

Uwe Haberlandt, Bora Shehu, Luisa Thiele, Winfried Willems, Henrike Stockel, Thomas Deutschländer, Thomas Junghänel & Jennifer Ostermüller

Methodische Untersuchungen für eine Neufassung der regionalisierten Starkregenstatistik KOSTRA-DWD

Methodological investigations for updating the regionalised extreme rainfall statistics KOSTRA-DWD

Der Beitrag präsentiert Ergebnisse eines Forschungsprojektes, welches sich mit methodischen Untersuchungen für eine Überarbeitung der Starkregenstatistik für Deutschland beschäftigt hat. Zuerst wurde eine Zeitreihenanalyse durchgeführt, um die Stationarität der Daten zu prüfen. Dann wurden unterschiedliche Methoden zur lokalen Extremwertstatistik und zur Regionalisierung verglichen, Ersteres speziell zum Ausgleich über die Dauerstufen. Die Datenbasis für den finalen Ansatz bildeten lange Reihen mit durchschnittlich 65 Beobachtungsjahren von 133 registrierenden Stationen. Zusätzlich wurden ca. 1.200 kürzere Reihen registrierender Stationen einbezogen. Es wurden Fehlerabschätzungen auf Basis von Kreuzvalidierungen gemacht und Unsicherheitsanalysen mithilfe von nicht-parametrischem Bootstrapping und von Simulationsexperimenten durchgeführt. Die Zeitreihenanalyse hat gezeigt, dass Instationaritäten speziell durch Sprünge zu Zeitpunkten der Sensorwechsel bei den Messgeräten auftreten. Trends wurden nur in geringem Umfang festgestellt. Die Reihen wurden daraufhin auf Sprünge bereinigt, und es wurde mit der stationären Extremwertstatistik gearbeitet. Als beste lokale Schätzmethode wurde das dauerstufenübergreifende Verfahren nach KOUTSOYIANNIS et al. (1998) auf Basis der Allgemeinen Extremwertverteilung GEV identifiziert. Als bestes Regionalisierungsverfahren hat sich Kriging der Parameter unter Verwendung der Informationen aus den kurzen Reihen als externe Drift erwiesen. Sowohl die Fehler der Ausgleichsmethode als auch die der Regionalisierung liegen jeweils unter 10 %. Die Unsicherheiten des Gesamtverfahrens haben deutschlandweit eine Spannweite von \pm 15 % bis 60 %. Der Deutsche Wetterdienst bereitet auf Basis dieser Ergebnisse eine Neufassung der Bemessungsniederschläge nach KOSTRA-DWD vor.

Schlagwörter: Starkregenstatistik, Regionalisierung, Unsicherheit, KOSTRA-DWD

Results from a research project are presented that dealt with methodological investigations regarding the renewal of the heavy rainfall statistics for Germany. First, a time series analysis was carried out testing the stationarity of the data. Then different methods for local extreme value analyses and for regionalisation were compared, the former especially regarding the smoothing over durations. The data basis for the final approach was the network with 133 recording stations of long observations with 65 years on average. In addition, data from recording stations with shorter observations (about 1,200 stations) were used. The errors were estimated based on cross-validation; the uncertainties were assessed by means of non-parametric bootstrapping and simulation experiments. The time series analysis showed that non-stationarities occur especially due to jumps at times when the instruments changed. Only a small portion of trends was found. The data was homogenised with a jump correction and a stationary extreme value analysis was applied. The best local estimation method was found to be the approach according to KOUTSOYIANNIS et al. (1998) using the GEV and estimating precipitation simultaneously over the durations. The best regionalisation method was kriging of the parameters using additional information from the short series as external drift. Both, the errors from the smoothing method as well as from regionalisation are smaller than 10 %. The uncertainties of the total approach show a range from \pm 15 % to 60 % throughout Germany. The German Weather Service is preparing a new version of KOSTRA-DWD based on these results.

Keywords: heavy rainfall statistics, regionalisation, uncertainty

1 Einführung

Bemessungsniederschläge stellen eine wesentliche Grundlage für die Dimensionierung von wasserwirtschaftlichen Bauwerken zum Schutz vor Hochwasser und Überflutungen in ländlichen und urbanen Räumen dar. Seit den 1980er Jahren liefert in Deutschland die "Koordinierte Starkniederschlagsregionalisierung und -auswertung" vom Deutschen Wetterdienst, kurz KOSTRA-DWD, Bemessungsniederschläge für unterschiedliche Anwendungszwecke (MALITZ & ERTEL, 2015). Diese wurde regelmäßig an neue Zeitperioden angepasst. Jedoch gab es kaum Anpassungen hinsichtlich neuer methodischer wissenschaftlicher Erkenntnisse. Zudem steht inzwischen ein dichtes Messnetz an automatisch registrierenden Niederschlagsstationen zur Verfügung, welches bisher nur partiell für die Extremwertstatistik genutzt wurde.

Ziel der vorliegenden Untersuchung war es deshalb, eine neue optimale Methodik für die flächendeckende Bereitstellung von Bemessungsniederschlägen in Deutschland zu finden, welche möglichst alle verfügbaren Stationsmessungen aus unterschiedlichen Quellen berücksichtigen sollte. Die neue Methodik sollte den aktuellen Stand der Wissenschaft adäquat berücksichtigen, objektiv, nachvollziehbar und praxistauglich sein. In einem entsprechenden Forschungsprojekt wurden unterschiedliche Methoden verglichen und ein optimales Gesamtverfahren identifiziert. Im Folgenden wird auszugsweise über die Ergebnisse dieses Forschungsprojektes berichtet. Detailliertere Ausführungen können auch JUNGHÄNEL et al. (2022) und SHEHU et al. (2022a, 2022b) entnommen werden.

Die lokale Extremwertstatistik für den Niederschlag wird in der deutschen Bemessungspraxis nach DWA-A 531 (2012) durch-

geführt und war entsprechend bisher auch so in KOSTRA-DWD implementiert. Internationale neue Erkenntnisse gibt es diesbezüglich jedoch bei der Auswahl der Verteilungsfunktionen (KOUTSOYIANNIS, 2004a, 2004b), den Parameterschätzverfahren und beim Ausgleich der Starkniederschläge über die Dauerstufen (FISCHER & SCHUMANN, 2018; KOUTSOYIANNIS et al., 1998; VAN DE VYVER, 2018).

Die Regionalisierung der punktuellen Starkregenstatistik für un beobachtete Orte wird entweder durch direkte Übertragung der Quantile, durch Übertragung der Parameter der Verteilungsfunktionen (JOHNSON & SHARMA, 2017; MADSEN, 2009; CERESSETTI et al., 2012) oder durch regionale Schätzung der Verteilungen auf Basis homogener Regionen mit dem Indexverfahren (BURN, 2014; HOSKING & WALLIS, 1997) vorgenommen. Die Regionalisierung der Parameter anstelle der Quantile ist zu bevorzugen, da bei Letzterem die Ordnungsprobleme größer sind und damit ein stärkerer Ausgleich über die Dauerstufen erforderlich wird. Die Übertragung der Parameter kann zum Beispiel durch geostatistische Verfahren, Regression, Splines oder Softcomputing-Ansätze durchgeführt werden. Am häufigsten wird in der internationalen Bemessungspraxis allerdings die Indexmethode verwendet.

Für die Quantifizierung der Unsicherheit müssen sowohl die lokale Extremwertschätzung (OVEREEM, 2008) als auch die Regionalisierung (UBOLDI, 2014) berücksichtigt werden. Insbesondere für Letzteres sind nur relativ wenig Untersuchungen in der Literatur zu finden. Methodisch werden hier häufig Bootstrapping-Verfahren oder Bayes'sche Methoden angewandt.

International wird zunehmend die Nutzung von Radardaten für die Extremwertstatistik propagiert. Auch in Deutschland gibt es entsprechende Anstrengungen (WINTERRATH, 2017). Wegen der noch hohen Unsicherheiten der Niederschlagssschätzung aus Radar, der noch relativ kurzen Beobachtungsperioden sowie einiger noch fehlender methodischer Grundlagen konnten die Wetterradardaten für die Ableitung der Extremwertstatistik im Forschungsvorhaben noch nicht direkt genutzt werden.

Im Kapitel 2 werden zunächst das Untersuchungsgebiet und die verwendeten Datensätze vorgestellt. Dies ist notwendig, da das methodische Vorgehen davon abhängt. Im Kapitel 3 werden dann die Methodiken der lokalen Extremwertstatistik, der Regionalisierung, der Unsicherheitsschätzung und der Gütebewertung beschrieben. Im Kapitel 4 werden die Ergebnisse erläutert, ebenfalls untergliedert in lokale Schätzung, Regionalisierung und Unsicherheit. Im Kapitel 5 werden Kartendarstellungen der Starkniederschläge und von deren Unsicherheiten diskutiert. Der Artikel schließt im Kapitel 6 mit Schlussfolgerungen und einem Ausblick.

2 Untersuchungsgebiet und Daten

2.1 Verfügbare Daten

Die Untersuchungen erfolgen für ganz Deutschland, wobei zunächst Niederschlagsdaten aus drei verschiedenen Kollektiven des Deutschen Wetterdienstes (DWD) einbezogen werden (Abb. 1). Dies betrifft die Langzeitbeobachtungen registrierender Messstellen (LS) mit 133 Stationen, die kürzeren Reihen registrierender Messstellen (SS) mit 1.261 Stationen (ohne LS-Standorte) und das Messnetz der täglich meldenden Messstellen (DS) mit 4.068 Stationen (ohne LS- und SS-Standorte).

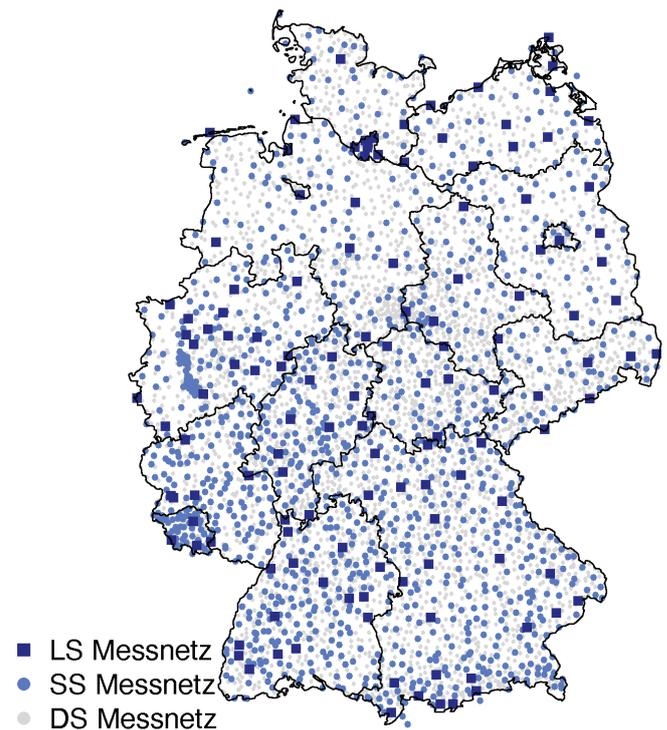


Abbildung 1

Verwendete Niederschlagskollektive in Deutschland.
Variations of terrestrial water storage (TWS) in Germany from four GRACE/Rainfall networks used in Germany.

