₄ -Tag, SWS-Tag
22	92005	N-NTZ1	BHI-Neustrelitz	BHI-N	NO _x -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
23	92105	NS-N	Neustrelitz	MD	SWS-Tag
24	93059	N-PW1	Pasewalk KHI	BHI-N	SO ₂ -Tag

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
25	93312	NB-N	Neubrandenburg	MD	SWS-Tag
26	93318	n-NBG3	Neubrandenburg R-Kochstr	BHI-N	SO ₂ -Tagesmittel, SO ₂ -Zeit
27	93417	N-BG1	Neubrandenburg Möwenstr	BHI-N	SO ₂ -Tag
28	93518	N-BG2	Neubrandenburg Industriegelände	BHI-N	SO ₂ -Tag
29	95141	N-ANKL	Anklam KHI	BHI-N	SO ₂ -Tag
30	100415	F-WPS	Schwedt WPS	PKSCHW	SO ₂ -Tag
31	100418	F-SCHW	Schwedt WPS	PKSCHW	NO _x -Tag, SO ₂ -Tag
32	104504	UE-N	Ückermünde	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
33	114251	HN-M	Hanum	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
34	120144	M-HISK	KHI Schönebeck	BHI-M	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
35	120214	M-OSCH	Oschersleben Pharm.	BHI-M	SO ₂ -Tag
36	120735	MG-M	Magdeburg	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
37	120740	M-PFEI	Magdeburg Pfeiff.Stiftungen	BHI-M	SO ₂ -Tag
38	120837	M-HIMA	KHI-Magdeburg	BHI-M	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Zeit
39	120839	M-MGLD	Magdeburg Labor	BHI-M	CO-Zeit, NO-Zeit, NO ₂ -Zeit, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
40	120938	M-MNEU	Magdeburg Neustadt	BHI-M	NO-Zeit, NO ₂ -Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit
41	121006	M-SOBG	Sommerschenburg	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
42	121011	UM-M	Ummendorf	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
43	121651	M-AHBU	Burg Jacobistraße	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
44	121825	M-ASHA	Haldensleben AGR	BHI-M	SO ₂ -Tag
45	123124	GA-M	Gardelegen	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
46	123651	M-KIST	Stendal Kindergarten	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
47	123751	M-HIST	KHI Stendal	BHI-M	SO ₂ -Tag
48	125444	SE-M	Seehausen/Altmark	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Termin, SWS-Tag
49	132533	P-06BR	Brandenburg	BHI-P	SO ₂ -Tagesmittel, SO ₂ -Zeit
50	135449	NR-P	Neuruppin	MD	SWS-Tag

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
51	140511	P-07LW	Luckenwalde KHI	BHI-P	SO ₂ -Tag
52	142205	PR-P	Potsdam SOR	MD	O ₃ -Zeit
53	142302	P-01PD	Potsdam BHI	BHI-P	NO ₂ -Tag, NO _x -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
54	142304	PO-P	Potsdam	MD	NO-Tag, NO ₂ -Tag, O ₃ -Termin, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
55	142332	SF-P	Berlin-Schönefeld	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
56	142404	P-02PD	Potsdam DRK	BHI-P	SO ₂ -Tagesmittel, SO ₂ -Zeit
57	142415	P-03TE	Teltow KPK	BHI-P	SO ₂ -Tag
58	142631	B-ADHF	Berlin Adlershof	BHI-B	SWS-Tag
59	142735	B-KOEP	Berlin Köpenick	BHI-B	SO ₂ -Tagesmittel, SO ₂ -Zeit
60	142739	B-MSEE	Berlin Müggelsee	BHI-B	SWS-Tag
61	142832	B-TRO	Berlin TRO	BHI-B	SWS-Tag
62	143025	T2-B	Berlin Alex 200m	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
63	143035	B-KADF	Berlin Kaulsdorf	BHI-B	SWS-Tag
64	143049	F-HENN	Hennickendorf	BHI-F	FL-Tag
65	143124	B-KFZI	Kraftf.-IMM Berlin	WTZ-AP	CO-Zeit
66	143125	FT-B	Berlin Alex	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
67	143126	B-PAST	Berlin Parochialstr.	BHI-B	CO-Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Zeit
68	143127	B-PAR4	Berlin Parochialstr. 4.Etage	BHI-B	-
69	143128	B-FRHN	Berlin Friedrichshain	BHI-B	SO ₂ -Tagesmittel, SO ₂ -Zeit
70	143134	B-BIDF	Berlin Biesdorf	BHI-B	-
71	143209	ST-P	Staaken	MD	SO ₂ -Tag
72	143225	B-PRBE	Berlin Prenzlauer Berg	BHI-B	CO-Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Zeit
73	143227	B-LIBG	Berlin Lichtenberg	BHI-B	SWS-Tag
74	143232	B-MARZ	Berlin Marzahn	BHI-B	SO ₂ -Tagesmittel, SO ₂ -Zeit
75	143325	B-PRBG	Berlin Prenzlauer Berg	BHI-B	SO ₂ -Tagesmittel, SO ₂ -Zeit
76	143327	B-NOST	Berlin Nordmarkstraße	BHI-B	SWS-Tag

