

Kurzzeit-Prognose von PM₁₀-Konzentrationen in der Luft unter Verwendung maschineller Lernverfahren

Von der Fakultät Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur
Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Vorgelegt von
Dipl.-Geogr. Daniel Nicklaß
aus Stuttgart

Hauptberichter: Prof. Dr. W. D. Blümel

Mitberichter: Prof. Dr. G. Baumbach

Tag der mündlichen Prüfung: 15. März 2010

Institut für Geographie der Universität Stuttgart

2010

Mein Dank gilt

Prof. Dr. rer. nat. Blümel vom Institut für Geographie und Prof. Dr.-Ing. Gün-ter Baumbach vom Institut für Verfahrenstechnik und Dampfkesselwesen (IVD) der Universität Stuttgart für die Betreuung der Dissertation.

Meinen MARQUIS Projektpartnern Dipl.-Met. Helmut Scheu-Hachtel, Dipl.-Met. Corinna Humpert-Zerulla, Markus Klein und Jürgen Kassens von der Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW), Dr. Achim Lohmeyer und Dipl.-Geoökol. Christine Sörgel vom Ingenieurbüro Lohmeyer, Prof. Dr. Leo Wanner von der Universität Pompeu Fabra, Barcelona, Dr. Bernd Bohnet von IS² für die Inspiration zu dieser Arbeit und ihre Unterstützung.

Dr. Jutta Holst und Prof. Dr. Helmut Mayer vom Meteorologischen Institut der Uni Freiburg (MIF), Dr. Ulrich Damrath, Ulrich Pflüger und Dipl.-Met. Hans-Joachim Koppert vom Deutschen Wetterdienst (DWD) für die meteorologische Unterstützung.

Meinen ehemaligen Kollegen Dr. Peter Bickel, Dr. Anna Frey, Dr. Jörg Haigis, Dipl.-Inf. Volker Klotz, Dipl.-Geogr. Renate Köble, Dipl.-Geoökol. Ulrike Kummer, Dipl.-Ing. Jochen Lambauer, Dr. Pauline Midgley, Dipl.-Math. Steffen Nitter, Dipl. Wi.-Ing. Kathrin Ohlau, Dr. Heiko Pfeiffer, Dr. Thomas Pregger, Dipl.-Ing. Philipp Preiss, Dr. Stefan Reis, Ralph Schelle und Dr. Jochen Theloke vom Institut für Energiewirtschaft und Rationelle Energieanwendung (IER) der Universität Stuttgart.

Meiner Frau Ute und meinen Töchtern Kay und Nele für ihre Liebe und Unterstützung.

"The best qualification of a prophet is to have a good memory."

George Savile, Marquis of Halifax

Inhaltsverzeichnis

1	Einleitung	1
1.1	Stand der Forschung	2
1.2	Aufgabenstellung	5
2	Grundlegende Aspekte zum Thema Feinstaub	7
2.1	Entstehung und Quellen von Feinstaub	7
2.2	Auswirkung von Feinstaub auf die menschliche Gesundheit	9
2.3	Gesetzliche Richtlinien bezüglich PM_{10}	11
3	Vorstellung der angewandten Methodik	14
3.1	Vorstellung der verwendeten Verfahren zur Prognose von PM_{10} -Konzentrationen	14
3.1.1	Die multiple lineare Regression als Referenzverfahren	16
3.1.2	Das instanzbasierte Lernen mit den k -nächsten-Nachbarn	17
3.1.3	Das Multilayer Perzeptron – ein künstliches neuronales Netz	18
3.1.4	Die Regression mittels Support-Vektor-Maschine	21
3.2	Methodik zur Optimierung der Attributauswahl und der Modellparameter	23
3.2.1	Optimierung der Attributauswahl mittels Genetischer Algorithmen	24
3.2.2	Optimierung der Modellparameter mittels Rasteroptimierung und Evolutionsstrategien	25

3.3	Methodik zur Abschätzung der Leistungsfähigkeit eines Prognosemodells	26
3.3.1	Methodik der Kreuzvalidierung zur Qualitätssicherung	27
3.3.2	Beschreibung der Kennzahlen und Bewertungsgrundlagen zur Abschätzung der Leistungsfähigkeit von Modellen	28
4	Auswahl und Vorbereitung der Eingangsdaten	33
4.1	Auswahl der zur Prognose geeigneten Attribute	34
4.1.1	Meteorologiedaten	35
4.1.2	Schadstoffkonzentrationen	43
4.1.3	Attribute zur Emissionsannäherung	44
4.2	Vorbereitung der Eingangsdaten (Preprocessing)	46
5	Erstellung und Optimierung von PM₁₀-Prognosemodellen und Abschätzung ihrer Leistung	48
5.1	Training und Leistungsabschätzung der Modelle	49
5.2	Optimierung der Attributzusammenstellung	52
5.3	Optimierung der Modellparameter	56
5.3.1	Parameteroptimierung beim instanzbasierten Lernen mit den <i>k</i> -nächsten-Nachbarn im Detail	60
5.3.2	Parameteroptimierung für ein Multilayer-Perzeptron im Detail	61
5.3.3	Parameteroptimierung für eine Support-Vektor-Maschine im Detail	63
5.4	Sensitivitätsanalyse bezüglich unsicherer meteorologischer Vorhersagen	64
6	Vorstellung der Ergebnisse	69
6.1	Ergebnisse der Prognose von PM ₁₀ -Tagesmittelwerten unter perfekten Bedingungen	69
6.2	Ergebnisse der Sensitivitätsanalyse – Abschätzung der Modellperformance unter realen Bedingungen	77
6.3	Abschätzung des Einflusses einzelner Attribute auf die Prognoseergebnisse	83

7 Kritische Diskussion und Ausblick	88
7.1 Kritische Diskussion	88
7.1.1 Qualität der PM ₁₀ -Prognose	88
7.1.2 Optimierung von Attributzusammenstellung und Modellparametern	90
7.1.3 Stabilität der Prognosemodelle bei Ausfällen von Messungen oder meteorologischen Vorhersagen	91
7.1.4 Physikalische Nachvollziehbarkeit der Modelle	91
7.1.5 Zeitaufwand bei Erstellung und Anwendung der Modelle	92
7.1.6 Reproduzierbarkeit von Modellen und Ergebnissen	93
7.1.7 Erweiterbarkeit der Modelle um neue Messdaten und weitere Attribute	94
7.1.8 Änderung der Rahmenbedingungen an einer Messstation	94
7.2 Ausblick	95
8 Zusammenfassung	97
A Vergleich der EURAD Prognosen für PM₁₀ mit Messdaten der LUBW	113
B Detaillierte Ergebnisse der Optimierung der Attributzusammenstellung	116
C Detaillierte Ergebnisse der PM₁₀-Prognose	120
D Detaillierte Ergebnisse der Sensitivitätsanalyse bezüglich unsicherer meteorologischer Vorhersagen	144

Abkürzungsverzeichnis

BC	Station Biberach
CO	Kohlenstoffmonoxid
DWD	Deutscher Wetterdienst
ES	Evolutionsstrategie
FAR	False Alarm Rate
FR-M	Station Freiburg-Mitte
GA	Genetischer Algorithmus
H_{mix}	Höhe der Mischungsschicht
IA	Index of Agreement
IBk	Instanzbasiertes Lernen mit den k -nächsten-Nachbarn
KA-NW	Station Karlsruhe-Nordwest
kNN	k -Nearest-Neighbour-Methode
LUBW	Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg
LuQx	deutscher Kurzzeit-Luftqualitätsindex
LR	multiple lineare Regression
MAE	Mean Absolut Error
MLP	Multilayer-Perzeptron
NMVOG	Non Methane Volatile Organic Compound
NH_3	Ammoniak
NN	Neuronales Netzwerk
NO_2	Stickstoffdioxid
NO_x	Stickstoffoxide
O_3	Ozon
$PM_{2.5}$	Particulate Matter $< 2,5 \mu m$
PM_{10}	Particulate Matter $< 10 \mu m$

POD	Probability of Detection
RMSE	Root Mean Squared Error
RSP	Respirable Suspended Particulate < 3,5 μm
S-BC	Station Stuttgart-Bad Cannstatt
SO ₂	Schwefeldioxid
SVM	Support-Vektor-Maschine
SVR	Support-Vektor-Regression
TSP	Total Suspended Particulate < 100 μm
VOC	Volatile Organic Compound
WHO	World Health Organization
YALE	<u>Y</u> et <u>A</u> nother <u>L</u> earning <u>E</u> nvironment

Abstract

Fine particulate matter (PM₁₀) in ambient air once again became an important issue of clean air policy from the beginning of the 1990s when it became apparent that even relatively low concentrations of PM₁₀ can cause serious health problems.

In order to inform the competent authorities and the general public about air quality and to assist them in taking measures to avoid health problems, it is necessary to establish a system which predicts the PM₁₀-concentration for the following day. For that purpose some established forecasting models such as the multiple linear regression are compared to innovative machine-learning techniques: the support-vector-machine, the multilayer-perceptron and the k -nearest-neighbour method. The respective performance of all models is compared, based on the same pool of data which contains measurements from four stations in Baden-Württemberg. As the principle contribution on human health problems caused by high air pollution levels, this study puts it's main focus on the correct prediction of high PM₁₀-concentrations.

The basic concept of machine-learning techniques is to extract information from data automatically, by computational and statistical methods. For the present study this means that these machine-learning algorithms automatically extract the relationships between the PM₁₀-concentrations and influencing factors such as meteorological parameters (windspeed, precipitation, ...) or the emission situation from historical measurement data. The information derived on relationships is integrated into a model which can now be applied to predict PM₁₀-concentrations using meteorological forecasts.

The use of meteorological forecasts for PM₁₀-modelling is associated with major uncertainties. Therefore the performances of all applied models are analysed using forecasted meteorological parameters of different quality. This sensitivity analysis shows how well the models perform under real world conditions.

The implementation of machine-learning techniques always necessitates some optimization steps. As well as the fine tuning of the model parameters, the input dataset also has to be optimized. The use of all available (more or less important) influencing factors does not result in the best model performance. In fact an optimized set of factors leads to the best results. These very complex optimization steps are realised using genetic algorithms, evolutionary strategies and raster optimization to minimise the effort and computing time.

Kapitel 1

Einleitung

In urbanen Gebieten wird die Luftqualität von einer Reihe von Luftschadstoffen beeinträchtigt. Neben gasförmigen Schadstoffen wie den Stickstoffoxiden spielen Schwebstoffe eine große Rolle. Sie waren in Verbindung mit einer Anreicherung von Schwefeldioxid beispielsweise für den London-Smog in den 50er Jahren des 20. Jahrhunderts verantwortlich, bei dem tausende von Menschen umkamen [11].

Galten die mit der Staubbelastung der Luft verbundenen Gefahren und Beeinträchtigungen durch nationale und europäische Umweltrichtlinien Ende der 80er und Anfang der 90er als weitgehend gebannt, so werden nach heutigem Kenntnisstand bereits im Vergleich zu früher relativ geringen Konzentrationen feiner Stäube negative gesundheitliche Auswirkungen zugeschrieben. Erneut in den Blick der Öffentlichkeit geriet die Feinstaubproblematik im Jahre 2005 als nach der Einführung des EU-Grenzwerts für PM_{10} (Stäube mit einem aerodynamischen Durchmesser $< 10 \mu m$) von $50 \mu g/m^3$ im Tagesmittel dieser Grenzwert in deutschen Städten häufig überschritten wurde [32, 110].

Um die zuständigen Behörden und die Bevölkerung über die Luftqualität informieren [116] und gegebenenfalls Maßnahmen zur Vermeidung von Gesundheitsschäden veranlassen zu können, ist es notwendig ein System aufzubauen, das in der Lage ist, PM_{10} -Konzentrationen für den folgenden Tag zu prognostizieren. Hierzu sollen etablierten Prognosemethoden verschiedene innovative maschinelle Lernverfahren gegenübergestellt und die jeweiligen, auf gleicher Datenbasis erzielten Prognoseleistungen miteinander verglichen werden. Da die größte Belastung für die menschliche Gesund-

heit durch Konzentrationsspitzen mit besonders hohen PM_{10} -Werten ausgehen, liegt hier das besondere Augenmerk der Prognose.

1.1 Stand der Forschung

Die Prognose von Luftschadstoffen ist schon lange Zeit ein Ziel der Forschung. Dementsprechend gibt es bereits mehrere Ansätze zur Prognose von PM_{10} -Konzentrationen. Dazu gehören zum einen die der *klassischen Ausbreitungsmodelle* wie z.B. EURAD (Europäisches Ausbreitungs- und Depositionsmodell) und THOR (THOR weather and air pollution forecast and scenario management system), die unter Verwendung von Emissionskatastern, Meteorologiedaten und -modellen für ihr Abbildungsgebiet flächendeckende Prognosen erstellen [51, 15]. Dabei können durch die großräumige Einbeziehung von Meteorologiedaten neben lokalen Emissionsquellen auch Ferntransporte mitbetrachtet werden. Beim Vergleich der verhältnismäßig grob auflösenden Rasterinformationen, die diese Systeme liefern, mit den an einer Messstation tatsächlich gemessenen Konzentrationen fällt auf, dass gerade die Konzentrationsspitzen nur sehr unbefriedigend wiedergegeben werden (siehe Abb. 1.1 bzw. Anhang A) [65]. Bei einer maximalen Rasterauflösung errechnet z.B. EURAD Prognosen, die für ein Gebiet von $5 \times 5 \text{ km}^2$ repräsentativ sind und für einzelne Punkte innerhalb dieses Rasters nur Näherungswerte sein können.

Um Schadstoffprognosen zu erstellen, die mit den Konzentrationen an Messstationen in Städten vergleichbar sind, muss also ein anderer Ansatz gewählt werden. Eine Möglichkeit besteht darin, die Zusammenhänge zwischen dem zu prognostizierenden Luftschadstoff und den Einflussparametern wie z.B. Meteorologiedaten zu nutzen. Dazu verwendet man *statistische Verfahren*, die es erlauben, eben jene Zusammenhänge aus historischen Daten zu extrahieren und für Vorhersagen einzusetzen.

Ein klassischer statistischer Ansatz ist die *Lineare Regressionsanalyse*, wie sie schon z.B. von Pfeiffer [80] für NO_2 , CO und NMVOC beschrieben wird. Ein unter anderem an der Universidade Nova de Lisboa [72] und in der Stadt Graz [48] eingesetztes System für die Prognose von PM_{10} arbeitet mit einer Linearen Multiregressionsanalyse unter Einbeziehung lokaler Meteorologiedaten. Lohmeyer et al. bzw. Sörgel haben diese Methode auf Stationen in Sachsen-Anhalt [61] und in der Testregion Baden-Württemberg [105] angewandt. Die dabei erzielten Ergebnisse sind zufriedenstellend,

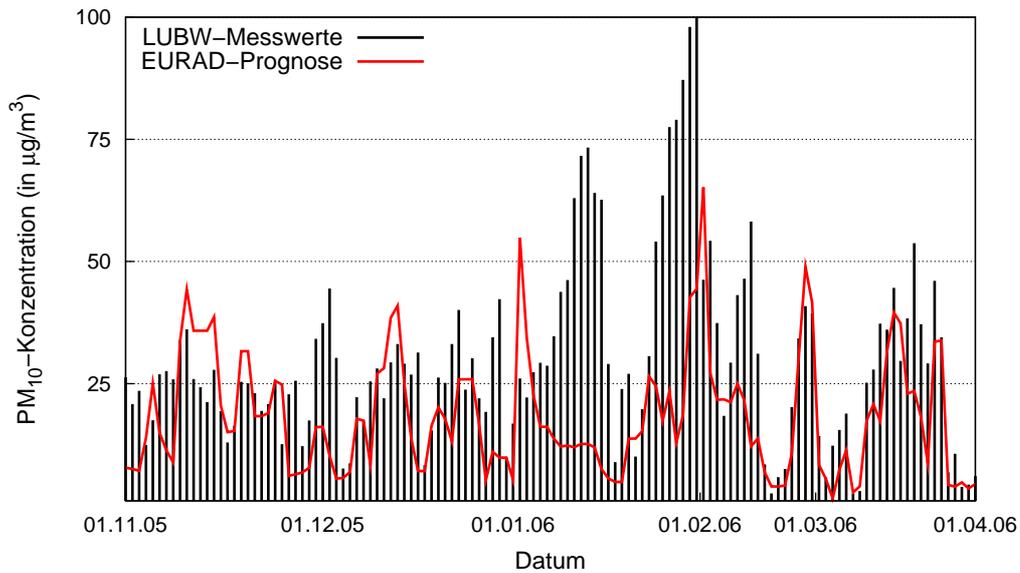


Abbildung 1.1: Vergleich der EURAD 12-h-Prognose für PM_{10} mit den Messwerten der Station Biberach

obwohl zumeist nur eine limitierte Anzahl von Einflussparametern berücksichtigt wird. Ein Manko dieser Arbeiten liegt jedoch darin, dass keine meteorologischen Vorhersagen zur Abschätzung der Prognosegenauigkeit mit einbezogen werden. Es kommen lediglich Messdaten – sozusagen eine „perfekte Prognose“ – zum Einsatz. Gerade die meteorologischen Vorhersagen sind aber mit großen Unsicherheiten behaftet. Somit ist die tatsächliche Leistung dieser Modelle für Baden-Württemberg und Sachsen-Anhalt ungeklärt. Außerdem eignet sich die Lineare Regression per se nur für die Analyse von Zusammenhängen linearer Natur. Da jedoch die Zusammenhänge zwischen Meteorologie und Feinstaubkonzentration nicht als linear angenommen werden können, scheint eine Methodik angebracht, die auch nichtlineare Beziehungen berücksichtigt.

Die Erkennung nichtlinearer Zusammenhänge leisten zahlreiche *maschinelle Lernverfahren*. Das in der Luftreinhaltung wohl gebräuchlichste dieser Verfahren ist das *künstliche neuronale Netz (NN)*. Es wurde in der Vergangenheit schon zur Prognose verschiedenster Luftschadstoffkonzentrationen eingesetzt. Beispiele finden sich in der Literatur für Ozon [27, 10, 14, 27, 74], Stickstoffoxide [17, 19] und Schwefeldioxid bzw. Kohlenmonoxid [12, 13, 75, 83, 109]. Auch für die Vorhersage von Feinstaubkonzentrationen verschiedenster Korngrößen finden NN Verwendung. Lu et al. [63, 62] beschreiben die Prognose von so genannten Respirable Suspended Particula-

tes (RSP, Korngröße $<3,5 \mu m$), Kamiski die von Total Suspended Particulates [52] (TSP, Korngröße $<100 \mu m$). Perez et al. [76, 77] sagen mittels NN die Konzentration von Feinstaub der Fraktion $<2,5 \mu m$ voraus und untersuchen dabei explizit die Leistungsfähigkeit von NN im Vergleich zur Linearen Regression. Sie kommen ebenso wie Foxall et al. [36] und Agirre et al. [2] zu dem Schluss, dass NN der Linearen Regression bei der Prognose von $PM_{2.5}$ bzw. NO_2 überlegen ist.

Auch für die Prognose von PM_{10} wurden NN bereits mehrfach eingesetzt, wobei die Ergebnisse recht unterschiedlich ausfallen. Die Arbeiten von Kolehmainen et al. [56] liefern eher unbefriedigende Ergebnisse, wobei allerdings nur ein limitierter Satz an Einflussfaktoren zum Einsatz kam. So wurde z.B. der Niederschlag als einer der wichtigsten Faktoren nicht berücksichtigt. Ähnliche Ergebnisse liefern auch Hooyberghs et al. [47] – er stellt jedoch vor allem den großen Einfluss der Mischungsschichthöhe auf die Feinstaubkonzentration heraus. Kukkonen et al. [58] sprechen NN auf dem jetzigen Entwicklungsstand sogar gänzlich die Fähigkeit zur Prognose von PM_{10} ab, wobei ihre Versuche wie bei EURAD auch auf die flächenhafte Konzentrationsverteilung von Feinstaub hin zielen. Eine positive Beurteilung von NN bei der Vorhersage von PM_{10} findet sich bei Chelani et al. und Klingner et al. [18, 55], die jedoch auch nur einen Teil der möglicherweise geeigneten Einflussfaktoren auf Konzentrationen von PM_{10} miteinbeziehen. So wird z.B. die Mischungsschichthöhe als wichtiger Parameter zur Vorhersage vollständig außer Acht gelassen. Partanen et al. [73] führen einen Vergleich der Ergebnisse eines NN und eines klassischen Ausbreitungsmodells durch, der zu Gunsten des NN ausfällt.

Ein weiteres maschinelles Lernverfahren, dass in der Luftreinhaltung jedoch bisher nur vereinzelt Anwendung gefunden hat, ist das *k-Nearest-Neighbour-Verfahren*. Bisher wurden mit diesem Verfahren Versuche unternommen, Ozon zu prognostizieren. Athanasiadis et al. [8] und Efraimidou et al. [28] erzeugten nur leidlich gute Vorhersagen, dies jedoch für qualitative Aussagen über die Ozonbelastung und nicht für numerische Konzentrationen. Deutlich bessere Ergebnisse erzielten Singh et al. [99] bei der Prognose von meteorologischen Parametern. Auch die Vorhersage zeitlicher Verläufe in anderen Disziplinen wie den Wirtschaftswissenschaften erbrachte Erfolg versprechende Ergebnisse [39, 100]. Bei Lohmeyer et al. [61] finden sich erste Ergebnisse zur Prognose von PM_{10} -Konzentrationen. Wiederum werden jedoch die Unsicherheiten bei den meteorologischen Vorhersagen außer Acht gelassen und

die Feinstaubprognosen nur mit Messwerten der meteorologischen Einflussfaktoren durchgeführt. Eine Abschätzung der Leistungsfähigkeit unter realen Bedingungen ist somit nicht möglich.

Die *Support-Vektor-Maschine (SVM)* ist ein vergleichsweise junges, maschinelles Lernverfahren, das sich gut zur Prognose von zeitlichen Verläufen eignet [71, 70, 34]. Eine Studie der Universität Xi'an berichtet über den Einsatz verschiedener maschineller Lernverfahren zur Prognose von Respirable Suspended Particulates [115, 63]. Dabei werden u.A. auch Ergebnisse einer SVM mit denen eines NN verglichen. Selbst unter Verwendung eines sehr limitierten Satzes an meteorologischen Einflussfaktoren, bestehend aus lediglich fünf verschiedenen Parametern, scheinen vor allem die Ergebnisse der SVM durchaus Erfolg versprechend. Einen ähnlichen Vergleich, diesmal für PM_{10} -Grenzwertüberschreitungen, findet man bei Raimondo et al. [88] – erneut mit guten Ergebnissen. Er beschreibt darüber hinaus die große Bedeutung einer Optimierung der Einflussfaktoren-Zusammenstellung bei maschinellen Lernverfahren. Ähnlich sehen dies auch Zickus et al. [122]. Sie legen dar, dass verschiedene Lernverfahren aus den geeigneten Einflussfaktorensatz teilweise dieselben, teilweise aber auch völlig unterschiedliche Faktoren auswählen, um ihre jeweils maximale Leistung zu erbringen.

1.2 Aufgabenstellung

Ziel der Arbeit ist die Erstellung eines Systems zur Kurzfrist-Prognose von PM_{10} -Konzentrationen für den nächsten Tag für vier ausgewählte Messstationen der LUBW (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg). Dabei soll ein PM_{10} -Tagesmittelwert prognostiziert werden, da dies hinsichtlich der EU-Gesetzgebung und der damit verbundenen Grenzwerte für 24-h-Mittel sinnvoll erscheint. Außerdem wird dieser Tagesmittelwert zur Berechnung des deutschen Kurzzeit-Luftqualitätsindex (LuQx) herangezogen. Prognoseziele sind deshalb der Tagesmittelwert der PM_{10} -Konzentrationen an einer Messstation und damit verbunden die Überschreitungen des EU-Grenzwertes sowie die Werte des deutschen Kurzzeit-Luftqualitätsindex.

Für die Prognose werden mehrere maschinelle Lernverfahren eingesetzt, die in der Lage sind aus Erfahrungen – in diesem Fall bestehend aus historischen Messreihen – auf zukünftige Situationen zu schließen. „The best qualification of a prophet is to

have a good memory.“ Dieses Zitat des Marquis von Halifax, George Savile, trifft also in besonderem Maße auf die maschinellen Lernverfahren zu. Die Aufstellung expliziter Regeln über die Zusammenhänge zwischen PM_{10} und den verschiedenen Einflussfaktoren aus historischen gemessenen Daten entfällt dabei und wird von den maschinellen Lernverfahren übernommen.

Diese in anderen Bereichen bewährten Lernverfahren sind für die Vorhersage von PM_{10} -Konzentrationen teilweise noch völlig unerforscht, teilweise aber auch unter stark eingeschränkten Rahmenbedingungen untersucht worden. Deshalb sollen die unterschiedlichen Methoden und deren Ergebnisse in dieser Arbeit unter einheitlichen, geeigneten Rahmenbedingungen miteinander verglichen und bewertet werden. Die Nutzung derselben Ausgangsdaten zur Modellbildung und Leistungsabschätzung für alle Methoden ist dabei von entscheidender Wichtigkeit. Es stellt für die Prognose von PM_{10} -Konzentrationen eine Neuerung dar, verschiedenste maschinelle Lernverfahren auf einer solchen einheitlichen und breiten Datenbasis miteinander zu vergleichen.

Die Prognoseleistung aller Modelle ist stark abhängig von der Güte der Eingangsdaten. Eine Eingangsgröße der Modelle bilden die meteorologischen Vorhersagen. Da diese mit großen Unsicherheiten behaftet sind, werden die Modelle einer Sensitivitätsanalyse unterzogen. Dabei wird untersucht, inwieweit sich von der Realität abweichende meteorologische Vorhersagen auf die Güte der PM_{10} -Prognose auswirken. Erst diese Sensitivitätsanalyse ermöglicht es, Aussagen über die tatsächliche Prognoseleistung der einzelnen erstellten Modelle zu treffen.

Die Ziele der Arbeit lassen sich wie folgt zusammenfassen:

- Bereitstellung und Aufbereitung einer möglichst breiten Datenbasis für die PM_{10} -Prognose für vier Messstationen der LUBW
- Erstellung und Optimierung von Modellen zur Prognose von PM_{10} -Tagesmittelwerten mittels maschineller Lernverfahren basierend auf identischen Eingangsdaten
- Sensitivitätsanalyse bezüglich unsicherer meteorologischer Eingangsdaten (Vorhersagen)
- Vergleich und Bewertung der verschiedenen Modelle und ihrer Leistungsfähigkeit

Kapitel 2

Grundlegende Aspekte zum Thema Feinstaub

2.1 Entstehung und Quellen von Feinstaub

Staub ist ein komplexes Gemisch aus festen und flüssigen Partikeln. Diese unterscheiden sich hinsichtlich ihrer Größe, Form, Farbe, chemischen Zusammensetzung, physikalischen Eigenschaften und ihrer Herkunft bzw. Entstehung. PM₁₀, in dieser Arbeit als Feinstaub bezeichnet, ist dabei die Fraktion mit einem aerodynamischen Durchmesser kleiner 10 μm . Grundsätzlich wird zwischen *primären* und *sekundären Partikeln* unterschieden. Erstere werden als primäre Emissionen direkt in die Atmosphäre abgegeben, Letztere entstehen durch luftchemische Prozesse aus gasförmig emittierten Vorläufersubstanzen [111].

Die Hauptursache für primäres PM₁₀ sind anthropogene Quellen wie der Verkehr, Feuerungsanlagen und industrielle Produktionsprozesse, wobei sich ihr Anteil am Gesamtstaub (TSP) je nach Sektor deutlich unterscheidet (siehe Abbildung 2.1). Die Bestandteile der primären Partikel reichen dabei von Ruß und unverbrannten Kohlenwasserstoffen über Metalle bis hin zu Sulfaten aus Verbrennungsprozessen. Dazu kommen vielfältige Partikel, die aus der mechanischen Bearbeitung von z.B. Holz und Metallen, dem Abrieb von Bremsen, Reifen und Straßen sowie dem Umschlag von staubenden Gütern stammen [87]. Natürlich entstandene Primärpartikel spielen hingegen in Baden-Württemberg laut Angaben der Landesanstalt für Umwelt, Messungen und Naturschutz lediglich eine untergeordnete Rolle [64].

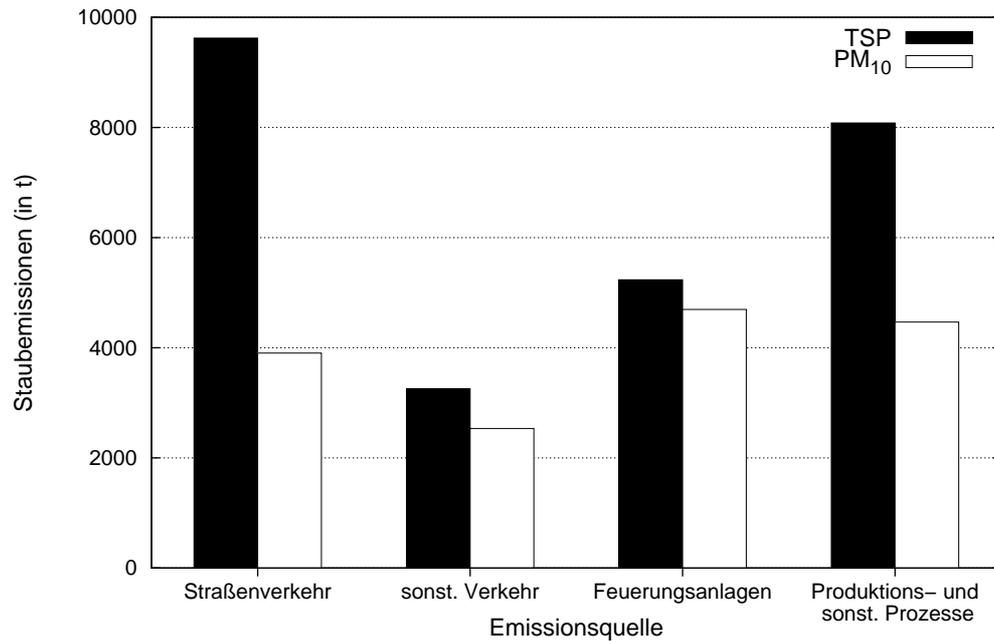


Abbildung 2.1: Emissionen für Gesamtstaub (TSP) und PM₁₀ im Jahr 2004 in Baden-Württemberg nach Quellgruppen [106]

Sekundär entstandenes PM₁₀ macht mit 35% einen erheblichen Anteil des gesamten Feinstaubes aus (siehe Tabelle 2.1). Sekundäre Partikel entstehen durch chemische Prozesse in der Atmosphäre aus gasförmigen Vorläufern wie Ammoniak (NH₃), flüchtigen organischen Verbindungen (VOC), Stickstoffoxiden (NO_x) und Schwefeldioxid (SO₂). Hauptquellen des Ammoniaks sind die Landwirtschaft, insbesondere die Tierhaltung, und daneben auch der Straßenverkehr. VOCs werden vor allem bei der Verwendung von Lösemitteln, aber auch in der Landwirtschaft und im Straßenver-

Tabelle 2.1: Anteile von primärem und sekundärem PM₁₀ im städtischen Hintergrund in Deutschland für 2001 [69]

Staubart	Anteil	davon	
		anthropogen	natürlich
primäre Partikel aus Verbrennung	35%	35%	0%
primäre Partikel aus sonstigen Quellen	30%	20%	10%
sekundäre Partikel	35%	30%	5%
gesamt	100%	85%	15%

kehr emittiert. Stickstoffoxide stammen überwiegend aus dem Straßenverkehr und Kraftwerken. Letztere sind neben dem Hausbrand auch Hauptemittenten von Schwefeldioxid [26].

2.2 Auswirkung von Feinstaub auf die menschliche Gesundheit

In mehreren *epidemiologischen Studien* wurden sowohl in Europa [104, 108, 113] als auch in den USA [84, 85, 95, 97, 96] die Auswirkungen von PM_{10} - und auch $PM_{2.5}$ -Konzentrationen in der Außenluft auf die Gesundheit untersucht. Dabei wurde ein signifikanter Zusammenhang zwischen der atmosphärischen Feinstaubbelastung und der *Morbidität* und *Mortalität* beim Menschen festgestellt. Als morbide *Kurzzeiteffekte* sind u.a. Veränderungen der Lungenfunktionsparameter und die Zunahme von Atemwegssymptomen zu nennen. Ebenso steigen durch kurzzeitige Exposition der Gebrauch von Medikamenten und die Zahl der Krankenhausaufenthalte infolge Atemwegserkrankungen an (siehe Tabelle 2.2). Risikogruppen sind dabei vor allem Asthmatiker und Kinder. Nach Studien in 29 europäischen und 20 nordamerikanischen Städten geht man von einer Steigerung der Mortalität um 0,62% bzw. 0,46% bei einem Anstieg des Tagesmittelwerts um je $10 \mu\text{g}/\text{m}^3$ PM_{10} aus [53, 92, 118]. Zwei Studien außerhalb Europas und den USA gehen mit 0,5% bzw. 0,49% für asiatische Städte von ähnlichen Werten aus [20, 44].

Ausgedehnte Langzeitstudien [85, 25] und neuere Analysen der dabei erhobenen Daten [57, 84] belegen zudem, dass ebenso die langfristige Exposition gegenüber Feinstaub Auswirkungen auf die menschliche Gesundheit hat (siehe Tabelle 2.3). Wo die Effekte durch kurzfristige Exposition durchaus reversibel sind, scheinen die *Langzeiteffekte* zu einer Schwächung des Immunsystems gegenüber Infektionen und somit langfristig zu einer erhöhten Mortalität innerhalb der gesamten Bevölkerung zu führen.

Neben den in epidemiologischen Studien untersuchten Zusammenhängen zwischen Feinstaubbelastung und Morbidität bzw. Mortalität betrachten zahlreiche *toxikologische Studien* die Wirkung von Feinstaub auf die Gesundheit. Jedoch erschwert die Vielfalt an Zusammensetzung und Struktur von Partikeln die Bewertung seiner toxikologischen Bedeutung. Grundsätzlich muss man aber von *spezifischen* und *un-*

Tabelle 2.2: Abschätzung der relativen Gesundheitsrisiken verbunden mit einem Anstieg der PM₁₀-Konzentration um 10 µg/m³ im Tagesmittel (Kurzzeiteffekte) nach WHO [117, 118]

	relatives Risiko durch PM₁₀ (95% Konfidenzintervall)
Medikamenteneinsatz	1,0305 (1,0201 – 1,0410)
Hustensymptome	1,0356 (1,0197 – 1,0518)
Abnahme von Lungen- funktionsparametern	1,0305 (1,0201 – 1,0410)
Veränderungen des max. expiratorischen Flusses (rel. zum Durchschnitt)	-0,13% (-0,17% – -0,09%)
atemwegsbedingte Kran- kenhausaufenthalte	1,0080 (1,0048 – 1,0112)

Tabelle 2.3: Abschätzung der relativen Risiken durch langfristige Exposition gegenüber PM_{2.5} bzw. PM₁₀ verbunden mit einem Anstieg der Konzentration um 10 µg/m³ nach WHO [117, 118]

	rel. Risiko durch PM_{2.5} (95% Konfidenzintervall)	rel. Risiko durch PM₁₀ (95% Konfidenzintervall)
Tod (nach [25])	1,14 (1,04 – 1,24)	1,10 (1,03 – 1,18)
Tod (nach [86])	1,07 (1,04 – 1,11)	—
Bronchitis (nach [24])	1,34 (0,94 – 1,99)	1,29 (0,96 – 1,83)
Änderung eines Lungen- funktionsparameters bei Kindern (nach [89])	-1,9% (-3,1% – -0,6%)	-1,2% (-2,3% – -0,1%)
Änderung eines Lungen- funktionsparameters bei Erwachsenen (nach [1])	— (—)	-1,0% (—)

spezifischen Wirkungen ausgehen. Dabei spricht man von spezifischer Wirkung, wenn Partikel Substanzen mit toxischen Eigenschaften beinhalten oder diese der Partikeloberfläche anhaften. Dabei sind vor allem Schwermetalle wie Blei, Cadmium und Quecksilber und die zum Teil kanzerogenen PAKs (polyzyklische aromatische Kohlenwasserstoffe) zu nennen. Unter einer unspezifischen Wirkung versteht man den bloßen Fremdkörperreiz, der durch inerte Partikel ausgelöst wird. Allerdings können diese Partikel wiederum als Träger für schädliche organische und anorganische Verbindungen dienen [87].

Zunehmend in den Fokus der Forschung geraten die *ultrafeinen Stäube*. Eine entscheidende Rolle spielt bei ihnen scheinbar die Anzahl- bzw. die Oberflächenkonzentration. Ultrafeinen Partikeln schreibt man durch ihre besondere Lungengängigkeit eine hohe Toxizität zu. Sie sollen zu erhöhtem Risiko für Erkrankungen der Herzkranzgefäße und Herzinfarkt [98, 79] bzw. zu einer Beeinträchtigung der Herzfunktion [40] führen. Einen weitergehenden Überblick aus der Wirkungsforschung bieten beispielsweise Peters et al. und Sloss [79, 101].

2.3 Gesetzliche Richtlinien bezüglich PM₁₀

Seit dem 1. Januar 2005 gilt in Deutschland der verschärfte Grenzwert von $50 \mu\text{g}/\text{m}^3$ für die Konzentration von PM₁₀ in der Luft. Den Grundstein für diese Änderung legte am 27. September 1996 die *europäische Luftqualitätsrahmenrichtlinie* 96/62/EC [31, 45]. Diese hat zum Ziel, weitestgehend unabhängig von bestimmten Emittenten gewisse Luftqualitätszielwerte zu erreichen und die Öffentlichkeit über die Luftqualität zu informieren.

Im nächsten Schritt wurden in den so genannten *Tochterraichtlinien* konkrete Grenzwerte für einzelne Komponenten festgelegt. Die Grenzwerte für PM₁₀ werden in der Richtlinie 99/30/EC vom 22. April 1999 [32] spezifiziert – sie sind aus der Tabelle 2.3 ersichtlich. Erstmals bezieht sich damit ein Grenzwert für Staub nicht mehr auf den Gesamtstaub, sondern lediglich auf die Fraktion mit einem aerodynamischen Durchmesser kleiner $10 \mu\text{g}/\text{m}^3$. So soll den Ergebnissen aus epidemiologischen und toxikologischen Studien Rechnung getragen werden. Wenn die geforderten Werte nicht eingehalten werden können, sind von den zuständigen Stellen Luftreinhaltepläne an-

zufertigen und umzusetzen. Diese sollen Maßnahmen zur Senkung der Emissionen enthalten, sodass die Grenzwerte in Zukunft eingehalten werden können.

Die Integration dieser europäischen Richtlinien in deutsches Recht erfolgte mit der Novellierung der *22. Verordnung zur Durchführung des Bundesimmissionsschutzgesetzes (BImSchV)* [21] vom 11. September 2002. Aktuell werden die erlaubten Maximalwerte in deutschen Großstädten an stark belasteten Tagen jedoch deutlich überschritten. So wurde im Jahr 2005 an insgesamt 60 deutschen Messstationen der Grenzwert mehr als 35-mal überschritten. Der Spitzenreiter ist dabei die Station „Am Neckartor“ in Stuttgart mit zu hohen PM_{10} -Konzentrationen an 187 Tagen [110]. 2006 wurden bereits im April an einer zweistelligen Zahl deutscher Stationen mehr als 35 Überschreitungen erfasst, darunter wiederum Stationen in Städten wie Stuttgart, München und Berlin [16].

Im nächsten Schritt sollen durch die Einführung neuer, noch schärferer Grenzwerte ab 2015 die Bemühungen um eine saubere Luft in Europa weiter vorangetrieben werden [33]. Jedoch behält sich die europäische Kommission mit einer Klausel vor, ggf. die zukünftigen Grenzwerte zu ändern bzw. Grenzwerte für andere Partikelfractionen wie $PM_{2.5}$ zu erlassen [41]. Basieren sollen diese Änderungen auf neuen Erkenntnissen bezüglich der Auswirkungen auf Gesundheit und Umwelt bzw. technischen Neuerungen und Erfahrungen mit der Anwendung der aktuellen Richtlinie.

Tabelle 2.4: Grenzwerte für PM₁₀ nach der EU-Richtlinie 99/30/EC [32]

Bezeichnung	Grenzwert	Toleranzmarge	Zeitpunkt, bis zu dem der Grenzwert zu erreichen ist
Stufe 1			
24-Stunden-Grenzwert	50 µg/m ³ PM ₁₀ dürfen nur 35-mal im Jahr überschritten werden.	50% bei Inkrafttreten dieser Richtlinie. Lineare Reduzierung am 1.1.2001 und alle 12 Monate bis auf 0% am 1.1.2005.	1.1.2005
Jahres-durchschnitts-grenzwert	40 µg/m ³ PM ₁₀	20% bei Inkrafttreten dieser Richtlinie. Lineare Reduzierung am 1.1.2001 und alle 12 Monate bis auf 0% am 1.1.2005.	1.1.2005
Stufe 2*			
24-Stunden-Grenzwert	50 µg/m ³ PM ₁₀ dürfen nur 7-mal im Jahr überschritten werden.	Aus Daten abzuleiten. Gleichwertig mit Grenzwert der Stufe 1.	1.1.2010
Jahres-durchschnitts-grenzwert	20 µg/m ³ PM ₁₀	50% am 1.1.2005. Lineare Reduzierung alle 12 Monate bis auf 0% am 1.1.2010	1.1.2010

*Grenzwerte sollen im Lichte neuer Informationen bezüglich Auswirkungen auf Gesundheit und Umwelt, technischen Möglichkeiten und Erfahrungen mit der Anwendung der Grenzwerte aus Stufe 1 in den Mitgliedsstaaten überprüft werden.

Kapitel 3

Vorstellung der angewandten Methodik

Die in dieser Arbeit angewandte Methodik lässt sich in mehrere Bereiche unterteilen. Im Vordergrund stehen zunächst die unterschiedlichen maschinellen Lernverfahren, die zur Prognose von PM_{10} -Konzentrationen herangezogen werden. Anschließend wird die Herangehensweise bei der Modelloptimierung erläutert. Da die Hauptzielsetzung der Arbeit neben einer möglichst genauen Feinstaubprognose eine Gegenüberstellung verschiedener Modelle ist, wird auch auf die Methodik zum Leistungsvergleich sorgfältig eingegangen. Zum Abschluss des Kapitels folgt ein kurzer Abriss über das Preprocessing der verwendeten Daten.

3.1 Vorstellung der verwendeten Verfahren zur Prognose von PM_{10} -Konzentrationen

Die grundlegende Aufgabenstellung bei *maschinellen Lernverfahren* ist es, Zusammenhänge zwischen Beobachtungen X und den Reaktionen eines Systems Y zu lernen, um für neue Beobachtungen X' Aussagen treffen zu können. Dies beinhaltet z.B. die Vorhersage der darauffolgenden Reaktion Y' [102].

Im Fall der PM_{10} -Prognose bestehen die Beobachtungen (oder auch Instanzen) aus einem Vektor mit verschiedenen Attributen. Mögliche Attribute sind z.B. meteorologische Größen wie Windgeschwindigkeit oder Niederschlagsmenge (sie werden in

Abschnitt 4.1 näher beschrieben). Aus diesem Grund kann man auch von einem Beobachtungsvektor \vec{X} sprechen. Jedem dieser Vektoren ist eine Reaktion des Systems zugeordnet – in diesem Fall die zu prognostizierende PM_{10} -Konzentration.

Die Erstellung eines PM_{10} -Prognosemodells lässt sich der obigen Notation folgend schematisch so beschreiben [38]:

- Grundlage: Trainingsdaten (Beobachtungen \vec{X} , Reaktionen Y)

$$T = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_n, y_n)\}$$

- Gesucht: approximierende Funktion (Modell)

$$\hat{f}(\vec{x}) = \hat{y} \approx y$$

- Ziel: Gute Vorhersage von Testdaten bzw. unbekanntem Daten

$$G = \{(\vec{x}'_1, y'_1), (\vec{x}'_2, y'_2), \dots, (\vec{x}'_m, y'_m)\}$$

Grundsätzlich ist die Vorgehensweise bei der Modellbildung und der Prognose für alle im Folgenden vorgestellten Lernverfahren dieselbe. Ein Teil der zur Verfügung stehenden Beobachtungen und die dadurch verursachten Reaktionen werden herangezogen, um ein Modell zu trainieren. *Training* bedeutet, dass mittels eines Lernverfahrens Zusammenhänge zwischen \vec{X} und Y erkannt und in einem Modell zusammengefasst werden. Das Modell wird während des Trainingsvorgangs immer weiter verbessert. Da bei diesem Lernprozess sowohl die Eingabe \vec{X} als auch die korrekte Ausgabe Y zur Verfügung stehen, spricht man vom *überwachten Lernen* (*supervised learning*) [60]. Ein trainiertes Modell sollte nun in der Lage sein, für neue Beobachtungen eine Prognose zur Reaktion des Systems anzustellen. Dies lässt sich überprüfen, indem man das Prognosemodell auf Beobachtungen, die nicht Teil der Trainingsdaten waren, anwendet und die Prognoseergebnisse mit den korrekten Reaktionen des Systems vergleicht (siehe Abschnitt 3.3). Der Lernprozess im Detail unterscheidet sich für die verschiedenen maschinellen Lernverfahren. In den folgenden Abschnitten sollen diese vorgestellt werden.

Softwareseitig kommen zur Modellerstellung zum einen YALE (Yet Another Learning Environment) [35, 68] für die multiple lineare Regression, das instanzbasierte Lernen und die Support-Vektor-Maschine, zum anderen QNET v2000 [114] für das Multilayer-Perzeptron zum Einsatz.

3.1.1 Die multiple lineare Regression als Referenzverfahren

Die lineare Regressionsanalyse ist ein häufig eingesetztes statistisches Verfahren, um Beziehungen zwischen einer abhängigen Variable und einer oder mehreren unabhängigen Variablen festzustellen. Bei Verwendung mehrerer Unabhängiger spricht man von einer *multiplen linearen Regression (LR)* [42]. Sie wird in verschiedenen deutschen Bundesländern bereits zur Prognose von PM_{10} eingesetzt [48, 61] und dient deshalb in dieser Arbeit als Referenzverfahren.

Beim Erstellen einer LR geht man davon aus, dass y von mehreren erklärenden Variablen x_n abhängt und erhält folgende Form [42]:

$$y = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_n * x_n + \epsilon \quad (3.1)$$

y stellt dabei den *Regressand*, die abhängige Variable, dar. α ist ein konstantes Glied. β_n ist der *Regressionskoeffizient* des n -ten Regressors (x_n), also der unabhängigen Variablen. ϵ repräsentiert eine Störgröße. Sie stellt nicht erfassbare Einflüsse dar. Um eine Regressionsfunktion zu erstellen, gilt es ϵ zu minimieren. Dies geschieht anhand der *Methode der kleinsten Quadrate (Least Square Method)*. Dabei wird die Summe der quadrierten Abweichungen der Funktion zu den tatsächlichen Beobachtungen minimiert. Als Ergebnis erhält man z.B. für die PM_{10} -Prognose ein Modell der Form:

$$PM_{10} = PM_{10Vortag} - 2 * Windgeschwindigkeit - 4 * Niederschlag + \dots \quad (3.2)$$

Grundbedingung für den Einsatz der LR ist, dass die Daten in ihrer Grundgesamtheit *normalverteilt* vorliegen. In der Praxis wird jedoch auch eine angenäherte Normalverteilung als ausreichend angesehen [91]. Um diese herzustellen, werden die Variablen in eine solche Form transformiert. Ebenso sollten alle Variablen unabhängig voneinander sein. Dieses Ideal lässt sich jedoch in der Praxis nicht ohne weiteres umsetzen, da viele der eingesetzten meteorologischen Kennzahlen nicht gänzlich unabhängig sind. Aus diesem Grund wurde eine Annäherung an das Unabhängigkeitsprinzip durchgeführt, die in diesem Fall ausreichen sollte [80].

Wie der Name schon sagt ist die LR auf die Beschreibung linearer Zusammenhänge beschränkt. Bei der Prognose von PM_{10} ist allerdings davon auszugehen, dass z.B. die Zusammenhänge zwischen meteorologischen Faktoren und der PM_{10} -

Konzentration nichtlinear sind. Um diese abbilden zu können, sollen maschinelle Lernverfahren zum Einsatz kommen, die dazu besser geeignet scheinen.

3.1.2 Das instanzbasierte Lernen mit den k -nächsten-Nachbarn

Das *Instanzbasierte Lernen mit den k -nächsten-Nachbarn (IBk)* ist eine Weiterentwicklung der *k -Nearest-Neighbour-Methode (kNN)*. Aus diesem Grunde sei zunächst die kNN beschrieben. Sie stellt eines der grundlegenden Verfahren zur Klassifizierung von Daten dar [9, 121], eignet sich jedoch auch zur numerischen Vorhersage.

Beim Trainingsprozess werden alle Trainingsdaten als Vektoren abgespeichert. Soll für eine neue Beobachtung eine Prognose erstellt werden, berechnet die kNN den Abstand aller Trainingsvektoren zum neuen Beobachtungsvektor. Der Abstand wird dabei mit der Euklidischen Distanz berechnet (siehe Formel 3.3).

$$d = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2} \quad (3.3)$$

Die k Trainingsvektoren mit dem geringsten Abstand nennt man die k -nächsten-Nachbarn – k bezeichnet dabei die Anzahl. Aus dem Mittel der Reaktionen des Systems der k -nächsten-Nachbarn ergibt sich der Prognosewert (s. Abb. 3.1). Nachteile dieses Verfahrens sind die Sensitivität gegenüber einem Rauschen der Trainingsdaten und der enorme rechnerische Aufwand bei großen Trainingsdatensätzen – es muss der Abstand jeder Trainingsinstanz zur neuen Beobachtung errechnet werden. Aus diesem Grund fand eine Weiterentwicklung der kNN zum IBk statt [3].

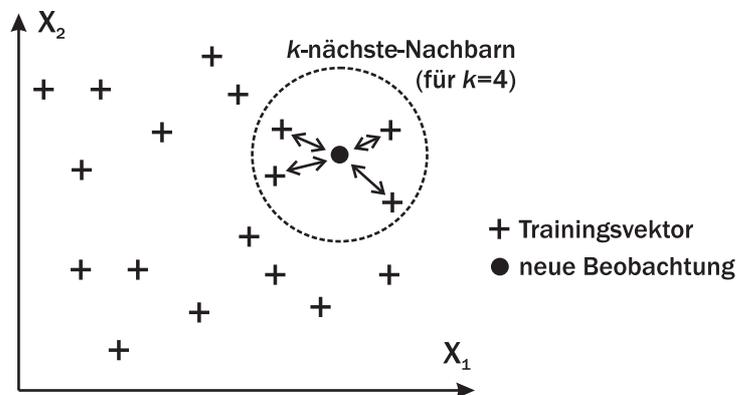


Abbildung 3.1: Darstellung der k -Nearest-Neighbour-Methode

IBk unterscheidet sich von kNN dadurch, dass nicht alle Trainingsinstanzen zur Modellbildung abgespeichert werden, sondern nur ausgewählte. Ähnliche Instanzen werden zusammengefasst. Dieses Vorgehen verkürzt die Berechnungszeit bei der Prognose, da jetzt nicht mehr so viele Distanzen ermittelt werden müssen. Ein weiterer Unterschied zwischen IBk und kNN besteht darin, dass die k nächsten Nachbarn gewichtet – nach ihrem Abstand zu einer neuen Beobachtung – in die Prognose eingehen. Ein Nachbar, der doppelt so weit von der neuen Beobachtung entfernt ist wie ein anderer, trägt nur mit halbem Gewicht zur Berechnung des Prognosewertes bei.

3.1.3 Das Multilayer Perzeptron – ein künstliches neuronales Netz

Ein *künstliches Neuronales Netz (NN)* ist die Kombination einer komplizierten, nicht-linearen Funktion mit einem einfachen Lern-Algorithmus, der durch iterative oder rekursive Vorgehensweise aus vorhandenen Eingangs- und gewünschten Ausgangswerten alle Parameter der Funktion zu bestimmen versucht [90].

Grundsätzlich setzt sich ein neuronales Netz aus einzelnen, miteinander verknüpften *Neuronen* (s. Abb. 3.2) zusammen. Jedes Neuron besteht dabei aus einer beliebigen Anzahl von Eingängen (zusammengefasst zu einem Eingangsvektor \vec{x}), die jeweils mit einem Gewicht versehen sind (Gewichtsvektor \vec{w}). Die Produkte aus Eingaben und zugehörigem Gewicht werden durch eine *Übertragungsfunktion* aufsummiert. Diese Summe wird an die *Aktivierungsfunktion* weitergeleitet, die die Ausgabe des Neurons bestimmt, welche dann an beliebig viele Neuronen weitergegeben wird [60].

NN werden nach ihrem Aufbau, man spricht von der *Netztopologie*, in unterschiedliche Typen eingeteilt. Sie können sich aus mehreren einzelnen Neuronen zusammensetzen, die jeweils untereinander verbunden sind, wie z.B. bei einem *Kohonen-Netz*. Sie können jedoch auch aus verschiedenen Schichten von Neuronen bestehen,

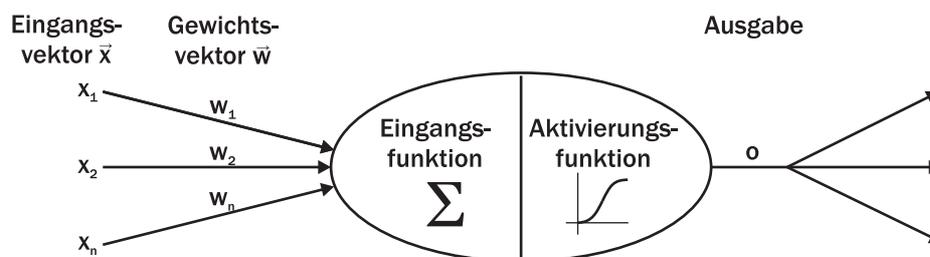


Abbildung 3.2: Schematische Darstellung eines künstlichen Neurons

wobei die Neuronen einer Schicht nicht miteinander, sondern nur mit der nachfolgenden Schicht verbunden sind. Eine solche Topologie besitzt das *Multilayer Perzeptron (MLP)*, welches in dieser Arbeit zur PM_{10} -Prognose verwendet wird. Im Detail besteht es aus einer *Eingabeschicht*, einer oder mehreren *verdeckten Schichten* und der *Ausgabeschicht* (s. Abb. 3.3) [37]. Jede Schicht besteht dabei aus einer Anzahl von Neuronen, die ihren Input von der vorhergehenden Schicht erhalten und ihren Output an die nächste weitergeben. Es handelt sich also um ein Feedforward Netz. Eine Rückkopplung (Feedback) zu vorhergehenden Schichten wie beim so genannten Elman-Netz [29] oder ein Überspringen von Schichten (Shortcut) ist dabei nicht vorgesehen. Lediglich die Ausgabeschicht gibt ihren Output nicht an eine folgende Schicht weiter, sondern liefert das Ergebnis eines NN.