Es wurden alle verfügbaren Messungen aller Stationen vom Beginn der Aufzeichnung bis Ende 2020 verwendet. Damit ausreichend Überlappung besteht, wurden für die Stationen mit langen Reihen (LS) mindestens 40 Jahre Beobachtungslänge gefordert und für die Stationen mit kurzen Reihen (SS) und die der Tageswerte (DS) mindestens 10 Jahre. Für das finale Verfahren verblieben die beiden Kollektive LS und DS. Daten für das LS-Kollektiv liegen für die meisten Stationen zwischen 1951 und 2020 vor, für das SS-Kollektiv zwischen 2000 und 2020. Die hier verwendete LS-Kerndatenbasis repräsentiert sowohl eine zeitliche Erweiterung der aktuellen KOSTRA-Datenbasis um 10 Jahre in die Gegenwart als auch teilweise eine Erweiterung in die Vergangenheit (Nachdigitalisierung). Zudem wurden mehrere lange Reihen neu erschlossen und die Anzahl der Stationen mit langen Reihen um ca. 50 % erweitert.

In Abbildung 2 ist die Beobachtungslänge für die verfügbaren Stationen der drei Kollektive dargestellt. Das LS-Kollektiv hat mit durchschnittlich 65 Jahren Beobachtungslänge zeitlich die beste Repräsentanz, das SS-Kollektiv ist dagegen mit durchschnittlich 18 Jahren Beobachtungslänge zeitlich deutlich beschränkt, jedoch räumlich repräsentativer. Das DS-Kollektiv mit den Tagesstationen ist für Dauerstufen kleiner 24 h nicht direkt nutzbar, enthält aber aufgrund der zeitlichen Länge und räumlichen Dichte wichtige Informationen, die für Disaggregation bzw. Regionalisierung genutzt werden können.

2.2 Datenaufbereitung und Homogenisierung

Zunächst wurden die Niederschlagsdaten auf Fehler geprüft und wenn nötig bereinigt. Um die täglichen Niederschläge für die

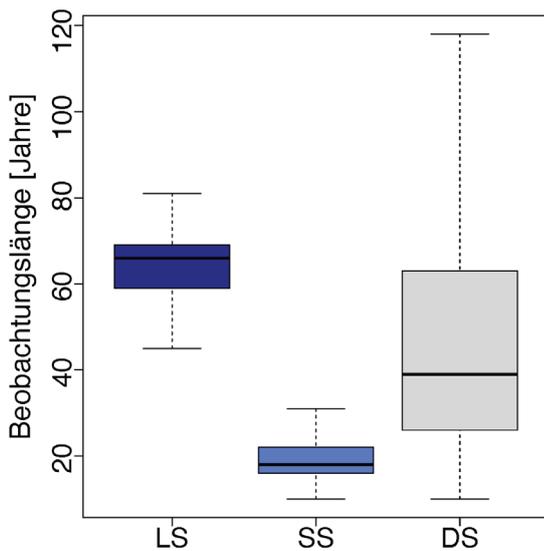


Abbildung 2
 Beobachtungslänge der Niederschlagsmessungen.
Observation length of precipitation measurements.

Regionalisierung nutzen zu können, wurden diese mithilfe eines multiplikativen Kaskadenmodells (OLSSON, 1998; LISNIAK et al., 2013) disaggregiert. Für die Parametrisierung des Disaggregationsmodells wurden hier erstmalig flächendeckend vorliegende Radardaten verwendet. Dies hat im Vergleich zur Verwendung von hoch aufgelösten Stationsdaten, die nur unregelmäßig

verteilt vorliegen, den Vorteil, dass keine Regionalisierung der Disaggregationsparameter erforderlich ist.

Die Stichprobenbildung erfolgte hier mit jährlichen Serien klassisch nach DWA (2012). Basierend auf Niederschlagsdaten mit einer Auflösung von 5 min wurden jährliche Maxima für die Dauerstufen $D = 5 \text{ min}, 10 \text{ min}, 15 \text{ min}, 30 \text{ min}, 1 \text{ h}, 2 \text{ h}, 6 \text{ h}, 12 \text{ h}, 1 \text{ d}, 2 \text{ d}, 3 \text{ d}$ und 7 d gebildet. Für die Dauerstufen $D = 5 \text{ min}, 10 \text{ min}$ und 15 min wurde eine Intervallkorrektur gegen Unterschätzung angewandt (DWA, 2012).

Die Prüfung der Stationarität erfolgte auf Basis der 133 Stationen mit langen Reihen (LS-Kollektiv). Es wurden unterschiedliche Untersuchungen durchgeführt und verschiedene Tests angewandt (JUNGHÄNEL et al., 2022; WILLEMS et al., 2023, in diesem Heft). Eine Entscheidung wurde schließlich mithilfe der auf die jährlichen Maxima angewandten instationären Extremwertanalyse (IEA) unter Verwendung der Allgemeinen Extremwertverteilung GEV (generalized extreme value) getroffen (COLES, 2001). Diese wurde so erweitert, dass über das Akaike-Informationskriterium eine Auswahl des am besten geeigneten Extremwertmodells aus den möglichen Klassen "stationär", "trendartig-instationär" oder "sprungartig-instationär" möglich ist (Abb. 3). Es wurde festgestellt, dass die Zeitreihen der kurzen Dauerstufen von $D = 5 \text{ min}$ bis $D = 30 \text{ min}$ größtenteils instationär sind, wobei die Instationaritäten maßgeblich durch positive Sprünge bedingt sind. Die Sprünge finden zu den Zeitpunkten des Sensorwechsels zwischen analogen und digitalen Messgeräten statt. Ab einer Dauerstufe von $D = 1 \text{ h}$ sind die Zeitreihen überwiegend stationär. Es ist zu vermuten, dass dies gerätetechnisch bedingt ist. Die früher verwendeten klassischen Regenschreiber konnten nur eine deutlich geringere maximale Intensität aufzeichnen im

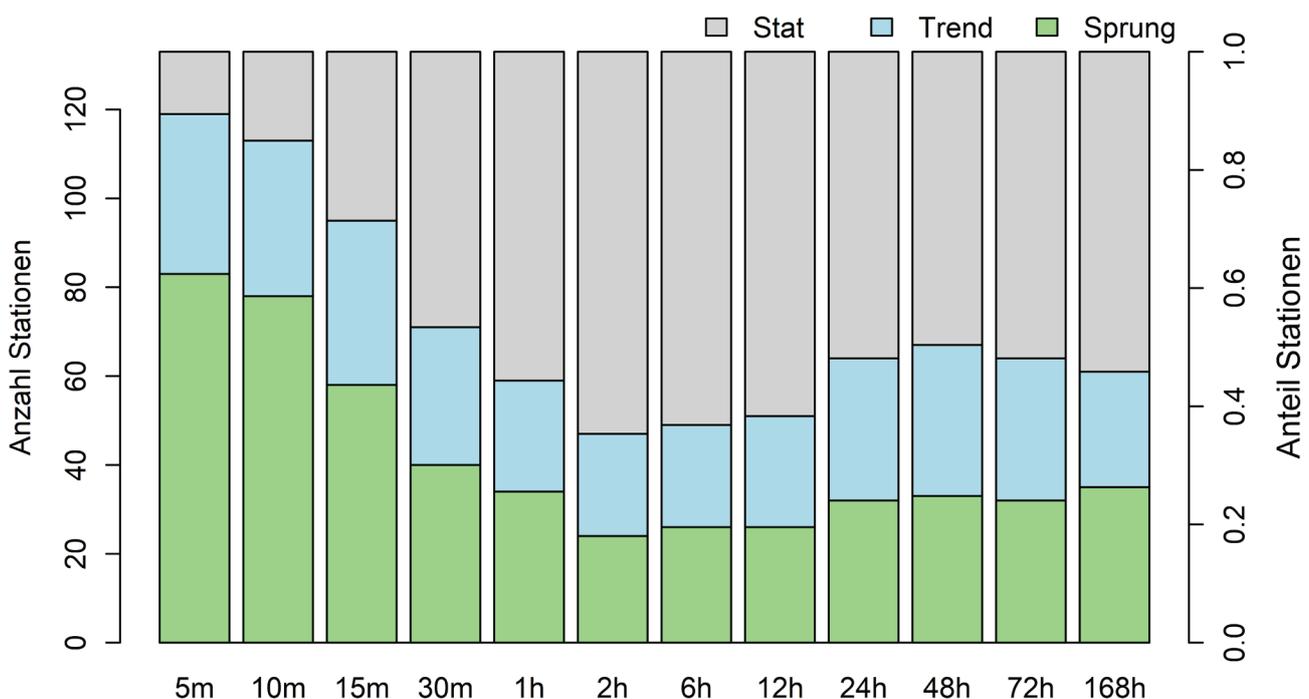


Abbildung 3
 Ergebnisse der erweiterten instationären Extremwertanalyse.
Results of the extended non-stationary extreme value analyses.

Vergleich zu den modernen Geräten, z. B. nach dem Wägeprinzip, was zur einer "Deckelung" der Messung von Extremereignissen mit kurzen Dauerstufen führt und positive Sprünge verursachen kann (WILLEMS et al., 2023, in diesem Heft). Trends, sowohl positive wie negative, sind in lediglich 15 bis 30 % der Serien enthalten. Aus diesem Grund wurde entschieden, eine stationäre Extremwertanalyse unter vorheriger Entfernung der Sprünge zu machen. Die Entfernung der Sprünge bzw. Homogenisierung der Daten erfolgte durch Anhebung der Niederschläge vor der Sprungstelle auf das Niveau nach dem Sprung. Ein positiver Nebeneffekt dieser endbetonten Sprungkorrektur der langen Reihen (LS-Kollektiv) ist die dadurch gebildete Konsistenz mit den kurzen Reihen aus dem SS-Kollektiv in den jüngeren Jahren des Auswertungszeitraums.