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
77	143406	P-08FA	Falkensee KPK	BHI-P	SO ₂ -Tag
78	143428	B-WSEE	Berlin Weissensee	BHI-B	SWS-Tag
79	143625	B-PANK	Berlin Pankow	BHI-B	SWS-Tag
80	143830	BU-B	Berlin Buch FIB	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SWS-Tag
81	143930	B-BUCH	Berlin Buch	BHI-B	SWS-Tag
82	144515	P-04OB	Oranienburg KHI	BHI-P	SO ₂ -Tag
83	145921	ZE-P	Zehdenick	MD	SWS-Tag
84	151307	LI-F	Lindenberg	MD	NO-Tag, NO ₂ -Tag, O ₃ -Zeit, SO ₂ -Tag, SWS-Tag
85	152133	F-FFO	Frankfurt/Oder	BHI-F	SO ₂ -Tag, SWS-Tag
86	153333	MN-F	Manschnow	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
87	162419	LN-E	Leinefelde	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SWS-Tag
88	163457	H-STOL	Stolberg/Harz Rathaus	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
89	164550	M-HKZW	Rübeland Zementwerk	BHI-M	CO-Zeit
90	164639	SK-M	Schierke	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
91	164757	M-BLA	Blankenburg Rathaus	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
92	164837	BR-M	Brocken	MD	NO ₂ -Tag, SO ₂ -Tag, SWS-Tag
93	165146	WE-M	Wernigerode	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
94	165147	M-ENWE	Wernigerode Energie	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
95	170002	E-ERFR	Erfurt Rieth	BHI-E	SO ₂ -Tagesmittel, SO ₂ -Zeit
96	170131	E-APOL	Apolda	BHI-E	SO ₂ -Tag
97	171257	H-WFES	Weisenfels, Engelsstraße	BHI-H	-
98	171258	H-WEIS	Weisenfels	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
99	171358	H-WFMS	Weisenfels Merseburger Str	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
100	171753	H-BBED	Braunsbedra	MIOELL	SO ₂ -Tagesmittel, SO ₂ -Zeit
101	171806	H-GKAY	Grosskayna Nordstr.	Leuna	H ₂ S-Zeit
102	172106	H-BDFR	Bad Frankenhausen	BHI-H	-
103	172255	H-Mers	Merseburg	Leuna	SO ₂ -Tagesmittel, SO ₂ -Zeit
104	172259	H-MECH	Merseburg Chr. Str.	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
105	172318	AT-H	Artern	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
106	172359	H-SKOS	Schkopau Obersch.	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
107	172459	H-SCHK	Schkopau	BUNA	SO ₂ -Tagesmittel, SO ₂ -Zeit
108	172509	H-KULP	Kulpenberg	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
109	172658	H-HASH	Halle Silberhöhe	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
110	172659	H-AMM	Ammendorf	BUNA	SO ₂ -Tagesmittel, SO ₂ -Zeit
111	172758	H-HASD	Halle-Süd	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
112	172855	H-HNEO	Halle Neustadt Obersch.	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
113	172858	H-HAVS	Halle Voss-Str.	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
114	172955	H-HNEU	Halle Neustadt	BUNA	SO ₂ -Tagesmittel, SO ₂ -Zeit
115	172958	H-HAFP	Halle Frankenplatz	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Termin
116	173057	H-HABS	Halle Burgstraße	BHI-H	NO _x -Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Termin
117	173157	HA-H	Halle Kröllwitz	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
118	173730	H-GROE	Grossoner Pos	MANSFK	SO ₂ -Tagesmittel, SO ₂ -Zeit
119	173931	H-HETK	Hettstedt Kreiskrankenhaus	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
120	174558	H-KOEM	Köthen Markt 1	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
121	174709	H-Qued	Quedlinburg	BHI-H	-
122	175234	M-FMAS	Stassfurt FMA	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
123	175303	M-HIHA	KHI Halberstadt	BHI-M	SO ₂ -Tagesmittel, SO ₂ -Zeit
124	180309	H-ZEI	Zeitz	HYDRWZ	SO ₂ -Tagesmittel, SO ₂ -Zeit
125	180408	H-ZONU	Zeitz Nuschke Straße	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
126	180705	H-NDF	Naundorf	BKKDEU	SO ₂ -Tagesmittel, SO ₂ -Zeit
127	180830	L-HIBO	Borna	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
128	180848	L-COLD	Colditz	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
129	181223	L-HBOE	Böhlen	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
130	181226	L-ROET	Roetha	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
131	182130	L-MOEL	Mölbis	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
132	181444	L-GRI	Grimma	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
133	181805	H-BDUE	Bad Dürrenberg	Leuna	H ₂ S-Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit
