

# Satelliten-basierte Schadstoffkartierung: ML-Anwendungen und Analyse von atmosphärischen Einflussfaktoren

Masterarbeit im Studienfach Physik

*Elias Papke*

Abgegeben am  
18. Juni 2025

angefertigt am  
DEUTSCHEN ZENTRUM FÜR LUFT- UND RAUMFAHRT E.V.  
OBERPFAFFENHOFEN  
und am  
INSTITUT FÜR PHYSIK DER UNIVERSITÄT AUGSBURG  
ATMOSPÄRENFERNERKUNDUNG

Erstgutachter: Prof. Dr. Michael Bitter  
Zweitgutachter: Dr. Andreas Hörner

Betreuerin: Dr. Jana Handschuh  
Betreuer: Dr. Frank Baier





# Abstract

The large-scale monitoring of air pollution is of central importance for public health and the effective implementation of measures to improve air quality. It is therefore essential to be able to distinguish which changes in air quality are caused by human activity and which can be attributed to weather phenomena. This work investigates the influence of meteorological factors on ground-level air pollution, as well as the spatial and temporal variability of its distribution.

To this end, the concentrations of  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  and  $\text{O}_3$  near the ground in an area containing Germany and parts of neighboring countries are derived using the Random Forest algorithm. It processes satellite observations, meteorological parameters and other auxiliary variables together with in-situ measurements of the pollutants.

The trained Models provide reliable values for ground-level  $\text{PM}_{2.5}$  ( $R^2 = 0.76$ ) and  $\text{NO}_2$  ( $R^2 = 0.79$ ) and very good results for  $\text{O}_3$  ( $R^2 = 0.91$ ). Accuracy deviates when the test and training data are spatially or temporally separated. Methods of explainable artificial intelligence (XAI) are used to evaluate the significance of the individual model parameters. These show that the accuracy of the predictions for  $\text{PM}_{2.5}$  depends mainly on temporal information while  $\text{NO}_2$  relies on spatial and  $\text{O}_3$  on meteorological information. Heavy precipitation and strong winds are among the most important  $\text{PM}_{2.5}$ - and  $\text{NO}_2$ -reducing environmental factors, with air temperature and intense solar radiation greatly increasing  $\text{O}_3$ -concentration. An investigation of the weekend effect on ozone production and the  $\text{O}_3$ - $\text{NO}_2$  model dependence suggest a shift from  $\text{NO}_x$ -sensitive conditions in winter to  $\text{NO}_x$ -saturated conditions in summer.

Local Indicators of Spatial Association (LISA) are used to identify hot- and coldspots in the derived pollution data within the study area. Urban regions are hotspots for  $\text{PM}_{2.5}$  and  $\text{NO}_2$  while being coldspots for  $\text{O}_3$ . Additionally, the surrounding geography and vegetation may play a role in the spatial distribution depending on the pollutant.



# Kurzzusammenfassung

Die großflächige Überwachung der Luftverschmutzung ist von zentraler Bedeutung für die öffentliche Gesundheit und das effektive Umsetzen von Maßnahmen zur Verbesserung der Luftqualität. Dabei ist es essenziell unterscheiden zu können, welche Veränderungen in der Luftqualität Ursache menschlichen Zutuns sind und welche sich auf Wetterphänomene zurückführen lassen. In dieser Arbeit wird daher der Einfluss meteorologischer Faktoren auf die Schadstoffsituation am Boden, sowie die räumliche und zeitliche Variabilität der Schadstoffverteilung näher untersucht.

Dazu wird mithilfe des Random Forest Algorithmus jeweils die bodennahe  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ -, und  $\text{O}_3$ -Konzentration in Deutschland und Teilen der angrenzenden Länder abgeleitet. Es werden tägliche Satellitendaten zusammen mit meteorologischen Parametern und weiteren Hilfsvariablen sowie in-situ Messungen der Schadstoffe verarbeitet.

Die Modelle liefern zuverlässige Werte für die Bodenkonzentrationen von  $\text{PM}_{2.5}$  ( $R^2 = 0.76$ ) und  $\text{NO}_2$  ( $R^2 = 0.79$ ) und sehr gute Ergebnisse für  $\text{O}_3$  ( $R^2 = 0.91$ ) mit abweichenden Genauigkeiten bei räumlicher oder zeitlicher Trennung von Test- und Trainingsdaten. Zur Bewertung der Signifikanz der einzelnen Modellparameter werden Methoden der erklärbaren Künstlichen Intelligenz (XAI) eingesetzt. Diese zeigen auf, dass die Genauigkeit der Vorhersagen für  $\text{PM}_{2.5}$  hauptsächlich von zeitlichen, für  $\text{NO}_2$  von räumlichen und für  $\text{O}_3$  von meteorologischen Informationen abhängt. Starke Regenfälle und Winde gehören zu den bedeutendsten  $\text{PM}_{2.5}$ - und  $\text{NO}_2$ -reduzierenden Umweltfaktoren, wobei die  $\text{O}_3$ -Konzentration vor allem durch hohe Lufttemperaturen und intensive Sonneneinstrahlung steigt. Die Untersuchung des Wochenendeffektes auf die Ozonproduktion und der  $\text{O}_3$ - $\text{NO}_2$ -Modellabhängigkeit weist auf einen Wechsel von  $\text{NO}_x$ -gesättigten Bedingungen im Winter zu  $\text{NO}_x$ -sensiblen Bedingungen im Sommer hin.

Lokale Indikatoren räumlicher Autokorrelation (LISA) wurden eingesetzt um Hot- und Coldspots in den flächenhaften Vorhersagen der Schadstoffverteilung zu identifizieren. Hotspots in den Bodenkonzentrationen von  $\text{PM}_{2.5}$  und  $\text{NO}_2$  sowie Coldspots von  $\text{O}_3$  sind dabei in den Stadtregionen zu beobachten, wobei je nach Schadstoff die umgebende Geographie und Vegetation eine Rolle in der räumlichen Verteilung spielen.



# Inhaltsverzeichnis

<b>Abstract</b>	<b>iii</b>
<b>Kurzzusammenfassung</b>	<b>v</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Grundlagen</b>	<b>5</b>
2.1 Kinetik chemischer Reaktionen . . . . .	5
2.2 Feinstaub . . . . .	6
2.3 Stickoxide . . . . .	8
2.4 Ozon . . . . .	10
2.5 Satellitenmessungen . . . . .	15
2.5.1 Aerosol Optische Dicke . . . . .	15
2.5.2 Troposphärisches NO <sub>2</sub> . . . . .	17
<b>3 Daten</b>	<b>19</b>
3.1 Satellitendaten . . . . .	19
3.2 Stationsdaten für bodennahe Schadstoffwerte . . . . .	20
3.3 Meteorologische Felder . . . . .	21
3.4 Ortsbezogene Daten . . . . .	21
<b>4 Methodik</b>	<b>23</b>
4.1 Random Forest Modell . . . . .	23
4.2 Modellaufbau . . . . .	24
4.3 Modelltestverfahren . . . . .	25
4.4 Modellanalysemethoden . . . . .	28
4.4.1 Feature Importance Analyse . . . . .	28
4.4.2 Partielle Abhängigkeiten . . . . .	29
4.5 Analysemethoden für flächenhafte Daten . . . . .	29
4.5.1 Vorhersage und Weiterverarbeitung der Flächendaten . . . . .	29
4.5.2 Datenaufteilung und Wetterlagenklassifikation . . . . .	30
4.5.3 Globale räumliche Autokorrelation . . . . .	31
4.5.4 Local Indicators of Spatial Association (LISA) . . . . .	32
<b>5 Ergebnisse</b>	<b>35</b>
5.1 Modellleistung . . . . .	35
5.2 Modell- und Variablenabhängigkeiten . . . . .	36
5.3 Räumliche Verteilung und zeitliche Entwicklung . . . . .	41
5.4 Saisonale Schadstoffverteilung . . . . .	47
5.5 Schadstoffverteilung bei verschiedenen meteorologischen Bedingungen . . . . .	48
5.5.1 Auswirkung der Windbedingungen . . . . .	48
5.5.2 Auswirkung des atmosphärischen Feuchtegehalts . . . . .	53
5.6 Wochenendeffekt . . . . .	54
5.7 Saisonale Abhängigkeiten zwischen Ozonkonzentration und NO <sub>2</sub> . . . . .	55

<b>6</b>	<b>Diskussion</b>	<b>57</b>
6.1	Modellgenauigkeit . . . . .	57
6.2	Physikalische Auslegung der Modellabhängigkeiten . . . . .	58
6.3	Ursachen räumlicher und saisonaler Variabilität der Schadstoffverteilung . .	63
6.4	Bestimmung des Ozonproduktionsregimes . . . . .	65
6.5	Irreguläre Einflüsse auf die Schadstoffentwicklung . . . . .	67
6.6	Einfluss des Feuchtegehalts der Atmosphäre . . . . .	68
6.7	Schadstofftransport . . . . .	70
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>71</b>
	<b>Literatur</b>	<b>75</b>
	<b>Abbildungsverzeichnis</b>	<b>91</b>
	<b>Tabellenverzeichnis</b>	<b>95</b>
	<b>Anhang</b>	<b>97</b>
A	Abbildungen flächenhafter Schadstoffmittelwerte . . . . .	98
B	Abbildungen zur Hot- und Coldspotanalyse . . . . .	103
C	Zusätzliche Abbildungen . . . . .	109
	<b>Danksagung</b>	<b>111</b>

# 1 Einleitung

Luftverschmutzung zählt zu den drängendsten gesellschaftlichen Herausforderungen des 21. Jahrhunderts und kann sowohl schwerwiegende Folgen für die Gesundheit der betroffenen Bevölkerung, als auch für empfindliche Ökosysteme und das Klima haben. Zu den bedeutendsten Luftschadstoffen gehören Feinstaub ( $\text{PM}_{2.5}$ ), Stickstoffdioxid ( $\text{NO}_2$ ) und bodennahes Ozon ( $\text{O}_3$ ), welche mit einer Vielzahl von Gesundheitsschäden in Verbindung gebracht werden.  $\text{PM}_{2.5}$  beschreibt besonders feine Partikel, die mit ihrer Größe von weniger als  $2.5\text{ }\mu\text{m}$  leicht in den menschlichen Organismus eindringen können [1] und so bei anhaltender Exposition für Atemwegs- und Herz-Kreislauf-Erkrankungen verantwortlich sein können und das Risiko für andere ernste Erkrankungen wie z.B. Lungenkrebs erhöhen [2].  $\text{NO}_2$ , das vor allem aus Verbrennungsprozessen stammt, kann die Atemwege schädigen und Erkrankungen wie z.B. Asthma verschärfen [3]. Zudem trägt es wesentlich zur Bildung von Ozon bei [4], welches eine ähnlich reizende Wirkung auf die Atemwege hat [5] und Schäden an Pflanzen verursachen kann [6]. Schadstoffe wie  $\text{NO}_2$  oder auch Schwefeldioxid ( $\text{SO}_2$ ) sind zudem zu großen Teilen für die Bildung von saurem Regen verantwortlich, welcher empfindliche Ökosysteme schädigen kann [7][3].

Die Weltgesundheitsorganisation (WHO) schätzt, dass im Jahr 2019 etwa 4.9 Millionen vorzeitige Todesfälle auf die Verschmutzung der Außenluft in urbanen Regionen zurückzuführen waren [8] und fast alle Menschen gesundheitsgefährdenden Mengen an Luftschadstoffen ausgesetzt waren [4].

Als Reaktion auf die zunehmende Bedrohung durch Luftverschmutzung wurden in den letzten Jahrzehnten die globalen und regionalen Anstrengungen zur Verringerung der Schadstoffemissionen deutlich verstärkt. Die Europäische Union verordnete in den Jahren 1996 und 2008 neue gesetzlich verbindliche Obergrenzen für Luftschadstoffe wie  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$ ,  $\text{NO}_2$ ,  $\text{SO}_2$  und  $\text{O}_3$  [9][10]. Seit 2008 wurden in vielen großen deutschen Städten Umweltzonen mit schärferen Abgasvorschriften für Fahrzeuge eingeführt, die vor allem auf eine Verringerung der Feinstaub- und Stickoxidbelastung abzielen [11]. Im Jahr 2018 wurden erneut verschärfte Fahrverbote für Dieselfahrzeuge angeordnet [12]. Obwohl  $\text{PM}_{2.5}$ - und  $\text{NO}_2$ -Konzentrationen in Europa über die letzten Jahre zurückgingen, werden vor allem in dicht besiedelten Gebieten wie Städten noch immer regelmäßig Schadstoffwerte gemessen, die über die von der WHO und der EU festgelegten Grenzwerte hinausgehen [13][14].

Um zukünftig die mit Luftverschmutzung verbundenen Gesundheitsrisiken besser einschätzen zu können und effektivere Maßnahmen zur Verbesserung der Luftqualität zu erarbeiten und umzusetzen, ist ein besseres Verständnis der Prozesse, welche zu hohen Bodenkonzentrationen von  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  führen, notwendig. Hierfür werden räumlich und zeitlich hochaufgelöste Daten benötigt, welche die Schadstoffsituation am Boden möglichst genau beschreiben. Europa kann dazu auf ein umfassendes Netz von in-situ Messstationen zurückgreifen [15], welche stündliche Informationen zur Luftzusammensetzung liefern und es ermöglichen, die zeitliche Entwicklung der Luftqualität an den Stationsstandorten zu überwachen. Allerdings sind diese Messstationen zahlenmäßig begrenzt räumlich, inhomogen verteilt und überwiegend in Städten konzentriert. Dadurch ergeben sich großflächige Lücken in den Schadstoffinformationen, deren Schließung eine

zentrale Herausforderung der umwelt- und gesundheitsorientierten Forschung darstellt.

Satellitenmessungen sind hier von zentraler Bedeutung, denn sie können wichtige flächen-deckende Informationen zur horizontalen Schadstoffverteilung liefern. Aufgrund der festen Orbits polar-umlaufender Satelliten sind diese jedoch in ihrer zeitlichen Auflösung meist auf eine Messung pro Tag beschränkt. Die Messung stützt sich auf die Reflexion von Sonnenlicht durch den Erdboden, weshalb durch Wolken zusätzliche Lücken in den Daten entstehen können. Zum einen verhindern diese die Beobachtung des darunterliegenden Atmosphärenbereichs, zum anderen besitzen sie eine im Vergleich zur Erdoberfläche deutlich stärkere Albedo [16]. Satellitenbeobachtungen liefern in der Regel Informationen über Schadstoffkonzentration in einer Luftsäule. Ohne Kenntnis des Höhenprofils der Schadstoffverteilung innerhalb der Atmosphäre ist es daher nur schwer möglich auf die Bodenkonzentration eines Schadstoffes zu schließen. Die vertikale Verteilung der Luftschadstoffe wird mitunter stark von meteorologischen Parametern wie Luftfeuchtigkeit, Grenzschichthöhe und Sonneneinstrahlung, sowie von der Stärke der lokalen Emissionen bestimmt. Diese Zusammenhänge sind aufgrund der Vielzahl der in der Atmosphäre wirkenden Prozesse sehr komplex und nur schwer zu quantifizieren. In den letzten Jahren hat sich maschinelles Lernen (ML) als leistungsfähiges Instrument erwiesen, um bodennahe Schadstoffkonzentrationen aus Satellitendaten herzuleiten, da es die Möglichkeit bietet, die Beziehung zwischen Umwelteinflüssen, Satellitenmessungen und Bodenkonzentrationen schnell und effektiv zu erlernen. Hierbei können bodenbasierte in-situ Messungen als Trainingsdaten fungieren und das Modell ggf. um diverse Hilfsvariablen erweitert werden, welche die verschiedenen Umwelteinflüsse sowie das menschliche Wirken beschreiben. Derartige Verfahren sind zudem, im Gegensatz zu Alternativen wie Chemische Transportmodelle (CTMs), nicht an die Verfügbarkeit von geeigneten Randbedingungen geknüpft [17] und werden in ihrer horizontalen Auflösung nur durch die Auflösung der Satellitenbeobachtungen und der Hilfsparameter beschränkt.

Unter den bereits erfolgreich zur Vorhersage von  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  eingesetzten ML-Modellen findet sich eine Vielzahl unterschiedlicher Algorithmen, wie Gradient Boosting (GB) [18, 19, 20, 21, 22, 23, 24, 25, 26], Random Forest (RF) [20, 22, 23, 25, 27, 28, 29, 30, 31, 32, 33] und Künstliche Neuronale Netze (ANN) [22, 23, 33, 34, 35], sowie komplexere bzw. hybride Modelle [26, 36, 37, 38, 39]. Durch die Verarbeitung meteorologischer Daten in den Modellen, welche einen entscheidenden Beitrag zur akkuraten Vorhersage liefern [40], bringen die Modelle die Luftverschmutzung am Boden mit den jeweiligen Wetter- und Umweltbedingungen in Verbindung. Diese können die Luftqualität zum Teil stark beeinflussen, was das effektive Umsetzen von Maßnahmen zur Verbesserung der Luftqualität erschweren kann. Der Einfluss meteorologischer Faktoren auf die Bodenkonzentration von Luftschadstoffen ist daher Bestandteil aktueller Forschung und wird regelmäßig von Studien thematisiert [41, 42, 43, 44, 45].

Ziel dieser Arbeit ist es, zum besseren Verständnis der treibenden Faktoren von Luftverschmutzung, den Einfluss atmosphärischer Bedingungen auf die Luftqualität am Boden zu untersuchen und gegebenenfalls von anthropogenen Einflüssen abzugrenzen.

Dazu sollen mithilfe von ML-Methoden satellitenbasierte Datensätze für die Bodenkonzentration der Luftschadstoffe  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  generiert werden, wofür ein RF-Ansatz nach dem Vorbild von Handschuh et al. [27] zum Einsatz kommt. Unter Zuhilfenahme meteorologischer Parameter und ergänzenden Hilfsdatensätzen zum örtlichen menschlichen Einfluss und zur geographischen Beschaffenheit wird je ein RF-Modell pro Schadstoff trainiert, durch das im Anschluss flächenhafte Vorhersagen für die Bodenkonzentration getroffen werden.



---

Unter Verwendung der abgeleiteten Daten soll die Analyse der Einflussfaktoren der bodennahen Schadstoffbelastung durchgeführt werden.

Diese besteht zum einen daraus, den Einfluss einzelner meteorologischer, räumlicher und zeitlicher Eingangsvariablen auf die Modellgenauigkeit zu quantifizieren und zu erklären. Methoden der erklärbaren künstlichen Intelligenz (Explainable AI - XAI) stellen dabei ein wichtiges Hilfsmittel dar, zusammen mit der räumlichen und zeitlichen Trennung von Trainings- und Testdaten der Modelle.

Zum anderen soll die räumliche und zeitliche Variabilität der Schadstoffbelastung in Deutschland und den daran angrenzenden Gebieten analysiert werden. Besonderes Augenmerk wird dabei auf Bedingungen für das Auftreten und die Ausprägung lokaler Schadstoff-Hotspots bzw. Coldspots, sowie Trends in der Schadstoffentwicklung gelegt.

Des Weiteren sollen die beobachteten Korrelationen zu Wetterbedingungen vor einem physikalischen Hintergrund analysiert werden. Dazu werden die durch die Modelle erlernten Zusammenhänge zwischen den meteorologischen Parametern bzw. den eingesetzten Hilfsdatensätzen und dem Schadstoffaufkommen am Boden untersucht und mit physikalischen und chemischen Prozessen in der Atmosphäre verglichen.

Im folgenden Kapitel (Kapitel 2) werden zunächst die theoretischen Grundlagen zu den betrachteten Schadstoffen, sowie Kerninformationen zu Satellitenmessungen vermittelt. Im Anschluss folgt eine Erläuterung der verwendeten Datenquellen und eine Beschreibung der zur Datenaufbereitung vorgenommenen Schritte (Kapitel 3). Der Modellaufbau sowie die wichtigsten Analysemethoden werden in Kapitel 4 zusammengefasst, gefolgt von der Präsentation der Ergebnisse in Kapitel 5. Zuletzt werden diese in Hinblick auf die oben angeführten Kernfragen der Arbeit interpretiert und diskutiert (Kapitel 6).



## 2 Grundlagen

Zur Analyse der Bodenkonzentrationen von Luftschadstoffen wie Feinstaub (PM<sub>2.5</sub>), Stickoxiden (NO<sub>x</sub>) und Ozon (O<sub>3</sub>) ist ein generelles Verständnis der Entstehungsprozesse von Luftverschmutzung notwendig. Im folgenden Kapitel werden daher grundlegende Eigenschaften der Schadstoffe, sowie deren Hauptemissionsquellen und chemischen Entstehungs- bzw. Umwandlungsprozesse in der Troposphäre beschreiben.

### 2.1 Kinetik chemischer Reaktionen

Grundlegend kann die Entwicklung einer chemischen Verbindung in der Atmosphäre über die Masseerhaltung beschrieben werden. Für die Stoffmengendichte  $n_A(\mathbf{r}, t)$  einer Substanz A in einem beliebigen Luftpaket muss stets

$$\frac{dn_A}{dt} = (F_{\text{in}} - F_{\text{out}}) + (P - L) \quad (2.1)$$

gelten, wobei  $F_{\text{in}}$  und  $F_{\text{out}}$  die pro Zeiteinheit und pro Luftvolumen zuströmende und abströmende Stoffmenge beschreibt.  $P$  ist die Summe der Raten aller Stoffquellen, während  $L$  analog die Summe der Senken darstellt [46]. Ist das Windströmungsfeld  $\mathbf{U}(\mathbf{r}, t) = (u, v, w)$ , nach dem sich die Luftmassen am Ort  $\mathbf{r}$  in der Atmosphäre bewegen, bekannt, lässt sich Gleichung (2.1) wie folgt formulieren [47]:

$$\frac{dn_A}{dt} = -\nabla \cdot (n_A \mathbf{U}) + P - L. \quad (2.2)$$

Ist die Strömung  $\mathbf{U}$  divergenzfrei, so gilt für den Advektionsterm  $-\nabla \cdot (n_A \mathbf{U}) = -\mathbf{U} \nabla n_A$ . Wenn die Stoffkonzentration Windaufwärts vom Ort  $\mathbf{R}$  (Richtung  $-\mathbf{U}(\mathbf{R})$ ) ansteigt, dann ist  $-\mathbf{U}(\mathbf{R}) \nabla n_A(\mathbf{R}) > 0$  und es kommt zu einem Zustrom stoffreicherer Luft, wodurch die Stoffdichte im Luftvolumen zunimmt. Bei Windstille oder wenn das Luftvolumen die gesamte Atmosphäre einschließt findet kein Netto-Teilchentransport statt. Wird zusätzlich davon ausgegangen, dass sich das System im Gleichgewichtszustand befindet ( $\frac{dn_A}{dt} = 0$ ), müssen die Produktions- und Verlustraten der Quellen und Senken des Stoffs A gleich sein ( $P = L$ ). Die Verweilzeit eines solchen Stoffes ist dann durch

$$\tau_A = \frac{n}{P} = \frac{n}{L} \quad (2.3)$$

gegeben [46]. Bei chemischen Vorgängen, sind die Produktions- und Verlustraten eines Stoffes durch die Reaktionsrate

$$v_C = -\frac{d[A]}{dt} = k_{A+B}[A][B] \quad (2.4)$$

für die Reaktion zweiter Ordnung



gegeben, wobei  $[A]$  und  $[B]$  die Konzentrationen der jeweiligen Stoffe darstellen [48]. Die Stofflebenszeit  $\tau_A$  aus Gleichung (2.3) ist dann durch die Reaktionsgeschwindigkeitskonstante  $k$  und die Konzentration der Reaktanten gegeben:

$$\tau_A = \frac{[A]}{k_{A+B}[A][B]} = \frac{1}{k_{A+B}[B]}. \quad (2.5)$$

## 2.2 Feinstaub

Aerosole bezeichnen im Allgemeinen in der Luft schwebende flüssige oder feste Partikel welche sich abhängig von ihrer Größe und Aerodynamik über längere Zeiträume in der Atmosphäre aufhalten können ohne zu Boden zu fallen. Die chemische und physikalische Zusammensetzung der Partikel kann dabei je nach Quelle sehr unterschiedlich ausfallen. Typischen Komponenten sind z.B. Ruß, Rauch, Mineralstaub und kleinste Flüssigkeitstropfen. Bei besonders kleinen Partikelgrößen wird von Feinstaub gesprochen. Dieser wird in Abhängigkeit vom Partikeldurchmesser häufig nach den folgenden Kategorien unterteilt:

- PM<sub>10</sub>: Partikel < 10 µm
- PM<sub>2.5</sub>: Partikel < 2.5 µm
- Ultrafeinstaub (PM<sub>0.1</sub>): Partikel < 0.1 µm

Für diese Einteilung gilt meist der aerodynamische Durchmesser als Bewertungskriterium. Er beschreibt den Durchmesser eines kugelförmigen Partikels mit definierter Dichte und mit derselben Sinkgeschwindigkeit wie das betrachtete Partikel. Aufgrund ihrer geringen Partikelgröße können PM<sub>2.5</sub> und PM<sub>0.1</sub> besonders tief in den menschlichen Organismus eindringen und sind daher besonders gesundheitsgefährdend.

### Entstehung

Bei der Entstehung von Feinstaub wird zwischen primären und sekundären Schadstoffquellen unterschieden. Primäre Quellen geben Feinstaubpartikel direkt an die Umgebung ab, wie beispielsweise bei Verbrennungsprozessen. Der Großteil des auf diese Weise freigesetzten PM<sub>2.5</sub> ist auf anthropogene Aktivitäten zurückzuführen. In Deutschland waren im Jahr 2022 Haushalte und Kleinverbraucher, in deren Heizungsanlagen Feinstaub freigesetzt wird, mit einem Anteil von 28% an den Gesamtemissionen die größte PM<sub>2.5</sub>-Quelle. Der Straßenverkehr war mit 25% zweitgrößter Verursacher. Hier entstehen Feinstäube zum Beispiel bei der Verbrennung von Dieselmotoren oder als Abrieb von Reifen und Bremsen. Zudem können bereits abgesetzte Staubpartikel durch Fahrzeuge und Wind aufgewirbelt werden und so erneut zur Feinstaubbelastung beitragen. Weitere Quellen beinhalten Direktmissionen aus Industrieprozessen (25%) wobei hier auch Vorläufersubstanzen für sekundäre Aerosole emittiert werden [1, 49].

Auf natürliche Weise entstehen primäre Aerosole z.B. in Form von Salzaerosolen in Küstenregionen durch Gischt [50] oder als biologisches organisches Material (Pollen, Sporen, etc.) durch Vegetation. Bodenerosion in trockenen Gebieten oder Katastrophen wie Waldbrände und Vulkanausbrüche können zudem große Mengen Mineralstaub bzw. Asche und Staub in die Atmosphäre ausstoßen, welche unter geeigneten Bedingungen über große Distanzen transportiert werden können (wie z.B. Sahara-Staub nach Europa).

Je nach Gebiet und Art der Emissionen, kann ein Großteil des Feinstaubs aus sekundären Aerosolen bestehen, die durch chemische Prozesse in der Atmosphäre aus reaktiven Vorläufersubstanzen gebildet werden [51, 52, 53]. Diese Vorläuferstoffe schließen unter anderem

flüchtige organische Verbindungen (VOC), Stickdioxid ( $\text{NO}_2$ ), Schwefeldioxid ( $\text{SO}_2$ ) und Ammoniak ( $\text{NH}_3$ ) ein, die durch chemische Prozesse in der Atmosphäre in Feststoffe umgewandelt werden können. Die Reaktionsprodukte lagern sich leicht an Kondensationskernen in der Atmosphäre an und verursachen so die Bildung von weiterem  $\text{PM}_{2.5}$  [54].

Durch die verzögerte Reaktion können die Vorläufersubstanzen davor über große Entfernungen transportiert werden und somit auch örtlich versetzt zur Luftverschmutzung beitragen [55].

Als Beispiel für sekundäre Aerosolbildung kann die Reaktion von Ammoniak mit Stickoxiden beschrieben werden. Ammoniak wird dabei in Industriellen Prozessen, oder durch Einsatz von Düngemitteln in der Landwirtschaft von Böden freigesetzt [55]. In einem ersten Schritt reagiert Stickstoffdioxid mit Hydroxylradikalen ( $\text{OH}$ ) wie in Reaktionsgleichung (R 2.2) zu Salpetersäure

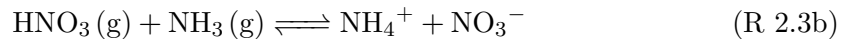


wobei M ein für die Reaktion benötigter Stoßpartner ist.

Abhängig von der relativen Luftfeuchtigkeit kann anschließend wie in (R 2.3a) bzw. (R 2.3b) Ammoniumnitrat ( $\text{NH}_4\text{NO}_3$ ) als sekundäres anorganisches Aerosol durch die Weiterreaktion mit Ammoniak gebildet werden [46, 56]:



für trockene Bedingungen in denen Ammoniumnitrat in fester Form auftritt, und

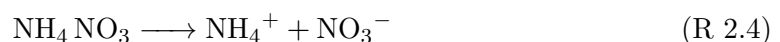


bei Überschreiten der zur Deliqueszenz<sup>1</sup> nötigen relativen Feuchtigkeit mit Ammoniumnitrat als Lösung. Das Gleichgewicht der Reaktionen ist für hohe Feuchtigkeitwerte stark zur Aerosolphase verschoben [46, 57].

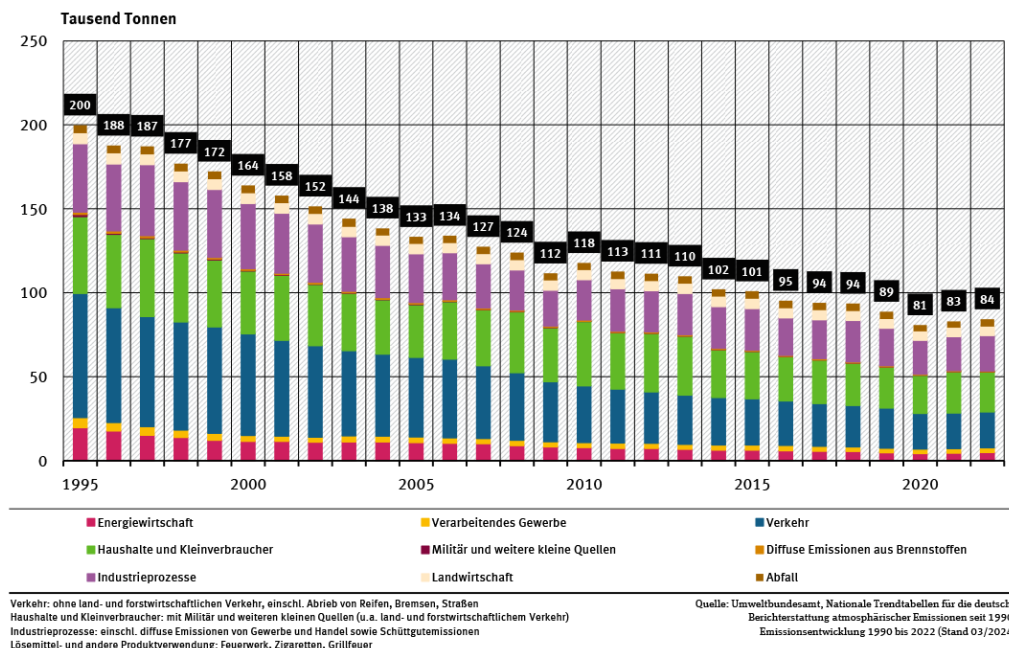
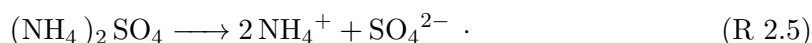
Ähnlich verläuft auch die Bildung von Ammoniumsulfat ( $(\text{NH}_4)_2\text{SO}_4$ ) aus  $\text{SO}_2$  und  $\text{NH}_3$ , welches gemeinsam mit Ammoniumnitrat einen großen Teil der Masse des Sekundärstaubs (40%) ausmacht [55]. Zudem relevant sind sekundäre Aerosole, die aus flüchtigen organischen Substanzen (VOC) entstehen. Quellen von VOCs werden in Kapitel 2.4 näher betrachtet. Zusätzlich zu den anthropogenen Quellen sind Vorläufergase auch in der Natur, z.B. in vulkanischen Gasen ( $\text{SO}_2$ ) oder Bränden ( $\text{NO}_x$ ) zu finden.

## Abbaumechanismen und Umwelteinflüsse

Neben den genannten Quellen und Prozessen hängt die Feinstaubkonzentration auch von verschiedenen Umwelteinflüssen ab. So kann Niederschlag das Auswaschen der Grenzschicht und somit die Deposition von PM aus der Atmosphäre erwirken. Nasse Oberflächen unterbinden zudem, besonders in städtischen Regionen, das (Wieder-)Aufwirbeln von Staub durch Verkehr, Wind und Materialumschlag. Daher folgt auf Niederschläge häufig eine zeitbeschränkte Verbesserung der Luftqualität in Bezug auf Feinstaub [58]. Bei Aerosolen mit hoher Konzentration von Ammoniumnitrat oder Ammoniumsulfat können sich diese Bestandteile gut im Regenwasser auflösen und durch die Bildung von Nitrat- und Sulfationen so zur Säuerung von Gewässern beitragen:



<sup>1</sup>Deliqueszenz beschreibt die Aufnahme von Feuchte aus der Luft durch eine feste Substanz, wodurch diese eine Lösung bildet.



**Abbildung 2.1:** Entwicklung der Menge des menschenproduzierten PM<sub>2.5</sub> und der Anteile der einzelnen Feinstaubquellen in Deutschland von 1995 bis 2022. Ein andauernder Abfall der gesamten Feinstaubemissionen und speziell der Emissionen durch Verkehr und Haushalte ist erkennbar. (Abbildung entnommen aus [1])

## Entwicklung

In den vergangenen Jahrzehnten sind die anthropogenen Feinstaubemissionen in Deutschland und der EU deutlich zurückgegangen (siehe Abb. 2.1). Dabei liegt in fast allen Emissionskategorien eine Verringerung des PM<sub>2.5</sub>-Ausstoßes vor. Rückgänge im Verkehrssektor sind zum einen auf technische Maßnahmen wie das Verbauen von Dieselpartikelfiltern und generelle Verbesserung der Abgaseffizienz von Verbrennungsmotoren durch schärfere Abgasgesetze zurückzuführen. Zum anderen bewirken Einschränkungen des Verkehrsaufkommens in Städten wie z.B. die Anpassung von Geschwindigkeitsbegrenzungen oder Einführung von Umweltzonen eine Verringerung der PM<sub>2.5</sub>-Emissionen.

Dem entgegen stehen Prognosen zur möglichen globalen Zunahme von natürlichen PM-Quellen im Zuge des Klimawandels. Dazu gehören erhöhte Staubemissionen durch die Ausbreitung von Wüsten, sowie ein durch Trockenheit bedingtes vermehrtes Auftreten von großflächigen Wald- und Buschbränden und Erosion. Häufigere Hitze- und Dürreereignisse beeinflussen zudem die Menge und Zusammensetzung von durch Pflanzen emittierten biogenen VOCs, welche zur Bildung sekundärer Aerosole beitragen können [49][59].

## 2.3 Stickoxide

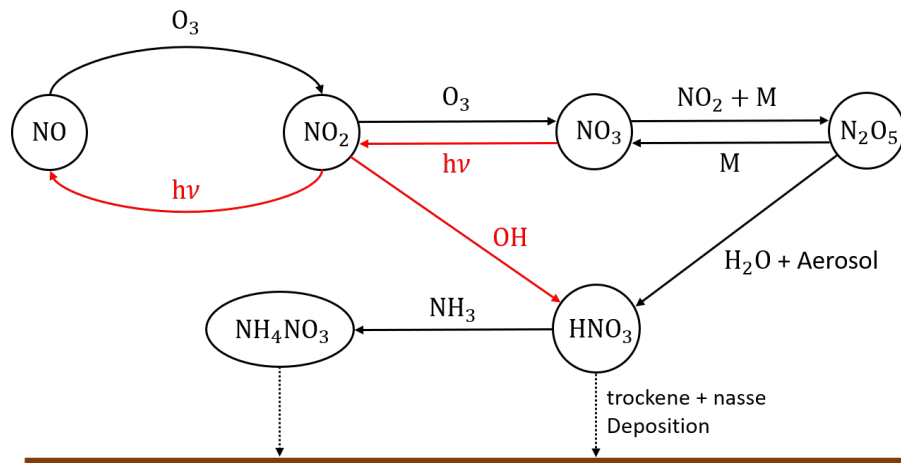
Stickstoffdioxid (NO<sub>2</sub>) und Stickstoffmonoxid (NO) stellen die für die Luftverschmutzung relevantesten nitrose Gase dar und werden unter der Bezeichnung Stickoxide (NO<sub>x</sub> = NO + NO<sub>2</sub>) zusammengefasst. Beide Stoffe werden häufig gemeinsam emittiert und durch chemische Prozesse in der Atmosphäre schnell ineinander umgewandelt.

## Entstehung

Die primären Quellen von  $\text{NO}_x$  sind hauptsächlich menschengemachte Verbrennungsprozesse, wobei unterschiedliche Bildungsmechanismen von Bedeutung sind. Zur thermischen Bildung von  $\text{NO}_x$  kommt es, wenn die Verbrennungsluft besonders hohe Temperaturen ( $\geq 1200^\circ\text{C}$ ) erreicht. Dabei werden molekularer Stickstoff und Sauerstoff aus der Luft in atomare Radikale gespalten und bilden so hauptsächlich  $\text{NO}$ , welches nach Abkühlung des Gases weiter zu  $\text{NO}_2$  oxidiert. Dazu kommt die Umsetzung der im Brennstoff gebundenen Stickstoffanteile zu  $\text{NO}_x$ . Diese ist stark von der Menge des gebundenen Stickstoffs und somit vom Brennstoff selbst abhängig [60].

Eine Quelle von Stickoxiden ist der Betrieb von Verbrennungsmotoren im Straßenverkehr. Mit 36% am Gesamtaufkommen leisteten die Verkehrsemissionen im Jahr 2022 den größten Beitrag zur  $\text{NO}_x$ -Luftverschmutzung, gefolgt von Abgasen aus Kraftwerken zur Energieerzeugung (26%) und aus Haushalten und Kleinverbrauchern (13%). Zudem kann in landwirtschaftlichen Böden  $\text{NO}_x$  durch Denitrifizierung gebildet und emittiert werden, wobei Nitrat durch Microorganismen in Stickstoff und  $\text{NO}$  umgewandelt wird [61].

In vergleichsweise geringen Mengen wird Stickstoffdioxid außerdem natürlich gebildet, z.B. durch die hohen Temperaturen bei Blitzeinschlägen oder bei Waldbränden.



**Abbildung 2.2:** Reaktionskette zum Abbau von  $\text{NO}_x$  in der Troposphäre. Rot markierte Übergänge finden nur tagsüber statt.

## Abbaumechanismen und Umwelteinflüsse

Grundlegend für das Verständnis der Chemie innerhalb der Troposphäre ist die Bildung von Hydroxyl-Radikalen ( $\text{OH}$ ).  $\text{OH}$  stellt, für den Abbau schlecht wasserlöslicher Stoffe, das wichtigste atmosphärische Oxidationsmittel dar [62]. Für seine Bildung wird ein energetisch angeregtes  $\text{O}(^1\text{D})$  Sauerstoffradikal benötigt, wie es beispielsweise bei der Spaltung von Ozon durch kurzwellige Strahlung entsteht (siehe (R 2.16b)) bevor es durch Stoß mit einem weiteren Atom in den Grundzustand zurückfallen kann. Bei der Kollision von  $\text{O}(^1\text{D})$  mit einem Wassermolekül entstehen zwei  $\text{OH}$ -Radikale [46]:



Da die Verfügbarkeit von  $\text{OH}$ , sowie das Ozon- $\text{NO}_x$ -Gleichgewicht von photochemischen Prozessen abhängen, müssen die Reaktionsmechanismen bei Tag und Nacht getrennt betrachtet werden. In beiden Fällen stellt Salpetersäure ( $\text{HNO}_3$ ) das Endprodukt der

Reaktionskette dar. Diese kann dann entweder durch trockene oder nasse Deposition (z.B. durch Lösung in Regenwasser) aus der Luft entfernt werden, oder wie in Kapitel 2.2 beschrieben, sekundäre Aerosole bilden.

Bei Tag befindet sich Stickstoffdioxid im Gleichgewicht mit  $O_3$  und  $NO$ . Der bedeutendste Abbaumechanismus ist hier die Reaktion mit  $OH$  (siehe hierzu Abb. 2.2):



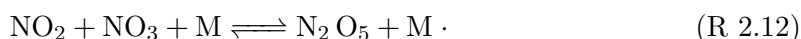
Bei Nacht werden weitere Stickstoff-Sauerstoff-Verbindungen relevant. So reagiert  $NO_2$  schnell mit Ozon unter Bildung von Nitrat-Radikalen ( $NO_3$ ) und molekularem Sauerstoff ( $O_2$ ):



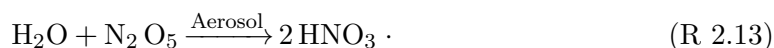
Unter Tageslicht besitzen diese Radikale nur eine sehr geringe Lebensdauer, da sie schnell durch Photolyse in  $NO$  und  $NO_2$  zurückgeführt werden, oder mit  $NO$  reagieren [46]:



Durch die geringe Verfügbarkeit von  $NO$  bei Nacht wird die Bildung von  $NO_3$  in relevanten Mengen ermöglicht, sodass diese wiederum mit  $NO_2$  reagieren um Distickstoffpentoxid ( $N_2O_5$ ) zu bilden [46]:



Durch die Feuchtigkeit in der Luft kann anschließend durch heterogene Hydrolyse von  $N_2O_5$  Salpetersäure entstehen [63]:



Der gesamte Reaktionsweg bei Nacht ist in Abb. 2.2 durch die schwarzen Pfeile gekennzeichnet.

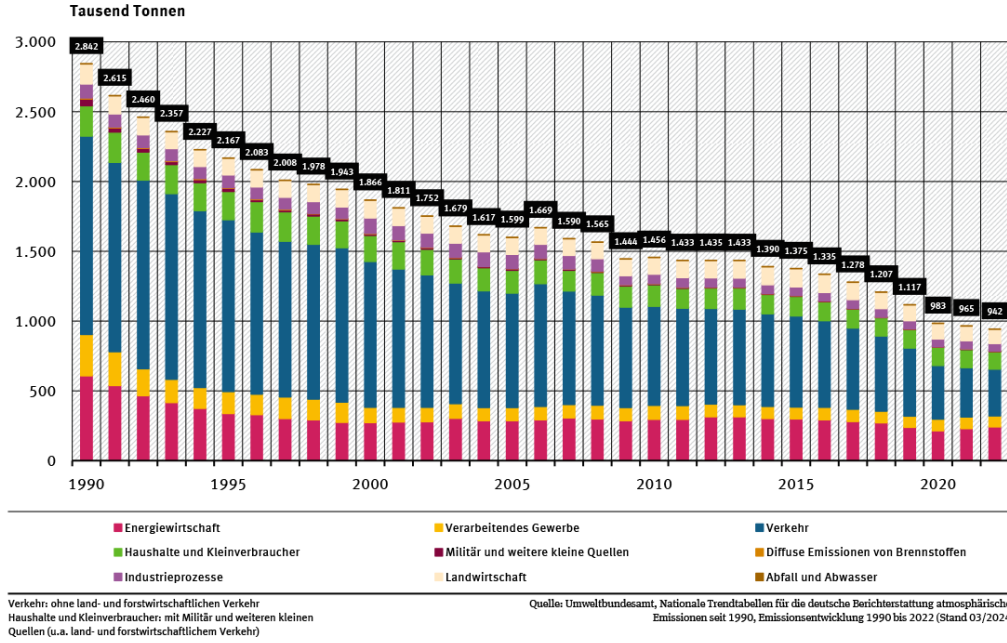
## Entwicklung

Die absolute Menge an  $NO_2$ -Emissionen zeigte in Deutschland über die letzten Jahrzehnte einen rücklaufenden Trend, wie in Abb. 2.3 erkenntlich ist. Die Umsetzung effektiver Maßnahmen zur Emissionsminderung haben hierbei vor allem zu einer Verringerung der Verkehrsemissionen und dem  $NO_x$ -Ausstoß durch Kraftwerke geführt. Primäre Maßnahmen die den Feuerungsprozess in Kraftwerken betreffen, z.B. Abgasrückführung oder Verringerung der Temperatur an geeigneten Stellen des Verbrennungsprozesses, können die Entstehung von Stickoxiden mindern. Sekundärmaßnahmen richten sich auf die Nachbehandlung des Abgases, unter anderem in Dieselfahrzeugen. Hier wird z.B. selektive katalytische Reduktion eingesetzt, in der durch Zugabe von Ammoniak und mithilfe eines Katalysators die Stickoxide in elementaren Stickstoff und Wasser umgesetzt werden [65].

## 2.4 Ozon

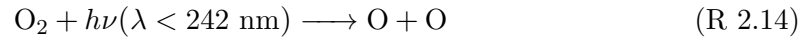
Ozon ist ein aus drei Sauerstoffatomen zusammengesetztes Molekül und wichtiger Bestandteil der Erdatmosphäre. Für seine Entstehung sind Sauerstoffradikale notwendig, welche



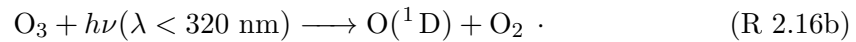
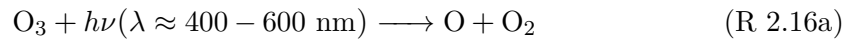


**Abbildung 2.3:** Entwicklung der NO<sub>x</sub> Emissionen und der Anteile der einzelnen Quellkategorien in Deutschland von 1995 bis 2022. (Abbildung entnommen aus [64])

hauptsächlich durch photochemische Reaktionen gebildet werden. In der Stratosphäre wird Ozon unter Einstrahlung von UV-Licht ständig gemäß den Reaktionen (R 2.14) und (R 2.15) gebildet.



Hierbei wird zunächst molekularer Sauerstoff (O<sub>2</sub>) durch hochenergetische UV-C Strahlung in atomaren Sauerstoff (O) gespalten (siehe (R 2.14)). Die freien Sauerstoffradikale können anschließend je mit einem Sauerstoffmolekül weiterreagieren und ein Ozonmolekül bilden, wobei zusätzlich noch ein Stoßparameter (z.B. N<sub>2</sub> oder O<sub>2</sub>) zum Energie- und Impulsübertrag benötigt wird (vgl. (R 2.15)). Ozonmoleküle zeigen besonders starke Absorptionen für Strahlung zwischen 240 nm und 320 nm (UV-B), wobei gemäß (R 2.16b) ein potenziell angeregtes O(<sup>1</sup>D) Sauerstoffradikal abgespalten wird [46].



Dies kann wiederum durch Stöße mit z.B. N<sub>2</sub> oder O<sub>2</sub> zurück in den Grundzustand übergehen:



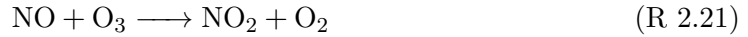
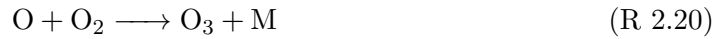
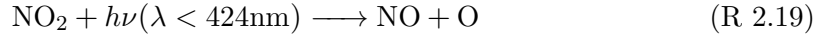
Der molekulare Sauerstoff kann anschließend über (R 2.15) zurück zu Ozon kombiniert werden. Zuletzt ist es möglich, dass O und O<sub>3</sub> unter Bildung von zwei O<sub>2</sub> Molekülen rekombinieren und so Ozon abgebaut wird:



Die Mechanismen für Produktion und Abbau von Ozon aus Gleichungen (R 2.14-2.18) werden als Chapman-Zyklus bezeichnet.

Mittels dieser Reaktionen kommt es in der Stratosphäre zu einer Ansammlung von Ozon und zur Bildung der Ozonschicht, welche die untere Atmosphäre vor hochenergetischer Strahlung schützt. Aufgrund der verringerten Strahlung laufen die oben genannten photochemischen Reaktionen in der Troposphäre in wesentlich kleineren Raten ab. Ozon besitzt zudem keine für die Außenluftverschmutzung relevanten primären Emissionsquellen, kann jedoch in Bodennähe durch den bereits in Kapitel 2.3 erwähnten  $\text{NO}_x$ -Ozon-Bildungszyklus durch Einwirkung weiterer Luftschadstoffe und Spurengase entstehen.

Dieser kann unter Vernachlässigung weiterer Substanzen, die in das Gleichgewicht eingreifen, wie folgt beschrieben werden [66]:



$\text{NO}_2$  wird bei Sonneneinstrahlung gemäß (R 2.19) durch Photolyse mit Wellenlängen  $\lambda < 424 \text{ nm}$  in Stickstoffmonoxid ( $\text{NO}$ ) überführt, wobei atomarer Sauerstoff freigesetzt wird. Durch diesen entsteht wiederum zusammen mit dem Luftsauerstoff Ozon (siehe (R 2.10) und (R.2.15)). Anschließend kann das gebildete  $\text{O}_3$  mit  $\text{NO}$  reagieren um erneut  $\text{NO}_2$  zu bilden, wodurch sich die Substanzen in einem Gleichgewicht befinden.

Die Reaktionsgeschwindigkeit von (R 2.21) ist durch  $k_{\text{NO}+\text{O}_3}$  gegeben, während die Rückreaktionen (R 2.19) und (R 2.20) mit dem photolytischen Ratenkoeffizient  $j_{\text{NO}_2}$  ablaufen, der die Reaktionsrate bestimmt. Nach Gleichung (2.2) und (2.4) gilt, falls sich das System aus den durch die Reaktionsgleichung (R 2.22) zusammengefassten Reaktionen im Gleichgewicht befindet:

$$0 = k_{\text{NO}+\text{O}_3}[\text{NO}][\text{O}_3] - j_{\text{NO}_2}[\text{NO}_2]. \quad (2.6)$$

Die Gleichgewichtskonzentrationen von Ozon,  $\text{NO}$  und  $\text{NO}_2$  sind dann

$$[\text{O}_3] = \frac{j_{\text{NO}_2}[\text{NO}_2]}{k_{\text{NO}+\text{O}_3}[\text{NO}]} \quad (2.7a)$$

$$[\text{NO}] = \frac{j_{\text{NO}_2}[\text{NO}_2]}{k_{\text{NO}+\text{O}_3}[\text{O}_3]} \quad (2.7b)$$

und

$$[\text{NO}_2] = \frac{k_{\text{NO}+\text{O}_3}[\text{NO}][\text{O}_3]}{j_{\text{NO}_2}}. \quad (2.7c)$$

Die Gleichgewichtskonstante  $K$  beschreibt das Verhältnis der Substanzen auf der linken und rechten Seite der Reaktionspfeile in (R 2.22).

$$K = \frac{[\text{NO}][\text{O}_3]}{[\text{NO}_2]} = \frac{j_{\text{NO}_2}}{k_{\text{NO}+\text{O}_3}} \quad (2.8)$$

wobei die Gleichungen (2.7a), (2.7b) und (2.7c) eingesetzt wurden. Die Reaktionsgeschwindigkeit  $k_{\text{NO}+\text{O}_3}$  beträgt etwa  $1.9 \times 10^{-14} \frac{\text{cm}^3}{\text{molecules}}$  [67, 68, 69, 70]. Für den photolytischen Ratenkoeffizient wird

$$j_{\text{NO}_2} = 0.0135 \times \exp\left(-0.360 \frac{1}{\cos(\chi)}\right) \frac{1}{\text{s}} \approx 8.91 \times 10^{-3} \frac{1}{\text{s}}$$

[71] angenommen, wobei ein Zenitwinkel von  $\chi = 30^\circ$  gewählt wurde. Mit diesen Werten ergibt sich für die Gleichgewichtskonstante in Gl. (2.8)

$$K = \frac{8.91 \times 10^{-3} \frac{1}{\text{s}}}{1.9 \times 10^{-14} \frac{\text{cm}^3}{\text{molecules}}} \approx 4.69 \frac{\text{molecules}}{\text{cm}^3} = 16.3 \text{ ppb} \quad (2.9)$$

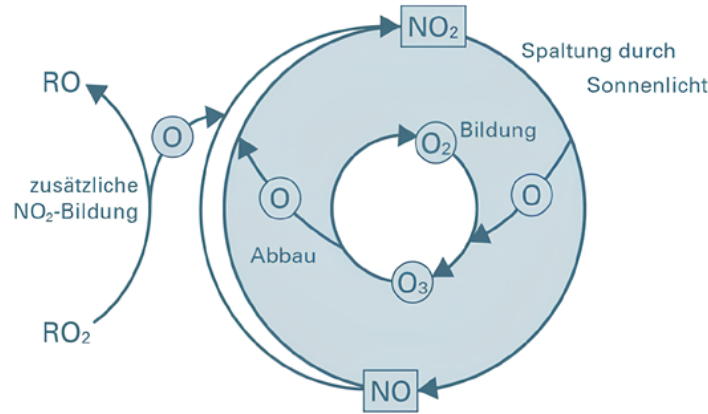
bei einem Luftdruck von  $p = 1013.25 \text{ hPa}$  und einer Temperatur  $T = 298 \text{ K}$ . Um die Lebenszeit von Ozon  $\tau_{\text{O}_3}$  im Gleichgewicht mit NO und NO<sub>2</sub> in der Troposphäre abzuschätzen wird Gl. (2.5) herangezogen. Ein Beispielwert für die Stickstoffmonoxidkonzentration liefern Aneja et al. [72] (Mittelwert von NO August 1991 in Raleigh, North Carolina) mit  $[\text{NO}]_{\text{ppb}} = 6.1 \text{ ppb}$  [72]. Es ergibt sich eine Lebensdauer von

$$\begin{aligned} \tau_{\text{O}_3} &= \frac{1}{k_{\text{NO}+\text{O}_3}[\text{NO}]} \\ &= \frac{1}{k_{\text{NO}+\text{O}_3}[\text{NO}]_{\text{ppb}} \frac{p}{k_{\text{B}}T}} \\ &= \frac{298 \text{ K} \times (1.38 \times 10^{-23} \frac{\text{J}}{\text{K}})}{(6.1 \text{ molecules} \times 10^{-9}) \times (1.9 \times 10^{-14} \frac{\text{cm}^3}{\text{molecules}}) \times (1013.25 \text{ hPa})} \\ &\approx 5.8 \text{ min} \end{aligned} \quad (2.10)$$

Die Lebensdauer eines Stickstoffdioxidmoleküls beträgt das Inverse des photolytischen Ratenkoeffizienten:

$$\tau_{\text{NO}_2} = \frac{1}{j_{\text{NO}_2}} \approx 1.87 \text{ min.} \quad (2.11)$$

In der Realität wird der beschriebene NO<sub>x</sub>-Ozon-Zyklus durch weitere Stoffe gestört,



**Abbildung 2.4:** : Reaktionskette des Ozon-NO<sub>x</sub>-Zyklus mit Abbau von NO durch zusätzliche VOC-Radikale (Abbildung entnommen aus [73])

wie in Abb. 2.4 dargestellt ist. Neben den in Kapitel 2.3 genannten Wegen zum NO<sub>x</sub>-Abbau, wirkt sich die Reaktion von NO<sub>x</sub> mit Flüchtigen organischen Verbindungen (VOC) besonders stark auf das Ozon-NO<sub>x</sub>-Gleichgewicht aus. VOCs sind Kohlenwasserstoffe (RH) und gelangen sowohl durch menschliches Zutun, als auch auf natürlichen Wegen in die Atmosphäre. In urbanen Gebieten sind vor allem der Einsatz von Lösungsmitteln (z.B. in Farben und Lacken), sowie das Verdunsten oder die unvollständige Verbrennung von Treibstoff als VOC-Quelle entscheidend. Der Großteil der VOCs in der Erdatmosphäre sind jedoch von Tieren und Pflanzen emittierte, sogenannte biogene VOCs (BVOC). Pflanzen

emittieren die größte Menge BVOCs, wobei die relevanteste Substanz Isopren darstellt. Die Menge der Emissionen wird von Temperatur und Sonneneinstrahlung beeinflusst, da diese die Verflüchtigung begünstigen.

