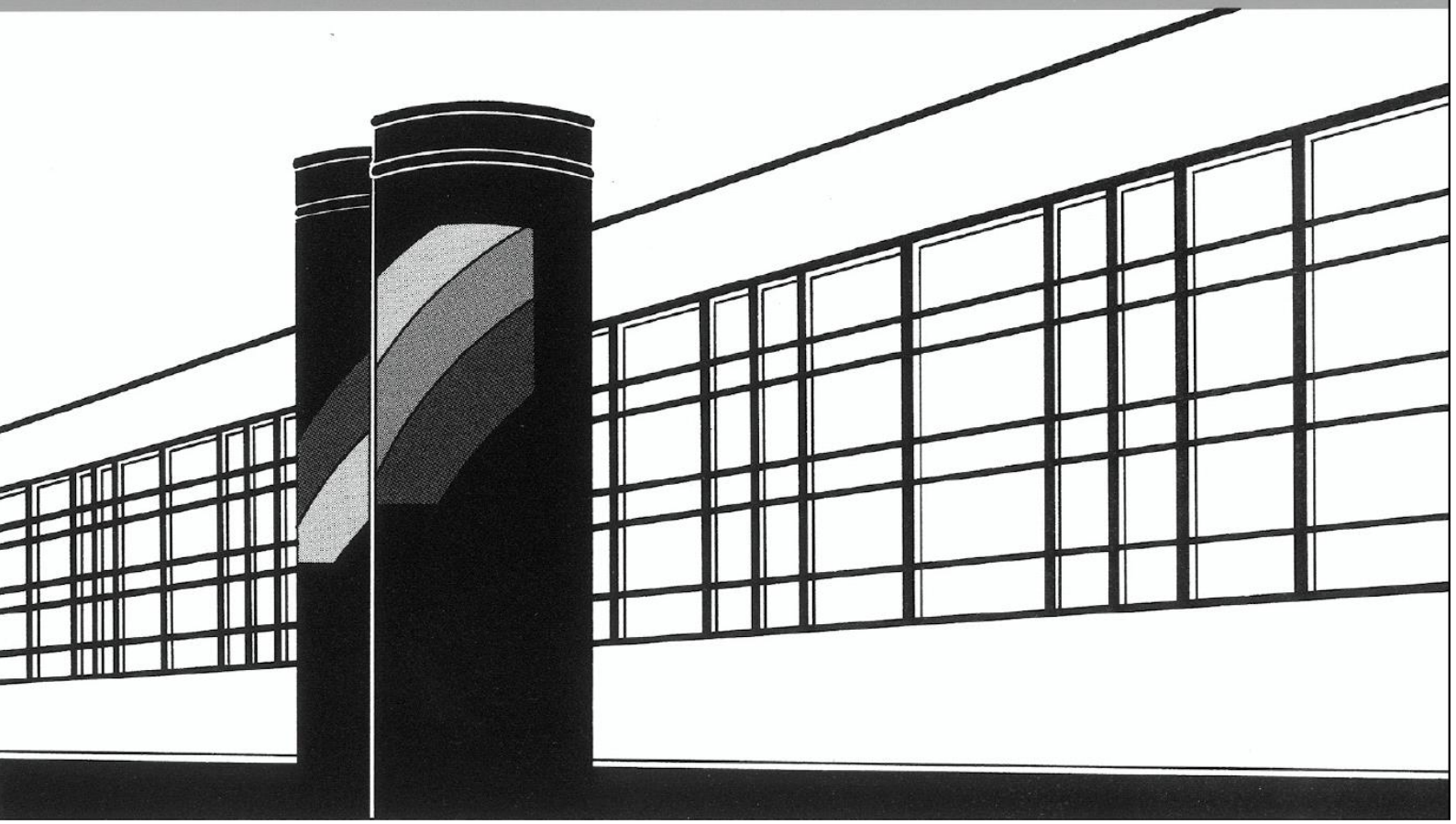
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

# Mitteilungen



Heft 254 Suroso

## Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields

# Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields

von der Fakultät Bau- und Umweltingenieurwissenschaften der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von

**Suroso**

aus Wonogiri, Indonesia

Hauptberichter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy

Mitberichter: Prof. Dr. ir. Niko Verhoest

Tag der mündlichen Prüfung: 25.07.2017

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2017

Heft 254

# Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields

von
Suroso

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

**D93 Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields**

**Bibliografische Information der Deutschen Nationalbibliothek**
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://www.d-nb.de abrufbar

**Suroso:**
Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystem-modellierung, 2017

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 254)
Zugl.: Stuttgart, Univ., Diss., 2017
ISBN 978-3-942036-58-0

NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen



Herausgegeben 2017 vom Eigenverlag des Instituts für Wasser- und Umweltsystem-modellierung

Druck: Document Center S. Kästl, Ostfildern

# Acknowledgement

In the name of Allah and all praise is due to Allah.

I would like to express my sincere gratitude to Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy for his brilliant guidance and wonderful support during my research and also for his excellent recommendation. I would like to thank Prof. Dr. Ir. Niko Verhoest for his great support, in particular for the improving comments and the time for my defense. I would also like to thank Prof. Dr. rer. nat. Jörg W. Metzger for his time, especially for my defense and Dr. -Ing. Gabriele Hartmann as the course Director of the ENWAT International Doctoral Program.

I would like to appreciate to all my colleagues at the Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart for helps, supports, discussion and friendships, especially to Dr. Jochen Seidel, Mrs. Astrid Lemp, Tobias Mosthaf, Faizan Anwar, Thomas Mller, Dirk Schlabing, Ning Wang, Iman Fatehi, Jieru Yan, Micha Eisele, Naibin Song, Ehsan Modiri, and many more.

I would like to thank the rector, dean, and head of civil engineering department of Jenderal Soedirman University and my colleagues who have supported my study.

I would like to thank the German Academic Exchange Service (DAAD) and the German Federal Ministry of Education and Research (BMBF) for providing the financial support.

Finally, I am extremely grateful to my parents (especially for my mother, almarhumah Ibu Sugiyem), my wife Ningrum, my daughter Dhana, and my son Daffa for their supports, patience, understandings and prayers.

# Contents

# List of Figures

# Abstract

Hydrologic system analysis plays an important role in various projects in water resources related issues, such as floods and drought control, as well as in planning, design, and operation of such projects. The role of this system is to produce output from given inputs. Precipitation is one of the major inputs of hydrologic systems. In reality, precipitation data are unfortunately frequently inadequate observation datasets, regarding both record length and completeness due to issues such as a small number of spatial observations and instrument error. Insufficient precipitation data regarding time series length and spatial coverage can lead to serious problems in hydrological analysis, resulting in either underestimation or overestimation of hydrological design values.

To overcome the lack of high quality of precipitation observation data, spatial and temporal models of precipitation are required to fill in missing values, extend the length of data, and interpolate and simulate spatial data at unobserved locations. Many precipitation models have been developed over the past half-century by hydrologists in order to bridge the gap between the need for high-quality precipitation data and the lack of available data in reality. Most precipitation models are developed with either an explicit or implicit underlying Gaussian dependence assumption, which can bias the estimation of reality. One of the main characteristics of Gaussian models is that the observation data are assumed to exhibit symmetric dependence structures, for instance, between low and high values.

The first goal of this study is to empirically investigate the behaviour of the spatial dependence of precipitation fields. This would determine whether the Gaussian assumption is fulfilled in regard to the symmetric spatial dependence structure between low and high precipitation values. The second target is to then quantify the consequences of an asymmetric spatial dependence on the spatial extremes of areal precipitation amounts, where are frequently required for hydrological design.

In order to complete the first goal of this study, an asymmetry function which can incorporate zero precipitation amounts is introduced on the basis of empirical bivariate copulas. Copulas are new tools for multivariate modelling which have been broadly implemented into precipitation applications over the last decade. Copulas are multivariate distributions with uniform marginal distributions used to describe the dependence structure between random variables without information on the univariate marginal distributions. The asymmetric function is then used for the investigations. Investigations of asymmetric spatial dependence are carried out in the regions of Bavaria, Baden-Württemberg and Singapore.

In order to achieve the second target, the symmetric Gaussian dependence based models are evaluated in the context of spatial extremes of areal precipitation amounts over a regular grid and compared to the asymmetric spatial dependence based models. Both models are

implemented in the regions of Singapore and Bavaria using daily precipitation. Gaussian copulas are chosen to represent the symmetric spatial dependence based models because the model is very popular and simple where the dependence structure is completely determined by the correlation coefficient matrix. The V-copulas are selected to represent the asymmetric spatial dependence based models which are constructed from Gaussian copulas through a non-monotonic transformation with the parameters m and k.

Both Gaussian and V-copulas are fitted to the empirical copulas using standard maximum likelihood methods, where zero precipitation amounts are treated as latent variables of a continuous distribution. Zero-inflated precipitation data frequently occur, especially at higher time resolutions (e.g. hourly or daily scales). Precipitation is modelled using a continuous distribution. Dry locations correspond to values not exceeding a threshold in the continuous distribution.

Investigation results prove that precipitation events tend to follow the positive asymmetric spatial dependence structure, in particular at short separating distances. This implies that precipitation with higher intensities tends to be more spatially correlated than lower intensities. This is very interesting since spatial interpolation is commonly carried out using nearby points. Consequently, spatial precipitation models based on symmetric Gaussian dependence could result in underestimation of actual precipitation extremes.

The V-transformed normal copulas provide a possible solution to model the natural processes of precipitation which follow asymmetric spatial dependence structures reasonably well within high multidimensional problems. Empirical investigations focusing on the spatial extremes of gridded areal precipitation amounts reveal that the Gaussian copulas frequently exhibit lower spatial extremes of mean areal gridded precipitation values than the V-copulas. This is an indication that extreme precipitation occurrences, which typically behave in a clustering manner, cannot be modelled reasonably by Gaussian copulas. As a result, Gaussian copulas would yield an underestimation of flood risks and should therefore be implemented with care in the wider practice of flood designs.

# Kurzfassung

Die Analyse hydrologischer Systeme spielt bei unterschiedlichsten Projekten, die sich mit Wasserressourcen verbundenen Fragestellungen beschäftigen, eine entscheidende Rolle. Das können z.B. Fragestellungen im Hochwasserschutz, bei wasserwirtschaftlichen Manahmen in Trockenzeiten oder für die Planung, das Entwerfen und den Betrieb von wasserwirtschaftlichen Projekten sein. Hydrologische Systeme wandeln Eingangsparameter in Ausgangsparameter um, dabei stellt Niederschlag eine der wesentlichen Eingangsgröen dar. Leider sind die vorhandenen Messdaten von Niederschlag oft unzureichend, da Messreihen zu kurz und unvollständig sind sowie in zu geringer räumlicher Dichte vorliegen. Gründe dafür sind z.B. Messfehler oder eine geringe Dichte des Niederschlagsmessnetzes. Das unzureichende Vorliegen von Niederschlagsdaten kann zu ernsthaften Schwierigkeiten bei der Analyse von hydrologischen Systemen führen, was zu Unter- oder berschätzungen von hydrologischen Bemessungswerten führen kann.

Um dem Mangel an hochwertigen Niederschlagsdaten zu überwinden, werden Modelle für die räumliche und zeitliche Modellierung von Niederschlag benötigt. Diese können vorliegende Niederschlagsmessreihen verlängern und Fehlwerte ersetzen sowie durch räumliche Interpolation und Simulation Niederschlagsdaten an unbeobachteten Orten erzeugen. Viele Niederschlagsmodelle wurden im letzten halben Jahrhundert von Hydrologen entwickelt, um die Kluft zwischen dem Bedarf an hochwertigen Niederschlagsdaten und dem geringen Vorkommen an Messdaten zu überbrücken. Bei den meisten verwendeten Niederschlagsmodellen wird entweder explizit oder implizit eine Gausche räumliche Abhängigkeit angenommen, was zu verzerrten Niederschlagsschätzungen führen kann. Eine Hauptcharakteristik von Gau-Modellen ist die symmetrische Abhängigkeitsstruktur zwischen räumlich verteilten Niederschlagswerten, z.B. zwischen hohen und niedrigen Werten.

Das erste Ziel dieser Arbeit ist die empirische Untersuchung der räumlichen Abhängigkeit von Niederschlagsfeldern. Damit soll festgestellt werden, ob die Annahme einer symmetrischen räumlichen Abhängigkeitsstruktur zwischen niedrigen und hohen Niederschlagswerten erfüllt wird. Das zweite Ziel ist anschlieend die Quantifizierung von extremen Gebietsniederschlägen unter Annahme einer asymmetrischen räumlichen Abhängigkeit. Diese werden häufig bei hydrologischen Bemessungen benötigt.

Um das erste Ziel dieser Arbeit zu erreichen, wird eine Asymmetrie Funktion auf der Basis von empirischen Copulas eingeführt, mit der es auch möglich ist Niederschlagswerte von 0 mm mit einzubeziehen. Copulas stellen ein neues Mittel für die multivariate Modellierung dar, die im vergangenen Jahrzehnt eine breite Verwendung bei der Modellierung von Niederschlägen gefunden haben. Copulas sind multivariate Verteilungen mit gleichverteilten Randverteilungen, die für die Beschreibung der Abhängigkeitsstruktur

zwischen Zufallsvariablen verwendet werden. Die Untersuchungen bezüglich einer asymmetrischen räumlichen Abhängigkeit werden für Bayern, Baden-Württemberg und Singapur durchgeführt.

Für die Erreichung des zweiten Ziels dieser Arbeit, werden extreme Gebietsniederschläge die mit einem Gau-Modell erzeugt werden mit Niederschlägen eines Modells verglichen, welches eine asymmetrische räumliche Abhängigkeit abbilden kann. Beide Modelle werden in Singapur und Bayern mit Tagesniederschlagswerten getestet. Gau-Copulas werden verwendet um symmetrische räumliche Abhängigkeiten abzubilden, da sie weit verbreitet und einfach zu verwenden sind. Bei Gau-Copulas wir die Abhängigkeitsstruktur zudem vollständig durch die Korrelationsmatrix bestimmt. V-Copulas werden verwendet um asymmetrische räumliche Abhängigkeiten abzubilden. Sie werden basierend auf Gau-Copulas durch nicht-monotone Transformationen erzeugt und mit den Parametern m und k beschrieben.

Beide Copula Modelle, Gau- und V-Copula, werden an empirische Copulas mittels Maximum-Likelihood-Methoden angepasst. Für die Modellierung der Niederschlagswerte ungleich Null wird eine kontinuierliche Verteilungsfunktion verwendet. Niederschlagswerte von 0 mm werden als latente Variablen von kontinuierlichen Verteilungsfunktionen behandelt, die einen Grenzwert der Verteilungsfunktion nicht überschreiten.

Die Untersuchungsergebnisse zeigen, dass Niederschlagsereignisse einer positiven asymmetrischen räumlichen Abhängigkeitsstruktur folgen, insbesondere für kleine räumliche Distanzen. Das bedeutet, dass hohe Niederschlagswerte räumlich stärker korreliert sind als niedrige Niederschlagswerte. Demzufolge könnte die Anwendung von räumlichen Niederschlagsmodellen, die symmetrische Abhängigkeitsstrukturen verwenden, zu Unterschätzungen von Niederschlagsextremen führen.

V-transformierte Copulas ermöglichen die Modellierung von natürlichen Prozessen wie Niederschlag, die einer asymmetrischen räumlichen Abhängigkeit folgen. Sich auf extremen Gebietsniederschlag konzentrierende empirische Untersuchungen zeigen, dass Gau-Copulas häufig zu niedrigeren Extremwerten als V-Copulas führen. Dies zeigt, dass extreme Niederschläge, die typischerweise eine Clusterbildung aufweisen, mit Gau-Copulas nicht angemessen modelliert werden können. Gau-Copulas würden daher im Gegensatz zu V-Copulas zu einer Unterschätzung von Hochwasserrisiken führen und sollten mit Vorsicht im Hochwasserschutz eingesetzt werden.

# 1 Introduction

## 1.1 Background and motivations

Hydrologic system analysis plays a key role in a variety of projects in water resources related issues, for example, floods and drought control in various project steps, such as planning, design, and operation. A hydrologic system is simply defined as a structure or volume in space surrounded by a boundary that receives water and other inputs, operates on them internally, and produces them as output (Chow, V.T., Maidment, D.R., and Mays, L.W., 1988). In other words, the role of this system is producing output from given inputs. This output, then, depends on the behavior of the inputs, the physical laws involved, and also the nature of the system itself (Clarke, 1973; Dooge, 1968).

Precipitation is one of the primary inputs of hydrologic systems and thus plays a crucial role in the prediction of the output of the systems. Precipitation is highly variable in space and time and shows a stochastical process behaviour. These are the most important reasons why most of hydrologic systems are characterised as stochastic processes, causing it to be a central topic in stochastic hydrology (Bardossy, 1998). In reality, precipitation data are unfortunately frequently inadequate observation datasets, regarding record length and completeness due to issues such as a small number of spatial observations and instrument error. Insufficient precipitation data concerning time series length and spatial coverage can lead to serious problems in hydrological analysis, resulting in either underestimation or overestimation of hydrological design values.

An estimation of hydrological design values derived from the poor quality of precipitation datasets needs serious attention. To overcome the lack of high quality of precipitation observation data, spatial and temporal models of precipitation are required in order to fill in missing values, extend the length of data, and interpolate and simulate spatial data at unobserved locations. In fact, many spatial and temporal models of precipitation fields have been developed over the past half-century by many hydrologists implemented in many different regions around the world with a large number of different approaches to bridge the gap between the need of high-quality precipitation data and the poor available data in reality.

However, numerous spatio-temporal models of precipitation fields are developed with either an explicit or implicit underlying Gaussian dependence assumption. The models developed based on the concept of single/multiple linear regression and its derivations assume that the observation data are normally distributed in time or space, such as autoregressive type models, generalized linear model, kriging type models and many other models. A model assuming the underlying Gaussianity is very popular to be adopted in a variety of applications due to its ease of computation, but this can bias the estimation significantly

from the reality either under- or overestimation if the Gaussian dependence assumption is difficult to be fulfilled. One of the main characteristics of the Gaussian models is that the observation data are assumed to exhibit symmetrical dependence structures, for instance, between low and high values of quantiles.

Precipitation has widely been recognized to show a skewed distribution in univariate cases and is more clearly detected at the higher time resolutions such as hourly and daily time scales. Indeed, for some general non-Gaussian based models, for instance, exponential, Gamma, Gumbel, Weibull, Pareto distribution functions and the extension models such as generalized extreme models; this can be successfully implemented to address the skewed distribution of precipitation at the univariate cases. In contrast, modeling skewed precipitation amounts at multivariate senses is still challenging due to its complexity. Few models have been built to overcome the asymmetric behaviour of precipitation, but many of them are still bivariate. Consequently, deeper empirical investigations of multivariate dependence structure of precipitation amounts are required as a basis for selecting the optimal model.

Another unique characteristic of precipitation is zero-inflated data frequently occurring in high time resolution ranging from daily, hourly, to even minute-by-minute precipitation events (Bárdossy and Pegram, 2009, 2013; Bárdossy and Plate, 1992; Serinaldi, 2008, 2009). This unique topic, dealing with zero-inflated data, remains challenging due to the intermittent property of precipitation in space and time. The dry events with zero precipitation, usually occurring with a high probability, are modeled by discrete distribution (namely, zero or non-zero values), while rainy events describing the precipitation amounts can be modelled using continuous distributions (Bárdossy and Plate, 1992). Even though handling zero-inflated precipitation data by simultaneously modeling discrete and continuous processes has been conducted commonly for the univariate cases using, for example, truncated distributions, implementation for multivariate cases is, however, still problematic (Bárdossy, 2011; Bárdossy and Pegram, 2013).

Recently, copulas, new tools for multivariate modeling, have been popularly implemented into precipitation applications over the last decade (Bárdossy and Pegram, 2009, 2013; Serinaldi, 2009; Vernieuwe et al., 2015). Copulas are multivariate distribution with uniform marginal distributions used to describe the dependence structure between random variables without requiring information on the univariate marginal distributions. Copulas have some major advantages in comparison to the traditional approaches. Analysis of the marginal distribution can be done separately from the joint distribution. Copulas allow the marginal distributions of data to be obtained from different families among other marginals. The marginal distributions can also be a different family from the joint distribution. Copulas are, therefore, invariant to monotonic transformations of the marginal variables. Any monotonic transformation, such as logarithms transformation, a Box-Cox transformation, or a normal score transformation, does not influence the copula (Bárdossy, 2006; Bárdossy and Li, 2008; Nelsen, 2006).

## 1.2 Research goals

The main goals of this research are to investigate empirically the behaviour of the asymmetric spatial dependence of the precipitation fields, and validate whether the Gaussian dependence assumption is fulfilled to the symmetric spatial dependence structure between low and high values of precipitation amounts and then to quantify the consequences of the asymmetric Gaussian dependence based-copula models on the spatial extremes of areal precipitation amounts, which are frequently required for hydrological design. The behaviour of the spatial Spearman's rank correlations of precipitation fields used for the basis of the spatial copula models will be investigated by incorporating zero precipitation amounts which are treated as censored variables of a continuous distribution function.

## 1.3 Structure of the thesis

This thesis is organized into nine chapters with the following structure:

Chapter 1 describes background and motivations why this study is important to be conducted. The main research goals are also expressed in this chapter.

Chapter 2 presents a description of study locations where are located in different climate regions, namely Bavaria, Baden-Württemberg, and Singapore. Characteristics of precipitation data sets and gauge stations, which are available in the study locations, are also explained in this chapter, including typical dominating weather and climate pattern in the regions.

Chapter 3 exhibits an important evaluation of some common parametric univariate distributions (namely, Weibull and gamma distribution) and a non-parametric model (gamma kernel distribution) to be fitted to the precipitation observation for each precipitation event at time step separately. This study is conducted in the regions of Singapore and Bavaria using three different evaluation methods, namely Anderson-Darling tests, Kolmogorov-Smirnov test, and Cramer-Von Misses test.

Chapter 4 explains the basic concept of an empirical dependence measures using copulas, such as Spearman's rank correlation and asymmetry function. A brief introduction to copulas theory in particular on the bivariate cases is presented in this chapter as well. Nevertheless, the spatial copula models for more than two variates are discussed in detail in chapter 7.

Chapter 5 reveals some interesting empirical results of the spatial rank correlation functions of precipitation fields at a variety of timescales conducted on the basis of empirical bivariate copulas by incorporating zero-inflated data treated as censored variables of the censored bivariate Gaussian copulas. The empirical bivariate copulas are constructed from the concepts of the geostatistical and pairwise approach for the regions of Singapore, Bavaria, and Baden-Württemberg.

Chapter 6 exposes new interesting empirical findings regarding the asymmetric spatial dependence structure of precipitation fields at a variety of time resolutions from hourly to

monthly aggregation levels in the regions of Singapore, Bavaria, and Baden-Württemberg on the basis of empirical bivariate copulas which are built using the concepts of the geostatistical and pairwise approach. A unique characteristic of zero-inflated precipitation is certainly taken into account for this investigation. In addition, conditional investigation of asymmetric spatial dependence restricted to days with atmospheric circulation patterns (CPs) is also discussed. An investigation of asymmetric spatial dependence using the simulated precipitation data sets from the physical precipitation model, convection-resolving modeling (CRM), is also presented in this chapter.

Chapter 7 describes the theory of spatial copula models extensively which is also an extension of copula theory introduced in the previous chapters, namely chapter 4 and chapter 5. Moreover, two different copula models, namely the symmetric Gaussian dependence based-model (Gaussian copulas) and the asymmetric dependence based-model (V-transformed Gaussian copulas) are discussed and implemented in the regions of Singapore and Bavaria.

Chapter 8 demonstrates some experiments of the spatial stochastic simulation of gridded daily precipitation using both Gaussian and V-copulas located in Singapore and Bavaria. Some new investigation results regarding the consequences of both models on the spatial extremes of areal precipitation amounts are discussed in this chapter.

Chapter 9 concludes some important findings of this thesis.

# 2 Data and study locations

This study uses high-quality precipitation data sets regarding record length, spatial coverage, and temporal resolutions at the different regions, especially in different geographical locations and different climatic regions, namely, Germany and Singapore. In Germany, this study is carried out in two areas of the Southern Germany; the state of Baden-Württemberg and the state of Bavaria. The investigation areas are shown in Figure 2.1.

## 2.1 Singapore

**Gauge stations network** Precipitation data sets analysed in this study consist of 30 precipitation gauge stations with the high-quality data sets covering a small scale of inter-gauge distances from 1.9 km to 38.7 km as presented in Figure 2.1. The temporal precipitation resolution available in the region is hourly which is recorded completely simultaneously from 30 gauge stations during the period of 1980-2010. In other words, there are no missing data. The precipitation data sets with high quality is provided by the National Environment Agency of Singapore (NEA) which can be assessed on-line through http://www.nea.gov.sg.

In order to investigate the impact of different temporal scales of precipitation fields on the statistical characteristics of interests, hourly precipitation amounts are then aggregated at higher scales, such as the 2-hourly, 3-hourly, 4-hourly, 6-hourly, and 12-hourly aggregation levels, and also accumulated into daily, 5-daily, 10-daily, 15-daily, and monthly values as well. The precipitation gauge stations are classified into different classes ranging from 5 km to 40 km with an increment of 5 km to investigate the influence of a different group of distances on the statistics of interest. All possible combination pairs of the precipitation gauge stations at various locations spaced a constant distance of 5 km are shown in Figure 2.1.

One can see that the number of possible pairs given group distances is more than 30 for the space distance less than 25 km, whereas the number of possible pairs with inter-gauge distance 30 km, 35 km, and 40 km is slightly less than 30, namely, 13, 10, and 4, respectively. Thus, precipitation analysis in Singapore using the geostatistical approach is carried out for the space distance less than 25 km only and the distance more than 25 km are not considered in this study.

**Statistics of precipitation amounts** Singapore is described as a very small, flat and relatively homogeneous region with respect to topography with an area of $\sim 718$ km$^2$ as depicted in Figure 2.1. This country lies on the north of the equator extending from $1.168^o$ to

$1.488^oN$ and from $103.68^o$ to $104.098^oE$. The country is located in the Southeastern Asia and exactly on an island offshore the southern end of the Malayan Peninsula separated between Indonesia and Malaysia. Precipitation in Singapore is quite high with the annual average of precipitation accumulation around 2430 mm and the annual average of probability of having wet day about 51% due to its location near to the equator, which is characterized with a typically tropical climate (Mandapaka and Qin, 2013).

The probability of having wet day occurs in SON (55.8%) and then followed by MAM (53.4%), DJF (51.7%), and JJA (46.6%). The percentage of wet days at any station in the months SON roughly ranges from 41-60% and approximately varies between 46-62% for the months MAM. The proportions of wet days in DJF and JJA are around 39-57% and 38-52%, respectively.

At the hourly scale, the largest percentage of wet hours occurs in DJF (9.24%) and then followed by SON (7.79%), MAM (6.71%), and JJA (6.32%). The percentage of wet hours at any station in DJF roughly ranges from 7.5-10.3% and approximately varies between 5.9-8.8% for SON. The proportions of wet hours in MAM and JJA roughly ranges between 4.7-7.7% and 4.6-7.3%, respectively.

At the monthly scale, the largest portion of wet months occurs in the DJF (99.41%) and then followed by SON (99.33%), MAM (99.31%), and JJA (99.29%). The percentage of wet months at any station in DJF roughly ranges from 93.2-100% and approximately varies between 92-100% for SON. The proportion of wet months in MAM and JJA roughly ranges 91.6-100% and 92-100%, respectively.

Precipitation analysis of the simultaneous spatial observations of all 30 precipitation gauge stations on a given same day during the period of 1980-2010 reveals that the proportion of wet days is slightly lower, roughly around 0.4-0.5. Here, a wet day is defined as a day at which more than 70% of all 30 gauge stations are wet with precipitation depth of more than 0.1 mm during that day. A wet event, an aggregation period, for example from hourly to monthly scale, is defined as an event or an aggregation period at which more than 70% of all 30 gauge stations are wet with precipitation depth of more than 0.1 mm during a time step. The proportions of wet events to the total events for different temporal scales from hourly to monthly are given in Figure 2.2. For hourly scale, for example, the proportion of wet hours approximately ranges from 3%-5.4% across seasons. At the monthly scale, it certainly yields 100% wet months because precipitation occurs every month.

Most of the precipitation events are mainly characterized as tropical convective precipitation which lasts on the short duration less than 1 hour with high intensity (Chatterjea, 1998) occurring in the afternoon and low intensity in the late evening (Beck et al., 2015). The average of hourly precipitation from all stations in DJF reveals the highest value of intensity; around 0.306 mm/hour and followed by SON (0.299 mm/hour), MAM (0.297 mm/hour), and JJA (0.245 mm/hour). The mean hourly precipitation at any station in DJF roughly ranges from 0.285 to0.325 mm and approximately varies between 0.253-0.322 mm for SON. The mean hourly precipitation in MAM and JJA roughly ranges 0.225-0.342 mm and 0.187-0.273 mm, respectively.

The average daily precipitation from all stations in DJF reveals the highest value of intensity;

around 7.4 mm/day and followed by MAM (7.2 mm/day), SON (7.1 mm/day), and JJA (5.9 mm/day). The mean daily precipitation at any station in the months of DJF roughly ranges from 6.8-7.8 mm and approximately varies between 5.4-8.2 mm for MAM. The average daily precipitation in SON and JJA roughly ranges 5.9-7.6 mm and 4.5-6.6 mm, respectively.

The average monthly precipitation from all stations in MAM reveals the highest value of intensity; around 220.36 mm/month) and followed by DJF (219.85 mm/month), SON (214.76 mm/month), and JJA (179.98 mm/month). The mean monthly precipitation at any station in the months of MAM roughly ranges from 165-250 mm and approximately varies between 202-231 mm for the months of DJF. The mean monthly precipitation in SON and JJA roughly ranges 179-230 mm and 136-201 mm, respectively.