3 Methodik

Für die Identifizierung eines optimalen Verfahrens zur flächendeckenden Bereitstellung von Bemessungsniederschlägen wurden unterschiedliche Methoden für die lokale und regionale Schätzung verglichen. Das Gesamtverfahren gliedert sich in drei Schritte: 1. Lokale Extremwertstatistik, 2. Regionalisierung und 3. Unsicherheitsschätzung. Für alle Schritte wurden Alternativen betrachtet. Basis für die Untersuchungen bildet das homogenisierte LS-Kollektiv mit den zeitlich repräsentativen langen Beobachtungen. Für die Regionalisierung wurde geprüft, inwieweit die Kollektive SS und DS signifikante Zusatzinformationen liefern. Im Folgenden werden die wesentlichen Schritte und die am meisten relevanten Verfahren methodisch erläutert.

3.1 Lokale Extremwertstatistik

Referenzmethode

Für die lokale Extremwertstatistik wurde die Allgemeine Extremwertverteilung (GEV) verwendet, welche sich grundsätzlich als am besten geeignet für die Beschreibung von Starkniederschlägen erwiesen hat (KOUTSOYIANNIS, 2004a, 2004b):

$$F(x; \mu, \sigma, \gamma) = \begin{cases} \exp\{-[1 + \gamma(x + \mu)/\sigma]^{1/\gamma}\}, & 1 + \gamma(x - \mu)/\sigma > 0, \gamma \neq 0 \\ \exp\{-\exp[-(x - \mu)/\sigma]\}, & \gamma = 0 \end{cases}, \quad (1)$$

wobei μ der Lageparameter mit $\mu \in \mathbb{R}$, σ der Skalenparameter mit $\sigma > 0$ und γ der Formparameter mit $\gamma \in \mathbb{R}$ ist. Falls $\gamma > 0$ ist, weist die Verteilung ein "Heavy Tail"-Verhalten auf. Die Parameterschätzung erfolgte mit der Methode der L-Momente, wobei der Formparameter alternativ mitgeschätzt oder mit $\gamma = 0.10$ konstant gehalten wird. Letzteres wird bei kurzen Reihen (≤ 50 Jahre) empfohlen, da bei diesen Stichprobenumfängen eine zuverlässige Schätzung des Formparameters nicht möglich ist (KOUTSOYIANNIS, 2004b).

Die separate Schätzung der Verteilungen für jede Dauerstufe führt zu Ordnungsproblemen. Je nach Datenlage kann nicht immer gewährleistet werden, dass das an einer Station geschätzte Quantil gegebener Jährlichkeit umso größer ist, je höher die zugrunde liegende Dauerstufe ist. Dies kann mit Korrekturverfahren erzwungen werden, von denen die im Nachfolgenden vorgestellten zwei Verfahren hier verglichen wurden.

Einstufiges Verfahren nach Koutsoyiannis

Das einstufige Verfahren nach KOUTSOYIANNIS et al. (1998) wird dauerstufenübergreifend für jede Station einzeln angewandt. Für die Parameterschätzung schlagen KOUTSOYIANNIS et al. (1998)

mehrere Methoden vor. Bei der hier betrachteten, robusten Methode werden zunächst die Niederschlagsintensitäten $i_{D,l}$ aus den jährlichen Serien für die Dauern $D = 1, \dots, k$ und die Jahre $l = 1, \dots, n$ berechnet. Diese werden für alle Dauerstufen mit

$$y_{D,l} = i_{D,l} \cdot b_D \quad (2)$$

skaliert, wobei b_D ein dauerstufenabhängiger Skalierungsfaktor ist. Dieser muss so bestimmt werden, dass eine einheitliche Wahrscheinlichkeitsverteilung für die skalierten Intensitäten y_D unabhängig von der Dauerstufe D gültig ist:

$$F(y) \equiv F(y_D) \quad \forall D. \quad (3)$$

Der Skalierungsfaktor wird als Funktion der Dauer ausgedrückt

$$b_D = (D + \theta)^\eta, \quad (4)$$

wobei θ und η unbekannte Parameter sind. Das Ziel besteht darin, θ und η so zu bestimmen, dass Gl. (3) so gut wie möglich erfüllt ist. Dazu wird die Test-Statistik des Kruskal-Wallis-Tests verwendet (HELSEL et al., 2020):

$$KW = \frac{12}{m(m+1)} \sum_{D=1}^k n_D \left(\bar{r}_D - \frac{m+1}{2} \right)^2 \quad (5)$$

wobei m der Rang aller $y_{D,l}$ und unabhängig von D ist, n_D die Anzahl der Jahre und \bar{r}_D der mittlere Rang für die Stichprobe D .

Mit diesem Test werden k -Stichproben (hier die skalierten Intensitäten $y_{D,l}$ aus den k -Dauerstufen) auf Ähnlichkeit hinsichtlich ihrer Zugehörigkeit einer einheitlichen Verteilungsfunktion verglichen. Die Teststatistik KW des Testes ist umso kleiner, je ähnlicher die Stichproben sind. Dies kann man sich zunutze machen, indem man die gesuchten Parameter durch numerische Minimierung der Teststatistik findet:

$$KW(\theta, \eta) \rightarrow \text{Min}. \quad (6)$$

Dabei sind die Gl. (4), (2) und (5) iterativ unter Nutzung eines beliebigen numerischen Optimierungsverfahrens anzuwenden. Anschließend wird die GEV an die auf diese Weise dauerstufenübergreifend generalisierten Intensitäten y mithilfe der L-Momentenmethode angepasst. Somit sind maximal fünf Parameter zu schätzen, die drei Parameter μ , σ und γ der GEV und die zwei Parameter θ und η für die Skalierung. Alternativ wird hier wiederum der Formparameter für alle Dauerstufen auf $\gamma = 0.10$ fixiert, was die Anzahl der Parameter um eins verringert. Mit diesem Vorgehen der gemeinsamen Schätzung über alle Dauerstufen treten keine Ordnungsprobleme mehr auf. Zudem ist die Parameterschätzung für die GEV wegen der größeren Stichprobe robuster. Dies gilt insbesondere bei Fixierung des Formparameters.

Zweistufiges Verfahren nach Fischer & Schumann

Beim zweistufigen Verfahren nach FISCHER & SCHUMANN (2018) in der hier umgesetzten Form werden zuerst separat für jede Dauerstufe GEV-Verteilungen mit der Methode der L-Momente angepasst und anschließend deren Parameter über Funktionen in Abhängigkeit der Dauer ausgeglichen:

$$\mu(D) = \frac{\alpha_\mu}{D^{\beta_\mu}}, \quad \sigma(D) = \frac{\alpha_\sigma}{D^{\beta_\sigma}}, \quad \frac{\sigma}{\gamma} = \alpha + \beta \cdot D, \quad (7)$$

wobei D die Dauerstufe, μ , σ und γ die GEV-Parameter sowie α und β die Funktionsparameter sind. Zum Vergleich mit der Methode nach Koutsoyiannis wird hier alternativ der Formparameter für alle Dauerstufen auf $\gamma = 0.10$ fixiert. Bei diesem Verfahren sind entweder sechs oder vier Parameter zu schätzen. Auch bei dieser Methode treten keine Ordnungsprobleme auf.

3.2 Regionalisierung

Ziel der Regionalisierung ist die Übertragung der lokalen Schätzungen an den Stationsstandorten auf beliebige Orte, hier auf ein deutschlandweites Raster mit einer Auflösung von 5 km x 5 km. Dabei können Quantile, Momente oder Parameter regionalisiert werden. Voruntersuchungen haben ergeben, dass die Güte für die Schätzung der Extremniederschläge für alle drei Ansätze ähnlich ist, es jedoch zu signifikanten Ordnungsproblemen bei der Interpolation der Quantile kommt und die Interpolation der Momente eine Anwendung der Methode nach Koutsoyiannis erschwert. Somit werden hier im Folgenden nur Parameterinterpolationen betrachtet. Für die Regionalisierung der Parameter werden drei bewährte Verfahren verglichen: Ordinary Kriging (OK), Kriging mit Externer Drift (KED) und die Indexmethode.

Für die beiden geostatistischen Verfahren OK und KED werden jeweils die Parameter der als optimal identifizierten lokalen Methode unabhängig voneinander räumlich interpoliert. Für das Indexverfahren werden nicht die Parameter interpoliert, sondern regional gültige Verteilungsfunktionen der Starkniederschläge gebildet. Die drei Regionalisierungsverfahren werden im Folgenden kurz charakterisiert. Ausführliche Beschreibungen können der jeweils zitierten Literatur entnommen werden.

Ordinary Kriging

Ordinary Kriging (OK) ist die in der Geostatistik am häufigsten angewendete stationäre Interpolationsmethode (GOOVAERTS, 1997). Hierbei wird die intrinsische Hypothese verwendet. Diese verlangt einerseits die Konstanz des Erwartungswertes $E[Z]$ innerhalb des Interpolationsraums und andererseits, dass die Varianz des Inkrements $Z(u)-Z(u+h)$ nur von der Entfernung h und nicht vom Ort u abhängt. Die Funktion $\gamma(h)$ wird Variogramm genannt:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{u_i-u_j=h} (Z(u_i) - Z(u_j))^2, \tag{8}$$

wobei $N(h)$ die Anzahl der Punktpaare ist, die durch die Entfernung h getrennt sind. An dieses empirische Variogramm wird mithilfe der Methode der kleinsten Quadrate ein theoretisches Variogramm angepasst. Hier wird das sphärische Variogramm-Modell verwendet. Für die Interpolation wird die Zielvariable Z an einem unbekanntem Ort u_0 durch den gewichteten Durchschnitt aller Z an den n bekannten, nahe gelegenen Orten u_i geschätzt:

$$Z'(u_0) = \sum_{i=0}^n \lambda_i \cdot Z(u_i). \tag{9}$$

Die Gewichte λ_i werden mit dem Kriging-Gleichungssystem als Funktion der Kollektivkonfiguration und des theoretischen Variogramms bestimmt.