Der Reaktionsmechanismus von VOCs bis zur Bildung von Ozon ist im Folgenden vereinfacht für RH beschrieben [74]:



Die gesamte Bildungsrate von Ozon ist durch die Verfügbarkeit von Hydroxyl-Radikalen (OH) limitiert und somit proportional zur Reaktionsrate von (R 2.23). Weitere Quellen und Senken für Radikale sind u.a. die Reaktionen (R 2.6) bzw. (R 2.7). Der VOC-Reaktionsmechanismus konkurriert dabei mit der NO<sub>2</sub>-Abbaureaktion (R 2.7) um die zur Verfügung stehenden Hydroxyl-Radikale.

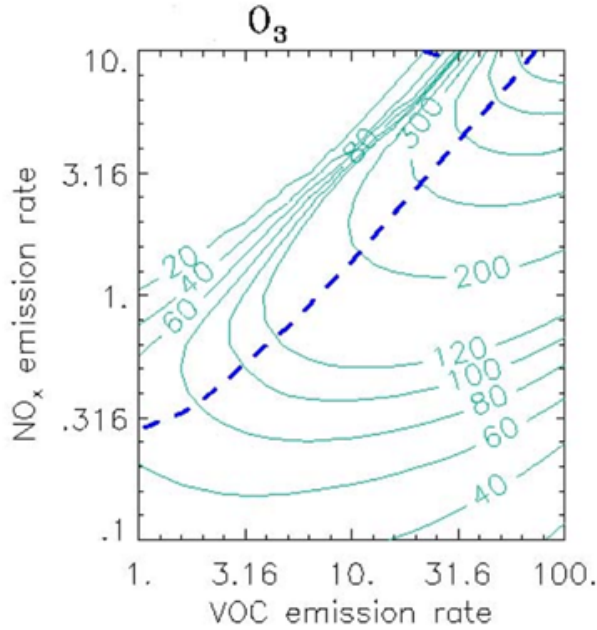
Durch (R 2.24) wird NO<sub>2</sub> aus NO gebildet, ohne den in (R 2.21) nötigen Abbau von Ozon. Durch das Brechen des NO<sub>x</sub>-O<sub>3</sub>-Zyklus kann die Ozonkonzentration [O<sub>3</sub>] weiter ansteigen (vgl. Abb. 2.4). Ist das Verhältnis der Konzentrationen [VOC]/[NO<sub>x</sub>] (korrigiert für relative Reaktivität der Substanzen mit OH) hoch, so erhöht sich auch die pro NO<sub>x</sub> gebildete Ozonmenge, bei konstanter Verfügbarkeit von OH. Die absolute Produktionsrate ist jedoch wegen (R 2.26) noch immer abhängig von der Stickoxidkonzentration [NO<sub>x</sub>] (vgl. Gl. (2.7a)), sodass das System in diesem Zustand sehr sensibel auf NO<sub>x</sub> reagiert [46]. Für [VOC]/[NO<sub>x</sub>] ≪ 1 reagiert OH verstärkt mit NO<sub>2</sub> (R 2.7), weshalb durch den schnellen Abbau von NO<sub>x</sub> und das intakte Gleichgewicht (R 2.22) nur geringe Mengen an Ozon entstehen.

Es existieren weitere Mechanismen, die die Bedeutung der Konzentrationsverhältnisse [VOC]/[NO<sub>x</sub>] für die Ozonproduktion bestärken, hier jedoch nicht berücksichtigt wurden. Die Abhängigkeit der Ozonproduktion von VOC- und NO<sub>x</sub>-Konzentrationen kann in einem Ozon-Isoplethen-Diagramm wie in Abb. 2.5 dargestellt werden. Für eine feste Menge VOCs existiert eine NO<sub>x</sub>-Konzentration, für welche das meiste Ozon produziert wird (blaue Linie in Abb. 2.5). Unterhalb dieses optimalen [VOC]/[NO<sub>x</sub>] Verhältnisses (links der blauen Linie in Abb. 2.5) führt eine Erhöhung der NO<sub>x</sub>-Konzentration zu einer Ozonverringerung und das System ist NO<sub>x</sub>-gesättigt. Für höhere Konzentrationen von VOCs (rechts der blauen Linie) wird durch mehr NO<sub>x</sub> auch mehr O<sub>3</sub> gebildet. Die Beziehung zwischen [O<sub>3</sub>] und NO<sub>2</sub>-Emissionen ist somit direkt von der Konzentration der VOCs in der Troposphäre abhängig [75].

## Entwicklung

Mithilfe der oben genannten Reaktionsmechanismen lässt sich die räumliche und zeitliche Verteilung von Ozon erklären. In den letzten Jahrzehnten konnte, wie in Abb. 2.6 zu sehen ist, generell ein leichter Anstieg der Ozonkonzentration in Deutschland gemessen werden. Zudem waren die mittleren Ozonwerte in ländlichen Regionen systematisch höher als in den Städten.

Die räumlich ungleiche Verteilung ist dabei wahrscheinlich auf den Titrationseffekt zurückzuführen welcher den Ozonabbau durch Stickstoffmonoxid gemäß (R 2.21) beschreibt. In der Nähe starker NO-Quellen, z.B. in verkehrsdichten Bereichen oder in der



**Abbildung 2.5:** Isoplethen der maximalen Ozonkonzentration (ppb) als Funktion der  $\text{NO}_x$  und VOC Emissionsrate ( $10^{12} \text{ molecules cm}^{-2} \text{ s}^{-1}$ ) aus theoretischen 0-d Berechnungen. Die blaue Linie trennt die VOC-sensitiven von den  $\text{NO}_2$ -sensitiven Bereichen. (Abb. entnommen aus [75])

unmittelbaren Nähe von starken Punktquellen (z.B. Kraftwerke) kann dies vor allem bei geringer Sonneneinstrahlung (im Winter oder nachts) zu einer Netto-Umwandlung von Ozon in  $\text{NO}_2$  führen. Dadurch kann die Ozonbildung vorerst unterdrückt werden. Weiter windabwärts, kann die Bildung durch transportierte Vorläufersubstanzen verzögert stattfinden, wodurch die Ozonwerte zunehmen [75]. Dies ist auch an Gleichung (2.7a) zu erkennen, die auf eine inverse Proportionalität zwischen  $\text{O}_3$ - und  $\text{NO}$ -Konzentration hindeutet. Die tendenzielle Abnahme von  $\text{NO}_x$ -Emissionen (vgl. Abb. 2.3) sorgt daher für eine leichte Schwächung des Titrationseffekts, was einen Beitrag zum Anstieg der  $\text{O}_3$ -Jahresmittelwerte in Abb. 2.6 leistet. Dieser ist jedoch auch für Sommermonate erkenntlich, weswegen noch weitere mögliche Ursachen in Betracht gezogen werden müssen [76]. Der wahrscheinlichste Verursacher ist die überregionale Zunahme des troposphärischen Ozons. Grund hierfür könnte die globale Zunahme der Produktion von Ozon-Vorläufersubstanzen, insbesondere in Asien, und der anschließende Transport nach Mitteleuropa sein [76, 4]. Die Menge der nicht-Methan VOC-Emissionen (NMVOCs) in Deutschland ist dagegen seit 1990 um bis zu 73% gesunken [77].

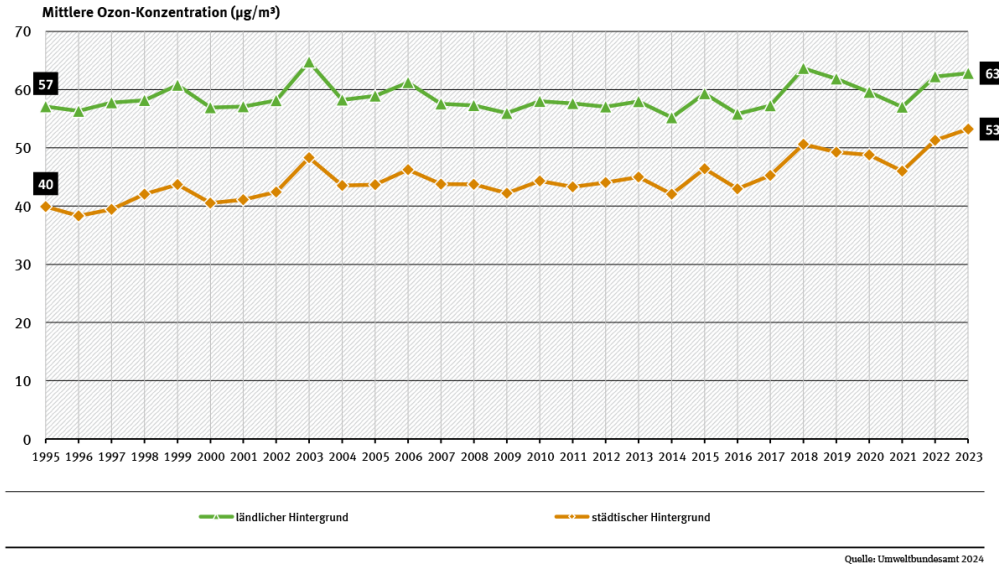
## 2.5 Satellitenmessungen

### 2.5.1 Aerosol Optische Dicke

Die Aerosol Optische Dicke (AOD) ist ein Maß dafür, wie stark das Sonnenlicht beim Durchlaufen einer Luftsäule in der Atmosphäre durch Aerosole gestreut und absorbiert wird. Mathematisch ist diese Abschwächung durch das Lambert-Beer'sche Gesetz beschrieben, nach welchem die Strahlungsintensität beim Durchlaufen eines optisch dicken Mediums exponentiell abnimmt. Eine anfängliche Intensität von  $I_0(\lambda)$  wird während des Strahlungstransports durch Wechselwirkung mit den Partikeln in der Atmosphäre auf

$$I(\lambda) = I_0(\lambda)e^{-\tau(\lambda)} \quad (2.12)$$

abgeschwächt, wobei  $\tau(\lambda)$  die wellenlängenabhängige AOD der vertikalen Luftsäule bezeichnet [78].



**Abbildung 2.6:** Gemessene Jahresmittelwerte des ländlichen und städtischen Hintergrunds von Ozon in Deutschland von 1995 bis 2023. (Abb. entnommen aus [4])

Die durch die Strahlungstransportgleichung beschriebene Übertragung von Licht berücksichtigt im Allgemeinen Streuung, Absorption und Emission. Für eine vereinfachte Betrachtung kann der Effekt der Emission vernachlässigt werden und die AOD in einer Luftsäule als Integral vom Boden ( $z_0$ ) bis zum oberen Rand der Atmosphäre ( $z_1$ ) entlang der Säule wie folgt beschrieben werden:

$$\tau = \int_{z_0}^{z_1} \beta_E(z) dz. \quad (2.13)$$

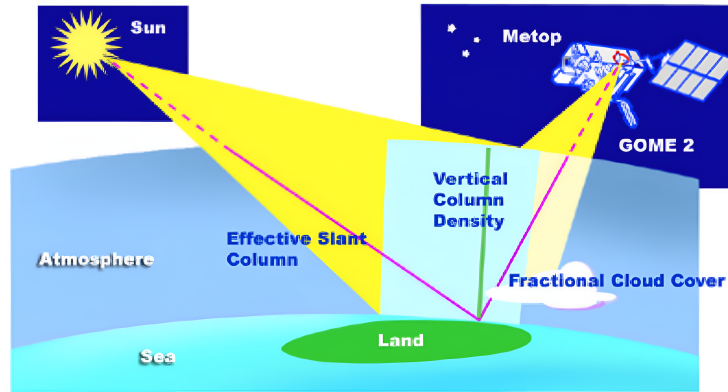
Dabei ist  $\beta_E(z)$  der aus Streuungs- und Absorptionskoeffizient zusammengesetzte Extinktionskoeffizient und beschreibt, wie stark die Aerosole in einer bestimmten Höhe  $z$  das Licht schwächen.

In dieser Arbeit kommen flächenbezogene, satellitenbasierte AOD-Daten zum Einsatz welche aus Messungen der MODIS-Instrumente (Moderate Resolute Imaging Spectroradiometer) auf den NASA-Satelliten Terra und Aqua hervorgehen. Die Spektorradiometer der Satelliten messen die Intensität des von der Erdoberfläche und der Atmosphäre reflektierten Sonnenlichts in verschiedenen Spektralbereichen. Zur Bestimmung der AOD muss die Reflexion der Erdoberfläche bekannt sein. Hierzu kommen Algorithmen zum Einsatz, welche das Reflexionsvermögen verschiedener Landoberflächen (z.B. Wüste oder Vegetation) berücksichtigen. Nun können die gemessenen Reflexionsdaten mit den erwarteten Strahlungsintensitäten von der Oberfläche verglichen werden um die AOD zu erhalten. Dazu wird die AOD im Strahlungstransportmodell solange angepasst, bis das simulierte Absorptionsverhalten dem gemessenen entspricht. Da sich die Satellitenmessung auf das vom Boden reflektierte Sonnenlicht stützt, ist sie anfällig für starke Schwankungen in der Albedo vor allem auf hellen Oberflächen wie z.B. bei Bedeckung durch Schnee oder in stark bebauten Regionen. Zudem wird das Retrieval durch Wolken gestört, weshalb diese identifiziert und maskiert werden müssen [79].

Die Zusammenhänge zwischen Satelliten-AOD und der  $PM_{2.5}$ -Konzentration am Boden sind in der Regel sehr komplex und die Umrechnung zwischen den Größen ist mit verschiedenen Schwierigkeiten verbunden. Zum einen bezieht sich die AOD auf die in der gesamten Luftsäule befindliche Aerosolmenge, wobei deren vertikale Verteilung nicht bekannt ist. Dies



macht es nur schwer mögliche Staub in hohen Luftschichten von bodennahem Feinstaub zu trennen. Zum anderen umfasst  $\text{PM}_{2.5}$  Aerosoltypen mit unterschiedlichen Zusammensetzungen und Partikelgrößenverteilungen, deren optische Eigenschaften verschieden ausfallen können. Das Streu- und Absorptionsvermögen, welches die gemessene AOD bestimmt ist daher nicht mit der Masse der Aerosole gleichzusetzen. Meteorologische Bedingungen nehmen dabei sowohl Einfluss auf die vertikale Staubverteilung, als auch auf die Beschaffenheit der Aerosole, weshalb diese bei der Gewinnung von  $\text{PM}_{2.5}$ -Bodenkonzentrationen aus AOD-Satellitendaten berücksichtigt werden müssen.



**Abbildung 2.7:** Veranschaulichung der durch Absorption gemessenen schrägen Säulendichte und der errechneten vertikalen Säulendichte während einer Satellitenmessung. (Abb. entnommen aus [80])

### 2.5.2 Troposphärisches $\text{NO}_2$

Zur Messung von  $\text{NO}_2$ -Konzentrationen in vertikalen Luftsäulen kommen vor allem Spektrometer zum Einsatz. Instrumente wie das TROPOspheric Monitoring Instrument (TROPOMI) an Bord des Satelliten Sentinel-5P der ESA, sind gut zur Überwachung atmosphärischer Spurengaskonzentrationen wie  $\text{NO}_2$  geeignet. TROPOMI ist ein Nadir ausgerichtete Spektrometer, dass das von der Erdoberfläche und Atmosphäre gestreute Sonnenlicht einfängt und darin Messungen im ultravioletten, sichtbaren, nahen Infrarot und kurzwelliges Infrarotlicht durchführt [81]. Zur Messung von  $\text{NO}_2$  kommt dabei der sichtbare 400 – 496 nm Kanal des Instruments zum Einsatz [82]. Für die Ableitung der vertikalen  $\text{NO}_2$ -Säulenkonzentration wird zunächst die gesuchte schräge Säulendichte (Slanted Column Density, SCD) ermittelt, wozu Differenzielle Optische Absorptions Spektroskopie (DOAS) verwendet wird. Diese liefert die Absolute Menge an  $\text{NO}_2$  entlang des effektiven Lichtwegs von der Sonne durch die Atmosphäre bis zum Satelliten (siehe Abb. 2.7). Die durch das Lambert-Beer'sche Gesetz beschriebene Absorption ist dabei vereinfacht durch

$$I(\lambda) = I_0(\lambda)e^{-\sum_i \sigma_i(\lambda)N_{s,i}} \quad (2.14)$$

gegeben.  $I(\lambda)$  ist das gemessene und  $I_0(\lambda)$  das absorptionsfreie Intensitätsspektrum.  $\sigma_i(\lambda)$  ist der bekannte Absorptions-Wirkungsquerschnitt und  $N_{s,i}$  die gesuchte SCD des Gases  $i$ . Ein Anpassungsalgorithmus sucht nach der Lösungsmenge für die Gasmenge in der schrägen Säulendichte  $N_{s,i}$  in Gleichung (2.14) für die sich die geringste quadratische Abweichung zu den gemessenen Absorptionsspektren ergibt [83]. Im Anschluss wird die schräge Säulendichte ( $N_s$ ) in einen stratosphärischen ( $N_s^{\text{strat}}$ ) und einen troposphärischen ( $N_s^{\text{trop}}$ ) Anteil aufgeteilt. Dies geschieht auf Basis eines chemischen Transportmodells, welches die

Satellitenbeobachtungen assimiliert und die beiden Komponenten der SDC trennt [82]. Zuletzt werden die SCD-Komponenten in vertikale stratosphärische und troposphärische NO<sub>2</sub>-Säulen (Vertical Columns, VC) überführt [83, 82]. Die aus der DOAS-Messung gewonnene schräge Säule (Slanted Column, SC) hängt von der auf dem Lichtweg vorhandenen Luftmenge ab. Letztere hängt von vielen Faktoren ab, einschließlich der Betrachtungsgeometrie und der vertikalen Verteilung des Absorbers. Der Einfluss des Lichtwegs wird dann durch den Luftmassenfaktor (Air Mass Factor, AMF) ausgedrückt der nach seiner Definition das Verhältnis der beiden Säulenarten beschreibt [84]:

$$\text{AMF} = \frac{\text{SC}}{\text{VC}}. \quad (2.15)$$

Die Luftmassenfaktoren werden aus einer Kombination aus tabellierten AMF-Daten und Modellinformationen zur vertikalen NO<sub>2</sub>-Verteilung hergeleitet. Die tabellierten AMFs sind dabei abhängig von der Höhe, Satellitengeometrie, Wolkenabdeckung sowie der Erhebung und Albedo der Erdoberfläche [83].



## 3 Daten

Die Daten, die für das Training der Modelle und die Abschätzung der bodennahen Schadstoffkonzentrationen herangezogen wurden, umfassen Messdaten aus Umweltüberwachungsstationen, Satellitendaten, meteorologische Daten aus atmosphärischen Modellen, sowie Daten zur Bevölkerungsdichte und zur Bebauung. Die betrachtete Zeitspanne umfasst die Jahre 2018 bis 2022, wobei für die Jahre 2021 und 2022 keine Daten zum troposphärischen  $\text{NO}_2$  vorliegen. Als Untersuchungsgebiet wurde der Bereich zwischen  $2^\circ$  und  $16^\circ$  östlicher Länge und  $46^\circ$  und  $56^\circ$  nördlicher Breite gewählt. Damit ist die gesamte Fläche Deutschlands und der Beneluxstaaten abgedeckt und Teilbereiche der umgebenden Staaten (Frankreich, Dänemark, Polen, Tschechien, Österreich, Slowenien, Italien, Schweiz).

Die meteorologischen Daten, Satelliten-AOD, sowie einige weitere Satellitendaten (LST, LC, NDVI) entsprechen den Daten von Handschuh et al. [27] und wurden zum Zweck dieser Arbeit zur Verfügung gestellt. Diese Datensätze wurden auf einem einheitlichen Gitter mit  $0.01^\circ \times 0.01^\circ$  horizontaler Auflösung für das Untersuchungsgebiet zur Verfügung gestellt. Tabelle 3.1 zeigt alle Variablen, die für die Modellentwicklung genutzt wurden.

### 3.1 Satellitendaten

Als Grundlage für die Feinstaubvorhersage dient die AOD. Es werden Daten des Moderate Resoluten Imaging Spektroradiometer (MODIS) an Bord der NASA-Satelliten Terra und Aqua genutzt, welche das untersuchte Gebiet näherungsweise um 10:30 Uhr und 13:30 Uhr Ortszeit überfliegen. Das verwendete AOD-Produkt (MCD19A2) basiert auf dem MAIAC-Retrieval-Algorithmus und kombiniert Terra- und Aqua-Messungen in einer räumlichen Auflösung von 1 km. Es werden AOD-Daten bei 550 nm verwendet, weil die Extinktionen durch  $\text{PM}_{2.5}$  in diesem Wellenlängenbereich gut sichtbar ist [85, 86].

Weitere eingesetzte MODIS-basierte Satellitendaten umfassen die Oberflächentemperatur (LST), Land Cover (LC), sowie den Normalized Difference Vegetation Index (NDVI). Zur Verbesserung der Datenabdeckung, wurden die täglichen LST-Daten mit einer ursprünglichen Auslösung von 1 km zu Monatsmittelwerten auf dem  $0.01^\circ \times 0.01^\circ$  Gitter kombiniert. Es wurden monatliche NDVI Datensätze der Produkte MYD11A1 und MOD11A1 mit einer räumlichen Auflösung von 1 km verwendet [87]. Daten zu LC sind je als Jahresmittelwerte verfügbar und stammen aus dem MODIS-Produkt MCD12Q1, mit einer Auflösung von 0.5 km. Es wurde das „LC\_Typ1“ Band verwendet, welches die Landbedeckung in 18 Klassen unterteilt [88]. Als Indikator für Stickstoffdioxid am Boden dient das Datenprodukt der troposphärischen  $\text{NO}_2$ -Säule das im Rahmen des S-VELD Projekts am Deutschen Zentrum für Luft und Raumfahrt (DLR) gewonnen wurden [89]. Die Daten basieren auf Messungen des TROPOspheric Monitoring Instrument (TROPOMI) Spektrometer an Bord des Copernicus-Satelliten Sentinel-5 precursor (S5P), der das Untersuchungsgebiet zwischen 13:00 und 14:00 Uhr lokaler Zeit [90] abtastet. Bei  $46^\circ$  nördlicher Breite deckt S5P in einem Orbit ein etwa  $33.6^\circ$  ( $2600 \text{ km}$ ) breites Band ab, wobei diese mit 14 Orbits pro Tag,  $25^\circ$  (Latitude) auseinanderliegen. Messungen aus aufeinanderfolgenden Überflügen zwischen 10:30 Uhr und 13:30 Uhr wurden zusammengefasst. Die Daten besitzen eine horizontale Auflösung von  $3.5 \text{ km} \times 5.5 \text{ km}$ . Die Verfügbarkeit der Daten beschränkt sich

**Tabelle 3.1:** Zielvariablen und mögliche Eingangsvariablen für die maschinellen Lernmodelle. Satellitendaten sind mit <sup>sat.</sup>, Stationsmessungen mit <sup>in-situ</sup> und Daten aus Wettermodelle mit <sup>m</sup> markiert.

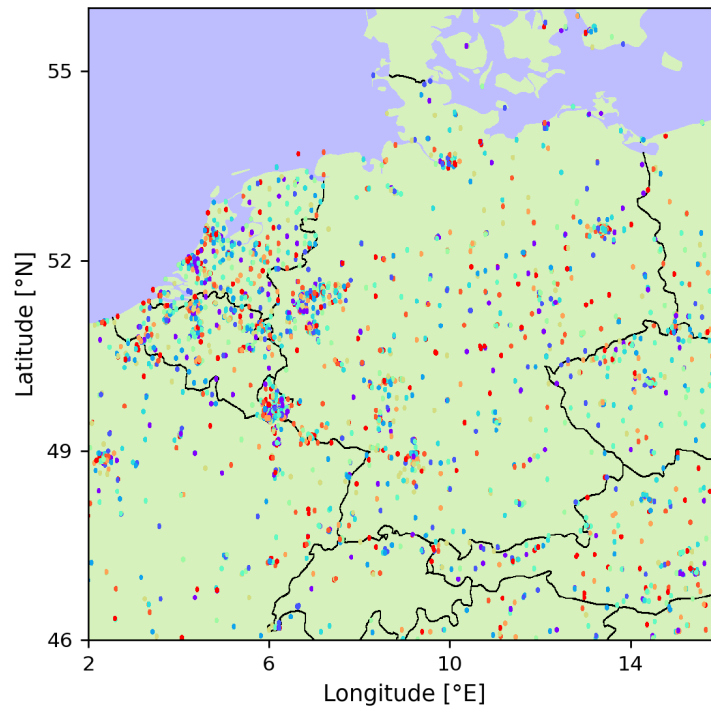
Bezeichnung	Name	Einheit
PM2.5 <sup>in-situ</sup>	PM <sub>2.5</sub> Feinstaubkonzentration	µg/m <sup>3</sup>
NO2 <sup>in-situ</sup>	NO <sub>2</sub> Bodenkonzentration	µg/m <sup>3</sup>
O3 <sup>in-situ</sup>	O <sub>3</sub> Bodenkonzentration	µg/m <sup>3</sup>
AOD <sup>sat.</sup>	Aerosol Optische Dicke	
TCNO2 <sup>sat.</sup>	NO <sub>2</sub> in troposphärischer Säule	µmol/m <sup>2</sup>
LST <sup>sat.</sup>	Oberflächentemperatur	K
Year	Jahr	
Month	Monat	
DoY	Jahrestag	
DoW	Wochentag	
NDVI <sup>sat.</sup>	Normierter-Differenz-Vegetationsindex	
LC <sup>sat.</sup>	Landbedeckung	
alt	Höhe über Meeresspiegle	m
pop_dens	Bevölkerungsdichte	1/km <sup>2</sup>
rd_type__(1-5)	Straßendichte (Typ 1-5)	m/km <sup>2</sup>
BLD <sup>m</sup>	Grenzschicht-Energiedissipation	J/m <sup>2</sup>
BLH <sup>m</sup>	Grenzschichthöhe	m
CAPE <sup>m</sup>	Labilitätsenergie	J/kg
DSR <sup>m</sup>	Direkte Sonneneinstrahlung	J/m <sup>2</sup>
RH <sup>m</sup>	Relative Feuchte	%
SP <sup>m</sup>	Luftdruck am Boden	hPa
SSR <sup>m</sup>	Sonneneinstrahlung am Boden (abwärtsgerichtet)	J/m <sup>2</sup>
STR <sup>m</sup>	Wärmestrahlung am Boden (abwärtsgerichtet)	J/m <sup>2</sup>
T2m <sup>m</sup>	Temperatur in 2 m Höhe	K
TCO3 <sup>m</sup>	O <sub>3</sub> in atmosphärischer Säule	kg/m <sup>2</sup>
TCWV <sup>m</sup>	Wasserdampf in atmosphärischer Säule	kg/m <sup>2</sup>
TD2m <sup>m</sup>	Taupunkt Temperatur in 2 m Höhe	K
TP <sup>m</sup>	Niederschlagshöhe	m
Vis <sup>m</sup>	Sichtweite	m
Wind <sup>m</sup>	Windgeschwindigkeit	m/s

auf die Jahre 2018 bis 2020, wodurch der Betrachtungszeitraum für die Schadstoffe NO<sub>2</sub> und O<sub>3</sub> auf diese Zeitspanne beschränkt ist.

## 3.2 Stationsdaten für bodennahe Schadstoffwerte

Die meisten europäischen Länder besitzen Messtationen zur in-situ Überwachung der Luftqualität, deren Daten durch die European Environment Agency (EEA) öffentlich zugänglich sind [91]. Von Interesse waren hier Messreihen für PM<sub>2.5</sub>, NO<sub>2</sub> und O<sub>3</sub> mit stündlicher Auflösung von allen Messstationen im Untersuchungsgebiet, ungeachtet von Stationstyp (Hintergrund, Industrie oder Verkehr) oder Umgebung (ländlich, urban, suburban). Insgesamt befinden sich 648 Stationen aus Deutschland (311), Belgien (84), Frankreich (74), Niederlande (47), Österreich (45), Tschechien (43), Polen (16), Italien (10), Schweiz (9), Dänemark (5) und Luxemburg (4) im betrachteten Gebiet. Diese sind, wie in Abb. 3.1 dargestellt, über das Gebiet verteilt. Angepasst an die Überflugzeit der Satelliten (Terra, Aqua, Sentinel 5P), wurden die hier verwendeten Stationsdaten durch Mittelung der

stündlichen in-situ Messwerte eines Tages zwischen 10:30 Uhr und 14:30 Uhr gewonnen.



**Abbildung 3.1:** Lage der Messstationen im Untersuchungsgebiet. Einfärbung der einzelnen Punkte ist zufällig gewählt.

### 3.3 Meteorologische Felder

Meteorologische Variablen, welche die Verteilung von Luftschadstoffen am Boden, sowie ihre Entstehung beeinflussen können, wurden ebenfalls in das Modell integriert (vgl. Tabelle 3.1). Es wurden Daten von numerischen Wettervorhersagen des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) verwendet. Dem hochauflösenden 10 Tage Vorhersagedatensatz Set I - HRES wurden die Daten um 12:00 Uhr entnommen, um die Überflugzeit der verschiedenen Satelliten möglichst gut anzupassen. Die ursprünglich mit  $0.1^\circ$  aufgelösten Daten wurden auf das gemeinsame  $0.01^\circ$ -Gitter übertragen. [92]

### 3.4 Ortsbezogene Daten

Zusätzlich zu den oben genannten Daten, wurden zur besseren Schätzung der Bodenemissionen von  $\text{NO}_2$ ,  $\text{PM}_{2.5}$  und entsprechenden Vorläufersubstanzen weitere Datensätze herangezogen.

Zum einen wurde die Straßendichte als Maß für Verkehrsemissionen betrachtet. Dazu wurde die Variable „road\_density“ aus dem Datensatz des Global Roads Inventory Project (GRIP4) mit einer Auflösung von  $0.083^\circ$  (5 arcmin) [93] verwendet, auf das  $0.01^\circ$ -Gitter skaliert und dabei für den gesamten Untersuchungszeitraum von 2018 bis 2022 als konstant angenommen. Die Straßendichte ist nach Größe der Straßen in fünf Kategorien aufgeteilt ((1) Autobahnen und Schnellstraßen, (2) primäre, (3) sekundäre, (4) tertiäre und (5) lokale Straßen). Die Dichte jedes Straßentyps wurde als eigene Variable betrachtet.

Dicht besiedelte Gebiete weisen in der Regel erhöhte Verkehrsemissionen und stärkere Heizaktivitäten auf. Die Stärke der Luftverschmutzung kann folglich vor allem in Bezug auf

NO<sub>2</sub> und PM<sub>2.5</sub> mit der Bevölkerungsdichte korreliert sein [94] welche daher eine geeignete Hilfsvariable darstellt. Aus diesem Grund wurde Schätzungen zur Bevölkerungsdichte aus der Global Human Settlement Layer Datensatz (GHS-POP R2023A) herangezogen [95]. Die Daten wurden von der ursprünglichen Auflösung von 100 m dem gemeinsamen Gitter (0.01°x0.01°) angepasst und analog zur Straßendichte zwischen 2018 und 2022 als konstant angenommen.

Durch verschiedene Faktoren nimmt auch die Höhenlage an einem bestimmten Ort Einfluss auf die Luftqualität. Durch das digitale Höhenmodell der TanDEM-X Mission mit 90 m Auflösung [96] wurde nach Anpassung an das oben genannte Gitter die geographische Höhe ebenfalls als zeitlich konstante Variable hinzugenommen.

## 4 Methodik

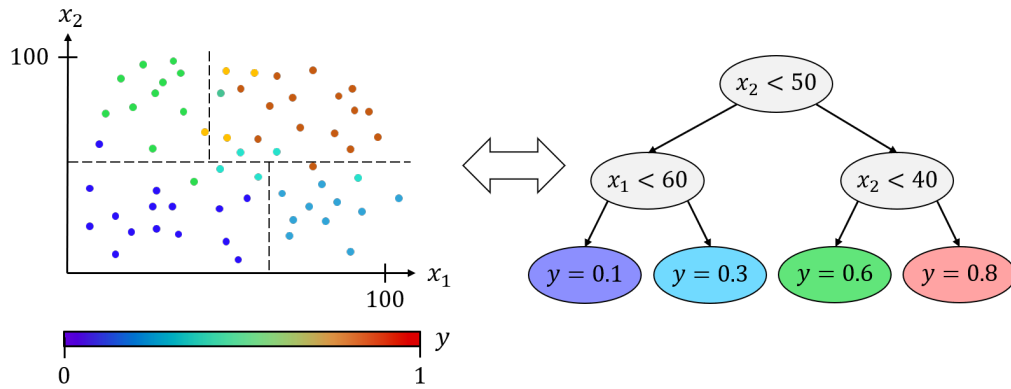
### 4.1 Random Forest Modell

Maschinelles Lernen (Machine Learning, ML) beruht auf der Entwicklung spezieller Algorithmen und statistischen Modellen zur Datenverarbeitung. Diese besitzen die Fähigkeit, aus Daten zu lernen und auf deren Grundlage Entscheidungen oder Vorhersagen zu treffen, ohne explizit dafür programmiert worden zu sein. Das Ergebnis des Lernprozesses ist ein Modell, welches eine Funktion zwischen den Eingangsvariablen und der Zielvariablen darstellt. Diese soll den statistischen Zusammenhang zwischen den beiden dabei möglichst gut beschreiben. Das hier eingesetzte Modell zur Vorhersage von  $\text{NO}_2$ -,  $\text{PM}_{2.5}$ - und  $\text{O}_3$ -Bodenkonzentrationen ist der Random Forest (RF) Algorithmus, welcher erstmals 2001 von Breiman [97] vorgeschlagen wurde. Das Random Forest Modell ist ein Ensemble aus einzelnen Entscheidungsbäumen zur Regression oder Klassifizierung (CART-Algorithmus), die in Abb. 4.1 veranschaulicht sind. Es ermöglicht die Vorhersage einer Zielvariable auf Basis mehrerer unabhängiger Eingangsvariablen (Merkmale). Im Folgenden soll dabei nur auf die Anwendung des RF für Regressionsprobleme eingegangen werden.

Entscheidungsbäume bestehen aus Knoten und Blättern. Eine Dateninstanz (Vektor aus allen Eingangsvariablen) wird an den Knoten jeweils mit einem Schwellenwert für ein bestimmtes Merkmal verglichen und entsprechend eingeordnet. Nach mehrmaliger Einordnung erreicht die Instanz einen Endpunkt (Blatt), dem ein fester Vorhersagewert zugeteilt ist, wobei die bei den Entscheidungsknoten betrachteten Merkmale und Schwellenwerte, sowie die Vorhersagewerte an den Blättern im Training bestimmt werden. Dazu wird wie in Abb. 4.1 ein Trainingsdatensatz in einzelne Bereiche aufgeteilt, die je ein Blatt repräsentieren und deren Vorhersagewert jeweils dem Mittelwert aller enthaltenen Datenpunkte entspricht. Schwellenwert und Merkmal für die Aufteilung werden dabei so gewählt, dass der mittlere quadratische Fehler der Datenpunkte zu ihrem Vorhersagewert über alle Bereiche minimal wird. Der Datensatz kann solange aufgeteilt werden, bis der Baum eine festgelegte maximale Tiefe (*max\_depth*) erreicht hat oder bis die Datenpunkte in einem Bereich eine minimale Anzahl erreichen.

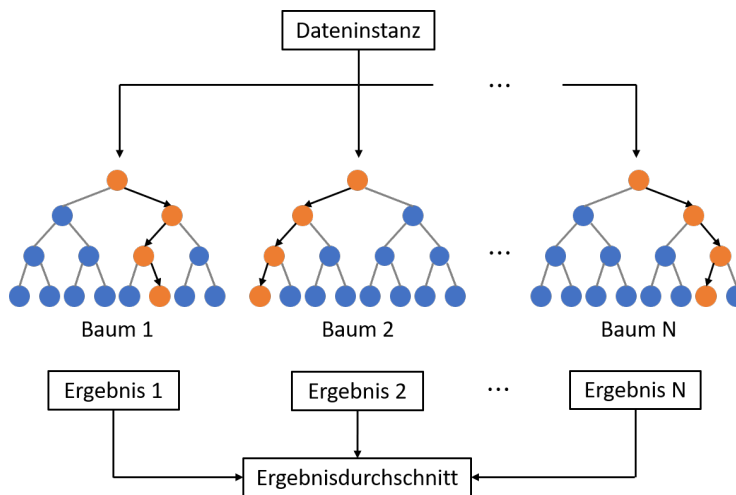
Die einzelnen Bäume werden auf zufälligen Stichproben der Trainingsmenge trainiert (bagging). Zusätzlich dazu steht beim Training eines Baumes für jeden Knoten nur eine zufällig gewählte Teilmenge aller Eingangsvariablen zur Auswahl, welche *max\_features* Merkmale enthält. Dadurch wird sichergestellt, dass sich die einzelnen Bäume des Ensembles hinreichend voneinander unterscheiden, was eine positive Wirkung auf die Fähigkeit des Modells zur Generalisierung hat [98]. Zur Vorhersage wird den einzelnen Bäumen, wie in Abb. 4.2 gezeigt, eine Dateninstanz übergeben und diese von jedem Baum individuell ausgewertet. Anschließend werden die Vorhersagen aller Bäume gemittelt wodurch sich aus den vielen schwachen Vorhersagen der Einzelbäume eine gute Vorhersage des Ensembles ergibt.

Die zufällige Auswahl von Merkmalen und Datenmengen im Training der einzelnen Bäume begrenzt die Gefahr der Überanpassung des Modells an die Trainingsdaten, vor allem im Fall von hochdimensionalen Datensätzen [98]. Die Architektur des RF ermöglicht es, die einzelnen Bäume unabhängig voneinander zu trainieren bzw. auszuwerten, wodurch



**Abbildung 4.1:** Training eines Entscheidungsbaums mit zwei Eingangsvariablen ( $x_1$  und  $x_2$ ) und der Zielvariable  $y$ . Der entstehende Baum hat eine Tiefe von 2.

sich Training und Auswertung einfach parallelisieren und die benötigte Rechenzeit ggf. stark reduzieren lässt. Ein weiterer Vorteil des Random Forest liegt darin, dass keine Notwendigkeit besteht, die Eingangs- und Zielvariablen zu skalieren oder zu normalisieren, was seine Anwendung erleichtert.



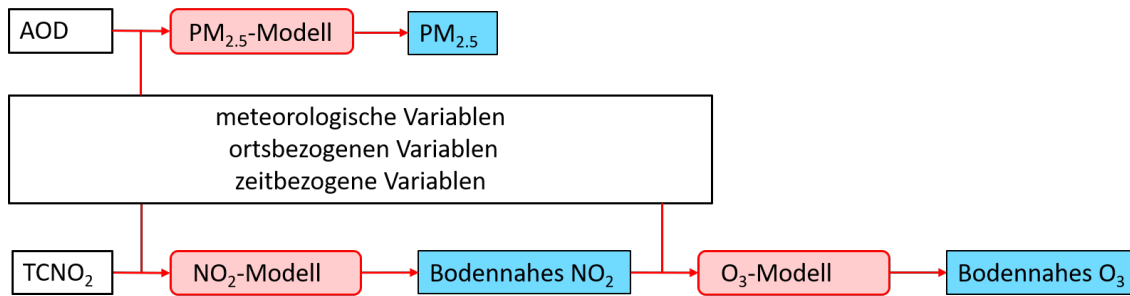
**Abbildung 4.2:** Aufbau eines Random Forest Modells aus  $N$  Entscheidungsbäumen. Die Bäume sagen unabhängig voneinander einen Wert für die Dateninstanz voraus. Das Gesamtergebnis ergibt sich aus dem Mittelwert der Einzelprediktionen.

## 4.2 Modellaufbau

Es wurde je ein Random Forest Modell mit den Zielvariablen PM2.5, NO<sub>2</sub> und O<sub>3</sub> und 31 potenziellen Eingangsvariablen (siehe Tab. 3.1) entwickelt. Dabei wurden auch Hilfsvariablen, wie das Jahr (Year), der Jahrestag (DoY) und der Monat (Month), sowie der Wochentag (DoW) generiert, welche dem Modell die zeitliche Einordnung der Daten ermöglicht. Zu Beginn werden alle flächenbezogenen Daten an den Positionen der Messstationen gefiltert und so räumlich und zeitlich den entsprechenden Stationsmessungen zugeordnet. Dabei wird jeweils der Wert des Gitterpunktes gewählt, der der Station am nächsten liegt. Der entstandene Trainingsdatensatz enthält ca. 1.2 Mio. Dateninstanzen für die Jahre 2018 bis 2022. TCNO<sub>2</sub> ist, wie schon in Kap. 3.1 erwähnt, nur bis zum 31.12.2020 verfügbar; die Dateninstanz für die NO<sub>2</sub>- und O<sub>3</sub>-Modelle fallen dementsprechend kleiner aus.

Zum Training werden den Modellen der Trainingsdatensatz mit den Stationsmessungen als

Zielvariable und meteorologischen Variablen, Satellitendaten sowie örtlichen und zeitlichen Hilfsgrößen als Eingangsvariablen übergeben. Da Stickstoffdioxid eine wichtige Vorläufersubstanz von Ozon darstellt, geht der Vorhersagewert des  $\text{NO}_2$ -Modells, wie in Abb. 4.3 gezeigt, als Merkmal in das Ozonmodell ein. Die Möglichkeit,  $\text{NO}_2$  als Vorläufersubstanz sekundärer Aerosole in das  $\text{PM}_{2.5}$ -Modell einfließen zu lassen, wurde ebenfalls getestet. Allerdings konnte keine signifikante Verbesserung der Vorhersagen festgestellt werden. Da nur vollständige Dateninstanzen vom Modell verwertet werden können, wodurch mehr Eingangsvariablen die Datenabdeckung potenziell verringern, wurde für dieses Modell auf die Hinzunahme der Variable verzichtet.



**Abbildung 4.3:** Genereller Modellaufbau der  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ -, und  $\text{O}_3$ -Modelle. Die Modelle für  $\text{NO}_2$  und  $\text{PM}_{2.5}$  beziehen jeweils Satellitenmessungen ( $\text{TCNO}_2$  bzw.  $\text{AOD}$ ) als Teil ihrer Eingangsvariablen. Die Vorhersagen des  $\text{NO}_2$ -Modells werden weiterverwertet und dienen als Eingangsvariable für das Modell zur Ozonvorhersage.

Die Auswahl der geeigneten Eingangsvariablen für jedes Modell ist ausschlaggebend für die Vorhersageleistung des Modells. Viele Variablen liefern zwar zusätzliche Informationen als Basis für die Vorhersagen, erhöhen jedoch den Rechenaufwand und verkleinern im Fall von lückenhaften Satellitendaten auch die verfügbare Trainingsdatenmenge. Das bestimmen der optimalen Eingangsvariablen wird mithilfe von Sequenzieller Feature Selektion (SFS) umgesetzt. Dabei wird, ausgehend vom kompletten Datensatz mit allen potenziellen Merkmalen, abwechselnd ein Merkmal entfernt und die resultierende Modellgenauigkeit mit der vorherigen verglichen. Die Variablenmenge, für die sich die besten Resultate erzielen ließ wird übernommen und der Auswahlprozess erneut durchgeführt, bis sich keine Verbesserung in der Modellgenauigkeit mehr einstellt. Die auf diese Weise ermittelten Eingangsvariablen für das jeweilige Modell sind in Tabelle 4.1 dargestellt.

Zur praktischen Implementierung des Random Forest Modells wurde die `RandomForestRegressor` Objektklasse der python-Bibliothek `scikit-learn` (v.1.3.0) verwendet [99]. Die optimale Einstellung der Modellparameter wurde empirisch ermittelt und ist in Tabelle 4.2 gezeigt.

## 4.3 Modelltestverfahren

Bei komplexen Modellen kommt es leicht zu einer Überanpassung an die Trainingsdaten, wodurch sie unbekannte aber gleich verteilte Daten nur noch schlecht vorhersagen können. Das Ziel ist es, einen möglichst generellen Zusammenhang zwischen Eingangsvariablen und Zielvariable zu finden, wofür eine strikte Trennung von Test- und Trainingsdaten notwendig ist.

Die Genauigkeit der RF-Modelle wurde durch 10-fache Kreuzvalidierung (CV10) bestimmt. Dabei wird der Datensatz zufällig in zehn möglichst gleich große Teilmengen

**Tabelle 4.1:** Optimale Eingangsparameter ermittelt durch Sequential Feature Selection, zeilenweise angeordnet als Satellitendaten, meteorologische Parameter, ortsfeste Merkmale und zeitliche Variablen.

Zielvar.	Optimale Eingangsvar.	Ungenutzte Var.	Anzahl
PM2.5	AOD,		22
	BLD, BLH, CAPE, LC, RH, SP, SSR, STR, TD2m, TP, Vis, T2m, TCO3,	DSR, LST, TCWV, Wind	
	alt, rd_type_1, rd_type_5, pop_dens,	rd_type_2/3/4	
NO2	DoY, DoW, Month, Year,		20
	TCNO2,	BLD, DSR, RH, SST, TCNO3, TCWV, TD2m, Vis, Wind	
	BLH, CAPE, LC, LST, NDVI, SP, STR, T2m, TP,		
O3	alt, rd_type_1, rd_type_2, rd_type_3, rd_type_4, rd_type_5, pop_dens,		18
	DoY, DoW, Year,	Month	
	NO2_pred,		
O3	BLD, BLH, LST, SP, SSR, T2m, TCWV, RH, Vis, TCO3,	CAPE, DST, LC, NDVI, STR, TD2m, TP, Wind	18
	alt, rd_type_4,	rd_type_1/2/3/5	
	pop_dens, DoY, DoW, Month, Year		

**Tabelle 4.2:** Auflistung der optimalen Modellparameter für die Random Forest Modelle.

Modellparameter	Bedeutung	Wert
<i>n_estimators</i>	Anzahl der Bäume im Modell	400
<i>max_features</i>	Anzahl der berücksichtigten Merkmale für bei der Aufspaltung eines Baumes	5
<i>bootstrap</i>	Ob Bootstrapping genutzt werden soll	True
<i>max_depth</i>	Die maximale Tiefe eines Baumes	20
<i>criterion</i>	Funktion zur Messung der Qualität einer Aufspaltung	„squared_error“
<i>n_jobs</i>	Anzahl der Threads zur Parallelisierung	-1



gespalten. Im Anschluss wird jede der Teilmengen genau einmal als Testdatensatz designiert, während mit den übrigen neun Datenmengen das Modell trainiert wird. Für jeden Testdatensatz werden Vorhersagen durch das trainierte Modell getroffen, sodass nach Abschluss aller Trainingsdurchläufe genau eine Vorhersage für jeden Datenpunkt im gesamten Datensatz vorliegt. Nun wird die Genauigkeit des Modells durch vergleichen der Testdatenvorhersagen mit ihren jeweiligen Zielwerten ermittelt. Dazu wird das Bestimmtheitsmaß (Determinationskoeffizient,  $R^2$ )

$$R^2 = 1 - \frac{\sum_i^{N_T} (y_i - \hat{y}_i)^2}{\sum_i^{N_T} (y_i - \bar{y})^2} \quad (4.1)$$

berechnet, wobei  $\hat{y}_i$  die Modellvorhersagewerte und  $y_i$  die Zielwerte aus den Trainingsdaten darstellen.  $\bar{y}$  ist der Mittelwert der Zielwerte und  $N_T$  die Anzahl der Testdaten, außerdem gilt  $-\infty < R^2 < 1$ . Je höher das Bestimmtheitsmaß, desto akkurater sind die Vorhersagen des Modells.  $R^2 = 1$  bedeutet eine komplette Übereinstimmung der Vorhersagen mit den Zielwerten, während  $R^2 = 0$  bedeutet, dass das Modell genauso schlechte Werte liefert, wie eine Vorhersage durch den Mittelwert.

Eine weitere hilfreiche Metrik stellt der mittlere quadratische Fehler (MSE) bzw. die Quadratwurzel des mittleren quadratischen Fehlers (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N_T} \sum_i^{N_T} (y_i - \hat{y}_i)^2} \quad (4.2)$$

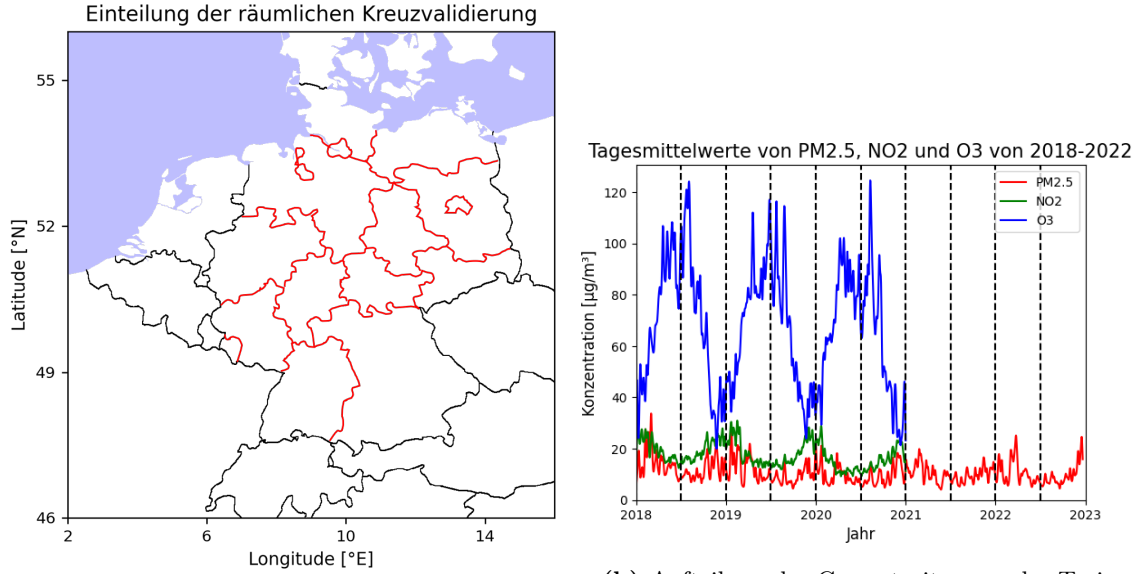
der Vorhersagen dar. Die tendenzielle Richtung des Fehlers ist, durch den mittleren Bias der Vorhersagen gegeben:

$$\text{Bias} = \frac{1}{N_T} \sum_i^{N_T} (\hat{y}_i - y_i) \quad (4.3)$$

Dieser gibt an, wie stark das Modell die Zielwerte im Mittel überschätzt (Bias > 0) bzw. unterschätzt (Bias < 0).

Zusätzlich zur oben genannten Kreuzvalidierung (CV10) werden die Modelle einer räumlichen und zeitlichen Kreuzvalidierung unterzogen, bei der jeweils die Daten der Stationen, bzw. der Messtage zufällig auf die einzelnen Untermengen aufgeteilt werden. Um im Weiteren eine noch stärkere Trennung zwischen Test- und Trainingsdaten zu erzielen wird der Datensatz im Anschluss nicht zufällig aufgeteilt, sondern so, dass sich zeitlich und räumlich nahe Datenpaare möglichst in der selben Teilmenge befinden. Zur räumlichen Kreuzvalidierung werden alle Stationen innerhalb Deutschlands nach Bundesländern und außerhalb nach dem jeweiligen Staat gruppiert, wodurch sich insgesamt 26 Teilmengen unterschiedlicher Größe ergeben (siehe Abb. 4.4a). Zur zeitlichen Kreuzvalidierung wird der Untersuchungszeitraum, wie in Abbildung 4.4b gezeigt in zehn gleich große Zeitabschnitte unterteilt. Da  $\text{NO}_2$  und  $\text{O}_3$  nur bis einschließlich 2020 vorhergesagt werden können, liegen hier effektiv nur 6 Abschnitte vor.

Neben dieser *geordneten* räumlichen bzw. zeitlichen Kreuzvalidierung wird zudem eine geringfügig stärker randomisierte Variante umgesetzt. Dabei werden alle Dateninstanzen räumlich nach der zugehörigen Messstation bzw. zeitlich nach den Messtagen gruppiert. Die einzelnen Gruppen werden anschließend *zufällig* auf die zehn Teilmengen der CV10 verteilt und analog zu dieser ausgewertet.