If areal mean precipitation amounts are calculated using wet events which precipitation occurs at minimum 70% of all stations, the average areal mean hourly precipitation roughly ranges from 3 to 5 mm per hour for all seasons. At the daily scale, the average of areal mean precipitation varies between 11-13 mm/day across seasons while at the monthly scale, the average of areal mean precipitation ranges from 179 to 220 mm per month in the whole year. The detailed descriptions of areal mean precipitation are shown in Figure 2.3.

The Figure 2.3 depicts not only the average values of areal mean precipitation over all wet time events, but also other descriptions of basic statistics, such as, the average, the standard deviation, minimum value, and maximum value. The statistic values of areal mean precipitation are presented on different time scales from hourly to monthly and different seasons; DJF, MAM, JJA, and SON.

At sub-daily scales, hourly precipitation in DJF and SON presents slightly lower average values of areal mean precipitation; around 3.64 mm and 4.13 mm, respectively. In contrast, hourly precipitation in MAM and JJA exhibits a slightly greater areal mean precipitation roughly about 4.76 mm and 4.33 mm, respectively. The range of average values of areal mean hourly precipitation for DJF, SON, MAM, and JJA are 0.08-32 mm, 0.1-50 mm, 0.1-34 mm, and 0.1-58 mm, respectively.

In contrast, on the daily scale, precipitation in DJF and MAM presents slightly higher average values of areal mean precipitation; around 15.76 mm and 14.54 mm, respectively. Daily precipitation in JJA and SON exhibits slightly smaller values roughly about 14.14 mm and 13.27 mm, respectively. The average values of areal mean daily precipitation for DJF, MAM, JJA and SON roughly ranges from 0.4-232 mm, 0.27-190 mm, 0.17-105 mm, and 0.13-170 mm, respectively.

At the monthly scale, precipitation in DJF and MAM presents slightly higher average values of areal mean precipitation (around 220 mm and 220 mm, respectively). Precipitation in JJA and SON exhibits slightly smaller values of areal mean precipitation (roughly about 200 mm and 215 mm, respectively). The averages values of areal mean monthly precipitation for DJF, MAM, JJA, and SON varies between 12-805 mm, 65-454 mm, 50-350 mm, and 8-350 mm, respectively.

**Global climate regime** In fact, the precipitation mechanism in Singapore is triggered not only by the tropical climatic regions due to its location close to the equator resulting the local convective precipitation, but also strongly influenced by the global climate regime, namely, the Asian Monsoon system. The Asian Monsoon system is basically driven by the movement of the sun bringing predominant atmospheric flow direction which follows an annual cycle.

The Asian Monsoon system in Singapore is classified into two different monsoon seasons including the Northeast Monsoon and the Southwest Monsoon, which are separated by two inter-monsoonal periods (Beck et al., 2015; Fong and Ng, 2012). As a result, there are some seasonal effects of climates in general.

1. The Northeast Monsoon arises from December to early March (the months of DJF). This Monsoon generates a constant north-eastern airflow transporting moisture from the Chinese Sea into the area of Singapore due to the pressure gradient between Asia (Siberia and Tibet) with the high pressure and the Inter-Tropical Convergence Zone (ITCZ) south of Singapore with the low pressure.
2. The first Inter-monsoon period lasts from the late March until May (the months of MAM). The ITCZ moves closer to Singapore toward the north bringing increased shower and thunderstorm activity over the region. Temperatures start to rise over Thailand and China that causing the pressure over the Asian continent begins to decrease. Meanwhile, the surface pressure over Australia starts to increase generating the south-easterly flow.
3. The Southwest Monsoon arises from June to September (the months of JJAS). The ITCZ is situated north of Singapore while the Australian land mass is relatively cooler than the South-East Asian Sea. The pressure gradient between high pressure over Australia and the ITCZ with low pressure generates the South-West Monsoon leading to south-south-easterly wind directions in the surroundings of Singapore.
4. The second Inter-monsoon period occurs from October to November (the months of ON). The ITCZ begins to move back towards the equator bringing increased shower and thunderstorm activity over the region.

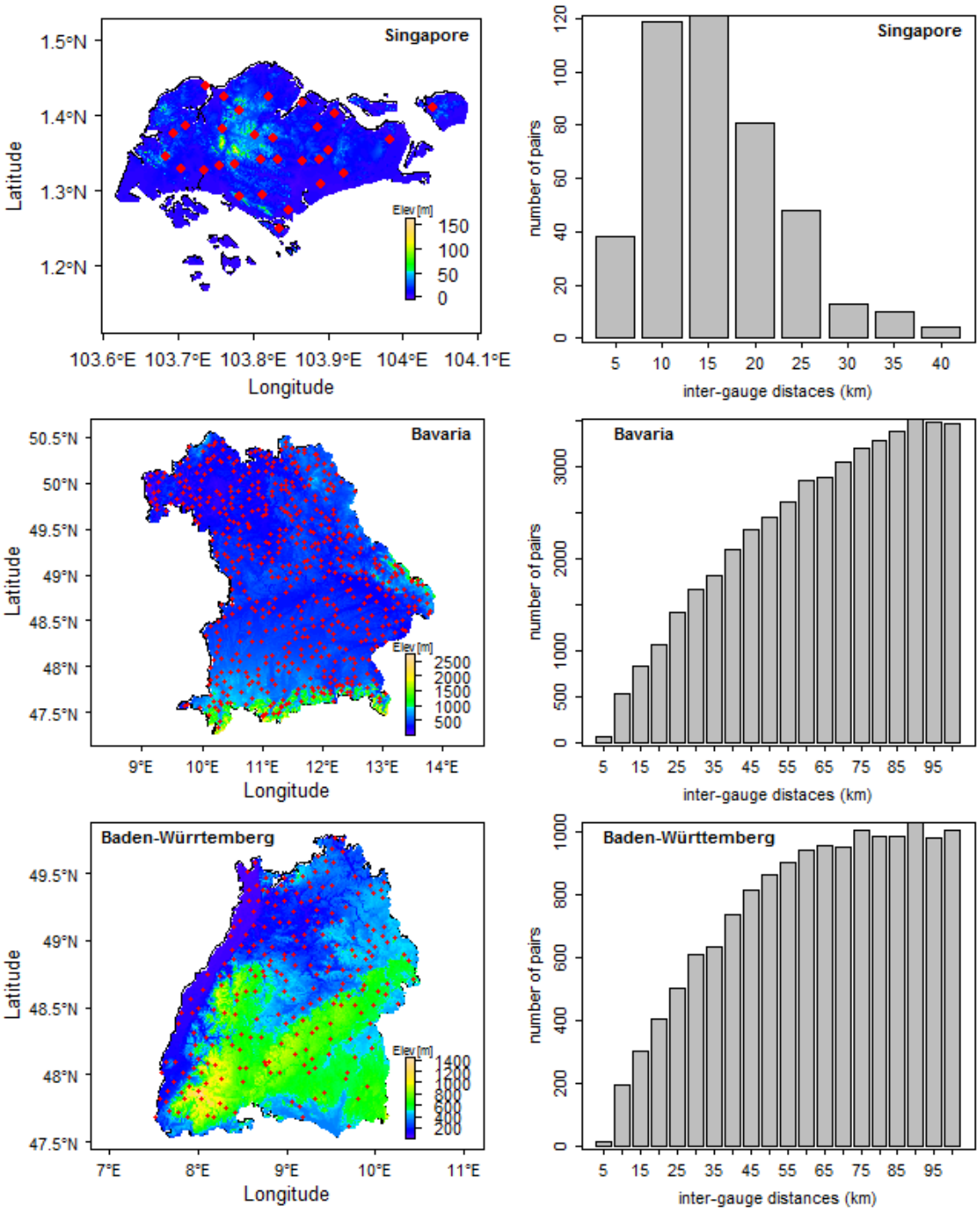



Figure 2.1: Study location: Singapore, Bavaria, and Baden-Württemberg

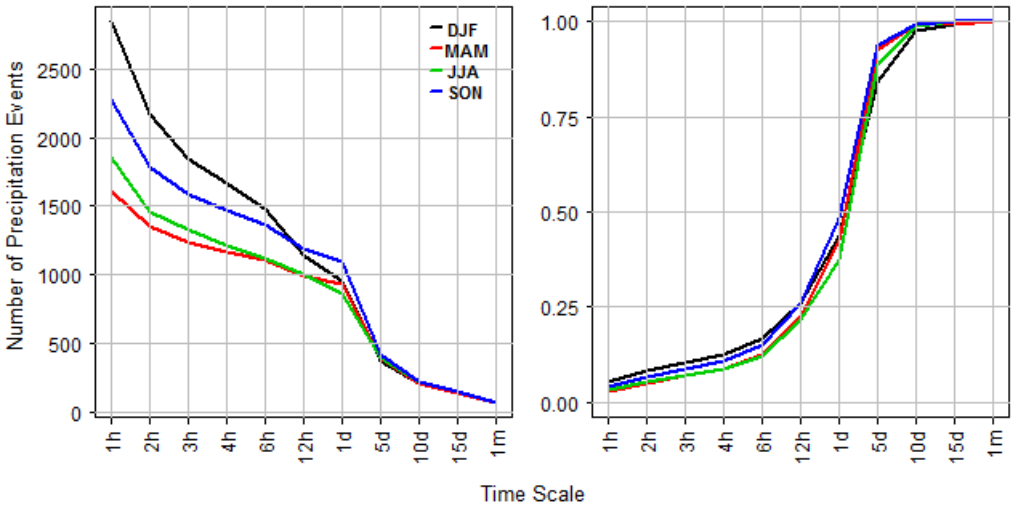



Figure 2.2: Number of precipitation events (hourly (1h) to monthly (1m)) with the probability of wet stations $\geq 0.7$ in Singapore. Left figure: number of wet events. Right figure: proportion of wet events to the total precipitation events.

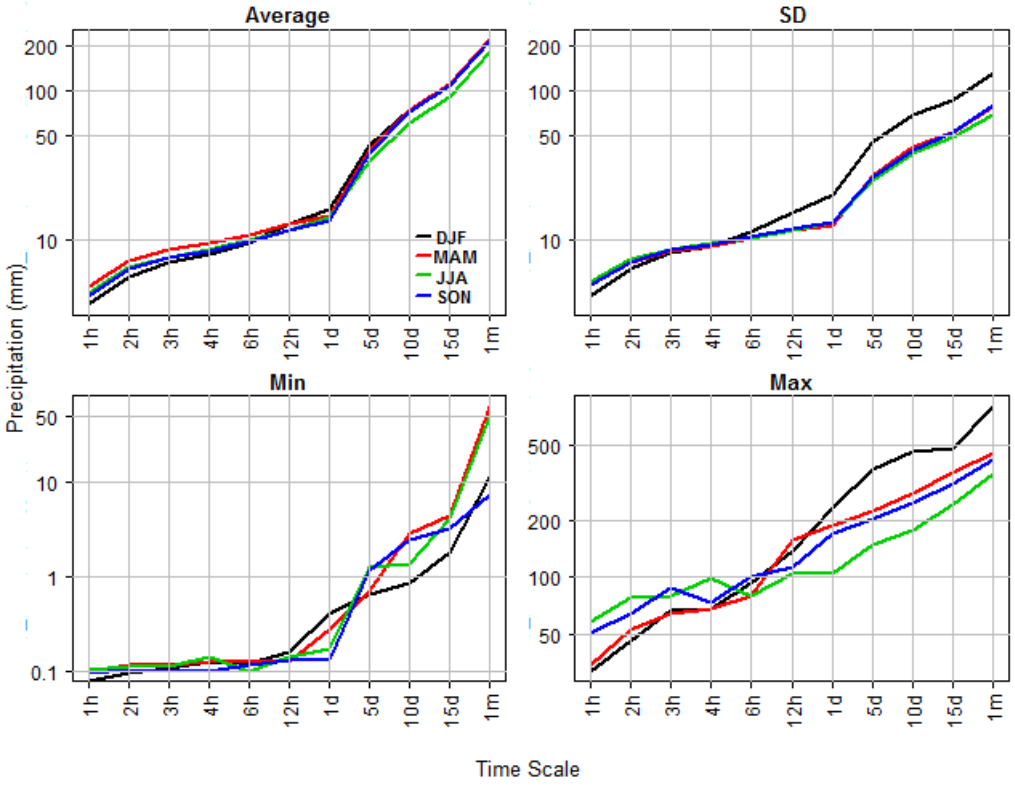



Figure 2.3: Areal mean precipitation of wet events (average, standard deviation, minimum and maximum in mm logarithmic scale) in Singapore.

## 2.2 Germany

**Gauge stations network** The precipitation observation data set considered within this study at the region of Baden-Württemberg consists of 236 precipitation gauge stations. The gauge stations cover a wide range of inter-gauge distances from 3.7 km to 277.6 km. The temporal precipitation resolution used in this study is daily recorded in the period of 1951 to 2001. The number of missing daily precipitation records, which are not used for further investigation, is less than 3% of the length of data sets (1951-2001) because it does not provide complete simultaneous spatial observations at the same time. Nevertheless, this is not essential because this study focuses more on the spatial analysis. In the other region (Bavaria), precipitation used in this study is recorded at 573 precipitation gauge stations. The precipitation stations cover a broad range of inter-gauge distances from 1.19 km to 391.23 km. Similar to the region of Baden-Württemberg, daily temporal resolution measured during the period of 1951-2001 is considered for this study. The number of missing precipitation records, which is not considered from the next analysis, is also less than 3% of the length of data sets (1951-2001). Precipitation observation data sets in both Bavaria and Baden-Württemberg are provided by the German Weather Services.

For further analysis, the daily precipitation data sets in both the regions are then aggregated into higher time scales such as 5-daily, 10-daily, 15-daily and monthly aggregation levels. This is done to study the effect of the increase of temporal scales of precipitation fields on the parameters of interests. Similarly, to see the impact of the increase of spatial distance on the statistical characteristic of interests, the mutual distances between stations are taken ranging from 5 to 100 km with an increment of 5 km. All possible combination pairs of the precipitation gauge stations at different locations, binned within 5 km interval, is presented in Figure 2.1.

One can see that the number of possible pairs of gauge stations at the regions of Baden-Württemberg and Bavaria is generally more than 30 for all different distance groups except for the distance of 5 km in the area of Baden-Württemberg. The number of possible pairs at the region of Baden-Württemberg is 13, 195, 300, 400 for the inter-gauge distances of 5, 10, 15, and 20 km, respectively, and then it increases dramatically to 1003 pairs at the inter-gauge distance of 100 km. The precipitation gauge station density at the region of Bavaria is significantly higher than the region of Baden-Württemberg. The number of possible pairs at the region of Bavaria is 51, 524, 834, and 1054 for the inter-gauge distances of 5, 10, 15, and 20 km, respectively, and then it increases to 3645 at the inter-gauge distance of 100 km. These indicate that the precipitation gauge network in both regions have a sufficient number of gauge pairs at a variety of inter-gauge distances, which should be sufficient for further analysis.

**Statistics of precipitation amounts in Baden-Württemberg** Baden-Württemberg is situated in the north of the equator extending from $7.52^o$ to $10.43^oE$ and from $47.62^o$ to $49.77^oN$ on the southwest of Germany with the area approximately 35 751 km$^2$. Baden-Württemberg is characterized by hilly to mountainous topography ranging from 75 to 1493 m above Mean Sea Water Level (MSWL) with a high forest cover of 39% (Hesse, 2010). Orographic precip-

itation frequently occurs in the higher region (Bárdossy and Pegram, 2013). Elevations at the gauge station, which are used in this study in the region of Baden-Württemberg, ranges from 96 to 1490 m from MSWL with the elevation mean 433.14 m. This points out how high variability of the topographic features spreads out over this region, especially in the Black Forest, being a mountain range with a south-north orientation with elevations up to 1500 m above mean sea level (Warrach-Sagi et al., 2013).

Typical precipitation processes dominating in the region of Baden-Württemberg exhibit seasonal behaviour in an entire year and strongly influenced by orographic precipitation due to the mountainous region in the Black Forest. The average of daily precipitation from all stations in JJA reveals the highest value of intensity; around 3.06 mm/day and followed by MAM (2.49 mm/day), DJF (2.46 mm/day), and SON (2.24 mm/day). The mean daily precipitation at any station in the months of JJA roughly ranges 2.1-5.4 mm and approximately varies between 1.7-5.3 mm for the months of MAM. The mean daily precipitation in the DJF and SON roughly ranges from 1.2 to 5.3 mm and from 1.5 to 5.1 mm, respectively.

The average monthly precipitation from all stations in JJA reveals the highest value of intensity; around 94.37 mm/month and followed by MAM (75.92 mm/month), DJF (73.53 mm/month), and SON (67.44 mm/month), respectively. The mean monthly precipitation at any station in the months of JJA roughly ranges 65.09-166.41 mm and approximately varies between 50.18-161.45 mm for MAM. The mean monthly precipitation in DJF and SON roughly ranges 35.70-198.53 mm and 44.83-152.83 mm, respectively.

The areal mean daily precipitation in Baden-Württemberg calculated using wet events or wet aggregation periods as mentioned roughly ranges 5.33-6.97 mm per day, whereas the average areal mean monthly precipitation varies between 67.44-94.37 mm in the whole year as shown in Figure 2.5a.

On the daily scale, precipitation in JJA and SON presents slightly greater average values of areal mean precipitation, namely, around 6.97 mm and 5.9 mm, respectively. Daily precipitation in MAM and DJF exhibits slightly lower values of areal mean precipitation roughly about 5.41 mm and 5.33 mm, respectively. The average values of areal mean daily precipitation for JJA, SON, MAM, and DJF are 0.42-33.21 mm, 0.5-36.64 mm, 0.47-34.76 mm, and 0.41-38.40 mm, respectively.

In the monthly scale, precipitation in JJA and MAM presents slightly higher average values of areal mean precipitation (around 94.37 mm and 75.92 mm, respectively). Precipitation in DJF and SON exhibits slightly smaller values of areal mean precipitation (roughly about 73.53 mm and 67.44 mm, respectively). The average values of areal mean precipitation for JJA, MAM, DJF, and SON are 20.29-202.49 mm, 18.20-197.79 mm, 8.71-209.84 mm, and 6.88-181.02 mm, respectively.

There are three characteristic patterns of large-scale flow associated with heavy precipitation in the region of Baden-Württemberg (Wulfmeyer et al., 2011):

1. Forced/frontal: typically a frontal line with embedded convection in a region of large-scale lifting. The precipitation occurs initiated by the large-scale forcing, with orographic modification of the flow and surface fluxes playing a secondary role.

2. Forced/non-frontal: synoptic-scale lifting, but no surface front, so that convection breaks out over a wider area. There is significant low-level flow, so orographic forcing is likely to be important, and surface fluxes may also play a role.

3. Air-mass convection (non-forced/non-frontal): occurring in a region of high pressure with no evidence of large-scale forcing at upper or lower levels. Since the low-level flow is also weak, surface fluxes may be dominant in initiating convection.

**Statistics of precipitation amounts in Bavaria** The state of Bavaria lies to the north of the equator extending from $9.12^o$ to $13.78^oE$ and from $47.33^o$ to $50.45^oN$ located in the southeast of Germany with an area of 70 547 km$^2$. In contrast to Baden-Württemberg, Bavaria is characterized as a very structured topography ranging from flat terrain in the northern region to the high mountains in the southern region, namely, the Alps and the Swabian Jura which has a southwesterly-northeasterly orientation and elevations up to 1,000 m above MSWL characterized by steep orography at its boundaries and a high plateau (Bárdossy and Pegram, 2013; Warrach-Sagi et al., 2013). The precipitation gauge stations at the region of Bavaria are elevated ranging from 110 to 1832 m from MSWL, with a mean elevation of 481.26 m. Like other regions in Germany, Bavaria has a typical mid-latitude moderate climate, characterized by a westerly flow with rainfall associated with frontal systems in winter and more convective precipitation in summer (Warrach-Sagi et al., 2013; Wulfmeyer et al., 2011).

Precipitation in Bavaria is mainly characterized by seasonal variations and mostly influenced by the Alpine precipitation climatology. Bavaria, which is located in the northern Alpine region, exhibits convective precipitation during the time period of May-September (Frei and Schär, 1998). This is why the average daily precipitation from all stations in JJA reveals the highest value of intensity; around 3.44 mm/day and followed by MAM (2.29 mm/day), SON (2.219 mm/day), and DJF (2.216 mm/day). The mean daily precipitation at any station in the months of JJA roughly ranges 2.02-8.74 mm and approximately varies between 1.35-6.14 mm for the months of MAM. The mean daily precipitation in SON and DJF roughly ranges 1.43-5.71 mm and 1.28-6.44 mm, respectively.

The average monthly precipitation from all stations in JJA presents the highest value of intensity; around 105.81 mm/month and followed by MAM (69.94 mm/month), SON (66.96 mm/month), and DJF (64.96 mm/month). The mean monthly precipitation at any station in the months of JJA roughly ranges 62.40-268.55 mm and approximately varies between 41.32-185.38 mm for the months of MAM. The mean monthly precipitation in SON and DJF roughly ranges 43.53-171.78 mm and 37.43-189.91 mm, respectively.

The areal mean precipitation in Bavaria calculated using wet events roughly ranges 4.85-7.58 mm per day, whereas the average areal mean monthly precipitation ranges 64-106 mm per month in the whole year as shown in Figure 2.5b. At the daily scale, precipitation in JJA and SON presents slightly greater average values of areal mean precipitation, namely, around 7.58 mm and 6.05 mm, respectively. Daily precipitation in MAM and DJF exhibits slightly lower values of areal mean precipitation roughly about 5.31 mm and 4.85 mm, respectively. The average values of areal mean daily precipitation for JJA, SON, MAM, and DJF ranges between 1.13-42.65 mm, 0.75-30 mm, 0.53-34.25 mm, and 0.36-23.23 mm, respectively.

At the monthly scale, precipitation in JJA and MAM presents slightly higher average values of areal mean precipitation; around 105.82 mm and 69.95 mm, respectively, whereas precipitation in SON and DJF exhibits slightly smaller values of areal mean precipitation roughly about 66.96 mm and 64.97 mm, respectively. The average values of areal mean precipitation for the JJA, MAM, SON, and DJF are 30.07-240.33 mm, 10.65-159.41 mm, 5.69-177.13 mm, and 10.99-173.26 mm, respectively.

## 2.3 Summary

- This study is located in different climate regions, namely, Singapore ($\sim$ 718 km$^2$), Bavaria ($\sim$ 70 547 km$^2$) and Baden-Württemberg ($\sim$ 35 751 km$^2$).
- The temporal resolution of the precipitation data available in the region of Singapore is hourly, recorded completely simultaneously from 30 gauge stations during the period of 1980-2010.
- The temporal resolution of the precipitation data available available in the regions of Bavaria (573 stations) and Baden-Württemberg (236 stations) is daily, measured during the period of 1951-2001 with missing data less than 3% of time series length.
- The average precipitation intensity of the study areas is 0.3 mm/hour (for Singapore), 2 mm/day (for Bavaria) and 3 mm/day (for Baden-Württemberg).

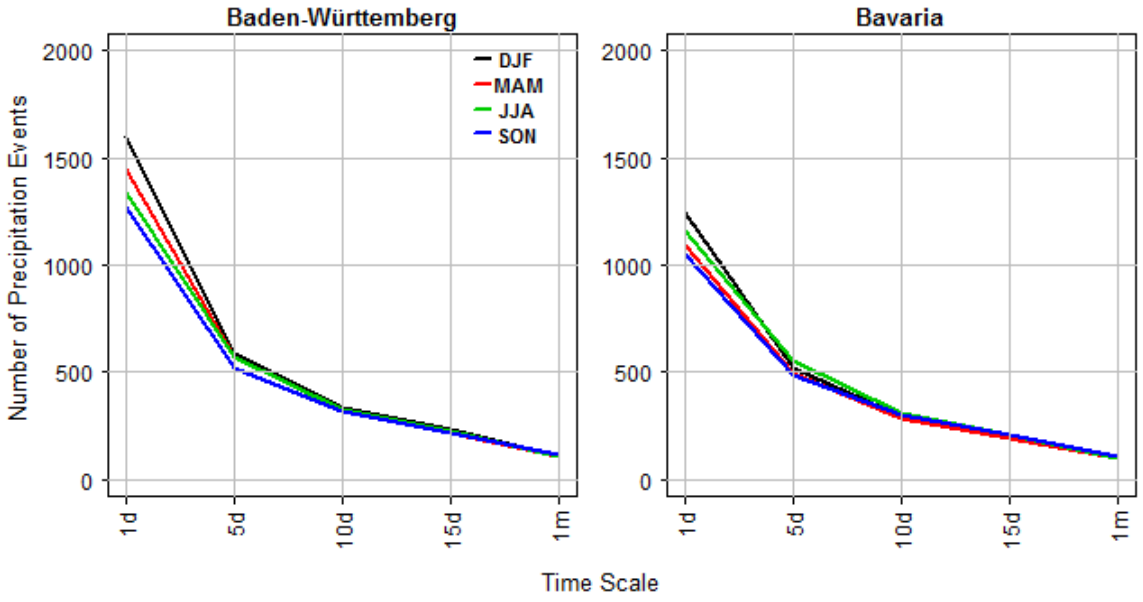
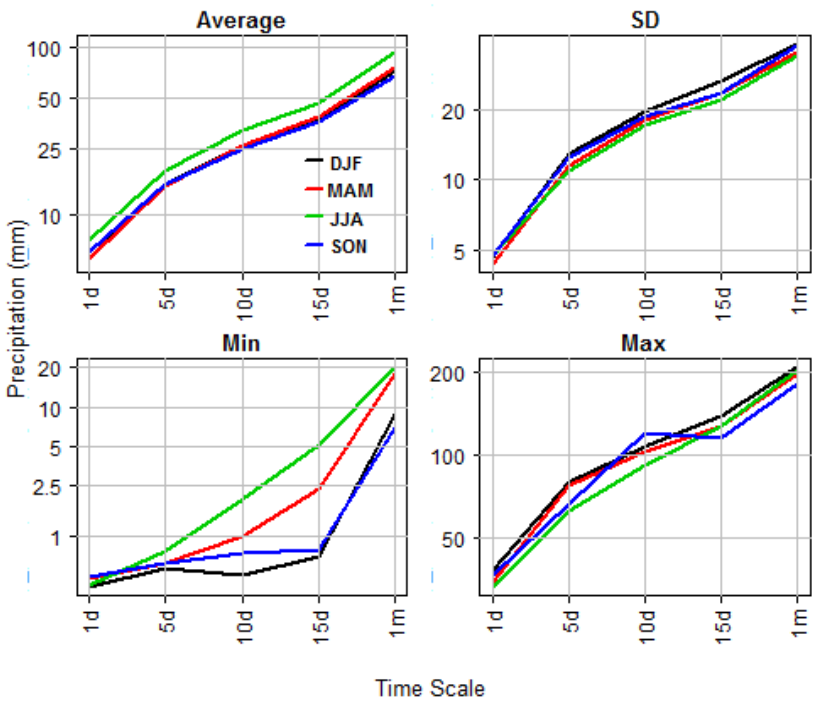
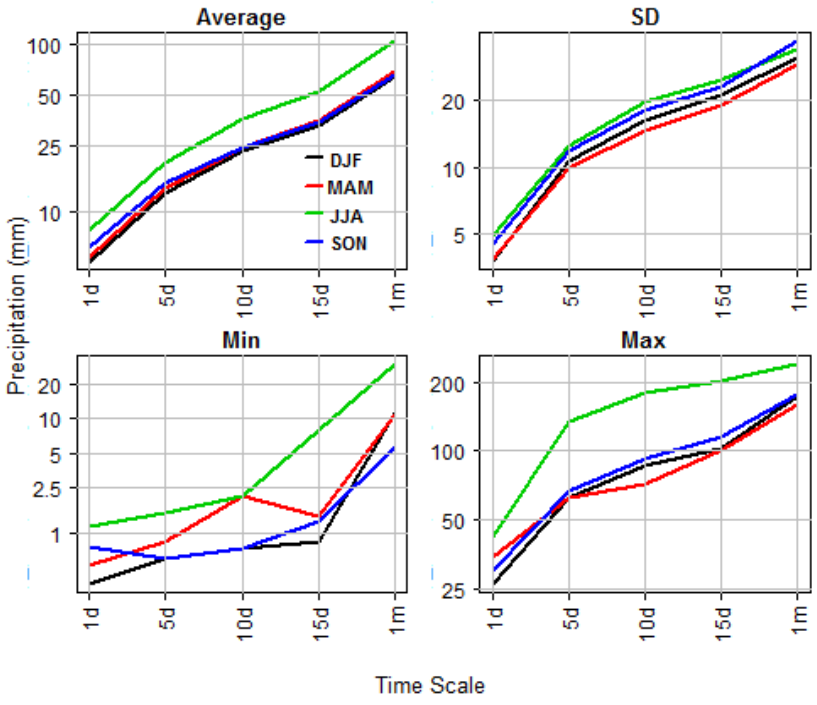



Figure 2.4: The number of events (daily (1d) to monthly (1m)) with the probability of wet stations $\geq 0.7$ in Bavaria and Baden-Württemberg.




(a) Baden-Württemberg




(b) Bavaria

Figure 2.5: Areal mean precipitation of wet events (average, standard deviation, minimum and maximum in mm logarithmic scale) in Baden-Württemberg (a) and Bavaria (b).

# 3 Investigating univariate spatial precipitation

## 3.1 Introduction

Classical geostatistics based spatial interpolations and random field simulations assume an underlying random function at a region of interest to be intrinsically stationary in space. This means that the spatial relation of two sample observation points does not depend on their absolute geographical location itself, but only on their relative locations at a given distance. Moreover, the expected value is assumed to be same over the whole region of interest. This assumption yields a consequence that the marginal distribution in the spatial domain is also assumed to be same for each observation point location (Bárdossy, 2006). In this context, a marginal is fitted for each (fixed) selected time step independently. No temporal stationary is assumed.