Kriging mit externer Drift

Beim Kriging mit externer Drift (KED) werden eine oder mehrere Zusatzvariablen bei der Interpolation berücksichtigt. Im Gegen-

satz zum Ordinary Kriging (OK) wird die Stationaritätsannahme bezüglich des Erwartungswertes verworfen. Bei Verwendung einer Zusatzvariable Y wird angenommen, dass der Erwartungswert der Zielvariablen $E[Z]$ an einem Ort u linear von Y abhängt:

$$E[Z(u) | Y_1(u), Y_2(u), \dots, Y_m(u)] = b_0 + \sum_{k=1}^m b_k Y_k(u). \tag{10}$$

Das Variogramm des KED wird hier vereinfacht wie beim OK berechnet. Die Gewichte werden mit dem Kriginggleichungssystem bestimmt, welches zusätzlich noch von Y abhängt. Die Interpolation erfolgt mit Gleichung (9). Als Zusatzvariable Y können verschiedene Standortcharakteristika (z. B. geografische Merkmale) oder Niederschlagsinformationen aus anderen Quellen verwendet werden. Es hat sich jedoch gezeigt, dass Standortcharakteristika kaum mit den Parametern der lokalen Verfahren korrelieren (SHEHU et al., 2022a). Auch Untersuchungen von Radardaten zeigten, dass insbesondere bei kurzen Dauerstufen kaum noch ein Zusammenhang zwischen geografischen Merkmalen und der Häufigkeit des Auftretens von Starkniederschlag besteht (WINTERRATH et al., 2017). Aus diesem Grund werden hier nur die interpolierten Parameter aus den kurzen (SS) und aus den disaggregierten täglichen Reihen (DS) als Zusatzinformation verwendet.

Das R-Paket "gstat" wird hier genutzt, um Variogramme anzupassen und die Variablen mit OK oder KED zu interpolieren (PEBESMA, 2004).

Indexmethode

Für die Regionalisierung wird zusätzlich die Indexmethode nach HOSKING & WALLIS (1997) verwendet. Durch die Bündelung von Informationen in statistisch-homogenen Regionen kann eine robustere Schätzung von Starkniederschlägen vorgenommen und innerhalb der Regionen die Informationen auf beliebige unbeobachtete Punkte übertragen werden. Die Bestimmung von homogenen Regionen wird mit der von HOSKING & WALLIS (1997) vorgeschlagenen Clustermethode nach WARD (1963) mithilfe von Standortmerkmalen (z. B. Breitengrad, Längengrad, Geländehöhe, mittlerer Jahresniederschlag und Sonnenscheindauer) durchgeführt. Dabei wurden neun homogene Regionen identifiziert (SHEHU et al., 2022a). Die weitere Schätzung erfolgt separat für alle homogenen Regionen. Zuerst werden für alle Stationen j und Dauerstufen D die jährlichen Serien bestimmt. Dann werden für jede Station nach KOUTSOYIANNIS et al. (1998) die Parameter θ und η und damit die generalisierten Intensitäten i ermittelt. Der Index Ψ ergibt sich als Mittelwert dieser Intensitäten. Anschließend werden die lokalen L-Momente $L_{k,j}$ für die Ordnungen $k = 1, \dots, 3$ und alle Stationen j geschätzt und die regionalen L-Momente L_k aus den lokalen L-Momenten, gewichtet mit der Reihenlänge n_j an den Stationen j , bestimmt:

$$L_k = \sum_{j=1}^N n_j L_{k,j} / \sum_{j=1}^N n_j. \tag{11}$$

Die Allgemeine Extremwertverteilung (GEV) wird über die regionalen L-Momente angepasst. Schließlich können mithilfe der regionalen Mittelwerte von θ und η , den regionalen GEV-Parametern und dem Index Ψ die Quantile für beliebige Dauerstufen ermittelt werden. Die vorherige geostatistische Interpolation des Indexes Ψ , entweder mit OK oder mit KED, gestattet eine Übertragung der Extremwertstatistik auf unbeobachtete Punkte innerhalb der Region.

Das R-Paket "ImomRFA" wird für die Anwendung der Indexmethode verwendet (HOSKING & WALLIS, 1997)

3.3 Bestimmung der Verfahrensfehler

Für die Abschätzung des Fehlers der verwendeten Verfahren muss der Schätzwert Z* mit einem Referenzwert Z verglichen werden. Folgende Fehlermaße werden für die Vergleiche im Kapitel 4 verwendet:

der prozentuale systematische Fehler

$$PBIAS [\%] = \frac{Z_i^* - Z_i}{Z_i} \cdot 100, \tag{12}$$

der prozentuale relative Standardfehler

$$RMSE [\%] = 100 \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (Z_i^* - Z_i)^2} / \bar{Z}, \tag{13}$$

das Nash-Sutcliffe-Effizienzkriterium

$$NSE[-] = 1 - \frac{\sum_{i=1}^n (Z^* - Z)^2}{\sum_{i=1}^n (Z^* - \bar{Z})^2} \tag{14}$$

und ein Kriterium, welches für die Interpolation die relative Verbesserung zu einem Referenzverfahren beschreibt

$$Perf_{impr} [\%] = 100 \cdot \frac{(Perf_{ref} - Perf_{new})}{Perf_{ref}}, \tag{15}$$

wobei als Gütekriterien, hier Perf_{impr} => RMSE (Gl. 13) und Perf_{impr} => PBIAS (Gl. 12), verwendet werden.

Für die Bestimmung der Verfahrensfehler wird ausschließlich das Kollektiv mit den langen Reihen LS verwendet. Für die Evaluation der lokalen Ausgleichsmethode im Kapitel 4.1 wird als Referenzwert Z das direkt geschätzte Niederschlagsquantil der GEV und als Schätzwert Z* das mit dem Ausgleichsverfahren entweder nach KOUTSOYIANNIS et al. (1998) oder nach FISCHER & SCHUMANN (2018) ermittelte Niederschlagsquantil verwendet. Für den Vergleich der Regionalisierungsverfahren im Kapitel 4.2 wird eine Kreuzvalidierung durchgeführt. Als Referenzwert Z wird jetzt die lokale ausgeglichene Schätzung verwendet, als Schätzwert Z* das mit der Interpolation erhaltene Niederschlagsquantil.

3.4 Bestimmung der Verfahrensunsicherheiten

Neben der Bestimmung der Verfahrensfehler ist es wichtig, die Unsicherheiten der Verfahren zu ermitteln. Im Gegensatz zur Fehlerermittlung wird hier nicht mit Referenzwerten verglichen, sondern die interne Variabilität der Ergebnisse, hauptsächlich bedingt durch unsichere Stichproben, betrachtet. Die Unsicherheit wird durch nicht-parametrisches Bootstrapping (DAVISON & HINKLEY, 2005) analysiert. Dabei werden unterschiedliche Expe-

rimente zur Ermittlung verschiedener Unsicherheiten durchgeführt:

- 1) Unsicherheit durch die lokale Parameterschätzung der GEV,
- 2) Unsicherheit der Variogrammschätzung für die Interpolation,
- 3) Unsicherheit bei der OK-Interpolation der Hilfsvariable Y,
- 4) Unsicherheit bei der KED-Interpolation der Zielvariable Z,
- 5) Gesamtunsicherheit mit Propagation der Einzelunsicherheiten.

Die Quantifizierung der Unsicherheit erfolgt durch Ermittlung des normalisierten 95%-Konfidenzintervalls der Schätzung:

$$nCI95\% [-] = \frac{(Z_{0,975}^* - Z_{0,025}^*)}{\bar{Z}^*}, \tag{16}$$

wobei im Zähler das 95%-Konfidenzintervall der geschätzten Zielgröße Z* aus dem Bootstrapping bzw. den wiederholten Simulationen und im Nenner der Mittelwert über alle ermittelten Zielwerte stehen.

Für 1) erfolgt ein 100-maliges Resampling mit Zurücklegen (d. h. jeder Wert kann mehrfach gezogen werden) aus den lokalen Stichproben unter Beibehaltung der jeweiligen beobachteten Reihenlänge mit wiederholter Anpassung der GEV. Für 2) wird die Variogrammschätzung 100-mal wiederholt, wobei aus der gemeinsamen Stichprobe der SS- und LS-Messnetze mit 1.394 Stationen jeweils 100 Stationen zufällig gezogen werden. Für 3) und 4) wird eine bedingte sequentielle Gauss'sche Simulation (SGS) durchgeführt. Methodische Details dazu können SHEHU & HABERLANDT (2022b) entnommen werden. Für 5) werden schließlich alle wesentlichen Einzelunsicherheiten zu einer Gesamtunsicherheit propagiert. Die Quantifizierung der verschiedenen Unsicherheiten mit Gl. (12) erfolgt primär für das Kollektiv LS mit den langen Reihen, wofür beim SGS ebenfalls mit Kreuzvalidierung gearbeitet wird. Die komplette Unsicherheitsanalyse wird nur für die final am besten geeignete Methodik durchgeführt (Kap. 4.3).

Das R-Paket "gsta" wird hier genutzt, um die SGS-Simulationen sowohl für OK als auch für KED durchzuführen (PEBESMA, 2004).

4 Ergebnisse

4.1 Lokale Extremwertstatistik

Zunächst wird untersucht, welche der beiden Ausgleichsmethoden unter Berücksichtigung unterschiedlicher Annahmen für den Shape-Parameter besser geeignet ist. Die Tabelle 1 zeigt die hier diskutierten Varianten.

Die Abbildung 4 zeigt die Ergebnisse des Vergleiches der Ausgleichsmethoden anhand des RMSE. Beide Verfahren haben eine ähnliche Güte mit mittleren Fehlern unter 10 %. Dabei ist FS

Tabelle 1 Ausgleichsmethoden, die für die lokale Statistik untersucht wurden. <i>Smoothing methods, which were analysed for local statistics.</i>			
Ausgleichsmethode	Formparameter	Abkürzung	Optimierte Parameter
Koutsoyiannis (KO)	stationsweise individuell aber konstant über alle Dauerstufen	KO.CON	μ, σ, γ, θ, η
Koutsoyiannis (KO)	für alle Stationen auf γ = 0,1 festgelegt	KO.FIX	μ, σ, θ, η
Fischer/Schumann (FS)	originäre Fischer-Schumann-Methode	FS.RLM	α _μ , β _μ , α _σ , β _σ , α, β
Fischer/Schumann (FS)	für alle Stationen auf γ = 0,1 festgelegt	FS.FIX	α _μ , β _μ , α _σ , β _σ