Das MLP ist theoretisch in der Lage, jede kontinuierliche Funktion anzunähern [90], was der Grund für die Wahl dieser Netztopologie für die PM_{10} -Prognose ist. Außerdem wurden mit diesem Netz bereits mehrfach Erfolge bei der Prognose von Luftschadstoffen erzielt (siehe Kapitel 1.1). Grundsätzlich ist auch ein *Elman-Netz* zur Prognose von Zeitreihen geeignet, da es mit seiner Fähigkeit zur Rückkopplung eine Art Gedächtnis hat. Es kann also z.B. einen 24-h-Prognosewert berechnen und diesen bei einer anschließenden 48-h-Prognose miteinbeziehen. Da in dieser Arbeit jedoch nur eine Prognose für die nächsten 24-h durchgeführt wird, lässt sich dieser Vorteil nicht nutzen.

Zum Lernen wird der *Backpropagation-Algorithmus* verwendet. Dazu werden zunächst die Gewichte zufällig festgelegt. Den folgenden Lernprozess kann man in drei alternierend ablaufende Schritte unterteilen [60]:

1. Der Forward-Pass

Der Eingabeschicht wird ein Inputdatensatz aus der Trainingsmenge zur Verfügung gestellt. Anschließend wird für jedes Neuron der ersten verdeckten Schicht durch die Aktivierungsfunktion und die Ausgabefunktion der Output berechnet. Dieser wird an die nächste Schicht weitergeleitet, in der ebenso verfahren wird, bis schließlich die Ausgabeschicht erreicht und ein Ausgabevektor berechnet ist.

2. Die Bestimmung des Fehlers

Die vom Netz gelieferte Ausgabe wird mit der korrekten (erwarteten) Ausgabe verglichen und daraus eine Fehlerfunktion berechnet.

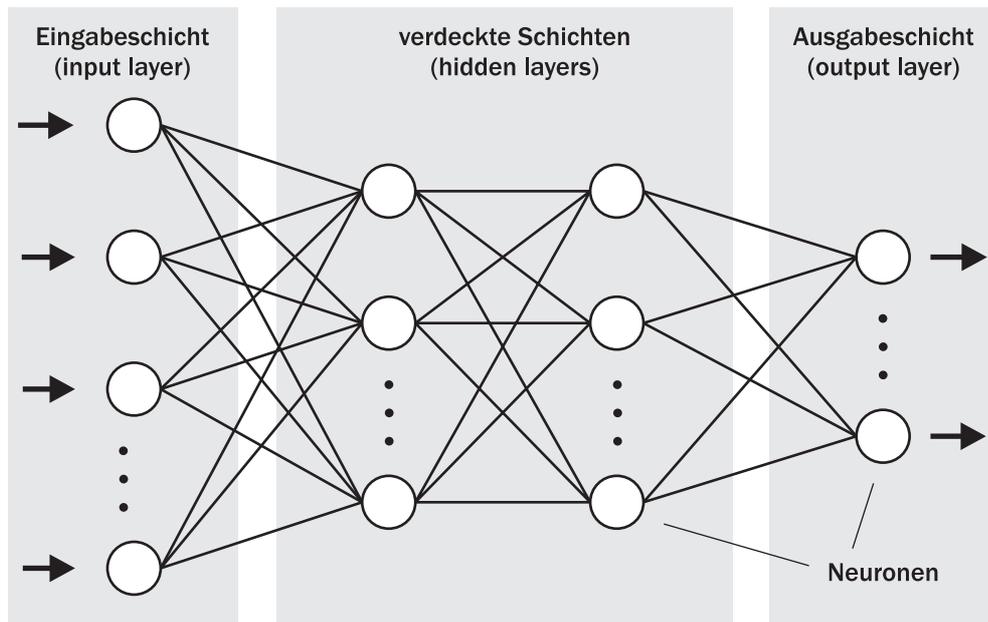


Abbildung 3.3: Schematische Darstellung eines Multilayer-Perzeptrons

3. Der Backward-Pass

Anhand der erstellten Fehlerfunktion werden die Gewichte rückwärtsgerichtet von der Ausgabeschicht zur Eingabeschicht verändert. Die Änderungen einer Schicht dienen also als Basis für die Gewichtsänderung der vorhergehenden Schicht.

Diese drei Schritte werden so lange durchgeführt, bis ein zufriedenstellendes Fehlerminimum erreicht ist. Der Backpropagation-Algorithmus sucht also durch die iterative Änderung von Gewichten ein idealerweise globales Minimum einer Fehlerfunktion. Diese kann vereinfacht durch das Auftragen des Fehlers auf der y-Achse und der Gewichte auf der x-Achse dargestellt werden. Beim Start des Trainings an einer beliebigen Stelle dieser Fehlerkurve wird durch den so genannten Gradientenabstieg das Erreichen eines Minimums angestrebt.

Ein Problem des MLP besteht darin, dass beim Training des Netzes die Lerndaten nur Stützpunkte der anzunähernden Funktion sein können. Später muss das trainierte Netz allerdings auch unbekannte Daten beurteilen, die zwischen diesen Stützpunkten liegen. Wird das Training zu lange durchgeführt, approximiert das Netz nicht mehr die Funktion, sondern beginnt die Trainingsdaten auswendig zu lernen – man spricht vom *Overfitting*. Um dies zu verhindern, wird schon während des Lernprozesses

ein Teil der Daten zur Überprüfung der Leistungsfähigkeit des Netzes herangezogen. Wird beim Training die Leistung des Netzes schlechter, so ist dies ein Hinweis auf Overfitting und der Lernvorgang sollte abgebrochen werden.

3.1.4 Die Regression mittels Support-Vektor-Maschine

Eine *Support-Vektor-Maschine (SVM)* bietet zwei grundsätzliche Möglichkeiten des maschinellen Lernens an. Dies ist zum einen die Klassifikation als die ursprüngliche Form einer SVM. Ziel einer *klassifizierenden SVM* ist es, eine möglichst einfache Trennfläche zwischen zwei Klassen zu finden. Diese Trennfläche wird von so genannten Support-Vektoren eindeutig beschrieben [112]. Die Klassifizierung kann man z.B. zur Vorhersage von PM_{10} -Grenzwertüberschreitungen anwenden, wobei das Ergebnis 1 (Überschreitung) oder -1 (keine Überschreitung) lautet. Zum anderen kann die SVM auch zur Regression und somit zur numerischen Vorhersage genutzt werden. Man spricht in diesem Fall auch von der *Support-Vektor-Regression (SVR)*. Ziel einer SVR ist, eine Kurve zu finden, die den Zusammenhang zwischen Beobachtungen X und Reaktionen des Systems Y gut interpoliert, dabei aber berücksichtigt, dass die Beobachtungen fehlerbehaftet sein können.

In beiden Fällen, Klassifikation und Regression, geht es nun darum zu bewerten, wie sehr die Abweichungen zwischen der Vorhersage und der tatsächlichen Reaktion des Systems zu bestrafen sind. Dies geschieht anhand einer Kostenfunktion.

Im Falle der SVR ist um die Regressionsfunktion ein Puffer ε als Toleranzmarge gelegt. Innerhalb dieser Marge gelegene Abweichungen werden nicht als Fehler erachtet. Für die Abweichungen ζ , die über die Marge hinausgehen, werden die Kosten mithilfe der *Kostenfunktion* f_{Kosten} ermittelt (siehe Abbildung 3.4). Sie heißt aufgrund ihrer Toleranzmarge auch ε -unempfindliche Kostenfunktion. Durch Minimierung der Gesamtkosten lässt sich während des Trainings schließlich die optimale Regressionsfunktion bestimmen. Diese Funktion wird wiederum von den Support-Vektoren beschrieben.

Ist die Optimierung der Kosten für lineare Zusammenhänge relativ einfach, so kann sie für nichtlineare Fragestellungen eine Herausforderung sein. Zumal es im Falle der PM_{10} -Prognose nicht nur um die Abhängigkeit der PM_{10} -Konzentration von einem Attribut, sondern von einer Vielzahl solcher Attribute geht. Aus diesem Grunde

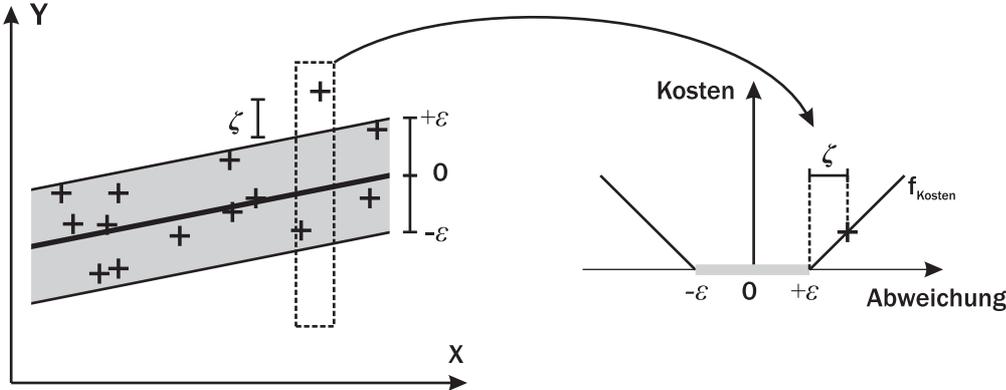


Abbildung 3.4: Darstellung einer Kostenfunktion mit Toleranzpuffer (nach [103])

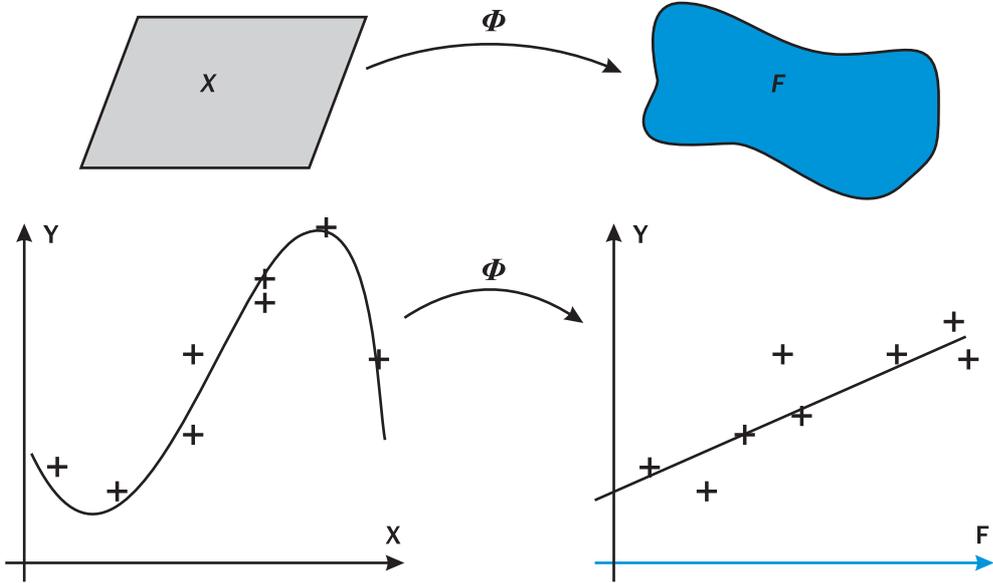


Abbildung 3.5: Transformation von Beobachtungen mittels der Funktion ϕ in einen höherdimensionalen Raum F nach [102])

verwenden SVM den so genannten *Kern-Trick*. Die Idee hinter dem Kern-Trick ist, die Beobachtungen mittels einer Funktion ϕ in einen höherdimensionalen Raum F zu überführen, wie dies Abbildung 3.5 zeigt. In einem Raum ausreichend hoher Dimensionalität – im Zweifelsfall unendlich – lässt sich wiederum eine lineare Funktion konstruieren. Bei der Rücktransformation in den niedrigerdimensionalen Raum wird aus der Regressionsgeraden eine nichtlineare, unter Umständen sogar nicht zusammenhängende Funktion, welche die Beobachtungen approximiert.

Bei diesem Vorgang stellen sich zwei Probleme: Die Hintransformation ist enorm rechenlastig und die Darstellung der Trennfläche im niedrigdimensionalen Raum im Allgemeinen äußerst komplex und damit praktisch unbrauchbar. Mit dem Kern-Trick ist es möglich, Hin- und Rücktransformation umzusetzen, ohne sie rechnerisch tatsächlich ausführen zu müssen. Für die Prognose mittels einer trainierten SVM ist es lediglich nötig, die Summe von Skalarprodukten der Support-Vektoren und eines Eingangsvektors, einer neuen Beobachtung, zu bilden.

3.2 Methodik zur Optimierung der Attributauswahl und der Modellparameter

Die Optimierung der Attributauswahl und der Modellparameter ist bei maschinellen Lernverfahren von besonderer Bedeutung. Nicht die Verwendung aller zur Verfügung stehenden Attribute führt zu den besten Ergebnissen, sondern vielmehr deren *optimale Kombination* [88] – dabei gilt es, den physikalischen Zusammenhang der Attribute zur PM_{10} -Konzentration zu beachten (siehe Abschnitt 4.1). Die Anzahl der möglichen Attributkombinationen bei n Attributen beträgt 2^n . Geht man, wie in dieser Arbeit, von 20 Attributen aus, so ergeben sich alleine daraus mehr als 1 Mio. Kombinationsmöglichkeiten. Darüber hinaus variieren die optimalen Parametereinstellungen der verschiedenen Modelle in Abhängigkeit von der Attributauswahl. Je nach Modell können dies ein oder mehrere z.T. realzahlige Parameter sein wie z.B. der Kostenparameter C und der Kern-Typ bei den Support-Vektor-Maschinen oder die Anzahl der Knoten und der verdeckten Schichten bei den Neuronalen Netzen.

Betrachtet man die Vielzahl der Kombinationsmöglichkeiten, die sich aus Attribut- und Parameterauswahl ergeben, zeigt sich, dass eine Optimierung in endli-

cher Zeit nicht trivial ist. Aus diesem Grunde werden verschiedene Ansätze verfolgt, um in vertretbarer Zeit möglichst gute Ergebnisse zu finden.

3.2.1 Optimierung der Attributauswahl mittels Genetischer Algorithmen

Ziel einer Optimierung der Attributauswahl muss sein, einen großen Suchraum möglichst schnell und effektiv zu durchschreiten und eine möglichst gute Attributkombination zu finden. Abbildung 3.6 zeigt den schematischen Ablauf einer Optimierung mittels eines *Genetischen Algorithmus (GA)*, wie er in der vorliegenden Arbeit zur Lösung des beschriebenen Problems genutzt wird [93]. Jedes Attribut wird dabei durch ein binäres Gen repräsentiert. Dabei drückt der Wert eines jeden Gens, das so genannte Allel, aus, ob ein Attribut verwendet wird (*1*) oder nicht (*0*). Alle Gene zusammen bilden ein Chromosom.

Der Optimierungsvorgang startet mit der Zusammenstellung einer Reihe zufällig erzeugter *Chromosomen*, die man *Startpopulation* bzw. Generation 0 nennt. Diese Chromosomen werden nach einem vorgegebenen Schema bewertet. Im vorliegenden Fall wird anhand des Root Mean Squared Errors (RMSE) abgeschätzt, wie gut die Prognose der PM₁₀-Konzentrationen bei Verwendung der ausgewählten Attribute ist (auf den der Bewertung zugrundeliegenden Leistungsparameter wird in Kapitel 3.3.2 näher eingegangen). Gemäß des so genannten *Heiratsschemas* wird nun selektiert, welche Elemente einer Generation zur Erzeugung von Nachkommen herangezogen werden sollen. Dabei werden die Chromosomen mit einer zur Güte ihrer Bewertung proportionalen Wahrscheinlichkeit ausgewählt. Somit wird erreicht, dass durch Fortpflanzung vor allem gut bewerteter Chromosomen von Generation zu Generation eine Steigerung der Bewertungsgüte stattfindet. Es werden jedoch auch schlechter bewertete Chromosome weitervererbt, wenn auch mit einer geringeren Wahrscheinlichkeit. Mittels *Rekombination*, dem so genannten *Crossover*, werden nun Nachkommen der aktuellen Generation erzeugt. Um eine zu schnelle Einengung auf ein vielleicht nur lokales Optimum zu vermeiden, werden einige der Chromosomen im nächsten Schritt mutiert. Es werden also einzelne Allele geändert, um eine gewisse Divergenz und Inhomogenität innerhalb der Populationen zu gewährleisten. Die Mutation spielt jedoch gegenüber der Rekombination bei Genetischen Algorithmen eine eher untergeordnete Rolle [93].

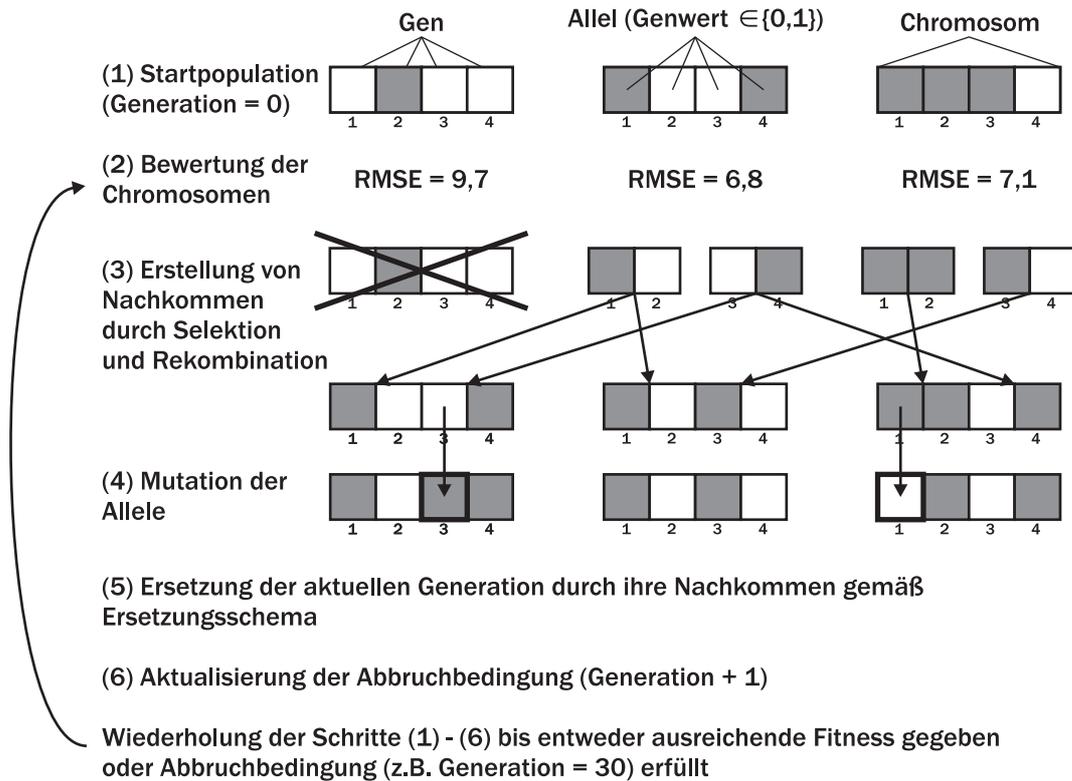


Abbildung 3.6: Schematische Darstellung einer Attributoptimierung mittels Genetischem Algorithmus

Die entstandenen Nachkommen ersetzen gemäß eines gegebenen *Ersetzungsschemas* vollständig oder auch nur teilweise die alte Generation und bilden somit die nächste – die Generationszählung wird um eins hochgesetzt. Ist die Abbruchbedingung, z.B. eine maximale Anzahl von Generationen oder eine ausreichende Fitness (heißt in diesem Fall eine bestimmte Prognosequalität), erreicht, wird die Optimierung beendet. Ist dies nicht der Fall, werden die vorangegangenen Schritte ab der Bewertung wiederholt.

3.2.2 Optimierung der Modellparameter mittels Rasteroptimierung und Evolutionsstrategien

Anders als bei der vorher beschriebenen Optimierung der Attributzusammenstellung, bei der die Verwendung einzelner Attribute durch binäre Gene repräsentiert werden kann, ist dies bei der Optimierung der Modellparameter nicht möglich. Die Parame-

terwerte für die betrachteten Modelle sind entweder reelle oder natürliche Zahlen oder sie lassen sich zumindest durch solche ausdrücken (nominale Parameter).

Ist der Parameterwert eine natürliche Zahl wie z.B. die Anzahl der nächsten Nachbarn k beim IBk-Verfahren, so ist dessen Optimierung relativ einfach. Nach der Vorabbestimmung eines vernünftigen Wertespektrums werden für alle Möglichkeiten die Leistungskennzahlen des Modells berechnet. Die aussichtsreichste Parametereinstellung wird später zur Modellbildung verwendet. Ist mehr als ein natürlichzahliger Parameter zu optimieren, sind alle sich daraus ergebenden möglichen Kombinationen zu überprüfen. Diesen Vorgang nennt man *Rasteroptimierung*. Verwendet wird sie, um z.B. bei einer SVM den bestgeeigneten Typ der verwendeten SVM und des Kerns zu bestimmen.

Der Optimierungsablauf bei reellzahligen Parametern ist der Attributoptimierung mittels Genetischer Algorithmen aus Abschnitt 3.2.1 in vielerlei Hinsicht sehr ähnlich. Man spricht bei der Optimierung reellzahliger Parameter jedoch von einer *Evolutionsstrategie (ES)*. Aus einer zufällig gewählten Startpopulation an Chromosomen, deren Allele in diesem Fall der Menge der reellen Zahlen entstammen, werden durch Bewertung, Selektion, Rekombination und Mutation immer neue Generationen gebildet. Dabei steigt die Fitness der Generationen, also die unter ihrer Verwendung zu erzielende Prognosequalität, an, bis eine Abbruchbedingung erfüllt ist. Stehen bei der Optimierung mittels GA jedoch Selektions- und Rekombinationsmethoden im Vordergrund, so ist bei ES vor allem die Mutation der entscheidende Faktor. Für detailliertere Informationen wird auf weiterführende Literatur verwiesen [93].

3.3 Methodik zur Abschätzung der Leistungsfähigkeit eines Prognosemodells

Ziel dieser Arbeit ist ein Vergleich verschiedener Modelle zur Prognose von PM_{10} -Konzentrationen. Dazu muss die Leistungsfähigkeit der verschiedenen Modelle abgeschätzt werden. Um die Vergleichbarkeit der Ergebnisse gewährleisten zu können, muss bei der Erstellung und Leistungsabschätzung der einzelnen Modelle zum einen immer derselbe Datensatz verwendet werden. Zum anderen müssen für alle Modelle dieselben Leistungskennzahlen berechnet werden. Die folgenden Unterkapitel stellen

die Methodik vor, die die Vergleichbarkeit der Ergebnisse der verschiedenen Modelle und somit deren Leistungsfähigkeit gewährleistet.

3.3.1 Methodik der Kreuzvalidierung zur Qualitätssicherung

Wie bereits in den Abschnitten über die Methoden zur Modellbildung (Kap. 3.1) beschrieben, wird von den zur Verfügung stehenden Daten immer ein Teil der Instanzmenge stichprobenartig zum Training und ein anderer Teil zum Testen des Modells verwendet. Man spricht hierbei von der *Holdout Methode*. Üblicherweise werden etwa ein Drittel für das Testen und die restlichen zwei Drittel für das Training reserviert [121]. Die Problematik dieser Methode liegt darin, dass man nicht genau weiß, ob die Teilmengen repräsentativ sind, ob also die unterschiedlichen Klassen (hier die unterschiedlichen PM_{10} -Konzentrationsbereiche) in beiden Teilmengen im gleichen Verhältnis vorliegen. Im schlimmsten Fall sind im Trainings- oder Testdatensatz Instanzen einer Klasse gar nicht vorhanden. Keines der Modelle kann unter diesen Voraussetzungen korrekte Ergebnisse für Beispiele dieser Klasse liefern. Bei der Auswahl der Test- und Trainingsmenge ist es also wichtig, darauf zu achten, dass jede Klasse in beiden Teilmengen korrekt repräsentiert wird. Um dies zu gewährleisten, wird eine *systematische Stichprobe* in Verbindung mit einer *indirekten Stratifikation* vorgenommen [91]. Dabei wird die Instanzmenge zunächst nach einem entscheidenden Parameter sortiert (hier wiederum die PM_{10} -Konzentration). Darauf folgt eine Auswahl jeder q -ten Instanz des sortierten Datensatzes. Hierbei ist q die Anzahl der zu nehmenden Stichproben.

Da die Kombination aus systematischer Stichprobe und indirekter Stratifikation die Repräsentativität verbessern, aber nicht vollständig gewährleisten kann und somit immer noch eine große Unsicherheit besteht, muss ein weiterer Ansatz zur Gewährleistung der Vergleichbarkeit der Modelle zum Einsatz kommen: die so genannte *Kreuzvalidierung* [121]. Dabei wird die Gesamtmenge der Daten in mehrere annähernd gleiche Teile unterteilt. Nun werden mehrere Modelle mit einem Lernverfahren erzeugt (z.B. mehrere Neuronale Netze), wobei reihum immer eine der Teildatenmengen beim Modelltraining weggelassen wird. Diese dient später zum Testen und zur Leistungsabschätzung des Modells. Nachdem alle Modelle erstellt sind, mittelt man die Leistungskennzahlen der einzelnen Modelle und erhält dadurch eine gute Abschätzung der theoretischen Leistungsfähigkeit des angewandten Lernverfahrens.

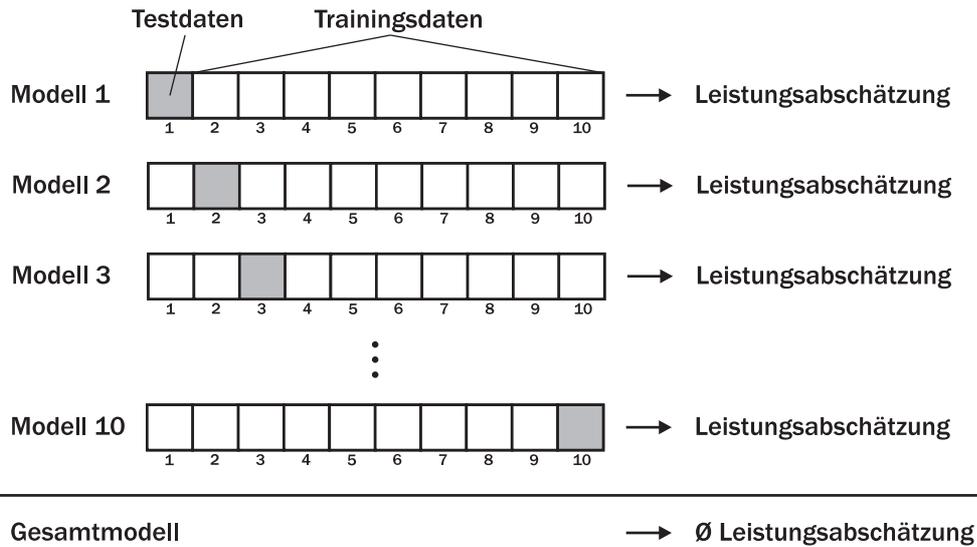


Abbildung 3.7: Schematische Darstellung einer zehnfachen Kreuzvalidierung

Die eventuellen Leistungsschwankungen der einzelnen Modelle, die durch ungenaue Stratifikation entstehen können, werden durch die Mittelung der einzelnen Ergebnisse ausgeglichen. Gerade weil die Datenbasis zur Modellerstellung für diese Arbeit – vor allem was seltene Wetterlagen angeht – relativ klein ist (sie umfasst Messdaten aus etwa 5 Jahren), lassen sich so belastbare Ergebnisse erzielen, weil immer ein größtmöglicher Teil der Daten zum Training herangezogen wird. Dies verbessert die Leistungsfähigkeit des Modells in der Regel.

In der vorliegenden Arbeit soll die Instanzmenge für alle Prognosemethoden in zehn annähernd gleiche Datensätze unterteilt werden. Man kann die Methode also als *zehnfache Kreuzvalidierung* bezeichnen (siehe Abbildung 3.7).

3.3.2 Beschreibung der Kennzahlen und Bewertungsgrundlagen zur Abschätzung der Leistungsfähigkeit von Modellen

Um die Abschätzung der Leistungsfähigkeit der einzelnen Modelle vergleichbar zu machen, müssen sinnvolle Leistungskennzahlen vereinbart werden. Dabei sind einerseits klassische statistische Qualitätsmaße zur Bewertung des Erfolgs der numerischen Vorhersage zu betrachten. Die Auswahl der hier verwendeten basiert auf den Vorschlägen der U.S. Environmental Protection Agency [30] zur Prognose von Ozon und $PM_{2,5}$ und wurde um weitere, für das konkrete Anwendungsgebiet sinnvoll erschein-

nende Leistungskennzahlen ergänzt. Andererseits sollen auch fachgebietsspezifische Bewertungsmöglichkeiten betrachtet werden.

In den folgenden Abschnitten werden die in dieser Arbeit verwendeten Kenngrößen beider Kategorien vorgestellt. Dabei sei p der Prognosewert für die Testinstanzen. Der tatsächliche Wert sei o (von Observation). \bar{p} bzw. \bar{o} seien die Mittel der Prognose- bzw. Messwerte.

Korrelationskoeffizient

Der *Korrelationskoeffizient* r (s. Gleichung 3.4) beschreibt den linearen Zusammenhang zwischen zwei Variablen, hier dem Prognose- und dem gemessenen Wert [91, 42]. Er ist dimensionslos und kann einen Wert zwischen 1 und -1 annehmen, wobei 1 eine perfekte und -1 eine perfekte negative Korrelation beschreibt. Geht der Wert gegen 0, so besteht kein linearer Zusammenhang zwischen den beiden Werten. Bei sinnvollen Prognosemethoden sollte also ein möglichst hoher positiver Wert erzielt werden [121].

$$r = \frac{S_{PO}}{S_P S_O}, \text{ wobei gilt: } S_{PO} = \frac{\sum_{i=1}^n (p_i - \bar{p})(o_i - \bar{o})}{n - 1},$$

$$S_P = \frac{\sum_{i=1}^n (p_i - \bar{p})^2}{n - 1},$$

$$S_O = \frac{\sum_{i=1}^n (o_i - \bar{o})^2}{n - 1} \tag{3.4}$$

Mit dem Korrelationskoeffizienten alleine kann jedoch keine Aussage über die eigentliche Modellgüte getroffen werden, wie von Willmott gut dargestellt wird [119]. Er soll als eines der klassischen Qualitätsmaße dennoch betrachtet werden. Besser geeignet scheinen die im Folgenden vorgestellten Kennzahlen.

Mean Absolute Error (MAE)

Beim *Mean Absolute Error* (*mittlerer absoluter Fehler – MAE*) (s. Gleichung 3.5) behandelt man alle Abweichungen eines Modells unabhängig von ihrem Vorzeichen. Die Beträge der Abweichungen werden aufsummiert und durch die Anzahl der Instanzen geteilt. Sehr große und sehr kleine Abweichungen werden somit gleich gewichtet

und die Dimension des Ergebnisses entspricht der des zu prognostizierenden Wertes [30, 121]. Je kleiner der MAE, desto besser das Prognosemodell.

$$MAE = \frac{1}{n} \left(\sum_{i=1}^n |p_i - o_i| \right) \quad (3.5)$$

Root Mean Squared Error (RMSE)

Der *Root Mean Squared Error (RMSE)* oder auch *Wurzel aus dem mittleren quadratischen Fehler* (s. Gleichung 3.6) ist ein sehr gebräuchliches Maß, um die Abweichung eines geschätzten von dem tatsächlichen Wert zu berechnen. Anders als beim MAE werden die Abweichungen aber vor dem Addieren quadriert. Somit werden im Gegensatz zum MAE große Abweichungen stärker gewichtet als geringfügige [121]. Anschließend wird die Wurzel gezogen, sodass der RMSE dieselbe Dimension besitzt wie der Prognosewert. Dies scheint für das vorliegende Anwendungsgebiet sinnvoll, da große Abweichungen gleichbedeutend mit deutlich verstärktem gesundheitlichem Risiko sein können. Kleine Prognosefehler hingegen bergen dieses Risiko nicht. Außerdem könnten bei starker Überschätzung der PM₁₀-Konzentration teure kurzfristige Minderungsmaßnahmen ausgelöst werden, die zu dem Zeitpunkt unangebracht sind.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - o_i)^2} \quad (3.6)$$

Index of Agreement (IA)

In den letzten Jahren haben Modellierer zunehmend den *Index of Agreement (IA)* nach Willmott als Maß zur Bewertung und zum Vergleich von Modellen verwendet [119, 120]. Er gibt durch einen dimensionslosen Wert an, wie gut die gemessenen Werte vorhergesagt werden konnten. Der Index of Agreement d ist dabei kein Korrelationsmaß, sondern stellt dar, inwieweit die Prognose abweichungsfrei ist. Dabei kann er Werte von 0 bis 1 annehmen, wobei 1 für ein perfektes Modell steht.

$$d = 1 - \frac{\sum_{i=1}^n (p_i - o_i)^2}{\sum_{i=1}^n (|p_i - \bar{o}| + |o_i - \bar{o}|)^2} \quad (3.7)$$

Überschreitung des EU-Grenzwerts für PM₁₀

Handelt es sich bei den bisher vorgestellten Qualitätsmaßen um klassische statistische Kenngrößen, so müssen ebenfalls fragestellungsspezifische Bewertungsgrundlagen festgelegt werden. Als Grundlage dient deshalb der bereits in Kapitel 2.3 ausführlich beschriebene europäische Grenzwert für PM₁₀ von 50 µg/m³ im Tagesmittel. Es ist von besonderer Bedeutung, die Überschreitung dieses Grenzwertes richtig zu prognostizieren. Betrachtet wird deshalb, wie viele Überschreitungen richtig erkannt werden und wie oft fälschlicherweise eine solche vorhergesagt wird. Dazu werden die folgenden zwei Kennzahlen berechnet (zum besseren Verständnis der Formeln ist Tabelle 3.1 zu beachten):

- Die False Alarm Rate (FAR)

Die *False Alarm Rate (FAR)* gibt an, wie viel Prozent der prognostizierten Grenzwertüberschreitungen nicht eingetreten sind – kleinere Werte weisen also auf ein gutes Modell hin [30]. Zur Berechnung dividiert man die Fehlalarme (b) durch die Anzahl aller prognostizierten Grenzwertüberschreitungen ($b+d$). 33% würde bedeuten, dass jede dritte vorhergesagte Überschreitung nicht eingetreten ist.

$$FAR = \frac{b}{(b+d) * 100} \quad (3.8)$$

- Die Probability of Detection (POD)

Mit der *Probability of Detection (POD)* wird berechnet, wie gut ein Modell Grenzwertüberschreitungen prognostizieren kann. Die POD betrachtet dabei nur richtig prognostizierte Überschreitungen (d) im Verhältnis zu tatsächlich gemessenen (c, d) und lässt Fehlalarme außen vor. Eine POD von 66% würde

Tabelle 3.1: Kontingenztafel für eine Zwei-Kategorien-Prognose (Grenzwertüberschreitung "ja"/"nein")

		Prognose	
		nein	ja
Messung	nein	a	b
	ja	c	d

bedeuten, dass zwei Drittel der gemessenen Überschreitungen richtig vorhergesagt wurden.

$$POD = \frac{d}{(c + d) * 100} \quad (3.9)$$

Der deutsche Kurzzeit-Luftqualitätsindex (LuQx)

Da der EU-Grenzwert lediglich eine Schwelle darstellt, die die PM₁₀-Konzentrationen je nach Station eher selten überschreiten, dient als fachspezifische Bewertungsgrundlage zusätzlich der *deutsche Kurzzeit-Luftqualitätsindex (LuQx)* [67, 59]. Er basiert auf dem sechsstufigen deutschen Schulnotensystem und bewertet Schadstoffkonzentrationen u.a. von Feinstaub (PM₁₀)¹ von sehr gut bis sehr schlecht (s. Tab. 3.2). Verbunden mit diesen Bewertungsstufen sind Gesundheitshinweise für die Bevölkerung. Es ist also auch von höchstem Interesse, ob die prognostizierten Werte im richtigen Intervall des LuQx liegen und somit der entsprechende Gesundheitshinweis gegeben werden kann.

Tabelle 3.2: Deutscher Kurzzeit-Luftqualitätsindex (LuQx) für PM₁₀

Indexklasse	Bewertung	PM ₁₀ -Konzentration (24-h-Mittel in µg/m ³)
1	sehr gut	≤ 9,9
2	gut	10,0 – 19,9
3	befriedigend	20,0 – 34,9
4	ausreichend	35,0 – 49,9
5	schlecht	50,0 – 99,9
6	sehr schlecht	≥ 100,0

¹Außerdem werden noch Konzentrationen von Ozon (O₃), Stickstoffdioxid (NO₂), Schwefeldioxid (SO₂) und Kohlenstoffmonoxid (CO) in die Wertung mit einbezogen.

Kapitel 4

Auswahl und Vorbereitung der Eingangsdaten

In der vorliegenden Arbeit werden die Daten der Schadstoffmessstationen der *Landesanstalt für Umweltschutz Baden Württemberg (LUBW)* Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt betrachtet. Bei allen vier Stationen handelt es sich um so genannte *städtische Hintergrundstationen* (siehe Abbildung 4.1). Die Standorte solcher Stationen sind so gewählt, dass sie von Verkehrsemissionen möglichst unbeeinflusst sind und somit die Grundbelastung einer Stadt darstellen. Sie sind also am ehesten repräsentativ für eine größere Fläche, was der Grund für die Auswahl von Hintergrundstationen für die PM_{10} -Prognosen ist. Weitere Stationstypen sind regionale Hintergrundstationen und Stationen an lokalen Quellen (Hot Spots) wie verkehrsreichen Straßen.

Wichtig für die Prognose von PM_{10} mittels maschineller Lernverfahren ist, dass *langjährige Messreihen* von Luftschadstoffen und Meteorologie vorliegen. Dies ist für die genannten vier Stationen der Fall, da an ihnen neben den Luftschadstoffen auch die meteorologische Daten festgehalten werden und es keine Standortveränderung in den letzten Jahren gab. Eine Umstellung der PM_{10} -Messung vom gravimetrischen Messverfahren auf das β -Absorptionsverfahren [41] und die damit verbundene Inkonsistenz in den Datenreihen wurde von der LUBW rechnerisch korrigiert.

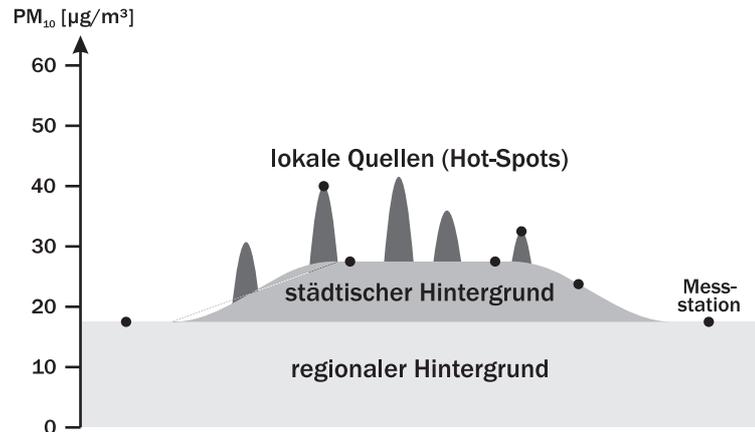


Abbildung 4.1: Darstellung verschiedener Messstationstypen

4.1 Auswahl der zur Prognose geeigneten Attribute

Die zur Prognose geeigneten Attribute müssen grundsätzlich zwei Anforderungen erfüllen:

1. Sie müssen zum Zeitpunkt der Prognose verfügbar sein.
2. Sie müssen in direktem physikalischen Zusammenhang mit der Feinstaubbelastung der Luft stehen. Es sollen also die PM₁₀-Konzentrationen oder zumindest ein Teil davon durch sie erklärbar sein. *Schein-* und *Trivialkorrelationen* müssen ausgeschlossen werden [91].

Die Attribute lassen sich dabei in drei Kategorien einteilen. Zunächst die gemessenen Werte: Dies sind sowohl Meteorologie- als auch Schadstoffmesswerte, die an den vorgestellten Messstationen erhoben werden. Es handelt sich dabei also grundsätzlich um Daten aus der Vergangenheit. Dazu kommen meteorologische Prognosen für die Tage oder Stunden, für die die PM₁₀-Konzentration vorhersagt werden soll. Die dritte Gruppe bilden Daten, mit denen man die Emissionssituation annähern kann. Das können zum Beispiel der Wochentag oder Tagestypen sein. Weitere Erläuterungen hierzu in Abschnitt 4.1.3.

4.1.1 Meteorologiedaten

Niederschlag

In der Literatur gilt Niederschlag als der meteorologische Haupteinflussfaktor auf die Konzentration von Partikeln in der Luft [4, 7, 6, 54]. Dabei wurden die den Veröffentlichungen zugrundeliegenden Studien zum Teil auch speziell für Baden-Württemberg, dem Untersuchungsgebiet dieser Arbeit, durchgeführt.

Niederschläge führen zu einer *Auswaschung* von Staub aus der Luft. Erkennbar ist dies schon bei Betrachtung von Abbildung 4.2, in der die PM_{10} -Konzentration gegen die Niederschlagsmenge aufgetragen ist. Dabei tritt deutlich eine Verschiebung hin zu höheren Immissionen an Tagen ohne Niederschlag hervor. Auch die Menge des Niederschlags spielt an den meisten Stationen eine Rolle. Je mehr Niederschlag fällt, desto geringer sind die Staubkonzentrationen (s. Abb. 4.3). Im Vergleich zum Niederschlagsereignis an sich ist die Menge des Niederschlags jedoch von untergeordneter Bedeutung. An der Station Stuttgart-Bad Cannstatt sinkt die PM_{10} -Konzentration bei geringen Niederschlägen (<1 mm) im Mittel um etwa $9 \mu\text{g}/\text{m}^3$. Der Unterschied zwischen den PM_{10} -Konzentrationen bei sehr kleinen und sehr großen Niederschlagsmengen beträgt lediglich $5 \mu\text{g}/\text{m}^3$.

Neben den täglichen Niederschlägen spielen auch *langanhaltende Niederschlagsereignisse* und Phasen der Trockenheit eine Rolle für die PM_{10} -Konzentrationen – dies wird aus Abbildung 4.4 ersichtlich. An je mehr Tagen hintereinander Niederschläge zu verzeichnen sind, desto geringer ist die durchschnittliche Staubkonzentration. Dies ist damit zu erklären, dass langanhaltende Niederschlagsereignisse am Boden abgelagerte Partikel abspülen und somit deren potenzielle Aufwirbelung verhindern und dass Schwebstaub über mehrere Tage aus der Atmosphäre ausgewaschen wird. An den Stationen Freiburg-Mitte und Karlsruhe-Nordwest tritt dieser Effekt vor allem dann auf, wenn die Niederschläge länger als vier Tage andauern. Für die Stationen Biberach und Stuttgart-Bad Cannstatt ist die beschriebene Auswirkung dagegen nicht zu beobachten. Bei langanhaltenden Trockenperioden steigt hingegen an allen betrachteten Stationen die PM_{10} -Konzentration an. Durch die Austrocknung der Oberflächen und der darauf abgelagerten Partikel können diese leichter aufgewirbelt werden. Außerdem entstehen durch Trockenheit von Böden natürlich auch neue Partikel. Da sowohl längere Phasen der Trockenheit als auch der

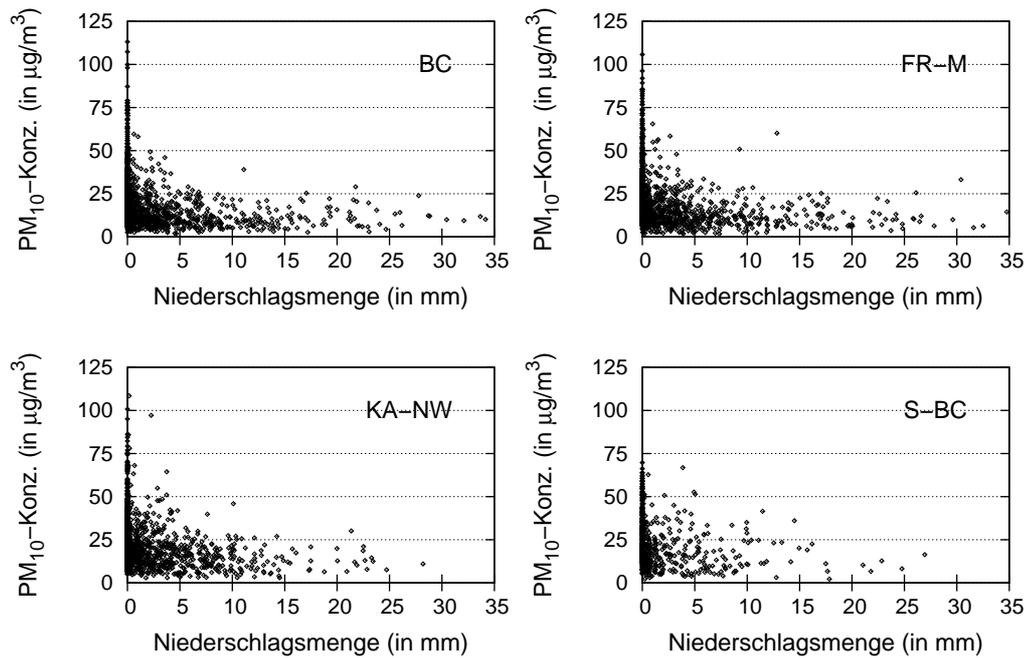


Abbildung 4.2: Gegenüberstellung der PM₁₀-Tagesmittelwerte und der Niederschlagssummen für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Niederschläge selten auftreten, werden in der Abbildung Perioden, die mehr als vier Tage andauern, zusammengefasst.

Der Niederschlag geht in die Prognose über die Attribute tägliche *Niederschlagsmenge* und *Anzahl der aufeinanderfolgenden Tage ohne bzw. mit Niederschlag* ein. Außerdem wird die summierte Niederschlagsmenge für die Zeitfenster 0–6 Uhr, 6–12 Uhr, 12–18 Uhr und 18–24 Uhr berechnet. Damit wird der Tatsache Rechnung getragen, dass Niederschlag, der z.B. erst um 23 Uhr fällt, den Tagesmittelwert der PM₁₀-Konzentration in geringerem Maße abschwächt als Niederschlag, der früher am Tag fällt. Die Auswaschung von Partikeln aus der Luft bleibt zwar uhrzeitunabhängig dieselbe, jedoch gehen weniger Stundenwerte mit geringerer PM₁₀-Konzentration in die Berechnung des Tagesmittelwerts ein.

Wind

Hohe *Windgeschwindigkeiten* bewirken eine rasche *Ausbreitung* von Schadstoffen [43]. Wind sorgt dafür, dass Feinstaubpartikel in der Luft verteilt und damit verdünnt

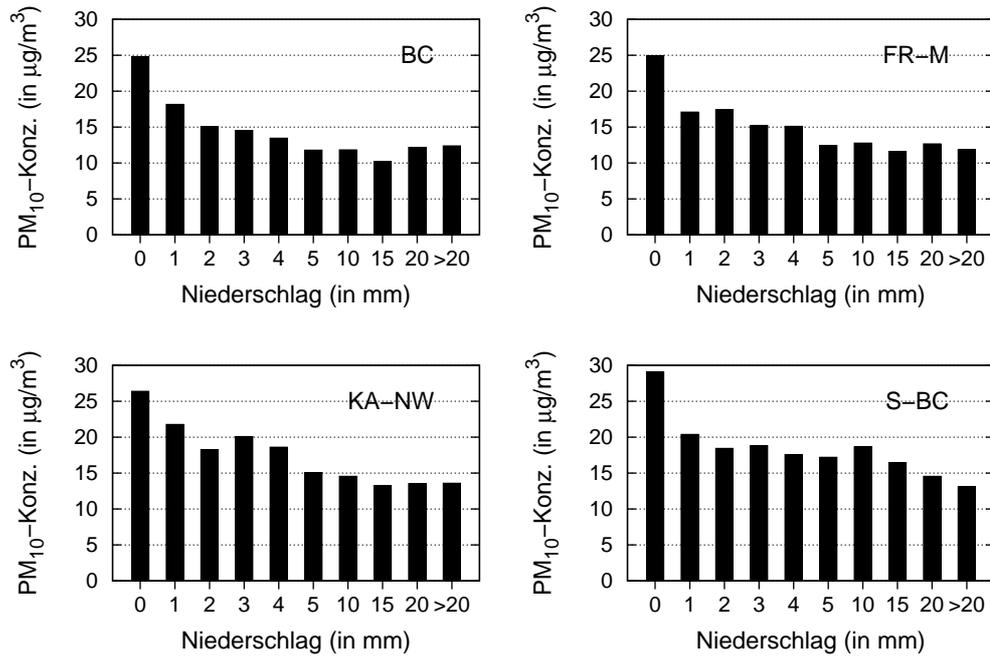


Abbildung 4.3: Durchschnittliche PM₁₀-Tagesmittelwerte abhängig von der Niederschlagsmenge an den Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

werden. Der entscheidende Faktor ist die Windgeschwindigkeit – sie ist in etwa proportional zu ihrer verdünnenden Wirkung [11]. Abbildung 4.5 zeigt, dass an allen Stationen maximale Schadstoffkonzentrationen nur in Verbindung mit geringen Windgeschwindigkeiten auftreten. Bei starkem Wind wird dagegen der PM₁₀-Grenzwert von $50\mu\text{g}/\text{m}^3$ praktisch nicht überschritten. Dabei ist der Windgeschwindigkeitsschwellwert, über dem solch hohe Konzentrationen nicht mehr vorkommen, an jeder Station verschieden.

Neben der Windgeschwindigkeit hat auch die *Windrichtung* Einfluss auf die Luftqualität. Die Ausbreitung von Luftschadstoffen erfolgt in Richtung des Windes. Liegt also eine starke Staubemissionsquelle in der Richtung, aus der der Wind kommt, so sind auch an einer in der Abgasfahne gelegenen Messstation höhere PM₁₀-Konzentrationen zu verzeichnen. Dies haben bereits mehrere Studien für PM₁₀ gezeigt [4, 7, 41, 54, 78]. Besonders deutlich wird dieser Effekt an der Station Karlsruhe-Nordwest bei Winden aus südlichen bzw. südöstlichen Richtungen (s. Abb. 4.6). Aber auch für Freiburg-Mitte ist ein deutlicher Konzentrationsanstieg bei südwestlichen

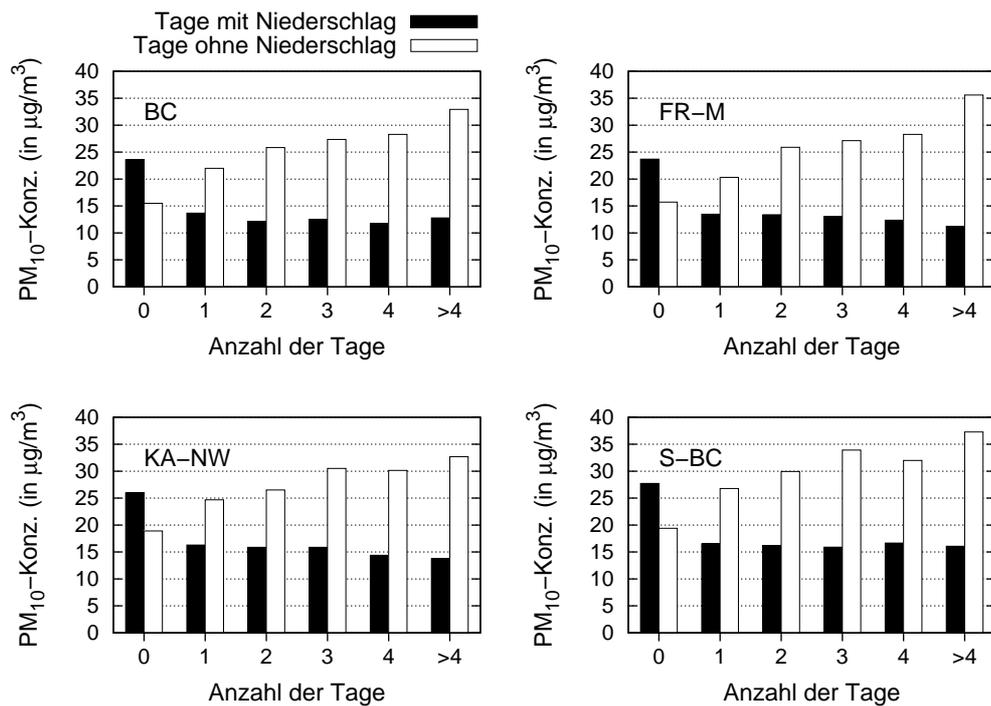


Abbildung 4.4: Durchschnittliche PM₁₀-Tagesmittelwerte bei anhaltenden Niederschlags- und Trockenheitsereignissen für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Winden ersichtlich. An den anderen Stationen bringen Südostwinde die höchsten PM₁₀-Konzentrationen mit sich.

Für die Erstellung der Prognosemodelle kommen die durchschnittliche Windgeschwindigkeit und die Hauptwindrichtung zum Einsatz.

