

Dissertation

**Identifizierung und Modellierung von räumlichen
Abhängigkeiten mit Anwendung auf deterministische
und probabilistische Windvorhersagen**

Marc Hüsich

01. Oktober 2019

Gutachter: Prof. Dr. Walter Krämer

Dr. Uwe Ligges

Technische Universität Dortmund

Fakultät Statistik

Dissertation zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

Vorgelegt der Fakultät Statistik der Technischen Universität Dortmund

Dortmund, 01. Oktober 2019

Danksagung

Ich bedanke mich bei Prof. Dr. Walter Krämer für die Betreuung meiner Promotion und für die langjährige Förderung als studentische Hilfskraft und als wissenschaftlicher Mitarbeiter. Dr. Uwe Ligges danke ich für die Begutachtung dieser Dissertation und für hilfreiche Anregungen in meiner Promotionsphase. Prof. Dr. Joachim Kunert danke ich für die Übernahme des Vorsitzes der Prüfungskommission und Prof. Dr. Jörg Rahnenführer für die Protokollführung während meiner Disputation.

Außerdem danke ich Dr. Lueder von Bremen und Bruno Schyska vom DLR-Institut für Vernetzte Energiesysteme in Oldenburg vielmals für die wissenschaftliche Zusammenarbeit in den letzten drei Jahren. Der stetige Austausch von Ideen und Anregungen hat meine Arbeit sehr bereichert und war immer sehr spannend und informativ.

Ein großer Dank gilt auch meinen Arbeitskollegen am Lehrstuhl für Wirtschafts- und Sozialstatistik, auf die ich mich immer in allen Belangen verlassen konnte. Die Zeit am Lehrstuhl wird mir in sehr positiver Erinnerung bleiben.

Ich danke zudem dem Europäischen Zentrum für mittelfristige Wettervorhersage für die Bereitstellung von deterministischen und probabilistischen Windgeschwindigkeitsvorhersagen. Der Universität Hamburg danke ich für die Ausgabe von Windgeschwindigkeitsmessungen am Wettermast in Hamburg. Dem Bundesministerium für Wirtschaft und Energie (BMWi) sowie dem Projektträger Jülich (PtJ) danke ich für die Bereitstellung von Windgeschwindigkeitsmessungen an den Forschungsplattformen FINO2 und FINO3.

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Tabellenverzeichnis	vi
Mathematische Notation	ix
1 Einleitung	1
2 Problemstellung	5
2.1 Problemstellung aus energiemeteorologischer Sicht	6
2.2 Problemstellung aus statistischer Sicht	9
2.3 Ziele der Dissertation	10
2.4 Software	11
3 Identifizierung räumlicher Abhängigkeitsstrukturen mit Hilfe eines Clusteringverfahrens für räumlich-zeitliche Daten	13
3.1 Motivation	13
3.2 Beschreibung des Clusteringverfahrens	16
3.3 Anwendung: Identifizierung von räumlichen Abhängigkeitsstrukturen bei deterministischen Windkraftvorhersagefehlern	24
3.4 Erweiterung: Kombination von Clustering-Ergebnissen aus mehreren Teilregionen	33
3.5 Diskussion und Fazit	38

4	Identifizierung von räumlichen Abhängigkeiten mit Copulas und Generalisierten Additiven Modellen	41
4.1	Motivation	41
4.2	Methodik	43
4.3	Anwendung: Identifizierung von räumlichen Abhängigkeitsstrukturen bei deterministischen Windkraftvorhersagefehlern	55
4.4	Diskussion und Fazit	74
5	Bewertung von multivariaten probabilistischen Windvorhersagen unter Berücksichtigung räumlicher Abhängigkeitsstrukturen	77
5.1	Motivation	77
5.2	Datengrundlage und deskriptive Analyse	80
5.3	Methodik	85
5.4	Ergebnisse	92
5.5	Sensitivitätsanalyse für multivariate Bewertungsregeln	99
5.6	Diskussion und Fazit	103
6	Zusammenfassung	105
	Literaturverzeichnis	109

Abbildungsverzeichnis

3.1	Studienregion bei der durchgeführten Analyse mit Angaben zur zugrundeliegenden Modellorographie (Höhe über NN in Metern eines bestimmten Gitterpunktes)	25
3.2	Clusteringergebnis für 24-Stunden-Vorhersagefehler ($\varrho = 0.7$)	28
3.3	Clusteringergebnis für 72-Stunden-Vorhersagefehler ($\varrho = 0.7$)	29
3.4	Clusteringergebnis für 24-Stunden-Vorhersagefehler ($\varrho = 0.9$)	30
3.5	Clusteringergebnis für 72-Stunden-Vorhersagefehler ($\varrho = 0.9$)	31
3.6	Standorte die beim Clustering für die 24- und 72-Stunden-Vorhersagefehler keinem Cluster zugeordnet werden ($\varrho = 0.9$)	32
3.7	Clusteranalyse für 12-Stunden-Vorhersagefehler an Standorten in Österreich und der Schweiz ($\varrho = 0.7$)	34
3.8	Grenzcluster, die sich bei separaten Clusteranalysen ($\varrho = 0.7$) für 12-Stunden-Vorhersagefehler in Österreich und der Schweiz ergeben	35
3.9	Clusteranalyse für 12-Stunden-Vorhersagefehler an Standorten in Österreich und der Schweiz ($\varrho = 0.7$) mit der Erweiterung aus Abschnitt 3.4	37
4.1	Schätzung der Parameter einer t-Copula (wahre Parameter: $\rho = 0.7$ und $\nu = 7$) für $T = 500$ simulierte Pseudo-Beobachtungen	48
4.2	Schätzung der Parameter einer t-Copula (wahre Parameter: $\rho = 0.7$ und $\nu = 7$) für $T = 10\,000$ simulierte Pseudo-Beobachtungen	49
4.3	Zeitreihe der 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)	56
4.4	Empirische Autokorrelationen der 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)	57
4.5	Empirische Autokorrelationen der quadrierten 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)	58

4.6	Streudiagramm der 24-Stunden-Vorhersagefehler an den Standorten s_{17058} und s_{17075} (gelegen innerhalb einer Distanz von etwa 140 Kilometern in Norddeutschland)	59
4.7	Streudiagramm rangtransformierter 24-Stunden-Vorhersagefehler an den Standorten s_{17058} und s_{17075} (gelegen innerhalb einer Distanz von etwa 140 Kilometern in Norddeutschland)	60
4.8	Simulierte Pseudo-Beobachtungen aus einer t-Copula und einer Gauß-Copula mit Parametern, die anhand von 24-Stunden-Vorhersagefehlern an den beiden Standorten s_{17058} und s_{17075} geschätzt wurden	61
4.9	Geschätzte Werte $\hat{\rho}_{MPL}$ des Parameters ρ einer t-Copula für 24-Stunden-Vorhersagefehler, dargestellt gegenüber der räumlichen Distanz der Standortpaare.	63
4.10	Geschätzte Werte $\hat{\nu}_{MPL}$ des Parameters ν einer t-Copula für 24-Stunden-Vorhersagefehler, dargestellt gegenüber der räumlichen Distanz der Standortpaare.	64
4.11	Modellierung der Schätzwerte $\hat{\rho}_{MPL}$ des Parameters ρ einer t-Copula (geschätzt für $B = 478\,638$ Standortpaare) für die 24-Stunden-Vorhersagefehler mit einem GAM basierend auf der räumlichen Distanz h in Kilometern, der maximalen Höhe über dem Meeresspiegel in Metern (MaxH) und dem Orographie-Index (OI)	66
4.12	Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ und GAM-Modellvorhersagen $\hat{\rho}_{GAM}$ für den Parameter ρ einer t-Copula, welche die Abhängigkeitsstruktur zwischen Vorhersagefehlern am Standort s_{17058} (in Norddeutschland) und an benachbarten Standorten beschreibt	68
4.13	Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ und GAM-Modellvorhersagen $\hat{\rho}_{GAM}$ für den Parameter ρ einer t-Copula, welche die Abhängigkeitsstruktur zwischen Vorhersagefehlern am Standort s_{32787} (in der Schweiz) und an benachbarten Standorten beschreibt	68
4.14	Modellvorhersagen des Parameters ρ einer t-Copula für die 24-Stunden-Vorhersagefehler mit dem angepassten GAM basierend auf der räumlichen Distanz, der maximalen Höhe und dem Orographie-Index (OI)	70
4.15	Modellvorhersagen des Parameters ρ einer t-Copula und des Randabhängigkeitskoeffizienten λ für 24-Stunden- und 72-Stunden-Vorhersagefehler	72

4.16	72-Stunden-Vorhersagefehler in der Region um den Standort s_{17058} am 04.02.2011 um 12:00 Uhr UTC	73
5.1	Lage der drei verschiedenen Standorte, für die Windgeschwindigkeitsmessungen und Vorhersagen vorliegen	80
5.2	Histogramme der gemessenen Windgeschwindigkeiten an den Standorten Hamburg, FINO2 und FINO3	82
5.3	Deterministische und Ensemble-Windgeschwindigkeitsvorhersage für FINO2 ausgegeben am 25. Juni 2011 um 00:00 Uhr UTC für verschiedene Laufzeiten (3 Std. - 120 Std.) verglichen mit beobachteten Windgeschwindigkeiten	83
5.4	Empirische Kreuzkorrelationen (Lag 0) zwischen Vorhersagefehlern der Kontrollvorhersage an den Standorten FINO2 und FINO3 für verschiedene Vorhersagehorizonte	84
5.5	Continuous Ranked Probability Skill Score (CRPSS) für EMOS-Vorhersagen verglichen mit dem originalen (raw) Ensemble für die Station FINO2 . . .	93
5.6	Multivariate Skill-Scores für EMOS-Vorhersagen verglichen mit dem originalen (raw) Ensemble (gemeinsame Bewertung für die Standorte FINO2 und FINO3)	95
5.7	Multivariate Skill-Scores für ECCQ-Vorhersagen verglichen mit EMOS-Vorhersagen (gemeinsame Bewertung für die Standorte FINO2 und FINO3)	96
5.8	Multivariate Skill-Scores für ECCQ-Vorhersagen verglichen mit EMOS-Vorhersagen (gemeinsame Bewertung für die Standorte FINO2, FINO3 und Hamburg)	98
5.9	Simulierte Skill-Scores für bivariate ECCQ-artige Vorhersagen mit einer festen Anzahl von $M = 51$ Ensemble-Mitgliedern verglichen mit EMOS-artigen Vorhersagen für variierende räumliche Abhängigkeiten	100
5.10	Simulierte Skill-Scores für bivariate ECCQ-artige Vorhersagen mit einer variierenden Anzahl von M Ensemble-Mitgliedern verglichen mit EMOS-artigen Vorhersagen für eine festgelegte räumliche Abhängigkeitsstruktur .	102

Tabellenverzeichnis

3.1	Resultierende Anzahl an Clustern für verschiedene Vorhersagehorizonte und Korrelationsstärken	27
3.2	Laufzeit des Clusteringverfahrens CorClustST in Sekunden für verschiedene Werte von ϵ und N für die 24-Stunden-Vorhersagefehler ($\varrho = 0.9$)	39
4.1	Übersicht populärer bivariater Copula-Familien mit wichtigen Eigenschaften	46

Mathematische Notation

B	Anzahl an gewählten räumlichen Punktpaaren zur Copula-Schätzung
$\beta_{0,q}, \beta_{1,q}, \dots$	Parameter eines Generalisierten Additiven Modells
C	Copula
$\mathcal{C}_1, \dots, \mathcal{C}_{N_C}$	Räumliche Cluster resultierend aus der Clusteranalyse
$\delta_1, \dots, \delta_Q$	Copula-Parameter
Δ_q	Parameterraum des Copula-Parameters δ_q
ϵ	Grenzwert für die räumliche ϵ -Umgebung bei der Clusteranalyse
$f_{k,q}$	Glatte Funktionen der erklärenden Variablen im Generalisierten Additiven Modell
F_{Z_s}	Verteilung der räumlichen Zufallsvariable $Z(s)$
$F_{Z_{s_i}, Z_{s_j}}$	Gemeinsame Verteilung der räumlichen Zufallsvariablen $Z(s_i)$ und $Z(s_j)$
$\mathcal{F}_{s,t}$	Probabilistische Vorhersage der Windgeschwindigkeit am Standort s zum Zeitpunkt t
g_q	Link-Funktion, die bei der Modellierung mit Generalisierten Additiven Modellen verwendet wird
G_1, \dots, G_K	Geographische Einflussvariablen zur Modellierung der räumlichen Abhängigkeitsstruktur
γ_1, γ_2	Grenzwerte für die Bestimmung von inneren und äußeren Grenzclustern bei der Clusteranalyse

h	Räumliche Distanz zwischen zwei Standorten
$H(s)$	Höhe über NN des Standorte s
K	Anzahl der räumlichen Einflussvariablen zur Modellierung der räumlichen Abhängigkeitsstruktur
$\ell(\boldsymbol{\delta})$	Copula Pseudo-Log-Likelihood-Funktion
λ_L, λ_U	Unterer bzw. oberer Randabhängigkeitskoeffizient
M	Anzahl an Ensemble-Mitgliedern bei einer meteorologischen Ensemble-Vorhersage
$\mu_{s,t}$	Mittelwertparameter einer trunkierten Normalverteilung (verwendet als prob. Vorhersage für die Windgeschwindigkeit am Standort s zum Zeitpunkt t)
N	Anzahl an räumlichen Standorten im Datensatz
N_C	Anzahl an gefundenen Clustern bei der Clusteranalyse
\mathcal{N}^0	Trunkierte Normalverteilung mit einem unterem Grenzwert von 0
ν	Parameter für die Freiheitsgrade einer t-Copula
$O = \{o_1, \dots, o_N\}$	Geordnete Liste von Standorten beim Clustering nach ihrer Anzahl an räumlich-zeitlichen Nachbarn
ϱ	Grenzwert für die räumlich-zeitliche ϱ -Umgebung bei der Clusteranalyse
ρ	Korrelationsparameter einer t-Copula bzw. Gauß-Copula
Q	Anzahl der Copula-Parameter
s_1, \dots, s_N	Räumliche Standorte im Datensatz
\mathcal{S}	Menge der räumlichen Standorte im Datensatz
$\sigma_{s,t}^2$	Varianzparameter einer trunkierten Normalverteilung (verwendet als prob. Vorhersage für die Windgeschwindigkeit am Standort s zum Zeitpunkt t)

$S_{s,t}^2$	Empirische Varianz der Ensemble-Vorhersagen am Standort s zum Zeitpunkt t
t	Zeitpunkt
t_ν	t -Verteilung mit ν Freiheitsgraden
T	Anzahl an Zeitpunkten im Datensatz
τ	Vorhersagehorizont der Windvorhersage
$u(s, t)$	Rangtransformierte Pseudo-Beobachtung am Standort s zum Zeitpunkt t
$X_1(s, t), \dots, X_M(s, t)$	Ensemble-Vorhersagen für die Windgeschwindigkeit am Standort s zum Zeitpunkt t
Ξ	Menge der Standorte, die bei der räumlichen Clusteranalyse keinem Cluster zugeordnet werden
$y(s, t)$	Beobachtete Windgeschwindigkeit am Standort s zum Zeitpunkt t
$Y(s, t)$	Räumlich-zeitliche Zufallsvariable für die Windgeschwindigkeit am Standort s und Zeitpunkt t
$z(s, t)$	Beobachtete Windkraftvorhersagefehler am Standort s zum Zeitpunkt t
$Z(s)$	Räumliche Zufallsvariable für Vorhersagefehler am Standort s (zeitunabhängig)
$\zeta_s(m)$	Rang des Ensemble-Mitglieds m unter allen Ensemble-Mitgliedern am Standort s für einen festgelegten Zeitpunkt

1 Einleitung

In vielen Naturwissenschaften wie etwa der Hydrologie oder der Meteorologie nimmt die statistische Analyse von räumlich-zeitlichen Datensätzen eine wichtige Rolle ein. So werden statistische Verfahren beispielsweise dazu eingesetzt, um räumliche Unterschiede bei Niederschlagsmengen und Temperaturveränderungen zu analysieren (siehe zum Beispiel Carrera-Hernández und Gaskin, 2007) oder um regionale Wetterextremereignisse zu modellieren (siehe zum Beispiel Coles, 1993). In empirischen Analysen ist es oftmals von großer Bedeutung, Abhängigkeiten zwischen den Beobachtungen an verschiedenen Standorten zu analysieren und geeignet zu modellieren. Im Zuge von immer größer werdenden Datenmengen müssen die verwendeten statistischen Verfahren dabei sehr effizient sein und gleichzeitig auch eine hohe Aussagekraft gewährleisten. In dieser Dissertation beleuchten wir einen Anwendungsfall bei dem räumliche Abhängigkeitsstrukturen von besonderer Relevanz sind und legen den Fokus auf Windgeschwindigkeits- bzw. Windleistungsvorhersagen. Die fluktuierende Erzeugung aus Windkraftanlagen kann anders als bei konventionellen Energieträgern im Vorfeld nicht exakt geplant werden. Energiehändler und Übertragungsnetzbetreiber greifen für die Planung bzw. die Steuerung von Produktionsmengen deshalb auf Vorhersagen der zukünftigen Windeinspeisung zurück. Meist sind diese Prognosen von deterministischer Natur, das heißt sie erlauben im Vorfeld keine Einschätzung über die vorliegende Vorhersageunsicherheit. Diese deterministischen Prognosen (im statistischen Sinne auch als Punktprognosen bezeichnet) müssen daher möglichst akkurat sein, da bei großen Vorhersagefehlern hohe Kosten für Energiemarktteilnehmer resultieren können oder ein Risiko für die Netzsicherheit entstehen kann. Besonders hohe

Vorhersagefehler für eine gesamte Region ergeben sich, wenn Prognosefehler gleichzeitig an benachbarten Standorten auftreten und somit wenn eine ausgeprägte räumliche Abhängigkeitsstruktur in den Vorhersagefehlern vorhanden ist.

Unter Verwendung von Vorhersagen des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) untersuchen wir deshalb in dieser Dissertation, wie sich räumliche Abhängigkeitsstrukturen bei deterministischen Vorhersagefehlern in verschiedenen Regionen in Europa und für verschiedene Vorhersagehorizonte unterscheiden. Für die Analyse verwenden wir zunächst ein eigens entwickeltes Clusteringverfahren (CorClustST), das erste Einblicke in räumliche Abhängigkeitsstrukturen bei räumlich-zeitlichen Datensätzen ermöglicht. Das Verfahren wurde in der Masterarbeit von Hüscher (2016) entwickelt und hier auch bereits für die Analyse von räumlichen Abhängigkeiten bei deterministischen Vorhersagefehlern benutzt. Im Rahmen dieser Dissertation wird das Clusteringverfahren insofern erweitert, dass auch eine Clusteranalyse von großen räumlich-zeitlichen Datensätzen ermöglicht wird. Die gesamte Methodik sowie die zentralen Ergebnisse sind in folgendem Artikel enthalten:

Hüscher, M., B. U. Schyska, L. von Bremen (2018). „CorClustST - Correlation-based clustering of big spatio-temporal datasets“, *Future Generation Computer Systems*. URL: <https://doi.org/10.1016/j.future.2018.04.002>.

Für eine noch detailliertere Analyse von räumlichen Abhängigkeitsstrukturen schlagen wir in dieser Dissertation außerdem einen copulabasierten Ansatz vor, mit dem insbesondere der Einfluss verschiedener geographischer Bedingungen gezielter untersucht werden kann. Mit Hilfe der beiden vorgeschlagenen Methoden charakterisieren wir die räumliche Abhängigkeitsstruktur von deterministischen Windleistungsvorhersagefehlern für unterschiedliche geographische Gegebenheiten und zeigen bedeutende Unterschiede für verschiedene Regionen in Europa auf.

Neben deterministischen Windvorhersagen werden in der Praxis zunehmend auch probabilistische Vorhersagen eingesetzt, die hilfreiche Aussagen über die Vorhersageunsicherheit ermöglichen. Dazu wird meist auf meteorologische Ensemble-Vorhersagen zurückgegriffen,

die ein Bündel von einzelnen Vorhersagen darstellen. Diese sollen die zugrundeliegende Vorhersageunsicherheit möglichst zuverlässig abbilden. Mit Hilfe von solchen probabilistischen Vorhersagen für mehrere Standorte kann zum Beispiel im Vorfeld das Risiko eingeschätzt werden, dass sehr große Abweichungen von der deterministischen Vorhersage gemeinsam für die betrachteten Standorte auftreten. Dazu müssen die Vorhersagen jedoch die zugrundeliegende räumliche Abhängigkeitsstruktur adäquat berücksichtigen. Im abschließenden Teil dieser Dissertation widmen wir uns deshalb der Frage, wie die Qualität multivariater probabilistischer Windgeschwindigkeitsvorhersagen unter Berücksichtigung räumlicher Abhängigkeitsstrukturen sinnvoll bewertet werden kann. Dabei gehen wir auf potenzielle Schwierigkeiten ein, die speziell bei unterschiedlichen Strukturen von probabilistischen Vorhersagen entstehen.

Diese Dissertation ist wie folgt strukturiert: In Kapitel 2 wird die Problemstellung aus energiemeteorologischer und statistischer Perspektive ausführlich dargelegt. Außerdem werden die Ziele der Dissertation und die für die Analysen verwendete statistische Software benannt. Kapitel 3 stellt das eigens entwickelte Clusteringverfahren für räumlich-zeitliche Datensätze vor und wendet die Methode für die Identifizierung von räumlichen Abhängigkeiten bei deterministischen Windleistungsvorhersagefehlern an. In Kapitel 4 wird daraufhin ein copulabasiertes Verfahren zur Identifizierung von räumlichen Abhängigkeiten vorgeschlagen, mit dem die Abhängigkeitsstruktur der Vorhersagefehler noch detaillierter beschrieben wird. Kapitel 5 beschäftigt sich dann mit der Bewertung von multivariaten probabilistischen Windvorhersagen unter Berücksichtigung von räumlichen Abhängigkeitsstrukturen. In Kapitel 6 fassen wir die Ergebnisse dieser Arbeit zusammen und diskutieren Ideen für weiterführende Forschungsarbeiten.

2 Problemstellung

Die Analyse von räumlichen Abhängigkeitsstrukturen ist in vielen Wissenschaftsfeldern von hoher Bedeutung. Insbesondere in der Meteorologie, aber auch in anderen Naturwissenschaften liegen Datensätze häufig in einer hohen räumlichen und zeitlichen Auflösung vor. Statistische Verfahren für die Identifizierung und Modellierung räumlicher Abhängigkeiten müssen deshalb die besondere Struktur des zugrundeliegenden Datenmaterials berücksichtigen. Einerseits sollten die Verfahren möglichst zuverlässige Aussagen über räumliche Abhängigkeitsstrukturen ermöglichen, gleichzeitig müssen die Methoden bei sehr großen Datensätzen aber auch sehr effizient sein. Diese Dissertation beschäftigt sich aus diesem Grund mit statistischen Methoden, die speziell für die Modellierung und Identifizierung von räumlichen Abhängigkeitsstrukturen geeignet sind. Der Anwendungsfokus liegt in dieser Dissertation auf Windkraft- bzw. Windgeschwindigkeitsvorhersagen, die im Zeitalter erneuerbarer Energien eine sehr hohe Relevanz für den Energiesektor besitzen. Im folgenden Unterabschnitt erläutern wir zunächst, warum die Identifizierung und Modellierung räumlicher Abhängigkeitsstrukturen bei Windvorhersagen aus energiemeteorologischer Sicht bedeutend ist. Aus der energiemeteorologischen Problemstellung ergeben sich Anforderungen an statistische Verfahren, die im darauffolgenden Unterabschnitt detailliert erläutert werden. Es folgt ein kurzer Überblick über die Ziele dieser Dissertation. Im abschließenden Teil des Kapitels wird außerdem die Software benannt, die für die Analysen in den darauffolgenden Kapiteln verwendet wird.

2.1 Problemstellung aus energiemeteorologischer Sicht

Die Transformation des Energiesektors nimmt in vielen Industriestaaten immer mehr Gestalt an. So deckten erneuerbare Energiequellen in Deutschland im Jahr 2018 laut dem Umweltbundesamt (2019) bereits „fast 38 Prozent des gesamten deutschen Bruttostromverbrauchs“. Es ist zu erwarten, dass sich dieser Anteil in den kommenden Jahren aufgrund von ausgesprochenen Klimaschutzzielen noch weiter erhöhen wird. Unter den erneuerbaren Energieträgern nimmt die Windenergie in Deutschland eine tragende Rolle ein. Nach Angaben von Statista (2019) lag der Anteil der Stromerzeugung aus Windkraftanlagen an der gesamten Bruttostromerzeugung aus erneuerbaren Energiequellen im Jahr 2018 bei etwa 49 Prozent, und damit deutlich vor anderen erneuerbaren Energieträgern wie der Photovoltaik. Anders als bei konventionellen Energieträgern wie der Braunkohle oder der Kernenergie ist die Stromproduktion aus Windkraftanlagen jedoch nicht exakt im Voraus planbar, sondern hängt stark von den vorherrschenden Wetterbedingungen ab. An Energiebörsen wie z.B. der EEX (European Energy Exchange) in Leipzig werden Strommengen für einen bestimmten Tag aber auf sogenannten Day-Ahead-Märkten bereits am Vortag oder auf Terminmärkten sogar bis zu sechs Jahre im Voraus gehandelt (siehe EHA, 2017). Die zukünftige Produktion aus Windkraftanlagen muss daher im Vorfeld möglichst exakt prognostiziert werden, da fehlende Strommengen meist durch den Zukauf von Strom aus in der Regel teureren Energiequellen ausgeglichen werden müssen. Fehlprognosen können daher zu hohen Kosten für einzelne Energiemarktteilnehmer führen. Auch für Übertragungsnetzbetreiber ist die vorhergesagte Stromproduktion aus Windkraftanlagen von enormer Bedeutung. Die Netzbetreiber müssen das Stromangebot und die Stromnachfrage stets im Gleichgewicht halten um die Netzstabilität zu gewährleisten (siehe EnergieAgentur.NRW, 2017). Vorhersagefehler können hier zu einem kurzfristigen Ungleichgewicht führen, weshalb die Übertragungsnetzbetreiber in diesen Fällen sogenannte Regelernergie benötigen um fehlende oder überschüssige Produktionsmengen auszugleichen (siehe Bundesnetzagentur, 2019).

Aufgrund der hohen praktischen Relevanz widmen sich neben kommerziellen Dienstleistern auch viele aktuelle Forschungsarbeiten der Frage, wie möglichst zuverlässige Windleistungsvorhersagen für einzelne Windkraftanlagen oder gesamte Regionen getroffen werden können. Diese Vorhersagen beruhen meist auf meteorologischen Windgeschwindigkeitsprognosen, die von Wetteranbietern wie dem Deutschen Wetterdienst (DWD) oder dem Europäischen Zentrum für mittelfristige Wettervorhersage (ECMWF) für verschiedene Zeithorizonte ausgegeben werden. Dabei unterscheidet man häufig zwischen deterministischen und probabilistischen Vorhersagen. Deterministische Vorhersagen können im statistischen Sinne als Punktvorhersagen für einen spezifischen Zeitpunkt interpretiert werden. Diese werden von den Wetterdienstleistern meist auf einem räumlichen Gitter für ein gesamtes Gebiet bereitgestellt. Die deterministische Vorhersage liegt oftmals in einer sehr hohen räumlichen Auflösung vor und liefert daher in der Regel eine hohe Informationsdichte. So bietet zum Beispiel das deterministische Wettermodell HRES des Europäischen Zentrums für mittelfristige Wettervorhersage laut ECMWF (2019) die zuverlässigste und höchstauflöste Vorhersage unter allen Modellen des Wetterdienstleisters. Die Vorhersagequalität deterministischer Vorhersagen nimmt jedoch intuitiv mit zunehmendem Vorhersagehorizont ab, was generell durch höhere Vorhersagefehler sichtbar wird. In einer empirischen Studie von Focken et al. (2002) wird außerdem angemerkt, dass die räumliche Korrelation von Vorhersagefehlern für längere Vorhersagehorizonte ausgeprägter ist. Eine ausgeprägte räumliche Abhängigkeitsstruktur von Vorhersagefehlern kann mit hohen aggregierten Vorhersagefehlern in gesamten Regionen einhergehen, die sowohl für den Energiehandel als auch für Übertragungsnetzbetreiber ein Risiko darstellen. Allerdings beschränkt sich die genannte Studie auf einzelne Windkraftanlagen in Deutschland und gibt keinen genauen Einblick, in wie weit geographische Faktoren wie beispielsweise die Orographie (Geländebedingungen) die räumliche Abhängigkeit von Vorhersagefehlern beeinflussen. Aus diesem Grund verwenden wir in dieser Dissertation großflächige Daten des ECMWF, um detailliertere Erkenntnisse über die räumliche Abhängigkeitsstruktur von deterministischen Windleistungsvorhersagefehlern zu erhalten.

Ein Nachteil deterministischer Vorhersagen ist, dass diese keine Informationen über die vorliegende Vorhersageunsicherheit liefern. Aus diesem Grund werden im Sinne des Risikomanagements häufig meteorologische Ensemble-Vorhersagen berücksichtigt, die auch als probabilistische Prognosen bezeichnet werden. Ein Ensemble-System besteht dabei aus einem Bündel von einzelnen Vorhersagen (sogenannte Ensemble-Mitglieder), mit denen die Vorhersageunsicherheit möglichst zuverlässig abgebildet werden soll. So bietet beispielsweise das Europäische Zentrum für mittelfristige Wettervorhersage neben der deterministischen Prognose (HRES) zusätzlich ein Ensemble-System mit 51 Mitgliedern an (siehe ECMWF, 2019). Dieses besteht den Angaben nach aus einer Kontrollvorhersage, die im Vergleich zum deterministischen Modell zwar eine geringere räumliche Auflösung besitzt, aber die gegenwärtigen atmosphärischen Bedingungen ebenfalls bestmöglich abbildet. Bei den 50 weiteren Ensemble-Mitgliedern werden laut ECMWF (2019) jeweils die Anfangszustände des Wettermodells sowie die Modellphysik leicht verändert, sodass eine Reihe von Szenarien der zukünftigen Wetterentwicklung bereitgestellt werden kann. Die Ensemble-Vorhersagen werden wie die deterministische Vorhersage ebenfalls auf einem räumlichen Gitter veröffentlicht. Mit Hilfe von qualitativ hochwertigen Ensemble-Vorhersagen kann zum Beispiel das Risiko abgeschätzt werden, dass hohe Abweichungen von der deterministischen Vorhersage gemeinsam in einer bestimmten Region auftreten und somit zu einem hohen aggregierten Prognosefehler führen. Dazu ist es natürlich notwendig, dass die Vorhersagen die räumliche Abhängigkeitsstruktur zwischen Beobachtungen an verschiedenen Standorten auch adäquat repräsentieren. In der Praxis stellt sich oft die Frage, wie die Qualität verschiedener Ensemble-Vorhersagen unter Berücksichtigung räumlicher Abhängigkeitsstrukturen sinnvoll miteinander verglichen werden kann. Diese Dissertation soll deshalb auch einen Einblick liefern, welche Probleme bei der Bewertung von Ensemble-Vorhersagen im Falle von ausgeprägten räumlichen Abhängigkeitsstrukturen auftreten können und wie Fehlinterpretationen in empirischen Studien vermieden werden können.

2.2 Problemstellung aus statistischer Sicht

Die meteorologischen Windvorhersagen werden von den Wetteranbietern, wie im vorherigen Abschnitt beschrieben, meist auf einem hochaufgelösten räumlichen Gitter für bestimmte Regionen zur Verfügung gestellt. Für die Analyse von räumlichen Abhängigkeitsstrukturen bei deterministischen Vorhersagefehlern verwenden wir zum Beispiel Vorhersagedaten für das gesamte europäische Festland, die vom ECMWF für einen Zeitraum von über fünf Jahren bereitgestellt werden. Dadurch liegt für die Analyse eine Vielzahl an Datenpunkten vor, weshalb statistische Verfahren zur Identifizierung von räumlichen Abhängigkeitsstrukturen spezifische Anforderungen erfüllen müssen. In einem ersten Schritt einer statistischen Analyse wird das Datenmaterial oftmals zunächst deskriptiv untersucht. Falls räumlich-zeitliche Daten jedoch in einer sehr hohen räumlichen Auflösung vorliegen, gestaltet sich eine deskriptive Analyse von räumlichen Abhängigkeitsstrukturen eher schwierig, da der Fokus meist nur auf einzelne Standorte oder Teilregionen gelegt werden kann. Einen Ausweg bieten hier Clusteringverfahren, mit denen verschiedene Standorte mit ähnlichen Eigenschaften effizient in Gruppen eingeteilt werden können. Herkömmliche Clusteringverfahren wurden jedoch nicht explizit für räumlich-zeitliche Daten entwickelt, weshalb diese die spezielle Charakteristik dieser Daten nicht berücksichtigen. Dies ist insbesondere aus rechentechnischen Gründen problematisch. Es besteht deshalb ein Bedarf an alternativen Clusteringalgorithmen, die erste Einblicke in die räumliche Abhängigkeitsstruktur der Daten ermöglichen.

Für eine detaillierte Analyse der räumlichen Abhängigkeitsstruktur von Windleistungsvorhersagefehlern wird zudem ein flexibles statistisches Verfahren benötigt, mit dem die Auswirkungen unterschiedlicher geographischer Gegebenheiten genauer untersucht werden können. Speziell bei meteorologischen Daten ist zu erwarten, dass räumliche Abhängigkeitsstrukturen stark von den geographischen Bedingungen wie zum Beispiel der Orographie beeinflusst werden. Klassische geostatistische Verfahren zur Analyse und Modellierung von räumlichen Abhängigkeiten treffen jedoch oftmals sehr restriktive Annahmen,

wie etwa, dass die Abhängigkeitsstruktur nur von der räumlichen Distanz der Standorte beeinflusst wird. Aus diesem Grund soll ein Schwerpunkt dieser Dissertation auch auf der Entwicklung eines statistischen Verfahrens liegen, mit dem räumliche Abhängigkeitsstrukturen bei räumlich-zeitlichen Daten für unterschiedliche geographische Gegebenheiten charakterisiert werden können.

Ein weiteres Augenmerk soll außerdem auf die Bewertung von meteorologischen Ensemble-Vorhersagen für mehrere Standorte gelegt werden. Dafür müssen aus statistischer Sicht zunächst sinnvolle Bewertungsregeln für multivariate probabilistische Vorhersagen definiert werden. Bei den zugrundeliegenden meteorologischen Ensemble-Vorhersagen muss zudem beachtet werden, dass diese zum Teil eine systematische Verzerrung oder eine zu geringe Vorhersagevarianz aufweisen. Um die Qualität der Vorhersagen zu verbessern werden deshalb statistische Nachverarbeitungsverfahren eingesetzt. Die nachverarbeiteten Vorhersagen weisen im Vergleich zu den originalen Ensemble-Vorhersagen eine andere Struktur auf, da diese zum Beispiel in Form von kompletten Wahrscheinlichkeitsverteilungen ausgegeben werden. Beim Vergleich der Qualität von probabilistischen Vorhersagen ist deshalb insbesondere von Interesse, unter welchen Umständen gängige Bewertungsregeln fehlspezifizierte räumliche Abhängigkeitsstrukturen zuverlässig identifizieren können und wie sich unterschiedliche Ensemble-Strukturen auf die Bewertung auswirken.

2.3 Ziele der Dissertation

Aus der energiemeteorologischen und der statistischen Problemstellung ergeben sich folgende Zielsetzungen, die im Rahmen dieser Dissertation verfolgt werden:

1. Die Entwicklung statistischer Verfahren zur Identifizierung von räumlichen Abhängigkeitsstrukturen in räumlich-zeitlichen Datensätzen.
2. Die Identifizierung von räumlichen Abhängigkeitsstrukturen bei deterministischen Windleistungsvorhersagefehlern für unterschiedliche geographische Bedingungen.

3. Die Bewertung von multivariaten probabilistischen Windvorhersagen unter Berücksichtigung von räumlichen Abhängigkeitsstrukturen.

In den folgenden Kapiteln 3 und 4 liegt der Fokus auf der Bearbeitung der ersten beiden Zielstellungen. Zunächst wird dazu in Kapitel 3 ein Clusteringverfahren für große räumlich-zeitliche Datensätze vorgeschlagen. In Kapitel 4 wird daraufhin ein copulabasiertes Verfahren für die detaillierte Identifizierung von räumlichen Abhängigkeitsstrukturen vorgestellt. Beide Methoden werden zur Identifizierung von Abhängigkeitsstrukturen bei deterministischen Windleistungsvorhersagefehlern in Europa angewandt. Kapitel 5 widmet sich dann der Bewertung von Ensemble-Windgeschwindigkeitsvorhersagen und der Frage, ob fehlspezifizierte räumliche Abhängigkeitsstrukturen in den Vorhersagen mit gängigen Bewertungsregeln auch im Falle von unterschiedlichen Ensemble-Strukturen zuverlässig identifiziert werden können.

2.4 Software

Für die Analysen in den folgenden Kapiteln verwenden wir die statistische Programmiersprache R (R Core Team, 2018). Zudem greifen wir auf folgende R-Pakete zurück: Räumliche Visualisierungen werden unter Zuhilfenahme der Pakete `ggplot2` (Wickham, 2016), `sp` (Pebesma und Bivand, 2005), `rgdal` (Bivand et al., 2017), `gridExtra` (Auguie, 2017), `ggmap` (Kahle und Wickham, 2013), `maps` (Becker et al., 2017) und `ggalt` (Rudis et al., 2017) erstellt. Das Paket `fields` (Nychka et al., 2015) wird für die Berechnung von räumlichen Distanzen verwendet. Für copulabasierte Analysen werden die Pakete `copula` (Hofert et al., 2017) und `VineCopula` (Schepsmeier et al., 2017) benutzt. Die Pakete `lhs` (Carnell, 2018) und `FNN` (Beygelzimer et al., 2013) werden für die Auswahl von geeigneten Standortpaaren in Kapitel 4 verwendet. Für die Anpassung von Generalisierten Additiven Modellen nutzen wir das Paket `mgcv` (Wood, 2017). Für die Analyse in Kapitel 5 werden zudem die Pakete `truncnorm` (Trautmann et al., 2014), `scoringRules` (Jordan et al., 2017) und `ensembleMOS` (Yuen et al., 2018) verwendet.

