

Albert-Ludwigs-Universität Freiburg
Institut für Hydrologie

Simon Klaffl

Kombinierte Wahrscheinlichkeiten (*joint probabilities*)
von Starkniederschlägen und
Vorfeuchtebedingungen in Baden - Württemberg

Diplomarbeit unter der Leitung von
Prof. Dr. Markus Weiler
Koreferent: Dr. Christoph Külls

Freiburg i. Br., März 2011

Ehrenwörtliche Erklärung:

Hiermit erkläre ich, dass die Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt wurde.

Ort, Datum

Unterschrift

Tabellenverzeichnis	iii
Abbildungsverzeichnis	iv
Anhangsverzeichnis.....	vii
Summary	viii
Zusammenfassung	ix
1 Einleitung.....	1
1.1 Fragestellung	2
1.2 Zielsetzung	3
1.3 Übersicht / Stand der Forschung.....	4
2 Datengrundlage	11
2.1 Niederschlagsdaten.....	11
2.2 Bodenfeuchtedaten.....	12
2.3 Zusätzliche Daten	13
3 Methodik & Vorgehensweise	17
3.1 Datenvorbereitung	17
3.2 Erstellung der Stichprobenkollektive	18
3.2.1 Jährliche und partielle Serie	19
3.2.2 Saisonale Einteilung	20
3.2.3 Zwischenfazit	22
3.3 Korrelationsanalyse.....	23
3.3.1 Kendall's Tau.....	24
3.3.2 Autokorrelation des Niederschlags	27
3.4 Univariate Randverteilungen	28
3.4.1 Anpassung theoretischer Verteilungsfunktionen.....	30
3.4.1.1 Maximum Likelihood Estimation.....	30
3.4.1.2 Kolmogorov - Smirnov Test	31
3.4.2 Generalisierte Extremwertverteilung (Niederschlag)	32
3.4.3 Beta-Verteilung (Bodenfeuchte)	33

3.5 Joint probabilities	34
4 Ergebnisse	41
4.1 Korrelationsanalyse.....	41
4.1.1 Korrelation von N und θ_v	41
4.1.2 Autokorrelation des Niederschlags	45
4.2 Anpassung der Randverteilungen.....	47
4.2.1 Niederschlag	47
4.2.2 Bodenfeuchte.....	49
4.2.3 Univariate Wiederkehrintervalle	53
4.3 Joint probabilities	57
5 Diskussion	63
5.1 Korrelationen zwischen N und θ	63
5.2 Randverteilungen.....	65
5.3 Joint probabilities	68
6 Fazit & Ausblick	71
Literaturverzeichnis.....	72
Anhang	79

Tabellenverzeichnis

Tabelle 2.1: Die zur exemplarischen Darstellung verschiedener Grafiken ausgewählten Modellflächen.	14
Tabelle 3.1: Kritische Werte D_n für $d_{\max}$ des Kolmogorov-Smirnov Tests bei gegebenem Signifikanzniveau α . (SCHÖNWIESE, 2006).....	32
Tabelle 4.1: Momente der Häufigkeitsverteilungen von τ über die Gesamtfläche.	44
Tabelle 4.2: Quantile der Momente und Parameter der GEV über alle Teilflächen ($n = 65\,645$).....	48
Tabelle 4.3: Quantile der Momente und Parameter der Beta-Verteilung über alle Teilflächen ($n = 65\,645$). Hier ist zu beachten, dass E und Sd in %FK angegeben sind, während sich α und β auf das Intervall $[0,1]$ der Betaverteilung beziehen.	51

Abbildungsverzeichnis

Abbildung 1.1 a: Bivariate Dichtefunktion $f(q,v)$ für maximale Abflussrate Q und Gesamtvolumen V eines Hochwasserereignisses, b: Konturplot von $f(q,v)$. (YUE & RASMUSSEN, 2002)	6
Abbildung 1.2 a: Bivariate, kumulative Wahrscheinlichkeitsverteilung $F'(Q \geq q, V \geq v)$ für maximale Abflussrate Q und Gesamtvolumen V eines Hochwasserereignisses, b: Konturplot von F' . (YUE & RASMUSSEN, 2002) ..	7
Abbildung 1.3: Vergleich von beobachteten (+), kontinuierlich simulierten (o) und durch Monte Carlo Simulation von Einzelereignissen erzeugte Häufigkeitsverteilungen der maximalen Abflussrate (graue Linien), sowie die angepasste GEV der kontinuierlichen Simulation (schwarze Linie) mit 95% Konfidenzintervallen (blau gestrichelte Linien). (KJELDTSEN et al., 2010)	9
Abbildung 2.1: CORINE - Landnutzungsklassen (Waboa, 2007).....	13
Abbildung 2.2: Zur Regionalisierung der Niederschlagsdaten benutzte Stationen des DWD. (WABOA, 2007)	15
Abbildung 2.3: Feldkapazitätsklassen der Böden in Baden-Württemberg und Lage der ausgewählten Teilflächen A – F aus Tabelle 2.1. (WABOA, 2007) ..	16
Abbildung 3.1: Niederschlags- und Bodenfeuchtezeitreihen, exemplarisch für das Jahr 1984, Teilfläche B.	18
Abbildung 3.2: Saisonalität der Starkniederschläge in Quartalsdarstellung. (WABOA, 2007)	20
Abbildung 3.3: Box-Plot des Jahresganges der Bodenfeuchte an den 6 ausgewählten Stationen A - F (siehe Tabelle 2.1) über den Zeitraum von 39 Jahren.	21
Abbildung 3.4: Häufigkeitsverteilungen der Bodenfeuchte-Tageswerte über die gesamte Zeitreihe (a) und jeweils getrennt für das hydrologische Sommer- (b) und Winterhalbjahr (c) an den Teilflächen A und D (siehe Tabelle 2.1).	22
Abbildung 3.5: Einteilung des X-Y Wahrscheinlichkeitsraumes zur Veranschaulichung der kombinierten, bivariaten Wahrscheinlichkeiten zweier Variablen X und Y. (YUE & RASMUSSEN, 2002)	36
Abbildung 3.6: Gemeinsame Wiederkehrintervalle $T'(x, y)$ gegen Abhängigkeitsparameter m des Gumbel-Logistic-Modells. $T'(x, y) \sim m$ für $T_X = T_Y = 2, 5, \dots, 100$ Jahre.	40
Abbildung 4.1: Scatterplot der saisonalen $N - \theta_0$ Stichproben für die Punkte A, C und E (siehe Tabelle 2.1).	42
Abbildung 4.2: Rangkorrelationskoeffizient τ von saisonalen Niederschlagsmaxima und Vorfeuchte für Sommer (links) und Winter	

(rechts). Schraffierte Flächen zeigen signifikante Korrelationen bei $\alpha = 5\%$.

43

- Abbildung 4.3: Häufigkeitsverteilung des Rangkorrelationskoeffizienten τ (graue Balken) und angenommene Normalverteilung mit $\mu = 0$ gemäß der Nullhypothese $H_0: \tau = 0$ (durchgezogene Linie). a) für Sommer und Winter ohne Bedingung, b) mit Bedingung $N_{t-1} < 5$ mm ("Vortag trocken"). 44
- Abbildung 4.4: Autokorrelation 1. Grades ρ_1 des Niederschlags gegen Rangkorrelation τ von Niederschlag N und Vorfeuchte θ für Sommer (links) und Winter (rechts). 45
- Abbildung 4.5: Autokorrelationskoeffizient ρ ($k = 1$ Tag) der Niederschlags - Tagessummen über 39 Jahre, aufgeteilt in Sommer (links) und Winter (rechts)..... 46
- Abbildung 4.6: Empirische Wahrscheinlichkeitsverteilung ($\times$) und GEV ($-$) der saisonalen Niederschlagsmaxima mit μ = Lage- (*loc*), σ = Streu- (*scale*) und ξ = Formparameter (*shape*) sowie dem p-Wert des KS-Tests für die Punkte A, C und D..... 47
- Abbildung 4.7: Häufigkeitsverteilung der p-Werte des K-S Tests für die Anpassung der GEV an die empirischen Verteilungen der saisonalen Niederschlagsmaxima für Sommer (links) und Winter (rechts). Die Nullhypothese wurde bei $\alpha = 0.05$ für keine der Teilflächen abgelehnt. .. 48
- Abbildung 4.8: Räumliche Verteilung von Erwartungswert und Standardabweichung der GEV für saisonale Niederschlagsmaxima im Sommer (links) und Winter (rechts). 49
- Abbildung 4.9: Empirische Wahrscheinlichkeitsverteilung ($\times$) und Beta-Verteilung ($-$) der Vorfeuchtebedingungen θ_v mit den Parametern $\alpha = \text{shape1}$ und $\beta = \text{shape2}$ sowie dem p-Wert des KS-Tests für die Punkte A, B und E. 50
- Abbildung 4.10: Häufigkeitsverteilung der p-Werte des K-S Tests für die Anpassung der Beta-Verteilung an die Vorfeuchtebedingungen θ_v für Sommer (links) und Winter (rechts). H_0 wurde bei $\alpha = 0.05$ im Sommer für insgesamt 0.8%, im Winter für ca. 20.2% der Gesamtfläche abgelehnt. 51
- Abbildung 4.11: Räumliche Verteilung von Erwartungswert E und Standardabweichung Sd der Beta-Verteilung für die Vorfeuchtebedingungen θ_v im Sommer (links) und Winter (rechts)..... 52
- Abbildung 4.12: Häufigkeitsverteilung der Wiederkehrintervalle T' [a] über die Gesamtfläche für $N > 40$ und 50 mm/24h (Sommer) und $N > 30$ und 50 mm/24h (Winter). 53
- Abbildung 4.13: Häufigkeitsverteilung der Wiederkehrintervalle T' [a] über die Gesamtfläche für $\theta_v > 60$ und 70 %FK (Sommer) und $\theta_v > 80$ und 90 %FK (Winter). 54
- Abbildung 4.14: Jährlichkeiten des Niederschlags für $N > 40$ und 50 mm/24h im Sommer (links) und $N > 30$ und 40 mm/24h im Winter (rechts). 55

Abbildung 4.15: Wiederkehrintervalle der Vorfeuchtebedingungen $\theta_v > 60$ und 70 %FK im Sommer (links) und $\theta_v > 80$ und 90 %FK im Winter (rechts)...	56
Abbildung 4.16: Empirische Überschreitungswahrscheinlichkeiten $F'(x,y)$ für die Punkte B, D und E.	58
Abbildung 4.17: Theoretische Überschreitungswahrscheinlichkeiten $F'(x,y)$ für die Punkte B, D und E mit BV und GEV als Randverteilungen.....	58
Abbildung 4.18: Gemeinsame Wiederkehrintervalle $T'(X > x, Y > y)$ aus den Randverteilungen BV für θ_v und GEV für N für die Punkte B, D und E. ...	59
Abbildung 4.19: Empirische, interpolierte Dichtefunktionen (<i>density traces</i>) der gemeinsamen Jährlichkeiten für jeweils 6 Kombinationen aus θ_v und N über alle Teilflächen.	59
Abbildung 4.20: Gemeinsame Wiederkehrintervalle von N und θ_v jeweils für $N > 40$ / > 50 mm sowie $\theta_v > 60$ / $> 70\%$ FK im Sommer.....	60
Abbildung 4.21: Gemeinsame Wiederkehrintervalle von N und θ_v jeweils für $N > 30$ / > 40 mm sowie $\theta_v > 80$ / $> 90\%$ FK im Winter.	61

Anhangsverzeichnis

Anhang 1: Zu Abbildung 4.1: Scatterplot der saisonalen $N - \theta_v$ Stichproben...	79
Anhang 2: Zu Abbildung 4.6: Empirische Wahrscheinlichkeitsverteilung ($\times$) und GEV ($-$) der saisonalen Niederschlagsmaxima mit μ = Lage- (loc), σ = Streu- (scale) und ξ = Formparameter (shape), sowie dem p-Wert des KS-Tests.	80
Anhang 3: Zu Abbildung 4.9: Empirische Wahrscheinlichkeitsverteilung ($\times$) und Beta-Verteilung ($-$) der Vorfeuchtebedingungen θ_v mit den Parametern α = shape1 und β = shape2 sowie dem p-Wert des KS-Tests.	81
Anhang 4: Zu Abbildung 4.16 : Empirische Überschreitungswahrscheinlichkeiten $F'(x,y)$.	82
Anhang 5: Zu Abbildung 6.1: Theoretische Überschreitungswahrscheinlichkeiten $F'(x,y)$ mit BV und GEV als Randverteilungen.	83
Anhang 6: Zu Abbildung 4.18: Gemeinsame Wiederkehrintervalle $T'(X > x, Y > y)$ aus den Randverteilungen BV für θ_v und GEV für N .	84
Anhang 7: Empirische Quantile (= Unterschreitungswahrscheinlichkeiten) der Vorfeuchtestichproben ($n = 39$) des Sommerhalbjahres.	85
Anhang 8: Empirische Quantile (= Unterschreitungswahrscheinlichkeiten) der Vorfeuchtestichproben ($n = 39$) des Winterhalbjahres.	86

Summary

This work is dealing with the development of a method to estimate the pre-event soil moisture conditions for event-based runoff modelling via bivariate frequency analysis. The input data consists of 39-year time series of measured and regionalized daily rainfall amounts and daily soil moisture conditions, modelled by the groundwater-recharge model GWN-BW for the whole area of Baden-Württemberg.

In advance, some basic concepts and common applications of bivariate frequency analysis are presented.

In a first step, the sample time series were selected by extracting the absolute seasonal daily rainfall maximum N for each of the time series for every model plot and then taking the soil moisture value θ_0 of the previous day as pre-event soil moisture. For these data samples, a Kendall rank-correlation analysis was performed but no significant correlations were found for the major part of the area. However, some diffuse correlation patterns could be detected at the 5% significance level, but these were mostly caused by strong autocorrelations in the rainfall time series.

As theoretical marginal distributions for the seasonal maximum rainfall series and the pre-event soil moisture conditions, the Beta-Distribution (BV) and the Generalized Extreme Value Distribution (GEV) were selected and fitted to the empirical distribution functions via Maximum Likelihood Estimation, respectively. The goodness of fit was tested by means of the Kolmogorov-Smirnov Test, and the null-hypothesis was only rejected for the Beta-Distribution in some minor parts of the country in winter and some randomly distributed areas in summer at the 5% significance level. For the GEV no test was rejected.

From these distribution functions, first-order moments and the univariate return periods for some representative values of rainfall and pre-event soil moisture were calculated and their spatial distribution was examined. On a large scale view, there is a clear connection between mean annual precipitation and average pre-event soil moisture, while the fine scale patterns cannot be clearly linked to certain physical or climatic factors.

The joint return periods, stochastically calculated for the independent case, proved to be very ambiguous and unfortunately not provided the expected information gain.

Keywords: joint probability, joint return period, conditional probability distribution, initial conditions, pre-event soil moisture, seasonal maximum precipitation, frequency analysis, statistical hydrology

Zusammenfassung

Diese Arbeit befasst sich mit der Entwicklung einer Methode, um mittels bivariater Statistik eine Abschätzung der Vorfeuchtebedingungen für ereignisbasierte Niederschlags-Abfluss – Modelle vornehmen zu können. Die verwendeten Eingangsdaten bestehen aus 39-jährigen Zeitreihen (1971 - 2009) regionalisierter 24h - Niederschlagshöhen und modellierter Tageswerte der Bodenfeuchte. Diese stammen aus Simulationen des Grundwassermodells GWN-BW, welche über die Fläche Baden-Württembergs durchgeführt wurden.

Vorab wurden einige Anwendungsmöglichkeiten sowie die theoretischen Grundlagen und Konzepte der (bivariaten) Extremwertanalyse vorgestellt und besprochen.

Im ersten Schritt der Analyse wurden die Stichprobenkollektive erstellt, indem aus jeder Zeitreihe der einzelnen Modellflächen die absoluten Maxima der Niederschlagshöhe N jeweils für das hydrologische Sommer- und Winterhalbjahr zusammen mit den Bodenfeuchte - Tageswerten θ_v des jeweiligen Vortags ausgewählt wurden. In einer Rangkorrelationsanalyse nach Kendall konnte für die erhaltenen N - θ_v Serien eine weitgehende Unabhängigkeit festgestellt werden. Geringe Flächenanteile wiesen hier zwar diffuse, aber signifikante Korrelationsmuster auf, diese können jedoch größtenteils mit dem zeitlichen Niederschlagsverhalten (Autokorrelation) begründet werden.

Anschließend wurden als theoretische Randverteilungsfunktionen die Generalisierte Extremwertverteilung (GEV) für die N - Serien und die Beta-Verteilung (BV) für die θ_v - Serien ausgewählt und mittels *Maximum Likelihood Estimation* an die empirischen Verteilungen angepasst. Die Güte der Anpassung wurde mithilfe des Kolmogorov-Smirnov Tests überprüft.

Aus diesen Verteilungsfunktionen wurden dann zunächst die univariaten Jährlichkeiten bestimmter Niederschlags und Vorfeuchtwerte ermittelt und deren räumliche Verteilung untersucht. Großskalig betrachtet konnte hier eine enge Verbindung der mittleren Vorfeuchtebedingungen mit den mittleren Niederschlagshöhen beobachtet werden, während sich die kleinräumigen Muster nicht eindeutig bestimmten Faktoren zuordnen lassen und allgemein ein hohes Maß an Emergenz aufweisen.

Die gemeinsamen (bivariaten) Jährlichkeiten (*joint probabilities*), welche stochastisch für den Fall unabhängiger Variablen für N und θ_v berechnet wurden, erwiesen sich als nur sehr eingeschränkt aussagekräftig und erbrachten leider nicht den erwarteten Informationsgewinn.

1 Einleitung

Die quantitative Erfassung und Vorhersage von Hochwasserereignissen sowie die Ermittlung entsprechender Bemessungswerte zur Planung wasserwirtschaftlicher Anlagen und der Entwicklung von Schutzmaßnahmen gehören mitunter zu den wichtigsten und schwierigsten Aufgaben der Hydrologie. Die neben den physikalisch begründeten Ansätzen wie der *Probable Maximum Precipitation* (PMP) Methode (DYCK & PESCHKE, 1995; FRANCHINI et al., 1996) angewendeten stochastischen Konzepte stützen sich auf die statistische und wahrscheinlichkeitstheoretische Analyse aller Informationen über die in der Vergangenheit beobachteten Hochwasserereignisse. Dabei ist je nach Aufgabenstellung die Abschätzung geeigneter Kennwerte wie beispielsweise Scheitelabfluss, Gesamtvolumen, Maximalpegel oder Hochwasserdauer für eine bestimmte Eintritts- bzw. Überschreitungswahrscheinlichkeit P_{ii} und dem daraus abgeleiteten Wiederkehrintervall T von besonderem Interesse. Betrachtet man diese als unabhängige Zufallsvariablen aus einer stationären und homogenen Grundgesamtheit und es existieren für einen Gewässerabschnitt ausreichend lange Beobachtungszeitreihen ($> 20 - 30$ Jahre), sodass von einer repräsentativen Stichprobe ausgegangen werden darf, kann dem aus der Zeitreihe entnommenen Hochwasserkollektiv eine Verteilungsfunktion (empirisch oder theoretisch) zugeordnet bzw. angepasst werden. Daraus können Größenordnung und Wahrscheinlichkeit der sowohl innerhalb des durch Beobachtungen belegten Zeitraum liegenden, als auch – durch Extrapolation einer theoretischen Verteilungsfunktion – der außerhalb des Beobachtungszeitraumes stattfindenden, sehr seltenen und großen Hochwasserereignissen abgeschätzt werden. Die Voraussetzungen *Stationarität* und *Homogenität* bedeuten hier, dass die Grundgesamtheit über den gesamten Inter- und Extrapolationszeitraum unverändert bleibt und einem einheitlichen hydrologischem Regime entstammt (DYCK & PESCHKE, 1995).

Stehen, wie es häufig der Fall ist, für ein Einzugsgebiet jedoch gar keine oder nur kurze und unzureichende Beobachtungsreihen zur Verfügung, oder kann

das Kriterium der Stationarität aufgrund von Landnutzungsänderungen oder zunehmenden wasserbaulichen Eingriffen in das hydrologische Regime während der Datenerhebung nicht erfüllt werden, so können durch den Einsatz geeigneter Niederschlags-Abfluss-Modelle auch wesentlich häufiger zur Verfügung stehende Niederschlagsdaten oder deren Extremwertverteilungen verwendet werden, um die daraus resultierenden Hochwasserereignisse unter Berücksichtigung verschiedener Einzugsgebietsparameter abzuleiten (BERAN & SUTCLIFFE, 1972; NANDAKUMAR et al., 1994; RAHMAN et al., 2002).

Kontinuierliche Niederschlags-Abfluss-Modelle sind zwar in der Lage, die Vorbedingungen eines Hochwasserereignisses aus der laufenden Simulation zu ermitteln, d.h. in einer Initialisierungsphase werden die Zustandsvariablen solange berechnet, bis sie von den willkürlich gewählten Anfangswerten unabhängig sind, sie benötigen meist jedoch umfassende und vollständige Zeitreihen meteorologischer Eingangsdaten und detaillierte Informationen über die hydrodynamischen und physikalischen Eigenschaften von Boden, Vegetation und Gewässernetz. Daher werden für operationelle Echtzeit- und Vorhersageanwendungen in der Praxis bevorzugt ereignisbasierte Abflussmodelle eingesetzt (BERTHET et al., 2009). Diese besitzen einen entscheidenden Vorteil gegenüber den kontinuierlich arbeitenden Modellen: Sie benötigen nur Eingangsdaten auf Ereignisskala, da sie auch nur die unmittelbar an der Hochwasserentstehung beteiligten Prozesse berücksichtigen. Dies erfordert zwar insgesamt weniger Parameter und erleichtert somit auch die Kalibrierung, ein großer Nachteil besteht hier jedoch darin, dass die Initialisierung der Anfangsbedingungen durch ein einziges Set von mehr oder weniger willkürlich gewählten, als repräsentativ erachteten Eingangsvariablen vorgenommen werden muss (TRAMBLAY et al., 2010).

1.1 Fragestellung

Die häufig zur Modellierung von Bemessungshochwässern aus Bemessungsniederschlägen verwendeten Ansätze (*design event* - oder *design flood approach*)

basieren letztlich auf dem funktionalen Zusammenhang $\text{Direktabfluss} = \text{Starkniederschlag} - \text{Gebietsrückhalt}$. Die größten Schwierigkeiten ergeben sich folglich aus der Bestimmung des aktuellen Gebietsrückhaltes (i. e. dem Füllungszustand der Gebietsspeicher zu Beginn des Ereignisses). Oft wird dabei die probabilistische Natur der damit verbundenen Eingangsparameter bzw. -variablen (z.B. Bodenfeuchte) – mit Ausnahme der Niederschlagshöhe – nicht oder nur ungenügend berücksichtigt, indem beispielsweise vorausgesetzt wird, dass für ein Starkniederschlagsereignis mit einem Wiederkehrintervall $T > 1$ Jahr die Sättigung des Gebietes prinzipiell weitgehend erreicht ist (DYCK & PESCHKE, 1995). Es ist allgemein bekannt, dass ein Niederschlagsereignis einer bestimmten Höhe, welches während einer Trockenperiode als irrelevant betrachtet werden kann, zu einer feuchteren Jahreszeit, wenn große Teile des Einzugsgebietes bereits gesättigt sind, ein gefährliches Hochwasser erzeugen kann. Betrachtet man nun die höchst unterschiedlichen möglichen Vorfeuchteverteilungen in einem Einzugsgebiet, so wird ersichtlich, dass es generell nicht möglich ist, einer bestimmten Niederschlagshöhe ein einhergehendes Hochwasser mit der selben Jährlichkeit zuzuordnen. Eine derart vereinfachte Herangehensweise birgt die Gefahr erheblicher systematischer Fehler in der Abschätzung der Hochwasserkennwerte für bestimmte Wiederkehrintervalle T . Daher erscheint es sinnvoll, die Wahrscheinlichkeiten extremer Niederschlagsereignisse in Kombination mit den Wahrscheinlichkeiten der potentiellen Vorfeuchtebedingungen zu betrachten, um möglicherweise präzisere Aussagen über Ausmaß und Jährlichkeit der daraus resultierenden Hochwasserereignisse treffen zu können (RAHMAN et al., 2002; YUE & RASMUSSEN, 2002; MARTINA et al., 2005; KJELDSSEN et al., 2010).

1.2 Zielsetzung

Das Ziel dieser Studie soll eine großräumige Analyse der bivariaten Verteilungen modellierter Vorfeuchtebedingungen und regionalisierten 24h- Niederschlagshöhen sowie die Identifikation möglicher Zusammenhänge zwischen Vorfeuchte und extremen Niederschlägen sein.

Dazu werden, unter Anwendung verschiedener Vorbedingungen und unter Berücksichtigung saisonaler Effekte, zunächst geeignete Stichprobenkollektive aus den Niederschlags- und Bodenfeuchtezeitreihen extrahiert und anschließend die jeweiligen Merkmalsausprägungen auf gegenseitige Korrelation beziehungsweise Unabhängigkeit untersucht. In einem weiteren Schritt werden den Extremwertstichproben beider Variablen anhand ihrer empirischen Wahrscheinlichkeitsverteilung und mittels statistischer Anpassungs- und Testverfahren, geeignete theoretische Verteilungsfunktionen (Randverteilungen) zugeordnet und angepasst. Aus diesen univariaten Randverteilungen kann nun, je nach Korrelationsstruktur der beiden Variablen, ein bivariates Verteilungsmodell entwickelt werden, um die Wahrscheinlichkeiten für das gemeinsame Auftreten bestimmter Merkmalsausprägungen (*joint* – oder *conditional probabilities*) bzw. das gleichzeitige Überschreiten zuvor definierter Schwellenwerte zu ermitteln.

Der Reiz dieses *joint-probability* Ansatzes liegt darin, dass vorhandene Daten und die bisher verwendeten Hochwassermodelle weiterhin in vollem Umfang eingesetzt werden können, da diese lediglich innerhalb eines hierfür angepassten Rahmens von Wahrscheinlichkeitsverteilungen angewendet werden. Damit könnte in der Zukunft möglicherweise eine Verbesserung in der Abschätzung extremer Hochwässer mit verhältnismäßig geringem Aufwand erreicht werden (RAHMAN et al., 1998).

1.3 Übersicht / Stand der Forschung

In einer der ersten zu diesem Thema veröffentlichten Arbeiten gelang es EAGLESON (1972) mithilfe des Ansatzes der kinematischen Welle und ohne Verwendung von Abflusszeitreihen, die Häufigkeitsverteilung der Durchflussraten von Hochwasserereignissen aus den Randverteilungen von Niederschlagsdauer und -intensität abzuschätzen, wobei diese als unabhängige und exponentiell verteilte Zufallsvariablen betrachtet wurden. Aus den daraus ermittelten Überschreitungswahrscheinlichkeiten konnten dann die theoretischen Wiederkehrintervalle dieser Hochwasserereignisse abgeleitet werden, wobei sich gute

Übereinstimmungen mit gemessenen Abflusszeitreihen von drei verschiedenen Einzugsgebiete in Connecticut, USA ergaben. BERAN (1973) setzte ein Abflussmodell auf Basis eines *catchment wetness index (CWI)* mit einem *unit hydrograph* als Routingkomponente ein, und testete die so erhaltenen Hochwasser-Häufigkeitsverteilungen und Wiederkehrintervalle hinsichtlich ihrer Sensitivität auf zufällige, empirischen Verteilungen folgenden Variationen des CWI sowie der Parameter Niederschlagshöhe und -dauer. Den größten Einfluss hatten hier Änderungen des CWI auf die resultierende Verteilung der Hochwasserereignisse. Es zeigte sich jedoch auch, dass die simulierte Hochwasserverteilung eine wesentlich flachere Form, d.h. eine größere Varianz aufwies als die beobachtete.

Seitdem wurden eine ganze Reihe von Studien zu hydrologischen Fragestellungen mit bi- und multivariaten Extremwertanalysen veröffentlicht, deren Großteil sich jedoch mit den *joint probabilities* von meist eng korrelierten, "artverwandten" Variablen wie Niederschlagsdauer und -intensität oder maximaler Abflussrate und Gesamtvolumen eines Hochwasserereignisses beschäftigte. RAYNAL-VILLASENOR & SALAS (1987), SINGH & SINGH (1991), BACCHI et al. (1994), YUE et al. (1999; 2001) und (YUE, 2001) untersuchten verschiedene multivariate Modelle mit Exponential- Gamma- oder Gumbel – Randverteilungen auf ihre Fähigkeit hin, Wahrscheinlichkeiten und Wiederkehrintervalle von hydrologischen Ereignissen mit mehreren korrelierten Zufallsvariablen als Prädiktoren beschreiben zu können. Die größten Probleme in der praktischen Anwendung dieser Ansätze ergeben sich erstens in der Erfassung und Parametrisierung der Korrelationsstruktur, bzw. des funktionalen Zusammenhangs (linear oder nicht-linear) der Variablen untereinander, und zweitens aus der Notwendigkeit, für beide Variablen gleiche Randverteilungstypen oder zumindest Verteilungsfunktionen der selben Familie (z.B. Gumbel- und Exponentialverteilung) einsetzen zu müssen (JOE, 1993; BALISTROCCHI & BACCHI, 2011).

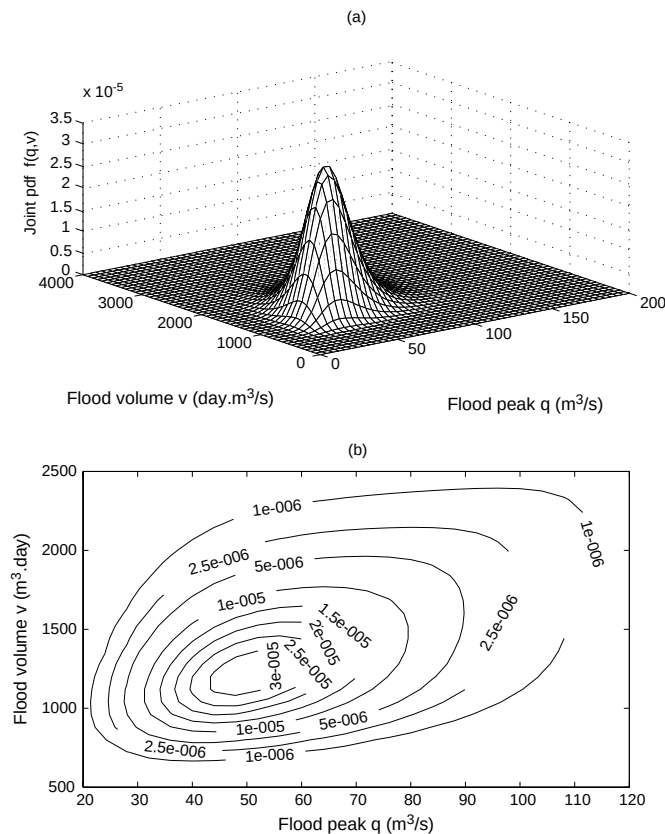


Abbildung 1.1 a: Bivariate Dichtefunktion $f(q,v)$ für maximale Abflussrate Q und Gesamtvolumen V eines Hochwasserereignisses, b: Konturplot von $f(q,v)$. (YUE & RASMUSSEN, 2002)

In Abbildung 1.1 ist eine solche bivariate Dichtefunktion $f(x,y)$ (*joint probability density function, JPDF*) mit Gumbelverteilungen für beide Rand- oder Marginalverteilungen $f(x)$, $f(y)$ zweier korrelierter Variablen X und Y (hier die maximale Abflussrate Q und Gesamtvolumen V eines Hochwasserereignisses) beispielhaft dargestellt. Abbildung 1.2 zeigt die dazugehörige kumulative Wahrscheinlichkeitsverteilung (*joint cumulative distribution function, JCDF*) für Ereignisse bei denen Q und V gleichzeitig einen bestimmten Schwellenwert der Merkmalsausprägung q bzw. v überschreiten. YUE & RASMUSSEN (2002) geben hier - nach Kenntnis des Autors - als einzige einen allgemeinen Überblick über die verschiedenen Konzepte und Interpretationsmöglichkeiten von multivariaten Wahrscheinlichkeitsverteilungen (*joint probabilities*) in Hinblick auf hydrologische Fragestellungen, insbesondere der Bedeutung von Wiederkehrintervallen in Zusammenhang mit multivariater Extremwertanalyse (vgl. Kapitel 3.5).