Mischungsschichthöhe

Neben der horizontalen Durchmischung der Atmosphäre durch Wind ist auch die vertikale Durchmischung von besonderer Bedeutung für die Luftbelastung. Sie ist abhängig von der Stabilität der Schichtung im unteren Teil der Troposphäre [43, 11]. Ist die Schichtung nicht stabil, können sich Luftverunreinigungen sehr schnell vertikal ausbreiten und werden so verdünnt. Der *vertikale Austausch* kann jedoch auch durch *Sperrschichten* unterbunden werden. Sie sind durch einen Anstieg der Lufttemperatur mit zunehmender Höhe – einer so genannten *Inversion* – gekennzeichnet, was dem normalen Bild eines Temperaturabfalls mit steigender Höhe widerspricht. Solche Sperrschichten verlangsamen oder unterbinden sogar den vertikalen Austausch von Luftmassen. Den Bereich unterhalb der Inversionsschicht bezeichnet man als Mi-

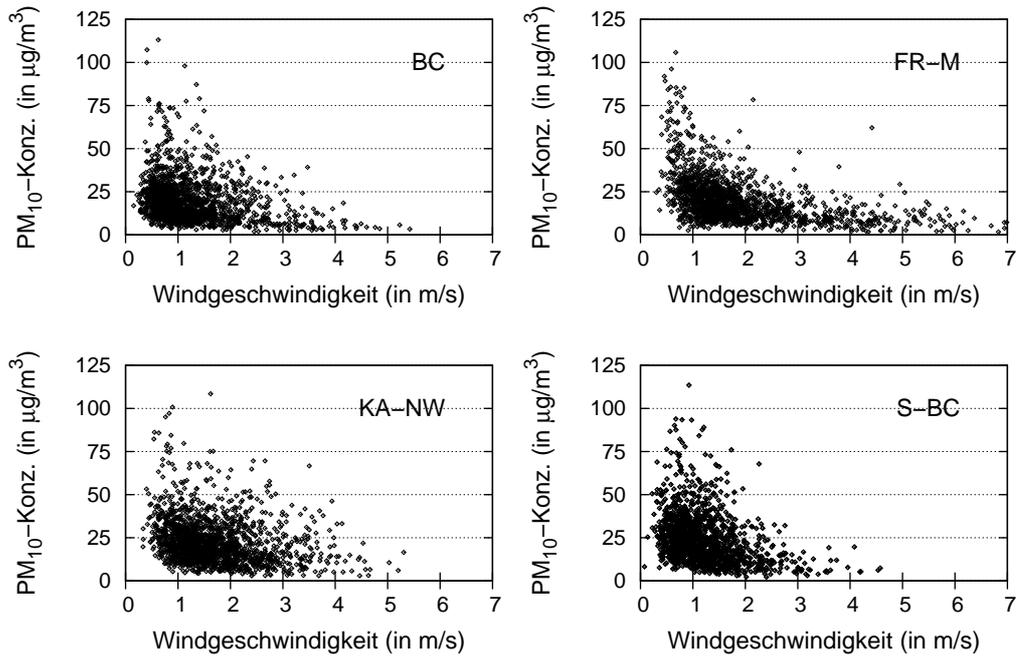


Abbildung 4.5: Gegenüberstellung der PM₁₀-Tagesmittelwerte und der durchschnittlichen Windgeschwindigkeit für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

schungsschicht, da sich nur in dieser Schicht am Boden freigesetzte Luftschadstoffe frei ausbreiten können. Ist diese Mischungsschicht nur sehr dünn, ist auch das Volumen, in dem sich ein Schadstoff ausbreiten kann, gering. Es findet also weniger Verdünnung statt, als bei einer dicken Mischungsschicht [107]. Anke et al. [7] haben in einer Studie festgestellt, dass in Baden-Württemberg die Mischungsschicht für die PM₁₀-Konzentration von Bedeutung ist, wenn auch in geringerem Ausmaß als die horizontale Ausbreitung durch den Wind.

Um die vertikale Durchmischung zu quantifizieren, findet für die Modellbildung die *absolute Höhe der Mischungsschicht* (H_{mix}) Verwendung. Sie wird jedoch nicht flächendeckend gemessen, sondern nur modelliert. Dazu wurden die Vertikalprofile der Lufttemperatur aus dem Lokal-Modell (LM) des Deutschen Wetterdienstes (DWD) für Baden-Württemberg auf 1-h-Basis verwendet [94]. Die Gitternetzweite beläuft sich auf etwa 3.5 km in der Horizontalen. In der Vertikalen standen die untersten 16 Modelllevels zur Verfügung, welche die unteren 2200 m der Atmosphäre abdecken. Die Luftmessstationen wurden jeweils dem nächstgelegenen Gitternetzpunkt zugeordnet.

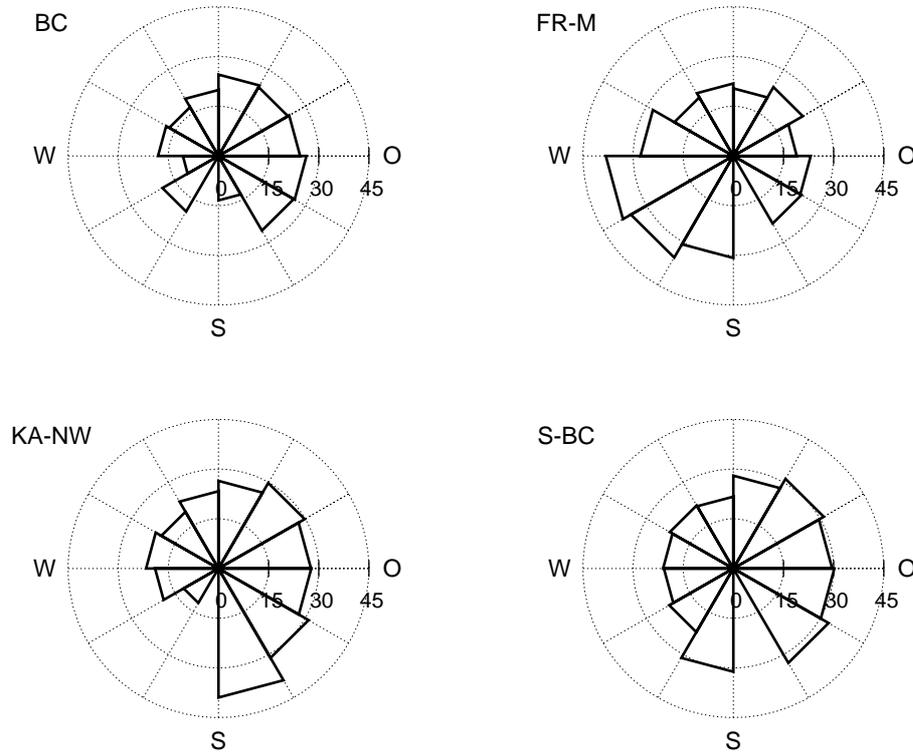


Abbildung 4.6: Tägliche durchschnittliche PM_{10} -Konzentration nach Windrichtungssektoren für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Hierbei ergab sich aufgrund der horizontalen Auflösung des LM ein räumlicher Fehler von maximal 1.8 km.

Für die Berechnung der Mischungsschichthöhe wurden die Vertikalprofile der potenziellen Temperatur, von der Bodenschicht ausgehend, nach Inversionsschichten durchsucht. Um mögliche Messfehler und direkte Einflüsse der Oberfläche auszuschließen, wurden die Bodenwerte der Vertikalprofile verworfen. Für das automatisierte Erkennen einer Inversion wurden drei ähnliche Parametrisierungsverfahren aus der Literatur verwendet [66, 82], die sich jedoch in den gewählten Grenzwerten für den Vertikalgradienten der potenziellen Temperatur und der Inversionsstärke unterscheiden. Als Mischungsschichthöhe wurde anschließend der Mittelwert aus den nach den drei Verfahren bestimmten Einzelwerten von H_{mix} definiert. Diese Berechnungen der Mischungsschichthöhen wurden vom Meteorologischen Institut der Universität Freiburg durchgeführt und zur Verfügung gestellt.

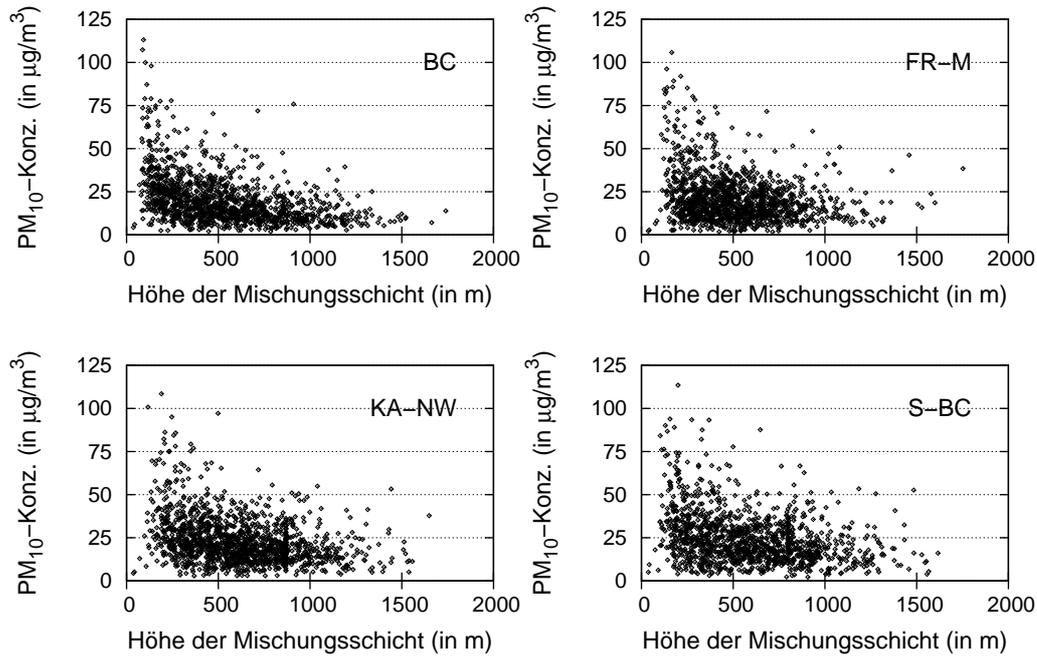


Abbildung 4.7: Gegenüberstellung der PM₁₀-Tagesmittelwerte und der durchschnittlichen Höhe der Mischungsschicht für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Neben der absoluten Mischungsschichthöhe werden auch so genannte Mischungsschicht-Stufen berechnet, die binär codiert sind. Sie geben an, ob die Mischungsschicht unterhalb einer gewissen Grenze liegt oder nicht. Der vertikale Bereich wurde dabei von 0 bis 500 m gewählt, wobei der Abstand der einzelnen Stufen 50 m beträgt.

Relative Luftfeuchtigkeit und Globalstrahlung

Aus einer Studie in Baden-Württemberg [5] geht hervor, dass oftmals Perioden hoher relativer Luftfeuchtigkeit und niedriger Globalstrahlung mit hohen PM₁₀-Konzentrationen einhergehen. Niedrige Globalstrahlung verbunden mit hoher Luftfeuchtigkeit ist Indikator einer starken Nebelbildung. Dies wiederum weist auf eine schlechte Luftaustauschsituation hin, was eine Verdünnung der Schadstoffkonzentrationen in der Luft verhindert. Hohe PM₁₀-Konzentrationen können also nicht nur, wie bereits beschrieben, durch anhaltende Trockenheit entstehen, sondern auch durch extrem feuchte Wetterlagen verursacht werden. Dies ist auch aus der Abbildung 4.8

ersichtlich. Für die Station Freiburg-Mitte scheint es sogar, dass an Tagen mit hoher Luftfeuchtigkeit häufiger extrem hohe PM_{10} -Konzentrationen auftreten als an weniger feuchten Tagen.

Um die durch diesen Effekt verursachten Perioden hoher PM_{10} -Konzentrationen prognostizieren zu können, werden die tägliche durchschnittliche Globalstrahlung und die relative Luftfeuchtigkeit betrachtet. Solche Prozesse, die von mehreren meteorologischen Parametern abhängig sind, sollten von den maschinellen Lernverfahren deutlich besser erkannt werden als von der Linearen Regression. Diese ist nur in der Lage, lineare Zusammenhänge zwischen einzelnen Attributen und der PM_{10} -Konzentration zu erkennen.

An den betrachteten Stationen wird die relative Luftfeuchtigkeit nicht erfasst. Sie lässt sich jedoch mithilfe der Magnusformel aus der Temperatur und dem Taupunkt berechnen (vereinfacht dargestellt in Formel 4.1).

$$r(T, T_{T_{aopt.}}) = 100 * SDD(T_{T_{aopt.}}) / SDD(T), \text{ wobei gilt :} \\ SDD(T) = 6.1078 * 10^{((7.5 * T) / (237.3 + T))} \quad (4.1)$$

Temperatur

Auch die Lufttemperatur hat einen Einfluss auf die PM_{10} -Konzentration. Da austauscharme Situationen wie Inversion oder starker Nebel immer mit niedrigen Temperaturen einhergehen, kann die Tagesmitteltemperatur einen Hinweis geben. Hohe PM_{10} -Konzentrationen kommen in Baden-Württemberg zumeist im Winter vor. Deutlich geht dies aus Abbildung 4.9 hervor – die höchsten PM_{10} -Konzentrationen entstehen nur bei Tagesmitteltemperaturen unter $5^{\circ}C$. Die Temperatur kann also auch als Indikator für die kalte Jahreszeit dienen.

Neben der Tagesmitteltemperatur werden auch Tagesminimum, -maximum und die Differenz zwischen diesen Temperaturwerten bereitgestellt. Es gilt zu untersuchen, ob eines dieser Attribute oder eine Kombination daraus die oben genannten Situationen besser beschreiben.

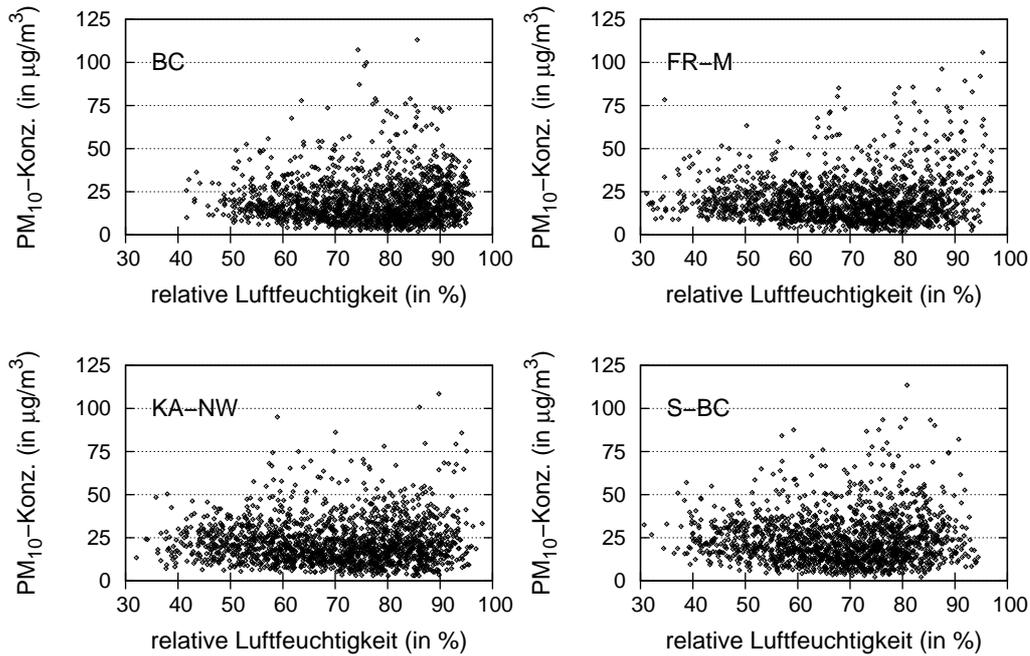


Abbildung 4.8: Gegenüberstellung der PM₁₀-Tagesmittelwerte und der durchschnittlichen relativen Luftfeuchtigkeit für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

4.1.2 Schadstoffkonzentrationen

PM₁₀-Konzentrationen

Die PM₁₀-Konzentrationen der vergangenen Tage sind selbstverständlich auch ein entscheidender Faktor bei der Prognose. Sie geben Aufschluss, von welchem Niveau aus sich diese Werte für die kommenden Stunden und Tage entwickeln. Wird z.B. eine niedrige PM₁₀-Konzentration am Vortag gemessen, so hat ein Niederschlagsereignis am Prognosetag weit geringere Auswirkungen als bei Messung hoher Werte.

NO₂-Konzentrationen

Stickstoffoxide sind mitunter entscheidende Vorläufersubstanzen bei der Bildung sekundärer Aerosole (siehe Kapitel 2.1). Außerdem stammt NO₂ teilweise aus denselben Quellen wie PM₁₀ (z.B. Dieselfahrzeuge). Aus diesem Grunde sollen die NO₂-Konzentrationen des Vortags für die Prognose der aktuellen PM₁₀-Konzentrationen herangezogen werden. Aus der Abbildung 4.10 geht deutlich dieser Zusammenhang

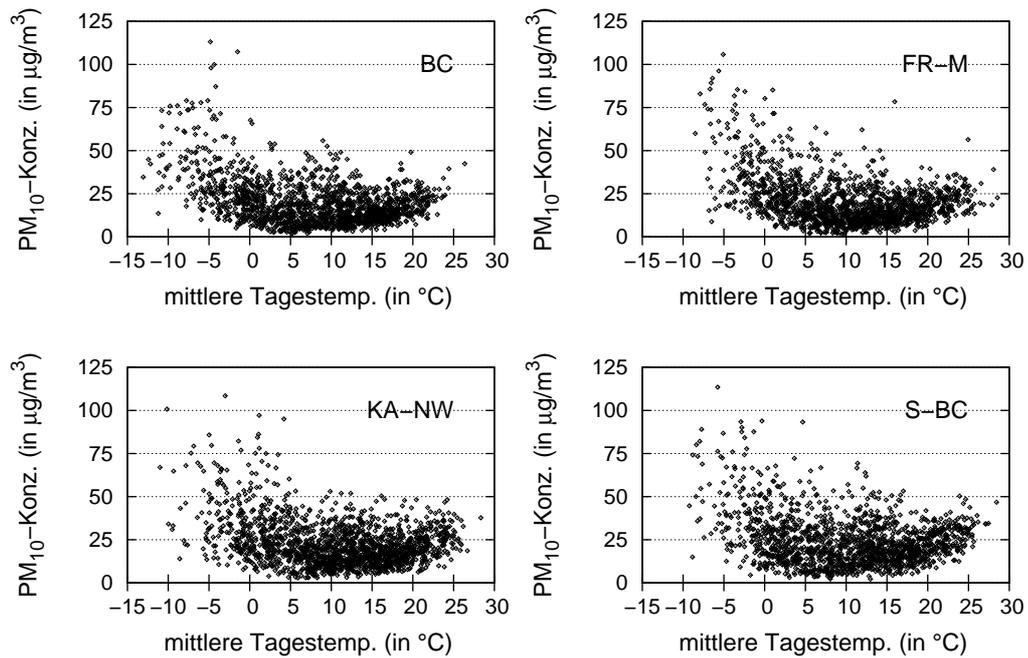


Abbildung 4.9: Gegenüberstellung der PM₁₀-Tagesmittelwerte und der mittleren Tagestemperatur für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

hervor – besonders offensichtlich für die Stationen Freiburg-Mitte und Karlsruhe-Nordwest. Werden am Vortag hohe Stickstoffdioxidkonzentrationen gemessen, so ist auch für den aktuellen Tag mit einer hohen Feinstaubbelastung der Luft zu rechnen. Maximalwerte der PM₁₀-Konzentration kommen ohne hohe Konzentrationen der Vorläufersubstanz NO₂ am Vortag praktisch nicht vor. Natürlich kann eine hohe Konzentration am Vortag auch auf eine länger andauernde austauscharme Situation hinweisen. Es muss sich also zeigen, ob dieses Attribut zur Verbesserung der PM₁₀-Prognose beitragen kann oder nicht.

4.1.3 Attribute zur Emissionsannäherung

Neben den meteorologischen Kennzahlen ist auch die *Emissionssituation* in der Umgebung einer Station für die Feinstaubkonzentration relevant – wenn auch in weit geringerem Maße, wie Untersuchungen u.a. von Anke et al. [4, 6] gezeigt haben. Für die vorliegende Prognose stehen jedoch weder ein zeitlich hochauflösendes Emissionskataster noch Verkehrszählungen zur Verfügung. Aus diesem Grund muss ein Ansatz

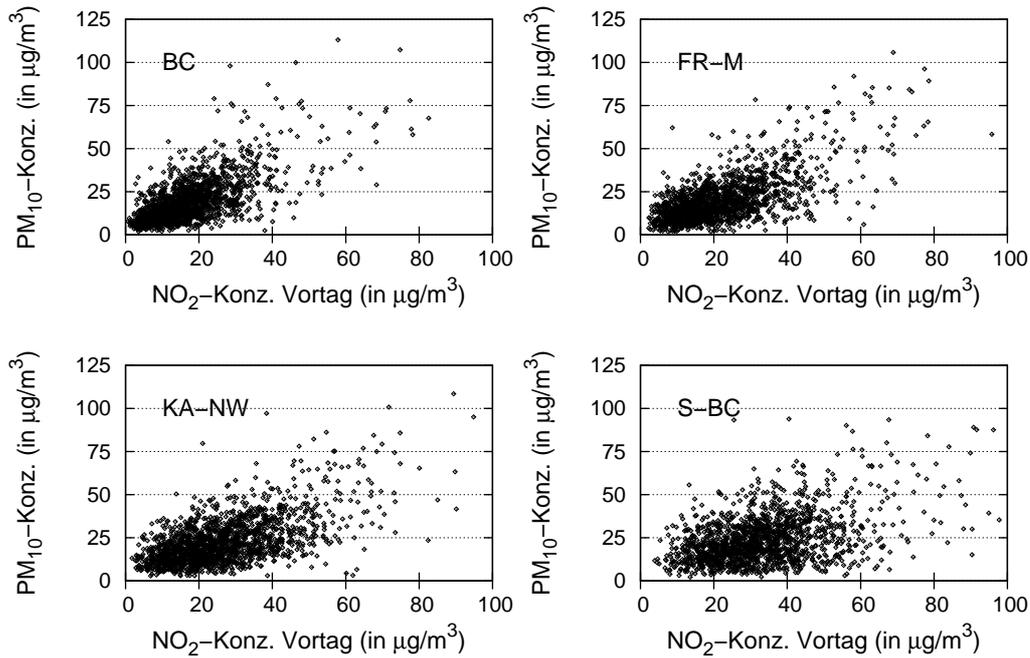


Abbildung 4.10: Gegenüberstellung der aktuellen PM₁₀- und der NO₂-Tagesmittelwerte des Vortags für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

gefunden werden, um die relevanten Emissionen wie Industrieaktivität und Straßenverkehr annähern zu können.

Dazu wird der Wochengang der PM₁₀-Konzentration analysiert (s. Abb. 4.11). Deutlich fällt dabei die Absenkung der Staubkonzentrationen am Samstag und vor allem am Sonntag auf. Sie ist auf das abgeschwächte Verkehrsaufkommen und die verminderten Industrieprozesse am Wochenende zurückzuführen. Klingner et al. [54] zeigt dies auch sehr deutlich am Beispiel von Stationen in Dresden.

Um diese Überlegungen nun in die geplanten Modelle mit einfließen zu lassen, werden zwei zusätzliche Parameter aus den Daten berechnet. Zum einen wird eine ID für die Wochentage eingeführt, wobei Feiertage eine eigene ID erhalten. Der zweite Parameter ist für die bloße Unterscheidung zwischen Werktag, Samstag und Sonntag bzw. Feiertag vorgesehen. Die späteren Untersuchungen werden zeigen, inwiefern sich diese Parameter zur Emissionsannäherung für die Prognose von PM₁₀ eignen.

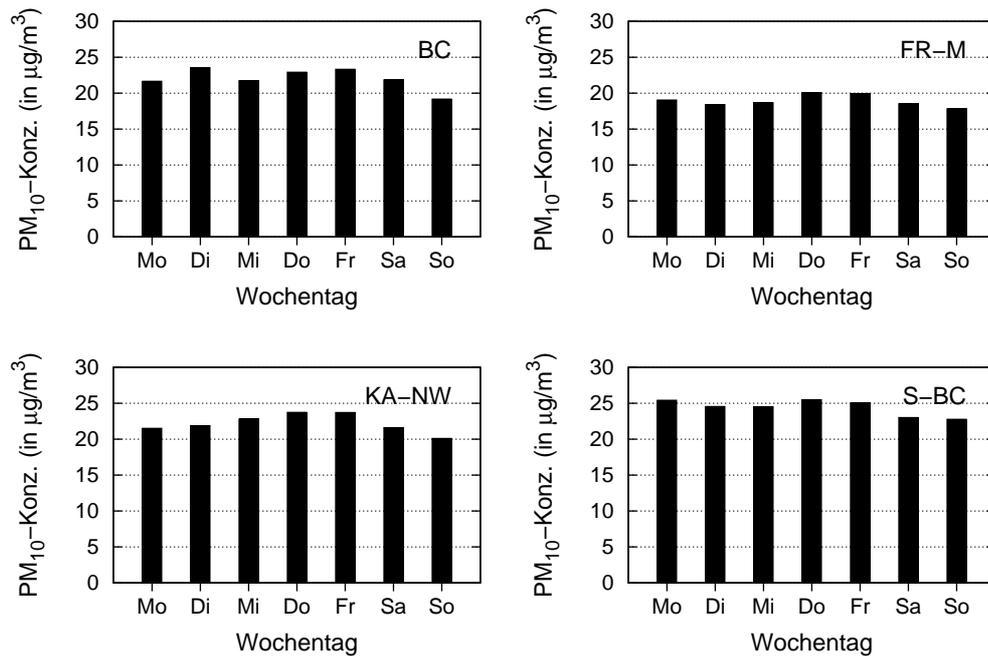


Abbildung 4.11: Durchschnittliche PM₁₀-Konzentrationen nach Wochentag für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

4.2 Vorbereitung der Eingangsdaten (Preprocessing)

Zur Erstellung von PM₁₀-Vorhersagemodellen ist die Qualität der Eingangsdaten von entscheidender Wichtigkeit. Oftmals treten jedoch Werte auf, die sich von den anderen Werten der Datenmenge abheben, so genannte *Ausreißer*. Dabei kann es sich zum einen um Mess-, Rechen- oder Schreibfehler handeln [42, 91], zum anderen können Ausreißer auch durch besondere Ereignisse verursacht sein. Ein gutes Beispiel hierfür ist das mit den Neujahrsfeiern verbundene Feuerwerk am 1. Januar gegen Mitternacht. Durch die Feuerwerkskörper werden in kürzester Zeit große Partikelmengen an die Luft abgegeben. Etwas zeitversetzt lässt sich dies als Spitze in der Konzentrationskurve feststellen, wie Abbildung 4.12 zeigt. Im Vergleich dazu verläuft der Tagesgang an Sonn- und Feiertagen im Januar (ohne Neujahr) relativ eben. Ein anderes Beispiel wäre eine temporäre Baustelle in der Nähe einer Messstation, die über einen gewissen Zeitraum erhöhte PM₁₀-Konzentrationen verursacht.

Die Beseitigung der Ausreißer im Eingangsdatensatz wird in mehreren Schritten durchgeführt [80]:

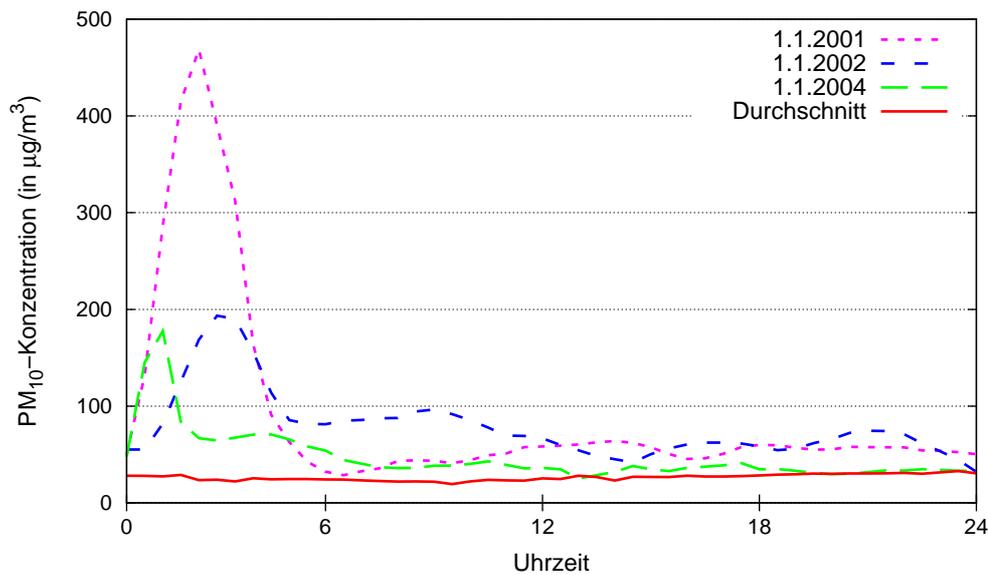


Abbildung 4.12: Vergleich des Tagesgangs der PM₁₀-Konzentration an Neujahrstagen und des durchschnittlichen Tagesgangs an Sonn- und Feiertagen im Januar (ohne Neujahr) für die Station Karlsruhe-Nordwest

- Im ersten Schritt werden *unmögliche Werte*, wie negative Schadstoffkonzentrationen oder Windrichtungen $> 360^\circ$ entfernt. Dazu werden die höchsten und niedrigsten Werte eines jeden Attributs überprüft.
- Als zweite Maßnahme folgt eine *visuelle Kontrolle* der Tagesgänge. Dabei werden u.a. grobe Messfehler oder der Anstieg der PM₁₀-Konzentrationen an Neujahr ersichtlich.
- Im dritten Schritt werden die Daten auf *starke Schwankungen* aufeinanderfolgender Werte untersucht. Als ausreißerverdächtig gelten dabei Differenzen, die größer als der Mittelwert der geprüften Variable sind. Als verdächtig identifizierte Schwankungen werden einer gesonderten visuellen Prüfung unterzogen und bei Bedarf eliminiert.
- Im vierten und letzten Schritt werden Tage, an denen nicht alle Attribute und die PM₁₀-Konzentration vorhanden sind, von der Modellerstellung ausgeschlossen.

Kapitel 5

Erstellung und Optimierung von PM₁₀-Prognosemodellen und Abschätzung ihrer Leistung

Die Erstellung von Modellen zur Prognose von PM₁₀-Konzentrationen erfolgt in drei aufeinanderfolgenden Schritten, die für jedes der verwendeten Lernverfahren grundsätzlich identisch sind: Zunächst werden die *Parametereinstellungen* des Modells optimiert. Basierend auf diesen optimierten Modellparametern wird anschließend die beste *Attributzusammenstellung* für dieses Modell ermittelt. Im letzten Schritt findet das *Modelltraining* statt. Mittels trainierter Modelle kann schließlich die eigentliche Prognose durchgeführt und deren Ergebnis zur Leistungsabschätzung herangezogen werden.

Im Folgenden werden diese einzelnen Schritte vorgestellt. Zum besseren Verständnis erfolgt dies in umgekehrter chronologischer Reihenfolge, da das Modelltraining und die Leistungsabschätzung integraler Bestandteil sowohl der Parameter- als auch der Attributoptimierung sind. Ebenso ist eine Attributoptimierung Teil der Parameteroptimierung. Alle Trainings- und Optimierungsschritte werden mit der Software YALE [68] durchgeführt. Lediglich beim Multilayer-Perzeptron (MLP) kommt zusätzlich die Software Qnet2000 [114] zum Einsatz.

Für die vorliegende Arbeit standen leider keine historischen meteorologischen Prognosedaten zur Verfügung. Aus diesem Grund wurden zur Leistungsabschätzung

der Modelle anstelle prognostischer Daten historische Messwerte verwendet. Da meteorologische Vorhersagen mit einer gewissen Unsicherheit behaftet sind, ist zu erwarten, dass die Leistung der Modelle durch den Einsatz von Messdaten anstelle von Prognosen besser sein wird. Aus diesem Grund schließt sich der Modellentwicklung eine kritische Auseinandersetzung mit Sensitivitäten bezüglich ungenauer meteorologischer Prognosedaten an (siehe Kap. 5.4).

5.1 Training und Leistungsabschätzung der Modelle

Als Training bezeichnet man den Vorgang, bei dem ein Modell mittels eines Lernverfahrens aus einem Trainingsdatensatz Zusammenhänge zwischen der PM₁₀-Konzentration und den vorgestellten Attributen erlernt. Das Ergebnis dieses Lernprozesses ist ein Modell, welches für unbekannte Daten eine PM₁₀-Prognose liefern kann.

Beim Training kann man die verwendeten Lernverfahren in zwei Gruppen teilen: Zur ersten Gruppe gehört die multiple lineare Regression (LR), das instanzbasierte Lernen mit den k -nächsten-Nachbarn (IBk) und die Support-Vektor-Maschine (SVM). Wenn man basierend auf einem dieser Verfahren ein Modell mit einem Datensatz trainiert, so ist das Ergebnis auch bei mehrfacher Wiederholung immer dasselbe. Der Vorteil dieser Verfahren ist, dass man manuell nicht in den Trainingsprozess eingreifen muss.

Anders dagegen beim Multilayer-Perzeptron (MLP). Wenn man ein Modell basierend auf diesem Verfahren mit demselben Datensatz mehrfach trainiert, erhält man als Ergebnis unterschiedliche Modelle mit unterschiedlicher Prognoseleistung. Der Grund hierfür ist die Zufallsbelegung der Gewichte beim Start des Backpropagation-Algorithmus. Von diesem zufällig gewählten Punkt wird nun für eine Fehlerfunktion ein Minimum gesucht. Es ist aber nicht gesagt, dass das absolute Minimum gefunden wird. Häufig endet der Lernprozess auch in einem lokalen Minimum. In der Praxis bedeutet das, dass man mit jedem Trainingsdatensatz mehrere MLPs erstellt, das mit der besten Prognoseleistung manuell auswählt und für die spätere Prognose nutzt.

Ein weiterer manueller Schritt beim Erstellen eines MLPs ist der Abbruch des Trainings, sobald es Hinweise auf ein Overfitting gibt. Beim Overfitting lernt das Netz ab einer gewissen Anzahl von Trainingsiterationen die Trainingsdaten auswendig, statt

die Zusammenhänge zwischen den Attributen und der PM₁₀-Konzentration abzubilden. Ein so übertrainiertes Netz liefert schlechtere Prognoseergebnisse als ein Netz, bei dem das Training rechtzeitig abgebrochen wurde. Um den richtigen Zeitpunkt für den Trainingsabbruch zu finden, wird ein kleiner Teil der Trainingsdaten beim Training außen vor gelassen. Nach jedem Trainingsdurchlauf wird für diese Testdaten eine Prognose vorgenommen. Solange der Root Mean Squared Error (RMSE), also die Wurzel aus dem mittleren quadratischen Fehler, für die Prognoseergebnisse sinkt, wird weitertrainiert, steigt er an, sollte das Training abgebrochen werden. In Abbildung 5.1 wird der RMSE gegen die Zahl der Trainingsdurchgänge aufgetragen und der optimale Abbruchzeitpunkt markiert. Zunächst fällt der RMSE sehr stark, das Modell verbessert sich durch die ersten Iterationen also deutlich. Ab ca. 1000 Iterationen wird das Gefälle der Fehlerkurve geringer und sie findet ihr absolutes Minimum bei etwa 7500 Iterationen. Danach fällt die Prognoseleistung ab, denn der RMSE wird wieder größer. Da ab diesem Zeitpunkt von Overfitting ausgegangen werden muss, wird das Training bei besagten 7500 Durchläufen abgebrochen.

In der vorliegenden Arbeit wird das Modelltraining immer in Kombination mit einer Prognose und einer Leistungsabschätzung durchgeführt, da es ja zunächst nicht darum geht, ein Modell für den operationellen Betrieb zu erstellen, sondern die Leistung verschiedener Modelle zu vergleichen und somit das bestgeeignete Modell zu finden. Generell ist jedoch die Vorgehensweise für die spätere Erstellung eines operationellen Modells dieselbe wie oben beschrieben. Vereinfacht ist dies in Abbildung 5.2 dargestellt. Demnach werden alternierend nach dem Schema der zehnfachen Kreuzvalidierung $\frac{9}{10}$ der zur Verfügung stehenden Daten zum Modelltraining genutzt. Mittels des so trainierten Modells (beim MLP das beste aus den erstellten Modellen) wird nun für das verbliebende $\frac{1}{10}$ der Daten eine Prognose angestellt. Deren Ergebnisse werden mit den gemessenen PM₁₀-Konzentrationen verglichen und die Leistungskennzahlen berechnet (siehe Abschnitt 3.3). Sind für alle zehn Datenteilmengen diese Kennzahlen ermittelt, so werden aus ihnen die jeweiligen Mittelwerte und somit die Abschätzung der Gesamtleistung eines Modells für eine Station ermittelt.

Wie wichtig diese Kreuzvalidierung ist, zeigt Tabelle 5.1. Hier sind die Leistungskennzahlen für vier Teildatenmengen aufgetragen. Obwohl die einzelnen Teildatenmengen bereits stratifiziert sind, die PM₁₀-Konzentrationsverteilung in allen Teilen also gleichmäßig ist, sind die Unterschiede enorm. Würde man zur Leistungs-

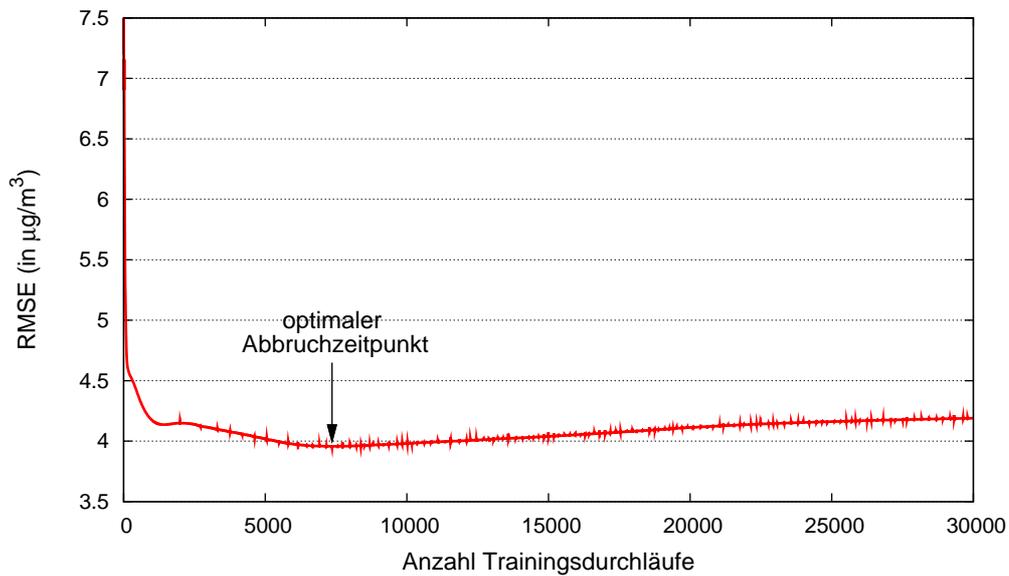


Abbildung 5.1: Entwicklung des RMSE der Testdaten beim Training eines MLPs (Station Biberach)

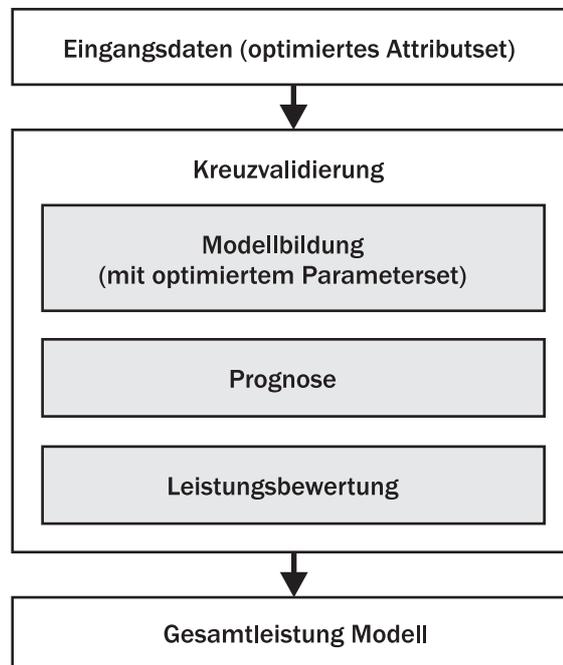


Abbildung 5.2: Schematische Darstellung des Modelltrainings

Tabelle 5.1: Leistungskennzahlen des MLP für vier verschiedene Datenteilmengen (Station Stuttgart-Bad Cannstatt)

		1	2	3	4
Korrelationskoeffizient (r)		0,88	0,87	0,84	0,84
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	5,3	5,2	6,0	6,2
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	6,9	7,4	8,1	8,1
Index of Agreement (IA)		0,93	0,93	0,91	0,91
False Alarm Rate (FAR)	%	33	40	0	29
Probability of Detection (POD)	%	57	43	50	63
LuQx richtig prognostiziert	%	58	65	58	62

abschätzung eines Modells eine zufällig gewählte und somit unstratifizierte Teilmenge wählen, wären die Abweichungen vorraussichtlich noch größer. Käme nicht einmal eine Kreuzvalidierung zum Einsatz, wäre das Ergebnis nicht belastbar und somit mit Ergebnissen anderer Modelle auch nicht vergleichbar.

5.2 Optimierung der Attributzusammenstellung

Eine Möglichkeit, die Prognosequalität von Modellen zu verbessern, ist die Optimierung der Attributzusammenstellung, da nicht die Verwendung aller Attribute, sondern eines optimalen Attributsets zu den besten Prognoseergebnissen führt. Die grundlegende Vorgehensweise ist dabei für alle Lernverfahren identisch. Sie ist in Abbildung 5.3 dargestellt. Mittels eines Genetischen Algorithmus (GA) wird die Attributzusammenstellung immer wieder verändert (siehe Kapitel 3.2.1). Mit jedem dieser Attributsets wird nachfolgend ein Modelltraining, verbunden mit Prognose und Leistungsabschätzung durchgeführt, wie im vorigen Abschnitt erläutert. Die optimale Attributkombination ist durch die beste Prognosequalität, also den geringsten RMSE gekennzeichnet und wird zur endgültigen Modellbildung verwendet.

Ein Nachteil des GA ist, dass unklar bleibt, ob eine Lösung das Optimum darstellt oder nur eine gute Annäherung bietet. Aus diesem Grund wird die Attributoptimierung mehrfach durchgeführt. Es zeigt sich jedoch, dass die Ergebnisse für diese Aufgabenstellung fast immer identisch sind oder zumindest nur sehr geringfügig variieren.

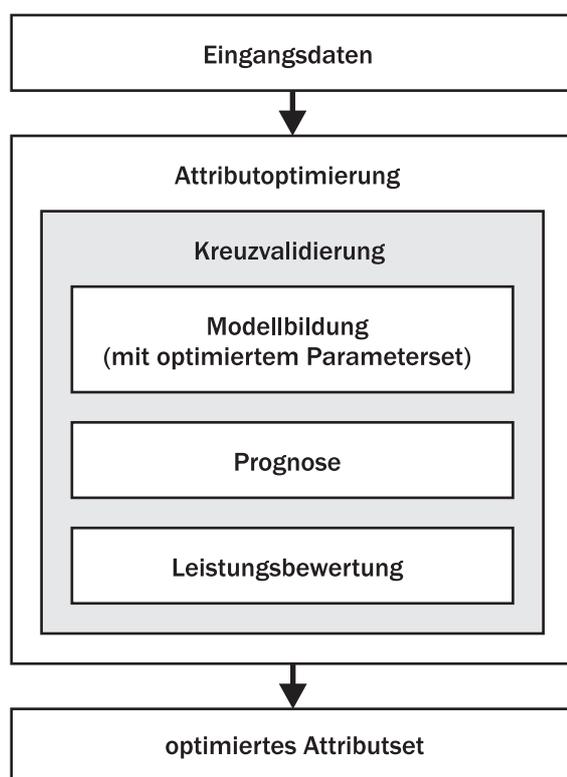


Abbildung 5.3: Schematische Darstellung einer Attributoptimierung

Die Verbesserungen bei der Prognosegenauigkeit, die durch die Attributoptimierung erzielt werden, seien am Beispiel des Einsatzes einer SVM für die Station Karlsruhe-Nordwest exemplarisch erläutert:

Abbildung 5.4 zeigt deutlich, dass die Verwendung aller Attribute nicht die besten Prognoseergebnisse liefert. Die Prognosetendenz weicht in diesem Fall teilweise deutlich von der Tendenz der gemessenen PM₁₀-Konzentrationen ab (z.B. um den 04.09.2002). Sobald die absolute Mischungsschichthöhe nicht mehr als Attribut herangezogen wird, verbessert sich die Prognosequalität. Die SVM scheint also dieses Attribut nicht gewinnbringend nutzen zu können. Gleiches gilt für das Attribut Wochentag. Wird es nicht genutzt sieht man z.B. eine Verbesserung um den 20.9.2002. Schritt für Schritt ergibt sich so durch das Weglassen der von der SVM nicht gewinnbringend nutzbaren Attribute eine Verbesserung der Prognoseergebnisse, die je nach Attribut mal größer, mal kleiner ausfällt.

Um den Eindruck, der durch die Betrachtung des kurzen Abschnitts im August und September 2002 entsteht, zu untermauern, liefert Tabelle 5.2 eine objektivere Sicht der Dinge. Die beschriebene Abweichung der Prognosetendenz von den gemessenen PM₁₀-Konzentrationen bei Verwendung der absoluten Mischungsschichthöhe manifestiert sich in einem stark ansteigenden Korrelationskoeffizienten beim Weglassen dieses Attributs. Ebenso lässt sich beim weiteren Optimieren der Attributzusammenstellung ein stetig steigender IA beobachten, der die Gesamtleistung des Modells betrachtet.

Deutlich wird auch, dass die Minimierung des RMSE nicht immer die Optimierung aller Leistungskennzahlen bedeutet. Gerade im Fall der False Alarm Rate (FAR – fälschlicherweise prognostizierte EU-Grenzwertüberschreitungen) und der Probability of Detection (POD – richtig prognostizierte EU-Grenzwertüberschreitungen) zeigen andere Attributkombinationen geringfügig bessere Ergebnisse. Ein minimaler RMSE stellt aber zumeist einen guten Mittelweg für alle Leistungskennzahlen dar.

Die Optimierung der Attributzusammenstellung läuft wie erwähnt für die anderen verwendeten Lernverfahren identisch ab. Dies bedeutet allerdings nicht, dass die Ergebnisse dieselben sind. Für die Station Karlsruhe-Nordwest gibt Tabelle 5.2 Aufschluss über die jeweils beste Attributkombination für die verschiedenen Lernverfahren. Interessant ist dabei, dass es einige Attribute gibt, die von allen Lernverfahren zur PM₁₀-Prognose genutzt werden. Das sind z.B. die Niederschlagssumme oder die

KAPITEL 5. ERSTELLUNG UND OPTIMIERUNG VON PM₁₀-PROGNOSEMODELLEN UND ABSCHÄTZUNG IHRER LEISTUNG

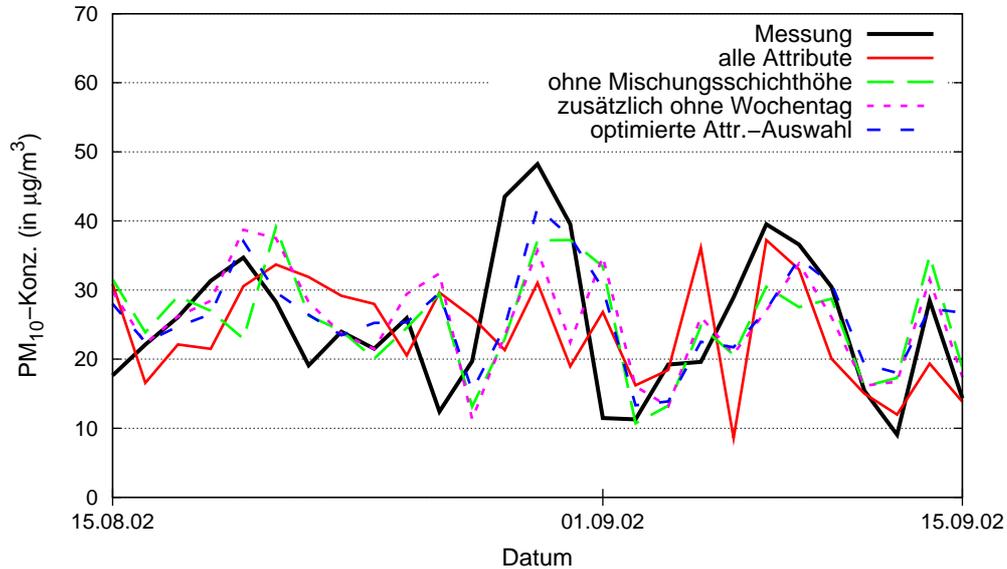


Abbildung 5.4: Vergleich verschiedener Attributkombinationen bei der Prognose von PM₁₀-Tagesmittelwerten mittels SVM (Station Karlsruhe-Nordwest)

Tabelle 5.2: Vergleich der Leistungskennzahlen bei der Prognose von PM₁₀-Tagesmittelwerten mittels SVM mit verschiedenen Attributkombinationen (Station Karlsruhe-Nordwest)

		Prognose nutzt alle Attribute	P. ohne Mischungs- schichthöhe	P. zusätzlich ohne Wochentag	P. mit opt. Attribut- kombination
Korrelationskoeffizient (r)		0,66	0,78	0,81	0,84
Mean Absolute Error (MAE)	µg/m ³	7,9	6,4	6,3	5,5
Root Mean Squared Error (RMSE)	µg/m ³	10,4	8,8	8,3	7,6
Index of Agreement (IA)		0,79	0,88	0,89	0,91
False Alarm Rate (FAR)	%	17	29	24	20
Probability of Detection (POD)	%	56	67	71	70
LuQx richtig prognostiziert	%	48	56	57	62

PM₁₀-Konzentration des Vortags. Andere Attribute werden nur von einzelnen oder wenigen Verfahren genutzt, wie z.B. die NO₂-Konzentration des Vortags, die nur das MLP zur Vorhersage verwendet, oder das Temperaturmaximum, welches die LR und das IBk nutzen.

Die Ergebnisse der Attributoptimierung für eine Station lassen sich anders als die der Parameteroptimierung nicht auf andere Stationen übertragen, d.h. sie muss für jede Station aufs neue durchgeführt werden (detaillierte Ergebnisse für jede Station finden sich in Anhang B). Eine Zusammenfassung aller Ergebnisse bietet Tabelle 5.2. Deutlich treten mehrere Muster hervor:

Es gibt eine Gruppe von Attributen, die an allen vier Stationen und von allen Lernverfahren zur PM₁₀-Prognose genutzt werden. Dazu zählen die Windgeschwindigkeit, die Niederschlagssumme und die PM₁₀-Konzentration des Vortags. Dieses Muster deckt sich mit den Aussagen der in Abschnitt 4.1 erwähnten Studien, nach denen Wind und Niederschlag die Haupteinflussfaktoren für den PM₁₀-Gehalt in der Luft sind.

Eine weitere Gruppe von Attributen wird vom jeweiligen Lernverfahren an allen Stationen verwendet. Die Anzahl dieser „Hauptattribute“ variiert von drei beim IBk, über sechs bei der SVM bis hin zu acht bei der LR und neun beim MLP. Alle übrigen Attribute werden nur hin und wieder von den Lernverfahren herangezogen. Besonders hoch ist die Fluktuation bei den Mischungsschichthöhenstufen – keine dieser Stufen wurde von einem Lernverfahren mehr als zwei Mal verwendet.

5.3 Optimierung der Modellparameter

Die Optimierung der Modellparameter stellt neben der Attributoptimierung eine weitere Möglichkeit dar, die Prognoseleistung eines Modells zu verbessern. Ziel dieser Optimierung ist es, ein Set aus Modellparametern zu finden, mit dem ein Modell bei der PM₁₀-Prognose die besten Ergebnisse erzielt. Das bedeutet in diesem Fall, dass der RMSE möglichst klein wird. Dazu kommen wie in Abschnitt 3.2.2 beschrieben die Rasteroptimierung bzw. Evolutionsstrategien zum Einsatz.

Die Parameteroptimierung stellt den zeitaufwändigsten Teil bei der Modellerstellung dar. Der Grund hierfür wird klar, wenn man den schematischen Aufbau

KAPITEL 5. ERSTELLUNG UND OPTIMIERUNG VON
 PM₁₀-PROGNOSEMODELLEN UND ABSCHÄTZUNG IHRER LEISTUNG

Tabelle 5.3: Optimierte Attributzusammenstellung nach Modellen für die Station Karlsruhe-Nordwest

	LR	IBk	MLP	SVM
Tagtyp	●	—	●	●
Windrichtung Ø	●	—	—	—
Windgeschwindigkeit Ø	●	●	●	●
Taupunkttemperatur Ø	●	—	●	●
Temperatur Ø	●	●	●	●
Temperatur-Maximum	●	●	—	—
Temperatur-Minimum	—	—	●	—
Δ Temp.-Max./Min.	—	—	●	—
Niederschlagssumme	●	●	●	●
PM ₁₀ Vortag Ø	●	●	●	●
NO ₂ Vortag Ø	—	—	●	—
Tage ohne Niederschlag	●	—	●	●
Tage mit Niederschlag	—	●	●	●
relative Luftfeuchtigkeit Ø	●	●	●	—
Mischungsschicht < 150m	●	—	—	—
Mischungsschicht < 300m	●	—	—	●
Mischungsschicht < 350m	—	—	—	●
Mischungsschicht < 400m	—	●	—	—
Mischungsschicht < 450m	—	—	●	—
Mischungsschicht < 500m	●	—	—	●

● verwendetes Attribut
 — nicht verwendetes Attribut

KAPITEL 5. ERSTELLUNG UND OPTIMIERUNG VON
 PM₁₀-PROGNOSEMODELLEN UND ABSCHÄTZUNG IHRER LEISTUNG

Tabelle 5.4: Zusammenfassung der Verwendungshäufigkeit einzelner Attribute nach Modellen an allen Stationen

	LR	IBk	MLP	SVM
Wochentag	—	—	—	1
Wochentag gewichtet	—	—	1	—
Tagtyp	4	—	4	3
Windrichtung Ø	4	1	—	—
Windgeschwindigkeit Ø	4	4	4	4
Taupunkttemperatur Ø	2	—	4	4
Temperatur Ø	4	2	4	4
Temperatur-Maximum	4	2	2	1
Temperatur-Minimum	—	2	3	1
Δ Temp.-Max./Min.	—	—	2	2
Globalstrahlung	1	—	1	—
Niederschlagssumme	4	4	4	4
PM ₁₀ Vortag Ø	4	4	4	4
NO ₂ Vortag Ø	1	1	4	—
Tage ohne Niederschlag	4	1	4	4
Tage mit Niederschlag	—	2	4	1
relative Luftfeuchtigkeit Ø	1	1	3	—
Mischungsschichthöhe Ø	2	1	—	—
Mischungsschichthöhe Ø <150m	2	—	—	—
Mischungsschichthöhe Ø <200m	1	—	—	—
Mischungsschichthöhe Ø <250m	—	—	1	—
Mischungsschichthöhe Ø <300m	2	—	—	1
Mischungsschichthöhe Ø <350m	—	—	—	2
Mischungsschichthöhe Ø <400m	—	1	—	1
Mischungsschichthöhe Ø <450m	—	—	1	—
Mischungsschichthöhe Ø <500m	1	—	—	1
insgesamt verwandte Attribute	47	26	50	38

einer Parameteroptimierung betrachtet (s. Abb. 5.5), der für alle verwendeten Lernverfahren identisch ist.