Thus, the marginal distribution is defined as the univariate spatial distribution from all spatial precipitation observations points for a given time step.

Many past investigations on the statistical description of precipitation are based on time series observations (Dai et al., 2014; Gyasi-Agyei and Pegram, 2014; Mao et al., 2015; Vogl et al., 2012; Vrac et al., 2007). These were even applied to spatial modelling (Dai et al., 2014; Davison et al., 2012; Gyasi-Agyei and Pegram, 2014; Haddad and Rahman, 2011; Mao et al., 2015), and each grid or every gauge has different marginal distributions (Mao et al., 2015; Vogl et al., 2012).

In hydrological applications, some univariate distribution functions are commonly used to model for precipitation intensity, such as normal distribution, exponential distribution, Weibull distribution, and Gamma distribution (Dai et al., 2014; Mao et al., 2015; Vogl et al., 2012), generalized Pareto distribution (Mao et al., 2015), log-normal (Dai et al., 2014), Gumbel (Gyasi-Agyei and Pegram, 2014), and generalized extreme value distribution (GEV) (Dai et al., 2014; Davison et al., 2012). Recently, Bárdossy and Pegram (2013) introduced the Gamma kernel as the best fit of marginal precipitation intensity. The parameters of the distribution functions are commonly estimated using a standard maximum likelihood approach. The Kolmogorov-Smirnov test or the Chi-Squared test are generally used for the standard goodness of fit tests.

The objective of this chapter is to investigate a suitable univariate distribution function to model the marginal distributions of all spatial observation points for each time step separately at different temporal scales (sub-daily and super-daily) and in different climate regions (Bavaria and Singapore). Some parametric theoretical distribution functions, such as

Weibull and Gamma distribution will be tested to find the best fit marginal model. However, the disadvantage of the parametric approach is that it is not flexible enough to describe the highly variable distributions from time step to time step (Bárdossy and Pegram, 2013). Thus, a non-parametric approach, namely Gamma kernel distribution, will also be tested for selection of the good marginal models.

The precipitation occurrences which are analysed in this study are all selected spatial precipitation events on a given time step at which are more than 0.7 of all gauge stations in the region of interest are wet with a precipitation with a depth of more than 0.1 mm during the time step. In Singapore for example, a precipitation occurrence on a given time event during the period of 1980-2010 is considered in this analysis if at least 21 of 30 precipitation gauge stations are receiving precipitation with more than 0.1 mm during the time step. In Bavaria, for example, a precipitation occurrence on a given time event during the period of 1951-2001 is considered in this analysis if at least 401 of 573 precipitation gauge stations are monitoring rainfall with a precipitation depth of more than 0.1 mm.

## 3.2 Parametric models

The investigations of the model that best fits the univariate spatial precipitation are conducted in Singapore and Bavaria using selected univariate parametric models (the Gamma and Weibull distribution) at the different time scales ranging from hourly to monthly for Singapore and from daily to monthly for Bavaria. The Gamma and Weibull distributions are basically generalizations of the exponential distribution. Both models involve a certain integral known as the Gamma function $\Gamma(k, \theta)$.

### 3.2.1 Gamma distribution

A Gamma type distribution is characterized by the following probability density function.

$$f(x) = \begin{cases} \frac{x^{\alpha-1}\exp^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

$\Gamma(\alpha)$ is the Gamma Function defined by the integral:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \tag{3.2}$$

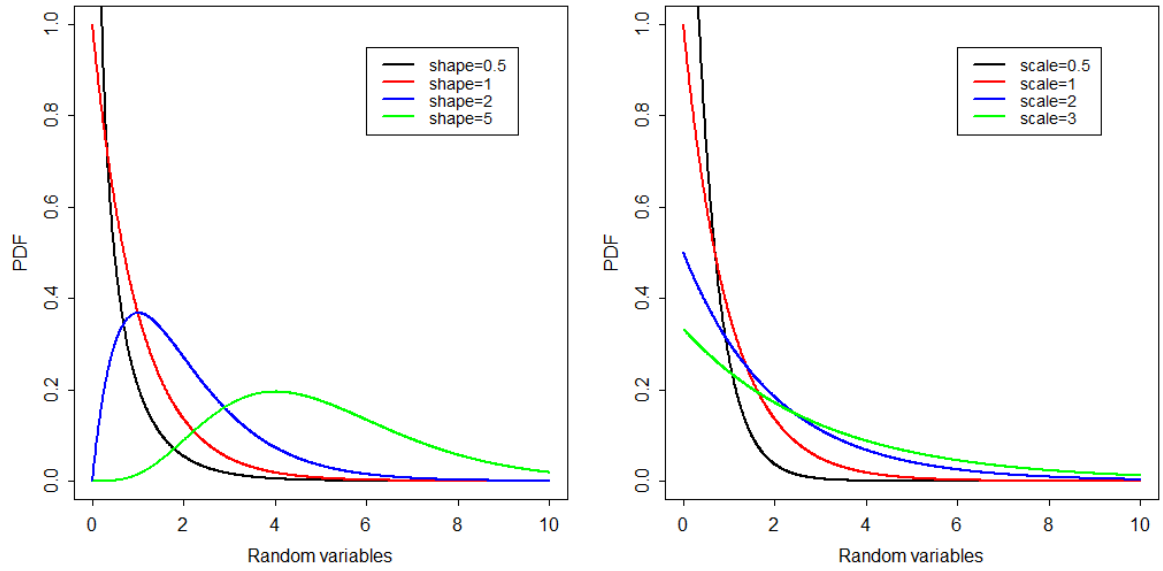
Where $\alpha$ and $\beta$ are both greater than zero. The parameter $\alpha$ is called the shape parameter, while the second parameter, $\beta$, is the scale parameter.

The shape parameter describes the form of the curve. The shape parameter $\alpha$ implies that changing the value of $\alpha$ will change the shape of the distribution. Distributions with a low shape parameter are positively skewed, and as the shape value increases the distribution

curve becomes more symmetrical. This is simply illustrated on the graph 3.1a using a variety of shape parameters $\alpha$ with given scale parameter $\beta$=1.

In contrast, changing the value of scale parameter $\beta$ does not affect the shape of the distribution. Instead, the scale parameter $\beta$ describes how spread out the distributions are. The larger the scale parameter $\beta$, the more spread out the distribution is. The smaller the scale parameter $\beta$, the more compressed the distribution. This is illustrated on the graph 3.1b using a variety of scale parameters $\beta$ with given shape parameter $\alpha = 1$.

For the shape parameter $\alpha \leq 1$, the density function has a maximum at $x = 0$ and is strongly skewed, and then it reduces to the exponential distribution at $\alpha = 1$, which is in fact a special kind of Gamma distribution in which $\alpha = 1$. For larger values of the shape parameter $\alpha$, the density function exhibits a single maximum at $x = \beta(\alpha - 1)$ and still continues to exhibit skewness until the shape parameter $\alpha$ is quite large (Wilks, 1990).




(a) Parameter of scale $\beta$=1

(b) Parameter of shape $\alpha$=1

Figure 3.1: Probability of density function of Gamma distribution with a variety of parameters combinations.

The mean and the variance of the Gamma distribution can be calculated using the following formulas, respectively.

$$E(X) = \alpha\beta \tag{3.3}$$

$$Var(X) = \alpha\beta^2 \tag{3.4}$$

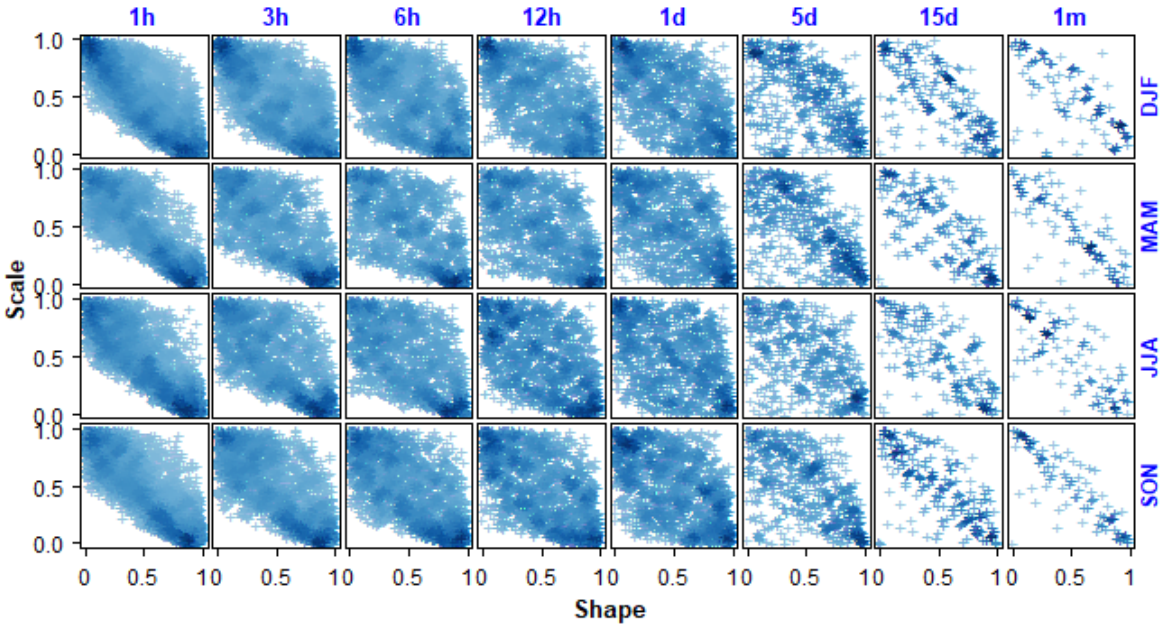
The Gamma distribution is an attractive model, especially for precipitation applications, due to several reasons (Husak et al., 2007; Wilks, 1990). First, the Gamma distribution is relatively straightforward utilizing only two parameters (namely, the parameters of shape and scale), but it offers a flexible representation of a variety of distribution shapes. The Gamma

distribution ranges from exponential-decay forms for shape values near one to nearly normal forms for shape values beyond 20. This flexibility allows for the Gamma distribution to be fit to any number of precipitation regimes with reasonable accuracy, while other distributions may fit only a single rainfall regime. Second, the Gamma distribution is bounded on the left at zero, and it is capable of mimicking non-negative precipitation amounts in a reasonable way. This is an important point to avoid negative values, especially for precipitation events with high variability and low mean values. Finally, the Gamma distribution is mainly characterized by its positively skewed characteristic because it has a long tail to the right of the distribution. Thus, non-zero probability of extremely high precipitation amounts can still be captured by using Gamma distribution. Besides, the Gamma distribution parameters can be scaled to describe precipitation amounts at a variety of time-scales, from hourly to monthly Ison et al. (1971).

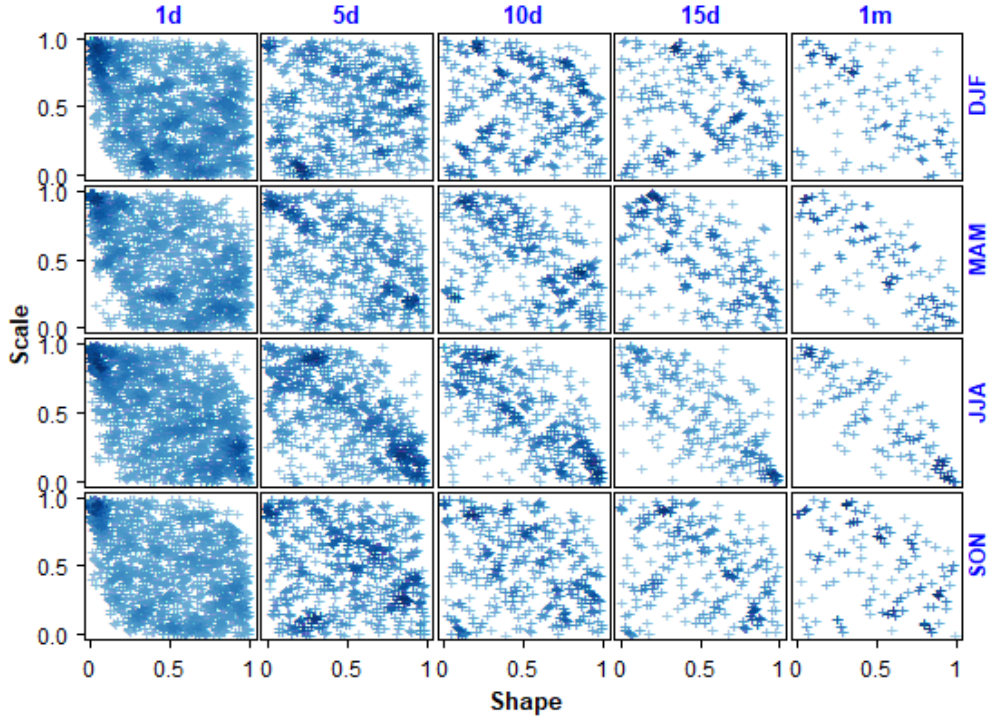
Furthermore, parameters of the Gamma distribution can be used to describe precipitation amounts in a variety of applications, such as the mean and the standard deviation (Husak et al., 2007). However, interpretation of the distribution parameters of the Gamma distribution requires some understanding of the distribution properties. Unlike the normal distribution, where a parameter, such as the mean or the standard deviation, can directly provide an intuitive understanding of some aspect of the distribution, the Gamma distribution requires both shape and scale parameters to be interpreted together. For example, precipitation occurrences with similar shape values, but different scale values, might have very different probability density functions than other precipitation events. Understanding of the distribution of the parameters of shape and scale is, therefore, of importance when describing precipitation events reasonably.

In many practical applications, the simultaneous interpretation of the parameters of shape and scale is challenging. The linear relationship between both parameters is not obviously found and not easy to be interpreted together. In this study, however, both parameters, which are obtained by fitting the Gamma distribution to the empirical distribution of precipitation amounts at a given time step for all selected precipitation occurrences, are scaled in the same unit ranging [0,1] on the uniform distribution using the rank transformation. This procedure will be discussed in detail in the section of copulas as given in section 4.2. As mentioned, the selected precipitation occurrences used in this study are the precipitation values on a given time step at which more than 0.7 of all gauge stations in the region of interest are wet with a precipitation depth of more than 0.1 mm.

Interestingly, overall, there is a reasonably good visual agreement of linear dependencies between the parameters of shape and scale describing precipitation occurrences as presented in Figure 3.2. Figure 3.2 shows multiple scatter plots between the parameters of shape and scale from the Gamma distribution at a variety of time scales from all time periods of different seasons in Singapore and Bavaria. The units are ranked variables scaled on the range (0, 1) using a rank transformation. In Singapore, (Figure 3.2a), there is a strong linear dependency between shape parameters and scale parameters across temporal scales and seasons, but it is a negative linear relationship, which is an consequence of the fact that the product of both shape and scale must equal the mean. If the scale parameter decreases then the shape parameter must increase, in order to preserve a constant mean (Husak et al., 2007).




(a) Singapore




(b) Bavaria

Figure 3.2: Scatter Plots of parameters of Gamma distribution (shape and scale) at a variety of time scales in different seasons in the regions of Singapore (Fig a) and Bavaria (Fig b).

In this case, a high shape value means that the precipitation occurs more spatially symmetrically distributed, implying that the probability of "drier" areas than the average is approximately equal to the probability of "wetter" areas than the average. In general, all spatial observation points are typically received consistent precipitation amounts. In addition, it also results in less variance in the distribution function. In contrast, a large-scale value means that the variance is quite large in comparison to the mean resulting in a more positively skewed distribution function (Husak et al., 2007). Similar to Singapore, there is also a negative linear dependency between the parameters of shape and the parameters of scale across temporal scales and seasons in Bavaria but more scattered as shown in Figure 3.2b.

### 3.2.2 Weibull distribution

A Weibull type distribution is characterized by the following distribution function.

$$F(x) = 1 - \exp^{-(\frac{x}{\beta})^{\alpha}} \tag{3.5}$$

The probability density function of a Weibull random variable can be expressed as follows.

$$f(x) = \begin{cases} \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1}\exp^{-(\frac{x}{\beta})^{\alpha}} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

The mean and the variance of the Weibull distribution can be calculated using the following equations, respectively.

$$E(X) = \beta\Gamma(1 + \frac{1}{\alpha}) \tag{3.7}$$

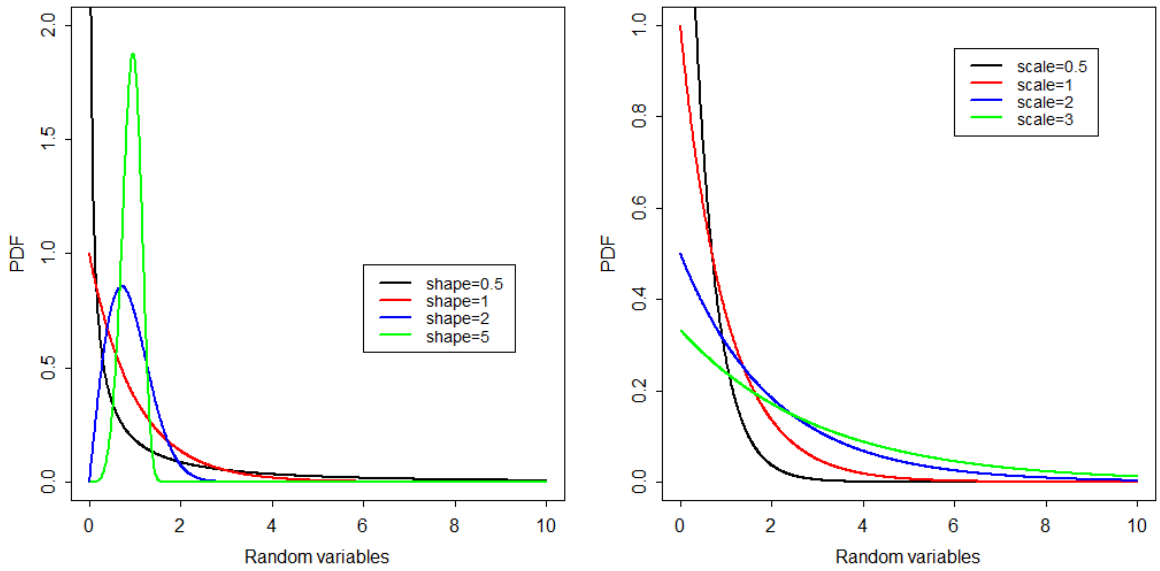
$$Var(X) = \beta^2\Big(\Gamma(1 + \frac{2}{\alpha}) - \big(\Gamma(1 + \frac{1}{\alpha})\big)^2\Big) \tag{3.8}$$

Where $\Gamma$ is the Gamma function as defined previously in Equation 3.2, parameters $\alpha$ and $\beta$ are both greater than zero. The parameter $\alpha$ is also called the shape parameter and the second parameter $\beta$, is named as the scale parameter, which are similar terms to the Gamma distribution's properties.

The shape parameter controls the form of the distribution implying that changing the value of $\alpha$ will change the shape of the distribution. Distributions with a low shape parameter are positively skewed, and as the shape value increases the distribution curve becomes more symmetrical. This can be explained using a density function with a variety of shape parameters $\alpha$ with given scale parameter $\beta = 1$ as presented in Figure 3.3a. For the shape parameter close to 3.6, the Weibull distribution is closely approximating to the Gaussian. The exponential distribution is a special case of the Weibull distribution, for which the shape parameter $\alpha = 1$. The Weibull distribution is increasingly strongly skewed as the shape parameter decreases toward zero (Wilks, 1989). In the other words, the shape parameter governs the tails asymptotic behavior. The distribution belongs to the sub-exponential family with a tail heavier than the exponential one for $\alpha < 1$, while for $\alpha > 1$ the distribution is characterized as hyper-exponential with a tail thinner than the exponential (Papalexiou et al., 2013).

In contrast to the shape parameter, changing the value of scale does not affect the shape of the distribution. Instead, the scale parameter controls the spread of the distribution. The larger the scale, the more spread out the distribution. The smaller the scale parameter, the more compressed the distribution. This is simply illustrated on graph 3.3b using a variety of scale parameters $\beta$ with given shape parameter $\alpha = 1$. For the shape parameter $\alpha \leq 1$, the density function has a maximum at $x = 0$ and is strongly skewed, and then it reduces to the exponential distribution at $\alpha = 1$, which is in fact a special kind of Weibull distribution for which $\alpha = 1$. For larger values of the shape parameter, the density function exhibits a single maximum at $x = \beta(\alpha - 1)$ and still continues to exhibit skewness until the shape parameter is quite large (Wilks, 1989). The density function for the Weibull distribution drops off much more quickly for shape parameter $\alpha > 1$ or slowly (for $\alpha < 1$) than the Gamma distribution. In the case where $\alpha = 1$, both Weibull and Gamma distributions reduce to the exponential distribution.

In practical hydrological applications, the scale parameter of the Weibull distribution reflects the average of precipitation intensity at the specific event in the area of interest, while the shape parameter represents the precipitation variability within the region of interest given an event or the width of the rainfall intensity distribution. A small value of the shape parameter indicates a wide distribution because of a large scattering of the precipitation intensities (contradiction to Gamma distribution). This reflects a tendency toward briefer and more predominantly convective precipitation due to high localized precipitation intensities (Casas et al., 2004; Llasat, 2001; Pinto et al., 2013). In contrast, greater values of the shape parameter indicate a greater tendency toward steadier precipitation derived from large-scale processes (Wilks, 1989).

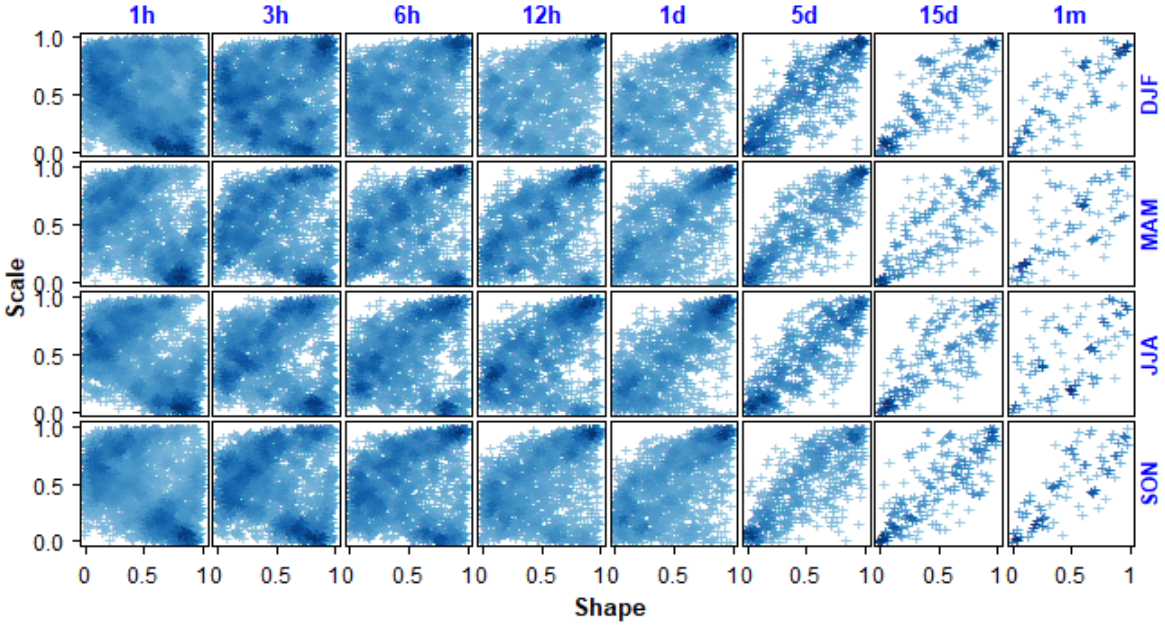



(a) Parameter of scale $\beta$=1

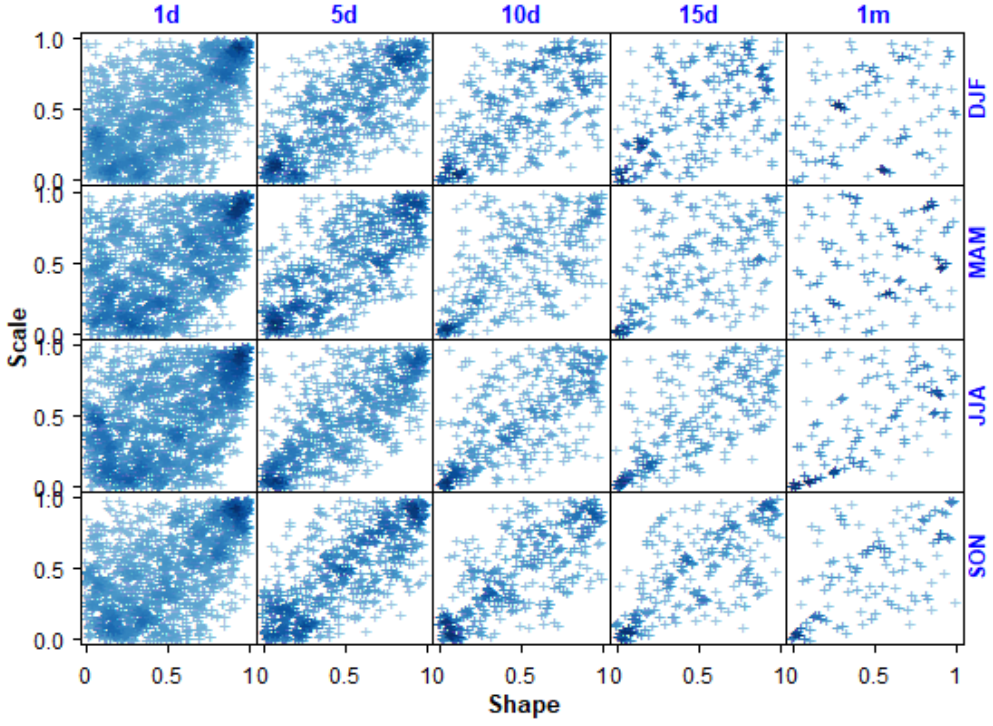
(b) Parameter of shape $\alpha$=1

Figure 3.3: Probability of density function of Weibull distribution with a variety of parameters combinations.

Figure 3.4 shows multiple scatter plots between the parameters of shape and scale from the Weibull distribution at a variety of timescales from all time periods in different seasons and different regions (Singapore and Bavaria). The units are scaled into the uniform domain (0,1) using rank transformation. In Singapore (Figure 3.4a), most precipitation values at different time scales and different seasons show positive linear dependencies between the shape parameters and scale parameters. Only a few cases present contradictory behavior, such as precipitation in the DJF season (with time resolution hourly and 3-hours), in the MAM season (1h), JJA (1h), and in the SON season (1h). Furthermore, the systematic pattern of positive linear dependencies between the shape parameters and scale parameters occur in Bavaria with all different temporal scales and seasons (Figure 3.4b).




(a) Singapore




(b) Bavaria

Figure 3.4: Scatter Plots of parameters of Weibull distribution (Shape-Scale) at various time scales in the regions of Singapore (Fig a) and Bavaria (Fig b).

## 3.3 Non-parametric model

The non-parametric approach is another alternative for fitting marginal distribution of precipitation amounts when spatial variability of precipitation amount is clearly detected in a region of interest. In this study, the Gamma kernel distribution is chosen because it can model positive precipitation amount directly (without transformation) in a more reasonably way, whereas the traditional symmetrical kernels are not able to model non-negative random variables.

Let $x_1, x_2, ..., x_n$ be a random sample from a distribution with an unknown probability density function which is defined on $[0, \inf]$. The standard kernel density estimator for $f$ can be written as follows.

$$f(x) = \frac{1}{nh}\Sigma\Big\{k\Big\{\frac{(x - x_i)}{h}\Big\}\Big\} \tag{3.9}$$

Where $k$ and $h$ are the kernel function and the smoothing bandwidth, respectively.

The kernel function $k$ is usually symmetric and is regarded as less important than the smoothing bandwidth. As precipitation amounts are positive, traditional symmetrical kernels are inappropriate; therefore, Gamma kernels are suggested (Bárdossy and Pegram, 2013):

$$k_{\frac{x}{b}+1,b}(t) = \frac{t^{(x/b)\exp(-t/b)}}{b^{x/b+1}\Gamma(x/b+1)} \tag{3.10}$$

Where the parameter $b(t_j)$ defines the bandwidth and is estimated using the leave one out maximum likelihood method (Bárdossy and Pegram, 2013). Figure 3.5 presents density function of Gamma kernel for a variety of bandwidth that models to a parametric Gamma distribution.

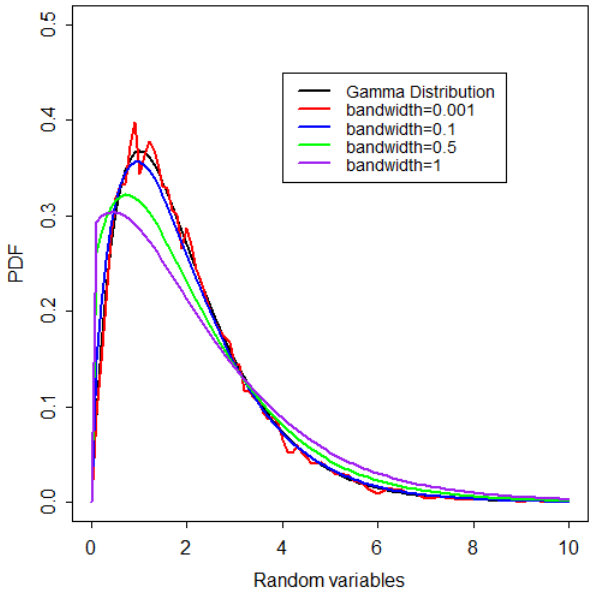



Figure 3.5: Probability density function of Gamma kernel with a variety of bandwidth fitted to the parametric Gamma distribution (shape=2, scale=1).