(a) Aufteilung des Untersuchungsgebiets bei der geordneten räumlichen Kreuzvalidierung nach Bundesländern und Staaten. Alle Stationen eines Untergebiets bilden jeweils eine Teilmenge der Kreuzvalidierung.

(b) Aufteilung des Gesamtzeitraums der Trainingsdaten in 10 zusammenhängende Zeitabschnitte zur (geordneten) zeitlichen Kreuzvalidierung. Dargestellt sind die geglätteten Tagesmittelwerte für PM<sub>2.5</sub>, NO<sub>2</sub> und O<sub>3</sub> für den gesamten Untersuchungszeitraum.

**Abbildung 4.4:** Darstellung der Datenaufteilung für die räumliche und zeitliche Kreuzvalidierung.

## 4.4 Modellanalysemethoden

Anders als bei einzelnen Entscheidungsbäumen, lässt sich der Entscheidungsmechanismus eines Random Forest Modells, aufgrund der hohen Anzahl der enthaltenen Bäume, nur noch schwer durch betrachten der einzelnen Knoten verstehen. Für die Analyse der modellgenerierten Daten ist es jedoch sinnvoll, dass die vom RF-Modell erlernten Zusammenhänge zwischen Eingangs- und Zielvariablen dargestellt werden können. Dazu wird das Modell im Folgenden mittels Methoden der Explainable AI untersucht.

### 4.4.1 Feature Importance Analyse

Um die Vorhersage des Modells zu begründen und um festzustellen, welche Merkmale (Features) besonders zur Vorhersagefähigkeit des Modells beitragen, wird eine Feature Importance (FI) Analyse durchgeführt. Die Methode misst die Relevanz eines Merkmals durch Veränderungen in der Modellleistung, wenn dem entsprechenden Merkmal der Informationsgehalt entzogen wird. Dazu werden die Werte einer einzelnen Eingangsvariable  $x_i$  des Testdatensatzes zufällig untereinander vermischt, neu den Instanzen zugeordnet und anschließend die Modellperformance  $R^2_{\text{perm}}$  auf dem resultierenden Datensatz evaluiert. Dies führt zu einer Erhöhung des mittleren Vorhersagefehlers, bzw. einer Verringerung des Bestimmtheitsmaßes (Mean Decrease in Accuracy, MDA)

$$\text{MDA} = R^2 - R^2_{\text{perm}} \quad (4.4)$$

welche umso stärker ausfällt, je relevanter das Merkmal für die Modellvorhersage ist. Die Berechnung der MDA ist abhängig von der zufälligen Neuverteilung der Datenpunkte. Der dadurch verursachte statistische Fehler wird durch mehrmaliges durchführen der

Berechnung (hier fünf Durchläufe) und anschließender Mittelwertbildung minimiert.

Das zur Bestimmung der Feature Importance verwendete Datenset ist dabei nicht zwingend festgelegt und kann Interessensabhängig gewählt werden. Um festzulegen mit welchen Features die besten Vorhersagen möglich sind ist eine Auswertung auf den Testdatensatz sinnvoll. Die Verwendung des Trainingsdatensatzes ist jedoch dann zu bevorzugen, wenn von Interesse ist, auf welche Eigenschaften das Modell seine Vorhersagen stützt, unabhängig davon, ob diese zutreffend sind oder nicht. Im Folgenden wird die FI stets auf dem Testdatensatz ermittelt.

#### 4.4.2 Partielle Abhängigkeiten

Da die FI-Analyse nur auf die Bedeutungsstärke eines Merkmals für die Modellvorhersage, nicht aber die Art der Abhängigkeit zwischen Eingangsvariablen und Zielvariable aufzeigt, ist es im Zuge der Modellanalyse hilfreich, die vom Modell erlernten Zusammenhänge direkt zu untersuchen. Hierfür werden die partiellen Abhängigkeiten des Modells in sog. Partial Dependence Plots (PDP) veranschaulicht, welche die mittlere marginale Wirkung eines oder mehrerer Merkmale auf das Ergebnis aufzeigen, während die Effekte aller anderen Merkmale neutralisiert werden.

Sei  $M(x_0, x_1, x_2, \dots) = M(\vec{X})$  die erlernte Modellfunktion, abhängig von allen Eingangsvariablen  $x_i$  und  $N$  die Anzahl der Instanzen im Trainingsdatensatz  $\{\vec{X}_n\}$ . Dann ist der im PDP gezeigte Zusammenhang für das Feature  $x_i$  durch

$$M^*(x_i) = \frac{1}{N} \sum_{n=0}^N M(x_{n,0}, \dots, x_{n,i-1}, x_i, x_{n,i+1}, \dots, x_N) \quad (4.5)$$

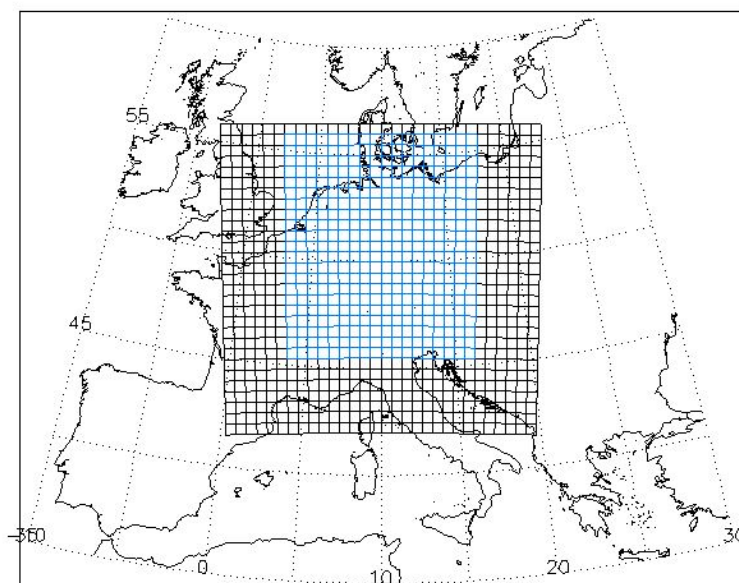
gegeben. Da für die Berechnung eines einzelnen Funktionswertes das Modell für alle  $N$  Testinstanzen ausgewertet werden muss, kann die Darstellung eines PDP, je nach Auflösung, bei großen Datenmengen sehr zeitaufwendig sein.

### 4.5 Analysemethoden für flächenhafte Daten

Im Anschluss an das Training und die Validierung der Modelle werden flächenhafte Schadstoffdaten generiert und zur Interpretation weiterverarbeitet.

#### 4.5.1 Vorhersage und Weiterverarbeitung der Flächendaten

Zunächst werden die gegitterten Eingangsvariablen für jeden Tag zusammengetragen und mithilfe eines auf der gesamten Trainingsmenge trainierten RF-Modells für  $\text{PM}_{2.5}$  und  $\text{NO}_2$  die Bodenkonzentration der entsprechenden Luftschadstoffe für jeden Gitterpunkt einzeln vorhergesagt. Mit der vorhergesagten  $\text{NO}_2$ -Konzentration werden daraufhin analog die Vorhersagen der Ozonwerte generiert. Die entstandenen Datensätze umfassen für  $\text{PM}_{2.5}$  alle Tage von 01.01.2018 bis zum 31.12.2022 bzw. bis zum 31.12.2020 für Stickstoffdioxid und Ozon und besitzen die gleiche räumliche Ausdehnung wie die in Kapitel 3 beschriebenen Flächendaten. An Gitterpunkten, an denen z.B. aufgrund von Wolkenbedeckung nicht alle Eingangsvariablen vorhanden sind, können keine Modellvorhersagen getroffen werden, wodurch diese leer bleiben.



**Abbildung 4.5:** Klassifizierungsgebiete für die objektive Wetterlagenklassifikation des DWD. Zur Bestimmung der Wetterlage für Deutschland wird nur der blaue Ausschnitt des Gitters benötigt, welcher sich sehr gut mit dem Untersuchungsgebiet dieser Arbeit deckt.

#### 4.5.2 Datenaufteilung und Wetterlagenklassifikation

Die täglichen Schadstoffdaten, sollen zur Analyse auf Basis ihrer meteorologischen Gesamtsituation in verschiedene Gruppen eingeteilt werden. Im Anschluss wird durch Mittelwertbildung innerhalb der Gruppen ein generelles Bild der Schadstoffverteilung gewonnen.

Eine erste Einteilung der Daten geschieht nach der Jahreszeit, die sowohl Einfluss auf anthropogene Schadstoffemissionen, als auch auf meteorologische Verhältnisse nimmt. Für den Beginn von Frühling, Sommer, Herbst und Winter wurde dabei jeweils der 21. Tag des März, Juni, September und Dezember gewählt.

Für eine direktere Einteilung nach meteorologischen Verhältnissen wird der Datensatz nach der vorherrschenden Großwetterlage (GWL) gruppiert, welche die über mehrere Tage stabile Wetterlage über einem Großraum beschreibt. Es wird nach der objektiven Wetterlagenklassifikation des DWD eingeteilt, welche im Gegensatz zu andern Klassifikationsmethoden (z.B. Hess Brezowsky) eindeutig numerisch nachvollziehbar ist, und keinen Interpretationsspielraum lässt [100]. Das Klassifikationsgebiet ist in Abb. 4.5 zu sehen, wobei bei der Klassifikation für Deutschland nur der blaue Ausschnitt benötigt wird, welcher sich gut mit dem hier betrachteten Untersuchungsgebiet deckt [101]. Das Verfahren verarbeitet das Geopotential, die Temperatur, die relative Feuchte, sowie zonale und meridionale Windkomponenten in unterschiedlichen Druckhöhen der Troposphäre. Daraus erfolgt die Einteilung in die Wetterklassen nach den folgenden Kriterien [102]:

1. **Windindex:** Wenn die Windrichtung an mehr als  $2/3$  der gewichteten Datenpunkte in einen festgelegten Sektor von  $90^\circ$  fällt liegt eine vorherrschende Windrichtung vor. Diese wird dann der am nächsten gelegenen Hauptwindrichtung, also Südost (SO), Südwest (SW), Nordost (NO) oder Nordwest (NW), zugeteilt, die den Windindex bestimmt. Findet sich keine vorherrschende Windrichtung wird der Windindex „XX“ zugeteilt.

2. **Zyklonalitätsindex bei 950 hPa:** Der Zyklonalitätsindex beschreibt das Vorzeichen der mittleren Krümmung des Geopotentials im Klassifikationsgebiet bei 950 hPa (untere Troposphäre). Negative Werte erhalten den Index „A“ (antizyklonal), positive werden mit „Z“ (zyklonal) beschrieben. Die Zyklonalität gibt Auskunft über die Lage von Hoch- und Tiefdruckgebieten in der aktuellen Wetterlage.
3. **Zyklonalitätsindex bei 500 hPa:** Wird wie in 2. berechnet, jedoch bei 500 hPa in der mittleren Troposphäre.
4. **Feuchteindex:** Der Atmosphäre wird entweder der Index feucht (F) oder trocken (T) zugeordnet, auf Basis dessen, ob der mittlere Wassergehalt der Atmosphäre über oder unter dem 30-Tage-Mittelwert um den entsprechenden Tag liegt.

Mit der Zusammensetzung der Kennung durch

$$\text{Wind (5)} \times \text{Zyklonalität 950 hPa (2)} \times \text{Zyklonalität 500 hPa (2)} \times \text{Feuchte (2)}$$

lässt sich so zwischen 40 verschiedenen Wetterlagen unterscheiden. Der DWD stellt einen Datensatz mit der Klassifizierung aller Tage von 1979 bis zum aktuellen Datum zur Verfügung [103].

Zusätzlich zur Einteilung der Vorhersagedaten nach der objektiven Großwetterklassifikationen, wird nach den vier Klassifikationsparametern, Windklasse, Zyklonalitätsklassen und Feuchteklasse, aufgeteilt. Die relativen Häufigkeiten der einzelnen Wetterlagen unterscheiden zwischen den Jahreszeiten zum Teil stark [103], was die Trennung der Wettereffekte von den jahreszeitlich bedingten Emissionsänderungen erschwert. Zur Minderung des Problems werden die Gruppen für die Großwetterlagen und Indizes weiter nach Jahreszeit aufgespalten.

### 4.5.3 Globale räumliche Autokorrelation

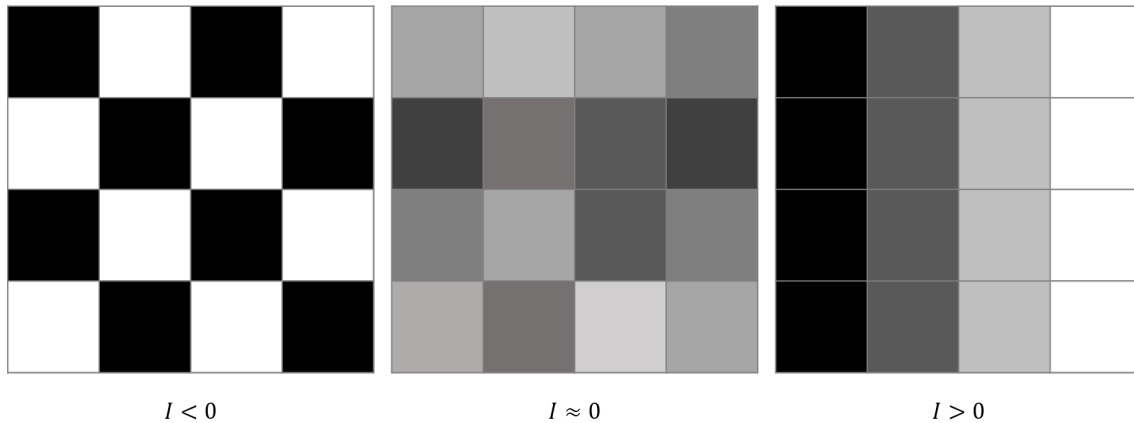
Das Aufkommen hoher Schadstoffwerte in Bodennähe ist an räumliche Merkmale sowie großflächige Wetterbedingungen geknüpft. Zudem kommt es durch Wind und Diffusion zum Austausch von Schadstoffen, wodurch sich die Luftverunreinigungen benachbarter Gebiete gegenseitig beeinflussen können. Im Folgenden wird daher auf ein Verfahren eingegangen, durch das die räumliche Korrelationen in den flächenhaften Mittelwerten ermittelt werden kann. Diese ist ausschlaggebend für die Bildung lokaler Hotspots und Coldspots, welche im Folgeabschnitt (4.5.4) untersucht werden. Eine Quantifizierung der Clusterbildung in einer räumlichen Datenverteilung ist durch die globale Moran's  $I$  Statistik beschrieben. Für  $N$  zusammenhängende Gitterpunkte  $i$ , wie sie in Abb. 4.6 dargestellt sind, lässt sich die Stärke der räumlichen Autokorrelation über

$$I = \frac{N}{W} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (4.6)$$

berechnen, wobei  $x_i$  den Gruppenmittelwert am Gitterpunkt  $i$  und  $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$  den Mittelwert aller Schadstoffwerte im Gebiet und der Gruppe angibt [104]. Die Komponente  $(\mathbf{Z})_i = z_i = d_i - \bar{x}$  ist dabei Teil eines Vektors der die Abweichungen der Gitterpunkte vom Mittelwert enthält. Moran's  $I$  ist ein Maß dafür, wie ähnlich sich im Mittel benachbarte Gebiete sind, wobei die Nachbarschaft eines Gebietes über die  $N \times N$  Gewichtsmatrix  $\hat{\mathbf{W}}$  festgelegt ist. Ihre Komponenten  $w_{ij}$ , mit  $w_{ij} \geq 0$  und  $w_{ii} = 0$ , beschreiben, wie stark die Verknüpfung zweier Gitterpunkte gewichtet werden soll. Punkte mit positiver Gewichtung müssen dabei nicht direkt aneinandergrenzen. In Gl. (4.6) wird durch die Summe aller

Gewichte  $W = \sum_{i=1}^N \sum_{j=1}^N w_{ij}$  geteilt; der Term  $N/W$  kann jedoch vernachlässigt werden, wenn die Gewichtsmatrix  $\hat{W}$  reihenweise oder zeilenweise normalisiert ist, d.h. wenn  $\sum_{i=1}^N w_{ij} = 1$  gilt.

Generell kann anhand des globalen Morans  $I$ -Index, welcher innerhalb des Wertebereichs  $-1 < I < 1$  liegt, zwischen drei groben Zuständen unterschieden werden. Für  $I > 1$  liegt positive räumliche Autokorrelation vor, was bedeutet, dass hohe Werte (im Vergleich zum Mittelwert) tendenziell von ebenfalls hohen Werten umgeben sind und niedrige Werte sich im Gegensatz in Bereichen mit weiteren niedrigen Werten finden. Für  $I \approx 0$  sind die Werte mehr oder weniger zufällig auf das Gebiet verteilt. Es kann dann davon ausgegangen werden, dass sich die einzelnen Knotenpunkte kaum gegenseitig beeinflussen. Negative Korrelationen  $I < 0$  zeigen die Tendenz hoher und niedriger Werte sich benachbart anzuordnen (vgl. Abb. 4.6).



**Abbildung 4.6:** Beispielhafte räumliche Verteilungen einer Größe mit verschieden starker globaler Autokorrelation. Links: Abwechselnde Verteilung / Negative räumliche Autokorrelation ( $I < 0$ ), Mitte: Zufällige Verteilung / Keine räumliche Autokorrelation ( $I = 0$ ), Rechts: Zusammenhängende Verteilung / positive räumliche Autokorrelation ( $i > 0$ ).

#### 4.5.4 Local Indicators of Spatial Association (LISA)

Wenn Schadstoffquellen, Schadstoffsenken oder meteorologische Eigenschaften stark lokalisiert sind, begünstigt dies die Bildung von örtlich begrenzten Bereichen mit deutlich besserer oder schlechterer Schadstoffbilanz als in ihrer Umgebung. Da sich bei der Identifizierung solcher Hotspots bzw. Coldspots nicht auf die subjektive Einschätzung durch visuelle Bewertung der Flächendaten verlassen werden kann, wird diese durch die Berechnung lokaler Indikatoren räumlicher Autokorrelation (LISA) verwirklicht. Ist die Gewichtsmatrix  $\hat{W}$  zeilenweise normalisiert, so kann Gl. (4.6) durch die Verwendung des lokalen Morans  $I_i$  formuliert werden [104]:

$$I = \sum_{i=1}^N \frac{I_i}{N} . \quad (4.7)$$

Dabei sind

$$I_i = \frac{z_i}{m_2} \sum_{j=1}^N w_{ij} z_j \quad (4.8)$$

die LISA, mit dem zweiten Moment

$$m_2 = \sum_{i=1}^N \frac{z_i^2}{N} . \quad (4.9)$$

Die Nullhypothese der Moran-Statistik geht davon aus, dass die beobachteten Autokorrelationen durch zufällige Verteilung der Werte auf die Gitterpunkte im Untersuchungsgebiet entstehen. Durch die zufällige Permutation der Werte an den Datenpunkten (abgesehen von Punkt  $i$ ) werden Vergleichswerte  $I_{i,\text{sim}}$  simuliert, die auf der Zufallsannahme beruhen. Liegt  $I_i$  nah an den simulierten  $I_{i,\text{sim}}$ -Werten, kann die Nullhypothese nicht abgelehnt werden und es ist wahrscheinlich, dass die beobachtete Verteilung das Ergebnis zufälliger räumlicher Prozesse darstellt. Liegt  $I_i$  jedoch weit außerhalb der Vergleichswerte, so kann die Nullhypothese abgelehnt werden, da der Datensatz eine größere räumliche Gruppierung aufweist als durch Zufallsprozesse erklärt werden kann. Die Wahrscheinlichkeit für das Zutreffen der Nullhypothese ist durch den aus den Simulationen berechneten  $p$ -Wert gegeben. Vor Beginn der Clusteranalyse wird daher ein Schwellenwert  $p_{\text{max}}$  festgelegt, unterhalb dessen der  $p$ -Wert als statistisch signifikant gelten soll.

Die lokale Autokorrelation  $I_i$  gibt an, wie stark der Wert am Gitterpunkt  $i$  mit der durch die Gewichte  $w_{ij}$  definierten Nachbarschaft korreliert ist. Dabei lassen sich die statistisch signifikanten räumlichen Ausreißer in vier Gruppen einteilen, die in Tabelle 4.3 gezeigt werden.

**Tabelle 4.3:** Abhängig von seiner Nachbarschaft, kann ein Punkt einer der vier Gruppen (HL, HH, LL, LH) zugeordnet werden. Wichtig ist dabei die Unterscheidung zwischen HH als Orte bei denen hohe Werte von hohen Werten umgeben sind, und LL bei denen wiederum niedrige Werte von niedrigen Werten umgeben sind. Bei HL bzw. LH liegt der Punktwert ungewöhnlich stark über bzw. unter denen der Nachbarn.

Mit $p < p_{\text{max}}$	$\sum_{j=1}^N w_{ij} z_j < 1$	$\sum_{j=1}^N w_{ij} z_j > 1$
$z_i > 0$	HL ( $I_i < 0$ )	HH ( $I_i > 0$ )
$z_i < 0$	LL ( $I_i > 0$ )	LH ( $I_i < 0$ )

HH bzw. LL beschreiben Gebiete mit positiven LISA ( $I_i > 0$ ), in denen sich ungewöhnlich hohe bzw. niedrige Schadstoffwerte sammeln und entsprechen damit den gesuchten Hot- bzw. Coldspots. HL- Bereiche werden durch vereinzelte hohe Werte charakterisiert, welche von niedrigen Werten umgeben sind. Analog bezeichnet LH vergleichsweise geringe Schadstoffwerte in einer Nachbarschaft aus hohen Werten. Das beschriebene Verfahren zum Finden möglicher Hotspots und Coldspots in den Schadstoffverteilungen wird nun auf die erhaltenen flächenbezogenen Gruppenmittelwerte aus Abschnitt 4.5.2 angewendet. Die verwendete Gewichtungsmatrix ist zeilenweise normalisiert und die Gewichtungstärke nimmt mit dem Abstand  $d_{ij}$  zweier Gitterpunkte ab:

$$w_{ij} = \begin{cases} 1/d_{ij} & \text{für } i \neq j \text{ benachbart} \\ 0 & \text{für } i = j \end{cases} . \quad (4.10)$$

Für die Berechnung der  $I_i$  wurde die `esda.Moran_Local` Funktion der Python-Bibliothek PySAL-esda [105] genutzt und als Grenzwert für die Signifikanz wurde  $p \leq 0.005$  gewählt. Im Anschluss wird noch die Größe aller gefundenen Hot- und Coldspots ermittelt, wozu jeweils die Anzahl der zusammenhängenden, statistisch signifikanten räumlichen Ausreißer unter Vernachlässigung diagonalen Verbindungen ermittelt wird. Damit lässt sich im Zweifelsfall in der Auswertung die genaue Größe der betroffenen Flächen vergleichen. Zur Verbesserung der visuellen Klarheit wird auf die Darstellung von Hotspots mit weniger als 250 Gitterpunkten ( $\sim 70\text{km}^2$ ) verzichtet.





## 5 Ergebnisse

Im Folgenden wird auf die Resultate des Modelltrainings, sowie auf die Vorhersage der flächenbezogenen Schadstoffe eingegangen. Zudem wird die Lage von Hot- und Coldspots in der Schadstoffverteilung unter verschiedenen saisonalen und meteorologischen Bedingungen dargestellt.

### 5.1 Modellleistung

Tabelle 5.1 und Abbildung 5.1 fassen die Ergebnisse der 10-fachen Kreuzvalidierung für die PM<sub>2.5</sub>, NO<sub>2</sub> und O<sub>3</sub> Modelle zusammen. Die entwickelten Schadstoffmodelle konnten die gemessenen Bodenkonzentrationen von PM<sub>2.5</sub> und NO<sub>2</sub> mit  $R^2$ -Werten von  $R^2 = 0.76$  bzw.  $R^2 = 0.79$  im Test vorhersagen. Der mittlere quadratische Fehler beträgt für die beiden Modelle 32% bzw. 43% des Durchschnitts und der mittlere Bias 0.1  $\mu\text{g}/\text{m}^3$  (0.8%) bzw. 0.19  $\mu\text{g}/\text{m}^3$  (1.1%). Die Validierung des Ozonmodells zeigt mit einem relativen Fehler von ca. 13% und einem Bias von  $-0.04 \mu\text{g}/\text{m}^3$  die beste Übereinstimmung zwischen Vorhersagen und Messungen. Mit einem Bestimmtheitsmaß von  $R^2 = 0.91$  ist die Genauigkeit deutlich höher als die der anderen beiden Modelle.

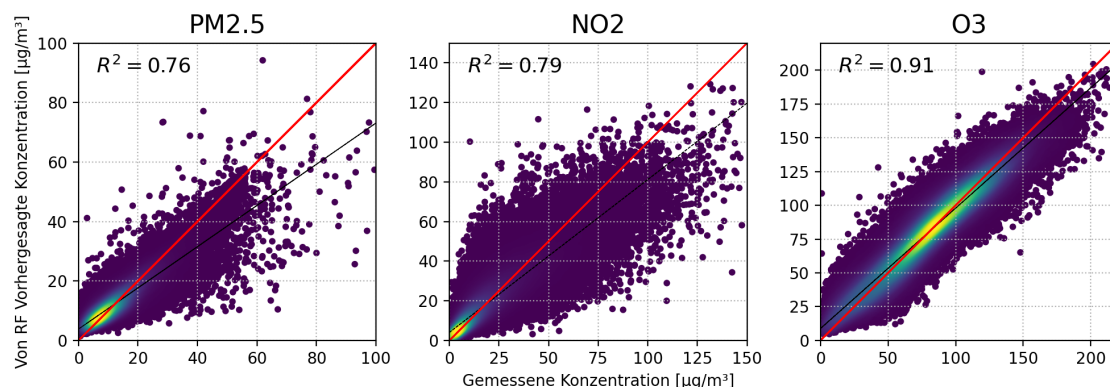
Den Streudiagrammen aus Abb. 5.1 ist die generelle Verteilung der Mess- und Vorhersagedaten zu entnehmen. Die schwarze Ausgleichsgerade kreuzt für alle Modelle die Winkelhalbierende und besitzt eine Steigung  $m < 1$ , was erkenntlich macht, dass alle drei Modelle geringe Schadstoffwerte systematisch über- und hohe Schadstoffwerte systematisch unterschätzen.

**Tabelle 5.1:** Bestimmtheitsmaße, Mittelwerte, quadratischer Fehler, Bias, Standardabweichung und verwendete Datenmenge der PM<sub>2.5</sub>-, NO<sub>2</sub>-, und O<sub>3</sub>-Modelle nach abgeschlossenem Training.

	$R^2$	MEAN [ $\mu\text{g}/\text{m}^3$ ]	RMSE [ $\mu\text{g}/\text{m}^3$ ]	Bias [ $\mu\text{g}/\text{m}^3$ ]	STD [ $\mu\text{g}/\text{m}^3$ ]	Datenmenge
PM <sub>2.5</sub>	0.76	12.15	3.85	0.10	6.22	125100
NO <sub>2</sub>	0.79	17.09	7.38	0.19	14.0	202770
O <sub>3</sub>	0.91	80.87	10.16	-0.04	31.15	128214

Im Anschluss wurde die in Kap. 4.3 beschriebene räumliche und zeitliche Kreuzvalidierung durchgeführt um die Einsetzbarkeit des Modells in unbekannten Regionen bzw. Zeiträumen zu ermitteln. Die entsprechenden Ergebnisse finden sich in Tabelle 5.2. Es zeigt sich, dass durch die zeitliche Trennung der Testdaten von den Trainingsdaten vor allem das PM<sub>2.5</sub>-Modell, mit einer Verringerung von  $R^2 = 0.76$  auf  $R^2 = 0.51$  (*zufällig*) bzw.  $R^2 = 0.35$  (*geordnet*), einen Großteil seiner Vorhersagefähigkeit einbüßt, wohingegen räumliche Kreuzvalidierung die Ergebnisse nur geringfügig verschlechtert. Im Vergleich dazu zeigte das NO<sub>2</sub>-Modell hier eine deutlich geringere Übereinstimmung mit den Messdaten an Orten auf denen es nicht trainiert wurde ( $R^2 = 0.31$  und  $R^2 = 0.36$ ) und verliert dagegen weniger Modellgenauigkeit bei zeitlicher Datentrennung ( $R^2 = 0.72$  bzw.  $R^2 = 0.78$ ).

Die räumliche und zeitliche Kreuzvalidierung des Ozonmodells kann dabei einerseits nur auf die Eingangsvariablen des Ozonmodells angewandt werden, wobei die vorhergesagten



**Abbildung 5.1:** Streudiagramm für die Vorhersagen der Modelle für  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  im Vergleich zu den Stationsmessungen. Die Winkelhalbierende des Quadranten ist in Rot und die Ausgleichsgerade durch die gestreuten Daten in schwarz aufgetragen.

$\text{NO}_2$ -Werte mit der Genauigkeit von  $R^2 = 0.79$  in das Modell eingehen (siehe Tab. 5.2  $\text{O}_3(\text{NO}_2^*)$ ).

Andererseits können diese auf das gesamte gestaffelte Modell angewandt werden, wodurch sich durch die (durch räumliche CV) verschlechterte  $\text{NO}_2$ -Vorhersage, welche in das  $\text{O}_3$ -Modell eingeht, ein zusätzlicher Fehler ergibt (Tab. 5.2  $\text{O}_3(\text{NO}_2)$ ). Die Ozonvorhersage verschlechtert sich jedoch in beiden Fällen nur geringfügig.

**Tabelle 5.2:** Bestimmtheitsmaße der Modelle unter verschiedenen Testverfahren zur zeitlichen und räumlichen Kreuzvalidierung. Die Gruppierung gibt an, dass jeweils die Daten einzelner Tage, Stationen, Regionen oder 6 Monaten stets gemeinsam in den Test- bzw. Traininsdaten vorkommen. Das Modell  $\text{O}_3(\text{NO}_2^*)$  wurde mit  $\text{NO}_2$ -Stationsdaten trainiert, während  $\text{O}_3(\text{NO}_2)$  mit den durch das  $\text{NO}_2$ -Modell vorhergesagten Eingangsdaten trainiert und getestet wurde.

Aufteilung: Gruppierung:	Bestimmtheitsmaß ( $R^2$ )				
	zufällig keine	zeitlich		räumlich	
		Tage (CV10)	6 Monate	Stationen (CV10)	Regionen
$\text{PM}_{2.5}$	0.76	0.51	0.35	0.73	0.61
$\text{NO}_2$	0.79	0.78	0.72	0.31	0.36
$\text{O}_3(\text{NO}_2)$	0.91	0.86	0.82	0.90	0.87
$\text{O}_3(\text{NO}_2^*)$	0.91	0.86	0.81	0.88	0.85

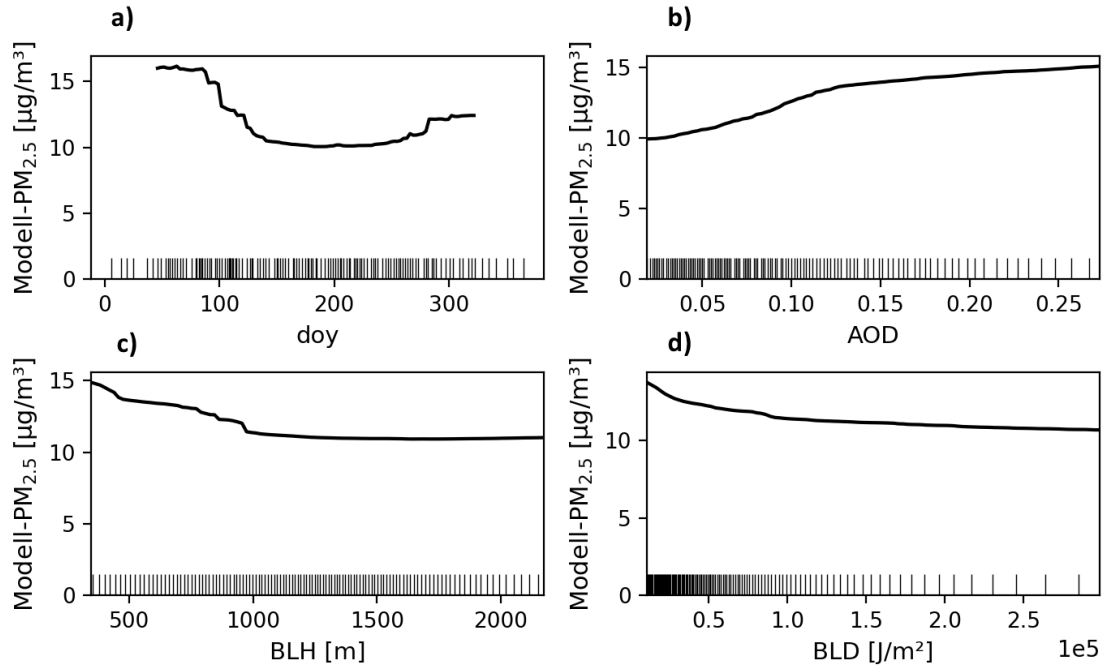
## 5.2 Modell- und Variablenabhängigkeiten

Die Genauigkeiten der Modelle für die einzelnen Schadstoffe reagieren unterschiedlich auf die Verfügbarkeit bestimmter Merkmale. Die Ergebnisse der Feature Importance Analyse für die verwendeten Variablen sind in Abb. 5.2 dargestellt.

### $\text{PM}_{2.5}$

Die Feature Importance Werte des  $\text{PM}_{2.5}$ -Modells in Abb. 5.2 a) zeigen, dass die Rolle des wichtigsten Merkmals auf eine nicht-physikalischen Hilfsvariable, den Jahrestag (DoY), fällt, der eine mittlere Abnahme der Modellgenauigkeit (MDA) von über 50% aufweist. An zweiter Stelle ( $\sim 30\%$ ) steht die Aerosol Optische Dicke (AOD) gefolgt von einer zeitlichen Hilfsvariable (Month). Ebenfalls bedeutsam für die Vorhersage sind Parameter zu meteorologischen Rahmenbedingungen wie Grenzschichthöhe (BLH), und -dissipation





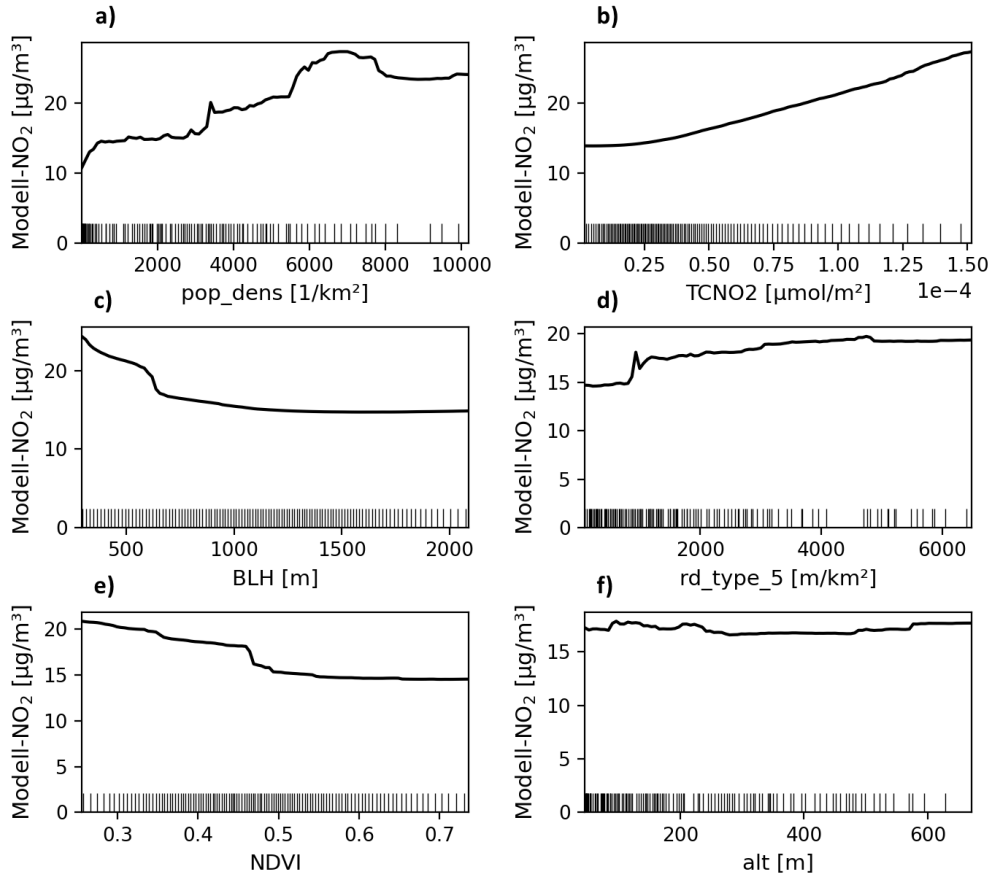
**Abbildung 5.3:** Partielle Abhängigkeiten des Modells zur  $PM_{2.5}$ -Vorhersage von den Variablen a) Jahrestag, b) Aerosol Optische Dichte, c) Grenzschichthöhe und d) Grenzschicht-Energiedissipation.

Auf Basis der FI werden die PDPs des  $NO_2$ -Modells für die sechs wichtigsten Variablen `pop_dens`, `TCNO2`, `BLH`, `rd_type_5`, `NDVI` und `alt` ermittelt und in den Abbildungen 5.4 a)-f) dargestellt. Es zeigen sich generell höhere  $NO_2$ -Konzentrationen bei höherer Bevölkerungsdichte, allerdings kann diese Beziehung für sehr dicht besiedelte Gebiete umgekehrt werden, wie in Abb. 5.4 a) zu sehen ist. Auch `TCNO2` und die Dichte kleiner, lokaler Straßen (`rd_type_5`) sind positiv mit der Zielvariable korreliert; hohe Werte für `NDVI` und `BLH` führen dagegen zu einem Absinken der vorhergesagten  $NO_2$ -Konzentration. Für die geographische Höhe (`alt`) zeigt sich im PDP (Abb. 5.4) kein eindeutiger Zusammenhang zum Vorhersagewert.

### $O_3$

Abbildung 5.2 c) ist zu entnehmen, dass für die bodennahen Ozonkonzentration die meteorologischen Variablen für die Temperatur in 2 m Höhe (`T2m`) mit über 30% MDA und SSR mit 19% die wichtigsten Vorhersageparameter darstellen. Darauf Folgt der Jahrestag (`DoY`) mit 13% MDA. Die Bodenkonzentration von Stickstoffdioxid, vorhergesagt durch das  $NO_2$ -Modell, ist eine wichtige Vorläufersubstanz von Ozon, in der Liste der wichtigsten Merkmale aber lediglich an vierter Stelle mit etwa 11% MDA. Weitere Variablen mit mittlerer Wichtigkeit sind hauptsächlich meteorologische Merkmale wie die `RH`, `LST`, `BLD`, `TCWV`, etc. Raumbezogene Merkmale wie die Bevölkerungsdichte, die geographische Höhe oder die Straßentypendichten besitzen entweder besonders geringe FI-Werte oder wurden bereits bei der Auswahl der Merkmale (Kap. 4.2) ausgeschlossen. Interessant ist zudem, dass der Ozongehalt der Atmosphärensäule (`TCO3`) ebenfalls unwichtig für die Vorhersage ist.

Die partielle Abhängigkeit des Modells von der Temperatur (`T2m`), der Sonneneinstrahlung (`SSR`), den Vorhergesagten  $NO_2$ -Werten (`NO2_pred`), der `DoY`-Variable, der Grenzschichthöhe (`BLH`) sowie der relativen Luftfeuchtigkeit (`RH`) wird berechnet und in Abbildung 5.5 a)-f) dargestellt. `T2m` und `SSR` sind jeweils positiv mit den Modellvor-

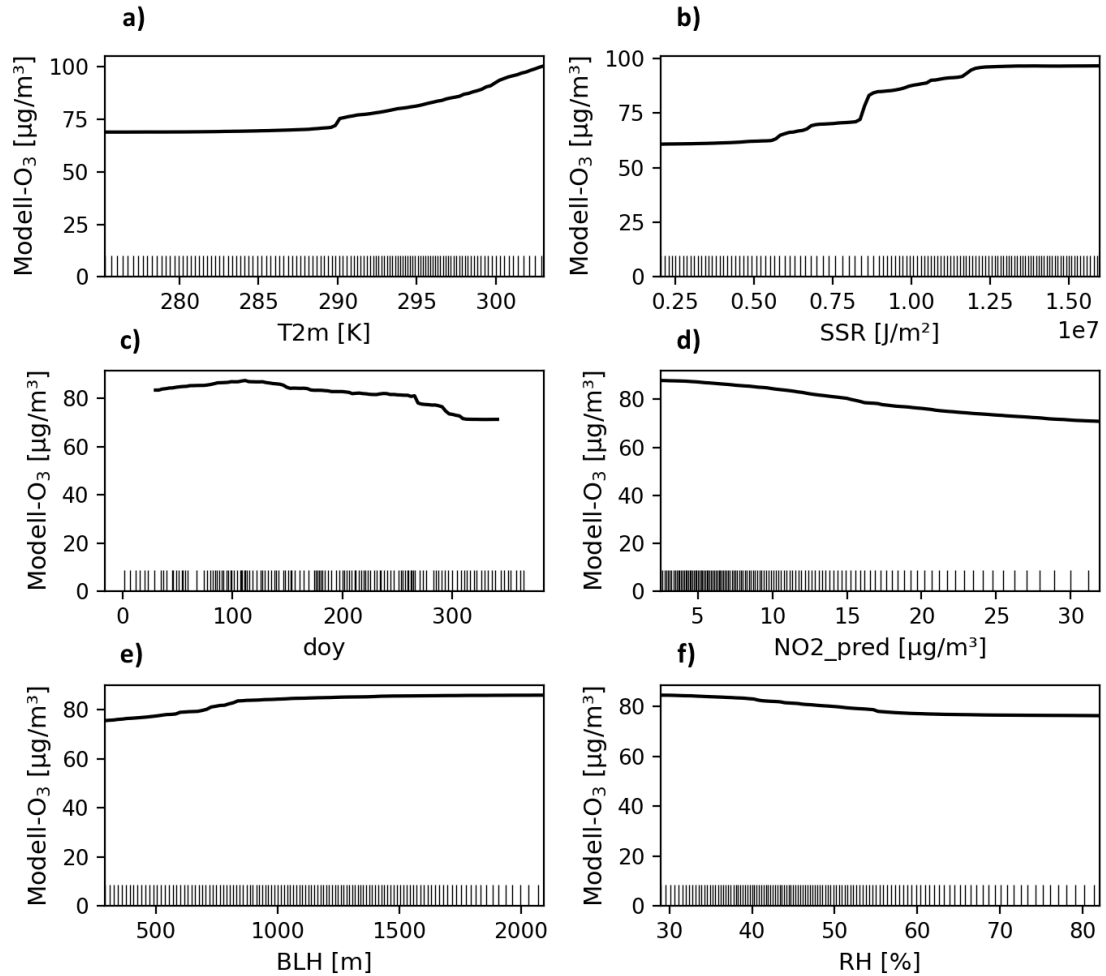


**Abbildung 5.4:** Partielle Abhängigkeiten des Modells zur NO<sub>2</sub>-Vorhersage von den Variablen a) Bevölkerungsdichte, b) NO<sub>2</sub>-Gehalt der troposphärischen Luftsäule c) Höhe der Grenzschicht, d) Dichte des Straßentyps 5 (lokale Straßen), e) Vegetationsindex und f) geographische Höhe.

hersagen korreliert, d.h. hohe Temperaturen und hohe Sonneneinstrahlung sind im Mittel für hohe Ozonkonzentrationen verantwortlich. Eine Erhöhung der NO<sub>2</sub>-Werte hingegen sorgt für eine Abnahme des bodennahen Ozons wie in Abb. 5.5 d) erkennbar ist. Ähnliches gilt für die Beziehung zwischen relativer Luftfeuchtigkeit und Ozon, allerdings ist der absenkende Einfluss von RH auf vorhergesagtes O<sub>3</sub> relativ gering. BLH zeigt eine schwache positive Korellation mit Ozon. Der O<sub>3</sub>-Jahresgang, mit einem maximum in der frühen Mitte des Jahres, ist ebenfalls erkennbar (vgl. Abb. 5.5 c)), allerdings fällt dieser wesentlich schwächer aus als in den Stationsmessungen in Abb. 5.15a erkennbar ist.

### Einfluss der Variablenkorrelation auf FI und PDP

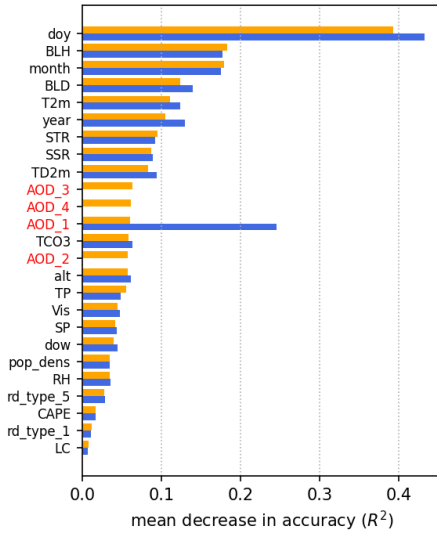
Für die Interpretation der Feature Importance gilt, dass die statistische Abhängigkeit der Merkmale untereinander die berechneten MDA-Werte beeinflussen kann. Der Extremfall kann durch das duplizieren eines Eingangsfeatures in einem Modell für PM<sub>2.5</sub> (mit allen verfügbaren Merkmalen) veranschaulicht werden. Die AOD-Variable wird dazu einem sonst identischen Modell mehrmals als Eingangsvariable übergeben. Anschließend wird erneut die Feature Importance ermittelt. Es ist zu erkennen, dass die in Abb. 5.6 in blau gezeigten ursprünglichen MDA-Werten der Variable AOD\_1, bei hinzufügen der identischen Merkmalen AOD\_2, AOD\_3 und AOD\_4 in das Modell, auf diese verteilt werden. Insofern sollte die FI für eine korrelierte Eingangsvariable nie im Vakuum betrachtet werden und stets zusammen mit Blick auf die anderen Modellvariablen interpretiert werden. Dazu wurden die



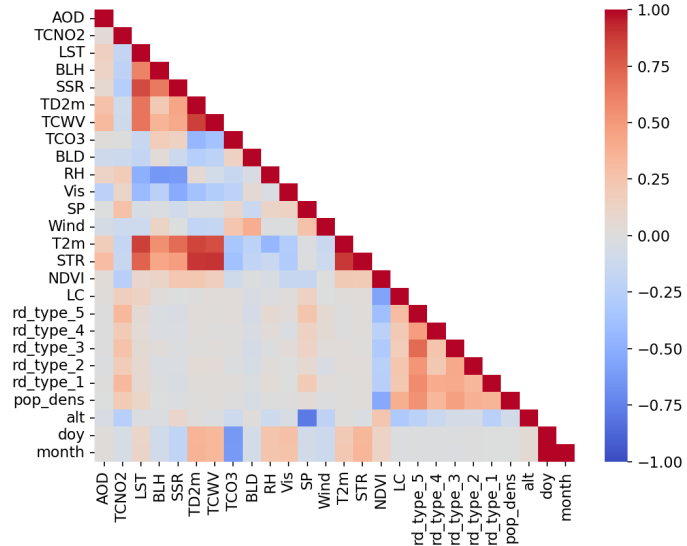
**Abbildung 5.5:** Partielle Abhängigkeiten des Modells zur O<sub>3</sub>-Vorhersage von den Variablen a) Temperatur in 2 m Höhe, b) Stärke der Sonneneinstrahlung auf die Erdoberfläche, c) Jahrestag, d) vom NO<sub>2</sub>-Modell vorhergesagte NO<sub>2</sub>-Bodenkonzentration, e) Grenzschichthöhe und f) relative Luftfeuchtigkeit.

Pearson-Korrelationskoeffizienten zwischen den einzelnen Eingangsvariablen bestimmt und in Abb. 5.7 dargestellt. Starke lineare Korrelationen finden sich z.B. zwischen Variablen die mit der eintreffenden Strahlungsleistung zusammenhängen. Temperaturparameter (T2m, TD2m und LST) sind klar von der Stärke der Sonneneinstrahlung (SST, STR) abhängig und beeinflussen wiederum den Wassergehalt der Atmosphäre (TCWV, RH) weshalb die beteiligten Variablen stark korreliert sind. Dazu kommen offensichtliche Korrelationen, wie zwischen DoY und Month, zwischen geographischer Höhe (alt) und Luftdruck (SP) und zwischen Bevölkerungsdichte und Straßendichte, besonders für kleine Straßen. Zu beachten ist bei den obigen Betrachtungen, dass sich unter den Eingangsvariablen mit Sicherheit auch nichtlinear korrelierte Merkmale befinden, welche durch die Berechnung der linearen Korrelationen nicht erkenntlich werden, aber dennoch Einfluss auf die FI nehmen können. Dies ist unter anderem zwischen DoY bzw. Month und Variablen mit stark ausgeprägtem Jahresgang zu vermuten, bei denen sich ein maximum oder minimum in der Mitte des Jahres zeigt (z.B. SSR, T2m, usw.).

Bei der Interpretation der PDP muss ebenfalls auf Korrelationen zwischen den Variablen geachtet werden, da die Berechnung der partiellen Abhängigkeiten von der vollständigen Unabhängigkeit der Merkmale ausgeht. Um den Einfluss dessen zu prüfen, wurden Modelle



**Abbildung 5.6:** Vergleich der Feature Importance eines Modells mit verschiedenen Eingangsvariablen mit einem Modell mit mehreren identischen Eingangsvariablen (AOD\_1-4). Die Information der einzelnen Merkmale ist redundant, wodurch sich die MDA-Wertung auf diese aufteilt.

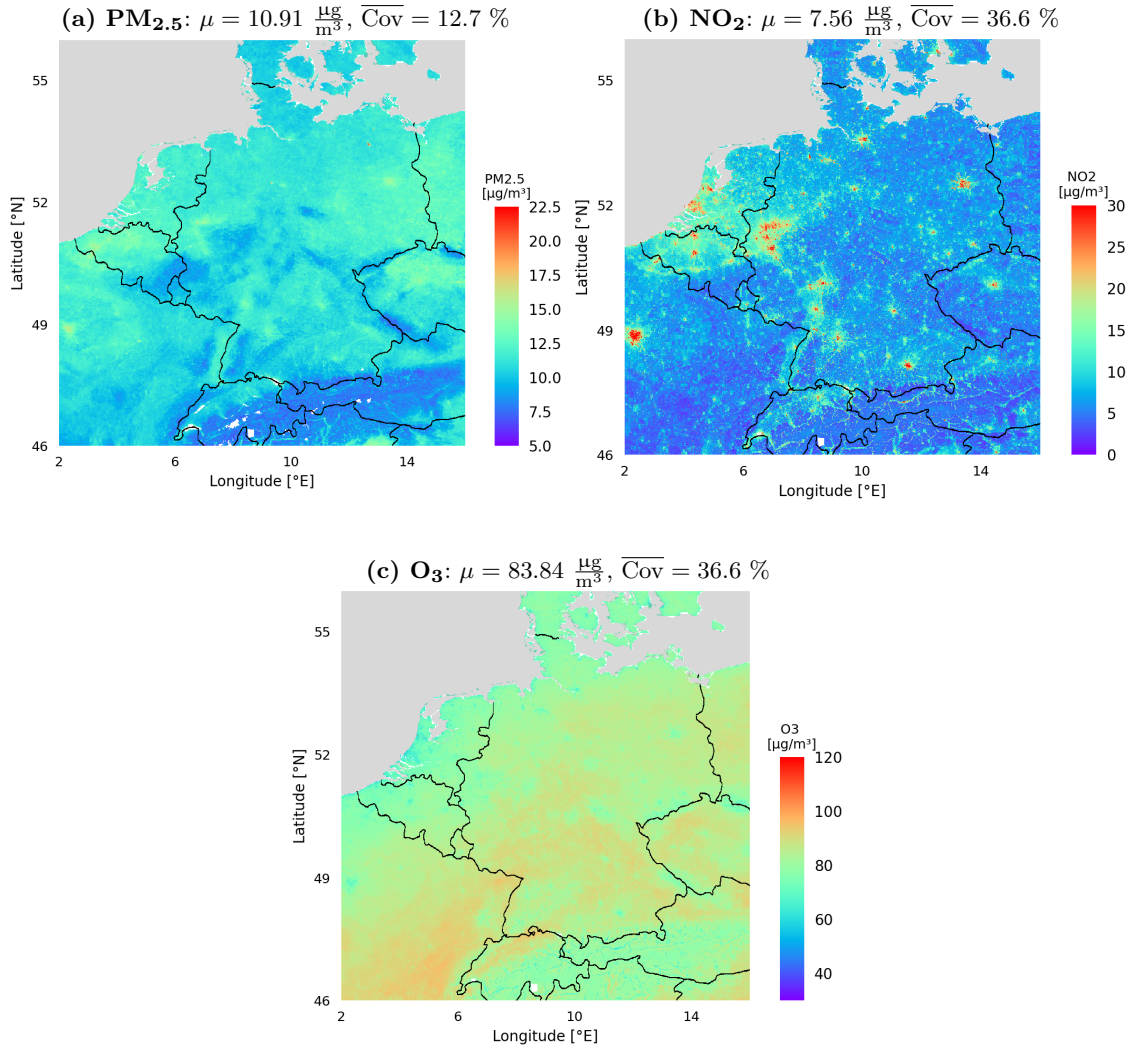


**Abbildung 5.7:** Pearson Korrelation zwischen einigen Eingangsvariablen. Es werden nur Variablen angezeigt, für welche bedeutende Korrelationen bestehen.

trainiert, in denen für jedes Merkmal alle in Abb. 5.7 erkenntlich stark damit korrelierten Variablen aus dem Modell entfernt wurden, auch wenn dadurch ein geringfügig größerer Modellfehler entsteht. Anhand der neuen Modelle konnte festgestellt werden, dass die partiellen Abhängigkeiten der um korrelierte Merkmale reduzierten Modelle mit denen der ursprünglichen Modelle sehr gut übereinstimmen. Die PDP wurde daher ungeachtet der starken Korrelationen dazu genutzt, die wichtigsten physikalischen Zusammenhänge zwischen Eingangsvariablen und Zielvariablen der  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Modelle zu untersuchen.