Da die optimale Parametereinstellung abhängig von der Attributzusammensetzung ist, muss für jeden zu prüfenden Parametersatz während der Parameteroptimierung immer auch eine Attributoptimierung stattfinden. So wird die beste Attributzusammenstellung für die gegebenen Parameter bestimmt. Bei der Attributoptimierung wiederum muss für jedes zu überprüfende Attributset eine zehnfache Kreuzvalidierung durchgeführt werden. Das optimierte Parameterset, also das mit der besten Leistungsbewertung, wird schließlich als Ausgabe gespeichert. Vereinfacht stellt sich der Ablauf der Parameteroptimierung wie folgt dar:

1. Eingangsdatensatz wird gelesen
2. „Start“-Parameterset wird gesetzt
3. „Start“-Attributset wird gesetzt
4. 1.–10. Modell wird mittels Kreuzvalidierung alternierend mit $\frac{9}{10}$ der Eingangsdaten trainiert und auf das restliche $\frac{1}{10}$ angewendet. Die Prognoseresultate werden bewertet.
⇒ Durchschnitt ergibt Gesamtleistung des Modells bei gegebenem Parameter- und Attributset
5. Attributset wird variiert und Schritt 4 wiederholt, bis Attributset optimal für Parametereinstellungen (beste Gesamtleistung)
6. Parameterset wird variiert und Schritt 4 und 5 wiederholt, bis Parameterset optimal (beste Gesamtleistung)
7. optimales Parameterset wird gespeichert

Bereits die Attributoptimierung inklusive der Kreuzvalidierung benötigt einige tausend Modellläufe. Dieser Vorgang muss nun für jeden getesteten Parametersatz durchgeführt werden, die Anzahl der Modellläufe für die Attributoptimierung multipliziert sich also mit der Anzahl der getesteten Parametersätze. Aus diesem Grund muss bei der Parameteroptimierung sehr überlegt vorgegangen werden, um den zeitlichen Rahmen nicht zu sprengen.

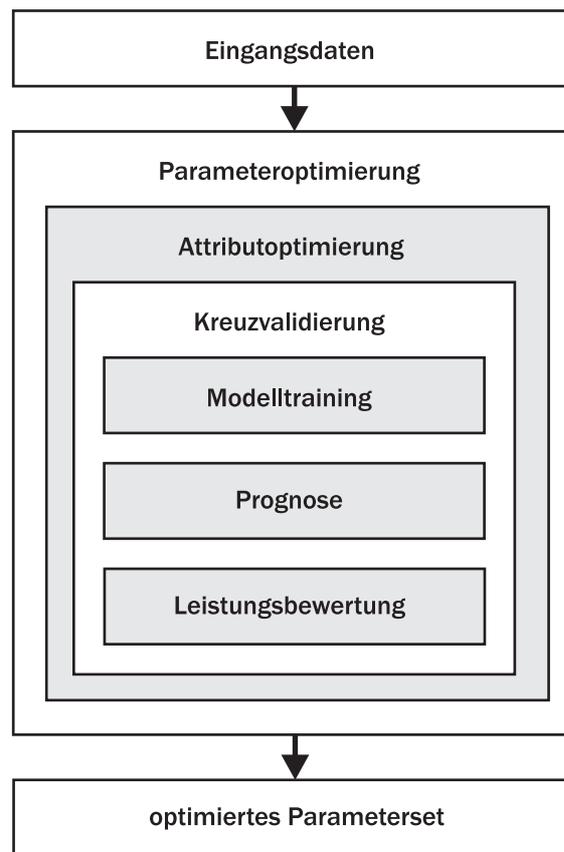


Abbildung 5.5: Schematische Darstellung einer Parameteroptimierung

Für die verschiedenen Lernverfahren müssen unterschiedliche Parameter optimiert werden: Gilt es für das IBk und das MLP lediglich einen bzw. drei ganzzahlige Parameter zu optimieren, so sind es für die SVM bis zu sechs zum Teil realzahlige. Die sich daraus ergebenden Kombinationsmöglichkeiten sind unendlich. Bei der LR sind keine Parameter zu optimieren.

5.3.1 Parameteroptimierung beim instanzbasierten Lernen mit den k -nächsten-Nachbarn im Detail

Beim IBk soll der Parameter k , der die Anzahl der nächsten Nachbarn angibt, bestimmt werden. Dazu wird das Wertespektrum für die Rasteroptimierung spezifiziert. Nach den bei der PM₁₀-Prognose gemachten Erfahrungen liegt das Optimum etwa zwischen 9 und 11 (s. Tab. 5.5). Da der Rechenaufwand beim IBk vergleichsweise gering ist, kann man zur Sicherheit jedoch auch einen größeren Wertebereich eingeben.

Die Auswirkungen der Optimierung des Parameters k lassen sich aus Abbildung 5.6 ersehen. Es ist deutlich zu erkennen, dass der Kurvenverlauf bei $k=1$ deutlich unruhiger ist als bei höheren Werten von k . Dies bestätigt auch der schlechte Korrelationskoeffizient in Tabelle 5.3.1. Grund hierfür ist, dass nur der nächste Nachbar zur Prognose herangezogen wird. Einzelne historische PM₁₀-Konzentrationen, die unter Umständen nicht repräsentativ sind, nehmen hier einen großen Einfluss. Anders verhält es sich, wenn der Mittelwert der 5, 9 oder 15 nächsten Nachbarn als Prognose gewählt wird. Gleichzeitig stellt man aber fest, dass eben diese Mittelwertbildung bei den höchsten Konzentrationen auch von Nachteil sein kann. Die hohen Konzentrationen werden mit $k=1$ am besten abgebildet, da sie in den Trainingsdaten selten vorkommen. Tritt z.B. eine hohe Konzentration, verbunden mit einer in der Vergangenheit nur einmal aufgetretenen meteorologischen Situation erneut ein, so werden neben diesem einen sehr ähnlichen Fall unweigerlich auch vier weniger ähnliche Fälle zur Prognose herangezogen. Dies verschlechtert die Prognose.

Insgesamt zeigt Tabelle 5.3.1 am Beispiel der Station Stuttgart-Bad Cannstatt jedoch deutlich, dass $k=9$ in fast allen Belangen überlegen ist. Die bereits erwähnte bessere POD bei $k=1$ wird mit einer deutlich schlechteren FAR erkauft und weist somit darauf hin, dass dieses Modell insgesamt eher überschätzt.

5.3.2 Parameteroptimierung für ein Multilayer-Perzeptron im Detail

Für das MLP müssen die Anzahl der Zwischenschichten und Zwischenschichtneuronen sowie die Art der Aktivierungsfunktion angepasst werden. Dabei sind die ersten beiden Parameter natürliche Zahlen, der letzte Parameter ist nominal. Auch hier kann

Tabelle 5.5: Ergebnisse der Modellparameteroptimierung für Instanzbasiertes Lernen mit den k -nächsten-Nachbarn

Station	k
Biberach	9
Freiburg-Mitte	10
Karlsruhe-Nordwest	11
Stuttgart-Bad Cannstatt	9

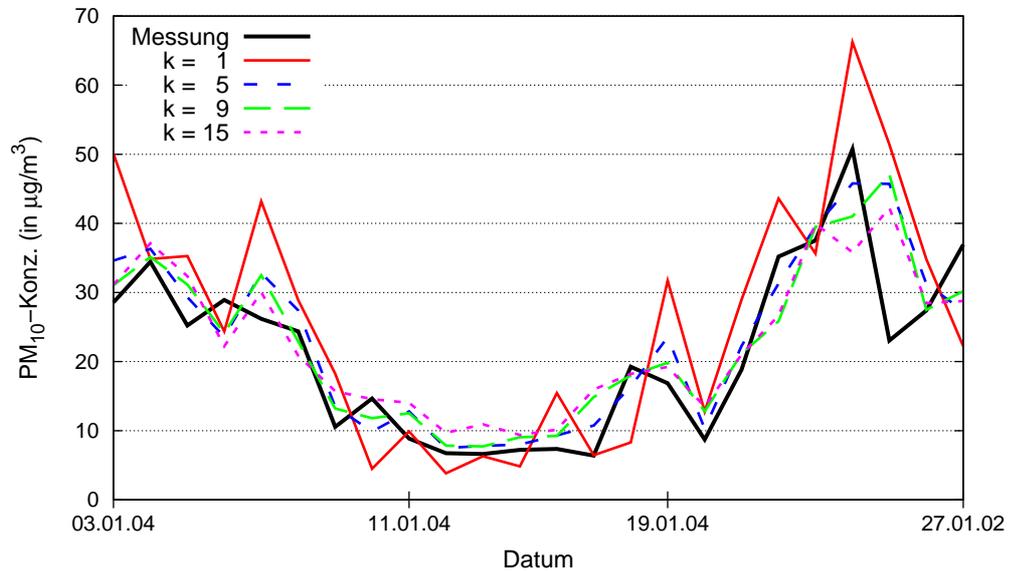


Abbildung 5.6: Vergleich verschiedener Parametereinstellungen (k) bei der Prognose von PM₁₀-Tagesmittelwerten mittels IBk (Station Stuttgart-Bad Cannstatt)

Tabelle 5.6: Vergleich der Leistungskennzahlen bei der Prognose von PM₁₀-Tagesmittelwerten mittels IBk mit unterschiedlichem k (Station Stuttgart-Bad Cannstatt)

		$k=1$	$k=5$	$k=9$	$k=15$
Korrelationskoeffizient (r)		0,72	0,82	0,83	0,82
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	8,1	6,3	6,2	6,4
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	11,0	8,6	8,4	8,6
Index of Agreement (IA)		0,85	0,89	0,89	0,88
False Alarm Rate (FAR)	%	43	27	23	24
Probability of Detection (POD)	%	55	49	47	41
LuQx richtig prognostiziert	%	49	59	60	58

die Rasteroptimierung eingesetzt werden. Es hat sich gezeigt, dass sich zur Prognose von PM₁₀ vor allem eine sigmoide Ausgabefunktion eignet. Sie wird deshalb für alle Stationen eingesetzt. Ebenso kommt für alle Stationen dieselbe Netzwerktopologie mit einer Zwischenschicht und fünf Zwischenschichtneuronen zum Einsatz. Die Unterschiede bei der Verwendung einer geringfügig höheren oder niedrigeren Anzahl von Neuronen sind vernachlässigbar.

Im Gegensatz zu den anderen Lernverfahren kommt nach den bereits beschriebenen Optimierungsmethoden beim MLP noch eine weitere Parameteroptimierung mittels der *Learn Rate Control (LRC)* von Qnet2000 zum Einsatz [114]. Sie optimiert dynamisch während des Trainingsprozesses den Parameter η – die *Lernrate*. η bestimmt, wie groß die Änderungen der Gewichte beim Backward-Pass des Back-Propagation-Algorithmus. Ein hohes η beschleunigt den Lernprozess, jedoch u.U. auf Kosten der Modellgüte. Ein kleines η lässt das Modell langsam lernen. Die Learn Rate Control optimiert die Lernrate in Abhängigkeit vom Zustand des Lernprozesses und verbessert so die Modellqualität und die Geschwindigkeit des Trainingsprozesses.

5.3.3 Parameteroptimierung für eine Support-Vektor-Maschine im Detail

Die Parameteroptimierung für eine SVM stellt die größte Herausforderung dar, da es sowohl realzahlige als auch numerische Parameter zu bestimmen gibt. Aus diesem Grund muss man sich an die optimale Parametereinstellung in zwei Stufen herantasten:

In der ersten Stufe gilt es, mit der Rasteroptimierung die beiden numerischen Parameter, den Typ der SVM und die Kern-Funktion festzulegen. Dazu werden alle

Tabelle 5.7: Ergebnisse der Modellparameteroptimierung für ein Multilayer-Perzeptron

Parameter	Einstellung
Aktivierungsfunktion	sigmoid
Zwischenschichten	1
Zwischenschichtneuronen	5
Lernrate (η)	dynamisch (LRC)

möglichen Kombinationen aus SVM-Typen (3 Varianten), Kern-Funktionen (5 Varianten) berechnet und die Modelleistung bewertet – der Werteraum für die realzahligen Parameter wird dabei in groben Intervallen durchschritten. Hsu [49] schlägt für die Intervalle z.B. $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ und $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$ vor. Hat man auf diesem Wege die numerischen Parameter bestimmt, werden in einem zweiten Schritt die realzahligen mittels Evolutionsstrategie optimiert.

Bei der Optimierung der Modellparameter für eine SVM haben sich vor allem die Parameter SVM-Typ, die Kern-Funktion und die Werte für den Kostenfaktor C und γ als entscheidend herausgestellt. Der Parameter nu hat deutlich geringeren Einfluss auf das Prognoseergebnis – beim RMSE liegen die Auswirkungen einer Veränderung dieses Parameters lediglich im Hundertstelbereich.

Beim Vergleich der Ergebnisse der Parameteroptimierung für eine SVM fällt auf, dass an allen Stationen ähnliche Ergebnisse erzielt werden. Aus diesem Grund wird eine Standard-Parametereinstellung für alle Stationen gewählt, die aus Tabelle 5.8 zu entnehmen ist.

5.4 Sensitivitätsanalyse bezüglich unsicherer meteorologischer Vorhersagen

Den größten Unsicherheitsfaktor bei der Prognose von PM₁₀-Konzentrationen stellt der Einsatz von meteorologischen Vorhersagen als Attribut dar. In den vorigen Abschnitten wurden für die Erstellung und Leistungsabschätzung der Prognosemodelle gemessene historische meteorologische Daten eingesetzt. Somit wurde die potenzielle

Tabelle 5.8: Ergebnisse der Modellparameteroptimierung für eine Support-Vektor-Maschine

Parameter	Einstellung
SVM-Typ	nu-SVR
Kern-Funktion	rbf
Kostenparameter (C)	250
γ	0,001
nu	0,4
Toleranzmarge (ε)	entfällt bei diesem SVM-Typ

Leistungsfähigkeit eines Modells bewertet, da die Meteorologie für den Prognosetag gemessen und somit immer richtig war – sozusagen eine perfekte meteorologische Vorhersage. Dies entspricht jedoch nicht der Realität bei der PM₁₀-Prognose für den nächsten Tag. Vielmehr müssen hier modellierte meteorologische Vorhersagen zum Einsatz kommen, die jedoch mit einer großen Unsicherheit belegt sind.

Für die betrachteten Messstationen liegen leider keine historischen Meteorologievorhersagen vor, die zur Abschätzung der tatsächlichen Prognoseleistung herangezogen werden können. Auch lässt sich keine generelle Aussage treffen, wie groß die Abweichungen zwischen meteorologischen Vorhersagen und der tatsächlich gemessenen Meteorologie sind. Dies ist in großem Maße von der Lage der Station und dem verwendeten meteorologischen Prognosemodell, das die Daten berechnet, abhängig.

Jacob [50] spricht in einer Vergleichsstudie mehrerer großskaliger Modelle von mittleren Abweichungen bei der täglichen Niederschlagsmenge von etwa 25-50%. Aus einem Vergleich gemessener und mit dem lokalen Modell des DWD COSMO-EU prognostizierter Meteorologiedaten an der Wetterstation Stuttgart-Echterdingen für das Jahr 2006 ergeben sich RMSE von 1,7-2,3° für die Temperatur und von 1,7 m/s für die Windgeschwindigkeit (siehe Abbildungen 5.7 und 5.8). Ebenso kann man der Abbildung entnehmen, dass jedes dritte bis vierte vorhergesagte Niederschlagsereignis nicht eintritt. Sehr selten hingegen kam es zu Niederschlägen, ohne dass sie vorhergesagt wurden. Ähnliche Werte finden sich in der Arbeit von Hoffmann [46] für Stationen in Sachsen.

Basierend auf diesen Zahlen, die nicht zwingend repräsentativ für die in dieser Arbeit betrachteten LUWB-Stationen sein müssen, wurde ein einfaches Schema entwickelt, um abzuschätzen, wie sich die Modelle bei fehlerhaften Meteorologievorhersagen verhalten. Wichtig ist dabei neben der absoluten Veränderung der Prognoseleistung vor allem auch die Tendenz.

Ausgehend von den gemessenen historischen Meteorologiedaten werden drei Datensätze mit Abweichungen von bis zu $\pm 10\%$, $\pm 25\%$ und $\pm 50\%$ erstellt. Um dies zu erreichen wird jedes Meteorologieattribut mit einer Zufallszahl zwischen 0,9 – 1,1, 0,75 – 1,25 bzw. 0,5 – 1,5 multipliziert. Einzige Ausnahme bildet die Niederschlagsmenge. Neben der bereits beschriebenen Methode Abweichungen zu simulieren wird jede fünfte, vierte bzw. dritte Niederschlagsmenge auf null gesetzt. Dies wirkt sich

dann auch auf die Attribute aus, die die Tage ohne bzw. mit Niederschlagsereignissen angeben.

Die Prognose unter Verwendung der abweichenden Meteorologiedatensätze erfolgt wiederum nach dem zehnfachen Kreuzvalidierungsschema. Die eingesetzten Modelle werden dabei weiterhin mit $\frac{9}{10}$ der gemessenen historischen Daten trainiert. Lediglich der Testdatensatz, also das restliche $\frac{1}{10}$ beinhaltet die abweichenden meteorologischen Attribute.

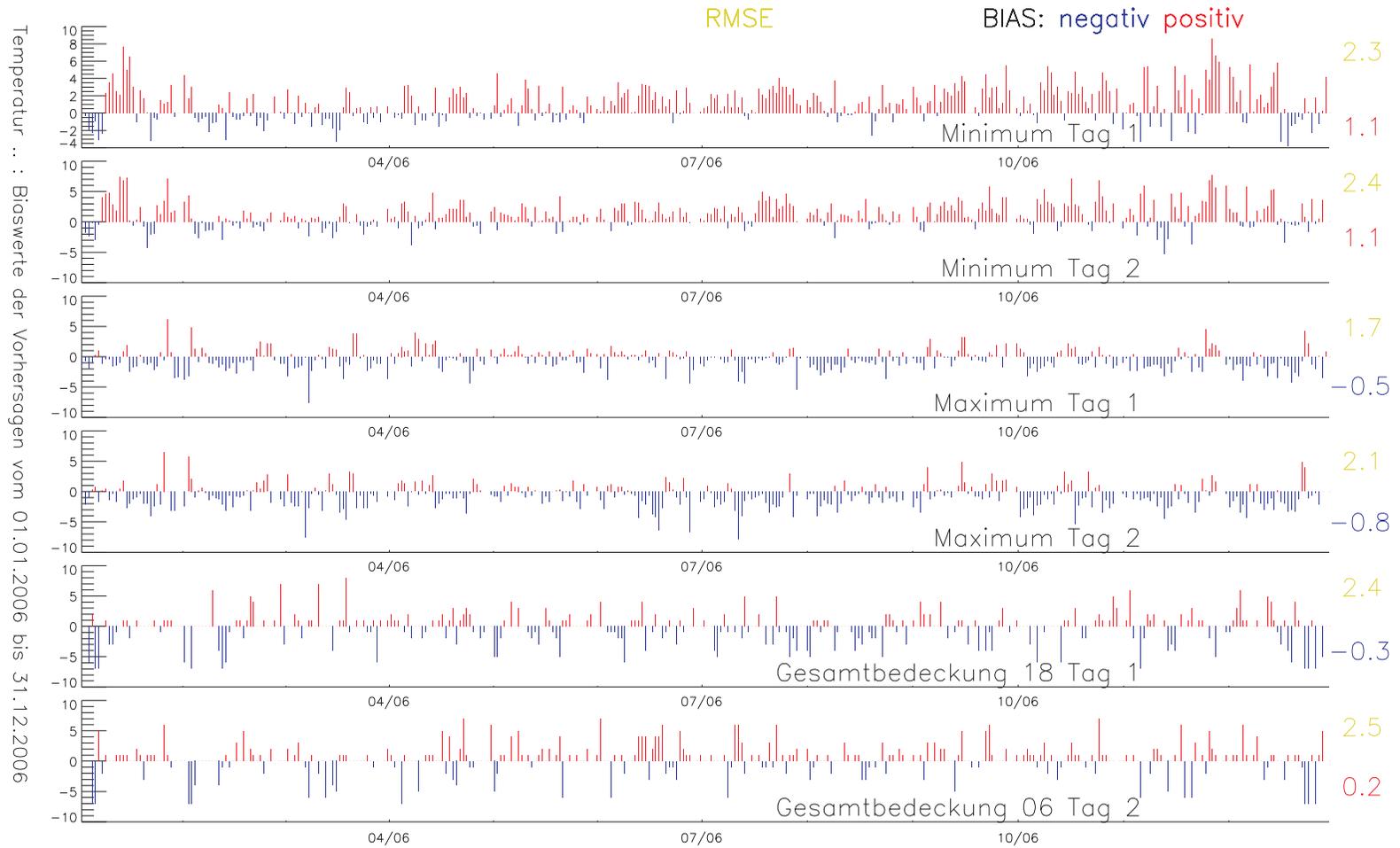


Abbildung 5.7: Verifikation der Prognosen für Bodenwetter (Temperatur und Gesamtbedeckung) an der DWD-Station Stuttgart-Echterdingen für 2006 [22]

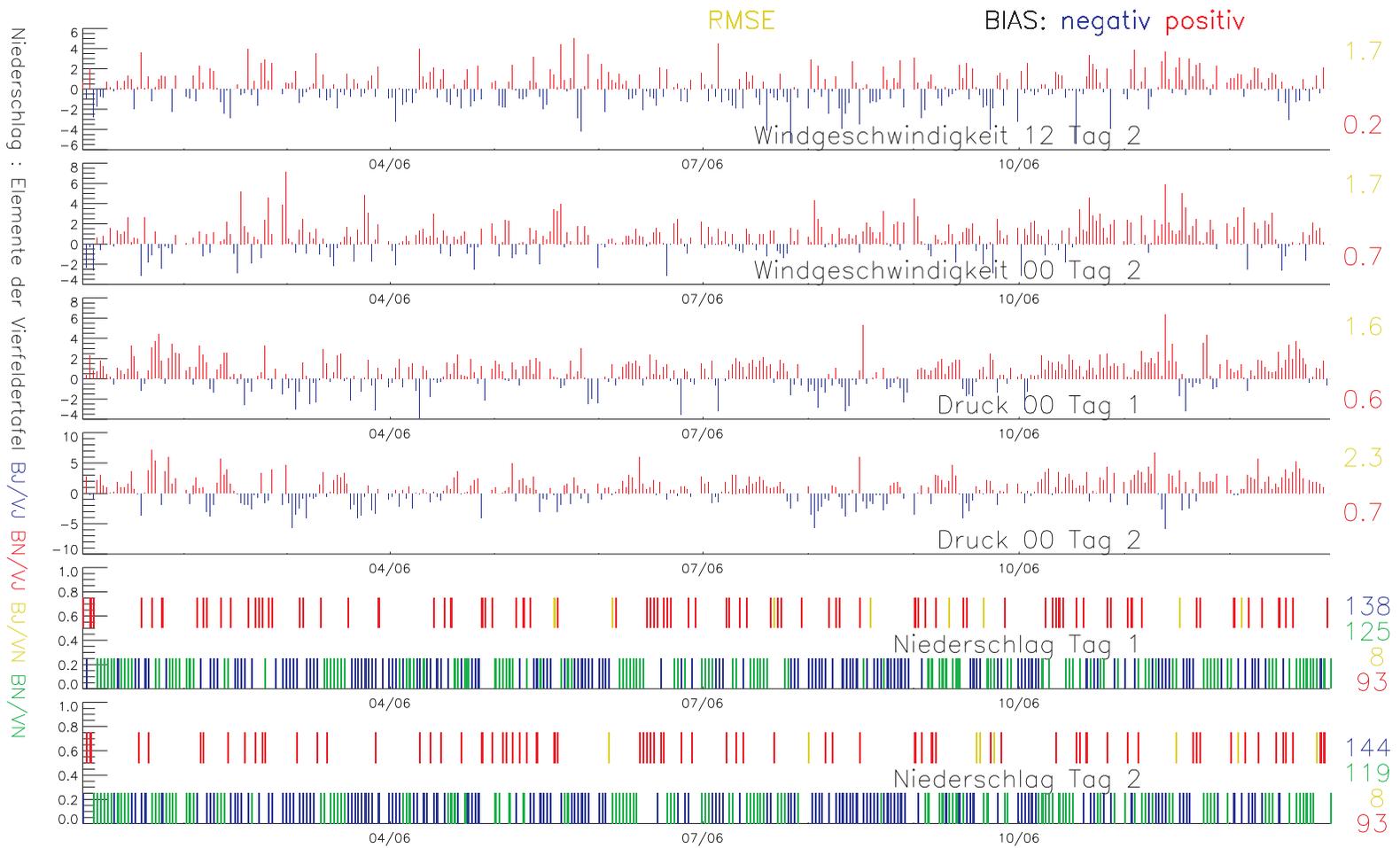


Abbildung 5.8: Verifikation der Prognosen für Bodenwetter (Windgeschwindigkeit, Druck und Niederschlag) an der DWD-Station Stuttgart-Echterdingen für 2006 [22]

Kapitel 6

Vorstellung der Ergebnisse

Die Vorstellung der Ergebnisse ist in drei Abschnitte unterteilt: In den ersten beiden sollen die Prognoseergebnisse der verschiedenen Modelle miteinander verglichen werden. Dazu werden zunächst werden die Ergebnisse der Prognosen vorgestellt, die auf gemessenen Meteorologiedaten beruhen. Anschließend folgt die Präsentation der Prognoseergebnisse, für die von der Realität abweichenden meteorologischen Daten verwendet wurden. Erstere dienen zur Abschätzung der potenziellen Leistungsfähigkeit der Prognosemodelle, Letztere zu Abschätzung der Leistung unter realen Bedingungen unter Berücksichtigung von Unsicherheiten der meteorologischen Vorhersagen. Im dritten Abschnitt wird auf den Einfluss der einzelnen Attribute auf das Prognoseergebnis eingegangen.

6.1 Ergebnisse der Prognose von PM₁₀-Tagesmittelwerten unter perfekten Bedingungen

Perfekte Bedingungen bei der Kurzzeit-Prognose von PM₁₀-Konzentrationen bedeuten, dass anstelle von meteorologischen Vorhersagen gemessene Meteorologiedaten eingesetzt werden. Da die Vorhersagen mit großen Unsicherheiten behaftet sind, lässt sich mit diesen Ergebnissen, bei deren Generierung die Unsicherheiten ausgeblendet wurden, die potenzielle Leistungsfähigkeit der Modelle abschätzen. Zunächst werden die Ganglinien der gemessenen und der prognostizierten PM₁₀-Konzentrationen be-

trachtet. Sie sind sehr gut dazu geeignet, die verschiedenen Modellergebnisse im Detail zu vergleichen.

Als erstes fällt auf, dass alle Modelle im Vergleich zu den Messungen zumindest subjektiv gute Ergebnisse liefern (vollständige Ganglinien aller Stationen und Modelle finden sich im Anhang C, Abbildung C.1 - C.16). Über weite Strecken verlaufen die Prognosekurven der verschiedenen Modelle parallel oder zumindest von der Tendenz her identisch. Die Amplitude variiert dabei aber je nach Modell stellenweise sehr deutlich, wie man in Abbildung 6.1 am 04.03.05 sieht; hier liegen zwischen den Ergebnissen des Multilayer-Perzeptrons (MLP) und der multiplen Linearen Regression (LR) knapp $20 \mu\text{g}/\text{m}^3$. Die Prognosen des Instanzbasierten Lernens mit den k -nächsten-Nachbarn (IBk) und die der Support-Vektor-Maschine (SVM) liegen an diesem Tag zwischen denen der des MLP und der LR. Das Ansteigen oder Absinken der gemessenen PM_{10} -Konzentration wird jedoch von allen Modellen erkannt, wenn auch unterschiedlich gut wiedergegeben, also mehr oder weniger unter- bzw. überschätzt. Über den gesamten betrachteten Zeitabschnitt prognostizieren, je nach Station, das MLP und die SVM die gemessenen PM_{10} -Konzentrationen am besten. Allerdings bilden auch diese zwei Modelle nicht jede Konzentrationsspitze gut ab.

Charakteristisch ist ebenfalls, dass nur sehr wenige Ereignisse mit hoher Konzentration von den Modellen gar nicht abgebildet werden. Tritt dieser Fall doch ein, dann wird das Ereignis meist von allen Modellen nicht erkannt, wie in Abbildung 6.2 am 13.04.2002 zu sehen. Für diesen Tag unterschätzen alle Modelle die gemessene Konzentration um $22 \mu\text{g}/\text{m}^3$ (LR) bis $30 \mu\text{g}/\text{m}^3$ (SVM), wobei die Tendenz vom 12. zum 13.04.2002 als fallend prognostiziert wird. Die Konzentration steigt in Wirklichkeit jedoch an. Ein gleichermaßen von allen Modellen nicht wahrgenommenes Ereignis weist darauf hin, dass entweder die zur Prognose herangezogenen Attribute es nicht erklären können oder keines der Modelle in der Lage ist, den Wirkungsmechanismus zwischen Attributen und Konzentrationen in vollem Umfang abzubilden.

Verlaufen die Prognosekurven über weite Strecken parallel zu denen der gemessenen Konzentrationen, so scheint in manchen Zeiträumen ein Versatz von einem Tag zwischen Prognose und Messung vorzuliegen; dies ist als drittes charakteristisches Merkmal zu bezeichnen. Abbildung 6.3 zeigt einen solchen Zeitversatz am Beispiel zweier Konzentrationsspitzen, die von allen Modellen zwar unterschätzt, aber dennoch wahrgenommen werden. Auf die gemessenen Spitzenwerte vom 07. und 17.01.02

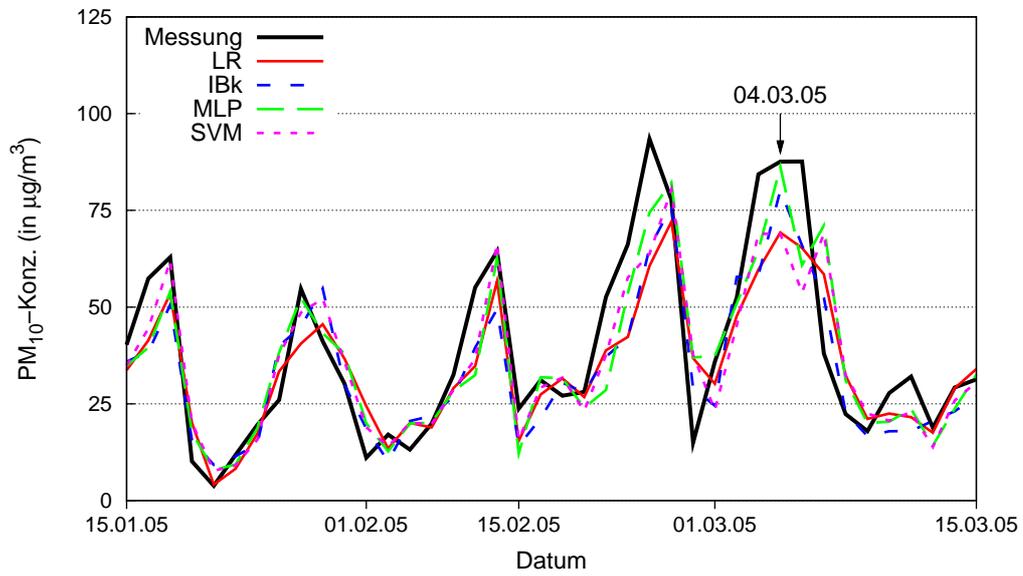


Abbildung 6.1: Vergleich der PM_{10} -Prognoseergebnisse der verschiedenen Modelle für einen Zeitraum im Winter 2005 an der Station Stuttgart-Bad Cannstatt

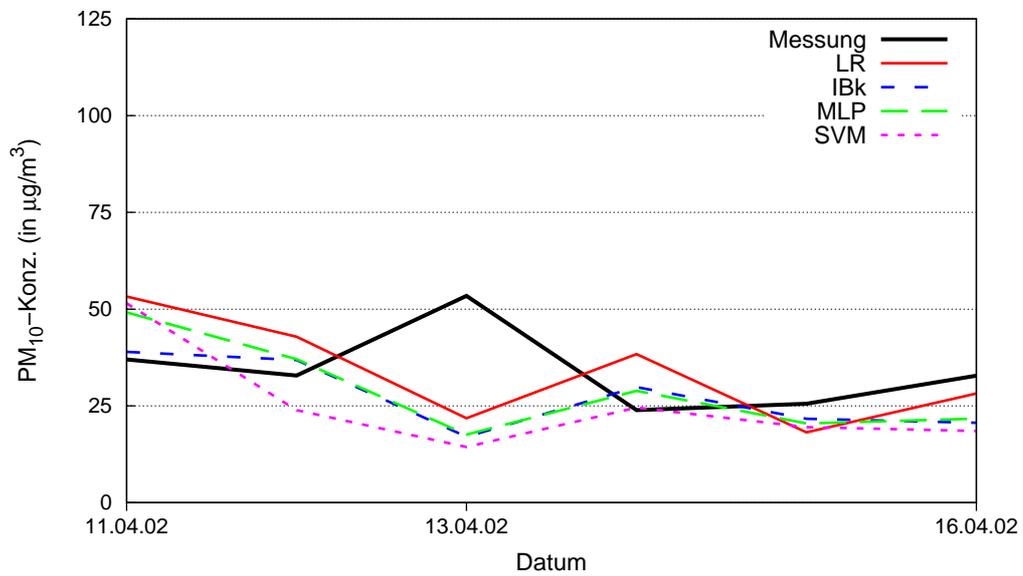


Abbildung 6.2: Vergleich der PM_{10} -Prognoseergebnisse der verschiedenen Modelle für einen Zeitraum im Frühjahr 2002 an der Station Stuttgart-Bad Cannstatt

folgen am nächsten Tag trotz fallender Messwerte die Spitzenwerte der Prognosen. Ähnliches lässt sich auch für die Konzentrationsssenken vom 08. und 13.01.02 feststellen; hier folgt auf die niedrige gemessene Konzenzentration am nächsten Tag ein lokales Prognoseminimum. Eine Erklärung für den Zeitversatz kann der große Einfluss der gemessenen Vortageskonzentration auf den Prognosewert sein (siehe Abschnitt 6.3). Können die Modelle für den Prognosetag Einflüsse, die die PM_{10} -Konzentration senken würden, nicht richtig abbilden oder fehlen diese Attribute gar, so spiegelt sich der gemessene Vortageswert in der Prognose überproportional wider.

Bei Datensätzen, die je nach Station bis zu 1200 Einzeltage umfassen, an denen mal das eine, mal das andere Modell besser abschneidet, eignen sich die Ganglinien schwerlich als Bewertungskriterium für eine allgemeine Leistungsabschätzung eines Modells. Um zu bewerten, wie gut ein Modell grundsätzlich arbeitet, wurden deshalb Scatterplots erstellt. Diese lassen zwar keine Aussagen über die Situation an einzelnen Tagen zu, wie dies bei den Ganglinien der Fall ist, jedoch können sie ein besseres Gesamtbild vermitteln.

Die Scatterplots weisen für alle Modelle einen guten linearen Zusammenhang zwischen den gemessenen PM_{10} -Konzentrationen und den Prognosewerten auf. Im Bereich der geringen bis mittleren Kozenztrationen bis etwa $40 \mu\text{g}/\text{m}^3$ zeigt sich z.B. an der Station Stuttgart-Bad Cannstatt (s. Abbildung 6.4) bei allen Modellen ein relativ ausgeglichenes Bild mit geringen Unter- und Überschätzungen der gemessenen Konzentrationen – abgesehen von wenigen Ausreißern. Darüber fächert die Punktwolke deutlich auf, wobei die Modelle dazu neigen, hohe Konzentrationen zu unterschätzen. Es scheint jedoch, dass das MLP mit den schmalsten Punktwolken die besten Prognoseergebnisse liefert. Dies lässt sich auch an den anderen Stationen feststellen (siehe Abbildungen C.17 bis C.19).

Wie bereits beim Betrachten der Ganglinien, fällt auch an den Scatterplots in Abbildung 6.4 auf, dass die meisten Ereignisse mit hoher Konzentration von allen Modellen abgebildet werden, wenn auch mit unterschiedlicher Genauigkeit. Deutlich wird dies z.B. an Markierung 1 – dem Ereignis mit der im Betrachtungszeitraum höchsten gemessenen Konzentration ($114 \mu\text{g}/\text{m}^3$) an der Station Stuttgart-Bad Cannstatt. Alle Modelle bilden die Konzentrationsspitze ab; sie unterschätzen in ihrer Prognose den tatsächlich gemessenen Wert jedoch, wobei das MLP die geringste Abweichung von $15 \mu\text{g}/\text{m}^3$ aufweist. Es folgen die LR, das IBk und die SVM mit 16, 39 bzw. $43 \mu\text{g}/\text{m}^3$.

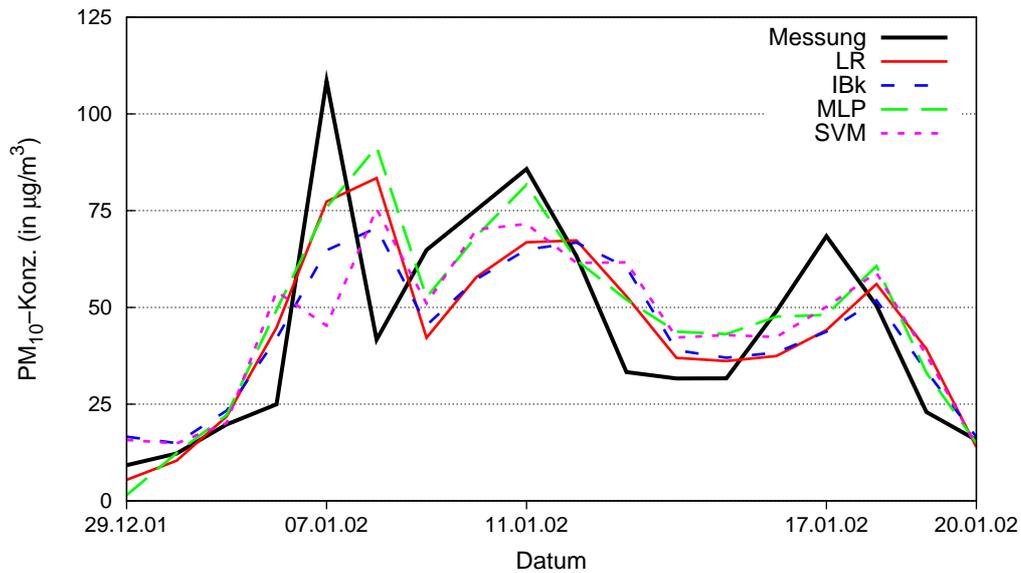


Abbildung 6.3: Vergleich der PM_{10} -Prognoseergebnisse der verschiedenen Modelle für einen Zeitraum im Winter 2001 an der Station Karlsruhe-Nordwest

Außerdem zeigt die Abbildung eines der seltenen Ereignisse, das von keinem der Modelle abgebildet wird (Markierung 2). Statt $53 \mu\text{g}/\text{m}^3$ und somit einer Überschreitung des Europäischen Grenzwerts prognostizieren alle Modelle etwa $15\text{-}20 \mu\text{g}/\text{m}^3$ – ein Wert, der deutlich unter dem Jahresdurchschnitt an dieser Station liegt. Grund dafür könnte eine außergewöhnliche Emissionssituation wie z.B. ein Verkehrsstau, eine Baustelle oder ein Brand im Umfeld der Messstation sein, die das Ereignis verursacht.

Die Erkenntnisse über die Leistungsfähigkeit der Modelle, die aus den Ganglinien und den Scatterplots gewonnen werden können, sind eher qualitativ und somit schwer vergleichbar. Zur exakten Einschätzung der Prognosegüte der einzelnen Modelle eignen sich am besten die in Abschnitt 3.3 vorgestellten Leistungskennzahlen. Diese sind in Tabelle 6.1 aufgetragen – ergänzt um die Kennzahlen der Persistenz, bei der angenommen wird, dass die Konzentration von heute der von morgen entspricht. Die Persistenzannahme stellt das einfachste denkbare Prognoseverfahren dar und dient zur Überprüfung, ob der Aufwand der Prognose mit maschinellen Lernverfahren lohnt. Bei Hervorhebung des jeweils besten Resultats ergibt sich ein eindeutiges Bild: das MLP ist an allen betrachteten Stationen das Modell mit der besten Prognoseleistung. Im Detail betrachtet lassen sich noch einige weiterreichende Aussagen treffen:

Tabelle 6.1: Leistungskennzahlen der verschiedenen Prognosemodelle für die Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Biberach		Persistenz	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,81	0,87	0,88	0,89	0,88
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	6,1	5,0	4,8	4,6	4,7
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	8,7	6,9	6,8	6,3	6,8
Index of Agreement (IA)		0,90	0,93	0,93	0,94	0,93
False Alarm Rate (FAR)	%	36	21	9	20	27
Probability of Detection (POD)	%	64	61	66	80	73
LuQx richtig prognostiziert	%	57	63	63	65	66

Freiburg-Mitte		Persistenz	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,79	0,86	0,87	0,88	0,87
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	6,4	5,1	5,1	4,7	4,8
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	9,3	7,2	7,1	6,7	7,1
Index of Agreement (IA)		0,88	0,92	0,92	0,93	0,93
False Alarm Rate (FAR)	%	40	26	13	20	24
Probability of Detection (POD)	%	58	51	47	65	54
LuQx richtig prognostiziert	%	57	62	62	65	65

Karlsruhe-Nordwest		Persistenz	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,76	0,83	0,81	0,84	0,84
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	7,1	5,8	6,1	5,5	5,4
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	9,7	7,8	8,1	7,4	7,4
Index of Agreement (IA)		0,87	0,90	0,88	0,91	0,91
False Alarm Rate (FAR)	%	34	18	16	16	18
Probability of Detection (POD)	%	66	69	60	73	69
LuQx richtig prognostiziert	%	52	59	57	61	61

Stuttgart-Bad Cannstatt		Persistenz	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,76	0,84	0,83	0,86	0,84
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	7,5	6,2	6,2	5,6	5,8
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	10,4	8,2	8,4	7,6	8,0
Index of Agreement (IA)		0,87	0,90	0,89	0,92	0,91
False Alarm Rate (FAR)	%	41	23	23	19	26
Probability of Detection (POD)	%	54	45	47	63	53
LuQx richtig prognostiziert	%	54	58	60	62	61

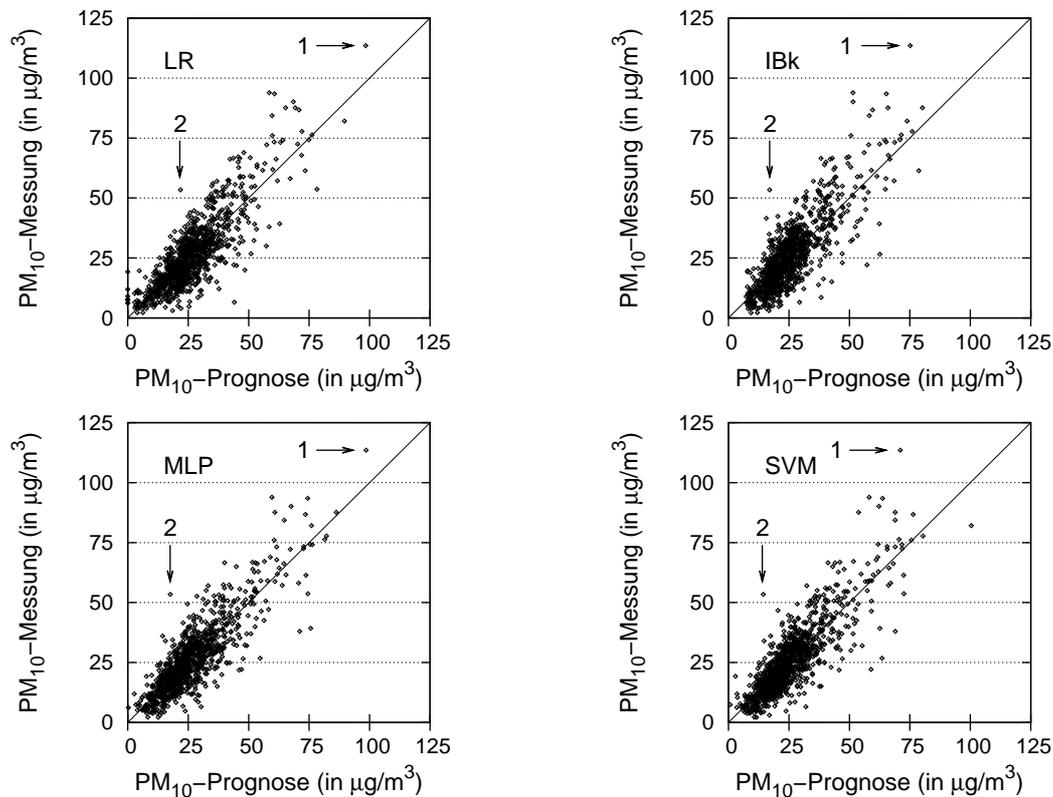


Abbildung 6.4: Vergleich der Messdaten mit den Prognosedaten der verschiedenen Modelle für die Station Stuttgart-Bad Cannstatt

- Die guten Korrelationskoeffizienten aller Modelle an allen Stationen unterstützen die Beobachtung, dass die Prognosekurven zueinander weitgehend parallel bzw. tendenziell der Messkurve entsprechend verlaufen. Außerdem weisen die Modelle, die seltener einen Zeitversatz zwischen tatsächlichem Ereignis und Prognose zeigen, einen besseren Korrelationskoeffizienten auf. Das sind vornehmlich die MLPs und an der Station Karlsruhe-Nordwest auch die SVM. Alle anderen Modelle fallen – wenn auch zumeist nur geringfügig – ab.
- Eine Aussage über die Abweichungen der Prognose gegenüber der Messung lässt sich mithilfe des Mean Absolute Error (MAE) und des Root Mean Squared Error (RMSE) treffen, die beide in $\mu\text{g}/\text{m}^3$ angegeben werden. Betrachtet man den MAE so liegt das MLP an drei Stationen, die SVM lediglich an der Station Karlsruhe-Nordwest vorne. Die durchschnittliche Abweichung zwischen Prognose und Messung ist bei diesen Modellen an der jeweiligen Station also am geringsten. Am schlechtesten schneiden die LR und das IBk mit dem höchsten

MAE an jeweils zwei Stationen ab. Gewichtet man durch Verwendung des RMSE große Abweichungen stärker als geringfügige, was für die PM₁₀-Prognose sinnvoll scheint, ist das MLP an allen Stationen führend – der RMSE also am kleinsten. Wiederum sind die Abweichungen beim LR und beim IBk am größten. Auch der Index of Agreement, der die Abweichungen der Modellergebnisse von den Messungen darstellt, unterstreicht diese Beobachtungen.

- Beziffert man die Unterschiede des RMSE und des MAE beim jeweils besten und schlechtesten Modell einer Station, so liegen sie unter $1 \mu\text{g}/\text{m}^3$, was zunächst relativ gering erscheint. Da die Prognosekurven aller Modelle über weite Strecken einigermaßen parallel verlaufen, ist dies nicht verwunderlich. Die deutlichen Unterschiede, die sich an einzelnen Tagen ergeben, werden durch die Vielzahl der Tage mit geringen Unterschieden verschleiert – beim MAE mehr, beim RMSE weniger.
- Dass es deutliche Unterschiede bei der Prognosegenauigkeit der verschiedenen Modelle gibt, besonders an Tagen mit hohen PM₁₀-Konzentrationen, zeigt die Probability of Detection (POD). Sie gibt an, wie viele Überschreitungen des EU-Grenzwertes für PM₁₀ richtig erkannt werden. Das MLP erkennt bis zu 80% aller Grenzwertüberschreitungen und liegt damit an allen Stationen deutlich vorn. Am schlechtesten schneiden wiederum die LR und das IBk ab, die SVM liegt im Mittelfeld. Die POD darf jedoch nicht unabhängig von der False Alarm Rate (FAR) betrachtet werden. Die FAR zeigt an, wieviel Prozent der prognostizierten Grenzwertüberschreitungen falsch sind, die Vergleichsmessung also unter dem Schwellwert von $50 \mu\text{g}/\text{m}^3$ liegt. Ein Modell, das die PM₁₀-Konzentrationen generell stark überschätzt und so eine hohe POD erreicht, müsste auch bei der FAR hohe und damit schlechte Werte zeigen. Dies ist beim MLP jedoch nicht der Fall. An zwei von vier der betrachteten Stationen besitzt es die beste FAR, an den anderen liegt sie auf Platz zwei. Auffällig ist, dass das sonst so schlecht abscheidende IBk an drei Stationen die beste FAR besitzt. Erkauft wird diese FAR jedoch mit einer schlechten POD, was darauf hinweist, dass das IBk hohe Werte eher unterschätzt. Deshalb prognostiziert es zwar selten fälschlicherweise eine Grenzwertüberschreitung, im Gegenzug dazu aber auch selten tatsächliche Überschreitungen.

- Der deutsche Kurzzeit-Luftqualitätsindex (LuQx) wird von allen Modellen ähnlich gut wiedergegeben – dennoch liegen MLP und SVM wiederum leicht vorne. Die geringen Unterschiede von maximal vier Prozent zwischen dem besten und schlechtesten Modell sind durch die großen Intervalle der einzelnen Indexklassen zu erklären. Die Intervalle betragen 10-50 $\mu\text{g}/\text{m}^3$ (s. Tab. 3.2), die Modelle weichen laut MAE im Schnitt nur um etwa 1 $\mu\text{g}/\text{m}^3$ voneinander ab. Oftmals liegen deshalb unterschiedlich genaue Prognosen im selben Intervall und somit in der selben Indexklasse. Eine genauere Gegenüberstellung des aus Prognosen und Messungen bestimmten LuQx zeigen die Tabellen C.1 bis C.4.

6.2 Ergebnisse der Sensitivitätsanalyse – Abschätzung der Modellperformance unter realen Bedingungen

Nachdem die bisher vorgestellten Ergebnisse die potenzielle Leistungsfähigkeit der Modelle für „perfekte“ meteorologische Vorhersagen darstellen, werden im Folgenden die Ergebnisse der Sensitivitätsanalyse bezüglich unsicherer meteorologischer Vorhersagen betrachtet. So wird versucht, die Modellperformance unter realen Bedingungen abzuschätzen – dabei kann die Realität durch fehlende historische Meteorologievorhersagen jedoch nur angenähert werden (siehe Abschnitt 5.4).

Erwartungsgemäß sind die Ergebnisse deutlich schlechter, wenn anstelle der meteorologischen Messwerte von den Messungen abweichende Daten zur PM_{10} -Prognose genutzt werden. Interessant ist hierbei neben den absoluten Werten der Leistungskennzahlen auch deren Entwicklung bei zunehmend ungenauen meteorologischen Vorhersagen. Außerdem gewinnt der Vergleich der Modellergebnisse mit der Persistenzannahme an Bedeutung und es stellt sich folgende Frage: Wird ein Punkt erreicht, an dem die Prognose mittels maschineller Lernverfahren keine Vorteile mehr gegenüber der simplen Annahme, dass die PM_{10} -Konzentration morgen dieselbe wie die heutige sei?

In Abbildung 6.5 ist die Entwicklung des RMSE bei steigender Abweichung der meteorologischen Vorhersagen von links nach rechts aufgetragen. Bei 0% Abweichung, also der „perfekten“ Vorhersage, bietet das MLP an allen Stationen die besten Prognoseergebnisse. Dasselbe ist der Fall für Abweichungen von bis zu $\pm 10\%$ bzw. $\pm 25\%$. Erst bei Abweichungen, die bis zu $\pm 50\%$ von der gemessenen Meteorologie betragen,

ist die Leistung anderer Modelle an drei Stationen geringfügig besser. Es bietet sich jedoch kein einheitliches Bild, da an zwei Stationen die SVM, an einer weiteren das IBk den besten RMSE liefert. Außerdem muss berücksichtigt werden, dass die Unterschiede an den Stationen Freiburg-Mitte und Karlsruhe-Nordwest zwischen dem besten und dem schlechtesten Modell bei $\pm 50\%$ Abweichung (IBk bzw. SVM) nur minimal ausfallen. Lediglich an der Station Stuttgart-Bad Cannstatt ist das MLP etwas deutlicher abgeschlagen.

Hervorzuheben ist bei der Entwicklung des RMSE, wie stabil vor allem das IBk an allen betrachteten Stationen ist. Auch bei steigenden Abweichungen in der meteorologischen Vorhersage verschlechtert sich der RMSE nur wenig. Ähnlich stabil sind nur das MLP und die SVM an einzelnen Stationen wie z.B. Karlsruhe-Nordwest. Es ist zu vermuten, dass bei noch größeren Abweichungen das IBk auch an anderen Stationen einen besseren RMSE liefern könnte, als die anderen betrachteten Modelle.

