

Trendanalyse NH₃-Immissionsmessungen in der Schweiz

Schlussbericht

Im Auftrag des Bundesamt für Umwelt (BAFU)

Michel Philipp, René Locher

Winterthur, 2. Dezember 2010

Zürcher Hochschule für Angewandte Wissenschaften
Institut für Datenanalyse und Prozessdesign
Rosenstrasse 3
Postfach
8401 Winterthur

michel.philipp@zhaw.ch
rene.locher@zhaw.ch

Impressum

Auftraggeber:

Bundesamt für Umwelt (BAFU),
Abt. Luftreinhaltung und NIS, CH-3003 Bern

Das BAFU ist ein Amt des Eidg. Departements für Umwelt, Verkehr, Energie und Kommunikation (UVEK).

Auftragnehmer:

Institut für Datenanalyse und Prozessdesign (idp), ZHAW

Autoren:

Michel Philipp, Dr. René Locher

Hinweis:

Diese Studie/dieser Bericht wurde im Auftrag des Bundesamtes für Umwelt (BAFU) verfasst.
Für den Inhalt ist allein der Auftragnehmer verantwortlich.

Inhaltsverzeichnis

1. Zusammenfassung	3
2. Einleitung	6
2.1. Ausgangslage	6
2.2. Ziele der Arbeit	6
2.3. Konzeptionelles Vorgehen	7
3. Zur Verfügung gestellte Daten	8
3.1. NH ₃ -Messreihen	8
3.2. Meteodaten	9
3.2.1. Modifizierte Meteogrößen	9
3.3. Standortinformationen	10
3.4. Zuordnung der Meteostationen zu NH ₃ -Standorten	11
4. Modellbildung	12
4.1. Qualitative Beschreibung der Zusammenhänge	12
4.2. Einfaches statistisches Modell	13
4.3. Erweitertes statistisches Modell	13
4.4. Variablenselektion	14
4.5. Linearisierung von erklärenden Variablen	14
4.6. Robuste Regression und Residuenanalyse	16
4.7. Korrelierte Residuen	18
4.7.1. Zeitlich korrelierte Residuen	18
4.7.2. Räumlich korrelierte Residuen	19
4.8. Bestes Modell	21
4.8.1. Resultate für Luzerner Standorte	23
5. Simulation	24
5.1. Szenarien	24
5.2. Übersicht über Simulationsverfahren	24
5.3. Details zur Berechnung der Simulation	25
5.4. Hinweise zu P-Werten und zur Interpretation der Resultate	25
5.5. Simulationsresultate	26
6. Literatur	28
A. Standorttabelle	29
B. Standorte in den Modellen	30
C. Interaktionen zum Abfangen von gekoppelten Einflüssen	31
C.1. Interaktion zwischen Standort und Monat	32
C.2. Interaktion zwischen Standort und Saison	33
C.3. Interaktion zwischen Meteovariablen und Höhenstationen	33
C.4. Interaktion zwischen Meteovariablen und Saison	34
C.5. Unterscheidung aussergewöhnlicher Standorte	34

Abbildungsverzeichnis

1.1. Exemplarische Messungen und Simulationen	5
3.1. Messplan der Standorte mit Priorität 1	8
3.2. Standorte der Messstationen	10
4.1. Partielle Residuen-Plots	15
4.2. Residuenanalyse	17
4.3. Residuenanalyse im Detail	17
4.4. Residuen gegen Zeit	19
4.5. Korrelationen der Standorte	20
4.6. Korrelationen der Standorte reduziert	21
5.1. P-Werte aus den Simulationen	27

Tabellenverzeichnis

5.1. Zusammenfassung der Simulationsresultate	26
---	----

1. Zusammenfassung

Ammoniak wird nach der Emission in die Atmosphäre verfrachtet, umgewandelt und schliesslich trocken oder nass in naturnahe Ökosysteme eingetragen und trägt damit wesentlich zur Überdüngung und Versauerung dieser Systeme bei. Deshalb wird international versucht, die Ammoniakemissionen zu begrenzen. In der Schweiz wurden ebenfalls entsprechende Massnahmen beschlossen.

Um einen Überblick über die Immissionsituation in der Schweiz zu erhalten, wurde das aus 16 Standorten bestehende Messnetz ab Jahr 2004 auf 27 und ab Jahr 2008 auf 44 Standorte erweitert. Das erweiterte Messnetz wird gemeinsam vom Bundesamt für Umwelt (BAFU), der OSTLUFT und der Kantone Bern, Freiburg, Graubünden, Luzern, Thurgau und Zug betrieben.

Die Messergebnisse dienen unterschiedlichen Zwecken. Einerseits können damit NH_3 -Immissionskarten (punktuell) überprüft werden, welche auf physikalisch-chemischen Modellierungen der NH_3 -Emissionen und deren Ausbreitung in der Atmosphäre beruhen. Andererseits kann damit indirekt die Wirksamkeit der bereits getroffenen Massnahmen verfolgt werden, indem man die Konzentrationsverläufe von NH_3 über mehrere Jahre verfolgt.

Bis dato konnte noch kein negativer Trend in den NH_3 -Immissionen festgestellt werden. Dies ist darauf zurückzuführen, dass bis vor kurzem noch zu wenige und zu wenig lange Messreihen existierten und die Minderungsmassnahmen erst langsam zu greifen beginnen. Zudem sind die Immissionskonzentrationen generell sehr stark von der Witterung abhängig, so dass zuerst mit statistischen Verfahren die Witterungseffekte aus den Konzentrationsverläufen herausgerechnet werden müssen, damit kleinere Veränderungen überhaupt detektiert werden können.

Das Bundesamt für Umwelt erteilte deshalb dem IDP den Auftrag, mit statistischen Methoden die folgenden Fragen zu beantworten:

- (i) Kann bereits ein Trend festgestellt werden: Schweizweit und über das ganze Jahr oder wenigstens in einzelnen Grossregionen während bestimmten Jahreszeiten?
- (ii) Wenn nein: Wie lange muss gemessen werden, bis eine allfällige Konzentrationsminderung von -2% pro Jahr bzw. -20% in 10 Jahren festgestellt werden kann?
- (iii) Wie viel länger müsste man messen, wenn nur halb so viele Standorte vorhanden wären?

Zur Beantwortung dieser Fragen wurde ein statistisches Regressionsmodell entwickelt, welches die NH_3 -Konzentrationen als Funktion des Standortes und der herrschenden Witterung beschreibt. Die Entwicklung des Modells basierte auf den Messdaten von insgesamt 40 Standorten, welche das Bundesamt für Umwelt selektierte. Alle Standorte waren von landwirtschaftlichen Emissionen beeinflusst und wiesen eine Messreihe von mindestens zwei Jahren Länge auf. Die 8 längsten Messreihen waren 10 Jahre lang.

Eine ausgedehnte Prüfung des entwickelten Regressionsmodells zeigte, dass 9 dieser Standorte aussergewöhnliche Abweichungen vom Modell (d.h. erhöhte Varianz in den Residuen) aufwiesen. Es handelte sich dabei um Standorte die um oder über 800 m.ü.M liegen, welche einen anderen saisonalen Verlauf (z.B. längere Winter) zeigten oder welche an atypischen Standorten (z.B. am Waldrand) platziert waren. Weiter zeigte diese Analyse, dass die Messwerte vor 2003 systematische Abweichungen aufwiesen. Der Grund dafür ist unklar.

In den Gemeinden Wauwil, Eschenbach und Schüpflheim werden aussergewöhnlich viele Standorte beprobt. Durch ihre räumliche Nähe sind diese Messungen verstärkt den gleichen Emissionsquellen ausgesetzt und korrelieren deshalb besonders hoch miteinander. Aus Gründen der

Repräsentativität für die landwirtschaftlich beeinflussten Gebiete im ganzen Schweizer Mittelland wurden die Passivsammler auf je einen Standort pro Dorf reduziert, so dass im Modell noch 17 Standorte enthalten waren. Dies hatte gleichzeitig den positiven Nebeneffekt, dass sich damit die räumliche Korrelation zwischen den einzelnen Standorten reduzieren liess und sich damit die Überprüfung des Trends vereinfachte bzw. dieser damit erst möglich wurde.

Das gesamtschweizerische Regressionsmodell erklärt die NH_3 -Konzentrationen mit den folgenden Grössen, geordnet nach abnehmender Bedeutung:

1. Standort
2. Monat des Jahres (Saison)
3. Monatliche Abweichung der Temperatur vom 10-jährigen Mittel
4. Monatliche Abweichung der Anzahl Trockentage vom 10-jährigen Mittel
5. Monatliche Abweichung der Windgeschwindigkeit vom 10-jährigen Mittel
6. Trend (Zeitliche Veränderung)

Die NH_3 -Immissionen steigen gemäss Modell mit steigender Lufttemperatur und Erhöhung der mittleren Anzahl Trockentage und sinken mit zunehmenden Wind, wie dies auf Grund luft-hygienischer Überlegungen erwartet wird. Die wichtigsten Faktoren für die Bestimmung der Immissionshöhe sind der Standort und der Monat des Jahres. Die Einflüsse der monatlichen Abweichung der Witterung vom langjährigen Mittel ist gering, da der grösste Einfluss der Witterung bereits im Monatseffekt enthalten ist. In den vorliegenden Daten kann auch mit dem statistischen Modell noch kein signifikanter Trend beobachtet werden, was auf Grund des Massnahmenplans für die NH_3 -Reduktionsmassnahmen auch nicht zu erwarten war. Auch für das Gebiet des Kantons Luzern konnte kein negativer Trend detektiert werden.

Das gefundene statistische Modell wurde nun im zweiten Schritt dazu verwendet, (realistische) Szenarien mit künstlich eingefügten Trends für die Zukunft zu simulieren. Die Abhängigkeiten von Standort und Saison wurden aus der Analyse der gemessenen Daten entnommen, die mittleren Abweichungen der Witterung vom langjährigen Mittel und die zufälligen Abweichungen der Messwerte vom Modell (= Residuen) wurden ebenfalls auf geeignete Art und Weise zufällig aus dem Originalmodell übernommen. Dies ergab für die Standorte simulierte Zeitreihen der Zukunft, wie sie exemplarisch in Abb. 1.1 dargestellt sind. Pro Szenario wurde die Simulation jeweils 500 Mal wiederholt und pro Simulation die Steigung des Trends und dessen P-Wert aus den Zeitreihen berechnet.

Die Analyse der Verteilungen der P-Werte zeigte, dass ein Trend von -1% pro Jahr nach 10 Jahren Messzeit signifikant erkennbar sein wird, wenn für die Analyse mindestens 7 nicht oder nur wenig korrelierende Standorte verwendet werden. Eine Analyse nach 5 Jahren wird nicht empfohlen. Wenn sie trotzdem durchgeführt wird, so müssen mindestens 12 nicht korrelierende Standorte ausgewertet werden und der erwartete Trend muss stärker als -2% pro Jahr sein. Trotzdem muss bei nur 5 Jahren Messdaten immer damit gerechnet werden, dass allfällig beobachtete Ab- und Zunahmen während ein oder zwei Jahren auch durch Sondereffekte verursacht wurden und nicht durch eine ursächliche Reduktion der Emissionen. Wenn diese Sondereffekte zufällig am Ende bzw. am Anfang einer Beobachtungsreihe lägen, so könnte irrtümlich auf eine Zu- bzw. Abnahme der Immissionen geschlossen werden. Nach zehnjähriger und noch längerer Beobachtungszeit können diese zufälligen Effekte praktisch ausgeschlossen werden.

Mit der hier vorgestellten Methode können auch kompakte Teilgebiete der Schweiz, welche mindestens die Grösse des Kantons Luzern aufweisen, untersucht werden, sofern mindestens 7 Standorte während mindestens 10 Jahren beprobt werden. Allerdings werden die gewonnenen Aussagen weniger stringent als für die Gesamtschweiz ausfallen, da die Messungen an diesen Standorte mit

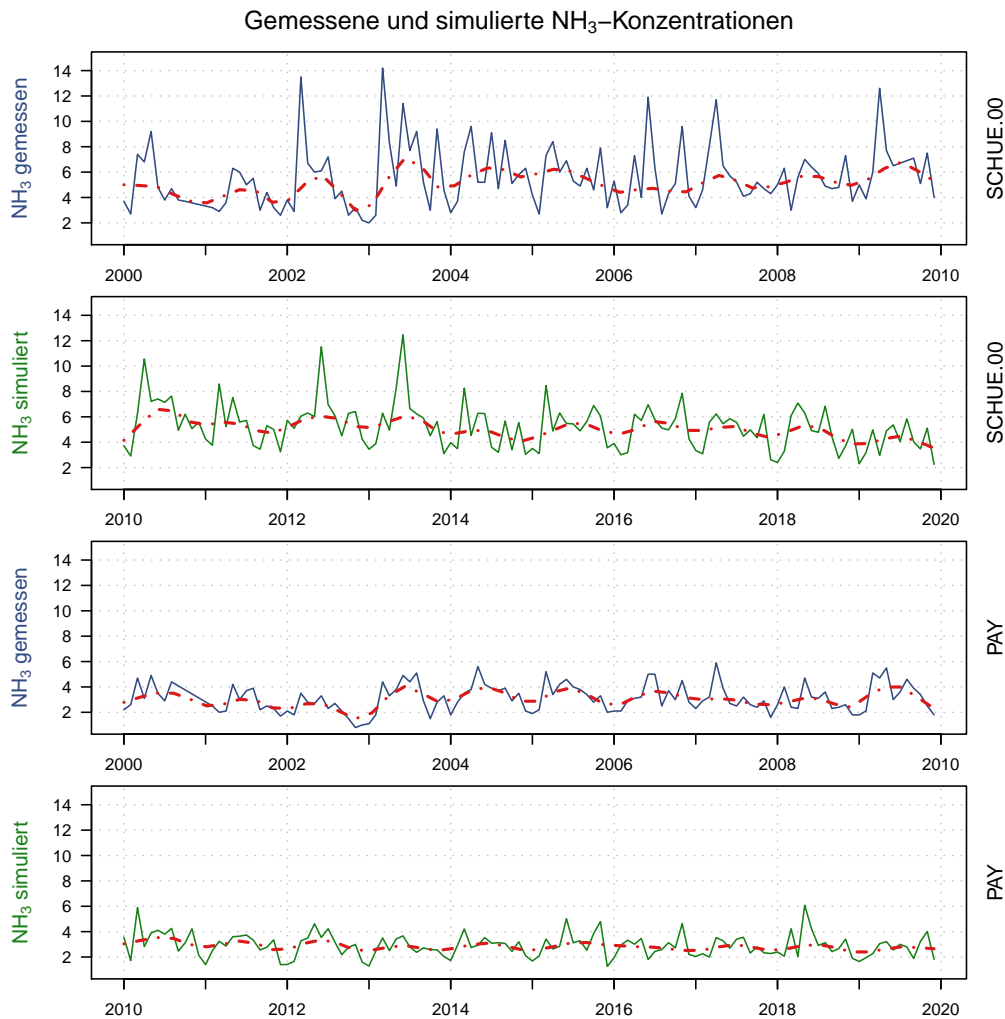


Abbildung 1.1: Reale Messungen und exemplarische Simulationen der Standorte in Schüpffheim (SCHUE.00) und Payerne (PAY). Von Auge ist der in den simulierten Reihen künstlich eingefügte, statistisch signifikante, negative Trend von -2% pro Jahr kaum erkennbar.

grosser Wahrscheinlichkeit untereinander stark korreliert sind und damit die hier angewandten Signifikanztests versagen.

