

PROFESSUR FÜR HYDROLOGIE
der Albert-Ludwigs-Universität Freiburg i. Br.

Janek Jenne

Vergleich und Analyse verschiedener Verfahren zur Biaskorrektur von Niederschlagsradar-Daten

Referent: Prof. Dr. Markus Weiler
Korreferent: Prof. Dr. Dirk Schindler

Masterarbeit unter Leitung von Prof. Dr. Markus Weiler
Freiburg i. Br., November 2023

PROFESSUR FÜR HYDROLOGIE
der Albert-Ludwigs-Universität Freiburg i. Br.

Janek Jenne

Vergleich und Analyse verschiedener Verfahren zur Biaskorrektur von Niederschlagsradar-Daten

Referent: Prof. Dr. Markus Weiler
Korreferent: Prof. Dr. Dirk Schindler

Masterarbeit unter Leitung von Prof. Dr. Markus Weiler
Freiburg i. Br., November 2023

Inhaltsverzeichnis

I	Inhaltsverzeichnis	I
II	Abbildungsverzeichnis	III
III	Tabellenverzeichnis	IV
IV	Abbildungsverzeichnis des Anhangs	V
V	Tabellenverzeichnis des Anhangs	V
	Zusammenfassung	VII
	Abstract	IX
1	Einleitung	1
1.1	Hintergrund von Biaskorrekturverfahren	1
1.1.1	Hintergrund	1
1.1.2	Definition des Bias und der Biaskorrektur	1
1.1.3	Verschiedene Arten der Biaskorrektur	2
1.1.4	Anwendungsgebiete	4
1.2	Nutzen von Niederschlagsradaren	4
1.2.1	Funktionsweise eines Niederschlagsradars	4
1.2.2	Weiterentwicklungen und Korrekturen	5
1.2.3	Potentiale von Radarprodukten	6
1.2.4	Vor- und Nachteile von Radardaten	7
1.2.5	Untersuchungen von Extremniederschlägen mittels Radardaten	8
2	Problemstellung und Zielsetzung	11
3	Methoden und Vorgehensweise	13
3.1	Datengrundlage	13
3.1.1	Radardaten	13
3.1.2	Stationsdaten	14
3.2	Methoden	16
3.2.1	Qualitätskontrolle	16
3.2.2	Untersuchung der Klimatologie	17
3.2.3	Berechnung der kumulativen Niederschlagssummen	17
3.2.4	Quantifizierung des Bias	18
3.2.5	Erstellung der Trainings- und Kontrollgruppen	19
3.2.6	Biaskorrekturverfahren	20
3.2.7	Sensitivitätstests	22
3.3	Anwendungsbeispiele	22
3.3.1	Korrektur der realen Events	22
3.3.2	Berechnung von Bemessungsniederschlägen	23

4	Resultate	25
4.1	Klimatologie	25
4.2	Bias	27
4.2.1	Häufigkeitsverteilungen	27
4.2.2	Räumliche Muster	29
4.2.3	Zusammenhänge mit anderen Parametern	31
4.2.4	Bias von realen Events	33
4.3	Biaskorrektur - Verfahren 1	34
4.4	Biaskorrektur - Verfahren 2	42
4.5	Biaskorrektur - Direkter Vergleich	47
4.6	Biaskorrektur - Sensitivitätstests	52
4.7	Anwendungsbeispiele - Korrektur realer Events	60
4.8	Anwendungsbeispiele - Berechnung von Bemessungsniederschlägen	62
5	Diskussion	65
5.1	Klimatologie	65
5.2	Bias	65
5.3	Biaskorrektur	68
5.4	Anwendungsbeispiele	74
5.5	Abschließende Einschätzung	74
6	Schlussfolgerungen	77
7	Datensätze	79
	Literatur	87
	Anhang	89

Abbildungsverzeichnis

Abb. 3.1	Karten der Niederschlagsstationen und der Wetterradare des Deutschen Wetterdienstes	14
Abb. 3.2	Darstellung von drei der fünf verwendeten Biastypen	18
Abb. 3.3	Räumliche Verteilung der Stationen einer jeden in dieser Arbeit verwendeten Trainings- und Kontrollgruppe	19
Abb. 4.1	Beispiel der klimatologischen Analyse (1)	25
Abb. 4.2	Beispiel der klimatologischen Analyse (2)	26
Abb. 4.3	Anzahl der Extremereignisse während des Untersuchungszeitraums basierend auf den Stationsdaten	26
Abb. 4.4	Histogramm des Bias (Stat-Px) einer jeden Dauerstufe für den zweiten Event-Rang	27
Abb. 4.5	Histogramm des Bias (Stat-Px) der Ränge 1, 5, 10 und 30 bezogen auf die Dauerstufe D10	28
Abb. 4.6	Vergleich des Biases von Stat-Px, Mean4 und Mean9 des fünften Ranges (D10)	29
Abb. 4.7	Räumliche Verteilung der Niederschläge und des Bias für den ersten Event-Rang (D10)	30
Abb. 4.8	Zusammenhang zwischen Bias und Stationshöhe	31
Abb. 4.9	Zusammenhang zwischen Bias und Richter-Klasse	32
Abb. 4.10	Zusammenhang zwischen Bias und dem jeweiligen Radarturm	33
Abb. 4.11	Beispiel des Bias zweier realer Events	34
Abb. 4.12	Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D10)	35
Abb. 4.13	Verfahren 1: ECDFs für den siebten Event-Rang	36
Abb. 4.14	Verfahren 1: Q-Q-Plot für den 75. Event-Rang (D10)	37
Abb. 4.15	Verfahren 1: Unterschiede in den ECDFs zwischen den Gruppen für den 20. Event-Rang (D10)	38
Abb. 4.16	Verfahren 1: Korrekturfaktoren für alle Event-Ränge, Quantile und Dauerstufen	39
Abb. 4.17	Anwendung der mithilfe von Verfahren 1 ermittelten Korrekturfaktoren auf alle Rasterzellen	40
Abb. 4.18	Verfahren 1: Vergleich der korrigierten Radar-Niederschlagskarten der drei Gruppen	41
Abb. 4.19	Verfahren 2: Q-Q-Plot für den siebten Event-Rang (D10)	42
Abb. 4.20	Verfahren 2: ECDFs für den siebten Event-Rang	43
Abb. 4.21	Verfahren 2: Korrekturfaktoren für alle Dauerstufen	44
Abb. 4.22	Anwendung der mithilfe von Verfahren 2 ermittelten Korrekturfaktoren auf alle Rasterzellen	45
Abb. 4.23	Verfahren 2: Vergleich der korrigierten Radar-Niederschlagskarten der drei Gruppen	46
Abb. 4.24	Vergleich des verbleibenden Bias beider Verfahren für die Kontrollgruppe bei Rang 10	48
Abb. 4.25	Vergleich des verbleibenden Bias beider Verfahren für die Trainingsgruppe bei Rang 10	49

Abb. 4.26	Vergleich der korrigierten Radarkarten beider Verfahren bei Rang 10	50
Abb. 4.27	Vergleich der korrigierten Radarkarten beider Verfahren bei Rang 75	51
Abb. 4.28	Sensitivitätstest 2: Q-Q-Plot für den siebten Event-Rang (D10)	52
Abb. 4.29	Sensitivitätstest 2: ECDFs für den siebten Event-Rang	53
Abb. 4.30	Sensitivitätstest 2: Q-Q-Plot für den ersten Event-Rang (D180)	54
Abb. 4.31	Anteil der Sommer- und Winterevents für die Stations- und Radardaten (Stat-Px)	55
Abb. 4.32	Sensitivitätstest 3: ECDFs für den 129. Event-Rang	55
Abb. 4.33	Sensitivitätstest 4 ECDFs für den 15. Event-Rang	56
Abb. 4.34	Vergleich der Korrekturfaktoren aller Sensitivitätstests	57
Abb. 4.35	Differenzen der ECDFs für Rang 10 (D10)	58
Abb. 4.36	Histogramme der Stations- und Radardaten für Rang 10 (D10)	58
Abb. 4.37	Korrektur des realen Ereignisses vom 17. Juni 2020 (Station 567)	60
Abb. 4.38	Korrektur des realen Ereignisses vom 31. Mai 2018 (Station 970)	61
Abb. 4.39	Korrektur des realen Ereignisses vom 6. August 2019 (Station 5029)	61
Abb. 4.40	Vergleich des berechneten 20-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 1	63
Abb. 4.41	Vergleich des berechneten 20-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 2	64
Abb. 7.1	Karte der verwendeten und verworfenen Niederschlagstationen	95
Abb. 7.2	Höhenverteilung der drei Trainings- und Kontrollgruppen	96
Abb. 7.3	Niederschlagsintensitäten des 95., 99., und 99,9. Quantils	97
Abb. 7.4	Histogramm des relativen Bias (Stat-Px) der Ränge 1, 5, 10 und 30 bezogen auf die Dauerstufe D10	98
Abb. 7.5	Räumliche Verteilung der Niederschläge und des Bias für den 30. Event- Rang (D10)	99
Abb. 7.6	Zusammenhang zwischen relativem Bias und Stationshöhe	100
Abb. 7.7	Zusammenhang zwischen dem relativen Bias und dem jeweiligen Radarturm	101
Abb. 7.8	Beispiel für einen zunehmenden bzw. abnehmenden Trend mit größer wer- dender Entfernung vom Radarstandort	102
Abb. 7.9	Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D60)	103
Abb. 7.10	Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D180)	104
Abb. 7.11	Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D360)	105
Abb. 7.12	Verfahren 1: Differenzen der Korrekturfaktoren für die zweite Gruppe . . .	106
Abb. 7.13	Verfahren 1: Differenzen der Korrekturfaktoren für die dritte Gruppe . . .	106
Abb. 7.14	Verfahren 2: Q-Q-Plot für den 75. Event-Rang (D10)	107
Abb. 7.15	Vergleich des verbleibenden Bias beider Verfahren für die Kontrollgruppe bei Rang 75	108
Abb. 7.16	Vergleich des verbleibenden Bias beider Verfahren für die Trainingsgruppe bei Rang 9	109
Abb. 7.17	Sensitivitätstest 2: ECDFs aller Dauerstufen für Rang 125	110
Abb. 7.18	Sensitivitätstest 4: Q-Q-Plot für den 15. Event-Rang (D360)	111
Abb. 7.19	Vergleich der Sensitivitätstests bezüglich ihrer räumlichen Performance für Rang 5 (D60)	112
Abb. 7.20	Kumulierte Differenzen des ECDF-Wertes zum vorigen für Rang 10 (D10)	113
Abb. 7.21	Korrektur der realen Ereignisse bei Station 555 und 684	114
Abb. 7.22	Korrektur der realen Ereignisse bei Station 704 und 2319	115
Abb. 7.23	Korrektur der realen Ereignisse bei Station 3376 und 5763	116
Abb. 7.24	Vergleich des berechneten 50-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 1	117
Abb. 7.25	Vergleich des berechneten 50-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 2	118

Tabellenverzeichnis

Tab. 3.1	Warnstufen des DWDs für Starkregenereignisse (nach DWD, 2023)	17
Tab. 4.1	Anzahl der Stationen pro Richterklasse (DWD, 2023)	32
Tab. 4.2	Anzahl der im Datum übereinstimmenden Werte pro Dauerstufe	55
Tab. 7.1	Koordinaten der Wetterradar-Standorte des Deutschen Wetterdienstes (nach Winterrath et al., 2017 und DWD, 2018)	94
Tab. 7.2	Koeffizienten der Korrektur (nach Richter, 1995, S. 37, 67, verändert)	94

Zusammenfassung

Diese Arbeit hatte zum Ziel neue Verfahren zur Biaskorrektur von Radar-Daten zu entwickeln, zu testen und zu vergleichen. Die Nutzung von Radardaten zur flächenhaften Ermittlung von Niederschlägen hat für meteorologische und hydrologische Anwendungen neue Möglichkeiten geschaffen. Durch ihre große räumliche Abdeckung und eine hohe räumliche Auflösung sind sie ein wichtiges Hilfsmittel geworden, um kleinskalige Extremereignisse besser untersuchen zu können. Entgegen diesen räumlichen Vorteilen weisen die radarbasierten Niederschlagsdaten jedoch einen Bias gegenüber den Messungen der Niederschlagsstationen auf. Dieses Bias äußert sich meist in Form einer Unterschätzung der Niederschlagsintensität der Stationsniederschläge, gerade für sehr große Intensitäten. Gerade für die extremen Starkregenereignisse stellt diese Unterschätzung ein gewichtiges Problem dar, was die Schätzung von Bemessungsniederschlägen für die Anwendung in der Praxis zusätzlich erschwert. Daher sind umfangreiche Korrekturen notwendig, um das Bias zu beseitigen oder zumindest zu verringern. Entgegen der häufig eingesetzten Interpolationsverfahren, die meist auf Kriging basieren, wurde in dieser Arbeit ein empirischer Quantile-Mapping-Ansatz angewandt. Der räumliche Fokus der Arbeit lag dabei auf gesamt Deutschland. Anstelle von realen Extremereignissen wurde auf Basis der Stations- und Radardaten jeweils ein Datensatz der 150 größten Niederschlagssummen einer Dauerstufe pro Station bzw. zugehöriger Raster-Zelle geschaffen. In einem ersten Schritt wurde das Bias ausgiebig untersucht, anschließend wurden zwei verschiedene Verfahren auf Basis des Quantile-Mapping-Ansatz entwickelt und verglichen, mit dem Ziel den berechneten Ereignis-Datensatz zu korrigieren. Die erste Variante stellte einen rangbasierten Ansatz dar, bei dem die berechneten Niederschlagssummen zuerst gemäß ihrer Niederschlagsintensitäten in Ränge eingeteilt und diese einzeln korrigiert wurden. Die zweite Variante korrigierte alle 150 Ereignisse auf einmal. Erwartet wurde eine bessere Performance des rangbasierten Verfahrens. Auch in dieser Arbeit wurde größtenteils eine Unterschätzung der Stationsniederschläge durch die Radardaten festgestellt. An einigen Standorten war auch eine Überschätzung feststellbar. Beide entwickelten Biaskorrekturverfahren reduzierten das Bias zwischen beiden Datensätzen. Während das rangbasierte Verfahren insbesondere bei einer ausreichenden Datenverfügbarkeit als bessere Wahl erscheint, ist das nicht-rangbasierte Verfahren bei einer geringen Anzahl an Daten vorzuziehen. Insgesamt zeigte sich jedoch, dass weitere Analysen ausstehen, um die Performance beider Verfahren besser quantifizieren zu können.

Stichworte: Bias; Biaskorrektur; Quantile-Mapping; Radar; Station; Extremniederschlag; Deutschland

Abstract

The aim of this study is the development, testing and comparison of new methods for correcting the bias of weather radar data. The usage of weather radar data to detect spatial patterns of rainfall has opened new possibilities for meteorological and hydrological uses. They became a valuable tool to analyze small-scaled extreme events due to their wide spatial coverage and their high spatial resolution. In contrast to their spatial benefits radar data often contain a bias when compared to weather stations. In most cases this results in an underestimation of precipitation, especially for those of high intensity. This underestimation represents a major issue for estimating design storms. As a result extensive corrections are necessary to reduce or remove this bias. Contrary to the usually used interpolation methods like kriging, this study examines the application of an empirical quantile mapping approach to the whole area of Germany. Instead of focusing on real-time events, a dataset consisting of the 150 highest precipitation sums of one duration was created for each weather station. A second dataset of equal size was created for the corresponding radar data. At first the existing bias was thoroughly analyzed. Subsequently two methods based on the quantile mapping were developed and compared to correct the given dataset. For the first method the computed precipitation sums were ranked according to their intensity. Each rank was corrected individually. The second method corrected all 150 events at once without ranking the data. It was expected, that the rank-based method outperforms the second method. In this study an underestimation of precipitation was detected for most stations when using radar data. However at some locations an overestimation was found. Both methods are able to reduce the bias in radar data. While the rank-based approach is helpful if enough data is available, the second method should be preferred if the data availability is poor. Finally it was proven that more analyses are necessary to quantify the performance of both methods.

Keywords: Bias; Bias correction; Quantile Mapping; Radar; Gauge; Extreme precipitation; Germany

1 Einleitung

1.1 Hintergrund von Biaskorrekturverfahren

1.1.1 Hintergrund

Im Bereich der Umweltwissenschaften ist der Begriff der Biaskorrektur in den letzten Jahren und Jahrzehnten besonders in der Klimatologie bekannt geworden. Insbesondere im Bereich der Klimamodellierung wurden verschiedene Verfahren entwickelt, getestet und angewandt.

Die Ursprünge der Biaskorrektur im Bereich der Klimatologie liegen dabei in der numerischen Wettervorhersage (Maraun, 2016). Diese Korrektur findet in der statistischen Methode der *model output statistics* statt, die wiederum eine Weiterentwicklung des Verfahrens der perfekten Prognose (perfect prognosis) ist. Das Verfahren der *perfect prognosis* nimmt die Ausgabegrößen (Outputs) des numerischen Wettermodells als wahr an. Aus Beobachtungsdaten zwischen einem Prädiktor und einer Variablen, die vorhersagbar ist, wird eine statistische Beziehung abgeleitet. Diese Beziehung wird auf die vorhergesagten Ausgabegröße des Modells angewandt, um so eine Vorhersage des Prädiktors zu ermöglichen (AMS, 2012b). Das Verfahren der *model output statistics* hingegen ist eine Postprocessing-Technik, um die Fehler der Wettervorhersage zu verringern. Der Modelloutput wird also als nicht perfekt angesehen und meist mithilfe von multilinearen Regressionsmethoden korrigiert. Bei diesem Verfahren wird eine statistische Beziehung zwischen der Ausgabegröße des numerischen Modells und einer gemessenen Wettervariablen bestimmt. Der Unterschied zwischen beiden ist, dass bei *perfect prognosis* die statistische Beziehung ausschließlich zwischen zwei empirisch gemessenen Größen bestimmt wird, bei *model output statistics* zwischen einer gemessenen und einer modellierten Größe (AMS, 2012a). Bereits hier bestand die Idee (weitere) Messdaten zu verwenden, um den systematischen Fehler zu verringern.

Heutzutage wird die Biaskorrektur insbesondere verwendet, um die Outputs globaler und regionaler Klimamodelle zu korrigieren, damit diese für Klimaauswirkungsstudien verwendet werden können. Hierbei wird die systematische Abweichung zwischen einer simulierten und einer gemessenen klimatischen Größe für historische Daten bestimmt und anschließend diese Beziehung auf modellierte, prognostizierte Daten angewandt, unter der Annahme, dass diese Beziehung auch für diese gilt (Maraun, 2016).

1.1.2 Definition des Bias und der Biaskorrektur

Die Fragestellung der Biaskorrektur stammt dabei ursprünglich aus der Mathematik bzw. dem Teilgebiet der Statistik. Das Wort „Bias“ stammt aus dem Englischen und bedeutet so viel wie Verzerrung oder Voreingenommenheit. Das Bias bezeichnet dabei die Verzerrung einer Schätzfunktion, z. B. des Stichprobenmittels. Sie gibt die Abweichung des Erwartungswertes der Schätzfunktion vom wahren Parameterwert der Grundgesamtheit an. Im Rahmen des Konzeptes der Maximum-Likelihood stellt sie die Differenz zwischen dem Parameterwert, bei dem die Likelihood-Funktion ihren maximalen Wert aufweist und dem tatsächlichen, wahren Wert dar (Bohm und Zech, 2020, S. 51, 63).

Der Begriff des Bias ist dabei eng verwandt mit dem des systematischen Fehlers, teilweise wird dieser selbst als Bias bezeichnet (vgl. Hedderich und Sachs, 2020). Systematische Fehler, die zu einem Bias führen, können verschiedene Ursachen haben, die meist vermeidbar sind. Aufbauend auf Sackett, 1979 geben Hedderich und Sachs verschiedene Grundtypen des systematischen Fehlers an (Hedderich und Sachs, 2020, S. 14–15):

1. Systematische Abweichung eines Messwertes vom wahren Wert durch fehlerhafte Kalibrierung des Messgerätes.
2. Systematische Verzerrung eines statistischen Maßes vom wahren Wert durch Fehler während der Messung, Erfassung oder Berechnung (z. B. durch nicht anwendbare Formeln).
3. Systematisch verfälschte Aussagen durch Fehler im Studiendesign und in der Interpretation und Darstellung der daraus gewonnenen Daten.

Ist der systematische Fehler (oder Teile dessen) bekannt, kann versucht werden diesen zu korrigieren. Eine Biaskorrektur sieht also korrekte Daten vor, mithilfe derer die modellierten Daten, die verzerrt/biased sind, korrigiert werden können. Im Idealfall sollten Modelle so spezifiziert werden, dass eine Biaskorrektur nicht notwendig ist (Maraun, 2016).

Im weiteren Verlauf dieser Arbeit ist mit Bias die Abweichung zweier Datensätze gemeint. Die Verfahren der Biaskorrektur werden hier als Korrekturmethode dieser Datensätze definiert und sollten daher als statistische Postprocessing-Verfahren aufgefasst werden. Verfahren zur Kalibrierung der Messgeräte, der Datenaufbereitung oder ähnlichem werden folgend nicht als Methoden der Biaskorrektur verstanden.

1.1.3 Verschiedene Arten der Biaskorrektur

Im Folgenden soll eine kurze Übersicht über verschiedene Biaskorrekturverfahren gegeben werden, die in dieser Arbeit verwendeten Methoden werden im Kapitel 3.2.6 „Biaskorrekturverfahren“ ausführlicher erläutert. Eine Übersicht und kurze Beschreibung der folgend genannten Korrekturverfahren kann dem Anhang entnommen werden (vgl. Anhang 7)

Abhängig von der Fachdisziplin und der Fragestellung wurden unterschiedliche Biaskorrekturverfahren entwickelt. Ein Großteil der folgend besprochenen Verfahren stammt dabei aus der Klimatologie und der Modellierung von Klimaveränderungen. Der Fragestellung geschuldet werden drei verschiedene Datensätze unterschieden: die gemessenen sowie modellierten Daten einer historischen Klimareferenzperiode und die prognostizierten Daten einer zukünftigen Periode. Trotz ihrer Entwicklung im Bereich der Klimamodellierung, können die Methoden auf andere Disziplinen und Fragestellungen übertragen werden, nötig dafür ist nur die Unterscheidung der Daten in Mess- und Modelldaten für die Kontrolle sowie Modelldaten für die Prognose.

Gemäß Maraun, 2016 können die Biaskorrekturverfahren anhand verschiedener Kriterien unterteilt werden. Eine Einteilung unterscheidet die zu korrigierende Komponente einer Verteilung. So gibt es Verfahren, die ausschließlich den Mittelwert korrigieren, andere berücksichtigen zudem die Standardabweichung oder verschiedene Quantile der Verteilung. Weiter schlägt Maraun die Unterteilung in Trend-erhaltende und Trend-nicht-erhaltende Verfahren oder die Abgrenzung von univariaten zu bi- bzw. multivariaten Methoden vor. J. Chen et al., 2013 unterteilten die Korrekturverfahren einerseits in Mittelwert-basierte und Verteilungs-basierte Methoden, andererseits in Methoden mit und ohne Korrektur der Frequenz an Regentagen. Andere Klassifizierungen unterscheiden parametrische und nicht-parametrische Verfahren oder variable (die auch prognostizierte Statistiken, wie Mittelwert etc. miteinbeziehen) und konstante Verfahren (Watanabe et al., 2012). Auch eine Unterteilung in direkte und indirekte Verfahren ist möglich. Bei ersteren stellen die abhängige und unabhängige Variable dieselbe Größe dar, z. B. die Niederschlagssumme, bei indirekten nicht. Indirekte Verfahren können weiter in lineare und nichtlineare Methoden unterteilt werden (Themekl et al., 2011). Zudem nennt Themekl et al. die Option punktweise von flächigen Methoden zu unterscheiden.

Zu den einfachsten direkten Methoden gehören die Mittelwert-basierten Verfahren wie *Linear scaling*, *Delta-Methode* (vgl. Maraun, 2016) und *Local intensity scaling* (z. B. Themekl et al., 2011). Während diese Methoden für Temperaturen als additive Formel konzipiert wurden, werden für Niederschläge meist multiplikative Formeln (wegen nicht möglicher negativer Werte) genutzt und damit relative Änderungen betrachtet. Das Verfahren der *Power transformation* (z. B. Le-

ander und Buishand, 2007) stellt eine Erweiterung der vorigen dar, das neben dem Mittelwert auch die Varianz, meist in Form der Standardabweichung, korrigiert.

Die weiteren direkten Verfahren sind Verteilungs-basiert und korrigieren neben dem Mittelwert und der Varianz auch einzelne Quantile bzw. Perzentile und damit auch die Frequenz der Regentage. Da ihr Fokus auf der Korrektur von Quantilen liegt werden diese Verfahren oftmals als **Quantile-Mapping**-Verfahren bezeichnet. Trotz der meist sehr ähnlichen oder gar gleichen Formulierung wurden verschiedenste Bezeichnungen verwendet (vgl. Teutschbein und Seibert, 2012). Andere Bezeichnungen sind u. a. „*probability mapping*“ (z. B. Block et al., 2009), „*quantile-quantile mapping*“ (z. B. Boé et al., 2007), „*statistical downscaling*“ (z. B. Piani et al., 2010a), „*histogram equalization*“ (z. B. Rojas et al., 2011) oder „*daily translation*“ (z. B. Mpelasoka und Chiew, 2009). Die Grundidee des Quantile-Mappings ist die Verwendung der kumulierten Verteilungsfunktionen, um eine quantilbezogene Korrektur zu ermöglichen. Dabei werden die kumulierten Verteilungsfunktionen der historischen Mess- und Modelldaten zueinander ins Verhältnis gesetzt. Daraus wird eine Transferfunktion bestimmt, die die Korrektur der Quantile der prognostizierten Modelldaten erlaubt. Da sie direkt die prognostizierte Zeitreihe nutzt und korrigiert kann die Methode als ein quantilbezogenes *linear scaling* aufgefasst werden (Cannon et al., 2015). Die Quantile-Mapping-Verfahren umfassen u. a. *Quantile-Mapping auf Basis der Gamma-Verteilung*, *Cumulative distribution function transform (CDF-t)*, *Equidistant quantile matching (EDCDFm)* (vgl. Pierce et al., 2015), *Quantile-Mapping auf Basis einer hergeleiteten Transferfunktion* (z. B. Piani et al., 2010b), *Empirisches Quantile-Mapping*, *Detrended quantile mapping*, *Quantile Delta Mapping* (vgl. Cannon et al., 2015). Während *EDCDFm* eine additive Korrektur vornimmt, wird beim *quantile delta mapping* eine multiplikative Korrektur angewandt (Cannon et al., 2015).

Bei den indirekten Methoden haben Prädiktor und Antwortvariable nicht dieselbe physikalische Größe, stehen aber in einem Zusammenhang zueinander. Die linearen Methoden basieren meist auf linearen Regressionsmodellen. Zu ihnen gehören z. B. die *Multiple lineare Regressionsmodelle*, *Multiple linearen Regressionsmodelle mit Randomisierung*, *Analogon-Methode* oder die *Nächste-Nachbarn-Analogon-Methode*. Während lineare Regressionmodelle grundsätzlich auch für die Biaskorrektur angewendet werden können, werden die Analogon-Verfahren meist für die Analyse von Modellunsicherheiten verwendet und können damit eher zur Schätzung der Unsicherheit eines Korrekturverfahrens verwendet werden (Themekl et al., 2011).

Von einigen Autoren wurde bemängelt, dass univariate Biaskorrekturverfahren nur an einem Ort zu einer Zeit funktionieren und damit die räumliche und zeitliche Struktur der korrigierten Zeitreihe nicht adäquat wiedergeben (Vrac und Friederichs, 2015; Maraun, 2016). Aus diesem Grund wurden multivariate Biaskorrekturverfahren entwickelt. Vrac und Friederichs, 2015 stellen zwei bivariate Methoden vor: die *bivariate conditional correction* und eine Version der *Shaake-shuffle-Methode*.

Im Bereich der Klimamodellierung liegt der Fokus der Biaskorrekturverfahren besonders auf dem zeitlichen Aspekt und dem Trend über Jahrzehnte. Bei der Korrektur von Radardaten steht die Korrektur der Information an einem bestimmten Ort zu einer bestimmten Zeit im Vordergrund. Bei Radardaten ist besonders die Verteilung des Bias im Raum, also das ortsabhängige Bias von Interesse. Da als Korrekturdatensatz oftmals Stationsdaten zum Einsatz kommen, wird auch von Merging- bzw. Aneichungs-Verfahren gesprochen. Oftmals enthalten diese Korrekturverfahren Interpolationstechniken, um (korrigierte) Informationen für den gesamten Raum zu erhalten. Da diese nicht immer voneinander getrennt werden können, werden im Folgenden Methoden vorgestellt, die beides enthalten können.

McKee und Binns, 2016 geben eine Übersicht über Merging-Methoden von Radar- und Stationsdaten. Sie unterscheiden Biaskorrekturverfahren von Verfahren, die nicht nur das Bias reduzieren, sondern auch die Varianz der Residuen minimieren. Weitere Übersichten finden sich z. B. in Goudenhoofdt und Delobbe, 2009 und in Hanchoo Wong et al., 2012. Zu diesen zählen beispielsweise *Mean field bias correction*, *Static local bias correction and range dependent adjustment*, *Brandes spatial adjustment*, *Local bias correction with ordinary kriging*, *Range-dependent*

bias correction (vgl. Goudenhoofd und Delobbe, 2009), *Range dependent mean field bias correction*, *Local bias correction* (vgl. Hanchoo Wong et al., 2012), *Weighted multiple regression model* (z. B. Gilewski und Nawalany, 2018) oder *Conditional mean field bias correction* (z. B. Thorndahl et al., 2017)

Die Methoden *Bayesian data combination*, *Conditional merging*, *Kriging with external drift* und *Statistical objective analysis* werden von McKee und Binns, 2016 als Techniken genannt, die neben der Biaskorrektur die Residuen-Varianz minimieren.

Andere Interpolationsverfahren, die zusammen mit der Biaskorrektur angewandt werden können sind z. B. *multiquadratic surface fitting* (z. B. Cole und Moore, 2008), die *Inverse Distanzwichtung* (z. B. Tesfagiorgis et al., 2011), die *Thiessen-Polygon-Methode* oder der *Singularity approach* (vgl. McKee und Binns, 2016).

1.1.4 Anwendungsgebiete

Der Großteil der Biaskorrekturverfahren wurde im Bereich der Klimamodellierungen entwickelt, meist für die Größen Temperatur und Niederschlag. Einige Analysen beziehen sich dabei auf einzelne Länder (z. B. Maurer und Hidalgo, 2008, Michelangeli et al., 2009, Themeßl et al., 2011, Themeßl et al., 2012, Piani und Haerter, 2012, Cannon et al., 2015, Vrac und Friederichs, 2015), andere auf eine kontinentale (z. B. Li et al., 2010, Piani et al., 2010a, Piani et al., 2010b) oder globale Skala (z. B. Watanabe et al., 2012, Hempel et al., 2013). Eine Vielzahl an Studien verwendete Klimamodellierungen (mit Korrekturverfahren) für hydrologische Auswirkungsstudien (z. B. Dettinger et al., 2004, Boé et al., 2007, Leander und Buishand, 2007, Leander et al., 2008, Block et al., 2009, Mpelasoka und Chiew, 2009, Rojas et al., 2011, J. Chen et al., 2013, Fang et al., 2015, Pierce et al., 2015). Hashino et al., 2007 nutzte drei verschiedene Biaskorrekturverfahren zur Vorhersage von Abflussvolumina in den USA.

Der andere Hauptteil der Studien wendete Biaskorrekturverfahren im Rahmen der Korrektur von Radar-Niederschlagsdaten an, um diese mithilfe von Stations-Niederschlagsmessungen zu korrigieren. Oftmals handelte es sich hierbei um das Bodenradar-Netz einzelner Länder (z. B. Smith und Krajewski, 1991, Anagnostou et al., 1999, Villarini et al., 2008, Germann et al., 2009, Goudenhoofd und Delobbe, 2009, Hanchoo Wong et al., 2012, Zhang et al., 2011, Yoo et al., 2014, J. Chen et al., 2013, Ozkaya und Akyurek, 2019), andere Studien fokussierten sich auf satellitenbasierte Wetterradare (z. B. Tesfagiorgis et al., 2011, Wang et al. 2021). Wieder andere Analysen beschäftigen sich mit der Echtzeit-Korrektur von Radar-Niederschlagsdaten (z. B. D.-J. Seo et al., 1999, D.-J. Seo und Breidenbach, 2002, Chumchean et al., 2006).

1.2 Nutzen von Niederschlagsradaren

Während Radarsysteme ursprünglich während des Zweiten Weltkrieges zur Luftraumüberwachung entwickelt wurden, stellte man den störenden Effekt des Niederschlags auf die Messqualität fest. Aus der Idee, diesen Effekt gezielt zur Überwachung des Wetters einzusetzen, entstanden die ersten Wetterradare (Berne und Krajewski, 2013).

Folgend sollen einige Funktionsweisen und Charakteristika von Radardaten näher beleuchtet werden. Eine Darstellung der für die Arbeit verwendeten Radare folgt im Kapitel „3.1.1 Radardaten“.

1.2.1 Funktionsweise eines Niederschlagsradars

Ein Wetterradar tastet aktiv unter Energieeinsatz mithilfe elektromagnetischer Wellen die Atmosphäre ab. Dabei soll die entfernungsbezogene Erfassung von Objekten in der Atmosphäre ermöglicht werden. Dies wird auch an der Bezeichnung *RADAR* deutlich, die eine englische Abkürzung für „Radio Detection and Ranging“ darstellt. Das Verfahren kann als kontinuierliches Verfahren mit zeitlich und räumlich hoher Auflösung klassifiziert werden. Da nicht direkt der

Niederschlag, sondern elektromagnetische Wellen gemessen werden, stellt das Verfahren eine indirekte Niederschlagsmessung dar (Winterrath et al., 2017).

Die eigentliche Antenne, die durch ein Radom (Radarkuppel) vor Witterungseinflüssen geschützt wird, ist in Azimut und Elevation beweglich. Die Bewegung des Radars in diese Richtungen ist Teil der festgelegten Scanstrategie eines jeden Radars. Für die Niederschlagsmessungen folgt die Antenne dem Verlauf der Orographie vor Ort und bestimmten Höhenwinkeln (Elevationen) (Winterrath et al., 2017). Je nach Bauart haben Wetterradare eine radiale Auflösung von 100 bis zu 1000 Meter, die Weite des Radarkegels beträgt meist etwa 1 Grad. Ein Radar nutzt Wellen im Mikrowellen- oder Radiowellen-Bereich. Abhängig vom verwendeten Teilbereich werden verschiedene Frequenzbänder unterschieden. Die am häufigsten genutzten sind das S-Band, C-Band und X-Band, mit jeweils unterschiedlicher Wellenlänge und Frequenz (Berne und Krajewski, 2013).

Beim Messvorgang erzeugt ein Radar hochfrequente Impulse elektromagnetischer Wellen, die durch einen Reflektor gebündelt und anschließend in eine bestimmte Raumrichtung gesendet werden. Das Messverfahren stellt also einen Volumenscan der Atmosphäre dar. Ein Teil dieser Strahlung wird an Hydrometeoren (Niederschlagsteilchen) reflektiert und ein Anteil dieses reflektierten Signals am Radar detektiert und verstärkt (Winterrath et al., 2017). Durch die Richtung und die Laufzeit der Welle kann die Position der Hydrometeore bestimmt werden. Die Intensität des reflektierten Signals gibt Aufschluss über die Intensität des Niederschlagsereignisses. Die einzelnen Pulse werden gemittelt, um Effekte positiver oder negativer Interferenz der Wellen zu beseitigen. Gemessen wird bedingt durch das Raumvolumen eine Größenverteilung mehrerer Hydrometeore. Hydrometeore können dabei Wolkentröpfchen, Regentropfen, Schneeflocken, Eiskristalle oder Hagelkörner sein. Die Größe beträgt etwa 0,1 bis 10 mm für Regentropfen, bis zu 25 mm für Schneeflocken, bis zu 50 mm für Hagelkörner (Berne und Krajewski, 2013).

Das gemessene reflektierte Signal wird als Reflektivitätsfaktor (Z) festgelegt und daran eine optimale Wellenlänge angepasst, die die Detektierbarkeit maximieren und störende Dämpfungseffekte minimieren soll. Dieser Faktor wird so festgelegt, dass er nur von den Eigenschaften der Hydrometeore und nicht von denen des genutzten Radars abhängt. Die aus diesem Kalibrierungsvorgang gewonnene Beziehung zwischen Reflektivitätsfaktor und Niederschlagsrate (R) wird Z-R-Beziehung genannt. Dabei wird ein potentieller Zusammenhang angenommen (Berne und Krajewski, 2013). Für eine genauere und realistischere Bestimmung der Niederschlagsmenge werden die bestimmten Radarniederschläge an Stationsniederschläge angeeicht (Winterrath et al., 2017).

1.2.2 Weiterentwicklungen und Korrekturen

Im Laufe der Zeit kam es zu einigen technischen Verbesserungen der Radartechnologie. Zwei Weiterentwicklungen haben dabei zu einer deutlichen Verbesserung der Datenqualität geführt, die Einführung von Doppler-Radaren und die von Dual-Pol-Radaren (Winterrath et al., 2017). Bei der Verwendung von Doppler-Radaren, die sich den Doppler-Effekt zunutze machen, kann zwischen beweglichen Hydrometeoren und nicht-beweglichen (Stör)Objekten unterschieden werden, was zur Qualitätsverbesserung des Radarproduktes beiträgt. Auch kann so auf Windverhältnisse geschlossen werden (Berne und Krajewski, 2013, Winterrath et al., 2017). Dual-Pol-Radare messen, im Gegensatz zu den lange üblichen Single-Pol-Geräten, in zwei Polarisationssebenen. Mithilfe dieser polarimetrischen Messungen können mikrophysikalischen Eigenschaften der Niederschläge untersucht werden. Besonders die Form der Hydrometeore ist dabei wichtig, da viele Hydrometeore nicht kugelförmig, sondern abgeflacht sind und damit verschiedene Polarisationssebenen besitzen. So kann die Form Aufschluss über die Phase und Größenverteilung der Hydrometeore geben. Zudem helfen polarimetrische Messungen meteorologische von nicht-meteorologischen Signalen (z. B. Insekten, Vögel, Flugzeuge, Gebäude) zu unterscheiden (Berne und Krajewski, 2013, Grazioli et al., 2015, Winterrath et al., 2017).

Bedingt durch das Messverfahren mittels elektromagnetischer Wellen müssen die Rohdaten auf gewisse Phänomene und Unsicherheiten korrigiert werden, bevor sie über die Z-R-Beziehung

in Niederschlagsraten umgewandelt werden. Folgend werden die wichtigsten Quellen für Messunsicherheit aufgeführt, eine Beschreibung der Korrekturverfahren findet nicht statt, kann aber z. B. Winterrath et al., 2017 entnommen werden.

Eine Unsicherheit stellt *ground clutter* dar, damit sind unerwünschte starke Reflexionen des Radarstrahls an der Topographie, wie z. B. Gebäuden, Bergen, hohen Türmen, Windturbinen, aber auch an Flugzeugen, Vögeln und Insekten, gemeint (Berne und Krajewski, 2013, Winterrath et al., 2017, Thorndahl et al., 2017). Eine teilweise oder vollständige Verdeckung von Gebieten durch Hindernisse wird als *beam blocking* oder *Negativspeichen* bezeichnet. Diese Gebiete sind für den Radarstrahl nicht oder nur schwer zugänglich, was zu einer Reduzierung des Signals führt. Nicht-bewegliche Objekte können mithilfe von Doppler-Radaren herausgefiltert werden (Berne und Krajewski, 2013, Winterrath et al., 2017). *Positivspeichen* sind Gebiete, die durch externen Energieeintrag z. B. von WLAN-Sendern oder von direkter Sonnenstrahlung bei Sonnenaufgang oder -untergang, erhöhte Signalwerte aufweisen (Winterrath et al., 2017).

Weitere Unsicherheiten sind die entfernungs- und höhenabhängige Signalabschwächung, die verschiedene Ursachen hat. Grundsätzlich wird mit zunehmender Entfernung der Radarkegel und damit das gemessene Volumen größer. Bei gleicher Energiemenge wird das reflektierte Signal so schwächer. Zusätzlich wird ein Teil des Radarstrahls an Hydrometeoren gestreut oder absorbiert, das reflektierte Signal muss also, um diese Signalabschwächung korrigiert werden, um Unterschätzungen der Niederschlagsrate zu vermeiden (Berne und Krajewski, 2013). Durch die Elevationswinkel der Messungen breitet sich der Radarstrahl nicht rein horizontal aus, sondern steigt auch vertikal an. Bei weiten Entfernungen befindet sich der Radarstrahl in Höhen von 1000 bis 2000 Metern (Berne und Krajewski, 2013). Wenn Niederschläge mit geringer vertikaler Ausdehnung nur teilweise oder gar nicht mehr erfasst werden, wird dies als *Overshooting* bezeichnet. Zusätzlich führt die Variabilität des Niederschlags in vertikaler Richtung zu größeren Unterschieden zum Niederschlag in Bodennähe. Da mit größerer Höhe der Anteil an festen Hydrometeoren zunimmt und diese einen geringeren Anteil des Radarstrahls reflektieren, wird das Signal systematisch unterschätzt (Berne und Krajewski, 2013, Winterrath et al., 2017).

Generell haben Luftfeuchte und Lufttemperatur Einfluss auf die Ausbreitung des Radarsignals und können die Lokalisierung der Hydrometeore erschweren. Eine weitere Abschwächung des Signals findet bei großer Niederschlagsintensitäten, z. B. bei Starkregen und schmelzendem Schneefall statt (Thorndahl et al., 2017). Bei polarimetrischen Radaren ist dies aber ein geringeres Problem (Berne und Krajewski, 2013). Ein besonderes Phänomen, das *bright band* genannt wird, tritt auf, wenn sich in dem abgetasteten Volumen eine Schmelzschicht befindet, was oft zu einer Überschätzung des Niederschlags führt Fabry und Zawadzki, 1995. Grundsätzlich kumulieren sich die abschwächenden Effekte auf, wodurch die Abschwächung bei größeren Distanzen und stärkeren Niederschlagsintensitäten größer ist (Berne und Krajewski, 2013).

Damit die zu erwartenden Niederschlagscharakteristiken erkennbar sind, ist ein räumlich dichtes Netz an Radaren nötig (Winterrath et al., 2017). Eine weitere Unsicherheit stellt die Parametrisierung der Z-R-Beziehung dar, die auf der Tröpfchenverteilung und Tröpfchenanzahl (raindrop size distribution) basiert (Berne und Krajewski, 2013, Winterrath et al., 2017).

1.2.3 Potentiale von Radarprodukten

Eins der größten Potentiale von Radardaten liegt in der flächenhaften Verfügbarkeit von Daten. Mit dem Ziel die Niederschläge einer Region besser untersuchen zu können, wurde in einigen Studien aus Reflektivitätsdaten selbst ein Radar-Datensatz erstellt (z. B. Smith et al., 2012, Goudenhoofd und Delobbe, 2016). Eine andere Möglichkeit stellt die Nutzung bereits prozessierter Daten in Form fertiger Niederschlagsschätzungen dar. Durch die, über die Jahre wachsende, Verfügbarkeit von Radarmessungen stieg das Interesse Radar-Niederschlagsschätzungen für klimatologischen Analysen zu verwenden. Saltikoff et al., 2019 geben eine Übersicht über Radarprodukte einzelner Länder und die Potentiale dieser für klimatologische Untersuchungen. Ein Beispiel einer solchen Radar-Klimatologie ist das von Fairman Jr et al., 2015 entwickelte Produkt für Großbritannien und Irland. Neben der klimatologischen Nutzung gibt es weitere

Anwendungspotentiale. Einige Beispiele hierfür sind die Untersuchung der Regen-Erosivität (z. B. Fischer et al., 2018, Auerswald et al., 2019) und der räumlichen Ausdehnung von Niederschlagsereignissen in Deutschland (z. B. Lengfeld et al., 2019) oder die Nutzung polarimetrischer Radarmessungen für die Klassifizierung von Hydrometeoren in der Schweiz (z. B. Grazioli et al., 2015).

Da die Niederschlagsschätzung mittels Wetterradaren ein indirektes Messverfahren darstellt, werden Radardaten zur Plausibilisierung bzw. Validierung oftmals mit Stationsmessungen verglichen (z. B. Kitchen und Blackall, 1992, Jessen et al., 2005, Cole und Moore, 2008, B.-C. Seo und Krajewski, 2010, Hazenberg et al., 2011, Meißner et al., 2012, Abon et al., 2016). Andere Autoren verglichen Bodenradare und Wettersatelliten (z. B. Crisolago et al., 2018, Ramsauer et al., 2018, Pejčic et al., 2018, Pejčic et al., 2022) oder alle drei Verfahren (z. B. Gilewski und Nawalany, 2018, Kidd et al., 2012). Im folgenden Abschnitt soll auf einige dieser Aspekte eingegangen werden.

1.2.4 Vor- und Nachteile von Radardaten

Noch vor Wetterradaren sind Niederschlagsstationen die wohl häufigste und bewährteste Technik Niederschläge zu messen. Zu den größten Vorteilen zählen die direkte Messweise der Niederschlagsrate und die langen Messzeiträume von teils mehr als 100 Jahren (Lengfeld et al., 2020). Einer der größten Nachteile von Stationsmessungen hingegen ist die oftmals geringe räumliche Dichte an Niederschlagsstationen mit einer stündlichen oder besseren zeitlichen Auflösung (Saltikoff et al., 2019). Hinzu kommt, dass Niederschlagsmessungen einer Station nur bis zu einer bestimmten Entfernung vom Stationsstandort für die Fläche repräsentativ sind. Diese autokorrelative Distanz genannte Entfernung nimmt für längere Zeiträume zwar meist zu (Villarini et al., 2008, Lengfeld et al., 2020), ist für Starkniederschläge aber oftmals klein (Lengfeld et al., 2020).

Lengfeld et al., 2020 fanden, dass zwischen 2001 und 2018 nur 17,3 % der stündlichen Starkniederschläge in Deutschland durch Stationsdaten erfasst wurden, im Vergleich zu 81,8 % der täglichen Ereignisse. Für Großbritannien zeigte sich ein ähnliches Ergebnis, im Jahr 2014 wurde nur 36,6 % der stündlichen Ereignisse, aber 52 % der täglichen Ereignisse sowohl von Niederschlagsstationen als auch Radaren detektiert (Lengfeld et al., 2020). Generell stellten sie fest, dass mit abnehmender Ereignisdauer zwischen einer Stunde und 72 Stunden die Anzahl der nicht-erfassten Ereignisse durch das Stationsnetzwerk zunimmt. Der Grund hierfür ist, dass die räumliche Ausdehnung der Ereignisse abnimmt (Lengfeld et al., 2020).

Während stündliche Starkniederschläge meist durch lokale konvektive Systeme entstehen, werden tägliche Niederschlagsereignisse größtenteils durch große Frontensysteme verursacht. Am Beispiel von Norddeutschland zeigt sich, dass größere konvektive Ereignisse vermehrt in Gebieten entstehen, die durch die Orographie abgeschattet sind (Nordost-Deutschland) wohingegen Nordwest-Deutschland sehr von Fronten geprägt ist (Lengfeld et al., 2019). Die typische mittlere horizontale Ausdehnung von täglichen Niederschlagsereignissen beträgt in Deutschland dabei 68 Kilometer, die von stündlichen Ereignissen im Mittel nur 17 Kilometer (Lengfeld et al., 2019). Das entspricht ungefähr der mittleren Distanz zwischen Niederschlagsstationen des Deutschen Wetterdienstes. Damit kann nicht zwingend gewährleistet werden, dass kleine Starkregenereignisse durch die Stationen erfasst werden (Lengfeld et al., 2019). Besonders für kurzzeitige Extremniederschläge von einer Stunde bis sechs Stunden Dauer ist die Nutzung von Radarmessungen damit eine Notwendigkeit (Lengfeld et al., 2020).

Ein Vorteil von Radardaten ist die großflächige räumliche Abdeckung (Saltikoff et al., 2019), wodurch sie das Bindeglied zwischen Stationsnetzwerken und Satelliten-Produkten darstellen (Lengfeld et al., 2020). Außerdem zeichnen sich Niederschlagsradare durch eine relativ kleinskalige räumliche und zeitliche Auflösung aus (Berne und Krajewski, 2013, Thorndahl et al., 2017, Saltikoff et al., 2019). Dies ermöglicht sowohl die Erfassung der vertikalen und horizontalen Ausdehnung von Wolken und Gewitterzellen (Winterrath et al., 2017) als auch die Beschreibung ihrer Dynamik, Variabilität und Zuggeschwindigkeit (Berne und Krajewski, 2013). Des Weiteren führt die Nutzung von Niederschlagsradaren zur Bereitstellung mehrerer neuer Parameter (Saltikoff et al., 2019). Mit diesen können die typische räumliche Ausdehnung, die Dauer und

Niederschlagstypen untersucht werden. Insbesondere die Dual-Polarisations-Technologie erlaubt die Analyse der Mikrophysik von Hydrometeoren, wie z. B. ihrer Größe, Form, Orientierung und Phase (Lengfeld et al., 2020). Da es sich bei dem Messverfahren um einen Volumenscan handelt, helfen Wetterradare bei der Beschreibung und Erforschung drei-dimensionaler Prozesse und Strukturen von Wetterereignissen, wie z. B. Konvektion, die Bildung von Stauwolken, Hagel, Hurrikans oder Kollisions-Koaleszenz-Prozessen in warmen Wolken. Dabei lassen sich diese Charakteristika durch Wetterradare mit höherer Auflösung und Genauigkeit beschreiben als mit Stationen oder Wettersatelliten (Saltikoff et al., 2019, Lengfeld et al., 2020).

Diese Vorteile führten dazu, dass Radardaten die Echtzeitüberwachung und das Nowcasting des Wetters, insbesondere von Extremereignissen ermöglichten (Thorndahl et al., 2017, Winterrath et al., 2017), die Verbesserung von Warnsystemen (von z. B. Hagel, Tornados oder Schneestürmen) erlaubten (Saltikoff et al., 2019) und für die Modellierung und Vorhersage von Sturzfluten oder für die Beschreibung der hydrometeorologischen Begebenheiten bei Hochwassern genutzt werden können (Berne und Krajewski, 2013). Aber auch das urbane Wassermanagement kann auf diese Weise mit hochaufgelösten Daten unterstützt werden (Berne und Krajewski, 2013). Für Gebiete ohne hydrologischen Pegel stellen Radardaten die einzige Möglichkeit dar, kleinskalige Extremereignisse für die Hochwasservorhersage zu erfassen, wie sich am Beispiel der USA zeigte (Lengfeld et al., 2020).

Neben all diesen Vorteilen ist die Nutzung von Niederschlagsradaren aber auch mit einigen Nachteilen verbunden. Für klimatologische Analysen ist eine Reprozessierung zur Homogenisierung der Zeitreihen nötig (Lengfeld et al., 2020), um veränderte Radarstandorte, Änderungen der räumlichen Abdeckung des Radarschirms oder Anpassung der Prozessierungstechniken zu berücksichtigen, was in vielen Ländern noch nicht durchgeführt wurde (Saltikoff et al., 2019). Oftmals besteht (wenn überhaupt) nur ein eingeschränkter, meist kostenpflichtiger Zugang zu Radardaten (Saltikoff et al., 2019). Die Radar-Zeitreihen sind meist recht kurz, insbesondere die von Dual-Pol-Radaren (Winterrath et al., 2017, Saltikoff et al., 2019, Lengfeld et al., 2020), die meist ein geringeres Bias aufweisen als die Single-Pol-Radare (Berne und Krajewski, 2013). Zudem ist das Radarnetz teils sehr klein (Berne und Krajewski, 2013), was die Unterschätzung des Niederschlags in Regionen mit weiter Distanz zum Radar verstärkt (Fairman Jr et al., 2015).

Auch können bei Radarmessungen eine räumliche und zeitliche Struktur der Unsicherheiten bestehen. Eines der größten Probleme der Niederschlagsradare ist die Unterschätzung des Niederschlags in bestimmten Orten oder Situationen. Besonders bei Schneefall als auch in Bergregionen bzw. Regionen mit komplexer Topographie wird der Niederschlag unterschätzt. Schließlich werden auch oftmals Starkniederschläge, z. B. aufgrund eines nassen Radoms, unterschätzt, sodass die Radardaten ohne Biaskorrektur nur sehr limitiert verwendbar sind (Berne und Krajewski, 2013). Ein möglicher Grund hierfür kann eine konstante Z-R-Beziehung sein, die Veränderungen in der Regentropfengröße bei zunehmender Niederschlagsintensität nur mangelhaft abbilden kann (Schleiss et al., 2020). Meist wird in Studien eine Unterschätzung des Stationsniederschlags durch die Radardaten festgestellt (z. B. Hazenberg et al., 2011, Smith et al., 2012, Fairman Jr et al., 2015, Goudenhoofd und Delobbe, 2016, Schleiss et al., 2020), teilweise auch eine Überschätzung (z. B. Fairman Jr et al., 2015, Haberlandt und Berndt, 2016).

1.2.5 Untersuchungen von Extremniederschlägen mittels Radardaten

Wie gefährlich Starkniederschläge sein können, zeigte sich beispielsweise am 29. Mai 2016 in Braunsbach, als Starkniederschläge zu einer Sturzflut führten, die große Schäden anrichtete (LGRB, 2023). Daher ist es nicht verwunderlich, dass sich bereits mehrere Autoren mit der Nutzung von Radardaten für die Vorhersage und Analyse von Starkniederschlägen auseinandergesetzt haben. Um das Starkregenrisiko besser verstehen zu können führte der Deutsche Wetterdienst (DWD) eine Reprozessierung der Radardaten durch, u. a. um die Erfassung des Vorkommens und Auftretens konvektiver Extremniederschläge besser verstehen zu können. Diese Radarklimatologie (RADKLIM) wurde bereits in einigen Studien verwendet (Winterrath et al., 2017).

Lengfeld et al., 2021 nutzten den RADKLIM-Datensatz zur Erstellung eines Kataloges von

Starkniederschlägen zwischen 2001 und 2020 in Deutschland. Dabei untersuchten sie Ereignisse von einstündiger bis 72-stündiger Dauer. Die Studie ergab, dass die Anzahl an Events und die betroffene Fläche sehr zwischen den Jahren variiert. Trotz der insgesamt geringeren Niederschläge im Jahr 2018, war dieses Jahr das mit den meisten Starkniederschlägen. Während langfristige Ereignisse mehr der Orographie folgen, sind kurzfristige zufällig über Deutschland verteilt. Die Studie war Teil des „Starkregen-Projekts“ des Gesamtverbandes Deutscher Versicherer (GDV) und wurde vom DWD ausgeführt. Im Fachbericht dieses Projektes nennt der GDV insbesondere Starkregenereignisse mit Dauerstufen von maximal neun Stunden als die Ereignisse mit der größten Schadenshäufigkeit. Langanhaltende Dauerregen führen dagegen vermehrt zu Schäden durch Flussausuferungen. Insbesondere Ereignisse mit Dauern zwischen zwei und drei Stunden verursachten die größte mittlere Schadenshäufigkeit (GDV, 2019).

Lengfeld et al., 2020 untersuchten kleinskalige Starkniederschläge in Deutschland, Großbritannien und den USA mithilfe von Radardaten und stellen Potenziale und Herausforderungen bei der Nutzung von Radardaten dar. Dabei betonten sie besonders den Nutzen von Radardaten für die Analyse stündlicher Extremereignisse. Insgesamt zeigte sich bei den Extremniederschlägen eine hohe räumliche und zeitliche Variabilität, oftmals hatten diese aber eine Ausdehnung von weniger als einem Kilometer und eine Dauer von unter einer Stunde.

Winterrath et al., 2019 analysierten Starkniederschläge in Deutschland mithilfe der vom DWD entwickelten Radarklimatologie (RADKLIM). Sie stellten fest, dass extreme Starkregenereignisse nicht zwingend an die Topographie gebunden sind, sondern überall auftreten können.

Auch Pöschmann et al., 2021 verwendeten den RADKLIM-Datensatz, um die Beziehung zwischen Niederschlagshöhe und Dauer bei Extremniederschlägen in Deutschland zu untersuchen. Hierbei stellte sich heraus, dass die Beziehung zwischen Niederschlagshöhe und -dauer im Gegensatz zu Stationsdaten keinem klaren Skalierungsverhalten folgt. Bei Radardaten waren drei verschiedene Phasen feststellbar. Unterhalb einer Stunde steigt die Niederschlagshöhe stark an, bei Dauern zwischen einer Stunde und einem Tag bleibt sie weitestgehend konstant, bei Dauern über einem Tag steigt sie wiederum schwach an.

Für Nordrhein-Westfalen untersuchten Jessen et al., 2005 fünf Starkregenereignisse mithilfe von Stationsdaten und Rohdaten von drei C-Band-Radaren und einem X-Band-Radar. Hierfür wendeten sie vier Korrekturmethode an, eine eventbasierte Mean-field-bias-Korrektur und drei Arten einer Korrektur-Matrix. Für diese wurde für jeden Stationsstandort Korrekturfaktoren bestimmt und diese anschließend mithilfe der drei nächsten Stationen basierend auf der quadratischen reziproken Distanz räumlich interpoliert. Eine zweite Variante berücksichtigte nicht nur das Pixel am Stationsstandort, sondern das Mittel aus diesem und den acht umliegenden Pixeln, die letzte Variante ist die Anwendung der Korrektur-Matrix nach Korrektur der Reflektivitätsmessungen. In ihrer Analyse stellten sie fest, dass die Nutzung des Mittels aus neun Pixeln bei sehr kleinskaligen Events problematisch ist, da diese so stark unterschätzt werden.

Neben der zeitlichen und räumlichen Analyse von Starkregenereignissen mithilfe von Radardaten wurde bereits von mehreren Autoren versucht Radardaten für eine Extremwertanalyse zu verwenden, meist mit der Anwendung eines Biaskorrekturverfahrens verbunden. Panziera et al., 2016 nutzten Niederschlagsradardaten für eine Extremwertanalyse, um Grenzwerte für ein neues automatisches Warnsystem für Starkniederschläge in der Schweiz zu gewinnen. Es zeigte sich, dass die Jährlichkeiten der Ereignisse einer großen räumlichen Variabilität unterliegen und dass die räumliche Verteilung dieser Ereignisse sehr von der Dauerstufe abhängt.

Overeem et al., 2009a nutzten Niederschlagsradardaten für eine Extremwertanalyse von Starkniederschlägen in den Niederlanden. Dazu korrigierten sie den Bias mithilfe von Stationsdaten. Als Korrekturverfahren wurde ein Korrektur-Feld entwickelt, indem zu jedem Pixel die Abweichung zum mean field bias berechnet wurde. Dabei stellten sie eine grundsätzliche Eignung der Radardaten zur Untersuchung des Verhältnisses zwischen Niederschlagshöhe und -dauer fest.

Rabiei und Haberlandt, 2015 verwendeten Quantile-Mapping zur Biaskorrektur von Wetterradardaten im Raum Hannover. Dazu korrigierten sie das Radarbild jedes Zeitschritts mithilfe der Stationsdaten unter der Annahme, dass das räumliche Bias, und nicht das zeitliche Bias, in

den Radardaten dominiert. Anschließend wandten sie mit conditional merging und kriging with external drift zwei Interpolationsverfahren an. Dabei stellten sie eine bessere Performance im Winter als im Sommer fest. Generell fanden sie, dass die Nutzung von Quantile-Mapping die Qualität der Radardaten als auch die Performance des Interpolationsverfahrens verbesserte. Da in ihren Analysen die Radarqualität im Winter besser war als im Sommer, ermittelten sie eine größere Verbesserung der Radarqualität nach der Biaskorrektur im Sommer. Daraus schlossen sie, dass die Biaskorrektur im Sommer wichtiger sei. Diesen saisonalen Unterschied begründen sie mit den unterschiedlichen Niederschlagstypen, da konvektive Ereignisse vermehrt im Sommer und Frontensysteme vermehrt im Winter überwiegen.

Haberlandt und Berndt, 2016 verwendeten Radardaten zur Berechnung von Bemessungsniederschlägen für Starkregenereignisse. Die Radar-Reflektivitäten korrigierten sie nach einem einfachen Schema. Dabei verglichen sie Stationsmessungen, Radarmessungen und biaskorrigierte Radarmessungen. Als Merging- bzw. Interpolationsverfahren kamen die Nächste-Nachbarn-Methode, ordinary kriging und conditional merging zum Einsatz. Verglichen mit den Stationsdaten führte die Verwendung der unkorrigierten Radardaten zu einer Überschätzung, die Verwendung der korrigierten Radardaten zu einer Unterschätzung der Starkniederschläge. Die Überschätzung begründen sie mit der wohl mangelhaften Fehlerkorrektur der Radar-Reflektivitäten, die Unterschätzung hingegen mit dem Glättungseffekt der Merging- bzw. Interpolationstechniken.

Für Baden-Württemberg nutzten Hänsler und Weiler, 2022 biaskorrigierte Radardaten zur Berechnung von Bemessungsniederschlägen für Starkregenereignisse. Der Fokus lag auf kurzen bis mittellangen konvektiven Starkregenereignissen, weswegen nur Daten der Sommermonate ausgewählt wurden. Ein Event mit einer Jährlichkeit von 100 Jahren stellte dabei die maximale Wiederkehrperiode dar. Um eine ausreichend lange Zeitreihe der Radardaten zu erhalten, wurde ein Regional-Subsampling-Verfahren angewandt. Nach einer Extremwertanalyse von 350 berechneten Bemessungsniederschlägen, wurde diese mithilfe von Quantile-Mapping korrigiert.