## 3.4 Evaluation

### 3.4.1 Methods

The evaluation of a suitable distribution or goodness of fit statistics (GOF) aims to measure the distance between the fitted parametric distribution and the empirical distribution. For any given null distribution (theoretical distribution), the GOF is used in order to test whether that the data were likely generated via the null distribution.. The null distribution here is the probability distribution of the test statistic when the null hypothesis is true. In this case, the null hypothesis, which says that theoretical distributions (namely, Gamma, Weibull, and Gamma kernel) and empirical distributions come from the same population, is used with the significance level of 95%. The GOF in this investigation is evaluated using three different measures, namely, Kolmogorov-Smirnov test, Anderson-Darling tests, and Cramer-Von Misses statistical test.

#### 3.4.1.1 Anderson-Darling test

The $k$-sample Anderson-Darling test is a non-parametric statistical procedure that tests the hypothesis that $k$ independent samples with sample sizes $n_1, ..., n_k$ arose from a common unspecified distribution function $F(x)$ (Scholz and Stephens, 1987). The hypothesis is that the populations from which two or more groups of data were drawn are identical. Each group should be an independent random sample from a population. The test statistics are essentially based on a doubly weighted sum of integrated squared differences between the empirical distribution functions of the individual samples and that of the pooled sample. One weighting adjusts for the possibly different sample sizes, and the other is the integration placing more weight on the tail differences of the compared distributions. The AD test is the generalization of the classical Anderson-Darling goodness of fit test (Scholz and Stephens, 1987), and it is used to test the hypothesis that $k$ independent samples belong to the same population without specifying their common distribution function or making any assumption on the parent distribution (Viglione et al., 2007).

The classical Anderson-Darling GOF statistical test is formulated by the following equations in order to test the hypothesis that a random sample $x_1, x_2, ...x_n$, with empirical distribution $F_n(x)$, comes from a continuous population with completely specified distribution function $F(x)$ (Scholz and Stephens, 1987).

$$AD = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x) \tag{3.11}$$

The $k$-sample Anderson Darling test statistics is then defined (Scholz and Stephens, 1987; Viglione et al., 2007) based on the comparison between local and regional empirical distribution functions as follows:

$$AD = \sum_{i=1}^{k} n_i \int_{-\infty}^{\infty} \frac{(\hat{F}_i(x) - H_N(x))^2}{H_N(x)(1 - H_N(x))} dH_N(x) \tag{3.12}$$

where $\hat{F}_i(x)$ is the empirical distribution function of the $i^{th}$ sample (local) distribution function, and $H_N(x)$ is that of joint sample of all $N = n_1 + ... + n_k$ observations (regional).

The AD test function used in this study is available in R (R Core Team, 2013) by the kSamples package (Scholz et al., 2012).

#### 3.4.1.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov (KS) test is by far the most popular method of all the methods for non-parametric goodness-of-fit tests. The KS test is widely used in many applications for statistical testing for the equality of two arbitrary distributions (empirical or theoretical). The idea behind the KS test is fairly simple. Given are the cumulative distribution function $F_n(x)$ of the continuous null distribution, and the empirical distribution function $F(x)$ of the observed data. The null hypothesis, where two distributions $F_n(x)$ and $F(x)$ are identical, is denoted $H_0 : F_n(x) = F(x)$ (Arnold and Emerson, 2011; Massey Jr, 1951).

The KS test measures the Kolmogorov distance between the two distributions, which is defined as the maximum possible value of difference between the two cumulative distributions $|F_n(x) - F(x)|$. This distance can be visualized in graphs as the largest vertical distance between the two cumulative distribution functions. Sufficiently large Kolmogorov distance provides sufficiently low p-values, and thus rejects $H_0$. In other words, high p-values of the test indicate that the distributions of covariates from two matched samples are not statistically different to each other. The KS statistics test measures can be formulated as follows (Massey Jr, 1951; Wilcox, 2005).

$$D_n = \sup_x |F_n(x) - F(x)| \tag{3.13}$$

By the Glivenko-Cantelli theorem, if the sample comes from distribution $F(x)$, then $D_n$ converges to 0 almost surely in the limit when $n$ goes to infinity. In practice, the statistic requires a relatively large number of data points to properly reject the null hypothesis. Critical values for the $D$ statistic is practically calculated using the following formula (Soláková et al., 2013):

$$\begin{aligned} D &= \max(D^+, D^-) \\ D^+ &= \max_{i=1,...,n} (\frac{i}{n} - F(x_i)) \\ D^- &= \max_{i=1,...,n} (F(x_i) - \frac{i-1}{n}) \end{aligned} \tag{3.14}$$

Where $x_i$ denotes the $i^{th}$ order statistic of the random sample and $F(x_i) = P\{X < i\}$ for the distribution that is being fit.

The classical Kolmogorov-Smirnov type non-parametric distance tests have good power properties. Unfortunately, the asymptotic distributions of the test statistics under the null hypotheses are unknown, because they depend on the underlying distribution of the data. In this study, a bootstrap strategy is adopted to overcome this problem (Abadie, 2002). The

Bootstrap based Kolmogorov-Smirnov test is highly recommended because, unlike the standard test, it provides correct coverage even when there are point masses in the distributions being compared. The detail of this strategy is found in Abadie (2002) as follows.

1. The KS statistic for the original samples $X_1, ..., X_n$ is computed
2. Resample n observations $\hat{X_1}, ..., \hat{X_n}$ from $X_1, ..., X_n$ with replacement
3. Divide $\hat{X_1}, ..., \hat{X_n}$ into two samples: $\hat{X_{1,1}}, ..., \hat{X_{1,n1}}$ given by $n_1$ the first elements of $\hat{X_1}, ..., \hat{X_n}$, and $\hat{X_{0,1}}, ..., \hat{X_{n,n_0}}$ given by the $n_0$ last elements of $\hat{X_1}, ..., \hat{X_n}$. Use the two generated samples to compute the KS statistic
4. Repeat Step 2-3, B times. Note that $n_0$ and $n_1$ are constant across bootstrap repetitions
5. Calculate the p-values of the tests by resampling from the pooled data set $X_1, ..., X_2$ we approximate the distribution of our test statistics when $F_1 = F_0$.

The KS test function used in this study is available in R (R Core Team, 2013) by the Matching package (Sekhon, 2011).

#### 3.4.1.3 Cramer-Von Misses test

The Cramer-von Mises (CVM) test is a measure of mean squared difference between the empirical $F_n(x)$ and hypothetical CDF $F_n(x)$ (Anderson, 1962; Baringhaus and Franz, 2004; Laio, 2004). The CVM criterion for testing that a sample $X_1, ..., X_n$ has been drawn from a specified continuous distribution $F(x)$ is given as follows.

$$CVM = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 dx \tag{3.15}$$

It practically can be estimated using the following formula.

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n} (F_i - \frac{2i-1}{2n})^2 \tag{3.16}$$

The CVM statistic takes into account all of the ordered data points and uses a half point correction. It can also be said that under $H_0$, the distribution of $F$ is uniform on $(0, 1)$. This fact could be exploited to make calculations easier. The critical values for CVM have been calculated by Pettitt and Stephens (1976). This function performs the CVM test and tests for normality. It takes a dataset as its argument and returns a p-value. The null hypothesis $H_0$ will be rejected for p-values $< 0.05$.

The CVM test function used in this study is available in R (R Core Team, 2013) by the cramer package (Franz, 2006).

### 3.4.2 Results

Three different statistical tests, namely, the KS test, the AD test, and the CVM test, have been conducted to evaluate 3 selected distribution functions (namely, Gamma, Weibull, and Gamma kernel distribution) for fitting distributions of precipitation amounts from all selected precipitation occurrences during the time period of 1980-2010 (for Singapore) and the time period of 1951-2001 (for Bavaria) where a precipitation amount at every time step is treated separately. The precipitation occurrences, which are analysed in this study, are selected precipitation values on a given time step at which more than 0.7 of all gauge stations in the region of interest receive precipitation as mentioned.

The null hypothesis is that theoretical (namely, Gamma, Weibull, and Gamma kernel) and empirical distributions are the same with a significance level of 95%. The acceptance rates that the null hypothesis accepts can then be calculated over all time periods, as shown in Figure 3.6, using 3 different statistical tests at a variety of different temporal scales in the region of Singapore (Figure 3.6a) and Bavaria (Figure 3.6b). Due to requiring a huge computation effort, Gamma kernel distribution is fitted only to daily precipitation amounts in both regions as shown in Figure 3.7.

**Case study in Singapore** Overall, both Gamma and Weibull distributions could reasonably model univariate precipitation amounts spatially from the observation points. The acceptance rates that both models successfully fit to the empirical distributions of precipitation amounts are roughly above 97%, on the whole, for all different statistical tests. This is an indication that only below 3% of selected all precipitation events which are analysed in this study fail in rejecting the null hypothesis due to p-values being less than 0.05. This applies for all different temporal scales ranging from hourly to monthly in all different seasons as presented in Figure 3.6a.

However, the Weibull distribution performs slightly better in the higher acceptance rates than the Gamma distribution. In DJF for example, the Weibull distribution presents a better acceptance rate than the Gamma distribution at eight different time scales (1h, 3h, 4h, 6h, 12h, 5d, 10d, and 15d) for the KS test and five different time scales (2h, 3h, 4h, 12h, and 10d) for the AD test, and six different time scales (1h, 4h, 12h, 5d, 10d, 15d) for the CVM test. In MAM for another example, the Weibull distribution presents a better acceptance rate than the Gamma distribution at eight different time scales (1h, 2h, 4h, 6h, 12h, 1d, 5d and 10d) for the KS test and seven different time scales (1h, 2h, 4h, 6h, 12h, 1d, and 10d) for the AD test, and five different time scales (1h, 2h, 4h, 6h, and 1d) for the CVM test.

In the JJA season, the Weibull distribution presents a better acceptance rate than the Gamma distribution at eight different time scales (1h, 2h, 3h, 4h, 12h, 5d, 10d and 15d) for the KS test and two different time scales (12h and 1d) for the AD test, and three different time scales (2h, 3h, and 1d) for the CVM test. In the SON season in the last example, the Weibull distribution has a better acceptance rate than the Gamma distribution at eight different time scales (1h, 3h, 4h, 6h, 12h, 1d, 10d and 15d) for the KS test and eight different time scales (1h, 3h, 4h, 6h, 12h, 1d, 10d and 15d) for the AD test, and five different time scales (1h, 4h, 12h, 1d, and 10d) for the CVM test.

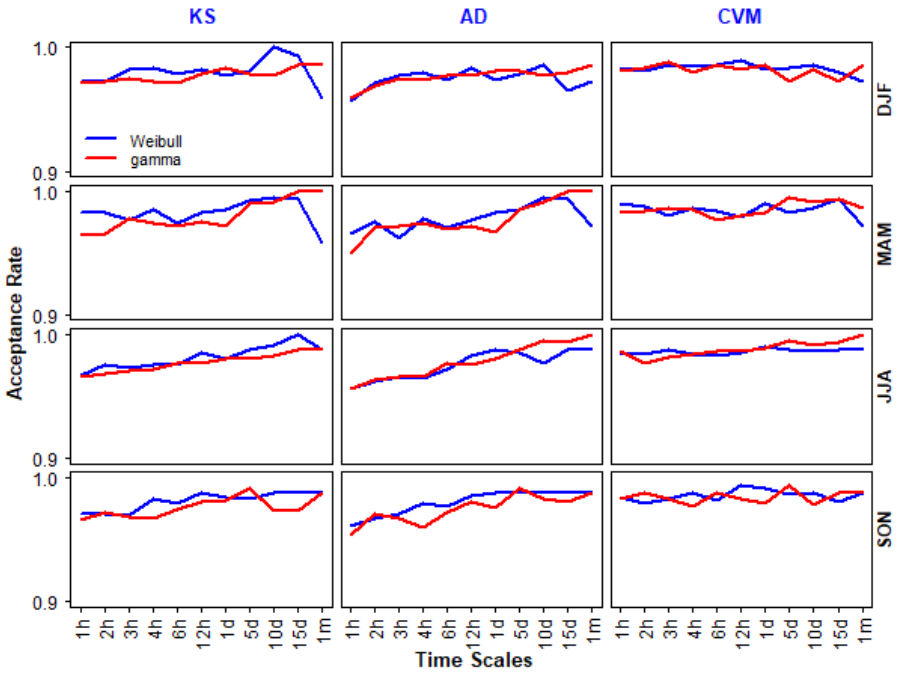
On a daily scale, the Weibull distribution also fits well in comparison to the Gamma kernels as shown in Figure 3.7a and the Gamma kernels perform even worst among other models for all seasons based on all statistical tests.

**Case study in Bavaria** Overall, both Gamma and Weibull distributions could not model the univariate precipitation values from the observation points in Bavaria as good as implemented in Singapore. Both models generally exhibit acceptance rates below 75% based on all different statistical tests at a variety of different temporal scales ranging from daily to monthly scale. This is an indication that for a lot of events the null hypothesis, stating that both theoretical model and empirical distributions come from the same distribution, should be rejected due to the p-values smaller than 0.05 as presented in Figure 3.6b. The acceptance rates decrease with increasing temporal scales. For example, the acceptance rates on a daily scale roughly range from 65% to 77% and then decrease dramatically to 10%-50% for a monthly scale.

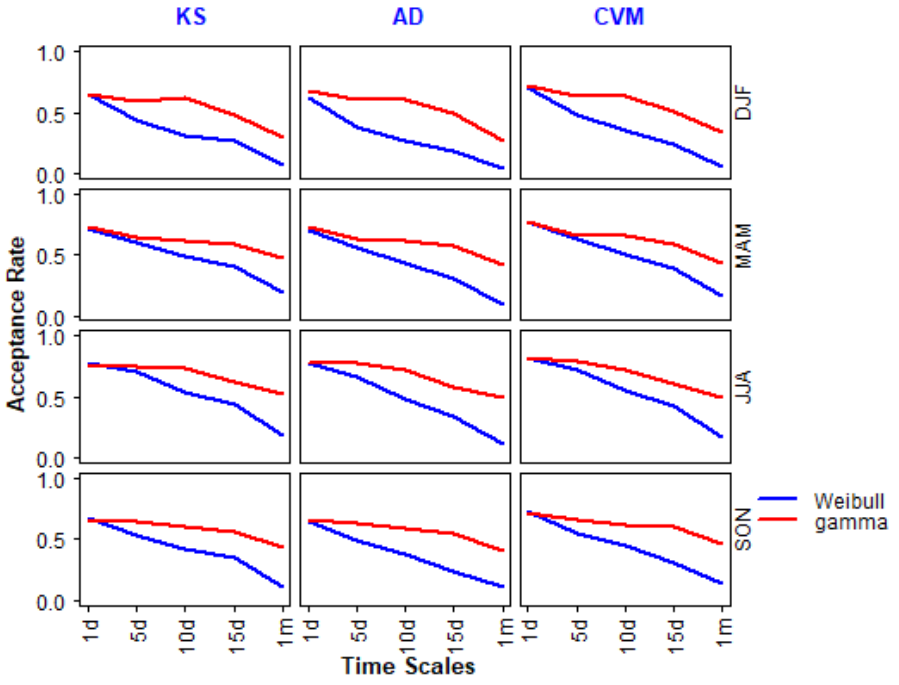
On daily scale, Weibull presents a better acceptance rate than the Gamma On a daily scale, Weibull presents a better acceptance rate than the Gamma distribution in summer and autumn (for the KS and the CVM tests). The rest of the seasons, the Gamma distribution exhibits higher acceptance rates than the Weibull distribution. For all higher time scales, the Gamma distribution exhibits higher acceptance rates than the Weibull distribution. However, in contrast to Singapore, Gamma kernels are superior for all seasons based on all statistical tests as shown in Figure 3.7b with the acceptance rates above 90%, while other models (Weibull and Gamma distributions) present an acceptance rate below 70%.

## 3.5 Summary

- The suitable distribution function applied to univariate marginal distribution functions of all spatial observation points for each precipitation event separately are investigated in the regions of Singapore and Bavaria using the Weibull, Gamma, and Gamma kernel distribution function.
- The precipitation occurrences analyzed in this study are precipitation values on a given time step at which more than 0.7 of all stations in the region of interest receive precipitation depth more than 0.1 mm.
- The best fit of marginal distribution is the Weibull distribution (for Singapore) and Gamma kernel function (for Bavaria).

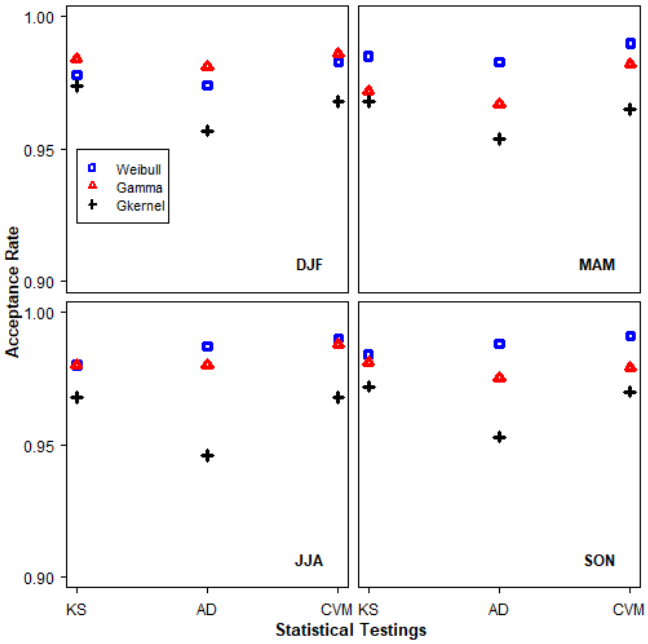



(a) Singapore

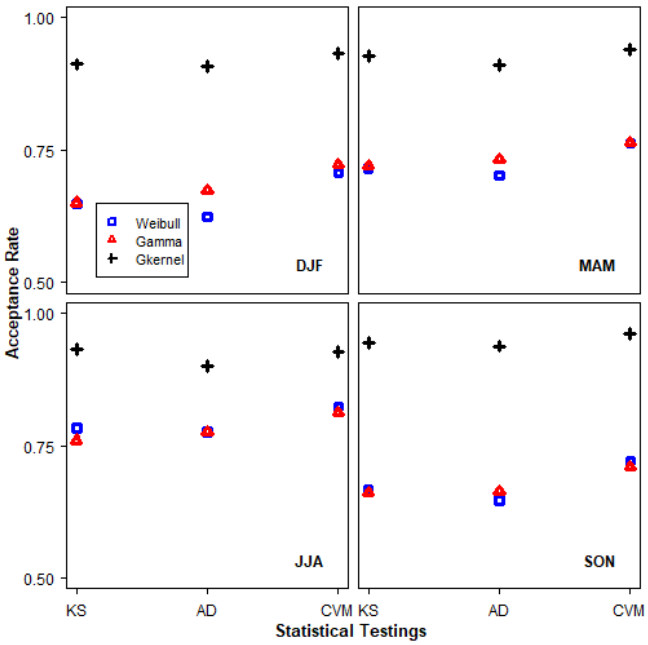



(b) Bavaria

Figure 3.6: Acceptance rates of fitting the univariate distributions of Weibull (blue lines) and Gamma (red lines) to precipitation amounts at a variety of temporal scales in Singapore (Fig a) and Bavaria (Fig b). Panels from left to right represent KS, AD, and CVM statistical tests. Panels from top to bottom represent the seasons of DJF, MAM, JJA, and SON.




(a) Singapore




(b) Bavaria

Figure 3.7: Acceptance rates of fitting the univariate distributions of Weibull (blue lines), Gamma (red lines), and Gamma kernel (black lines) to daily precipitation amounts in Singapore (Fig a) and Bavaria (Fig b) in different seasons of DJF, MAM, JJA, and SON.

# 4 Dependence measures based on bivariate copulas

## 4.1 Introduction

Over the past century, Pearson's correlation coefficient (PCC) has been broadly used in a variety of applications to quantify the strength and direction (decreasing or increasing) of a linear relationship between two variables. In fact, this correlation was the first formal correlation measure of linear dependency between a pair of random variables (Lee Rodgers and Nicewander, 1988) and still the most widely used measure of dependence in various applications.

The PCC between two random variables $X$ and $Y$, which is first developed in 1895, can be calculated using the following formula:

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}} \tag{4.1}$$

where $\rho$ is the PCC, $\mathrm{var}(X)$ and $\mathrm{var}(Y)$ are the variances of $X$ and $Y$, and $\mathrm{cov}(X,Y)$ is the covariance of $X$ and $Y$. The $\mathrm{cov}(X,Y)$ is computed as follows.

$$\begin{aligned} \mathrm{cov}(X,Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(X,Y) - E(X)E(Y) \end{aligned} \tag{4.2}$$

The square root of the variances of $X$ and $Y$ in Equation 4.1, which are also called the standard deviations of $X$ and $Y$, are used as the denominator to adjust the scales of the covariance to have equal units. The correlation coefficient is a covariance which is rescaled by the standard deviations. Because the covariance depends on the scales of measurement for $X$ and $Y$, it is often not a useful descriptive measure of dependence between two variables. When the covariance is divided by the two standard deviations, the range of the covariance is rescaled to the interval between -1 and +1 because the absolute value of the numerator is less than or equal to the denominator. Thus, the interpretation of correlation as a measure of a relationship becomes more tractable than the covariance, and different correlations are more easily compared (Lee Rodgers and Nicewander, 1988).

The PCC is invariant to linear transformations of either variable. The PCC is positive values in the case of an increasing linear association while negative values of $\rho$ indicate a decreasing linear association. If the correlation coefficient $\rho$ is zero, it implies the absence of a linear association between the two variables.

Despite its simplicity, the PCC has several disadvantages due to some limitations. These make that the PCC is frequently interpreted incorrectly. Firstly, the PCC is much more influenced by outliers, which may represent incorrect measurements or extreme events, than by observations near the mean. Consequently, the existence of extreme events can lead to a biased estimate of the correlation (Habib and Krajewski, 2001).

Secondly, the PCC could be very sensitive to the data which exhibit non-normal distribution (Kowalski, 1972) because this measure of linear dependence involves the underlying hypothesis that the data follow a bivariate normal distribution (Serinaldi, 2008). If the normality assumption is difficult to be fulfilled, the measure of dependence can lead to high bias in estimation. For very skewed distributions, for example, the Pearson's sample correlation coefficient can give a poor estimation of the population value (Habib and Krajewski, 2001).

Thirdly, the PCC is only accurate if the correlation relationship is linear and contemporaneous (Anscombe, 1973). Anscombe (1973) also identifies four data sets with the same correlation and other statistical properties (such as skew and mean) that can have completely different distributions of data.

Therefore, even though the PCC is a very useful statistical measure, it is limited by the assumptions underlying the estimation and should only be applied when those assumptions are satisfied. Non-parametric rank measures of dependence, such as Spearman's and Kendall's dependence measure free from distributional assumptions are suggested as possible alternatives to the PCC.

Precipitation amounts which are characterized by highly skewed distributions with zero-inflated data needs another spatial dependence measure as an alternative. The Spearman's rank correlation function is one of the alternatives of dependence measures to replace the Pearson correlation function because of being free from the distribution assumption (Habib and Krajewski, 2001). Furthermore, he Spearman's rank correlation is also less influenced by outliers and it can capture underlying nonlinear relationships between two variables.

## 4.2 Bivariate copulas

A copula was originally introduced by Sklar (1959) as a useful method for deriving joint distributions given the marginals, especially when dealing with non-normal distributions. The method is also likely to be beneficial when the marginals can be stipulated definitely, but the joint distribution is very tricky to be constructed (Cameron et al., 2004). Even though copulas are not new, it has recently attracted great attention. The name "copula" is basically from "couple" to emphasize the manner which a copula combines a joint distribution function to the marginals (Nelsen, 2006). In other words, a copula is a function that links a multidimensional distribution to its marginals.

The bivariate copula is defined as a joint bivariate distribution with the univariate marginals in standard uniform U(0,1), and it is expressed mathematically as follows.

$$C : [0,1]^2 \rightarrow [0,1] \tag{4.3}$$

A bivariate distribution function with given marginal distributions in the original domain $H(x_1, x_2)$ can be modeled by a bivariate copula function $C$ given marginal distributions in the uniform domain $U(0, 1)$ with the help of Sklar's theorem (Nelsen, 2006; Sklar, 1959). The marginal distributions in the uniform distribution U(0,1) can be carried out by a transformation using its cumulative distribution function $F_1(x_1)$ and $F_2(x_2)$ or the rank transformation of the original data values. Hence, analysis of marginal distributions can be done separately from the joint distribution without requiring simultaneous analysis. This means the parameters estimation of marginal distributions and joint distribution in the one-stage maximum likelihood estimation is not necessary. The relationship between the joint bivariate distribution function $H$ and the copula function $C$ can be formulated as follows.

$$H(x_1, x_2) = C\Big(F_1(x_1), F_2(x_2)\Big) \tag{4.4}$$

One can see that a copula function $C$ describes the dependence structure of bivariate distribution function $H$ independently from the marginal distributions. This flexibility of selecting the family of marginal distributions is another benefit implementing copula in multivariate problems, while the traditional multivariate models require all univariate marginal distributions coming from the same family of marginal distributions.

If marginal distribution functions $F_1$ and $F_2$ are all continuous meaning that the bivariate distribution function $H$ is also continuous, then the copula $C$ is unique (Nelsen, 2006). The bivariate copula function $C$ conversely can be constructed from a joint bivariate distribution function $H$ (Nelsen, 2006) as follows.

$$C(u_1, u_2) = H\Big(F_1^{-1}(u_1), F_1^{-1}(u_2)\Big) \tag{4.5}$$

Where $F_1^{-1}, ..., F_n^{-1}$ are the inverses of marginal distributions.

Theoretically, the copula density $c$ can be defined by the density of bivariate distribution divided by the product of densities of the univariate marginals in this manner.

$$c(u_1, u_2) = \frac{h\Big(F_1^{-1}(u_1), F_2^{-1}(u_2)\Big)}{\prod_{i=1}^{2} f_i\Big(F_i^{-1}(u_i)\Big)} \tag{4.6}$$

Where $h$ is multivariate distribution density function corresponding to $H$ and $f_i$ are density functions of marginal distributions corresponding to $F_i$

For an empirical bivariate copula, its densities can be constructed from the 2-dimensional histogram of a scatter plot of the marginals $u_1$ and $u_2$. If data are few, the densities of empirical bivariate copula can also be estimated via a 2-dimensional kernel density estimation (Guthke, 2013).

Figure 4.1.a illustrates the scatter plot of the marginal distribution in the original domain. It can be seen visually that there is an obvious correlation between the variables. But this different strength of dependence for various percentiles can hardly be seen. Figure 4.1.b and Figure 4.1.c describe the process of transformation of the marginal distribution into

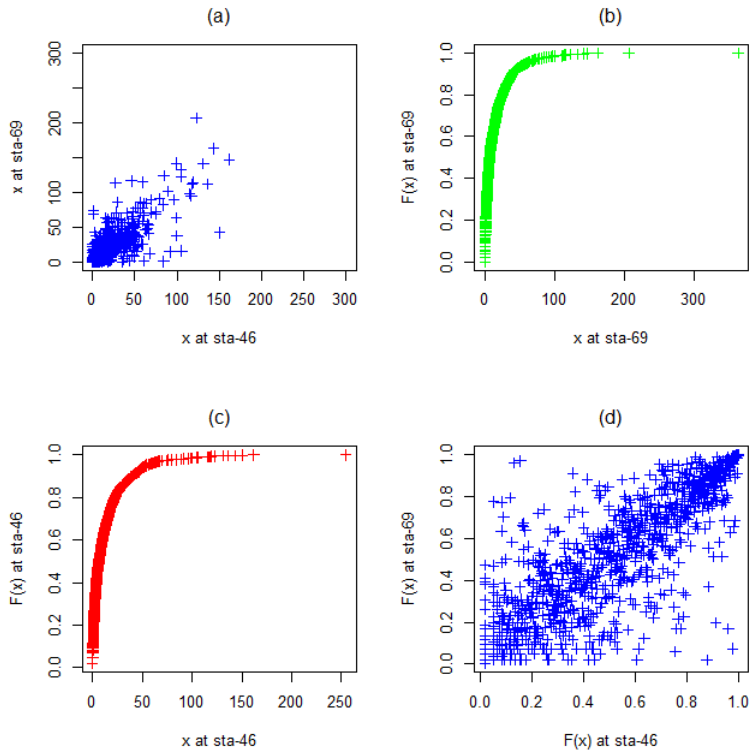



Figure 4.1: Empirical bivariate distributions in original and uniform domain between two precipitation stations (sta-46 and sta-69) for non-zero daily precipitation in the region of Singapore.

the uniform distribution by the empirical distribution function. In Figure 4.1.d shows the scatter plot of the marginal distribution in the uniform distribution (0,1). It can be seen that high values have a stronger dependence than low values because the point density is higher in the top right corner than in the lower left one. This is why the scatter plots with uniform marginal are more suitable to describe the dependence structure. Besides, the description of dependency is standardized, because measures of dependence, like the correlation coefficient, can now be calculated without the influence of marginal distribution.

## 4.3 The Spearman's rank correlation

Spearman's rank correlation (Spearman, 1904) is defined as a non-parametric measure of rank correlation for evaluating the degree of monotonic association between two independent variables. The Spearman's rank correlation between two variables is similar to the Pearson correlation except that it operates on the rank values of those two variables rather than the raw data (Gautheir, 2001). When those two variables have a similar rank, the Spearman's rank correlation between two variables will be high, and vice versa.