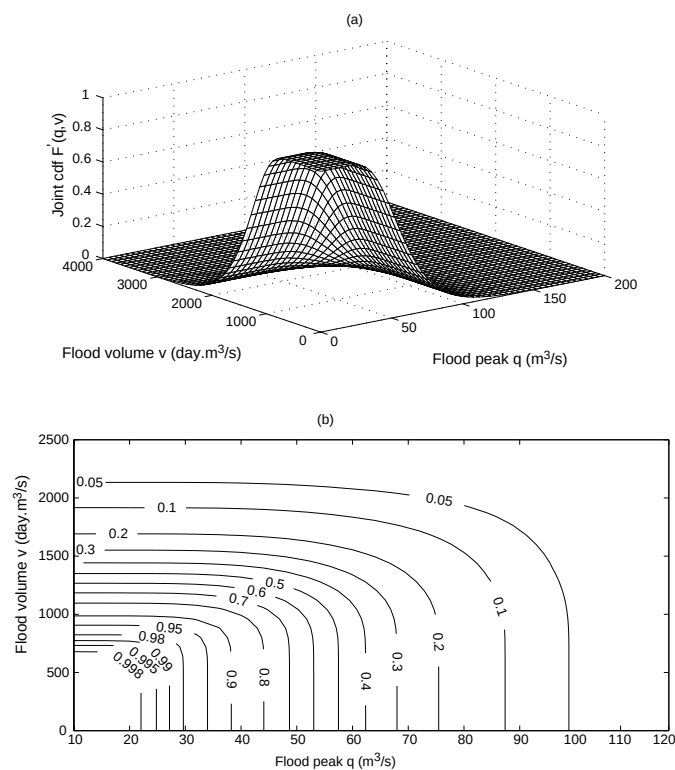


Abbildung 1.2 a: Bivariate, kumulative Wahrscheinlichkeitsverteilung $F'(Q \geq q, V \geq v)$ für maximale Abflussrate Q und Gesamtvolumen V eines Hochwasserereignisses, b: Konturplot von F' . (YUE & RASMUSSEN, 2002)

Eine weitere in der Literatur oft behandelte Methode zur Beschreibung von *joint probabilities* besteht in der Anwendung von sogenannten Copulas (u.a. ZHANG & SINGH, 2006; 2007; BACCHI & BALISTROCCHI, 2010; BALISTROCCHI & BACCHI, 2011). Aufgrund ihrer mathematischen Komplexität und der Tatsache, dass man bei den in dieser Studie untersuchten Variablen prinzipiell von gegenseitiger Unabhängigkeit ausgehen kann, sollen diese Verfahren hier nur kurz erwähnt werden. Eine Copula - Funktion C ist in der Lage, eine bi- oder auch multivariate Verteilungsfunktion $H(x, y, \dots)$ zweier oder mehrerer Zufallsvariablen $X, Y, \dots$ mit deren Randverteilungen $F(x), G(y), \dots$ (welche hier beliebigen Familien angehören können) in der Form $H(x, y, \dots) = C(F(x), G(y), \dots)$ zu verknüpfen und dabei auch deren Korrelationsstruktur weitgehend zu erfassen. Für eine tiefgehende Erläuterung und weiterführende Literatur zum Thema Copulas siehe SCHWEIZER & SKLAR (1983), NELSEN (1991) und DURANTE & SEMPI (2010).

Obwohl ein allgemeiner Konsens darin besteht, dass die Vorfeuchtebedingungen in einem Einzugsgebiet zu den entscheidenden Kontrollgliedern bei der Abflussbildung gehören und prinzipiell als Zufallsvariable behandelt werden sollten, wird dieser Tatsache meist nur in Form einiger verschiedener Bodenfeuchteklassen (*Antecedent Moisture Classes, AMC*), *Antecedent Precipitation Indices (API)*, u. A. in Kombination mit der SCS – *Curve Numbers* Methode (z. B. FRANCHINI et al., 1996; DE MICHELE & SALVADORI, 2002; MARTINA et al., 2005) oder beispielsweise durch Abschätzung des Speicherzustands aus dem Basisabfluss (LONGOBARDI et al., 2003), Rechnung getragen. In konkretem Bezug auf die oben genannte Zielsetzung existieren in der zugänglichen Literatur kaum Studien, in denen die verschiedenartigen Vorfeuchte- oder Speicherparameter der untersuchten Einzugsgebiete explizit als Zufallsvariablen behandelt werden. Eine allgemeine Analyse von Niederschlags- und Vorfeuchtevariablen im Sinne ihrer *joint probabilities* konnte in dieser Form während der Literaturrecherche nicht gefunden werden. In einer teilweise mit diesem Thema verwandten Publikation konnten TRAMBLAY et al. (2010) diesbezüglich nachweisen, dass Modelsimulationen des Bodewassergehaltes aus dem mit GWN-BW vergleichbaren, von *Météo-France* entwickelten Wasserhaushaltsmodell SIM durchaus als Prädiktoren geeignet sind (besser als die o. g. AMC oder API), um die Anfangsbedingungen ereignisbasierter Abflussmodelle festzulegen.

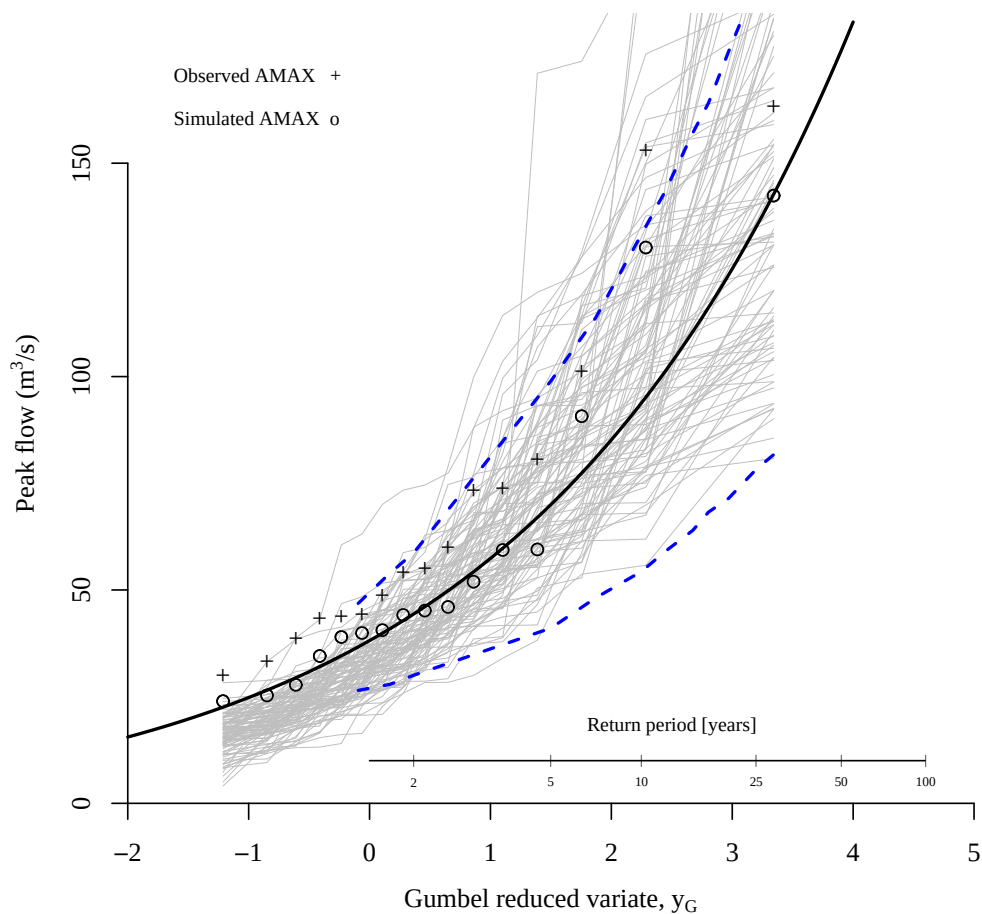


Abbildung 1.3: Vergleich von beobachteten (+), kontinuierlich simulierten (o) und durch Monte Carlo Simulation von Einzelereignissen erzeugte Häufigkeitsverteilungen der maximalen Abflussrate (graue Linien), sowie die angepasste GEV der kontinuierlichen Simulation (schwarze Linie) mit 95% Konfidenzintervallen (blau gestrichelte Linien). (KJELDEN et al., 2010)

Lediglich KJELDEN et al. (2010) setzten das von MOORE (2007) vorgestellte, ereignisbasierte *Probability Distribution Model* (PDM) ein, um in einer Monte Carlo Simulation eine große Anzahl an Abflussereignissen mit den *joint probabilities* verschiedener Eingangsvariablen für das Einzugsgebiet des Blyth River in Nordost-England zu erzeugen. Die Parameter für die Vorbedingungen (u. a. Niederschlagsdauer, Niederschlagsintensität und Bodenfeuchtedefizit) wurden dabei mithilfe der zuvor durch Anpassung an beobachtete Daten ermittelten Verteilungsfunktionen stochastisch modelliert. Durch eine separate stochastische Simulation der Zeitspanne zwischen den Abflussereignissen konnten die Einzelereignisse zu 100 "kontinuierlichen" Zeitreihen (oder *event strings*) mit je-

weils 16 Jahren Länge zusammengesetzt werden. Die grauen Linien in Abbildung 1.3 zeigen die empirischen Hochwasserverteilungen (jährliche Maxima der Abflussrate) aus der Monte Carlo Simulation, die Kreise die Spitzenabflüsse einer kontinuierlichen Kontrollsimulation, die schwarze durchgezogene Linie eine daran angepasste Extremwertverteilung (EV1 = Gumbelverteilung) mit 95% Konfidenzintervall (blaue gestrichelte Linie). Die Kreuze stehen für die beobachteten Maxima. Für die kontinuierliche Simulation lässt sich hier eine deutliche, systematische Unterschätzung der Abflussraten über die gesamte Spanne gegebener Jährlichkeiten feststellen. Auch unter Berücksichtigung der 95% Konfidenzintervalle kann die beobachtete Hochwasserverteilung offenbar nicht verlässlich wiedergegeben werden. Obwohl sich auch für die aus der Monte Carlo Simulation erzeugten Verteilungen eine tendenzielle Unterschätzung – besonders für die geringeren Jährlichkeiten – erkennen lässt, zeigt die Grafik sehr deutlich die extreme Varianz der möglichen Hochwasserverteilungen, die sich aus den verschiedenen Realisationen der zusätzlichen zufallsverteilten Eingangsvariablen ergibt.

In Anbetracht dieser Ergebnisse könnte die folgende Analyse also als Basis für eine zukünftige, vergleichbare Studie in Baden-Württemberg dienen.

2 Datengrundlage

Die für die Analyse vorliegenden Bodenfeuchte- und Niederschlagsdaten bestehen aus 39-jährigen Zeitreihen (1971 - 2009) aus den für die gesamte Fläche Baden-Württembergs durchgeführten Simulationen des Grundwasserneubildungsmodells GWN-BW. Dieses Modell basiert auf dem an der ETH Zürich entwickelten und von ARMBRUSTER (2002) erweiterten Verdunstungsmodell TRAIN bzw. TRAIN-GWN und ist ein deterministisches, flächendifferenziertes Modell zur Berechnung der aktuellen Evapotranspiration, zur Simulation des Bodenwasserhaushaltes sowie zur Bestimmung der unterhalb der durchwurzelten Bodenzone gebildeten Sickerwassermenge auf Tagesbasis. Dabei werden modellinterne Rechenschritte in Zusammenhang mit Strahlungsbilanz und Sickerwasserbildung in deutlich höherer zeitlicher Auflösung (z.B. Stundenbasis) abgearbeitet (MORHARD, 2005; 2009).

Die Datensätze der Simulationsergebnisse liegen in Form von zwei Matrizen, für Niederschlag und Bodenfeuchte mit jeweils 65 645 Zeilen für die räumlichen Bezugspunkte und 14 184 Spalten für die Tageswerte vom 01.01.1971 bis 30.10.2009 vor. Die Bezugspunkte entsprechen durchschnittlich etwa 0.54 km² großen Polygonflächen, welche durch Verschneidung von Leitbodenform und insgesamt 14 Landnutzungsklassen (siehe Abbildung 2.1) erzeugt wurden (WABOA, 2007).

2.1 Niederschlagsdaten

Die meteorologischen Eingangsdaten (i. e. Stationswerte des Niederschlags als Tagessummen in mm) können mit GWN-BW nach verschiedenen Verfahren auf die Fläche übertragen werden. In den vorliegenden Simulationsergebnissen wurde die räumliche Interpolation der Niederschlagszeitreihen durch ein Segmentverfahren realisiert, bei welchem jeweils die zum Zielpunkt nächstgelegene Station innerhalb von sechs Kreissegmenten mit je 60° Öffnungswinkel ausgewählt wurde. Die in den Segmenten gefundenen Stationswerte wurden einer

inversen Gewichtung nach dem Quadrat der Distanz zum Zielpunkt unterzogen. Für Parameter, bei denen von einer Höhenabhängigkeit ausgegangen werden kann (z.B. Niederschlag, Lufttemperatur) wird zusätzlich der Höhengradient zwischen den umliegenden Beobachtungspunkten analysiert und bei deutlicher Abhängigkeit zwischen Merkmalsausprägung und Geländehöhe diese zusätzlich zum Ergebnis der Distanzgewichtung berücksichtigt. Dabei wird die Gewichtung des aus der Höhenlage des Zielpunktes geschätzten Wertes um so stärker, je stärker die Korrelation zwischen Merkmalsausprägung und Geländehöhe ausfällt (LUBW, 2006; MORHARD, 2009). Die zur Regionalisierung verwendeten Niederschlagsstationen des DWD sind in Abbildung 2.2 zu sehen.

2.2 Bodenfeuchtedaten

Die Bodenfeuchte θ ist in Prozent bzw. als Faktor $[0,1]$ des volumetrischen Wassergehalts der entsprechenden Bodenart bei Feldkapazität angegeben (% FK). Die Feldkapazität FK gibt an, welche Wassermenge (in mm/m Bodentiefe) ein frei drainierender Boden über längere Zeit speichern kann. In der Regel entspricht dies dem Wassergehalt des Bodens etwa ein bis zwei Tage nach ausreichender Wasserzufuhr, bei fehlender Evapotranspiration kann dieser Wert dann über einen längeren Zeitraum näherungsweise erhalten bleiben.

Die hier verwendete Einheit hat den Vorteil, den Sättigungsgrad der jeweiligen Bezugsfläche direkt, ohne Berücksichtigung der Bodenart und seiner Speicherefähigkeit angeben zu können, während der tatsächliche Wassergehalt nur mit Kenntnis des drainierbaren Porenvolumens bzw. des volumetrischen Wassergehalts dieser Bodenart bei Feldkapazität bestimmt werden kann. Ein anderer Vorteil dieser Einheit hinsichtlich ereignisbasierter Niederschlag-Abfluss-Modellierung besteht darin, dass bei einem Wert von 100% FK das Einsetzen von für diese Fläche typischen Abflussbildungsmechanismen (Tiefenperkolation, Zwischenabfluss, Oberflächenabfluss etc.) angenommen werden kann (DYCK & PESCHKE, 1995). Abbildung 2.3 zeigt die räumliche Verteilung der Feldkapazität in qualitativen Klassen.

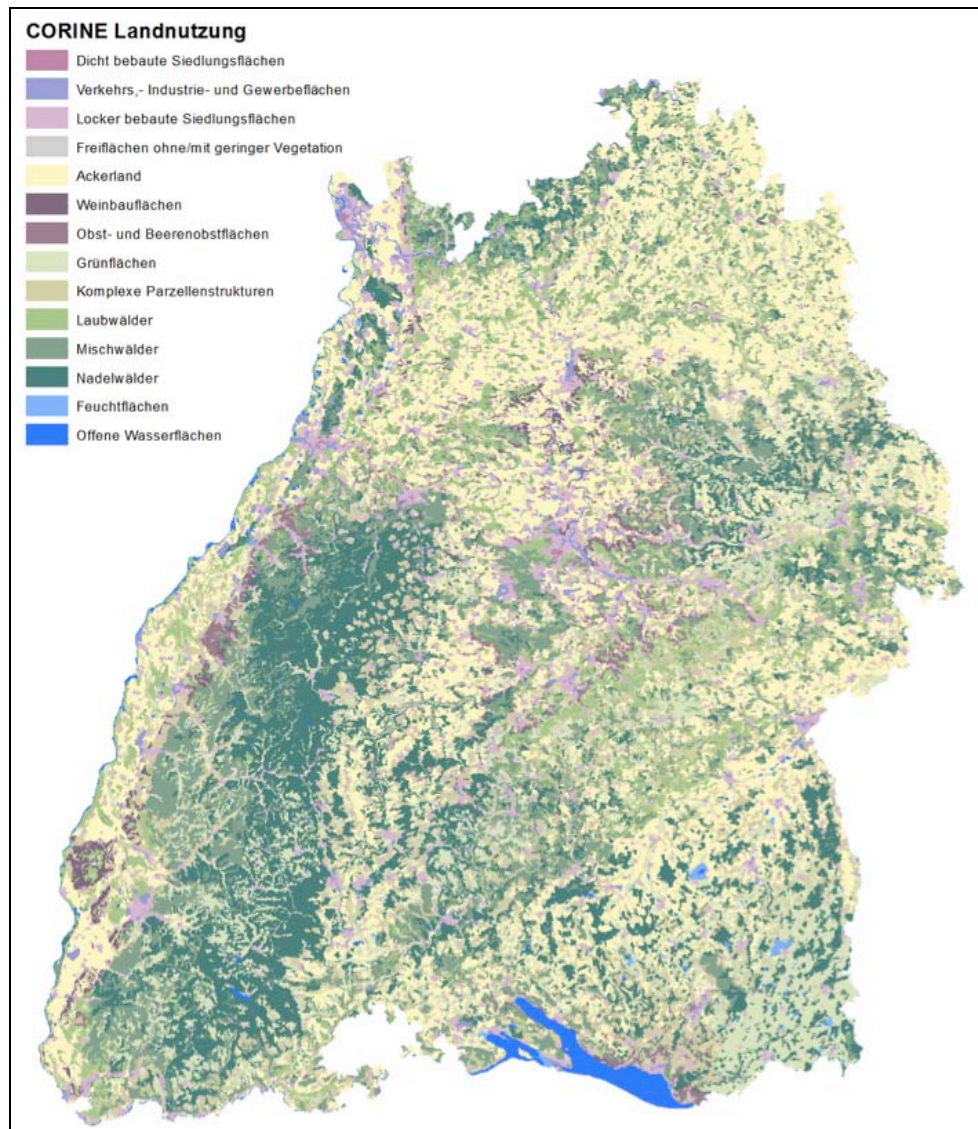


Abbildung 2.1: CORINE - Landnutzungsklassen (Waboa, 2007).

2.3 Zusätzliche Daten

Die weiteren in dieser Arbeit verwendeten Daten und Karten stammen aus dem Räumlichen Informations- und Planungssystem (RIPS) der Landesanstalt für Umweltschutz Baden-Württemberg. Die Projektgruppe "Wasser- und Bodenatlas Baden-Württemberg" (WABOA, 2007) erarbeitet flächendifferenzierte Landesübersichtskarten aus verschiedenen hydrologischen, meteorologischen und geografischen Themenbereichen auf einer einheitlichen Grundlage. Zur Ermittlung möglicher gebietsspezifischer Einflussfaktoren auf die untersuchten Nie-

derschlags - Vorfeuchte Korrelationen und Verteilungen wurden verschiedene thematische Karten aus diesem Werk ausgewählt.

Der Übersichtlichkeit halber wurde versucht, eine Reihe möglichst repräsentativer und aussagekräftiger Modellflächen (in Abb. 2.3 mit Buchstaben A - F markiert) auszuwählen, die im Folgenden zur detaillierten grafischen Darstellung von beispielsweise zeitlichen Verläufen oder Häufigkeitsverteilungen, welche nicht in Kartenform darstellbar sind, dienen sollen. Tabelle 2.1 gibt eine Zusammenfassung über einige wichtige Kennwerte (Feldkapazität des Bodens, Landnutzung, potentielle Verdunstung und Höhe über Meeresniveau) der ausgewählten Modellflächen. Die Saisonalität des Niederschlags gibt an, in welchem Quartal durchschnittlich der größte Anteil der jährlichen Niederschlagssumme gefallen ist. Die Spalten "GIS_ID" und "R_ID" dienen lediglich zur Referenzierung der jeweiligen Fläche in den beigefügten GIS- und R-Dateien.

Tabelle 2.1: Die zur exemplarischen Darstellung verschiedener Grafiken ausgewählten Modellflächen.

ID	GIS_ID	R_ID	Feldkapazität		Landnutzung	N - Saison.	H. ü. NN [m]	V _{Pot} [mm]
			[mm]	Klasse				
A	57168	7927	260-520	mittel-hoch	Laubwälder	Sommer	ca. 450	525-550
B	32918	11673	260-390	mittel	Ackerland	Sommer	ca. 150	625-650
C	4803	31516	N/A	stark wechselnd	Ackerland	ausgegl.	ca. 130	625-650
D	36349	42178	30-390	sehr gering-mittel	Mischwälder	Sommer	ca. 750	450-475
E	3081	58304	30-260	sehr gering - gering	Grünland	Winter	ca. 1020	450-475
F	50262	65519	260-390	mittel	Nadelwald	Sommer	ca. 500	550-575

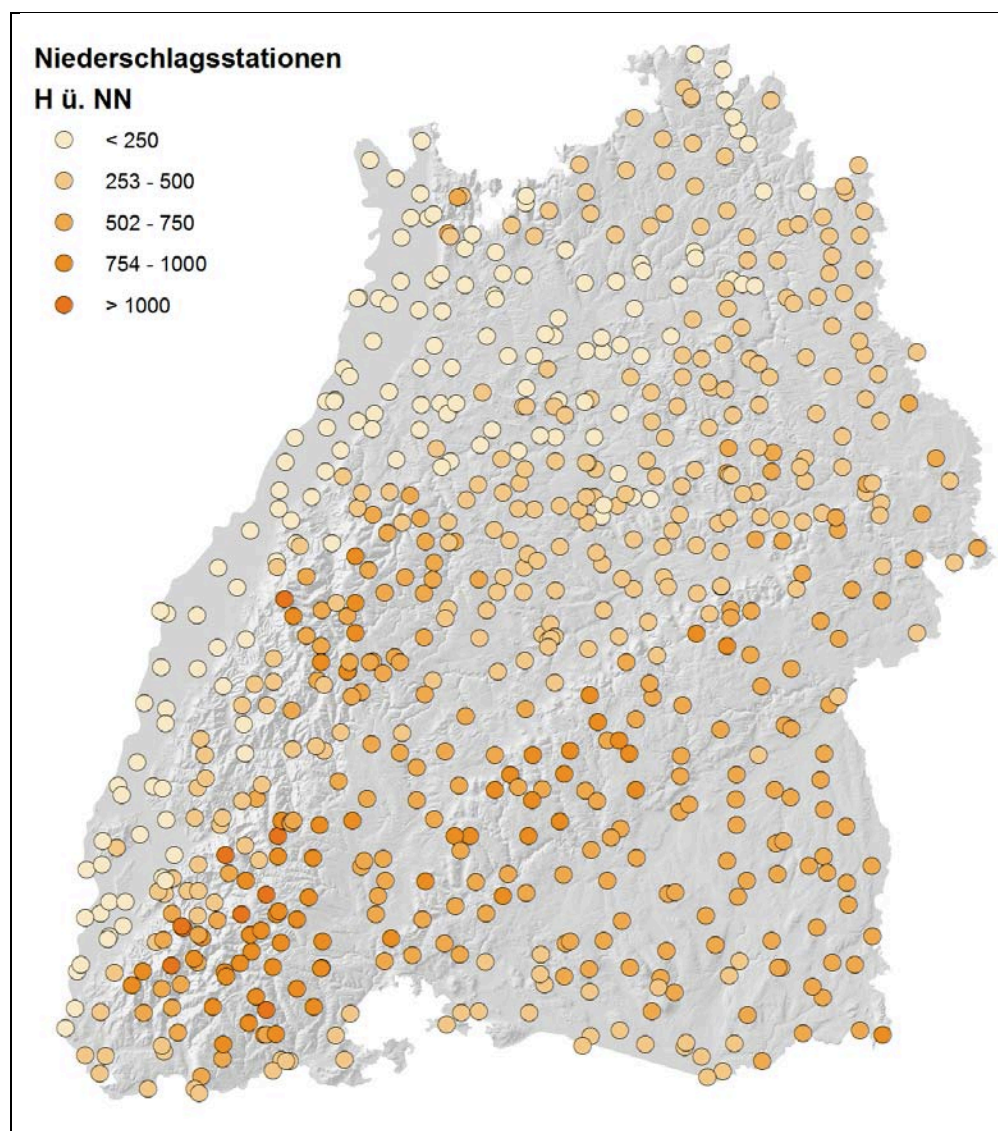


Abbildung 2.2: Zur Regionalisierung der Niederschlagsdaten benutzte Stationen des DWD. (WABOA, 2007)

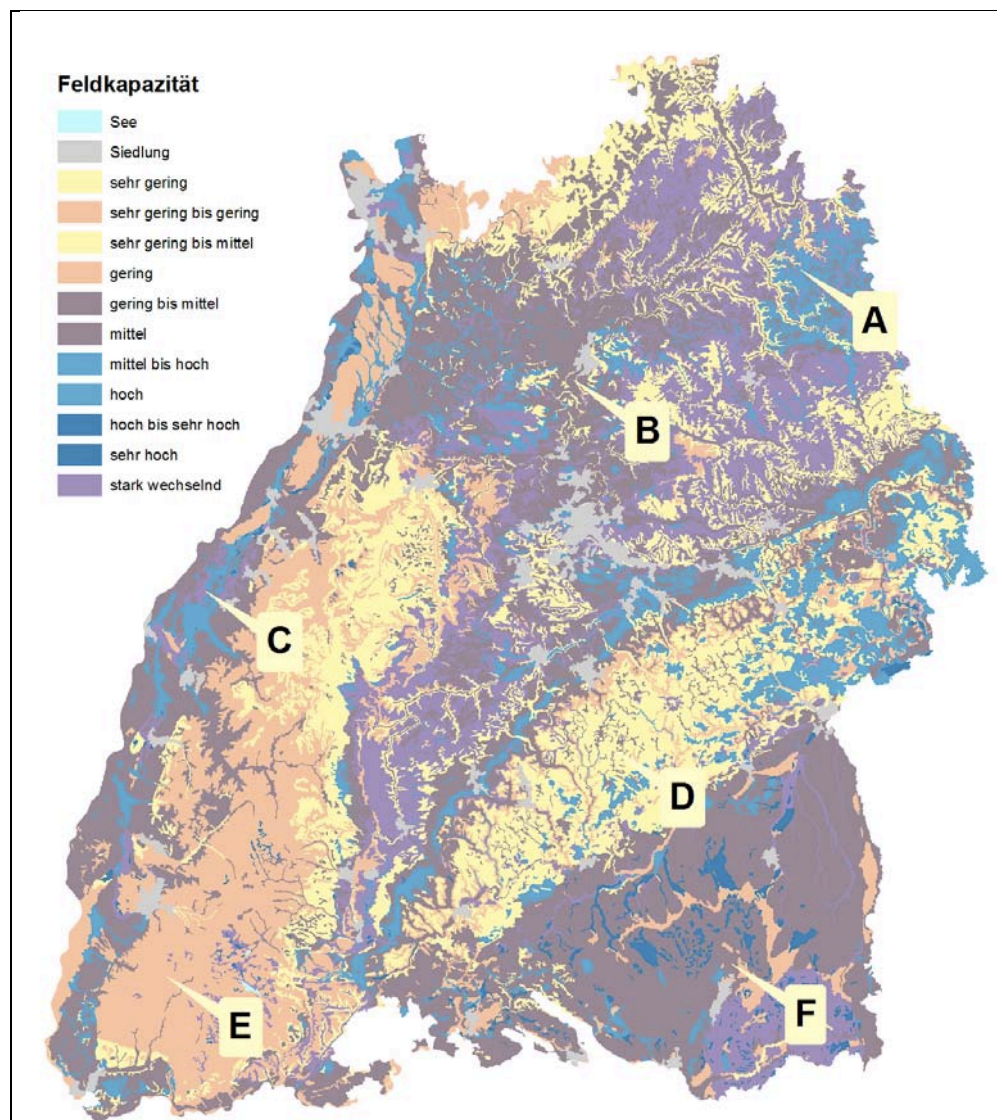


Abbildung 2.3: Feldkapazitätsklassen der Böden in Baden-Württemberg und Lage der ausgewählten Teilflächen A – F aus Tabelle 2.1. (WABOA, 2007)

3 Methodik & Vorgehensweise

Die statistische Analyse wurde nahezu ausschließlich mithilfe der freien, unter der *GNU General Public License* veröffentlichten Statistiksoftware *R* und den optionalen, auf den Servern des *Comprehensive R Archive Network* (CRAN) verfügbaren Erweiterungspaketen durchgeführt (siehe CRAWLEY, 2007; R, 2010). Für die räumliche Darstellung der (Teil-) Ergebnisse und die Erstellung des Kartenmaterials wurde das von der Firma ESRI (*Environmental Systems Resource Institute*) entwickelte geografische Informationssystem ArcGIS Desktop 10 (ESRI, 2010) verwendet.

3.1 Datenvorbereitung

Da jeder der beiden Datensätze über 930 Mio. Elemente ($14\,184 \times 65\,645$) enthält und deshalb in *R* nur mit enormem Arbeitsspeicherbedarf am Stück verarbeitet werden könnte, besteht der erste Schritt der Datenbearbeitung darin, die beiden großen Matrizen mit N (mm) und θ (%FK) in 65 645 Einzeldateien mit den jeweils korrespondierenden N und θ - Zeitreihen für jeden Punkt aufzuteilen und diese anschließend nacheinander (als *Batch*) zu verarbeiten. In diesem Schritt wurde gleichzeitig auch die zeitliche Verzögerung (*time-lag*) vorgenommen, indem die N -Zeitreihe gegenüber der θ - Zeitreihe um einen Tag nach vorne verschoben wurde. Dadurch wird in den folgenden Analyseschritten automatisch der Niederschlagswert des Tages t mit der Bodenfeuchte des Tages $t-1$, also der Vorfeuchte θ_v verglichen. Dabei fallen jeweils die ersten und letzten Wertepaare aus der Zeitreihe, da für diese kein korrespondierender N bzw. θ - Wert vorhanden ist. Dieser Schritt ist nötig, da der Niederschlag des Tages t im Modell bereits in die Berechnung der Bodenfeuchte des selben Tages t eingegangen ist und eine Analyse dieser Wertepaare letztlich nur die modellinterne Berechnung von θ aus dem N - Input reproduzieren würde. Fehlende oder Null-Werte werden jeweils während der jeweiligen Rechenoperationen identifiziert und entfernt bzw. entsprechend gehandhabt.

3.2 Erstellung der Stichprobenkollektive

Die im folgenden Kapitel erlangten Erkenntnisse bestimmen in hohem Maße die weitere Entwicklung der Methodik. Deshalb wurde zu Gunsten von Verständlichkeit und Konsistenz entschieden, einige wichtige Teilergebnisse bereits in diesem Kapitel vorweg zu nehmen und in einer kurzen Synthese am Ende des Kapitels zusammen zu führen, um daraus die weitere Vorgehensweise zu entwickeln.

Die extremwertstatistische Analyse der Eingangsdaten erfordert zunächst die Bildung geeigneter Stichprobenkollektive aus den kontinuierlichen Zeitreihen jeder einzelnen Polygonfläche. Dazu werden jeweils die Extremwerte des Niederschlags mit der korrespondierenden Bodenfeuchte des Vortags ausgewählt (siehe Kap. 3.1). Um die für die Frequenzanalyse grundsätzlich vorausgesetzten Bedingungen eines Poisson - Prozesses (siehe u.a. HAAN, 1982; MANIAK, 1997) zu erfüllen, müssen für beide Variablen jeweils zeitlich unabhängige Wertepaare gefunden werden. Da die zeitliche Auflösung der Daten mit 24h - Intervallen nicht ausreicht, um einzelne Niederschlagsereignisse voneinander abzugrenzen, werden dazu hier die Tagessummen eingesetzt (im Folgenden oft auch als "Ereignisse" bezeichnet).

Die zeitliche Entwicklung der Bodenfeuchte folgt der des Niederschlags prinzipiell mit kurzer Verzögerung (im Allgemeinen $< 24h$) (siehe Abbildung 3.1). Es kann also davon ausgegangen werden, dass für unabhängige Niederschlagsereignisse auch die Vorfeuchtwerte θ_v untereinander unabhängig sind.

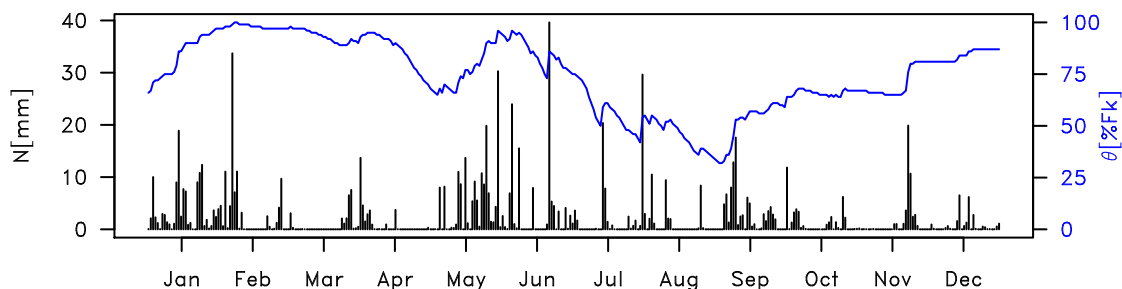


Abbildung 3.1: Niederschlags- und Bodenfeuchtezeitreihen, exemplarisch für das Jahr 1984, Teilfläche B.