Ein ganz anderes Bild zeigt die LR. Vor allem an den Stationen in Karlsruhe und Stuttgart verschlechtert sich der RMSE mit zunehmenden Abweichungen dramatisch. Der RMSE steigt hier um fast $3 \mu\text{g}/\text{m}^3$ an – zum Vergleich sind es beim MLP lediglich $0,6 \mu\text{g}/\text{m}^3$. Die LR ist zumindest an diesen Stationen völlig ungeeignet, bei größeren Abweichungen der meteorologischen Vorhersagen akzeptable Prognosen zu liefern. In beiden Fällen liefert bei einer Abweichung der Meteorologie von bis zu $\pm 50\%$ sogar die Persistenz-Annahme bessere Prognoseergebnisse.

Neben dem RMSE ist vor allem die möglichst exakte Prognose hoher PM_{10} -Konzentrationen und insbesondere die Vorhersage von EU-Grenzwertüberschreitungen von Bedeutung. Aus diesem Grund ist die Entwicklung der POD in Abbildung 6.6 aufgetragen. Zeigte sich beim RMSE zumindest bei einer Abweichung der meteorologischen Vorhersagen von bis zu $\pm 50\%$ ein uneinheitliches Bild, so ist es bei der POD eindeutig. Das MLP zeigt an allen Stationen die besten Resultate – egal ob die Abweichung $\pm 0\%$ oder $\pm 50\%$ beträgt. Schien das IBk durch seine Konstanz beim RMSE bei großen Abweichungen noch eine interessante Alternative zu sein, so ist es bei der Prognose von Überschreitungen des EU-Grenzwerts sehr weit abgeschlagen. Die SVM liefert die zweitbeste Vorhersage von Grenzwertüberschreitungen, erreicht jedoch nie das hohe Niveau des MLP. Da Überschreitungen des Grenzwerts für PM_{10} oftmals über mehrere Tage aufgrund von längeranhaltenden Inversions- oder sonstigen stabilen Wetterlagen auftreten,

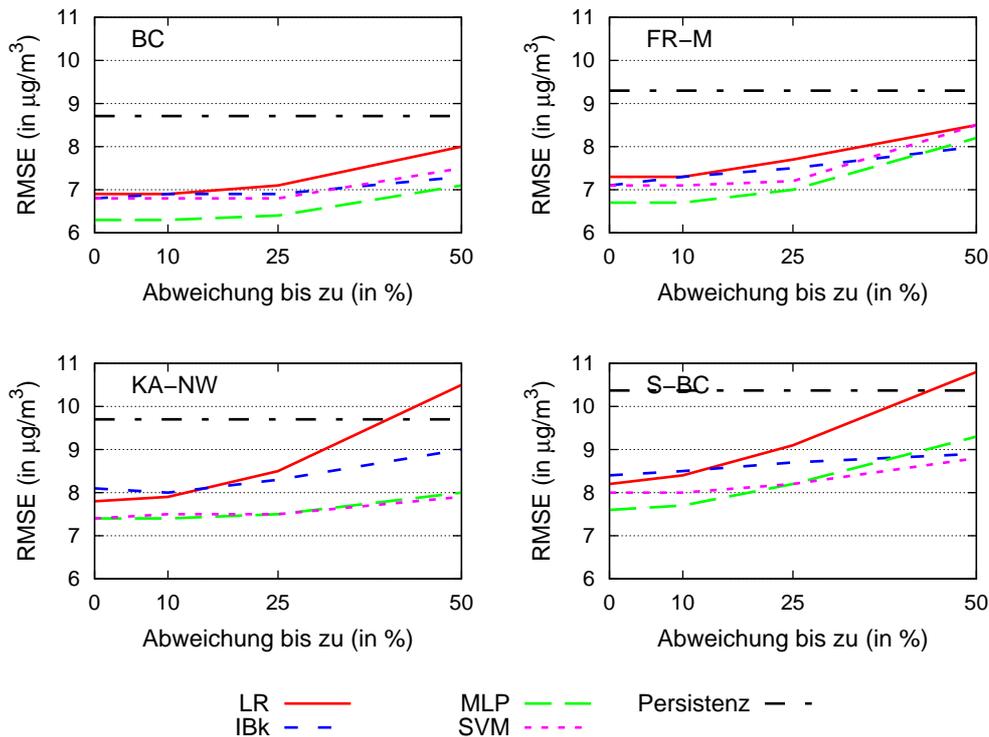


Abbildung 6.5: Veränderung der PM_{10} -Prognosegenauigkeit (RMSE) verschiedener Modelle bei Abweichung der meteorologischen Prognosedaten für Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

schneidet auch die Persistenz-Annahme bei der POD relativ gut ab. Auch bei „perfekten“ Meteorologievorhersagen schlägt sie an den Stationen Freiburg-Mitte und Stuttgart-Bad Cannstatt bis auf das MLP alle anderen Modelle.

Die POD sollte nie unabhängig von der FAR betrachtet werden, da sehr gute Werte bei der Prognose von Grenzwertüberschreitungen mit einer hohen Anzahl von Fehlalarmen einhergehen können. Deutlich wird dies bei der Persistenz-Annahme, wie Abbildung 6.7 zeigt. Sie liefert zwar recht gute Ergebnisse bei der POD, dafür aber auch die schlechtesten bei der FAR. Grund hierfür ist, dass bei der Persistenz-Annahmen am ersten Tag nach jeder Episode von Grenzwertüberschreitungen ein Fehlalarm folgt, da ja der Wert des Vortags als Prognose übernommen wird.

Beim MLP, das bei der POD als bestes Modell abgeschnitten hat, tritt das Phänomen einer guten POD bei einer schlechten FAR lediglich an der Station Freiburg-Mitte auf. Das Modell neigt hier bei steigenden Abweichungen der meteorologischen Vorhersagen zum Überschätzen der PM_{10} -Konzentrationen und somit zu einer hohen

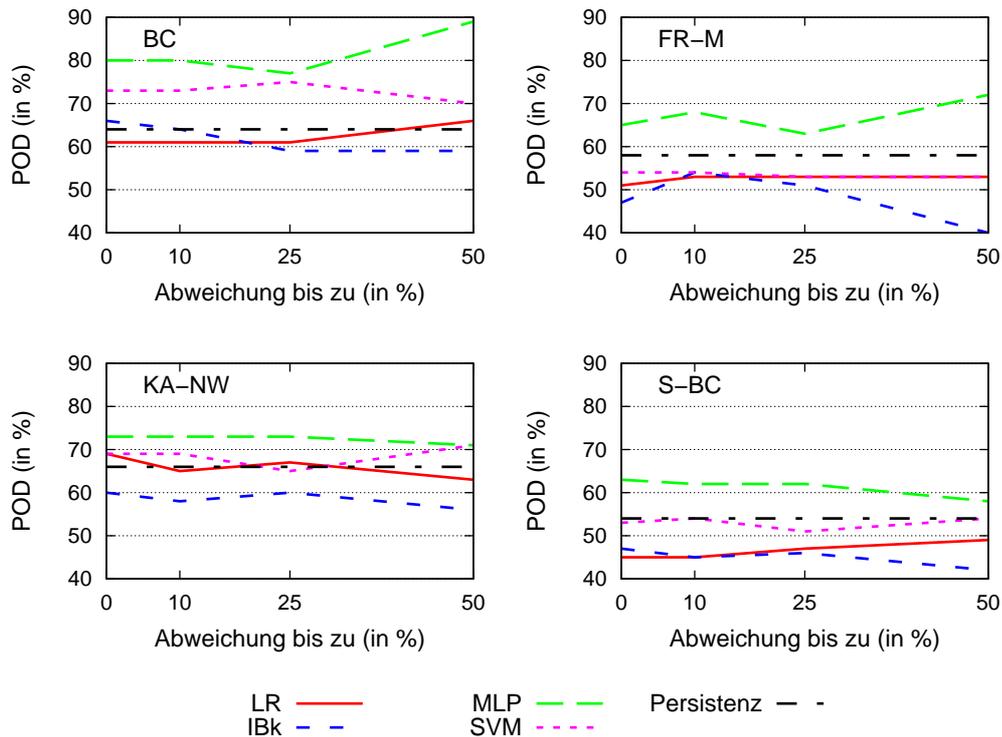


Abbildung 6.6: Veränderung der PM_{10} -Prognosegenauigkeit (POD) verschiedener Modelle bei Abweichung der meteorologischen Prognosedaten für Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

FAR bei einer ebenfalls hohen POD. An allen anderen Stationen zeigt das MLP jedoch immer eine gute FAR bei gleichzeitig hervorragender POD. Wie schon beim Einsatz gemessener Meteorologiedaten zeigt das IBk auch bei abweichender Meteorologie eine gute FAR. Verbunden mit seiner schlechten POD ist dies aber eher auf eine grundsätzliche Unterschätzung hoher PM_{10} -Konzentrationen des Modells zurückzuführen und somit kein Argument für dessen Einsatz. Die SVM und die LR verursachen im Vergleich mit den anderen Modellen deutlich mehr Fehlalarme.

Tabelle 6.2 zeigt in der Übersicht eine Zusammenfassung, welches Modell bei welcher Abweichung der meteorologischen Vorhersagen die beste Leistung bietet (im Detail sind die Leistungskennzahlen im Anhang D dargestellt). Dabei werden die Platzierungen der einzelnen Modelle an den verschiedenen Stationen addiert. Je kleiner das Resultat, desto besser ist die Leistungskennzahl dieses Modells über alle Stationen.

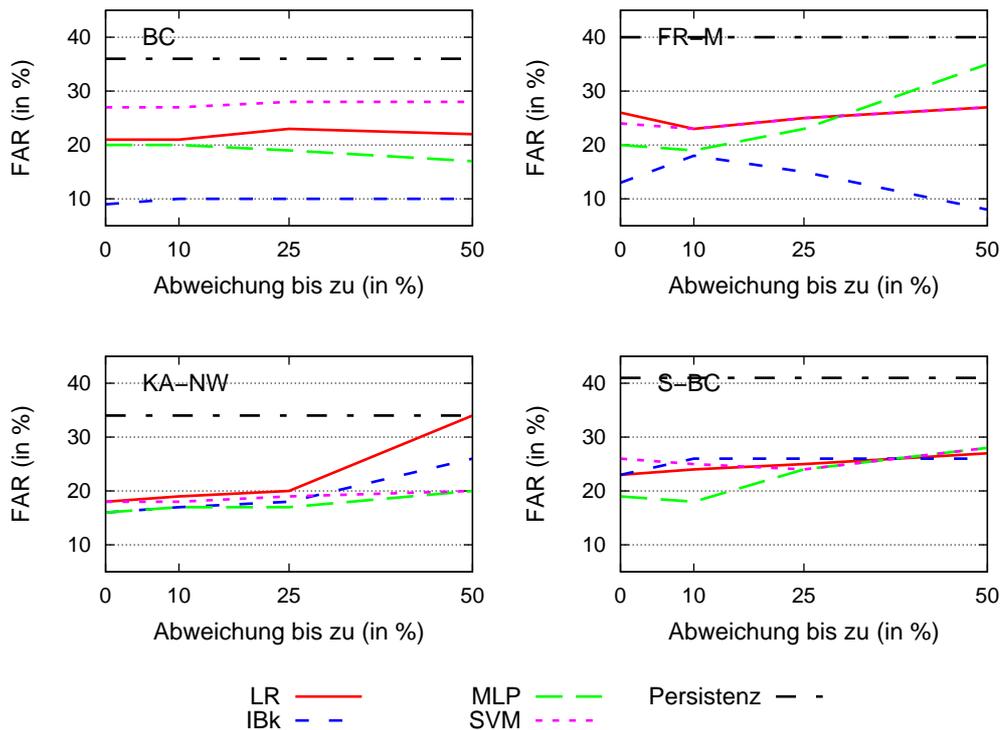


Abbildung 6.7: Veränderung der PM_{10} -Prognosegenauigkeit (FAR) verschiedener Modelle bei Abweichung der meteorologischen Prognosedaten für Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

Auch, wenn das MLP nicht an allen Stationen und in allen Kategorien die besten Werte erzielt hat, so wird doch deutlich, dass es selbst bei abweichenden meteorologischen Vorhersagen im Allgemeinen die beste Prognoseleistung erbringt. Lediglich in wenigen Ausnahmefällen stellt das MLP in der Zusammenfassung nur das zweitbeste Modell dar. In den wichtigsten Kategorien, dem RMSE und der Vorhersage von Grenzwertüberschreitungen (POD), ist dies jedoch nicht der Fall. Zurückzuführen ist die Führungsposition unter anderem auf die Stabilität des MLP. Ist es im Einzelfall an einer Station und in einer Kategorie nicht führend, so belegt es meist zumindest den zweiten Platz. Die übrigen Modelle zeigen bei guter Leistung an einer Station sehr schlechte Leistungen an einer anderen. Beispielsweise ist das IBk bei $\pm 10\%$ Abweichung an drei Stationen bei der FAR führend, an der vierten Station bietet es jedoch die meisten Fehllarme (siehe Tabellen D.1 bis D.4). Solche „Ausfälle“ zeigt das MLP nicht.

Tabelle 6.2: Zusammenfassung der Platzierungen der einzelnen Modelle an allen Stationen bei abweichenden meteorologischen Vorhersagen (Summe der Platzierungen: < bedeutet besser)

keine Abweichung (Messung)	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)	13	12	4	7
Mean Absolute Error (MAE)	13	13	5	7
Root Mean Squared Error (RMSE)	14	12	4	7
Index of Agreement (IA)	11	13	4	6
False Alarm Rate (FAR)	12	5	6	14
Probability of Detection (POD)	13	14	4	8
LuQx richtig prognostiziert	13	13	5	5

±10% Abweichung	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)	12	13	4	7
Mean Absolute Error (MAE)	14	13	5	7
Root Mean Squared Error (RMSE)	13	14	4	8
Index of Agreement (IA)	12	14	4	6
False Alarm Rate (FAR)	12	7	6	13
Probability of Detection (POD)	14	12	4	8
LuQx richtig prognostiziert	13	13	5	6

±25% Abweichung	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)	16	11	4	6
Mean Absolute Error (MAE)	16	12	6	5
Root Mean Squared Error (RMSE)	16	12	4	6
Index of Agreement (IA)	13	14	4	5
False Alarm Rate (FAR)	13	8	6	11
Probability of Detection (POD)	10	16	4	9
LuQx richtig prognostiziert	15	13	7	5

±50% Abweichung	LR	IBk	MLP	SVM
Korrelationskoeffizient (r)	15	8	6	7
Mean Absolute Error (MAE)	15	9	7	7
Root Mean Squared Error (RMSE)	15	8	8	8
Index of Agreement (IA)	14	11	4	6
False Alarm Rate (FAR)	11	6	10	10
Probability of Detection (POD)	12	15	4	7
LuQx richtig prognostiziert	15	9	7	8

6.3 Abschätzung des Einflusses einzelner Attribute auf die Prognoseergebnisse

Einer der wichtigsten Faktoren für die Akzeptanz von Modellen ist neben den Ergebnissen deren Nachvollziehbarkeit bzw. Transparenz. Aus diesem Grund wird auch oftmals die LR zur Prognose von PM_{10} -Konzentrationen genutzt. Anhand der aus einer LR resultierenden Linearen Regressionsgleichung kann man den Einfluss der verschiedenen Attribute wie Windgeschwindigkeit und Niederschlagssumme direkt ablesen. Werden die Attribute vor der Modellerstellung auf Werte zwischen 0 und 1 normiert, so erhält man aus dem Betrag des jeweiligen Regressionskoeffizienten den Anteil der einzelnen Attribute am Prognoseergebnis (siehe Tabelle 6.3).

Maschinelle Lernverfahren gelten dagegen oft als Black-Boxes. Dennoch ist es möglich, solche Modelle transparenter zu machen. Mit dem so genannten *Input Node Interrogator* der eingesetzten Software Qnet2000 [114] ist es möglich, beim MLP den *durchschnittlichen Anteil* eines jeden Knotens der Eingabeschicht auf die Modellausgabe zu berechnen. Man erhält somit für jedes Attribut den durchschnittlichen prozentualen Einfluss auf die prognostizierte PM_{10} -Konzentration (siehe Tabelle 6.3).

Der Unterschied zwischen den Anteilen der Attribute bei der LR und dem MLP besteht darin, dass es sich beim MLP um durchschnittliche Werte handelt. Sind die Anteile bei der LR statisch in der Linearen Regressionsgleichung festgelegt, so sind sie beim MLP dynamisch. Ein MLP ist im Prinzip eine komplizierte nichtlineare Funktion. Dabei bestimmt unter anderem die Aktivierungsfunktion die Ausgabe eines jeden Neurons und gibt somit vor, wie stark ein Attribut in die nächste Schicht des MLP bzw. letztendlich in das Prognoseergebnis eingeht. Da keine lineare, sondern eine sigmoide Aktivierungsfunktion eingesetzt wird, ist auch der Grad der Weitergabe eines Attributs nicht linear, sondern vom Attributwert abhängig und somit dynamisch. Im Extremfall bedeutet dies, dass ein Attribut mit einem geringen durchschnittlichen prozentualen Anteil am Prognoseergebnis durch eine besondere Ausprägung an einem einzelnen Tag einen deutlich größeren Anteil haben kann. Für die Praxis bedeutet dies, dass Attribute mit einem Anteil von wenigen Prozent bei der LR ohne größere Auswirkungen auf die Qualität der Ergebnisse weggelassen werden können. Beim MLP muss durch Weglassen solcher Attribute mit geringem durchschnittlichen An-

Tabelle 6.3: Prozentualer Anteil der einzelnen Attribute am Prognoseergebnis bei der Linearen Regression

		BC	FR-M	KA-NW	S-BC
Tagtyp	%	5	3	3	3
Windrichtung \emptyset	%	3	7	2	1
Windgeschwindigkeit \emptyset	%	9	12	7	7
Taupunkttemperatur \emptyset	%	—	3	11	—
Temperatur \emptyset	%	26	18	29	28
Temperatur-Maximum	%	16	15	11	24
Globalstrahlung	%	3	—	—	—
Niederschlagssumme	%	2	2	2	3
PM ₁₀ Vortag \emptyset	%	30	29	21	23
NO ₂ Vortag \emptyset	%	—	—	—	2
Tage ohne Niederschlag	%	3	5	3	2
relative Luftfeuchtigkeit \emptyset	%	—	—	6	—
Mischungsschichthöhe \emptyset	%	3	3	—	—
Mischungsschichthöhe $\emptyset < 150\text{m}$	%	—	—	1	2
Mischungsschichthöhe $\emptyset < 200\text{m}$	%	—	3	—	—
Mischungsschichthöhe $\emptyset < 300\text{m}$	%	—	—	2	3
Mischungsschichthöhe $\emptyset < 450\text{m}$	%	—	—	—	—
Mischungsschichthöhe $\emptyset < 500\text{m}$	%	—	—	2	—

(eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen)

Tabelle 6.4: Durchschnittlicher prozentualer Anteil der einzelnen Attribute am Prognoseergebnis beim Multilayer-Perzeptron

		BC	FR-M	KA-NW	S-BC
Wochentag gewichtet	%	1	—	—	—
Tagtyp	%	1	2	3	2
Windgeschwindigkeit $\bar{\varnothing}$	%	8	13	12	8
Taupunkttemperatur $\bar{\varnothing}$	%	3	10	4	10
Temperatur $\bar{\varnothing}$	%	20	9	7	9
Temperatur-Maximum	%	13	16	—	—
Temperatur-Minimum	%	—	5	4	11
Δ Temp.-Max./Min.	%	—	—	4	10
Globalstrahlung	%	—	—	—	2
Niederschlagssumme	%	7	3	10	7
PM ₁₀ Vortag $\bar{\varnothing}$	%	37	25	40	28
NO ₂ Vortag $\bar{\varnothing}$	%	2	5	4	1
Tage ohne Niederschlag	%	3	5	5	3
Tage mit Niederschlag	%	5	5	5	1
relative Luftfeuchtigkeit $\bar{\varnothing}$	%	—	3	2	7
Mischungsschichthöhe $\bar{\varnothing}$ <250m	%	1	—	—	—
Mischungsschichthöhe $\bar{\varnothing}$ <450m	%	—	—	2	—

(eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen)

teil am Prognoseergebnis u. U. mit deutlichen Qualitätseinbußen an einzelnen Tagen gerechnet werden.

Des weiteren gibt es erste Ansätze, auch die den SVM zugrundeliegenden Regeln zu extrahieren [23]. Es ist also möglich, mit einem weiteren Analyseschritt auch für maschinelle Lernverfahren eine ähnliche Transparenz wie bei der LR zu erreichen. Somit müssen den Schlussfolgerungen von Lohmeyer et al. [61], die Transparenz sei ein großer Vorteil der LR, revidiert werden.

Vergleicht man die Anteile der Attribute am Prognoseergebnis bei LR und MLP, fällt zunächst die große Bedeutung der Vortages-PM₁₀-Konzentration auf. In allen Modellen und an allen Stationen gehört sie zu den wichtigsten Einflussgrößen der Prognose, wobei man die Anteile bei den unterschiedlichen Modellen aus genannten Gründen nicht 1:1 vergleichen kann. Die Feinstaubkonzentration des Vortages bildet sozusagen die Grundlage für die weitere Entwicklung der Luftqualitätssituation des Prognosetages. Tritt am Prognosetag eine austauscharme Situation auf – z.B. geprägt durch schwache Winde bzw. Windstille oder durch eine Inversion – bleibt die PM₁₀-Konzentration auf dem Niveau des Vortags bzw. steigt darüberhinaus an. Treten hingegen am Prognosetag stärkere horizontale und/oder vertikale Luftbewegungen oder auch Niederschläge auf, so wird die PM₁₀-Konzentration in der Luft verdünnt bzw. ausgewaschen. Von der Vortageskonzentration wird also je nach Grad der Verdünnung oder Auswaschung ein gewisser Anteil abgezogen. Der große Einfluss der PM₁₀-Konzentration des Vortags erklärt auch den in Abschnitt 6.1 angesprochenen Zeitversatz von Konzentrationsspitzen zwischen Prognose- und Messkurven. Fehlen dem Prognosemodell in bestimmten Situationen die reduzierenden Einflussfaktoren als Input oder können sie nicht richtig interpretiert werden, so werden basierend auf den hohen Vortageswerten zu hohe PM₁₀-Konzentrationen prognostiziert.

Der zweite wichtige Faktor für die Prognosemodelle ist die Temperatur, bei der LR und dem MLP jedoch in unterschiedlicher Ausprägung. Bei der LR dominiert durchweg an allen Stationen der Einfluss der Durchschnitts- und der Maximaltemperatur eines Tages. Beim MLP spielen je nach Station neben den genannten auch das Temperatur-Minimum und die Schwankungsbreite der Temperatur im Tagesverlauf eine entscheidende Rolle. Insgesamt sind jedoch alle Temperaturwerte als Indikator für austauscharme Situationen anzusehen, die in Baden-Württemberg hauptsächlich in der kalten Jahreszeit in Form von Inversionswetterlagen auftreten. Die unterschied-

liche Berücksichtigung der Temperaturattribute bei den verschiedenen Modellen und Stationen weist lediglich darauf hin, dass dieselbe Aussage mal aus dem einen, mal aus dem anderen Attribut besser ausgelesen werden kann.

Addiert man die Anteile der Temperaturattribute (inklusive Taupunkttemperatur) so erhält man für beide Modelle an allen Stationen ähnliche Werte von 36 bis 52%. Einzige Ausnahme bildet die Station Karlsruhe-Nordwest. Hier liegt der Anteil für die LR bei 51%, für das MLP jedoch nur bei 19%. Statt dessen zieht das MLP vor allem die Niederschläge und die Tage ohne und mit Niederschlägen zur Prognose heran. Die LR nutzt den Zusammenhang zwischen Niederschlägen und PM_{10} -Konzentrationen nicht im gleichen Maße. Beim Vergleich mit anderen Stationen spiegelt sich dieser deutliche Unterschied in der Nutzung von Attributen jedoch nicht signifikant im Root Mean Squared Error (RMSE) wider.

Die Windgeschwindigkeit spielt sowohl für das MLP als auch die LR eine wichtige Rolle. Mit Anteilen zwischen 7 und 13% stellt sie neben den PM_{10} -Konzentrationen des Vortags und der Temperatur den drittwichtigsten Faktor dar. Relativ geringen Einfluss auf beide Modelltypen haben die Höhe der Mischungsschicht und die emissionsannähernden Attribute Wochentag und Tagtyp. Eine mögliche Erklärung ist, dass eine bodennahe Inversion als konzentrationserhöhender Faktor über die Temperaturattribute für die Modelle besser abzubilden ist als über die Mischungsschichthöhe. Wenn man davon ausgeht, dass die Emissionssituation implizit in den Schadstoffkonzentrationen des Vortages enthalten ist, lässt sich auch der geringe Einfluss von Wochentag und Tagtyp erklären.

Kapitel 7

Kritische Diskussion und Ausblick

7.1 Kritische Diskussion

7.1.1 Qualität der PM₁₀-Prognose

Eine Bewertung der Prognosequalität der verschiedenen Modelle wird bereits in Kapitel 6 vorgenommen. Als bestes Modell zur punktuellen Prognose von PM₁₀ an Messstationen hat sich dabei das Multilayer-Perzeptron (MLP) erwiesen. Dabei waren dessen Prognoseergebnisse signifikant besser, sowohl als die des Referenzverfahrens, der Linearen Regression (LR), als auch der weiteren untersuchten Verfahren, der Support-Vektor-Maschine (SVM) und des Instanzbasierten Lernens mit den k -nächsten-Nachbarn (IBk). Auch im Vergleich zu den Prognosen des Europäischen Ausbreitungs- und Depositionsmodells EURAD liefert das MLP die besseren Ergebnisse – natürlich muss man dabei aber die Ausrichtung von EURAD auf großräumige PM₁₀-Prognosen sehen.

Herausragend ist das MLP vor allem deshalb, weil es auch bei von der Realität abweichenden meteorologischen Eingangsparametern im Vergleich die besten Prognoseergebnisse liefert. Es wird zwar an manchen Stationen von Modellen in einzelnen Kategorien in seiner Prognoseleistung übertroffen; über alle Bewertungskategorien hinweg gesehen erweist es sich jedoch als stabilstes Modell, das an keiner Station schlechte Ergebnisse erzeugt.

Bei Betrachtung der Prognoseergebnisse eines Modells an den verschiedenen Stationen fällt auf, dass die Schwankungsbandbreite sehr groß ist. Beim Root Mean Squa-

red Error (RMSE) reicht sie z.B. beim MLP von $6,3 \mu\text{g}/\text{m}^3$ an der Station Biberach bis zu $7,6 \mu\text{g}/\text{m}^3$ an der Station Stuttgart-Bad Cannstatt. Bei der Anwendung der Methodik auf andere Stationen sind ähnliche oder zum Teil auch noch größere Unterschiede zu erwarten. Dabei sind die für die Prognoseleistung ausschlaggebenden Faktoren für alle Modelle gleichermaßen:

- die Verfügbarkeit aller relevanten Attribute,
- die Größe und Repräsentativität des Satzes an historischen Messdaten und
- die Qualität der meteorologischen Vorhersagen.

Um eine PM_{10} -Konzentration an einer Station prognostizieren zu können, ist es zunächst wichtig, dass alle relevanten (Haupt-)Einflussfaktoren verfügbar sind. Dies sind in der Regel Windgeschwindigkeit und Niederschläge. Je nach Station können jedoch eine Vielzahl weiterer dazukommen – eine ausführliche Beschreibung dazu findet sich in Kapitel 4.1. Die in der vorliegenden Arbeit betrachteten Stationen erfüllen diese Voraussetzung, da an ihrem Standort nicht nur Luftschadstoffe, sondern auch die Meteorologie gemessen wird. Für Stationen, an denen dies nicht der Fall ist, müssen meteorologische Daten von benachbarten Stationen herangezogen werden, was die Prognoseleistung vorraussichtlich verschlechtern wird. Vor allem in dicht besiedelten Stadtgebieten können die lokalen Meteorologieverhältnisse extrem unterschiedlich sein und die Prognose dementsprechend negativ beeinflussen.

Neben der Verfügbarkeit der relevanten Attribute ist auch der Umfang der verfügbaren Zeitreihen von großer Bedeutung. Grundsätzlich gilt, je mehr historische Daten verfügbar sind, desto besser kann ein Modell die physikalischen Zusammenhänge erlernen und später abbilden. Wichtig ist dabei vor allem, dass die historischen Daten möglichst alle an einer Station auftretenden Variationen der Meteorologie und der Emissionssituationen abdecken. Steht nur ein kleiner, nicht repräsentativer, Datensatz zum Trainieren eines Modells zur Verfügung, können in Situationen, die im Trainingsdatensatz nicht vorhanden sind, u.U. keine hinreichend guten Prognoseergebnisse erzielt werden.

Neben den beiden erstgenannten Faktoren, die sich auf die Modellbildung auswirken, beeinflusst die Genauigkeit der meteorologischen Vorhersagen die Prognosequalität bei der Anwendung des Modells maßgeblich. Je besser die benötigten me-

eteorologischen Attribute vorhergesagt werden, desto besser ist in der Regel auch das Ergebnis der PM_{10} -Prognose, wie die Ergebnisse der Sensitivitätsanalyse zeigen (siehe Abschnitt 6.2). Die Qualität der meteorologischen Vorhersagen bildet den größten Unsicherheitsfaktor bei der Prognose von PM_{10} mittels maschineller Lernverfahren.

7.1.2 Optimierung von Attributzusammenstellung und Modellparametern

Die weitgehend automatisierte Optimierung von Attributzusammenstellung und Modellparametern mit Genetischen Algorithmen, Evolutionsstrategien und Rasteroptimierung hat sich als ausgezeichnetes Mittel erwiesen, die Prognosequalität aller Modelle zu verbessern. Die optimale Auswahl lässt sich sonst aus der Vielzahl von möglichen Attributkombinationen und Parametereinstellungen nur mit sehr großem Aufwand treffen. Die Wahl eines Standard-Attributsets für alle Stationen sollte nicht getroffen werden. Zu deutlich ist die Verbesserung der Prognosequalität, die durch stationsabhängige Attribute erzielt werden kann. Wenn jedoch alle Attribute zur Modellbildung herangezogen werden, verschlechtert sich die Performance erheblich.

Keine manuellen Eingriffe bei der Optimierung erfordern die LR und das IBk. Bei der SVM muss die Attributzusammenstellung nach den Optimierungsläufen geringfügig angepasst werden, dafür konnten Standardwerte für die Parametereinstellungen gefunden werden, die sich auch für weitere Stationen verwenden lassen sollten. Unglücklicherweise sind vor allem beim besten Modell, dem MLP die manuellen Eingriffe bei der Optimierung relativ aufwendig. Das liegt vor allem an der Verwendung zweier unterschiedlicher Softwareprodukte. Die Attribut- und die erste Parameteroptimierung wird mit YALE [68] durchgeführt. Die abschließende Modellerstellung erfolgt aufgrund der dort integrierten, sehr effektiven Funktion *Learn Rate Control* zur Optimierung zweier weiterer Modellparameter, mit Qnet2000 [114] (siehe Abschnitt 5.3.2). Die deutliche Verbesserung der Prognosequalität rechtfertigt jedoch diesen weiteren Optimierungsschritt.

Besondere Bedeutung gewinnt die Optimierung der Attributzusammenstellung und der Parametereinstellungen vor allem vor dem Hintergrund einer Anwendung der Methodik auf weitere Stationen. Wie bereits beschrieben sind nicht alle Attribute an allen Stationen gleichermaßen verfügbar. Darüber hinaus können in anderen Regionen oder auch an anderen Messstationstypen weitere Attribute hinzukommen (z.B.

Verkehrsmengen an Straßenstationen) – die ermittelten standortunabhängigen Attribute (siehe Tabelle 5.2) können hier nur eine Orientierung bieten. Denkbar ist auch, dass sich damit die optimalen Parametereinstellungen eines Modells verändern. Eine Überprüfung ist vor allem für das MLP und die SVM nötig, da hier für die Arbeit Standardeinstellungen für alle Stationen getroffen werden.

7.1.3 Stabilität der Prognosemodelle bei Ausfällen von Messungen oder meteorologischen Vorhersagen

Ausfälle von Messungen oder meteorologischen Vorhersagen sind für die Prognosemodelle ein großes Problem. Der Ausfall eines einzelnen Attributs bedeutet für die LR, die SVM und das MLP, dass keine Prognose durchgeführt werden kann. Das IBk hingegen ist unempfindlich gegenüber Messausfällen. Bei der dem IBk zugrundeliegenden Berechnung des Abstands zwischen Trainingsvektor und dem neuen Beobachtungsvektor ist es unerheblich, ob alle Attribute im Beobachtungsvektor vorhanden sind oder nicht. Fehlt eines oder mehrere der Attribute, so wird/werden er/sie bei der Abstandsberechnung einfach nicht berücksichtigt. Es eignet sich daher gut als Backup-System, das, falls z.B. ein MLP aufgrund eines fehlenden Mess- oder meteorologischen Vorhersagewerts nicht eingesetzt werden kann, trotzdem eine Prognose liefert – wenn auch von geringerer Qualität.

Falls einige gemessene oder vorhergesagte Attribute besonders häufig ausfallen ist es auch denkbar, parallel ein oder mehrere weitere MLP zu trainieren, bei denen diese Attribute nicht miteinbezogen werden. Bei der Anzahl der Attribute muss man jedoch sehr genau auswählen, für wie viele Attributkombinationen man diesen Aufwand betreibt.

7.1.4 Physikalische Nachvollziehbarkeit der Modelle

Die physikalische Nachvollziehbarkeit ist entscheidend für die Akzeptanz von Modellergebnissen. Wichtig ist dabei, dass die Eingangsattribute in einem physikalisch eindeutigen Zusammenhang mit den zu prognostizierenden PM_{10} -Konzentrationen stehen. Da alle Modelle dieselben Attribute zur Verfügung haben und deren Relevanz in zahlreichen Studien nachgewiesen wurde (siehe Abschnitt 4.1) sind sie alle

physikalisch nachvollziehbar. Was die Transparenz der verschiedenen Modelle angeht, bestehen jedoch Unterschiede.

Die LR ist auf den ersten Blick am einfachsten nachzuvollziehen. In der Regressionsgleichung sind bei Normierung der Attributwerte die Haupteinflussfaktoren auf die PM_{10} -Konzentration leicht abzulesen. Hingegen sind alle verwendeten maschinellen Lernverfahren grundsätzlich als *Black Box* zu bezeichnen, da man die Berechnungen der Prognosen nicht ohne weiteres nachvollziehen kann. Der Vorteil, im Vorfeld keine Regeln für die Zusammenhänge zwischen Attributen und der PM_{10} -Konzentration formulieren zu müssen, wendet sich hier scheinbar zum Nachteil. Jedoch gibt es für manche der maschinellen Lernverfahren Analysemethoden, um den Einfluss einzelner Attribute auf die PM_{10} -Konzentration zu bestimmen.

Die für das MLP verwendete Software *Qnet2000* [114] lässt mit dem *Input Node Interrogator* eine Abschätzung der des prozentualen Einflusses der verschiedenen Attribute am Prognoseergebnis zu. Anders als bei Lohmeyer [61] beschrieben, bietet die LR also keine größere Transparenz als das MLP. Des weiteren gibt es erste Ansätze, um auch aus SVM [23] die ihnen zugrundeliegenden Regeln zu extrahieren. Es lässt sich also auch hier – verbunden mit einem weiteren Analyseschritt – eine ähnliche Transparenz erreichen, wie dies bei der LR gegeben ist.

7.1.5 Zeitaufwand bei Erstellung und Anwendung der Modelle

Grundsätzlich sind alle eingesetzten Lernverfahren relativ schnell, vergleicht man sie mit numerischen Prognosemodellen wie z.B. EURAD. Dennoch sind erhebliche Unterschiede in der Geschwindigkeit der Lernverfahren zu verzeichnen. Zunächst muss man zwei unterschiedliche Schritte betrachten – (1) die Erstellung eines Modells, die den größten Zeitaufwand benötigt, und (2) den Einsatz eines trainierten Modells.

Die Erstellung beinhaltet die Parameter- und Attributoptimierung und das Modelltraining. Den geringsten Zeitaufwand für die Optimierungsvorgänge und das Modelltraining, etwa 5 Minuten, benötigt das IBk, da bei der Optimierung (fast) nicht manuell nachbearbeitet werden muss. Ähnlich schnell ist auch die LR, allerdings ist bei ihr noch eine aufwendigere Datenvorbereitung vonnöten, da bei diesem Ansatz normalverteilte Eingangsdaten vorausgesetzt werden. Deutlich langsamer sind das MLP und vor allem die SVM. Verursacht wird dies durch die nötige manuelle Nach-

bearbeitung bei der Attributoptimierung und bei der SVM durch die Optimierung der Vielzahl von Modellparametern. Die Parameteroptimierung kann bei der SVM bis zu mehrere Tagen in Anspruch nehmen. Durch die gefundenen standardisierten Parametereinstellungen ist diese aufwändige Optimierung jedoch nicht für jede neue Station nötig. Der Zeitvorteil der LR und des IBk wird durch diese Standardeinstellungen größtenteils wett gemacht. Das den Optimierungsschritten folgende Modelltraining dauert beim MLP und bei der SVM wiederum nur wenige Minuten.

Sind die Modelle trainiert und werden sie zur Vorhersage der PM_{10} -Konzentrationen für den nächsten Tag genutzt, berechnen sie alle das Ergebnis in Sekundenbruchteilen. Bei der operationellen Anwendung unterscheiden sich die Modelle in punkto Geschwindigkeit also nicht.

7.1.6 Reproduzierbarkeit von Modellen und Ergebnissen

Bei der Reproduzierbarkeit von Modellen unterscheiden sich die eingesetzten Methoden deutlich voneinander. Auf der einen Seite stehen die SVM, das IBk und die LR. Die der jeweiligen Modellbildung zugrundeliegenden Optimierungsprobleme weisen alle eine global eindeutige Lösung auf. Die mit diesen Methoden erstellten Modelle sind somit reproduzierbar – liefern bei gleich bleibenden Trainingsdaten jeweils identische Modelle und somit auch identische PM_{10} -Prognosen.

Anders ist dies beim MLP. Durch die Zufallsbelegung der Gewichte beim zum Training eingesetzten *Backpropagation Algorithmus* ist nicht sichergestellt, dass das gesuchte Fehlerminimum das globale darstellt. Vielmehr findet das MLP zumeist ein lokales Optimum, das sich mehr oder weniger vom globalen unterscheidet. Das Modell kann also nicht exakt reproduziert werden. In Folge dessen sind auch die Prognoseergebnisse bei gleich bleibenden Trainingsdaten unterschiedlich. Die Unterschiede sind dabei entgegen anderen Untersuchungen mit MLP [81] bei der PM_{10} -Prognose deutlich ausgeprägt. Sie gehen teilweise so weit, dass ein „schlecht“ trainiertes MLP von anderen Methoden wie der SVM übertroffen wird. Aus diesem Grund sollten immer mehrere MLP mit denselben Eingangsdaten trainiert werden, von denen dann das beste zur Prognose verwendet wird.

7.1.7 Erweiterbarkeit der Modelle um neue Messdaten und weitere Attribute

Bei allen vorgestellten Verfahren ist die Aufnahme neuer für die PM_{10} -Konzentrationen relevanter Einflussgrößen problemlos möglich. Auch können weitere historische Messreihen zum Training der Modelle verwendet werden, was in der Regel die Prognosequalität verbessern sollte. Vor allem im ersten Fall empfiehlt sich jedoch, die Schritte der Optimierung der Attributkombination erneut durchzuführen. So sollte untersucht werden, ob sich die neuen Attribute positiv auf die Prognosequalität auswirken und ob dadurch andere Attribute nicht mehr zur Prognose herangezogen werden müssen.

7.1.8 Änderung der Rahmenbedingungen an einer Messstation

Die Verfügbarkeit von möglichst langjährigen Messreihen ist für eine gute Prognose von größter Bedeutung, da beim Lernprozess Zusammenhänge zwischen PM_{10} -Konzentration und Attributen ermittelt werden. Ein großes Problem stellt daher die Änderung der Rahmenbedingungen an einer Station dar. Das können zum einen stationsbezogene Veränderungen wie eine Verlegung des Messorts oder eine Umstellung der Messmethodik sein. Zum anderen sind aber auch Verschiebungen in der Emissionssituation wie eine Änderung des Straßenverlaufs oder des Verkehrsaufkommens – sowohl kurzfristig durch eine Baustelle als auch langfristig z.B. durch politische Maßnahmen – möglich. Nicht zuletzt sind auch Veränderungen des Klimas von großer Bedeutung.

Ändern sich die Rahmenbedingungen an einer Messstation, so sind die aus historischen Daten extrahierten Zusammenhänge zwischen PM_{10} -Konzentration und Meteorologie bzw. Emissionssituation unter Umständen nicht mehr gültig. Lässt sich eine Umstellung der Messmethodik noch relativ einfach mit einem Korrekturfaktor ausgleichen, so ist dies bei Änderungen des Standorts, der Emissionssituation oder der Wetter- und Strömungssituation deutlich schwieriger. Eine Änderung der Emissionssituation kann durch Änderungen von Straßenverläufen oder Fahrverbote hervorgerufen werden, die Wetter- bzw. Strömungssituation kann durch den Bau oder Abriss von Gebäuden verändert werden. Es liegen jedoch derzeit keine Untersuchungen vor, wie

man solche Veränderungen ausgleichen kann. Denkbar sind wiederum Korrekturfaktoren, die jedoch von Fall zu Fall stark variieren dürften.

7.2 Ausblick

Auch wenn das Multilayer-Perzeptron sehr gute Prognoseergebnisse für PM_{10} -Konzentrationen liefert, es sind dennoch zahlreiche Erweiterungen und zusätzliche Untersuchungen denkbar:

Die vorgestellte Prognose ist bisher auf den folgenden Tag beschränkt. Denkbar sind auch Prognosen für mehrere Tage, die jedoch einhergehend mit steigenden Unsicherheiten bei den meteorologischen Vorhersagen ungenauer werden sollten. Außerdem verlieren auch die eingesetzten PM_{10} - und NO_2 -Messwerte des Vortags an Bedeutung für längerfristige Prognosen. Für mehrtägige PM_{10} -Prognosen ist der Einsatz spezieller künstlicher Neuroner Netze wie das Elman-Netz [29] zu untersuchen. Es ist mittels Rückkopplungen in der Netzarchitektur in der Lage, Zeitreihen gut vorherzusagen.

Um die Prognosequalität weiter zu verbessern, ist es denkbar, jahreszeitabhängige Modelle zu bilden. Da sehr hohe PM_{10} -Konzentrationen vor allem in den Wintermonaten auftreten – bestimmte Mechanismen also nur in der kalten Jahreszeit wirken – kann es sinnvoll sein, Einzelmodelle für Winter und Sommer zu bilden. Diese sind weniger komplex, können aber u.U. bessere Prognoseergebnisse erzielen. Erste Versuche mit der multiplen linearen Regression [61] haben hier schon Erfolge gezeigt.

Da sich die Untersuchungen in der vorliegenden Arbeit auf Messstationen des städtischen Hintergrunds mit Meteorologieerfassung beschränken, ist es von großem Interesse, die vorgestellte Methodik auf weitere auszudehnen. Denkbar sind Stationen mit anderer Emissionscharakteristik, wie z.B. straßennahe Stationen, bei denen dann auch eventuell vorhandene Verkehrsmengenerfassungen mit einzubeziehen sind. Außerdem gilt es zu untersuchen, wie gut die Prognosequalität an Stationen ist, an denen die Meteorologie nicht erfasst wird und deshalb Meteorologiedaten benachbarter Stationen herangezogen werden müssen. Dazu ist auch der Einsatz historischer meteorologischer Vorhersagen anstelle einer Sensitivitätsanalyse interessant.

Auch für Stationen, an denen sich die Rahmenbedingungen wie die Emissionssituation oder Klima- und Strömungsverhältnisse geändert haben, könnte ein Konzept

erarbeitet werden, wie die Prognosemodelle ggf. angepasst werden müssten. Damit wäre es auch möglich, Minderungsmaßnahmen wie Fahrverbote zu simulieren und deren Minderungspotenzial abzuschätzen.

Hinsichtlich der für 2015 geplanten Einführung neuer EU-Grenzwerte für besonders feine Partikel mit einem aerodynamischen Durchmesser kleiner $2,5 \mu m$ ($PM_{2.5}$) [33] ist es auch von großem Interesse, die Prognosen diese Partikel auszudehnen. Sie sind laut der Weltgesundheitsorganisation WHO [118] deutlich schädlicher für die menschliche Gesundheit als PM_{10} da sie durch ihre geringe Größe leichter in tiefere Bereiche des Atemtrakts vordringen können. An den meisten Messstationen fehlen jedoch noch die Voraussetzungen, um die ultrafeinen Partikel getrennt von den gröberen zu erfassen.

Kapitel 8

Zusammenfassung

Feinstaub (PM_{10}) ist als Luftschadstoff mit der Einführung des EU-Grenzwerts von $50 \mu\text{g}/\text{m}^3$ im Tagesmittel im Jahre 2005 in den Blick der Öffentlichkeit geraten [32]. Länger ist jedoch schon bekannt, dass PM_{10} schädliche Auswirkungen unter anderem auf die menschliche Gesundheit hat. Um auf diese Gefährdung mit Maßnahmen reagieren und die Bevölkerung informieren zu können, ist eine möglichst genaue Vorhersage der PM_{10} -Konzentrationen nötig.

Ziel der Arbeit war es, ein Modell zur Kurzfrist-Prognose von PM_{10} -Konzentrationen für den nächsten Tag für die vier Stationen Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt des Luftmessnetzes der Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW) zu erstellen. Dabei wurden drei maschinelle Lernverfahren mit dem in Baden-Württemberg bereits operationell eingesetzten Modell der multiplen Linearen Regression (LR) [61] verglichen. Die eingesetzten maschinellen Lernverfahren sind das Multilayer-Perzeptron (MLP) – ein künstliches neuronales Netz –, die Support-Vektor-Maschine (SVM) sowie das Instanzbasierte Lernen mit den k -nächsten-Nachbarn (IBk). Diese Verfahren konnten auf anderen Gebieten bzw. bei der PM_{10} -Prognose unter limitierenden Rahmenbedingungen vielversprechende Ergebnisse liefern. Grundlage für die Modellerstellung waren historische Luftschadstoff- und Meteorologiemessdaten, die an den Stationen erhoben worden waren. Basierend auf diesen Daten sind die maschinellen Lernverfahren in der Lage, die Zusammenhänge zwischen der PM_{10} -Konzentration und den meteorologischen und sonstigen Einflussparametern (Attributen) abzuleiten und in ein Modell zu übertragen. Ein solches Modell kann dann unter Berücksichtigung

von meteorologischen Vorhersagen zukünftige Feinstaubkonzentrationen prognostizieren.

Als klarer Sieger aus dem Vergleich geht das Multilayer-Perzeptron hervor. Es liefert hervorragende Ergebnisse und ist nicht nur der Referenzmethode, der LR, sondern auch den beiden anderen betrachteten maschinellen Lernverfahren in Bezug auf die Prognosequalität deutlich überlegen. Liegt der RMSE bei der LR je nach Station zwischen 6,9 und 8,2 $\mu\text{g}/\text{m}^3$, so beträgt er beim MLP lediglich zwischen 6,3 und 7,6 $\mu\text{g}/\text{m}^3$. Noch deutlicher wird der Unterschied bei der Prognose besonders hoher und somit besonders relevanter PM_{10} -Konzentrationen. Ist die LR in der Lage, maximal 69% der Überschreitungen des EU-Grenzwerts von 50 $\mu\text{g}/\text{m}^3$ zu erkennen, so prognostiziert das MLP bis zu 80% richtig. Ebenso übertrifft das MLP das etablierte Europäische Ausbreitungs- und Depositionsmodell EURAD in der Prognose punktueller PM_{10} -Konzentrationen, das einen RMSE von z.T. über 20 $\mu\text{g}/\text{m}^3$ aufweist (allerdings für einen anderen Betrachtungszeitraum).

Auch bei der Untersuchung, wie die verschiedenen Modelle auf abweichende meteorologische Vorhersagen reagieren, stellt sich das MLP als stabilstes heraus. Es wurden Abweichungen von bis zu $\pm 10\%$, $\pm 25\%$ und $\pm 50\%$ simuliert. Dabei erhöhte sich der RMSE bei der LR auf maximal 10,8 $\mu\text{g}/\text{m}^3$, beim MLP lediglich auf 9,3 $\mu\text{g}/\text{m}^3$.

Sollte die Vorhersage der vom Modell benötigten Meteorologiedaten ausfallen, kann das MLP keine PM_{10} -Prognose bilden – ebensowenig das Referenzverfahren der LR und die SVM. Deshalb bietet es sich an, das IBk als Backup-System einzusetzen. Seine Prognoseleistung reicht zwar nicht an die des MLP heran, jedoch ist sie besser als bei der LR und dazu unempfindlich gegenüber Ausfällen der meteorologischen Vorhersagen. Mit dem Fehlen einzelner oder mehrerer Meteorologiedaten geht zwar eine Verschlechterung der Prognosequalität einher, jedoch lässt sich zumindest eine Abschätzung der durchschnittlichen PM_{10} -Konzentration des nächsten Tages treffen.

Neben der Wahl des geeigneten maschinellen Lernverfahrens zu Prognose von PM_{10} -Konzentrationen zeigt sich, dass auch die Optimierung der Modellparameter und ihrer Eingangsdaten von großer Bedeutung für die Prognosequalität sind. Für keines der Modelle – auch nicht für das favorisierte MLP – lässt sich ein für alle Stationen geeigneter einheitlicher Satz an Attributen (Einflussfaktoren auf die PM_{10} -Konzentrationen) finden, mit dem optimale Prognoseergebnisse erzielt werden. Allerdings zeigen sich einige Attribute wie Windgeschwindigkeit und Niederschlagssumme

als standort- und modellunabhängig. Diese können als Anhaltspunkt für die Ausdehnung der Methodik auf andere Stationen dienen. Darüber hinaus kommen aber je nach Station und Modell weitere Attribute zum Einsatz, die die jeweilige Prognosequalität verbessern. Diese müssen von Station zu Station neu bestimmt werden; dazu eignen sich Genetische Algorithmen hervorragend. Mit ihrer Hilfe kann eine große Anzahl von möglichen Attributkombinationen sehr schnell nach dem Optimum durchsucht werden.

Anders ist dies bei der Optimierung der Modellparameter. Hier konnten für das MLP Standardeinstellungen gefunden werden, die zumindest für andere Stationen des städtischen Hintergrunds geeignet sein sollten. Die Parameteroptimierung entfällt also. Für Stationen mit anderer Emissionscharakteristik wie z.B. straßennahe Hot-Spot-Stationen ist es jedoch angebracht, diese Standardeinstellungen erneut zu überprüfen. Dies kann mit einer Rasteroptimierung der Parameter in Kombination mit einem Genetischen Algorithmus für die Attributoptimierung erfolgen. Beim Einsatz des IBk als Backup-Methode ist dieser kombinierte Optimierungsschritt für jede neue Station nötig. Allerdings lässt sich der zu untersuchende Parameterraum durch die in der Arbeit gewonnenen Erkenntnisse eng eingrenzen, was die Optimierungszeit drastisch verkürzt.

Literaturverzeichnis

- [1] ACKERMANN-LIEBRICH, U., LEUENBERGER, P., SCHWARTZ, J., SCHINDLER, C., MONN, C., BOLOGNINI, G., BONGARD, J., BRÄNDLI, O., DOMENIGHETTI, G., ELSASSER, S., GRIZE, L., KARRER, W., KELLER, R., KELLER-WOSSIDLO, H., KÜNZLI, N., MARTIN, B., MEDICI, T., PERRUCHOU, A., SCHONI, M., TSCHOPP, J., VILLIGER, B., WÜTHRICH, B., ZELLWEGER, J. und ZEMP, E.: *Lung function and long-term exposure to air pollutants in Switzerland*. American journal of respiratory and critical care medicine, 155(1):122–129, 1997.

- [2] AGIRRE-BASURKO, E., IBARRA-BERASTEGI, G. und MADARIAGA, I.: *Regression and Multilayer Perceptron-Based Models to Forecast Hourly O₃ and NO₂ Levels in the Bilbao Area*. Environmental Modelling & Software, 21(4):430–446, 2006.

- [3] AHA, D., KIBLER, D. und ALBERT, M.: *Instance-Based Learning Algorithms*. Machine Learning, 6(1):37–66, 1991.

- [4] ANKE, K., KLINGNER, M., SÄHN, E. und LÖSCHAU, G.: *Screeningverfahren und Immissionsmodelle zur Auswertung von PM₁₀-Langzeitmessungen*. Gefahrstoffe Reinhaltung der Luft, 63(5):201–208, 2003.

- [5] ANKE, K., SÄHN, E., JOHN, D. und KLINGNER, M.: *Hierarchische Gliederung von Überschreitungsepisoden – erweiterte und detaillierte Analyse für Schwebstaub (PM₁₀)*. Techn. Ber., Fraunhofer Institut für Verkehrs- und Infrastruktursysteme (IVI), Dresden, 2005.

- [6] ANKE, K., SÄHN, E. und KLINGNER, M.: *Statistische Quellgruppenanalyse für die PM₁₀-Belastung in sächsischen Ballungsräumen*. Techn. Ber., Fraunhofer Institut für Verkehrs- und Infrastruktursysteme (IVI), Dresden, 2004.
- [7] ANKE, K., SÄHN, E., KLINGNER, M. und WOLF, U.: *Intensität und Nachwirkung meteorologischer Einflussfaktoren auf die PM₁₀-Konzentration*. Gefahrstoffe - Reinhaltung der Luft, 65(1/2):41–48, 2005.
- [8] ATHANASIADIS, I., KARATZAS, K. und P.MITKAS: *Classification techniques for air quality forecasting*. In: *Proceedings of the 17th European Conference on Artificial Intelligence Riva del Garda, Italy / Aug 28th - Sept 1st, 2006*.
- [9] BACKHAUS, K., ERICHSON, B., PLINKE, W., SCHUCHARD-FISCHER, C. und WEIBER, R.: *Multivariate Analysemethoden – Eine Anwendungsorientierte Einführung*. Springer, Berlin, 4. Aufl., 1987.
- [10] BALAGUER-BALLESTER, E., CAMPS-VALLS, G., CARRASCO-RODRÍGUEZ, J., SORIA-OLIVAS, E. und TASCÓN, S. DESVALLE: *Effective 1-day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks*. Ecological Modelling, 156:27–41, 2002.
- [11] BAUMBACH, G.: *Luftreinhaltung*. Springer, Berlin, 3. Aufl., 1993.
- [12] BENVENUTO, F. und MARANI, A.: *Neural Networks for Environmental Problems: Data Quality Control and Air Pollution Nowcasting*. Global Nest, 2(3):281–292, 2000.
- [13] BENVENUTO, F. und MARANI, A.: *Nowcasting of urban air pollutants by neural networks*. Il Nuovo Cimento, 23(6):567–586, 2000.
- [14] BOZJAR, M., MLAKAR, P. und GRAŠIĆ, B.: *Neural networks based Ozone forecasting*. In: *Proceedings from the 9th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, 2004.
- [15] BRANDT, J. und CHRISTENSEN, J.H.: *THOR Integrated Air Pollution Forecast and Management System.*, 2007. <http://www2.dmu.dk/AtmosphericEnvironment/thor/index.htm>.