The dependence between pairs of points expressed by the rank-correlation function can be

illustrated as follows. Let $X_1$ and $X_2$ be continuous random variables whose copula is $C$. Then, the Spearman's rank correlation $\rho$ for $X_1$ and $X_2$ is given by Nelsen (2006) as a result of the bivariate central moment of order two.

$$\rho(x_1, x_2) = \frac{E(U_1, U_2) - E(U_1)E(U_1)}{\sqrt{var(U_1)}\sqrt{var(U_1)}} = \frac{E(U_1, U_2) - \frac{1}{4}}{\frac{1}{12}} = 12E[(U_1 - 0.5)(U_2 - 0.5)] \quad (4.7)$$

$$\begin{aligned} \rho(x_1, x_2) &= \frac{E(U_1, U_2) - \frac{1}{4}}{\frac{1}{12}} \\ &= 12E[U_1, U_2] - 3 \\ &= 12 \int_0^1 \int_0^1 c(u_1 u_2) du_1 du_2 - 3 \end{aligned} \quad (4.8)$$

where 0.5 is the univariate first-order moments (the mean values) and $\frac{1}{12}$ is the univariate second order central moments (variance) of the univariate uniform distribution. As a consequence, the Spearman's rank correlation for a pair of continuous random variables $X_1$ and $X_2$ is identical to the Pearson's product-moment correlation coefficient for the grades of $X_1$ and $X_2$; the random variables $u_1 = F(X_1)$ and $u_2 = F(X_2)$.

The Spearman's rank correlation in Equation 4.7 and Equation 4.8 can also be approximated using parameter of the bivariate Gaussian copula underlying assumption that the data follow the bivariate Gaussian copula.

Let $X_1$ and $X_2$ be continuous random variables whose copula is $C$ where $X_1 > 0$ and $X_2 > 0$. The bivariate Gaussian copula $C$ and its density $c$ can be formulated in the following equations.

$$H(X_1 > 0, X_2 > 0) = C\Big(F_1(x_1), F_2(x_2)\Big) \quad (4.9)$$

$$C(u_1, u_2) = \Phi^2\Big(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\Big) \quad (4.10)$$

$$c(u_1, u_2 \mid \theta) = \frac{\phi^2\Big(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta\Big)}{\phi_1\Big(\Phi^{-1}(u_1)\Big)\phi_2\Big(\Phi^{-1}(u_2)\Big)} \quad (4.11)$$

where $\theta$ is the linear correlation coefficient ($\rho$), $u_1$ and $u_2$ are transformed random variables of $X_1$ and $X_2$ in the uniform distribution ranging from 0-1, $\Phi^{-1}$ is the inverse of the cdf of the standard normal distribution $N(0, 1)$ and $\Phi^2$ in this case is the cdf of a bivariate normal distribution with mean 0 and covariance matrix $\Sigma$, $\phi^2$ is the PDF of a bivariate normal distribution with mean 0 and covariance matrix $\Sigma$.

For data sets containing zeros, it can be modeled by a mixed discrete and continuous distribution as introduced in Bárdossy and Pegram (2009, 2013) and Serinaldi (2009). A specific treatment of the zeros can be conducted by assuming that they behave like censored values

of a continuous distribution. It assumes that $P_0$ is the probability of dry events and the positive part is described with a continuous distribution function as introduced by Bárdossy and Pegram (2009).

Based on Equations 4.9 and 4.10, which are basically the formulas for describing bivariate copulas for both being different from zero $(X_1 > 0, X_2 > 0)$. However, if one marginal of pairs is either zero or non-zero $(X_1 > 0, X_2 \geq 0)$ or $(X_1 \geq 0, X_2 > 0)$ with the probability of zeros $P_0$, the bivariate copulas yield as follows.

$$H(X_1 = 0, X_2 > 0) = C\Big(P_0, F_2(x_2)\Big) \tag{4.12}$$

$$C(u_1 = P_0, u_2 = F_2(x_2)) = \Phi^2\Big(\Phi^{-1}(P_0), F_1^{-1}(u_2)\Big) \tag{4.13}$$

$$c(u_1, u_2 \mid \theta) = \int_0^{P_0} \frac{\phi^2\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta\big)}{\phi\big(\Phi^{-1}(u_2)\big)} du_1 \phi\big(\Phi^{-1}(u_2)\big) \tag{4.14}$$

If both pairs of marginal equal to zero $(X_1 = 0, X_2 = 0)$ and the probability of zeros for the marginals $X_1$ and $X_2$ are $P_{01}$ and $P_{02}$, respectively, then bivariate copulas can be formulated as follows.

$$H(X_1 = 0, X_2 = 0) = C\Big(P_{01}, P_{02}\Big) \tag{4.15}$$

$$C(u_1 = P_{01}, u_2 = P_{02}) = \Phi^2\Big(\Phi^{-1}(P_{01}), \Phi^{-1}(P_{02})\Big) \tag{4.16}$$

$$c(u_1, u_2 \mid \theta) = \int_0^{P_{01}} \int_0^{P_{02}} \frac{\phi^2\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta\big)}{\phi\big(\Phi^{-1}(u_1)\big)\phi\big(\Phi^{-1}(u_2)\big)} du_1 du_2 \tag{4.17}$$

## 4.4 Asymmetric dependence

Both the Pearson's correlation and the Spearman's rank correlation are capable of quantifying a linear association measure between two random variables with their assumptions and limitations as discussed in section 4.2. Indeed, the Spearman's rank correlation is suggested as a good alternative to substitute the Pearson's rank correlation, but in fact, it formally also depends upon only one correlation coefficient describing the entire dependence structure of the two random variables.

However, in some cases, quantification of the degree of association not on the whole distribution of multiple variables, but on the extreme values is more interesting. The dependence structure between the extremely high values of two random variables may be different from that of the extremely low values. In other words, the dependence between extreme high values might be stronger than mid values or even low values, or vice versa. The scatter plot of the empirical copula can still detect a possible clustering dependence structure between two random variables in a particular part, either upper or lower tail visually.

In the practical application of hydrology, precipitation extremes occurring simultaneously at multiple locations can lead to a severe impact on risk assessment and mitigation strategies. Estimation of the dependence structure in the occurrence of extreme events, both upper and lower tails, is therefore required. A precipitation model which is not capable of capturing the joint extreme events gives either underestimation or overestimation of design values.

The upper tail dependence (UTD) coefficient $\lambda_u$ and lower tail dependence (LTD) $\lambda_L$ (Joe, 1997) have been used widely to measure the mutual occurrence of extreme events in a variety of applications. However, in the spatial statistics the main interest is usually not at the very high extremes, but in the location of values that exceed certain thresholds (Bárdossy, 2006). Empirical investigations present that the UTD shows bad performances with a strong bias estimation (Serinaldi et al., 2014).

Instead of focusing on the tail dependencies, Bárdossy (2006) expressed the non symmetrical dependence by the ratio of the joint probability of not exceeding a quantile $u$ and exceeding the quantile $1-u$. This ratio expresses whether higher values have a stronger dependence than the low values $(A_1(u) > 1)$ or the opposite $(A_1(u) < 1)$, or if the dependence is the same as formulated below.

$$A_1(u) = \frac{2u - 1 + C(1-u, 1-u)}{C(u,u)} \tag{4.18}$$

Based on the combination of two third order moments, Bárdossy and Li (2008); Li (2010) described a different non-symmetrical measure which characterizes the difference in the dependence between high or low percentiles as given the following formula:

$$\begin{aligned} A_2(u_1,u_2) &= E[(U_1-0.5)(U_2-0.5)^2 + (U_1-0.5)^2(U_2-0.5)] \\ &= \int_0^1 \int_0^1 (u_1-0.5)(u_2-0.5)^2 + (u_1-0.5)^2(u_2-0.5)c(u_1,u_2)du_1du_2 \end{aligned} \tag{4.19}$$

$A_2$ is positive for a stronger dependence of high values and $A_2$ will be negative for a stronger dependence of low values. If percentiles above and below 0.5 exhibits a symmetric dependence, $A_2$ will be zero (Li, 2010) and (Guthke, 2013).

Both formulas above (Equations 4.18 and 4.19 ) are not applicable if datasets consist of zero-inflated data which frequently happens in precipitation fields, especially for shorter time scales (daily and sub-daily) as shown in a scatter plot of precipitation amount at two locations in Figure 4.2. On the left panel, a scatter plot with marginal distribution on the original domain in mm is shown, while on the right panel, the scatter plot is made on the copula domain. Random variable $X_1$ has the probability of zeros being 0.16, and random variable $X_2$ has the probability of zeros equals to 0.10. The red lines indicate the quantile threshold $q_0$ to distinguish frequencies of events between the upper-right corner and the lower-left corner.

In order to handle this problem, a new simple method considering zero precipitation modified based on Equation 4.20 is adopted. This method is just a comparison of the mass of empirical density copula between the upper-right part and lower-left part. To distinguish

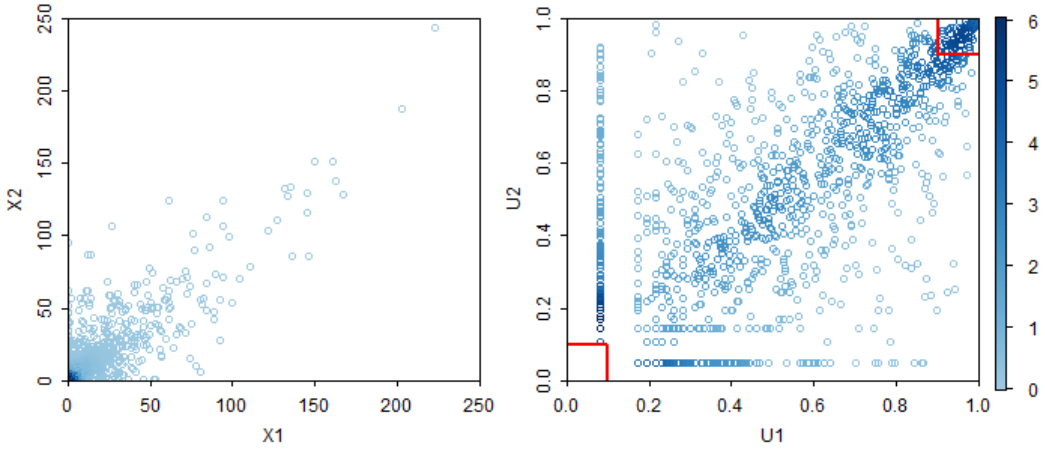



Figure 4.2: Scatter plot of empirical bivariate copula with zero-inflated data.

those parts, a threshold $q_0$ with different quantiles, such as 0.1, 0.2, 0.3, 0.4, and 0.5 can be adopted. The threshold $q_0$ must be bigger than the probability of dry events. The formula is illustrated below.

$$A_3(u_1, u_2) = \frac{\int_{1-q_0}^{1}\int_{1-q_0}^{1} c(u_1, u_2)du_1du_2 - \int_{0}^{q_0}\int_{0}^{q_0} c(u_1, u_2)du_1du_2}{\int_{1-q_0}^{1}\int_{1-q_0}^{1} c(u_1, u_2)du_1du_2 + \int_{0}^{q_0}\int_{0}^{q_0} c(u_1, u_2)du_1du_2} \tag{4.20}$$

where $q_0$ is a threshold to distinguish mass of empirical density copula between the upper-right part and lower-left part of bivariate copulas. $q_0$ must be bigger than the probability of zeros.

In order to demonstrate the asymmetry function $A_3$ to measure the non-symmetrical dependence structure, 500 Monte Carlo based stochastic simulations using three different families of copulas, namely, Gumbel, Gaussian, and Clayton copulas are made. The Gaussian copula exhibits the symmetric dependence structure. The Gumbel copula has more densities concentrated in the upper tails, and the Clayton copula is recognized as copula model with the heavy concentration of probability near (0,0) (Nelsen, 2006).

The bivariate Gumbel copula density can be written as follows.

$$c_(u_1, u_2) = \exp\left[-\left((-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\right)^{1/\theta}\right] \tag{4.21}$$

where $1 < \theta < \infty$ is the parameter controlling the dependence. Perfect dependence is obtained if $\theta \to \infty$, while $\theta \to 1$ implies independence.

The bivariate Clayton copula density can formulated as follows.

$$c_(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta} \tag{4.22}$$

where $0 < \theta < \infty$ is the parameter controlling the dependence. Perfect dependence is obtained if $\theta \to \infty$, while $\theta \to 0$ implies independence.

Each simulation is carried out for a variety of number of samples $N$, namely $N$= 30, 50, 100, 500, 1000, and 5000 pairs. Figure 4.3 shows an example of a scatter plot of 500 random variables simulated from three families of copulas consisting of a Gumbel copula with parameter $\theta = 3$ on the left panel, a Gaussian copula with parameter $\rho = 0.8$ on the middle panel, and a Clayton copula with parameter $\theta = 2$ on the right panel.

Every simulation, the asymmetry function $A_3$ is calculated using the formula 4.20 for different quantiles $q_0$ as the threshold, namely $q_0$= 0.1, 0.2, 0.3, 0.4, and 0.5. The basic idea is to calculate the difference of frequencies/densities between the upper tail and lower tail as marked with the red lines on Figure 4.3.

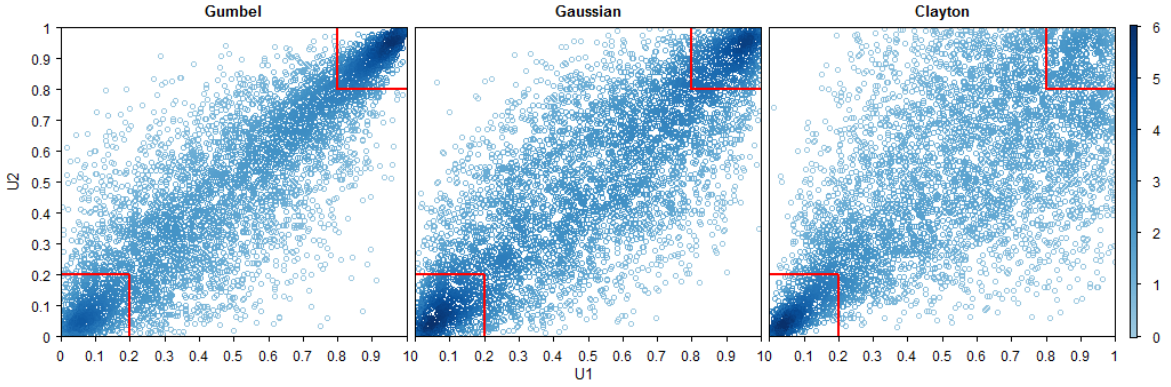



Figure 4.3: Scatter plot of simulated data based on a Gumbel copula having $\theta = 3$, a Gaussian copula with $\rho = 0.8$, and Clayton copula with $\theta = 2$ consisting of 500 pairs.

Figure 4.4 depicts the asymmetry function $A_3$ for a variety of the simulated sample of sizes $N$ using the three families of copulas above and different the quantile thresholds, namely $q_0$= 0.1, 0.2, 0.3, 0.4, and 0.5. The horizontal axis defines the quantile thresholds and the primary vertical axis represents the asymmetry function $A_3$. Every quantile threshold $q_0$, there is box-plot of the asymmetry function $A_3$ from 500 stochastic simulations created from the Monte Carlo approach which is available in R (R Core Team, 2013) by the copula package (Hofert et al., 2017).

One can see that the Gumbel copulas exhibit clearly the positive asymmetric dependence on the whole in particular for the quantile thresholds $q_0$ 0.1 and 0.2. The Gaussian copula present symmetric dependence consistently for all the quantile thresholds. The Clayton copula generally present the negative asymmetric dependence especially for the quantile thresholds $q_0$ 0.1, 0.2 and 0.3. Overall, each copula family exhibits its pattern more consistently in accordance with increasing samples of size.

## 4.5 Summary

- Copula is a joint distribution function with the univariate marginals in uniform distribution U(0,1).
- The Pearson's correlation has some drawbacks, such as being restricted to normality assumption, assuming a linear relationship, and being sensitive to outliers.
- The Spearman's rank correlation is one of the alternatives which could resolve these matters.
- The Spearman's rank correlation also offers a flexibility of incorporating zeros variables which are treated as latent variables underlying a continuous distribution function using the censored bivariate Gaussian copulas, for example.
- An asymmetric function measuring the degree of dependence structure between high and low values by incorporating zero precipitation amounts is introduced using the empirical bivariate copulas.

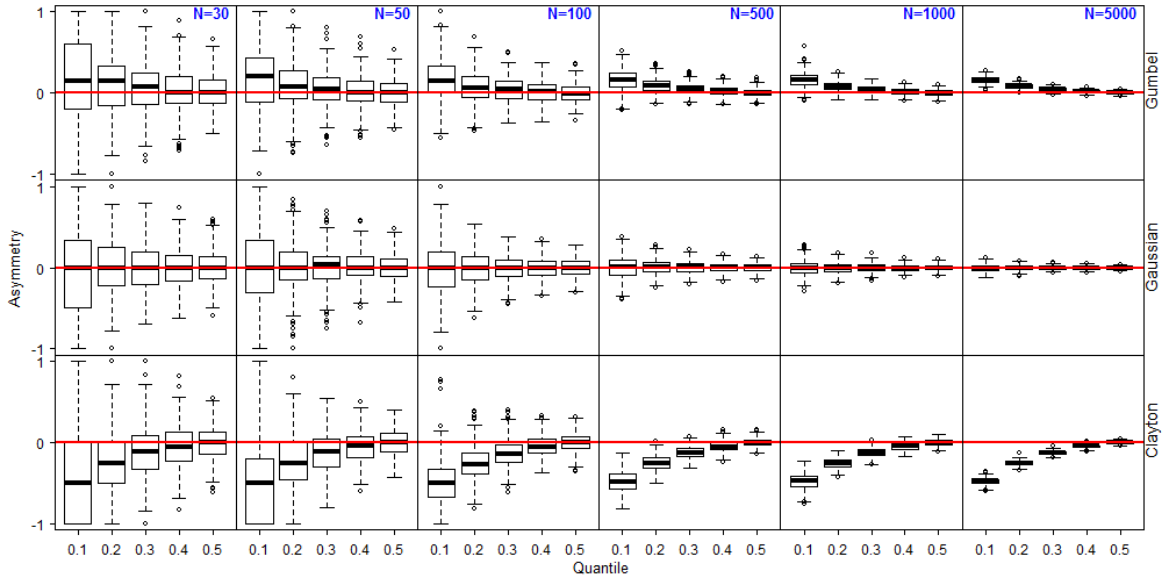



Figure 4.4: Simulated asymmetry function using Gumbel copula $\theta = 3$ (top row), Gaussian copula $\rho = 0.8$ (middle row), and Clayton copula $\theta = 2$ (bottom row) with 500 Monte Carlo simulations.

# 5 Investigating spatial rank correlation using censored bivariate Gaussian copulas

## 5.1 Introduction

Spatial modelling requires the fundamental assumption that observations in different locations are usually highly spatially correlated to each other in close distance and have a weaker spatial dependence in far distance. This assumption is then reflected to modelling the spatial correlation function. The spatial correlation coefficient is equal to 1 near at zero distance and then continuously decreases with increasing distance within the region of interest until a certain distance. The distance, where there is no longer a decrease of spatial dependence, is called correlation length. In many practical applications, all observations only within this correlation length and the nearest neighbours are considered in the estimation of the spatial model (Sen, 2009).

There are some spatial correlation functions in literature, such as, exponential, spherical, and Gaussian or other models, which are widely implemented in a variety of spatial analyses. Figure 5.1 shows examples of the theoretical spatial correlation functions for the exponential, spherical, and Gaussian model with a domain equal to the radius of the area. The exponential, spherical, and Gaussian spatial correlation functions can be formulated as shown in Equation 5.1, Equation 5.2, and Equation 5.3, respectively.

- Exponential

$$\rho(h) = \exp\left(\frac{-h}{a}\right) \tag{5.1}$$

- Spherical

$$\rho(h) = \begin{cases} 1 - \frac{3h}{2a} + \frac{h^3}{2a^3} & \text{if } h \leq a \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

- Gaussian

$$\rho(h) = \exp\left(\frac{-h^2}{a^2}\right) \tag{5.3}$$

The variable $a$ in the equations above represents the range parameter. The range parameter $a$ is obtained by fitting the model to the available data. For the Spherical spatial correlation function, the correlation length is the same as the range parameter since the correlation coefficient at the distance range is close to zero. For the exponential spatial correlation function,

however, the range parameter $a$ is not called the correlation length because the correlation coefficient at the distance range is 0.37 or $\exp(-1)$. The exponential spatial correlation function is used for this study because it is very simple to implement for a large number of precipitation realizations. The correlation length is, therefore, three times the range parameter because correlation coefficient at the distance $3\times$ range is close to zero or $\exp(-3)$.

The spatial correlation functions in this chapter are calculated using Spearman's rank correlation coefficient on the basis of empirical bivariate copulas. The empirical bivariate copulas are constructed by using two different approaches, namely, the geostatistical approach and the pairwise approach, which are explained in the subsequent sections. The concept of geostatistical approach has been widely applied to the precipitation fields for solving the spatial interpolation issues, especially used for the rainfall-runoff processes where both precipitation and runoff require the same time event. The pair-wise approach containing simultaneous temporal observations is useful for infilling missing precipitation observations. The exponential spatial correlation function is chosen to characterize the Spearman's rank correlations over distances of interest.

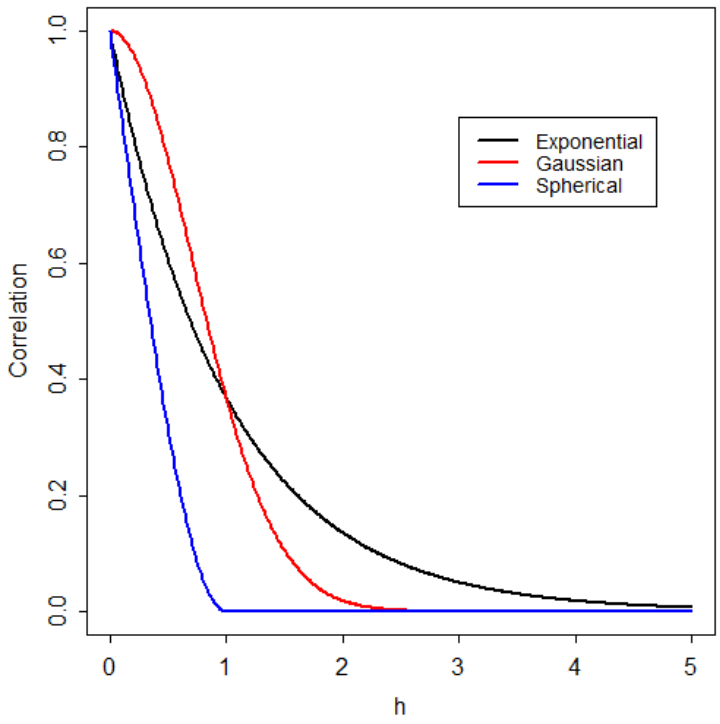



Figure 5.1: The spatial correlation functions as a function of distance $h$ with range $a = 1$ using different models (exponential, Gaussian, and Spherical).

## 5.2 Maximum likelihood estimation of the censored bivariate Gaussian copulas

The Spearman's rank correlation is calculated using the parameter of the Pearson correlation coefficient of the censored bivariate Gaussian copula as presented in chapter 4. The Pearson correlation in the copula domain is basically same as the Spearman's rank correlation. The Spearman's rank correlation can also be estimated using the correlation parameter of the bivariate Gaussian copula underlying assumption that the data follow the bivariate Gaussian copula. When the data also contain zeros and the zeros are treated as latent variables of a continuous copula distribution, the bivariate Gaussian copula is the choice of approach for estimating Spearman's rank correlation.

The Spearman's rank correlation can be estimated with censored bivariate copulas by numerically maximizing the likelihood function. The likelihood function consists of three parts, namely, both marginals equal to non-zeros (Equation 4.11), both marginals equal to either zero or non-zero (Equation 4.14) and both marginals equal to zero (Equation 4.17).

The likelihood function of the censored bivariate copula above for estimating the parameter $\theta$ can be calculated using the following formula (Bárdossy, 2011). Here, the parameter $\theta$ represents the Spearman's rank correlation.

$$
\begin{aligned}
L(x_1, x_2 \mid \theta) = & \prod_{(x_1,x_2)\in I_1} \frac{\phi^2\big(\Phi^{-1}(u_1 = F_1(x_1)), \Phi^{-1}(u_2 = F_2(x_2)) \mid \theta\big)}{\phi\big(\Phi^{-1}(u_1 = F_1(x_1))\big)\phi\big(\Phi^{-1}(u_2 = F_2(x_2))\big)} \times \\
& \prod_{(x_1,x_2)\in I_2} \int_{u_1=0}^{u_1=P_{01}} \int_{u_2=0}^{u_2=P_{02}} \frac{\phi^2\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2) \mid \theta\big)}{\phi\big(\Phi^{-1}(u_1)\big)\phi\big(\Phi^{-1}(u_2)\big)} du_1 du_2 \times \\
& \prod_{(x_1,x_2)\in I_3} \int_{u_1=0}^{u_1=P_{01}} \frac{\phi^2\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2 = F_2(x_2)) \mid \theta\big)}{\phi\big(\Phi^{-1}(u_2 = F_2(x_2))\big)} du_1 \phi\big(\Phi^{-1}(u_2)\big)
\end{aligned}
\tag{5.4}
$$

where $\phi^2$ is the bivariate normal density with mean 0 and covariance matrix $\Sigma$, $\Phi^{-1}$ is the inverse of the CDF of the standard normal distribution $N(0,1)$ and $\phi$ is the density of the standard normal distribution $N(0,1)$.

Here, the set $I_1$ contains pairs of gauges with both variables equal to non-zero precipitation $(x_1 > 0, x_2 > 0)$. The set $I_2$ consists of zero pairs $(x_1 = 0, x_2 = 0)$ and $P_{01}$ and $P_{02}$ are the probability of zeros for the marginal distributions. In the set $I_3$, pairs are listed which consist of non-zero and zero precipitation amounts $(x_1 > 0, x_2 = 0)$ or $(x_1 = 0, x_2 > 0)$. The variables of $u_1$ and $u_2$ are uniformly distributed and range from 0 to 1. They are derived from the rank transformation of the random variable if $x_1 \neq 0$ and $x_2 \neq 0$, respectively.

## 5.3 Geostatistical context

### 5.3.1 Approach and methodology

#### 5.3.1.1 Isotropic assumption

As mentioned above, spatially distributed data values at different locations are frequently not completely random, but usually exhibit some form of spatial correlation structure, where close points in space tend to be highly correlated. A phenomenon spreads in space and exhibits a certain spatial correlation structure is called as term "regionalized". If $f(x)$ is the value at the point $x$ of a characteristic $f$ of this phenomenon then $f(x)$ is called "regionalized variable". Geostatistics can be defined as "the application of probabilistic methods to regionalized variables" which designates a function displayed in real space in order to exhibit a structure of spatial correlation (Matheron, 1971).

In the context of geostatistics, the regionalized variables are assumed to be a single realization of a random variable, which is characterized by its distribution function. The distribution function that completely describes a spatially distributed variable can thus be defined by assigning an $n$-dimensional probability distribution $F_n$ to the set of points in the domain of interest. Using only one realization for counting frequencies to find a multivariate probability distribution $F$ that fits to the data is not possible without an additional assumption about the spatial process, that is the spatial strong stationarity (Guthke, 2013).

Under the assumption of the spatial strong stationarity, the hypothesis that the random function is dependent on the location is relaxed towards a dependence of the spatial configuration of the regionalized variables. In other words, the random function remains the same if a certain configuration of points is only shifted by a length of separating distance $h$ (this also called translation invariant). The bivariate distribution function can now be inferred by counting frequencies of values in the same configuration (Guthke, 2013) and the spatial correlation model finally can be estimated.

The spatial correlations are calculated under the isotropic assumption. The isotropic assumption means that the spatial correlation structure is the same for all directions. The isotropic spatial correlation depends only on the length of the separating distance between two points and not on their relative directions. The isotropic assumption is commonly made in any practical spatial application. This is because the interpretation of the correlation behavior becomes simple and the estimation process of the correlation coefficient is also easy (Maity and Sherman, 2012).

The spatial correlations derived in the context of spatial geostatistics require regionalization in order to construct an empirical bivariate distribution. The spatial Spearman's rank correlations are calculated based on underlying bivariate normal distribution on the unit domain U(0,1). This censored bivariate Gaussian copula is fitted to the empirical bivariate copulas.

The empirical bivariate copulas for the spatial precipitation observations $z$ at any given location $s$ of $n$ gauges $z(s_1), ..., z(s_n)$ can be constructed using the procedures given by Bárdossy (2006). The procedure is conducted for each selected time interval of precipitation values

separately due to the assumption of a single realization of a spatial random variable as follows.

1. Precipitation data from all spatial observation points at any given time step are transformed into the uniform distribution $u(0,1)$ using the empirical marginal distribution function $F_z(z(s_i))$ or the normalized ranks from the observation points as follows.

$$u = \frac{n - R(z(s_i)) - \frac{1}{2}}{n} \tag{5.5}$$

   where $z(s_i)$ is a univariate random variable at location $s_i$, $n$ is the number of data points, $R(z(S_i))$ is the rank of $z(s_i)$ in the set $z(s_1), ..., z(s_n)$, $z_{s_1}$ is the maximum value, and $z_{s_n}$ is the minimum value.

2. For any given length of the separating distance $h$, all gauge stations are grouped with similar lags $h$ and the set of pairs $S(h)$ of gauge stations can be created. The set of pairs $S(h)$ consists of the empirical distribution function values of the precipitation variables of any pair at locations separated by the length of distance $h$. For the isotropic assumption, the length of distance $h$ does not depend on any direction, no matter whether the two gauges at separating length of distance $h$ are located in the north-south direction, the west-east direction, or any other directions. The set of pairs $S(h)$ are formulated as follows.

$$S(h) = c\Big(F_z\big(z(s)\big), F_z\big(z(s+h)\big)\Big) \tag{5.6}$$

   Again here, $S(h)$ is a set of points in the unit square. Note that $S(h)$ is by definition symmetrical regarding the major axis $u_1 = u_2$ of the unit square, namely, if $(u_1, u_2) \in S(h)$, then $(u_2, u_1) \in S(h)$.