Je weniger die Messungen untereinander korrelieren, um so höher ist die Qualität der Trendanalyse. Minimale Korrelationen haben die beprobten Standorte, wenn sie möglichst weit auseinander in verschiedenen Geländekammern liegen. Die Messwerte können auf wenigen Metern stark variieren. Deshalb sollte der Standort über die ganze Beobachtungsperiode am exakt gleichen Ort betrieben werden und es sollten keine grossen Emissionsquellen in unmittelbarer Nähe des Standortes vorhanden sein. Damit diese Anforderung über 10 oder mehr Jahre mit Sicherheit erfüllt ist, können bei Bedarf auch mehrere Passivsammler in der Nähe aufgestellt werden. Für die Analyse wird dann aber nur derjenige Standort berücksichtigt, welcher den obigen Kriterien am besten entspricht.

2. Einleitung

2.1. Ausgangslage

Ammoniak wird nach der Emission in die Atmosphäre verfrachtet, umgewandelt und schliesslich trocken oder nass in naturnahe Ökosysteme eingetragen und trägt damit wesentlich zur Überdüngung und Versauerung dieser Systeme bei [1]. Deshalb wird international versucht, die Ammoniakemissionen zu begrenzen. In der Schweiz wurden ebenfalls entsprechende Massnahmen beschlossen.

Um einen Überblick über die Immissionsituation in der Schweiz zu erhalten, wurde das aus 16 Standorten bestehende Messnetz ab Jahr 2004 auf 27 und ab Jahr 2008 auf 44 Standorte erweitert. Das erweiterte Messnetz wird gemeinsam vom Bundesamt für Umwelt (BAFU), der OSTLUFT und der Kantone Bern, Freiburg, Graubünden, Luzern, Thurgau und Zug betrieben [2].

Die Messergebnisse dienen unterschiedlichen Zwecken. Einerseits können damit NH_3 -Immissionskarten (punktuell) überprüft werden, welche auf physikalisch-chemischen Modellierungen der NH_3 -Emissionen und deren Ausbreitung in der Atmosphäre beruhen. Andererseits kann damit indirekt die Wirksamkeit der bereits getroffenen Massnahmen verfolgt werden, indem man die Konzentrationsverläufe von NH_3 über mehrere Jahre verfolgt.

Bis dato konnte gemäss dem Bericht über Ammoniak-Immissionsmessungen in der Schweiz [3] der Forschungsstelle für Umweltbeobachtung (FUB) noch kein negativer Trend in den NH_3 -Immissionen festgestellt werden. Dies ist darauf zurückzuführen, dass bis vor kurzem noch zu wenige und zu wenig lange Messreihen existierten und die Minderungsmassnahmen erst langsam zu greifen beginnen. Zudem sind die Immissionskonzentrationen generell sehr stark von der Witterung abhängig, so dass zuerst mit statistischen Verfahren die Witterungseffekte aus den Konzentrationsverläufen herausgerechnet werden müssen, damit kleinere Veränderungen überhaupt detektiert werden können.

2.2. Ziele der Arbeit

In dieser Arbeit sollen mittels statistischer Methoden die folgenden Fragen beantwortet werden:

- (i) Kann bereits ein Trend festgestellt werden: Schweizweit und über das ganze Jahr oder wenigstens in einzelnen Grossregionen während bestimmten Jahreszeiten?
- (ii) Wenn nein: Wie lange muss gemessen werden, bis eine allfällige Konzentrationsminderung von -2% pro Jahr festgestellt werden kann?
- (iii) Wie viel länger müsste man messen, wenn nur halb so viele Standorte vorhanden wären?

Zur Beantwortung der obigen Fragen werden die folgenden Voraussetzungen gemacht:

- Die zukünftige Variabilität der Daten entspricht der vergangenen.
- Die beprobten Standorte sind für die Schweiz (oder wenigstens für eine bestimmte Region bzw. Saison) repräsentativ.
- Die Umsetzung der Massnahmen zur Minderung der NH_3 -Emissionen erfolgt synchron in der ganzen Schweiz oder wenigstens in einzelnen Grossregionen.
- Der Trend ist monoton gleichförmig abnehmend.

2.3. Konzeptionelles Vorgehen

Um die Fragestellungen zu einem allgemeinen Trend in den NH_3 -Immissionen beantworten zu können, müssen die Messungen mit Hilfe eines statistischen Modells von witterungs- und standortabhängigen Einflüssen bereinigt werden. Das Vorgehen bei der Modellbildung entsprach einem iterativen Prozess aus den Schritten Datenaufbereitung, Modellanpassung und Residuenanalyse (= Prüfen der statistischen Modellvoraussetzungen). Das Ziel war erreicht, sobald ein Modell gefunden worden war, welches die Daten nach den verschiedenen Kriterien best möglichst beschrieb (= bestes Modell). Aus dem besten Modell wurde direkt der von allen Störgrößen bereinigte Trend abgelesen (Fragestellung i). Das Modell wurde anschliessend in einem zweistufigen Verfahren dazu verwendet, um die weiteren Fragestellungen zu bearbeiten:

- Simulation der NH_3 -Konzentration unter einem künstlich vorgegebenen, negativen Trend von -1% bzw. -2% pro Jahr für die nächsten 5 bzw. 10 Jahre.
- Schätzen des Trends und des dazugehörigen P-Wertes aus den simulierten Prognosewerten.

Dieser Vorgang wurde 500 mal wiederholt, um die Verteilung der daraus geschätzten Trends und der dazugehörigen P-Werte zu bestimmen und damit die Frage (ii) zu beantworten. Mit einer reduzierten Auswahl an Standorten wurde die Fragestellung (iii) nach dem gleichen Vorgehen bearbeitet.

3. Zur Verfügung gestellte Daten

Zur Beantwortung der Fragestellungen aus Kapitel 2.2 hatte das BAFU dem IDP Daten von NH_3 -Immissionsmessungen¹ verschiedener Standorte in der Schweiz, sowie Meteodaten und Informationen zu den Messstandorten zur Verfügung gestellt.

3.1. NH_3 -Messreihen

Die Messreihen der NH_3 -Immissionskonzentrationen standen dem IDP als Perioden- und als Monatsmittel zur Verfügung. Die periodisch gemessenen Werte waren von der FUB zu Monatsmitteln aggregiert worden, wobei monatsübergreifende Periodenmittel entsprechend ihrem Anteil an Messtagen in den betroffenen Monaten gewichtet worden waren. Das BAFU teilte die gelieferten NH_3 -Werte gemäss ihrer Standorte in 3 Prioritätsstufen ein (vgl. Kap. 3.3).

Einerseits begannen die Messreihen zu unterschiedlichen Zeitpunkten. Andererseits gab es einige wenige fehlende Messwerte (diese wurden absichtlich nicht ersetzt oder interpoliert). Beide Informationen können dem Messplan in Abb. 3.1 pro Messstation entnommen werden.

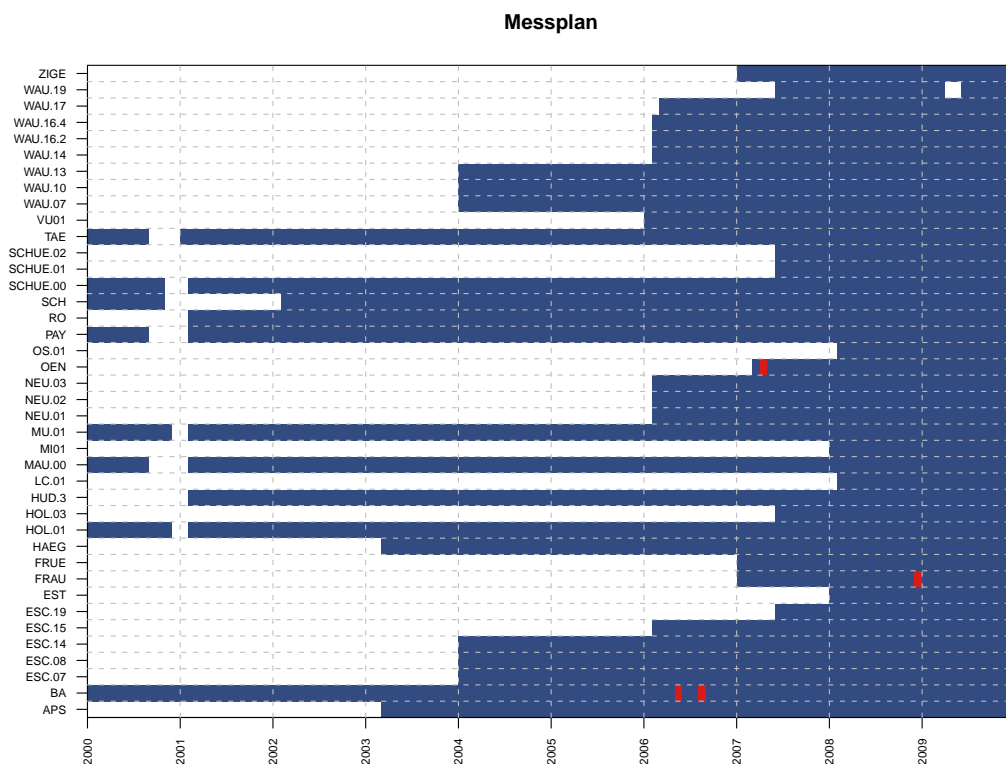


Abbildung 3.1: Messplan der Standorte mit Priorität 1: Während den blau eingefärbten Perioden waren von den entsprechenden Standorten Monatsmittel verfügbar. An den rot eingefärbten Stellen waren von der zugehörigen Meteostation nicht genügend Meteo-Daten verfügbar um ein Monatsmittel zu bestimmen.

¹Die Messungen der Ammoniakimmissionen wurden von der Forschungsstelle für Umweltbeobachtung (FUB) durchgeführt. Methodische Überlegungen zum Messverfahren [3] und zu den Messstandorten können dem Bericht [2] bzw. der Standortdatenbank [4] der FUB entnommen werden.

3.2. Meteodaten

Die Meteodaten stammten von Meteo-Schweiz- und NABEL-Stationen. Sie wurden vorgängig zu Tageswerten aggregiert und beinhalteten die Messgrössen

- Tagesmittel des Dampfdrucks 2m über Boden in [hPa],
- Tagesmaximum der Lufttemperatur 2m über Boden in [°C],
- Tagesminimum der Lufttemperatur 2m über Boden in [°C],
- Tagesmittel der Lufttemperatur 2m über Boden in [°C],
- Niederschlag pro Kalendertag in [mm],
- Tagesmittel der relativen Luftfeuchtigkeit 2m über Boden in [%] und
- Tagesmittel der Windgeschwindigkeit (skalar) in [m/s].

Das IDP aggregierte die täglichen Meteodaten auf Monatswerte, damit sie mit der Auflösung der NH₃-Immissionsmessungen übereinstimmten. Unter der Bedingung, dass in einem Monat mehr als 80% der Tageswerte vorhanden waren, wurden stetige Grössen wie Temperatur und Windgeschwindigkeit und Zählraten wie Anzahl Regentage oder Anzahl Trockentage gemittelt. Die Zählraten wurden anschliessend mit der Anzahl Tage des Monats multipliziert (= einfache Imputation der Daten).

3.2.1. Modifizierte Meteogrössen

Auf Grund lufthygienischer Überlegungen sollten gewisse Transformationen der Meteogrössen, im Folgenden Meteovariablen genannt², die Immissionen besser erklären als die ursprünglich gemessenen Grössen. Z.B. ist es entscheidend wie der Regen im Monat auf die einzelnen Tage verteilt ist. Der Grund liegt im Düngeverhalten der Bauern, welche u.a. auf Grund der Nässe des Bodens entscheiden, ob sie Jauche ausbringen oder nicht. Es wurden folgende Anpassungen der Meteovariablen getestet:

- Differenz³ von Tagesmaximum und Tagesminimum gemittelt über einen Monat.
- Anzahl Tage pro Monat mit einer Tagestemperaturdifferenz < 3 [C°].
- Anzahl Tage pro Monat mit einer Tagestemperaturdifferenz < 5 [C°].
- Anzahl Trockentage, an denen die Regenmenge 0 ist.
- Anzahl Trockentage, jeweils gerechnet ohne den ersten Tag einer Trockenperiode⁴ (im Folgenden mit ATT abgekürzt).
- Anzahl Trockentage mit einer Regenmenge im Intervall [0, 0.1] mm.
- Anzahl Tage mit einer Regenmenge im Intervall [0.1, 1] mm.

²Die in die Regression eingehenden Grössen werden in der Statistik "Variablen" genannt. In der Modellgleichung wird die Variable auf der linken Seite des Gleichheitszeichens als Zielvariable bezeichnet und soll mit Hilfe der Variablen auf der rechten Seite, den sogenannten erklärenden Variablen, erklärt werden.

³Bei über die ganze Nacht andauernden Inversionen kühlt die Luft weniger stark ab als wenn keine Inversion herrscht, so dass bei Inversion geringere Temperaturdifferenzen zwischen Tag und Nacht erwartet werden. Es wurde getestet, ob die mittlere Temperaturdifferenz allenfalls als Mass für die Anzahl und Dauer der Inversionen an einem bestimmten Standort verwendet werden kann.

⁴Direkt nach einer Regenperiode ist der Boden immer noch feucht, was einen anderen Effekt auf die NH₃-Emission hat, als trockener Boden.

- Anzahl Tage mit einer Regenmenge im Intervall $[1, 10]$ mm.
- Anzahl Tage mit einer Regenmenge ≥ 10 mm.
- Anzahl Kältetage mit einer mittlere Tagestemperatur < 0 [C°].

