

TEXTE

47/2019

Ableitung sichererer Trendaussagen zur Entwicklung der Luftqualität in Deutschland

Abschlussbericht

TEXTE 47/2019

Umweltforschungsplan des
Bundesministeriums für Umwelt,
Naturschutz und nukleare Sicherheit

Forschungskennzahl 3717522540
UBA-FB 002741

Ableitung sicherer Trendaussagen zur Entwicklung der Luftqualität in Deutschland

Abschlussbericht

von

Jörg Schmidtke, Kerstin Schmidt
BioMath GmbH, 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

 /umweltbundesamt.de

 /umweltbundesamt

Durchführung der Studie:

BioMath GmbH
Friedrich - Barnewitz - Str. 8
18119 Rostock

Abschlussdatum:

Oktober 2018

Redaktion:

Fachgebiet II 4.2 Beurteilung der Luftqualität
Susan Kessinger

Publikationen als pdf:

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

ISSN 1862-4804

Dessau-Roßlau, Mai 2019

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

Kurzbeschreibung

Zur Beurteilung der Trends der Luftschadstoffe Ozon, Stickstoffdioxid (NO₂) und Feinstaub (PM₁₀) in Deutschland standen räumlich und zeitliche inhomogene Messnetze zur Verfügung. Die Inhomogenitäten beruhten einerseits auf der ungleichmäßigen geografischen Verteilung und andererseits auf der Schließung, Neueinrichtung oder zwischenzeitlichen Außerbetriebnahme von Messstationen.

Im Rahmen dieses Projektes wurde eine Methodik entwickelt, die unter Berücksichtigung der Inhomogenitäten, die Trends der Luftschadstoffe Ozon, NO₂ und PM₁₀ zuverlässig berechnet. Mit Hilfe dieser Methode wird zunächst eine Mittelwertzeitreihe geschätzt und danach durch lineare Regression der Trend bestimmt. Die Schätzung der Mittelwertzeitreihe erfolgt durch einen Differenzialgleichungsansatz.

Diese Methode wurde in einem R-Programm umgesetzt und kann auch für andere Luftschadstoffe, die sich ähnlich wie Ozon, NO₂ und PM₁₀ verhalten, verwendet werden.

Abstract

To assess trends of the air pollutants ozone, nitrogen oxide (NO₂) and particulate matter (PM₁₀) in Germany, spatially and temporally inhomogeneous monitoring networks were available. The inhomogeneity was caused both by irregular geographic distribution and closure, new installation or decommissioning of monitoring stations.

Within the context of this project, a method was developed to reliably calculate trends of the air pollutants ozone, NO₂ and PM₁₀ by taking into account data inhomogeneity. This method estimates a mean times series and determines the trend by linear regression. Mean time series are estimated by a differential equation approach.

This method was implemented in an R program and also can be applied to other air pollutants developing comparably to ozone, NO₂ and PM₁₀.

Inhaltsverzeichnis

Abbildungsverzeichnis	10
Tabellenverzeichnis	11
Abkürzungsverzeichnis	12
Zusammenfassung	13
Summary	26
1 Einleitung	38
2 Material	38
2.1 Daten	38
3 Analyse	42
3.1 Deskriptive	42
3.1.1 Ozon	43
3.1.1.1 Ozon - Anzahl aller Stationen	46
3.1.1.2 Ozon - Anzahl durchgängig messender Stationen bis 2016	47
3.1.1.3 Ozon - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016	48
3.1.1.4 Ozon - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)	50
3.1.1.5 Ozon - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)	52
3.1.1.6 Ozon - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)	54
3.1.1.7 Ozon - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)	56
3.1.1.8 Ozon - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden	58
3.1.1.9 Ozonwerte in $\mu\text{g}/\text{m}^3$ - alle Messreihen und durchgängig messende Messreihen von 1990 - 2016	60
3.1.1.10 Ozon - Messreihen der jährlichen Mittelwerte in $\mu\text{g}/\text{m}^3$	65
3.1.1.11 Ozon - Boxplot der Mittelwerte ($\mu\text{g}/\text{m}^3$) je Messreihe je Kategorie	66
3.1.1.12 Ozon - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund	67
3.1.1.13 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund	69
3.1.1.14 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah	71
3.1.1.15 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah	73

3.1.1.16	Ozon - Trend: Anstieg der Messreihen ($\mu\text{g}/\text{m}^3$ je Jahr) durchgängig vs. nicht durchgängig messend	75
3.1.2	NO2	77
3.1.2.1	NO2 - Anzahl aller Stationen	79
3.1.2.2	NO2 - Anzahl durchgängig messender Stationen bis 2016	80
3.1.2.3	NO2 - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016	81
3.1.2.4	NO2 - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)	83
3.1.2.5	NO2 - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)	84
3.1.2.6	NO2 - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)	85
3.1.2.7	NO2 - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)	86
3.1.2.8	NO2 - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden	87
3.1.2.9	NO2-Werte in $\mu\text{g}/\text{m}^3$ - alle Messreihen und durchgängig messende Messreihen von 1990 - 2016	89
3.1.2.10	NO2 - Messreihen der jährlichen Mittelwerte in $\mu\text{g}/\text{m}^3$	94
3.1.2.11	NO2 - Boxplot der Mittelwerte in $\mu\text{g}/\text{m}^3$ je Messreihe je Kategorie	95
3.1.2.12	NO2 - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund	97
3.1.2.13	NO2 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund	99
3.1.2.14	NO2 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah	101
3.1.2.15	NO2 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah	103
3.1.2.16	NO2 - Trend: Anstieg der Messreihen ($\mu\text{g}/\text{m}^3$ je Jahr) durchgängig vs. nicht durchgängig messend	105
3.1.3	PM10	107
3.1.3.1	PM10 - Anzahl aller Stationen	109
3.1.3.2	PM10 - Anzahl durchgängig messender Stationen bis 2016	110
3.1.3.3	PM10 - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016	111
3.1.3.4	PM10 - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)	113
3.1.3.5	PM10 - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)	114

3.1.3.6	PM10 - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)	115
3.1.3.7	PM10 - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)	116
3.1.3.8	PM10 - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden	117
3.1.3.9	PM10-Werte in $\mu\text{g}/\text{m}^3$ - alle Messreihen und durchgängig messende Messreihen von 1990 - 2016	119
3.1.3.10	PM10 - Messreihen der jährlichen Mittelwerte in $\mu\text{g}/\text{m}^3$	124
3.1.3.11	PM10 - Boxplot der Mittelwerte ($\mu\text{g}/\text{m}^3$) je Messreihe je Kategorie	125
3.1.3.12	PM10 - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund	126
3.1.3.13	PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund	128
3.1.3.14	PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah	130
3.1.3.15	PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah	132
3.1.3.16	PM10 - Trend: Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe durchgängig vs. nicht durchgängig messend	134
3.2	Methoden der Deskriptiven.....	136
3.2.1	Rangordnung der Messreihen.....	136
3.2.2	lineares Regressionsmodell zur Trendberechnung.....	136
3.3	Ergebnisse der deskriptiven Auswertung.....	138
3.4	Induktiv.....	138
3.4.1	Modell der jährlichen Differenzen.....	139
3.4.1.1	Herleitung des Modells	139
3.4.1.2	Trendberechnung	141
3.4.2	Ozon - Zeitraum 2000-2016.....	142
3.4.2.1	Ozon - ländlicher Hintergrund	145
3.4.2.2	Ozon - städtischer Hintergrund	146
3.4.2.3	Ozon - verkehrsnah	147
3.4.2.4	Ozon - industrienah	148
3.4.3	NO ₂ - Zeitraum 2000-2016.....	149
3.4.3.1	NO ₂ - ländlicher Hintergrund	151
3.4.3.2	NO ₂ - städtischer Hintergrund	152
3.4.3.3	NO ₂ - verkehrsnah	153
3.4.3.4	NO ₂ - industrienah	154
3.4.4	PM10 - Zeitraum 2000-2016.....	155

3.4.4.1	PM10 - ländlicher Hintergrund	157
3.4.4.2	PM10 - städtischer Hintergrund	158
3.4.4.3	PM10 - verkehrsnah	159
3.4.4.4	PM10 - industrienah	160
3.4.5	Ergebnis der induktiven Auswertung.....	161
4	Programm zur Berechnung der Modell-Mittelwertzeitreihen	162
4.1	R-Programm.....	162
4.1.1	Daten/Messreihen der Messtationen	162
4.1.2	R-Programmanpassungen.....	163
4.1.3	Ausgabedateien.....	163
4.1.4	Programmablauf	164
4.2	Berechnungen mit dem R-Programm.....	164
4.2.1	Ozon von 1991 bis 2017	165
4.2.2	Ozon von 2000 bis 2017	166
4.2.3	Ozon von 2008 bis 2017	167
4.2.4	NO2 von 1991 bis 2017	168
4.2.5	NO2 von 2000 bis 2017	169
4.2.6	NO2 von 2008 bis 2017	170
4.2.7	PM10 von 1991 bis 2017	171
4.2.8	PM10 von 2000 bis 2017.....	172
4.2.9	PM10 von 2008 bis 2017.....	173
5	Literaturrecherche	174
6	Anhänge.....	175
6.1	Review Protokoll zur Literaturrecherche.....	175
6.1.1	Review-Strategie	175
6.1.1.1	Hintergrund	175
6.1.1.2	Ziel des Reviews	176
6.1.1.3	Methode	176
6.1.2	Review Dokumentation.....	178
6.1.2.1	Suchergebnisse	178
6.1.2.2	Eignungsprüfung der identifizierten Dokumente	178
6.1.2.3	Extraktion von Informationen und Klassifikation	179
6.2	R-Programmcode	189
7	Quellenverzeichnis.....	193

Abbildungsverzeichnis

Abbildung 1:	Ozon - Karte mit allen Messreihen von 1990-2016.....	39
Abbildung 2:	NO ₂ - Karte mit allen Messreihen von 1990-2016.....	40
Abbildung 3:	PM ₁₀ - Karte mit allen Messreihen von 1990-2016.....	41
Abbildung 4:	Beispiel - Messreihen und berechnete Mittelwertzeitreihe nach Modell.....	141
Abbildung 5:	Jahresmittelwertzeitreihen des Modells für Ozon.....	144
Abbildung 6:	Jahresmittelwertzeitreihen des Modells für NO ₂	150
Abbildung 7:	Jahresmittelwertzeitreihen des Modells für PM ₁₀	156
Abbildung 8:	Verarbeitungsschema R-Programm.....	162

Tabellenverzeichnis

Tabelle 1:	Kategorisierung.....	38
Tabelle 2:	Ozon - Anzahl aller Messreihen	39
Tabelle 3:	NO ₂ - Anzahl aller Messreihen	40
Tabelle 4:	PM ₁₀ - Anzahl aller Messreihen	41
Tabelle 5:	Messreihen UBA-Auswahl	42
Tabelle 6:	abgeleitete Variablen	42
Tabelle 7:	Schema Rangordnung.....	136
Tabelle 8:	Deskriptive Übersicht und Bewertung.....	138
Tabelle 9:	Messwerte	140
Tabelle 10:	Differenzen	140
Tabelle 11:	Wichtung.....	141
Tabelle 12:	Berechnung.....	141

Abkürzungsverzeichnis

Ozon	Luftschadstoff bestehend aus drei Sauerstoffatomen, gemessen in $\mu\text{g}/\text{m}^3$
NO₂	Luftschadstoff Stickstoffdioxid, gemessen in $\mu\text{g}/\text{m}^3$
PM₁₀	Luftschadstoff Feinstaub mit einem aerodynamischen Durchmesser von $\leq 10 \mu\text{m}$, gemessen in $\mu\text{g}/\text{m}^3$
KI	Konfidenzintervall
Eol	Exchange of Information; Exchange on Information decision (Eol) on Air; 97/101/EC: Council Decision of 27 January 1997 establishing a reciprocal exchange of information and data from networks and individual stations measuring ambient air pollution within the Member States

Zusammenfassung

Für die Trendanalyse von Luftschadstoffen stellte das UBA Daten von Ozon, NO₂ und PM₁₀ zur Verfügung. Die Daten beinhalteten Jahresmittelwerte der jeweiligen Schadstoffe in µg/m³ über den Zeitraum von 1990 bis 2016, im Nachgang auch für das Jahr 2017.

Ein Datensatz bezog sich auf eine Messstation und beinhaltete Jahresmittelwerte für den Zeitraum, in dem diese Station in Betrieb war. Jeder Station waren geografischen Koordinaten im Koordinatensystem WGS84 zugeordnet. Weiterhin war jede Station in Stationsumgebung und Art der Station nach EoI (Exchange of Information) klassifiziert. Für die spätere Analyse wurde mit Hilfe dieser Klassifikation eine Kategorisierung jeder Station in

- ▶ ländlicher Hintergrund
- ▶ städtischer Hintergrund
- ▶ verkehrsnah
- ▶ industrienah

durchgeführt.

Die deskriptive Auswertung der Daten der drei Luftschadstoffe umfasste die Beschreibung der vorliegenden Messreihen nach den Gesichtspunkten:

- ▶ Anzahl und Verteilung der Stationen
- ▶ Verlauf der Werte der Messreihen
- ▶ Aussagen linearer Trends der Messreihen mit einfachen Modellen

Die **Anzahl der Stationen** wurde einerseits für alle vorhandenen Stationen und andererseits für ausschließlich durchgängig messende Stationen analysiert. Unter durchgängig messenden Stationen wurden Stationen verstanden, die von Betriebsbeginn bis Betriebsende in jedem Betriebsjahr einen gültigen Jahresmittelwert besaßen. Die Analysen erfolgten für alle Kategorien (ländlicher Hintergrund, städtischer Hintergrund, verkehrsnah, industrienah) zusammen und für jede Kategorie einzeln.

Um die Stationen und somit auch die Messreihen besser charakterisieren zu können, wurden für jede Station die folgenden abgeleiteten Variablen berechnet.

- ▶ Anzahl Messwerte der Station
- ▶ Jahr in dem die Station den ersten Messwert erzeugte (Messbeginn)
- ▶ Jahr in dem die Station den letzten Messwert erzeugte (Messende)
- ▶ Anzahl Jahre zwischen Messbeginn und Messende (Betriebszeit)
- ▶ Anzahl Jahre der Betriebszeit in denen keine Messwerte erzeugt wurden
- ▶ Bewertung der Messreihe

Basierend auf diesen Variablen wurden Balken-, Liniendiagramme und Boxplots erstellt und die Variablen tabellarisch gelistet. Weiterhin wurden die Stationen auf Karten visualisiert und die Anzahl der Stationen bezüglich der Bundesländer/ Betreiber grafisch dargestellt.

Für die Beschreibung des **Verlaufs der Werte der Messreihen** wurden diese in Liniendiagrammen getrennt nach durchgängig messend, nicht durchgängig messend und getrennt nach den Kategorien dargestellt. Weiterhin wurden je Messreihe statistische Maßzahlen berechnet und diese in Boxplots visualisiert. Die Maßzahlen wurden tabellarisch gelistet.

Für die **Berechnung der Trends** wurden die Messreihen in eine Rangordnung gebracht. Die Hauptkriterien für diese Rangordnung waren die Anzahl der Werte der Messreihe und der Zeitraum, in dem die Messwerte erhoben wurden. Die Trendberechnung erfolgte schrittweise: im ersten Schritt wurde der

Trend für alle durchgängig gemessenen Reihen berechnet und in den folgenden Schritten wurde jeweils die im Rang nächste Messreihe in die Trendberechnung mit aufgenommen. Die Trendberechnung erfolgte mit einem linearen Regressionsmodell, in dem der Parameter Anstieg und dessen 95% Konfidenzintervall geschätzt wurden. Der Anstieg entspricht der jährlichen Zunahme bzw. Abnahme der Messwerte.

Mit Hilfe des Konfidenzintervalls konnte getestet werden, ob der Anstieg ungleich Null ist, d.h. ob der Trend statistisch signifikant ist.

Das lineare Regressionsmodell wurde einerseits mit den Messwerten und andererseits mit den jährlichen Differenzen der Messwerte gerechnet. Die Ergebnisse der schrittweisen Trendberechnung wurden in Liniendiagrammen und zusammenfassend in Konfidenzintervalldarstellungen visualisiert.

Die Deskriptive zeigte das Spektrum der Heterogenität der Stationen und deren Messreihen. Insgesamt konnte für alle Schadstoffe festgestellt werden, dass die Daten des ländlichen Hintergrundes eine gute Basis für die Trendberechnung bildeten. Für Trendaussagen der Kategorie industrienah ist zu berücksichtigen, dass die Anzahl der Messreihen bei allen Schadstoffen äußerst gering war. In der Kategorie verkehrsnah gab es für NO₂ und PM₁₀ einen sehr starken Anstieg der Anzahl der Stationen ab dem Jahr 2005. Ferner streuten in allen Kategorien die Messreihen von PM₁₀ von 1990 bis 2003 untereinander sehr stark.

In der folgenden Tabelle sind die Charakteristiken der Luftschadstoffdaten zusammengestellt. Die Homogenität wurde auf einer Skala mit drei Stufen bewertet („+“ = gut, „+/-“ = neutral, „-“ = schlecht) bewertet. Die Anstiege beziehen sich auf den Zeitraum von 1990 – 2016 und wurden mit den Differenzen der Messwerte berechnet.

Kategorie	Eigenschaft	Ozon	NO ₂	PM ₁₀
	Flächendeckend in DE durchgängig messend	ab 1995	ab 2000	ab 2000
ländlicher Hintergrund	proz. Anteil der Stationen je Schadstoff	25,4 %	11,9 %	14,0 %
	Homogenität des Stationsbetriebes	+	+	+
	Homogenität der Messreihen	+/-	+/-	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,12	-0,26*	-0,46*
städtischer Hintergrund	proz. Anteil der Stationen je Schadstoff	55,4 %	36,1 %	40,5 %
	Homogenität des Stationsbetriebes	+/-	+/-	+/-
	Homogenität der Messreihen	+	+	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,23*	-0,52*	-0,86*
verkehrsnah	proz. Anteil der Stationen je Schadstoff	12,6 %	45,8 %	35,1 %
	Homogenität des Stationsbetriebes	-	-	-
	Homogenität der Messreihen	+/-	-	-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,36*	-0,67*	-1,22*
industrienah	proz. Anteil der Stationen je Schadstoff	6,6 %	6,2 %	10,4 %
	Homogenität des Stationsbetriebes	+/-	+/-	+/-
	Homogenität der Messreihen	+/-	+/-	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,10	-0,43*	-0,89*

* signifikant auf dem Niveau von 5%

In der deskriptiven Auswertung zeigte sich, dass die Heterogenität der Messreihen sowohl bezüglich der Anzahl der Messwerte als auch bezüglich des Messniveaus einen sehr starken Einfluss auf den Verlauf der Jahresmittelwertzeitreihen hat.

Das Modell der jährlichen Differenzen reduzierte jede Messreihe auf ihre jährlichen Veränderungen. Die jährlichen Veränderungen sind normierte Werte, die unabhängig vom Messniveau sind. Damit konnten für die Bildung der Jahresmittelwertzeitreihen neben den durchgängig messenden Reihen auch die Messreihen berücksichtigt werden, die nicht für jedes Jahr einen Messwert besitzen. Die Jahresmittelwertzeitreihe für das Modell wurden wie folgt berechnet:

1. je Messreihe wurden die jährlichen Differenzen gebildet
2. die jährlichen Differenzen wurden je Jahr gemittelt
3. die so gebildeten Mittelwerte wurden für jedes Jahr sukzessive über die vorherigen Jahre summiert
4. zu den summierten Mittelwerten je Jahr wurde der gewichtete Mittelwert des letzten Jahres addiert

In der weiteren induktiven Auswertung der drei Luftschadstoffe wurden die Jahresmittelwertzeitreihen je Kategorie für

- ▶ die Rohdaten, d.h. für alle Messreihen,
- ▶ die UBA-Auswahl von Messreihen und
- ▶ das Modell der jährlichen Differenzen

berechnet. Weiterhin wurden

- ▶ die Trends der Zeitreihen bestimmt,
- ▶ die Trends auf Signifikanz getestet (einerseits mit Hilfe der linearen Regression und andererseits mit dem Mann-Kendall Test) und
- ▶ die Zeitreihen untereinander deskriptiv verglichen.

Die durch das Modell der Differenzen berechnete standardisierte Mittelwertzeitreihe ist für die Schätzung der Trends der Luftschadstoffe Ozon, NO₂ und PM₁₀ sehr gut geeignet. Das Modell bezieht alle Messreihen eines Zeitintervalls, die mindestens zwei aufeinanderfolgende Messwerte besitzen, mit ein. Weiterhin wird durch die Differenzbildung aufeinanderfolgender Messwerte das Messniveau der Messreihen eliminiert. Somit werden die vorhandenen Inhomogenitäten der Daten berücksichtigt.

Durch Addition einer Konstanten wird die standardisierte Mittelwertzeitreihe auf das vorhandene Messniveau der Daten transformiert. Als Konstante ist der gewichtete Mittelwert des letzten Messzeitpunktes gut geeignet. Der Vergleich der so berechneten Mittelwertzeitreihen mit den Mittelwertzeitreihen der Rohdaten und der UBA-Auswahl zeigt für alle Schadstoffe in den Kategorien ländlicher Hintergrund, städtischer Hintergrund und industrienah im Mittel geringe Abweichungen. Die Abweichungen zwischen Modell und Rohdaten sind in der Kategorie verkehrsnah etwas stärker ausgeprägt, hier insbesondere am Anfang des Zeitintervalls. Die Ursache dafür liegt an der Tatsache, dass zunehmend die Messungen am Ende des Zeitintervalls auf einem höheren Niveau stattfanden. Somit ergibt sich durch den gewichteten Mittelwert am Ende des Zeitintervalls, dass das Modell am Anfang des Zeitintervalls höhere Werte (bei Ozon niedrigere Werte) schätzt. Allerdings ist die Abweichung zwischen Modell und UBA-Auswahl deutlich kleiner.

Alle Mittelwertzeitreihen wurden mit Hilfe der Parameterschätzung der linearen Regression und mit dem Mann-Kendall Test auf Signifikanz des Trends getestet. Sowohl bezüglich einer Signifikanz von 0,05 als auch von 0,01 zeigten beide Tests die gleichen Testergebnisse.

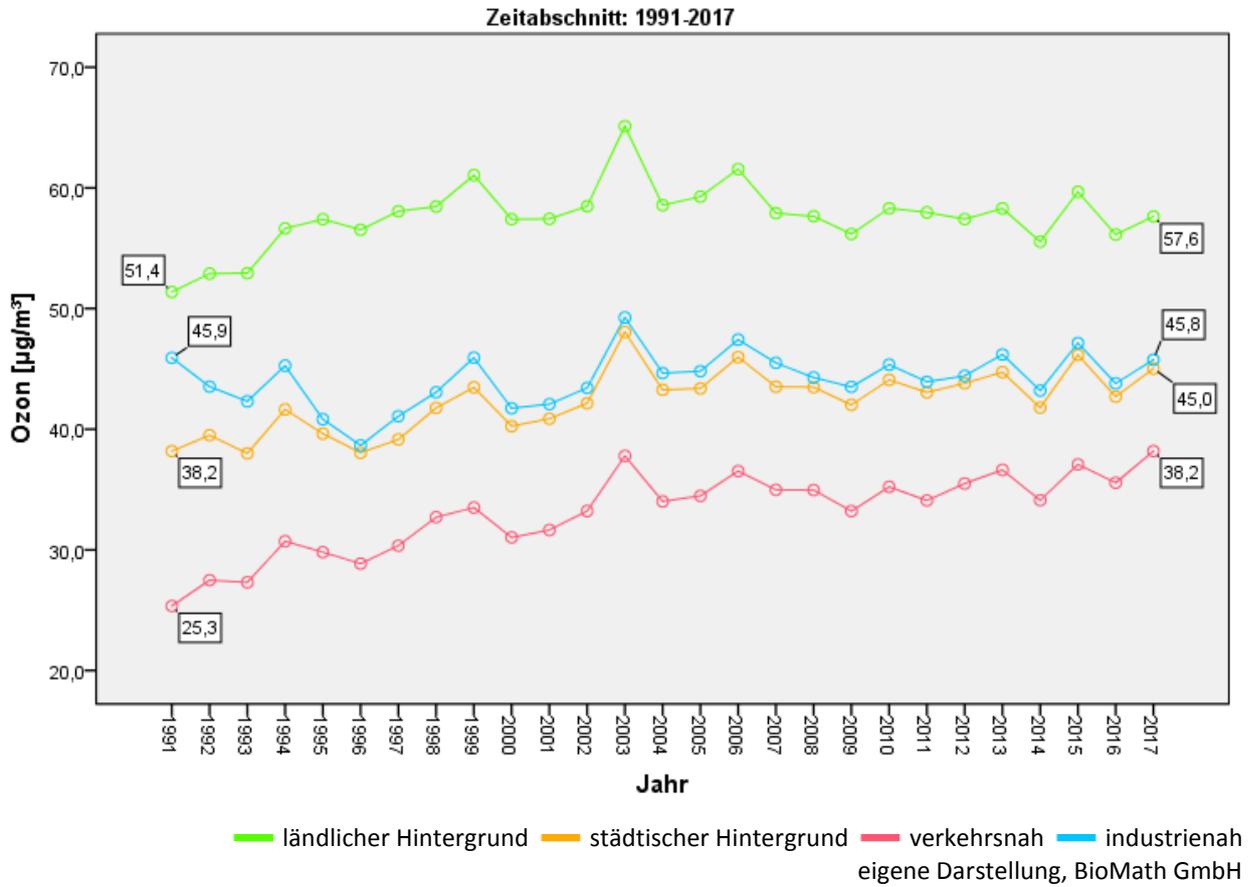
Die Parameterschätzung der linearen Regression ist sowohl für die Schätzung des Trends als auch für den Test auf Signifikanz des Trends geeignet. Die Voraussetzungen für die Anwendung der Regressionsanalyse, insbesondere der Normalverteilung des Fehlers (und nicht der Daten), sind weitestgehend erfüllt. Es ist nicht notwendig den Mann-Kendall Test, der nur Ränge verwendet, heranzuziehen.

Insgesamt kann festgestellt werden, dass die Mittelwertzeitreihen des Modells und der UBA-Auswahl zwar recht dicht beieinander liegen, aber unterschiedliche Stärken des Trends aufweisen.

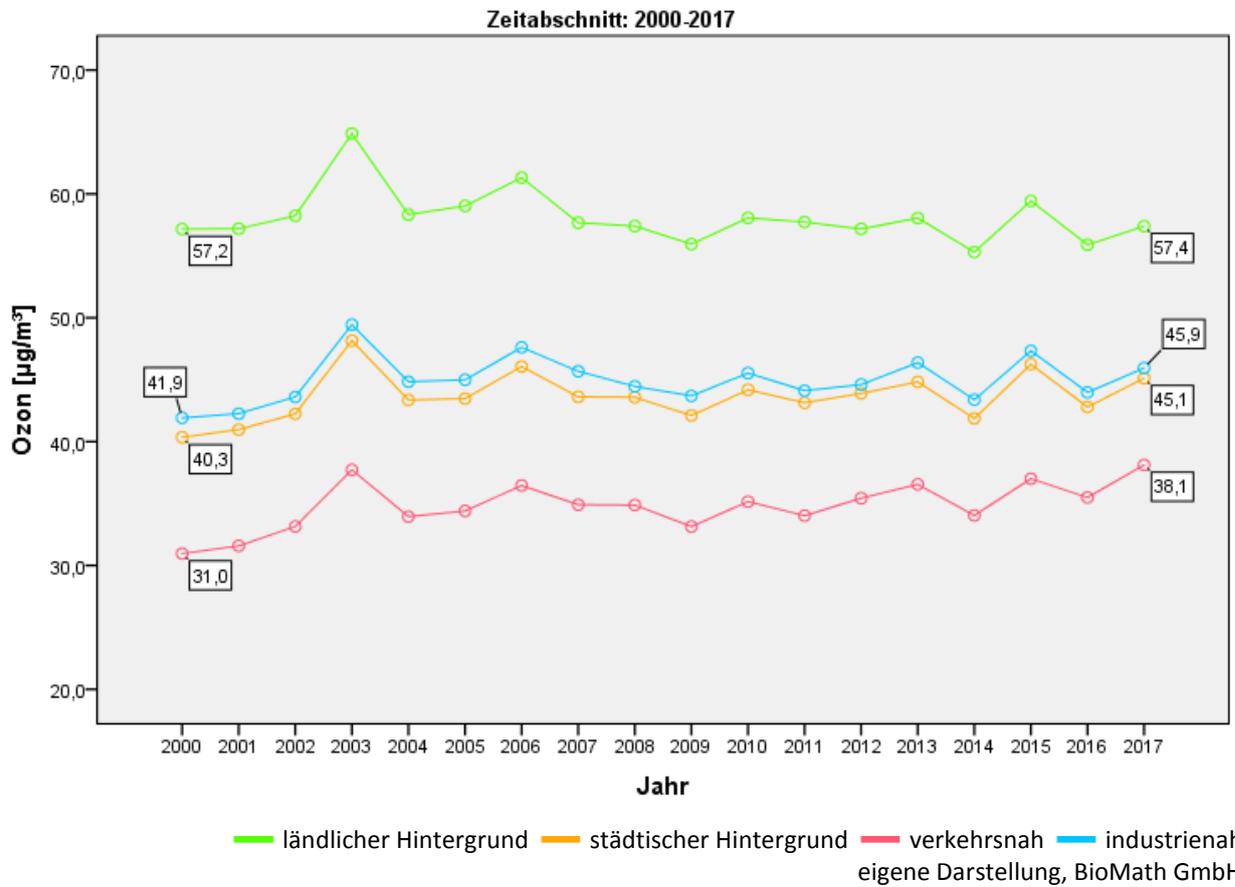
Die Berechnungen des Modells wurden in einem R-Programm umgesetzt und können auch für andere Luftschadstoffe, die sich ähnlich wie Ozon, NO₂ und PM₁₀ verhalten, verwendet werden.

Für die Zeitschnitte von 1991 bis 2017, 2000 bis 2017 und 2008 bis 2017 wurden folgende Mittelwertzeitreihen berechnet.

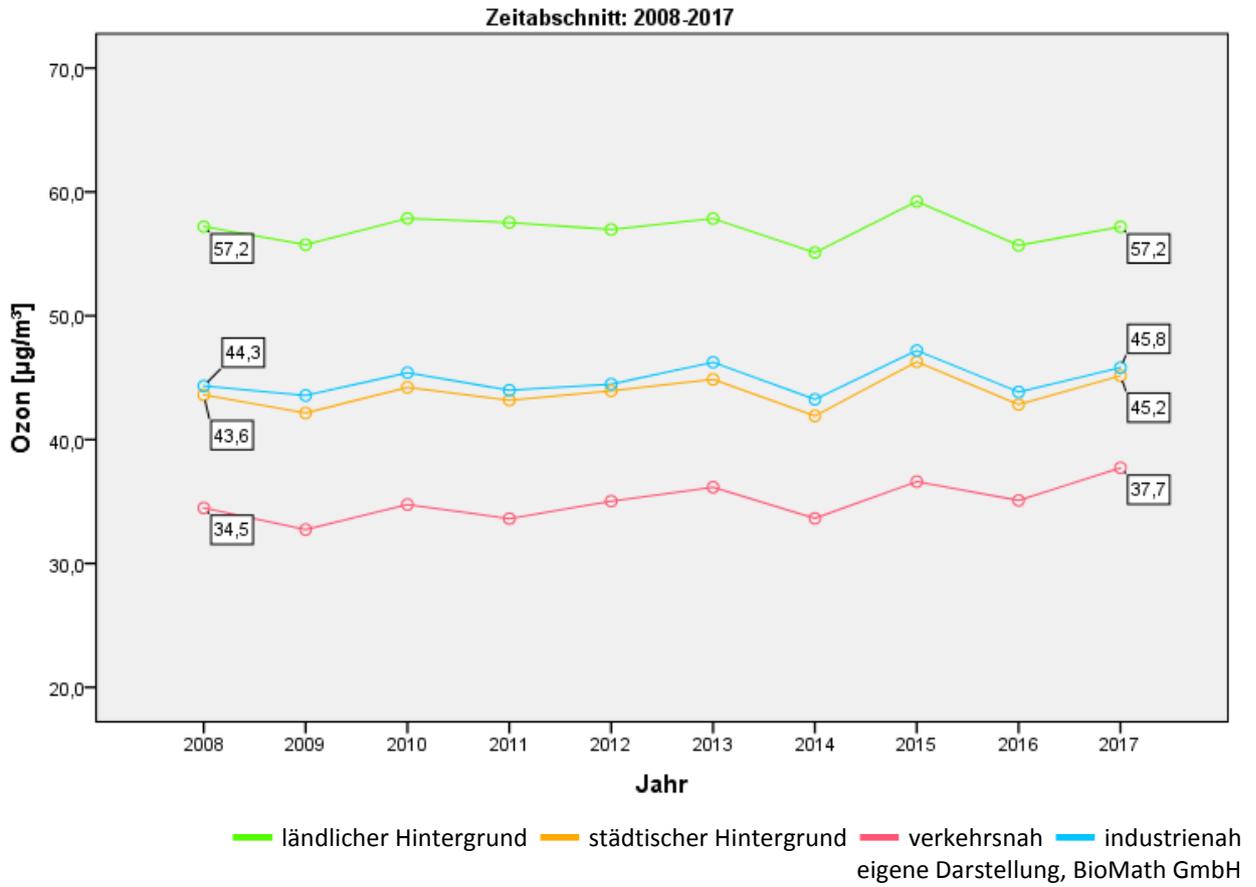
Ozon



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	0,1070	-0,0255	0,2395	Nein
städtischer Hintergrund	0,2247	0,1280	0,3213	Ja
verkehrsnahe	0,3593	0,2692	0,4494	Ja
industrienah	0,1062	-0,0022	0,2145	Nein

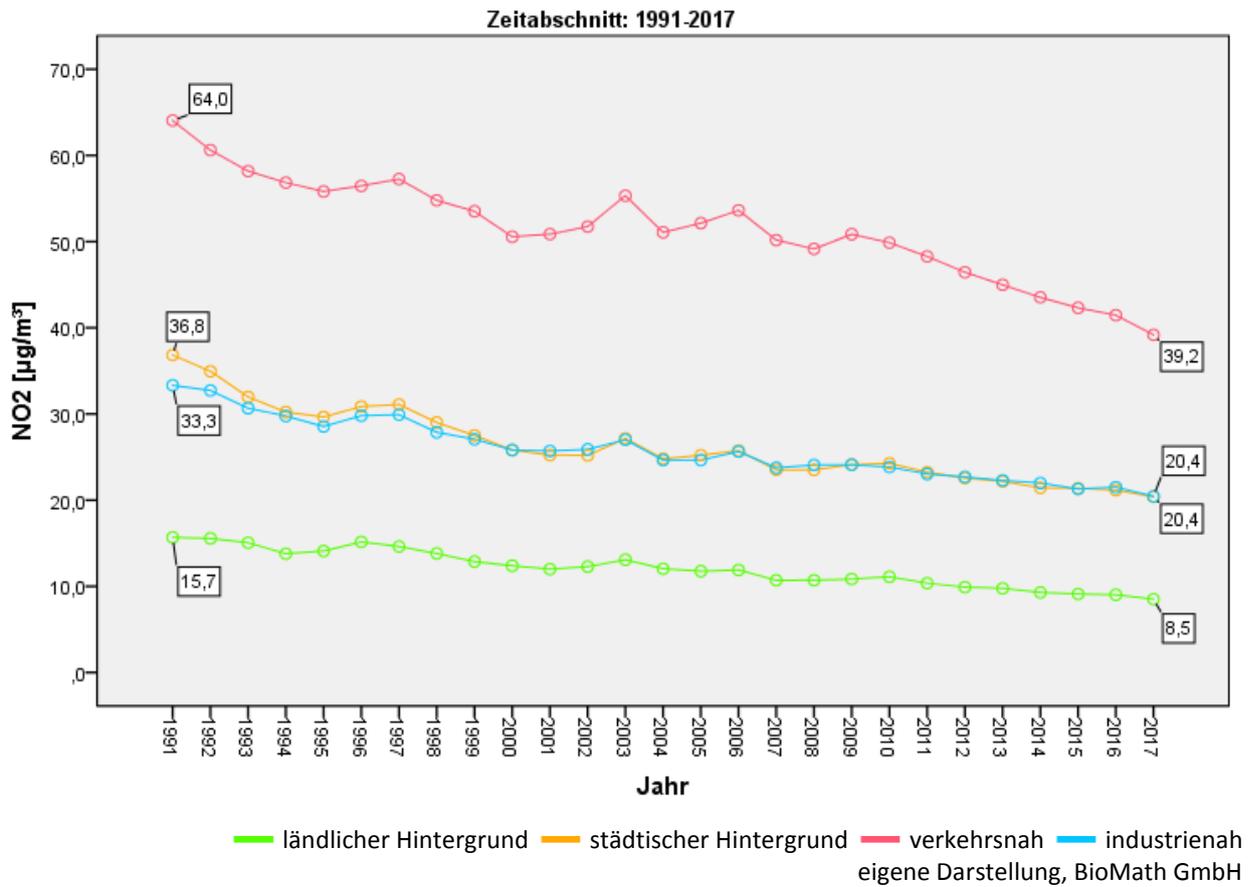


Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,1434	-0,3452	0,0583	Nein
städtischer Hintergrund	0,0964	-0,0886	0,2813	Nein
verkehrsnah	0,2143	0,0588	0,3697	Ja
industrienah	0,0707	-0,1136	0,2550	Nein

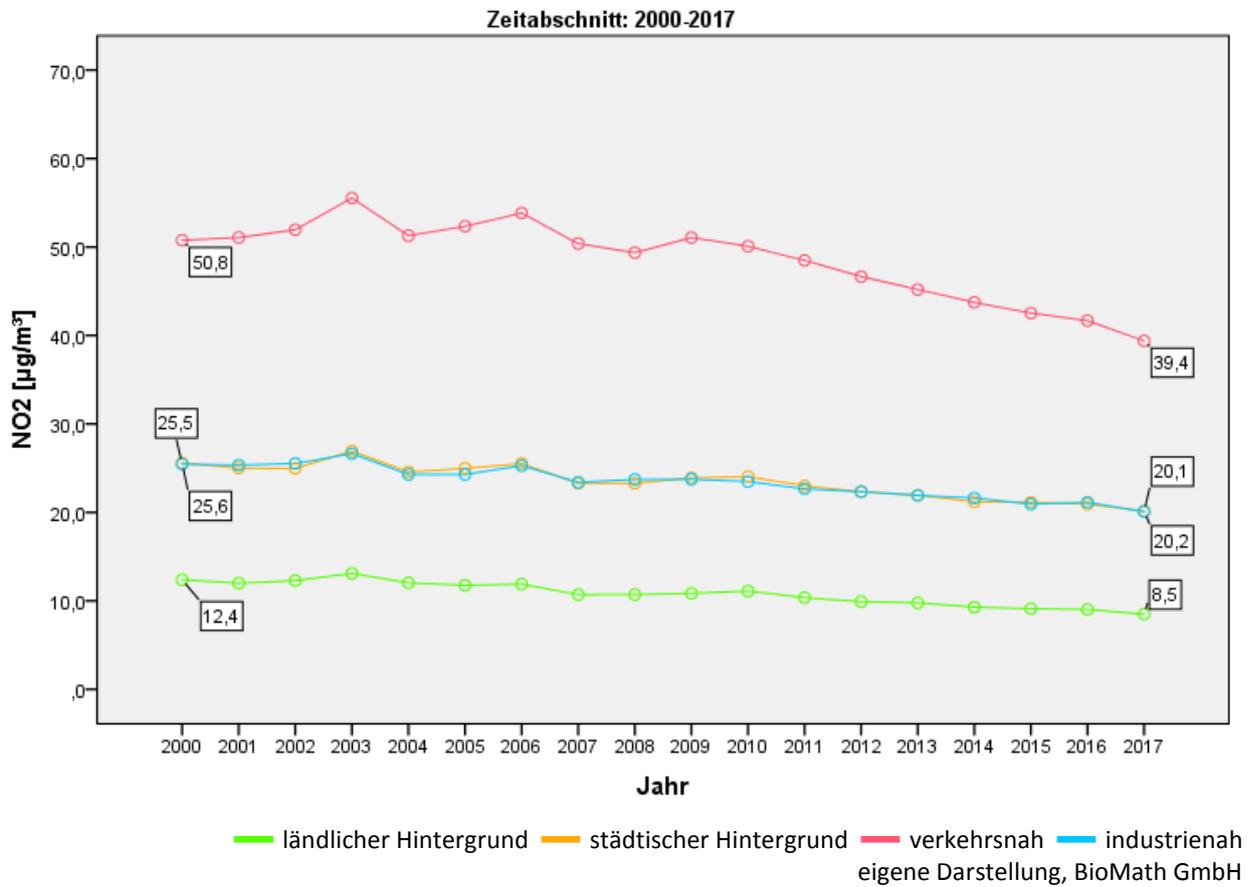


Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,0002	-0,3324	0,3320	Nein
städtischer Hintergrund	0,1583	-0,1882	0,5048	Nein
verkehrsnah	0,3397	0,0421	0,6373	Ja
industrienah	0,1449	-0,1827	0,4724	Nein

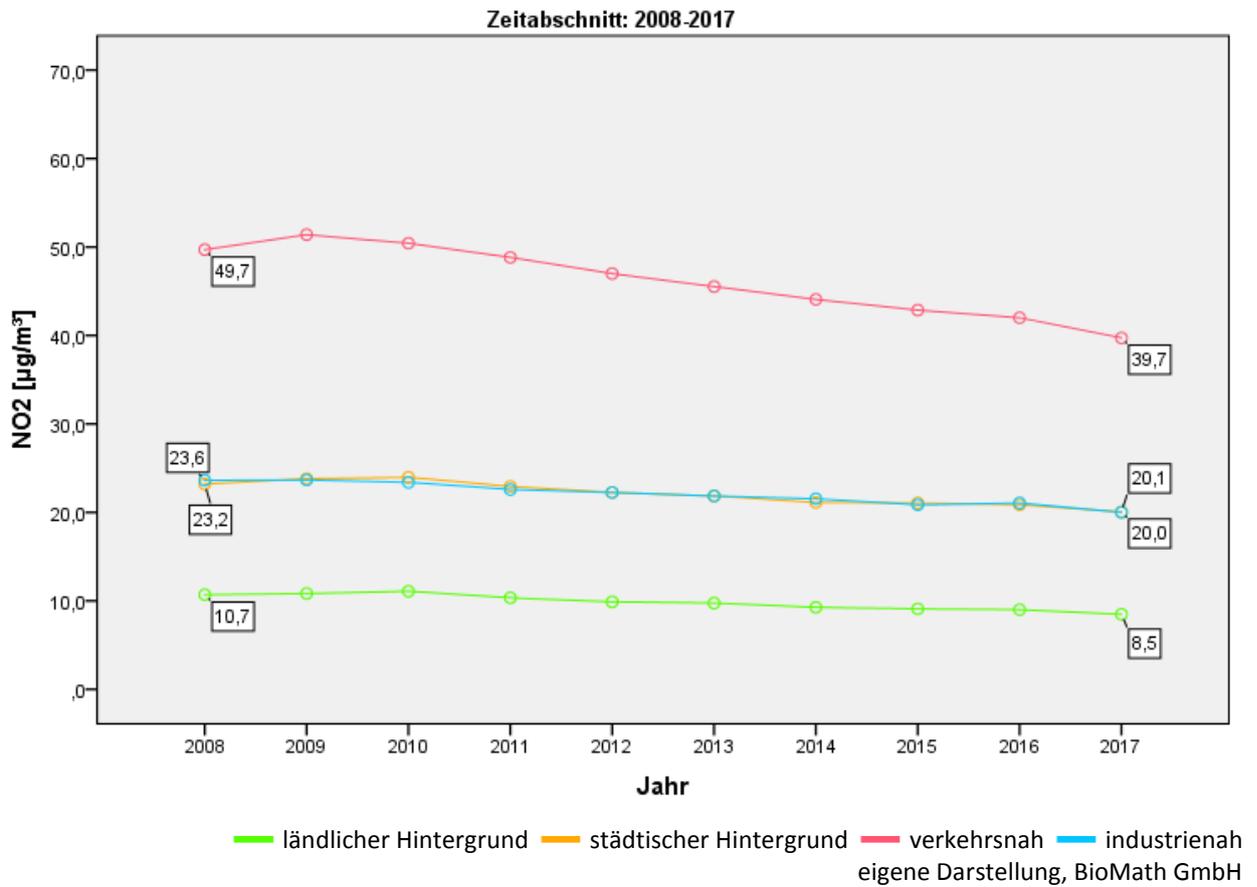
NO2



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2636	-0,2883	-0,2390	Ja
städtischer Hintergrund	-0,5086	-0,5839	-0,4333	Ja
verkehrsnahe	-0,7029	-0,8077	-0,5980	Ja
industrienah	-0,4267	-0,4740	-0,3793	Ja

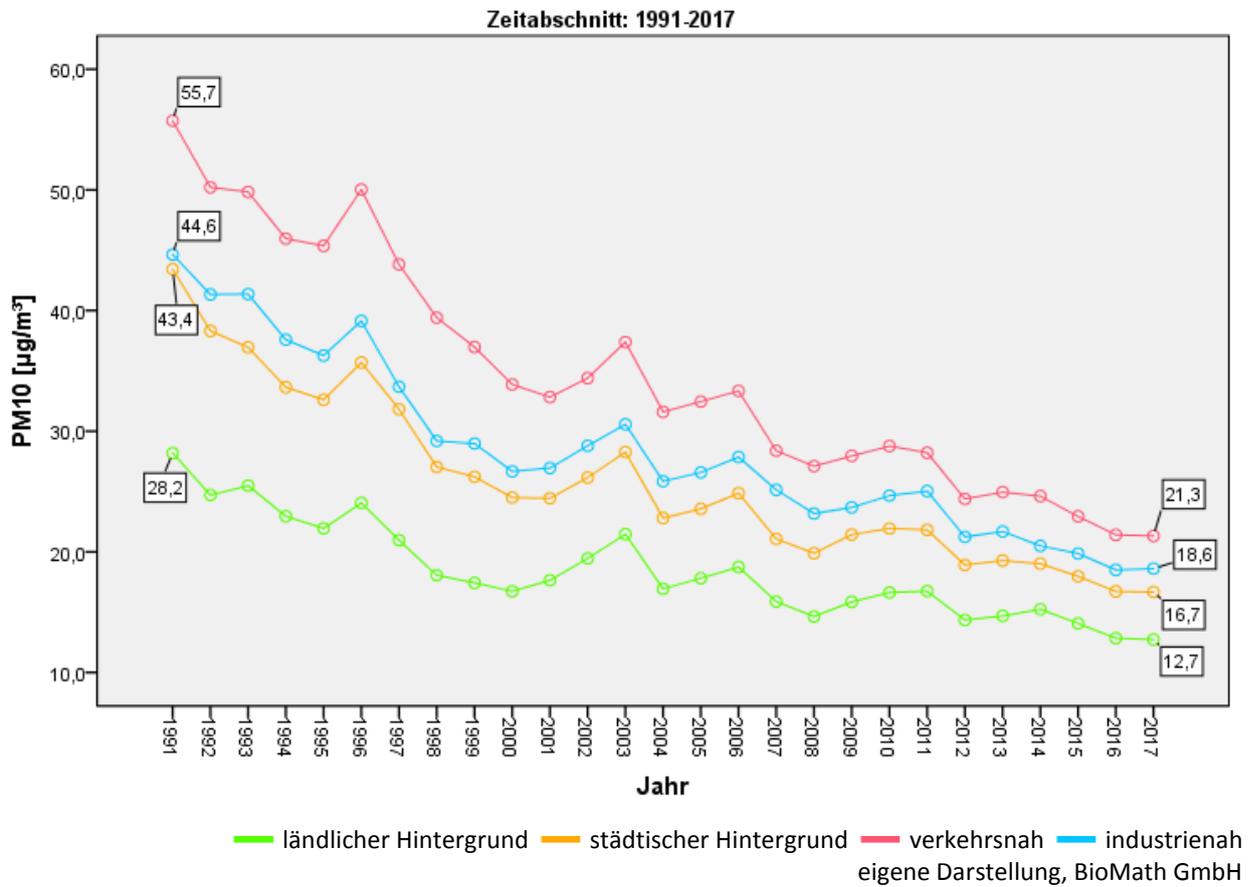


Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2408	-0,2802	-0,2014	Ja
städtischer Hintergrund	-0,3280	-0,4001	-0,2560	Ja
verkehrsnah	-0,7301	-0,9474	-0,5128	Ja
industrienah	-0,3272	-0,3846	-0,2697	Ja

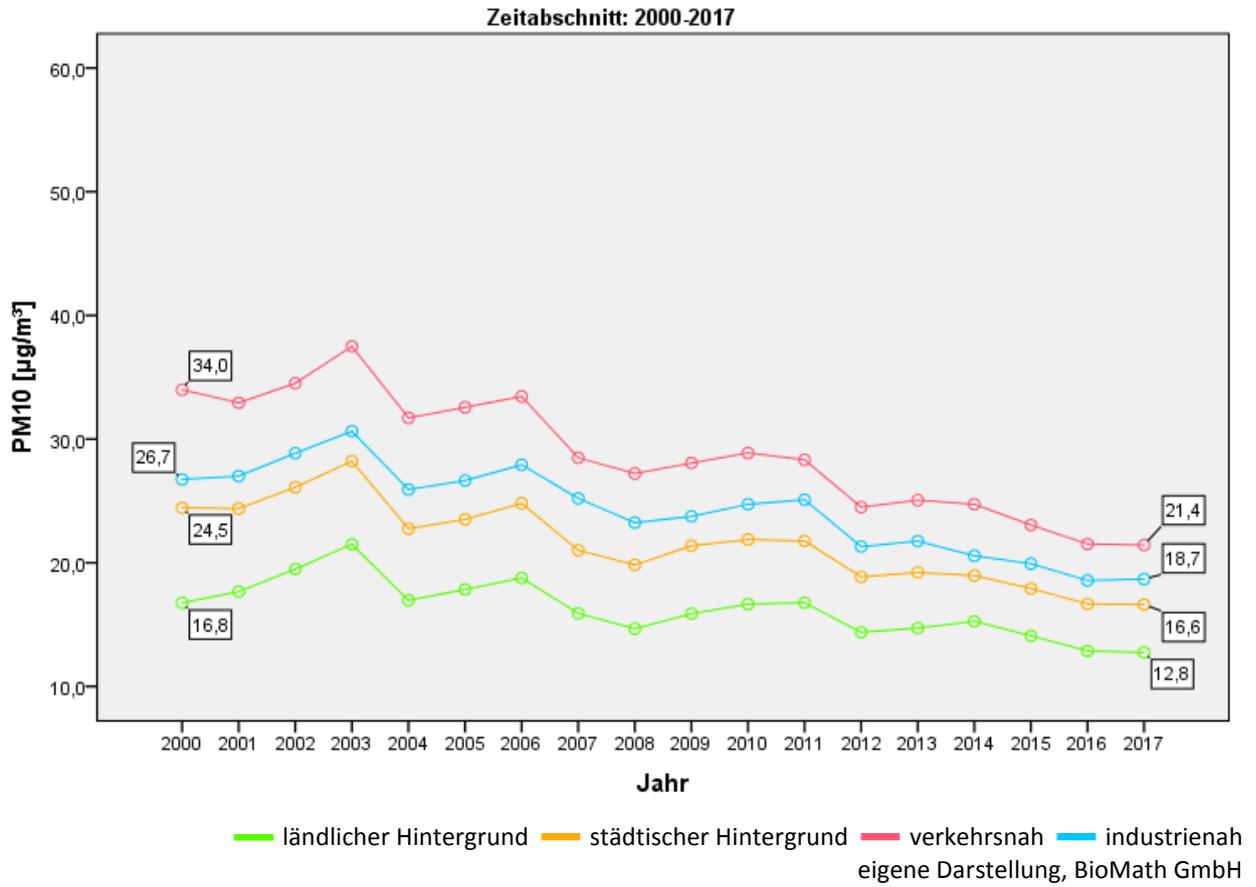


Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2787	-0,3443	-0,2131	Ja
städtischer Hintergrund	-0,4198	-0,5322	-0,3075	Ja
verkehrsnah	-1,2669	-1,5178	-1,0159	Ja
industrienah	-0,4077	-0,4696	-0,3457	Ja

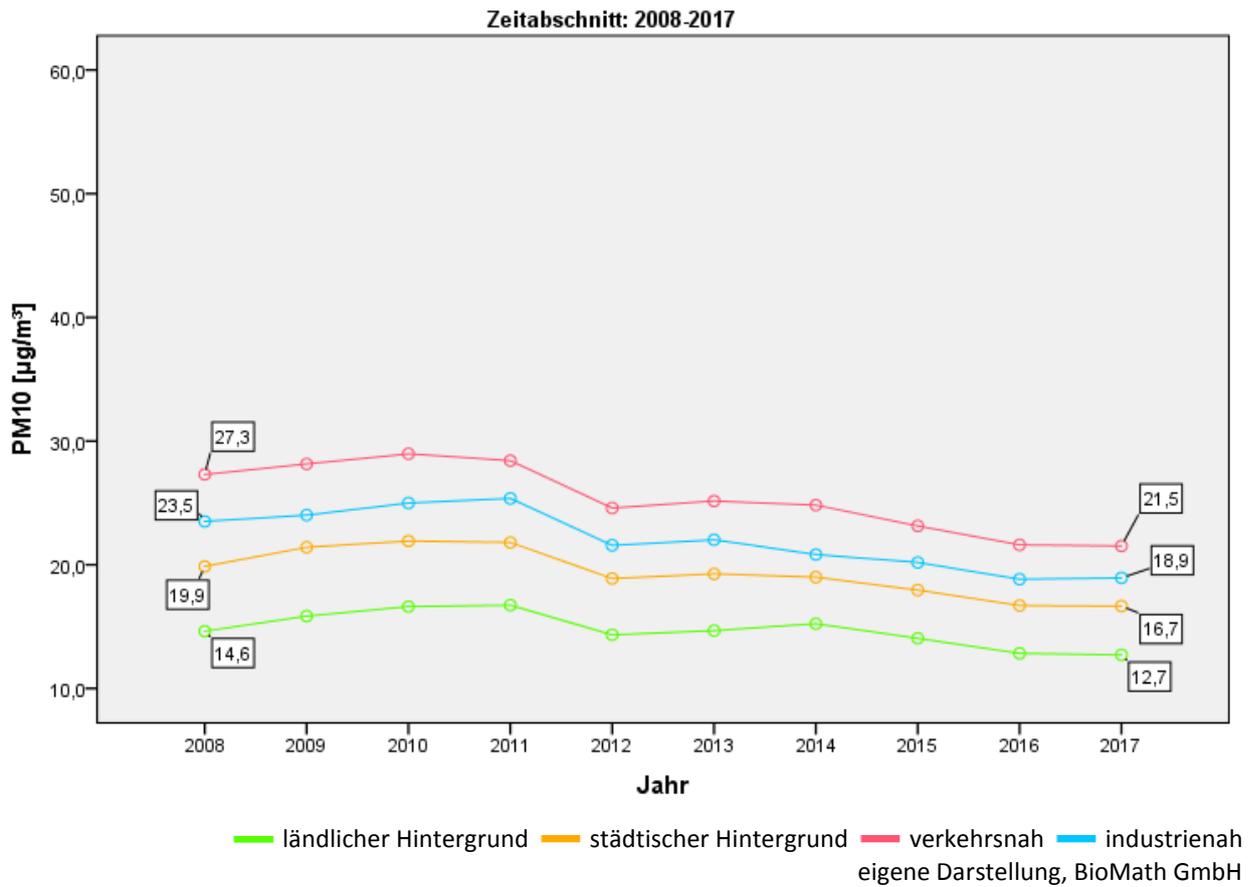
PM10



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,4601	-0,5529	-0,3673	Ja
städtischer Hintergrund	-0,8403	-0,9742	-0,7064	Ja
verkehrsnah	-1,1945	-1,3403	-1,0486	Ja
industrienah	-0,8751	-1,0061	-0,7440	Ja



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,3402	-0,4761	-0,2042	Ja
städtischer Hintergrund	-0,5496	-0,6928	-0,4065	Ja
verkehrsnah	-0,8400	-1,0058	-0,6742	Ja
industrienah	-0,5990	-0,7467	-0,4513	Ja



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,3351	-0,5901	-0,0800	Ja
städtischer Hintergrund	-0,5438	-0,8174	-0,2702	Ja
verkehrsnahe	-0,8317	-1,1367	-0,5267	Ja
industrienah	-0,6942	-0,9929	-0,3955	Ja

Summary

The Federal Environmental Agency (UBA) provided ozone, NO₂ and PM₁₀ data, comprising annual means in µg/m³ of the respective air pollutants, for the time period from 1990 to 2016, later also for 2017.

One data record referred to one monitoring station and contained annual means for the time period the station was operating. Geographical coordinates (coordinate system WGS84) were assigned to each station. Furthermore, each station was classified according to the station's surrounding and EoI (Exchange of Information) type of station. This classification was used to assign each station into one of the categories

- ▶ rural background
- ▶ urban background
- ▶ urban-near traffic
- ▶ industry.

Descriptive analysis of the three air pollutants data considered the following aspects:

- ▶ number and distribution of monitoring stations
- ▶ development of values within measurement series
- ▶ statements on linear trends in measurement series by simple models

The **number of monitoring stations** was analysed for all existing stations, and additionally for all continuously measuring stations. Continuously measuring stations were defined to be stations having a valid mean value for each year between start and end of operation. The analysis was performed for all categories (rural and urban background, traffic and industry) together, as well as for each single category.

To better characterise the stations and the measurement series, for each station the following variables were calculated:

- ▶ number of values measured at the station
- ▶ year of first measurement (start of operation)
- ▶ year of last measurement (end of operation)
- ▶ number of years between start and end of operation (operation period)
- ▶ number of years with no measurements
- ▶ evaluation of the measurement series

Based on these variables, bar and line charts as well as boxplots were created, and the variables were listed in tables. Furthermore, the stations were visualised on maps, and the number of stations per Federal State/ operator were plotted.

To describe the development of values within measurement series, for all categories the values of both continuously and not continuously measuring stations were presented in line charts. Statistical parameters per measurement series were calculated and plotted in boxplots. The parameters were listed in tables.

To calculate the trends, the measurement series were ranked. The mean criteria for ranking were the number of values in the measurement series and the time period of measuring. Trend calculation was performed in several steps: in a first step the trends were calculated for only continuous measurement series, and in the following steps the next ranked measurement series were included, respectively.

Trends were calculated applying a linear regression model estimating the slope and its 95% confidence interval. The slope corresponds to the annual increase/ decrease of the measurements.

The confidence interval corresponds to a test whether the slope is different from zero, i.e. whether the trend is statistically significant.

The linear regression model was applied both to the measurements and to the annual measurement differences. The results of stepwise trend calculation were displayed in line graphs and summarised in confidence interval diagrams.

The descriptive analysis shows the spectrum of heterogeneity of stations and their measurement series. Generally, the rural background data for all pollutants form a sound basis for calculating trends. For industrial areas, the number of measurement values for all pollutants was very low. In the category rural-near traffic, from the year 2005 there was a strong rise in the number of stations measuring NO₂ and PM₁₀. Moreover, from 1990 to 2003 in all categories the PM₁₀ measurement series are subject to considerable variance.

The following Table summarises the characteristics of the air pollutant data. Homogeneity was assessed on a three-level- scale („+“ = good, „+/-“ = neutral „-“ = poor). The slopes correspond to the time period 1990 to 2016 and were calculated on the basis of measurement differences.

category	character	ozone	NO ₂	PM ₁₀
	nationwide continuous measurement in DE	from 1995	from 2000	from 2000
rural background	percentage of stations per pollutant	25,4 %	11,9 %	14,0 %
	homogeneity of station operation	+	+	+
	homogeneity of measurement series	+/-	+/-	+/-
	slope in µg/m ³ per year (measurement diff.)	+0,12	-0,26*	-0,46*
urban background	percentage of stations per pollutant	55,4 %	36,1 %	40,5 %
	homogeneity of station operation	+/-	+/-	+/-
	homogeneity of measurement series	+	+	+/-
	slope in µg/m ³ per year (measurement diff.)	+0,23*	-0,52*	-0,86*
urban-near traffic	percentage of stations per pollutant	12,6 %	45,8 %	35,1 %
	homogeneity of station operation	-	-	-
	homogeneity of measurement series	+/-	-	-
	slope in µg/m ³ per year (measurement diff.)	+0,36*	-0,67*	-1,22*
industry	percentage of stations per pollutant	6,6 %	6,2 %	10,4 %
	homogeneity of station operation	+/-	+/-	+/-
	homogeneity of measurement series	+/-	+/-	+/-
	slope in µg/m ³ per year (measurement diff.)	+0,10	-0,43*	-0,89*

* significant at 5% level

The descriptive analysis demonstrated the very strong influence of the heterogeneity of the measurement series – both concerning the number of measurements and the measurement level – on the course of annual means’ time series.

The model of measurement differences reduced each measurement series to its annual changes. The annual changes are normed values and independent from the measurement level. Thereby for generating annual means’ time series not only continuous but also non-continuous measurement series could be included. The annual means’ time series for the model were calculated as follows:

1. annual differences per measurement series were calculated

2. the annual differences per year were averaged
3. for each year, these means were summarised successively over the previous years
4. for each year, the weighted mean of the last year was added to the summarised means

In the ensuing inductive analysis of the three air pollutants the annual means' time series per category were calculated for

- ▶ raw data, i.e. for all measurement series
- ▶ UBA selection of measurement series
- ▶ the model of annual measurement differences

Furthermore, the

- ▶ trends of time series were calculated,
- ▶ trends were tested for significance (by applying both linear regression and Mann-Kendall test)
- ▶ time series among each other were compared descriptively.

The mean time series standardised by the model of annual measurement differences are very well suited for estimating trends of the air pollutants ozone, NO₂ and PM₁₀. The model includes all measurement series of a time period, if there are at least two consecutive measurements. Moreover, by calculating differences of consecutive measurements the measurement level is eliminated. Thereby existing inhomogeneities of data are considered.

By adding a constant term, the standardised mean time series is transformed to the measurement level of the data. As a constant, the weighted mean of the last measurement timepoint is suitable. A comparison of mean time series calculated in such a way with mean time series of raw data and the UBA selection shows low deviations for all pollutants in the categories rural background urban background and industry. In the category rural-near traffic, the deviations between model and raw data are higher, especially at the beginning of the time period. This is caused by the fact that at the end of the time period measurements progressively were performed at a higher level. By working with weighted means at the end of the time period the model estimated higher values (for ozone: lower values) at the beginning of the time period. However, the deviation between model and UBA selection is considerably smaller.

All mean time series were tested for significance by analysing the parameter estimations of the linear regression and applying the Mann-Kendall test. Both methods showed same test results both at 0.05 and 0.01 significance level.

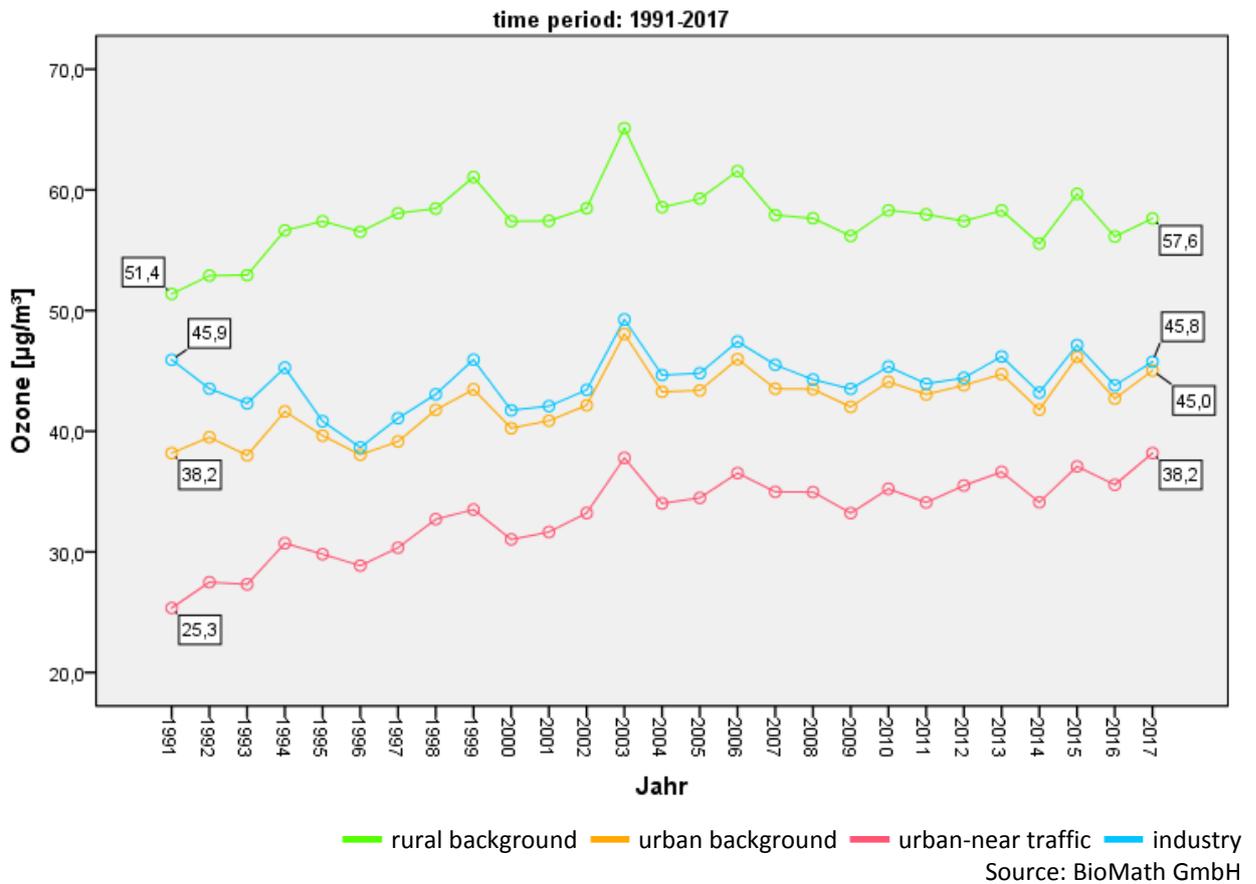
The parameter estimation of the linear regression is suited for estimations trends as well as testing the significance of trends. The requirements for applying regression analysis, mainly the normal distribution of the residuals (and not of the data) are fulfilled. There is no need to use the Mann-Kendall test which only considers ranks.

In summary, the mean time series of the model and the UBA selection are close to each other but show different strengths of the trends.

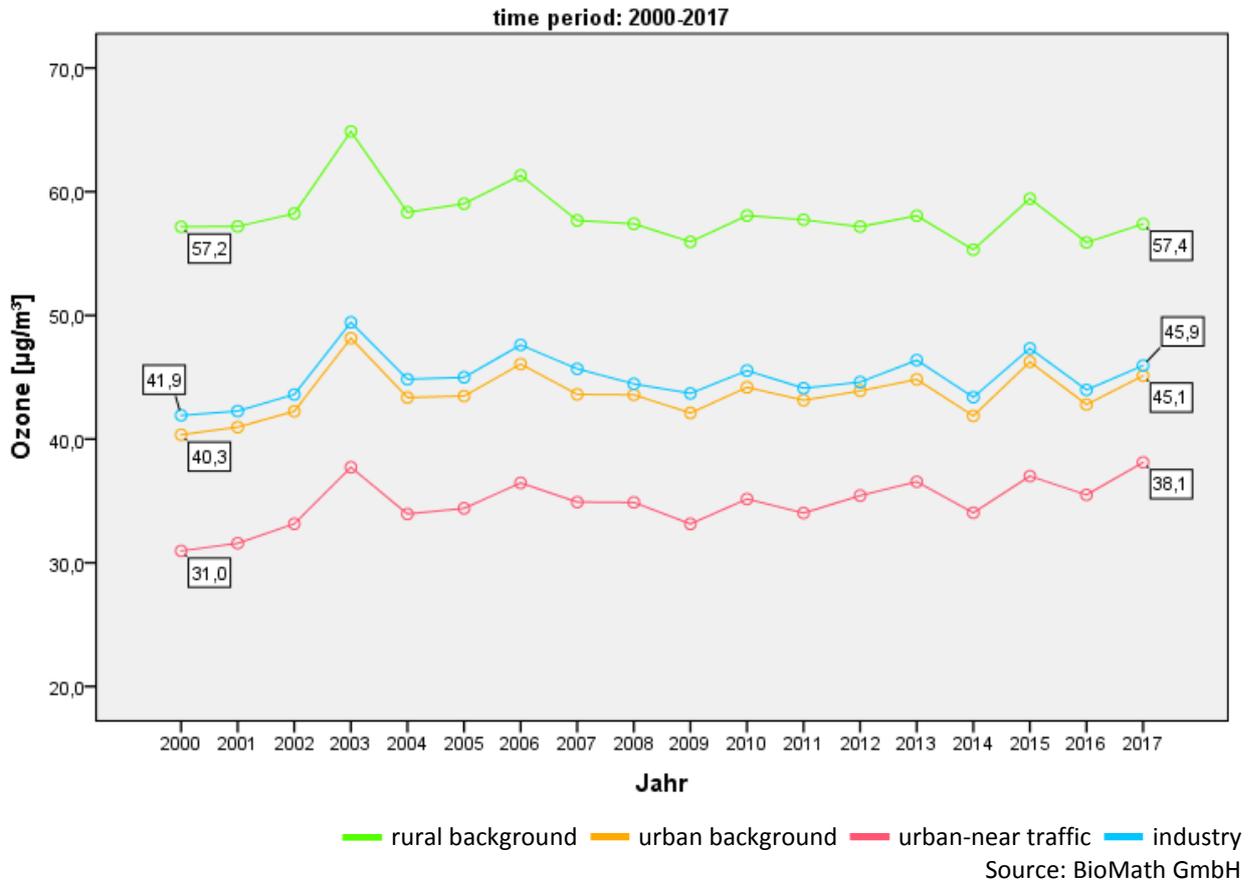
The model calculations were implemented in an R program and also can be applied to other air pollutants developing comparably to ozone, NO₂ and PM₁₀.

For the time periods 1991 to 2017, 2000 to 2017 and 2008 to 2017 the following mean time series were calculated.

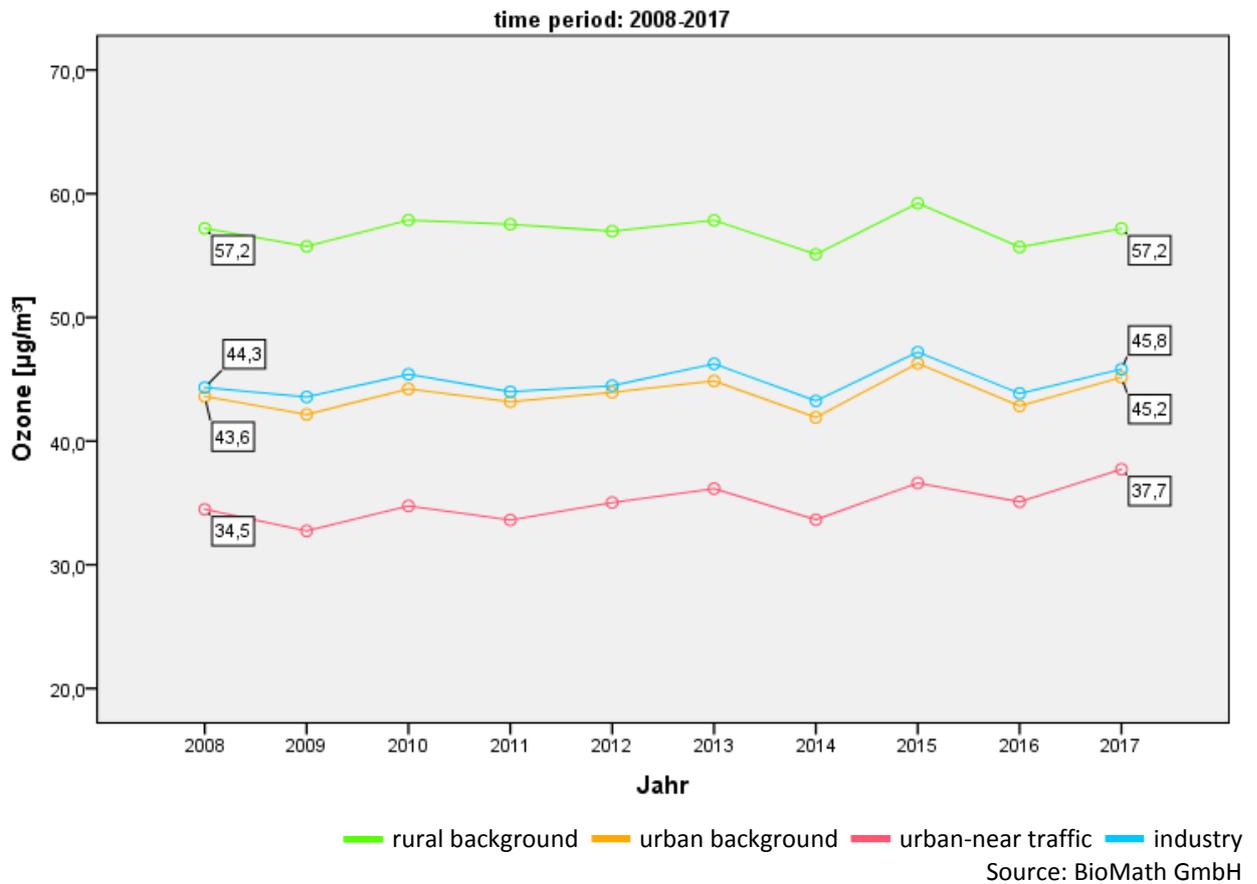
Ozone



category	slope	95% CI lower	95% CI upper	P<0,05
rural background	0,1070	-0,0255	0,2395	no
urban background	0,2247	0,1280	0,3213	yes
urban-near traffic	0,3593	0,2692	0,4494	yes
industry	0,1062	-0,0022	0,2145	no

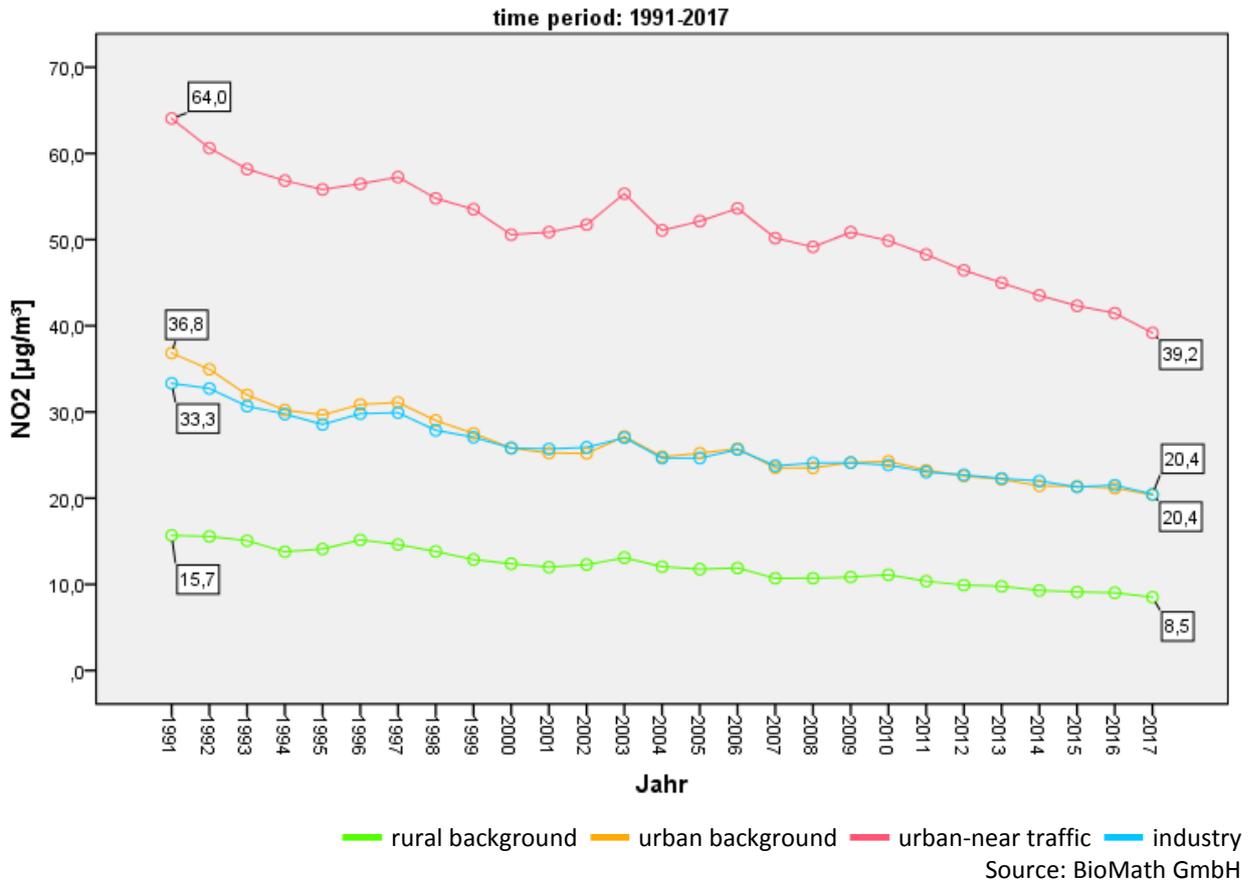


category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,1434	-0,3452	0,0583	no
urban background	0,0964	-0,0886	0,2813	no
urban-near traffic	0,2143	0,0588	0,3697	yes
industry	0,0707	-0,1136	0,2550	no

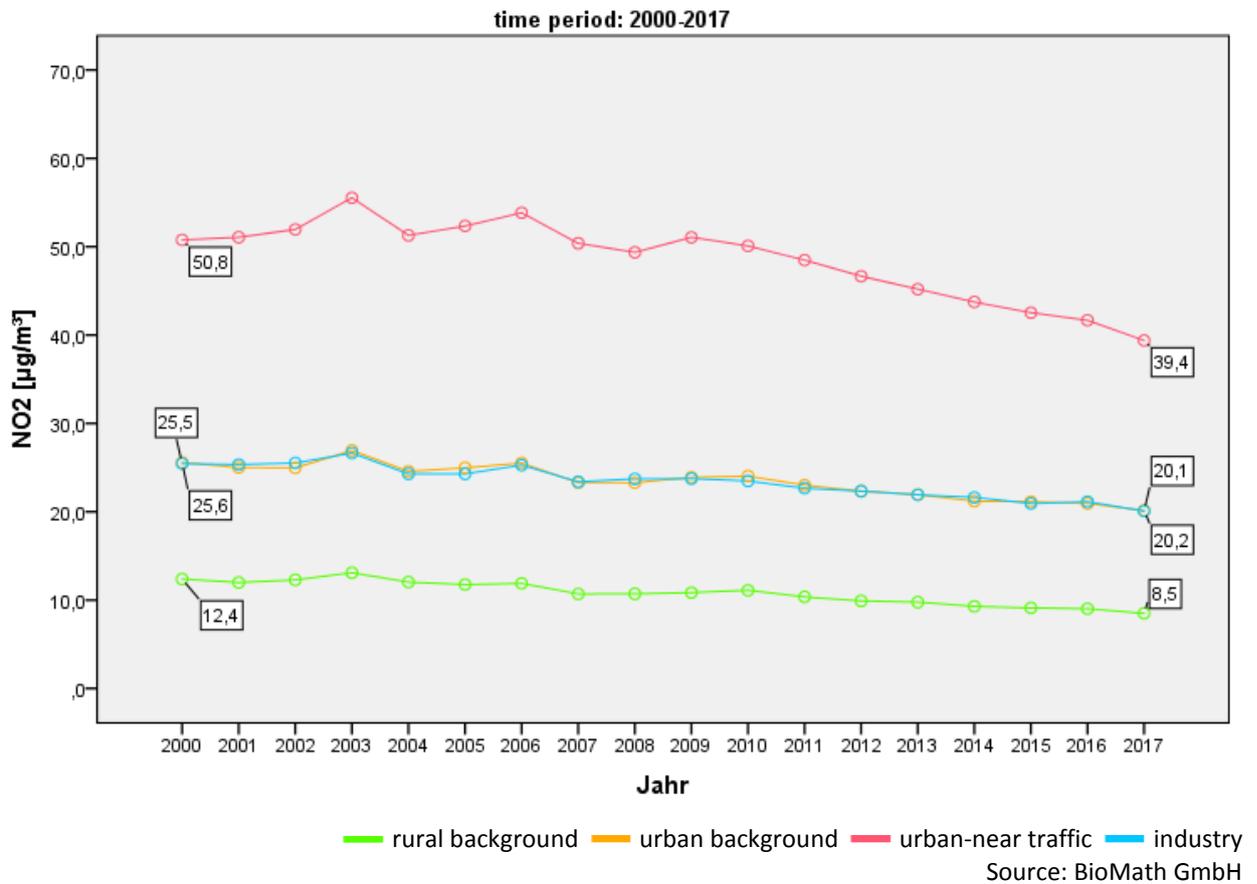


category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,0002	-0,3324	0,3320	no
urban background	0,1583	-0,1882	0,5048	no
urban-near traffic	0,3397	0,0421	0,6373	yes
industry	0,1449	-0,1827	0,4724	no

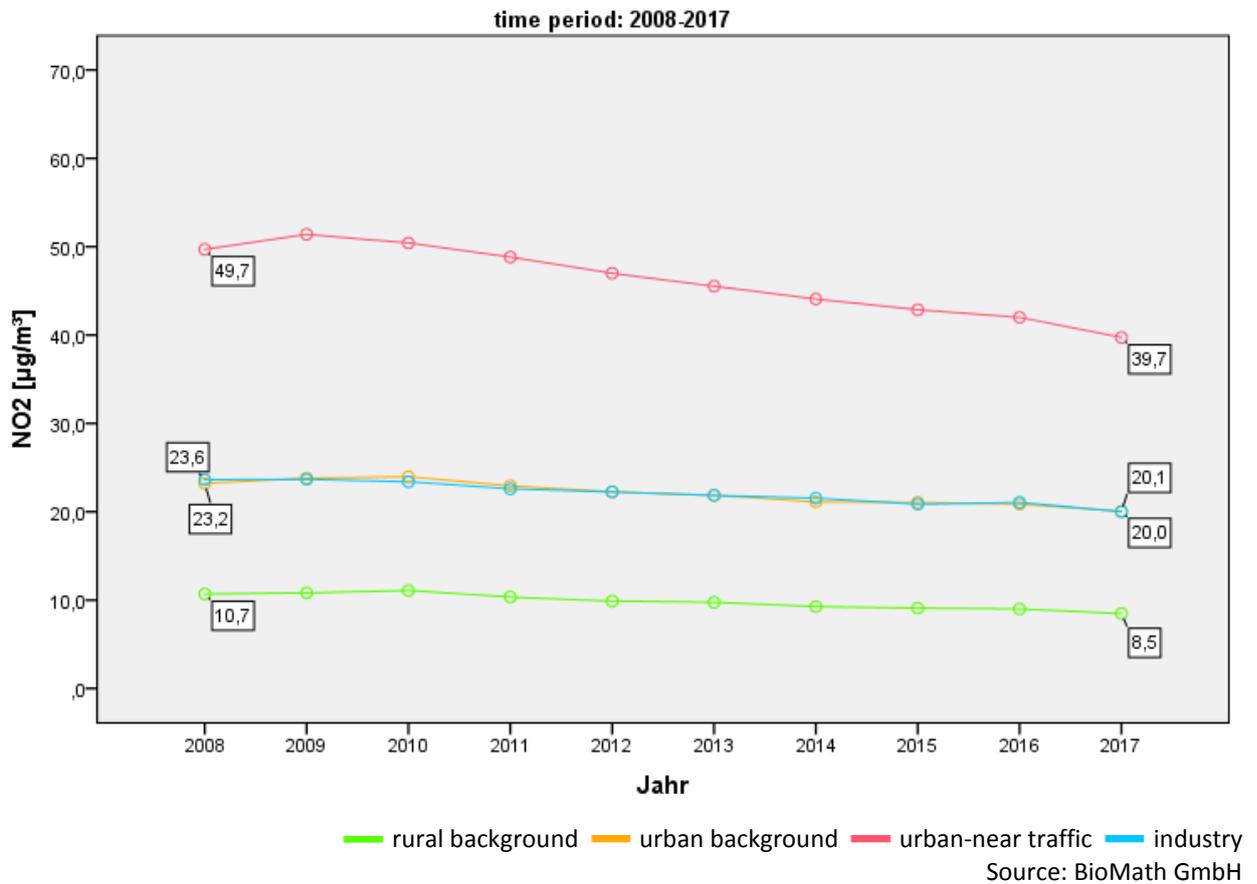
NO2



category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,2636	-0,2883	-0,2390	yes
urban background	-0,5086	-0,5839	-0,4333	yes
urban-near traffic	-0,7029	-0,8077	-0,5980	yes
industry	-0,4267	-0,4740	-0,3793	yes

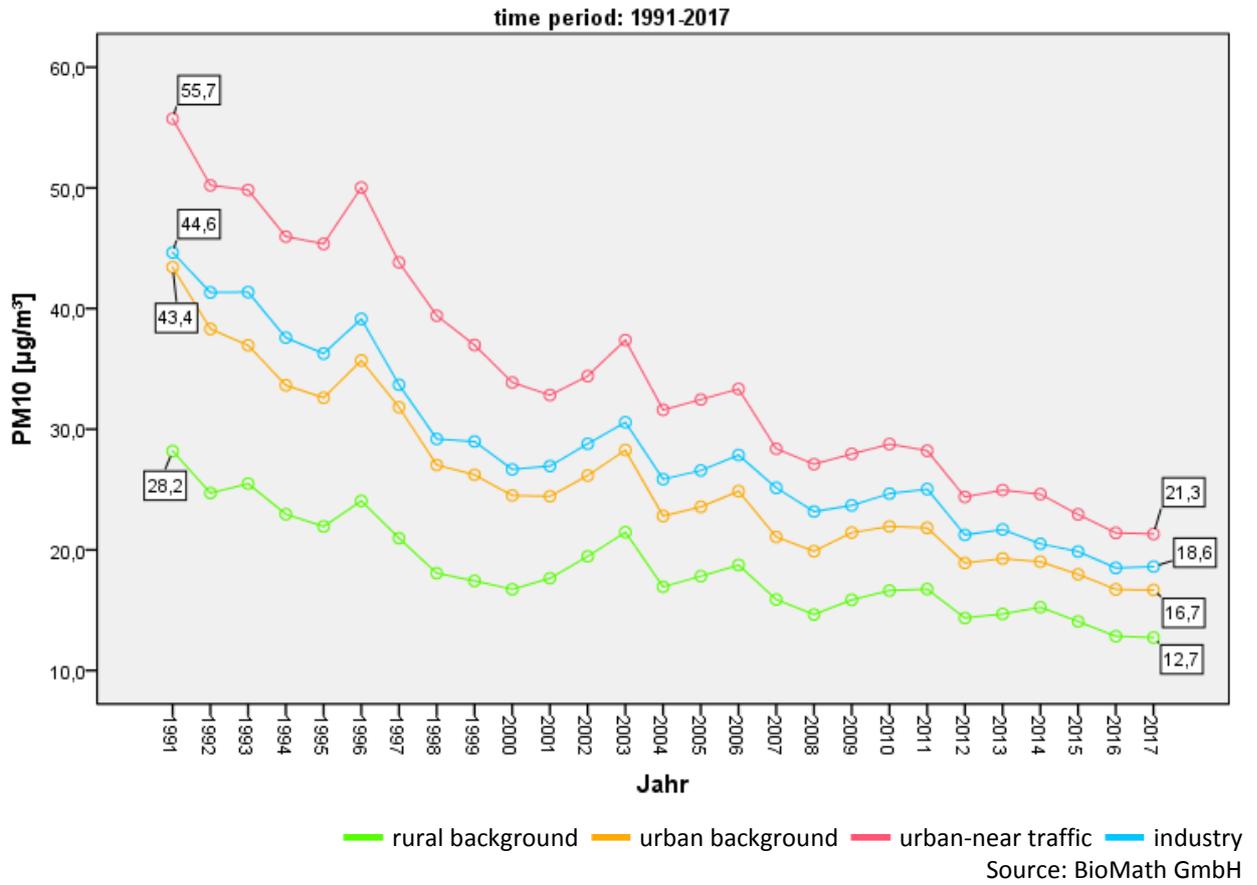


category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,2408	-0,2802	-0,2014	yes
urban background	-0,3280	-0,4001	-0,2560	yes
urban-near traffic	-0,7301	-0,9474	-0,5128	yes
industry	-0,3272	-0,3846	-0,2697	yes

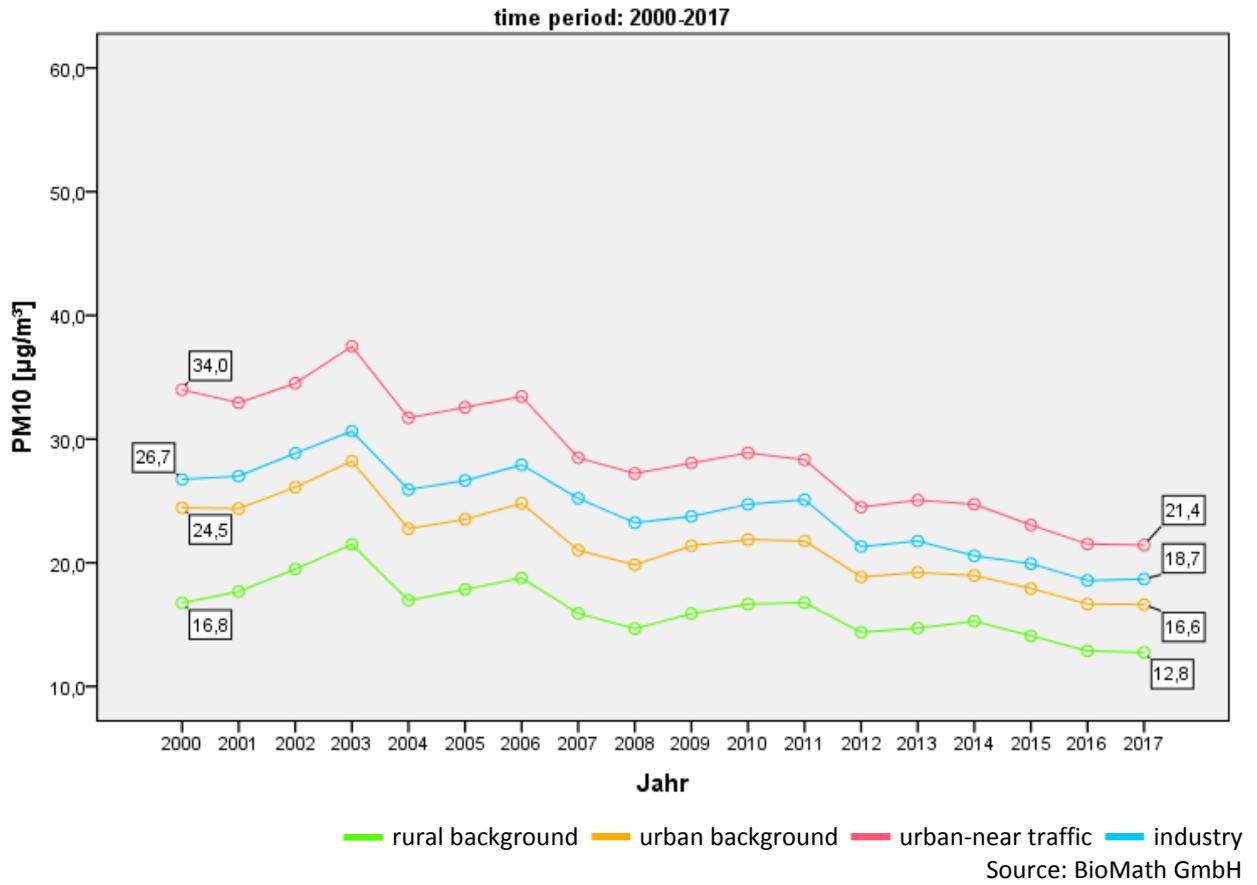


category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,2787	-0,3443	-0,2131	yes
urban background	-0,4198	-0,5322	-0,3075	yes
urban-near traffic	-1,2669	-1,5178	-1,0159	yes
industry	-0,4077	-0,4696	-0,3457	yes

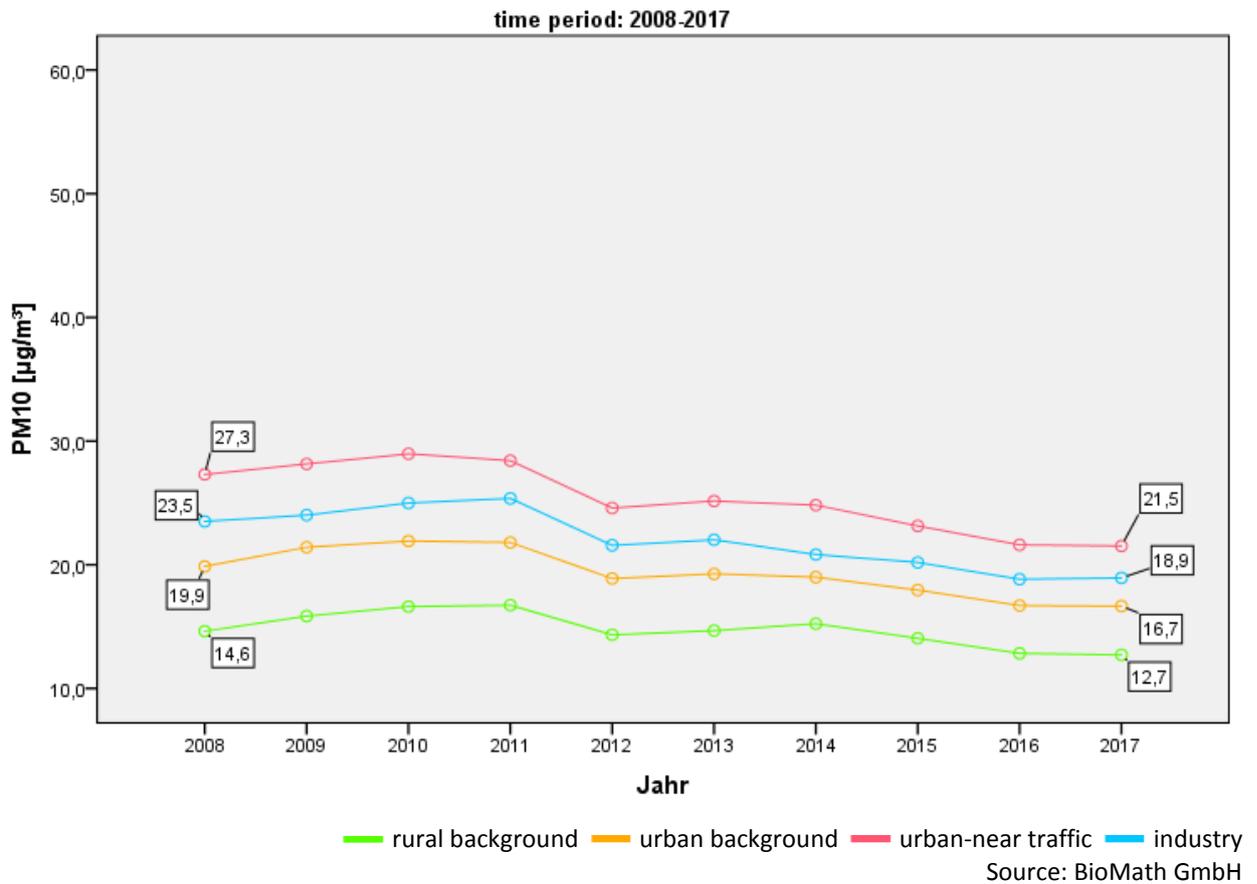
PM10



category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,4601	-0,5529	-0,3673	yes
urban background	-0,8403	-0,9742	-0,7064	yes
urban-near traffic	-1,1945	-1,3403	-1,0486	yes
industry	-0,8751	-1,0061	-0,7440	yes



category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,3402	-0,4761	-0,2042	yes
urban background	-0,5496	-0,6928	-0,4065	yes
urban-near traffic	-0,8400	-1,0058	-0,6742	yes
industry	-0,5990	-0,7467	-0,4513	yes



category	slope	95% CI lower	95% CI upper	P<0,05
rural background	-0,3351	-0,5901	-0,0800	yes
urban background	-0,5438	-0,8174	-0,2702	yes
urban-near traffic	-0,8317	-1,1367	-0,5267	yes
industry	-0,6942	-0,9929	-0,3955	yes

1 Einleitung

Zur Beurteilung der Luftqualität in Deutschland stehen dem Umweltbundsamt Daten aus räumlich und zeitlich inhomogenen Messnetzen der 16 Bundesländer und des Umweltbundsamtes zur Verfügung. Die Inhomogenität beruht einerseits auf der ungleichmäßigen Verteilung und andererseits auf der Schließung, Neueinrichtung oder zwischenzeitlichen Außerbetriebnahme von Messstationen. Damit ist es problematisch, sichere Trendaussagen zur kurz- und langfristigen Entwicklung der Luftqualität, die als repräsentativ für Deutschland angesehen werden können, zu treffen.

Im Rahmen dieses Projektes soll eine Methodik entwickelt werden, mit der sichere Aussagen zu kurz- und langfristigen mittleren Trends für die Luftschadstoffe Ozon, Stickstoffdioxid und Feinstaub getroffen und objektiv beurteilt werden können.

2 Material

2.1 Daten

Das UBA stellte am 13.11.2017 Daten von Ozon und am 21.11.2017 Daten von NO₂ und PM₁₀ zur Verfügung. Die Daten lagen im Excelformat vor und beinhalteten Jahresmittelwerte der jeweiligen Schadstoffe in µg/m³ über den Zeitraum von 1990 bis 2016.