### 5.3 Räumliche Verteilung und zeitliche Entwicklung

Aufgrund der eingeschränkten Verfügbarkeit von Satellitendaten sind die täglichen Flächendaten teilweise sehr lückenhaft. Die Betrachtung von (Jahres-) Mittelwerten ist deshalb sinnvoll. Die absolute Abdeckung unterscheidet sich aufgrund der unterschiedlichen Methoden zur Datengewinnung zwischen  $\text{PM}_{2.5}$  mit 12.7% und  $\text{NO}_2$  bzw.  $\text{O}_3$  mit 36.6% deutlich. Die Abbildungen 5.8 a)-c) stellen die Mittelwerte der vorhergesagten Schadstoffkonzentrationen über den jeweiligen gesamten Untersuchungszeitraum dar und Abb. 5.10 zeigt Histogramme der Werte. Zusätzlich wird die in Kapitel 4.5.4 beschriebene Analyse zur Identifikation von Hot- und Coldspots an den Daten angewandt (siehe Abb. 5.9). Zur besseren Orientierung sind (nur) in diesen Karten einige der größten Städte des Untersuchungsgebiets markiert. Um zu erkennen, ob sich die Schadstoffverteilung während des Messzeitraums verändert, werden außerdem die jeweiligen Jahresmittelwerte gebildet. Um die mittlere Belastung der Bevölkerung durch  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  abschätzen zu können, muss die ungleiche Verteilung der Bevölkerung auf das Gesamtgebiet berücksichtigt werden. Dazu werden, wie in Gleichung 5.1 gezeigt, die flächenbezogenen Mittelwerte des



**Abbildung 5.8:** Darstellung der über den gesamten Untersuchungszeitraum (PM<sub>2.5</sub>: 2018 bis 2022, NO<sub>2</sub> und O<sub>3</sub>: 2018 bis 2020) gemittelten PM<sub>2.5</sub>-, NO<sub>2</sub> und O<sub>3</sub>-Modellvorhersagen für die Schadstoffbodenkonzentration. Es ist jeweils der Mittelwert über das Gesamtgebiet ( $\mu$ ) und die mittlere Abdeckung der enthaltenen Tage ( $\overline{\text{Cov}}$ ) angegeben

Schadstoffs  $X$  über den Untersuchungszeitraum mit der Bevölkerungsdichte multipliziert, über die Gesamtfläche summiert und anschließend durch die Gesamtbevölkerung geteilt:

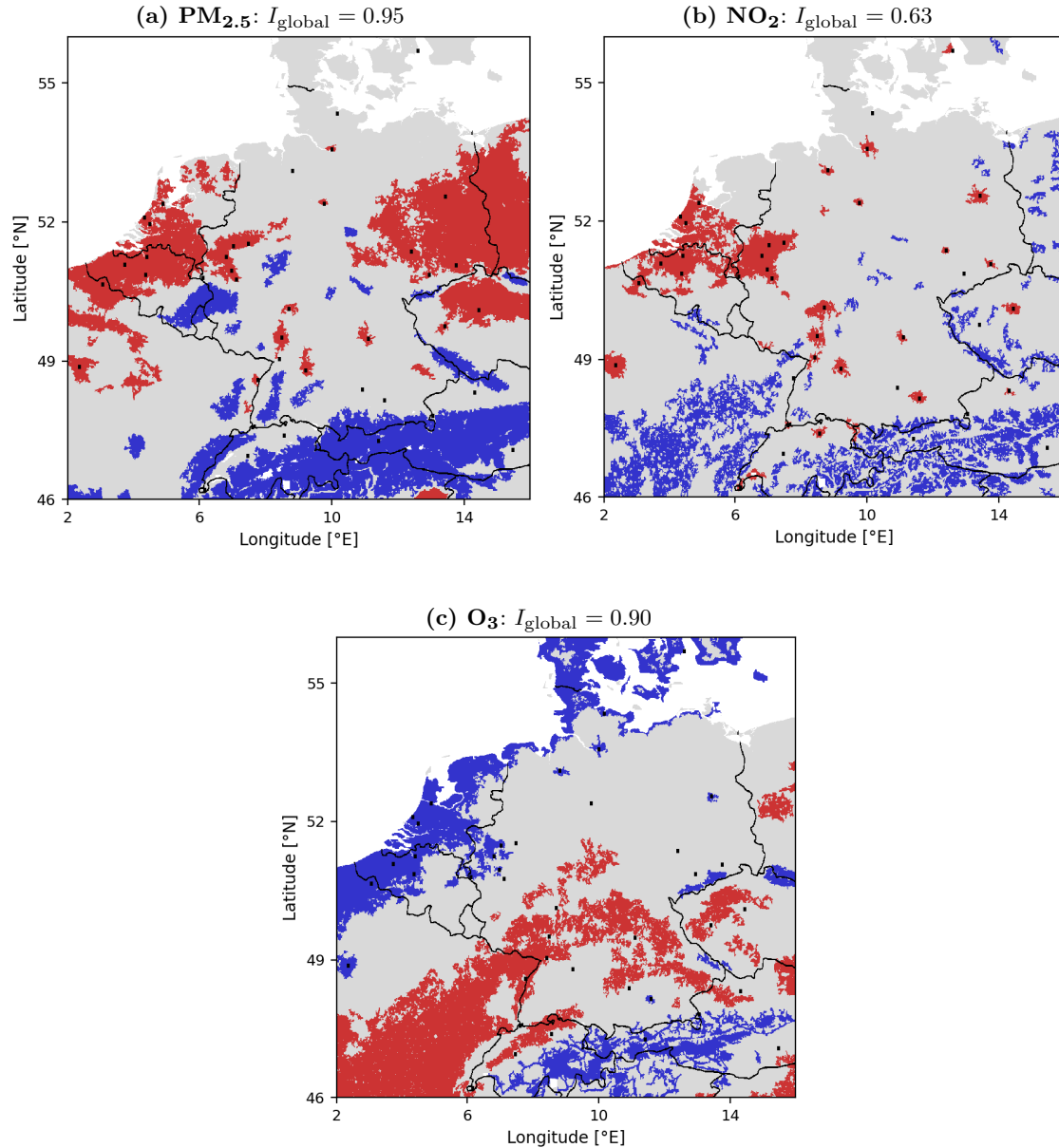
$$\overline{EXP}_X = \frac{\sum_{i,j} (\text{pop\_dens})_{i,j} \times \overline{X}_{i,j}}{\sum_{i,j} (\text{pop\_dens})_{i,j}}. \quad (5.1)$$

Dabei ist  $\overline{EXP}_X$  die mittlere Belastung der Bevölkerung durch den Schadstoff  $X$ ,  $(\text{pop\_dens})_{i,j}$  die Bevölkerungsdichte und  $\overline{X}_{i,j}$  der zeitliche Mittelwert jeweils am Gitterpunkt  $(i,j)$ .

**Tabelle 5.3:** Jahresmittelwerte der vorhergesagten Schadstoffkonzentrationen über das Gesamte Untersuchungsgebiet.

Jahr:	2018	2019	2020	2021	2022
PM <sub>2.5</sub> [ $\mu\text{g}/\text{m}^3$ ]	11.95	11.20	10.35	10.42	10.53
NO <sub>2</sub> [ $\mu\text{g}/\text{m}^3$ ]	7.53	8.15	7.00	-	-
O <sub>3</sub> [ $\mu\text{g}/\text{m}^3$ ]	88.13	81.48	81.86	-	-

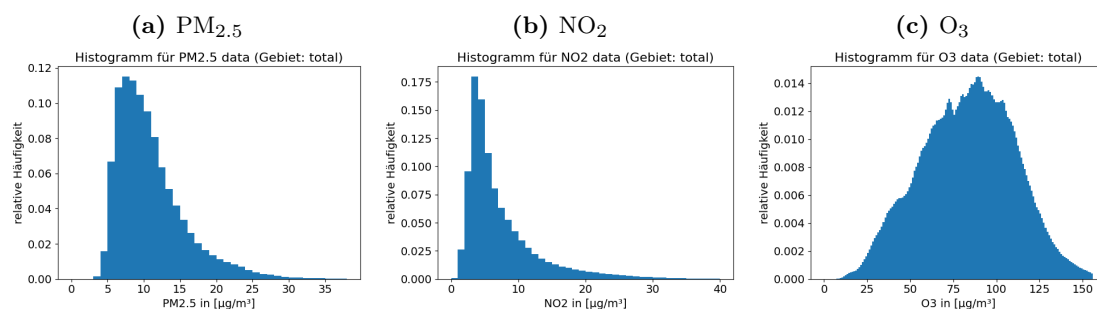




**Abbildung 5.9:** LISA in den Gesamtmittelwerten der  $\text{PM}_{2.5}$ -,  $\text{NO}_2$  und  $\text{O}_3$ -Modellvorhersagen. Rote Flächen zeigen Hotspots (HH) auf und blaue Flächen signalisieren Coldspots (LL), wobei nur Hot- bzw. Coldspots mit einer zusammenhängenden Fläche von mehr als 250 Gitterpunkten ( $\sim 70\text{km}^2$ ) gezeigt werden.  $I_{\text{global}}$  gibt die Stärke der globalen Autokorrelation an.

### $\text{PM}_{2.5}$

Die vorhergesagten  $\text{PM}_{2.5}$ -Mittelwerte (Abb. 5.8a) betragen im gesamten Untersuchungsgebiet im Durchschnitt  $10.91 \mu\text{g}/\text{m}^3$ , variieren jedoch mäßig mit der geographischen Lage zwischen dem Minimum von  $5.21 \mu\text{g}/\text{m}^3$  und dem Maximum von  $26.89 \mu\text{g}/\text{m}^3$ . Die Feinstaubwerte fallen in den Alpen generell deutlich geringer aus als in flachen Regionen, mit Ausnahme großer und gut erschlossener Täler (z.B. Inntal oder Etschtal) für die vergleichsweise hohe Werte bestehen bleiben. Zudem lassen sich in der unmittelbaren Nähe größerer europäischer Ballungsräume wie Paris, Berlin, Köln, Düsseldorf, aber auch im industriellen und bevölkerungsreichen Ruhrgebiet erhöhte  $\text{PM}_{2.5}$ -Werte erkennen. Die LISA-Analyse in Abb. 5.9a liefert neben den bereits genannten Gebieten großflächige Hotspots zwischen den



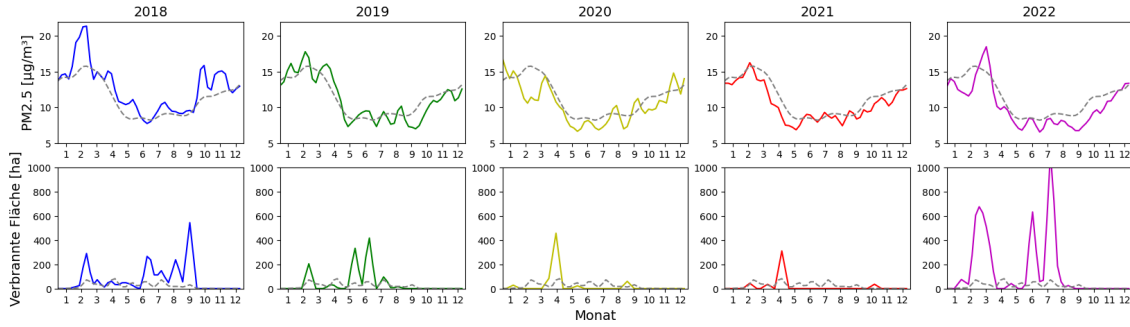
**Abbildung 5.10:** Histogramme der Verteilung der Flächendaten auf dem Gesamtgebiet für a)  $\text{PM}_{2.5}$ , b)  $\text{NO}_2$  und c)  $\text{O}_3$ .

besonders Bevölkerungsreichen Teilen der Niederlande (Randstad) und Brüssel, sowie für weite Teile Ostdeutschlands, Polens und Tschechiens. Naturbelassene Gebiete wie Harz, Schwarzwald, Bayerischer Wald, usw. sind dagegen klar als Regionen mit deutlich geringerer Belastung zu erkennen. Für die Verteilung von  $\text{PM}_{2.5}$  ergibt sich eine mittlere Belastung der Bevölkerung von  $11.96 \mu\text{g}/\text{m}^3$ .

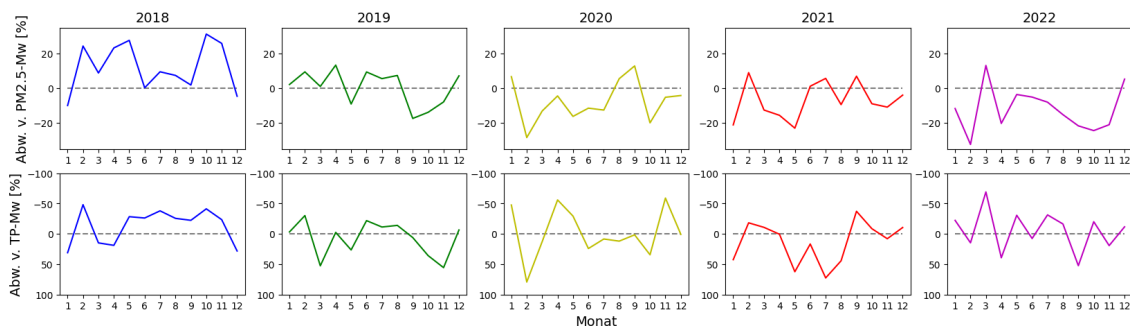
Für  $\text{PM}_{2.5}$  wurden Daten für fünf Jahre betrachtet, deren Mittelwerte in Tab. 5.3 und flächenhafte Verteilungen in Abbildung A.1 zu finden sind. 2018 stellt sich ein Jahresmittelwert von  $11.95 \mu\text{g}/\text{m}^3$  ein, welcher im Jahr 2019 um  $0.75 \mu\text{g}/\text{m}^3$  (auf  $11.20 \mu\text{g}/\text{m}^3$ ) und 2020 erneut um  $0.85 \mu\text{g}/\text{m}^3$  (auf  $10.35 \mu\text{g}/\text{m}^3$ ) abnimmt. Die Folgejahre erfahren dagegen einen geringen Anstieg des  $\text{PM}_{2.5}$ -Mittels auf  $10.42 \mu\text{g}/\text{m}^3$  (2021) und  $10.53 \mu\text{g}/\text{m}^3$  (2022). Der anfängliche Rückgang von  $\text{PM}_{2.5}$  ist in den Flächendaten sehr deutlich zu erkennen und dabei nicht nur in Ballungsräumen, sondern Großflächig ausgeprägt. In Norddeutschland und den Niederlanden liegt laut den Karten in Abb. A.1 eine deutlich geringere Grundbelastung für die letzten Jahre 2020 bis 2022 im Vergleich zu 2018 und 2019 vor, wobei die stärkste Verbesserung der Luftqualität im Jahre 2020 erkennbar ist.

Abbildung 5.11 zeigt den Verlauf des  $\text{PM}_{2.5}$ -Mittelwertes zusammen mit Daten zu Brandaktivitäten in Zentraleuropa. Letztere sind Daten aus Betrachtungen von MODIS-Satellitendaten des EFFIS [106]. Sie enthalten die pro Woche verbrannte Fläche aller überwiegend innerhalb des Untersuchungsgebiets liegenden Länder (Deutschland, Tschechien, Schweiz, Österreich, Belgien und Niederlande). Die  $\text{PM}_{2.5}$ -Daten wurden zur Unterdrückung des Wochengangs ebenfalls als Wochenmittelwerte aufgenommen. Es zeigen sich merkliche Brandereignisse zwischen Februar und März in 2018, 2019 und 2022, sowie im Juni und Juli 2022. Weitere Brandflächen konnten von Juni bis September 2018, im Mai und Juni 2019 und im April der Jahre 2020 und 2021 beobachtet werden. Die Feinstaubwerte schwanken dabei überwiegend zufällig und relativ geringfügig, um den fünfjährigen Mittelwert. Auffällig sind jedoch die besonders hohen  $\text{PM}_{2.5}$ -Werte im Januar/Februar 2018 sowie die besonders geringe Belastung im Februar/März 2020. Etwas schwächer, aber merklich unter dem Mittel, liegen auch die  $\text{PM}_{2.5}$ -Werte von September bis November 2019 und Februar bis Mai 2021.

Des Weiteren wurden in Abb. 5.12 jeweils die Verläufe der Monatsmittelwerte des vom  $\text{PM}_{2.5}$ -Modell vorhergesagten Feinstaubs und der Niederschlagsmenge im Untersuchungsgebiet als Abweichungen vom fünfjährigen Mittelwert aufgetragen. Neben den in Abb. 5.11 beobachteten  $\text{PM}_{2.5}$ -Schwankungen zeigen sich für den Niederschlag besonders geringe Werte im Februar, Juli und Oktober 2018, im Januar und April 2020, sowie im März 2022. Starke Niederschläge können im März, Oktober und November 2019, Februar 2020, Mai und Juli 2021 und im September 2022 beobachtet werden.



**Abbildung 5.11:** Mit einem Binomialfilter ( $1/4 [1, 2, 1]$ ) geglättete Wochenmittelwerte für PM<sub>2.5</sub> und die verbrannte Fläche in den Ländern Deutschland, Tschechien, Schweiz, Österreich, Belgien und Niederlande. Die grauen Linien geben für PM<sub>2.5</sub> den Mittelwert zwischen 2018 und 2022 und für die verbrannte Fläche den Mittelwert zwischen 2006 und 2024 an.



**Abbildung 5.12:** Der mittlere Niederschlag und Monatsmittelwerte für PM<sub>2.5</sub> auf dem Untersuchungsgebiet für die Jahre 2018 bis 2022. Die Daten sind aufgetragen als Abweichung vom zwischen 2018 und 2022 gemessenen Mittelwert. Die Niederschlagswerte wurden zur besseren Interpretierbarkeit invertiert aufgetragen.

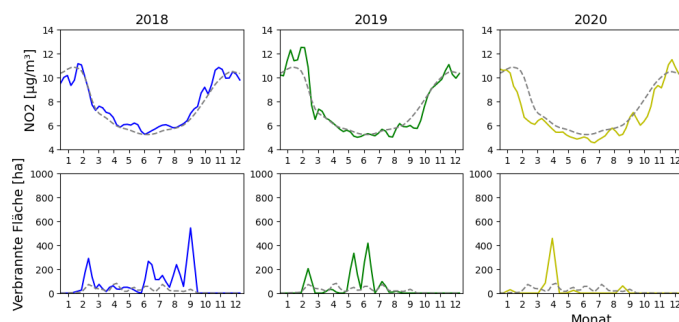
## NO<sub>2</sub>

Der Mittelwert der bodennahen Stickstoffdioxidbelastung liegt zwischen Anfang 2018 und Ende 2020 bei  $7.56 \mu\text{g}/\text{m}^3$ . Der Abb. 5.8b ist klar zu entnehmen, dass NO<sub>2</sub> im Vergleich zu PM<sub>2.5</sub> oder Ozon wesentlich inhomogener verteilt ist, wobei nahezu jede größere Stadt mit deutlich höheren Schadstoffwerten als in den ländlichen Regionen auffällt. In Abb. 5.9b sind dementsprechende Hotspots zu erkennen. Die Inhomogenität in der NO<sub>2</sub>-Verteilung zeigt sich sowohl an der schmalen und asymmetrischen Werteverteilung in Abb. 5.10b, als auch an der Spanne der Extremwerte, welche mit maximalen Stickstoffdioxidkonzentrationen von  $71.17 \mu\text{g}/\text{m}^3$  und Minima von  $1.62 \mu\text{g}/\text{m}^3$  die größte unter den untersuchten Schadstoffen darstellt. Wie schon bei PM<sub>2.5</sub> finden sich in den Alpen und vielen naturbelassenen Regionen vergleichsweise niedrige NO<sub>2</sub>-Werte, wobei sich Gebirgstäler durch höhere Werte deutlicher von ihrer Umgebung abheben. Generell sind die Coldspots in Abb. 5.9b zahlreicher und gleichmäßiger über das Untersuchungsgebiet verteilt als bei den andern Schadstoffen, fallen allerdings auch wesentlich kleiner aus, wodurch viele unter die gesetzte Mindestgröße von  $\sim 70 \text{ km}^2$  fallen. Die am stärksten verschmutzten Städte sind Paris, Hamburg, München, Frankfurt am Main, Berlin und Prag, wobei die bedeutendste großflächige Verschmutzung auf den Westen der Niederlande, Nordbelgien und das Ruhgebiet begrenzt sind. Im Durchschnitt ist jeder Einwohner im Gesamtgebiet über den Untersuchungszeitraum einer NO<sub>2</sub>-Konzentration von  $16.77 \mu\text{g}/\text{m}^3$  ausgesetzt.

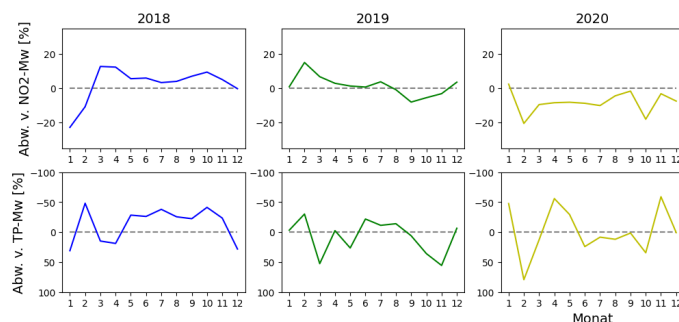
Die zeitliche Entwicklung ist anhand von Tab. 5.3 und der Abbildungen A.2 an den entsprechenden Jahresmittelwerten zu erkennen. Analog zu PM<sub>2.5</sub> ist ein allgemeines

Absinken der  $\text{NO}_2$ -Konzentration im Jahr 2020 ( $7.0 \mu\text{g}/\text{m}^3$ ) im Vergleich zu den Vorjahren (2018:  $7.53 \mu\text{g}/\text{m}^3$ , 2019:  $8.15 \mu\text{g}/\text{m}^3$ ) zu vermerken. Davon abgesehen ähneln sich die Schadstoffkarten der drei Jahre in ihrer räumlichen Verteilung stark.

Analog zu  $\text{PM}_{2.5}$  sind in den Abbildungen 5.13 und 5.14 die wöchentlichen bzw. monatlichen  $\text{NO}_2$ -Werte zwischen 2018 und 2020 gegen die verbrannte Fläche und den Neiderschlag aufgetragen. Der Verlauf der  $\text{NO}_2$ -Konzentration zeigt vereinzelte Erhöhungen zwischen März und November 2018, sowie einen deutlichen Anstieg im Januar und Februar 2019. Werte unter dem Dreijahresdurchschnitt finden sich im Januar 2018, zwischen August und Oktober in 2019, sowie besonders im Februar 2020 mit ebenfalls leicht erhöhten Niederschlagswerten für den Rest des Jahres.



**Abbildung 5.13:** Mit einem Binomialfilter ( $1/4[1, 2, 1]$ ) geglättete Wochenmittelwerte für  $\text{NO}_2$  und die verbrannte Fläche in den Ländern Deutschland, Tschechien, Schweiz, Österreich, Belgien und Niederlande. Die grauen Linien geben für  $\text{NO}_2$  den Mittelwert zwischen 2018 und 2020 und für die verbrannte Fläche den Mittelwert zwischen 2006 und 2024 an.



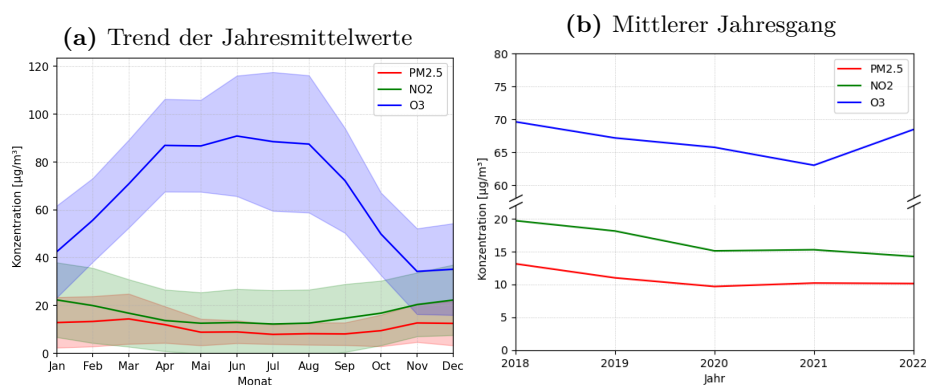
**Abbildung 5.14:** Der mittlere Niederschlag und Monatsmittelwerte für  $\text{NO}_2$  auf dem Untersuchungsgebiet für die Jahre 2018 bis 2020. Die  $\text{NO}_2$ -Daten sind aufgetragen als Abweichung vom zwischen 2018 und 2020 gemessenen Mittelwert. Die Niederschlagswerte beziehen sich auf den Mittelwert zwischen 2018 und 2022 und wurden zur besseren Interpretierbarkeit invertiert aufgetragen.

## $\text{O}_3$

Ozon ist von den betrachteten Schadstoffen am gleichmäßigsten im Untersuchungsgebiet verteilt (siehe Abb. 5.8c und Abb. 5.10c) und liegt im Mittel mit einer Konzentration von  $83.84 \mu\text{g}/\text{m}^3$  vor. Die Extremwerte belaufen sich auf  $96.69 \mu\text{g}/\text{m}^3$  und  $54.73 \mu\text{g}/\text{m}^3$ , womit

der Maximalwert dichter am Mittelwert liegt als das Minimum. Südlich gelegene Gebiete sind mit Ausnahme der Alpen nach Abbildung 5.8c und Abb. 5.9c stärker belastet als der Norden, wobei sich in den Küstenregionen der Nordsee die geringsten Werte und größten Coldspots befinden. Im Inland sind Coldspots erneut in den Alpen und vegetationsdichten Gebieten zu finden, insbesondere im Bayerischen Wald und im Riesengebirge (Polen/Tschechien) (siehe Abb. 5.9c), aber auch generell verringerte Werte im Schwarzwald, Thüringer Wald u. Thüringer Schiefergebirge (siehe Abb. 5.8c). Abweichend von  $\text{PM}_{2.5}$  und  $\text{NO}_2$  zeigen sich für die dort erwähnten Ballungsräume, also Großstädte und das Ruhrgebiet, sowie für Täler in den Alpen wesentliche lokale Absenkungen der Ozonbelastung. Speziell in Paris, Bremen, Hamburg, Berlin und München, sowie über dem Ruhrgebiet lassen sich Coldspots erkennen, während sich Hotspots großflächig über den Südwesten Frankreichs und Mitteldeutschland erstrecken. Ebenso kann im Nordböhmisches Becken ein Hotspot erkannt werden. Die mittlere Belastung der Bevölkerung im Untersuchungsgebiet beträgt  $77.73 \mu\text{g}/\text{m}^3$ .

Die Ozonmittelwerte für 2018 liegen mit  $88.13 \mu\text{g}/\text{m}^3$  deutlich über den Mittelwerten der anderen Jahre, welche  $81.48 \mu\text{g}/\text{m}^3$  für 2019 und  $81.86 \mu\text{g}/\text{m}^3$  für 2020 betragen (siehe Tab. 5.3 und Abb. A.3). Ebenfalls auffällig in Abb. A.3b ist die vergrößerte Niedrigozonfläche an der Nordseeküste der Niederlande im Jahr 2019.



**Abbildung 5.15:** Schadstoffwerte der  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Stationsmessungen: a) Jahresgang der Messdaten als Monatsmittelwerte. Die eingefärbten Flächen beschreiben die Standardabweichung innerhalb des Monats. b) Entwicklung der Jahresmittelwerte der Stationsdaten von 2018 bis 2022.

## 5.4 Saisonale Schadstoffverteilung

Angesichts des merklichen Jahresganges der Messwerte für  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  (vgl. Abb. 5.15a) wird eine Separierung der flächenbezogenen Schadstoffwerte nach den Jahreszeiten durchgeführt. Die räumliche Verteilung ist in Abb. 5.16 dargestellt, Mittel- und Extremwerte sind in Tab. 5.4 zu finden und die auf Hot- und Coldspots untersuchten Jahreszeitenmittel sind in Abb. B.1, B.2 und B.3 enthalten.

Für  $\text{PM}_{2.5}$  und Stickstoffdioxid liegen die Schadstoffbelastungen in Winter 39% bzw. 44% über dem Gesamtdurchschnitt (aller Jahre), während im Sommer ein Rückgang der Schadstoffe um 20% bzw. 25% zu beobachten ist. Für  $\text{NO}_2$  lässt sich zudem im Herbst ebenfalls eine 26%-ige Erhöhung und im Frühling eine 20%-ige Verringerung der Konzentrationen feststellen. Die jahreszeitliche Verteilung der Ozonkonzentrationen verhält

**Tabelle 5.4:** Räumliche Mittel- und Extremwerte in den flächenhaften Mittelwerten der Schadstoffwerte für die einzelnen Jahreszeiten.

	PM <sub>2.5</sub> [ $\mu\text{g}/\text{m}^3$ ]			NO <sub>2</sub> [ $\mu\text{g}/\text{m}^3$ ]			O <sub>3</sub> [ $\mu\text{g}/\text{m}^3$ ]		
	Mittel	Min.	Max.	Mittel	Min.	Max.	Mittel	Min.	Max.
Gesamt	10.91	5.21	26.88	7.56	1.62	71.17	83.84	54.73	96.69
Winter	15.10	4.74	45.55	10.92	2.52	72.64	60.18	30.31	80.43
Frühling	11.16	4.64	34.84	6.09	1.17	60.13	98.96	64.32	110.46
Sommer	8.73	4.37	18.52	5.69	0.60	78.84	103.02	60.33	126.28
Herbst	10.99	4.65	25.60	9.51	1.71	71.41	55.53	32.46	73.91

sich qualitativ umgekehrt zur vorhandenen NO<sub>2</sub>-Menge, mit deutlich stärkerer Belastung im Frühling und im Sommer (18 bzw. 23% über dem Jahresmittel) und geringeren Werten im Herbst und Winter (34 bzw. 28% unter dem Jahresmittel). Die räumliche Verteilung von NO<sub>2</sub> erfährt über die Jahreszeiten keine signifikante Veränderung, abgesehen davon, dass sich im Winter die Werte vor allem um die Ballungsräume erhöhen und sich die entsprechenden Hotspots vergrößern (vgl. Abb. 5.16 und B.2). Zudem entfällt aufgrund der großflächig geringen NO<sub>2</sub>-Konzentration im ländlichen Bereich im Sommer und Frühling der entsprechende Coldspot in den Alpen (vgl. Abb. B.2c und B.2b). PM<sub>2.5</sub> erfährt in der kalten Jahreszeit einen relativ gleichmäßigen Anstieg, in sämtlichen Regionen, wobei im Norden von Tschechien für alle Jahreszeiten ein Hotspot vorliegt (vgl. Abb. B.1). Für Ozon ist die bereits in Kap. 5.3 beobachtete geringere Bodenkonzentration innerhalb von Städten im Winter und Herbst deutlich stärker ausgeprägt als im Sommer und Frühling. Es können im Sommer nur noch für wenige Städte (z.B. Paris und Bremen), dafür aber für die Küsten- und Alpenregionen größere Coldspots ausgemacht werden. Im Südwesten des Untersuchungsgebiets, bzw. im Südosten im Frühling, breiten sich die größten Ozonhotspots über das Flachland um die Alpen aus. Gebirge und naturbelassene Gebiete, die in den warmen Jahreszeiten mit leicht verringerten Ozonkonzentrationen auffallen, verändern sich über das Jahr nur geringfügig und liegen im Winter eher über dem Saisondurchschnitt (vgl. Abb. 5.16).

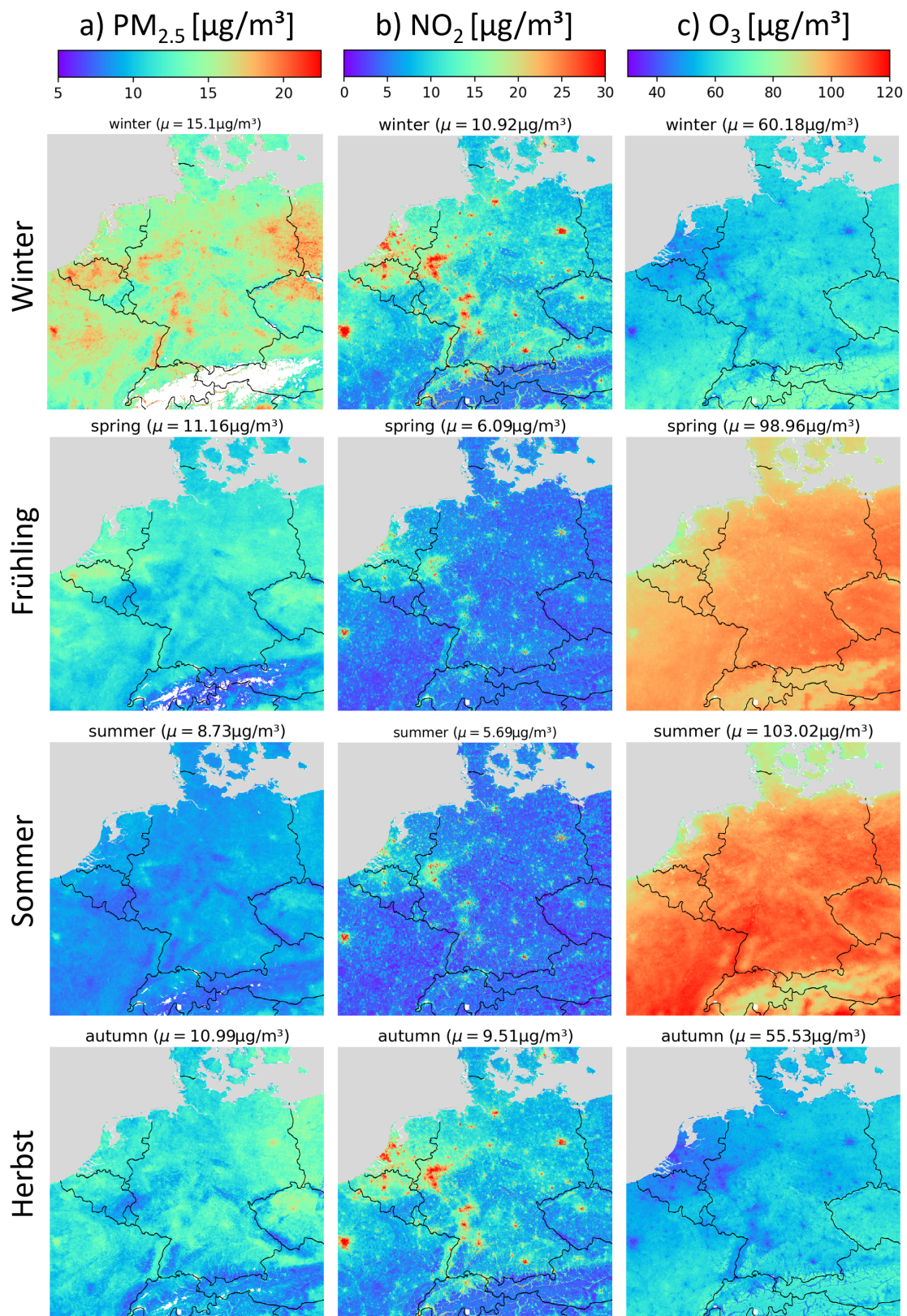
## 5.5 Schadstoffverteilung bei verschiedenen meteorologischen Bedingungen

In Folgenden soll durch Einteilung der Tage im Untersuchungszeitraum nach bestimmten meteorologischen Kriterien der Einfluss von Wetterverhältnissen auf die von den RF-Modellen vorhergesagte Luftqualität ersichtlich werden.

### 5.5.1 Auswirkung der Windbedingungen

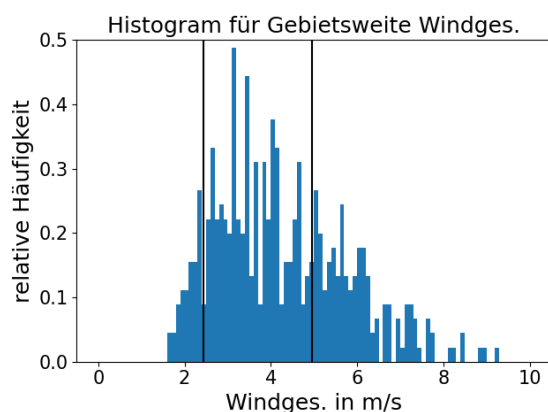
Zunächst wird die Schadstoffkonzentration für verschiedene Windverhältnisse untersucht. Ein Maß für die Windstärke an einem Tag im gesamten Untersuchungsgebiet wird durch Mittelung der tagesspezifischen ECMWF-Daten für den Wind erhalten. Anschließend werden die Daten so aufgeteilt, dass jeweils ca. 15% der Tage mit den höchsten und niedrigsten mittleren Windgeschwindigkeiten zusammengefasst werden, wie in Abb. 5.17 dargestellt ist. Dabei muss eingeräumt werden, dass die so erhaltenen mittleren Windgeschwindigkeiten im Winter Tendenziell höher ausfallen als im Sommer, und somit saisonale Unterschiede bei der Interpretation berücksichtigt werden müssen. Zusätzlich dazu soll der Einfluss der Windrichtung untersucht werden, wozu Großwetterlagenklassifikationen aus Kap. 4.5.2



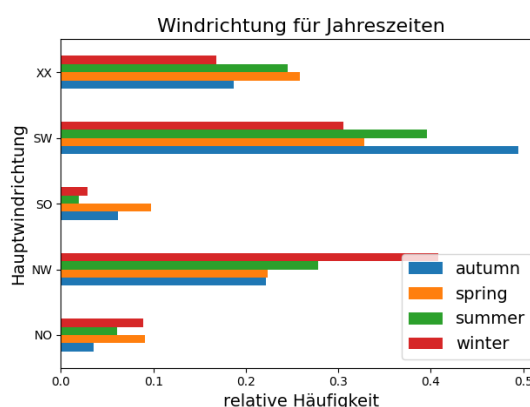


**Abbildung 5.16:** Jahreszeitliche Mittelwerte der Modellvorhersagen des bodennahen a) PM<sub>2.5</sub>, b) NO<sub>2</sub> und c) O<sub>3</sub>.

zum Einsatz kommen. Die Vorhersagen werden dazu über ihre Windindizes nach Hauptwindrichtungen gruppiert und jeweils ein Mittelwert gebildet. Dabei ist wiederum auf die ungleiche Verteilung der Hauptwindrichtungen in den einzelnen Jahreszeiten zu achten, die in Abb. 5.18 zu sehen ist. Die häufigste Windrichtung ist die Südwestlage mit 38% aller Tage, gefolgt von Nordwest mit 28% und unbestimmt (XX) mit 21%. Nordöstliche und südöstliche Anstromrichtungen treten nur an 7% bzw. 6% der Tage auf. Im Winter ist Nordwesten die häufigste Windrichtung, während zu den übrigen Jahreszeiten verstärkt Südwestwinde vorliegen.



**Abbildung 5.17:** Aufteilung der Daten nach Stärke der mittleren Windgeschwindigkeit pro Tag. Der linke und rechte Extrembereich jenseits der schwarzen vertikalen Linien enthält jeweils ca. 15% der Daten.



**Abbildung 5.18:** Verteilung der Hauptwindrichtung aus den Großwetterlagen auf die Jahreszeiten.

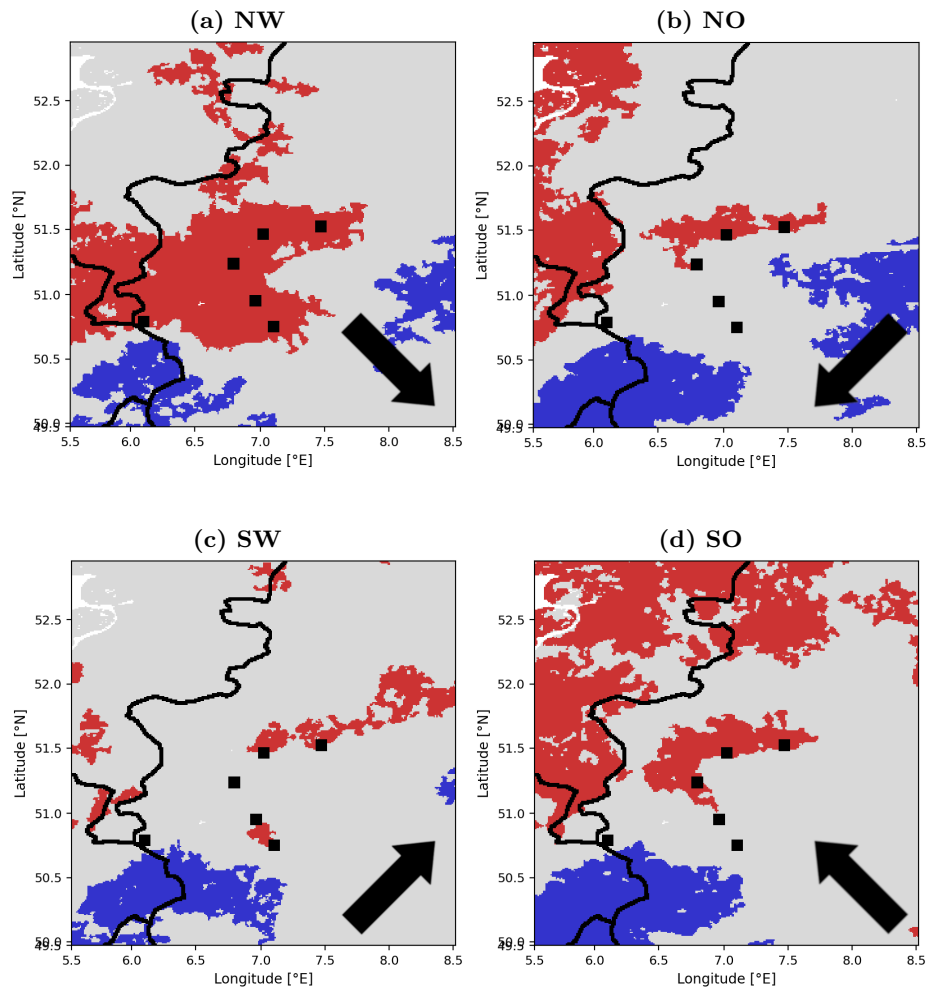
## PM<sub>2.5</sub>

Der Bereich mit mittleren Windstärken (siehe Abb. A.4b) zwischen 2.5 und 5 m/s weist ähnliche Feinstaubmittelwerte und -verteilungen auf wie die Gesamtmittelwerte in Abb. 5.8a. Zwischen Stark- und Niedrigwindniveau ist jedoch ein deutlicher Unterschied zu erkennen. Bei geringen Windstärken stellt sich im Untersuchungsgebiet ein erhöhter Mittelwert von  $11.89 \mu\text{g}/\text{m}^3$  ein; bei höheren Windgeschwindigkeiten fällt dieser unter den Durchschnitt auf  $9.03 \mu\text{g}/\text{m}^3$  ab. Hohe Windstärken sind oft mit stärkerer Wolkenbedeckung verbunden, weshalb die Datenabdeckung in diesem Bereich deutlich schlechter ausfällt und Artefakte wie in Abb. A.4c entstehen. Wenn davon ausgegangen wird, dass der dadurch entstandene Fehler gering ist, kann geschlossen werden, dass die Region von der Nordseeküste in Belgien und den Niederlanden, bis zur Rhein-Ruhr-Region bei hohen Windstärken zu Coldspots in der Feinstaubkonzentration werden (vgl. Abb. B.4c). Dieselbe Region stellt bei niedrigen Windstärken den stärksten PM<sub>2.5</sub> Hotspot im Gesamtgebiet dar (vgl. Abb. B.4c). Die Windrichtung hat ebenfalls starken Einfluss auf die PM<sub>2.5</sub>-Belastung. Die höchsten Feinstaubwerte finden sich im Mittel für Südostwind ( $13.05 \mu\text{g}/\text{m}^3$ ), wobei Nordwestwind mit  $9.23 \mu\text{g}/\text{m}^3$  die niedrigste Belastung mit sich bringt (Abb. A.5). Dieser Unterschied ist vermutlich auf die Saisonalität des Windes zurückzuführen. Die Metropolregion Rhein-Ruhr ist als markanter Hotspot gut für die Untersuchung von Änderungen in der PM<sub>2.5</sub>-Verteilung mit der Windrichtung geeignet.

Beim Vergleich der Hotspots im Gesamtmittelwert (Abb. 5.9a) mit denen in Abb. 5.19 lässt sich der Einfluss der Windrichtung auf die PM<sub>2.5</sub>-Verteilung erkennen. Der vom Ruhrgebiet und den Städten Köln, Bonn und Düsseldorf ausgehende großflächige Hotspot fällt für die



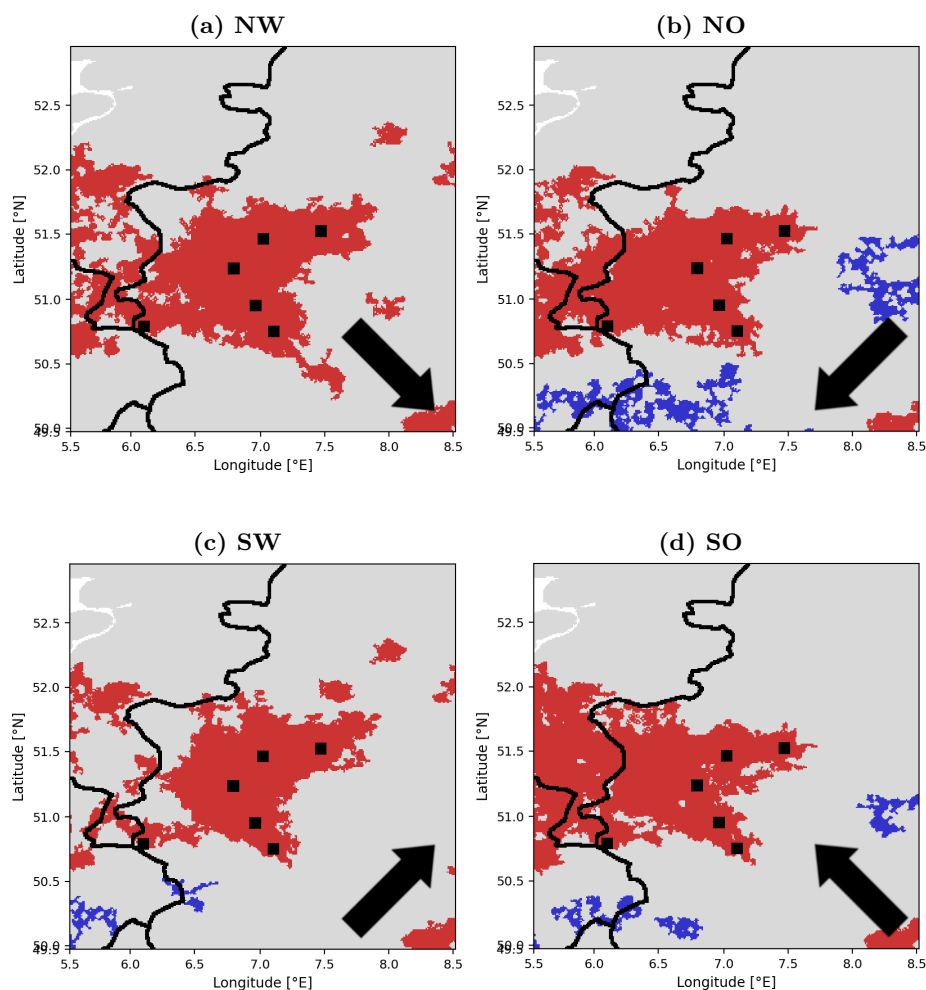
Hauptwindrichtung NW besonders groß aus. Des Weiteren ist in den Abbildungen 5.19 a) (NW) c) (SW) und d) (SO) zu erkennen, dass sich dieser tendenziell verstärkt Abwärts der jeweiligen Windrichtung ausbreitet.



**Abbildung 5.19:** Ausschnitte der  $PM_{2.5}$ -Hotspots über dem Ruhrgebiet bei nordwestlicher, nordöstlicher, südwestlicher und südöstlicher Anstromrichtung. Die Pfeile kennzeichnen die jeweilige Hauptwindrichtungen nach der Wetterlagenklassifikation des DWD.

## $NO_2$

Ähnlich wie bei  $PM_{2.5}$  sind für  $NO_2$  nur die Windextreme von Interesse die sich analog mit niedrigen Werten von  $6.81 \mu\text{g}/\text{m}^3$  bei starkem Wind und hohen Werten von  $7.90 \mu\text{g}/\text{m}^3$  bei schwachem Wind auszeichnen. Die  $NO_2$ -Konzentration nimmt bei zunehmenden Windgeschwindigkeiten vor allem in den Ballungsräumen deutlich ab und die Hotspots verkleinern sich (vgl. Abb. A.6c und B.6c). Anders als bei  $PM_{2.5}$  hat die Hauptwindrichtung für Stickstoffdioxid keinen signifikanten Einfluss auf die mittlere Bodenkonzentration (vgl. Abb. A.7). Allerdings lässt sich in Abb. 5.20b, 5.20d und 5.20c (NO, SO und SW) für die Hotspots rund um das Ruhrgebiet erneut eine Verstärkung der Schadstoffreichen Gebiete windabwärts des Ballungsgebiets erkennen.



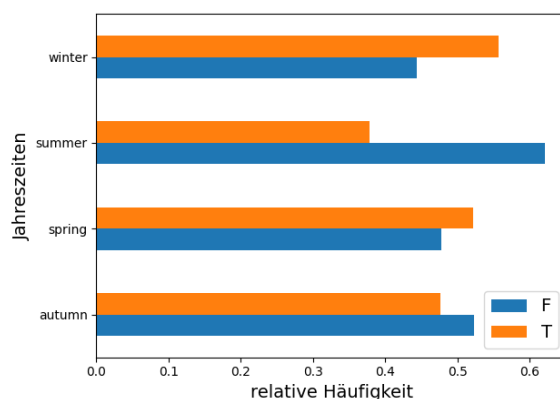
**Abbildung 5.20:** Ausschnitte der  $\text{NO}_2$ -Hotspots über dem Ruhrgebiet bei nordwestlicher, nordöstlicher, südwestlicher und südöstlicher Anstromrichtung. Die Pfeile kennzeichnen die jeweilige Hauptwindrichtungen nach der Wetterlagenklassifikation des DWD.

### $\text{O}_3$

Die Ozonkonzentrationen zeigen ähnliche Windabhängigkeiten wie  $\text{PM}_{2.5}$  (Abb. A.8 u. A.9). Sie ist besonders gering für hohe Windgeschwindigkeiten ( $72.42 \mu\text{g}/\text{m}^3$ ) und die Windindizes NO ( $79.76 \mu\text{g}/\text{m}^3$ ), NW ( $79.16 \mu\text{g}/\text{m}^3$ ), und SW ( $81.29 \mu\text{g}/\text{m}^3$ ). Dagegen fällt sie höher aus bei niedrigen Windgeschwindigkeiten ( $92.18 \mu\text{g}/\text{m}^3$ ) und südöstlicher und unbestimmter Anstromrichtung (SO:  $90.09 \mu\text{g}/\text{m}^3$ , XX:  $89.00 \mu\text{g}/\text{m}^3$ ). Für die nordwestliche und südwestliche Hauptwindrichtung vergrößern sich die Coldspots an der Nord- bzw. Ostseeküste, während im südlichen Inland die im Gesamtmittel beobachteten Ozonhotspots zu finden sind (siehe Abb. A.9b und A.9d). Bei Nordostwind ist im Osten, also in Österreich, Tschechien und Teilen Ostdeutschlands und Polens ein Ozonminimum im Inland zu erkennen (vgl. Abb. A.9e), wohingegen Ostdeutschland, Polen und Tschechien bei Südostwind besonders hohe Ozonkonzentrationen aufweisen (vgl. Abb. A.9c).

### 5.5.2 Auswirkung des atmosphärischen Feuchtegehalts

Der Wassergehalt der Atmosphäre spielt eine wichtige Rolle in den chemischen Prozessen der Schadstoffe, weshalb im Folgenden nach Unterschieden zwischen den Feuchteklassen (trocken (T) und feucht (F)) der Großwetterlagen gesucht wird. Da hierbei mögliche Einflüsse der Jahreszeitlich bedingten Emissions- und Temperaturunterschiede verhindert werden sollen, die sich aus der in Abb. 5.21 gezeigten ungleichen Verteilung der Feuchteindizes zwischen Sommer und Winter ergeben, wird der Datensatz im Folgenden zusätzlich nach Jahreszeiten aufgespalten.



**Abbildung 5.21:** Relative Häufigkeit der Tage deren Großwetterlage als feucht oder trocken eingestuft wurde, in Abhängigkeit der Jahreszeit.