3. From the set of pairs $S(h)$, the spatial copula function $C_S$ is implemented as a function of the length of distance $h$ extended from Equation 4.4. For any two selected quantiles $u_1$ and $u_2$, Equation 4.4 becomes as follows.

$$\begin{aligned} C_s(u_1, u_2; h) &= Prob\Big(F_z\big(Z(s)\big) \leq u_1, F_z\big(Z(s+h)\big) \leq u_1\Big) \\ &= C\Big(u_1 = F_z\big(Z(s)\big), u_2 = F_z\big(Z(s+h)\big)\Big) \end{aligned} \tag{5.7}$$

   Here, $F_z$ is the marginal distribution of the variable $Z$ and is supposed to be the same for each location $s$. Any two locations separated by the length of distance $h$ are related to each other by the bivariate distribution function whose dependence is described by the bivariate copula. As with variograms, it is assumed that this copula does not depend on the location $x$ but only on the separating distance $h$.

4. The unit square is divided into $k$ x $k$ grids. The empirical copula density can then be constructed by calculating the empirical frequencies on each grid resulting in two dimensional histograms of the marginals $u_1$ and $u_2$.

Since the empirical bivariate copula density was calculated, the Spearman's rank correlation coefficient can be estimated via fitting of the censored bivariate Gaussian copulas to the empirical bivariate copulas by maximizing numerically the likelihood function as given by Equation 5.4. The fitted parameter of the censored Gaussian copulas is Spearman's rank correlation. The detailed procedure of estimating Spearman's rank correlation can be described as follows.

1. The uniformly distributed values $u_1$ and $u_2$ are transformed into the normal domain using the inverse of the standard normal distribution.
2. The likelihood function of the censored bivariate normal distribution, as shown in Equation 5.4, is maximized numerically to get the parameter $\theta$.

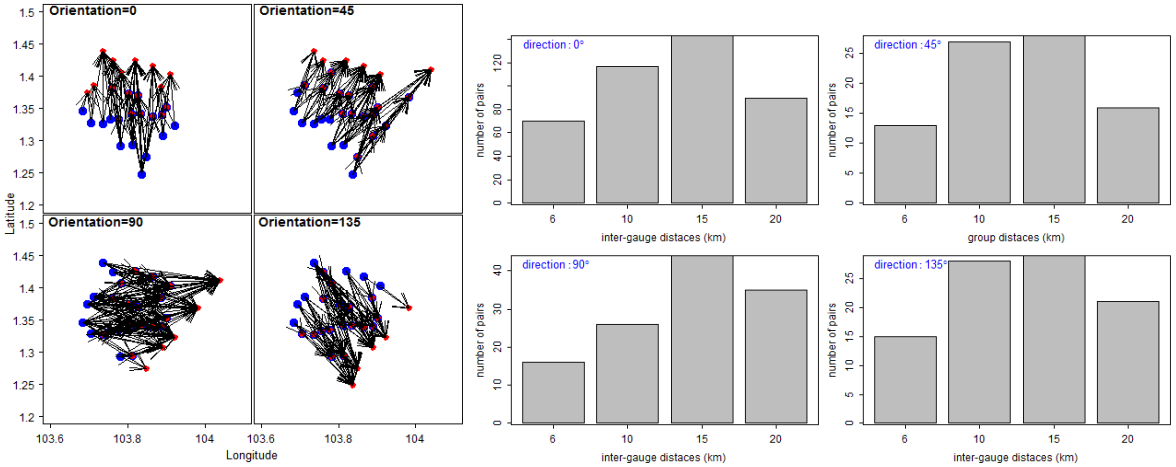
#### 5.3.1.2 Anisotropic assumption

The assumption of isotropy is frequently applied to describe the spatial correlation structure because the interpretation of the correlation behavior is simple and the process of estimation is straightforward as mentioned. The assumption of isotropy, however, can give a bias estimation when not appropriate to the data. If the spatial correlation structure is not isotropic, then the anisotropic assumption is suggested. The anisotropic assumption means that the spatial correlation between any two observations depends not only on the distances between those sites but also on their relative directions (Maity and Sherman, 2012), thus different spatial correlation structures are calculated for the different directions.

In general, some procedures of estimating of the spatial correlation function under the anisotropic assumption in the case of geostatistical approach is similar to the isotropic assumption above. For example, for each selected time interval, the set of pairs $S(h)$ which consists of the empirical distribution function values of the interested variables at different locations separated by the length of distance vector $h$ should be determined. However, unlike the isotropic assumption where all directions are considered to be the same for estimating the spatial correlation function, the anisotropic approach in this study proposes four different orientations that can be taken into account for estimating the spatial correlation function, for example, $0^o$, $45^o$, $90^o$ and $135^o$ from the North direction with a tolerance of $22.5^o$. By definition, the spatial correlations at those directions are the same as the directions $180^o$, $225^o$, $270^o$, and $315^o$ from the direction of north, respectively, due to the symmetrical condition. For each selected time interval, a different correlation function is estimated for each direction. In principle, for each direction, this anisotropic approach employs the same procedure for estimating the spatial correlation function as the isotropic approach above.

The available number of pairs of gauge stations at any given distance vector $h$ is a crucial issue for the geostatistical approach since the inferential processes for this approach use the data sets collected from those pairs. Just for example in the region of Singapore, all possible combinations of gauge pairs for the directions of $0^o$, $45^o$, $90^o$ and $135^o$ or directions of $180^o$, $225^o$, $270^o$, and $315^o$ from the directions of North are shown in Figure 5.2a, and the total number of all possible pairs for each direction at any given distance vector $h$ is presented in Figure 5.2b.

The configuration and number of pairs in each direction at any given distance vector $h$ can be different from the isotropic approach. For example, in the direction of $0^o$ from the North, the total number of pairs within the distance groups of 6 km, 10 km, 15 km, and 20 km is 70, 117, 143, and 90, respectively. Those numbers are sufficient to be used for estimating the spatial correlation functions. For the directions of $45^o$ and $90^o$, the total number of pairs within the distance classes of 6 km, 10 km, 15 km, and 20 km is 13, 27, 28, 16 for the direction of $45^o$, respectively, and 16, 26, 44, and 35 for the direction of $90^o$, respectively. Then, the total number of possible pairs for the direction of $135^o$ within inter-gauge distances of 6 km, 10 km, 15 km, and 20 km is 15, 28, 29, 21, respectively. The total number of all possible gauge pairs at certain directions in Singapore might not be sufficient for investigating the correlation behavior accurately. Nevertheless, this is an alternative to the isotropic approach for the comparison of correlation behaviors.




(a) All possible combinations of pairs

(b) Number of all possible combinations of pairs

Figure 5.2: All possible combinations of gauge pairs with different orientations, namely, $0^o$, $45^o$, $90^o$ and $135^o$ from the direction of North in the Singapore region.

### 5.3.2 Application in Singapore

A spatial correlation analysis for many realizations of precipitation amounts as well as different time scales is explored in this study on the basis of the isotropic and the anisotropic assumption.

**Effect of spatial scales** Figure 5.3 shows multiple box plots of the spatial correlation functions within the different distance groups (namely, 5 km, 10 km, and until 25 km) and various timescales from hourly to monthly for a given season (namely, DJF, MAM, JJA, or SON). The influence of inter-gauge distances on Spearman's rank correlation is clearly seen as expected from literature. Overall, the correlation functions decline with increasing distances and can be fitted reasonably using simple theoretical negative exponential correlation functions which are marked with the red lines as shown in Figure 5.3. Furthermore, this

trend applies not only to one specific time scale or any season but also to all time scales and all seasons. This guarantees that the established concept of the theoretical spatial correlation function is still reasonably applied to the real applications of precipitation fields in the region of Singapore.

However, there is some evidence that the correlation functions of some precipitation occurrences do not decrease according to the increase of inter-gauge distances. On the contrary, the correlation functions increase in accordance with increasing distances. This phenomenon mainly occurs for the higher temporal resolutions, such as daily or even higher ones because the correlations exhibit higher variabilities at the smaller time scales. For example, the correlation functions at the hourly scale are more variable than those at the monthly scale. Thus, there is still a chance that a practical application will disagree with the established theory, despite the too small probability. This might also be caused by large scale features, which could not be simply incorporated, resulting in higher correlations for the larger distances.

In addition, the correlation lengths, which are derived from the fitted negative exponential correlation functions, are very small only roughly ranging from 12 to 15 km at all different temporal scales and across different seasons. The correlation length can be defined as the maximum distance between points where the spatial relationships between those points still exist. Using the negative exponential function, the correlation length where the spatial correlation coefficient is $exp(-3) =\sim 0.05$ can be practically calculated from the tipple of the range parameters of the negative exponential functions which roughly range from 4 to 5 km for the case of Singapore.

**Effect of time scales** In contrast to the effect of distances between points at different locations on the spatial rank correlations, which presents an inversely proportional relationship, the increases of time aggregations commonly are directly proportional to the spatial rank correlations of precipitation occurrences. However, the increasing trends of the spatial rank correlations with the time scales are generally very low, with a small slope as depicted in Figure 5.4. This general characteristic is more obviously seen, especially at the distance group of 5 km. The increasing trends of the spatial rank correlation in temporal aggregation become negligible beyond the distance of 5 km. For example, even though there is an increase of the correlation coefficient with increasing time intervals, especially in comparison the sub-daily time scales to the super-daily timescales at the distance classes above 10 km, it is very small and almost negligible.

The spatial rank correlations of precipitation occurrences generally become less variable over time series observations at the higher time scales, as mentioned previously. This overall trend applies systematically for all different temporal aggregations from an hourly scale to a monthly scale. Interestingly, the lowest degrees of the variability of the spatial rank correlations arise at the inter-gauge distance of 15 km, as shown clearly in Figure 5.4. This means that the spatial rank correlation exhibits higher fluctuations at the inter-gauge distances below 15 km and tend to be less variable when approaching the distance of 15 km. The rank correlation coefficient will turn into the higher uncertainty again at the inter-gauge distances above 15 km. The distance of 15 km is a reasonable distance for the lowest degree

of uncertainty of the correlation function because the correlation length of the distance of 15 km is approximated to be $\sim$0. In other words, the correlation becomes nearly zero at the inter-gauge distance of 15 km which is supported evidently by most precipitation events resulting in less variability. The correlation will exhibit higher uncertainty again at the inter-gauge distance above 15 km because it presents a negative relationship between the correlation and the inter-gauge distance. It is clearly seen in Figure 5.4 where the empirical correlations at the group distance 20 and 25 km are negative, in general.

The findings fairly differ from the study conducted by Mandapaka and Qin (2013) reporting that the correlation coefficients (here is the traditional Pearson correlation) increase significantly with the increase of time scales. But, notice that there are some different fundamental approaches from the study that was carried out by Mandapaka and Qin (2013). They used a pair-wise approach consisting of simultaneous time series observations for creating the bivariate distributions. In addition, the conventional Pearson correlation is adopted to quantify the spatial dependence structure of the precipitation fields. This pair-wise approach will be discussed separately in section 5.4, but instead of using the Pearson correlation that Mandapaka and Qin (2013) implemented, the Spearman rank correlation will be applied in this study on the basis of the empirical bivariate copulas using the censored bivariate Gaussian copulas.

**Effect of seasons** Investigations of the spatial rank correlations are conducted in each different season separately to detect whether the spatial rank correlation behavior is also affected by seasonal periods. Based on the result of an empirical investigation, the estimations of the spatial rank correlation considering different seasons generally do not produce significant influences on the behaviors of the spatial correlation function. This is due to the small values of the correlation lengths, which only roughly range from 12-15 km, and the correlations also present high fluctuation over time and space on the whole. Consequently, the different values are too small compared to the spatial resolutions for distinguishing the spatial correlation behaviors in different seasons.

At the hourly scale, for instance, despite the fact there is a systematic pattern that the correlation length in DJF and SON are slightly more consistent than the ones in the months of MAM and JJA, but it is quite a minor difference if compared to the spatial resolutions; around 12.9 km and 13.2 km for the DJF and SON seasons, respectively, and 12.8 km and 12.8 km for the DJF and SON seasons, respectively. At the 12-hour temporal scale for an another example, the correlation distances for both DJF and SON roughly range between 13.8 km and 14 km, whereas the correlation lengths in the others seasons, the MAM and JJA season, are about 13.2 km and 13.4 km, respectively.

Those similar patterns above also apply to the other different temporal scales; considering different seasons for estimating the spatial rank correlation function does not change the general correlation characteristics, but only reveals a slightly different magnitude of the spatial correlation lengths. There is no significant difference of the spatial correlation levels between different seasons, this was also confirmed by the previous study carried out by Beck et al. (2015); Mandapaka and Qin (2013) stating that the statistical differences between early morning and afternoon precipitation are much stronger than the differences between

the seasons, which is typical for equatorial regions with a diurnal cycle in precipitation occurrences.

**Effect of anisotropic assumption** In order to investigate the effects of different directions of inter-gauge distances of any two points at different locations on the Spearman's rank correlation functions, the direction of inter-gauge distances are classified into four groups, namely, the direction of $0^o$, $45^o$, $90^o$ and $135^o$ calculated from the direction of North. For each direction, the distances between gauge stations are then grouped into several classes, namely, 0-6 km, 6-10 km, 10-15 km, and 15-20 km considering a minimum number of gauge pairs. The Spearman's rank correlation calculated for each direction of $0^o$, $45^o$, $90^o$ and $135^o$ are averaged over all events at any given vector of distances yielding a single realization of observation points over spatial domain. The values are then projected on both axes. Projections to the axis of E-W direction are conducted by multiplying between the vector of distance between two gauge points of interest by $\sin(\theta)$, where $\theta$ denotes angle between two gauges and the North direction. In addition, the Spearman's rank correlations are also projected to the axis of N-S direction using the distance between two gauge points of interest multiplied by $\cos(\theta)$. The projected correlations to both axes are then interpolated over the whole spatial domain of interest, resulting in isolines of the correlation function.

Figure 5.5 shows multiple plots of isolines of the correlation function in the two-dimensional spaces of inter-gauge distances using the geostatistical approach at different time scales from hourly scale to monthly scale in the region of Singapore. The figure attempts to describe the characteristics of the anisotropic spatial correlation of precipitation. Similar to the isotropic approach, the rank correlations are generally quite low, with the correlation length around 15 km, similarly as when estimated by using isotropic approach. Overall, the anisotropic correlations are independent of the different temporal scales and different seasons. It means that estimating a correlation function using different time scales and different seasons will yield a similar result in general.

However, there is a slightly different characteristic of correlation functions in terms of direction, in particular within a small radius. Overall, the correlations within the North-East to the South-West direction tends to be slightly higher than those in the northwest to the southeast direction at the same distance of interest. This is a bit different from Beck et al. (2015) because of the use of different approaches. When using a small number of gauge pairs, this finding might also be questionable, and further investigation with more pairs needs to be conducted.

For further practical applications, the isotropic approach might be still quite reasonable because the effect of anisotropy on the spatial correlation only exists in the condition where the inter-gauge distances are very short, in particular smaller than 12 km, and, consequently, the anisotropic correlation will be very low beyond this range (Beck et al., 2015). In addition, the total number of gauge stations in this region is basically not sufficient to capture the spatial anisotropy correlation using a regionalization approach for the smaller distances.

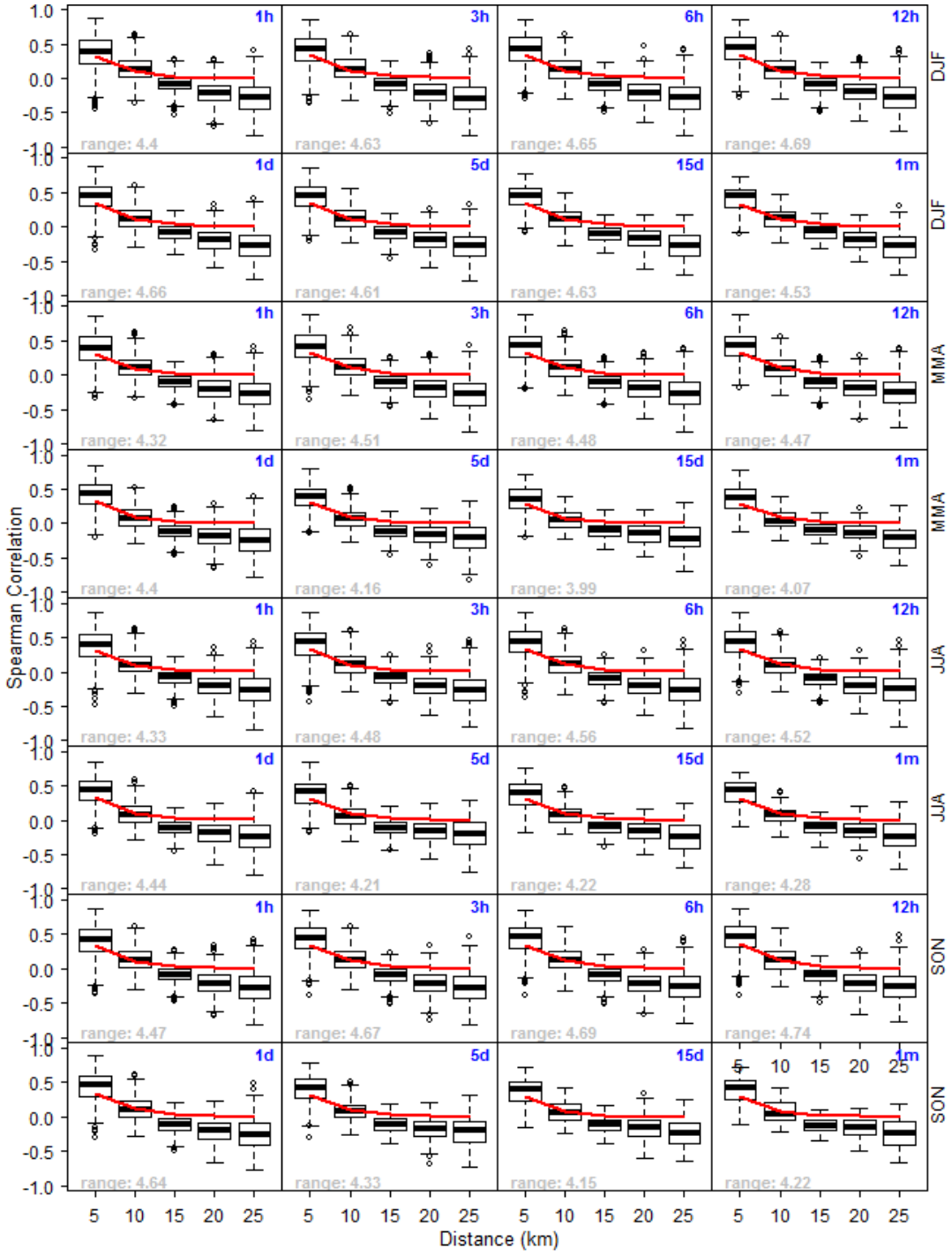



Figure 5.3: The Spearman's rank correlation functions over distances using the geostatistical approach in the region of Singapore. The vertical axes represent the rank correlation. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). The horizontal axes represent distances (from 5km to 25 km). Each panel represents time scale (hourly (1h) to monthly (1m).

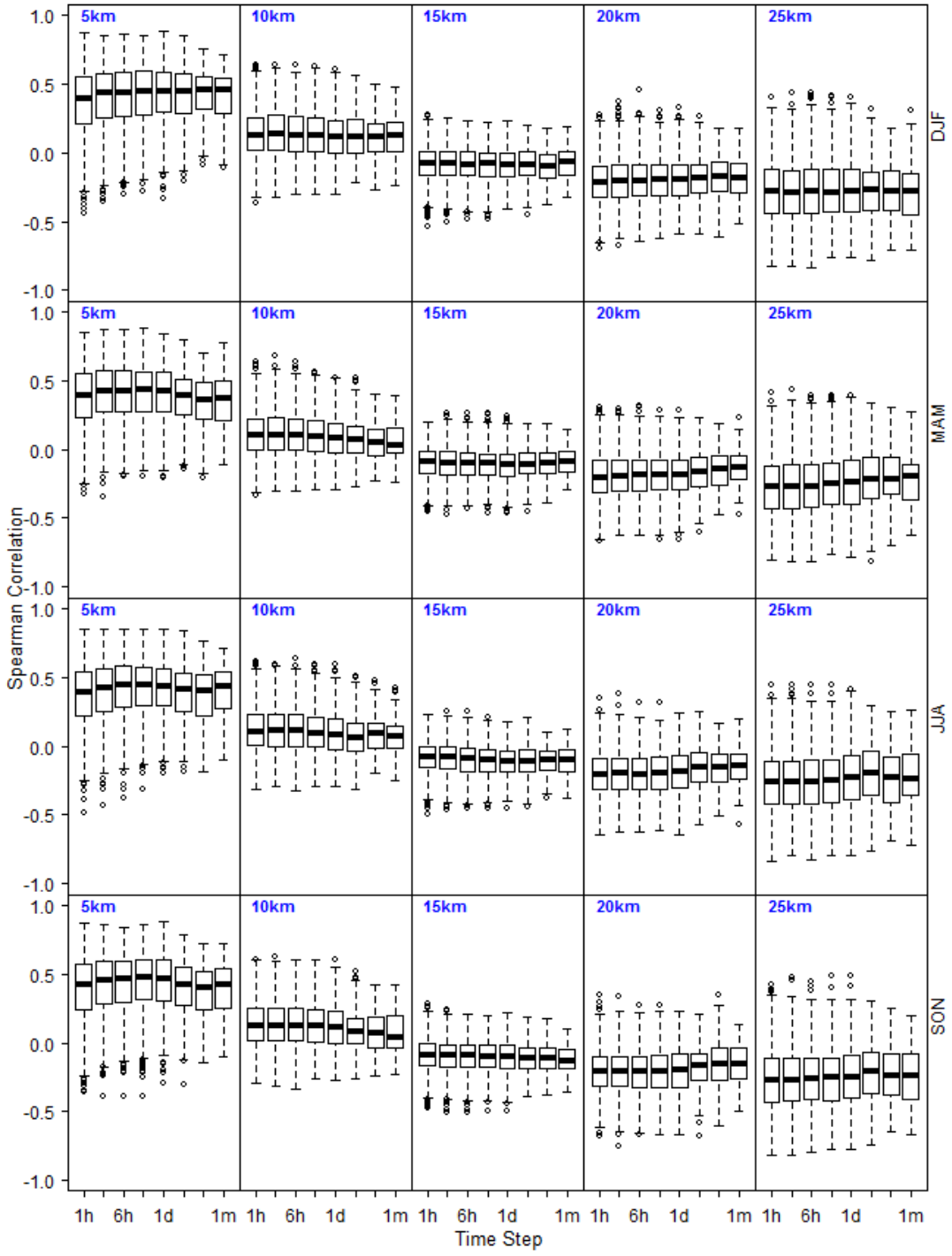



Figure 5.4: The Spearman's rank correlation functions over time scales using the geostatistical approach in the region of Singapore. The vertical axes represent the rank correlation. The horizontal axes represent a variety of time scales from hourly to monthly. Panel from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent distances (5 km to 25 km).

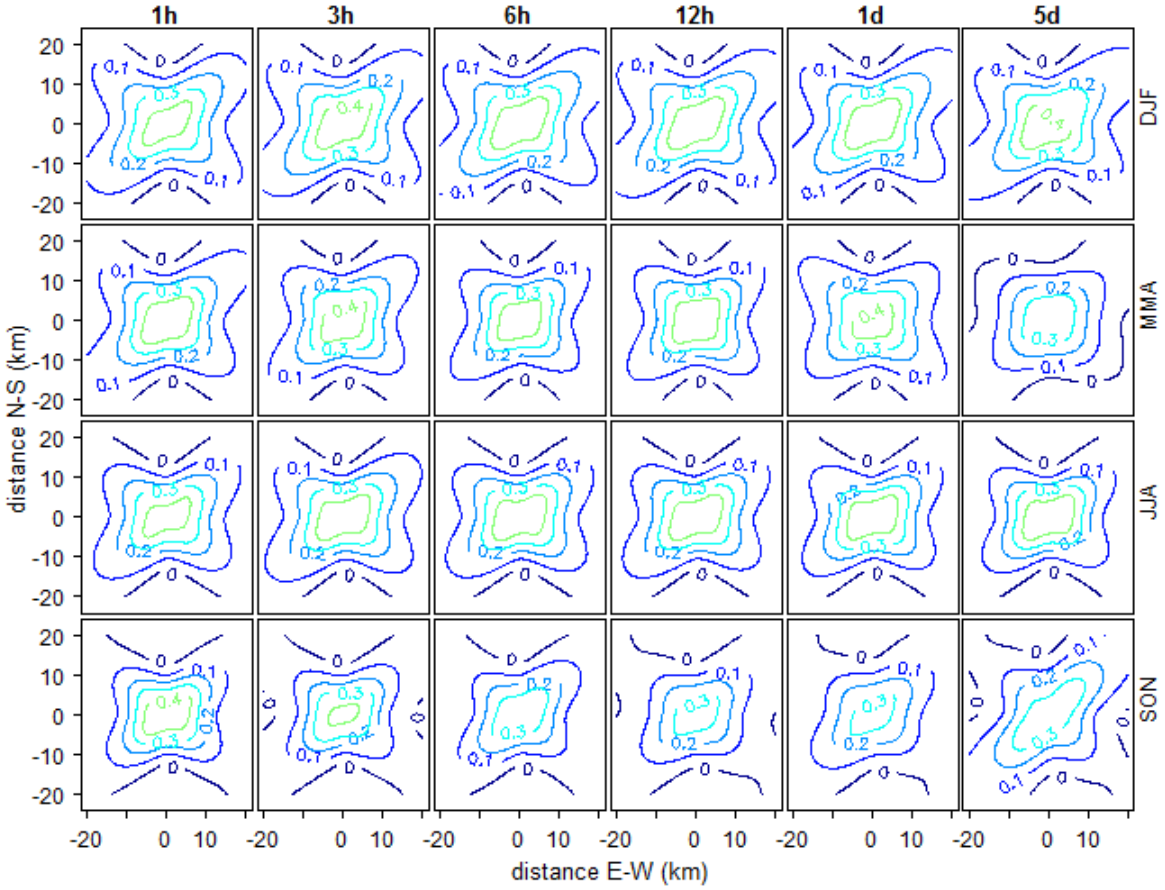



Figure 5.5: Isocorrelation lines in the two-dimensional spaces of distances using the geostatistical approach in Singapore. The vertical axes represent the lag distances in North and South direction in km. The horizontal axes stand for the lag distances in East-West direction. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right show the different time scales from hourly to 5-days.

### 5.3.3 Application in Bavaria and Baden-Württemberg

Empirical investigation results of the spatial correlation functions in the region of Bavaria and Baden-Württemberg are reported in Figure 5.6 and Figure 5.7. Both figures present the Spearman's rank correlations in the region of Bavaria and Baden-Württemberg at a variety of time scales from daily to monthly scale and also in the four different seasons (namely, DJF, MAM, JJA, and SON) based on the geostatistical approach. Figure 5.6 describes in detail the effects of the various distances on the spatial rank correlation functions, while Figure 5.7 stresses the impact of different time scales on the spatial rank correlation functions.

**Effect of spatial scales** Overall, it can be clearly seen that, despite only a few fluctuations in comparison to Singapore, the spatial Spearman's rank correlations of precipitation occurrences in both regions (Bavaria and Baden-Württemberg) decrease smoothly over the inter-gauge distances at the different locations and also fit reasonably using the negative exponential correlation functions which are marked with the red lines as shown in Figure 5.6. In contrast to Singapore with correlation lengths of $\sim$15 km, on the whole, the spatial correlation lengths in both Bavaria and Baden-Württemberg are significantly roughly higher by more than 60 km for all cases.

However, there is a strong empirical evidence that the spatial rank correlation function for Baden-Württemberg exhibits different patterns than those correlations for Bavaria. The correlation lengths in Baden-Württemberg roughly range from 60 to 84 km, whereas the correlation lengths for Bavaria roughly vary between 105 and 294 km due to the different correlation recession rate obtained from the fitted negative exponential distribution. This might be due to the different characteristics of the topography in both regions.

On one hand, the topographic features in Bavaria are relatively more structured varying from flat terrain in the northern part to the high mountains in the southern region. On the other hand, the topographic features in Baden-Württemberg are characterized as mountainous regions with a high variability of elevations spread out over this region. As a result, the spatial Spearman's rank correlations of precipitation occurrences in Baden-Württemberg also exhibit higher fluctuations over distances. This indicates that the high variability of the correlations spreads over the region, which frequently occurs at the lower time scales, for example at the daily scale as shown in Figure 5.6.

In addition, a lot of outliers are also detected in Baden-Württemberg indicating that many precipitation events in some locations are quite rare from many other places. Those outliers are mainly detected at the inter-gauge distance of 5 km and the daily scale. Some outliers are also detected for the higher time scales and the greater distances but only a few cases. In general, outliers might be due to incorrect measurements or coming from a different population where it represents a rare event.

Outliers here are, therefore, simply defined as observations that are numerically distant from the rest of the data; a data point located outside the fences (whiskers) of the box-plot, for example, observations that fall outside by 1.5 times the interquartile (IQR) range above the upper quartile (the third quartile, $Q_3$) and below the lower quartile (the first quartile, $Q_1$).

IQR is defined as the difference between the third quartile and the first quartile $Q_3 - Q_1$.

Many outliers arise at the distance of 5 km and particularly in the daily scale because many precipitation events in those cases exhibit very low positive relationships or even negative correlations. This is then extremely contradictory with other majoring precipitation events which present the higher rank correlations at the close distances. This might be also due to the fact that the number of pairs in this group is not sufficient to capture this phenomenon in comparison to the further distance groups, whereas the variability of topography in Baden-Württemberg exhibits a significantly pronounced, especially in the southern part, or it might be due to local spatial extent, for instance, an effect of orographic precipitation.