3.3. Standortinformationen

Die Informationen für NH₃-Messstandorte beinhalteten neben Namen, Kurzbezeichnungen⁵, Koordinaten, Standorthöhen und weiteren standortspezifischen Informationen eine vom BAFU erstellte Priorisierung, nach welcher die Messreihen untersucht werden sollen. Alle berücksichtigten Standorte waren primär durch die Landwirtschaft beeinflusst. Priorität 1 erhielten die Standorte, welche in ganzjährig bewirtschafteten Landwirtschaftszonen des Schweizer Mittellandes und der Voralpen lagen und deren Messperiode länger als zwei Jahre war. Priorität 2 erhielten analoge Standorte im Tessin und in der Region Basel. Mit Priorität 3 wurden Alpweiden versehen, welche nur während der warmen Jahreszeit bewirtschaftet wurden. Damit standen dem IDP 40 Standorte mit Priorität 1, 9 Standorte mit Priorität 2 und 7 Standorte mit Priorität 3 zur Verfügung. Für den vorliegenden Bericht wurden nur Standorte mit Priorität 1 berücksichtigt. In Abb. 3.2 sind vollständigshalber alle gelieferten NH₃-Messstationen auf einer Schweizerkarte eingetragen.

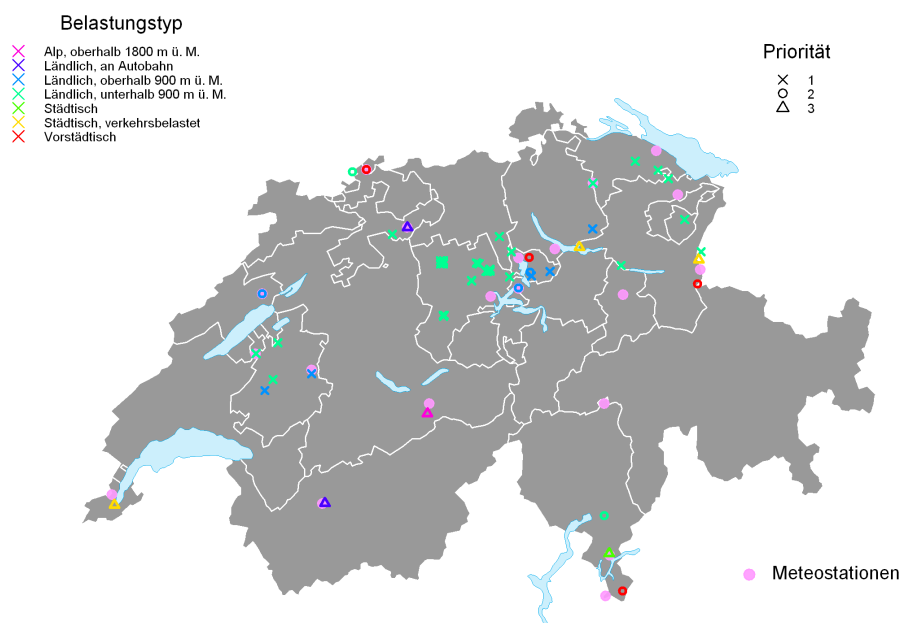


Abbildung 3.2: Standorte der NH₃-Messstationen nach Priorität und Belastungstyp und Meteo-Messstationen.

Detaillierte Beschreibungen der NH₃-Standorte (u.a. Beschreibung der umliegenden Natur und Landwirtschaft, Bilder und Standortkarten) wurden in einer Standortdatenbank zusammenge-

⁵Im vorliegenden Bericht werden die Messstandorte mit diesem Code identifiziert, vgl. Anhang A.

stellt, auf welche das IDP ebenfalls Zugriff hatte. Von den Meteo-Messstationen waren Namen, Kurzbezeichnungen Koordinaten und Standorthöhen verfügbar (vgl. auch Anhang A).

3.4. Zuordnung der Meteostationen zu NH₃-Standorten

Die Zuordnung der Meteostandorte zu den NH₃-Messstationen wurde vom BAFU vorgenommen. Dabei wurde darauf geachtet, dass sich die Meteostation in der Nähe der NH₃-Messtation und vor allem in ähnlicher Höhenlage befand.

Die Zuteilung wurde in Absprache mit dem BAFU für die Standorte BA, FRAU und MU.01 modifiziert, da die ursprünglich gewählten Meteostationen HOE und CHZ viele fehlende Tageswerte aufwiesen. Für BA wurde neu die Meteostation RIG, für FRAU und MU.01 die Meteostation LUZ eingesetzt.

4. Modellbildung

Die Modellbildung war ein komplexer, iterativer Prozess, dessen Vorgehen hier schematisch beschrieben wird. Zu Gunsten der Übersicht kommen Fachausdrücke vor, welche erst in den entsprechenden Unterkapiteln näher erläutert werden.

1. Qualitative Beschreibung der Zusammenhänge zwischen den NH_3 -Konzentrationen und den übrigen (Mess-)Grössen, welche zur Verfügung standen (Kap. 4.1).
2. Herleitung eines einfachen linearen Modells (Kap. 4.2).
3. Ausbau zum generalisierten additiven Modell (Kap. 4.3).
4. Selektion der erklärenden Variablen inkl. Überprüfung auf allfällige Interaktionen zwischen erklärenden Variablen (Kap. 4.4).
5. Linearisierung der Variablen (Kap. 4.5).
6. Robustifizierung des Modells und Residuenanalyse⁶. Evtl. weglassen von nicht dazu passenden Beobachtungen (Kap. 4.6).
7. Überprüfung der zeitlichen und räumlichen Korrelationen. Reduktion der Korrelationen durch Ausdünnung der Standorte. Weiter mit Schritt 4., sofern nicht bereits bestes Modell gefunden.

Die Schritte 4. bis 7. wurden iterativ durchlaufen, bis die Analyse in allen Aspekten konsistent war. Auf jeder dieser Stufen wurde eine Residuenanalyse vorgenommen. Besonders zentral war aber die Residuenanalyse des robusten Modells (6.). In den folgenden Unterkapiteln werden die vorgenommenen Analysen kommentiert, wobei aus Gründen der Übersichtlichkeit die iterativen Arbeiten nur exemplarisch beschrieben werden.

4.1. Qualitative Beschreibung der Zusammenhänge

Auf Grund lufthygienischer Überlegungen kann davon ausgegangen werden, dass die Höhe der Immissionen u.a. abhängt von

- (i) der Frequenz und Art der Ausbringung der Jauche,
- (ii) der Dichte, Grösse und Bauart von Ställen,
- (iii) der Distanz der Messstation zu Ställen, Gewässern und Pflanzen,
- (iv) der Temperatur,
- (v) den Windverhältnissen,
- (vi) der Boden- und Luftfeuchtigkeit.

Diese Liste von Einflüssen kann grob in standortspezifische (i) - (iii) und meteorologische (iii) - (vi) Einflüsse unterteilt werden, wobei auch (iii) - (vi) teilweise als standortspezifisch betrachtet werden muss. Mit Hilfe statistischer Regressionsmethoden erlaubten die vorhandenen Daten, einen von diesen Störgrössen bereinigten Trend zu schätzen.

⁶Streng genommen wird die Residuenanalyse (in geringerem Umfang) bei jedem Schritt ab Nr. 4. durchgeführt.

4.2. Einfaches statistisches Modell

Die obigen Überlegungen führten zum vereinfachten statistischen Modell, einem sogenannten Regressionsmodell:

$$\log(\text{NH}_3_t^{(s)}) = \beta_0 + \text{Standort}^{(s)} + \text{Monat}_t + \beta_{\text{Trend}} \cdot t + \beta_T \cdot \Delta T_t^{(10J)} + \beta_{WS} \cdot \Delta WS_t^{(10J)} + \beta_{ATT} \cdot \Delta ATT_t^{(10J)} + \varepsilon_t \quad (4.1)$$

Das Modell kann wie folgt interpretiert werden: Die logarithmierten NH_3 -Konzentrationen sind bis auf einen zufälligen Fehler ε_t linear abhängig vom Standort s , der Saison, der Meteo und der Zeit t . Die Ammoniakkonzentration NH_3 geht logarithmiert ins Modell ein. Damit beeinflussen die einzelnen erklärenden Variablen die Konzentration multiplikativ wie wir das aus physikalisch-chemischen Gründen auch erwarten würden. Gleichzeitig liegt damit auch die Fehlerstruktur näher an der Normalverteilung und es ist sichergestellt, dass bei Prognosen mit diesem Modell nie negative Konzentrationen auftreten können⁷.

Der Standort ist eine Faktorvariable, deren Faktorstufen den einzelnen Standorten s entsprechen und implizit alle standortspezifischen Abhängigkeiten abbildet.

Die Saison ist ebenfalls eine Faktorvariable, deren Faktorstufen hier in der Regel dem Monat des Jahres (also Januar, Februar etc.) entsprechen. Es wurden allerdings auch Modelle untersucht, in welchen verschiedene Monate, z.B. die Wintermonate, zu einer einzigen Faktorstufe zusammengefasst wurden.

Die Meteogrößen sind explizit als monatliche Abweichungen vom 10-jährigen Monatsmittel bzw. Monatssumme im Modell enthalten. Z.B. entspricht $\Delta T^{(10J)}$ für $t = \text{Januar } 2009$ der Differenz zwischen dem Monatsmittel des Januars 2009 und dem Mittel aller Januarmonate der Jahre 2000 bis 2009. Die 10-jährigen Monatsmittel wurden jeweils pro Meteostandort bestimmt. Dahinter steckt die Idee, dass die NH_3 -Konzentration in einer gewissen Periode regional von der durchschnittlichen Witterung abweichen kann⁸. Im Gegensatz dazu steckt die standortspezifische Abweichung von der gesamtschweizerischen Witterung⁹ implizit in der Variable Standort. Der generelle gesamtschweizerische Meteoverlauf und die davon abhängigen Düngeaktivitäten der Bauern ist implizit in der Variable Saison enthalten.

4.3. Erweitertes statistisches Modell

Das im vorhergehenden Kapitel beschriebene Modell scheint zwar bereits relativ komplex, bildet aber die Wirklichkeit noch zu wenig genau ab. Einerseits sind die potentiellen Einflüsse der Variablen auf den Logarithmus der NH_3 -Konzentrationen nicht zwingend linear, andererseits können die Einflüsse einzelner Variablen von anderen abhängen. Beide Phänomene wurden untersucht. Während die nichtlinearen Phänomene im endgültigen Modell enthalten sind, wurden die Interaktionen als für nicht notwendig erachtet. Die Interaktionen sind zwar durchaus vorhanden, sind aber nicht sehr ausgeprägt. Da Interaktionen die Allgemeingültigkeit des Modells sehr stark einschränken¹⁰ und die Anzahl Daten zur Beschreibung dieser Interaktionen trotz der Grösse des Datensatzes zu klein ist, wurde auf diese im endgültigen Modell verzichtet. Detailliertere Angaben zu den Interaktionen finden sich im Anhang C. Im Folgenden wird

⁷Begründung: $e^x \geq 0$ für beliebige reelle x .

⁸Beispielsweise aussergewöhnlich viel Regen im August und September 2006 in der Westschweiz aber nicht in der übrigen Schweiz

⁹Z.B. ist die mittlere Jahrestemperatur u.a. eine Funktion der Höhe ü. Meer und der Orographie.

¹⁰Wenn z.B. Interaktionen zwischen dem Standort und einer Meteovariablen auftreten, so kann der Einfluss der Meteovariablen nur noch in Zusammenhang mit einem spezifischen Standort diskutiert werden. Eine allgemeine Abhängigkeit der Immission von dieser Meteovariablen ist dann im Extremfall gar nicht mehr beschreibbar.

nur auf die Untersuchung der nichtlinearen Phänomene mittels sogenannten GAM-Modellen¹¹ eingegangen.

Wie der Name sagt, sind GAMs eine Generalisierung der gewöhnlichen linearen Regression. Wenn man in der Gleichung für die lineare Regression 4.1 die Koeffizienten der linearen Terme der kontinuierlichen Variablen $\beta_{(j)} \cdot x^{(j)}$ durch glatte Funktionen $g_{(j)}(x^{(j)})$ ersetzt, erhält man das entsprechende GAM-Modell:

$$\log(NH_3^t) = \beta_0 + Standort^{(s)} + Monat_t + g_{Trend}(t) + g_T(\Delta T^{(10J)}) + g_{WS}(\Delta WS^{(10J)}) + g_{ATT}(\Delta ATT^{(10J)}) + \varepsilon_t \quad (4.2)$$

Mit sogenannten partiellen Residuenplots kann der Zusammenhang zwischen der Zielvariablen und der Variablen $x^{(j)}$ dargestellt werden, wobei jeweils die Effekte der übrigen Variablen herausgerechnet werden. Die Kurven in den einzelnen Teilbildern entsprechen gerade der Form der glatten Funktionen $g_{(j)}$ bei den kontinuierlichen Variablen bzw. den Stufen der Faktorvariablen (vgl. Abb. 4.1). Der Wertebereich, welcher die Funktion in der Vertikalen abdeckt, entspricht gerade dem Einfluss der entsprechenden Variablen auf die Zielvariable.

4.4. Variablenselektion

Die generalisierten additiven Modelle wurden in dieser Arbeit verwendet, um (i) die beste Auswahl aus den verfügbaren Meteovariablen zu treffen¹² und (ii) allfällige nichtlineare Strukturen zu detektieren.

Es wurden vorwiegend Modelle miteinander verglichen, welche sich in den Meteovariablen Temperatur und Niederschlag unterscheiden (vgl. Kapitel 3.2.1). Der Selektionsprozess ergab (nach einigen Iterationen) als bestes Modell das GAM (4.2). Dieses Modell beschreibt die Daten gut, weist nur wenige nichtlineare Strukturen auf und kann lufthygienisch sinnvoll interpretiert werden. Die partiellen Residuen-Plots sind in Abb. 4.1 dargestellt.

4.5. Linearisierung von erklärenden Variablen

Im Idealfall werden robuste Methoden zur Endüberprüfung des besten Modells und zur Schätzung des Trends eingesetzt. Begründung: Bei den konventionellen Verfahren können bereits einzelne Ausreisser die Analyse komplett wertlos machen. Wenn das Modell mehr als eine erklärende Variable enthält, kann dies in der Regel nicht einmal mit einer Residuenanalyse entdeckt werden. Um ein vollständig parametrisierbares Modell zu erhalten, wurde das GAM-Modell zuerst in ein gewöhnliches lineares Modell überführt, um es anschliessend robust analysieren zu können.