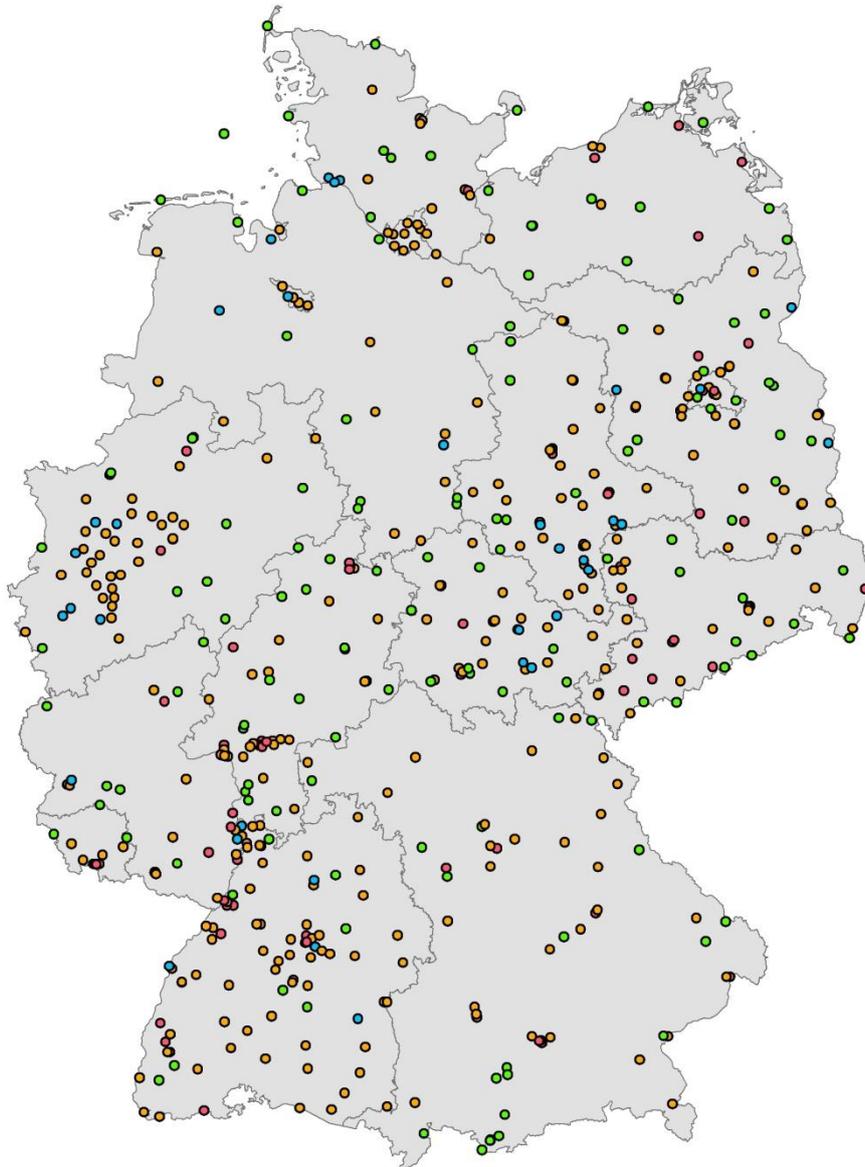
Ein Datensatz bezog sich auf eine Messstation und beinhaltete Jahresmittelwerte für den Zeitraum, in dem diese Station in Betrieb war. Jeder Station waren geografischen Koordinaten im Koordinatensystem WGS84 zugeordnet. Weiterhin war jede Station in Stationsumgebung und Art der Station nach EoI klassifiziert. Für die spätere Analyse wurde mit Hilfe dieser Klassifikation eine Kategorisierung jeder Station durchgeführt.

Tabelle 1: Kategorisierung

Kategorie	Farbe	Stationsumgebung nach EoI	Art der Station nach EoI
ländlicher Hintergrund		Hintergrund Hintergrund Hintergrund Hintergrund	ländlich stadtnah ländlich regional ländlich abgelegen ländliches Gebiet
städtischer Hintergrund		Hintergrund Hintergrund	städtisches Gebiet vorstädtisches Gebiet
verkehrsnahe		Verkehr	
industrienah		Industrie	

Die Messstationen verteilten sich, unabhängig vom Messzeitraum, bezüglich der vier Kategorien wie folgt:

Abbildung 1: Ozon - Karte mit allen Messstationen von 1990-2016

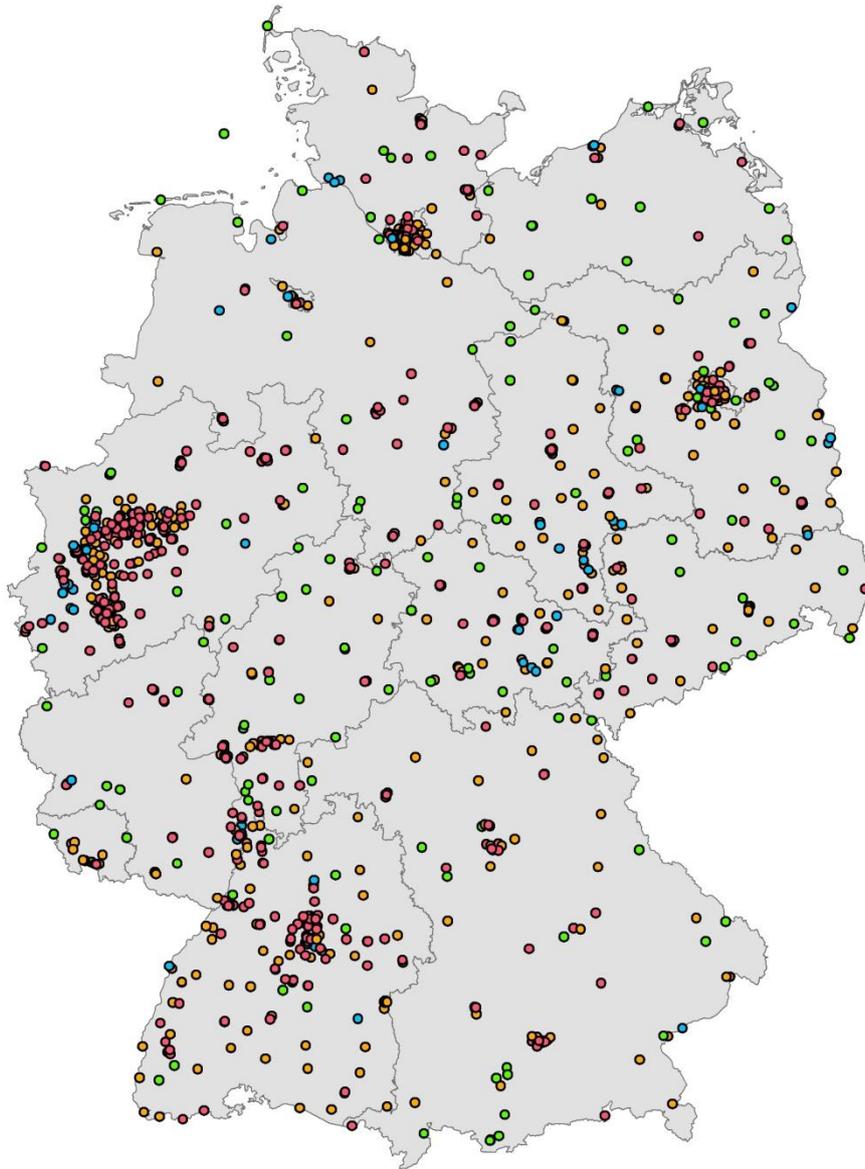


eigene Darstellung, BioMath GmbH

Tabelle 2: Ozon - Anzahl aller Messstationen

Kategorie	Anzahl aller Stationen
ländlicher Hintergrund	139
städtischer Hintergrund	303
verkehrsnahe	69
Industrienah	36
Gesamt	547

Abbildung 2: NO₂ - Karte mit allen Messstationen von 1990-2016

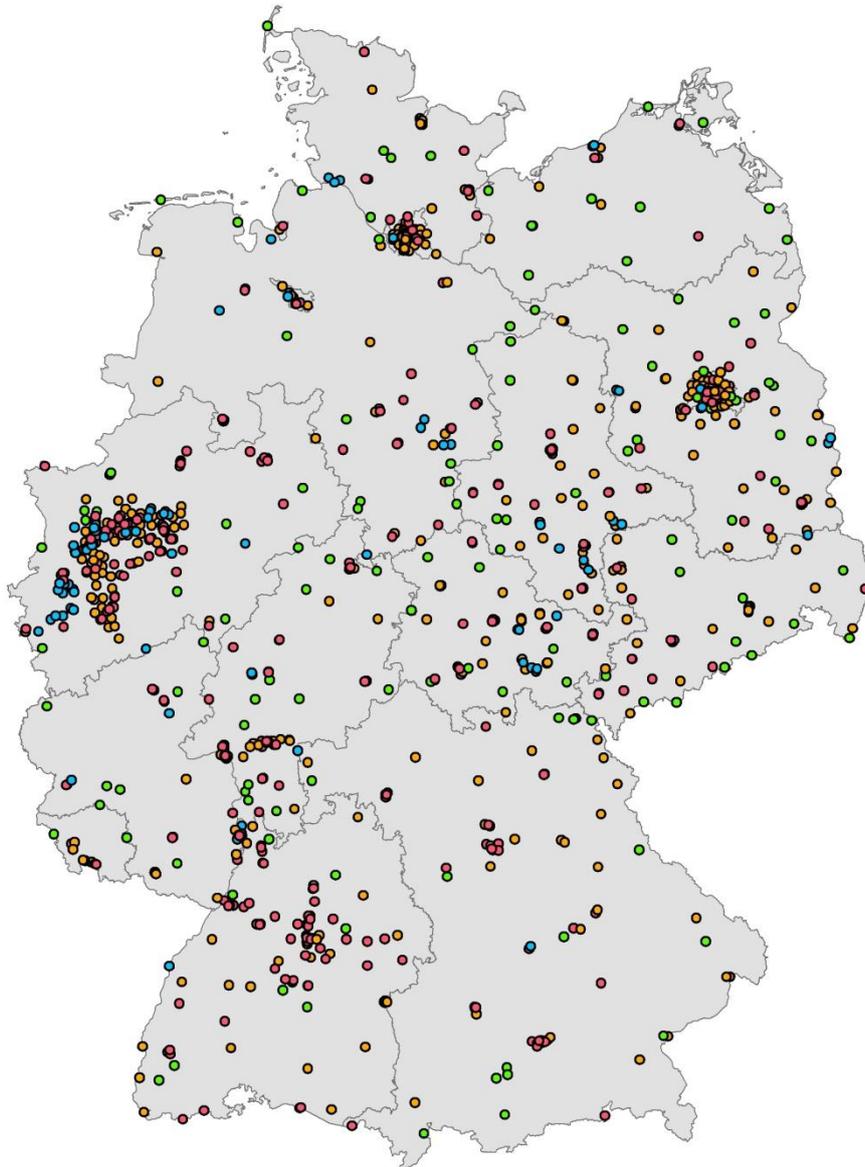


eigene Darstellung, BioMath GmbH

Tabelle 3: NO₂ - Anzahl aller Messstationen

Kategorie	Anzahl aller Stationen
ländlicher Hintergrund	133
städtischer Hintergrund	401
verkehrsnahe	509
Industrienah	69
Gesamt	1112

Abbildung 3: PM10 - Karte mit allen Messstationen von 1990-2016



eigene Darstellung, BioMath GmbH

Tabelle 4: PM10 - Anzahl aller Messstationen

Kategorie	Anzahl aller Stationen
ländlicher Hintergrund	128
städtischer Hintergrund	371
verkehrsnahe	322
Industrienah	95
Gesamt	916

Weiterhin übermittelte das UBA am 03.04.2018 seine Auswahl von Messreihen für die Berechnung der Jahresmittelwertzeitreihen je Kategorie für den Zeitraum von 2000 bis 2016. Die Auswahl wurde so

getroffen, dass alle durchgängigen Messreihen, aber mindestens 30 % der am stärksten besetzten Messreihen, je Kategorie ausgewählt wurden. Das 30 %-Kriterium wurde für NO₂ der Kategorie verkehrsnah auf Grund der starken Inhomogenität nicht angewandt.

Tabelle 5: Messreihen UBA-Auswahl

Kopfzeile	Kopfzeile	Kopfzeile	Kopfzeile
Ozon	ländlicher Hintergrund	49 von 122	durchgängige Messreihen
	städtischer Hintergrund	97 von 257	durchgängige Messreihen
	verkehrsnah	16 von 52	Messreihen mit mindestens 10 gültigen Werten
	industrienah	10 von 32	durchgängige Messreihen
NO ₂	ländlicher Hintergrund	54 von 112	durchgängige Messreihen
	städtischer Hintergrund	126 von 307	durchgängige Messreihen
	verkehrsnah	36 von 467	durchgängige Messreihen
	industrienah	19 von 62	Messreihen mit mindestens 10 gültigen Werten
PM ₁₀	ländlicher Hintergrund	37 von 102	Messreihen mit mindestens 16 gültigen Werten
	städtischer Hintergrund	96 von 273	durchgängige Messreihen
	verkehrsnah	92 von 293	Messreihen mit mindestens 11 gültigen Werten
	industrienah	26 von 85	Messreihen mit mindestens 10 gültigen Werten

3 Analyse

3.1 Deskriptive

Die deskriptive Auswertung der Daten der drei Luftschadstoffe umfasst die Beschreibung der vorliegenden Messreihen nach den Gesichtspunkten:

- ▶ Anzahl und Verteilung der Stationen
- ▶ Verlauf der Werte der Messreihen
- ▶ Aussagen linearer Trends der Messreihen mit einfachen Modellen

Die **Anzahl der Stationen** wurde einerseits für alle vorhandenen Stationen und andererseits für ausschließlich durchgängig messende Stationen analysiert. Unter durchgängig messenden Stationen werden Stationen verstanden, die von Betriebsbeginn bis Betriebsende in jedem Betriebsjahr einen gültigen Jahresmittelwert besaßen. Die Analysen erfolgten für alle Kategorien (ländlicher Hintergrund, städtischer Hintergrund, verkehrsnah, industrienah) zusammen und für jede Kategorie einzeln.

Um die Stationen und somit auch die Messreihen besser charakterisieren zu können, wurden für jede Station abgeleitete Variablen berechnet.

Tabelle 6: abgeleitete Variablen

Variable	Bedeutung
NJahr	Anzahl Messwerte der Station
MinJahr	Jahr in dem die Station den ersten Messwert erzeugte (Messbeginn)
MaxJahr	Jahr in dem die Station den letzten Messwert erzeugte (Messende)
DiffJahr	Anzahl Jahre zwischen Messbeginn und Messende (Betriebszeit)
NJahrFehl	Anzahl Jahre der Betriebszeit in denen keine Messwerte erzeugt wurden
MaxMinFehl	Kennzeichnet den Messbeginn, das Messende und die Anzahl der Jahre, in denen kein Messwert erzeugt wurde.

Basierend auf diesen Variablen wurden Balken-, Liniendiagramme und Boxplots erstellt und die Variablen tabellarisch gelistet. Weiterhin wurden die Stationen auf Karten visualisiert und die Anzahl der Stationen bezüglich der Bundesländer/ Betreiber grafisch dargestellt.

Für die Beschreibung des **Verlaufs der Werte der Messreihen** wurden diese in Liniendiagrammen bezüglich durchgängig messend, nicht durchgängig messend und bezüglich der Kategorien dargestellt. Weiterhin wurden je Messreihe statistische Maßzahlen berechnet und diese in Boxplots visualisiert. Die Maßzahlen wurden tabellarisch gelistet.

Für die **Berechnung der Trends** wurden die Messreihen in eine Rangordnung gebracht. Das Hauptkriterium für diese Rangordnung war die Anzahl der Werte der Messreihe und der Zeitraum, in dem die Messwerte erhoben wurden. Mit weiteren Nebenkriterien wurde die Rangordnung so verfeinert, dass eine möglichst ausbalancierte schrittweise Trendberechnung durchgeführt werden konnte (siehe 3.2.1). D.h. die Messreihen auf den ersten Rängen sind durchgängig messende Reihen, die auf den letzten Rängen sind Reihen mit den wenigsten Messwerten.

Die Trendberechnung erfolgte schrittweise: im ersten Schritt wurde der Trend für alle durchgängig messenden Reihen berechnet und in den folgenden Schritten wurde jeweils die im Rang nächste Messreihe in die Trendberechnung mit aufgenommen. Die Trendberechnung erfolgte mit einem linearen Regressionsmodell, in dem der Parameter Anstieg und dessen 95% Konfidenzintervall geschätzt wurden. Der Anstieg entspricht der jährlichen Zunahme bzw. Abnahme der Messwerte.

Mit Hilfe des Konfidenzintervalls kann getestet werden, ob der Anstieg ungleich Null ist, d.h. ob der Trend statistisch signifikant ist. Signifikanz auf dem Niveau α liegt genau dann vor, wenn das $(1-\alpha)$ Konfidenzintervall die Null nicht miteinschließt. Mit dem 95% Konfidenzintervall kann somit eine Signifikanz auf dem Niveau von 5% festgestellt werden.

Das lineare Regressionsmodell wurde einerseits mit den Messwerten und andererseits mit den jährlichen Differenzen der Messwerte gerechnet (siehe 3.2.2). Die Ergebnisse der schrittweisen Trendberechnung wurden in Liniendiagrammen und zusammenfassend in Konfidenzintervalldarstellungen visualisiert. Um die Trendberechnung bezüglich der Messwerte und bezüglich der Differenzen der Messwerte zu unterscheiden, wird dafür im Weiteren der Begriff *Datenbasis Messwerte* und *Datenbasis Differenzen* verwendet.

Anmerkung zur Trendberechnung

Im Hinblick darauf, dass eine Methode zur sicheren Trendaussage entwickelt werden soll, wurde mit der Modellrechnung für jährliche Differenzen ein erster Ansatz verfolgt. In diesem Ansatz werden die Messwerte der Reihen auf ihre jährliche Veränderung (= „den Trend“) reduziert. Damit sind die Messreihen unabhängig von ihrem Messniveau. D.h. Verfälschungen von Trends, weil Messreihen auf einem höheren oder niedrigeren Niveau als im Mittel sind, können damit ausgeschlossen werden.

3.1.1 Ozon

Anzahl und Verteilung der Stationen

Der überwiegende Teil der Messstationen gehört mit 55,4 % zum städtischen Hintergrund, gefolgt vom ländlichen Hintergrund mit 25,4 %. Die Kategorien verkehrsnah mit 12,6 % und industrienah mit 6,6 % sind deutlich geringer vertreten. Betrachtet man nur die durchgängig messenden Stationen von 1990 - 2016 so sind es gerade mal 34 von 547 Stationen, d.h. 6,2 % (siehe Abbildung 1 und 3.1.1.2).

Von 1990 bis 1997 steigt die Anzahl der Stationen stark an, hält sich dann 5 Jahre lang etwa auf diesem Niveau und fällt dann leicht wieder ab. Dieser Verlauf ist auch in den einzelnen Kategorien zu verzeichnen. Nur bei der Kategorie industrienah ist dieser Verlauf auf Grund der sehr geringen Anzahl von Stationen nicht so deutlich (siehe 3.1.1.1).

Insgesamt zeigen die jeweils durchgängig messenden Stationen vom Jahr i ($i > 1990$) bis 2016 eine kontinuierliche Zunahme. Dieses widerspiegelt sich in den Kategorien städtischer und ländlicher Hintergrund. Die Kategorie verkehrsnah zeigt ein ganz anderes Verhalten. Hier gibt es nur sehr wenige Stationen, die durchgängig messen, dazu aber überproportional viele Stationen, die nicht durchgängig messen. Z.B. Im Jahr 1997 gibt es 4 Stationen, die bis 2016 durchgängig und 46 Stationen die nicht durchgängig messen. Die Kategorie industrienah ähnelt bezüglich des Verhältnisses durchgängig / nicht durchgängig messend eher dem städtischen und ländlichen Hintergrund, allerdings mit deutlich weniger Stationen (siehe 3.1.1.2 und 3.1.1.3).

Im Jahr 1990 gab es nur Stationen im westlichen Teil Deutschlands und West-Berlin. Bis zum Jahre 1995 nahm die Anzahl der Stationen im östlichen Teil Deutschland deutlich zu. Das erklärt den schon genannten starken Anstieg der Anzahl von Stationen von 1990 bis 1997. Ab 1995 gibt es in gesamt Deutschland eine flächendeckende Anzahl von durchgängig messenden Stationen (siehe 3.1.1.4 bis 3.1.1.7).

Inbetriebnahme und Außerbetriebnahme von Stationen sind heterogen über den Zeitabschnitt 1990 - 2016 verteilt. Das betrifft alle Kategorien. Die Betriebszeiten schwanken von 1 Jahr bis zu 27 Jahren und sind im Mittel etwa 20 Jahre. Nur 0,5 % der Stationen waren lediglich 1 Jahr in Betrieb. Knapp 70 % der Stationen lieferten in jedem Betriebsjahr einen Wert. Bei etwas über 20 % der Stationen fehlte ein Wert während der Betriebszeit. Würde man diesen einen fehlenden Wert auffüllen, so stellen über 90 % der Stationen vollständige Messreihen in ihren Betriebszeiten bereit (siehe 3.1.1.8).

Verlauf der Werte der Messreihen

Die Werte der Messreihen betragen im Mittel $45,6 \mu\text{g}/\text{m}^3$ und schwanken zwischen $10,8$ und $105 \mu\text{g}/\text{m}^3$. Betrachtet man nur die durchgängig messenden Reihen, so zeigen diese innerhalb der Kategorien städtischer und ländlicher Hintergrund Verläufe auf unterschiedlichen Niveaus. Z.B. sind bei den durchgängig messenden Reihen der Kategorie ländlicher Hintergrund im Liniendiagramm drei Niveau-Gruppen zu erkennen (im Mittel etwa 40 , 60 und etwas über $80 \mu\text{g}/\text{m}^3$). In der Kategorie industrienah gibt es nur eine nicht durchgängig messende Reihe auf einem Niveau im Mittel um $26 \mu\text{g}/\text{m}^3$. Alle Messreihen der Kategorie verkehrsnah sind auf einem einheitlichen Niveau (siehe 3.1.1.9). Messreihen mit hohen Abweichungen zum mittleren Niveau sind im Boxplot gut erkennbar - 4 Reihen ländlicher Hintergrund, 1 Reihe städtischer Hintergrund (siehe 3.1.1.11).

Die mittleren Messreihen der durchgängig und nicht durchgängig messenden Reihen unterscheiden sich deutlich in den Kategorien städtischer Hintergrund, verkehrsnah und industrienah. Hier sind die Mittelwerte der durchgängig messenden Reihe immer niedriger als die der nicht durchgängig messenden Reihen. In der Kategorie ländlicher Hintergrund gibt nicht so starke Abweichungen, ab 1999 sind sie sich sehr ähnlich (siehe 3.1.1.10).

Trends der Messreihen

In der schrittweisen Trendberechnung gibt es in allen Kategorien Unterschiede im Verlauf der Anstiege zwischen der Datenbasis der Messwerte und der der Differenzen der Messwerte.

In der Kategorie ländlicher Hintergrund sind die Abweichungen relativ gering. Die Anstiege sind fast immer positiv knapp über Null aber nur ab Rang 449 signifikant für die Datenbasis der Messwerte (siehe 3.1.1.12).

In der Kategorie städtisch verkehrsnah sind die Anstiege der Datenbasis der Differenzen der Messwerte immer gleichmäßig niedriger als die der Datenbasis der Messwerte. Dieses steht im Zusammenhang mit dem höheren Niveau der nicht durchgängig messenden Reihen. Diese Reihen verzerren den Anstieg ins Positive. Alle Anstiege sind positiv und signifikant (siehe 3.1.1.13).

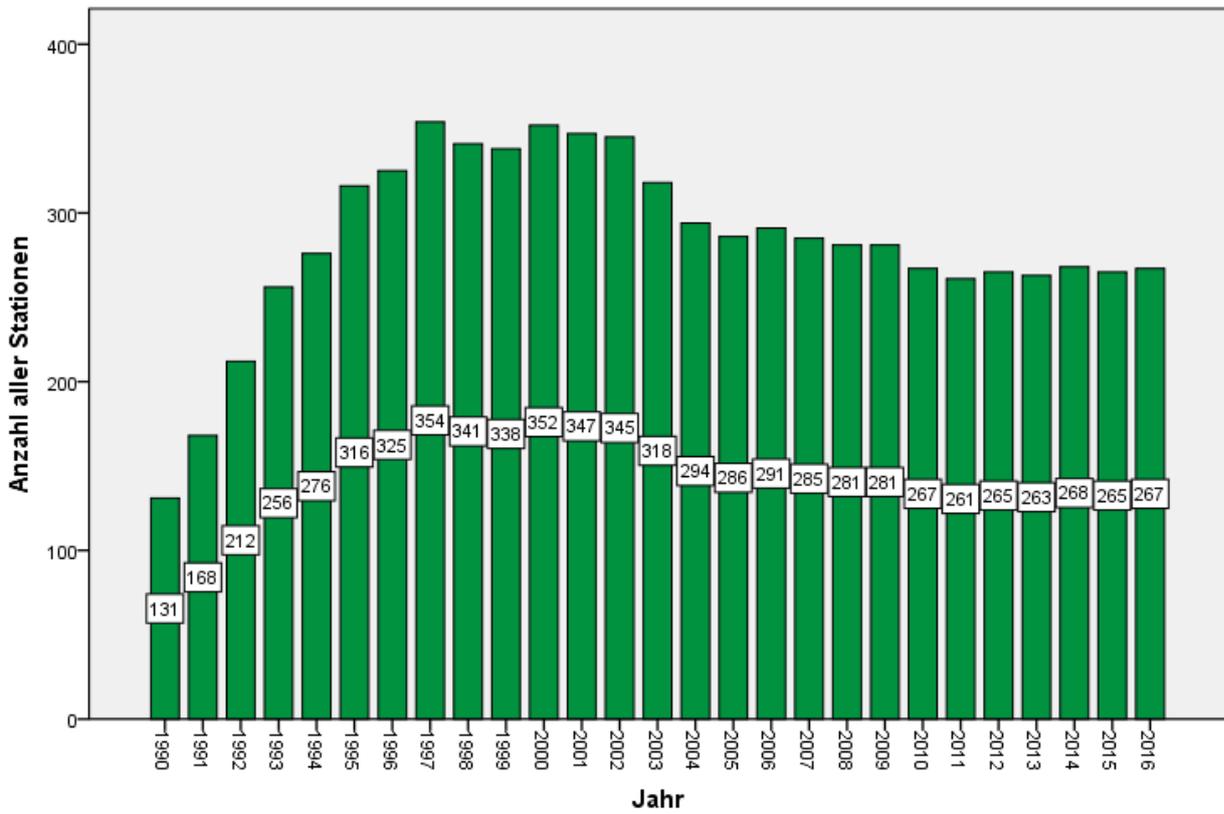
In der Kategorie verkehrsnah gibt es bis zum Rang 140 starke Abweichungen der Anstiege der Datenbasen. Die Ursache hierfür ist, dass nur 4 Messreihen bis zu diesem Rang in die Berechnung einfließen.

Bis zum Rang 356 sind die Unterschiede der Anstiege sehr gering. In den folgenden Rängen kommt es dann zu einem auseinanderdriften der Anstiege. Die Ursache dafür ist die zu geringe Anzahl von Messwerten der hinzukommenden Reihen. Alle Anstiege sind positiv und signifikant (siehe 3.1.1.14).

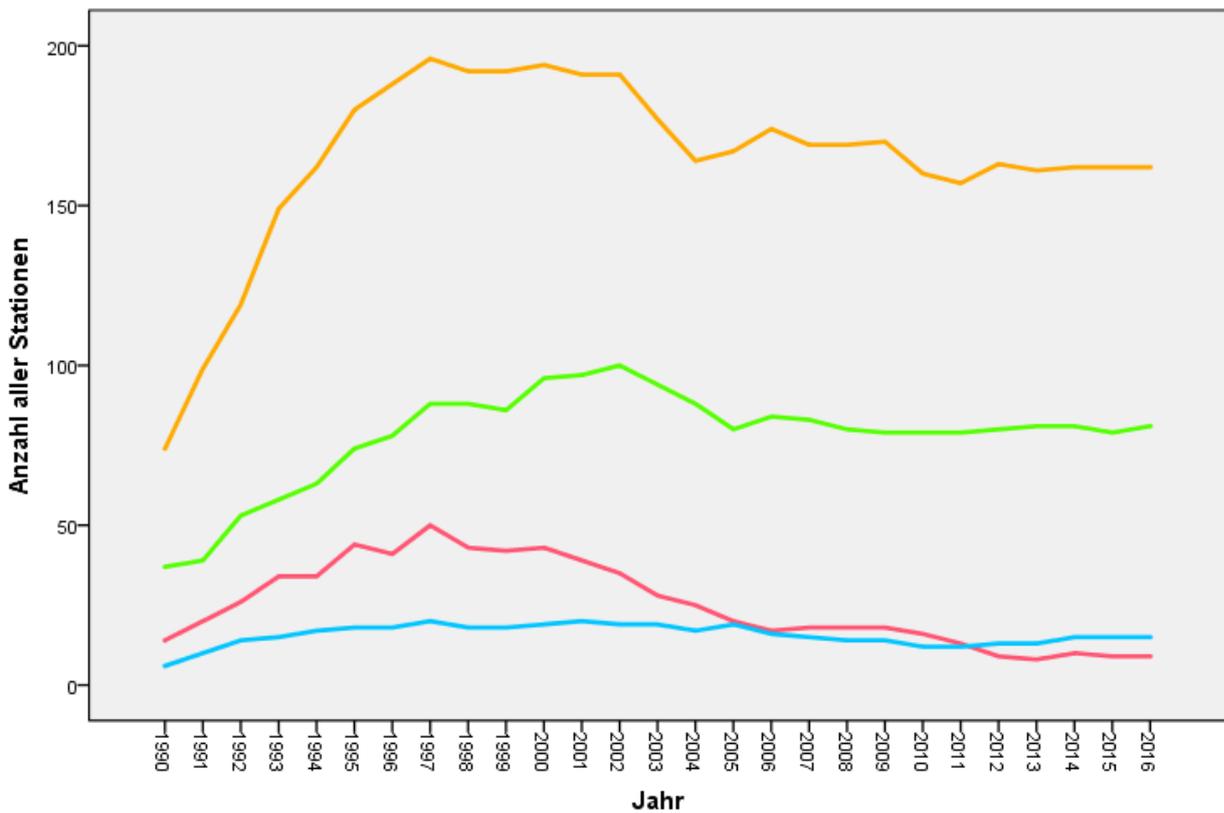
In der Kategorie industrienah gehen die Anstiege der beiden Datenbasen ab Rang 103 gleichmäßig auseinander. Die Anstiege der Datenbasis der Differenzen der Messwerte sind dann immer gleichmäßig niedriger als die der Datenbasis der Messwerte. Auch hier ist die Ursache die zu geringe Anzahl von Messwerten der hinzukommenden Reihen. Alle Anstiege sind positiv und signifikant (siehe 3.1.1.15).

Die Ergebnisse der Trendberechnung (Messwerte vs. Differenzen und durchgängig messend vs. nicht durchgängig messend) sind in den Fehlerbalkendiagrammen mit 95% Konfidenzintervallen zusammengefasst (siehe 3.1.1.16).

3.1.1.1 Ozon - Anzahl aller Stationen

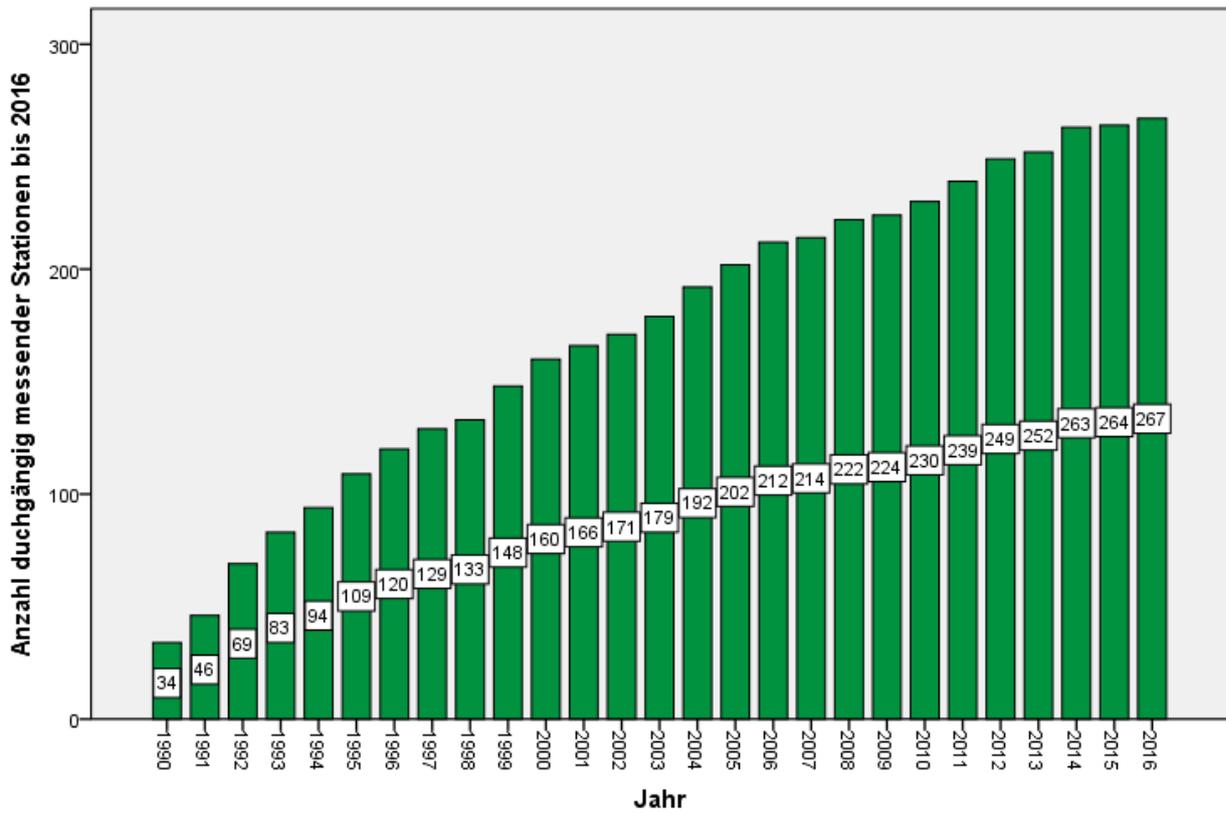


eigene Darstellung, BioMath GmbH

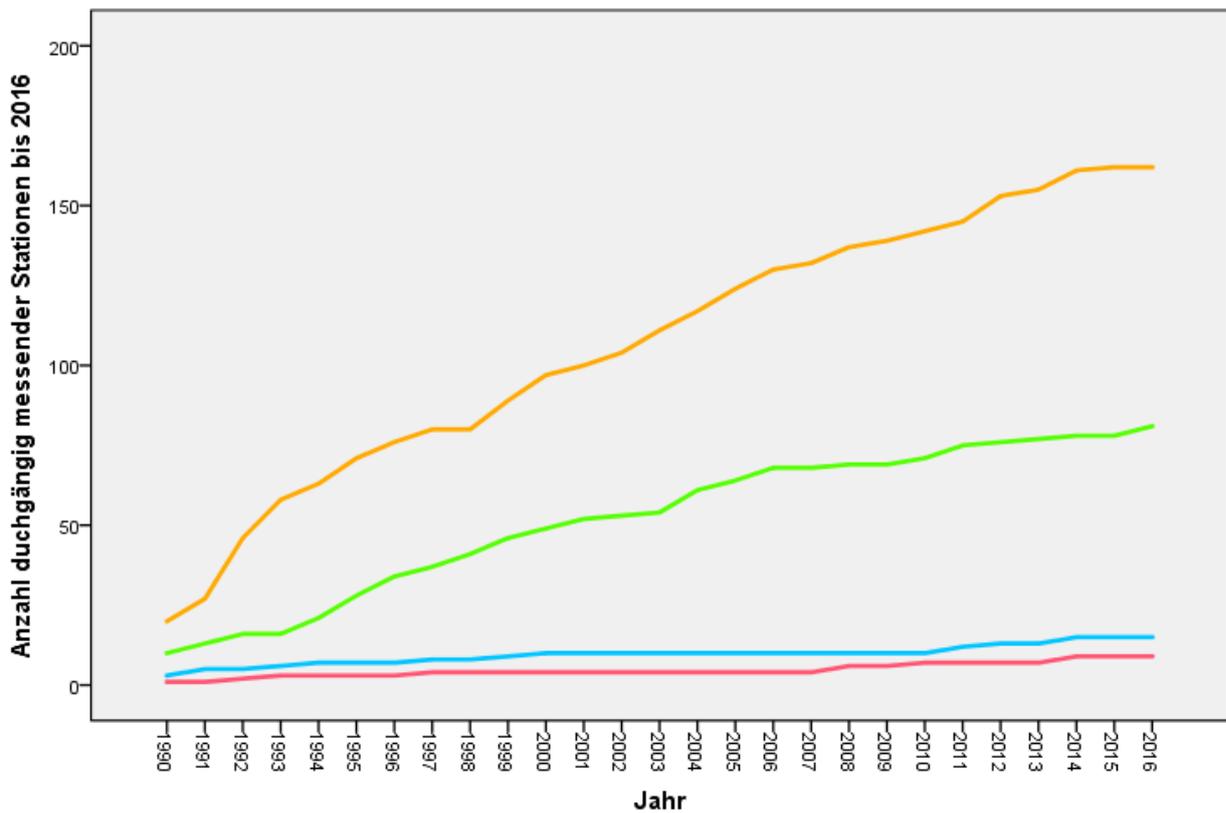


— ländlicher Hintergrund — städtischer Hintergrund — verkehrsnahe — industrienah
eigene Darstellung, BioMath GmbH

3.1.1.2 Ozon - Anzahl durchgängig messender Stationen bis 2016

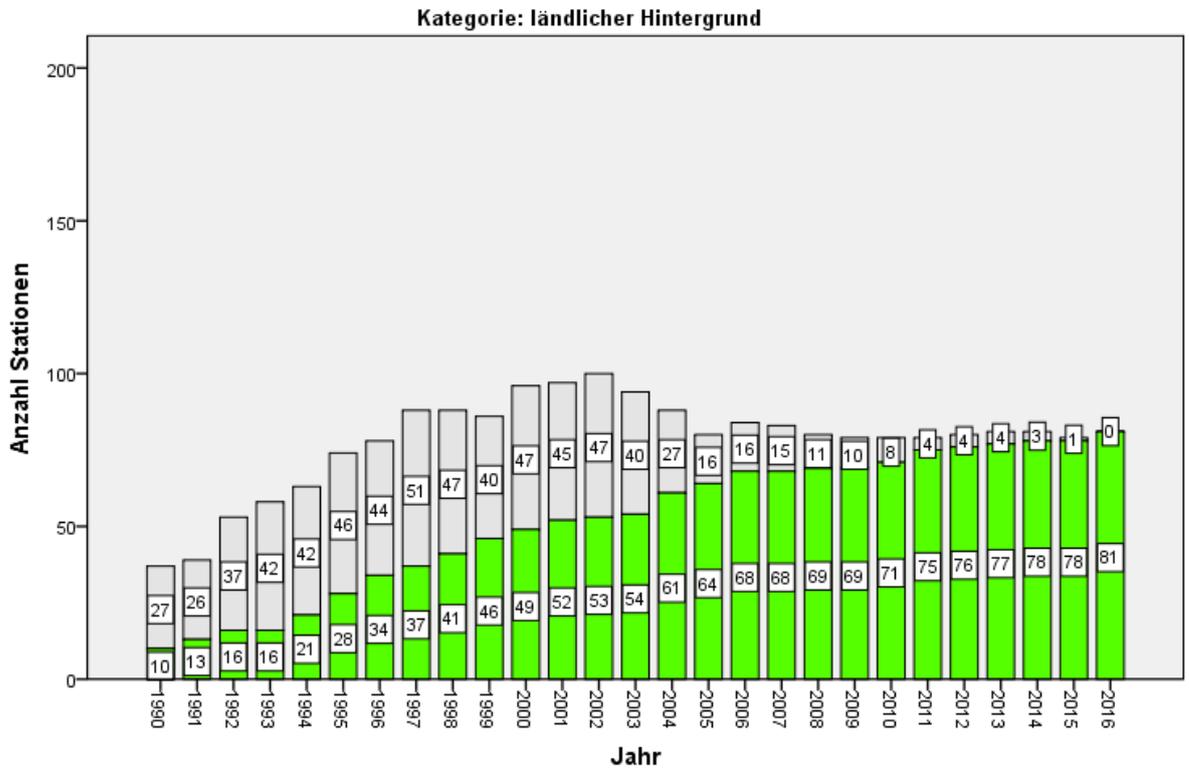


eigene Darstellung, BioMath GmbH

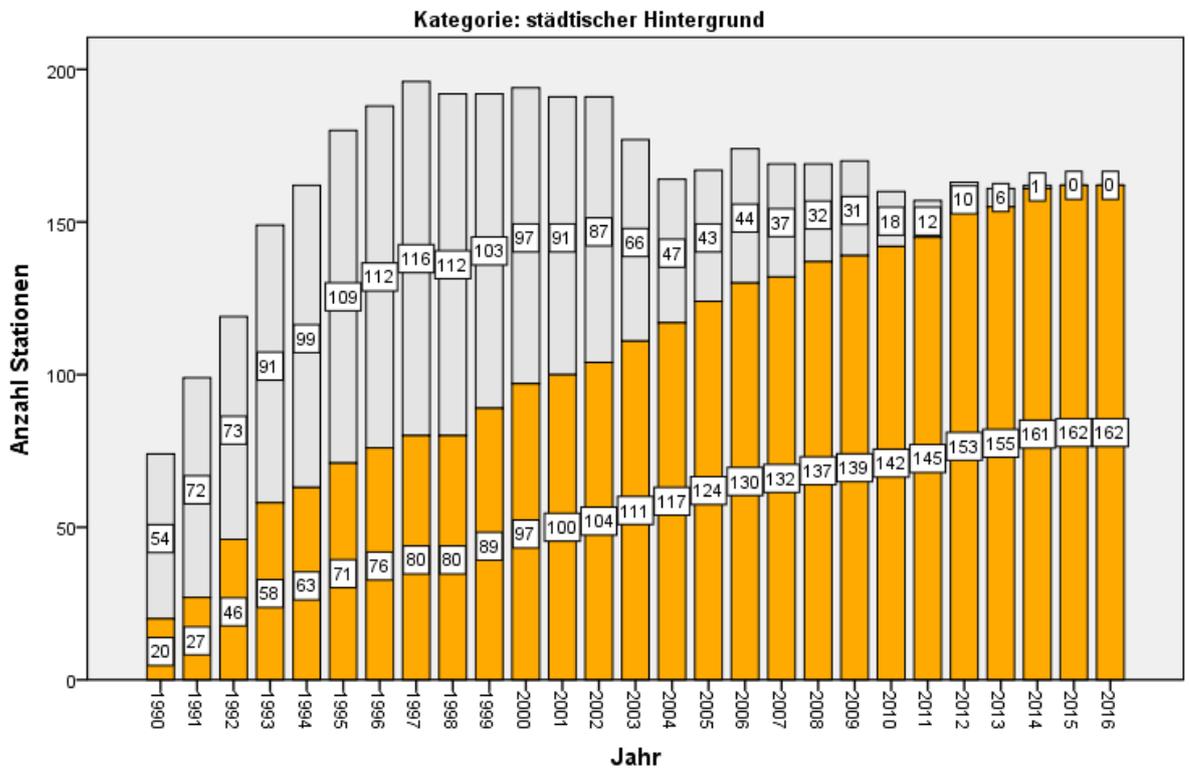


— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnah
 — industrienah
 eigene Darstellung, BioMath GmbH

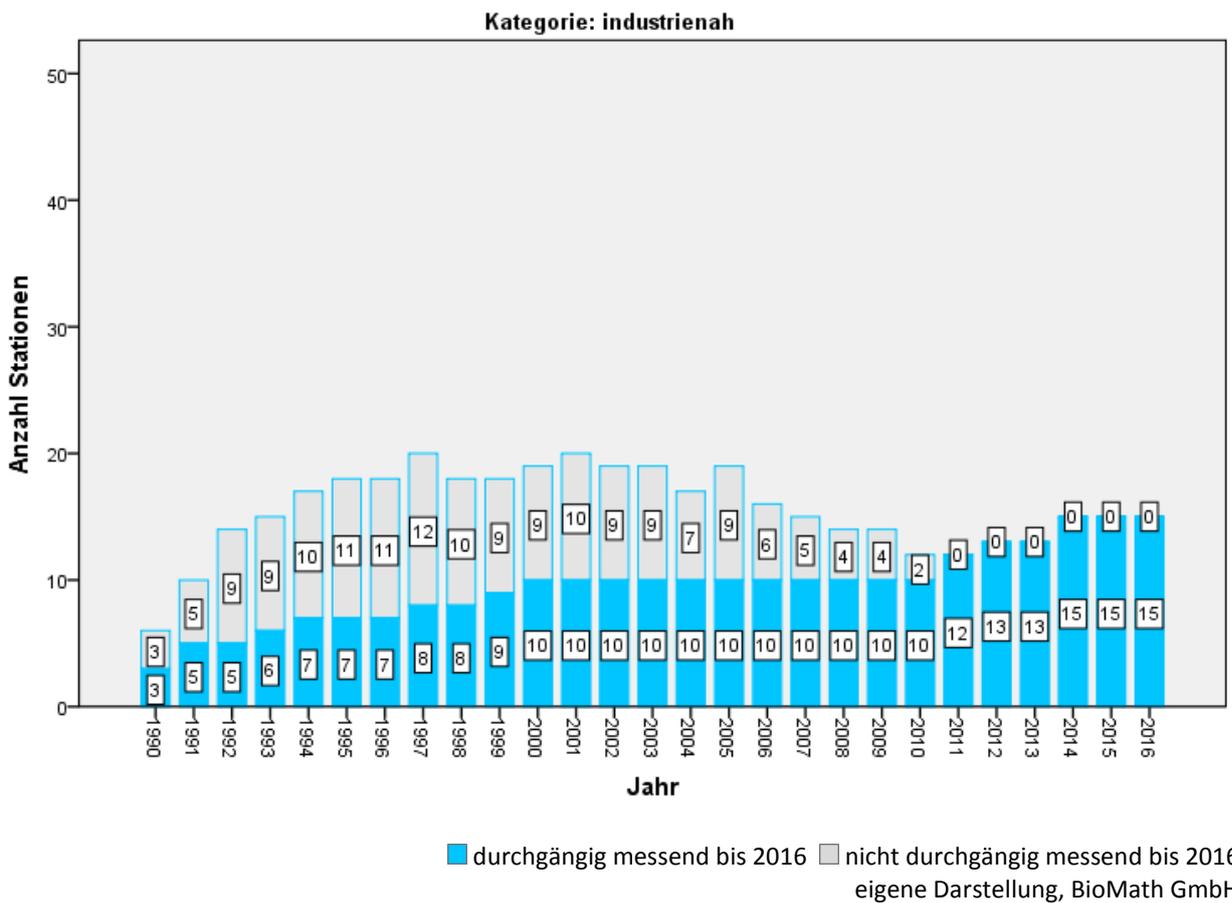
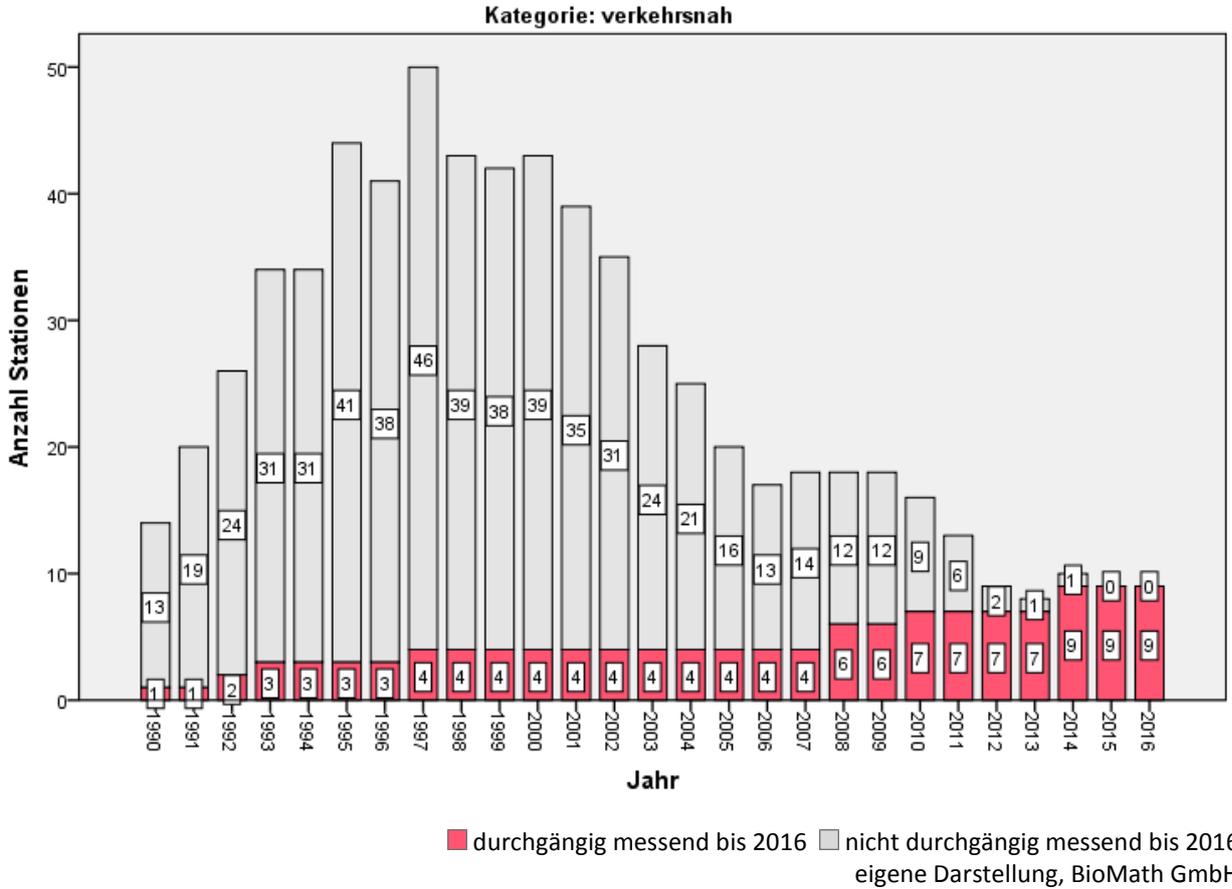
3.1.1.3 Ozon - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016



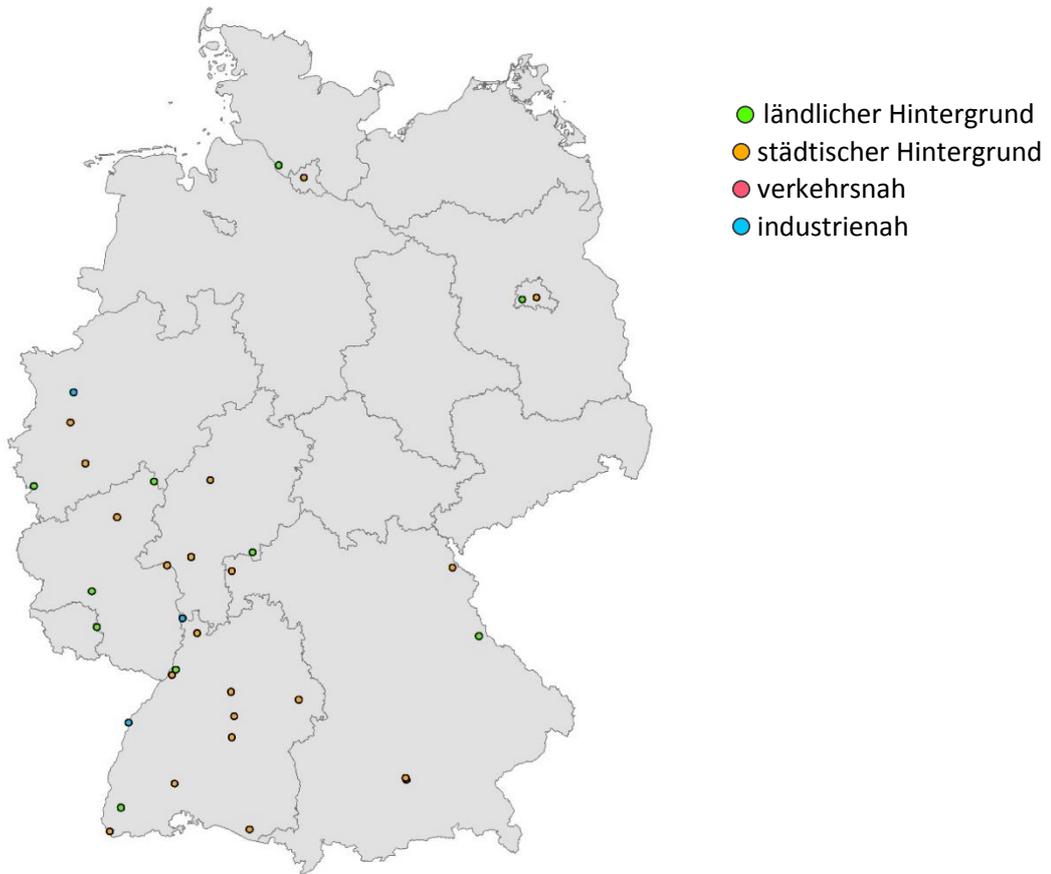
■ durchgängig messend bis 2016 ■ nicht durchgängig messend bis 2016
eigene Darstellung, BioMath GmbH



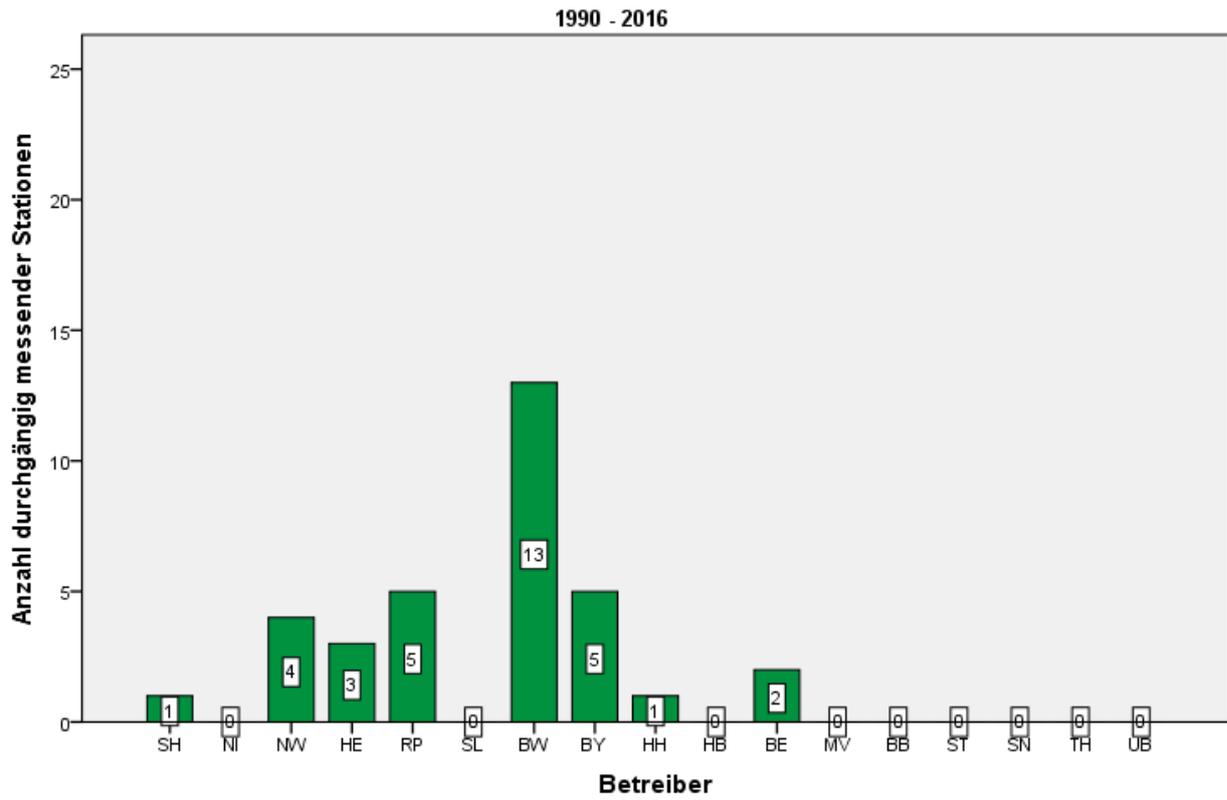
■ durchgängig messend bis 2016 ■ nicht durchgängig messend bis 2016
eigene Darstellung, BioMath GmbH



3.1.1.4 Ozon - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)

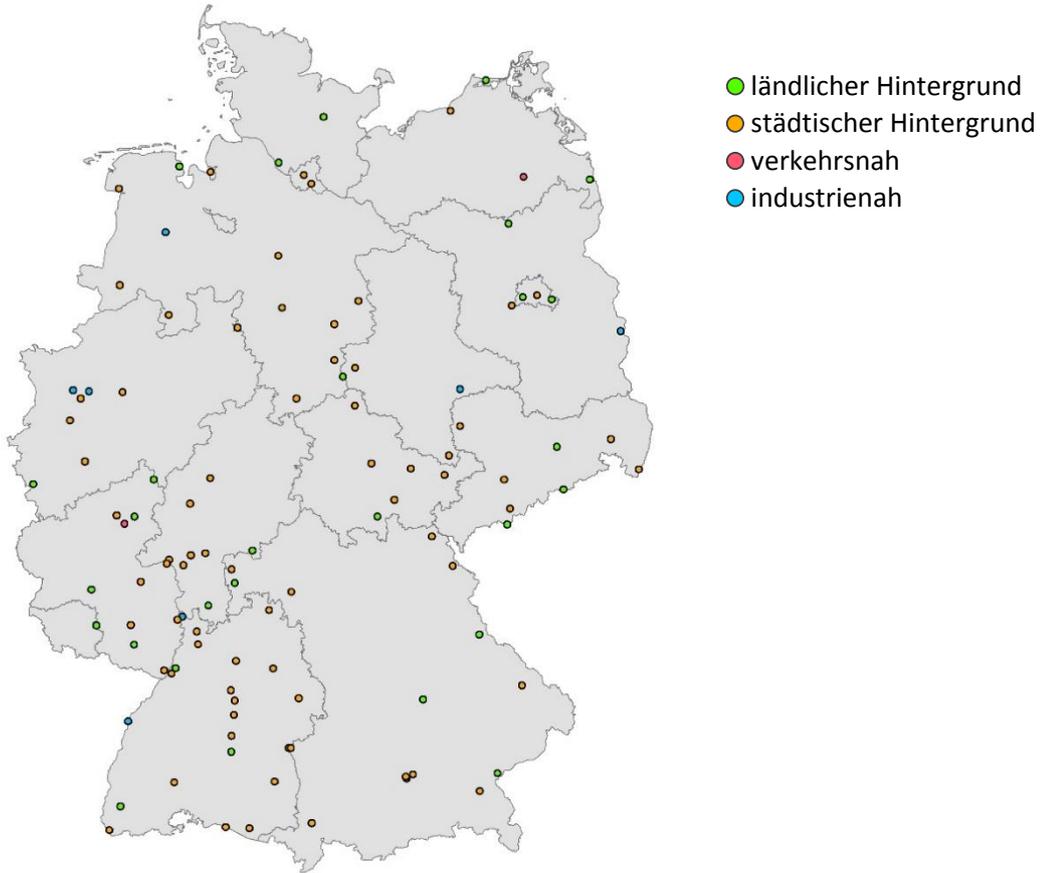


eigene Darstellung, BioMath GmbH

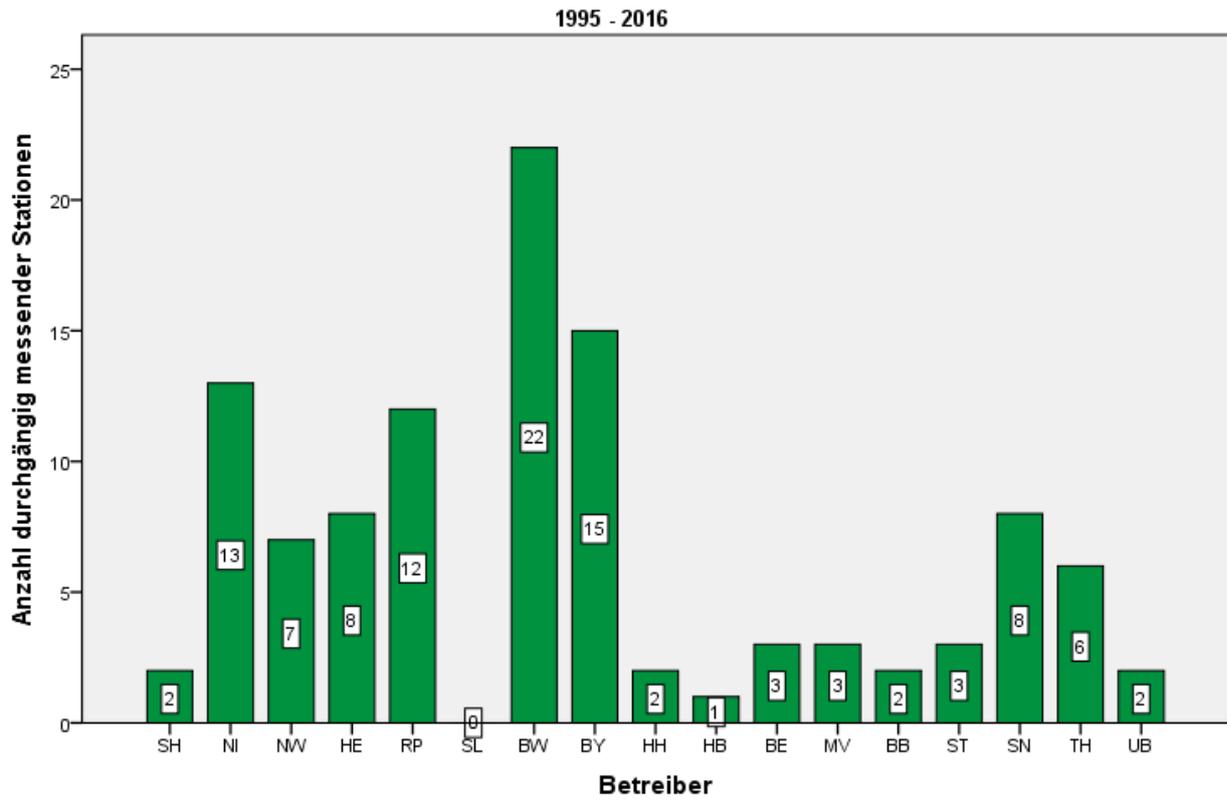


eigene Darstellung, BioMath GmbH

3.1.1.5 Ozon - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)

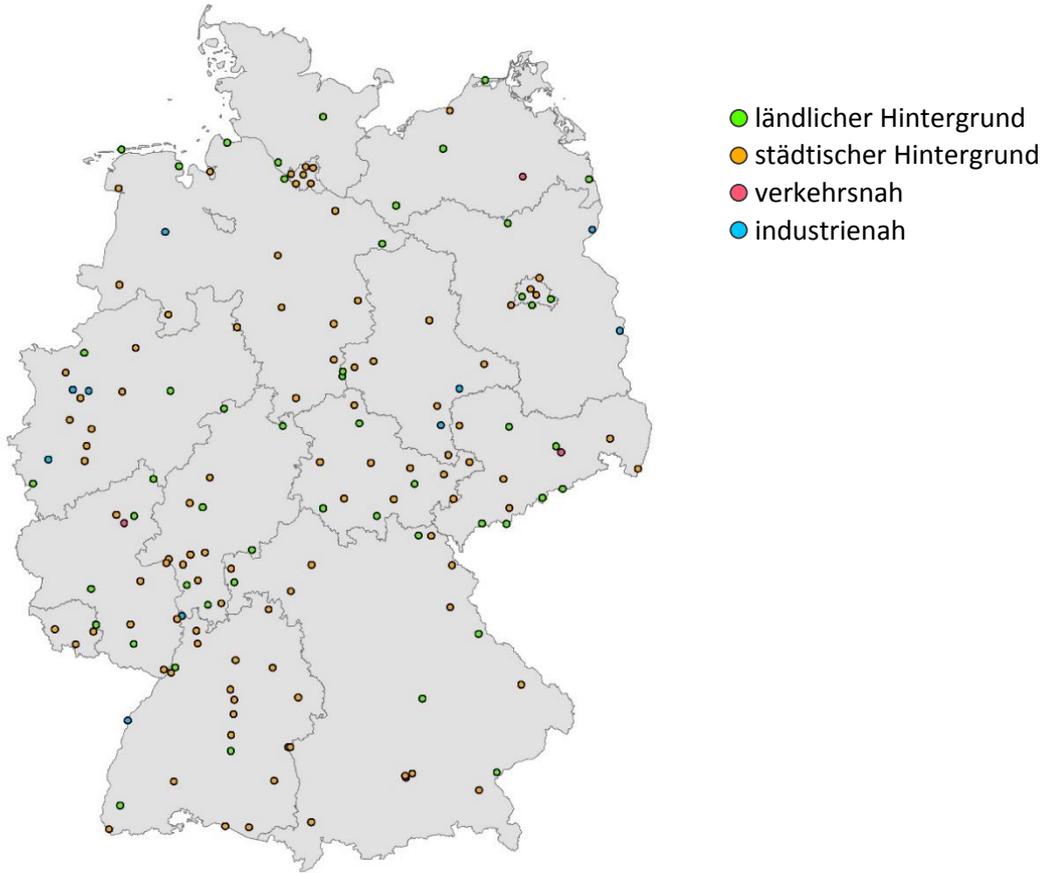


eigene Darstellung, BioMath GmbH

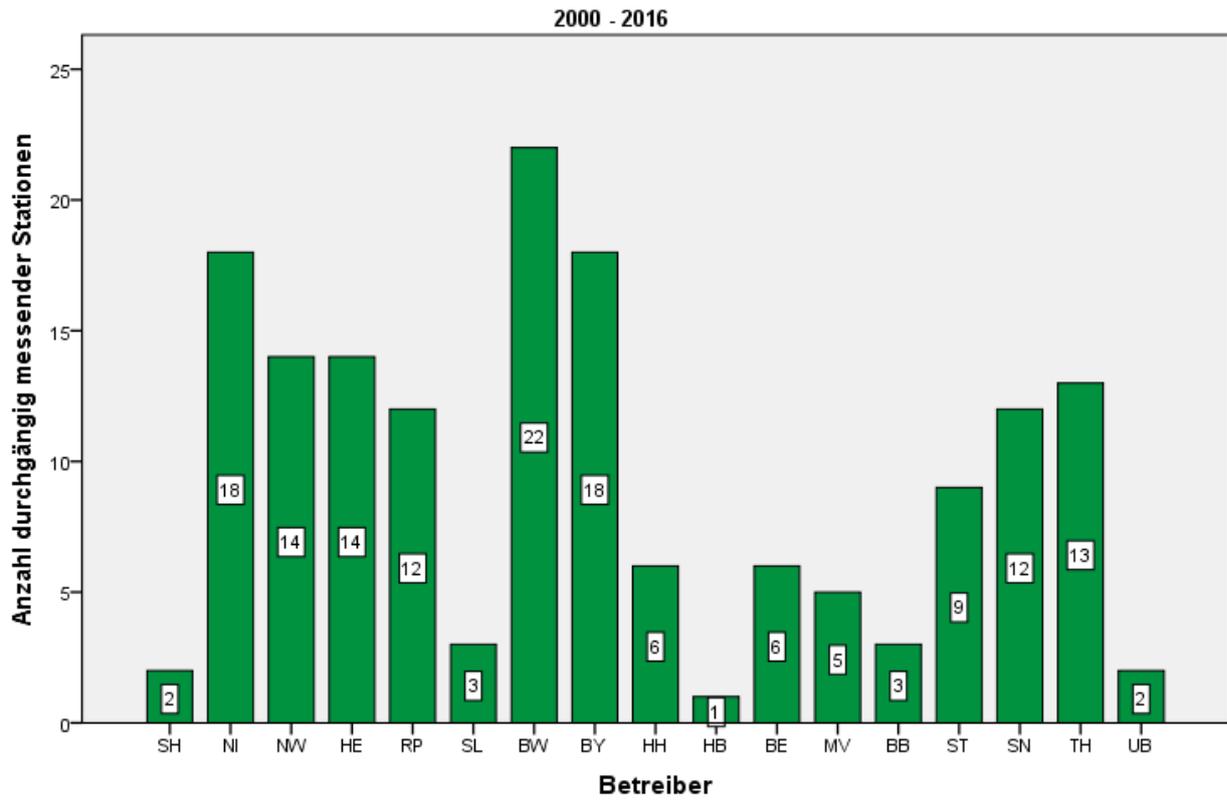


eigene Darstellung, BioMath GmbH

3.1.1.6 Ozon - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)

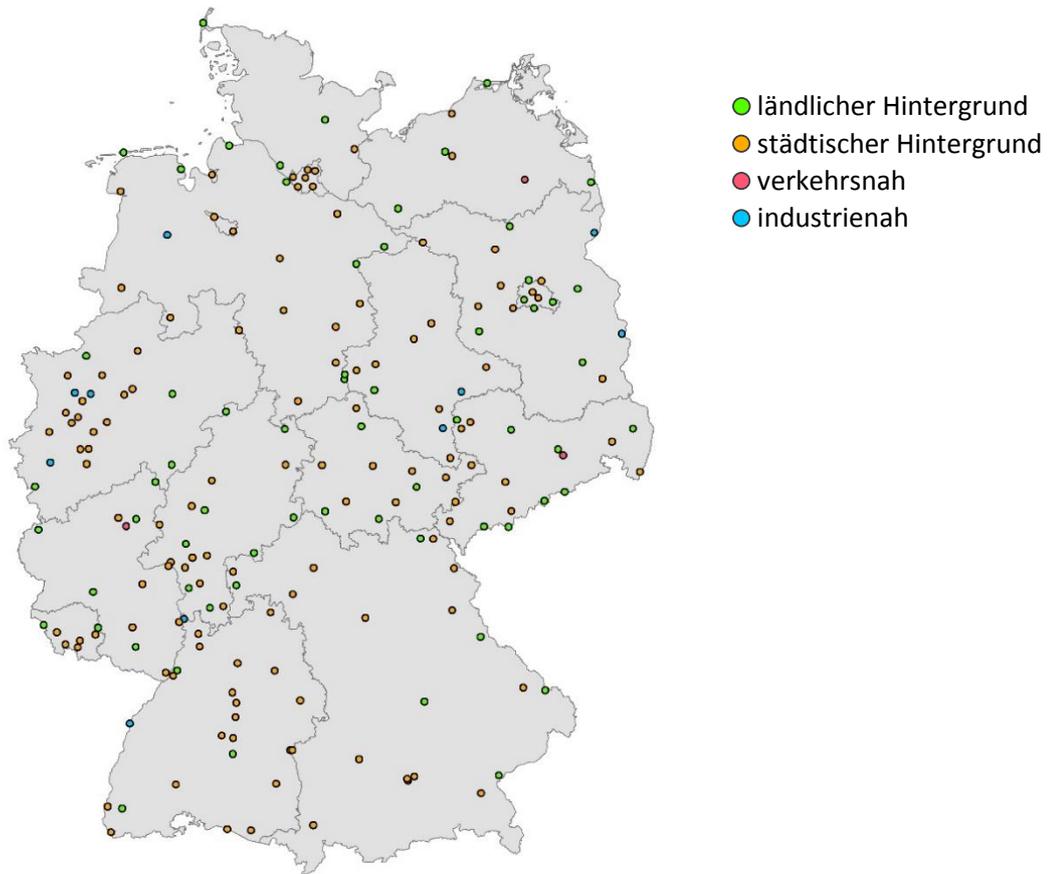


eigene Darstellung, BioMath GmbH

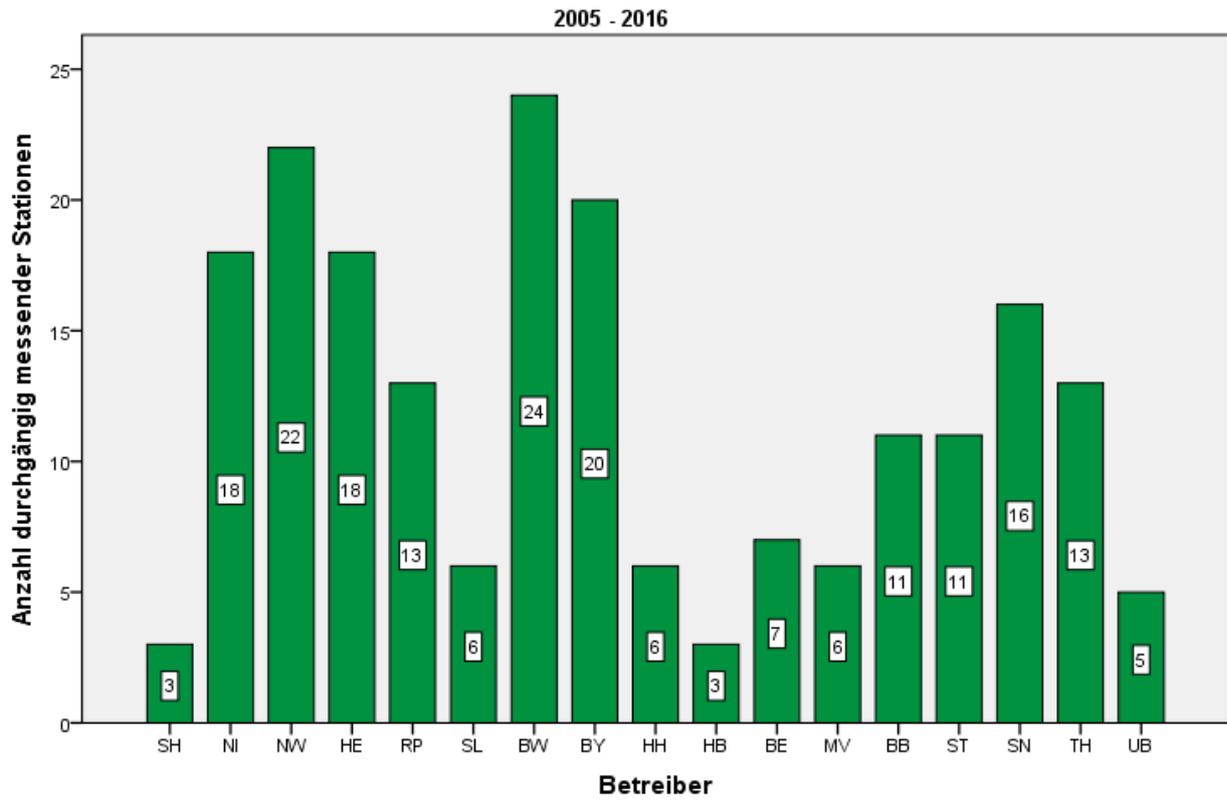


eigene Darstellung, BioMath GmbH

3.1.1.7 Ozon - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)

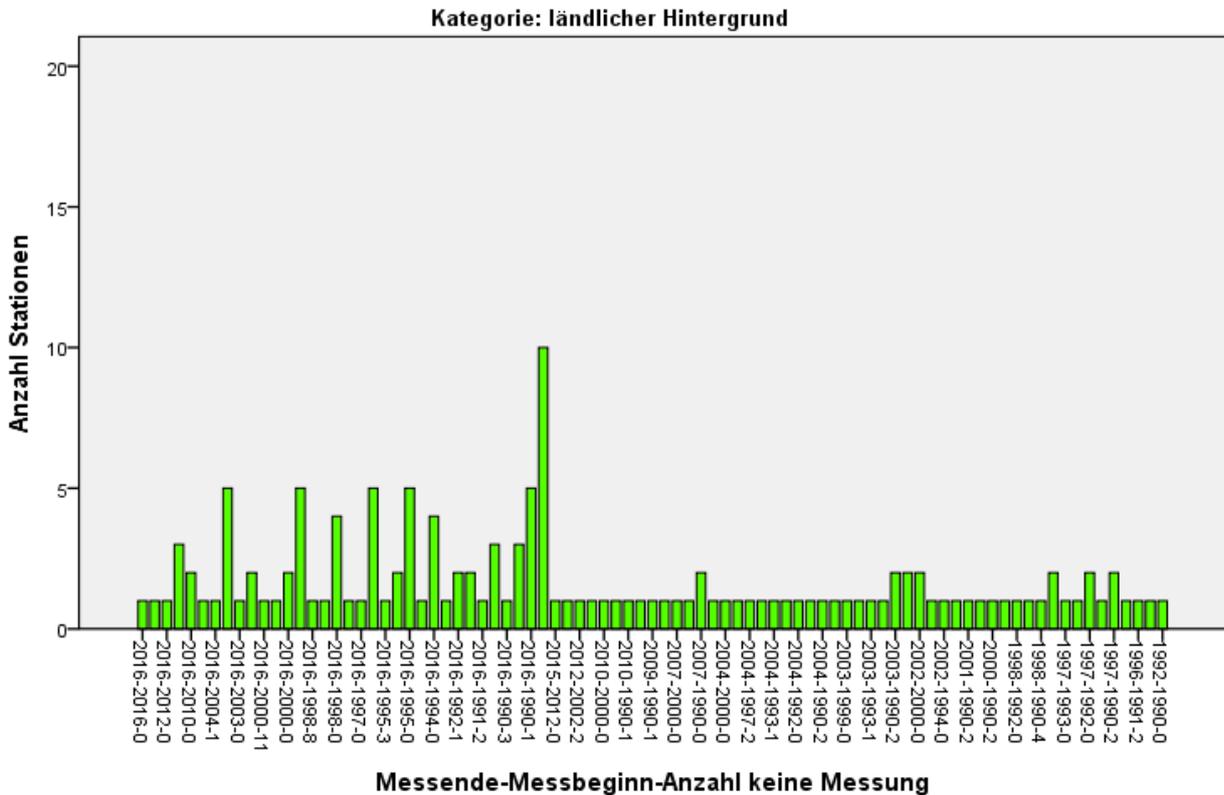


eigene Darstellung, BioMath GmbH

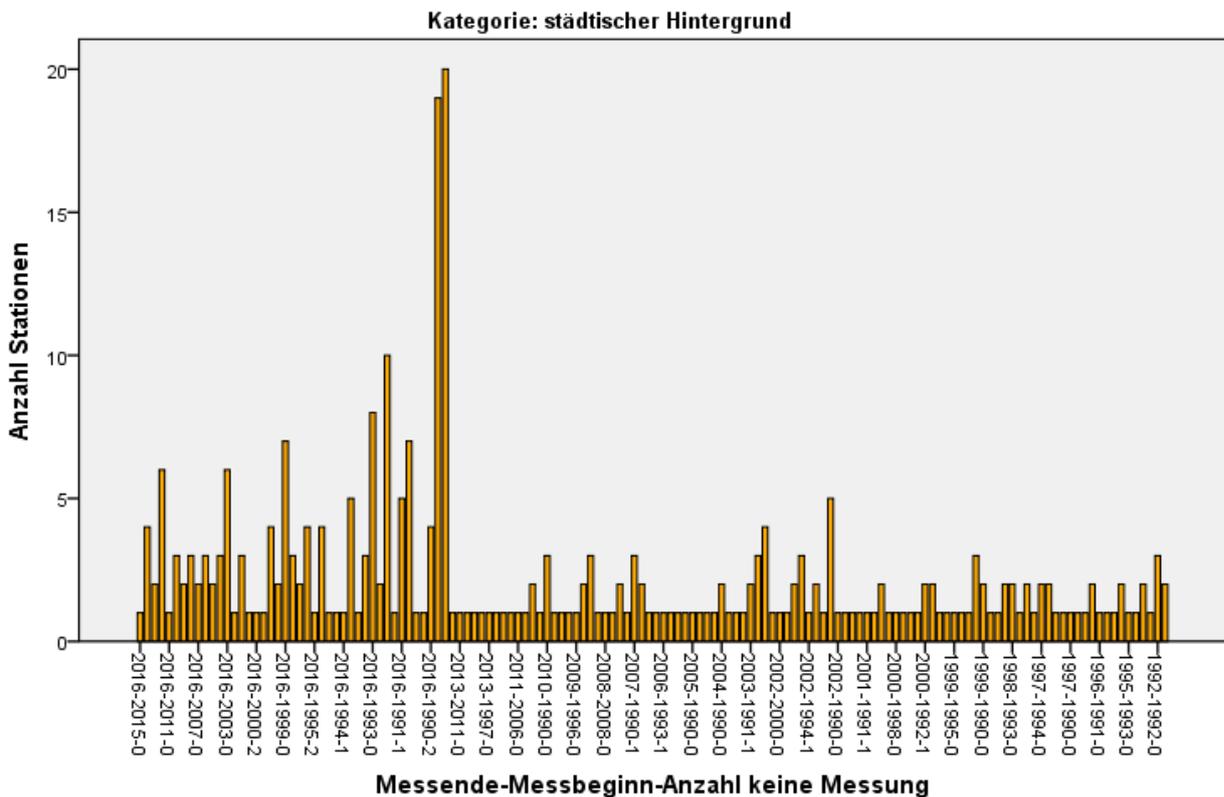


eigene Darstellung, BioMath GmbH

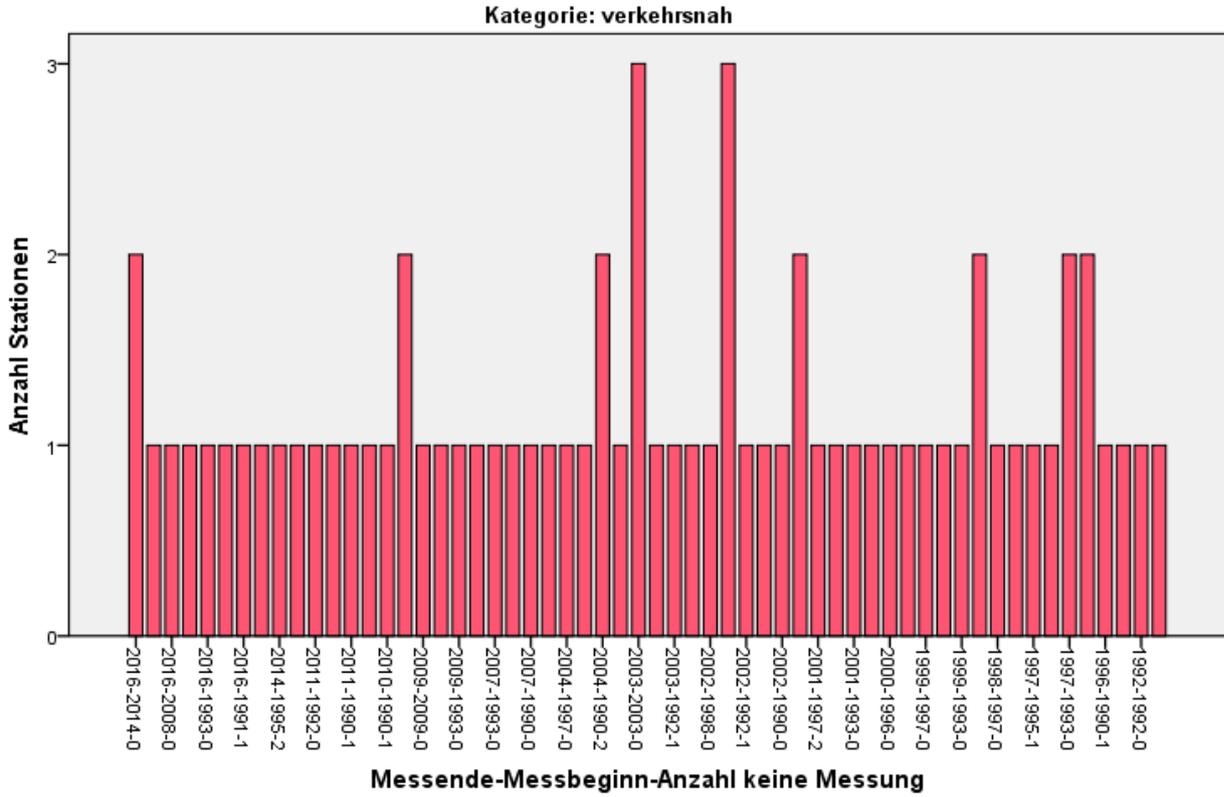
3.1.1.8 Ozon - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden



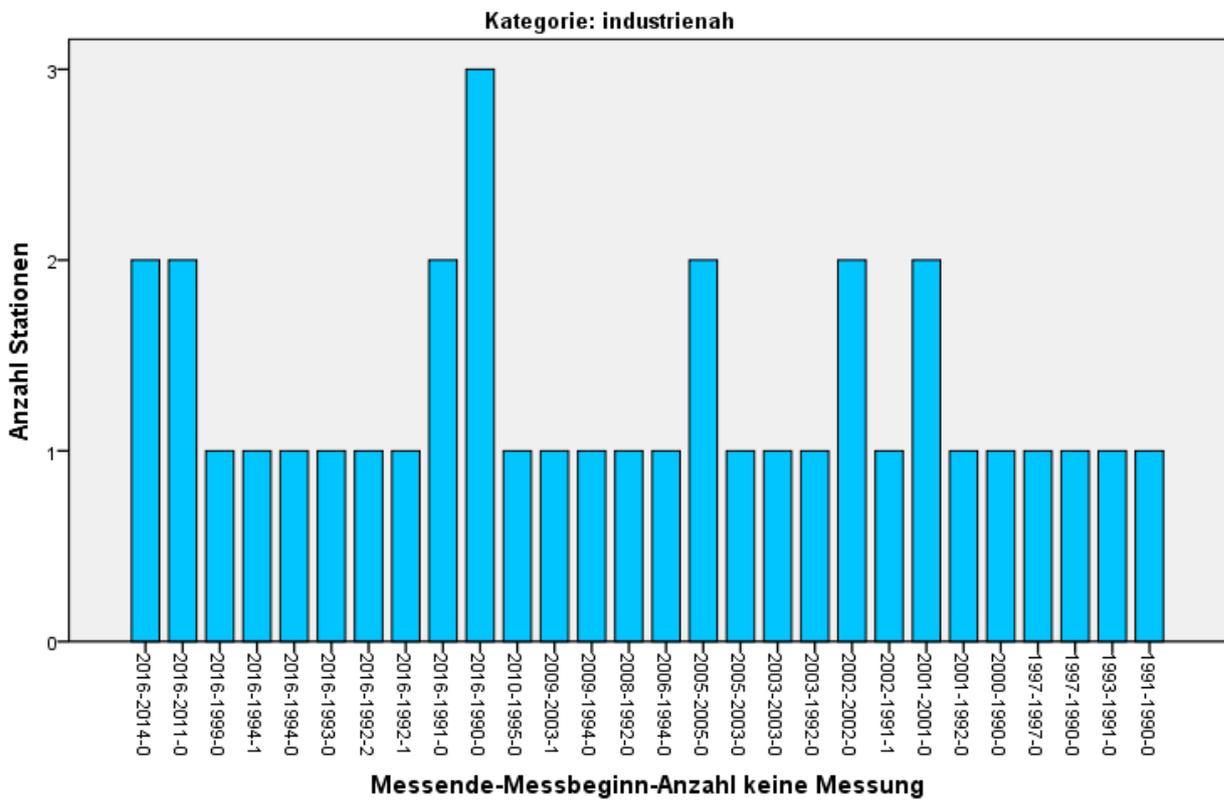
eigene Darstellung, BioMath GmbH



eigene Darstellung, BioMath GmbH

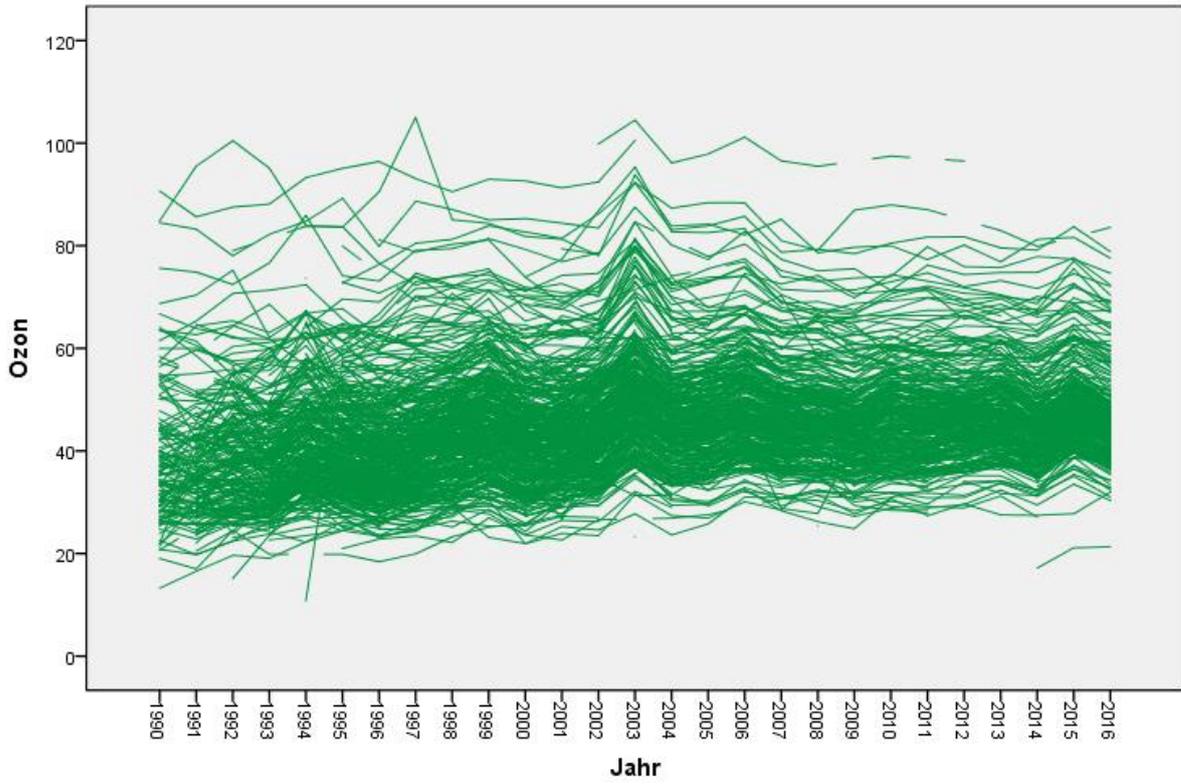


eigene Darstellung, BioMath GmbH

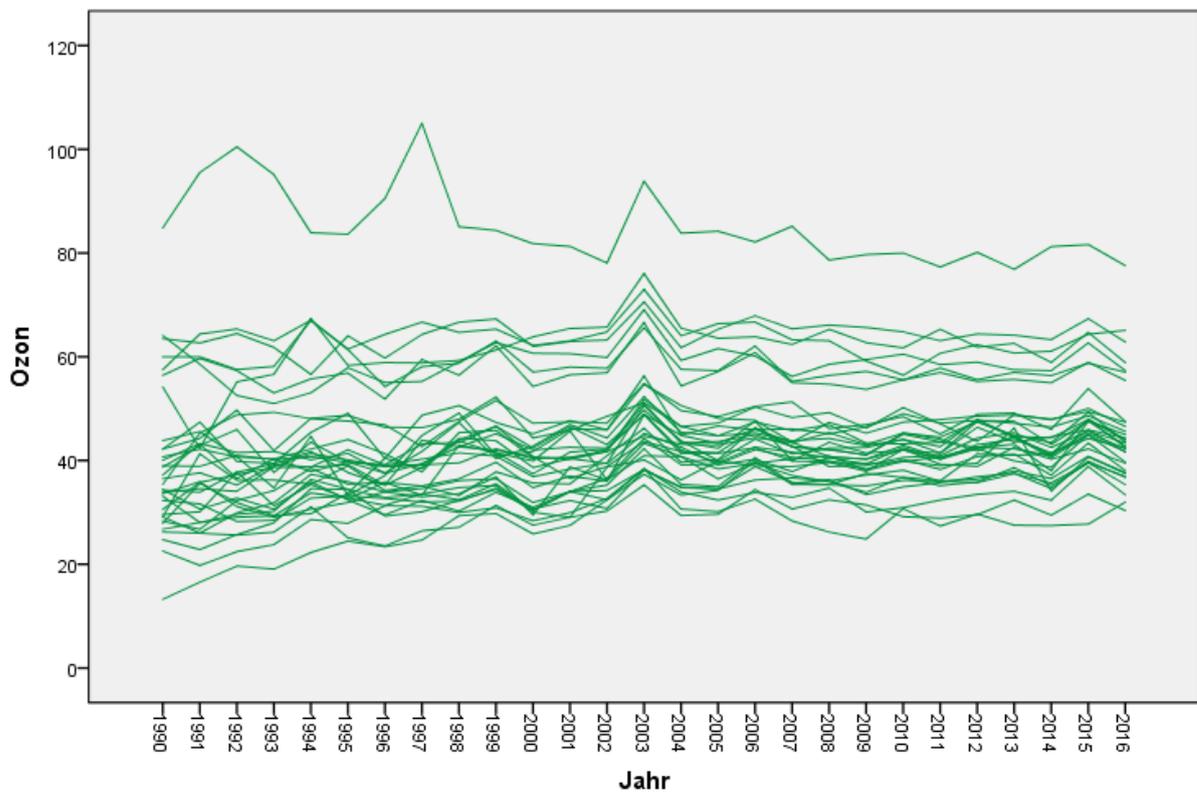


eigene Darstellung, BioMath GmbH

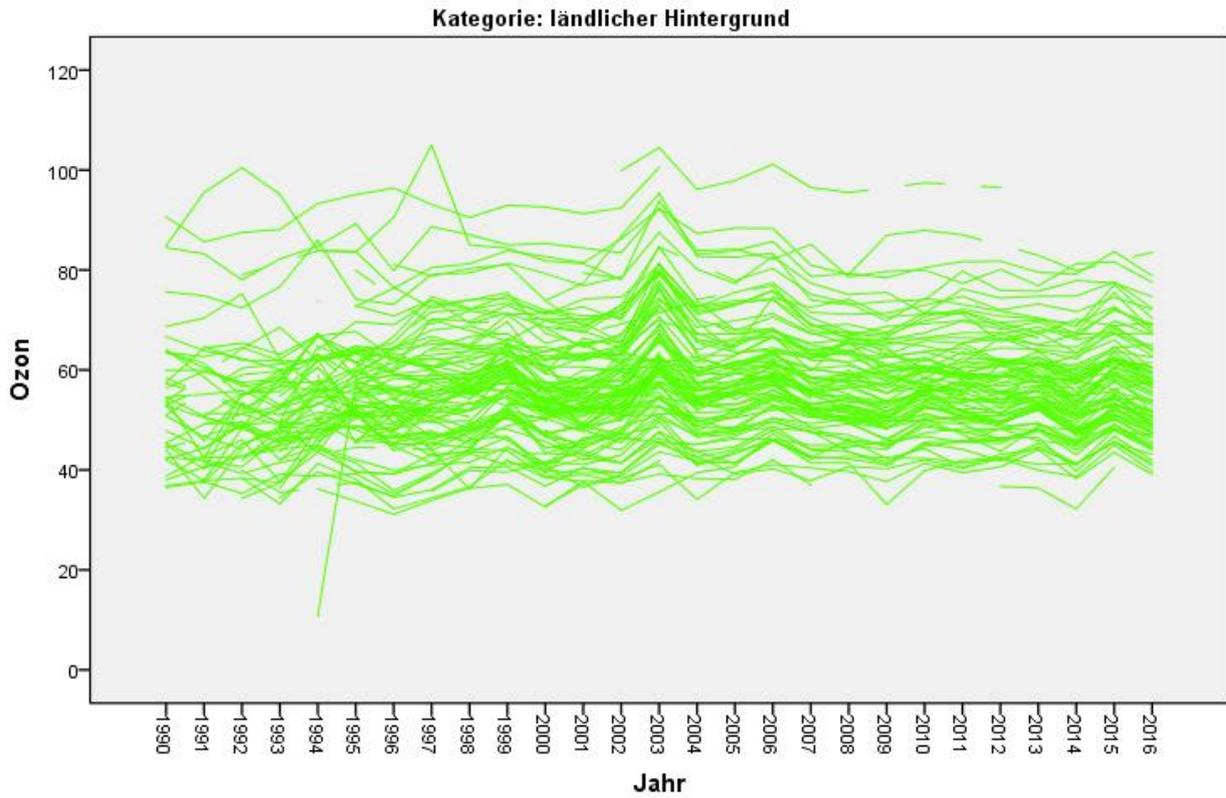
3.1.1.9 Ozonwerte in $\mu\text{g}/\text{m}^3$ - alle Messreihen und durchgängig messende Messreihen von 1990 - 2016