2 Problemstellung und Zielsetzung

Das große Schadpotential von Extremniederschlägen war bereits mehrfach der Grund, warum Analysen dieser veranlasst wurden (z. B. Jessen et al., 2005, Winterrath et al., 2017).

Aufgrund ihrer direkten Messung des Niederschlags sind die Daten der Niederschlagsstationen auf den ersten Blick vorzuziehen, doch durch den teils sehr lokalen Charakter der Starkniederschläge und ihre kurze Dauer (De Lannoy et al., 2005) werden diese von den Niederschlagsstationen nicht zwingend erfasst (Lengfeld et al., 2020).

Trotz der technischen Fortschritte stellten viele Studien eine Unterschätzung der Niederschlagsrate im Vergleich zu Stationen fest, insbesondere bei Starkregen (z. B. Hazenberg et al., 2011, Smith et al., 2012, Fairman Jr et al., 2015, Goudenhoofd und Delobbe, 2016, Schleiss et al., 2020), was wohl zu einem nicht unwichtigen Anteil an dem unterschiedlichen Messvolumen liegt (Schleiss et al., 2020).

Radardaten sind also ideal geeignet, um die räumlich teils sehr kleinen Extremereignisse zu detektieren und zu monitoren, als indirekte Messung sind aber umfangreiche Korrekturen nötig, um die Radardaten nutzen zu können (Berne und Krajewski, 2013).

Da die Vorhersage und Analyse von Starkregenerereignissen wichtig sind, um die Gefahr für den Menschen zu verringern und große Schäden zu vermeiden, müssen Verfahren entwickelt werden den bestehenden Bias zu reduzieren und bestenfalls zu beseitigen. Idealerweise kombiniert ein solches Verfahren die Messgenauigkeit der Stationen bezüglich der Niederschlagsrate mit der besseren räumlichen und zeitlichen Auflösung der Niederschlagsradare.

Für die Analyse dieser seltenen Ereignisse kommt erschwerend hinzu, dass die Zeitreihen der Radardaten meist noch sehr kurz sind und es daher fraglich ist, ob genügend Extremereignisse mithilfe der Radardaten erfasst wurden (Lengfeld et al., 2020).

Trotzdem versuchten einige Studien Radardaten zur Berechnung von Bemessungsniederschlägen zu verwenden (Overeem et al., 2009a, Haberlandt und Berndt, 2016, Hänsler und Weiler, 2022). Zumeist wurden dafür direkt die Reflektivitätsdaten der jeweiligen Radarstandorte und eine fixe Parametrisierung der Z-R-Beziehung verwendet (Jessen et al., 2005, Overeem et al., 2009a, Rabiei und Haberlandt, 2015 und Haberlandt und Berndt, 2016), nur selten kamen dabei RADKLIM, als vom DWD bereits reprozessiertes Radarprodukt zum Einsatz (Hänsler und Weiler, 2022).

Während Overeem et al., 2009a die grundsätzliche Eignung von biaskorrigierten Radardaten für die Berechnung von Extremwertstatistiken feststellte, kamen Haberlandt und Berndt, 2016 zu dem Schluss, dass ihr entwickeltes Verfahren so keinen zusätzlichen Nutzen zur Berechnung von Bemessungsniederschlägen darstellt. Gleichzeitig stellte Overeem et al., 2009a fest, dass sein (hauptsächlich auf einer mean-field-bias-Korrektur basierendes) Verfahren noch nicht ausreicht, um das Bias für Dauern unter einem Tag zu korrigieren.

Entgegen der häufig zur Korrektur von Radardaten verwendeten Verfahren, wie der Mean-field-bias-Korrektur oder der verschiedenen Kriging-Varianten, nutzten Rabiei und Haberlandt, 2015 und Hänsler und Weiler, 2022 zur Biaskorrektur einen Quantile-Mapping-Ansatz basierend auf einer theoretischen Verteilungsfunktion. Ein Vorteil dieses Verfahrens ist, dass das räumliche Muster des Radarbilds nicht verändert wird (Rabiei und Haberlandt, 2015, Hänsler und Weiler, 2022).

Rabiei und Haberlandt, 2015 konzentrierten sich auf die stündliche Korrektur eines jeden Radarbildes, ihr Ziel lag hauptsächlich darin ein Verfahren zu entwickeln, dass einer Echtzeit-Korrektur der Radardaten nahekommt. Obwohl die Biaskorrektur selbst ohne eine Interpolation auskommt, dienten die biaskorrigierten Radardaten in ihrer Analyse als zusätzliche Information

für mehrere Interpolationsverfahren. Auch umfasste der betrachtete Zeitraum von 2004 bis 2008 nur relativ wenig Jahre.

Die Studie von Hänsler und Weiler, 2022 hingegen diente primär der Berechnung von Bemessungsniederschlägen mithilfe von biaskorrigierten Radardaten, um so von der räumlichen Abdeckung der Radarmessungen profitieren zu können. Der Fokus lag also nicht auf der Korrektur realer Extremereignisse, sondern auf der Erstellung eines Datensatzes von Niederschlagsereignissen, die für eine Extremwertanalyse und schließlich für die Berechnung von Bemessungsniederschläge verwendet werden kann.

Durch die Anwendung einer *regional frequency analysis* wird, ähnlich wie bei Overeem et al., 2009a, das Problem der recht kurzen Radarzeitreihen umgangen. Somit basiert die Analyse nicht auf nur wenigen Stationen mit einer ausreichend langen Zeitreihe, deren Resultate folgend auf die ganze Fläche interpoliert werden muss. Ein entscheidender Vorteil dieses Verfahrens ist es also, dass es die Berechnung von Bemessungsniederschlägen erlaubt ohne eine räumliche Interpolation anwenden zu müssen, die oftmals eine der Hauptfehlerquellen bei der Erstellung eines regionalen Datensatzes mithilfe von Radardaten ist (Hänsler und Weiler, 2022).

Die Ergebnisse dieser Studie ergaben räumliche Muster, die mit der theoretischen Erwartung mehr übereinstimmten als reine auf Stationsdaten basierende Interpolationsprodukte. Trotzdem zeigt sich gerade im Hinblick auf die Biaskorrektur, dass diese noch nicht ausreicht, um das Bias vollständig zu beseitigen. Als einer der Gründe wurde dabei die additive Korrektur der Niederschläge genannt, und vorgeschlagen direkt die Radarniederschläge statt der Parameter der Generalisierten Paretoverteilung zu korrigieren (Hänsler und Weiler, 2022).

Eine weitere Einschränkung ist, dass sich die hier vorgestellten Studien zur Biaskorrektur von Radardaten auf eine gewisse Region, wie Hannover (Rabiei und Haberlandt, 2015, Haberlandt und Berndt, 2016) oder auf ein Bundesland, wie Baden-Württemberg (Hänsler und Weiler, 2022) fokussieren.

Diese Wissenslücken sollen mit dieser Arbeit weiter untersucht werden, der räumliche Fokus liegt dabei auf ganz Deutschland. Das Hauptziel dieser Arbeit ist es mehrere Biaskorrekturverfahren zu entwickeln, zu testen und miteinander zu vergleichen, um das Bias zwischen Radar- und Stationsdaten zu reduzieren. Hierbei knüpft die Arbeit an die Vorarbeiten von Hänsler und Weiler, 2022 an. Im Fokus der Arbeit stehen Starkniederschläge von maximal sechs-stündiger Dauer. Dazu werden „künstliche“ Niederschlagsereignisse mithilfe des Konzepts der kumulativen Summen aus den Stations- und Radardaten berechnet.

In einem ersten Schritt soll das bestehende Bias zwischen beiden Produkten für ganz Deutschland untersucht und Zusammenhänge zu anderen Parameter analysiert werden, und anschließend verschiedene Biaskorrekturverfahren, basierend auf dem Verfahren des Quantile-Mappings, angewandt und verglichen werden. Entgegen der bisherigen Analysen wird hierfür ein empirischer Ansatz verwendet. In einem letzten Schritt soll eine Extremwertanalyse der korrigierten Stationen durchgeführt werden.

Es wird angenommen, dass zwischen den Extremereignissen basierend auf Radar- bzw. Stationsdaten ein Bias besteht und die Radardaten die Stationsdaten in der Niederschlagsintensität unterschätzen. Weiter wird davon ausgegangen, dass die Anwendung der entwickelten Biaskorrekturverfahren zu einer Reduzierung des Bias führt. Schließlich wird erwartet, dass eine Unterteilung der Niederschlagsereignisse anhand ihrer Eventgröße eine bessere Performance aufweist als die Korrektur aller Events auf einmal.

3 Methoden und Vorgehensweise

3.1 Datengrundlage

Im Rahmen dieser Arbeit wurden sowohl aufbereitete Daten von Niederschlagsradaren als auch Daten von Niederschlagsstationen verwendet. Diese Daten sind georeferenziert, die Angaben zu Breiten- und Längengraden sind auf das geodätische Koordinatensystem WGS84 bezogen. Sofern nicht anders angegeben sind die in der Arbeit dargestellten Karten mit der auf das WGS84-Koordinatensystem basierten World-Mercator-Projektion (EPSG-Code: 3395) dargestellt. Zur Darstellung der europäischen (und auch deutschen) Landesgrenzen wurde die Shapefile-Datei „Staaten 2020 1:1 000 000, Stand 03.04.2020“ von eurostat/GISCO (vgl. D1) verwendet. Sämtliche Analysen der Arbeit wurden mithilfe der Programmiersprache Python durchgeführt, dazu wurde die Python-Distribution *Anaconda* genutzt. Die digitale Version enthält neben dem Code zur Berechnung der Analyse-Ergebnisse eine `environment.yml`-Datei mit der die verwendete Umgebung in *Anaconda* hergestellt werden kann.

3.1.1 Radardaten

Bei den verwendeten Radardaten handelt es sich um ein Radar-Niederschlagsprodukt des Unternehmens Kachelmannwetter, das im Rahmen des AVOSS-Projektes des Lehrstuhls zur Verfügung stand (vgl. D9).

Dieses Radarprodukt hat eine räumliche Auflösung von 1 km x 1 km und deckt beinahe ganz Deutschland ab. Die zeitliche Auflösung beträgt fünf Minuten und umfasst den Zeitraum von 2006 bis 2022. Diese Niederschlagsdaten entstammen den Reflektivitätsprodukten der 17 Wetterradare des Deutschen Wetterdienstes mit einem jeweiligen Radius des Radarschirms von 128 km (Winterrath et al., 2017). Jeder Zyklus einer Radarmessung beginnt dabei mit dem bodennahen Niederschlagsscan, der dem Verlauf der Orographie folgt. Anschließend findet der Volumenscan statt, der aus den Messungen für die Höhenwinkel 5,5°, 4,5°, 3,5°, 2,5°, 1,5°, 0,5° und weiter für die Höhenwinkel 8,0°, 12,0°, 17,0° und 25,0° besteht (Winterrath et al., 2017). Folgend wurden die Reflektivitäts-Daten dieser Radare prozessiert, Störsignale entfernt und nicht-lokal über einen radar-physikalischen Zusammenhang betrachtet. Ein Optimierungsalgorithmus angeeignet (Weiler et al., 2019).

Die Abbildung 3.1 (rechts) zeigt die Standorte der Wetterradare und ihre räumliche Abdeckung. Weitere Informationen zu dem jeweiligen Standort kann Tabelle 7.1 entnommen werden. Die Koordinaten der Radarstandorte wurden dabei aus Winterrath et al., 2017 und DWD, 2018 übernommen. Da es im Laufe der Jahre zu einigen Änderungen der Radartechniken und teilweise auch Änderungen des Radarstandortes kam, wurden der Einfachheit halber die aktuellsten Informationen aus DWD, 2018 für die Analysen verwendet. Einzig das Niederschlagsradar Emden-Knock bildet eine Ausnahme, da dieses 2018 durch das Ausfallsicherungsradar Borkum ersetzt wurde, welches 2021 durch ein stationäres Radar ebenfalls auf Borkum ersetzt wurde (Weigl, 2018, Weigl, 2021). Da ein Großteil des Untersuchungszeitraums vor dem Jahr 2018 liegt, wurde in diesem Fall die Koordinaten des Radars Emden-Knock verwendet. Eine ausführliche Auflistung der Radarhistorie und Änderungen der Messsysteme und Informationen zu den aktuellen Radarstandorten kann Winterrath et al., 2017 und DWD, 2018 entnommen werden.

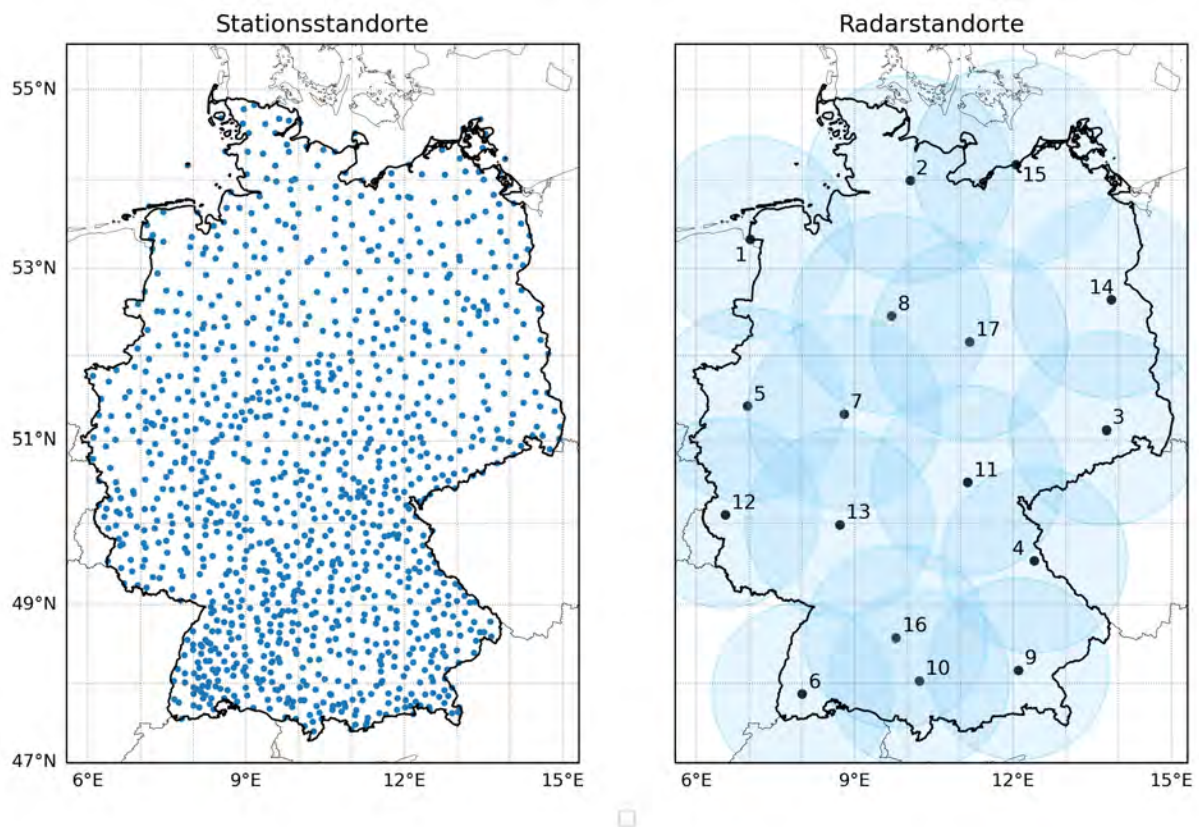


Abbildung 3.1: Karten der Niederschlagsstationen (links) und der Wetterradare (rechts) des Deutschen Wetterdienstes. Die Radarstandorte sind: 1: Emden-Knoock, 2: Boostedt, 3: Dresden-Klotzsche, 4: Eisberg, 5: Essen-Bredeney, 6: Feldberg, 7: Flechtdorf, 8: Hannover, 9: Isen, 10: Memmingen, 11: Neuhaus am Rennweg, 12: Neuheilenbach, 13: Offenthal, 14: Prötzel, 15: Rostock, 16: Türkheim, 17: Ummendorf. Der Radarschirm ein Standortes umfasst jeweils 128 km. (©EuroGeographics bezüglich der Verwaltungsgrenzen)

3.1.2 Stationsdaten

Die genutzten Stationsdaten stammen aus dem Stations-Netzwerk des Deutschen Wetterdienstes (DWD), haben eine zeitliche Auflösung von zehn Minuten und umfassen insgesamt 1172 Stationen (Stand: 09.2023). Diese werden über das *Climate Data Center* des DWDs zur Verfügung gestellt. Auch die Metadaten der 10-minütigen Niederschlagsstationen entstammen dem *Climate Data Center* des Deutschen Wetterdienstes bezogen (vgl. D4). Die Niederschlagsdaten wurden in aufbereiteter Form über die Professur-eigene App *WeatherDB* (vgl. 7) bezogen. Die Daten umfassen dabei die historischen und aktuellen 10-minütigen Daten des DWDs (vgl. D2 und D3).

Die Rohdaten des DWD unterliegen, je nach Qualitätsniveau einer mehr oder weniger ausführlichen Qualitätskontrolle des Wetterdienstes, sie sind jedoch weder homogenisiert, noch fand eine Korrektur des Benetzungs- und Windfehlers nach Richter, 1995 statt (vgl. Winterrath et al., 2017). Die Verwendung der *WeatherDB*-App erlaubt die Nutzung von Niederschlagsdaten, auf die ein mehrstufiges Aufbereitungsverfahren angewandt wurde. Dieses Verfahren umfasste den Download der Rohdaten vom *Climate Data Center*, eine Qualitätskontrolle der Daten, das Füllen der Zeitreihe und die Anwendung der Richter-Korrektur ¹.

Der Download der Daten begann frühestens ab dem 01.01.1994. Zur Qualitätskontrolle der Niederschlagsdaten wurde die aggregierte Tagessumme der 10-minütigen Daten mit dem gemessenen Tageswert des DWDs verglichen. Die betroffenen 10-minütigen Daten wurden in den zwei

¹Eine Beschreibung der Methode ist unter <https://weatherdb.readthedocs.io/en/latest/Methode.html> (Stand 10.2023) zu finden.

gelöscht. Zum einen falls die aggregierte Tagessumme ein Niederschlag von 0 mm ergab, die gemessene Tagessumme aber nicht. Zum anderen falls die gemessene Tagessumme doppelt so groß war als die aggregierte und falls drei aufeinanderfolgende identische 10-minütigen Niederschlagswerte bestanden, die das Qualitätsniveau des DWD der Stufe 3 nicht erreichten. Die letzte Bedingung beruht dabei auf der Annahme, dass bei einer Messgenauigkeit der Stationen von 0,01 mm die identische Werte durch die Teilung von z. B. stündlichen Werten zustande kam.

Zum Füllen der Datenlücken wurden Daten der Nachbarstationen verwendet. Die Nachbarstationen wurden dabei in der Reihenfolge der geringsten horizontalen Entfernung ausgewählt. Dazu wurde ein Regionalisierungsverfahren angewandt, dass das HYRAS-Raster der mehrjährigen Niederschläge des Zeitraums 1991-2020 des DWDs verwendete. Da dieses eine Auflösung von 1 km² hat wurde es mit dem Digitalen Geländemodell (DEM25) auf eine Auflösung von 25 m² präzisiert. Für jede Rasterzelle und die umliegenden Nachbarzellen (Fenster) wurde folgend eine lineare Regression zwischen mehrjährigem Niederschlagswert und Höhe definiert. Die Größe des Fensters startete bei 5 km² und wurde sukzessive um 1 km auf beiden Seiten vergrößert, bis das Fenster größer als 13 km² war oder die Standardabweichung größer als 4 m ist. Die bestimmte Regression für das Fenster wurde dann auf die DEM25-Zellen innerhalb der 1 km² HYRAS-Zelle angewandt, um neue mehrjährige Niederschlagsfaktoren zu berechnen. Dabei wurde je ein Regionalisierungsfaktor für den Sommer (April bis September) und ein Faktor für den Winter (Oktober bis März) berechnet.

Die Fülldaten wurden anschließend gemäß Formel (3.1) berechnet:

$$N_{fillup} = \begin{cases} N_{neighbor} \cdot \frac{N_{station,ma_winter}}{N_{neighbor,ma_winter}} & \text{if } month \in [4 : 9] \\ N_{neighbor} \cdot \frac{N_{station,ma_sommer}}{N_{neighbor,ma_sommer}} & \text{if } month \notin [4 : 9] \end{cases} \quad (3.1)$$

Dabei ist N_{fillup} die Niederschlagswerte zum Füllen, $N_{station}$ und $N_{neighbour}$ die Niederschlagswerte der Station bzw. Nachbarstation, die Indices ma_winter und ma_sommer stehen für das langjährige Mittel der Winter- bzw. Sommermonate.

Schließlich wurden die 10-minütige Niederschlagswerte an die Tagessummen angepasst. Dazu wurde aus den gefüllten 10-minütigen Werten eine aggregierte Tagessumme verglichen und der Quotient aus dieser mit der gemessenen Tagessumme gebildet und dieser auf die 10-minütigen Werte angewandt, sodass die Differenz zur täglichen Messung relativ addiert wurde.

In einem letzten Schritt wurde auf die 10-minütigen Niederschlagswerte die auf Richter, 1995 basierende Korrektur angewandt, um Benetzungs- und Windverluste zu korrigieren. Hierfür wird für jede Station die Horizontabschirmung aus dem DEM25 oder dem DEM80 berechnet. Dazu wurden Rasterzellen auf einer 75 km langen Linie zur Station ausgewählt, der Punkt mit dem größten Winkel zur Station wurde als Horizontabschirmung (Winkel) für diese Linie gewählt. Von Norden bis Süden wurde dies für mehrere Linien mit einem Abstand von 3° durchgeführt. Die Richter-Horizontabschirmung wurde anschließend mit Formel (3.2) berechnet:

$$H' = 0,15 \cdot H_{S-SW} + 0,35 \cdot H_{SW-W} + 0,35 \cdot H_{W-NW} + 0,15 \cdot H_{NW-N}. \quad (3.2)$$

Dabei ist $H_{S-SW;SW-W;W-NW;NW-N}$ die Horizontabschirmung in ° in dem angegebenen Teil der Himmelsrichtungen und H' die Richter-Horizontabschirmung in °. Mit diesem wurde für jede Station die Richter-Expositions-Klasse definiert.

$$\Delta N = b * N^E, \quad (3.3)$$

Mit der Niederschlagsdifferenz ΔN [mm/h], dem Korrekturfaktor b [-], dem Niederschlagswert N [mm/h] und dem Exponenten E [-], und den täglichen Temperaturwerten (zur Bestimmung der Niederschlagsart) wurde der tägliche Richter-korrigierte Niederschlagswert nach Formel (3.3) und den Parameterwerten in Tabelle 7.2 berechnet. Schließlich wurde diese tägliche Korrektur als Block auf die 10-minütigen Messwerte größer 0 mm angewandt.

Weitere über das Climate Data Center bezogene Daten, die nicht über *WeatherDB* bezogen wurden, sind die 5-minütigen und stündlichen Niederschlagsmessungen (D5, D6), sowie die KOSTRA-DWD-2020V-Daten D7).

Der KOSTRA-DWD-2020V-Datensatz stellt ein Rasterprodukt dar, das zum Großteil auf dem Stationsnetz des DWDs basiert, verwendete RADKLIM-Daten spielen dabei nur eine untergeordnete Rolle (Junghänel et al., 2022). Dieses Produkt hat eine räumliche Auflösung von 5 km x 5 km und gibt die Bemessungsniederschläge (oder Starkniederschlagshöhen) für Jährlichkeiten von einem bis zu 100 Jahren an (Junghänel et al., 2023).

3.2 Methoden

3.2.1 Qualitätskontrolle

In einem ersten Schritt wurden die ungefüllten, gefüllten und Richter-korrigierten Stationsdaten miteinander auf die Eignung für die folgenden Analysen verglichen. Schließlich wurden die Richter-korrigierten Niederschlagsdaten der Stationen ausgewählt. Zum einen konnte durch das Regionalisierungsverfahren die Anzahl an Fehlwerten im Datensatz beseitigt werden, sodass eine möglichst gute Datengrundlage für die Bildung der Bemessungsniederschläge gewährleistet war. Zum anderen bietet die Nutzung der Richter-korrigierte Daten den Vorteil, dass die Messungenauigkeiten der Station durch Windüberwehungen verringert wurden. Die verbleibende Unsicherheit liegt somit vielmehr an der Qualität der Radardaten. Da durch die Richter-Korrektur, die Werte nach oben korrigiert wurden, hatte dies unter der Annahme, dass die Radarniederschläge die Stationsniederschläge unterschätzen, eine Vergrößerung des Bias zur Folge. Im Hinblick auf das Biaskorrekturverfahren konnte so die Wahrscheinlichkeit einer zu geringen Korrektur der Werte verringert werden.

Eine Homogenisierung der Stationszeitreihe fand nicht statt, allerdings wurde der Einfluss der Änderung des Stationsstandortes auf die Niederschlagswerte untersucht. Dazu wurden zuerst mithilfe der Stations-Metadaten diejenigen Stationen identifiziert, bei denen im betrachteten Zeitraum mindestens ein räumlicher Stationsversatz vorkam. Anschließend fand eine Gruppierung der betroffenen Stationen je nach Anzahl der Versätze statt. Da die Zeiträume vor und nach dem Versatz teils eine sehr unterschiedliche Anzahl an Messwerte aufwiesen und damit keine vergleichbare Grundgesamtheit vorlag, wurde auf eine quantitative Analyse in Form eines Median-Tests oder ähnlichen Trend-Tests verzichtet, und nur eine visuelle, qualitative Analyse durchgeführt. Für jeden Teilzeitraum einer Station wurden ausschließlich Niederschlagswerte (> 0 mm) verwendet und visuell in Form von Boxplots dargestellt, um das Vorhandensein eines Trends in den Medianen der Teilzeiträume zu prüfen. Da ein sichtbarer Trend nicht nur durch den räumlichen Versatz, sondern auch klimatisch bedingt sein kann, wurde die betroffene Station mit den drei nächsten umliegenden Stationen verglichen. Diese wurden entsprechend in dieselben Teilbereiche aufgeteilt und auf einen Trend hin untersucht. Nur wenn alle drei umliegenden Stationen einen gegenüber der betroffenen Station relevanten abweichenden Trend aufwiesen, wurde die Station für die Biaskorrektur nicht weiterverwendet.

Da sich im Laufe der weiteren Analysen herausstellte, dass einige Bemessungsniederschläge extrem hohe Niederschlagssummen aufwiesen, wurden die zwanzig größten Ereignisse der 10-minütigen Niederschlagsdaten untersucht. Dazu wurde die 10-minütigen Rohdaten mit den gefüllten und Richter-korrigierten Daten verglichen. Falls dabei keine nennenswerte Abweichung aufzufinden war, wurden die 10-minütigen Niederschlagsdaten sowohl mit 5-minütigen und stündlichen Niederschlagssummen des DWDs verglichen (vgl. D5 und D6). Falls eine gewichtige Abweichung festgestellt wurde und eine extreme, unplausible Erhöhung des Niederschlagswertes zustande kam, wurde die betroffene Station verworfen. In dem Fall, dass die 5-minütigen und stündlichen Niederschlagswerte aus Fehlwerten bestanden oder aber keine für diese Zeiten entsprechenden Niederschlagssummen aufwiesen und gleichzeitig ein geringes Qualitätsniveau von Seiten des DWDs angegeben wurde, wurden die betroffenen Stationen ebenfalls nicht weiter verwendet. Von den untersuchten 20 Ereignissen wurden insgesamt acht Ereignisse und damit acht

Stationen verworfen.

Tabelle 3.1: Warnstufen des DWDs für Starkregenereignisse (nach DWD, 2023)

Warnstufe	Niederschlag [l/m ²]	
	60-minütig	360-minütig
Stufe 2 Starkregen	15–25	20–35
Stufe 3 Heftiger Starkregen	25–40	35–60
Stufe 4 Extrem heftiger Starkregen	> 40	> 60

In einem letzten Schritt wurde geprüft, ob für jede verbleibende Station alle notwendigen Radarniederschläge gegeben sind. Falls um die Rasterzelle des Radardatensatzes, in der sich die Station befindet, nicht alle acht umliegenden Rasterzellen vorhanden, wurden auch diese Stationen verworfen. Dies betraf ausschließlich drei Stationen in der Nähe der Deutschen Landesgrenze, die nicht mehr vollständig von dem jeweiligen Radarschirm erfasst wurden. Eine Karte der genutzten und verworfenen Stationen ist im Anhang zu finden (vgl. Abbildung 7.1)

3.2.2 Untersuchung der Klimatologie

Zur weiteren Plausibilisierung der Stationsdaten und der klimatologischen Beschreibung des betrachteten Zeitraums wurden einfache klimatologische Analysen durchgeführt. Neben der stationsweisen Berechnung von Jahressummen und der Bestimmung der minimalen, mittleren und maximalen Jahressumme wurden die langjährigen arithmetischen Monatsmittel berechnet. Diesen lag der Zeitraum von 2006 bis 2022 zugrunde. Zuletzt wurden die 10-minütigen Stationsdaten zu Tageswerten aggregiert, um den Anteil an Regentagen ($> 0,1$ mm Niederschlag) zu bestimmen.

Da der Fokus der Arbeit auf kurzfristigen Extremereignissen liegt, wurden die Stationsdaten, zusätzlich zu den klimatologischen Analysen, auf die Häufigkeit von Extremereignissen untersucht. Als Kennzahl wurden die Starkregenwarnstufen des Deutschen Wetterdienstes verwendet (vgl. DWD, 2023 und Tabelle 3.1). Die Warnstufen wurden jedoch getrennt für die 60-minütige und 360-minütige Dauer berechnet. Zuerst wurde für jeden 10-Minuten-Schritt auf die entsprechende Dauerstufe aggregiert und anschließend geprüft, welche der resultierenden Niederschlagssummen die Kriterien für die Starkregenwarnstufen erfüllen. Zuletzt wurde sichergestellt, dass von dieser Auswahl, nur Events mit einem zeitlichen Abstand von über 48 Stunden zum vorigen bzw. nächsten Event berücksichtigt werden, um die mehrfache Zählung des gleichen Events zu vermeiden. Zum Vergleich wurden jeweils das 95ste, 99ste und 99,9ste Quantil der beiden Dauerstufen für jede Station berechnet.

3.2.3 Berechnung der kumulativen Niederschlagssummen

Da in der Praxis oftmals Bemessungsniederschläge z. B. für ingenieurtechnische Fragestellungen benötigt werden, sollten aus den 10- bzw. 5-minütigen Stations- und Radardaten Niederschlagshöhen einer bestimmten Dauerstufe berechnet werden. Um kurzfristige Starkniederschläge abzubilden, wurden Niederschlagsintensitäten von 10-minütiger, 60-minütiger, 180-minütiger und 360-minütiger Dauer berechnet.

Die verwendete Methode zur Berechnung der Niederschlagssummen entspricht dabei der von Hänslers und Weiler, 2022 entwickelten. Hierfür wurden für jeden Zeitschritt, sofern genügend Werte zur Verfügung standen, die Daten entsprechend der Dauerstufe aggregiert. Daraus ergab sich ein Datensatz für jede Dauerstufe, der z. B. alle möglichen 60-minütigen Niederschlagswerte des Zeitraums enthielt. Anschließend wurde für jede Dauerstufe die maximale Niederschlagshöhe pro Tag und das dazugehörige Datum ermittelt. Aus diesen täglichen maximalen Niederschlagswerten wurden schließlich die 150 größten Werte einer jeden Station bzw. eines jeden Pixels bestimmt. Dabei kam wiederum ein zeitlicher Puffer von 48 Stunden zwischen den Niederschlagssummen zum Einsatz. Diese 150 Niederschlagshöhen pro Dauerstufe und Station (folgend auch

als Ereignisse oder Events bezeichnet) wurden für die weiteren Analysen verwendet, und meist als Event-Rang bezeichnet.

3.2.4 Quantifizierung des Bias

Vor der Entwicklung und Anwendung der Biaskorrekturverfahren, wurde sich ein Bild über das bestehende Bias zwischen den Stations- und Radarevents gemacht. Obwohl in Wirklichkeit beide Messungen fehlerbehaftet sind, werden die Stationsniederschläge als „wahrer Niederschlag“ bzw. als die beste Schätzung (*ground truth*) angesehen, was in der direkten Messung des Niederschlags begründet ist. Anhand der Stationsniederschläge wurde so die Korrektheit der Radarniederschläge ermittelt. Dazu wurden die berechneten maximalen Niederschlagsintensitäten genutzt. Das absolute Bias wurde daher gemäß Formel (3.4) und das relative Bias gemäß Formel (3.5) definiert.

$$bias_{i; abs} = N_{i; rad} - N_{i; stat} \quad (3.4)$$

$$bias_{i; rel} = \frac{N_{i; rad} - N_{i; stat}}{N_{i; stat}}. \quad (3.5)$$

Hier steht $bias_{abs}$ für das absolute Bias, $bias_{rel}$ für das relative Bias, N_{rad} für den Radarniederschlag, N_{stat} für den Stationsniederschlag und der Index i für die jeweilige Station.

Da die Stationsdaten als Punktdaten, die Radardaten als Raster-Datensatz abgespeichert sind, musste zuerst jedem Stationsstandort die entsprechende Rasterzelle zugeordnet werden. Dazu wurde die Rasterzelle ausgewählt, die die geringste geodätische Distanz zum jeweiligen Stationsstandort aufweist. Diese Rasterzelle wurde als *Stations-Pixel* bzw. *Stat-Px* benannt. Für den Fall, dass ein Niederschlagsereignis nicht vollständig in einer Rasterzelle liegt, wurden auch die acht umliegenden Rasterzellen mitberücksichtigt (siehe Abbildung 3.2). Insgesamt wurden fünf verschiedene Biastypen berechnet: das Bias des Stations-Pixels *Stat-Px*, das mittlere Bias der vier nächsten Pixel *Mean4*, das mittlere Bias der neun Pixel *Mean9* sowie zum Vergleich das minimale Bias und das maximale Bias der neun Pixel.

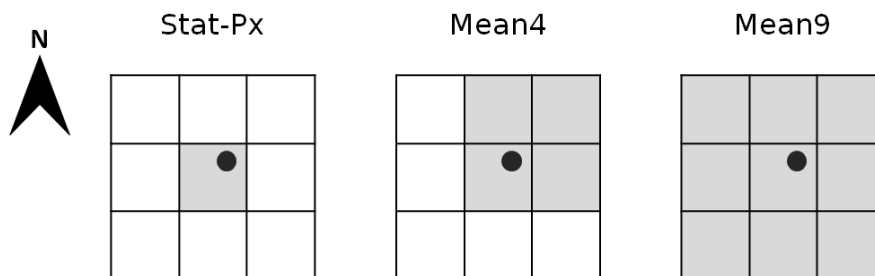


Abbildung 3.2: Darstellung von drei der fünf verwendeten Biastypen: Das die Station beinhaltende Pixel (*Stat-Px*), die vier nächstgelegenen Pixel *Mean4* und alle umliegenden Pixel *Mean9*.

Zur weiteren Analyse des Bias wurden potentielle Zusammenhänge des absoluten und des relativen Bias mit anderen Parametern untersucht. Diese Analysen beschränkten sich meist auf das Stations-Pixel, weil die zusätzlichen Informationen meist nur für den Stationsstandort bekannt waren. Zum einen wurde die Korrelation zwischen Bias und Stationshöhe analysiert, zum anderen ein potentieller Zusammenhang zwischen dem Bias und der Richterklasse, als indirektes Maß für den Windeinfluss bzw. den Windschatten der Station, untersucht. Zuletzt fand eine Analyse des Bias pro Radarturm statt. Dazu wurden diejenigen Stationen, die sich im Bereich des Radarschirms befanden, dem jeweiligen Radarturm zugeordnet. Da sich die Abdeckung des Radarschirms teils überschneidet, wurden einige Stationen mehreren Radartürmen zugeordnet.

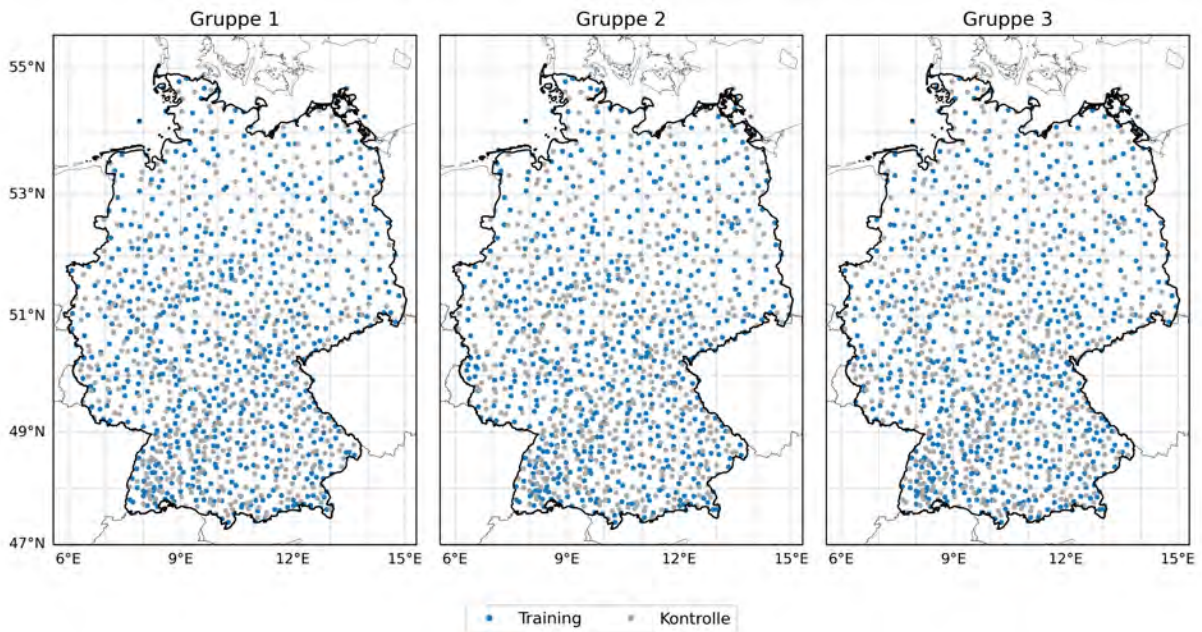


Abbildung 3.3: Räumliche Verteilung der Stationen einer jeden in dieser Arbeit verwendeten Trainings- und Kontrollgruppe (©EuroGeographics bezüglich der Verwaltungsgrenzen)

Für jeden Radarturm wurde weiter die Korrelation zwischen dem Bias und der (dreidimensionalen) Distanz zwischen Radar- und Stationsstandort für jedes der 150 Ereignisse analysiert und folgend ein allgemeiner Trend pro Radarstandort abgeschätzt.

Als Ergänzung zu diesen „künstlichen“ Ereignissen fand eine exemplarische Untersuchung des Bias von realen Niederschlagsereignissen statt. Dazu wurden neun reale Extremniederschlagsereignisse im Zeitraum von 2006 bis 2022 ausgewählt. Die Eventdauer wurde als der Zeitraum mit Niederschlägen größer 0,1 mm definiert.

3.2.5 Erstellung der Trainings- und Kontrollgruppen

Der im Korrekturverfahren gefundene Zusammenhang zwischen Stations- und Radardaten soll auf die gesamte Radar-Fläche angewandt werden. Unter der getroffenen Annahme, dass die Stationsniederschläge „wahr“ sind, stehen in der Fläche jedoch nur punktuell an den Stationsstandorten Daten zur Verfügung, die einen direkten Vergleich zwischen Stations- und Radarniederschlägen erlauben. Damit die Qualität des Korrekturverfahrens abgeschätzt werden konnte, wurden daher die genutzten Niederschlagsstationen in eine Trainings- und eine Kontrollgruppe aufgeteilt. Zur Auswahl einer entsprechenden Kontrollstation zu jeder Trainingsstation, wurde eine höhengewichtete Distanz verwendet, ähnlich dem Verfahren des Wasserhaushaltsmodells LARSIM (LARSIM-Entwicklergemeinschaft, 2023, S. 21–22). Zuerst wurde dazu die geodätische Entfernung einer jeden Station zu allen anderen Stationen berechnet, anschließend die Gewichtungsfaktoren nach Formel (3.4) ermittelt und diese mit der Entfernung multipliziert.

$$h_{fac} = 1 + \frac{\Delta h^{par_1}}{par_2} \quad (3.6)$$

Dabei ist h_{fac} der dimensionslose Gewichtungsfaktor, ΔH die Höhendifferenz der zwei Stationen in m und par_1 und par_2 zwei dimensionslose Parameterwerte. Für diese Arbeit wurden die Parameterwerte $par_1 = 1$ und $par_2 = 100$ gewählt.

Anschließend wurde sukzessive zufällig eine Station gewählt und zu dieser innerhalb eines Radius von 40 km zufällig eine entsprechende Station selektiert. Falls innerhalb dieses Radius keine verbleibende Station mehr vorhanden war, wurde zufällig eine der 50 nächsten Stationen aus-

gewählt. Für die gewählten Parameterwerte wurde sich nach Testen verschiedener Alternativen entschieden. Damit sollte gewährleistet werden, dass einerseits die Stationen der Trainings- und Kontrollgruppe randomisiert ausgewählt werden und keine räumlichen Muster sichtbar sind und andererseits, dass die Höhenverteilung der Stationen in beiden Gruppen ähnlich ist (vgl. 7.2).

Die Auswahl der Stationen der Kontrollgruppe ist entscheidend für die folgenden Korrekturverfahren und damit für die Qualität der Korrektur der Radardaten. Um die Unsicherheit bedingt durch die Stationsauswahl für die Gruppenbildung abzuschätzen, wurde die Aufteilung in Trainings- und Kontrollgruppe dreimal durchgeführt (siehe Abbildung 3.3). Insgesamt standen somit drei Trainingsgruppen und drei Kontrollgruppen für die weiteren Analysen zur Verfügung. Die drei Trainings- bzw. Kontrollgruppen waren jeweils gleich groß, bedingt durch das Verfahren kamen aber überlappenden Stationen vor, die in zwei oder allen drei Gruppen enthalten waren.

Für die weiteren Analysen wird Gruppe 1 als Referenz festgelegt, Gruppe 2 und 3 dienen dazu etwaige Abweichungen zu quantifizieren und darzustellen. Wenn nicht anders angegeben beziehen sich die Analysen und Grafiken somit auf Gruppe 1.

3.2.6 Biaskorrekturverfahren

Grundlegendes Vorgehen

Das Verfahren des Quantile Mappings nutzt, wie in der Einleitung bereits erwähnt, die kumulierte Verteilungsfunktionen, um eine Transferfunktion zu bestimmen. Hier wurde sich für ein empirisches Quantile-Mapping-Verfahren entschieden, wodurch keine Parameterschätzung notwendig war. Die Grundidee des verwendeten empirischen Quantile-Mapping-Verfahrens kann mit Formel (3.7) ausgedrückt werden:

$$N_{stat} = ECDF_{stat, trng}^{-1}(ECDF_{rad, trng}(N_{rad})) \quad (3.7)$$

Dabei ist N die Niederschlagsintensität in [mm/Dauerstufe], $ECDF()$ die empirische kumulierte Häufigkeitsverteilung (ECDF) und $ECDF()^{-1}$ die inverse empirische kumulierte Häufigkeitsverteilung. Die Indices $stat$ und rad symbolisieren die Stations- bzw. Radardaten und $trng$ die Trainingsgruppe. Mit dieser Transferfunktion können die ECDFs der Stations- und Radardaten miteinander in Verbindung gesetzt werden.

Für die praktische Berechnung wurde eine bestimmte Anzahl an Wahrscheinlichkeitsklassen (Bins) festgelegt. Für diese wurden Quantile für die der Größe nach sortierten Stations- und Radardaten der Trainingsgruppen und für die Radardaten der Kontrollgruppe berechnet. Die verwendeten Radardaten beschränkten sich dabei auf das Stations-Pixel. Der Berechnung der Quantile lag dabei Formel (3.8) zu Grunde, die der Methode 7 in (Hyndman und Fan, 1996) entspricht.

$$k = p \cdot (n - 1) + 1 \quad (3.8)$$

In der Formel entspricht p der Wahrscheinlichkeit zur Berechnung des Quantils, sodass p % der Datenpunkte kleiner oder gleich des Quantilwertes sind, n der Anzahl an Datenpunkten und k einem (virtuellen) Index (*plotting position*). Falls dieser Index nicht ganzzahlig ist, also keinem Index des sortierten Datensatzes entspricht, wird der Quantilwert aus dem vorigen und nächsten Datenpunkt gemäß dem nicht-ganzzahligen Teil des Index interpoliert.

Somit konnte jedem Radar-Quantil der Trainingsgruppe ein Stations-Quantil der Trainingsgruppe mit der gleichen Wahrscheinlichkeit zugeordnet werden. Diese Datenpaare können zueinander ins Verhältnis gesetzt werden, um damit Korrekturfaktoren zu bestimmen.

Um abrupte Änderungen der Korrekturfaktoren zu verhindern, wurde die Verteilung der Korrekturfaktoren geglättet. Dazu wurde ein gleitender Mittelwert unter Berücksichtigung des vorigen und nächsten Wertes verwendet. Für die Glättung des ersten und letzten Wertes der Korrekturfaktoren wurde der zweite bzw. vorletzte Wert doppelt gewichtet. Diese geglätteten Korrekturfaktoren wurden folgend gemäß Formel (3.9) auf die Kontrollgruppe der Radardaten angewandt. Dazu wurde die Niederschlagswerte der Kontrollgruppe der Radardaten gemäß der

Quantil-Obergrenzen in Klassen eingeteilt und diese mit den jeweiligen Korrekturfaktoren multipliziert.

$$\begin{aligned}
N_{p; rad, ctrl_corr} &= N_{p; rad, ctrl} \cdot C_p \\
N_{p; rad, ctrl_corr} &= N_{p; rad, ctrl} \cdot \frac{Q_{p; stat, trng}}{Q_{p; rad, trng}} \\
&\text{mit } N_p = Q_{p-1; rad, ctrl} < N \leq Q_{p; rad, ctrl}
\end{aligned} \tag{3.9}$$

N steht dabei für alle Niederschlagswerte, N_p für die Niederschlagswerte die in eine Quantilkategorie der Wahrscheinlichkeit p fallen, deren Werte also kleiner oder gleich dem Quantilwert dieser Klasse Q_p aber größer als der vorige Quantilwert Q_{p-1} sind. C_p ist der Korrekturfaktor einer Quantilkategorie, die Indices *stat*, *trng* und *rad*, *trng* stehen für die Stations- bzw. Radardaten der Trainingsgruppen und *rad*, *ctrl* und *rad*, *ctrl_corr* für die Radardaten bzw. korrigierten Radardaten der Kontrollgruppe.

Um die Leistungsfähigkeit der Korrektur zu bewerten, wurde anschließend das verbleibende absolute Bias $bias_{rem, abs}$ und relative Bias $bias_{rem, rel}$ für jede Station i mithilfe der Stationsdaten der Kontrollgruppe berechnet (vgl. (3.10) und (3.11)).

$$bias_{i; rem, abs} = N_{i; rad_corr} - N_{i; stat} \tag{3.10}$$

$$bias_{i; rem, rel} = \frac{N_{i; rad_corr} - N_{i; stat}}{N_{i; stat}} \tag{3.11}$$

In einem zweiten Schritt wurden alle Radar-Rasterzellen nach derselben Formel korrigiert. Auch hier wurden zuerst die gleichen Quantile aller Radarwerte berechnet und anschließend die entsprechenden Korrekturfaktoren angewandt.

Verfahren 1: Rangbasierte Korrektur

Bei diesem Verfahren fand eine rangbasierte Korrektur pro Dauerstufe statt. Durch die Auswahl von 150 Niederschlagsereignissen standen pro Station 150 Ereignisse zur Verfügung die sich in ihrer Niederschlagsintensität unterschieden. Für insgesamt 1000 Wahrscheinlichkeitsklassen (Bins), beginnend bei 0.1 % bis hin zu 100 %, wurden Quantile berechnet. Dabei wurde jedoch nur jeweils einer der 150 Event-Ränge des Datensatzes ausgewählt und die zugehörigen 1154 Niederschlagswerte für die Korrektur genutzt. Pro Station wurde so jeweils Durchlauf nur ein Wert auf einmal verwendet. Dies hatte zur Folge, dass insgesamt 150 verschiedene Sets an Korrekturfaktoren vorlagen. Da weiter jeweils drei Versionen der Trainings- bzw. Kontrollgruppen vorlagen, belief sich die Gesamtanzahl damit auf 3×150 verschiedenen Sets für eine Dauerstufe. Diese Sets können anschließend als dreidimensionales Feld dargestellt werden, um den Zusammenhang zwischen Korrekturfaktor, dem jeweiligen Quantil und dem Event-Rang untersuchen zu können. Die Idee hinter diesem Ansatz ist es, für eine gewisse Niederschlagsstärke ein eigenes Set an Korrekturfaktoren zu erhalten. Dieses Verfahren kann so dem no-pooling-Konzept zugeordnet werden, da für jeden Rang eine eigene Korrektur stattfindet und zwischen den Rängen keine Informationen geteilt bzw. kombiniert werden.

Verfahren 2: Gemeinsame Korrektur aller Events

Das zweite Verfahren kann als Gegenteil zum ersten aufgefasst werden und dem complete-pooling-Konzept zugeordnet werden, da hier alle Informationen zwischen den Rängen kombiniert werden. Bei diesem Verfahren wurden alle 150 Ereignisse einer Station gemeinsam als Niederschlagsdaten für die Korrektur verwendet. Eine Unterscheidung zwischen den Rängen fand also nicht statt. Auch hier wurden insgesamt 1000 Wahrscheinlichkeitsklassen (Bins), beginnend bei 0.1 % bis zu 100 %, für die Quantilberechnung verwendet. Durch das Zusammenfassen aller Ränge kann das Set an Korrekturfaktoren nur als zweidimensionale Kurve dargestellt und somit nur der

Zusammenhang zwischen Korrekturfaktor und Quantil betrachtet werden.

3.2.7 Sensitivitätstests

Zusätzlich zum Vergleich der beiden Verfahren, wurde die Sensitivität des Verfahrens mithilfe bestimmter vergleichender Analysen abgeschätzt. Dabei wurde der Einfachheit halber jeweils das Verfahren 2 als Referenzverfahren verwendet. Zu diesen Tests gehören die Untersuchung des:

1. Einfluss der Anzahl an Bins: Da die Anzahl der Wahrscheinlichkeitsklassen im Verfahren festgelegt werden muss, wurde der Einfluss der Anzahl dieser Bins untersucht. Als Vergleich zur Referenz wurde das Verfahren mit nur 100 Bins durchgeführt.
2. Einfluss des Radar-Rasters: Hierfür wurde vergleichend die Analyse mit dem Mittel der nächsten vier Pixel (mean4) durchgeführt.
3. Einfluss der Jahreszeit: Bei diesem Test wurden nur solche Events verwendet, die im Sommer stattfanden. Mit Sommer werden hierbei die Monate von April bis einschließlich September bezeichnet. Dazu wurde zuerst der Anteil der Sommer- bzw. Winter-Events für die Stations- und Radardaten bestimmt, und anschließend nur die Ereignisse für die Korrektur genutzt, die sich bei beiden im Sommer ereigneten.
4. Einfluss des zeitlichen Zusammenhangs: Da die Auswahl der 150 Ereignisse für die Stations- als auch Radardaten unabhängig voneinander geschieht, ist ein zeitlicher Zusammenhang zwischen beiden nicht zwingend gegeben. Dieser Test berücksichtigt daher nur diejenigen Events einer Station und des zugehörigen Radarpixels, die in ihrem Datum übereinstimmen. Dabei müssen die übereinstimmenden Events nicht zwingend den selben Event-Rang der

Aufgrund von in dieser Form nicht erwartenden Ergebnissen wurde eine weitere Analyse durchgeführt, die zum Ziel hatte die Quantile der zugrundeliegenden Stations- und Radardaten genauer zu analysieren. Hierbei wurde die ECDFs einer jeden Gruppe und eines jeden Ranges der Stations- und Radardaten verglichen. In einem ersten Schritt wurde die Differenz der beiden ECDFs berechnet, um einen Vergleich zu dem im Verfahren genutzten multiplikativen Ansatz (mithilfe des Quotienten aus Stations- und Radardaten) zu ermöglichen. Auch wurden jeweils für die Stations- und Radardaten die kumulierten Differenzen des aktuellen Wertes der ECDF zum vorigen Wert ermittelt. Aus der Verteilung der kumulierten Differenz konnte abgeschätzt werden, wie schnell die jeweilige ECDF ansteigt und somit in welchem Bereich der ECDFs die größten Unterschiede der zwei Datensets zu finden sind. Abschließend wurde die Häufigkeitsverteilung der Stations- und Radardaten untersucht und die empirische Varianz und Schiefe dieser Verteilungen als Lagemaße berechnet und miteinander verglichen.

3.3 Anwendungsbeispiele

3.3.1 Korrektur der realen Events

Folgend wurde versucht die gewonnenen Ergebnisse auf ihre Anwendbarkeit zu testen. Als ersten Anwendungsbeispiel wurde analysiert, inwiefern die Korrekturfaktoren beider Verfahren geeignet sind, die neun realen Extremereignisse zu korrigieren.

Dazu wurde für jedes reale Event (entsprechend der Berechnung der Niederschlagssummen) die maximale Niederschlagshöhe einer gewissen Dauerstufe berechnet. Diese Berechnungen fand für die Stations- und Radardaten gleichermaßen statt und wurde für alle Dauerstufen durchgeführt, für die im betrachteten Zeitfenster genügend Werte zur Verfügung standen. Anschließend wurde der entsprechende Korrekturfaktor bestimmt. Für Verfahren 1 musste dafür für die entsprechende Station zuerst der Event-Rang bestimmt werden. Dabei wurde der größtmögliche Event-Rang ausgewählt, für den die Niederschlagssumme des realen Events größer oder gleich der Niederschlagssumme des entsprechenden Event-Ranges war. Mithilfe des ausgewählten Event-Ranges

und den zugehörigen 1000 Quantilklassen wurde anschließend die größtmögliche Quantiklasse ausgewählt, in die die Niederschlagssumme des realen Events passte. Zuletzt wurde der zu dieser Klasse zugehörige Korrekturfaktor mit der Niederschlagssumme des realen Events multipliziert. Diese Korrektur wurde für das Stations-Pixel und auch alle acht umliegenden Radar-Rasterzellen durchgeführt, sodass zu jeder Radar-Niederschlagssumme auch die Angabe der Spannweite möglich war. Für Verfahren 2 verlief die Korrektur analog, nur die Auswahl des Event-Ranges entfiel.

3.3.2 Berechnung von Bemessungsniederschlägen

Das zweite Anwendungsbeispiel stellte die Berechnung von Bemessungsniederschlägen dar. Als Datengrundlage dienten die biaskorrigierten Radardaten einer jeden Raster-Zelle des Radar-Datensatzes. Von den 150 korrigierten Niederschlagsintensitäten pro Pixel wurden die Ereignisse mit einer Jährlichkeit größer gleich 1 ausgewählt, was hier den 17 größten Ereignissen entsprach. An diese wurde mithilfe der L-Momenten-Methode (mit dem Python-Package „lmoments3“) eine Generalisierte Pareto-Verteilung gefittet. Für jedes Raster-Pixel standen anschließend die Parameter *location*, *scale* und *shape* dieser drei-parametrischen Verteilung zur Verfügung. Zuletzt wurde für jedes Raster-Pixel mithilfe dieser Parameter der Bemessungsniederschlag eines 20- und eines 50-jährlichen Ereignisses berechnet. Zum Vergleich der Bemessungsniederschläge wurden die KOSTRA-DWD-2020V-Daten des DWD verwendet (siehe D7).

4 Resultate

4.1 Klimatologie

Die Untersuchung der Klimatologie des Untersuchungszeitraums auf Basis der Stationsdaten zeigt die zu erwartenden Charakteristiken. Bei Betrachtung der Jahressummen, fallen besonders Trockenjahre wie 2018 als auch regenreichere Jahre wie 2007 auf (siehe Abbildung 4.1). Gleichzeitig ist die Orographie mit höheren Niederschlägen in den Mittelgebirgslagen als auch im Alpenvorland erkennbar, jedoch zeigt sich, wie zu erwarten, das nicht in jedem Trocken- bzw. Regenjahr die minimalen bzw. maximalen Niederschlagssummen in denselben Regionen auftreten müssen.

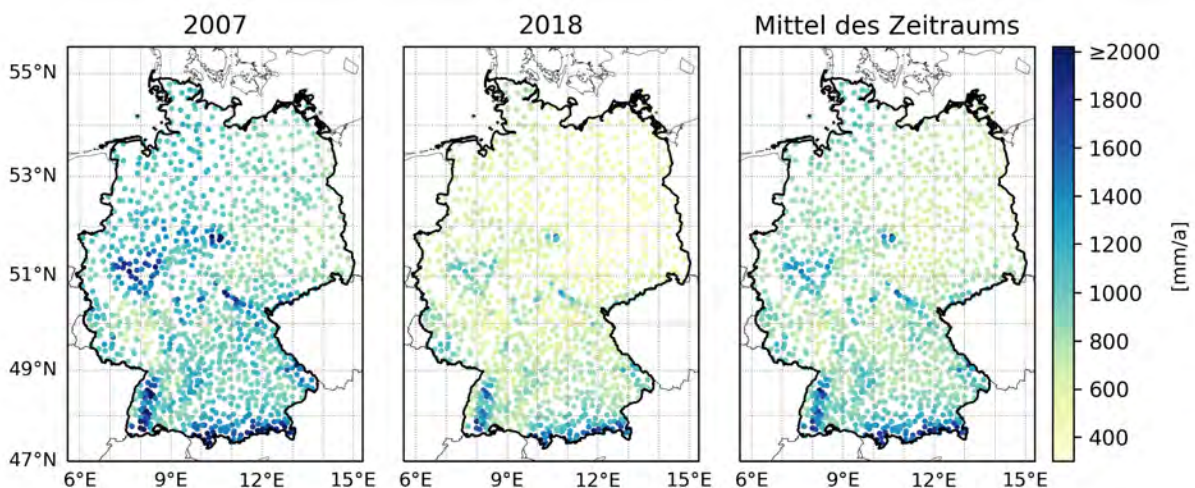


Abbildung 4.1: Beispiel der klimatologischen Analyse. Zu sehen ist der Jahresniederschlag des regenreichen Jahrs 2007, des Trockenjahrs 2018 und der mittlere Jahresniederschlag des Gesamtzeitraums. (©EuroGeographics bezüglich der Verwaltungsgrenzen)

Diese orographisch bedingten Niederschläge zeigen sich besonders in den langjährigen Monatsmitteln des Zeitraums von 2006 bis 2022 (siehe Abbildung 4.2). In den Sommermonaten sind die hohen Niederschläge im Bereich der Mittelgebirge und der Alpen sowie im Alpenvorland gut sichtbar. Die Wintermonate zeigen gerade in den Mittelgebirgen hohe Niederschlagssummen. Des Weiteren sind das ganze Jahr über, besonders jedoch im Winter, die unterschiedlichen Niederschlagssummen zwischen West- und Ostdeutschland erkennbar, während im Westen höhere Monatsmittel auftreten, sind die Niederschläge im Osten geringer. Dieser Einfluss der Kontinentalität ist insbesondere bei Untersuchung des Anteils an Regentagen während des Zeitraums sichtbar. Während in Nordwestdeutschland mehr Regentage vorkommen, ist der Anteil dieser in Nordostdeutschland geringer.

Auch hier ist der hohe Anteil an Regentagen im Bereich der Mittelgebirge und die geringe Anzahl an Regentagen nahe des Frankfurter Raums oder in Brandenburg und Thüringen erkennbar. Insgesamt schwankt der Anteil an Regentagen während des Untersuchungszeitraums zwischen dem Großteil der Stationen zwischen 40 % und 50 %.

Weiterführend wurde das Auftreten von Extremereignissen näher untersucht. Abbildung 4.3 zeigt die Anzahl der Extremereignisse nach den DWD Warnkriterien. Zwischen den 60-minütigen und 360-minütigen Karten kann ein zunehmender Einfluss der Orographie für länger dauernde Events beobachtet werden. Beim Vergleich der Warnstufen innerhalb einer Dauerstufe ist erkenn-

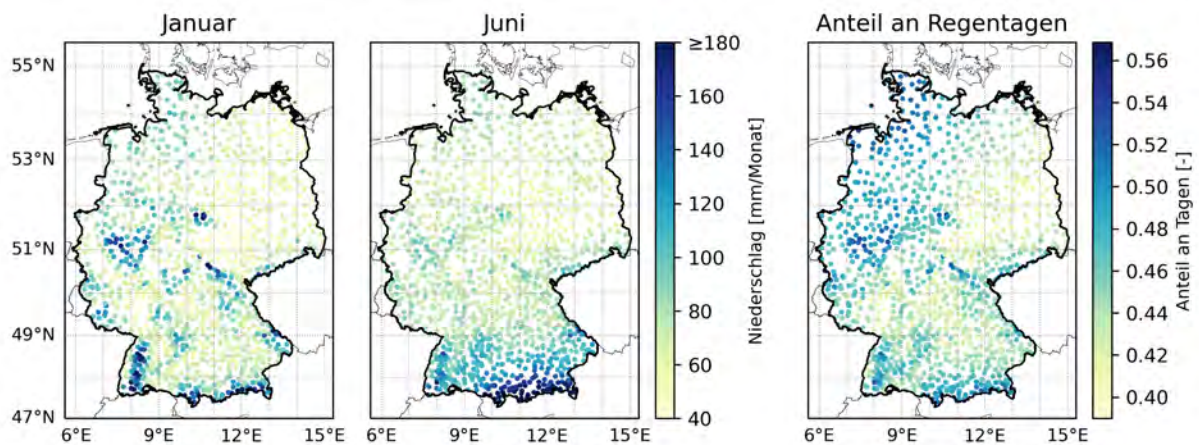


Abbildung 4.2: Beispiel der klimatologischen Analyse. Neben den langjährigen (2006–2022) Monatsniederschlägen des Januars und Julis ist der Anteil der Regentage bezogen auf den Gesamtzeitraum dargestellt. (©EuroGeographics bezüglich der Verwaltungsgrenzen)

bar, dass mit höherer Warnstufe und damit extremen Niederschlagssummen die Stationen mit einer hohen Anzahl an solchen Events zufälliger im Raum verteilt ist und nicht zwangsläufig in der Nähe eines Mittelgebirges zu finden ist.