3.2.1 Jährliche und partielle Serie

Die zur Auswahl der Extremwerte gängigen Ansätze bestehen zum einen in der Selektion der jährlichen Maxima (jährliche Serie oder *annual maximum peak*, AMP), zum anderen besteht die Möglichkeit, Niederschlagshöhen oberhalb eines geeigneten Schwellenwertes (partielle Serie oder *peak over threshold*, POT) zu erfassen (LANGE, 2008; LEONARD et al., 2008). Letzteres ist jedoch nur bei verhältnismäßig kurzen Zeitreihen (< 30 Jahre) üblich, da durch die Wahl des Schwellenwertes der Stichprobenumfang vergrößert und somit die Vertrauenswürdigkeit der abgeschätzten Verteilungsfunktion (im statistischen Sinne) verbessert werden kann.

Der Auswahl des Schwellenwertes ist besondere Aufmerksamkeit zu widmen. Dabei können sowohl rein mathematisch-statistische (z.B. die Verwendung bestimmter Quantile), als auch prozessorientierte Überlegungen (z.B. aus Abflussmodellen ermittelte mittlere effektive Niederschlagshöhen) zur Anwendung kommen. Generell ist hier zu beachten, den Schwellenwert hoch genug zu wählen um die Unabhängigkeit der Ereignisse (und damit die Gültigkeit eines Poisson – Modells) sicherzustellen, gleichzeitig jedoch niedrig genug, um den Stichprobenumfang ausreichend groß zu halten (HAAN, 1982; LANG et al., 1999). Sind die zu untersuchenden Zeitreihen ausreichend lang, wird im Allgemeinen die jährliche Serie (AMP) bevorzugt, da hier die Werte – mit Ausnahme der über den Jahreswechsel des hydrologischen Jahres vom 31.10 auf den 1.11. auftretenden Ereignisse – voneinander Unabhängig sind und auch die zwischenjährliche Varianz der Niederschlagshöhen in der Verteilung berücksichtigt werden (HAAN, 1982; SCHÖNWIESE, 2006).

Da die vorliegenden Datenreihen mit 39 Jahren ausreichend lang sind, werden für die Ermittlung der Niederschlagsextremwertverteilungen die jährlichen Serien eingesetzt, jedoch mit einer zusätzlichen Berücksichtigung saisonaler Effekte.

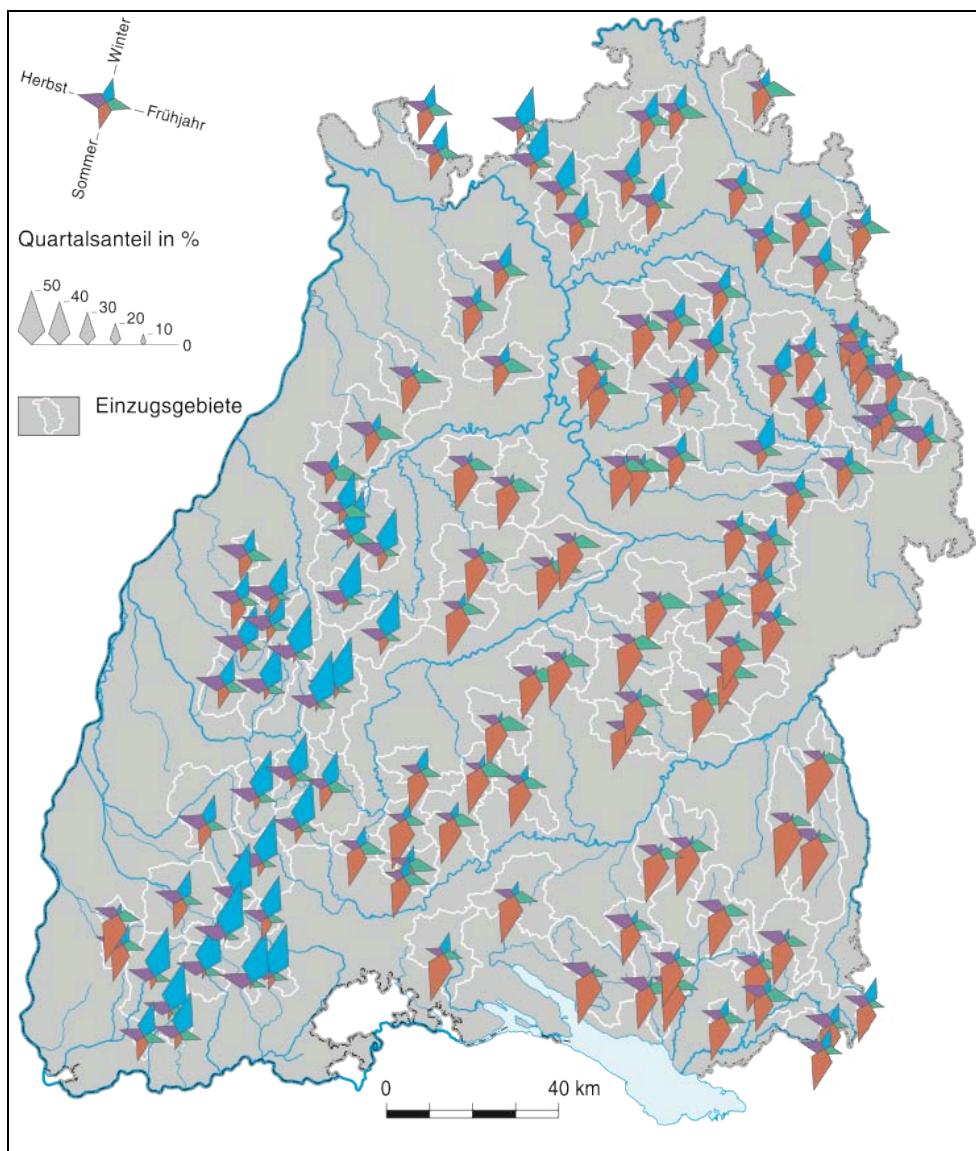


Abbildung 3.2: Saisonalität der Starkniederschläge in Quartalsdarstellung. (WABOA, 2007)

3.2.2 Saisonale Einteilung

Die in Abbildung 3.2 gezeigte Verteilung der Starkniederschläge weist eine von Norden nach Süden immer deutlichere Saisonalität über den gesamten Untersuchungsraum auf: Während die Starkniederschläge in den höher gelegenen Regionen des Schwarzwaldes oder dem Odenwald aufgrund ihrer vorwiegend frontalen Entstehung gehäuft im Winter auftreten, zeigen die Niederschläge in den tieferen Lagen eine stärkere Konzentration der Extremwerte in den Som-

mermonaten, was auf einen überwiegend konvektiven Charakter der Ereignisse schließen lässt (WABOA, 2007).

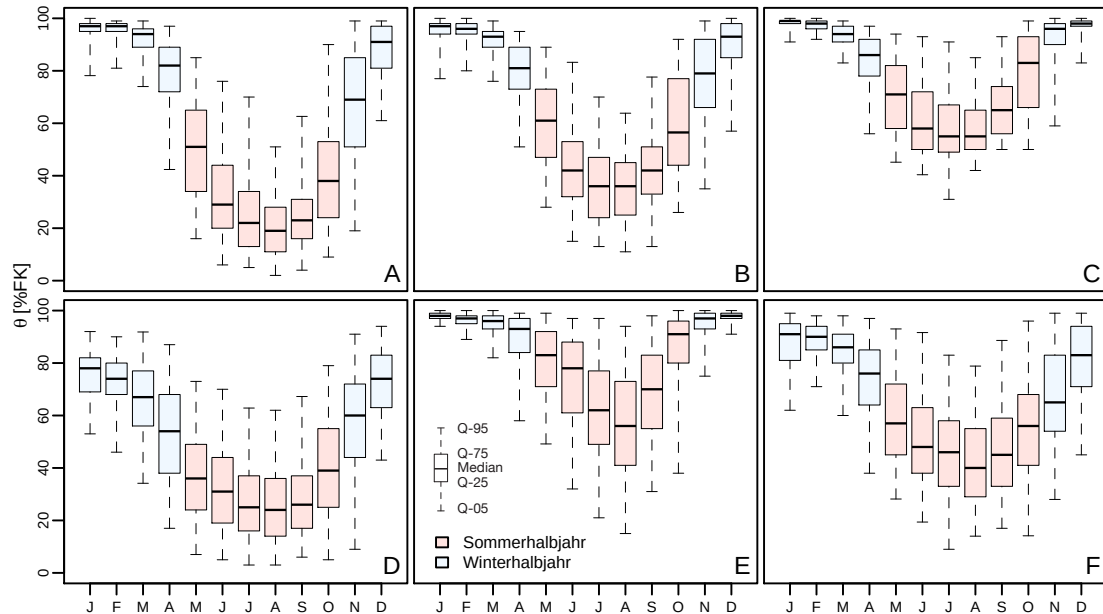


Abbildung 3.3: Box-Plot des Jahresganges der Bodenfeuchte an den 6 ausgewählten Stationen A - F (siehe Tabelle 2.1) über den Zeitraum von 39 Jahren.

Da der Bodenwassergehalt innerhalb der durchwurzelter Zone stark durch die aktuelle Evapotranspiration und diese wiederum stark von jahreszeitlich geprägten Größen wie z.B. der bodennahen Lufttemperatur und des Vegetationszustands beeinflusst wird, ist folglich auch für die korrespondierenden Bodenfeuchtezeitreihen von einer ausgeprägten Saisonalität auszugehen (RAVELO & DECKER, 1979; WILSON et al., 2004). Abbildung 3.3 zeigt Boxplots der monatlichen Verteilung der Bodenfeuchte über den gesamten Modellierungszeitraum. Alle ausgewählten Stationen zeigen in Hinblick auf ihren Median einen deutlichen Jahresgang, sowie weniger homogene, aber dennoch sichtbare Jahresgänge von Varianz und Schiefe der Verteilungen. Noch deutlicher wird dies, wenn man die Häufigkeitsverteilungen der Tageswerte der Bodenfeuchte über die gesamte Zeitreihe betrachtet. Diese zeigen über den gesamten Untersuchungsraum überwiegend bimodale Formen (siehe Abbildung 3.4 a), welche sich scheinbar aus zwei überlagerten, linksschiefen/rechtssteilen bzw. rechtsschiefen/linkssteilen Häufigkeitsverteilungen zusammensetzen. Wie die Histo-

gramme b und c in Abbildung 3.4 zeigen, ist eine gute Differenzierung der zugrundeliegenden Verteilungen durch eine einfache Einteilung der Zeitreihen in das hydrologische Sommer- (= 1. Mai - 31. Okt) und Winterhalbjahr (= 1. Nov - 31. Apr) möglich (DYCK & PESCHKE, 1995).

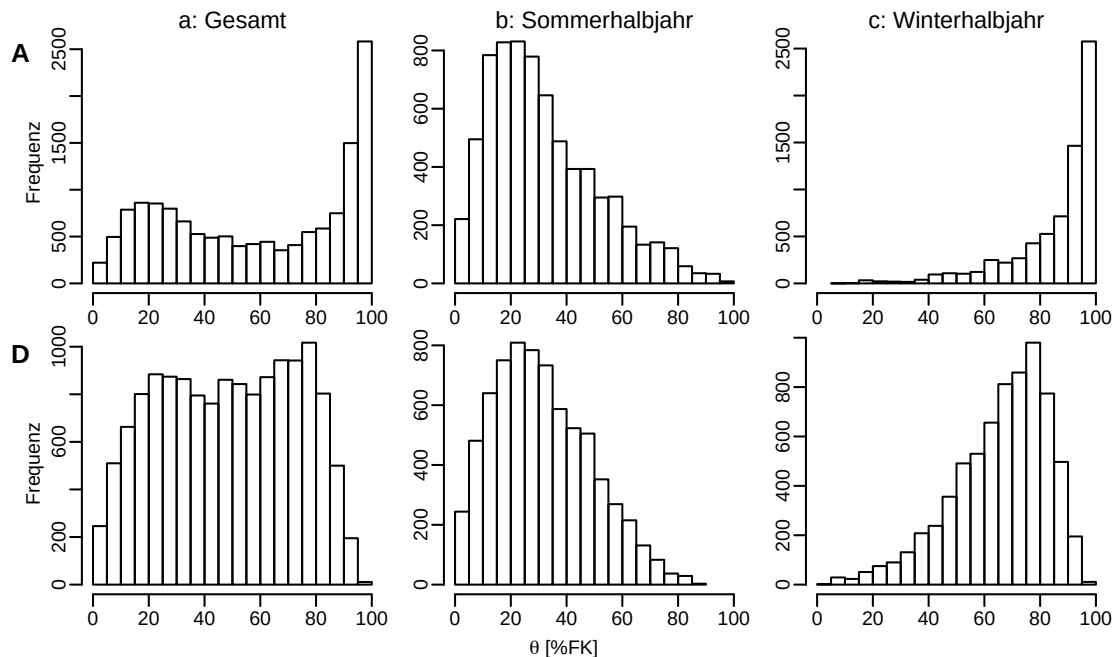


Abbildung 3.4: Häufigkeitsverteilungen der Bodenfeuchte-Tageswerte über die gesamte Zeitreihe (a) und jeweils getrennt für das hydrologische Sommer- (b) und Winterhalbjahr (c) an den Teilflächen A und D (siehe Tabelle 2.1).

3.2.3 Zwischenfazit

Aufgrund dieser Erkenntnisse wurde entschieden, wie von LANG et al. (1999) in solchen Fällen empfohlen, anstelle der jährlichen Niederschlagsmaxima jeweils die saisonalen Maxima für die Extremwertstichproben des Niederschlags zu verwenden und die Analyse getrennt vorzunehmen. Auf diese Weise lassen sich den erhaltenen empirischen Häufigkeitsverteilungen die passenden theoretischen Verteilungsfunktionen wesentlich einfacher und eindeutiger zuordnen. Ein weiterer Vorteil dieser Einteilung ist die garantierte Unabhängigkeit der Ereignisse, da grundsätzlich mindestens 6 Monate zwischen den Ereignissen liegen und auch die bei der Auswahl der jährlichen Maxima mögliche Erfassung von abhängigen Ereignissen über den Jahreswechsel ausgeschlossen werden kann. Bei dieser Vorgehensweise muss jedoch beachtet werden, dass sich die

daraus gewonnenen Überschreitungswahrscheinlichkeiten bzw. Wiederkehrintervalle der Extremereignisse auch jeweils nur auf die betreffende Saison beziehen. Die Jährlichkeit für ein Niederschlagsereignis einer bestimmten Höhe zu einem beliebigen Zeitpunkt des Jahres lässt sich hierauf folgend nach einfachen stochastischen Prinzipien aus den jeweiligen Einzelwahrscheinlichkeiten dieses Ereignisses für Sommer- und Winterhalbjahr berechnen. Vorausgesetzt die Sommer- und Wintermaxima X_S und X_W sind voneinander unabhängig, so gilt beispielsweise für die Wahrscheinlichkeit $P'_{SW}(X_S > x \vee X_W > x)$, dass also entweder X_S oder X_W oder beide in einem Jahr einen Schwellenwert x überschreiten:

$$P'_{SW} = P'(X_S) + P'(X_W) - P(X_S)P(X_W) \quad (3.2.1)$$

mit

$$P'(X) = 1 - P(X) \quad (3.2.2)$$

wobei $P(X)$ für die Unterschreitungswahrscheinlichkeit $P(X \leq x)$ und $P'(X)$ für die Überschreitungswahrscheinlichkeit $P(X > x)$ steht. Die Wahrscheinlichkeit $P'_{SW}(X_S > x \vee X_W > x \neg X_S > x \wedge X_W > x)$, dass nur einer der beiden x überschreitet, würde sich dann aus

$$P'_{SW} = P(X_S)P'(X_W) + P(X_W)P'(X_S) \quad (3.2.3)$$

ergeben (SCHÖNWIESE, 2006).

3.3 Korrelationsanalyse

Wie in Kapitel 1.3 bereits angesprochen wurde, besteht ein grundlegender Schritt bei der Ermittlung bivariater Wahrscheinlichkeitsverteilungen darin, die untersuchten Variablen auf Unabhängigkeit zu überprüfen, bzw. eventuelle Abhängigkeiten zu identifizieren und zu Beschreiben. Zu diesem Zweck wur-

den zunächst mögliche Korrelationen zwischen den extrahierten Niederschlagsmaxima und der korrespondierenden Vorfeuchte untersucht.

Prinzipiell kann dabei unter Betrachtung der beteiligten Prozesse von einer weitgehenden Unabhängigkeit der Variablen ausgegangen werden. Insofern lassen sich hier möglicherweise gefundene Korrelationen mit großer Wahrscheinlichkeit durch starke Autokorrelationseffekte der Niederschlagszeitreihe erklären. So ist zwar die Bodenfeuchte am Tag t kausal unabhängig von der Niederschlagshöhe N am Tag $t + 1$ und andersherum, jedoch kann N am Tag $t + 1$ durchaus eine starke Korrelation zu N am Tag t aufweisen, welche wiederum starken Einfluss auf die Bodenfeuchte dieses Tages hat. Deshalb wurde zur Kontrolle auch die Stärke der Autokorrelation des Niederschlags untersucht.

3.3.1 Kendall's Tau

Um die Stärke eines (linearen) Zusammenhangs zweier normalverteilter Variablen X und Y zu untersuchen, wird im allgemeinen der Produkt-Moment-Korrelationskoeffizient oder auch *Pearson's r* verwendet. Dieses parametrische Verfahren ist jedoch darauf beschränkt, ausschließlich lineare Zusammenhänge zu identifizieren und kann auch nur im Falle normalverteilter Variablen zuverlässige Ergebnisse liefern. Da für die Niederschlagsmaxima generell nicht mit einer Normalverteilung gerechnet werden kann, und auch die korrespondierenden Bodenfeuchtwerte bereits aufgrund ihrer beidseitigen Begrenzung im Intervall $[0,1]$ bzw. 0-100% keiner Normalverteilung angehören können (vgl. Abbildung 3.4), wurde für diese Untersuchung der Rangkorrelationskoeffizient nach KENDALL (1949) verwendet. Da er auf den Rängen der Daten anstelle der eigentlichen Werte basiert, ist dieser in der Lage, jede Form linearer oder nicht-linearer, monotoner Abhängigkeiten zu erfassen (REIMANN, 2008).

Zur Berechnung des Korrelationskoeffizienten τ nach Kendall werden zunächst alle Werte von X und Y nach Rängen geordnet, d.h. der kleinste Wert erhält den Rang 1, der zweitkleinste den Rang 2 usw. Im Falle einer positiven (negativen) monotonen Abhängigkeit zwischen den Wertepaaren werden mit wachsenden

Rängen von X auch die Ränge von Y häufiger zunehmen (abnehmen). Liegt Unabhängigkeit vor, werden die Y - Ränge in etwa gleich oft zu- und abnehmen, während X monoton ansteigt. Anschließend wird die Teststatistik S berechnet, indem von allen Kombination eines X - Y Wertepaares mit den folgenden Wertepaaren die Vorzeichenfunktion $\text{sgn}(m)$ der Steigung m zwischen diesen Punkten gebildet und aufsummiert wird:

$$S = P + M \quad (3.3.1)$$

wobei P die Summe positiver Steigungen $\text{sgn}(m) = 1$ und M die Summe der negativen Steigungen $\text{sgn}(m) = -1$ ist, mit

$$\text{sgn}(m) = \begin{cases} 1, & \text{wenn } m > 0 \\ 0, & \text{wenn } m = 0 \\ -1, & \text{wenn } m < 0 \end{cases} \quad (3.3.2)$$

Eine Datenreihe der Länge n ergibt also genau $n(n - 1)/2$ Wertepaare, deren Vorzeichenfunktion in S aufsummiert wird. Für eine perfekte positive monotone Korrelation würde sich also $S = n(n - 1)/2$ ergeben, für eine perfekte negative monotone Korrelation $S = -n(n - 1)/2$.

Teilt man nun S durch $n(n - 1)/2$, erhält man Kendall's τ , welches die Stärke der Korrelation als Wert zwischen -1 und 1 angibt:

$$\tau = \frac{S}{n(n - 1)/2} \quad (3.3.3)$$

τ stellt somit eine Wahrscheinlichkeit dar, d. h., es ist die Differenz zwischen der Wahrscheinlichkeit, dass die beobachteten Daten für beide Variablen in derselben Rangfolge vorliegen und der Wahrscheinlichkeit, dass die Daten für die beiden Variablen verschiedene Rangfolgen besitzen.

Die Nullhypothese H_0 des Tests besagt, dass die Variablen $X_i, Y_i, i = 1, 2, 3 \dots n$ unabhängig und gleichverteilt sind, also kein Trend vorliegt ($\tau = 0$). Die Alternativhypothese H_1 besagt, dass ein Trend ($\tau \neq 0$) vorliegt. Um die statistische Signifikanz von τ zu testen, wird S mit Werten verglichen, welche unter der Annahme einer richtigen Nullhypothese zu erwarten sind. Für kleine Stichproben ($n \leq 10$), können die kritischen Werte für S aus Tabellen entnommen werden (siehe HELSEL & HIRSCH, 2005). Für größere Stichproben ($n > 10$) kann die Verteilung von S gemäß der Nullhypothese gut durch eine Normalverteilung mit dem Erwartungswert $\mu_S = 0$ angenähert werden. Die Standardabweichung σ_S dieser Normalverteilung ist dann mit

$$\sigma_S = \sqrt{\frac{n(n-1)(2n+5) - \sum_{i=1}^n t_i(t_i-1)(2t_i+5)i}{18}} \quad (3.3.4)$$

gegeben, wobei die Summe in σ_S eine Korrektur für das Auftreten von gleichen Werten (*ties*) in der Datenreihe darstellt, mit t_i für die Anzahl der identischen Einträge pro Wert vom jeweiligen Umfang i . Einmalige Werte werden dabei als t_1 behandelt, sprich *ties* mit Umfang 1.

Mit σ_S lässt sich nun aus S die standardisierte Teststatistik Z ermitteln:

$$Z = \begin{cases} \frac{S-1}{\sigma_S} & , \text{wenn } S > 0 \\ 0 & , \text{wenn } S = 0 \\ \frac{S+1}{\sigma_S} & , \text{wenn } S < 0 \end{cases} \quad (3.3.5)$$

Bei einem zweiseitigen Test kann dann die Nullhypothese mit der Signifikanz α abgelehnt werden, wenn $|Z_S| > Z_{\alpha/2}$, wobei $Z_{\alpha/2}$ der Wert der Standardnormalverteilung mit einer Überschreitungswahrscheinlichkeit von $\alpha/2$ ist. Hier sei angemerkt, dass die Nullhypothese niemals Bewiesen werden kann. Es kann nur von ihrer Richtigkeit ausgegangen werden (d. h. sie wird nicht abgelehnt),

solange es keinen hinreichenden Grund gibt, vom Gegenteil auszugehen (HAAN, 1982; HELSEL & HIRSCH, 2005; REIMANN, 2008).

Die Korrelationsanalyse in dieser Arbeit wurde mit der Funktion `cor()` bzw. `cor.test()`, welche standardmäßig in *R* im Paket *stats* implementiert ist und nach den oben genannten Regeln, inklusive der Korrektur für *ties* arbeitet, durchgeführt (siehe CRAWLEY, 2007; R, 2010). Hier wird jedoch anstelle der Entscheidung, die Nullhypothese bei einem Signifikanzniveau α anzunehmen oder abzulehnen, ein sogenannter p-Wert (*p-value*) ausgegeben. Dieser entspricht bei einem zweiseitigen Test der doppelten Wahrscheinlichkeit, ein gleich großes oder größeres Z als das errechnete (für $Z > 0$) oder ein gleich kleines oder kleineres Z (für $Z < 0$) zu erhalten.

Der p-Wert kann also mit dem Signifikanzniveau selbst bzw. der tatsächlichen Irrtumswahrscheinlichkeit verglichen werden, eine vermeintlich unzutreffende Nullhypothese zurückzuweisen, obwohl sie richtig ist (Fehler 1. Art). In diesem Falle würde dies also bedeuten, den Variablen eine Korrelation zu unterstellen, die in Wirklichkeit nicht existiert. Je kleiner also der p-Wert ist, desto unwahrscheinlicher ist es, einen solchen Fehler zu begehen. Je nach Fragestellung werden hier im Allgemeinen kritische Werte für α bzw. den p-Wert von ≤ 0.1 ("signifikant"), ≤ 0.05 ("stark signifikant") und ≤ 0.01 ("äußerst signifikant") eingesetzt (u. a. SCHÖNWIESE, 2006; REIMANN, 2008).

3.3.2 Autokorrelation des Niederschlags

Der Autokorrelationskoeffizient gibt Auskunft darüber, inwieweit die Werte einer Zeitreihe von den vorhergehenden Werten dieser Reihe beeinflusst werden. Dazu wird die Datenreihe jeweils mit wachsendem zeitlichem Versatz (k) mit sich selbst korreliert, wobei die Zeitreihen gleichzeitig um den Wert von k verkürzt werden müssen. Für die Abhängigkeit der Werte zum Vorhergehenden Wert beträgt k also 1, man spricht dann von Autokorrelation 1. Grades. Für $k = 0$ ergibt sich gezwungenermaßen ein Koeffizient von 1. Alternativ lässt sich auch eine mittlere Dauer (als vielfaches der Einheit von t , z. B. Tage) angeben, bis der

Korrelationskoeffizient einen festgelegten Schwellenwert (meist Null) erreicht bzw. über/unterschritten hat oder um einen bestimmten Faktor gedämpft wurde. Dieses Dauer wird als Persistenzzahl oder *memory* bezeichnet, da die Zeitreihendaten innerhalb dieses Intervalls näherungsweise von den vorhergehenden Abhängig sind und somit eine gewisse "Erhaltungsneigung" besteht. Für den Autokorrelationskoeffizienten k -ten Grades ρ_k gilt in Anlehnung an den Produkt-Moment-Korrelationskoeffizienten nach Pearson:

$$\rho_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.3.6)$$

wobei x_i die Beobachtung zum Zeitpunkt i und $\bar{x}$ das arithmetische Mittel aus n Beobachtungen ist (HAAN, 1982; HELSEL & HIRSCH, 2005; SCHÖNWIESE, 2006).

Die Autokorrelation wurde in *R* mit der Funktion `acf()` im Paket `stats` berechnet (CRAWLEY, 2007; R, 2010).

3.4 Univariate Randverteilungen

Zur Ermittlung der kombinierten Wahrscheinlichkeiten für das gemeinsame Auftreten bestimmter Merkmalsausprägungen zweier Variablen, ist es zunächst nötig, deren univariate Verteilungen (auch Rand- oder Marginalverteilungen) zu beschreiben. Die empirischen kumulativen Verteilungsfunktionen der Stichprobenkollektive lassen sich mittels *plotting positions* bestimmen. Dazu werden den Daten auf- oder absteigende Ränge m zugeordnet und die Über- oder Unterschreitungswahrscheinlichkeit $P'(X)$ oder $P(X)$ (je nach auf- oder absteigenden Rängen, vgl. Formel 3.2.2) eines Wertes als Anteil der Datenwerte, welche kleiner oder größer sind als x mittels

$$P(X) = \frac{m}{n} \quad (3.4.1)$$

berechnet, wobei n den Stichprobenumfang bezeichnet. Diese Form der Berechnung, auch bekannt als CALIFORNIA - Formel, hat allerdings die Eigenschaft, dem kleinsten bzw. größten Wert der Stichprobe die Wahrscheinlichkeiten 0 bzw. 1 zuzuweisen. D. h. es wird vorausgesetzt, dass der kleinst- und größtmögliche Wert der Grundgesamtheit bereits gemessen wurde und Teil der Stichprobe ist. Da diese Annahme wohl nur in den allerwenigsten Fällen erfüllt sein kann, bietet hier u. a. die Formel nach WEIBULL mit

$$P(X) = \frac{m}{n-1} \quad (3.4.2)$$

eine Alternative. Diese lässt immer eine "Restwahrscheinlichkeit" offen, mit welcher der größte oder kleinste Wert noch über- oder unterschritten werden kann. Diese Art der Berechnung ist für die Extremwerte des Niederschlags generell gut geeignet, da hier nicht davon ausgegangen werden kann, dass der größtmögliche Wert auch Teil der Stichprobe ist. Für die Bodenfeuchtwerte θ gestaltet sich Auswahl geeigneter *plotting positions* dagegen wesentlich schwieriger, da sich diese innerhalb des Intervalls $[0, 1]$ bewegen. Hier sollte für Maximalwerte $x < 1$ zwar eine gewisse Überschreitungswahrscheinlichkeit $P'(X > x) > 0$ gegeben sein (bzw. für Minimalwerte $x > 0$ sollte $P(X < x) > 0$ sein), für Werte von $\theta = 1$ bzw. $\theta = 0$ müssen $P(\theta < 0)$ bzw. $P'(\theta > 1)$ jedoch 0 sein, um sinnvolle Ergebnisse zu erhalten.

Ein weiterer Nachteil der empirischen Verteilungsfunktionen für diese Analyse ist, dass die erhaltenen Wahrscheinlichkeiten in den Extrembereichen zunehmend unzuverlässig werden und keine Aussagen über Wiederkehrintervalle getroffen werden können, die über den Beobachtungszeitraum hinaus gehen. Für die großräumige Berechnung der *joint probabilities* würde deshalb nur ein sehr kleines Intervall von N und θ - Werten existieren, für welche sich über das gesamte Untersuchungsgebiet sinnvolle und zuverlässige Wiederkehrintervalle berechnen ließen. Um diese für eine Reihe verschiedener Kombinationen von N und θ abschätzen zu können ist es nötig, passende stetige Verteilungsfunktionen zu finden, mit deren Hilfe die gemeinsamen Jährlichkeiten über eine aus-

reichend große Bandbreite inter- und teilweise auch extrapoliert werden können (HAAN, 1982; DYCK & PESCHKE, 1995).

3.4.1 Anpassung theoretischer Verteilungsfunktionen

Es steht eine Vielzahl von Verteilungsfunktionen zur Anwendung bei extremwertstatistischen Frequenzanalysen zur Verfügung. Eine erste Eingrenzung der potentiell tauglichen Verteilungsfunktionen kann zwar bereits anhand allgemeiner Eigenschaften der Variablen und deren empirischer Verteilungen vorgenommen werden. Da es aber nicht möglich ist, einer hydrologischen Variable bzw. einem Prozess von vorneherein eine bestimmte Verteilungsfunktion zuzuordnen, müssen diese letztlich nach statistischen Gütemaßen aus geeigneten Anpassungstests bestimmt werden.

Die Anpassung der theoretischen Verteilungsfunktionen an die Datenreihen wurde mittels des Verfahrens der "optimalen Mutmaßlichkeit", oder besser bekannt als *Maximum Likelihood Estimation* vorgenommen. Die Güte der Anpassung wurde anschließend für einige Stichproben graphisch überprüft und durch den Kolmogorov-Smirnov-Test (KS – Test) für alle Teilflächen bestimmt.

3.4.1.1 Maximum Likelihood Estimation

Die Anpassung der Verteilungsparameter mittels *Maximum Likelihood Estimation* (MLE) lässt sich am einfachsten folgendermaßen ausdrücken: "Ist die Verteilungsfunktion der Grundgesamtheit bekannt, für welche Parameter ($E(X)$ und $Var(X)$) dieser Verteilung sind dann die vorliegenden Stichprobendaten am Wahrscheinlichsten?" (LANGE, 2008) Um dies zu bestimmen, wird die "Mutmaßlichkeitsfunktion" oder *Likelihood* – Funktion

$$L(\Theta) = f(x_1, x_2, \dots, x_n; \Theta) = \prod_{i=1}^n f(x_i; \Theta) \quad (3.4.3)$$

gebildet, welche die Wahrscheinlichkeit für das Stichprobenergebnis als Funktion des unbekannten Parameter(-vektors) Θ angibt. Der Schätzwert $\hat{\Theta}$ für Θ ergibt sich dann als Lösung von

$$\frac{\partial L(\Theta)}{\partial \Theta} = 0 \quad \text{bzw.} \quad \frac{\partial \ln L(\Theta)}{\partial \Theta} = 0 \quad \text{und} \quad \frac{\partial^2 L(\Theta)}{\partial \Theta^2} < 0 \quad (3.4.4)$$

d. h. es wird das Maximum der *Likelihood*-Funktion $L(\Theta)$ bzw. der *Log-Likelihood*-Funktion $\ln L(\Theta)$ durch numerische Iteration bestimmt und nach Θ aufgelöst (SCHÖNWIESE, 2006).