- [16] BRUCKMANN, P., WURZLER, S. und MILDENBERGER, K.: *Die Episode hoher Staubkonzentrationen im Januar und Februar 2006*. Immissionsschutz, (2):64–72, 2006.
- [17] CAPPÀ, C., ANFOSSI, D., GROSÀ, M. und NATALE, P.: *Short-term prediction of urban NO₂ pollution by means of artificial neural networks*. International Journal of Environment and Pollution, 15(5):483–496, 2001.
- [18] CHELANI, A., AJGHATE, D. und HASAN, M.: *Prediction of Ambient PM₁₀ and Toxic Metals using Artificial Neural Networks*. Journal of the Air & Waste Management Association, 52(7):805–810, 2002.
- [19] CHELANI, A. und HASAN, M.: *Forecasting nitrogen dioxide concentration in ambient air using artificial neural networks*. International Journal of Environmental Studies, 58:487–499, 2001.
- [20] COHEN, A., ANDERSON, H. und OSTRO, B.: *Mortality impacts of urban air pollution*. In: EZZATI, M., LOPEZ, A., RODGERS, A. und MURRAY, C. J. L. (Hrsg.): *Comparative quantification of health risks – global and regional burden of disease attributable to selected major risk factors*, S. 1353–1434. World Health Organisation, Genf, 2004.
- [21] DEUTSCHER BUNDESTAG: *22. Verordnung zur Durchführung des Bundes-Immissionsschutzgesetzes (BImSchV)*. Bundesgesetzblatt, 66, 2002.
- [22] DEUTSCHER WETTERDIENST: *Verifikationsergebnisse zum COSMO-EU*. DWD, 2007.
- [23] DIEDERICH, J. und BARAKAT, N.: *Hybrid Rule-Extraction from Support Vector Machines*. In: *Proceedings of the 2004 IEEE*, 2004.
- [24] DOCKERY, D., CUNNINGHAM, J., DAMOKOSH, A., NEAS, L., SPENGLER, J., KOUTRAKIS, P., WARE, J., RAIZENNE, M. und SPEIZER, F.: *Health effects of acid aerosols on North American children – respiratory symptoms*. Environmental health perspectives, 104:500–505, 1996.
- [25] DOCKERY, D., POPE, A., XU, X., SPENGLER, J., WARE, J., FAY, M. und FERRIS, B.: *An association between air pollution and mortality in six U.S. cities*. The New England Journal of Medicine, 329(24):1753–1759, 1993.

- [26] DRECHSLER, S., AHRENS, D., LUMPP, R., HOLST, T. und UHRNER, U.: *Hohe Feinstaubbelastung als Folge sekundärer Aerosolbildung*. Immissionsschutz, (2):56–62, 2006.
- [27] DUTOT, A., RYNKIEWICZ, J., STEINER, F. und RUDE, J.: *A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions*. Environmental Modelling & Software, 22(9):1261–1269, 2007.
- [28] EFRAIMIDOU, M., KANAKI, M., ATHANASIADIS, I., MITKAS, P. und KARATZAS, K.: *Data mining air quality data for Athens, Greece*. In: TOCHTERMANN, K. und SCHARL, A. (Hrsg.): *Proceedings of the 20th International Conference on Informatics for Environmental Protection – Managing Environmental Knowledge*, S. 505– 508, 2006.
- [29] ELMAN, J.: *Finding Structure in Time*. Cognitive Science, 14:179–211, 1990.
- [30] EPA, U.: *Guidelines for Developing an Air Quality (Ozone and PM_{2.5}) Forecasting Program*. Techn. Ber., U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 2003.
- [31] EUROPARAT: *Council Directive 1996/62/EC of 27 September 1996 on ambient air quality assessment and management*. Official Journal of the European Communities, L 296:55–63, 1996.
- [32] EUROPARAT: *Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air*. Official Journal of the European Communities, L 163:41–60, 1999.
- [33] EUROPÄISCHES PARLAMENT: *Verbesserte Luftqualität für Europa*, 2007. <http://www.europarl.de/>.
- [34] FERNÁNDEZ, R.: *Predicting time series with a local support vector regression machine*. . In: *Advanced Course on Artificial Intelligence (ACAI'99)*, 1999.
- [35] FISCHER, S., KLINKENBERG, R. und MIERSWA, I.: *YALE 3.4 Yet Another Learning Environment - User Guide, Operator Reference, Developer Tutorial*. Universität Dortmund, 2006.

- [36] FOXALL, R., CAWLEY, G., DORLING, S. und MANDIC, D.: *Error Functions for Prediction of Episodes of Poor Air Quality*. In: DORRONSORO, J. R. (Hrsg.): *International Conference on Artificial Neural Networks*, S. 1031–1036, Berlin, 2002. Springer.
- [37] GARDNER, M. und DORLING, S.: *Artificial neural networks (The multilayer Perceptron) - a review of applications in the atmospheric sciences*. *Atmospheric environment*, 32(14-15):2627–2636, 1998.
- [38] GIESE, M.: *Maschinelles Lernen: Klassische Ansätze I*. Vorlesungsskript. Max-Planck-Gesellschaft, 2002.
- [39] GIMENO-ILLA, J., BÉJAR-ALONSO, J. und SÀNCHEZ-MARRÉ, M.: *Nearest-Neighbours for Time Series*. *Applied Intelligence*, 20(1):21–35, 2004.
- [40] GODLESKI, J., SIOUTAS, C., KALTER, M., CARALANO, P. und KOUTRAKIS, P.: *Death from inhalation of concentrated ambient particles in animal models of pulmonary disease*. In: *Second colloquium on particulate air pollution and human health*, S. 136–143, Park City, Utah, 1996.
- [41] HAINSCH, A.: *Ursachenanalyse der PM10-Immission in urbanen Gebieten am Beispiel der Stadt Berlin*. Dissertation, TU Berlin, 2003.
- [42] HARTUNG, J., ELPELT, B. und KLÖSENER, K. H.: *Statistik – Lehr- und Handbuch der angewandten Statistik*. R. Oldenbourg, München, 5. Aufl., 1986.
- [43] HÄCKEL, H.: *Meteorologie*. Ulmer, Stuttgart, 3. Aufl., 1993.
- [44] HEI: *Health effects of outdoor air pollution in developing countries of Asia – a literature review*. Techn. Ber. 15, The Health Effects Institute International Oversight Committee, Boston, MA, 2004.
- [45] HILGERS, H. und WÖLLENSTEIN, N.: *Der aktuelle Begriff – Feinstaub-Richtlinie*. Techn. Ber. 14, Wissenschaftliche Dienste des deutschen Bundestages, 2005.
- [46] HOFFMANN, T.: *Statistische Modelle zur täglichen Prognose von Feinstaubkonzentrationen Überblick der verfügbaren Systeme, Sensitivitätsstudie und Anwendungsprogrammierung für das System ProFet*. Diplomarbeit, TU Dresden, Fakultät Forst-,Geo- und Hydrowissenschaften, 2007.

- [47] HOOYBERGHS, J., MENSINK, C., DUMONT, G., FIERENS, F. und BRASSEUR, O.: *A neural network forecast for daily average PM10 concentrations in Belgium*. Atmospheric Environment, 39(18):3279–3289, 2005.
- [48] HÖRMANN, S. und STADLOBER, E.: *Modellierung und Prognose der Feinstaubkonzentration in Graz-Mitte für die Zeitperiode Herbst/Winter - Endbericht*. Techn. Ber., Umweltamt Graz, 2004.
- [49] HSU, C., CHANG, C. und LIN, C.: *A Practical Guide to Support Vector Classification*. Techn. Ber., Department of Computer Science – National Taiwan University, 2007.
- [50] JACOB, D., HURK, B. V. DEN, ANDRÆ, U., ELGERED, G., FORTELIUS, C., GRAHAM, L., JACKSON, S., KARSTENS, U., KÖPKEN, C., LINDAU, R., PODZUN, R., ROCKEL, B., RUBEL, F., SASS, B., SMITH, R. und YANG, X.: *A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period*. Meteorology and Atmospheric Physics, 77:19–43, 2001.
- [51] JAKOBS, H., MEMMESHEIMER, M. und EBEL, A.: *Daily Forecast of Air Quality over Europe with the EURAD Model System*. Geophysical Research Abstracts, 7(2), 2005.
- [52] KAMINSKI, W., SKRZYPSKI, J. und STRUMILLO, P.: *Forecasting of air pollution in urban areas by means of artificial neural networks*. In: SUCHAROV, L. und BREBBIA, C. A. (Hrsg.): *Urban Transport VI - Urban transport and the Environment for the 21st Century*, S. 115–124. WIT Press, 2000.
- [53] KATSOUYANNI, K., TOULOUMI, G., SAMOLI, E., GRYPARIS, A., TERTRE, A. L., MONOPOLIS, Y., ROSSI, G., ZMIROU, D., BALLESTER, F., BOUMGHAR, A., ANDERSON, H. R., WOJTYNIAK, B., PALDY, A., BRAUNSTEIN, R., PEKKANEN, J., SCHINDLER, C. und SCHWARTZ, J.: *Confounding and Effect Modification in the Short-Term Effects of Ambient Particles on Total Mortality – Results from 29 European Cities within the APHEA2 Project*. Epidemiology, 12:521–531, 2001.
- [54] KLINGNER, M., SÄHN, E., ANKE, K., HOLST, T., ROST, J., MAYER, H. und AHRENS, D.: *Reduktionspotentiale verkehrsbeschränkender Maßnahmen*

- in Bezug zu meteorologisch bedingten Schwankungen der PM₁₀- und NO_X-Immissionen.* Gefahrstoffe - Reinhaltung der Luft, 66(7/8):326–334, 2006.
- [55] KLINGNER, M., VOIGTLÄNDER, M., ANKE, K. und SÄHN, E.: *PM₁₀ Vorhersagemodell.* Techn. Ber., Fraunhofer Institut für Verkehrs- und Infrastruktursysteme (IVI), Dresden, 2006.
- [56] KOLEHMAINEN, M., MARTIKAINEN, H., HILTUNEN, T. und RUUSKANEN, J.: *Forecasting air quality parameters using hybrid neural network modelling.* Environmental monitoring and Assessment, 65(1-2):277–286, 2000.
- [57] KREWSKI, D., BURNETT, R. T., GOLDBERG, M. S., HOOVER, K., SIEMIATYCKI, J., JERRETT, M., ABRAHAMOWICZ, M. und WHITE, W. H.: *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality.* Techn. Ber., Health Effects Institute (HEI), Cambridge, 2000.
- [58] KUKKONEN, J., PARTANEN, L., KARPPINEN, A., RUUSKANEN, J., JUNNINEN, H., KOLEHMAINEN, M., NISKA, H., DORLING, S. R., CHATTERTON, T., FOXALL, R. und CAWLEY, G.: *Extensive evaluation of neural network models for the prediction of NO₂ and PM₁₀ concentrations, compared with a deterministic modelling system and measurements in central Helsinki.* Atmospheric environment, 37(32):4539–4550, 2003.
- [59] LFU: *Jahresbericht 2000/2001.* Techn. Ber., Landesamt für Umweltschutz Baden-Württemberg, Karlsruhe, 2002.
- [60] LIPPE, W.: *Einführung in SoftComputing – Vorlesungsskript.* Techn. Ber., Institut für Informatik, Westfälische Wilhelms Universität Münster, 2004.
- [61] LOHMEYER, A., DÜRING, I., GIERETH, M., HOFFMANN, T., NICKLASS, D., KLEIN, M., SCHEU-HACHTTEL, H., SÖRGEL, C. und WANNER, L.: *Kurzfrist – Feinstaub - Prognosemodelle für MARQUIS und LÜSA Sachsen-Anhalt.* Gefahrstoffe - Reinhaltung der Luft, 67(7/8), 2007.
- [62] LU, W., FAN, H., LEUNG, A. und WONG, J.: *Analysis of pollutant levels in central Hong Kong applying neural network method with particle swarm optimization.* Environmental Monitoring and Assessment, 79:217–230, 2002.

- [63] LU, W., WANG, W., WANG, X., XU, Z. und LEUNG, A.: *Using improved neural network model to analyze RSP, NO_X and NO₂ levels in urban air in Mong Kok, Hong Kong*. Environmental Monitoring and Assessment, 87(3):235–254, 2003.
- [64] LUBW: *Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg – Luftqualitätsindizes LuQx und LaQx*, 2007. <http://www.lubw.baden-wuerttemberg.de/servlet/is/20146/>.
- [65] MARQUIS: *Report on air quality and weather assessment and interpretation models*. Techn. Ber., MARQUIS Konsortium, 2005.
- [66] MARIK, F., FISCHER, K., MCDONALD, T. und SAMSON, P.: *Comparison of methods for estimating mixing height used during the 1992 Atlanta Field Intensive*. Journal of Applied Meteorology, 34(8):1802–1814, 1995.
- [67] MAYER, H.: *Indizes zur human-biometeorologischen Bewertung der thermischen und lufthygienischen Komponente der Klimas*. Gefahrstoffe - Reinhaltung der Luft, 66(4):165–174, 2006.
- [68] MIERSWA, I., WURST, M., KLINKENBERG, R., SCHOLZ, M. und EULER, T.: *YALE: Rapid Prototyping for Complex Data Mining Tasks*. In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. ACM Press, Philadelphia, 2006.
- [69] MÜLLER, J.: *Source Identification of Airborne Particulate Matter by use of Impactor Measurements*. In: *Proceedings of the 15th Conference of the Aerosol Society*, S. 75–78, Manchester, 2004.
- [70] MÜLLER, K., SMOLA, A., RÄTSCH, G., SCHÖLKOPF, B., KOHLMORGEN, J. und VAPNIK, V.: *Predicting time series with support vector machines*. In: *Proceedings of the International Conference on Artificial Neural Networks*, S. 999. Springer, 1997.
- [71] MUKHERJEE, S., OSUNA, E. und GIROSI, F.: *Nonlinear prediction of chaotic time series using a support vector machine*. In: *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing 7*, S. 511–519, Amelia Island, 1997.

- [72] NETO, J., TORRES, P., FERREIRA, F. und BOAVIDA, F.: *Lisbon Air Quality Forecast Using Statistical Methods*. In: *Proceedings of the AQM conference*, Bd. 1, S. 591–597, Istanbul, 2005.
- [73] PARTANEN, L., KUKKONEN, J., KARPPINEN, A., RUUSKANEN, J., PATA-MA, T., KOLEHMAINEN, M., DORLING, S., FOXALL, R. und CAWLEY, G.: *Inter-Comparison of Neural Network, Statistical and Deterministic Models for Predicting the Concentrations of NO₂ and PM₁₀ in Urban Air*. S. 74–77, 2003.
- [74] PASTOR-BÁRCENAS, O., SORIA-OLIVAS, E., MARTÍN-GUERRERO, J., CAMPS-VALLS, G., CARRASCO-RODRÍGUEZ, J. und VALLE-TASCÓN, S. DEL: *Unibased sensitivity analysis and pruning techniques in neural networks for surface ozone modelling*. *Ecological Modelling*, 182:149–158, 2005.
- [75] PELLICCIONI, A. und POLI, U.: *Use of neural net models to forecast atmospheric pollution*. *Environmental Monitoring and Assessment*, 65(1-2):297–304, 2000.
- [76] PEREZ, P. und REYES, J.: *Predictability of particulate air pollution time series*, 2007. <http://fisica.usach.cl/grupos/ambiente/paper2.pdf>.
- [77] PEREZ, P., TRIER, A. und REYES, J.: *Prediction of PM_{2.5} concentrations several hours in advance using neural networks in Santiago, Chile*. *Atmospheric Environment*, 34(8):1189–1196, 2000.
- [78] PESCH, M.: *Beitrag zur Ursachenanalyse der Feinstaubbelastung in Berlin*. Techn. Ber., Technische Universität Berlin, 2006.
- [79] PETERS, A., SCHULZ, H., KREYLING, W. G. und WICHMANN, H. E.: *Staub und Staubinhaltsstoffe – Feine und ultrafeine Partikel*. In: WICHMANN, H. E., SCHLIPKOETER, H. W. und FÜLGRAFF, G. (Hrsg.): *Handbuch der Umweltmedizin*. Ecomed, 14. Aufl., 1998.
- [80] PFEIFFER, H.: *Methode zur Prognose von Luftverunreinigungen im Bereich Verkehrsreicher Straßen*. Diplomarbeit, Universität Stuttgart, 2001.
- [81] PFEIFFER, H.: *Neural Modelling of the Spatial Distribution of Air Pollutants. A new method developed considering as example Cyprus*. Dissertation, Universität Stuttgart, 2005.

-
- [82] PIRINGER, M., BAUMANN, K. und LANGER, M.: *Summertime mixing heights at Vienna, Austria, estimated from vertical soundings and by a numerical model*. *Boundary Layer Meteorol*, 89(1):25–45, 1998.
- [83] POKROVSKI, O. und KWOK, R.: *Modelling of meteorological factor impacts on urban air pollution species by fuzzy logic approach*. In: SPORTISSE, B. (Hrsg.): *Air Pollution Modelling and Simulation*, S. 513–526. Springer, 2002.
- [84] POPE, C., BURNETT, R., THUN, M., CALLE, E., KREWSKI, D., ITO, K. und THURSTON, G.: *Lung Cancer, Cardiopulmonary Mortality and Long-term Exposure to Fine Particulate Air Pollution*. *The Journal of the American Medical Association*, 287(9):1132–1141, 2002.
- [85] POPE, C., SCHWARTZ, J. und RANSOM, M.: *Daily mortality and PM10 pollution in Utah Valley*. *Archives of Environmental Health*, 47(3):211–217, 1992.
- [86] POPE, C., THUN, M., NAMBOODIRI, M., DOCKERY, D., EVANS, J., SPEIZER, F. und HEATH, C.: *Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults*. *American journal of respiratory and critical care medicine*, 151(3):669–674, 1995.
- [87] PREGGER, T.: *Ermittlung und Analyse der Emissionen zur Minderung primärer anthropogener Feinstäube in Deutschland*. Dissertation, Universität Stuttgart, 2006.
- [88] RAIMONDO, G., MONTUORI, A., MONIACI, W., PASERO, E. und ALMKVIST, E.: *A Machine Learning Tool to Forecast PM10 Level*. In: *Proceedings of the 5th Conference on Artificial Intelligence Applications to Environmental Science*, 2007.
- [89] RAIZENNE, M., NEAS, L., DAMOKOSH, A., DOCKERY, D., SPENGLER, J., KOUTRAKIS, P., WARE, J. und SPEIZER, F.: *Health effects of acid aerosols on North American children – pulmonary function*. *Environmental health perspectives*, 104(5):506–514, 1996.
- [90] ROJAS, R.: *Theorie der neuronalen Netze – Eine systematische Einführung*. Springer, Berlin, 4. Aufl., 1996.

- [91] SACHS, L.: *Angewandte Statistik – Anwendung statistischer Methoden*. Springer, Berlin, 9. Aufl., 1999.
- [92] SAMET, J., ZEGER, S., KELSALL, J., XU, J. und KALKSTEIN, L.: *Does weather confound or modify the association of particulate air pollution with mortality? – An analysis of the Philadelphia data, 1973–1980*. Environmental Research, 77:9–19, 1998.
- [93] SCHÖNEBURG, E., HEINZMANN, F. und FEDDERSEN, S.: *Genetische Algorithmen und Evolutionsstrategien*. Addison-Wesley, Bonn, 1994.
- [94] SCHRODIN, R.: *Quarterly Report of the German NWP-System - Part 2: Description of the german NWP-System*. Techn. Ber., Deutscher Wetterdienst, Offenbach a. M., 2004.
- [95] SCHWARTZ, J.: *Particulate air pollution and chronic respiratory disease*. Environmental Research, 62:7–13, 1993.
- [96] SCHWARTZ, J. und DOCKERY, D. W.: *Increased mortality in Philadelphia associated with daily air pollution concentrations*. American Review of Respiratory Disease, 145:600–604, 1992.
- [97] SCHWARTZ, J. und DOCKERY, D. W.: *Particulate air pollution and daily mortality in Steubenville, Ohio*. American Journal of Epidemiology, 125:12–19, 1992.
- [98] SEATON, A., MACNEE, W., DONALDSON, K. und GOODEN, D.: *Particulate air pollution and acute health effects*. Lancet, 345(8943):167–178, 1995.
- [99] SINGH, D., GANJU, A. und SINGH, A.: *Weather prediction using nearest-neighbour model*. Current Science, 88(8):1283–1289, 2005.
- [100] SINGH, S.: *Fuzzy nearest neighbour method for time-series forecasting*. In: *Proc. 6th European Congress on Intelligent Techniques and Soft Computing (EU-FIT'98), Aachen, Germany,, Bd. 3, S. 1901–1905*, 1998.
- [101] SLOSS, L.: *The importance of PM10/2.5 emissions*. Techn. Ber., IEA Clean Coal Centre, 2004.
- [102] SMOLA, A.: *Lernen mit Kernen*. Dissertation, Universität Berlin, 1998.

-
- [103] SMOLA, A. und SCHÖLKOPF, B.: *A Tutorial on Support Vector Regression*. Techn. Ber., NeuroCOLT, 1998.
- [104] SPIX, C. und WICHMANN, H.E.: *Daily mortality and air pollutants: findings from Köln, Germany*. Journal of Epidemiology and Community Health, 50(1):52–58, 1996.
- [105] SÖRGE, C.: *Development of a PM₁₀ concentration forecast model by multi-regression summary for LUBW monitoring station Freudenstadt*. Techn. Ber., Ingenieurbüro Lohmeyer, Radebeul, 2006.
- [106] STATISTISCHES LANDESAMT BADEN-WÜRTTEMBERG: *Gesamt-Staub- und Feinstaub(PM10)-Emissionen in Baden-Württemberg 1990 bis 2004 nach Verursachersektoren*, 2007. <http://www.statistik-portal.de/UmweltVerkehr/Landesdaten/11a07.asp>.
- [107] STEISLINGER, B.: *Einfluß von Temperaturinversionen auf Konzentration und Verteilung von Luftverunreinigungen*. VDI, Düsseldorf, 1994.
- [108] SUNYER, J., CASTELLSAGUE, J., SAEZ, M., TOBIAS, A. und ANTO, J. M.: *Air pollution and mortality in Barcelona*. Journal of Epidemiology and Community Health, 50(1):76–80, 1996.
- [109] TURIAS, I., GONZÁLEZ, F. und GALINDO, P.: *Application of neural techniques to the modelling of time-series of atmospheric pollution data in the 'Campo de Gibraltar' region*. In: *Procc. of the 8th International Conference on Engineering Application of Neural Network (EANN'03)*, S. 9–16, 2003.
- [110] UBA: *Homepage des Umweltbundesamts Deutschland*, 2007. <http://www.env-it.de/luftdaten/start.fwd>.
- [111] UMWELTBUNDESAMT ÖSTEREICH: *Homepage des Umweltbundesamts Österreich*, 2007. <http://www.umweltbundesamt.at/umweltschutz/luft/luftschadstoffe/staub/>.
- [112] VAPNIK, V. N.: *Statistical Learning Theory*. John Wiley & Sons, New York, 1998.
- [113] VERHOEFF, A., HOEK, G. und WIJEN, H. V.: *Air pollution and daily mortality in Amsterdam*. Epidemiology, 7:225–230, 1996.

- [114] VESTASERVICES: *QnetV2000 – Neural Network Modeling for Windows*, 2007.
<http://www.qnetv2k.com>.
- [115] WANG, W., XU, Z. und LU, W.: *Three improved neural network models for air quality forecasting*. *Engineering Computations*, 20(2):192–210, 2003.
- [116] WANNER, L., NICKLASS, D., BOHNET, N., N. B. A., BRONDER, J., FERREIRA, F., FRIEDRICH, R., KARPPINEN, A., LAREAU, F., LOHMEYER, A., PANIGHI, A., PARISIO, S., SCHEU-HACHTEL, H. und SERPA, J.: *From Measurement Data to Environmental Information - MARQUIS - A Multimodal AiR Quality Information Service for the General Public*. In: SWAYNE, A. und HREBICEK, J. (Hrsg.): *Environmental Software Systems – Dimensions of Environmental Informatics*, Bd. 7, 2007.
- [117] WHO: *Air Quality Guidelines for Europe 2nd Edition*. World Health Organisation, 2000.
- [118] WHO: *Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide – Global update 2005 – Summary of risk assessment*. Techn. Ber., World Health Organisation, 2006.
- [119] WILLMOTT, C.: *On the validation of models*. *Physical geography*, 2:184–194, 1981.
- [120] WILLMOTT, C., ACKLESON, S., DAVIS, R., FEDDEMA, J., KLINK, K., LEGATES, D., DONNELL, J. und ROWE, C.: *Statistics for the evaluation and comparison of models*. *Journal of Geophysical Research*, 90:8995–9005, 1985.
- [121] WITTEN, I. und FRANK, E.: *Data Mining – Praktische Werkzeuge und Techniken für das maschinelle Lernen*. Hanser, München, 2000.
- [122] ZICKUS, M., GREIG, A. und NIRANJAN, M.: *Comparison of Four Machine Learning Methods for Predicting PM10 Concentrations in Helsinki, Finland*. *Water, Air and Soil Pollution: Focus*, 2(5-6):717–729, 2002.

Anhang A

Vergleich der EURAD Prognosen für PM_{10} mit Messdaten der LUBW

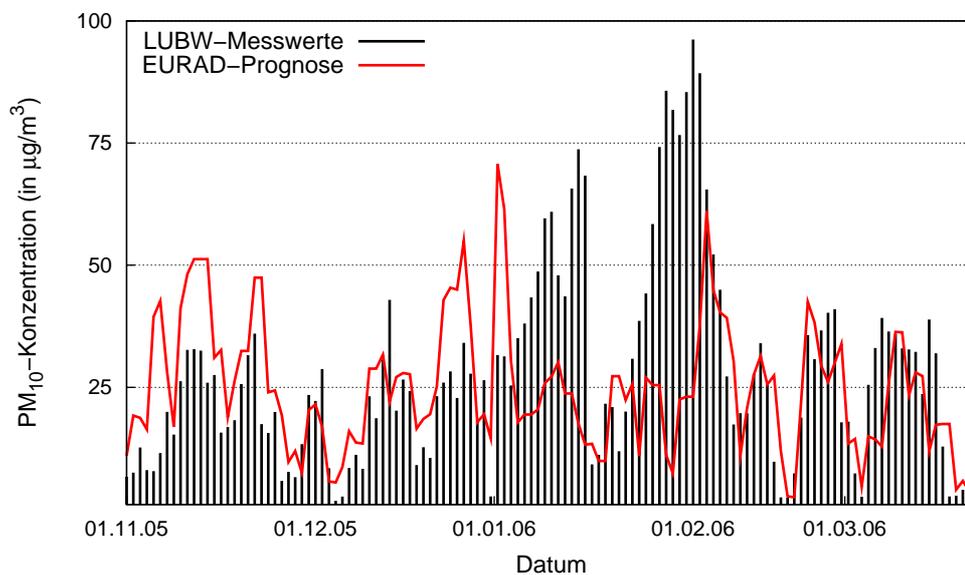


Abbildung A.1: Vergleich der EURAD 12-h-Prognose für PM_{10} mit den Messwerten der Station Freiburg-Mitte

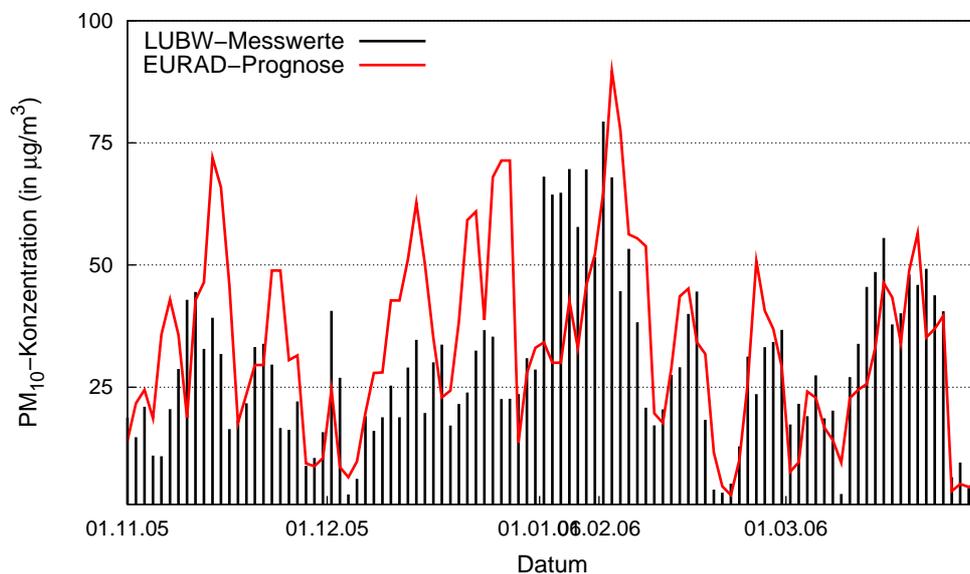


Abbildung A.2: Vergleich der EURAD 12-h-Prognose für PM₁₀ mit den Messwerten der Station Karlsruhe-Nordwest

Tabelle A.1: Leistungskennzahlen der 12-h-Prognose für PM₁₀ des EURAD Modells für verschiedene Stationen

		BC	FR-M	KA-NW	S-BC
Korrelationskoeffizient (r)		0,37	0,26	0,55	0,51
Mean Absolute Error (MAE)	µg/m ³	14,8	14,4	13,0	12,9
Root Mean Squared Error (RMSE)	µg/m ³	21,7	21,5	17,6	16,9
Index of Agreement (IA)		0,57	0,51	0,72	0,69
False Alarm Rate (FAR)	%	100	86	79	78
Probability of Detection (POD)	%	0	7	36	11
LuQx richtig prognostiziert	%	37	37	43	31

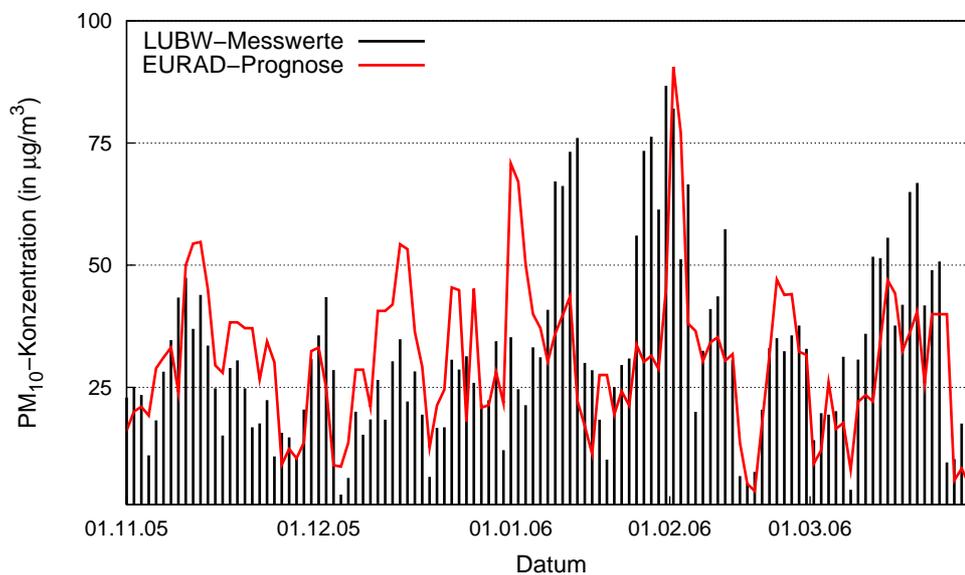


Abbildung A.3: Vergleich der EURAD 12-h-Prognose für PM₁₀ mit den Messwerten der Station Stuttgart-Bad Cannstatt

Anhang B

Detaillierte Ergebnisse der Optimierung der Attributzusammenstellung

ANHANG B. DETAILLIERTE ERGEBNISSE DER OPTIMIERUNG DER
ATTRIBUTZUSAMMENSTELLUNG

Tabelle B.1: Optimierte Attributzusammenstellung nach Modellen für die Station Biberach

	LR	IBk	MLP	SVM
Wochentag	—	—	—	●
Wochentag gewichtet	—	—	●	—
Tagtyp	●	—	●	—
Windrichtung \emptyset	●	—	—	—
Windgeschwindigkeit \emptyset	●	●	●	●
Taupunkttemperatur \emptyset	—	●	●	●
Temperatur \emptyset	●	—	●	—
Temperatur-Maximum	●	—	●	●
Δ Temp.-Max./Min.	—	—	—	●
Globalstrahlung	●	—	—	—
Niederschlagssumme	●	●	●	●
PM ₁₀ Vortag \emptyset	●	●	●	●
NO ₂ Vortag \emptyset	—	—	●	—
Tage ohne Niederschlag	●	—	●	●
Tage mit Niederschlag	—	●	●	—
Mischungsschichthöhe \emptyset	●	—	—	—
Mischungsschicht <250m	—	—	●	—

ANHANG B. DETAILLIERTE ERGEBNISSE DER OPTIMIERUNG DER
ATTRIBUTZUSAMMENSTELLUNG

Tabelle B.2: Optimierte Attributzusammenstellung nach Modellen für die Station
Freiburg-Mitte

	LR	IBk	MLP	SVM
Tagtyp	●	—	●	●
Windrichtung \emptyset	●	●	—	—
Windgeschwindigkeit \emptyset	●	●	●	●
Taupunkttemperatur \emptyset	●	●	●	●
Temperatur \emptyset	●	●	●	●
Temperatur-Maximum	●	●	●	—
Temperatur-Minimum	—	●	●	●
Niederschlagssumme	●	●	●	●
PM ₁₀ Vortag \emptyset	●	●	●	●
NO ₂ Vortag \emptyset	—	—	●	—
Tage ohne Niederschlag	●	—	●	●
Tage mit Niederschlag	—	—	●	—
relative Luftfeuchtigkeit \emptyset	—	—	●	—
Mischungsschichthöhe \emptyset	●	—	—	—
Mischungsschicht <200m	●	—	—	—

ANHANG B. DETAILLIERTE ERGEBNISSE DER OPTIMIERUNG DER
ATTRIBUTZUSAMMENSTELLUNG

Tabelle B.3: Optimierte Attributzusammenstellung nach Modellen für die Station Stuttgart-Bad Cannstatt

	LR	IBk	MLP	SVM
Tagtyp	●	—	●	●
Windrichtung \emptyset	●	—	—	—
Windgeschwindigkeit \emptyset	●	●	●	●
Taupunkttemperatur \emptyset	—	—	●	●
Temperatur \emptyset	●	—	●	●
Temperatur-Maximum	●	—	—	—
Temperatur-Minimum	—	●	●	—
Δ Temp.-Max./Min.	—	—	●	●
Globalstrahlung	—	●	●	—
Niederschlagssumme	●	●	●	●
PM ₁₀ Vortag \emptyset	●	●	●	●
NO ₂ Vortag \emptyset	●	●	●	—
Tage ohne Niederschlag	●	●	●	●
Tage mit Niederschlag	—	—	●	—
relative Luftfeuchtigkeit \emptyset	—	—	●	—
Mischungsschicht < 150m	●	—	—	—
Mischungsschicht < 300m	●	—	—	—
Mischungsschicht < 350m	—	—	—	●
Mischungsschicht < 400m	—	—	—	●

Anhang C

Detaillierte Ergebnisse der PM₁₀-Prognose

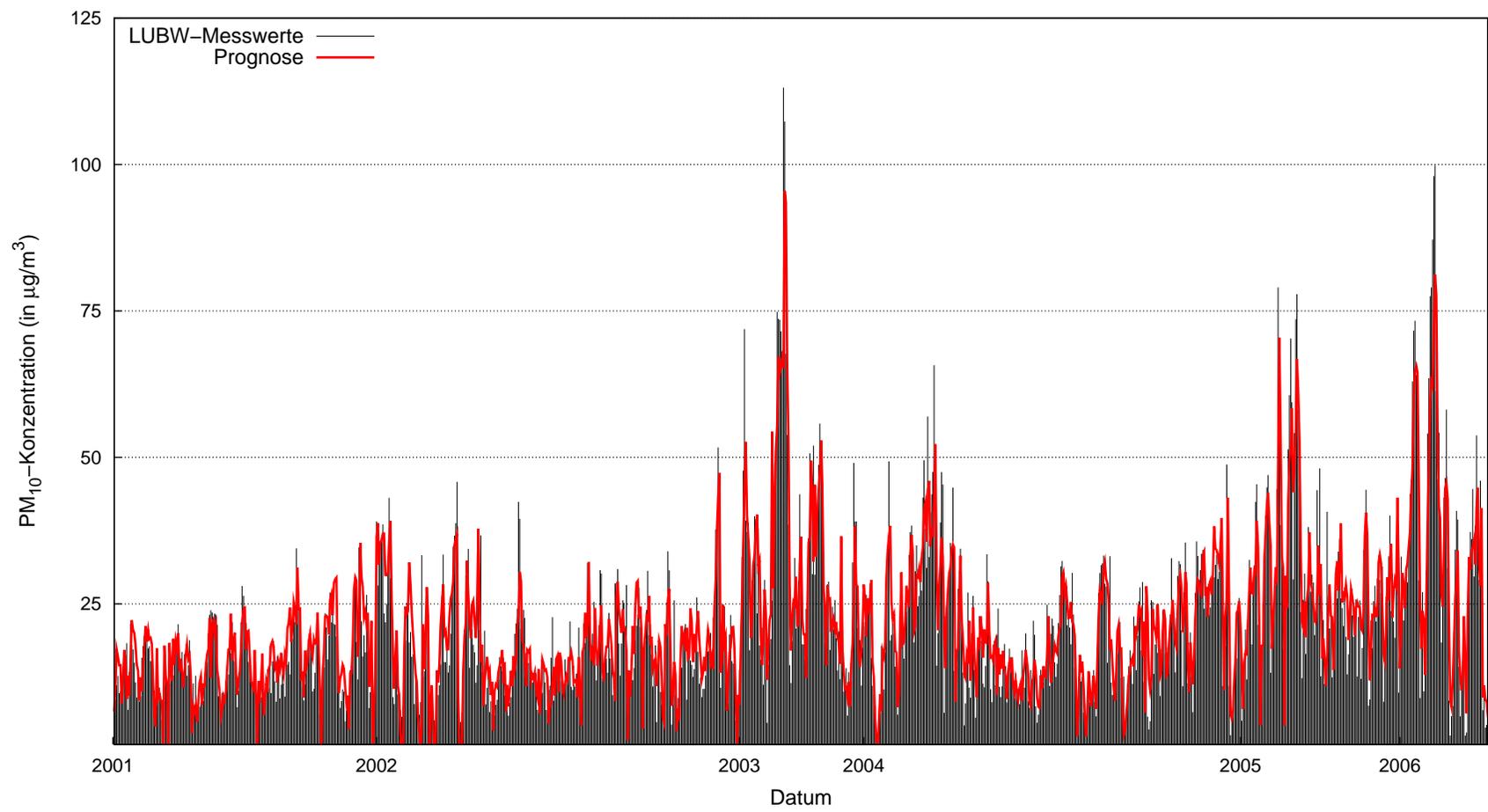


Abbildung C.1: Vergleich der PM₁₀-Prognose mittels linearer Regression mit den Messwerten der Station Biberach

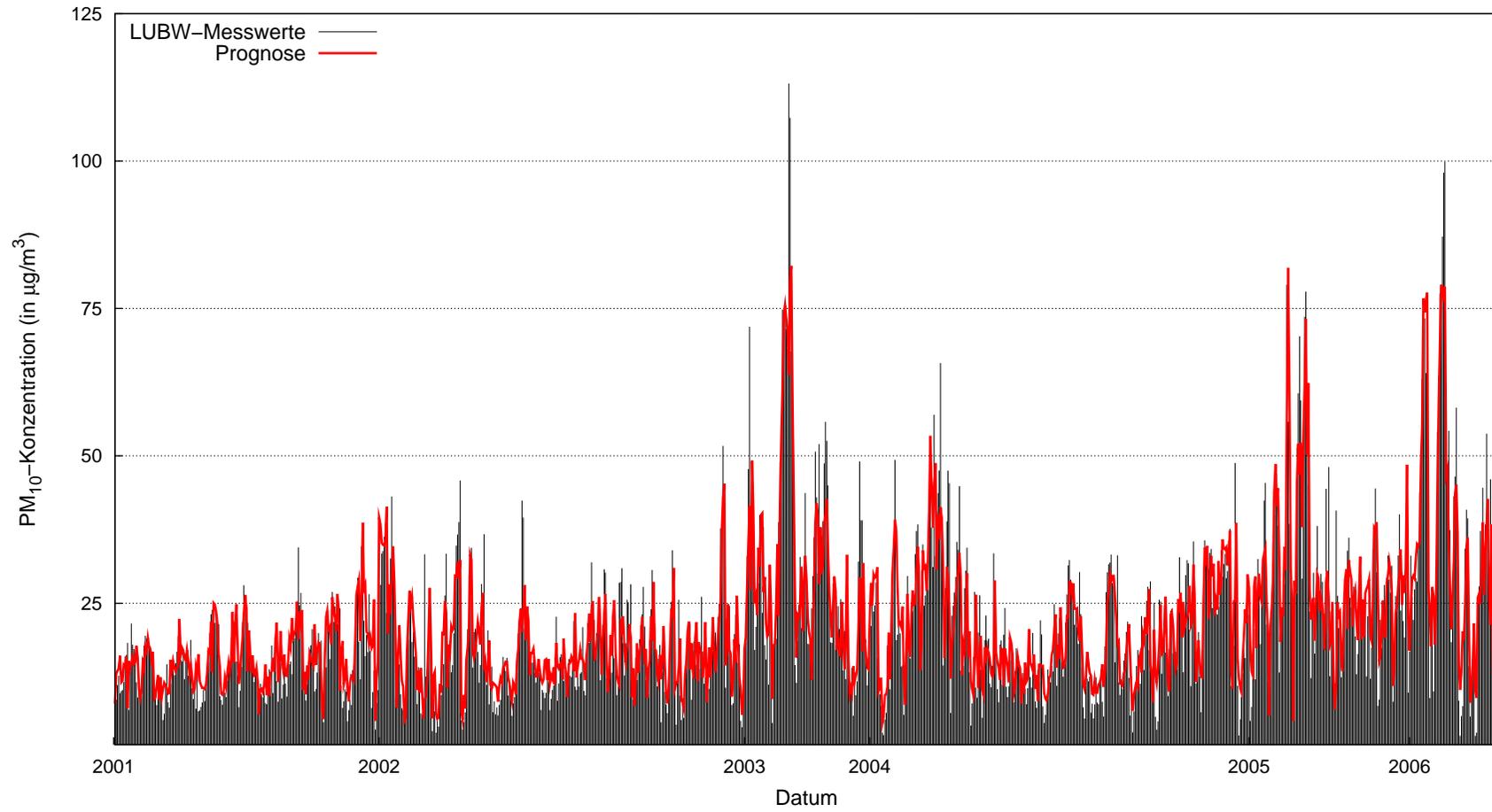


Abbildung C.2: Vergleich der PM₁₀-Prognose mittels der k-Nearest-Neighbour Methode mit den Messwerten der Station Biberach

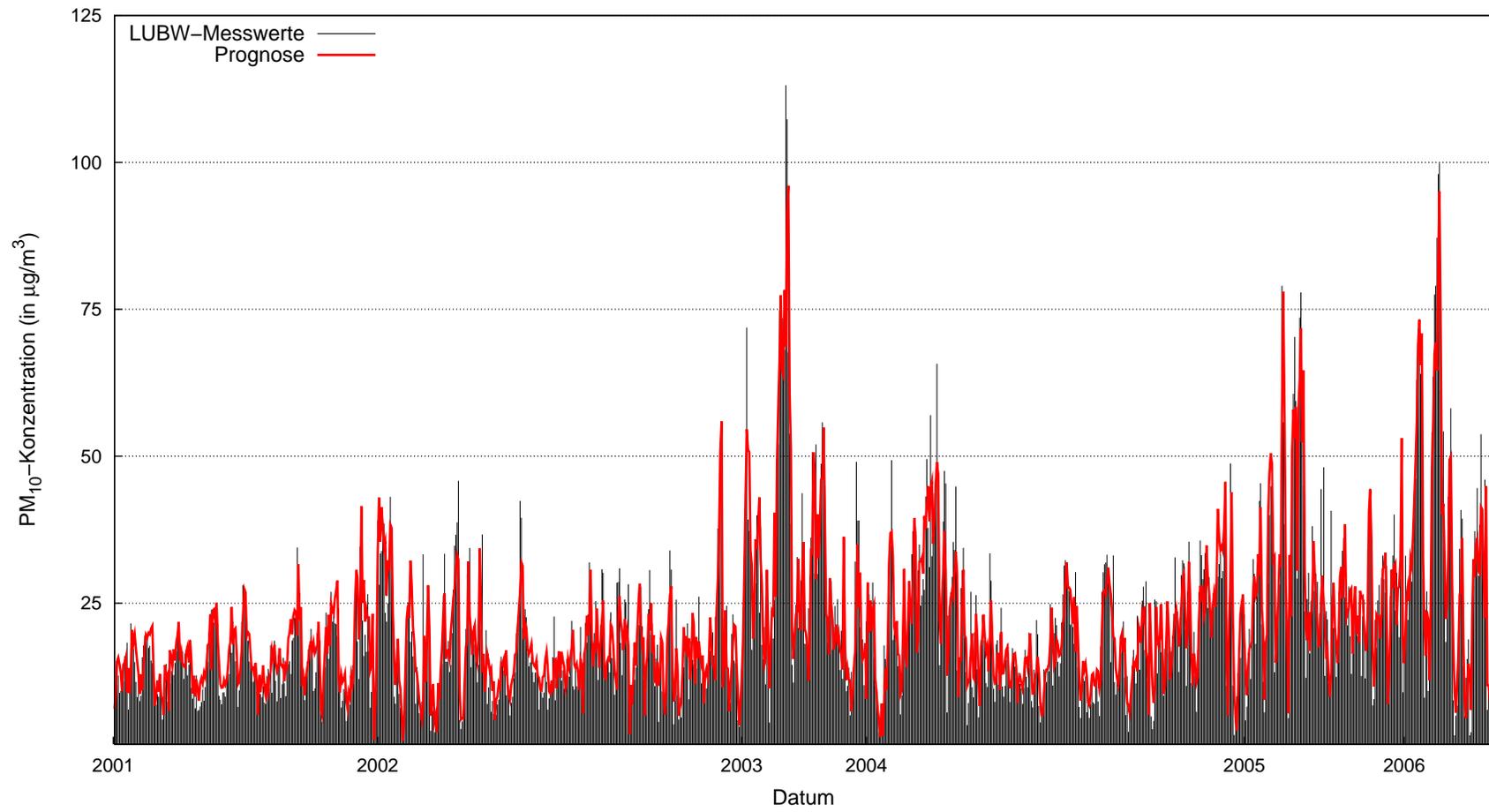


Abbildung C.3: Vergleich der PM₁₀-Prognose mittels Multilayer-Perzepton mit den Messwerten der Station Biberach

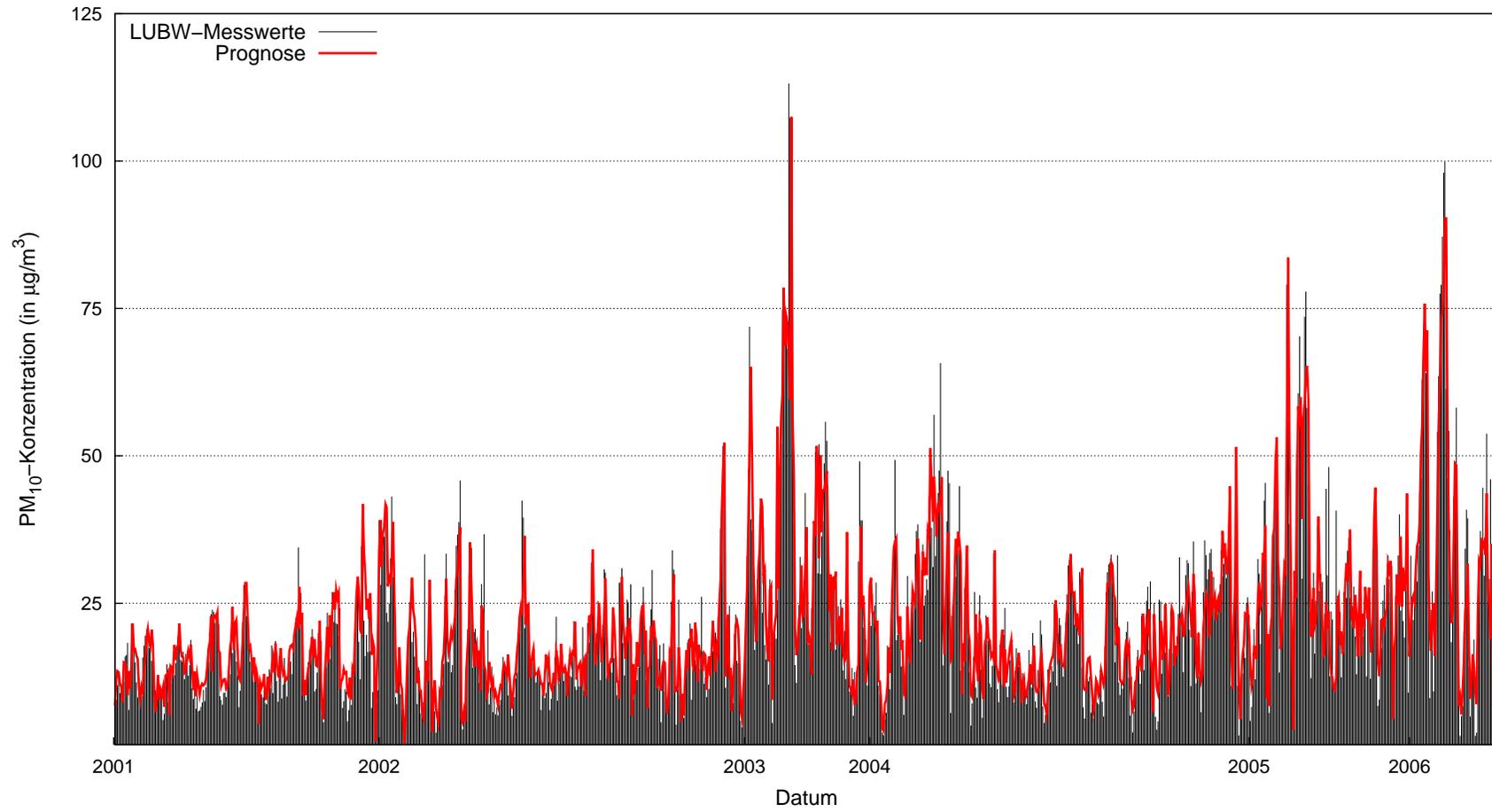


Abbildung C.4: Vergleich der PM₁₀-Prognose mittels Support-Vektor-Maschine mit den Messwerten der Station Biberach

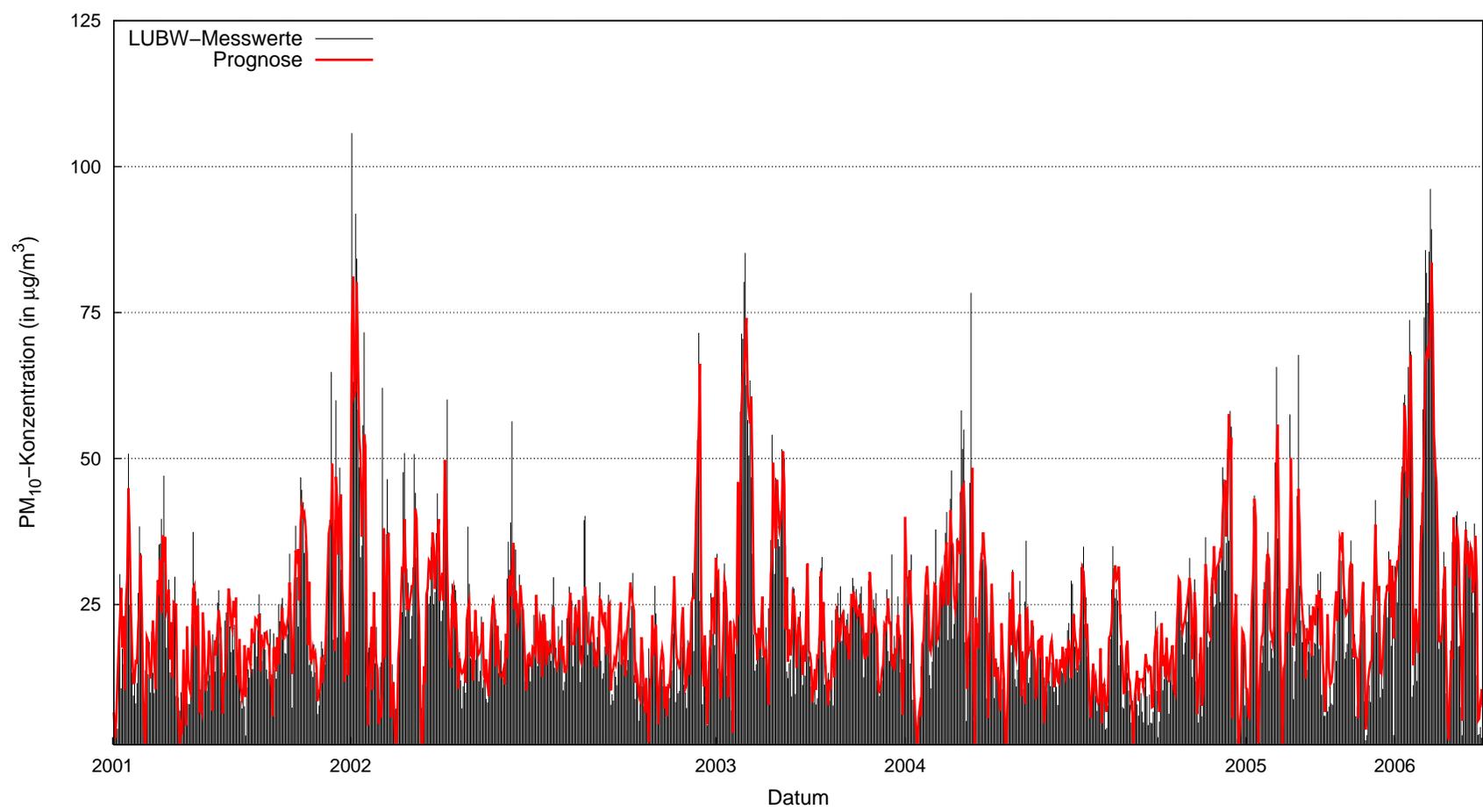


Abbildung C.5: Vergleich der PM₁₀-Prognose mittels linearer Regression mit den Messwerten der Station Freiburg-Mitte