Die Analyse der 95-%-, 99-%- und 99,9-%-Quantile zeigt keine so deutlichen Ergebnisse, beim 99,9-%-Quantil ist jedoch auch eine zufälligerere Verteilung der Stationen mit hohen Niederschlagswerten im Raum erkennbar (vgl. Abbildung 7.3). Insgesamt zeigt sich, dass die Niederschlagswerte der einzelnen Quantile deutlich unterhalb der DWD-Warnkriterien liegen.

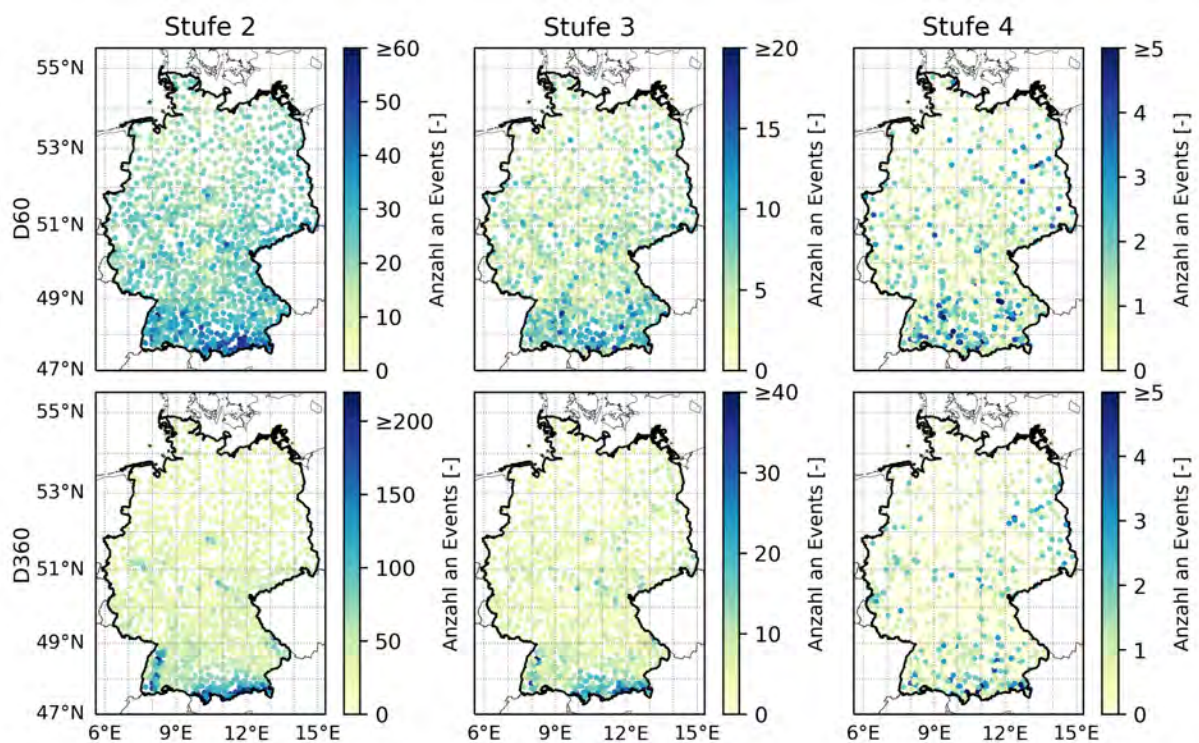


Abbildung 4.3: Anzahl der Extremereignisse während des Untersuchungszeitraums basierend auf den Stationsdaten. Die Analyse fand getrennt für die Dauerstufen D60 und D360 statt. Zwischen zwei Ereignissen musste mindestens ein zeitlicher Abstand von 48 Stunden liegen. (©EuroGeographics bezüglich der Verwaltungsgrenzen)

4.2 Bias

4.2.1 Häufigkeitsverteilungen

Grundsätzlich ist bei einem Großteil der Stationen ein negatives Bias erkennbar. Das bedeutet, dass die Radarniederschläge an den meisten Stationsstandorten die Stationsniederschläge unterschätzen. Einige Standorte weisen hingegen einen positiven Bias und damit eine Überschätzung der Stationsniederschläge durch die Radarniederschläge auf.

Mit zunehmender Eventdauer (Dauerstufe) nimmt die Spannbreite des absoluten Bias zu, wie es durch die größeren Intensitäten zu erwarten ist (siehe Abbildung 4.4). Gleichzeitig nimmt jedoch auch der Anteil der Stationen zu, die einen positiven Bias aufweisen. Während bei D10 der Großteil der Häufigkeitsverteilung einen negativen Bias aufweist, beträgt dieser Anteil bei D360 nur noch in etwa die Hälfte der Häufigkeitsverteilung.

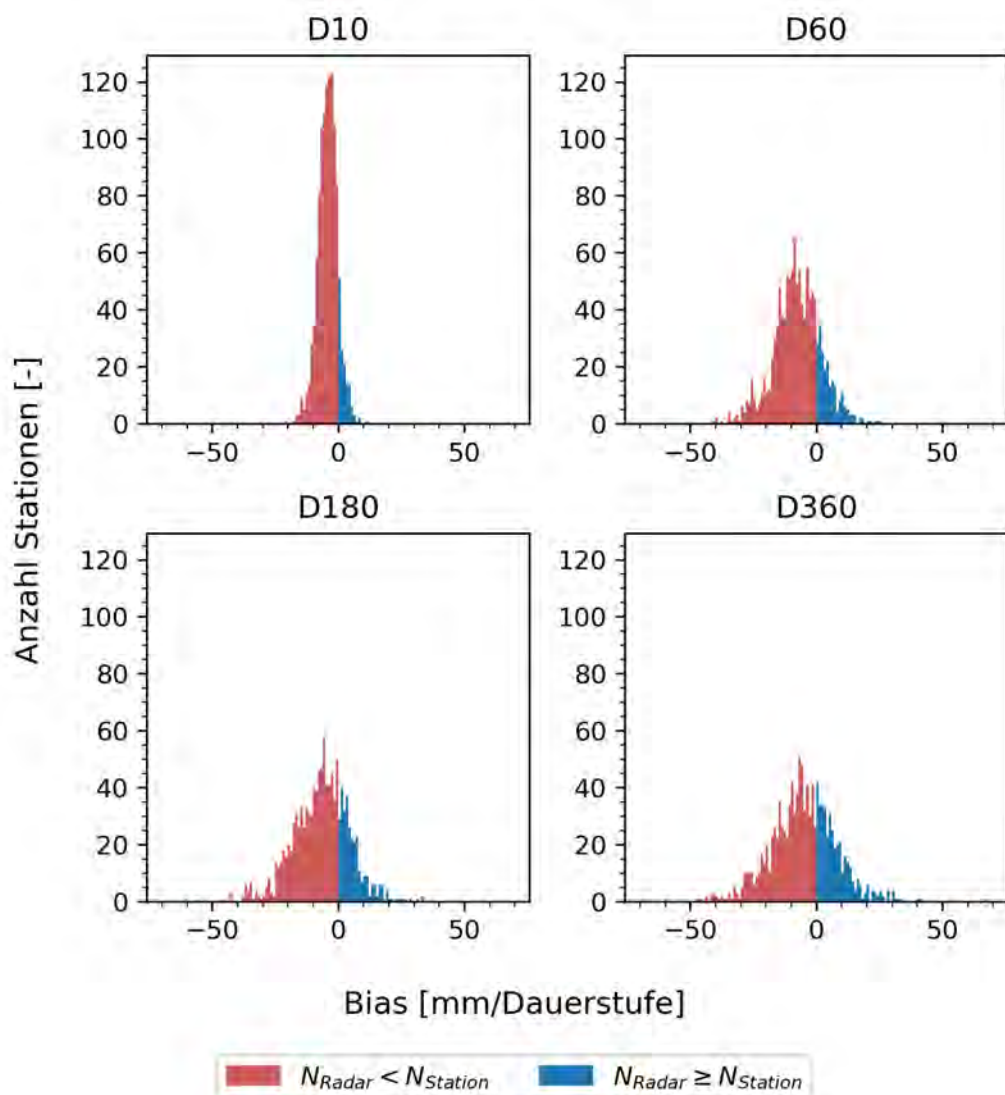


Abbildung 4.4: Histogramm des Bias (Stat-Px) einer jeden Dauerstufe für den zweiten Event-Rang

Werden verschiedene Event-Ränge miteinander verglichen, lässt sich eine Abnahme der Spannbreite des absoluten Bias (siehe Abbildung 4.5), nicht aber des relativen Bias (siehe Abbildung 7.4), mit zunehmendem Event-Rang feststellen. Dies ist schlüssig, da ein größerer Rang einer kleineren Niederschlagsintensität entspricht. Zudem wird die Häufigkeitsverteilung des absoluten und relativen Bias mit größerem Rang spitzer. Der Anteil an Stationen mit positivem Bias ist

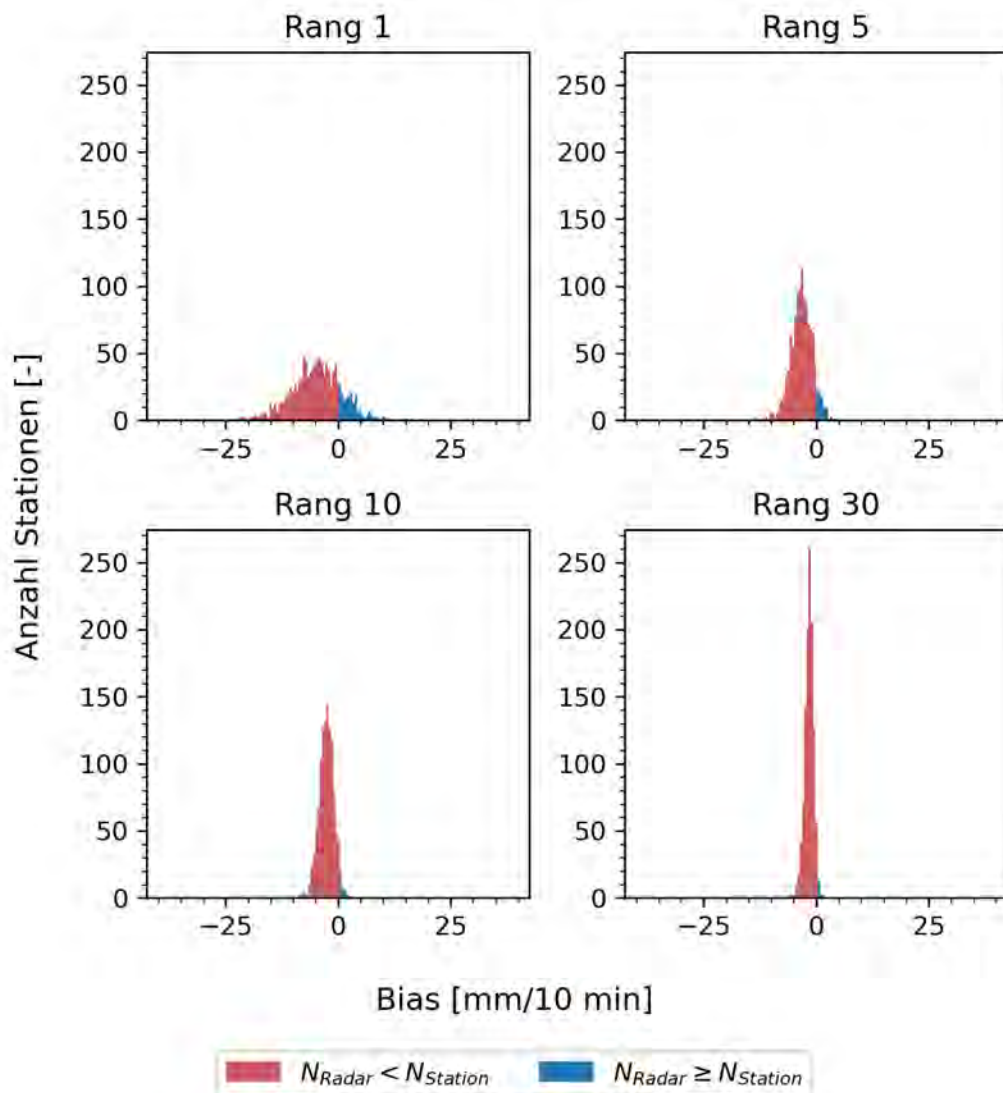


Abbildung 4.5: Histogramm des Bias (Stat-Px) der Ränge 1, 5, 10 und 30 bezogen auf die Dauerstufe D10

für die ersten zwei bis fünf Event-Ränge, und damit den extremsten Ereignissen, am größten und nimmt folgend mit zunehmenden Rang ab.

Zuletzt wurde das Bias der verschiedenen Biastypen untersucht. Die Gegenüberstellung der verschiedenen Biastypen lässt hier nur einen geringen Einfluss des Biastyps auf die Höhe und Spannbreite des Bias erkennen. So ist zum einen die Spannbreite des Bias sehr ähnlich, zum anderen weisen auch die Häufigkeiten (Anzahl an Stationen) eines bestimmten Bias zwischen Stat-Px, Mean4 und Mean9 nur geringe Unterschiede auf. Teilweise ist die Spannbreite des Bias bei Stat-Px etwas größer und die Häufigkeitsverteilung etwas flacher als bei Mean4 und Mean9. Ein klarer Trend ist jedoch nicht zu erkennen.

Die Analyse des Bias ergab somit, dass insbesondere die Dauerstufe und der Event-Rang den größten Einfluss auf das Bias haben, der Biastyp hingegen einen in dieser Arbeit vernachlässigbar geringen.

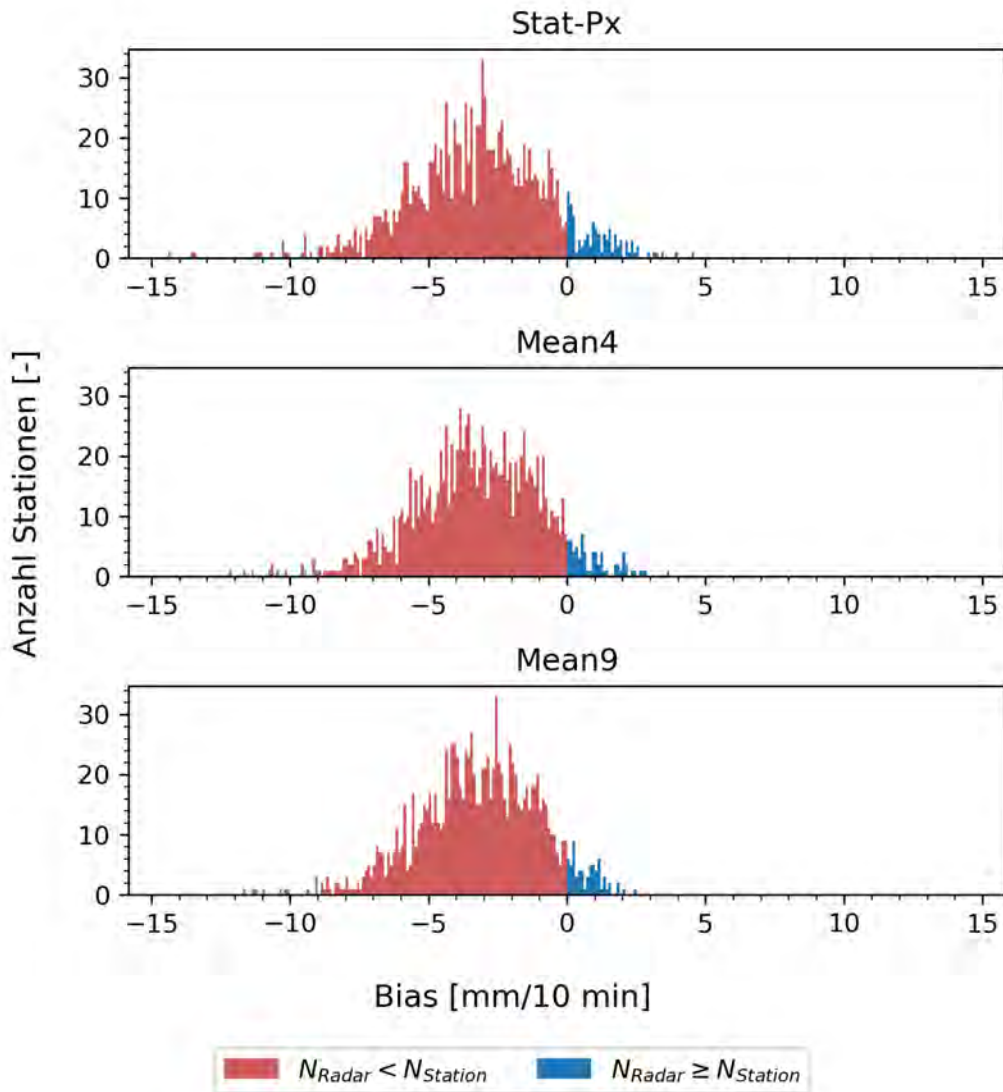


Abbildung 4.6: Vergleich des Biases von Stat-Px, Mean4 und Mean9 des fünften Ranges (D10)

4.2.2 Räumliche Muster

Neben der Untersuchung dieser Biasstatistiken wurde die räumliche Verteilung der Niederschläge und des Bias betrachtet (siehe Abbildung 4.7). Verglichen mit den Stationsniederschlägen zeigt die Radarkarte meist geringere Niederschlagsintensitäten an. Die Karte der Radarniederschläge von D10 zeigt für den ersten Event-Rang in Mitteleuropa ein Muster mit sehr hohen Niederschlagsintensitäten, das in der Karte der Stationsniederschläge nicht erkennbar ist. Hierbei handelt es sich wohl um ein Artefakt, was bei der Prozessierung der Radardaten noch nicht entfernt wurde.

Grundsätzlich fallen beim Betrachten der Radarkarte „Strahlen“ mit geringerem Niederschlag auf, was auf Negativspeichen hindeutet. Die prominenteste Negativspeiche ist im Bereich des Feldbergs zu sehen, weitere sind besonders Nord- und Nordostdeutschland zu erkennen. Zudem ist gerade bei Ostfriesland aber auch in im Osten Brandenburgs erkennbar, dass sich einige Radarschirme überlagern. Dies ist wohl auf die (meist zeitweise) Veränderung des Radarstandortes auf ein Ausfallsicherungsradar während Wartungsarbeiten oder Umrüstungen zurückzuführen, da so über den Zeitraum von beiden Standorten Messergebnisse vorliegen.

Während für einige Regionen Deutschlands die Karten der Stations- und Radarniederschläge ähnliche Muster aufweisen, zeigen sich in anderen Regionen auch durchaus Abweichungen. Da-

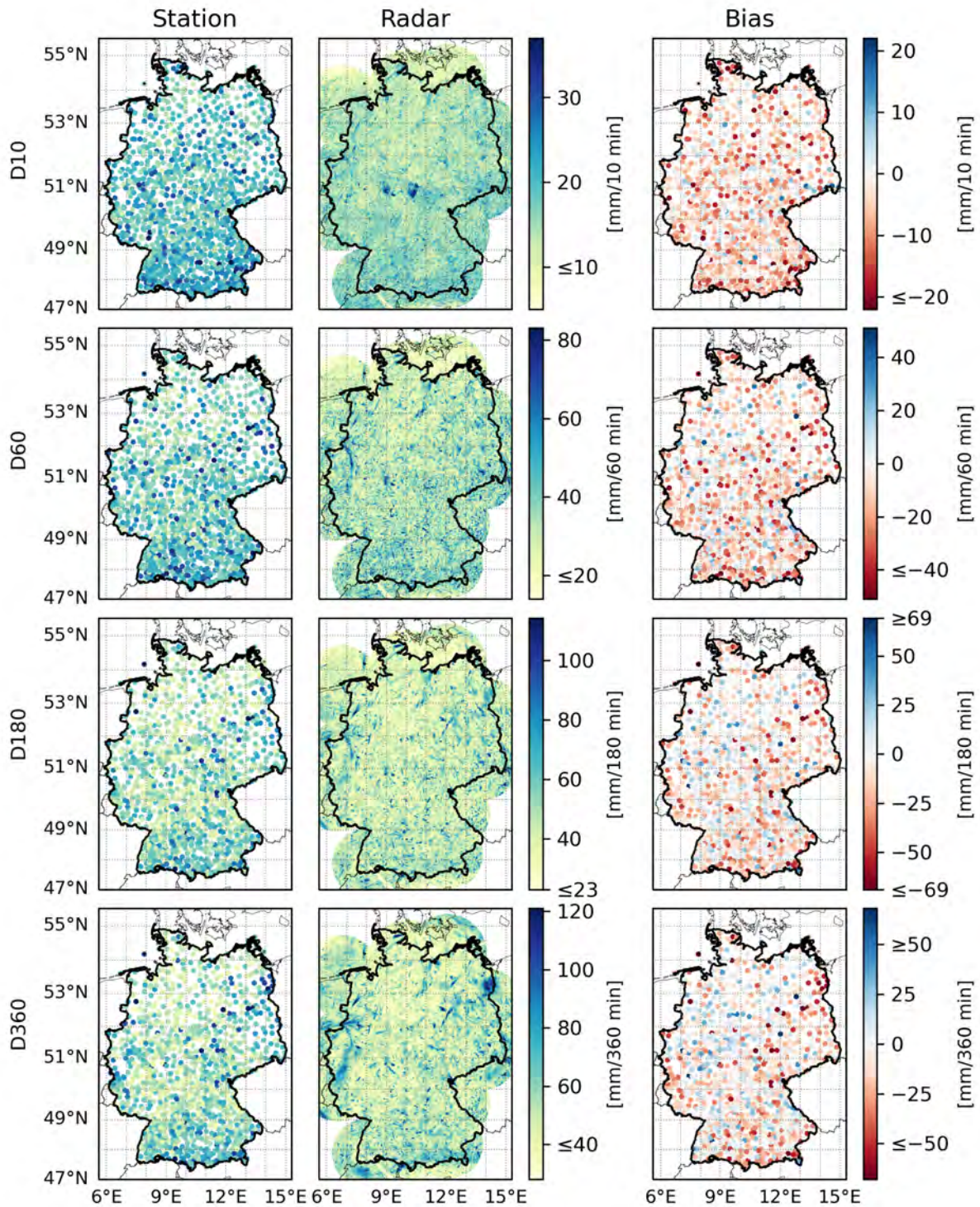


Abbildung 4.7: Räumliche Verteilung der Niederschläge und des Bias für den ersten Event-Rang (D10). Gut zu erkennen ist das zufällige Muster der Niederschläge. (©EuroGeographics bezüglich der Verwaltungsgrenzen)

bei ist dies von Rang und Dauerstufe abhängig. Für die ersten Ränge und kleinen Dauerstufen scheint die räumliche Übereinstimmung zwischen Stations- und Radarprodukt schlechter auszufallen als für größere Ränge (vgl. Abbildung 7.5). Je größer der Event-Rang ist, desto mehr stimmen die räumlichen Muster beider Datensätze überein. Gleichzeitig folgt das Muster auch immer mehr der Orographie der Mittelgebirge, mit den höchsten Niederschlägen im Schwarzwald, dem Alpenvorland und dem Alpenraum. Auch im Thüringer Wald, im Harz und teilweise

im Süderbergland, Vogelsberg und Rhön sind höhere Niederschlagsintensitäten zu sehen. Dagegen sind bei den ersten Event-Rängen die Stationen mit hohen Niederschlägen recht zufällig über Deutschland verteilt. Bei den Dauerstufen D60, D180 und D360 zeigt sich der gleiche Trend, hier folgt das Muster jedoch schon bei den kleineren und mittleren Rängen der Orographie. Dieses räumliche Muster der Niederschläge paust sich grundsätzlich auch auf das Muster des absoluten und relativen Bias durch.

Bei Betrachtung des Bias fällt auf, dass mit zunehmenden Rang die Anzahl der Stationen mit positivem Bias abnimmt. Weiter ist der Anteil der Stationen mit positivem Bias bei den größeren Dauerstufe wie D180 und D360 größer als bei den kleinen Dauerstufen. Generell ist eine Zunahme des positiven Bias mit zunehmender Dauerstufe erkennbar. Bei D180 und D360 ist der Anteil der Stationen mit positivem Bias auch für die größeren Ränge höher als für D10 oder D60. Auch beim positiven Bias zeigt sich für die ersten Ränge kein deutliches räumliches Muster, für die letzten Ränge der großen Dauerstufen fallen insbesondere die Tallagen zwischen den Gebirgen in Mitteldeutschland, das Tiefland des nördlichen Oberrheins und Teile Bayerns, Brandenburgs und Mecklenburg-Vorpommerns mit vermehrten positivem Bias auf. Schließlich wurde die Abhängig des räumlichen Musters vom verwendeten Biastyp untersucht. Zwischen den absoluten und relativen Biaskarten von Stat-Px, Mean4 und Mean9 sind jedoch genauso wie bei den Histogramme kaum Unterschiede erkennbar.

4.2.3 Zusammenhänge mit anderen Parametern

Die Betrachtung der Beziehung zwischen Stationshöhe und Bias zeigt insbesondere bei den ersten Rängen keinen eindeutigen Trend, mit zunehmenden Rang ist ein leicht negativer Trend mit zunehmender Stationshöhe sichtbar, insgesamt ist die Streuung jedoch sehr groß (siehe Abbildung 4.8). Auch hier zeigt sich, dass es mehr Standorte mit einem negativen Bias als solche mit positivem Bias gibt, bei den größeren Dauerstufen der Anteil der Standorte mit negativem Bias jedoch abnimmt. Während bei den ersten Rängen der 150 Events die Streuung im Diagramm relativ homogen ist, wird die Streuung bei den weiteren Event-Rängen für größere Stationshöhen breiter. Insgesamt zeigt sich eine größere Streuung im negativen Biasbereich als im positiven. Das relative Bias zwischen den Dauerstufen weist nur geringe Unterschiede in der Streuung auf (vgl. 7.6). Die gleiche Analyse mit Mean4, Mean9, Min oder Max als Biastyp (nicht abgebildet) zeigt ähnliche Muster.

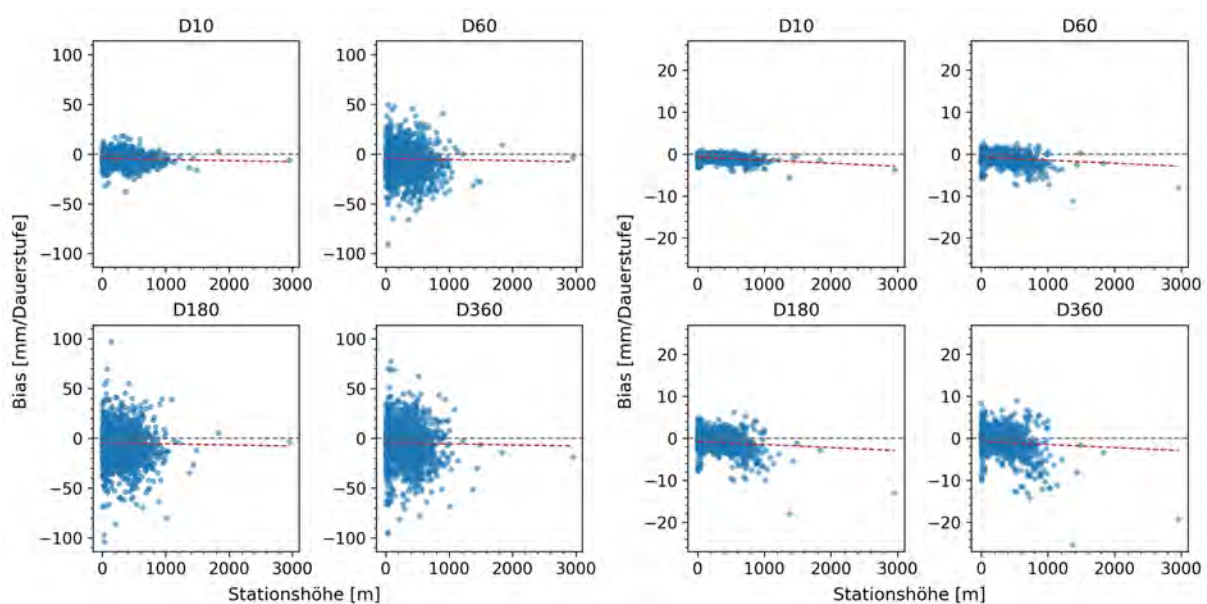


Abbildung 4.8: Zusammenhang zwischen Bias und Stationshöhe. Dargestellt sind der erste (links) und 75. Event-Rang (rechts)

Für alle 150 Events und alle Dauerstufen weisen die stark geschützten Standorte im Vergleich zu den gering geschützten vermehrt ein negatives Bias auf, Überschätzungen der Niederschläge durch die Radardaten treten also weniger auf.

Tabelle 4.1: Anzahl der Stationen pro Richterklasse (DWD, 2023)

Richter-Klasse	Anzahl an Stationen
ungeschützt	833
leicht geschützt	225
geschützt	65
stark geschützt	31

Während bei den ersten Event-Rängen kaum ein Trend sichtbar ist, ist besonders bei D180 und D360 mit zunehmendem Event-Rang eine Abnahme des absoluten sowie relativen Bias von ungeschützter hin zur stark geschützter Richter-Klasse erkennbar (siehe Abbildung 4.9). Werden die Daten der 150 Events nicht einzeln, sondern zusammen untersucht, ist eine Zunahme der Spannweite des Bias von ungeschützten hin zu stark geschützten Standorten erkennbar.

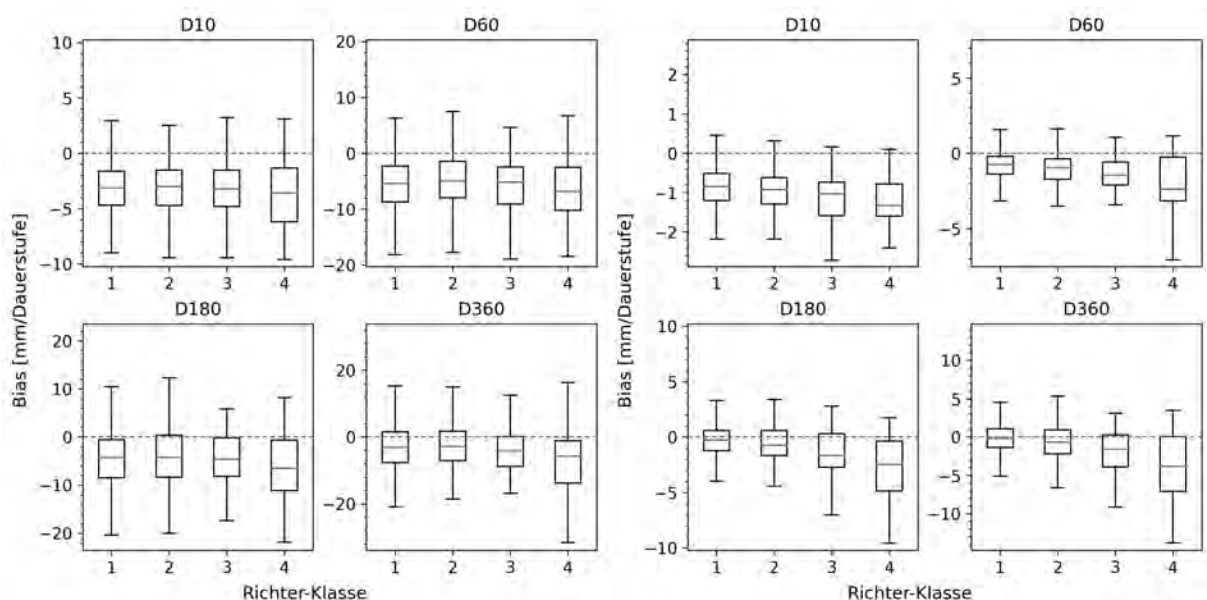


Abbildung 4.9: Zusammenhang zwischen Bias und Richter-Klasse. Dargestellt sind der fünfte (links) und 80. Event-Rang (rechts). Die Richter-Klassen sind 1: ungeschützter, 2: leicht geschützt, 3: geschützter, 4: stark geschützter Standort.

In einem letzten Schritt wurde das ortsbezogene Bias pro Radarturm und der Einfluss der Entfernung einer Niederschlagsstation vom Radarturm auf das Bias untersucht. Für die 10-minütige Dauerstufe ist der Großteil der Bias-Werte pro Radarstandort negativ. Für die weiteren Dauerstufen nimmt der Anteil der positiven Bias-Werte zu. Diese Entwicklung zeigt sich für jeden Radarstandort (siehe Abbildung 4.10). Bei Betrachtung des absoluten Bias fallen die Radartürme Feldberg (6), Isen (9) und Memmingen (10) mit einer großen Spannweite des Bias auf, bei dem relativen Bias fallen besonders die Radarstandorte Boostedt (2), Feldberg (6), Rostock (15) und Türkheim (16) auf (siehe Abbildung 7.7).

Bei der Untersuchung des Einflusses der Distanz zum Radarstandort auf das Bias sind teils sehr unterschiedliche Trends erkennbar (vgl. Abbildung 7.8). Zu beachten ist dabei, dass grundsätzlich eine große Streuung auftritt, was die Signifikanz dieser Trends in Frage stellt. Bei den meisten Radartürmen ist ein Anstieg des Bias mit zunehmender Entfernung vom Radarturm, bei wenigen Türmen jedoch auch eine Abnahme des Bias mit zunehmender Entfernung feststellbar. Die Anzahl der Radartürme mit einem über die Distanz abnehmenden Bias ist bei größeren

Dauerstufen (D180, D360) geringfügig größer als bei den kleineren (D10, D60). Schließlich zeigen einige Radarstandorte auch keinen klar erkennbaren Trend (neutral) oder einen wechselnden Trend zwischen den 150 Events auf. Zu Beachten ist jedoch, dass dies nur den überwiegenden Trend der 150 Events darstellt. Oftmals weichen gerade die etwa ersten 15 Event-Ränge von den restlichen ab.

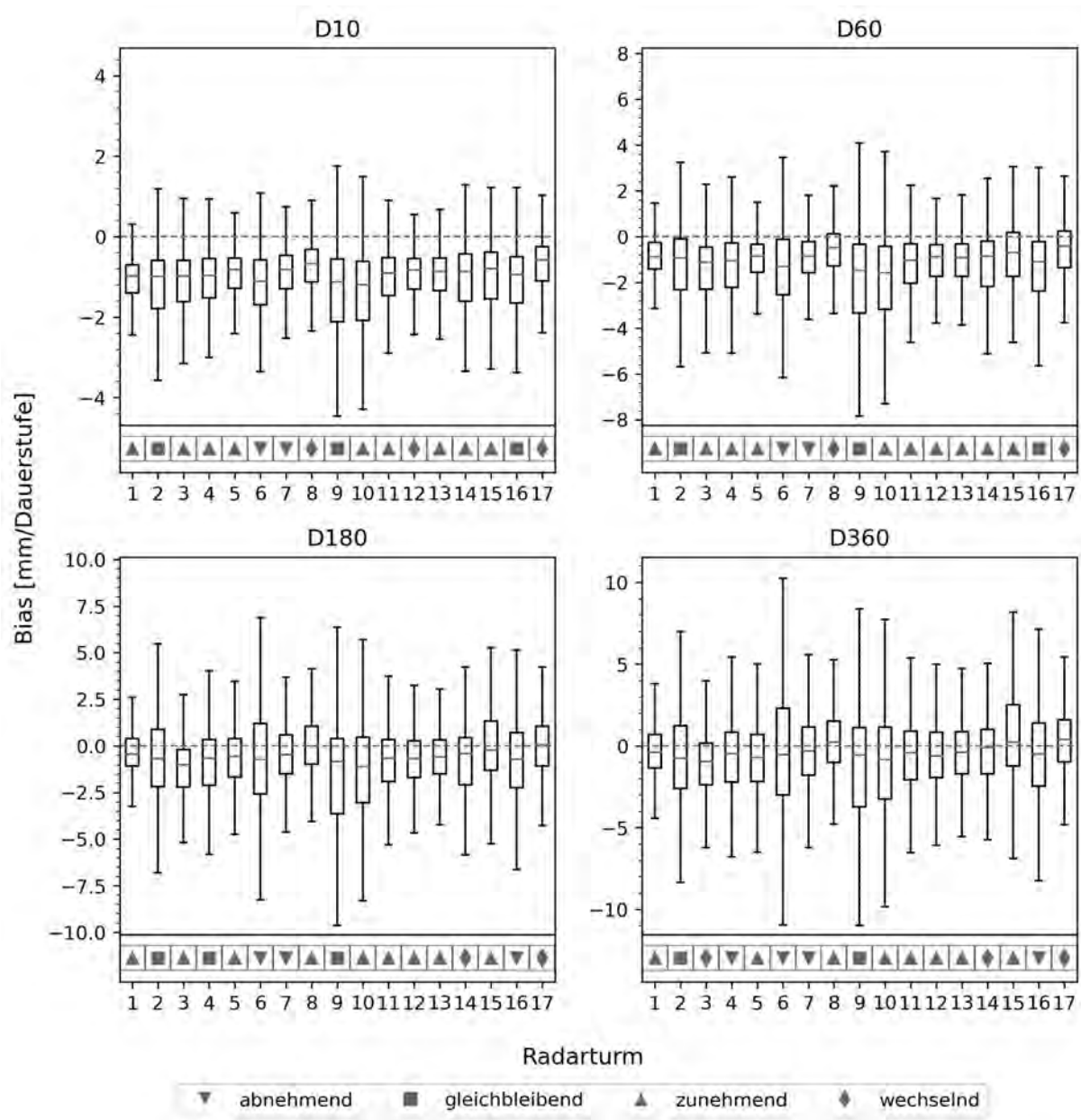


Abbildung 4.10: Zusammenhang zwischen Bias und dem jeweiligen Radarturm. Die Symbole unterhalb der Grafik stellen Indikatoren dar, ob sich das Bias mit zunehmender Entfernung verändert. Zur Erläuterung der Radar-Nr. siehe Tabelle 7.1

4.2.4 Bias von realen Events

Da die betrachteten Niederschlagsereignisse kumulative Summen darstellen, wurde zuletzt beispielhaft das Bias von realen Extremereignissen betrachtet. Dazu wurde je ein Starkregenereignis von insgesamt neun Stationen untersucht. Dabei zeigen sich große Unterschiede zwischen den neun Stationen. Teilweise zeigt sich eine große Unterschätzung der Stationsniederschläge durch die Radardaten (v. a.. Station 555, 970 oder 3376), teilweise eine mehr oder weniger gute Über-

einstimmung (Station 2319, 567, 684). Während bei einigen Stationen die Niederschlagssummen zumindest noch in der Spannweite der neun Radar-Pixel liegen (Station 684), trifft auch dies bei anderen Stationen nicht zu. Neben den Abweichungen in der Niederschlagssumme ist bei manchen Stationen ein zeitlicher Versatz zwischen dem Stations- und dem Radar-Ereignis erkennbar (z. B. Station 5029, 5763).

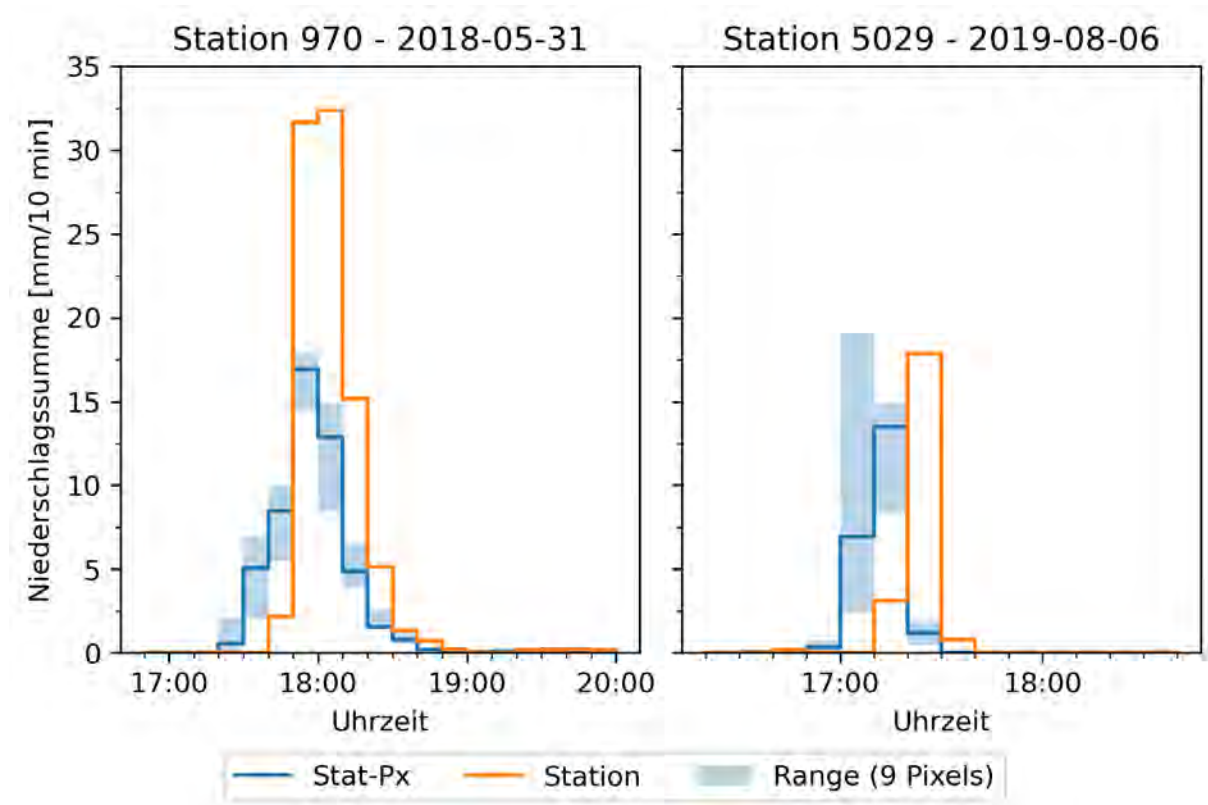


Abbildung 4.11: Beispiel des Bias zweier realer Events. Neben dem Niederschlagswert des Stations-Pixels ist die Spannweite der neun Pixel zu sehen.

4.3 Biaskorrektur - Verfahren 1

Anhand von Quantile-Quantile-Plots (Q-Q-Plots) und der empirischen kumulativen Verteilungsfunktionen (ECDFs) soll die Performance des Verfahrens abgeschätzt werden. Nach Korrektur der Radardaten der Trainingsgruppe zeigt sich erwartungsgemäß eine sehr gute Übereinstimmung zwischen Stations- und korrigierten Radardaten (siehe Abbildung 4.12). Auch die Korrektur der Kontrollgruppe führt grundsätzlich zu einer besseren Übereinstimmung der zwei Datensets. Im Vergleich zur korrigierten Trainingsgruppe sind jedoch insbesondere für die kleinen und großen Quantile vermehrt Abweichungen erkennbar, während der mittlere Bereich eine gute Übereinstimmung aufweist.

Für die großen Quantile (hohe Niederschlagsintensitäten) führte die Korrektur teilweise auch zu einer „Überkorrektur“ der Werte, sodass die Stationsniederschläge nun überschätzt werden. Für die kleinen Quantile zeigt sich, dass die Korrektur nicht vollständig ausreicht, um das Bias vollständig zu beseitigen, hier werden die Stationsniederschläge weiterhin unterschätzt. Die empirischen kumulativen Verteilungsfunktionen (ECDFs) weisen dieselben Merkmale auf: die Korrektur führt zu einer deutlichen Reduzierung des Bias, die größten Abweichungen bestehen weiterhin für die kleinsten und größten Quantile bzw. Unterschreitungswahrscheinlichkeiten (siehe Abbildung 4.13). Gut erkennbar ist das mit größerer Dauerstufe abnehmende Bias (vor der Korrektur). Während das Verhältnis zwischen Stations- und Radarniederschläge für die ersten

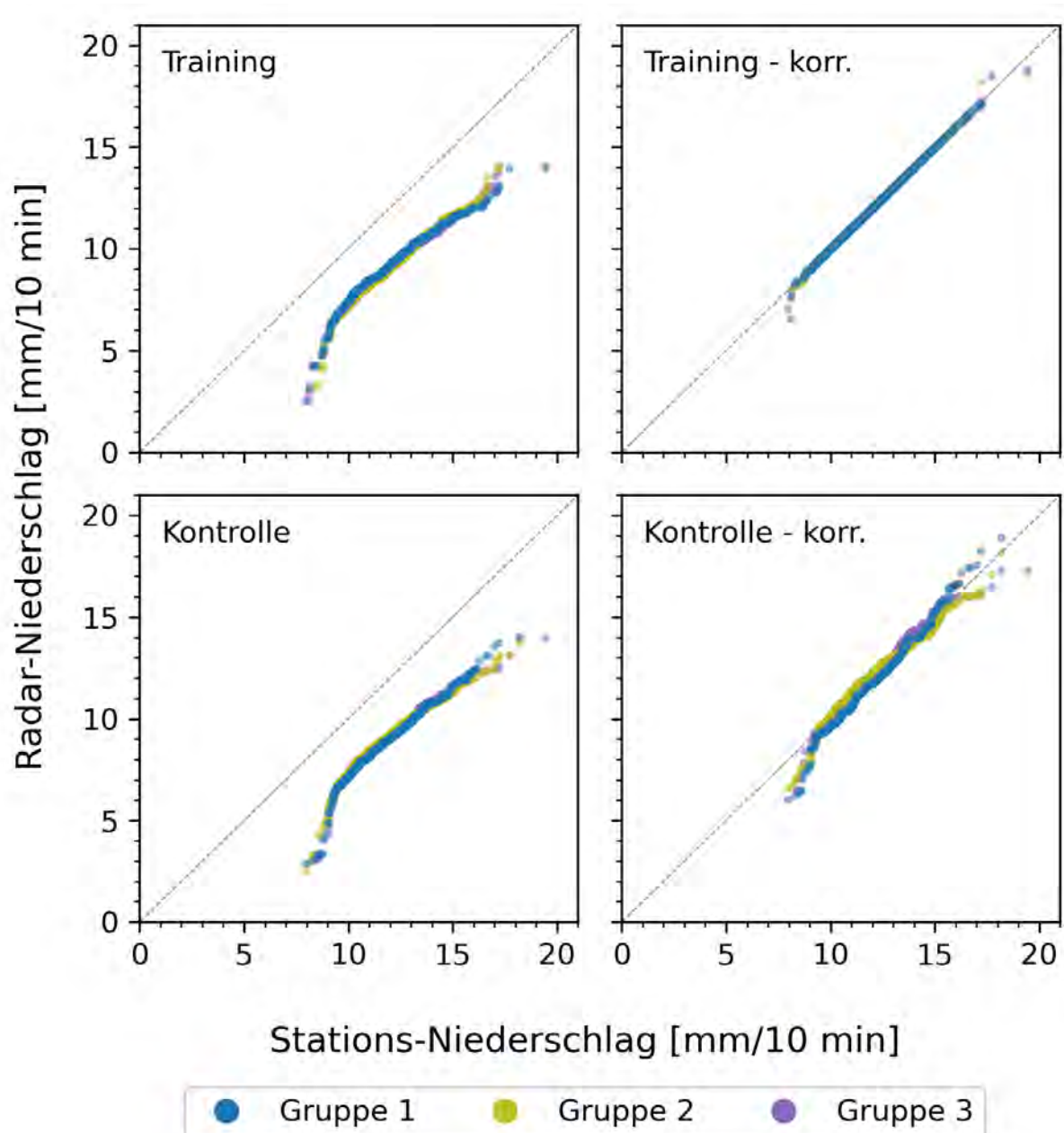


Abbildung 4.12: Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D10). Dargestellt sind alle drei Gruppen.

Event-Ränge einem eher linearen Verlauf folgt, zeigt sich bei den späteren Rängen ein eher S-förmiger Verlauf (siehe Abbildung 4.14). Bei Betrachtung der verschiedenen Event-Ränge zeigen sich dieselben Charakteristiken, die Abweichungen für die großen Quantile reduzieren sich (absolut) gesehen mit zunehmenden Rang und damit abnehmender Eventstärke.

Zwischen den Dauerstufen lassen sich weder in den Q-Q-Plots noch in den ECDFs signifikanten Abweichungen in der Performance der Korrektur feststellen (vgl. Abbildungen 4.13 sowie 7.9, 7.10, 7.11). Insgesamt sind die Unterschiede zwischen den drei Gruppen recht gering, Unterschiede sind hauptsächlich für die größten Quantile der Kontrollgruppe sichtbar (siehe Abbildung 4.15).

Bei den Radardaten zeigen sich größere Differenzen zwischen den drei Gruppen als bei den Stationsdaten und bei der Kontrollgruppe größere Differenzen als bei der Trainingsgruppe. Die größten Unterschiede zeigen sich dabei bei den kleinsten und größten Niederschlagsintensitäten (Unterschreitungswahrscheinlichkeiten). Dieses Muster zeigt sich für alle Event-Ränge. Bei der Analyse der relativen Differenzen wird deutlich, dass die Unterschiede zwischen den Gruppen für

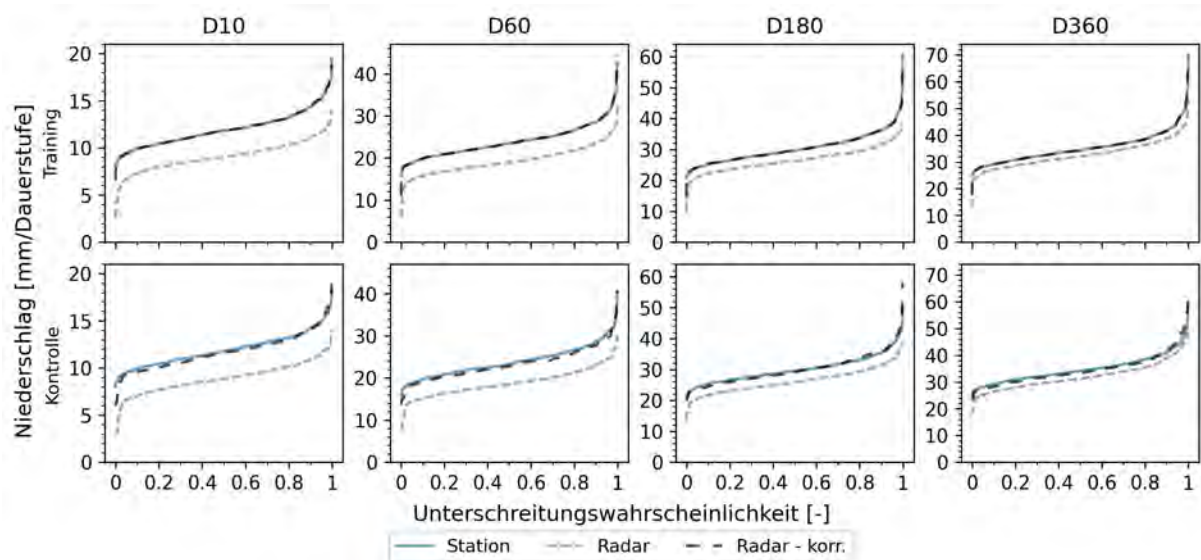


Abbildung 4.13: Verfahren 1: ECDFs für den siebten Event-Rang

kleine Niederschlagsintensitäten relativ gesehen am größten sind.

Abbildung 4.16 zeigt die Verteilung der Korrekturfaktoren, bezogen auf die Unterschreitungswahrscheinlichkeiten der zugehörigen ECDFs.

Auffallend sind besonders die hohen Werte für kleine Niederschlagsintensitäten, die deutlich größer ausfallen als die restlichen Korrekturfaktoren. Aber auch für die größten Niederschlagsintensitäten ist ein Anstieg des Korrekturfaktors erkennbar, wobei diese nicht so hohe Werte wie für die kleinen Intensitäten erreichen. Dieser Anstieg ist für die längeren Dauerstufen (D180 und D360) stärker ausgeprägt als für die kürzeren (D10 und D60). Beim Vergleich der Event-Ränge ist eine Zunahme des Korrekturfaktors für kleine Unterschreitungswahrscheinlichkeiten mit zunehmendem Rang feststellbar. Das bedeutet also, dass die niedrige Niederschlagsintensitäten relativ gesehen bei den letzten Event-Ränge stärker korrigiert werden, als bei den ersten Event-Ränge. Des Weiteren ist mit zunehmender Dauerstufe eine Reduzierung der Faktoren erkennbar.

Auch die Unterschiede zwischen den Gruppen sind für die kleinen Niederschlagsintensitäten (Unterschreitungswahrscheinlichkeiten) größer als für die restlichen Bereiche der Verteilung (siehe Abbildung 7.12 und 7.13). Für die ersten Event-Ränge sind diese Gruppenunterschiede stärker ausgeprägt als für die späteren Ränge.

Weiter wurde die Performance des Verfahrens für die Anwendung auf die Gesamtfläche Deutschlands untersucht. Entsprechend der Erwartung führte die Anwendung der Biaskorrektur zu einer Erhöhung der Radarniederschläge (siehe Abbildung 4.17). Die stärksten Korrekturen sind dabei in Gebieten feststellbar, die bereits vor der Korrektur hohe Werte aufwiesen. Das räumliche Muster ist vor und nach der Korrektur dasselbe, alleinig die Niederschlags-Intensitäten wurden durch die Korrektur angepasst.

Um den Einfluss der drei Trainings- und Kontrollgruppen zu untersuchen, wurden die zugehörigen Sets an Korrekturfaktoren flächig auf ganz Deutschland angewandt und die Ergebnisse miteinander verglichen. Dabei fallen sowohl Flächen mit einer positiven als auch Flächen mit einer negativen Abweichung auf (siehe Abbildung 4.18). Insgesamt sind die Gruppenunterschiede für den Großteil Deutschlands jedoch recht gering. Auch die Gruppenunterschiede sind in den Bereichen, die vor der Korrektur hohe Niederschlagsintensitäten aufwiesen stärker ausgeprägt.