Die Anpassung der Verteilungsfunktionen mittels MLE wurde in *R* mithilfe der Funktionen `fitdist()` im Paket `fitdistrplus` und `gev.fit()` im Paket `ismev` durchgeführt (siehe CRAWLEY, 2007; COWPERTWAIT & METCALFE, 2009; R, 2010).

3.4.1.2 Kolmogorov - Smirnov Test

Um zu testen, ob die gewählten theoretischen Verteilungsfunktionen die empirischen Verteilungen der Stichproben für alle 65 645 Modellflächen hinreichend gut wiedergeben können, wird ein flexibles und effizientes Testverfahren benötigt, mit dessen Hilfe die Güte der Anpassung für alle ausgewählten Verteilungsfunktionen gleichermaßen bestimmt werden kann. Für kontinuierliche, nicht-normalverteilte Zufallsvariablen bietet sich hier der Kolmogorov-Smirnov Test (KS-Test) an, welcher als Teststatistik d die Distanz zwischen der theoretischen kumulativen Verteilungsfunktion $F(x_i)$ und der empirischen Verteilung $P(x_i)$ an jedem Punkt $i = 1 \dots n$ berechnet und das Maximum der absoluten Differenz $d_{\max}$ ermittelt:

$$d_{\max} = \max \left(\left| F(x_i) - P(x_i) \right| \right) \quad (3.4.5)$$

Die Nullhypothese H_0 des Tests besagt hier, dass $F(X) = P(X)$, dass also die Stichprobenverteilung $P(X)$ einer Grundgesamtheit mit der theoretischen Ver-

teilung $F(X)$ entstammt. Die Alternativhypothese H_1 lautet dementsprechend $F(X) \neq P(X)$. Die Entscheidung, die Nullhypothese anzunehmen oder abzulehnen kann hier wieder auf zwei verschiedene Arten getroffen werden. Zum einen kann $d_{\max}$ mit tabellierten kritischen Werten verglichen werden, welche für ein vorgegebenes Signifikanzniveau α nicht überschritten werden dürfen (siehe Tabelle 3.1). Andererseits kann auch hier (wie es in R der Fall ist) wieder ein p -value ermittelt werden (vgl. Kapitel 3.3.1), der die Wahrscheinlichkeit $P(d_{\max} < D_n)$ mittels einer speziell dafür entwickelten, multimodalen Spline-Verteilungsfunktion angibt (näheres dazu siehe MARSAGLIA et al., 2003). Auch hier wird die Nullhypothese zugunsten der Alternativhypothese abgelehnt, wenn der p -value das gewählte Signifikanzniveau (z. B. $\alpha = 0.05$) unterschreitet (HAAN, 1982; LANGE, 2008).

Tabelle 3.1: Kritische Werte D_n für $d_{\max}$ des Kolmogorov-Smirnov Tests bei gegebenem Signifikanzniveau α . (SCHÖNWIESE, 2006)

α	0.20	0.10	0.05	0.01	0.001
D_n	$1.073 \sqrt{n}$	$1.224 \sqrt{n}$	$1.358 \sqrt{n}$	$1.628 \sqrt{n}$	$1.040 \sqrt{n}$

3.4.2 Generalisierte Extremwertverteilung (Niederschlag)

Zur Modellierung der Häufigkeitsverteilung der saisonalen 24h - Niederschlagsmaxima stehen vorerst eine Reihe von Verteilungsfunktionen zur Auswahl. Nach einer ersten graphischen Eingrenzung potentiell geeigneter Verteilungsfunktionen wurde schließlich die Generalisierte Extremwertverteilung (GEV) ausgewählt. Die GEV ist die verallgemeinerte Form der Extremwertverteilungen der Typen I – III (Gumbel, Fréchet und Weibull) und besitzt durch ihre 3-parametrische Form eine enorme Flexibilität (JENKINSON, 1955; HAAN, 1982). Die kumulative Verteilungsfunktion der GEV lautet

$$P(X; \mu, \sigma, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (3.4.6)$$

mit $\mu \in [-\infty, +\infty]$ als Lageparameter (*location*), $\sigma < 0$ als Streuparameter (*scale*) und $\xi \in [-\infty, +\infty]$ als zusätzlichen Formparameter (*shape*), welcher die Form des oberen *tailings*, also der Wahrscheinlichkeiten im oberen Extrembereich der Verteilungsfunktion beeinflusst. Erwartungswert $E(X)$ und Standardabweichung $Sd(X)$ der GEV können mithilfe der Gammafunktion Γ folgendermaßen berechnet werden:

$$E(X) = \mu - \frac{\sigma}{\xi} + \frac{\sigma}{\xi} g_1 \quad (3.4.7)$$

$$Sd(X) = \sqrt{\frac{\sigma^2}{\xi^2} (g_2 - g_1^2)} \quad (3.4.8)$$

wobei $g_k = \Gamma(1 - k\xi)$ und $k = [1, 2]$ (JENKINSON, 1955; COLES, 2001; VILLARINI et al., 2011).

3.4.3 Beta-Verteilung (Bodenfeuchte)

Wie in Kapitel 3.4 bereits erläutert, ist das wichtigste Kriterium für die Verteilungsfunktion der Bodenfeuchte eine links- und rechtsseitige Begrenzung im Intervall $[0, 1]$, außerdem sollte sie in ihrer Form möglichst flexibel sein. Diese Anforderungen werden von der Beta-Verteilung (BV) erfüllt, welche auch von RAVELO & DECKER (1979) zur Modellierung der Bodenfeuchte empfohlen wird. Die Beta-Verteilung ist mit der Gammaverteilung verwandt und kann sowohl extrem links- oder rechtsschiefe Glockenformen als auch exponentiell verlaufende Kurvenformen annehmen. Die 2-parametrische Form ist beidseitig auf das Intervall $[0, 1]$ begrenzt, mit zwei zusätzlichen Parametern lässt sie sich jedoch auch auf jedes beliebige Intervall erweitern (was in diesem Fall nicht nötig ist). Die Formel für die kumulative Verteilungsfunktion der BV mit den Parametern α und β lautet nach RAVELO & DECKER (1979) und HAAN (1982)

$$P(X; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx \quad (3.4.9)$$

wobei $B(\alpha, \beta)$ die Betafunktion darstellt, welche mit der Gammafunktion Γ folgendermaßen in Verbindung steht:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad (3.4.10)$$

Erwartungswert $E(X)$ und Standardabweichung $Sd(X)$ der Beta-Verteilung lassen sich aus α und β berechnen:

$$E(X) = \frac{\alpha}{\alpha + \beta} \quad (3.4.11)$$

$$Sd(X) = \sqrt{\frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}} \quad (3.4.12)$$

3.5 Joint probabilities

Für die univariate Frequenzanalyse hydrologischer (Extrem-) Ereignisse sind die grundlegenden Konzepte bezüglich Überschreitungswahrscheinlichkeiten und Wiederkehrintervallen im Allgemeinen weitgehend bekannt und bewährt (MANIAK, 1997; YUE & RASMUSSEN, 2002). Für Ereignisse, welche durch mehrere (evtl. korrelierte) Zufallsvariablen beschrieben werden müssen, ist es nicht offensichtlich, was unter einem Ereignis mit einem bestimmten Wiederkehrintervall zu verstehen ist. YUE & RASMUSSEN (2002) geben hier den im Folgenden kurz zusammengefassten, einführenden Überblick:

Für den univariaten Fall lassen sich die Wahrscheinlichkeiten der Messwerte eindeutig zuordnen. Die kumulative Verteilungsfunktion $F_X(x)$ der Zufallsvariable X gibt an, mit welcher Wahrscheinlichkeit X einen gegebenen Wert x nicht überschreitet:

$$F_X(x) = \Pr[X \leq x] \quad (3.5.1)$$

Daraus lässt sich (wie in Formel 3.2.2 bereits ausgedrückt) die Überschreitungswahrscheinlichkeit $F'_X(x) = \Pr[X \geq x]$ als Komplementärwahrscheinlichkeit von $F_X(x)$ berechnen:

$$F'_X(x) = 1 - F_X(x) \quad (3.5.2)$$

Die entsprechenden Wiederkehrintervalle $T_X(x)$ lassen sich dann jeweils durch Bildung des Kehrwertes aus der Überschreitungswahrscheinlichkeit bilden.

$$T_X(x) = \frac{1}{F'_X(x)} = \frac{1}{1 - F_X(x)} \quad (3.5.3)$$

Folglich lässt sich $F_X(x)$ auch durch $T_X(x)$ darstellen:

$$F_X(x) = \frac{T_X(x) - 1}{T_X(x)} \quad (3.5.4)$$

Die physikalische Bedeutung der Wiederkehrintervalle bleibt zwar für den bivariaten Fall die gleiche wie für den univariaten Fall, dennoch ist es nicht eindeutig, wie die kombinierten Wahrscheinlichkeiten den bivariaten Ereignissen mit den Variablen X und Y zugeordnet werden sollen. Um dies zu Veranschaulichen, kann man sich den gesamten "Wahrscheinlichkeitsraum" von X und Y als Ebene der Fläche 1 vorstellen, welche (wie in Abbildung 3.5 dargestellt), in 4 Quadranten unterteilt ist.

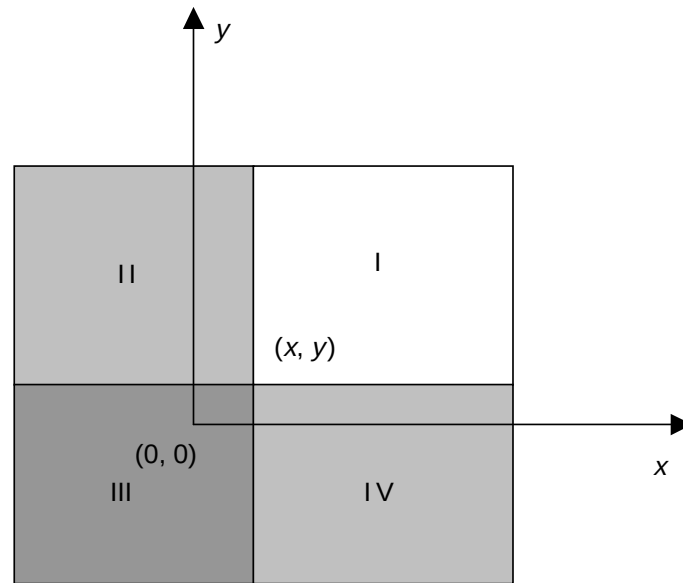


Abbildung 3.5: Einteilung des X-Y Wahrscheinlichkeitsraumes zur Veranschaulichung der kombinierten, bivariaten Wahrscheinlichkeiten zweier Variablen X und Y. (YUE & RASMUSSEN, 2002)

Die Wahrscheinlichkeiten, dass ein bivariates Ereignis (X, Y) in einen der Quadranten I – IV fällt, lassen sich folgendermaßen ausdrücken ($f(x,y)$ bezeichnet dabei die theoretische bivariate Dichtefunktion der Grundgesamtheit):

Quadrant I:

$$F_I(x,y) = \Pr[X > x, Y > y] = \int_x^{+\infty} \int_y^{+\infty} f(x,y) \, dx \, dy \quad (3.5.5)$$

$$= F'(x,y)$$

Dies ist also die Wahrscheinlichkeit, dass beide Variablen X und Y gleichzeitig die Werte x und y überschreiten.

Quadrant II:

$$F_{II}(x,y) = \Pr[X \leq x, Y > y] = \int_{-\infty}^x \int_y^{+\infty} f(x,y) \, dx \, dy$$

$$\begin{aligned}
&= \int_{-\infty}^x \left[\int_{-\infty}^{+\infty} f(x,y) \, dy - \int_{-\infty}^y f(x,y) \, dy \right] dx \\
&= F_X(x) - \int_{-\infty}^x \int_{-\infty}^y f(x,y) \, dx \, dy
\end{aligned} \tag{3.5.6}$$

Quadrant III:

$$\begin{aligned}
F_{III}(x,y) &= \Pr[X \leq x, Y \leq y] \\
&= F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(x,y) \, dx \, dy
\end{aligned} \tag{3.5.7}$$

Quadrant IV:

$$\begin{aligned}
F_{IV}(x,y) &= \Pr[X > x, Y \leq y] = \int_x^{+\infty} \int_{-\infty}^y f(x,y) \, dx \, dy \\
&= \int_{-\infty}^y \left[\int_{-\infty}^{+\infty} f(x,y) \, dx - \int_{-\infty}^x f(x,y) \, dx \right] dy \\
&= F_Y(y) - F(x,y)
\end{aligned} \tag{3.5.8}$$

Daraus folgt

$$\begin{aligned}
&\Pr[-\infty < X < +\infty, -\infty < Y < +\infty] = \\
&= F_I(x,y) + F_{II}(x,y) + F_{III}(x,y) + F_{IV}(x,y) = 1
\end{aligned} \tag{3.5.9}$$

und damit

$$F'(x,y) = 1 + F(x,y) - F_X(x) - F_Y(y) \tag{3.5.10}$$

Das gemeinsame Wiederkehrintervall des Ereignisses $(X > x, Y > y)$ kann also durch

$$T(x, y) = \frac{1}{1 - F(x, y)} \quad \text{mit } F(x, y) = \Pr[X \leq x, Y \leq y] \quad (3.5.11)$$

definiert werden. $T(x, y)$ stellt den Fall dar, dass entweder x oder y oder beide ($X > x \vee Y > y \vee X > x \wedge Y > y$) überschritten werden. Das gemeinsame Wiederkehrintervall $T'(x, y)$ eines Ereignisses, für das x und y gleichzeitig überschritten werden ($X > x \wedge Y > y$) kann mit

$$T'(x, y) = \frac{1}{F'(x, y)} = \frac{1}{1 + F(x, y) - F_X(x) - F_Y(y)} \quad (3.5.12)$$

bestimmt werden.

Sind die betrachteten Variablen voneinander unabhängig, so kann deren kombinierte Verteilungsfunktion durch

$$F(x, y) = F_X(x)F_Y(y) = \frac{(T_X - 1)(T_Y - 1)}{T_X T_Y} \quad (3.5.13)$$

beschrieben werden. Folglich können die bivariaten (kombinierten) Wiederkehrintervalle $T(x, y)$ und $T'(x, y)$ aus den univariaten T_X und T_Y berechnet werden:

$$T(x, y) = \frac{T_X T_Y}{T_X + T_Y - 1} = \frac{T_X}{\frac{T_X - 1}{T_Y} + 1} = \frac{T_Y}{\frac{T_Y - 1}{T_X} + 1} \quad (3.5.14)$$

$$T'(x, y) = T_X T_Y \quad (3.5.15)$$

Da $T_X \geq 1$ und $T_Y \geq 1$, lassen sich folgende Ungleichungen formulieren:

$$T(x, y) \leq \min(T_X, T_Y) \leq \max(T_X, T_Y) \leq T'(x, y) \quad (3.5.16)$$

Im Falle von $T_X = T_Y = T$ lassen sich $T(x, y)$ und $T'(x, y)$ vereinfacht durch

$$T(x, y) \approx T/2 \quad (3.5.17)$$

$$T'(x, y) = T^2 \quad (3.5.18)$$

abschätzen. Für $T_X = T_Y = 100$ Jahre würde sich also $T(x, y) = 50.25$ Jahre und $T'(x, y) = 10\,000$ Jahre ergeben. Die Eintrittswahrscheinlichkeit des Ereignisses mit beiden $X > x$ und $Y > y$ ist also sehr viel geringer als die eines Ereignisses bei dem nur eine der beiden Variablen ($X > x \vee Y > x$) einen Grenzwert überschreitet. Für die folgende Analyse wird ausschließlich $F'(x, y)$ bzw. $T'(x, y)$ d.h. der für Quadrant I gültige Fall betrachtet.

Die bedingten Wahrscheinlichkeiten $F(x|y)$, also die Wahrscheinlichkeiten für ein Ereignis mit $X > x$ für ein gegebenes $Y = y$, entsprechen bei unabhängigen Variablen den jeweiligen univariaten Randwahrscheinlichkeiten, also

$$F(x|y) = F_X(x) \quad \text{und} \quad F(y|x) = F_Y(y) \quad (3.5.19)$$

sowie

$$T(x|y) = T_X(x) \quad \text{und} \quad T(y|x) = T_Y(y) \quad (3.5.20)$$

Der Fall abhängiger Variablen soll hier nicht genauer Ausgeführt werden, da er für diese Studie nicht relevant ist. Allgemein lässt sich jedoch sagen, dass die gemeinsamen Wiederkehrintervalle mit zunehmender gegenseitiger Abhängigkeit dramatisch abnehmen. Abbildung 3.6 zeigt die gemeinsamen Wiederkehrintervalle $T'(x, y)$ für $T_X = T_Y$ gegen den Abhängigkeitsparameter von m . Dieser stammt aus dem von YUE & RASMUSSEN (2002) verwendeten, bivariaten Gumbel-Logistic Modell und stellt die Stärke der Korrelation als Wert zwischen 1 und 5 dar, wobei 1 für Unabhängigkeit und 5 für perfekte Korrelation steht.

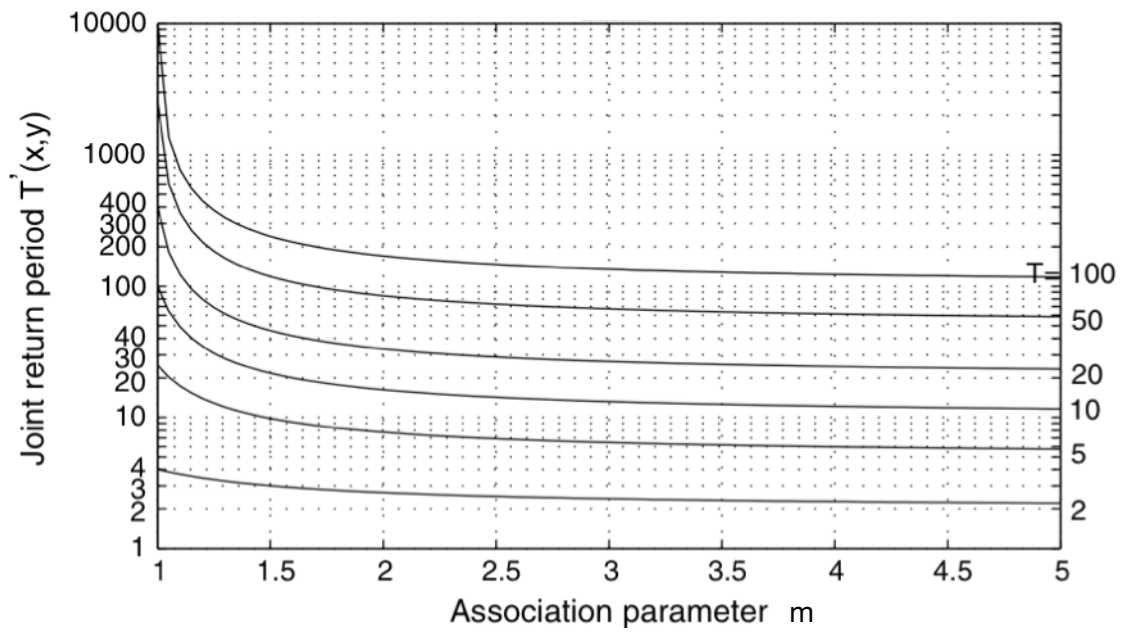


Abbildung 3.6: Gemeinsame Wiederkehrintervalle $T'(x, y)$ gegen Abhängigkeitsparameter m des Gumbel-Logistic-Modells. $T'(x, y) \sim m$ für $T_x = T_y = 2, 5 \dots 100$ Jahre.

4 Ergebnisse

Da aufgrund der saisonalen Differenzierung in Sommer- und Winterhalbjahr (Kap. 3.2.3) zur Darstellung der in Tabelle 2.1 ausgewählten 6 Teilflächen jeweils 12 Einzelgrafiken nötig sind, werden im Folgenden meist nur 2 - 4 dieser Punkte dargestellt. Die Grafiken für die übrigen Punkte sind mit Referenz auf die jeweilige Abbildung im Anhang zu finden.

4.1 Korrelationsanalyse

Die Korrelationsanalyse wurde zunächst mit den in Kapitel 3.1 und 3.2 besprochenen Extremwertstichproben aus den je 39 saisonalen Niederschlagsmaxima am Tag t und den korrespondierenden Bodenfeuchtwerten am Vortag ($t - 1$) durchgeführt. Um neben der Autokorrelation selbst zusätzliche Hinweise auf den in Kapitel 3.3 angedeuteten Einfluss der Niederschlagsverhältnisse am Tag ($t - 1$) auf die Bodenfeuchte zu erhalten, wurden zum Vergleich alternative Stichproben mit der Bedingung $N_{t-1} < 5$ mm ausgewählt (im Folgenden als "Vortag trocken" oder "VT-Stichproben" bezeichnet), sprich es werden zunächst die Niederschlagstageswerte, für die am Vortag weniger als 5 mm Niederschlag gefallen ist (und die klimatischen Verhältnisse in Bezug auf den Bodenwasserhaushalt also als "trocken" angesehen werden können) ausgewählt und daraus die saisonalen Maxima entnommen.

4.1.1 Korrelation von N und θ_v

Für einen ersten Eindruck über die gewonnenen Stichproben sind in Abbildung 4.1 die jährlichen Niederschlags- und Bodenfeuchtwerte für die Punkte A, C und E gegeneinander dargestellt. Abbildung 4.2 zeigt die räumliche Verteilung der Korrelationskoeffizienten, die Bereiche mit signifikanten Korrelationen sind für $\alpha = 5\%$ ($-0.183 > \tau > 0.183$) schwarz schraffiert hervorgehoben. Für dieses Signifikanzniveau weisen im Sommer insgesamt 14.6% und im Winter 10.1% der

Gesamtfläche Baden-Württembergs signifikante, vorwiegend positive Korrelationen auf.

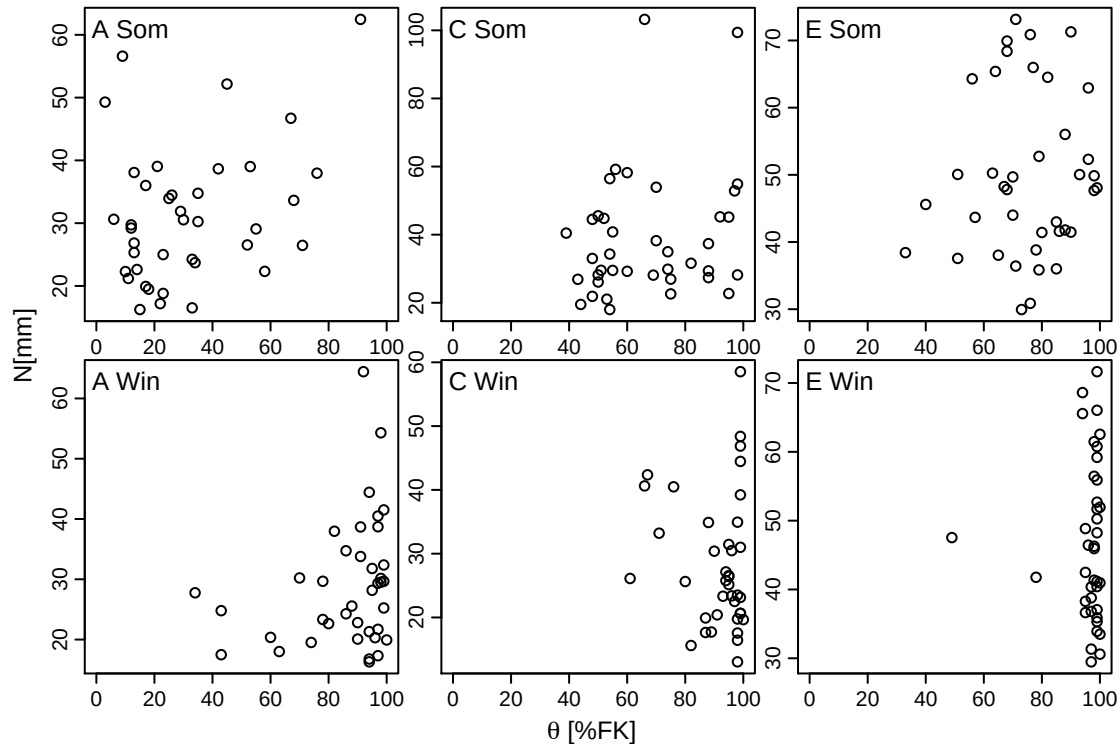


Abbildung 4.1: Scatterplot der saisonalen $N - \theta_v$ Stichproben für die Punkte A, C und E (siehe Tabelle 2.1).

Für die zusätzlichen VT-Stichproben ergeben sich mit ca. 8.4% im Sommer und ca. 7.9% im Winter deutlich geringere Flächenanteile mit signifikanten Korrelationen. Gemäß der Nullhypothese ($H_0: \tau = 0$) des Korrelationstests nach Kendall lässt sich die Häufigkeitsverteilung von τ für unabhängige Variablen mit einer Normalverteilung ($\mu = 0$, σ siehe Formel 3.3.4) beschreiben. Angenommen, die Stichproben aller Teilflächen stammen aus der selben Grundgesamtheit, sollte diese Normalverteilung auch die empirische Verteilung von τ über die Fläche beschreiben können. Abbildung 4.3 zeigt die Histogramme der τ -Werte jeweils für die Sommer- und Winterstichproben (a) und zum Vergleich für die VT-Stichproben (b), in Tabelle 4.1 sind die Momente der empirischen Verteilungen sowie der unter H_0 angenommenen Verteilung von a und b zusammengefasst.

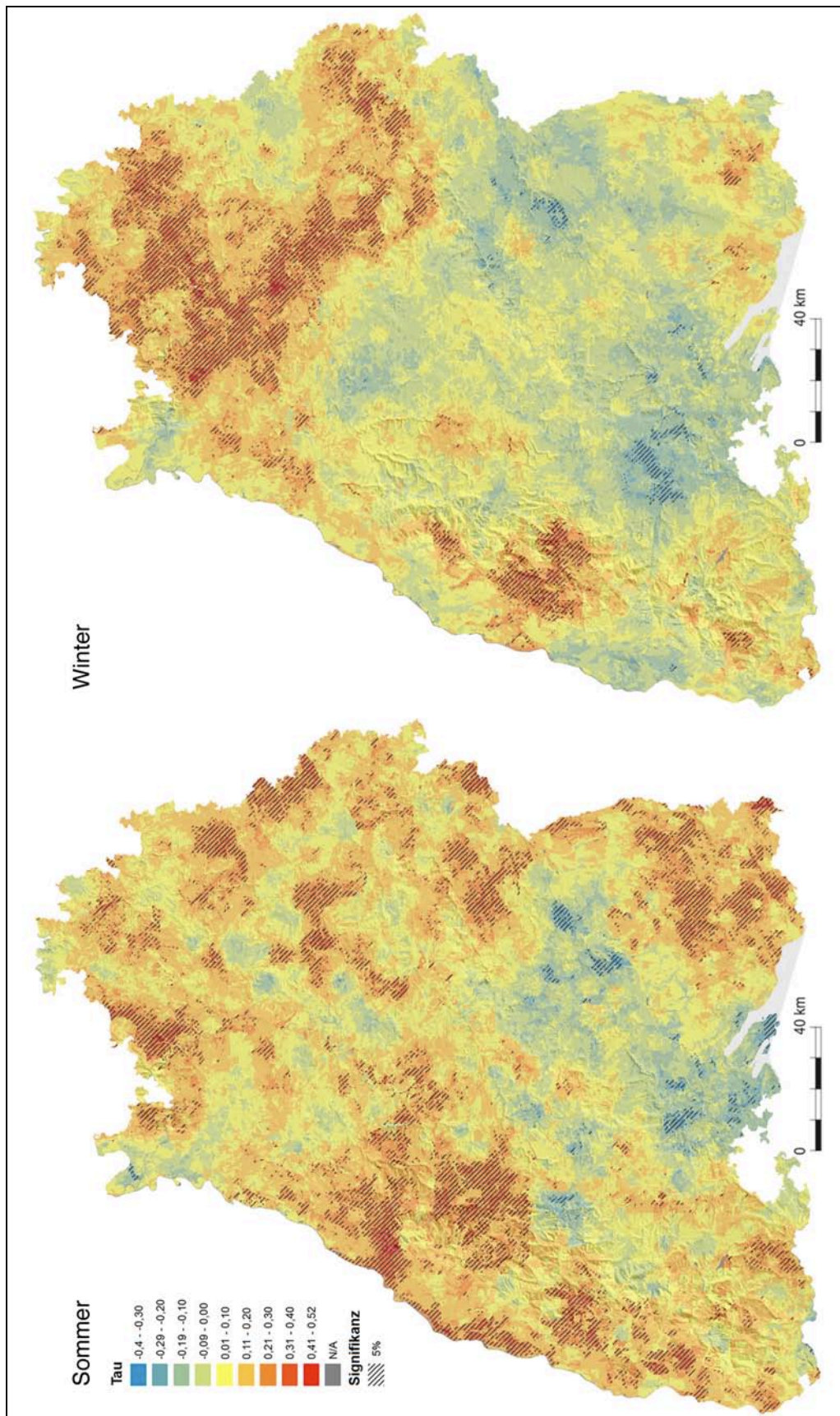


Abbildung 4.2: Rangkorrelationskoeffizient τ von saisonalen Niederschlagsmaxima und Vorfeuchte für Sommer (links) und Winter (rechts). Schraffierte Flächen zeigen signifikante Korrelationen bei $\alpha = 5\%$.

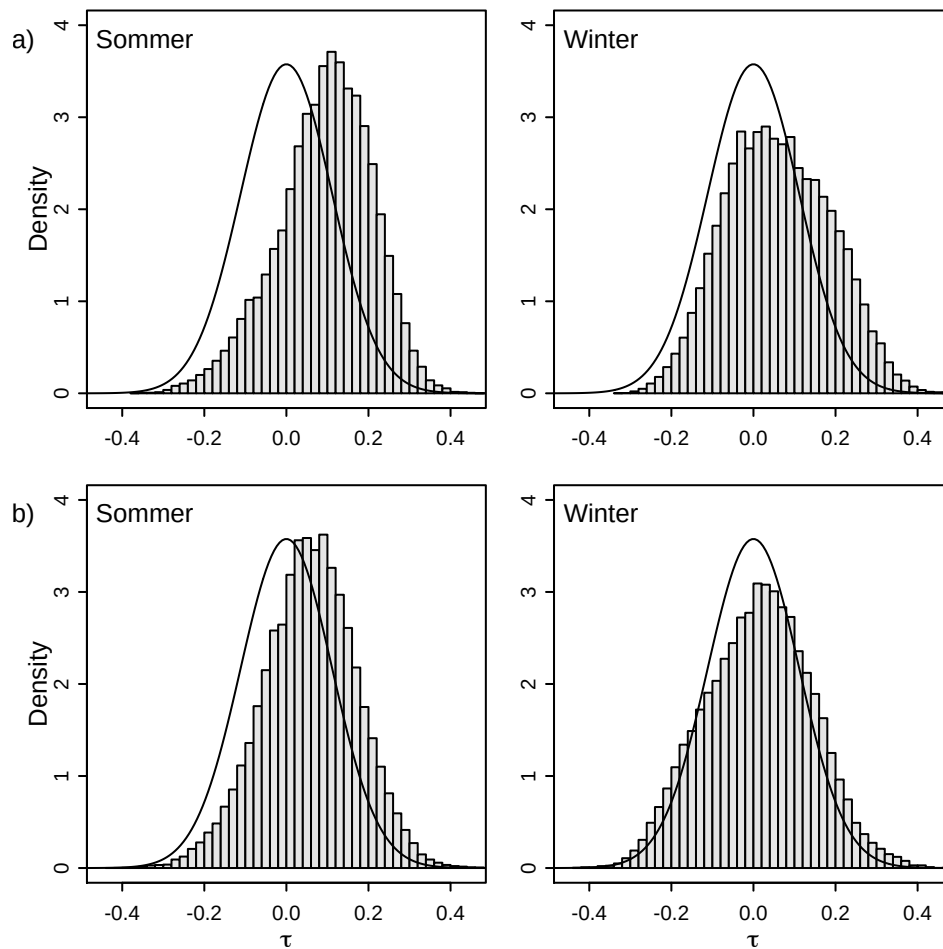


Abbildung 4.3: Häufigkeitsverteilung des Rangkorrelationskoeffizienten τ (graue Balken) und angenommene Normalverteilung mit $\mu = 0$ gemäß der Nullhypothese $H_0: \tau = 0$ (durchgezogene Linie). a) für Sommer und Winter ohne Bedingung, b) mit Bedingung $N_{t-1} < 5$ mm ("Vortag trocken").

Tabelle 4.1: Momente der Häufigkeitsverteilungen von τ über die Gesamtfläche.