Auf Grund der partiellen Residuen-Plots wurde das (nach einigen Iterationen erhaltene)¹³ beste Modell für 17 Standorte wie folgt stückweise linearisiert. Die Konfidenzintervalle der partiellen Residuen wurden dabei grosszügig ausgelegt, da es sich um (leicht) korrelierte Daten handelt (vgl. Kap. 4.7), welche zu engen Konfidenzintervallen führen können.

¹¹= Generalized Additive Models = generalisierte additive Modelle

¹²d.h. die Variablen mit grossem Einfluss auf die NH_3 -Konzentration zu finden, deren partielle Residuenplots zusätzlich eine physikalisch-chemisch sinnvolle Kurvenform aufweisen.

¹³Die Begründung für die Reduktion der Standorte und der Beobachtungsperiode wird im Kapitel 4.6 erläutert.

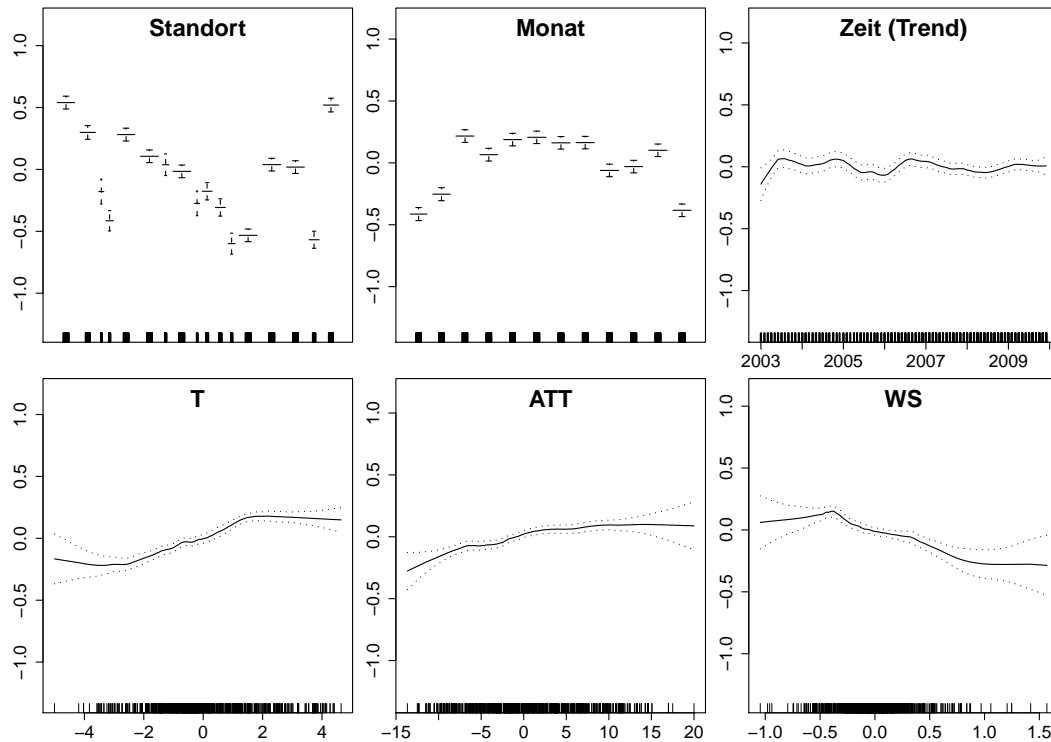


Abbildung 4.1: Partielle Residuen-Plots des besten Modells 4.2 unter Berücksichtigung der selektierten 17 Standorte zwischen Anfang 2003 und Ende 2009. Auf der x -Achse sind die Werte der Variablen und auf der y -Achse der partielle Einfluss auf die NH_3 -Konzentration dargestellt. Der grösste Einfluss stammt von der Variablen Standort. Die saisonale Abhängigkeit der Immissionen kann dem mittleren Bild oben entnommen werden. Die generell höheren Werte im Sommer sind auf die erhöhte Temperatur und die erhöhten Düngeaktivitäten der Landwirte in dieser Jahreszeit zurückzuführen. Die kleinen lokalen Maxima im März, Oktober und November sind auf das erste Ausbringen von Jauche im Frühjahr (Leeren des vollen Jauchetanks) und das letzte Ausbringen von Jauche vor Winterbeginn (Leeren des Jauchetanks vor dem Winter) zurückzuführen. Mit zunehmender Temperatur (T), zunehmender Anzahl Trockentage (ATT) und abnehmender Windgeschwindigkeit (WS) nimmt die NH_3 -Konzentration zu, wie man dies aus lufthygienischen Überlegungen erwarten würde. Man beachte, dass der Haupteffekt dieser drei Meteorvariablen indirekt im Saisoneffekt steckt und die drei letzten Figuren nur den Effekt der aktuellen Abweichung vom langjährigen saisonalen Mittel darstellt. Die geschätzten Kurven sind glatt und können leicht linearisiert werden. Über die Beobachtungsperiode wird kein Trend detektiert. Die kleineren systematischen Abweichungen von der Nulllinie sind auf kleinere Unzulänglichkeiten des Modells zurückzuführen.

$$\Delta T^{(10J)} = \begin{cases} -3, & \text{wenn } \Delta T^{(10J)} < -3 \\ 1.5, & \text{wenn } \Delta T^{(10J)} > 1.5 \\ \Delta T^{(10J)}, & \text{sonst} \end{cases} \quad (4.3)$$

$$\Delta ATT^{(10J)} = \begin{cases} 3.5, & \text{wenn } \Delta ATT^{(10J)} > 3.5 \\ \Delta ATT^{(10J)}, & \text{sonst} \end{cases} \quad (4.4)$$

$$\Delta WS^{(10J)} = \begin{cases} -0.5, & \text{wenn } \Delta WS^{(10J)} < -0.5 \\ 0.8, & \text{wenn } \Delta WS^{(10J)} > 0.8 \\ \Delta WS^{(10J)}, & \text{sonst} \end{cases} \quad (4.5)$$

4.6. Robuste Regression und Residuenanalyse

Die Parameter des linearisierten Modells wurden mittels robuster linearer Regression¹⁴ geschätzt. Dazu wurde die Funktion `rlm` aus dem R-Package¹⁵ MASS mit der M-Schätzmethode verwendet. Die M-Methode ist zwar robust, aber nur wenn die Daten keine Hebelpunkte enthalten. In unserem Fall war dies jedoch kein Problem, da die Daten keine gefährlichen Hebelpunkte enthielten und damit die M-Methode vollauf genügte. Die MM-Methoden, welche auch mit Hebelpunkten robuste Ergebnisse liefern, sind leider in R 2.10.1 noch so implementiert, dass sie in Modellen mit Faktorvariablen häufig scheitern.

Die Residuenanalyse des robusten Modells ist sehr zentral. Mit dieser wird überprüft, ob alle implizit getroffenen Modellvoraussetzungen erfüllt sind. Wenn nicht, muss ein weiterer Iterationsschritt angehängt werden. Als Residuen r_i werden die (zufälligen) Abweichungen der gemessenen Daten von den mit dem Modell geschätzten Daten bezeichnet. Diese sind eine Schätzung für den theoretischen Fehler des Modells, z.B. ε_i in 4.2.

Die Fehler sollten gemäss den Modellvoraussetzungen (i) normalverteilt mit Erwartungswert¹⁶ 0 und konstanter Varianz σ^2 , also $r_i \sim \mathcal{N}(0, \sigma^2)$, und (ii) unkorreliert¹⁷ $Cov(r_i, r_j) = 0$ für $i \neq j$ sein. Diese Voraussetzungen und weitere Hinweise auf eine schlechte Modellanpassung werden in der Residuenanalyse mit Hilfe grafischer Methoden geprüft.

Die Residuen des robustifizierten Modells 4.2 mit den ursprünglichen 40 Standorten erfüllten die Anforderungen in den Aspekten nicht optimal (vgl. Abb. 4.2), welche hier und im Kapitel 4.7 besprochen werden. Bei tiefen NH_3 -Konzentrationen wurden zu viele negative Residuen gefunden. Zudem waren im tiefen Wertebereich auch die Varianzen generell höher als bei höheren Konzentrationen. Letzteres ist vor allem auf Rundungseffekte zurückzuführen: Die tiefen Immissionswerte wurden leider nur mit zwei signifikanten Stellen angegeben¹⁸.

Zur vertieften Analyse der Abweichungen der Residuen von der Modellvoraussetzung (i) wurden die Residuen in Boxplots aufgetragen. Einmal wurden die Residuen nach Standort gruppiert, einmal nach dem Monat des Jahres, in welchem die Beobachtung statt fand (Abb. 4.3). Daraus war ersichtlich, dass die auffälligen Residuen zu wenigen Standorten gehörten und dass die positiven Ausreisser eher im Sommer und die negativen eher im Winter auftraten. Diese liessen sich in zwei Gruppen unterteilen:

- BA, LC.01, OS.01, RO, ZIGE, FRUE sind Standorte die in einer Höhe um oder über 800 m.ü.M liegen.
- HUD.3, MU.01, SCH sind Standorte in aussergewöhnlicher Umgebung und sind zudem im Vergleich zum Umland ebenfalls leicht erhöht, auch wenn sie deutlich unterhalb von 800 m liegen. Gemeinsam ist ihnen, dass sie an einem Waldrand oder mindestens in einer Waldlichtung liegen und sich in der näheren Umgebung kein normaler Bauernbetrieb befindet. HUD.3 befindet sich in einem Naturschutzgebiet, MU.01 neben einer Baumschule und SCH am Hang oberhalb einer Talebene.

Dieses aussergewöhnliche Verhalten lässt sich mindestens für die höher gelegenen Standorte plausibel erklären. Diese Standorte liegen im Winter häufiger oberhalb der Inversionsgrenzen

¹⁴Für eine Motivation vgl. vorhergehendes Kap. 4.5.

¹⁵R ist die Open-Source-Software für statistische Analysen, welche für diese Arbeit verwendet wurde:
<http://www.r-project.org/>.

¹⁶Der Erwartungswert ist der erwartete Wert, wenn über unendlich viele Beobachtungen gemittelt würde.

¹⁷Die Grösse des Fehlers der einen Beobachtung darf nicht mit der Grösse des Fehlers einer anderen Beobachtung zusammenhängen.

¹⁸Aus statistischen Gründen sollten Messungen generell mit mindestens drei signifikanten Stellen abgespeichert werden, auch wenn dadurch die Genauigkeit der Messung übertroffen wird. Stärker gerundet werden sollte erst in der Zusammenfassung der Resultate.

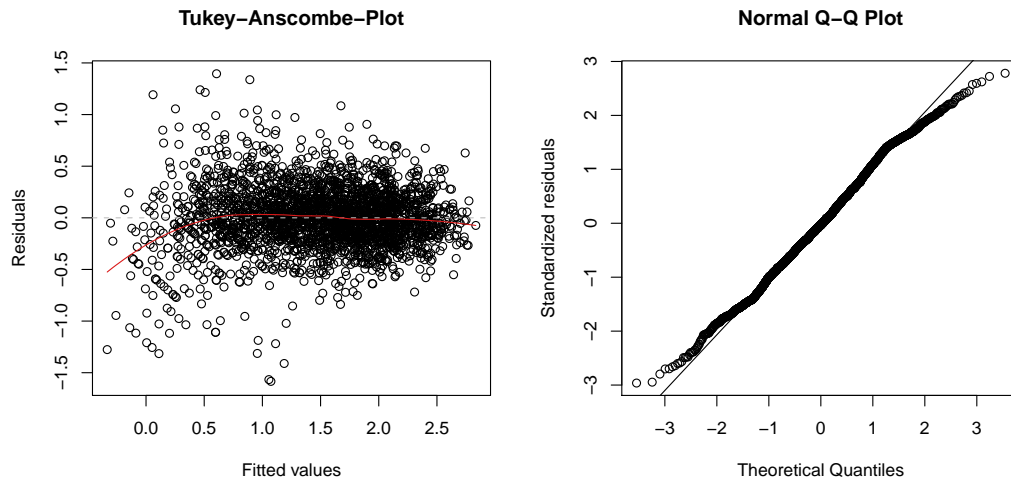


Abbildung 4.2: Die Residuenanalyse basierend auf dem Modell mit linearisierten Daten zwischen 2000 und 2009 und allen 40 Standorten zeigt, dass unser Modell 4.2 die Voraussetzungen der linearen Regressionsrechnung noch nicht ausreichend erfüllt: Die Residuen streuen bei tiefen Werten mehr (Tukey-Anscombe-Plot, linke Grafik). Die Varianz der Residuen ist also nicht konstant. Dies liegt daran, dass die NH_3 -Messwerte in diesem Bereich an die Nachweisgrenze des Messverfahrens geraten und zudem die Daten nur auf zwei signifikante Stellen abgespeichert wurden (Rundungseffekt). Ausserdem zeigt der rot eingezeichnete Glätter durch die Residuen an, dass der Erwartungswert bei tiefen Erwartungswerten nicht wie im Idealfall 0 sondern negativ ist. Gemäss dem QQ-Plot (rechte Grafik), sind die Residuen nicht perfekt normalverteilt. Die Verteilung der Residuen hat leicht kürzere Schwänze als jene der Normalverteilung (sichtbar durch die Abweichungen von der Gerade).