**Effect of time scales** The effect of temporal scales on the Spearman's rank correlation is illustrated explicitly in Figure 5.7. On the whole, there is a slight increase in the spatial Spearman's rank correlation as a result of increasing the temporal aggregations. This trend is more obviously detected in Bavaria rather than those which are located in Baden-Württemberg.

In Bavaria, the pattern that the spatial rank correlations increase with the time scales seems to be more pronounced at the smaller distances, for example at distances from 5 km to 15 km. Similar pattern apply for all different seasons, no matter which of the seasons, DJF, MAM, JJA, or SON. For the months of MAM and JJA, this pattern is also applicable to the distance of 50 km. For the seasons of SON and DJF, the Spearman's rank correlations increase with increasing time scales, from daily to the 10-15 days scales. For the monthly scale, the correlations, then, either remain stable or decreases slightly at a distance above 15 km.

Increasing temporal aggregations from a daily to a monthly scale also cause the lower variability of the spatial Spearman's rank correlations. This trend is valid for all distances ranging from 5 km to 50 km in the region of Bavaria as shown in Figure 5.7. Outliers of the spatial rank correlation also decrease with the increasing time scales as well as increasing distances. At the distances from 5 to 15 km, for example, the Spearman's rank correlation consists of a few outliers which are more concentrated on the distance of 5 km. At the distances of more than 15 km, there are very few or almost negligible outliers. Notice that, all the outliers are more concentrated below the lower boundary.

In contrast to the region of Bavaria, discussion about the impact of time aggregations on the characteristics of the Spearman's rank correlation in the region of Baden-Württemberg is more demanding. The correlations grow slightly as a result of the increase of temporal scales which mostly occur at distances of 5-15 km. At distances greater than 15 km, the Spearman's rank correlation tend to stay steady or decline gradually on the whole.

Overall, temporal aggregation scales yield less uncertainty and smaller outliers of the parameter of interest without influence by the seasons and distances. However, most outliers are clustered at the close distances of 5-10 km which are dominated by the values lying below the lower boundary. At the distances above 15 km, there are some outliers, especially in the months of DJF and JJA, which are dominated by the correlations lying above the upper boundary of outliers.

**Effect of anisotropic assumption** Similar to those results which are conducted in Singapore, in order to detect the influences of different directions of the distance vectors on the Spearman's rank correlation functions, four different classes of directions of inter-gauge distances are made, namely, the direction of $0^o$, $45^o$, $90^o$ and $135^o$ from the North. For each direction, the inter-gauge distances are then grouped into several distance classes from 5 to 60 km with the increment of 5 km. A reasonable minimum number of gauge pairs for each direction and for a given distance vector at different locations is considered in order to estimate the spatial correlation functions in a reasonable way.

Figure 5.8 shows multiple plots of the iso-rank correlation lines in the two-dimensional domains of distances for precipitation occurrences at different time scales from hourly to monthly scales using the geostatistical approach. The analyses are conducted separately in different seasons for both regions, Bavaria and Baden-Württemberg. The figure attempts to describe the characteristics of the spatial correlation of precipitation on the basis anisotropic assumption. Each subplot corresponds to the time scale of interest from daily to monthly. The mean values of the Spearman's rank correlations over time periods for each direction of $0^o$, $45^o$, $90^o$ and $135^o$ are projected on both axes as in the case of Singapore. The isolines of the correlation function in two-dimensional spaces can be created by spatial interpolation over the whole region of interest.

Overall, the spatial correlations on the basis anisotropic assumption for Bavaria show significantly higher values than those which are located in Baden-Württemberg. On the radius of 60 km, the rank correlations roughly range from 0.3-0.4 for Bavaria and $\leq$0.1 for Baden-Württemberg. This pattern is consistent with the isotropic assumption.

The spatial correlation patterns seem to increase due to aggregation of temporal scales. The increase of the spatial correlation is more significant for Bavaria. The spatial rank correlations for Baden-Württemberg do not depend on the different seasons. This is because the correlation functions in different seasons exhibit similar values. In contrast, the spatial rank correlations for Bavaria in the seasons of MAM and JJA produce greater correlations than those which are located in Baden-Württemberg.

In Bavaria, the rank correlations within the North-West to the South-East direction tend to be slightly higher than those which are located in the North-East to the South-West direction or other directions. The fact applies only in the season of DJF for all different time scales. For the other seasons (MAM, JJA, SON), the contours of spatial correlations are approximately rectangular indicating that the rank correlations within the North-West to the South-East direction tend to be equal to those which are located in the North-East to the South-West direction for all different time scales but greater than the directions North to West or East to West.

In contrast, the spatial rank correlations in Baden-Württemberg exhibit superiority in the direction of the North-East to the South-West for all timescales and seasons, on the whole. This might be due to the mountainous region in the Black Forest, with a South-North orientation, which contributes significantly to typical precipitation processes in Baden-Württemberg.

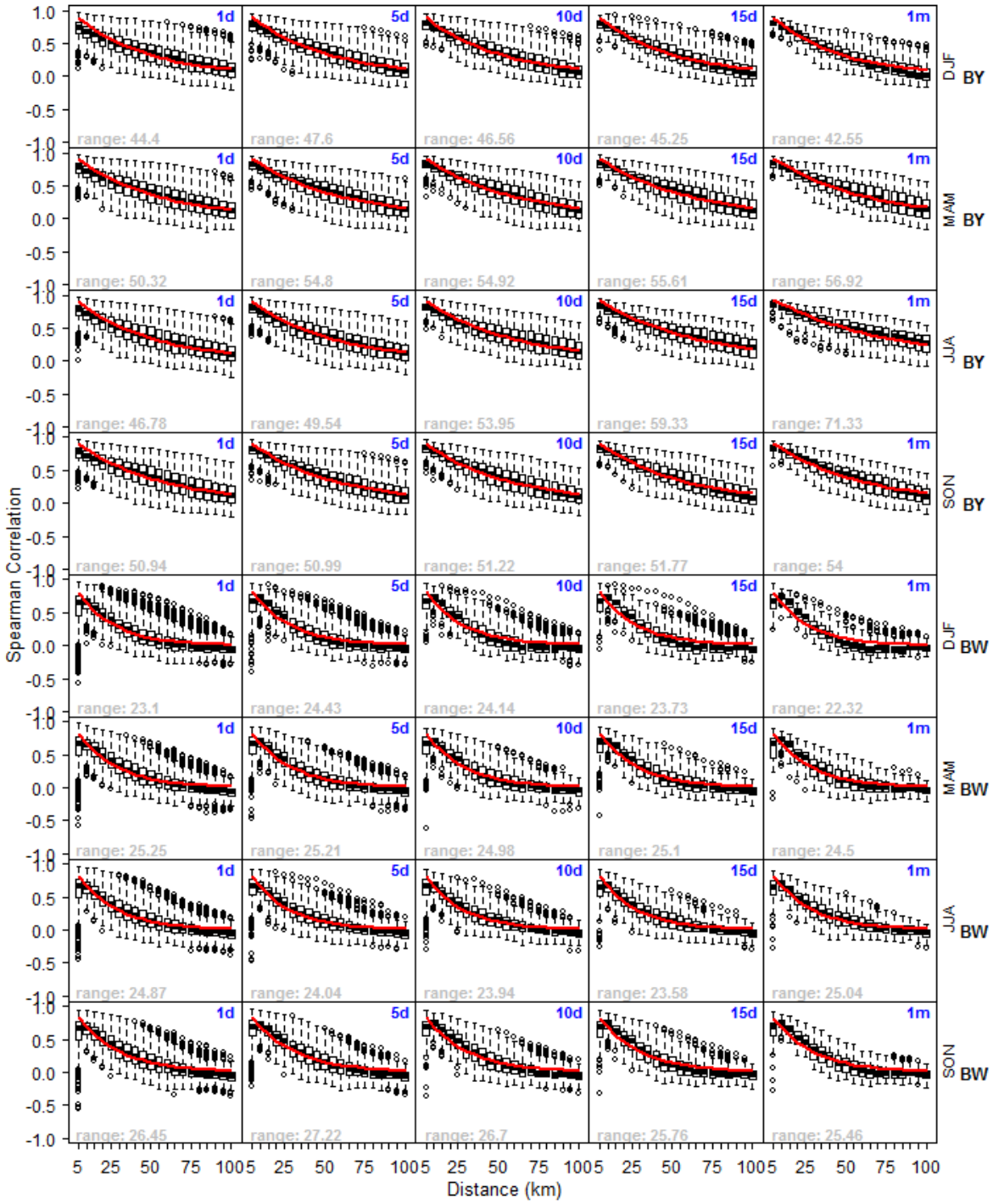



Figure 5.6: The Spearman's rank correlation functions over distances using the geostatistical approach in the regions of Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the rank correlation. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON) and regions (BY, BW). The horizontal axes represent distances (from 5 to 100 km). Panels from left to right represent the time scale (daily (1d) to monthly (1m).

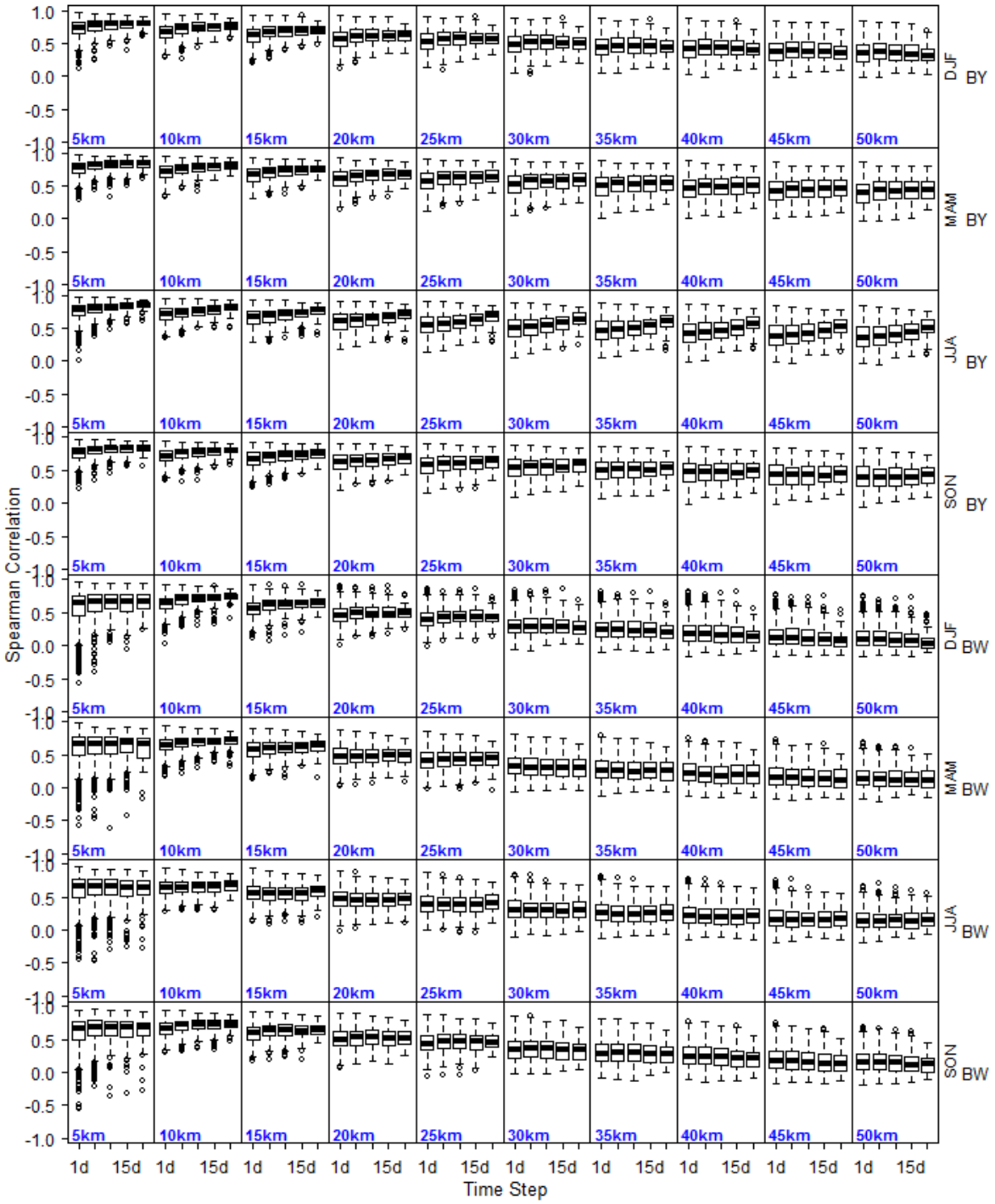



Figure 5.7: The Spearman's rank correlation functions over time scales using the geostatistical approach in the regions of Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the rank correlation. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON) and regions (BY, BW). The horizontal axes represent a variety of time scales from daily to monthly. Panels from left to right represent distances from 5 km to 50 km.

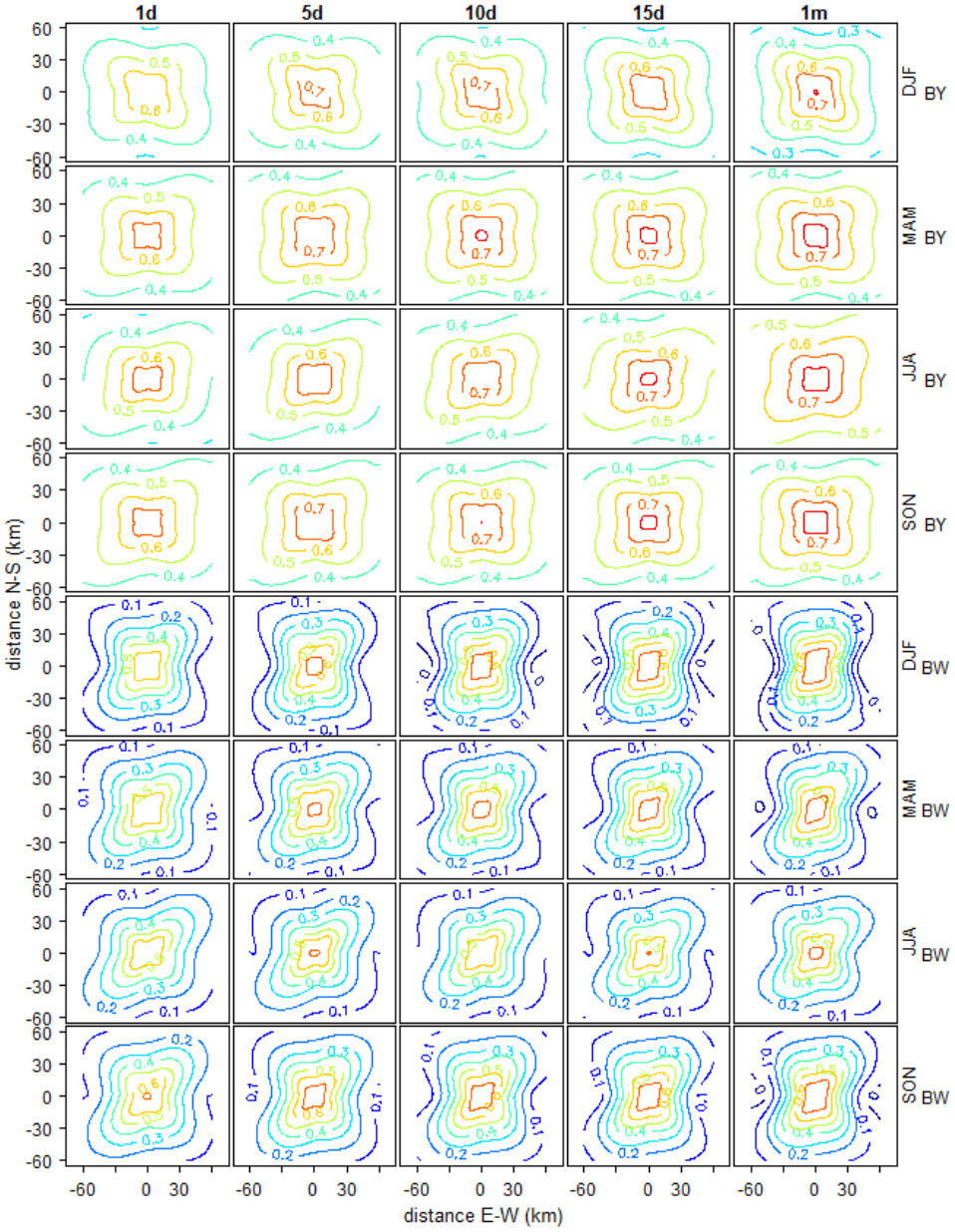



Figure 5.8: Isocorrelation lines in the two-dimensional spaces of distances using the geostatistical approach on diffent time scales in Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the lag distances in North and South direction in km unit. The horizontal axes stand for the lag distances in East-West direction. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON) and regions (BY, BW). Panels from left to right represent the time scale (daily (1d) to monthly (1m).

## 5.4 Pairwise context

### 5.4.1 Approach and methodology

#### 5.4.1.1 Isotropic assumption

Unlike the geostatistical approach where precipitation values at each time step are treated separately using the assumption of a single realization of spatial random functions based on the regionalization approach, in the context of pairwise-based approach here, precipitation amounts are analyzed between two gauge stations at the exact geographical locations containing simultaneous temporal observations. Empirical bivariate distribution functions are then constructed from each pair of gauge stations consisting of contemporaneous precipitation observations. The total number of empirical bivariate distribution functions equals all possible pairwise combinations of the precipitation gauge stations at different locations, which can be calculated by using the formula $\frac{n(n-1)}{2}$, where $n$ is the number of stations in the region of interest. Every pair of gauge stations is then investigated for estimating the Spearman's rank correlation without taking into account its relative orientation between two points, i.e. the isotropic assumption is followed.

Estimations of the Spearman's rank correlation are conducted underlying the Gaussian copulas distribution which is fitted to the empirical bivariate copulas. The estimation of the empirical bivariate copulas from all pairs of rain gauge stations containing time series precipitation data sets can be done with the following procedures. First, precipitation data at each gauge station are transformed into marginal uniform distribution ranging from 0-1, i.e. $U(0,1)$, using the empirical marginal distribution function or rank transformation. The unit squares are divided into $k$ x $k$ grids, where $k$ is an integer number. The empirical copula density can then be constructed by calculating the empirical frequencies on the grid. Finally, the Spearman's rank correlation coefficient is estimated by using the parameter of the bivariate Gaussian copulas which are best fitted to the empirical bivariate copulas.

Similar to the geostatisical approach, the bivariate Gaussian copulas are fitted to the empirical bivariate copulas considering zero values of precipitation using the censored bivariate normal distribution with a mean equal zero and standard deviation of 1 as explained previously in section 5.2. Precipitation is always characterized by the intermittent process with the combination of zeros (dry events) and non-zero precipitation (wet) occurrences. There are four different kinds of possible pairs with respect to the zeros, namely, (1) positive values at both stations $(X_1 > 0, X_2 > 0)$, (2 and 3) positive value at one station and a null value at the other $(X_1 > 0, X_2 = 0)$ and vice versa $(X_1 = 0, X_2 > 0)$, (4) no rain at both stations $(X_1 = 0, X_2 = 0)$, where $X_1$ and $X_2$ are the precipitation values collected at two rain gauges, station-1 and station-2, respectively.

According to previous studies on quantifying the spatial dependence between contemporaneous pairs of rainfall observations, taking into account the intermittent nature of the rain and the skewness of the data (Ha and Yoo, 2007; Habib and Krajewski, 2001; Serinaldi, 2008), only case-1 (only positive values at both rain gauges) provides correlation estimates useful for the characterization of the rainfall fields on the basis of the Pearson's correlation.

However, the recent studies conducted by Bárdossy and Pegram (2009, 2013) suggested that incorporating zero into modeling is a more realistic approach and better mimics nature. This study, however, considers all cases except case-4 in order to take into account zero-inflated precipitation data that commonly occur in the higher temporal resolutions.

#### 5.4.1.2 Anisotropic assumption

The anisotropic assumption is made in order to investigate the effects of different relative directions between pairs of gauge stations as well as its length of distance on the Spearman's rank correlation functions. All the spatial correlations between a pair of precipitation gauge stations obtained from the isotropic assumption are plotted in a two-dimensional domain according to the distance between the stations in the East-West (E-W) direction as the vertical axis (the x-axis) and the South-North (S-N) direction as the horizontal axis (the y-axis). The anisotropic assumption can be carried out with two approaches described below.

The first approach; all the spatial correlations between a pair of gauge stations at different locations with its relative orientation are projected into both axes; the x-axis and the y-axis. Projections to the x-axis (E-W direction) are conducted by multiplying the distance of two points by $\sin(\theta)$, where $\theta$ denotes the angle between two gauges and the North direction. In addition, the Spearman's rank correlations are projected to the y-axis (N-S direction) using the distance two gauge points $\times \cos(\theta)$. The projected correlation functions to both axes are then interpolated over the whole spatial domain at the region of interest yielding isolines of the correlation functions.

Another approach is that the first station of each pair is shifted to the origin of the graph and the correlation value is plotted at the end of the distance vector linking the stations. The correlation function between the pairs of stations 1 and 2 and the correlation between 2 and 1 are the same, so the resulting plot is point symmetrical (Beck et al., 2015). Similarly, the plotted correlation functions are then used to interpolate over the two-dimensional spaces of the lag distances in East-West and North-South directions resulting in isolines of the correlation functions.

### 5.4.2 Application in Singapore

Similarly to the geostatistical approach discussed in section 5.3, the effect of different distances as well as different temporal aggregations on the spatial Spearman's rank correlation functions are conducted but using the pairwise-based approach. Empirical investigation results are presented in Figure 5.9 and Figure 5.10. Figure 5.9 and Figure 5.10 depict the spatial rank correlation functions of precipitation values at different time scales from hourly to monthly scale as well as different seasons using the pairwise approach in the region of Singapore. Figure 5.9 highlights the effect of various distances on the Spearman's rank correlation, whereas Figure 5.10 emphasizes the impact of different time scales on the correlations.

**Effect of spatial scales** The spatial Spearman's rank correlations decline gradually as the result of the increment distances corresponding to the stations located at different locations and fit well enough using the negative exponential functions which are marked with the red lines as shown in Figure 5.9. The spatial correlation behavior above is similar to those which are conducted based on the geostatistical approach.

However, the spatial correlation functions carried out based on the pairwise approach exhibit significantly higher values in comparison to the geostatistical approach. The correlation lengths based on the pairwise approach roughly range from 12-255 km, whereas the correlation lengths which are conducted based on the geostatistical approach only roughly range from 12-15 km as presented in section 5.3.

Unlike the geostatistical approach, where the correlations are nearly independent of the seasons, empirical evidence proves that the spatial correlation functions conducted on the basis of the pairwise approach are influenced by the seasonal periods. For example, in the months of DJF and MAM, the correlation lengths roughly vary between 21 and 60 km, and 12 and 30 km for hourly and 12-hours scale, respectively. By contrast, in the months of JJA and SON, despite of the correlation lengths at the time scale of hourly and 12-hours presenting similar behavior for both seasons, namely, ranging from 15 to 43 km and 15 to 41 km, respectively, the values are significantly different from those which are estimated above in the seasons of DJF and MAM. For the higher temporal scales, such as from daily to monthly scales, the correlation lengths exhibit significantly different values for all different seasons, namely, ranging from 53 to 103 km, and from 46 to 110 km for the seasons of JJA and SON and roughly varying between 66 and 213 km and between 33 and 60 km for the seasons of DJF and MAM, respectively.

Another characteristic dissimilarity of the spatial rank correlation functions carried out on the basis the geostatistical and pairwise approach in term of the aspect of spatial scale is that the variability of the spatial rank correlation function based on the pairwise approach increases with the increase of distance scales. This is contradictory to the geostatistical approach where the increment of distance scales brings less uncertainty of the Spearman's rank correlation at least until a specific distance point, for example, the distance of 15 km.

**Effect of time scales** The impact of different temporal scales on the Spearman's rank correlation functions which are estimated based on the pairwise approach are obviously seen in Figure 5.9. Overall, the Spearman's rank correlations increase dramatically because of the increase of temporal scales from hourly to monthly scale. This tendency is more decisive than those which are calculated by the geostatistical approach. This pattern applies not only to the specific distances of the gauge pairs which are located at different locations but also to all distances from 5 to 25 km. This increasing trend of the correlations applies for all different seasons no matter which season of the year.

The variability of the Spearman's rank correlation function generally decreases with time scales. This is consistent with the geostatistical approach where the increment of temporal scales brings less uncertainty of the correlation. However, the degree of variability is substantially small in comparison to the geostatistical approach. Outliers are found more

frequently as the temporal scales increase, which slightly differs from the geostatistical approach.

**Effect of anisotropic assumption** Figure 5.11 shows isolines of spatial correlation in the two-dimensional spaces of distances at different timescales from hourly to 5-days scale and in different seasons using the pairwise approach in Singapore in order to detect spatial anisotropy correlation. Overall, the anisotropic spatial correlation functions are pronounced only for the DJF season, during which the North-East Monsoon occurs. Thus, the spatial rank correlation in the North-East to South-West direction is the largest among other directions.

The Northeast Monsoon occurs from December to early March when the cooling of the air masses over Siberia and Tibet leads to a high-pressure zone over Asia generating a constant north-eastern airflow transporting moisture from the Chinese Sea into the area. This means that the correlations are higher orthogonal to the prevailing wind direction. It is an indication that some of the events have a frontal character so that the stations in the North-West to the South-East configuration are on the same line relative to the flow field and thus receive precipitation within the time Beck et al. (2015). For other seasons (MAM, JJA, and SON), the spatial correlations exhibit isotropic behavior because of the cylindrical isolines of the correlation functions implying that the rank correlations are the same in all directions for any given distance, on the whole.

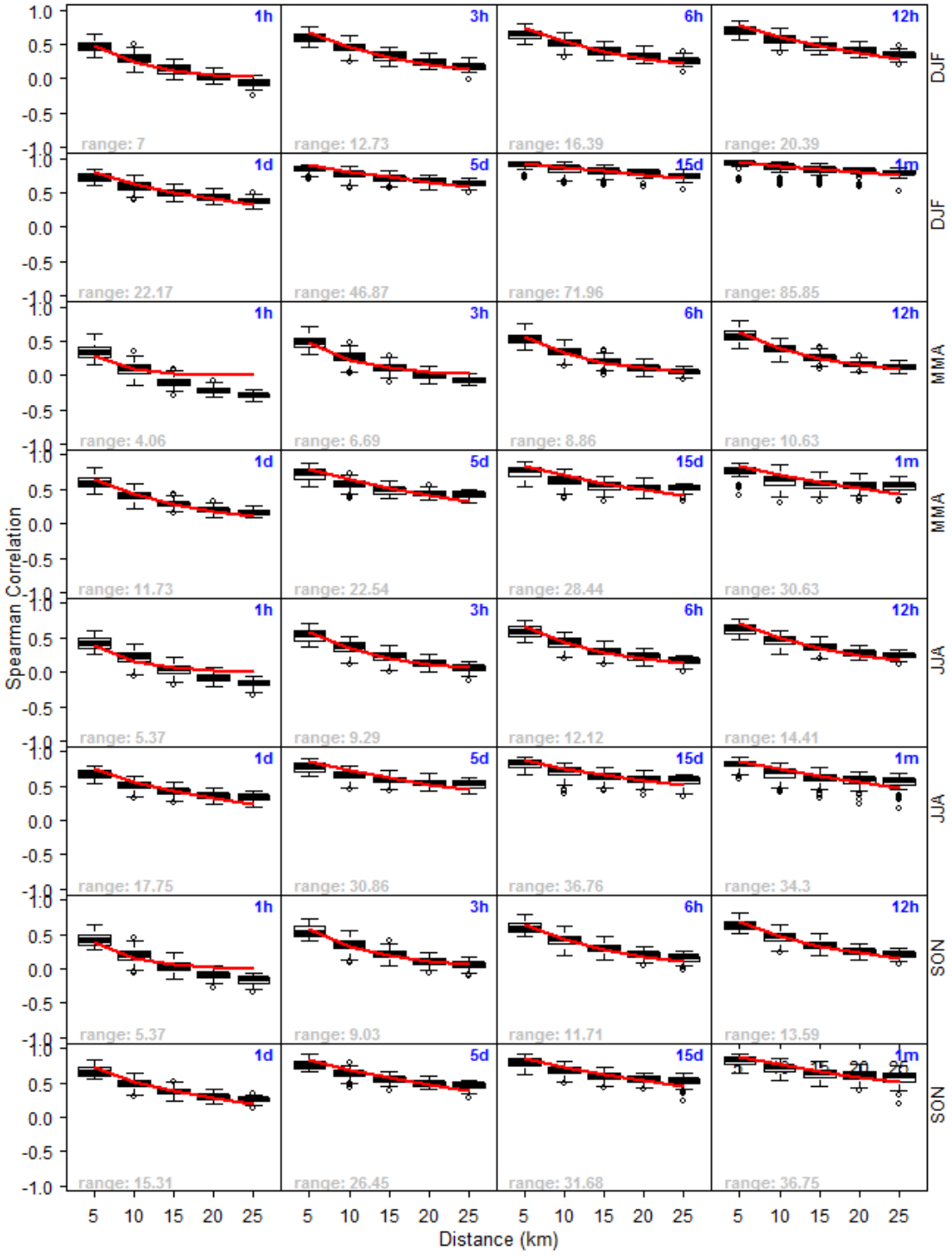



Figure 5.9: The Spearman's rank correlation functions over distances using the pairwise approach in the regions of Singapore. The vertical axes represent the rank correlation. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). The horizontal axes represent distances (from 5 km to 25 km). Each panel represents the time scale (hourly (1h) to monthly (1m).