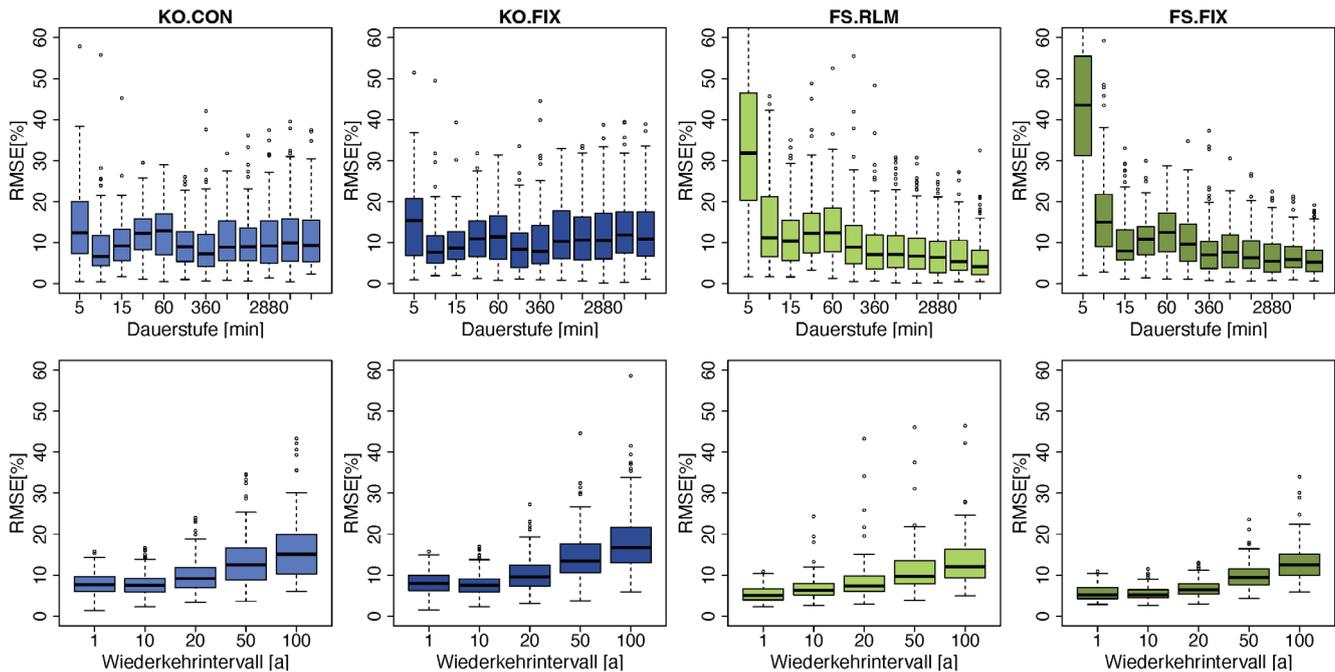


Abbildung 4

Abweichung der Ausgleichsschätzung Koutsoyiannis (KO) vs. Fischer/Schumann (FS) von der lokalen Dauerstufen-spezifischen Schätzung der Niederschlagsquantile für 133 lange Reihen; oben: gemittelt über alle Wiederkehrintervalle; unten: gemittelt über alle Dauerstufen.

Deviation of the generalised estimation of Koutsoyiannis (KO) vs. Fischer/Schumann (FS) from the duration-specific local estimation of the precipitation quantiles for 133 long series. Upper row: averaged over all return periods; lower row: averaged over all durations.

für lange Dauerstufen etwas im Vorteil und KO für kurze Dauerstufen, wobei hier insbesondere die Dauerstufen $D = 5$ min und $D = 10$ min auffallen. Wird über alle verwendeten Dauerstufen gemittelt, erhalten höhere Dauerstufen mehr Gewicht und FS ist leicht im Vorteil, wobei sich die typische Zunahme des Fehlers mit dem Wiederkehrintervall zeigt. Die Fixierung des Shape-Parameters für alle Stationen auf 0,1 verschlechtert die Modellgüte kaum.

Die Abbildung 5 zeigt die lokale Unsicherheit für beide Verfahren im Vergleich. Es wird ersichtlich, dass das originale FS-Verfahren die höchste Unsicherheit aufweist, partiell der Tatsache geschuldet, dass es die höchste Anzahl an Parametern benötigt. Die beiden Versionen FS.FIX und KO.FIX mit fixiertem Shape-Parameter sind wiederum sehr ähnlich und zeigen die geringste Unsicherheit mit leichten Vorteilen für KO.FIX. Für die weitere Bearbeitung und Regionalisierung wird die Ausgleichsmethode nach KOUTSOYIANNIS et al. (1998) mit fixiertem Shape-Parameter ausgewählt, da sie die bessere Güte für die wichtigen kurzen Dauerstufen hat und insgesamt eine etwas geringere Unsicherheit aufweist.

4.2 Regionalisierung

Als Basisdatensatz für die Regionalisierung werden die 133 Stationen aus dem Kollektiv mit langen Reihen verwendet (LS). Als Zusatzinformationen für OK und KED dienen Daten aus den anderen beiden Kollektiven SS und DS. Die Tabelle 2 zeigt die wesentlichen untersuchten Methoden-Daten-Kombinationen.

Als Basisverfahren werden zunächst die beiden einfachen Verfahren Ordinary Kriging OK[LS] und das Index-Verfahren INDEX[LS]

nur unter Verwendung der Daten aus dem Kollektiv mit den langen Reihen gegenübergestellt. Die Tabelle 3 zeigt einerseits, dass die absolute Interpolationsgüte insgesamt sehr gut ist und andererseits, dass diese beiden Verfahren eine ähnliche Leistung

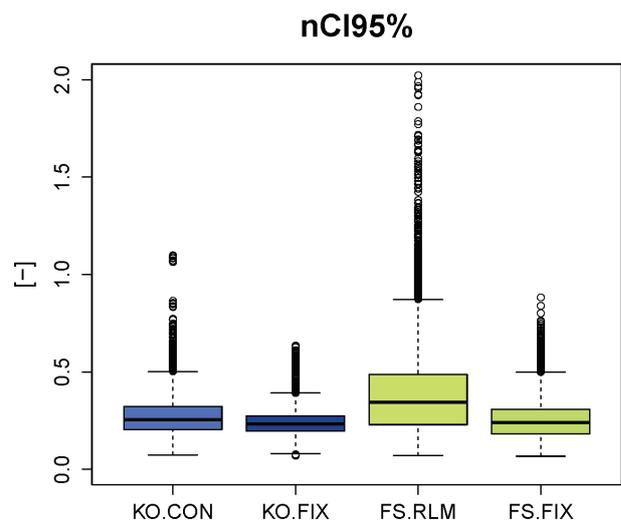


Abbildung 5

Unsicherheit der Ausgleichsschätzung der Niederschlagsquantile gemittelt über alle 133 Stationen, Dauerstufen und Wiederkehrintervalle $T = 1, 10, 20, 50, 100$ a dargestellt als normalisiertes 95%-Konfidenzintervall nCI95% (siehe Gl. 16).

Uncertainty of the generalised estimation of the precipitation quantiles, averaged over all 133 stations, duration levels and return periods $T = 1, 10, 20, 50, 100$ yr shown as normalised confidence interval nCI95% (see Eq. 16).

Tabelle 2
Überblick über die untersuchten Methoden-Daten-Kombinationen.
Overview of the analysed method-data combinations.

Symbol	Beschreibung
Univariate Benchmark-Methoden	
OK[LS]	Ordinary Kriging Interpolation mit langen Reihen LS
INDEX[LS]	Nur LS als Eingabe, Index interpoliert mit OK[LS]
Multivariate Methoden	
OK[LS+SS]	Ordinary Kriging Interpolation von LS und SS
KED[LS SS]	KED Interpolation von LS mit OK[SS] als externe Drift
KED[LS DS]	KED Interpolation von LS mit OK[DS] als externe Drift
KED[LS+SS DS]	KED Interpolation von LS und SS mit OK[DS] als externe Drift
KED[LS SS+DS]	KED Interpolation von LS mit OK[SS] und OK[DS] als externe Drift

aufweisen, wobei OK leichte Vorteile gegenüber dem Index-Verfahren hat. Als bestes Referenzverfahren wird somit hier OK[LS] ausgewählt. Im Weiteren werden hier nur die geostatistischen multivariaten Ansätze diskutiert. Für Ergebnisse bezüglich der multivariaten Index-Verfahren wird auf SHEHU et al. (2022a) verwiesen.

Die multivariaten Ansätze benutzen Zusatzinformationen aus den beiden Kollektiven SS bzw. DS und müssen sich an dem

Basisverfahren OK[LS] messen lassen. Die Abbildung 6 zeigt die Verbesserung der Interpolation im Vergleich zu OK[LS] für verschiedene multivariate Methoden bei Hinzunahme der kurzen bzw. der disaggregierten Reihen. Für das Wiederkehrintervall von $T = 1$ a ist die beste Methode KED[LS|SS+DS], welche Daten aus allen drei Kollektiven integriert. Für die größeren Wiederkehrintervalle $T > 1$ a ergibt sich die stärkste Verbesserung mit KED[LS|KS]. Es ist etwas überraschend, dass die Tagesniederschläge DS in dieser Untersuchung keinen Mehrwert für die Niederschlagsschätzung mit Wiederkehrintervallen $T > 1$ a haben. Für längere Dauerstufen ist mit Einbeziehung der Tageswerte zwar eine gute Interpolationsgüte zu erreichen, aber für kurze mit $D < 6$ h gibt es signifikante systematische Unterschätzungen (nicht dargestellt). Die Verwendung unterschiedlicher Datensätze für verschiedene Dauerstufen wäre eine Alternative, verkompliziert jedoch die gesamte Prozedur. Da für die Bemessung die höheren Wiederkehrintervalle am meisten relevant sind und ein möglichst sparsames Verfahren hinsichtlich Aufwands und Datenbereitstellung wichtig ist, wird als insgesamt am besten geeignete Methode KED[LS|SS] ausgewählt.

Die Tabelle 4 zeigt die absoluten Fehler des besten Verfahrens KED[LS|SS]. Die niedrigen Interpolationsfehler demonstrieren die hohe Güte des Verfahrens. Im Vergleich mit Tabelle 3 wird nochmal ersichtlich, welche Verbesserung durch Hinzunahme der kurzen Reihen mit KED[LS|SS] gegenüber OK[LS] erzielt werden konnte.

Tabelle 3
Güte der Kreuzvalidierung als Median über 133 LS-Standorten für die beiden Benchmark-Methoden OK[LS] und INDEX[LS].
Performance of cross-validation as median over 133 LS locations for the two benchmark methods OK[LS] and INDEX[LS].