**Tabelle 5.5:** Mittlere Schadstoffkonzentrationen über die Gesamtfläche des Beobachtungsgebiets, jeweils für feuchte (F) und trockene (T) Wetterbedingungen

Feuchteindex:	PM <sub>2.5</sub> [µg/m <sup>3</sup> ]		NO <sub>2</sub> [µg/m <sup>3</sup> ]		O <sub>3</sub> [µg/m <sup>3</sup> ]	
	T	F	T	F	T	F
Gesamt	11.16	10.53	7.90	7.16	79.25	89.14
Winter	15.20	14.69	11.35	9.87	59.90	60.85
Frühling	11.34	10.83	6.15	5.98	96.62	102.11
Sommer	7.66	9.43	5.44	5.85	94.94	108.33
Herbst	10.96	10.99	9.63	9.33	55.23	55.99

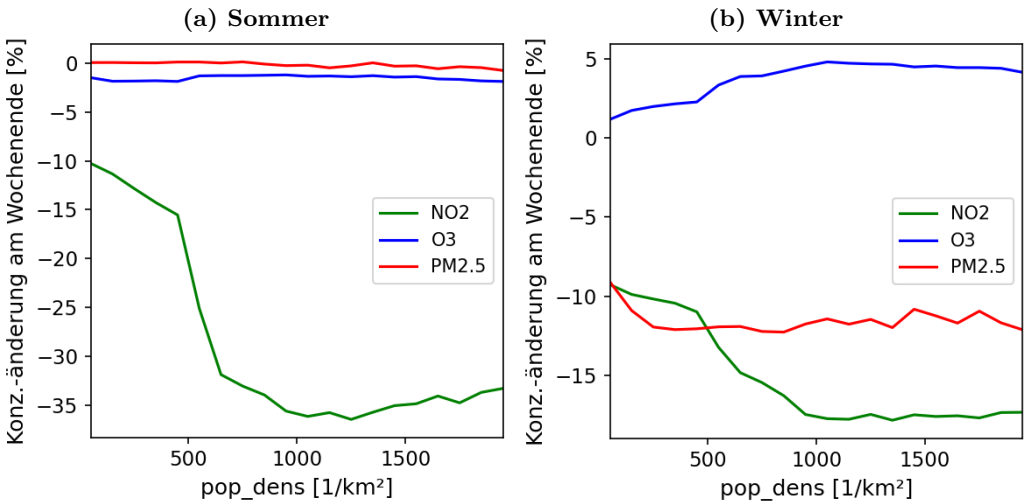
Im Winter und Frühling lassen sich leichte Erhöhungen der Feinstaubkonzentration bei trockenen Bedingungen erkennen, wohingegen im Sommer in solchen Wetterlagen wesentlich weniger Feinstaub vorliegt (siehe Tabelle 5.5). Die räumliche Verteilung der Daten zeigt wenig Auffälligkeiten für die einzelnen Bereiche. Lediglich im Herbst sind für feuchte Verhältnisse in Ostdeutschland, Polen und Tschechien die Feinstaubwerte sichtbar gegenüber denen der trockenen erhöht. Für NO<sub>2</sub> liegt nur während der kalten Jahreszeit ein verringerter Mittelwert für als feucht klassifizierte Tage vor, wobei sich im Bezug auf die relative räumliche Ausdehnung wenig verändert. Die Ozonwerte weisen für feuchte Konditionen generell, besonders jedoch im Sommer, einen Anstieg auf. Es ist anzumerken, dass die Abdeckung durch Satellitendaten in trockenen Bedingungen mit 42% höher ausfällt als in feuchten Bedingungen (32%).

### 5.6 Wochenendeffekt

Die Veränderungen der anthropogenen Aktivitäten im Verlauf einer Woche finden größtenteils unabhängig von anderen Umwelteinflüssen statt. Der Unterschied in den Schadstoffkonzentrationen zwischen Wochenende (Samstag und Sonntag) und dem Rest der Woche (Montag bis Freitag) kann daher direkt auf den Menschen zurückgeführt werden und wird im Folgenden untersucht.

**Tabelle 5.6:** Mittlere Schadstoffkonzentrationen über die Gesamtfläche des Beobachtungsgebiets, jeweils von Montag bis Freitag und von Samstag und Sonntag für verschiedene Jahreszeiten.

Wochentag	PM <sub>2.5</sub> [µg/m <sup>3</sup> ]		NO <sub>2</sub> [µg/m <sup>3</sup> ]		O <sub>3</sub> [µg/m <sup>3</sup> ]	
	Mo-Fr	Sa-So	Mo-Fr	Sa-So	Mo-Fr	Sa-So
Gesamt	10.88	10.99	7.79	6.97	83.85	83.86
Winter	15.58	14.16	11.20	10.16	59.94	60.80
Frühling	11.03	11.57	6.25	5.70	98.59	99.88
Sommer	8.73	8.74	5.87	5.19	103.45	101.91
Herbst	10.99	11.03	9.94	8.52	54.75	57.33



**Abbildung 5.22:** Mittlerer Änderung der NO<sub>2</sub>, O<sub>3</sub> und PM<sub>2.5</sub>-Konzentration am Wochenende gegenüber dem jeweiligen jahreszeitlichen Mittel des (a) Sommers und (b) Winters. Die Stärke des Effekts ist für Gebiete mit verschiedenen Bevölkerungsdichten dargestellt.

Tabelle 5.5 zeigt, dass die NO<sub>2</sub>-Konzentration an Wochenenden ca. 8% unterhalb des Gesamtmittelwertes liegt. Der Unterschied zwischen Wochenenden und Tagen unter der Woche fällt in den einzelnen Jahreszeiten unterschiedlich stark aus, wobei vor allem die Unterschiede in den Extrema (Sommer und Winter) interessant sind. Im Sommer beträgt dieser ca. -12% des Jahreszeitenmittelwerts und -10% bzw. -15% im Winter bzw. Herbst. Abb. 5.22 zeigt, dass die relative Stärke der Abnahme tendenziell mit der Bevölkerungsdichte der Region steigt, was auf unterschiedlich starke Effekte in ländlichen und urbanen Regionen schließen lässt.

Für bodennahes Ozon fällt der Wochenendeffekt im Sommer mit -1.5% geringfügig negativ aus, ist für den Winter (+1.4%) und Herbst (+4.6%) jedoch positiv. Der Effekt ist im Sommer größtenteils unabhängig von der Bevölkerungsdichte (vgl. Abb. 5.22a), wobei

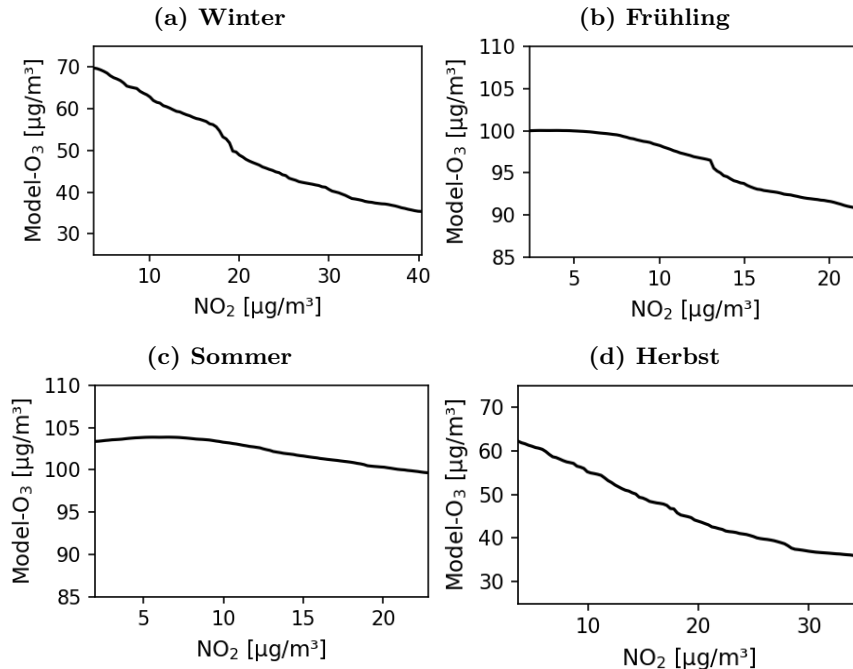
er im Winter analog zu NO<sub>2</sub> mit dieser in seiner Stärke zunimmt (vgl. Abb. 5.22b).

Zur Vollständigkeit wurden auch die Feinstaubdaten entsprechend aufgeteilt, wobei sich hier für den Winter ein Wochenendeffekt von -9.4% ergibt, der anfänglich mit der Bevölkerungsdichte wächst. Dagegen fällt dieser im Frühling geringfügig positiv aus (+4.8%) und ist vernachlässigbar klein zu den übrigen Jahreszeiten.

## 5.7 Saisonale Abhängigkeiten zwischen Ozonkonzentration und NO<sub>2</sub>

Die aus Kapitel 2.4 bekannten Beziehungen zwischen Ozon- und Stickoxidkonzentrationen sollen im Folgenden untersucht werden. Dazu wird für jede Jahreszeit auf den entsprechenden Daten je ein Ozonmodell mit den gleichen Hyperparametern und Eingangsvariablen wie das Gesamtmodell trainiert. Für diese saisonalen Modelle wird ihm Anschluss die partielle Abhängigkeit zwischen O<sub>3</sub> und NO<sub>2</sub> ausgewertet.

Die saisonalen Ozonmodelle besitzen aufgrund der geringeren Trainingsdatenmenge mit  $R^2 = 0.8$  im Winter und 0.82 im Sommer eine etwas geringere Modellgenauigkeit. Die partiellen Abhängigkeiten zur Vorläufersubstanz NO<sub>2</sub> sind in Abb. 5.23 dargestellt, wobei auf die Verwendung unterschiedlicher Achsenlimits hingewiesen sei. Für Winter und Herbst ist die Relation zwischen O<sub>3</sub> und NO<sub>2</sub> in Abbildung 5.23a und 5.23d am eindeutigsten und streng monoton fallend. Im Sommer ist für geringe NO<sub>2</sub>-Werte zunächst ein Anstieg des Ozons mit zusätzlichem NO<sub>2</sub> zu erkennen, ab Konzentrationen über 6 µg/m<sup>3</sup> sinkt die Ozonkonzentration jedoch ab (vgl. Abb. 5.23c). Das Verhalten im Frühling zeigt ebenfalls eine Antikorrelation der beiden Schadstoffe, diese ist allerdings deutlich schwächer ausgeprägt als im Winter bzw. Herbst (vgl. Abb. 5.23b).



**Abbildung 5.23:** Partielle Abhängigkeiten des Vorhergesagten Ozons vom bodennahen NO<sub>2</sub> (NO2\_pred). Es wurden separate Modelle für die Daten aus (a) Winter, (b) Frühling, (c) Sommer und (d) Herbst trainiert.



## 6 Diskussion

Im folgenden Teil der Arbeit sollen die Ergebnisse aus Kapitel 5 näher untersucht und mithilfe der in den Grundlagen (Kapitel 2) vermittelten Kenntnisse über die troposphärische Schadstoffentwicklung interpretiert werden.

### 6.1 Modellgenauigkeit

Die Genauigkeiten der drei Random Forest Modelle zur Vorhersage von bodennahen PM<sub>2.5</sub>-, NO<sub>2</sub>-, und O<sub>3</sub>-Konzentrationen aus Satellitendaten und weiteren unterstützenden Variablen wurde ermittelt. Der grundlegend sehr geringe Bias und die hohe Bestimmtheit ( $R^2$ ) lassen darauf schließen, dass alle drei Modelle gut für die Vorhersage bodennaher Schadstoffkonzentrationen geeignet sind. Da  $R^2$ -Werte in der Regel den besseren Vergleichswert für Regressionsmodelle darstellen [107] wird im Folgenden primär auf diese eingegangen, um Modelle zu vergleichen.

Demnach lässt sich für das trainierte PM<sub>2.5</sub>-Modell die geringste Modellgenauigkeit unter den drei Schadstoffmodellen feststellen. Es liefert ähnlich gute Ergebnisse wie vergleichbare Modelle in den Untersuchungen von Handschuh et al. [28][27] und Hu et al. [108]. Letztere stellten zudem einen positiven Einfluss von Faltungsebenen auf die Modellgenauigkeit fest. Bei einem solchen Ansatz erhält das Modell nicht nur Daten für den entsprechenden Gitterpunkt und Zeitpunkt, sondern zusätzlich Informationen über die räumliche und zeitliche Umgebung, z.B. erhält das Modell zusätzliche Informationen über nahegelegene PM<sub>2.5</sub>-Messungen (vgl. Hu et al. [108]) oder zu Messungen am Vortag. Die zeitliche Kreuzvalidierung zeigt klar, dass zeitliche Trennung der Daten einen starken Einfluss auf die Modellgenauigkeit ausübt, wobei die Verschlechterung der Vorhersagen für die striktere Trennung stärker ausfällt. Die räumliche Kreuzvalidierung konnte dabei mit wesentlich kleineren Verschlechterungen noch vergleichsweise gute Bestimmtheitsmaße hervorbringen. Beim Einsatz des Modells auf Daten außerhalb des Untersuchungszeitraums muss daher mit erheblichen Verringerungen in der Genauigkeit der Vorhersagen gerechnet werden, was die universelle Anwendbarkeit des Verfahrens einschränkt. Das sich PM<sub>2.5</sub>-Vorhersagen in ML-Modellen durch zeitliche Kreuzvalidierung in der Regel stärker verschlechtern als durch räumliche Kreuzvalidierung deckt sich mit den Beobachtungen bestehender Studien z.B. von Zhu et al. [109] und Handschuh et al. [110].

Auch für die Vorhersage von oberflächennahen NO<sub>2</sub>-Konzentrationen mit RF-Modellen finden sich ähnliche Genauigkeiten in bestehenden Studien wie der von M. Li et al. [111]. Bessere Ergebnisse konnten wiederum von Cao [112] mithilfe eines Faltungs-Neuralen Netzes und moderat veränderten Vorhersagevariablen erreicht werden. Balamurugan et al. [19] setzten dagegen ein Gradient-Boosting-Tree Modell (GBT) sehr ähnlicher Datengrundlage wie in diese Arbeit zur Vorhersage von bodennahem NO<sub>2</sub> ein und konnten Bestimmtheitsmaße von  $R^2 = 0.88$  erreichen. Unterschiede im Untersuchungszeitraum und vor allem im Untersuchungsgebiet könnten für die stark unterschiedlichen Modellgenauigkeiten verantwortlich sein. Die am dichtesten besiedelten Länder Europas, also Belgien und die Niederlande, welche auch die höchste Dichte an Messstationen aufweisen, fielen nicht in den Trainingsbereich von Balamurugan et al. [19] (ausschließlich Deutschland) und

könnten die Beziehung zwischen Eingangsvariablen und  $\text{NO}_2$  deutlich komplexer ausfallen lassen als bei der Betrachtung eines einzelnen Landes.  $\text{NO}_2$ -Modelle für unterschiedliche Regionen Mitteleuropas können sich daher stark voneinander unterscheiden. Dies erklärt auch die Ergebnisse der räumlichen Kreuzvalidierung, die eine deutliche Verschlechterung bei der Trennung von Test- und Trainingsdaten nach (Bundes-) Ländern zeigt. ML-basierte Stickstoffdioxidvorhersagen scheinen, ungeachtet des verwendeten Modelltyps, starke Verringerungen in der Modellgenauigkeit bei räumlich getrennter Auswertung zu erfahren und bei zeitlicher Trennung größtenteils unbeeinflusst zu bleiben, wie der Vergleich mit anderen Untersuchungen zeigt [19][112]. Ein Einsatz des Modells auf Gebieten außerhalb des Untersuchungsraums, ohne weitere Trainingsdaten für diese Flächen, ist daher nur mit deutlich geringerer Präzision möglich.

Das Ozonmodell zeigt besonders gute Genauigkeiten und schneidet von den drei Modellen am besten ab. Es stellt eine leichte Verbesserung gegenüber früheren Studien dar, in denen, im Kontrast zu dieser Arbeit, kein gestaffelter Aufbau aus  $\text{NO}_2$ - und  $\text{O}_3$ -Modellen angesetzt wurde (vgl. W. Wang et al. [31]). Ähnlich gute Vorhersagewerte für bodennahes Ozon konnten durch ein Ensemble-Modelle von X. Liu et al. [37] oder mittels GBT-Modell von Balamurugan et al. [19] erzielt werden. Bei letzteren kam, wie hier, ein gestaffelter Modelllaufbau mit Nutzung des vorhergesagten  $\text{NO}_2$ -Werts und mit vergleichbaren Datenquellen zum Einsatz. Die Modellleistung scheint sich bei räumlicher und zeitlicher Kreuzvalidierung nur geringfügig zu verändern, was sich mit den Ergebnissen von Balamurugan et al. [19] deckt und für die Übertragbarkeit des Modells auf andere Zeiträume und Gebiete spricht.

Es konnte festgestellt werden, dass alle drei hier trainierten Modelle dazu neigen, niedrige Extremwerte in den Vorhersagen zu überschätzen und hohe Extremwerte zu unterschätzen. Dies ist ebenfalls an den zum Vergleich herangezogenen Modellen von Handschuh et al. [28][27], Cao [112], Balamurugan et al. [19] und M. Li et al. [111] zu finden. Dies könnte eine direkte Folge der entscheidungsbaumbasierten Vorhersagemethoden sein, in der Wertebereiche durch ihren Mittelwert approximiert werden und so Extremwerte nicht gut repräsentiert werden können.

Modellansätze, in denen räumliche- oder zeitliche Faltungen der Daten vorgenommen wurden (vgl. Hu et al. [108] und Sun et al. [113]) waren dem hier verwendeten Ansatz tendenziell überlegen. Tatsächlich bleiben durch die unabhängige Vorhersage der Schadstoffkonzentrationen pro Gitterpunkt, Informationen aus benachbarten Gebieten unberücksichtigt, die z.B. durch den Transport von Schadstoffen oder deren Vorläufersubstanzen Einfluss auf die Schadstoffkonzentration nehmen können. Besonders eine Berücksichtigung der Windrichtung bei der Faltung und der Nähe des Gitterpunkts zu starken stationären Schadstoffquellen wie Flughäfen oder Kraftwerken könnte die Informationslage des Vorhersagemodells deutlich verbessern. Eine zeitliche Faltung der Eingangsdaten könnte vor allem für langlebigere Schadstoffe, deren Konzentration sich über den Zeitraum von Tagen ändert, hilfreiche Informationen bereitstellen.

## 6.2 Physikalische Auslegung der Modellabhängigkeiten

Im Folgenden werden die Ergebnisse der Feature Importance Analyse interpretiert und ggf. anhand der partiellen Abhängigkeiten überprüft, ob die erlernten Zusammenhänge das entsprechenden Modells physikalisch sinnvoll sind.



### PM<sub>2.5</sub>

Die wichtigste Variable für die Vorhersage von bodennahen Feinstaubwerten ist durch den Jahrestag (DoY) gegeben. Dieser kann aufgrund der starken Korrelation zusammen mit dem drittichtigsten Merkmal (Month) gemeinsam betrachtet werden (vgl. Abb. 5.7). Die Wichtigkeit dieser Variablen ist einerseits darin zu begründen, dass PM<sub>2.5</sub> durch Unterschiede in den menschenverursachten Emissionen, aber auch durch meteorologische Bedingungen eine starke Saisonalität besitzt. Zeitabhängige Merkmale können Muster im Jahresgang der Luftverschmutzung an das Modell weitergeben und so die Vorhersage deutlich verbessern. Dies ist auch am Verlauf der partiellen Abhängigkeit zwischen PM<sub>2.5</sub> und DoY zu erkennen (vgl. Abb. 5.3 a)). Der zum Teil unstetige Verlauf ist ein Anzeichen darauf, dass das PM<sub>2.5</sub>-Modell den Jahrestag nutzt um besondere Feinstaubereignisse in den Trainingsdaten zu erlernen. Aufgrund der begrenzten Anzahl an Messjahren, lassen sich durch in Kombination von DoY- mit der Year-Variable einzelne Zeitepisoden identifizieren und die jeweiligen Feinstaubverhältnisse vorhersagen, ohne einen direkten Zusammenhang mit den physikalischen Verursachern oder der AOD herstellen zu müssen. Dies deckt sich mit den Ergebnissen der temporalen Kreuzvalidierung, bei der die in den Daten vorhandenen Muster durch zeitliche Trennung nicht mehr erkannt werden und es so zu einer Verschlechterung der Modellgenauigkeit kommt.

Die AOD wird insgesamt als zweitwichtigster und unter den physikalischen Merkmalen als wichtigster Vorhersageparameter identifiziert. Hohe PM<sub>2.5</sub>-Konzentrationen tragen direkt zu der Aerosolmenge in der vertikalen Atmosphärensäule bei, deren Absorptionsstärke durch die AOD charakterisiert wird, weshalb die streng monoton steigende partielle Modellabhängigkeit (Abb. 5.3 b)) einen durchaus sinnvollen physikalischen Zusammenhang zwischen PM<sub>2.5</sub> und AOD darstellt.

Auch die Höhe der Grenzschicht wirkt sich als eines der wichtigsten Merkmale auf die bodennahe Feinstaubkonzentration aus. Innerhalb der Grenzschicht steht die Atmosphäre in direkter Wechselwirkung mit der Erdoberfläche. Durch Reibungskräfte verursachte Turbulenz sowie Konvektion spielen eine entscheidende Rolle für die Ausbreitung und dem Transport atmosphärischer Spurenstoffe [114][115]. Eine hohe Schichthöhe erlaubt eine großräumigere Verteilung der Staubpartikel und resultiert so in einer geringeren PM<sub>2.5</sub>-Bodenkonzentration, was sich ebenfalls im Verhalten des Modells widerspiegelt (Abb. 5.3 c)) [116]. Die saisonale Abhängigkeit der Grenzschichthöhe fließt dabei zudem in die Funktion der DoY-Variable ein. In den Wintermonaten ist die Grenzschicht oft niedriger [117], was die Feinstaubkonzentration erhöht, wobei die Emissionen durch Heizungsaktivitäten zusätzlich verstärkt werden. Tageszeitliche Schwankungen der Grenzschichthöhe können vom Modell jedoch nicht berücksichtigt werden, da lediglich ein einzelner Zeitschritt pro Tag verarbeitet wurde.

Auch die Beschaffenheit der Grenzschicht spielt eine Rolle für die PM<sub>2.5</sub>-Konzentration, da sie zusätzliche Informationen zur vertikalen Verteilung der Aerosole liefert. Bei hoher turbulenter Diffusivität werden die Luftmassen innerhalb der Schicht besser durchmischt, was dazu führt, dass Staubpartikel schneller und effektiver verteilt werden und weniger Feinstaub in Bodennähe vorliegt [118]. Als Parameter für diese Beschaffenheit wurde die Variable BLD betrachtet, die die Energiedissipation durch turbulente Vorgänge in der Grenzschicht beschreibt. Auch hier hat das Modell einen physikalisch sinnvollen Zusammenhang gelernt (vgl. Abb. 5.3 d)).

Generell konnte festgestellt werden, dass sich das PM<sub>2.5</sub>-Modell durch Hinzufügen von Daten mit geographischem Bezug wie Bevölkerungsdichte oder Straßendichte kaum verbessern ließ. Dies widerspricht der Erwartung, dass PM<sub>2.5</sub> zum Teil an ortsfeste Quellen

wie urbane Emissionen gebunden ist.

In feuchter Umgebung kann es zu hygroskopischem Wachstum von Feinstaubpartikeln kommen, bei dem Aerosole Feuchtigkeit aus der Luft anziehen und so ihre chemischen Eigenschaften und Absorptionseigenschaften ändern [119]. Dies kann sich direkt auf die Beziehung zwischen  $\text{PM}_{2.5}$ -Konzentration und der gemessenen AOD auswirken weshalb es nicht den Erwartungen entspricht, dass wichtige Indikatoren des Prozesses wie die relative Luftfeuchtigkeit (RH) eine derart geringe Rolle in der Modellvorhersage einnehmen. Die Wichtigkeit von RH für die  $\text{PM}_{2.5}$ -Vorhersage ist dabei in machen Studien significant [118], wird jedoch von anderen als schwach befunden [110, 120]. Es ist denkbar, dass der Einfluss der Luftfeuchtigkeit durch andere korrelierte Merkmale wie TP, BLH oder SSR in das Modell eingeht.

## **$\text{NO}_2$**

Durch das Fehlen signifikanter natürlicher Quellen wird die örtliche  $\text{NO}_2$ -Belastung hauptsächlich durch anthropogene Prozesse gesteuert, wobei Verkehrsemissionen den größten Beitrag zu den Stickoxidemissionen liefern. Es ist daher nicht ungewöhnlich, dass die wichtigste Eingangsvariable in der Vorhersage von bodennahem Stickstoffdioxid die Bevölkerungsdichte ist. In bevölkerungsreichen Gebieten erhöht sich das Verkehrsaufkommen und somit auch die Emissionen, weshalb der Anstieg der  $\text{NO}_2$ -Vorhersagewerte mit der Bevölkerungsdichte in Abb. 5.4 a) sinnvoll ist. Bei 7000 Einwohnern pro  $\text{km}^2$  erreicht die mittlere  $\text{NO}_2$ -Konzentration jedoch einen Hochpunkt und steigt danach nicht weiter. Als möglicher Grund dafür kommt in Frage, dass besonders dicht besiedelte Gebiete häufig gut durch öffentlichen Nahverkehr erschlossen sind und der Bedarf bzw. das Aufkommen an Individualverkehr dadurch nicht weiter zunimmt.

Das Einbeziehen der Straßendichte sollte ebenfalls Aufschluss über die Verkehrsemissionen liefern. In der Nähe von Straßen, deren Dichte z.B. durch `rd_type_5` (lokale Straßen) beschrieben ist, zeigt sich eine höhere  $\text{NO}_2$ -Belastung, was ebenfalls den Erwartungen entspricht. Die verschiedenen Straßendichten sind merklich untereinander und mit der Bevölkerungsdichte korreliert, weshalb ihre Wichtigkeit von der Feature Importance wahrscheinlich unterschätzt wird.

Ein Großteil des in der Troposphäre enthaltenen  $\text{NO}_2$  befindet sich in bodennähe [121], weshalb der troposphärische Stickstoffdioxidgehalt direkt mit der Bodenkonzentration korreliert. Der Vorhersagewert in Abb. 5.4 b) steigt dementsprechend bei verstärkten  $\text{NO}_2$ -Konzentrationen in den Satellitendaten.

Einen wichtigen Hinweis darauf, wieviel des troposphärischen  $\text{NO}_2$  am Boden vorhanden ist, liefert wie schon bei  $\text{PM}_{2.5}$  die Höhe der Grenzschicht. Auch hier erlaubt eine hohe BLH eine bessere Verteilung des Schadstoffs und sorgt so für eine Verringerung der Bodenkonzentration, was mit dem Verhalten in Abb. 5.4 c) übereinstimmt.

Die fünftwichtigste Variable ist durch den NDVI gegeben, welcher eine Verringerung der  $\text{NO}_2$ -Werte für höhere Bedeckungen mit Vegetation zeigt. Dies könnte einerseits der Fall sein, da der NDVI negativ mit der Bevölkerungsdichte korreliert ist, andererseits kann Vegetation auch direkt als Senke für  $\text{NO}_2$  dienen, da Pflanzen während der Photosynthese  $\text{NO}_2$  über die Blätter in ihren Stoffwechsel aufnehmen können und es so aus der Luft entfernen [122].

Die Beziehung zwischen Höhenlage (alt) und Schadstoffbelastung ist in Abb. 5.4 f) nicht eindeutig identifizierbar. Während (Boden-) Erhöhungen relativ zur umgebenden Geographie durch stärkeren Luftstrom die  $\text{NO}_2$ -Konzentration theoretisch verringern und lokale Absenkungen (Täler) die Wahrscheinlichkeit auf Inversionslagen erhöhen und den

Abtransport von Luftschadstoffen generell erschweren, lässt sich nur schwer eine Aussage über den Einfluss der absoluten geographischen Höhe treffen.

Insgesamt ist auffällig, dass zeitliche Variablen kaum eine Rolle in der Vorhersage von  $\text{NO}_2$  spielen, obwohl auch für  $\text{NO}_2$  ein Jahresgang in Abb. 5.15a erkennbar ist. Es ist möglich, dass dieser durch die saisonalen Änderungen im NDVI oder der Grenzschichthöhe bereits ausreichen charakterisiert ist. Da die zeitliche Kreuzvalidierung jedoch gute Ergebnisse liefert, kann zum einen vermutet werden, dass sich nur wenige Episoden im Untersuchungszeitraum finden, die sich durch Modell-relevante Besonderheiten in der Stickstoffdioxidkonzentration auszeichnen. Zum anderen ist der Transport von  $\text{NO}_2$  aufgrund der im Vergleich zu  $\text{PM}_{2.5}$  erhöhten Reaktivität und somit verringerten mittleren Lebensdauer [123, 124], nur auf kleinen Skalen möglich. Dadurch mitteln sich diese Effekte über das Untersuchungsgebiet raus.

### $\text{O}_3$

Lufttemperatur und Sonneneinstrahlung bestimmen im Wesentlichen, wie schnell chemische und photochemischen Prozesse in der Atmosphäre ablaufen. Durch das Fehlen direkter Emissionsquellen ist die Bodenkonzentration von Ozon direkt von diesen Größen abhängig. Die steigende partielle Abhängigkeit des Modells von diesen beiden wichtigsten Vorhersagevariablen (vgl. Abb. 5.5 b), 5.5 a)) beschreibt die Beschleunigung der entsprechenden Bildungsprozesse und ist somit physikalisch sinnvoll. Hinzu kommt, dass hohe Temperaturen die Freisetzung biogener und anthropogener VOCs fördern, was zusätzlich zur Erhöhung des Ozons beiträgt.

Als wichtiger Bestandteil der Ozon-Chemie ist  $\text{NO}_2$  ebenfalls unter den wichtigsten Eingangsvariablen zu finden. Die genaue Interpretation der partiellen Abhängigkeit wird in Kap. 6.4 näher erläutert.

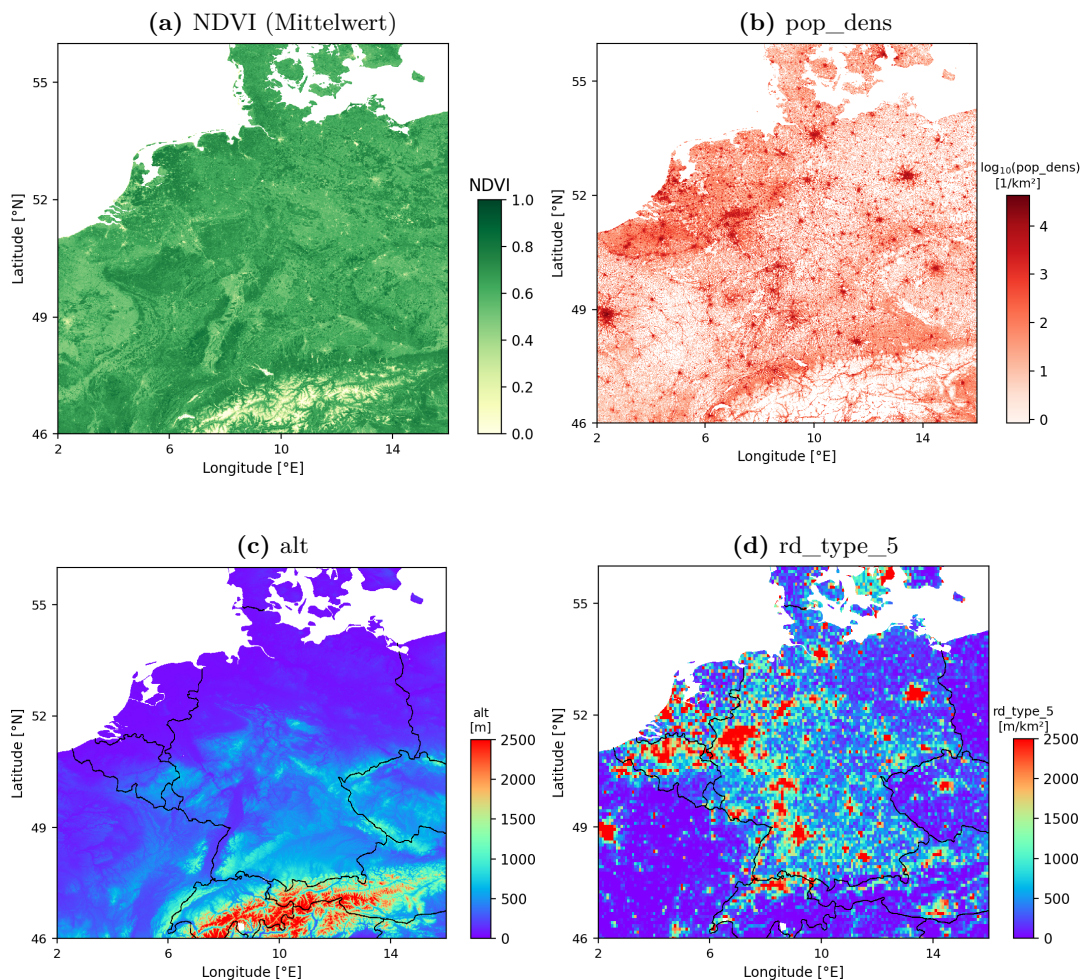
Es ist wahrscheinlich, dass sich die Abhängigkeit zwischen der Jahrestagsvariable (DoY) und  $\text{O}_3$  größtenteils aufgrund der saisonalen Variationen von  $\text{NO}_2$ , T2m, und SSR ergeben. Die Abhängigkeit von der Hilfsvariable lässt sonst auf keinen physikalischen Zusammenhang schließen.

Ebenfalls wichtig ist die relative Luftfeuchtigkeit, wobei die  $\text{O}_3$ -Konzentration des Modells mit steigender Luftfeuchtigkeit leicht absinken (vgl. Abb. 5.5 f)). Die in der Luft enthaltene Feuchtigkeit kann sich sowohl positiv als auch negativ auf die Ozonkonzentration auswirken. Bei der Produktion von Hydroxylradikalen durch die Reaktion (R 2.6) reagiert die Feuchtigkeit in der Luft mit Sauerstoffradikalen, welche sonst über (R 2.17) und (R 2.15) zur Bildung von Ozon beitragen könnten. Dadurch kann die Konzentration des bodennahen Ozons in speziell in Regionen mit geringen  $\text{NO}_x$ -Konzentrationen stark negativ mit der relativen Feuchtigkeit korreliert sein [125][126]. In stark verschmutzten Umgebungen, mit hohen  $\text{NO}_x$ -Konzentrationen, stellen sich konkurrierende Reaktionen ein, die den Ozon-reduzierenden Effekt der Luftfeuchtigkeit ausgleichen oder sogar Umkehren können [127][128]. Zum einen führt eine verstärkte Bildung von OH-Radikalen zu einer Beschleunigung der Ozonbildung aus VOCs (vgl. (R 2.23)-(R 2.26)). Zum anderen erhöht sich auch die Abbaurate der Vorläufersubstanz  $\text{NO}_2$  zu Salpetersäure nach Reaktion (R 2.7), was zu einer komplizierten Beziehung zwischen Feuchtigkeit und Ozonkonzentration führt. Der vom Modell erlernte  $\text{O}_3$ -RH-Zusammenhang ist unter bestimmten atmosphärischen Bedingungen denkbar, resultiert jedoch höchstwahrscheinlich aus der negativen Korrelation von relativer Luftfeuchtigkeit mit T2m und SSR.

Die Wichtigkeit der Grenzschichthöhe ist bei Ozon unter den untersuchten Schadstoffen am geringsten. Unter Vernachlässigung anderer Faktoren, wie des Transports aus höheren

Atmosphärenschichten, ist ein ähnlicher physikalischer Zusammenhang zwischen BLH und Ozonkonzentration am Boden zu erwarten wie für andere Schadstoffe. Dafür ist wieder die durch eine hohe Grenzschicht hervorgerufene bessere Durchmischung verantwortlich. Eine niedrige Grenzschicht sollte zudem zur Folge haben, dass sich VOCs und andere Vorläufersubstanzen ebenfalls stärker an der Oberfläche sammeln, wodurch die Ozonkonzentration erhöht wird. Da die BLH jedoch positiv mit der Sonneneinstrahlung und der Temperatur, sowie negativ mit der relativen Luftfeuchtigkeit korreliert, erlernt das Modell einen Zusammenhang, der den physikalischen Erwartungen widerspricht.

Generell hängt die Ozonvorhersage gemäß der Feature Importance nur geringfügig von Zeitvariablen (DoY, Month) und fast gar nicht von Ortsvariablen (pop\_dens) ab. Das bestätigen auch die Ergebnisse der zeitlichen bzw. räumlichen Kreuzvalidierung, die nur sehr geringe Verschlechterungen in der Modellgenauigkeit hervorrufen. Die Abhängigkeiten beschränken sich daher hauptsächlich auf die zugrundeliegenden Meteorologischen Bedingungen und die Konzentration von NO.



**Abbildung 6.1:** Darstellung einiger räumlicher Variablen: (a) NDVI als Mittelwert über den gesamten Untersuchungszeitraum, (b) Bevölkerungsdichte, (c) geographische Höhe und (d) Dichte lokaler Straßen.

## 6.3 Ursachen räumlicher und saisonaler Variabilität der Schadstoffverteilung

Die Verteilung der Schadstoffe über das Untersuchungsgebiet ergibt mehrere Hotspots und Coldspots die zum Teil auf menschliche, aber auch natürliche Ursachen zurückgeführt werden.

### PM<sub>2.5</sub>

Die mittlere Feinstaubbelastung der Bevölkerung liegt leicht über der Belastung der Untersuchungsfläche, was mit der beobachteten Erhöhung der PM<sub>2.5</sub>-Konzentration in Ballungsräumen übereinstimmt. Verantwortlich dafür sind menschenverursachte Emissionen wie z.B. aus Straßenverkehr und Energiewirtschaft. Die markantesten Variationen in der räumlichen Verteilung entstehen jedoch an naturdichten Regionen (vgl. Abb. 6.1a), was mit dem geringen Beitrag des NDVI zur Modellperformance im Widerspruch steht. Der NDVI wurde von S. Guo et al. [129] als einen der wichtigsten Einflussfaktoren auf PM<sub>2.5</sub> beschreiben. Vegetation kann dabei durch verschiedene Effekte eine mindernde Wirkung auf die Feinstaubkonzentration haben. Einerseits kommt es durch eine hohe Pflanzendichte zu einer verstärkten (trockenen) Disposition von PM<sub>2.5</sub>, was jedoch stark abhängig von der Art der Pflanzen ist [130]. Blattoberflächen von Bäumen können durch die gute Durchströmung der Baumkronen mehr Feinstaub aufnehmen als niedrig wachsende Pflanzen wie Gras [131]. Zusätzlich beeinflusst besonders hohe Vegetation den Transport von Feinstaub durch die Luft. Vor allem in urbanen Gebieten bilden Gewächse effektive Barrieren, die den Transportweg von PM<sub>2.5</sub> am Boden erschwert, wodurch sich starke PM<sub>2.5</sub>-Quellen von menschlichen Lebensräumen trennen lassen und so windabwärts liegende Regionen von hohen Belastungen bewahrt werden [132]. Wälder sind aus diesen Gründen bekannte Senken für atmosphärische Schadstoffe wie PM<sub>2.5</sub>, was sich in der räumlichen Verteilung der PM<sub>2.5</sub>-Konzentrationen widerspiegelt.

Generell kann beim Vergleich der Feinstaubkarten mit der geographischen Höhenlage festgestellt werden, dass sich Coldspots in der PM<sub>2.5</sub> Konzentration bevorzugt in bergigen Regionen finden. Dagegen ist für flache oder regional abgesenkte Gebiete wie Norddeutschland, die Beneluxländer, Nordtschechien oder entlang des Rhein tendenziell mehr PM<sub>2.5</sub> vorhergesagt. In den Alpen zeigt sich die flächen- und intensitätsmäßig größte Absenkung der Feinstaubbelastung (vgl. 5.8a). Neben der vergleichbar geringen Bevölkerungs- und Straßendichte in Gebirgsregionen (vgl. 6.1b und 6.1d) sorgen diese zusammen mit dichten Wäldern um die Region (vgl. Abb. 6.1a) für die geringe Luftverschmutzung mit PM<sub>2.5</sub>. Berge stellen bedeutende Hindernisse für den Luftstrom dar und verringern somit die interregionale Ausbreitung bzw. den Abtransport von Schadstoffen, was vor allem in Gebirgs- und Flusstälern zu einer Ansammlung von PM<sub>2.5</sub> führt [133]. Die hohe Belastung nördlich von Prag, welche über das gesamte Jahr zu beobachten ist kann unter anderem dadurch erklärt werden, dass hier ein Gebiet mit starken (wirtschaftlichen) Emissionen von Gebirgen umschlossen wird. Im Nordböhmischen Becken befinden sich mehrere Braunkohle-Tagebaustätten, und die entsprechenden Braunkohlekraftwerke<sup>2</sup> welche potentiell einen erheblichen Beitrag zur lokalen Feinstaubbelastung leisten.

Generell erhöhte Feinstaubwerte im Winter lassen sich neben der bereits erwähnten Saisonalität von meteorologischen Bedingungen wie BLH vor allem in Städten auf ver-

---

<sup>2</sup>Braunkohlekraftwerke in (Nord-)böhmischen Becken: Kraftwerk Prunéřov, Kraftwerk Chvaletice, Kraftwerk Mělník. Braunkohle-Tagebau im Nordböhmischen Becken: Tagebau Nástup-Tušimice, Tagebau Vršany, Tagebau lom Bílina.

stärkte Heizemissionen und verstärkten Kraftwerksbetrieb wegen erhöhtem Strombedarf zurückführen. Ostdeutschland und Polen sind vor allem in Herbst und im Winter verstärkt belastet. Mögliche Gründe hierfür sind die Lage einiger der emissionsreichsten Kraftwerke Europas in der Nähe der deutsch-polnischen Grenze [134], Agraremissionen oder möglicherweise Langstreckentransport von Feinstaub und Vorläufersubstanzen aus Osteuropa, was in Kap 6.7 weiter diskutiert wird.

## NO<sub>2</sub>

Da die Verweilzeit von NO<sub>2</sub> in der Troposphäre wesentlich geringer ausfällt als die von Feinstaubpartikeln und kaum natürliche Quellen vorhanden sind, sind die Hotspots für NO<sub>2</sub> stärker um die Ballungsräume lokalisiert als die des Feinstaubes. Sie sind, wie schon in Kap. 6.2 diskutiert, eine Folge der starken Verkehrsemissionen in den jeweiligen Regionen. Die große Bedeutung menschlicher Einflüsse als NO<sub>2</sub>-Prädiktor wird zudem durch die Ähnlichkeit der Verteilung des flächenhaften NO<sub>2</sub> mit der Bevölkerungs- und Straßendichte ersichtlich. Die mittlere Bevölkerungsbelastung fällt aufgrund der stark lokalisierten NO<sub>2</sub>-Hotspots deutlich höher aus als der Flächenmittelwert. Aus denselben Gründen wie schon bei PM<sub>2.5</sub> verringern Gebirge die NO<sub>2</sub>-Belastung, während Täler stärker verschmutzt sind. Außerdem stellt sich eine ähnliche Beziehung zum NDVI ein, wobei die Abbauprozesse von NO<sub>2</sub> durch Vegetation bereits in Kap. 6.2 erläutert wurden. Im Gegensatz zu PM<sub>2.5</sub> bleibt die Position der NO<sub>2</sub> Hot- und Coldspots über das Jahr relativ konstant, was aufgrund der Wichtigkeit der statischen Eingangsvariablen wie Bevölkerungs- und Straßendichte nicht verwunderlich ist. Dass in den Herbst- und Wintermonaten jedoch deutlich mehr Stickstoffdioxid und tendenziell größere Hotspots vorliegen, ist zum einen den geänderten meteorologischen Bedingungen geschuldet. Beispielhaft hierfür sind die niedrige Grenzschichthöhe oder die geringere Sonneneinstrahlung, welche das NO<sub>2</sub>-O<sub>3</sub>-Gleichgewicht durch eine Verlangsamung der photolytischen Prozesse zugunsten von NO<sub>2</sub> verschiebt. Zum anderen, kommt es im Winter zu erhöhtem Heiz- und Kraftwerkbetrieb, wobei durch kalte Temperaturen zusätzlich die Abgaseffizienz von Verbrennungsmotoren negativ beeinflusst wird, was eine Erhöhung der NO<sub>2</sub>-Belastung nach sich zieht.

## O<sub>3</sub>

Obwohl NO<sub>2</sub> als nötige Vorläufersubstanz von O<sub>3</sub> gilt, ist die tatsächliche Beziehung zwischen den Konzentrationen der beiden Gase nicht vollständig durch eine einfache Partielle Abhängigkeit (Abb. 5.5d) erklärbar. Das Zusammenspiel mit anderen Substanzen wie VOCs oder OH lässt viele meteorologische Einflüsse auf die Ozonkonzentration zu, wie die FI-Analyse für Ozon aufzeigt. Die stärkere Sonneneinstrahlung im Sommer beschleunigt zum einen die Bildungsprozesse von O<sub>3</sub> aus NO<sub>2</sub> nach Gl. (R 2.22) und erhöht so großflächig die Konzentration des Schadstoffs. Gleichzeitig wird die photolytische Bildung von atomarem Sauerstoff nach Gleichung (R 2.19) durch verstärkte Sonneneinstrahlung begünstigt. Dies führt zusammen mit dem erhöhten atmosphärischen Wassergehalt im Sommer nach (R 2.6) zu einer höheren Konzentration von OH-Radikalen, welche sich an der Ozonproduktion und am NO<sub>2</sub>-Abbau beteiligen. Des Weiteren begünstigen hohe Temperaturen und intensive Einstrahlung den Übergang von anthropogenen VOCs in die Atmosphäre, sowie die Freisetzung biogener VOCs durch Vegetation, welche zur chemischen Bildung von troposphärischem Ozon führen (siehe Gl. (R 2.23)-(R 2.26)). Mit der Sommersonnenwende (Sommeranfang) am 21.06. wird die Anzahl der mittleren täglichen Sonnenstunden maximal, wobei Juli und August für gewöhnlich die heißesten Monate des Jahres sind. Im Gegensatz dazu trifft zur Wintersommersonnenwende (21.12.) im Mittel

die geringste Strahlungsmenge auf die Erdoberfläche. Diese Faktoren erklären, weshalb sich im Frühling und im Sommer deutlich mehr Ozon in der unteren Atmosphäre befindet als im Winter oder Herbst.

Die flächenhafte Verteilung von Ozon aus der Modellvorhersage zeigt höhere Konzentrationen in ländlichen Regionen als in Städten und deckt sich mit den Messungen des Ozontrends in Abb. 2.6. Dies führt dazu, dass die mittlere Schadstoffbelastung pro Person für  $O_3$  unter dem Flächenmittelwert liegt. Im Gegensatz zum  $NO_2$ -Modell gehen in das Ozonmodell kaum räumliche Variablen wie Bevölkerungs- und Straßendichte ein, die dem Modell eine Einordnung eines Pixels nach urban oder rural ermöglichen. Der dennoch stark betonte Unterschiede zwischen Stadt und Land sind daher v. a. auf die ungleiche Verteilung von  $NO_2$  zurückzuführen. Durch die starken  $NO_x$ -Emissionen in den Ballungsräumen trägt möglicherweise der Titrationseffekt (vgl. Kap. 2.4) zur Verringerung der Ozonwerte in Städten bei. Dies erklärt ebenfalls, weshalb sich die Ozon-Coldspots in den Städten im Winter vergrößern, da zu dieser Jahreszeit die  $NO_x$ -Emissionen deutlich ansteigen und den Titrationseffekt und damit den Unterschied zwischen Stadt und Land verstärken. Auch wenn die Modellanalyse keine starken Zusammenhänge zwischen Ozon und NDVI vermuten lässt, da der Parameter nicht Teil der ausgewählten Variablen ist, kann vor allem in den südlichen Regionen ein räumlicher Zusammenhang zwischen NDVI und Ozonkonzentration beobachtet werden (vgl. Abb. 6.1a). Erhöhte Ozonwerte in vegetationsreichen Gebieten (Schwarzwald, Bayerischer Wald, Naturpark Vosges du Nord) und den Alpen im Winter und Herbst sind vermutlich ebenfalls auf die Schwächung des Titrationseffekts durch die geringeren  $NO_x$ -Emissionen an diesen Stellen verantwortlich. Die Umkehrung der Korrelation im Frühling und Sommer weist auf eine mögliche Änderung der  $NO_x$ - $O_3$ -Beziehung hin, welche in Kap. 6.4 näher erläutert wird. Vor allem zu den warmen Jahreszeiten stellt sich an Küstenregionen ein ausgeprägtes Land-See-Windsystem ein. Durch die am Tag wesentlich schnellere Erwärmung der Luftmassen über dem Land stellt sich zur untersuchten Tageszeit generell ein Seewind ein, der Seeluft ins Landesinnere transportiert. Mit Ausnahme von Emissionen aus der Schifffahrt ist diese Luft aufgrund fehlender Quellen frei von Vorläufersubstanzen wie (biogenen) VOCs und Stickoxiden, wodurch sich entlang der Nord- und Ostseeküsten des Untersuchungsgebiets besonders ozonarme Regionen herausbilden.

## 6.4 Bestimmung des Ozonproduktionsregimes

Ein zentraler Faktor in der Ozonbildung ist, gemäß der Feature Importance Analyse und theoretischen Betrachtung, die Verfügbarkeit von Vorläufersubstanzen wie  $NO_2$  und VOCs. Das Verhältnis zwischen  $NO_2$ -Konzentration und gebildetem Ozon ist, wie in Abb. 2.5 erkennbar, im Allgemeinen stark nichtlinear und zusätzlich abhängig vom Zusammenspiel mit weiteren reaktiven Stoffen in der Atmosphäre, sowie den meteorologischen Rahmenbedingungen. Es wird grob zwischen  $NO_x$ -gesättigten Bedingungen, unter denen zusätzliche Stickoxide eine Verringerung der Ozonkonzentrationen verursachen, und  $NO_x$ -sensiblen Bedingungen (VOC-gesättigt) unterschieden, die sich durch eine positive Korrelation der beiden Schadstoffe auszeichnen. Ist die Ableitung

$$\frac{d[O_3]}{d[NO_2]}$$

deutlich positiv, so bestehen  $NO_2$ -sensible Bedingungen, wobei bei deutlich negativen Werten eine  $NO_2$ -Sättigung vorliegt. Sillman et al. [135] setzten das Verhältnis der jeweiligen Konzentrationen  $[O_3]$  zu  $[NO_y]$  (reaktive Stickstoffspezies inklusive NO und  $NO_2$ ) zur

Bestimmung der Ozonproduktionsverhaltens ein. Die  $\text{NO}_2$ -Konzentrationen fallen aufgrund der erhöhten primären Emissionen durch menschliche Aktivitäten, wie z.B. Pendelverkehr, an den Wochenenden (Sa-So) merklich geringer aus als zu den üblichen Arbeitstagen (Mo-Fr). Effekte durch wechselnde Wetterbedingungen können bei Betrachtung des Wochenendeffekts durch Mittelung über hinreichend viele Untersuchungstage als Ursache für Ozonschwankungen ausgeschlossen werden. Somit eignet sich die aus dem Wochenendeffekt resultierende Änderung in der Ozon- und Stickstoffdioxidkonzentration gut zur Bestimmung der vorherrschenden Ordnung, was auch der Einsatz in bestehenden Studien zeigt [136, 137], denn sie liefern die Größen  $d[\text{O}_3]$  und  $d[\text{NO}_2]$ . Für die Jahreszeiten Winter und Herbst sind die Ergebnisse aus Kapitel 5.7 zum Teil durch den Titrationseffekt zu erklären, wiesen jedoch stark auf eine  $\text{NO}_x$ -Sättigung im Untersuchungsgebiet hin, da

$$\begin{aligned} \left( \frac{d[\text{O}_3]}{d[\text{NO}_2]} \right)_{\text{Winter}} &\approx \frac{[\text{O}_3]_{\text{Sa-So}} - [\text{O}_3]_{\text{Mo-Fr}}}{[\text{NO}_2]_{\text{Sa-So}} - [\text{NO}_2]_{\text{Mo-Fr}}} \\ &= \frac{60.80 \frac{\mu\text{g}}{\text{m}^3} - 59.94 \frac{\mu\text{g}}{\text{m}^3}}{10.16 \frac{\mu\text{g}}{\text{m}^3} - 11.20 \frac{\mu\text{g}}{\text{m}^3}} \\ &\approx -0.827 < 0 \end{aligned} \quad (6.1)$$

Im Sommer können anhand der partiellen Abhängigkeit (Abb. 5.23c) beide Regime beobachtet werden, allerdings spricht die positive Reaktion der Ozonkonzentration auf die sinkenden  $\text{NO}_2$ -Werte am Wochenende für eine generelle Einordnung der Bedingungen als  $\text{NO}_x$ -sensibel:

$$\begin{aligned} \left( \frac{d[\text{O}_3]}{d[\text{NO}_2]} \right)_{\text{Sommer}} &\approx \frac{[\text{O}_3]_{\text{Sa-So}} - [\text{O}_3]_{\text{Mo-Fr}}}{[\text{NO}_2]_{\text{Sa-So}} - [\text{NO}_2]_{\text{Mo-Fr}}} \\ &= \frac{101.91 \frac{\mu\text{g}}{\text{m}^3} - 103.45 \frac{\mu\text{g}}{\text{m}^3}}{5.19 \frac{\mu\text{g}}{\text{m}^3} - 5.87 \frac{\mu\text{g}}{\text{m}^3}} \\ &\approx 2.26 > 0 . \end{aligned} \quad (6.2)$$

Die wahrscheinlichste Ursache für den Wechsel der Ordnung ist das Zusammenspiel zwischen den Konzentrationen von  $\text{NO}_x$  und VOCs, welches zu unterschiedlichen Beziehungen zwischen  $\text{NO}_2$  und Ozon führen kann. Die blaue Linie in Abb. 2.5 trennt die beiden Ordnungen nach dem vorliegenden Verhältnis voneinander. Die  $\text{NO}_2$ -Emissionen fallen im Sommer deutlich geringer aus, während gleichzeitig mehr VOCs und OH-Radikale freigesetzt werden (siehe Kap. 6.3), was ein höheres  $[\text{VOC}]/[\text{NO}_x]$ -Verhältnis und damit  $\text{NO}_x$ -sensible Bedingungen hervorruft. In den kalten Jahreszeiten ist dieses wiederum wesentlich geringer, was einen Wechsel zur  $\text{NO}_x$ -Sättigung hervorruft. Bestehende Studien, wie die von X. Jin et al. [138], Beekmann et al. [139] und Balamurugan et al. [19], beschreiben ähnliche Jahreszeitliche Änderung der Abhängigkeit von bodennahem Ozon in Europa mit  $\text{NO}_x$ -gesättigten Bedingungen im Winter und je nach Region gemischten Bedingungen im Sommer. X. Jin et al. [138] halten zudem einen Trend zu verstärkten  $\text{NO}_x$ -sensiblen Bedingungen fest, vermutlich hervorgerufen durch die Umsetzung von Maßnahmen zur Abgasreduzierung. In den Vorhersagedaten konnten keine klaren Unterschiede im Ozonbildungsverhalten zwischen ländlichen und urbanen Regionen festgestellt werden, obwohl aufgrund der deutlich erhöhten Stickstoffdioxidkonzentrationen in den Städten und biogenen VOCs-Emissionen in vegetationsreichen Gebieten ein durch  $\text{NO}_x$  begrenztes System auf dem Land wahrscheinlicher scheint. Daraus kann geschlussfolgert werden, dass der Einfluss von biogenen VOCs auf die Ozonbildung im Untersuchungsgebiet relativ gering ausfällt, und die VOC-Emissionen durch anthropogene Quellen in den Ballungsräumen ähnlich zunehmen wie die Stickoxidemissionen. Curci et al. [140] schätzen den Zuwachs von



O<sub>3</sub> durch biogene Emissionen in Europa im Sommer auf  $\sim 5\%$ , was diese Vermutung unterstützt, jedoch kann auch nicht ausgeschlossen werden, dass eine korrekte Differenzierung zwischen städtischen und ländlichen Bedingungen über die Möglichkeiten der verwendeten Modelle und Methoden hinausgeht.