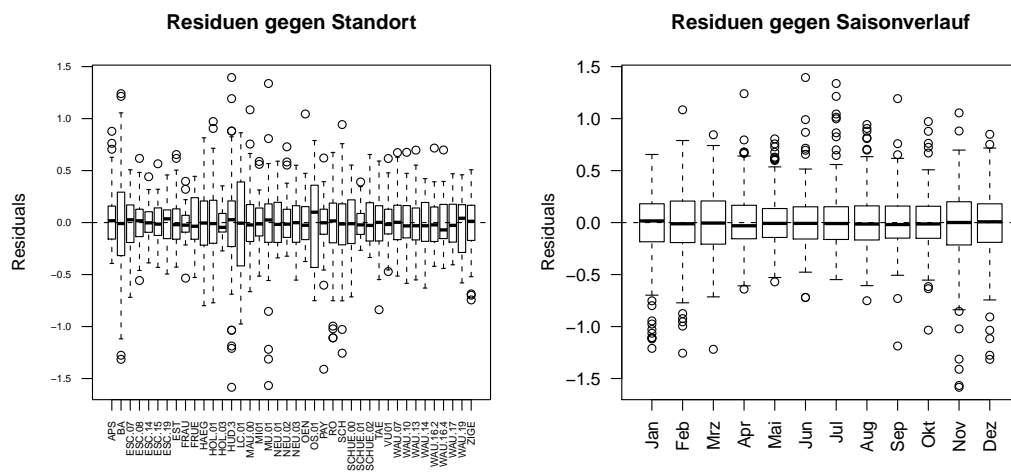


Abbildung 4.3: Boxplots der Residuen basierend auf dem Modell mit linearisierten Daten zwischen 2000 und 2009 und allen 40 Standorten: Links wurden die Residuen pro Standort und rechts pro Monat des Jahres zusammengefasst, in welchem die dazugehörigen Messungen gemacht wurden. Punkte, welche ausserhalb der Whisker (= gestrichelte Linien mit horizontalem Abschluss) liegen, gelten als Ausreisser. Die Whisker und die Box sind Masse für die Verteilung der Residuen von diesem Standort. Gemäss der linken Grafik lässt sich das Problem der unregelmässigen Streuung in 4.2 mit wenigen Standorten erklären. Diese haben gemäss der rechten Grafik im Winter tendenziell tiefere und im Sommer tendenziell höhere Residuen, im Vergleich zum Rest der Standorte.

und weisen zu diesen Zeitpunkten tiefere Luftschadstoffwerte auf als die Standorte unterhalb der Inversionsgrenze. Zudem sind die Winter in der Höhe länger, so dass dort im Winterhalbjahr die Wiesen generell weniger lange gedüngt werden können, was ebenfalls zu tendenziell tieferen Schadstoffwerten führt. Da diese Eigenheiten der höher gelegenen Standorte im Modell nicht explizit berücksichtigt wurden - es gibt keine Interaktionen zwischen Saisonverlauf und Standort - treten die oben beschriebenen Strukturen in den Residuen auf. Die hier gemachten Überlegungen treffen mindestens teilweise auch auf die drei tiefer gelegenen aussergewöhnlichen Standorte zu, da die Inversionsgrenzen je nach Geländekammer auf unterschiedlicher Höhe liegen.

Um das Modell möglichst einfach zu halten, d.h. keine Interaktionen berücksichtigen zu müssen, wurden diese Standorte aus dem endgültigen Modell ausgeschlossen. Zudem handelt es sich bei den ausgeschlossenen Standorten eher um Spezialfälle, welche sich schlecht für generelle Aussagen im Schweizer Mittelland eignen und somit für das zu untersuchende Problem nicht repräsentativ sind. Dieses Vorgehen entledigte uns gleichzeitig auch von den sehr tief liegenden Messungen, welche nahe des unteren Einsatzbereichs der Passivsammler liegen und welche (bei der Angabe von nur zwei signifikanten Stellen für tiefe Konzentrationen) zusätzlich Rundungseffekten unterliegen.

4.7. Korrelierte Residuen

Neben den systematischen Abweichungen der Residuen im tiefen Konzentrationsbereich wurden bei der Analyse der verbleibenden 31 Standorte auch substantielle (vor allem positive) Korrelationen gefunden. Dies hat zur Folge, dass die Fehler des statistischen Modells falsch geschätzt werden, wenn diese Korrelationen nicht adäquat berücksichtigt werden. Substantielle Korrelationen sind auch häufig ein Zeichen dafür, dass das beschreibende Modell nicht alle relevanten erklärenden Variablen enthält oder der funktionale Zusammenhang nicht richtig erfasst wurde. Natürlich gibt es aber auch wenige Beispiele, in welchen die Korrelationen inhärent mit dem Problem verbunden sind, z.B. wenn quantenphysikalische Experimente ausgewertet werden.

Lufthygienische Überlegungen legen nahe, dass in diesen Auswertungen substantielle Korrelationen auftreten. Die relevanten physikalischen (nichtlinearen) Prozesse laufen auf der Minuten- und Stundenskala ab. Es standen aber nur Monatsmittel¹⁹ der Konzentrationen zur Verfügung. Damit ist von vorne herein klar, dass mit räumlichen (und evtl. zeitlichen) Korrelationen zu rechnen ist.

4.7.1. Zeitlich korrelierte Residuen

Mit 17 räumlich nur wenig korrelierten Standorten (vgl. Kapitel 4.7.2) wurde der zeitliche Verlauf der Residuen über die 10 Messjahre untersucht (Abb. 4.4). Die Variabilität der Residuen vor 2003 scheint höher zu sein. Ebenso sind dort deutliche Abweichungen von der Nulllinie zu beobachten. Die genauen Ursachen dafür sind unklar. Zum einen wurde in der Zeit vor 2004 mit einem anderen Passivsammlertyp gemessen, welcher möglicherweise anders auf Witterungsphänomene reagierte²⁰. Zum andern wurde vor 2003 erst an sehr wenigen Standorten gemessen. Ferner wurde im Laufe der ersten Jahre die Platzierung der Passivsammler besser standardisiert. Zusätzlich zu diesen messtechnischen Veränderungen gab es in diesen Jahren ausserordentliche meteorologische Ereignisse:

- überdurchschnittlich viele und heftige Niederschläge im Jahr 2002 [5],

¹⁹14-Tagesmittel lagen zwar auch vor, wurden aber wegen des erhöhten Aufwands nicht untersucht, da die zusätzliche zeitliche Auflösung das Modell vermutlich nur wenig verbessern würde.

²⁰Die Messdaten vor 2004 wurden durch das BAFU via Parallelmessungen mit den nachfolgenden Messungen homogenisiert, wobei jedoch allfällige Witterungsphänomene nicht berücksichtigt wurden.

- ein viel zu warmes Jahresende 2002 und
- ein extrem heisser Sommer 2003, welcher bereits Mitte Mai begann [6].

Weil die Variabilität der Daten vor 2003 nicht eindeutig erklärt werden konnte, wurden diese für die Bestimmung des besten Modells weggelassen. Wenn die Strukturen in den Residuen mit der Geschichte der Messtechnik zu tun hätten, würden ähnliche Abweichungen in zukünftigen Daten nicht mehr auftreten. Da dies jedoch nicht absolut sicher ist, muss bis auf weiteres auch in zukünftigen Daten mit analogen Strukturen in den Residuen gerechnet werden.

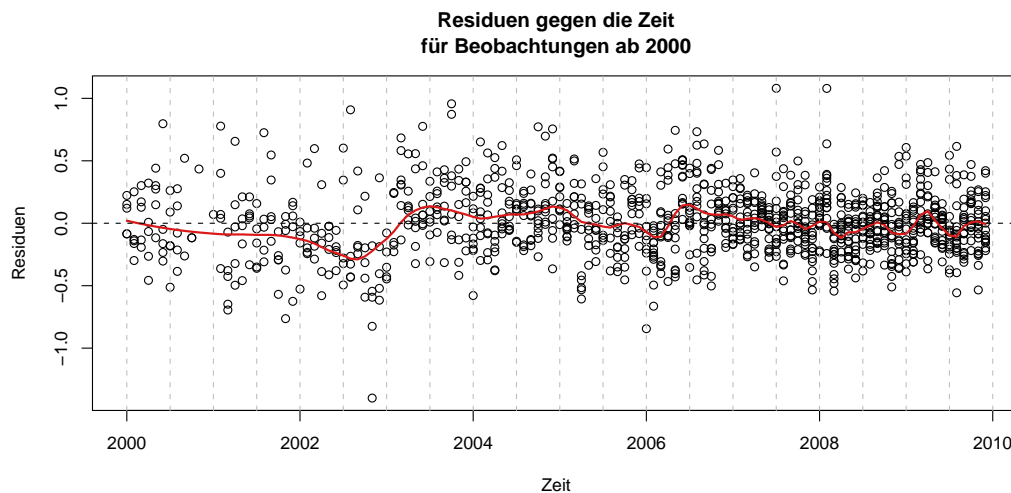


Abbildung 4.4: Die Grafik zeigt die Residuen gegen die Zeit des linearisierten Modells mit 17 Standorten für Beobachtungen ab 2000. Die rote Linie ist eine Glättung, die den allgemeinen Verlauf der Residuen angibt. Vor 2003 ist eine deutlich erhöhte Variabilität und systematische Abweichungen der Residuen von der Nulllinie erkennbar.

4.7.2. Räumlich korrelierte Residuen

Nach Einschränkung auf die Messjahre 2003–2009, wurde die räumliche Korrelation der Standorte untersucht. Wie erwartet korrelierten Standorte untereinander um so mehr, je näher beieinander diese lagen (vgl. Abb. 4.5): Die im besten Modell enthaltenen Grössen können den Einfluss der herrschenden Witterung und der landwirtschaftlichen Emissionen recht gut abbilden. Es sind aber trotzdem Natur gegebene Unzulänglichkeiten (z.B. zu lange Mittelungszeit, vgl. oben) im Modell vorhanden, welche zu Abweichungen zwischen den realen Messungen und den Schätzungen des Modells führen. Da die Witterung in einem kleinen Gebiet homogener ist als in einem grossen, werden auch die Abweichungen auf Grund der nicht perfekten Abbildung des Witterungseinflusses im Modell in einem kleinen Gebiet eher in die gleiche Richtung gehen als in einem grossen. Dieser Effekt wird durch die Wirkung der (lokalen) Emissionen verstärkt, da sich diese im Falle von NH_3 ebenfalls nur kleinräumig auswirken. Zudem korrelieren in einem Gebiet Aktivitäten einzelner Landwirte zeitlich stark, da ihr Verhalten von der Witterung abhängig sind (z.B. Ernteaktivitäten) und die offiziellen Empfehlungen den Landwirten nahe legen, Jauche nur bei bestimmten Witterungskonstellationen auszubringen. Weiter sind sich auf Grund der Topographie, der Witterung und der politischen Einheit (z.B. Kanton) auch die einzelnen Gutsbetriebe und damit die landwirtschaftlichen Emissionen in einem eng begrenzten Gebiet in der Regel ähnlicher.

Die daraus resultierende Korrelationsstruktur ist ziemlich komplex, so dass diese trotz der relativ hohen Anzahl von 40 Standorten ungenügend abgebildet werden konnte. Weiter war zu beachten, dass die Resultate, welche aus der Analyse gewonnen werden, möglichst allgemeine Aussagekraft erhalten sollten. Deshalb wurde die Modellierung der Korrelationen verworfen und die Strategie der systematischen Reduktion nahe beieinander liegender Standorte gewählt. Aus dem gleichen Grund wurde auch die Mittelung der korrelierenden Standorte verworfen.

Es wurden deshalb systematisch diejenigen Standorte entfernt, welche hohe Korrelationen mit anderen Standorten in der Nähe aufwiesen. Bei der Auswahl wurde darauf geachtet, dass in jeder betroffenen Region der Standort mit der längsten Messung beibehalten werden konnte. In Wauwil wurde WAU.13, in Schüpfheim SCHUE.00 und in Eschenbach ESC.08 im Datensatz belassen. Die räumlichen Korrelationen zwischen den übrigbleibenden 17 Standorten sind gemäss Abb. 4.6 nun deutlich geringer und weisen keine prominenten Strukturen mehr auf²¹.

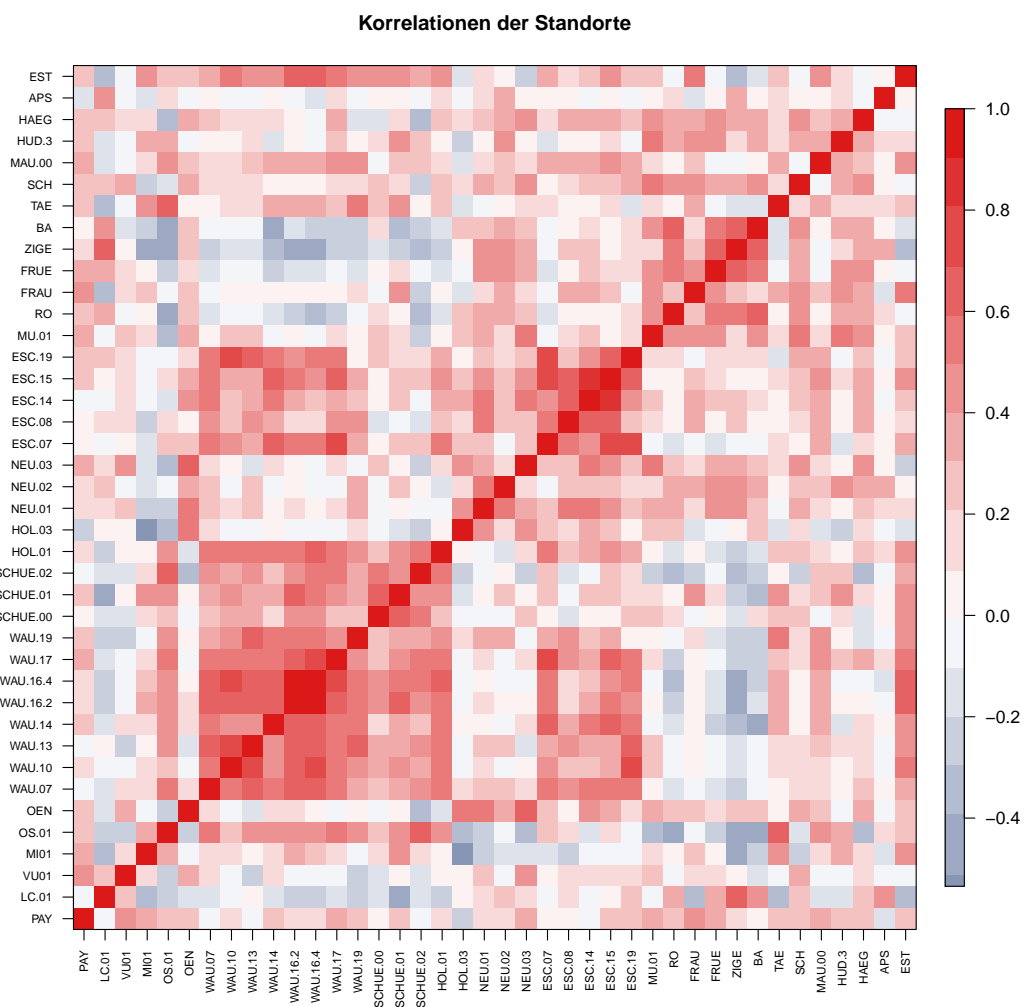


Abbildung 4.5: Der Matrixplot stellt die Korrelationen zwischen den Residuen von je zwei Standorten dar. Die Residuen stammen vom linearisierten Modell ((4.6)) mit den Daten aller 40 Standorte zwischen 2003 und 2009. Die Standorte sind geografisch von Westen nach Osten sortiert. Hohe positive (rot) und negative (blau) Korrelationen sind unerwünscht, da sie die Genauigkeit der Parameterschätzung beeinträchtigen. Wir finden hohe räumliche Korrelationen an den Standorten in Wauwil (WAU), Schüpfheim (SCHUE) und Eschenbach (ESC). Die Standorte BA, ZIGE und FRAU korrelieren ebenfalls. Sie gehören aber zu den aussergewöhnlichen Standorten und sind im besten Modell ohnehin nicht enthalten (vgl. Kap. 4.6).