	T = 1 a	T = 10 a	T = 20 a	T = 50 a	T = 100 a
OK[LS]					
RMSE [%]	10.12	9.58	9.89	10.45	10.65
PBIAS [%]	2.00	1.90	1.60	1.90	1.90
NSE [-]	0.970	0.977	0.973	0.973	0.972
INDEX[LS]					
RMSE [%]	10.16	10.21	10.55	10.62	10.69
PBIAS [%]	1.90	1.90	2.10	1.80	1.50
NSE [-]	0.970	0.972	0.971	0.971	0.969

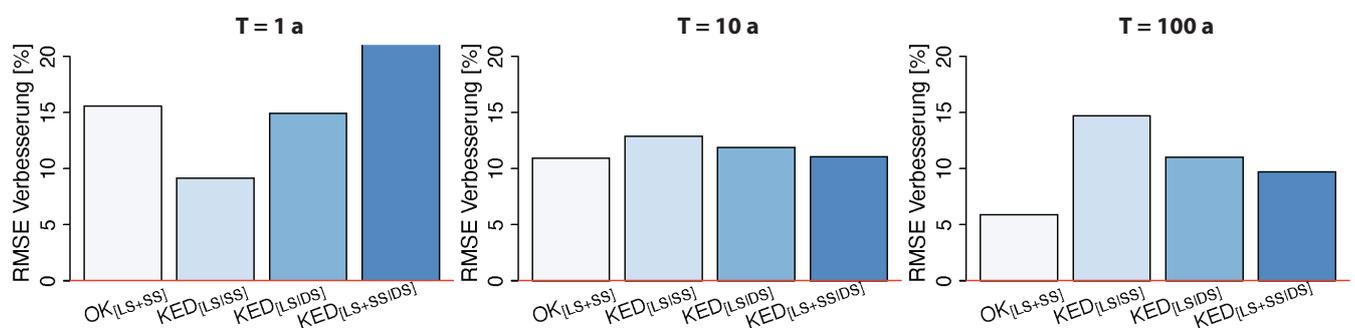


Abbildung 6
Verbesserung der Interpolation im Vergleich zur Referenzmethode OK[LS] für verschiedene Methoden bei Hinzunahme der kurzen Reihen [SS] bzw. der disaggregierten Reihen [DS].
Improvement of the interpolation compared to the benchmark method OK[LS] for different approaches using short series [SS] and/or disaggregated series [DS] additionally.

Tabelle 4

Güte der Kreuzvalidierung als Median über 133 LS-Standorte für die beiden Benchmark-Methoden OK[LS] und INDEX[LS].
Performance of cross-validation as median over 133 LS locations for the two benchmark methods OK[LS] and INDEX[LS].

	T = 1 a	T = 10 a	T = 20 a	T = 50 a	T = 100 a
KED[LS SS]					
RMSE [%]	8.11	8.06	8.24	8.46	8.86
PBIAS [%]	1.00	1.00	0.80	1.00	0.80
NSE [-]	0.982	0.981	0.979	0.979	0.980

4.3 Unsicherheitsschätzung

Die verschiedenen Unsicherheitsquellen für die Methode KED[LS|SS] wurden bereits im Kapitel 3.4 aufgelistet. Dementsprechend werden insgesamt 5 Simulations-Experimente durchgeführt, um die Unsicherheit zu quantifizieren. In Experiment 1 bis 4 werden einzelne Unsicherheiten analysiert und in Experiment 5 die Gesamtunsicherheit des Verfahrens. Für die Ermittlung der Einzelunsicherheiten werden jeweils 100 Wiederholungen zur Parameterschätzung gerechnet und daraus das normalisierte Konfidenzintervall (Gl. 12) bestimmt. Die Abbildung 7 zeigt die Quantifizierung der Unsicherheit für die 133 Stationen mit langen Reihen als Breite des normierten Konfidenzintervalls. Die kleinste Unsicherheit wird durch die VarioGRAMMSchätzung verursacht (Exp. 2). Ebenfalls folgt eine relativ geringe Unsicherheit aus der Interpolation der Zusatzvariable mit OK[SS] (Exp. 3) vor Anwendung des KED. Als Nächstes folgt größenordnungsmäßig die lokale Unsicherheit für die Anpassung der GEV (Exp. 1). Die lokale Schätzung zeigt den Effekt der Koutsoyiannis-Methode, welche zu einem parabolischen Verlauf der Unsicherheit über die Dauerstufen mit den geringsten Werten im Zentrum führt. Die eindeutig größte Unsicherheit resultiert aus der finalen Interpolation der Parameter mit KED[LS|SS] (Exp. 4). Hier ist ein sinusförmiger Verlauf der Unsicherheiten über die Dauern zu erkennen. Ursache dafür ist vermutlich die für die SGS-Simulation erforderliche Transformation der Parameter in die Normalverteilung, welche die Korrelation zwischen den Parametern verstärkt und diese nicht-linearen Effekte bewirkt.

Für die Abschätzung der Gesamtunsicherheit (Exp. 5) werden direkt nur die beiden größten Unsicherheitskomponenten, die

lokale Unsicherheit und die Interpolationsunsicherheit vom KED, kombiniert. Dazu werden 100 x 100 = 10.000 wiederholte Parameterschätzungen gemacht. Indirekt werden damit auch die beiden anderen Unsicherheiten berücksichtigt, da 10.000 Variogrammschätzungen und ebenso vielen Interpolationen der Zusatzvariable nötig sind.

Die Gesamtunsicherheit hängt kaum vom Wiederkehrintervall ab, aber stark und nicht-linear von der Dauer und zeigt einen Verlauf, wie man ihn aus der Kombination der Experimente 1 und 4 erwarten würde. Nach Abbildung 7 variiert die Gesamtunsicherheit im Median aller Stationen zwischen 30 % (± 15 % um den Mittelwert) bei $d = 1.440$ min und 50 % (± 25 % um den Mittelwert) bei $D = 15$ min für die Starkregenschätzung mit einem Wiederkehrintervall von $T = 10$ a. Im Gegensatz zu den geringen Verfahrensfehlern von < 10 % (Tab. 4) ist mit signifikanten Unsicherheiten zu rechnen. Diese resultierten prinzipiell aus dem beschränkten Informationsgehalt der Stichproben hinsichtlich des Extremwertverhaltens der Niederschläge. Längere Zeitreihen und ein dichteres Messnetz könnten zukünftig die Unsicherheiten reduzieren.

5 Räumlich Darstellungen

Nach Interpolation der vier Parameter $\mu, \sigma, \theta, \eta$ mit KED[LS|SS] auf ein Raster von 5 km x 5 km können für beliebige Punkte Starkniederschläge als Funktion von Dauer D und Wiederkehrintervalle T abgeleitet werden. Die Abbildung 8 zeigt beispielhaft Kartendarstellungen der räumlichen Niederschlagsverteilung in Deutschland für ausgewählte D und T .

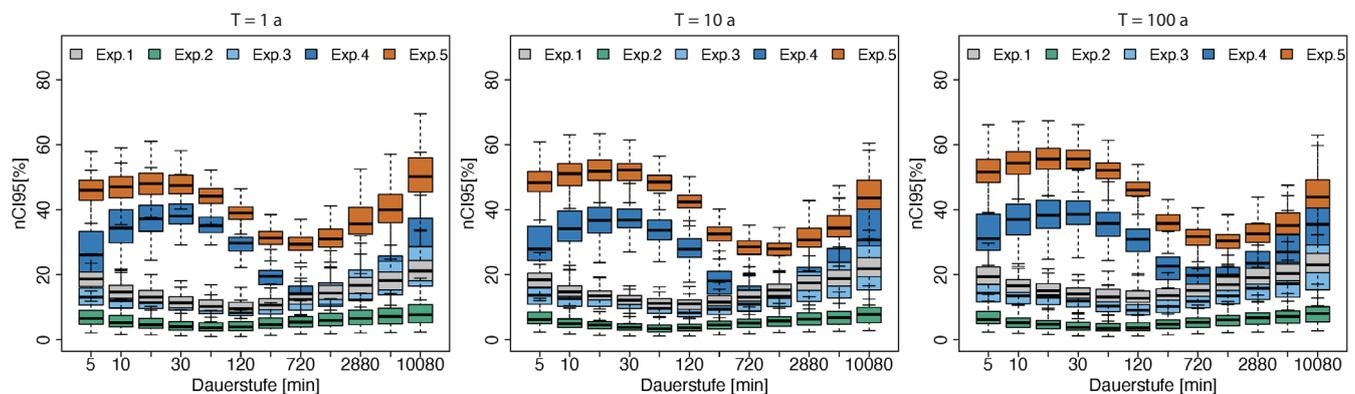


Abbildung 7

Unsicherheiten für die Schätzung der Starkniederschläge an den Standorten der 133 Stationen mit langen Reihen in Abhängigkeit von der Dauer für ausgewählte Wiederkehrintervalle für verschiedene Komponenten des Gesamtverfahrens entsprechend der Simulationsexperimente 1 bis 5.
Uncertainty for the estimation of precipitation at the 133 locations with long series depending on duration for selected return periods for different components of total procedure according to simulation experiments 1 to 5.