3 Identifizierung räumlicher Abhängigkeitsstrukturen mit Hilfe eines Clusteringverfahrens für räumlich-zeitliche Daten

3.1 Motivation

Im Zeitalter exponentiell wachsender Datenmengen gewinnen effiziente Algorithmen, die Muster in sehr großen Datensätzen erkennen können, in vielen Forschungsgebieten zunehmend an Bedeutung. Eine zentrale Rolle in empirischen Analysen nehmen dabei Clusteringverfahren ein (siehe Kaufman und Rousseeuw, 1990; Everitt et al., 2011; Ericson und Pallickara, 2013), mit denen systematisch Gruppen von Objekten mit ähnlichen Eigenschaften identifiziert werden können. Diese Verfahren eignen sich insbesondere dazu, große Datensätze vor zu verarbeiten oder zu komprimieren. Vor allem in meteorologischen oder anderen naturwissenschaftlichen Anwendungen liegen Daten oft in einer relativ hohen räumlichen und zeitlichen Auflösung vor, weshalb diese eine spezielle Charakteristik aufweisen. Populäre Clusteringverfahren, wie der k -means-Algorithmus (siehe Hartigan und Wong, 1979) oder hierarchische Clusteringansätze (siehe Sibson, 1973; Defays, 1977; Ward, 1963), wurden jedoch nicht speziell für solche räumlich-zeitlichen Datensätze ent-

wickelt und berücksichtigen daher nicht die besonderen Eigenschaften des zugrundeliegenden Datenmaterials. Eine Übersicht über häufig verwendete Clusteringmethoden bei räumlich-zeitlichen Datensätzen liefert die Studie von Kisilevich et al. (2010). Laut den Autoren basieren die meisten Methoden, die zur Clusteranalyse von räumlich-zeitlichen Daten angewandt werden, auf dichte-basierten Clusteringverfahren wie z.B. dem Algorithmus DBSCAN (siehe Ester et al., 1996). Dieser klassifiziert alle Objekte hinsichtlich der Anzahl an benachbarten Objekten, die ähnliche Eigenschaften bezüglich der beobachteten Werte aufweisen. Birant und Kut (2007) erweitern den Algorithmus DBSCAN speziell für räumlich-zeitliche Datensätze. Deren Algorithmus ST-DBSCAN berücksichtigt zusätzlich zur Ähnlichkeit der beobachteten Werte auch die räumliche Distanz zwischen zwei Standorten, und erlaubt dadurch eine effizientere Einteilung von Standorten in verschiedene Cluster.

Auch wenn in jüngerer Zeit weitere Clusteringverfahren (wie z.B. der von Agrawal et al. (2016) vorgeschlagene Algorithmus ST-OPTICS) speziell für räumlich-zeitliche Datensätze entwickelt wurden, existiert noch immer ein Bedarf an alternativen Clusteringverfahren, die verschiedene Schwächen von häufig verwendeten Methoden berücksichtigen: Verglichen mit dem klassischen k -means-Algorithmus haben die Algorithmen ST-DBSCAN und ST-OPTICS zum Beispiel den Vorteil, dass die Anzahl an Clustern im Vorfeld nicht festgelegt werden muss, was zu einer höheren Flexibilität führt. Ein Nachteil dieser dichte-basierten Verfahren ist hingegen, dass diese nicht direkt bedeutungsvolle Clusterzentren finden, die ein bestimmtes Cluster repräsentieren. Dies ist insbesondere für eine gezielte Datenreduktion problematisch. Außerdem müssen für diese Algorithmen im Vorfeld Kontrollparameter festgelegt werden, um eine möglichst optimale Zuordnung von Objekten hinsichtlich verschiedener Kriterien zu ermöglichen. Im Falle von unterschiedlich festgelegten Kontrollparametern, die z.B. bei der Durchführung mehrerer Clusteranalysen für unterschiedliche Szenarien notwendig sein könnten, sind die resultierenden Clusteringergebnisse jedoch nicht mehr miteinander vergleichbar.

In diesem Kapitel wird auf Basis obiger Überlegungen ein Clusteringansatz vorgestellt (CorClustST), der die besondere Struktur von räumlich-zeitlichen Datensätzen aufgreift und insbesondere einen sinnvollen Vergleich von Clusteringergebnissen für verschiedene Szenarien ermöglicht. Die Methode basiert dabei auf empirischen Kreuzkorrelationen zwischen den Beobachtungszeitreihen an verschiedenen Standorten. Die resultierenden Cluster sind leicht zu interpretieren - unabhängig von der Anzahl an Zeitpunkten sowie von der zugrundeliegenden beobachteten Zielgröße. Da die Beobachtungen oftmals nur bis zu einer gewissen räumlichen Distanz korrelieren, ist es außerdem nicht notwendig, Korrelationen zwischen allen möglichen räumlichen Standortpaaren zu berechnen. Um die Berechnungszeit und den Speicherbedarf des Algorithmus zu minimieren, werden deshalb nur Paare von Standorten betrachtet, die innerhalb einer gewissen räumlichen Distanz zueinander liegen. Der vorgeschlagene Algorithmus basiert auf einigen grundlegenden Ideen der Verfahren ST-DBSCAN und ST-OPTICS, modifiziert diese jedoch insofern, dass eine gute Vergleichbarkeit von Clusteringergebnissen für verschiedene Szenarien oder Teilregionen ermöglicht wird.

Die verschiedenen Schritte der Clusteringstrategie werden detailliert in Abschnitt 3.2 beschrieben. Daraufhin wird das Verfahren in Abschnitt 3.3 für eine Clusteranalyse für deterministische Windkraftvorhersagefehler in Europa verwendet. Die ersten drei Abschnitte dieses Kapitels entsprechen zu großen Teilen den Ausführungen in der Masterarbeit von Hüscher (2016, S. 13-18, S. 34-42). Da aus rechnerischen Gründen und aufgrund von Speicherplatzbeschränkungen eine Clusteranalyse für sehr große Datensätze problematisch sein kann, wird in Abschnitt 3.4 außerdem eine Erweiterung vorgestellt, mit der Clusteringergebnisse aus verschiedenen Teilregionen kombiniert werden können. Der abschließende Abschnitt 3.5 vergleicht das vorgestellte Clusteringverfahren mit anderen Clusteringmethoden und ordnet die gewonnenen Erkenntnisse kritisch ein. Die Methodik und die Ergebnisse dieses Kapitels sind auch in Hüscher et al. (2018) enthalten.

3.2 Beschreibung des Clusteringverfahrens

3.2.1 Definitionen

Die Beschreibung des Clusteringverfahrens basiert auf einer Menge von N Standorten $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$. Weiterhin sei angenommen, dass Beobachtungen $z(s, t)$ für alle räumlichen Standorte $s \in \mathcal{S}$ für T verschiedene Zeitpunkte $t = 1, \dots, T$ vorliegen. Bevor die Methodik im Detail beschrieben wird, werden zunächst einige notwendige Definitionen eingeführt. Diese basieren in weiten Teilen auf den Ausführungen von Birant und Kut (2007). Definition 3.1 beschreibt zunächst die generelle Idee einer räumlichen Clusteranalyse. Standorte die bei der Clusteranalyse keinem Cluster zugeordnet werden, werden im Folgenden in der Menge Ξ zusammengefasst.

Definition 3.1 *Der Prozess, bei dem die Menge von Standorten \mathcal{S} in N_C disjunkte Cluster $\mathcal{C}_i \subseteq \mathcal{S}$, ($i = 1, 2, \dots, N_C$) mit der Eigenschaft $\cup_{i=1}^{N_C} \mathcal{C}_i \cup \Xi = \mathcal{S}$ aufgeteilt wird, wird als **Clustering** bezeichnet.*

Für das Clusteringverfahren werden außerdem verschiedene Nachbarschaftsumgebungen definiert, die eine Einteilung von räumlichen Standorten in verschiedene Cluster ermöglichen. In einem räumlich-zeitlichen Kontext können sich verschiedene Standorte zum einen in räumlicher Hinsicht und zum anderen hinsichtlich der beobachteten Werte ähneln. Definition 3.2 führt zunächst die sogenannte räumliche ϵ -Nachbarschaft eines Standortes ein. Diese enthält alle räumlichen Nachbarn, die innerhalb einer festgelegten Distanz ϵ zum jeweiligen Standort liegen.

Definition 3.2 Die **räumliche ϵ -Nachbarschaft** eines Standortes s sei definiert als

$$\text{SpatNeigh}(s) = \{q \in \mathcal{S} \mid \text{dist}(s, q) \leq \epsilon\},$$

wobei $\text{dist}(s, q)$ die räumliche Distanz zwischen den Standorten s und q bezeichne. Standorte, die innerhalb der räumlichen ϵ -Nachbarschaft liegen, werden als **räumliche Nachbarn** von s bezeichnet.

Da nicht nur die räumliche Distanz verschiedener Standorte von Interesse ist, sondern insbesondere Gemeinsamkeiten der Standorte bezüglich der beobachteten Werte, werden außerdem empirische Kreuzkorrelationen zwischen den Beobachtungszeitreihen an verschiedenen Standorten bei der Clusteranalyse berücksichtigt. In Definition 3.3 wird die räumlich-zeitliche ϱ -Nachbarschaft eines Standortes unter Verwendung der empirischen Korrelation nach Pearson definiert (siehe Pearson, 1895).

Definition 3.3 Die **räumlich-zeitliche ϱ -Nachbarschaft** eines Standortes s sei definiert als

$$\text{SpatTempNeigh}(s) = \{q \in \text{SpatNeigh}(s) \mid \text{cor}(s, q) > \varrho\},$$

wobei

$$\text{cor}(s, q) := \frac{\sum_{t=1}^T (z(s, t) - \bar{z}_s)(z(q, t) - \bar{z}_q)}{\sqrt{\sum_{t=1}^T (z(s, t) - \bar{z}_s)^2} \sqrt{\sum_{t=1}^T (z(q, t) - \bar{z}_q)^2}}$$

die empirische Kreuzkorrelation (zum Lag 0) nach Pearson zwischen den Beobachtungszeitreihen an den Standorten s und q bezeichne. Weiterhin sei \bar{z}_s das arithmetische Mittel der beobachteten Werte am Standort s . Standorte, die innerhalb dieser Nachbarschaft liegen, werden als **räumlich-zeitliche Nachbarn** von s bezeichnet.

Mit den oben eingeführten Definitionen können die verschiedenen Schritte des Clusteringverfahrens im Folgenden detailliert beschrieben werden.

3.2.2 Beschreibung der Clusteringstrategie

Der erste Schritt des Clusteringverfahrens ist zugleich der rechenintensivste Teil. Zunächst werden für alle Standorte $s \in \mathcal{S}$ empirische Kreuzkorrelationen zwischen der Beobachtungszeitreihe am Standort s und den Beobachtungszeitreihen aller räumlicher Nachbarn berechnet. Die Distanz ϵ für die ϵ -Nachbarschaft sollte dabei so gewählt werden, dass einerseits eine akzeptable Rechenzeit resultiert und andererseits ausreichend viele räumliche Nachbarn berücksichtigt werden. Da alle paarweisen Kreuzkorrelationen unabhängig voneinander berechnet werden können, ist eine parallele Implementierung dieses Schrittes einfach in gängigen Programmiersprachen umsetzbar. Dies macht den Algorithmus daher auch für große räumlich-zeitliche Datensätze performant.

Um die weiteren Schritte des Clusteringverfahrens durchführen zu können, wird zunächst für alle Standorte im Datensatz die Anzahl an räumlichen Nachbarn bestimmt, zu denen eine empirische Kreuzkorrelation größer als ein im Vorfeld gewählter Grenzwert ϱ vorliegt. Diese Standorte bezeichnen wir als räumlich-zeitliche Nachbarn (siehe Definition 3.3). Beispielsweise kann ein Grenzwert von $\varrho = 0.9$ gewählt werden, um mit Hilfe der Clusteranalyse Regionen mit stark ausgeprägten Kreuzkorrelationen identifizieren. Ein Grenzwert von $\varrho = 0.7$ kann hingegen festgelegt werden, um auch moderatere Kreuzkorrelationen bei der Analyse zu berücksichtigen. Die im Datensatz enthaltenen Standorte werden daraufhin in absteigender Ordnung bezüglich der Anzahl ihrer räumlich-zeitlichen Nachbarn angeordnet. Diese Anordnung wird in einer Liste $O = \{o_1, \dots, o_N\}$ gespeichert.

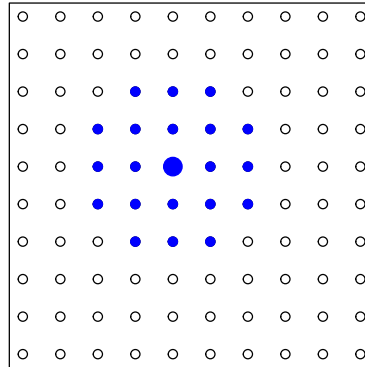
Schritt 1 Für alle Standorte $s \in \mathcal{S}$: Berechne Pearsons empirische Kreuzkorrelation $cor(s, q)$ zu allen räumlichen Nachbarn $q \in \text{SpatNeigh}(s)$ für eine festgelegte Distanz ϵ . Ordne die Standorte absteigend hinsichtlich ihrer Anzahl an räumlich-zeitlichen Nachbarn $|\text{SpatTempNeigh}(s)|$ für einen festgelegten Wert ϱ an, und speichere diese Anordnung in einer Menge $O = \{o_1, \dots, o_N\}$.

Im Folgenden werden die Standorte hinsichtlich ihrer Anzahl an räumlich-zeitlichen Nachbarn in Cluster aufgeteilt. Die verschiedenen Schritte des Clusteringansatzes werden mit Hilfe eines fortlaufenden Beispiels (ein Gitter mit 10×10 Standorten) visualisiert. Zunächst wird der räumliche Standort mit der größten Anzahl an räumlich-zeitlichen Nachbarn als das Clusterzentrum des Clusters \mathcal{C}_1 gewählt und alle räumlich-zeitlichen Nachbarn werden ebenfalls diesem Cluster zugeordnet.

Schritt 2 *Wähle den Standort*

$$o_1 = \operatorname{argmax}_{s \in \mathcal{S}} |\operatorname{SpatTempNeigh}(s)| \in O$$

für einen festgelegten Wert ϱ als Clusterzentrum des Clusters \mathcal{C}_1 . Alle räumlich-zeitlichen Nachbarn $q \in \operatorname{SpatTempNeigh}(o_1)$ werden dem Cluster \mathcal{C}_1 zugeordnet.



Für alle weiteren geordneten Standorte $o_i \in O$ wird daraufhin folgende Clusteringprozedur iterativ angewandt: Falls o_i nicht zu einem Cluster gehört, und falls mehr als 50% der räumlich-zeitlichen Nachbarn von o_i ebenfalls keinem Cluster angehören, wird o_i als Clusterzentrum eines Clusters gewählt. Durch die Restriktion, dass mehr als die Hälfte der Nachbarn von o_i keinem Cluster angehören sollen, wird garantiert, dass Standorte die Nahe an der Grenze eines bereits existierenden Clusters liegen, nicht als Zentrum eines neuen Clusters gewählt werden. Ein deutlich kleinerer Wert als 50% würde zu neuen Clusterzentren nahe der Grenze von existierenden Clustern führen. In diesem Falle wäre

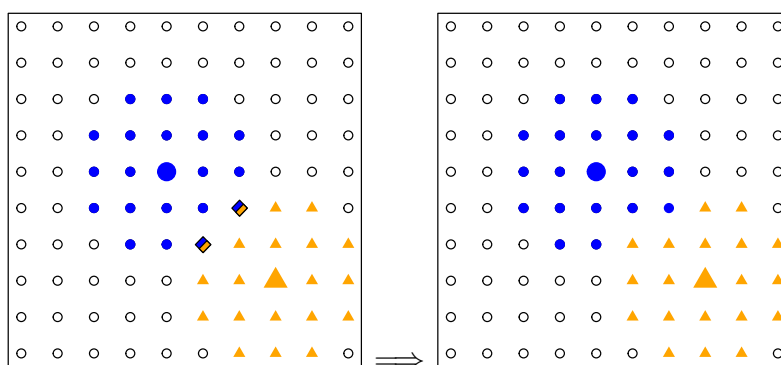
es schwierig, Standorte die den räumlich-zeitlichen ρ -Umgebungen beider Clusterzentren angehören, zu einem der Cluster zuzuordnen. Auf der anderen Seite würde ein deutlich höherer Wert als 50% zu vielen nicht-überlappenden Clustern führen. Dadurch könnten viele sehr kleine Cluster in den Regionen zwischen zwei größeren Clustern entstehen. Aus diesen Gründen erscheint ein Wert von 50% als guter Kompromiss, um eine sinnvolle Aufteilung der Standorte in verschiedene Cluster zu ermöglichen.

Die aktuelle Clusterbezeichnung wird daraufhin um den Wert eins erhöht, und o_i wird zum entsprechenden Cluster zugeordnet. Für alle räumlich-zeitlichen Nachbarn von o_i wird daraufhin überprüft, ob sie bereits einem Cluster angehören. Falls ein Nachbar noch nicht zu einem Cluster gehört, wird er direkt dem aktuellen Cluster zugeordnet. Falls ein Nachbar hingegen bereits zu einem Cluster gehört, wird überprüft ob die Kreuzkorrelation zwischen der Beobachtungszeitreihe des Nachbarn und der des Standortes o_i größer ist als die Kreuzkorrelation zu seinem bisherigen Clusterzentrum. In diesem Falle wird der entsprechende Nachbar ebenfalls dem aktuellen Cluster zugeordnet. Daraufhin wird der nächste Punkt o_{i+1} verarbeitet und das Verfahren wird iterativ wiederholt. Diese Prozedur wird in den Clusteringschritten 3 und 4 zusammengefasst.

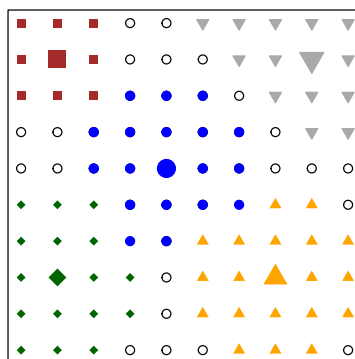
Schritt 3 Wähle den nächsten Standort

$$o_2 = \operatorname{argmax}_{s \in \{S \setminus o_1\}} |\operatorname{SpatTempNeigh}(s)| \in O$$

(für den mehr als 50% der räumlich-zeitlichen Nachbarn noch keinem Cluster angehören) als Clusterzentrum des Clusters \mathcal{C}_2 . Für alle Standorte $q \in \operatorname{SpatTempNeigh}(o_2)$: Ordne q dem Cluster \mathcal{C}_2 zu, falls q entweder noch keinem Cluster angehört, oder falls die Kreuzkorrelation zwischen den Beobachtungszeitreihen der Standorte q und o_2 größer ist als die Kreuzkorrelation zwischen der Zeitreihe des Standortes q und der Zeitreihe seines bisherigen Clusterzentrums.

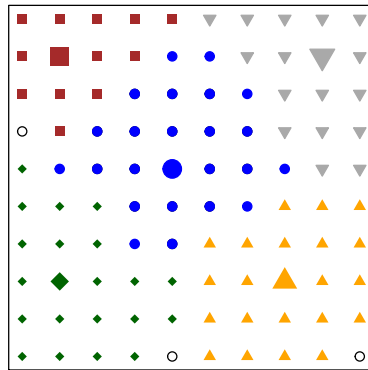


Schritt 4 Wiederhole Schritt 3 für o_3, \dots, o_N bis alle Standorte in O verarbeitet wurden.



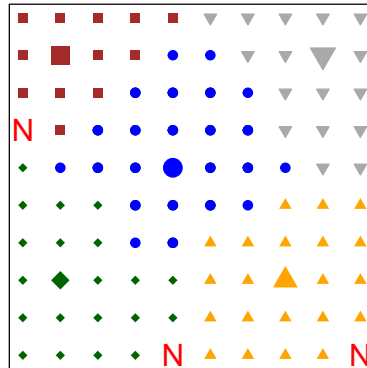
Aufgrund der Struktur des Clusteringverfahrens werden möglicherweise bestimmte Standorte nicht zu einem Cluster zugeordnet, obwohl sie räumlich-zeitliche Nachbarn besitzen. Dies ist beispielsweise dann möglich, wenn mehr als die Hälfte der räumlich-zeitlichen Nachbarn eines Standortes bereits einem Cluster angehören und der Standort deshalb nicht als Clusterzentrum in Betracht kommt, und wenn der Standort außerdem in keiner ϱ -Nachbarschaft eines anderen Clusterzentrums liegt. Um eine einfachere Interpretation der Ergebnisse zu ermöglichen, sollen jedoch nur die Standorte keinem Cluster angehören, die zu keinem anderen räumlichen Standort eine Kreuzkorrelation größer als der festgelegte Wert ϱ aufweisen. Deshalb werden die oben angesprochenen Standorte, die räumlich-zeitliche Nachbarn besitzen, im nächsten Schritt dem Cluster zugeordnet zu dem der ähnlichste räumlich-zeitliche Nachbar (hinsichtlich der Kreuzkorrelation) angehört.

Schritt 5 Standorte, die räumlich-zeitliche Nachbarn besitzen, aber nach Schritt 4 noch keinem Cluster angehören, werden dem Cluster zugeordnet, dem der ähnlichste räumlich-zeitliche Nachbar (hinsichtlich der Kreuzkorrelation) angehört.



Für höhere festgelegte Werte von ϱ ist es wahrscheinlich, dass einige Standorte keine räumlich-zeitlichen Nachbarn besitzen. Diese Standorte werden im abschließenden Schritt 6 des Clusteringverfahrens als nicht zugeordnete Standorte gekennzeichnet.

Schritt 6 Standorte, die nach Schritt 5 noch keinem Cluster \mathcal{C}_i (mit $i = 1, \dots, N_C$) angehören, werden als nicht zugeordnete Standorte gekennzeichnet.



Der vorgestellte Clusteringansatz erlaubt es, verschiedene Clusteringergebnisse zu vergleichen und hinsichtlich der im Vorfeld festgelegten Korrelationsstärke zu interpretieren. Die Methode ermöglicht dadurch insbesondere, Regionen mit starken und wenig ausgeprägten räumlichen Korrelationsstrukturen zu identifizieren. Der Algorithmus kann daher ein hilfreiches Verfahren darstellen, um wertvolle Informationen über die Abhängigkeitsstrukturen in räumlich-zeitlichen Datensätzen zu gewinnen. Im folgenden Abschnitt verwenden wir das Verfahren dazu, um den Einfluss des Vorhersagehorizonts auf die räumliche Abhängigkeitsstruktur von deterministischen Windkraftvorhersagefehlern in Europa zu untersuchen.

3.3 Anwendung: Identifizierung von räumlichen Abhängigkeitsstrukturen bei deterministischen Windkraftvorhersagefehlern

Wie bereits in Kapitel 2 beschrieben, sind zuverlässige Windgeschwindigkeits- und Windkraftvorhersagen von hoher Bedeutung für den gesamten Energiesektor. In diesem und im folgenden Kapitel liegt der Fokus zunächst auf den Vorhersagefehlern von deterministischen Windkraftvorhersagen. Bezüglich der Vorhersagefehler ist wohlbekannt, dass die Qualität der Vorhersagen für längere Vorhersagehorizonte abnimmt. Zum Beispiel zeigt Lange (2005) für eine Testanlage am Standort Hilkenbrook (Deutschland), dass die Güte von Windgeschwindigkeits- und Windkraftprognosen für längere Vorhersagehorizonte bis zu 48 Stunden stetig abnimmt. In der empirischen Literatur wurde bislang allerdings nur wenig diskutiert, dass auch die räumliche Korrelation von Vorhersagefehlern mit zunehmendem Vorhersagehorizont ansteigt. Auf diese Tatsache wird beispielsweise in der Studie von Focken et al. (2002) hingewiesen. Laut den Autoren resultieren die höheren räumlichen Korrelationen aus wachsenden systematischen Fehlern für längere Vorhersagehorizonte. In einer Fallstudie für Westdänemark zeigen Tastu et al. (2011) zudem, dass Windkraftvorhersagefehler für sehr kurze Vorhersagehorizonte nur eine geringe räumliche Korrelation aufweisen. Es ist daher anzunehmen, dass räumliche Abhängigkeitsstrukturen bei Vorhersagefehlern für längere Vorhersagehorizonte insgesamt deutlich ausgeprägter sind als für kurze Vorhersagehorizonte. Insbesondere in Regionen, in denen eine große Anzahl an Windkraftanlagen stationiert ist, stellen gemeinsam auftretende Vorhersagefehler (und damit hohe aggregierte Vorhersagefehler) ein potenzielles Risiko dar.

Um diesen Aspekt genauer zu untersuchen, verwenden wir den vorgestellten Clusteringalgorithmus im Folgenden um erste Erkenntnisse über die räumliche Abhängigkeitsstruktur von Windkraftvorhersagefehlern in Europa zu gewinnen. Die Analyse basiert auf deterministischen Vorhersagefehlern für das europäische Festland für den Zeitraum von April

2010 bis Februar 2016. Die Datenaufbereitung erfolgte in der Masterarbeit von Hüsck (2016), in der die Datenstruktur und die Verarbeitung bereits genauer beschrieben wurden. Das berücksichtigte Studiengebiet (visualisiert in Abbildung 3.1) umfasst 32 europäische Länder, in denen zumindest eine geringe Anzahl an installierter Windleistung vorhanden ist (ausgewählt basierend auf den Statistiken in EWEA, 2016). Die Windkraftvorhersagen resultieren dabei aus deterministischen Windgeschwindigkeitsvorhersagen (in m/s) in einer Höhe von 100 Metern des Europäischen Zentrums für mittelfristige

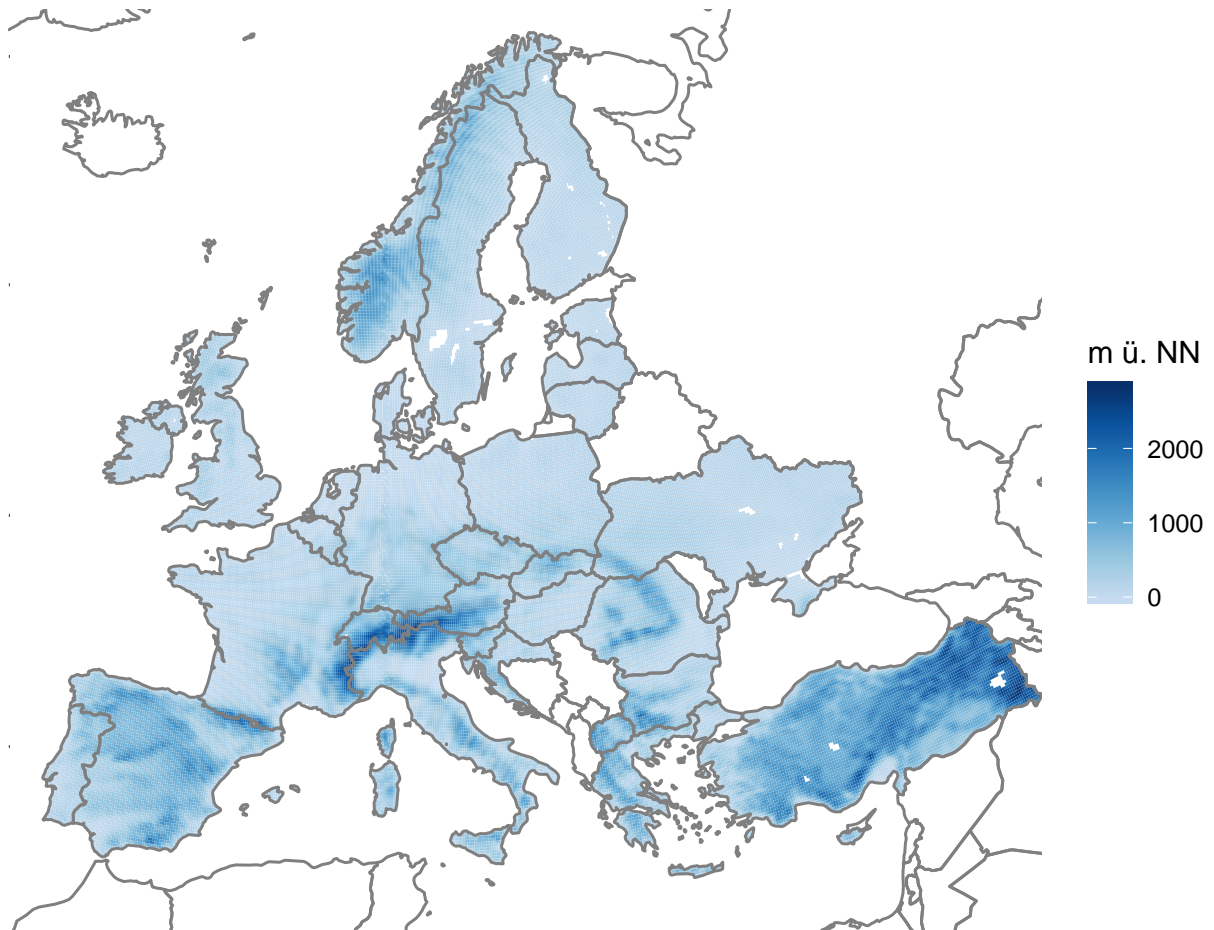


Abbildung 3.1: Studienregion bei der durchgeführten Analyse mit Angaben zur zugrundeliegenden Modellorographie (Höhe über NN in Metern eines bestimmten Gitterpunktes)

Wettervorhersage (siehe ECMWF, 2019), die auf einem räumlichen Gitter zur Verfügung gestellt werden. Für die Umwandlung der Windgeschwindigkeiten in Windleistung an einem spezifischen Gitterpunkt wird eine von Späth et al. (2015) vorgeschlagene regionale Leistungskurve verwendet. Solche Leistungskurven transformieren die Windgeschwindigkeit in (normalisierte) Windleistung im Intervall $[0, 1]$, wobei ein Wert von 1 bedeutet, dass eine Windkraftanlage Leistung in Höhe der Nennleistung erbringen kann. In der regionalen Leistungskurve wird laut Späth et al. (2015) zusätzlich berücksichtigt, dass bei der Produktion von Windstrom Verluste auftreten, die sich beispielsweise durch Übertragungsverluste oder Stillstandszeiten ergeben. Bei der Analyse werden insgesamt sechs unterschiedliche Vorhersagehorizonte betrachtet (12, 24, 36, 48, 60 und 72 Stunden). Die Vorhersagen werden vom ECMWF für alle Vorhersagehorizonte zweimal am Tag ausgegeben. Die Vorhersagefehler für einen bestimmten Vorhersagehorizont τ werden für alle Standorte durch Subtrahieren der tatsächlichen Windleistung (verfügbar in Form einer 0-Stunden Vorhersage) von der im Voraus erstellten τ -Stunden Prognose berechnet.

Die erstellten Vorhersagefehlerdatensätze umfassen für jeden Vorhersagehorizont insgesamt Daten für $N = 49968$ Standorte (Gitterpunkte) und für $T = 4316$ äquidistante Zeitpunkte. Bei einer große Anzahl an Datenpunkten kann eine Clusteranalyse mit herkömmlichen Algorithmen, wie z.B. hierarchischen Clusterverfahren, schnell sehr rechenintensiv werden. Oftmals muss für die Clusteranalyse eine Distanzmatrix mit $\frac{N(N-1)}{2}$ Einträgen gespeichert werden, was für gewöhnliche Computer eine Herausforderung darstellen kann. Der vorgeschlagene Clusteringansatz hingegen vermeidet diese Problematik indem nur Paare von Standorten betrachtet werden, die innerhalb einer gewissen räumlichen Distanz zueinander liegen. Wir führen das Clustering im Folgenden für zwei verschiedene Korrelationsstärken durch. In einem ersten Szenario wählen wir $\rho = 0.7$ (moderate Korrelation) und in einem zweiten Szenario setzen wir $\rho = 0.9$ (hohe Korrelation). Für diese relativen hohen Werte für ρ ist es unwahrscheinlich, dass Standortpaare in einer sehr großen Distanz existieren, deren Beobachtungszeitreihen eine Kreuzkorrelation größer als der festgesetzte Grenzwert aufweisen. Eine eigenständige Analyse zeigt zudem, dass Vor-

hersagefehler für eine Distanz von mehr als 600 Kilometern nahezu unkorreliert sind. Aus diesem Grund setzen wir den Wert von ϵ für die Clusteranalyse auf 600 Kilometer. Die Clusteranalyse wird mit dieser Einstellung für alle sechs Vorhersagehorizonte (12 Stunden bis 72 Stunden) für beide Korrelationsstärken durchgeführt.

Um einen ersten Eindruck über die unterschiedlichen Clusteringergebnisse zu erhalten, vergleichen wir in Tabelle 3.1 zunächst die Anzahl der Cluster für die verschiedenen Vorhersagehorizonte und Korrelationsstärken. Der Vergleich zeigt, dass der Vorhersagehorizont einen deutlichen Einfluss auf die räumliche Korrelation von Windkraftvorhersagefehlern hat. Im moderaten Korrelationsszenario ($\rho = 0.7$), verringert sich die Anzahl der Cluster von 3189 für die 12-Stunden-Vorhersagefehler auf 1259 für die 72-Stunden-Vorhersagefehler. Im Falle $\rho = 0.9$ ergeben sich 10269 verschiedene Cluster für die 12-Stunden-Vorhersagefehler und 5597 Cluster für die 72-Stunden-Vorhersagefehler.

Tabelle 3.1: Resultierende Anzahl an Clustern für verschiedene Vorhersagehorizonte und Korrelationsstärken

Korrelationsstärke	Vorhersagehorizont					
	12 Std.	24 Std.	36 Std.	48 Std.	60 Std.	72 Std.
$\rho = 0.7$	3189	2633	2178	1847	1503	1259
$\rho = 0.9$	10269	8900	7879	7078	6247	5597

Um verschiedene Regionen in Europa hinsichtlich der räumlichen Korrelation von Vorhersagefehlern zu untersuchen, fokussieren wir uns auf einen Vergleich der Ergebnisse für die 24-Stunden- und die 72-Stunden-Vorhersagefehler. In den Abbildungen 3.2 und 3.3 werden zunächst die resultierenden Cluster für die moderate Korrelationsstärke ($\rho = 0.7$) visualisiert. Beim Vergleich der beiden Clusteringergebnisse fällt auf, dass die Cluster für die 72-Stunden-Vorhersagefehler im gesamten europäischen Festland deutlich größer ausfallen als für die 24-Stunden-Vorhersagefehler. Außerdem werden große Unterschiede zwischen den einzelnen europäischen Regionen deutlich. Die größten Cluster findet man in flachen Gebieten in Nordeuropa, Deutschland, Frankreich und Osteuropa. In diesen Gebieten erhöht sich die Größe der einzelnen Cluster beim Vergleich der beiden Vorhersa-

gehizonte sehr deutlich. Dieses Ergebnis hat eine hohe praktische Relevanz, da in diesen Regionen eine verhältnismäßig große Anzahl an installierter Windleistung vorhanden ist. Eine hohe Korrelation von Windkraftvorhersagefehlern für nahe gelegene Standorte kann unter Umständen mit hohen kumulierten Vorhersagefehlern einhergehen. Die kleinsten Cluster treten in Gebirgsregionen wie den Alpen, den Pyrenäen und den Karpaten auf sowie in den höher liegenden Gegenden in Südeuropa und Norwegen.



Abbildung 3.2: Clusteringergebnis für 24-Stunden-Vorhersagefehler ($\rho = 0.7$)

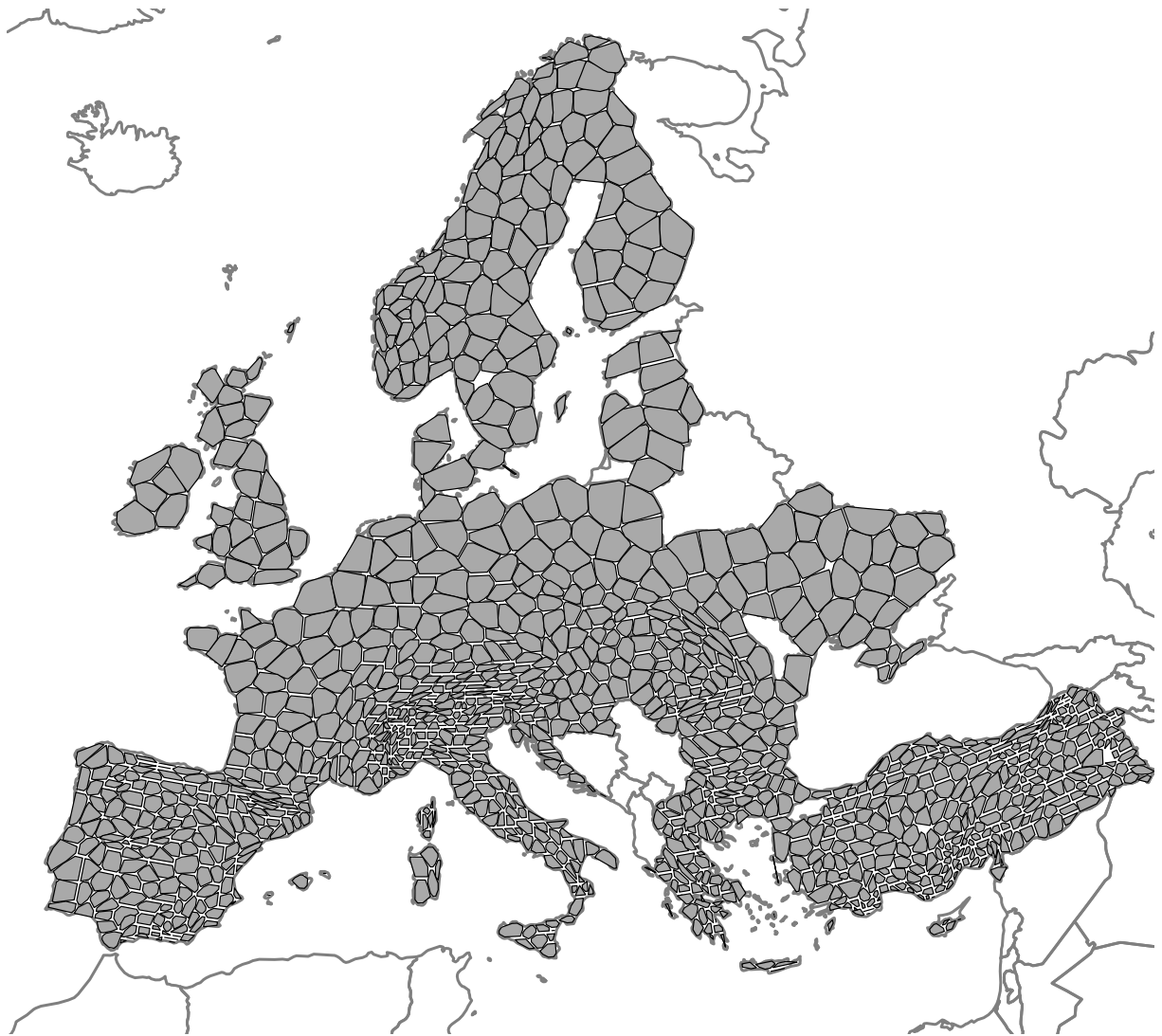


Abbildung 3.3: Clusteringergebnis für 72-Stunden-Vorhersagefehler ($\varrho = 0.7$)

Ähnliche Ergebnisse ergibt die Clusteranalyse für die hohe Korrelationsstärke ($\rho = 0.9$). Die Abbildungen 3.4 und 3.5 zeigen ein ähnliches Bild wie für das moderate Korrelationszenario, wobei die Cluster im Vergleich insgesamt deutlich kleiner ausfallen. In den flachen Gebieten in Nordeuropa mit einer hohen Anzahl an installierter Windleistung ist die Korrelationsstruktur deutlich ausgeprägter als in Gegenden in Südeuropa oder in den Gebirgsregionen.