Stichprobe	Saison	empirische Verteilung			
		μ	σ	Q_{05}	Q_{95}
a) ohne Bedingung	Sommer	0.09	0.12	-0.11	0.27
	Winter	0.06	0.13	-0.14	0.27
b) "Vortag trocken"	Sommer	0.06	0.11	-0.13	0.24
	Winter	0.01	0.13	-0.20	0.22
theoretische Verteilung gemäß $H_0 (\mu=0)$					
		0	0.11	-0.18	0.18

4.1.2 Autokorrelation des Niederschlags

Die Autokorrelationsanalyse für die Tagessummen des Niederschlags wurde wie in Kapitel 3.3.2 beschrieben für die kontinuierlichen, in Sommer- und Winterhalbjahr eingeteilten Zeitreihen durchgeführt. Die Inkonsistenz, welche sich aus dieser Aufteilung jeweils an den wechseln zwischen den Halbjahren ergibt wurde zugunsten der Rechenzeit vernachlässigt. Da hier auf jedes "falsche" Wertepaar (jeweils 31.10 + 01.05 und 31.04 + 01.11) mindestens 182 "richtige" Paare kommen, scheint diese Vernachlässigung vertretbar zu sein. Um den Einfluss der Autokorrelation 1. Grades ρ_1 (vgl. Kap. 3.3) auf die Korrelation von Niederschlag und Vorfeuchte τ einschätzen zu können, sind diese in Abbildung 4.4 gegeneinander dargestellt, eine weitere Korrelationsanalyse zwischen den beiden Koeffizienten ergab für beide Jahreszeiten stark signifikante Abhängigkeiten (p-Wert $\ll 0.01$) mit $\tau = 0.21$ im Sommer und $\tau = 0.12$ im Winter. Abbildung 4.5 zeigt die räumliche Verteilung des Autokorrelationskoeffizienten ρ_k für $k = 1$ Tag.

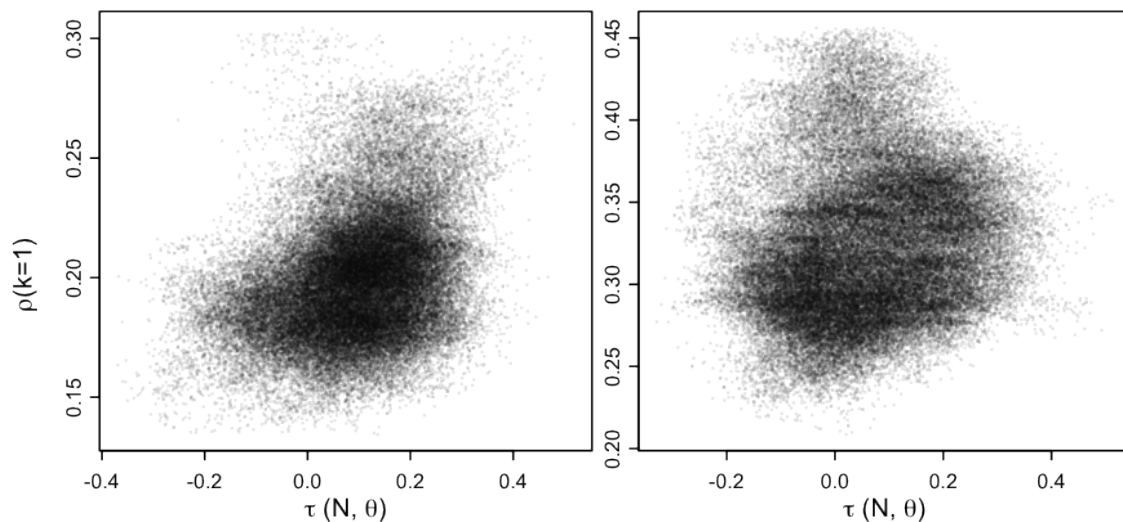


Abbildung 4.4: Autokorrelation 1. Grades ρ_1 des Niederschlags gegen Rangkorrelation τ von Niederschlag N und Vorfeuchte θ für Sommer (links) und Winter (rechts).

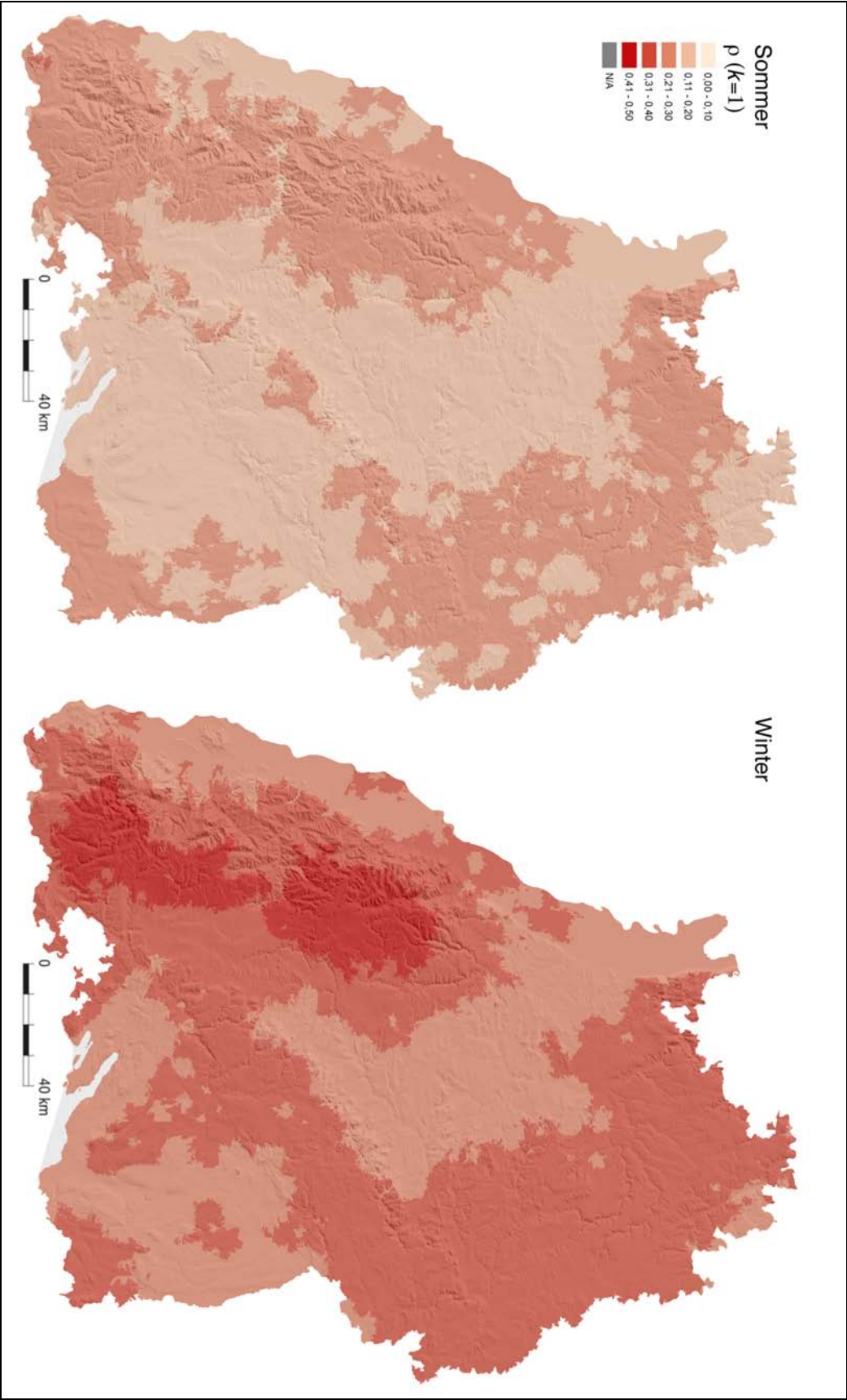


Abbildung 4.5: Autokorrelationskoeffizient $p(k = 1 \text{ Tag})$ der Niederschlags -Tagessummen über 39 Jahre, aufgeteilt in Sommer (links) und Winter (rechts).

4.2 Anpassung der Randverteilungen

Die Anpassung der theoretischen Verteilungsfunktionen (GEV für den Niederschlag und BV für die Bodenvorfeuchte) wurde, wie in Kapitel 3.4 beschrieben, mittels MLE durchgeführt und die Güte der Anpassung anschließend durch den Kolmogorov-Smirnov Test überprüft.

4.2.1 Niederschlag

Die in Abbildung 4.6 exemplarisch für die Punkte A, C und D dargestellte Anpassung der Generalisierten Extremwertverteilung an die saisonalen Niederschlagsmaxima lieferte insgesamt sehr gute Ergebnisse. Die Häufigkeitsverteilung der p-Werte des KS-Tests in Abbildung 4.7 zeigt, dass die Nullhypothese bei $\alpha = 0.05$ für keine der 65 645 Teilflächen abgelehnt wurde. Die mittleren p-Werte betragen im Sommer 0.86 und im Winter 0.84.

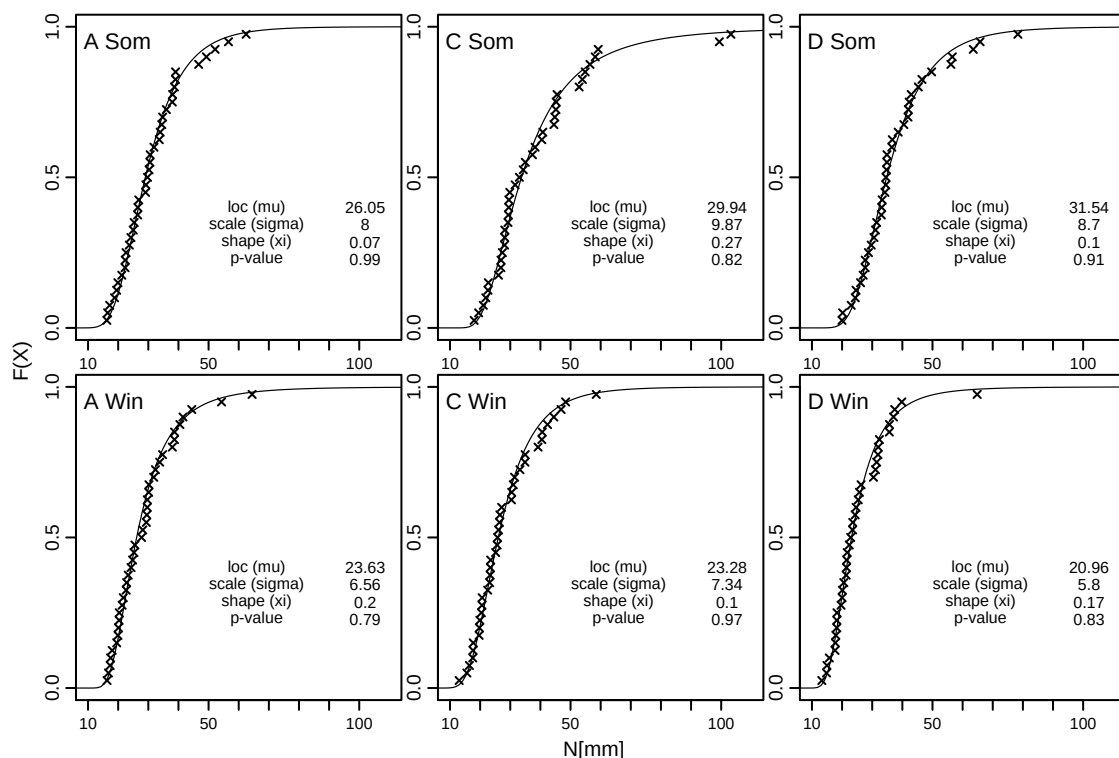


Abbildung 4.6: Empirische Wahrscheinlichkeitsverteilung (x) und GEV (—) der saisonalen Niederschlagsmaxima mit μ = Lage- (*loc*), σ = Streu- (*scale*) und ξ = Formparameter (*shape*) sowie dem p-Wert des KS-Tests für die Punkte A, C und D.

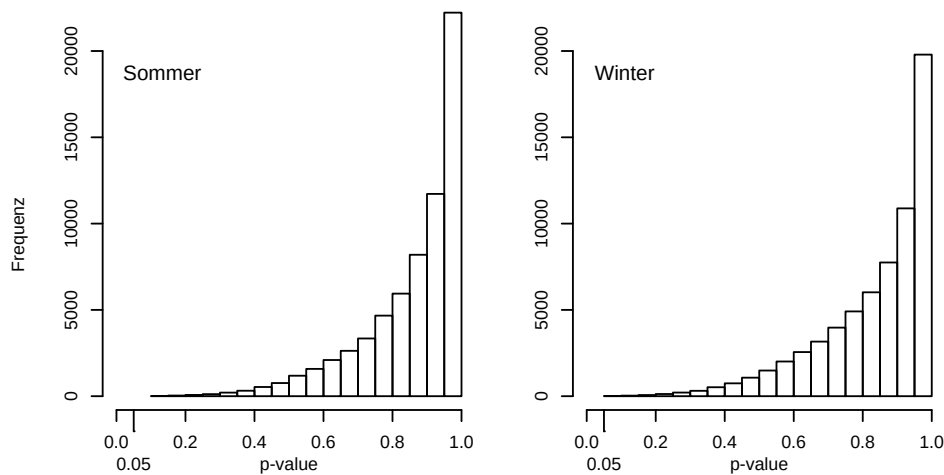


Abbildung 4.7: Häufigkeitsverteilung der p-Werte des K-S Tests für die Anpassung der GEV an die empirischen Verteilungen der saisonalen Niederschlagsmaxima für Sommer (links) und Winter (rechts). Die Nullhypothese wurde bei $\alpha = 0.05$ für keine der Teilflächen abgelehnt.

In Tabelle 4.2 ist die Verteilung der Parameter der GEV sowie deren Erwartungswert und Standardabweichung (im Folgenden mit E und Sd bezeichnet um Verwechslung mit Lage- und Streuparameter μ und σ der GEV zu vermeiden) über alle Teilflächen aufgeführt. Abbildung 4.8 zeigt die räumliche Verteilung von E und Sd , jeweils für das Sommer- und Winterhalbjahr.

Tabelle 4.2: Quantile der Momente und Parameter der GEV über alle Teilflächen ($n = 65\,645$).

Saison	Moment/ Parameter	Min.	Q ₀₅	Median	Q ₉₅	Max.	Mittelw.	Sd
Sommer	E [mm]	27.1	30.8	35.3	46.4	65.1	36.5	4.8
	Sd [mm]	7.1	9.2	12.4	22.5	519.1	14.0	8.8
	μ (loc)	21.0	24.9	29.6	39.8	52.7	30.6	4.5
	σ (scale)	4.2	6.3	8.2	10.8	15.7	8.3	1.4
	ξ (shape)	-0.5	-0.1	0.1	0.4	0.9	0.1	0.2
Winter	E [mm]	18.3	22.7	27.4	45.9	78.4	29.5	7.4
	Sd [mm]	4.7	7.2	10.6	18.1	142.7	11.4	4.0
	μ (loc)	16.0	18.7	22.6	38.9	65.3	24.5	6.5
	σ (scale)	3.0	5.1	6.5	11.1	20.6	7.0	2.0
	ξ (shape)	-0.8	-0.1	0.1	0.3	0.7	0.1	0.1

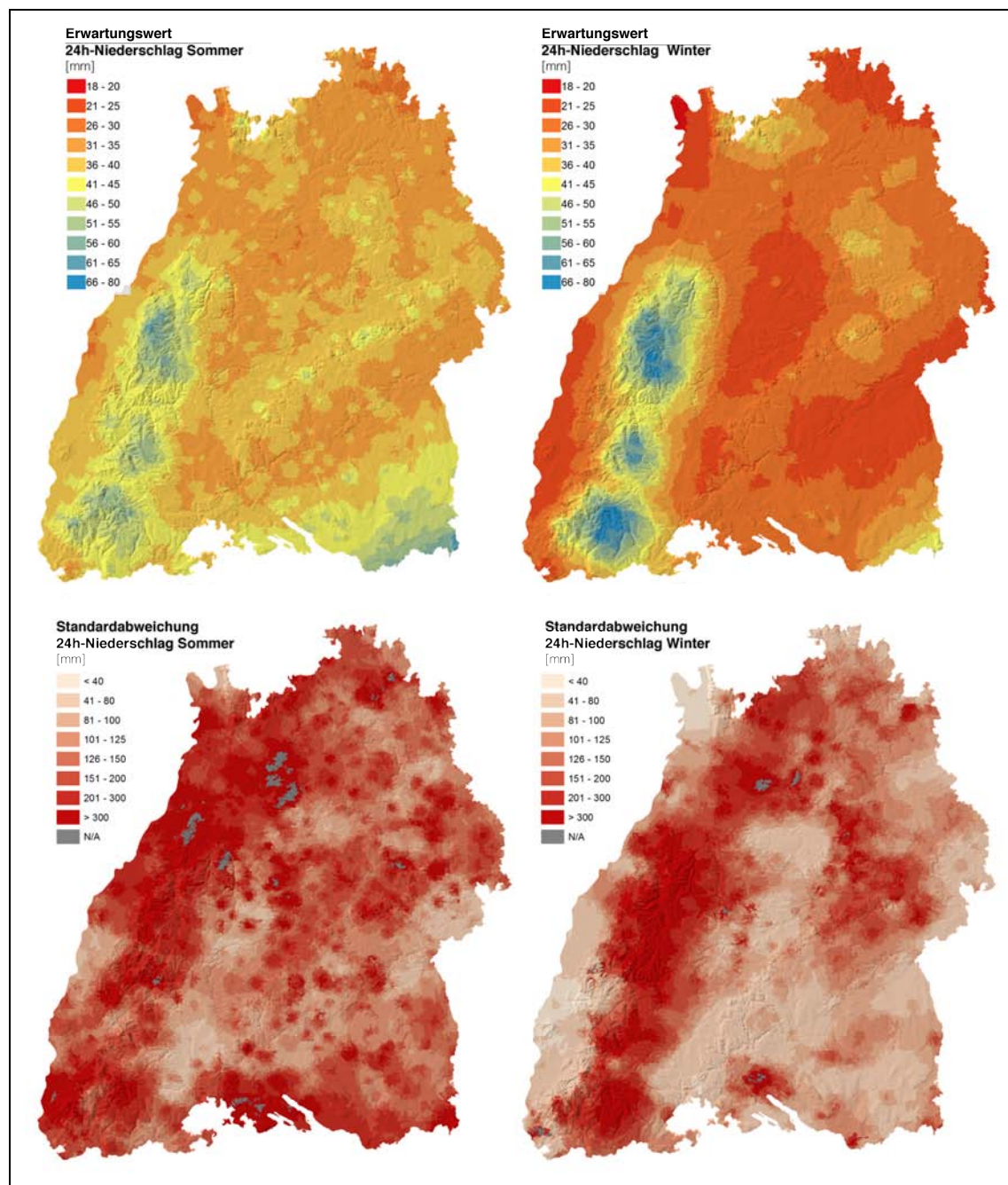


Abbildung 4.8: Räumliche Verteilung von Erwartungswert und Standardabweichung der GEV für saisonale Niederschlagsmaxima im Sommer (links) und Winter (rechts).

4.2.2 Bodenfeuchte

Der Vergleich zwischen den empirischen Verteilungen und den daran angepassten Beta-Verteilungen ist in Abbildung 4.9 dargestellt. Die Anpassung der Beta-

Verteilung erzielt, wie aus Abbildung 4.10 ersichtlich wird, insgesamt weniger gute, aber dennoch akzeptable Ergebnisse.

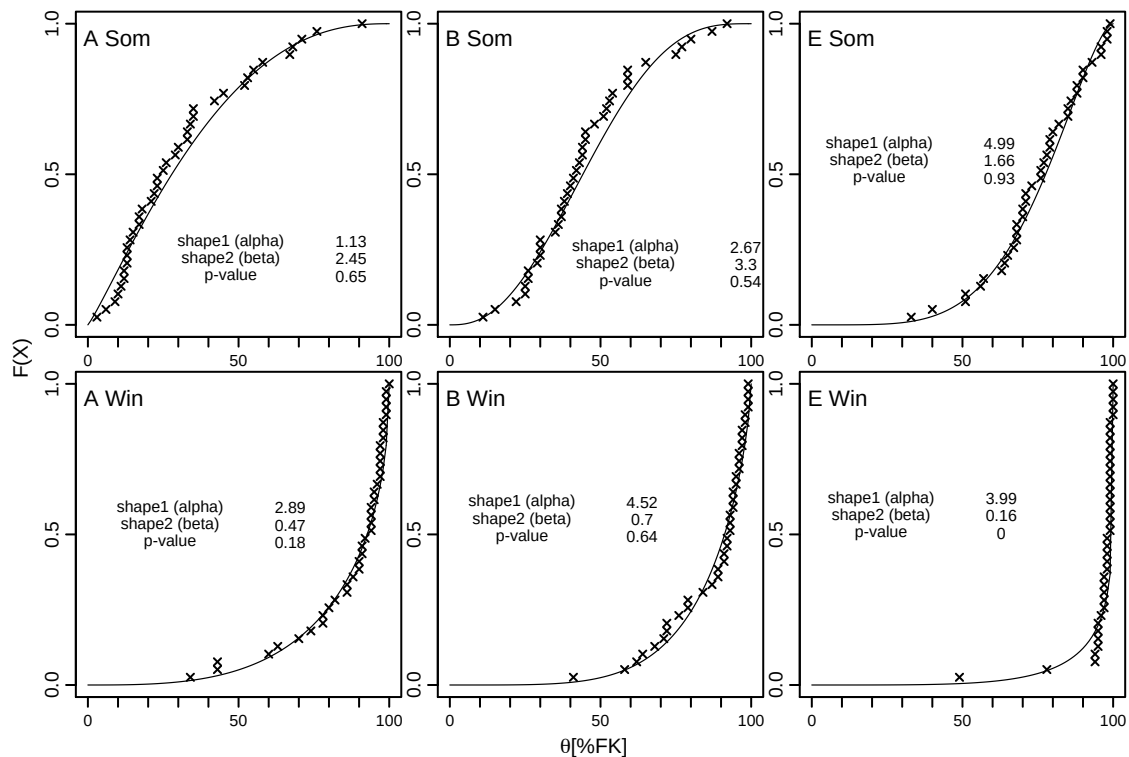


Abbildung 4.9: Empirische Wahrscheinlichkeitsverteilung (x) und Beta-Verteilung (—) der Vorfeuchtebedingungen θ_v mit den Parametern $\alpha = \text{shape1}$ und $\beta = \text{shape2}$ sowie dem p-Wert des KS-Tests für die Punkte A, B und E.

Besonders für das Winterhalbjahr konnte die Nullhypothese des KS-Tests für insgesamt 1120 Teilflächen, was etwa 20.2% der Gesamtfläche entspricht, nicht bestätigt werden. Diese Flächen befinden sich hauptsächlich in Gebieten mit sehr hohem Erwartungswert (sprich im Winter annähernd dauergesättigten Flächen mit $\theta > 90\%FK$) und sehr geringer Varianz (siehe Abbildung 4.11 rechts) wie dem Hochschwarzwald, dem südlichem Odenwald sowie dem westlichen Voralpenland nördlich des Bodensees. Durch die hier häufig auftretenden *ties* (gleiche Werte) wird die Anpassung der stetigen Verteilungsfunktion zwangsläufig unsicherer (vgl. Abbildung 4.9, Punkt E -Winter).

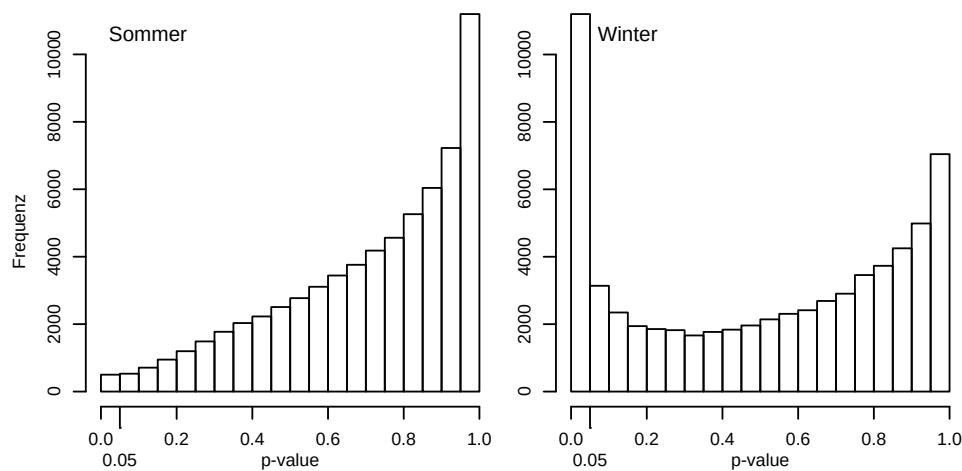


Abbildung 4.10: Häufigkeitsverteilung der p-Werte des K-S Tests für die Anpassung der Beta-Verteilung an die Vorfeuchtebedingungen θ_v für Sommer (links) und Winter (rechts). H_0 wurde bei $\alpha = 0.05$ im Sommer für insgesamt 0.8%, im Winter für ca. 20.2% der Gesamtfläche abgelehnt.

Für das Sommerhalbjahr wurde insgesamt nur für 501 Teilflächen ($< 0.8\%$ der Gesamtfläche) ein p-Wert < 0.05 ermittelt, welche ohne erkennbare Muster über die Landesfläche verstreut liegen. Eine stichprobenhafte Sichtung dieser Teilflächen zeigt, dass die θ_v -Verteilungen hier annähernd einer Gleichverteilung entsprechen, welche die Beta-Verteilung nur ungenügend beschreiben kann. Die Verteilung von Mittel- bzw. Erwartungswert E , Standardabweichung Sd sowie der beiden Parameter α und β über die Fläche sind in Tabelle 4.3 zusammengefasst, Abbildung 4.11 zeigt die räumliche Verteilung von E und Sd .

Tabelle 4.3: Quantile der Momente und Parameter der Beta-Verteilung über alle Teilflächen ($n = 65\,645$). Hier ist zu beachten, dass E und Sd in %FK angegeben sind, während sich α und β auf das Intervall $[0,1]$ der Betaverteilung beziehen.

Saison	Moment/ Parameter	Min.	Q ₀₅	Median	Q ₉₅	Max.	Mittelw.	Sd
Sommer	E [%FK]	17.2	27.4	44.7	64.6	90.0	45.2	11.6
	Sd [%FK]	5.2	13.5	18.3	23.4	29.2	18.3	3.1
	α (shape 1)	0.55	1.22	2.56	6.42	49.32	3.09	2.20
	β (shape 2)	0.54	1.59	3.29	6.35	43.65	3.58	1.78
Winter	E [%FK]	42.0	63.2	84.1	94.0	99.2	45.2	11.6
	Sd [%FK]	1.3	8.4	14.7	21.3	25.4	18.3	3.0
	α (shape 1)	1.26	2.15	4.11	8.54	48.66	4.62	2.30
	β (shape 2)	0.06	0.26	0.82	2.10	8.28	0.96	0.61

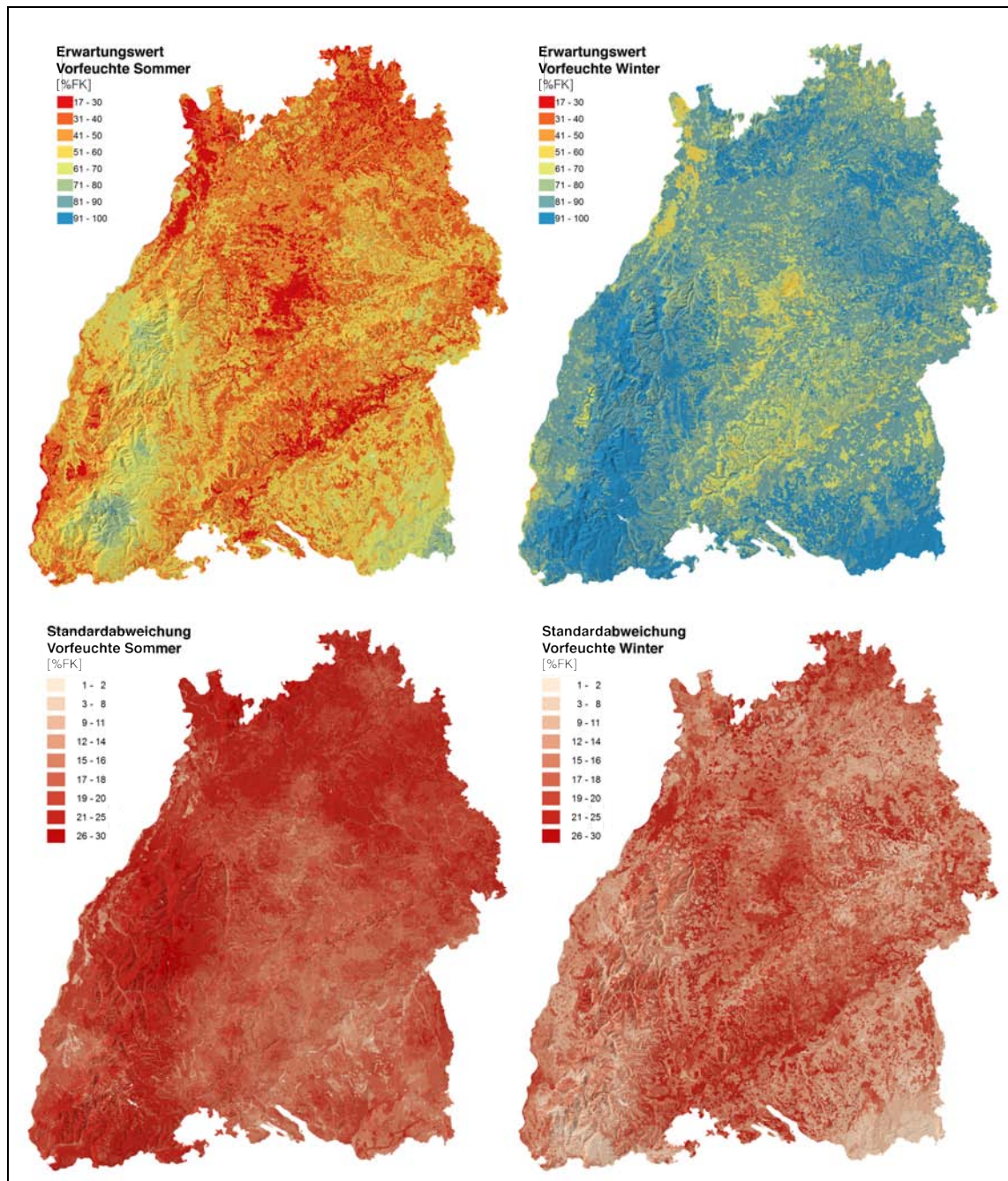


Abbildung 4.11: Räumliche Verteilung von Erwartungswert E und Standardabweichung Sd der Beta-Verteilung für die Vorfeuchtebedingungen θ_v im Sommer (links) und Winter (rechts).

4.2.3 Univariate Wiederkehrintervalle

Um die räumlichen Muster der später (in Kapitel 4.3) erhaltenen kombinierten Wahrscheinlichkeiten bzw. Wiederkehrintervalle Interpretieren zu können, sollen hier zum Vergleich die univariaten Wiederkehrintervalle T' für gegebene Werte von Niederschlag und Vorfeuchte dargestellt werden. Um die Jährlichkeiten nicht zu weit außerhalb des Mess- bzw. Modellierungszeitraumes extrapolieren zu müssen, wurden hier für Niederschlag und Vorfeuchte jeweils zwei Werte pro Saison nahe des Erwartungswertes (vgl. Tabelle 4.1 und 4.2) ausgewählt, für welche auch die berechneten kombinierten Wiederkehrintervalle – zumindest überwiegend – in einem sinnvollen und vertretbarem Rahmen $T' < 1000a$ bleiben, was mit rund 31×31 Jahren noch gut innerhalb des univariaten Beobachtungs- bzw. Modellierungszeitraumes beider Variablen liegt.