## 6.5 Irreguläre Einflüsse auf die Schadstoffentwicklung

Um Modellvorhersagen besser interpretieren zu können und zukünftige Modelle zu verbessern, ist es wichtig festzustellen, inwiefern diese auf zeitlich oder räumlich isolierte Schwankungen in den Schadstoffemissionen oder den meteorologischen Rahmenbedingungen reagieren.

Ist die Vorhersage der bodennahen Schadstoffe besonders abhängig von nichtphysikalischen Variablen wie DoY, DoW, Year etc., so ist dies ein Hinweis darauf, dass die übrigen (physikalischen) Variablen die Schadstoffquellen und relevanten meteorologischen Rahmenbedingungen wohlmöglich nicht ausreichend beschreiben. Besonders die PM<sub>2.5</sub>-Vorhersage stützt sich in diesem Fall auf zeitlich nahe Schadstoffmessungen aus dem Modelltraining anstatt auf Satellitenmessungen der AOD oder Meteorologie, wie anhand der Feature Importance Analyse in Kap. 6.2 deutlich wird. Im Verlauf der durchschnittlichen PM<sub>2.5</sub>-Belastung des Untersuchungsgebiets finden sich viele Zeitabschnitte in denen die Konzentration deutlich unter bzw. über dem für den Jahresabschnitt typischen Durchschnitt liegt. Diese decken sich zu großen Teilen mit Episoden mit besonders viel bzw. besonders wenig Niederschlag (siehe Abb. 5.12), welcher unter anderem für das Auswaschen von Feinstaub aus der Luft verantwortlich ist und ein Wiederaufwirbeln der Partikel auf nassen Oberflächen unterbindet (siehe Kapitel 2.2). Da bei Regenwetter die Satellitenbeobachtung durch Wolken verhindert wird, liegen während dieser Ereignisse keine Trainingsdaten für die Modelle vor, weshalb dieser starke Zusammenhang zwischen Niederschlag und Feinstaubkonzentration nicht in der Feature Importance bzw. den PDPs zu finden ist. Der Mangel an Trainingsdaten zu Niederschlagsmengen und die dennoch starke zeitliche Korrelation mit Regenereignissen ist ein weiterer Hinweis darauf, dass das PM<sub>2.5</sub>-Modell die fehlenden Informationen zu wichtigen Merkmalen durch die Zeitvariablen DoY und Year ausgleicht (siehe Abschnitt 6.2). Der Niederschlag ist daher trotz seiner geringen FI als eine der wichtigsten bestimmenden Faktoren für PM<sub>2.5</sub> anzusehen. In bestehenden Untersuchungen zu Einflussfaktoren von PM<sub>2.5</sub> wurde dieser ebenfalls als besonders wichtig und als stark reduzierend für die Feinstaubkonzentration befunden [141, 142, 143, 144]. Für NO<sub>2</sub> lassen sich in Abb. 5.14 ebenfalls Ähnlichkeiten zwischen der Abweichung der Schadstoffkonzentration und der (negativen) Niederschlagsmenge feststellen. Nasse Deposition von NO<sub>2</sub> stellt den treibenden Faktor hinter diesem Zusammenhang dar. Dabei löst sich NO<sub>2</sub> in Regentropfen und wird mit diesen zu Boden getragen, was es effektiv aus der Atmosphäre entfernt. Regenfälle sind demnach zuverlässige Indikatoren für eine Reduzierung der NO<sub>2</sub>-Bodenkonzentration, was ebenfalls von bestehenden Studien bestätigt wird [145, 146]. Dabei bestehen geringfügig konkurrierende Effekte in Form von erhöhter NO<sub>2</sub>-Produktion durch Blitze während Gewittern. Diese werden zwar in der Regel durch den auswaschenden Effekt von Regenfällen überschattet [145], können aber einen Beitrag zur Ozonproduktion leisten [147].

Während Trockenepisoden erhöht sich zudem die Gefahr für Waldbrände, welche potenziell viel PM<sub>2.5</sub> und NO<sub>2</sub> freisetzen. Ein Vergleich der wöchentlichen Brandfläche mit den mittleren NO<sub>2</sub>-Vorhersagen (Siehe Abb. 5.13) zeigt keine nennenswerten Zusammenhänge, allerdings werden in den Monaten Februar und März der Jahre 2018 deutlich höhere Feinstaubwerte vorhergesagt die, wie in Abb. 5.11 erkenntlich, mit Brandereignissen ein-

hergehen. Bei der Betrachtung muss allerdings eingeräumt werden, dass auf mindestens genauso viele der in den Daten verzeichneten Brände keine merkliche Erhöhung des  $\text{PM}_{2.5}$  folgt, weshalb ein kausaler Zusammenhang zwischen den Ereignissen unwahrscheinlich ist oder dieser an weitere Bedingungen geknüpft ist. Kleinere Brände könnten außerdem zu wenig Schadstoffe emittieren, um im flächenmäßigen Mittelwert einen merklichen Unterschied zu machen, weshalb zusätzlich die jeweilige Brandstelle untersucht werden muss. Mithilfe der vom DWD bereitgestellten Kartendarstellungen des Waldbrandindex (Monatsmittel) [148] werden Orte und Monate mit besonders starken Bränden innerhalb Deutschlands mit den entsprechenden Vorhersagen für  $\text{PM}_{2.5}$  und  $\text{NO}_2$  sowie TCNO<sub>2</sub>-Daten abgeglichen. Auch hier kann kein eindeutiger Zusammenhang festgestellt werden.

Ein weiterer Einflussfaktor im Feinstaubaufkommen ist der gelegentlich nach Europa transportierte Saharastaub. Die Abb. C.1 [149] des DWD zeigt die Staubereignisse in Deutschland der letzten Jahre (am Hohenpeißenberg). Zwischen 2018 und 2022, ist vor allem der Monat März des Jahres 2022 von starken Staubereignissen betroffen. Diese überschneiden sich zeitlich mit der deutlichen Erhöhung der vorhergesagten  $\text{PM}_{2.5}$ -Bodenkonzentration (vgl. Abb. 5.11) und kommen somit ebenfalls als mögliche Ursache in Frage. Es ist unklar, wie genau sich der Saharastaub aufgrund seiner Zusammensetzung und vertikalen Verteilung in der Atmosphäre auf den Zusammenhang zwischen AOD und  $\text{PM}_{2.5}$  am Boden auswirkt. Innerhalb der Stationsmessungen ist eine ähnliche Erhöhung festzustellen, allerdings bleibt diese in den entsprechenden AOD-Werten aus. Dies legt nahe, dass die Unterschiede in der Vorhersage nicht auf Basis der Satellitendaten, sondern wenn dann durch zeitliche Zuordnung zu dem Ereignis zustande kommen oder eine Folge meteorologischer Bedingungen darstellen.

Innerhalb des Untersuchungszeitraums liegen außerdem die im Jahr 2020 im Zuge der COVID-19 Pandemie von vielen Regierungen eingeführten Lockdown Maßnahmen zur Kontaktbeschränkung mit dem Ziel die Ausbreitung des Coronavirus zu bremsen. Folglich haben sich mit dem Beginn der Maßnahmen im März 2020 auch die Aktivitätsmuster der betroffenen Bevölkerung und dadurch die damit verbundenen anthropogenen Emissionen stark verändert. Der Einfluss auf die Luftqualität ist in den minimalen Vorhersagewerten für  $\text{NO}_2$  und  $\text{PM}_{2.5}$  für das Jahr 2020 erkenntlich, was bereits von einer Vielzahl bestehender Studien beobachtet wurde [136, 150, 151, 152]. Die Reaktion der vorhergesagten Schadstoffwerte auf plötzliche Veränderungen in den Emissionen unter statistisch unveränderten meteorologischen Bedingungen zeigt, dass sich die generierten Schadstoffkarten potenziell dazu eignen, um die Effektivität von Umweltrichtlinien zu testen. Wenn sich beispielsweise durch das Einführen strengerer Abgasrichtlinien eine tatsächliche Verringerung der Emissionen einstellt, kann diese durch die Modelle in einer Verbesserung der Luftqualität beobachtet werden.

## 6.6 Einfluss des Feuchtegehalts der Atmosphäre

### $\text{PM}_{2.5}$

Der Wassergehalt der Atmosphäre spielt eine wichtige Rolle bei der Bildung sekundärer Aerosole. Hygroskopisches Wachstum (siehe, Kap. 6.2) beeinflusst den Wassergehalt und lässt die absolute Masse speziell von sekundären anorganischen Aerosolen wie Nitraten und Sulfaten ansteigen. Die Stationsmessungen geben dagegen die Trockenaerosolmasse wieder. Bei hohen relativen Flüchtigkeiten wird zudem verstärkt Ammoniumnitrat gebildet [153], da die Reaktionsrate der Reaktion (R 2.3b) von der OH-Konzentration abhängt. Deswegen ist generell mit einer positiven Korrelation zwischen dem Wassergehalt in

der Atmosphäre und der gemessenen  $\text{PM}_{2.5}$ -Konzentration zu rechnen. Ein Anstieg von  $\text{PM}_{2.5}$  wird in den Sommermonaten beobachtet, jedoch zeigt sich im Winter - und in abgeschwächter Form auch im Frühling - ein gegensätzlicher Zusammenhang. Viele Studien haben bereits die Beziehung zwischen  $\text{PM}_{2.5}$  und Luftfeuchtigkeit untersucht, wobei sowohl positive als auch negative Korrelationen festgestellt wurden [143, 144]. Die bedeutendsten saisonalen Unterschiede des bodennahen  $\text{PM}_{2.5}$  liegen neben der Stärke der Emissionen in der chemischen Zusammensetzung der sekundären anorganischen Aerosole [154]. Die unterschiedliche Reaktion des Modells auf die Luftfeuchtigkeit im Winter und im Sommer könnte damit im Zusammenhang stehen, allerdings konnte keine konkrete Erklärung auf Basis der theoretischen Betrachtungen gefunden werden.

## **$\text{NO}_2$**

Die  $\text{NO}_2$ -Konzentration zeigt für Frühling, Sommer und Winter ein qualitativ ähnliches Verhalten wie  $\text{PM}_{2.5}$ . Der bedeutendste Abbaumechanismus für  $\text{NO}_2$  besteht in der Umwandlung zu Salpetersäure über die Reaktionswege in Abb. 2.2. Die Erhöhung des Wassergehalts der Atmosphäre erleichtert die Bildung von OH-Radikalen nach Gleichung (R 2.6). Damit wird sowohl in Gleichung (R 2.7) als auch in Gleichung (R 2.13) der  $\text{NO}_2$ -Abbau durch feuchte Bedingungen beschleunigt, was zu einer Verringerung der Bodenkonzentration führen sollte. Im Winter kann beobachtet werden, dass sich die Modellvorhersage mit zunehmender Feuchtigkeit erhöht, jedoch zeigt sich im Sommer ein gegenteiliges Verhalten. Dies kann Ursache der Antikorrelation zwischen relativer Luftfeuchtigkeit und der Grenzschichthöhe sein (vgl. Abb. 5.7), durch die das Modell höhere Schadstoffbelastungen für feuchte Wetterlagen vorhersagt. Ferner liegt in den Sommermonaten ein Großteil des  $\text{NO}_x$  als NO vor und  $\text{NO}_2$ -Emissionen sind zu dieser Zeit generell geringer. Es ist daher möglich, dass der Einfluss des  $\text{NO}_2$ -Abbaueffekts durch die Feuchtigkeit nur im Winter besonders stark ausgeprägt ist und die Modellvorhersage aus diesem Grund im Sommer von den Erwartungen abweicht.

Eine Erhöhung des atmosphärischen Wassergehalts fördert die Wolkenbildung und geht daher in der Regel mit einer dichteren Wolkendecke über dem Untersuchungsgebiet einher, was auch an der geringeren Abdeckung für als feucht klassifizierte Tage erkennbar ist. Als Folge des Schattenwurfs durch die Wolkendecke nimmt die mittlere Bestrahlungsstärke am Boden ab. Da starke Sonneneinstrahlung die Lebensdauer von  $\text{NO}_2$  durch Beschleunigung der photolytischen Reaktionen erheblich verkürzt (vgl. Gl. (2.11)), könnte eine Erhöhung des Wassergehalts der Atmosphäre somit gerade im Sommer, einen Anstieg der  $\text{NO}_2$ -Konzentration zur Folge haben, wie auch in den Modellvorhersagen beobachtet wird.

## **$\text{O}_3$**

Die Auswirkungen auf die Ozonkonzentration sind dabei erheblich komplexer. Einerseits verstärkt die Sonneneinstrahlung die Bildung von Ozon aus  $\text{NO}_2$ , sowie die Freisetzung von VOCs. Andererseits läuft die Ozonbildung im Sommer, wie in Kap. 6.4 erläutert wurde, zum Teil unter  $\text{NO}_x$ -sensitiven Bedingungen ab und reagiert dann positiv auf den Anstieg der  $\text{NO}_2$ -Konzentration. Dazu kommt, dass bei starker Bestrahlung auch die Lebenszeit der  $\text{O}_3$ -Moleküle verringert wird, da diese nach Gleichung (R 2.16a) und (R 2.16b) abgebaut werden können. So ist eine Erhöhung der Ozonkonzentration durch teilweise Beschattung einer Fläche und der Erhalt des entstandenen Ozons durch Bewölkung denkbar, wodurch sich die positive Korrelation zur Feuchte erklären lässt.

## 6.7 Schadstofftransport

Neben der Position der Quellen und Senken bestimmt vor allem der Transport von Schadstoffen deren räumliche Verteilung. Bei geringen Windgeschwindigkeiten ist die Durchmischung der Luftmassen in der unteren Atmosphäre besonders schlecht, weshalb Schadstoffe länger in der Nähe ihrer Quellen verweilen und die Luftverschmutzung, vor allem in Gebieten mit starken Emissionen, steigt. Bei hohen Windgeschwindigkeiten ist die untere Atmosphäre besser durchmischt, was besonders in den Ballungsräumen zu einer Verringerung der vorliegenden Schadstoffwerte führen sollte. Die  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Modelle bestätigen dieses Verhalten durch die vorhergesagten Schadstoffwerte, welche für höhere Windgeschwindigkeiten deutlich geringer ausfallen und wodurch die Windgeschwindigkeit als eine der wichtigsten Einflussfaktoren hervorgeht. Die starke negative Korrelation zwischen  $\text{PM}_{2.5}$  und Windstärke bzw.  $\text{NO}_2$  und Windstärke findet sich ebenfalls in der Literatur [143, 155, 156, 157, 158]. Gegensätzlich zu den Modellergebnissen wurden bereits direkte positive Korrelationen zwischen Windgeschwindigkeit und bodennahem Ozon Tu et al. [159], Borhani et al. [160] und Yildizhan et al. [161] festgestellt. Allersdings spielen dabei die Windrichtung und Konzentration von Vorläufersubstanzen vor Ort und Windaufwärts eine entscheidende Rolle, weshalb die Beobachtungen je nach untersuchtem Gebiet stark unterschiedlich ausfallen können.

Auch wenn für verschieden starke mittlere Windgeschwindigkeiten offensichtlich Unterschiede in der vom Modell vorhergesagten Schadstoffkonzentration und -verteilung zu beobachten sind, fällt der Einfluss der Wind-Variable im RF-Modell nur gering aus. Sie wurde bei der Variablenauswahl für keines der Modelle als gute Vorhersagevariable erkannt. Eine mögliche Erklärung dafür ist die positive Korrelation zwischen Windstärke und BLD (vgl. Abb. 5.7). Außerdem ist, wie schon die Feuchte, eine hohe Windgeschwindigkeit mit einer dichteren Wolkendecke verbunden, wie anhand der Abdeckungen der Modelldaten in Abb. A.4, A.6 und A.8 erkennlich ist. Diese hat wiederum eine Reduzierung der Sonneneinstrahlung zur Folge, was die Ozonproduktion schwächt.

Durch den Abtransport von Schadstoffen kann die Luftqualität zwar lokal verringert werden, allerdings trägt dieser windabwärts der Quellen zur Erhöhung der Luftverschmutzung bei. Dies kann eindeutig durch das Verhalten des  $\text{PM}_{2.5}$ - und  $\text{NO}_2$ -Hotspots im Ruhrgebiet und die umliegenden Regionen beobachtet werden, die je nach Hauptwindrichtung unterschiedlich über die Umgebung ausgedehnt sind. An Tagen an denen ein Gebiet windabwärts der größten Städte liegt erfährt dieses eine deutlich höhere Luftverschmutzung als an anderen Tagen. Für Ozon ist der Effekt zu schwach, um eine merkliche Veränderung der Hotspots zu erkennen, die auf eine Abhängigkeit von der Windrichtung hinweisen könnte.

## 7 Zusammenfassung und Ausblick

Ziel dieser Arbeit war die Identifikation und Analyse der wichtigsten meteorologischen Einflussfaktoren auf die Konzentration von  $\text{PM}_{2.5}$ ,  $\text{NO}_2$  und  $\text{O}_3$  in Bodennähe. Zusätzlich sollte die räumlichen und zeitlichen Variabilität der Luftqualität über Deutschland und Teilen der angrenzenden Länder untersucht werden. Zu diesem Zweck wurden zunächst die fundamentalen Bildungs- und Abbauprozesse in der Atmosphäre sowie Grundlagen zur Satellitenmessung vermittelt (Kapitel 2). Daraufhin wurden die verarbeiteten Daten, sowie das eingesetzte maschinelle Lernmodell (ML-Modell) und die Analysemethoden vorgestellt (Kapitel 3 und 4). Zuletzt wurden die erhaltenen Ergebnisse präsentiert (Kapitel 5) und mit Hinsicht auf die zentralen Ziele der Arbeit interpretiert (Kapitel 6).

Es wurden drei Random Forest Modelle entwickelt, die zur Ableitung von bodennahen  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Konzentrationen über dem Untersuchungsgebiet dienten. Als Datengrundlage wurden Satellitendaten zur Aerosol Optischen Dicke (MAIAC) und troposphärischem  $\text{NO}_2$  (TROPOMI) mit Schadstoffinformationen aus in-situ Messungen, meteorologischen Variablen, sowie Daten zur menschlichen Präsenz und geographischen Beschaffenheit verwendet. Wie in ähnlichen Studien zuvor wurde für die Ozonvorhersage ein gestaffeltes Modell entwickelt, das die eigenen  $\text{NO}_2$ -Vorhersagen als Vorläufersubstanzen berücksichtigt.

Die Modelle wurden zunächst unterschiedlichen Testverfahren unterzogen, um ihre allgemeine Leistung und Genauigkeit zu beurteilen. Bei 10-facher Kreuzvalidierung konnten generell gute Bestimmtheitsmaße ( $R^2$ ) von 0.76, 0.79 und 0.91 für die  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Modelle erreicht werden. Teils geringere Genauigkeiten unter räumlicher und zeitlicher Kreuzvalidierung und Vergleiche mit anderen Studien lassen darauf schließen, dass die Wahl des Untersuchungsgebiets und des Untersuchungszeitraums die Modellleistung merklich beeinflussen kann.

Anhand von Feature Importance Analysen konnten die für die Genauigkeit der RF-Modelle bedeutendsten Eingangsvariablen identifiziert werden. Für die Feinstaubvorhersage stellten sich Zeitvariablen wie der Jahrestag, sowie die Satelliten-AOD als die wichtigsten Informationsquellen heraus. Dagegen stützt sich das  $\text{NO}_2$ -Modell auf räumliche Variablen zur örtlichen Bevölkerungsdichte und die entsprechenden Satellitendaten zum  $\text{NO}_2$ -Gehalt der troposphärischen Luftsäule. Für Ozon sind hauptsächlich meteorologische Parameter in Form von Sonneneinstrahlung und Lufttemperatur relevant, allerdings leistet auch die Menge des verfügbaren  $\text{NO}_2$  einen entscheidenden Beitrag zur Modellgenauigkeit. Mithilfe von Partial Dependence Plots (PDPs) konnten die von den Modellen erlernten Abhängigkeiten zwischen den Schadstoffkonzentrationen und den verarbeiteten Variablen visualisiert und untersucht werden. Die meisten Abhängigkeiten entsprechen dabei qualitativ den erwarteten physikalischen oder chemischen Prozessen in der Atmosphäre.

Die räumliche Variabilität betreffend, konnten durch das bilden lokaler Indikatoren räumlicher Autokorrelation (LISA) auf den flächenhaften Vorhersagen der Schadstoffe, Hotspots von  $\text{PM}_{2.5}$  und  $\text{NO}_2$  in urbanen und industriereichen Regionen identifiziert werden. Gebirge und vegetationsreiche Gebiete decken sich dabei mit lokalen Absenkungen in den Konzentrationen der beiden Schadstoffe. Gegenteilig führte der Titrationseffekt,

vor allem in Herbst und im Winter, zur Ausbildung von Ozon-Coldspots in den meisten größeren Städten und es besteht ferner ein Nord-Süd-Gradient in der Ozonverteilung mit tendenziell größeren Konzentrationen in südlich gelegenen Regionen.

Die saisonalen Schwankungen der Schadstoffmittelwerte über dem Untersuchungsgebiet stellen die stärkste zeitliche Variabilität der  $\text{PM}_{2.5}$ -,  $\text{NO}_2$ - und  $\text{O}_3$ -Konzentrationen dar. In den Wintermonaten weisen Feinstaub und Stickstoffdioxid die höchsten Bodenkonzentrationen auf, hervorgerufen durch verstärkte anthropogene Emissionen und als Folge einer tendenziell niedrigeren und stabileren atmosphärischen Grenzschicht, die zur Anreicherung von Schadstoffen in Bodennähe führt. Ozon zeigt dagegen die minimale Konzentration in den Herbst- und Wintermonaten mit einem klaren Maximum im Sommer, was die Folge höherer Temperatur und stärkerer Sonneneinstrahlung ist, welche die chemischen Prozesse zur Ozonbildung und die Freisetzung von Vorläufersubstanzen wie flüchtigen organischen Verbindungen (VOCs) beschleunigen. Durch Untersuchung des Wochenendeffekts und der zusätzlichen Analyse der partiellen Abhängigkeiten saisonal trainierter Modelle konnte auf dem betrachteten Gebiet ein genereller Übergang von VOC-sensitiven Bedingungen der Ozonproduktion in den Wintermonaten, hin zu geringfügig  $\text{NO}_x$ -sensitiven Bedingungen im Sommer festgestellt werden.

Zu den wichtigsten meteorologischen Einflussfaktoren gehören neben der bereits erwähnten Lufttemperatur und Intensität der Sonneneinstrahlung für  $\text{O}_3$  sowie Grenzschichthöhe für  $\text{PM}_{2.5}$  und  $\text{NO}_2$  auch die Windstärke. Starke Winde sorgen für eine gute Durchmischung der unteren Atmosphäre und reduzieren so die Bodenkonzentration aller drei Schadstoffe erheblich. Auch konnte durch eine Betrachtung des  $\text{PM}_{2.5}$ - und  $\text{NO}_2$ -Hotspots über dem Ruhrgebiet bei verschiedenen Windrichtungen der Transport von Schadstoffen in windabwärts gelegenen Regionen beobachtet werden. Daher gilt die Windrichtung je nach der relativen Lage zu starken Schadstoffemissionen, als einflussreicher meteorologischer Parameter.

Daten zur Niederschlagsmenge wurden durch die Feature Importance für alle Schadstoffe als relativ unwichtig befunden. Vergleiche der zeitlichen Verläufe der mittleren Niederschlagsmenge und des mittleren Feinstaubes auf dem Untersuchungsgebiet zeigen jedoch eine starke Antikorrelation zwischen den Daten auf. Wegen mangelnden Satellitendaten bei Regenfällen aufgrund von Wolkenbedeckung, gehen diese kaum in das Modelltraining mit ein. Dieser „Schönwetterbias“ des Modells ist eine der zentralen Herausforderungen beim Arbeiten mit Satellitendaten und ML-Modellen.

Eine weitere Herausforderung stellen die Abhängigkeiten zwischen den einzelnen Eingangsvariablen dar. Es konnte beobachtet werden, dass starke Korrelationen zwischen den Modellvariablen die Interpretation der Feature Importance Werte erschweren können. Durch sie lassen sich Veränderungen in den Schadstoffdaten nur schwer eindeutig einem der beteiligten meteorologischen Parametern zuschreiben.

Neben der Bewältigung der genannten Herausforderungen ist eine Erweiterung und Verbesserung der in dieser Arbeit eingesetzten Modelle und Analyseverfahren in der Zukunft erstrebenswert.

Die Schadstoffe  $\text{NO}_2$  und  $\text{O}_3$  aber vor allem  $\text{PM}_{2.5}$  werden aufgrund ihrer gesundheitlichen Relevanz in der Literatur bevorzugt diskutiert. Im Rahmen zukünftiger Arbeiten sind weitere Stoffe wie Schwefeldioxid ( $\text{SO}_2$ ), Kohlenstoffmonoxid ( $\text{CO}$ ) und Formaldehyd ( $\text{HCHO}$ ) jedoch ebenfalls von potentiell Interesse. Sie beteiligen sich zum einen erheblich an der Bildung sekundärer Aerosole [162, 163]. Zum anderen ist  $\text{HCHO}$  eng mit der Konzentration von nicht-Methan VOCs (NMVOCs) verbunden [164], die zusammen mit  $\text{NO}_x$ ,

---

CO und Methan ( $\text{CH}_4$ ) eine der wichtigsten Vorläufersubstanzen von troposphärischem Ozon darstellen [165][166]. Die genannten Stoffe sind als TROPOMI-Datenprodukt verfügbar [167], weshalb die Analyse und Modellumsetzung analog zu denen für  $\text{NO}_2$  in dieser Arbeit durchgeführt werden könnte. Zusätzlich könnten die abgeleiteten Schadstoffdaten als Vorläufersubstanzen in die Vorhersagemodelle für  $\text{PM}_{2.5}$  und  $\text{O}_3$  eingehen und ihr Einfluss entsprechend analysiert werden.

Auf Basis der Variablenabhängigkeiten und den Ergebnissen anderer wissenschaftlicher Studien [108, 112], wird außerdem davon ausgegangen, dass sich die Modellgenauigkeit durch die räumliche und zeitliche Faltung der Eingangsdaten noch verbessern lässt. Dabei könnten dem Modell für die Vorhersage der Schadstoffwerte an einem Gitterpunkt zusätzliche Daten zu Schadstoffquellen, Bodenmessungen oder Geographie aus der Umgebung übergeben werden. Diese könnten weiterführend unter Einbezug der Windrichtung verarbeitet werden, um den regionalen Transport von Schadstoffen zu berücksichtigen. Daten zu vorangegangenen Tagen oder meteorologischen Ereignissen beinhalten ebenfalls Informationen die potenziell für die Vorhersage nützlich sind. Viele der Schwächen des gewählten Modellaufbaus (z.B. fehlende Berücksichtigung des Transports, fehlende Trainingsdaten durch Wolkenbedeckung während Regenfällen, usw.), die durch die isolierte Betrachtung einzelner Gitterpunkte entstehen, könnten dadurch überwunden oder abgeschwächt werden. Die genaueren Modellvorhersagen könnten neben zuverlässigeren Schadstoffkarten einen noch besseren Einblick in die zugrundeliegenden atmosphärischen Prozesse liefern.





# Literatur

- [1] Umweltbundesamt (UBA). *Emission von Feinstaub der Partikelgröße PM<sub>2,5</sub>*. 2024. URL: <https://www.umweltbundesamt.de/daten/luft/luftschadstoff-emissionen-in-deutschland/emission-von-feinstaub-der-partikelgroesse-pm25> (besucht am 9. Feb. 2025).
- [2] World Health Organization (WHO). *Air Quality Guidelines*. 2005. URL: <https://iris.who.int/bitstream/handle/10665/107823/9789289021920-eng.pdf> (besucht am 10. Mai 2025).
- [3] United States Environmental Protection Agency (EPA). *Basic Information about NO<sub>2</sub>*. 2024. URL: <https://www.epa.gov/no2-pollution/basic-information-about-no2> (besucht am 10. Mai 2025).
- [4] Umweltbundesamt (UBA). *Ozon-Belastung*. 2024. URL: <https://www.umweltbundesamt.de/daten/luft/ozon-belastung> (besucht am 9. Feb. 2025).
- [5] World Health Organization (WHO). *illions of people still breathe unhealthy air: new WHO data*. 2022. URL: <https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data> (besucht am 10. Mai 2025).
- [6] United States Environmental Protection Agency (EPA). *Ecosystem Effects of Ozone Pollution*. 2024. URL: <https://www.epa.gov/ground-level-ozone-pollution/ecosystem-effects-ozone-pollution> (besucht am 10. Mai 2025).
- [7] M. M. M. F. Jion, J. N. Jannat, M. Y. Mia, M. A. Ali, M. S. Islam, S. M. Ibrahim, S. C. Pal, A. Islam, A. Sarker, G. Malafaia, M. Bilal und A. R. M. T. Islam. „A critical review and prospect of NO<sub>2</sub> and SO<sub>2</sub> pollution over Asia: Hotspots, trends, and sources“. In: *Science of The Total Environment* 876 (2023), S. 162851. ISSN: 0048-9697. DOI: <https://doi.org/10.1016/j.scitotenv.2023.162851>.
- [8] World Health Organization (WHO). *Ambient (outdoor) air pollution*. 2024. URL: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health) (besucht am 10. Mai 2025).
- [9] European Comission. *Zero Pollution Action Plan*. 2021. URL: [https://environment.ec.europa.eu/strategy/zero-pollution-action-plan\\_en](https://environment.ec.europa.eu/strategy/zero-pollution-action-plan_en) (besucht am 17. Mai 2025).
- [10] European Parleament. *DIRECTIVE 2008/50/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 May 2008 on ambient air quality and cleaner air for Europe*. 2015. URL: <https://faolex.fao.org/docs/pdf/eur80016.pdf> (besucht am 17. Mai 2025).
- [11] Deutsche Bundestag. *Kurzinformation - Umweltzonen: Ausgestaltung, Überwachung*. 2021. URL: <https://www.bundestag.de/resource/blob/851600/f425e9a35138c0f0290b17a323e1a223/WD-8-056-21-pdf-data.pdf> (besucht am 17. Mai 2025).

- [12] Deutsche Bundestag. *Dokumentation – Diesel-Fahrverbot*. 2018. URL: <https://www.bundestag.de/resource/blob/546680/c5d56b986d304e60b1ac7d33931b47ab/WD-7-007-18-pdf.pdf> (besucht am 17. Mai 2025).
- [13] European Environment Agency (EEA). *Europe's air quality status 2024*. 2024. URL: <https://www.eea.europa.eu/publications/europes-air-quality-status-2024> (besucht am 16. Mai 2025).
- [14] European Environment Agency (EEA). *Exceedance of air quality standards in Europe*. 2025. URL: <https://www.eea.europa.eu/en/analysis/indicators/exceedance-of-air-quality-standards> (besucht am 17. Mai 2025).
- [15] European Environment Agency (EEA). *European Air Quality Index: current air quality information at your finger tips*. 2017. URL: <https://www.eea.europa.eu/highlights/european-air-quality-index-current> (besucht am 16. Mai 2025).
- [16] T. Wagner, A. Richter, C. von Friedeburg, M. Wenig und U. Platt. „Case Studies for the Investigation of Cloud Sensitive Parameters as Measured by GOME“. In: *Sounding the Troposphere from Space: A New Era for Atmospheric Chemistry*. Hrsg. von P. Borrell, P. M. Borrell, J. P. Burrows und U. Platt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, S. 199–210. ISBN: 978-3-642-18875-6. DOI: 10.1007/978-3-642-18875-6\_24.
- [17] T.-H. Liu, F.-T. Jeng, H.-C. Huang, E. Berge und J. S. Chang. „Influences of initial conditions and boundary conditions on regional and urban scale Eulerian air quality transport model simulations“. In: *Chemosphere - Global Change Science* 3.2 (2001), S. 175–183. ISSN: 1465-9972. DOI: [https://doi.org/10.1016/S1465-9972\(00\)00048-9](https://doi.org/10.1016/S1465-9972(00)00048-9).
- [18] A. C. Just, K. B. Arfer, J. Rush, M. Dorman, A. Shtein, A. Lyapustin und I. Kloog. „Advancing methodologies for applying machine learning and evaluating spatiotemporal models of fine particulate matter (PM<sub>2.5</sub>) using satellite data over large regions“. In: *Atmospheric Environment* 239 (2020), S. 117649. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2020.117649>.
- [19] V. Balamurugan, J. Chen, A. Wenzel und F. N. Keutsch. „Spatiotemporal modeling of air pollutant concentrations in Germany using machine learning“. In: *Atmospheric Chemistry and Physics* 23.17 (2023), S. 10267–10285. DOI: 10.5194/acp-23-10267-2023.
- [20] J. Chen, H. Dong, Z. Zhang, B. Quan und L. Luo. „High-Spatiotemporal-Resolution Estimation of Ground-Level Ozone in China Based on Machine Learning“. In: *Atmosphere* 15.1 (2024). ISSN: 2073-4433. DOI: 10.3390/atmos15010034.
- [21] M. Lee, L. Lin, C.-Y. Chen, Y. Tsao, T.-H. Yao, M.-H. Fei und S.-H. Fang. „Forecasting Air Quality in Taiwan by Using Machine Learning“. In: *Scientific Reports* 10 (März 2020). DOI: 10.1038/s41598-020-61151-7.
- [22] E. Koçak. „Comprehensive evaluation of machine learning models for real-world air quality prediction and health risk assessment by AirQ+“. In: *Earth Science Informatics* 18 (Juni 2025). DOI: 10.1007/s12145-025-01941-7.
- [23] W. Y. Hong. „Meteorological variability and predictive forecasting of atmospheric particulate pollution“. In: *Scientific Reports* 14 (Jan. 2024). DOI: 10.1038/s41598-023-41906-8.

- 
- [24] N. Ahmad, C. Lin, A. K. H. Lau, J. Kim, T. Zhang, F. Yu, C. Li, Y. Li, J. C. H. Fung und X. Q. Lao. „Estimation of ground-level NO<sub>2</sub> and its spatiotemporal variations in China using GEMS measurements and a nested machine learning model“. In: *Atmospheric Chemistry and Physics* 24.16 (2024), S. 9645–9665. DOI: 10.5194/acp-24-9645-2024.
- [25] A. Fania, A. Monaco, E. Pantaleo, T. Maggipinto, L. Bellantuono, R. Cilli, A. Lacalamita, M. La Rocca, S. Tangaro, N. Amoroso und R. Bellotti. „Estimation of Daily Ground Level Air Pollution in Italian Municipalities with Machine Learning Models Using Sentinel-5P and ERA5 Data“. In: *Remote Sensing* 16 (März 2024), S. 1206. DOI: 10.3390/rs16071206.
- [26] L. Mampitiya, N. Rathnayake, L. P. Leon, V. Mandala, H. M. Azamathulla, S. Shelton, Y. Hoshino und U. Rathnayake. „Machine Learning Techniques to Predict the Air Quality Using Meteorological Data in Two Urban Areas in Sri Lanka“. In: *Environments* 10.8 (2023). ISSN: 2076-3298. DOI: 10.3390/environments10080141.
- [27] J. Handschuh, T. Erbertseder und F. Baier. „Systematic Evaluation of Four Satellite AOD Datasets for Estimating PM<sub>2.5</sub> Using a Random Forest Approach“. In: *Remote Sensing* 15.8 (2023). ISSN: 2072-4292. DOI: 10.3390/rs15082064.
- [28] J. Handschuh, T. Erbertseder, M. Schaap und F. Baier. „Estimating PM<sub>2.5</sub> surface concentrations from AOD: A combination of SLSTR and MODIS“. In: *Remote Sensing Applications: Society and Environment* 26 (2022), S. 100716. ISSN: 2352-9385. DOI: <https://doi.org/10.1016/j.rsase.2022.100716>.
- [29] J. Wei, W. Huang, Z. Li, W. Xue, Y. Peng, L. Sun und M. Cribb. „Estimating 1-km-resolution PM<sub>2.5</sub> concentrations across China using the space-time random forest approach“. In: *Remote Sensing of Environment* 231 (2019), S. 111221. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2019.111221>.
- [30] K. de Hoogh, A. Saucy, A. Shtein, J. Schwartz, E. A. West, A. Strassmann, M. Puhon, M. Rösli, M. Stafoggia und I. Kloog. „Predicting Fine-Scale Daily NO<sub>2</sub> for 2005–2016 Incorporating OMI Satellite Data Across Switzerland“. In: *Environmental Science & Technology* 53.17 (2019). PMID: 31415154, S. 10279–10287. DOI: 10.1021/acs.est.9b03107. eprint: <https://doi.org/10.1021/acs.est.9b03107>.
- [31] W. Wang, X. Liu, J. Bi und Y. Liu. „A machine learning model to estimate ground-level ozone concentrations in California using TROPOMI data and high-resolution meteorology“. In: *Environment International* 158 (2022), S. 106917. ISSN: 0160-4120. DOI: <https://doi.org/10.1016/j.envint.2021.106917>.
- [32] H. Li, Y. Yang, J. Jin, H. Wang, K. Li, P. Wang und H. Liao. „Climate-driven deterioration of future ozone pollution in Asia predicted by machine learning with multi-source data“. In: *Atmospheric Chemistry and Physics* 23.2 (2023), S. 1131–1145. DOI: 10.5194/acp-23-1131-2023.
- [33] Y.-C. Liang, Y. Maimury, A. H.-L. Chen und J. R. C. Juarez. „Machine Learning-Based Prediction of Air Quality“. In: *Applied Sciences* 10.24 (2020). ISSN: 2076-3417. DOI: 10.3390/app10249151.
- [34] K. L. Chan, E. Khorsandi, S. Liu, F. Baier und P. Valks. „Estimation of Surface NO<sub>2</sub> Concentrations over Germany from TROPOMI Satellite Observations Using a Machine Learning Method“. In: *Remote Sensing* 13.5 (2021). ISSN: 2072-4292. DOI: 10.3390/rs13050969.
-

- [35] M. Ghahremanloo, Y. Lops, Y. Choi und B. Yeganeh. „Deep Learning Estimation of Daily Ground-Level NO<sub>2</sub> Concentrations From Remote Sensing Data“. In: *Journal of Geophysical Research: Atmospheres* 126.21 (2021). e2021JD034925 2021JD034925, e2021JD034925. DOI: <https://doi.org/10.1029/2021JD034925>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JD034925>.
- [36] Z. Zhao, J. Wu, F. Cai, S. Zhang und Y.-G. Wang. „A hybrid deep learning framework for air quality prediction with spatial autocorrelation during the COVID-19 pandemic“. In: *Scientific Reports* 13.1 (Jan. 2023), S. 1015.
- [37] X. Liu, Y. Zhu, L. Xue, A. R. Desai und H. Wang. „Cluster-Enhanced Ensemble Learning for Mapping Global Monthly Surface Ozone From 2003 to 2019“. In: *Geophysical Research Letters* 49.11 (2022). e2022GL097947 2022GL097947, e2022GL097947. DOI: <https://doi.org/10.1029/2022GL097947>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL097947>.
- [38] N. Zaini, E. Lee Woen, A.-M. Ali Najah Ahmed, M. Abdul Malek und M. F. Chow. „PM<sub>2.5</sub> forecasting for an urban area based on deep learning and decomposition method“. In: *Scientific Reports* 12 (Okt. 2022). DOI: 10.1038/s41598-022-21769-1.
- [39] X. Cheng, W. Zhang, A. Wenzel und J. Chen. „Stacked ResNet-LSTM and CORAL model for multi-site air quality prediction“. In: *Neural Computing and Applications* 34 (Aug. 2022). DOI: 10.1007/s00521-022-07175-8.
- [40] N. Kadir, M. Faisal und F. Kurniawan. „Analysis of the Impact of Meteorological Factors on Predicting Air Quality in South Tangerang City using Random Forest Method“. In: *Applied Information System and Management (AISM)* 7 (Sep. 2024). DOI: 10.15408/aism.v7i2.38466.
- [41] G. T. H. Nguyen, L. T. La, H. Hoang-Cong und A. H. Le. „An exploration of meteorological effects on PM<sub>2.5</sub> air quality in several provinces and cities in Vietnam“. In: *Journal of Environmental Sciences* 145 (2024), S. 139–151. ISSN: 1001-0742. DOI: <https://doi.org/10.1016/j.jes.2023.07.020>.
- [42] Y. Liu, Y. Zhou und J. Lu. „Exploring the relationship between air pollution and meteorological conditions in China under environmental governance“. In: *Scientific Reports* 10 (Sep. 2020), S. 14518. DOI: 10.1038/s41598-020-71338-7.
- [43] A. Rodríguez-Sánchez, J. Santiago, M. Vivanco, B. Sanchez, E. Rivas, A. Martilli und F. Martín. „How do meteorological conditions impact the effectiveness of various traffic measures on NO<sub>x</sub> concentrations in a real hot-spot?“ In: *Science of The Total Environment* 954 (2024), S. 176667. ISSN: 0048-9697. DOI: <https://doi.org/10.1016/j.scitotenv.2024.176667>.
- [44] K. Do, M. Mahish, A. K. Yeganeh, Z. Gao, C. L. Blanchard und C. E. Ivey. „Emerging investigator series: a machine learning approach to quantify the impact of meteorology on tropospheric ozone in the inland southern California“. In: *Environ. Sci.: Atmos.* 3 (8 2023), S. 1159–1173. DOI: 10.1039/D2EA00077F.
- [45] C. Zhao, Z. Lin, L. Yang, M. Jiang, Z. Qiu, S. Wang, Y. Gu, W. Ye, Y. Pan, Y. Zhang, T. Wang, Y. Jia und Z. Chen. „A study on the impact of meteorological and emission factors on PM<sub>2.5</sub> concentrations based on machine learning“. In: *Journal of Environmental Management* 376 (2025), S. 124347. ISSN: 0301-4797. DOI: <https://doi.org/10.1016/j.jenvman.2025.124347>.

- 
- [46] J. Seinfeld und S. Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley, 2012. ISBN: 9781118591369.
- [47] D. Jacob. *Introduction to Atmospheric Chemistry*. Princeton University Press, 1999. ISBN: 9781400841547.
- [48] K. Laidler. *Chemical Kinetics*. Harper & Row, 1987. ISBN: 9780060438623.
- [49] R. S. Sokhi, N. Moussiopoulos, A. Baklanov, J. Bartzis, I. Coll, S. Finardi, R. Friedrich, C. Geels, T. Grönholm, T. Halenka, M. Ketzel, A. Maragkidou, V. Matthias, J. Moldanova, L. Ntziachristos, K. Schäfer, P. Suppan, G. Tsegas, G. Carmichael, V. Franco, S. Hanna, J.-P. Jalkanen, G. J. M. Velders und J. Kukkonen. „Advances in air quality research – current and emerging challenges“. In: *Atmospheric Chemistry and Physics* 22.7 (2022), S. 4615–4703. DOI: 10.5194/acp-22-4615-2022.
- [50] D. C. Blanchard und A. H. Woodcock. „THE PRODUCTION, CONCENTRATION, AND VERTICAL DISTRIBUTION OF THE SEA-SALT AEROSOL“. In: *Annals of the New York Academy of Sciences* 338.1 (1980), S. 330–347. DOI: <https://doi.org/10.1111/j.1749-6632.1980.tb17130.x>. eprint: <https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.1980.tb17130.x>.
- [51] J. Liu, J. H. Christensen, Z. Ye, S. Dong, C. Geels, J. Brandt, A. Nenes, Y. Yuan und U. Im. „Impact of meteorology and aerosol sources on PM<sub>2.5</sub> and oxidative potential variability and levels in China“. In: *Atmospheric Chemistry and Physics* 24.18 (2024), S. 10849–10867. DOI: 10.5194/acp-24-10849-2024.
- [52] Q. Wang, P. Wang, H. Chen, J. Ma, Y. Jia, C. Wang, L. Qiao, Q. Fu, A. Mellouki und L. Li. „Unraveling Contributions from Combustion, Secondary, Traffic, and Dust Sources through Particle Mass Size Distribution Measurement“. In: *Aerosol and Air Quality Research* 24.11 (2024), S. 240135. ISSN: 2071-1409. DOI: 10.4209/aaqr.240135.
- [53] C. Tomasi und A. Lupi. „Primary and Secondary Sources of Atmospheric Aerosol“. In: *Atmospheric Aerosols*. John Wiley & Sons, Ltd, 2017. Kap. 1, S. 1–86. ISBN: 9783527336449. DOI: <https://doi.org/10.1002/9783527336449.ch1>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527336449.ch1>.
- [54] G. P. Brasseur, J. J. Orlando und G. S. Tyndall. *Atmospheric chemistry and global change*. eng. Topics in environmental chemistry. OUP, 1999. ISBN: 9780195105216.
- [55] Umweltbundesamt (UBA). *Feinstaubbelastung in Deutschland (2009)*. 2009. URL: <https://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/3565.pdf> (besucht am 9. Feb. 2025).
- [56] F. Guo und S. Xie. „Formation Mechanisms of Secondary Sulfate and Nitrate PM<sub>2.5</sub>“. In: *Progress in Chemistry* 35.9, 1313 (2023), S. 1313–1326. DOI: 10.7536/PC221201.
- [57] L. Chen, Y. Sun, X. Wu, Y. Zhang, C. Zheng, X. Gao und K. Cen. „Unit-based emission inventory and uncertainty assessment of coal-fired power plants“. In: *Atmospheric Environment* 99 (2014), S. 527–535. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2014.10.023>.
- [58] A. Strobl. *Feinstaubkonzentration in Abhängigkeit des Niederschlages an baustellenbeeinflussten Immissionsmessstellen*. 2009. URL: [https://www.tirol.gv.at/fileadmin/themen/sicherheit/emissionen-sicherheitstechnik-anlagen/downloads/PM10\\_Reduktion\\_durch\\_Niederschlag.PDF](https://www.tirol.gv.at/fileadmin/themen/sicherheit/emissionen-sicherheitstechnik-anlagen/downloads/PM10_Reduktion_durch_Niederschlag.PDF) (besucht am 17. Mai 2025).
-

- [59] E. Pfannerstill. *Biogenic Emissions and Air Quality Impacts*. 2025. URL: <https://www.fz-juelich.de/en/research/research-fields/young-scientists/young-investigator-groups/biogenic-emissions-and-air-quality-impacts> (besucht am 20. Mai 2025).
- [60] nited States Environmental Protection Agency (EPA), Clean Air Technology Center. *Nitrogen Oxides (NO<sub>x</sub>), Why and How They Are Controlled*. 1999. URL: <https://www3.epa.gov/ttn/catc/dir1/fnoxdoc.pdf> (besucht am 11. Feb. 2025).
- [61] I. Lagzi, R. Mészáros, G. Gelybó und Á. Leelddoto. *Atmospheric Chemistry*. Jan. 2013.
- [62] Deutscher Wetterdienst (DWD). *Peroxy-Radikale - "Waschmittel" der Atmosphäre*. URL: [https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung\\_atmosphaere/spurengase/inh\\_nav/rox.html](https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung_atmosphaere/spurengase/inh_nav/rox.html) (besucht am 9. Feb. 2025).
- [63] Y. Chen, R. Wolke, L. Ran, W. Birmili, G. Spindler, W. Schröder, H. Su, Y. Cheng, I. Tegen und A. Wiedensohler. „A parameterization of the heterogeneous hydrolysis of N<sub>2</sub>O<sub>5</sub> for mass-based aerosol models: improvement of particulate nitrate prediction“. In: *Atmospheric Chemistry and Physics* 18.2 (2018), S. 673–689. DOI: 10.5194/acp-18-673-2018.
- [64] Umweltbundesamt (UBA). *Stickstoffoxid-Emissionen*. 2024. URL: <https://www.umweltbundesamt.de/daten/luft/luftschadstoff-emissionen-in-deutschland/stickstoffoxid-emissionen> (besucht am 9. Feb. 2025).
- [65] L. Han, S. Cai, M. Gao, J.-y. Hasegawa, P. Wang, J. Zhang, L. Shi und D. Zhang. „Selective Catalytic Reduction of NO<sub>x</sub> with NH<sub>3</sub> by Using Novel Catalysts: State of the Art and Future Prospects“. In: *Chemical Reviews* 119 (Aug. 2019). DOI: 10.1021/acs.chemrev.9b00202.
- [66] G. Michalski, S. K. Bhattacharya und G. Girsch. „NO<sub>x</sub> cycle and the tropospheric ozone isotope anomaly: an experimental investigation“. In: *Atmospheric Chemistry and Physics* 14.10 (2014), S. 4935–4953. DOI: 10.5194/acp-14-4935-2014.
- [67] R. A. Borders und J. W. Birks. „High-precision measurements of activation energies over small temperature intervals: curvature in the Arrhenius plot for the reaction nitric oxide + ozone .fwdarw. nitrogen dioxide + oxygen“. In: *The Journal of Physical Chemistry* 86.17 (1982), S. 3295–3302. DOI: 10.1021/j100214a007. eprint: <https://doi.org/10.1021/j100214a007>.
- [68] P. C. Moonen, J. N. Cape, R. L. Storeton-West und R. McCOLM. „Measurement of the NO + O<sub>3</sub> Reaction Rate at Atmospheric Pressure Using Realistic Mixing Ratios“. In: *Journal of Atmospheric Chemistry* 29.3 (März 1998), S. 299–314.
- [69] J. W. Birks, B. Shoemaker, T. J. Leck und D. M. Hinton. „Studies of reactions of importance in the stratosphere. I. Reaction of nitric oxide with ozone“. In: *The Journal of Chemical Physics* 65.12 (Dez. 1976), S. 5181–5185. ISSN: 0021-9606. DOI: 10.1063/1.433059. eprint: [https://pubs.aip.org/aip/jcp/article-pdf/65/12/5181/18904181/5181\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/65/12/5181/18904181/5181_1_online.pdf).
- [70] G. W. Ray und R. T. Watson. „Kinetics of the reaction nitric oxide + ozone .fwdarw. nitrogen dioxide + oxygen from 212 to 422 K“. In: *The Journal of Physical Chemistry* 85.12 (1981), S. 1673–1676. DOI: 10.1021/j150612a015. eprint: <https://doi.org/10.1021/j150612a015>.

- 
- [71] D. Parrish, P. Murphy, D. Albritton und F. Fehsenfeld. „The measurement of the photodissociation rate of NO<sub>2</sub> in the atmosphere“. In: *Atmospheric Environment* (1967) 17.7 (1983), S. 1365–1379. ISSN: 0004-6981. DOI: [https://doi.org/10.1016/0004-6981\(83\)90411-0](https://doi.org/10.1016/0004-6981(83)90411-0).
- [72] V. P. Aneja, D.-S. Kim und W. L. Chameides. „Trends and analysis of ambient NO, NO<sub>y</sub>, CO, and ozone concentrations in raleigh, north carolina“. In: *Chemosphere* 34.3 (1997), S. 611–623. ISSN: 0045-6535. DOI: [https://doi.org/10.1016/S0045-6535\(96\)00393-1](https://doi.org/10.1016/S0045-6535(96)00393-1).
- [73] Bayerisches Landesamt für Umwelt (LfU). *Bodennahes Ozon und Sommersmog*. 2020. URL: [https://www.lfu.bayern.de/buerger/doc/uw\\_47\\_bodennahes\\_ozon.pdf](https://www.lfu.bayern.de/buerger/doc/uw_47_bodennahes_ozon.pdf) (besucht am 9. Feb. 2025).
- [74] S. Sillman und D. He. „Some theoretical results concerning O<sub>3</sub>-NO<sub>x</sub>-VOC chemistry and NO<sub>x</sub>-VOC indicators“. In: *Journal of Geophysical Research: Atmospheres* 107.D22 (2002), ACH 26-1-ACH 26-15. DOI: <https://doi.org/10.1029/2001JD001123>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2001JD001123>.
- [75] S. Sillman. *Overview: Tropospheric ozone, smog and ozone-NO<sub>x</sub>-VOC sensitivity*. URL: <https://public.websites.umich.edu/~sillman/ozone.htm> (besucht am 9. Feb. 2025).
- [76] Deutscher Wetterdienst (DWD). *Deutscher Wetterdienst, Bodennahes Ozon (O<sub>3</sub>)*. URL: [https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung\\_atmosphaere/spurengase/inh\\_nav/o3\\_node.html](https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung_atmosphaere/spurengase/inh_nav/o3_node.html) (besucht am 9. Feb. 2025).
- [77] Umweltbundesamt (UBA). *Luftschadstoff-Emissionen in Deutschland*. 2024. URL: <https://www.umweltbundesamt.de/daten/luft/luftschadstoff-emissionen-in-deutschland> (besucht am 9. Feb. 2025).
- [78] Deutschen Wetterdienst (DWD). *Aerosol Optische Dicke*. URL: [https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung\\_atmosphaere/aerosol/inh\\_nav/aod\\_node.html](https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung_atmosphaere/aerosol/inh_nav/aod_node.html) (besucht am 19. Mai 2025).
- [79] L. A. Remer, Y. J. Kaufman, D. Tanré, S. Mattoo, D. A. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. C. Levy, R. G. Kleidman, T. F. Eck, E. Vermote und B. N. Holben. „The MODIS Aerosol Algorithm, Products, and Validation“. In: *Journal of the Atmospheric Sciences* 62.4 (2005), S. 947–973. DOI: 10.1175/JAS3385.1.
- [80] European Space Agency (ESA). *GOME-2 overview*. URL: [https://www.esa.int/var/esa/storage/images/esa\\_multimedia/images/2002/01/measurement\\_principle\\_gome-2\\_collecting\\_solar\\_radiation/9202130-5-eng-GB/Measurement\\_principle\\_GOME-2\\_collecting\\_solar\\_radiation\\_pillars.jpg](https://www.esa.int/var/esa/storage/images/esa_multimedia/images/2002/01/measurement_principle_gome-2_collecting_solar_radiation/9202130-5-eng-GB/Measurement_principle_GOME-2_collecting_solar_radiation_pillars.jpg) (besucht am 18. Mai 2025).
- [81] European Space Agency (ESA). *Tropomi*. URL: [https://www.esa.int/Applications/Observing\\_the\\_Earth/Copernicus/Sentinel-5P/Tropomi](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Tropomi) (besucht am 15. Mai 2025).
- [82] J. van Geffen, K. F. Boersma, H. Eskes, M. Sneep, M. ter Linden, M. Zara und J. P. Veefkind. „S5P TROPOMI NO<sub>2</sub> slant column retrieval: method, stability, uncertainties and comparisons with OMI“. In: *Atmospheric Measurement Techniques* 13.3 (2020), S. 1315–1335. DOI: 10.5194/amt-13-1315-2020.