Abbildung 3.4: Clusteringergebnis für 24-Stunden-Vorhersagefehler ($\rho = 0.9$)



Abbildung 3.5: Clusteringergebnis für 72-Stunden-Vorhersagefehler ($\varrho = 0.9$)

Dies zeigt auch eine Analyse der Standorte, die bei der Clusteranalyse zu keinem Cluster zugeordnet werden weil sie keine räumlich-zeitlichen Nachbarn nach Definition 3.3 besitzen. Diese Gebiete zeichnen sich durch eine weniger stark ausgeprägte Korrelationsstruktur im Vergleich zu anderen Regionen aus. Abbildung 3.6 vergleicht die Standorte, die bei der Clusteranalyse für die hohe Korrelationsstärke ($\varrho = 0.9$) keinem Cluster zugeordnet werden für die 24-Stunden- und die 72-Stunden-Vorhersagefehler. Für die 24-

Stunden-Vorhersagefehler werden hauptsächlich Standorte im Hochgebirge, wie z.B. in den französischen Alpen, den Apenninen in Italien oder im Taurusgebirge in der Türkei, keinem Cluster zugeordnet. Außerdem gibt es mehrere Standorte nahe der Küsten, die ebenfalls keinem Cluster zugeordnet werden. Die Standorte in den Küstengebieten sind bei Betrachtung der Grafik für die 72-Stunden-Vorhersagefehler überwiegend nicht mehr vorhanden. Allerdings existiert immer noch eine relativ große Anzahl an nicht zugeordneten Standorten in den französischen Alpen sowie in den Gebirgsregionen in Griechenland und in der Türkei. Die Anzahl der nicht zugeordneten Standorte reduziert sich insgesamt im Vergleich deutlich für längere Vorhersagehorizonte. Dies resultiert aus einer generell stärker ausgeprägten räumlichen Korrelationsstruktur.

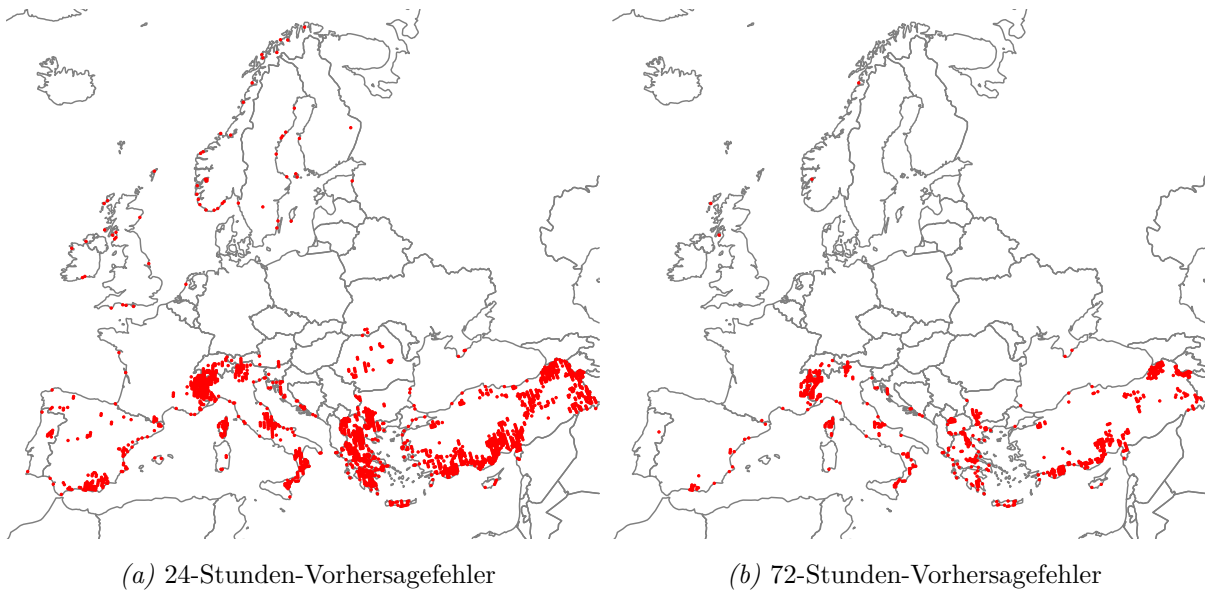


Abbildung 3.6: Standorte die beim Clustering für die 24- und 72-Stunden-Vorhersagefehler keinem Cluster zugeordnet werden ($\varrho = 0.9$)

3.4 Erweiterung: Kombination von Clustering-Ergebnissen aus mehreren Teilregionen

Die Anwendung im vorherigen Abschnitt demonstriert, dass der vorgestellte Algorithmus eine effiziente Methode zur Identifizierung von Abhängigkeitsstrukturen in räumlich-zeitlichen Datensätzen darstellt. Allerdings können bei der Anwendung des Verfahrens Speicherplatzbeschränkungen bei noch größeren Datenmengen zum Problem werden, insbesondere wenn sich die Anzahl der räumlichen Standorte deutlich erhöht. Im Beispiel der Clusteranalyse für Windkraftvorhersagefehler könnte dies zum Beispiel dann zum Problem werden, wenn große Länder wie Russland ebenfalls bei der Analyse berücksichtigt werden. Des Weiteren erhöht sich auch die räumliche Auflösung meteorologischer Vorhersagen stetig, damit eine noch höhere Vorhersagequalität erreicht werden kann. Beispielfähig hat das Europäische Zentrum für mittelfristige Wettervorhersage im März 2016 den horizontalen Abstand des verwendeten Gitters von 16 auf 9 Kilometer reduziert (siehe ECMWF, 2016). Um Probleme zu vermeiden die aufgrund von Speicherplatzbeschränkungen entstehen könnten, wird in diesem Abschnitt eine Erweiterung des Clusteringverfahrens präsentiert. Die Erweiterung basiert auf der Idee, dass das komplette Studiengebiet in mehrere Teilregionen unterteilt werden kann, für die dann separate Clusteranalysen durchgeführt werden. Da die Beobachtungen an sehr weit entfernten Standorten oftmals nicht stark korrelieren, sollten Cluster die weit entfernt von einer anderen Teilregion liegen von der Unterteilung der Studienregion unberührt bleiben. Als Beispiel führen wir eine Clusteranalyse für 12-Stunden-Windkraftvorhersagefehler für Standorte in Österreich und in der Schweiz durch. Dabei wählen wir ein moderates Korrelationsszenario ($\rho = 0.7$). Der Wert ϵ für die räumliche ϵ -Nachbarschaft wird erneut auf 600 Kilometer festgelegt. Die Clusteranalyse wird einmal für beide Regionen gemeinsam durchgeführt und einmal für beide Länder separat. Die resultierenden Cluster werden in Abbildung 3.7 verglichen. Die Cluster der separaten Clusteranalysen, die im Vergleich zum gemeinsamen Clustering unterschiedlich sind, sind dabei in rot markiert.

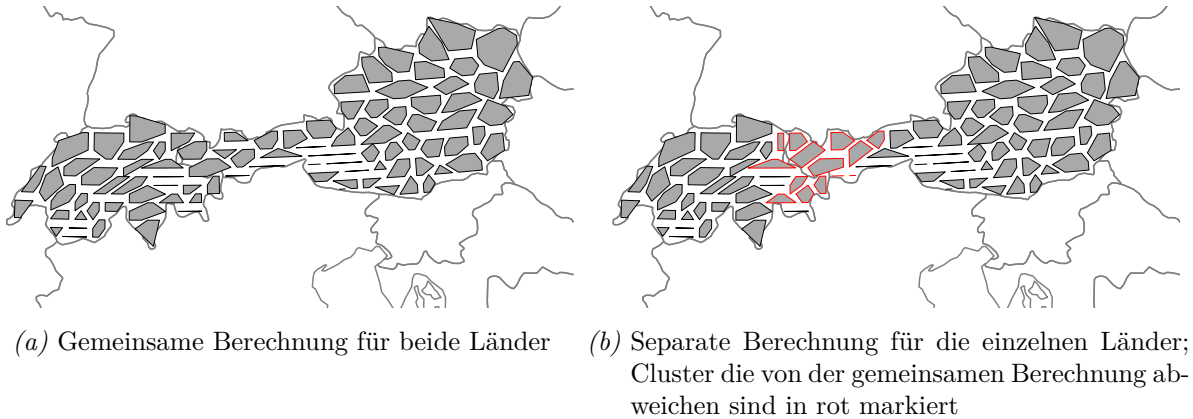


Abbildung 3.7: Clusteranalyse für 12-Stunden-Vorhersagefehler an Standorten in Österreich und der Schweiz ($\rho = 0.7$)

Der Vergleich zeigt, dass sich die Clusteringstruktur nur in der Region nahe der Grenze der beiden Länder unterscheidet. Um die Clustering-Ergebnisse der beiden Teilregionen zu kombinieren, müssen daher nur die Standorte erneut verarbeitet werden, die zu einem Cluster nahe der Grenze gehören. Aus diesem Grund definieren wir für die Erweiterung des Clusteringverfahrens sogenannte innere Grenzcluster (Cluster bei denen sich mindestens ein Standort innerhalb einer Distanz γ_1 zum nächstliegenden Standort der anderen Teilregion befindet) und äußere Grenzcluster (Cluster die nicht zu den inneren Grenzclustern gehören und bei denen sich mindestens ein Standort innerhalb einer Distanz γ_2 zum nächstliegenden Standort der anderen Teilregion befindet). Dadurch können die Standorte ermittelt werden, für die ein zusätzliches Clustering durchgeführt wird. Während die Standorte in den äußeren Grenzclustern ihre Clusterzugehörigkeiten aus den separaten Clusteranalysen behalten, werden die Bezeichnungen der Standorte die zu den inneren Grenzclustern gehören gelöscht. Für das Beispiel von separaten Clusteranalysen für 12-Stunden-Vorhersagefehler in Österreich und der Schweiz zeigt Abbildung 3.8 die resultierenden inneren und äußeren Grenzcluster für zwei unterschiedliche Werte für γ_1 und γ_2 . Für die weitere Analyse wird im nächsten Schritt ein neuer Datensatz erstellt, der die Beobachtungszeitreihen und Clusterbezeichnungen für die Standorte enthält, die zu

den inneren oder äußeren Grenzclustern gehören. Für diese Standorte wird daraufhin mit Hilfe des Verfahrens aus Abschnitt 3.2 eine zusätzliche Clusteranalyse durchgeführt.

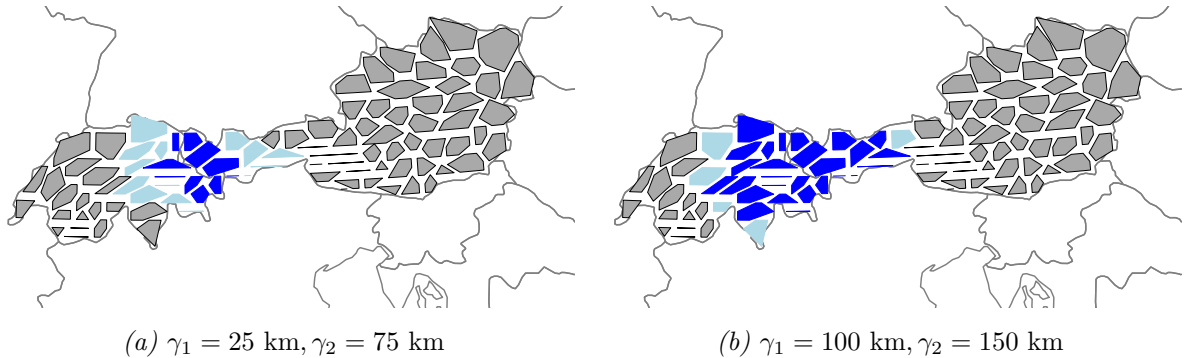


Abbildung 3.8: Grenzcluster, die sich bei separaten Clusteranalysen ($\varrho = 0.7$) für 12-Stunden-Vorhersagefehler in Österreich und der Schweiz für zwei verschiedene Einstellungen der Parameter γ_1 und γ_2 ergeben. Innere Grenzcluster sind in dunkelblau, äußere Grenzcluster in hellblau markiert.

Um die Erweiterung des Clusteringverfahrens im Allgemeinen zu beschreiben, nehmen wir an, dass die Menge an räumlichen Standorten \mathcal{S} in R verschiedene Teilregionen $\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(R)}$ aufgeteilt wird. Weiterhin sei die Anzahl an verfügbaren Zeitpunkten für alle Teilregionen $\mathcal{S}^{(r)}$ ($r = 1, \dots, R$) identisch, während sich die Anzahl der Standorte $N^{(r)}$ auch unterscheiden kann. Für jede Teilregion wird zunächst eine separate Clusteranalyse mit dem in Abschnitt 3.2 vorgeschlagenen Algorithmus durchgeführt. Dies resultiert in einer Anzahl von $N_{\mathcal{C}}^{(r)}$ Clustern für die entsprechende Teilregion $\mathcal{S}^{(r)}$. Bevor die Cluster der unterschiedlichen Teilregionen kombiniert werden können, müssen zunächst die Clusterbezeichnungen angepasst werden, um Cluster mit gleichen Bezeichnungen aus den verschiedenen Teilregionen zu vermeiden. Die Cluster aus der ersten Teilregion $\mathcal{S}^{(1)}$ behalten ihre Bezeichnungen. Die Cluster aus der zweiten Teilregion $\mathcal{S}^{(2)}$ erhalten die Bezeichnungen $C_{N_{\mathcal{C}}^{(1)}+1}, \dots, C_{N_{\mathcal{C}}^{(1)}+N_{\mathcal{C}}^{(2)}}$ und die Cluster aus $\mathcal{S}^{(3)}$ die Bezeichnungen $C_{N_{\mathcal{C}}^{(1)}+N_{\mathcal{C}}^{(2)}+1}, \dots, C_{N_{\mathcal{C}}^{(1)}+N_{\mathcal{C}}^{(2)}+N_{\mathcal{C}}^{(3)}}$. Dieses Verfahren wird für alle weiteren Teilregionen entsprechend angepasst. Daraufhin werden die folgenden Schritte durchgeführt, um ein Clusteringergebnis für die gesamte Studienregion zu erhalten:

Schritt 1 Bestimme die Menge an inneren Grenzstandorten

$$s_{inner}^{(r)} = \{s \in \mathcal{S}^{(r)} \mid \min_{q \in \{\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(R)}\} \setminus \mathcal{S}^{(r)}} \text{dist}(s, q) \leq \gamma_1\}$$

für jede Teilregion $\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(R)}$. Speichere die gesamte Liste an inneren Grenzstandorten in einer Menge $s_{inner} = \cup_{r \in \{1, \dots, R\}} s_{inner}^{(r)}$ und die entsprechenden Clusterbezeichnungen in einer Menge \mathcal{C}_{inner} .

Schritt 2 Bestimme die Menge an äußeren Grenzstandorten

$$s_{outer}^{(r)} = \{s \in \mathcal{S}^{(r)} \mid \gamma_1 < \min_{q \in \{\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(R)}\} \setminus \mathcal{S}^{(r)}} \text{dist}(s, q) \leq \gamma_2\}$$

für jede Teilregion $\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(R)}$. Speichere die gesamte Liste an äußeren Grenzstandorten in einer Menge $s_{outer} = \cup_{r \in \{1, \dots, R\}} s_{outer}^{(r)}$ und die entsprechenden Clusterbezeichnungen in einer Menge \mathcal{C}_{outer} .

Schritt 3 Erstelle eine neue Teilregion \mathcal{S}_{edge} mit den Standorten, die zu einem Cluster aus \mathcal{C}_{inner} oder \mathcal{C}_{outer} gehören, sowie den Standorten in der Menge s_{inner} , die bei der separaten Clusteranalyse keinem Cluster zugeordnet wurden.

Schritt 4 Entferne alle Clusterbezeichnungen der Standorte aus \mathcal{S}_{edge} , die zu einem Cluster aus \mathcal{C}_{inner} gehören, und wende die Clusteringstrategie aus Abschnitt 3.2 für die neue Teilregion an. Dies führt zu neuen Clustern, die durch neue Clusterbezeichnungen gekennzeichnet werden.

Schritt 5 Kombiniere die Clusteringergebnisse für die Grenzregionen aus Schritt 4 mit den Clusterbezeichnungen der restlichen Standorte. Dies führt zu einem finalen Clusteringergebnis.

Wir testen den oben vorgeschlagenen Ansatz erneut für eine Clusteranalyse für 12-Stunden-Vorhersagefehler in Österreich und der Schweiz. Die Clusteringergebnisse, die zunächst für beide Länder separat ermittelt wurden, werden nun mit Hilfe der beschriebenen Erweiterung kombiniert. Die Kombination der Ergebnisse wird einmal mit relativ kleinen Distanzen von $\gamma_1 = 25$ km und $\gamma_2 = 75$ km durchgeführt, und einmal mit größeren Distanzen von $\gamma_1 = 100$ km und $\gamma_2 = 150$ km. Die Ergebnisse sind in Abbildung 3.9 visualisiert.

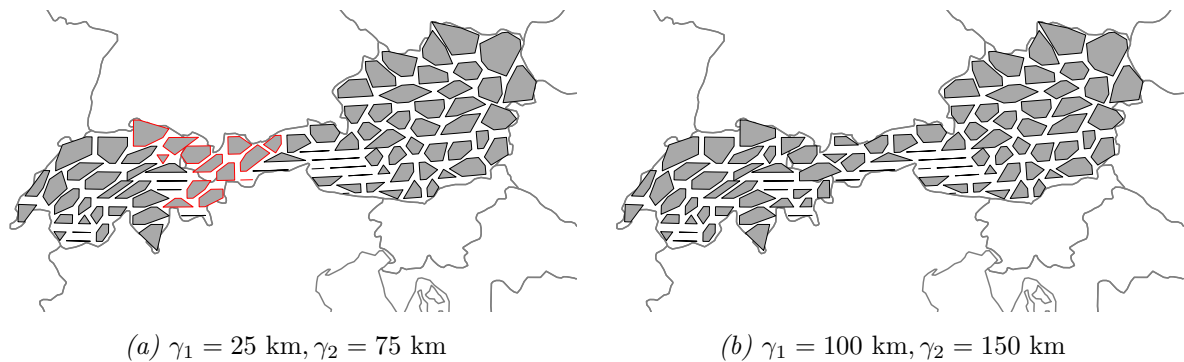


Abbildung 3.9: Clusteranalyse für 12-Stunden-Vorhersagefehler an Standorten in Österreich und der Schweiz ($\varrho = 0.7$) mit der Erweiterung aus Abschnitt 3.4. Cluster die vom Ergebnis der gemeinsamen Clusteranalyse für beide Länder abweichen sind in rot markiert.

Die Cluster die für kleine Werte der Parameter γ_1 und γ_2 resultieren, unterscheiden sich im Grenzbereich der beiden Länder immer noch im Vergleich zur gemeinsamen Clusteranalyse. Wenn die Grenzwerte für die Distanzen jedoch erhöht werden, ähneln die resultierenden Cluster immer mehr denen der gemeinsamen Clusteranalyse. Für Werte von $\gamma_1 = 100$ km und $\gamma_2 = 150$ km ist die Clusteringstruktur bereits identisch zur Clusteringstruktur, die sich bei der gemeinsamen Berechnung ergibt. Generell heben beide Berechnungsarten aber unabhängig der Wahl von γ_1 und γ_2 die gleichen Regionen hervor, in denen hohe oder weniger stark ausgeprägte Korrelationsstrukturen vorhanden sind. Die Distanzen γ_1 und γ_2 sollten im Vorfeld mit Bezug zum Korrelationsgrenzwert ϱ festgelegt werden. Da ein hoher Wert von ϱ zu insgesamt kleineren Clustern und deshalb zu einer eher ähnlicheren Struktur in den Grenzregionen für beide Berechnungsarten führt,

können in diesem Fall insgesamt kleinere Werte für die beiden Parameter gewählt werden als bei niedrigen Werten von ϱ . Mit Hilfe der vorgestellten Erweiterung kann die Clusteranalyse parallel für verschiedene Teilregionen durchgeführt werden, was insbesondere für sehr große räumlich-zeitliche Datensätze von Vorteil ist. Insbesondere bei möglichen Speicherplatzproblemen bietet die vorgeschlagene Erweiterung daher eine Möglichkeit, eine Clusteranalyse auch für eine große Studienregion mit einer hohen Anzahl an räumlichen Standorten durchzuführen.

3.5 Diskussion und Fazit

Das in diesem Kapitel vorgestellte Verfahren zur Clusteranalyse von räumlich-zeitlichen Datensätzen ermöglicht eine effiziente Identifizierung von räumlichen Abhängigkeitsstrukturen. Im Vergleich zu herkömmlichen Clusteringverfahren besitzt die Strategie einige nennenswerte Vorteile: Anders als zum Beispiel beim populären k -means-Algorithmus muss die Anzahl der Cluster im Vorfeld nicht spezifiziert werden. Beim k -means-Algorithmus wird eine möglichst optimale Anzahl an Clustern meist mit heuristischen Kriterien wie dem Elbow-Kriterium (siehe Thorndike, 1953) bestimmt, weshalb ein Clustering mehrmals für verschiedene Werte von k durchgeführt werden muss. Dies kann speziell bei sehr großen Datenmengen problematisch werden. Der vorgeschlagene Algorithmus CorClustST besitzt insgesamt eine geringe Komplexität und benötigt insbesondere im Vergleich zu hierarchischen Clusteringverfahren nur wenig Speicherplatz. Das Clustering ist vor allem dann sehr effizient, wenn ein geringer Wert für den Parameter ϵ gewählt wird. In diesem Fall wird nur eine geringe Anzahl an räumlichen Nachbarn für einen spezifischen Standort berücksichtigt, weshalb insgesamt weniger Standortpaare verarbeitet werden. Die Komplexität des Verfahrens ist in diesem Fall vergleichbar mit der von anderen Algorithmen wie ST-DBSCAN, die ebenfalls speziell für räumlich-zeitliche Datensätze entwickelt wurden. Das vorgestellte Verfahren ist zudem leicht zu parallelisieren, da die Kreuzkorrelationen zwischen den Standortpaaren im ersten und rechenintensivsten Schritt der Methode unab-

hängig voneinander berechnet werden können. Zusammengefasst erlauben die relativ kurzen Berechnungszeiten des Algorithmus auch ein Clustering von großen räumlich-zeitlichen Datensätzen. Dies wird beispielsweise in Tabelle 3.2 demonstriert, in der die Laufzeiten des Algorithmus für eine Clusteranalyse von 24-Stunden-Windkraftvorhersagefehlern für verschiedene Werte von ϵ und für unterschiedliche Anzahlen an Standorten N verglichen werden. Der Grenzwert für die räumlich-zeitliche Nachbarschaft wird dabei auf $\varrho = 0.9$ festgelegt. Sogar für einen relativ hohen Wert von $\epsilon = 600$ Kilometern erhöht sich die Laufzeit für eine größere Anzahl an Standorten N nur moderat. Die Berechnungszeiten fallen außerdem noch geringer aus, wenn die Kreuzkorrelationen im ersten Schritt des Verfahrens parallel berechnet werden.

Tabelle 3.2: Laufzeit des Clusteringverfahrens CorClustST in Sekunden für verschiedene Werte von ϵ und N für die 24-Stunden-Vorhersagefehler ($\varrho = 0.9$)

	$N = 1000$	$N = 10000$	$N = 20000$	$N = 49968$
$\epsilon = 100$ km	66.69	875.53	1438.81	3684.51
$\epsilon = 300$ km	208.59	5709.20	8981.60	21422.19
$\epsilon = 600$ km	236.71	16094.08	29004.67	74527.64

Auch bezüglich der Interpretation der Ergebnisse bietet das Verfahren einige Vorteile: Zum einen werden bei der Clusteranalyse bedeutungsvolle Clusterzentren bestimmt, die die größte Anzahl an räumlich-zeitlichen Nachbarn in einer bestimmten Region besitzen. Diese Clusterzentren können daher als repräsentative Standorte für ein Cluster angesehen werden. Insbesondere für hohe Werte von ϱ können diese Standorte deshalb auch für eine gezielte Datenreduktion verwendet werden. Des Weiteren ist es in vielen Anwendungsfällen (wie z.B. bei der Analyse von Windkraftvorhersagefehlern) von besonderem Interesse, räumliche Abhängigkeitsstrukturen für verschiedene Szenarien (wie z.B. unterschiedliche Vorhersagehorizonte) oder verschiedene Regionen zu vergleichen. Hier grenzt sich der vorgeschlagene Algorithmus stark von anderen Clusteringverfahren ab. Anders als herkömmliche Algorithmen hat das Verfahren nicht das Ziel, ein möglichst optimales Clusteringergebnis hinsichtlich bestimmter Kriterien zu erzielen. Um ein solches Ziel zu

erreichen, müssen bei anderen Verfahren verschiedene Kontrollparameter im Vorfeld festgelegt werden, was eine Vergleichbarkeit für verschiedene Szenarien (bei unterschiedlich gewählten Kontrollparametern) schwierig macht. Der vorgeschlagene Algorithmus erlaubt hingegen eine einfachere Interpretation von verschiedenen Clusteringergebnissen durch eine Fixierung des Grenzwertes ϱ für die räumlich-zeitliche Nachbarschaftsumgebung. Das Verfahren ist daher als eine effiziente deskriptive Methode zu betrachten, mit der räumliche Korrelationsstrukturen sinnvoll für verschiedene Regionen und Szenarien miteinander verglichen werden können.

Mit Hilfe der vorgestellten Clusteringstrategie haben wir in Abschnitt 3.3 erste Erkenntnisse über die räumliche Abhängigkeitsstruktur von deterministischen Windkraftvorhersagefehlern in Europa gewonnen. Wie im Vorfeld erwartet, sind räumliche Korrelationen bei Vorhersagefehlern insbesondere für längere Vorhersagehorizonte ausgeprägt. Weiterhin unterscheiden sich die Korrelationsstrukturen auch sehr stark hinsichtlich geographischer Gegebenheiten. Insbesondere in flachen Gebieten mit einer hohen Anzahl an installierter Windleistung liegen ausgeprägte räumliche Korrelationen vor. Diese Erkenntnis ist vor allem deshalb interessant, weil bei ausgeprägten räumlichen Abhängigkeitsstrukturen zu erwarten ist, dass sich Vorhersagefehler an naheliegenden Standorten kumulieren und zu einem hohen aggregierten Vorhersagefehler in einem gesamten räumlichen Gebiet führen. Die korrelationsbasierte Analyse aus diesem Kapitel lässt allerdings noch keine detaillierten Rückschlüsse zu, ob auch sehr große Vorhersagefehler häufig gemeinsam in einem räumlichen Kontext auftreten. Im folgenden Kapitel wird deshalb ein copulabasierter Ansatz vorgestellt, der weiterführende Erkenntnisse über die Abhängigkeitsstruktur von deterministischen Windkraftvorhersagefehlern ermöglicht.

4 Identifizierung von räumlichen Abhängigkeiten mit Copulas und Generalisierten Additiven Modellen

4.1 Motivation

Das im vorherigen Kapitel beschriebene Clusteringverfahren ermöglicht eine erste, effiziente Identifizierung von räumlichen Abhängigkeitsstrukturen bei räumlich-zeitlichen Datensätzen. Für die Analyse von räumlichen Abhängigkeiten bei Windkraftvorhersagefehlern konnten so bereits einige interessante Erkenntnisse gewonnen werden, die in diesem Kapitel noch genauer untersucht werden sollen. Besonders relevant für den Energiesektor sind dabei Vorhersagefehler, die in Gebieten mit einer hohen Anzahl an installierter Windleistung gleichzeitig auftreten und deshalb zu einem hohen aggregierten Vorhersagefehler für die gesamte Region führen. Insbesondere ist hier interessant, ob speziell sehr große Vorhersagefehler häufig gemeinsam in einem räumlichen Kontext auftreten, was ein potenzielles Risiko für Übertragungsnetzbetreiber oder Energiehändler bedeuten würde. Korrelationsbasierte Verfahren wie die Methode aus Kapitel 3 liefern für diese Fragestellung keine explizite Antwort, weshalb alternative Verfahren für eine detailliertere Analyse der Abhängigkeitsstruktur benötigt werden.

Die Analyse und Modellierung von räumlichen Abhängigkeitsstrukturen ist eigentlich eine traditionelle Aufgabe der Geostatistik. Ein klassischer Ansatz für eine Untersuchung von räumlichen Abhängigkeiten ist zum Beispiel eine sogenannte Semivariogramm-Analyse (für Details wird auf die Ausführungen von Schabenberger und Gotway (2004, S. 133-214) verwiesen). Für eine solche Analyse wird jedoch oft unterstellt, dass die Abhängigkeit der Beobachtungen an verschiedenen Standorten nur von der räumlichen Distanz der Standorte beeinflusst wird (siehe Schabenberger und Gotway, 2004, S. 27). Vor allem bei meteorologischen Anwendungen ist diese Annahme meist nicht erfüllt, da andere Einflussfaktoren wie z.B. geographische Gegebenheiten einen nicht zu vernachlässigenden Einfluss auf die Abhängigkeitsstruktur haben. Folgt man den Ausführungen von Gräler (2014), liegt ein weiterer Nachteil populärer geostatistischer Verfahren auf der Hand: Oftmals unterstellen die verwendeten Verfahren implizit eine Normalverteilungsannahme an die Daten, die jedoch in vielen Anwendungsfällen verletzt ist. Aus diesem Grund wurden in der empirischen Literatur in den letzten Jahren vermehrt copulabasierte Ansätze vorgeschlagen, da diese eine flexiblere Modellierung von räumlichen Abhängigkeiten ermöglichen. Beispielsweise schlägt Gräler (2014) die sogenannte räumliche Vine-Copula für die Modellierung und Interpolation von räumlichen Zufallsfeldern vor, und benutzt die Methode für die Interpolation von Notfallszenariodaten. In einer angewandten Arbeit verwenden Durocher et al. (2016) das räumliche Copula-Modell von Bárdossy (2006) und Bárdossy und Li (2008) für Hochwasservorhersagen. Krupskii und Genton (2017) benutzen Faktor-Copula-Modelle zur Modellierung von räumlich-zeitlichen Abhängigkeiten und für die Vorhersage von Windgeschwindigkeiten in den Niederlanden. Der Fokus dieser Arbeiten liegt jedoch eher auf räumlichen Interpolationen und Vorhersagen als auf der Identifizierung von räumlichen Abhängigkeitsstrukturen unter Berücksichtigung von unterschiedlichen geographischen Gegebenheiten.

In diesem Kapitel stellen wir eine flexible Methode vor, die speziell für die Identifizierung und den Vergleich von räumlichen Abhängigkeitsstrukturen für unterschiedliche Szenarien (wie zum Beispiel unterschiedliche Vorhersagehorizonte bei Windkraftvorhersagen) und

verschiedene Regionen geeignet ist. Das Verfahren basiert auf der grundlegenden Idee von Vatter und Chavez-Demoulin (2015), die die Parameter einer geeigneten Copula-Familie (oder ein in Relation stehendes Abhängigkeitsmaß) mit Hilfe eines Generalisierten Additiven Modells unter Berücksichtigung von verschiedenen Einflussfaktoren modellieren. Diese Idee haben wir speziell für die Identifizierung von räumlichen Abhängigkeitsstrukturen aufgegriffen, wobei wir uns wie im vorherigen Kapitel erneut auf die Analyse von Abhängigkeitsstrukturen für unterschiedliche Standortpaare fokussieren und dabei auch geographische Gegebenheiten dieser Paare berücksichtigen. Mit Hilfe des vorgeschlagenen Ansatzes können räumliche Abhängigkeitsstrukturen detaillierter analysiert werden.

Das Kapitel ist wie folgt strukturiert: Abschnitt 4.2 führt zunächst die grundlegenden Konzepte für die Modellierung von Abhängigkeitsstrukturen mit Copulas ein. Daraufhin wird die vorgeschlagene Methodik für die Identifizierung von räumlichen Abhängigkeiten mit Copulas und Generalisierten Additiven Modellen detailliert erläutert. In Abschnitt 4.3 verwenden wir das Verfahren, um genauere Erkenntnisse über die räumliche Abhängigkeitsstruktur von deterministischen Windleistungsvorhersagefehlern für unterschiedliche geographische Gegebenheiten zu gewinnen. Die Vor- und Nachteile der Methodik werden in Abschnitt 4.4 ausführlich diskutiert. Das Kapitel schließt mit einer kritischen Einordnung der Ergebnisse.

4.2 Methodik

Wie im vorherigen Kapitel betrachten wir erneut den Fall, dass ein Datensatz mit Beobachtungswerten $z(s, t)$ (z.B. Windkraftvorhersagefehler) für eine Menge von N unterschiedlichen Standorten $\mathcal{S} = \{s_1, \dots, s_N\}$ sowie für T verschiedene Zeitpunkte vorliegt. Im Folgenden fokussieren wir uns auf die Analyse der Abhängigkeitsstruktur von Beobachtungen an verschiedenen Standortpaaren s_i und s_j aus der Menge \mathcal{S} . Im nächsten Unterabschnitt wird zunächst erläutert, wie diese Abhängigkeitsstrukturen für verschiedene Standortpaare mit Hilfe von Copulas modelliert werden können. Daraufhin stellen wir

eine Strategie vor, die es ermöglicht aus geschätzten Copulas gezielte Informationen über die räumliche Abhängigkeitsstruktur für unterschiedliche geographische Gegebenheiten zu extrahieren.

4.2.1 Hintergrund: Modellierung von Abhängigkeiten mit Copulas

Für eine detaillierte Beschreibung des Verfahrens führen wir in einem ersten Schritt stetige räumliche Zufallsvariablen $Z(s_1), \dots, Z(s_N)$ für die verschiedenen Standorte ein. In diesem Unterabschnitt liegt der Fokus zunächst auf der Modellierung von Abhängigkeiten zwischen den Beobachtungen an zwei festgelegten Standorten s_i und s_j , ohne dafür bereits geographische Eigenschaften wie z.B. die räumliche Distanz dieser Standorte zu berücksichtigen.

Für die weitere Erläuterung bezeichnen wir mit $F_{Z_{s_i}}$ und $F_{Z_{s_j}}$ die Verteilungen der Zufallsvariablen $Z(s_i)$ und $Z(s_j)$. Außerdem sei $F_{Z_{s_i}, Z_{s_j}}$ die gemeinsame Verteilung der beiden Zufallsvariablen. Nach dem Theorem von Sklar (1959) existiert für stetige Zufallsvariablen $Z(s_i)$ und $Z(s_j)$ eine eindeutige bivariate **Copula** $C : [0, 1]^2 \rightarrow [0, 1]$, welche die folgende Eigenschaft erfüllt:

$$F_{Z_{s_i}, Z_{s_j}}(\mathbf{y}) = C(F_{Z_{s_i}}(y_1), F_{Z_{s_j}}(y_2)), \quad \forall \mathbf{y} = [y_1, y_2]^\top \in \mathbb{R}^2. \quad (4.1)$$

Die gemeinsame Verteilung kann demnach durch die univariaten Randverteilungen und die entsprechende Copula C charakterisiert werden. Die Copula enthält dabei Informationen über die Abhängigkeitsstruktur der beiden Zufallsvariablen (siehe Joe, 2014, S. 7-9). Im Anwendungsfall ist man bei gegebenen Beobachtungsdaten (die als Realisationen von unabhängig identisch verteilten Zufallsvariablen aufgefasst werden können) für zwei verschiedene Standorte s_i und s_j natürlich daran interessiert, eine Copula zu finden, welche die Abhängigkeitsstruktur zwischen den beobachteten Werten bestmöglich beschreibt. Dabei existiert eine Vielzahl an verschiedenen bivariaten Copula-Familien, mit denen unterschiedlichste Abhängigkeitsstrukturen modelliert werden können. Eine Übersicht über

einige populäre bivariate Copula-Familien ist in Tabelle 4.1 zusammengestellt. Neben den Parameterräumen der verschiedenen Copula-Familien ist ebenso angegeben, ob die Copula-Familien die Modellierung von ausgeprägten Randabhängigkeiten erlauben. Die Modellierung von Randabhängigkeiten ist insbesondere dann relevant, wenn sehr große Beobachtungen häufig gemeinsam für verschiedene Merkmale auftreten. Im Bereich der Finanzökonomie wird beispielsweise häufig beobachtet, dass sehr hohe positive und negative Renditen oftmals gleichzeitig für verschiedene Aktien auftreten (siehe zum Beispiel Malevergne und Sornette, 2003). In einem räumlichen Kontext sind ausgeprägte Randabhängigkeiten ebenfalls von besonderem Interesse, wie zum Beispiel für eine Untersuchung ob sehr große Windleistungsvorhersagefehler häufig gemeinsam in einem räumlichen Kontext auftreten. Aus diesem Grund definieren wir den **unteren Randabhängigkeitskoeffizienten** zwischen den Zufallsvariablen $Z(s_i)$ und $Z(s_j)$ als

$$\lambda_L := \lim_{u \rightarrow 0^+} \mathbb{P}(Z(s_j) \leq F_{Z_{s_j}}^{-1}(u) \mid Z(s_i) \leq F_{Z_{s_i}}^{-1}(u)) \quad (4.2)$$

und den **oberen Randabhängigkeitskoeffizienten** als

$$\lambda_U := \lim_{u \rightarrow 1^-} \mathbb{P}(Z(s_j) > F_{Z_{s_j}}^{-1}(u) \mid Z(s_i) > F_{Z_{s_i}}^{-1}(u)). \quad (4.3)$$

Folgend den Ausführungen von Joe (2014, S. 62) können diese Koeffizienten auch direkt mit Hilfe der Copula C ausgedrückt werden:

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} \quad \text{und} \quad \lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}. \quad (4.4)$$

Unter den in Tabelle 4.1 angegebenen Copula-Familien, ist die **t-Copula-Familie** die einzige, die gleichzeitig die Modellierung von ausgeprägten oberen und unteren Randabhängigkeiten erlaubt. Die Copula-Dichte einer bivariaten **t-Copula** mit Korrelationsparameter

$\rho \in (-1, 1)$ und $\nu > 2$ Freiheitsgraden ist definiert als

$$c_{\rho, \nu}(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \frac{\Gamma((\nu+2)/2) \Gamma(\nu/2)}{\Gamma^2((\nu+1)/2)} \frac{\left(1 + \frac{y_1^2 + y_2^2 - 2\rho y_1 y_2}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2}}{\left(1 + \frac{y_1^2}{\nu}\right)^{-(\nu+1)/2} \left(1 + \frac{y_2^2}{\nu}\right)^{-(\nu+1)/2}} \quad (4.5)$$

mit $y_1 = \mathbf{t}_\nu^{-1}(u_1)$ und $y_2 = \mathbf{t}_\nu^{-1}(u_2)$, wobei $\mathbf{t}_\nu^{-1}(\cdot)$ die Quantilsfunktion einer univariaten t -Verteilung mit ν Freiheitsgraden bezeichne (siehe Joe, 2014, S. 181). Im Falle der t -Copula mit Parametern ρ und ν sind der untere und der obere Randabhängigkeitskoeffizient identisch (siehe Demarta und McNeil, 2005). Laut den Autoren ergibt sich der Randabhängigkeitskoeffizient bei einer t -Copula zu

$$\lambda = \lambda_L = \lambda_U = 2\mathbf{t}_{\nu+1}(-\sqrt{\nu+1}\sqrt{1-\rho}/\sqrt{1+\rho}), \quad (4.6)$$

wobei $\mathbf{t}_{\nu+1}(\cdot)$ die Verteilungsfunktion einer t -Verteilung mit $\nu+1$ Freiheitsgraden bezeichne. Wie in Tabelle 4.1 dokumentiert, weisen einige populäre Copulas wie die Gauß-Copula oder die Frank-Copula im Gegensatz zur t -Copula keine ausgeprägten Randabhängigkeiten auf. Die Familie der Gumbel-Copulas erlaubt hingegen die Modellierung von oberen Randabhängigkeiten, aber nicht die Modellierung von unteren Randabhängigkeiten. Je nach Struktur der vorliegenden Daten ist daher eine Copula-Familie zu wählen, mit der die Abhängigkeitsstruktur zwischen den Beobachtungen geeignet modelliert werden kann.