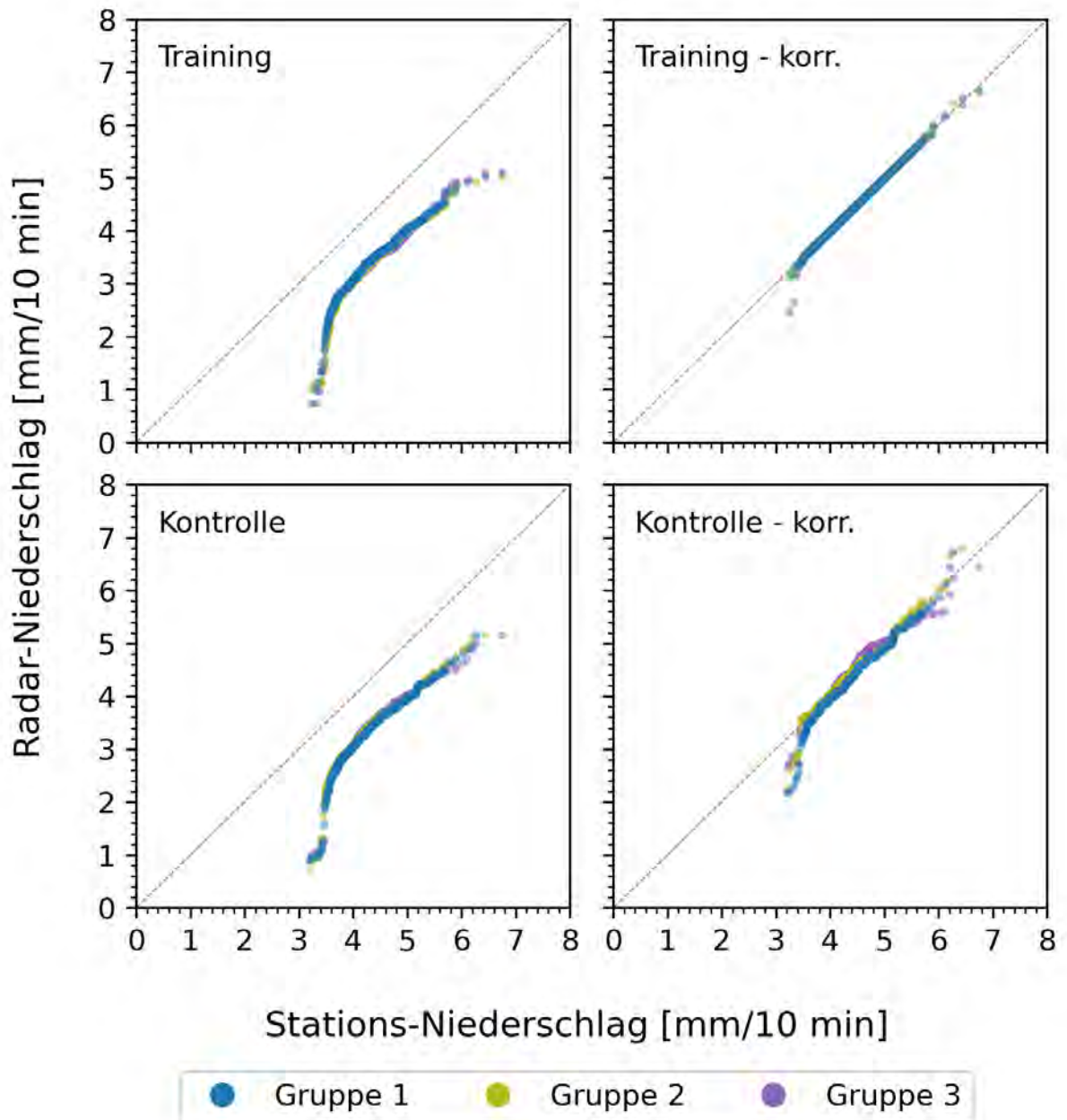


Abbildung 4.14: Verfahren 1: Q-Q-Plot für den 75. Event-Rang (D10)

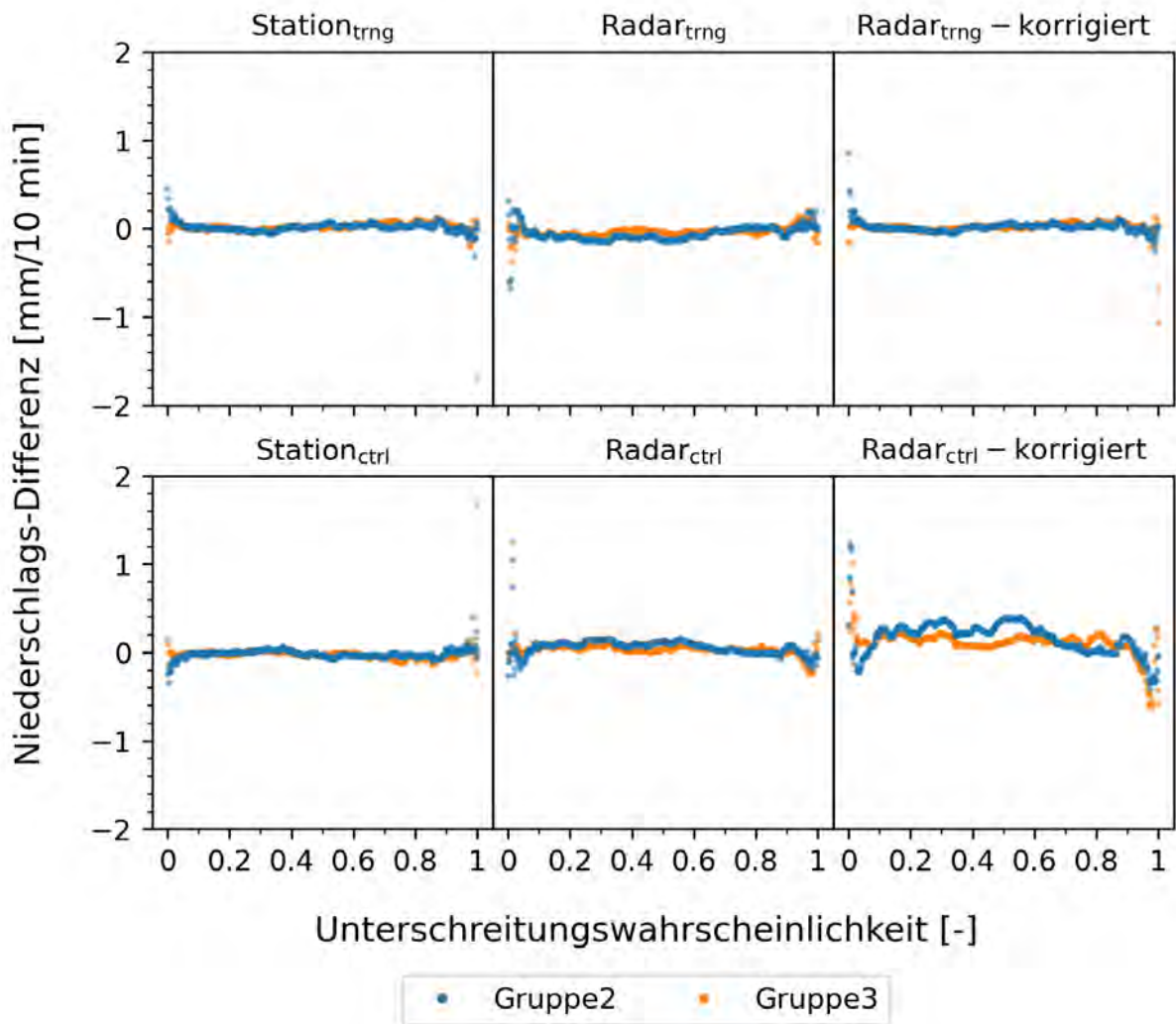


Abbildung 4.15: Verfahren 1: Unterschiede in den ECDFs zwischen den Gruppen für den 20. Event-Rang (D10) verglichen mit Gruppe 1

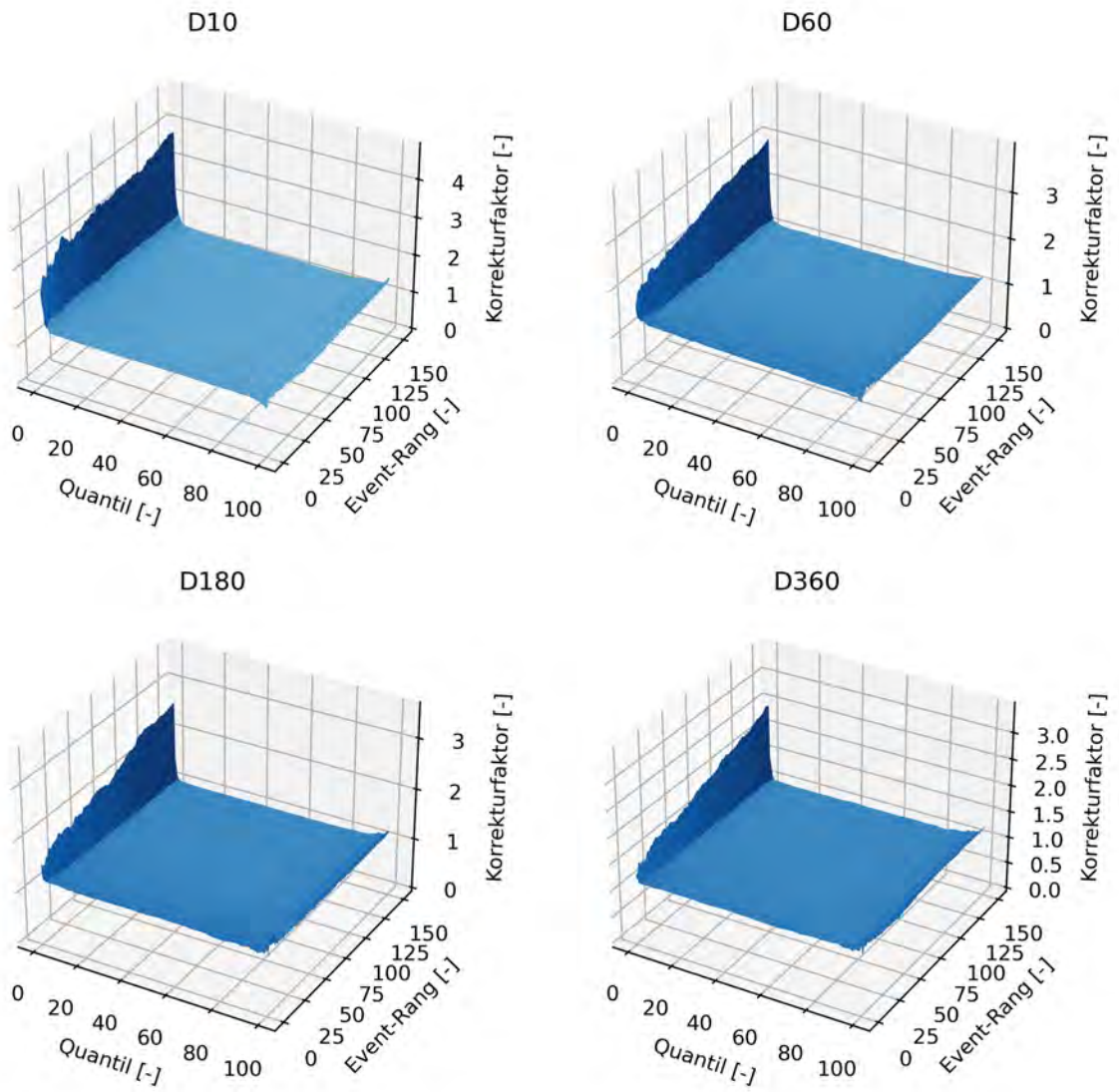


Abbildung 4.16: Verfahren 1: Korrekturfaktoren für alle Event-Ränge, Quantile und Dauerstufen

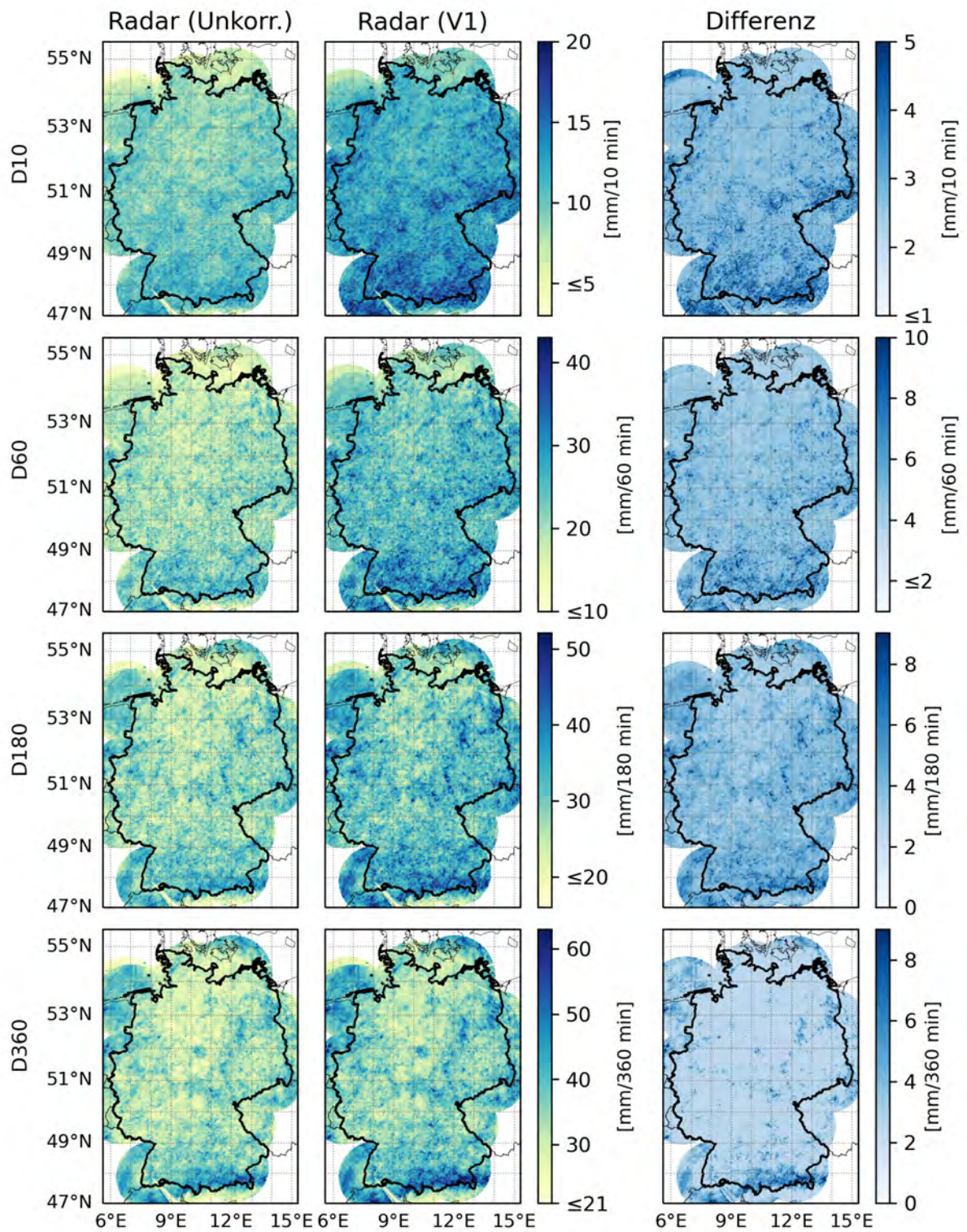


Abbildung 4.17: Anwendung der mithilfe von Verfahren 1 ermittelten Korrekturfaktoren auf alle Rasterzellen (©EuroGeographics bezüglich der Verwaltungsgrenzen)

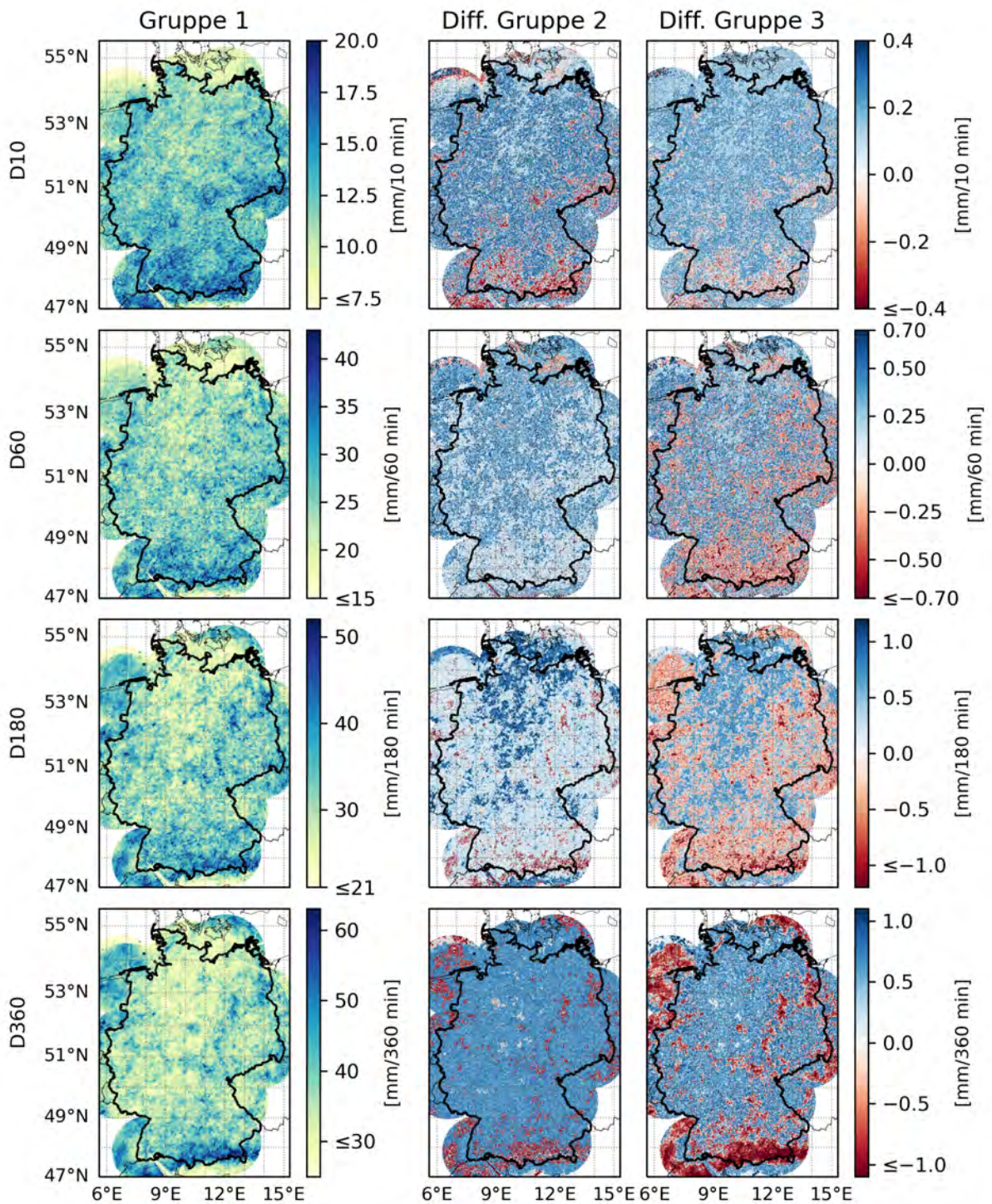


Abbildung 4.18: Verfahren 1: Vergleich der korrigierten Radar-Niederschlagskarten der drei Gruppen (©EuroGeographics bezüglich der Verwaltungsgrenzen)

4.4 Biaskorrektur - Verfahren 2

Bei der Betrachtung der Q-Q-Plots und ECDFs von Verfahren 2, zeigt sich insgesamt ein ähnlicher Verlauf zu Verfahren 1, jedoch sind auch einige Unterschiede erkennbar.

Die Korrektur der Trainingsgruppe und die Korrektur der Kontrollgruppe führen meist zu ähnlich guten bzw. schlechten Übereinstimmungen zwischen den Radardaten und den Stationsdaten. Grundsätzlich kann Verfahren 2 die kleinen Niederschlagsintensitäten nur unzureichend korrigieren (siehe Abbildung 4.19). Bei allen Dauerstufen und dem Großteil aller Ränge unterschätzen die korrigierten Radardaten im Bereich der niedrigeren Niederschlagsintensitäten die Stationsniederschläge. Auch bei Betrachtung der ECDF fallen die gleichen Unterschiede auf (siehe Abbildung ??). Besonders der schlechtere Fit der korrigierten Radardaten der Trainingsgruppe an die Stationsdaten fällt auf.

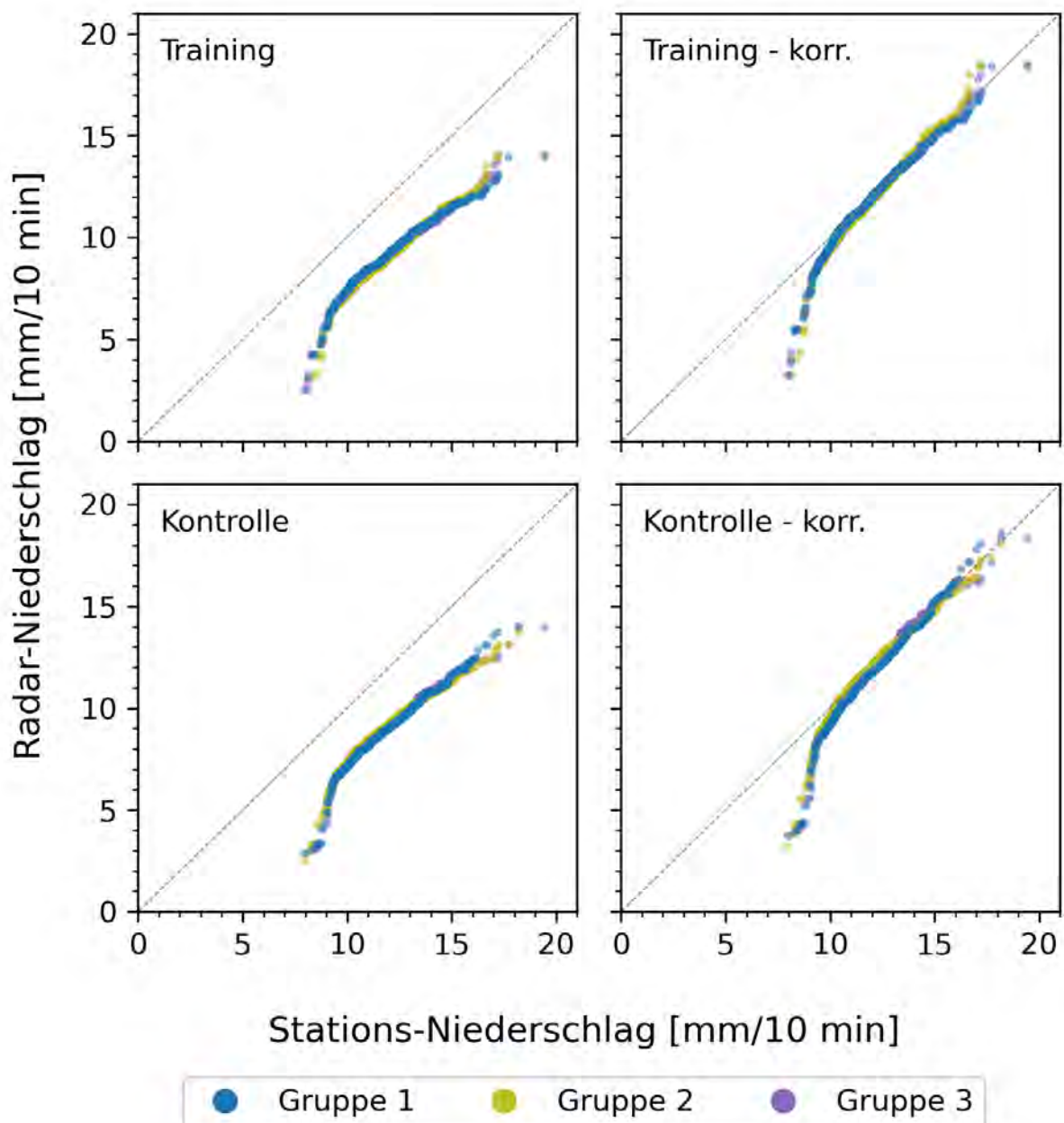


Abbildung 4.19: Verfahren 2: Q-Q-Plot für den siebten Event-Rang (D10). Dargestellt sind alle drei Gruppen.

Während die großen Quantile der mittleren Event-Ränge am besten korrigiert werden (vgl.

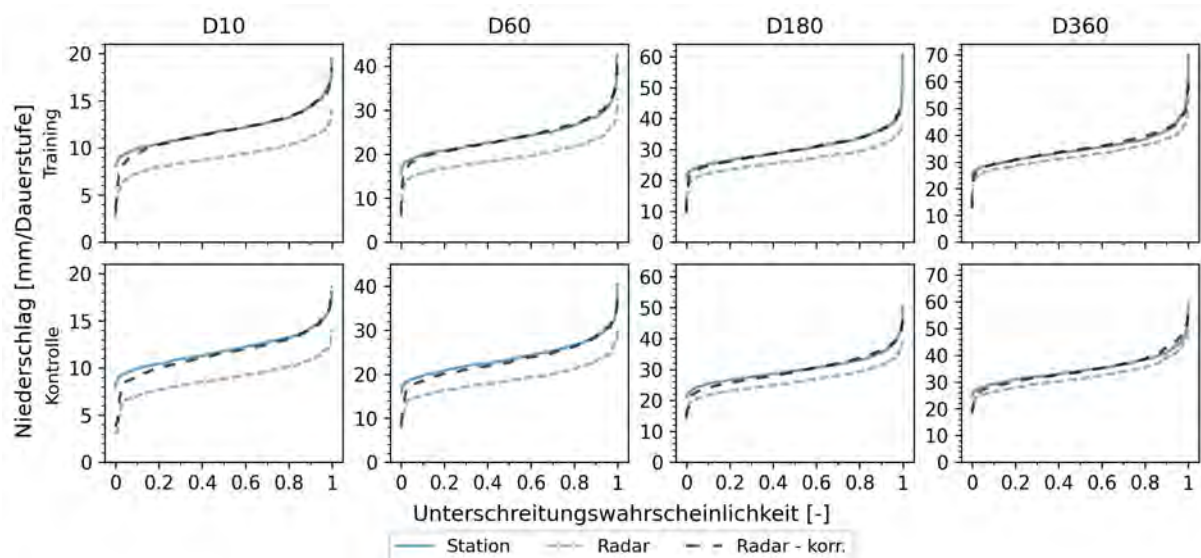


Abbildung 4.20: Verfahren 2: ECDFs für den siebten Event-Rang

Abbildung 7.14), kommt es bei den ersten Rängen teils zu einer Überschätzung des Stationsniederschlags und bei den letzten Rängen meist zu einer Unterschätzung. Insbesondere bei D180 und D360 ist dies ausgeprägt. Die Unterschätzung der Stationsniederschläge durch die korrigierten Radardaten für die größten Quantile der späteren Ränge ist auch gut in den ECDFs erkennbar.

Begründet durch die verhältnismäßig geringen Korrekturfaktoren für mittlere Quantile, findet bei D180 und besonders bei D360 in diesem Bereich nur eine geringfügige Korrektur statt. Jedoch weist dieser Bereich für diese Dauerstufen und für die mittleren und späteren Ränge bereits vor der Korrektur ein geringes Bias zu den Stationsdaten auf. Die Unterschiede zwischen den Gruppen sind bei den ersten Rängen am stärksten ausgeprägt.

Verglichen mit Verfahren 1 ist die Korrektur der Trainingsgruppe bei Verfahren 2 schlechter, insbesondere für die kleinen und großen Quantile. Allerdings weist Verfahren 2, bis auf die ersten Ränge, etwas geringere Differenzen zwischen den drei Gruppen auf als Verfahren 1.

Für große Quantile ist bei Verfahren 1 festzustellen, dass einige Gruppen „überkorrigiert“ werden, während bei Verfahren 2 für diese Quantile oftmals keine zu starke, sondern eine passende oder teils zu schwache Korrektur stattfindet.

Der Verlauf der Korrekturfaktoren ähnelt dem des ersten Verfahrens. Auch hier sind die größten Korrekturfaktoren im Bereich der kleinsten Niederschlagsintensitäten zu finden und weiter ein Anstieg der Faktoren für die größeren Niederschlagsintensitäten festzustellen. Dieser Anstieg ist stärker ausgeprägt als bei Verfahren 1 und zeigt sich hauptsächlich bei den Dauerstufen D60, D180 und D360, jedoch kaum bei D10.

Ebenso ist eine Abnahme der Korrekturfaktoren bei größeren Dauerstufen (z. B. D360) gegenüber den kleineren Dauerstufe (z. B. D10) zu erkennen.

Die Faktoren für die kleinsten Unterschreitungswahrscheinlichkeiten sind bei Verfahren 2 geringer als die entsprechenden Faktoren der ersten Ränge des ersten Verfahrens (siehe Abbildung 4.21). Dies trifft auf alle Dauerstufen zu. Meist entsprechen diese eher Faktoren im vorderen mittleren Bereich der Event-Ränge von Verfahren 1. So betragen beispielsweise für D10 die größten Korrekturfaktoren von Verfahren 2 etwa 3.0, während die entsprechenden Faktoren von Verfahren 1 einen Bereich von etwa 2.4 bis 4.0 abdecken. Im Vergleich zu Verfahren 1 sind die Unterschiede im Verlauf der Korrekturfaktoren zwischen den Gruppen deutlich geringer ausgeprägt.

Auch bei Verfahren 2 führte die Anwendung der mithilfe des Quantile-Mapping-Verfahrens ermittelten Korrekturfaktoren auf die gesamte Fläche Deutschlands zu einer Erhöhung der Niederschlagsintensitäten. Ebenso wie bei Verfahren 1 wurden die Regionen mit bereits großen

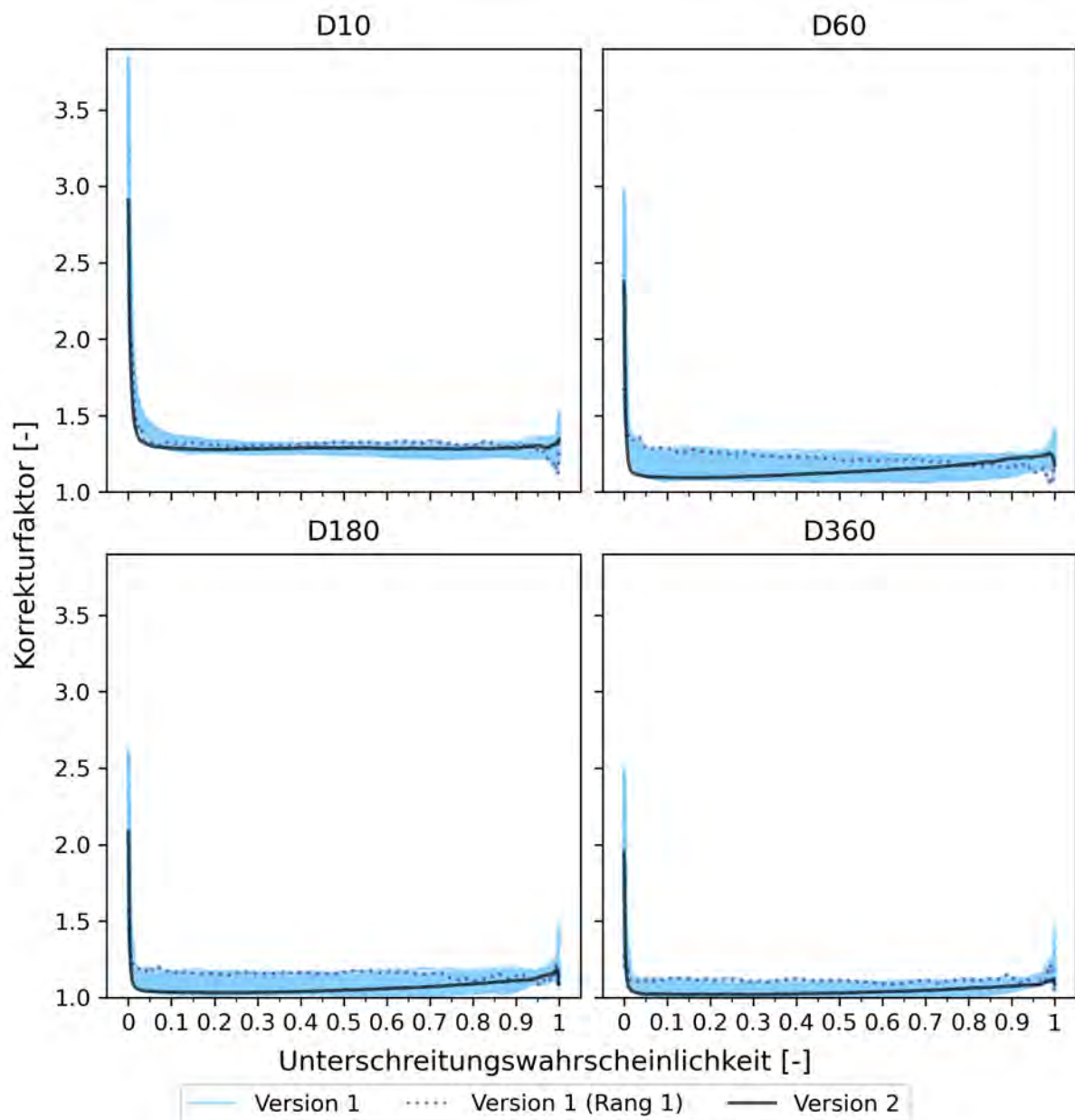


Abbildung 4.21: Verfahren 2: Korrekturfaktoren für alle Dauerstufen. In hellblau sind alle 150 Sets an Korrekturfaktoren von Verfahren 1 abgebildet, die gepunktete dunkelblaue Linie zeigt den Verlauf der Korrekturfaktoren bei dem ersten Event-Rang von Verfahren 1. Dargestellt ist nur Gruppe 1.

Niederschlägen am stärksten korrigiert (siehe Abbildung 4.22).

Die Differenz zu den unkorrigierten Radarniederschlägen weist im Vergleich zu Verfahren 1 jedoch eine etwas niedrigere Spannbreite auf, was sich auch für die anderen Ränge zeigt.

Auch hier sind Flächen mit positiver und negativer Abweichung zu sehen. Im Vergleich zu Verfahren 1 sind die Gruppenunterschiede jedoch insgesamt geringer, was u. a. an der geringeren Spannbreite der Differenz sichtbar wird (siehe Abbildung 4.23).

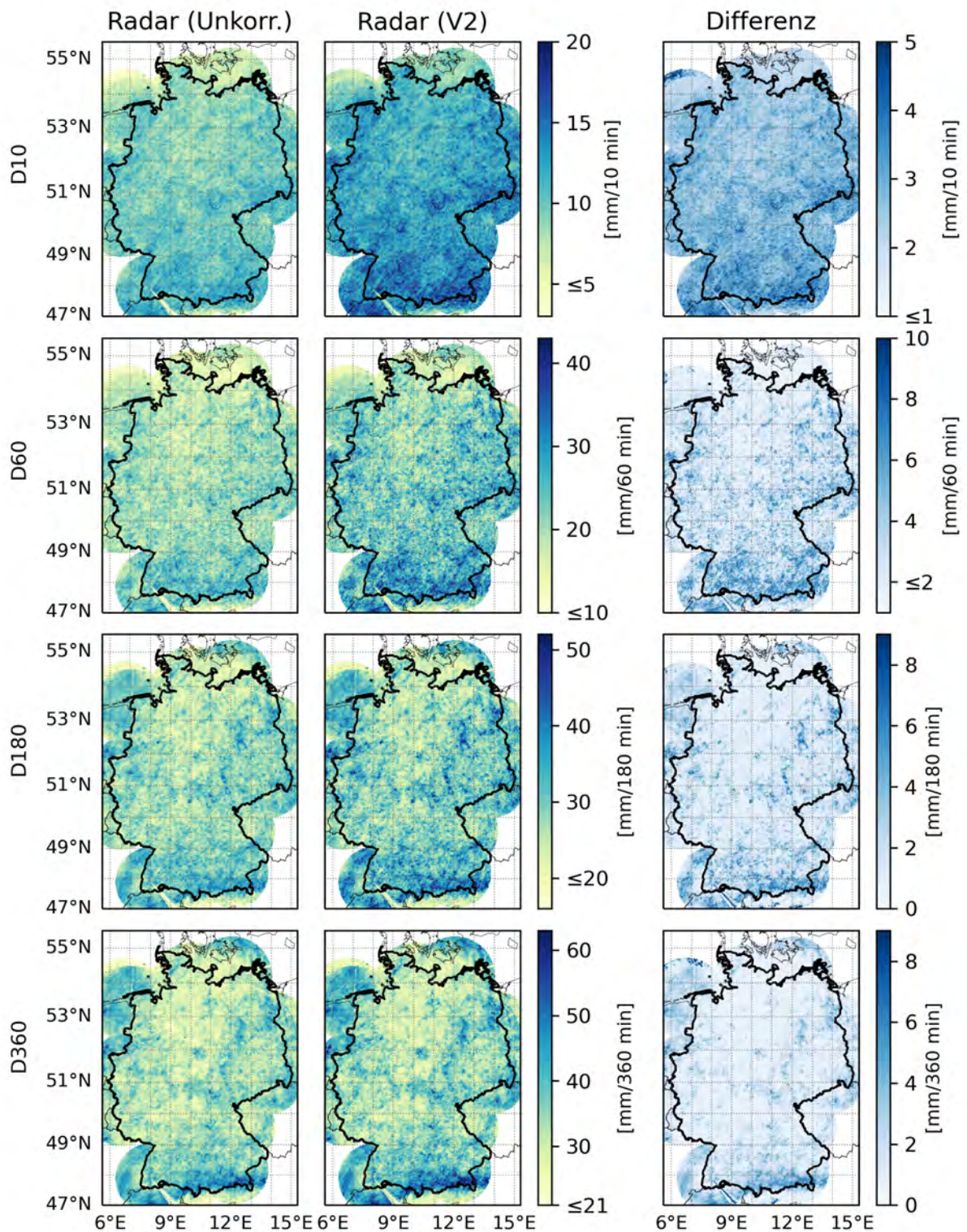


Abbildung 4.22: Anwendung der mithilfe von Verfahren 2 ermittelten Korrekturfaktoren auf alle Rasterzellen (©EuroGeographics bezüglich der Verwaltungsgrenzen)

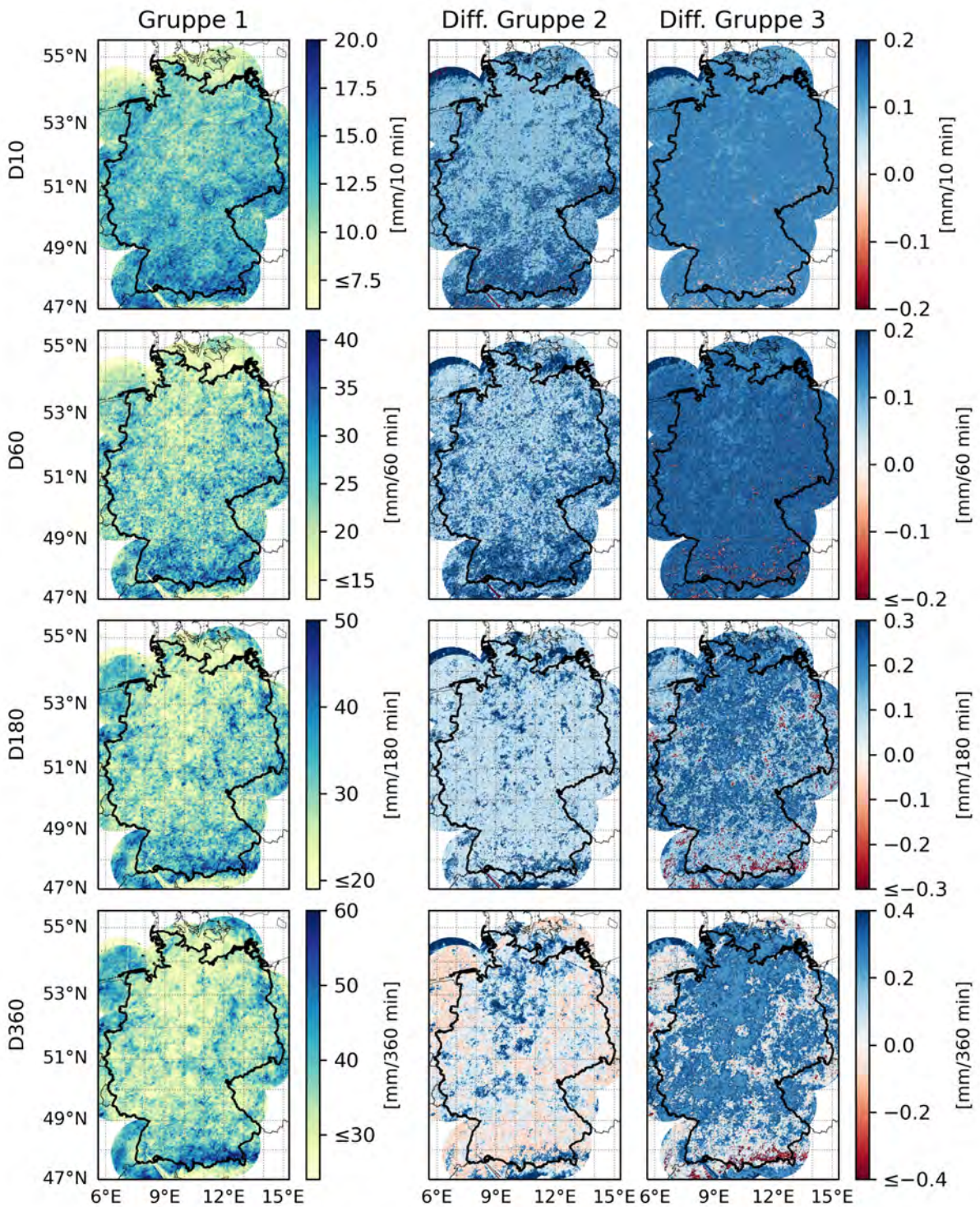


Abbildung 4.23: Verfahren 2: Vergleich der korrigierten Radar-Niederschlagskarten der drei Gruppen (©EuroGeographics bezüglich der Verwaltungsgrenzen)

4.5 Biaskorrektur - Direkter Vergleich

Da die Karten der mittels Verfahren 2 korrigierten Radarniederschläge denen von Verfahren 1 sehr ähnlich sind, wurde folgend ein direkter Vergleich der beiden durchgeführt. Zum einen wurde ein Vergleich des verbleibenden Bias nach der Korrektur an den Stationsstandorten durchgeführt, zum anderen die Differenz beider Verfahren für die gesamte Fläche Deutschland verglichen.

Die Untersuchung des verbleibenden Bias ergab, dass das Bias nach der Korrektur insgesamt geringer ist, was an dem geringeren Anteil an Rot-Tönen in der Karte zu erkennen ist. Dies zeigt sich besonders für D10 (siehe Abbildung 4.24). Allerdings kommen auch einige Stationen vor, die nun nach der Korrektur ein positives Bias haben, was auf eine „Überkorrektur“ der Niederschlagswerte hindeutet. Zwischen der Trainings- und Kontrollgruppe sind kaum Unterschiede feststellbar (siehe Abbildungen 4.24 und 4.25). Bei den größeren Event-Rängen ist bei der Trainingsgruppe im Vergleich zur Kontrollgruppe ein homogeneres Muster erkennbar, einzelne Stationen mit großem positiven oder negativem Bias sind weniger aufzufinden (siehe Abbildungen 7.15 und 7.16). Beim direkten Vergleich der Karten von Verfahren 1 und Verfahren 2 fällt die Ähnlichkeit der beiden auf, insgesamt sind nur minimale Unterschiede zu sehen. Werden die korrigierten Radarkarten für die gesamte deutsche Fläche verglichen, lassen sich mehr Unterschiede feststellen.

Für die ersten Event-Ränge werden die Regionen mit geringer Niederschlagsintensität bei Verfahren 2 stärker korrigiert als bei Verfahren 1 (siehe Abbildung 4.26). Bereiche mit großer Intensität werden jedoch gleich stark oder geringer korrigiert als bei Verfahren 1. Diese Flächen mit einer geringeren Intensität stellen jedoch einen nicht zu unterschätzenden Anteil der Gesamtfläche dar, und sind insbesondere für Norddeutschland von Bedeutung. Dieser Trend verändert sich leicht mit größer werdendem Event-Rang. Für die großen Ränge korrigieren beiden Verfahren die kleinen Niederschlagsintensitäten ähnlich stark. Im Bereich der mittleren bis großen Intensitäten ist die stärkere Korrektur bei Verfahren 2 besser zu sehen (siehe Abbildung 4.27). Die Bereiche mit den größten Niederschlagsintensitäten werden hier wiederum stärker von Verfahren 1 korrigiert. Dies stimmt weitestgehend mit dem Verlauf der Korrekturfaktoren überein. Da bei Verfahren 2 gerade für die größeren Dauerstufen die Korrekturfaktoren im Bereich der größeren Quantile schon bei einem kleineren Quantil-Wert ansteigen als bei Verfahren 1, wird derselbe Anteil der Fläche Deutschland stärker korrigiert. Dabei ist jedoch anzumerken, dass dieser Vergleich nur für die erste Gruppe durchgeführt wurde. Inwieweit diese Wechselbeziehung immer derart ausgeprägt ist, bleibt gerade im Hinblick darauf, dass die Gruppenunterschiede bei Verfahren 1 stärker ins Gewicht fallen, abzuklären.

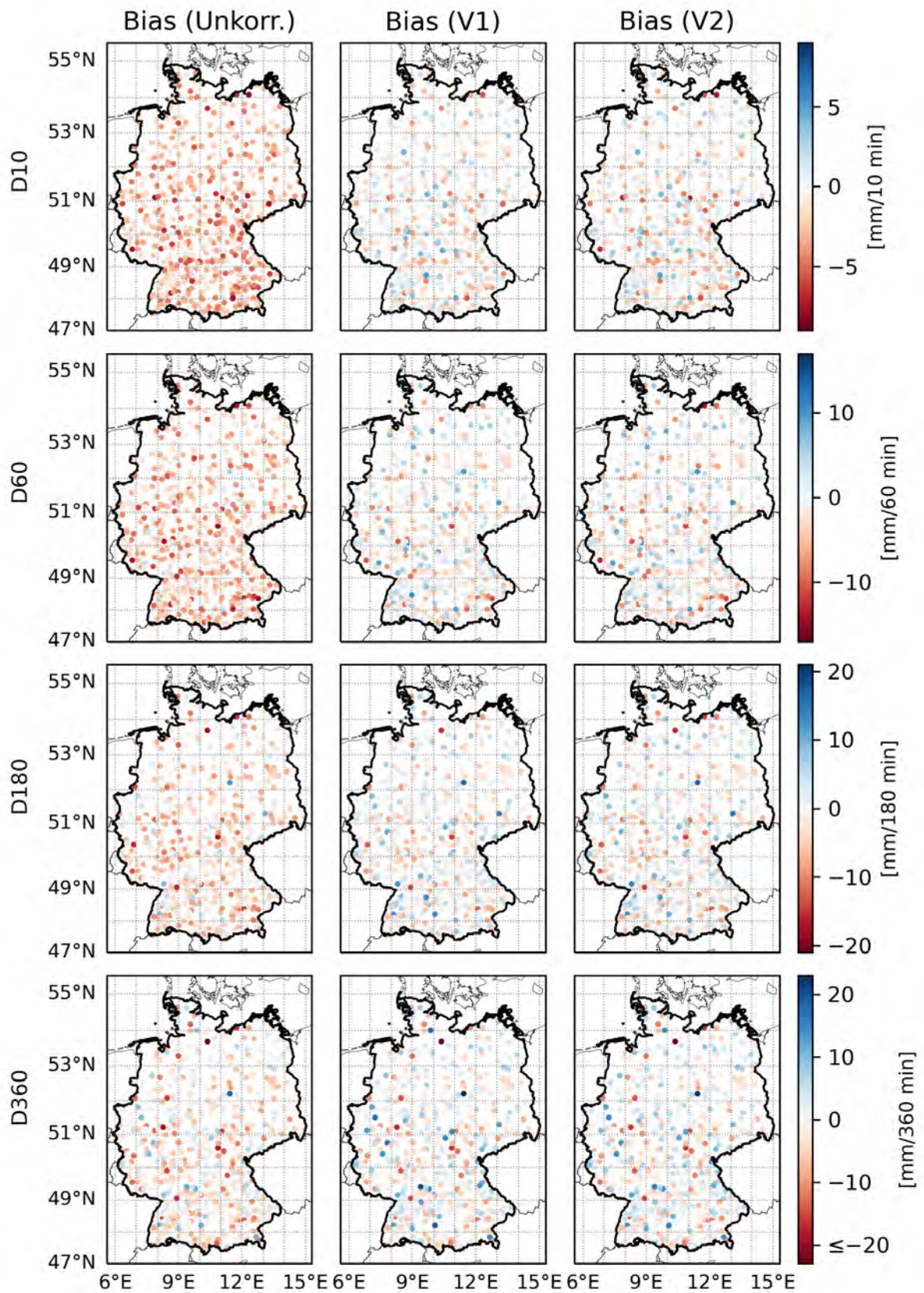


Abbildung 4.24: Vergleich des verbleibenden Bias beider Verfahren für die Kontrollgruppe bei Rang 10 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

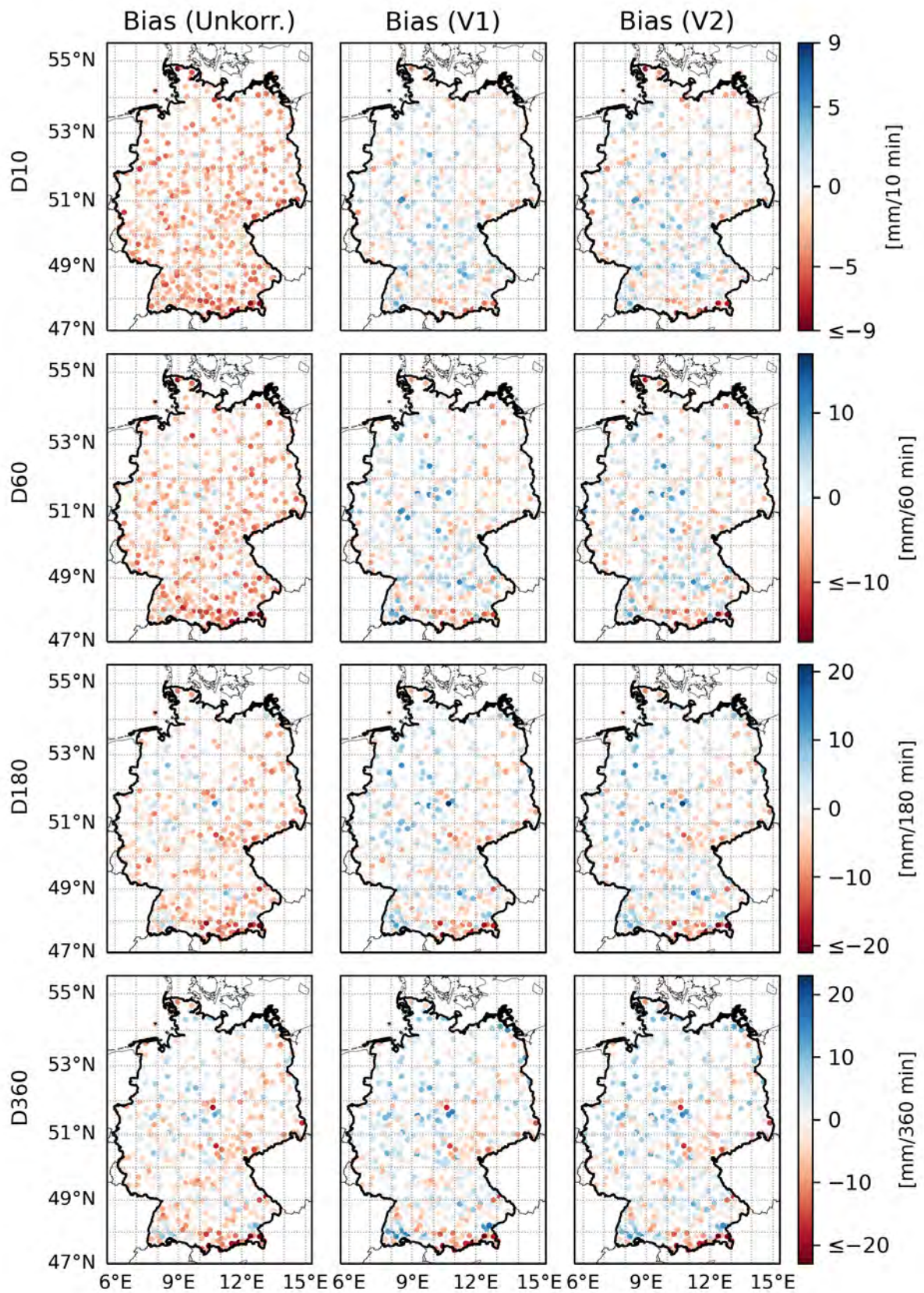


Abbildung 4.25: Vergleich des verbleibenden Bias beider Verfahren für die Trainingsgruppe bei Rang 10 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

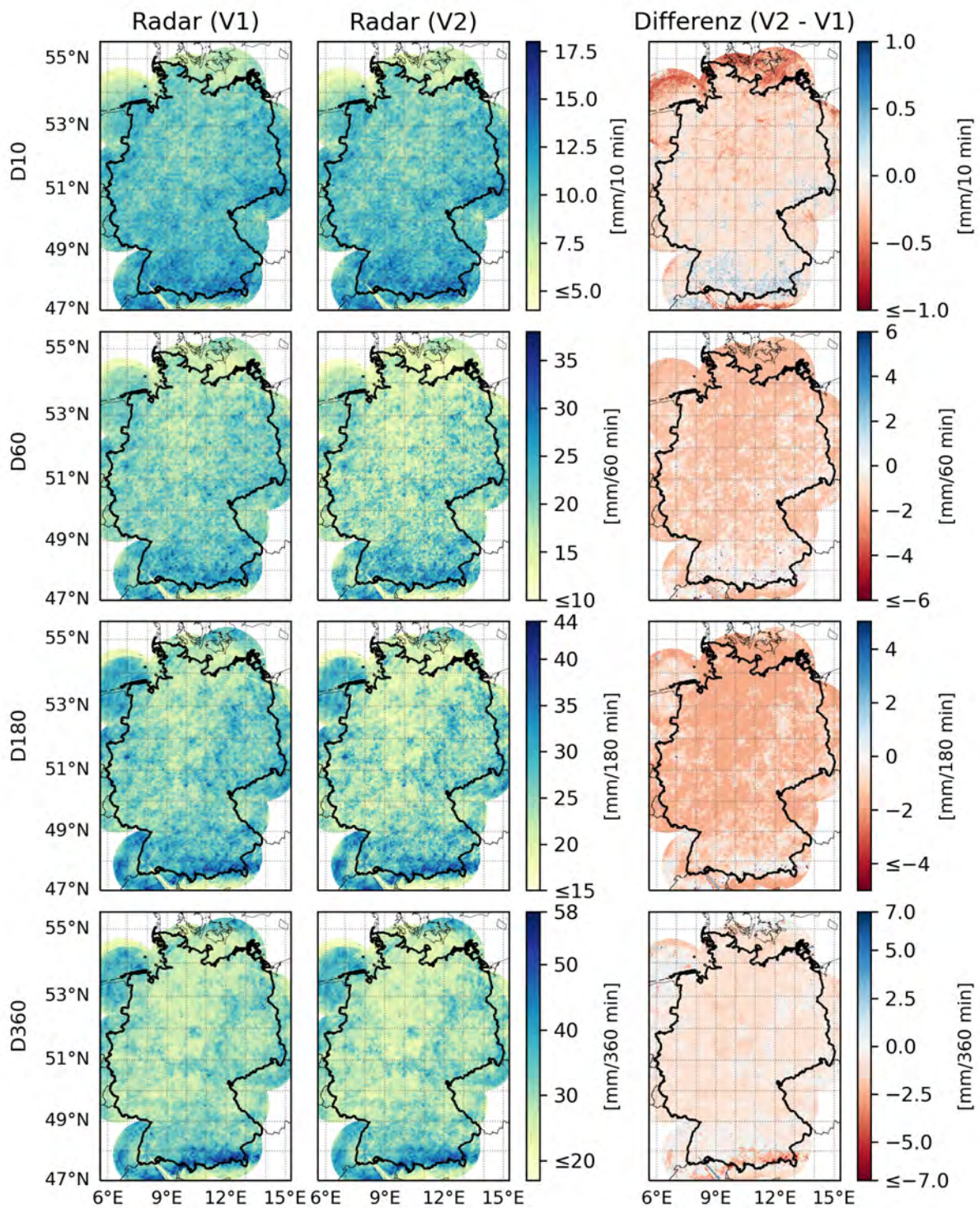


Abbildung 4.26: Vergleich der korrigierten Radarkarten beider Verfahren bei Rang 10
 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

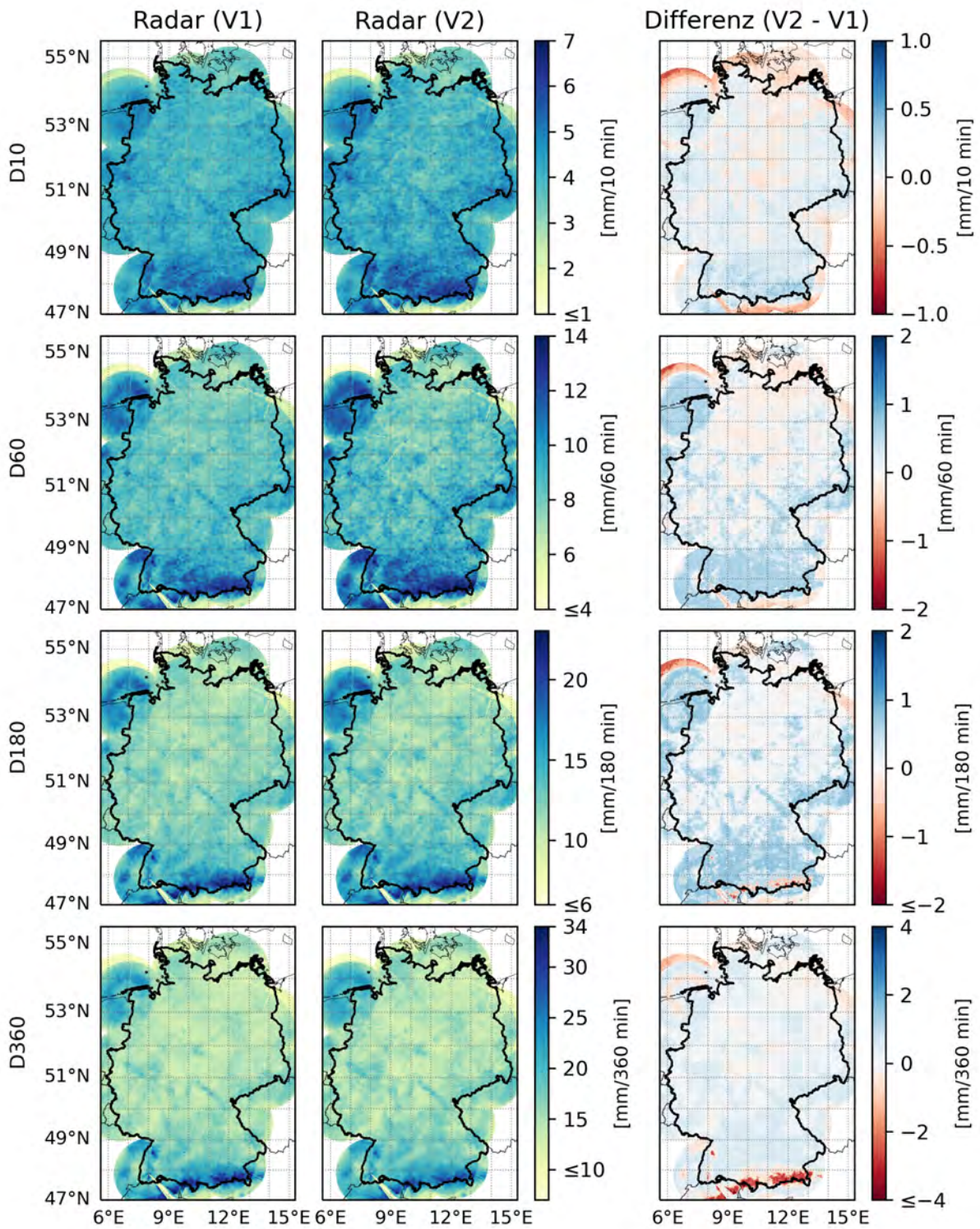


Abbildung 4.27: Vergleich der korrigierten Radarkarten beider Verfahren bei Rang 75
 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

4.6 Biaskorrektur - Sensitivitätstests

Folgend wurde untersucht, inwiefern bestimmte Annahmen und Parameterwerte einen Einfluss auf das Korrekturergebnis haben. Dazu gehörte die Anzahl an Quantilklassen (Sensitivitätstest 1), die Verwendung von Mean4 statt Stat-Px (Sensitivitätstest 2), die Verwendung von ausschließlich Sommererevents (Sensitivitätstest 3) und die Verwendung von ausschließlich im Datum übereinstimmenden Events (Sensitivitätstest 4). Der Einfachheit halber und weil nicht für jeden Test genügend Werte für eine rangbasierte Betrachtung vorhanden waren, wurden diese Tests im Vergleich zu Verfahren 2 durchgeführt.

Die Q-Q-Plots des Sensitivitätstest 1 zeigen dabei insgesamt ein ähnliches Aussehen wie bei Verfahren 2. Für D10 und D180 sind aber bei den ersten Event-Rängen mehr Unterschiede feststellbar als bei Verfahren 2, insbesondere für die große Niederschlagsintensitäten (siehe Abbildung 4.28).

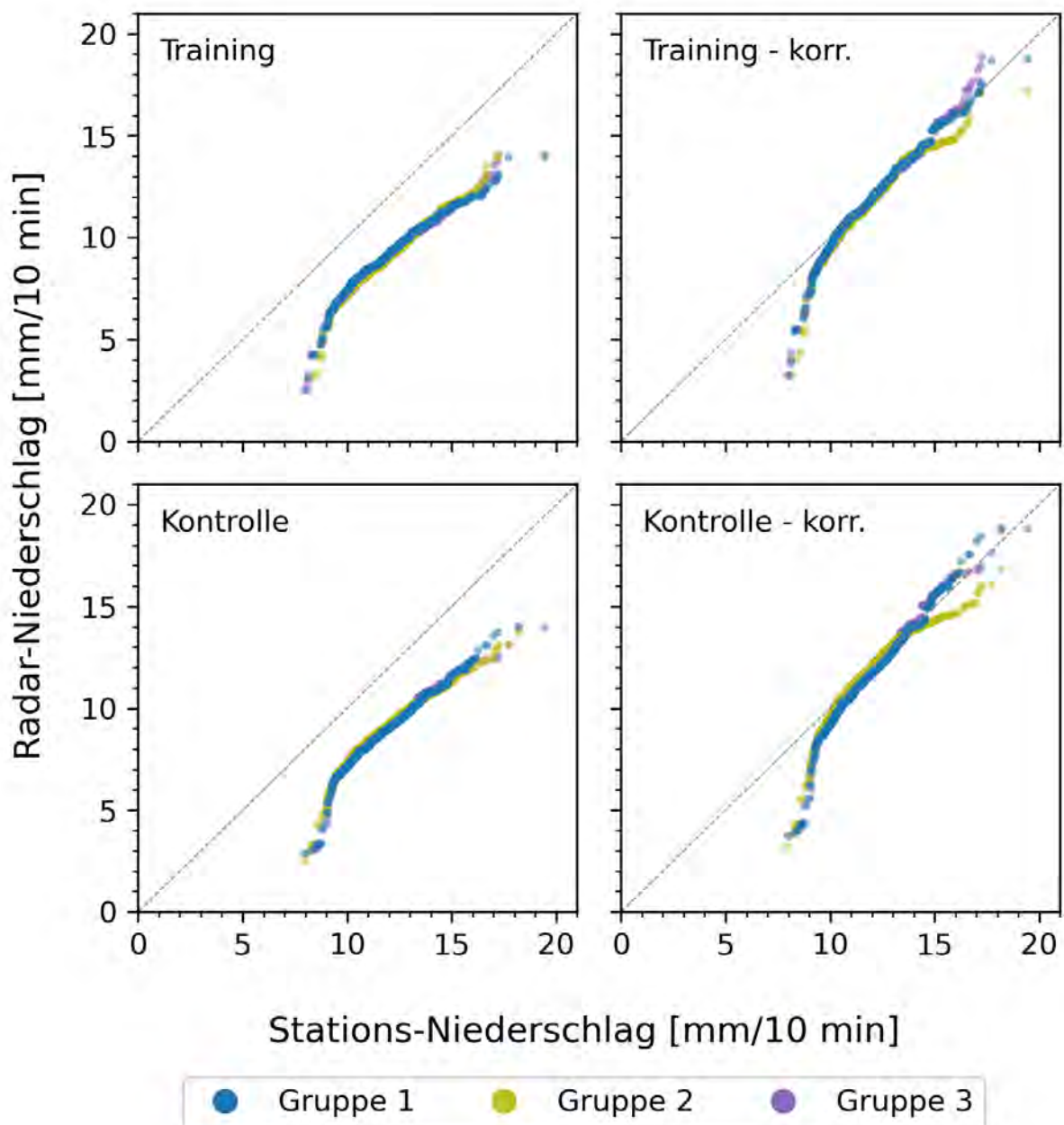


Abbildung 4.28: Sensitivitätstest 2: Q-Q-Plot für den siebten Event-Rang (D10)

Bedingt durch die geringeren Korrekturfaktoren für die ersten Quantile, werden diese weniger

stark korrigiert, was besonders bei den späteren Event-Rängen (aufgrund der kleineren Spannweite der Werte) erkennbar ist. Der Fit ist somit im Bereich kleiner Quantile schlechter, ansonsten dem von Verfahren 2 recht ähnlich. Dies zeigt sich auch bei den ECDFs (siehe Abbildung 7.17).

Im Vergleich zu Verfahren 2 fällt beim Sensitivitätstest 2 auf, dass die unkorrigierten Radardaten der Trainings- und Kontrollgruppe die Stationsdaten im Bereich der großen Niederschlags-Quantile stärker unterschätzen, während die Unterschätzung für die kleinsten Quantile etwas geringer ausfällt (siehe Abbildung 4.29).

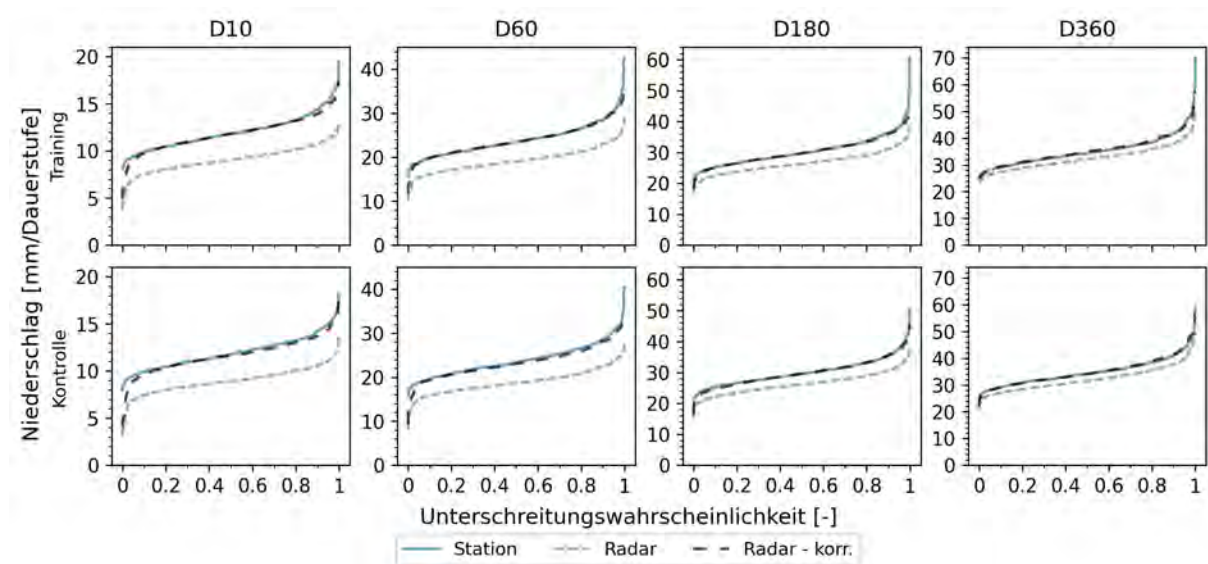


Abbildung 4.29: Sensitivitätstest 2: ECDFs für den siebten Event-Rang

Gleichzeitig zeigt sich bei den korrigierten Daten für die großen Quantile seltener eine zu starke Korrektur als bei Verfahren 2, für manche Event-Ränge werden diese Quantile auch nach der Korrektur noch unterschätzt. Die Korrektur der Trainings- und Kontrollgruppe ist dabei ähnlich gut. Bei Betrachtung der ECDFs fällt der bessere Fit bei den kleinsten Quantilen für die ersten Event-Ränge auf, während für die letzten Ränge ein etwas schlechterer Fit festzustellen ist. Sowohl in den QQ-Plots als auch in den ECDFs sind bei den Dauerstufen D180 und D360 für die ersten Event-Ränge Sprünge zu sehen, die bei Verfahren 2 nicht vorkommen (siehe Abbildung 4.30). Beim Vergleich der drei Trainings- und Kontrollgruppen fallen diese wiederum auf, was verdeutlicht, dass diese nicht für jede Gruppe vorkommen, ansonsten sind die Gruppen sehr ähnlich.