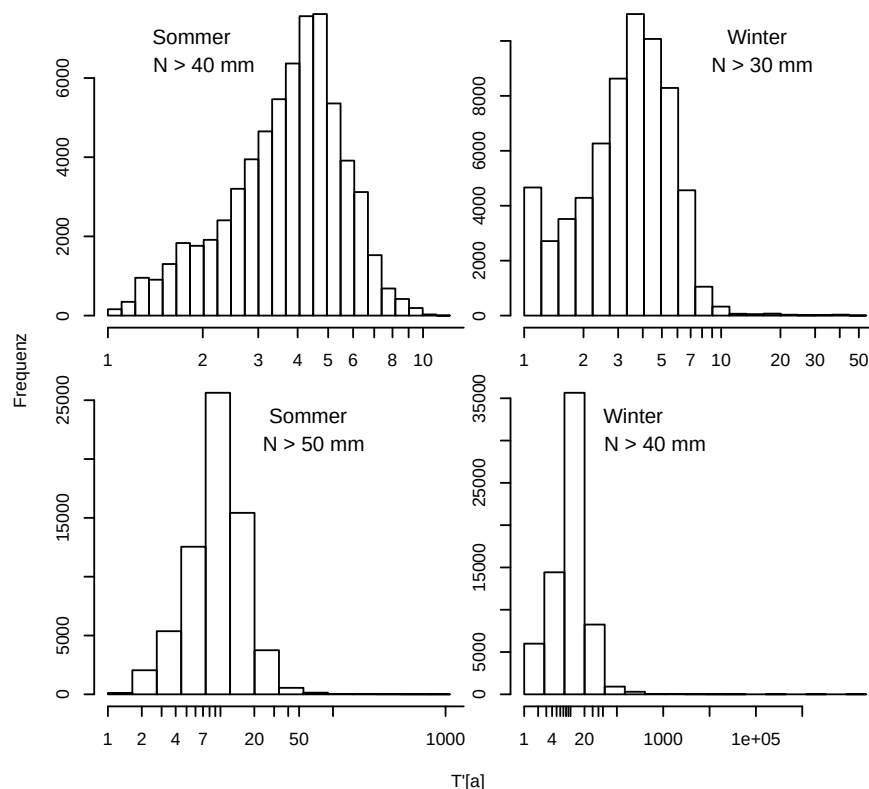


Abbildung 4.12: Häufigkeitsverteilung der Wiederkehrintervalle T' [a] über die Gesamtfläche für $N > 40$ und 50 mm/24h (Sommer) und $N > 30$ und 50 mm/24h (Winter).

Für den Niederschlag wurden hier die Werte 40 und 50 mm/24h im Sommer sowie 30 und 40 mm/24h im Winter gewählt. Für die Bodenfeuchte entsprechend 60 und 70%FK im Sommer sowie 80 und 90% FK im Winter. Die Häufigkeitsverteilungen der daraus resultierenden Wiederkehrintervalle sind in Abbildung 4.12 für den Niederschlag und in Abbildung 4.13 für die Vorfeuchte dargestellt. Die räumlichen Darstellungen jeweils in Abbildung 4.14 für den Niederschlag und Abbildung 4.15 für die Vorfeuchte. Besonders auffällig ist hier das Muster der einzelnen Niederschlagsstationen, welche sich deutlich durch nahezu kreisrunde Cluster verringerter Jährlichkeiten durchzeichnet. Dies ist durch das Interpolationsverfahren mittels inverser Distanzgewichtung zurückzuführen, da hier die mittleren gemessenen Niederschlagshöhen mit zunehmender Distanz zu der jeweiligen Station rasch abnehmen und damit letztlich zu stark erhöhten Jährlichkeiten zwischen den Stationen führen.

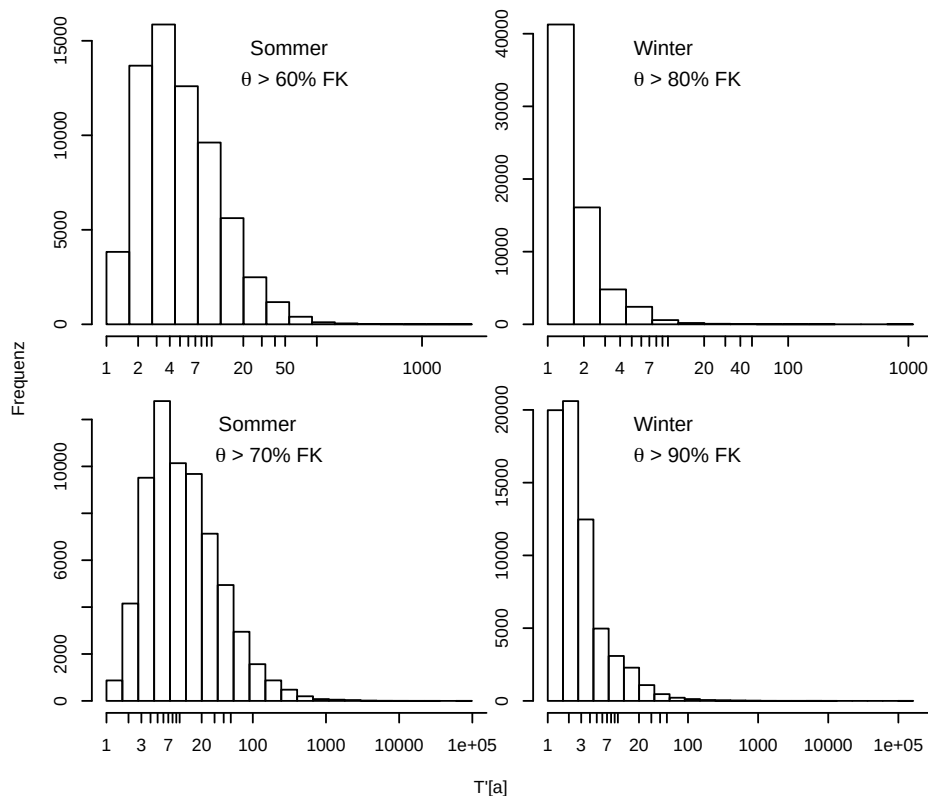


Abbildung 4.13: Häufigkeitsverteilung der Wiederkehrintervalle T' [a] über die Gesamtfläche für $\theta_v > 60$ und 70 %FK (Sommer) und $\theta_v > 80$ und 90 %FK (Winter).

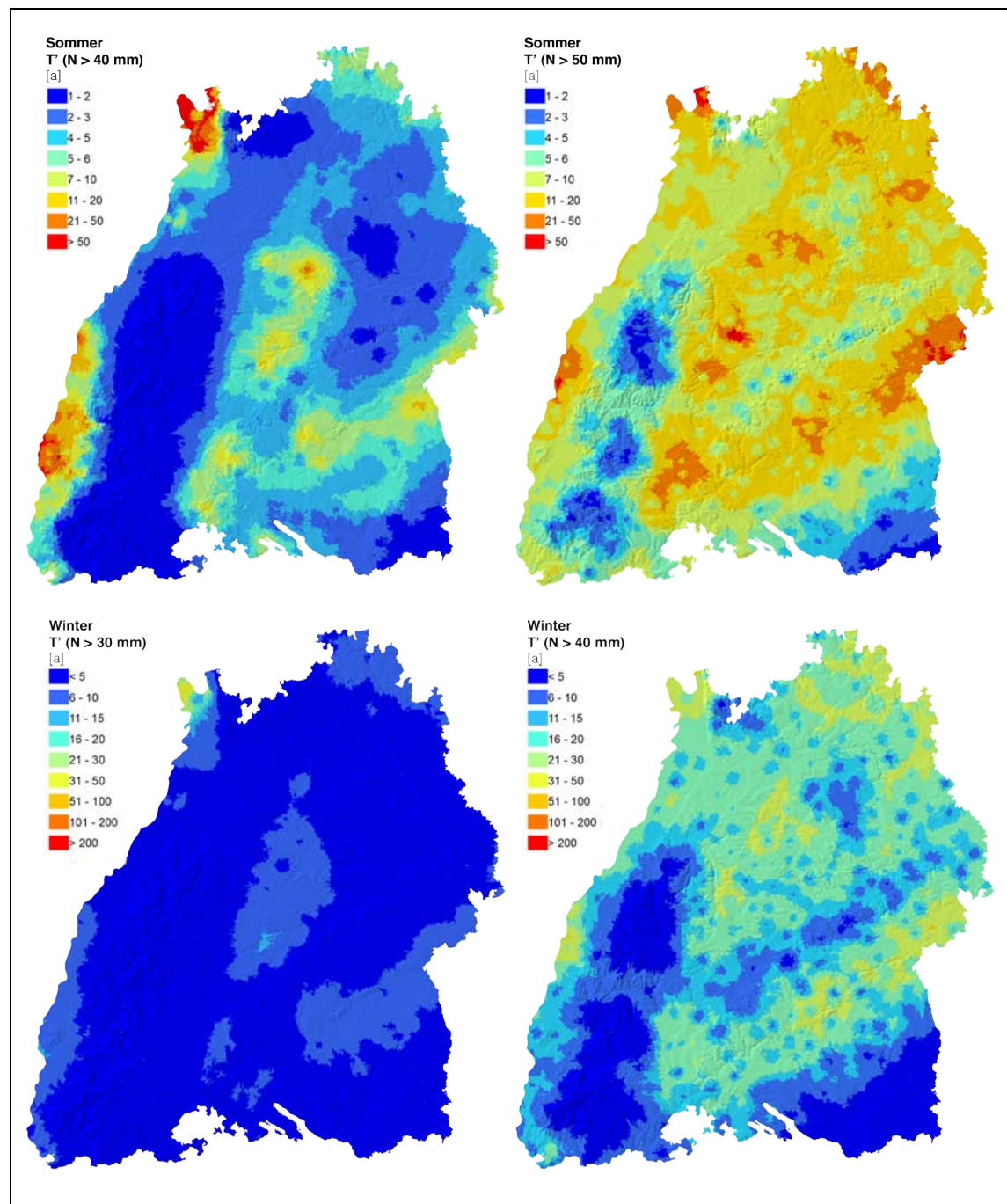


Abbildung 4.14: Jährlichkeiten des Niederschlags für $N > 40$ und 50 mm/24h im Sommer (links) und $N > 30$ und 40 mm/24h im Winter (rechts).

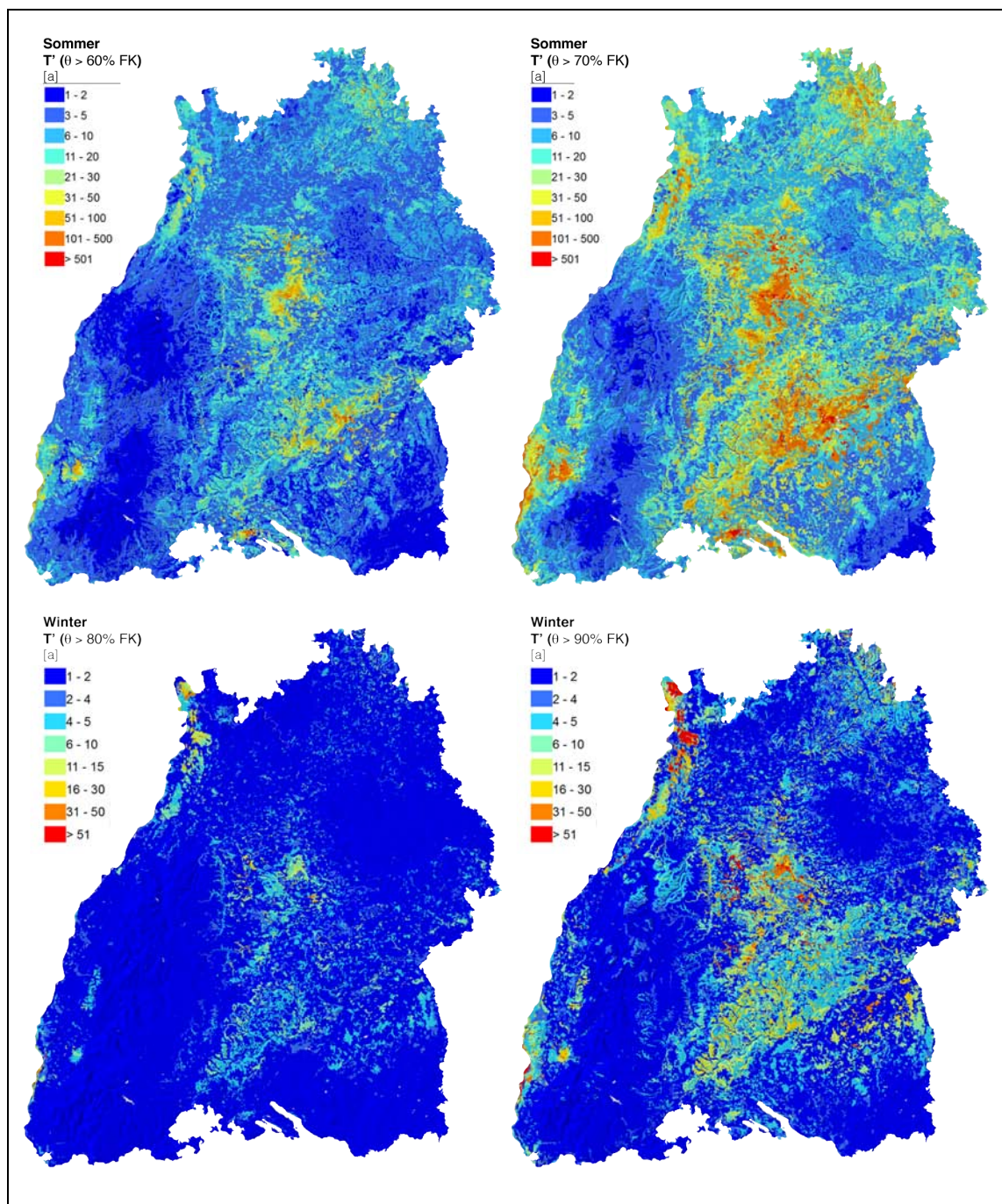


Abbildung 4.15: Wiederkehrintervalle der Vorfeuchtebedingungen $\theta_v > 60$ und 70 %FK im Sommer (links) und $\theta_v > 80$ und 90 %FK im Winter (rechts).

4.3 Joint probabilities

Wie aus Abbildung 1.2 b bereits hervorgeht, existieren bei einer bivariaten Wahrscheinlichkeitsverteilung theoretisch unendlich viele Kombinationen der zwei Variablen für eine gegebene gemeinsame Über- bzw. Unterschreitungswahrscheinlichkeit (die Iso-Linien mit gleicher Wahrscheinlichkeit können auch als *Pareto-Fronten* bezeichnet werden). Von daher können für die räumliche Darstellung in Kartenform nur die Wiederkehrintervalle für einen gegebene Kombination von N und θ_v - Werten verwendet werden. Um einen Eindruck von den möglichen Formen der erhaltenen bivariaten Verteilungen zu bekommen, sind in Abbildung 4.16 zunächst die empirischen, gemeinsamen Überschreitungswahrscheinlichkeiten $F'(x,y)$ (aus Formel 3.5.5 und 3.5.13) für die Punkte B, D und E in perspektivischer Darstellung, in Abbildung 4.17 die entsprechenden theoretischen Verteilungen mit BV und GEV als Randverteilungen dargestellt. Die daraus berechneten Wiederkehrintervalle (siehe Formel 3.5.15) sind dann in Abbildung 4.18 als Konturplot dargestellt, um eine bessere Auflösung bei den geringen bis mittleren Jährlichkeiten zu erhalten.

Die Schwierigkeiten bei der Darstellung der *joint probabilities* besteht in der enormen Spannweite der Wiederkehrintervalle bereits für Kombinationen relativ geringer Niederschlags- und Vorfeuchtwerte. Deshalb wurde bei der Erstellung der Karten die Intervallbildung weitgehend nach Quantilen durchgeführt und extrem hohe Jährlichkeiten in oftmals sehr weiten Klassengrenzen zusammengefasst, um auch eine Vergleichbarkeit der unterschiedlichen Wertekombinationen und Jahreszeiten untereinander zu ermöglichen. Um die gesamte Spannweite aller möglichen Wiederkehrintervalle und das Verhalten ihrer Häufigkeitsverteilungen für verschiedene N und θ_v - Werte betrachten zu können, sind in Abbildung 4.19 die interpolierten, empirischen Dichtefunktionen (*density traces*) der Wiederkehrintervalle für die in Kapitel 4.2.3 festgelegten $N - \theta_v$ Kombinationen über alle Teilflächen dargestellt.

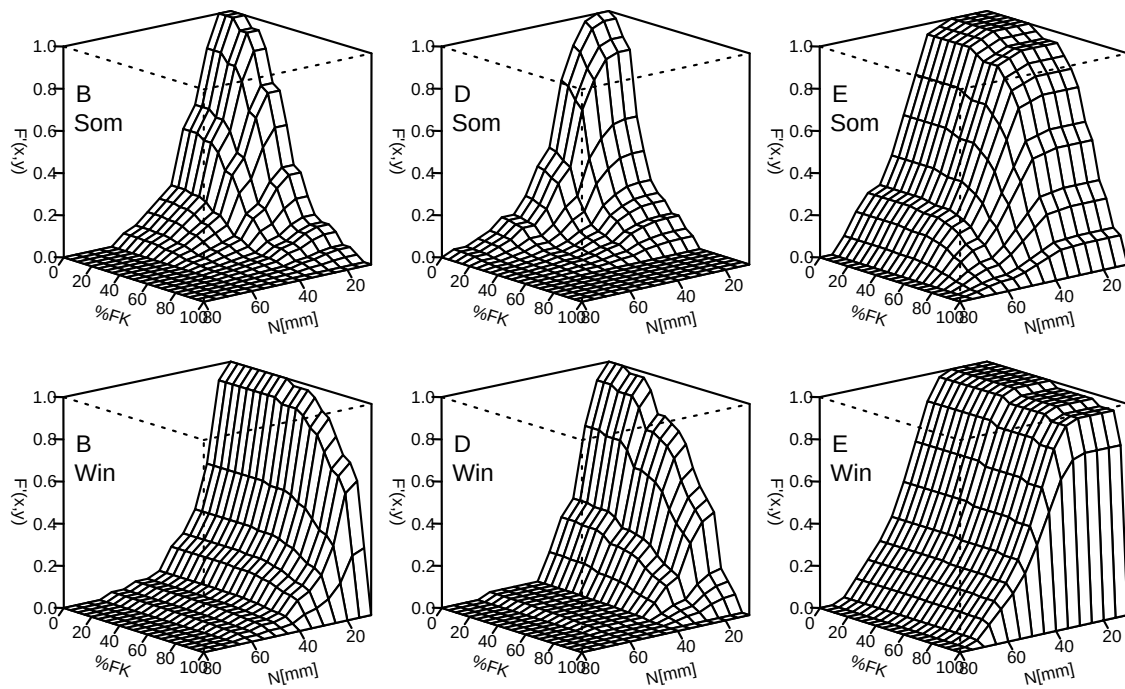


Abbildung 4.16: Empirische Überschreitungswahrscheinlichkeiten $F'(x,y)$ für die Punkte B, D und E.

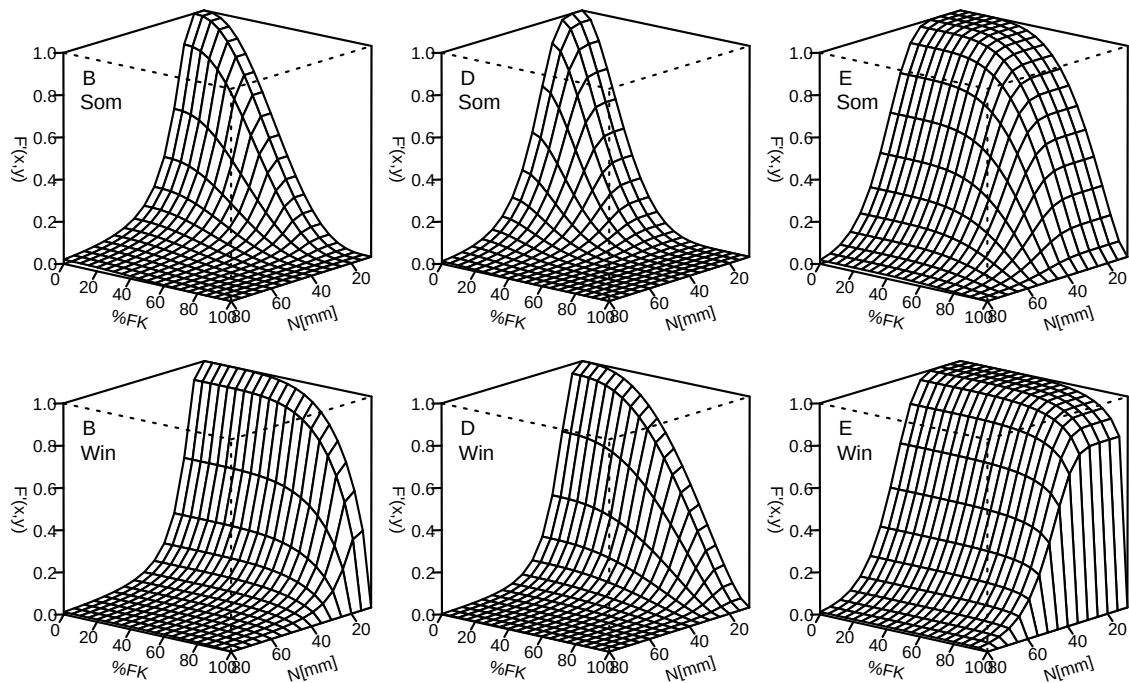


Abbildung 4.17: Theoretische Überschreitungswahrscheinlichkeiten $F(x,y)$ für die Punkte B, D und E mit BV und GEV als Randverteilungen.

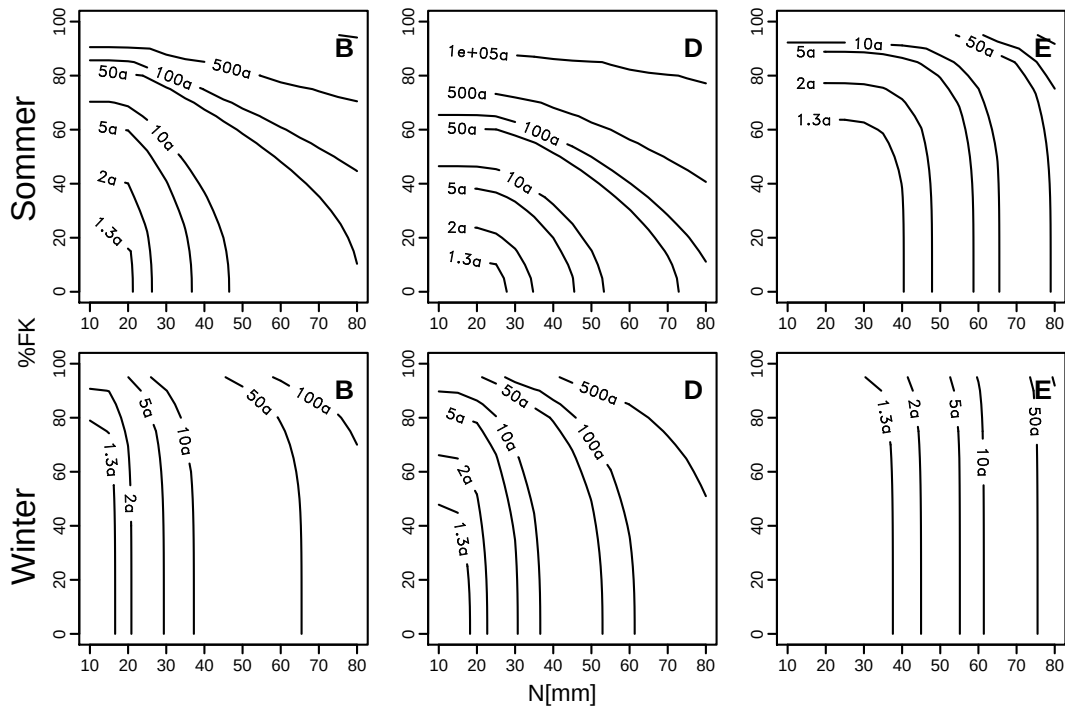


Abbildung 4.18: Gemeinsame Wiederkehrintervalle $T'(X > x, Y > y)$ aus den Randverteilungen BV für θ_v und GEV für N für die Punkte B, D und E.

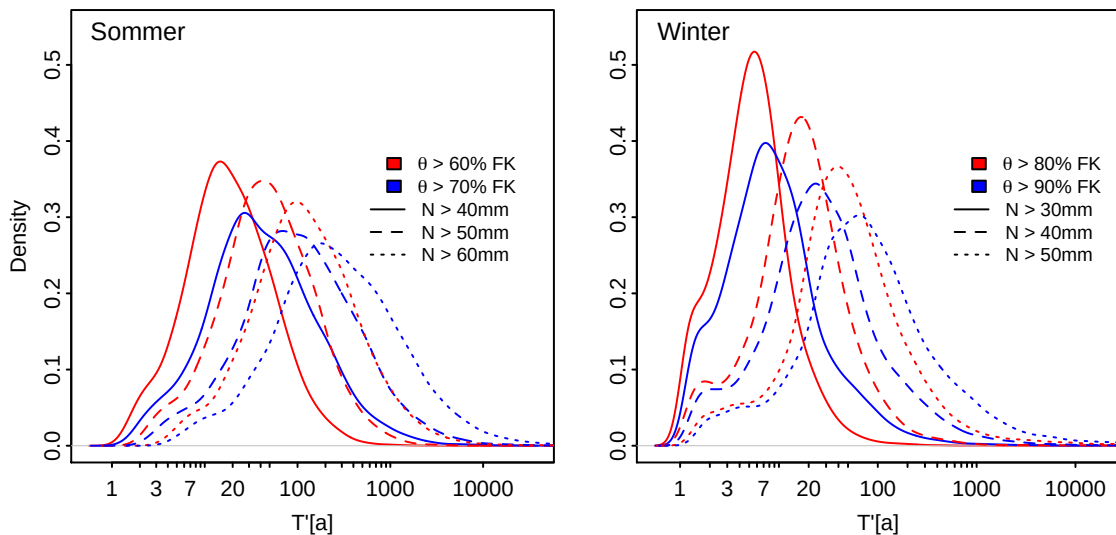


Abbildung 4.19: Empirische, interpolierte Dichtefunktionen (*density traces*) der gemeinsamen Jährlichkeiten für jeweils 6 Kombinationen aus θ_v und N über alle Teilflächen.

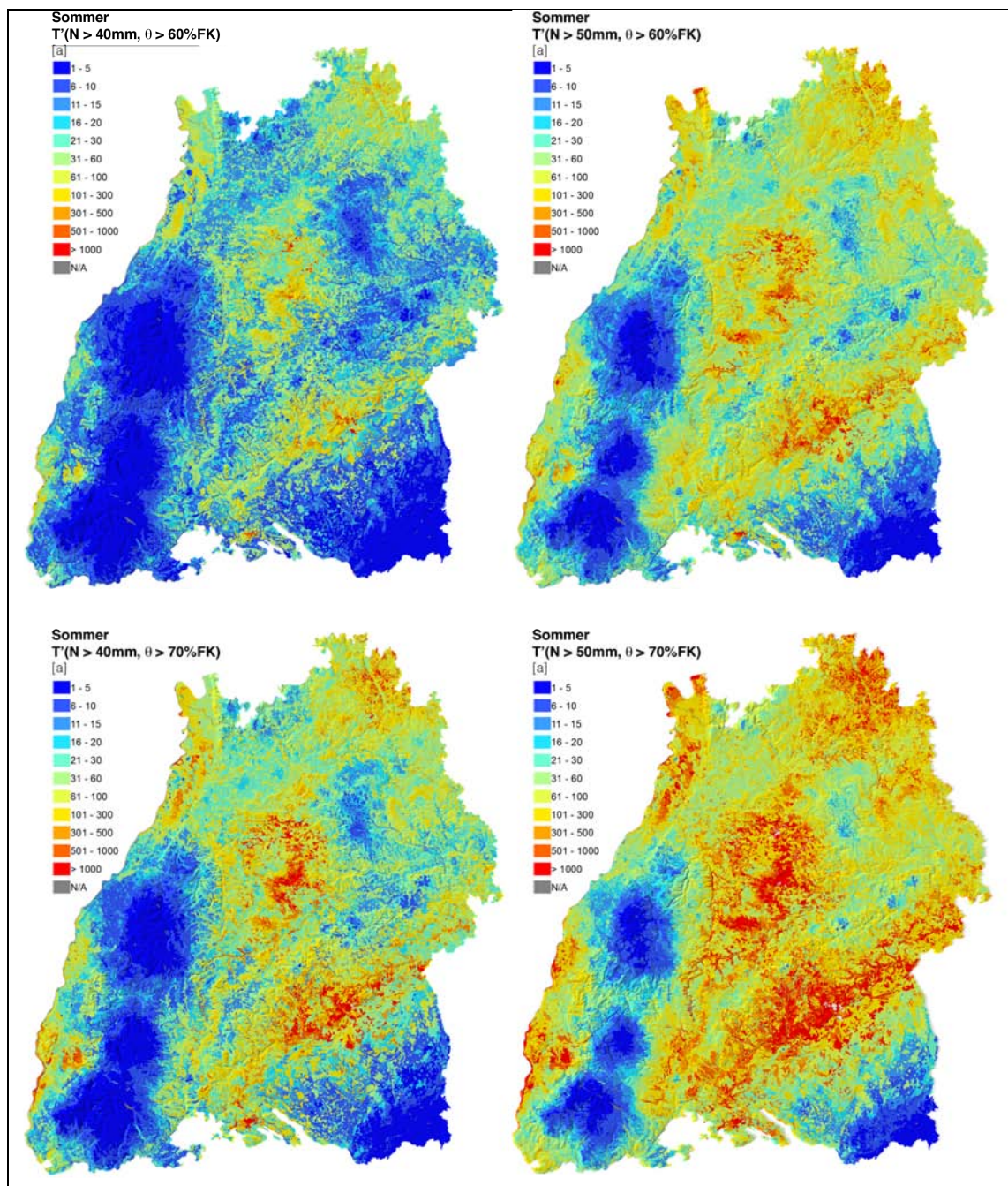


Abbildung 4.20: Gemeinsame Wiederkehrintervalle von N und θ_v jeweils für $N > 40 / > 50 \text{ mm}$ sowie $\theta_v > 60 / > 70\%\text{FK}$ im Sommer.

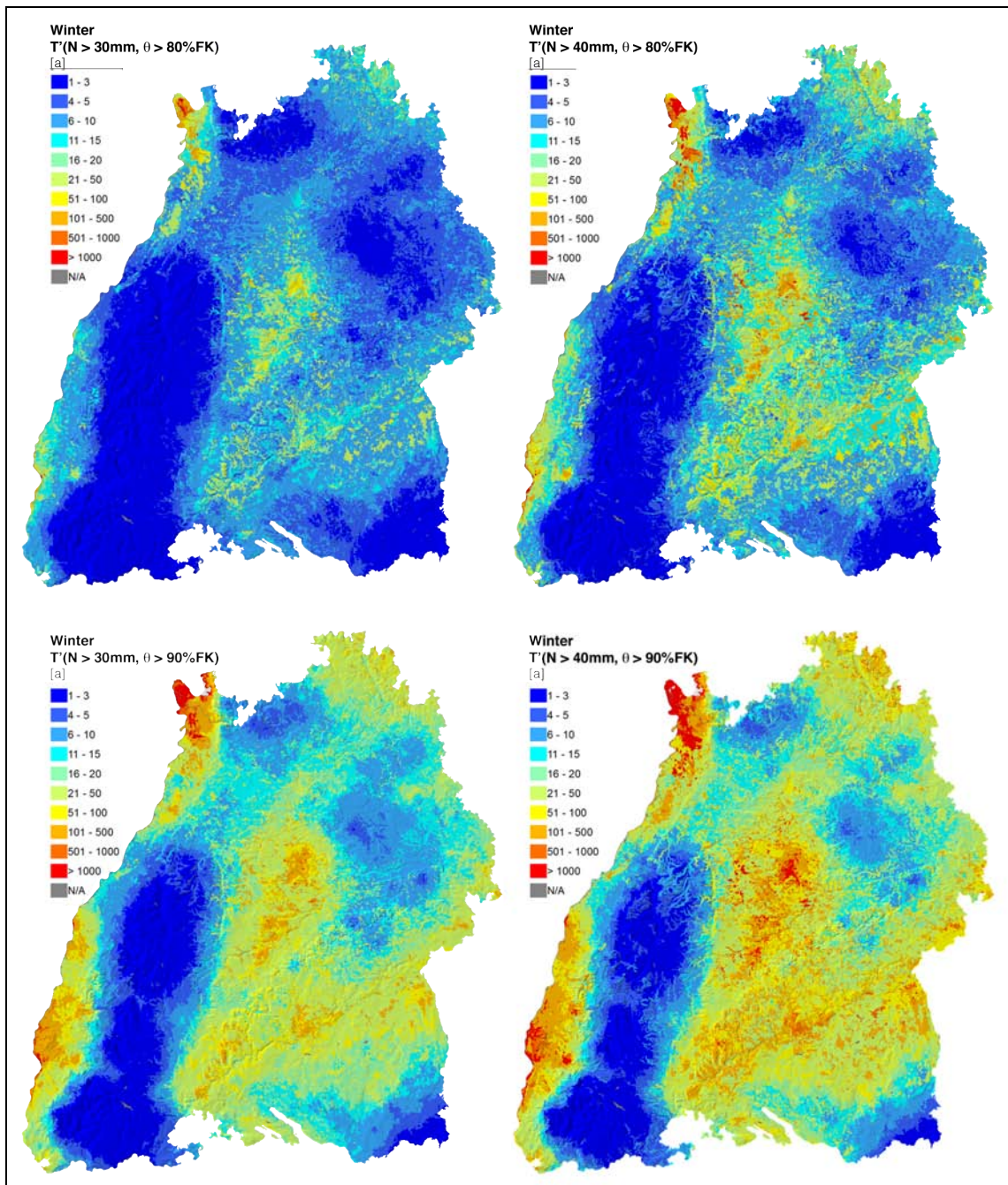


Abbildung 4.21: Gemeinsame Wiederkehrintervalle von N und θ_v jeweils für $N > 30 / > 40$ mm sowie $\theta_v > 80 / > 90\%\text{FK}$ im Winter.