Tabelle 4.1: Übersicht populärer bivariater Copula-Familien mit wichtigen Eigenschaften (Quellen: Joe, 2014, S. 159-222; Schepsmeier et al., 2017)

Copula-Familie	Parameter	Randabhängigkeitskoeffizienten
Gauß-Copula	$\rho \in (-1, 1)$	$\lambda_L = \lambda_U = 0$
t -Copula	$\rho \in (-1, 1),$ $\nu > 2$	$\lambda_L = \lambda_U = 2\mathbf{t}_{\nu+1}(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}})$
Frank-Copula	$\theta \in \mathbb{R} \setminus \{0\}$	$\lambda_L = \lambda_U = 0$
Gumbel-Copula	$\theta \in [1, \infty)$	$\lambda_L = 0, \lambda_U = 2 - 2^{1/\theta}$
Clayton-Copula	$\theta \in [-1, \infty) \setminus \{0\}$	$\lambda_L = 2^{-1/\theta}, \lambda_U = 0$
Joe-Copula	$\theta \in [1, \infty)$	$\lambda_L = 0, \lambda_U = 2 - 2^{1/\theta}$
Unabh.-Copula	-	$\lambda_L = \lambda_U = 0$

Nach Bestimmung einer geeigneten Copula-Familie können im nächsten Schritt die Parameter der ausgewählten Copula für die gegebenen Daten geschätzt werden. Wir verwenden dazu den von Genest et al. (1995) vorgeschlagenen Maximum-Pseudo-Likelihood-Ansatz. Dieser hat den Vorteil, dass die univariaten Randverteilungen im Vorfeld nicht parametrisch spezifiziert werden müssen. Stattdessen basiert die Schätzmethode auf rangtransformierten Pseudo-Beobachtungen

$$u(s_i, t) = \frac{1}{T+1} \cdot \text{rank}(z(s_i, t)), \quad t = 1, \dots, T, \quad (4.7)$$

wobei $\text{rank}(z(s_i, t))$ den Rang der Beobachtung $z(s_i, t)$ unter allen realisierten Beobachtungen am Standort s_i bezeichne. Die Copula-Parameter, die wir in einem Parametervektor $\boldsymbol{\delta} = [\delta_1, \dots, \delta_Q]^\top$ zusammenfassen, werden daraufhin durch Maximierung der Pseudo-Log-Likelihood-Funktion

$$\ell(\boldsymbol{\delta}) = \sum_{t=1}^T \log [c_{\boldsymbol{\delta}}(u(s_i, t), u(s_j, t))] \quad (4.8)$$

geschätzt. In Gleichung (4.8) bezeichnet $c_{\boldsymbol{\delta}}$ die zugehörige Dichtefunktion der Copula. Der Maximum-Pseudo-Likelihood-Schätzer ist laut Genest et al. (1995) unter bestimmten Regularitätsbedingungen konsistent und asymptotisch normalverteilt. Die resultierenden Schätzwerte fassen wir im Folgenden als $\hat{\boldsymbol{\delta}}_{MPL} = [\hat{\delta}_{1,MPL}, \dots, \hat{\delta}_{Q,MPL}]^\top$ zusammen.

Hinsichtlich der geschätzten Parameter muss beachtet werden, dass für eine relativ geringe Anzahl an Beobachtungsdaten eine ausgeprägte Verzerrung in den Schätzwerten vorhanden ist. Dies wird beispielhaft mit Hilfe einer Simulationsstudie mit 10 000 Wiederholungen für die bivariate t -Copula deutlich. In jeder Wiederholung simulieren wir $T = 500$ Pseudo-Beobachtungen aus einer t -Copula mit Korrelationsparameter $\rho = 0.7$ und $\nu = 7$ Freiheitsgraden und schätzen die Parameter daraufhin mit der Maximum-Pseudo-Likelihood-Methode. Dazu verwenden wir die Funktionen `BiCopSim` und `BiCopEst` aus dem Paket `VineCopula` (siehe Schepsmeier et al., 2017) in R. Im Optimierungsprozess wird ein Wert von $\nu = 30$ als maximaler Wert für den Freiheitsgradparameter gesetzt,

da eine t-Copula für größere Werte von ν fast nicht mehr von einer Gauß-Copula zu unterscheiden ist (siehe Schepsmeier et al., 2017). Die simulierten Verteilungen von $\hat{\rho}_{MPL}$ und $\hat{\nu}_{MPL}$ werden in Abbildung 4.1 verglichen. Während die Verteilung von $\hat{\rho}_{MPL}$ symmetrisch um den wahren Wert $\rho = 0.7$ ist, ist die Verteilung von $\hat{\nu}_{MPL}$ rechtsschief und die Parameterschätzungen weisen eine relativ hohe Varianz auf. Für eine höhere Anzahl an verfügbaren Beobachtungen werden die Verzerrung und die Varianz der Parameterschätzungen kleiner. Dies wird in Abbildung 4.2 deutlich, in der simulierte Verteilungen von $\hat{\rho}_{MPL}$ und $\hat{\nu}_{MPL}$ für Schätzungen basierend auf $T = 10\,000$ Pseudo-Beobachtungen verglichen werden. Für eine große Anzahl an verfügbaren Daten ist auch die simulierte Verteilung von $\hat{\nu}_{MPL}$ symmetrisch um den wahren Wert $\nu = 7$ und ähnelt einer Normalverteilung. Für kleine Stichproben muss die Verzerrung und die hohe Varianz der

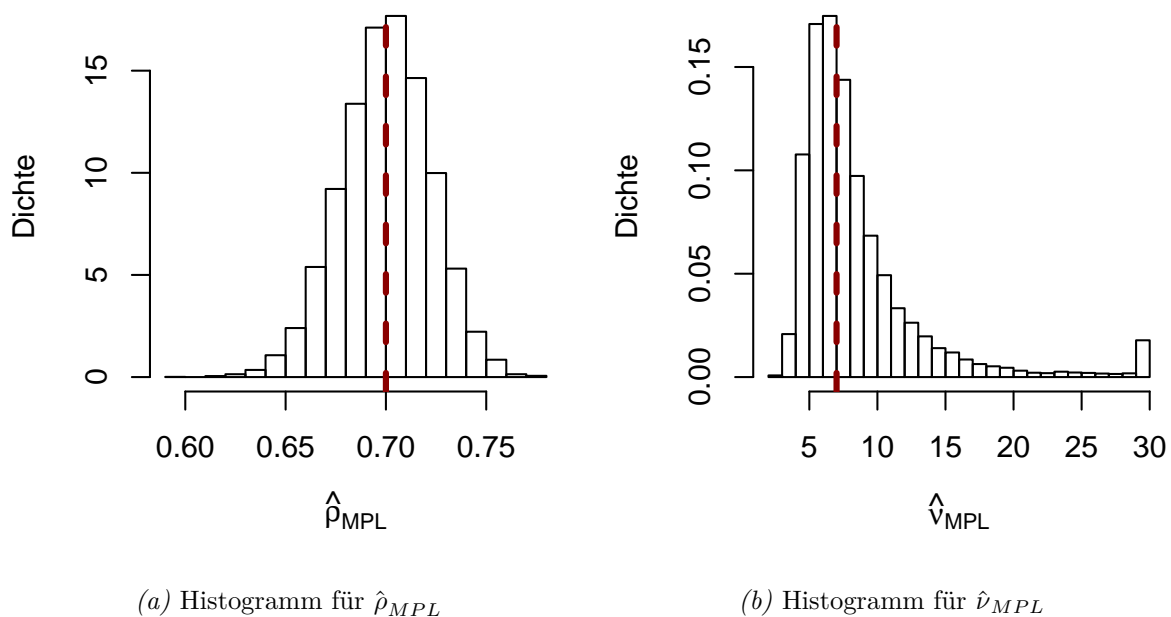


Abbildung 4.1: Schätzung der Parameter einer t-Copula (wahre Parameter: $\rho = 0.7$ und $\nu = 7$). In einer Monte-Carlo-Simulation mit 10 000 Wiederholungen werden in jeder Wiederholung zunächst $T = 500$ Pseudo-Beobachtungen aus der wahren Copula simuliert und die Parameter daraufhin mit Hilfe des Maximum-Pseudo-Likelihood-Ansatzes geschätzt.

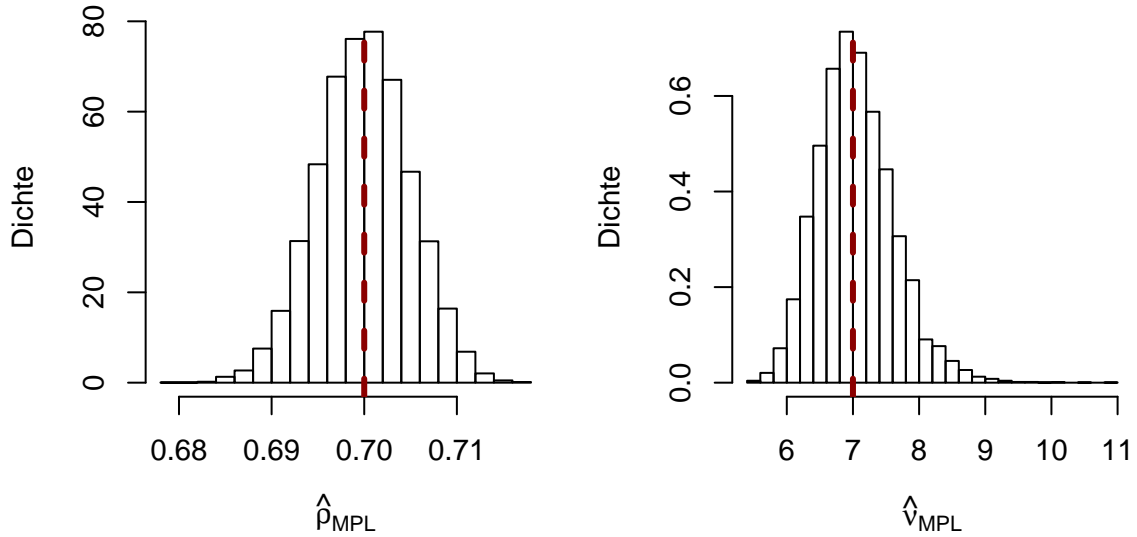
(a) Histogramm für $\hat{\rho}_{MPL}$ (b) Histogramm für $\hat{\nu}_{MPL}$

Abbildung 4.2: Schätzung der Parameter einer t-Copula (wahre Parameter: $\rho = 0.7$ und $\nu = 7$). In einer Monte-Carlo-Simulation mit 10 000 Wiederholungen werden in jeder Wiederholung zunächst $T = 10\,000$ Pseudo-Beobachtungen aus der wahren Copula simuliert und die Parameter daraufhin mit Hilfe des Maximum-Pseudo-Likelihood-Ansatzes geschätzt.

Parameterschätzer hingegen beachtet werden.

Für die Wahl einer geeigneten Copula-Familie verwenden wir das Informationskriterium nach Akaike (AIC). Für Pseudo-Beobachtungen $u(s_i, t)$ und $u(s_j, t)$ für Zeitpunkte $t = 1, \dots, T$ und eine (angepasste) Q -parametrische Copula mit Dichte $c_{\hat{\delta}}$ ist dieses definiert als

$$AIC = -2 \sum_{t=1}^T \log [c_{\hat{\delta}}(u(s_i, t), u(s_j, t))] + 2Q. \quad (4.9)$$

Nach der Schätzung der Copula-Parameter mit dem Maximum-Pseudo-Likelihood-Ansatz für verschiedene Copula-Familien wird abschließend die Copula-Familie ausgewählt, für die das AIC minimal ist (siehe Schepsmeier et al., 2017).

4.2.2 Beschreibung räumlicher Abhängigkeitsstrukturen mit Copulas und Generalisierten Additiven Modellen

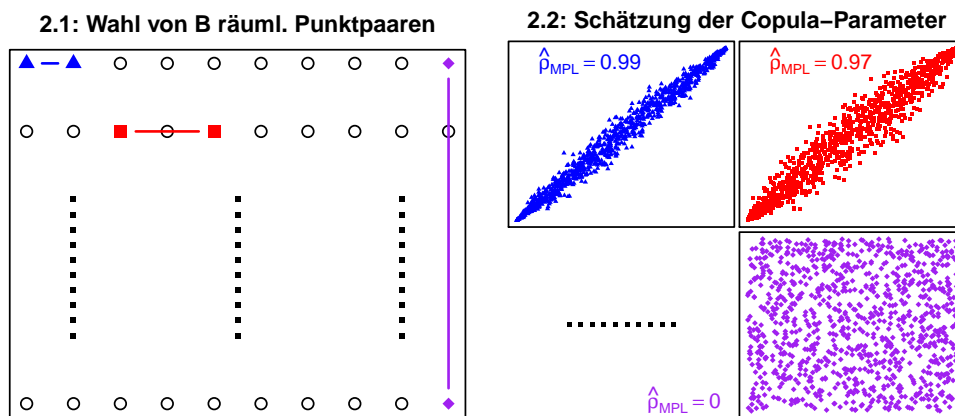
Aufbauend auf den Grundlagen aus dem vorherigen Abschnitt stellen wir nun eine intuitive Methode vor, mit der räumliche Abhängigkeitsstrukturen hinsichtlich unterschiedlicher geographischer Einflussfaktoren charakterisiert werden können. Dabei betrachten wir erneut verschiedene Paare von Standorten, für die wir jeweils separat eine geeignete Copula schätzen, welche die Abhängigkeitsstruktur zwischen den Beobachtungen bestmöglich charakterisiert. Daraufhin modellieren wir die geschätzten Copula-Parameter für diese Standortpaare mit Hilfe von Generalisierten Additiven Modellen, um systematische Unterschiede in den Parameterschätzungen bezüglich geographischer Einflussfaktoren zu identifizieren. Um bei einer solchen Analyse zu aussagekräftigen Ergebnissen zu gelangen, ist es notwendig, dass Standortpaare aus dem gesamten Studiengebiet bei der Analyse berücksichtigt werden. Die verschiedenen Schritte des Verfahrens werden im Folgenden detailliert beschrieben. Zum besseren Verständnis werden Schritt 2 und Schritt 3 der vorgeschlagenen Methode mit Hilfe eines fiktiven Beispiels mit simulierten Daten aus einparametrischen Gauß-Copulas visualisiert. Dabei variiert der Korrelationsparameter ρ der Gauß-Copulas nur bezüglich der räumlichen Distanz h zwischen zwei Standorten.

Schritt 1 Wähle für verschiedene Standortpaare $\{s_i, s_j\}$ mit Hilfe von Akaikes Informationskriterium (AIC) die Copula-Familie aus, welche die Abhängigkeitsstruktur zwischen den Beobachtungen an den Standorten bestmöglich charakterisiert.

Im ersten Schritt des Verfahrens wird zunächst überprüft, ob es eine geeignete Copula-Familie gibt, welche die bivariate Abhängigkeitsstruktur zwischen den Beobachtungen für verschiedene Paare $\{s_i, s_j\}$ geeignet modelliert. Dazu wird zunächst eine zufällige Auswahl an Standortpaaren (z.B. 1000 Paare) getroffen, für die jeweils die Parameter von verschiedenen Copula-Familien mit Hilfe des Maximum-Pseudo-Likelihood-Ansatzes geschätzt werden. Um lange Rechenzeiten zu vermeiden, reicht hier die Auswahl einer relativ kleinen Anzahl an Standortpaaren. Basierend auf dem Informationskriterium nach

Akaike (AIC) wird für jedes Standortpaar dann die Copula-Familie ausgewählt, für die das kleinste AIC vorliegt. Falls für die verschiedenen Paare sehr viele unterschiedliche Copula-Familien ausgewählt werden, sollte nach Schritt 1 eine detaillierte Analyse der selektierten Copulas für die verschiedenen Standortpaare erfolgen. Im Folgenden gehen wir jedoch davon aus, dass eine einzige Copula-Familie die Abhängigkeitsstruktur zwischen den Beobachtungen für verschiedene Standortpaare geeignet modelliert.

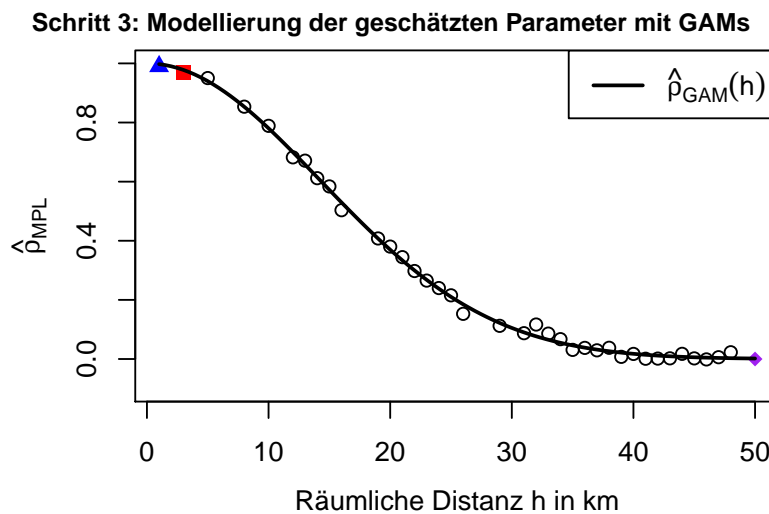
Schritt 2 Wähle B verschiedene Standortpaare aus, die unterschiedliche geographische Gegebenheiten geeignet repräsentieren. Schätze die Parameter der nach Schritt 1 ausgewählten Copula-Familie mit Hilfe des Maximum-Pseudo-Likelihood-Ansatzes für alle B Paare und erhalte geschätzte Parametervektoren $\hat{\delta}_{MPL}^{(1)}, \dots, \hat{\delta}_{MPL}^{(B)}$.



Nach der Wahl einer geeigneten Copula-Familie in Schritt 1 wird im nächsten Schritt eine größere Anzahl an B verschiedenen Standortpaaren ausgewählt. Dabei sollte die Auswahl der Standortpaare so vorgenommen werden, dass die gewählten Paare verschiedene geographische Gegebenheiten (erklärende Einflussgrößen) geeignet repräsentieren, die für die Analyse relevant sind. Falls beispielsweise nur die räumliche Distanz zwischen zwei Standorten für die Abhängigkeitsstruktur relevant erscheint, könnten z.B. B Zufallszahlen aus einer stetigen Gleichverteilung im Intervall $[h_{lower}, h_{upper}]$ gezogen werden. Dabei sei h_{lower} die minimale Distanz zwischen zwei Standorten im Datensatz und h_{upper} die maximale Distanz zwischen zwei Standorten, die für die Analyse von Interesse ist. Für alle resultie-

renden Zufallszahlen kann dann das Standortpaar ausgewählt werden, bei dem der Unterschied zwischen ihrer räumlichen Distanz und der jeweiligen Zufallszahl am geringsten ist. Für alle so ausgewählten Standortpaare werden daraufhin relevante räumliche Einflussfaktoren, wie z.B. die räumliche Distanz, in einem zusätzlichen Datensatz gespeichert. Außerdem werden die Parameter der in Schritt 1 ausgewählten Copula-Familie mit Hilfe des Maximum-Pseudo-Likelihood-Verfahrens für alle Standortpaare geschätzt. Dies führt zu B geschätzten Parametervektoren $\hat{\delta}_{MPL}^{(1)}, \dots, \hat{\delta}_{MPL}^{(B)}$. Im Falle von zusätzlichen Einflussfaktoren (wie z.B. orographischen Merkmalen) empfehlen wir die Verwendung einer Latin-Hypercube-Stichprobe (siehe Carnell, 2018) zur Wahl der verschiedenen Standortpaare. Dadurch wird garantiert, dass der gesamte Wertebereich verschiedener erklärender Variablen bei der Auswahl der Standortpaare gleichmäßig berücksichtigt wird. Damit eine ausreichende Repräsentation von verschiedenen geographischen Gegebenheiten gesichert ist, sollte außerdem eine hohe Anzahl an verschiedenen Standortpaaren (z.B. $B > 100\,000$) für die Copula-Schätzung gewählt werden.

Schritt 3 *Modelliere die geschätzten Parameter $\hat{\delta}_{q,MPL}^{(1)}, \dots, \hat{\delta}_{q,MPL}^{(B)}$ aus Schritt 2 in Abhängigkeit von verschiedenen geographischen Einflussvariablen G_1, \dots, G_K mit separaten Generalisierten Additiven Modellen für $q = 1, \dots, Q$.*



Mit den geschätzten Copula-Parametern und den geographischen Informationen zu den verschiedenen Standortpaaren kann nun der Einfluss von verschiedenen geographischen Einflussfaktoren auf die räumliche Abhängigkeitsstruktur analysiert werden. Basierend auf den Ideen von Vatter und Chavez-Demoulin (2015) schlagen wir vor, die verschiedenen geographischen Informationen als erklärende Variablen G_1, \dots, G_K zu verwenden und die geschätzten Copula-Parameter mit Generalisierten Additiven Modellen zu modellieren. Aufbauend auf Wood (2017) und Hastie und Tibshirani (1990) spezifizieren wir ein **Generalisiertes Additives Modell** (GAM) der Form

$$g_q \left\{ \mathbb{E}(\hat{\delta}_{q,MPL}^{(b)}) \right\} = \beta_{0,q} + \underbrace{\sum_{k=1}^K f_{k,q}(G_k^{(b)})}_{\text{Haupteffekte}} + \underbrace{f_{K+1,q}(G_1^{(b)}, G_2^{(b)}) + \dots}_{\text{Interaktionen}}, \quad b = 1, \dots, B \quad (4.10)$$

für jeden Copula-Parameter der in Schritt 1 gewählten Copula-Familie und somit für $q = 1, \dots, Q$. Aus Übersichtlichkeitsgründen wird hier nicht zwischen den Parameterschätzern als Zufallsvariablen und den Schätzwerten aus Schritt 2 unterschieden. Da die Copula-Parameterschätzer, wie in Abschnitt 4.2.1 diskutiert, unter Regularitätsbedingungen asymptotisch normalverteilt sind, wird bei der Modellierung für die erklärende Variable eine Normalverteilungsannahme unterstellt. Dies sollte zumindest für Datensätze mit einer hohen Anzahl an Beobachtungen (wie z.B. in der Anwendung in Abschnitt 4.3) eine valide Annahme darstellen. Die Funktionen $g_q : \Delta_q \rightarrow \mathbb{R}$ in Gleichung (4.10) bezeichnen im Vorfeld festgelegte Link-Funktionen, die die Parameterschätzwerte aus den Parameterräumen Δ_q in die reelle Zahlenebene transformieren. Die Funktionen $f_{k,q}(G_k^{(b)})$ werden als sogenannte glatte Funktionen der erklärenden Variablen gewählt. Beispielsweise können dazu Spline-Funktionen verwendet werden, die mit Hilfe von Basiserweiterungen der Form

$$f_{k,q}(G_k^{(b)}) = \sum_{l=1}^{L_k} \beta_{kl,q} \cdot \zeta_{kl,q}(G_k^{(b)}) \quad (4.11)$$

mit im Vorfeld festgelegten Basisfunktionen $\zeta_{kl,q}(G_k^{(b)})$ dargestellt werden. Ein Vorteil von

Generalisierten Additiven Modellen ist, dass glatte Interaktionen der erklärenden Variablen (wie z.B. $f_{K+1,q}(G_1^{(b)}, G_2^{(b)})$) ebenfalls berücksichtigt werden können (für Details wird auf die Ausführungen von Wood (2017, S. 227-237) verwiesen). Die Modellformel in Gleichung (4.10) enthält zum einen sogenannte Haupteffekte der erklärenden Variablen und berücksichtigt ebenfalls Interaktionen der erklärenden Variablen. Dies erlaubt es einerseits, den Einfluss von wichtigen Faktoren wie der räumlichen Distanz zweier Standorte auf die Parameterschätzwerte zu untersuchen und ermöglicht außerdem, bestimmte Interaktionseffekte zu analysieren. So kann z.B. untersucht werden, ob sich die räumliche Distanz von Standortpaaren für unterschiedliche orographische Gegebenheiten gleichermaßen auf die Abhängigkeitsstruktur auswirkt. Nach Ersetzung aller glatten Funktionen durch Basiserweiterungen wie in Gleichung (4.11) und nach Einführung von Bedingungen zur Identifizierbarkeit der Modellparameter entspricht das GAM einem klassischen Generalisierten Linearen Modell (GLM) mit einer relativ großen Modellmatrix (aufgrund der hohen Anzahl an verschiedenen Basisfunktionen). Der komplette Parametervektor β des GAM wird mit Hilfe einer iterativen, penalisierten Kleinste-Quadrate-Methode (PIRLS) geschätzt (siehe Wood, 2017, S. 251). Für methodische Details und die Wahl geeigneter Spline-Funktionen verweisen wir auf die ausführlichen Erläuterungen von Wood (2017).

Nachdem Generalisierte Additive Modelle für alle verschiedenen Copula-Parameter angepasst wurden, kann der Einfluss von verschiedenen geographischen Einflussfaktoren auf die räumliche Abhängigkeitsstruktur der Beobachtungen untersucht werden. Mit den angepassten Modellen können zum Beispiel Prognosen $\hat{\delta}_{GAM}(G_1, \dots, G_M)$ für die Modellparameter für festgelegte Werte der räumlichen Einflussfaktoren $G_1(s_i, s_j), \dots, G_K(s_i, s_j)$ getroffen werden. Die Abhängigkeitsstruktur kann dadurch rein mit den verfügbaren, geographischen Merkmalen zweier Standorte beschrieben werden. Da verschiedene Abhängigkeitsmaße (wie z.B. Randabhängigkeitskoeffizienten) für viele Copula-Familien direkt mit Hilfe der Copula-Parameter ausgedrückt werden können, liefern die angepassten Modelle eine direkte Vergleichsmöglichkeit von Abhängigkeitsstrukturen für unterschiedliche geographische Gegebenheiten sowie für verschiedene Szenarien.

4.3 Anwendung: Identifizierung von räumlichen Abhängigkeitsstrukturen bei deterministischen Windkraftvorhersagefehlern

Im Folgenden verwenden wir das vorgestellte Verfahren für die Identifizierung von räumlichen Abhängigkeiten bei Windkraftvorhersagefehlern in Europa. Für die Analyse nutzen wir erneut die Datensätze mit deterministischen Windkraftvorhersagefehlern für das europäische Festland im Zeitraum von April 2010 bis Februar 2016, die bereits im vorherigen Kapitel verwendet wurden. Aus Übersichtlichkeitsgründen beschränken wir uns in diesem Kapitel jedoch auf einen Vergleich der Abhängigkeitsstrukturen für 24-Stunden-Vorhersagefehler und 72-Stunden-Vorhersagefehler.

4.3.1 Deskriptive Analyse und Wahl einer geeigneten Copula-Familie

Bevor wir das Verfahren aus Abschnitt 4.2 zur Identifizierung von räumlichen Abhängigkeiten anwenden, führen wir zunächst eine deskriptive Zeitreihenanalyse durch. Dies erlaubt uns einige Annahmen zu überprüfen, die bei der weiteren Modellierung implizit getroffen werden. Für diese deskriptive Analyse betrachten wir zunächst nur die Vorhersagefehler der 24-Stunden-Prognose. Beispielhaft wählen wir den Gitterpunkt s_{17058} (gelegen in Norddeutschland) aus der Menge aller Standorte aus und visualisieren die beobachtete Vorhersagefehlerzeitreihe am entsprechenden Standort in Abbildung 4.3. Es fällt auf, dass die Vorhersagefehler über die Zeit um den Wert Null schwanken und dass bei einigen Zeitpunkten auffällig große Vorhersagefehler vorliegen. Allerdings sind auf den ersten Blick keine Zeitfenster mit systematisch höheren oder niedrigeren Vorhersagefehlern zu erkennen. Eine detailliertere Untersuchung der Abhängigkeitsstruktur über die Zeit erlaubt die Betrachtung von empirischen Autokorrelationsfunktionen. Abbildung 4.4 zeigt die empirischen Autokorrelationen für die Vorhersagefehlerzeitreihe am Standort s_{17058} . Wie an der Grafik zu erkennen ist, sind für die betrachtete Zeitreihe keine ausgeprägten Autokorrela-

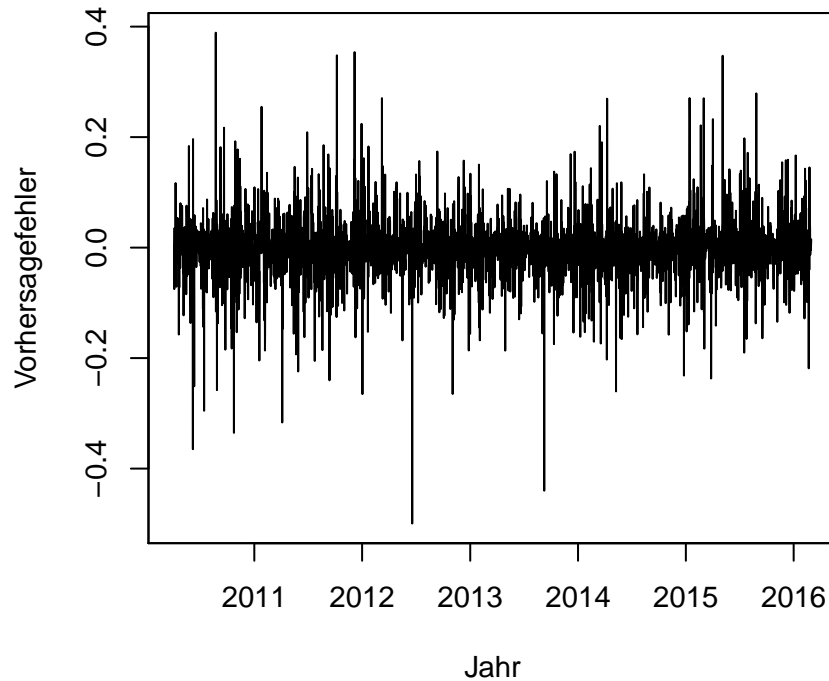


Abbildung 4.3: Zeitreihe der 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)

tionen zu verzeichnen. Die Vorhersagefehler scheinen daher über die Zeit unkorreliert zu sein. Ein ähnliches Bild bietet die Betrachtung der empirischen Autokorrelationen für die quadrierten Vorhersagefehler am entsprechenden Standort in Abbildung 4.5. Auch hier ist keine deutliche Abhängigkeitsstruktur zu erkennen. Vergleichbare Ergebnisse ergeben sich bei der Analyse von Vorhersagefehlerzeitreihen für andere Standorte. Bei einer Anwendung des in Abschnitt 4.2 vorgestellten Verfahrens wird implizit angenommen, dass die Beobachtungen an einem bestimmten Standort als Realisationen von unabhängig, identisch verteilten Zufallsvariablen angesehen werden können. Aufgrund der nur gering vorhandenen zeitlichen Abhängigkeitsstrukturen verzichten wir für die Analyse auf eine Nutzung von klassischen Zeitreihenverfahren, mit denen systematische Abhängigkeitsstrukturen im Vorfeld bereinigt werden können. Die Annahme unabhängiger Vorhersagefehler kann aus meteorologischer Sicht auch dadurch gerechtfertigt werden, dass zwischen

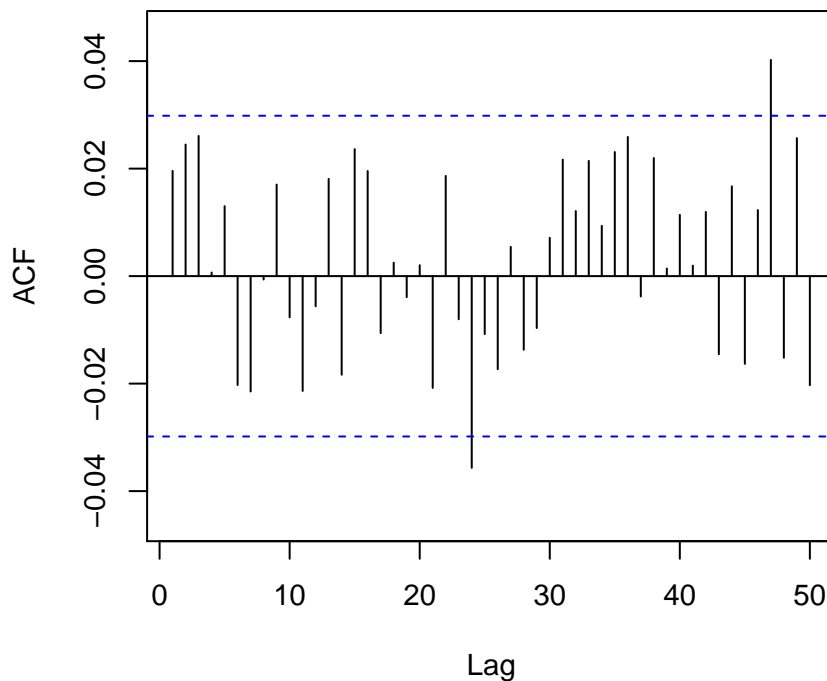


Abbildung 4.4: Empirische Autokorrelationen der 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)

zwei Beobachtungen ein relativ langer Zeitabstand von 12 Stunden liegt, und dass jeder Vorhersagefehler aus einer neuen Initialisierung des deterministischen Vorhersagemodells des Wetteranbieters ECMWF stammt. Die Annahme identisch verteilter Vorhersagefehler über die Zeit ist aus meteorologischer Sicht hingegen kritisch zu hinterfragen, weil grundsätzlich im Zeitverlauf Wetterperioden mit einer höheren oder geringeren Vorhersageunsicherheit zu erwarten sind. Die Ergebnisse mit der vorgestellten Methodik müssen deshalb auch in diesem Hinblick kritisch diskutiert werden.