²¹Vgl. Anhang B für eine Auflistung der Standorte in den getesteten Modellen.

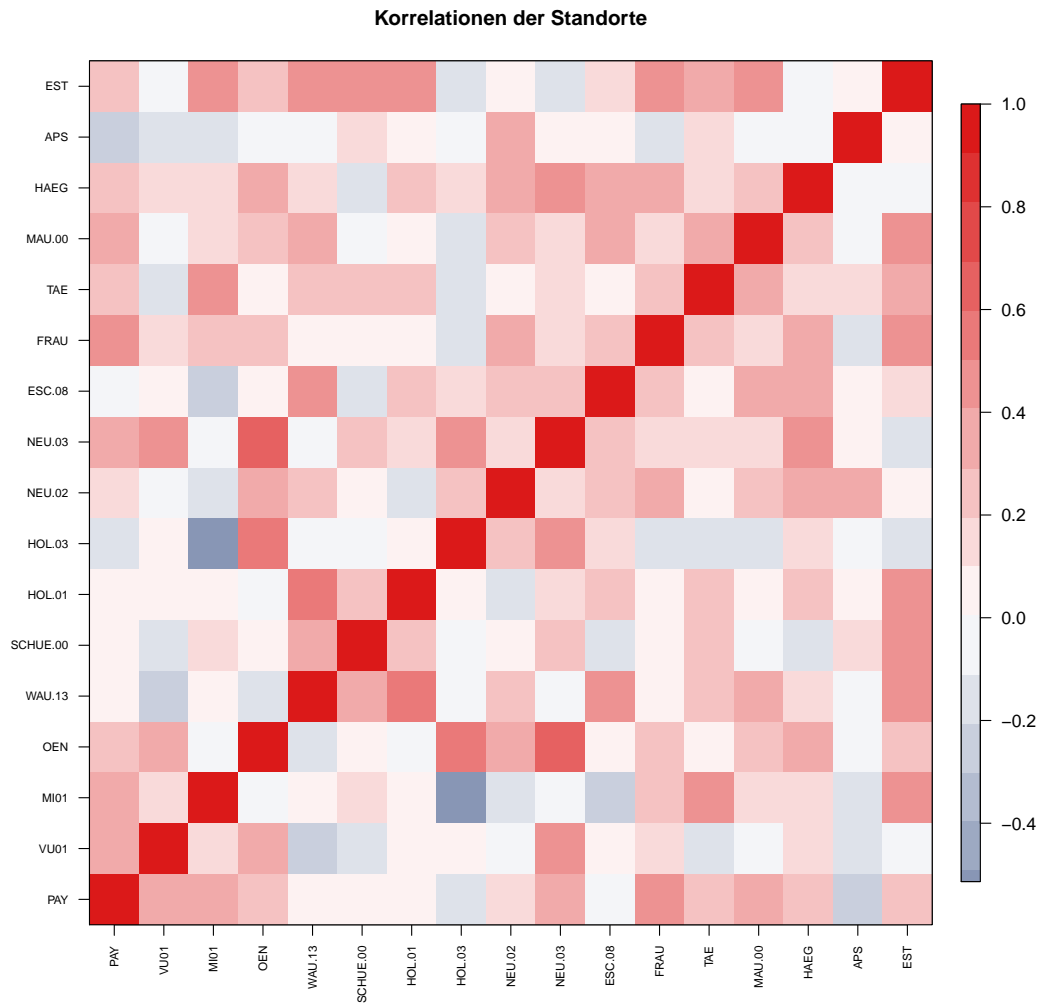


Abbildung 4.6: Im Matrixplot über die 17 Standorte des linearisierten Modells sind die Korrelationen tiefer als in Abb. 4.5 mit allen Standorten der Priorität 1. Korrelationen können in seltenen Fällen auch rein zufällig entstehen, wie z.B. zwischen NEU.03 und OEN.

4.8. Bestes Modell

Mit den ausgewählten 17 Standorten (vgl. Anhang A) konnte ein gesamtschweizerisches Modell ohne zeitliche und wenig räumlicher Korrelation gebaut werden, welches die Daten von 2003 bis 2009 gut beschreibt.

$$\log(NH_3_t^{(s)}) = \beta_0 + \text{Standort}^{(s)} + \text{Monat}_t + g_{Trend}(t) + g_T(\Delta T^{(10J)}) + g_{WS}(\Delta WS^{(10J)}) + g_{ATT}(\Delta ATT^{(10J)}) + \varepsilon_t \quad (4.6)$$

mit

$$\begin{aligned} \Delta T^{(10J)} &= \begin{cases} -3, & \text{wenn } \Delta T^{(10J)} < -3 \\ 1.5, & \text{wenn } \Delta T^{(10J)} > 1.5 \\ \Delta T^{(10J)}, & \text{sonst} \end{cases} \\ \Delta ATT^{(10J)} &= \begin{cases} 3.5, & \text{wenn } \Delta ATT^{(10J)} > 3.5 \\ \Delta ATT^{(10J)}, & \text{sonst} \end{cases} \\ \Delta WS^{(10J)} &= \begin{cases} -0.5, & \text{wenn } \Delta WS^{(10J)} < -0.5 \\ 0.8, & \text{wenn } \Delta WS^{(10J)} > 0.8 \\ \Delta WS^{(10J)}, & \text{sonst} \end{cases} \end{aligned}$$

Der P-Wert des Trends war mit 0.57 sehr viel grösser als 0.05. Es konnte somit kein signifikanter Trend festgestellt werden²². Die stark signifikanten Koeffizienten der Meteorvariablen belegen, dass im Vergleich zum 10-jährigen Monatsmittel

- überdurchschnittliche Temperaturen,
- überdurchschnittlich viele Trockentage und
- unterdurchschnittliche Windgeschwindigkeit

die NH_3 -Konzentration erhöhen, vorausgesetzt die übrigen Variablen bleiben konstant. Der Standardfehler der Residuen beträgt 0.21.

²²Im besten Modell, welches die Daten ausreichend beschreibt, ist somit die Variable t bzw. Zeit nicht enthalten. Da im Folgenden aber Daten mit signifikanten Trends simuliert werden sollen, wurde hier diese Variable im Modell belassen.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.469911	8.153649	0.793	0.428
...				
time	-0.002314	0.004064	-0.569	0.569
T	0.103615	0.005727	18.091	< 2e-16 ***
ATT	0.017457	0.001699	10.274	< 2e-16 ***
WS	-0.293518	0.030115	-9.747	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2117 on 985 degrees of freedom

Multiple R-squared: 0.8177, Adjusted R-squared: 0.812

F-statistic: 142.5 on 31 and 985 DF, p-value: < 2.2e-16

4.8.1. Resultate für Luzerner Standorte

Weiter wurde das Modell 4.6 auf alle Luzerner Standorte angewendet. Da diese Standorte untereinander korrelieren ist eine saubere Interpretation der P-Werte nicht möglich! Da der Trend das falsche Vorzeichen hat, kann aber mit Sicherheit gesagt werden, dass auch in Luzern noch keine Abnahme der NH_3 -Immissionen über die letzten 10 Jahre statt gefunden hat. Analysen von kompakten Kantonen im Schweizerland, welche mindestens die Ausdehnung des Kanton Luzerns haben, können auf die gleiche Art und Weise (und mit den selben Vorbehalten) analysiert werden.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.564e+01	8.047e+00	-1.944	0.05219 .
...				
time	8.605e-03	4.010e-03	2.146	0.03209 *
T	1.234e-01	5.269e-03	23.427	< 2e-16 ***
ATT	2.242e-02	1.533e-03	14.623	< 2e-16 ***
WS	-5.056e-01	3.566e-02	-14.179	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2037 on 1175 degrees of freedom

Multiple R-squared: 0.8026, Adjusted R-squared: 0.7965

F-statistic: 132.7 on 36 and 1175 DF, p-value: < 2.2e-16

5. Simulation

Das in Kapitel 4 gebaute Modell beschreibt die NH_3 -Monatsmittel in der Schweiz als Funktion des Standortes, der Saison, der Meteorvariablen und der Zeitvariablen zwischen 2003 und 2009. Unter der Voraussetzung, dass der Zusammenhang zwischen NH_3 -Monatsmittel und den erklärenden Variablen auch in der Zukunft gültig ist, kann das an die vergangenen Daten angepasste Modell verwendet werden, um verschiedene zukünftige Szenarien mit unterschiedlichen Trends und von unterschiedlicher Dauer zu simulieren. Die Ergebnisse geben dann direkt Antwort auf die in Kap. 2.2 beschriebenen Fragen.

5.1. Szenarien

Die folgenden Szenarien wurden mit einen künstlichen, negativen Trend von

- A. -2% pro Jahr während 5 Jahren,
- B. -2% pro Jahr während 10 Jahren,
- C. -1% pro Jahr während 5 Jahren und
- D. -1% pro Jahr während 10 Jahren

für

- 1. 17 Standorte (ohne aussergewöhnliche und räumlich korrelierte)
- 2. 12 der 17 Standorte
- 3. 7 der 17 Standorte

simuliert.

5.2. Übersicht über Simulationsverfahren

Ein Simulationsvorgang bestand jeweils aus den folgenden Schritten:

- (i) Berechnung der Erwartungswerte über die nächsten 5 bis 10 Jahre mit den Koeffizienten des besten Modells gemäss Kapitel 4.8. Ausgenommen davon war der Trend, welcher gemäss den Szenarien in Kapitel 5.1 variiert wurde.
- (ii) Addition eines zufälligen Fehlers.
- (iii) Die so erhaltenen Raum-Zeit-Reihen wurden wie zuvor die realen Messdaten ausgewertet, d.h. es wurden alle Koeffizienten des statistischen Modells geschätzt. Für die Trend-schätzung interessierte dann aber nur der Koeffizient des Trends und der dazu gehörende P-Wert.

Die Schritte wurden für jedes Szenario 500-Mal wiederholt, um die Verteilung des geschätzten Trends und deren P-Werte aus Schritt (iii) zu bestimmen. Man beachte, dass die simulierten Raum-Zeit-Reihen keine Prognose im Sinne einer Wetterprognose darstellen sondern einzig die Fehlerstruktur der realen Daten richtig wieder geben. Diese Fehlerstruktur alleine bestimmt nämlich, wie lange für die sichere Detektion eines Trends gemessen werden muss.

5.3. Details zur Berechnung der Simulation

Für die Prognose wurden die Daten zuerst vom geschätzten Trend (auch wenn er nicht signifikant war), bereinigt. Gemäss dem Satz von Taylor gilt für kleine Δx

$$\log(x \pm \Delta x) \approx \log(x) \pm \frac{\Delta x}{x} \quad (5.1)$$

Wenn Δx die absolute Zu- bzw. Abnahme der NH_3 -Konzentration über einen bestimmten Zeitraum ist, so bezeichnet $\frac{\Delta x}{x}$ die prozentuale Zu- bzw. Abnahme in dieser Zeit. Damit gilt für die trendkorrigierte logarithmierte Konzentration

$$\log(\text{NH}_3^*) \approx \log(\text{NH}_3) \pm \text{Trend} \cdot t \quad (5.2)$$

Als Grundlage für die Simulationen wurde das beste Modell für die trendbereinigten Daten der 17 Standorte von 2003 bis 2009 verwendet (im folgenden Originalmodell genannt). Für die Prognose eines neuen Datenpunktes mit bestimmtem Standort s und Zeitpunkt t wurde dabei wie folgt vorgegangen: Standort- und Saisoneffekt sowie die Koeffizienten für die Meteorvariablen wurden aus dem Originalmodell übernommen. Der Koeffizient für den Trend wurde gemäss Szenario eingesetzt. Die Werte für die Meteorvariablen und für den zufälligen Fehler wurden nach dem Zufallsprinzip aus dem Originaldatensatz bzw. der Originalauswertung entnommen, indem dort ein Standort s^* und ein Zeitpunkt t^* zufällig ausgewählt und die dazugehörigen Meteorvariablen und das Residuum zur Berechnung des neuen Datenpunktes mit Standort s und Zeitpunkt t verwendet wurden. Damit wird jede Meteorvariable und jedes Residuum von allen Standorten und Zeitpunkten des trendbereinigten Datensatzes mit gleicher Wahrscheinlichkeit gezogen, unabhängig von der Anzahl Standorte und der Länge des simulierten Datensatzes. Die Residuen der Simulationen sind folglich unkorreliert und vom Standort unabhängig und sind deshalb nur geeignet, den Datensatz mit den 17 (fast) unkorrelierten Standorten oder einen daraus abgeleiteten Subset korrekt zu simulieren. Ferner werden die simulierten Verläufe zeitlich weniger korrelieren als die Originaldaten, da zeitliche Korrelationen der Meteorvariablen bei diesem Verfahren nicht berücksichtigt werden. Dies ist jedoch von keiner praktischen Relevanz, da die Korrelationsstruktur der Residuen durch diesen Effekt überhaupt nicht beeinflusst wird.

5.4. Hinweise zu P-Werten und zur Interpretation der Resultate

Pro Szenario resultieren 500 geschätzte Trends. Anhand ihrer P-Werte wird die Nullhypothese (der geschätzte Trend ist Null) geprüft. Ein P-Wert unterhalb von 0.05 bedeutet z.B., dass bei der gegebenen Datenlage die Nullhypothese des geschätzten Trends auf einem Signifikanzniveau von 5% verworfen werden kann. Der Trend kann dann als signifikant verschieden von Null betrachtet werden. Mit den Simulationen wird die Verteilung der P-Werte bei unkorrelierten Daten untersucht.

Die hier verwendeten Regressionen berücksichtigen keine Korrelationen in den Fehlern der Beobachtungen. Sind die Fehler der Beobachtungen korreliert (vgl. Kapitel 4.7), so werden die Fehler in den geschätzten Koeffizienten falsch geschätzt. Damit sind die von der Regression ausgegebenen P-Werte unzuverlässig und es können nicht signifikante Terme im Modell als signifikant erscheinen und umgekehrt.

5.5. Simulationsresultate

Die Verteilungen der P-Werte aus den Simulationen wurden in Abb. 5.1 und in Tab. 5.1 zusammengefasst. Bei einer Messdauer von 10 Jahren an 17 untereinander unkorrelierten Standorten kann ein negativer Trend von bereits -1% pro Jahr mit Sicherheit erkannt werden. Auch bei einer Reduktion der Anzahl Standorte auf 12 oder 7 wird ein Trend von -1% sicher detektiert.