134	181923	L-HIEK	EK Leipzig	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
135	181925	LP-L	Leipzig-Stadt	MD	NO-Zeit, NO ₂ -Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
136	182001	H-LEJA	Leuna Jahnweg	BHI-H	H ₂ S-Zeit
137	182002	H-LEUK	Leuna Kindergrippe	Leuna	SO ₂ -Tagesmittel, SO ₂ -Zeit
138	182019	L-LPWE	Leipzig West	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
139	182020	L-LPZ3	Leipzig Hyg 3	BHI-L	-
140	182021	L-LPSW	Leipzig-Südwest	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
141	182022	L-LPZ3	Leipzig	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Zeit
142	182023	L-LPZ2	Leipzig Hyg 2	BHI-L	CO-Zeit
143	182024	L-LPSO	Leipzig Südost	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
144	182100	H-MRDS	Merseburg Rathaus	Leuna	SO ₂ -Tagesmittel, SO ₂ -Zeit
145	182225	L-LPNO	Leipzig-Nord	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
146	182445	L-WUR	Wurzen	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
147	182503	H-DASP	Leuna-DASPIG	Leuna	-
148	182738	L-EILB	Eilenburg	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
149	183309	H-LABG	Landsberg	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
150	183221	L-DELI	Delitzsch	BHI-L	SO ₂ -Tagesmittel
151	183719	H-BILE	Bitterfeld Leninstraße	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
152	183816	H-SAND	Sandersdorf	CKB	NO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
153	183819	H-CKB	Bitterfeld	CKB	CL-Tag, NO ₂ -Tag, NO ₂ -Termin, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
154	183820	H-BIRA	Bitterfeld Ratsw	CKB	NO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
155	183918	H-GREP	Greppin	CKB	NO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
156	184017	H-Wolf	Wolfen	FKWOLF	CS ₂ -Tag, H ₂ S-Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
157	184116	H-WONO	Wolfen-Nord	BHI-H	-
158	184144	H-BDSG	Bad Schmiedeberg	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
159	185016	H-DESS	Dessau	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
160	185239	H-WITT	Wittenberg	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
161	185336	H-PIWN	Piesteritz Waldstraße	STWP	NO _x -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
162	185339	WT-H	Wittenberg MD	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
163	190039	D-FREI	Freital	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
164	190040	D-FDRK	Freit, DRK	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
165	190105	K-MOOS	Moosheim	TUSFTK	SO ₂ -Tag
166	190144	DS-D	Dresden Südhöhe	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
167	190245	D-DER	Dresden	BHI-D	NO ₂ -Tag, NO _x -Tag, PB-Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag, SWS-Zeit
168	190247	ST-D	Dresden-Striesen	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
169	190345	DN-D	Dresden Neustadt	MD	O ₃ -Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit
170	190445	D-DREF	Dresden F	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
171	190446	D-DRE2	Dresden Buchw.Str.	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
172	190449	WH-D	Dresden W. Hirsch	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
173	190640	MR-D	Radebeul Messwagen	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
174	190740	D-DRLA	Dresden-Land	BHI-D	SO ₂ -Tag
175	190741	RA-D	Wahnsdorf/Radebeul	MD	NO-Tag, NO-Zeit, NO ₂ -Tag, NO ₂ -Zeit, O ₃ -Termin, O ₃ -Zeit, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SO ₄ -Tag, SWS-Tag
176	190810	L-DOEB	Döbeln	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit
177	191028	D-MEI	Meissen	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
178	191030	D-MEI2	Mei Container	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
179	191806	OS-L	Oschatz	MD	SO ₂ -Tag, SWS-Tag
180	191818	D-RISE	Riesa Sche/Eng Str.	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
181	191819	D-Rie	Riesa	BHI-D	SO ₂ -Tag
182	191832	D-GRO	Großenhain	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
183	192427	D-GRRA	Gröditz Rathaus	ZWGROE	SO ₂ -Tag
184	192527	D-GROE	Gröditz Rathaus	ZWGROE	SO ₂ -Tag
185	192528	D-GRLA	Gröditz Landambulatorium	BHI-D	SO ₂ -Tag
186	192852	C-SHDR	Schwarzheide- Ruhland	SWSCHW	CL-Tag
187	192952	C-SHDW	Schwarzheide West	SWSCHW	CL-Tag
188	192954	C-SHDO	Schwarzheide Ost	SWSCHW	CL-Tag