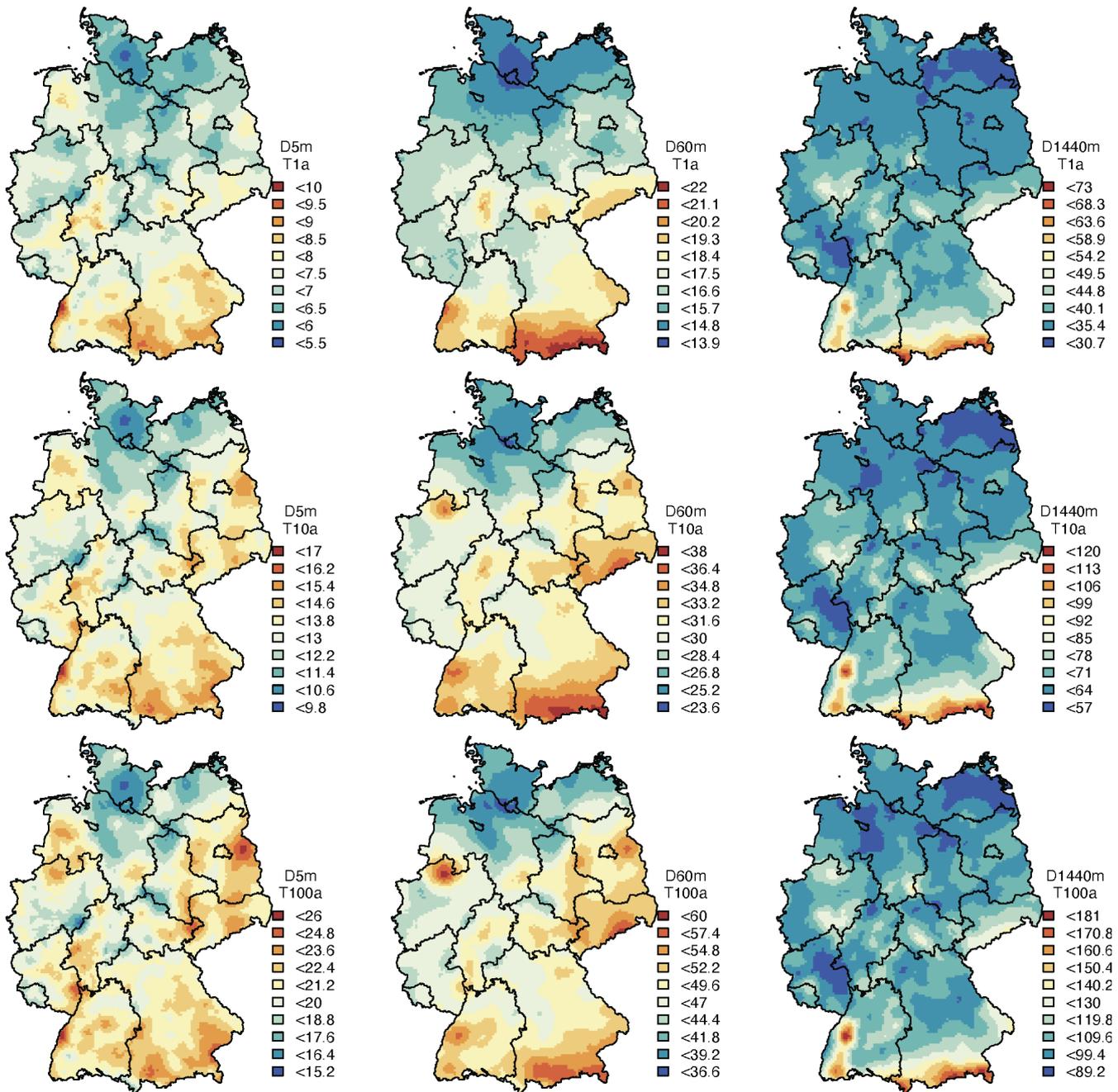


Abbildung 8
 Räumliche Verteilung der Starkniederschläge in Deutschland in mm/D für ausgewählte Dauerstufen D und Wiederkehrintervalle T.
 Spatial distribution of heavy rainfall in Germany in mm/D for different duration levels D and return periods T.

Es ist zu erkennen, dass die räumliche Variabilität mit zunehmender Dauer abnimmt. Für kurze Dauerstufen, vor allem $D = 5$ min, können Extremwerte überall auftreten. Es gibt keine Kopplung an die Topographie. Für Tageswerte mit $D = 24$ h sind hohe Niederschläge vor allem in den Gebirgen zu finden. Die räumlichen Muster für gleiche Dauerstufen, aber unterschiedliches Wiederkehrintervall sind dagegen sehr ähnlich. Nur die Intensitäten sind unterschiedlich. Allerdings fällt bei der Dauerstufe von $D = 1$ h und Wiederkehrintervallen $T > 1$ a ein lokales Maximum bei Münster auf. Dieses resultiert aus dem im Jahre 2014 aufgetretenen sehr seltenen Starkregenereignis und hat signifikanten Ein-

fluss auf die Anpassung der Verteilung. Das zeigt die Problematik der Behandlung von singulären Extremwerten in Stichproben zur Ableitung von Bemessungsniederschlägen auf. Ein grundsätzliches Ausschließen solcher Werte würde zur Unterschätzung, ein Einschließen dagegen zur Überschätzung der Niederschläge führen. Eine Entscheidung kann mithilfe der Unsicherheitsanalysen getroffen werden.

Die Abbildung 9 zeigt die räumliche Verteilung verschiedener Unsicherheiten für die Schätzung der Niederschläge für unterschiedliche Dauerstufen und das Wiederkehrintervall von

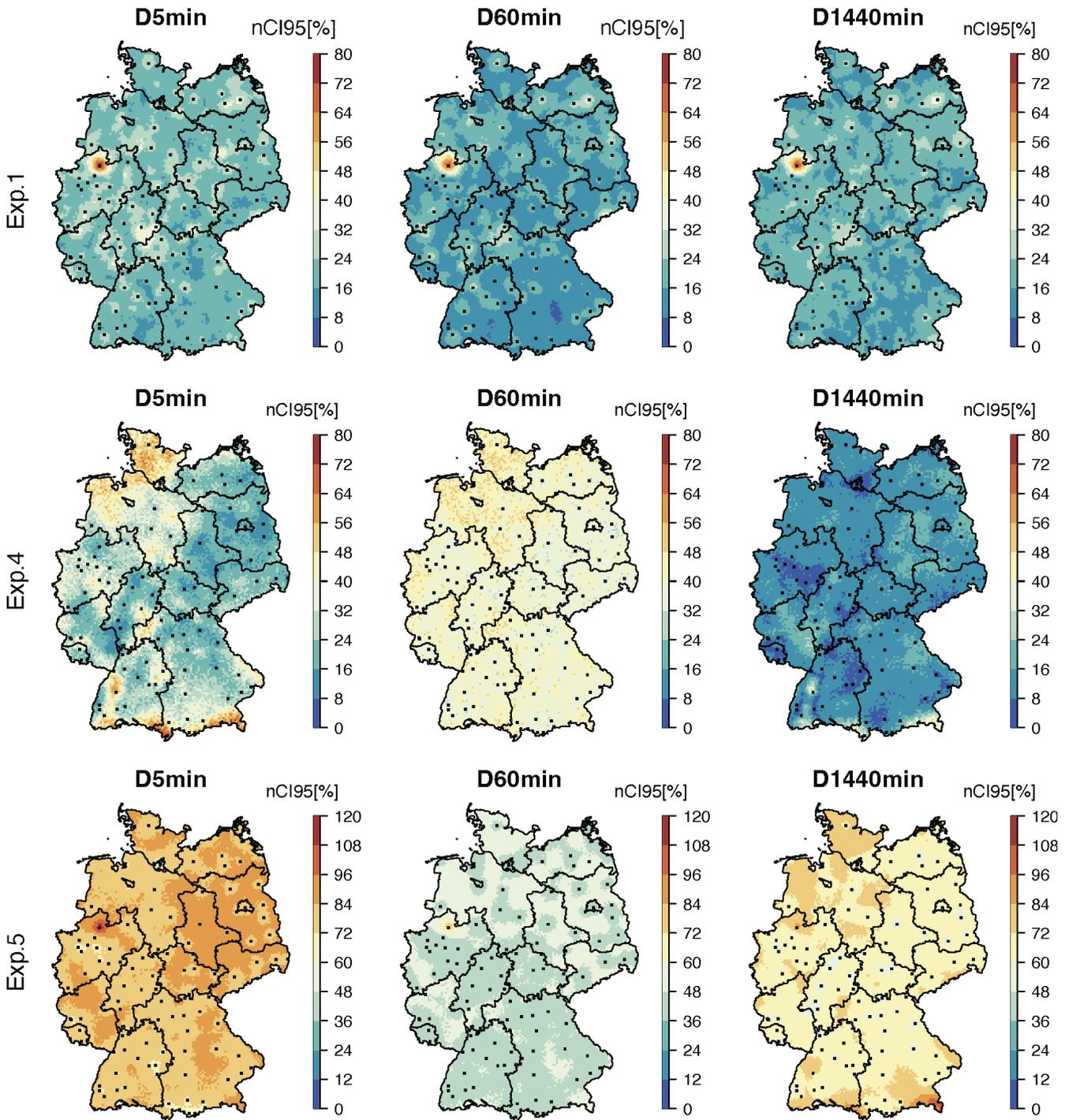


Abbildung 9

Räumliche Verteilung verschiedener Unsicherheiten bei der Niederschlagsschätzung in Deutschland für verschiedene Dauerstufen und ein Wiederkehrintervall von $T = 100$ a; Exp. 1: lokale Unsicherheit (obere Reihe), Exp. 4: Unsicherheit aus der Regionalisierung (mittlere Reihe), Exp 5: Gesamtunsicherheit (untere Reihe). Die Unsicherheit ist dargestellt als normalisiertes 95%-Konfidenzintervall $nCI95\%$ (siehe Gl. 16).

Spatial distribution of uncertainty for the estimation of heavy precipitation in Germany for different duration levels and a return period of $T = 100$ a; Exp. 1: local uncertainty (upper row), Exp. 4: uncertainty from regionalisation (middle row), Exp. 5: total uncertainty (lower row). The uncertainty is shown as normalised confidence interval $nCI95\%$ (see Eq. 16).

$T = 100$ Jahren. Für die Darstellung der lokalen Unsicherheit (Exp. 1) wurden die 100 lokalen Parameterrealisationen separat mit KED[LS][SS] regionalisiert und die Variabilität für jedes Raster ermittelt. Für die Darstellung der regionalen Unsicherheit (Exp. 4) liefert die sequentielle Gauss'sche Simulation (SGS) direkt

100 Realisationen der räumlichen Verteilung. Für die Gesamtunsicherheit (Exp. 5) werden 100 lokale Parametersätze mit jeweils 100 SGS-Simulationen verbunden und liefern 10.000 Karten, aus denen dann die Gesamtvariabilität des Verfahrens abgeleitet werden kann.

Während die räumliche Verteilung der lokalen Unsicherheit (Exp. 1) sehr ähnlich für alle Dauerstufen ist, zeigen die regionale und die Gesamtunsicherheit unterschiedliche Muster für jede Dauerstufe. In den Karten der lokalen Unsicherheit (Exp. 1) fällt deutlich ein Maximum bei Münster auf, welches aus der Berücksichtigung des dort aufgetretenen singulären Ereignisses resultiert. Dieses Maximum ist in den Karten der aus der Regionalisierung resultierenden Unsicherheit (Exp. 4) nicht zu sehen. Es tritt jedoch bei Kombination der lokalen und regionalen Unsicherheit in den Karten der Gesamtunsicherheit (Exp. 5) wieder auf. Damit ist nachgewiesen, dass dieses singuläre Ereignis eine extrem hohe lokale Unsicherheit verursacht. Es sollte für die Erstellung der deutschlandweiten Niederschlagsstatistik ausgeschlossen werden. Insgesamt sind für die Bemessungsniederschläge in Deutschland Unsicherheiten zwischen

± 30 bis 60 % für $D = 5$ min,
 ± 15 bis 45 % für $D = 1$ h und
 ± 20 bis 50 % für $D = 1$ Tag

anzunehmen. Die geringsten Unsicherheiten treten an den Stationsstandorten der langen Reihen und häufig auch in deren Nähe auf, die höchsten da, wo die Messnetze spärlich sind.