Nach nur 5 Jahren Messzeit wird ein Trend nur noch als signifikant nachgewiesen, wenn dieser -2% oder negativer ist und wenn mindestens 12 unkorrelierte Standorte zur Verfügung stehen. Unabhängig davon läuft man bei nur 5 Jahren Beobachtungszeit Gefahr, dass man eine zufällige Abweichung zu Beginn oder am Ende der Beobachtungszeit fälschlicherweise als signifikanten Trend interpretiert. Letzteres kann nicht ausgeschlossen werden, da die Interpretation der Abweichungen der Residuen des besten Modells vor 2003 nicht eindeutig ist (vgl. Kapitel 4.7.1)

Anzahl Standorte	Abnahme Total	Abnahme pro Jahr	Abnahme erkennbar?	Wsk. $P\text{-Wert} < 0.05$
17	-20%	-2% in 10 Jahren	ja	1
17	-10%	-1% in 10 Jahren	ja	1
17	-10%	-2% in 5 Jahren	ja	0.99
17	-5%	-1% in 5 Jahren	nein	0.56
12	-20%	-2% in 10 Jahren	ja	1
12	-10%	-1% in 10 Jahren	ja	1
12	-10%	-2% in 5 Jahren	ja	0.95
12	-5%	-1% in 5 Jahren	nein	0.45
7	-20%	-2% in 10 Jahren	ja	1
7	-10%	-1% in 10 Jahren	ja	0.96
7	-10%	-2% in 5 Jahren	teilweise	0.76
7	-5%	-1% in 5 Jahren	nein	0.28

Tabelle 5.1: Zusammenfassung der Resultate aus den Simulationen von 7, 12 und 17 unkorrelierten Standorten für 5 und 10 Messjahre bei unterschiedlich starker Abnahme von -1% bzw. -2% pro Jahr. In der letzten Spalte wurde die Wahrscheinlichkeit berechnet, dass der P-Wert eines geschätzten Trends unter der Signifikanzgrenze von 0.05 liegen wird.

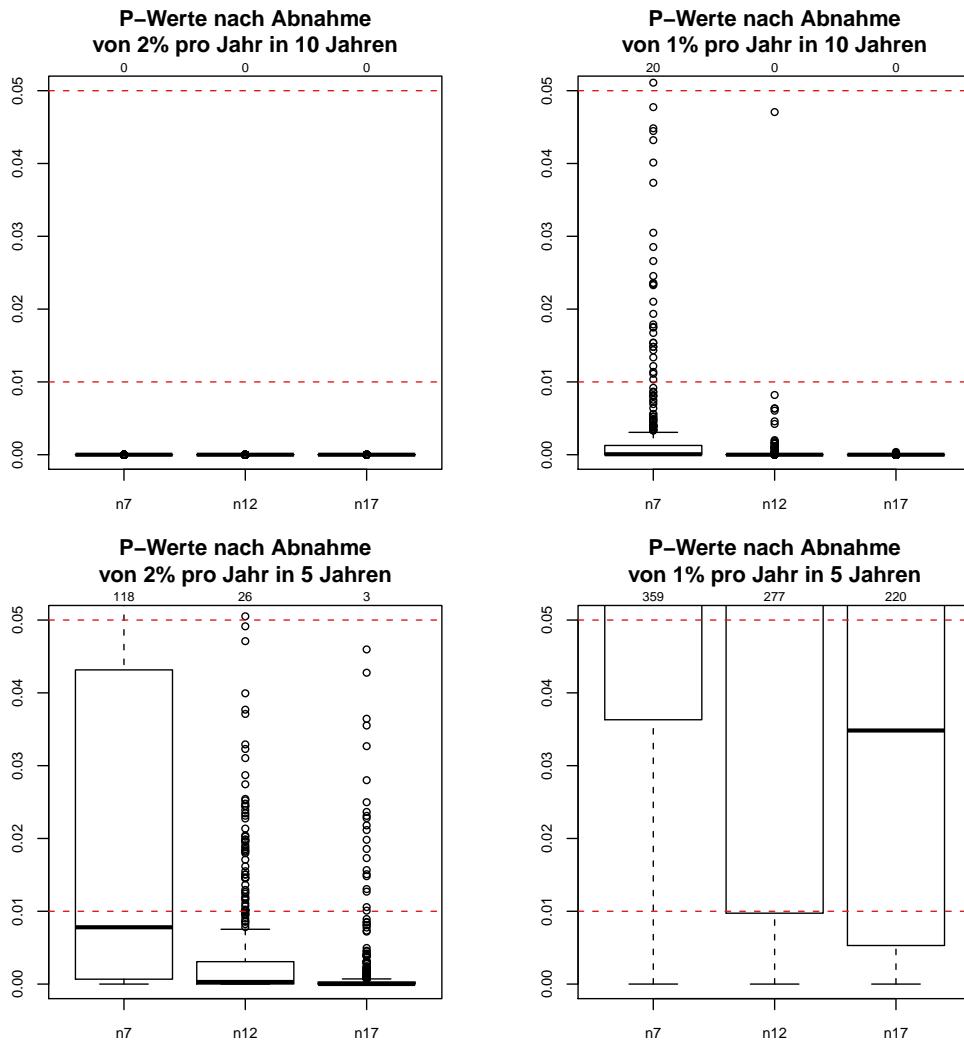


Abbildung 5.1: Boxplots der P-Werte für die Simulationen mit 7, 12 und 17 unkorrelierten Standorten. Die (rot) gestrichelten Linien markieren die in den Naturwissenschaften üblichen Verwerfungsgrenzen der Nullhypothese bei -1% und -5% . Die Nummern auf der oberen Achse jeder Grafik bezeichnen die Anzahl Werte, welche oberhalb der 5% -Grenze liegen und nicht mehr im Anzeigebereich der Grafik liegen.

6. Literatur

- [1] Stickstoffhaltige Luftschadstoffe in der Schweiz. Status-Bericht der Eidg. Kommission für Lufthygiene (EKL); Schriftenreihe Umwelt Nr. 384, S. 168; Bundesamt für Umwelt, Bern, 2005
- [2] Ammoniak-Immissionsmessungen in der Schweiz 2000 - 2009, im Auftrag des Bundesamtes für Umwelt (BAFU), der OSTLUFT und der Kantone Luzern, Freiburg, Zug, Bern, Thurgau und Graubünden; Lotti Thöni, Eva Seitler; FUB, Juli 2009
- [3] Ammoniak-Immissionsmessungen in der Schweiz - Sammel- und Messmethoden; Eva Seitler, Lotti Thöni; FUB, Juni 2009
- [4] Beschrieb der einzelnen NH₃-Passivsammlerstandorte; Auszug für 2009 aus der NH₃-Standortdatenbank der FUB, betrieben im Auftrag des BAFU; 2010
- [5] Unwetterschäden in der Schweiz im Jahre 2002; Christoph Hegg, Marielle Fraefel, Eva Frick, Franziska Schmid, Alexandre Badoux; WSL Eidg. Forschungsanstalt, Abt. Wasser, Erd- und Felsbewegungen, 2003
- [6] Arbeitsbericht: Die extreme Sommerhitze im aussergewöhnlichen Witterungsjahr 2003; Stephan Bader; MeteoSchweiz, 2004

A. Standorttabelle

NH ₃ -Messstationen				Meteo-Stationen		
Code	Standort	Kanton	Höhe	Code	Standort	Höhe
PAY	Payerne	VD	489	PAY	Payerne	490
TAE	Tänikon	TG	539	TAE	Tänikon	539
MAU.00	Mauren 0	TG	439	GUT	Güttingen	440
SCHUE.00	Schüpfheim 0	LU	735	LUZ	Luzern	454
BA	Bachtel	ZH	930	HOE	Hörnli	1132
RO	Root Michaelskreuz	LU	791	CHZ	Cham	440
HUD.3	Hudelmoos 3	TG	520	GUT	Güttingen	440
MU.01	Muri 1	AG	470	CHZ	Cham	440
SCH	Schänis	SG	630	GLA	Glarus	517
*HOL.00	Holderhus 0	LU	590	LUZ	Luzern	454
FRUE	Früebüel	ZG	980	RIG	Rigi-Seebodenalp	1031
ZIGE	Zigerhüttli	ZG	989	RIG	Rigi-Seebodenalp	1031
LC.01	Le Châtelard 1	FR	980	PLF	Plaffeien	1042
OS.01	Oberschrot 1	FR	900	PLF	Plaffeien	1042
OEN	Oensingen	SO	450	HAE	Härkingen	431
VU01	Vuisternens-en-Ogoz	FR	850	PLF	Plaffeien	1042
FRAU	Kloster Frauental	ZG	395	CHZ	Cham	440
MI01	Misery	FR	607	PAY	Payerne	490
NEU.01	Neudorf 1	LU	735	LUZ	Luzern	454
NEU.02	Neudorf 2	LU	735	LUZ	Luzern	454
NEU.03	Neudorf 3	LU	735	LUZ	Luzern	454
HOL.01	Holderhus 1	LU	590	LUZ	Luzern	454
HOL.03	Holderhus 3	LU	600	LUZ	Luzern	454
ESC.07	Eschenbach 7	LU	500	LUZ	Luzern	454
ESC.08	Eschenbach 8	LU	495	LUZ	Luzern	454
ESC.14	Eschenbach 14	LU	540	LUZ	Luzern	454
ESC.15	Eschenbach 15	LU	520	LUZ	Luzern	454
ESC.19	Eschenbach 19	LU	495	LUZ	Luzern	454
EST	Eschen	FL	441	VAD	Vaduz	457
HAEG	Häggenchwil	SG	555	GUT	Güttingen	440
APS	Ap.-Steinegg	AI	820	STG	St.Gallen	776
SCHUE.01	Schüpfheim 1	LU	705	LUZ	Luzern	454
SCHUE.02	Schüpfheim 2	LU	700	LUZ	Luzern	454
WAU.07	Wauwil 07	LU	501	LUZ	Luzern	454
WAU.10	Wauwil 10	LU	501	LUZ	Luzern	454
WAU.13	Wauwil 13	LU	501	LUZ	Luzern	454
WAU.14	Wauwil 14	LU	501	LUZ	Luzern	454
WAU.16.2	Wauwil 16.2	LU	499	LUZ	Luzern	454
WAU.16.4	Wauwil 16.4	LU	499	LUZ	Luzern	454
WAU.17	Wauwil 17	LU	501	LUZ	Luzern	454
WAU.19	Wauwil 19	LU	501	LUZ	Luzern	454

* Wurde durch HOL.01 ersetzt. Da sich HOL.00 auf unterschiedlicher Höhe am gleichen Ort befand, konnten die Daten umgerechnet werden.

B. Standorte in den Modellen

	Modell mit $n =$ Anzahl Standorte				
	$n = 40$	$n = 31$	$n = 17$	$n = 12$	$n = 7$
PAY	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
TAE	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
MAU 00	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
SCHUE 00	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
BA	<i>x</i>				
RO	<i>x</i>				
HUD 3	<i>x</i>				
MU 01	<i>x</i>				
SCH	<i>x</i>				
FRUE	<i>x</i>				
ZIGE	<i>x</i>				
LC 01	<i>x</i>				
OS 01	<i>x</i>				
OEN	<i>x</i>	<i>x</i>	<i>x</i>		
VU01	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
FRAU	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
MI01	<i>x</i>	<i>x</i>	<i>x</i>		
NEU 01	<i>x</i>	<i>x</i>			
NEU 02	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
NEU 03	<i>x</i>	<i>x</i>	<i>x</i>		
HOL 01	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
HOL 03	<i>x</i>	<i>x</i>	<i>x</i>		
ESC 07	<i>x</i>	<i>x</i>			
ESC 08	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
ESC 14	<i>x</i>	<i>x</i>			
ESC 15	<i>x</i>	<i>x</i>			
ESC 19	<i>x</i>	<i>x</i>			
EST	<i>x</i>	<i>x</i>	<i>x</i>		
HAEG	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
APS	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
SCHUE 01	<i>x</i>	<i>x</i>			
SCHUE 02	<i>x</i>	<i>x</i>			
WAU 07	<i>x</i>	<i>x</i>			
WAU 10	<i>x</i>	<i>x</i>			
WAU 13	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
WAU 14	<i>x</i>	<i>x</i>			
WAU 16.2	<i>x</i>	<i>x</i>			
WAU 16.4	<i>x</i>	<i>x</i>			
WAU 17	<i>x</i>	<i>x</i>			
WAU 19	<i>x</i>	<i>x</i>			

Die Spalte $n = 17$ bezeichnet die Standorte, welche für das beste Modell gemäss Kapitel 4.8 verwendet wurden.

C. Interaktionen zum Abfangen von gekoppelten Einflüssen

Bei einer Interaktionen hängt der Einfluss einer Variablen auf die Zielgrösse von anderen Variablen ab. Beispielsweise liegt eine Interaktion zwischen der Standorthöhe und der Temperatur vor, wenn die 10-jährige Monatsdifferenz der Temperatur oberhalb von 800 Metern einen anderen Einfluss auf die NH_3 -Konzentration hat, als unterhalb.

Unterschiedliches saisonales Verhalten der Standorte kann mit Interaktionen abgebildet werden. Dadurch erhält das Modell quasi zusätzliche erklärende Variablen, welche wie gewohnt auf ihre Relevanz überprüft werden müssen. Ein Modellgütekriterium heisst BIC von Schwarz. Gesucht wird das Modell mit dem kleinstem Wert

$$BIC = -2 \cdot \ln L + p \cdot \ln(n)$$

wobei p die Anzahl Modellparameter, n die Anzahl Beobachtungen und L die Loglikelihood der Modellanpassung bezeichnet. Das BIC bestraft durch den zweiten Term Modelle mit zu vielen (überflüssigen) Parametern und bevorzugt kleinere Modelle.

Es wurde versucht die standortspezifischen Saisonverläufe auf verschiedene Arten mit Interaktionen abzubilden, welche in den folgenden Unterkapiteln dokumentiert sind. Für die Schätzung der Modellparameter wurden jeweils die linearisierten Daten aller 40 Standorte mit Priorität 1 zwischen 2003 und 2009 verwendet. Der BIC-Wert für das Modell ohne Interaktionen beträgt -6255.091 und kann für den direkten Vergleich mit den folgenden Modellen verwendet werden.