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
189	193054	C-SHDS	Schwarzheide Schipkau	SWSCHW	CL-Tag
190	193835	DK-C	Doberlug-Kirchhain	MD	SWS-Tag
191	200156	D-GLA	Görlitz Land	BHI-D	SO ₂ -Tag
192	200540	D-LOE2	Löbau Bebelstraße	BHI-D	SO ₂ -Tag
193	200640	D-LOE	Löbau	BHI-D	SO ₂ -Tag
194	200811	D-BI	Bischofswerda	BHI-D	SO ₂ -Tag
195	200959	D-GST2	Görlitz Waggonbau	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
196	201055	D-GST	Görlitz Stadt	BHI-D	SO ₂ -Tag
197	201057	GO-D	Görlitz	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
198	201125	D-BAU2	Bautzen Wenden Graben	BHI-D	SO ₂ -Tag
199	201126	D-BAU	Bautzen	BHI-D	SO ₂ -Tag
200	201606	D-KA2	Kamenz Pl. D. Befreiung	BHI-D	SO ₂ -Tag
201	201706	D-KA	Kamenz	BHI-D	SO ₂ -Tag
202	201849	D-NIE	Niesky	BHI-D	SO ₂ -Tag
203	202526	C-DREI	Dreiweibern	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
204	202542	C-PUB	Publick	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
205	202648	C-RIE	Rietschen	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
206	203243	C-BMUS	Bad Muskau	BHI-C	SO ₂ -Tagesmittel, SO ₂ -Zeit
207	204420	C-COHY	Cottbus Hygiene	BHI-C	SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
208	204435	C-KUN	Kunersdorf	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
209	204520	C-COLI	Cottbus Liebknechtstraße	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
210	204623	C-MER	Merzdorf	ORGREB	SO ₂ -Tagesmittel, SO ₂ -Zeit
211	204719	CO-C	Cottbus	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
212	204739	C-FSAC	Forst-SACRO	KBKJAE	SO ₂ -Tagesmittel, SO ₂ -Zeit
213	204931	C-HEI	Heinersbrück	KBKJAE	SO ₂ -Tagesmittel, SO ₂ -Zeit
214	205126	C-PEI	Peitz	KBKJAE	-
215	205638	C-DEU	Deulowitz	KBKJAE	SO ₂ -Tagesmittel, SO ₂ -Zeit
216	212644	U-HIHI	Hildburghausen, KHI	BHI-U	SO ₂ -Tagesmittel, SO ₂ -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
217	213423	ME-U	Meiningen	MD	NO-Tag, NO ₂ -Tag, O ₃ -Zeit, SO ₂ -Tag, SWS-Tag
218	213425	U-MEHI	Meiningen, KHI	BHI-U	NO _x -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
219	213641	U-SUHI	Suhl, Hygieneinstitut	BHI-U	SO ₂ -Termin, SWS-Tag
220	213809	KA-U	Kaltennordheim	MD	O ₃ -Termin, O ₃ -Zeit, SO ₂ -Tag, SO ₂ -Termin, SWS-Tag
221	213940	U-ZMWL	Zella-Mehlis WLAG	BHI-U	SO ₂ -Tag
222	213946	SM-U	Schmücke	MD	NO-Tag, NO ₂ -Tag, O ₃ -Termin, O ₃ -Zeit, SO ₂ -Tag, SWS-Tag
223	214155	U-ILTG	Ilmenau Techn. Glas	BHI-U	SO ₂ -Tag
224	214327	U-SMHI	Schmalkalden, KHI	BHI-U	SO ₂ -Tag
225	214907	U-MEBS	Merkers, Berufsschule	BHI-U	-
226	214913	U-BASA	Bad Salzungen	KBW	SO ₂ -Tagesmittel, SO ₂ -Zeit
227	214921	U-BLHG	Bad Liebenstein	BHI-U	SO ₂ -Tag
228	215057	E-ARNS	Arnstadt	BHI-E	SO ₂ -Tag, SWS-Tag
229	215110	U-TORT	Tiefenort, Bez. Suhl	KBW	SO ₂ -Tagesmittel, SO ₂ -Zeit
230	215128	IN-U	Großer Inselsberg	MD	O ₃ -Termin, SO ₂ -Tag, SO ₂ -Termin, SWS-Tag
231	215206	U-SPRI	Springen, Bez. Sul	KBW	SO ₂ -Tagesmittel, SO ₂ -Zeit
232	215422	E-RUHL	Ruhla	BHI-E	-
233	215742	E-GOTH	Gotha	BHI-E	SO ₂ -Tag
234	215958	EB-E	Erfurt-Bindersleben	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
235	222110	U-SOHI	Sonneberg, KHI	BHI-U	SO ₂ -Tag
236	222311	SB-U	Sonneberg-Neufang	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
237	222728	BG-G	Brennersgrün	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
238	223008	NH-U	Neuhaus/Rennweg	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
239	223108	U-NERA	Neuhaus, Apotheke	BHI-U	NO _x -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
240	223449	SZ-G	Schleiz	MD	SO ₂ -Tagesmittel, SO ₂ -Zeit
241	224325	U-SA	Saalfeld	BHI-U	SO ₂ -Tagesmittel, SO ₂ -Zeit
242	225417	BB-E	Bad Berka	MD	SWS-Tag
243	225635	G-Jena	KHI-Jena-Stadt	BHI-G	SO ₂ -Tagesmittel, SO ₂ -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
244	225801	E-ERFK	Erfurt Karth. Str.	BHI-E	SO ₂ -Tagesmittel, SO ₂ -Zeit