6 Schlussfolgerungen und Ausblick

Im Ergebnis der vorliegenden Untersuchungen wurde ein neues Verfahren zur regionalisierten Starkregenstatistik für Deutschland vorgeschlagen. Dieses basiert auf einer lokalen Extremwertanalyse nach KOUTSOYIANNIS et al. (1998) mit anschließender Interpolation der Parameter mit der Methode des Krigings mit Externer Drift (KED) räumlich bezogen auf ein Raster von 5 km x 5 km. Es erlaubt die Ableitung von Bemessungsniederschlägen für beliebige Dauerstufen von $D = 5$ min bis $D = 7$ d und Wiederkehrintervalle zwischen $T = 1$ a und $T = 100$ a. Dabei werden Niederschlagsinformationen aus allen registrierenden Stationen von langen und kurzen Beobachtungsreihen einbezogen. Die Methodik nutzt wissenschaftlich bewährte Verfahren, ist objektiv, klar und nachvollziehbar und zeigt nur geringe Fehler. Erstmals für Deutschland wurde auch eine komplette, alle wichtigen Quellen umfassende quantitative Unsicherheitsanalyse der regionalisierten Starkregenstatistik durchgeführt. Basierend auf den Ergebnissen der Untersuchungen erstellt der DWD eine neue Starkregenstatistik für Deutschland. Notwendige Anschlussuntersuchungen umfassen die Ermittlung von Bemessungsniederschlägen mit Wiederkehrintervallen $T > 100$ a, die Bestimmung von maximal möglichen Niederschlägen und die Flächenabminderung von Punktniederschlägen zur Bestimmung von Gebietsniederschlägen. Ebenfalls wären Untersuchungen wünschenswert, die sich mit Änderungen der Starkregenstatistik als Folge der Klimaänderung beschäftigen. Abschließend ist zu bemerken, dass die vorliegenden Ergebnisse auch eine gute Referenz für künftige Extremwertschätzungen auf Basis von Radardaten liefern.

7 Conclusions and outlook

With these investigations a new method for regionalised heavy rainfall statistic in Germany is proposed. It is based on the local extreme values analysis of KOUTSOYIANNIS et al. (1998) with subsequent interpolation of the parameters on a 5 km x 5 km grid using KED kriging. It allows the derivation of design rainfall for any duration between $D = 5$ min and $D = 7$ d and return periods

between $T = 1$ yr and $T = 100$ yr. All rainfall data from recording stations with short and long observation periods are included. The approach uses scientifically well proven methods, is objective, clear, reproducible and shows only small errors. For the first time in Germany a complete uncertainty analysis of the regionalised design rainfall has been carried out including all relevant sources of uncertainty. Based on the results of this study the German Weather Service is providing a new heavy rainfall statistics for Germany. Necessary subsequent research includes the estimation of rainfall for return periods $T > 100$ yr, the estimation of maximum probable rainfall and the derivation of area reduction factors to estimate areal rainfall from point values. In addition, investigations of the climate impact on the extreme rainfall statistics would be useful. Finally, it should be noted, that these results provide a good reference for future extreme value statistics based on weather radar data.

Datenverfügbarkeit

Die Analysen basieren weitgehend auf frei verfügbaren Niederschlagsdaten des Deutschen Wetterdienstes (DWD) (https://www.dwd.de/DE/klimaumwelt/cdc/cdc_node.html oder auf Anfrage) sowie auf Daten von Partnern, wie Landesbehörden, Wasserverbänden oder kommunalen Wasserbetrieben, die in der Regel nicht frei verfügbar sind. Die entwickelten Softwareprogramme können bei Dr.-Ing. Bora Shehu angefragt werden. Die erstellte Starkregenstatistik KOSTRA-DWD-2020 wird vom DWD veröffentlicht und ist frei zugänglich.

Danksagung

Wir danken der Länderarbeitsgemeinschaft Wasser (LAWA) und dem Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz für die Förderung des Vorhabens MUNSTAR. Außerdem danken wir dem Hauptausschuss Hydrologie und Wasserbewirtschaftung der Deutschen Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V. (DWA AG HW-1.1 Niederschlag) für die begleitende Unterstützung des Projektes. Außerdem wird den zwei anonymen Gutachtern gedankt, die mit ihren nützlichen Hinweisen zur Verbesserung des Artikels beigetragen haben.

Anschriften der Verfasser

Prof. Dr.-Ing. Uwe Haberlandt
 Dr.-Ing. Bora Shehu
 Luisa Thiele, M.Sc.
 Leibniz Universität Hannover
 Institut für Hydrologie und Wasserwirtschaft
 Appelstraße 9a, 30167 Hannover
 haberlandt@iww.uni-hannover.de

Dr. Winfried Willems
 Henrike Stockel, M.Sc.
 Ingenieurhydrologie, Angewandte
 Wasserwirtschaft und Geoinformatik IAWG
 Alte Landstr. 12 – 14, 85521 Ottobrunn
 willems@iawg.de

Dr. Thomas Deutschländer
 Thomas Junghänel
 Jennifer Ostermüller
 Deutscher Wetterdienst
 Frankfurter Straße 135, 63067 Offenbach
 thomas.deutschlaender@dwd.de

Literaturverzeichnis

- BURN, D.H. (2014): A framework for regional estimation of intensity-duration-frequency (IDF) curves. *Hydrol. Process.*, 28(14). DOI:10.1002/hyp.10231.
- CERESSETTI, D., E. URSU, J. CARREAU, S. ANQUETIN, J.D. CREUTIN, L. GARDES, S. GIRARD & G. MOLINIÉ (2012): Evaluation of classical spatial-analysis schemes of extreme rainfall. *Nat. Hazards Earth Syst. Sci.*, 12(11), 3229–3240. DOI:10.5194/nhess-12-3229-2012.
- COLES, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer, London
- DAVISON, A.C. & D.V. HINKLEY (2005): *Bootstrap methods and their applications*. Cambridge University Press, New York, 582 pp.
- DWA-A531 (2012): Arbeitsblatt: Starkregen in Abhängigkeit von Wiederkehrzeit und Dauer. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V., Hennef.
- FISCHER, S. & A.H. SCHUMANN (2018): Berücksichtigung von Starkregen in der Niederschlagsstatistik. *Hydrol. und Wasserbewirtschaftung*, 62(4), 221–240. DOI:10.5675/HyWa.
- GOOVAERTS, P. (1997): *Geostatistics for natural resources evaluation*. Oxford University Press, New York, Oxford, 483 pp.
- HELSEL, D.R., R.M. HIRSCH, K.R. RYBERG, S.A. ARCHFIELD & E.J. GILROY (2020): *Statistical methods in water resources*. 4-A3, Reston, VA.
- HOSKING, J.R.M. & J.R. WALLIS (1997): *Regional Frequency Analysis*. Cambridge University Press.
- KOUTSOYIANNIS, D., D. KOZONIS & A. MANETAS (1998): A mathematical framework for studying rainfall intensity-duration frequency relationships. *J. Hydrol.*, 206(1–2), 118–135. DOI:10.1016/S0022-1694(98)00097-3.
- KOUTSOYIANNIS, D. (2004a): Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrol. Sci. J.*, 49(4), 575–590. DOI:10.1623/hysj.49.4.575.54430.
- KOUTSOYIANNIS, D. (2004b): Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrol. Sci. J.*, 49(4), 591–610. DOI:10.1623/hysj.49.4.591.54424.
- JOHNSON, F. & A. SHARMA (2017): Design Rainfall. in *Handbook of Applied Hydrology*, edited by V. P. Singh, pp. 125–3 to 125-13, McGraw-Hill, New York.
- JUNGHÄNEL, T., F. BÄR, T. DEUTSCHLÄNDER, U. HABERLANDT, I. OTTE, B. SHEHU, H. STOCKEL, K. STRICKER, L.-B. THIELE & W. WILLEMS (2022): *Methodische Untersuchungen zur Novellierung der Starkregenstatistik für Deutschland (MUNSTAR)*. Synthesebericht. 95 pp.
- LISNIAK, D., J. FRANKE & C. BERNHOFER (2013): Circulation pattern based parameterization of a multiplicative random cascade for disaggregation of observed and projected daily rainfall time series. *Hydrol. Earth Syst. Sci.*, 17(7), 2487–2500. DOI:10.5194/hess-17-2487-2013, 2013.
- MALITZ, G. & H. ERTEL (2015): *KOSTRA-DWD-2010: Starkniederschlags-höhen für Deutschland. Bezugszeitraum 1951 bis 2010*. Offenbach am Main: Deutscher Wetterdienst.
- MADSEN, H., K. ARNBJERG-NIELSEN & P.S. MIKKELSEN (2009): Update of regional intensity-duration-frequency curves in Denmark: Tendency towards increased storm intensities. *Atmos. Res.*, 92(3). DOI:10.1016/j.atmosres.2009.01.013.
- OLSSON, J. (1998): Evaluation of a scaling cascade model for temporal rainfall disaggregation. *Hydrol. Earth Syst. Sci.*, 2(1), 19–30. DOI:10.5194/hess-2-19-1998.
- OVEREEM, A., A. BUIHAND & I. HOLLEMAN (2008): Rainfall depth-duration-frequency curves and their uncertainties. *J. Hydrol.*, 348(1–2), 124–134. DOI:10.1016/j.jhydrol.2007.09.044.
- PEBESMA, E.J. (2004): Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30: 683–691. DOI:10.1016/J.Cageo.2004.03.012.
- SHEHU, B., W. WILLEMS, H. STOCKEL, L. THIELE & U. HABERLANDT (2022a): Regionalisation of Rainfall Depth-Duration-Frequency curves in Germany. *Hydrol. Earth Syst. Sci. Discuss.* [preprint]. DOI: 10.5194/hess-2022-118, in review.
- SHEHU, B. & U. HABERLANDT (2022b): Uncertainty estimation of regionalised depth-duration-frequency curves in Germany. *Hydrol. Earth Syst. Sci. Discuss.* [preprint]. DOI: 10.5194/hess-2022-254, in review.
- UBOLDI, F., A.N. SULIS, C. LUSSANA, M. CISLAGHI & M. RUSSO (2014): A spatial bootstrap technique for parameter estimation of rainfall annual maxima distribution. *Hydrol. Earth Syst. Sci.*, 18(3), 981–995. DOI:10.5194/hess-18-981-2014.
- VAN DE VYVER, H. (2018): A multiscaling-based intensity-duration-frequency model for extreme precipitation. *Hydrol. Process.*, 32(11). DOI:10.1002/hyp.11516, 2018.
- WARD, J.H. (1963): Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58: 236–244. DOI:10.1080/01621459.1963.10500845.
- WILLEMS, W., H. STOCKEL, U. HABERLANDT, B. SHEHU, T. DEUTSCHLÄNDER & T. JUNGHÄNEL (2023): *Untersuchungen zur Stationarität extremer Niederschläge in Deutschland (pers. Kommunikation, in Vorbereitung)*.
- WINTERRATH, T., C. BRENDEL, M. HAFER, T. JUNGHÄNEL, A. KLAMETH, E. WALAWENDER, E. WEIGL & A. BECKER (2017): *Erstellung einer radargestützten Niederschlagsklimatologie. Berichte des Deutschen Wetterdienstes*. Offenbach.