Für die Anwendung des vorgestellten Verfahrens überprüfen wir in einem ersten Schritt zunächst, ob eine geeignete Copula-Familie existiert, mit der die Abhängigkeitsstruktur zwischen den Beobachtungen für verschiedene Standortpaare geeignet modelliert werden kann. Um die Wahl einer geeigneten Copula-Familie zu demonstrieren, betrachten wir zunächst beispielhaft die Gitterpunkte s_{17058} und s_{17075} , die innerhalb einer Distanz

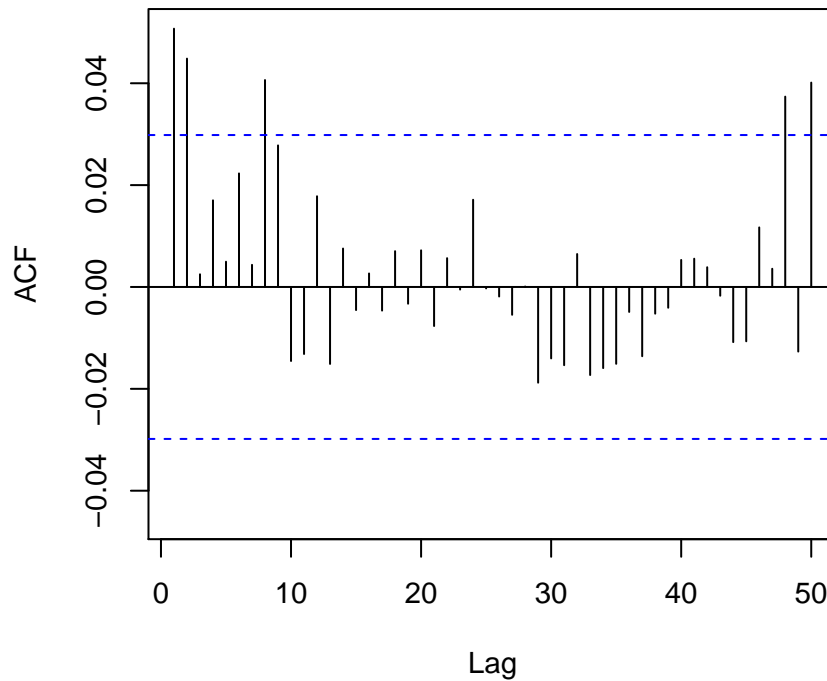


Abbildung 4.5: Empirische Autokorrelationen der quadrierten 24-Stunden-Vorhersagefehler am Standort s_{17058} (gelegen in Norddeutschland)

von etwa 140 Kilometern in Norddeutschland liegen. Abbildung 4.6 zeigt ein Streudiagramm der 24-Stunden-Vorhersagefehler an den beiden Standorten. Bei Betrachtung des Streudiagramms zeigt sich, dass die Vorhersagefehler an den beiden Standorten positiv korrelieren. Insbesondere ist auffällig, dass sehr große negative oder positive Vorhersagefehler gleichzeitig an beiden Standorten auftreten. Dies wird auch bei Betrachtung eines Streudiagramms rangtransformierter Beobachtungen (berechnet wie in Gleichung (4.7)) in Abbildung 4.7 deutlich. Im Falle von Bindungen (die vereinzelt bei den Vorhersagefehlern auftreten) werden den entsprechenden Beobachtungen zufällige Ränge vergeben. Die sich so ergebende empirische Copula zeigt eine Anhäufung von Beobachtungen an den Rändern der Wertebereiche und charakterisiert somit das gemeinsame Auftreten von sehr großen Vorhersagefehlern in einem räumlichen Kontext. Wir verwenden im Weiteren die Funktion `BiCopSelect` aus dem Paket `VineCopula` (siehe Schepsmeier et al., 2017) in R um die An-

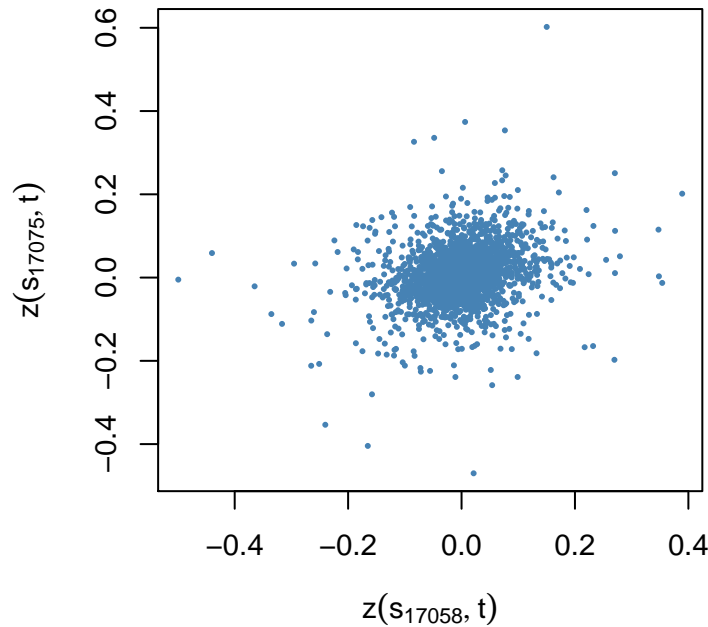


Abbildung 4.6: Streudiagramm der 24-Stunden-Vorhersagefehler an den Standorten s_{17058} und s_{17075} (gelegen innerhalb einer Distanz von etwa 140 Kilometern in Norddeutschland)

passungsgüte verschiedener Copula-Familien an die Pseudo-Beobachtungen hinsichtlich des AIC-Kriteriums zu vergleichen. Die Funktion ermöglicht einen direkten Vergleich für knapp 40 verschiedene Copula-Familien. Für die in Abbildung 4.7 betrachteten Pseudo-Beobachtungen liefert eine t -Copula mit geschätzten Parametern von $\hat{\rho}_{MPL} \approx 0.307$ und $\hat{\nu}_{MPL} \approx 2.391$ den kleinsten AIC-Wert. Um zu überprüfen, ob die angepasste Copula die Abhängigkeitsstruktur auf adäquate Weise beschreibt, simulieren wir probeweise $T = 4316$ (Anzahl der Zeitpunkte im Datensatz) Pseudo-Beobachtungen aus einer t -Copula mit den entsprechenden Parametern. Diese sind in Abbildung 4.8a in einem Streudiagramm veranschaulicht. Beim Vergleich der simulierten Werte mit der empirischen Copula in Abbildung 4.7 wird deutlich, dass die t -Copula die Abhängigkeitsstruktur geeignet beschreibt. Insbesondere berücksichtigt die t -Copula, dass eine relativ hohe Anzahl an Beobachtungen an den Rändern des Wertebereichs gemeinsam an beiden Standorten auftreten. Bei Betrachtung simulierter Beobachtungen aus einer Gauß-Copula mit Parameter $\hat{\rho}_{MPL} \approx 0.307$ in

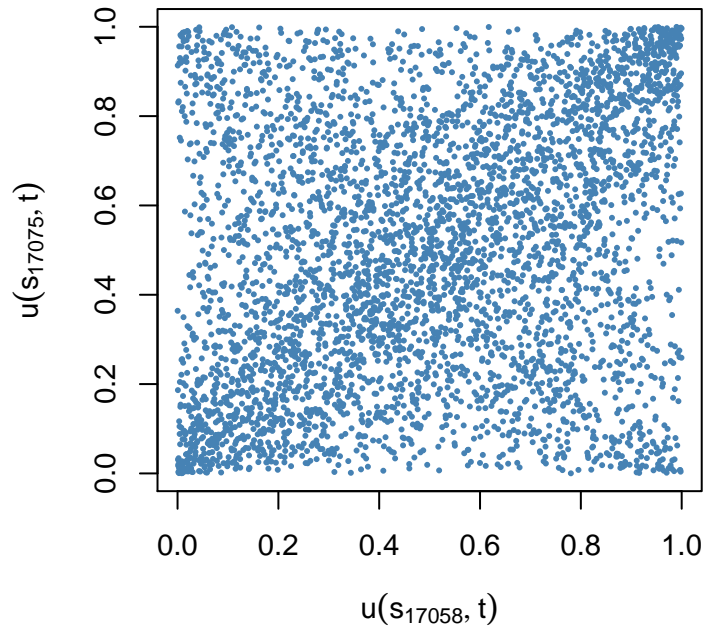
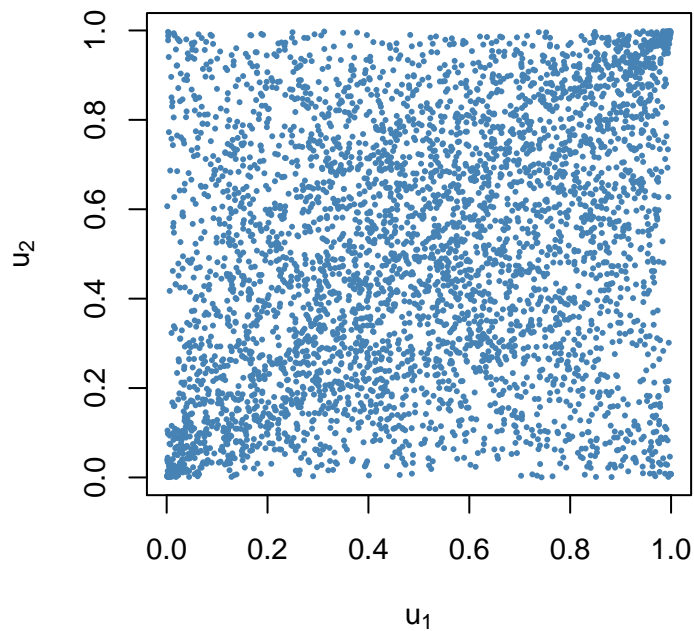
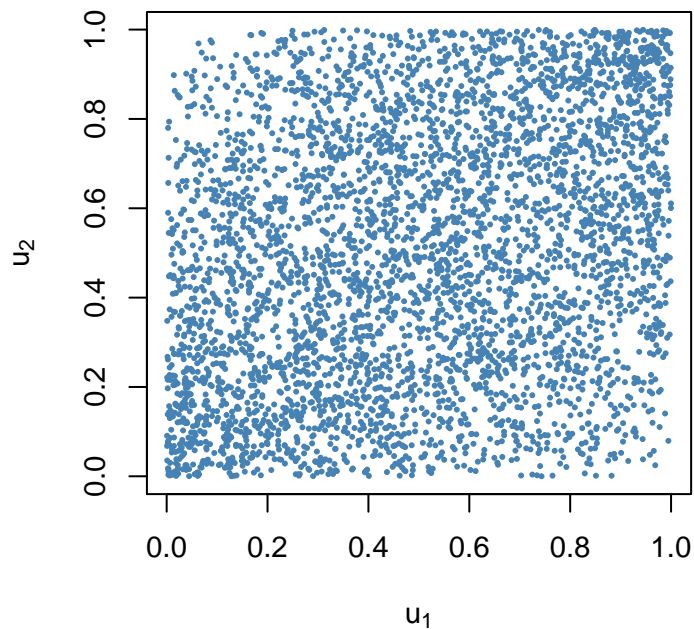


Abbildung 4.7: Streudiagramm rangtransformierter 24-Stunden-Vorhersagefehler an den Standorten s_{17058} und s_{17075} (gelegen innerhalb einer Distanz von etwa 140 Kilometern in Norddeutschland)

Abbildung 4.8b zeigt sich auch illustrativ, dass die t -Copula die vorliegende Abhängigkeitsstruktur zum Beispiel deutlich besser verkörpert als eine Gauß-Copula. Dies zeigt sich insbesondere daran, dass bei der Gauß-Copula die Abhängigkeiten in den Rändern deutlich schwächer repräsentiert werden. Um zu überprüfen, ob die Familie der t -Copulas für die Beschreibung der Abhängigkeitsstruktur auch für verschiedene Standortpaare $\{s_i, s_j\}$ geeignet ist, wählen wir mit einer Latin-Hypercube-Stichprobe abschließend noch eine kleine Anzahl von 1000 Standortpaaren aus, die verschiedene Distanzen zwischen zwei Standorten bis zu 600 Kilometern abdecken. Für alle 1000 Paare wird daraufhin die Copula-Familie bestimmt, für die sich das kleinste AIC ergibt. Hier wird bei insgesamt 995 Standortpaaren eine t -Copula ausgewählt. In den anderen fünf Fällen liegt das AIC für die beste Copula-Familie nur geringfügig unter dem AIC für eine t -Copula, was für eine ähnlich gute Anpassungsgüte der beiden Copulas spricht. Aus diesem Grund verwenden wir die Familie der t -Copulas für die weiterführende Analyse im folgenden Unterabschnitt.



(a) t-Copula mit Parametern $\hat{\rho}_{MPL} \approx 0.307$ und $\hat{\nu}_{MPL} \approx 2.391$



(b) Gauß-Copula mit Parameter $\hat{\rho}_{MPL} \approx 0.307$

Abbildung 4.8: Simulierte Pseudo-Beobachtungen aus einer t-Copula und einer Gauß-Copula mit Parametern, die anhand von 24-Stunden-Vorhersagefehlern an den beiden Standorten s_{17058} und s_{17075} geschätzt wurden

4.3.2 Anwendung des Verfahrens zur Identifizierung von räumlichen Abhängigkeitsstrukturen

Nach der Wahl einer geeigneten Copula-Familie kann nun das vorgestellte Verfahren angewandt werden, um Abhängigkeitsstrukturen bei Windkraftvorhersagefehlern in Hinblick auf unterschiedliche geographische Bedingungen zu charakterisieren. Dafür ist es notwendig, bedeutsame geographische Einflussfaktoren zu berücksichtigen, die für die räumliche Abhängigkeit von Vorhersagefehlern relevant sein könnten. Da bei der Clusteranalyse im vorherigen Kapitel deutliche Unterschiede zwischen Kreuzkorrelationen in flachen Gebieten und Gebirgsregionen sichtbar wurden, berücksichtigen wir neben der räumlichen Distanz zweier Standorte auch orographische Informationen. Dazu verwenden wir Angaben des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) zur zugrundeliegenden Modellorographie (Höheninformationen) des deterministischen Vorhersagemodells (abgebildet in Abbildung 3.1 in Kapitel 3, S. 25). Zum einen verwenden wir dabei die maximale Höhe über dem Meeresspiegel $\text{MaxH}(s_i, s_j)$ zweier Standorte. Da nur die maximale Höhe allerdings noch keine Information liefert, ob zwei Standorte durch ähnliche oder komplett unterschiedliche Geländebedingungen charakterisiert sind, definieren wir außerdem den sogenannten „Orographie-Index“

$$\text{OI}(s_i, s_j) := 1 - \frac{|H(s_i) - H(s_j)|}{\text{MaxH}(s_i, s_j)} \quad (4.12)$$

zweier Standorte s_i und s_j . Dabei bezeichne $H(s_i)$ die Höhe über dem Meeresspiegel des Standortes s_i . Der Orographie-Index ist auf das Intervall $[0, 1]$ beschränkt, wobei ein Wert von 1 bedeutet, dass zwei Standorte in einer exakt gleichen Höhe über dem Meeresspiegel liegen. Ein Wert nahe 0 kann hingegen je nach maximaler Höhe bedeuten, dass sich zwei Standorte durch komplett unterschiedliche Geländebedingungen auszeichnen. In Kombination mit der räumlichen Distanz und der maximalen Höhe zweier Standorte sollte der Orographie-Index eine gezielte Charakterisierung von Abhängigkeitsstrukturen zwischen Vorhersagefehlern für unterschiedliche geographische Gegebenheiten erlauben.

Für die Modellierung müssen zunächst Standortpaare identifiziert werden, die den gesamten Wertebereich der verschiedenen potenziellen Einflussgrößen (räumliche Distanz, maximale Höhe, Orographie-Index) geeignet abdecken. Für die Auswahl standardisieren wir die Werte der verschiedenen Einflussgrößen zunächst, damit ähnliche Wertebereiche für alle Variablen vorliegen. Daraufhin ziehen wir eine dreidimensionale Latin-Hypercube-Stichprobe mit einem Stichprobenumfang von 500 000 Beobachtungen mit Hilfe der Funktion `randomLHS` aus dem Paket `lhs` (siehe Carnell, 2018) in R, die die gesamten Wertebereiche der standardisierten Einflussgrößen abdeckt. Für alle 500 000 simulierten Werte bestimmen wir danach die Standortpaare in unserem Datensatz, bei denen die standardisierten geographischen Merkmale im Sinne der euklidischen Distanz am geringsten von den simulierten Werten abweichen. Da durch diese Auswahlstrategie bestimmte Standortpaare mehrfach ausgewählt werden, verbleiben endgültig $B = 478\,638$ eindeutige Standortpaare,

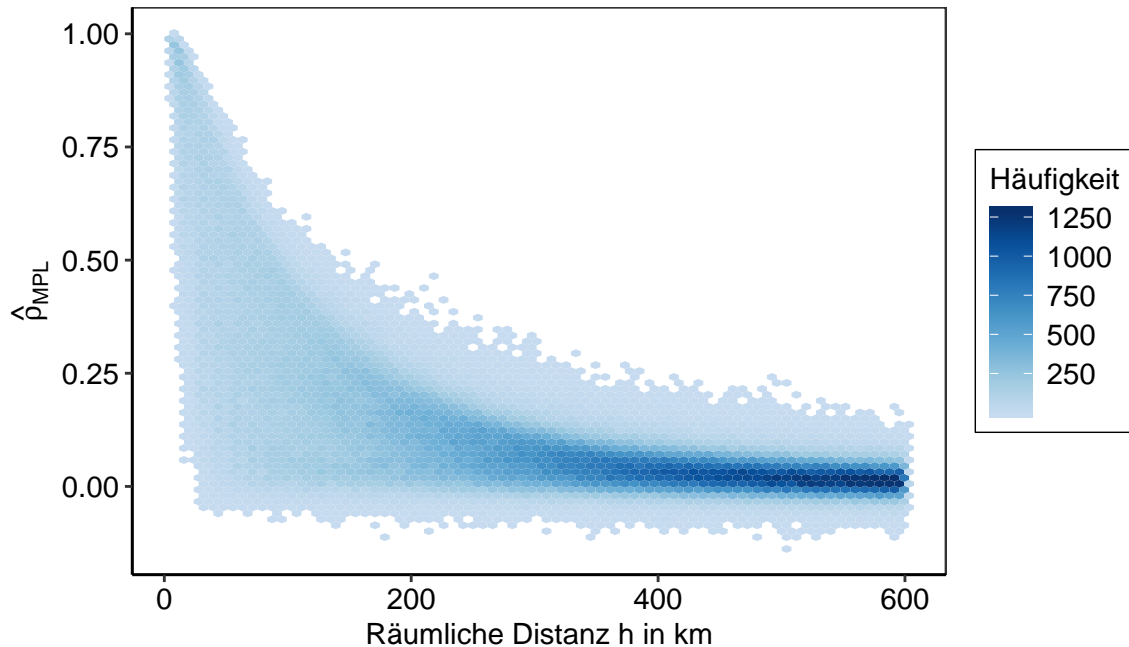


Abbildung 4.9: Geschätzte Werte $\hat{\rho}_{MPL}$ des Parameters ρ einer t-Copula für 24-Stunden-Vorhersagefehler, dargestellt gegenüber der räumlichen Distanz der Standortpaare. Berücksichtigt werden $B = 478\,638$ Standortpaare, die unterschiedliche räumliche Gegebenheiten repräsentieren.

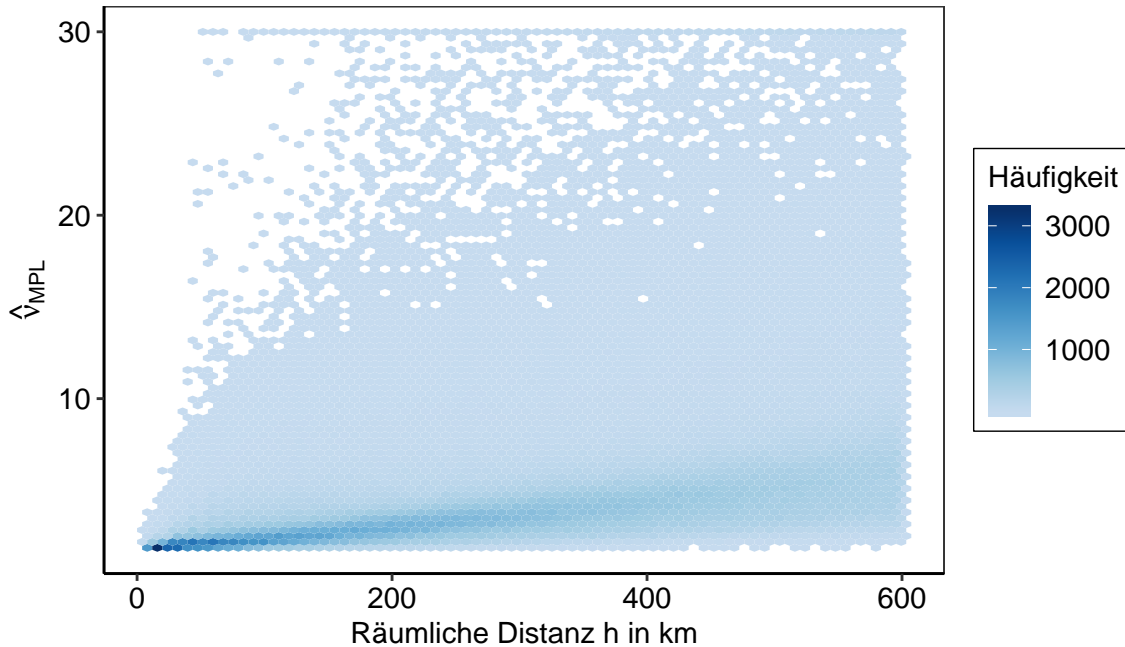


Abbildung 4.10: Geschätzte Werte $\hat{\nu}_{MPL}$ des Parameters ν einer t-Copula für 24-Stunden-Vorhersagefehler, dargestellt gegenüber der räumlichen Distanz der Standortpaare. Berücksichtigt werden $B = 478\,638$ Standortpaare, die unterschiedliche räumliche Gegebenheiten repräsentieren.

die verschiedene geographische Gegebenheiten gleichmäßig abdecken. Für diese Standortpaare schätzen wir die Parameter einer t-Copula mit dem Maximum-Pseudo-Likelihood-Ansatz. In diesem Abschnitt führen wir die Schätzung zunächst nur für die 24-Stunden-Vorhersagefehler durch. In den Abbildungen 4.9 und 4.10 werden die geschätzten Korrelationsparameter $\hat{\rho}_{MPL}^{(b)}$ und die geschätzten Freiheitsgrade $\hat{\nu}_{MPL}^{(b)}$ für $b = 1, \dots, B$ gegen die räumliche Distanzen der entsprechenden Standortpaare abgetragen. Bei Betrachtung der geschätzten Korrelationsparameter in Abbildung 4.9 ist es keine Überraschung, dass die Parameterschätzwerte mit größerer räumlicher Distanz zwischen zwei Standorten geringer ausfallen. Es zeigt sich aber auch, dass nur die räumliche Distanz die geschätzten Parameterwerte nicht auf eine adäquate Weise erklären kann. Speziell für geringe räumliche Distanzen zeigt die Grafik eine sehr hohe Variabilität in den resultierenden Schätzwerten für unterschiedliche Standortpaare. Betrachtet man die geschätzten Werte für den

Parameter ν bestätigt sich der Eindruck aus der Simulationsstudie in Abschnitt 4.2.1, dass die Parameterschätzwerte eine relativ große Varianz aufweisen. Abbildung 4.10 zeigt außerdem, dass die Parameterschätzwerte für größere räumliche Distanzen zwischen zwei Standorten höher ausfallen. Wie bereits in Gleichung (4.6) in Abschnitt 4.2.1 angemerkt, ist der Parameter ν insbesondere für die Randabhängigkeitsstruktur einer t-Copula relevant. Kleinere Werte von ν bedeuten dabei ausgeprägtere Randabhängigkeiten, weshalb das Ergebnis von höheren Schätzwerten bei größeren Distanzen aus meteorologischer Sicht sinnvoll erscheint. Das gemeinsame Auftreten von sehr hohen Vorhersagefehlern ist naturgemäß eher für solche Standortpaare zu erwarten, die räumlich nahe beieinanderliegen.

Zusammengefasst zeigt bereits die graphische Analyse, dass Modelle die nur die räumliche Distanz zweier Standorte verwenden, die Unterschiede in den Parameterschätzwerten nicht geeignet beschreiben können. Bei der Modellierung berücksichtigen wir deshalb auch mögliche Effekte von unterschiedlichen orographischen Gegebenheiten und spezifizieren Generalisierte Additive Modelle der Form

$$\begin{aligned} g_q \left\{ \mathbb{E}(\hat{\delta}_{q,MPL}^{(b)}) \right\} &= \beta_{0,q} + f_{1,q}(\text{Distanz}^{(b)}) + f_{2,q}(\text{MaxH}^{(b)}) + f_{3,q}(\text{OI}^{(b)}) \\ &+ f_{4,q}(\text{Distanz}^{(b)}, \text{MaxH}^{(b)}) + f_{5,q}(\text{Distanz}^{(b)}, \text{OI}^{(b)}) \\ &+ f_{6,q}(\text{MaxH}^{(b)}, \text{OI}^{(b)}) + f_{7,q}(\text{Distanz}^{(b)}, \text{MaxH}^{(b)}, \text{OI}^{(b)}) \end{aligned} \quad (4.13)$$

mit $b = 1, \dots, B$ für $q = 1, 2$, wobei $\hat{\delta}_{1,MPL}^{(b)} = \hat{\rho}_{MPL}^{(b)}$ und $\hat{\delta}_{2,MPL}^{(b)} = \hat{\nu}_{MPL}^{(b)}$. Als Link-Funktionen verwenden wir aufgrund der verschiedenen Parameterbereiche der Copula-Parameter $g_1(y) = \log\left(-\log\left(\frac{1-y}{2}\right)\right)$ und $g_2(y) = \log(y-2)$. Die flexible Modellformel in Gleichung (4.13) erlaubt die Berücksichtigung von Haupteffekten wie der räumlichen Distanz und ermöglicht zugleich die Betrachtung von Wechselwirkungen, wie etwa zwischen der maximalen Höhe zweier Standorte und dem Orographie-Index. Da anzunehmen ist, dass sich der Effekt unterschiedlicher geographischer Gegebenheiten für größere räumliche Distanzen eher weniger stark auf die Parameterschätzwerte auswirkt, werden außerdem Interaktionen zwischen der räumlichen Distanz und den anderen Einflussgrößen in der

Modellgleichung berücksichtigt. Für die glatten Funktionen verwenden wir im Rahmen dieser Analyse kubische P-Splines bzw. Tensorprodukte von kubischen P-Splines (für Details siehe Wood, 2017, S. 204-205).

Um die Anpassungsgüte der Modelle zu untersuchen, vergleichen wir in Abbildung 4.11 zunächst die Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ für Standortpaare mit einer maximalen Höhe im Intervall (90,110) und einem Orographie-Index im Intervall

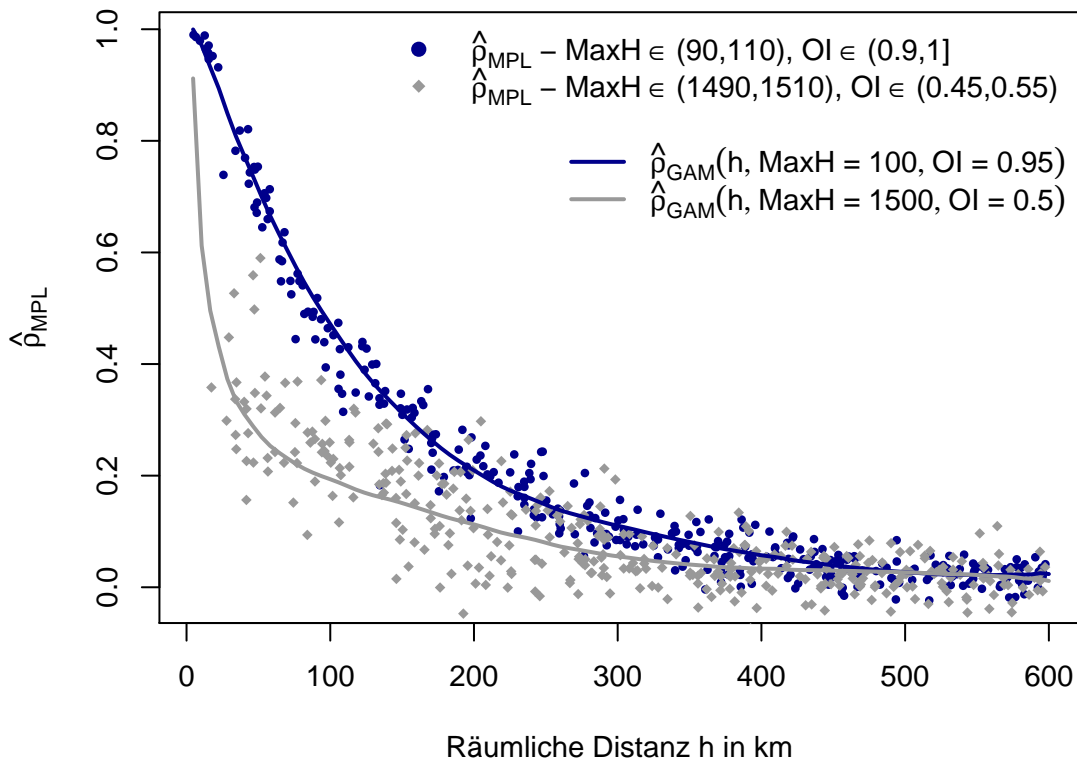


Abbildung 4.11: Modellierung der Schätzwerte $\hat{\rho}_{MPL}$ des Parameters ρ einer t-Copula (geschätzt für $B = 478\,638$ Standortpaare) für die 24-Stunden-Vorhersagefehler mit einem GAM basierend auf der räumlichen Distanz h in Kilometern, der maximalen Höhe über dem Meeresspiegel in Metern (MaxH) und dem Orographie-Index (OI). Die Modellvorhersagen mit dem angepassten GAM werden exemplarisch für zwei verschiedene Spezifizierungen für die maximale Höhe und den Orographie-Index verglichen.

(0.9, 1] mit den Schätzwerten für Standortpaare mit einer maximalen Höhe im Intervall (1490, 1510) und einem Orographie-Index im Intervall (0.45, 0.55). Es zeigt sich deutlich, dass die geschätzten Parameterwerte für die Standortpaare mit einer geringen maximalen Höhe und einem hohen Orographie-Index, also zwei Standorte die beide in einer eher flacheren Region liegen (illustriert in blau in Abbildung 4.11), für kurze räumliche Distanzen deutlich höher ausfallen als die Schätzwerte für die anderen Standortpaare. Die Modellvorhersagen $\hat{\rho}_{GAM}(h, \text{MaxH}, \text{OI})$ mit dem angepassten GAM variieren sichtbar für verschiedene Werte der maximalen Höhe und des Orographie-Index. Insgesamt passen die dargestellten Kurven gut zu den Maximum-Pseudo-Likelihood-Schätzwerten im Streudiagramm. Wie erwartet liegt für größere räumliche Distanzen kein nennenswerter Unterschied in den Parameterschätzwerten mehr vor, was vom Modell auch adäquat aufgegriffen wird.

Die hohe Anpassungsgüte des Modells zeigt sich auch bei einem räumlichen Vergleich von vorhergesagten Parameterschätzwerten in den Abbildungen 4.12 und 4.13. Hier werden die Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ und die Modellvorhersagen $\hat{\rho}_{GAM}$ mit dem angepassten Generalisierten Additiven Modell für zwei zentrale Standorte s_{17058} (gelegen in Norddeutschland) und s_{32787} (gelegen in der Schweiz) miteinander verglichen. Für beide zentralen Standorte wird der Parameter ρ einer bivariaten t-Copula mit dem Maximum-Pseudo-Likelihood-Ansatz für alle räumlichen Nachbarn geschätzt, die innerhalb einer Distanz von 600 Kilometern zum jeweiligen Standort liegen. Außerdem wird das angepasste GAM verwendet, um Modellvorhersagen $\hat{\rho}_{GAM}$ für alle betrachteten Standortpaare nur mit Hilfe von geographischen Informationen zu treffen. Der Vergleich zeigt, dass das angepasste Modell die unterschiedlichen Abhängigkeitsstrukturen in den verschiedenen Regionen angemessen abbildet. Insbesondere erkennt das Modell, dass in flachen Gebieten wie in Norddeutschland deutlich ausgeprägtere Abhängigkeitsstrukturen vorliegen als in Gebirgsregionen wie in den Alpen.

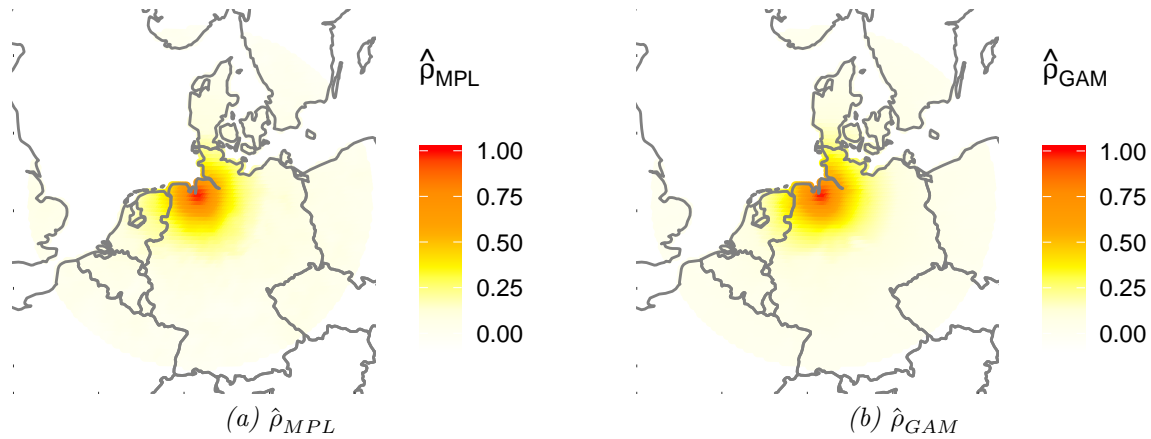


Abbildung 4.12: Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ und GAM-Modellvorhersagen $\hat{\rho}_{GAM}$ für den Parameter ρ einer t-Copula, welche die Abhängigkeitsstruktur zwischen Vorhersagefehlern am Standort s_{17058} (in Norddeutschland) und an benachbarten Standorten beschreibt

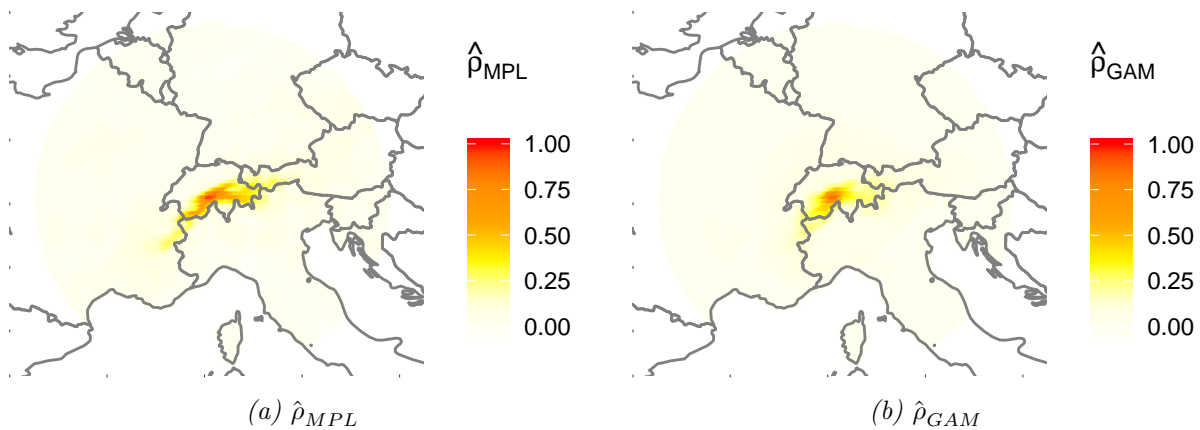


Abbildung 4.13: Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\rho}_{MPL}$ und GAM-Modellvorhersagen $\hat{\rho}_{GAM}$ für den Parameter ρ einer t-Copula, welche die Abhängigkeitsstruktur zwischen Vorhersagefehlern am Standort s_{32787} (in der Schweiz) und an benachbarten Standorten beschreibt

Die Anpassungsgüte des Generalisierten Additiven Modells an die Schätzwerte $\hat{\nu}_{MPL}$ des zweiten Copula-Parameters ν , ist im Vergleich zum Modell für die geschätzten Korrelationsparameter schlechter. Dies liegt vor allem an der hohen Varianz der Parameterschätzwerte, die bereits in Abbildung 4.10 sichtbar wurde. Das angepasste Modell greift hauptsächlich die ansteigenden Parameterschätzwerte für ν für größere räumliche Distanzen zwischen zwei Standorten auf, die auch aus meteorologischer Sicht sinnvoll erscheinen. Systematische Unterschiede hinsichtlich verschiedener geographischer Gegebenheiten können wir mit Hilfe des angepassten Modells jedoch nicht direkt identifizieren. Für die weitere Interpretation der Ergebnisse fokussieren wir uns deshalb auf die Interpretation des angepassten Generalisierten Additiven Modells an die Parameterschätzwerte $\hat{\rho}_{MPL}$.

In Abbildung 4.14 vergleichen wir beispielhaft die Modellvorhersagen mit dem angepassten GAM für verschiedene Einstellungen der geographischen Einflussgrößen (Räumliche Distanz, maximale Höhe und Orographie-Index) für die 24-Stunden-Vorhersagefehler. Dabei wird einmal eine räumliche Distanz von 20 Kilometern und einmal eine Distanz von 200 Kilometern fest gewählt, während die Werte für die maximale Höhe und den Orographie-Index variieren. Es ist nicht überraschend, dass die vorhergesagten Korrelationsparameter für Standorte, die innerhalb einer räumlichen Distanz von 20 Kilometern liegen, generell höher ausfallen als für Standorte, die eine räumliche Distanz von 200 Kilometern aufweisen. Der Vergleich zeigt weiter, dass die maximale Höhe und der Orographie-Index eine substantielle Auswirkung auf die Modellvorhersagen haben. Für zwei Standorte die in einer Entfernung von 20 Kilometern liegen, werden die höchsten Korrelationsparameter bei einer geringen maximalen Höhe und einem hohen Orographie-Index prognostiziert. Dies entspricht einem Paar von Standorten, die beide in einer eher flachen Gegend wie Norddeutschland oder Dänemark liegen. Für Standorte, die komplett unterschiedliche Geländebedingungen verzeichnen (also Standortpaare mit einer hohen maximalen Höhe und einem kleinen Orographie-Index) werden die geringsten Korrelationsparameter vorhergesagt. Auch wenn die prognostizierten Parameter für größere Werte der maximalen Höhe generell geringer ausfallen, werden besonders dann noch hohe Korrelationen erwar-

tet, wenn zwei Standorte ähnliche Geländebedingungen aufweisen. Der Orographie-Index ist also insbesondere dann relevant, wenn zumindest einer der beiden Standorte in einer Gebirgsregion liegt. Für zwei Standorte die in einer Distanz von 200 Kilometern liegen, werden ausgeprägte Korrelationsstrukturen vor allem für flachere Regionen prognostiziert. In Gebirgsregionen scheinen die Vorhersagefehler bei einer derart hohen Distanz hingegen tendenziell unkorreliert zu sein.

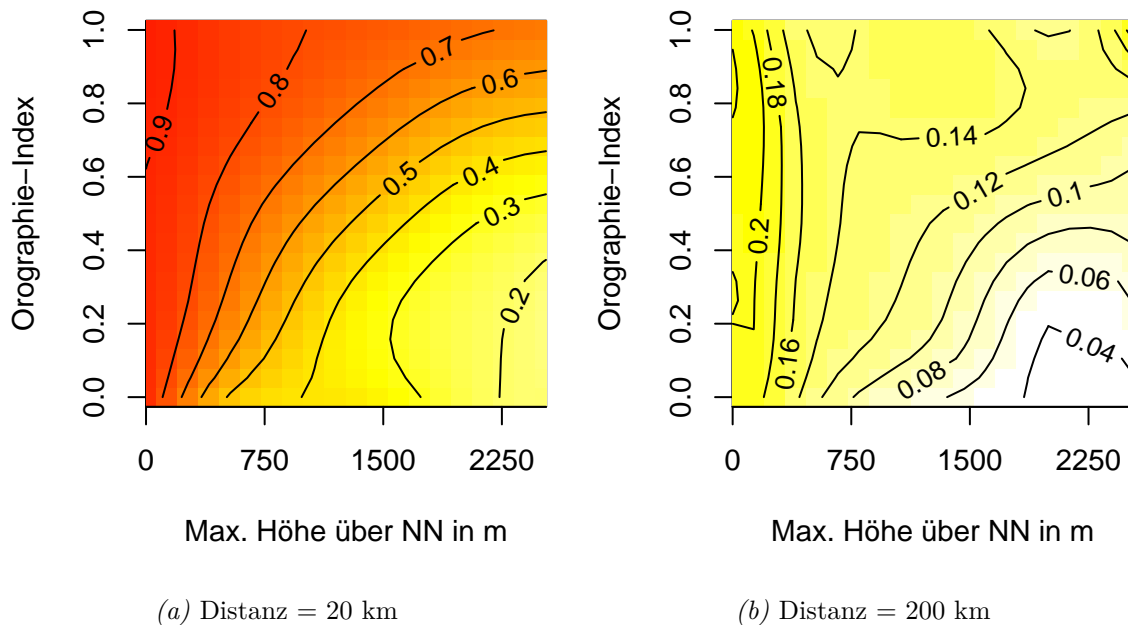


Abbildung 4.14: Modellvorhersagen des Parameters ρ einer t-Copula für die 24-Stunden-Vorhersagefehler mit dem angepassten GAM basierend auf der räumlichen Distanz, der maximalen Höhe und dem Orographie-Index (OI). Die Modellvorhersagen werden für festgelegte räumliche Distanzen von 20 und 200 Kilometern und variierende Werte für die maximale Höhe und den Orographie-Index visualisiert.