Die Untersuchung des Datums der Niederschlagsereignisse ergab, dass der überwiegende Anteil im Sommer stattfand. Mit zunehmenden Event-Rang und damit abnehmender Niederschlagsintensität nimmt der Anteil der Winterereignisse für alle Dauerstufen zu. Während für die ersten Ränge der Anteil der Winterereignisse bei D10 nur wenige Prozent beträgt, stellt dieser Anteil für die letzten Ränge bereits ein Fünftel aller Niederschlagswerte dar.

Abbildung 4.32 zeigt für jeden Event-Rang der Stationsdaten (links) und der Radardaten (Stat-Px) (rechts) den Anteil der Niederschlagsereignisse, die im Sommer stattfanden. Dieser Trend zeigt sich bei allen Dauerstufen, jedoch in unterschiedlichem Ausmaß. Grundsätzlich nimmt mit zunehmender Dauerstufe auch der Anteil der Winterereignisse zu, bei D360 beträgt der Anteil dieser für die letzten Ränge bereits mehr als zwei Drittel aller Werte. Zudem ist auffallend, dass der Anteil der Winterereignisse bei D180 und D360 schneller ansteigt und keiner linearen Zunahme mehr entspricht.

Zwischen den Stationsdaten und Radardaten lassen sich nur geringfügige Unterschiede feststellen. Während die Verteilung für D10 und D60 sehr ähnlich ist, weisen die Verteilungen der Radardaten für D180 und D360 einen etwas geringeren Anteil an Winterereignissen auf. Bei der Gegen-

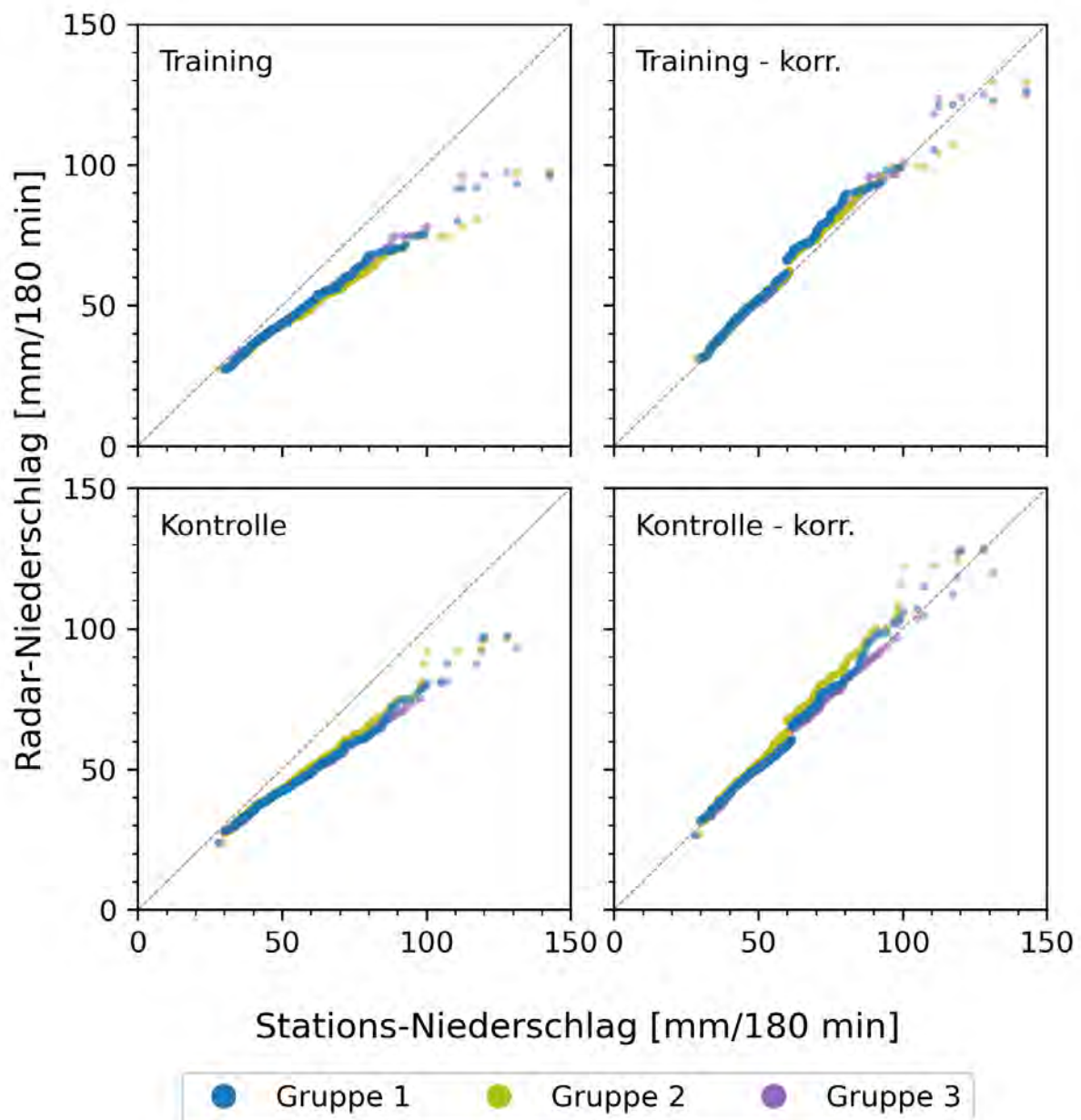


Abbildung 4.30: Sensitivitätstest 2: Q-Q-Plot für den ersten Event-Rang (D180). Gut erkennbar sind die Sprünge bei den korrigierten Daten.

überstellung der Radardaten des Stations-Pixels mit den umliegenden acht Radar-Rasterzellen konnten keine nennenswerten Unterschiede in den Verteilungen festgestellt werden.

Auch bei diesem Test ergab sich bei der Analyse der Q-Q-Plots und ECDFs ein sehr ähnliches Bild wie bei Verfahren 2. Bei den späteren Rängen sind weniger Datenpunkte zu sehen, wie es durch die Verteilung der Sommer-Ereignisse auch zu erwarten war. Dies zeigt sich vor allem bei den größeren Dauerstufen (D180 und D360).

Für die großen Quantile ist teils kein so einheitliches Muster zwischen den drei Gruppen erkennbar wie bei Verfahren 2, was besonders für die späteren Event-Ränge zutrifft.

Für die späteren Event-Ränge zeigt sich teils ein besserer, teils ein schlechterer Fit sowohl im Bereich der kleineren als auch bei den größten Quantilen (siehe Abbildung 4.32).

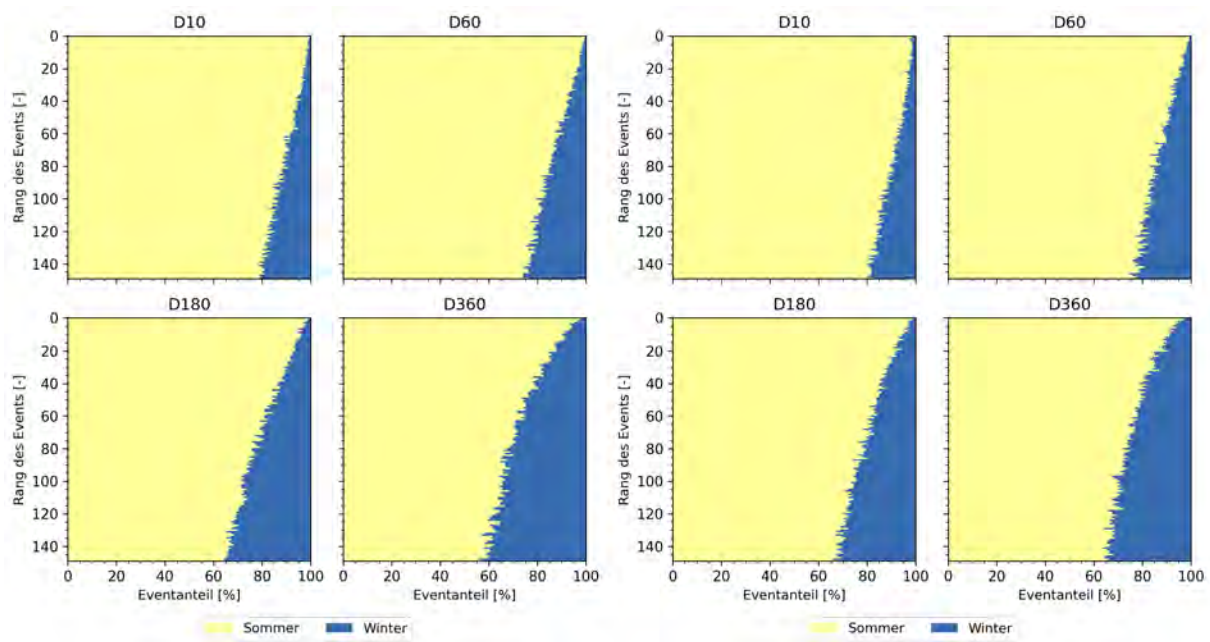


Abbildung 4.31: Anteil der Sommer- und Winterevents für die Stations- und Radardaten (Stat-Px)

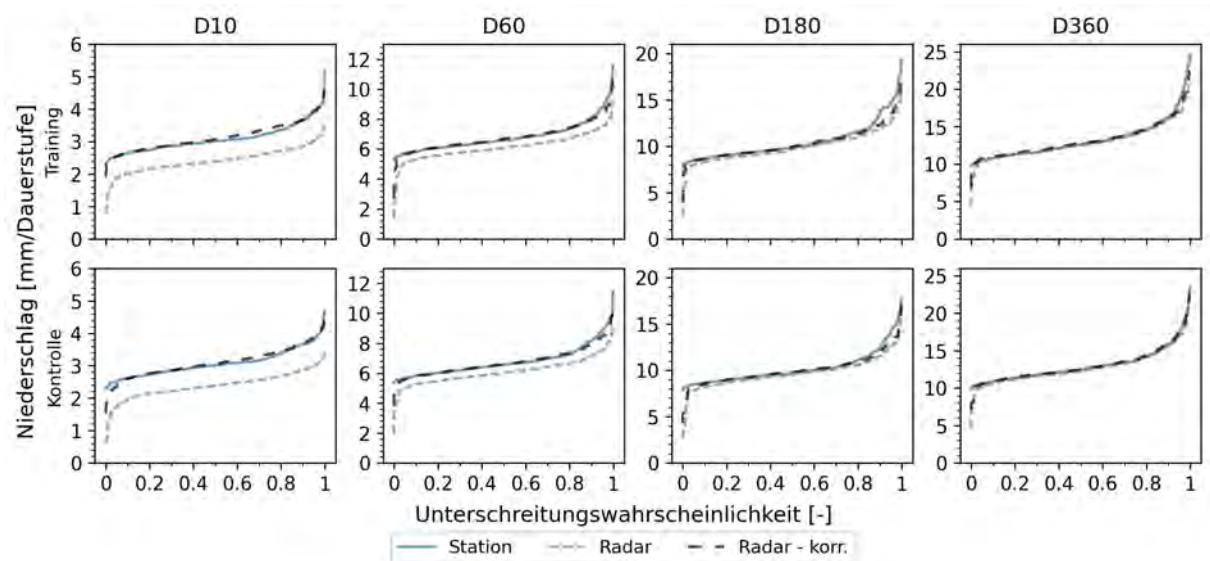


Abbildung 4.32: Sensitivitätstest 3: ECDFs für den 129. Event-Rang

Tabelle 4.2: Anzahl der im Datum übereinstimmenden Werte pro Dauerstufe

Biastyp	D10	D60	D180	D360	Alle Werte
Stat-Px	9512	9024	8296	8101	173100
Mean4	37952	36131	33525	32232	173100

Die Untersuchung der im Datum übereinstimmenden Ereignisse führte zu den größten Unterschieden im Vergleich zu Verfahren 2. Die Anzahl der Werte die für das Verfahren zur Verfügung standen ist verschwindend gering (siehe Tabelle 4.2).

Bei Betrachtung der Q-Q-Plots aber auch der ECDFs ist deutlich zu sehen, dass weniger Datenpunkte vorhanden sind als beim zweiten Verfahren, besonders im Bereich der kleinen Quantile,

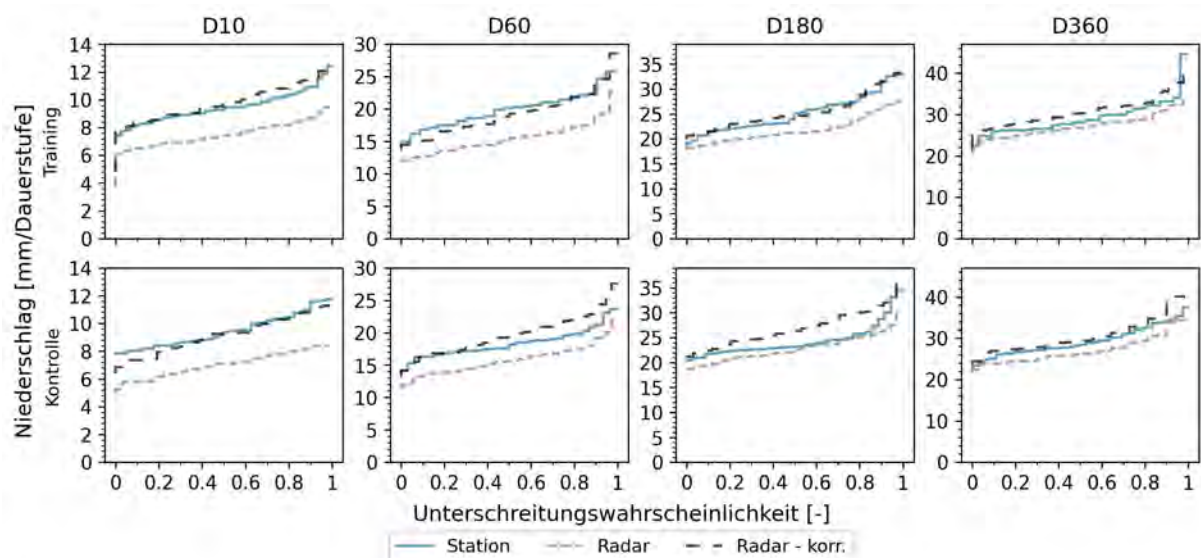


Abbildung 4.33: Sensitivitätstest 4 ECDFs für den 15. Event-Rang

teils auch bei den großen Quantilen (siehe Abbildung 7.18). Je nach Event-Rang ist eine unzureichende Korrektur, eine zu starke Korrektur oder eine mehr oder weniger passende Korrektur zu sehen. Werden die ECDFs betrachtet, kann eine nur sehr mangelhafte Korrektur festgestellt werden (siehe Abbildung 4.33). Die Streuung zwischen den drei Gruppen ist, besonders für die größten Quantile, größer als bei Verfahren 2. Der räumliche Vergleich der vier Tests offenbarte keine sichtbaren Unterschiede zwischen den durchgeführten Sensitivitätsanalysen (siehe Abbildung 7.19). Abweichungen zeigten sich auch nicht für die anderen Event-Ränge und Dauerstufen.

Da der entscheidende Unterschied die berechneten Korrekturfaktoren darstellten, wurden diese folgend miteinander verglichen (siehe Abbildung 4.34).

Bei der Verwendung von 100 statt 1000 Bins (Quantil-Klassen) zur Berechnung der Quantile, fallen die kleineren Werte für die kleinsten Quantile auf. Für die kleinsten Quantile betragen die Korrekturfaktoren in etwa die Hälfte der entsprechenden Faktoren von Verfahren 2. Wird nicht das Stations-Pixel allein als Grundlage des Radarniederschlags verwendet, sondern das Mittel der nächsten 4 Rasterzellen (Mean4), sind wiederum geringere Korrekturfaktoren für die kleinsten Quantile sichtbar. Im Vergleich zum Sensitivitätstest 1 sind die Unterschiede zu Verfahren 2 jedoch geringer. Die Korrekturfaktoren dieses Sensitivitätstest mit ausschließlich Sommer-Ereignissen ähneln grundsätzlich denen des zweiten Verfahrens. Im Bereich der kleinsten Quantile sind kleinere Korrekturfaktoren aufzufinden, die Differenz dieser zu denen des zweiten Verfahrens ist jedoch geringer als bei den Sensitivitätstest 1 und 2. Die ausschließliche Verwendung von im Datum übereinstimmenden Niederschlagsereignissen resultiert deutlich in Unterschieden zu Verfahren 2. Die Korrekturfaktoren der kleinsten Quantile sind wiederum niedriger als im Verfahren 2.

Da bei allen Verfahren und Sensitivitätstests die ersten Quantile die größten Korrekturfaktoren aufwiesen, wurde dies weiter untersucht. Dazu wurden die Differenzen zwischen den ECDFs betrachtet, um eine absolute Abweichung feststellen zu können. Beim Vergleich der größten Quantile eines Ranges mit den kleinsten Quantilen desselben Ranges, ist festzustellen, dass die Differenz der Stations- und Radardaten für diese größten Quantile bei den ersten Event-Rängen recht groß, für die mittleren Ränge relativ klein ist und für die letzten Ränge wieder ansteigt (siehe Abbildung 4.35). Dabei sind die Differenzen für die kleinsten Quantile bei jedem Rang, jeder Trainings- und Kontrollgruppe entsprechend hoch. Die Differenzen für die größten Quantile hingegen sind sehr von der jeweiligen Gruppe abhängig. Dabei sind die entsprechenden Trainings- und Kontrollgruppen zueinander gegenläufig. Ist bei der jeweiligen Trainingsgruppe ein großer Anstieg bei den letzten Quantilen zu erkennen, fällt der Anstieg für ebendiese Quantile bei der

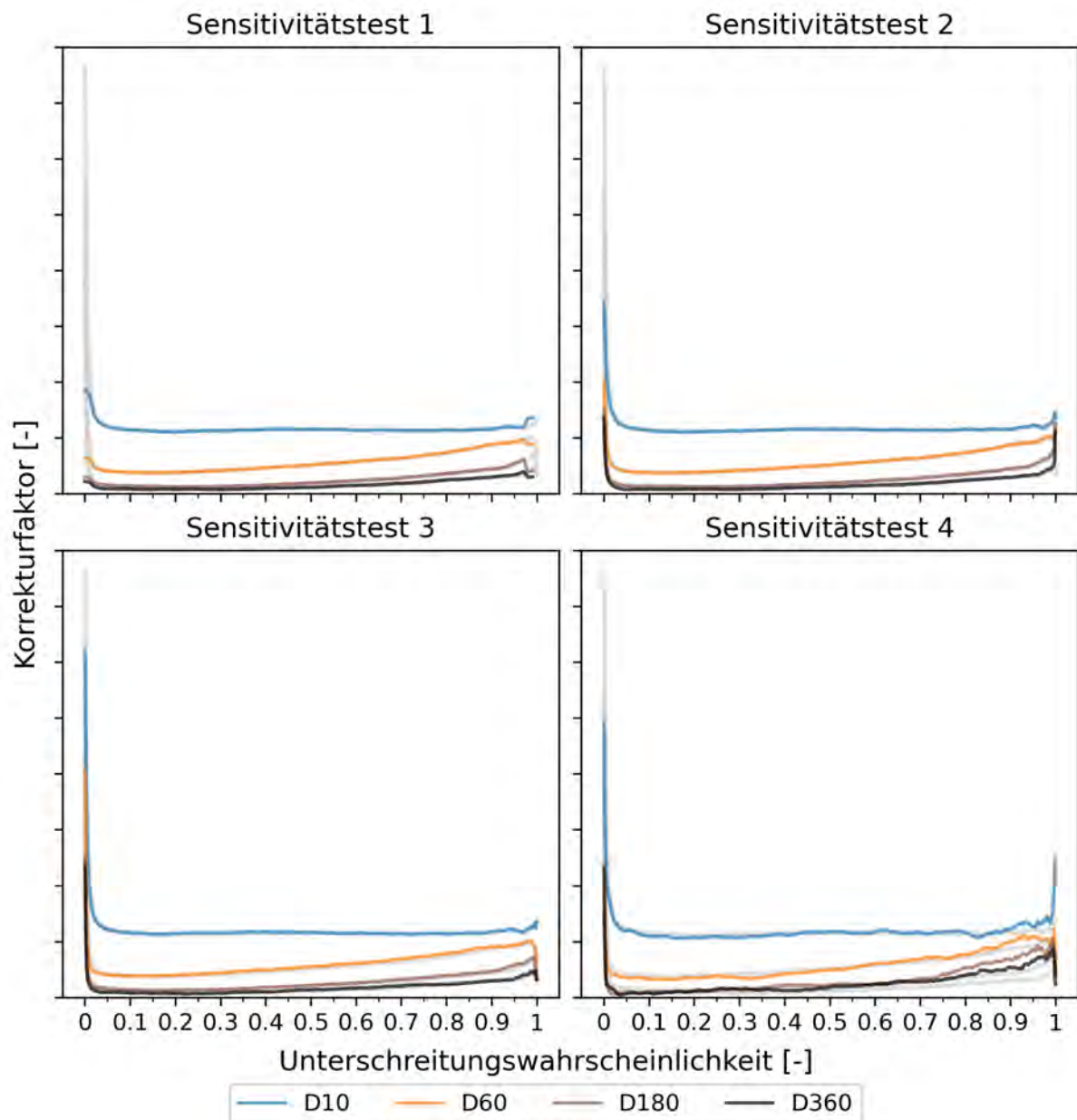


Abbildung 4.34: Vergleich der Korrekturfaktoren aller Sensitivitätstests. Hellgrau hinterlegt ist der Verlauf für Verfahren 2.

Kontrollgruppe schwächer aus, ist der Anstieg bei der Kontrollgruppe stärker, fällt er bei der entsprechenden Trainingsgruppe kleiner aus. Grundsätzlich zeigt sich dies bei allen Dauerstufen. Während für die ersten Event-Ränge die Differenzen der größten Quantile im Verhältnis größer sind als die der größten Quantile, kehrt sich dies meist ab den mittleren Event-Rängen um und bei D10 und D360 invertiert sich dies bei den letzten Rängen wiederum, sodass die größten Quantile wieder die größten Differenzen aufweisen.

Folgend wurden die Histogramme der Stations- und Radardaten eines Ranges und einer Gruppe betrachtet. Hier fällt auf, dass die Häufigkeitsverteilung der Stationsdaten mehr Werte im rechten Verteilungsrand aufweist und damit rechtsschief ist, die Verteilung der Radardaten hingegen über mehr Werte im linken Verteilungsrand verfügt und damit linksschief ist (siehe Abbildung 4.36). Dies zeigt sich mehr oder weniger für alle Event-Ränge, ist bei den mittleren Rängen jedoch nicht so stark ausgeprägt wie für die ersten und letzten. Auch die Betrachtung der kumulierten Differenzen des aktuellen Wertes der ECDF zum vorigen Wert zeigen diesen Trend, bei den

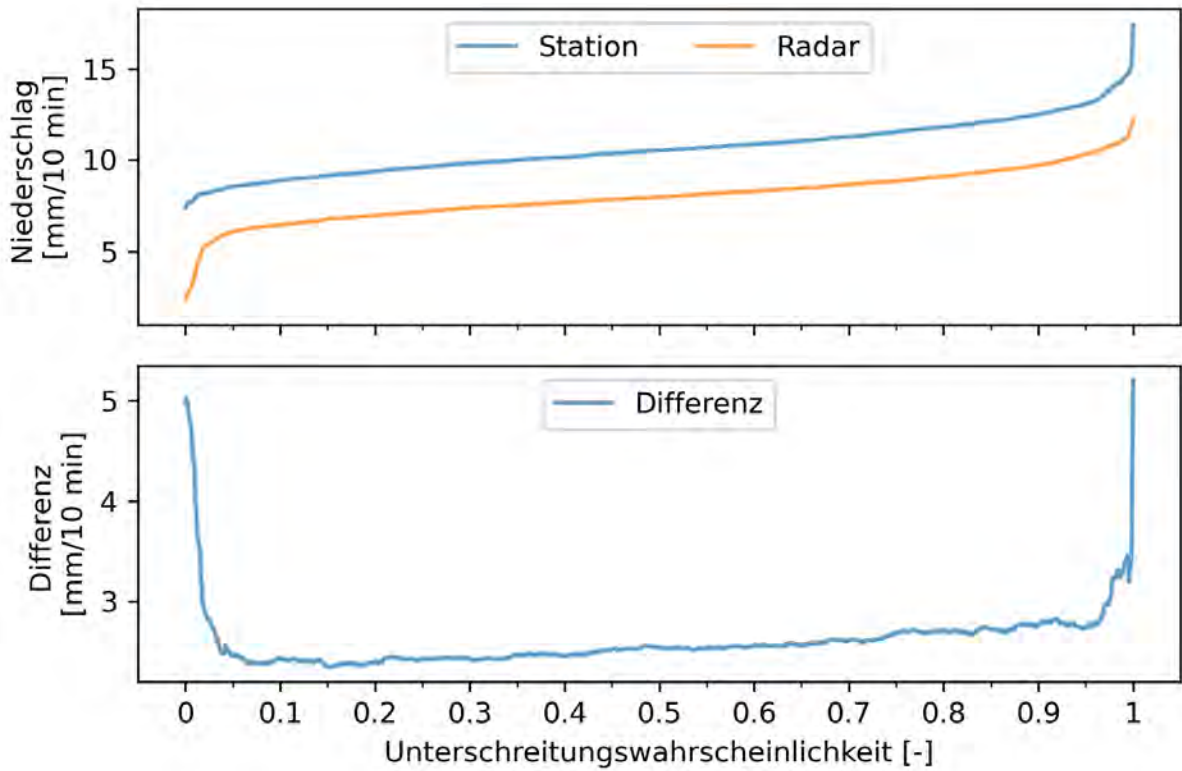


Abbildung 4.35: Differenzen der ECDFs für Rang 10 (D10)

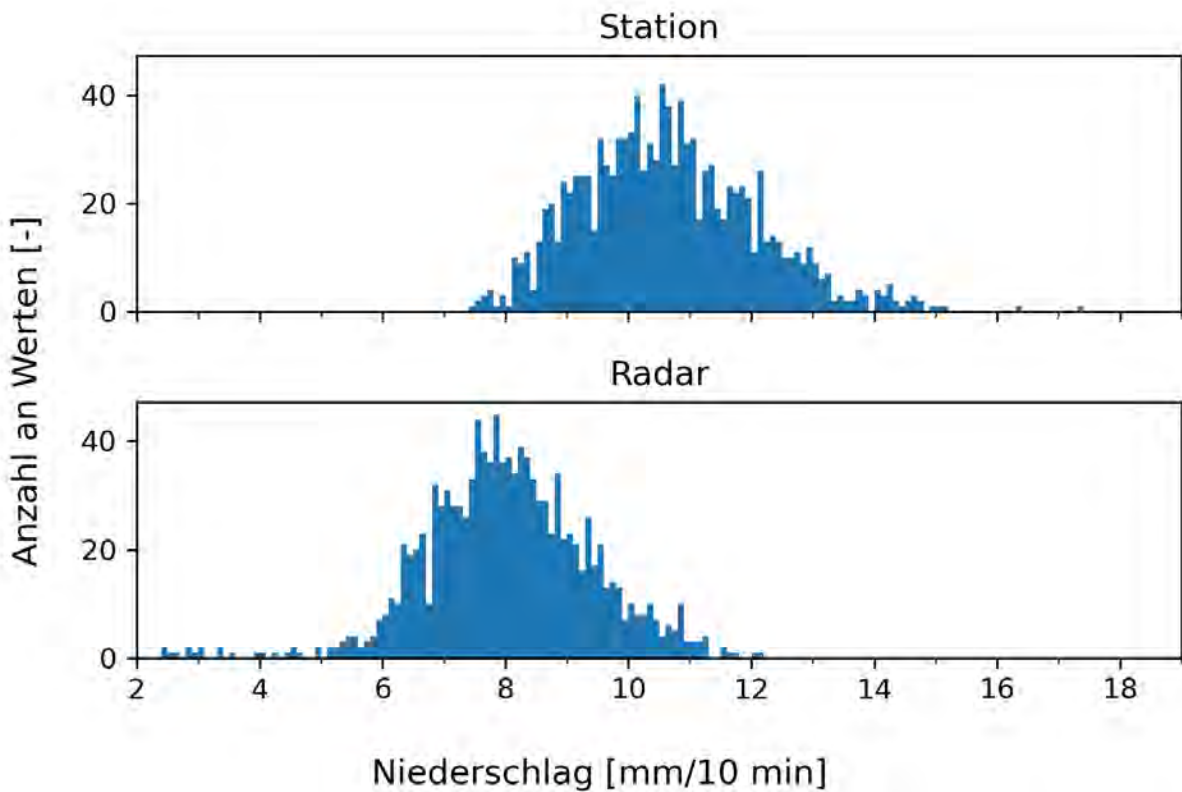


Abbildung 4.36: Histogramme der Stations- und Radardaten für Rang 10 (D10)

Stationsdaten ist der steilste Anstieg am rechten Rand der Verteilung (große Niederschlagswerte) zu sehen, bei den Radardaten dagegen am linken Rand der Verteilung (kleine Niederschlagswerte)

(siehe Abbildung 7.20).

Schließlich wurde die empirische Varianz und Schiefe der beiden Datensätze ermittelt. Diese Berechnung fand zum einen für jede Trainings- und Kontrollgruppen statt, zum anderen aber auch für alle Stations- und Radardaten unabhängig der Gruppe. Dabei zeigt sich, dass die empirische Varianz der Radardaten für alle Ränge, alle Gruppen und alle Dauerstufen kleiner ist als die der Stationsdaten. Für die Schiefe ergab sich kein so eindeutiges Resultat. Während die Stationsdaten immer eine positive Schiefe aufweisen und damit einer rechtsschiefen Verteilung folgen, ändert sich die Schiefe der Radardaten mit der Dauerstufe. Für D10 und D60 ist der überwiegende Anteil der Event-Ränge linksschief, wohingegen bei D180 nur noch die Kontrollgruppen der Radardaten linksschief sind und bei D360 kaum mehr Werte mit einer negativen Schiefe zu finden sind. Insgesamt sind die Stationsdaten jedoch generell rechtsschiefer als die Radardaten.

4.7 Anwendungsbeispiele - Korrektur realer Events

Im Großen und Ganzen fällt die Korrektur der Niederschlagssummen der realen Events nur sehr durchschnittlich aus. Bei keinem der untersuchten Beispiele konnte mithilfe der ermittelten Korrekturfaktoren alle Dauerstufen zufriedenstellend korrigiert werden. Während bei manchen Stationen die Korrektur nicht ausreichte, um zu einer guten Übereinstimmung mit den Stations-Summen zu führen (z. B. Station 970, 3376), fand bei anderen Stationen eine zu starke Korrektur statt (z. B. Station 567, 684, 5763), bei den restlichen Stationen (5029, 2319, 555, 704) führte die Korrektur zumindest für einige Dauerstufe zu einer besseren Übereinstimmung. Abbildung 4.37 zeigt die Korrektur des Ereignisses vom 17. Juni 2020 der Station 567. Für jede Dauerstufe findet eine „Überkorrektur“ der Niederschlagssumme statt. Nur bei D10 befindet sich die Niederschlagssumme der Stationsdaten noch in der Spannweite der Niederschlagssummen der neun Radarpixel. Bei D10 und D180 reduziert sich immerhin die Differenz zwischen Stations- und Radarsumme nach erfolgter Korrektur. Das Niederschlagsereignis vom 31. Mai 2018 an der Station 970 ist ein Beispiel für eine unzureichende Korrektur (siehe Abbildung 4.38). Beim Vergleich der 10-minütigen Niederschlagssummen (links) fällt die große Abweichung zwischen Radar- und Stationsdaten sofort auf. Auch die Anwendung der Korrekturfaktoren auf die maximale Niederschlagssumme führt zu keiner deutlichen Reduzierung dieses Bias.

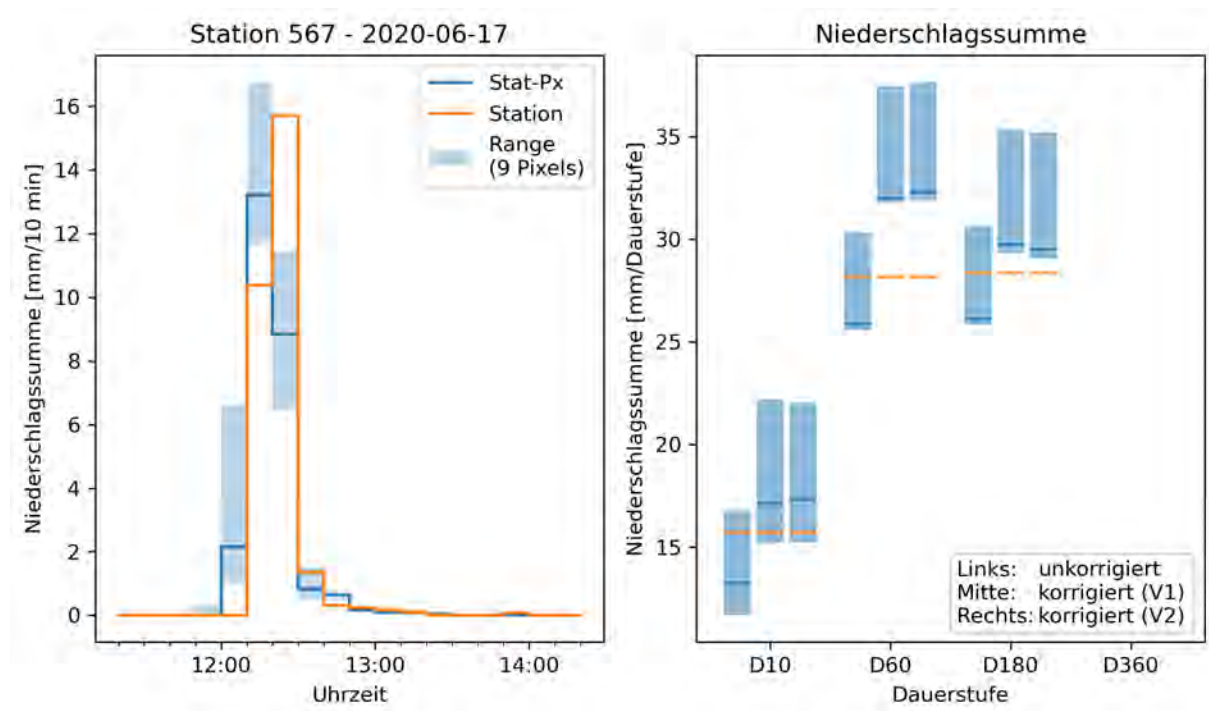


Abbildung 4.37: Korrektur des realen Ereignisses vom 17. Juni 2020 (Station 567). Zu beachten ist, dass nur die maximale Niederschlagssumme korrigiert wurde, nicht der zeitliche Verlauf.

Das bei Station 5029 gemessene Extremereignis vom 6. August 2019 ist ein Beispiel für eine zumindest teilweise erfolgreiche Korrektur (siehe Abbildung 4.39). Bei Dauerstufe D10 führt die Korrektur zu einer beinahe perfekten Übereinstimmung der Niederschlagssumme der Stations- und Radardaten. Für D60 ist diese Übereinstimmung bereits vor der Anwendung der Korrekturfaktoren gegeben, hier führen die Faktoren daher zu einer Überschätzung der Niederschlagssumme der Station.

Die hier gezeigten Abbildungen stellen jeweils das Ergebnis der Korrektur mithilfe der Faktoren aus der ersten der drei Gruppen dar. Für die zwei anderen Gruppen ergaben sich ähnliche Verhältnisse. Die restlichen Korrekturen der realen Events können dem Anhang entnommen werden (vgl. Abbildung 7.21, 7.22 und 7.23). Wenn die Niederschläge der neun Radar-Rasterzellen

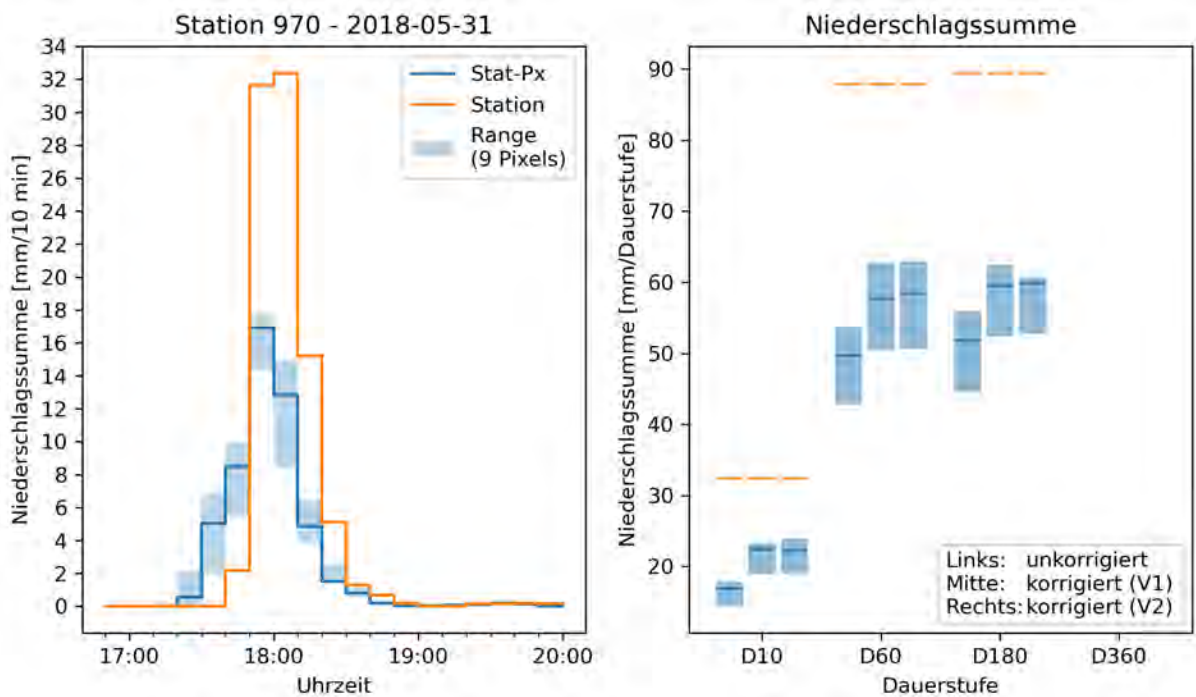


Abbildung 4.38: Korrektur des realen Ereignisses vom 31. Mai 2018 (Station 970)

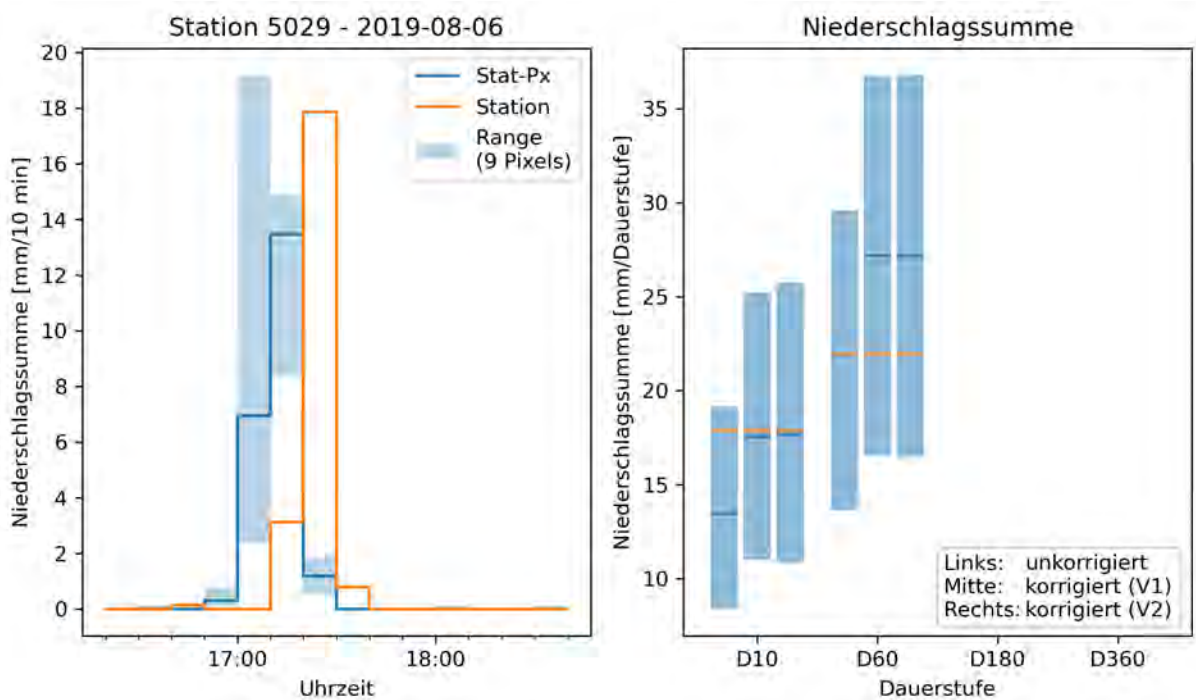


Abbildung 4.39: Korrektur des realen Ereignisses vom 6. August 2019 (Station 5029)

während des realen Events eine große Spannweite aufwiesen, zeigt sich auch bei der Korrektur der Niederschlagssumme eine (sehr) große Spannweite.

Zusammenfassend lassen sich einige Erkenntnisse festhalten:

1. Ein zeitlicher Versatz spielt bei der Betrachtung der (maximalen) Niederschlagssummen nur eine geringe Rolle.
2. Die Performance der Korrektur ist sehr von der Dauerstufe abhängig. Eine gute Performance

bei D10 bedeutet nicht zwingend eine gute Performance bei den anderen Dauerstufen.

3. Eine große Unterschätzung des Stationsniederschlags durch die Radardaten kann durch die Korrektur nicht behoben werden.
4. Im Falle einer Überschätzung des Stationsniederschlags durch die Radardaten sind die Korrekturfaktoren meist gänzlich unbrauchbar.
5. Die Unterschiede zwischen den Gruppen sind meist recht gering, sind bei den größeren Dauerstufen jedoch größer.
6. Die Ergebnisse der Korrektur mithilfe der Faktoren von Verfahren 1 und 2 sind sich grundsätzlich sehr ähnlich, es ist keine einheitliche Abweichung des einen Verfahrens vom anderen erkennbar.

4.8 Anwendungsbeispiele - Berechnung von Bemessungsniederschlägen

Als letztes Anwendungsbeispiel wurde ein 20-jährliche und 50 jährliche Ereignis für jede Raster-Zelle Deutschlands berechnet. Der Vergleich mit KOSTRA (KOSTRA-DWD-2020V) offenbart große Unterschiede zu den verwendeten Radardaten. Die unkorrigierten sowie korrigierten Radardaten unterschätzen die Bemessungsniederschläge von KOSTRA deutlich (siehe Abbildungen 4.40 und 4.41). Einige Pixel weisen jedoch auch deutlich größere bzw. kleine Niederschlagswerte auf, die Spannbreite ist also insgesamt größer. Während sich das räumliche Muster von KOSTRA für die erste Dauerstufe (D10) sehr von dem der radarbasierten Karten unterscheidet, stimmen die Muster für die größeren Dauerstufen und besonders für D360 besser überein.

Beide Verfahren erhöhen den Bemessungsniederschlag in den Gebieten, die bereits hohe Niederschlagswerten aufweisen. Beim 50-jährlichen Ereignis ist jedoch in manchen Bereichen eine Reduzierung im Vergleich zum unkorrigierten Bemessungsniederschlag zu sehen. Dies zeigt sich insbesondere bei Verfahren 2 und den Dauerstufen D180 und D360 (siehe Abbildungen 7.24 und 7.25). Dennoch führt Verfahren 1 grundsätzlich zu einer stärkeren Erhöhung der Niederschlagsintensitäten als Verfahren 2. Zwischen den drei Gruppen konnten keine Unterschiede festgestellt werden.

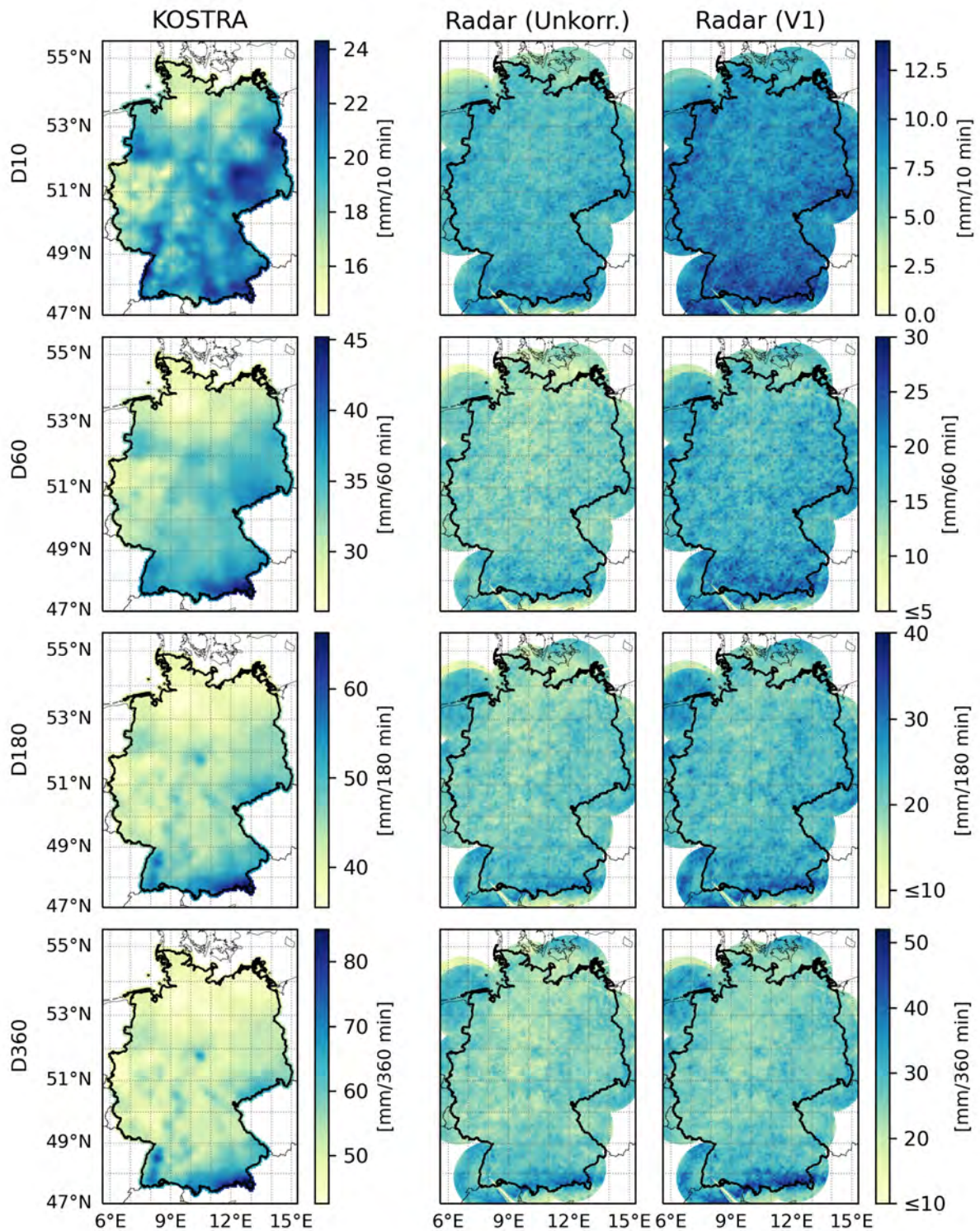


Abbildung 4.40: Vergleich des berechneten 20-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 1. Man beachte die anderen Achsenlimits bei der Karte der KOSTRA-Daten! (©EuroGeographics bezüglich der Verwaltungsgrenzen)

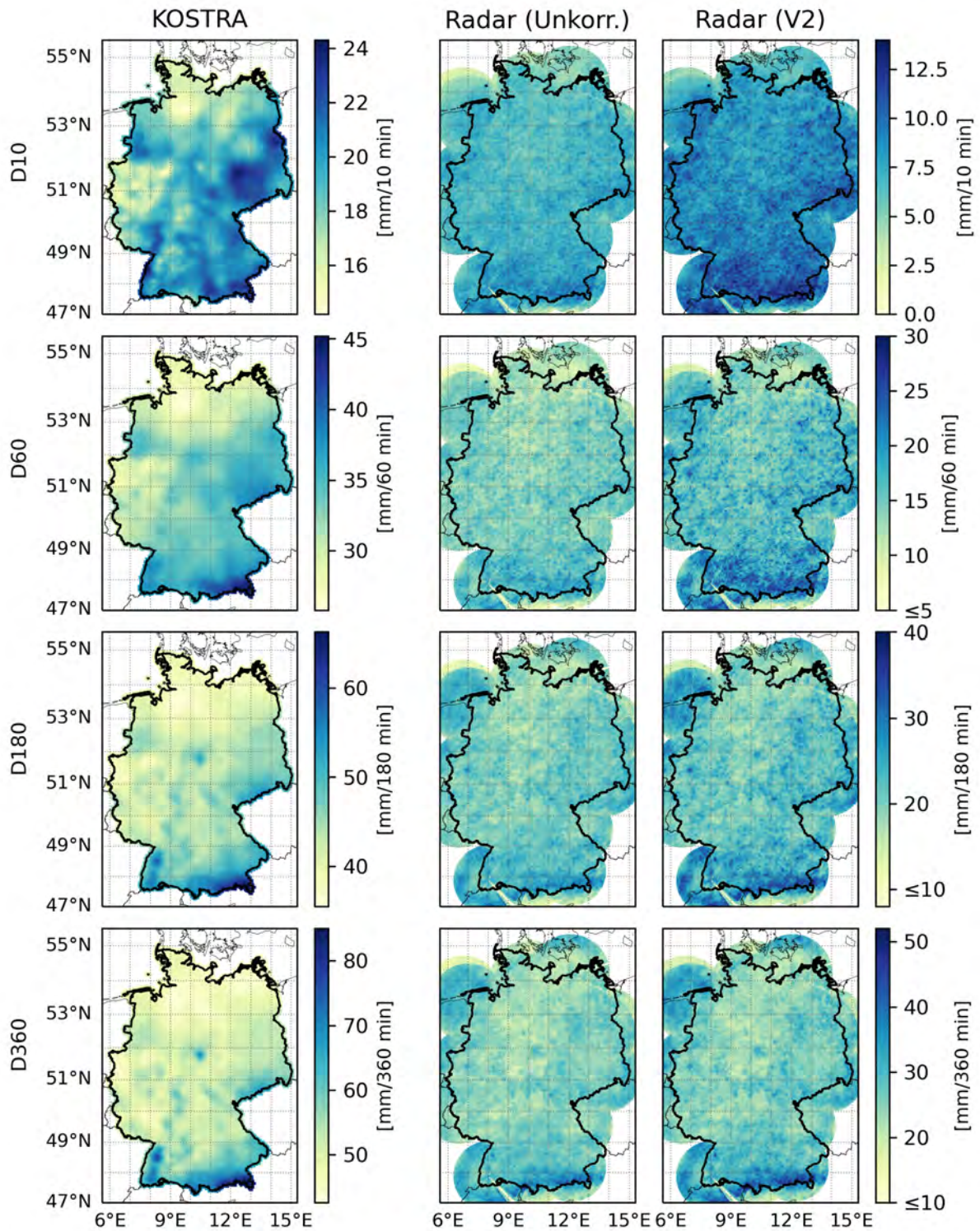


Abbildung 4.41: Vergleich des berechneten 20-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 2. Man beachte die anderen Achsenlimits bei der Karte der KOSTRA-Daten! (©EuroGeographics bezüglich der Verwaltungsgrenzen)

5 Diskussion

5.1 Klimatologie

Die Ergebnisse der klimatologischen Analysen ähneln denen anderer Studien (z. B. Winterrath et al., 2017). Der Anteil an Regentagen von etwa 50 % wurde auch in der Analyse der Niederschlagsdaten für die neue KOSTRA-Version festgestellt (Junghänel et al., 2022).

Die Untersuchung der Anzahl an Starkregenereignissen zeigte, die zufälliger räumlichere Verteilung der Events mit extremer Niederschlagsintensität (Warnstufe 3 und 4), während die Verteilung bei Warnstufe 2 weitestgehend den Mittelgebirgslagen Deutschlands entspricht. Dies deckt sich mit vergleichbaren Analysen der Literatur (Winterrath et al., 2017) und verdeutlicht, dass die extremen Starkregenereignisse meist konvektive Ereignisse darstellen, die überall auftreten können und nur sehr gering mit der Orographie korreliert sind (Junghänel et al., 2022).

Diese Analysen zeigen also die grundsätzliche Eignung der aufbereiteten und in der Arbeit verwendeten Stationsdaten an.

5.2 Bias

Beim Vergleich des räumlichen Musters von Stations- und Radarniederschlag wurde deutlich, dass diese beiden Datensätze bei einem größeren Rang räumlich besser übereinstimmen als bei einem kleinem Event-Rang. Genau wie die Niederschläge selbst, folgt auch das Bias mit zunehmendem Rang vermehrt der Topographie Deutschlands, wohingegen die ersten Ränge eine zufälliger räumliche Verteilung der extremen Niederschläge bzw. des Bias aufzeigen. Während für die letzten Event-Ränge und Dauerstufe die Regionen mit negativem Bias meist den Mittelgebirgslagen oder dem Alpenraum und Alpenvorland entsprachen, finden sich die Stationen mit positivem Bias vermehrt in den Tallagen. Bei der Analyse des Bias stellte sich grundsätzlich heraus, dass die Dauerstufe und der Event-Rang den größten Einfluss auf den absoluten sowie relativen Bias haben, der Einfluss des Biastyps (Stat-Px, Mean4, Mean9) hingegen recht gering ist.

Die Spannbreite des absoluten Bias nimmt mit zunehmender Dauerstufe zu und mit abnehmendem Event-Rang ab, was mit der Zunahme bzw. Abnahme der entsprechenden Niederschlagsintensitäten erklärt werden kann. Insgesamt zeigte sich jedoch (auch für das relative Bias), dass die Häufigkeitsverteilung des Bias für die ersten Ränge etwas spitzer wird, und das Bias somit generell homogener verteilt ist. Weiter wurde sichtbar, dass der Anteil an Stationen mit positivem Bias mit zunehmender Dauerstufe zunahm, sich mit zunehmendem Rang (und damit abnehmender Niederschlagsintensität) jedoch reduzierte.

Die Korrelationen des Bias mit der Stationshöhe, Richterklasse und dem Radarstandort ergaben unterschiedlich starke Wechselbeziehungen. Zwischen dem Bias und der Stationshöhe besteht mit zunehmendem Rang ein leicht negativer Trend. Die Analyse der Richterklasse ergab mit zunehmendem Rang einen Trend zu einem vermehrt negativen Bias mit zunehmendem Schutz der Station, was sich besonders bei D180 und D360 zeigt. Schließlich zeigte die Untersuchung des ortsbezogenen Bias pro Radarstandort, dass manche Standorte eine größere Spanne des Bias aufwiesen als andere. Die Analyse einer potentiellen Korrelation zwischen dem Bias und der Distanz der Niederschlagsstation zum entsprechenden Radarturm ergab keinen eindeutigen Trend.

Eine Annahme war das Vorhandensein eines negativen Bias. Dies konnte in dieser Untersuchung nachgewiesen werden. Die Analyse des absoluten und relativen Bias ergab ein überwiegend negatives Bias und damit einer Unterschätzung der Stationsniederschläge durch die Radardaten.

Damit stimmen die Ergebnisse mit vergleichbaren Analysen anderer Studien überein, die meist auch von einer Unterschätzung des Stationsniederschlags berichteten (z. B. Hazenberg et al., 2011, Smith et al., 2012, Fairman Jr et al., 2015, Goudenhoofdt und Delobbe, 2016, Schleiss et al., 2020, Hänslér und Weiler, 2022). Die von Haberlandt und Berndt, 2016 festgestellte Überschätzung des Stationsniederschlags erklärten sie selbst mit einer unzureichenden Korrektur der genutzten Radar-Reflektivitäten.

Eine der Ursachen dieser Abweichung kann mit der unterschiedlich guten räumlichen Abdeckung erklärt werden. Der zunehmende Einfluss der Orographie Deutschland auf die Starkniederschläge bei größeren Dauerstufen (besonders D180 und D360) kann mit den auftretenden Steigungsregen in den Mittelgebirgslagen oder im Alpenraum begründet werden. Mit zunehmender Dauer hat Steigungsregen bei Mittelgebirgslagen einen größeren Einfluss, besonders sehr kurzfristige Starkregenereignisse (D10) treten hingegen vermehrt lokal auf.

Auch Analysen des Radarprodukts RADKLIM zeigten, dass Starkniederschläge von kurzer Dauer grundsätzlich weniger der Topographie Deutschlands folgen. Da sich diese Analysen hauptsächlich auf die Berechnung von Bemessungsniederschlägen bezogen, wurde die Verlässlichkeit dieser Analysen aufgrund der kurzen Radar-Zeitreihe noch als eingeschränkt betrachtet (Winterrath et al., 2017, Junghänel et al., 2022). Die Unterschiede in den Niederschlagskarten der Stations- und Radardaten deckt sich jedoch auch mit der Erwartung, dass kleinskalige Extremereignisse nicht zwingend vom Stationsnetz erfasst werden (Lengfeld et al., 2020). Dies ist durch ihren kleinen Radius von 10–30 km² und ihre kurze Dauer von wenigen Minuten bis hin zu einer halben Stunde begründet (De Lannoy et al., 2005). Folgend kann es zu Unterschieden und damit zu einem Bias zwischen Stations- und Radarprodukt kommen.

Gleichzeitig muss jedoch auch beachtet werden, dass die abgebildeten Niederschlagsereignisse der Stations- und Radarkarte, durch die voneinander unabhängige Auswahl der 150 größten Niederschlagsereignisse, nicht zwangsläufig dasselbe Ereignis darstellen müssen.

Der verhältnismäßig große Einfluss der Dauerstufe auf das Bias ist nicht weiter verwunderlich, da durch die Unterteilung in Dauerstufen der Einfluss verschiedener Niederschlagstypen abgeschätzt werden kann. Bedingt durch die Zeitdauer, ist der Einfluss extremer konvektiver Ereignisse bei den kurzen Dauerstufen (insbesondere D10) größer als bei den größeren, was sich gerade bei den Karten des Bias zeigte. Die Unterteilung in Event-Ränge kann hingegen als eine feinere Untergliederung angesehen werden, die den Einfluss der Niederschlagsintensitäten pro Dauerstufe aufzeigt. Während die ersten Event-Ränge vermehrt extreme konvektive Starkniederschläge darstellen, nimmt dieser konvektive „Charakter“ der Niederschläge mit zunehmendem Rang ab.

Der verhältnismäßig große Anteil an Stationen mit positivem Bias und damit einer Überschätzung der Stationsniederschläge bei den ersten Rängen ist auffällig. Da die ersten Event-Ränge die Ereignisse mit einer sehr extremen Niederschlagsintensität darstellen ist eine Korrelation zwischen Niederschlagsintensität und Überschätzung denkbar.

Meist wird bei hohen Niederschlagsintensitäten mit einer Unterschätzung gegenüber den Stationsmessungen gerechnet. Grundsätzlich ist bei sehr hohen Niederschlagsintensitäten die Qualität der Radarmessung geringer (Berne und Krajewski, 2013). Teilweise wird dies mit der Abschwächung des reflektierten Signals eines Bereiches durch vorgelagerte Bereiche mit hoher Niederschlagsintensität (Jessen et al., 2005) begründet. Oftmals wird als Grund auch eine für hohe Niederschlagsintensitäten ungeeignete Z-R-Beziehung hervorgebracht, die die natürliche Variation der Regentropfen-Charakteristika (raindrop size distribution) mit steigender Intensität nicht vollständig berücksichtigt (Schleiss et al., 2020).

Eine Überschätzung der Niederschlagsintensitäten kann jedoch besonders bei schmelzendem Schneefall beobachtet werden, was zu dem Phänomen des *bright band* führen kann (Fabry und Zawadzki, 1995). Der in dieser Arbeit gefundene zunehmende Einfluss der Winterevents mit größerer Dauerstufe, erscheint auf dem ersten Blick als mögliche Ursache, allerdings nimmt der Anteil der Winterevents mit zunehmendem Event-Rang ab, der Anteil der Stationen mit positivem Bias hingegen zu. Auch spricht dagegen, dass zumindest für lange Dauerstufen die Stationen

mit positivem Bias vermehrt in den Tallagen aufzufinden sind.