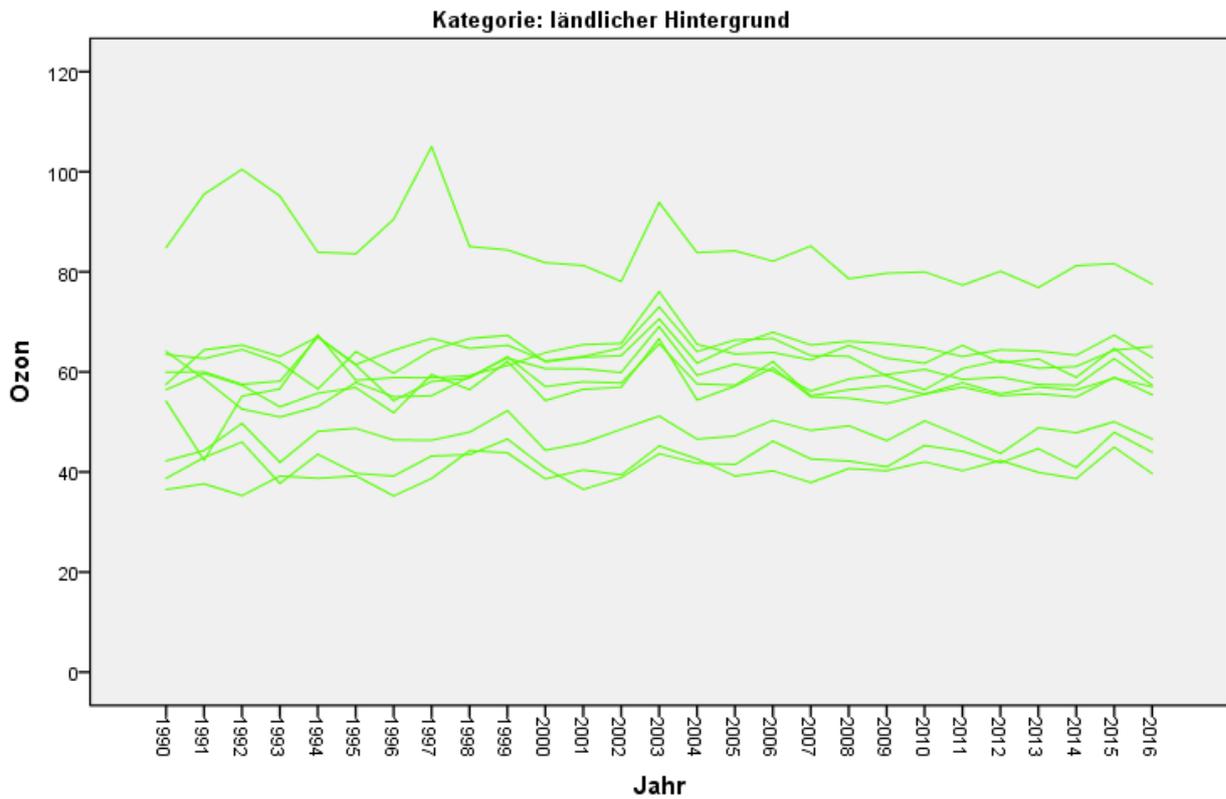
alle Messreihen
eigene Darstellung, BioMath GmbH



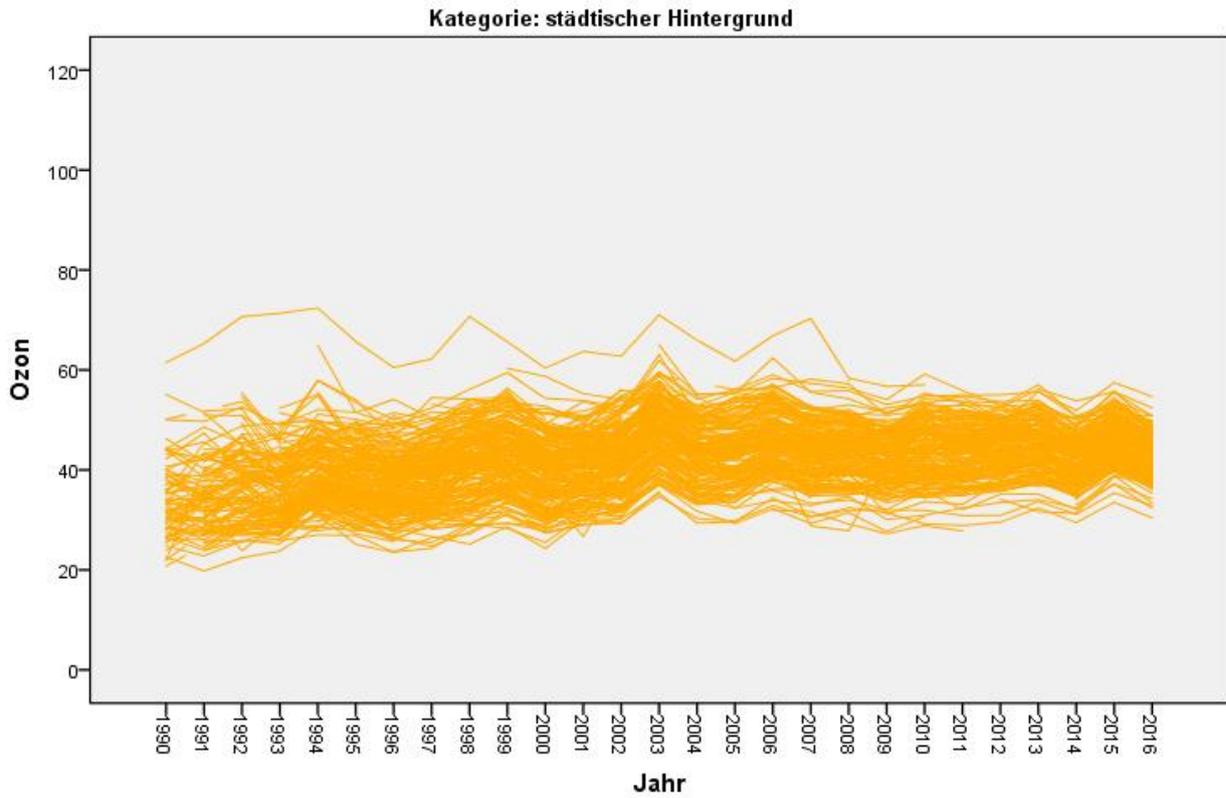
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



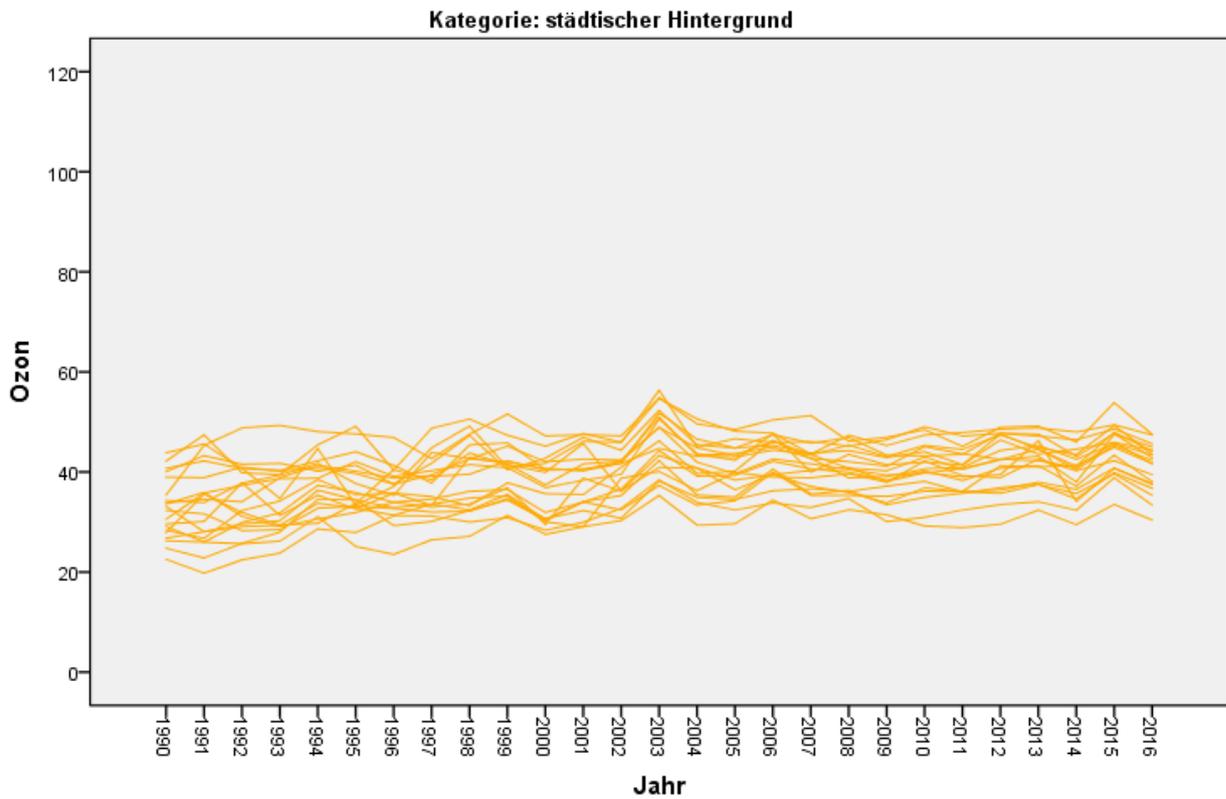
alle Messreihen
eigene Darstellung, BioMath GmbH



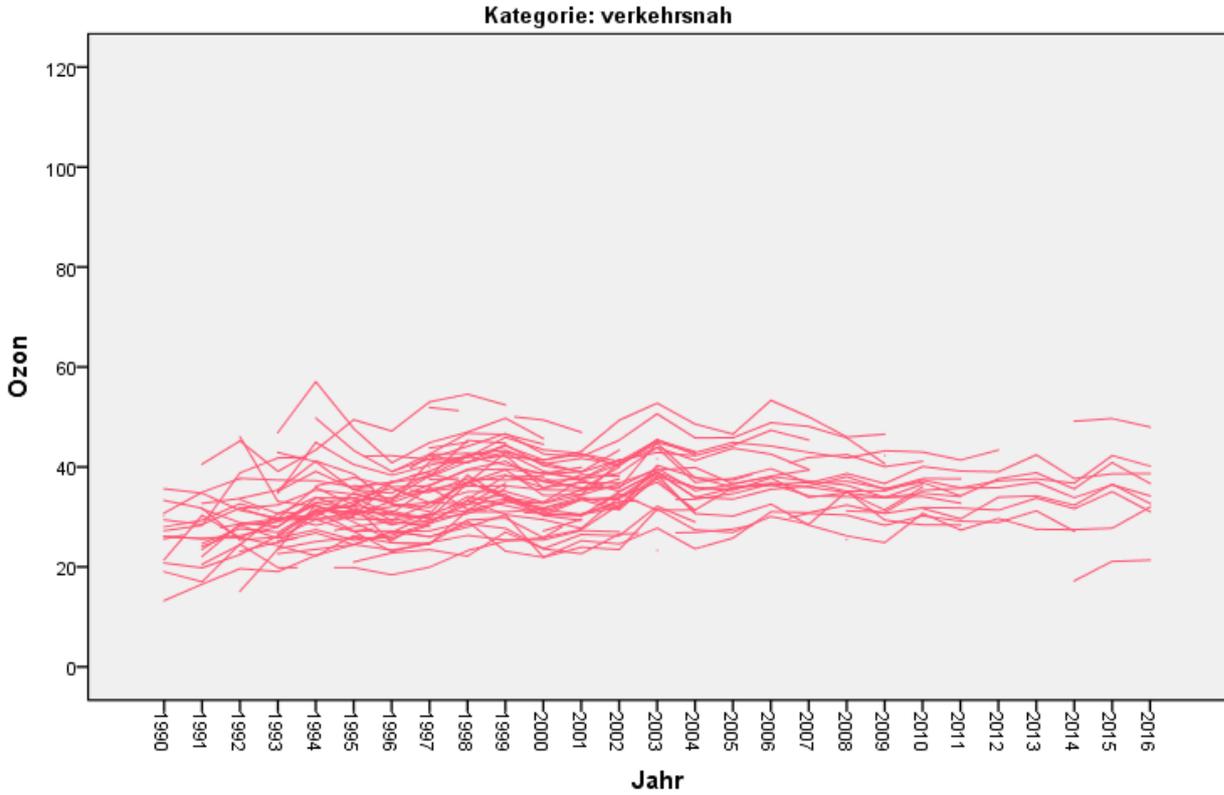
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



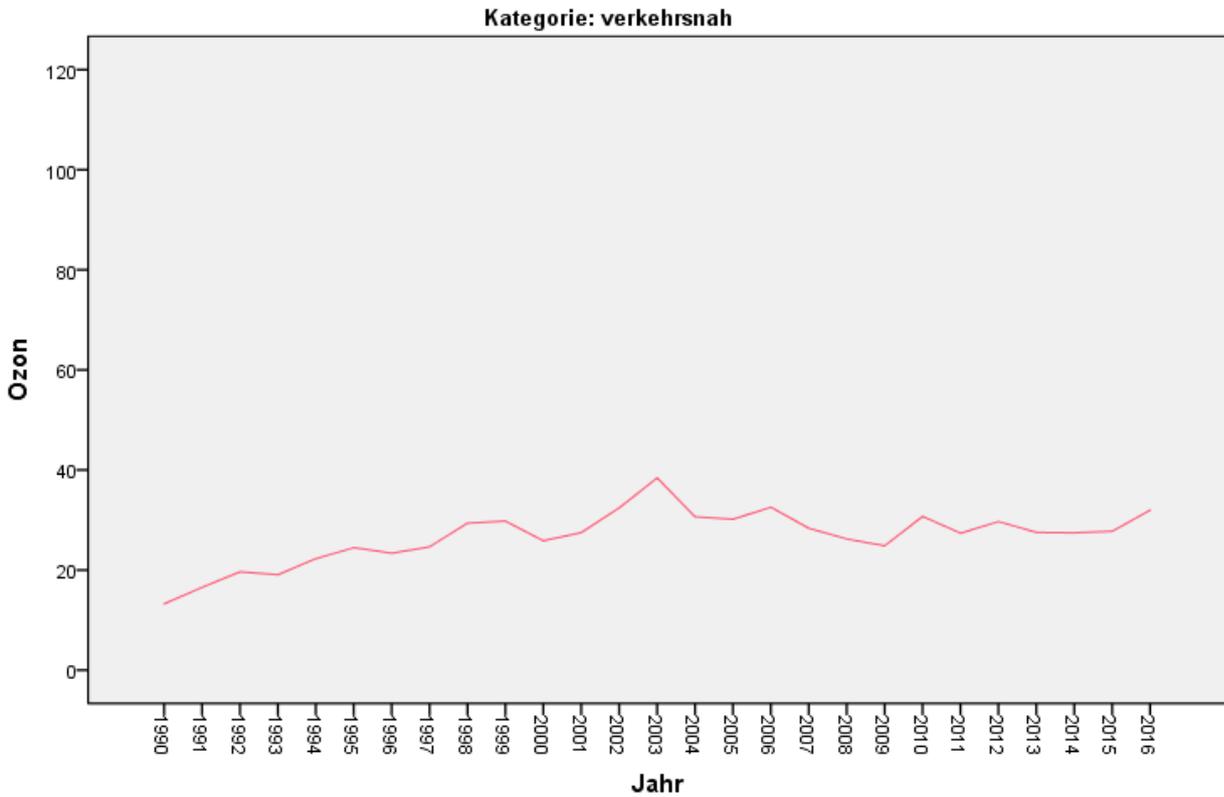
alle Messreihen
eigene Darstellung, BioMath GmbH



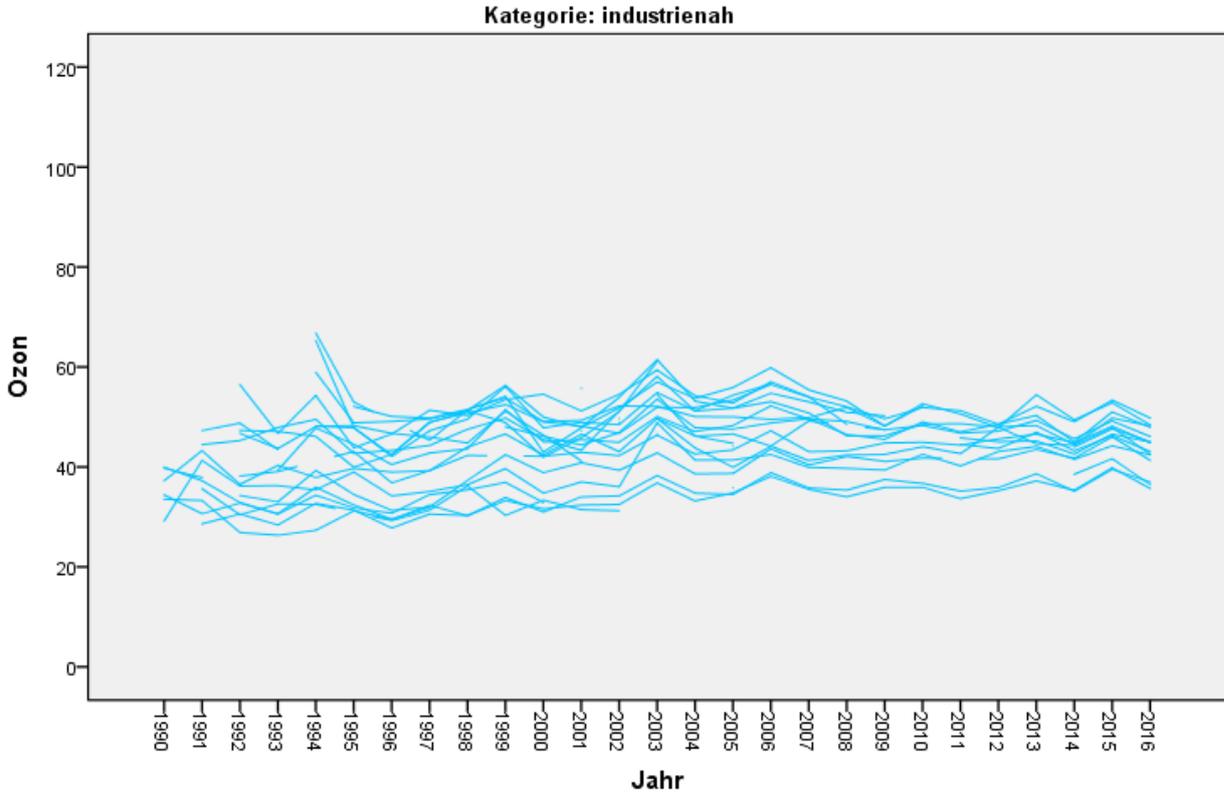
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



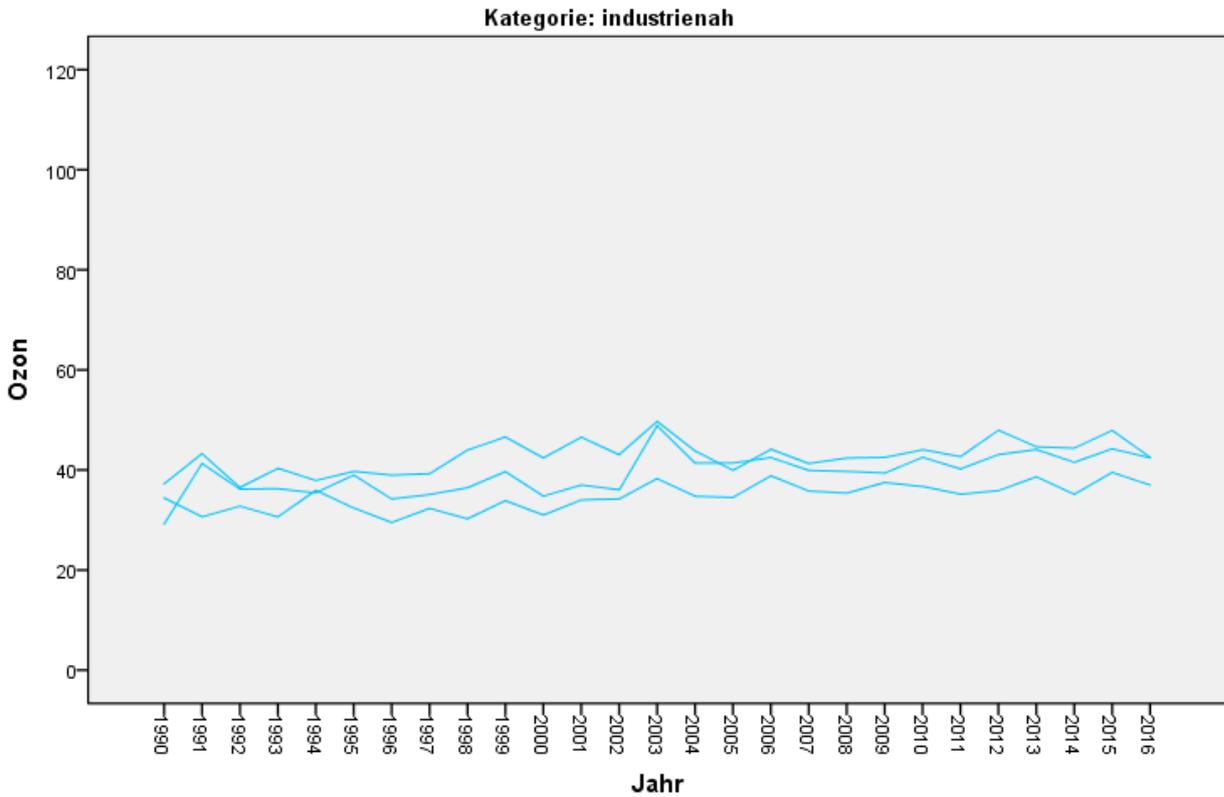
alle Messreihen
eigene Darstellung, BioMath GmbH



durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

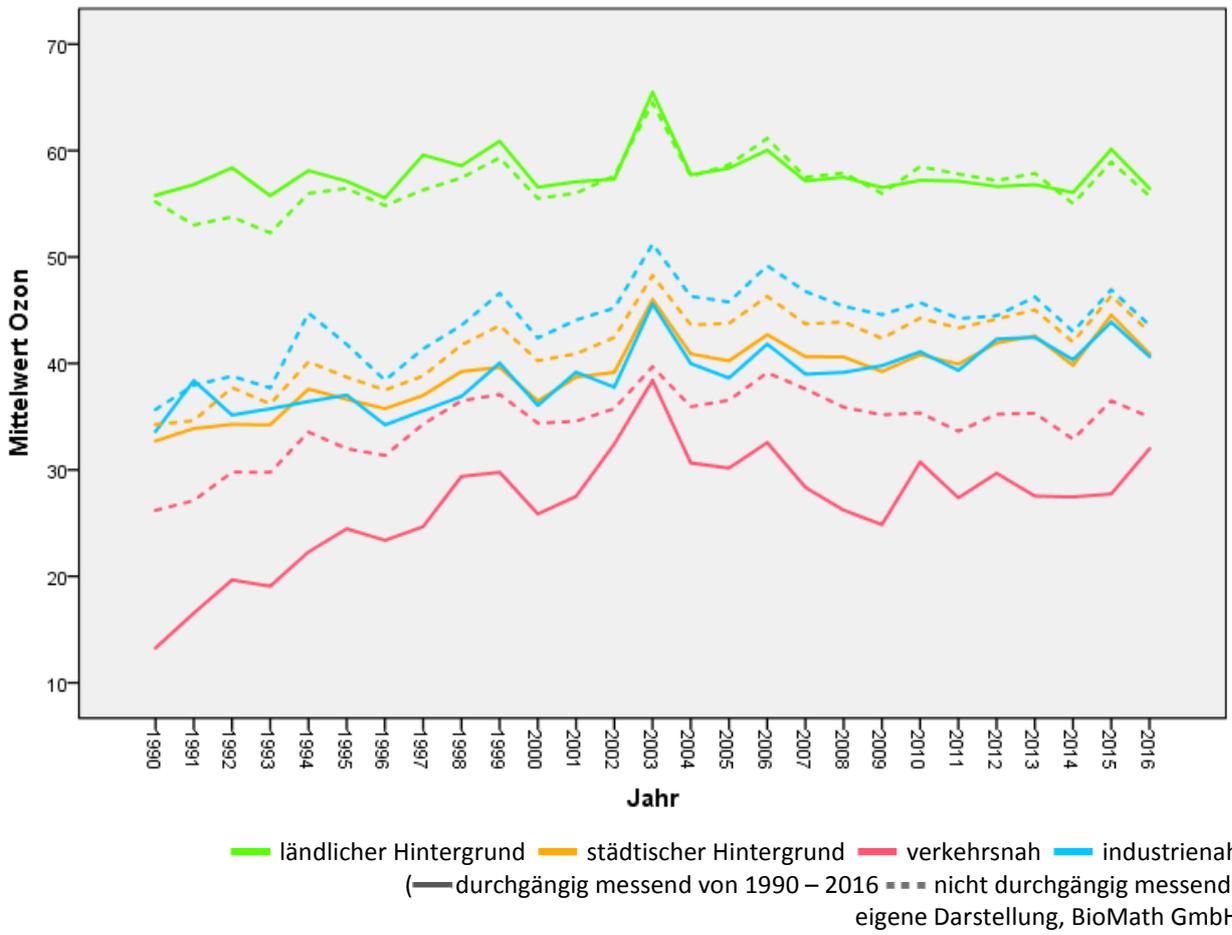


alle Messreihen
eigene Darstellung, BioMath GmbH

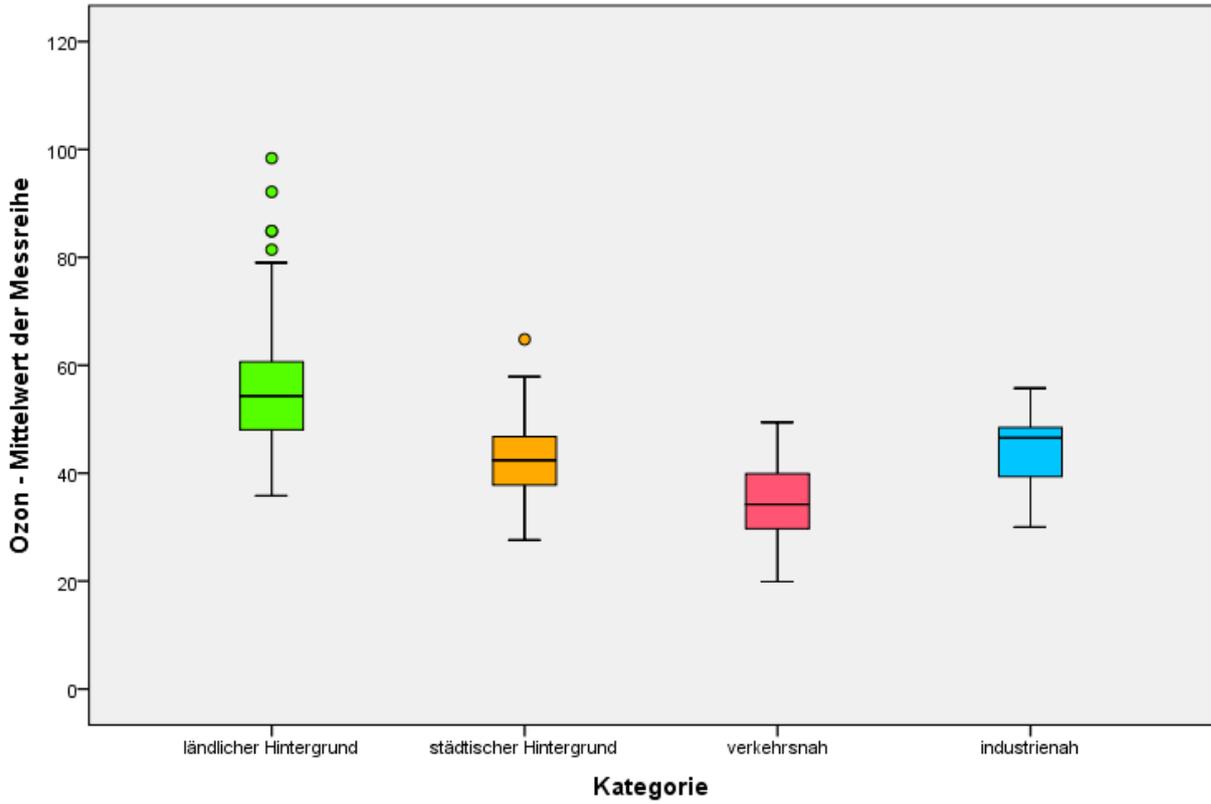


durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

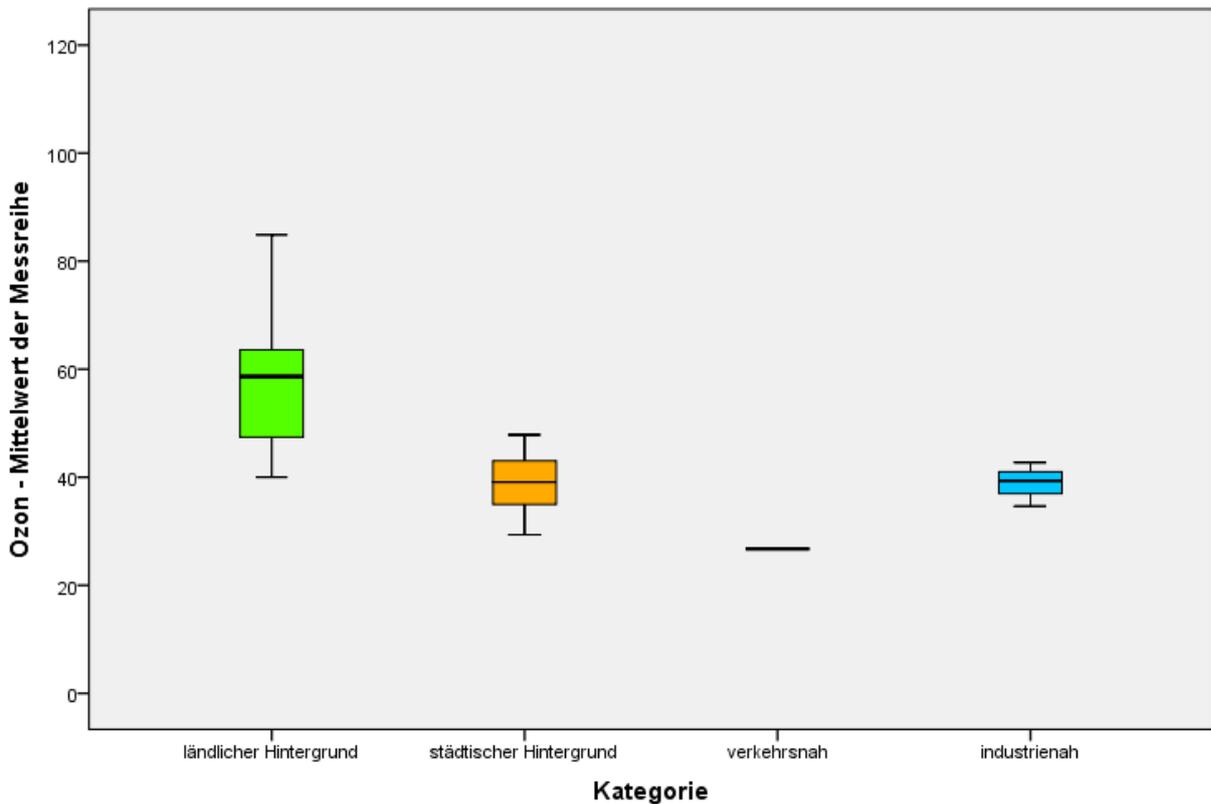
3.1.1.10 Ozon - Messreihen der jährlichen Mittelwerte in $\mu\text{g}/\text{m}^3$



3.1.1.11 Ozon - Boxplot der Mittelwerte ($\mu\text{g}/\text{m}^3$) je Messreihe je Kategorie

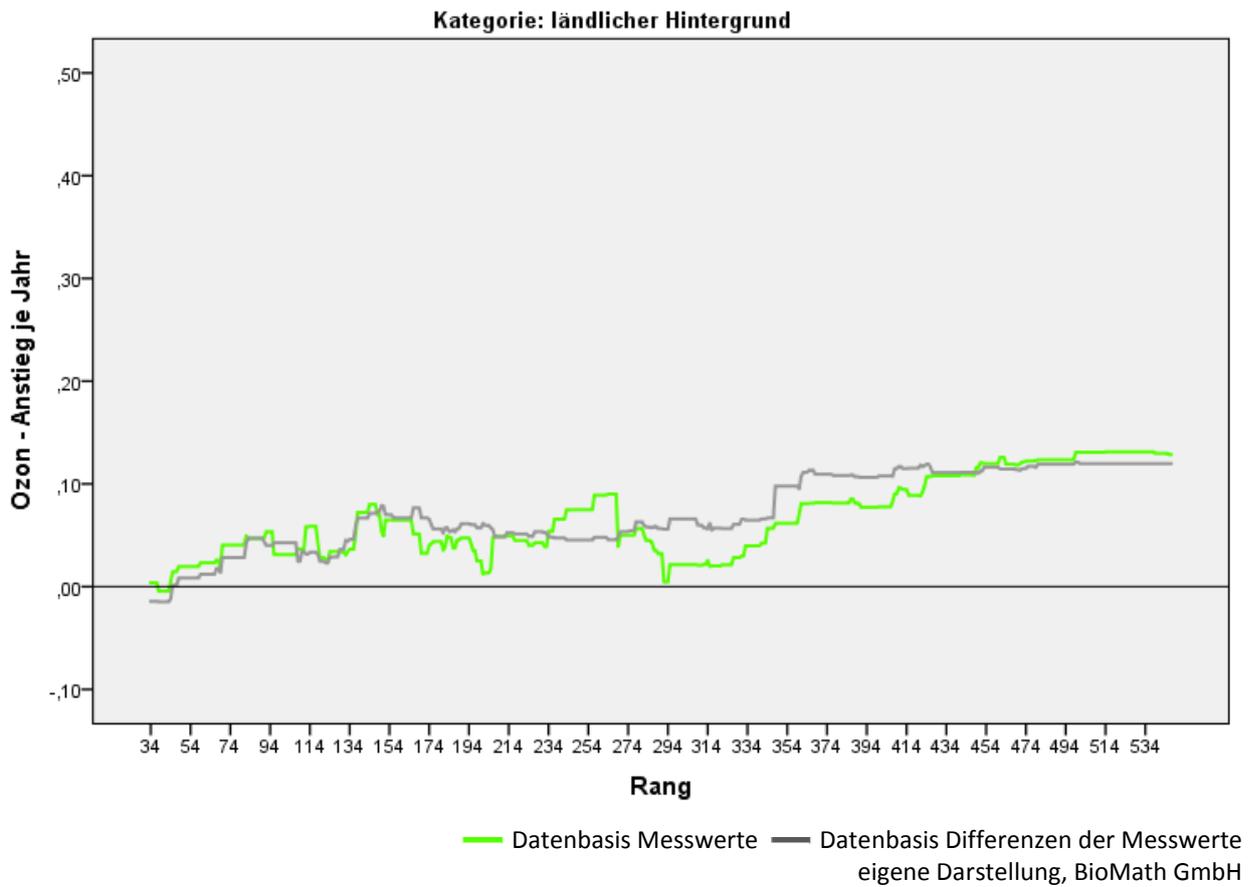


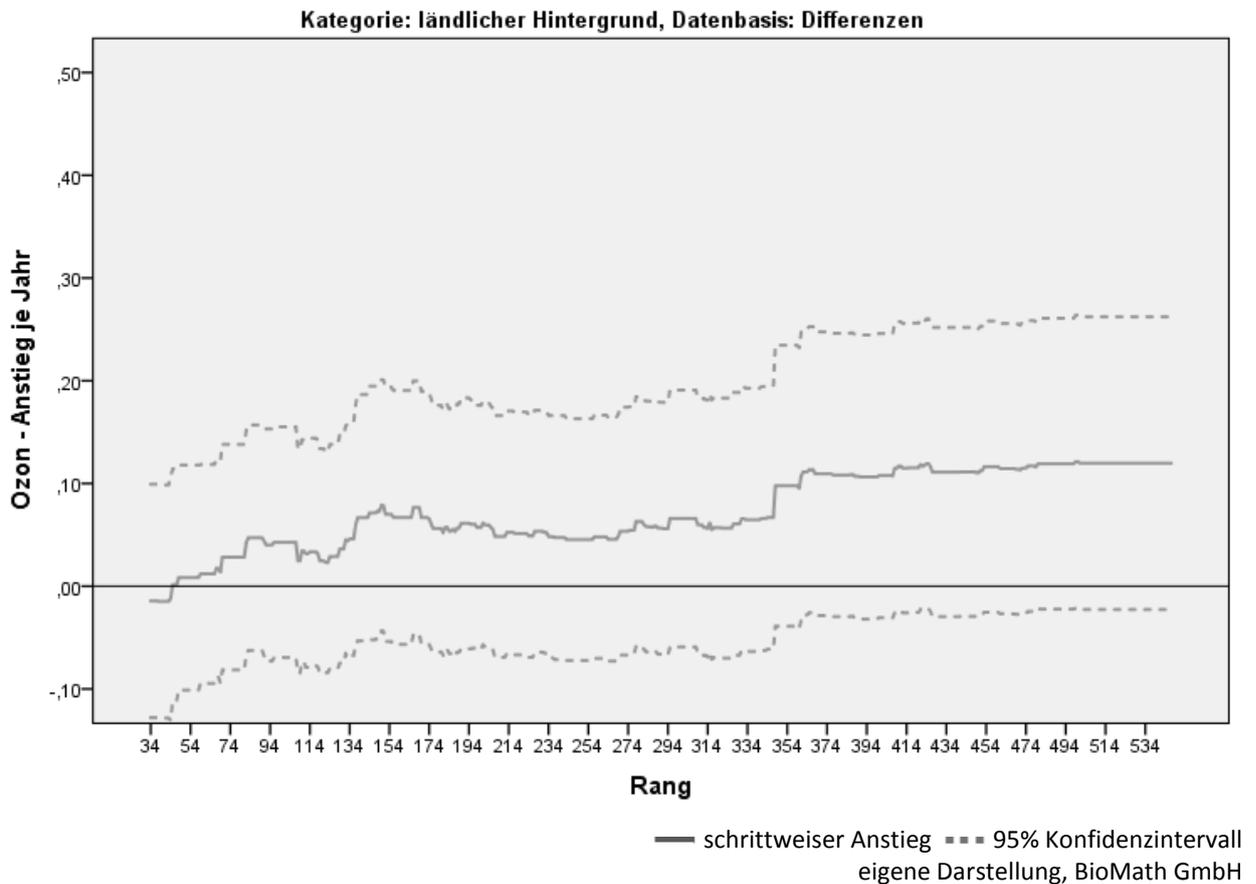
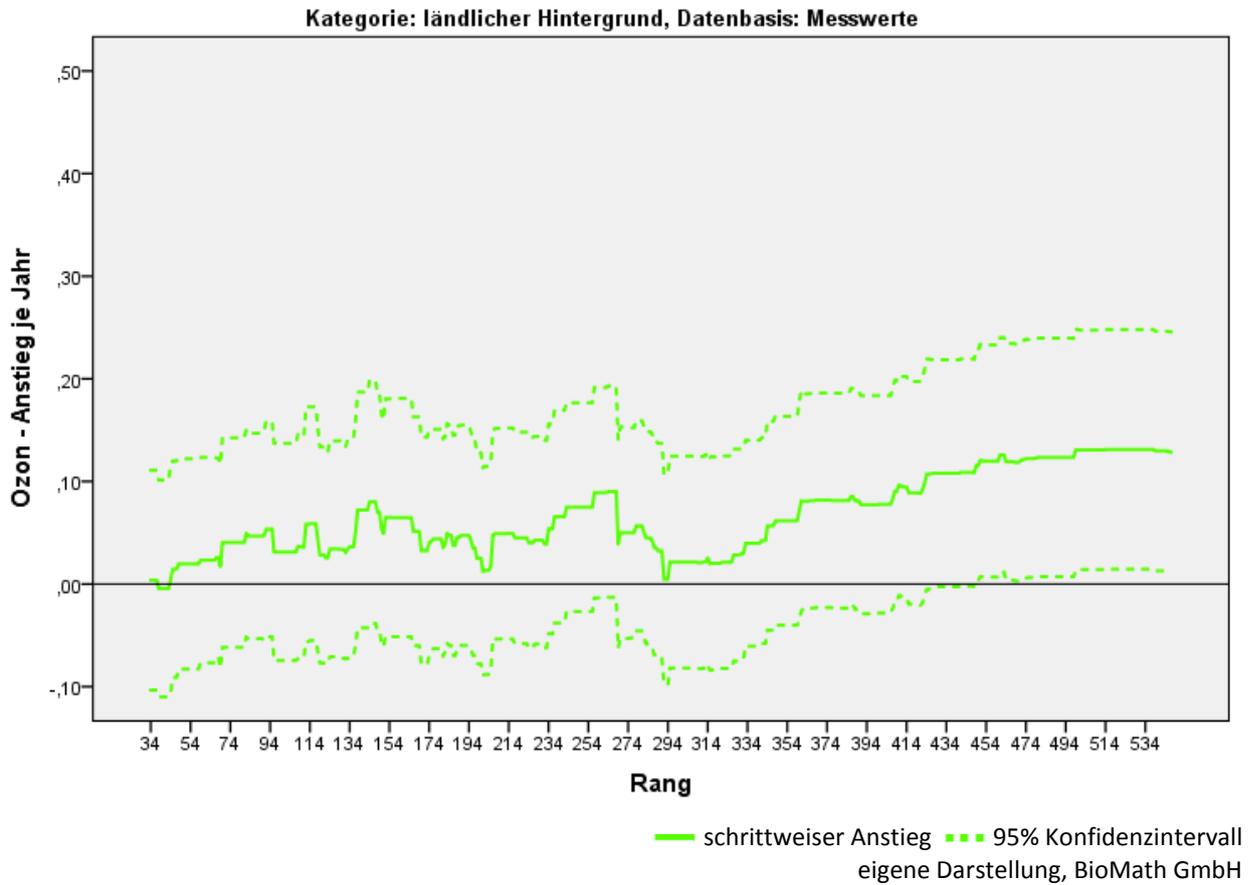
alle Messreihen
eigene Darstellung, BioMath GmbH



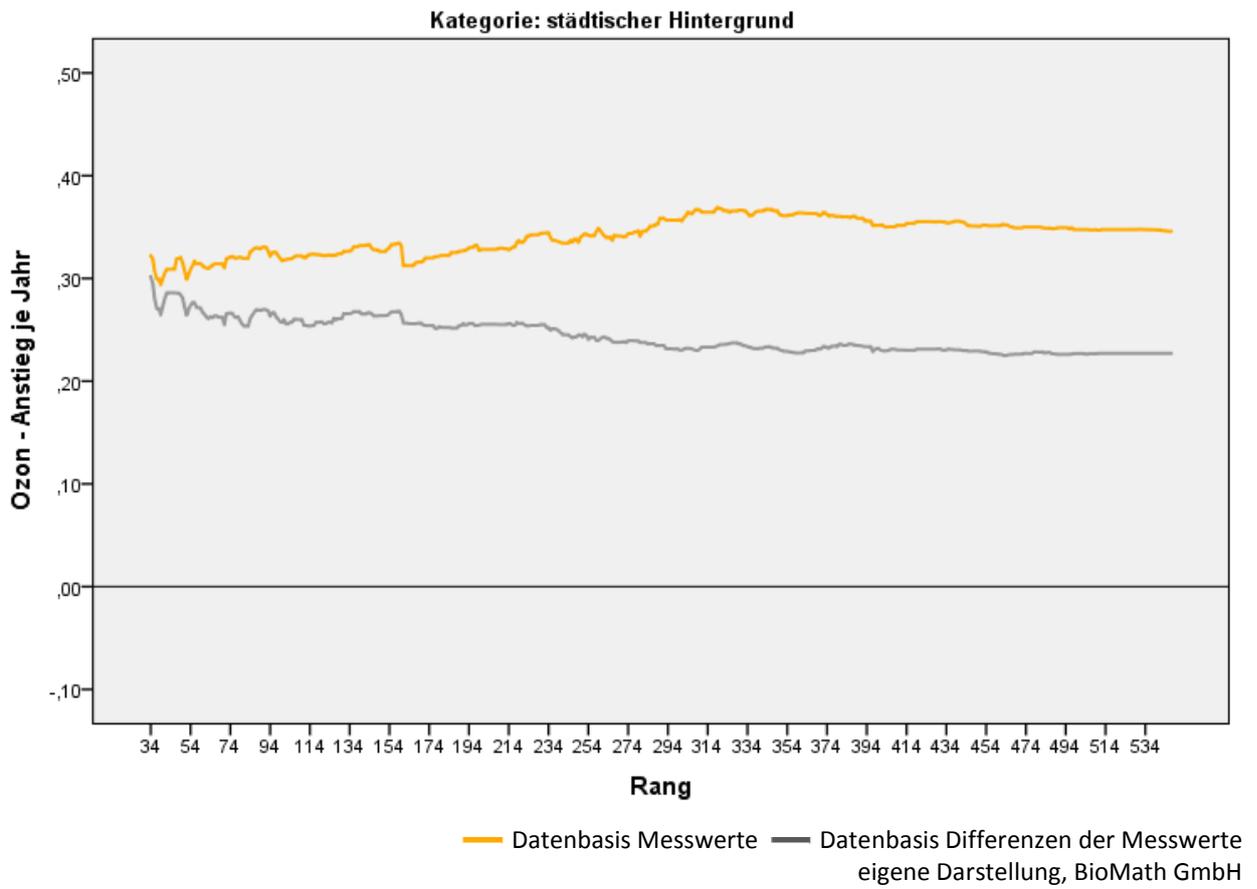
durchgängig messende Messreihen von 1990 - 2016
eigene Darstellung, BioMath GmbH

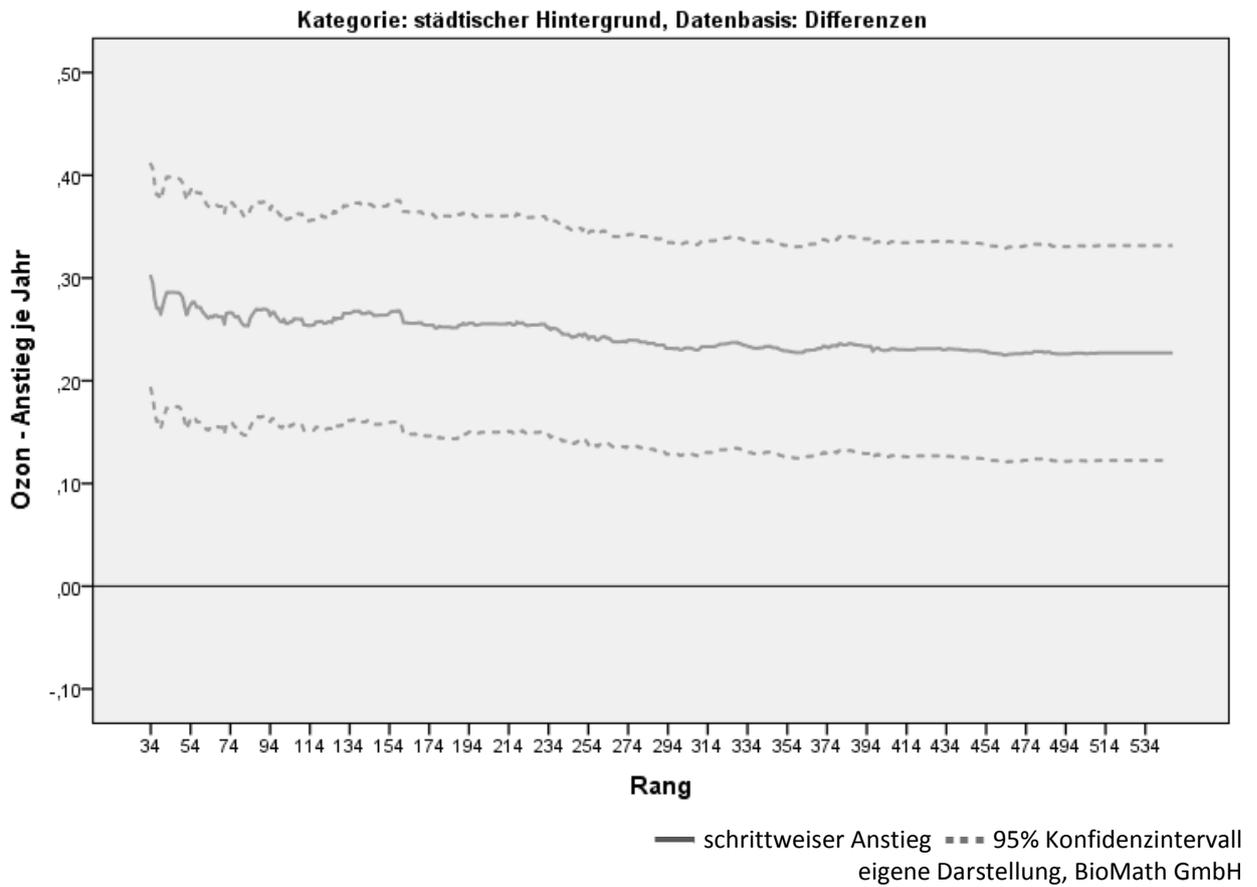
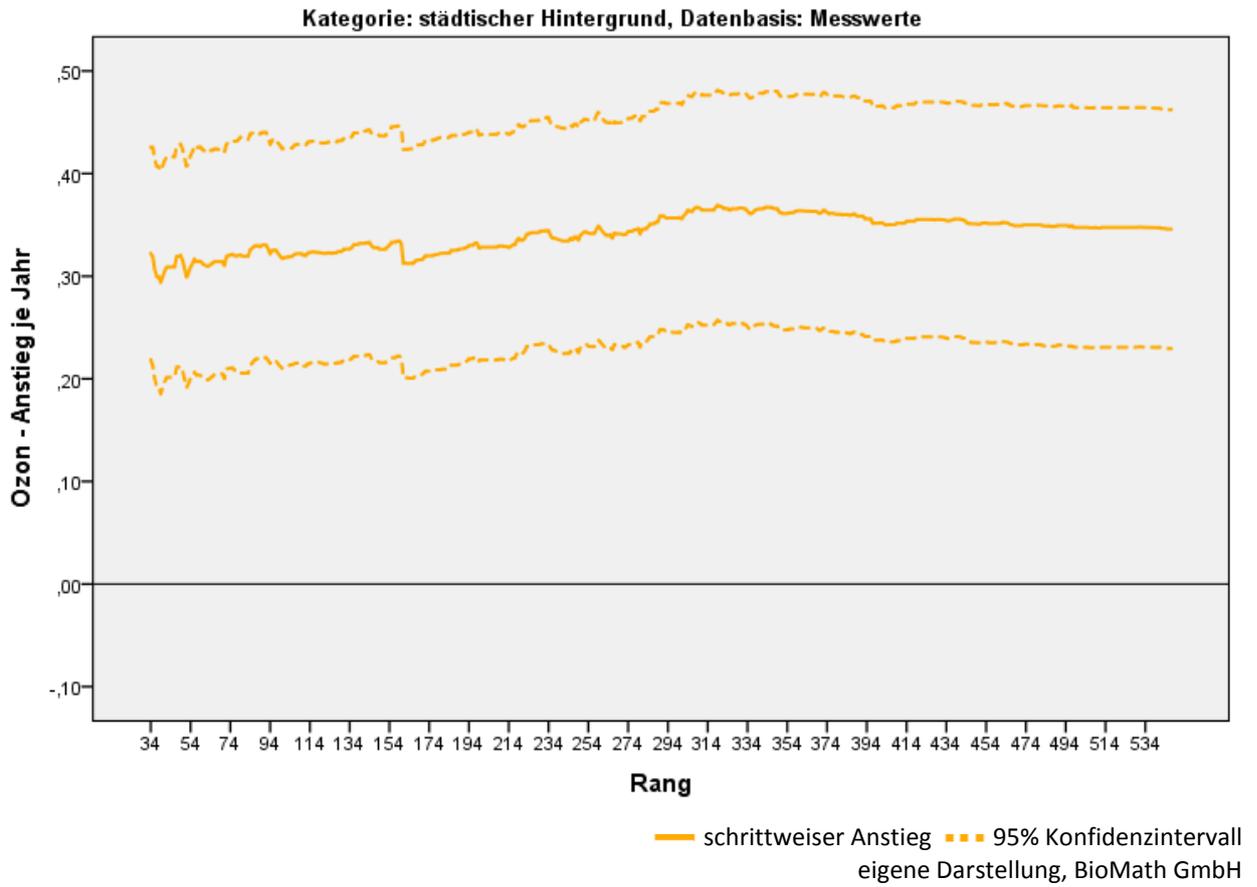
3.1.1.12 Ozon - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund



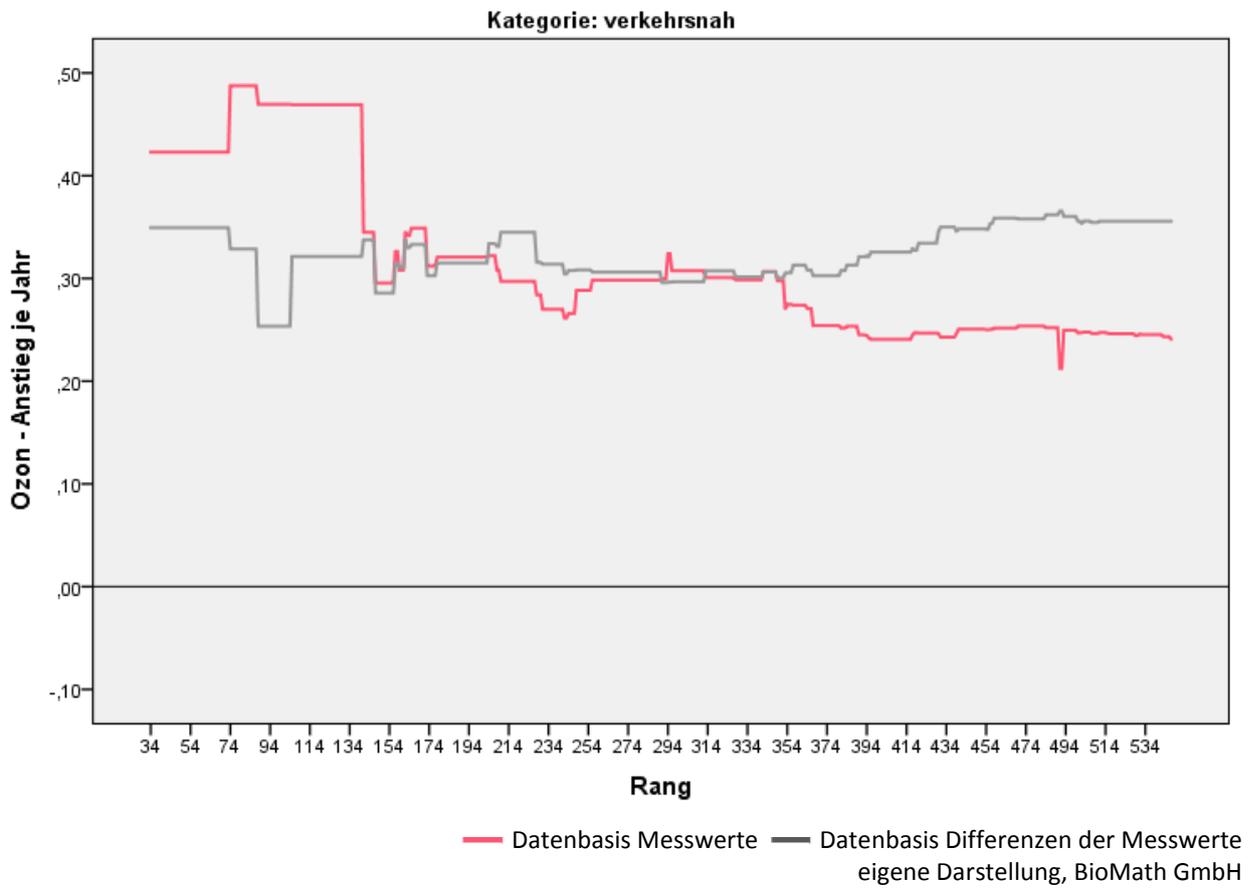


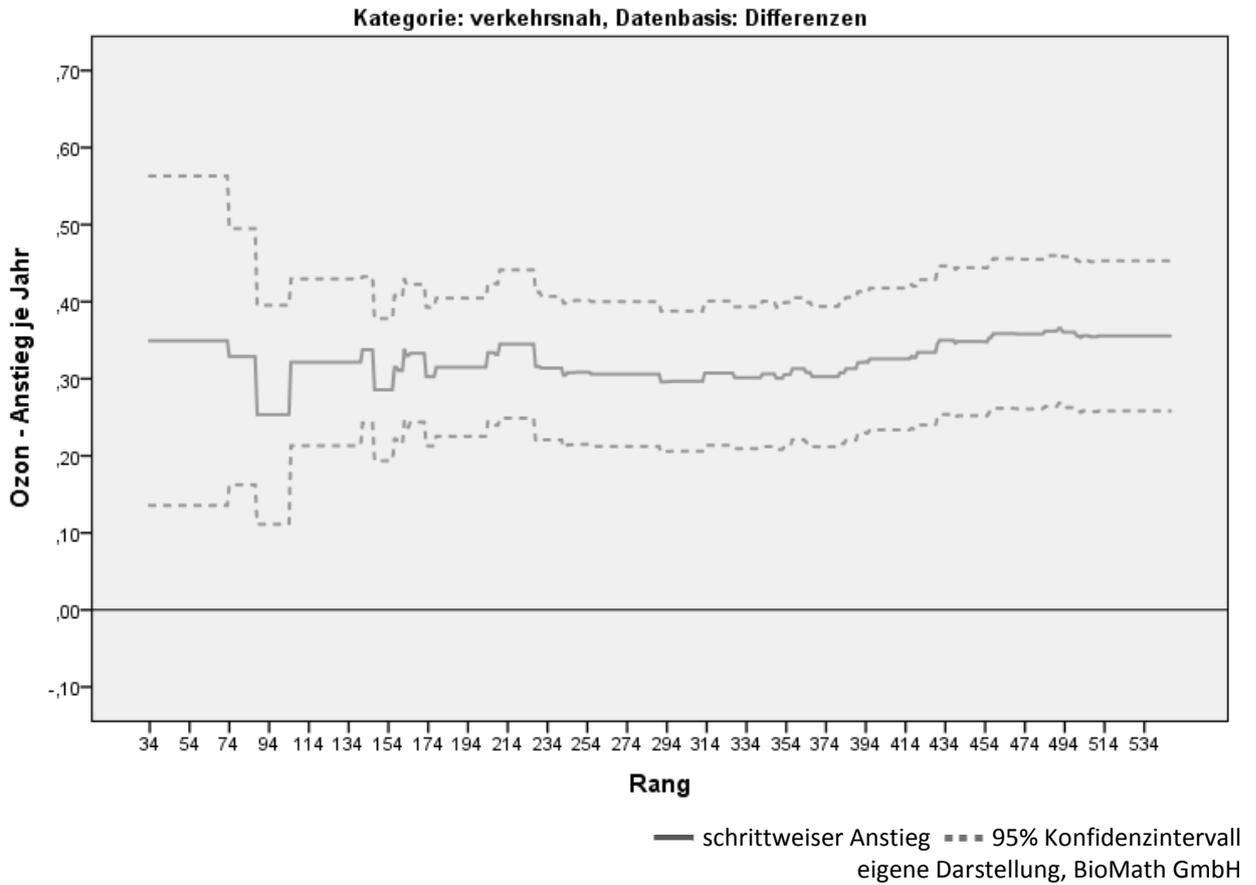
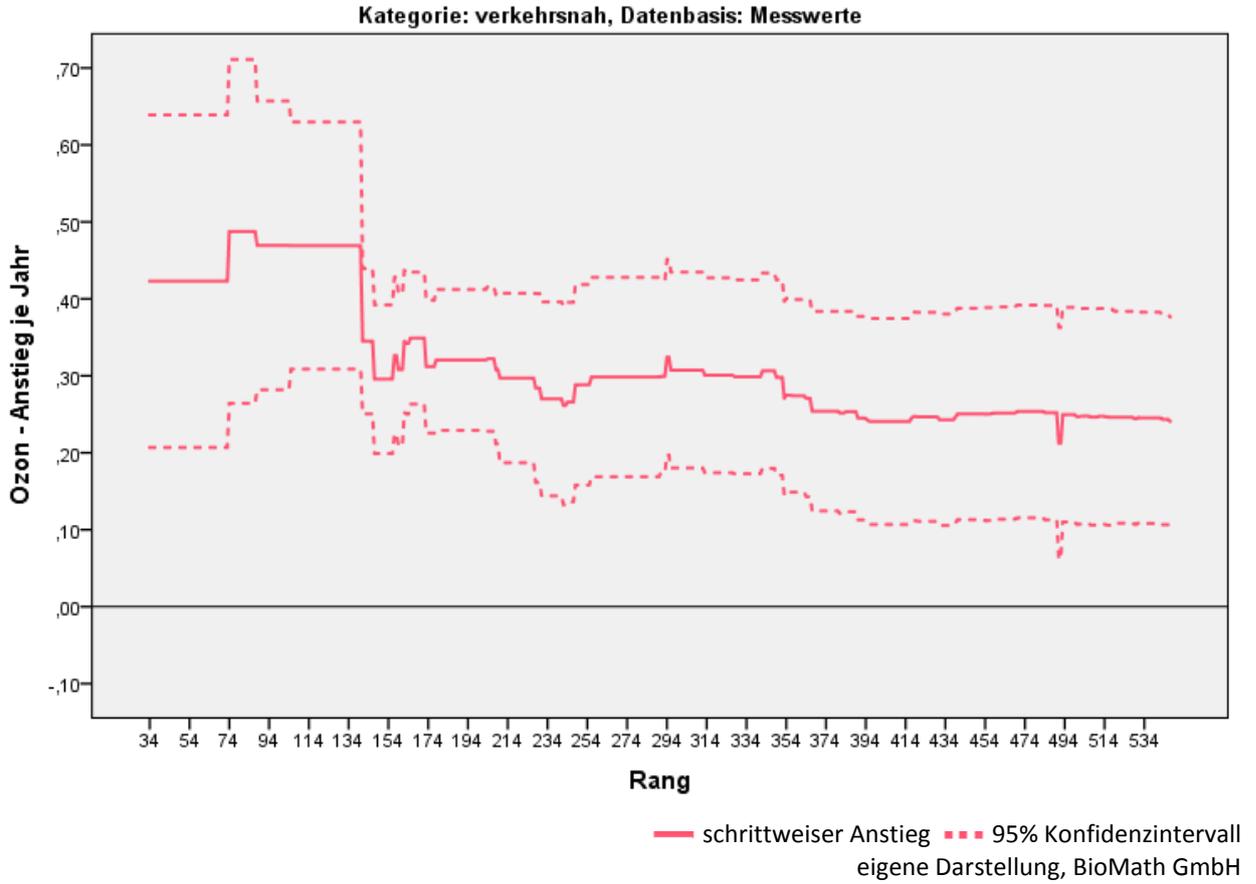
3.1.1.13 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund



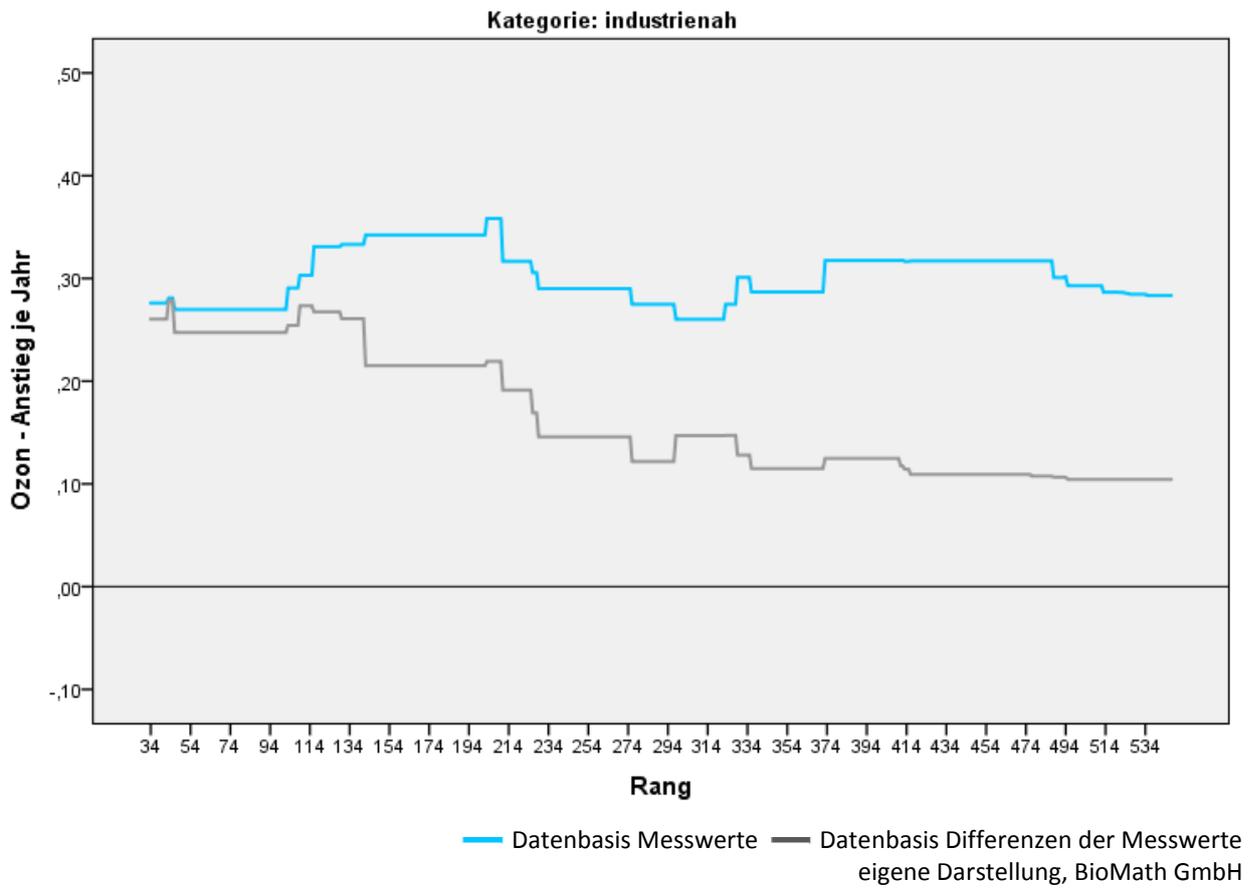


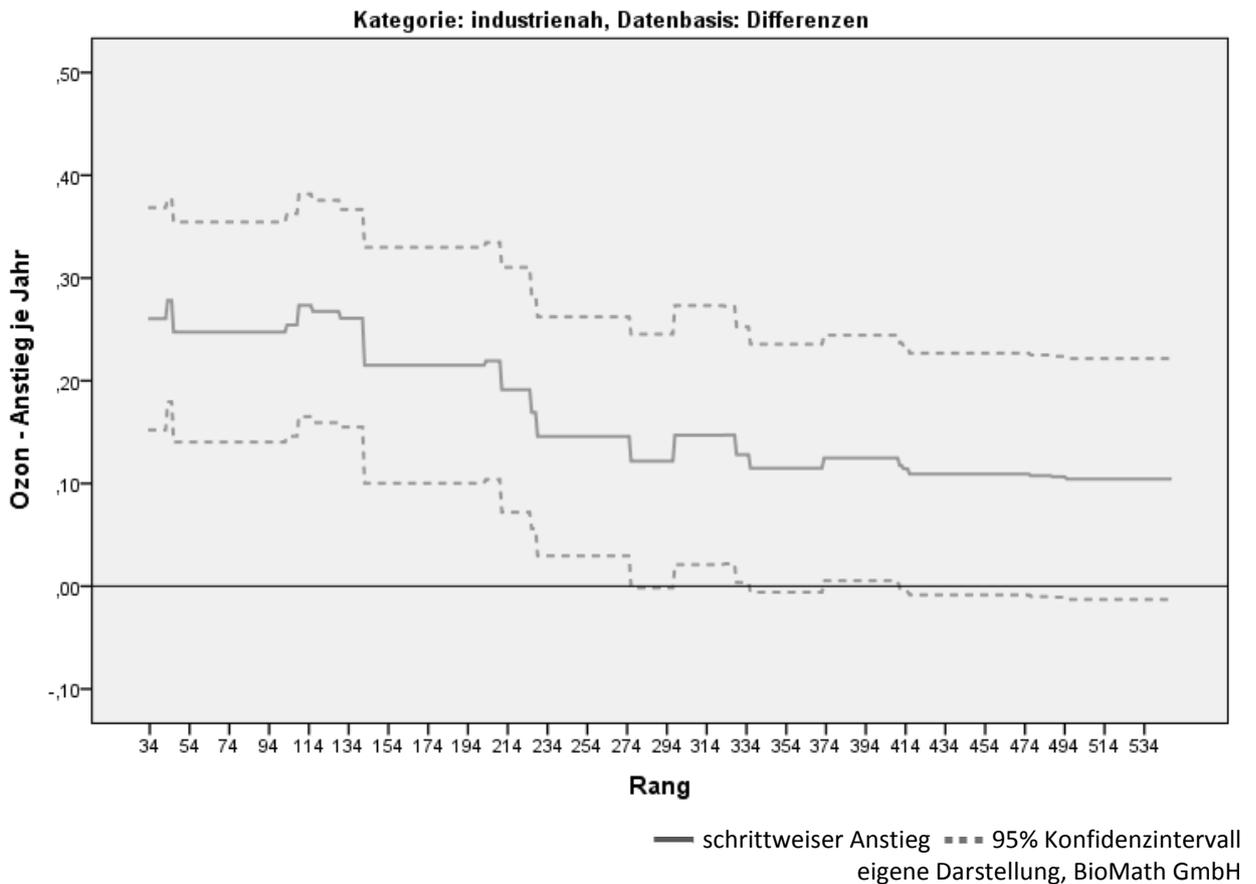
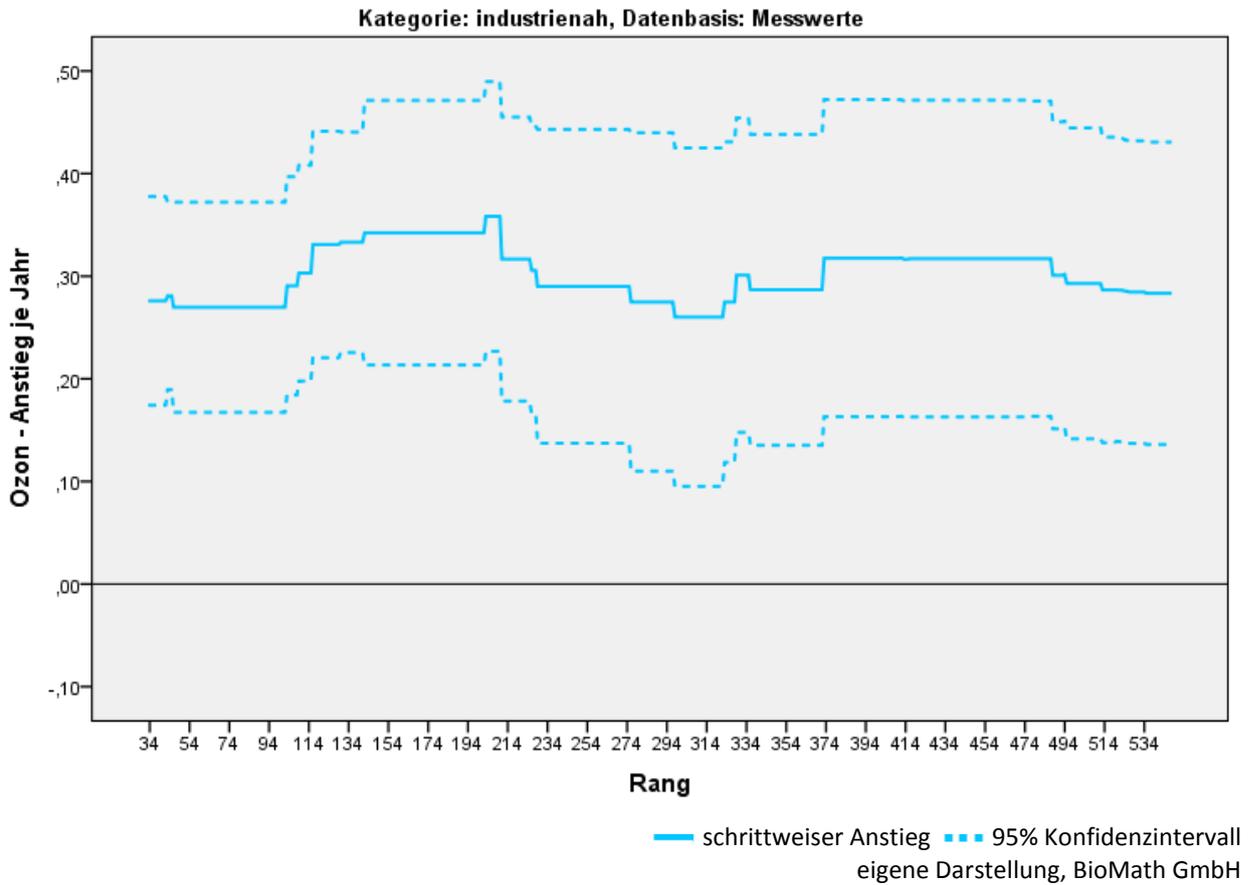
3.1.1.14 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah



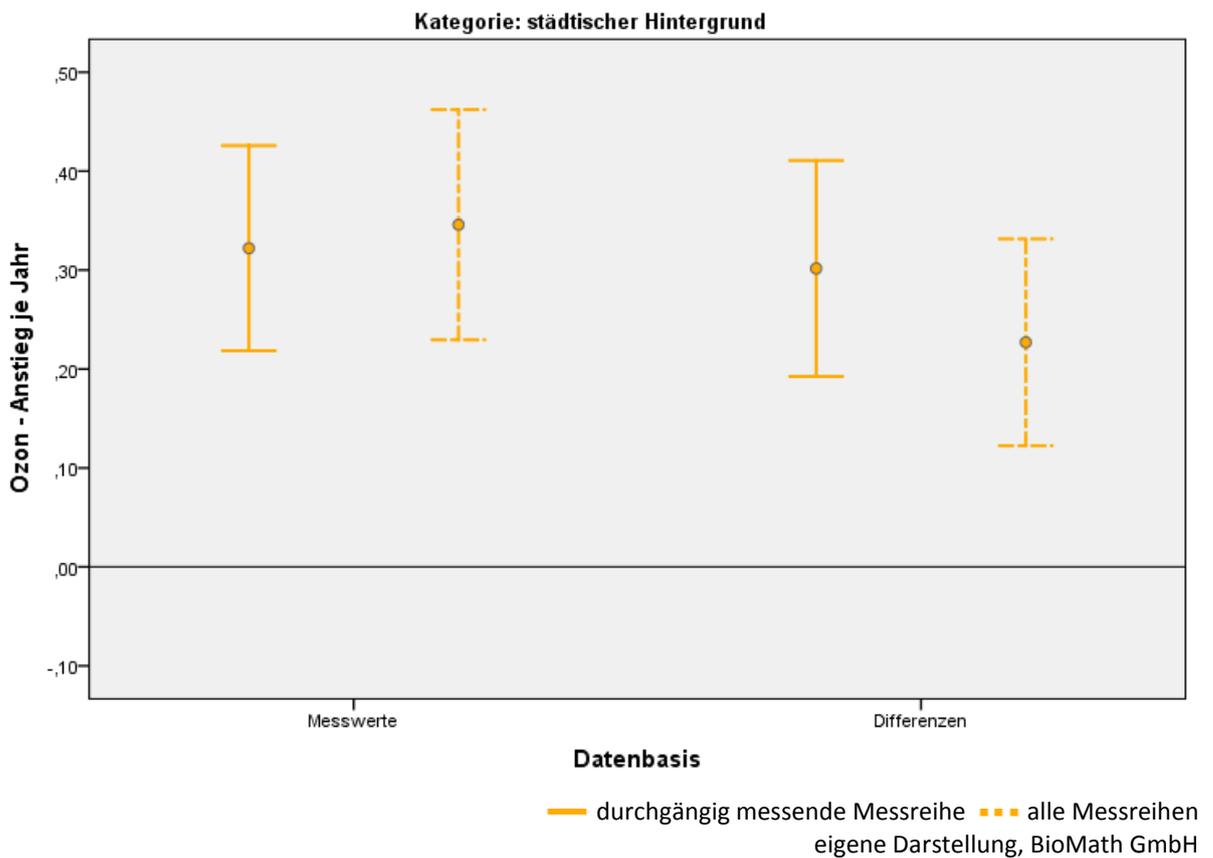
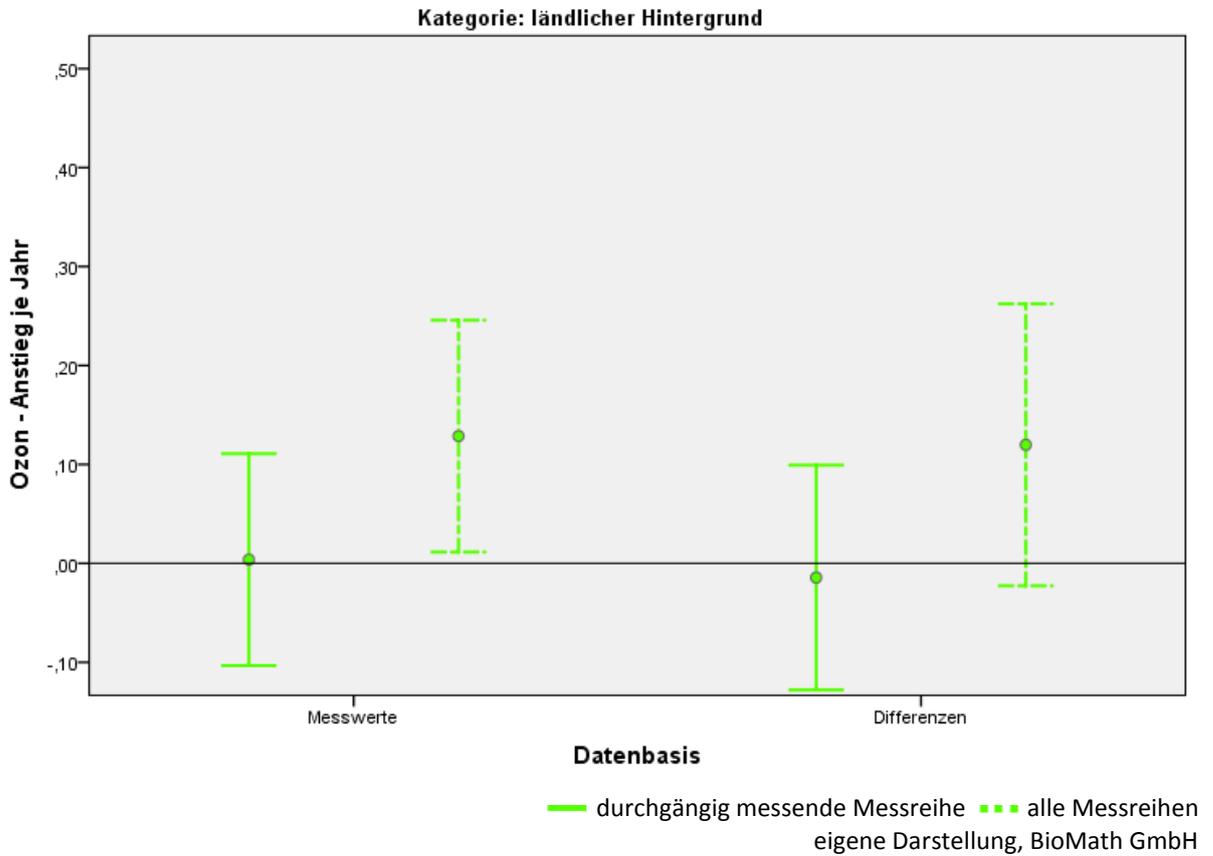


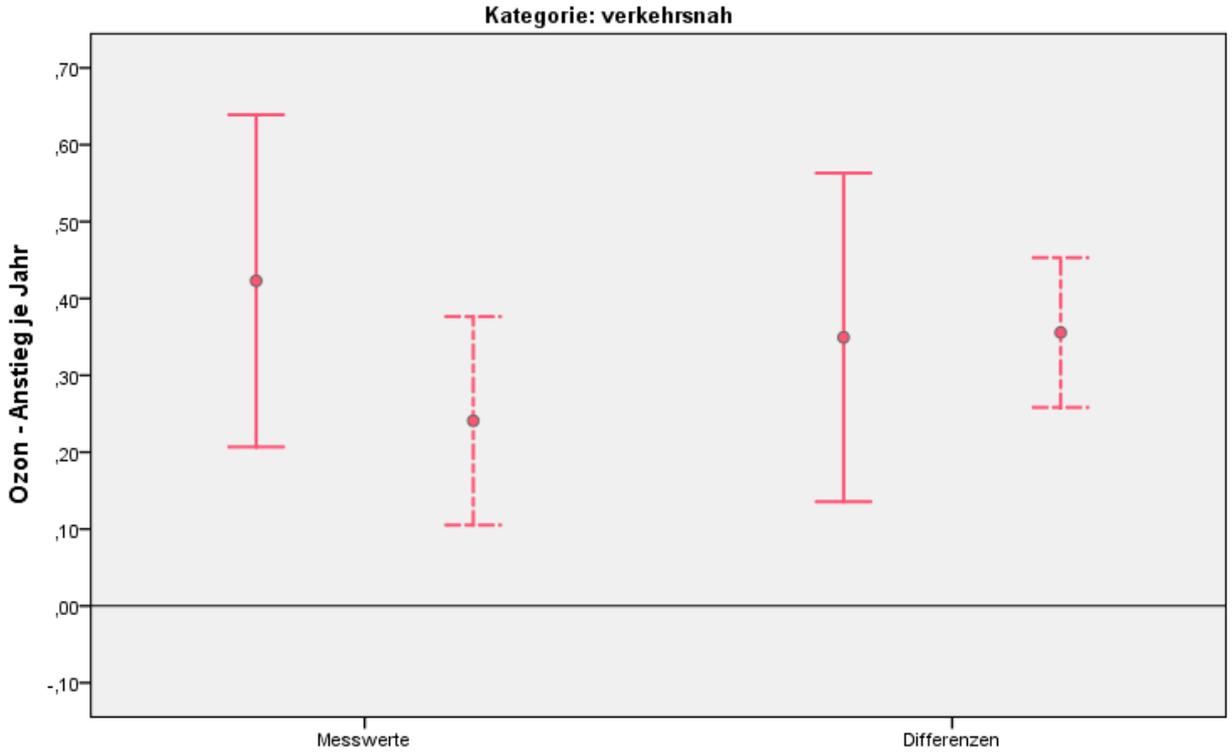
3.1.1.15 Ozon - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah





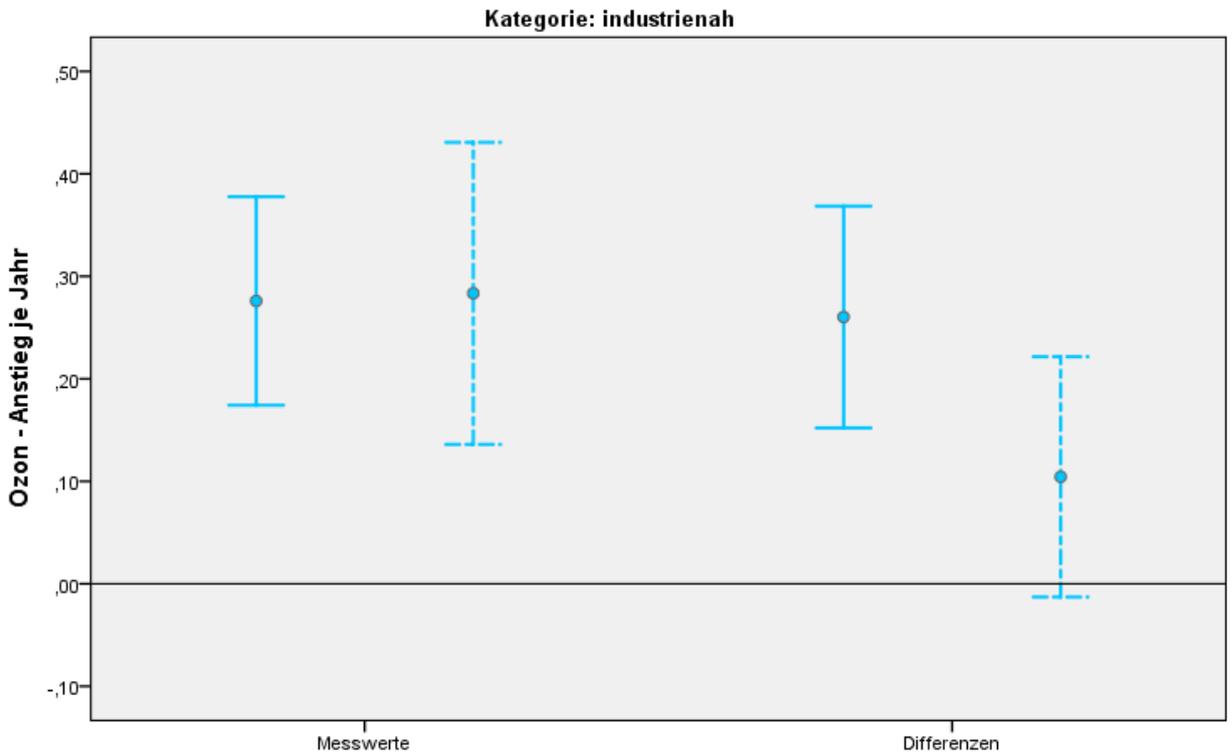
3.1.1.16 Ozon - Trend: Anstieg der Messreihen ($\mu\text{g}/\text{m}^3$ je Jahr) durchgängig vs. nicht durchgängig messend





Datenbasis

— durchgängig messende Messreihe - - - alle Messreihen eigene Darstellung, BioMath GmbH



Datenbasis

— durchgängig messende Messreihe - - - alle Messreihen eigene Darstellung, BioMath GmbH

3.1.2 NO₂

Anzahl und Verteilung der Stationen

Der überwiegende Teil der Messstationen gehört mit 45,8 % zur Kategorie verkehrsnah, gefolgt vom städtischen Hintergrund mit 36,1 %. Die Kategorien ländlicher Hintergrund mit 11,9 % und industrienahe mit 6,2 % sind deutlich geringer vertreten. Betrachtet man nur die durchgängig messenden Stationen von 1990 - 2016 so sind es gerade mal 75 von 1112 Stationen, d.h. 7,5 % (siehe Abbildung 2 und 3.1.2.2).

Von 1990 bis 1997 steigt die Anzahl der Stationen stark an, hält sich dann 5 Jahre lang etwa auf diesem Niveau, fällt danach leicht ab und steigt erneut an. Im Zeitraum von 2009 bis 2016 ist die Anzahl der Stationen nur um 10 gestiegen. Dieser Verlauf ist aber in den einzelnen Kategorien so nicht zu verzeichnen. Ab dem Jahr 2005 steigt die Anzahl der Stationen in der Kategorie verkehrsnah sehr stark an, wobei gleichzeitig die Anzahlen der Stationen der Kategorien ländlicher und städtischer Hintergrund fallen. (siehe 3.1.2.1).

Insgesamt zeigen die jeweils durchgängig messenden Stationen vom Jahr i ($i > 1990$) bis 2016 eine kontinuierliche Zunahme. Dieses widerspiegelt sich insbesondere in den Kategorien städtischer und ländlicher Hintergrund. Die Kategorie verkehrsnah zeigt eine deutlich stärkere Zunahme der Stationen ab 2005. (siehe 3.1.2.2 und 3.1.2.3).

Im Jahr 1990 gab es nur Stationen im westlichen Teil Deutschlands und West-Berlin. Bis zum Jahre 1995 nahm die Anzahl der Stationen im östlichen Teil Deutschland deutlich zu. Das erklärt den schon genannten starken Anstieg der Anzahl von Stationen von 1990 bis 1997. Nur in Sachsen-Anhalt konnten erst ab 1997 durchgängig messende Reihe generiert werden. Ab 2000 gibt es in gesamt Deutschland eine flächendeckende Anzahl von durchgängig messenden Stationen (siehe 3.1.2.4 bis 3.1.2.7).

Inbetriebnahme und Außerbetriebnahme von Stationen sind heterogen über den Zeitabschnitt 1990 - 2016 verteilt. Das betrifft alle Kategorien aber besonders die Kategorie verkehrsnah. Die Betriebszeiten schwanken von 1 Jahr bis zu 27 Jahren und sind im Mittel etwa 11 Jahre. Es waren 12,3 % der Stationen lediglich nur 1 Jahr in Betrieb. Etwa 85 % der Stationen lieferten in jedem Betriebsjahr einen Wert. Bei knapp 9 % der Stationen fehlte ein Wert während der Betriebszeit. Würde man diesen einen fehlenden Wert auffüllen, so stellen nicht ganz 90 % der Stationen vollständige Messreihen in ihren Betriebszeiten bereit (siehe 3.1.2.8).

Verlauf der Werte der Messreihen

Die Werte der Messreihen betragen im Mittel $30,6 \mu\text{g}/\text{m}^3$ und schwanken zwischen $1,8$ und $121,3 \mu\text{g}/\text{m}^3$. Alle Messreihen der Kategorien ländlicher Hintergrund, städtischer Hintergrund und industrienahe sind in sich einheitlich. Die Messreihen der Kategorie verkehrsnah sind sehr ungleichmäßig, insbesondere gibt es ab 2007 Messreihen die bei über $80 \mu\text{g}/\text{m}^3$ beginnen. (siehe 3.1.2.9). Messreihen mit hohen Abweichungen zum mittleren Niveau sind im Boxplot gut erkennbar - 7 Reihen ländlicher Hintergrund, 9 Reihen verkehrsnah (siehe 3.1.2.11).

Die mittleren Messreihen der durchgängig und nicht durchgängig messenden Reihen unterscheiden sich in den Kategorien städtischer Hintergrund und industrienahe. Hier sind die Mittelwerte der durchgängig messenden Reihe immer höher als die der nicht durchgängig messenden Reihen. In der Kategorie ländlicher Hintergrund gibt es kaum Abweichungen. In der Kategorie verkehrsnah unterscheiden sich die mittleren Messreihen der durchgängig und nicht durchgängig messenden Reihen deutlich voneinander. Die Abweichungen reichen von knapp 0 bis über $16 \mu\text{g}/\text{m}^3$ (siehe 3.1.2.10).

Trends der Messreihen

In der schrittweisen Trendberechnung gibt es in allen Kategorien Unterschiede im Verlauf der Anstiege zwischen den Datenbasen der Messwerte und denen der Differenzen der Messwerte.

In der Kategorie ländlicher Hintergrund sind die Abweichungen relativ gering. Ab Rang 689 gibt es einen Sprung zwischen den Trends der beiden Datenbasen. Ab diesem Rang kommen Messreihen mit 7 oder weniger Messwerten auf geringen Messniveau hinzu. Alle Anstiege sind negativ und signifikant (siehe 3.1.2.12).

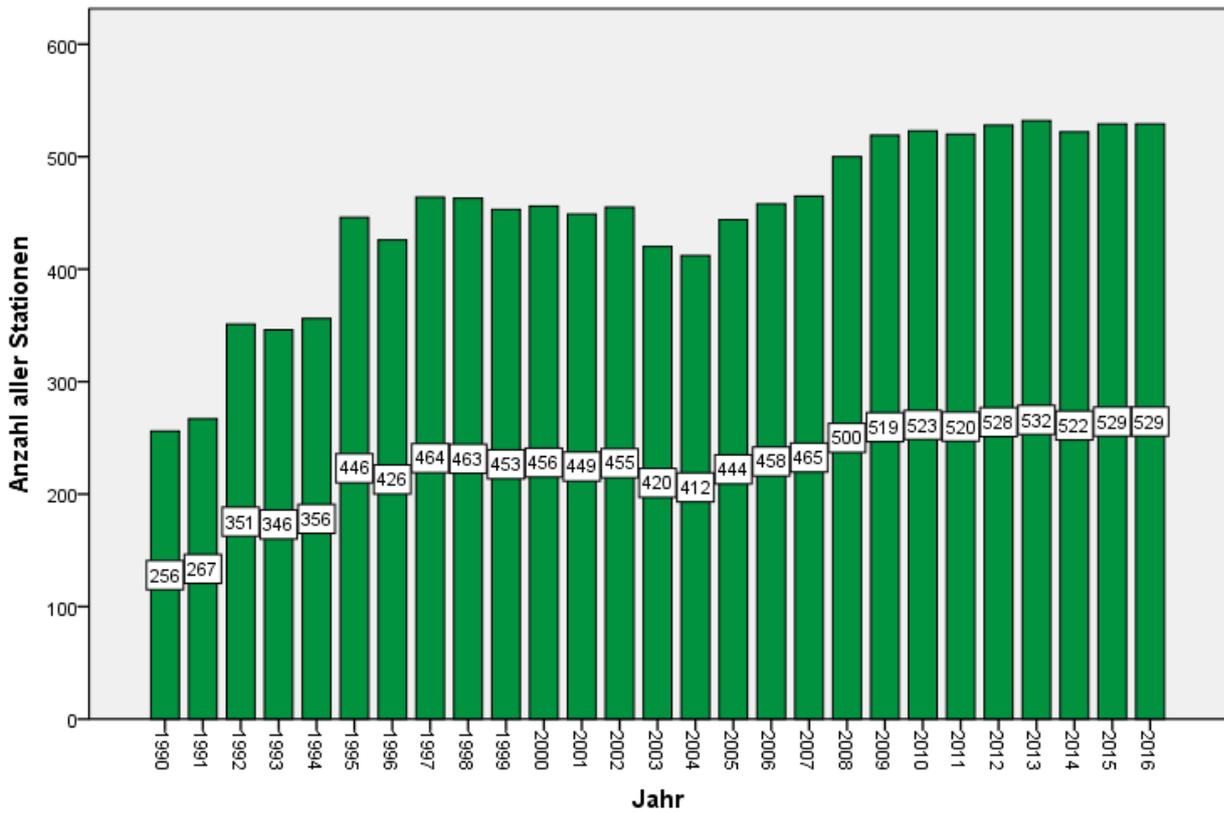
In der Kategorie städtischer Hintergrund driften die Anstiege der beiden Datenbasen ab Rang 184 auseinander. Die Anstiege der Datenbasis der Messwerte ist immer geringer als der der Differenzen der Messwerte. Alle Anstiege sind negativ und signifikant (siehe 3.1.2.13).

In der Kategorie verkehrsnah gibt es ab Rang 183 ein starkes und ab Rang 378 ein sehr starkes auseinanderlaufen der Anstiege der beiden Datenbasen. Ab Rang 583 bleibt die Abweichung der Anstiege zwischen den Datenbasen annähernd konstant bei $0,6 \mu\text{g}/\text{m}^3$ je Jahr. Die Anstiege der Datenbasis der Messwerte streben gegen $-0,07 \mu\text{g}/\text{m}^3$, die der Differenzen der Messwerte gegen $-0,67 \mu\text{g}/\text{m}^3$ je Jahr. Ab Rang 483 ist der Anstieg der Datenbasis der Mittelwerte nicht mehr signifikant. Alle Anstiege der Datenbasis der Differenzen der Messwerte sind signifikant. Die Ursache für dieses Verhalten liegt an der starken Heterogenität der Messreihen (siehe 3.1.2.14).

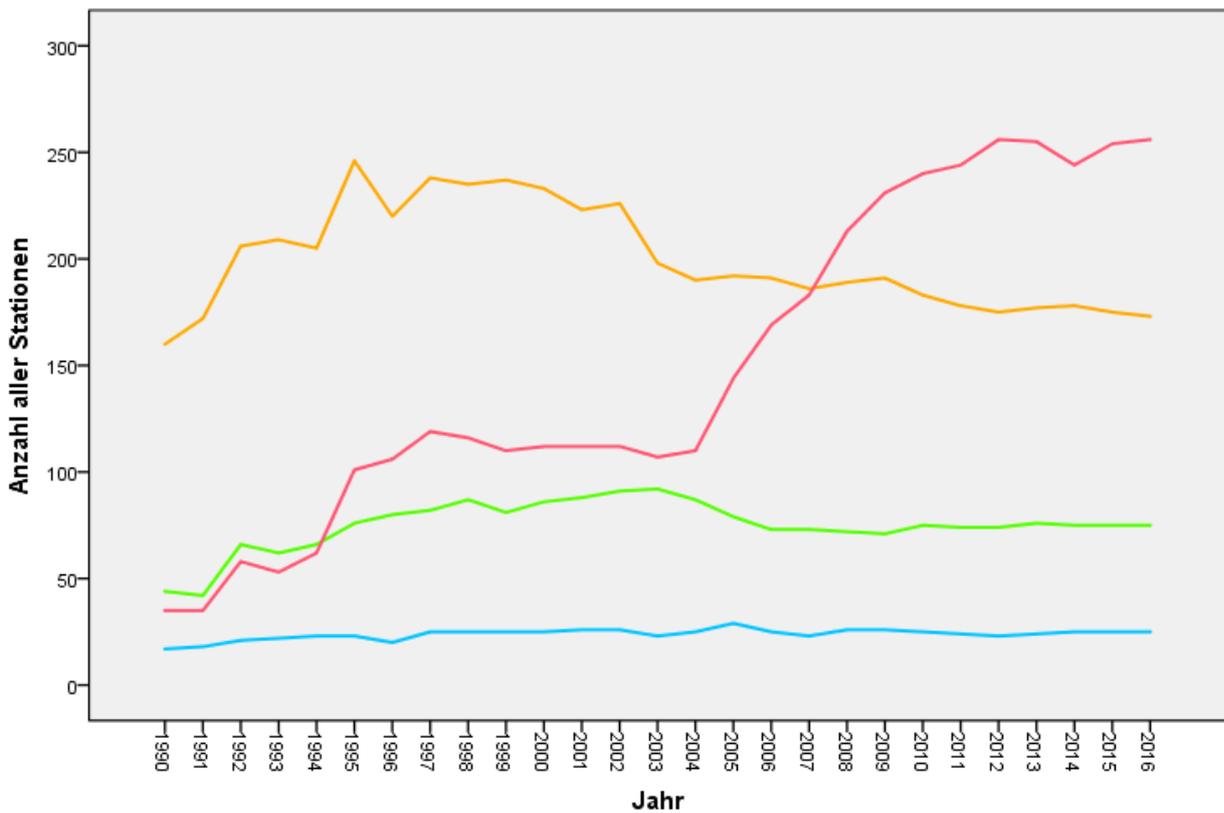
In der Kategorie industrienah zeigen die Anstiege der beiden Datenbasen einen ähnlichen Verlauf aber auf unterschiedlichem Niveau. Die Anstiege der Datenbasis der Differenzen der Messwerte sind gleichmäßig höher als die der Datenbasis der Messwerte. Alle Anstiege sind negativ und signifikant (siehe 3.1.2.15).