4.3.3 Vergleich der Ergebnisse für 24-Stunden und 72-Stunden Vorhersagefehler

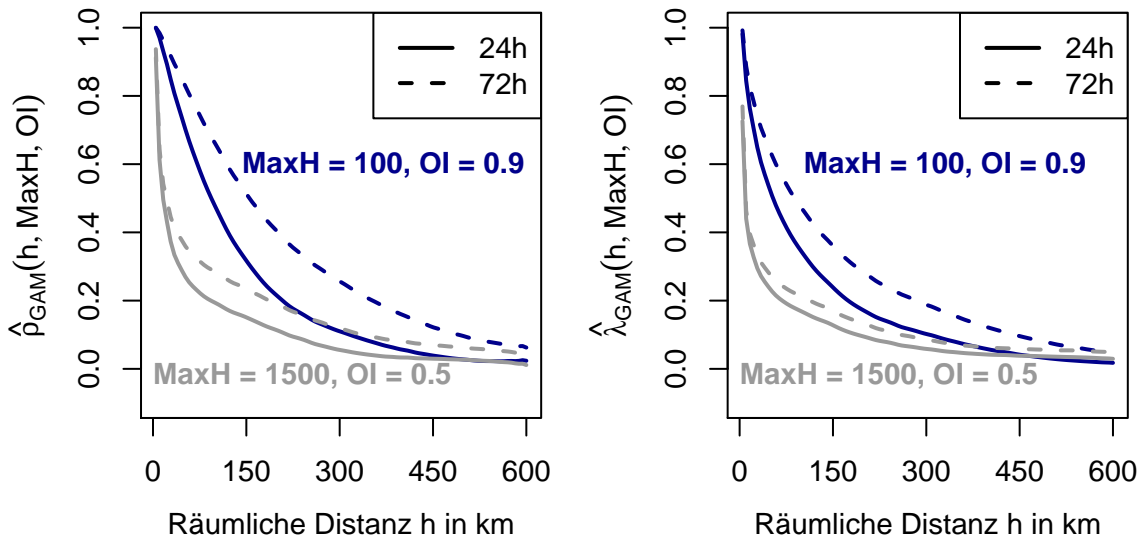
Die Clusteranalyse in Kapitel 3 hat bereits angedeutet, dass räumliche Korrelationsstrukturen bei Windkraftvorhersagefehlern für längere Vorhersagehorizonte ausgeprägter zu sein scheinen. Um diesen Aspekt noch genauer zu untersuchen, vergleichen wir im Folgenden die Ergebnisse für die 24-Stunden-Vorhersagefehler mit denen für einen längeren Vorhersagehorizont (72 Stunden). Wie bereits für die 24-Stunden-Vorhersagefehler stellt sich heraus, dass die Familie der t -Copulas die am besten geeignete Copula-Familie zur Beschreibung der Abhängigkeitsstruktur für verschiedene Standortpaare hinsichtlich des AICs ist. Wir schätzen deshalb die Parameter einer t -Copula erneut für die im vorherigen Abschnitt ausgewählten 478 638 Standortpaare mit Hilfe des Maximum-Pseudo-Likelihood-Ansatzes. Diese Schätzwerte werden daraufhin verwendet, um Generalisierte Additive Modelle für $\hat{\rho}_{MPL}$ und $\hat{\nu}_{MPL}$ wie in Gleichung (4.13) aufzustellen. Da die Ergebnisse denen für die 24-Stunden-Vorhersagefehler ähneln, verzichten wir auf eine detaillierte Beschreibung der einzelnen Schritte der Analyse. Stattdessen vergleichen wir direkt die geschätzten Abhängigkeitsstrukturen für die 24-Stunden-Vorhersagefehler und die 72-Stunden-Vorhersagefehler. Abbildung 4.15a zeigt zunächst die Modellvorhersagen des Korrelationsparameters ρ einer t -Copula für beide Vorhersagehorizonte für zwei unterschiedliche geographische Gegebenheiten. Der Vergleich zeigt erneut, dass die Orographie einen nennenswerten Einfluss auf die räumliche Abhängigkeitsstruktur von Windkraftvorhersagefehlern hat. Außerdem bestätigt sich der Eindruck aus der Clusteranalyse, dass die Korrelationsstruktur für längere Vorhersagehorizonte ausgeprägter ist.

Da das größte Risiko für den Energiesektor allerdings dann besteht, wenn sehr große Vorhersagefehler gemeinsam in einem räumlichen Kontext auftreten, ist es von besonderem Interesse die geschätzten Randabhängigkeiten für verschiedene Vorhersagehorizonte und unterschiedliche geographische Gegebenheiten zu analysieren. Wie bereits in Gleichung (4.6) in Abschnitt 4.2.1 diskutiert, ergibt sich der (obere und untere) Randabhängigkeits-

koeffizient bei einer t-Copula als

$$\lambda = \lambda_L = \lambda_U = 2t_{\nu+1}(-\sqrt{\nu+1}\sqrt{1-\rho}/\sqrt{1+\rho}),$$

wobei $t_{\nu+1}$ eine univariate t-Verteilung mit $\nu + 1$ Freiheitsgraden bezeichnet. Mit Hilfe der Modellvorhersagen $\hat{\rho}_{GAM}(h, \text{MaxH}, \text{OI})$ und $\hat{\nu}_{GAM}(h, \text{MaxH}, \text{OI})$ ist es daher möglich, auch den Randabhängigkeitskoeffizienten nur mit Hilfe der räumlichen Distanz h , der maximalen Höhe und dem Orographie-Index zu prognostizieren. Die resultierenden Plug-In-Vorhersagen $\hat{\lambda}_{GAM}(h, \text{MaxH}, \text{OI})$ des Randabhängigkeitskoeffizienten werden in Abbildung 4.15b verglichen. Die Grafik deutet darauf hin, dass Randabhängigkeiten insbesondere für zwei Standorte vorhanden sind, die innerhalb einer geringen räumlichen Distanz zueinander liegen. Relevant sind hier auch erneut unterschiedliche geographische Gegebenheiten, wobei ausgeprägtere Randabhängigkeiten eher in flacheren Gebieten (blaue Linien) als in Gebirgsregionen (graue Linien) zu erwarten sind. Außerdem scheinen



(a) Korrelationsparameter ρ

(b) Randabhängigkeitskoeffizient λ

Abbildung 4.15: Modellvorhersagen des Parameters ρ einer t-Copula und des Randabhängigkeitskoeffizienten λ für 24-Stunden- und 72-Stunden-Vorhersagefehler

auch die Randabhängigkeitsstrukturen für längere Vorhersagehorizonte ausgeprägter zu sein als für kurze Vorhersagehorizonte.

Angemerkt werden muss, dass die Ergebnisse für die Randabhängigkeiten insgesamt mit Vorsicht interpretiert werden müssen. Problematisch ist hier vor allem, dass die Maximum-Pseudo-Likelihood-Schätzwerte $\hat{\nu}_{MPL}$, auf denen die prognostizierten Werte $\hat{\lambda}_{GAM}$ basieren, eine relativ hohe Varianz aufweisen. Dies führt zu einer schlechten Anpassung des Generalisierten Additiven Modells und damit auch automatisch zu einer Verzerrung der prognostizierten Randabhängigkeitskoeffizienten. Außerdem ist zu beachten, dass die vorgeschlagene Methodik ausschließlich auf bivariaten Copulas für verschiedene Standortpaare basiert, weshalb keine direkten Aussagen über die gemeinsame Abhängigkeitsstruktur der Vorhersagefehler an mehreren Standorten getroffen werden können. Weiterführende deskriptive Analysen zeigen jedoch, dass sich die gefundenen Abhängigkeitsstrukturen auch direkt in unterschiedlich hohen, aggregierten Vorhersagefehlern in verschiedenen Regionen widerspiegeln. Wie in Abbildung 4.16 beispielhaft für die 72-Stunden-Vorhersagefehler am 04.02.2011 um 12:00 Uhr UTC illustriert, treten hohe aggregierte Vorhersagefehler insbesondere in flachen Gebieten wie z.B. in Norddeutschland und für längere Vorhersagehorizonte auf. In Gebirgsregionen und für kürzere Vorhersagehorizonte zeigt sich hingegen, dass aggregierte Vorhersagefehler aufgrund von geringeren räumlichen Abhängigkeitsstrukturen häufiger von räumlichen Glättungseffekten profitieren.

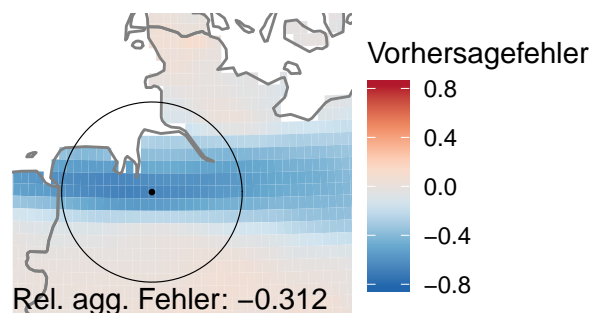


Abbildung 4.16: 72-Stunden-Vorhersagefehler in der Region um den Standort s_{17058} am 04.02.2011 um 12:00 Uhr UTC

4.4 Diskussion und Fazit

Die in diesem Kapitel vorgestellte Methode zur Identifizierung von räumlichen Abhängigkeitsstrukturen mit Copulas und Generalisierten Additiven Modellen ermöglicht eine detaillierte Charakterisierung von Abhängigkeitsstrukturen in Hinblick auf unterschiedliche geographische Gegebenheiten. Das Verfahren profitiert dabei insbesondere von einer hohen Flexibilität, da bei der Modellierung der paarweisen Abhängigkeiten für Standortpaare auf unterschiedlichste Copula-Familien zurückgegriffen werden kann. Dadurch können verschiedene Abhängigkeitsformen gezielt identifiziert werden. Das Verfahren erlaubt so im Vergleich zu korrelationsbasierten Methoden zusätzliche Einblicke in die Abhängigkeitsstruktur bei räumlich-zeitlichen Daten. So können zum Beispiel durch die Analyse der ausgewählten Copula-Familie Hinweise dafür gefunden werden, ob extreme Beobachtungen tendenziell gemeinsam in einem räumlichen Kontext auftreten. Das Verfahren ist außerdem dazu geeignet, räumliche Abhängigkeitsstrukturen in verschiedenen Regionen und für unterschiedliche Szenarien zu vergleichen. Durch die Wahl geographischer Einflussfaktoren bei der Modellierung der Parameterschätzwerte kann so beispielsweise der Einfluss unterschiedlicher Geländebedingungen auf die geschätzten Abhängigkeitsstrukturen untersucht werden.

Angemerkt werden muss, dass bei der Anwendung des Verfahren einige Annahmen getroffen werden, die kritisch diskutiert werden sollten. Zum einen wird implizit angenommen, dass die beobachteten Werte an einem bestimmten Standort als Realisationen von unabhängig und identisch verteilten Zufallsvariablen angesehen werden können. Diese Annahme ist insbesondere bei räumlich-zeitlichen Datensätzen oftmals schwer zu rechtfertigen. Im Anwendungsfall in Abschnitt 4.3 konnten über den Zeitverlauf nur schwach ausgeprägte Autokorrelationsstrukturen identifiziert werden. Falls in den Daten jedoch auffälliger Strukturen über die Zeit vorhanden sind, sind weitere Vorverarbeitungsschritte der Daten erforderlich, um weiterhin eine aussagekräftige Interpretation der Ergebnisse zu erlauben. Des Weiteren wird durch die Anpassung von Copulas für den gesamten beobachteten Zeit-

raum auch implizit unterstellt, dass sich die Abhängigkeitsstruktur über die Zeit nicht verändert. Eine Verletzung dieser Annahme kann ebenfalls zu einer möglichen Fehlinterpretation der Ergebnisse führen. So vergleichen zum Beispiel Manner und Reznikova (2012) die Anpassungsgüte verschiedener Copula-Familien für simulierte Daten aus bivariaten Gauß-Copulas mit zeitvariierendem Copula-Parameter ρ . Dabei stellt sich heraus, dass andere Copula-Familien als die Gauß-Copula im Sinne der Log-Likelihood-Statistik oftmals bevorzugt werden, wenn bei der Anpassung eine zeitkonstante Abhängigkeitsstruktur unterstellt wird. Laut den Autoren eignen sich in diesen Fällen zeitvariierende Copulas wie Stochastische Autoregressive Copulas (SCAR) oder Regime Switching Copulas (RSC) besser für die Modellierung. In zukünftigen Arbeiten soll das vorgeschlagene Verfahren daher so modifiziert werden, dass auch zeitveränderliche Abhängigkeitsstrukturen identifiziert werden können. Weiterhin ist zu beachten, dass die geschätzten Copula-Parameter bei einer geringen Anzahl an verfügbaren Beobachtungen eine hohe Varianz und eine Verzerrung aufweisen. In diesen Fällen ist die Aussagekraft der Ergebnisse daher nur gering. In zukünftigen Studien soll deshalb auch an einer Verbesserung der Methode für kleine Stichproben gearbeitet werden. Außerdem muss berücksichtigt werden, dass das Verfahren rein auf bivariaten Copulas basiert und deshalb keine Aussagen über die gemeinsame Abhängigkeitsstruktur der Beobachtungen an mehreren Standorten ermöglicht. Die Erkenntnisse aus der bivariaten Analyse können jedoch genutzt werden, um gezielte weiterführende Untersuchungen mit multivariaten Ansätzen durchzuführen.

Mit Hilfe des Verfahrens haben wir die räumliche Abhängigkeitsstruktur von deterministischen Windkraftvorhersagefehlern in Europa detailliert analysiert. Dabei hat sich die Erkenntnis aus dem vorherigen Kapitel bestätigt, dass räumliche Abhängigkeitsstrukturen bei Vorhersagefehlern insbesondere in flachen Regionen mit einer hohen Anzahl an installierter Windleistung zu verzeichnen sind. Die Auswahl der t -Copula-Familie als geeignete Copula-Familie hat zudem angedeutet, dass auch sehr große Vorhersagefehler häufig gemeinsam in einem räumlichen Kontext auftreten. Dies wird auch durch hohe aggregierte Vorhersagefehler sichtbar, die ein erhöhtes Risiko für den Energiesektor darstellen

können. Dieses Risiko kann jedoch durch die Verwendung von meteorologischen Ensemble-Vorhersagen bereits im Vorfeld gezielt eingeschätzt werden. Bei gemeinsamen Vorhersagen für mehrere Standorte ist es dabei erforderlich, dass die Ensemble-Vorhersagen die räumliche Abhängigkeitsstruktur geeignet abbilden. Das folgende Kapitel greift deshalb die Frage auf, wie die Qualität von verschiedenen Ensemble-Vorhersagen bzw. multivariaten probabilistischen Vorhersagen unter der Berücksichtigung von räumlichen Abhängigkeiten geeignet bewertet und verglichen werden kann.

5 Bewertung von multivariaten probabilistischen Windvorhersagen unter Berücksichtigung räumlicher Abhängigkeitsstrukturen

5.1 Motivation

Die Ergebnisse aus den beiden vorherigen Kapiteln verdeutlichen, dass räumliche Abhängigkeitsstrukturen bei deterministischen Windleistungsvorhersagefehlern insbesondere in windreichen Regionen auftreten. Dies kann in diesen Gebieten zum Teil mit hohen aggregierten Vorhersagefehlern einhergehen. Da akkurate regionale Windgeschwindigkeits- und Windleistungsvorhersagen eine hohe Bedeutung für den Energiesektor besitzen, widmen sich viele aktuelle Forschungsarbeiten der Frage, inwiefern probabilistische Vorhersagen eine zuverlässige Beurteilung der Vorhersageunsicherheit für eine gesamte Region bzw. für eine Vielzahl von räumlichen Standorten ermöglichen können (siehe zum Beispiel die Arbeiten von Cannon et al., 2017 und Späth et al., 2015). Die probabilistischen Vorhersagen basieren, wie in Kapitel 2 näher erläutert, dabei meist auf meteorologischen Ensemble-Prognosen. Diese werden von Wetteranbietern wie dem Europäischen Zentrum für mittelfristige Wettervorhersage für verschiedene Vorhersagehorizonte ausgegeben. Bei

der Beurteilung der Qualität dieser Ensemble-Vorhersagen fällt jedoch häufig auf, dass die Varianz der Ensemble-Mitglieder zu gering ist, um die tatsächlich vorliegende Unsicherheit geeignet abzubilden, oder dass eine systematische Verzerrung vorliegt (siehe z.B. Hamill und Colucci, 1997 oder Gneiting et al., 2005). Aus diesem Grund werden oft statistische Nachbearbeitungsverfahren wie zum Beispiel Ensemble Model Output Statistics (EMOS, siehe Gneiting et al., 2005 sowie Thorarinsdottir und Gneiting, 2010) oder Bayesian Model Averaging (BMA, siehe Sloughter et al., 2010) verwendet, mit denen die Qualität der originalen Ensemble-Vorhersagen in der Regel verbessert werden kann. Da diese populären Nachbearbeitungsverfahren jedoch hauptsächlich für Vorhersagen für einzelne Standorte geeignet sind, müssen diese Methoden modifiziert werden um eine hochwertige gemeinsame Vorhersage für mehrere Standorte zu ermöglichen (siehe dazu die Arbeiten von Möller et al., 2013; Hemri et al., 2015; Feldmann et al., 2015). Ein sehr populärer Ansatz ist, zunächst klassische Nachbearbeitungsverfahren wie EMOS separat für verschiedene Standorte zu verwenden und die dabei verlorengegangene räumliche Abhängigkeitsstruktur über sogenannte Ensemble-Copula-Coupling-Verfahren (siehe Schefzik et al., 2013) aus den originalen Ensemble-Vorhersagen zurückzugewinnen.

Aufgrund der Vielzahl an verfügbaren Nachbearbeitungsverfahren sowie einer großen Anzahl an Wetterdienstleistern, kommt es in der Praxis oft zu der Situation, dass die Qualität verschiedener probabilistischer Vorhersagen miteinander verglichen werden soll. Dies stellt insbesondere beim Vergleich von Ensemble-Vorhersagen mit unterschiedlichen Ensemble-Größen eine Herausforderung dar. Im univariaten Fall für einen einzelnen Standort wird die Qualität einer probabilistischen Windgeschwindigkeitsvorhersage dabei zum Beispiel häufig mit dem Continuous Ranked Probability Score (CRPS) bewertet. Falls eine Ensemble-Vorhersage vorliegt, die als unabhängig identisch verteilte Zufallsstichprobe aus einer bestimmten Wahrscheinlichkeitsverteilung interpretiert werden kann, muss jedoch beachtet werden, dass der gewöhnliche Schätzer des CRPS verzerrt ist (siehe Ferro et al., 2008 sowie Zamo und Naveau, 2018). Außerdem zeigen Zamo und Naveau (2018), dass der empirische CRPS auch sensitiv auf die Konstruktion eines Ensembles reagiert. Ein En-

semble das aus äquidistanten Quantilen einer Wahrscheinlichkeitsverteilung besteht, wird zum Beispiel laut den Autoren bei gleicher Ensemble-Größe in der Regel besser bewertet als eine reine Zufallsstichprobe aus der gleichen Verteilung.

Die oben angesprochenen Probleme sind auch bei der Bewertung von multivariaten probabilistischen Vorhersagen relevant. Aus diesem Grund beschäftigen wir uns in diesem abschließenden Kapitel mit der wichtigen Frage, wie die Qualität von verschiedenen multivariaten Ensemble-Vorhersagen unter Berücksichtigung von räumlichen Abhängigkeitsstrukturen sinnvoll miteinander verglichen werden kann - insbesondere mit Blick auf unterschiedliche Ensemble-Größen bzw. Ensemble-Strukturen. Dabei fokussieren wir uns auf zwei populäre Bewertungsregeln für multivariate Vorhersagen, den Energy Score (siehe Gneiting und Raftery, 2007) und den Variogramm Score (siehe Scheuerer und Hamill, 2015). In einer empirischen Analyse untersuchen wir, ob die verschiedenen Bewertungsregeln fehlspezifizierte räumliche Abhängigkeitsstrukturen auch bei unterschiedlichen Ensemble-Größen identifizieren können. Dazu verwenden wir Ensemble-Vorhersagen der Windgeschwindigkeit des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) und beobachtete Windgeschwindigkeiten an drei verschiedenen Standorten in Deutschland. Eine an die empirische Analyse anschließende Simulationsstudie erlaubt es, die Ergebnisse aus statistischer Sicht genauer zu diskutieren.

Das Kapitel ist folgendermaßen strukturiert: In Abschnitt 5.2 wird zunächst das verwendete Datenmaterial genauer beschrieben. Abschnitt 5.3 führt populäre statistische Nachbearbeitungsverfahren und Bewertungsregeln für probabilistische Vorhersagen ein. In der empirischen Analyse in Abschnitt 5.4 liegt der Fokus zunächst auf der Qualität von probabilistischen Vorhersagen für einzelne Standorte. Daraufhin vergleichen wir die Qualität von gemeinsamen Vorhersagen für mehrere Standorte mit Hilfe des Energy Scores und des Variogramm Scores. Die Sensitivität der verschiedenen Bewertungsregeln in Hinblick auf die Ensemble-Größe und unterschiedlich ausgeprägte räumliche Abhängigkeitsstrukturen wird in der Simulationsstudie in Abschnitt 5.5 detaillierter analysiert. Abschnitt 5.6 fasst die Ergebnisse zusammen und diskutiert Ideen für weiterführende Forschungsarbeiten.

5.2 Datengrundlage und deskriptive Analyse

Als Datengrundlage verwenden wir in diesem Kapitel Ensemble-Vorhersagen der Windgeschwindigkeit in 100 Metern Höhe des Europäischen Zentrums für mittelfristige Wettervorhersage sowie beobachtete Windgeschwindigkeiten (jeweils in m/s) an zwei Offshore-Forschungsplattformen (FINO2 und FINO3) in Deutschland sowie an einem Onshore-Standort (Wettermast Hamburg). Die Vorhersagen und Beobachtungen liegen dabei jeweils für den Zeitraum von Februar 2010 bis Juli 2014 vor. Die drei verschiedenen Standorte sind in Abbildung 5.1 visualisiert.

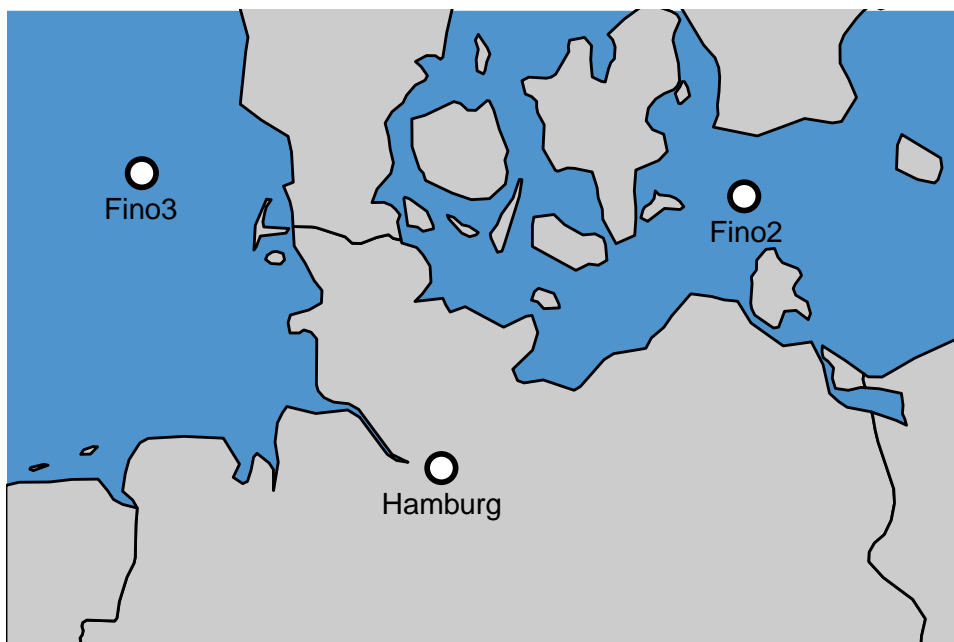


Abbildung 5.1: Lage der drei verschiedenen Standorte, für die Windgeschwindigkeitsmessungen und Vorhersagen vorliegen

Die Offshore-Plattform FINO2 liegt nördlich der Insel Rügen, während sich die Plattform FINO3 westlich von Sylt befindet. Die Distanz der beiden Offshore-Standorte beträgt ungefähr 382 Kilometer. Der Wettermast in Hamburg liegt im Vergleich etwa 258 Kilometer von FINO2 und rund 267 Kilometer von FINO3 entfernt.

Die Datensätze wurden bereits von Junk et al. (2014) verwendet und detailliert beschrieben. Die zugrundeliegenden Ensemble-Vorhersagen werden vom ECMWF ähnlich wie die deterministischen Vorhersagen zwei Mal am Tag um 00:00 Uhr UTC und um 12:00 Uhr UTC für verschiedene Vorhersagehorizonte (3 Std. - 120 Std.) ausgegeben. Das Ensemble-System besteht dabei wie in Kapitel 2 erläutert aus einer Kontrollvorhersage und 50 weiteren Ensemble-Mitgliedern, bei denen die Anfangsbedingungen bei der Initialisierung im Vergleich zur Kontrollvorhersage jeweils verändert werden. Die Kontrollvorhersage bezeichnen wir im Folgenden auch als deterministische Vorhersage. Um lokale Vorhersagen für die einzelnen Standorte zu erhalten, haben Junk et al. (2014) die ECMWF-Vorhersagen (die auf einem räumlichen Gitter ausgegeben werden) unter Verwendung der geographischen Koordinaten der Standorte interpoliert. Die Windmessungen sind laut den Autoren in einer 10-minütigen Auflösung verfügbar. Da der Wettermast in Hamburg jedoch nur Messungen in einer Höhe von 50 und 110 Metern bereitstellt, wurden die verfügbaren Messungen laut Junk et al. (2014) ebenfalls auf eine Höhe von 100 Metern linear interpoliert. Die Windgeschwindigkeiten am Standort FINO2 werden laut den Autoren in einer Höhe von 102 Metern gemessen und die Windgeschwindigkeiten am Standort FINO3 in einer Höhe von 100 Metern.

Bevor wir die statistische Methodik für die Bewertung und Nachverarbeitung von probabilistischen Windgeschwindigkeitsvorhersagen vorstellen, führen wir zunächst eine deskriptive Analyse der Daten durch. Diese erlaubt es uns, verschiedene Charakteristiken der gemessenen Windgeschwindigkeiten an den verschiedenen Standorten zu erkennen und die Qualität von probabilistischen bzw. deterministischen Vorhersagen für verschiedene Vorhersagehorizonte zu vergleichen. Abbildung 5.2 zeigt zunächst die Verteilung der gemessenen Windgeschwindigkeiten an den verschiedenen Standorten. Wie bereits von Junk et al. (2014) erwähnt, sind die Windgeschwindigkeiten an den beiden Offshore-Standorten FINO2 und FINO3 im Mittel deutlich höher als am Wettermast in Hamburg. Außerdem liegt an den beiden Offshore-Standorten eine deutlich höhere Varianz in den gemessenen Windgeschwindigkeiten vor.

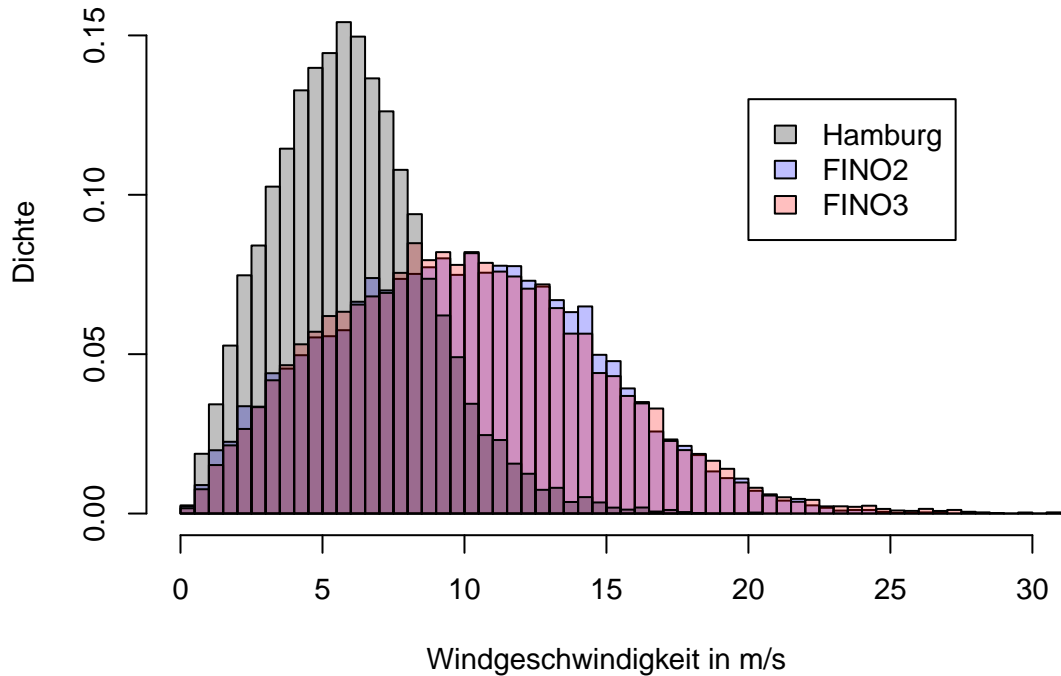


Abbildung 5.2: Histogramme der gemessenen Windgeschwindigkeiten an den Standorten Hamburg, FINO2 und FINO3

Die Qualität von deterministischen und probabilistischen Windgeschwindigkeitsvorhersagen für verschiedene Vorhersagehorizonte wird exemplarisch in Abbildung 5.3 für den Standort FINO2 illustriert. Dabei visualisieren wir die deterministische Kontrollvorhersage (blaue Linie) und die Ensemble-Vorhersagen (graue Linien), die jeweils am 25. Juni 2011 um 00:00 Uhr UTC ausgegeben wurden, gemeinsam mit den tatsächlich gemessenen Werten (rote Linie). Die Grafik illustriert sehr schön, dass die Varianz der Ensemble-Vorhersagen mit größerer Laufzeit ansteigt und dass sich die Abweichung zwischen der deterministischen Vorhersage und der Messung für längere Vorhersagehorizonte ebenfalls erhöht. Außerdem zeigt sich insbesondere für kurze Vorhersagehorizonte, dass die Varianz des Ensembles die zugrundeliegende Unsicherheit nicht adäquat repräsentiert. Für verschiedene Laufzeiten liegt die beobachtete Windgeschwindigkeit unterhalb oder oberhalb der Vorhersage aller Ensemble-Mitglieder. Ähnliche Ergebnisse ergeben sich für andere

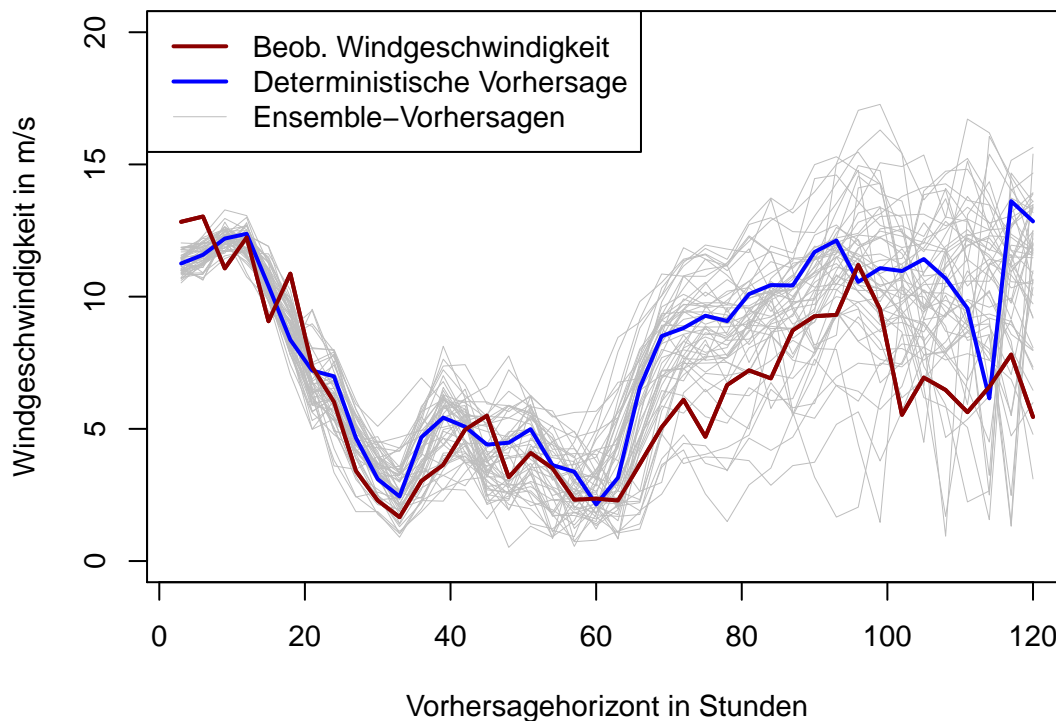


Abbildung 5.3: Deterministische und Ensemble-Windgeschwindigkeitsvorhersage für FINO2 ausgegeben am 25. Juni 2011 um 00:00 Uhr UTC für verschiedene Laufzeiten (3 Std. - 120 Std.) verglichen mit beobachteten Windgeschwindigkeiten

Zeitpunkte und die beiden weiteren Standorte im Datensatz. Eine Anwendung statistischer Nachverarbeitungsverfahren ist daher notwendig, um die Qualität der originalen Ensemble-Vorhersagen zu verbessern.

Der Fokus der Analyse liegt jedoch wie bereits in den vorherigen Kapiteln auf der Untersuchung von räumlichen Abhängigkeitsstrukturen. Hier soll in diesem Kapitel explizit betrachtet werden, wie sich räumliche Abhängigkeiten auf die Bewertung von multivariaten probabilistischen Windgeschwindigkeitsvorhersagen auswirken. Dazu ist es hilfreich sich zunächst einen ersten Einblick über die vorliegende räumliche Abhängigkeitsstruktur von deterministischen Vorhersagefehlern zu verschaffen. In Abbildung 5.4 stellen wir deshalb empirische Kreuzkorrelationen zum Lag 0 (berechnet wie in Definition 3.3 in Kapitel 3, S. 17) zwischen den Vorhersagefehlern der Kontrollvorhersage an den Standorten FINO2

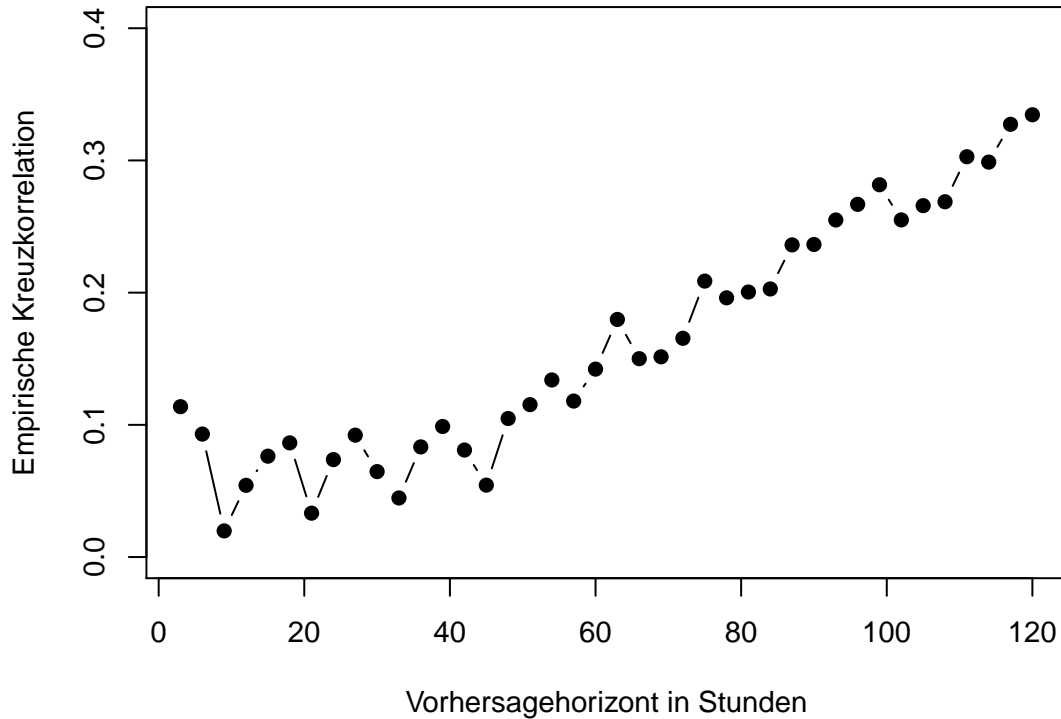


Abbildung 5.4: Empirische Kreuzkorrelationen (Lag 0) zwischen Vorhersagefehlern der Kontrollvorhersage an den Standorten FINO2 und FINO3 für verschiedene Vorhersagehorizonte

und FINO3 für unterschiedliche Vorhersagehorizonte dar. Wie bereits in den vorherigen Kapiteln stellt sich erneut heraus, dass die Kreuzkorrelationen zwischen den Vorhersagefehlerzeitreihen für längere Vorhersagehorizonte höher ausfallen. Es ist daher zu erwarten, dass Vorhersagefehler für längere Vorhersagehorizonte häufiger gleichzeitig an verschiedenen Standorten auftreten und zu einem hohen aggregierten Vorhersagefehler führen. Die verschiedenen Mitglieder einer Ensemble-Vorhersage sowie nachverarbeitete probabilistische Vorhersagen müssen daher das gemeinsame Verhalten für räumlich benachbarte Standorte auf geeignete Weise abbilden. Vorhersagen, welche die räumlichen Abhängigkeiten berücksichtigen, sollten speziell für längere Vorhersagehorizonte insgesamt besser bewertet werden als Vorhersagen, die räumliche Abhängigkeitsstrukturen nicht beachten. In Abschnitt 5.4 überprüfen wir, ob sich diese Vermutung bei den vorliegenden Daten auch

bei einem Vergleich von Ensemble-Vorhersagen mit unterschiedlichen Ensemble-Größen bestätigt. Die notwendige statistische Methodik für die Analyse wird im folgenden Abschnitt eingeführt.