Eine mögliche Ursache könnte daher sein, dass die verwendeten Z-R-Beziehungen die gemessenen Reflektivitäten in den Berg- und Tallagen nicht ausreichend unterscheiden. Zwar verwendet der DWD für sein Radarprodukt RADOLAN, zumindest in der Online-Aneichung verfeinerte von der Reflektivität abhängige Z-R-Beziehungen (Bartels et al., 2004), vielleicht reicht dies gerade bei großen Höhendifferenzen noch nicht aus.

Auffallend ist auch, dass der Anteil an Stationen mit positivem Bias mit der Dauerstufe zunimmt. Grundsätzlich wird erwartet, dass mit zunehmender Dauer räumlich ausgedehntere Niederschläge fallen, da konvektive Ereignisse meist von kürzerer Dauer sind (De Lannoy et al., 2005). Grundsätzlich ist aber zu beachten, dass über den Niederschlagsverlauf innerhalb dieses Zeitraums keine Aussagen getroffen werden können. So kann die berechnete Niederschlagssumme das Resultat mehrerer kurzer konvektiver Ereignisse, aber auch nur eines länger dauernden Ereignisses sein.

Das zusehends negativere Bias mit steigender Stationshöhe könnte etwa durch höhen- und entfernungsabhängige Fehler während der Radarmessung zustande kommen. Dass dieser Trend erst bei den späteren Rängen (mit geringerer Intensität) deutlicher sichtbar wird, weist darauf hin, dass dieser Trend bei den ersten Event-Rängen womöglich von anderen Prozessen (wie bright band) überdeckt wird. Insgesamt ist die Signifikanz dieses Trends durch die große Streuung anzuzweifeln.

Das größere Bias in den Mittelgebirgen sowie im Alpenraum und Alpenvorland deutet auf eine schlechtere Performance der Radarmessung im bergigen Terrain hin. Dies stimmt auch mit der Spannbreite des Bias einiger Radarstandorte überein. Denn der Großteil der Radarstandorte mit einer großen Spannbreite des Bias (Feldberg, Isen, Memmingen, Türkheim) befindet sich in Süddeutschland im Bereich des Schwarzwaldes, der Schwäbischen Alb oder des Alpenvorlands. Bei der Betrachtung des relativen Bias sind mit Boostedt und Rostock zwei weitere Standorte in Norddeutschland auffällig. Dies sind meist Radarstandorte, die größere Höhenunterschiede in der Topographie aufweisen, was die Radarmessung insgesamt durch verschiedene Beeinträchtigungen erschwert. Dazu gehören Ground Clutter, Negativ-Speichern, eine ausgeprägte vertikale Variabilität der Niederschlagsverhältnisse und auch häufiger Schneefall (Berne und Krajewski, 2013). Für Nordostdeutschland könnte der geringere Einfluss von Frontensystemen und damit die vermehrte Bildung konvektiver Ereignisse (Lengfeld et al., 2019) ein Grund für die höhere Spannbreite des relativen Bias sein.

Zu beachten ist jedoch auch, dass die Anzahl der Niederschlagsstationen pro Radarschirm im Norden niedriger ist als im Süden und dass manche Niederschlagsstationen mehreren Radarschirmen zuordnen lassen. Jedoch führte eine vergleichende Analyse, bei der jede Station nur dem nächsten Radarturm zugeordnet wurde zu ähnlichen Ergebnissen. Die abweichenden Trends des Bias gegenüber der Entfernung zum Radarturm können vielleicht mit der unterschiedlich erfolgreichen örtlichen Entfernungs- und Höhenkorrektur am Radarstandort begründet sein. In seiner Analyse fand Fairman Jr et al., 2015 hingegen, dass die Distanz zum Radar und die Differenz zwischen Stations- und Radarhöhe nur schwach mit dem Bias korrelieren. Er sah die Fehler/das Bias vielmehr als Folge der Umwandlung der gemessenen Reflektivitäten in Niederschlagsraten und der Korrekturen wegen Änderungen in der Orographie (Fairman Jr et al., 2015).

Die Zunahme des negativen Bias mit stärkerem Schutz vor Windüberwehungen ist auf den ersten Blick widersprüchlich. Zu beachten ist jedoch, dass das Verfahren nur eine gute Näherung darstellt (Richter, 1995) und als Proxy für den Windeinfluss womöglich nicht zwingend geeignet ist. Erwartet wird, dass mit zunehmendem Windschutz der Fehler abnimmt (Richter, 1995), dies ist jedoch auf den Messfehler der Niederschlagsstation bezogen und nicht auf die Messung des Radars. Da ein geschützter Standort eine Lage mit starkem Bewuchs oder Bebauung ist, kann dies einen Einfluss auf die Qualität der Radarmessung nehmen, weil bei mehr Hindernissen, wie Bauwerken oder Berggipfeln vorkommen, die zu einer Abschwächung des reflektierten Signals führen können, sodass die Niederschlagsintensität an diesem Standort unterschätzt wird (Winterrath et al., 2017). Obwohl zumeist versucht wird diese Beeinträchtigungen zu korrigieren (vgl.

Winterrath et al., 2017), ist diese nicht zwingend erfolgreich, wie die erkennbaren Negativspeichen in den Radarkarte dieser Arbeit aufzeigen.

Auch hier sollte jedoch berücksichtigt werden, dass den jeweiligen Richter-Klassen nicht die gleiche Anzahl an Stationen zugrunde liegt, sodass gerade bei der stark geschützten Klasse (31 Stationen) das ortsgebundene Bias im Verhältnis einen größeren Einfluss als bei den ungeschützten Klassen (mit über 800 Stationen) hat. Im Rahmen weiterer Analysen sollte, daher die Lage der (stark) geschützten Niederschlagsstationen genauer untersucht werden, um zu prüfen, ob ein eventueller Zusammenhang mit den Gebirgslagen besteht.

Insgesamt darf jedoch nicht vergessen werden, dass der Vergleich der Stations- und Radardaten generell schwierig ist und mit Unsicherheiten behaftet ist, was teilweise durch das unterschiedliche Messvolumen bedingt ist, das die Vergleichbarkeit beider Produkte erschwert (Schleiss et al., 2020).

Bei dieser Analyse muss zudem beachtet werden, dass Radar- und Stationsereignis nicht zwingend dasselbe Ereignis darstellen müssen, sodass das Auswahlverfahren der 150 größten Niederschlagssummen nicht vollständig als Ursache einer Abweichung ausgeschlossen werden kann. Zwischen den Stations- und Radarereignissen eines Ranges besteht daher kein zeitlicher Zusammenhang. Berechnet wird vielmehr eine Niederschlagsstatistik der größten Extremereignisse einer Dauerstufe pro Datensatz. Dies sollte gerade im Hinblick auf die Biaskorrekturverfahren berücksichtigt werden. Bei dieser Analyse konnten somit nur potentielle Beziehungen zwischen dem Bias und zeitlich möglichst invariablen Eigenschaften des Stations- bzw. Radarstandortes analysiert werden. Zusammenhänge zu finden, die den aktuellen Atmosphärenzustand mitberücksichtigen, bleibt damit die Aufgabe anderer Studien.

Trotzdem ergab die Analyse des Bias in dieser Studie einige Indizien über potentielle Zusammenhänge des Bias mit anderen Parametern. Als die größten Einflüsse können die Dauerstufe und der Event-Rang angesehen werden. Der Einfluss der Orographie Deutschland wird mit zunehmender Dauerstufe, besonders jedoch auch abnehmender Niederschlagsintensität (zunehmenden Event-Rang) sichtbar.

5.3 Biaskorrektur

Im Rahmen dieser Arbeit wurden zwei verschiedene Biaskorrekturverfahren auf Basis des empirischen Quantile-Mappings entwickelt. Für die Biaskorrektur stehen grundsätzlich eine Vielzahl an verschiedenen Methoden zur Verfügung, meist stellen diese auf kriging basierende Merging-Verfahren dar. Cole und Moore, 2008 nannte z. B. das *Multiquadratic surface fitting* als Methode, die besonders für Starkregen geeignet sei. Allerdings ist dafür eine hohe Dichte Niederschlagsstationen notwendig (Jessen et al., 2005). Auch ist unklar, ob für das Kriging die Verwendung angeeichter oder ungeeichte Radardaten besser ist (Junghänel et al., 2022). Bei der Erstellung eines regionalen Datensatzes mithilfe von Radardaten stellen Interpolationsverfahren jedoch meist eine der Hauptfehlerquellen dar (Hänsler und Weiler, 2022).

Das in dieser Arbeit genutzte Korrekturverfahren unterschied sich in einigen Punkten von anderen Studien. Während Overeem et al., 2009a und Haberlandt und Berndt, 2016 andere Typen an Biaskorrekturverfahren anwandten, nutzten Rabiei und Haberlandt, 2015 und Hänsler und Weiler, 2022 auch ein auf Quantile-Mapping basierendes Verfahren.

Im Unterschied zu den Analysen von Overeem et al., 2009a, Rabiei und Haberlandt, 2015 und Haberlandt und Berndt, 2016 wurden für diese Arbeit nicht direkt die Reflektivitätsdaten der Radarstandorte genutzt, sondern eine bereits prozessierte und auf Radardaten basierende quantitative Niederschlagsschätzung verwendet.

Dies erschwert den Vergleich mit diesen Studien, da potentielle Fehlerquellen in der Methodik (wie die Parametrisierung der Z-R-Beziehung oder die Entfernung von Störsignalen) nicht direkt miteinander verglichen werden können.

Zugleich lag der Hauptfokus der zuvor genannten Studien, bis auf Rabiei und Haberlandt, 2015, auf der Nutzbarkeit biaskorrigierter Radardaten für die Berechnung von Bemessungsniederschlä-

gen, wohingegen die hier vorgestellte Arbeit zum Hauptziel hatte neue Biaskorrekturverfahren zu entwickeln und zu testen.

Die Grundidee des Quantile-Mappings für die Korrektur von Radardaten liegt dabei in dem Miteinbeziehen der Informationen von anderen Orten und/oder anderen Zeiten. Damit ähnelt es etwas dem Analogon-Verfahren bei dem historische Atmosphärenzustände als Schätzwert für den aktuellen genutzt werden (Themekl et al., 2011). Die Vorteile des Quantile-Mappings liegen dabei in der Korrektur der ganzen kumulativen Häufigkeitsverteilung (CDF) und dem Erhalt der räumlichen Struktur (Rabiei und Haberlandt, 2015, Hänsler und Weiler, 2022). Dies konnte auch in der hier durchgeführten Arbeit bestätigt werden. Aufgrund des recht ähnlichen Korrekturverfahrens findet folgend ein Vergleich mit Haberlandt und Berndt, 2016 und insbesondere mit Hänsler und Weiler, 2022 statt.

Die Analyse von Rabiei und Haberlandt, 2015 fokussierte sich auf die Entwicklung eines Verfahrens zur Korrektur von stündlichen Niederschlagsdaten. Die biaskorrigierten Radardaten dienten letztlich als zusätzliche flächenhafte Informationen für das Interpolationsverfahren. Die Berechnung von Bemessungsniederschlägen und damit eine Schwerpunktsetzung auf die Statistik der Niederschläge fand nicht statt, daher wurden auch reale stündliche Radardaten verwendet. Interessanterweise verbesserte die Biaskorrektur nicht nur die Qualität der Radardaten, sondern auch die des Interpolationsverfahrens (Rabiei und Haberlandt, 2015). Bei den hier entwickelten Biaskorrekturverfahren war nicht die Nutzung für die Interpolation im Vordergrund, hier lag das Hauptaugenmerk hingegen vielmehr auf der Korrektur der Statistik der Extremniederschläge.

Der Hauptunterschied zwischen dem hier in dieser Arbeit verwendeten Verfahren und dem in Hänsler und Weiler, 2022 beschriebenen liegt in der Reihenfolge der Methoden. Hänsler und Weiler, 2022 fitteten zuerst eine Generalisierte Paretoverteilung (GPD) an die ausgewählten Niederschlagsereignisse der Radar-Rasterzellen und korrigierten anschließend die Parameter dieser Verteilung für ein jährliches Ereignis mittels Quantile-Mapping. Da nur die jährlichen Events zur Korrektur des Location-Parameters der GPD verwendet wurden und die zusätzliche Regenmenge zu allen Events der Extremwertanalyse addiert wurden, nimmt der relative Beitrag der Biaskorrektur für Bemessungsniederschläge mit höherer Jährlichkeit ab, sodass gerade bei den selteneren weiterhin eine Unterschätzung der Stationsniederschläge feststellbar war. Der entscheidende Nachteil dieses Verfahrens stellte somit die additive Korrektur dar (Hänsler und Weiler, 2022). In dieser Arbeit wurden hingegen in einem ersten Schritt die Radarniederschläge mittels Quantile-Mapping korrigiert, und anschließend als Anwendungsbeispiel für zwei Jährlichkeiten Bemessungsniederschläge berechnet. Somit wurde hier der Vorschlag von Hänsler und Weiler, 2022 aufgegriffen und direkt die Radarniederschläge statt der Parameter der GPD korrigiert. Dies hat den Vorteil, dass so keine Abhängigkeit von den Korrekturfaktoren besteht, die für schwächere Ereignissen berechnet wurden und auf stärkere Ereignisse nur unzureichend angewandt werden können (Hänsler und Weiler, 2022).

Während Rabiei und Haberlandt, 2015 und Hänsler und Weiler, 2022 für ihre Analysen eine theoretische Verteilungsfunktion im Rahmen des Quantile-Mapping-Verfahrens verwendeten, wurde sich in dieser Arbeit für einen empirischen Ansatz entschieden.

Erwartet wurde dabei eine Reduzierung des Bias nach Anwendung der Biaskorrektur. Beide entwickelten Verfahren führten auch zu einer Reduzierung des Bias. Von einer Reduzierung des Bias zwischen Stations- und Radardaten wurde auch von anderen Autoren nach Anwendung ihrer Korrekturmethode berichtet (Overeem et al., 2009a, Rabiei und Haberlandt, 2015, Haberlandt und Berndt, 2016, Hänsler und Weiler, 2022). Folgend soll auf verschiedene Einflüsse und Bewertungskriterien eingegangen werden.

Einfluss der Dauerstufe

Während bei Verfahren 1 keine gewichtigen Abweichungen in der Performance der Korrektur mit zunehmender Dauerstufe feststellbar waren, war bei Verfahren 2 die Korrektur der großen Event-Ränge für große Dauerstufen schlechter. Dies kann daran liegen, dass die Einteilung nach

Event-Rängen und damit eine feinere Unterteilung nach Niederschlagsintensität fehlt, sodass bei Verfahren 2 insgesamt die Spannbreite der Verteilungen größer ist und der Einfluss der Dauerstufe sichtbarer ist als bei Verfahren 1.

Einfluss der Trainings- und Kontrollgruppen

Die Anwendung von Verfahren 1 führte zu einer sehr guten Korrektur der Trainingsgruppe und erwartungsgemäß zu einer recht guten, aber etwas schlechteren Korrektur der Kontrollgruppe. Auch Hänslér und Weiler, 2022 stellten in ihrer Analyse eine fast perfekte Übereinstimmung der biaskorrigierten Radardaten mit den Stationsdaten für die Stationsstandorte fest. Verfahren 2 hingegen zeigte für Trainings- und Kontrollgruppe eine ähnlich gute bzw. schlechte Korrektur. Die Unterschiede zwischen beiden Gruppen waren dabei geringer. Die schlechtere Performance bei der Korrektur der Trainingsgruppen ist durch das Verfahren zur Berechnung der Quantile begründet. Während bei Verfahren 1 eine rangbasierte Berechnung der Korrekturfaktoren stattfindet und für jeden Rang ein eigenes Set an passenden Korrekturfaktoren vorhanden ist, wird bei Verfahren 2 dasselbe Set an Korrekturfaktoren auf alle Ränge gleichermaßen angewandt.

Einfluss der Event-Ränge

Verfahren 1 zeichnete sich dadurch aus, dass bei der Gegenüberstellung der korrigierten Radarwerten und der Stationswerte für die ersten Ränge ein eher linearer Verlauf, für die weiteren Ränge ein eher S-förmiger Verlauf sichtbar war. Bei Verfahren 2 konnte dieser lineare Verlauf bei den ersten Rängen nicht beobachtet werden. Hier fiel auf, dass die Werte im Bereich der großen Quantile für die ersten Event-Ränge teils zu stark korrigiert wurden, für die letzten Ränge dagegen nicht ausreichend korrigiert wurden. Dies zeigte sich besonders bei den größeren Dauerstufen (D180, D360). Auch dies kann mit dem einen Set an Korrekturfaktoren bei Verfahren 2 erklärt werden. Dieses stellt ungefähr den mittleren Zustand der Ränge dar. Insgesamt zeigt Verfahren 2 aus diesem Grund auch im Bereich der kleinen Quantile eine noch geringere Korrektur als Verfahren 1, bei den großen Quantilen zeigt sich meist eine passende oder zu geringe Korrektur aber seltener als bei Verfahren 1 eine zu starke Korrektur.

Einfluss der Quantile

Generell zeigten beide Verfahren die besten Korrekturergebnisse im mittleren Quantil-Bereich, die Korrektur des Randbereichs war schlechter. Während bei Verfahren 1 die größten Quantile eher „überkorrigiert“ wurden, war die Radar-Niederschlagswerte für Verfahren 2 in diesem Bereich entweder passend oder zu klein. Die kleinsten Quantile wurden bei beiden Verfahren eher mangelhaft korrigiert, bei Verfahren 1 war trotzdem der Fit insgesamt besser als bei Verfahren 2. Der Grund für die bessere Korrektur der mittleren Quantile könnte an den ähnlicheren Niederschlagswerten zwischen Stations- und Radardaten liegen. Dies zeigt sich auch dadurch, dass die Abweichungen für die mittelgroßen Quantile teilweise bereits vor der Korrektur geringer waren als die der kleinen und großen Quantile. Die kleinen und großen Niederschlagsintensitäten hingegen stellen den Randbereich der Niederschlagsverteilung eines Datensets dar, sodass Abweichungen zwischen Stations- und Radardaten wahrscheinlicher sind. Gleichzeitig führt dies dazu, dass der Randbereich der Verteilung empfindlicher auf die Auswahl der Stationen (und damit der Niederschlagswerte) bei der Auftrennung in Trainings- und Kontrollgruppen reagiert, sodass die Gruppenunterschiede meist besonders in diesen Bereichen größer sind. Die unzureichende Korrektur der kleineren Niederschlagsintensitäten (kleinen Quantile) kann teilweise mit dem Verlauf der Korrekturfaktoren begründet werden, da in diesem Bereich die größten Faktoren aufzufinden sind und diese demnach durch die angewandte Glättung am stärksten betroffen sind.

Einfluss der Gruppenunterschiede

Wurden die drei Trainings- bzw. Kontrollgruppen miteinander verglichen, zeigten sich für Verfahren 1 einige Differenzen, besonders für die großen Quantile, wohingegen in Verfahren 2 kaum Unterschiede der Gruppen feststellbar waren. Der größere Einfluss der drei Trainingsgruppen bei der Berechnung der Korrekturfaktoren und damit der Korrektur der Radar-Niederschlagswerte bei Verfahren 1 kann wohl auf die geringere Anzahl an Daten pro Berechnung zurückgeführt werden. Bei Verfahren 2 werden diese Gruppenunterschiede vermutlich durch die insgesamt größere Anzahl an Niederschlagswerten ausgeglichen.

Verbleibendes Bias

Die Analyse der verbleibenden Bias ergab, dass die Radardaten an manchen Stationen zu stark korrigiert wurden und nun teilweise ein positives Bias aufwiesen. Dabei muss jedoch beachtet werden, dass die verwendeten Verfahren die Informationen aller Stationsstandorte (der Trainingsgruppe) für die Korrektur nutzen. Korrigiert wird also die Niederschlagsstatistik, nicht standortbezogen.

Analyse der Korrekturfaktoren

Die Analyse der Korrekturfaktoren ergab bei beiden Verfahren sehr hohe Korrekturfaktoren für die ersten Quantile und wiederum ein Ansteigen der Faktoren im Bereich der großen Quantile. Dieser Anstieg ist bei den Dauerstufen D180 und D360 größer und zudem bei Verfahren 2 stärker ausgeprägt als bei Verfahren 1. Die Faktoren für die kleinsten Quantile sind bei Verfahren 2 kleiner als bei Verfahren 1. Bei beiden Verfahren werden mit zunehmender Dauerstufe die Korrekturfaktoren kleiner. Verfahren 1 weist größere Unterschiede zwischen den Gruppen auf als Verfahren 2. Die Korrekturfaktoren selbst weisen bei Verfahren 2 einen glatteren Verlauf auf, was darauf hindeutet, dass mehr (ähnliche) Datenpunkte zur Verfügung standen. Sanfte Übergänge des Korrekturfaktors sind wünschenswert, um Sprünge zu vermeiden. Dazu kann versucht werden die Anzahl der Quantile zu erhöhen, oder aber die Korrekturfaktoren in diesem Bereich stärker zu glätten. Auch die Mittelung der Korrekturfaktoren der drei Gruppen könnte dieses Problem vermindern.

Räumlicher Vergleich

Der räumliche Vergleich beider Verfahren zeigte, dass Verfahren 2 für die ersten Ränge die Bereiche mit geringer Niederschlagsintensität und für die späteren Ränge die mittelgroßen Intensitäten stärker korrigiert als Verfahren 1. Bei einer multiplikativen Korrektur ändert sich das räumliche Muster zwar grundsätzlich nicht, die Unterschiede der Radarpixel untereinander können jedoch durchaus verändert werden, das heißt nicht jedes Radar-Pixel wird gleich stark korrigiert.

Die Anwendung der ermittelten Korrekturfaktoren auf die gesamte Fläche Deutschlands und damit alle Radar-Pixel basiert letztlich auf der Annahme, dass die Verteilung der Starkniederschläge der Stationen für den gesamten Raum repräsentativ ist. Ob dies tatsächlich zutrifft, bleibt unklar. Das Grundproblem bei der Bewertung von biaskorrigierten Niederschlagsdaten bleibt das Fehlen eines Datensatzes zur Validierung (Hänsler und Weiler, 2022). Unter der Annahme, dass die Stationsdaten an ihrem Standort korrekte Werte aufweisen, ist zumindest dort eine direkte Bewertung möglich. Für die Regionen, die sich weiter von den Stationsdaten entfernt befinden ist dies nicht möglich.

Grundsätzlich ist ein Vergleich mit interpolierten Stationsprodukten denkbar, dies beinhaltet jedoch zwei Schwächen. Zum einen sind diese selbst aufgrund des Interpolationsverfahren mit Unsicherheiten behaftet, zum anderen können sich die verschiedenen Interpolationsprodukte

selbst voneinander unterscheiden. Wird das interpolierte Produkt nun als Referenz genutzt, kann das Biaskorrekturverfahren eine maximal so gute Performance aufweisen, wie das Interpolationsverfahren des Stationsproduktes. In gewisser Hinsicht sind Abweichungen vom interpolierten Stationsprodukt auch durchaus erwünscht, da meist davon ausgegangen wird, dass die räumlichen Muster der Radarprodukte, zumindest für kleine Dauern, meist zutreffender sind (Lengfeld et al., 2019). Schließlich muss beachtet werden, dass die Niederschlagsmessungen der Stationen selbst Unsicherheiten beinhalten können (Schleiss et al., 2020). Die fachliche Plausibilisierung der räumlichen Muster bleibt damit wohl weiter eine wichtige Aufgabe.

Analyse der Quantile

Da das Grundkonzept des Quantile-Mappings darin besteht die Quantile einer Verteilung gleichzusetzen, haben die zugrundeliegenden Verteilungen einen entscheidenden Einfluss. Während die empirische Varianz der verwendeten Radardaten kleiner ist als die der Stationsdaten, sind die Stationsdaten generell rechtsschiefer als die Stationsdaten. Der Grund dafür liegt wohl an den zugrundeliegenden Datenverteilung, was sich auch auf die berechneten Niederschlagssummen auswirkt. Wie die Analyse der Quantile zeigte, hat wohl sowohl die Varianz als auch die Schiefe der Häufigkeitsverteilungen der Stations- und Radardaten einen entscheidenden Einfluss auf die Verteilung der Korrekturfaktoren.

Grundsätzlich ist auch die Anzahl der Daten von Bedeutung. Bei Verfahren 1 stehen für die Berechnung der 1000 Quantile und damit auch der Korrekturfaktoren pro Rang jeweils 1154 Werte zur Verfügung, wohingegen bei Verfahren 2 die 1000 Quantile aus 173100 Werten berechnet werden. Somit wächst der Einfluss des Interpolationsverfahrens der Quantilberechnung je weniger Daten im Verhältnis zur Anzahl der Quantil-Klassen vorhanden sind.

Diese Datenlastigkeit des Verfahrens bringt Vorteile und Nachteile mit sich. Ein Vorteil des empirischen Quantile-Mappings ist das dieses Verfahren auf verschiedene Parameter anwendbar ist und keine Annahmen zur zugrundeliegenden Verteilungsfunktion getroffen werden muss (Thiemeßl et al., 2011). Ein entscheidender Nachteil ist jedoch die Sensitivität des Verfahrens gegenüber den zugrundeliegenden Daten. Wenn nur wenig Datenpunkte existieren, können somit weitere Unsicherheiten entstehen (Rabiei und Haberlandt, 2015).

Da die Bildung von Datenpaaren bzw. „Quantilpaaren“ mithilfe der Stations- und Radardaten notwendig sind, sollten diese denselben Zeitraum umfassen, was gerade im Hinblick auf Extremwertanalysen die Nutzbarkeit einschränkt. Hänslers und Weilers, 2022 äußerte zudem die Unsicherheit der Stationsdaten durch unterschiedlich lange Zeitreihen als ein Problem, das die Analyse beeinträchtigen kann. Im Rahmen dieser Arbeit sollte dies weniger problematisch sein, da durch die gefüllten und Richter-korrigierten Stationsdaten grundsätzlich gleich viele Daten vorhanden sein sollten. Trotzdem kann eben dieses Füllen der Fehlwerte durch den Regionalisierungsansatz bzw. die Anpassung an die Tagessumme weitere Unsicherheiten einführen. Daher war auch für diese Analyse eine Plausibilisierung der Stationsdaten ausweichlich. Unsicherheiten diesbezüglich können aber nicht vollständig ausgeschlossen werden.

Zudem ist eine Extrapolation nötig, wenn ein neuer größerer Extremwert auftritt, der bisher nicht in den Messdaten enthalten war. Oftmals wird hierfür schlicht der letzte Wert verwendet bzw. hier der letzte Korrekturfaktor auf den größten Wert angewandt (Boé et al., 2007),

Auch sind die ermittelten Korrekturfaktoren in dieser Form nur für den vorherrschenden Datensatz verwendbar. Zwar kann das Set an Korrekturfaktoren grundsätzlich auf einen anderen Datensatz mit gleich vielen Quantil-Klassen angewandt, der Bezug zum ursprünglichen Niederschlagswert des Quantils ist so aber nicht mehr gegeben. Falls nur ein Niederschlagswert gegeben ist (vgl. das Anwendungsbeispiel der realen Events) und dieser sehr von den Quantil-Werten abweicht, ist die Sinnhaftigkeit den zugehörigen Korrekturfaktor anzuwenden fraglich.

Sensitivitätstestanalysen

Der erste Sensitivitätstest diente der Abschätzung des Einflusses der Quantil-Klassen (Bins) auf die Korrekturmethode. Hierbei zeigte sich, dass für die ersten Quantile ein schlechterer Fit zustande kam, sonst aber sehr ähnliche Ergebnisse wie bei Verfahren 2 erkennbar waren. In ihrer Analyse verwendeten Hänsler und Weiler, 2022 100 diskrete Bins, diese stellten jedoch theoretische Quantile dar.

Lafon et al., 2013 testeten u. a. empirische Quantile-Mapping-Verfahren mit 25, 50, 75 und 100 Quantil-Klassen im Rahmen einer Analyse von täglichen Niederschlagsdaten eines regionalen Klimamodells. Dabei fanden sie heraus, dass mit größer werdender Anzahl der Quantil-Klassen, die zugrundeliegenden Häufigkeitsverteilungen besser abgebildet und korrigiert wurden. Wenn versucht wird die etwa 600 Werte einer Gruppe in 1000 Bins einzuteilen, bleiben manche der Bins grundsätzlich leer, was eine (lineare) Interpolation zur Berechnung der Quantile unumgänglich macht. Im Verhältnis zu 1000 Bins wird bei 100 Bins derselbe Korrekturfaktor auf mehr Werte angewandt.

Auch wenn die Ergebnisse dieser Analyse hauptsächlich Unterschiede für die kleinen Quantile eines Ranges feststellten, kann davon ausgegangen werden, dass eine größere Anzahl an Quantil-Klassen abrupte Änderungen der Korrekturfaktoren, die als Sprünge im Q-Q-Plot erkennbar sind, reduziert bzw. vermeidet. Dies zeigt sich auch im Verlauf der Korrekturfaktoren, für das erste Quantil ist der Faktor gegenüber Verfahren 2 deutlich reduziert, was mit der größeren Berechnung der Quantile in Verbindung gebracht werden kann.

Bei Verwendung des mittleren Radarniederschlags der vier nächsten Raster-Zellen (Mean4) statt nur des Stations-Pixels, wird für die ersten Event-Ränge sichtbar, dass die Radardaten der Trainings- und Kontrollgruppe die Stationswerte mehr unterschätzen als bei Verfahren 2. Dies kann wohl auf die Mittelung zurückgeführt werden. Diese ist auch die Ursache dass die kleinsten Quantile im Vergleich zu Verfahren 2 weniger unterschätzt werden, die großen Quantile dafür vermehrt.

Der abnehmende Anteil an Sommer-Events mit zunehmendem Event-Rang und zunehmender Dauerstufe deckt sich mit der Erwartung, dass gerade die kurzen und extremen Niederschlagsereignisse konvektiv bedingt sind und damit hauptsächlich im Sommer vorkommen. Die Performance der Korrektur ist insgesamt der des zweiten Verfahrens recht ähnlich. Bei den letzten Rängen zeigt sich wie zu erwarten eine geringere Anzahl an Daten. Für die kleinen Quantile der letzten Ränge ist dabei teilweise ein besserer Fit, teilweise auch ein schlechterer Fit zu erkennen.

Rabiei und Haberlandt, 2015 fanden in ihrer Analyse eine bessere Qualität der Radardaten im Winter als im Sommer, und folgerte, dass die Biaskorrektur im Sommer, wegen der vermehrt konvektiven Ereignisse, wichtiger sei. Hänsler und Weiler, 2022 wählten nur Ereignisse der Sommermonate (April bis Oktober), um den Vergleich mit anderen Datensätzen zu gewährleisten und um feste Niederschläge im Winter zu vermeiden. Aufgrund der insgesamt mangelnden Anzahl an Winterereignissen konnte eine ausschließliche Korrektur dieser Winterereignisse in dieser Analyse nicht durchgeführt werden.

Bei ausschließlicher Verwendung von im Datum übereinstimmender Events konnte eine deutlich schlechtere Performance der Korrektur festgestellt werden. Ursache hierfür ist schlicht die geringe Anzahl an Stations- und Radarereignissen, die in ihrem Datum übereinstimmen (Stat-Px.: ca. 5 %, Mean4: ca. 20

Abschließend lässt sich festhalten, dass die Anzahl der Bins einen recht großen Einfluss auf den Verlauf der Korrekturfaktoren haben, gefolgt von der Beschränkung auf die übereinstimmenden Events und der Nutzung von Mean4 als Grundlage der Radardaten. Die ausschließliche Verwendung von Sommerereignissen hatte einen recht geringen Einfluss, was aber auch daran liegen kann, dass sich der Großteil der Ereignisse ohnehin im Sommer ereignete. Angewandt auf die Gesamtfläche Deutschlands führte dies jedoch nur zu relativ geringen Unterschieden.

5.4 Anwendungsbeispiele

Die Betrachtung der neun realen Extremereignisse zeigte, dass die berechneten Korrekturfaktoren zur Korrektur der Niederschlagssummen realer Ereignisse meist nicht geeignet sind. Da das Verfahren zur Korrektur der Niederschlagsstatistik konzipiert wurde ist dies nicht weiter verwunderlich. Beim Vergleich der Stations- und Radardaten während des Extremereignissen konnte eine mehr oder weniger passende zeitliche Übereinstimmung festgestellt werden. Auch andere Analysen stellten eine recht gute Übereinstimmung fest, allerdings nicht für Dauern unter einer Stunde, was hauptsächlich auf die unterschiedlichen Messvolumen zurückgeführt wurde (Schleiss et al., 2020). Wichtiger als eine zeitliche Korrektur erscheint jedoch eine räumliche Korrektur (Rabiei und Haberlandt, 2015).

Die berechneten Bemessungsniederschläge für ein 20-jährliches und ein 50-jährliches Ereignis auf Basis der korrigierten Radar-Niederschlagsdaten zeigten sehr große Abweichungen zum KOSTRA-DWD-2020V-Produkt. Trotz der Unterschätzung der Niederschlagsintensitäten im Vergleich zu KOSTRA, führt die Anwendung der Biaskorrektur meist zu einer Erhöhung der Bemessungsniederschläge. Dabei stellte sich Verfahren 1 als wirksamer heraus als Verfahren 2. Jedoch konnten bei den größeren Dauerstufen beim 50-jährlichen Ereignis nach der Korrektur Bereiche mit geringeren Niederschlagsintensitäten festgestellt werden, was auf ein fehlerhaftes Fitten der Generalisierten-Pareto-Verteilung hindeutet.

Generell muss festgehalten werden, dass diese Berechnung einen rein exemplarischen Charakter hat. Eine genauere Abwägung, inwieweit die Nutzung der Generalisierten-Pareto Funktion mit drei Parametern für den kurzen betrachteten Zeitraum von 17 Jahren und die betrachteten Bemessungsniederschläge im Bereich von maximal 50 Jahren sinnvoll ist, wird hier nicht durchgeführt, sollte jedoch in weiteren Analysen durchgeführt werden.

Die Ergebnisse auf Basis der Radardaten berechneten Bemessungsniederschlägen stimmen mit anderen Studien überein. Die geringe Korrelation der ermittelten Bemessungsniederschlägen mit der Orographie für kleine Dauerstufen und die Zunahme dieser Korrelation mit steigender Dauerstufe wurde auch in vergleichenden Analysen festgestellt (Winterrath et al., 2017, Hänsler und Weiler, 2022, Junghänel et al., 2022).

Das räumliche Muster auf Basis der korrigierten Radardaten entspricht grundsätzlich den theoretischen Erwartungen, dass kleinskalige konvektive Ereignisse bei kleinen Dauerstufen und größere Frontensystemen bei größeren Dauerstufen überwiegen (Hänsler und Weiler, 2022). Auch die Unterschätzung in der Niederschlagsintensität deckt sich mit den Ergebnissen anderen Studien (Hänsler und Weiler, 2022, Junghänel et al., 2022), was hauptsächlich mit der kurzen Zeitreihe begründet wird (Junghänel et al., 2022). Da für die Praxis gerade Bemessungsniederschläge mit einer hoher Jährlichkeit von Bedeutung sind, ist der kurze Zeitraum für die Extremwertanalyse hinderlich, da die Berechnung bei kleinen Stichproben nur mit großer Ungenauigkeit möglich ist. Gerade die Werte einer hohen Jährlichkeit hängen von der Schiefe der Stichprobe ab (Junghänel et al., 2022).

Overeem et al., 2009a kamen in ihrer Studien zu dem Schluss, dass eine Extremwertanalyse mit Radardaten grundsätzlich möglich ist, wenn eine *regional frequency analysis* durchgeführt wird. Auch sie stellten jedoch in Frage, wie nützlich die Radardaten wegen des kurzen Zeitraums wirklich sind. Eine Lösungsfindung, die bessere räumliche Auflösung der Niederschlags-Radardaten mit der umfassenderen Zeitreihen der Stationen zu kombinieren ist bleibt damit weiterhin eine Aufgabe weiterer Analysen.

5.5 Abschließende Einschätzung

Zu Beginn der Analyse wurde angenommen, dass ein rangbasierte Quantile-Mapping-Ansatz (Verfahren 1) eine bessere Performance aufweist als eine gleichzeitige Anwendung auf alle Daten (Verfahren 2).

Diese Annahme konnte nicht vollständig bestätigt werden. Beide Verfahren führten zu einer

Reduktion des Bias, trotzdem ist jedes der beiden Verfahren in gewissen Situationen dem anderen überlegen. Der rangbasierte Ansatz zeichnet sich durch einen sehr guten Fit der Trainingsgruppe aus und ist in diesem Punkt Verfahren 2 überlegen. Auch der bessere Fit für die kleinen Quantile spricht für Verfahren 1. Für Verfahren 2 spricht die einfachere Berechnung und die geringeren Unterschiede zwischen den Gruppen. Dies spricht für eine geringere Sensitivität bezüglich der Stationsauswahl. Auch für die Korrektur der mittelgroßen bis großen Niederschlagsintensitäten kann Verfahren 2 die bessere Wahl sein. Grundsätzlich stellen die zugehörigen Korrekturfaktoren quasi eine Mittelung über die Ränge bei Verfahren 1 dar. Die Anwendung auf die Gesamtfläche offenbart einige interessante Unterschiede, auch wenn die Korrektheit dieser räumlichen Analysen weiter ungeklärt bleibt.

Die Wahl des Verfahrens orientiert sich damit an der Fragestellung und der Datenverfügbarkeit. Folgend seien einige Handlungsempfehlungen genannt. Wenn genügend Daten vorhanden sind und diese ausreichend unterschiedlich sind, sollte der rangbasierte Ansatz (Verfahren 1) verwendet werden, da so eine bessere Korrektur der einzelnen Niederschlagsintensitäten durch die Aufteilung in Ränge stattfinden kann. Auch wenn der Fokus ohnehin auf eine Abstufung der Intensitäten liegt, ist Verfahren 1 meist die bessere Wahl. Verfahren 1 ist zudem nützlich, um den Einfluss der Stationsauswahl besser zu untersuchen. Falls die Datenlage jedoch mangelhaft ist und nur wenige Daten zur Verfügung stehen und wenn die Abhängigkeit von den gewählten Stationen gering sein soll, ist Verfahren 2 zu bevorzugen. Auch die teilweise zu starke Korrektur bei Verfahren 1 kann für dieses Verfahren sprechen. Da der Fokus im Bereich der Extremniederschläge, insbesondere in der Praxis, meist auf der Korrektur der extremen Ereignisse liegt ist das Vorhandensein eines „Puffers“ sinnvoll, um etwaige Unsicherheiten besser ausgleichen zu können. Schließlich ist festzuhalten, dass das Verfahren sehr sensitiv auf Änderungen der Daten reagiert. Daher ist eine gute Qualität der Radar- und Stationsdaten umso wichtiger.

6 Schlussfolgerungen

Die Untersuchung des Bias zwischen Radar- und Stationsdaten offenbarte ein bestehendes Bias in Form einer Unterschätzung der Stationsniederschläge durch die Radardaten. Gleichzeitig wurde an manchen Stationsstandorten eine Überschätzung der Stationsniederschläge festgestellt. Die durchgeführte Arbeit zeichnet sich auch durch die deutschlandweite Analyse des Bias aus. Verschiedene Parameter wurden auf ihren Zusammenhang mit dem Bias untersucht. Hier könnten weitere Analysen ansetzen, um die Ursachen für die Unterschätzung oder Überschätzung räumlich besser zu lokalisieren.

Mit dem Ziel das bestehende Bias zu verringern bzw. zu beseitigen wurden zwei Biaskorrekturverfahren basierend auf einem empirischen Quantile-Mapping-Ansatz entwickelt. Das erste Verfahren stellte einen rangbasierten Ansatz dar, das zweite sah die gleichzeitige Korrektur aller verwendeten Ereignisse vor. Beide Verfahren führten erwartungsgemäß zu einer Reduktion des Bias.

Die Annahme, dass das rangbasierte Verfahren dem nicht-rangbasierten in der Performance der Korrektur überlegen ist, konnte nicht eindeutig festgestellt werden. Während eine gute Datenverfügbarkeit für den rangbasierten Ansatz spricht, ist die Anwendung des zweiten Verfahrens gerade bei einer geringeren Anzahl an Daten vorteilhaft. Die entwickelten Verfahren zeichnen sich somit durch eine große Sensitivität gegenüber den verwendeten Daten aus.

Eine Verbesserung der Qualität der Radarreflektivitäten selbst bzw. der daraus abgeleiteten Niederschlagssummen sollte das primäre Ziel sein, um das Bias möglichst gering zu halten. Trotz der bereits vom DWD verfeinerten Z-R-Beziehungen, kann auf Seiten der Betreiber versucht werden die Parametrisierung dynamischer zu gestalten und insbesondere im Bereich der großen Reflektivitätswerte zu optimieren.

Eine weitere Möglichkeit stellt eine Veränderung der stündlichen Aneichung der Radardaten an die Stationsdaten dar. Eine mehrmalige Aneichung innerhalb einer Stunde könnte gerade im Hinblick auf kurzfristige konvektive Ereignisse helfen, deren Dauer meist kürzer als eine Stunde sind (Schleiss et al., 2020).

Da oftmals eben diese Parametrisierung der Z-R-Beziehung als Ursache für das bestehende Bias erwähnt wird, könnten auch direkt die Reflektivitätsdaten statt der abgeleiteten Niederschlagssummen für ein Korrekturverfahren verwendet werden. Hierbei ist jedoch zu beachten, dass die Fehler in der Radarmessung (Positiv- bzw. Negativ-Speichen, Ground Clutter oder Beam Blocking) enthalten sein können. Auch müsste in diesem Fall selbstständig ein Komposit berechnet werden, wenn die Analyse nicht nur für einen Radarstandort durchgeführt wird. Gleichzeitig ist ein solches Verfahren vermutlich sehr auf den jeweiligen Radarstandort beschränkt und eine Übertragung auf andere Standorte könnte sich als schwierig erweisen.

Bezüglich der entwickelten Methode zur Biaskorrektur wäre es interessant die Performance beider Verfahren zu untersuchen, wenn ausschließlich Winterereignisse verwendet werden. Hierbei wäre es auch interessant zu sehen, ob der Anteil des positiven Bias für die großen Dauerstufen abnimmt.

Des Weiteren könnte getestet werden, ob die Verwendung von Datenverteilungen, die nicht rechts- oder linksschief sind zu einer Verbesserung der Korrektur führen.

Schließlich könnte versucht werden die Abhängigkeit des rangbasierten Verfahrens zu reduzieren. Dazu könnte ein genügend großes Set an Trainings- und Kontrollgruppen ausgewählt und die resultierenden Sets an Korrekturfaktoren gemittelt werden.

Im Hinblick auf die Berechnung der Bemessungsniederschläge könnte die Anwendung einer regional frequency analysis interessant sein, um validere Schätzungen dieser zu erhalten.

7 Datensätze

Die in der Arbeit verwendeten Datensätze sind entweder frei zugänglich oder wurden im Rahmen des Projektes zur Verfügung gestellt. Zu den verwendeten externen Datenquellen zählen:

- D1** Eurostat/GISCO (2023): Countries 2020, Version 03.04.2020, <https://gisco-services.ec.europa.eu/distribution/v2/countries/download/ref-countries-2020-01m.shp.zip>
- D2** DWD (2023a): Climate Data Center - 10-minütige Stationsmessungen des Niederschlages für Deutschland (Historische Daten), https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/precipitation/historical/
- D3** DWD (2023b): Climate Data Center - 10-minütige Stationsmessungen des Niederschlages für Deutschland (Aktuelle Daten), https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/precipitation/recent/
- D4** DWD (2023c): Climate Data Center - 10-minütige Stationsmessungen des Niederschlages für Deutschland (Metadaten), https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/precipitation/meta_data/
- D5** DWD (2023d): Climate Data Center - 5-minütige Stationsmessungen des Niederschlages für Deutschland (Historisch), https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/5_minutes/precipitation/historical/
- D6** DWD (2023e): Climate Data Center - Stündliche Stationsmessungen des Niederschlages für Deutschland (Historisch), https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/precipitation/historical/
- D7** DWD (2023f): Climate Data Center - KOSTRA 2020, https://opendata.dwd.de/climate_environment/CDC/grids_germany/return_periods/precipitation/KOSTRA/KOSTRA_DWD_2020/
- D8** Professur für Hydrologie - Albert-Ludwigs-Universität Freiburg i. Br. (2023): WeatherDB, <https://apps.hydro.uni-freiburg.de/de/weatherdb/> [Zugriff nur im Intranet der Universität Freiburg möglich]
- D9** Meteologix AG (2023): Radar SD, <https://kachelmannwetter.com/de>

Literatur

- Abon, C. C., Kneis, D., Crisologo, I., Bronstert, A., David, C. P. C. und Heistermann, M. (2016). “Evaluating the potential of radar-based rainfall estimates for streamflow and flood simulations in the Philippines”. In: *Geomatics, Natural Hazards and Risk* 7.4, S. 1390–1405. DOI: <https://doi.org/10.1080/19475705.2015.1058862>.
- Anagnostou, E. N., Krajewski, W. F. und Smith, J. (1999). “Uncertainty Quantification of Mean-Areal Radar-Rainfall Estimates”. In: *Journal of Atmospheric and Oceanic Technology* 16.2, S. 206–215. DOI: [https://doi.org/10.1175/1520-0426\(1999\)016<0206:UQOMAR>2.0.CO;2](https://doi.org/10.1175/1520-0426(1999)016<0206:UQOMAR>2.0.CO;2).
- Auerswald, K., Fischer, F. K., Winterrath, T. und Brandhuber, R. (2019). “Rain erosivity map for Germany derived from contiguous radar rain data”. In: *Hydrology and Earth System Sciences* 23.4, S. 1819–1832. DOI: <https://doi.org/10.5194/hess-23-1819-2019>.
- Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O. und Gerlach, N. (2004). *Projekt RADOLAN. Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer)*. Offenbach am Main, Deutschland: Deutscher Wetterdienst. 111 S. URL: https://www.dwd.de/DE/leistungen/radolan/radolan_info/abschlussbericht_pdf.pdf.
- Berne, A. und Krajewski, W. F. (2013). “Radar for hydrology: Unfulfilled promise or unrecognized potential?” In: *Advances in Water Resources* 51, S. 357–366. DOI: <https://doi.org/10.1016/j.advwatres.2012.05.005>.
- Block, P. J., Souza Filho, F. A., Sun, L. und Kwon, H.-H. (2009). “A Streamflow Forecasting Framework using Multiple Climate and Hydrological Models1”. In: *JAWRA Journal of the American Water Resources Association* 45.4, S. 828–843. DOI: <https://doi.org/10.1111/j.1752-1688.2009.00327.x>.
- Boé, J., Terray, L., Habets, F. und Martin, E. (2007). “Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies”. In: *International Journal of Climatology* 27.12, S. 1643–1655. DOI: <https://doi.org/10.1002/joc.1602>.
- Bohm, G. und Zech, G. (2020). *Statistische Datenanalyse für Physiker*. 1. Aufl. Berlin, Heidelberg, Deutschland: Springer Spektrum. 115 S. ISBN: 978-3-662-61391-7. DOI: <https://doi.org/10.1007/978-3-662-61391-7>.
- Cannon, A. J., Sobie, S. R. und Murdock, T. Q. (2015). “Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?” In: *Journal of Climate* 28.17, S. 6938–6959. DOI: <https://doi.org/10.1175/JCLI-D-14-00754.1>.
- Chen, H., Cifelli, R., Chandrasekar, V. und Ma, Y. (2019). “A Flexible Bayesian Approach to Bias Correction of Radar-Derived Precipitation Estimates over Complex Terrain: Model Design and Initial Verification”. In: *Journal of Hydrometeorology* 20.12, S. 2367–2382. DOI: <https://doi.org/10.1175/JHM-D-19-0136.1>.
- Chen, J., Brissette, F. P., Chaumont, D. und Braun, M. (2013). “Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America”. In: *Water Resources Research* 49.7, S. 4187–4205. DOI: <https://doi.org/10.1002/wrcr.20331>.
- Chumchean, S., Seed, A. und Sharma, A. (2006). “Correcting of real-time radar rainfall bias using a Kalman filtering approach”. In: *Journal of Hydrology* 317.1–2, S. 123–137. DOI: <https://doi.org/10.1016/j.jhydro1.2005.05.013>.
- Cole, S. J. und Moore, R. J. (2008). “Hydrological modelling using raingauge- and radar-based estimators of areal rainfall”. In: *Journal of Hydrology* 358.3, S. 159–181. DOI: <https://doi.org/10.1016/j.jhydro1.2008.05.025>.

- Crisologo, I., Warren, R. A., Mühlbauer, K. und Heistermann, M. (2018). “Enhancing the consistency of spaceborne and ground-based radar comparisons by using beam blockage fraction as a quality filter”. In: *Atmospheric Measurement Techniques* 11.9, S. 5223–5236. DOI: <https://doi.org/10.5194/amt-11-5223-2018>.
- De Lannoy, G. J. M., Verhoest, N. E. C. und De Troch, F. P. (2005). “Characteristics of rainstorms over a temperate region derived from multiple time series of weather radar images”. In: *Journal of Hydrology* 307.1, S. 126–144. DOI: <https://doi.org/10.1016/j.jhydrol.2004.10.007>.
- Dettinger, M. D., Cayan, D. R., Meyer, M. K. und Jeton, A. E. (2004). “Simulated Hydrologic Responses to Climate Variations and Change in the Merced, Carson, and American River Basins, Sierra Nevada, California, 1900–2099”. In: *Climatic Change* 62.1–3, S. 283–317. DOI: <https://doi.org/10.1023/B:CLIM.0000013683.13346.4f>.
- Fabry, F. und Zawadzki, I. (1995). “Long-Term Radar Observations of the Melting Layer of Precipitation and Their Interpretation”. In: *Journal of Atmospheric Sciences* 52.7, S. 838–851. DOI: [https://doi.org/10.1175/1520-0469\(1995\)052<0838:LTR00T>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<0838:LTR00T>2.0.CO;2).
- Fairman Jr, J. G., Schultz, D. M., Kirshbaum, D. J., Gray, S. L. und Barrett, A. I. (2015). “A radar-based rainfall climatology of Great Britain and Ireland”. In: *Weather* 70.5, S. 153–158. DOI: <https://doi.org/10.1002/wea.2486>.
- Fang, G. H., Yang, J., Chen, Y. N. und Zammit, C. (2015). “Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China”. In: *Hydrology and Earth System Sciences* 19.6, S. 2547–2559. DOI: <https://doi.org/10.5194/hess-19-2547-2015>.
- Fischer, F. K., Winterrath, T. und Auerswald, K. (2018). “Temporal- and spatial-scale and positional effects on rain erosivity derived from point-scale and contiguous rain data”. In: *Hydrology and Earth System Sciences* 22.12, S. 6505–6518. DOI: <https://doi.org/10.5194/hess-22-6505-2018>.
- Germann, U., Berenguer, M., Sempere-Torres, D. und Zappa, M. (2009). “REAL—Ensemble radar precipitation estimation for hydrology in a mountainous region”. In: *Quarterly Journal of the Royal Meteorological Society* 135.639, S. 445–456. DOI: <https://doi.org/10.1002/qj.375>.
- Gesamtverband der Deutschen Versicherungswirtschaft e. V. (2019). *Forschungsprojekt Starkregen, Summary November 2019*. Berlin: Gesamtverband der Deutschen Versicherungswirtschaft e. V. URL: <https://www.gdv.de/resource/blob/52868/c6d7ffceab5d13fc0f7659496ced6421/forschungsprojekt-starkregen-summary-download-data.pdf> (besucht am 26.08.2023).
- Gilewski, P. und Nawalany, M. (2018). “Inter-Comparison of Rain-Gauge, Radar, and Satellite (IMERG GPM) Precipitation Estimates Performance for Rainfall-Runoff Modeling in a Mountainous Catchment in Poland”. In: *Water* 11, 1665. DOI: <https://doi.org/10.3390/w10111665>.
- Goudenhoofd, E. und Delobbe, L. (2009). “Evaluation of radar-gauge merging methods for quantitative precipitation estimates”. In: *Hydrology and Earth System Sciences* 13.2, S. 195–203. DOI: <https://doi.org/10.5194/hess-13-195-2009>.
- (2016). “Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements”. In: *Journal of Hydrometeorology* 17.4, S. 1223–1242. DOI: <https://doi.org/10.1175/JHM-D-15-0166.1>.
- Grazioli, J., Tuia, D. und Berne, A. (2015). “Hydrometeor classification from polarimetric radar measurements: a clustering approach”. In: *Atmospheric Measurement Techniques* 8.1, S. 149–170. DOI: <https://doi.org/10.5194/amt-8-149-2015>.
- Haberlandt, U. und Berndt, C. (2016). “The value of weather radar data for the estimation of design storms – an analysis for the Hannover region”. In: *Proceedings of the International Association of Hydrological Sciences* 373, S. 81–85. DOI: <https://doi.org/10.5194/piahs-373-81-2016>.
- Hanchoo Wong, R., Weesakula, U. und Chumcheanb, S. (2012). “Bias correction of radar rainfall estimates based on a geostatistical technique”. In: *ScienceAsia* 38, S. 373–385. DOI: <https://doi.org/10.2306/scienceasia1513-1874.2012.38.373>.

- Hänsler, A. und Weiler, M. (2022). “Enhancing the usability of weather radar data for the statistical analysis of extreme precipitation events”. In: *Hydrology and Earth System Sciences* 26.19, S. 5069–5084. DOI: <https://doi.org/10.5194/hess-26-5069-2022>.
- Hashino, T., Bradley, A. A. und Schwartz, S. S. (2007). “Evaluation of bias-correction methods for ensemble streamflow volume forecasts”. In: 11.2, S. 939–950. DOI: <https://doi.org/10.5194/hess-11-939-2007>.
- Hazenberg, P., Leijnse, H. und Uijlenhoet, R. (2011). “Radar rainfall estimation of stratiform winter precipitation in the Belgian Ardennes”. In: *Water Resources Research* 47.2, W02507. DOI: <https://doi.org/10.1029/2010WR009068>.
- Hedderich, J. und Sachs, L. (2020). *Angewandte Statistik. Methodensammlung mit R*. 17. Aufl. Berlin, Heidelberg, Deutschland: Springer Spektrum. 1054 S. ISBN: 978-3-662-62294-0. DOI: <https://doi.org/10.1007/978-3-662-62294-0>.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J. und Piontek, F. (2013). “A trend-preserving bias correction – the ISI-MIP approach”. In: *Earth System Dynamics* 4.2, S. 219–236. DOI: <https://doi.org/10.5194/esd-4-219-2013>.
- Hyndman, R. J. und Fan, Y. (1996). “Sample Quantiles in Statistical Packages”. In: *The American Statistician* 50.4, S. 361–365. DOI: <https://doi.org/10.2307/2684934>.
- Jessen, M., Einfalt, T., Stoffer, A. und Mehlig, B. (2005). “Analysis of heavy rainfall events in North Rhine–Westphalia with radar and raingauge data”. In: *Atmospheric Research* 77.1, S. 337–346. DOI: <https://doi.org/10.1016/j.atmosres.2004.11.031>.
- Junghänel, T., Bär, F., Deutschländer, T., Haberlandt, U., Otte, I., Shehu, B., Stockel, H., Stricker, K., Thiele, L.-B. und Willems, W. (2022). *Methodische Untersuchungen zur Novellierung der Starkregenstatistik für Deutschland (MUNSTAR). Synthesebericht*. Techn. Ber. Deutscher Wetterdienst. 95 S. URL: https://www.dwd.de/DE/leistungen/kostra_dwd_rasterwerte/download/Synthesebericht_MUNSTAR_pdf.pdf (besucht am 30.10.2023).
- Junghänel, T., Ostermüller, J. und Deutschländer, T. (2023). *KOSTRA-DWD-2020 Anwendungshilfe*. Techn. Ber. Deutscher Wetterdienst. 24 S. URL: https://www.dwd.de/DE/leistungen/kostra_dwd_rasterwerte/download/kostra_dwd_2020_anwenderhilfe_pdf.pdf (besucht am 30.10.2023).
- Kidd, C., Bauer, P., Turk, J., Huffman, G. J., Joyce, R., Hsu, K.-L. und Braithwaite, D. (2012). “Intercomparison of High-Resolution Precipitation Products over Northwest Europe”. In: *Journal of Hydrometeorology* 13.1, S. 67–83. DOI: <https://doi.org/10.1175/JHM-D-11-042.1>.
- Kitchen, M. und Blackall, R. M. (1992). “Representativeness errors in comparisons between radar and gauge measurements of rainfall”. In: *Journal of Hydrology* 134.1, S. 13–33. DOI: [https://doi.org/10.1016/0022-1694\(92\)90026-R](https://doi.org/10.1016/0022-1694(92)90026-R).
- Lafon, T., Dadson, S., Buys, G. und Prudhomme, C. (2013). “Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods”. In: *International Journal of Climatology* 33.6, S. 1367–1381. DOI: <https://doi.org/10.1002/joc.3518>.
- Landesamt für Geologie, Rohstoffe und Bergbau (2023). *Sturzflutereignis bei Braunsbach im Mai 2016 | LGRBwissen*. Landesamt für Geologie, Rohstoffe und Bergbau. URL: <https://lgrbwissen.lgrb-bw.de/ingenieurgeologie/massenbewegungen/murgangaehnliche-ereignisse-fliessen/sturzflutereignis-bei-braunsbach-im-mai-2016> (besucht am 26.08.2023).
- LARSIM-Entwicklergemeinschaft (2023). *Das Wasserhaushaltsmodell LARSIM. Modellgrundlagen und Anwendungsbeispiele*. Techn. Ber. Hochwasserzentralen LUBW, BLfU, LfU RP, HL-NUG, BAFU. 260 S. URL: <http://www.larsim.info/dokumentation/LARSIM-Dokumentation.pdf> (besucht am 10.10.2023).
- Leander, R. und Buishand, T. A. (2007). “Resampling of regional climate model output for the simulation of extreme river flows”. In: *Journal of Hydrology* 332.3–4, S. 487–496. DOI: <https://doi.org/10.1016/j.jhydrol.2006.08.006>.
- Leander, R., Buishand, T. A., van den Hurk, B. J. J. M. und de Wit, M. J. M. (2008). “Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output”.

- In: *Journal of Hydrology* 351.3–4, S. 331–343. DOI: <https://doi.org/10.1016/j.jhydrol.2007.12.020>.
- Lengfeld, K., Kirstetter, P.-E., Fowler, H. J., Yu, J., Becker, A., Flamig, Z. und Gourley, J. (2020). “Use of radar data for characterizing extreme precipitation at fine scales and short durations”. In: *Environmental Research Letters* 15.8, 085003. DOI: <https://doi.org/10.1088/1748-9326/ab98b4>.
- Lengfeld, K., Walawender, E., Winterrath, T. und Becker, A. (2021). “CatRaRE: A Catalogue of radar-based heavy rainfall events in Germany derived from 20 years of data”. In: *Meteorologische Zeitschrift* 30.6, S. 469–487. DOI: <https://doi.org/10.1127/metz/2021/1088>.
- Lengfeld, K., Winterrath, T., Junghänel, T. und Becker, A. (2019). “Characteristic spatial extent of rain events in Germany from a radar-based precipitation climatology”. In: *Rainfall Monitoring, Modelling and Forecasting in Urban Environment. UrbanRain18: 11th International Workshop on Precipitation in Urban Areas. Conference Proceedings*. Hrsg. von N. Peleg und P. Molnar. Zürich, Schweiz: ETH Zürich, Institute of Environmental Engineering, S. 61–66. DOI: <https://doi.org/10.3929/ethz-b-000347543>.
- Li, H., Sheffield, J. und Wood, E. F. (2010). “Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching”. In: *Journal of Geophysical Research: Atmospheres* 115, D10. DOI: <https://doi.org/10.1029/2009JD012882>.
- Maraun, D. (2016). “Bias Correcting Climate Change Simulations - a Critical Review”. In: *Current Climate Change Reports* 2, S. 211–220. DOI: <https://doi.org/10.1007/s40641-016-0050-x>.
- Maurer, E. P. und Hidalgo, H. G. (2008). “Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods”. In: *Hydrology and Earth System Sciences* 12.2, S. 551–563. DOI: <https://doi.org/10.5194/hess-12-551-2008>.
- McKee, J. L. und Binns, A. D. (2016). “A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology”. In: *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 41.1–2, S. 186–203. DOI: <https://doi.org/10.1080/07011784.2015.1064786>.
- Meißner, D., Gebauer, S., Schumann, A. H., Pahlow, M. und Rademacher, S. (2012). “Analyse radarbasierter Niederschlagsprodukte als Eingangsdaten verkehrsbezogener Wasserstandsvorhersagen am Rhein”. In: *Hydrologie und Wasserbewirtschaftung* 56.1, S. 16–28. DOI: https://doi.org/10.5675/HyWa_2012_1_2.
- Michelangeli, P.-A., Vrac, M. und Loukos, H. (2009). “Probabilistic downscaling approaches: Application to wind cumulative distribution functions”. In: *Geophysical Research Letters* 36.11, L11708. DOI: <https://doi.org/10.1029/2009GL038401>.
- Mpelasoka, F. S. und Chiew, F. H. S. (2009). “Influence of Rainfall Scenario Construction Methods on Runoff Projections”. In: *Journal of Hydrometeorology* 10.5, S. 1168–1183. DOI: <https://doi.org/10.1175/2009JHM1045.1>.
- Overeem, A., Buishand, T. A. und Holleman, I. (2009a). “Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar”. In: *Water Resources Research* 45.10, W10424. DOI: <https://doi.org/10.1029/2009WR007869>.
- Overeem, A., Holleman, I. und Buishand, A. (2009b). “Derivation of a 10-Year Radar-Based Climatology of Rainfall”. In: *Journal of Applied Meteorology and Climatology* 48.7, S. 1448–1463. DOI: <https://doi.org/10.1175/2009JAMC1954.1>.
- Ozkaya, A. und Akyurek, Z. (2019). “Evaluating the use of bias-corrected radar rainfall data in three flood events in Samsun, Turkey”. In: *Natural Hazards* 98, S. 643–674. DOI: <https://doi.org/10.1007/s11069-019-03723-z>.
- Panziera, L., Gabella, M., Zanini, S., Hering, A., Germann, U. und Berne, A. (2016). “A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland”. In: *Hydrology and Earth System Sciences* 20.6, S. 2317–2332. DOI: <https://doi.org/10.5194/hess-20-2317-2016>.