Die Ergebnisse der Trendberechnung (Messwerte vs. Differenzen und durchgängig messend vs. nicht durchgängig messend) sind in den Fehlerbalkendiagrammen mit 95% Konfidenzintervallen zusammengefasst (siehe 3.1.2.16).

3.1.2.1 NO₂ - Anzahl aller Stationen

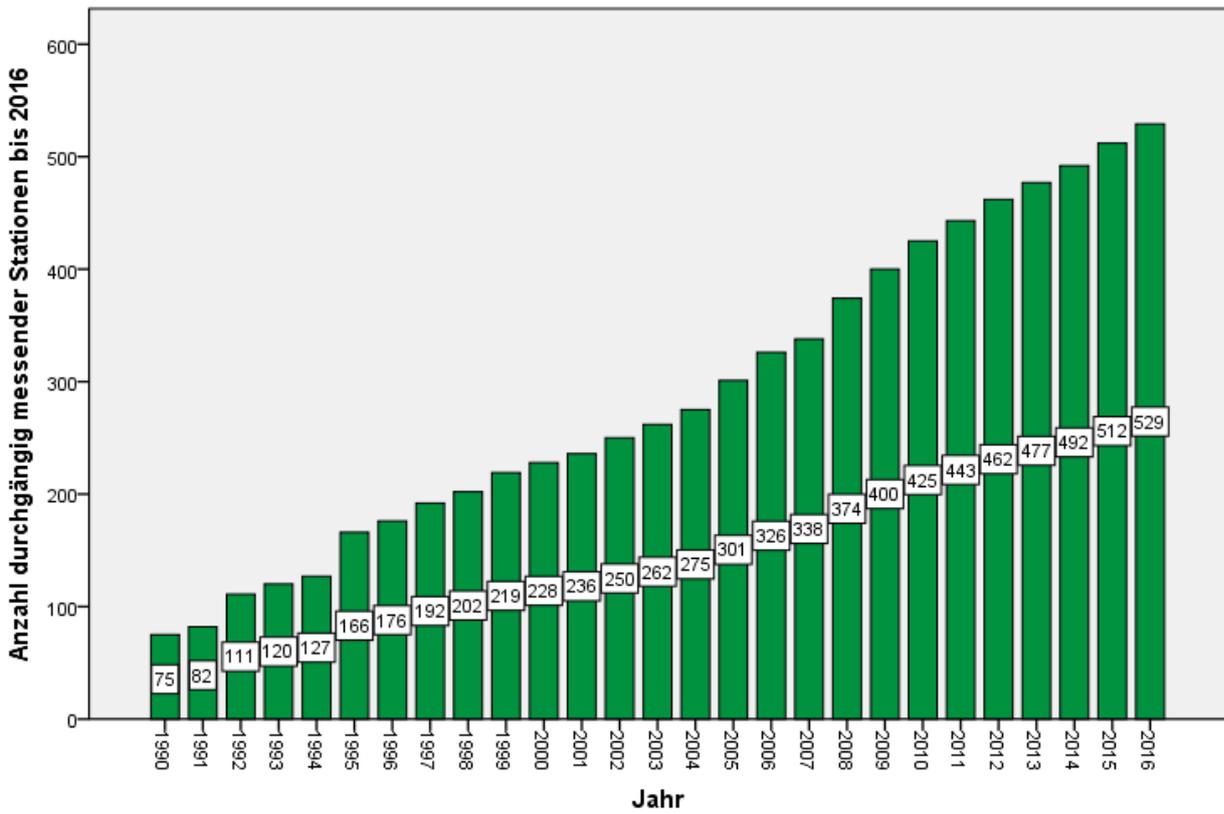


eigene Darstellung, BioMath GmbH

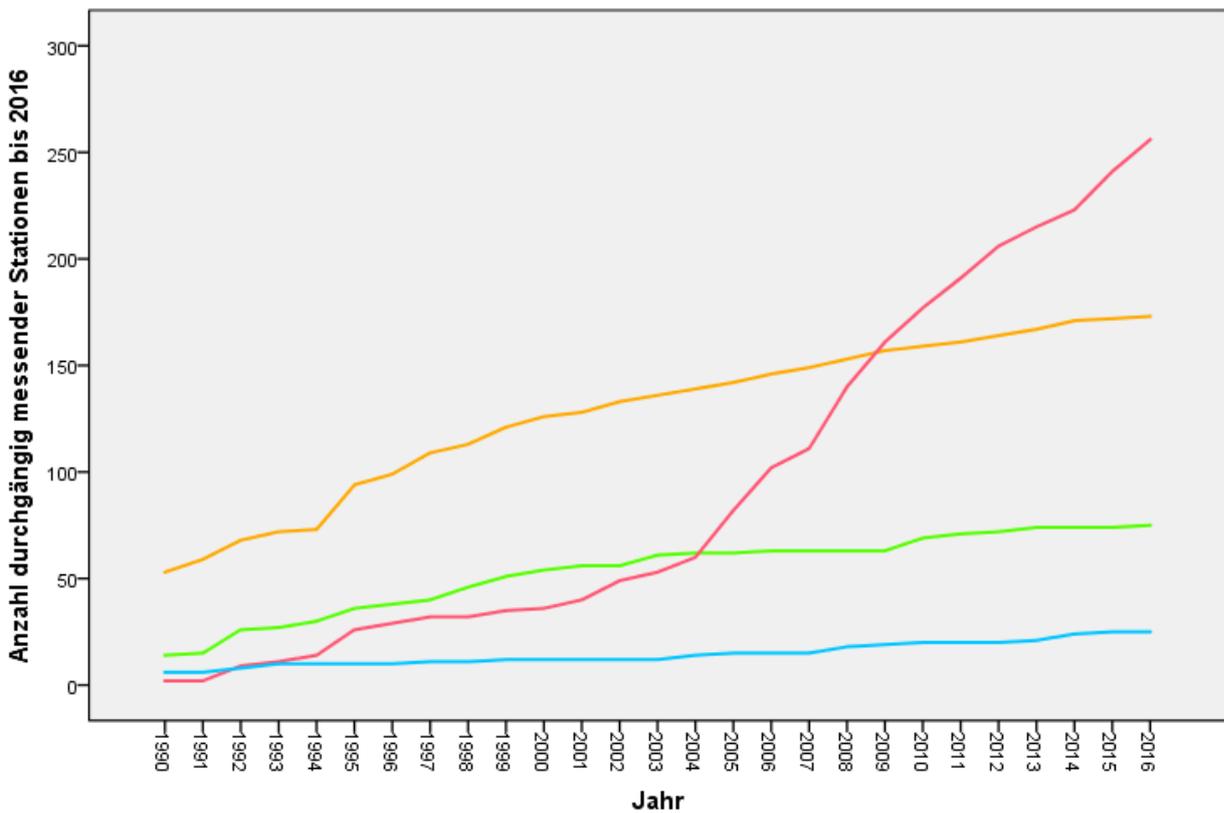


— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnahe
 — industrienah
 eigene Darstellung, BioMath GmbH

3.1.2.2 NO₂ - Anzahl durchgängig messender Stationen bis 2016

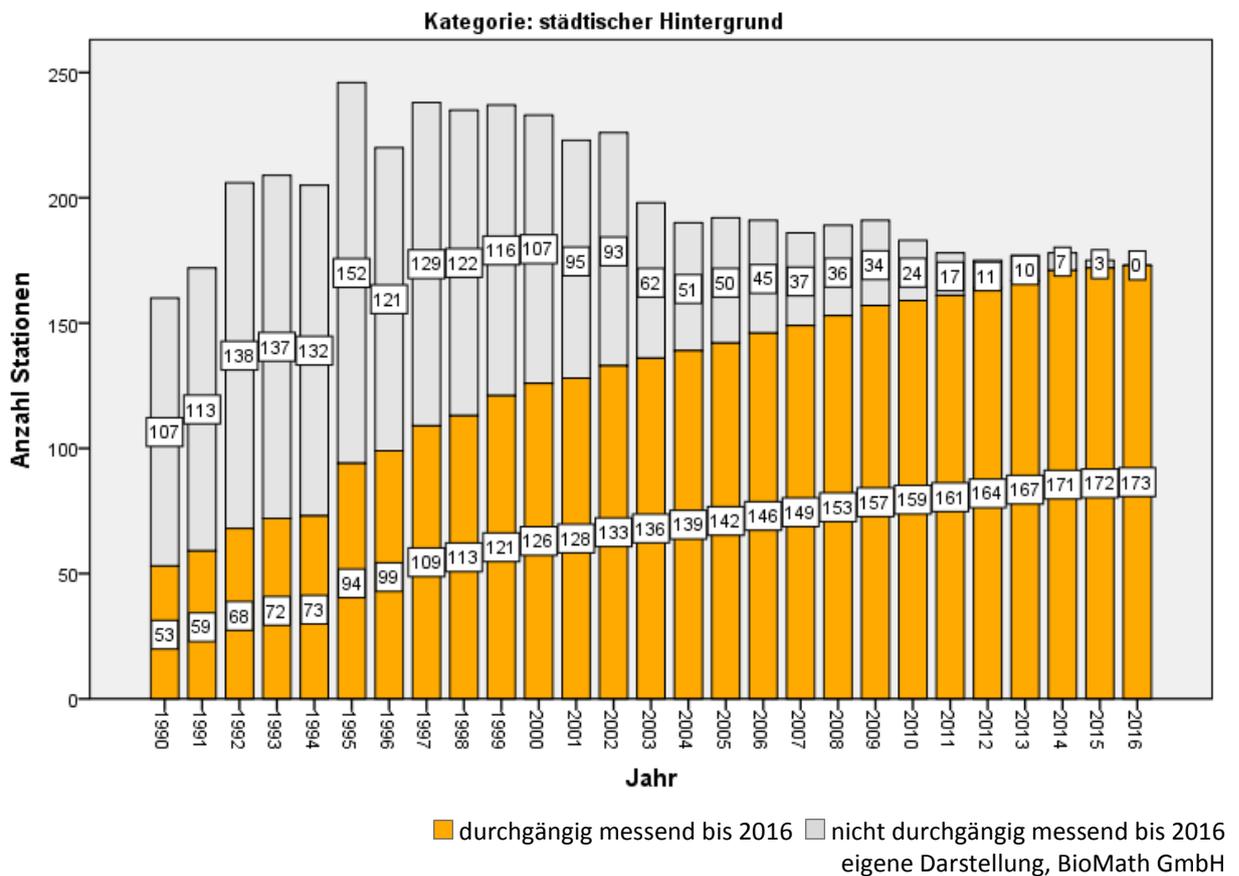
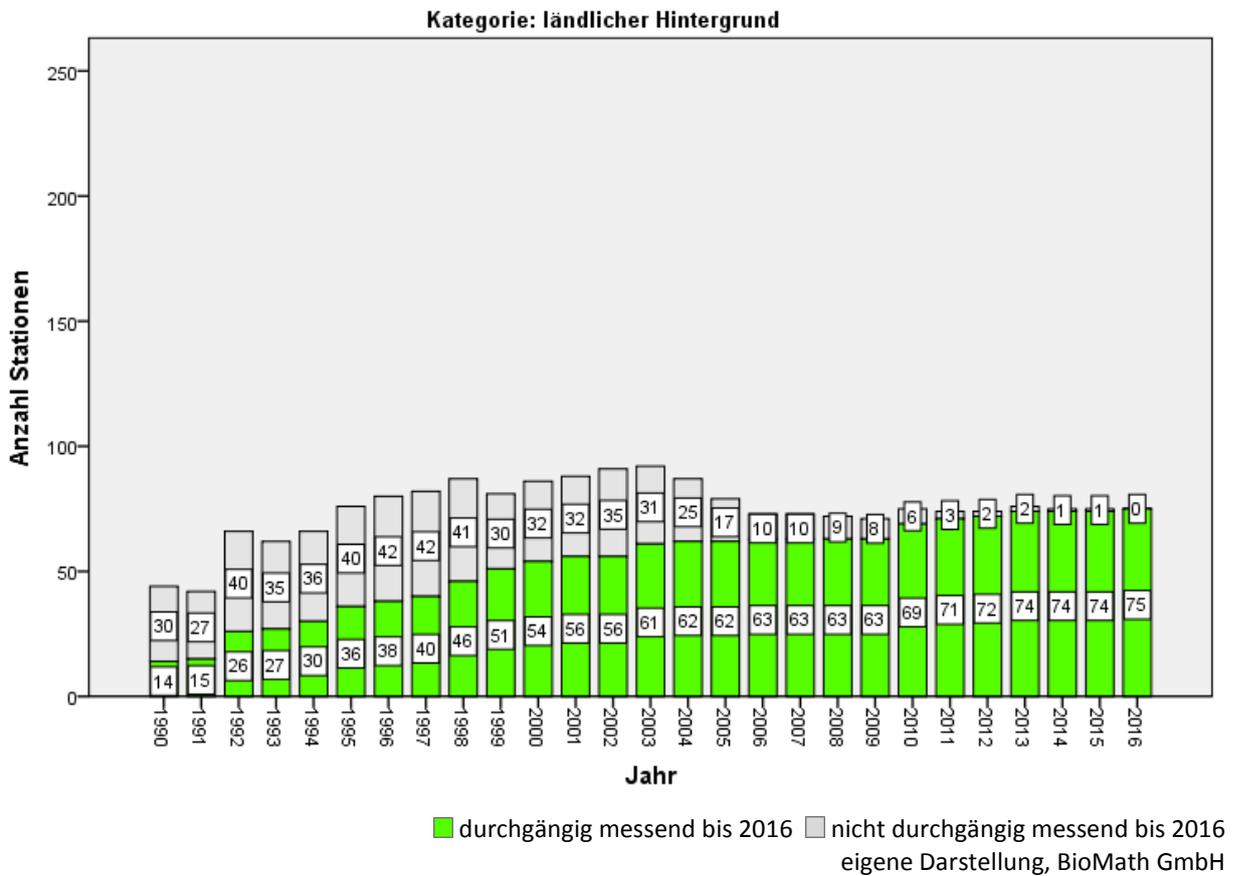


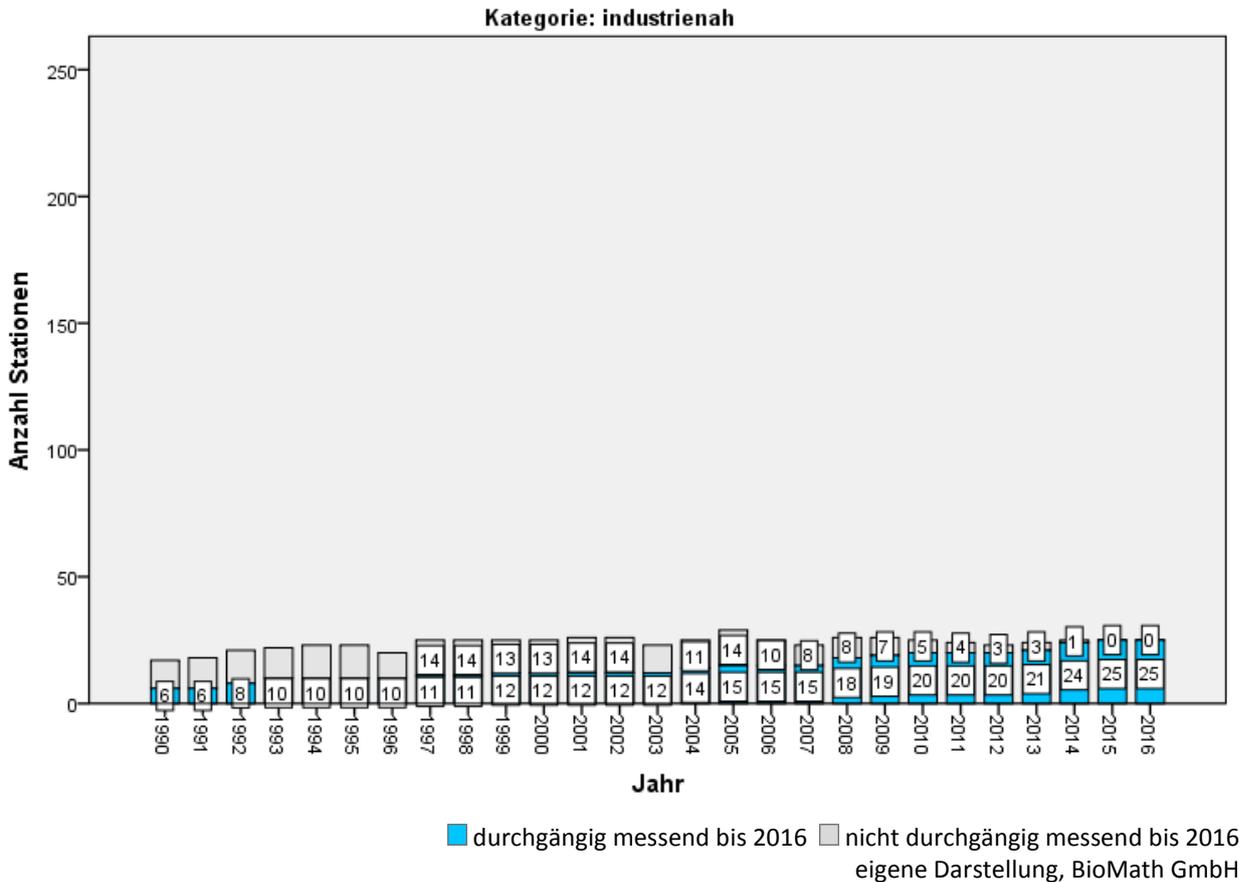
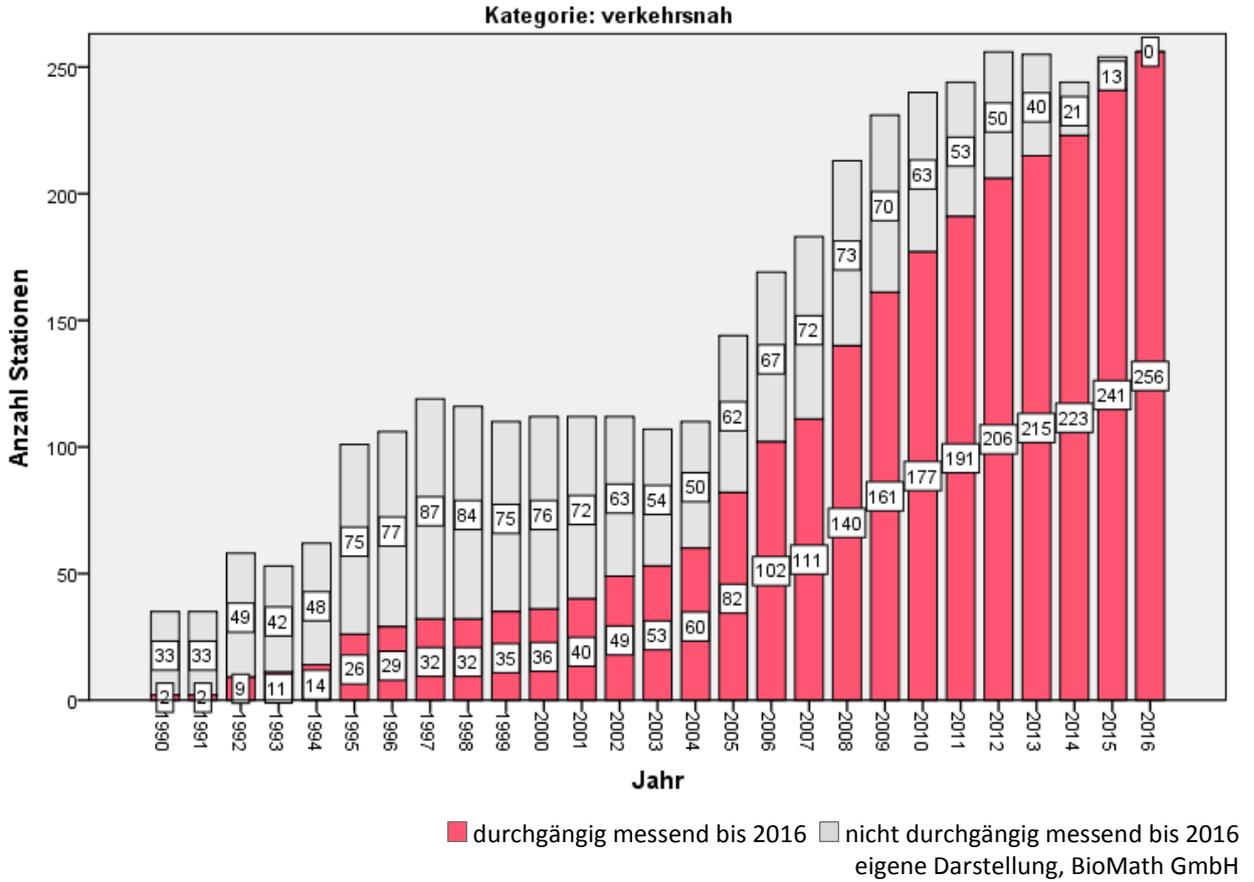
eigene Darstellung, BioMath GmbH



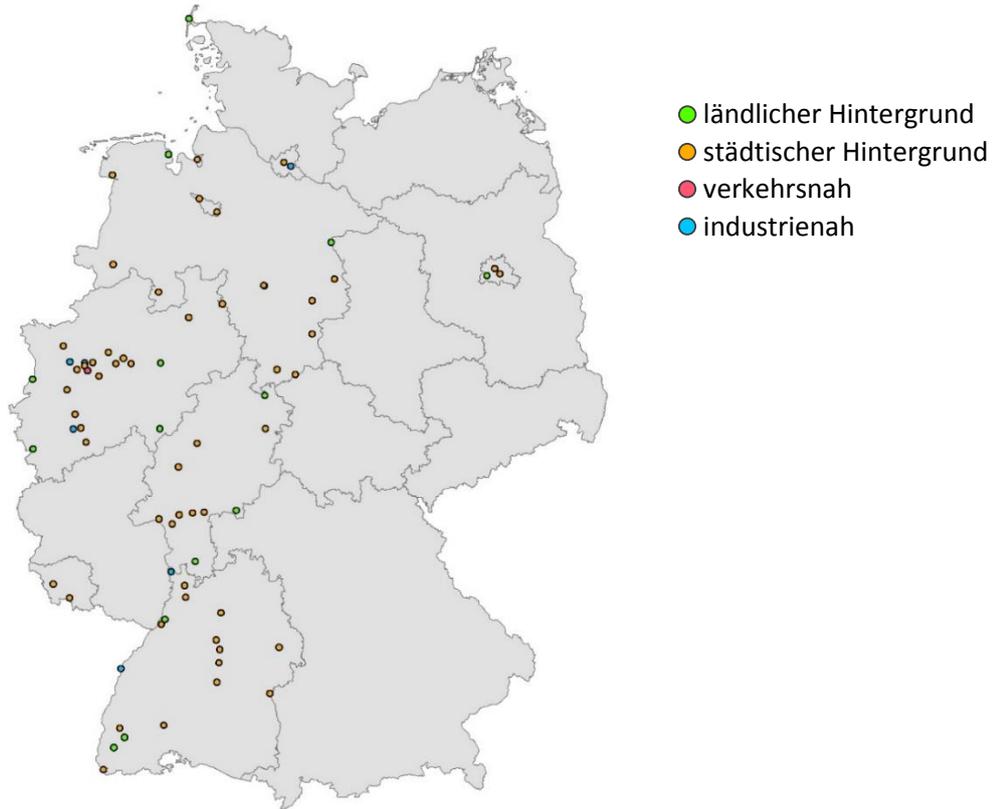
— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnahe
 — industrienah
 eigene Darstellung, BioMath GmbH

3.1.2.3 NO₂ - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016

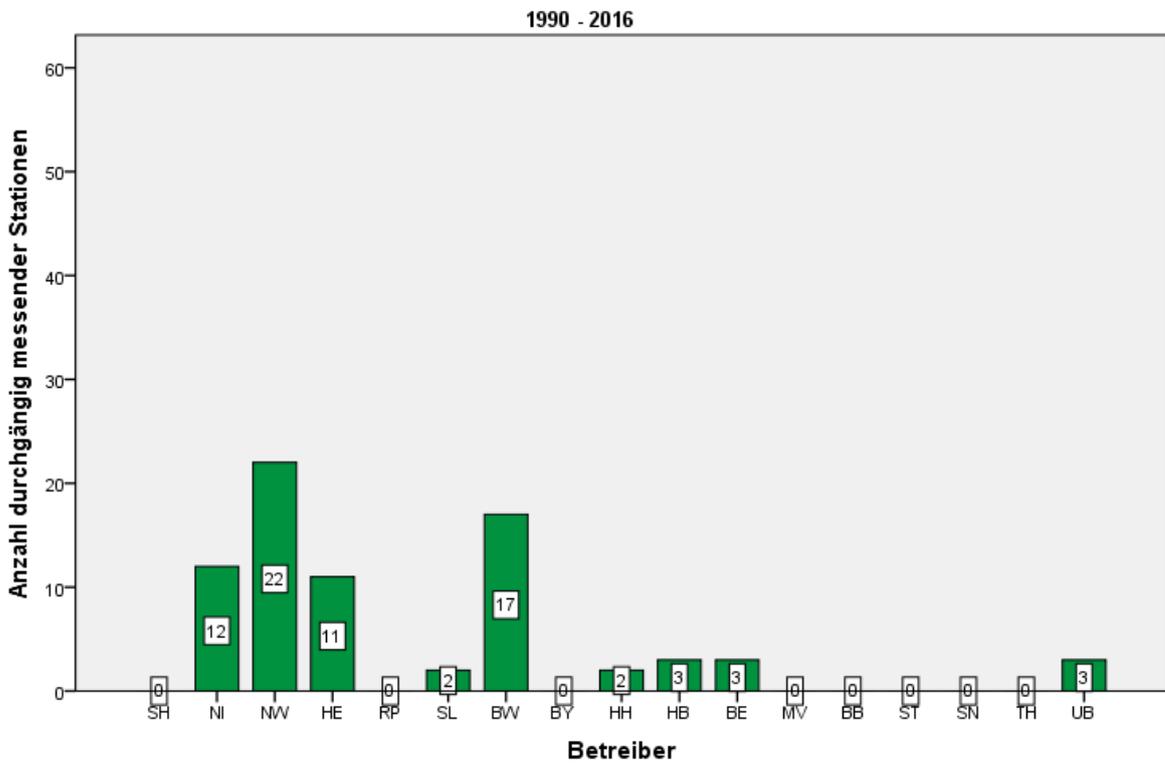




3.1.2.4 NO₂ - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)

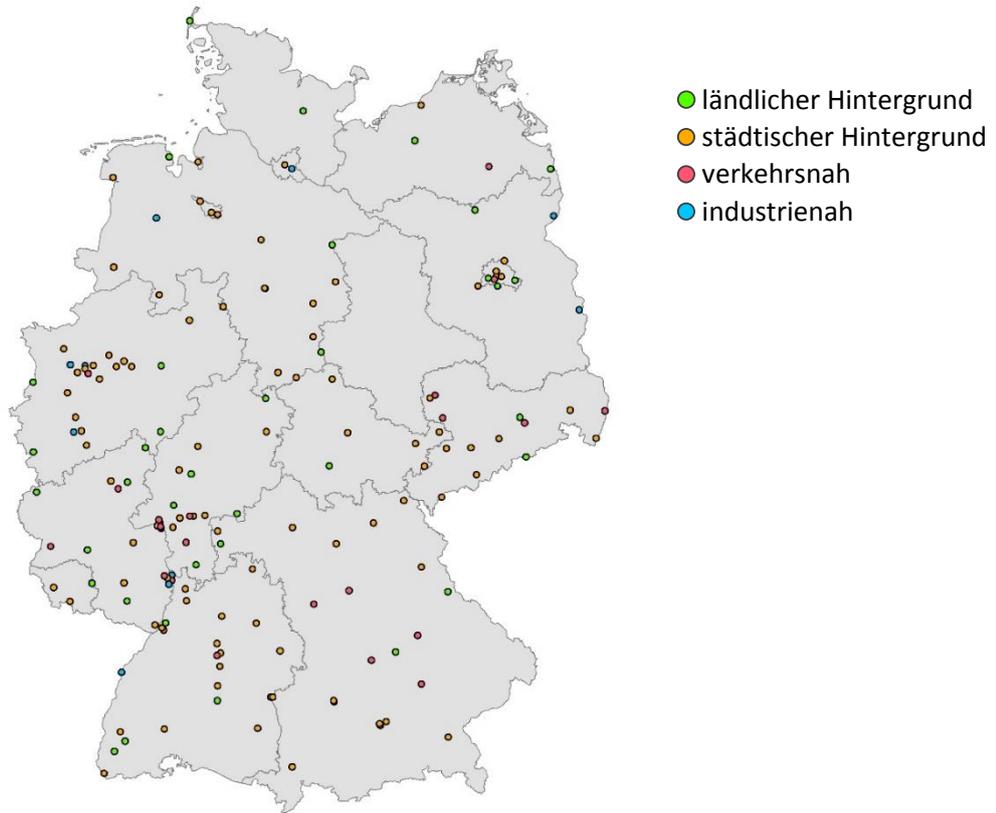


eigene Darstellung, BioMath GmbH

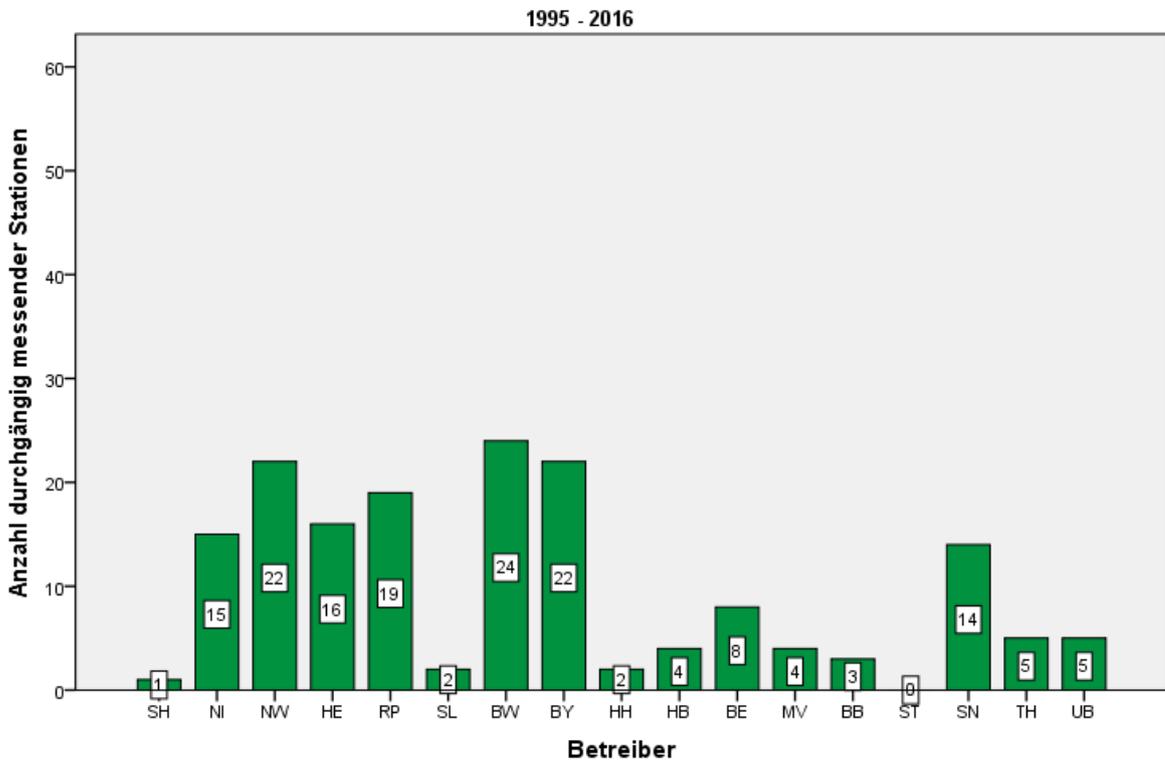


eigene Darstellung, BioMath GmbH

3.1.2.5 NO₂ - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)



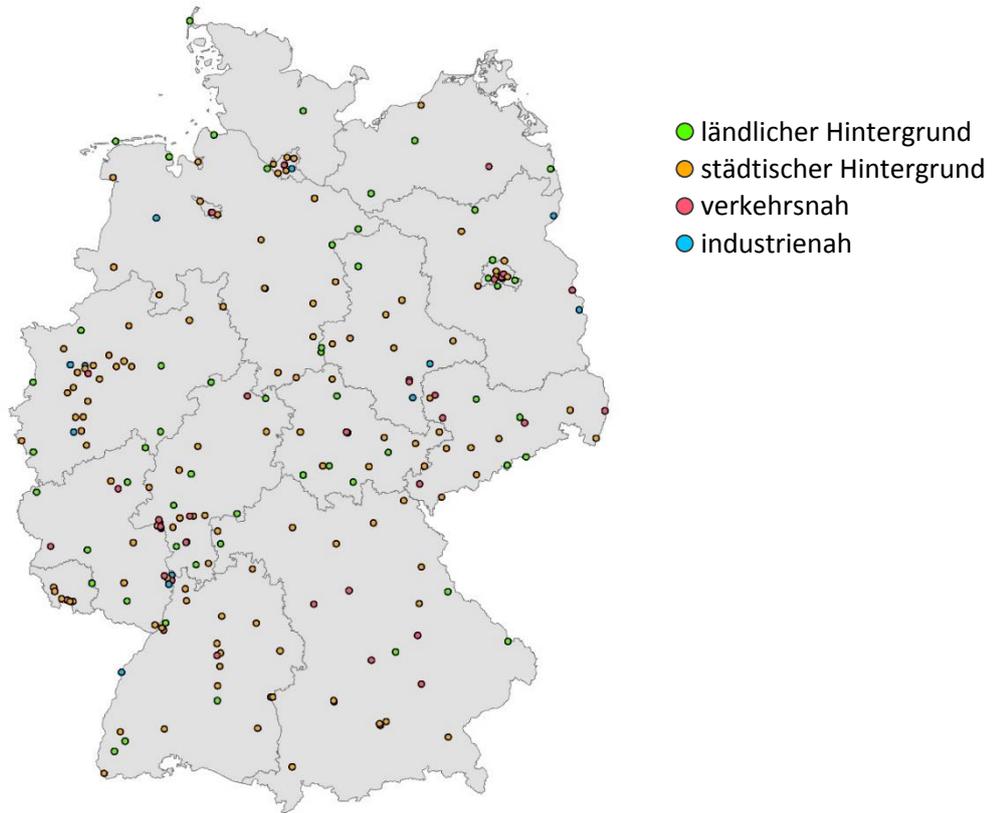
eigene Darstellung, BioMath



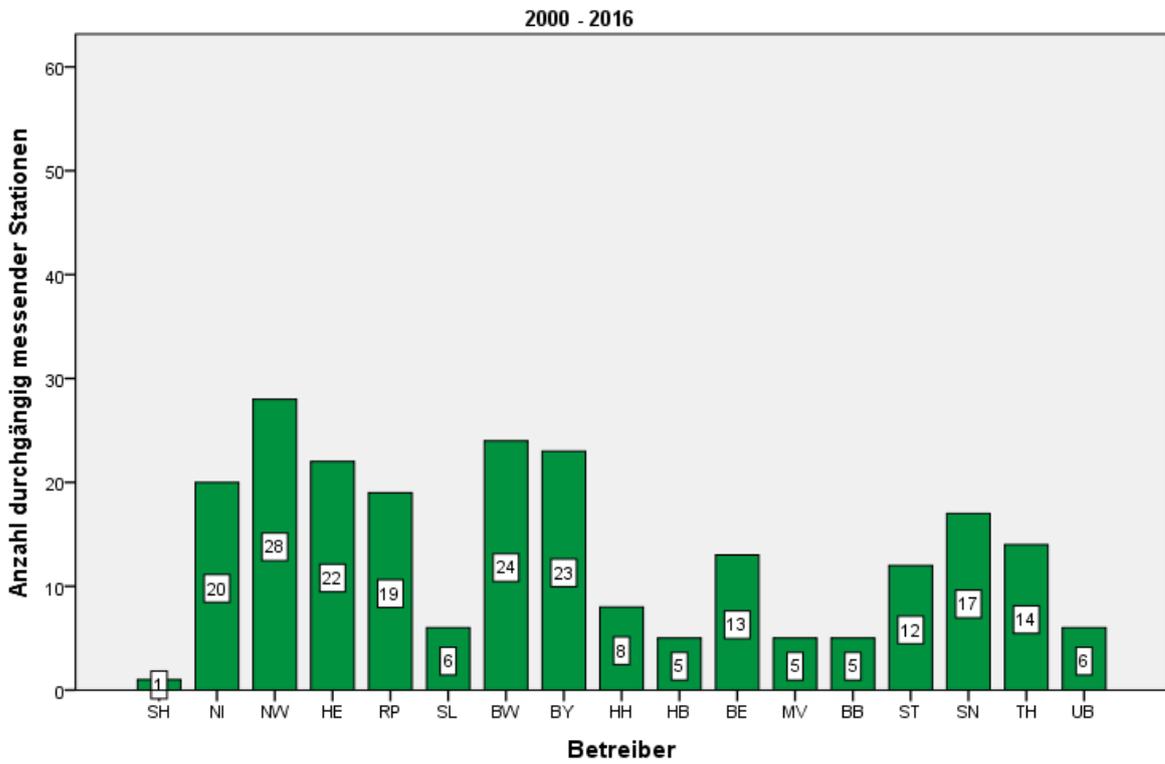
GmbH

eigene Darstellung, BioMath GmbH

3.1.2.6 NO₂ - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)



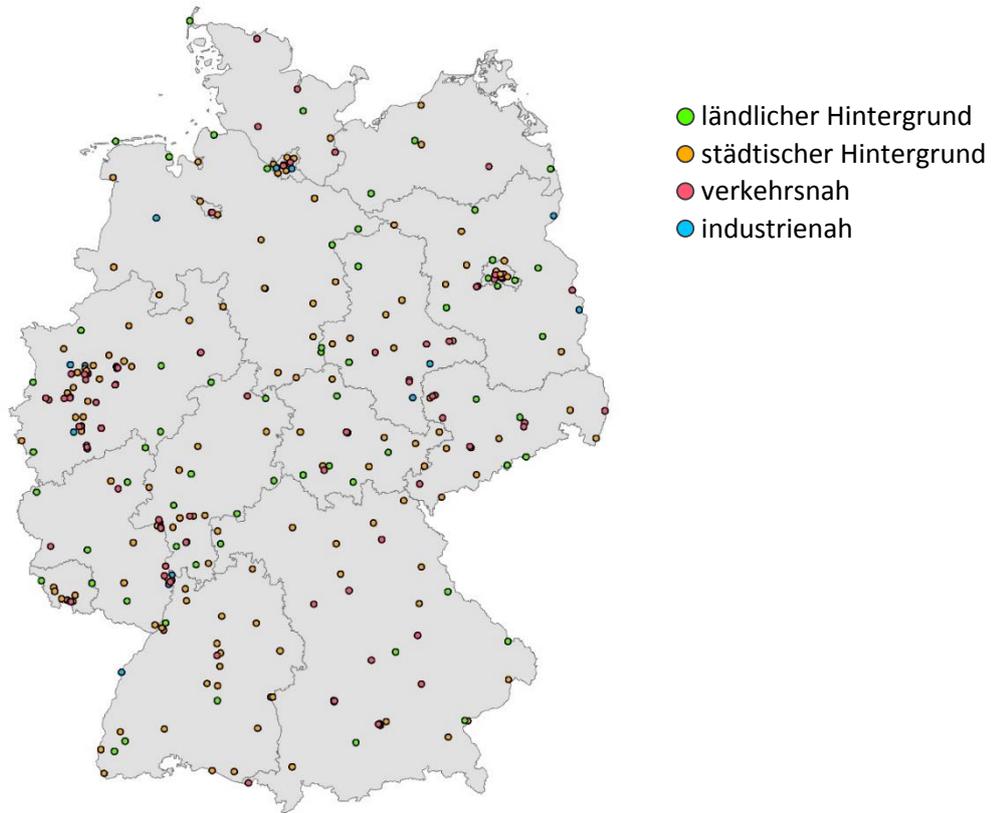
eigene Darstellung, BioMath



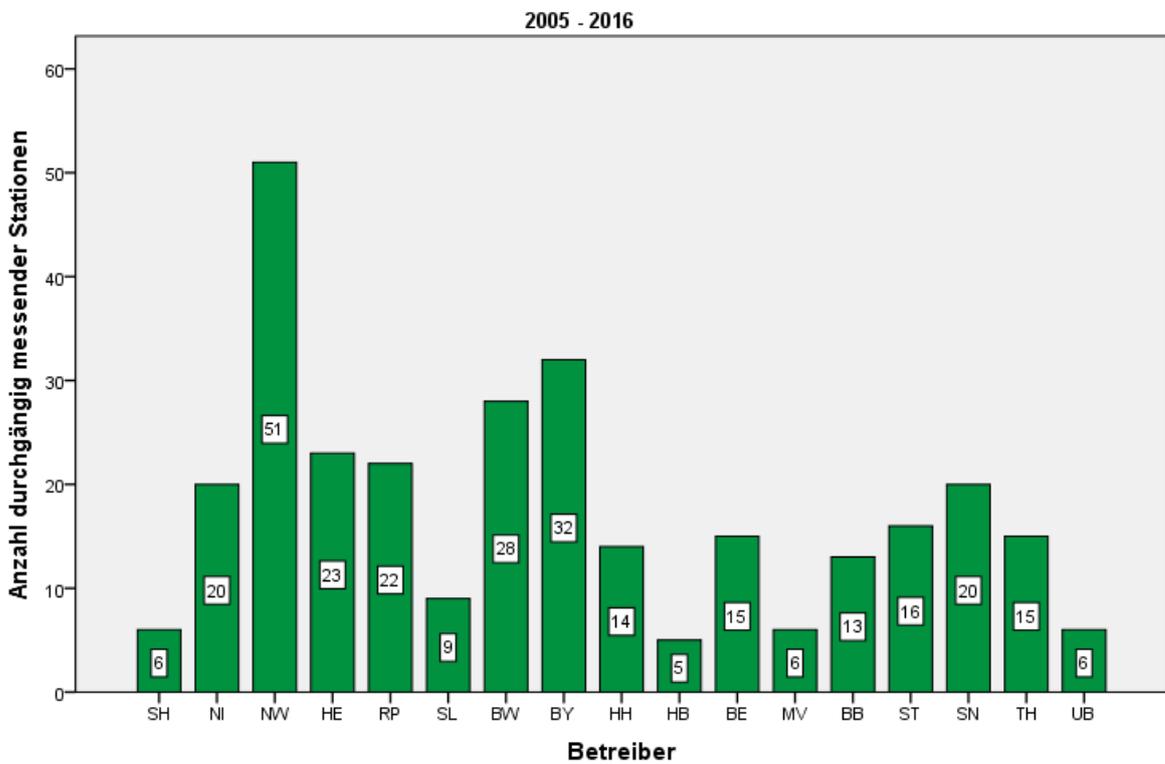
GmbH

eigene Darstellung, BioMath GmbH

3.1.2.7 NO₂ - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)



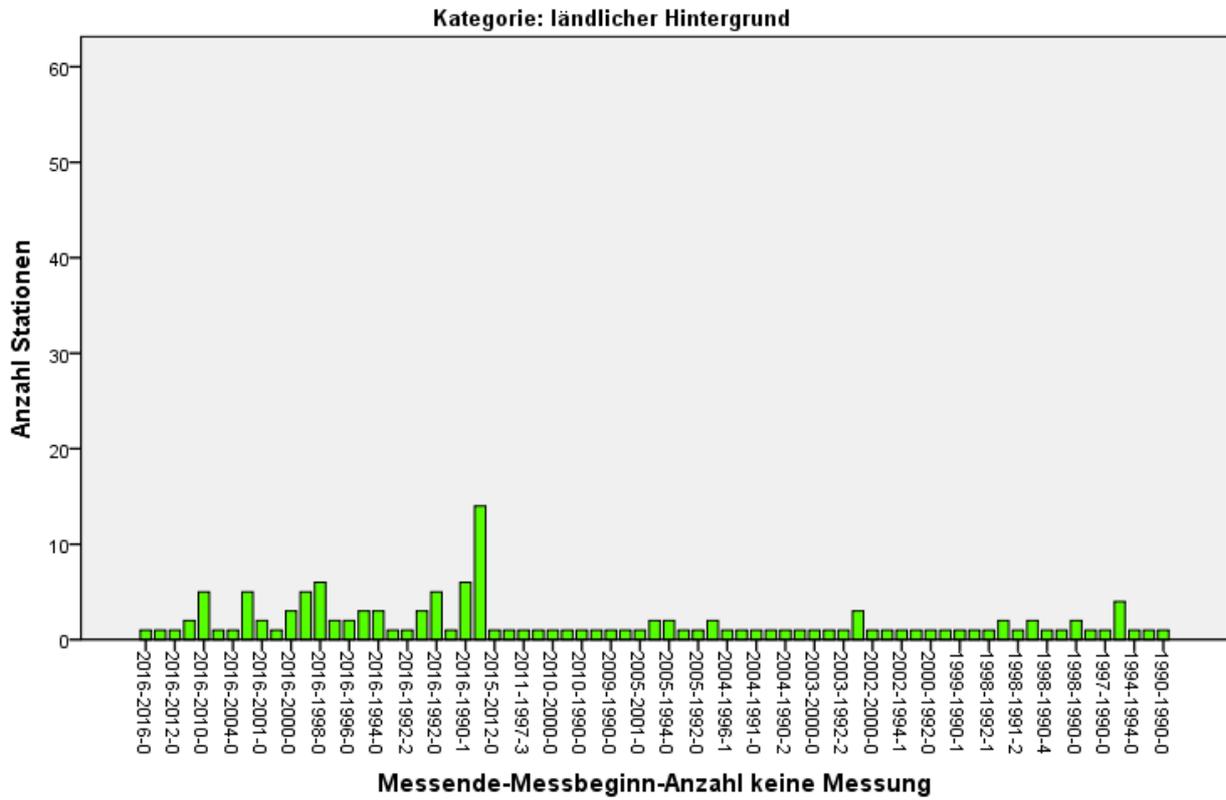
eigene Darstellung, BioMath



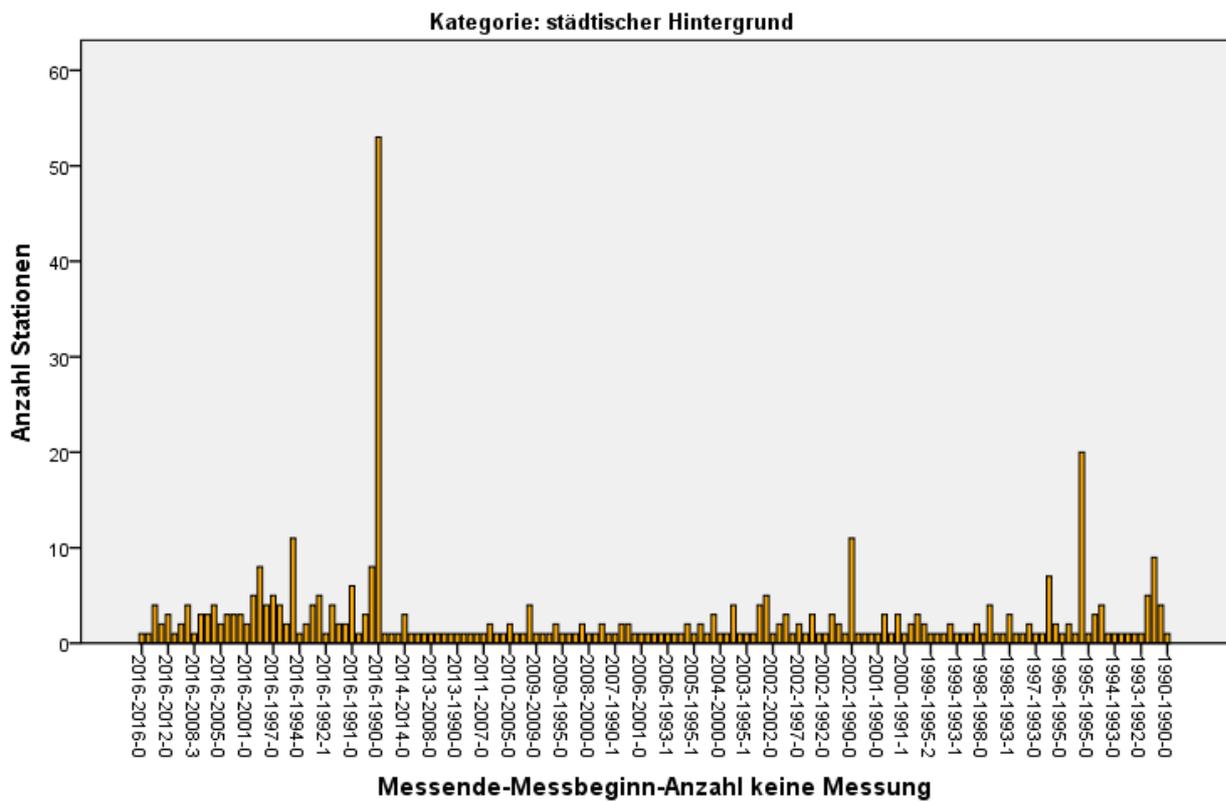
GmbH

eigene Darstellung, BioMath GmbH

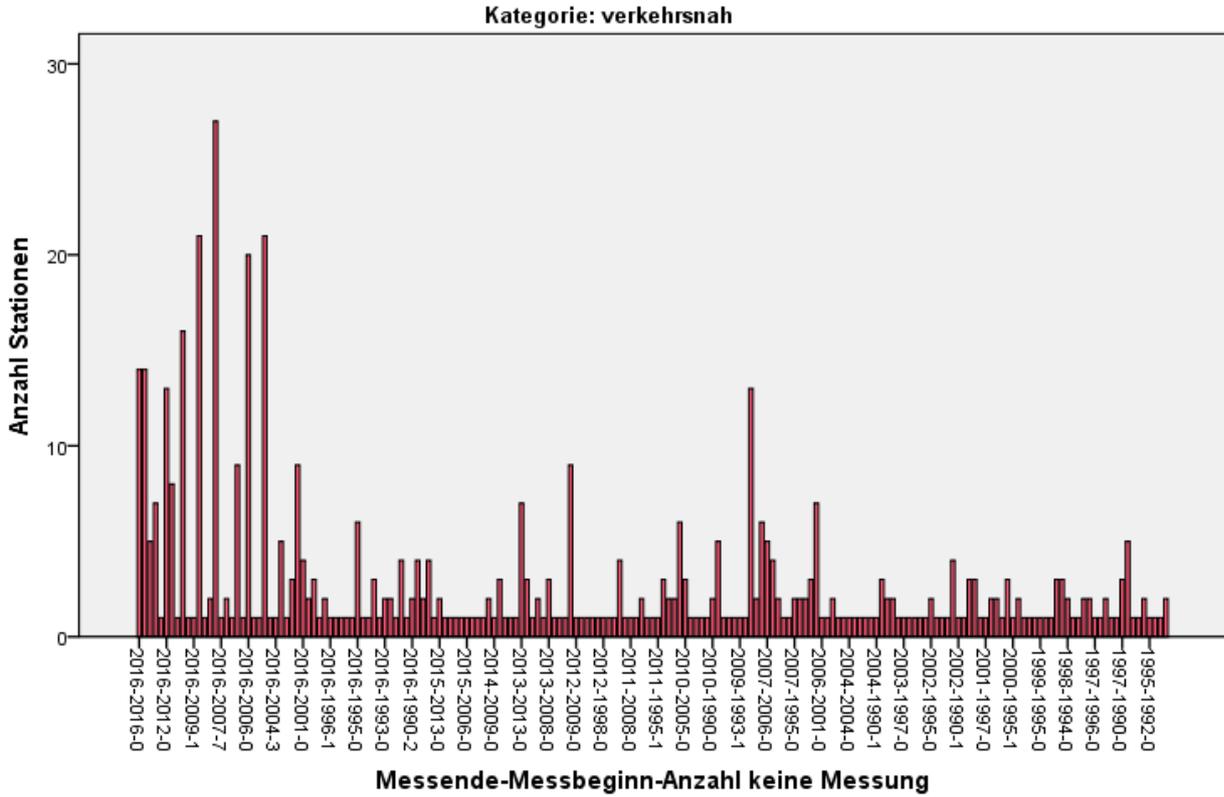
3.1.2.8 NO₂ - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden



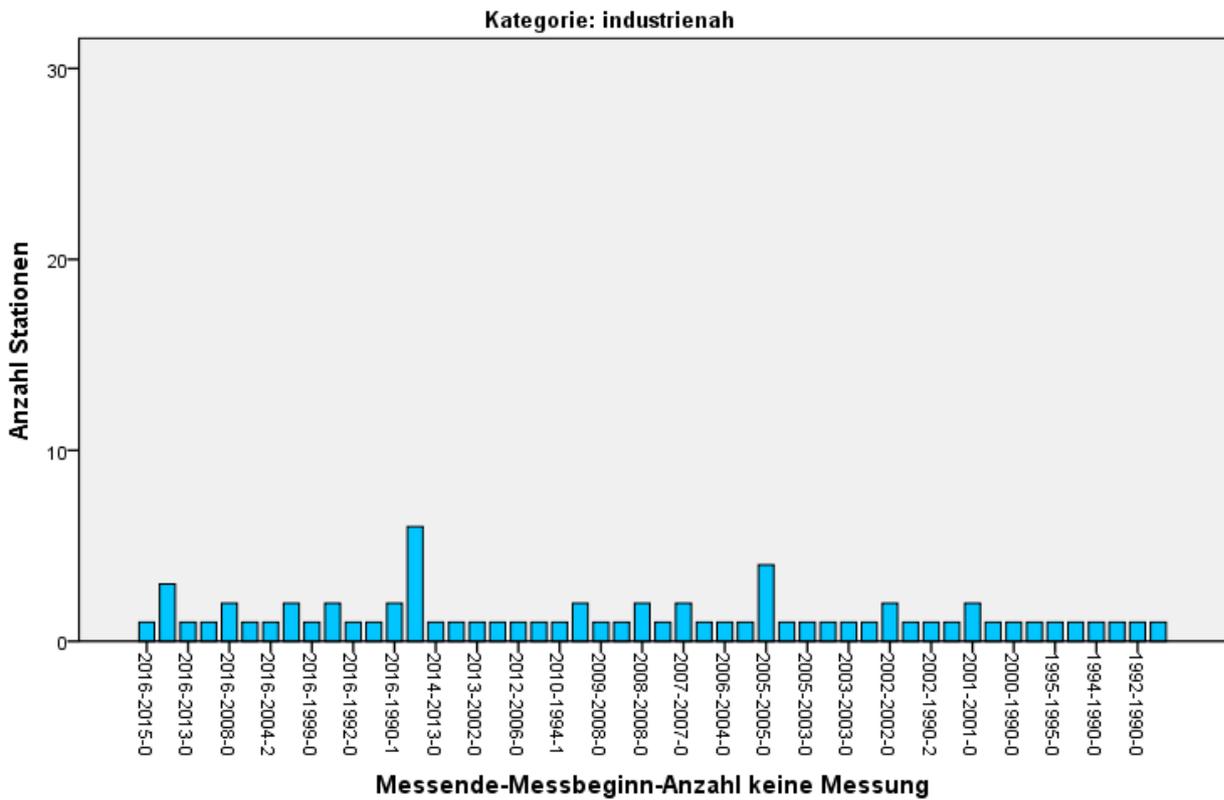
eigene Darstellung, BioMath GmbH



eigene Darstellung, BioMath GmbH

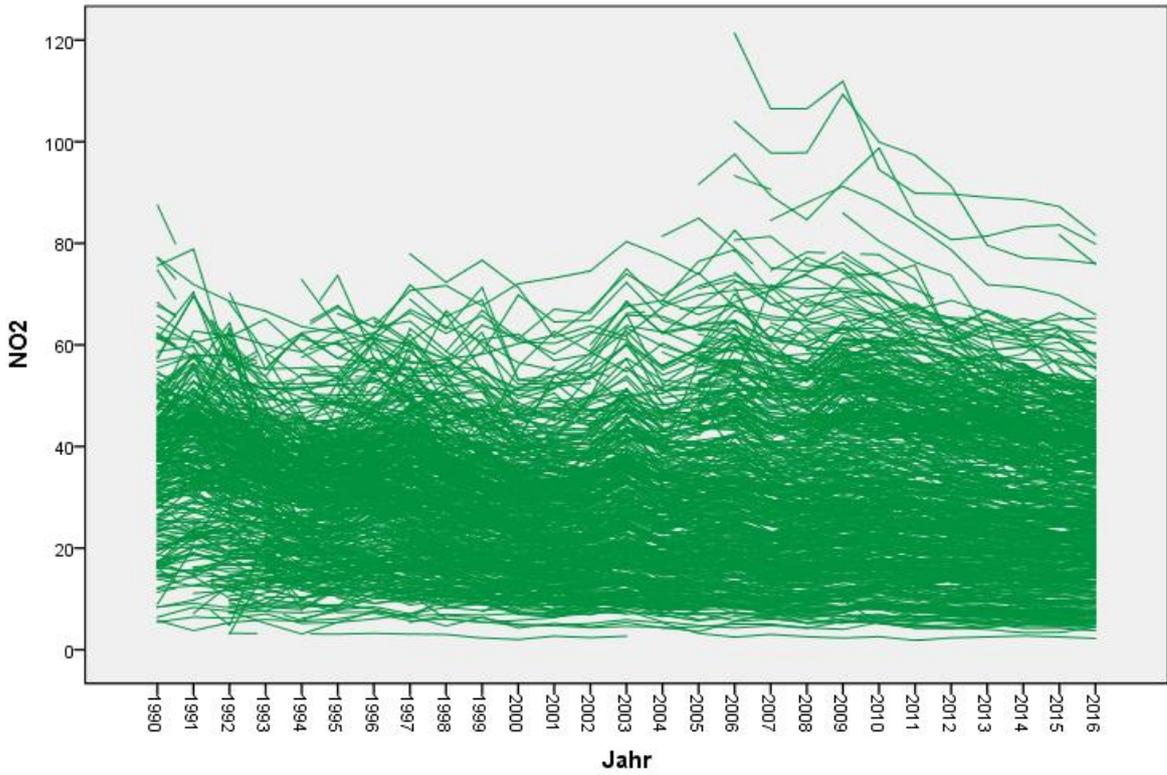


eigene Darstellung, BioMath GmbH

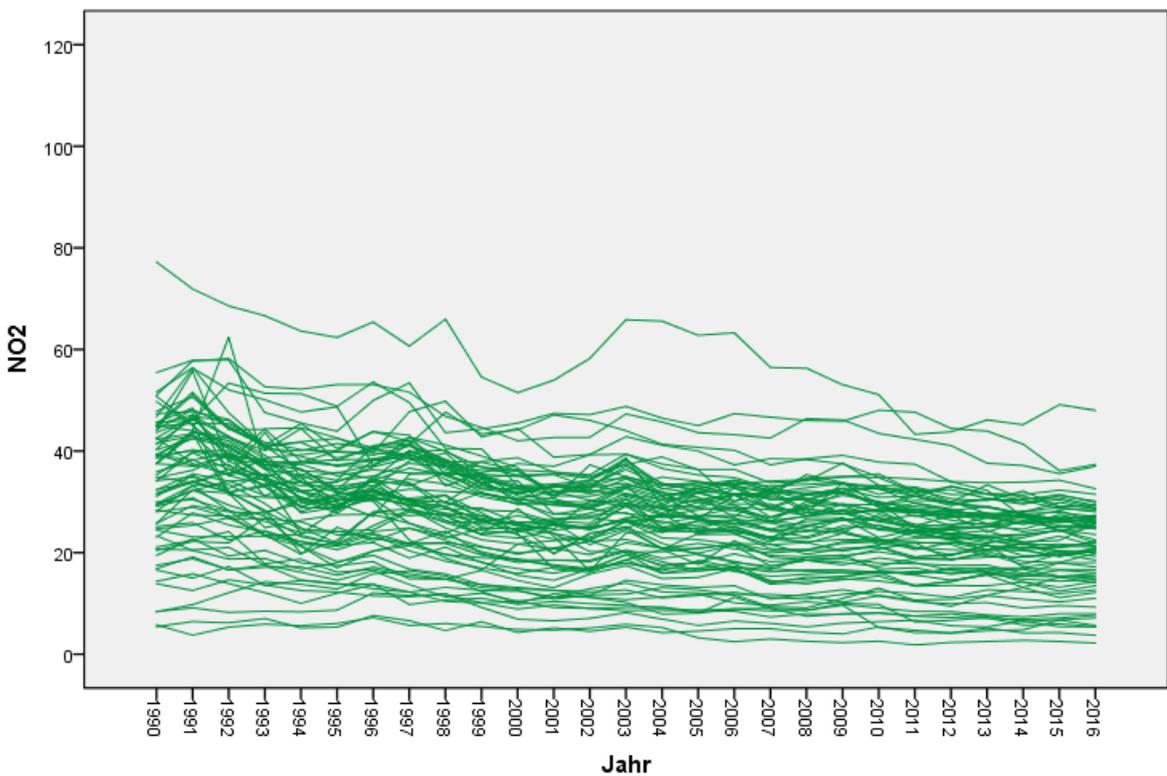


eigene Darstellung, BioMath GmbH

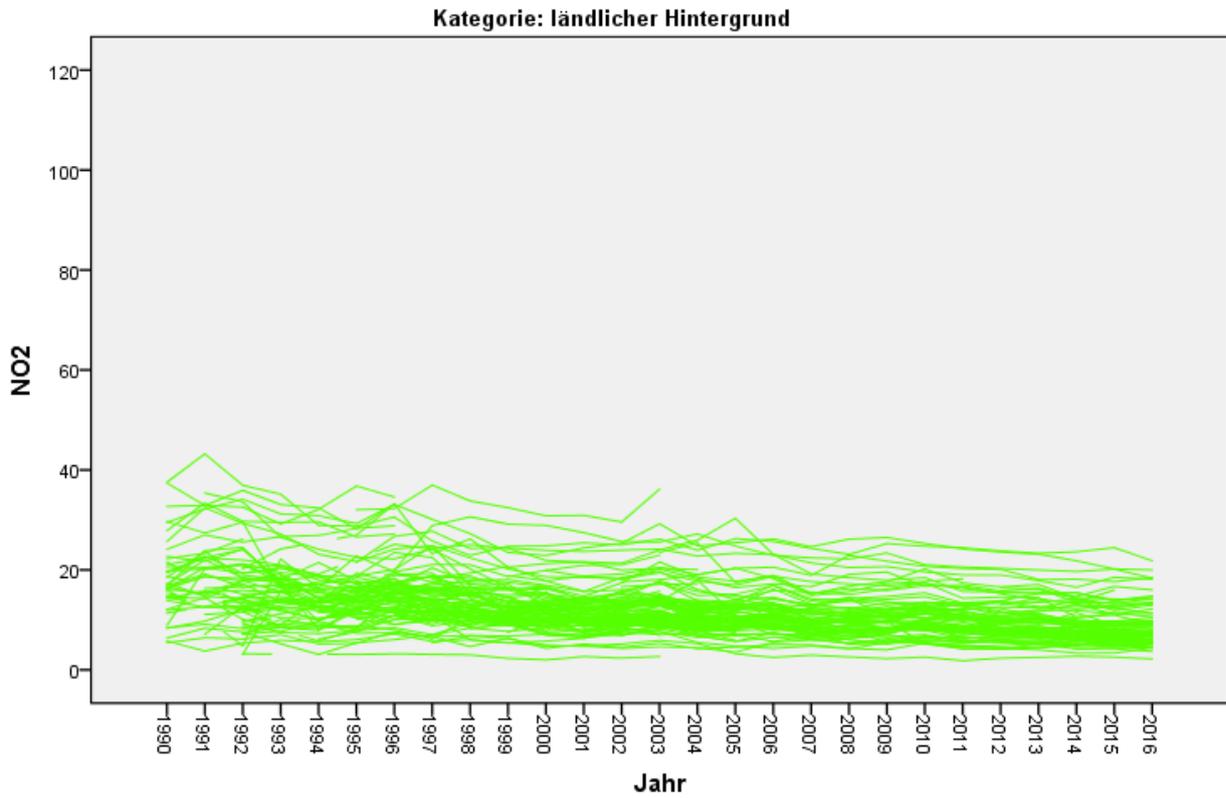
3.1.2.9 NO₂-Werte in µg/m³- alle Messreihen und durchgängig messende Messreihen von 1990 - 2016



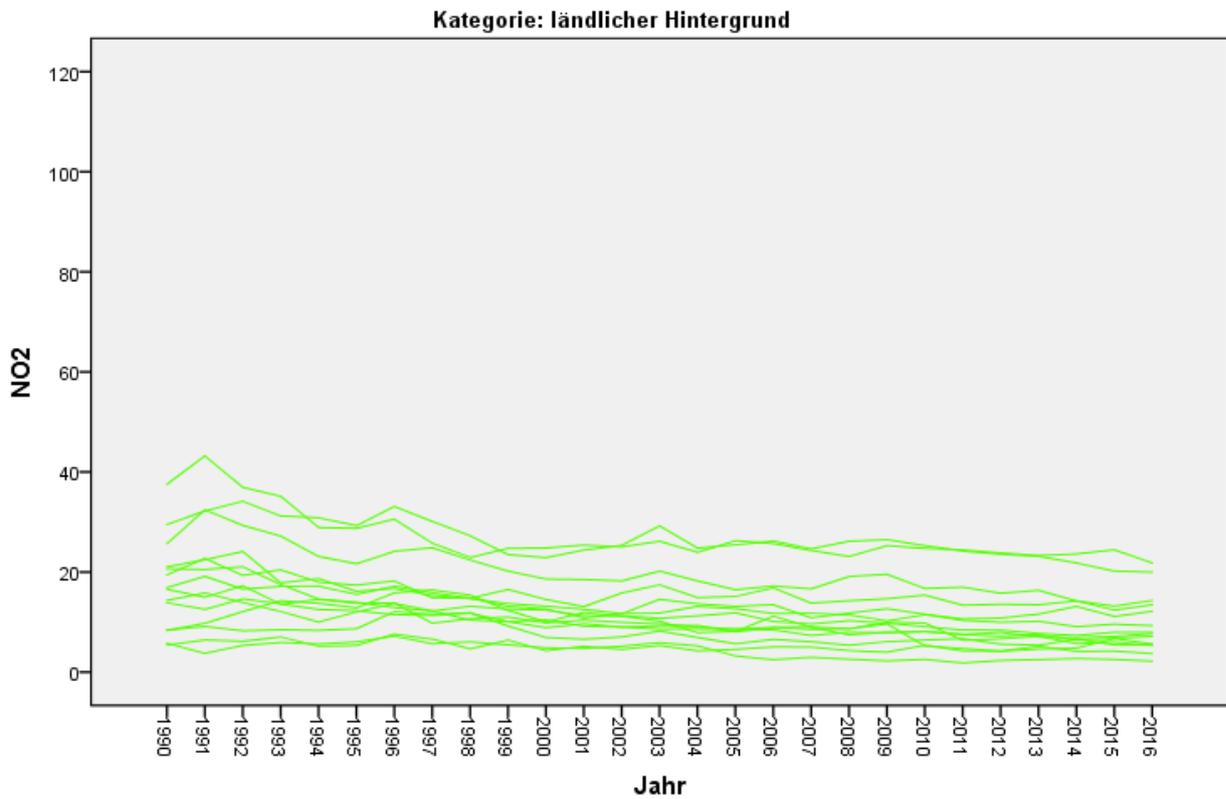
alle Messreihen
eigene Darstellung, BioMath GmbH



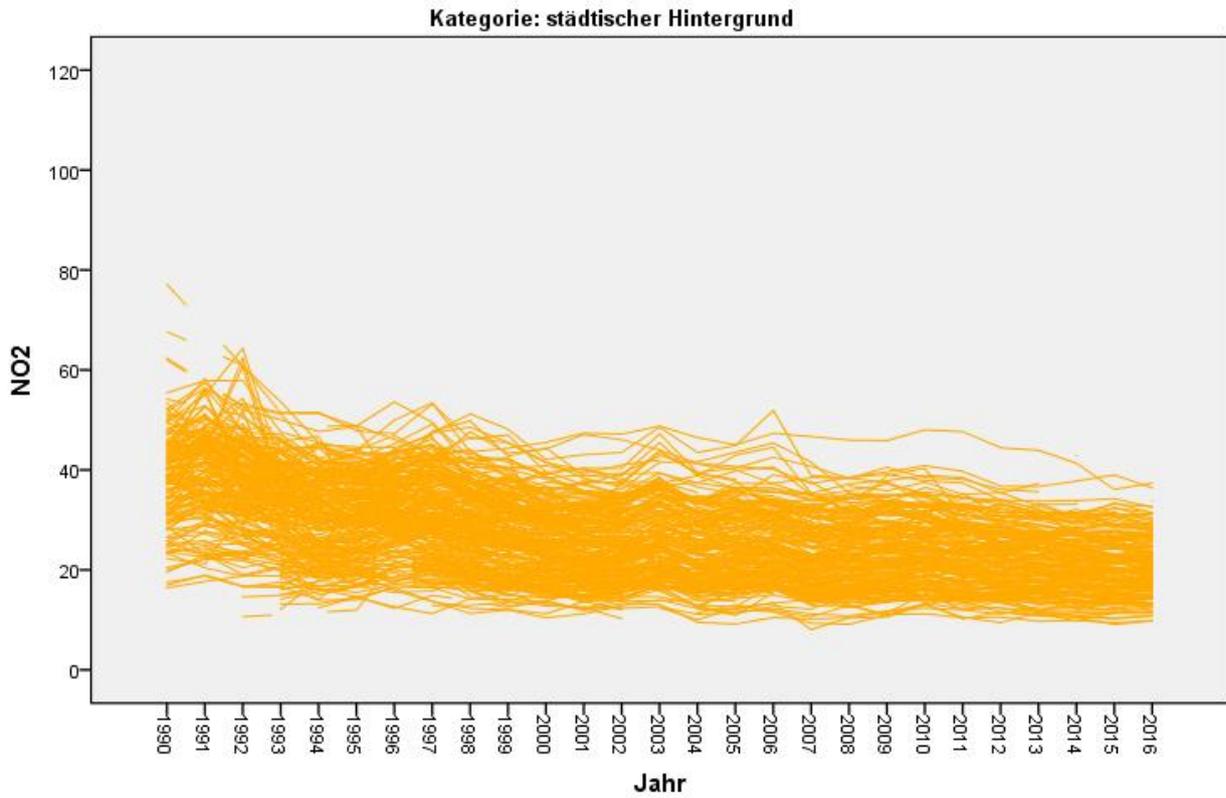
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



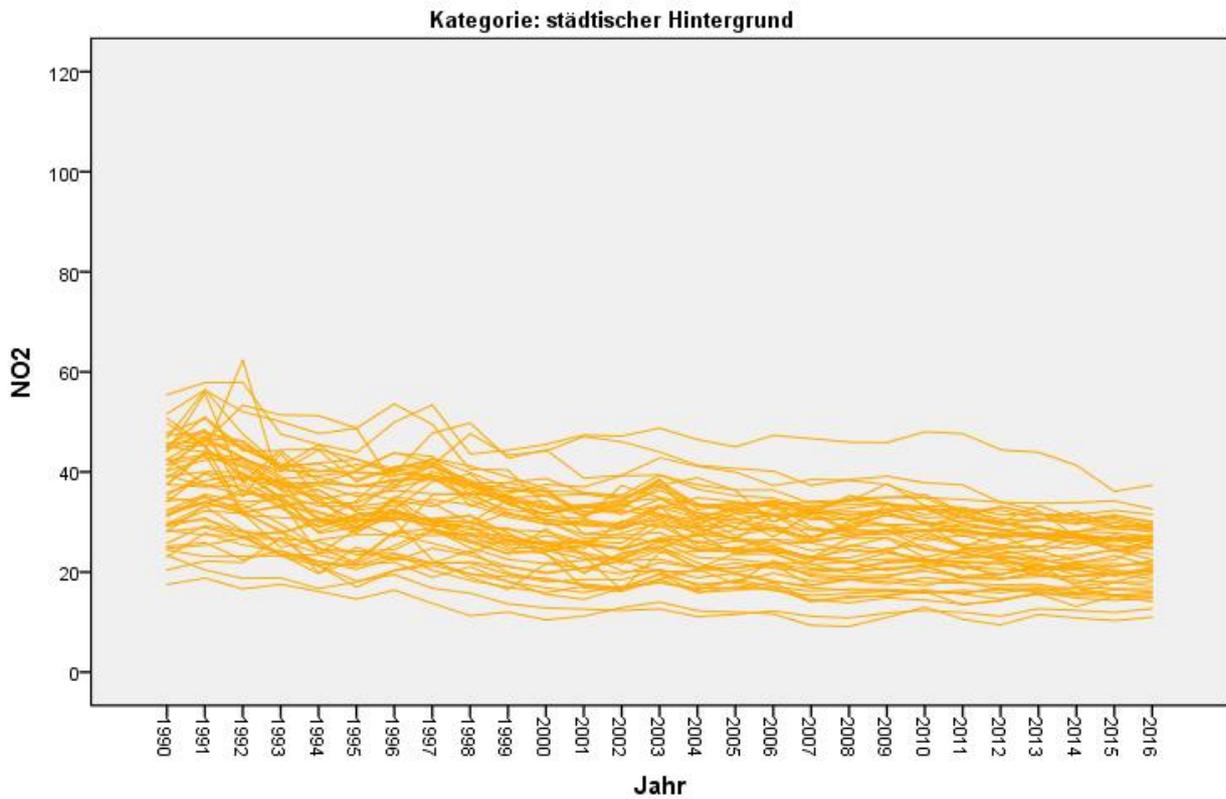
alle Messreihen
eigene Darstellung, BioMath GmbH



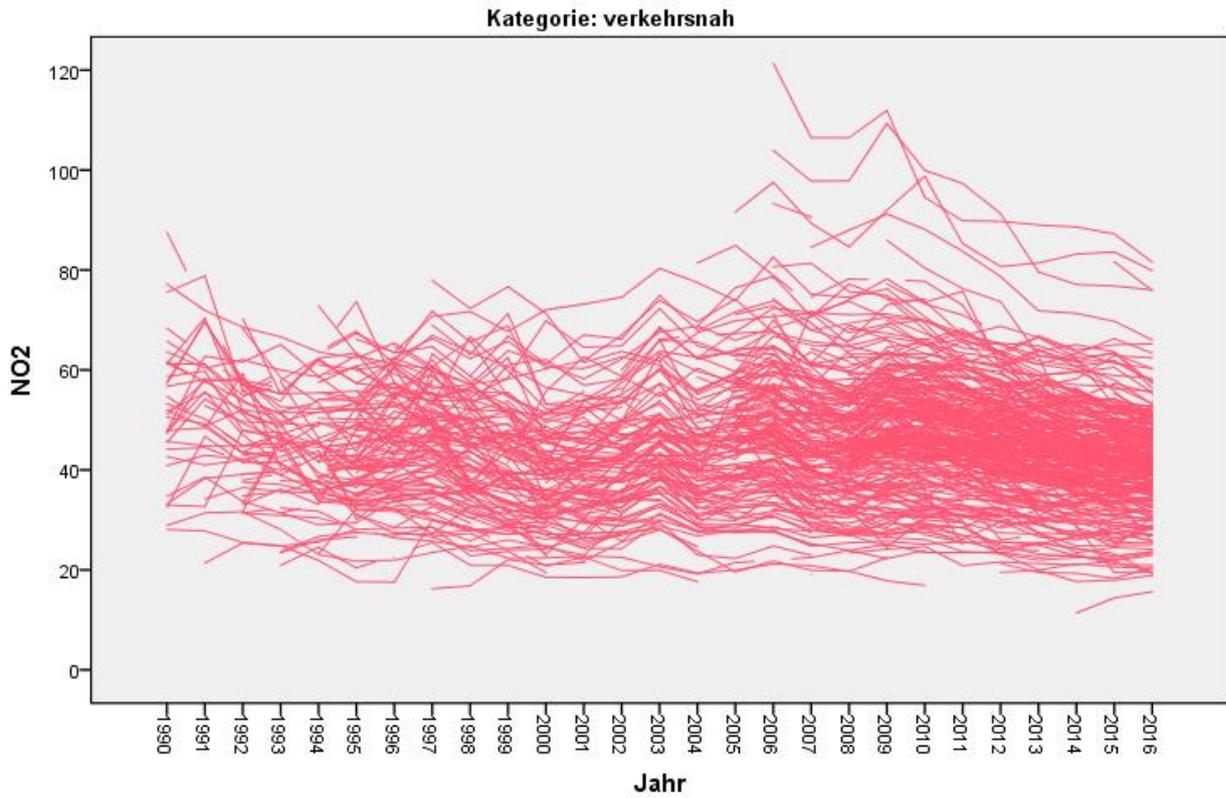
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



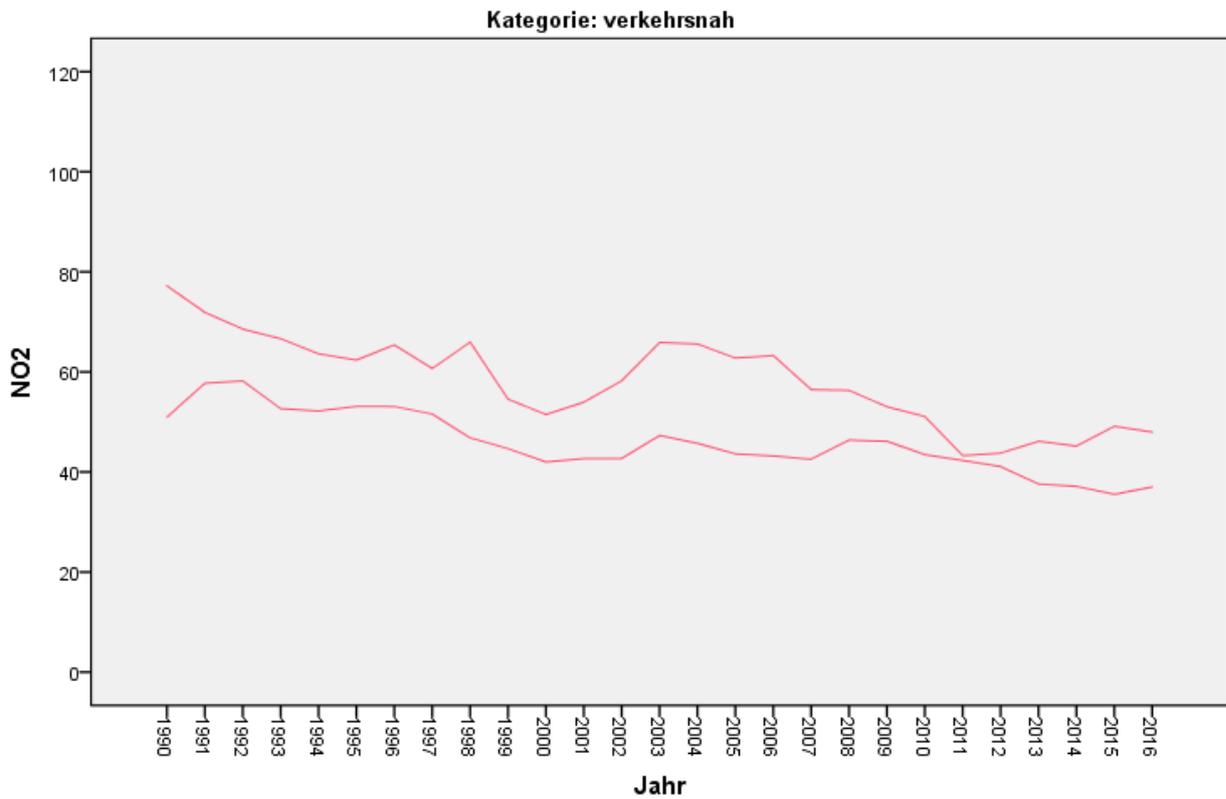
alle Messreihen
eigene Darstellung, BioMath GmbH



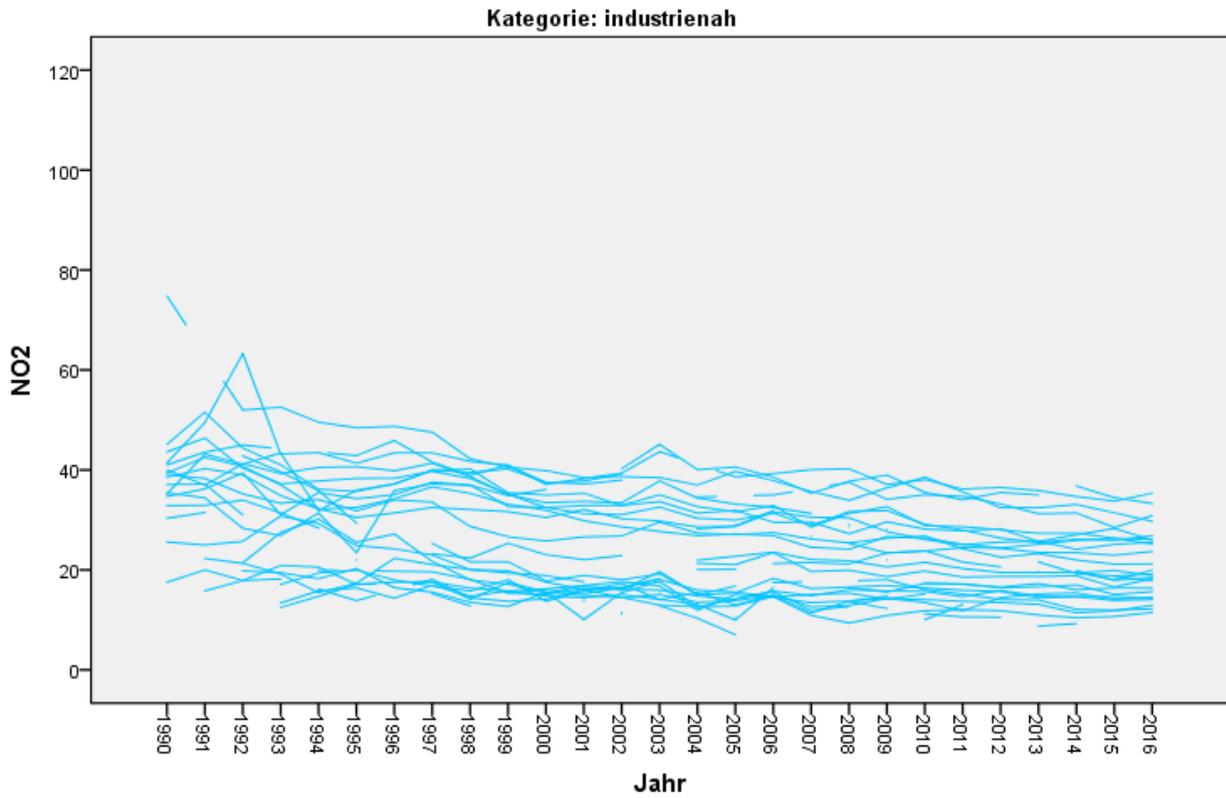
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



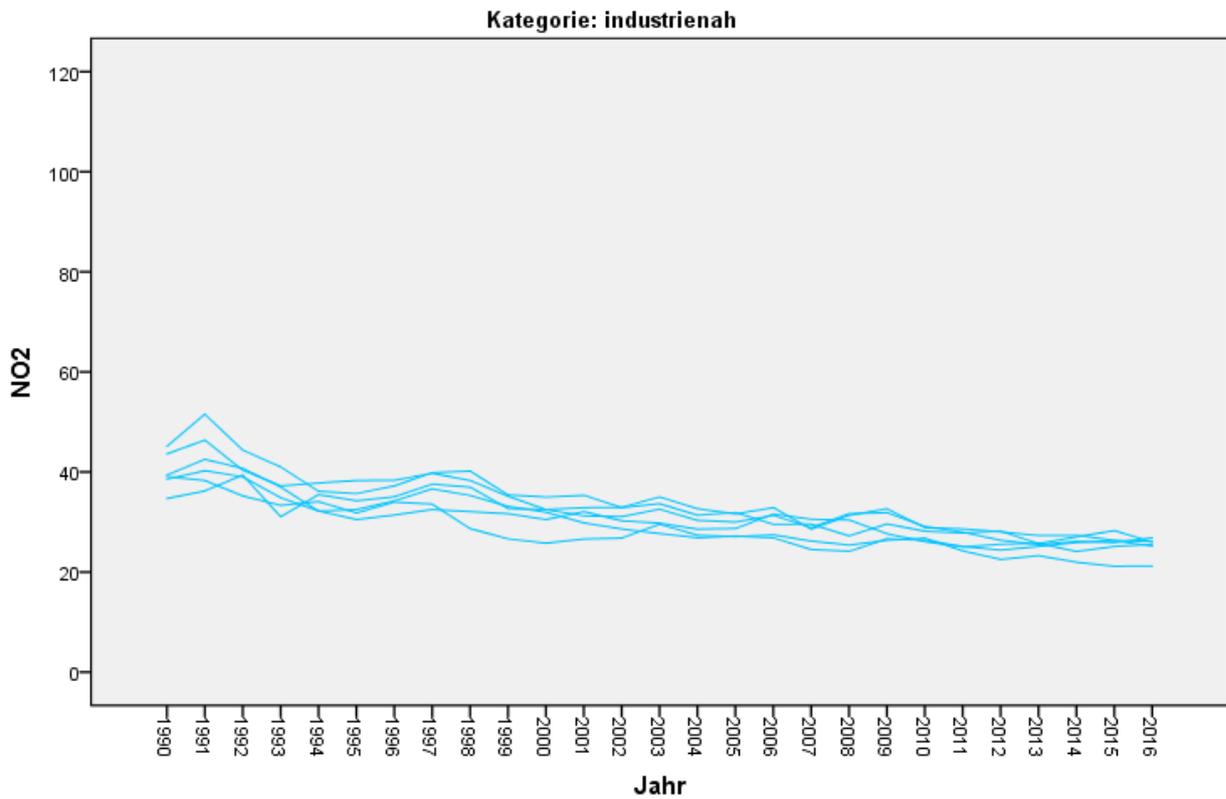
alle Messreihen
eigene Darstellung, BioMath GmbH



durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

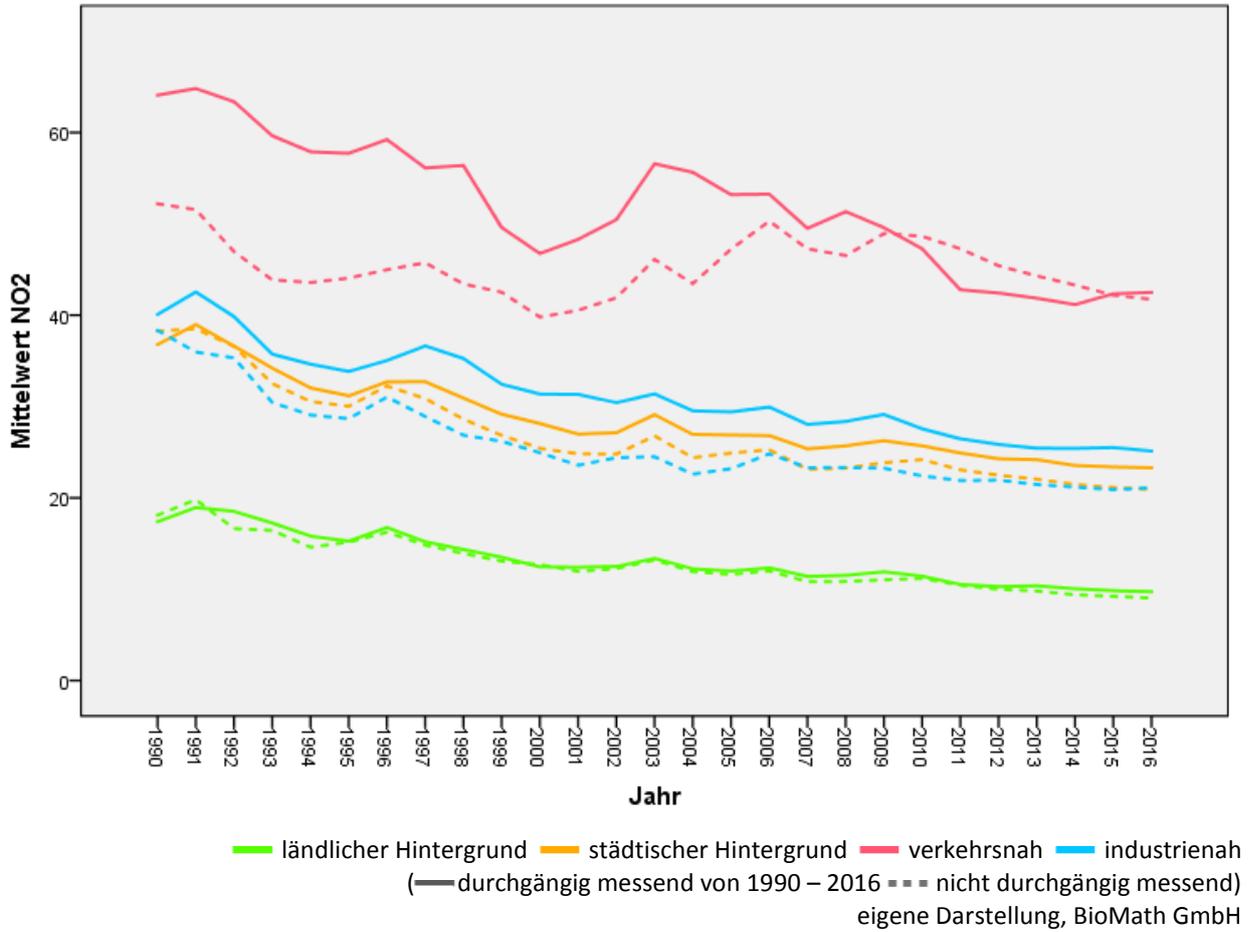


alle Messreihen
eigene Darstellung, BioMath GmbH

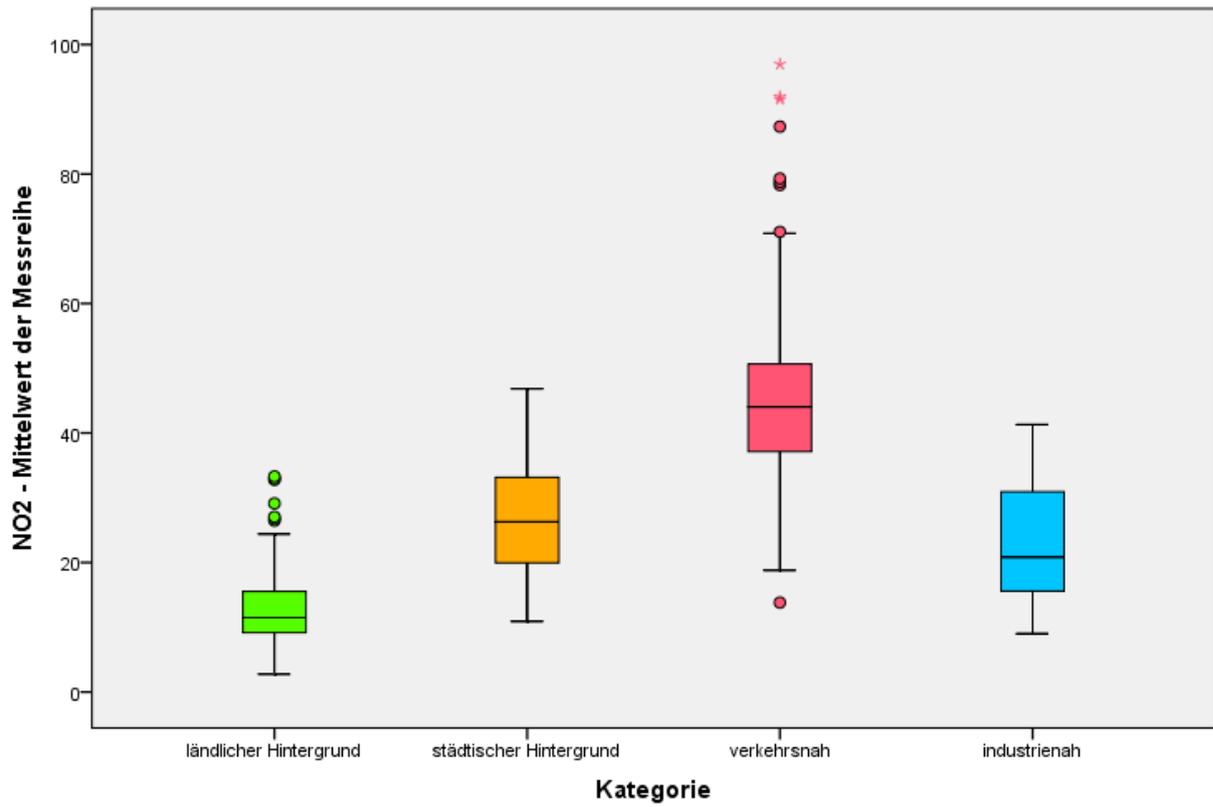


durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

3.1.2.10 NO₂ - Messreihen der jährlichen Mittelwerte in µg/m³

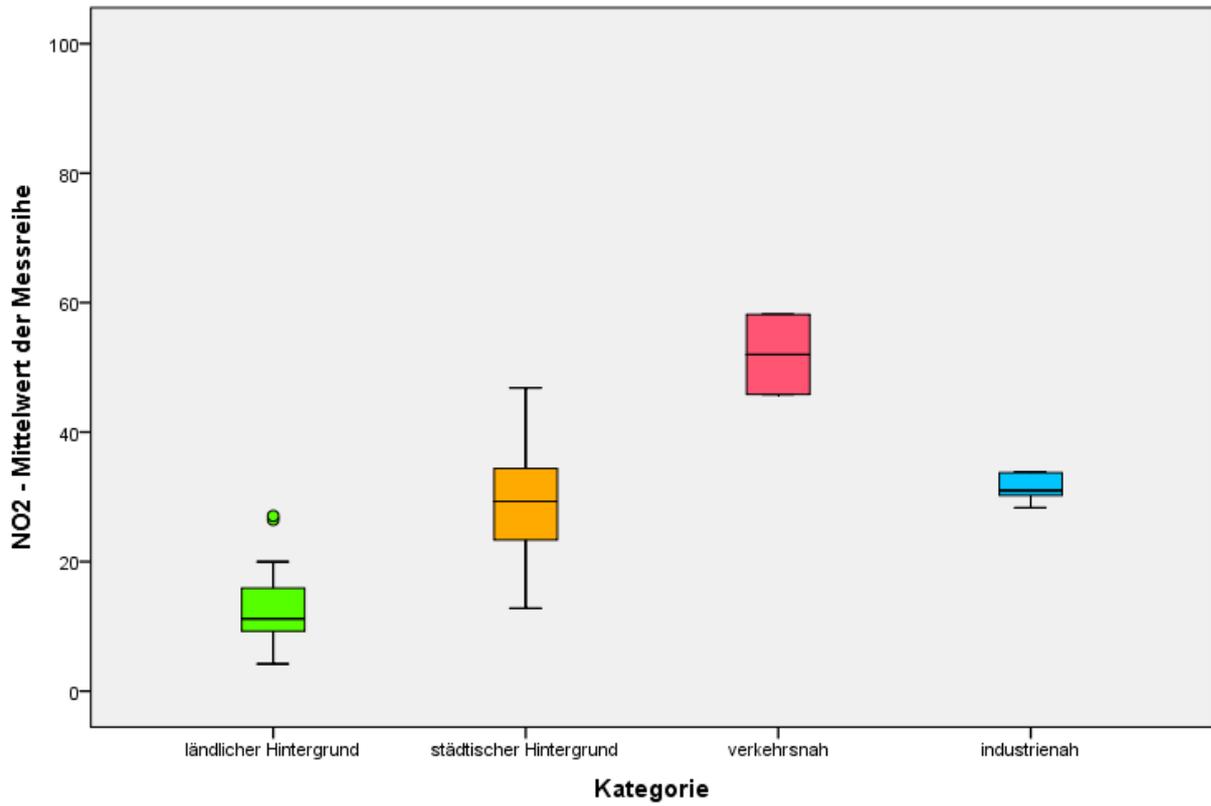


3.1.2.11 NO₂ - Boxplot der Mittelwerte in $\mu\text{g}/\text{m}^3$ je Messreihe je Kategorie



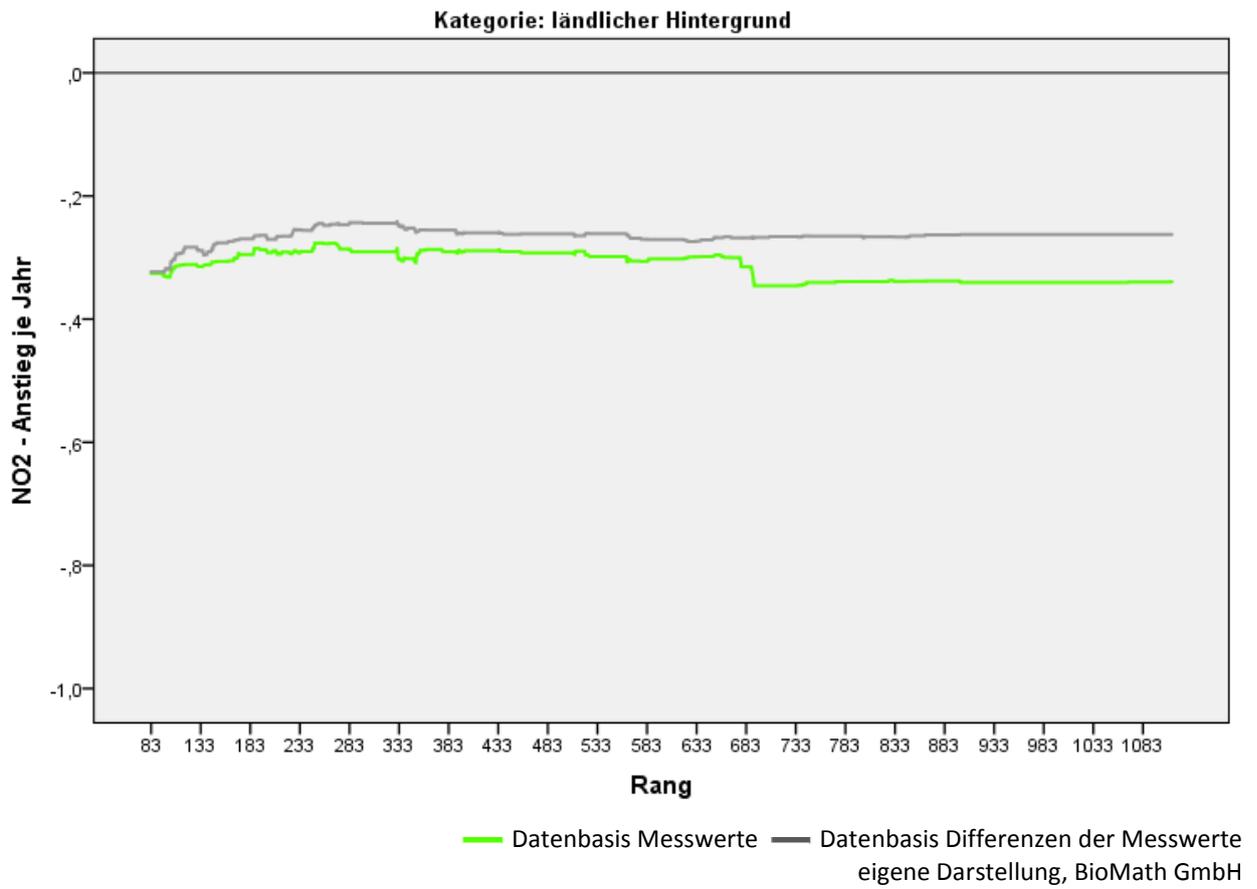
>

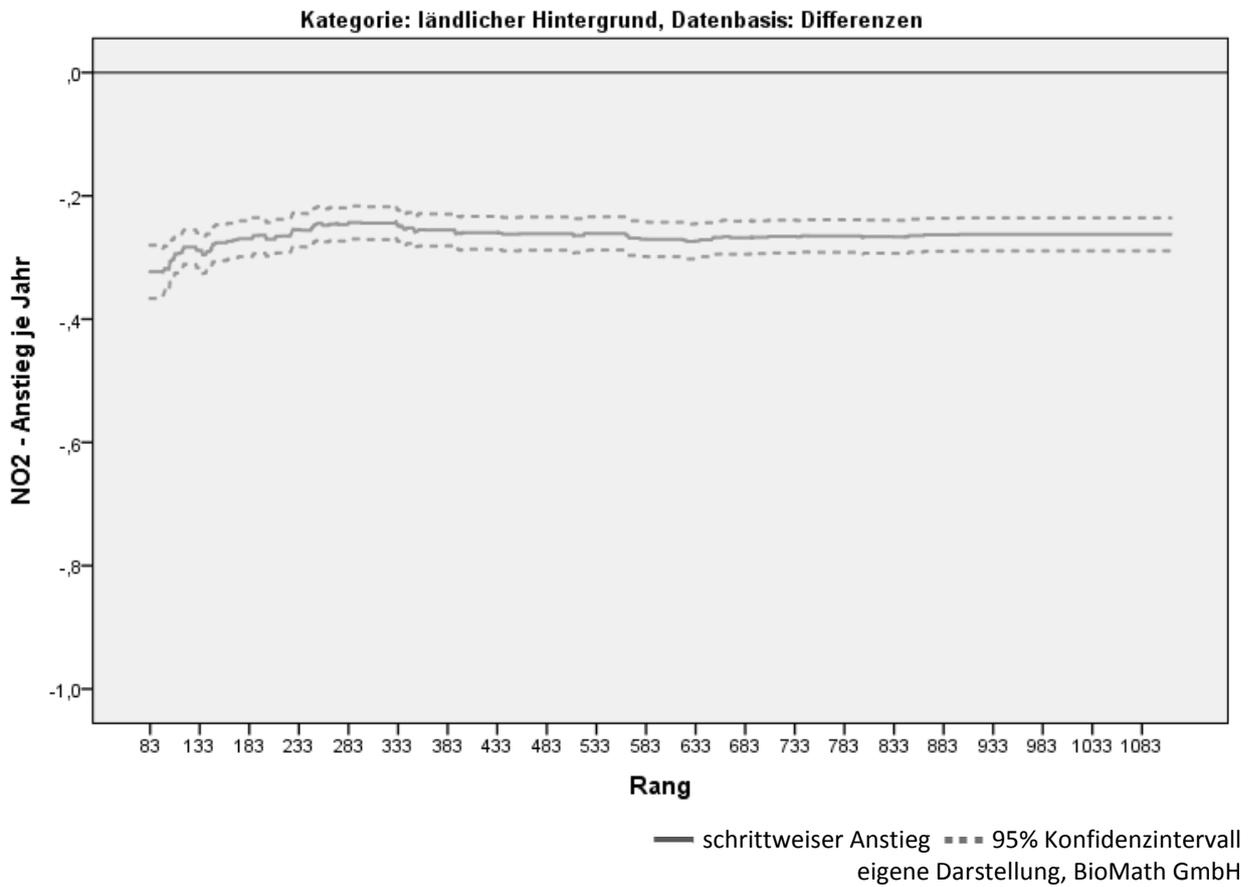
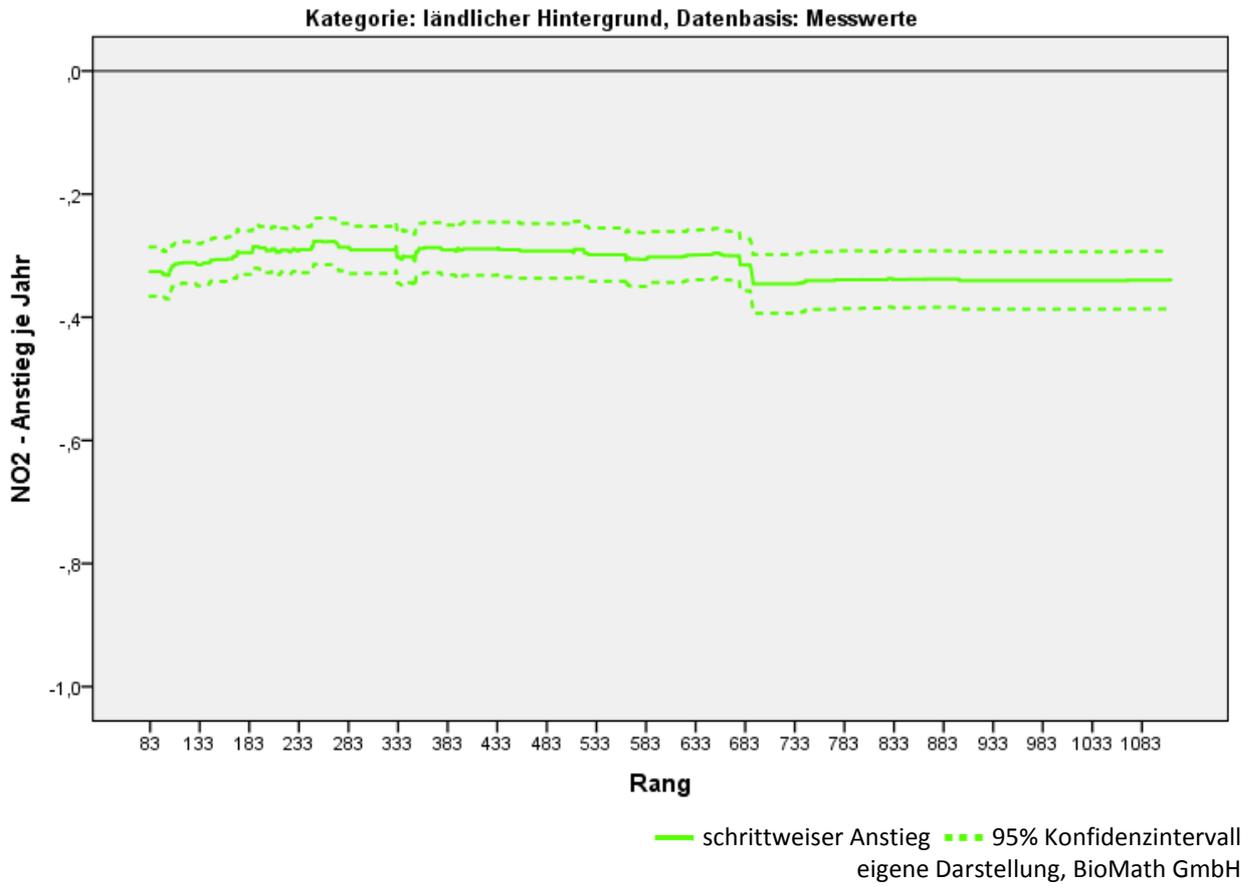
alle Messreihen
eigene Darstellung, BioMath GmbH