5.3 Methodik

Wir betrachten in diesem Kapitel probabilistische Vorhersagen für eine stetige Zufallsvariable $Y(s, t)$ (Windgeschwindigkeit) am Standort s und zum Zeitpunkt t , die entweder in Form einer stetigen Wahrscheinlichkeitsverteilung $\mathcal{F}_{s,t}$ oder in Form von M Ensemble-Mitgliedern $X_1(s, t), \dots, X_M(s, t) \sim \mathcal{F}_{s,t}$ vorliegt. Weiterhin bezeichne $y(s, t)$ die tatsächlich beobachtete Windgeschwindigkeit am Standort s zum Zeitpunkt t . Multivariate probabilistische Vorhersagen für N verschiedene Standorte s_1, \dots, s_N zum Zeitpunkt t werden im Folgenden als $\mathcal{F}_{s_1, \dots, s_N, t}$ bezeichnet. Falls multivariate Vorhersagen in Form von M verschiedenen Ensemble-Mitgliedern verfügbar sind, werden diese als $\mathbf{X}_1(t), \dots, \mathbf{X}_M(t)$ zusammengefasst, wobei $\mathbf{X}_i(t) = [X_i(s_1, t), \dots, X_i(s_N, t)]^\top$ dem Vorhersagevektor für das i -te Ensemble-Mitglied entspricht. Realisierte Windgeschwindigkeiten für verschiedene Standorte an einem bestimmten Zeitpunkt t fassen wir zudem in einem Beobachtungsvektor $\mathbf{y}_t = [y(s_1, t), \dots, y(s_N, t)]^\top$ zusammen.

Im nächsten Unterabschnitt werden zunächst bekannte Bewertungsregeln eingeführt, mit denen die Qualität von univariaten und multivariaten probabilistischen Vorhersagen bewertet werden kann. Daraufhin stellen wir das EMOS-Verfahren (Ensemble Model Output Statistics) als eine statistische Nachverarbeitungsmethode für die Verbesserung von Ensemble-Vorhersagen vor. Wir verzichten auf eine Erläuterung und Verwendung des BMA-Verfahrens (Bayesian Model Averaging), da dieses im Vergleich zu EMOS sehr ähnliche Ergebnisse erzielt aber gleichzeitig längere Rechenzeiten erfordert. Abschließend wird in diesem Abschnitt das ECCQ-Verfahren (quantilsbasiertes Ensemble-Copula-Coupling) erläutert, mit dem verbesserte multivariate probabilistische Vorhersagen unter Berücksichtigung von räumlichen Abhängigkeitsstrukturen getroffen werden können.

5.3.1 Bewertungsregeln für stetige probabilistische Vorhersagen

Im univariaten Fall für einen einzelnen Standort s wird eine stetige probabilistische Vorhersage $\mathcal{F}_{s,t}$ bezüglich einer Beobachtung $y(s, t)$ häufig mit dem **Continuous Ranked Probability Score** bewertet. Dieser ist nach Gneiting und Raftery (2007) definiert als

$$\text{CRPS}(\mathcal{F}_{s,t}, y(s, t)) = \int_{-\infty}^{\infty} (\mathcal{F}_{s,t}(x) - \mathbf{1}(y(s, t) \leq x))^2 dx. \quad (5.1)$$

Falls das erste Moment einer Zufallsvariable mit Verteilung $\mathcal{F}_{s,t}$ existiert, kann der CRPS laut Gneiting und Raftery (2007) ebenfalls als

$$\text{CRPS}(\mathcal{F}_{s,t}, y(s, t)) = \mathbb{E}(|X^{(1)}(s, t) - y(s, t)|) - \frac{1}{2}\mathbb{E}(|X^{(1)}(s, t) - X^{(2)}(s, t)|) \quad (5.2)$$

ausgedrückt werden, wobei $X^{(1)}(s, t), X^{(2)}(s, t) \stackrel{\text{u.i.v.}}{\sim} \mathcal{F}_{s,t}$. In meteorologischen Anwendungen ist eine probabilistische Vorhersage jedoch häufig nur in Form von M Ensemble-Mitgliedern $X_1(s, t), \dots, X_M(s, t)$ verfügbar. Folgend den Ausführungen von Jordan et al. (2017) wird die theoretische Verteilungsfunktion $\mathcal{F}_{s,t}(x)$ dann in der Regel durch die empirische Verteilungsfunktion $\hat{\mathcal{F}}_{s,t}(x) = \frac{1}{M} \sum_{i=1}^M \mathbf{1}\{X_i(s, t) \leq x\}$ der Ensemble-Mitglieder approximiert. Der empirische CRPS bei einer Ensemble-Vorhersage ergibt sich dann als

$$\text{CRPS}(\hat{\mathcal{F}}_{s,t}, y(s, t)) = \frac{1}{M} \sum_{i=1}^M |X_i(s, t) - y(s, t)| - \frac{1}{2M^2} \sum_{i=1}^M \sum_{j=1}^M |X_i(s, t) - X_j(s, t)|. \quad (5.3)$$

Falls die einzelnen Ensemble-Mitglieder $X_1(s, t), \dots, X_M(s, t)$ als unabhängig identisch verteilte Zufallsvariablen mit Vorhersageverteilung $\mathcal{F}_{s,t}$ angesehen werden können, muss beachtet werden, dass der empirische CRPS aus Gleichung (5.3) den wahren CRPS aus Gleichung (5.2) verzerrt schätzt (siehe Ferro et al., 2008 sowie Zamo und Naveau, 2018). Dies liegt an der Tatsache, dass sich in der Doppelsumme die Einträge für $i = j$ zu Null ergeben, weshalb der hintere Erwartungswert in Gleichung (5.2) nicht mehr erwartungstreu geschätzt wird. Diese Tatsache kann insbesondere dann problematisch werden, wenn En-

sembles mit unterschiedlicher Ensemble-Größe verglichen werden. Die Verzerrung ist laut Zamo und Naveau (2018) insbesondere für eine kleine Anzahl an Ensemble-Mitgliedern M ausgeprägt. Als Alternative zum verzerrten Schätzer kann laut den Autoren jedoch der sogenannte Faire Continuous Ranked Probability Score (FCRPS) verwendet werden, der durch

$$\text{FCRPS}(\widehat{\mathcal{F}}_{s,t}, y(s, t)) = \frac{1}{M} \sum_{i=1}^M |X_i(s, t) - y(s, t)| - \frac{1}{2M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M |X_i(s, t) - X_j(s, t)| \quad (5.4)$$

gegeben ist. Hier werden die entsprechenden Einträge im Vorfaktor nicht berücksichtigt. Der FCRPS ist daher im Falle von unabhängig identisch verteilten Ensemble-Mitgliedern unverzerrt (siehe Zamo und Naveau, 2018).

In empirischen Studien ist man meist an der Qualität von probabilistischen Vorhersagen über einen gesamten Zeitraum interessiert. Falls eine Menge von Vorhersagen $\mathbf{F} = \{\mathcal{F}_{s,1}, \dots, \mathcal{F}_{s,T}\}$ sowie Beobachtungen $\mathbf{y} = [y(s, 1), \dots, y(s, T)]^\top$ für verschiedene Zeitpunkte $t = 1, \dots, T$ vorliegen, wird für gewöhnlich der mittlere Continuous Ranked Probability Score berechnet (siehe Bradley und Schwartz, 2011):

$$\overline{\text{CRPS}}(\mathbf{F}, \mathbf{y}) = \frac{1}{T} \sum_{t=1}^T \text{CRPS}(\mathcal{F}_{s,t}, y(s, t)). \quad (5.5)$$

Nach Bradley und Schwartz (2011) kann die Qualität zweier verschiedener probabilistischer Vorhersagen $\mathbf{F}^{(1)}$ und $\mathbf{F}^{(2)}$ über einen gesamten Zeitraum dann mit sogenannten Skill-Scores verglichen werden, wie zum Beispiel dem Continuous Ranked Probability Skill Score

$$\text{CRPSS}_{\mathbf{F}^{(1)}/\mathbf{F}^{(2)}} = 1 - \frac{\overline{\text{CRPS}}(\mathbf{F}^{(1)}, \mathbf{y})}{\overline{\text{CRPS}}(\mathbf{F}^{(2)}, \mathbf{y})}. \quad (5.6)$$

Ein positiver Skill-Score bedeutet, dass die Vorhersage $\mathbf{F}^{(1)}$ insgesamt besser bewertet

wird als die Vorhersage $\mathbf{F}^{(2)}$.

Anders als im univariaten Fall gibt es für die Bewertung von multivariaten probabilistischen Vorhersagen mindestens zwei Bewertungsregeln, die häufig in empirischen Studien verwendet werden. In vielen Fällen wird die Qualität einer multivariaten probabilistischen Vorhersage $\mathcal{F}_{s_1, \dots, s_N, t}$ für N verschiedene Standorte bezüglich eines Beobachtungsvektors $\mathbf{y}_t = [y(s_1, t), \dots, y(s_N, t)]^\top$ mit dem Energy Score (ES) bewertet. Der Energy Score entspricht einer Generalisierung des CRPS für den multivariaten Fall und ist laut Gneiting und Raftery (2007) definiert als

$$\text{ES}(\mathcal{F}_{s_1, \dots, s_N, t}, \mathbf{y}_t) = \mathbb{E} \left(\|\mathbf{X}^{(1)}(t) - \mathbf{y}_t\| \right) - \frac{1}{2} \mathbb{E} \left(\|\mathbf{X}^{(1)}(t) - \mathbf{X}^{(2)}(t)\| \right), \quad (5.7)$$

wobei $\mathbf{X}^{(1)}(t), \mathbf{X}^{(2)}(t) \stackrel{\text{u.i.v.}}{\sim} \mathcal{F}_{s_1, \dots, s_N, t}$. Dabei steht $\|\cdot\|$ für die Euklidische Norm.

Falls eine multivariate probabilistische Vorhersage durch ein Ensemble $\mathbf{X}_1(t), \dots, \mathbf{X}_M(t)$ mit $\mathbf{X}_i(t) = [X_i(s_1, t), \dots, X_i(s_N, t)]^\top$ gegeben ist, wird für gewöhnlich der empirische Energy Score

$$\text{ES}(\hat{\mathcal{F}}_{s_1, \dots, s_N, t}, \mathbf{y}_t) = \frac{1}{M} \sum_{i=1}^M \|\mathbf{X}_i(t) - \mathbf{y}_t\| - \frac{1}{2M^2} \sum_{i=1}^M \sum_{j=1}^M \|\mathbf{X}_i(t) - \mathbf{X}_j(t)\| \quad (5.8)$$

für die Bewertung der Vorhersagequalität verwendet (siehe Jordan et al., 2017). Ähnlich wie beim empirischen CRPS ist der empirische Energy Score jedoch verzerrt, wenn die Ensemble-Mitglieder unabhängig und identisch verteilte Zufallsvektoren mit multivariater Verteilung $\mathcal{F}_{s_1, \dots, s_N, t}$ darstellen. Für den Vergleich von Ensembles mit unterschiedlichen Ensemble-Größen kann stattdessen der Faire Energy Score

$$\text{FES}(\hat{\mathcal{F}}_{s_1, \dots, s_N, t}, \mathbf{y}_t) = \frac{1}{M} \sum_{i=1}^M \|\mathbf{X}_i(t) - \mathbf{y}_t\| - \frac{1}{2M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \|\mathbf{X}_i(t) - \mathbf{X}_j(t)\| \quad (5.9)$$

verwendet werden, der unter anderem in ähnlicher Form von Wilks (2019, S. 444) benannt wird.

In jüngeren Studien (siehe Scheuerer und Hamill, 2015) wird häufig darauf hingewiesen, dass der Energy Score oftmals nicht sensitiv auf fehlspezifizierte Korrelationsstrukturen reagiert. Um dieses Problem anzugehen, schlagen Scheuerer und Hamill (2015) deshalb den Variogramm Score der Ordnung p vor, der sich in einer ungewichteten Version für ein Ensemble mit M Mitgliedern mit

$$\text{VS}^p(\hat{\mathcal{F}}_{s_1, \dots, s_N, t}, \mathbf{y}_t) = \sum_{i=1}^N \sum_{j=1}^N \left(|y(s_i, t) - y(s_j, t)|^p - \frac{1}{M} \sum_{k=1}^M |X_k(s_i, t) - X_k(s_j, t)|^p \right)^2 \quad (5.10)$$

ergibt. Laut Jordan et al. (2017) wird der Wert von p meist als $p = 0.5$ oder $p = 1$ gewählt. In unserer Analyse betrachten wir beispielhaft den Variogramm Score der Ordnung $p = 0.5$.

Ähnlich wie für den univariaten CRPS kann die Qualität zweier verschiedener multivariater Vorhersagen $\mathbf{F}^{(1)}$ und $\mathbf{F}^{(2)}$ über einen gesamten Zeitraum mit Skill-Scores verglichen werden. Diese bezeichnen wir im Folgenden als $\text{ESS}_{\mathbf{F}^{(1)}/\mathbf{F}^{(2)}}$ (Energy Skill Score), $\text{FESS}_{\mathbf{F}^{(1)}/\mathbf{F}^{(2)}}$ (Fairer Energy Skill Score) bzw. $\text{VSS}_{\mathbf{F}^{(1)}/\mathbf{F}^{(2)}}^{0.5}$ (Variogramm Skill Score).

5.3.2 Ensemble Model Output Statistics

Nachdem wir verschiedene Bewertungsregeln für probabilistische Vorhersagen eingeführt haben, kann nun ein statistisches Nachverarbeitungsverfahren beschrieben werden, mit dem die Qualität der originalen Ensemble-Vorhersagen verbessert werden kann. Über die letzten Jahre wurden dazu verschiedene Methoden entwickelt und in empirischen Studien verglichen, wie zum Beispiel Bayesian Model Averaging (BMA, siehe Sloughter et al., 2010) oder Ensemble Model Output Statistics (EMOS, siehe Gneiting et al., 2005). Das EMOS-Verfahren hat dabei den Vorteil, dass es eine recht einfache Modellstruktur aufweist und dass nur eine geringe Anzahl an Parametern geschätzt werden muss. Außerdem kann das Verfahren für verschiedene Zielgrößen wie etwa der Bodentemperatur,

für Niederschlagsmengen oder für Windgeschwindigkeiten verwendet werden. Dazu muss im Vorfeld zunächst eine passende Verteilungsfamilie gewählt werden, die für Vorhersagen der jeweiligen Zielgröße geeignet erscheint. Für die probabilistische Vorhersage für einen bestimmten Zeitpunkt werden die Parameter der ausgewählten Verteilungsfamilie dann in Abhängigkeit der originalen Ensemble-Vorhersagen und zusätzlichen Hilfsparametern ausgedrückt. Diese Hilfsparameter werden auf Basis von zurückliegenden Beobachtungen und Ensemble-Vorhersagen so geschätzt, dass der mittlere CRPS für vergangene Ensemble-Vorhersagen und Beobachtungen minimiert wird.

Speziell für Windgeschwindigkeitsvorhersagen oder andere nicht-negative Zielgrößen schlagen Thorarinsdottir und Gneiting (2010) einen EMOS-Ansatz basierend auf einer bei Null trunkeierten Normalverteilung vor: Die Methode benutzt dabei direkt die Informationen einer Ensemble-Vorhersage $X_1(s, t), \dots, X_M(s, t)$ für eine stetige und nicht-negative Zielvariable $Y(s, t)$. Weiterhin werden beobachtete Werte $y(s, 1), \dots, y(s, t-1)$ und Ensemble-Vorhersagen für zurückliegende Zeitpunkte berücksichtigt. Für einen bestimmten räumlichen Standort s schlagen die Autoren zum Zeitpunkt t eine probabilistische Vorhersage der Form

$$\hat{Y}(s, t) \sim \mathcal{N}^0[\mu_{s,t}, \sigma_{s,t}^2], \quad (5.11)$$

vor, wobei die Parameter $\mu_{s,t} = a + b \cdot (X_1(s, t) + \dots + X_M(s, t))$ und $\sigma_{s,t}^2 = c + d \cdot S_{s,t}^2$ mit $S_{s,t}^2 = \frac{1}{M} \sum_{k=1}^M (X_k(s, t) - \bar{X}_{s,t})^2$ von den einzelnen Ensemble-Mitgliedern abhängen. Die Parameter a, b, c und d werden durch Minimierung von

$$\frac{1}{\ell} \sum_{j=t-\ell}^{t-1} \text{CRPS} \left(\mathcal{N}^0[\mu_{s,j}, \sigma_{s,j}^2], y(s, j) \right) \quad (5.12)$$

für eine im Vorfeld spezifizierte Anzahl von ℓ vergangenen Zeitpunkten geschätzt. Dabei wird die Nebenbedingung berücksichtigt, dass die Parameter b, c und d größer als 0 sein müssen (siehe Thorarinsdottir und Gneiting, 2010).

5.3.3 Ensemble Copula Coupling

Wenn klassische Nachverarbeitungsverfahren wie EMOS separat für einzelne Standorte angewandt werden, enthalten die nachverarbeiteten probabilistischen Vorhersagen keine Informationen über die räumliche Abhängigkeitsstruktur zwischen den Beobachtungen an verschiedenen Standorten und sind daher nicht für eine gemeinsame Vorhersage nützlich. Mit Hilfe des von Schefzik et al. (2013) vorgeschlagenen quantilsbasierten Ensemble-Copula-Coupling-Verfahrens (ECCQ) können diese separaten EMOS-Vorhersagen jedoch mit wertvollen Informationen über die räumliche Abhängigkeitsstruktur kombiniert werden, die direkt aus den originalen Ensemble-Vorhersagen entnommen werden können.

Für jeden Zeitpunkt t schlagen die Autoren vor, zunächst aufsteigende Ränge der einzelnen (originalen) Ensemble-Mitglieder $\zeta_s(m) = \text{rank}(X_m(s, t))$ für $m = 1, \dots, M$ separat für alle Standorte $s = s_1, \dots, s_N$ zu bestimmen. Daraufhin wird für alle Standorte ein quantilsbasiertes Ensemble der Größe M (originale Ensemble-Größe) mit

$$\tilde{X}_1(s, t) = \mathcal{F}_{s,t}^{-1}\left(\frac{1}{M+1}\right), \dots, \tilde{X}_M(s, t) = \mathcal{F}_{s,t}^{-1}\left(\frac{M}{M+1}\right)$$

bestimmt, wobei $\mathcal{F}_{s,t}^{-1}(\cdot)$ der Quantilsfunktion der angepassten EMOS-Verteilung am Standort s zum Zeitpunkt t entspricht. Abschließend wird die ECCQ-Ensemble-Vorhersage für einen bestimmten Standort s generiert, indem die Quantile in der Rangordnung des originalen Ensembles angeordnet werden:

$$\hat{X}_1(s, t) = \tilde{X}_{\zeta_s(1)}(s, t), \dots, \hat{X}_M(s, t) = \tilde{X}_{\zeta_s(M)}(s, t).$$

Durch die Verwendung der Rangordnungen des originalen Ensembles wird die darin vorhandene räumliche Abhängigkeitsstruktur somit direkt auf die Quantile der EMOS-Verteilungen übertragen. Die ECCQ-Vorhersage kombiniert daher die potenziell bessere Vorhersagequalität der EMOS-Vorhersage mit den Informationen über die räumliche Abhängigkeitsstruktur aus dem originalen Ensemble.

5.4 Ergebnisse

Die deskriptive Analyse in Abschnitt 5.2 hat gezeigt, dass eine statistische Nachverarbeitung der originalen Ensemble-Vorhersagen notwendig ist, um eine zuverlässige Einschätzung der Vorhersageunsicherheit zu gewährleisten. Weiterhin hat die Analyse von Kreuzkorrelationen deterministischer Vorhersagefehler verdeutlicht, dass räumliche Abhängigkeitsstrukturen insbesondere für längere Vorhersagehorizonte ausgeprägt sind. Diese Abhängigkeiten müssen daher für sinnvolle gemeinsame Vorhersagen für mehrere Standorte berücksichtigt werden. Im Folgenden wenden wir zunächst das EMOS-Verfahren separat für einzelne Standorte an, und vergleichen die Qualität der nachverarbeiteten Vorhersagen mit der des originalen Ensembles. Daraufhin fokussieren wir uns auf die Generierung von multivariaten probabilistischen Vorhersagen für mehrere Standorte und deren Bewertung. Hier untersuchen wir insbesondere, ob die verschiedenen multivariaten Bewertungsregeln aus Abschnitt 5.3.1 fehlspezifizierte Abhängigkeitsstrukturen auch im Falle von Ensemble-Vorhersagen mit unterschiedlichen Ensemble-Größen zuverlässig identifizieren können.

5.4.1 Nachverarbeitung von Ensemble-Vorhersagen für einzelne Standorte

Wir wenden zunächst das in Abschnitt 5.3.2 vorgestellte EMOS-Verfahren separat für die einzelnen Standorte (Hamburg, FINO2 und FINO3) und für die einzelnen Vorhersagehorizonte (3 Std. - 120 Std.) an. Dabei wählen wir eine gleitende Periode von 200 vorausgegangenen Zeitpunkten (100 Tage) für die Anpassung von trunkierten Normalverteilungen gemäß (5.11) bzw. (5.12) auf Basis der Ensemble-Vorhersagen und zurückliegender Beobachtungen. Für die Bewertung der Vorhersagequalität ziehen wir für jeden Zeitpunkt 10 000 Zufallszahlen aus der angepassten Verteilung (diese sollten die EMOS-Verteilung gut approximieren) und berechnen den Continuous Ranked Probability Score wie in Gleichung (5.3). Den CRPS berechnen wir ebenfalls für die originale Ensemble-Vorhersage für

die entsprechenden Zeitpunkte. Daraufhin ermitteln wir für alle Vorhersagehorizonte den Continuous Ranked Probability Skill Score (CRPSS) über den gesamten Zeitraum für die EMOS-Vorhersagen verglichen mit dem originalen Ensemble. Die Anwendung des EMOS-Verfahrens bewirkt bei allen drei Standorten und über alle Vorhersagehorizonte hinweg eine Verbesserung der Vorhersagequalität hinsichtlich des CRPS. In Abbildung 5.5 sind beispielhaft die Skill-Scores für die Offshore-Plattform FINO2 für verschiedene Vorhersagehorizonte visualisiert. Wie zu sehen ist, verbessert sich die Vorhersagequalität mit dem EMOS-Verfahren insbesondere für sehr kurze Vorhersagehorizonte verglichen mit den originalen Ensemble-Vorhersagen. Dieses Ergebnis ist nicht überraschend, da bereits in der deskriptiven Analyse sichtbar wurde, dass die realisierten Windgeschwindigkeiten insbesondere für kurze Vorhersagehorizonte häufig höher oder niedriger ausfallen als von allen

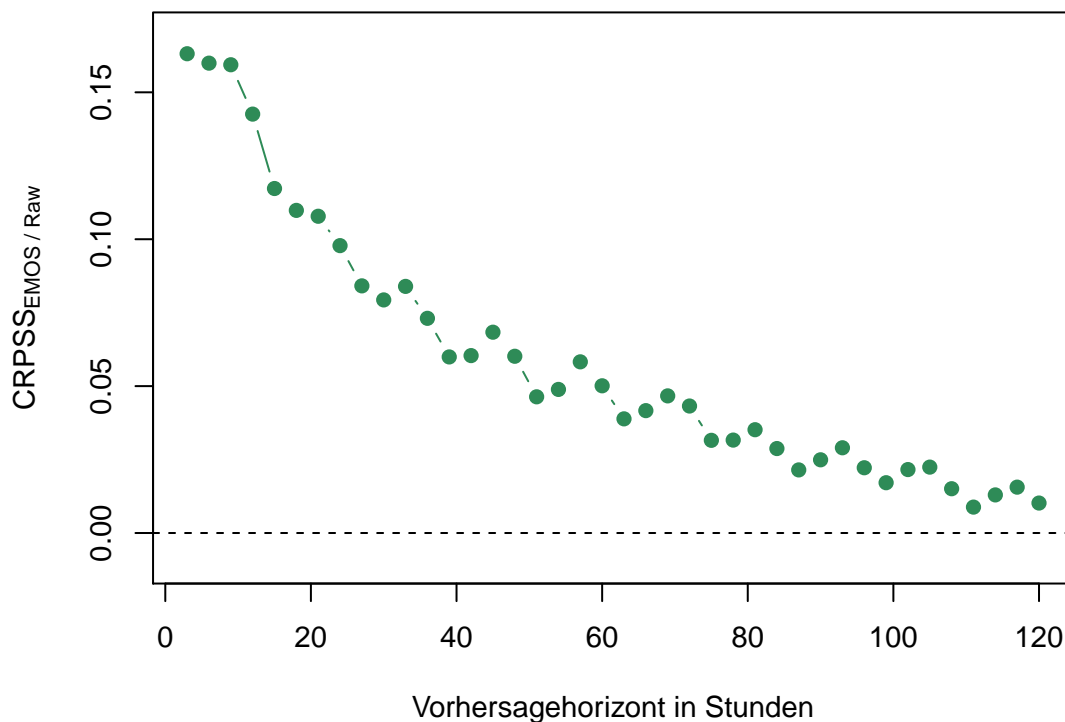


Abbildung 5.5: Continuous Ranked Probability Skill Score (CRPSS) für EMOS-Vorhersagen verglichen mit dem originalen (raw) Ensemble für die Station FINO2

Ensemble-Mitgliedern prognostiziert. Nach der Anwendung des EMOS-Verfahrens sind die nachverarbeiteten Prognosen besser kalibriert als die originalen Ensemble-Vorhersagen, da die vorliegende Unsicherheit in einer adäquateren Weise repräsentiert wird. Insgesamt bestätigen die Ergebnisse die Erkenntnisse von Junk et al. (2014) und demonstrieren, dass eine univariate Nachverarbeitung von Ensemble-Vorhersagen sinnvoll ist, um die Vorhersagequalität von probabilistischen Windgeschwindigkeitsvorhersagen zu verbessern.

5.4.2 Bewertung von gemeinsamen probabilistischen Vorhersagen für mehrere Standorte

Der Fokus unserer Analyse liegt jedoch nicht auf der univariaten Nachverarbeitung von Ensemble-Vorhersagen, sondern auf der Bewertung von gemeinsamen probabilistischen Windgeschwindigkeitsvorhersagen für die Standorte FINO2, FINO3 und Hamburg. Da sich die empirischen Windgeschwindigkeitsverteilungen an den beiden Offshore-Standorten FINO2 und FINO3 wie in Abbildung 5.2 gesehen stark von der Verteilung am Wettermast in Hamburg unterscheiden, fokussieren wir uns zunächst auf gemeinsame Vorhersagen für die beiden Offshore-Standorte.

Unter Verwendung der bekanntesten multivariaten Bewertungsregeln (Energy Score, Fairer Energy Score, Variogramm Score) vergleichen wir in Abbildung 5.6 zunächst die Qualität der für die einzelnen Standorte separat angepassten EMOS-Vorhersagen mit der Qualität der originalen Ensemble-Vorhersagen. Dabei werden die EMOS-Vorhersagen erneut für jeden Zeitpunkt durch 10 000 Zufallszahlen aus den entsprechenden univariaten EMOS-Verteilungen repräsentiert. Bei Betrachtung der Energy Skill Scores fällt auf, dass die EMOS-Vorhersagen für fast alle Vorhersagehorizonte im Mittel besser bewertet werden als die originalen Ensemble-Vorhersagen. Die Qualität der EMOS-Vorhersagen scheint nur für sehr große Vorhersagehorizonte ähnlich (Energy Skill Score) bzw. schlechter (Fairer Energy Skill Score) als beim originalen Ensemble zu sein. Die Differenzen zwischen dem Energy Score und dem Fairen Energy Score ergeben sich allein aus der Tatsache, dass

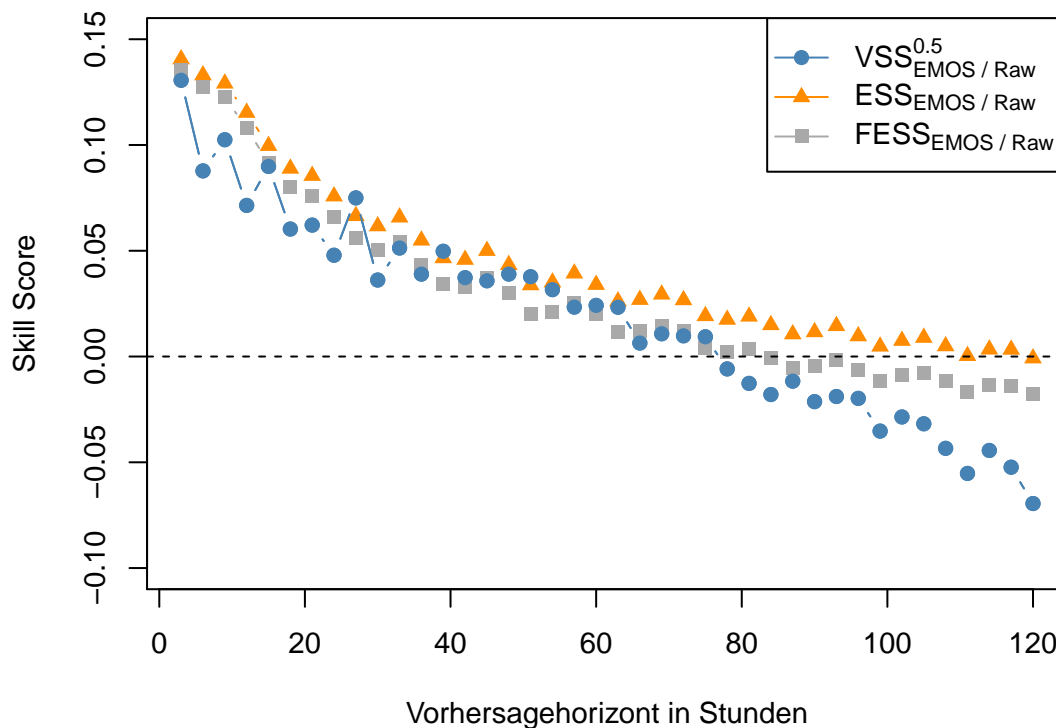


Abbildung 5.6: Multivariate Skill-Scores für EMOS-Vorhersagen verglichen mit dem originalen (raw) Ensemble (gemeinsame Bewertung für die Standorte FINO2 und FINO3)

sich die Ensemble-Größen bei der EMOS-Vorhersage (10 000 Ensemble-Mitglieder) und beim originalen Ensemble (51 Ensemble-Mitglieder) unterscheiden. Der Variogramm Score deutet insbesondere für längere Vorhersagehorizonte auf eine schlechtere Vorhersagequalität der separaten EMOS-Vorhersagen im Vergleich zum originalen Ensemble hin. Dieses Ergebnis ist nach Betrachtung der empirischen Kreuzkorrelationen zwischen Vorhersagefehlern an den Standorten FINO2 und FINO3 in Abbildung 5.3 nicht überraschend, da räumliche Abhängigkeitsstrukturen insbesondere für längere Vorhersagehorizonte relevant sind. Da die separaten EMOS-Vorhersagen keine Informationen über die räumliche Abhängigkeitsstruktur der Beobachtungen berücksichtigen, sollte die Qualität der Vorhersagen im Vergleich zum originalen Ensemble vor allem für lange Vorhersagehorizonte abnehmen. Dieses Phänomen wird insbesondere vom Variogramm Score identifiziert. Der Vergleich der Skill-Scores deutet also daraufhin, dass der Energy Score weniger stark auf

fehlspezifizierte Abhängigkeitsstrukturen reagiert als der Variogramm Score.

Wie bereits angemerkt, berücksichtigen die separaten EMOS-Vorhersagen nicht die räumliche Abhängigkeitsstruktur der Beobachtungen an den verschiedenen Standorten. Um die Informationen über die räumliche Abhängigkeitsstruktur aus dem ECMWF-Ensemble zurückzugewinnen, wenden wir deshalb die Methode des quantilsbasierten Ensemble-Copula-Couplings (ECCQ) an. In Abbildung 5.7 wird die Qualität der ECCQ-Vorhersagen mit der Qualität der EMOS-Vorhersagen erneut unter Verwendung der verschiedenen Skill Scores verglichen. Wie erwartet schlägt das ECCQ-Verfahren die EMOS-Vorhersage hinsichtlich des Variogramm Skill Scores insbesondere für längere Vorhersagehorizonte. Für kurze Vorhersagehorizonte (also für weniger ausgeprägte Abhängigkeiten) stellt sich hingegen heraus, dass die EMOS-Vorhersagen im Mittel einen besseren Variogramm Score

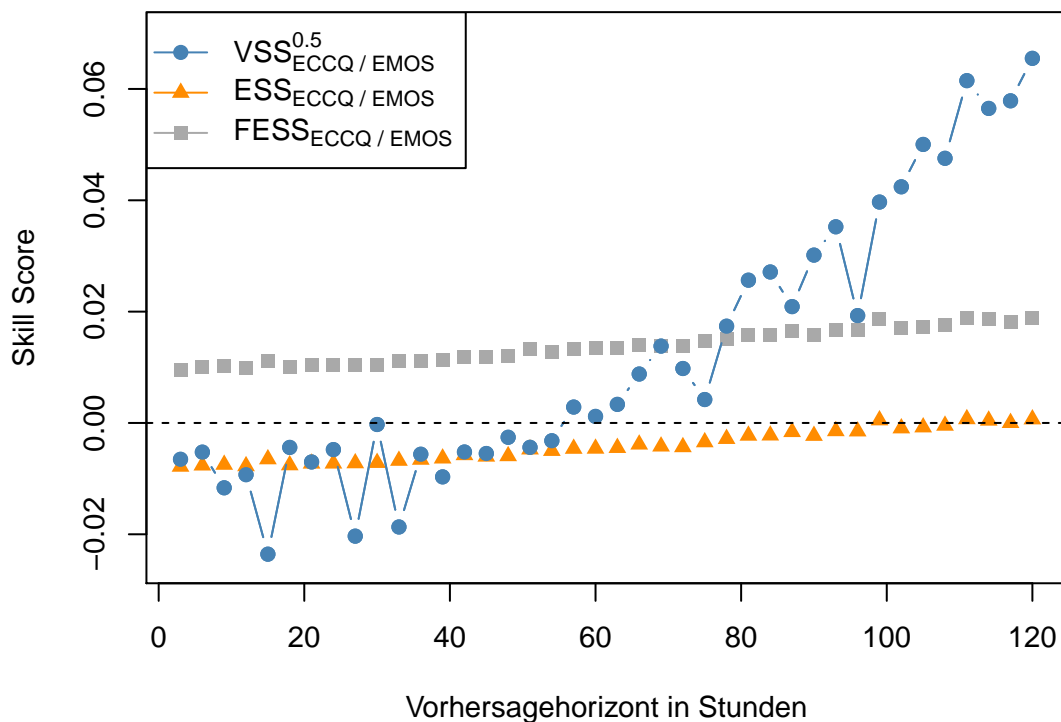


Abbildung 5.7: Multivariate Skill-Scores für ECCQ-Vorhersagen verglichen mit EMOS-Vorhersagen (gemeinsame Bewertung für die Standorte FINO2 und FINO3)

erzielen als die ECCQ-Vorhersagen. Bezüglich des Energy Skill Scores erzielen die ECCQ-Vorhersagen nur für sehr lange Vorhersagehorizonte eine ähnliche Vorhersagequalität wie die EMOS-Vorhersagen. Der Faire Energy Skill Score zeigt hingegen eine verbesserte Qualität im Vergleich zu EMOS für alle betrachteten Vorhersagehorizonte auf. Diese Ergebnisse sind auf den ersten Blick überraschend, können aber durch die unterschiedlichen Strukturen der ECCQ- und der EMOS-Vorhersagen erklärt werden. Während für die EMOS-Vorhersagen 10 000 Zufallszahlen aus den angepassten trunkierten Normalverteilungen gezogen werden, werden für die ECCQ-Vorhersagen systematisch 51 äquidistante Quantile für die einzelnen Standorte ausgewählt. Ähnlich wie von Zamo und Naveau (2018) für den Fairen CRPS gezeigt, ist der Faire Energy Score nur dann ein unverzerrter Schätzer für den wahren Energy Score, wenn die Ensemble-Vorhersage als eine unabhängig identisch verteilte Zufallsstichprobe aus der wahren Vorhersageverteilung angesehen werden kann. Die Wahl von 51 äquidistanten Quantilen repräsentiert hingegen keine Zufallsstichprobe, weshalb der Faire Energy Score in diesem Falle verzerrt ist. Diese Problematik wird in der Sensitivitätsanalyse in Abschnitt 5.5 detaillierter untersucht.

Insgesamt zeigen die Ergebnisse deutliche Unterschiede zwischen den verwendeten multivariaten Bewertungsregeln auf. Wenn die Vorhersagen für den Wettermast in Hamburg ebenfalls bei der Analyse berücksichtigt werden, ergeben sich ähnliche Ergebnisse wie bei der Analyse für die beiden Offshore-Standorte. Wie in Abbildung 5.8 zu sehen ist, zeigt der Variogramm Skill Score immer noch für lange Vorhersagehorizonte eine verbesserte Vorhersagequalität der ECCQ-Vorhersagen im Vergleich zu den EMOS-Vorhersagen. Die EMOS-Vorhersagen erzielen jedoch für fast alle Vorhersagehorizonte einen im Mittel besseren Energy Score als die ECCQ-Vorhersagen. Dies liegt insbesondere daran, dass hier die unterschiedlichen Ensemble-Größen nicht adäquat berücksichtigt werden. Beim Fairen Energy Score zeigt sich erneut eine verbesserte Vorhersagequalität für alle Vorhersagehorizonte, was wiederum durch die unterschiedlichen Strukturen der Ensembles begründet werden kann.