Die Abbildungen 4.20 und 4.21 zeigen schließlich die räumliche Darstellung der gemeinsamen Wiederkehrintervalle für jeweils 4 verschiedene Kombinationen (siehe Legende) aus N und θ_v im Sommer und Winter. Die Kombinationen aus N und θ_v wurden hier so gewählt, dass für jede Variable jeweils ein identisches

Wertepaar existiert, um die Änderungen der resultierenden räumlichen Muster den einzelnen Variablen zuordnen zu können.

5 Diskussion

Zunächst fällt allgemein auf, dass die für die Niederschlagsdaten eingesetzte Interpolationstechnik nicht geeignet ist, um damit eine anschließende Extremwertanalyse durchzuführen, da durch die Anwendung der inversen Distanzgewichte die Niederschlagshöhe mit zunehmendem Abstand zur Station konstant kleiner wird, was zu einer systematischen Unterschätzung der Niederschlagssummen und folglich zu einer Überschätzung der Jährlichkeiten für die Bereiche zwischen den Stationen führt. Dieser Effekt wird umso stärker, je häufiger in der entsprechenden Region konvektive Ereignisse mit geringer räumlicher Ausdehnung und hohen Niederschlagssummen stattfinden (vgl. Abbildung 4.14), also vorwiegend im Sommer in den mittleren bis tiefer gelegenen Regionen Baden-Württembergs. Nur in Regionen, in denen häufig lang anhaltende und weiträumige Ereignisse mit frontalem Charakter stattfinden, beispielsweise westlich des Schwarzwaldes und im Alpenvorland, wo diese durch häufig auftretende zyklonale West und Nordwest-Wetterlagen sowie orographische Effekte zustande kommen, zeigt sich ein etwas homogeneres Muster der Niederschlagsextremwerte trotz verhältnismäßig geringer Stationsdichte (STEINBRICH et al., 2005). Von daher wäre es hier grundsätzlich sinnvoller, die Extremwertanalyse zuerst für die einzelnen Stationsdaten durchzuführen und die Ergebnisse anschließend mittels eines erwartungstreuen Interpolationsverfahren zu regionalisieren.

5.1 Korrelationen zwischen N und θ

Die für die Niederschlags- und Vorfeuchtestichproben durchgeführte Korrelationsanalyse (siehe Kapitel 3.3 und 4.1, Abbildung 4.2) lieferte keine Eindeutig zu interpretierenden Ergebnisse. Generell zeigt sich hier besonders im Winter eine Konzentration der Flächen mit signifikanten positiven Korrelationen im Nordosten des Landes mit nur wenigen, vereinzelt Flächen mit negativen Korrelationskoeffizienten. Im Sommer sind die Muster der (ebenfalls vorwiegend posi-

tiven) Korrelationen insgesamt wesentlich heterogener. Für beide Jahreszeiten lassen sich hier weder Zusammenhänge mit regionalen Gebietseigenschaften wie Landnutzung, verschiedener Bodenparameter wie Feldkapazität, Wasserleitfähigkeit und Gründigkeit, noch durch klimatische Parameter wie potentieller und aktueller Verdunstung, mittlerer Jahrestemperatur oder Blitzhäufigkeit als grobes Maß für die Frequenz konvektiver Ereignisse eindeutig feststellen.

Bei der Interpretation dieser Ergebnisse muss jedoch berücksichtigt werden, dass je nach gewähltem Signifikanzniveau immer eine gewisse Irrtumswahrscheinlichkeit gegeben ist. Wird der Korrelationstest nach Kendall, wie es hier der Fall ist, für viele einzelne Stichproben durchgeführt, welche zumindest großräumig der selben Grundgesamtheit zugeordnet werden können, muss bei einem Signifikanzniveau von beispielsweise 5% eben auch mit einem Anteil zufälliger Korrelationen in der selben Größenordnung von etwa 5% gerechnet werden (HAAN, 1982). Dieser wird mit 14.6% im Sommer und 10.1% im Winter jedoch deutlich überschritten. Unter dieser Annahme sollten die Bereiche mit signifikanten Korrelationen (in Abbildung 4.2 durch schwarze Schraffur gekennzeichnet) allerdings auch über die Fläche mehr oder weniger zufällig und gleichverteilt liegen, was hier scheinbar nicht der Fall ist. Betrachtet man diese Flächen allerdings im direkten Vergleich mit den durch die Interpolation entstandenen Clustern der Niederschlagshöhen, wird deutlich, dass auch hier jeweils Bereiche gleicher Boden- und Landnutzungseinheiten rund um die Stationen betroffen sind, und somit das identische Zeitliche Verhalten des Niederschlags ausgehend von den jeweiligen Stationen sowie die modellbedingte, identische Entwicklung der Bodensättigung innerhalb dieser Einheiten zu dieser sichtbaren Clusterbildung führt.

Der in Abbildung 4.3a gezeigte Vergleich zwischen der unter der Nullhypothese des Tests zu erwartenden Verteilung und der tatsächlichen, räumlichen Verteilung von τ zeigt, dass hier immer noch von einem großen Anteil "zufällig" entstandener Korrelationen ausgegangen werden kann. Das Verhalten der Verteilung unter der Bedingung, dass am Vortag des Ereignisses keine relevante Niederschlagsmenge (> 5 mm) gefallen sein darf (Abbildung 4.3b), scheint die

anfängliche Vermutung zu bestätigen, dass auch das langfristige Autokorrelationsverhalten des Niederschlags teilweise zur Erklärung dieser scheinbaren Abhängigkeiten herangezogen werden kann. Dies erscheint auch logisch, da eine starke Autokorrelation des Niederschlags für vorwiegend lang anhaltende Ereignisse spricht. Die Wahrscheinlichkeit, dass es am Tag vor einem ausgewählten Extremereignis bereits geregnet hat, ist hier also besonders groß und damit folglich auch die Vorfeuchte an diesem Tag. Ein direkter Vergleich der Abbildungen 4.2 und 4.5 zeigt hier einen deutlichen räumlichen Zusammenhang zwischen den beiden Koeffizienten τ und ρ . In einer weiteren Überprüfung mittels Rangkorrelation ergab sich hier ein stark signifikanter ($\alpha < 0.01$) Zusammenhang für beide Jahreszeiten (siehe Abbildung 4.4).

Eine weitere, denkbare Erklärung für die positiven Korrelationen zwischen N und θ_v (sprich: Eine erhöhte Vorfeuchte hat zumindest tendenziell auch eine erhöhte maximale Niederschlagshöhe zur Folge) wären u. a. von RASMUSSEN (1967), ELTAHIR (1998), MENDÉZ-BAROSSO et al. (2009) und VAN DER ENT (2010) für große Teile Nord(ost)-Amerikas beschriebene Rückkopplungs- oder Recyclingmechanismen, welche durch erhöhte Evapotranspiration von feuchten Böden und dem daraus resultierenden erhöhten Wassergehalt in den unteren Troposphärenschichten auch zu höheren Niederschlagssummen führen können. Diese Mechanismen können jedoch mit dieser Datengrundlage und den hier eingesetzten Stichproben aus jeweils einem Jahresmaximum der 24-h Niederschlagssumme und der jeweils korrespondierenden Vorfeuchte kaum verifiziert werden. Hier wäre eine präzisere Analyse von – je nach Verfügbarkeit – kürzeren, jedoch zeitlich feiner aufgelösten Niederschlagszeitreihen sinnvoller, um auch zwischen frontalen und konvektiven Ereignissen differenzieren zu können, da diese Mechanismen offensichtlich nur für letztere von Bedeutung sind.

5.2 Randverteilungen

Die Anpassung der GEV an die Niederschlagsserien lieferte, wie in Abbildung 4.7 anhand der p-Werte des KS-Tests abzulesen ist, sehr gute Ergebnisse. Auf-

grund der großen Zahl an Teilflächen, kann hier nur der p-Wert als Maß für die Güte der Anpassungen beider Verteilungen (BV und GEV) herangezogen werden.

Für die Interpretation der daraus erhaltenen Überschreitungswahrscheinlichkeiten bzw. Jährlichkeiten ergeben sich für den Niederschlag die bereits erwähnten Probleme, welche sich durch die sehr inhomogene Regionalisierung der Niederschlagsdaten ergeben. Betrachtet man hier jedoch nur die für die einzelnen Stationen ermittelten Jährlichkeiten, ergibt sich eine gute, sowohl räumliche als auch quantitative Übereinstimmung der hier ermittelten 1- und 100-Jährlichen Niederschlagshöhen mit den in WABOA (2007) dargestellten Höhen der selben Jährlichkeiten. Die hier ermittelten 24-h Niederschlagssummen fallen dabei insgesamt etwas geringer aus, was sich in erster Linie durch die getrennte Betrachtung von Sommer- und Winterhalbjahr erklären lässt.

Die Anpassung der Betaverteilung verlief insgesamt weniger gut, aber dennoch zufriedenstellend für den allergrößten Teil der Landesfläche. Besonders im Winterhalbjahr waren hier, wie in Kapitel 4.2.2 bereits erwähnt, vor allem die Regionen Hochschwarzwald, südlicher Odenwald sowie das westliche Voralpenland von schwächeren Anpassungen mit abgelehnter Nullhypothese betroffen. Die in diesen Fällen vorliegenden Verteilungen entsprechen im Großen und Ganzen der in Abbildung 4.10 für den Punkt E im Winter gezeigten Form mit sehr hohem Erwartungswert (meist $> 90\%$) und oftmals vielen gleichen Werten (*ties*), was sich auch in einer stark verringerten Varianz bzw. Standardabweichung der Vorfeuchte in diesen Bereichen widerspiegelt.

Die räumliche Verteilung der Vorfeuchte - Erwartungswerte entspricht großflächig betrachtet dem, was unter Berücksichtigung des langzeitlichen Verhaltens der verschiedenen Bodenspeicher zu erwarten ist. Die entscheidenden Größen sind hier die Dauer seit dem letzten größeren Niederschlag, der Sättigungszustand der Bodenspeicher zu Beginn der darauffolgenden niederschlagsfreien Zeit und die Geschwindigkeit, mit der sich der Bodenspeicher innerhalb der durchwurzelter Zone wieder entleert. Für diese Entleerung können im Sommer

und insgesamt während der Vegetationsperiode vorwiegend Evaporation und Transpiration als entscheidende Kontrollglieder angenommen werden. Im Winter hingegen, wenn die Pflanzentranspiration auf ein Minimum sinkt und dem Boden auch durch eine temperaturbedingte, geringe potentielle Verdunstung kaum Wasser entzogen wird, spielen die flächenspezifischen Abflussbildungsmechanismen (Oberflächen- und Zwischenabfluss und Tiefenperkolation) die vorherrschende Rolle.

Diese Effekte lassen sich für einzelne, unscharf begrenzte Teilbereiche auch mit dem in WABOA (2007) vorliegenden Kartematerial bestätigen. Hier fallen im Sommer zunächst Flächen mit generell hohem Niederschlagsinput und geringer potentieller Verdunstung in den hochgelegenen Gebieten des Schwarzwaldes sowie des Alpenvorlandes, mit generell hoher Vorfeuchte ($< 80\%$) auf. Für die restliche Landesfläche ergibt sich ein wesentlich diffuseres Bild, wobei sich auch hier Bereiche geringer Vorfeuchte und hoher potentieller Verdunstung tendenziell abdecken. Insgesamt beeinflussen jedoch eine ganze Reihe weiterer unbekannter oder nur sehr grob klassifizierter Größen wie Bodenmächtigkeit, Durchlässigkeit und Speicherverhalten oder Vegetationszustand die erhaltenen räumlichen Muster, wobei jeder einzelne dieser Parameter eine untergeordnete Rolle spielt und erst das Produkt aller Einflussfaktoren die volle Variabilität zu erklären vermag. Dieses Phänomen wird oftmals auch als Emergenz bezeichnet. Besonders auffällig zeigen sich jedoch sowohl größere Siedlungsräume wie die Gegenden um Stuttgart, Karlsruhe oder Freiburg, welche durch hohe Anteile versiegelter Flächen und erhöhter potentieller Verdunstung über das gesamte Jahr stark verringerte Vorfeuchtebedingungen aufweisen, sowie gut durchlässige Böden deren Wasserhaushaltsregime vorwiegend durch vertikale Sickerwasserbewegungen geprägt ist, wie die der Gäulandschaften, den Hochflächen der Schwäbischen- und Fränkischen Alb sowie vereinzelt Teilen des nördlichen Alpenvorlands. Hier ist offensichtlich besonders im Winter die schnelle Drainage dieser Böden für die verringerten Vorfeuchten verantwortlich, was auch mit einem stark verringerten Verhältnis von aktueller zu potentieller Verdunstung aufgrund des verringerten Wasserangebots der oberen Bodenschichten in Ein-

klang steht. Eine umfassende Untersuchung aller möglicher Einflussfaktoren, beispielsweise durch eine Hauptkomponenten- oder Faktorenanalyse, würde im Idealfall letztlich nur den Input des eingesetzten Grundwassermodells reproduzieren. Von daher ist es auch nicht Überraschend, das die modellierten Vorfeuchtebedingungen großräumig betrachtet in erster Linie mit den mittleren Jahresniederschlagshöhen korrelieren und die zusätzlichen Modellparameter und Inputdaten für die sehr komplexen, kleinräumigen Strukturen verantwortlich sind.

5.3 Joint probabilities

Die gemeinsamen Wiederkehrintervalle der verschiedenen Niederschlag - Vorfeuchte - Kombinationen können grundsätzlich nicht mit Wiederkehrintervallen von Abflussereignissen verglichen werden, da für stetige Variablen theoretisch unendlich viele verschiedene Kombinationen gleicher Wahrscheinlichkeit zusammen mit anderen Einflussgrößen zu ähnlichen Abflussereignissen führen können.

Die in den Abbildungen 4.20 und 4.21 gezeigten gemeinsamen Wiederkehrintervalle sind räumlich nur schwer repräsentativ darstellbar, da hier immer zwei der drei Variablen N , θ_v und F' bzw. T' vorgegeben werden müssen, um die dritte zu erhalten. So könnte anstelle der gemeinsamen Jährlichkeiten T' für gegebene N und θ_v zwar beispielsweise auch jeweils ein Vorfeuchtwert dargestellt werden, welcher zusammen mit einer bestimmten Niederschlagshöhe in einer gegebenen Jährlichkeit resultiert (oder andersherum) – ohne einen direkten Vergleich mit modellierten oder gemessenen, aus den verschiedenen Kombinationen resultierenden Abflusswerten einzelner Pegeleinzugsgebiete ist hier kein nachvollziehbarer Informationsgewinn durch die einfache Multiplikation der Randverteilungen (vgl. Kapitel 3.5) erkennbar.

Da die gemeinsamen Wiederkehrintervalle letztlich nur das reziproke Produkt der univariaten Randverteilungen darstellen, lassen sich auch die erhaltenen Muster prinzipiell durch eine multiplikative Überlagerung der in den Abbil-

dungen 4.8 und 4.11 bzw. 4.14 und 4.15 dargestellten Muster der Momente und Jährlichkeiten erklären.

Hier lag offensichtlich über lange Zeit ein grundsätzlicher Denkfehler vor. Es wurde zu spät erkannt, dass die in Kapitel 4.1 aufgezeigte Unabhängigkeit zwischen N und θ_v einen bivariaten Ansatz für die betreffenden Variablen letzten Endes hinfällig werden lässt, da die Niederschlagshöhe eben keinen Einfluss auf die Form, sprich die Parameter der Vorfeuchteverteilung hat. Auch räumlich betrachtet ergeben sich für die gemeinsamen Jährlichkeiten letztlich keine neuen Muster für unterschiedliche Niederschlagshöhen und Vorfeuchtwerte (siehe Abbildungen 4.20 und 4.21), lediglich Größenordnung und Varianz der Jährlichkeiten nehmen rapide aber konstant zu, was sich besonders deutlich anhand der empirischen Dichtefunktionen in Abbildung 4.19 erkennen lässt.

Diese extrem hohen Varianzen machen eine praktische Interpretation der gemeinsamen Jährlichkeiten nahezu unmöglich. Es stellt sich die Frage, ob die Synthese der Einzelwahrscheinlichkeiten zweier unabhängiger Variablen den Informationsgehalt der daraus resultierenden, gemeinsamen Jährlichkeit als Maß für die zu erwartenden Vorfeuchtebedingungen innerhalb eines bestimmten Gebietes nicht sogar erheblich reduziert, da der bestehende Abstraktionsgrad dieses Maß sehr unanschaulich wirken lässt. Die Abhängigkeit der im Rahmen eines *joint probability* Ansatzes untersuchten Variablen ist also nicht wie zunächst vermutet, als einer von zwei prinzipiell möglichen Fällen zu verstehen, sondern vielmehr als eine grundlegende Voraussetzung, um das Verhalten der Verteilungsmomente über den gesamten Wertebereich beider Variablen interpretieren zu können.

Ein weiterer, im nachhinein als nicht zielführend erscheinender Teil der Vorgehensweise ist die Verwendung der Überschreitungswahrscheinlichkeiten F' der Vorfeuchtebedingungen. Hier ist es gerade in Hinblick auf die Initialisierung von Abflussmodellen möglicherweise sinnvoller, die Unterschreitungswahrscheinlichkeiten der fraglichen θ_v - Werte bzw. die maximalen, während einer

gewissen Zeitspanne zu erwartenden Werte, sprich die Quantile der θ_v - Verteilung einzusetzen.

Diesen Überlegungen zufolge sollte es also ausreichend sein, die Vorfeuchteverteilungen ohne Berücksichtigung der Niederschlagshöhe zu betrachten und lediglich die (binäre) Bedingung zu stellen, dass am Folgetag ein Niederschlagsereignis stattfindet. Dazu sind in Anhang 7 und 8 die empirischen 50, 80, 90 und 95% Quantile der hier verwendeten Vorfeuchteverteilung jeweils für das Sommer- und Winterhalbjahr abgebildet. Die hier erkennbaren Muster entsprechen – innerhalb der durch die Anpassung der BV bedingten Unsicherheiten – zwangsläufig denen der univariaten Wiederkehrintervalle in Abbildung 4.15, jedoch mit dem Unterschied, dass hier die maximalen Vorfeuchtwerte angegeben sind, welche mit einer gegebenen Wahrscheinlichkeit (i. e. innerhalb eines bestimmten Zeitraumes) möglicherweise erreicht, aber nicht überschritten werden. Unter dieser Betrachtung bliebe allerdings zu untersuchen, ob sich die Abschätzung dieser Quantile nicht möglicherweise durch ein darauf besser angepasstes Auswahlverfahren der Stichprobenkollektive mit vergrößertem Stichprobenumfang verbessern ließe, da in diesem Verfahren durch die Vorgabe der saisonalen Niederschlagsmaxima nur jeweils ein Wert pro Halbjahr eingegangen ist.

6 Fazit & Ausblick

Abschließend lässt sich sagen, dass diese Arbeit keineswegs als vollendet betrachtet werden kann. Sie ist vielmehr als Zwischenstand in der Entwicklung einer aussagekräftigen Methodik zur statistischen Abschätzung der Vorfeuchtebedingungen für ereignisbasierte Niederschlags-Abfluss Modelle zu verstehen.

Der hier verfolgte Ansatz über die joint probabilities der saisonalen Niederschlagsmaxima mit den korrespondierenden Vorfeuchtwerten erwies sich als nur sehr eingeschränkt nützlich, um die behandelte Fragestellung zu beantworten. Es konnte jedoch sehr eindrucksvoll gezeigt werden, welche extremen Auswirkungen die Berücksichtigung zusätzlicher Zufallsvariablen auf die Häufigkeitsverteilungen von Extremereignissen haben können. Eine (Unter-) Überschätzung der Vorfeuchtebedingungen führt letztlich auch zu einer systematischen (Unter-) Überschätzung der Abflussreaktionen auf einen bestimmten Niederschlagsinput und folglich zu einer (Über-) Unterschätzung der Jährlichkeiten entsprechender Hochwasserereignisse, wenn davon ausgegangen wird, dass die Jährlichkeit des Niederschlagsinputs für das resultierende Hochwasser erhalten bleibt.

Die enorme räumliche Heterogenität der gemeinsamen Jährlichkeiten zeigt auch, dass die Annahme von homogenen Vorbedingungen und gleichverteiltem Niederschlag bei der ereignisbasierten Modellierung von Extremereignissen je nach Skala zu erheblichen systematischen Fehlern führen kann. Um die Frage nach geeigneten Vorfeuchtebedingungen zur Modellinitialisierung beantworten zu können, ist eine Frequenzanalyse der univariaten, empirischen Vorfeuchteverteilungen offensichtlich ausreichend. Für eine Lösung mittels eines bivariaten Ansatzes wären beispielsweise die Vorfeuchte in Abhängigkeit zur der verstrichenen Zeit seit dem letzten Ereignis erfolgversprechend.

Literaturverzeichnis

- ARMBRUSTER V., 2002: Grundwasserneubildung in Baden - Württemberg. Freiburger Schriften zur Hydrologie. Institut für Hydrologie, Universität Freiburg, Freiburg i. Br. ISSN 0945-1609
- BACCHI B. und BALISTROCCHI M., 2010: Rainfall event variables: Dependence analysis by way of copulas, International Workshop: Advances in statistical hydrology, 23-25 May 2010, Taormina, Italy. http://www.risorseidriche.dica.unict.it/Sito_STAHY2010_web/pdf_papers/BacchiB_BalistrocchiM.pdf
- BACCHI B., BECCIU G. und KOTTEGODA N.T., 1994: Bivariate exponential model applied to intensities and durations of extreme rainfall. *Journal of Hydrology*, 155(1-2): 225-236.
- BALISTROCCHI M. und BACCHI B., 2011: Modelling the statistical dependence of rainfall event variables by a trivariate copula function. *Hydrol. Earth Syst. Sci. Discuss.*, 8(1): 429-481. DOI: 10.5194/hessd-8-429-2011
- BERAN M.A., 1973: Estimation of Design Floods and the Problem of Equating the Probability of Rainfall and Runoff, Symposium on the Design of Water Resources Projects with Inadequate Data. Unesco - WMO - IAHS, June 1973, Madrid, Spain, pp. 33-50.
- BERAN M.A. und SUTCLIFFE J.V., 1972: An index of flood-producing rainfall based on rainfall and soil moisture deficit. *Journal of Hydrology*, 17(3): 229-236.
- BERTHET L., ANDRÉASSIAN V., PERRIN C. und JAVELLE P., 2009: How crucial is it to account for the Antecedent Moisture Conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrol. Earth Syst. Sci. Discuss.*, 6(2): 1707-1736. DOI: 10.5194/hessd-6-1707-2009
- COLES S., 2001: An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. Springer-Verlag, London. ISBN: 1-85233-459-2
- COWPERTWAIT P.S.P. und METCALFE A.V., 2009: Use R! Introductory Time Series with R. Springer, Dordrecht Heidelberg London New York. ISBN 978-0-387-88697-8, DOI 10.1007/978-0-387-88698-5

- CRAWLEY M., 2007: The R book. Wiley, 950 pp. ISBN: 978-0-470-51024-7, <http://www.worldcat.org/isbn/9780470510247>
- DE MICHELE C.D. und SALVADORI G., 2002: On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition. *Journal of Hydrology*, 262(1-4): 245-258.
- DURANTE F. und SEMPI C., 2010: Copula theory: an introduction, *Copula Theory and Its Application*. Johannes Kepler Universität, Linz, pp. 3-31. http://fodok.jku.at/fodok/publikation.xsql?PUB_ID=34561
- DYCK S. und PESCHKE G., 1995: *Grundlagen der Hydrologie*, 3. Auflage. Verlag für Bauwesen GmbH, Berlin. ISBN: 3-345-00586-7
- EAGLESON P.S., 1972: Dynamics of flood frequency. *Water Resour. Res.*, 8(4): 878-898. DOI: 10.1029/WR008i004p00878
- ELTAHIR E.A.B., 1998: A Soil Moisture and Rainfall Feedback Mechanism: 1. Theory and observations. *Water Resour. Res.*, 34(4): 765-776. DOI: 10.1029/97wr03499
- ESRI, 2010: ArcGIS Desktop 10.4. Environmental Systems Resource Institute. http://www.esri.com/products/index.html#desktop_gis_panel
- FRANCHINI M., HELMLINGER K.R., FOUFOULA-GEORGIOU E. und TODINI E., 1996: Stochastic storm transposition coupled with rainfall - runoff modeling for estimation of exceedance probabilities of design floods. *Journal of Hydrology*, 175(1-4): 511-532.
- HAAN C.T., 1982: *Statistical Methods in Hydrology*, 3. Auflage. Iowa State University Press, Ames, 378 pp. ISBN: 0-8138-1510-X
- HELSEL D.R. und HIRSCH R.M., 2005: *Hydrologic Analysis and Interpretation, Techniques of Water-Resources Investigations of the United States Geological Survey*. U.S. Geological Survey. http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html
- JENKINSON A.F., 1955: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348): 158-171.
- JOE H., 1993: Parametric Families of Multivariate Distributions with Given Margins. *Journal of Multivariate Analysis*, 46(2): 262-282.

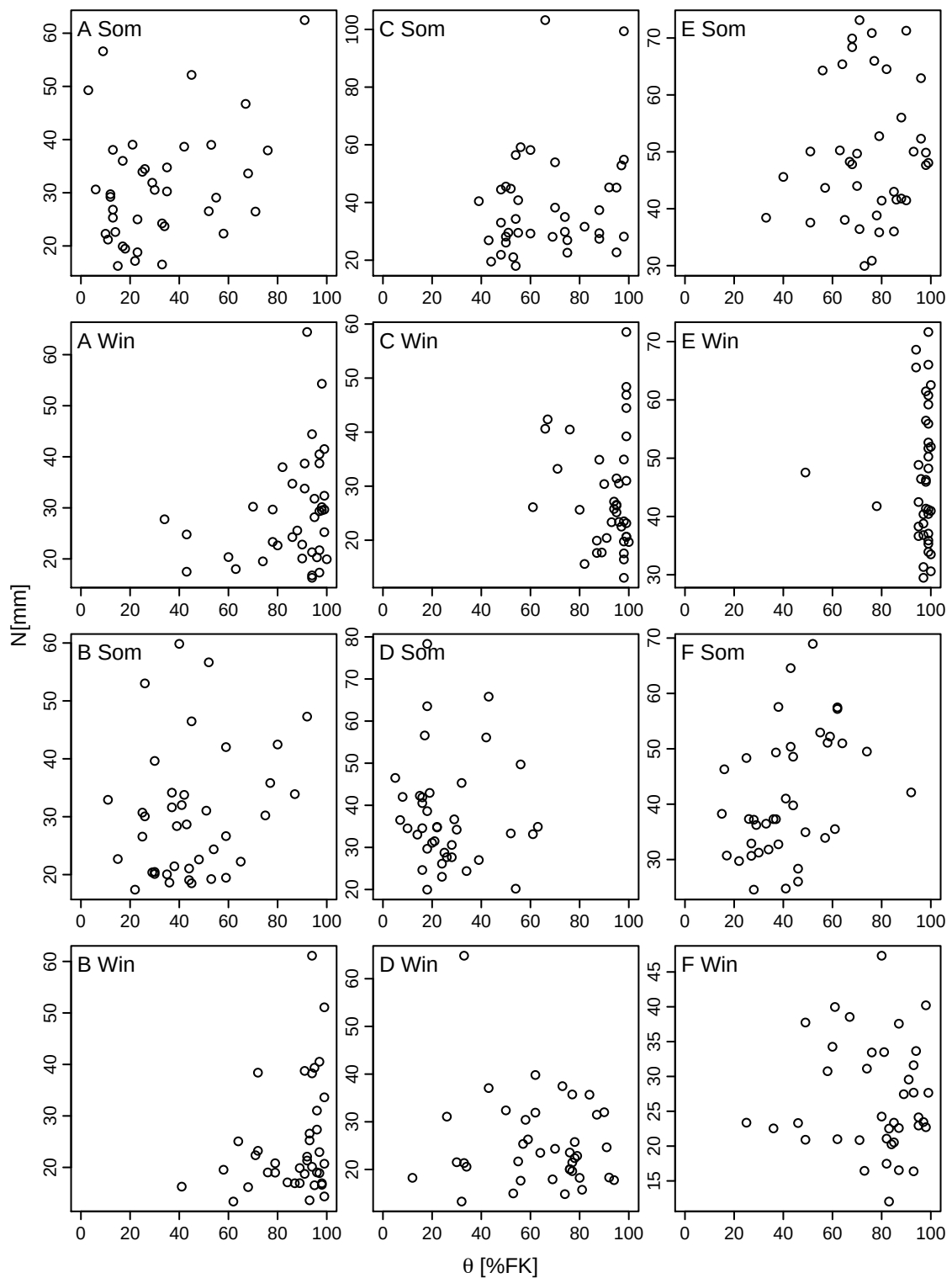
- KENDALL M.G., 1949: Rank and Product-Moment Correlation. *Biometrika*, 36(1-2): 177-193. DOI: 10.1093/biomet/36.1-2.177
- KJELDSEN T., SVENSSON C. und JONES D., 2010: A joint probability approach to flood frequency estimation using Monte Carlo simulation. In: C. Kirby (Editor), *Role of Hydrology in Managing Consequences of a Changing Global Environment*. British Hydrological Society Third International Symposium. British Hydrological Society, Newcastle, pp. 278-282. ISBN: 1903741173. <http://nora.nerc.ac.uk/11157/>
- LANG M., OUARDA T.B.M.J. und BOBÉE B., 1999: Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3-4): 103-117.
- LANGE J., 2008: Vorlesungsskript. Modul 5: Statistik & Regionalisierung - Extremwertstatistik. Wintersemester 2008/09. Institut für Hydrologie, Universität Freiburg, Freiburg i. Br.
- LEONARD M., METCALFE A. und LAMBERT M., 2008: Frequency analysis of rainfall and streamflow extremes accounting for seasonal and climatic partitions. *Journal of Hydrology*, 348(1-2): 135-147.
- LONGOBARDI A., VILLANI P., GRAYSON R.B. und WESTERN A.W., 2003: On The Relationship Between Runoff Coefficient And Catchment Initial Conditions, in: *Proceedings of MODSIM 2003*, pp. 867-872. http://www.mssanz.org.au/MODSIM03/Volume_02/A17/04_Longobardi.pdf
- LUBW, 2006: INTERREG III A Projekt MoNit - Modellierung der Grundwasserbelastung durch Nitrat im Oberrheingraben. (Kap. 1.5: Gesamtabfluss und Grundwasserneubildung aus Niederschlag). Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg, Karlsruhe. <http://www.lubw.baden-wuerttemberg.de/servlet/is/21697/>
- MANIAK U., 1997: *Hydrologie und Wasserwirtschaft: Eine Einführung für Ingenieure.*, 4. Auflage. Springer Verlag, Berlin-Heidelberg. ISBN-13: 978-3540200918
- MARSAGLIA G., TSANG W.W. und WANG J., 2003: Evaluating Kolmogorov's Distribution. *Journal of Statistical Software*, 8(18). <http://www.jstatsoft.org/v08/i18/>
- MARTINA M.L.V., TODINI E. und LIBRALON A., 2005: A Bayesian decision approach to rainfall thresholds based flood warning. *Hydrol. Earth Syst. Sci. Discuss.*, 2(6): 2663-2706. 10.5194/hessd-2-2663-2005