Die nachfolgend (in Anlehnung an die R-Nomenklatur) verwendete Kurzschreibweise für Modell-Formeln lautet am Beispiel des besten Modells 4.6:

$$\log(\text{NH}_3) \sim \text{Standort} + \text{Monat} + \text{Zeit} + \Delta T^{(10J)} + \Delta WS^{(10J)} + \Delta ATT^{(10J)} \quad (\text{C.1})$$

Ein anderes Verfahren zur Variablenselektion basiert auf der sogenannten ANOVA, in welcher die (Faktor-)Variable mit dem F- oder Chi-Test auf Signifikanz getestet wird. Dieses Verfahren führt bei Modellen mit vielen Faktorvariablen und Beobachtungen zu grösseren Modellen als die BIC-Methode, so dass diese Methode hier nur zu Dokumentationszwecken eingesetzt wurde. Die Entscheidungen basierten auf der Interpretation des BIC.

Eine Analyse der Varianzen mit obigem Modell und allen 40 Standorten ergab die nachfolgende Tabelle. Die Werte in der zweiten Spalte gibt an, um wie viel die Variable die Summe der Residuenquadrate reduziert, wenn sie ins Modell aufgenommen wird. Der P-Wert in der letzten Spalte gibt Auskunft über den F-Test, bei dem die Signifikanz der Variablen unter Berücksichtigung des Verlusts von Freiheitsgraden geprüft wird. In der untersten Zeile sollte die Zahl der übrig bleibenden Residuenquadrate möglichst klein sein und die Anzahl Freiheitsgrade möglichst nahe an die Anzahl der Beobachtungen kommen.

Analysis of Variance Table

Response: NH3log

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
site	39	588.56	15.091	288.8476	<2e-16	***
mon	11	102.41	9.310	178.1849	<2e-16	***
time	1	0.00	0.003	0.0511	0.8212	
T	1	52.84	52.840	1011.3525	<2e-16	***
ATT	1	16.99	16.994	325.2616	<2e-16	***
WS	1	10.12	10.118	193.6547	<2e-16	***

Residuals 2189 114.37 0.052

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

C.1. Interaktion zwischen Standort und Monat

Mit der Interaktion zwischen Standort und Monat erhält jeder Standort einen eigenen Saisonverlauf. Die stolze Zahl von 1760 Freiheitsgrade sind trotzdem kritisch, wenn man sich vergegenwärtigt, wieviele einzelnen Koeffizienten in diesem Modell geschätzt werden müssen. Pro (Standort \times Monate)-Kombination standen damit nur noch 5 Datenpunkte zur Verfügung! Das BIC betrug nur noch -4125.955 , weshalb diese Interaktion bei der Anzahl verfügbarer Beobachtungen nicht sinnvoll war.

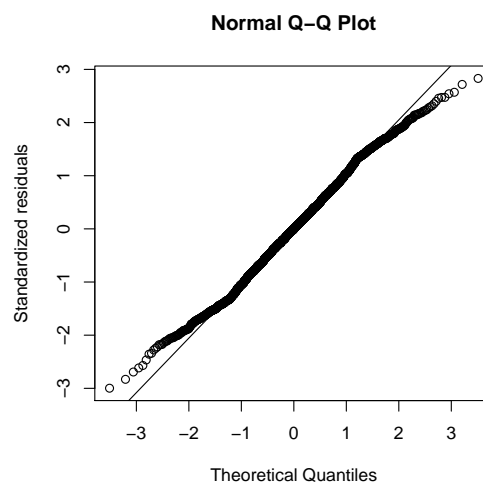
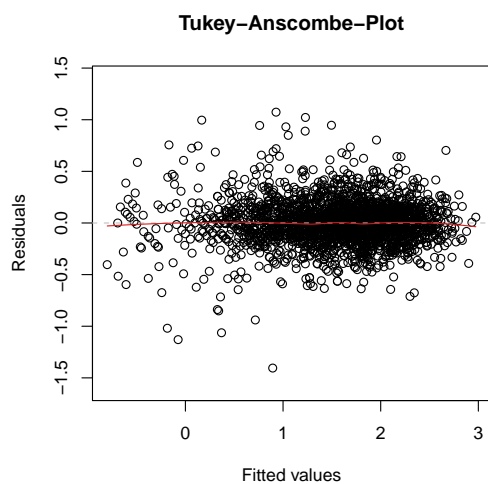
$$\log(NH_3) \sim \text{Standort} \times \text{Monat} + \text{Trend} + \Delta T^{(10J)} + \Delta WS^{(10J)} + \Delta ATT^{(10J)}$$

Analysis of Variance Table

Response: NH3log

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
site	39	607.11	15.567	405.4937	<2e-16	***
mon	11	105.54	9.595	249.9217	<2e-16	***
time	1	0.03	0.027	0.6939	0.405	
T	1	51.76	51.758	1348.2212	<2e-16	***
ATT	1	16.01	16.009	417.0027	<2e-16	***
WS	1	9.84	9.837	256.2429	<2e-16	***
site:mon	429	55.64	0.130	3.3786	<2e-16	***
Residuals	1760	67.57	0.038			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



C.2. Interaktion zwischen Standort und Saison

Etwas weniger Freiheitsgrade beanspruchte die Interaktion zwischen dem Standort und der Jahreszeit als 4-stufige Faktorvariable. Der BIC-Wert lag mit -5796.106 aber immer noch über dem Wert des Modells ohne Interaktionen, weshalb auch diese Interaktion als nicht signifikant betrachtet wurde.

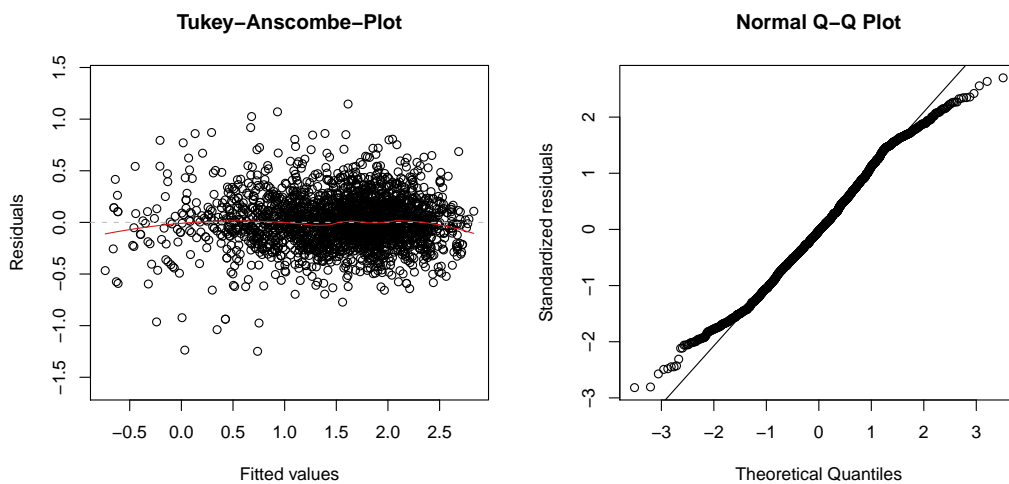
$$\log(NH_3) \sim \text{Standort} \times \text{Monat} + \text{Trend} + \Delta T^{(10J)} + \Delta WS^{(10J)} + \Delta ATT^{(10J)}$$

Analysis of Variance Table

Response: NH3log

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
site	39	605.46	15.525	334.7500	<2e-16	***
seas	3	101.68	33.894	730.8384	<2e-16	***
time	1	0.00	0.003	0.0672	0.7955	
T	1	50.41	50.412	1087.0106	<2e-16	***
ATT	1	18.94	18.943	408.4650	<2e-16	***
WS	1	9.72	9.720	209.5981	<2e-16	***
site:seas	117	33.24	0.284	6.1253	<2e-16	***
Residuals	2080	96.46	0.046			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



C.3. Interaktion zwischen Meteorvariablen und Höhenstationen

In diesem Modell wurden die zweifach Interaktionen

- (i) Temperatur \times Standorthöhe über 600 m.ü.M.
- (ii) Anzahl Trockentage \times Standorthöhe über 600 m.ü.M.
- (iii) Windgeschwindigkeit \times Standorthöhe über 600 m.ü.M.

geprüft. Nach schrittweisem Eliminieren der unbedeutenden Interaktionsterme mit dem BIC-Kriterium fielen alle Interaktionen aus dem Modell. Sie trugen also zu wenig zur Reduktion der Residuen bei, um im Modell zu verbleiben.

C.4. Interaktion zwischen Meteovariablen und Saison

In diesem Modell wurden die zweifach Interaktionen

- (i) Temperatur \times Jahreszeit
- (ii) Anzahl Trockentage \times Jahreszeit
- (iii) Windgeschwindigkeit \times Jahreszeit

geprüft. Nach schrittweisem Eliminieren der unbedeutenden Interaktionsterme mit dem BIC-Kriterium blieb (i) und (iii) im Modell enthalten. Der BIC-Wert betrug -6294.639 und war zwar tiefer als beim Modell ohne Interaktionen. Eine Analyse der Variabilitäten im Modell mit Interaktionen zeigte aber auch, dass der Anteil der Residuen (Spalte `Sum Sq` im R-Output), der durch die Interaktionen `T:Seas` und `WS:Seas` zusätzlich erklärt werden konnte, im Verhältnis zu den anderen Einflussgrößen sehr klein war. Da dieses Modell schwierig zu interpretieren ist, wurde auch auf diese Interaktionen verzichtet.

$$\log(NH_3) \sim \text{Standort} + \text{Monat} + \text{Trend} + \Delta T^{(10J)} + \left(\Delta T^{(10J)} \times \text{Jahreszeit} \right) \\ + \Delta WS^{(10J)} + \left(\Delta WS^{(10J)} \times \text{Jahreszeit} \right) + \Delta AAT^{(10J)}$$

Analysis of Variance Table

Response: NH3log

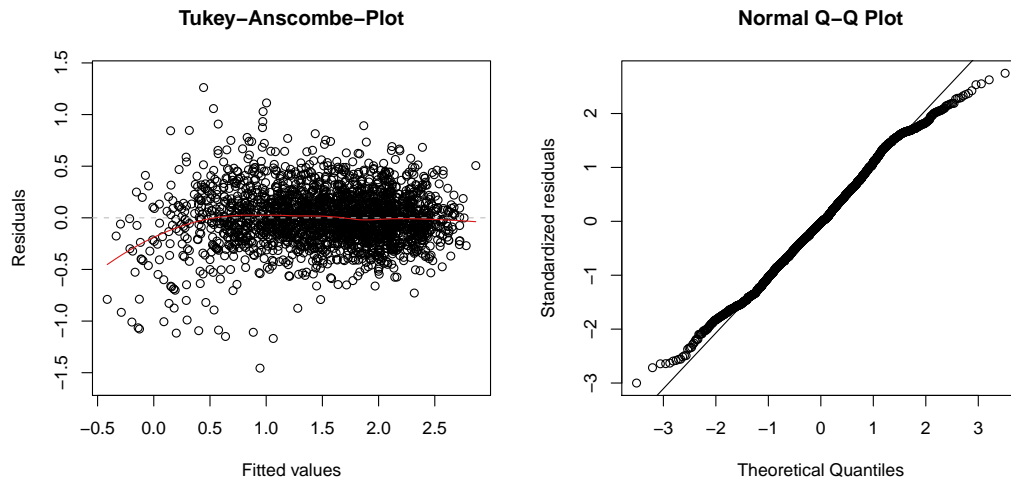
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
site	39	589.10	15.105	298.668	< 2.2e-16 ***
mon	11	102.96	9.360	185.080	< 2.2e-16 ***
T	1	52.44	52.439	1036.859	< 2.2e-16 ***
ATT	1	17.20	17.199	340.072	< 2.2e-16 ***
WS	1	10.20	10.195	201.591	< 2.2e-16 ***
T:seas	3	1.97	0.655	12.955	2.180e-08 ***
WS:seas	3	3.49	1.163	23.002	1.171e-14 ***
Residuals	2184	110.45	0.051		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

C.5. Unterscheidung aussergewöhnlicher Standorte

Eine weitere Möglichkeit war, die Unterschiede der aussergewöhnlichen Standorte in den Monaten November bis Februar zu modellieren. Dazu wurden die beiden Gruppen von ausserordentlichen Standorten je eine zusätzliche Variable notwendig:

- (i) für die sechs Standorte um oder über 800 m ü.M (*Inv*) und
- (ii) für die drei Standorte mit hoher Variabilität in den Messungen unterhalb 800 m (*Var*).



$$\begin{aligned} \log(\text{NH}_3) \sim & \text{Standort} + \text{Monat} + \text{Trend} + T^{(10J)} + \text{ATT}^{(10J)} + \text{WS}^{(10J)} + \text{Inv} + \text{Var} \\ & + (\Delta T^{(10J)} \times \text{Inv}) + (\Delta \text{ATT}^{(10J)} \times \text{Inv}) + (\Delta \text{WS}^{(10J)} \times \text{Inv}) \\ & + (\Delta T^{(10J)} \times \text{Var}) + (\Delta \text{ATT}^{(10J)} \times \text{Var}) + (\Delta \text{WS}^{(10J)} \times \text{Var}) \end{aligned}$$

Nach schrittweisem Eliminieren der unbedeutenden Terme mit dem BIC-Kriterium blieb nur die Interaktion zwischen $T^{(10J)}$ und Inv und (ohne Interaktion) Var im Modell enthalten. Der BIC-Wert für dieses Modell lag bei -6417.741 und war tiefer als das Modell ohne diese Interaktionen und zusätzlichen Variablen. Der Anteil der Residuenquadrate, der damit zusätzlich erklärt werden konnte war eher klein. Da mit dieser Lösung die Probleme der hohen Variabilität in den Standorten nur halbwegs behoben werden konnten, wurde das Weglassen der aussergewöhnlichen Standorte vorgezogen. Dies hatte den Vorteil, dass einerseits das Modell einfacher und andererseits auch die verbleibenden Standorte einheitlicher wurden. Beides trägt dazu bei, dass die Resultate übersichtlicher wurden und leichter zu interpretieren sind.

Analysis of Variance Table

Response: NH3log

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
site	39	614.38	15.753	326.121	< 2.2e-16	***
mon	11	109.43	9.948	205.949	< 2.2e-16	***
T	1	54.81	54.807	1134.597	< 2.2e-16	***
Inv	1	9.78	9.780	202.462	< 2.2e-16	***
Var	1	3.97	3.974	82.260	< 2.2e-16	***
ATT	1	17.14	17.137	354.763	< 2.2e-16	***
WS	1	9.19	9.187	190.185	< 2.2e-16	***
T:Inv	1	1.66	1.661	34.390	5.195e-09	***
Residuals	2187	105.64	0.048			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