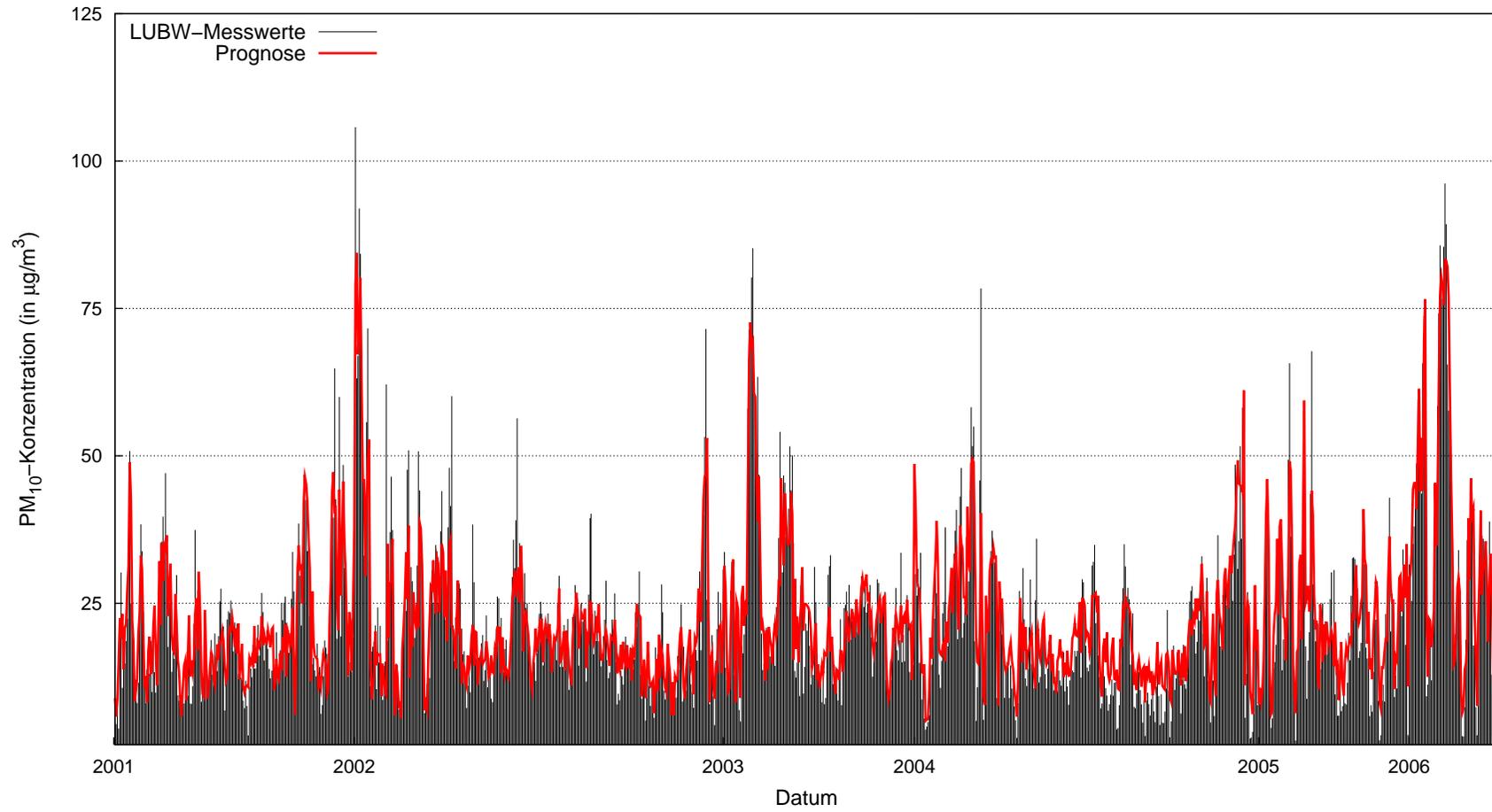


Abbildung C.6: Vergleich der PM₁₀-Prognose mittels der k-Nearest-Neighbour Methode mit den Messwerten der Station Freiburg-Mitte

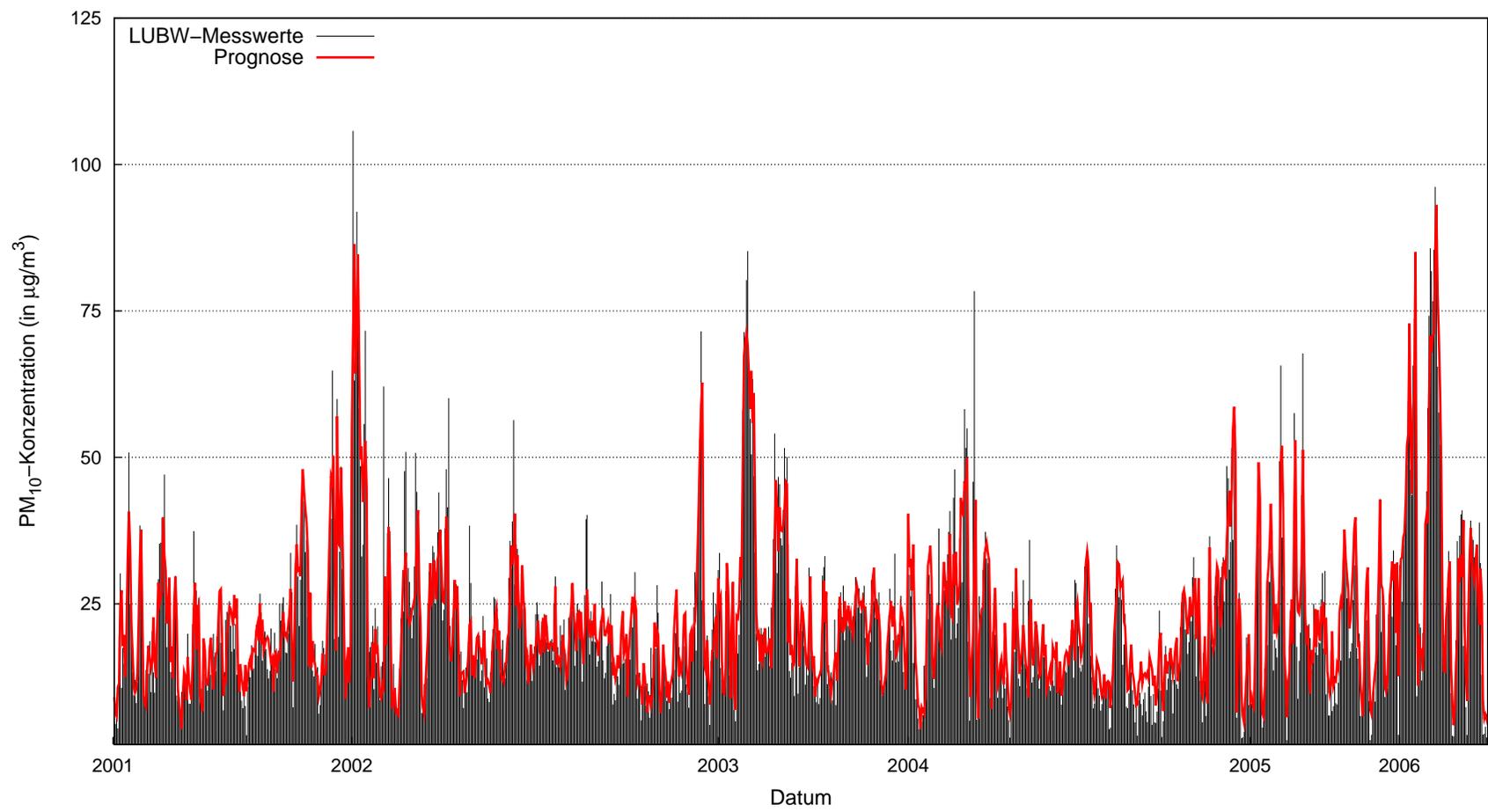


Abbildung C.7: Vergleich der PM₁₀-Prognose mittels Multilayer-Perzeptron mit den Messwerten der Station Freiburg-Mitte

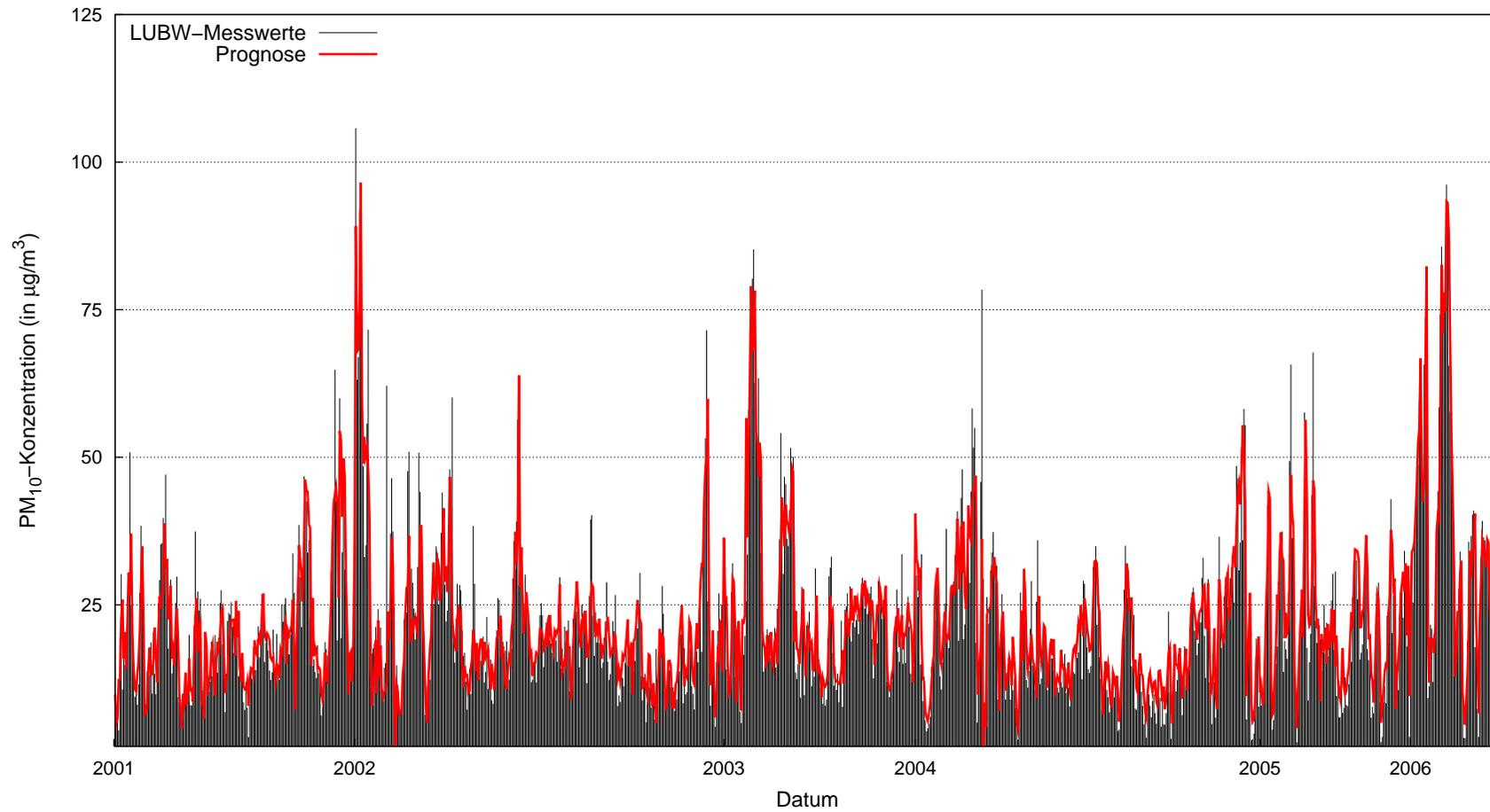


Abbildung C.8: Vergleich der PM₁₀-Prognose mittels Support-Vektor-Maschine mit den Messwerten der Station Freiburg-Mitte

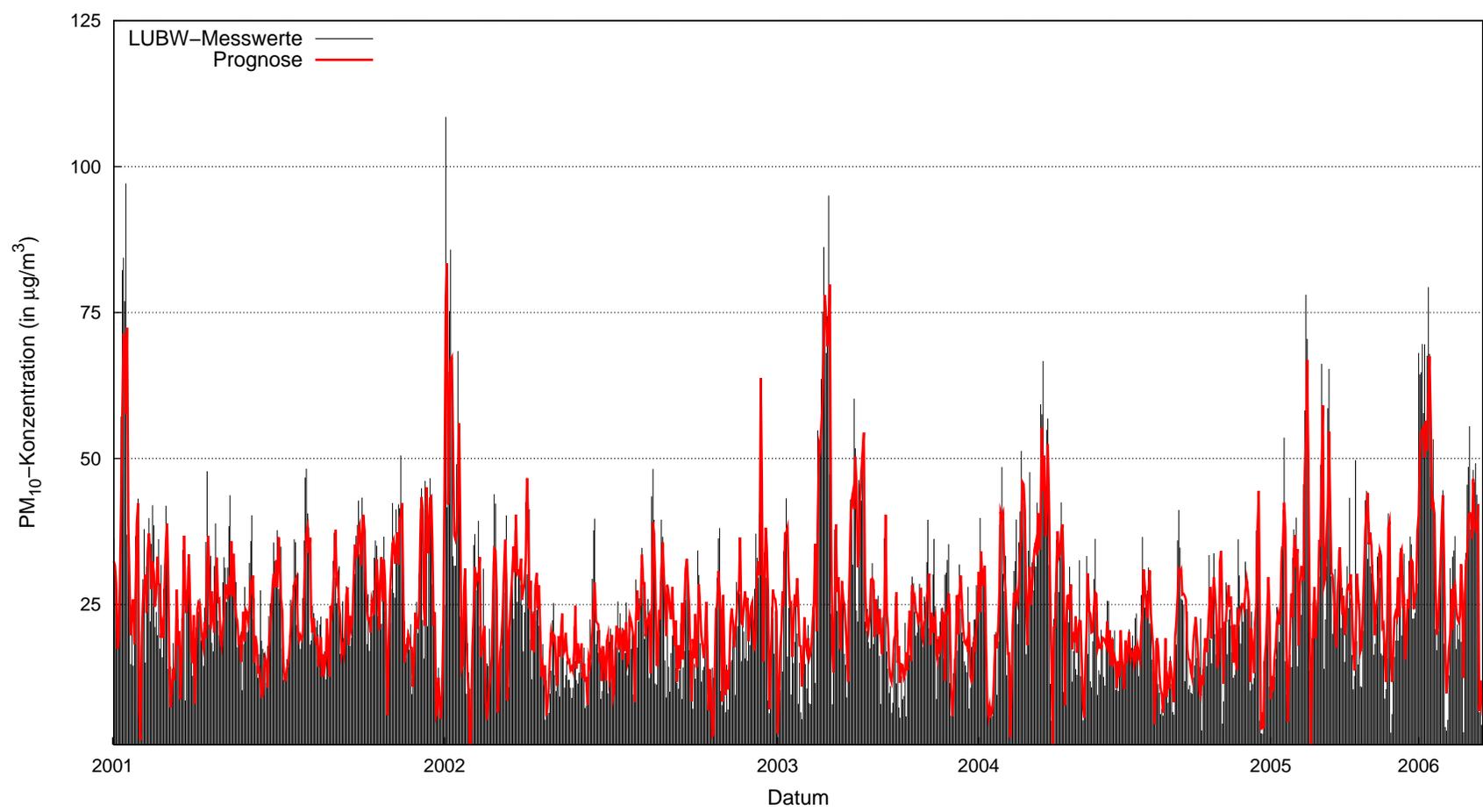


Abbildung C.9: Vergleich der PM₁₀-Prognose mittels linearer Regression mit den Messwerten der Station Karlsruhe-Nordwest

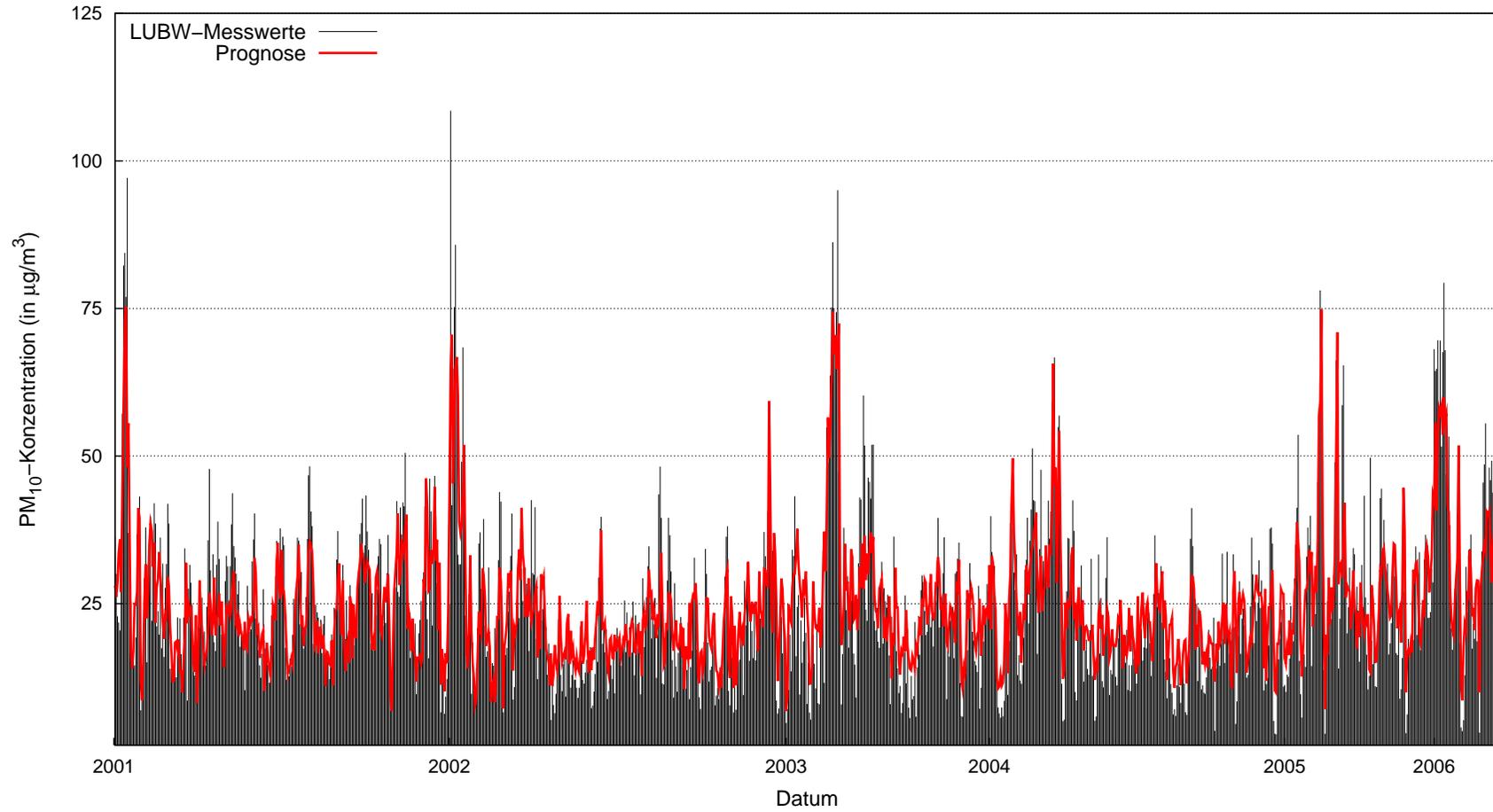


Abbildung C.10: Vergleich der PM₁₀-Prognose mittels der k-Nearest-Neighbour Methode mit den Messwerten der Station Karlsruhe-Nordwest

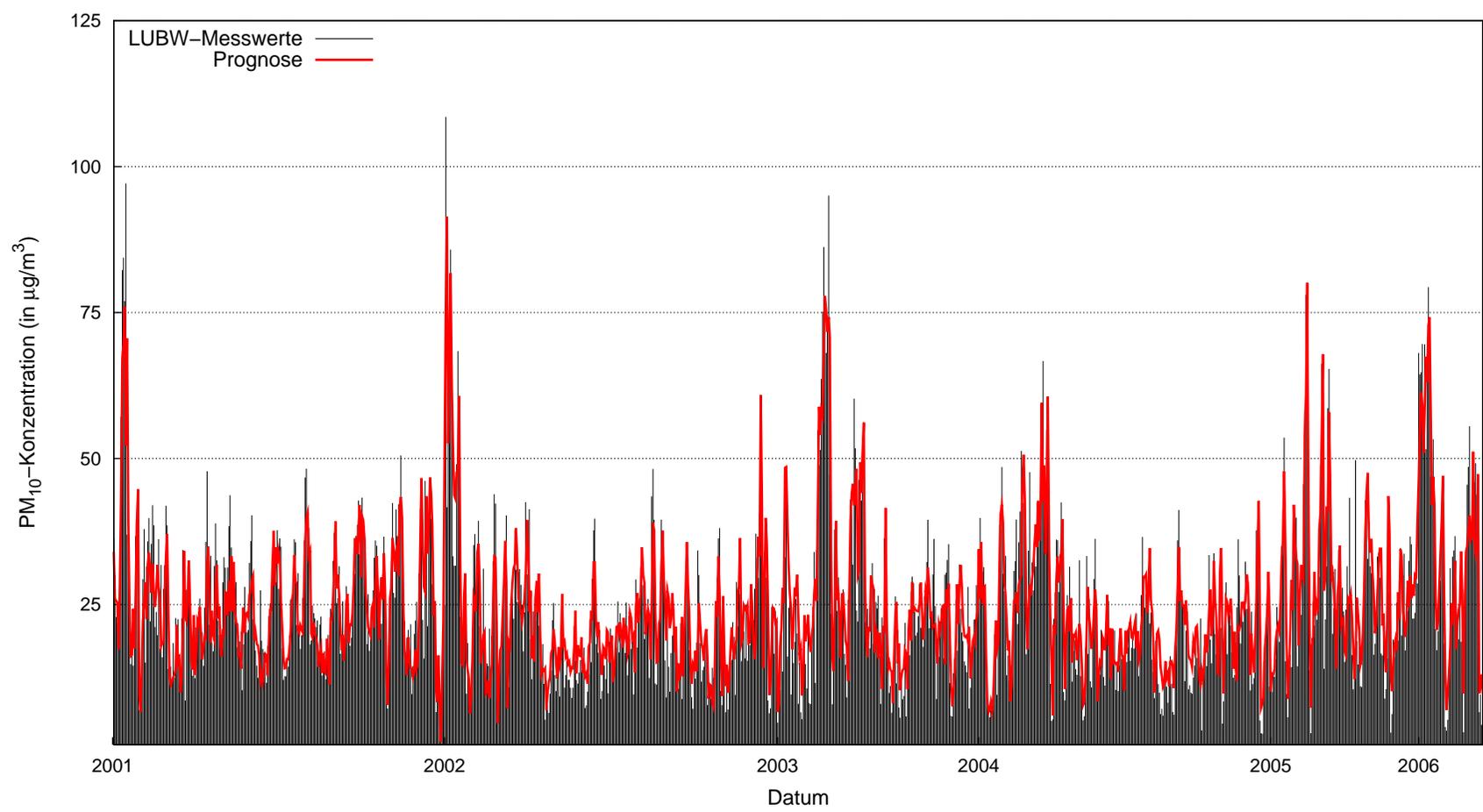


Abbildung C.11: Vergleich der PM₁₀-Prognose mittels Multilayer-Perzeptron mit den Messwerten der Station Karlsruhe-Nordwest

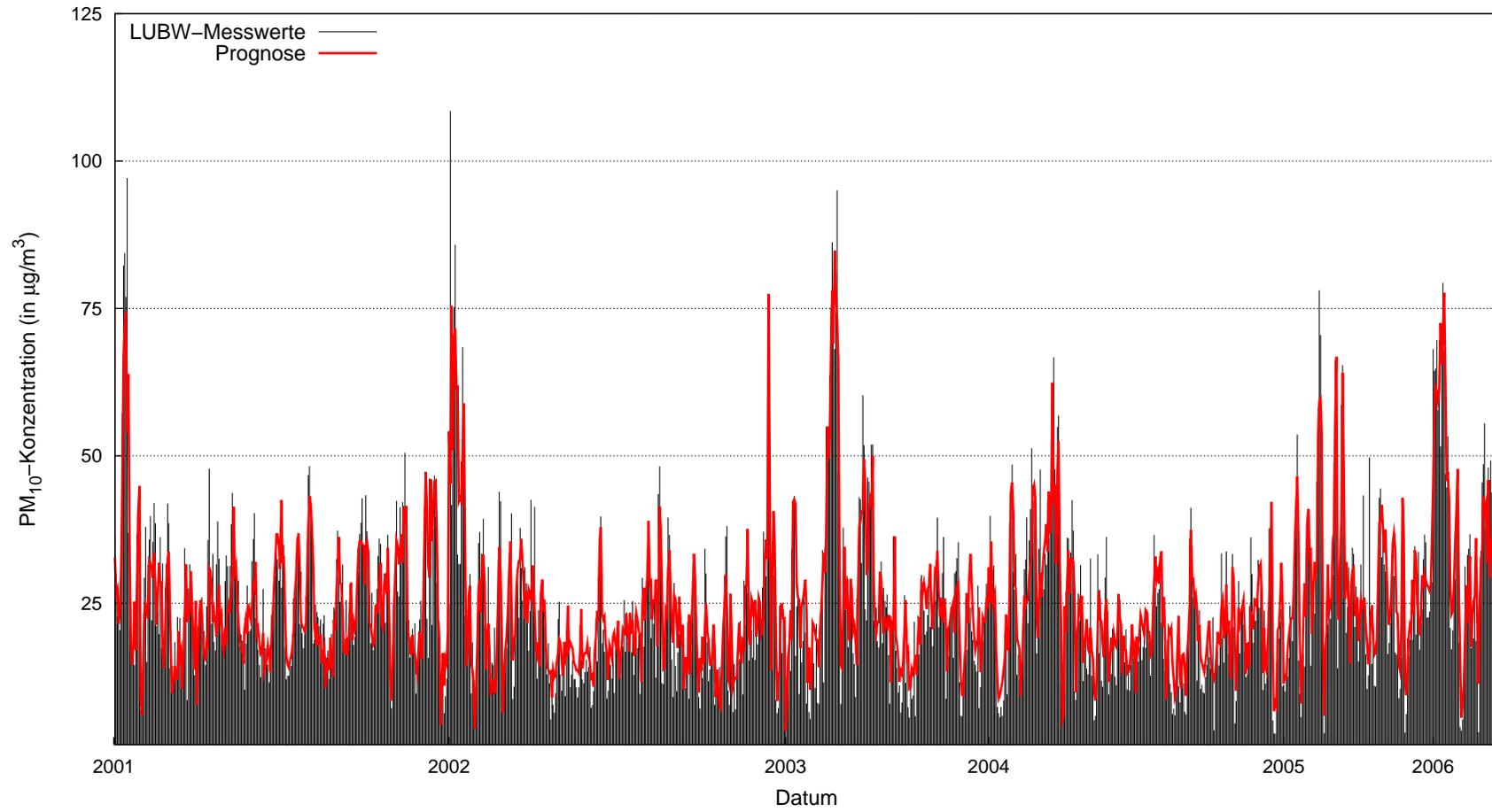


Abbildung C.12: Vergleich der PM₁₀-Prognose mittels Support-Vektor-Maschine mit den Messwerten der Station Karlsruhe-Nordwest

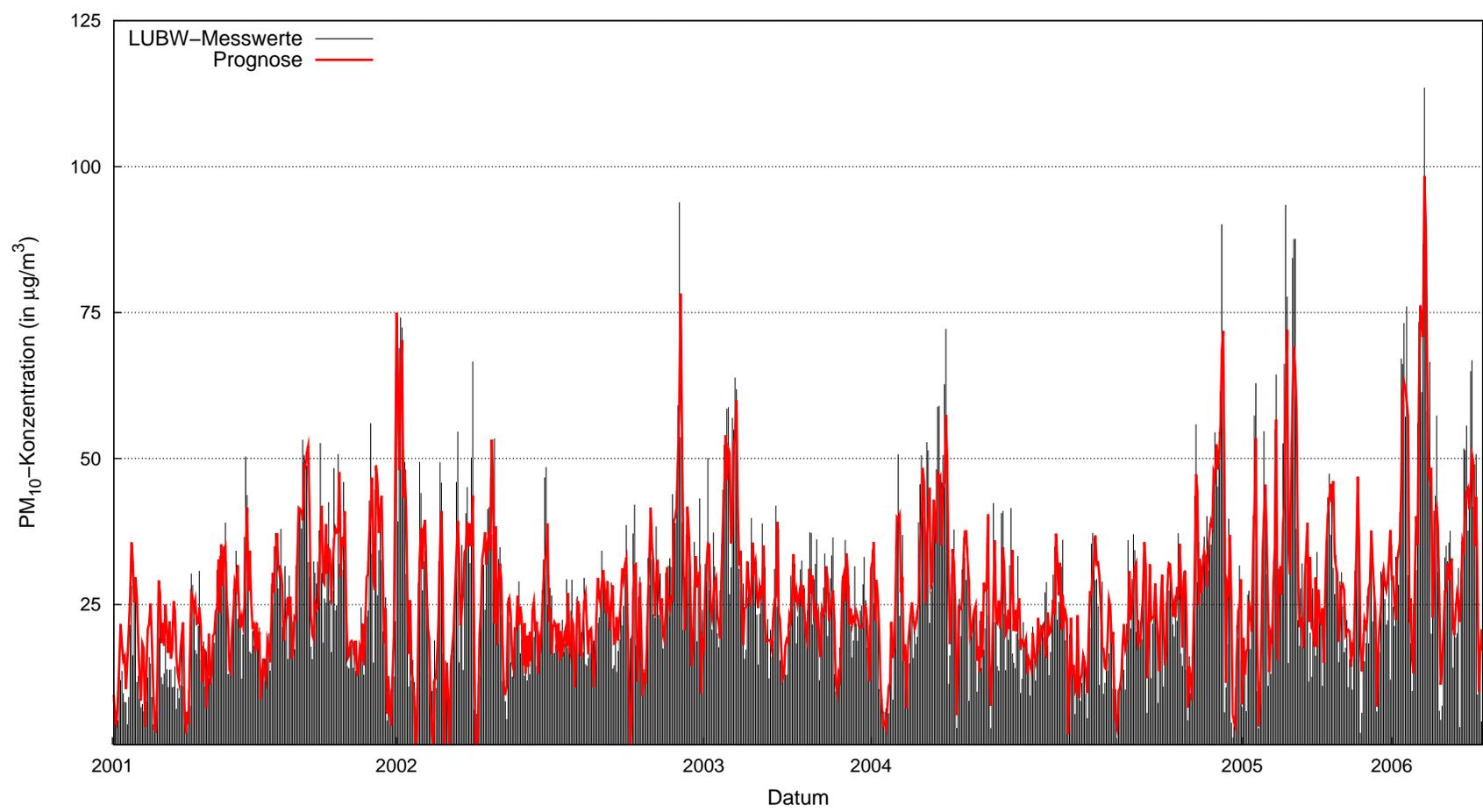


Abbildung C.13: Vergleich der PM₁₀-Prognose mittels linearer Regression mit den Messwerten der Station Stuttgart-Bad Cannstatt

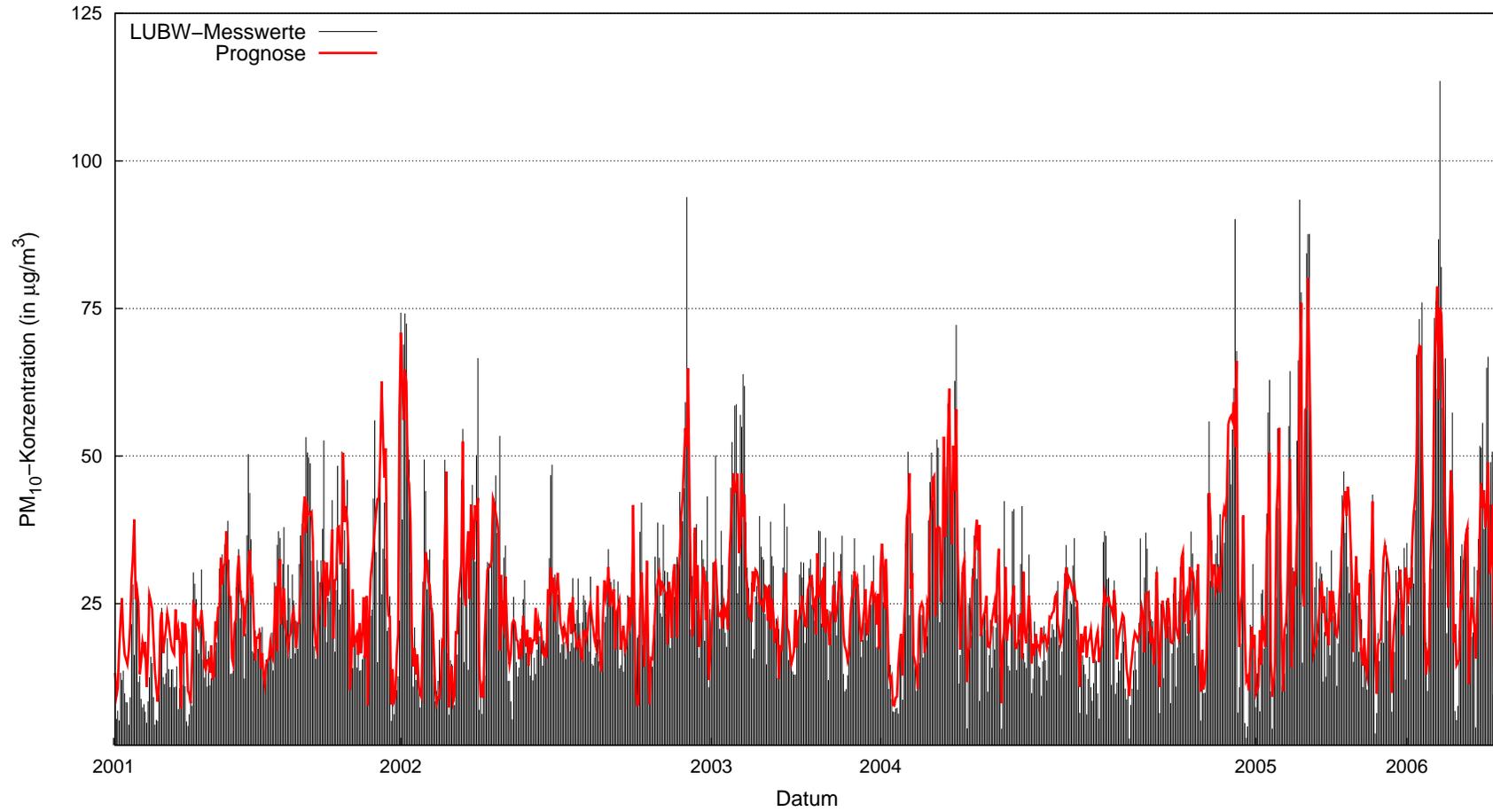


Abbildung C.14: Vergleich der PM₁₀-Prognose mittels der k-Nearest-Neighbour Methode mit den Messwerten der Station Stuttgart-Bad Cannstatt

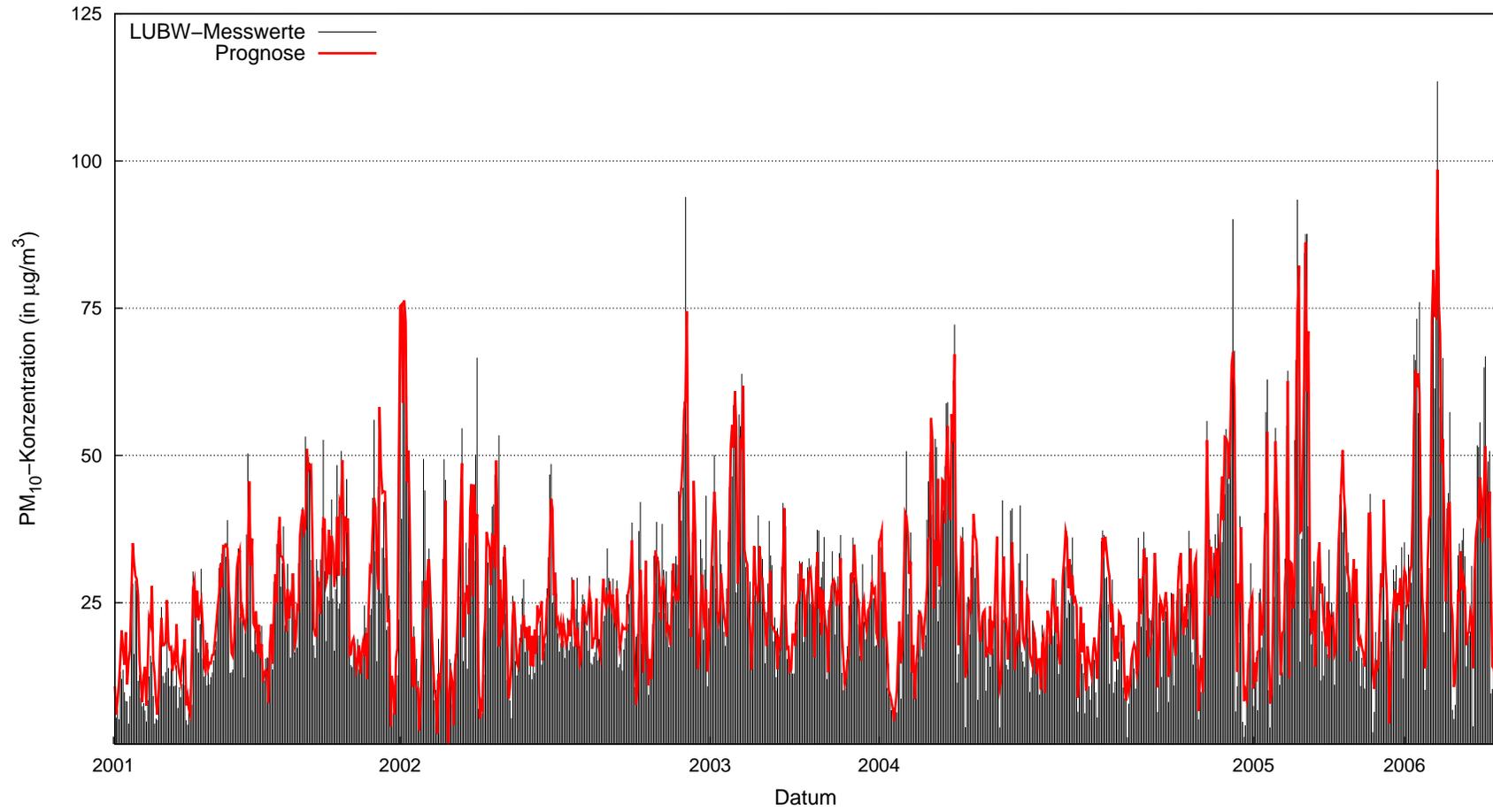


Abbildung C.15: Vergleich der PM₁₀-Prognose mittels Multilayer-Perzeptron mit den Messwerten der Station Stuttgart-Bad Cannstatt

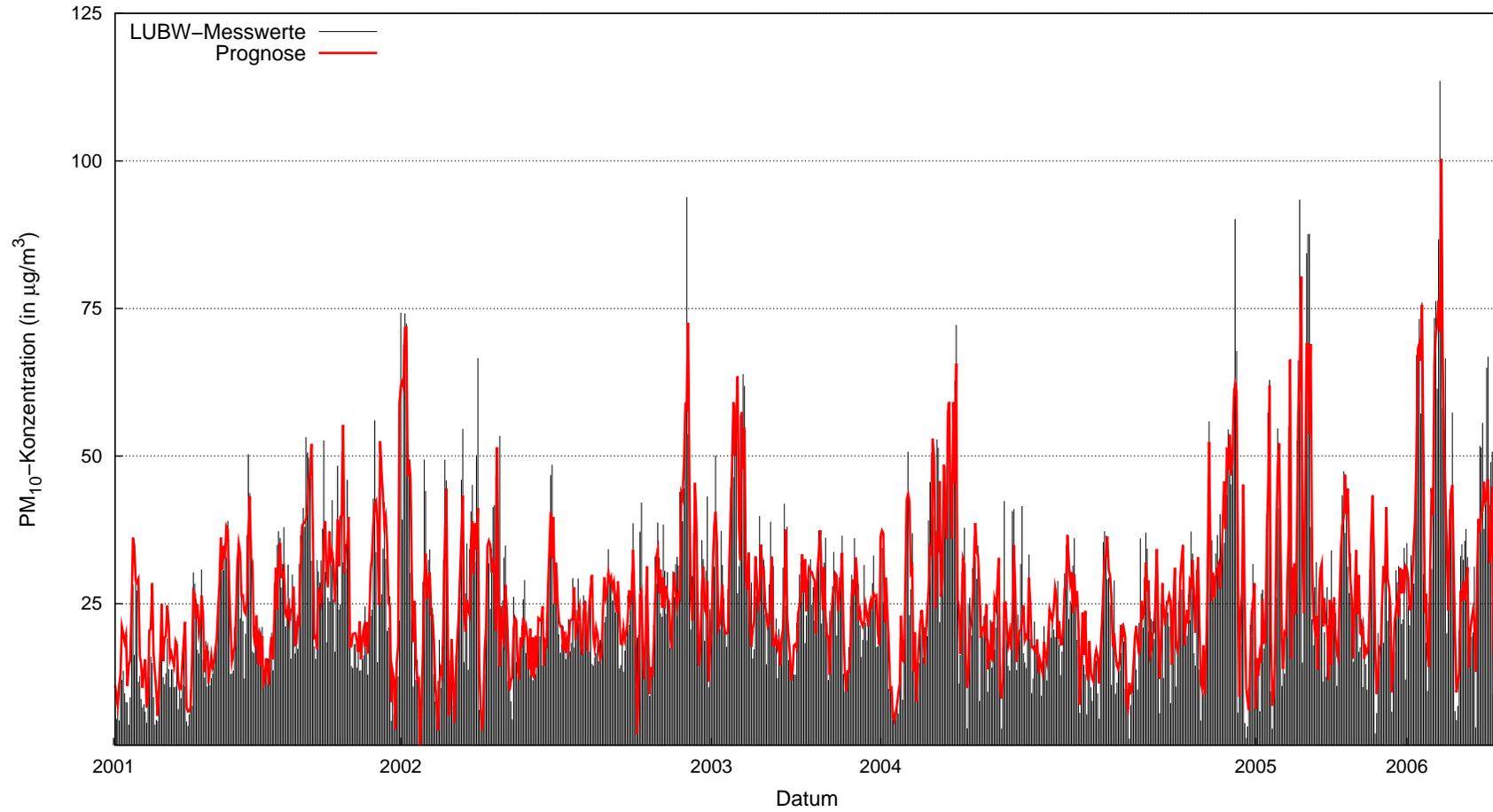


Abbildung C.16: Vergleich der PM₁₀-Prognose mittels Support-Vektor-Maschine mit den Messwerten der Station Stuttgart-Bad Cannstatt

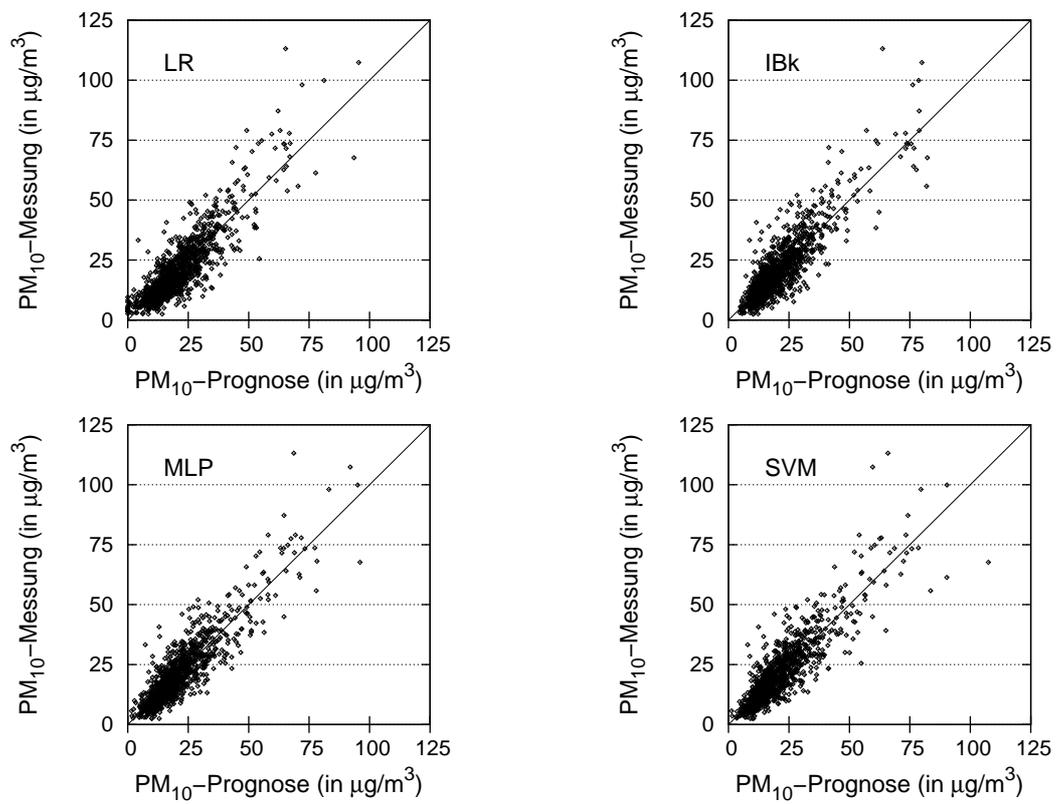


Abbildung C.17: Vergleich der Messdaten mit den Prognosedaten der verschiedenen Modelle für die Station Biberach

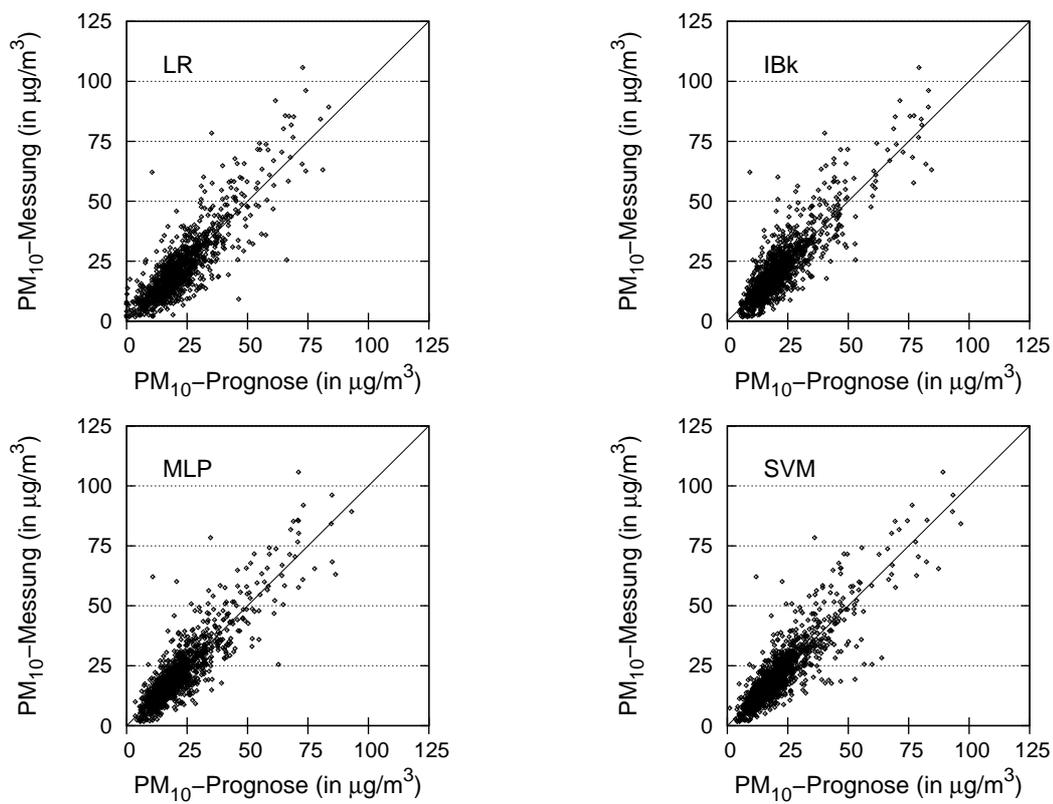


Abbildung C.18: Vergleich der Messdaten mit den Prognosedaten der Verschiedenen Modelle für die Station Freiburg-Mitte

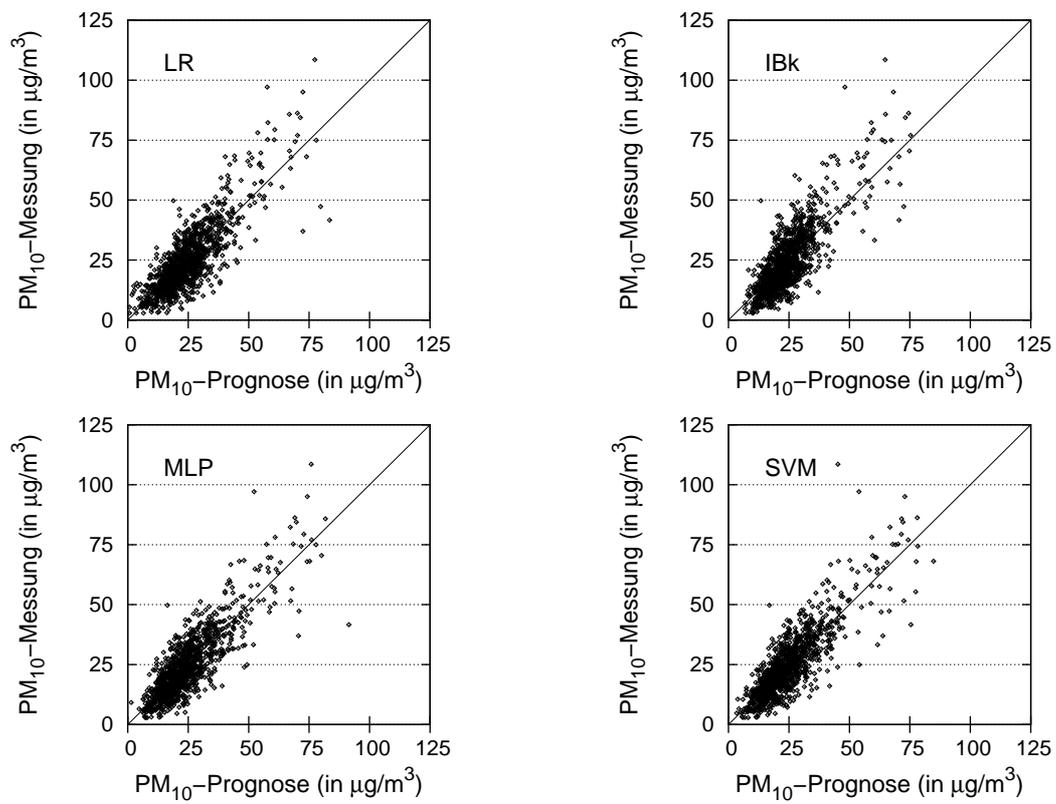


Abbildung C.19: Vergleich der Messdaten mit den Prognosedaten der Verschiedenen Modelle für die Station Karlsruhe-Nordwest

ANHANG C. DETAILLIERTE ERGEBNISSE DER PM₁₀-PROGNOSE

Tabelle C.1: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Biberach

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	9%(101)	7%(76)	1%(7)	—	—	—	17%(184)
	2	4%(40)	26%(289)	10%(113)	0%(1)	—	—	40%(443)
	3	0%(2)	7%(74)	22%(237)	2%(22)	0%(1)	—	31%(336)
	4	—	0%(2)	3%(37)	4%(43)	1%(6)	—	8%(88)
	5	—	—	0%(1)	1%(16)	2%(25)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		13%(143)	40%(441)	36%(395)	7%(82)	3%(34)	—	1095
korrekt		71%	66%	60%	52%	74%	—	63%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	6%(67)	10%(113)	0%(4)	—	—	—	17%(184)
	2	2%(24)	30%(329)	8%(89)	0%(1)	—	—	40%(443)
	3	0%(1)	8%(87)	21%(234)	1%(14)	—	—	31%(336)
	4	—	0%(3)	4%(47)	3%(35)	0%(3)	—	8%(88)
	5	—	—	0%(2)	1%(13)	2%(27)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		8%(92)	49%(532)	34%(376)	6%(63)	3%(32)	—	1095
korrekt		73%	62%	62%	56%	84%	—	63%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(90)	8%(93)	0%(1)	—	—	—	17%(184)
	2	2%(25)	31%(341)	7%(77)	—	—	—	40%(443)
	3	0%(2)	8%(93)	20%(216)	2%(25)	—	—	31%(336)
	4	—	0%(2)	4%(41)	3%(36)	1%(9)	—	8%(88)
	5	—	—	0%(2)	1%(7)	3%(33)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		11%(117)	48%(529)	31%(337)	6%(68)	4%(44)	—	1095
korrekt		77%	64%	64%	53%	75%	—	65%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(93)	8%(88)	0%(3)	—	—	—	17%(184)
	2	2%(19)	31%(342)	7%(81)	0%(1)	—	—	40%(443)
	3	0%(3)	8%(85)	20%(221)	2%(25)	0%(2)	—	31%(336)
	4	—	0%(3)	4%(37)	3%(38)	1%(10)	—	8%(88)
	5	—	—	0%(1)	1%(11)	3%(29)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		11%(115)	47%(518)	31%(343)	7%(75)	4%(43)	—	1095
korrekt		81%	66%	64%	51%	67%	—	66%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG C. DETAILLIERTE ERGEBNISSE DER PM₁₀-PROGNOSE