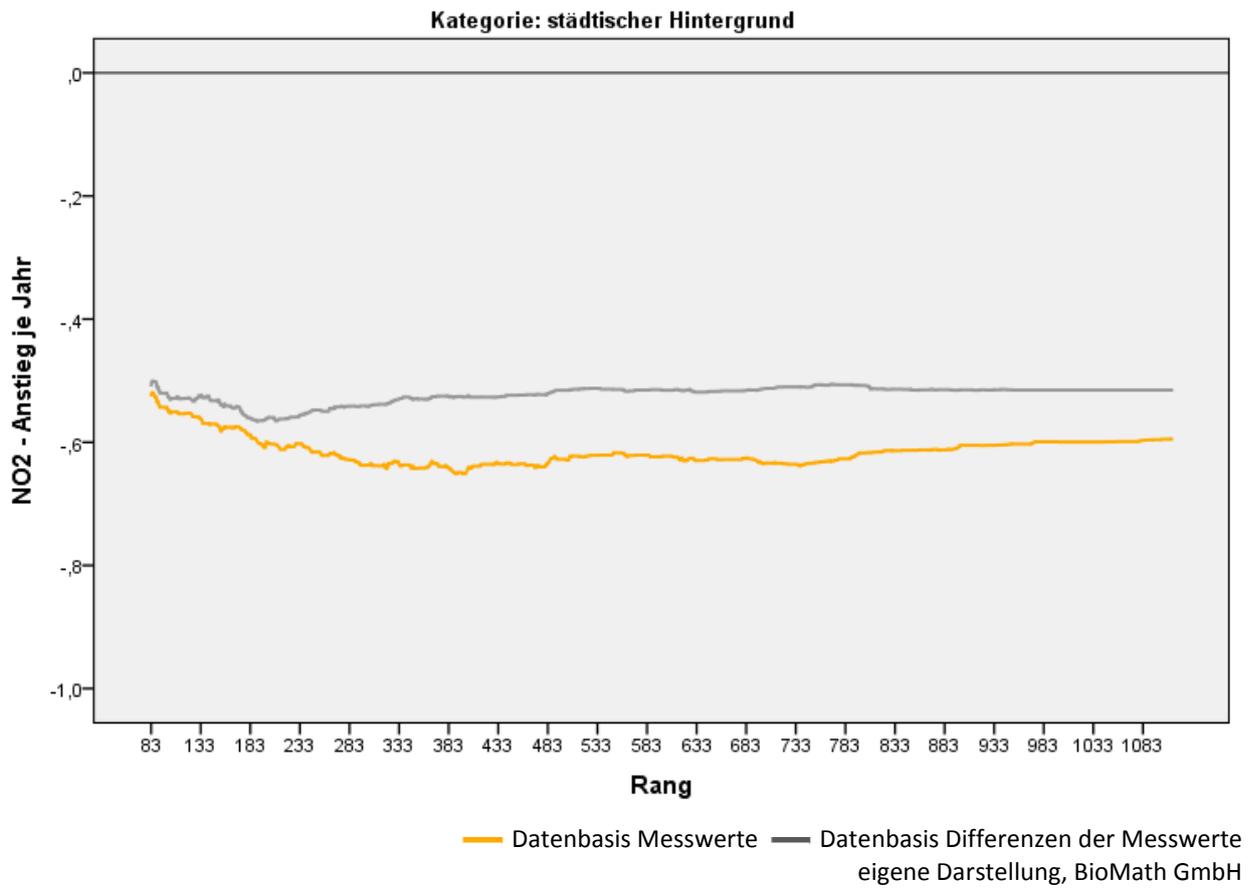
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

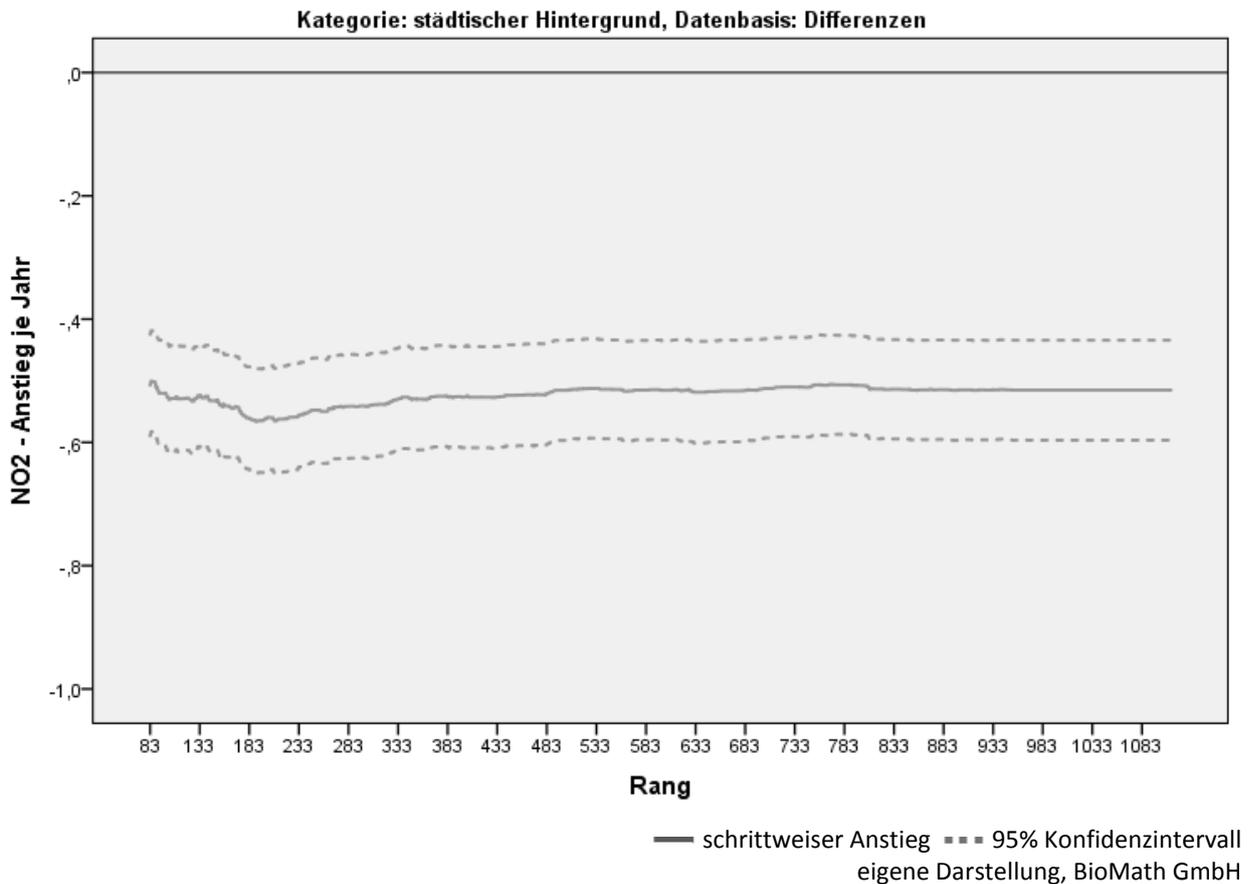
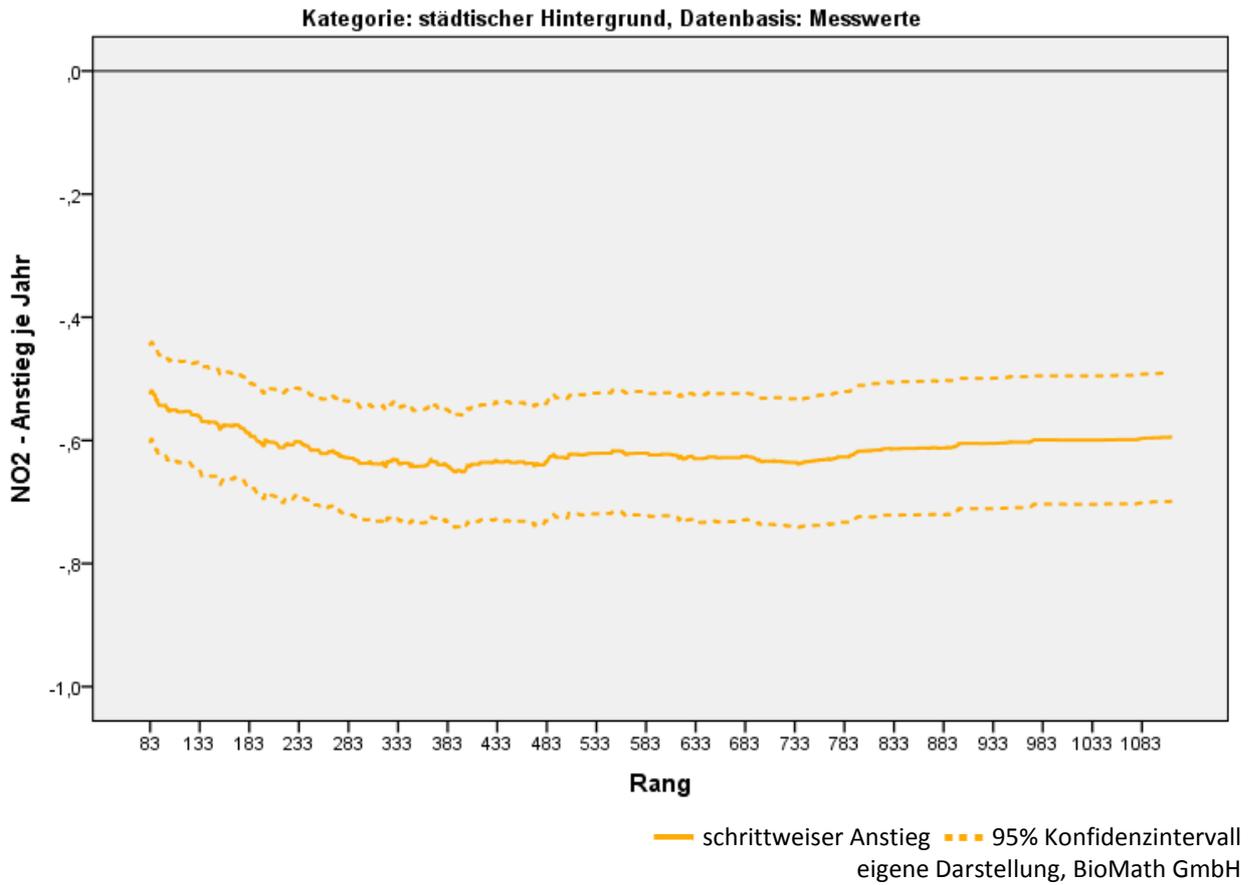
3.1.2.12 NO₂ - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund



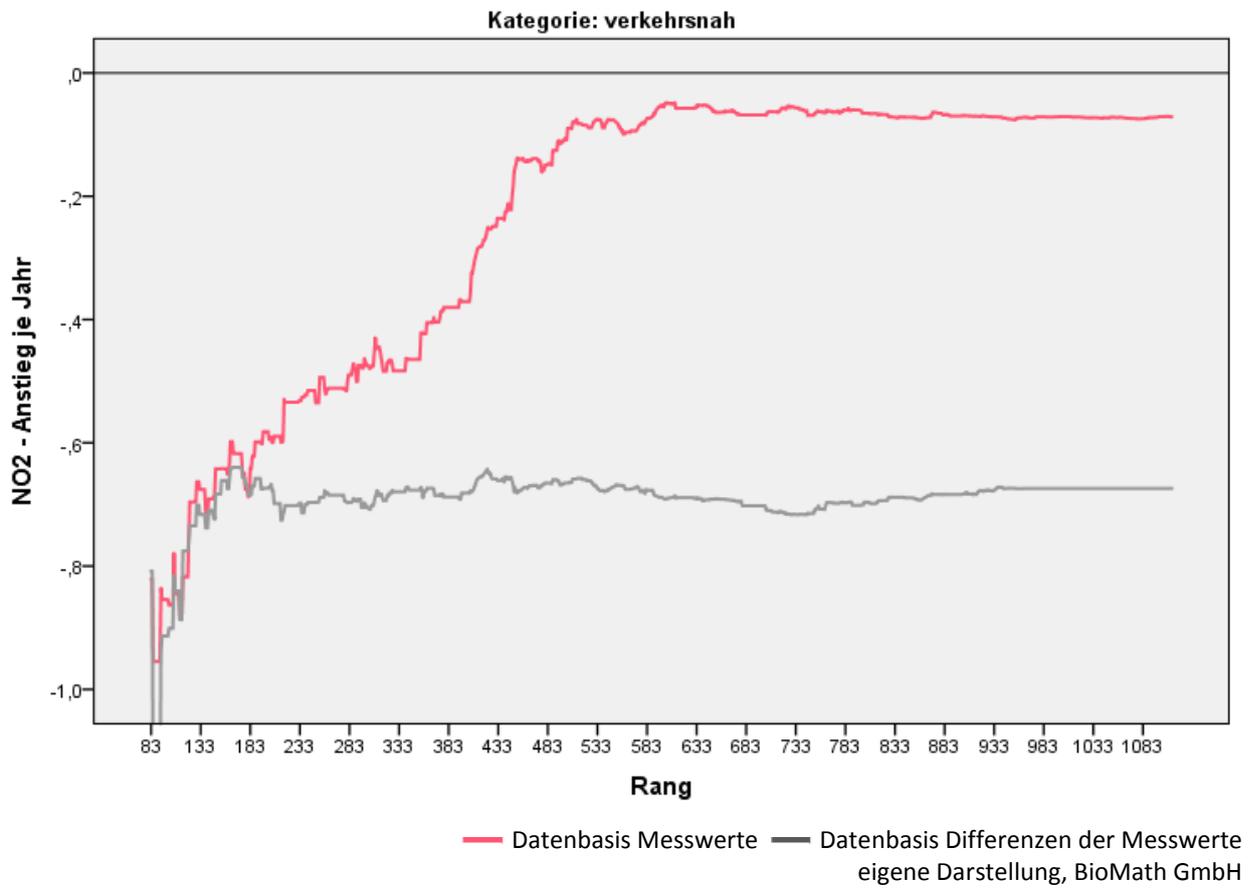


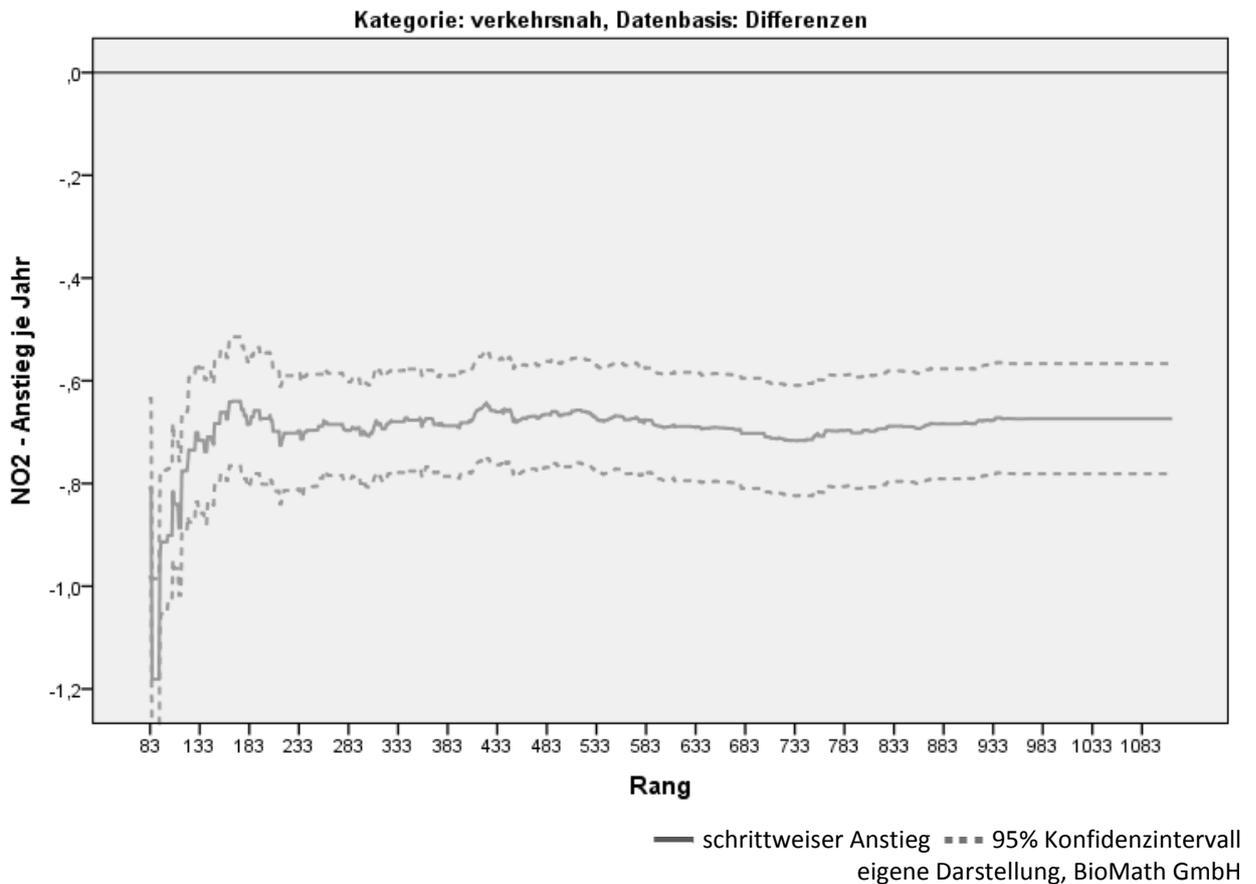
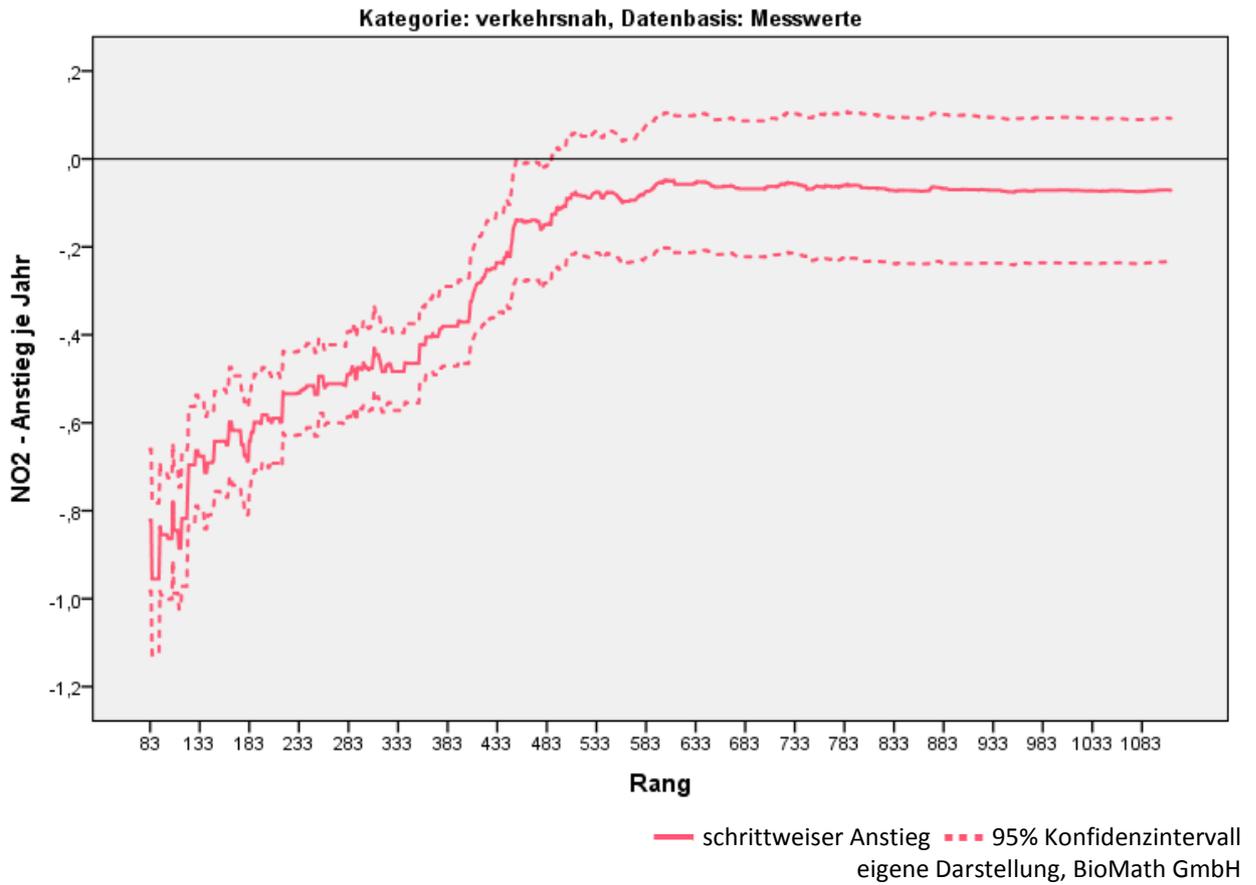
3.1.2.13 NO2 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund



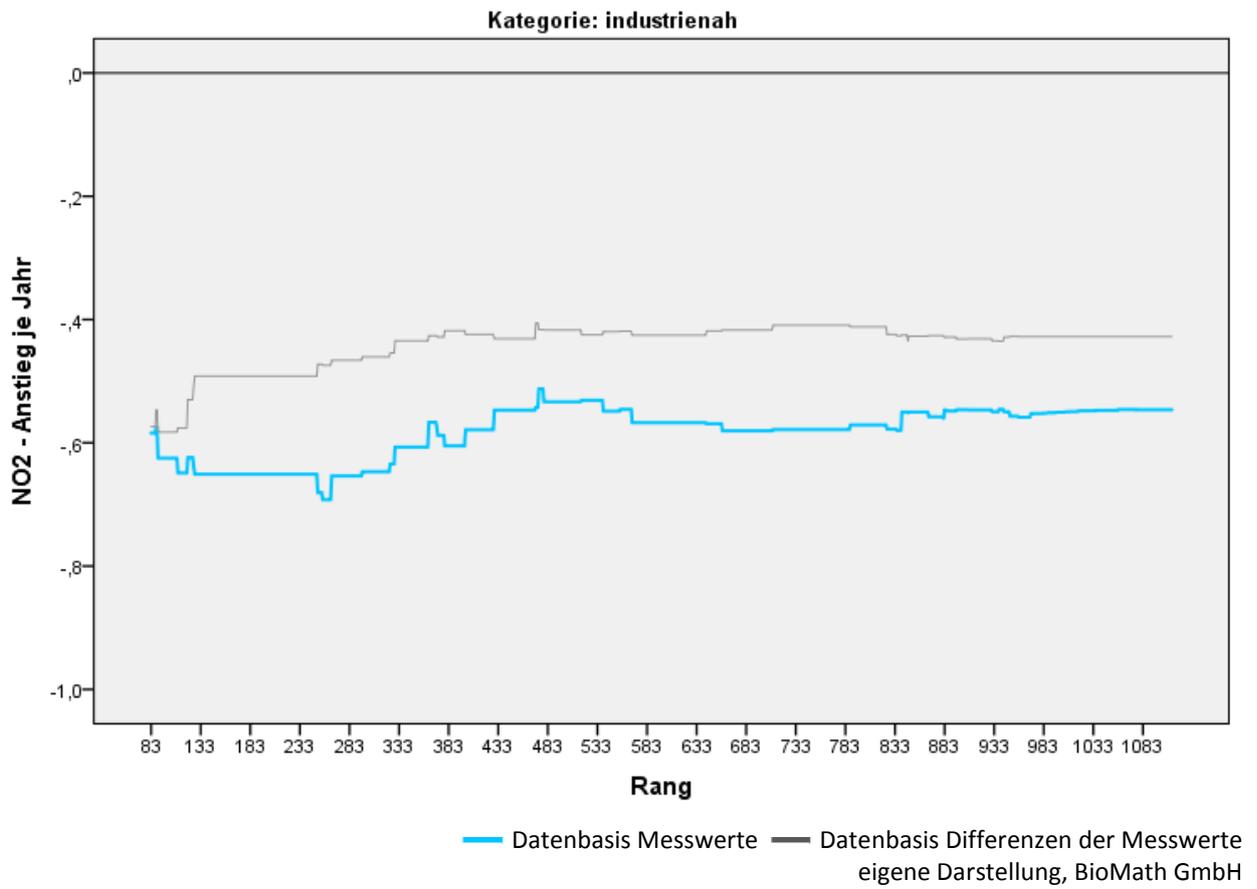


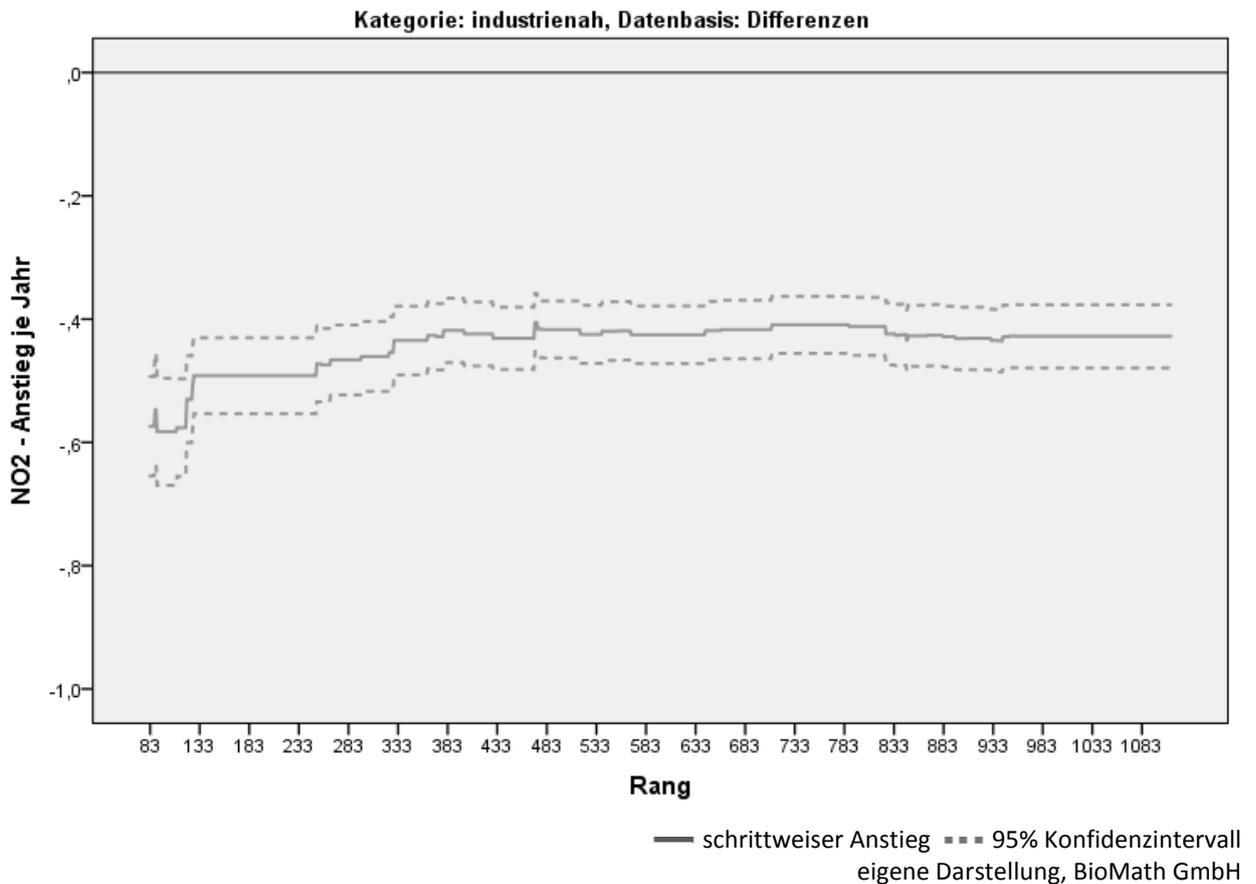
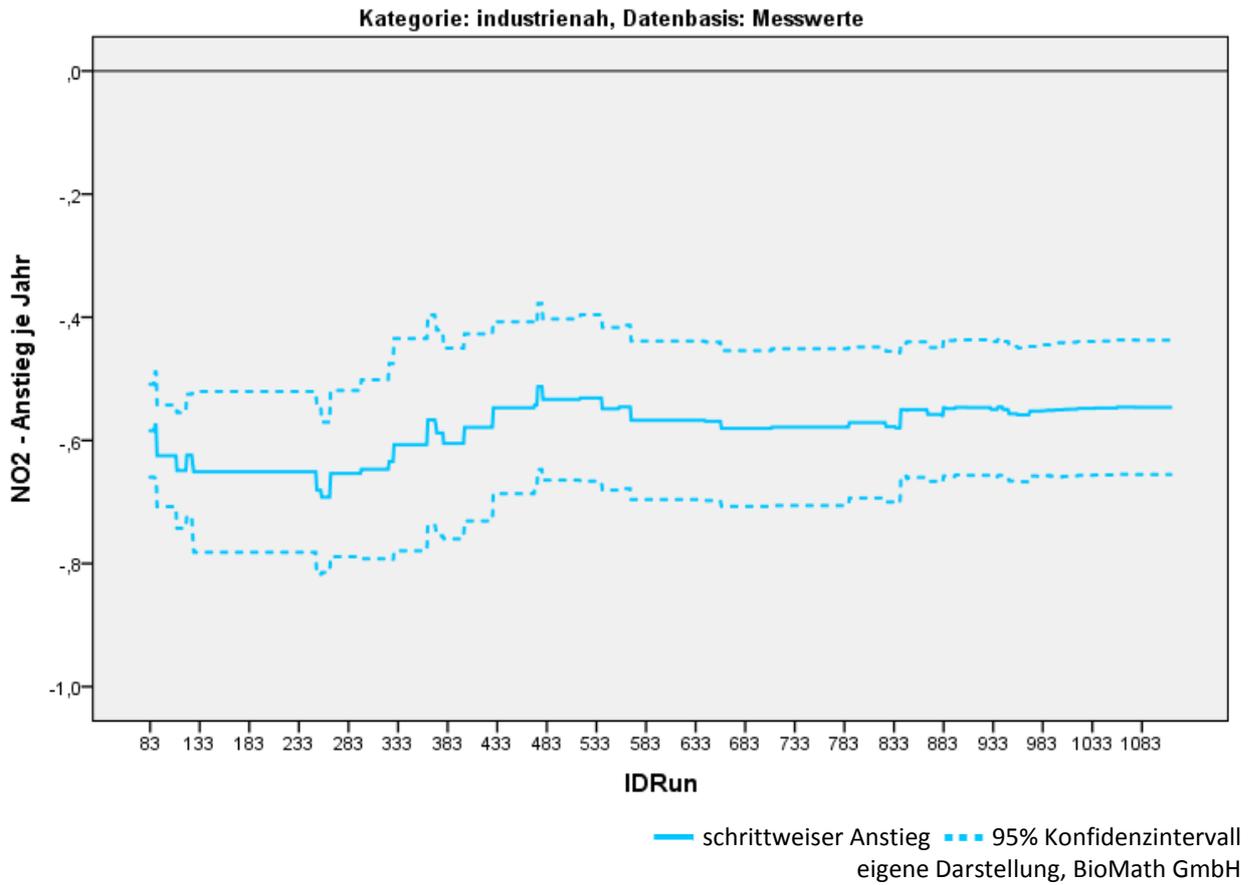
3.1.2.14 NO₂ - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah



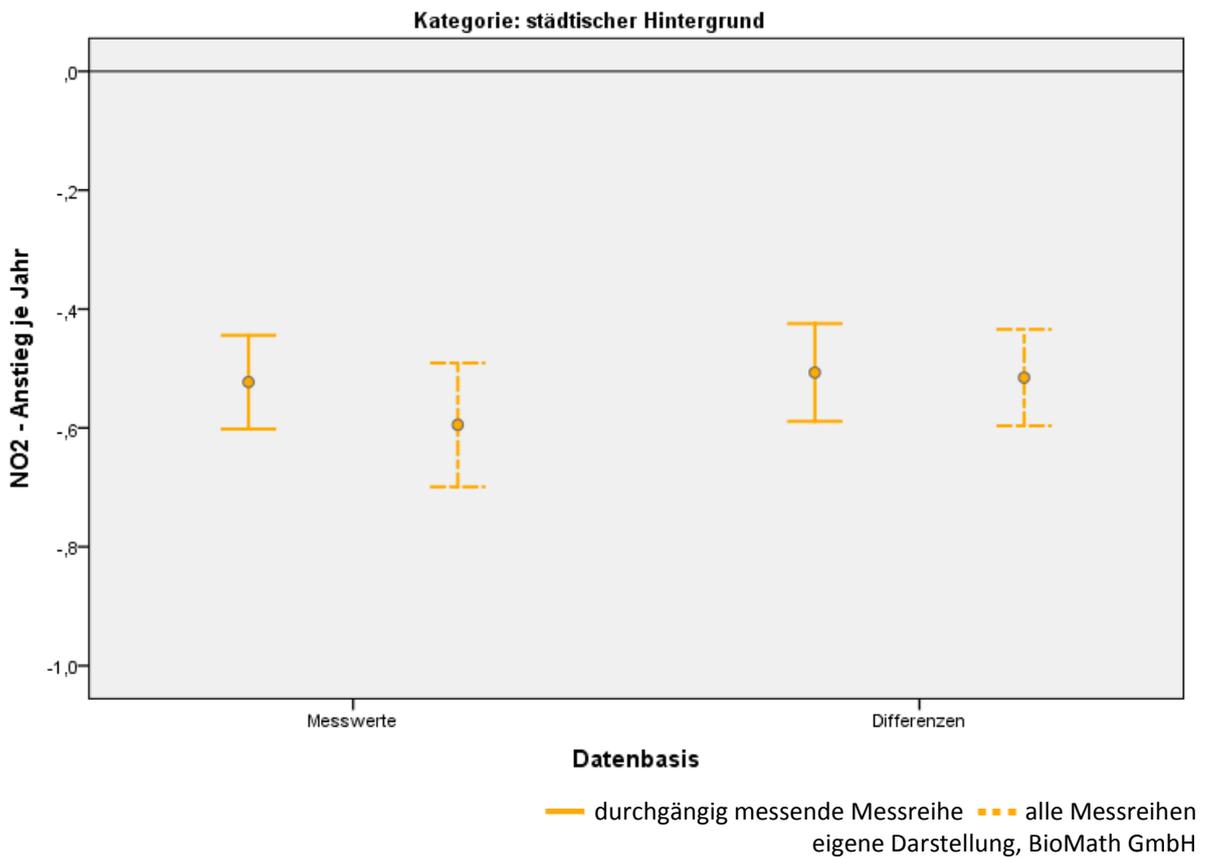


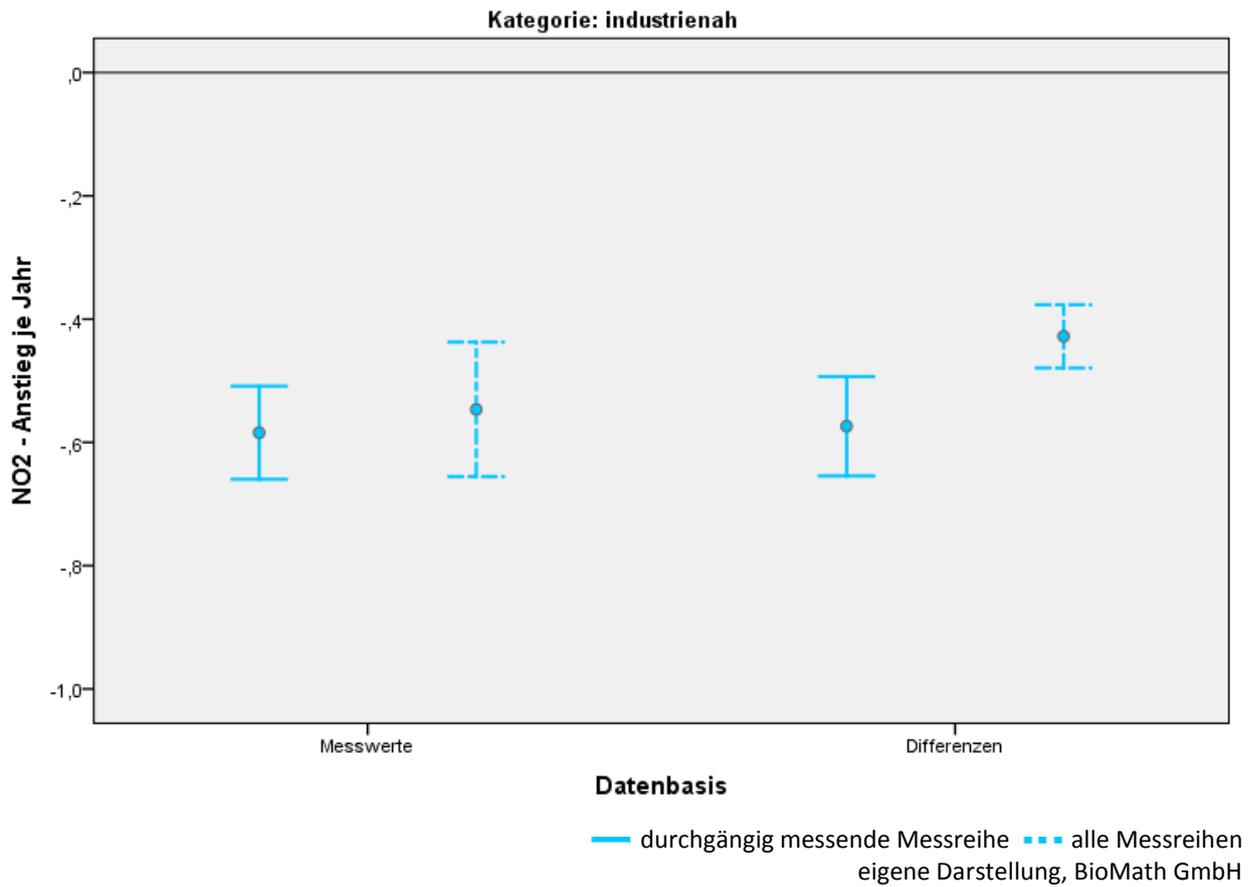
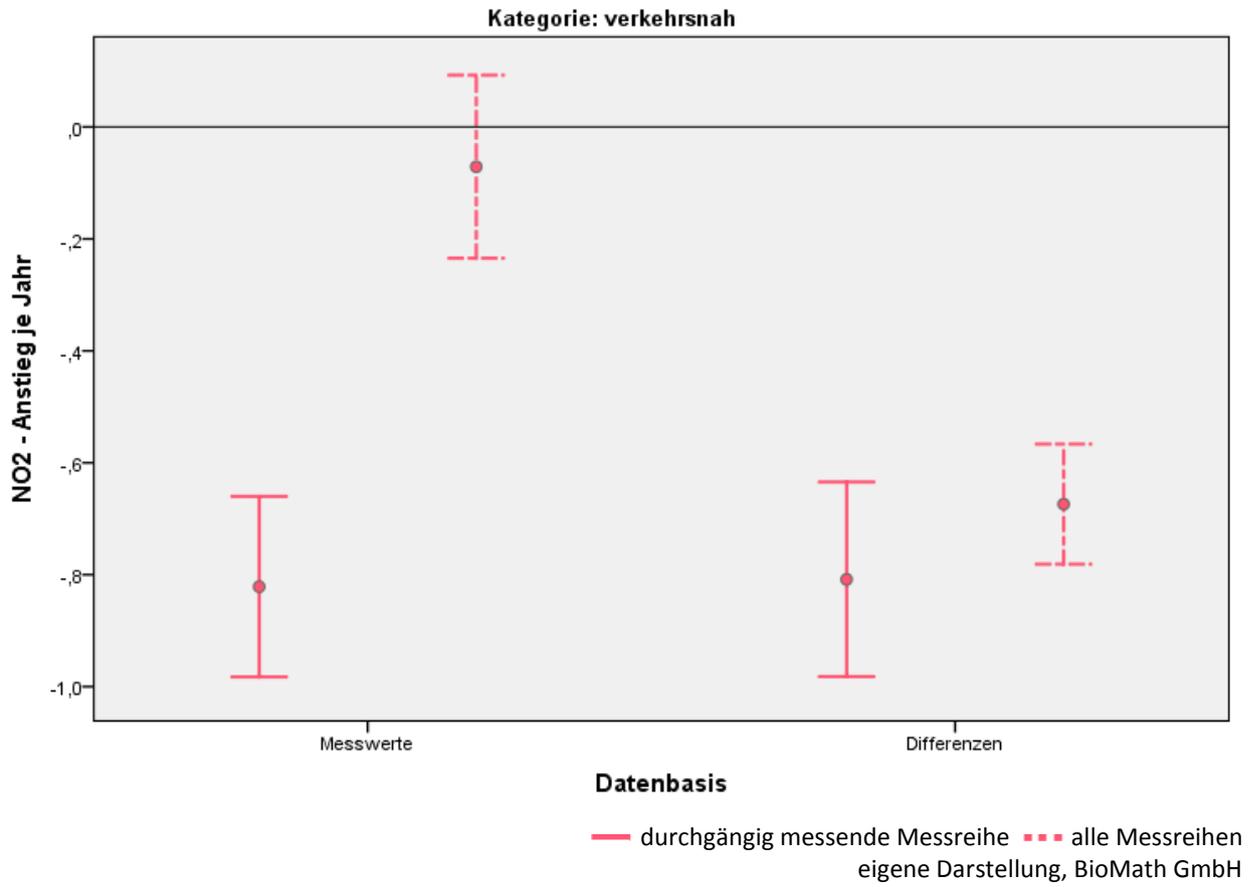
3.1.2.15 NO₂ - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah





3.1.2.16 NO₂ - Trend: Anstieg der Messreihen ($\mu\text{g}/\text{m}^3$ je Jahr) durchgängig vs. nicht durchgängig messend





3.1.3 PM10

Anzahl und Verteilung der Stationen

Der überwiegende Teil der Messstationen gehört mit 40,5 % zum städtischen Hintergrund, gefolgt von der Kategorie verkehrsnah mit 35,1 %. Die Kategorien ländlicher Hintergrund mit 14,0 % und industrienah mit 10,4 % sind deutlich geringer vertreten. Betrachtet man nur die durchgängig messenden Stationen von 1990 - 2016 so sind es gerade mal 29 von 916 Stationen, d.h. 3,2 % (siehe Abbildung 3 und 3.1.3.1).

Von 1990 bis 1995 steigt die Anzahl der Stationen stark an, fällt dann kurzzeitig leicht ab und hält sich bis 2003 auf einem einheitlichen hohen Niveau von etwa 370 Stationen. Ab 2004 kommt es erneut zu einem Anstieg, der ab 2011 leicht abfällt. Dieser Verlauf ist auch im ländlichen Hintergrund und in der Kategorie industrienah zu verzeichnen. Im städtischen Hintergrund kommt es nach 1995 zu einer kontinuierlichen Abnahme der Anzahl der Stationen. In der Kategorie verkehrsnah steigt die Anzahl der Stationen von 2005 bis 2007 sehr stark an und danach wieder ab. (siehe 3.1.3.1).

Insgesamt zeigen die jeweils durchgängig messenden Stationen vom Jahr i ($i > 1990$) bis 2016 eine kontinuierliche Zunahme. Dieses widerspiegelt sich im städtischen und ländlichen Hintergrund sowie in der Kategorie industrienah. Die Kategorie verkehrsnah zeigt ab 2005 eine deutlich stärkere Zunahme an durchgängig messende Stationen bis 2016. Im städtischen Hintergrund gibt es im Zeitraum von 1990 bis 2000 einen überproportional hohen Anteil an nicht durchgängig messenden Stationen (siehe 3.1.3.2 und 3.1.3.3).

Im Jahr 1990 gab es äußerst wenige durchgängig messende Stationen bis 2016 und das überwiegend in Nordrheinwestfalen, Hessen und Berlin. Bis zum Jahre 1995 nahm die Anzahl der durchgängig messenden Stationen deutlich zu. Ab 2000 gibt es in gesamt Deutschland eine flächendeckende Anzahl von durchgängig messenden Stationen (siehe 3.1.3.4 bis 3.1.3.7).

Inbetriebnahme und Außerbetriebnahme von Stationen sind heterogen über den Zeitabschnitt 1990 - 2016 verteilt. Das betrifft alle Kategorien. Die Betriebszeiten schwanken von 1 Jahr bis zu 27 Jahren und sind im Mittel etwa 11 Jahre. Es waren 10 % der Stationen lediglich nur 1 Jahr in Betrieb. Fast 70 % der Stationen lieferten in jedem Betriebsjahr einen Wert. Bei knapp 13 % der Stationen fehlte ein Wert während der Betriebszeit. Würde man diesen einen fehlenden Wert auffüllen, so stellen etwa 83 % der Stationen vollständige Messreihen in ihren Betriebszeiten bereit (siehe 3.1.3.8).

Verlauf der Werte der Messreihen

Die Werte der Messreihen betragen im Mittel $26,4 \mu\text{g}/\text{m}^3$ und schwanken zwischen $5,4$ und $103 \mu\text{g}/\text{m}^3$. Insgesamt kann festgestellt werden, dass sich der Verlauf der Messreihen von 1990 bis etwa 2003 von dem Verlauf der Jahre 2003 bis 2016 unterscheiden. Die Variabilität zwischen den Messreihen bis 2003 ist deutlich höher als zwischen den Messreihen von 2003 bis 2016. Die Ursache hierfür könnte das Messverfahren von PM10 sein. Denn von 1990 bis etwa 2000 wurde im Wesentlichen der Gesamtschwebstaub gemessen und dieser dann in PM10 umgerechnet (siehe 3.1.3.9). Messreihen mit hohen Abweichungen zum mittleren Niveau sind im Boxplot gut erkennbar – in der Kategorie verkehrsnah sind die Abweichungen am stärksten ausgeprägt (siehe 3.1.3.11).

Die mittleren Messreihen der durchgängig und nicht durchgängig messenden Reihen unterscheiden sich deutlich in den Kategorien städtischer Hintergrund, verkehrsnah und industrienah. Im städtischen Hintergrund und in der Kategorie industrienah sind die Mittelwerte der durchgängig messenden Reihen immer höher als die der nicht durchgängig messenden Reihen. Die Kategorie verkehrsnah zeigt diesbezüglich ein gegenläufiges Verhalten. Im ländlichen Hintergrund gibt es ab 1995 nicht so starke Abweichungen (siehe 3.1.3.10).

Trends der Messreihen

In der schrittweisen Trendberechnung gibt es in allen Kategorien Unterschiede im Verlauf der Anstiege zwischen der Datenbasis der Messwerte und der der Differenzen der Messwerte.

In der Kategorie ländlicher Hintergrund sind die Abweichungen insgesamt gering, auf den ersten Rängen etwas größer als auf folgenden Rängen. Ab Rang 571 driften die Anstiege der Datenbasis der Messwerte und die der Differenzen der Messwerte leicht auseinander. Alle Anstiege sind negativ und signifikant (siehe 3.1.3.12).

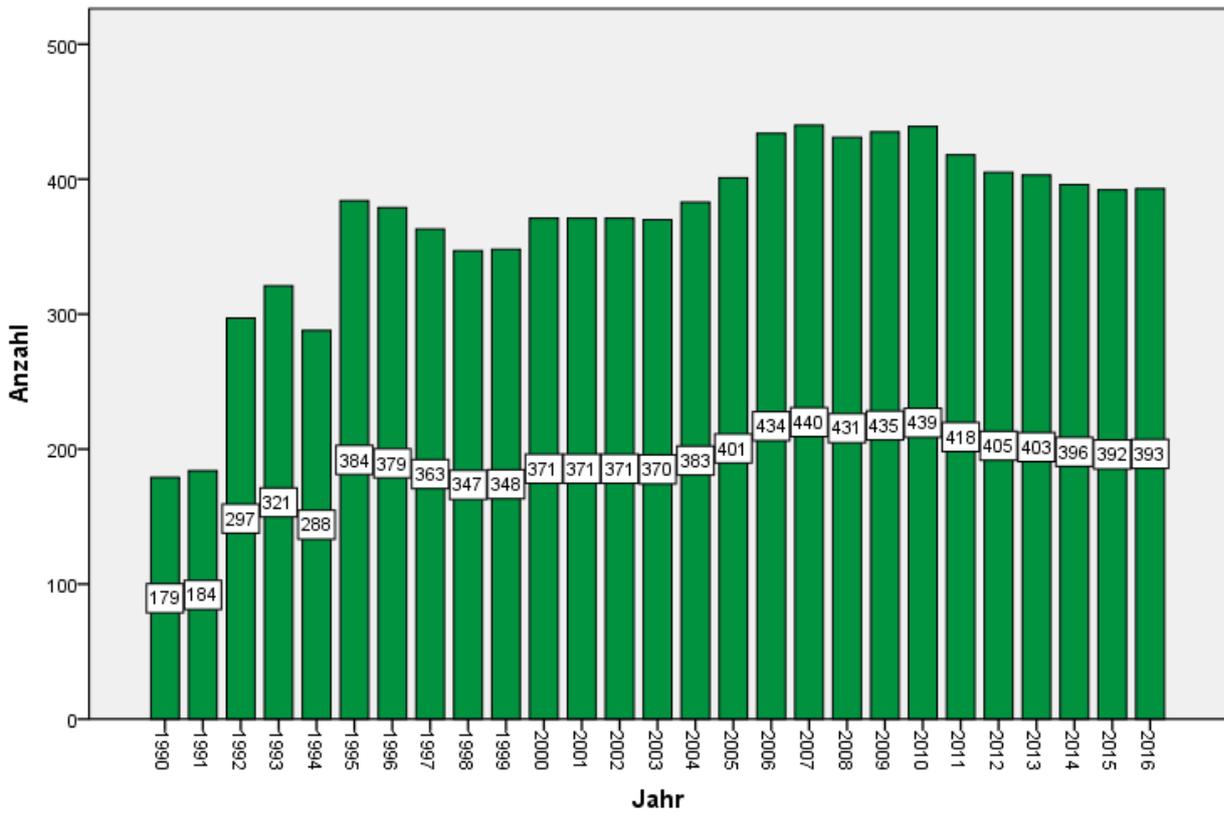
Auch im städtischen Hintergrund sind die Abweichungen zwischen den Anstiegen der Datenbasis der Messwerte und der der Differenzen der Messwerte gering. Ab Rang 191 driften die Anstiege der Datenbasis der Messwerte und die der Differenzen der Messwerte leicht auseinander. Alle Anstiege sind negativ und signifikant (siehe 3.1.3.13).

In der Kategorie verkehrsnah gibt es ab Rang 151 leichte und ab Rang 371 starke Abweichungen der Anstiege zwischen den Datenbasen. Zum Ende der schrittweisen Trendberechnung liegen die Unterschiede bei $0,36 \mu\text{g}/\text{m}^3$ je Jahr. Die Anstiege der Datenbasis der Messwerte sind ab Rang 151 immer höher als die der Datenbasis der Differenzen der Messwerte. Die Ursache der Abweichungen ist die starke Schwankung der Niveaus der hinzukommenden Messreihen. Z.B. haben ab Rang 371 die Messreihen 11 oder weniger Messwerte und schwanken im mittleren Niveau zwischen $18,6$ und $63,8 \mu\text{g}/\text{m}^3$. Alle Anstiege sind negativ und signifikant (siehe 3.1.3.14).

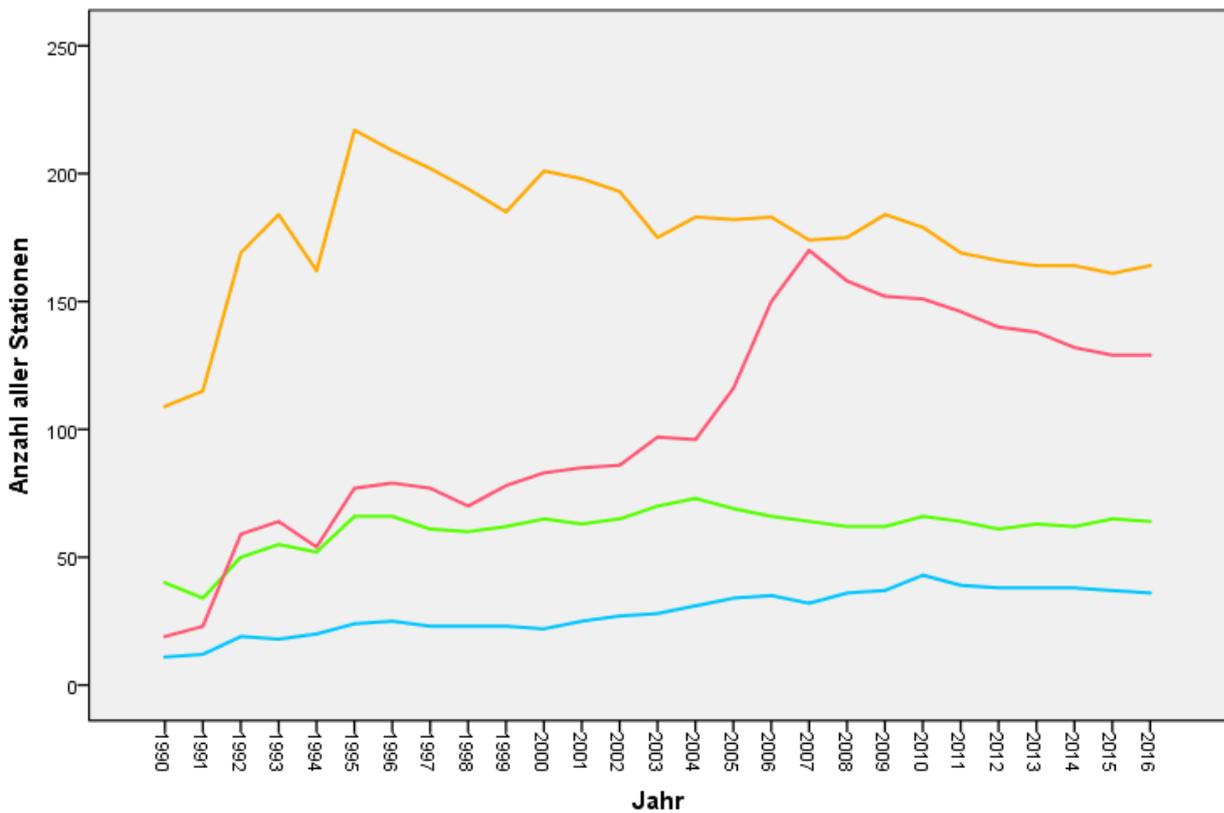
In der Kategorie industrienah sind die Anstiege der beiden Datenbasen gegenläufig. Bis zum Rang 411 ist der Anstieg der Datenbasis der Messwerte kleiner als der der Differenzen der Messwerte. Auf den folgenden Rängen bis 571 gibt es kaum Abweichungen, danach erfolgt zunächst ein leichtes, dann ab Rang 711 ein etwas stärkeres Auseinanderdriften der Anstiege in umgekehrter Richtung. Die Ursache liegt daran, dass ab Rang 711 die Messreihen nur noch 3 oder weniger Messwerte besitzen. Alle Anstiege sind negativ und signifikant (siehe 3.1.3.15).

Die Ergebnisse der Trendberechnung (Messwerte vs. Differenzen und durchgängig messend vs. nicht durchgängig messend) sind in den Fehlerbalkendiagrammen mit 95% Konfidenzintervallen zusammengefasst (siehe 3.1.3.16).

3.1.3.1 PM10 - Anzahl aller Stationen

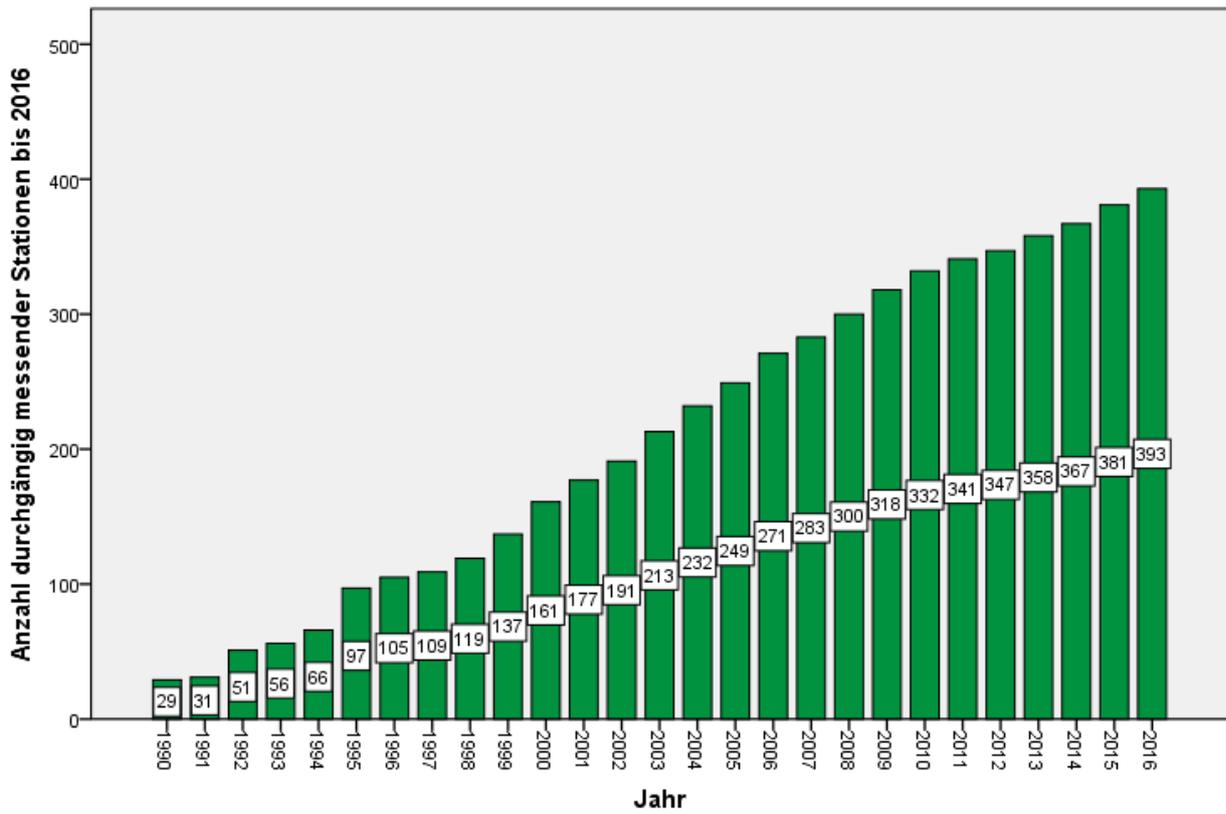


eigene Darstellung, BioMath GmbH

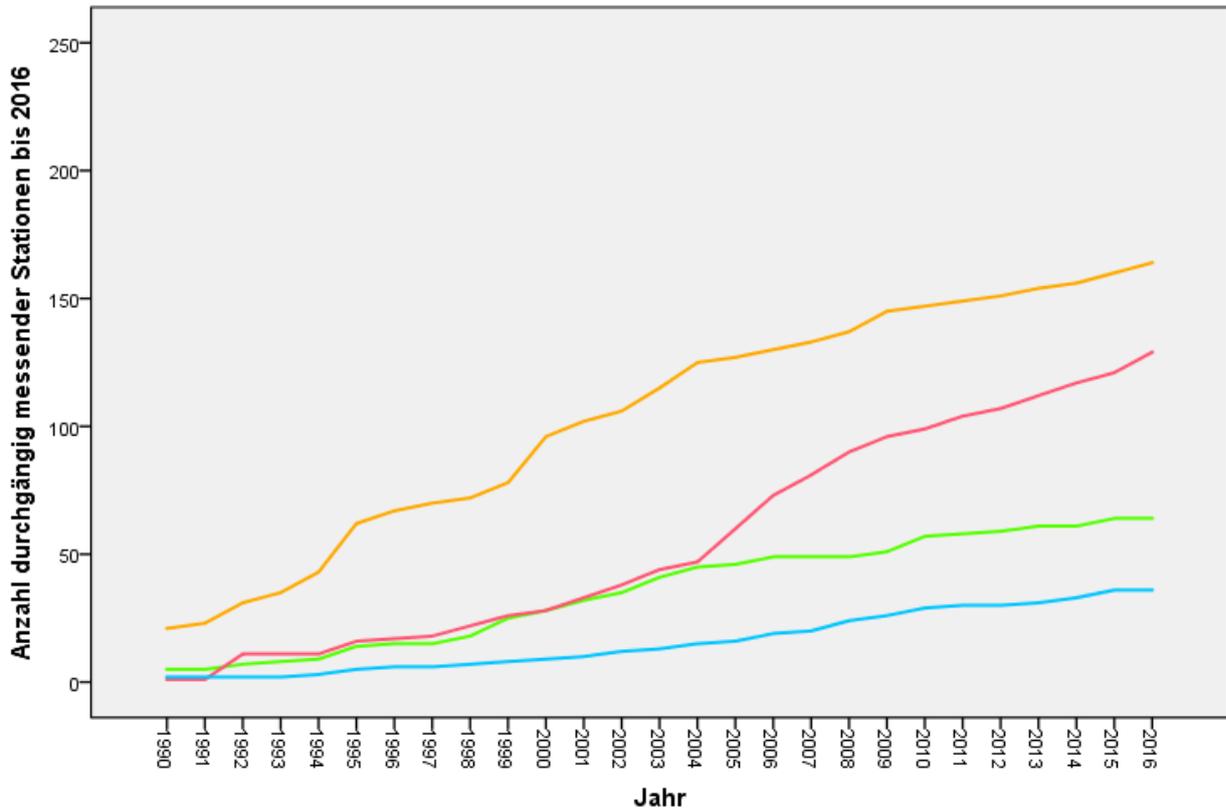


— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnah
 — industrienah
 eigene Darstellung, BioMath GmbH

3.1.3.2 PM10 - Anzahl durchgängig messender Stationen bis 2016

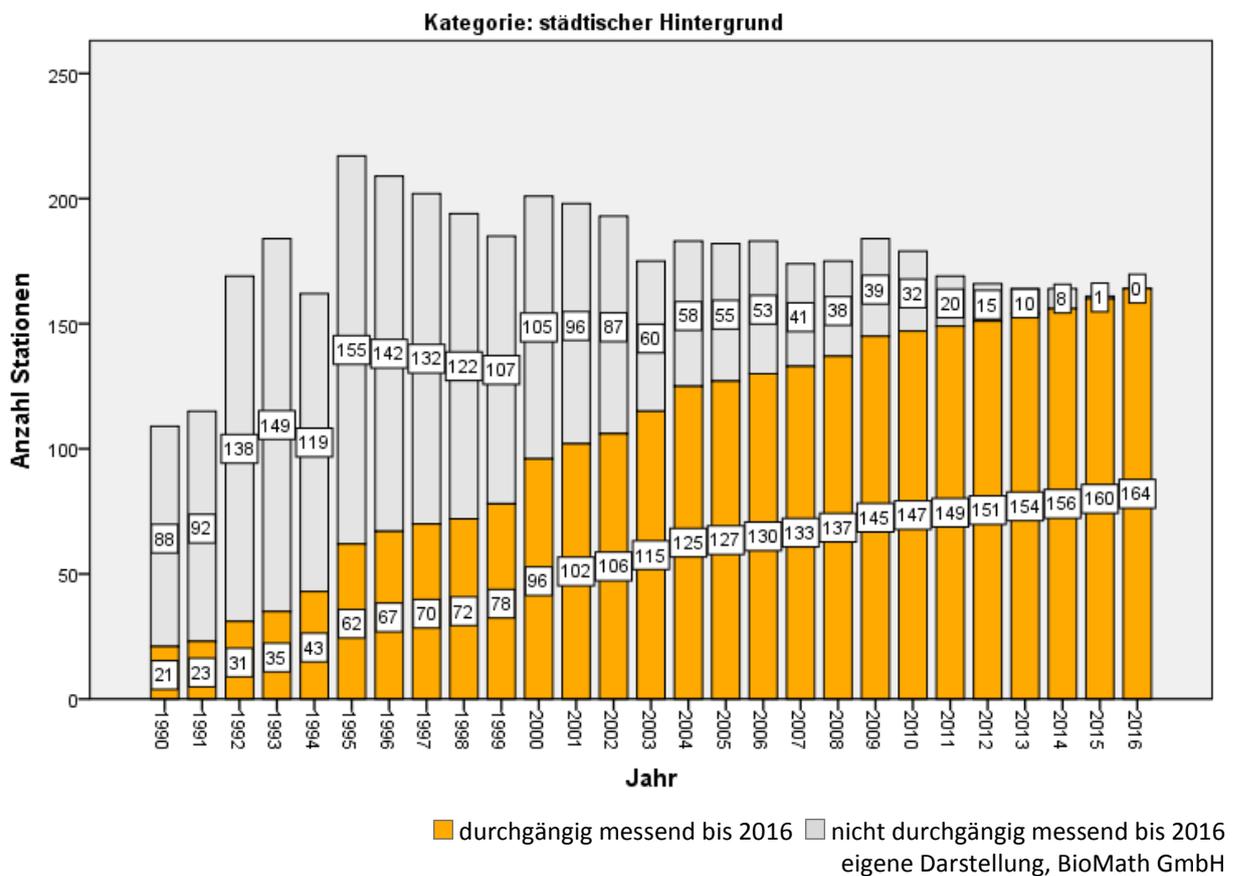
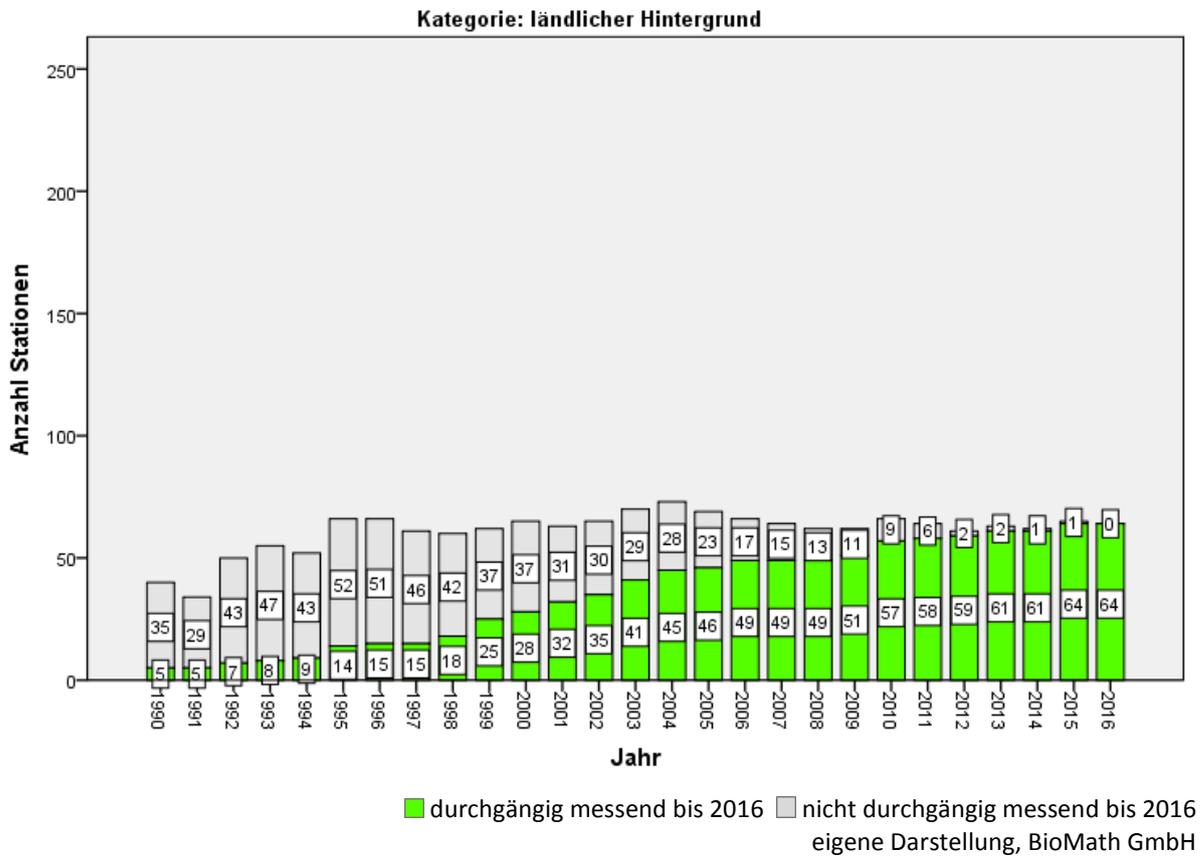


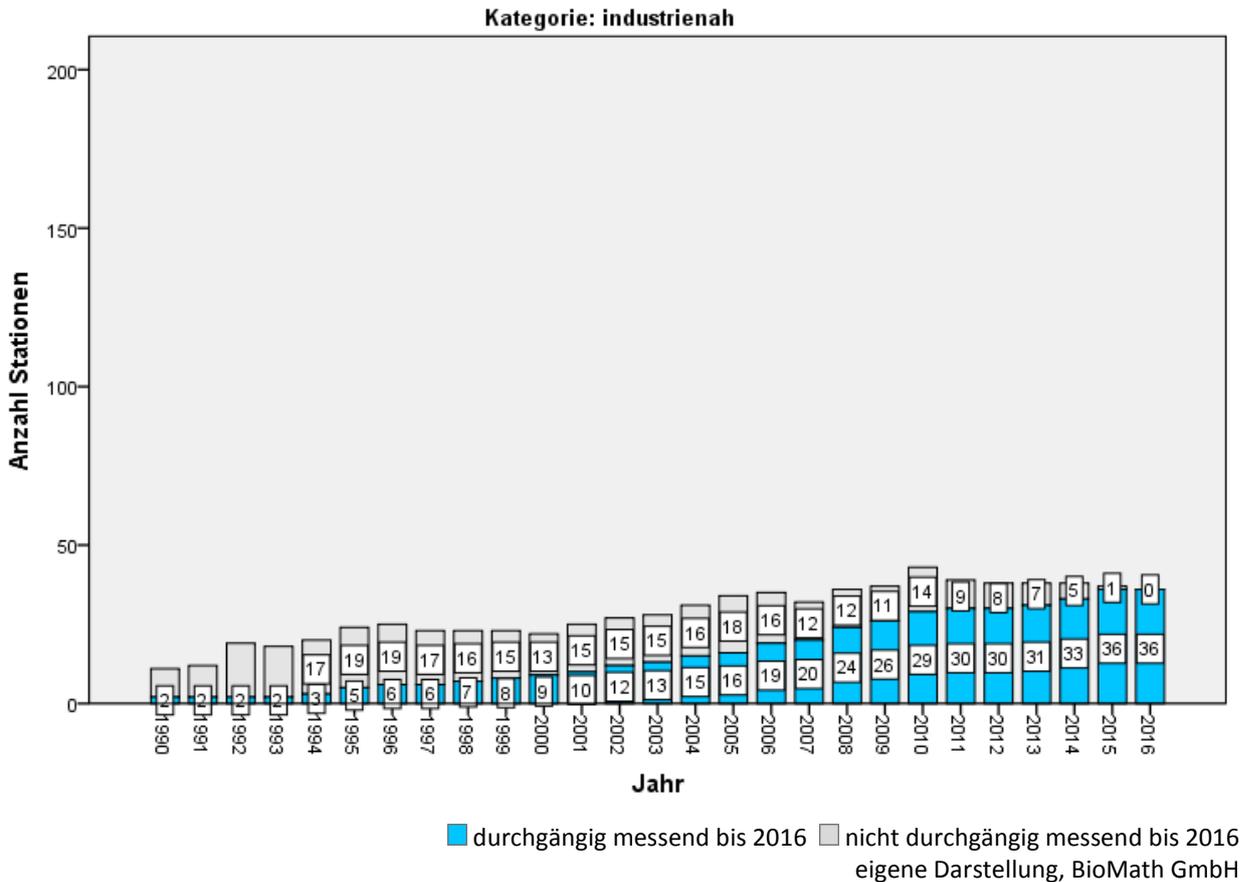
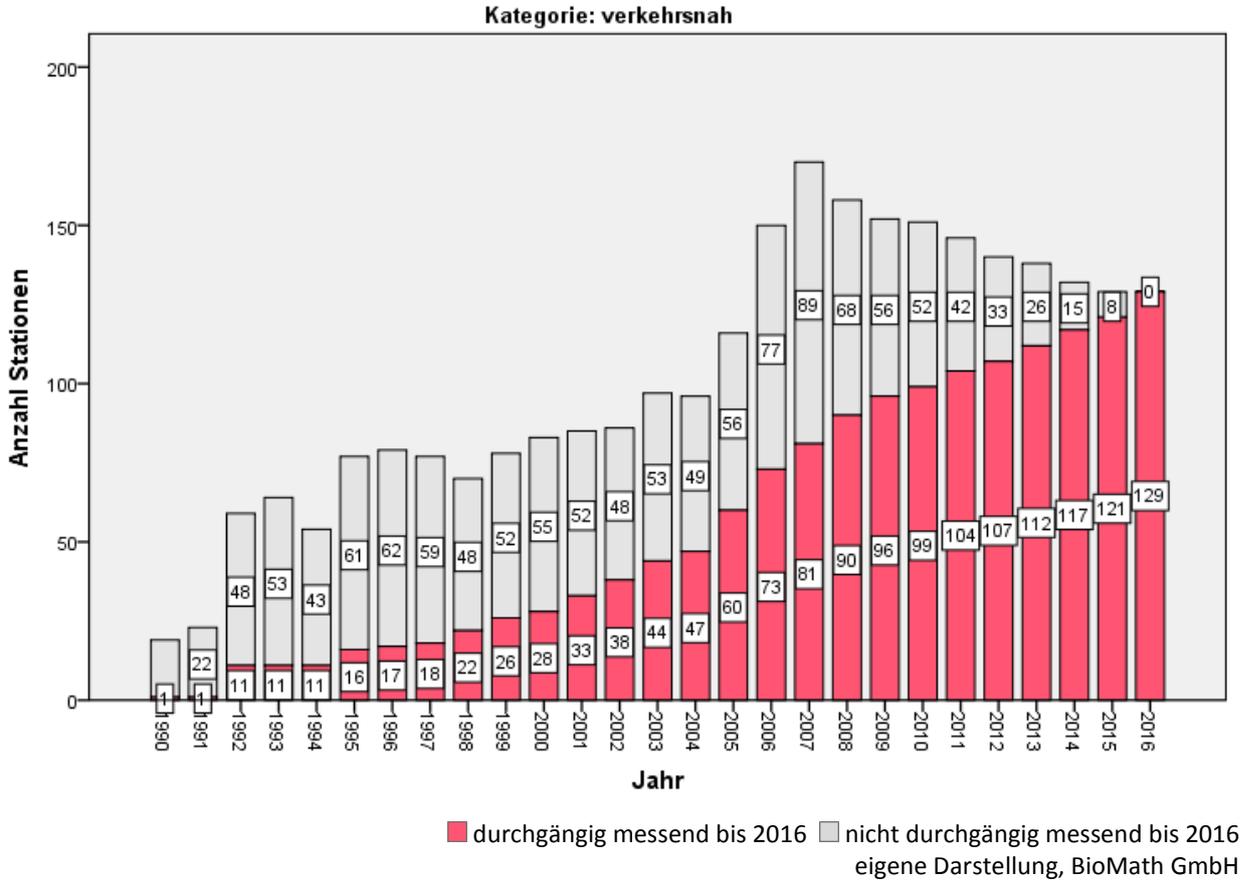
eigene Darstellung, BioMath GmbH



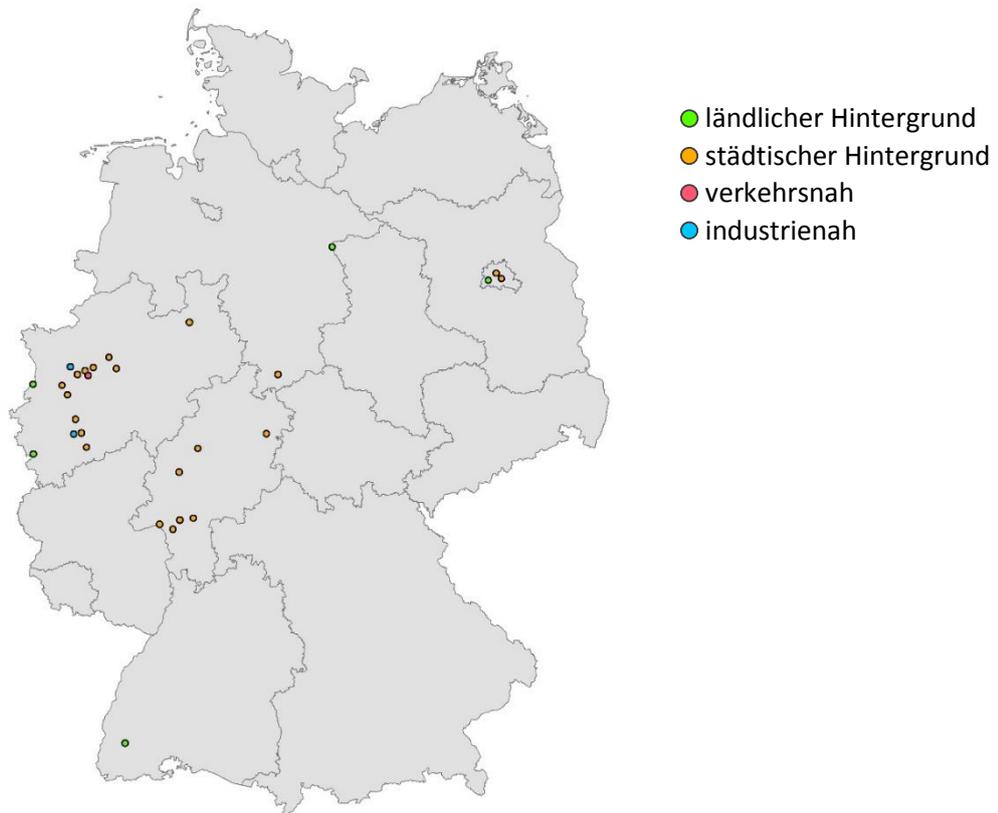
— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnah
 — industrienah
 eigene Darstellung, BioMath GmbH