- Pejcic, V., Garfias, P. S., Mühlbauer, K., Troemel, S. und Simmer, C. (2022). “Evaluation of Germany’s network radar composite rain products with GPM near surface precipitation estimations” (San Francisco, USA, 9.–13. Dez. 2019). DOI: <https://doi.org/10.1002/essoar.10501387.1>.
- Pejcic, V., Trömel, S., Mühlbauer, Kai, Saavedra, Pablo, Beer, Josephin und Simmer, Clemens (2018). “Synergy of GPM and ground-based radar observations for precipitation estimation and detection of microphysical processes”. In: *2018 19th International Radar Symposium (IRS)*. 2018 19th International Radar Symposium (IRS) (Bonn, Germany, 20.–22. Juni 2018). Institute of Electrical und Electronics Engineers, S. 1–8. ISBN: 978-1-5386-1269-9, 978-3-7369-9545-1 (Online). DOI: <https://doi.org/10.23919/IRS.2018.8447923>.
- Piani, C. und Haerter, J. O. (2012). “Two dimensional bias correction of temperature and precipitation copulas in climate models”. In: *Geophysical Research Letters* 39.20, L20401. DOI: <https://doi.org/10.1029/2012GL053839>.
- Piani, C., Haerter, J. O. und Coppola, E. (2010a). “Statistical bias correction for daily precipitation in regional climate models over Europe”. In: *Theoretical and Applied Climatology* 99.1–2, D10, S. 187–192. DOI: <https://doi.org/10.1007/s00704-009-0134-9>.
- Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S. und Haerter, J. O. (2010b). “Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models”. In: *Journal of Hydrology* 395.3–4, S. 199–215. DOI: <https://doi.org/10.1016/j.jhydrol.2010.10.024>.
- Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T. und Hegewisch, K. C. (2015). “Improved Bias Correction Techniques for Hydrological Simulations of Climate Change”. In: *Journal of Hydrometeorology* 16.6, S. 2421–2442. DOI: <https://doi.org/10.1175/JHM-D-14-0236.1>.
- Pöschmann, J. M., Kim, D., Kronenberg, R. und Bernhofer, C. (2021). “An analysis of temporal scaling behaviour of extreme rainfall in Germany based on radar precipitation QPE data”. In: *Natural Hazards and Earth System Sciences* 21.4, S. 1195–1207. DOI: <https://doi.org/10.5194/nhess-21-1195-2021>.
- Rabiei, E. und Haberlandt, U. (2015). “Applying bias correction for merging rain gauge and radar data”. In: *Journal of Hydrology* 522, S. 544–557. ISSN: 0022-1694. DOI: <https://doi.org/10.1016/j.jhydrol.2015.01.020>.
- Ramsauer, T., Weiß, T. und Marzahn, P. (2018). “Comparison of the GPM IMERG Final Precipitation Product to RADOLAN Weather Radar Data over the Topographically and Climatically Diverse Germany”. In: *Remote Sensing* 10.12, 2029. DOI: <https://doi.org/10.3390/rs10122029>.
- Richter, D. (1995). “Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-Niederschlagsmessers”. In: *Berichte des Deutschen Wetterdienstes*. Hrsg. von Deutscher Wetterdienst. Bd. 195. Offenbach am Main, Deutschland: Selbstverlag des Deutschen Wetterdienstes. ISBN: ISBN 978-3-88148-309-4. URL: <http://nbn-resolving.de/urn:nbn:de:101:1-201601274368>.
- Rojas, R., Feyen, L., Dosio, A. und Bavera, D. (2011). “Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations”. In: *Hydrology and Earth System Sciences* 15.8, S. 2599–2620. DOI: <https://doi.org/10.5194/hess-15-2599-2011>.
- Sackett, D. L. (1979). “Bias in analytic research”. In: *Journal of chronic diseases* 32.1–2, S. 51–63. DOI: [https://doi.org/10.1016/0021-9681\(79\)90012-2](https://doi.org/10.1016/0021-9681(79)90012-2).
- Saltikoff, E., Friedrich, K., Soderholm, J., Lengfeld, K., Nelson, B., Becker, A., Hollmann, R., Urban, B., Heistermann, M. und Tassone, C. (2019). “An Overview of Using Weather Radar for Climatological Studies: Successes, Challenges, and Potential”. In: *Bulletin of the American Meteorological Society* 100.9, S. 1739–1752. DOI: <https://doi.org/10.1175/BAMS-D-18-0166.1>.

- Schleiss, M., Olsson, J., Berg, P., Niemi, T., Kokkonen, T., Thorndahl, S., Nielsen, R., Ellerbæk Nielsen, J., Bozhinova, D. und Pulkkinen, S. (2020). “The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden”. In: *Hydrology and Earth System Sciences* 24.6, S. 3157–3188. DOI: <https://doi.org/10.5194/hess-24-3157-2020>.
- Seo, B.-C. und Krajewski, W. F. (2010). “Scale Dependence of Radar Rainfall Uncertainty: Initial Evaluation of NEXRAD’s New Super-Resolution Data for Hydrologic Applications”. In: *Journal of Hydrometeorology* 11.5, S. 1191–1198. DOI: <https://doi.org/10.1175/2010JHM1265.1>.
- Seo, D.-J., Breidenbach, J. P. und Johnson, E. R. (1999). “Real-time estimation of mean field bias in radar rainfall data”. In: *Journal of Hydrology* 223.3–4, S. 131–147. DOI: [https://doi.org/10.1016/S0022-1694\(99\)00106-7](https://doi.org/10.1016/S0022-1694(99)00106-7).
- Seo, D.-J. und Breidenbach, J. P. (2002). “Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements”. In: *Journal of Hydrometeorology* 3.2, S. 93–111. DOI: [https://doi.org/10.1175/1525-7541\(2002\)003<0093:RTCSN>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0093:RTCSN>2.0.CO;2).
- Smith, J. A., Baeck, M. L., Villarini, G., Welty, C., Miller, A. J. und Krajewski, W. F. (2012). “Analyses of a long-term, high-resolution radar rainfall data set for the Baltimore metropolitan region”. In: *Water Resources Research* 48.4, W04504. DOI: <https://doi.org/10.1029/2011WR010641>.
- Smith, J. A. und Krajewski, W. F. (1991). “Estimation of the Mean Field Bias of Radar Rainfall Estimates”. In: *Journal of Applied Meteorology and Climatology* 30.4, S. 397–412. DOI: [https://doi.org/10.1175/1520-0450\(1991\)030<0397:EOTMFB>2.0.CO;2](https://doi.org/10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2).
- Society, American Meteorological (25. Apr. 2012a). *Model output statistics – Glossary of Meteorology*. American Meteorological Society. URL: https://glossary.ametsoc.org/wiki/Model_output_statistics (besucht am 27.07.2023).
- (26. Jan. 2012b). *Perfect prognostic – Glossary of Meteorology*. American Meteorological Society. URL: https://glossary.ametsoc.org/wiki/Perfect_prognostic (besucht am 27.07.2023).
- Tesfagiorgis, K., Mahani, S. E., Krakauer, N. Y. und Khanbilvardi, R. (2011). “Bias correction of satellite rainfall estimates using a radar-gauge product – a case study in Oklahoma (USA)”. In: *Hydrology and Earth System Sciences* 15.8, S. 2631–2647. DOI: <https://doi.org/10.5194/hess-15-2631-2011>.
- Teutschbein, C. und Seibert, J. (2012). “Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods”. In: *Journal of Hydrology* 456–457, S. 12–29. DOI: <https://doi.org/10.1016/j.jhydrol.2012.05.052>.
- Themeßl, M. J., Gobiet, A. und Heinrich, G. (2012). “Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal”. In: *Climatic Change* 112.2, S. 449–468. DOI: <https://doi.org/10.1007/s10584-011-0224-4>.
- Themeßl, M. J., Gobiet, A. und Leuprecht, A. (2011). “Empirical-statistical downscaling and error correction of daily precipitation from regional climate models”. In: *International Journal of Climatology* 31.10, S. 1530–1544. DOI: <https://doi.org/10.1002/joc.2168>.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., Veldhuis, M.-C. ten, Arnbjerg-Nielsen, K., Rasmussen, M. R. und Molnar, P. (2017). “Weather radar rainfall data in urban hydrology”. In: *Hydrology and Earth System Sciences* 21.3, S. 1359–1380. DOI: <https://doi.org/10.5194/hess-21-1359-2017>.
- Villarini, G., Serinaldi, F. und Krajewski, W. F. (2008). “Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches”. In: *Advances in Water Resources* 31.12, S. 1674–1686. DOI: <https://doi.org/10.1016/j.advwatres.2008.08.002>.
- Vrac, Mathieu und Friederichs, Petra (2015). “Multivariate—Intervariable, Spatial, and Temporal—Bias Correction”. In: *Journal of Climate* 28, S. 218–237. DOI: <https://doi.org/10.1175/JCLI-D-14-00059.1>.

- Watanabe, S., Kanae, S., Seto, S., Yeh, P. J.-F., Hirabayashi, Y. und Oki, T. (2012). “Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models”. In: *Journal of Geophysical Research: Atmospheres* 117.D23, D23114. DOI: <https://doi.org/10.1029/2012JD018192>.
- Weigl, E. (27. Feb. 2018). *RADOLAN Information Nr. 46*. Deutscher Wetterdienst. URL: https://www.dwd.de/DE/leistungen/radolan/radolan_info/radolan_info_nr_46.pdf (besucht am 10. 10. 2023).
- (6. Mai 2021). *RADOLAN Information Nr. 51*. Deutscher Wetterdienst. URL: https://www.dwd.de/DE/leistungen/radolan/radolan_info/radolan_info_nr_51.pdf (besucht am 10. 10. 2023).
- Weiler, M., Hänsler, A., Zimmer, J. und Moser, M. (2019). “Nutzung von Radardaten im Starkregenrisikomanagement in Baden- Württemberg”. In: *Wasserwirtschaft* 12, S. 63–67. URL: <https://doi.org/10.1007/s35147-019-0311-4>.
- Wetterdienst, Deutscher (25. Apr. 2018). *Metadaten zu den Radaren des Radarverbunds des DWD*. URL: https://www.dwd.de/DE/derdwd/messnetz/atmosphaerenbeobachtung/_functions/HaeufigGesucht/koordinaten-radarverbund.pdf (besucht am 10. 10. 2023).
- (2023). *Warnkriterien*. URL: https://www.dwd.de/DE/wetter/warnungen_aktuell/kriterien/warnkriterien.html (besucht am 10. 10. 2023).
- Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Walawender, E., Weigl, E. und Becker, A. (2017). *Erstellung einer radargestützten Niederschlagsklimatologie*. Berichte des Deutschen Wetterdienstes 251. Offenbach am Main, Deutschland: Deutscher Wetterdienst. 71 S. ISBN: 978-3-88148-499-2. URL: <http://nbn-resolving.de/urn:nbn:de:101:1-20170908911>.
- Winterrath, T., Brendel, T., Junghänel, T., Klameth, A., Lengfeld, K., Walawender, E., Weigl, E., Hafer, E. und Becker, A. (2019). “An overview of the new radar-based precipitation climatology of the Deutscher Wetterdienst – data, methods, products”. In: *Rainfall Monitoring, Modelling and Forecasting in Urban Environment. UrbanRain18: 11th International Workshop on Precipitation in Urban Areas. Conference Proceedings*. Hrsg. von N. Peleg und P. Molnar. Zürich, Schweiz: ETH Zürich, Institute of Environmental Engineering, S. 132–137. DOI: <https://doi.org/10.3929/ethz-b-000347607>.
- Yoo, C., Park, C., Yoon, J. und Kim, J. (2014). “Interpretation of mean-field bias correction of radar rain rate using the concept of linear regression”. In: *Hydrological Processes* 28.19, S. 5081–5092. DOI: <https://doi.org/10.1002/hyp.9972>.
- Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Van Cooten, S., Kelleher, K., Kitzmiller, D., Ding, F., Seo, D.J., Wells, E. und Dempsey, C. (2011). “National Mosaic and Multi-Sensor QPE (NMQ) System: Description, Results, and Future Plans”. In: *Bulletin of the American Meteorological Society* 92.10, S. 1321–1338. DOI: <https://doi.org/10.1175/2011BAMS-D-11-00047.1>.

Anhang

Beschreibung verschiedener Arten der Biaskorrektur

Direkte Methoden

Linear scaling (LS): Hier wird das Verhältnis der Mittelwerte von Messdaten und historischen Modelldaten gebildet um ein Korrekturfaktor für die zukünftigen Modelldaten zu finden (vgl. Hashino et al., 2007; Teutschbein und Seibert, 2012; J. Chen et al., 2013; Fang et al., 2015; Maraun, 2016).

Local intensity scaling (LOCI): LOCI erweitert das *linear scaling*, indem vor der Berechnung des Korrekturfaktors ein Schwellenwert für Niederschläge definiert wird, sodass die Anzahl der Regentage in Modell- und Messdaten übereinstimmt (vgl. Themeßl et al., 2011; Teutschbein und Seibert, 2012; J. Chen et al., 2013; Fang et al., 2015).

Delta-Methode: Die Delta-Methode wiederum nutzt die prognostizierte Zeitreihe nicht direkt, sondern berechnet das Verhältnis der Mittelwerte zwischen modellierten historischen und zukünftigen Modelldaten. Die Messdaten werden anschließend mit diesem Signal multipliziert, um so korrigierte zukünftige Daten zu erhalten (vgl. Mpelasoka und Chiew, 2009; Teutschbein und Seibert, 2012; Watanabe et al., 2012; Maraun, 2016).

Power transformation: Hier wird ein exponentieller Zusammenhang angenommen. Als Maß wird der Variationskoeffizient genutzt, und der Exponent so bestimmt, dass sich die Variationskoeffizienten zwischen Messdaten und modellieren Kontrolldaten entsprechen (vgl. Leander und Buishand, 2007; Leander et al., 2008; Teutschbein und Seibert, 2012; Watanabe et al., 2012; Fang et al., 2015).

Quantile-Mapping auf Basis der Gamma-Verteilung: Diese häufig verwendete Methode wird auch „*distribution mapping*“ genannt (vgl. Teutschbein und Seibert, 2012). Für Niederschläge wird meist die Gamma-Verteilung verwendet. Dieses Verfahren basiert damit auf der Annahme, dass sowohl die historischen Modelldaten als auch die Messdaten dieser Verteilung folgen. Zuerst wird an die Messdaten und historischen Modelldaten jeweils eine Gamma-Verteilung gefittet. Anschließend wird die kumulierte Gamma-Verteilungsfunktion der historischen Modelldaten auf die prognostizierten Modelldaten angewandt und diese danach mithilfe der inversen kumulierten Gamma-Verteilungsfunktion der Messdaten transformiert. Dadurch findet eine quantilbezogene Korrektur statt (vgl. Li et al., 2010; Piani et al., 2010a; Teutschbein und Seibert, 2012; Watanabe et al., 2012; J. Chen et al., 2013; Fang et al., 2015; Pierce et al., 2015; Maraun, 2016).

Quantile-Mapping auf Basis einer hergeleiteten Transferfunktion: Das von Piani et al., 2010b entwickelte Verfahren umgeht diese a-priori Annahme einer Gamma-Verteilung, indem die Transferfunktion direkt aus den Daten selbst bestimmt wird. Dazu werden erst kumulierte empirische Verteilungsfunktionen der historischen Modell- und Messdaten gebildet und diese zueinander ins Verhältnis gesetzt. An dieses Verhältnis wird ein bestimmter Zusammenhang (z. B. eine lineare Gleichung) gefittet und diesen verwendet, um die prognostizierten Modelldaten zu korrigieren (vgl. Piani et al., 2010b; Watanabe et al., 2012).

Empirisches Quantile-Mapping: Wie der Name bereits andeutet, werden auch hier empirische kumulierte Verteilungsfunktionen gebildet und keinerlei Verteilung an die Daten gefittet.

Aus dem Verhältnis der historischen Model- und Messdaten, lassen sich so Biaskorrekturfaktoren ableiten, die für die Korrektur der prognostizierten Modelldaten genutzt werden können (vgl. Hashino et al., 2007; Mpelasoka und Chiew, 2009; Themeßl et al., 2011; Themeßl et al., 2012; J. Chen et al., 2013; Cannon et al., 2015; Fang et al., 2015; Vrac und Friederichs, 2015; Maraun, 2016).

Detrended quantile mapping: Die bisherigen Methoden beruhen auf der Annahme, dass das Bias zeitlich stationär ist und aus der Vergangenheit auf die Gegenwart bzw. Zukunft übertragen werden kann. Diese Annahme trifft (gerade im Bereich der Klimamodellierungen) nicht zwingend zu. Das bisherige Quantile-Mapping-Verfahren wird angepasst, indem die jeweilige (invers) kumulierte Verteilungsfunktion mit dem Verhältnis aus den langjährigen Mitteln der jeweiligen Zeireihe multipliziert wird (vgl. Cannon et al., 2015).

Cumulative distribution function transform (CDF-t): Beim üblichen *quantile mapping* wird zu einem gegebenen Wert das Quantil der kumulierten Verteilungsfunktion der historischen Modelldaten bestimmt und anschließend der entsprechende Wert für dieses Quantil bei der Verteilungsfunktion der Messdaten ermittelt. Bei der CDF-t-Methode wird die Reihenfolge der Operationen vertauscht. Hier wird zuerst der zu einem gegebenen Quantil der kumulierten Verteilungsfunktion der historischen Modelldaten gehörende Wert ermittelt und anschließend zu diesem Wert das entsprechende Quantil bei der Verteilungsfunktion der Messdaten bestimmt. Bei der CDF-t-Methode wird das zu einem bestimmten Wert gehörende Quantil geändert, wohingegen beim üblichen Quantile-Mapping der zu einem bestimmten Quantil gehörende Wert geändert wird. Während beim üblichen Quantile-Mapping-Verfahren also angenommen wird, dass der Fehler im Niederschlagswert in Zukunft bei diesem Wert noch derselbe ist, wird bei der CDF-t-Methode davon ausgegangen, dass der Fehler in dem Quantil bei ebendiesem Quantil in Zukunft noch derselbe ist (vgl. Michelangeli et al., 2009; Pierce et al., 2015; Vrac und Friederichs, 2015).

Equidistant quantile matching (EDCDFm): Auch diese Methode verfolgt eine etwas andere Strategie. Hier werden die kumulierten Verteilungsfunktionen von historischen und prognostizierten Modelldaten verglichen. Dabei wird pro Quantil die Differenz zwischen beiden ermittelt und auf die Messdaten angewandt. Damit entspricht diese Methode letztlich eine quantilbezogene *Delta-Methode* (vgl. Li et al., 2010; Watanabe et al., 2012; Pierce et al., 2015). Ein ähnliches Verfahren ist *daily scaling* (vgl. Mpelasoka und Chiew, 2009; J. Chen et al., 2013) oder *quantile delta mapping* (vgl. Cannon et al., 2015). Während EDCDFm eine additive Korrektur vornimmt, wird beim *quantile delta mapping* eine multiplikative Korrektur angewandt (Cannon et al., 2015).

Indirekte Methoden

Multiple lineare Regressionsmodelle: Bei dieser Methode werden für jede Station sinnvolle Variablen ausgewählt mit deren Hilfe, die Antwortvariable korrigiert werden kann (vgl. Themeßl et al., 2011).

Multiple linearen Regressionsmodelle mit Randomisierung: Als Erweiterung zu den *Multiplen-Linearen-Regressionsmodellen*, werden hier die klassifizierten Residuen aus der multiplen linearen Regression als zufälliger Fehler-Term addiert (vgl. Themeßl et al., 2011).

Analogon-Methode: Ein nichtlineares, stochastisches Verfahren stellt die Analogon-Methode dar. Bei dieser Resampling-Methode wird der Atmosphärenzustand eines Tages mit einer Sammlung historischer ähnlicher (analoger) Zustände verglichen. Dieser historische Zustand wird dann als Schätzwert für die Antwortvariable genutzt. Dieses Verfahren setzt damit allerdings voraus, dass es bereits Messungen eines vergleichbaren Zustands gibt (vgl. Themeßl et al., 2011).

Nächste-Nachbarn-Analogon-Methode: Dieser Ansatz stellt eine Weiterentwicklung des *Analogon-Verfahrens* dar. Hier wird aus den ähnlichsten Atmosphärenzuständen mithilfe des Nächsten-Nachbarn-Algorithmus zufällig ein Zustand ausgewählt. Dadurch soll die Variabilität erhöht werden (vgl. Themeßl et al., 2011).

Multivariate Methoden

Bivariate conditional correction: Diese Methode nutzt das mathematische Konzept der Copula. In dieser Methode wird erst eine univariate Biaskorrektur auf eine Variable (z. B. Temperatur) angewandt und die empirisch kumulierte Verteilungsfunktion gebildet. Anschließend werden zu dieser Variablen die zugehörigen Werte der anderen Variable (z. B. Niederschlag) zugeordnet und diese Paare in definierte Quantil-Intervalle gruppiert. Zuletzt findet eine univariate Biaskorrektur des zweiten Parameters (Niederschlag) für jede Gruppe/jeden Bin statt. Durch die Kopplung sollen die Abhängigkeiten zwischen den zwei Variablen reduziert werden (vgl. Piani und Haerter, 2012; Vrac und Friederichs, 2015).

Empirical copula-bias correction: Dieses von Vrac und Friederichs als *empirical copula-bias correction* bezeichnete Verfahren, soll nicht nur die intervariable Abhängigkeit, sondern auch die räumliche und zeitliche Abhängigkeit reduzieren. Für das Verfahren wird eine (univariate) biaskorrigierte Zeitreihe als Referenz benötigt. Zuerst werden die Ränge dieser Zeitreihe bestimmt. Anschließend werden auch die Ränge einer anderen Zeitreihe derselben Größe ermittelt. Die Ränge der zweiten, unkorrigierten Zeitreihe werden nun so geordnet, dass sie der Reihenfolge der Ränge der Referenz-Zeitreihe entsprechen. Korrigiert wird somit die Struktur der Ränge einer Variablen (vgl. Vrac und Friederichs, 2015).

Radarbezogene Methoden

Mean field bias correction: Dieses Verfahren ist das wohl einfachst denkbare. Aus dem Verhältnis zwischen Radar- und Stationsdaten wird ein Korrekturfaktor berechnet, der auf das ganze Radarfeld angewandt wird. Dabei besteht die Annahme, dass der Bias räumlich konstant ist (vgl. Goudenhoofd und Delobbe, 2009; McKee und Binns, 2016; Thorndahl et al., 2017; H. Chen et al., 2019). Letztlich entspricht diese Methode dem *linear scaling*. Der Biaskorrekturfaktor kann auf verschiedenen zeitlichen Skalen berechnet werden: z. B. auf täglicher (vgl. Smith et al., 2012) oder stündlicher Basis (vgl. Tesfagiorgis et al., 2011; Hanchoo Wong et al., 2012). Der stündliche mean field bias wird oftmals in Echtzeit-Anwendungen genutzt, meist mithilfe des Konzepts des Kalman-Filters (vgl. Smith und Krajewski, 1991; D.-J. Seo et al., 1999; D.-J. Seo und Breidenbach, 2002; Chumchean et al., 2006; Ozkaya und Akyurek, 2019). Als Alternative zur Berechnung pro Zeiteinheit, können auch alle Werte-Paare der Radar- und Stationsniederschläge gegeneinander aufgetragen und daran eine lineare Regression gefittet werden, um so eine Transferfunktion zu erhalten (vgl. Yoo et al., 2014).

Range dependent mean field bias correction: Diese einfache Erweiterung der *mean field bias correction* basiert auf Gruppenbildung. Hierbei werden die stündlichen Daten in verschiedene Gruppen mit unterschiedlicher Entfernung zum Radar eingeteilt und anschließend für jede Gruppe ein mittlerer Korrekturfaktor bestimmt (vgl. Hanchoo Wong et al., 2012).

Local bias correction: Bei dieser Methode von Hanchoo Wong et al., 2012 werden die stündlichen Stationsdaten einer gewissen Anzahl an Jahren gemittelt und diese Mittel mithilfe von Kriging räumlich interpoliert. Anschließend werden Klassen gebildet und für jede Klasse mit derselben mittleren Niederschlagsmenge ein Korrekturfaktor aus dem Verhältnis von Stations- und Radarniederschlägen bestimmt. Im Gegensatz zur Methode der *range dependent mean field correction* werden die Gruppen also nicht entsprechend der Entfernung zum Radar gewählt, sondern gemäß des langjährigen mittleren Niederschlags.

Static local bias correction and range dependent adjustment: Hier wird zuerst der *mean field bias* korrigiert und anschließend für jede Station das residuale/übriggebliebene jährliche gemittelte Bias ermitteln. Die daraus abgeleiteten Korrekturfaktoren werden dann mithilfe von Kriging räumlich interpoliert (vgl. Goudenhoofd und Delobbe, 2009).

Brandes spatial adjustment: Auch dieses Korrekturverfahren, geht von einem räumlich variablen Bias aus. Für jede Niederschlagsstation wird ein Korrekturfaktor aus dem Verhältnis von Stations- und Radardaten bestimmt. Auf diese wird ein Distanz-gewichtetes Interpolationsverfahren angewandt, um Korrekturfaktoren für das gesamte Radarfeld zu erhalten (vgl. Goudenhoofd und Delobbe, 2009; Overeem et al., 2009b; McKee und Binns, 2016).

Local bias correction with ordinary kriging: Diese Methode ähnelt *Brandes spatial adjustment*. Als räumliches Interpolationsverfahren wird jedoch ordinary kriging gewählt. Grundlage dafür ist das Model der räumlichen Kovarianz und deren Darstellung als Semivariogramm, aus dem die Gewichtungsfaktoren bestimmt werden (vgl. Goudenhoofd und Delobbe, 2009; McKee und Binns, 2016). Vergleichbar sind die Verfahren von Zhang et al., 2011 und H. Chen et al., 2019.

Range-dependent bias correction: Das Verfahren der *range-dependent bias correction* nimmt an, dass das Bias der Radardaten eine Funktion der Distanz vom Radarturm ist. Damit betont es das Problem des Overshootings. Der Korrekturfaktor wird auf der log-Skala berechnet und als Polynom zweiter Ordnung ausgedrückt (vgl. Goudenhoofd und Delobbe, 2009; McKee und Binns, 2016).

Weighted multiple regression model: Diese Methode von Gilewski und Nawalany, 2018, die auf einer linearen Regression basiert, ist ein vergleichbares Verfahren, berücksichtigt neben der Distanz zwischen Radarturm und Niederschlagsstation aber auch die minimale Höhe des Radarstrahls über dieser Station und die Höhe der Station in Meter über NHN.

Conditional mean field bias correction: Auch dieses bivariate Verfahren nimmt eine systematische, entfernungsabhängige Biasfunktion an. Dabei wird das Bias an einer Station als abhängig vom entsprechenden Radarwert an diesem Ort angesehen. Zur Ermittlung des bedingten Erwartungswertes kann als nicht-parametrisches Verfahren ein gleitender Mittelwert angewandt werden. Als parametrisches Verfahren kann die Copula-Regression, mit den Radar- und Stationsdaten als Randverteilungen, genutzt werden, um deren Zusammenhang zu bestimmen (vgl. Thorndahl et al., 2017; Villarini et al., 2008).

Bayesian data combination: Bei der Methode der *Bayesian data combination* werden die Niederschlagsstationen zuerst mithilfe von Block-kriging räumlich interpoliert und pro Gitterpunkt die Differenz zwischen Radarfeld und interpolierten Stationen berechnet. Anschließend wird das Bias-Feld mit einem experimentellen Variogramm gefittet, um ein geglättetes Bias-Feld zu erhalten. Innerhalb eines Bayesschen Ansatzes wird zuletzt ein Kalman-Filter angewandt, um die interpolierten Stationswerte mit dem modellierten Bias-Variogramm zu kombinieren (vgl. McKee und Binns, 2016). Eine andere Methode, die auch auf dem Bayesschen Ansatz beruht, ist in (H. Chen et al., 2019) beschrieben.

Conditional merging: Dieses Verfahren, auch *kriging with radar correction* genannt, beruht auf der Annahme, dass die Radarmessungen ein räumlich korrektes Feld von unbekanntem Niederschlagswerten und die Stationsmessungen ein räumlich unbekanntes Feld von wahren Niederschlagswerten widerspiegeln. Bei dieser Methode werden die Stationsdaten über ein räumliches Interpolationsverfahren (ordinary kriging) in das Radarfeld gefittet. Anschließend werden die Radarwerte über den Stationen identifiziert und räumlich interpoliert. Die Differenz zwischen interpoliertem Radarfeld und originalem Radarfeld ergibt ein Bias-Korrektur-Feld. Dieses wird schließlich auf das interpolierte Stationsfeld addiert (vgl. Goudenhoofd und Delobbe, 2009; Rabiei und Haberlandt, 2015; McKee und Binns, 2016).

Kriging with external drift: Bei dieser Korrekturmethode werden Radardaten als Hilfsgröße betrachtet. Dieses Interpolationsverfahren ähnelt dem ordinary kriging, das im Verfahren genutzte Mittel ist hier aber eine lineare Funktion des Radarfelds. Der addierte Term des zufälligen Fehlers entspricht hier den regionalisierten Stationsdaten. Der restliche lineare Term entspricht dem deterministischen Teil des Modells, in dem die Radardaten verwendet werden (vgl. Goudenhoofdt und Delobbe, 2009; Goudenhoofdt und Delobbe, 2016; Rabiei und Haberlandt, 2015; McKee und Binns, 2016).

Statistical objective analysis: Hier werden Radardaten als Hintergrundfeld betrachtet, das mithilfe der Messungen korrigiert wird. Für jede Station wird die Differenz zwischen Stations- und korrespondierendem Radarwert berechnet und ein Gewichtungsfaktor multipliziert. Dann wird für jeden Gitterpunkt des Radarfelds die Summe der gewichteten Differenzen zum Radarwert addiert. Die Gewichtungsfaktoren werden so bestimmt, dass die Varianz der Residuen des korrigierten Radarfeldes minimal wird (vgl. McKee und Binns, 2016).

Abkürzungsverzeichnis

DWD Deutscher Wetterdienst

D10 10-minütige Dauerstufe (Niederschlagssumme pro 10 Minuten)

D60 60-minütige Dauerstufe (Niederschlagssumme pro 60 Minuten)

D180 180-minütige Dauerstufe (Niederschlagssumme pro 180 Minuten)

D360 360-minütige Dauerstufe (Niederschlagssumme pro 360 Minuten)

Stat-Px Radar-Rasterzelle, die den Stationsstandort enthält

Mean4 Mittel der vier dem Stationsstandort nächsten Radar-Rasterzellen

Mean9 Radar-Rasterzelle des Stationsstandorts (Stat-Px) mit den acht umliegenden Radar-Rasterzellen

V1 Biaskorrektur-Verfahren 1: rangbasierte Korrektur der Events

V2 Biaskorrektur-Verfahren 2: gemeinsame Korrektur aller Events

Sensi-Test1 Sensitivitätstest 1: Test des Einflusses der Binanzahl (100 Bins statt 1000 Bins)

Sensi-Test2 Sensitivitätstest 2: Test des Einflusses des Radar-Rasters (Mean4 statt Stat-Px)

Sensi-Test3 Sensitivitätstest 3: Test des Einflusses der Jahreszeit (ausschließlich Sommererevents statt alle Events)

Sensi-Test4 Sensitivitätstest 4: Test des Einflusses des zeitlichen Zusammenhangs (ausschließlich übereinstimmende Events statt alle Events)

Tabelle 7.1: Koordinaten der Wetterradar-Standorte des Deutschen Wetterdienstes (nach Winterrath et al., 2017 und DWD, 2018)

Nr.	Standort	WMO-Nr.	Breitengrad [°] ¹	Längengrad [°] ¹	Höhe [m]
1	Emden-Knock	10204	53,3387	7,0238	58,0
2	Boostedt	10132	54,0044	10,0469	124,6
3	Dresden-Klotzsche	10488	51,1246	13,7686	263,4
4	Eisberg	10780	49,5407	12,4028	798,8
5	Essen-Bredeney	10410	51,4056	6,9671	185,1
6	Feldberg	10908	47,8736	8,0036	1516,1
7	Flechtendorf	10440	51,3112	8,802	628,0
8	Hannover	10339	52,4601	9,6945	97,7
9	Isen	10873	48,1747	12,1018	677,8
10	Memmingen	10950	48,0421	10,2192	724,4
11	Neuhaus am Rennweg	10557	50,5001	11,135	880,0
12	Neuheilenbach	10605	50,1097	6,5483	585,8
13	Offenthal	10629	49,9847	8,7129	245,8
14	Prötzel	10392	52,6487	13,8582	193,9
15	Rostock	10169	54,1757	12,0581	37,0
16	Türkheim	10832	48,5854	9,7827	767,6
17	Ummendorf	10356	52,1601	11,1761	185,0

¹ Alle Angaben zu Breiten- und Längengrad beziehen sich auf Referenzsystem WGS84.

Tabelle 7.2: Koeffizienten der Korrektur (nach Richter, 1995, S. 37, 67, verändert)

Nieder- schlagstyp	Temperatur [° C]	E	b			
			$H' = 2$ [1]	$H' = 5$ [2]	$H' = 9,5$ [3]	$H' = 16$ [4]
Regen (Sommer)	≥ 3.0	0,38	0,345	0,310	0,280	0,245
Regen (Winter)	≥ 3.0	0,46	0,340	0,280	0,240	0,190
Mischung	$-0,7 < T < 3.0$	0,55	0,535	0,390	0,305	0,185
Schnee	≤ -0.7	0,82	0,720	0,510	0,330	0,210

H' : Horizontabschirmung; Stationslage ¹ frei; ² leicht geschützt; ³ geschützt; ⁴ stark geschützt

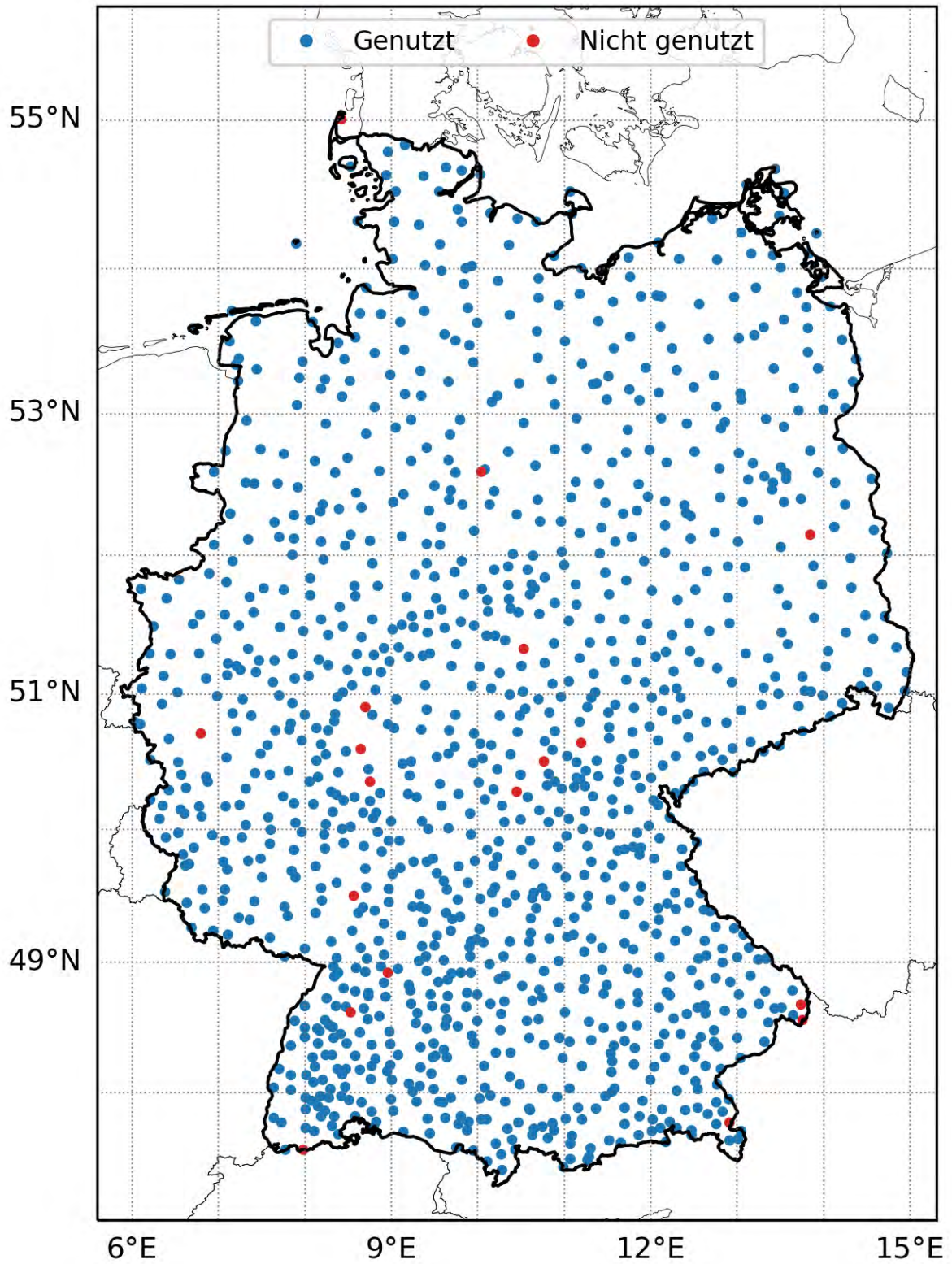


Abbildung 7.1: Karte der verwendeten und verworfenen Niederschlagstationen
 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

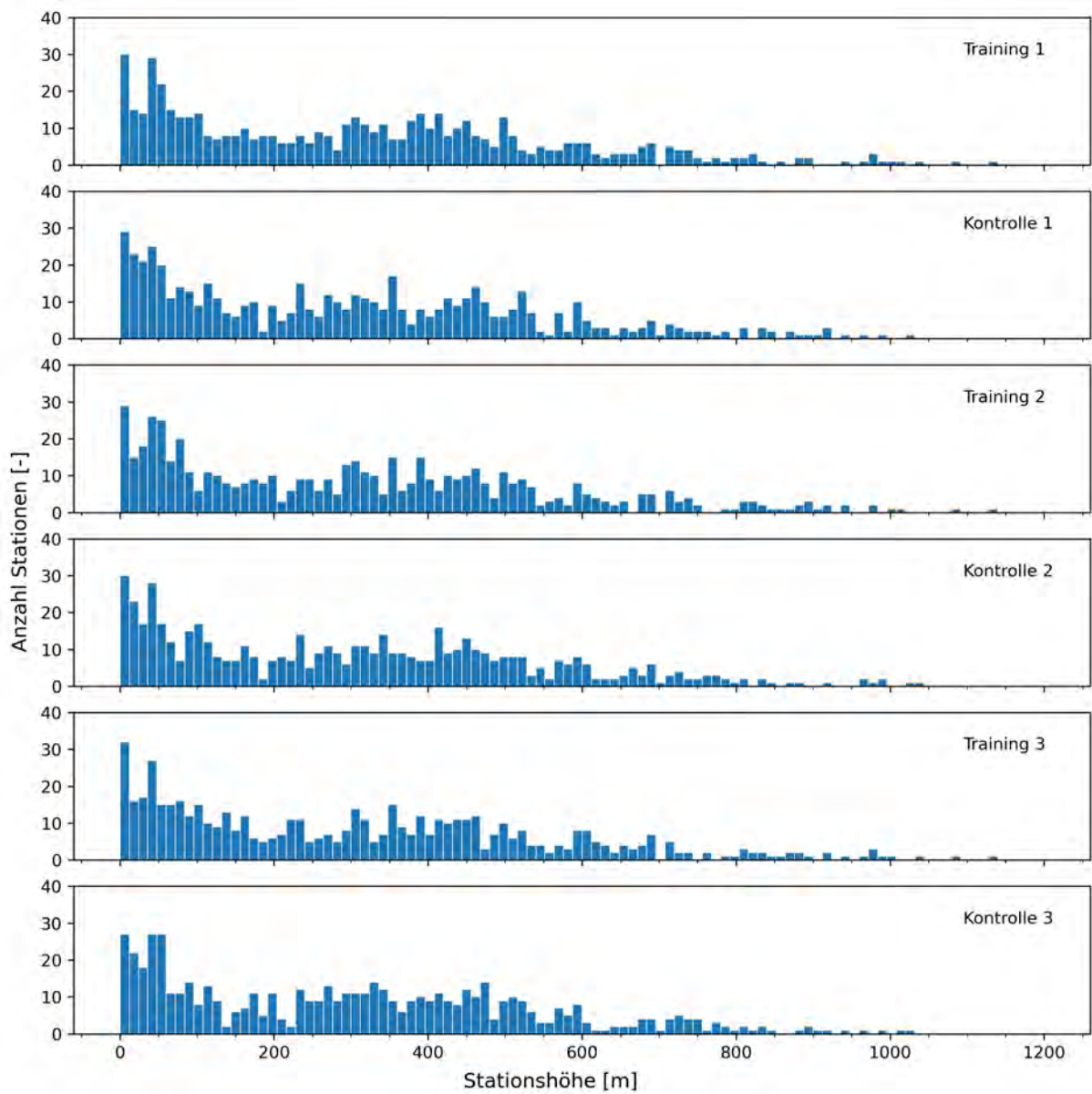


Abbildung 7.2: Höhenverteilung der drei Trainings- und Kontrollgruppen

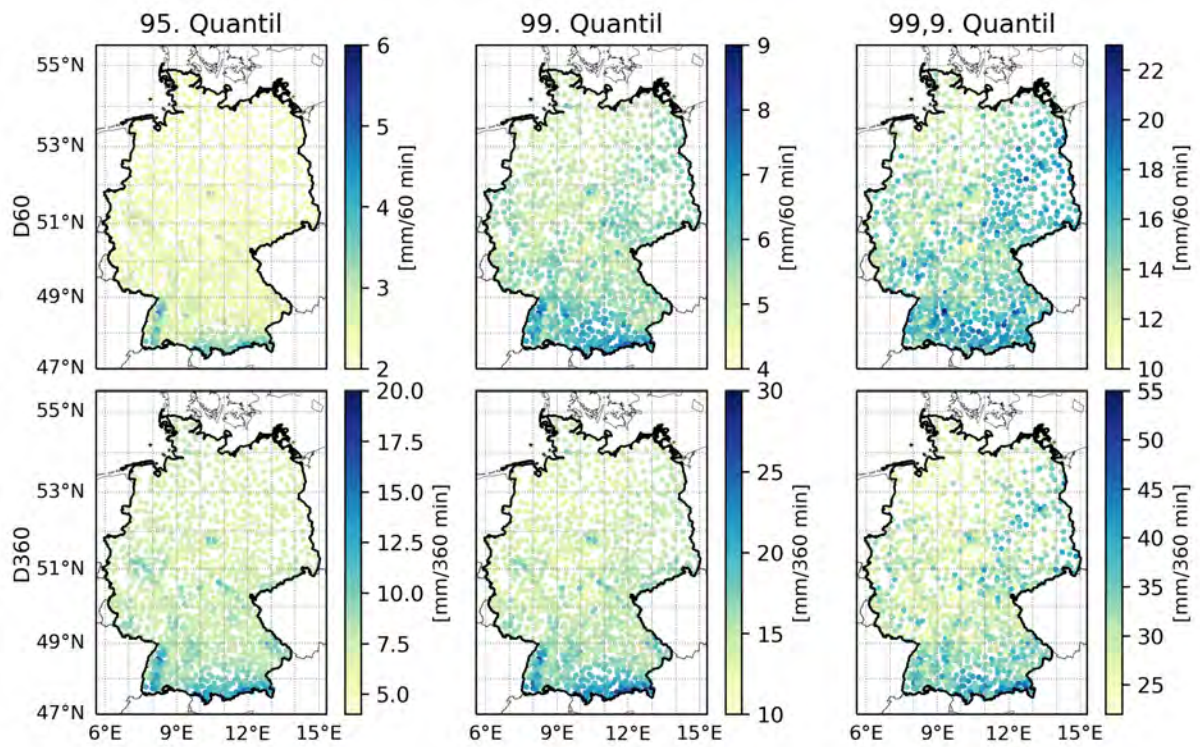


Abbildung 7.3: Niederschlagsintensitäten des 95., 99., und 99,9. Quantils
 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

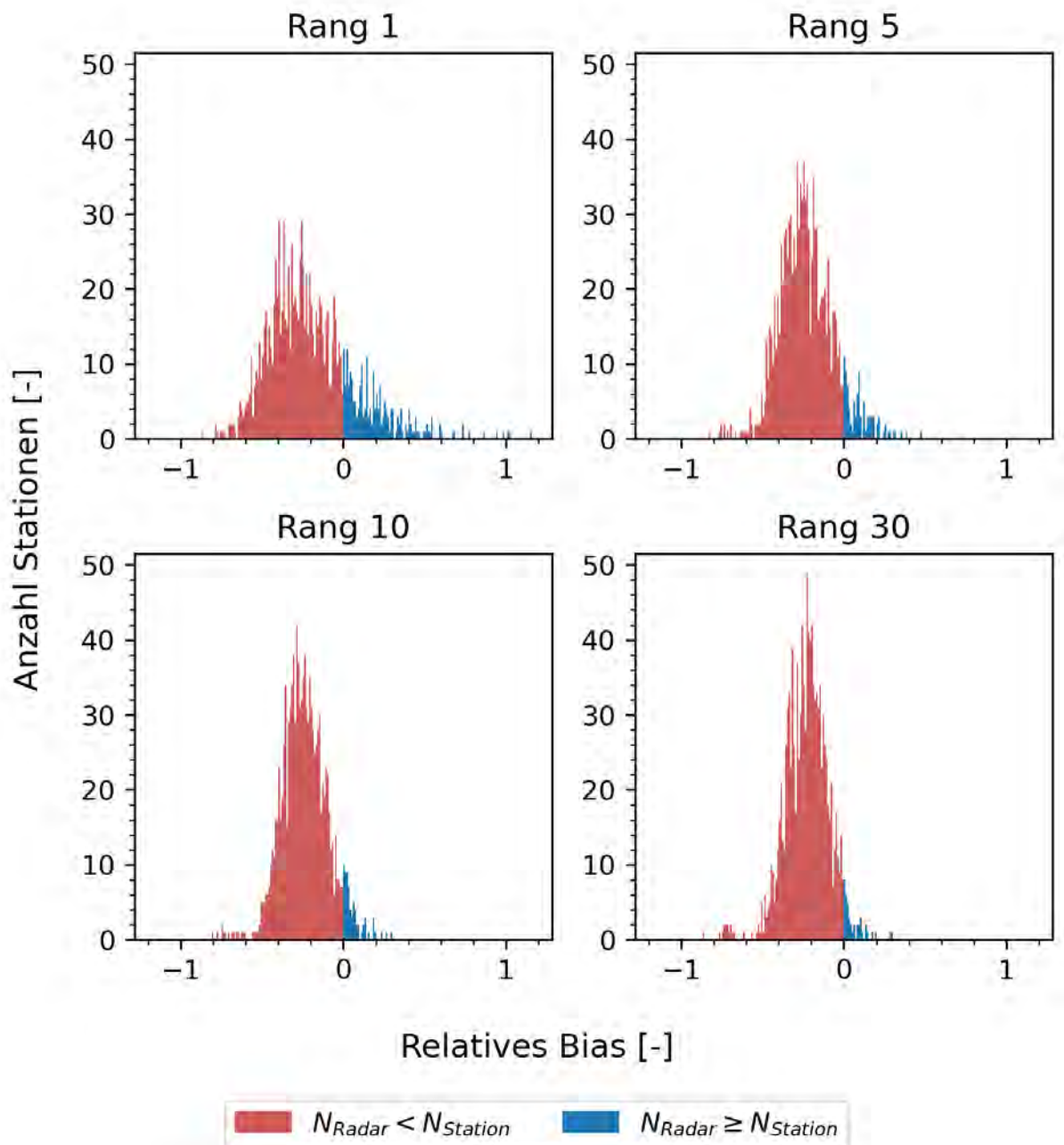


Abbildung 7.4: Histogramm des relativen Bias (Stat-Px) der Ränge 1, 5, 10 und 30 bezogen auf die Dauerstufe D10

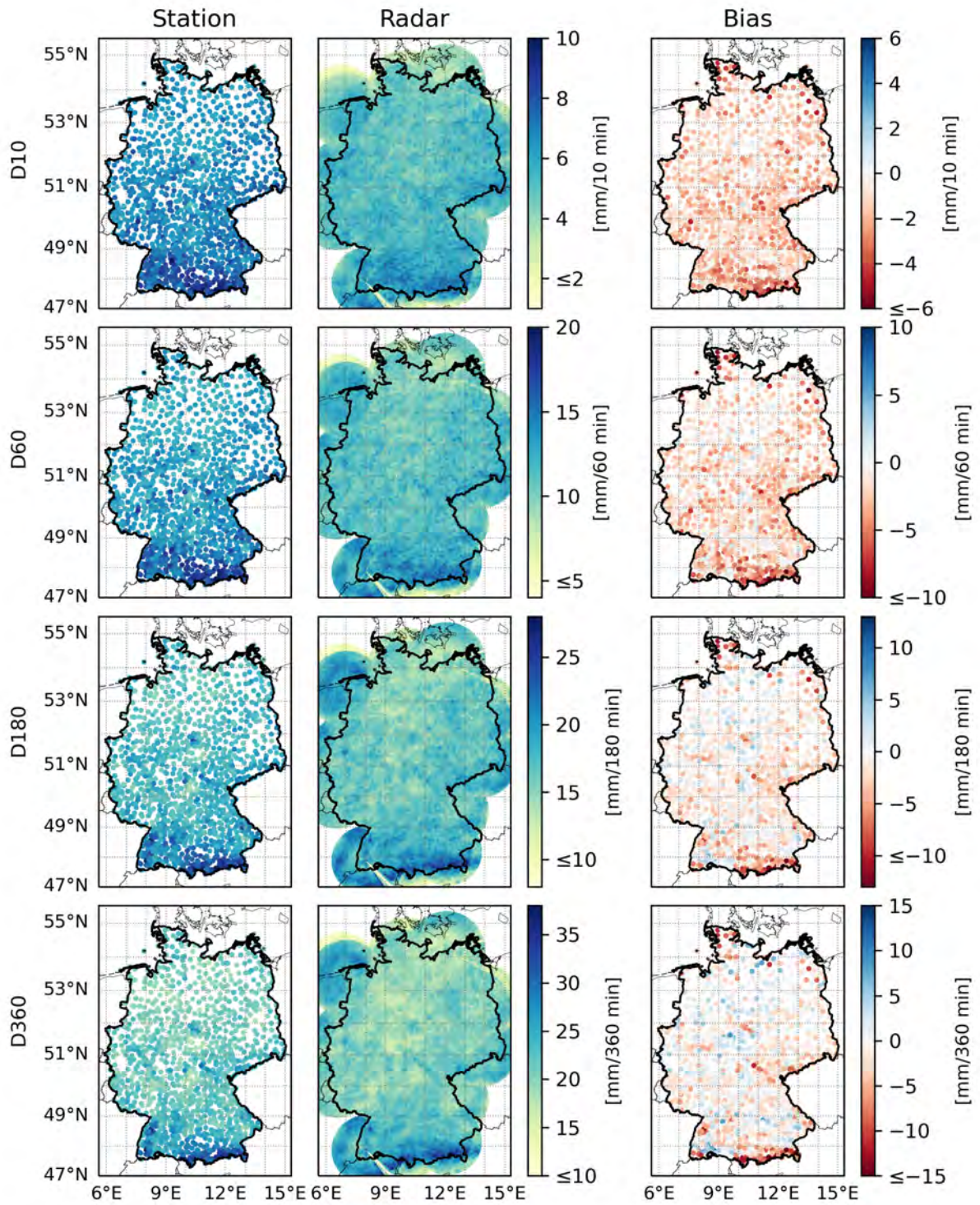


Abbildung 7.5: Räumliche Verteilung der Niederschläge und des Bias für den 30. Event-Rang (D10) (©EuroGeographics bezüglich der Verwaltungsgrenzen)

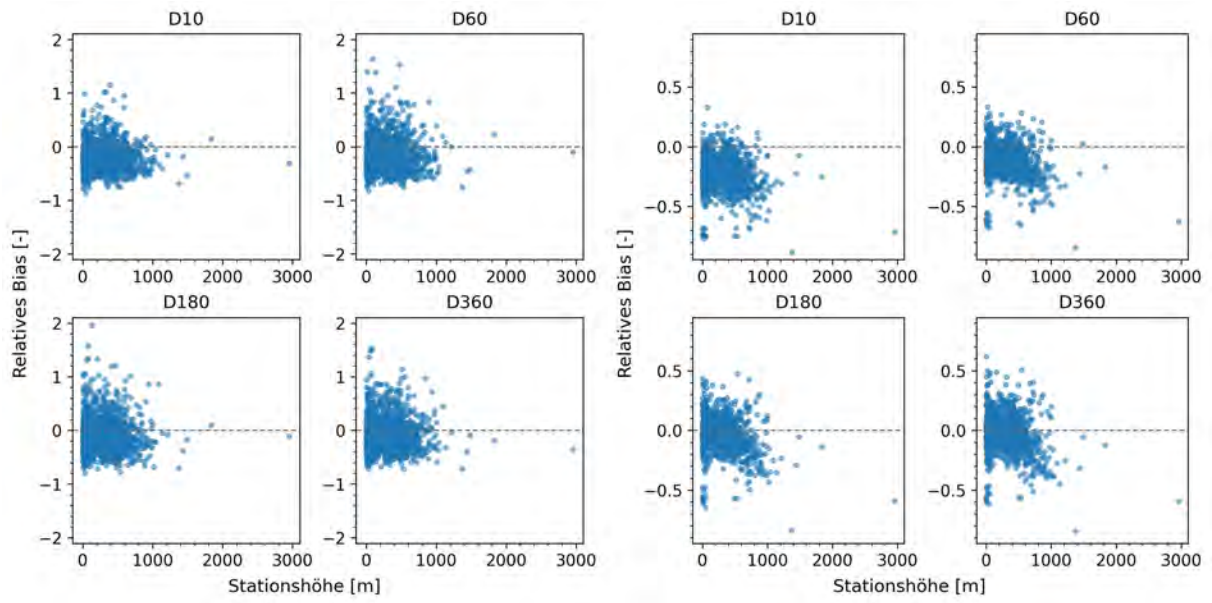


Abbildung 7.6: Zusammenhang zwischen relativem Bias und Stationshöhe. Dargestellt sind der erste (links) und 75. Event-Rang (rechts)

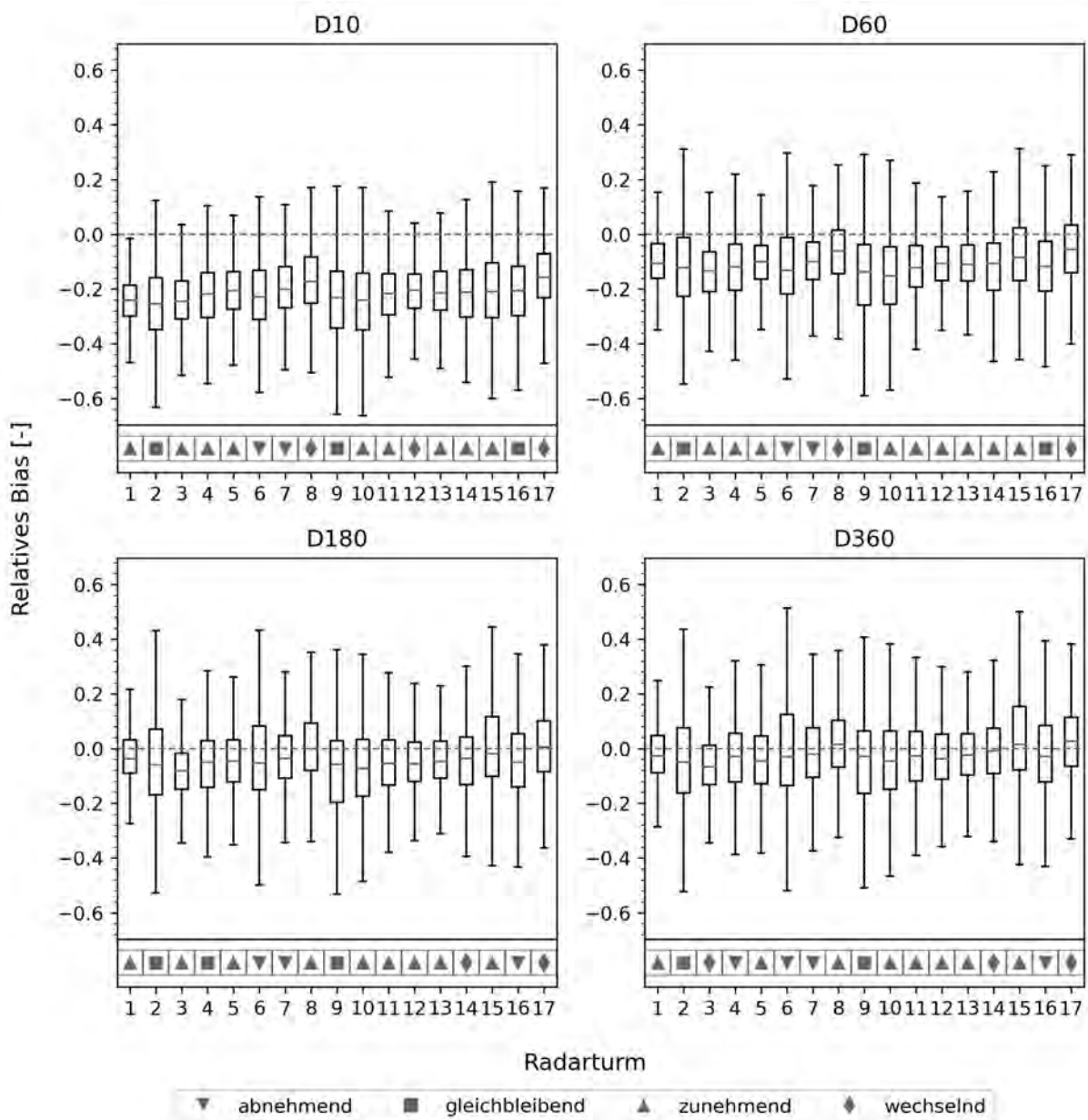


Abbildung 7.7: Zusammenhang zwischen dem relativen Bias und dem jeweiligen Radarturm. Die Symbole unterhalb der Grafik stellen Indikatoren dar, ob sich das Bias mit zunehmender Entfernung verändert. Zur Erläuterung der Radar-Nr. siehe Tabelle 7.1

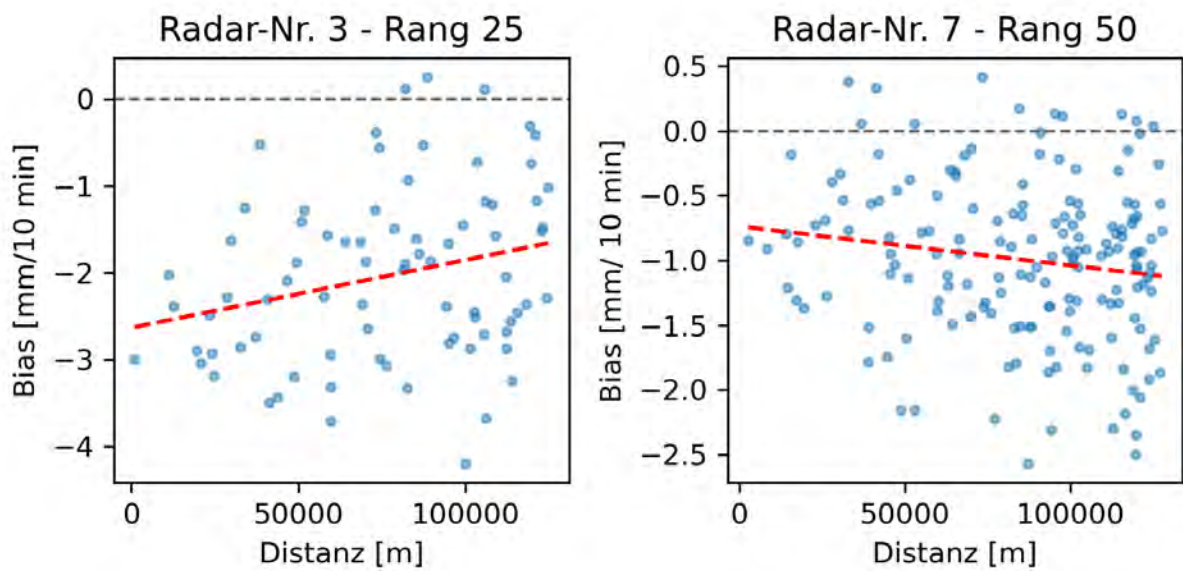


Abbildung 7.8: Beispiel für einen zunehmenden bzw. abnehmenden Trend mit größer werdender Entfernung vom Radarstandort. Man beachte jedoch die große Streuung.