Tabelle C.2: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Freiburg-Mitte

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(84)	6%(72)	1%(8)	0%(1)	—	—	15%(165)
	2	3%(30)	27%(301)	10%(109)	0%(5)	—	—	40%(445)
	3	0%(2)	7%(84)	22%(250)	3%(30)	0%(2)	—	33%(369)
	4	—	0%(2)	4%(40)	3%(37)	1%(6)	—	8%(86)
	5	—	0%(1)	0%(4)	2%(23)	2%(28)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		10%(116)	41%(460)	37%(411)	9%(96)	3%(39)	—	1122
korrekt		72%	65%	61%	39%	72%	—	62%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	5%(56)	9%(101)	1%(8)	—	—	—	15%(165)
	2	2%(19)	29%(322)	9%(102)	0%(2)	—	—	40%(445)
	3	—	9%(97)	22%(248)	2%(22)	0%(2)	—	33%(369)
	4	—	0%(4)	3%(39)	4%(41)	0%(2)	—	8%(86)
	5	0%(1)	—	0%(4)	2%(25)	2%(26)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		7%(76)	47%(524)	36%(401)	8%(90)	3%(31)	—	1122
korrekt		74%	61%	62%	46%	84%	—	62%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(76)	7%(82)	1%(7)	—	—	—	15%(165)
	2	3%(34)	28%(316)	8%(92)	0%(3)	—	—	40%(445)
	3	0%(1)	7%(82)	23%(260)	2%(24)	0%(2)	—	33%(369)
	4	—	0%(4)	3%(38)	3%(37)	1%(7)	—	8%(86)
	5	—	0%(1)	0%(5)	1%(14)	3%(36)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		10%(111)	43%(485)	36%(402)	7%(78)	4%(46)	—	1122
korrekt		68%	65%	65%	47%	78%	—	65%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(82)	7%(81)	0%(2)	—	—	—	15%(165)
	2	2%(28)	29%(321)	8%(90)	0%(5)	0%(1)	—	40%(445)
	3	0%(2)	8%(86)	22%(251)	2%(26)	0%(4)	—	33%(369)
	4	—	0%(3)	3%(38)	4%(40)	0%(5)	—	8%(86)
	5	—	0%(1)	0%(4)	2%(21)	3%(30)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		10%(112)	44%(492)	34%(385)	8%(92)	4%(41)	—	1122
korrekt		73%	65%	65%	43%	73%	—	65%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

Tabelle C.3: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Karlsruhe-Nordwest

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(39)	5%(52)	1%(11)	—	—	—	9%(102)
	2	2%(25)	19%(219)	12%(134)	0%(4)	—	—	34%(382)
	3	0%(1)	7%(84)	28%(321)	4%(40)	0%(1)	—	40%(447)
	4	—	0%(2)	7%(80)	5%(58)	1%(7)	—	13%(147)
	5	—	—	—	2%(17)	3%(35)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		6%(65)	32%(357)	48%(546)	11%(119)	4%(44)	—	1131
korrekt		60%	61%	59%	49%	80%	—	59%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	1%(11)	7%(76)	1%(15)	—	—	—	9%(102)
	2	1%(6)	20%(224)	13%(151)	0%(1)	—	—	34%(382)
	3	0%(1)	7%(79)	31%(347)	2%(19)	0%(1)	—	40%(447)
	4	—	0%(3)	9%(106)	3%(32)	1%(6)	—	13%(147)
	5	—	—	0%(5)	2%(17)	3%(30)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		2%(18)	34%(382)	55%(624)	6%(69)	3%(38)	—	1131
korrekt		61%	59%	56%	46%	79%	—	57%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(38)	5%(55)	1%(9)	—	—	—	9%(102)
	2	1%(11)	21%(240)	11%(129)	0%(2)	—	—	34%(382)
	3	0%(1)	8%(94)	28%(315)	3%(36)	0%(1)	—	40%(447)
	4	—	0%(3)	7%(77)	5%(60)	1%(7)	—	13%(147)
	5	—	—	0%(1)	1%(14)	3%(37)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		4%(50)	35%(392)	47%(531)	10%(112)	4%(46)	—	1131
korrekt		76%	61%	59%	54%	80%	—	61%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(30)	5%(62)	1%(10)	—	—	—	9%(102)
	2	1%(13)	21%(239)	11%(127)	0%(3)	—	—	34%(382)
	3	0%(1)	8%(86)	29%(324)	3%(34)	0%(2)	—	40%(447)
	4	—	0%(2)	7%(79)	5%(60)	1%(6)	—	13%(147)
	5	—	—	0%(1)	1%(15)	3%(36)	—	5%(52)
	6	—	—	—	0%(1)	—	—	0%(1)
total ¹		4%(44)	34%(389)	48%(541)	10%(113)	4%(44)	—	1131
korrekt		68%	61%	60%	53%	82%	—	61%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG C. DETAILLIERTE ERGEBNISSE DER PM₁₀-PROGNOSE

Tabelle C.4: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Stuttgart-Bad Cannstatt

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(40)	3%(33)	1%(8)	0%(1)	—	—	8%(82)
	2	2%(20)	16%(157)	13%(132)	0%(3)	—	—	31%(312)
	3	—	5%(54)	30%(304)	5%(49)	0%(3)	—	41%(410)
	4	—	0%(1)	7%(67)	5%(53)	1%(7)	—	13%(128)
	5	—	—	0%(4)	4%(38)	3%(33)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		6%(60)	24%(245)	51%(515)	14%(144)	4%(44)	—	1008
korrekt		67%	64%	59%	37%	75%	—	58%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(23)	5%(47)	1%(11)	0%(1)	—	—	8%(82)
	2	1%(11)	18%(183)	11%(115)	0%(3)	—	—	31%(312)
	3	—	7%(73)	30%(306)	3%(26)	0%(5)	—	41%(410)
	4	—	0%(1)	7%(68)	5%(53)	1%(6)	—	13%(128)
	5	—	0%(1)	0%(4)	3%(35)	3%(35)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		3%(34)	30%(305)	50%(504)	12%(118)	5%(47)	—	1008
korrekt		68%	60%	61%	45%	74%	—	60%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(29)	5%(48)	0%(5)	—	—	—	8%(82)
	2	1%(15)	19%(195)	10%(99)	0%(3)	—	—	31%(312)
	3	—	6%(62)	30%(301)	5%(46)	0%(1)	—	41%(410)
	4	—	0%(1)	6%(60)	6%(57)	1%(10)	—	13%(128)
	5	—	0%(1)	0%(5)	2%(22)	5%(47)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		4%(44)	30%(307)	47%(470)	13%(128)	6%(59)	—	1008
korrekt		66%	64%	64%	45%	80%	—	62%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(30)	5%(48)	1%(4)	—	—	—	8%(82)
	2	1%(14)	18%(193)	11%(103)	0%(2)	—	—	31%(312)
	3	—	7%(66)	30%(303)	3%(36)	0%(5)	—	41%(410)
	4	—	0%(1)	7%(68)	5%(50)	1%(9)	—	13%(128)
	5	—	0%(1)	0%(4)	3%(31)	3%(38)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		4%(44)	31%(309)	48%(482)	12%(119)	5%(53)	—	1008
korrekt		68%	62%	63%	42%	72%	—	61%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

Anhang D

Detaillierte Ergebnisse der Sensitivitätsanalyse bezüglich unsicherer meteorologischer Vorhersagen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.1: Leistungskennzahlen der verschiedenen Modelle bei Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$, $\pm 25\%$ bzw. $\pm 50\%$ für die Station Biberach

keine Abweichung		LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,87	0,88	0,89	0,88
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	5,0	4,8	4,6	4,7
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	6,9	6,8	6,3	6,8
Index of Agreement (IA)		0,93	0,93	0,94	0,93
False Alarm Rate (FAR)	%	21	9	20	27
Probability of Detection (POD)	%	61	66	80	73
LuQx richtig prognostiziert	%	63	63	65	66

$\pm 10\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,87 (—)	0,87 (—)	0,89 (—)	0,88 (—)
MAE	$\mu\text{g}/\text{m}^3$	5,0 (—)	4,9 (+0,1)	4,6 (—)	4,7 (—)
RMSE	$\mu\text{g}/\text{m}^3$	6,9 (—)	6,9 (+0,1)	6,3 (—)	6,8 (—)
IA		0,93 (—)	0,93 (—)	0,94 (—)	0,93 (—)
FAR	%	21 (—)	10 (—)	20 (—)	27 (—)
POD	%	61 (—)	64 (—2)	80 (—)	73 (—)
LuQx	%	63 (—1)	64 (—)	66 (+1)	65 (—1)

$\pm 25\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,86 (—0,01)	0,87 (—0,01)	0,89 (—0,01)	0,87 (—)
MAE	$\mu\text{g}/\text{m}^3$	5,3 (+0,3)	5,0 (+0,2)	4,7 (+0,1)	4,8 (+0,1)
RMSE	$\mu\text{g}/\text{m}^3$	7,1 (+0,2)	6,9 (+0,1)	6,4 (+0,1)	6,8 (—)
IA		0,92 (—0,01)	0,92 (—)	0,94 (—)	0,93 (—)
FAR	%	23 (+2)	10 (+1)	19 (—1)	28 (+1)
POD	%	61 (—)	59 (—7)	77 (—2)	75 (+2)
LuQx	%	62 (—2)	63 (—)	64 (—1)	65 (—1)

$\pm 50\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,83 (—0,05)	0,85 (—0,02)	0,87 (—0,03)	0,85 (—0,03)
MAE	$\mu\text{g}/\text{m}^3$	6,2 (+1,2)	5,4 (+0,6)	5,3 (+0,7)	5,4 (+0,7)
RMSE	$\mu\text{g}/\text{m}^3$	8,0 (+1,1)	7,3 (+0,5)	7,0 (+0,7)	7,5 (+0,7)
IA		0,90 (—0,02)	0,91 (—0,01)	0,93 (—0,01)	0,92 (—0,02)
FAR	%	22 (+1)	10 (+1)	17 (—3)	28 (+1)
POD	%	66 (+5)	59 (—7)	89 (+9)	70 (—2)
LuQx	%	57 (—6)	61 (—2)	62 (—4)	59 (—7)

* Δ bezeichnet die Abweichung zur Leistung mit perfekter Prognose.

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.2: Leistungskennzahlen der verschiedenen Modelle bei Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$, $\pm 25\%$ bzw. $\pm 50\%$ für die Station Freiburg-Mitte

keine Abweichung		LR IBk MLP SVM			
Korrelationskoeffizient (r)		0,86	0,87	0,88	0,87
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	5,1	5,1	4,7	4,8
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	7,3	7,1	6,7	7,1
Index of Agreement (IA)		0,92	0,92	0,93	0,93
False Alarm Rate (FAR)	%	26	13	20	24
Probability of Detection (POD)	%	51	47	65	54
LuQx richtig prognostiziert	%	62	62	65	65

$\pm 10\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,86 (-0,01)	0,86 (—)	0,88 (—)	0,87 (—)
MAE	$\mu\text{g}/\text{m}^3$	5,2 (+0,1)	5,2 (+0,1)	4,7 (—)	4,8 (—)
RMSE	$\mu\text{g}/\text{m}^3$	7,3 (—)	7,3 (+0,2)	6,7 (—)	7,1 (—)
IA		0,92 (—)	0,92 (-0,01)	0,93 (—)	0,93 (—)
FAR	%	23 (-3)	18 (+6)	19 (-1)	23 (-2)
POD	%	53 (+2)	54 (+7)	68 (+4)	54 (—)
LuQx	%	62 (—)	62 (—)	65 (—)	65 (—)

$\pm 25\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,84 (-0,02)	0,85 (-0,02)	0,87 (-0,01)	0,86 (-0,01)
MAE	$\mu\text{g}/\text{m}^3$	5,5 (+0,4)	5,4 (+0,3)	5,0 (+0,3)	4,9 (+0,1)
RMSE	$\mu\text{g}/\text{m}^3$	7,7 (+0,4)	7,5 (+0,4)	7,0 (+0,3)	7,2 (+0,1)
IA		0,91 (-0,01)	0,91 (-0,01)	0,93 (-0,01)	0,93 (—)
FAR	%	25 (-1)	15 (+2)	23 (+4)	25 (+1)
POD	%	53 (+2)	51 (+4)	63 (-2)	53 (-2)
LuQx	%	59 (-3)	60 (-2)	62 (-3)	63 (-1)

$\pm 50\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,81 (-0,05)	0,83 (-0,04)	0,83 (-0,05)	0,81 (-0,06)
MAE	$\mu\text{g}/\text{m}^3$	6,5 (+1,4)	5,9 (+0,8)	5,8 (+1,1)	6,5 (+1,7)
RMSE	$\mu\text{g}/\text{m}^3$	8,5 (+1,2)	8,0 (+0,9)	8,2 (+1,5)	8,5 (+1,4)
IA		0,89 (-0,03)	0,89 (-0,03)	0,91 (-0,02)	0,89 (-0,03)
FAR	%	27 (+1)	8 (-5)	35 (+15)	27 (+2)
POD	%	53 (+2)	40 (-7)	72 (+7)	53 (-2)
LuQx	%	55 (-8)	58 (-4)	59 (-6)	55 (-10)

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.3: Leistungskennzahlen der verschiedenen Modelle bei Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$, $\pm 25\%$ bzw. $\pm 50\%$ für die Station Karlsruhe-Nordwest

keine Abweichung		LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,83	0,81	0,84	0,84
Mean Absolute Error (MAE)	$\mu\text{g}/\text{m}^3$	5,8	6,1	5,5	5,4
Root Mean Squared Error (RMSE)	$\mu\text{g}/\text{m}^3$	7,8	8,1	7,4	7,4
Index of Agreement (IA)		0,90	0,88	0,91	0,91
False Alarm Rate (FAR)	%	18	16	16	18
Probability of Detection (POD)	%	69	60	73	69
LuQx richtig prognostiziert	%	59	57	61	61

$\pm 10\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,82 (-0,01)	0,82 (+0,01)	0,84 (-)	0,84 (-)
MAE	$\mu\text{g}/\text{m}^3$	5,9 (+0,1)	6,1 (-)	5,5 (-)	5,4 (-)
RMSE	$\mu\text{g}/\text{m}^3$	7,9 (+0,1)	8,0 (-0,1)	7,4 (-)	7,5 (+0,1)
IA		0,90 (-)	0,88 (-)	0,91 (-)	0,91 (-)
FAR	%	19 (+1)	17 (-)	17 (+2)	18 (-)
POD	%	65 (-4)	58 (-2)	73 (-)	69 (-)
LuQx	%	58 (-2)	57 (-)	59 (-2)	61 (-)

$\pm 25\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,79 (-0,04)	0,80 (-0,01)	0,84 (-0,01)	0,84 (-0,01)
MAE	$\mu\text{g}/\text{m}^3$	6,6 (+0,8)	6,4 (+0,3)	5,6 (+0,1)	5,6 (+0,2)
RMSE	$\mu\text{g}/\text{m}^3$	8,5 (+0,7)	8,3 (+0,2)	7,5 (+0,1)	7,5 (+0,1)
IA		0,88 (-0,02)	0,87 (-0,01)	0,91 (-)	0,91 (-)
FAR	%	20 (+2)	18 (+2)	17 (+2)	19 (+1)
POD	%	67 (-2)	60 (-)	73 (-)	65 (-4)
LuQx	%	54 (-6)	55 (-2)	61 (-)	59 (-2)

$\pm 50\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,70 (-0,12)	0,76 (-0,05)	0,82 (-0,03)	0,82 (-0,02)
MAE	$\mu\text{g}/\text{m}^3$	8,4 (+2,6)	6,8 (+0,7)	6,0 (+0,5)	5,9 (+0,5)
RMSE	$\mu\text{g}/\text{m}^3$	10,5 (+2,7)	9,0 (+0,9)	8,0 (+0,6)	7,9 (+0,5)
IA		0,84 (-0,06)	0,85 (-0,03)	0,90 (-0,01)	0,90 (-0,01)
FAR	%	34 (+16)	26 (+9)	20 (+4)	20 (+1)
POD	%	63 (-6)	56 (-4)	71 (-2)	71 (+2)
LuQx	%	44 (-15)	53 (-3)	58 (-3)	59 (-2)

* Δ bezeichnet die Abweichung zur Leistung mit perfekter Prognose.

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.4: Leistungskennzahlen der verschiedenen Modelle bei Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$, $\pm 25\%$ bzw. $\pm 50\%$ für die Station Stuttgart-Bad Cannstatt

keine Abweichung		LR	IBk	MLP	SVM
Korrelationskoeffizient (r)		0,84	0,83	0,86	0,84
Mean Absolute Error (MAE) $\mu\text{g}/\text{m}^3$		6,2	6,2	5,6	5,8
Root Mean Squared Error (RMSE) $\mu\text{g}/\text{m}^3$		8,2	8,4	7,6	8,0
Index of Agreement (IA)		0,90	0,89	0,92	0,91
False Alarm Rate (FAR) %		23	23	19	26
Probability of Detection (POD) %		45	47	63	53
LuQx richtig prognostiziert %		58	60	62	61

$\pm 10\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,83 (-0,01)	0,82 (—)	0,85 (-0,01)	0,84 (—)
MAE	$\mu\text{g}/\text{m}^3$	6,4 (+0,2)	6,3 (+0,1)	5,7 (+0,1)	5,8 (—)
RMSE	$\mu\text{g}/\text{m}^3$	8,4 (+0,2)	8,5 (+0,1)	7,7 (+0,1)	8,0 (—)
IA		0,90 (—)	0,89 (—)	0,92 (—)	0,91 (—)
FAR	%	24 (+2)	26 (+3)	18 (-1)	25 (—)
POD	%	45 (—)	45 (-3)	62 (-1)	54 (+1)
LuQx	%	58 (—)	58 (-2)	61 (-1)	60 (-1)

$\pm 25\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,80 (-0,04)	0,82 (-0,01)	0,84 (-0,02)	0,84 (-0,01)
MAE	$\mu\text{g}/\text{m}^3$	7,0 (+0,8)	6,5 (+0,3)	6,2 (+0,6)	6,0 (+0,2)
RMSE	$\mu\text{g}/\text{m}^3$	9,1 (+0,9)	8,7 (+0,3)	8,2 (+0,6)	8,2 (+0,2)
IA		0,89 (-0,02)	0,88 (-0,01)	0,91 (-0,01)	0,91 (—)
FAR	%	25 (+2)	26 (+2)	24 (+6)	24 (-2)
POD	%	47 (+3)	46 (-1)	62 (-1)	51 (-1)
LuQx	%	54 (-4)	58 (-2)	59 (-4)	60 (-1)

$\pm 50\%$		LR (Δ^*)	IBk (Δ^*)	MLP (Δ^*)	SVM (Δ^*)
r		0,73 (-0,11)	0,80 (-0,02)	0,79 (-0,07)	0,81 (-0,04)
MAE	$\mu\text{g}/\text{m}^3$	8,5 (+2,3)	6,7 (+0,5)	6,9 (+1,3)	6,5 (+0,7)
RMSE	$\mu\text{g}/\text{m}^3$	10,8 (+2,6)	8,9 (+0,5)	9,3 (+1,7)	8,8 (+0,8)
IA		0,85 (-0,05)	0,87 (-0,02)	0,89 (-0,03)	0,89 (-0,02)
FAR	%	27 (+5)	26 (+2)	28 (+9)	28 (+2)
POD	%	49 (+4)	42 (-5)	58 (-6)	54 (+1)
LuQx	%	48 (-10)	56 (-3)	55 (-7)	57 (-4)

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.5: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Biberach bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	9%(98)	7%(79)	1%(7)	—	—	—	17%(184)
	2	4%(44)	26%(282)	11%(115)	0%(2)	—	—	40%(443)
	3	0%(2)	6%(71)	22%(238)	2%(24)	0%(1)	—	31%(336)
	4	—	0%(2)	3%(37)	4%(43)	1%(6)	—	8%(88)
	5	—	—	0%(1)	1%(16)	2%(25)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		13%(144)	40%(434)	36%(398)	8%(85)	3%(34)	—	1095
korrekt		68%	65%	60%	51%	74%	—	63%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	6%(69)	10%(110)	0%(5)	—	—	—	17%(184)
	2	2%(22)	30%(331)	8%(89)	0%(1)	—	—	40%(443)
	3	0%(3)	8%(84)	21%(234)	1%(15)	—	—	31%(336)
	4	—	0%(3)	4%(45)	3%(37)	0%(3)	—	8%(88)
	5	—	—	0%(1)	1%(15)	2%(26)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		9%(94)	48%(528)	34%(374)	6%(68)	3%(31)	—	1095
korrekt		73%	63%	63%	54%	84%	—	64%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	9%(95)	8%(88)	0%(1)	—	—	—	17%(184)
	2	2%(26)	31%(338)	7%(79)	—	—	—	40%(443)
	3	0%(3)	8%(90)	20%(219)	2%(24)	—	—	31%(336)
	4	—	0%(2)	4%(40)	3%(37)	1%(9)	—	8%(88)
	5	—	—	0%(2)	1%(7)	3%(33)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		11%(124)	47%(518)	31%(341)	6%(68)	4%(44)	—	1095
korrekt		77%	65%	64%	54%	75%	—	66%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(85)	8%(96)	0%(3)	—	—	—	17%(184)
	2	2%(18)	31%(342)	7%(82)	0%(1)	—	—	40%(443)
	3	0%(3)	8%(83)	20%(221)	2%(28)	0%(1)	—	31%(336)
	4	—	0%(3)	4%(37)	3%(37)	1%(11)	—	8%(88)
	5	—	—	0%(1)	1%(11)	3%(29)	0%(1)	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		10%(106)	48%(524)	31%(344)	7%(77)	4%(43)	0%(1)	1095
korrekt		80%	65%	64%	48%	67%	0%	65%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.6: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Biberach bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 25\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(91)	8%(86)	1%(7)	—	—	—	17%(184)
	2	5%(55)	25%(278)	10%(109)	0%(1)	—	—	40%(443)
	3	0%(4)	6%(68)	22%(240)	2%(23)	0%(1)	—	31%(336)
	4	—	0%(3)	3%(38)	4%(40)	1%(7)	—	8%(88)
	5	—	—	0%(1)	1%(16)	2%(25)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		14%(150)	40%(435)	36%(395)	7%(80)	3%(35)	—	1095
korrekt		61%	64%	61%	50%	71%	—	62%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	6%(65)	10%(113)	1%(6)	—	—	—	17%(184)
	2	2%(22)	30%(332)	8%(88)	0%(1)	—	—	40%(443)
	3	0%(3)	8%(85)	21%(234)	1%(14)	—	—	31%(336)
	4	—	0%(4)	4%(49)	3%(32)	0%(3)	—	8%(88)
	5	—	—	0%(3)	1%(15)	2%(24)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		8%(90)	49%(534)	35%(380)	6%(62)	3%(29)	—	1095
korrekt		72%	62%	62%	52%	83%	—	63%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(93)	8%(89)	0%(2)	—	—	—	17%(184)
	2	3%(34)	29%(320)	8%(89)	—	—	—	40%(443)
	3	0%(2)	8%(88)	20%(220)	2%(26)	—	—	31%(336)
	4	—	0%(2)	4%(42)	3%(36)	1%(8)	—	8%(88)
	5	—	—	0%(2)	1%(8)	3%(32)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		12%(129)	46%(499)	32%(355)	6%(70)	4%(42)	—	1095
korrekt		72%	64%	62%	51%	76%	—	64%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	9%(98)	8%(83)	0%(3)	—	—	—	17%(184)
	2	3%(30)	30%(324)	8%(88)	0%(1)	—	—	40%(443)
	3	0%(2)	7%(82)	20%(224)	2%(27)	0%(1)	—	31%(336)
	4	—	0%(3)	3%(38)	3%(35)	1%(12)	—	8%(88)
	5	—	—	—	1%(11)	3%(30)	0%(1)	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total ¹		12%(130)	45%(492)	32%(353)	7%(74)	4%(45)	0%(1)	1095
korrekt		75%	66%	63%	47%	67%	0%	65%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.7: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Biberach bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 50\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(87)	9%(87)	1%(10)	—	—	—	17%(184)
	2	4%(71)	26%(237)	11%(130)	0%(5)	—	—	40%(443)
	3	0%(7)	8%(71)	21%(232)	2%(25)	0%(1)	—	31%(336)
	4	—	0%(3)	3%(37)	4%(41)	1%(7)	—	8%(88)
	5	—	—	0%(1)	1%(14)	2%(27)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		15%(165)	36%(398)	37%(410)	8%(85)	3%(37)	—	1095
korrekt		53%	60%	57%	48%	73%	—	57%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	5%(54)	11%(125)	0%(5)	—	—	—	17%(184)
	2	2%(17)	30%(331)	9%(94)	0%(1)	—	—	40%(443)
	3	0%(1)	9%(94)	20%(223)	2%(17)	—	—	31%(336)
	4	—	1%(6)	4%(47)	3%(33)	0%(2)	—	8%(88)
	5	—	—	0%(3)	1%(15)	2%(24)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		7%(72)	51%(556)	34%(372)	6%(66)	3%(29)	—	1095
korrekt		75%	60%	60%	50%	83%	—	61%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(85)	9%(97)	0%(2)	—	—	—	17%(184)
	2	4%(49)	27%(297)	9%(96)	0%(1)	—	—	40%(443)
	3	0%(4)	7%(82)	20%(222)	3%(28)	—	—	31%(336)
	4	—	0%(4)	4%(40)	3%(36)	1%(8)	—	8%(88)
	5	—	—	0%(1)	0%(4)	3%(37)	—	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		13%(138)	44%(480)	33%(361)	6%(69)	4%(47)	—	1095
korrekt		62%	62%	61%	52%	79%	—	62%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(78)	9%(96)	1%(10)	—	—	—	17%(184)
	2	4%(42)	26%(285)	11%(115)	0%(1)	—	—	40%(443)
	3	0%(4)	8%(85)	20%(221)	2%(23)	0%(3)	—	31%(336)
	4	—	0%(5)	4%(40)	3%(34)	1%(9)	—	8%(88)
	5	—	—	0%(3)	1%(10)	3%(28)	0%(1)	4%(42)
	6	—	—	—	—	0%(2)	—	0%(2)
total¹		11%(124)	43%(471)	36%(389)	6%(68)	4%(42)	0%(1)	1095
korrekt		63%	61%	57%	50%	67%	0%	59%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.8: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Freiburg-Mitte bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(83)	7%(75)	1%(6)	0%(1)	—	—	15%(165)
	2	3%(34)	26%(292)	10%(114)	0%(5)	—	—	40%(445)
	3	0%(2)	7%(84)	23%(254)	2%(27)	0%(2)	—	33%(369)
	4	—	0%(2)	3%(39)	3%(38)	1%(7)	—	8%(86)
	5	0%(1)	—	0%(5)	2%(21)	3%(29)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		11%(120)	40%(453)	37%(418)	8%(92)	3%(39)	—	1122
korrekt		69%	64%	61%	41%	74%	—	62%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	6%(62)	9%(96)	1%(7)	—	—	—	15%(165)
	2	2%(18)	29%(322)	9%(103)	0%(2)	—	—	40%(445)
	3	—	9%(97)	22%(244)	2%(25)	0%(3)	—	33%(369)
	4	—	0%(5)	3%(39)	3%(38)	0%(4)	—	8%(86)
	5	0%(1)	0%(1)	0%(3)	2%(21)	3%(30)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		7%(81)	46%(521)	35%(396)	8%(86)	3%(38)	—	1122
korrekt		77%	62%	62%	44%	79%	—	62%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(76)	7%(83)	1%(6)	—	—	—	15%(165)
	2	3%(36)	28%(314)	8%(92)	0%(3)	—	—	40%(445)
	3	0%(2)	7%(76)	24%(264)	2%(24)	0%(3)	—	33%(369)
	4	—	0%(3)	4%(40)	3%(37)	1%(6)	—	8%(86)
	5	—	0%(1)	0%(3)	1%(14)	3%(38)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		10%(114)	43%(477)	36%(405)	7%(78)	4%(48)	—	1122
korrekt		67%	66%	65%	47%	79%	—	65%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(84)	7%(78)	0%(3)	—	—	—	15%(165)
	2	3%(30)	29%(320)	8%(89)	0%(5)	0%(1)	—	40%(445)
	3	0%(2)	8%(87)	22%(250)	2%(26)	0%(4)	—	33%(369)
	4	—	0%(3)	3%(38)	4%(41)	0%(4)	—	8%(86)
	5	—	0%(1)	0%(4)	2%(21)	3%(30)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		10%(116)	44%(489)	34%(384)	8%(93)	4%(40)	—	1122
korrekt		72%	65%	65%	44%	75%	—	65%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.9: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Freiburg-Mitte bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 25\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(78)	7%(80)	1%(6)	0%(1)	—	—	15%(165)
	2	3%(39)	25%(280)	11%(121)	0%(5)	—	—	40%(445)
	3	0%(5)	8%(91)	21%(241)	3%(29)	0%(3)	—	33%(369)
	4	—	0%(4)	3%(39)	3%(36)	1%(7)	—	8%(86)
	5	0%(1)	—	0%(5)	2%(21)	3%(29)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		11%(123)	41%(455)	37%(412)	8%(92)	4%(40)	—	1122
korrekt		63%	62%	58%	39%	73%	—	59%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(49)	10%(107)	1%(9)	—	—	—	15%(165)
	2	2%(20)	28%(319)	9%(104)	0%(2)	—	—	40%(445)
	3	0%(1)	9%(102)	22%(243)	2%(21)	0%(2)	—	33%(369)
	4	—	0%(4)	4%(42)	3%(37)	0%(3)	—	8%(86)
	5	0%(1)	0%(1)	0%(3)	2%(23)	2%(28)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		6%(71)	48%(533)	36%(401)	7%(83)	3%(34)	—	1122
korrekt		69%	60%	61%	45%	82%	—	60%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(76)	7%(83)	1%(6)	—	—	—	15%(165)
	2	3%(32)	27%(307)	9%(103)	0%(3)	—	—	40%(445)
	3	0%(1)	8%(95)	22%(243)	2%(28)	0%(3)	—	33%(369)
	4	—	0%(2)	4%(41)	3%(34)	1%(9)	—	8%(86)
	5	—	0%(1)	0%(4)	1%(16)	3%(35)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		10%(109)	43%(488)	35%(397)	7%(81)	4%(47)	—	1122
korrekt		70%	63%	61%	42%	74%	—	62%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(79)	7%(83)	0%(3)	—	—	—	15%(165)
	2	3%(31)	28%(316)	8%(91)	1%(7)	—	—	40%(445)
	3	0%(4)	8%(86)	22%(249)	2%(26)	0%(4)	—	33%(369)
	4	—	0%(3)	4%(40)	3%(37)	1%(6)	—	8%(86)
	5	—	0%(1)	0%(5)	2%(21)	3%(29)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		10%(114)	44%(489)	35%(388)	8%(91)	4%(40)	—	1122
korrekt		69%	65%	64%	41%	73%	—	63%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.10: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Freiburg-Mitte bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 50\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(86)	6%(62)	1%(16)	0%(1)	—	—	15%(165)
	2	7%(73)	20%(219)	13%(149)	0%(4)	—	—	40%(445)
	3	1%(9)	7%(81)	21%(241)	3%(33)	0%(5)	—	33%(369)
	4	—	0%(5)	3%(35)	4%(40)	1%(6)	—	8%(86)
	5	0%(1)	0%(1)	0%(2)	2%(24)	3%(29)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		15%(168)	33%(368)	39%(443)	9%(102)	4%(41)	—	1122
korrekt		51%	60%	54%	39%	71%	—	55%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	5%(52)	9%(98)	1%(15)	—	—	—	15%(165)
	2	2%(26)	26%(291)	11%(127)	0%(1)	—	—	40%(445)
	3	—	9%(100)	22%(247)	2%(20)	0%(2)	—	33%(369)
	4	—	0%(5)	4%(44)	3%(37)	—	—	8%(86)
	5	—	0%(2)	1%(7)	2%(25)	2%(22)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		7%(78)	44%(496)	39%(440)	7%(83)	2%(25)	—	1122
korrekt		67%	59%	56%	45%	88%	—	58%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	7%(83)	6%(72)	1%(9)	0%(1)	—	—	15%(165)
	2	5%(59)	25%(277)	9%(104)	0%(5)	—	—	40%(445)
	3	0%(1)	9%(98)	20%(230)	3%(31)	1%(9)	—	33%(369)
	4	—	1%(6)	3%(36)	3%(31)	1%(13)	—	8%(86)
	5	—	0%(2)	0%(2)	1%(12)	4%(40)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		13%(143)	41%(455)	34%(381)	7%(80)	6%(63)	—	1122
korrekt		58%	61%	60%	39%	63%	—	59%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	8%(86)	6%(62)	1%(16)	0%(1)	—	—	15%(165)
	2	7%(73)	20%(219)	13%(149)	0%(4)	—	—	40%(445)
	3	1%(9)	7%(81)	21%(241)	3%(33)	0%(5)	—	33%(369)
	4	—	0%(5)	3%(35)	4%(40)	1%(6)	—	8%(86)
	5	—	0%(1)	0%(2)	2%(24)	3%(29)	—	5%(56)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		15%(168)	33%(368)	39%(443)	9%(102)	4%(41)	—	1122
korrekt		51%	60%	54%	39%	71%	—	55%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.11: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Karlsruhe-Nordwest bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(42)	4%(49)	1%(11)	—	—	—	9%(102)
	2	2%(28)	18%(207)	13%(144)	0%(3)	—	—	34%(382)
	3	0%(3)	7%(84)	27%(311)	4%(48)	0%(1)	—	40%(447)
	4	—	0%(1)	7%(77)	5%(62)	1%(7)	—	13%(147)
	5	—	—	—	2%(19)	3%(33)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		6%(73)	30%(341)	48%(542)	12%(132)	4%(42)	—	1131
korrekt		58%	61%	57%	47%	79%	—	58%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	1%(9)	7%(77)	1%(16)	—	—	—	9%(102)
	2	0%(5)	19%(218)	14%(158)	0%(1)	—	—	34%(382)
	3	0%(1)	6%(72)	31%(355)	2%(18)	0%(1)	—	40%(447)
	4	—	0%(4)	10%(109)	3%(29)	0%(5)	—	13%(147)
	5	—	—	0%(5)	2%(18)	3%(29)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		1%(15)	33%(371)	57%(643)	6%(66)	3%(36)	—	1131
korrekt		60%	59%	55%	44%	81%	—	57%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(32)	5%(62)	1%(8)	—	—	—	9%(102)
	2	1%(12)	21%(236)	12%(131)	0%(3)	—	—	34%(382)
	3	0%(1)	9%(101)	27%(307)	3%(36)	0%(2)	—	40%(447)
	4	—	0%(2)	7%(78)	5%(60)	1%(7)	—	13%(147)
	5	—	—	0%(1)	1%(14)	3%(37)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		4%(45)	35%(401)	46%(525)	10%(113)	4%(47)	—	1131
korrekt		71%	59%	58%	53%	79%	—	59%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(29)	6%(64)	1%(9)	—	—	—	9%(102)
	2	1%(12)	21%(241)	11%(127)	0%(2)	—	—	34%(382)
	3	0%(1)	8%(86)	29%(323)	3%(35)	0%(2)	—	40%(447)
	4	—	0%(2)	7%(81)	5%(58)	1%(6)	—	13%(147)
	5	—	—	0%(1)	1%(15)	3%(36)	—	5%(52)
	6	—	—	—	0%(1)	—	—	0%(1)
total¹		4%(42)	35%(393)	48%(541)	10%(111)	4%(44)	—	1131
korrekt		69%	61%	60%	52%	82%	—	61%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.12: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Karlsruhe-Nordwest bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 25\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(39)	4%(49)	1%(14)	—	—	—	9%(102)
	2	4%(49)	15%(166)	14%(162)	0%(5)	—	—	34%(382)
	3	0%(3)	9%(104)	27%(302)	3%(6)	0%(2)	—	40%(447)
	4	—	0%(5)	6%(67)	6%(68)	1%(7)	—	13%(147)
	5	—	—	0%(1)	2%(17)	3%(34)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		8%(91)	29%(324)	48%(546)	11%(126)	4%(44)	—	1131
korrekt		43%	51%	55%	54%	77%	—	54%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	1%(11)	7%(76)	1%(15)	—	—	—	9%(102)
	2	0%(3)	18%(205)	15%(173)	0%(1)	—	—	34%(382)
	3	—	7%(82)	31%(345)	2%(18)	0%(2)	—	40%(447)
	4	—	0%(3)	9%(105)	3%(34)	0%(5)	—	13%(147)
	5	—	—	0%(3)	2%(19)	3%(30)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		1%(14)	32%(366)	57%(641)	6%(72)	3%(38)	—	1131
korrekt		79%	56%	54%	47%	79%	—	55%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(33)	5%(61)	1%(8)	—	—	—	9%(102)
	2	1%(11)	22%(244)	11%(122)	0%(5)	—	—	34%(382)
	3	—	8%(96)	28%(315)	3%(34)	0%(2)	—	40%(447)
	4	—	0%(1)	7%(78)	5%(61)	1%(7)	—	13%(147)
	5	—	—	0%(1)	1%(14)	3%(37)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		4%(44)	35%(402)	47%(524)	10%(114)	4%(47)	—	1131
korrekt		75%	61%	60%	54%	79%	—	61%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(28)	6%(66)	1%(8)	—	—	—	9%(102)
	2	2%(17)	20%(228)	12%(135)	0%(2)	—	—	34%(382)
	3	—	8%(90)	28%(319)	3%(36)	0%(2)	—	40%(447)
	4	—	0%(2)	7%(77)	5%(62)	1%(6)	—	13%(147)
	5	—	—	0%(1)	2%(17)	3%(34)	—	5%(52)
	6	—	—	—	0%(1)	—	—	0%(1)
total¹		4%(45)	34%(386)	48%(540)	10%(118)	4%(42)	—	1131
korrekt		62%	59%	59%	53%	81%	—	59%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.13: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Karlsruhe-Nordwest bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 50\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(39)	3%(38)	2%(25)	—	—	—	9%(102)
	2	7%(74)	11%(125)	15%(168)	1%(15)	—	—	34%(382)
	3	2%(27)	8%(88)	22%(250)	7%(76)	1%(6)	—	40%(447)
	4	0%(2)	1%(6)	6%(70)	5%(58)	1%(11)	—	13%(147)
	5	—	—	—	2%(20)	3%(32)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		13%(142)	23%(257)	45%(513)	15%(169)	4%(50)	—	1131
korrekt		27%	49%	49%	34%	64%	—	44%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	1%(8)	6%(70)	2%(24)	—	—	—	9%(102)
	2	1%(7)	16%(182)	17%(189)	0%(4)	—	—	34%(382)
	3	—	6%(71)	31%(355)	2%(18)	0%(3)	—	40%(447)
	4	—	0%(2)	9%(106)	3%(32)	1%(7)	—	13%(147)
	5	—	—	0%(4)	2%(20)	2%(28)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		1%(15)	29%(325)	60%(678)	7%(74)	3%(39)	—	1131
korrekt		53%	56%	52%	43%	72%	—	53%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(29)	5%(63)	1%(10)	—	—	—	9%(102)
	2	1%(18)	21%(218)	11%(141)	0%(4)	0%(1)	—	34%(382)
	3	0%(1)	8%(102)	28%(308)	3%(35)	0%(1)	—	40%(447)
	4	—	0%(2)	7%(77)	5%(61)	1%(7)	—	13%(147)
	5	—	—	0%(1)	1%(15)	3%(36)	—	5%(52)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		4%(48)	34%(385)	47%(537)	10%(115)	4%(46)	—	1131
korrekt		60%	57%	57%	53%	78%	—	58%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(30)	5%(61)	1%(11)	—	—	—	9%(102)
	2	2%(17)	20%(223)	12%(137)	0%(5)	—	—	34%(382)
	3	0%(1)	8%(88)	28%(318)	3%(38)	0%(2)	—	40%(447)
	4	—	0%(2)	6%(73)	6%(64)	1%(8)	—	13%(147)
	5	—	—	0%(1)	1%(14)	3%(37)	—	5%(52)
	6	—	—	—	0%(1)	—	—	0%(1)
total¹		4%(48)	33%(374)	48%(540)	11%(122)	4%(47)	—	1131
korrekt		63%	60%	59%	52%	79%	—	59%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.14: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Stuttgart-Bad Cannstatt bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 10\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(41)	3%(33)	1%(7)	0%(1)	—	—	8%(82)
	2	2%(22)	16%(161)	12%(124)	0%(5)	—	—	31%(312)
	3	—	6%(64)	29%(296)	5%(46)	0%(3)	—	41%(410)
	4	—	0%(2)	6%(65)	5%(53)	1%(8)	—	13%(128)
	5	—	—	0%(5)	4%(37)	3%(33)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		6%(64)	26%(260)	49%(497)	14%(142)	4%(45)	—	1008
korrekt		64%	62%	60%	37%	73%	—	58%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(18)	5%(54)	1%(9)	0%(1)	—	—	8%(82)
	2	1%(9)	18%(177)	12%(123)	0%(3)	—	—	31%(312)
	3	—	7%(74)	30%(302)	3%(29)	0%(5)	—	41%(410)
	4	—	0%(3)	7%(68)	5%(50)	1%(7)	—	13%(128)
	5	—	0%(1)	0%(5)	4%(36)	3%(33)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		3%(27)	31%(309)	50%(507)	12%(119)	5%(46)	—	1008
korrekt		67%	57%	60%	42%	72%	—	58%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(33)	4%(44)	0%(5)	—	—	—	8%(82)
	2	1%(15)	19%(191)	10%(103)	0%(3)	—	—	31%(312)
	3	—	7%(70)	29%(288)	5%(51)	0%(1)	—	41%(410)
	4	—	0%(1)	6%(58)	6%(60)	1%(9)	—	13%(128)
	5	—	0%(1)	0%(5)	2%(23)	5%(46)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		5%(48)	30%(307)	46%(459)	14%(137)	6%(57)	—	1008
korrekt		69%	62%	63%	44%	81%	—	61%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(28)	5%(49)	0%(5)	—	—	—	8%(82)
	2	2%(16)	18%(184)	11%(110)	0%(2)	—	—	31%(312)
	3	—	6%(65)	30%(301)	4%(39)	0%(5)	—	41%(410)
	4	—	0%(2)	7%(67)	5%(50)	1%(9)	—	13%(128)
	5	—	0%(1)	0%(3)	3%(31)	4%(39)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total ¹		4%(44)	30%(301)	48%(486)	12%(122)	5%(54)	—	1008
korrekt		64%	61%	62%	41%	72%	—	60%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.15: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Stuttgart-Bad Cannstatt bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 25\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(39)	3%(30)	1%(11)	0%(2)	—	—	8%(82)
	2	3%(34)	15%(148)	12%(121)	1%(9)	—	—	31%(312)
	3	0%(4)	7%(72)	26%(267)	6%(64)	0%(3)	—	41%(410)
	4	—	0%(4)	6%(59)	6%(56)	1%(9)	—	13%(128)
	5	—	0%(1)	1%(8)	3%(31)	3%(35)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		8%(77)	25%(255)	46%(466)	16%(162)	5%(48)	—	1008
korrekt		51%	58%	57%	35%	75%	—	54%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(19)	5%(52)	1%(10)	0%(1)	—	—	8%(82)
	2	1%(12)	17%(169)	13%(128)	0%(3)	—	—	31%(312)
	3	—	6%(65)	31%(314)	3%(26)	0%(5)	—	41%(410)
	4	—	0%(1)	7%(72)	5%(48)	1%(7)	—	13%(128)
	5	—	0%(1)	1%(6)	3%(34)	3%(34)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		3%(31)	29%(288)	53%(530)	11%(112)	5%(47)	—	1008
korrekt		61%	59%	59%	43%	72%	—	58%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(32)	4%(43)	1%(7)	—	—	—	8%(82)
	2	2%(21)	17%(173)	11%(115)	0%(3)	—	—	31%(312)
	3	—	7%(67)	29%(289)	5%(51)	0%(3)	—	41%(410)
	4	—	0%(3)	6%(60)	5%(53)	1%(12)	—	13%(128)
	5	—	0%(1)	1%(8)	2%(20)	5%(46)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		5%(53)	28%(287)	48%(479)	13%(127)	6%(62)	—	1008
korrekt		60%	60%	60%	42%	74%	—	59%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(30)	5%(47)	0%(5)	—	—	—	8%(82)
	2	2%(19)	17%(176)	11%(114)	0%(3)	—	—	31%(312)
	3	0%(1)	6%(61)	30%(303)	3%(40)	0%(5)	—	41%(410)
	4	—	0%(3)	6%(60)	6%(58)	1%(7)	—	13%(128)
	5	—	0%(1)	1%(6)	3%(30)	4%(38)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		5%(50)	29%(288)	48%(488)	13%(131)	5%(51)	—	1008
korrekt		60%	61%	62%	44%	75%	—	60%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

Tabelle D.16: Gegenüberstellung des LuQx aus Messungen und Prognosen verschiedener Modelle für die Station Stuttgart-Bad Cannstatt bei einer Abweichung der meteorologischen Vorhersagen um bis zu $\pm 50\%$

		LuQx Prognose (LR)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	4%(41)	3%(26)	1%(13)	0%(2)	—	—	8%(82)
	2	6%(62)	11%(113)	12%(120)	2%(17)	—	—	31%(312)
	3	3%(28)	6%(60)	23%(234)	8%(82)	1%(6)	—	41%(410)
	4	0%(3)	0%(3)	6%(56)	6%(58)	1%(8)	—	13%(128)
	5	—	—	1%(8)	3%(31)	4%(36)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		13%(134)	20%(202)	43%(431)	19%(190)	5%(51)	—	1008
korrekt		31%	56%	54%	31%	71%	—	48%

		LuQx Prognose (IBk)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	2%(22)	5%(48)	1%(12)	—	—	—	8%(82)
	2	1%(11)	17%(171)	13%(127)	0%(3)	—	—	31%(312)
	3	0%(1)	8%(82)	29%(292)	3%(31)	0%(4)	—	41%(410)
	4	—	0%(4)	7%(68)	5%(49)	1%(7)	—	13%(128)
	5	—	—	1%(6)	4%(38)	3%(31)	—	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		3%(34)	30%(305)	50%(505)	12%(121)	4%(43)	—	1008
korrekt		65%	56%	58%	40%	72%	—	56%

		LuQx Prognose (MLP)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	5%(50)	4%(39)	1%(9)	—	—	—	8%(82)
	2	4%(39)	16%(160)	11%(107)	1%(6)	—	—	31%(312)
	3	0%(4)	9%(87)	25%(248)	6%(65)	1%(6)	—	41%(410)
	4	—	1%(7)	4%(42)	6%(56)	1%(10)	—	13%(128)
	5	—	—	1%(11)	2%(20)	4%(40)	0%(1)	7%(72)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		9%(93)	29%(293)	41%(417)	15%(147)	6%(57)	—	1008
korrekt		54%	55%	59%	38%	70%	—	55%

		LuQx Prognose (SVM)						
		1	2	3	4	5	6	total ¹
LuQx Messung	1	3%(28)	4%(45)	1%(9)	—	—	—	8%(82)
	2	2%(19)	16%(165)	12%(124)	0%(4)	—	—	31%(312)
	3	0%(1)	7%(66)	29%(288)	5%(51)	0%(4)	—	41%(410)
	4	—	0%(4)	6%(57)	5%(55)	1%(12)	—	13%(128)
	5	—	—	0%(5)	3%(30)	4%(39)	0%(1)	7%(75)
	6	—	—	—	—	0%(1)	—	0%(1)
total¹		5%(48)	28%(280)	48%(483)	14%(140)	6%(56)	0%(1)	1008
korrekt		58%	59%	60%	39%	70%	0%	57%

¹eventuelle Abweichungen in den Prozentsummen sind auf Rundungsfehler zurückzuführen

ANHANG D. DETAILLIERTE ERGEBNISSE DER
SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
METEOROLOGISCHER VORHERSAGEN

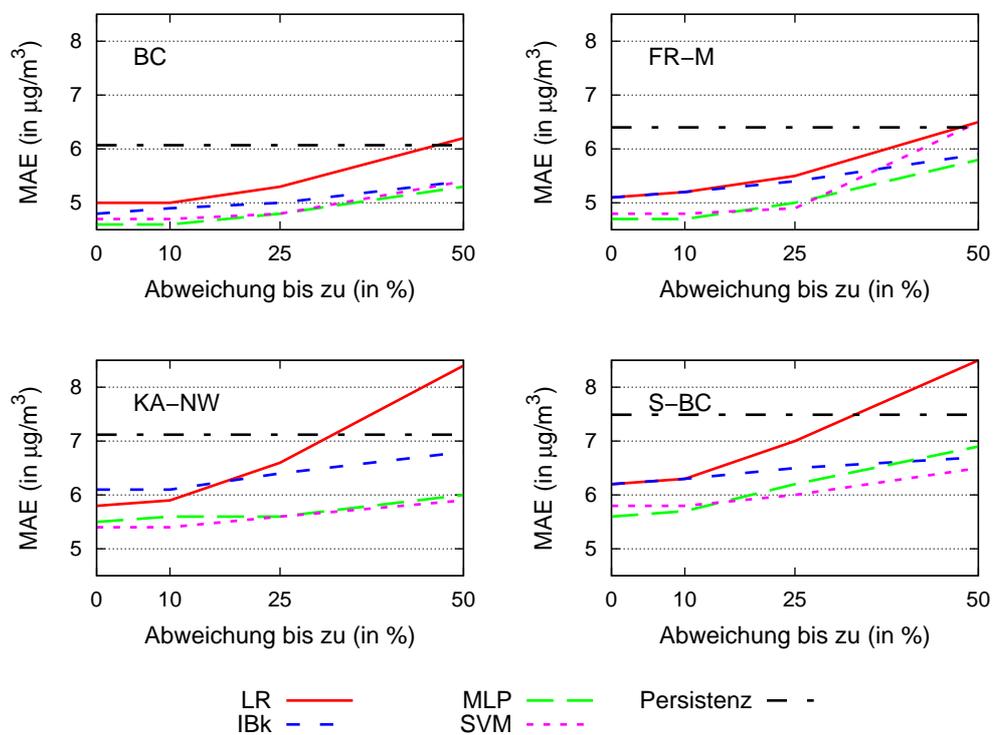


Abbildung D.1: Veränderung der PM_{10} -Prognosegenauigkeit (MAE) verschiedener Modelle bei Abweichung der meteorologischen Prognosedaten für Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt

ANHANG D. DETAILLIERTE ERGEBNISSE DER
 SENSITIVITÄTSANALYSE BEZÜGLICH UNSICHERER
 METEOROLOGISCHER VORHERSAGEN

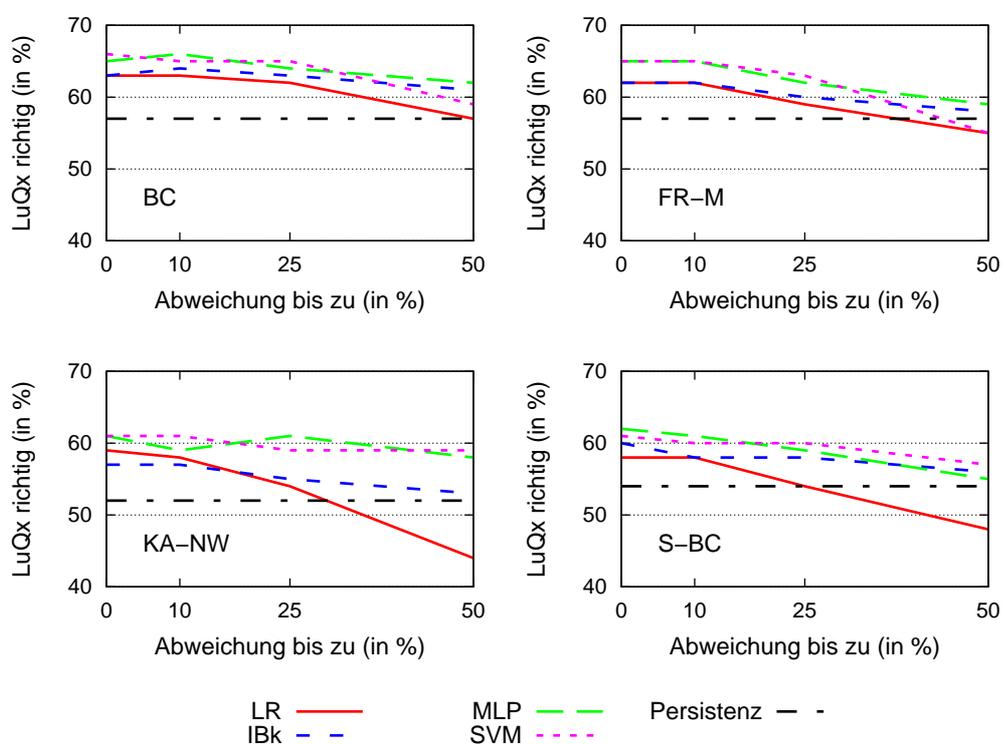


Abbildung D.2: Veränderung der PM₁₀-Prognosegenauigkeit (LuQx) verschiedener Modelle bei Abweichung der meteorologischen Prognosedaten für Biberach, Freiburg-Mitte, Karlsruhe-Nordwest und Stuttgart-Bad Cannstatt