245	225901	E-ERFN	Erfurt-Nord	BHI-E	SO ₂ -Tagesmittel, SO ₂ -Zeit
246	225902	E-ERFP	Erfurt Platz Pariser K.	BHI-E	NO _x -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
247	225920	E-WEIM	Weimar	BHI-E	SO ₂ -Tagesmittel, SO ₂ -Zeit
248	232636	K-CAFD	Carlsfeld	TUSFTK	SO ₂ -Tag
249	232637	CA-K	Carlsfeld MD	MD	O ₃ -Zeit, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
250	232657	FI-K	Fichtelberg	MD	O ₃ -Termin, O ₃ -Zeit, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SWS-Tag
251	232859	K-KRRO	Kretscham-Rothens.	TUSFTK	SO ₂ -Tag
252	232908	P2-K	Plauen2	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
253	233008	K-PLRA	Plauen Rathaus	HI-ZW	SO ₂ -Tagesmittel, SO ₂ -Zeit
254	233009	PL-K	Plauen	MD	SO ₂ -Tag, SO ₂ -Termin, SWS-Tag
255	233542	K-AUSB	Aue Straße d. Befreiung	HI-ZW	SO ₂ -Tagesmittel, SO ₂ -Zeit
256	233643	AU-K	Aue	MD	NO-Zeit, NO ₂ -Zeit, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
257	233912	G-GRZ	Greiz, KHI	BHI-G	SO ₂ -Tagesmittel, SO ₂ -Zeit
258	234157	K-GEL	Gelenau	TUSFTK	SO ₂ -Tag
259	234330	K-ZWIH	Zwickau Hygiene	HI-ZW	SO ₂ -Tagesmittel, SO ₂ -Zeit
260	234629	K-CROB	Crossen Brückenstraße	SDAGWK	SO ₂ -Tagesmittel, SO ₂ -Zeit
261	234843	K-HERA	Hohenstein-Ernstthal, Rath.	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
262	234852	KM-K	Karl-Marx-Stadt	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
263	234853	K-PM12	K-M-Stadt, Bretgasse	BHI-K	SWS-Zeit
264	234933	K-GLAU	Glauchau Marxstraße	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
265	234953	K-KMS	Karl-Marx-Stadt	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
266	235053	K-KGAS	K-M-Stadt, Gustav-Adolf-Str.	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
267	235111	G-TALI	Tagebau Lichtenberg	SDAGWK	SO ₂ -Tagesmittel, SO ₂ -Zeit
268	235157	K-EDH	K-M-Stadt Ebersdorf	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
269	235227	K-MEER	Meerane Poliklinik	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
270	235305	G-GERA	Gera	BHI-G	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
271	235308	GE-G	Gera-Leumnitz	MD	NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
272	235405	G-GEFE	Gera Friedrich-Engels-Str.	BHI-G	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit
273	235926	L-ALT	Altenburg	BHI-L	SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Zeit
274	242901	K-BAER	Bärenstein	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
275	243413	K-REI	Reitzenhain	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
276	243501	K-ANN	Annaberg-Buchholz, Rath.	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
277	243518	K-RUEB	Rübenau	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
278	243526	K-DND	Deutschneudorf	BHI-K	SO ₂ -Tagesmittel, SO ₂ -Zeit
279	243811	K-GEB	Gebirge	TUSFTK	SO ₂ -Tag
280	243909	MA-K	Marienberg	MD	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit, SWS-Tag
281	244230	K-CAE	Cämmerswalde	TUSFTK	SO ₂ -Tag
282	244413	K-REID	Reifland	TUSFTK	SO ₂ -Tag
283	244445	ZI-D	Zinnwald-Georgenfeld	MD	NO-Tag, NO ₂ -Tag, SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Termin, SO ₂ -Zeit, SO ₄ -Tag, SWS-Tag
284	244645	D-ALT2	Altenberg Galgental	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
285	244646	D-ALT	Altenberg Zinnerz	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
286	244648	D-GEIS	Geising, Landambulatorium	BHI-D	SO ₂ -Tag
287	244854	D-BREI	Breitenau	TUSFTK	SO ₂ -Tag
288	245056	D-GOBA	Gottleuba	TUSFTK	SO ₂ -Tag
289	245257	D-BAMB	Berggießhübel, Ambulatorium	BHI-D	SO ₂ -Tag
290	245357	D-BERG	Berggießhübel	TUSFTK	SO ₂ -Tag
291	245440	D-DIPP	Dippoldiswalde	BHI-D	SO ₂ -Tag
292	245459	D-HELL	Hellendorf, Bahrtal	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
293	245651	D-DOPL	Dohna, Plo. Höhe	BHI-D	FL-Tag
294	245652	D-DOKI	Dohna Kinderhort	BHI-D	FL-Tag