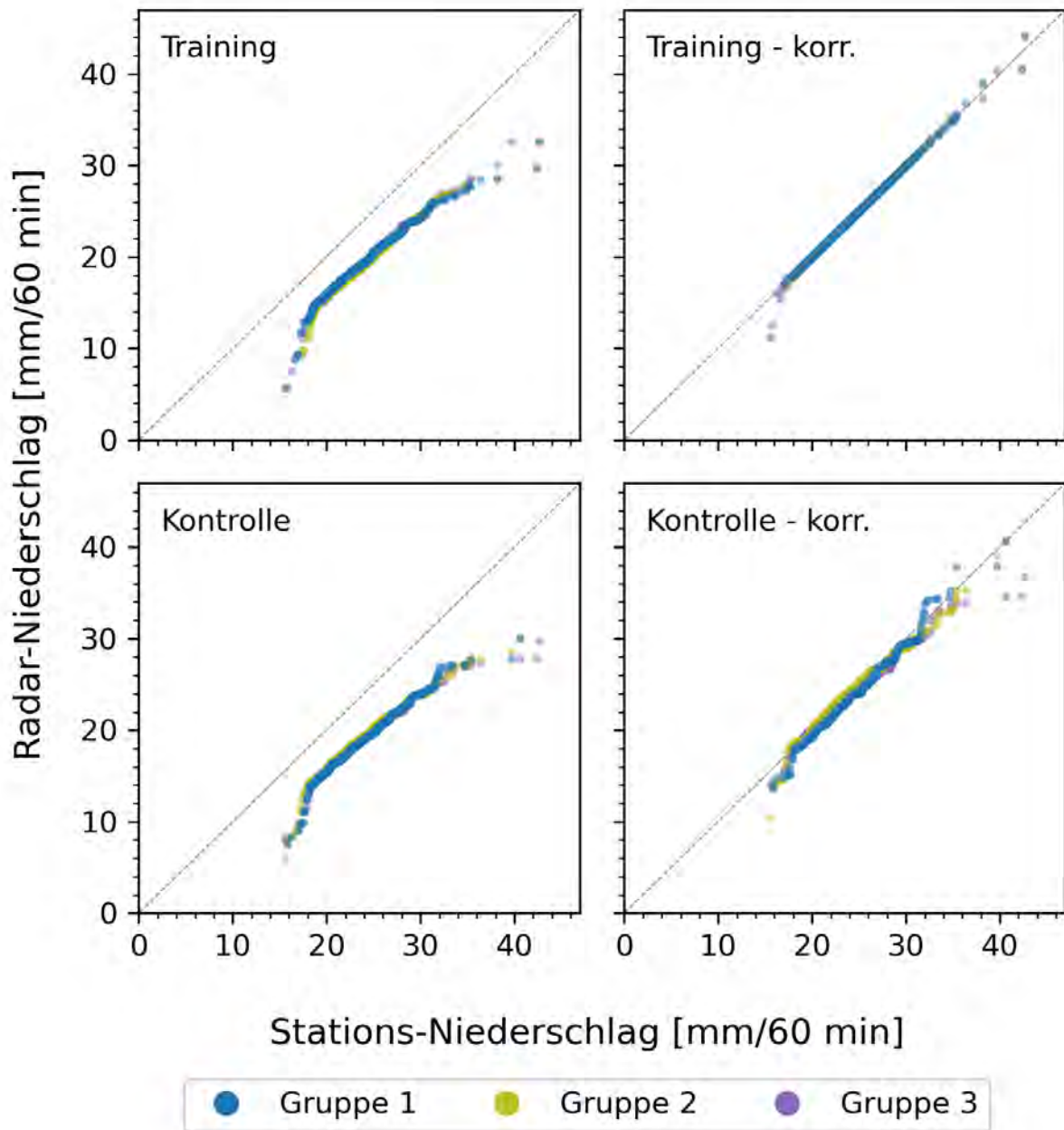


Abbildung 7.9: Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D60)

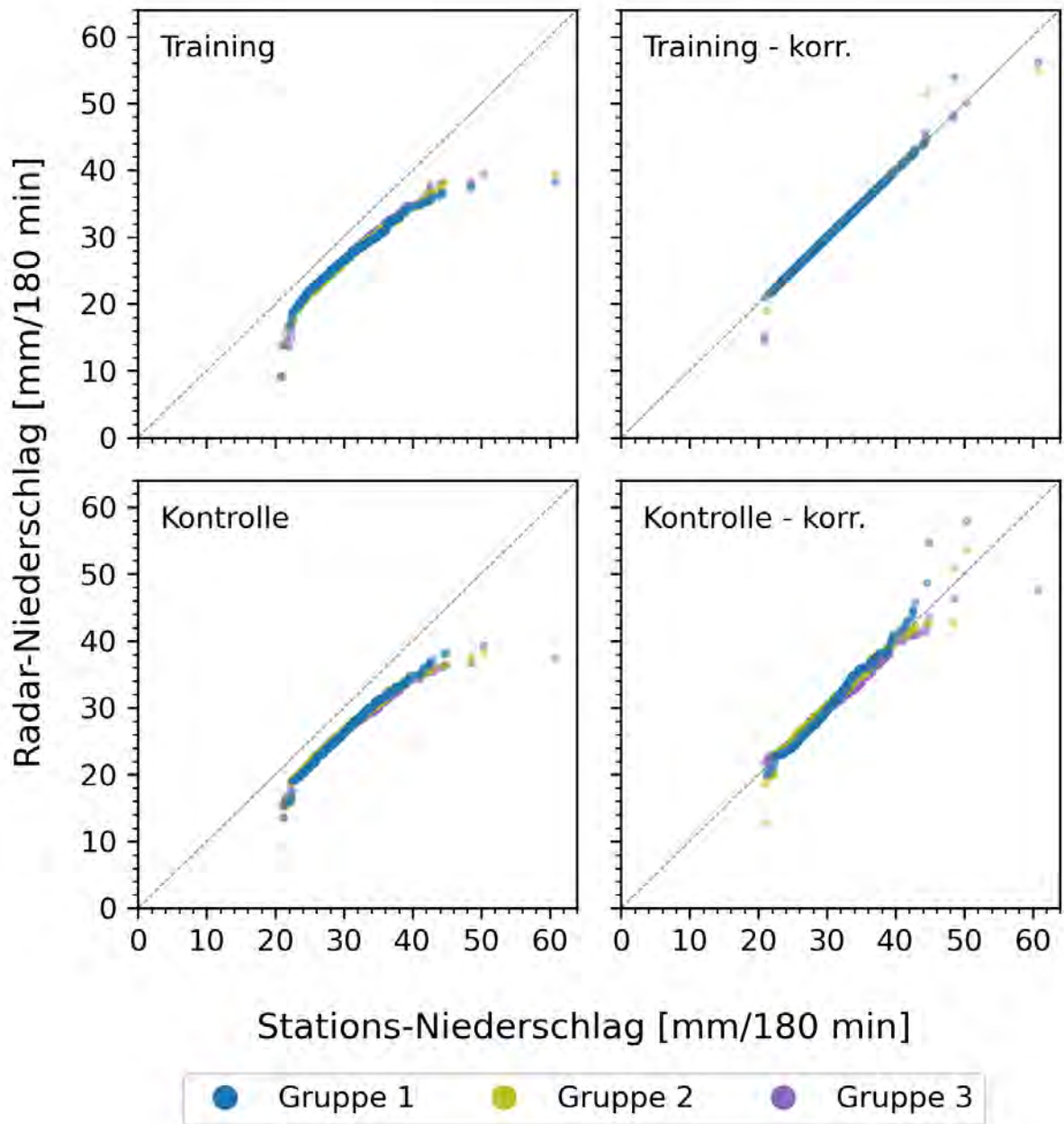


Abbildung 7.10: Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D180)

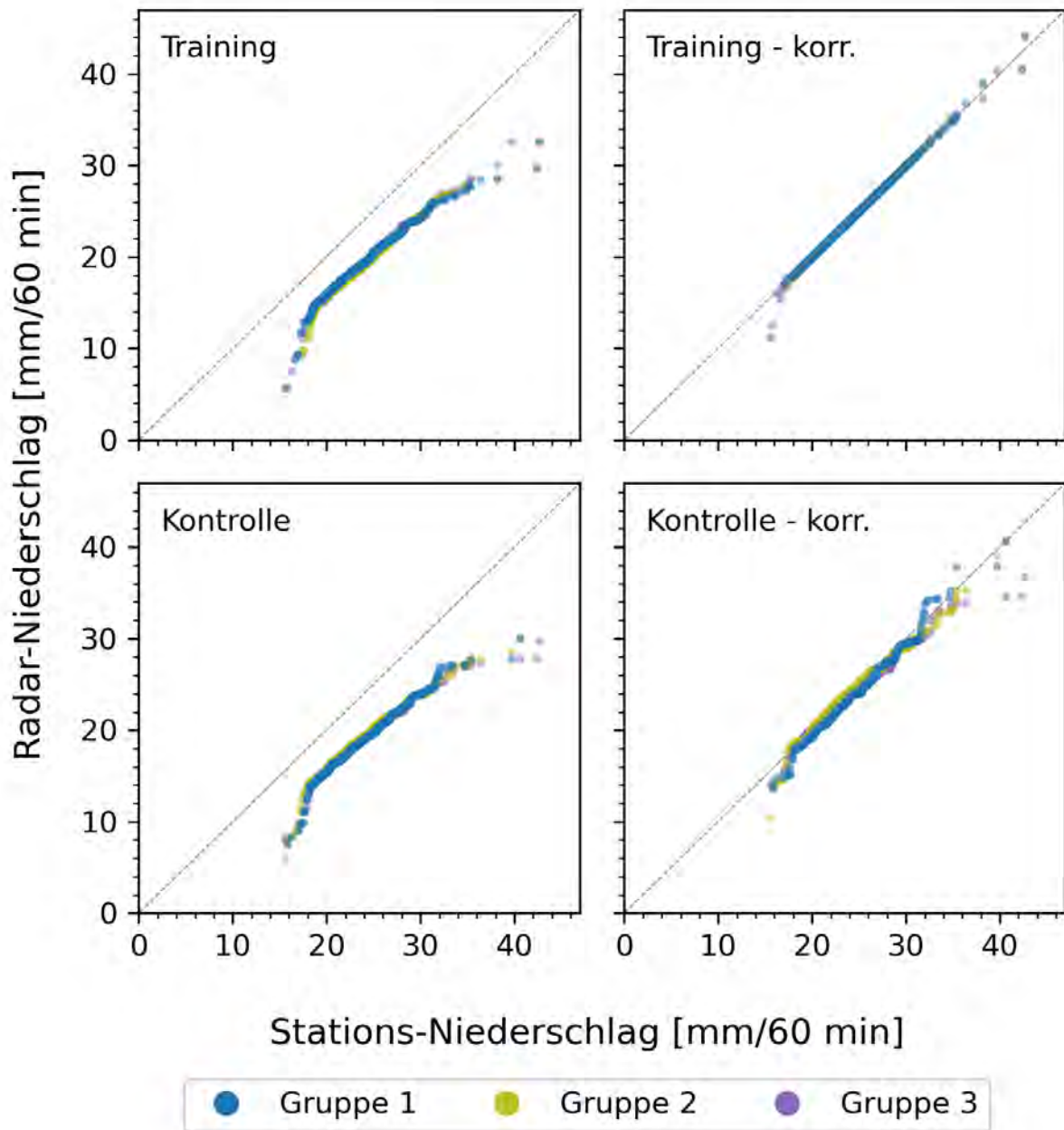


Abbildung 7.11: Verfahren 1: Q-Q-Plot für den siebten Event-Rang (D360)

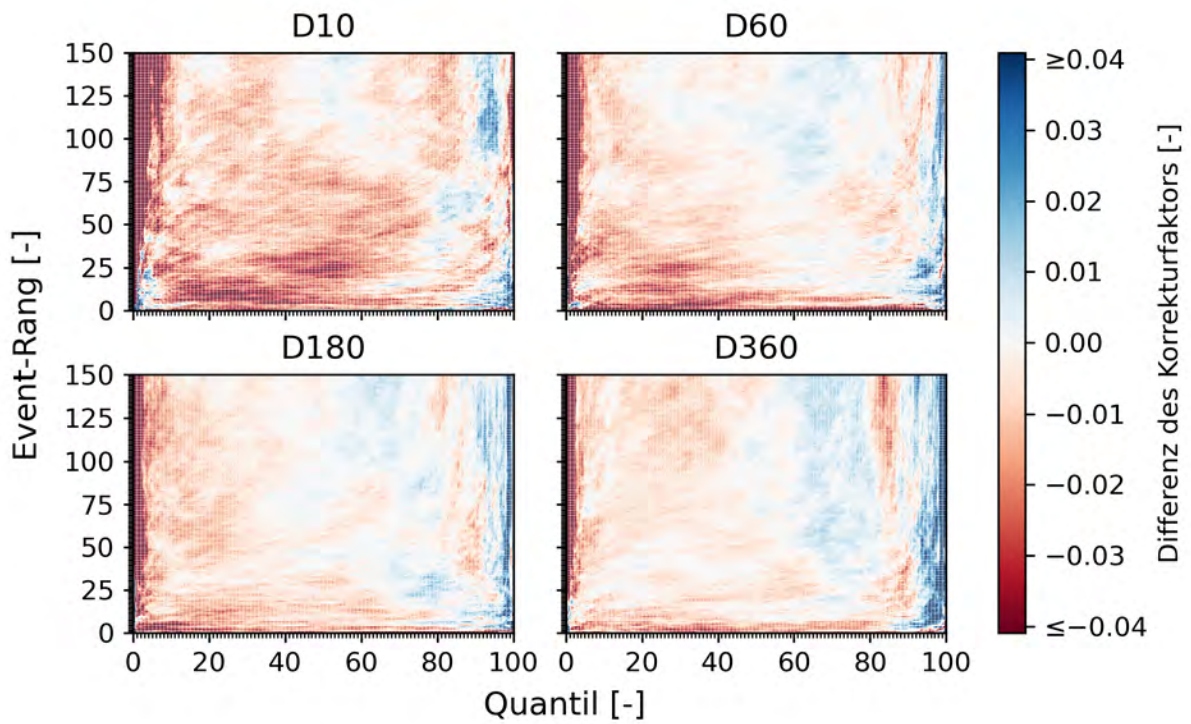


Abbildung 7.12: Verfahren 1: Differenzen der Korrekturfaktoren für die zweite Gruppe

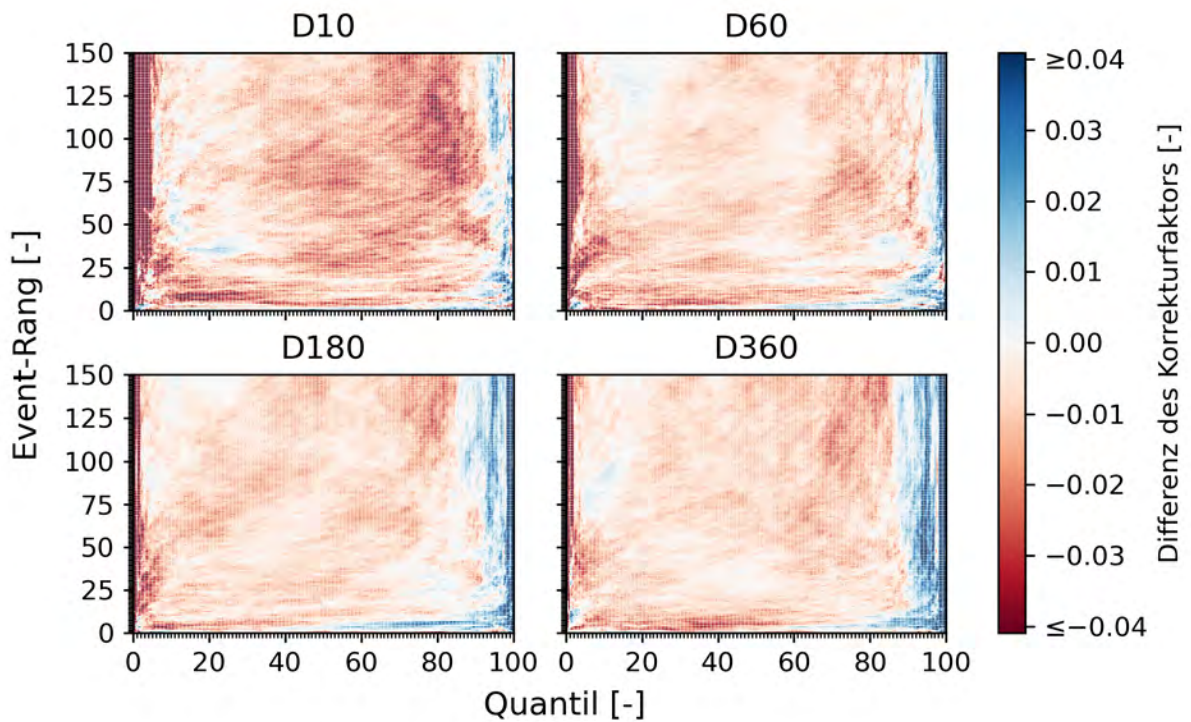


Abbildung 7.13: Verfahren 1: Differenzen der Korrekturfaktoren für die dritte Gruppe

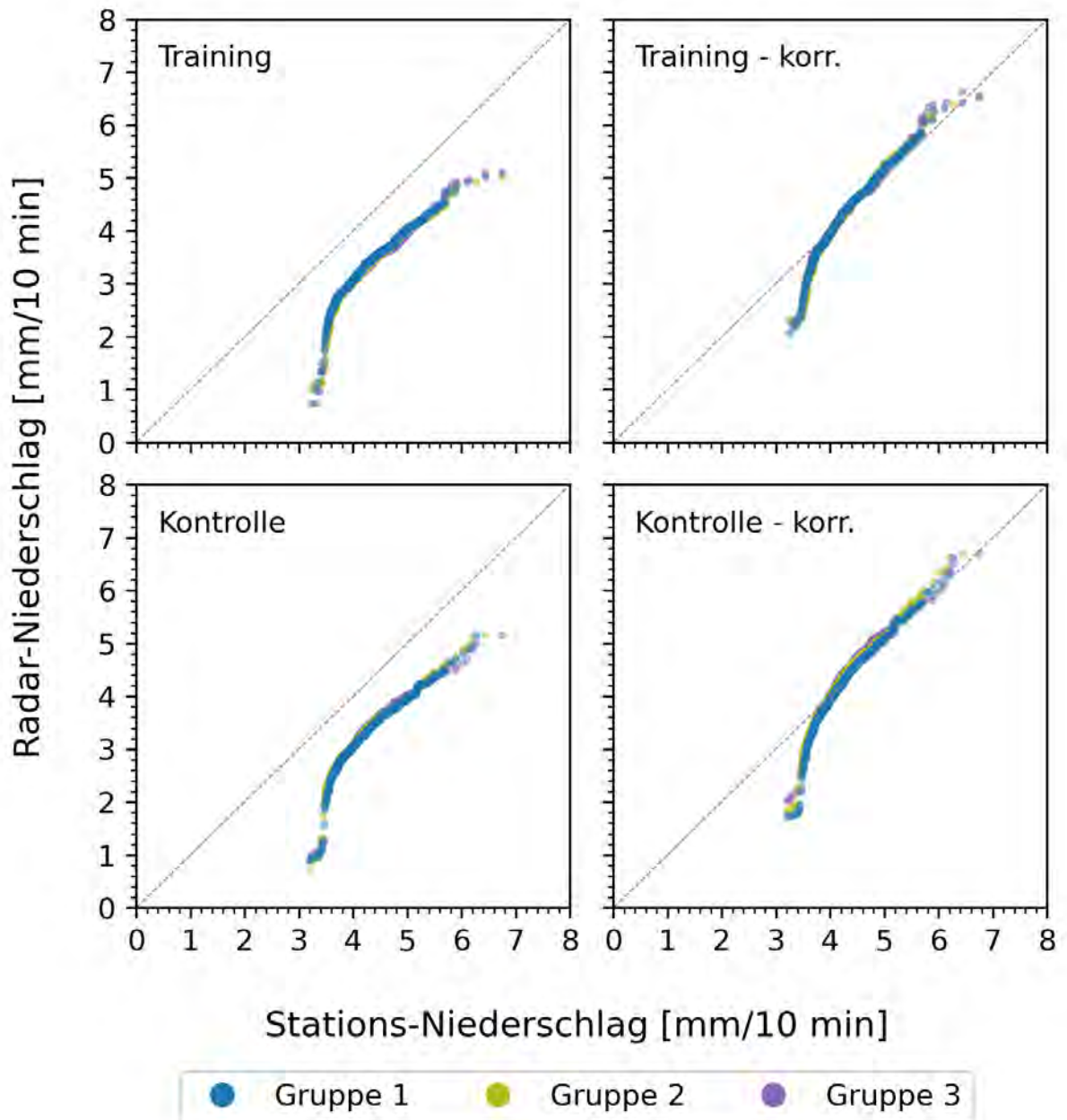


Abbildung 7.14: Verfahren 2: Q-Q-Plot für den 75. Event-Rang (D10)

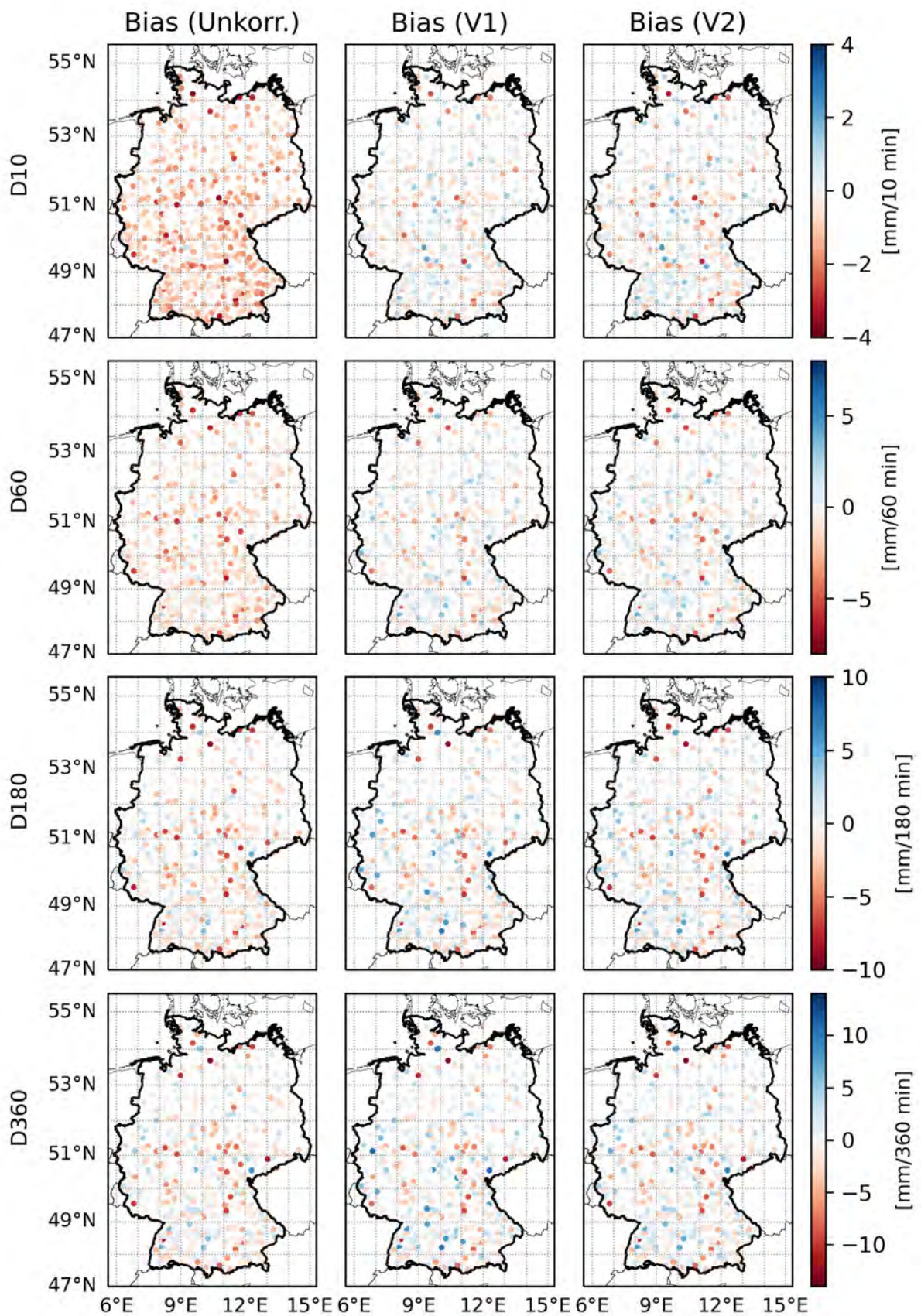


Abbildung 7.15: Vergleich des verbleibenden Bias beider Verfahren für die Kontrollgruppe bei Rang 75 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

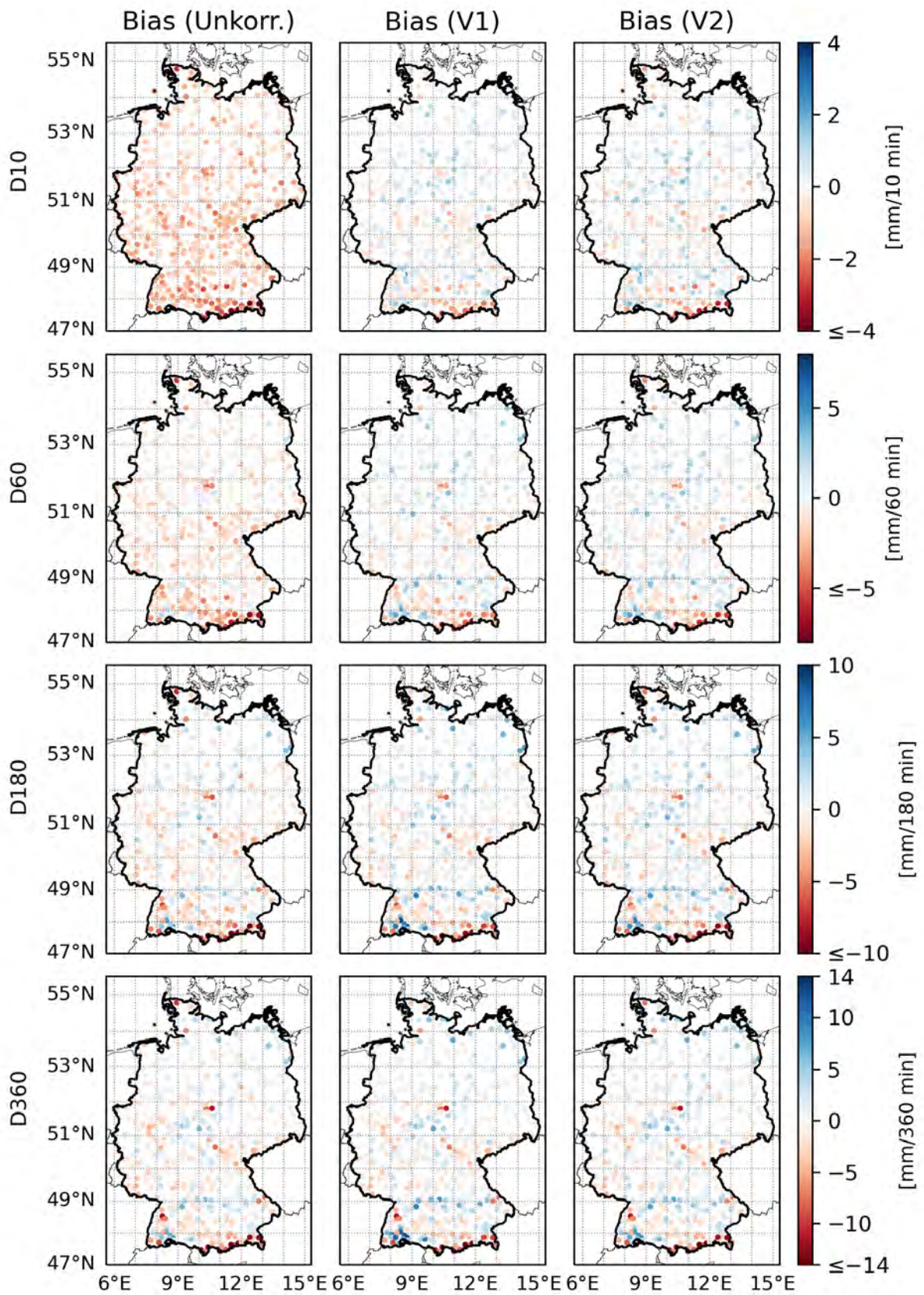


Abbildung 7.16: Vergleich des verbleibenden Bias beider Verfahren für die Trainingsgruppe bei Rang 75 (©EuroGeographics bezüglich der Verwaltungsgrenzen)

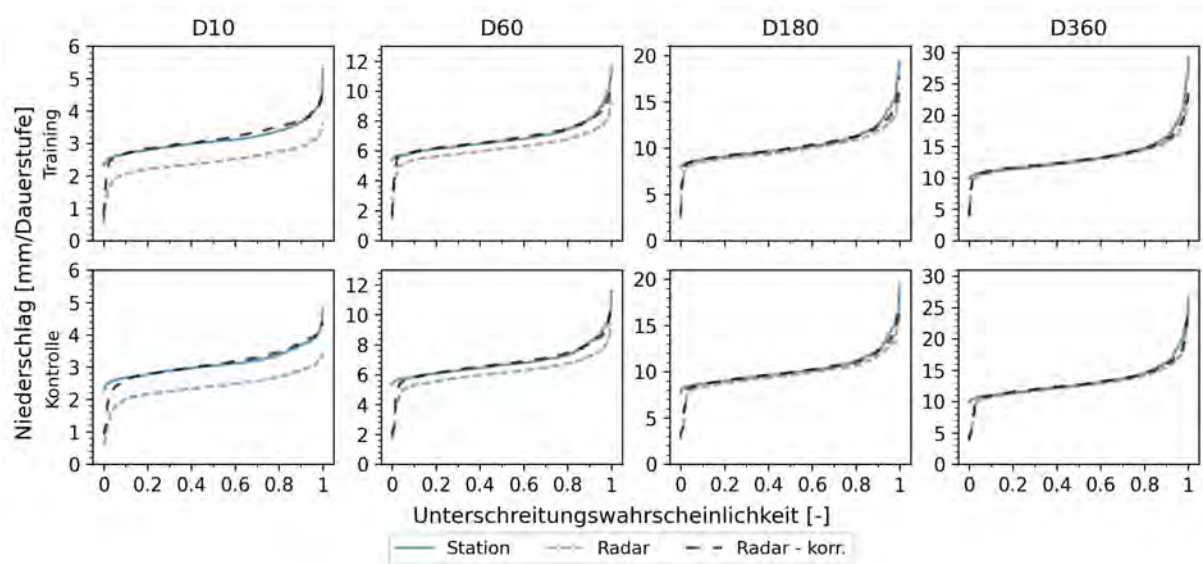


Abbildung 7.17: Sensitivitätstest 2: ECDFs aller Dauerstufen für Rang 125

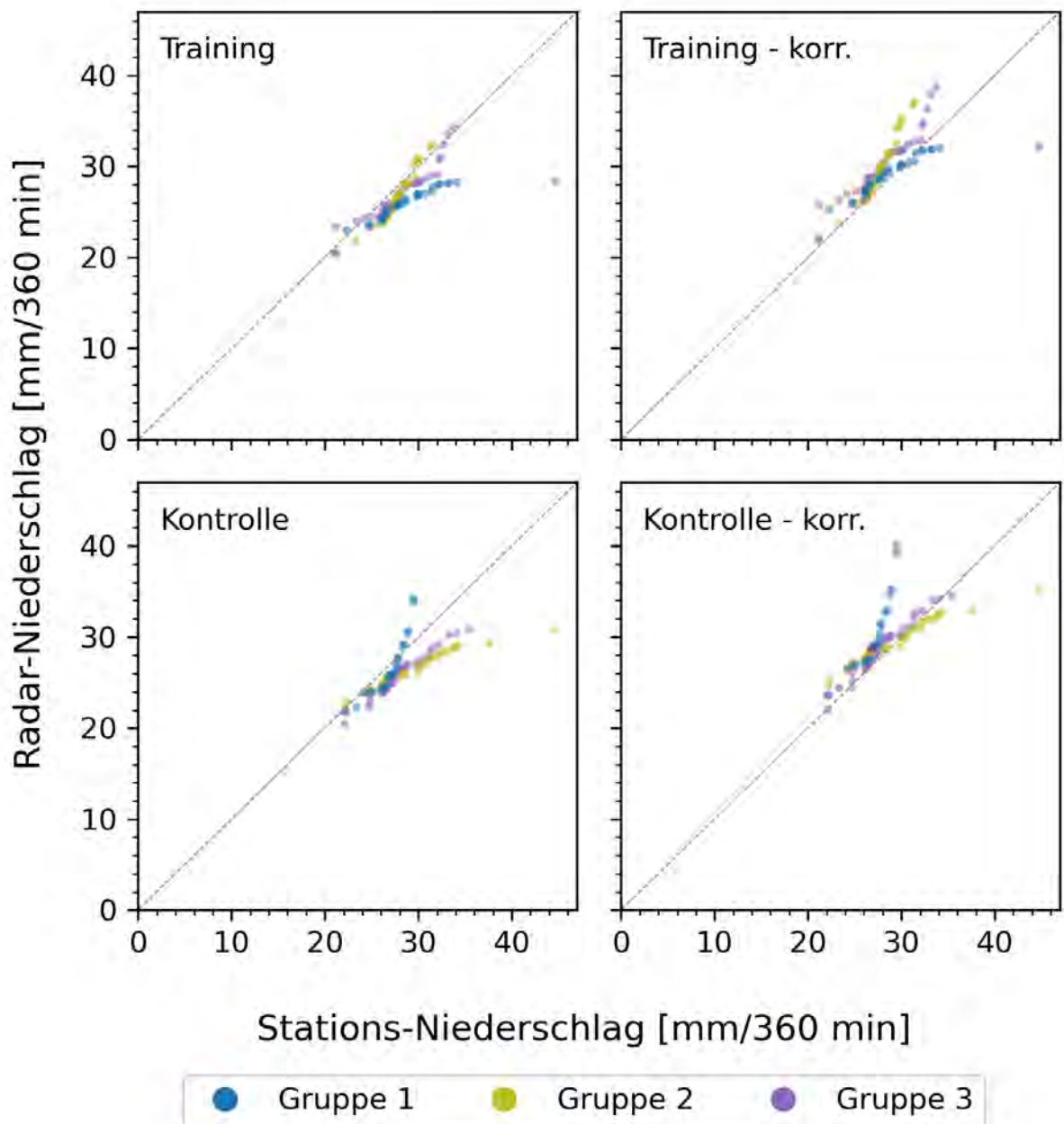


Abbildung 7.18: Sensitivitätstest 4: Q-Q-Plot für den 15. Event-Rang (D360). Die geringe Performance ist eindeutig sichtbar.

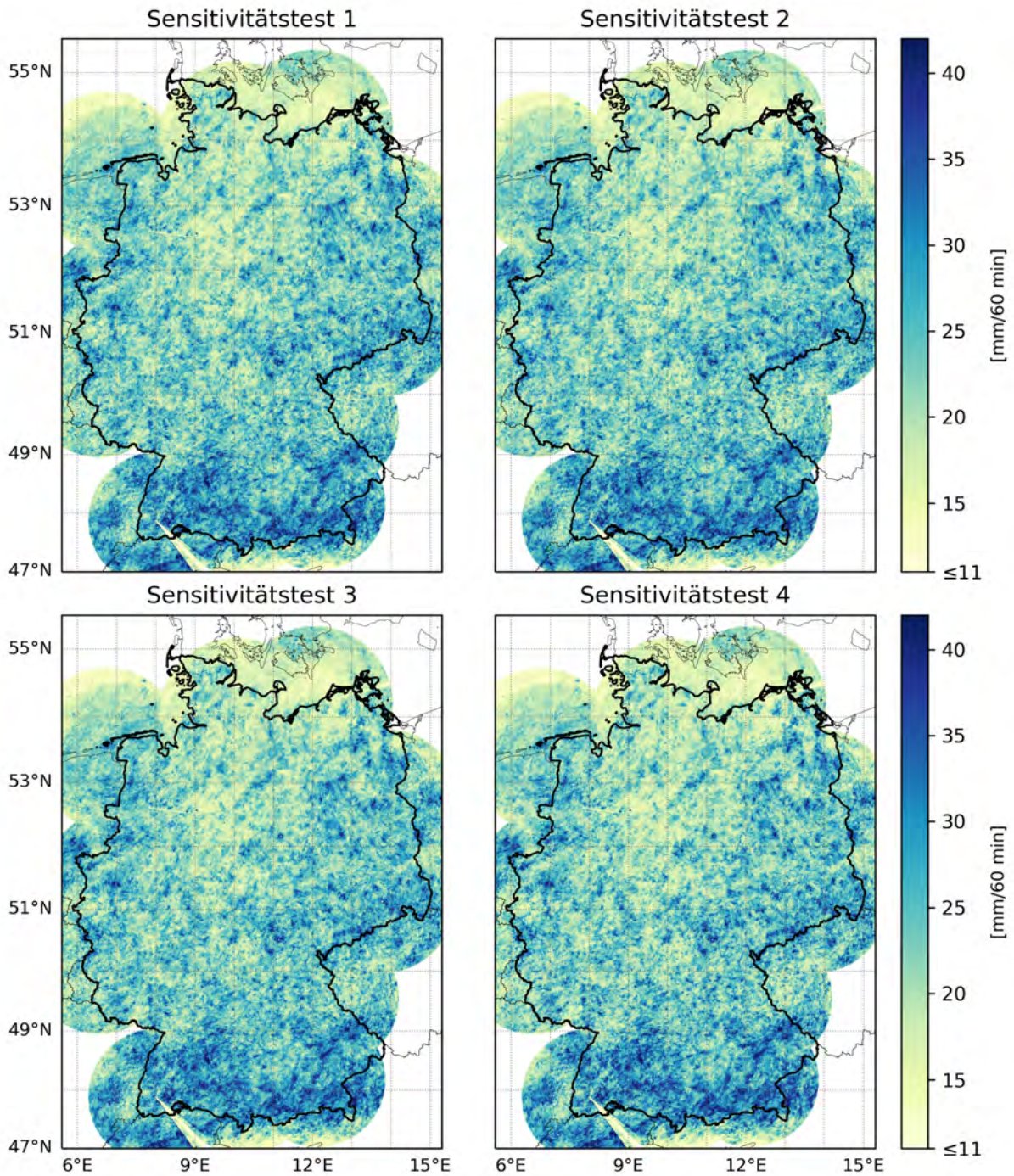


Abbildung 7.19: Vergleich der Sensitivitätstests bezüglich ihrer räumlichen Performance für Rang 5 (D60) (©EuroGeographics bezüglich der Verwaltungsgrenzen)

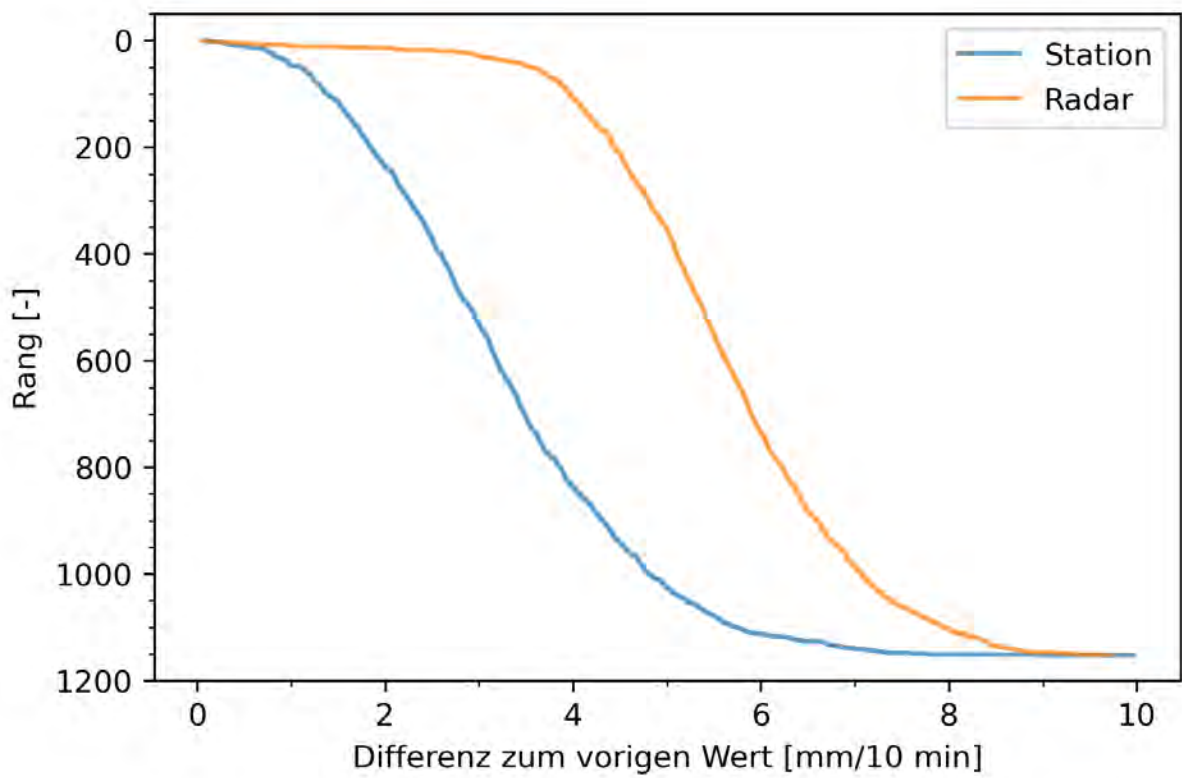


Abbildung 7.20: Kumulierte Differenzen des ECDF-Wertes zum vorigen für Rang 10 (D10)

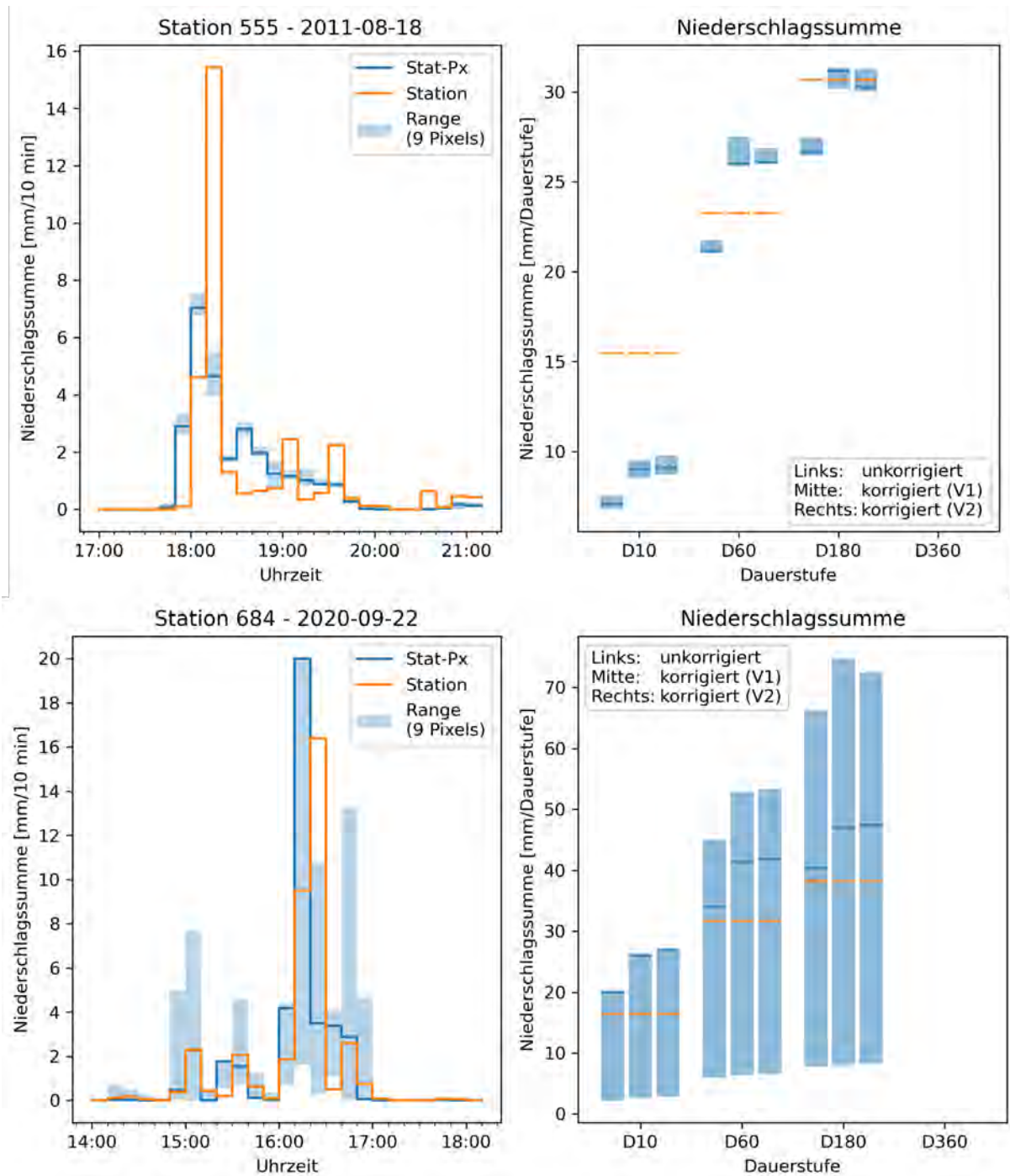


Abbildung 7.21: Korrektur der realen Ereignisse bei Station 555 und 684

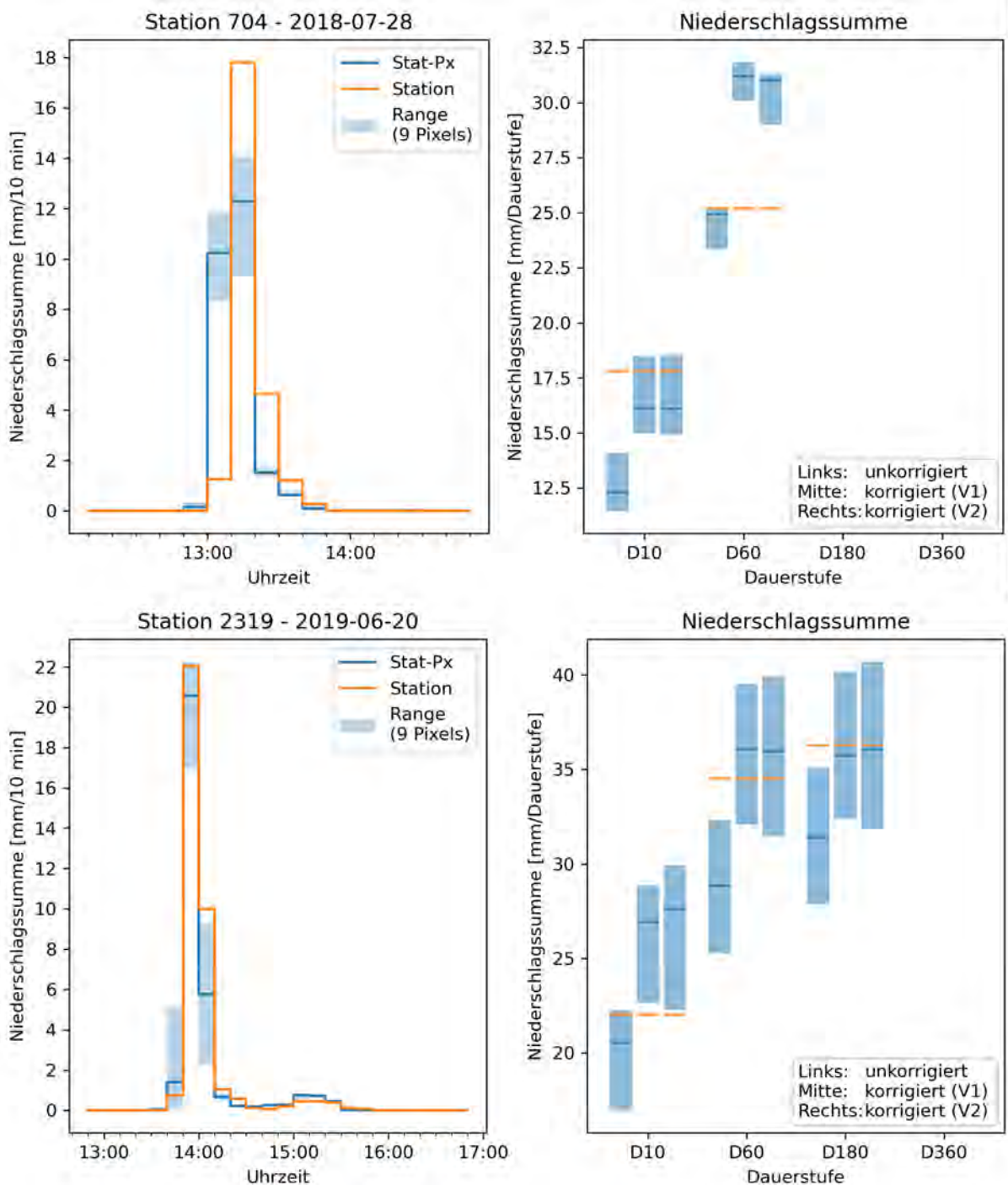


Abbildung 7.22: Korrektur der realen Ereignisse bei Station 704 und 2319

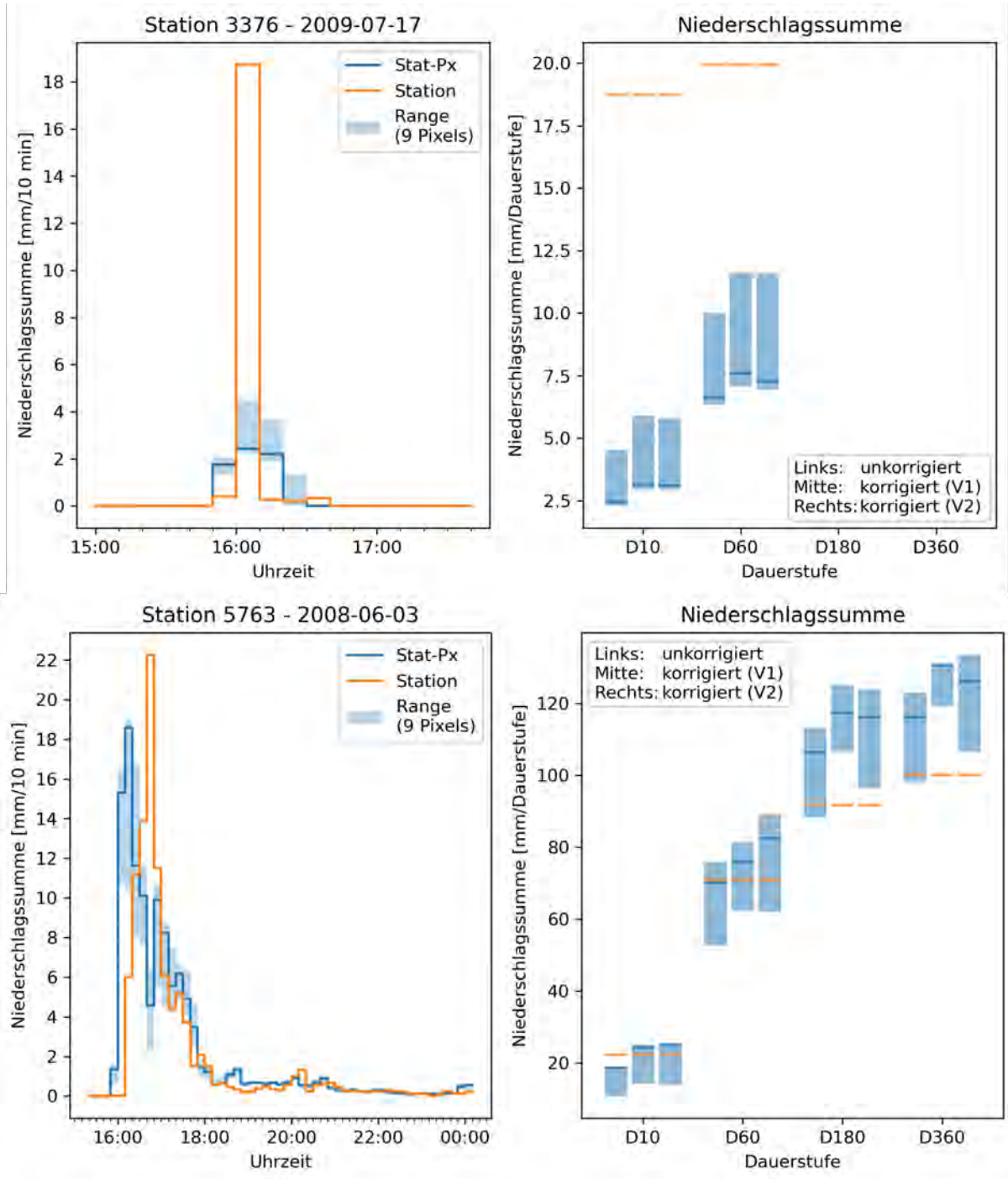


Abbildung 7.23: Korrektur der realen Ereignisse bei Station 3376 und 5763

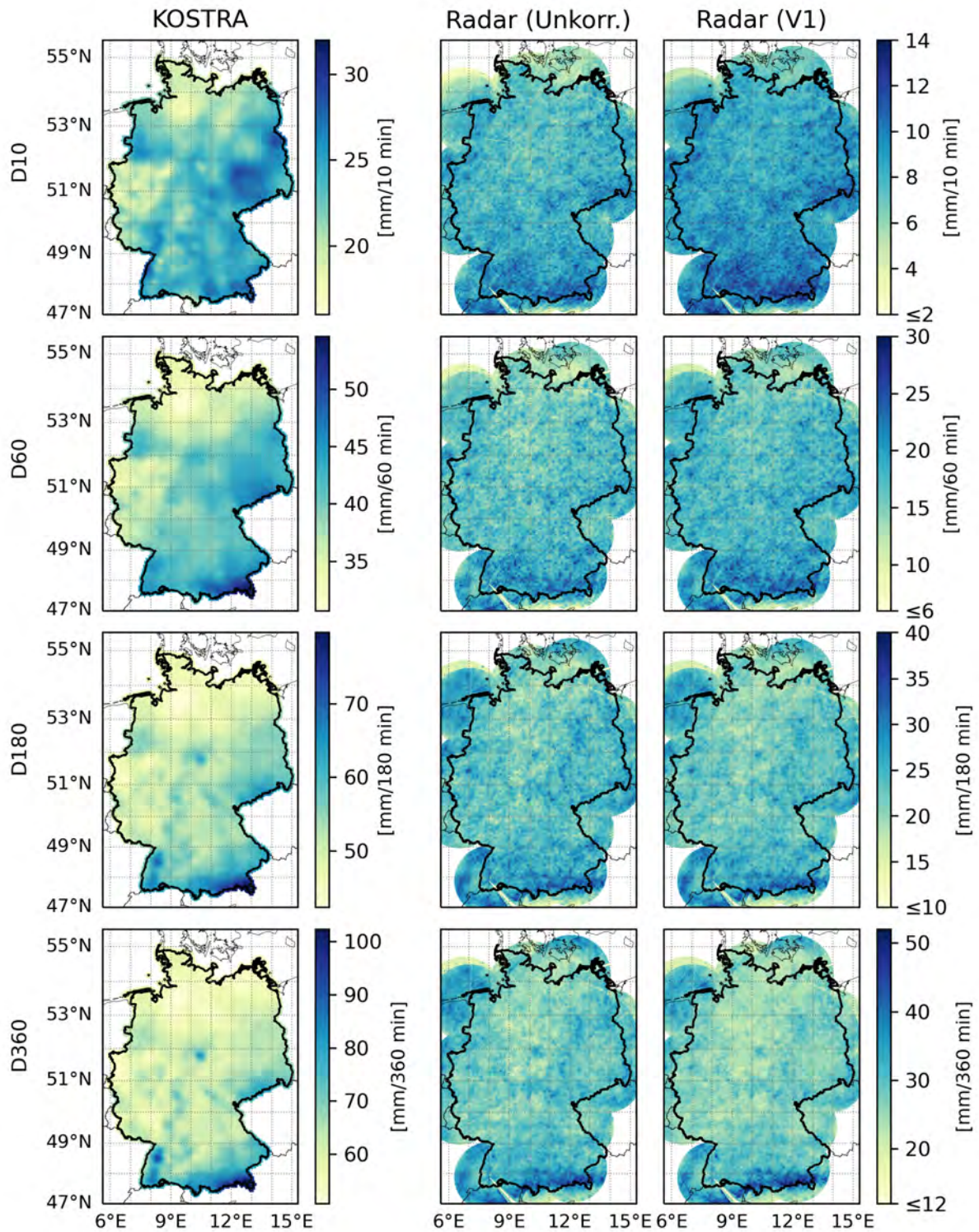


Abbildung 7.24: Vergleich des berechneten 50-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 1. Man beachte die anderen Achsenlimits bei der Karte der KOSTRA-Daten! (©EuroGeographics bezüglich der Verwaltungsgrenzen)

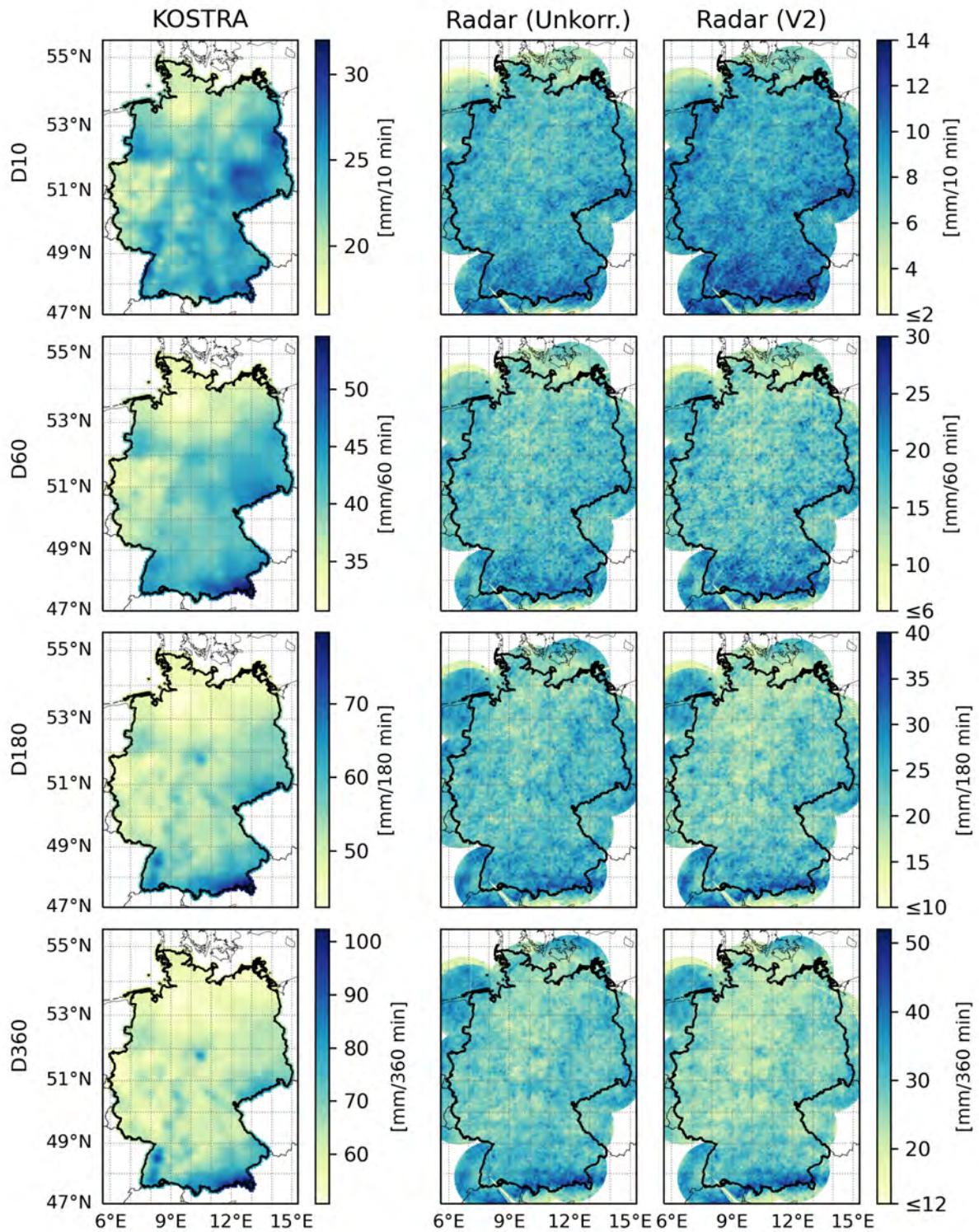


Abbildung 7.25: Vergleich des berechneten 50-jährlichen Bemessungsniederschlags mit dem KOSTRA-Produkt für Verfahren 2. Man beachte die anderen Achsenlimits bei der Karte der KOSTRA-Daten! (©EuroGeographics bezüglich der Verwaltungsgrenzen)

Erklärung zur Abgabe der Masterarbeit

Name des/der Studierenden

Hiermit versichere ich, dass

ich die eingereichte Masterarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Inhalte als solche kenntlich gemacht habe.

Die eingereichte Masterarbeit ist oder war weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens.

Datum, Ort

Unterschrift