3.1.3.3 PM10 - Anzahl durchgängig und nicht durchgängig messender Stationen bis 2016

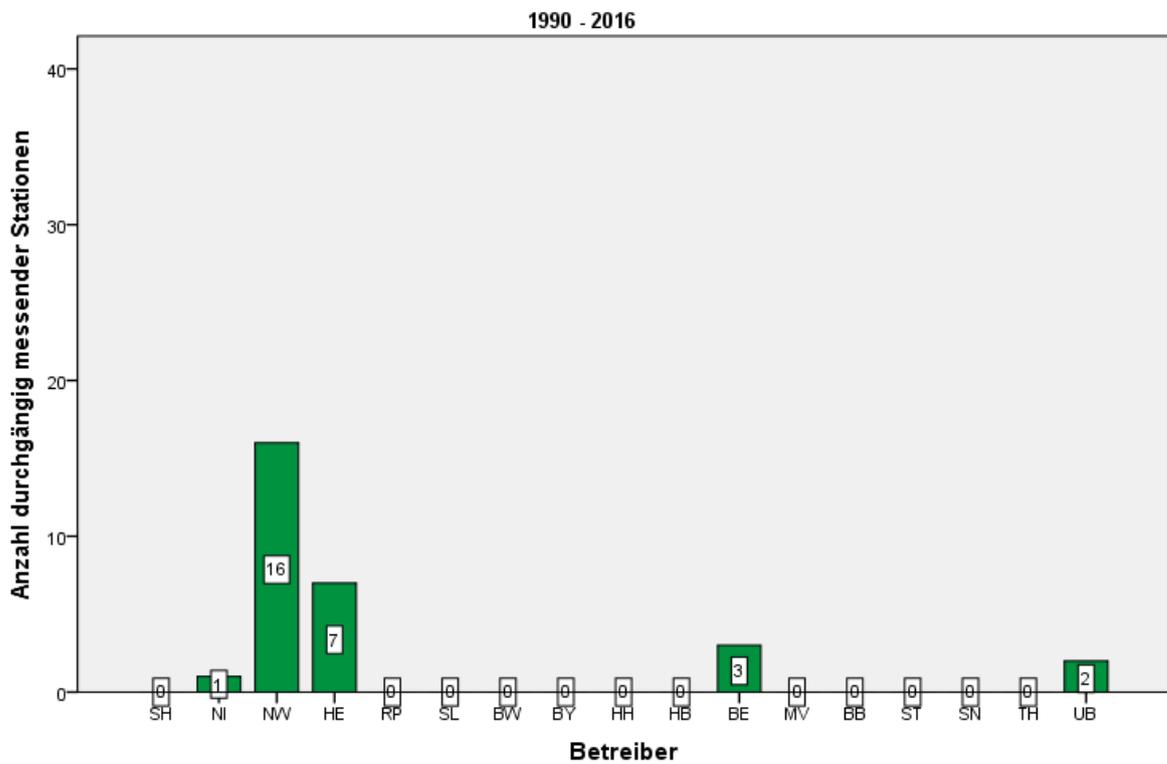




3.1.3.4 PM10 - Anzahl durchgängig messender Stationen 1990-2016 (Karte / Betreiber)



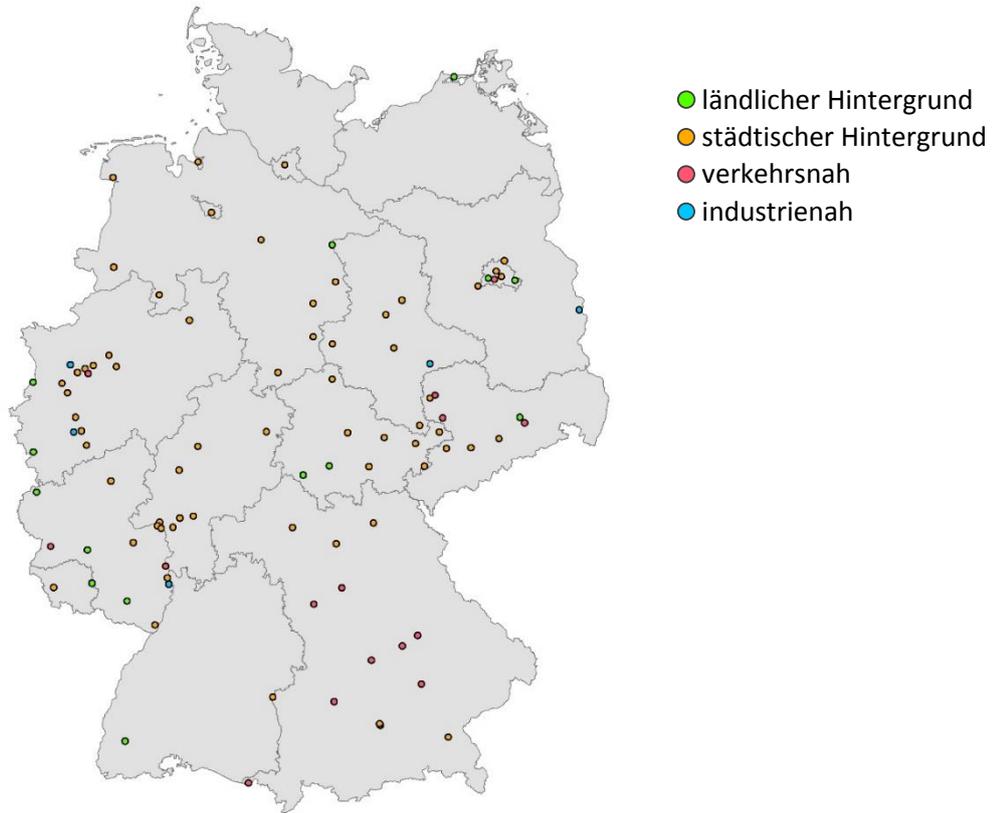
eigene Darstellung, BioMath



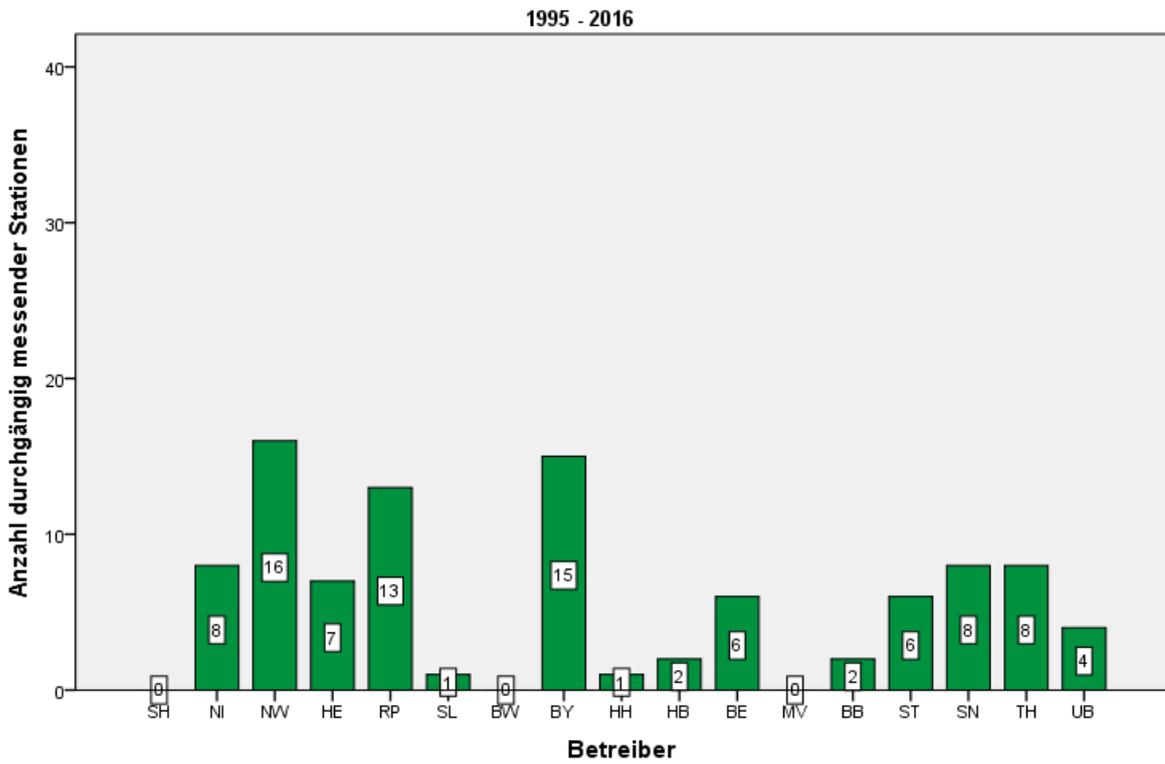
GmbH

eigene Darstellung, BioMath GmbH

3.1.3.5 PM10 - Anzahl durchgängig messender Stationen 1995-2016 (Karte / Betreiber)



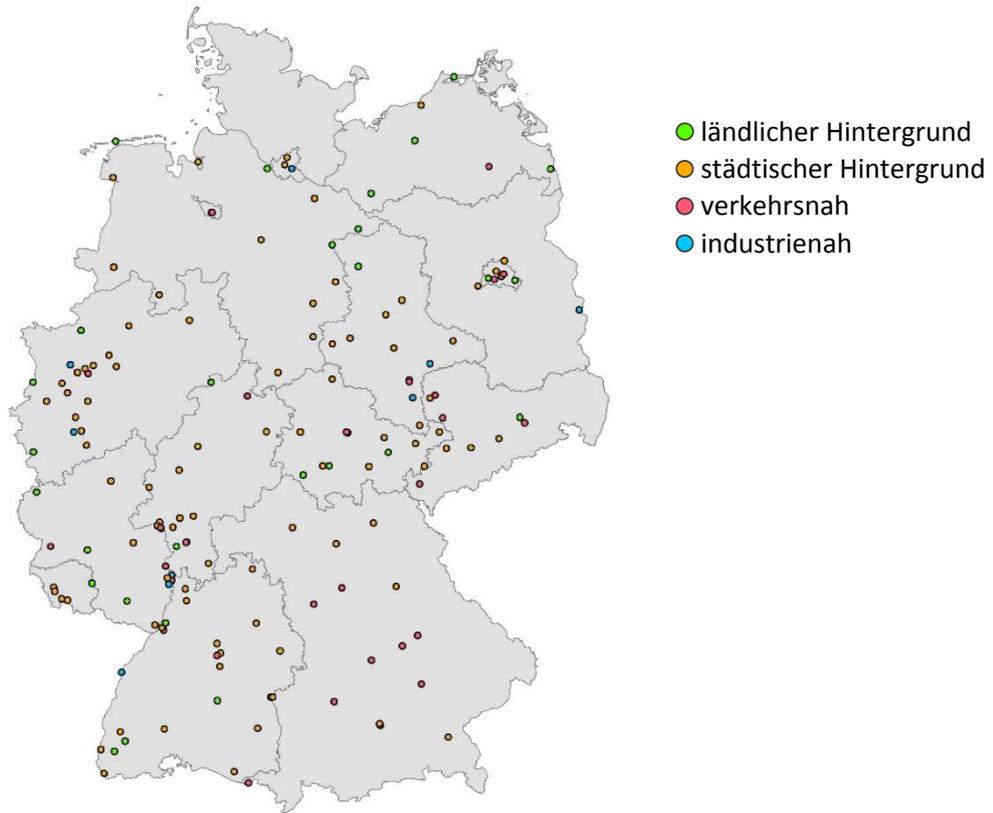
eigene Darstellung, BioMath



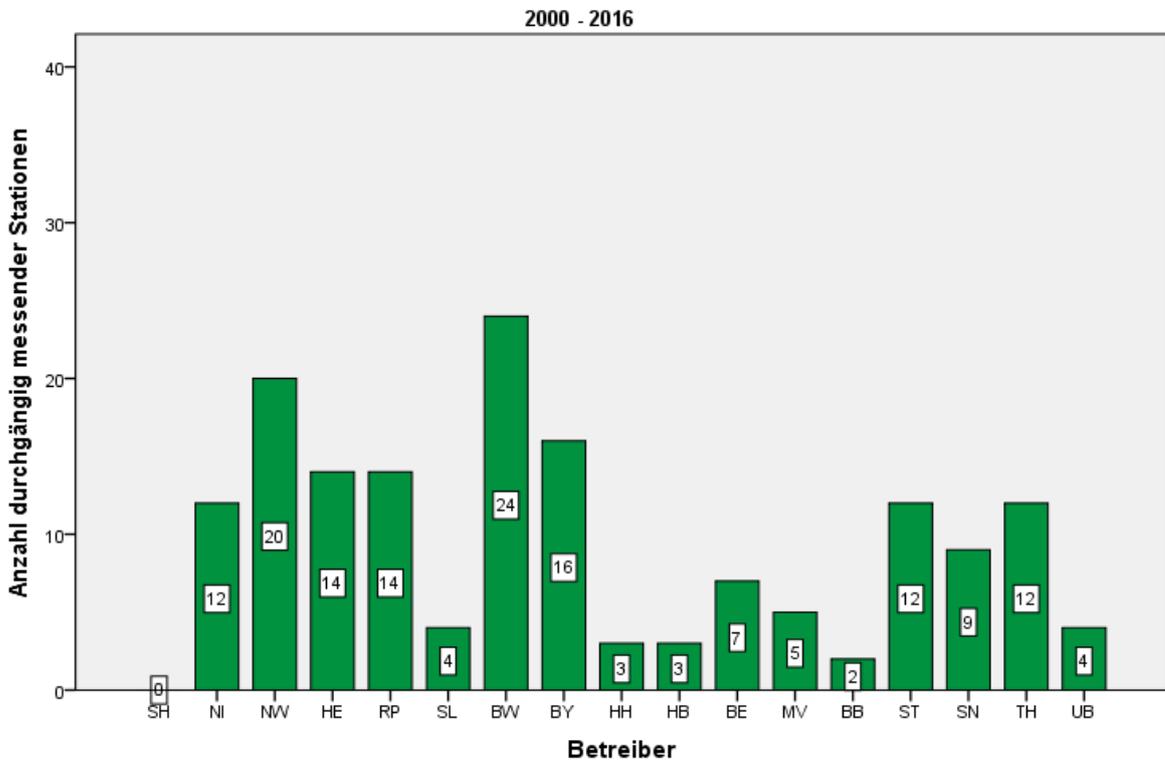
GmbH

eigene Darstellung, BioMath GmbH

3.1.3.6 PM10 - Anzahl durchgängig messender Stationen 2000-2016 (Karte / Betreiber)



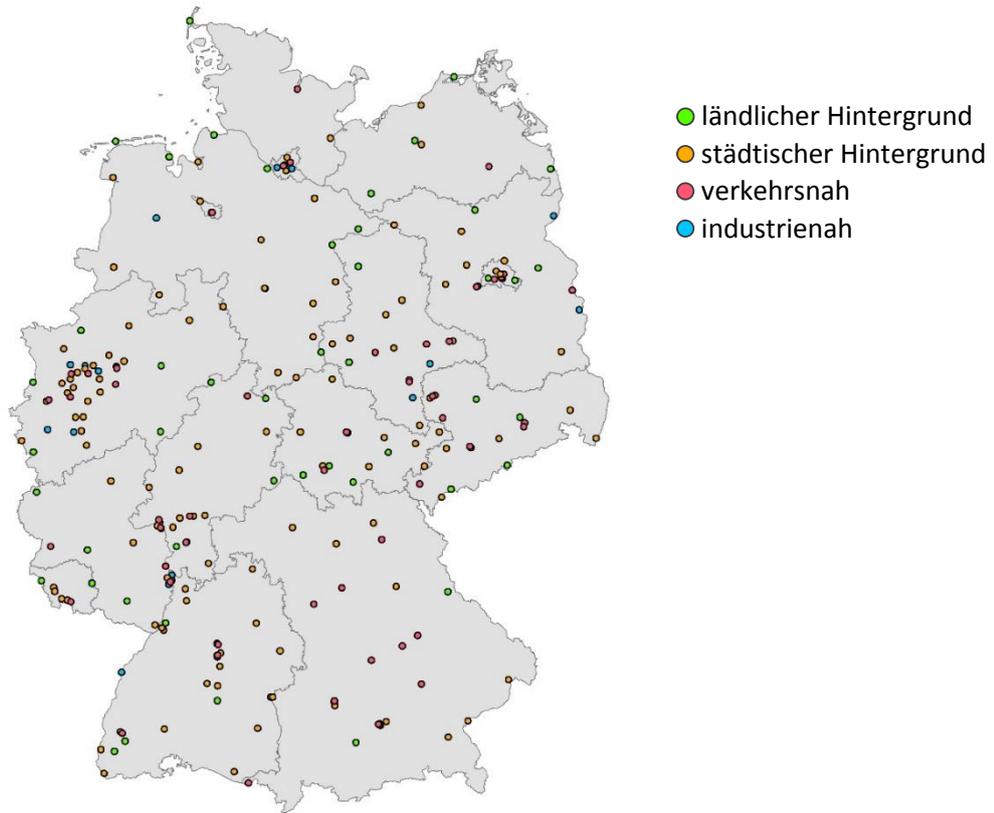
eigene Darstellung, BioMath



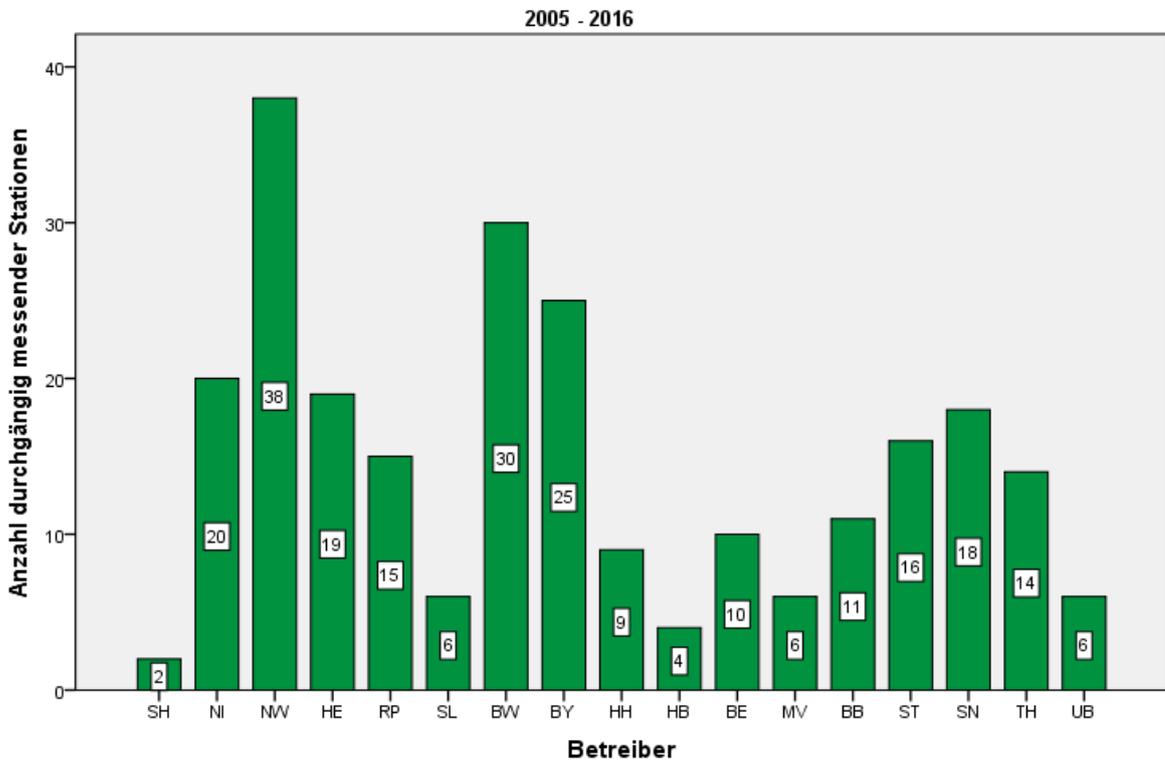
GmbH

eigene Darstellung, BioMath GmbH

3.1.3.7 PM10 - Anzahl durchgängig messender Stationen 2005-2016 (Karte / Betreiber)



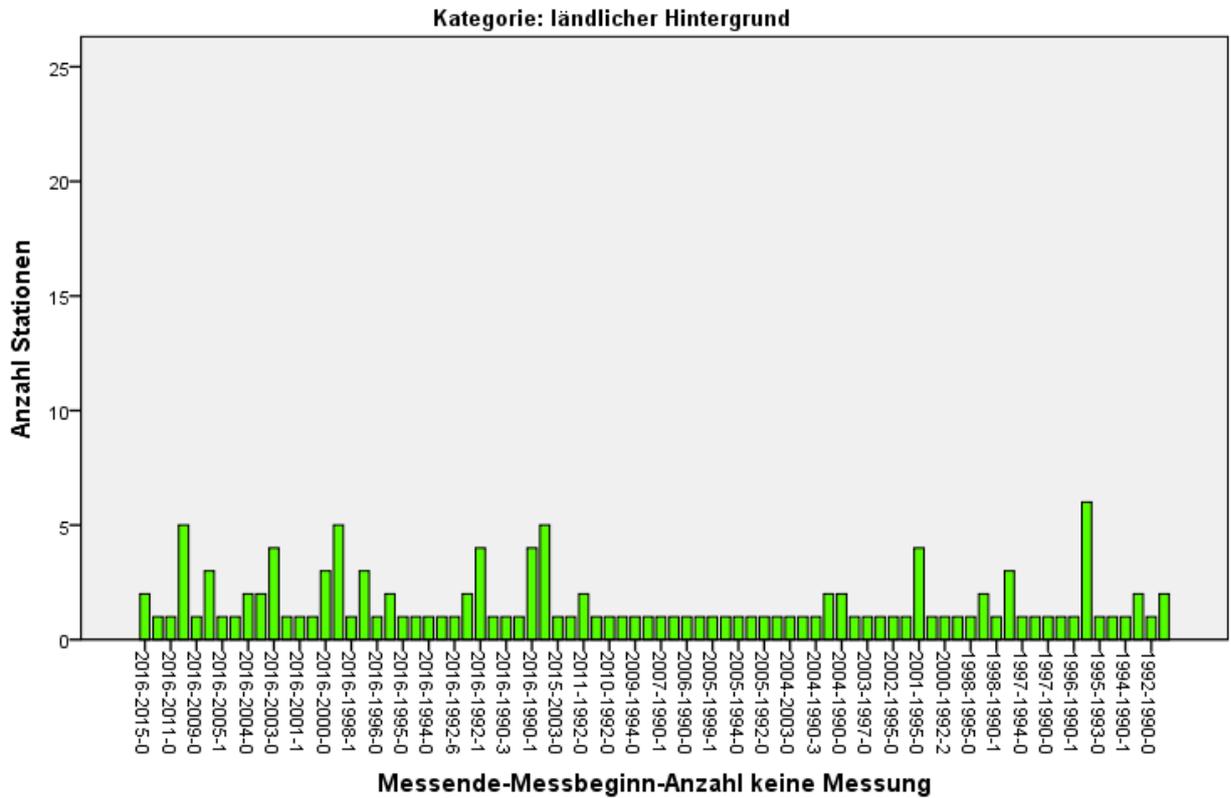
eigene Darstellung, BioMath



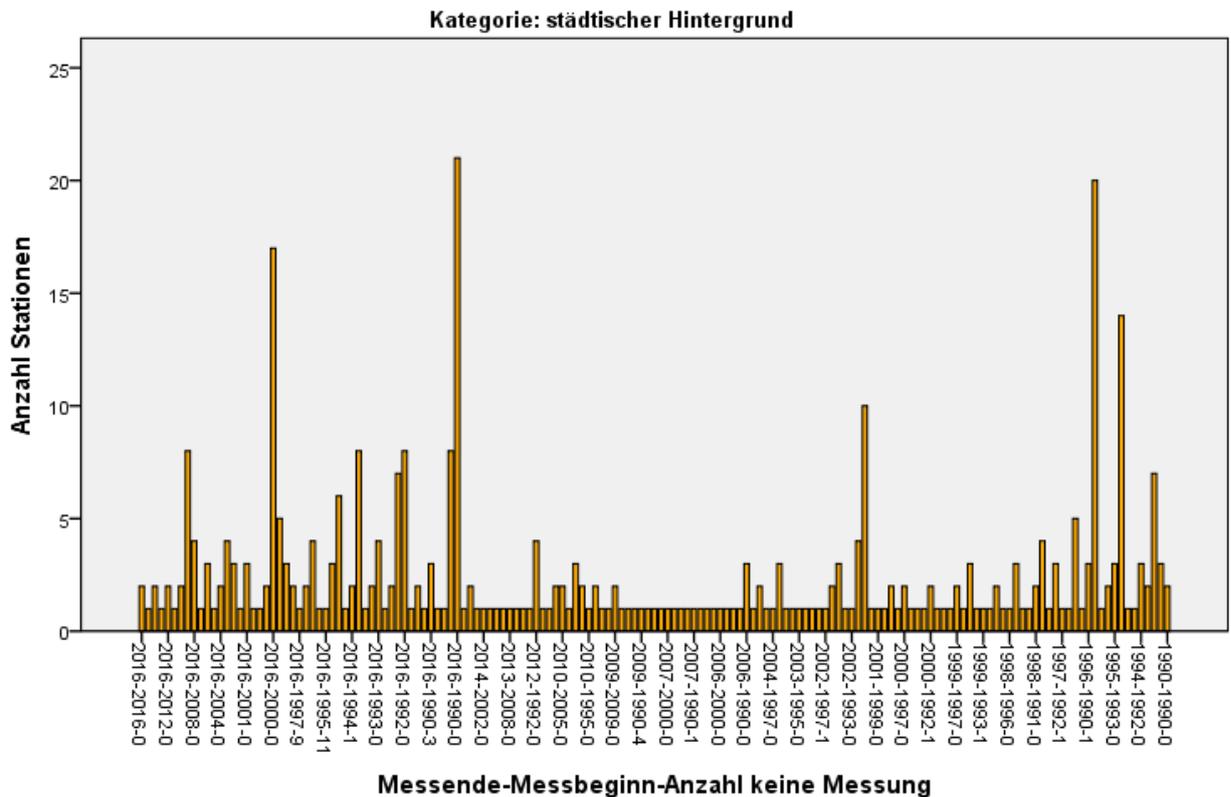
GmbH

eigene Darstellung, BioMath GmbH

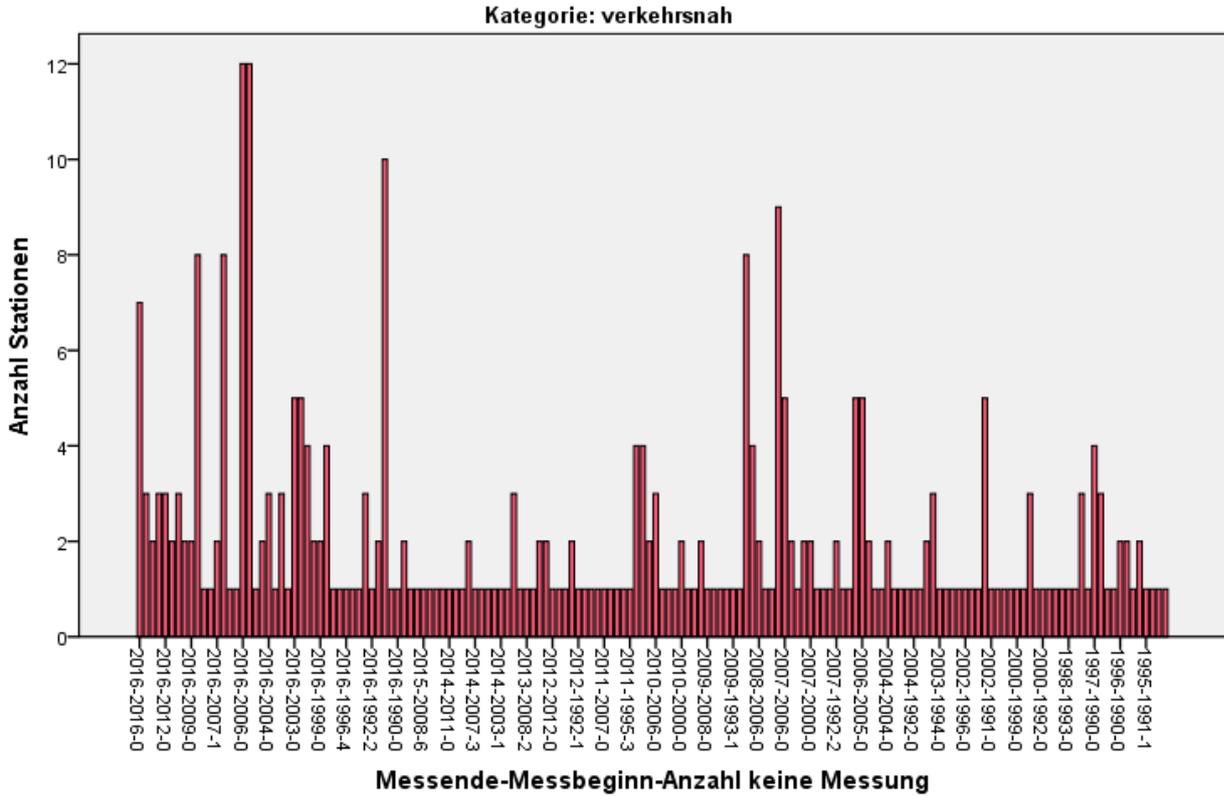
3.1.3.8 PM10 - Anzahl Stationen für Zeitabschnitte in denen Messstationen betrieben wurden



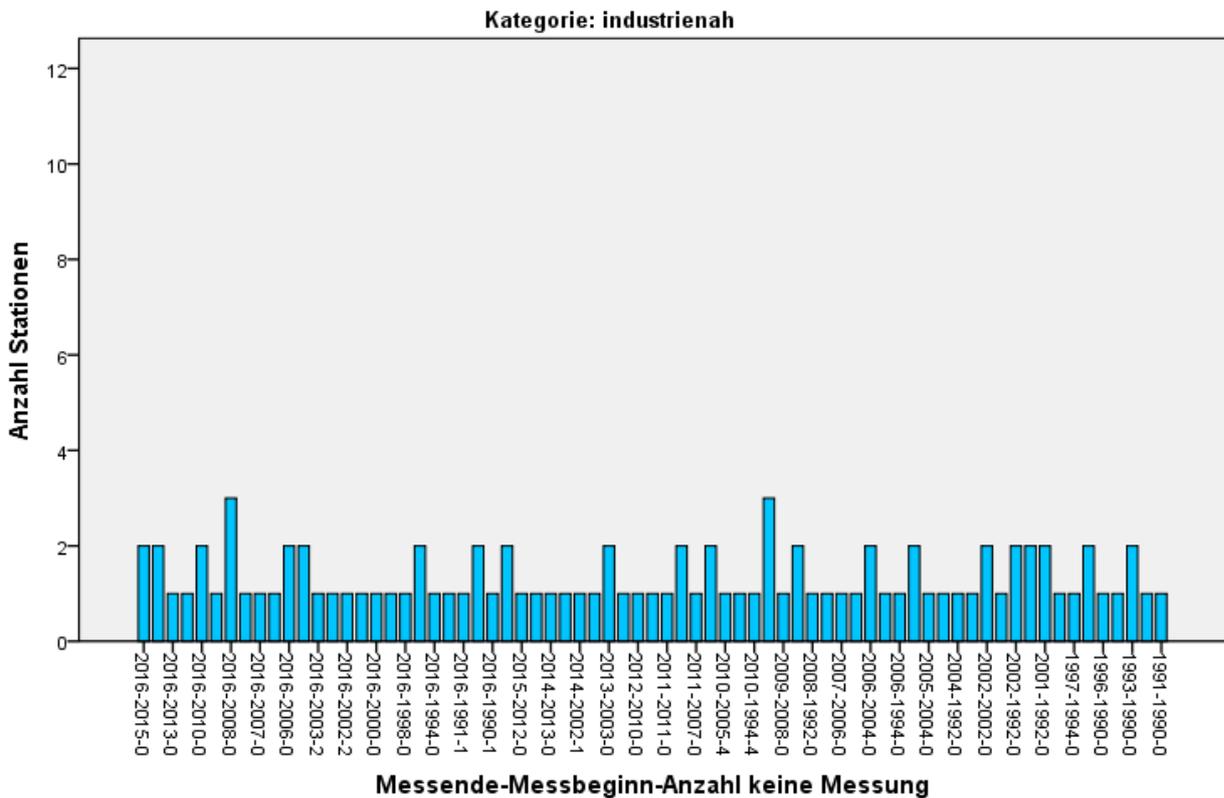
eigene Darstellung, BioMath GmbH



eigene Darstellung, BioMath GmbH

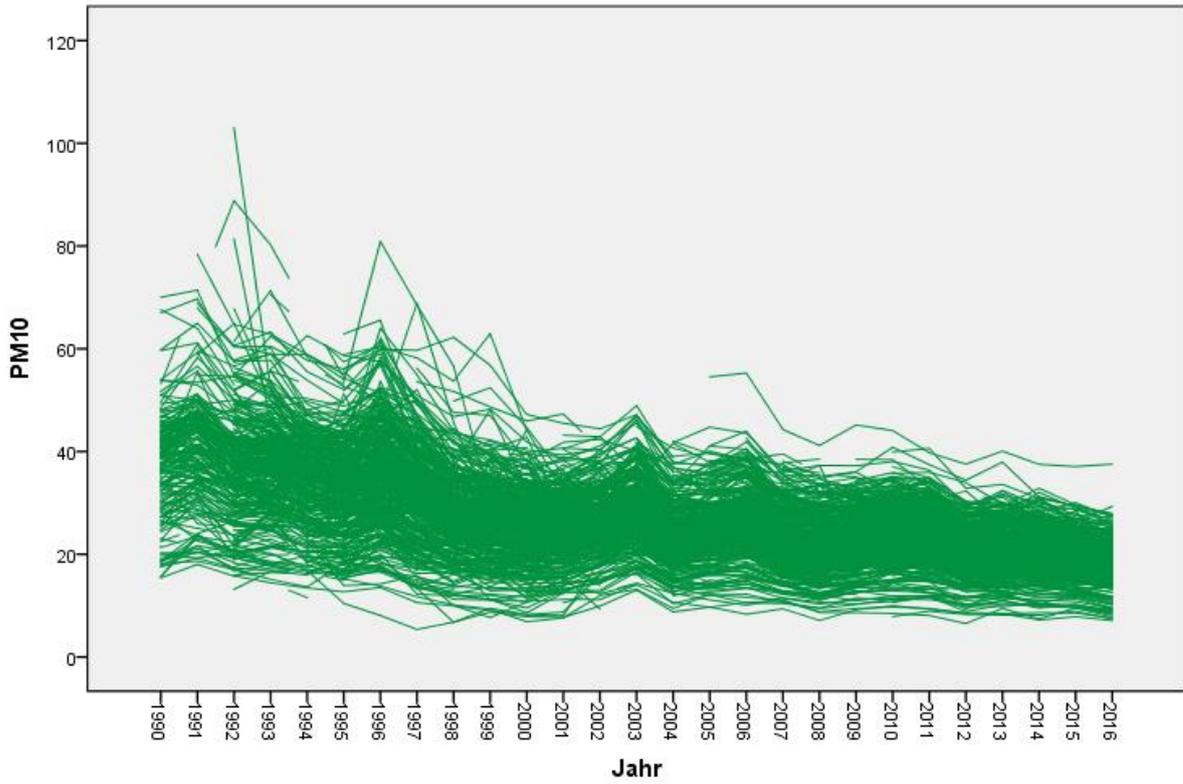


eigene Darstellung, BioMath GmbH

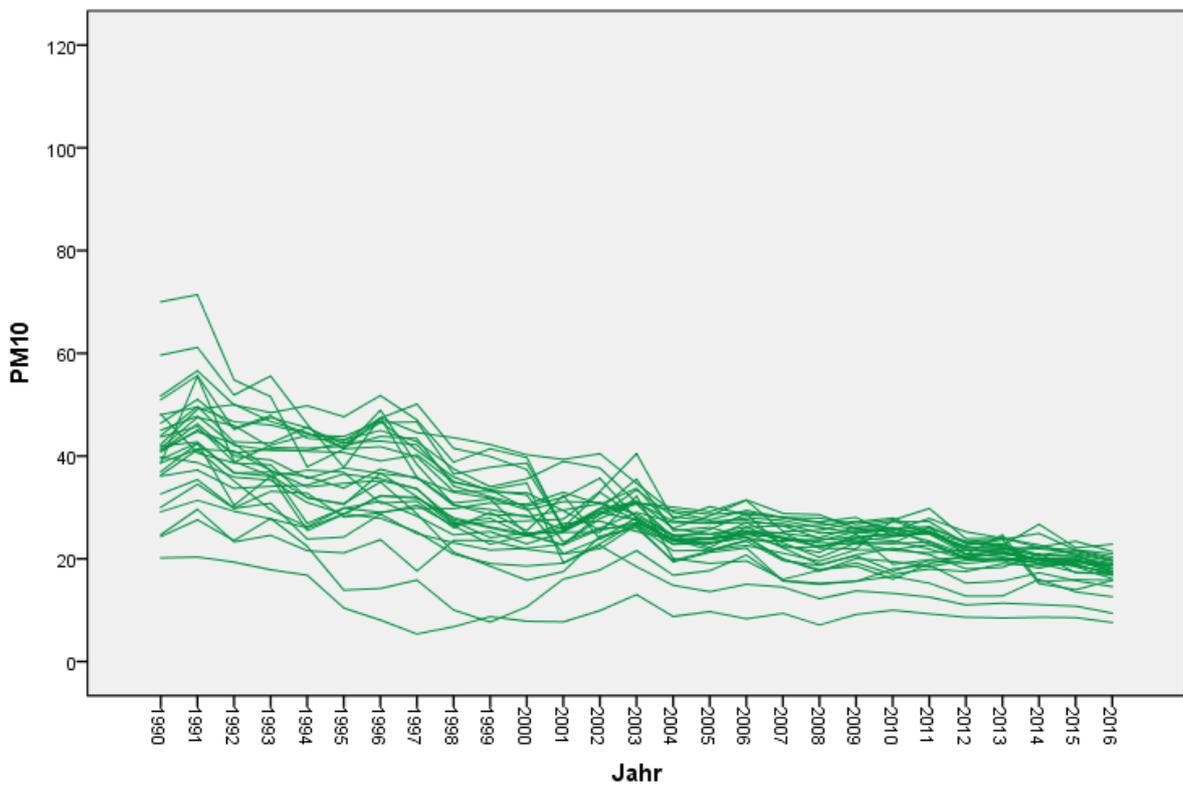


eigene Darstellung, BioMath GmbH

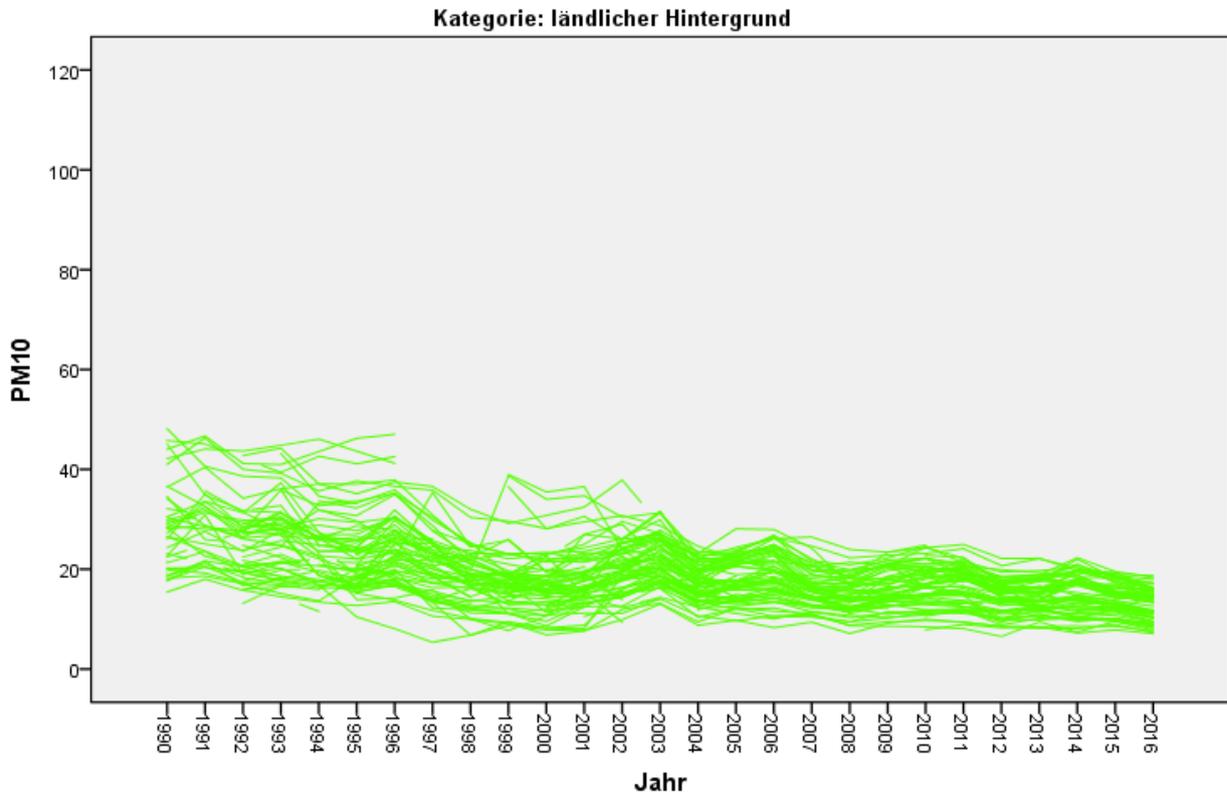
3.1.3.9 PM10-Werte in $\mu\text{g}/\text{m}^3$ - alle Messreihen und durchgängig messende Messreihen von 1990 - 2016



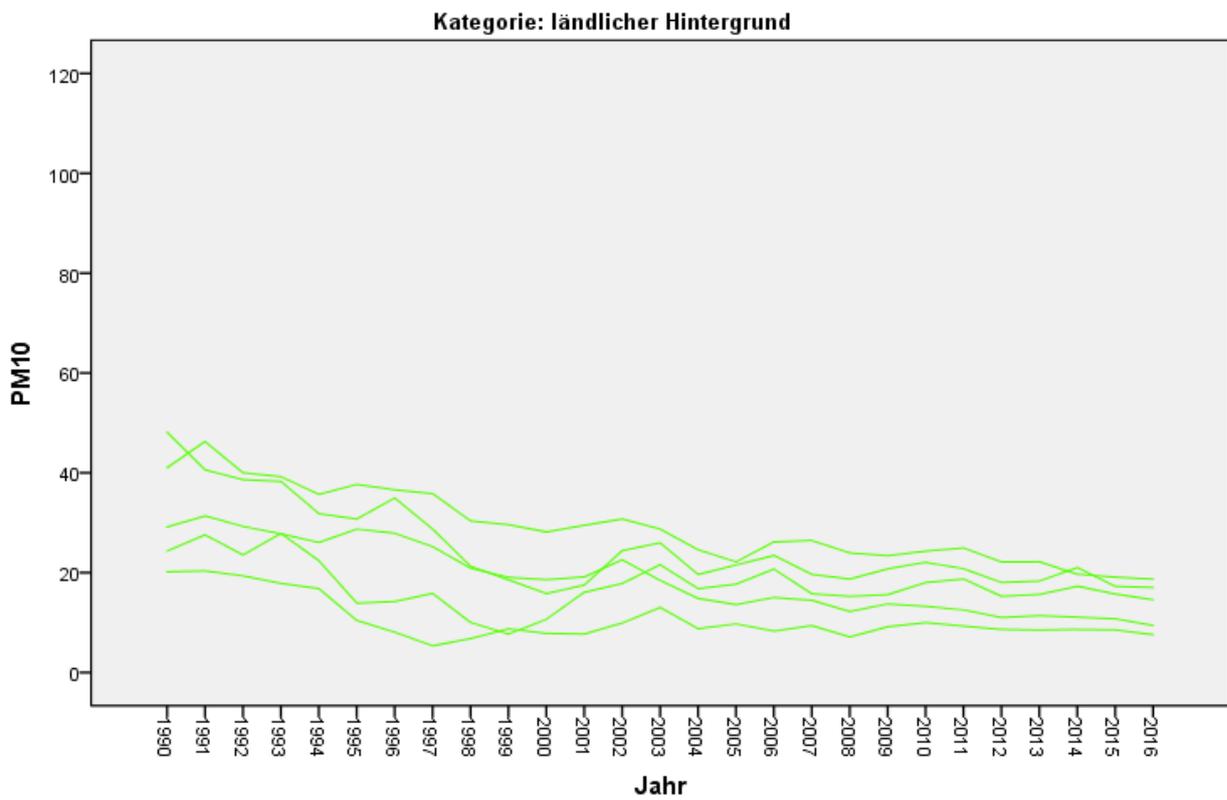
alle Messreihen
eigene Darstellung, BioMath GmbH



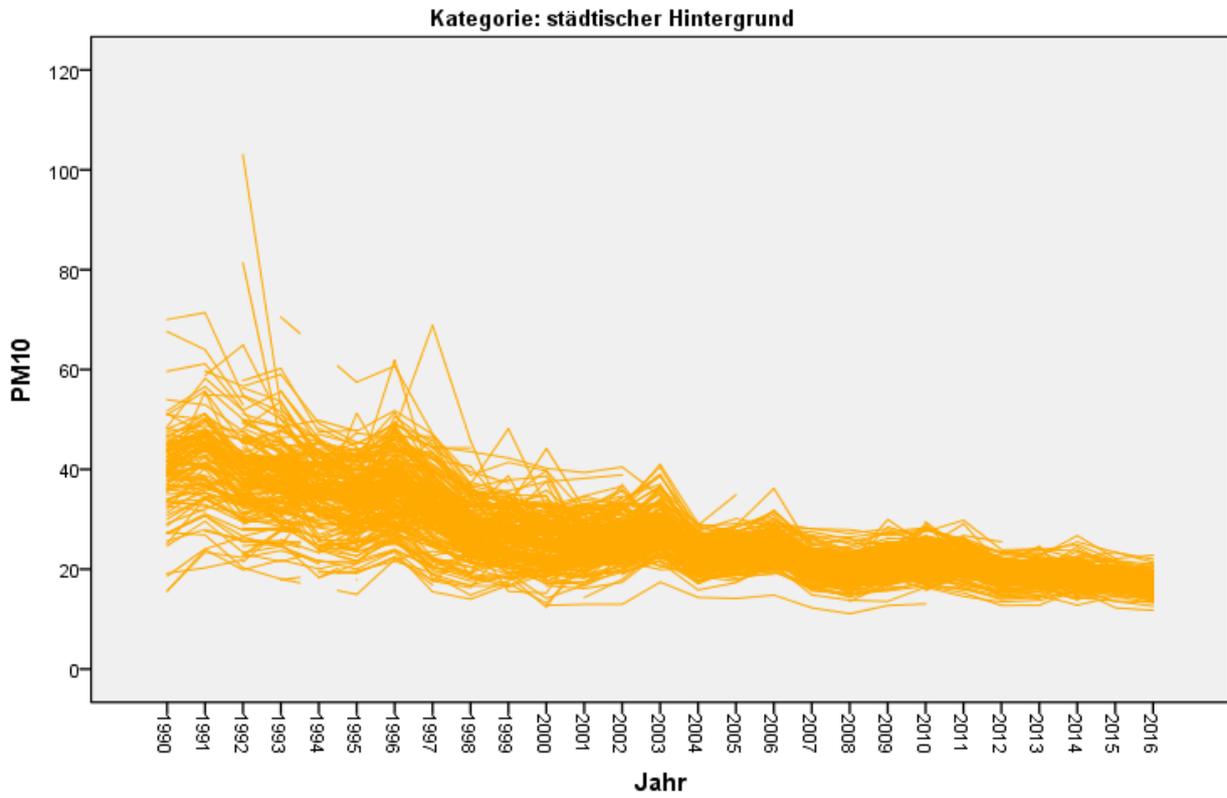
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



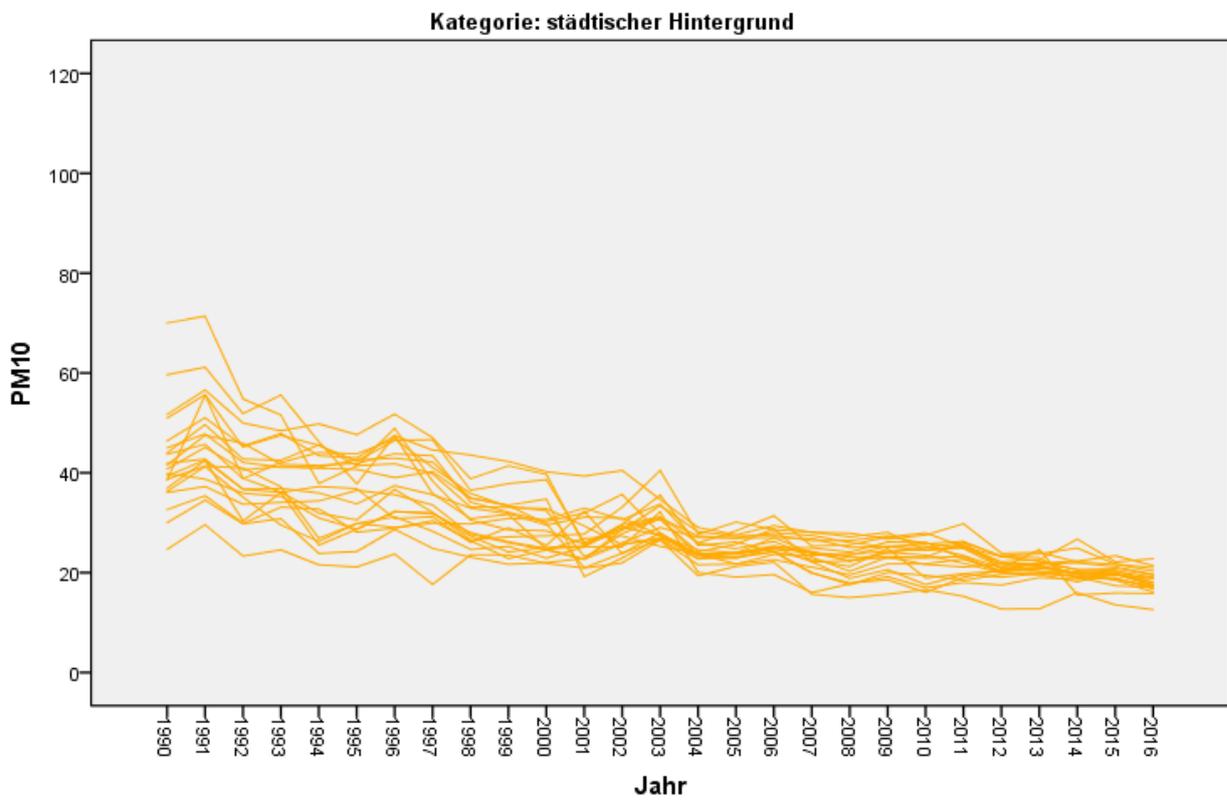
alle Messreihen
eigene Darstellung, BioMath GmbH



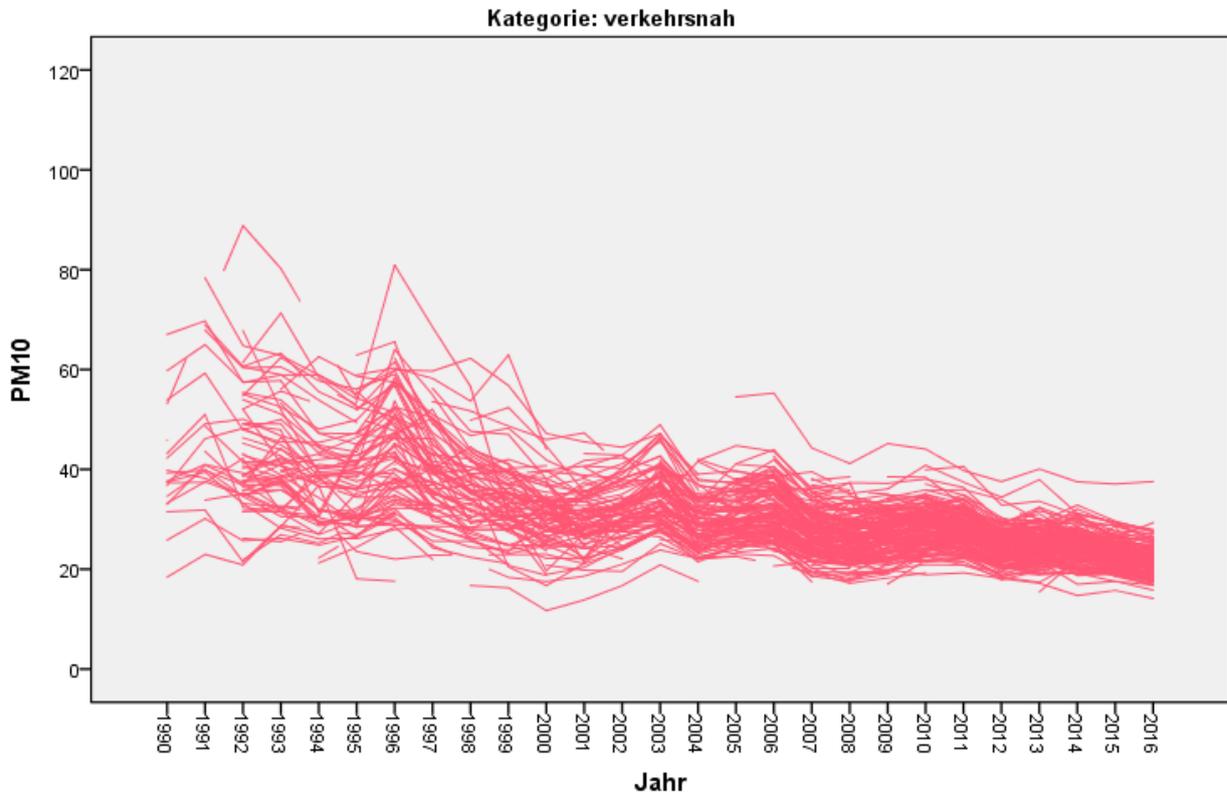
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



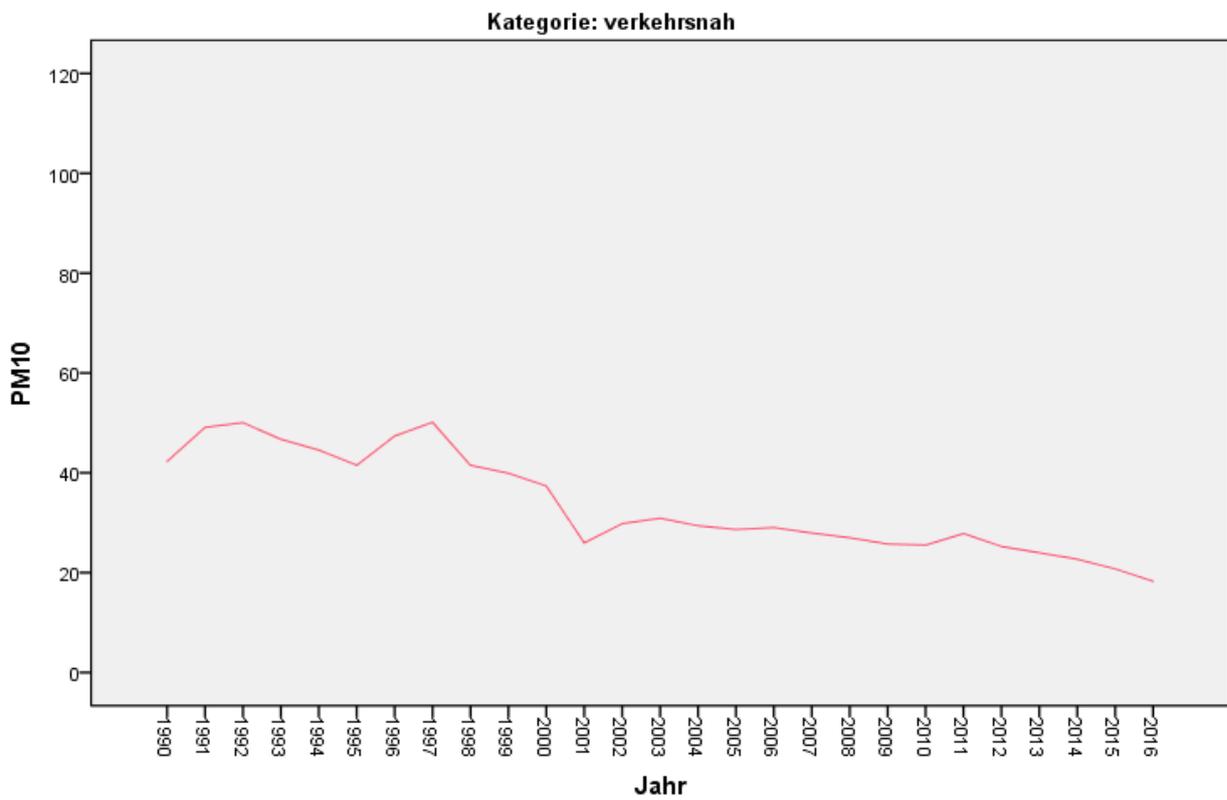
alle Messreihen
eigene Darstellung, BioMath GmbH



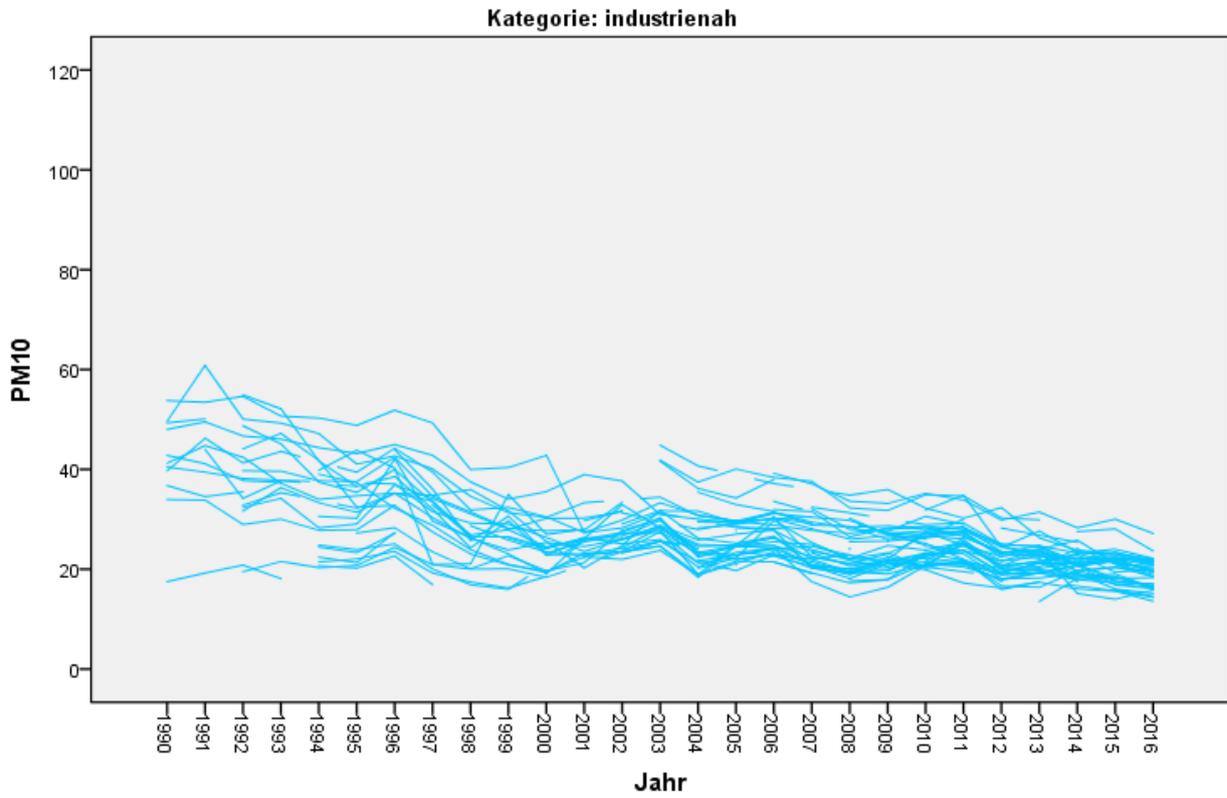
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH



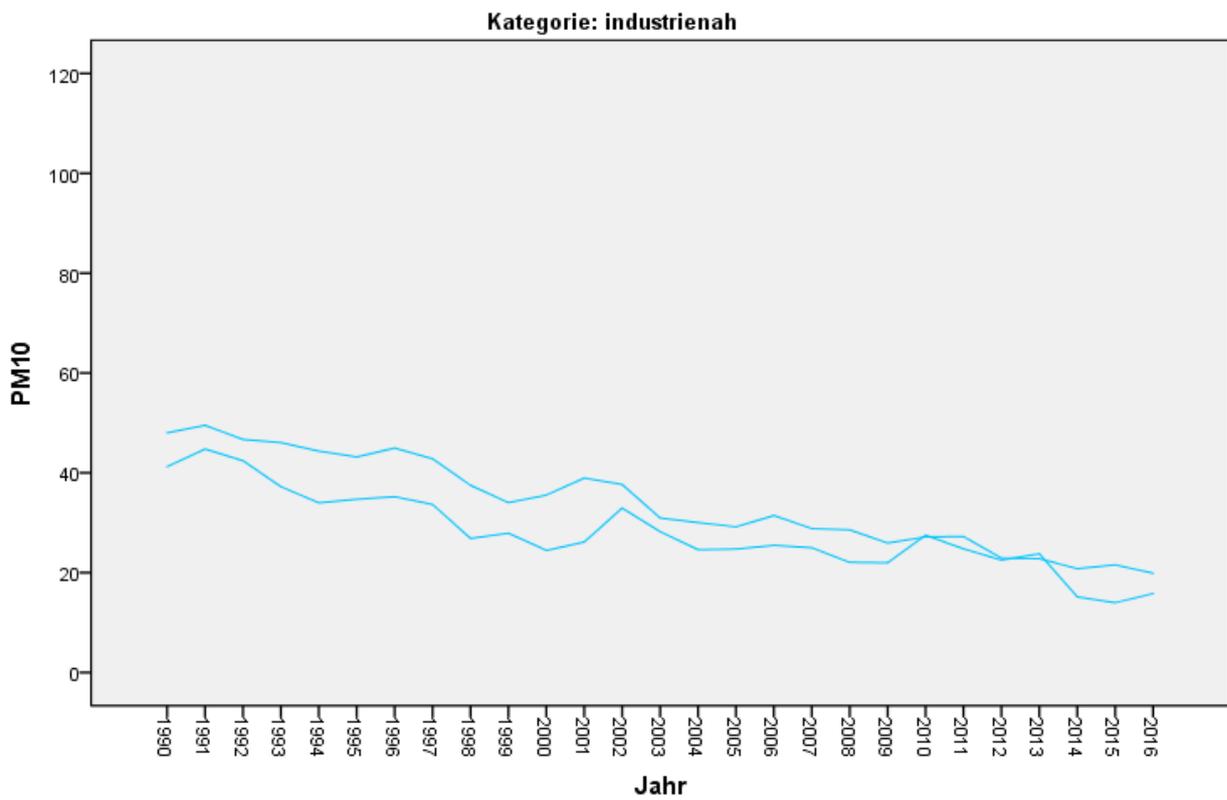
alle Messreihen
eigene Darstellung, BioMath GmbH



durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

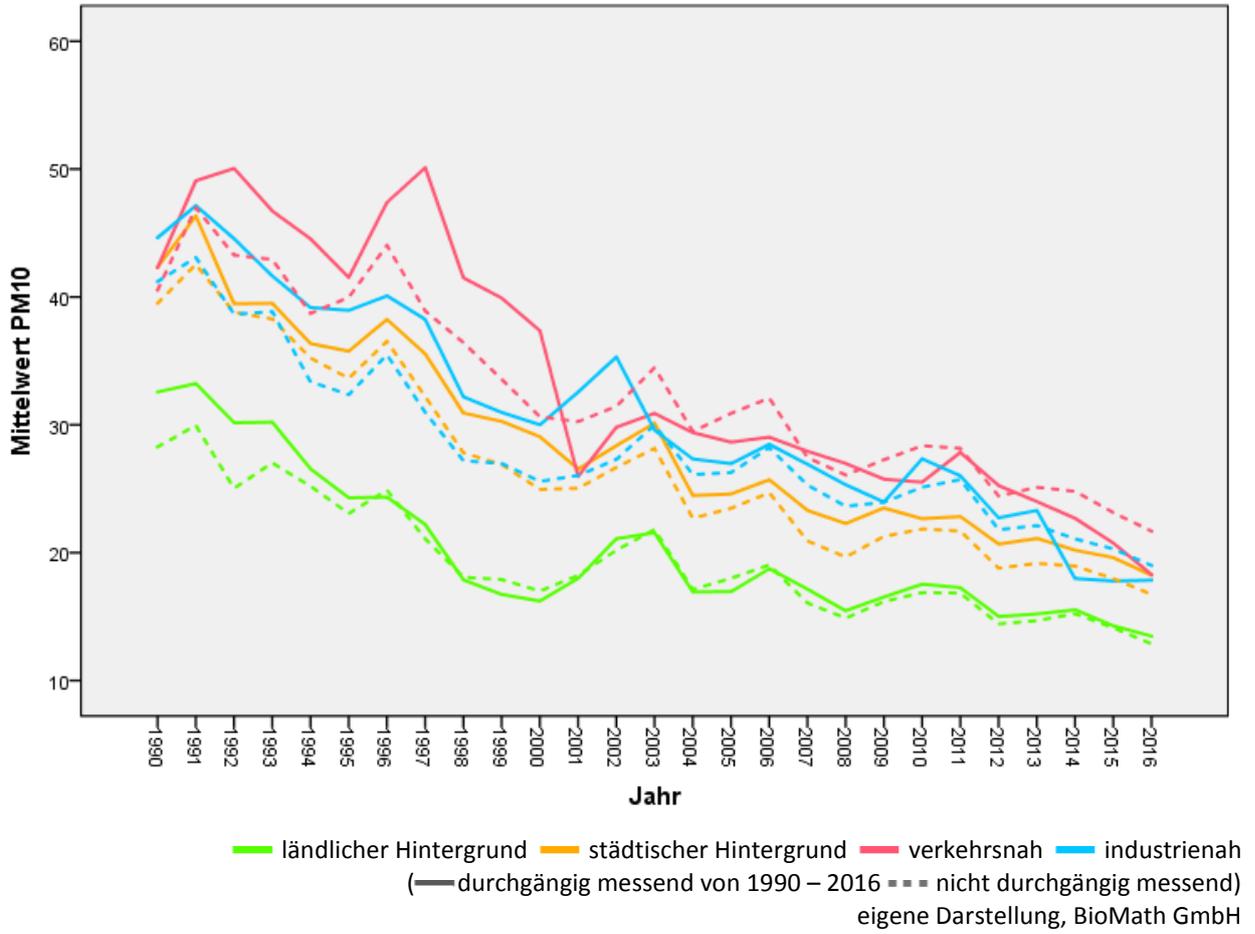


alle Messreihen
eigene Darstellung, BioMath GmbH

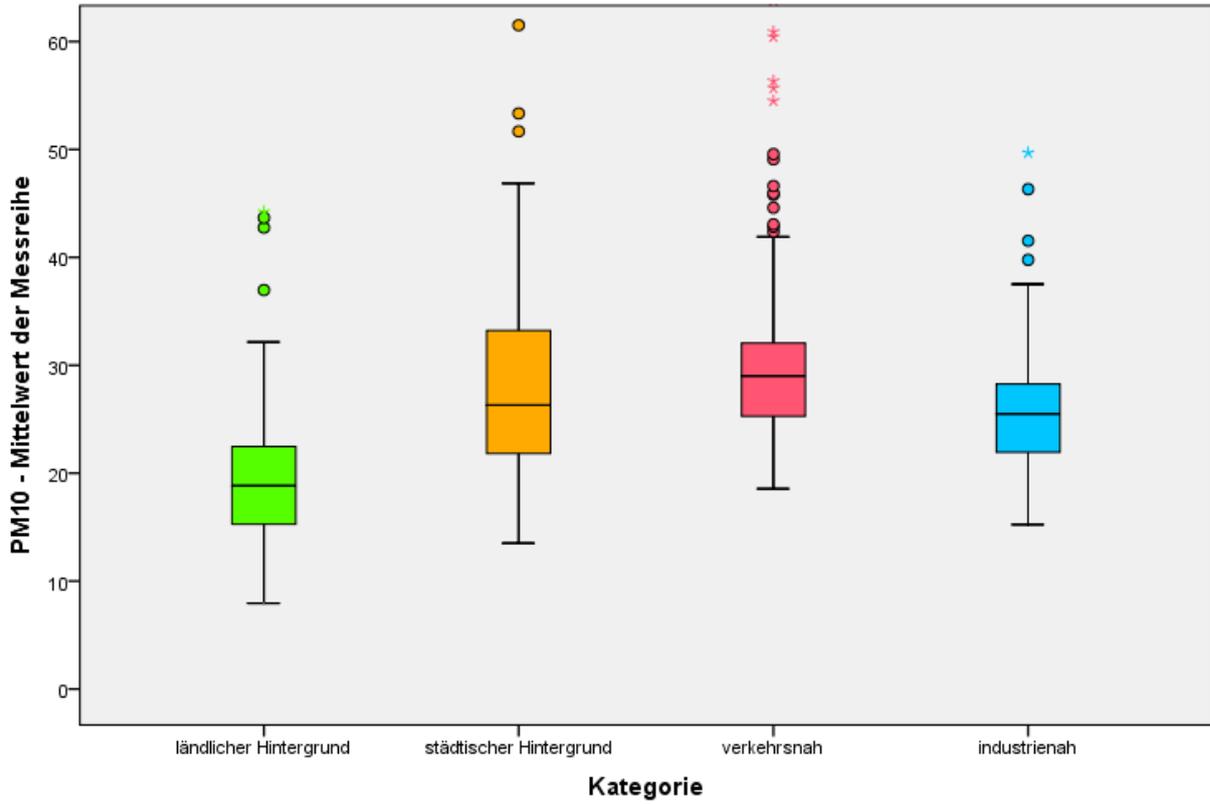


durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

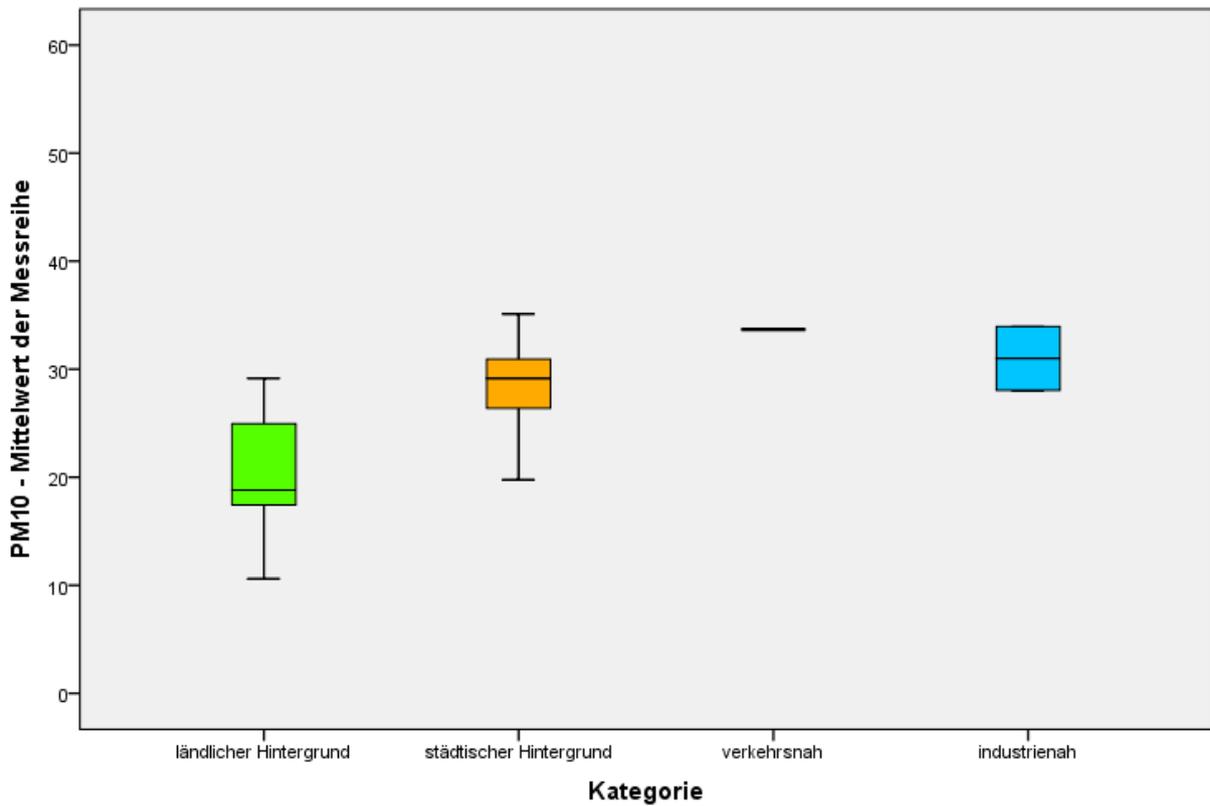
3.1.3.10 PM10 - Messreihen der jährlichen Mittelwerte in $\mu\text{g}/\text{m}^3$



3.1.3.11 PM10 - Boxplot der Mittelwerte ($\mu\text{g}/\text{m}^3$) je Messreihe je Kategorie

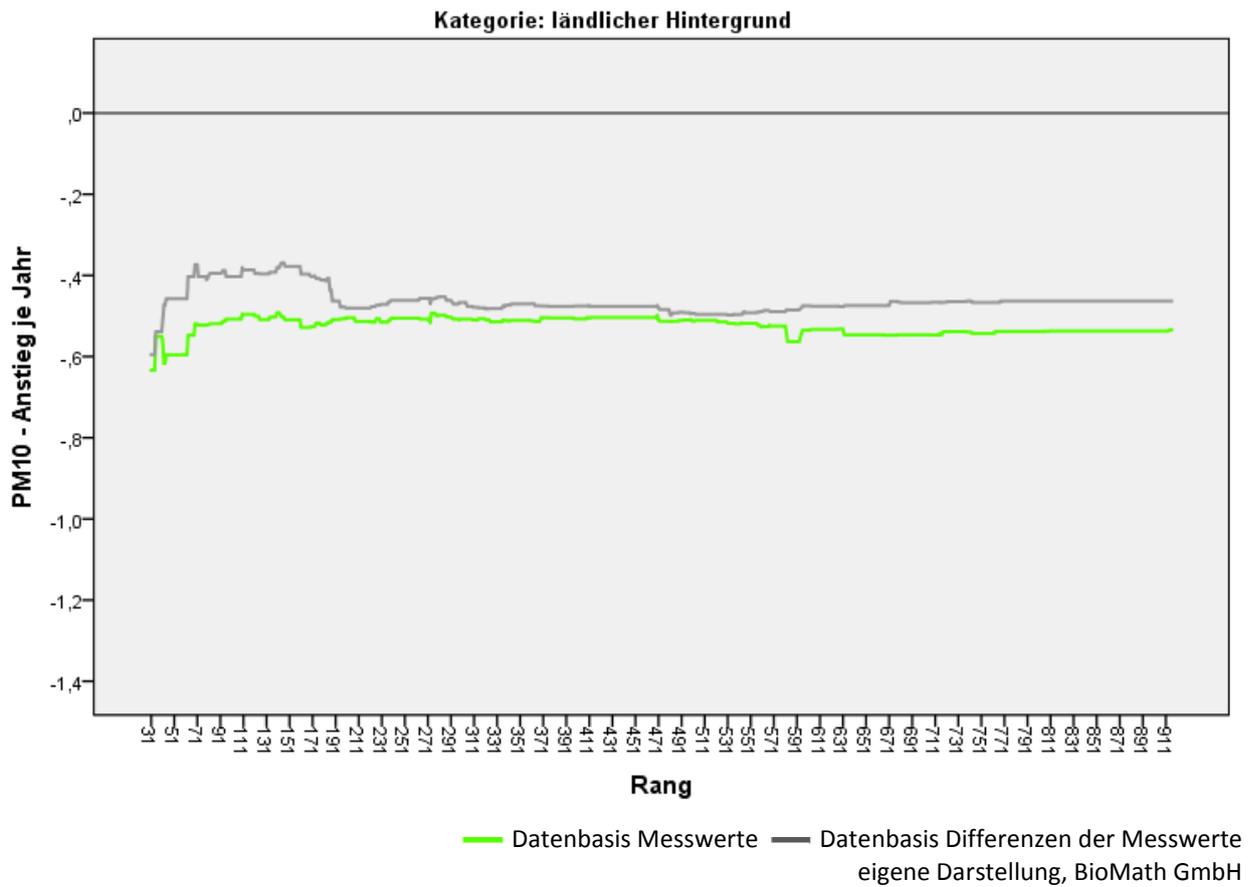


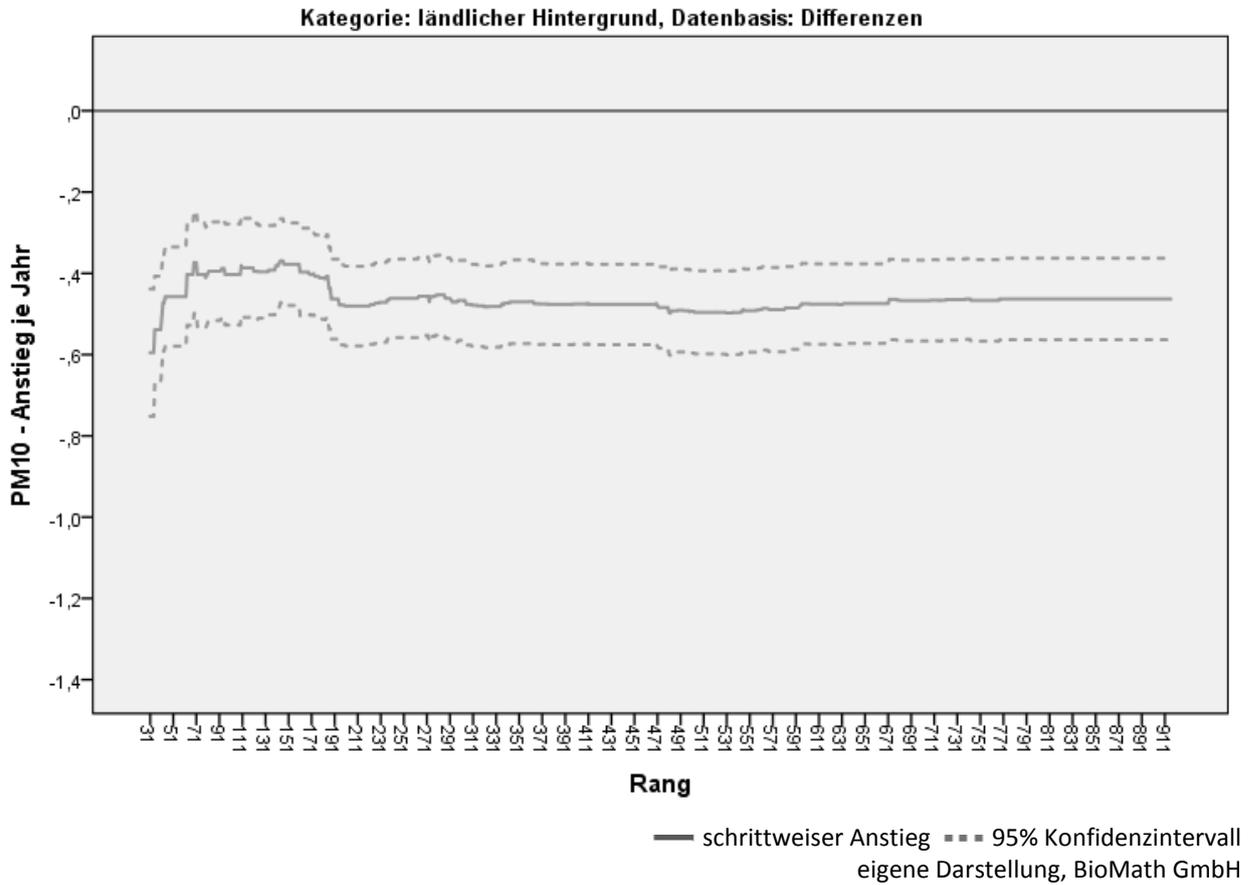
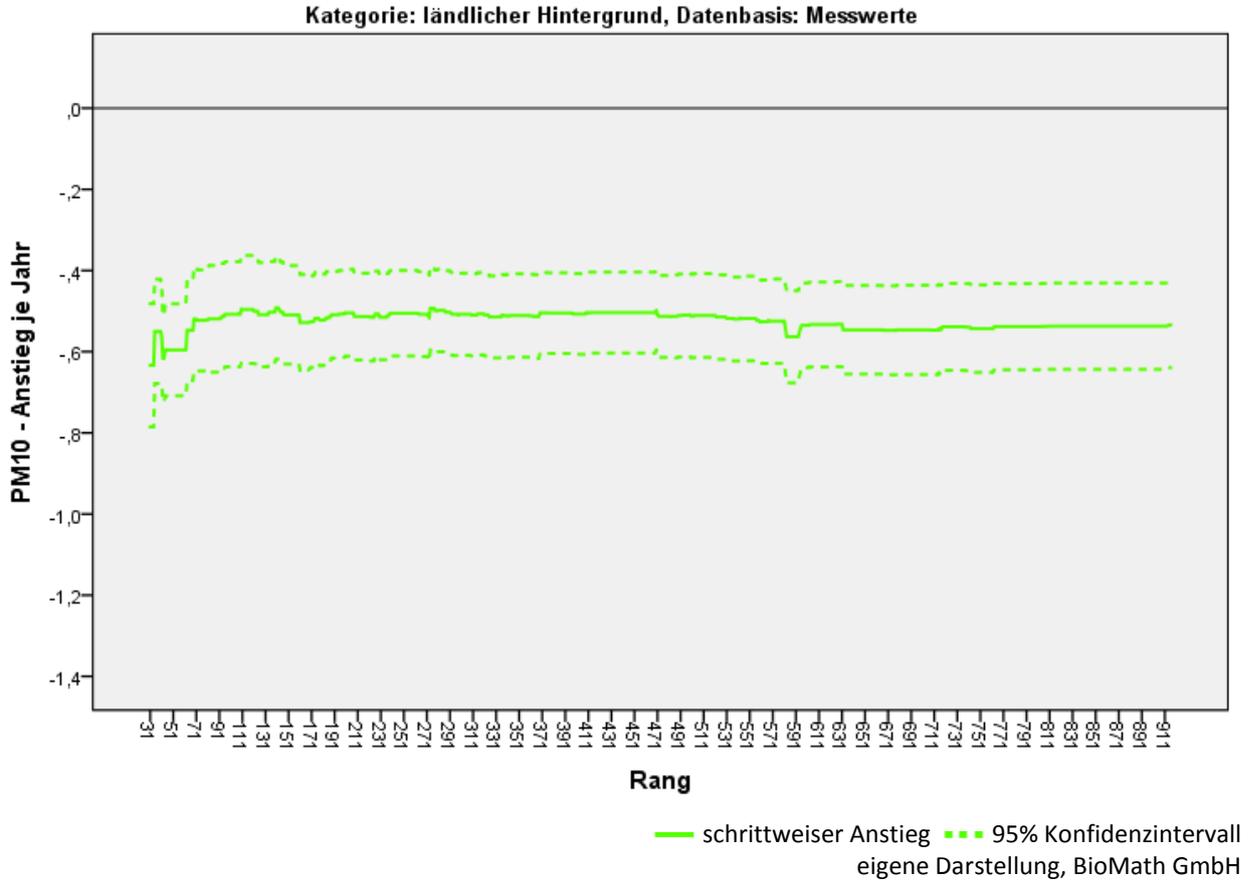
alle Messreihen
eigene Darstellung, BioMath GmbH



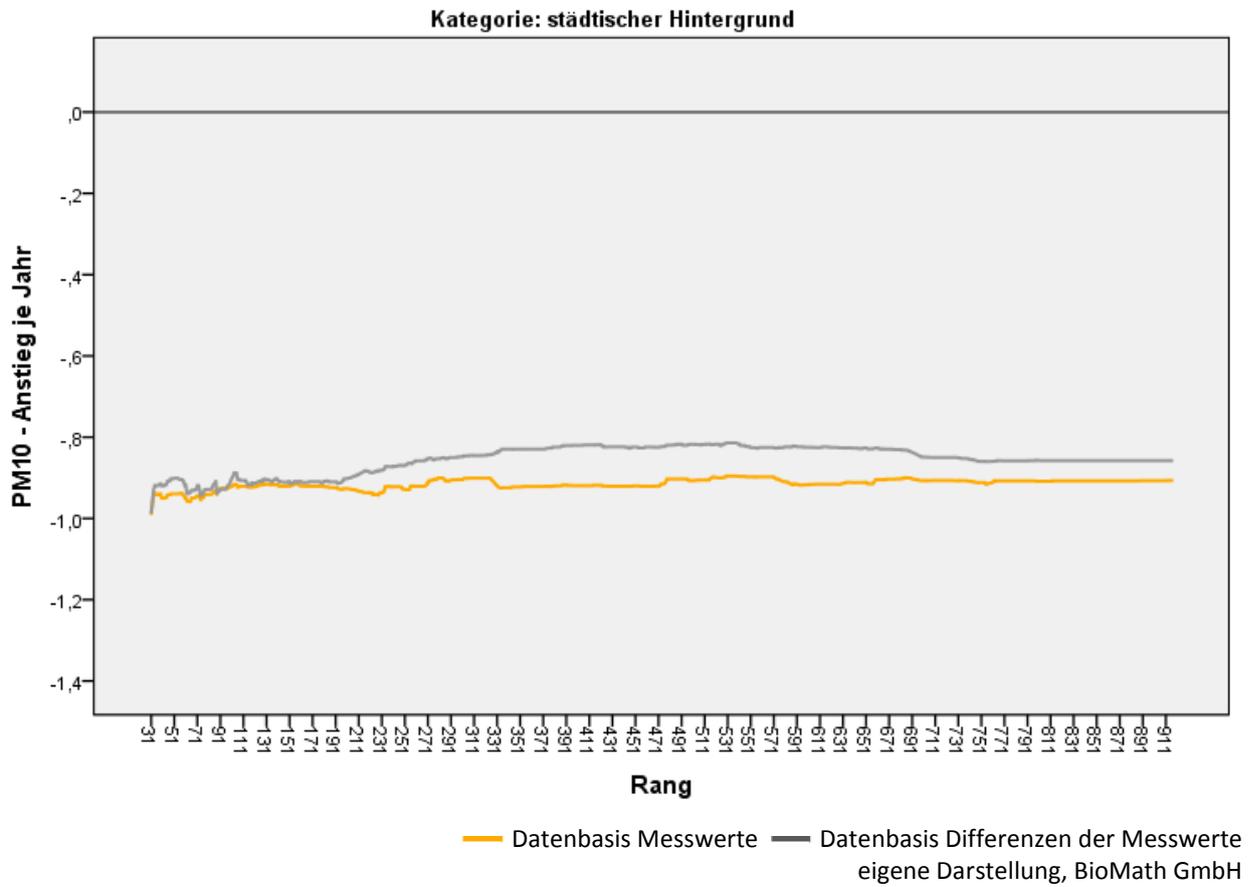
durchgängig messende Messreihen von 1990 – 2016
eigene Darstellung, BioMath GmbH

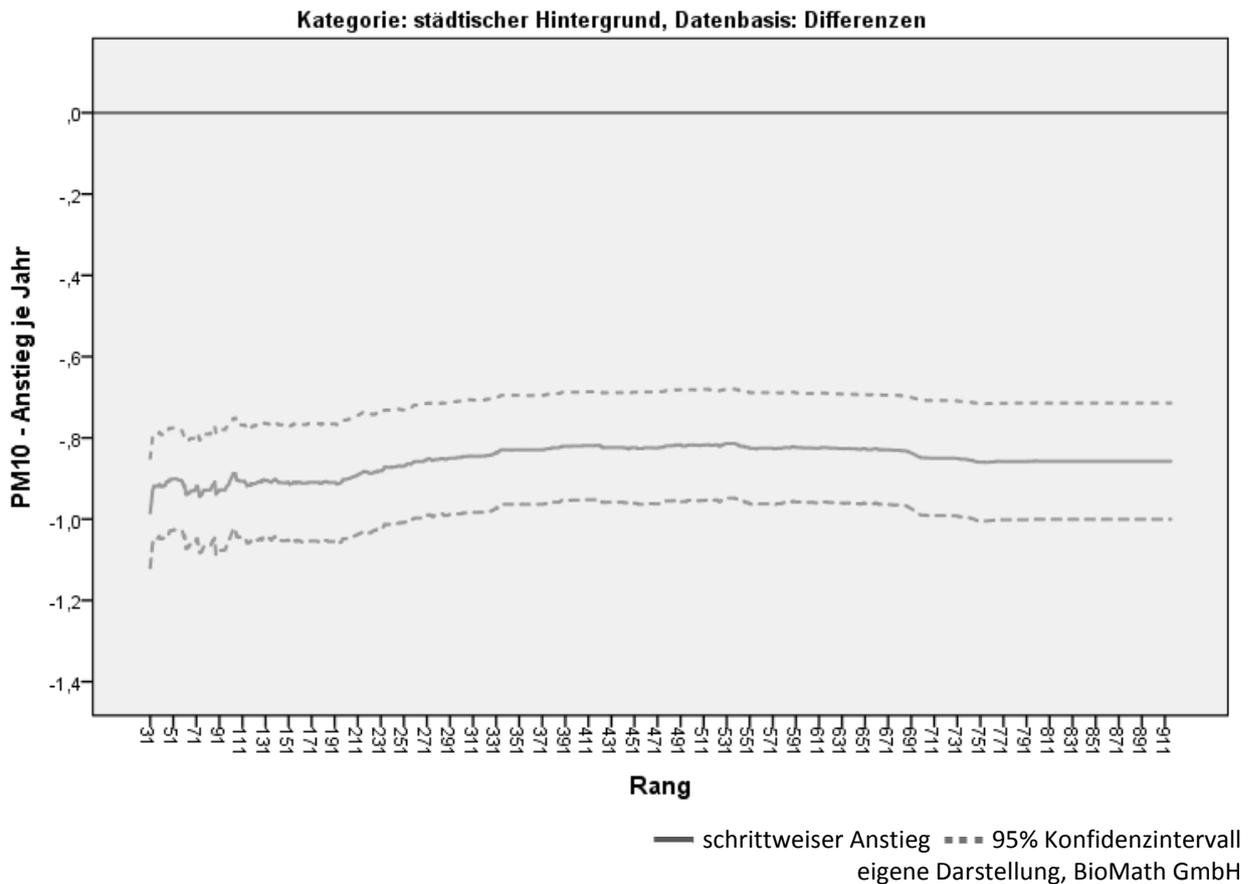
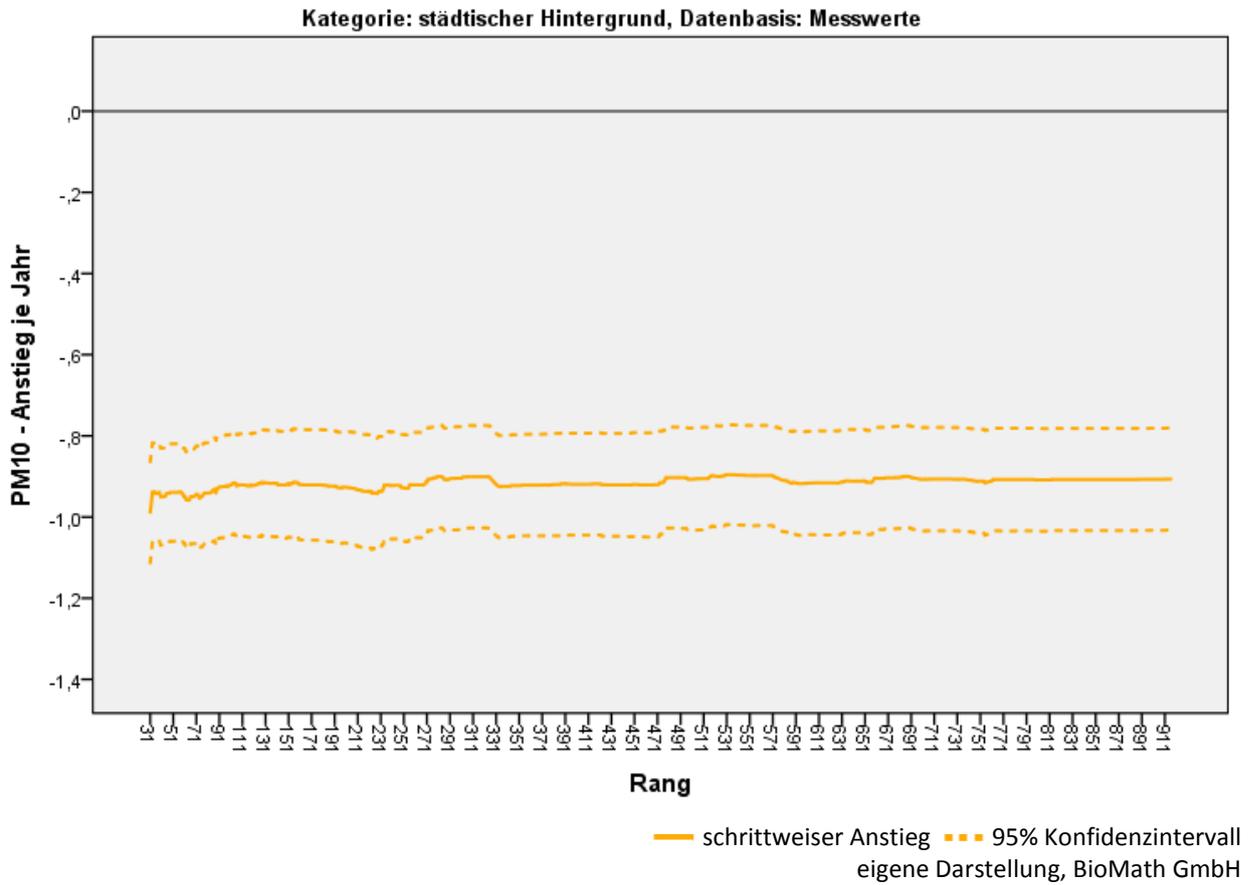
3.1.3.12 PM10 - Trend: schrittweise Anstiege ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe ländlicher Hintergrund



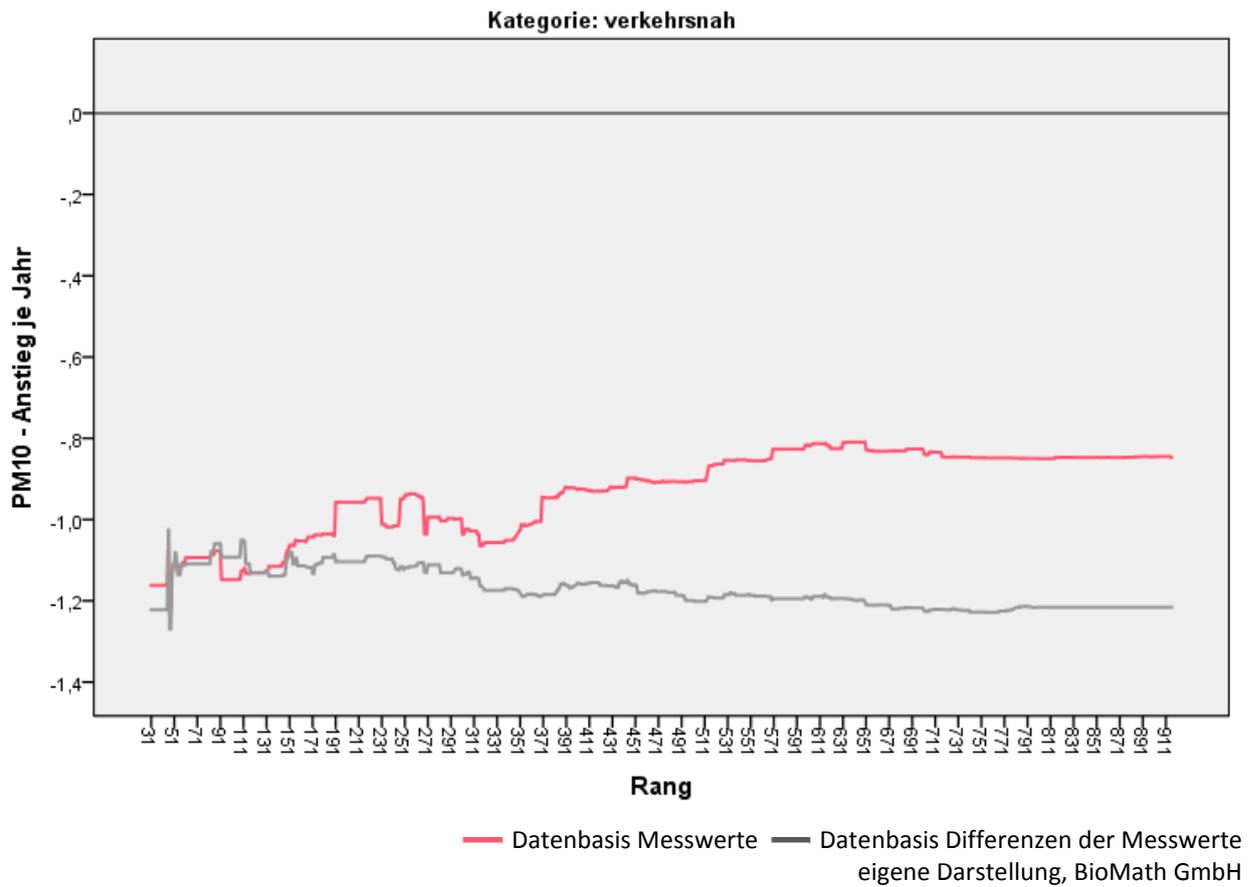


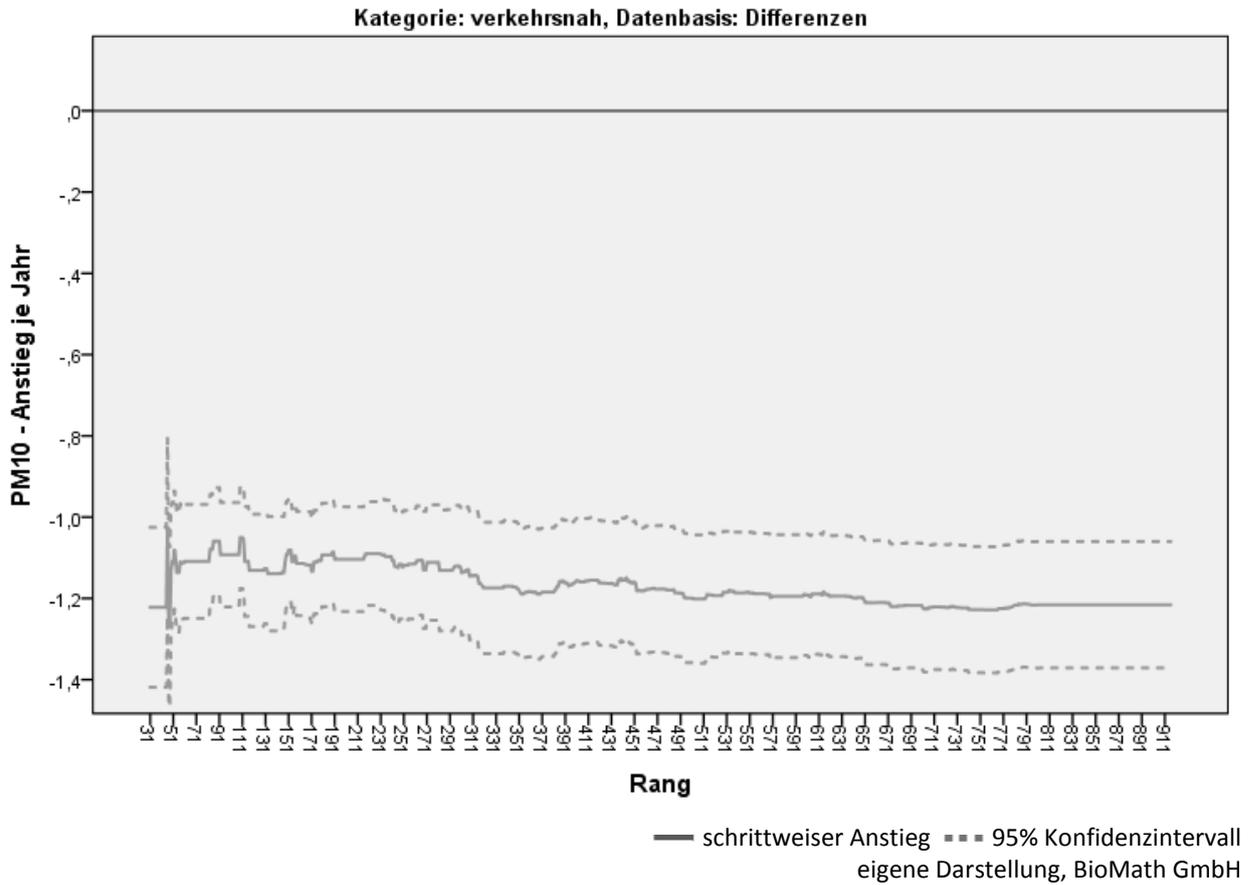
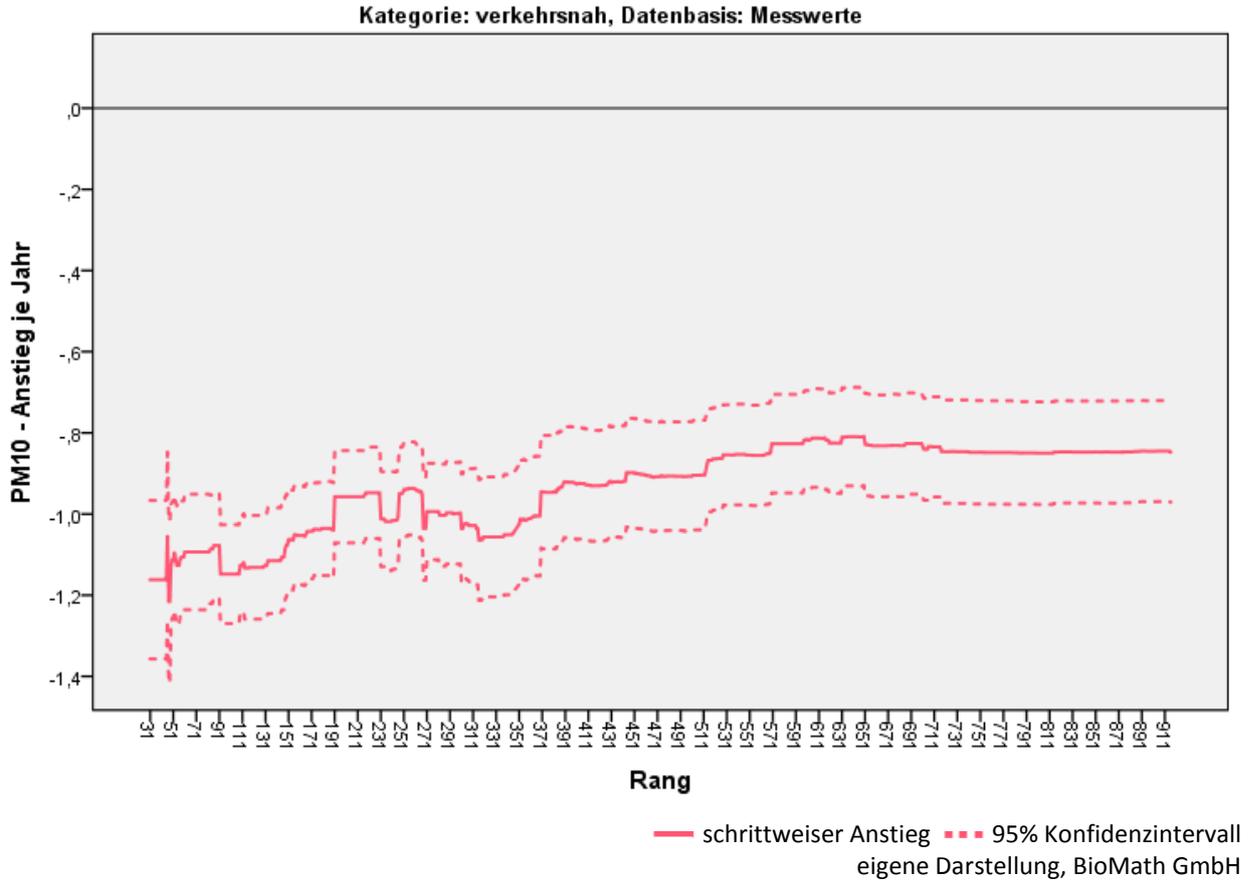
3.1.3.13 PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe städtischer Hintergrund