LN	Stat. Nr.	Kurzbez.	Stationsname	Betreiber	Messparameter
295	245653	D-DOIN	Dohna Investbar.	BHI-D	-
296	245750	D-DOBU	Dohna Burgberg	BHI-D	-
297	245751	D-DOWI	Dohna Winzerhaus	FLWDO	FL-Tag
298	245752	D-DOGM	Dohna Gärtnerei	FLWDO	
299	245753	D-KOET	Köttewitz	BHI-D	FL-Tag
300	245756	D-PIRK	Pirna KSW	KSWPI	CS ₂ -Tag, H ₂ S-Tag
301	245850	D-GORK	Gorknitz	FLWDO	FL-Tag
302	245851	D-DOSP	Dohna Sportplatz	BHI-D	FL-Tag
303	245852	D-HEID	Heidenau	FLWDO	-
304	245853	D-GROS	Grossedlitz	BHI-D	FL-Tag
305	245856	D-PIR	Pirna	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
306	245857	D-PIR2	Pirna Kreispoliklinik	BHI-D	SO ₂ -Tagesmittel, SO ₂ -Zeit
307	245951	D-DORA	Dohna Rathaus	BHI-D	-
308	245952	D-HEI	Heidenau Apotheke	BHI-D	SO ₂ -Tag
309	255001	D-BAHR	Bahratal	TUSFTK	SO ₂ -Tag
310	255048	D-Lue	Lückendorf	TUSFTK	SO ₂ -Tag
311	255141	D-WAZI	Waltersdorf/Zittau	TUSFTK	SO ₂ -Tag
312	255448	D-D-ZI	T1MP Zittau	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
313	255509	D-SCHA	Bad Schandau	BHI-D	SO ₂ -Tag
314	255549	D-ZITT	Zittau	BHI-D	SO ₂ -Tag
315	255816	D-SEB	Sebnitz	BHI-D	SO ₂ -Tag, SO ₂ -Tagesmittel, SO ₂ -Zeit
316	255937	D-NEUG	HKW Neugersdorf	HKWNEU	SO ₂ -Tagesmittel, SO ₂ -Zeit
317	271231	H1-HBS	Halle Burgstraße Innen	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
318	271232	H2-HBS	Halle Burgstraße Außen	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
319	273211	H1-HFP	Halle Frankenplatz Innen	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit
320	273212	H2-HFP	Halle Frankenplatz Außen	BHI-H	SO ₂ -Tagesmittel, SO ₂ -Zeit