- MENDÉZ-BAROSSO L.A., VIVONI E.R., WATTS C.J. und RODRÍGUEZ J.C., 2009: Seasonal and interannual relations between precipitation, surface soil moisture and vegetation dynamics in the North American monsoon region. *Journal of Hydrology*, 377: 59-70. DOI: 10.1016/j.jhydrol.2009.08.009
- MOORE R.J., 2007: The PDM rainfall-runoff model. *Hydrology and Earth System Sciences*, 11(1): 483-499.
- MORHARD A., 2005: Simulation von Bodenwasserhaushalt, Verdunstung und Sickerwassermenge - methodische Aspekte am Beispiel ausgewählter Projektanwendungen. Langfassung des Vortrages beim LfU Kolloquium am 13.07.2005. GIT HydroS Consult GmbH, Freiburg.
- MORHARD A., 2009: Kurzbeschreibung des Modells GWN-BW. Erweiterungen in Version 2.0. Unveröffentlichter Beitrag zur Modelldokumentation. GIT HydroS Consult GmbH, Freiburg.
- NANDAKUMAR N., MEIN R.G. und SIRIWARDENA L., 1994: Loss modelling for flood estimation: A review, Cooperative Research Centre for Catchment Hydrology, Monash University, Clayton, Australia. <http://www.catchment.crc.org.au/archive/pubs/1000147.html>
- NELSEN R.B., 1991: Copulas and association. In: G. Dall'Aglio (Editor), *Advances in probability distributions with given marginals. Mathematics and its Applications* (67). Kluwer Academic Publishers Group, Dordrecht, The Netherlands. ISBN: 0-7923-1156-6
- R, 2010: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria. <http://www.R-project.org/>
- RAHMAN A., HOANG T.M.T., WEINMANN P.E. und LAURENSEN E.M., 1998: Joint probability approaches to design flood estimation: A review, Cooperative Research Centre for Catchment Hydrology, Monash University, Clayton, Australia. <http://www.catchment.crc.org.au/archive/pubs/1000068.html>
- RAHMAN A., WEINMANN P.E., HOANG T.M.T. und LAURENSEN E.M., 2002: Monte Carlo simulation of flood frequency curves from rainfall. *Journal of Hydrology*, 256(3-4): 196-210.
- RASMUSSEN E.M., 1967: Atmospheric water vapor transport and the water balance of north america: Part I. Characteristics of the water vapor flux field. *Monthly Weather Review*, 95(7): 403-426. DOI: 10.1175/1520-0493(1967)095<0403:AWVTAT>2.3.CO;2

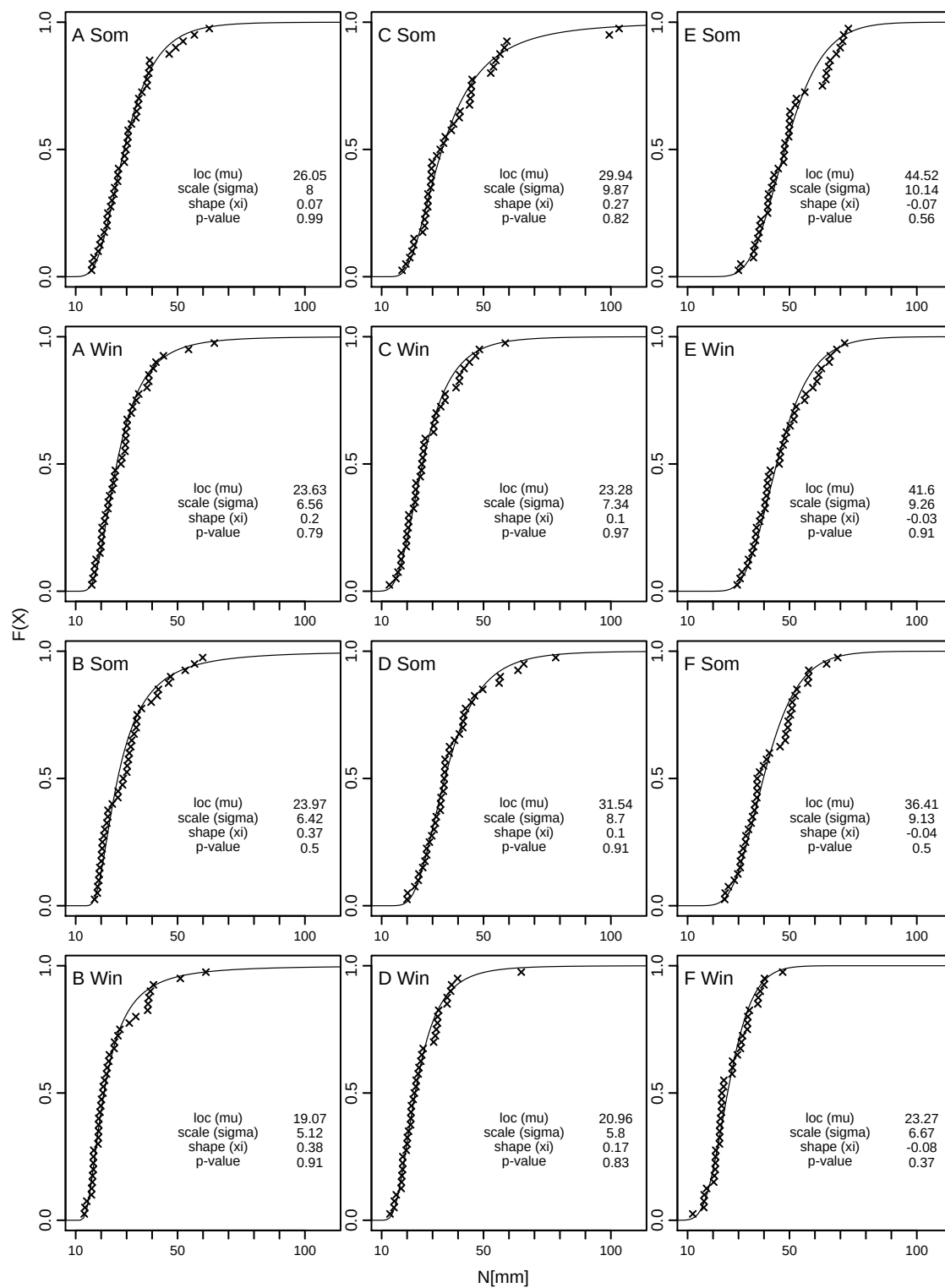
- RAVELO A.C. und DECKER W.L., 1979: The probability distribution of a soil moisture index. *Agricultural Meteorology*, 20(4): 301-312.
- RAYNAL-VILLASENOR J.A. und SALAS J.D., 1987: Multivariate extreme value analysis, *Water for the future: Hydrology in Perspective. Proceedings of the Rome Symposium, April 1987. IAHS Pub. No. 164.* http://iahs.info/redbooks/a164/iahs_164_0111.pdf
- REIMANN C., 2008: *Statistical data analysis explained: Applied environmental statistics with R.* John Wiley, Hoboken, N.J. ISBN: 9780470985816
- SCHÖNWIESE C.-D., 2006: *Praktische Statistik für Meteorologen und Geowissenschaftler*, 4. Auflage. Gebrüder Borntraeger, Berlin-Stuttgart, 302 pp.
- SCHWEIZER B. und SKLAR A., 1983: *Probabilistic metric spaces.* Elsevier New York, Holland. ISBN 0-444-00666-4
- SINGH K. und SINGH V., 1991: Derivation of bivariate probability density functions with exponential marginals. *Stochastic Hydrology and Hydraulics*, 5(1): 55-68. DOI: 10.1007/bf01544178
- STEINBRICH A., UHLENBROOK S., REICH T. und KOLOKOTRONIS V., 2005: Raumzeitliche Zusammenhänge zwischen Großwetterlagen und Starkniederschlägen in Baden-Württemberg. *Wasserwirtschaft*, 95(11): 14-19.
- TRAMBLAY Y., BOUVIER C., MARTIN C., DIDON-LESCOT J.-F., TODOROVIC D. und DOMERGUE J.-M., 2010: Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling. *Journal of Hydrology*, 387(3-4): 176-187.
- VAN DER ENT R.J., SAVENIJE H.H.G., SCHAEFLI B. und STEELE-DUNNE S.C., 2010: Origin and fate of atmospheric moisture over continents. *Water Resour. Res.*, 46(9). 10.1029/2010wr009127
- VILLARINI G., SMITH J.A., BAECK M.L., VITOLO R., STEPHENSON D.B. und KRAJEWSKI W.F., 2011: On the Frequency of Heavy Rainfall for the Midwest of the United States. *Journal of Hydrology*, In Press, Accepted Manuscript. DOI: 10.1016/j.jhydrol.2011.01.027
- WABOA, 2007: *Wasser- und Bodenatlas Baden-Württemberg.* Landesanstalt für Umweltschutz Baden-Württemberg und Ministerium für Umwelt und Verkehr Baden-Württemberg.

- WILSON D.J., WESTERN A.W. und GRAYSON R.B., 2004: Identifying and quantifying sources of variability in temporal and spatial soil moisture observations. *Water Resour. Res.*, 40(2). DOI: 10.1029/2003wr002306
- YUE S., 2001: A bivariate gamma distribution for use in multivariate flood frequency analysis. *Hydrological Processes*, 15(6): 1033-1045. DOI: 10.1002/hyp.259
- YUE S., OUARDA T.B.M.J. und BOBÉE B., 2001: A review of bivariate gamma distributions for hydrological application. *Journal of Hydrology*, 246(1-4): 1-18.
- YUE S., OUARDA T.B.M.J., BOBÉE B., LEGENDRE P. und BRUNEAU P., 1999: The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology*, 226(1-2): 88-100. DOI: 10.1016/S0022-1694(99)00168-7
- YUE S. und RASMUSSEN P., 2002: Bivariate frequency analysis: discussion of some useful concepts in hydrological application. *Hydrological Processes*, 16(14): 2881-2898.
- ZHANG L. und SINGH V.P., 2006: Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering*, 11(2): 150-164.
- ZHANG L. und SINGH V.P., 2007: Bivariate rainfall frequency distributions using archimedean copulas. *Journal of Hydrology*, 332(1-2): 93-109.

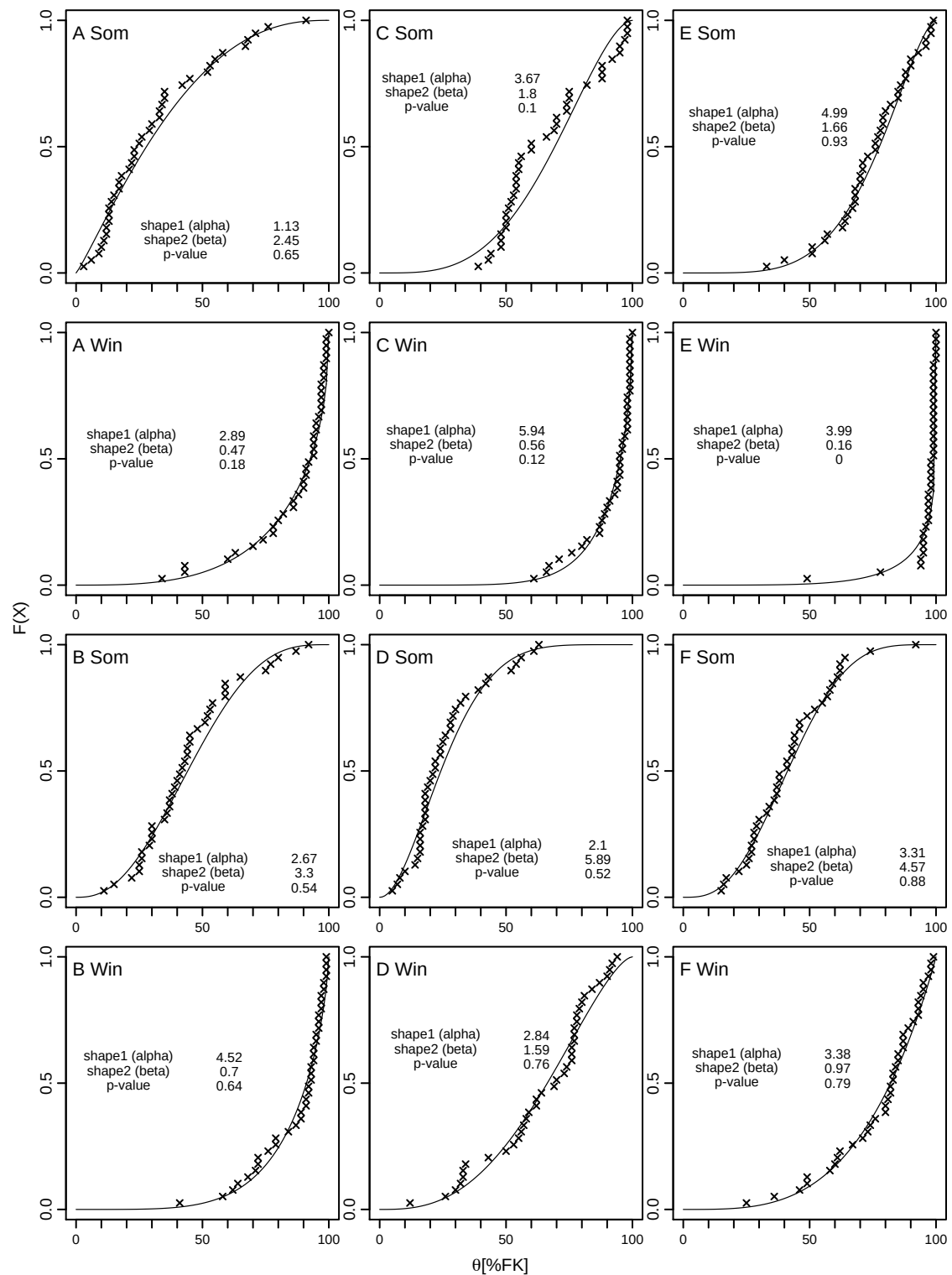
Anhang



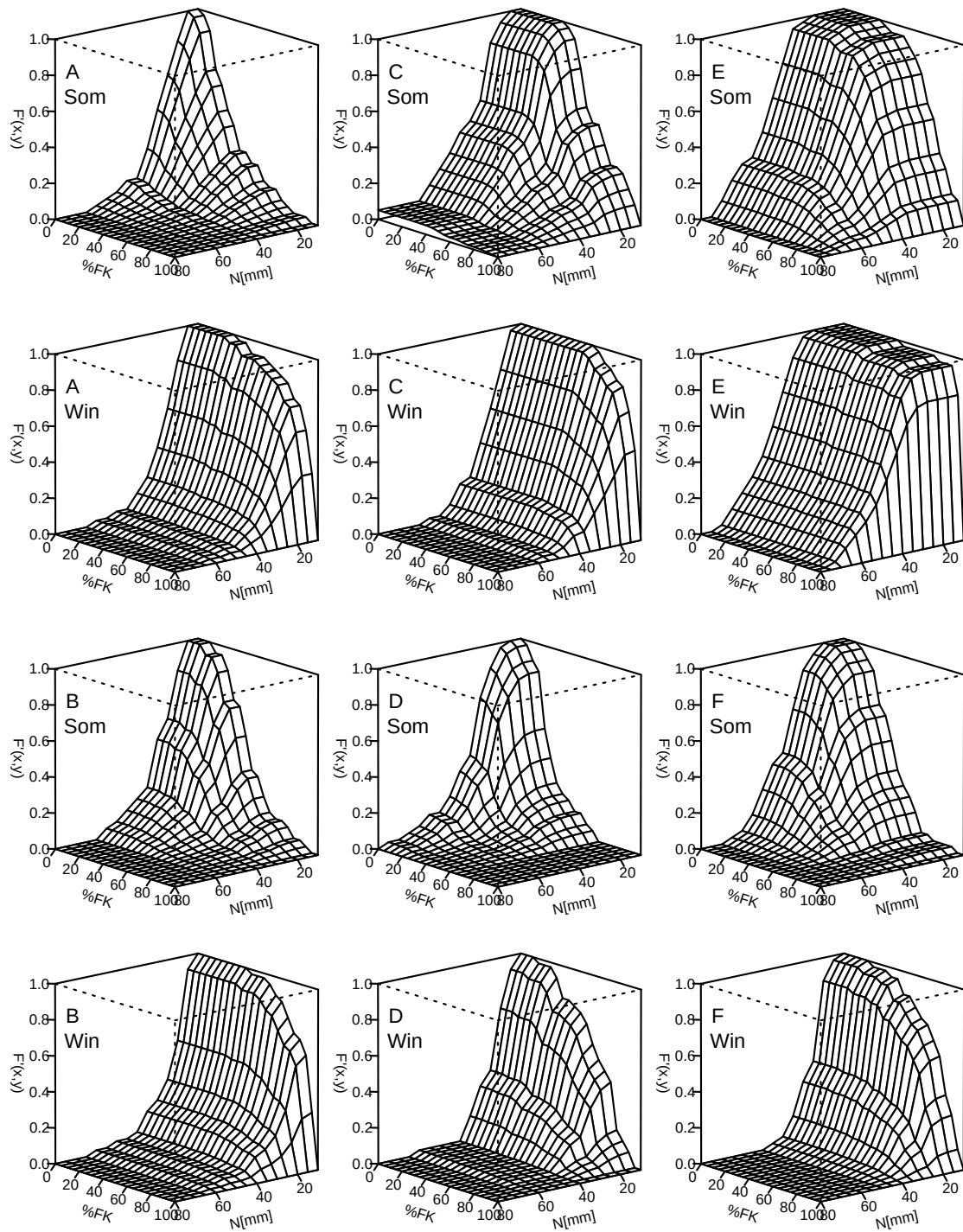
Anhang 1: Zu Abbildung 4.1: Scatterplot der saisonalen $N - \theta_v$ Stichproben.



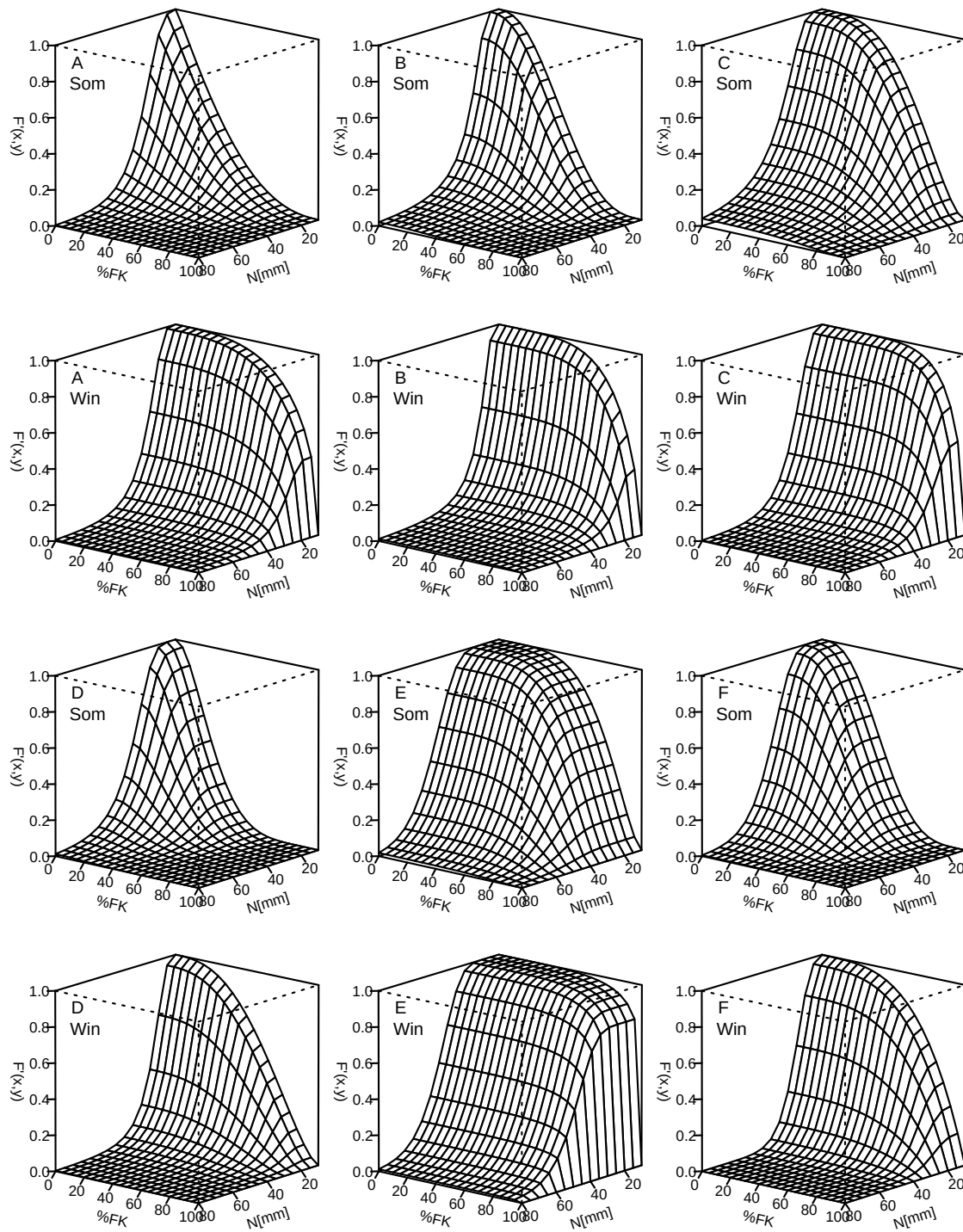
Anhang 2: Zu Abbildung 4.6: Empirische Wahrscheinlichkeitsverteilung (x) und GEV (—) der saisonalen Niederschlagsmaxima mit μ = Lage- (loc), σ = Streu- (scale) und ξ = Formparameter (shape), sowie dem p-Wert des KS-Tests.



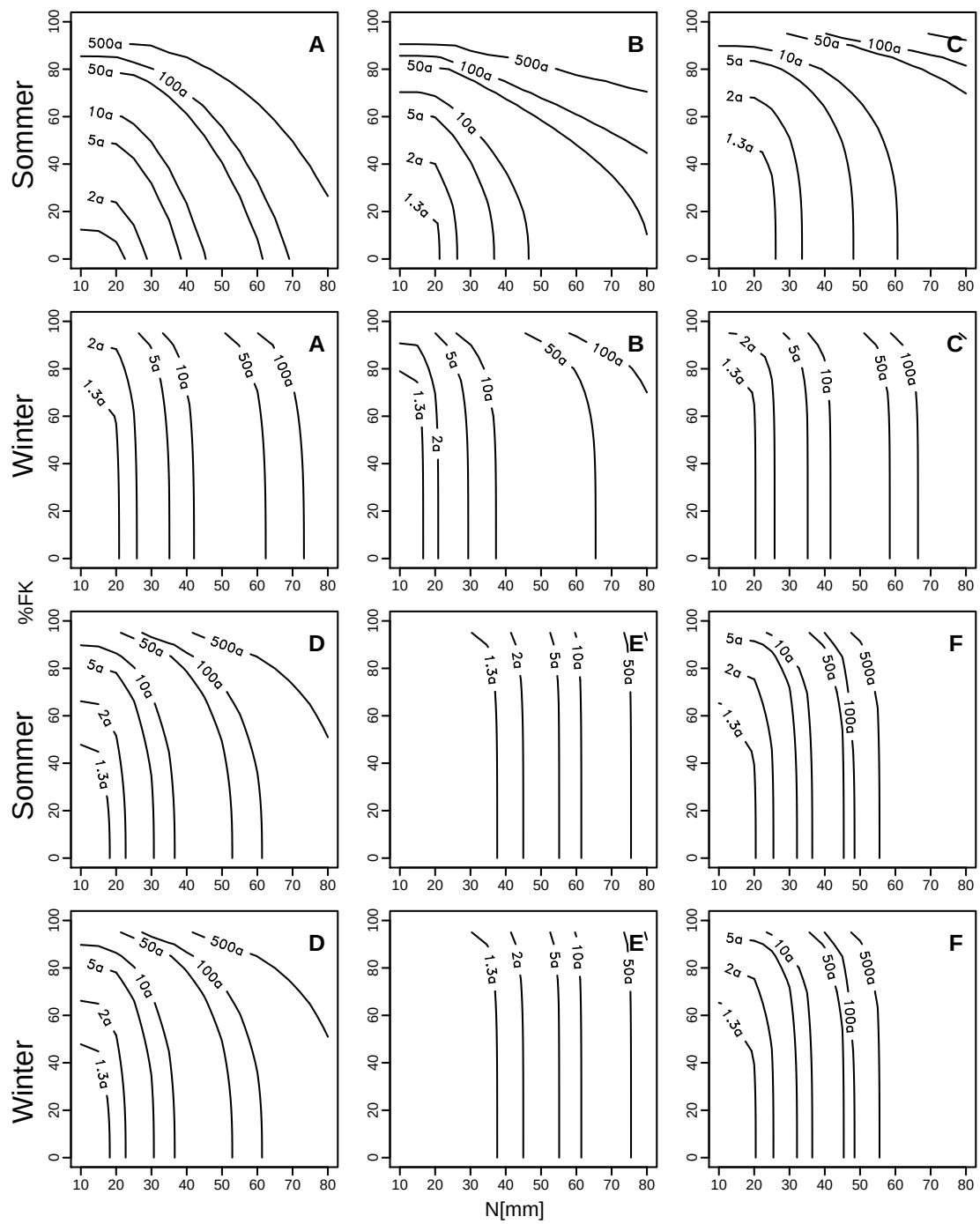
Anhang 3: Zu Abbildung 4.9: Empirische Wahrscheinlichkeitsverteilung (x) und Beta-Verteilung (—) der Vorfeuchtebedingungen θ_v mit den Parametern $\alpha = \text{shape1}$ und $\beta = \text{shape2}$ sowie dem p-Wert des KS-Tests.



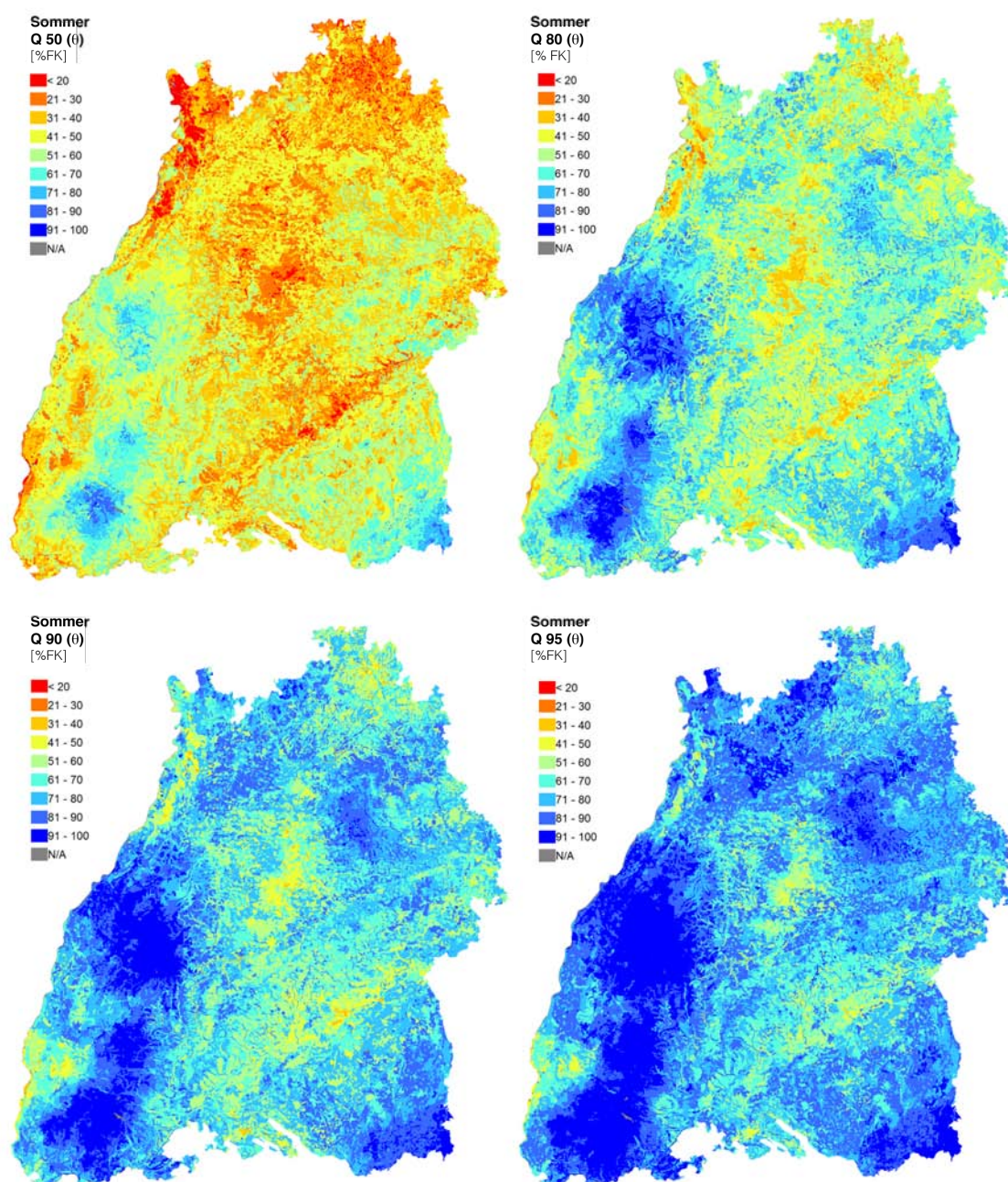
Anhang 4: Zu Abbildung 4.16 : Empirische Überschreitungswahrscheinlichkeiten $F'(x,y)$.



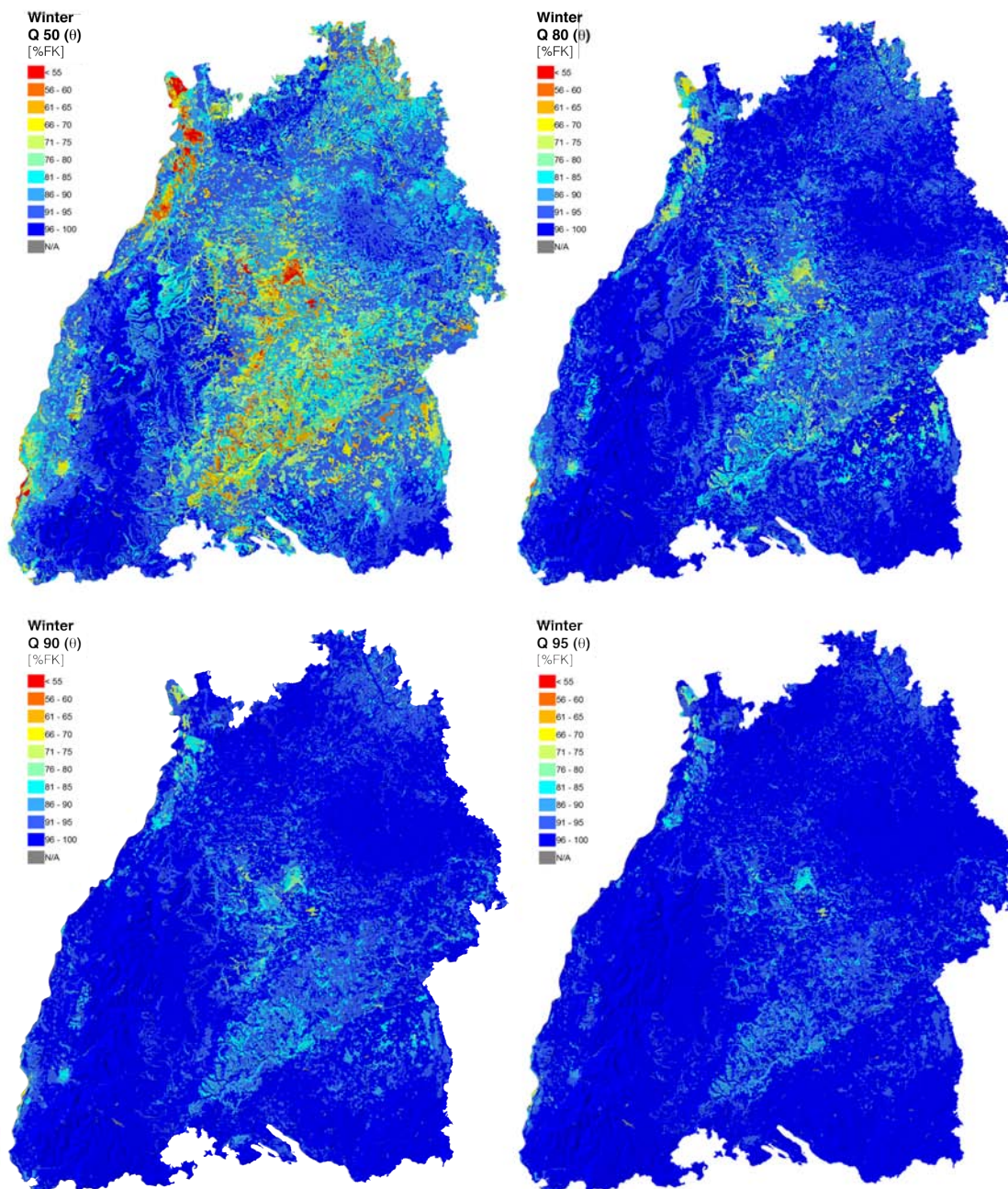
Anhang 5: Zu Abbildung 6.1: Theoretische Überschreitungswahrscheinlichkeiten $F'(x,y)$ mit BV und GEV als Randverteilungen.



Anhang 6: Zu Abbildung 4.18: Gemeinsame Wiederkehrintervalle $T'(X > x, Y > y)$ aus den Randverteilungen BV für θ_v und GEV für N .



Anhang 7: Empirische Quantile (= Unterschreitungswahrscheinlichkeiten) der Vorfeuchtestichproben ($n = 39$) des Sommerhalbjahres.



Anhang 8: Empirische Quantile (= Unterschreitungswahrscheinlichkeiten) der Vorfeuchtestichproben ($n = 39$) des Winterhalbjahres.