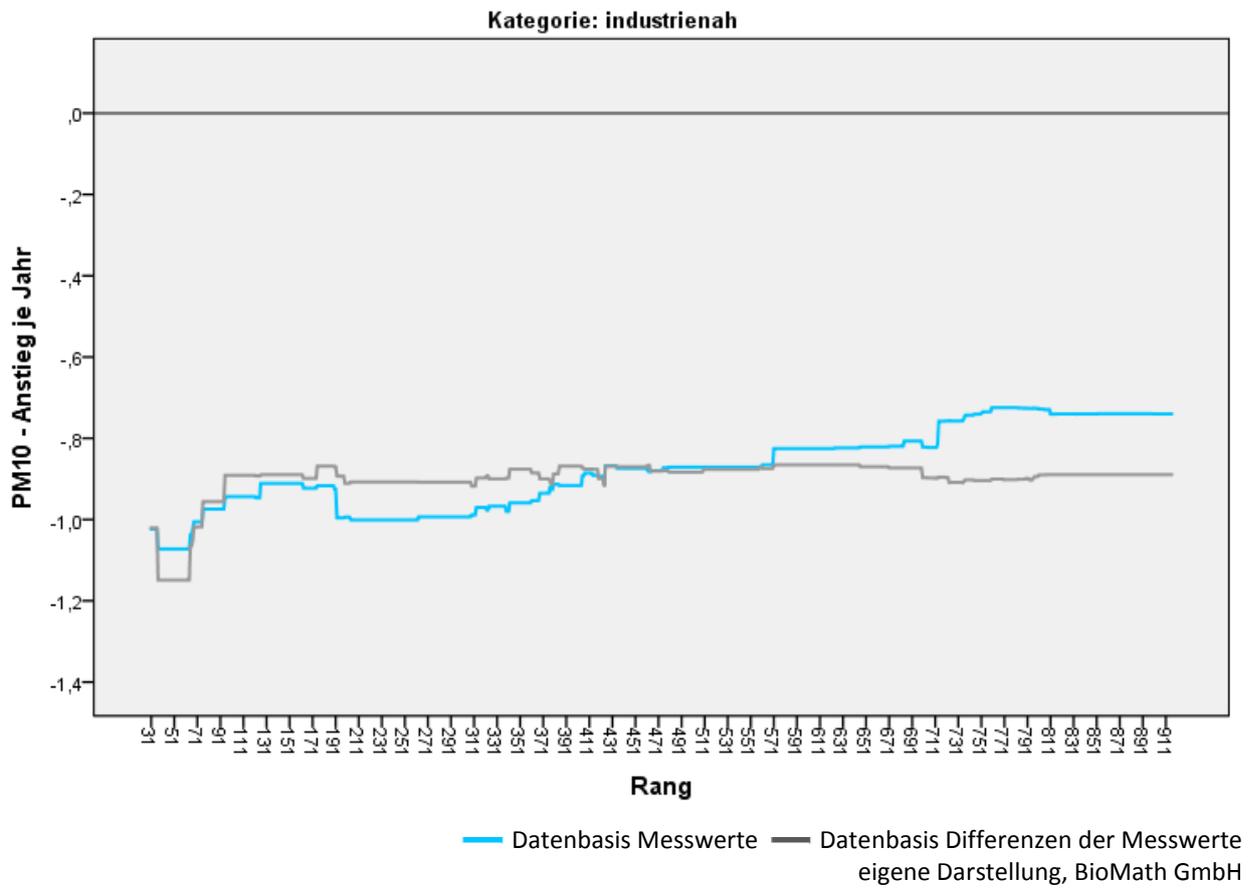


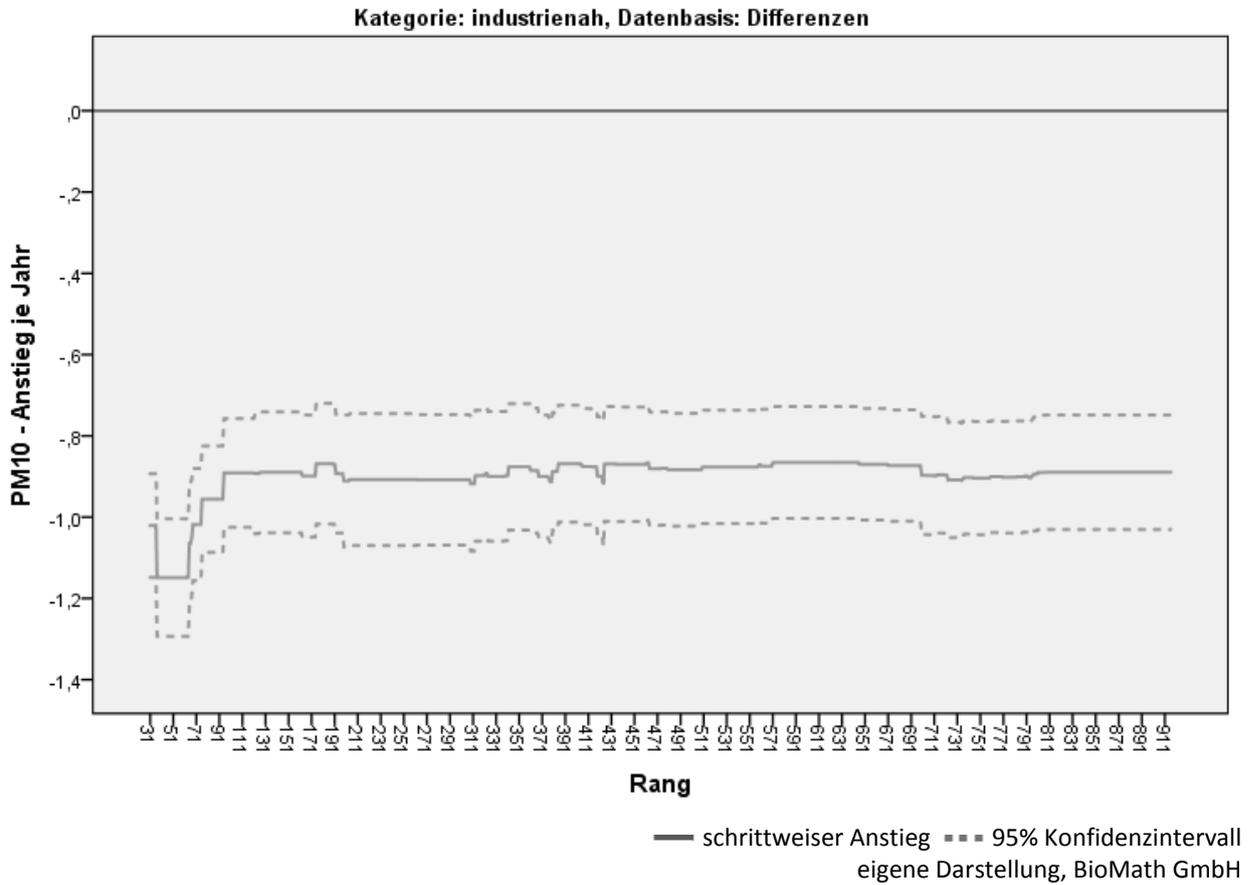
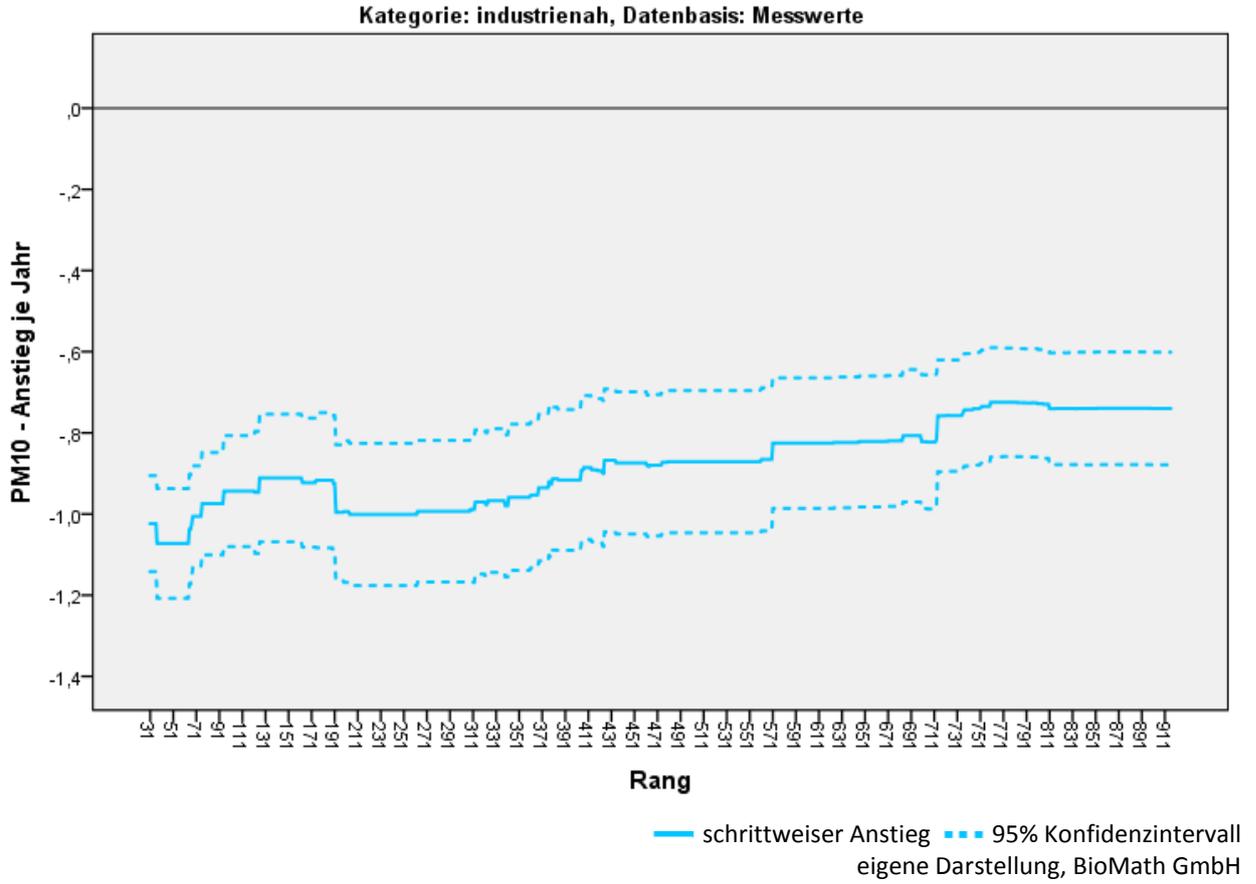
3.1.3.14 PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe verkehrsnah



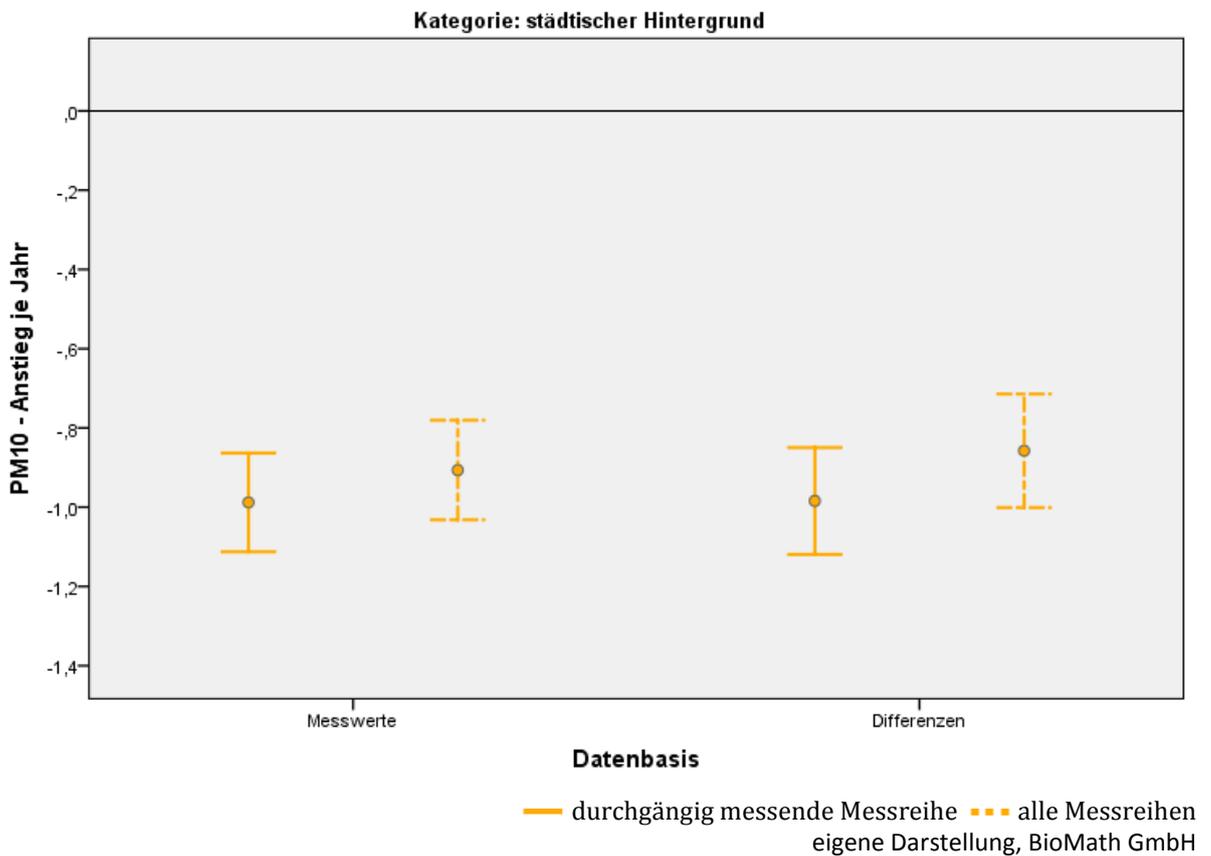
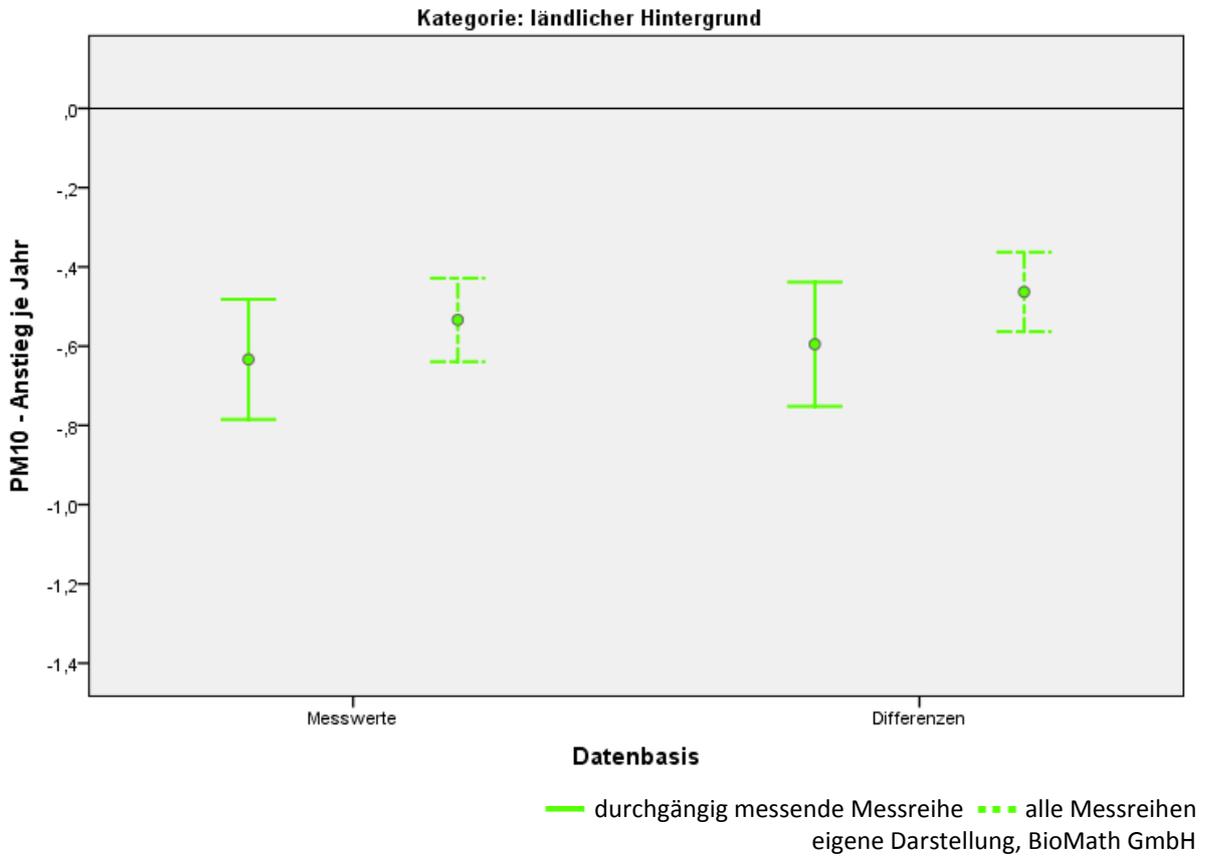


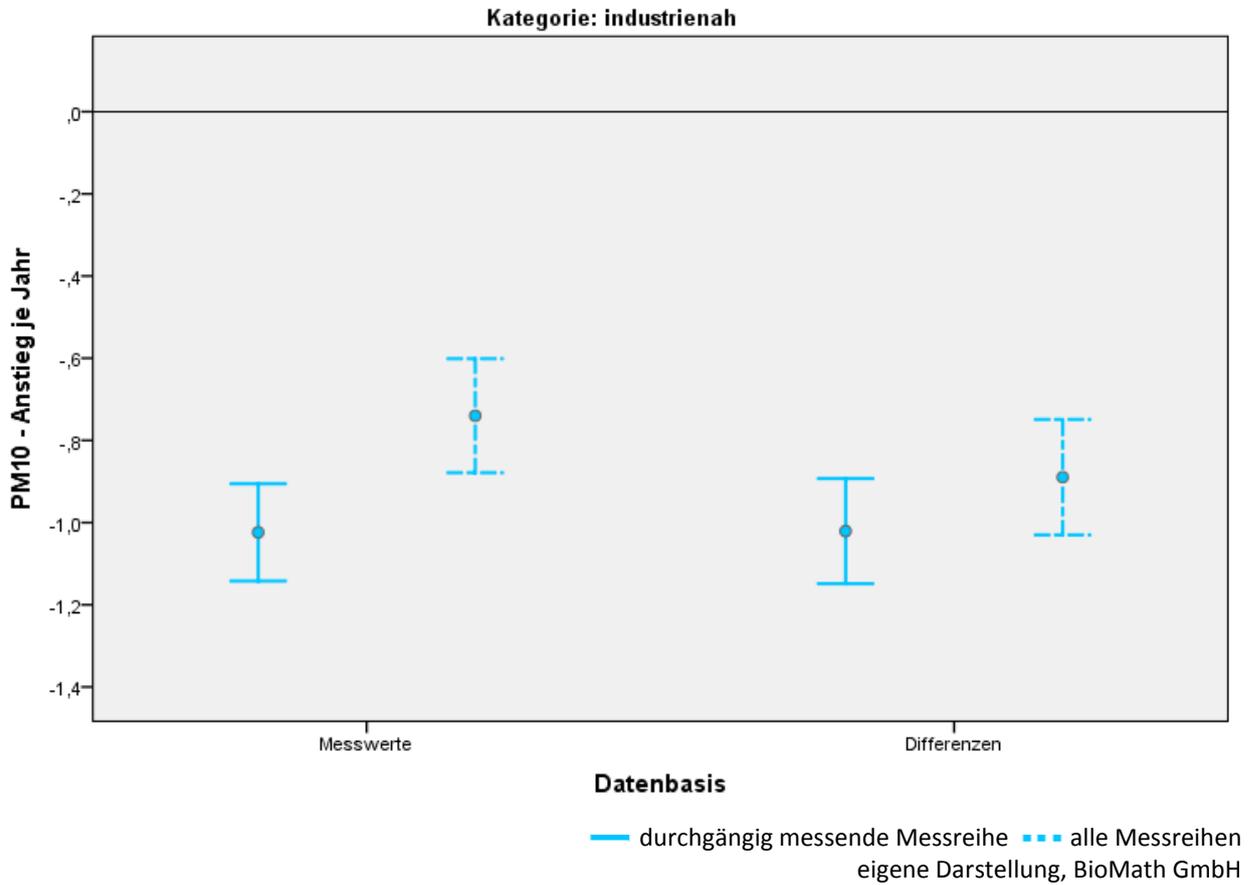
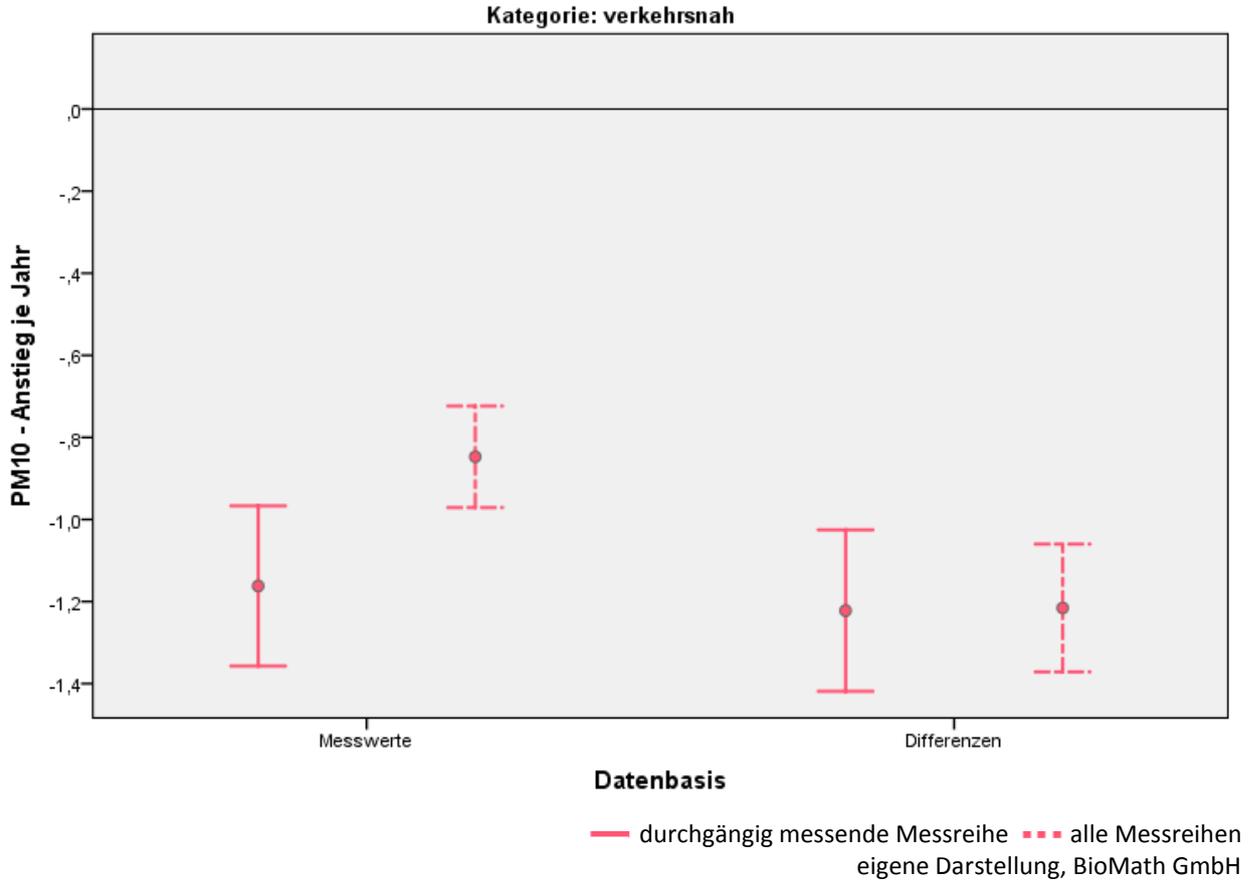
3.1.3.15 PM10 - Trend: schrittweiser Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe industrienah





3.1.3.16 PM10 - Trend: Anstieg ($\mu\text{g}/\text{m}^3$ je Jahr) der Messreihe durchgängig vs. nicht durchgängig messend





3.2 Methoden der Deskriptiven

3.2.1 Rangordnung der Messreihen

Für die schrittweise Berechnung der Trends wurden die Messreihen in eine Rangordnung gebracht. Ziel der Rangordnung ist, dass die sukzessive Trendberechnung möglichst wenig Sprünge aufweist. Dazu wurden die Messreihen nach dem folgenden Schema sortiert:

Tabelle 7: Schema Rangordnung

Stufe	Variable	Reihenfolge	Bedeutung
1	NJahr	absteigend	In der Reihenfolge ganz vorne stehen Messreihen mit möglichst vielen Messwerten. Damit ist sichergestellt, dass durchgängig messende Reihe am Anfang stehen.
2	DiffJahr	aufsteigend	Innerhalb von NJahr werden die Messreihen nach Betriebsjahren sortiert. Damit werden Messreihen, denen Messwerte in der Betriebszeit fehlen, so einsortiert, dass „möglichst durchgängig messende“ Reihen weiter vorne in der Reihenfolge stehen.
3	Abweichung	aufsteigend	Innerhalb von DiffJahr wird die Abweichung vom Betriebsbeginn und Betriebsende zu dem Jahr 2003 (= Mitte zwischen 1990 und 2016) bewertet – je näher die Mitte der Betriebszeit an 2003 ist, desto weiter vorne wird die Reihe einsortiert. Damit wird eine gewisse Gleichmäßigkeit für die schrittweise Berechnung der Trends gewährleistet.
4	MaxJahr	absteigend	Innerhalb von Abweichung werden Messreihen, die später enden, vor den Reihen einsortiert, die früher enden.
5	Stationscode	aufsteigende	Innerhalb von MaxJahr wird nach dem Stationscode sortiert. Diese Sortierung ist willkürlich gewählt, sichert aber die Eindeutigkeit und Nachvollziehbarkeit der Rangordnung.

3.2.2 lineares Regressionsmodell zur Trendberechnung

Der Trend wird mit einem linearen Regressionsmodell bestimmt.

$$m_t = \beta_1 t + \beta_2 \text{ mit } t = 1990, \dots, 2016$$

Dabei ist

m_t Messwert zum Zeitpunkt t , entsprechend der gewählten Datenbasis

β_1 Anstieg der Regressionsgeraden, entspricht der jährlichen Zunahme bzw. Abnahme (Trend)

β_2 Konstante

Sind die Messwerte die Datenbasis für die Berechnung des linearen Trends, so ist m_t der Mittelwert des Jahres t , gebildet aus den Messwerten der einbezogenen Messreihen.

Sind die Differenzen der Messwerte die Datenbasis, so wird m_t wie folgt bestimmt:

Für jede Messreihe mit den Messwerten y_i werden die jährlichen Differenzen $d_i = y_i - y_{i-1}$, beginnend mit dem Jahr 1991, gebildet.

Danach wird der Mittelwert \bar{d}_t der berechneten Differenzen je Jahr über die einbezogenen Messreihen bestimmt.

Aus diesen Mittelwerten wird durch rekursive Summation die Messreihe $m_t = \sum_{j=1991}^t \bar{d}_j$ für die lineare Trendberechnung erzeugt.

3.3 Ergebnisse der deskriptiven Auswertung

Die Deskriptive zeigt das Spektrum der Heterogenität der Stationen und deren Messreihen. Insgesamt kann für alle Schadstoffe festgestellt werden, dass die Daten des ländlichen Hintergrundes eine gute Basis für die Trendberechnung bilden. Für Trendaussagen der Kategorie industriennah ist zu berücksichtigen, dass die Anzahl der Messreihen bei allen Schadstoffen äußerst gering ist. In der Kategorie verkehrsnah gibt es für NO₂ und PM₁₀ einen sehr starken Anstieg der Anzahl der Stationen ab dem Jahr 2005. Ferner streuen in allen Kategorien die Messreihen von PM₁₀ von 1990 bis 2003 untereinander sehr stark.

In Tabelle 8 sind einige Charakteristiken der Luftschadstoffdaten zusammengestellt. Die Homogenität ist auf einer Skala mit drei Stufen bewertet („+“ = gut, „+/-“ = neutral, „-“ = schlecht) bewertet. Die Anstiege beziehen sich auf den Zeitraum von 1990 – 2016 und wurden mit der Datenbasis der Differenzen der Messwerte berechnet.

Tabelle 8: Deskriptive Übersicht und Bewertung

Kategorie	Eigenschaft	Ozon	NO ₂	PM ₁₀
	Flächendeckend in DE durchgängig messend	ab 1995	ab 2000	ab 2000
ländlicher Hintergrund	proz. Anteil der Stationen je Schadstoff	25,4 %	11,9 %	14,0 %
	Homogenität des Stationsbetriebes	+	+	+
	Homogenität der Messreihen	+/-	+/-	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,12	-0,26*	-0,46*
städtischer Hintergrund	proz. Anteil der Stationen je Schadstoff	55,4 %	36,1 %	40,5 %
	Homogenität des Stationsbetriebes	+/-	+/-	+/-
	Homogenität der Messreihen	+	+	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,23*	-0,52*	-0,86*
verkehrsnah	proz. Anteil der Stationen je Schadstoff	12,6 %	45,8 %	35,1 %
	Homogenität des Stationsbetriebes	-	-	-
	Homogenität der Messreihen	+/-	-	-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,36*	-0,67*	-1,22*
industriennah	proz. Anteil der Stationen je Schadstoff	6,6 %	6,2 %	10,4 %
	Homogenität des Stationsbetriebes	+/-	+/-	+/-
	Homogenität der Messreihen	+/-	+/-	+/-
	Anstieg in µg/m ³ je Jahr (Diff. der Messwerte)	+0,10	-0,43*	-0,89*

* signifikant auf dem Niveau von 5%

3.4 Induktiv

Die induktive Auswertung der drei Luftschadstoffe umfasst den Zeitraum von 2000 bis 2016. Für diesen Zeitraum wurden die Jahresmittelwertzeitreihen je Kategorie für

- ▶ die Rohdaten, d.h. für alle Messreihen,
- ▶ die UBA-Auswahl von Messreihen und
- ▶ das Modell der jährlichen Differenzen

berechnet. Weiterhin wurden

- ▶ die Trends der Zeitreihen bestimmt,
- ▶ die Trends auf Signifikanz getestet (einerseits mit Hilfe der linearen Regression und andererseits mit dem Mann-Kendall Test) und

- ▶ die Zeitreihen untereinander deskriptiv verglichen.

3.4.1 Modell der jährlichen Differenzen

In der deskriptiven Auswertung zeigte sich, dass die Heterogenität der Messreihen sowohl bezüglich der Anzahl der Messwerte als auch bezüglich des Messniveaus einen sehr starken Einfluss auf den Verlauf der Jahresmittelwertzeitreihen hat. Damit verbunden ist auch die Berechnung des Trends der Jahresmittelwertzeitreihen.

Das Modell der jährlichen Differenzen reduziert jede Messreihe auf ihre jährlichen Veränderungen. Die jährlichen Veränderungen sind normierte Werte die unabhängig vom Messniveau sind. Damit können für die Bildung der Jahresmittelwertzeitreihen neben den durchgängig messenden Reihen auch die Messreihen berücksichtigt werden, die nicht für jedes Jahr einen Messwert besitzen.

Die Jahresmittelwertzeitreihe für das Modell wird wie folgt berechnet:

1. je Messreihe werden die jährlichen Differenzen gebildet
2. die jährlichen Differenzen werden je Jahr gemittelt
3. die so gebildeten Mittelwerte werden für jedes Jahr sukzessive über die vorherigen Jahre summiert
4. zu den summierten Mittelwerten je Jahr wird der gewichtete Mittelwert des letzten Jahres addiert

3.4.1.1 Herleitung des Modells

Das Modell der jährlichen Differenzen basiert auf der Lösung einer gewöhnlichen, in diesem Fall einer linearen, Differentialgleichung

$$y'(t) = a$$

wobei $y(t)$ die Messwerte der Messreihe und a der Anstieg der Messreihe ist. Die Lösung der Differentialgleichung ist

$$y'(t) = a$$

$$\frac{dy}{dt} = a$$

$$dy = a \cdot dt$$

$$\int dy = \int a \cdot dt$$

$$y(t) = a \cdot t + const$$

wobei die Konstante $const$ frei wählbar ist.

Die Jahresmittelwertzeitreihe für das Modell wird dann wie folgt gebildet.

Es sei y_i ein Wert in einer Messreihe im Jahr i .

1. Die jährliche Differenz dieses Messwertes ist dann

$$d_i = y_i - y_{i-1}.$$

Diese Differenz ist ein Maß für die Änderung des Messwertes, rückwirkend um ein Jahr. Sie entspricht dem Differential $\frac{dy}{dt}$.

2. Der Mittelwert über die jährlichen Differenzen d_i von N Messreihen im Jahr i ist somit

$$\bar{d}_i = \frac{1}{N} \sum_{j=1}^N d_{i,j} .$$

Dieser Mittelwert stellt die mittlere Änderung aller Messwerte eines Jahres zum Vorjahr dar.

3. Für ein Jahr i aus einem jährlichen Zeitintervall $[a, b]$ wird die sukzessive Summe c_i gebildet.

$$c_i = \sum_{j=a}^i \bar{d}_j$$

Führt man diese Summation für jedes Jahr i des Zeitintervalls $[a, b]$ aus, erhält man einen standardisierten Kurvenverlauf, beginnend mit dem Wert \bar{d} des ersten Jahres des Zeitintervalls.

4. Für das letzte Jahr des Zeitintervalls $[a, b]$ wird ein gewichteter Mittelwert m bestimmt.

$$m = \frac{\sum_{j=1}^n y_j g_j}{\sum_{j=1}^n g_j} \text{ mit den Gewichten } g_j = k_j - \left(1 - \frac{k_j}{(b-a)}\right) \text{ wobei}$$

y_j die Messwerte des letzten Jahres des Zeitintervalls,

n die Anzahl der Messwerte im letzten Jahr des Zeitintervalls und

k_j die Anzahl der Werte der Messreihe j im Zeitintervall ist.

Die Wichtung berücksichtigt die Anzahl der Messwerte einer Messreihe. Durchgängig messende Reihen erhalten das größte Gewicht. Die Wichtung wird kleiner je weniger Messwerte die Reihe besitzt.

Abschließend wird der gewichtete Mittelwert m zu den sukzessiv gebildeten Summen c_i addiert.

$$z_i = c_i + m$$

Die Folge der Werte z_i über das Zeitintervall $[a, b]$ stellt die Jahresmittelwertzeitreihe dar.

Für eine reine Trendberechnung genügt es mit der Folge der sukzessiven Summen c_i zu rechnen - sie stellt den standardisierten Kurvenverlauf dar. Mit Hilfe des gewichteten Mittelwertes für das letzte Jahr kann die standardisierte Kurve in eine Jahresmittelwertzeitreihe transformiert werden.

Beispiel

Gegeben seien drei Messreihen für den Zeitraum 2000 bis 2005. Die Messreihe 1 ist durchgehend messend. Den anderen Messreihen fehlt jeweils ein Messwert.

Tabelle 9: Messwerte

Reihe	2000	2001	2002	2003	2004	2005
1	61,2	58,2	59,1	57,5	58,7	56,4
2		56,4	54,3	54,5	53,1	53,6
3	54,7	54,5	52,1		52,2	51,5

Die jährlichen Differenzen werden so gebildet, dass sie rückwirkend zu interpretieren sind. D.h. es gibt für das Jahr 2000 keine Differenz.

Tabelle 10: Differenzen

Reihe	2000	2001	2002	2003	2004	2005
1		-3,0	0,9	-1,6	1,2	-2,3

Reihe	2000	2001	2002	2003	2004	2005
2			-2,1	0,2	-1,4	0,5
3		-0,2	-2,4			-0,7

Tabelle 11: Wichtung

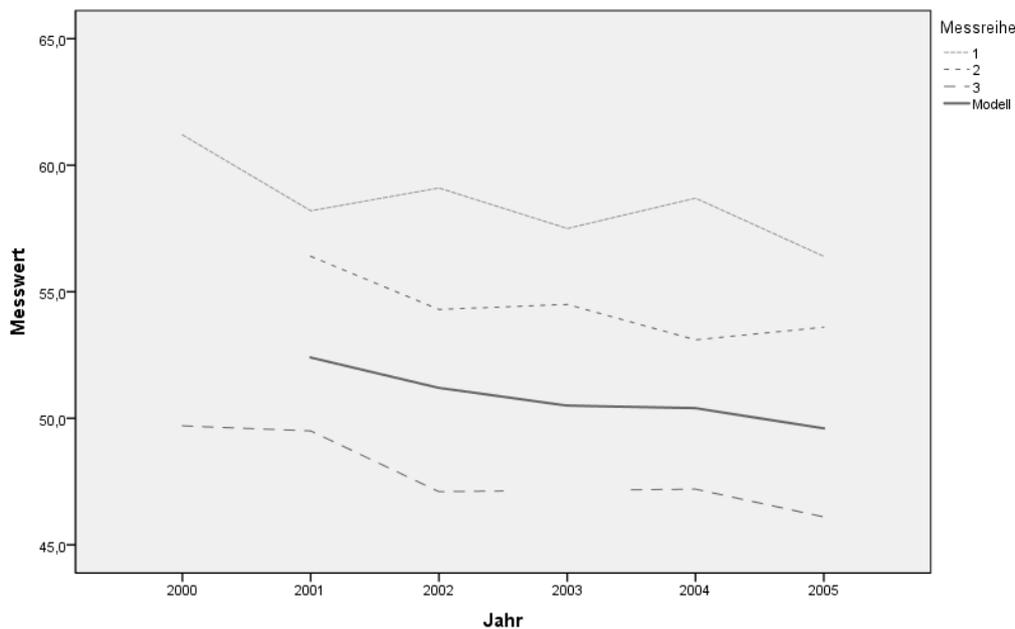
Reihe	Anzahl Messwerte	Länge Zeitintervall	Gewicht
1	6	6	6,2
2	5	6	5,0
3	5	6	5,0

Der gewichtete Mittelwert des Jahres 2005 ist 54,0.

Tabelle 12: Berechnung

	2000	2001	2002	2003	2004	2005
Mittelwert der Differenzen		-1,6	-1,2	-0,7	-0,1	-0,8
sukzessive Summation		-1,6	-2,8	-3,5	-3,6	-4,4
Addition des gewichteten Mittelwertes des Jahres 2005 (entspricht der Mittelwertzeitreihe nach Modell)		52,4	51,2	50,5	50,4	49,6

Abbildung 4: Beispiel - Messreihen und berechnete Mittelwertzeitreihe nach Modell



eigene Darstellung, BioMath GmbH

3.4.1.2 Trendberechnung

Die Trends wurden einerseits mit Hilfe der linearen Regression und andererseits mit dem Mann-Kendall Test für die Jahresmittelwertzeitreihen untersucht.

Lineare Regression

Werden die Punkte z_i einer Jahresmittelwertzeitreihe an eine lineare Funktion

$$z(t) = \beta_0 + \beta_1 t$$

angepasst, so ist der Schätzwert des Parameters β_1 der lineare Trend der Zeitreihe. Um zu ermitteln ob sich dieser lineare Trend signifikant von Null unterscheidet, führt man den Test mit Hilfe eines Konfidenzintervalls durch. Dazu wird zunächst für den Schätzwert des Parameters β_1 , zu einer vorgegebenen Irrtumswahrscheinlichkeit α , das Konfidenzintervall zum Niveau $(1 - \alpha)$ berechnet. Danach wird geprüft ob das Konfidenzintervall die Null überdeckt.

Das Konfidenzintervall zum Niveau $(1 - \alpha)$

überdeckt die Null \Rightarrow der lineare Trend unterscheidet sich nicht signifikant von Null

überdeckt nicht die Null \Rightarrow der lineare Trend unterscheidet sich signifikant von Null

Mann-Kendall Test

Der Mann-Kendal Test ist ein nicht-parametrischer Trend-Test. In der zugrunde liegenden Teststatistik τ wird zunächst jeder Messwert der Zeitreihe mit dem vorherigen Wert und allen nachfolgenden Werten verglichen und die Vorzeichen der Abweichungen ermittelt.

Seien y_j und y_k zwei Messwerte einer Zeitreihe so wird das Vorzeichen mit der Vorzeichenfunktion sgn wie folgt berechnet:

$$sgn(y_j - y_k) = \begin{cases} 1 & ,wenn (y_j - y_k) > 0 \\ 0 & ,wenn (y_j - y_k) = 0 \\ -1 & ,wenn (y_j - y_k) < 0 \end{cases}$$

Danach werden die Vorzeichen aufsummiert und standardisiert. Die so berechnete Teststatistik ist asymptotisch normalverteilt, d.h. die Signifikanz wird über die Standardnormalverteilung ermittelt.

Als nicht-parametrischer Vorzeichentest ist der Mann-Kendall Test unabhängig von der Verteilung der Daten. Allerdings ist er sehr empfindlich, wenn Autokorrelationen in der Zeitreihe vorhanden sind. Eine Autokorrelation liegt vor, wenn Messwerte von vergangenen Messwerten abhängig sind.

3.4.2 Ozon - Zeitraum 2000-2016

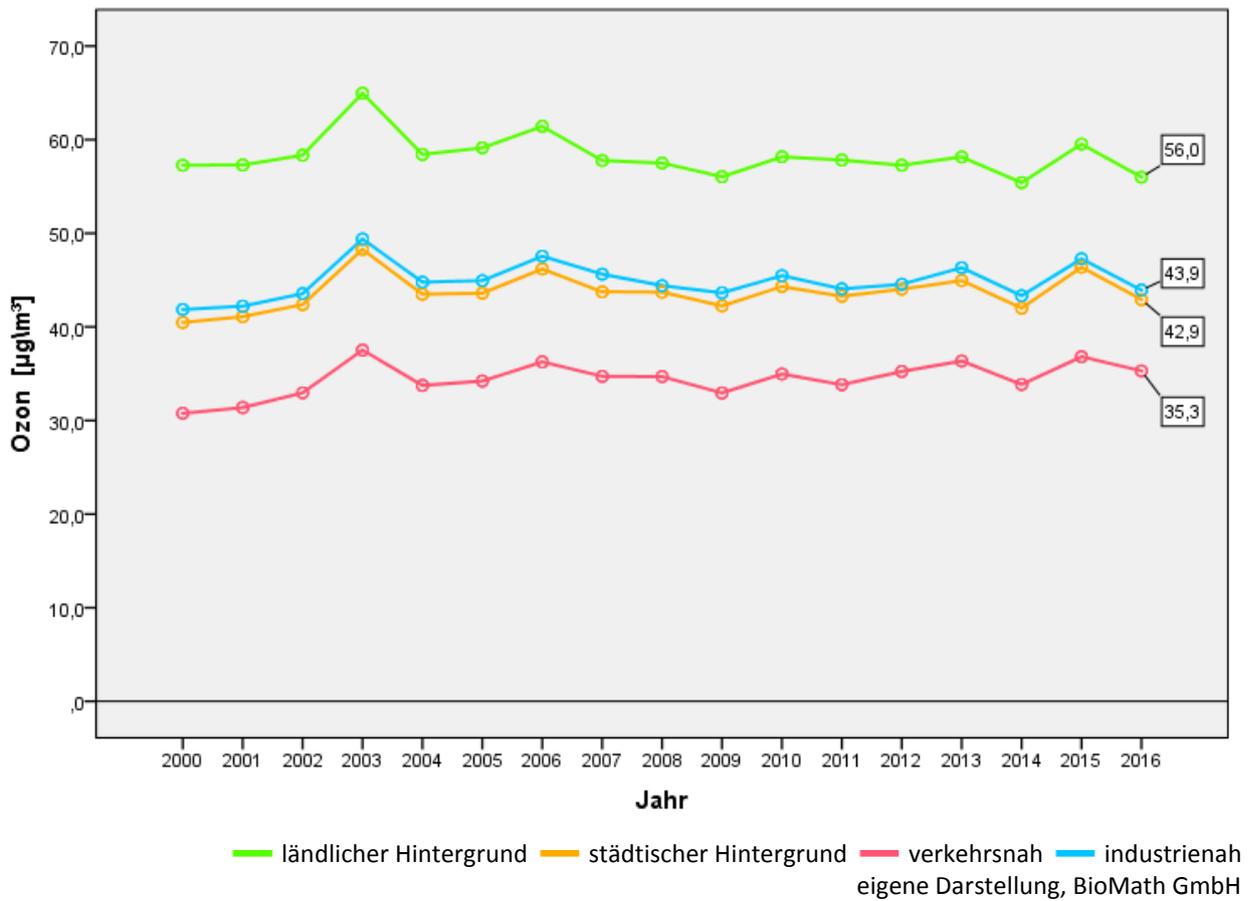
Die folgenden Aussagen beziehen sich auf den Zeitraum von 2000 bis 2016. Dazu wurden für die Modellrechnung alle Messreihen, die in dem Zeitraum von 1999 bis 2016 mindestens zwei aufeinanderfolgende Messwerte generierten, verwendet. Grundlage für die Berechnung der Mittelwertzeitreihe der Rohdaten waren alle Messreihen, die im Zeitraum von 2000 bis 2016 mindestens einen Messwert hatten.

Nur in der Kategorie verkehrsnah und nur für das Modell konnte ein signifikanter Anstieg der Ozonkonzentration um $0,18 \mu\text{g}/\text{m}^3$ je Jahr nachgewiesen werden (siehe 3.4.2.1, 3.4.2.2 und 3.4.2.4). In allen anderen Kategorien zeigten sich für das Modell, für die Rohdaten und für die UBA-Auswahl keine signifikanten Trends.

Die Abweichungen der Mittelwertzeitreihen zwischen Modell, UBA-Auswahl und Rohdaten war im Mittel immer kleiner als $0,71 \mu\text{g}/\text{m}^3$, d.h. sie lagen alle dicht beieinander.

In den Mittelwertzeitreihen der Kategorie verkehrsnah widerspiegelte sich die Heterogenität der Daten (siehe 3.4.2.3). Die Abweichung der Mittelwertzeitreihen zwischen Modell, UBA-Auswahl und Rohdaten betrug im Mittel bis zu $1,35 \mu\text{g}/\text{m}^3$. Die größten Abweichungen waren im Zeitraum von 2000 bis 2009, am stärksten im Jahr 2003, zu verzeichnen.

Abbildung 5: Jahresmittelwertzeitreihen des Modells für Ozon

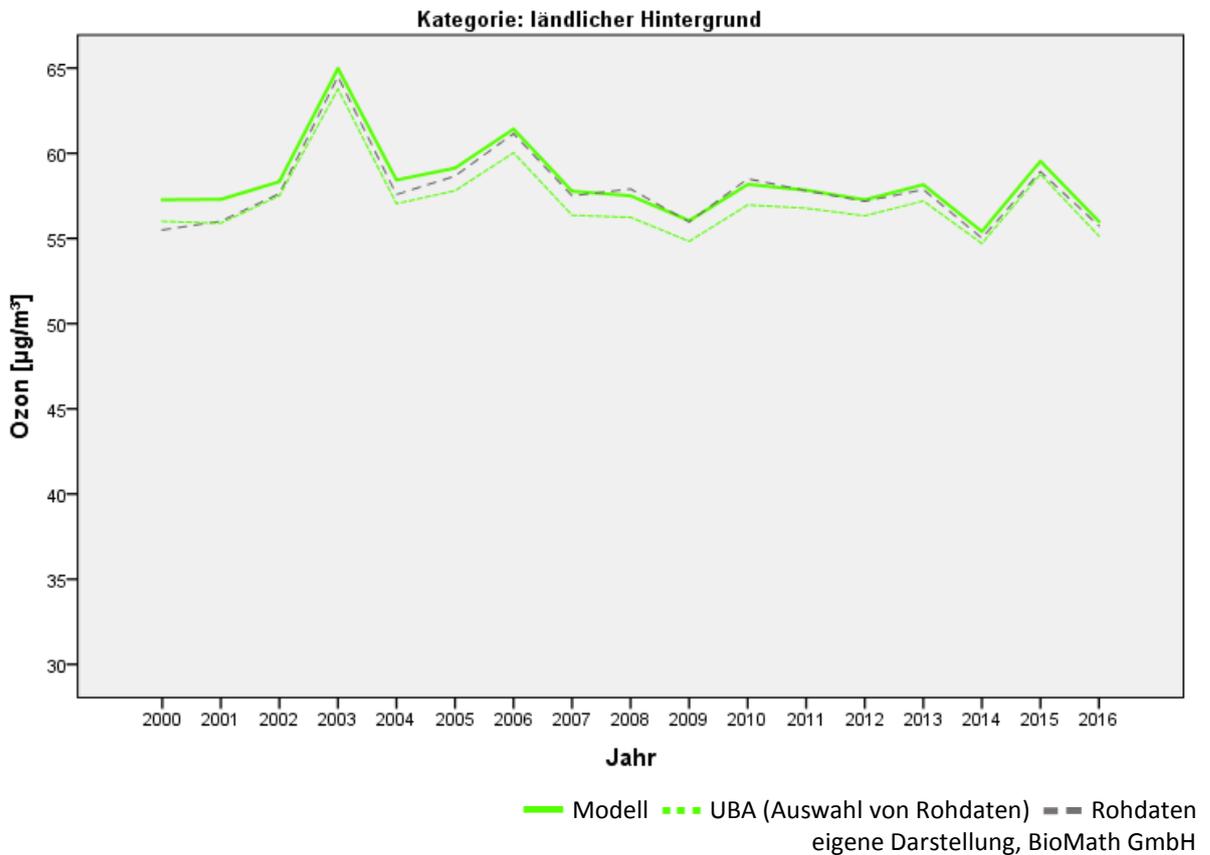


Das Modell weist nur in der Kategorie verkehrsnah einen signifikanten Trend von $0,18 \mu\text{g}/\text{m}^3$ je Jahr aus. In allen anderen Kategorien waren keine signifikanten Trends vorhanden.

Anmerkung

Die Messwerte des Jahres 2003 beeinflussen die Trendaussagen. Entfernt man die Werte für 2003 verstärkt sich der signifikante Anstieg der Ozonkonzentration in der Kategorie verkehrsnah auf $0,24 \mu\text{g}/\text{m}^3$ je Jahr. Bei allen anderen Kategorien wäre auch kein Trend zu verzeichnen.

3.4.2.1 Ozon - ländlicher Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	123	100,8	1452	100,8
UBA	49	40,2	833	57,8
Rohdaten	122	100,0	1441	100,0

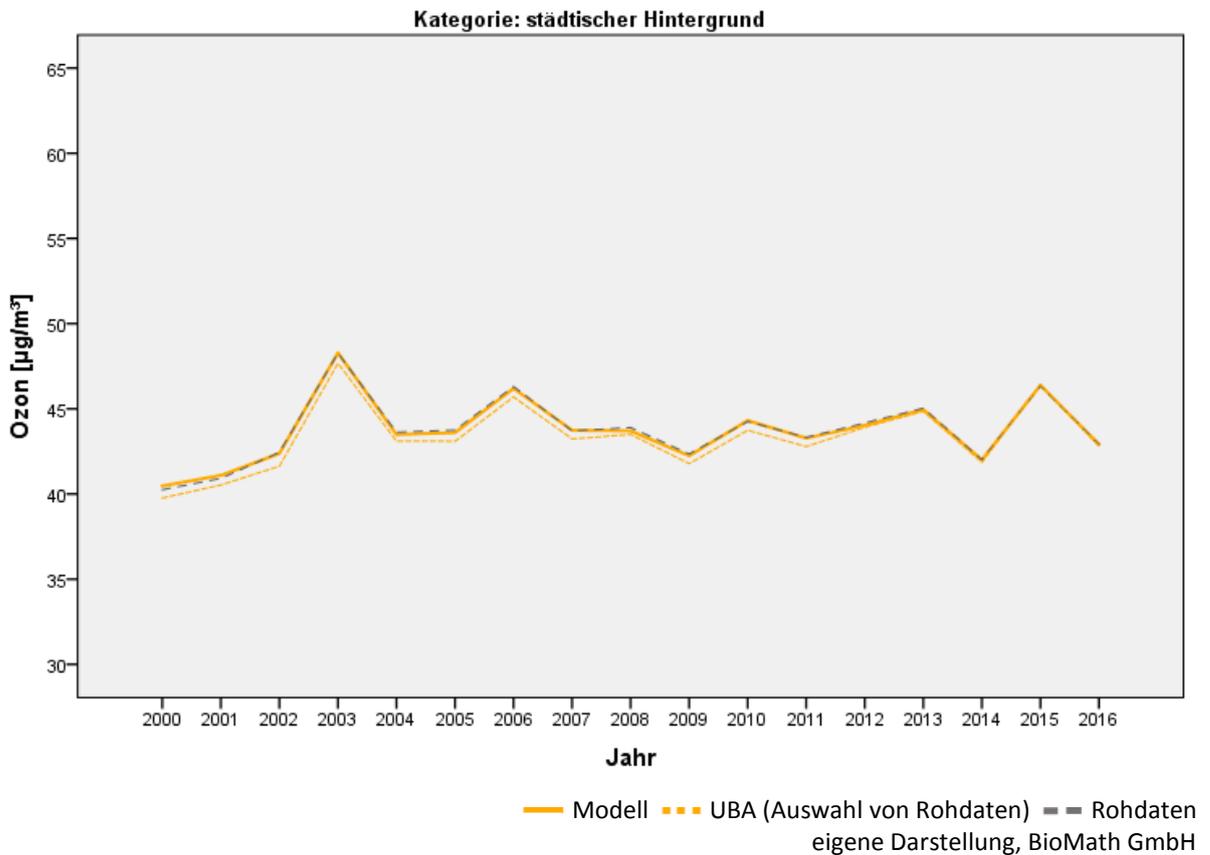
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,1912	0,2842	-0,1542	-0,3820	0,0736		-0,4691	0,1608	
UBA	-0,1471	0,41	-0,1225	-0,3501	0,1051		-0,4371	0,1921	
Rohdaten	-0,0588	0,7417	-0,0921	-0,3350	0,1509		-0,4279	0,2438	

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	0,42	-0,40	1,76
UBA - Rohdaten	-0,71	-1,67	0,51
Modell - UBA	1,13	0,70	1,41

3.4.2.2 Ozon - städtischer Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	254	98,8	2926	101,1
UBA	97	37,7	1649	57,0
Rohdaten	257	100,0	2893	100,0

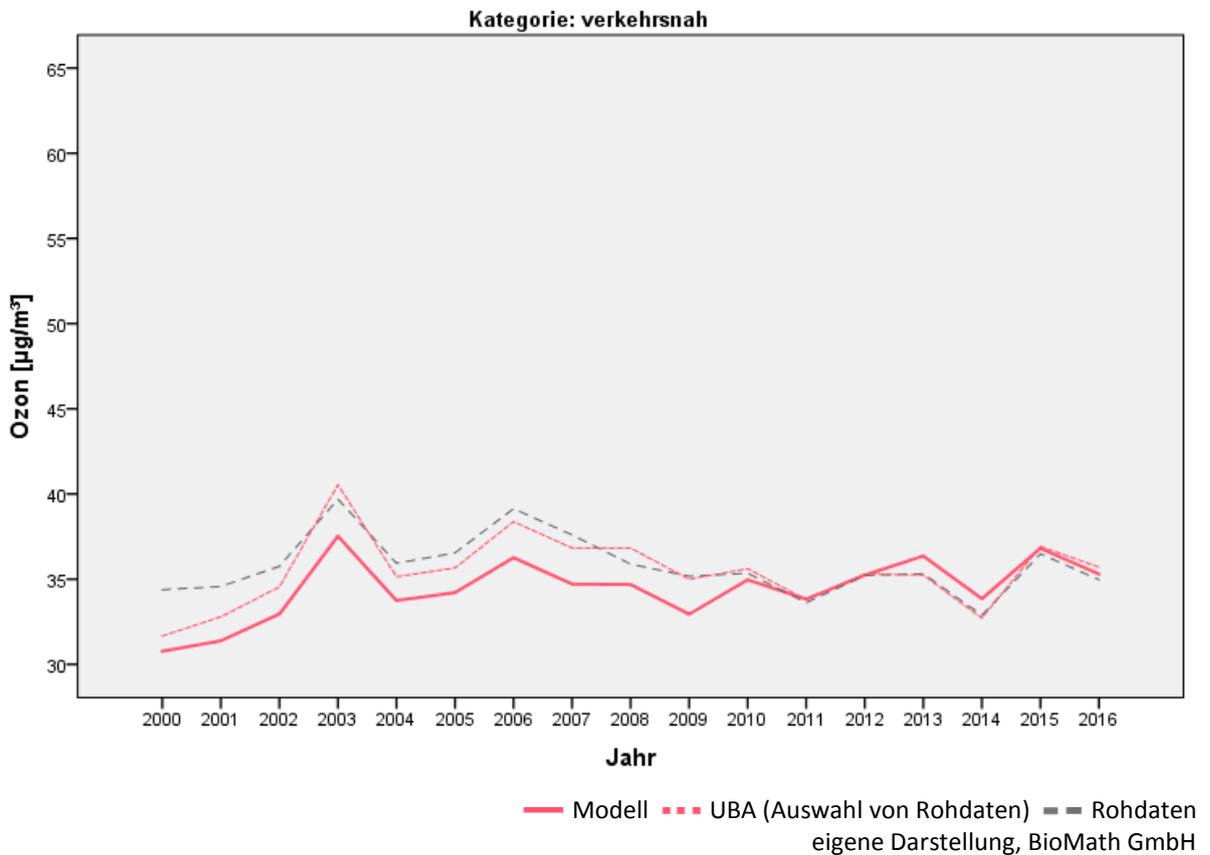
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	0,2206	0,2165	0,0823	-0,1260	0,2907		-0,2057	0,3704	
UBA	0,2941	0,0994	0,1221	-0,0883	0,3324		-0,1687	0,4129	
Rohdaten	0,2206	0,2165	0,0882	-0,1239	0,3002		-0,2050	0,3813	

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	-0,03	-0,19	0,22
UBA - Rohdaten	-0,42	-0,81	0,08
Modell - UBA	0,39	-0,06	0,74

3.4.2.3 Ozon - verkehrsnah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	44	84,6	333	99,4
UBA	16	30,8	209	62,4
Rohdaten	52	100,0	335	100,0

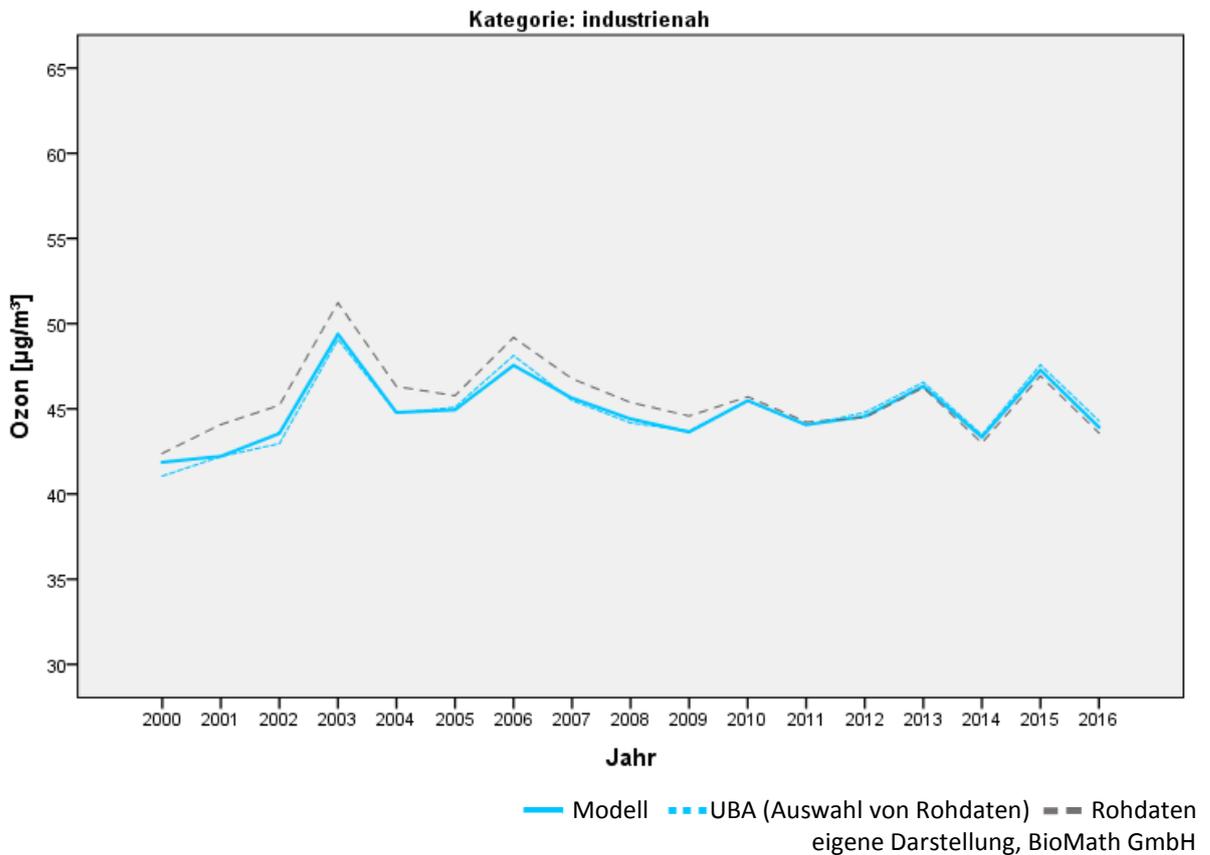
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	0,4265	0,0169	0,1819	0,0118	0,3520	Ja	-0,0532	0,4170	
UBA	0,1029	0,5641	0,0279	-0,2050	0,2607		-0,2941	0,3498	
Rohdaten	-0,1912	0,2842	-0,1046	-0,2868	0,0777		-0,3565	0,1474	

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	-1,35	-3,62	1,04
UBA - Rohdaten	-0,34	-2,72	0,94
Modell - UBA	-1,01	-3,02	1,14

3.4.2.4 Ozon - industrienah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	27	84,4	266	99,6
UBA	10	31,3	170	63,7
Rohdaten	32	100,0	267	100,0

Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	0,1324	0,4584	0,0628	-0,1455	0,2711		-0,2252	0,3508	
UBA	0,1912	0,2842	0,1080	-0,1112	0,3272		-0,1951	0,4110	
Rohdaten	-0,0882	0,6211	-0,0730	-0,3078	0,1618		-0,3976	0,2517	

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	-0,71	-1,87	0,34
UBA - Rohdaten	-0,72	-2,25	0,71
Modell - UBA	0,01	-0,57	0,81

3.4.3 NO₂ - Zeitraum 2000-2016

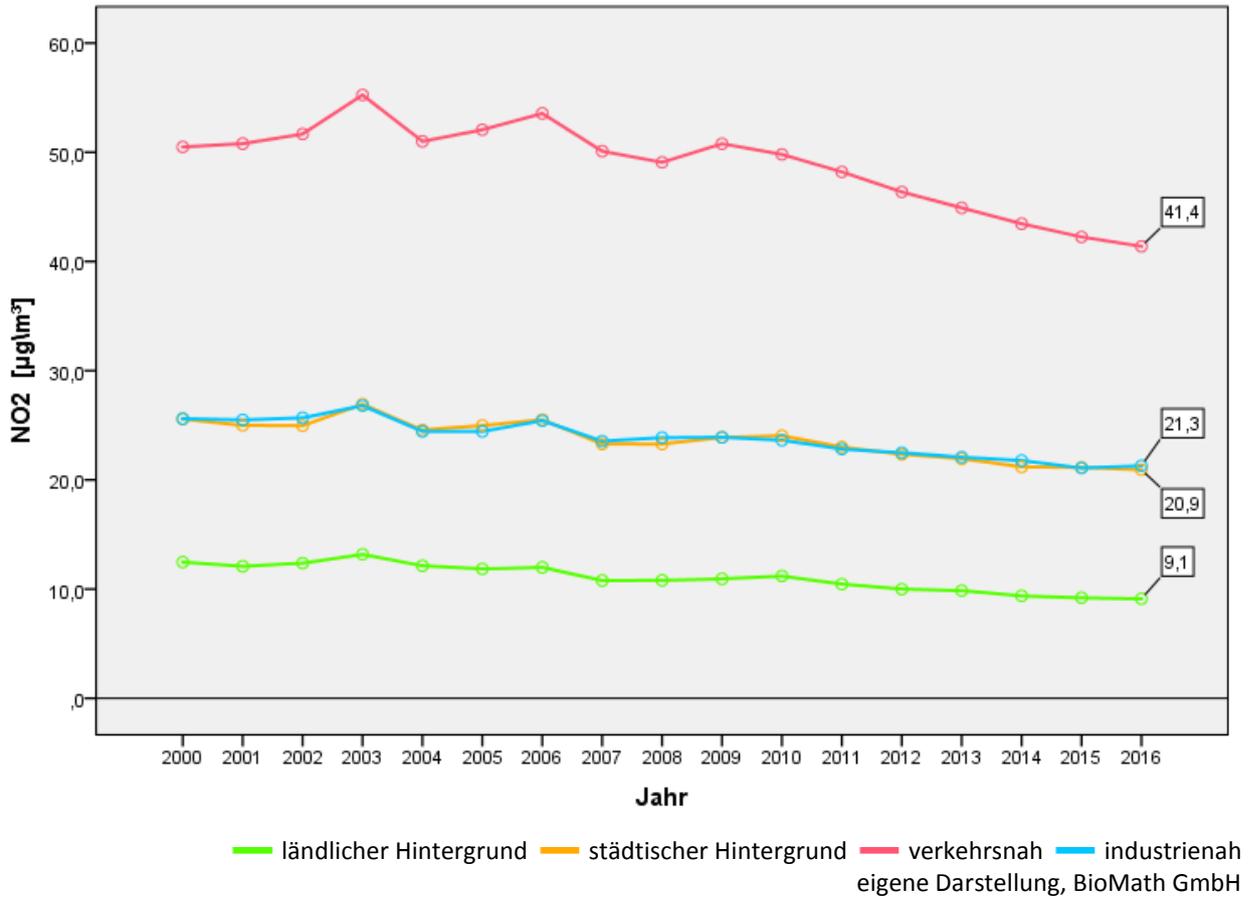
Die folgenden Aussagen beziehen sich auf den Zeitraum von 2000 bis 2016. Dazu wurden für die Modellrechnung alle Messreihen, die in dem Zeitraum von 1999 bis 2016 mindestens zwei aufeinanderfolgende Messwerte generierten, verwendet. Grundlage für die Berechnung der Mittelwertzeitreihe der Rohdaten waren alle Messreihen, die im Zeitraum von 2000 bis 2016 mindestens einen Messwert hatten.

In allen Kategorien konnte für das Modell und die UBA-Auswahl ein signifikantes Sinken der NO₂-Konzentration nachgewiesen werden (siehe 3.4.3.1 bis 3.4.3.4). Die Rohdaten zeigten im ländlichen und städtischen Hintergrund sowie in der Kategorie industrienah ebenfalls einen signifikant fallenden Trend. Aber in der Kategorie verkehrsnah wiesen die Rohdaten keinen signifikanten Trend aus.

In den Mittelwertzeitreihen der Kategorie verkehrsnah widerspiegelt sich ebenfalls die Heterogenität der Daten, insbesondere die unterschiedlichen Messniveaus der Reihen. Dies ist deutlich im Verlauf der Zeitreihe der Rohdaten zu erkennen (siehe 3.4.2.3). Die Mittelwertzeitreihe der UBA-Auswahl wich von den Rohdaten im Mittel um $0,78 \mu\text{g}/\text{m}^3$, die des Modells um $3,89 \mu\text{g}/\text{m}^3$ ab. Die Spannweite der Abweichungen war aber gleich groß. Die größten Abweichungen zu den Rohdaten waren im Zeitraum von 2000 bis 2006.

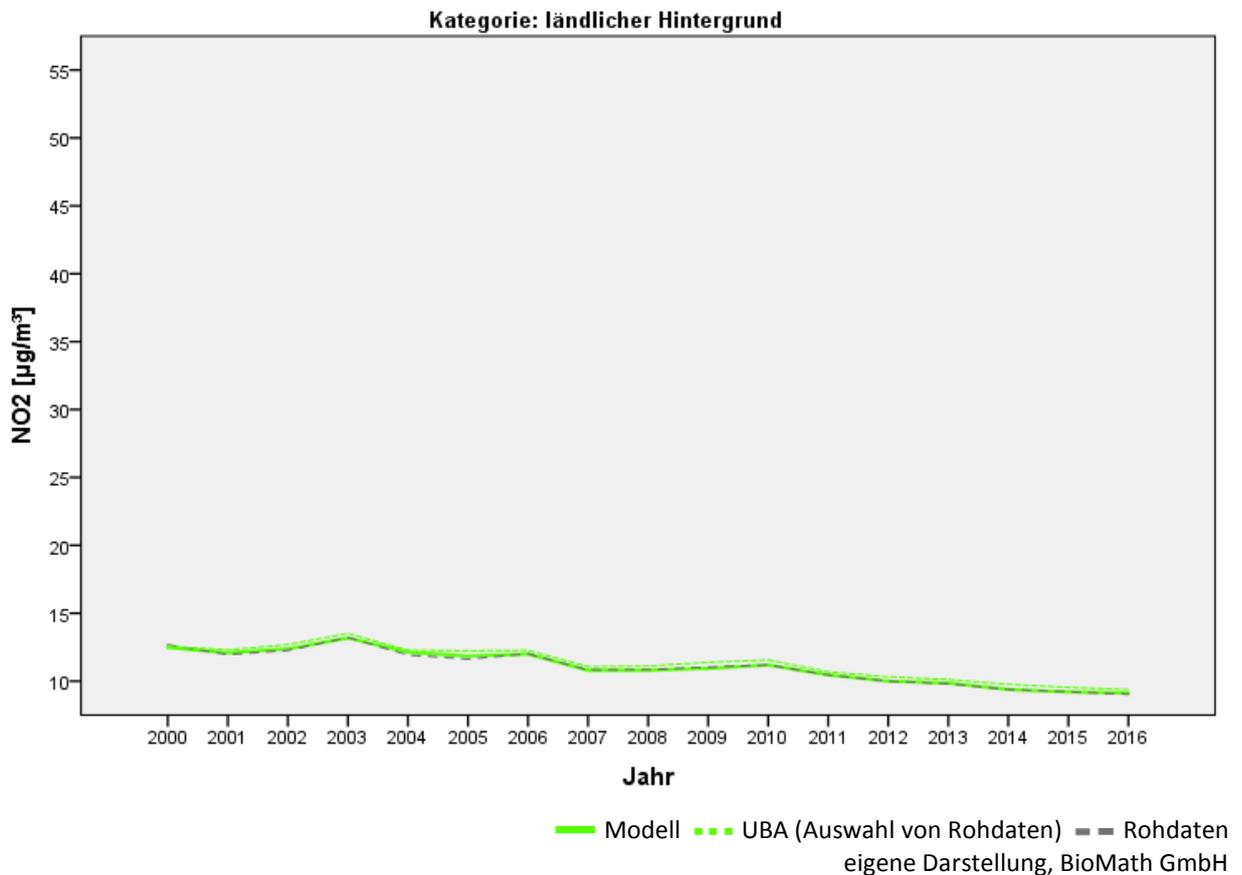
In der Kategorie industrienah lagen die Abweichungen der verschiedenen Mittelwertzeitreihen bei maximal $1,19 \mu\text{g}/\text{m}^3$, d.h. diese Unterschiede können als gering eingeschätzt werden.

Abbildung 6: Jahresmittelwertzeitreihen des Modells für NO₂



Das Modell weist einen signifikanten Trend von $-0,23 \mu\text{g}/\text{m}^3$ je Jahr für den ländlichen Hintergrund, von $-0,32 \mu\text{g}/\text{m}^3$ je Jahr für den städtischen Hintergrund, von $-0,66 \mu\text{g}/\text{m}^3$ je Jahr für die Kategorie verkehrsnahe und von $-0,31 \mu\text{g}/\text{m}^3$ je Jahr für die Kategorie industrienah aus.

3.4.3.1 NO2 - ländlicher Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	121	108,0	1378	103,1
UBA	54	48,2	918	68,7
Rohdaten	112	100,0	1336	100,0

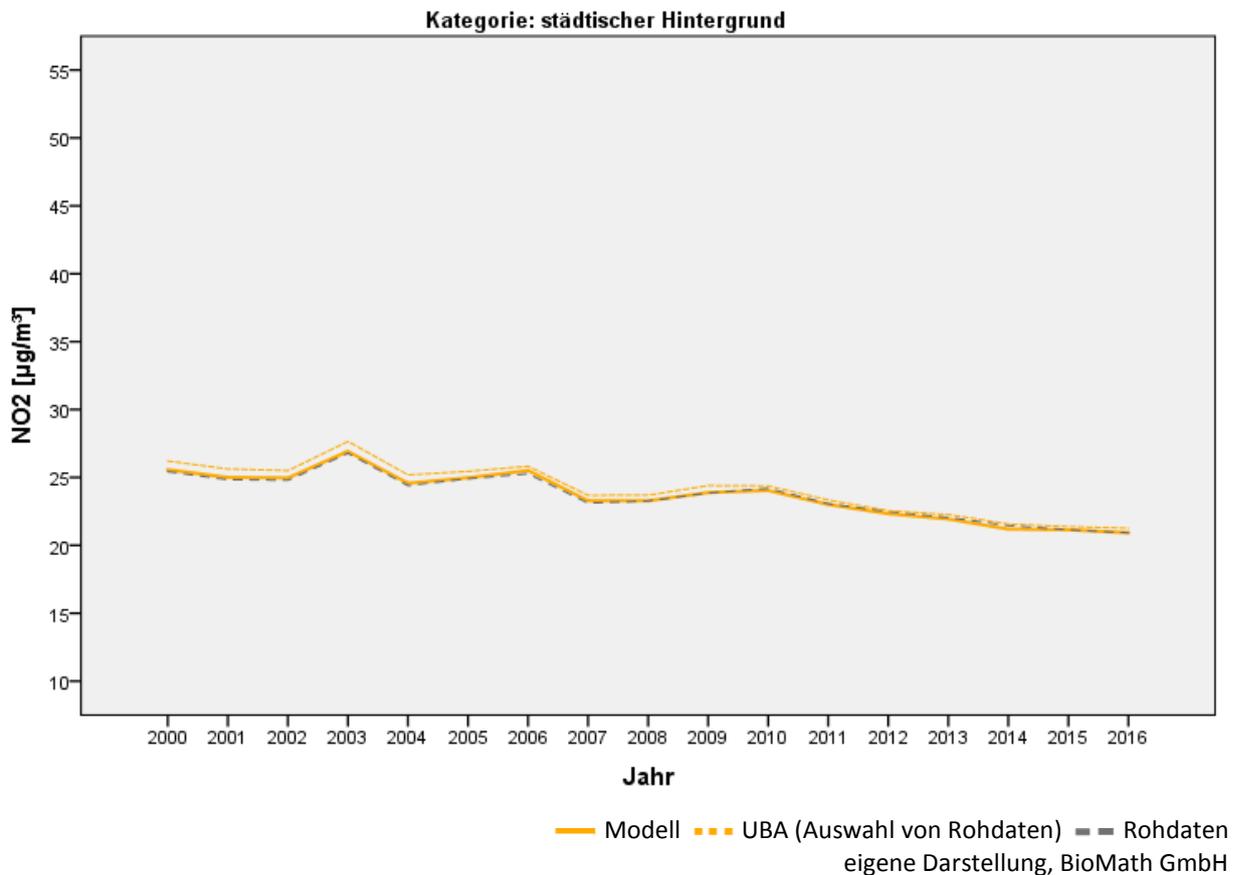
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,8235	<,0001	-0,2348	-0,2786	-0,1911	Ja	-0,2954	-0,1743	Ja
UBA	-0,8235	<,0001	-0,2279	-0,2755	-0,1803	Ja	-0,2937	-0,1621	Ja
Rohdaten	-0,8088	<,0001	-0,2337	-0,2786	-0,1889	Ja	-0,2957	-0,1717	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	0,02	-0,21	0,22
UBA - Rohdaten	0,31	-0,11	0,61
Modell - UBA	-0,29	-0,45	-0,09

3.4.3.2 NO₂ - städtischer Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	330	107,5	3425	105,1
UBA	126	41,0	2142	65,7
Rohdaten	307	100,0	3258	100,0

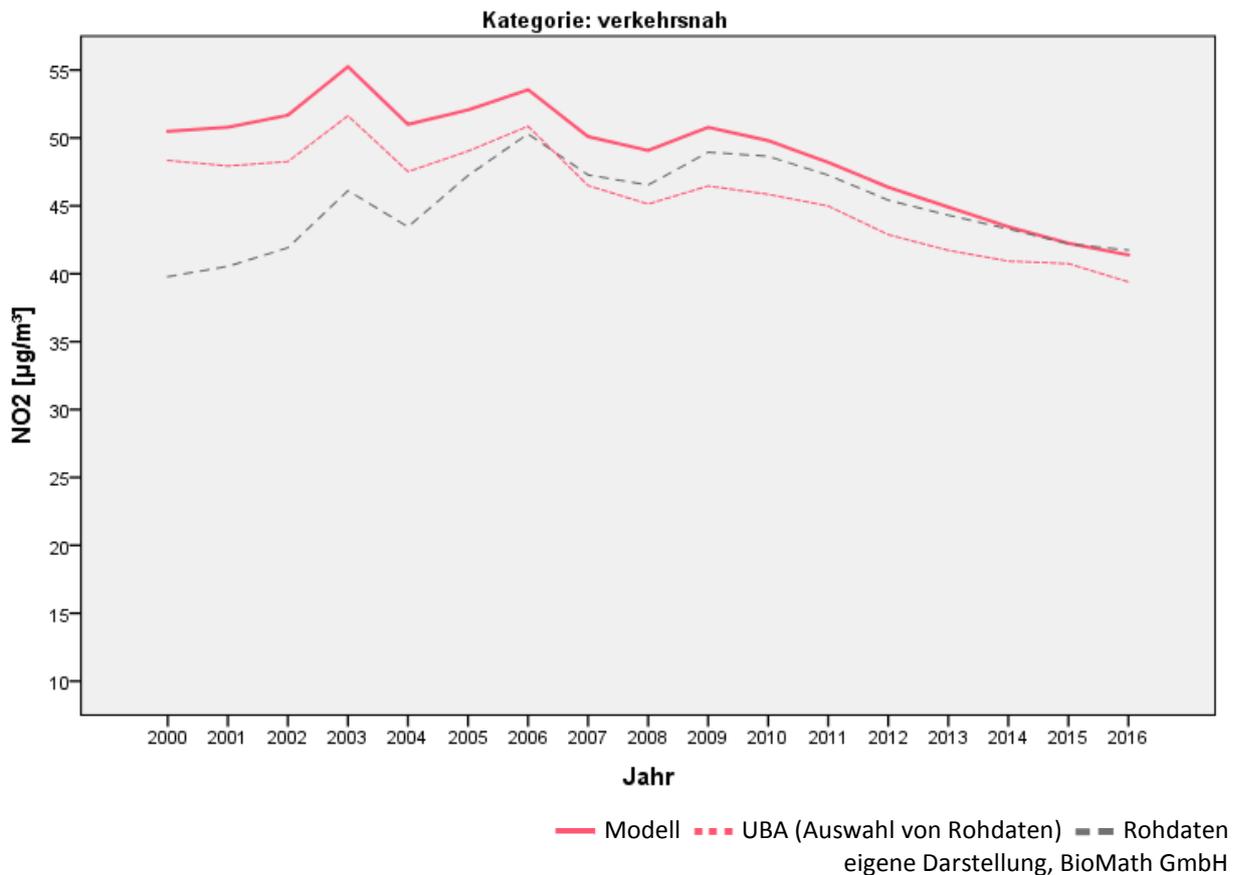
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,7941	<,0001	-0,3159	-0,3957	-0,2361	Ja	-0,4262	-0,2056	Ja
UBA	-0,8088	<,0001	-0,3402	-0,4209	-0,2596	Ja	-0,4517	-0,2287	Ja
Rohdaten	-0,7647	<,0001	-0,2943	-0,3743	-0,2143	Ja	-0,4049	-0,1837	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	0,03	-0,27	0,24
UBA - Rohdaten	0,47	0,07	0,87
Modell - UBA	-0,44	-0,73	-0,22

3.4.3.3 NO₂ - verkehrsnah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	389	83,3	3161	97,5
UBA	36	7,7	612	18,9
Rohdaten	467	100,0	3242	100,0

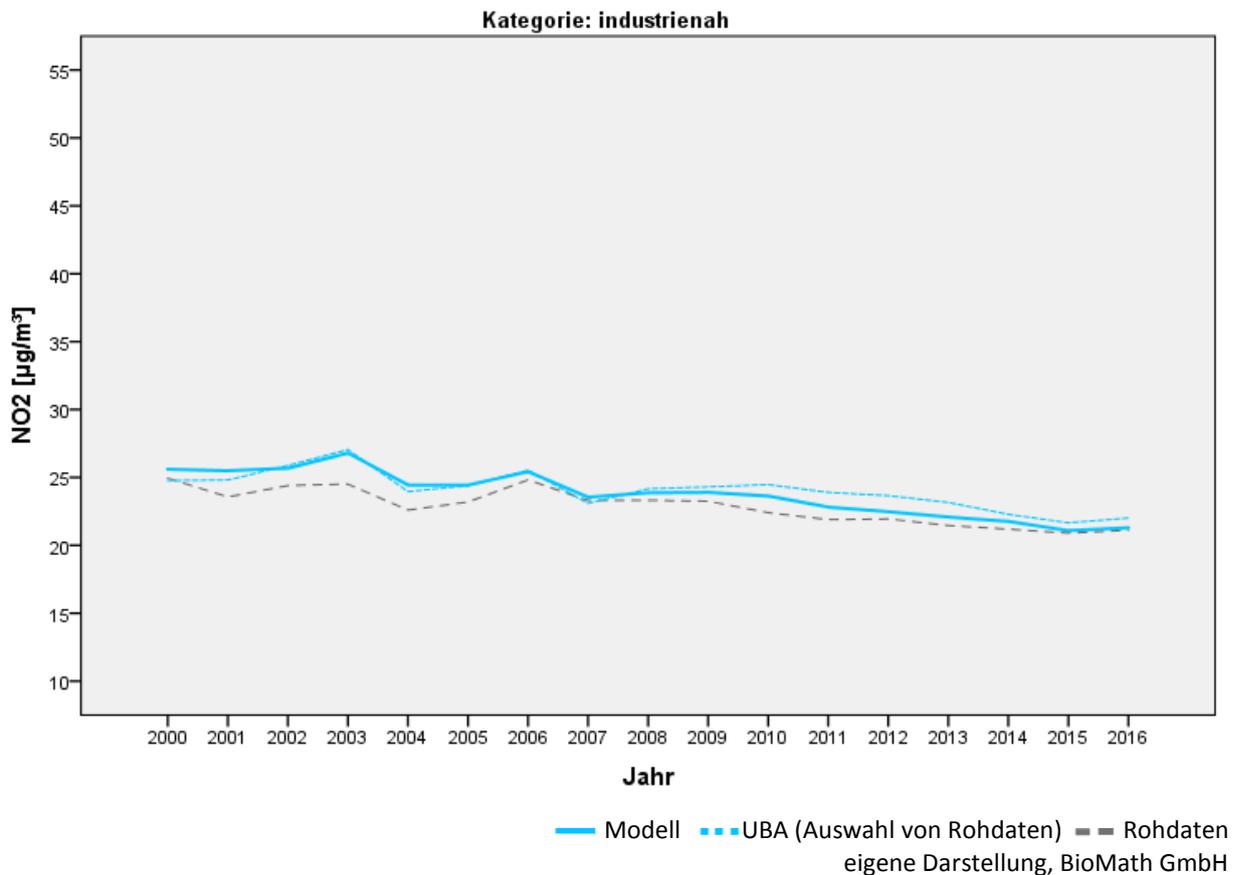
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,6912	0,0001	-0,6631	-0,8907	-0,4354	Ja	-0,9778	-0,3484	Ja
UBA	-0,7794	<,0001	-0,6280	-0,8142	-0,4418	Ja	-0,8854	-0,3706	Ja
Rohdaten	0,0294	0,8691	0,0803	-0,2580	0,4186		-0,3874	0,5480	

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	3,89	-0,35	10,70
UBA - Rohdaten	0,78	-2,80	8,55
Modell - UBA	3,11	1,50	4,32

3.4.3.4 NO₂ - industrienah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	50	80,6	425	100,0
UBA	19	30,6	290	68,2
Rohdaten	62	100,0	425	100,0

Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,8235	<,0001	-0,3145	-0,3771	-0,2519	Ja	-0,4010	-0,2279	Ja
UBA	-0,6324	0,0004	-0,2175	-0,3105	-0,1245	Ja	-0,3461	-0,0889	Ja
Rohdaten	-0,7353	<,0001	-0,2332	-0,2994	-0,1670	Ja	-0,3247	-0,1417	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	0,91	0,18	2,29
UBA - Rohdaten	1,19	-0,22	2,55
Modell - UBA	-0,28	-1,18	0,84

3.4.4 PM10 - Zeitraum 2000-2016

Die folgenden Aussagen beziehen sich auf den Zeitraum von 2000 bis 2016. Dazu wurden für die Modellrechnung alle Messreihen, die in dem Zeitraum von 1999 bis 2016 mindestens zwei aufeinanderfolgende Messwerte generierten, verwendet. Grundlage für die Berechnung der Mittelwertzeitreihe der Rohdaten waren alle Messreihen, die im Zeitraum von 2000 bis 2016 mindestens einen Messwert hatten.

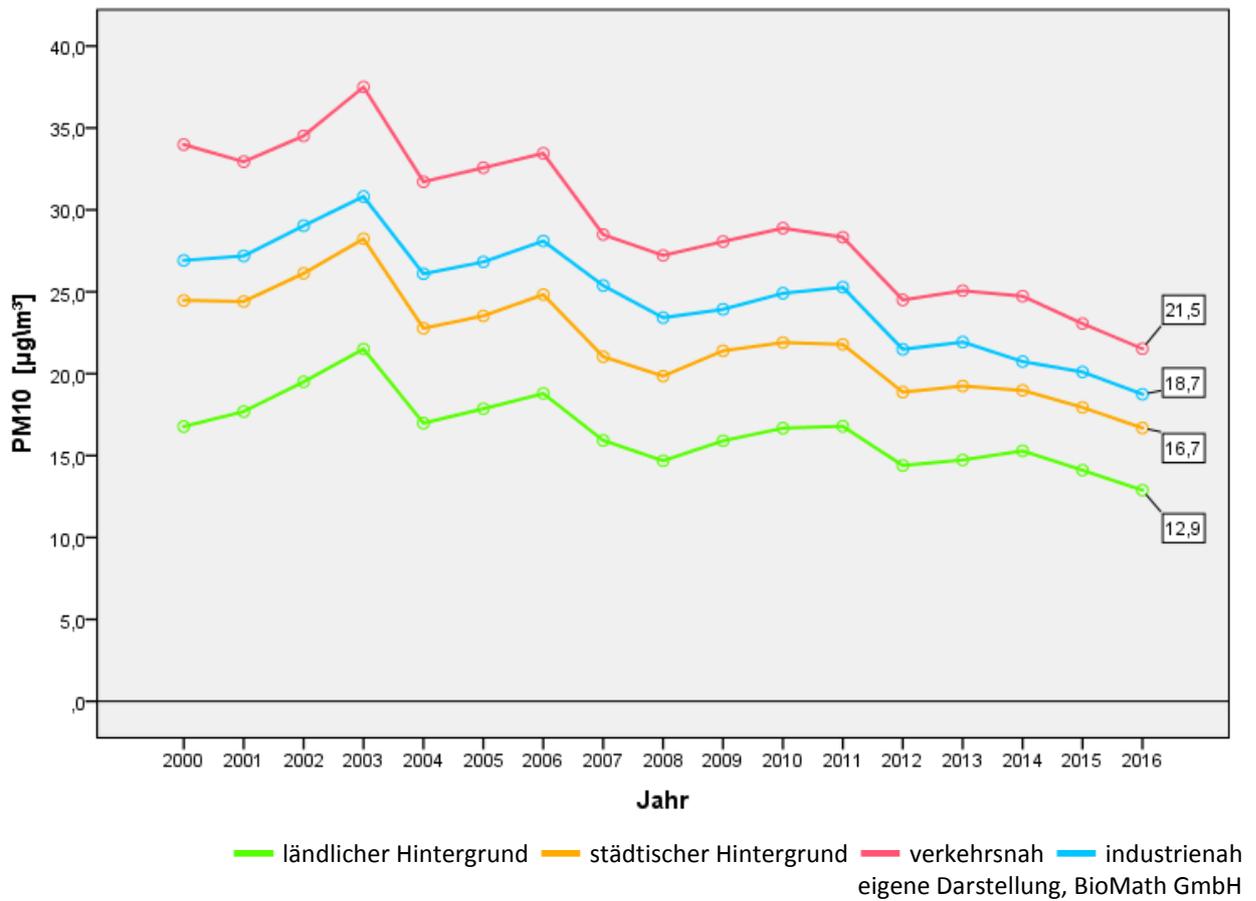
In allen Kategorien konnte für das Modell, der UBA-Auswahl und der Rohdaten ein signifikantes Sinken der PM10-Konzentration nachgewiesen werden (siehe 3.4.4.1 bis 3.4.4.4).

Für die Kategorien ländlicher Hintergrund und industrienah waren die Abweichungen der Mittelwertzeitreihen zwischen Modell, UBA-Auswahl und Rohdaten im Mittel kleiner als $0,4 \mu\text{g}/\text{m}^3$, d.h. sie lagen sehr dicht beieinander.

In der Kategorie städtischer Hintergrund lagen die Zeitreihen mit einer mittleren Abweichung von maximal $0,05 \mu\text{g}/\text{m}^3$ extrem dicht beieinander.

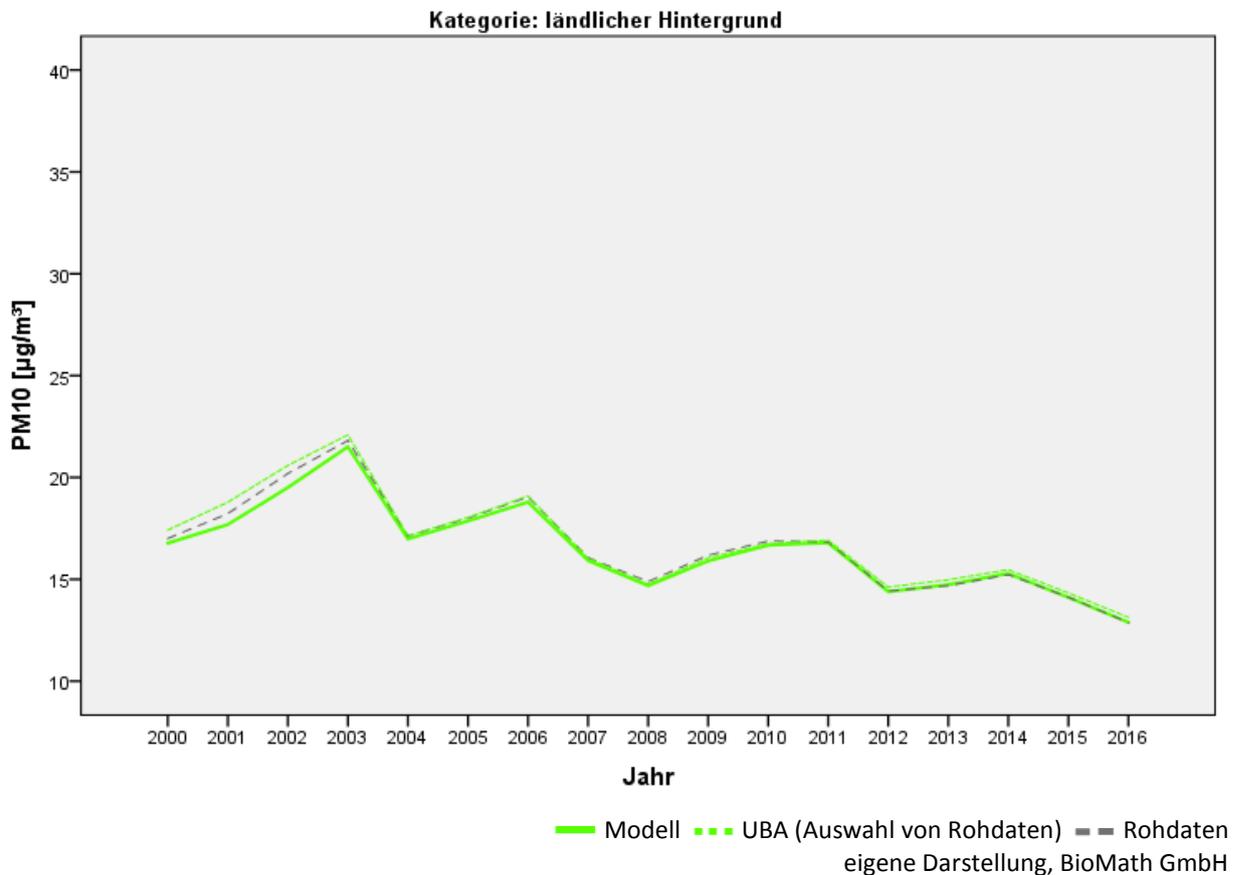
In der Kategorie verkehrsnah wichen die Zeitreihen im Mittel um maximal $1,2 \mu\text{g}/\text{m}^3$ ab, d.h. die Zeitreihen lagen dicht beieinander.

Abbildung 7: Jahresmittelwertzeitreihen des Modells für PM10



Das Modell weist einen signifikanten Trend von $-0,33 \mu\text{g}/\text{m}^3$ je Jahr für den ländlichen Hintergrund, von $-0,54 \mu\text{g}/\text{m}^3$ je Jahr für den städtischen Hintergrund, von $-0,84 \mu\text{g}/\text{m}^3$ je Jahr für die Kategorie verkehrsnah und von $-0,59 \mu\text{g}/\text{m}^3$ je Jahr für die Kategorie industrienah aus.

3.4.4.1 PM10 - ländlicher Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	113	110,8	1130	102,4
UBA	37	36,3	620	56,2
Rohdaten	102	100,0	1104	100,0

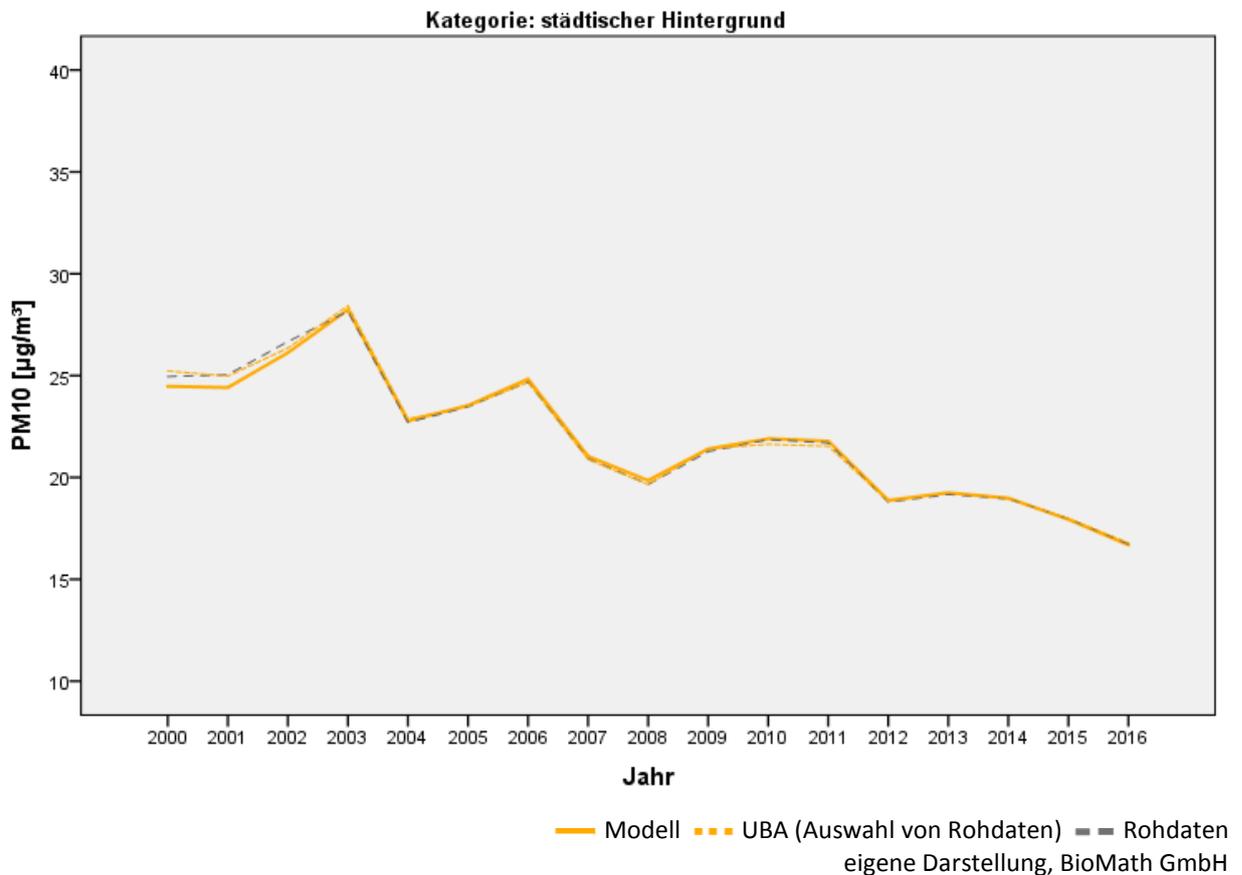
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,5882	0,0010	-0,3263	-0,4789	-0,1737	Ja	-0,5373	-0,1153	Ja
UBA	-0,6177	0,0005	-0,3675	-0,5247	-0,2102	Ja	-0,5849	-0,1500	Ja
Rohdaten	-0,6324	0,0004	-0,3572	-0,5134	-0,2011	Ja	-0,5731	-0,1414	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	-0,18	-0,69	0,06
UBA - Rohdaten	0,15	-0,12	0,56
Modell - UBA	-0,33	-1,10	-0,08

3.4.4.2 PM10 - städtischer Hintergrund



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	300	109,9	3117	103,4
UBA	96	35,2	1632	54,1
Rohdaten	273	100,0	3015	100,0

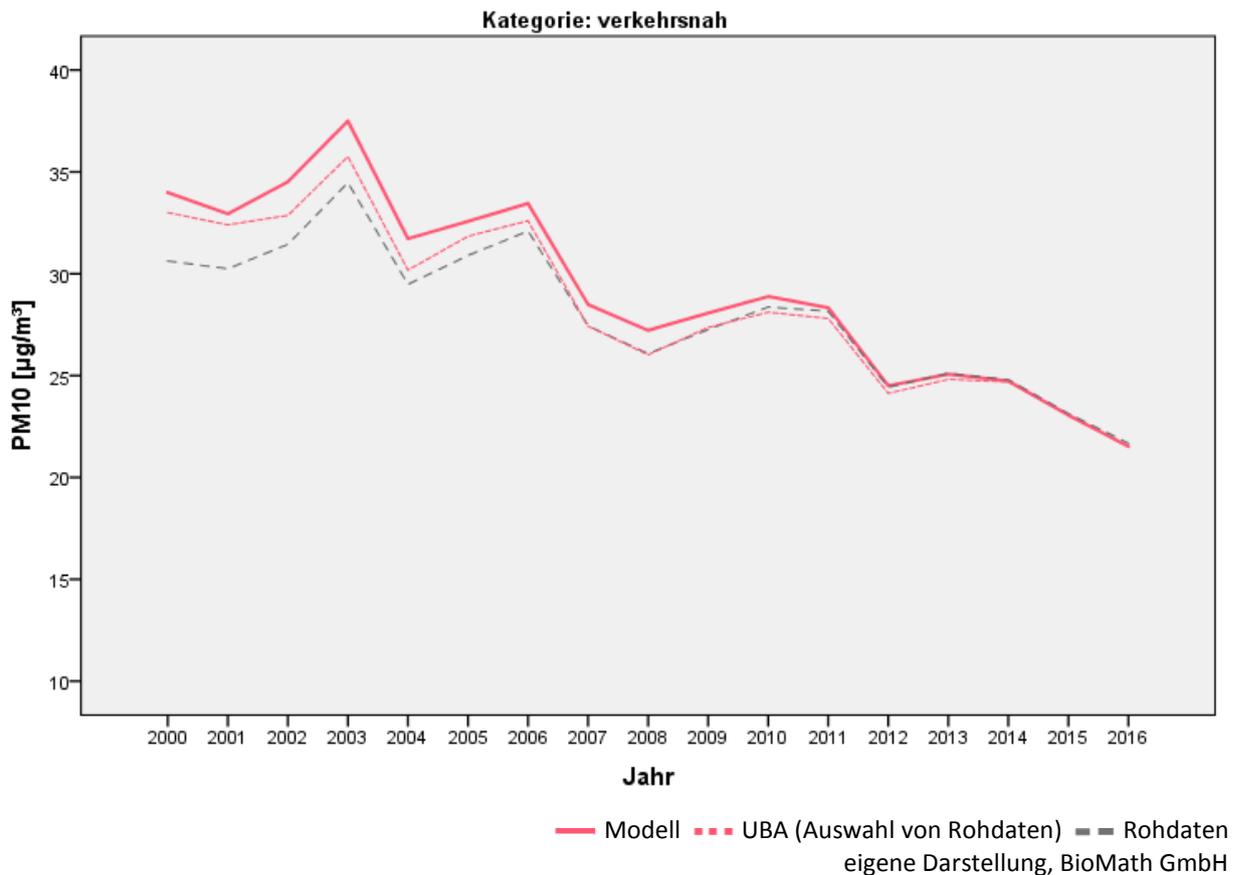
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,7059	<,0001	-0,5434	-0,7051	-0,3817	Ja	-0,7670	-0,3198	Ja
UBA	-0,7353	<,0001	-0,5747	-0,7303	-0,4190	Ja	-0,7898	-0,3595	Ja
Rohdaten	-0,7206	<,0001	-0,5703	-0,7281	-0,4126	Ja	-0,7884	-0,3523	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	-0,04	-0,63	0,19
UBA - Rohdaten	0,01	-0,32	0,28
Modell - UBA	-0,05	-0,75	0,27

3.4.4.3 PM10 - verkehrsnah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	241	82,3	2086	96,7
UBA	92	31,4	1285	59,5
Rohdaten	293	100,0	2158	100,0

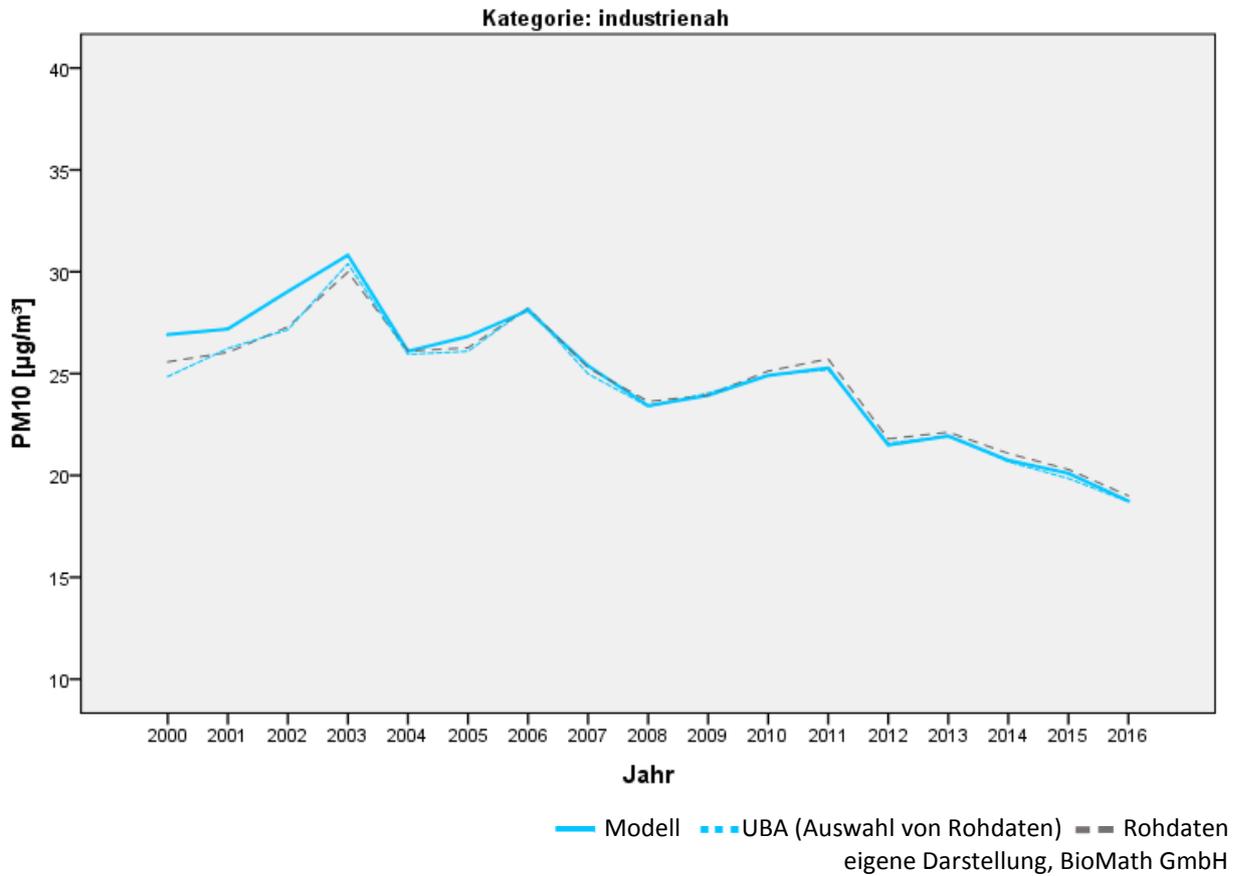
Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,7500	<,0001	-0,8356	-1,0231	-0,6481	Ja	-1,0948	-0,5763	Ja
UBA	-0,7500	<,0001	-0,7509	-0,9288	-0,5729	Ja	-0,9969	-0,5048	Ja
Rohdaten	-0,6765	0,0002	-0,5981	-0,7843	-0,4120	Ja	-0,8555	-0,3408	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	1,23	-0,16	3,36
UBA - Rohdaten	0,46	-0,35	2,38
Modell - UBA	0,76	-0,01	1,74

3.4.4.4 PM10 - industrienah



Anzahl

Quelle	Messreihen		Messwerte	
	Anzahl	Prozent	Anzahl	Prozent
Modell	70	82,4	559	97,0
UBA	26	30,6	373	64,8
Rohdaten	85	100,0	576	100,0

Trend

Quelle	Mann-Kendall		Lineare Regression						
	Kendal Tau	p-Wert	Anstieg	95%-KI unten	95%-KI oben	p<0,05	99%-KI unten	99%-KI oben	p<0,01
Modell	-0,7353	<,0001	-0,5883	-0,7548	-0,4218	Ja	-0,8184	-0,3581	Ja
UBA	-0,6324	0,0004	-0,4968	-0,6887	-0,3049	Ja	-0,7621	-0,2316	Ja
Rohdaten	-0,6324	0,0004	-0,4800	-0,6586	-0,3013	Ja	-0,7269	-0,2330	Ja

Abweichung

Differenz	Mittelwert	Min	Max
Modell - Rohdaten	0,20	-0,44	1,74
UBA - Rohdaten	-0,19	-0,72	0,41
Modell - UBA	0,39	-0,13	2,06

3.4.5 Ergebnis der induktiven Auswertung

Die durch das Modell der Differenzen berechnete standardisierte Mittelwertzeitreihe ist für die Schätzung der Trends der Luftschadstoffe Ozon, NO₂ und PM₁₀ sehr gut geeignet. Das Modell bezieht alle Messreihen eines Zeitintervalls die mindestens zwei aufeinanderfolgende Messwerte besitzen mit ein. Weiterhin wird durch die Differenzbildung aufeinanderfolgender Messwerte das Messniveau der Messreihen eliminiert. Somit werden die vorhandenen Inhomogenitäten der Daten berücksichtigt.