- [83] H. J. van Geffen J. Eskes, K. F. Boersma, H. Eskes und J. P. Veefkind. *TROPOMI ATBD of the total and tropospheric NO<sub>2</sub> data products*. 2022. URL: <https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-NO2-data-products> (besucht am 14. Mai 2025).
- [84] University of Bremen. *DOAS – Airmass Factor Calculator*. 2018. URL: <https://www.iup.uni-bremen.de/doas/airmassfactors.htm> (besucht am 14. Mai 2025).
- [85] J. Jin, B. Henzing und A. Segers. „How aerosol size matters in aerosol optical depth (AOD) assimilation and the optimization using the Ångström exponent“. In: *Atmospheric Chemistry and Physics* 23.2 (2023), S. 1641–1660. DOI: 10.5194/acp-23-1641-2023.
- [86] C. Zheng, C. Zhao, Y. Zhu, Y. Wang, X. Shi, X. Wu, T. Chen, F. Wu und Y. Qiu. „Analysis of influential factors for the relationship between PM<sub>2.5</sub> and AOD in Beijing“. In: *Atmospheric Chemistry and Physics* 17.21 (2017), S. 13473–13489. DOI: 10.5194/acp-17-13473-2017.
- [87] F. J. Gallego Pinilla. „A population density grid of the European Union“. In: *Population and Environment* 31 (Juli 2010), S. 460–473. DOI: 10.1007/s11111-010-0108-y.
- [88] D. Sulla-Menashe, J. M. Gray, S. P. Abercrombie und M. A. Friedl. „Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product“. In: *Remote Sensing of Environment* 222 (2019), S. 183–194. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2018.12.013>.
- [89] S. Liu, P. Valks, G. Pinardi, J. Xu, K. L. Chan, A. Argyrouli, R. Lutz, S. Beirle, E. Khorsandi, F. Baier, V. Huijnen, A. Bais, S. Donner, S. Dörner, M. Gratsea, F. Hendrick, D. Karagkiozidis, K. Lange, A. J. M. PETERS, J. Remmers, A. Richter, M. Van Roozendaal, T. Wagner, M. Wenig und D. G. Loyola. „An improved TROPOMI tropospheric NO<sub>2</sub> research product over Europe“. In: *Atmospheric Measurement Techniques* 14.11 (2021), S. 7297–7327. DOI: 10.5194/amt-14-7297-2021.
- [90] European Space Agency (ESA). *SentiWiki: Sentinel-5p Mission overview*. URL: <https://sentiwiki.copernicus.eu/web/s5p-mission> (besucht am 1. März 2025).
- [91] European Environment Agency (EEA). *Messstationsdaten*. (Ab dem 31.12.2024: <https://eeadmz1-downloads-webapp.azurewebsites.net/>). URL: <https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm> (besucht am 23. Dez. 2024).
- [92] European Centre for Medium-Range Weather Forecasts (ECMWF). *Atmospheric Model high resolution 15-day forecast (Set I - HRES)*. URL: <https://www.ecmwf.int/en/forecasts/datasets/set-i> (besucht am 1. März 2025).
- [93] J. Meijer, M. Huijbregts, K. Schotten und A. Schipper. „Global patterns of current and future road infrastructure“. In: *Environmental Research Letters* 13 (Juni 2018). DOI: 10.1088/1748-9326/aabd42.
- [94] R. Borck und P. Schrauth. „Population density and urban air quality“. In: *Regional Science and Urban Economics* 86 (2021), S. 103596. ISSN: 0166-0462. DOI: <https://doi.org/10.1016/j.regsciurbeco.2020.103596>.
- [95] A. Carioli, M. Schiavina, S. Freire und K. MacManus. *GHS-POP R2023A - GHS population grid multitemporal (1975-2030)*. Verfügbar: <https://geoservice.dlr.de/web/dataguide/tdm30/>. European Commission, Joint Research Centre (JRC), 2023.



- 
- [96] B. Wessel. *TanDEM-X Ground Segment – DEM Products Specification Document*. PID: <http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe>. EOC, DLR, Oberpfaffenhofen, Germany, 2013.
- [97] L. Breiman. „Random Forests“. In: *Machine Learning* 45 (Okt. 2001), S. 5–32. DOI: 10.1023/A:1010950718922.
- [98] X. Gao, Z. Ruan, J. Liu, Q. Chen und Y. Yuan. „Analysis of Atmospheric Pollutants and Meteorological Factors on PM<sub>2.5</sub> Concentration and Temporal Variations in Harbin“. In: *Atmosphere* 13.9 (2022). ISSN: 2073-4433. DOI: 10.3390/atmos13091426.
- [99] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot und É. Duchesnay. „Scikit-learn: Machine Learning in Python“. In: *Journal of Machine Learning Research* 12.85 (2011), S. 2825–2830.
- [100] Deutschen Wetterdienst (DWD). *Objektive Wetterlagenklassifikation*. URL: <https://www.dwd.de/DE/leistungen/wetterlagenklassifikation/wetterlagenklassifikation.html> (besucht am 5. März 2025).
- [101] Deutschen Wetterdienst (DWD). *Beschreibung des Verfahrens und der Wetterlagenklassen*. URL: <https://www.dwd.de/DE/leistungen/wetterlagenklassifikation/beschreibung.html> (besucht am 5. März 2025).
- [102] P. Bissolli und E. Dittmann. „The objective weather type classification of the German Weather Service and its possibilities of application to environmental and meteorological investigations“. In: *Meteorologische Zeitschrift* 10.4 (Okt. 2001), S. 253–260. DOI: 10.1127/0941-2948/2001/0010-0253.
- [103] Deutschen Wetterdienst (DWD). *WLK-Daten und zusätzliche Vorhersagedaten*. URL: [https://www.dwd.de/DE/leistungen/wetterlagenklassifikation/online\\_wlkvorhersage.txt](https://www.dwd.de/DE/leistungen/wetterlagenklassifikation/online_wlkvorhersage.txt) (besucht am 5. März 2025).
- [104] L. Anselin. „Local Indicators of Spatial Association—LISA“. In: *Geographical Analysis* 27.2 (1995), S. 93–115. DOI: <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4632.1995.tb00338.x>.
- [105] S. Rey und L. Anselin. „PySAL: A Python library of spatial analytical methods“. English (US). In: *Review of Regional Studies* 37.1 (2007), S. 5–27. ISSN: 0048-749X.
- [106] European Forest Fire Information System (EFFIS). *Statistics Portal: Seasonal Trends*. URL: <https://forest-fire.emergency.copernicus.eu/apps/effis.statistics/seasonaltrend> (besucht am 10. Mai 2025).
- [107] D. Chicco, M. Warrens und G. Jurman. „The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation“. In: *PeerJ Computer Science* 7 (Juli 2021), e623. DOI: 10.7717/peerj-cs.623.
- [108] X. Hu, J. H. Belle, X. Meng, A. Wildani, L. A. Waller, M. J. Strickland und Y. Liu. „Estimating PM<sub>2.5</sub> Concentrations in the Conterminous United States Using the Random Forest Approach“. In: *Environmental Science & Technology* 51.12 (2017). PMID: 28534414, S. 6936–6944. DOI: 10.1021/acs.est.7b01210. eprint: <https://doi.org/10.1021/acs.est.7b01210>.
-

- [109] S. Zhu, J. Tang, X. Zhou, P. Li, Z. Liu, C. Zhang, Z. Zou, T. Li und C. Peng. „Research progress, challenges, and prospects of PM<sub>2.5</sub> concentration estimation using satellite data“. In: *Environmental Reviews* 31.4 (2023), S. 605–631. DOI: 10.1139/er-2022-0125. eprint: <https://doi.org/10.1139/er-2022-0125>.
- [110] J. Handschuh, T. Erbertseder und F. Baier. „On the added value of satellite AOD for the investigation of ground-level PM<sub>2.5</sub> variability“. In: *Atmospheric Environment* 331 (2024), S. 120601. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2024.120601>.
- [111] M. Li, Y. Wu, Y. Bao, B. Liu und G. P. Petropoulos. „Near-Surface NO<sub>2</sub> Concentration Estimation by Random Forest Modeling and Sentinel-5P and Ancillary Data“. In: *Remote Sensing* 14.15 (2022). ISSN: 2072-4292. DOI: 10.3390/rs14153612.
- [112] E. Cao. „National ground-level NO<sub>2</sub> predictions via satellite imagery driven convolutional neural networks“. In: *Frontiers in Environmental Science* 11 (Dez. 2023). DOI: 10.3389/fenvs.2023.1285471.
- [113] J. Sun, J. Gong und J. Zhou. „Estimating hourly PM<sub>2.5</sub> concentrations in Beijing with satellite aerosol optical depth and a random forest approach“. In: *Science of The Total Environment* 762 (2021), S. 144502. ISSN: 0048-9697. DOI: <https://doi.org/10.1016/j.scitotenv.2020.144502>.
- [114] Z. Ndlovu. „Turbulent diffusion and air pollution: A comprehensive review of mechanisms, impacts, and modeling approaches“. In: *World Journal of Advanced Research and Reviews* 23 (Sep. 2024), S. 1511–1525. DOI: 10.30574/wjarr.2024.23.3.2784.
- [115] X. Zhang, Y. Yin, R. van der A, H. Eskes, J. van Geffen, Y. Li, X. Kuang, J. L. Lapierre, K. Chen, Z. Zhen, J. Hu, C. He, J. Chen, R. Shi, J. Zhang, X. Ye und H. Chen. „Influence of convection on the upper-tropospheric O<sub>3</sub> and NO<sub>x</sub> budget in southeastern China“. In: *Atmospheric Chemistry and Physics* 22.9 (2022), S. 5925–5942. DOI: 10.5194/acp-22-5925-2022.
- [116] M. Tursumbayeva, A. Kerimray, F. Karaca und D. A. Permadi. „Planetary Boundary Layer and its Relationship with PM<sub>2.5</sub> Concentrations in Almaty, Kazakhstan“. In: *Aerosol and Air Quality Research* 22.8 (2022), S. 210294. ISSN: 2071-1409. DOI: 10.4209/aaqr.210294.
- [117] P. Kalmus, C. O. Ao, K.-N. Wang, M. P. Manzi und J. Teixeira. „A high-resolution planetary boundary layer height seasonal climatology from GNSS radio occultations“. In: *Remote Sensing of Environment* 276 (2022), S. 113037. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2022.113037>.
- [118] L. Yang, H. Xu und S. Yu. „Estimating PM<sub>2.5</sub> Concentrations in Contiguous Eastern Coastal Zone of China Using MODIS AOD and a Two-Stage Random Forest Model“. In: *Journal of Atmospheric and Oceanic Technology* 38.12 (2021), S. 2071–2080. DOI: 10.1175/JTECH-D-20-0214.1.
- [119] W.-S. Won, R. Oh, W. Lee, S. Ku, P.-C. Su und Y.-j. Yoon. „Hygroscopic properties of particulate matter and effects of their interactions with weather on visibility“. In: *Scientific Reports* 11 (Aug. 2021). DOI: 10.1038/s41598-021-95834-6.
- [120] S. Sangpradid, T. Uttha und I. Aroonsri. „Estimates of PM<sub>2.5</sub> Concentration Based on Aerosol Optical Thickness Data Using Ensemble Learning with Support Vector Machine and Decision Tree“. In: *Environmental Research, Engineering and Management* 79 (Dez. 2023), S. 18–28. DOI: 10.5755/j01.erem.79.4.33913.

- 
- [121] R. P. Horner, E. A. Marais, N. Wei, R. G. Ryan und V. Shah. „Vertical profiles of global tropospheric nitrogen dioxide (NO<sub>2</sub>) obtained by cloud slicing the TROPOspheric Monitoring Instrument (TROPOMI)“. In: *Atmospheric Chemistry and Physics* 24.22 (2024), S. 13047–13064. DOI: 10.5194/acp-24-13047-2024.
- [122] D. J. Nowak, S. Hirabayashi, A. Bodine und E. Greenfield. „Tree and forest effects on air quality and human health in the United States“. In: *Environmental Pollution* 193 (2014), S. 119–129. ISSN: 0269-7491. DOI: <https://doi.org/10.1016/j.envpol.2014.05.028>.
- [123] S. Beirle, K. F. Boersma, U. Platt, M. G. Lawrence und T. Wagner. „Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space“. In: *Science* 333.6050 (2011), S. 1737–1739. DOI: 10.1126/science.1207824. eprint: <https://www.science.org/doi/pdf/10.1126/science.1207824>.
- [124] V. Shah, D. J. Jacob, K. Li, R. F. Silvern, S. Zhai, M. Liu, J. Lin und Q. Zhang. „Effect of changing NO<sub>x</sub> lifetime on the seasonality and long-term trends of satellite-observed tropospheric NO<sub>2</sub> columns over China“. In: *Atmospheric Chemistry and Physics* 20.3 (2020), S. 1483–1495. DOI: 10.5194/acp-20-1483-2020.
- [125] C. E. Johnson, W. J. Collins, D. S. Stevenson und R. G. Derwent. „Relative roles of climate and emissions changes on future tropospheric oxidant concentrations“. In: *Journal of Geophysical Research: Atmospheres* 104.D15 (1999), S. 18631–18645. DOI: <https://doi.org/10.1029/1999JD900204>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1999JD900204>.
- [126] X. Lu, L. Zhang, X. Yue, J. Zhang, D. A. Jaffe, A. Stohl, Y. Zhao und J. Shao. „Wildfire influences on the variability and trend of summer surface ozone in the mountainous western United States“. In: *Atmospheric Chemistry and Physics* 16.22 (2016), S. 14687–14702. DOI: 10.5194/acp-16-14687-2016.
- [127] S. Sillman und P. J. Samson. „Impact of temperature on oxidant photochemistry in urban, polluted rural and remote environments“. In: *Journal of Geophysical Research: Atmospheres* 100.D6 (1995), S. 11497–11508. DOI: <https://doi.org/10.1029/94JD02146>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/94JD02146>.
- [128] S. C. Kavassalis und J. G. Murphy. „Understanding ozone-meteorology correlations: A role for dry deposition“. In: *Geophysical Research Letters* 44.6 (2017), S. 2922–2931. DOI: <https://doi.org/10.1002/2016GL071791>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL071791>.
- [129] S. Guo, X. Tao und L. Liang. „Exploring Natural and Anthropogenic Drivers of PM<sub>2.5</sub> Concentrations Based on Random Forest Model: Beijing–Tianjin–Hebei Urban Agglomeration, China“. In: *Atmosphere* 14.2 (2023). ISSN: 2073-4433. DOI: 10.3390/atmos14020381.
- [130] S. S. Somvanshi und M. Kumari. „Comparative analysis of different vegetation indices with respect to atmospheric particulate pollution using sentinel data“. In: *Applied Computing and Geosciences* 7 (2020), S. 100032. ISSN: 2590-1974. DOI: <https://doi.org/10.1016/j.acags.2020.100032>.
-

- [131] D. Fowler, J. N. Cape, M. H. Unsworth, H. Mayer, J. M. Crowther, P. G. Jarvis, B. Gardiner und W. J. Shuttleworth. „Deposition of atmospheric pollutants on forests“. In: *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 324.1223 (1989), S. 247–265. DOI: [10.1098/rstb.1989.0047](https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.1989.0047). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.1989.0047>.
- [132] A. Diener und P. Mudu. „How can vegetation protect us from air pollution? A critical review on green spaces’ mitigation abilities for air-borne particles from a public health perspective - with implications for urban planning“. In: *Science of The Total Environment* 796 (2021), S. 148605. ISSN: 0048-9697. DOI: <https://doi.org/10.1016/j.scitotenv.2021.148605>.
- [133] Y. Wen, J. Xiao, J. Yang, S. Cai, M. Liang und P. Zhou. „Quantitatively Disentangling the Geographical Impacts of Topography on PM<sub>2.5</sub> Pollution in China“. In: *Remote Sensing* 14.24 (2022). ISSN: 2072-4292. DOI: [10.3390/rs14246309](https://doi.org/10.3390/rs14246309).
- [134] Ember-Energy. *EU power plant emissions - data tool*. 2023. URL: <https://ember-energy.org/latest-insights/eu-ets-2022/eu-power-plant-emissions-data-tool/> (besucht am 6. Juni 2025).
- [135] S. Sillman, R. Vautard, L. Menut und D. Kley. „O<sub>3</sub>-NO<sub>x</sub>-VOC sensitivity and NO<sub>x</sub>-VOC indicators in Paris: Results from models and Atmospheric Pollution Over the Paris Area (ESQUIF) measurements“. In: *Journal of Geophysical Research: Atmospheres* 108.D17 (2003). DOI: <https://doi.org/10.1029/2002JD001561>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2002JD001561>.
- [136] V. Balamurugan, J. Chen, Z. Qu, X. Bi, J. Gensheimer, A. Shekhar, S. Bhattacharjee und F. N. Keutsch. „Tropospheric NO<sub>2</sub> and O<sub>3</sub> Response to COVID-19 Lockdown Restrictions at the National and Urban Scales in Germany“. In: *Journal of Geophysical Research: Atmospheres* 126.19 (2021). e2021JD035440 2021JD035440, e2021JD035440. DOI: <https://doi.org/10.1029/2021JD035440>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JD035440>.
- [137] Y. H. Wang, B. Hu, D. S. Ji, Z. R. Liu, G. Q. Tang, J. Y. Xin, H. X. Zhang, T. Song, L. L. Wang, W. K. Gao, X. K. Wang und Y. S. Wang. „Ozone weekend effects in the Beijing–Tianjin–Hebei metropolitan area, China“. In: *Atmospheric Chemistry and Physics* 14.5 (2014), S. 2419–2429. DOI: [10.5194/acp-14-2419-2014](https://doi.org/10.5194/acp-14-2419-2014).
- [138] X. Jin, A. M. Fiore, L. T. Murray, L. C. Valin, L. N. Lamsal, B. Duncan, K. Folkert Boersma, I. De Smedt, G. G. Abad, K. Chance und G. S. Tonnesen. „Evaluating a Space-Based Indicator of Surface Ozone-NO<sub>x</sub>-VOC Sensitivity Over Midlatitude Source Regions and Application to Decadal Trends“. In: *Journal of Geophysical Research: Atmospheres* 122.19 (2017), S. 10, 439–10, 461. DOI: <https://doi.org/10.1002/2017JD026720>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD026720>.
- [139] M. Beekmann und R. Vautard. „A modelling study of photochemical regimes over Europe: robustness and variability“. In: *Atmospheric Chemistry and Physics* 10.20 (2010), S. 10067–10084. DOI: [10.5194/acp-10-10067-2010](https://doi.org/10.5194/acp-10-10067-2010).
- [140] G. Curci, M. Beekmann, R. Vautard, G. Smiatek, R. Steinbrecher, J. Theloke und R. Friedrich. „Modelling study of the impact of isoprene and terpene biogenic emissions on European ozone levels“. In: *Atmospheric Environment* 43.7 (2009). Natural and Biogenic Emissions of Environmentally Relevant Atmospheric Trace Constituents

- in Europe, S. 1444–1455. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2008.02.070>.
- [141] R. Wang, K. Cui, H.-L. Sheu, L.-C. Wang und X. Liu. „Effects of Precipitation on the Air Quality Index, PM<sub>2.5</sub> Levels and on the Dry Deposition of PCDD/Fs in the Ambient Air“. In: *Aerosol and Air Quality Research* 23.4 (2023), S. 220417. ISSN: 2071-1409. DOI: 10.4209/aaqr.220417.
  - [142] A. Ziernicka-Wojtaszek, Z. Zuśka und J. Kopcińska. „Assessment of the Effect of Meteorological Conditions on the Concentration of Suspended PM<sub>2.5</sub> Particulate Matter in Central Europe“. In: *Sustainability* 16.11 (2024). ISSN: 2071-1050.
  - [143] R. Zalakeviciute, J. López-Villada und Y. Rybarczyk. „Contrasted Effects of Relative Humidity and Precipitation on Urban PM<sub>2.5</sub> Pollution in High Elevation Urban Areas“. In: *Sustainability* 10.6 (2018). ISSN: 2071-1050.
  - [144] E. Zender-Świercz, B. Galiszewska, M. Telejko und M. Starzomska. „The effect of temperature and humidity of air on the concentration of particulate matter - PM<sub>2.5</sub> and PM<sub>10</sub>“. In: *Atmospheric Research* 312 (2024), S. 107733. ISSN: 0169-8095. DOI: <https://doi.org/10.1016/j.atmosres.2024.107733>.
  - [145] J.-M. Yoo, Y.-R. Lee, D. Kim, M.-J. Jeong, W. R. Stockwell, P. K. Kundu, S.-M. Oh, D.-B. Shin und S.-J. Lee. „New indices for wet scavenging of air pollutants (O<sub>3</sub>, CO, NO<sub>2</sub>, SO<sub>2</sub>, and PM<sub>10</sub>) by summertime rain“. In: *Atmospheric Environment* 82 (2014), S. 226–237. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2013.10.022>.
  - [146] A. Martin. „Estimated washout coefficients for sulphur dioxide, nitric oxide, nitrogen dioxide and ozone“. In: *Atmospheric Environment (1967)* 18.9 (1984). CACGP Symposium on Tropospheric Chemistry With Emphasis on Sulphur and Nitrogen Cycles and the Chemistry of Clouds and Precipitation, S. 1955–1961. ISSN: 0004-6981. DOI: [https://doi.org/10.1016/0004-6981\(84\)90373-1](https://doi.org/10.1016/0004-6981(84)90373-1).
  - [147] Y. Choi, J. Kim, A. Eldering, G. Osterman, Y. L. Yung, Y. Gu und K. N. Liou. „Lightning and anthropogenic NO<sub>x</sub> sources over the United States and the western North Atlantic Ocean: Impact on OLR and radiative effects“. In: *Geophysical Research Letters* 36.17 (2009). DOI: <https://doi.org/10.1029/2009GL039381>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL039381>.
  - [148] Deutschen Wetterdienst (DWD). *Deutscher Klimaatlas*. URL: [https://www.dwd.de/DE/klimaumwelt/klimaatlas/klimaatlas\\_node.html](https://www.dwd.de/DE/klimaumwelt/klimaatlas/klimaatlas_node.html) (besucht am 10. Mai 2025).
  - [149] Deutschen Wetterdienst (DWD). *Saharastaub*. URL: [https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung\\_atmosphaere/aerosol/inh\\_nav/saharastaub\\_node.html](https://www.dwd.de/DE/forschung/atmosphaerenbeob/zusammensetzung_atmosphaere/aerosol/inh_nav/saharastaub_node.html) (besucht am 1. März 2025).
  - [150] J. Barré, H. Petetin, A. Colette, M. Guevara, V.-H. Peuch, L. Rouil, R. Engelen, A. Inness, J. Flemming, C. Pérez García-Pando, D. Bowdalo, F. Meleux, C. Geels, J. H. Christensen, M. Gauss, A. Benedictow, S. Tsyro, E. Friese, J. Struzewska, J. W. Kaminski, J. Douros, R. Timmermans, L. Robertson, M. Adani, O. Jorba, M. Joly und R. Kouznetsov. „Estimating lockdown-induced European NO<sub>2</sub> changes using satellite and surface observations and air quality models“. In: *Atmospheric Chemistry and Physics* 21.9 (2021), S. 7373–7394. DOI: 10.5194/acp-21-7373-2021.

- [151] S. K. Grange, J. D. Lee, W. S. Drysdale, A. C. Lewis, C. Hueglin, L. Emmenegger und D. C. Carslaw. „COVID-19 lockdowns highlight a risk of increasing ozone pollution in European urban areas“. In: *Atmospheric Chemistry and Physics* 21.5 (2021), S. 4169–4185. DOI: 10.5194/acp-21-4169-2021.
- [152] S. Solberg, S.-E. Walker, P. Schneider und C. Guerreiro. „Quantifying the Impact of the Covid-19 Lockdown Measures on Nitrogen Dioxide Levels throughout Europe“. In: *Atmosphere* 12.2 (2021). ISSN: 2073-4433. DOI: 10.3390/atmos12020131.
- [153] D. J. Jacob und D. A. Winner. „Effect of climate change on air quality“. In: *Atmospheric Environment* 43.1 (2009). Atmospheric Environment - Fifty Years of Endeavour, S. 51–63. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2008.09.051>.
- [154] J. Duan, R.-J. Huang, Y. Li, Q. Chen, Y. Zheng, Y. Chen, C. Lin, H. Ni, M. Wang, J. Ovadnevaite, D. Ceburnis, C. Chen, D. R. Worsnop, T. Hoffmann, C. O’Dowd und J. Cao. „Summertime and wintertime atmospheric processes of secondary aerosol in Beijing“. In: *Atmospheric Chemistry and Physics* 20.6 (2020), S. 3793–3807. DOI: 10.5194/acp-20-3793-2020.
- [155] R. Harrison und J. Baker. „The wind speed dependence of the concentrations of airborne particulate matter and NO<sub>x</sub>“. In: *Atmospheric Environment* 44 (Apr. 2010), S. 1682–1690. DOI: 10.1016/j.atmosenv.2010.01.007.
- [156] T. Danek, E. Węglińska und M. Zareba. „The influence of meteorological factors and terrain on air pollution concentration and migration: a geostatistical case study from Krakow, Poland“. In: *Scientific Reports* 12 (Juni 2022), S. 1–22. DOI: 10.1038/s41598-022-15160-3.
- [157] X. Li, Y. Ma, Y. Wang, N. Liu und Y. Hong. „Temporal and spatial analyses of particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>) and its relationship with meteorological parameters over an urban city in northeast China“. In: *Atmospheric Research* 198 (2017), S. 185–193. ISSN: 0169-8095. DOI: <https://doi.org/10.1016/j.atmosres.2017.08.023>.
- [158] P. J. Asilevi, E. Dzidzorm, P. Boakye und E. Quansah. „Nitrogen dioxide (NO<sub>2</sub>) Meteorology and predictability for air quality management using TROPOMI Check for updates“. In: *npj Clean Air* 1 (März 2025). DOI: 10.1038/s44407-024-00003-4.
- [159] J. Tu, Z.-G. Xia, H. Wang und W. Li. „Temporal variations in surface ozone and its precursors and meteorological effects at an urban site in China“. In: *Atmospheric Research* 85.3 (2007), S. 310–337. ISSN: 0169-8095. DOI: <https://doi.org/10.1016/j.atmosres.2007.02.003>.
- [160] F. Borhani, M. Shafie-Pour-Motlagh, A. Stohl, Y. Rashidi und A. Ehsani. „Tropospheric Ozone in Tehran, Iran, during the last 20 years“. In: *Environmental Geochemistry and Health* 44 (Okt. 2021), S. 3615–3637. DOI: 10.1007/s10653-021-01117-4.
- [161] H. Yildizhan, M. Udristoiu, T. Pekdogan und A. Ameen. „Observational study of ground-level ozone and climatic factors in Craiova, Romania, based on one-year high-resolution data“. In: *Scientific Reports* 14 (Nov. 2024), S. 26733. DOI: 10.1038/s41598-024-77989-0.

- 
- [162] M. Claeys, B. Graham, G. Vas, W. Wang, R. Vermeylen, V. Pashynska, J. Cafmeyer, P. Guyon, M. O. Andreae, P. Artaxo und W. Maenhaut. „Formation of Secondary Organic Aerosols Through Photooxidation of Isoprene“. In: *Science* 303.5661 (2004), S. 1173–1176. DOI: 10.1126/science.1092805. eprint: <https://www.science.org/doi/pdf/10.1126/science.1092805>.
- [163] L. Xu, N. T. Tsona und L. Du. „Relative Humidity Changes the Role of SO<sub>2</sub> in Biogenic Secondary Organic Aerosol Formation“. In: *The Journal of Physical Chemistry Letters* 12.30 (2021). PMID: 34324359, S. 7365–7372. DOI: 10.1021/acs.jpcllett.1c01550. eprint: <https://doi.org/10.1021/acs.jpcllett.1c01550>.
- [164] G.-M. Oomen, J.-F. Müller, T. Stavrakou, I. De Smedt, T. Blumenstock, R. Kivi, M. Makarova, M. Palm, A. Röhling, Y. Té, C. Vigouroux, M. M. Friedrich, U. Frieß, F. Hendrick, A. Merlaud, A. Piters, A. Richter, M. Van Roozendaal und T. Wagner. „Weekly derived top-down volatile-organic-compound fluxes over Europe from TROPOMI HCHO data from 2018 to 2021“. In: *Atmospheric Chemistry and Physics* 24.1 (2024), S. 449–474. DOI: 10.5194/acp-24-449-2024.
- [165] S. Houweling, F. Dentener und J. Lelieveld. „The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry“. In: *Journal of Geophysical Research: Atmospheres* 103.D9 (1998), S. 10673–10696. DOI: <https://doi.org/10.1029/97JD03582>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/97JD03582>.
- [166] European Space Agency (ESA). *TROPOMI: Formaldehyde*. URL: <https://www.tropomi.eu/data-products/f/formaldehyde> (besucht am 4. Juni 2025).
- [167] European Space Agency (ESA). *TROPOMI: Data products*. URL: <https://www.tropomi.eu/data-products> (besucht am 4. Juni 2025).





# Abbildungsverzeichnis

2.1	Entwicklung der Menge des menschenproduzierten $\text{PM}_{2.5}$ und der Anteile der einzelnen Feinstaubquellen in Deutschland von 1995 bis 2022. Ein andauernder Abfall der gesamten Feinstaubemissionen und speziell der Emissionen durch Verkehr und Haushalte ist erkennbar. (Abbildung entnommen aus [1])	8
2.2	Reaktionskette zum Abbau von $\text{NO}_x$ in der Troposphäre. Rot markierte Übergänge finden nur tagsüber statt. . . . .	9
2.3	Entwicklung der $\text{NO}_x$ Emissionen und der Anteile der einzelnen Quellkategorien in Deutschland von 1995 bis 2022. (Abbildung entnommen aus [64]) . . . . .	11
2.4	: Reaktionskette des Ozon- $\text{NO}_x$ -Zyklus mit Abbau von $\text{NO}$ durch zusätzliche VOC-Radikale (Abbildung entnommen aus [73]) . . . . .	13
2.5	Isoplethen der maximalen Ozonkonzentration (ppb) als Funktion der $\text{NO}_x$ und VOC Emissionsrate ( $10^{12} \text{ molecules cm}^{-2} \text{ s}^{-1}$ ) aus theoretischen 0-d Berechnungen. Die blaue Linie trennt die VOC-sensitiven von den $\text{NO}_2$ -sensitiven Bereichen. (Abb. entnommen aus [75]) . . . . .	15
2.6	Gemessene Jahresmittelwerte des ländlichen und städtischen Hintergrunds von Ozon in Deutschland von 1995 bis 2023. (Abb. entnommen aus [4]) . . . . .	16
2.7	Veranschaulichung der durch Absorption gemessenen schrägen Säulendichte und der errechneten vertikalen Säulendichte während einer Satellitenmessung. (Abb. entnommen aus [80]) . . . . .	17
3.1	Lage der Messstationen im Untersuchungsgebiet. Einfärbung der einzelnen Punkte ist zufällig gewählt. . . . .	21
4.1	Training eines Entscheidungsbaums mit zwei Eingangsvariablen ( $x_1$ und $x_2$ ) und der Zielvariable $y$ . Der Entstehende Baum hat eine Tiefe von 2. . . . .	24
4.2	Aufbau eines Random Forest Modells aus $N$ Entscheidungsbäumen. Die Bäume sagen unabhängig voneinander einen Wert für die Dateninstanz voraus. Das Gesamtergebnis ergibt sich aus dem Mittelwert der Einzelvorhersagen. . . . .	24
4.3	Genereller Modellaufbau der $\text{PM}_{2.5}$ -, $\text{NO}_2$ -, und $\text{O}_3$ -Modelle. Die Modelle für $\text{NO}_2$ und $\text{PM}_{2.5}$ beziehen jeweils Satellitenmessungen (TCNO2 bzw. AOD) als Teil ihrer Eingangsvariablen. Die Vorhersagen des $\text{NO}_2$ -Modells werden weiterverwertet und dienen als Eingangsvariable für das Modell zur Ozonvorhersage. . . . .	25
4.4	Darstellung der Datenaufteilung für die räumliche und zeitliche Kreuzvalidierung. . . . .	28
4.5	Klassifizierungsgebiete für die objektive Wetterlagenklassifikation des DWD. Zur Bestimmung der Wetterlage für Deutschland wird nur der blaue Ausschnitt des Gitters benötigt, welcher sich sehr gut mit dem Untersuchungsgebiet dieser Arbeit deckt. . . . .	30

4.6	Beispielhafte räumliche Verteilungen einer Größe mit verschieden starker globaler Autokorrelation. Links: Abwechselnde Verteilung / Negative räumliche Autokorrelation ( $I < 0$ ), Mitte: Zufällige Verteilung / Keine räumliche Autokorrelation ( $I = 0$ ), Rechts: Zusammenhängende Verteilung / positive räumliche Autokorrelation ( $i > 0$ ). . . . .	32
5.1	Streudiagramm für die Vorhersagen der Modelle für $PM_{2.5}$ , $NO_2$ und $O_3$ im Vergleich zu den Stationsmessungen. Die Winkelhalbierende des Quadranten ist in Rot und die Ausgleichsgerade durch die gestreuten Daten in schwarz aufgetragen. . . . .	36
5.2	Feature Importance Werte aller verwendeten Merkmale für die $PM_{2.5}$ -, $NO_2$ -, und $O_3$ -Modelle. . . . .	37
5.3	Partielle Abhängigkeiten des Modells zur $PM_{2.5}$ -Vorhersage von den Variablen a) Jahrestag, b) Aerosol Optische Dicke, c) Grenzschichthöhe und d) Grenzschicht-Energiedissipation. . . . .	38
5.4	Partielle Abhängigkeiten des Modells zur $NO_2$ -Vorhersage von den Variablen a) Bevölkerungsdichte, b) $NO_2$ -Gehalt der troposphärischen Luftsäule c) Höhe der Grenzschicht, d) Dichte des Straßentyps 5 (lokale Straßen), e) Vegetationsindex und f) geographische Höhe. . . . .	39
5.5	Partielle Abhängigkeiten des Modells zur $O_3$ -Vorhersage von den Variablen a) Temperatur in 2 m Höhe, b) Stärke der Sonneneinstrahlung auf die Erdoberfläche, c) Jahrestag, d) vom $NO_2$ -Modell vorhergesagte $NO_2$ -Bodenkonzentration, e) Grenzschichthöhe und f) relative Luftfeuchtigkeit. . . . .	40
5.6	Vergleich der Feature Importance eines Modells mit verschiedenen Eingangsvariablen mit einem Modell mit mehreren identischen Eingangsvariablen (AOD_1-4). Die Information der einzelnen Merkmale ist redundant, wodurch sich die MDA-Wertung auf diese aufteilt. . . . .	41
5.7	Pearson Korrelation zwischen einigen Eingangsvariablen. Es werden nur Variablen angezeigt, für welche bedeutende Korrelationen bestehen. . . . .	41
5.8	Darstellung der über den gesamten Untersuchungszeitraum ( $PM_{2.5}$ : 2018 bis 2022, $NO_2$ und $O_3$ : 2018 bis 2020) gemittelten $PM_{2.5}$ -, $NO_2$ und $O_3$ -Modellvorhersagen für die Schadstoffbodenkonzentration. Es ist jeweils der Mittelwert über das Gesamtgebiet ( $\mu$ ) und die mittlere Abdeckung der enthaltenen Tage ( $\overline{Cov}$ ) angegeben . . . . .	42
5.9	LISA in den Gesamtmittelwerten der $PM_{2.5}$ -, $NO_2$ und $O_3$ -Modellvorhersagen. Rote Flächen zeigen Hotspots (HH) auf und blaue Flächen signalisieren Coldspots (LL), wobei nur Hot- bzw. Coldspots mit einer zusammenhängenden Fläche von mehr als 250 Gitterpunkten ( $\sim 70km^2$ ) gezeigt werden. $I_{global}$ gibt die Stärke der globalen Autokorrelation an. . . . .	43
5.10	Histogramme der Verteilung der Flächendaten auf dem Gesamtgebiet für a) $PM_{2.5}$ , b) $NO_2$ und c) $O_3$ . . . . .	44
5.11	Mit einem Binomialfilter ( $1/4 [1, 2, 1]$ ) geglättete Wochenmittelwerte für $PM_{2.5}$ und die verbrannte Fläche in den Ländern Deutschland, Tschechien, Schweiz, Österreich, Belgien und Niederlande. Die grauen Linien geben für $PM_{2.5}$ den Mittelwert zwischen 2018 und 2022 und für die verbrannte Fläche den Mittelwert zwischen 2006 und 2024 an. . . . .	45

5.12	Der mittlere Niederschlag und Monatsmittelwerte für PM <sub>2.5</sub> auf dem Untersuchungsgebiet für die Jahre 2018 bis 2022. Die Daten sind aufgetragen als Abweichung vom zwischen 2018 und 2022 gemessenen Mittelwert. Die Niederschlagswerte wurden zur besseren Interpretierbarkeit invertiert aufgetragen. . . . .	45
5.13	Mit einem Binomialfilter (1/4[1, 2, 1]) geglättete Wochenmittelwerte für NO <sub>2</sub> und die verbrannte Fläche in den Ländern Deutschland, Tschechien, Schweiz, Österreich, Belgien und Niederlande. Die grauen Linien geben für NO <sub>2</sub> den Mittelwert zwischen 2018 und 2020 und für die verbrannte Fläche den Mittelwert zwischen 2006 und 2024 an. . . . .	46
5.14	Der mittlere Niederschlag und Monatsmittelwerte für NO <sub>2</sub> auf dem Untersuchungsgebiet für die Jahre 2018 bis 2020. Die NO <sub>2</sub> -Daten sind aufgetragen als Abweichung vom zwischen 2018 und 2020 gemessenen Mittelwert. Die Niederschlagswerte beziehen sich auf den Mittelwert zwischen 2018 und 2022 und wurden zur besseren Interpretierbarkeit invertiert aufgetragen. . . . .	46
5.15	Schadstoffwerte der PM <sub>2.5</sub> -, NO <sub>2</sub> - und O <sub>3</sub> -Stationsmessungen: a) Jahresgang der Messdaten als Monatsmittelwerte. Die eingefärbten Flächen beschreiben die Standardabweichung innerhalb des Monats. b) Entwicklung der Jahresmittelwerte der Stationsdaten von 2018 bis 2022. . . . .	47
5.16	Jahreszeitliche Mittelwerte der Modellvorhersagen des bodennahen a) PM <sub>2.5</sub> , b) NO <sub>2</sub> und c) O <sub>3</sub> . . . . .	49
5.17	Aufteilung der Daten nach Stärke der mittleren Windgeschwindigkeit pro Tag. Der linke und rechte Extrembereich jenseits der schwarzen vertikalen Linien enthält jeweils ca. 15% der Daten. . . . .	50
5.18	Verteilung der Hauptwindrichtung aus den Großwetterlagen auf die Jahreszeiten. . . . .	50
5.19	Ausschnitte der PM <sub>2.5</sub> -Hotspots über dem Ruhrgebiet bei nordwestlicher, nordöstlicher, südwestlicher und südöstlicher Anstromrichtung. Die Pfeile kennzeichnen die jeweilige Hauptwindrichtung nach der Wetterlagenklassifikation des DWD. . . . .	51
5.20	Ausschnitte der NO <sub>2</sub> -Hotspots über dem Ruhrgebiet bei nordwestlicher, nordöstlicher, südwestlicher und südöstlicher Anstromrichtung. Die Pfeile kennzeichnen die jeweilige Hauptwindrichtung nach der Wetterlagenklassifikation des DWD. . . . .	52
5.21	Relative Häufigkeit der Tage deren Großwetterlage als feucht oder trocken eingestuft wurde, in Abhängigkeit der Jahreszeit. . . . .	53
5.22	Mittlerer Änderung der NO <sub>2</sub> , O <sub>3</sub> und PM <sub>2.5</sub> -Konzentration am Wochenende gegenüber dem jeweiligen jahreszeitlichen Mittel des (a) Sommers und (b) Winters. Die Stärke des Effekts ist für Gebiete mit verschiedenen Bevölkerungsdichten dargestellt. . . . .	54
5.23	Partielle Abhängigkeiten des Vorhergesagten Ozons vom bodennahen NO <sub>2</sub> (NO <sub>2</sub> _pred). Es wurden separate Modelle für die Daten aus (a) Winter, (b) Frühling, (c) Sommer und (d) Herbst trainiert. . . . .	55
6.1	Darstellung einiger räumlicher Variablen: (a) NDVI als Mittelwert über den gesamten Untersuchungszeitraum, (b) Bevölkerungsdichte, (c) geographische Höhe und (d) Dichte lokaler Straßen. . . . .	62
A.1	Flächenhafte PM <sub>2.5</sub> Mittelwerte der einzelnen Messjahre. . . . .	98
A.2	Flächenhafte NO <sub>2</sub> Mittelwerte der einzelnen Messjahre. . . . .	98

A.3	Flächenhafte O <sub>3</sub> Mittelwerte der einzelnen Messjahre. . . . .	99
A.4	Flächenhafte PM <sub>2.5</sub> -Mittelwerte für verschiedene mittlere Windstärken. . .	99
A.5	Flächenhafte PM <sub>2.5</sub> -Mittelwerte für verschiedene Hauptwindrichtungen. . .	100
A.6	Flächenhafte NO <sub>2</sub> -Mittelwerte für verschiedene mittlere Windstärken. . . .	100
A.7	Flächenhafte NO <sub>2</sub> -Mittelwerte für verschiedene Hauptwindrichtungen. . . .	101
A.8	Flächenhafte O <sub>3</sub> -Mittelwerte für verschiedene mittlere Windstärken. . . .	101
A.9	Flächenhafte O <sub>3</sub> -Mittelwerte für verschiedene Hauptwindrichtungen. . . .	102
B.1	LISA für die PM <sub>2.5</sub> -Mittelwerte der einzelnen Jahreszeiten. . . . .	103
B.2	LISA für die NO <sub>2</sub> -Mittelwerte der einzelnen Jahreszeiten. . . . .	104
B.3	LISA für die O <sub>3</sub> -Mittelwerte der einzelnen Jahreszeiten. . . . .	105
B.4	LISA für die PM <sub>2.5</sub> -Mittelwerte bei Verschiedenen mittleren Windstärken. .	105
B.5	LISA für PM <sub>2.5</sub> -Mittelwerte bei verschiedenen Hauptwindrichtungen. . . .	106
B.6	LISA für die NO <sub>2</sub> -Mittelwerte bei Verschiedenen mittleren Windstärken. . .	106
B.7	LISA für NO <sub>2</sub> -Mittelwerte bei verschiedenen Hauptwindrichtungen. . . . .	107
B.8	LISA für die O <sub>3</sub> -Mittelwerte bei Verschiedenen mittleren Windstärken. . . .	107
B.9	LISA für O <sub>3</sub> -Mittelwerte bei verschiedenen Hauptwindrichtungen. . . . .	108
C.1	Saharastaubmessungen am Hohenpeißenberg. [149] . . . . .	109

# Tabellenverzeichnis

3.1	Zielvariablen und mögliche Eingangsvariablen für die maschinellen Lernmodelle. Satellitendaten sind mit <sup>sat.</sup> , Stationsmessungen mit <sup>in-situ</sup> und Daten aus Wettermodelle mit <sup>m</sup> markiert. . . . .	20
4.1	Optimale Eingangsparameter ermittelt durch Sequential Feature Selection, zeilenweise angeordnet als Satellitendaten, meteorologische Parameter, ortsfeste Merkmale und zeitliche Variablen. . . . .	26
4.2	Auflistung der optimalen Modellparameter für die Random Forest Modelle. . . . .	26
4.3	Abhängig von seiner Nachbarschaft, kann ein Punkt einer der vier Gruppen (HL, HH, LL, LH) zugeordnet werden. Wichtig ist dabei die Unterscheidung zwischen HH als Orte bei denen hohe Werte von hohen Werten umgeben sind, und LL bei denen wiederum niedrige Werte von niedrigen Werten umgeben sind. Bei HL bzw. LH liegt der Punktwert ungewöhnlich stark über bzw. unter denen der Nachbarn. . . . .	33
5.1	Bestimmtheitsmaße, Mittelwerte, quadratischer Fehler, Bias, Standardabweichung und verwendete Datenmenge der PM <sub>2.5</sub> -, NO <sub>2</sub> -, und O <sub>3</sub> -Modelle nach abgeschlossenem Training. . . . .	35
5.2	Bestimmtheitsmaße der Modelle unter verschiedenen Testverfahren zur zeitlichen und räumlichen Kreuzvalidierung. Die Gruppierung gibt an, dass jeweils die Daten einzelner Tage, Stationen, Regionen oder 6 Monaten stets gemeinsam in den Test- bzw. Trainingsdaten vorkommen. Das Modell O <sub>3</sub> (NO <sub>2</sub> <sup>*</sup> ) wurde mit NO <sub>2</sub> -Stationsdaten trainiert, während O <sub>3</sub> (NO <sub>2</sub> ) mit den durch das NO <sub>2</sub> -Modell vorhergesagten Eingangsdaten trainiert und getestet wurde. . . . .	36
5.3	Jahresmittelwerte der vorhergesagten Schadstoffkonzentrationen über das gesamte Untersuchungsgebiet. . . . .	42
5.4	Räumliche Mittel- und Extremwerte in den flächenhaften Mittelwerten der Schadstoffwerte für die einzelnen Jahreszeiten. . . . .	48
5.5	Mittlere Schadstoffkonzentrationen über die Gesamtfläche des Beobachtungsgebiets, jeweils für feuchte (F) und trockene (T) Wetterbedingungen . . . . .	53
5.6	Mittlere Schadstoffkonzentrationen über die Gesamtfläche des Beobachtungsgebiets, jeweils von Montag bis Freitag und von Samstag und Sonntag für verschiedene Jahreszeiten. . . . .	54



# Anhang

## A Abbildungen flächenhafter Schadstoffmittelwerte

Abbildung A.1: Flächenhafte PM<sub>2.5</sub> Mittelwerte der einzelnen Messjahre.

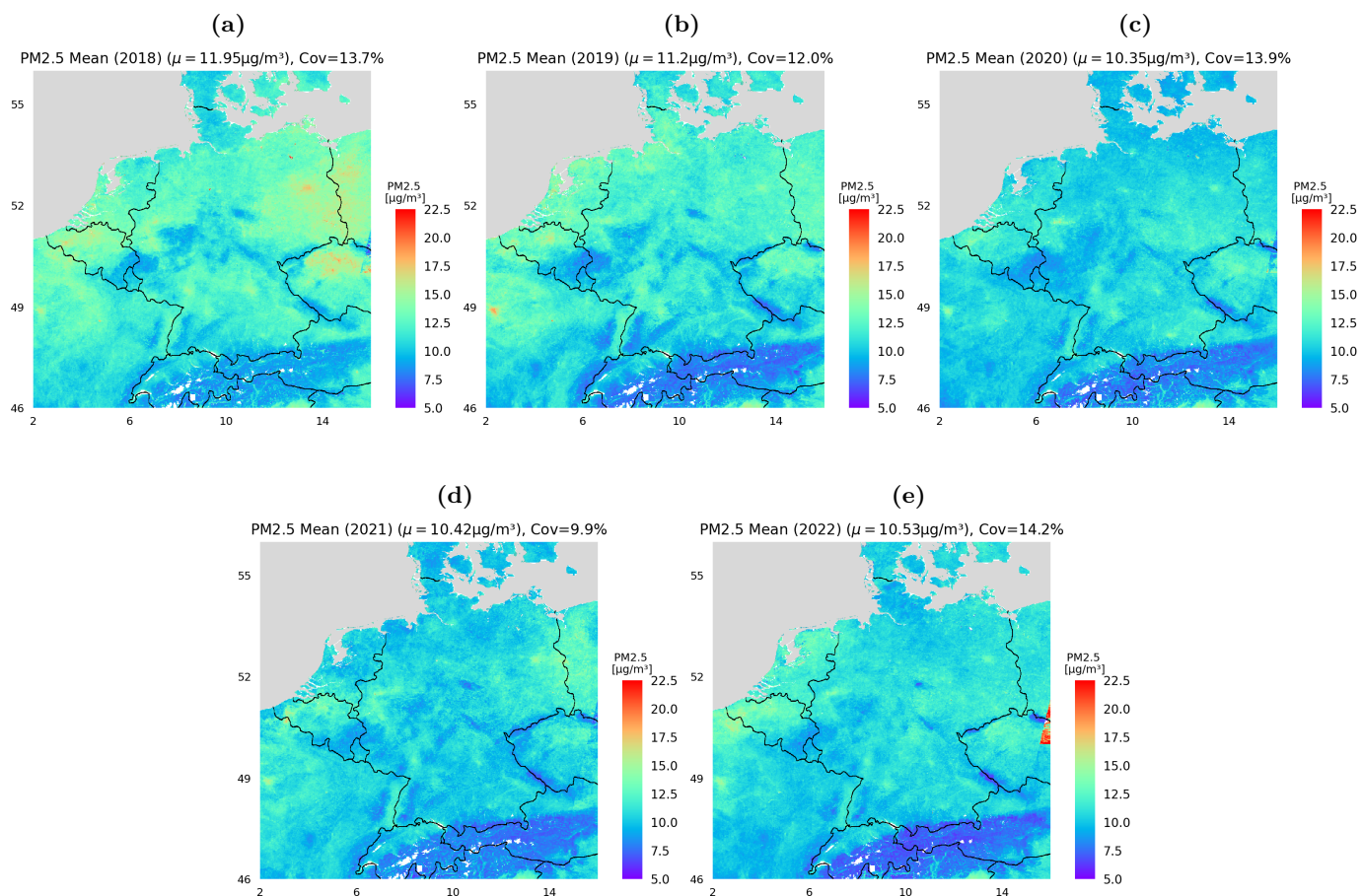


Abbildung A.2: Flächenhafte NO<sub>2</sub> Mittelwerte der einzelnen Messjahre.