A.2 Anlagedateien

Rohdatendateien.xlsx

All ASCII Typ-E (unbekannte Stationen).xlsx

All IBM Lotus.xlsx

Stationen (Metadaten)(Entfernungen).xlsx

Übersicht validierte finale Daten.xlsx

CL_Tag_Final.txt

CO_Zeit_Final.txt

CS2_Tag_Final.txt

FL_Tag_Final.txt

H2S_Tag_Final.txt

H2S_Zeit_Final.txt

NO_Tag_Final.txt

NO_Zeit_Final.txt

NO2_Tag_Final.txt

NO2_Termin_Final.txt

NO2_Zeit_Final.txt

NOx_Tag_Final.txt

NOx_Zeit_Final.txt

O3_Termin_Final.txt

O3_Zeit_Final.txt

PB_Tag_Final.txt

SO2_Tag_Final.txt

SO2_Tagesmittel_Final.txt

SO2_Termin_Final.txt

SO2_Zeit_Final.txt

SO4_Tag_Final.txt

SWS_Tag_Final.txt

SWS_Termin_Final.txt

SWS_Zeit_Final.txt

5 Quellenverzeichnis

Deutsche Demokratische Republik. TGL 39612 - Nutzung und Schutz der Atmosphäre. REINHALTUNG DER LUFT. Begriffe. TGL 39612.