Durch Addition einer Konstanten wird die standardisierte Mittelwertzeitreihe auf das vorhandene Messniveau der Daten transformiert. Als Konstante ist der gewichtete Mittelwert des letzten Messzeitpunktes gut geeignet. Der Vergleich der so berechneten Mittelwertzeitreihen mit den Mittelwertzeitreihen der Rohdaten und der UBA-Auswahl zeigt für alle Schadstoffe in den Kategorien ländlicher Hintergrund, städtischer Hintergrund und industrienah im Mittel geringe Abweichungen. Die Abweichungen zwischen Modell und Rohdaten sind etwas stärker in der Kategorie verkehrsnah ausgeprägt, hier insbesondere am Anfang des Zeitintervalls. Die Ursache dafür liegt an der Tatsache, dass zunehmend die Messungen am Ende des Zeitintervalls auf einem höheren Niveau stattfanden. Somit ergibt sich durch den gewichteten Mittelwert am Ende des Zeitintervalls, dass das Modell am Anfang des Zeitintervalls höhere Werte (bei Ozon niedrigere Werte) schätzt. Allerdings ist die Abweichung zwischen Modell und UBA-Auswahl deutlich kleiner.

Alle Mittelwertzeitreihen wurden mit Hilfe der Parameterschätzung der linearen Regression und mit dem Mann-Kendall Test auf Signifikanz des Trends getestet. Sowohl bezüglich einer Signifikanz von 0,05 als auch von 0,01 zeigten beide Tests die gleichen Testergebnisse.

Die Parameterschätzung der linearen Regression ist sowohl für die Schätzung des Trends als auch für den Test auf Signifikanz des Trends geeignet. Die Voraussetzungen für die Anwendung der Regressionsanalyse, insbesondere der Normalverteilung des Fehlers (und nicht der Daten), sind weitestgehend erfüllt. Dazu wurden unter anderem die standardisierten Residuen der Regression für die Mittelwertzeitreihen des Modells im Histogramm und P-P Diagramm¹ beurteilt. Es ist nicht notwendig den Mann-Kendall Test, der nur Ränge verwendet, heranzuziehen.

Insgesamt kann festgesellt werden, dass die Mittelwertzeitreihen des Modells und der UBA-Auswahl zwar recht dicht beieinander liegen, aber unterschiedliche Stärken des Trends aufweisen.

¹ Probability-Probability-Plot - stellt die erwarteten Wahrscheinlichkeiten (hier Normalverteilung) zu den beobachteten Wahrscheinlichkeiten dar. Wenn die Punkte annähernd auf der Diagonalen liegen, kann von einer Normalverteilung ausgegangen werden.

4 Programm zur Berechnung der Modell-Mittelwertzeitreihen

Für das Modell der jährlichen Differenzen (siehe 3.4.1) wurde ein R-Programm zur Berechnung der Mittelwertzeitreihen und zur Schätzung der Trends entwickelt.

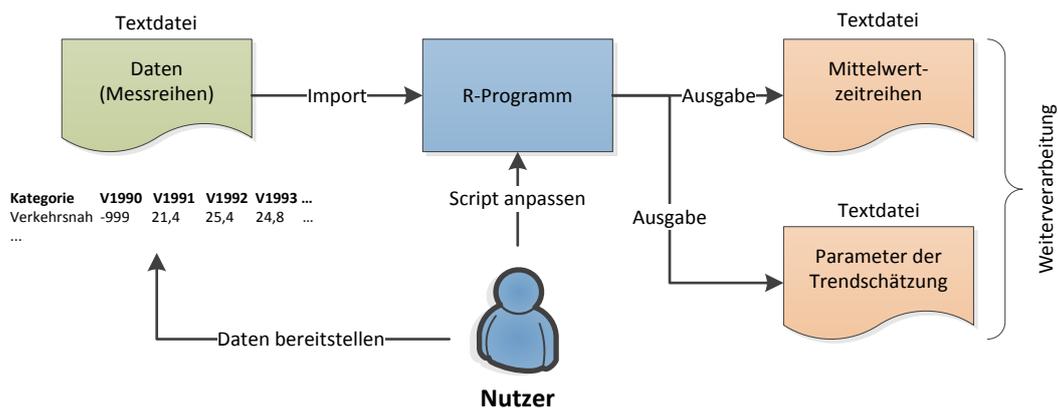
4.1 R-Programm

Unter der folgenden R-Umgebung wurde das Programm erstellt:

- ▶ R version 3.5.0 (2018-04-23)
- ▶ Plattform: x86_64-w64-mingw32/x64 (64-bit)
- ▶ Library data.table version 1.11.2 (Extension of data.frame)
- ▶ Library plyr (Tools for Splitting, Applying and Combining Data)

Als Dateneingabe benötigt das Programm eine Textdatei mit den Jahresmittelwerten der Messstationen. Im Programmabschnitt *Eingabe/Vorgaben* ist diese Datei dem Programm bekannt zu machen. Weiterhin ist in diesem Abschnitt der Zeitbereich für die Auswertung festzulegen. Ferner sind die Dateinamen für die Ausgabe der Mittelwertzeitreihen und der Parameter der Trends anzugeben. Die Ausgabedateien sind ebenfalls Textdateien, die weiterverarbeitet werden können (siehe Abbildung 8).

Abbildung 8: Verarbeitungsschema R-Programm



eigene Darstellung, BioMath GmbH

Der Zeichensatz jeder Textdatei ist vom Typ Windows-1252. Jede Zeile der Textdatei ist mit dem Steuerzeichen $0D_{Hex}^2$ und $0A_{Hex}$ abgeschlossen. In der ersten Zeile stehen die Variablennamen, in den folgenden Zeilen die Datensätze. Die Variablennamen und die Felder der Datensätze sind durch das Tabulatorzeichen (9_{Hex}) getrennt.

4.1.1 Daten/Messreihen der Messtationen

In dieser Textdatei muss die Variable *Kategorie* und für jedes Jahr die Variablen *V<jahr>* in aufsteigender Feld-Reihenfolge vorhanden sein. Da der Buchstabe *V* zur Identifizierung verwendet wird, dürfen keine anderen Variablennamen mit dem Buchstaben *V* beginnen. Alle weiteren Variablen werden ebenfalls eingelesen aber nicht weiterverarbeitet. Die Werte der Variable *Kategorie* wird als Textvariable, die der Messwerte als numerische Variable verarbeitet. Numerische Werte kleiner gleich Null werden als Fehlwerte interpretiert.

² Hexadezimal

Beispiel

Stationscode	V1992	V1993	V1994	Kategorie
DEBB001	-999	13,1609	13,2045	städtischer Hintergrund
DEBB003	25,3892	24,7781	25,5781	verkehrsnahe
DEBB006	-999	21,1701	19,4704	städtischer Hintergrund
DEBB026	19,8251	19,4231	18,2513	industrienah
DEBE027	20,443	20,8685	18,7011	ländlicher Hintergrund

4.1.2 R-Programmanpassungen

Im Programmabschnitt *Eingabe/Vorgabe* sind die Inhalte der folgenden Variablen festzulegen.

Variable	Inhalt
arbeitsverzeichnis	Arbeitsverzeichnis
datendateiname	Dateiname der Daten/Messreihen
ergebnisdateiname_modellmittelwerte	Dateiname für die berechneten Werte der Mittelwertzeitreihe
ergebnisdateiname_regression	Dateiname für die Parameter der Trendschätzung
jahr_von	Auswertungszeitraum: erstes Jahr*
jahr_bis	Auswertungszeitraum: letztes Jahr

* Das erste Jahr des Auswertungszeitraumes muss immer größer sein als das erste Jahr in den Daten.

Beispiel

```
# Arbeitsverzeichnis
arbeitsverzeichnis <- "D:/UBA-Luft/R"

# Dateinamen für die Eingabedaten und Ausgabedaten
datendateiname <- "Data/Ozon-Jahresmittelwerte_1990-2017.txt"
ergebnisdateiname_modellmittelwerte <- "Out/Ozon-Modell_2008-2017 (Modell).txt"
ergebnisdateiname_regression <- "Out/Ozon-Modell_2008-2017 (Regression).txt"

# Auswertungszeitraum
# ACHTUNG: Die Variable jahr_von muss immer größer sein als das erste Jahr in den
#           Eigabedaten.
#           Für das jahr_von wird die Differenz aus jahr_von - jahr_von-1 gebildet.

jahr_von <- 2008
jahr_bis <- 2017
```

4.1.3 Ausgabedateien

Die Ausgabedaten enthalten für jede Kategorie einen Datensatz mit den folgenden Variablen.

Datei	Variable	Inhalt
Zeitreihe	Kategorie	Bezeichnung der Kategorie
	gew.mw.sum.mw.diff.V<jahr>	Modell-Mittelwert des Jahres <jahr> (erstes Jahr)
	gew.mw.sum.mw.diff.V<jahr+1>	Modell-Mittelwert des Jahres <jahr+1>

	gew.mw.sum.mw.diff.V<jahr+n>	Modell-Mittelwert des Jahres <jahr+n> (letztes Jahr)

Datei	Variable	Inhalt
Trend	Kategorie	Bezeichnung der Kategorie
	steigung	Anstieg der linearen Regression (linearer Trend)
	unteres.95Konfidenzintervall	Unteres 95%-Konfidenzintervalls des Anstieges
	oberes.95Konfidenzintervall	Oberes 95%-Konfidenzintervalls des Anstieges

4.1.4 Programmablauf

Das Programm ist in drei Parts eingeteilt.

1. Eingaben/Vorgaben
2. Berechnung der Modell-Mittelwerte und Regressionsanalyse für den linearen Trend
3. Ausgabe der Modell-Mittelwerte und Regressionsparameter für den linearen Trend

Im ersten Part werden alle Variablen gelöscht und die Dateinamen sowie der Auswertungszeitraum in Programmvariablen gespeichert.

Zu Beginn des zweiten Parts werden zusätzlichen Bibliotheken geladen. Die weiteren Berechnungen erfolgen so, dass zunächst nur die Messwerte in die Tabelle *auswertungsdaten* übernommen werden. Diese Tabelle wird um Spalten Kategorie und um die Differenzen zwischen den Jahren erweitert. Danach erfolgt die Mittelwertbildung über die Kategorien und Speicherung dieser berechneten Werte in der Tabelle *ergebnisse*. Auch diese Tabelle wird um weitere Spalten erweitert. Bei der Spaltenerweiterung werden die Spaltennamen mit folgenden Präfixen versehen.

Präfix	Bedeutung
diff	Differenz
mw	Mittelwert
sum	Summe
gew	gewichtet

Mit Hilfe dieser Präfixnamen ergibt sich die Bedeutung der Variablen, z.B. bedeutet *sum.mw.diff.V2015* die Summe der Mittelwerte der Differenzen zwischen 2014 und 2015. Die Endergebnisse werden in der Tabelle *ergebnismodellmittelwerte* gespeichert.

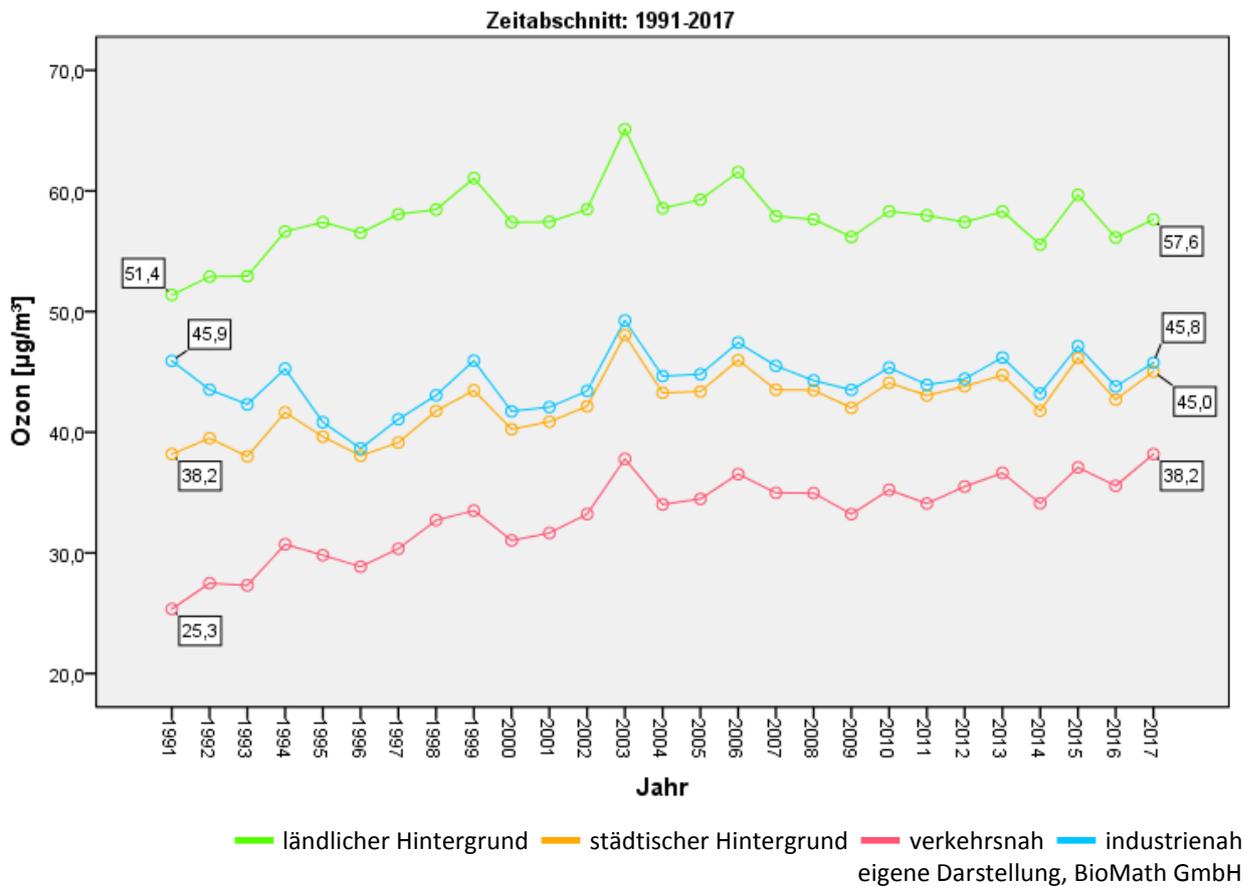
Für die lineare Regression wird die Tabelle *ergebnis* auf die Spalten der Modell-Mittelwerte reduziert, danach die Regression gerechnet und die Ergebnisse in der Tabelle *ergebnisregression* gespeichert.

Im dritten Part werden die Tabellen der Ergebnisse in Textdateien gespeichert.

4.2 Berechnungen mit dem R-Programm

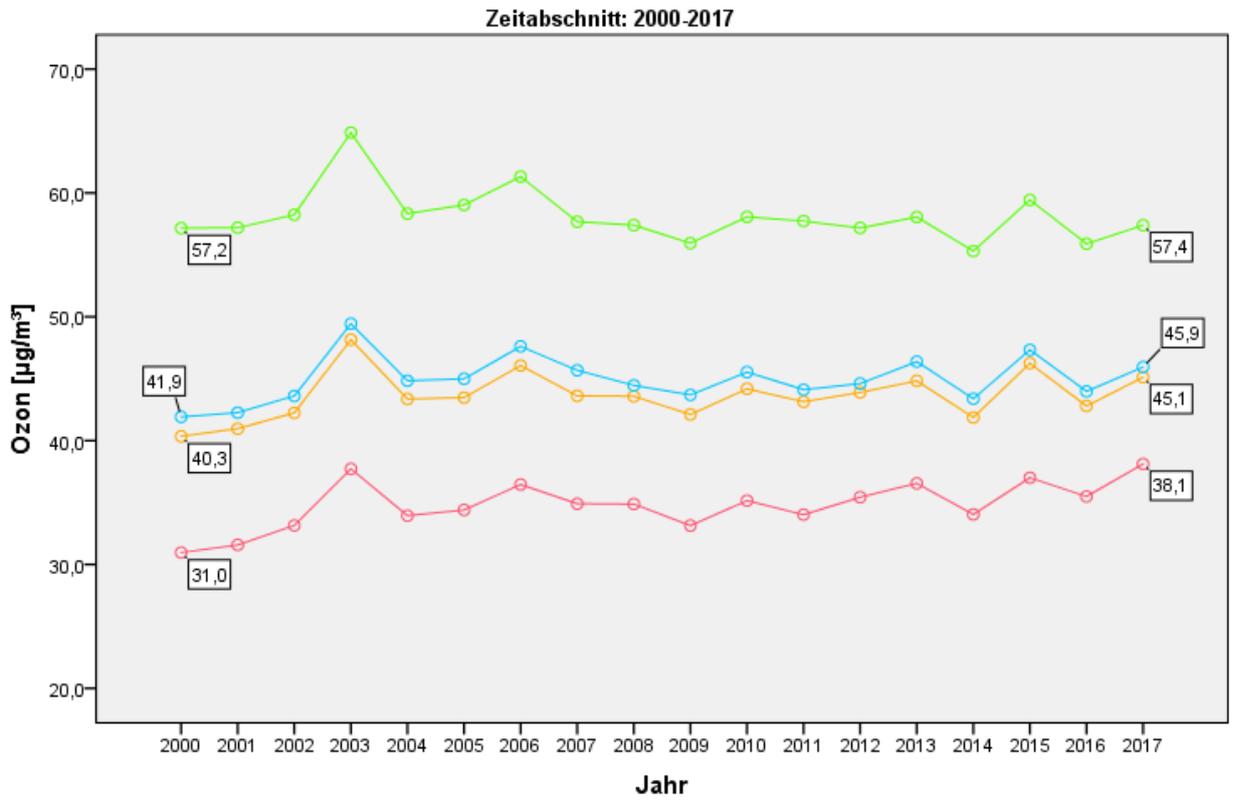
Mit dem R-Programm wurden die Jahresmittelwerte von Ozon, NO₂ und PM₁₀ für die Zeiträume 1991-2017 (langfristig), 2000-2017 (mittelfristig) und 2008-2017 (kurzfristig) ausgewertet.

4.2.1 Ozon von 1991 bis 2017



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	0,1070	-0,0255	0,2395	Nein
städtischer Hintergrund	0,2247	0,1280	0,3213	Ja
verkehrsnah	0,3593	0,2692	0,4494	Ja
industrienah	0,1062	-0,0022	0,2145	Nein

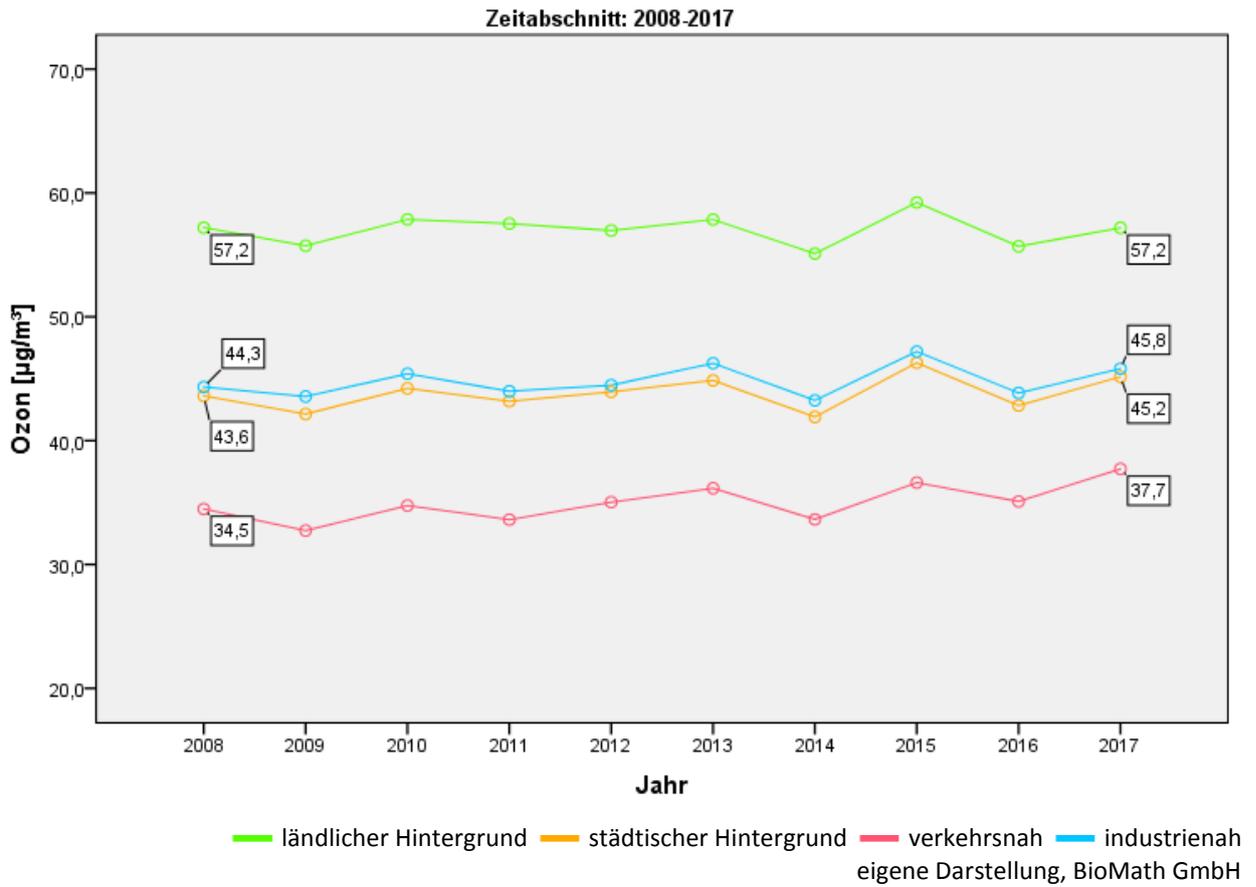
4.2.2 Ozon von 2000 bis 2017



— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnah
 — industrienah
 eigene Darstellung, BioMath GmbH

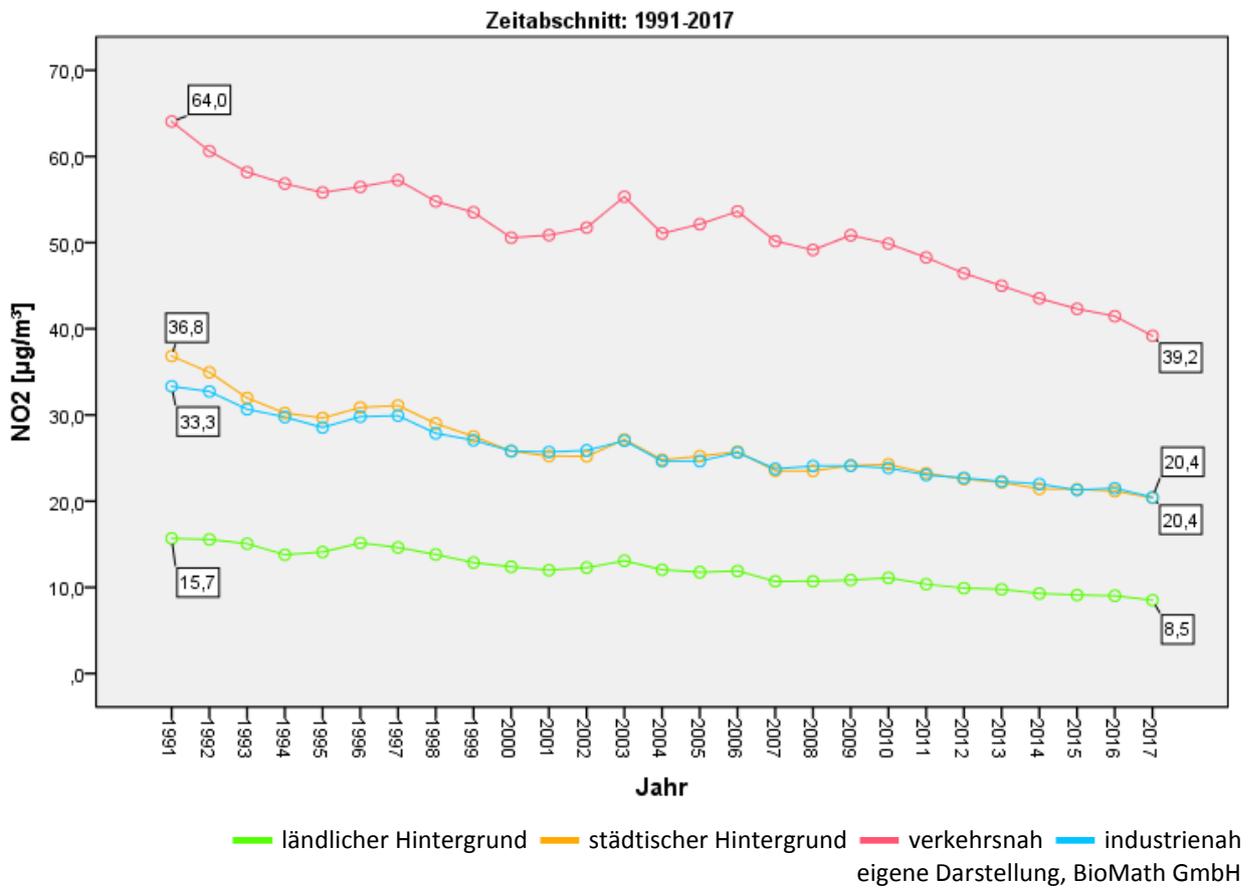
Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,1434	-0,3452	0,0583	Nein
städtischer Hintergrund	0,0964	-0,0886	0,2813	Nein
verkehrsnahe	0,2143	0,0588	0,3697	Ja
industrienah	0,0707	-0,1136	0,2550	Nein

4.2.3 Ozon von 2008 bis 2017



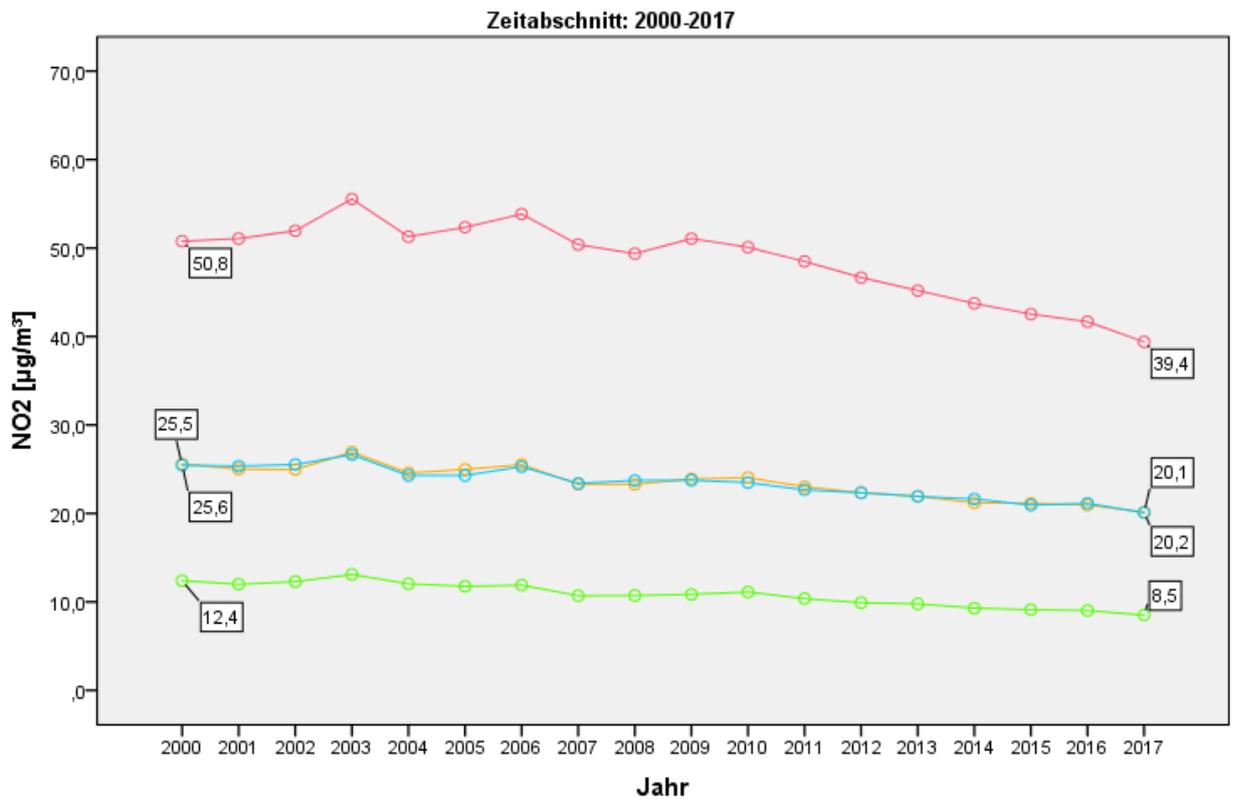
Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,0002	-0,3324	0,3320	Nein
städtischer Hintergrund	0,1583	-0,1882	0,5048	Nein
verkehrsnah	0,3397	0,0421	0,6373	Ja
industrienah	0,1449	-0,1827	0,4724	Nein

4.2.4 NO2 von 1991 bis 2017



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2636	-0,2883	-0,2390	Ja
städtischer Hintergrund	-0,5086	-0,5839	-0,4333	Ja
verkehrsnahe	-0,7029	-0,8077	-0,5980	Ja
industrienah	-0,4267	-0,4740	-0,3793	Ja

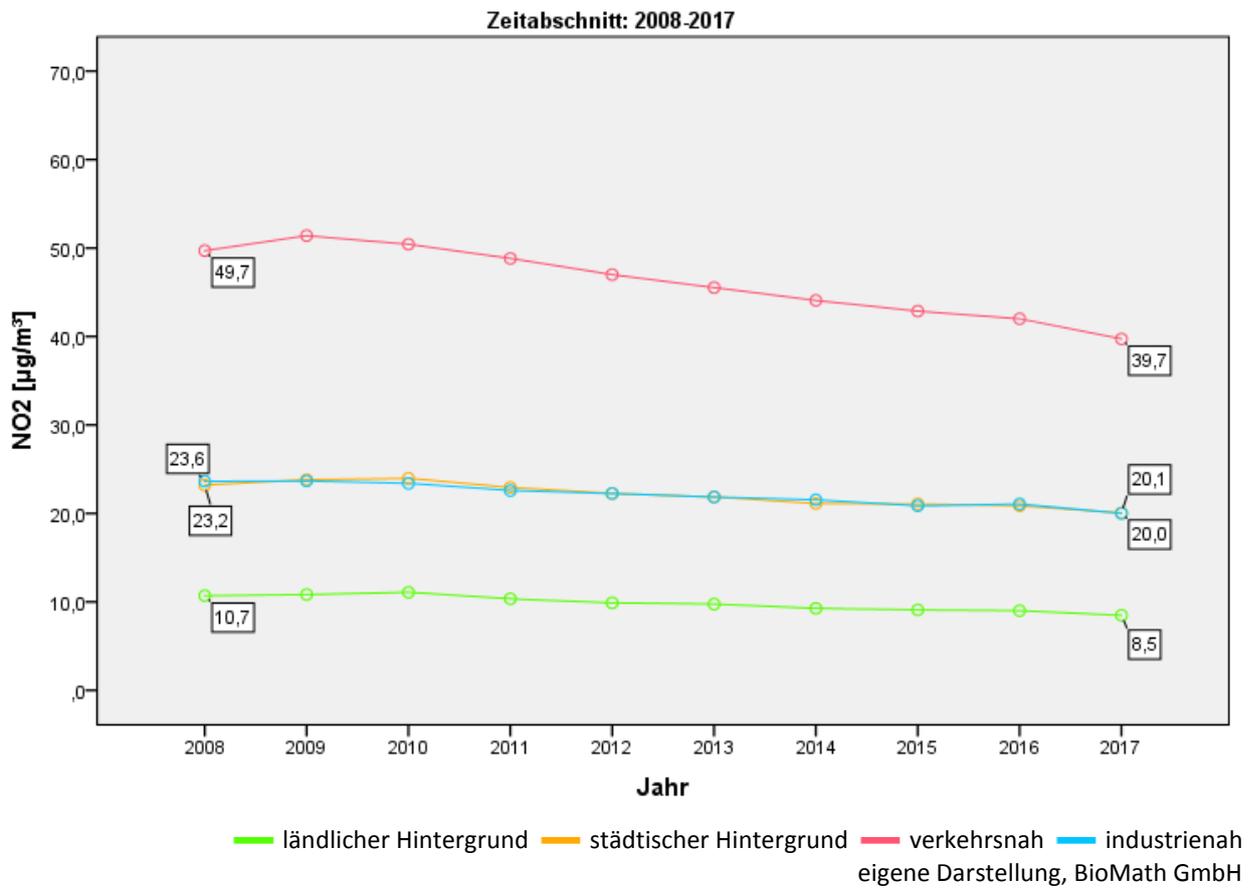
4.2.5 NO2 von 2000 bis 2017



— ländlicher Hintergrund
 — städtischer Hintergrund
 — verkehrsnah
 — industrienah
 eigene Darstellung, BioMath GmbH

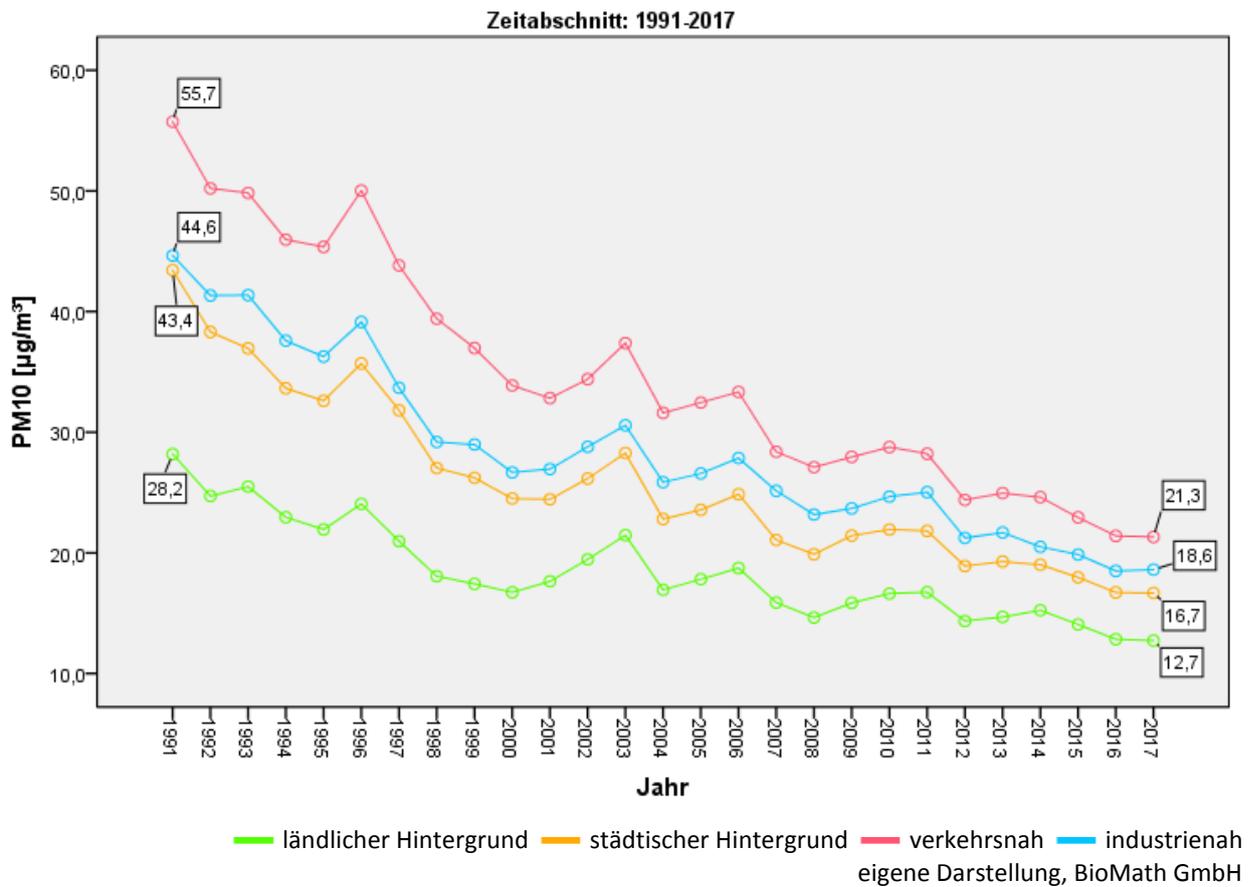
Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2408	-0,2802	-0,2014	Ja
städtischer Hintergrund	-0,3280	-0,4001	-0,2560	Ja
verkehrsnahe	-0,7301	-0,9474	-0,5128	Ja
industrienah	-0,3272	-0,3846	-0,2697	Ja

4.2.6 NO2 von 2008 bis 2017



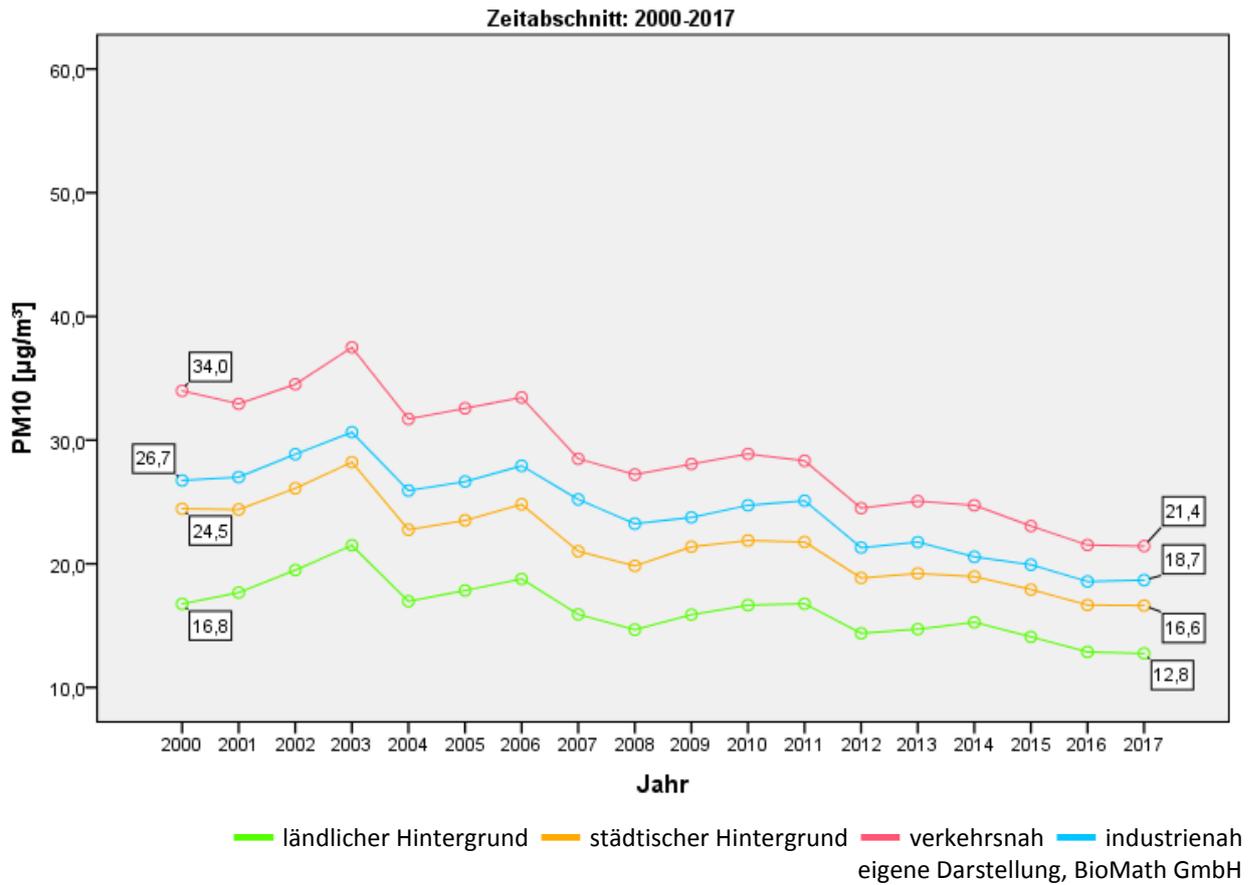
Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,2787	-0,3443	-0,2131	Ja
städtischer Hintergrund	-0,4198	-0,5322	-0,3075	Ja
verkehrsnah	-1,2669	-1,5178	-1,0159	Ja
industrienah	-0,4077	-0,4696	-0,3457	Ja

4.2.7 PM10 von 1991 bis 2017



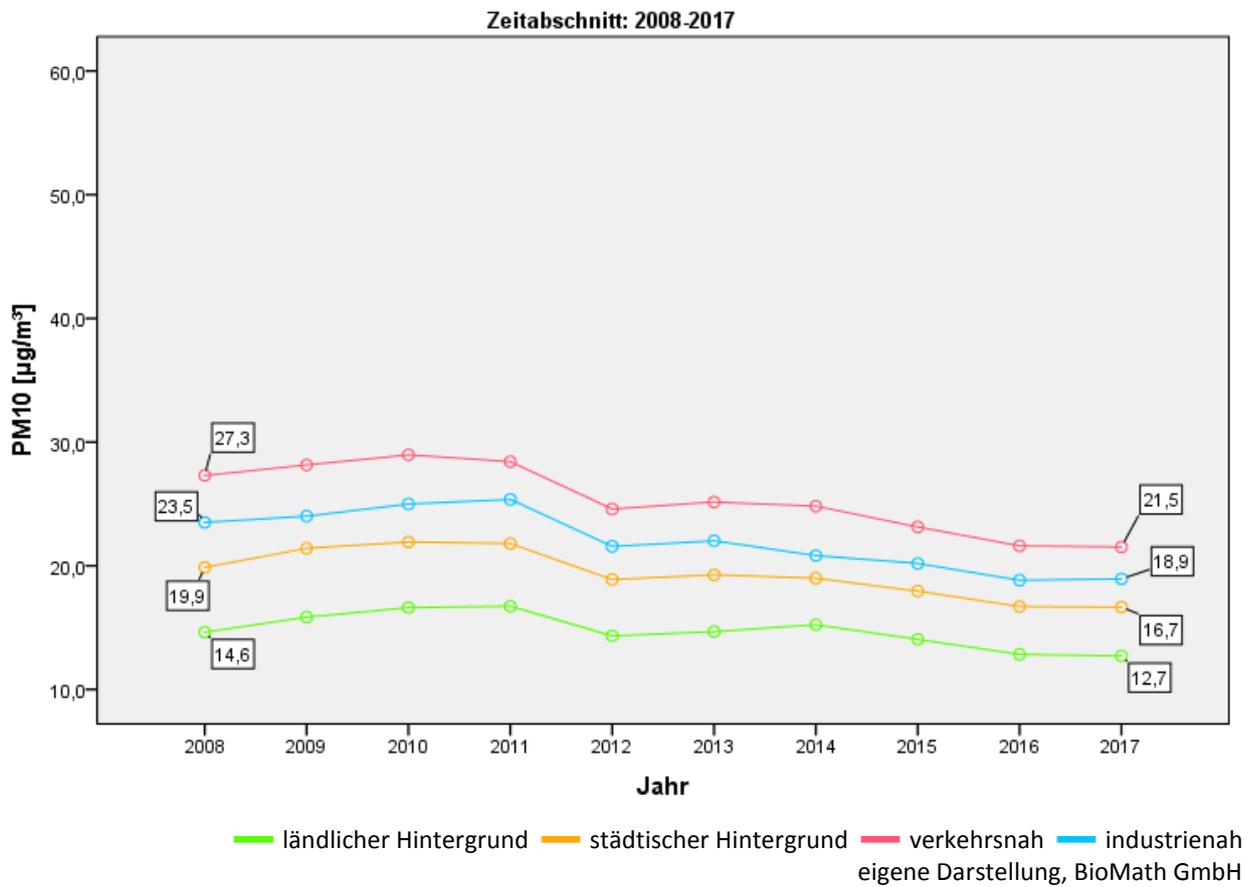
Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,4601	-0,5529	-0,3673	Ja
städtischer Hintergrund	-0,8403	-0,9742	-0,7064	Ja
verkehrsnahe	-1,1945	-1,3403	-1,0486	Ja
industrienah	-0,8751	-1,0061	-0,7440	Ja

4.2.8 PM10 von 2000 bis 2017



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,3402	-0,4761	-0,2042	Ja
städtischer Hintergrund	-0,5496	-0,6928	-0,4065	Ja
verkehrsnahe	-0,8400	-1,0058	-0,6742	Ja
industrienah	-0,5990	-0,7467	-0,4513	Ja

4.2.9 PM10 von 2008 bis 2017



Kategorie	Anstieg	95% KI unten	95% KI oben	P<0,05
ländlicher Hintergrund	-0,3351	-0,5901	-0,0800	Ja
städtischer Hintergrund	-0,5438	-0,8174	-0,2702	Ja
verkehrsnah	-0,8317	-1,1367	-0,5267	Ja
industrienah	-0,6942	-0,9929	-0,3955	Ja

5 Literaturrecherche

Die Literaturrecherche erfolgte in Form eines narrativen Reviews, angelehnt an die Richtlinien für Systematische Übersichtsarbeiten (Systematic Review). Die Suchfrage hieß: Welche methodischen Ansätze zur Analyse von Trends in der Luftqualität, insbesondere zur Berücksichtigung inhomogener Daten, sind veröffentlicht?

Zunächst wurde die Suchstrategie entwickelt, die eine Suche in einer elektronischen Literaturdatenbank (Scopus), im Internet (Google und Google Scholar) und manuell (Referenzlisten von Reviews, Schlüsselartikeln und ggf. deutschsprachige Webseiten von Zeitschriften und Organisationen, z.B. UBA) vorsah. Für die Suche wurden Suchbegriffe definiert und diese entsprechend der Regeln der Suchplattformen zu Suchgleichungen verknüpft. Die in den unterschiedlichen Quellen identifizierten Publikationen wurden anschließend auf Duplikate untersucht und die Titel und Abstracts auf ihre Eignung zur Beantwortung der Suchfrage geprüft.

Insgesamt konnten 181 Dokumente, die sich mit der Trendanalyse von Daten zur Luftqualität beschäftigen, identifiziert werden. Aus allen Dokumenten wurden folgende Daten zu den methodischen Ansätzen der Trendanalyse extrahiert:

- ▶ Region, in der die Daten erhoben wurden
- ▶ Messwerte: Parameter, Kategorie, Anzahl Stationen/ Umfang Daten,
- ▶ Ziel der Analyse, Methodik

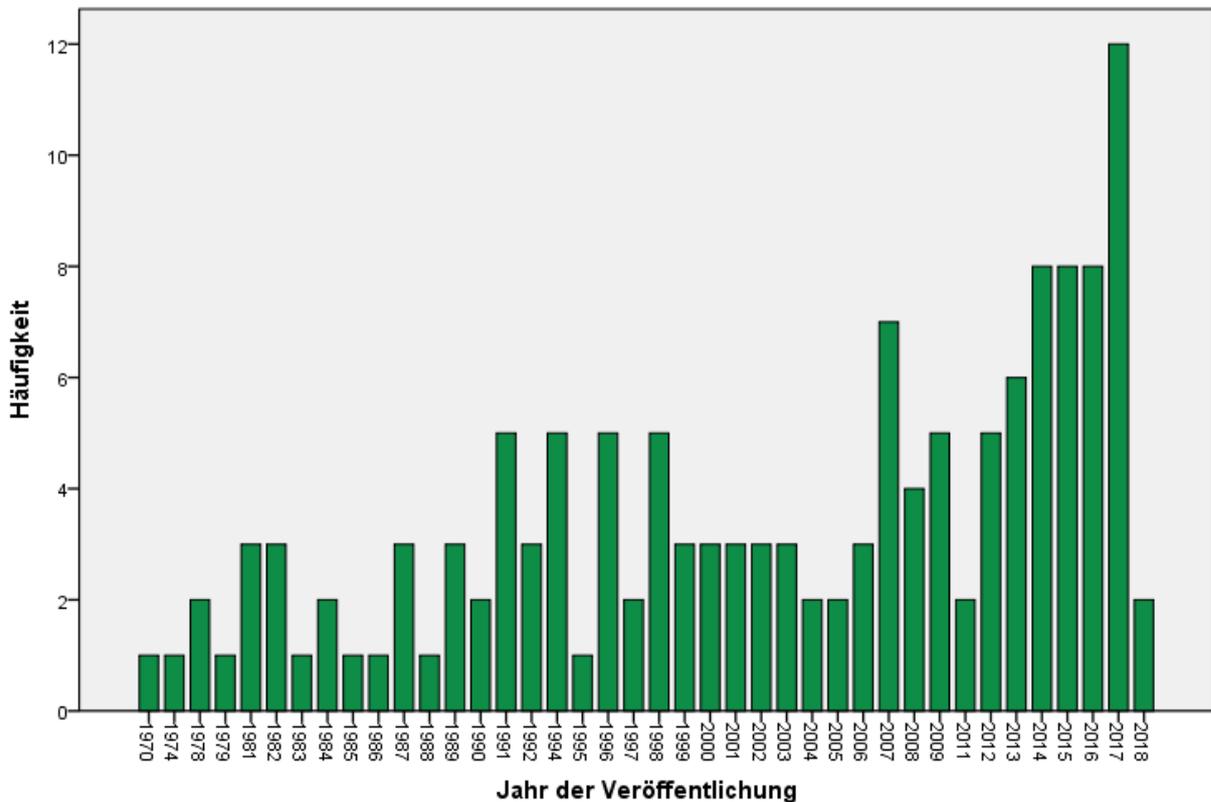
Abstract	Online DOI	Relevanz	Region	Region Kat	Parameter	Kategorien	Stationen/ Daten	Zeitraum	Zeitraum Dauer	Ziel Analyse	Methodik	Kommentar
Air quality management rega	https://10.10	1	Portugal	Portugal	PM10, PM2.5,	urban, suburbe	20 monitoring	2007 to 201	9	temporal pattern	statistical methods	
There have been many chang	https://10.10	1	Northeastern Unit	USA	PM2.5		NYS Departm	between 20	6	trend analyses	Mann-Kendall regression with	
Nablus city is an important u	https://10.10	1	Nablus, Palestine	Palastina	PM10, PM5.0,	urban and indu	four roadsides	December :	1	concentration di	?	
The Japanese government ha	https://10.10	1	Japan	Japan	sulphur dioxide (SO2)	carbon m	?	2002 to 201	9	evaluating air qu		
Presence of certain substanc	https://10.13	1	Bangalore, India	Indien	sulphur dioxide (SO2),	nitrogen state	pollution ?			identify air pollut		
A four-step adaptive ozone tr	https://10.10	1	global	global	ozone		four observati	1979 to 200	30	trend estimation	multivariate line	variations asso
Based on satellite remote se	https://10.60	1	North Pacific Oce	Nordpazifik	ozone		satellite remot	1979 to 201	35	long-term trends	multivariate linear regression m	
In the last 30 years ambient	(https://10.10	1	South Central U.S	USA	ozone		four South Ce	1986-2015	29	long-term trends	?	
Bangalore is one of India's fa	https://10.15	1	Bangalore, India	Indien	?		six specific lo	2006-2013	7	air pollution tren	?	
Trends in the emissions to ai	https://10.15	1	Ireland	Irland	sulphur dioxide, nitrogen oxid	?	?	1990-2009	19	trend analysis	logarithmic mean Divisia index	
The ozone radiometer GROM	https://10.51	1	Bern, Switzerland	Schweiz	ozone		radiometer GF	January 19	18	long-term trend	robust multilinear parametric tr	
Trends in the vertical distrib	https://10.51	1	global	global	ozone		?	periods before and afte		trends in the ver	?	
Western Iran, for example, K	https://	1	Kermanshah, Iran	Iran	sulphur dioxide, ozone, nitrogen	obtained from	2006-2011	5	determine the tr	?		
The puzzling increase of anni	https://10.10	1	Japan, Tokyo Met	Japan	ozone		monitoring dat	1990-2010	20	trends	linear regressions	
Equivalent effective stratosph	https://10.10	1	polar regions (70-5	Polargebiet	ozone		?	2001-2012	11	trend	regression of time series	
Information on air quality in u	https://10.10	1	Catania, Italy	Italien	NO2, O3, CO, SO2, PM10		three monitori	2010 to 201	14	analyze the leve	ranks six levels	
This paper details a new met	https://10.51	1	?	k.A.	ozone		Stratospheric	?		long-term trend	a new method non-uniform sp	
We describe a hierarchical st	https://10.51	1	global	global	ozone		satellite meas	1984-2011	27	trend	hierarchical sta	Model variables
Background. The chemical c	https://10.11	1	Bari province, Italy	Italien	PM10, CO, NOx (NO and NO	six monitoring	January 201	2	trends	multivariate statistical techniqu		
We analyze the temporal vari	https://	1	?	k.A.	PM10		country wide (ten years		10	temporal variati	Kolmogorov-Zurbenko filtering r	
In EU, a significant percenta	https://10.10	1	Florence (Italy)	Italien	SO2, CO, NO,	urban	all stations of	1993-2012	19	trends	linear regressive framework	
Fluctuations of surface ozone	https://	1	Delhi, India	Indien	ozone		?	?		fluctuations of s	time series decomposition tec	
Regulations of the Montreal F	https://10.10	1	global	global	ozone		SBUV/SBUV-	1979-2012	33	long-term chang	statistical model evaluating tec	
Data from the annual, season	https://10.33	1	Mexican metropol	Mexiko	CO, NO2, SO2, O3, and PM1	automatic air-	2000-2011	11	annual trends, c	?		
We have studied data from tv	https://10.51	1	global	global	ozone		?	1984-2011	27	trend analysis o	?	

eigene Darstellung, BioMath GmbH

Die 181 Artikel wurden deskriptiv analysiert:

Die meistgefundenen Autoren waren Johannes Staehelin (ETH Zürich, Schweiz - 10 Publikationen), Gregory C. Reinsel (Wisconsin, University, Madison, United States – 8 Publikationen), und Tiao, George C Chicago, University, IL, United States – 7 Publikationen).

Die Publikationen erschienen zwischen 1970 und 2018.



eigene Darstellung, BioMath GmbH

Die am häufigsten genannten Regionen der Datenerhebung und Trendanalyse waren: global (22 Publikationen), USA (21), Nördliche Hemisphäre (11), Deutschland (9) und die Schweiz (9).

Der mit Anstand am meisten analysierte Parameter war Ozon (107 Publikationen), gefolgt von Feinstaub (29) und NO₂ (25).

Nur 13 Publikationen machten eine Angabe zum Belastungsszenario (ländlicher Hintergrund: 5 Publikationen, urbaner Hintergrund: 8 Publikationen, verkehrsnah: 1, industrienah: 1).

Die Anzahl der Messstationen lag zwischen 1 und 650. Oft wurden Daten von größeren Monitoring Programmen (SAGEII, TOMS) ausgewertet. Der Zeitraum der Messungen lag zwischen 1 und 72 Jahren.

Als Methoden der Trendanalyse wurden vor allem genannt: Regression, Zeitreihenanalyse, Trendtests, Mittelwertberechnungen, Monte-Carlo Analyse, Gemischte Modelle. Zur Handhabung inhomogener Daten konnten wir keine weiterführenden detaillierten Informationen identifizieren.

Die Suchstrategie und Dokumentation der Suche sowie eine deskriptive Beschreibung der identifizierten Publikationen findet sich in Anhang 6.1.

6 Anhänge

6.1 Review Protokoll zur Literaturrecherche

6.1.1 Review-Strategie

6.1.1.1 Hintergrund

Zur Beurteilung der Luftqualität in Deutschland stehen dem Umweltbundsamt Daten aus räumlich und zeitlich inhomogenen Messnetzen der 16 Bundesländer und des Umweltbundsamtes zur Verfügung.

Die Inhomogenität beruht einerseits auf der ungleichmäßigen Verteilung und andererseits auf der Schließung, Neueinrichtung oder zwischenzeitlichen Außerbetriebnahme von Messstationen. Damit ist es problematisch, sichere Trendaussagen zur kurz- und langfristigen Entwicklung der Luftqualität, die als repräsentativ für Deutschland angesehen werden können, zu treffen.

Im Rahmen dieses Projektes soll eine Methodik entwickelt werden, mit der sichere Aussagen zu kurz- und langfristigen mittleren Trends für die Luftschadstoffe Ozon, Stickstoffdioxid und Feinstaub NO₂ getroffen und objektiv beurteilt werden können.

Zur Thematik der Analyse und Ableitung von Trends in der Luftqualität soll eine Literaturrecherche vorgenommen werden. Zudem sollen interne und externe UBA-Berichte zu bisherigen Auswertungen zu dem Thema berücksichtigt werden.

6.1.1.2 Ziel des Reviews

Review-Typ: Narrative Review

Narrative Reviews bieten meist einen breiten Überblick zu einem bestimmten Thema. Sie sind daher gut geeignet, sich schnell über den aktuellen Forschungsstand zu einem Thema zu informieren.

Fragestellung – Konzept und Schlüsselemente

Konzept/ Elemente	Kriterien	Potentielle Suchbegriffe
Ergebnis (Luftqualität)	Luftverschmutzung Luftqualität Feinstaub Stickstoffdioxid Ozon	Air pollution, air quality, particulate matter, NO ₂ , fine dust, ozone
Entwicklung (Trend)	Trendanalyse	Trend analysis

Primäre Suchfrage

Welche methodischen Ansätze zur Analyse von Trends in der Luftqualität, insbesondere zur Berücksichtigung inhomogener Daten, sind veröffentlicht?

6.1.1.3 Methode

Suchstrategie

Suchbegriffe und Suchgleichung

Die Suche erfolgt in den Titeln der Publikationen. Es werden keine Synonyme und verbundene Begriffe verwendet. Es werden Trunkierungen und Platzhaltersymbole eingesetzt. Suchbegriffe können mit Booleschen und Ähnlichkeitsoperatoren verbunden werden.

Da in Google/ Google Scholar keine umfangreichen Suchgleichungen (Verknüpfung der Suchbegriffe sowohl über OR, AND, NOT Operatoren) formuliert, keine Truncation- und Wildcard- oder Proximity-Funktionen genutzt werden können, müssen aus der Suchstrategie des Review-Protokolls einzelne Schlüsselworte, die nur über AND verknüpft werden, ausgewählt und unterschiedlich miteinander zu mehreren Suchen kombiniert werden.

Set	Feld	Suchgleichung	Konzept/ Elemente
-----	------	---------------	-------------------

#1	Title	(air AND (pollut* OR quality)	Ergebnis
#2	Title	particulate*	Ergebnis
#3	Title	NO2	Ergebnis
#4	Title	(fine AND dust)	Ergebnis
#5	Title	ozone	Ergebnis
#6	Title	#1 OR #2 OR #3 OR #4 OR #5	
#7	Title	(trend* AND analys*)	Entwicklung
#8	Title	#6 AND #7	

Vereinfachte Suche in Google/ Google Scholar

Der Suchbegriff Ozon wurde als relevantester Parameter der Luftqualitätsmessung ausgewählt.

Set	Feld	Suchgleichung	Konzept/ Elemente
#1	irgendwo auf der Seite	(trend* AND analys*)	Entwicklung
#2	irgendwo auf der Seite	ozone	Ergebnis
#3	irgendwo auf der Seite	#1 AND #2	

Sprache: Die Suche wird ausschließlich in Englisch durchgeführt.

Zeitraum der Veröffentlichung: Keine Einschränkung.

Suchbereiche/ Quellen

Elektronische Literaturdatenbanken: Scopus

Internet-Suche: die ersten 50 relevanten Ergebnisse in Google und Google Scholar

Manuelle Suchen: Referenzlisten von Reviews, Schlüsselartikeln und ggf. deutschsprachige Webseiten von Zeitschriften und Organisationen (Suchbegriffe in Deutsch: Trendanalyse Luftqualität)

Publikationen, die vom Auftraggeber zur Verfügung gestellt wurden

Auswahl relevanter Referenzen

Die Auswahl erfolgt in zwei Stufen:

Stufe 1: Sichtung der Titel und Abstracts

Stufe 2: Sichtung der Volltexte

nach den folgenden Kriterien:

Eignungs-/ Einschlusskriterien

Konzept	Kriterien	Einschlusskriterien
Informationen	Methodenbeschreibung	Methoden zur Analyse von Entwicklungen Trends und Ableitung von Vorhersagen, insbesondere zur Berücksichtigung inhomogener Daten

Zugang	Volltext	Volltext verfügbar
Format	Studien, Methodenbeschreibungen	Reviews werden eingeschlossen

6.1.2 Review Dokumentation

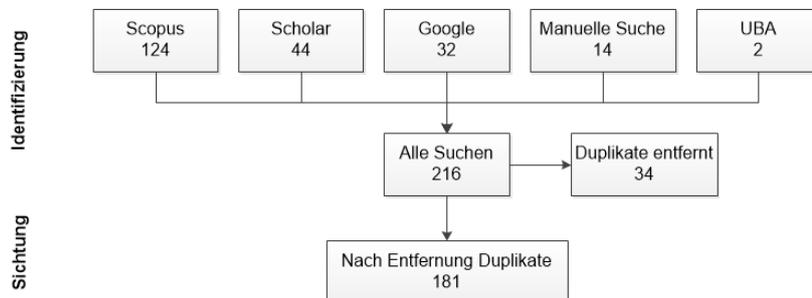
6.1.2.1 Suchergebnisse

Bei der Suche in Literaturdatenbanken und im Internet wurden insgesamt 200 Referenzen gefunden.

Quelle	Suchgleichung	Datum	Anzahl Treffer
Scopus	(TITLE ((air AND (pollut* OR quality)) OR particulate OR no2 OR (fine AND dust) OR ozone) AND TITLE (trend* AND analys*))	21.02.2018	124
Google Scholar	trend analysis ozone (erste 50 Ergebnisse)	22.3.2018	44
Google	trend analysis ozone (erste 50 Ergebnisse)	22.3.2018	32

Weiterhin hat der Auftraggeber zwei Dokumente übersandt, und bei der Suche in Referenzlisten oder anderweitig geprüfter Quellen wurden 14 Publikationen identifiziert.

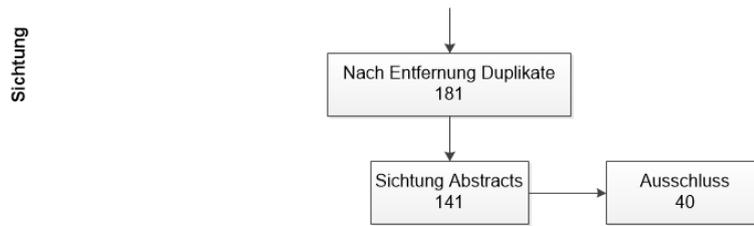
In allen Quellen zusammen wurden somit 216 Referenzen gefunden, darunter waren 34 Duplikate.



eigene Darstellung, BioMath GmbH

6.1.2.2 Eignungsprüfung der identifizierten Dokumente

Die verbliebenen 181 Publikationen wurden auf Erfüllung der Einschlusskriterien geprüft. Dazu wurden ausschließlich Titel und Abstract gesichtet. 40 Publikationen konnten ausgeschlossen werden, weil kein Abstract verfügbar war oder sie keine Methoden zur Analyse von Entwicklungen Trends und Ableitung von Vorhersagen beschrieben.



eigene Darstellung, BioMath GmbH

6.1.2.3 Extraktion von Informationen und Klassifikation

Aus den verbliebenen 141 Artikel wurden Informationen zu folgenden Punkten extrahiert:

Bibliografische Charakteristika: meistgefundene Autoren (mindestens 3 Artikel) und Erscheinungsjahr

Autor	Anzahl Artikel
Staehein, J.	10
Reinsel, G. C.	8
Tiao, G. C.	7
Bishop, L.	4
Kerr, J. B.	4
Miller, A. J.	4
Aneja, V. P.	3
Bodeker, G. E.	3
Cavender, K.	3
Claude, H.	3
Flynn, L. E.	3
Hollandsworth, S. M.	3
Kyrölä, E.	3
Laine, M.	3
Logan, J. A.	3
Reinsel, Gregory C.	3
Sather, M. E.	3
Stolarski, R. S.	3
Tiao, George C.	3
Umweltbundesamt	3
Wuebbles, D. J.	3
Zawodny, J.	3

Inhaltliche Charakteristika: Region, Parameter, Belastungsszenario, Anzahl der Messstationen und Zeitraum sowie statistische Methoden der Trendanalyse

Region	Anzahl	Prozent
global	22	15,6
USA	21	14,9
Nördliche Hemisphäre	11	7,8
Deutschland	9	6,4
Schweiz	9	6,4
Europa	6	4,3
Indien	6	4,3
China	5	3,5
Kanada	5	3,5
Italien	4	2,8
mehrere Regionen	3	2,1
Australien	2	1,4
Irland	2	1,4
Japan	2	1,4
Mexiko	2	1,4
Spanien	2	1,4
United Kingdom	2	1,4
Asien	1	,7
Chile	1	,7
Finnland	1	,7
Iran	1	,7
Korea	1	,7
Malaysia	1	,7
Niederlande	1	,7
Nordatlantik	1	,7
Nordpazifik	1	,7
Palästina	1	,7
Polargebiet	1	,7
Portugal	1	,7
Russland	1	,7
Südatlantik	1	,7
Taiwan	1	,7
Tibet	1	,7
keine Angabe	12	8,5
Total	141	100,0

Parameter	Anzahl	Prozent
-----------	--------	---------

Parameter	Anzahl	Prozent
ozone	92	63,1
PM10	2	1,4
sulphur dioxide	2	1,4
a pollutant such as ozone	1	,7
air pollutants - Air Quality Index (AQI)	1	,7
air pollution	1	,7
carbon monoxide (CO), nitrogen dioxide (NO2), ozone (O3) and sulfur dioxide (SO2)	1	,7
CO, NO2, SO2, O3, and PM10	1	,7
nitric acid, ozone, and hydrogen peroxide	1	,7
NO and NO(y) as well as O3 and CO	1	,7
NO2	1	,7
NO2, O3, CO, SO2, PM10	1	,7
NO2, PM 10 and ozone	1	,7
Nox, ozone	1	,7
oxidant	1	,7
ozone and nitrogen oxides (NO2 and NOx)	1	,7
ozone, Nox	1	,7
ozone, sukfur, nitrogen, particulate matter	1	,7
Particulate matter, ozone, nitrogen oxide	1	,7
PM 2.5	1	,7
PM-2.5	1	,7
PM10, carbon monoxide (CO), sulphur dioxide (SO2), ozone (O3), nitrogen mono- and dioxide (NOx), fine particles (PM2.5)	1	,7
PM10, CO, NOx (NO and NO2), Benzene and Toluene	1	,7
PM10, NO2	1	,7
PM10, NO2, Ozon	1	,7
PM10, PM2,5, NO2	1	,7
PM10, PM2.5, and PMc	1	,7
PM10, PM5.0, PM2.5, and PM1.0	1	,7
PM10, SO2	1	,7
PM2.5	1	,7
PM2.5 and surface ozone (O3)	1	,7
Ruß, Partikel, PM10	1	,7
Schwefeldioxid (SO2), Kohlenmonoxid (CO), Stickstoffmonoxid (NO), Stickstoffdioxid (NO2), Benzol, Toluol und m-/p-Xylol (BTX), Ozon (O3), Feinstaub (PM10), Feinstaub (PM2,5) und Ruß	1	,7
SO 2, dustfall, B[a]P, NO 2 and PM 10, Nox, CO, ozone pollution, total suspended particulate (TSP), Pb	1	,7

Parameter	Anzahl	Prozent
SO ₂ , NO _x , CO and PM _{2.5} , PM ₁₀	1	,7
SO ₂	1	,7
SO ₂ and particulate matter (PM)	1	,7
SO ₂ , CO, NO, NO ₂ , O ₃ and PM ₁₀	1	,7
So ₂ , NO ₂ , NO _x , PM ₁₀ , PM _{2,5} , C ₆ H ₆ , CO, O ₃	1	,7
sulfur dioxide (SO ₂) carbon monoxide (CO), suspended particulate matter (SPM), Nitrogen dioxide (NO ₂), and photochemical oxidants (OX)	1	,7
sulfur dioxide, ozone, nitrogen oxides, carbon monoxide and particulates	1	,7
sulphur dioxide (SO ₂), nitrogen dioxide (NO ₂) and respirable particulate matter (PM ₁₀)	1	,7
sulphur dioxide, nitrogen oxides, carbon monoxide, volatile organic compounds and ammonia	1	,7
total suspended particulates	1	,7
twenty-seven elements and ions	1	,7
keine Angabe	3	2,1
Total	141	100,0

Belastungsszenario	Anzahl	Prozent
urban	4	2,8
background mountain	1	,7
rural background	1	,7
urban and industrial center	1	,7
urban and rural	1	,7
urban background	1	,7
urban-near traffic, urban background and rural background	1	,7
urban, suburban, or rural	1	,7
urban, urban-background, rural	1	,7
urbanized	1	,7
keine Angabe	128	90,8
Total	141	100,0

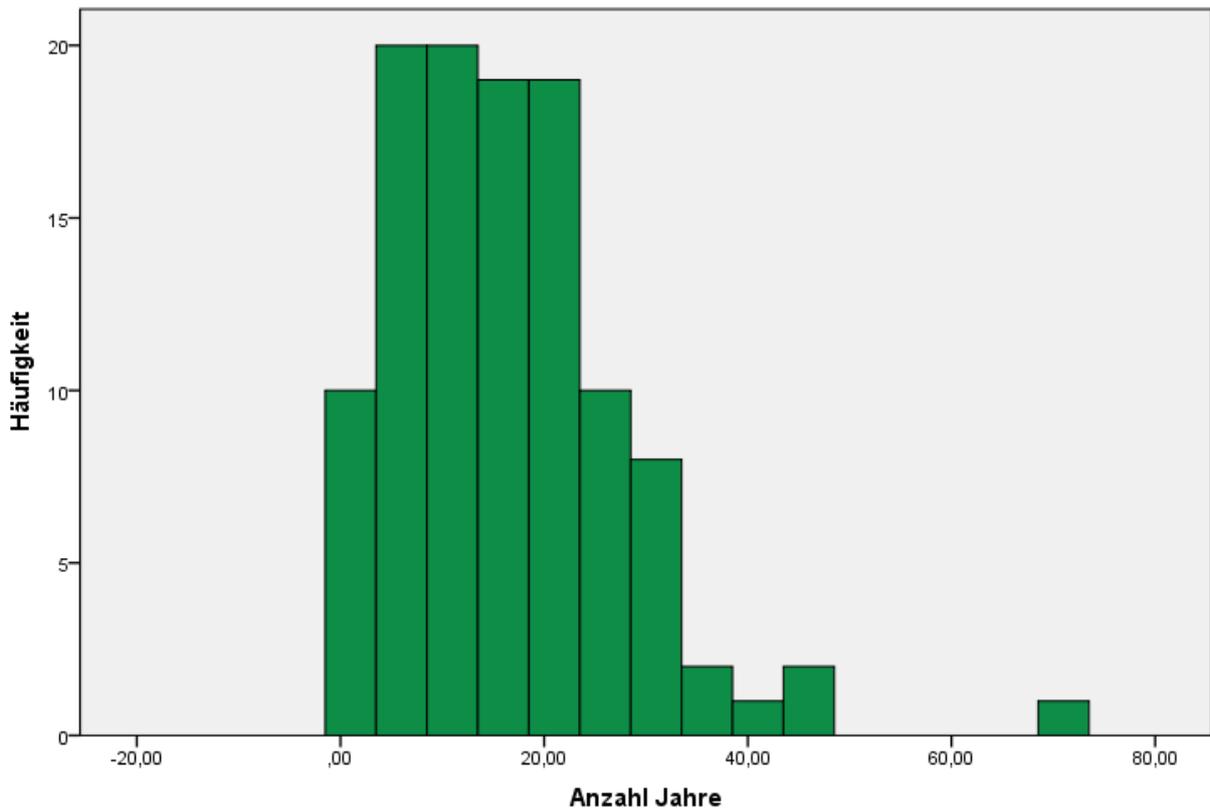
Daten/ Stationen	Anzahl	Prozent
'cohesive' SBUV(/2) total ozone dataset	1	,7
10 Umkehr stations	1	,7
11 locations in Eastern Canada	1	,7
11 monitoring sites	1	,7
13 stations	1	,7

Daten/ Stationen	Anzahl	Prozent
13 Umkehr stations	1	,7
14 stations	1	,7
158 harmonised monitoring stations	1	,7
20 monitoring stations	1	,7
29 monitoring stations	1	,7
29 Stationen	1	,7
29 stations in the former Soviet Union	1	,7
3 American monitoring stations	1	,7
35 Dobson stations	1	,7
36 ground-based Dobson recording stations	1	,7
36 Immissionsmessstationen	1	,7
5 Stationen	2	1,4
549 sites	1	,7
56 stations	1	,7
650 Messtationen	2	1,4
8 Stationen	1	,7
8 stations	1	,7
9 continuous monitoring stations	1	,7
a number of monitoring sites	1	,7
all Houston Area monitors	1	,7
all stations of urban air quality monitoring network	1	,7
automatic air-monitoring networks located in the MCMA (32 stations), GMA (8 stations), and MMA (5 stations)	1	,7
balloon soundings	1	,7
Brewer-Mast sonde	1	,7
cohesive data set from the SBUV-SBUV/2 satellite system	1	,7
collection of ground-based recording stations	1	,7
country wide datasets	1	,7
daily ozone emission dataset obtained from the World Ozone and Ultraviolet Data Centre (WOUDC)	1	,7
data from long-term monitoring stations	1	,7
data sets	1	,7
data, obtained from India Meteorological Department	1	,7
Dobson total column ozone data	1	,7
Dobson total ozone data, 29 stations	1	,7
Dobson total, Umkehr profile data	1	,7
downwind of major urban areas	1	,7
EMEP network	1	,7

Daten/ Stationen	Anzahl	Prozent
five stations	1	,7
four observational and reanalysis total ozone data sets	1	,7
four roadsides and four urban homes	1	,7
four sites	1	,7
four south central U.S. cities	1	,7
four South Central U.S. cities	1	,7
from the satellite Nimbus 7, corrected with ground remote sensing data	1	,7
German air quality monitoring network	1	,7
global collection of ground-based recording stations	1	,7
ground-based network and Nimbus 7 and NOAA 11	1	,7
Hohenpeissenberg	1	,7
hypothetical	1	,7
Italian Air Force's Vigna di Valle Observatory	1	,7
Kislovodsk High Mountain Station, Jungfrauoch	1	,7
Mace Head Atmospheric Research Station	1	,7
measurements from the Nimbus 7 TOMS	1	,7
measurements from TOMS onboard satellite Nimbus 7, Meteor 3 and Earth Probe	1	,7
measures taken by the City Council of Madrid	1	,7
monitoring data	1	,7
Naha (Okinawa Island), Japan	1	,7
National Air Monitoring Sites (NAMS) and other sites	1	,7
network of Dobson stations	1	,7
Nimbus-7 and Meteor-3 TOMS column ozone data	1	,7
Nimbus-7 TOMS total ozone measurements	1	,7
NYS Department of Environmental Conservation sites in Rochester, NY	1	,7
obtained from the Environmental Protection Agency	1	,7
ozonesonde stations	1	,7
radiometer GROMOS	1	,7
Raleigh and Charlotte, North Carolina	1	,7
records from EMEP (European Monitoring and Evaluation Programme) sites	1	,7
SAGE I/II ozone data	1	,7
satellite measurements by the Stratospheric Aerosol and Gas Experiment (SAGE) II and the Global Ozone Monitoring by Occultation of Stars (GOMOS) instruments	1	,7
satellite remote sensing datasets	1	,7
SBUV/SBUV-2 merged ozone data	1	,7
selected European stations	1	,7
semi-rural	1	,7

Daten/ Stationen	Anzahl	Prozent
Seoul and Taegu	1	,7
set of selected surface ozone (nine stations) and ozone vertical profile measurements (from six stations)	1	,7
seven monitoring sites (Nizamuddin, Ashok Vihar, Shahzada Baug, Shahadara, Janakpuri, Sirifort and ITO)	1	,7
several	1	,7
simulations	1	,7
six monitoring stations	1	,7
six specific locations	1	,7
some monitoring stations	1	,7
State of Saxony PM 2.5 monitoring values	1	,7
state pollution control board website	1	,7
Stratospheric Aerosol and Gas Experiment (SAGE) II ozone data set.	1	,7
thousands of monitoring sites around the world	1	,7
three monitoring stations, whose locations have not changed over the years	1	,7
TOMS	1	,7
TOMS and SBUV measurements	1	,7
TOMS data	1	,7
TOMS, Meteor-3	1	,7
Total Ozone Mapping Spectrometer (TOMS) on the Nimbus 7 satellite	1	,7
two high-elevation sites, one low-elevation site	1	,7
two monitoring sites	1	,7
two sites	1	,7
two specific sites Azusa and San Bernardino	1	,7
two stations downwind	1	,7
keine Angabe	39	27,6
Total	141	100,0

Mittelwert	N	Std. Abweichung	Minimum	Maximum	Median
16,1518	112	11,24679	1,00	72,00	15,0000



eigene Darstellung, BioMath GmbH

Methoden

a new method of regression for sparsely sampled data sets for use with time-series analysis

a statistical model: extreme value distribution, location parameter changes in time accordingly to a regression model.

analyzing extreme values

annual statistical multiple regression relationships applied to certain "fixed values" of the influencing meteorological parameters

bootstrap method, extreme value statistics and a nonparametric test

classical additive time series decomposition technique, adjustment was made using Kolmogorov-Zurbenko filters followed by stepwise re

correlation

Daniel trend test

different linear trend for each month of the year, regression random effects models, additional explanatory factors

frequency domain statistical mode

geometric means, moving average method and time series analysis

GEOS-CHEM model, a global three-dimensional model of the troposphere

global chemistry-climate model EMAC

GM (1,1) model

hierarchical statistical state space model

Methoden

first-order autoregression of the residuals [so-called AR(1) models]

Jahresmittelwerte

Kolmogorov-Zurbenko filtering method: time series are being decomposed into their short-, mid-, and long-term component. Trend analys

linear regression analysis

linear regression method

linear regression model

linear regression techniqu

linear regressions

linear regressive framework

linearer Regression

logarithmic mean Divisia index decomposition

long-range dependence time series model

MANN - KENDALL test

Mann-Kendall regression with Sen's slope to estimate the trends and seasonality, piecewise regression, multivariate linear regression

Mann-Kendall test

Mann-Kendall's (MK) trend analysis, autoregressive integrated moving average (ARIMA)

median time series and 25%, 75% quantiles

methodology to normalize air quality trends with respect to typical year meteorology

Mittelwerte

Monte Carlo analysis

multifunctional regression model

multiple linear regression model

multiple linear regression model incorporating meteorological parameters

multiple regression analysis

multiple regression models involving combinations of explanatory variables

Multivariate Adaptive Regression Splines (MARS) methodology and principal component analysis, bootstrapping.

multivariate adaptive regression splines methodology, block bootstrapping of the models' residuals

multivariate linear regression (MLR) and ensemble empirical mode decomposition (EEMD)

multivariate linear regression model

multivariate statistical techniques such as Principal Component Analysis (PCA) and Absolute Principal Component Scores (APCS)

non-homogenous Poisson process: account for the confounding effects of meterological conditions

order restricted inference and isotonic regression

percentile of daily mean concentration values corresponding to the xx highest daily mean.

ranks six levels

Methoden

Regressionsmodell

regression analysis

regression of time series

regression time series models that include seasonal trend, and other factors, random effects model for the individual station trend

regression time series models that include seasonal, trend, and other factors, random effects model for the individual station trend

regression time series models, which include seasonal and trend factors

regression-time series models, which include seasonal components, a linear trend term, a solar flux term, and an autocorrelated autor

review

robust multilinear parametric trend model which includes a linear term, the solar variability

spatial and temporal trend analysis using the generalized additive mixed model (GAMM)

spatial interpolation dispersion model IMMIKART

spectral analysis of data based on Fourier transform methods

statistical methods

statistical model containing quasibiennial, seasonal, and semiannual oscillations, a linear component, and a first-order autoregressi

statistical model evaluating the residual long-term variability due to combined effect of factors

statistical model, including terms for seasonal variation, linear trend, quasi-biennial oscillation, solar cycle and second-order a

statistical post processing (quality assessment)

theory of extreme values, bootstrap testing procedure

three-year moving average, linear regression, quadratic logistic regression of odd ratios, confidence interval estimation with Monte

time series analyses, spectral analysis and running averages

time series decomposition techniques

time series model

time series modeling

time series modelling

time series models incorporate seasonal, trend, and noise factors

traditional statistical time series trend, or hockey-stick algorithm

trend model with atmospheric explanatory variables

trend plots and box plots and with analysis of variance techniques

trend surface analysis, a technique of areal regression

two-dimensional (UIUC 2-D) chemical transport model of the global atmosphere

using atmospheric variables in a stepwise regression model

year-specific statistical models for meteorological adjustment

6.2 R-Programmcode

```
#####
# Projekt:      Ableitung sicherer Trendaussagen zur Entwicklung der Luftqualität in
Deutschland
#              UFOPLAN 2017 - FKZ 3717522540 AZ: 50 475/149
#
# Programmname: Modellrechnung jährlicher Differenzen V100-20181011.R
# Autor:       Jörg Schmidtke, BioMath GmbH
# Letzte Änderung: 11.10.2018
# PART 1:     Eingaben/Vorgaben
# PART 2:     Berechnung der Modell-Mittelwerte und Regressionsanalyse für den
#             linearen Trend
# PART 3:     Ausgabe der Modell-Mittelwerte und Regressionsparameter für den
#             linearen Trend
#####

#####
# Part 1: Eingaben/Vorgaben
#####

# Löschen aller Variablen
rm(list=ls())

# Arbeitsverzeichnis
arbeitsverzeichnis <- "D:/UBA-Luft/R"

# Dateinamen für die Eingabedaten und Ausgabedaten
datendateiname      <- "Data/Ozon-Jahresmittelwerte_1990-2017.txt"
ergebnisdateiname_modellmittelwerte <- "Out/Ozon-Modell_2008-2017 (Modell-
Mittelwerte).txt"
ergebnisdateiname_regression      <- "Out/Ozon-Modell_2008-2017 (Regression).txt"

# Auswertungszeitraum
# ACHTUNG: Die Variable jahr_von muss immer größer sein als das erste Jahr in den Eingabedaten.
#         Für das jahr_von wird die Differenz aus jahr_von - jahr_von-1 gebildet.

jahr_von <- 2008
jahr_bis <- 2017

#####
# Part 2: Berechnung der Modell-Mittelwerte und Regressionsanalyse für den linearen Trend
#####

# Systemvorbereitungen
setwd(arbeitsverzeichnis) # Working Directory
library(data.table)       # zum Daten managen
library(plyr)             # Umwandeln von Listen zu Tabellen

# Variablen für das Handling des Zeitabschnittes
jahr_diff      <- jahr_bis - jahr_von # Länge des Zeitabschnittes
jahr_von_minus1 <- jahr_von -1       # Beginn des Zeitabschnittes für die Berechnung
der Differenzen

# Import der Datendatei
rohdaten <- fread(datendateiname,
```

```

header = TRUE,          # 1. Zeile = Spaltenname
dec = ",",              # Komma als Dezimaltrennzeichen
stringsAsFactors = TRUE # Strings als Variablen mit Stufen

# Variable Kategorie als Faktor
rohdaten$Kategorie <- as.factor(rohdaten$Kategorie)

# nur Variablen der Messwerte der Jahre
spaltejahr <- substr(names(rohdaten),1,1) %in% c("V") # wenn erstes Zeichen des Spalten-
namens = V
jahrdaten <- rohdaten[, ..spaltejahr]                # Reduziere auf Jahresspalten

# Werte <= 0 intern auf N/A setzen
jahrdaten[jahrdaten <= 0] <- NA

# Auswertungsdaten von jahr_von-1 bis jahr_bis
auswertungsjahre <- c(jahr_von_minus1:jahr_bis)
spalteauswertungsjahr <- as.numeric(substr(names(jahrdaten),2,5)) %in% auswertungsjahre #
wenn Spaltenname als Zahl im Bereich JahrBeginn bis JahrEnde
auswertungsdaten <- jahrdaten[, ..spalteauswertungsjahr] #
Reduziere auf Auswahl

# Differenzen zwischen den Messwerten der Jahre, beginnend mit jahr_von-1
for (i in 2:(jahr_diff+2)){ # Schleife von 2 bis Anzahl der Jahre
  j <- i-1
  auswertungsdaten <- cbind(auswertungsdaten,
                           diff = auswertungsdaten[, ..i] - auswertungsdaten[, ..j])
}

# erste Spalte löschen, d.h. Variable jahr_von-1 entfernen
auswertungsdaten[,1] <- NULL

# Kategorie anfügen
auswertungsdaten <- cbind(auswertungsdaten, Kategorie=rohdaten$Kategorie)

# Anzahl Messjahre berechnen und in njahr speichern
spaltenv <- names(auswertungsdaten)[substr(names(auswertungsdaten), 1, 1) == "V"]
auswertungsdaten[ , njahr := Reduce(`+`, lapply(.SD, function(x) !is.na(x))), .SDcol =
spaltenv ]

# Daten mit mindestens einem Messjahr auswählen
auswertungsdaten <- auswertungsdaten[njahr>0,]

# Mittelwerte der Spalten je Kategorie
spaltendiff <- names(auswertungsdaten)[substr(names(auswertungsdaten), 1, 4) == "diff"]
ergebnis <- auswertungsdaten[ , lapply(.SD, mean, na.rm=TRUE), by="Kategorie",.SDcol =
spaltendiff]
names(ergebnis)[-c(1)] <- paste0("mw.",names(ergebnis)[-c(1)])

# Aufaddieren
spaltemwdiff <- substr(names(ergebnis),1,7) %in% c("mw.diff", "Kategor")
mwdifferenzen <- ergebnis[, ..spaltemwdiff]

# Spaltennamen vergeben
spaltennamensummwdiff <- paste0("sum.", names(mwdifferenzen))

for (i in 2:dim(mwdifferenzen)[2]){
  summe_i <- mwdifferenzen[, rowSums(.SD), by=Kategorie, .SDcols=c(2:i)]
  names(summe_i)[2] <- spaltennamensummwdiff[i]
  ergebnis <- merge(ergebnis, summe_i, by="Kategorie")
}

```

```

}

# Beginn: Gewichtete Mittelwerte des letzten Jahres berechnen

# Jahr als neue Variable
for (i in 1:(jahr_diff+1)){ # Schleife von 1 bis Anzahl der Jahre
  auswertungsdaten <- cbind(auswertungsdaten,
                             jahr = auswertungsdaten[, ..i]/auswertungsdaten[, ..i] + i-2
                             + jahr_von)
}

# jahrmin und jahrmax bestimmen
auswertungsdaten$zeile <- seq.int(nrow(auswertungsdaten))
spalten <- paste0("jahr.V",seq(from=jahr_von,to=jahr_bis,by=1))
auswertungsdaten[, c("jahrmin", "jahrmax") := .(min(.SD, na.rm=TRUE),
                                                  max(.SD, na.rm=TRUE)), by=zeile,
.SDcols=spalten]
# jahrdiff berechnen
auswertungsdaten[, jahrdiff := jahrmax-jahrmin]
# jahrdiff <=0 intern auf N/A setzen
auswertungsdaten[jahrdiff <= 0] <- NA
# gewicht berechnen
auswertungsdaten[, gewicht := njahr-(1-(njahr/jahrdiff))]
# gewicht <=0 intern auf N/A setzen
auswertungsdaten[gewicht <= 0] <- NA

# Variable des letzten Jahres
varletztesjahr <- paste0("V", jahr_bis)

# falls für das letzte Jahr kein Messwert vorliegt, dann gewicht intern auf N/A setzen
auswertungsdaten[, gewicht := gewicht * .SD/.SD, .SDcol=varletztesjahr]

# Produkt aus dem Gewicht und dem Messwert des letzten Jahres bestimmen
auswertungsdaten[, produkt := gewicht * .SD, .SDcol=varletztesjahr]
# Summation gewicht und produkt
GundP <- auswertungsdaten[, lapply(.SD, sum, na.rm=TRUE), by="Kategorie",
.SDcol=c("gewicht", "produkt")]
names(GundP)[2:3] <- paste0("sum.", names(GundP)[2:3])
ergebnis <- merge(ergebnis, GundP, by="Kategorie")
ergebnis[, mw.gewichtet := sum.produkt / sum.gewicht]

# Ende: Gewichtete Mittelwerte berechnen

# Berechnung der Modell-Mittelwerte
varsummwdiffletztesjahr <- paste0("sum.mw.diff.V", jahr_bis)
ergebnis[, sum_mw_diffletztesjahr := .SD, .SDcol=varsummwdiffletztesjahr]
spaltensummwdiff <- names(ergebnis)[substr(names(ergebnis), 1, 11) == "sum.mw.diff"]
ergebnis[, c(paste0("gew.mw.",spaltensummwdiff)) := mw.gewichtet + .SD -
sum_mw_diffletztesjahr, .SDcol=spaltensummwdiff]

# Ergebnis reduzieren auf Kategorie und Modell-Mittelwerte
spaltenmodellmittelwerte <- substr(names(ergebnis), 1, 6) %in% c("Katego", "gew.mw")
ergebnismodellmittelwerte <- ergebnis[, ..spaltenmodellmittelwerte]

# Daten für Regression vorbereiten
spaltenfuerreg <- substr(names(ergebnis), 1, 18) %in% c("gew.mw.sum.mw.diff", "Katego-
rie")
datenregression <- ergebnis[, ..spaltenfuerreg]
datenregressiontrans <- melt.data.table(data=datenregression,

```

```

        id.vars="Kategorie",
        value.name = "gew.mw.sum.mw.diff")
datenregressionstrans[, Jahr := as.numeric(substr(variable, 21, 24))]

# Regression pro Kategorie
kategorien <- levels(datenregressionstrans$Kategorie)
listeregressionen <- matrix(list(), length(kategorien), 3)
rownames(listeregressionen) <- kategorien
colnames(listeregressionen) <- c("reg", "summary", "steigung")

for (i in kategorien){

  # Teildatensatz pro Kategorie
  datenregressionstransprokategorie <- subset(datenregressionstrans, Kategorie==i)

  # Regression
  regression <- lm(data = datenregressionstransprokategorie,
                   formula = gew.mw.sum.mw.diff ~ Jahr)

  # Anstieg
  steigung <- data.table(coef(regression) ["Jahr"],
                        confint(regression, "Jahr", level=0.95))

  names(steigung) <- c("steigung", "unteres 95Konfidenzintervall", "oberes
95Konfidenzintervall")

  listeregressionen[[i, "reg"]] <- regression
  listeregressionen[[i, "summary"]] <- summary(regression)
  listeregressionen[[i, "steigung"]] <- steigung
}

# Zusammenführen der Regressionsergebnisse
ergebnisregression <- ldply(listeregressionen[, "steigung"], data.frame, .id="Kategorie")

#####
# Part 3: Ausgabe der Modell-Mittelwerte und Regressionsparameter für den linearen Trend
#####

# Exportieren der Modellmittelwerte
fwrite(x = ergebnismodellmittelwerte, file = ergebnisdateiname_modellmittelwerte,
sep="\t", dec = ",")

# Exportieren der Regressionsparameter
fwrite(x = ergebnisregression, file = ergebnisdateiname_regression, sep="\t", dec = ",")

#####
# Ende
#####

```

7 Quellenverzeichnis

Augustin, Colette u.a. (2016): Air pollution trends in the EMEP region between 1990 and 2012. Kjeller, Norway

Gilbert, Richard (1987): Statistical Methods for Environmental Pollution Monitoring. Canada

Hennemuth, Barbara u.a. (2013): Statistische Verfahren zur Auswertung von Klimadaten aus Modell und Beobachtung. Hamburg, Deutschland

Klose, Suan und Graff, Arno (2009): Trendanalyse von Jahresmittelwertzeitreihen von NO₂, Feinstaub und Ozon im Zeitraum von 1995 bis 2007 in Abhängigkeit von der Messnetzhomogenität. Düsseldorf, Deutschland

Schlittgen, Reiner und Streitberg, Bernd (2001): Zeitreihenanalyse. München Wien