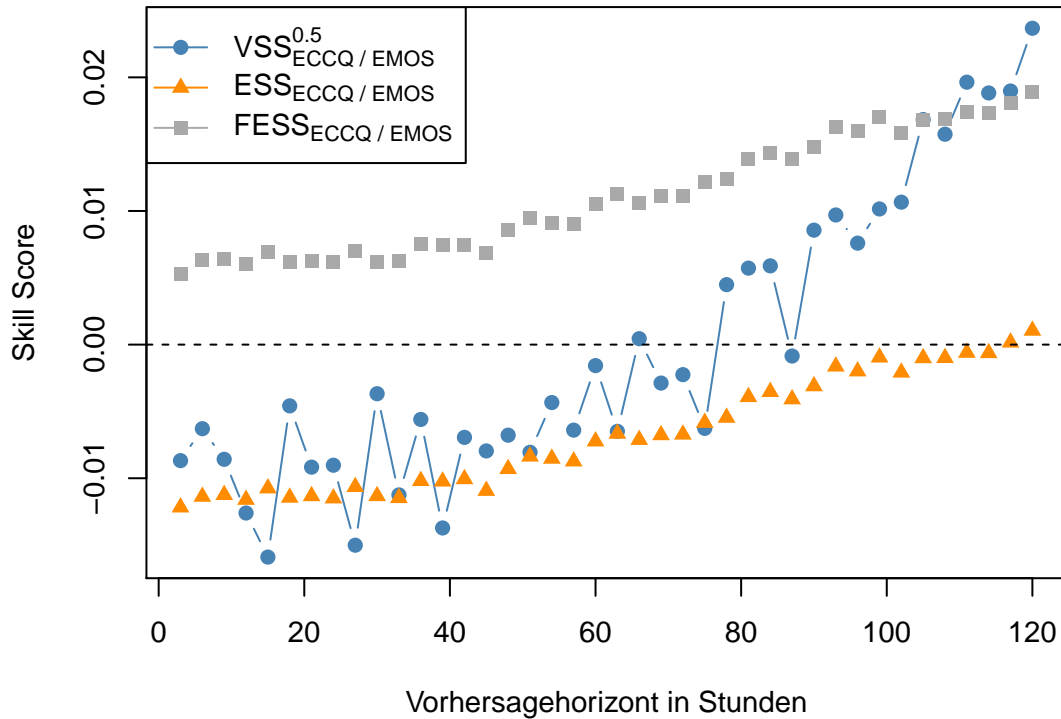


Abbildung 5.8: Multivariate Skill-Scores für ECCQ-Vorhersagen verglichen mit EMOS-Vorhersagen (gemeinsame Bewertung für die Standorte FINO2, FINO3 und Hamburg)

Die Ergebnisse werfen die Frage auf, ob die verschiedenen Bewertungsregeln tatsächlich in der Lage sind, fehlspezifizierte räumliche Abhängigkeitsstrukturen zuverlässig zu identifizieren, die in vielen empirischen Studien von hoher Relevanz sind. Insbesondere bei unterschiedlichen Ensemble-Größen oder unterschiedlichen Ensemble-Strukturen kann dies problematisch werden. Im folgenden Abschnitt versuchen wir deshalb mit Hilfe einer Sensitivitätsanalyse genauere Erklärungen für die Ergebnisse aus der empirischen Analyse zu finden. Dazu führen wir Simulationsstudien für die unterschiedlichen Bewertungsregeln durch, mit denen der Einfluss der Ensemble-Größe M auf die Bewertung von ECCQ-artigen Vorhersagen im Vergleich zu EMOS-artigen Vorhersagen für verschiedene räumliche Abhängigkeitsstrukturen analysiert werden kann.

5.5 Sensitivitätsanalyse für multivariate Bewertungsregeln

In der Sensitivitätsanalyse führen wir mehrere Monte-Carlo-Simulationen mit jeweils 100 000 Replikationen durch. In jeder Replikation wird zunächst ein Beobachtungsvektor $y = [y_1, y_2]^\top$ (z.B. Windgeschwindigkeiten an zwei Standorten) generiert. Dafür legen wir einen Zufallsvektor $Y = [Y_1, Y_2]^\top$ zugrunde, wobei $Y_1 \sim \mathcal{N}^0[10, 4]$ und $Y_2 \sim \mathcal{N}^0[11, 4]$ (bei Null) trunziert normalverteilte Zufallsvariablen mit unterschiedlichen Parametern darstellen. Die Abhängigkeitsstruktur zwischen Y_1 und Y_2 spezifizieren wir außerdem durch eine Gauß-Copula mit Korrelationsparameter ρ , der für die verschiedenen Simulationen variiert. Mit Hilfe der Copula und der Randverteilungen können entsprechende Beobachtungen in jeder Replikation simuliert werden. Außerdem simulieren wir eine EMOS-artige probabilistische Vorhersage indem wir 10 000 Zufallszahlen jeweils aus den beiden (wahren) Randverteilungen $\mathcal{N}^0[10, 4]$ und $\mathcal{N}^0[11, 4]$ ziehen, ohne jedoch die zugrundeliegende Abhängigkeitsstruktur zu berücksichtigen. Abschließend generieren wir eine ECCQ-artige Vorhersage. Dazu werden zunächst M Vektoren zufällig aus der zugrundeliegenden bivariaten Verteilung simuliert. Diese stellen eine fiktive Ensemble-Vorhersage dar, in der die Abhängigkeitsstruktur über die Gauß-Copula berücksichtigt ist. Daraufhin wenden wir das Ensemble-Coupling-Verfahren an, wofür wir zunächst jeweils M äquidistante Quantile der wahren Randverteilungen auswählen. Diese werden dann entsprechend der Rangstruktur des fiktiven Ensembles umgeordnet, sodass die simulierte Abhängigkeitsstruktur auf die Quantile übertragen wird. In jeder Replikation werden daraufhin der Energy Score, der Faire Energy Score und der Variogramm Score der Ordnung $p = 0.5$ berechnet. Abschließend werden Skill-Scores für die einzelnen Bewertungsregeln berechnet, wobei jeweils das arithmetische Mittel der Scores über alle 100 000 Replikationen verwendet wird. Die Simulationsstudien führen wir für unterschiedliche Abhängigkeitsstrukturen (kontrolliert durch den Parameter ρ der Gauß-Copula) und für verschiedene Ensemble-Größen M durch.

In Abbildung 5.9 werden zunächst die Skill-Scores der ECCQ-artigen Vorhersagen im Vergleich zu EMOS-artigen Vorhersagen für variierende Abhängigkeitsstrukturen visualisiert. Dabei ist die Ensemble-Größe auf eine Anzahl von $M = 51$ Ensemble-Mitgliedern (also ähnlich wie beim ECMWF-Ensemble) festgelegt. Der Parameter ρ der Gauß-Copula variiert hingegen von $\rho = 0$ bis $\rho = 0.5$. Für alle verschiedenen Parametereinstellungen wird jeweils eine separate Monte-Carlo-Simulation mit 100 000 Wiederholungen durchgeführt. Die simulierten Skill-Scores bestätigen die Ergebnisse, die in der empirischen Analyse in Abschnitt 5.4 festgestellt wurden. Bezüglich des Variogramm Scores erzielt das

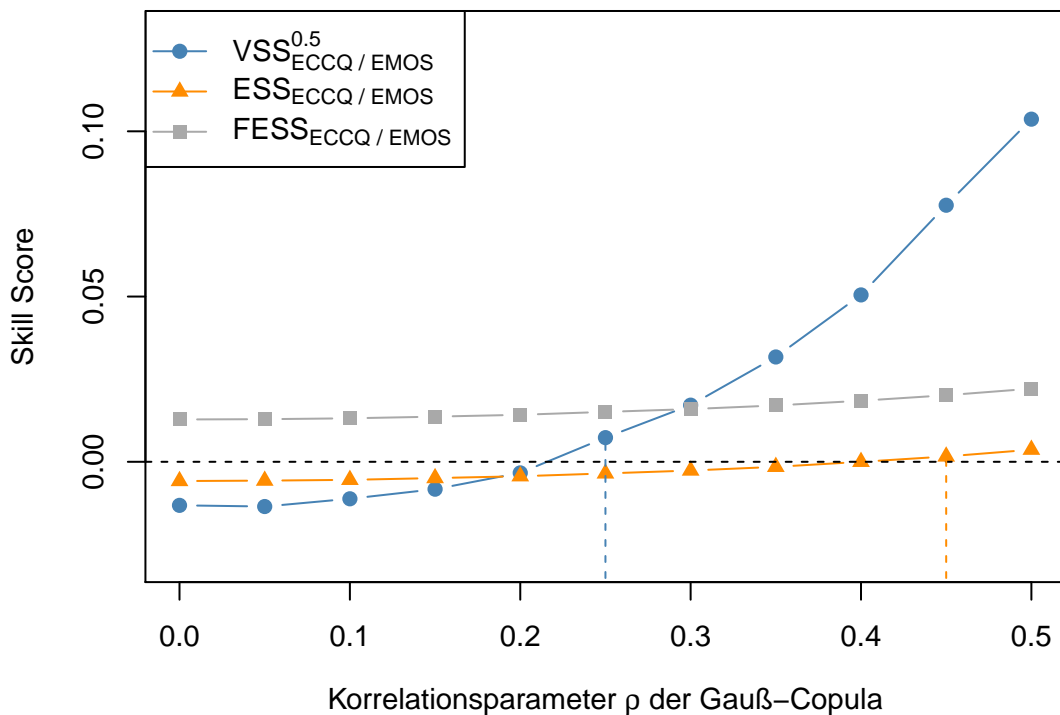


Abbildung 5.9: Simulierte Skill-Scores für bivariate ECCQ-artige Vorhersagen mit einer festgewählten Anzahl von $M = 51$ Ensemble-Mitgliedern verglichen mit EMOS-artigen Vorhersagen (bestehend aus 10 000 Ensemble-Mitgliedern). In der Simulation wird die wahre Abhängigkeitsstruktur zwischen zwei trunziert normalverteilten Zufallsvariablen durch eine Gauß-Copula mit Korrelationsparameter ρ spezifiziert. Für jeden Wert von ρ wird eine Monte-Carlo-Simulation mit 100 000 Replikationen durchgeführt. Die Skill-Scores werden daraufhin basierend auf gemittelten Scores über alle Replikationen berechnet.

ECCQ-Ensemble eine bessere Vorhersagequalität als die EMOS-artige Vorhersage (mit einer größeren Anzahl an Ensemble-Mitgliedern) für Werte von ρ größer oder gleich 0.25. Ein Korrelationsparameter größer oder gleich 0.45 und damit eine ausgeprägtere Abhängigkeitsstruktur ist hingegen für einen verbesserten mittleren Energy Score notwendig. Der Faire Energy Score zeigt hingegen für alle betrachteten Korrelationsparameter einen im Mittel besseren Wert für die ECCQ-artige Vorhersage als für die EMOS-artige Vorhersage (sogar für $\rho = 0$). Dieses Ergebnis kann durch die unterschiedlichen Strukturen der beiden Vorhersagen erklärt werden. Für den Fairen CRPS zeigen bereits Zamo und Naveau (2018), dass Ensembles bestehend aus äquidistanten Quantilen der zugrundeliegenden Wahrscheinlichkeitsverteilung im Mittel besser abschneiden als Ensembles, die eine reine Zufallsstichprobe aus der entsprechenden Verteilung darstellen. Die Simulationsstudie für den Fairen Energy Score zeigt, dass dieses Ergebnis auch im multivariaten Fall Bestand hält.

Der Einfluss der Ensemble-Größe M auf die Bewertung der verschiedenen Vorhersagen ist in Abbildung 5.10 dargestellt. In diesem Fall führen wir Simulationsstudien mit 100 000 Replikationen für unterschiedliche Ensemble-Größen von $M = 25$ bis $M = 500$ Ensemble-Mitgliedern durch und spezifizieren die Abhängigkeitsstruktur dabei mit einer Gauß-Copula mit festem Korrelationsparameter $\rho = 0.2$. Der Vergleich der Variogramm Skill Scores und der Energy Skill Scores illustriert erneut, dass sich die beiden Bewertungsregeln deutlich unterscheiden. Während der Variogramm Score den ECCQ-Vorhersagen eine im Vergleich zu den EMOS-Vorhersagen bessere Qualität für Ensemble-Größen größer oder gleich $M = 75$ bescheinigt, ist für den Energy Score im Gegensatz eine Ensemble-Größe von $M = 250$ oder höher notwendig. Der Faire Energy Score indiziert hingegen für alle Ensemble-Größen eine verbesserte Vorhersagequalität. Dies liegt erneut hauptsächlich an der unterschiedlichen Struktur der Ensembles und an der Verzerrung des Fairen Energy Scores, falls das Ensemble aus äquidistanten Quantilen der zugrundeliegenden Verteilung konstruiert wird. Ähnlich wie von Zamo und Naveau (2018) für den fairen CRPS gezeigt, ist diese Verzerrung insbesondere für kleine Ensemble-Größen ausgeprägt. Deshalb schnei-

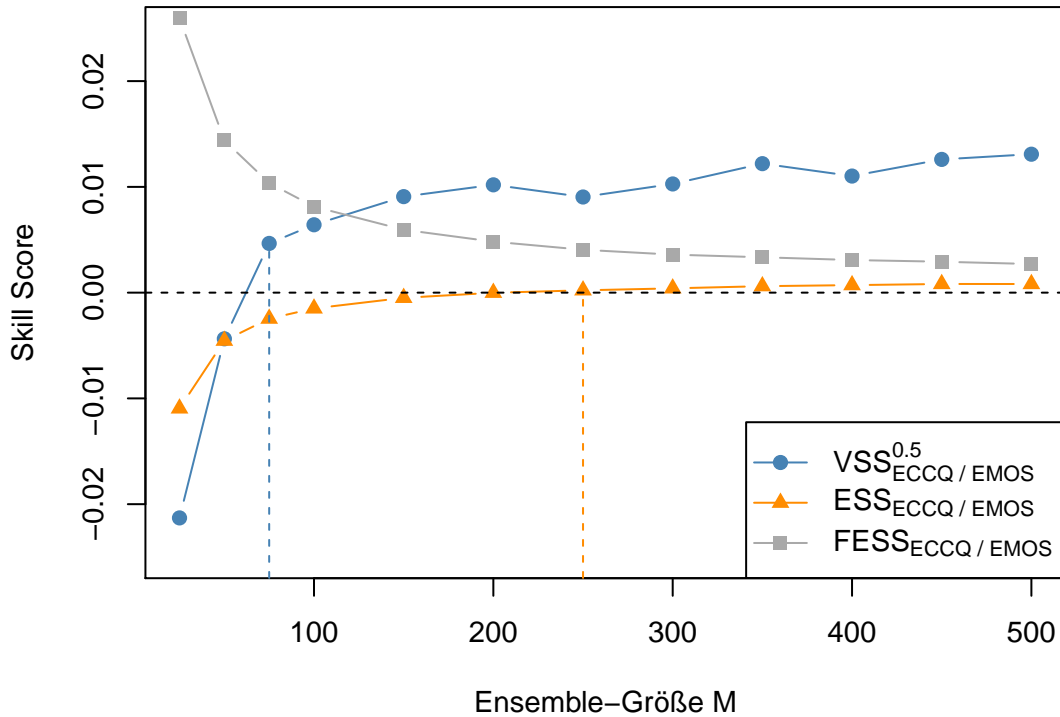


Abbildung 5.10: Simulierte Skill-Scores für bivariate ECCQ-artige Vorhersagen mit einer variierenden Anzahl von M Ensemble-Mitgliedern verglichen mit EMOS-artigen Vorhersagen (bestehend aus 10 000 Ensemble-Mitgliedern). In der Simulation wird die wahre Abhängigkeitsstruktur zwischen zwei trunziert normalverteilten Zufallsvariablen durch eine Gauß-Copula mit festgewähltem Korrelationsparameter $\rho = 0.2$ spezifiziert. Für jeden Wert von M wird eine Monte-Carlo-Simulation mit 100 000 Replikationen durchgeführt. Die Skill-Scores werden daraufhin basierend auf gemittelten Scores über alle Replikationen berechnet.

det das ECCQ-artige Ensemble hier im Vergleich deutlich besser ab als die EMOS-artige Vorhersage.

Insgesamt bestätigt die Sensitivitätsanalyse die Ergebnisse von Scheuerer und Hamill (2015), dass der Variogramm Score generell deutlich sensitiver auf fehlspezifizierte Abhängigkeitsstrukturen reagiert als der Energy Score. Auch wenn Ensembles mit unterschiedlichen Ensemble-Größen verglichen werden, identifiziert der Variogramm Score fehlspezifizierte Abhängigkeitsstrukturen noch relativ zuverlässig. Wie von Scheuerer und Hamill (2015) angesprochen, besitzt der Variogramm Score jedoch auch einige unattraktive Ei-

genschaften im Vergleich zum Energy Score. So wird eine gleichmäßig verzerrte Vorhersage (z.B. eine um ein Meter pro Sekunde zu hohe Windgeschwindigkeit an allen Standorten) vom Variogramm Score nicht bestraft, da diese Verzerrung durch das Differenzieren bei der Berechnung des Scores eliminiert wird. Anwender sollten deshalb in empirischen Studien möglichst verschiedene Bewertungskriterien verwenden, um die Qualität von unterschiedlichen probabilistischen Vorhersagen sinnvoll miteinander vergleichen zu können. Die Resultate zeigen außerdem, dass beim Vergleich von Ensemble-Vorhersagen mit unterschiedlichen Ensemble-Größen oder einer unterschiedlichen Struktur auch deren Auswirkung auf die Bewertungsregeln berücksichtigt werden muss. Andernfalls kann es leicht passieren, dass fehlspezifizierte Abhängigkeitsstrukturen nicht korrekt identifiziert werden und dass die Ergebnisse aus diesem Grund fehlinterpretiert werden.

5.6 Diskussion und Fazit

Die durchgeführten empirischen Analysen sowie die Sensitivitätsanalyse in diesem Kapitel verdeutlichen die Schwierigkeiten, die bei der Bewertung von multivariaten probabilistischen Windgeschwindigkeitsvorhersagen auftreten können. Zum einen zeigt sich, dass die beiden populärsten multivariaten Bewertungsregeln, der Energy Score und der Variogramm Score, sehr unterschiedlich auf fehlspezifizierte räumliche Abhängigkeiten in den Vorhersagen reagieren. Weiterhin zeigt sich, dass die Ensemble-Größe und die Ensemble-Struktur einen großen Einfluss auf die Bewertungsregeln haben. Vor allem bei einem Vergleich von Ensembles mit unterschiedlichen Ensemble-Strukturen, wie zum Beispiel beim Vergleich einer Zufallsstichprobe gegenüber einem Ensemble bestehend aus äquidistanten Quantilen, können leicht Fehlinterpretationen auftreten. Dies kann dazu führen, dass fehlspezifizierte Abhängigkeiten in multivariaten Vorhersagen nur dadurch nicht identifiziert werden, weil die verglichenen Ensemble-Vorhersagen durch unterschiedliche Ensemble-Größen oder Ensemble-Strukturen charakterisiert werden. Insgesamt zeigt das Kapitel daher, dass die Eigenschaften der Ensemble-Vorhersagen sowie die Unterschiede in den

Bewertungsregeln beachtet werden müssen, wenn die Qualität multivariater probabilistischer Vorhersagen sinnvoll eingeschätzt werden soll. Die Ergebnisse sind daher nicht nur für Windgeschwindigkeitsvorhersagen relevant, sondern für viele Anwendungen in denen die Qualität multivariater probabilistischer Prognosen verglichen werden soll.

In zukünftigen Arbeiten möchten wir die Analysen aus diesem Kapitel für eine größere Anzahl an räumlichen Standorten durchführen, um zu überprüfen ob sich die Ergebnisse für höhere Dimensionen verändern. Weiterhin stellt sich die Frage, inwiefern multivariate probabilistische Vorhersagen das in Kapitel 4 beschriebene Phänomen berücksichtigen, dass sehr große Abweichungen von der deterministischen Vorhersage häufig gemeinsam in einem räumlichen Kontext auftreten. Für die Beantwortung dieser Frage könnten zum Beispiel Bewertungsregeln verwendet werden, die speziell für die Bewertung von Quantilsvorhersagen (siehe Gneiting und Raftery, 2007) geeignet sind. Dadurch ließe sich gezielter beurteilen, ob multivariate probabilistische Vorhersagen die Wahrscheinlichkeit eines gemeinsamen Auftretens von solchen extremen Abweichungen korrekt einschätzen.

6 Zusammenfassung

Die vorliegende Dissertation befasst sich mit der Identifizierung und Modellierung von räumlichen Abhängigkeitsstrukturen bei Windgeschwindigkeits- und Windleistungsvorhersagen, die für den Energiesektor von zentraler Bedeutung sind. Einzuspeisende Strommengen werden bereits Tage im Voraus hinsichtlich des zu erwartenden Angebots und der erwarteten Nachfrage gehandelt und ausgerichtet. Bei ungeplanten Produktionsüberschüssen oder fehlenden Strommengen, die durch Prognosefehler der fluktuierenden Windleistung entstehen, können deshalb zusätzliche Kosten auf einzelne Energiemarktteilnehmer zukommen. Hohe aggregierte Vorhersagefehler für ein gesamtes räumliches Gebiet stellen zudem ein potenzielles Risiko für die Netzsicherheit dar. Diese treten insbesondere dann auf, wenn zwischen den Vorhersagefehlern an einzelnen Standorten eine ausgeprägte räumliche Abhängigkeitsstruktur besteht. Im ersten Teil dieser Dissertation analysieren wir deshalb zunächst mit Hilfe von statistischen Verfahren, wie sich die räumliche Abhängigkeitsstruktur von deterministischen Windleistungsvorhersagefehlern in Europa für verschiedene Vorhersagehorizonte und unterschiedliche geographische Bedingungen unterscheidet. Die Windleistungsvorhersagen basieren dabei auf deterministischen Windgeschwindigkeitsvorhersagen des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF). Aufgrund der hohen räumlichen Auflösung der Wettervorhersagen liegt für die Analyse eine Vielzahl an Datenpunkten vor. Dies stellt Herausforderungen an die verwendeten statistischen Verfahren, die einerseits aussagekräftige Ergebnisse produzieren müssen aber gleichzeitig auch eine angemessene Rechenzeit aufweisen sollten. Für die erste Analyse von räumlichen Abhängigkeitsstrukturen bei Windleistungsvorhersagefehlern

verwenden wir zunächst ein in der Masterarbeit von Hüsich (2016) entwickeltes Clusteringverfahren für räumlich-zeitliche Datensätze. Dieses Verfahren wird in dieser Dissertation zudem für eine Anwendung bei noch größeren Datensätzen sinnvoll erweitert. Die Clusteranalyse zeigt, dass räumliche Abhängigkeitsstrukturen in den Vorhersagefehlern für längere Vorhersagehorizonte ausgeprägter sind und dass hohe Abhängigkeiten vor allem in Regionen mit einer hohen Anzahl an installierter Windleistung auftreten. Für eine noch genauere Analyse der Abhängigkeitsstrukturen schlagen wir außerdem ein copulabasiertes Verfahren vor. Dieses ermöglicht eine gezieltere Charakterisierung räumlicher Abhängigkeiten in Hinblick auf unterschiedliche geographische Gegebenheiten. Die Anwendung des Verfahrens für die Vorhersagefehler bestätigt die Ergebnisse aus der Clusteranalyse und zeigt zudem, dass insbesondere sehr hohe Vorhersagefehler häufig gemeinsam an benachbarten Standorten auftreten.

Die identifizierten räumlichen Abhängigkeitsstrukturen illustrieren, dass durch die Verwendung von rein deterministischen Vorhersagen ein hohes Risiko für Energiemarktteilnehmer und Übertragungsnetzbetreiber entstehen kann. Um die Vorhersageunsicherheit bereits im Vorfeld besser abschätzen zu können, bieten sich probabilistische Vorhersagen an. Meist wird dazu auf meteorologische Ensemble-Vorhersagen zurückgegriffen, die ebenfalls von verschiedenen Wetteranbietern veröffentlicht werden. Eine sinnvolle probabilistische Vorhersage für mehrere Standorte muss einerseits die Vorhersageunsicherheit an einzelnen Standorten adäquat einschätzen, und gleichzeitig auch die räumliche Abhängigkeitsstruktur zwischen den Standorten berücksichtigen. Im abschließenden Teil dieser Dissertation untersuchen wir deshalb, wie die Qualität von Ensemble-Vorhersagen im Falle von ausgeprägten räumlichen Abhängigkeitsstrukturen sinnvoll bewertet werden kann. In einer empirischen Analyse zeigen wir, dass gängige Bewertungsregeln für multivariate probabilistische Vorhersagen sensitiv auf unterschiedliche Strukturen von Ensemble-Vorhersagen reagieren. Insbesondere der Einfluss der Ensemble-Größe darf in empirischen Analysen nicht vernachlässigt werden. Eine Simulationsstudie verdeutlicht, dass andernfalls fehlspezifizierte Abhängigkeitsstrukturen in den Vorhersagen von den Bewertungs-

regeln zum Teil nur aufgrund von unterschiedlichen Ensemble-Größen und verschiedenen Ensemble-Strukturen nicht korrekt identifiziert werden.

In weiterführenden Forschungsarbeiten könnte der Fokus verstärkt auf die Analyse von sehr großen Windleistungsvorhersagefehlern gelegt werden. Hier ist insbesondere interessant, welche direkten Auswirkungen sehr hohe aggregierte Vorhersagefehler auf die Netzstabilität und die Entwicklung von Strompreisen auf Energiemärkten haben. Unter Umständen könnten Informationen aus meteorologischen Ensemble-Vorhersagen ebenfalls dazu verwendet werden, probabilistische Vorhersagen für die Strompreisentwicklung auf Energiemärkten zu treffen. Dies wäre insbesondere im Sinne des Risikomanagements von Bedeutung. Auch im Kontext von probabilistischen Strompreisvorhersagen ist dann eine geeignete Bewertung der Vorhersagen wieder relevant. Die Erkenntnisse aus dem abschließenden Teil dieser Dissertation könnten für diese Untersuchungen hilfreich sein.

Literaturverzeichnis

- Agrawal, K., S. Garg, S. Sharma und P. Patel (2016). „Development and validation of OPTICS based spatio-temporal clustering technique“. *Information Sciences* 369, S. 388–401.
- Auguie, B. (2017). *gridExtra: Miscellaneous Functions for “Grid” Graphics*. R package version 2.3. URL: <https://CRAN.R-project.org/package=gridExtra>.
- Bárdossy, A. (2006). „Copula-based geostatistical models for groundwater quality parameters“. *Water Resources Research* 42(11).
- Bárdossy, A. und J. Li (2008). „Geostatistical interpolation using copulas“. *Water Resources Research* 44(7).
- Becker, R. A., A. R. Wilks, R. Brownrigg, T. P. Minka und A. Deckmyn (2017). *maps: Draw Geographical Maps*. R package version 3.2.0. URL: <https://CRAN.R-project.org/package=maps>.
- Beygelzimer, A., S. Kakadet, J. Langford, S. Arya, D. Mount und S. Li (2013). *FNN: Fast Nearest Neighbor Search Algorithms and Applications*. R package version 1.1. URL: <https://CRAN.R-project.org/package=FNN>.
- Birant, D. und A. Kut (2007). „ST-DBSCAN: An algorithm for clustering spatial–temporal data“. *Data & Knowledge Engineering* 60(1), S. 208–221.
- Bivand, R. S., T. Keitt und B. Rowlingson (2017). *rgdal: Bindings for the ‘Geospatial’ Data Abstraction Library*. R package version 1.2-16. URL: <https://CRAN.R-project.org/package=rgdal>.

- Bradley, A. A. und S. S. Schwartz (2011). „Summary Verification Measures and Their Interpretation for Ensemble Forecasts“. *Monthly Weather Review* 139(9), S. 3075–3089.
- Bundesnetzagentur (2019). *Regelenergie*. URL: https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Versorgungssicherheit/Engpassmanagement/Regelenergie/regelenergie-node.html (besucht am 27.08.2019).
- Cannon, D., D. Brayshaw, J. Methven und D. Drew (2017). „Determining the bounds of skilful forecast range for probabilistic prediction of system-wide wind power generation“. *Meteorologische Zeitschrift* 26(3), S. 239–252.
- Carnell, R. (2018). *lhs: Latin Hypercube Samples*. R package version 0.16. URL: <https://CRAN.R-project.org/package=lhs>.
- Carrera-Hernández, J. J. und S. J. Gaskin (2007). „Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico“. *Journal of Hydrology* 336(3), S. 231–249.
- Coles, S. G. (1993). „Regional Modelling of Extreme Storms Via Max-Stable Processes“. *Journal of the Royal Statistical Society: Series B (Methodological)* 55(4), S. 797–816.
- Defays, D. (1977). „An efficient algorithm for a complete link method“. *The Computer Journal* 20(4), S. 364–366.
- Demarta, S. und A. J. McNeil (2005). „The t Copula and Related Copulas“. *International Statistical Review* 73(1), S. 111–129.
- Durocher, M., F. Chebana und T. B. Ouarda (2016). „On the prediction of extreme flood quantiles at ungauged locations with spatial copula“. *Journal of Hydrology* 533, S. 523–532.
- ECMWF (2016). *New forecast model cycle brings highest-ever resolution*. URL: <http://www.ecmwf.int/en/about/media-centre/news/2016/new-forecast-model-cycle-brings-highest-ever-resolution> (besucht am 16.08.2019).
- ECMWF (2019). *Medium-range forecasts*. URL: <http://www.ecmwf.int/en/forecasts/documentation-and-support/medium-range-forecasts> (besucht am 16.08.2019).

- EHA (2017). *Wie läuft der Handel mit Strom*. URL: <https://www.eha.net/blog/details/handel-mit-strom.html> (besucht am 27.08.2019).
- EnergieAgentur.NRW (2017). *Big Data: Präzisere Wetter- und Leistungsprognosen verhelfen zu Netzstabilität*. URL: <https://www.energieagentur.nrw/blogs/erneuerbare/beitraege/big-data-praezisere-wetter-und-leistungsprognosen-verhelfen-zu-netzstabilitaet/> (besucht am 27.08.2019).
- Ericson, K. und S. Pallickara (2013). „On the performance of high dimensional data clustering and classification algorithms“. *Future Generation Computer Systems* 29(4). Special Section: Utility and Cloud Computing, S. 1024–1034.
- Ester, M., H.-P. Kriegel, J. Sander und X. Xu (1996). „A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise“. In: *KDD-96 Proceedings*. AAAI Press, S. 226–231.
- Everitt, B. S., S. Landau, M. Leese und D. Stahl (2011). *Cluster Analysis*. Fünfte Auflage. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester, Großbritannien.
- EWEA (2016). *Wind in power: 2015 European statistics*. URL: <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/EWEA-Annual-Statistics-2015.pdf> (besucht am 16.08.2019).
- Feldmann, K., M. Scheuerer und T. L. Thorarinsdottir (2015). „Spatial Postprocessing of Ensemble Forecasts for Temperature Using Nonhomogeneous Gaussian Regression“. *Monthly Weather Review* 143(3), S. 955–971.
- Ferro, C. A. T., D. S. Richardson und A. P. Weigel (2008). „On the effect of ensemble size on the discrete and continuous ranked probability scores“. *Meteorological Applications* 15(1), S. 19–24.
- Focken, U., M. Lange, K. Mönnich, H.-P. Waldl, H. G. Beyer und A. Luig (2002). „Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects“. *Journal of Wind Engineering and Industrial Aerodynamics* 90(3), S. 231–246.

- Genest, C., K. Ghoudi und L.-P. Rivest (1995). „A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions“. *Biometrika* 82(3), S. 543–552.
- Gneiting, T. und A. E. Raftery (2007). „Strictly Proper Scoring Rules, Prediction, and Estimation“. *Journal of the American Statistical Association* 102(477), S. 359–378.
- Gneiting, T., A. E. Raftery, A. H. Westveld III und T. Goldman (2005). „Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation“. *Monthly Weather Review* 133(5), S. 1098–1118.
- Gräler, B. (2014). „Modelling skewed spatial random fields through the spatial vine copula“. *Spatial Statistics* 10, S. 87–102.
- Hamill, T. M. und S. J. Colucci (1997). „Verification of Eta–RSM Short-Range Ensemble Forecasts“. *Monthly Weather Review* 125(6), S. 1312–1327.
- Hartigan, J. A. und M. A. Wong (1979). „Algorithm AS 136: A K-Means Clustering Algorithm“. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28(1), S. 100–108.
- Hastie, T. J. und R. J. Tibshirani (1990). *Generalized Additive Models*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, Boca Raton, USA.
- Hemri, S., D. Lisniak und B. Klein (2015). „Multivariate postprocessing techniques for probabilistic hydrological forecasting“. *Water Resources Research* 51(9), S. 7436–7451.
- Hofert, M., I. Kojadinovic, M. Maechler und J. Yan (2017). *copula: Multivariate Dependence with Copulas*. R package version 0.999-17. URL: <https://CRAN.R-project.org/package=copula>.
- Hüsch, M. (2016). „Analyzing Spatio-Temporal Dependence of Wind Power Forecast Errors“. Masterarbeit. Technische Universität Dortmund.
- Hüsch, M., B. U. Schyska und L. von Bremen (2018). „CorClustST - Correlation-based clustering of big spatio-temporal datasets“. *Future Generation Computer Systems*. URL: <https://doi.org/10.1016/j.future.2018.04.002>.

- Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, Boca Raton, USA.
- Jordan, A., F. Krüger und S. Lerch (2017). „Evaluating probabilistic forecasts with scoringRules“. *arXiv Preprint*. URL: <https://arxiv.org/abs/1709.04743>.
- Junk, C., L. von Bremen, M. Kühn, S. Späth und D. Heinemann (2014). „Comparison of Postprocessing Methods for the Calibration of 100-m Wind Ensemble Forecasts at Off- and Onshore Sites“. *Journal of Applied Meteorology and Climatology* 53(4), S. 950–969.
- Kahle, D. und H. Wickham (2013). „ggmap: Spatial Visualization with ggplot2“. *The R Journal* 5(1), S. 144–161.
- Kaufman, L. und P. J. Rousseeuw (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, USA.
- Kisilevich, S., F. Mansmann, M. Nanni und S. Rinzivillo (2010). „Spatio-temporal clustering“. In: *Data Mining and Knowledge Discovery Handbook*. Hrsg. von O. Maimon und L. Rokach. Zweite Auflage. Springer, New York, USA. Kap. 44, S. 855–874.
- Krupskii, P. und M. G. Genton (2017). „Factor copula models for data with spatio-temporal dependence“. *Spatial Statistics* 22, S. 180–195.
- Lange, M. (2005). „On the Uncertainty of Wind Power Predictions - Analysis of the Forecast Accuracy and Statistical Distribution of Errors“. *Journal of Solar Energy Engineering* 127(2), S. 177–184.
- Malevergne, Y. und D. Sornette (2003). „Testing the Gaussian copula hypothesis for financial assets dependences“. *Quantitative Finance* 3(4), S. 231–250.
- Manner, H. und O. Reznikova (2012). „A Survey on Time-Varying Copulas: Specification, Simulations, and Application“. *Econometric Reviews* 31(6), S. 654–687.
- Möller, A., A. Lenkoski und T. L. Thorarinsdottir (2013). „Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas“. *Quarterly Journal of the Royal Meteorological Society* 139(673), S. 982–991.

- Nychka, D., R. Furrer, J. Paige und S. Sain (2015). *fields: Tools for Spatial Data*. R package version 9.0. URL: <https://CRAN.R-project.org/package=fields>.
- Pearson, K. (1895). „Note on Regression and Inheritance in the Case of Two Parents“. *Proceedings of the Royal Society of London* 58, S. 240–242.
- Pebesma, E. J. und R. S. Bivand (2005). „Classes and methods for spatial data in R“. *R News* 5(2), S. 9–13.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Wien, Österreich. URL: <https://www.R-project.org/>.
- Rudis, B., B. Bolker und J. Schulz (2017). *ggalt: Extra Coordinate Systems, ‘Geoms’, Statistical Transformations, Scales and Fonts for ‘ggplot2’*. R package version 0.4.0. URL: <https://CRAN.R-project.org/package=ggalt>.
- Schabenberger, O. und C. A. Gotway (2004). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, Boca Raton, USA.
- Schefzik, R., T. L. Thorarinsdottir und T. Gneiting (2013). „Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling“. *Statistical Science* 28(4), S. 616–640.
- Schepsmeier, U., J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler und T. Erhardt (2017). *VineCopula: Statistical Inference of Vine Copulas*. R package version 2.1.2. URL: <https://CRAN.R-project.org/package=VineCopula>.
- Scheuerer, M. und T. M. Hamill (2015). „Variogram-Based Proper Scoring Rules for Probabilistic Forecasts of Multivariate Quantities“. *Monthly Weather Review* 143(4), S. 1321–1334.
- Sibson, R. (1973). „SLINK: an optimally efficient algorithm for the single-link cluster method“. *The Computer Journal* 16(1), S. 30–34.
- Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Publication de l’Institut de Statistique de l’Université de Paris.

- Sloughter, J. M. L., T. Gneiting und A. E. Raftery (2010). „Probabilistic Wind Speed Forecasting Using Ensembles and Bayesian Model Averaging“. *Journal of the American Statistical Association* 105(489), S. 25–35.
- Späth, S., L. von Bremen, C. Junk und D. Heinemann (2015). „Time-consistent calibration of short-term regional wind power ensemble forecasts“. *Meteorologische Zeitschrift* 24(4), S. 381–392.
- Statista (2019). *Verteilung der Stromerzeugung aus Erneuerbaren Energien in Deutschland nach Energieträger im Jahr 2018*. URL: <https://de.statista.com/statistik/daten/studie/173871/umfrage/stromerzeugung-aus-erneuerbaren-energien-in-deutschland/> (besucht am 27.08.2019).
- Tastu, J., P. Pinson, E. Kotwa, H. Madsen und H. A. Nielsen (2011). „Spatio-temporal analysis and modeling of short-term wind power forecast errors“. *Wind Energy* 14(1), S. 43–60.
- Thorarinsdottir, T. L. und T. Gneiting (2010). „Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression“. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173(2), S. 371–388.
- Thorndike, R. L. (1953). „Who belongs in the family?“. *Psychometrika* 18(4), S. 267–276.
- Trautmann, H., D. Steuer, O. Mersmann und B. Bornkamp (2014). *truncnorm: Truncated Normal Distribution*. R package version 1.0-7. URL: <https://CRAN.R-project.org/package=truncnorm>.
- Umweltbundesamt (2019). *Bilanz 2018: Anteil erneuerbarer Energien steigt auf 16,6 Prozent*. URL: <https://www.umweltbundesamt.de/presse/pressemitteilungen/bilanz-2018-anteil-erneuerbarer-energien-steigt-auf> (besucht am 27.08.2019).
- Vatter, T. und V. Chavez-Demoulin (2015). „Generalized additive models for conditional dependence structures“. *Journal of Multivariate Analysis* 141, S. 147–167.
- Ward, J. H. (1963). „Hierarchical Grouping to Optimize an Objective Function“. *Journal of the American Statistical Association* 58(301), S. 236–244.

- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Zweite Auflage. Springer, Cham, Schweiz.
- Wilks, D. S. (2019). *Statistical Methods in the Atmospheric Sciences*. Vierte Auflage. Elsevier, Amsterdam, Niederlande.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*. Zweite Auflage. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, Boca Raton, USA.
- Yuen, R., S. Baran, C. Fraley, T. Gneiting, S. Lerch, M. Scheuerer und T. Thorarinsdottir (2018). *ensembleMOS: Ensemble Model Output Statistics*. R package version 0.8.2. URL: <https://CRAN.R-project.org/package=ensembleMOS>.
- Zamo, M. und P. Naveau (2018). „Estimation of the Continuous Ranked Probability Score with Limited Information and Applications to Ensemble Weather Forecasts“. *Mathematical Geosciences* 50(2), S. 209–234.