Deutsche Demokratische Republik. TGL 39615/02 - Nutzung und Schutz der Atmosphäre. Manuelle Immissionsmethoden. Allgemeine Forderungen zur Probenahme. TGL 39615/02.

Deutsche Demokratische Republik. TGL 39615/03 - Nutzung und Schutz der Atmosphäre. Manuelle Immissionsmethoden. Bestimmung von Schwefeldioxid in der atmosphärischen Luft. TGL 39615/03.

Deutsche Demokratische Republik. TGL 39614/03 - Nutzung und Schutz der Atmosphäre. Manuelle Emissionsmethoden. Bestimmung von Stickoxiden durch alkalimetrische Titration. TGL 39614/03.

Feister, U. & Warmbt, W. (1987). Long-term measurements of surface ozone in the German Democratic Republic. *Journal of Atmospheric Chemistry*, 5(1). <https://doi.org/10.1007/BF00192500>

Fleischer, P. & Pohl, K. (1995). Aufbereitung der Immissionsmeßreihen des ehemaligen Meteorologischen Dienstes der DDR für wissenschaftliche Auswertungen.

Hůnová, I. (2020). Ambient Air Quality in the Czech Republic: Past and Present. *Atmosphere*, 11(2), 214. <https://doi.org/10.3390/atmos11020214>

Lahmann, E. & Werner, H. (1992). Immissionsüberwachung in der DDR : vorveröffentlichter Teilbericht zu "Feststellung und Bewertung von Immissionen - Leitfaden zur Immissionsüberwachung in Deutschland". Forschungsbericht 10402267. Umweltbundesamt.

Landesumweltamt Brandenburg. (1995). Luftqualität 1975 bis 1990 (Studien und Tagungsberichte). Zugriff am 07.06.2022.

Marquardt, W., Bruggemann, E. & Ihle, P. (1996). Trends in the composition of wet deposition: effects of the atmospheric rehabilitation in East-Germany. *Tellus B*, 48(3), 361–371. <https://doi.org/10.1034/j.1600-0889.1996.t01-2-00003.x>

Ratzlaff, U. (1990). Ermittlung von Wahrscheinlichkeitsdichteverteilungen für Konzentrationswerte in Luft- und Niederschlagsproben, insbesondere zur Risikoabschätzung bei Belastungen im Extrembereich. TU Dresden.

Teichert, F. & Warmbt, W. (1955). Ozonuntersuchungen am Meteorologischen Observatorium Wahnsdorf (Abhandlungen des Meteorologischen und Hydrologischen Dienstes der Deutschen Demokratischen Republik). Berlin.

Warmbt, W. (1966). Surface ozone and artificial β - activity in Dresden-Wahnsdorf. *Tellus*, 18(2-3), 441–450. <https://doi.org/10.1111/j.2153-3490.1966.tb00256.x>

Werner, H. (1992a). Wintersmog in der ehemaligen DDR. Teil I: Ursachen, Kriterien und Datengewinnung. *Staub - Reinh. d. Luft*, (52), 119–126.

Werner, H. (1992b). Wintersmog in der ehemaligen DDR. Teil II: Datenanalyse, Diskussion und Ausblick. *Staub - Reinh. d. Luft*, 52, 239–244.

Werner, H. & Meckel, U. (1989). Ambient air monitoring in the GDR and quality assurance measures taken by the State Sanitary Inspectorate. In L. J. Brassler & W. C. Mulder (Hrsg.), *Man and his ecosystem. Proceedings (Internationaler Kongreß: Reinhaltung der Luft, Bd. 8)*. Amsterdam: Elsevier.