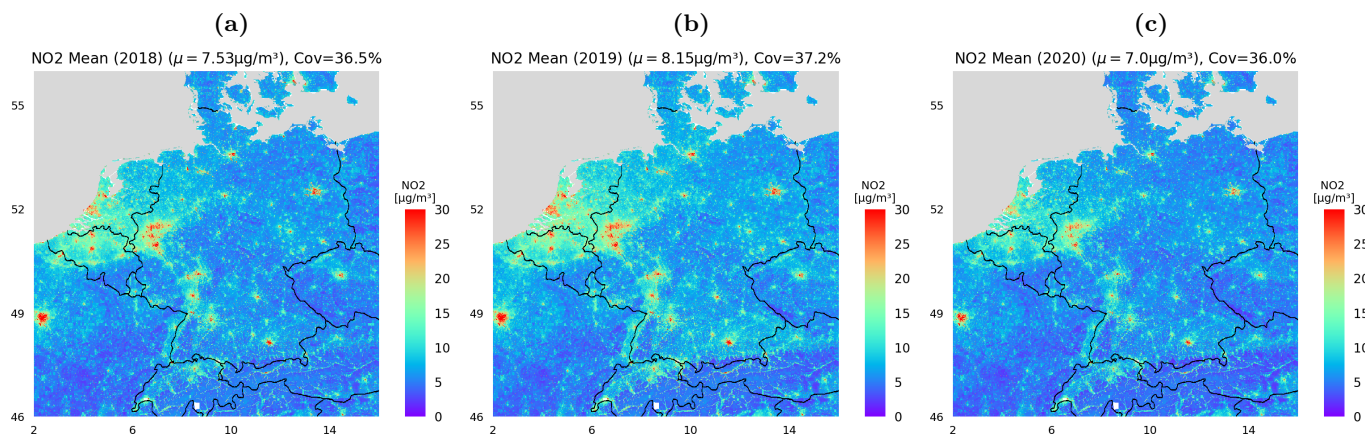




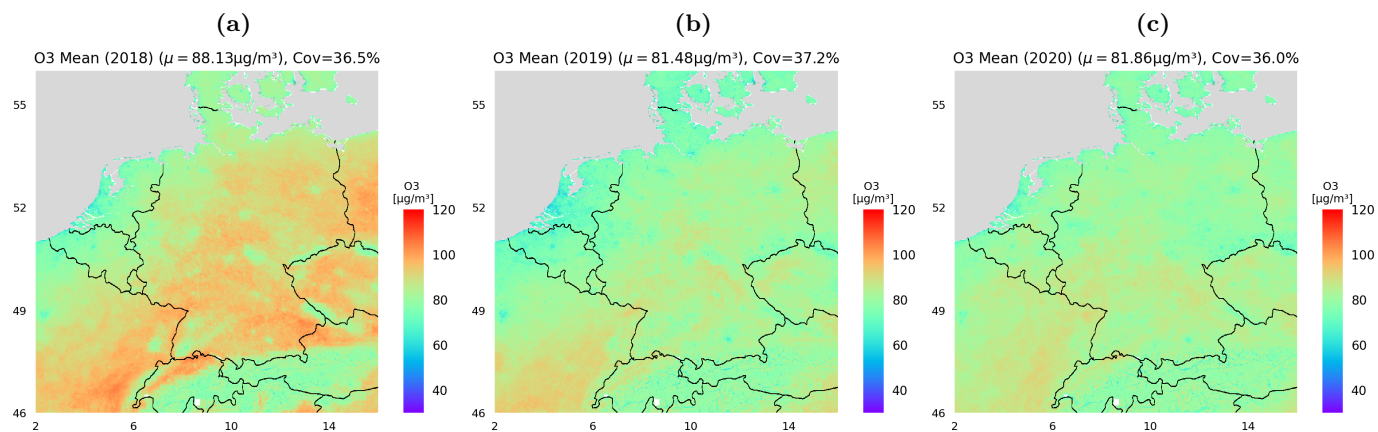
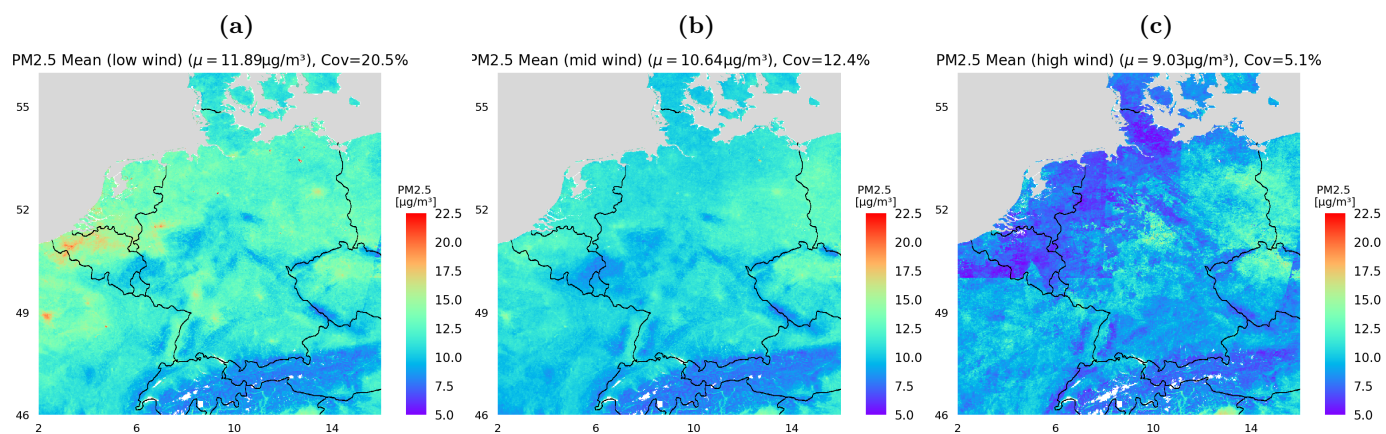
Abbildung A.3: Flächenhafte O<sub>3</sub> Mittelwerte der einzelnen Messjahre.Abbildung A.4: Flächenhafte PM<sub>2.5</sub>-Mittelwerte für verschiedene mittlere Windstärken.

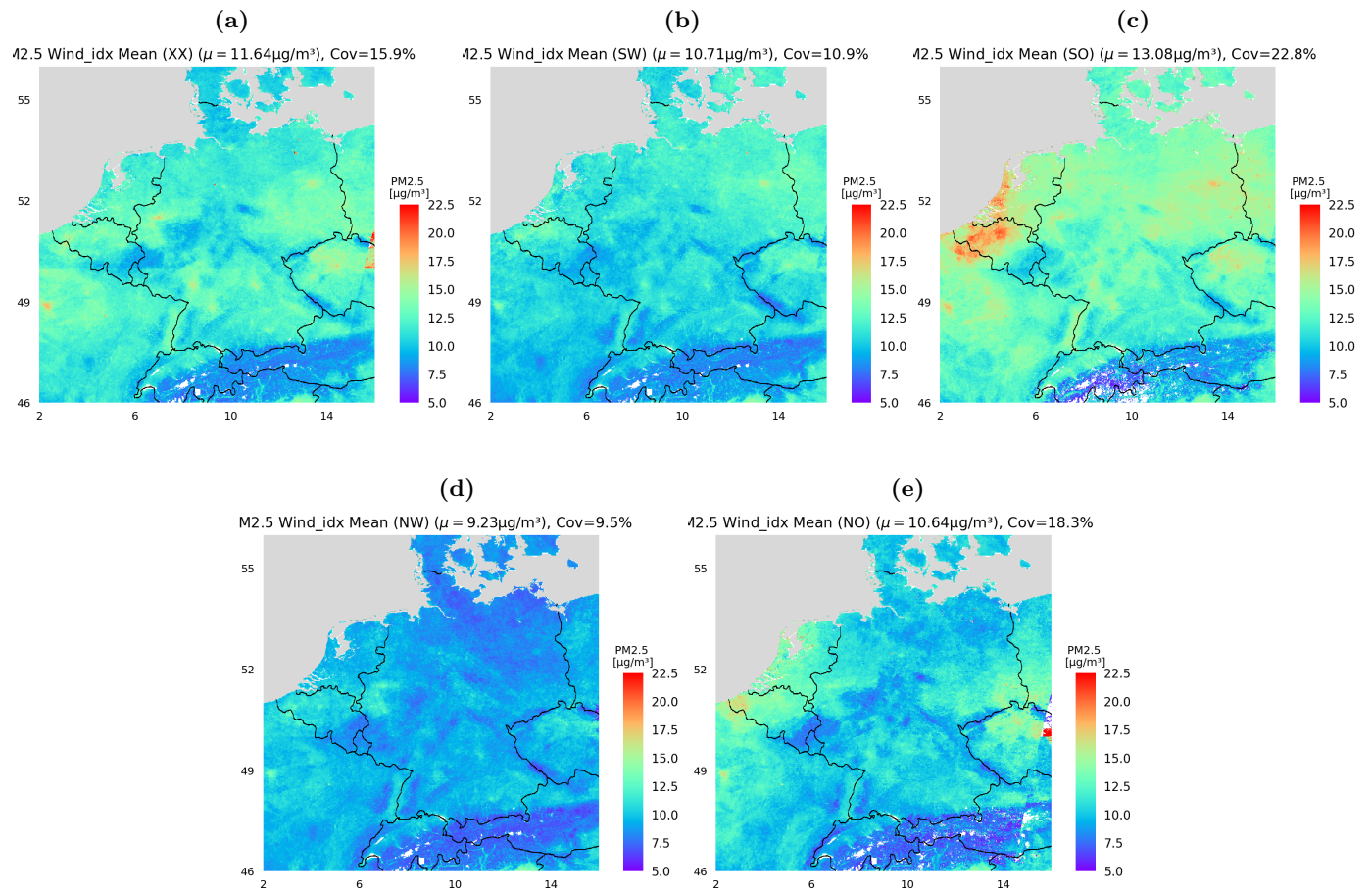
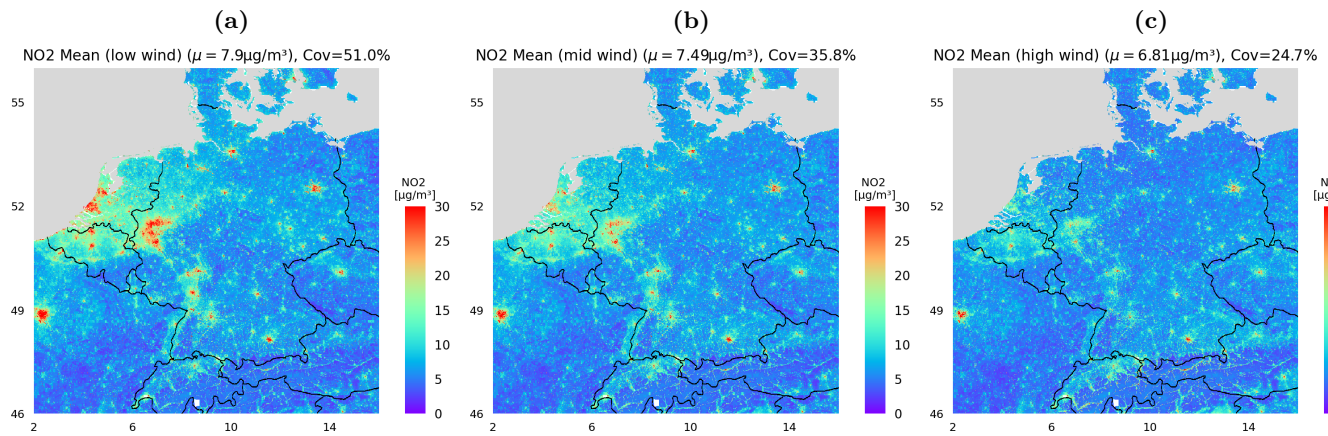
Abbildung A.5: Flächenhafte PM<sub>2.5</sub>-Mittelwerte für verschiedene Hauptwindrichtungen.Abbildung A.6: Flächenhafte NO<sub>2</sub>-Mittelwerte für verschiedene mittlere Windstärken.



Abbildung A.7: Flächenhafte NO<sub>2</sub>-Mittelwerte für verschiedene Hauptwindrichtungen.

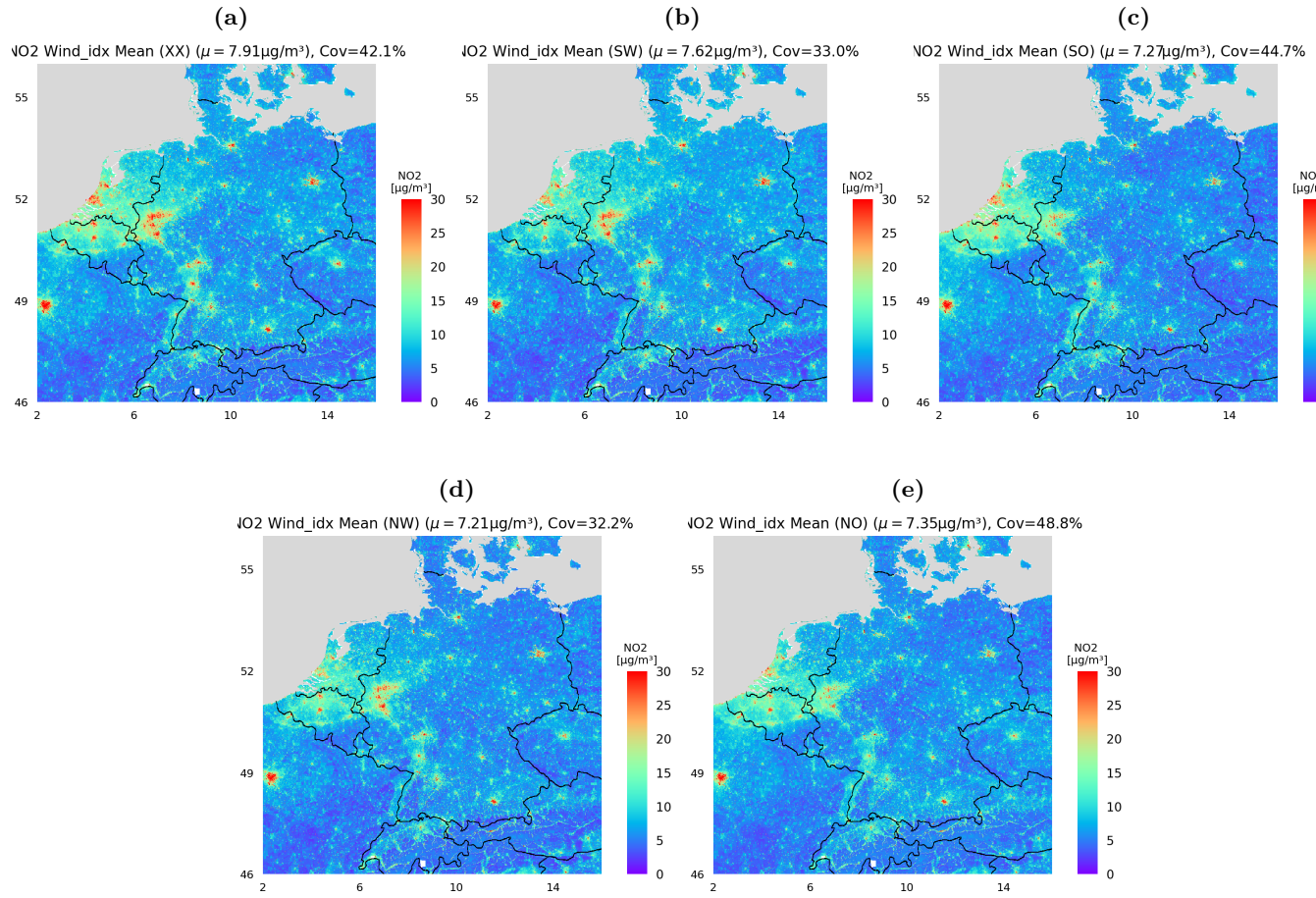
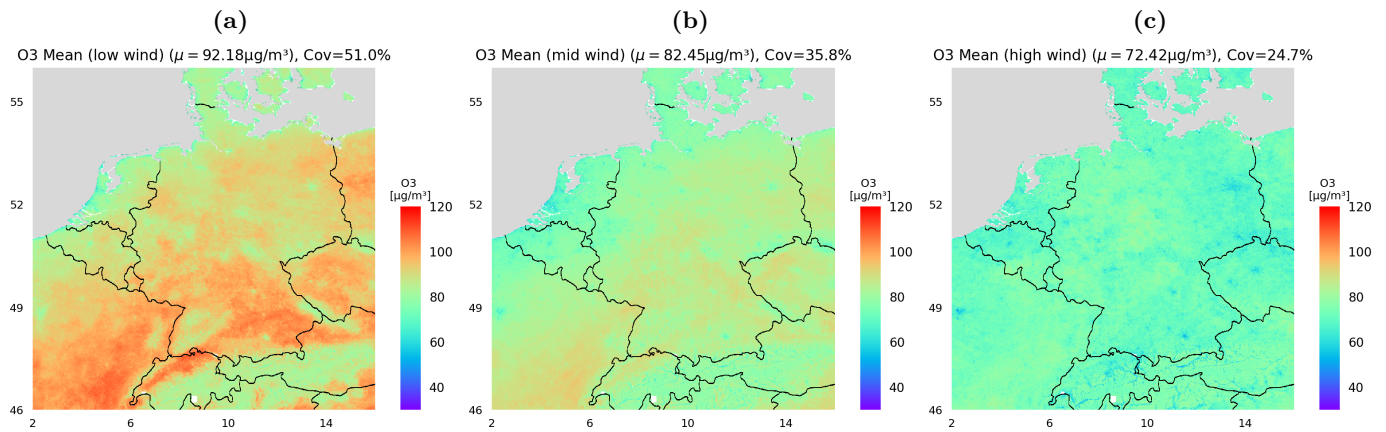
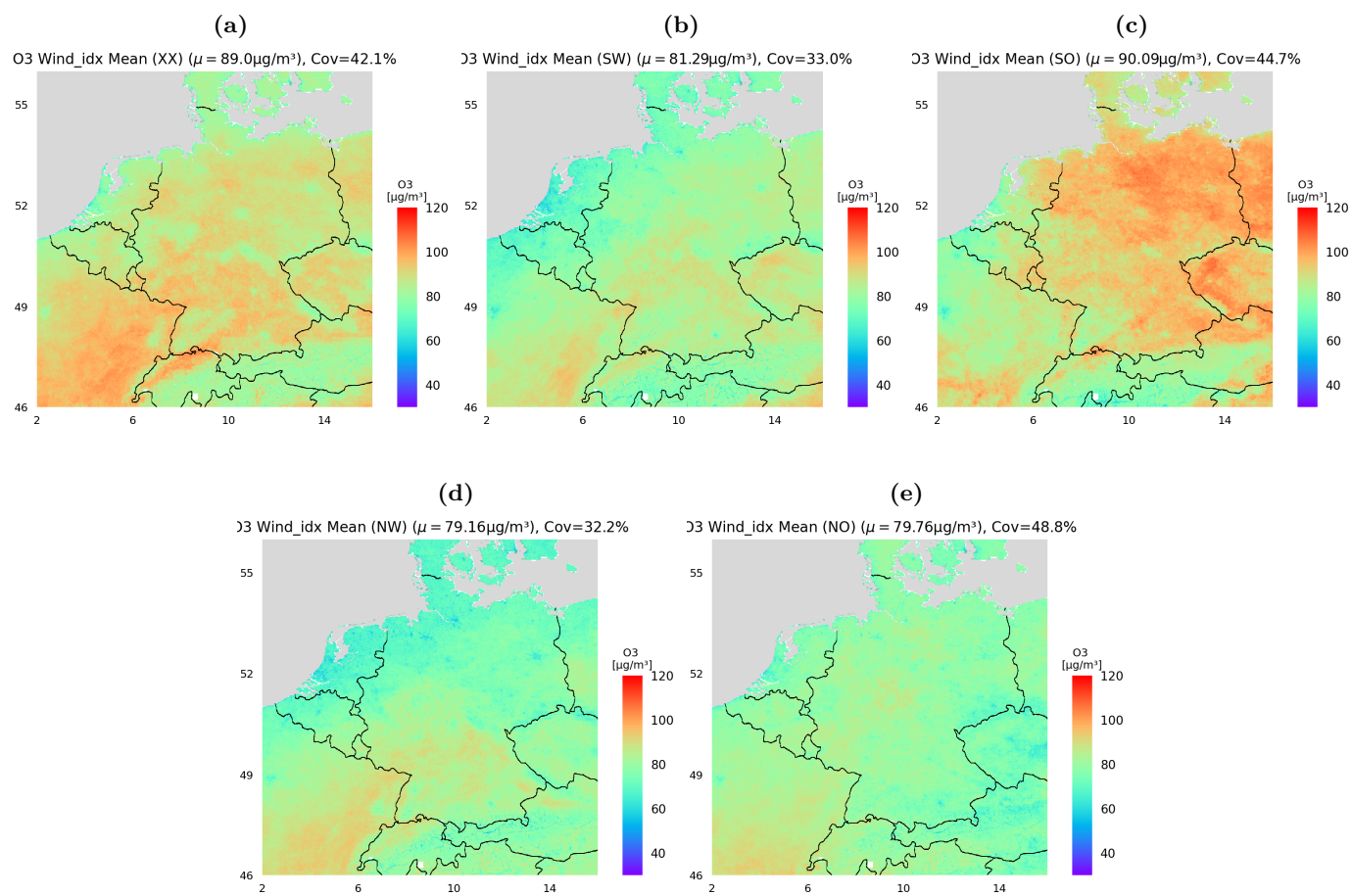
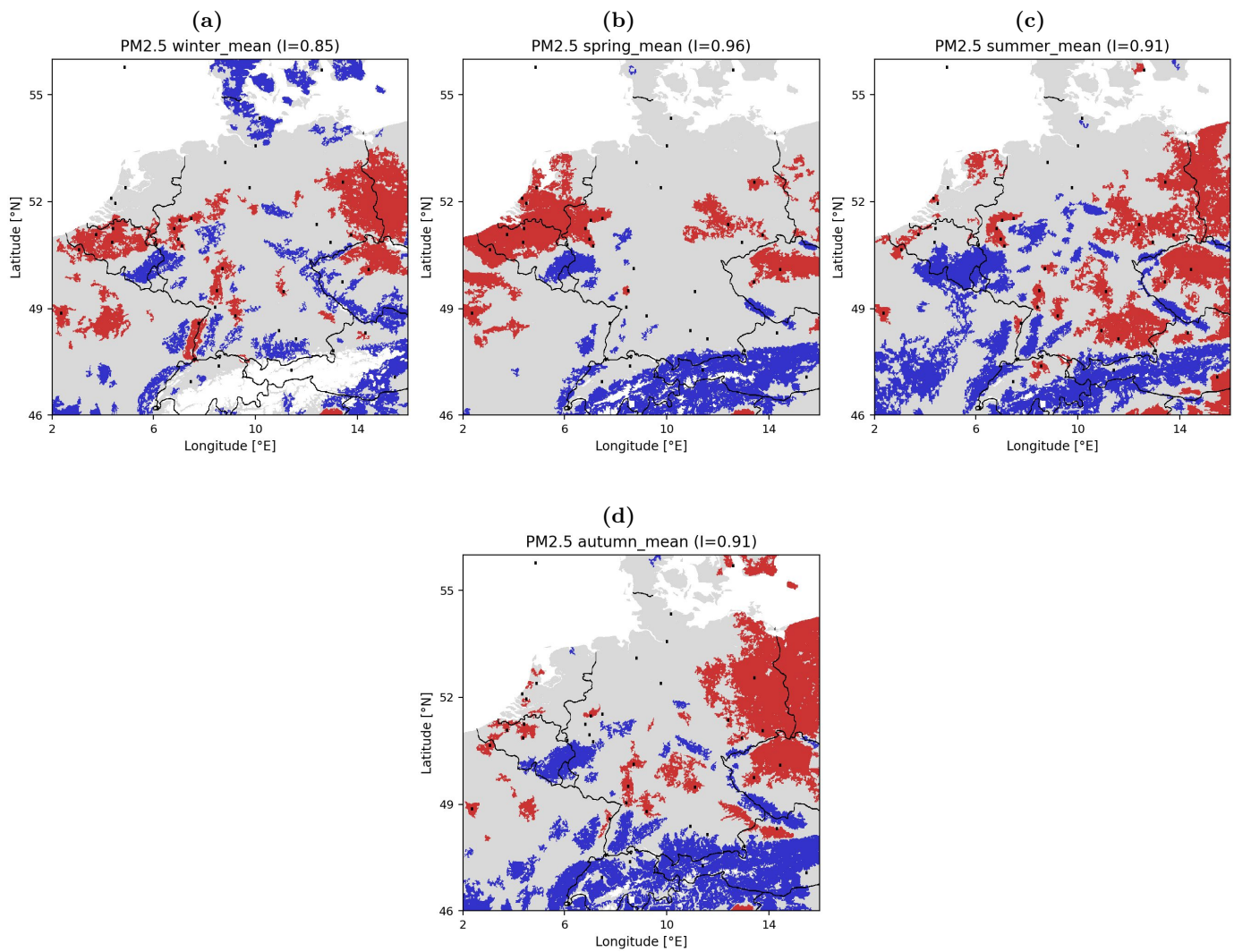


Abbildung A.8: Flächenhafte O<sub>3</sub>-Mittelwerte für verschiedene mittlere Windstärken.



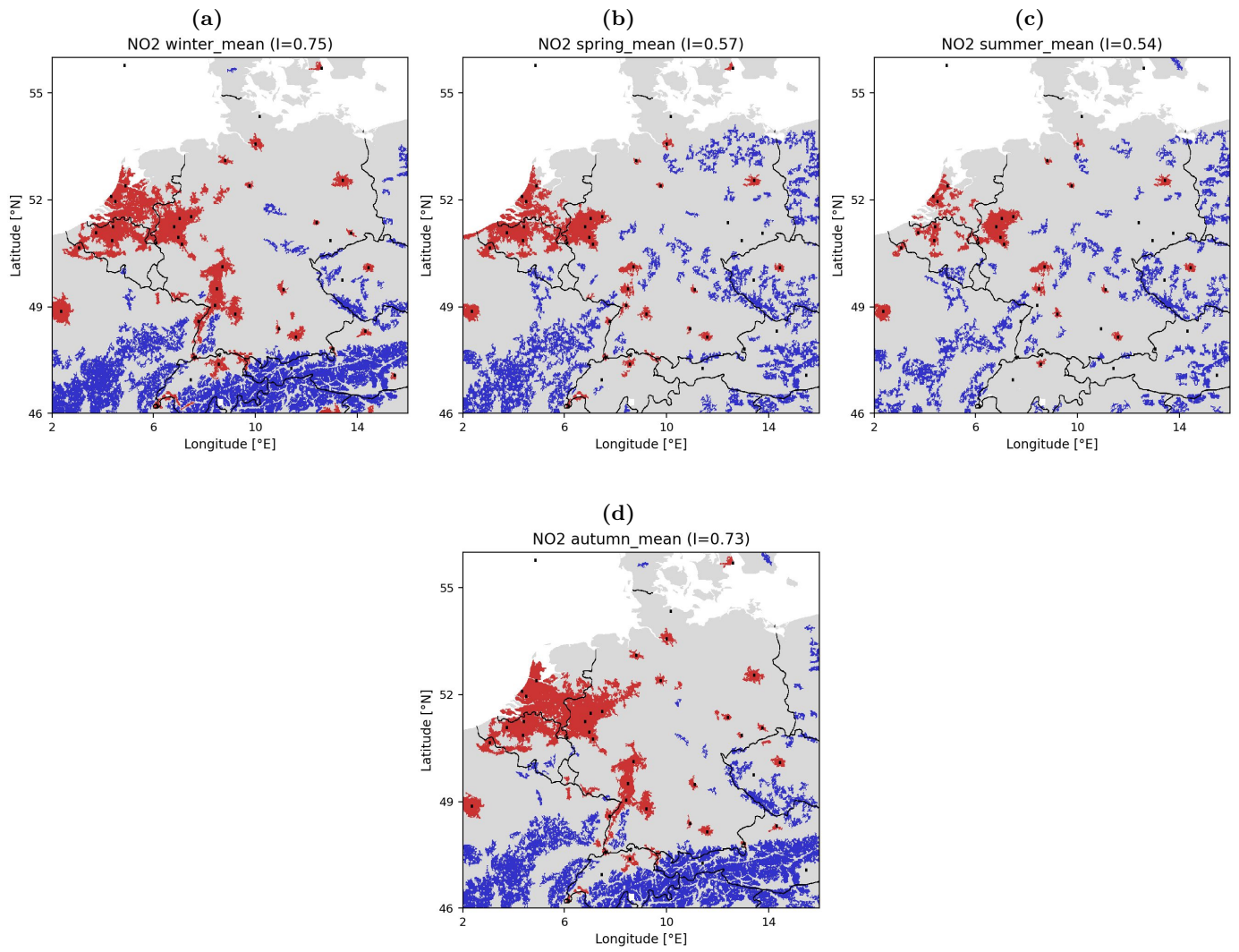
**Abbildung A.9:** Flächenhafte O<sub>3</sub>-Mittelwerte für verschiedene Hauptwindrichtungen.

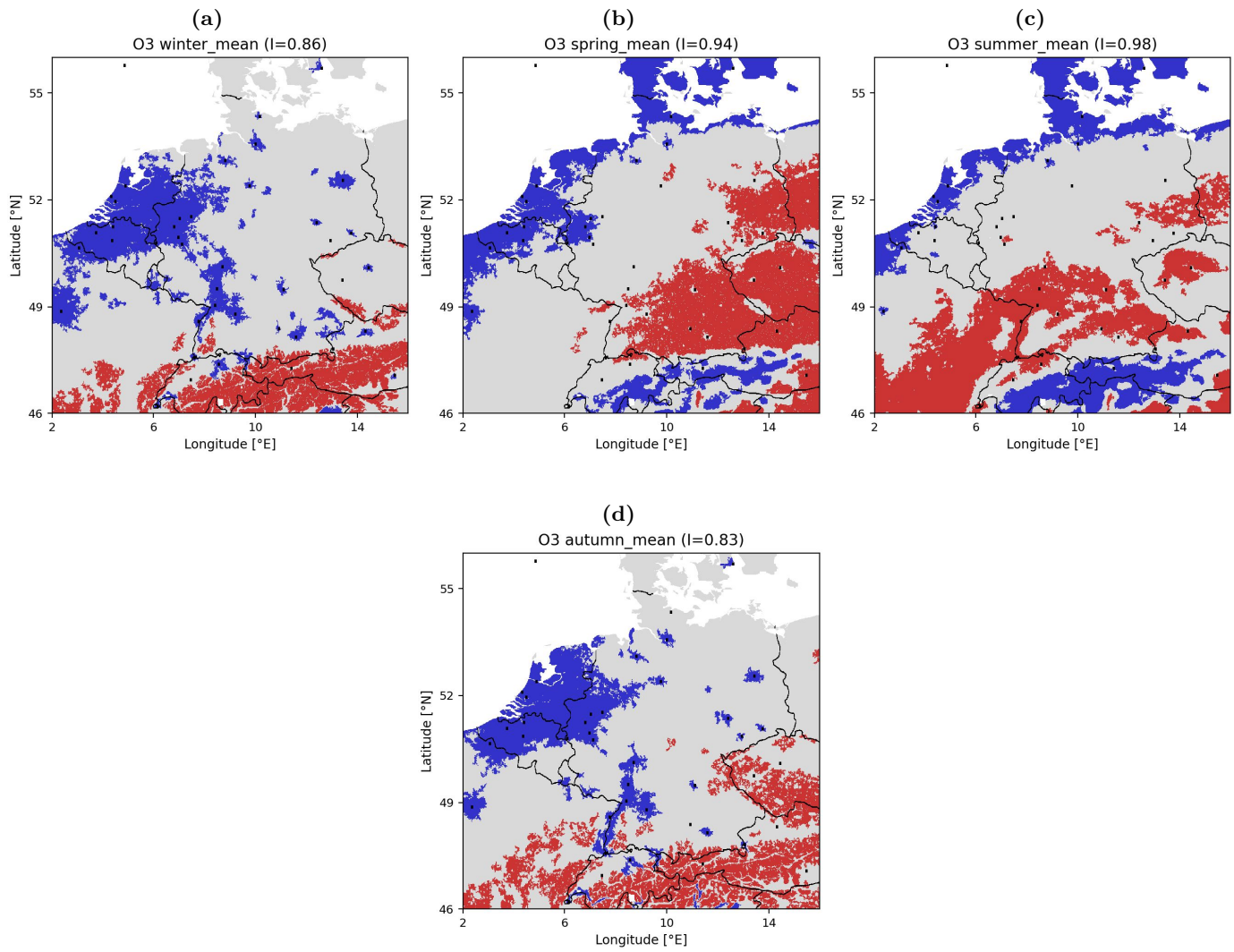
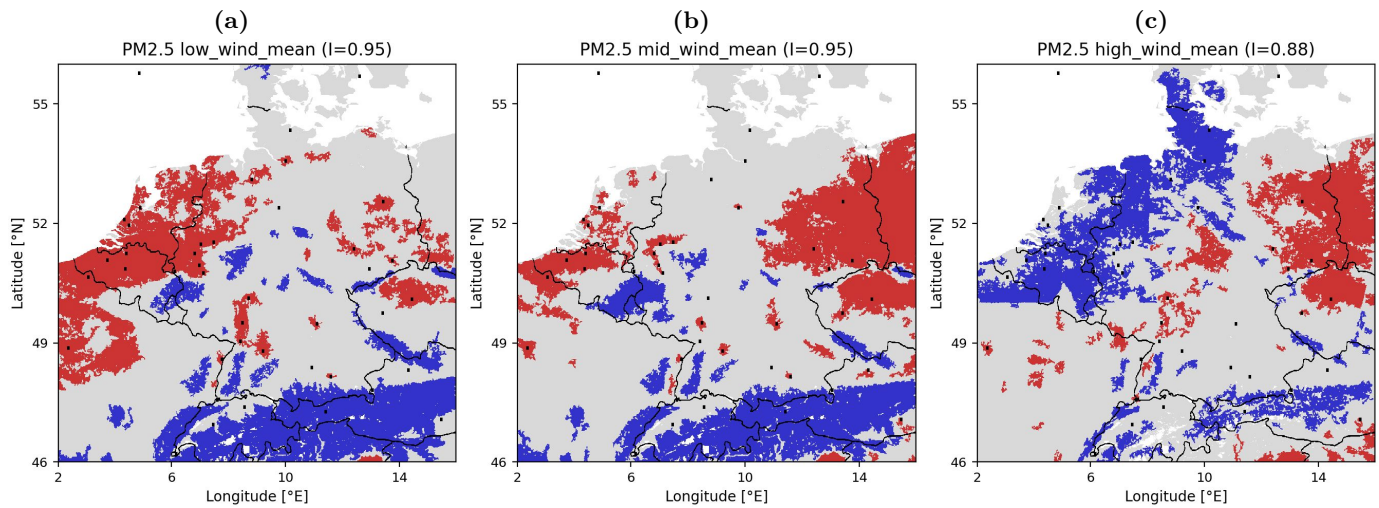
## B Abbildungen zur Hot- und Coldspotanalyse

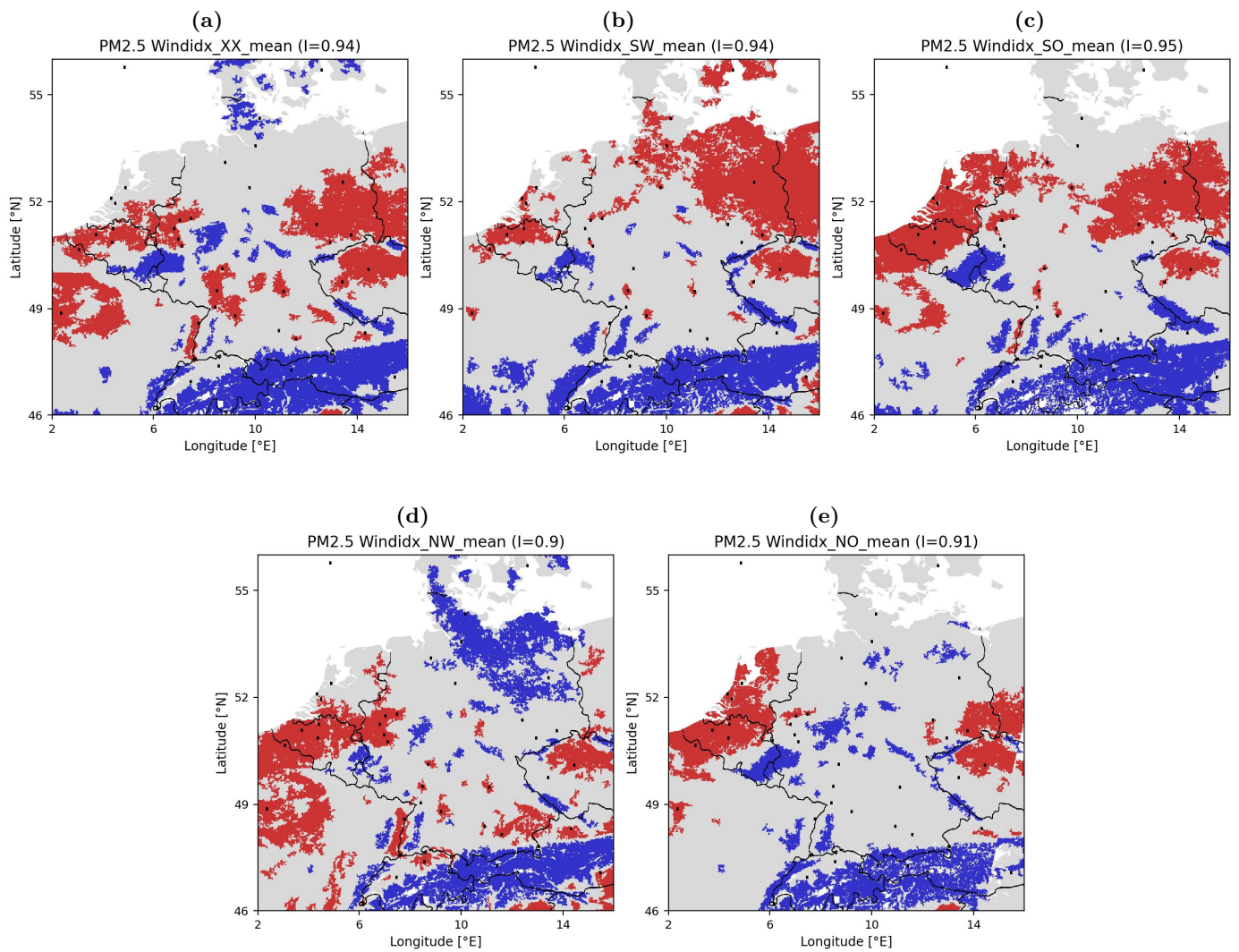
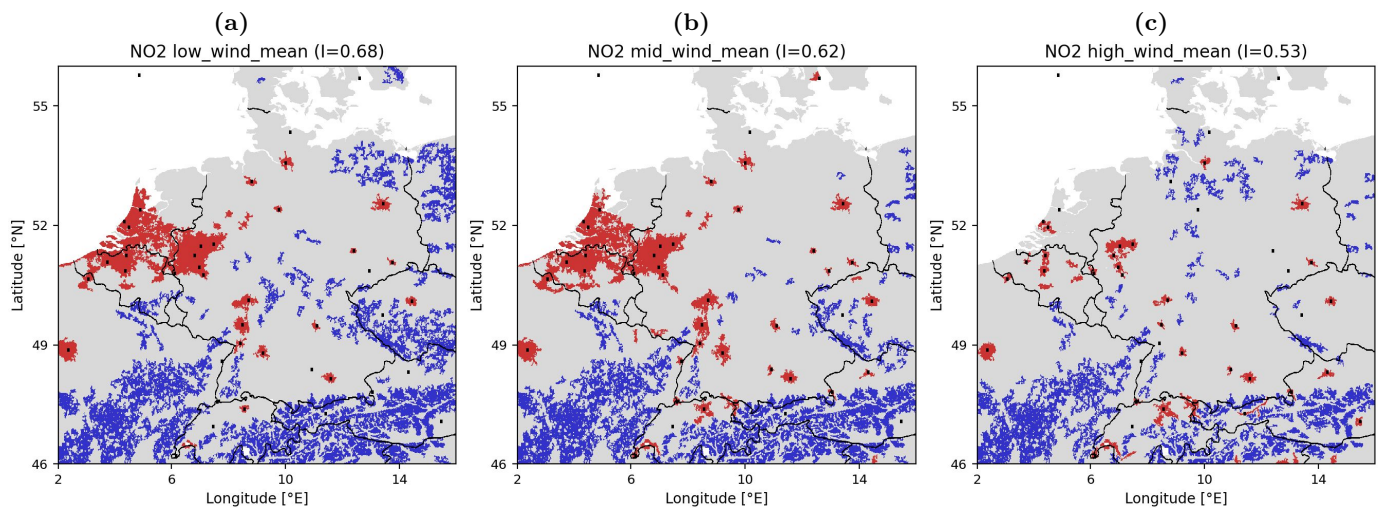
Abbildung B.1: LISA für die PM<sub>2.5</sub>-Mittelwerte der einzelnen Jahreszeiten.



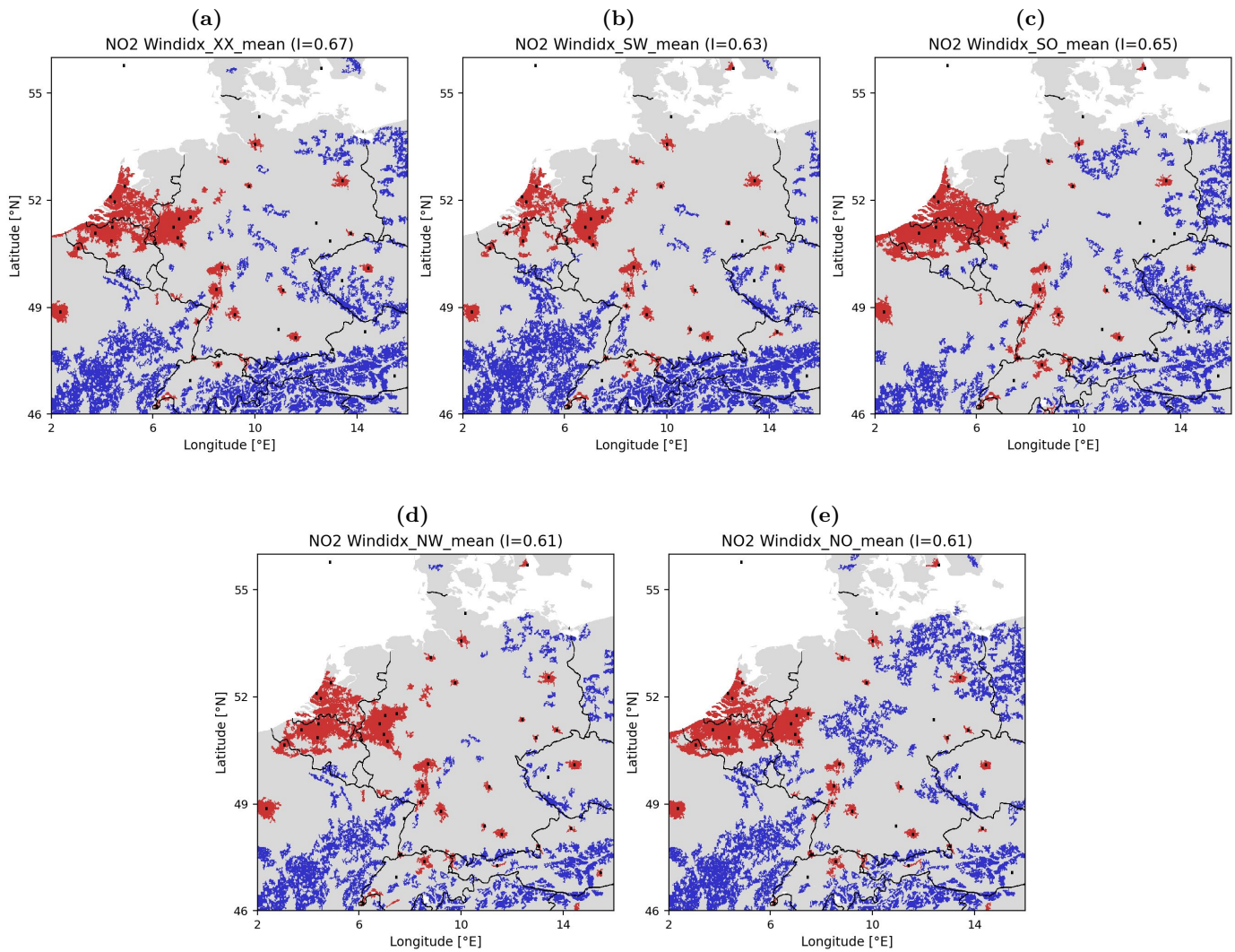
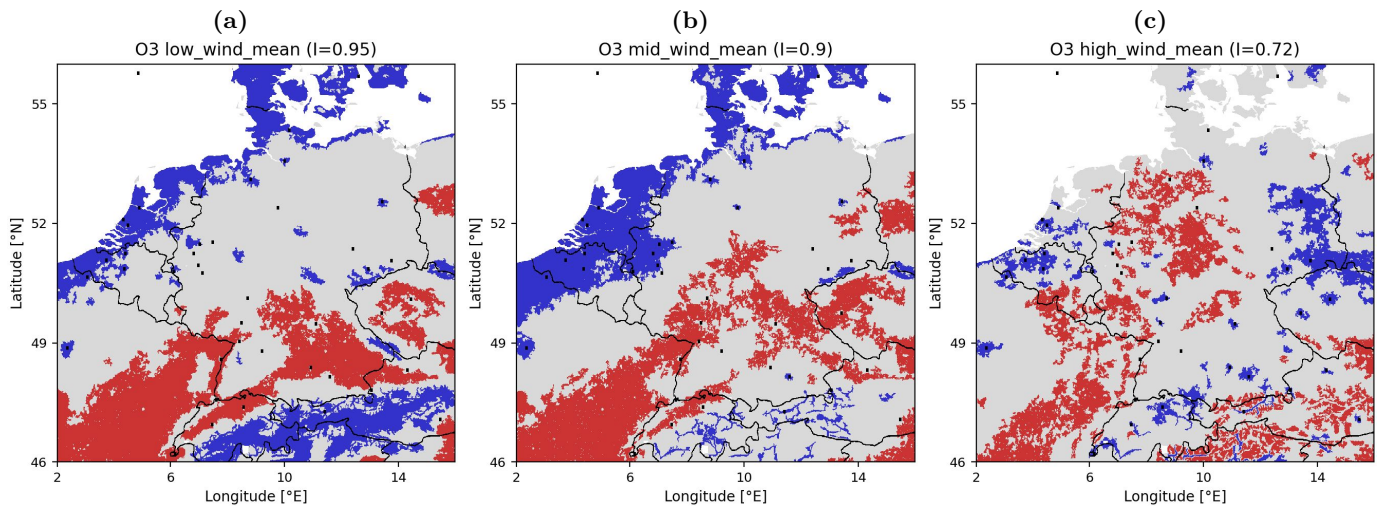
**Abbildung B.2:** LISA für die NO<sub>2</sub>-Mittelwerte der einzelnen Jahreszeiten.

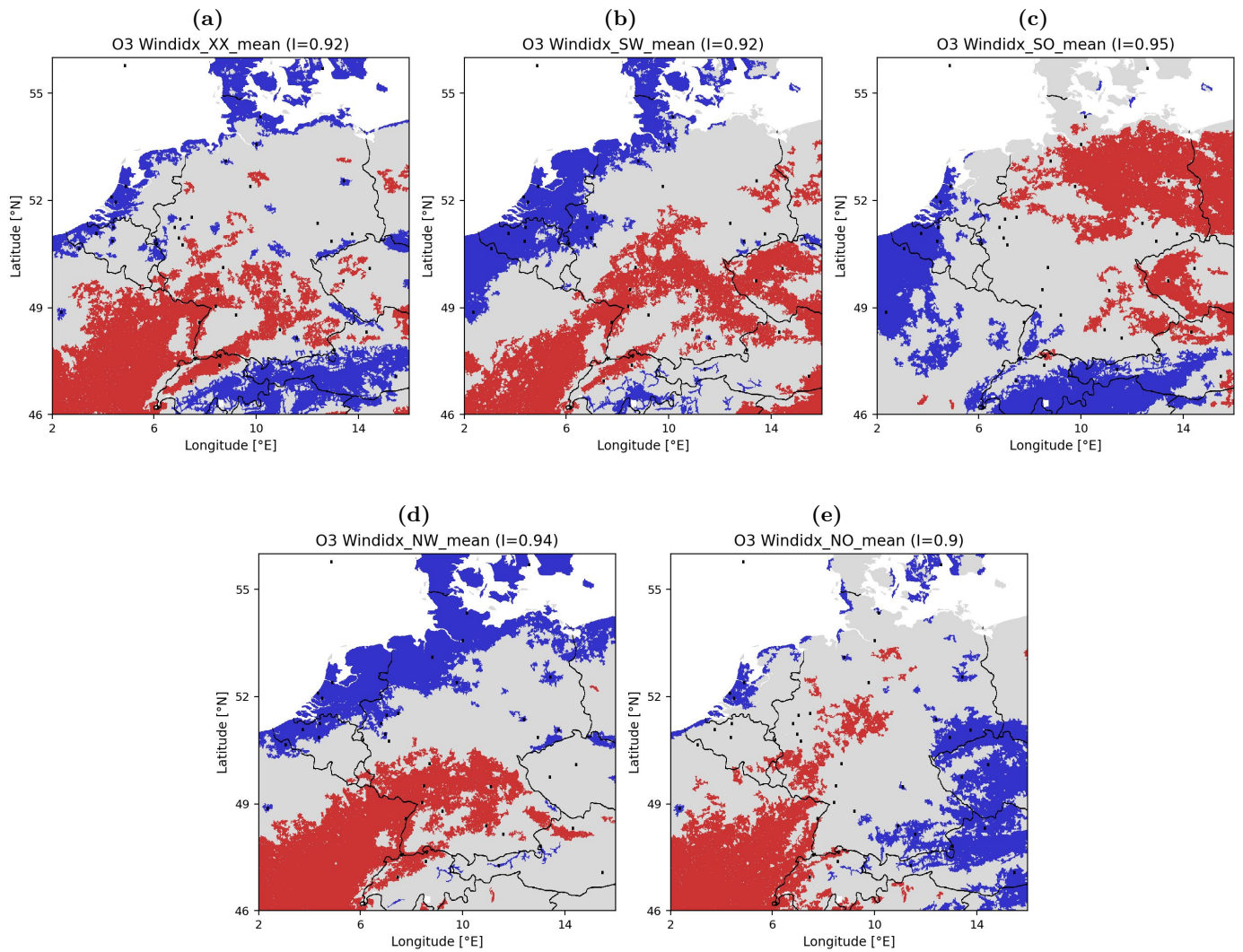


**Abbildung B.3:** LISA für die O<sub>3</sub>-Mittelwerte der einzelnen Jahreszeiten.**Abbildung B.4:** LISA für die PM<sub>2.5</sub>-Mittelwerte bei Verschiedenen mittleren Windstärken.

**Abbildung B.5:** LISA für PM<sub>2.5</sub>-Mittelwerte bei verschiedenen Hauptwindrichtungen.**Abbildung B.6:** LISA für die NO<sub>2</sub>-Mittelwerte bei Verschiedenen mittleren Windstärken.

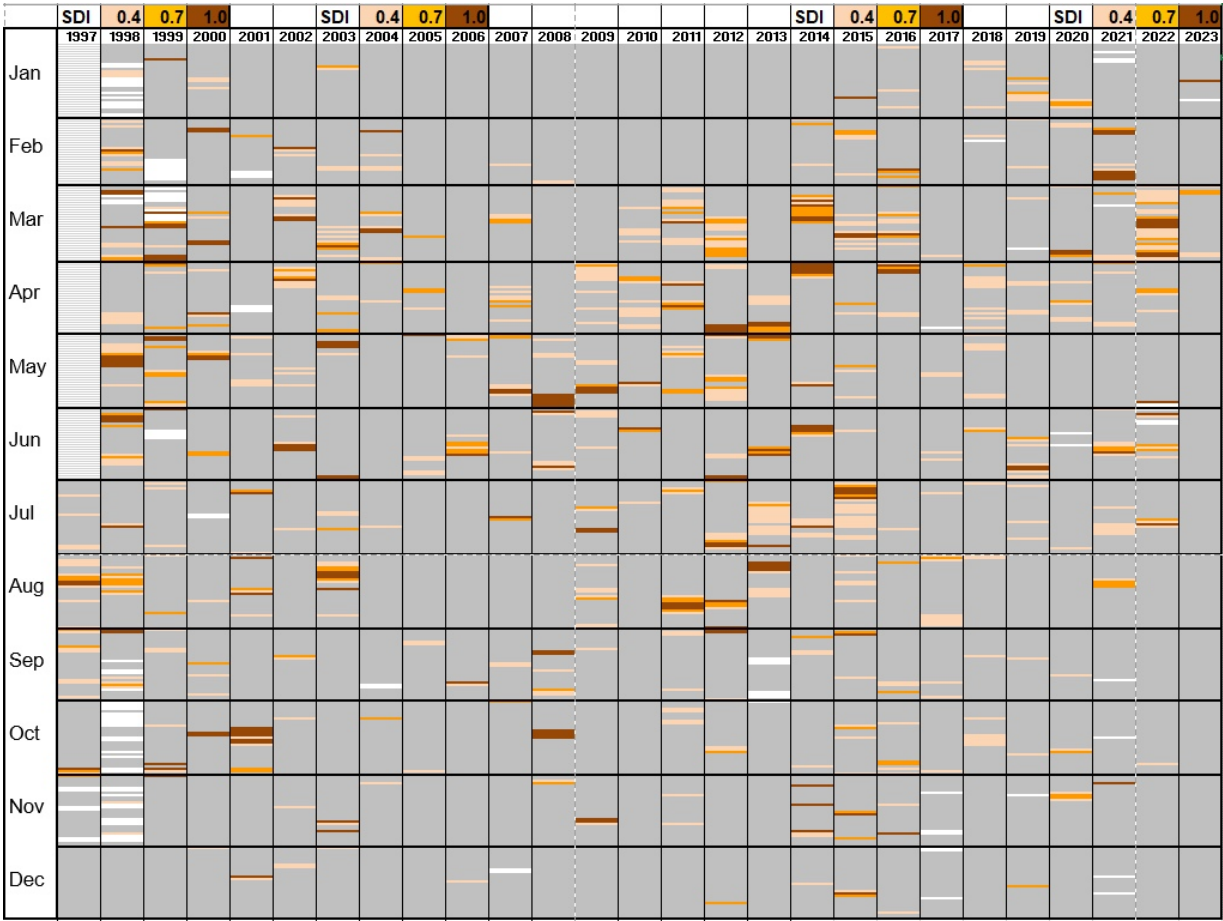


**Abbildung B.7:** LISA für NO<sub>2</sub>-Mittelwerte bei verschiedenen Hauptwindrichtungen.**Abbildung B.8:** LISA für die O<sub>3</sub>-Mittelwerte bei Verschiedenen mittleren Windstärken.

**Abbildung B.9:** LISA für O<sub>3</sub>-Mittelwerte bei verschiedenen Hauptwindrichtungen.

C Zusätzliche Abbildungen

Abbildung C.1: Saharastaubmessungen am Hohenpeißenberg. [149]





# Danksagung

Ich möchte mich Herzlich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Zunächst gilt mein Dank meinem Erstprüfer Prof. Dr. Michael Bittner, der mit seiner Vorlesung „Physik der Atmosphäre“ mein Interesse an der Arbeit im Bereich der Atmosphärenphysik geweckt hat und bereit war, mich für meine Abschlussarbeit am Deutschen Zentrum für Luft und Raumfahrt aufzunehmen.

Außerdem danke ich Dr. Andreas Hörner, für sein Interesse an meiner Arbeit und die Bereitschaft als Zweitgutachter zur Verfügung zu stehen.

Für ihre anhaltende Hilfsbereitschaft und meine freundliche Aufnahme in ihr Team möchte ich mich bei allen Mitgliedern des Teams Spurengase bedanken. Meine Betreuerin Dr. Jana Handschuh konnte mich stets mit ihren Ideen und neuen Perspektiven unterstützen. Ihr gilt mein besonderer Dank für die vielen Stunden wertvoller Beratung und das spenden der nötigen Motivation für die Überwindung zahlreicher Schwierigkeiten. Ebenso danke ich Dr. Frank Baier für seine Betreuung, seine Bereitschaft mich mit Fachwissen zu unterstützen und die Ermutigung meine Fortschritte mit dem Team zu Teilen.

Abschließend möchte ich ein Dankeschön an meine Eltern, Geschwistern und Freunde richten, die mir in dieser Zeit andauernd zur Seite standen und stets bereit waren sich meiner Sorgen anzunehmen.