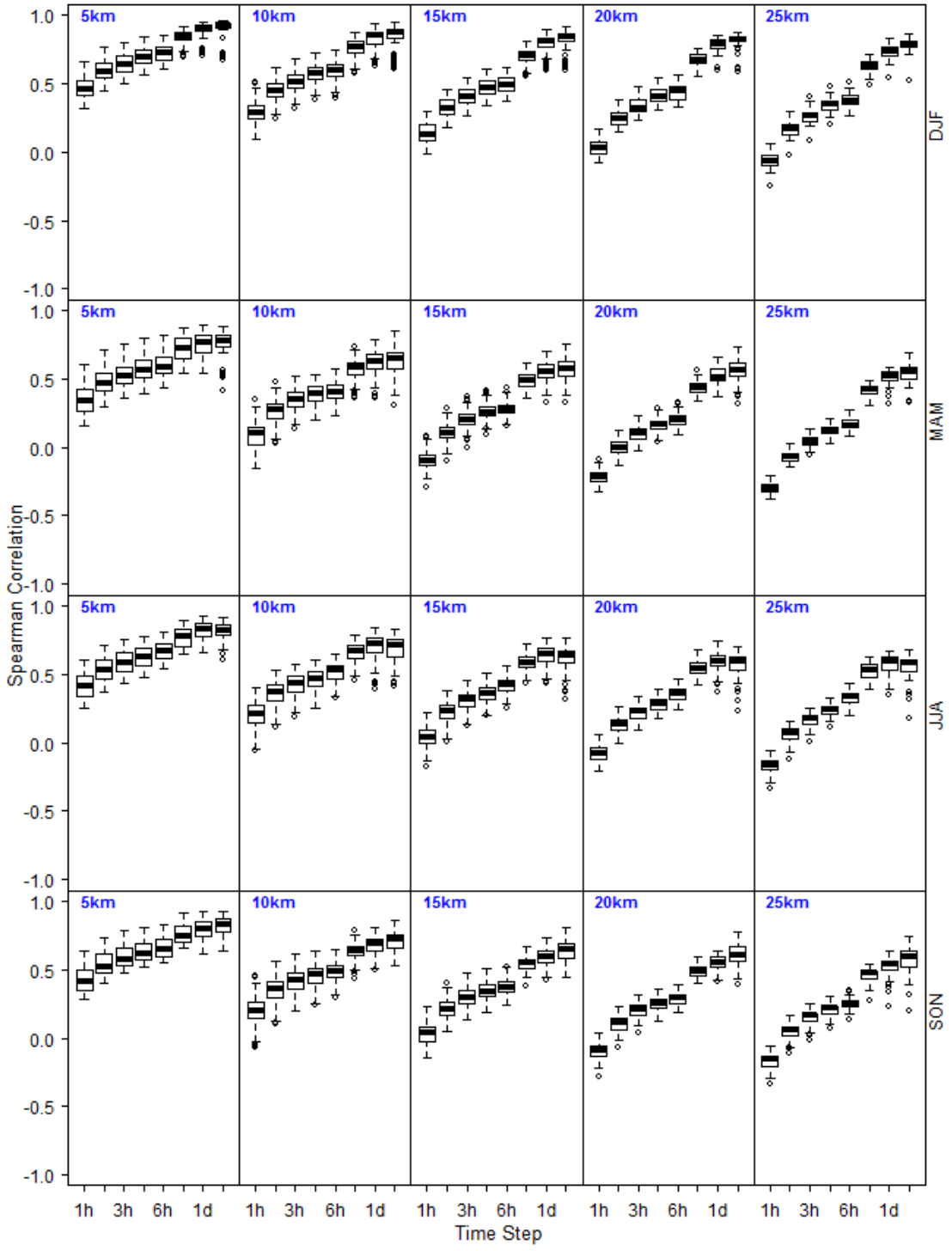



Figure 5.10: The Spearman's rank correlation functions over time scales using the pairwise approach in the region of Singapore. The vertical axes represent the rank correlation. The horizontal axes represent a variety of time scales from hourly to monthly. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panel from left to right represent distances (5 km to 25 km).

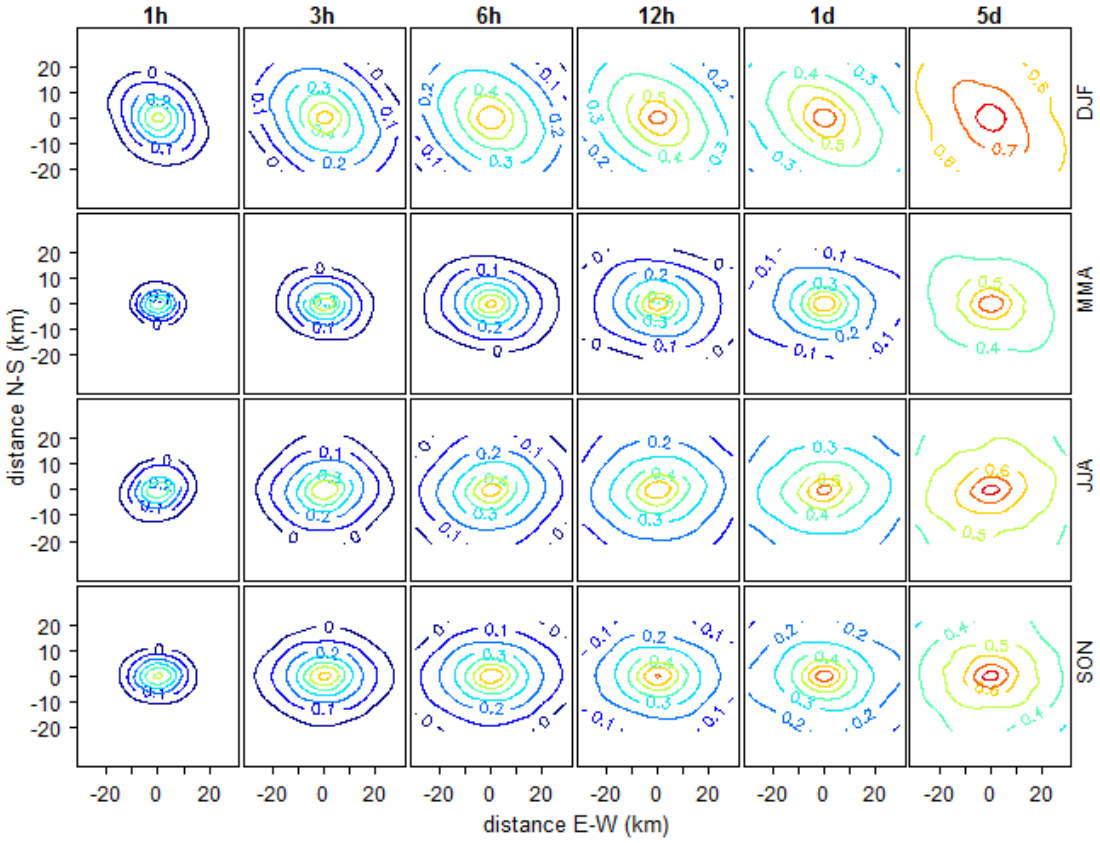



Figure 5.11: Isocorrelation lines in the two-dimensional spaces of distances using the pairwise approach in Singapore. The vertical axes represent the lag distances in North and South direction in km unit. The horizontal axes stand for the lag distances in East-West direction. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent time scales from hourly to 5-days.

### 5.4.3 Application in Bavaria and Baden-Württemberg

Figure 5.12 and Figure 5.13 depict the Spearman's rank correlation at a variety of time scales and in different seasons based on the pairwise approach in the regions of Bavaria and Baden-Württemberg. Figure 5.12 describes the effects of different distances on the Spearman's rank correlation functions and Figure 5.13 highlights the impact of various time scales on the correlation functions. Both figures consist of 40 multiple box plots of the correlation functions of precipitation occurrences at the different time scales from daily to monthly scale.

**Effect of spatial scales** Overall, the Spearman's rank correlations tend to decline with increasing distance scales from 5 to 100 km. This is because the negative exponential correlation function is also fitted quite well to the empirical rank correlations as shown in Figure 5.12 marked with the red lines. The parameter of the negative exponential correlation function is used to characterize the behavior of the spatial Spearman's rank correlation due to its simplicity.

Similar to the cases in the region of Singapore, the Spearman's rank correlation in both regions, Bavaria and Baden-Württemberg, which are estimated based on the pairwise approach, exhibit significantly greater values than those estimated by the geostatistical approach. The correlation lengths, which are estimated using the pairwise approach, roughly range from 249 to 1656 km and 273 to 1437 km across temporal scales and seasons for the regions of Bavaria and Baden-Württemberg, respectively, whereas the correlations which are estimated using the geostatistical approach, roughly vary between 105 and 294 km and 60 and 84 km for both regions, respectively. It means that the fitted theoretical exponential correlation functions based on the pairwise approach decrease very slowly in comparison to those estimated by the geostatistical approach.

Despite the small differences, the correlation lengths for Baden-Württemberg are generally higher than those located in Bavaria, especially for the smaller time scales. This contrasts with the results obtained from the geostatistical approach, where the correlation lengths for Baden-Württemberg are systematically lower in comparison to the case in Bavaria. For example at the daily scale, the correlation length for Baden-Württemberg for the season of DJF is 408 km, whereas, in Bavaria, the correlation length is slightly lower, that is 402 km. For another season of MAM, JJA, and SON, the correlation lengths are 366, 273, and 438 km for Baden-Württemberg and 312, 249, 405 km, for Bavaria. However, the effect of the seasons on the rank correlations tends to be similar to both regions for all different temporal scales.

The correlation lengths of the seasons of SON and DJF for both regions are systematically greater than those occurred in the seasons MAM and JJA. In addition, both regions exhibit a similar pattern that the variability of the rank correlation at each of the time scales increases along with the increase of inter-gauge distances. The greater variability of the entire data sets is represented by the wider of the distance covering the middle 50% of the data as shown in the box-plots, that is the difference between the 75th percentile and the 25th percentile. Similarly, the number of outliers increases because of increasing distances scales for all time scales and seasons. Most outliers arise below the lower boundary.

**Effect of time scales** The influence of increasing temporal aggregations on the Spearman's rank correlation incorporating different classes of inter-gauge distances and seasons is clearly seen in Figure 5.13. In general, there is a substantial increase in the magnitude of the spatial correlation with temporal aggregations. However, increasing this correlation trend is also followed by increasing its uncertainty on the whole. The number of outliers also increases substantially in accordance to the rise of temporal scales. Most outliers occur below the lower boundary and are more concentrated within the higher distance groups.

**Effect of anisotropic assumption** Figure 5.14 shows iso-correlation lines in the two-dimensional spaces of inter-gauge distances using the pairwise approach in Bavaria and Baden-Württemberg in order to detect the anisotropic spatial correlation. Within a radius of 60 km, the direction of anisotropic spatial correlations can be detected from South-West to North-East at the daily scales for all seasons and both regions except for the season of DJF in Bavaria where it seems to be the opposite way. Increasing time scales yield the pattern of iso-tropical spatial correlation for both regions within different seasons except for the season of DJF in Baden-Württemberg.

## 5.5 Summary

- Most precipitation amounts in Bavaria exhibit a greater spatial rank correlation than those occurring in Baden-Württemberg and the spatial rank correlation functions for Baden-Württemberg are higher than the ones for Singapore. The spatial correlation in Singapore is very low indicating that the high spatial variability exists caused by local convective precipitation mainly driven by vertical processes. In contrast, in Southern Germany, there is a factor of the large scale frontal system from the atmospheric variables influencing the local variable of precipitation mainly driven by horizontal processes.
- The spatial rank correlations calculated using the pairwise approach with simultaneous temporal occurrences exhibit a systematically higher value than using the geostatistical approach. This empirical evidence is consistent with different climate regions, in Singapore and the South of Germany. The reason for this is mainly caused by the driving processes with large spatial extent where the geostatistical approach assumes spatial stationary over the spatial domain of interest and therefore neglects the influence of the mean value.
- Spatial correlations depend on aggregation. The longer the time interval is, the higher the correlations are.
- The concept of geostatistical approach has been widely applied to the precipitation fields for solving the spatial interpolation issues, especially used for the rainfall-runoff processes where both precipitation and runoff require the same time event. The pairwise approach containing simultaneous temporal observations is useful for infilling missing precipitation observations.

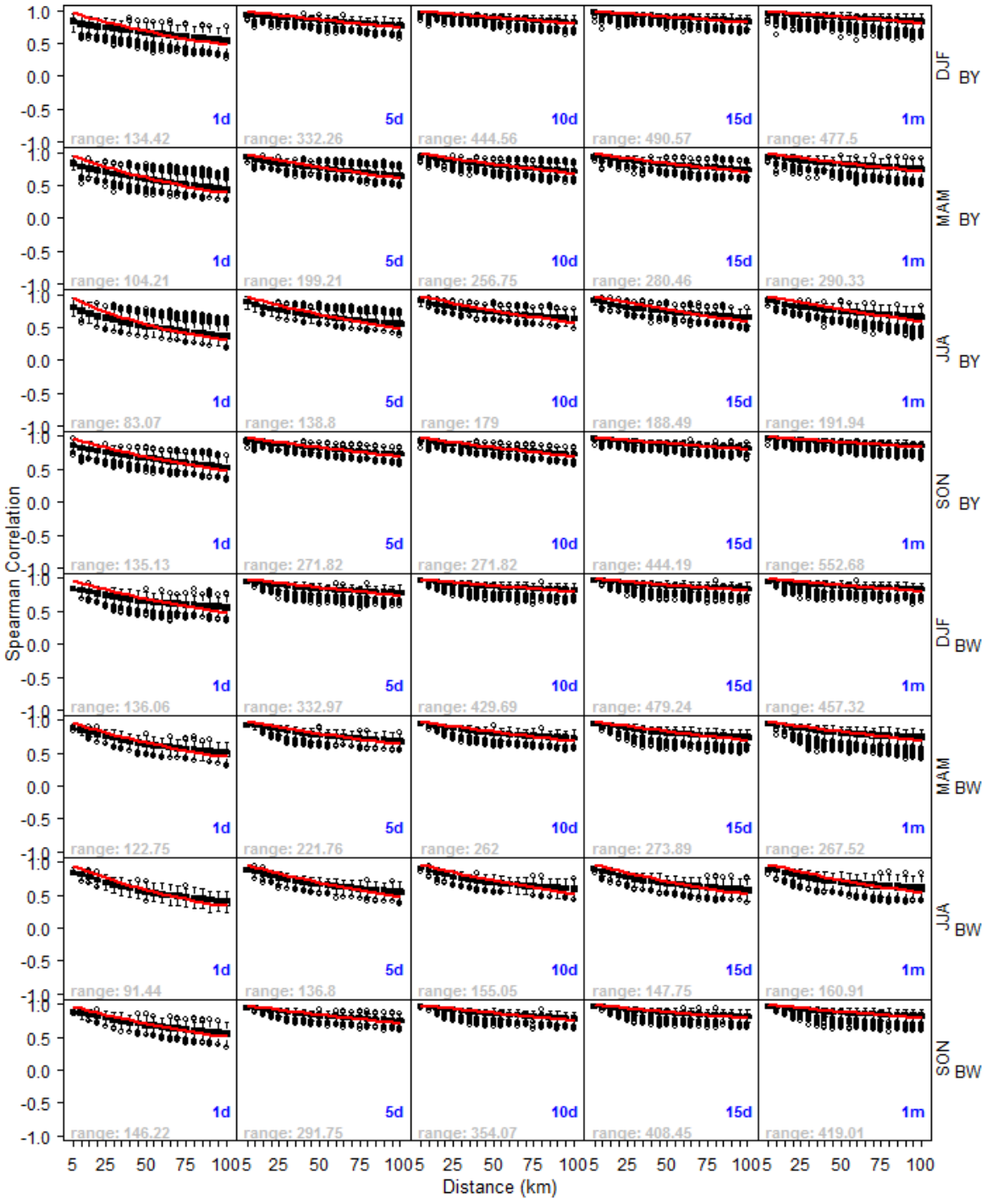



Figure 5.12: The Spearman's rank correlation functions over distances using the pairwise approach in the regions of Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the rank correlation. Panels from top to bottom represent the seasons (DJF, MAM, JJA, and SON) and regions (BY, BW). The horizontal axes represent distances (from 5 to 100 km). Panels from left to right represent the time scale (daily (1d) to monthly (1m).

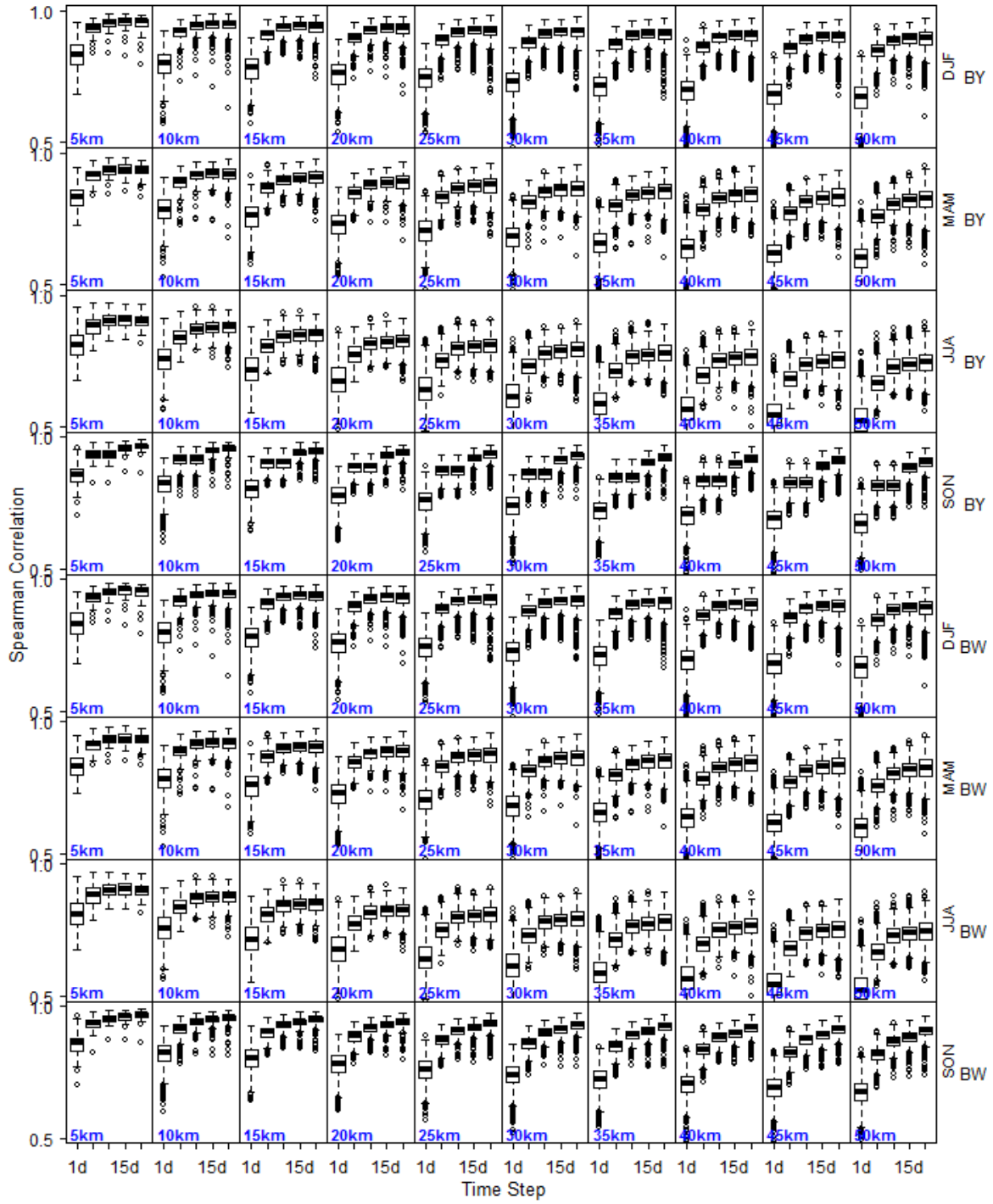



Figure 5.13: The Spearman's rank correlation functions over time scales using the pairwise approach in the regions of Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the rank correlation. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON) and regions (BY, BW). The horizontal axes represent a variety of time scales from daily to monthly. Panels from left to right represent distances from 5 km to 50 km.

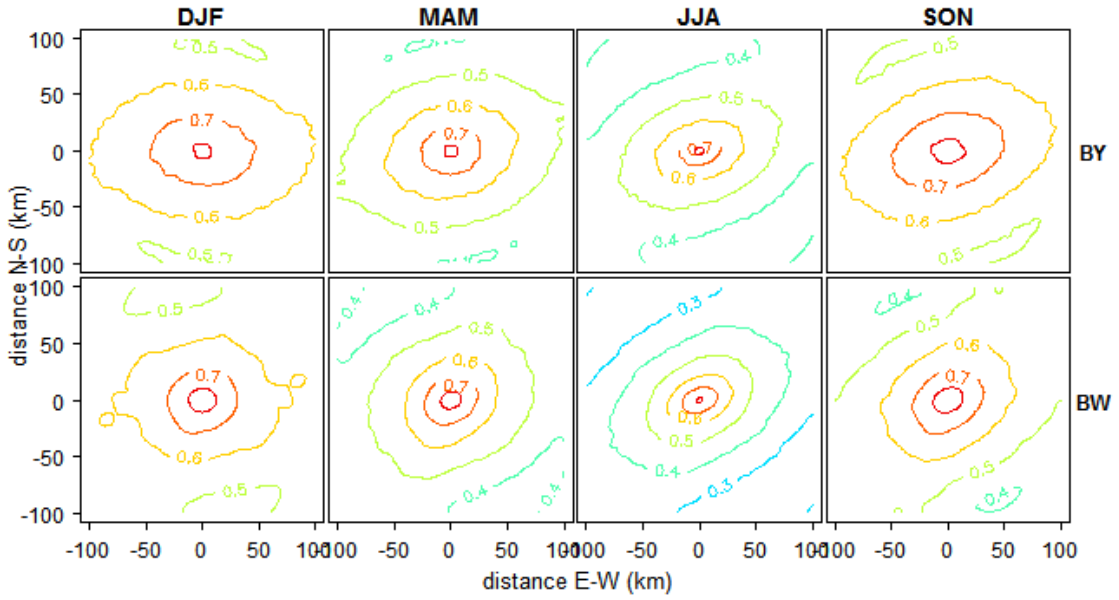



Figure 5.14: Isocorrelation lines of daily precipitation in the two-dimensional spaces of distances using the pairwise approach in Bavaria (BY) and Baden-Württemberg (BW). The vertical axes represent the lag distances in North and South direction in km unit. The horizontal axes stand for the lag distances in East-West direction. Panels from left to right represent seasons (DJF, MAM, JJA, and SON). Panels from top to bottom represent regions of BY and BW.

# 6 Investigating asymmetric spatial dependence using bivariate copulas

## 6.1 Introduction

Reliable spatial information on precipitation data sets is required for an accurate quantification of water resources planning and management since much empirical solid evidence shows that precipitation exhibits high spatial variability. Unfortunately, the spatial coverage of available precipitation data sets is frequently inadequate for representing the higher degree of spatial variability of precipitation. The low densities of precipitation gauge network due to a small number of stations installed in many parts of the world brings poor quality of spatial precipitation data sets. Hence, a decent spatial interpolation method is necessary in order to estimate ungauged areas.

There are a variety of spatial interpolation models that have been developed for precipitation applications over the last a half of century. Many of them are developed based on Gaussian assumption because they are derived from the concept of linear regression. One of the spatial models, which is popularly implemented in precipitation studies is Kriging. Kriging type models can be classified into stochastic models, which are able to deliver uncertainty of the interpolation quality directly with its variances, as precipitation behavior exhibits a higher uncertainty of spatial distribution of precipitation.

However, Kriging models offer the quality of interpolation, which is just a function of the observation density and the variogram models (Bárdossy, 2006). These functions can only describe the spatial dependence as an integral over the whole distribution of the parameter values. Anomalies in the observations and areas with high or low variabilities cannot, therefore, be considered for estimating the variances (Bárdossy and Li, 2008). These models also cannot detect properly the different percentile values, for example, extremes, which might have a different spatial dependence structure from the central values (Bárdossy, 2006). This issue is relevant to precipitation fields which are characterized skewed distribution, and variogram functions which can only be dominated by a few anomalously differing pairs, might occur.

Copula based model is a promising tool to overcome those drawbacks. Copulas are also flexible to analysis the marginal distributions separately from the joint distributions and allow them to be same or different from the distribution family. Furthermore, copulas are capable of expressing dependence structures on quantile scale, which is able to detect both non-linear and tail, or asymmetric dependence (Nelsen, 2006). Recently, spatial interpolation copula models have been implemented in the precipitation studies among few scientists

(Bárdossy and Pegram, 2013, 2014; Wasko et al., 2013).

Such spatial interpolation models with underlying Gaussian assumption are convenient for computation, but it could deviate from reality (either under- or overestimating), if the assumption is hard to fulfill. One of the characteristics of the Gaussian assumption is the symmetric dependence between low and high values. Thus, a deeper investigation of asymmetric spatial dependence on the characteristics of precipitation fields is of importance for a better development of spatial precipitation modeling.

The purpose of this chapter is to investigate the non-symmetrical spatial dependence of precipitation fields between high and low values with a given quantile threshold $q_0$ (such as quantile 10%, 20%, 30%, and 40%) on the basis of empirical bivariate copulas. In addition, to see deeper effect of different distance and temporal scales on its asymmetric dependence is also interesting within different climate regions. Furthermore, conditional investigation of asymmetric spatial dependence restricted to days with atmospheric circulation patterns (CPs) is carried out. RCM data sets are also investigated afterward to see to what extent, non-symmetrical dependence behavior is related to the underlying physical processes. Finally, temporal investigations simultaneous occurrences of high and low values of precipitation between two gauges are also investigated for all possible combinations of pairs.

## 6.2 Geostatistical approach

Bivariate copula constructed by the geostatistical approach used here is a similar approach as described in section 5.3 of chapter 5. The asymmetry function $A_3$ at any given certain of a time interval and a specific separating distance h as well as a different threshold $q_0$ is calculated using Equation 4.20 for each empirical bivariate copula as discussed in section 4.3. If the asymmetry function $A_3$ is positive, then the empirical bivariate copula of the field of precipitation presents the positive asymmetric spatial dependence. If $A_3$ is negative, then it presents the negative asymmetric spatial dependence. If $A_3$ is zero, then it shows the symmetric spatial dependence.

Based on the calculation of the asymmetry function for all time periods, it is then recapitulated how many positive, negative, and zero asymmetric spatial dependences exist. The Gaussian simulation is then applied to measure the degree of its uncertainty with a certain significant level, namely 90% of confidence interval. Theoretically, Gaussian simulations present the symmetric dependence between low and high values, but it, practically, might have either positive or negative asymmetric dependence. The range of non-symmetrical dependence from Gaussian simulation is called here as a confidence interval of estimations.

The detailed procedure for investigating the asymmetry function and the statistical testing for measuring a degree of uncertainty are described as follows.

1. For any selected time scale, season, and quantile threshold $q_0$, the empirical asymmetric dependence is calculated using Equation 4.20 on the basis of the empirical bivariate copulas using the geostatistical approach with a given separating distance, for example, 5 km, 10 km, and other distances with an increment of 5 km.

2. For any selected time scale, season, and quantile threshold $q_0$, precipitation amounts are transformed to the Gaussian domain using the normal score transformation.
3. Based on the statistical parameters obtained in step-2 above, multivariate normal distribution with $t$ times $n$ dimensions are generated, where $n$ is the number of gauge stations in the region of interest and $t$ is the length of precipitation time series.
4. For each realization from the $t$ Gaussian realizations, the asymmetric spatial dependence using Equation 4.20 is calculated and determined whether there is positive, negative, or zero asymmetric dependence.
5. For all $t$ Gaussian realizations obtained from the step-3, the number of events at which precipitation values present the positive, negative and zero asymmetric spatial dependence from step-4 is counted.
6. The steps 1-5 are repeated 100 times and based on the simulations, thus the quantile 5% and 95% for lower and upper limit of the confidence interval can be calculated for all parameters of interests.
7. The steps 1-6 are conducted repeatedly using a different quantile threshold (such as, $q_0$= 0.1, 0.2, 0.3, and 0.4) and a different season (such as, DJF, MAM, JJA, and SON), and in various time scales ranging from hourly to monthly for Singapore and from daily to monthly for both Bavaria and Baden-Württemberg.

### 6.2.1 Implementation in Singapore

Investigations of non-symmetrical spatial dependence based on the bivariate copulas using the geostatistical approach expose new findings in the Singapore region. The region of Singapore has been characterized as a homogeneous area regarding topography due to the flat land surface, even though it is also described as high variable in terms of spatial and temporal precipitation events due to the tropical climate region. The investigation findings are presented in Figure 6.1, which shows frequencies of precipitation occurrences that exhibit the positive and negative asymmetric dependences for the quantile threshold of 0.2 at the different temporal scales (from hourly to monthly) and at the different distance scales as well as in the different seasons (namely, DJF, MAM, JJA, SON).

**Effect of spatial scales** Overall, precipitation values which follow the positive asymmetric dependence decrease with the increase of distances and, on the contrary, precipitation amounts which behave the negative asymmetric dependence increase with the increase of distances. In this case, precipitation values exhibit the positive asymmetric dependence dominantly at close distances below 10 km, while precipitation amounts, which present the negative asymmetric dependence become more dominant at far distances above 20 km. Then, precipitation values, which showing the symmetric spatial dependence, lead significantly at distances between 10 and 20 km. This means that there is a transition phase from the positive asymmetric spatal dependence to the negative asymmetric spatial dependence through the Gaussian symmetric spatial dependence at the specific distances. In other

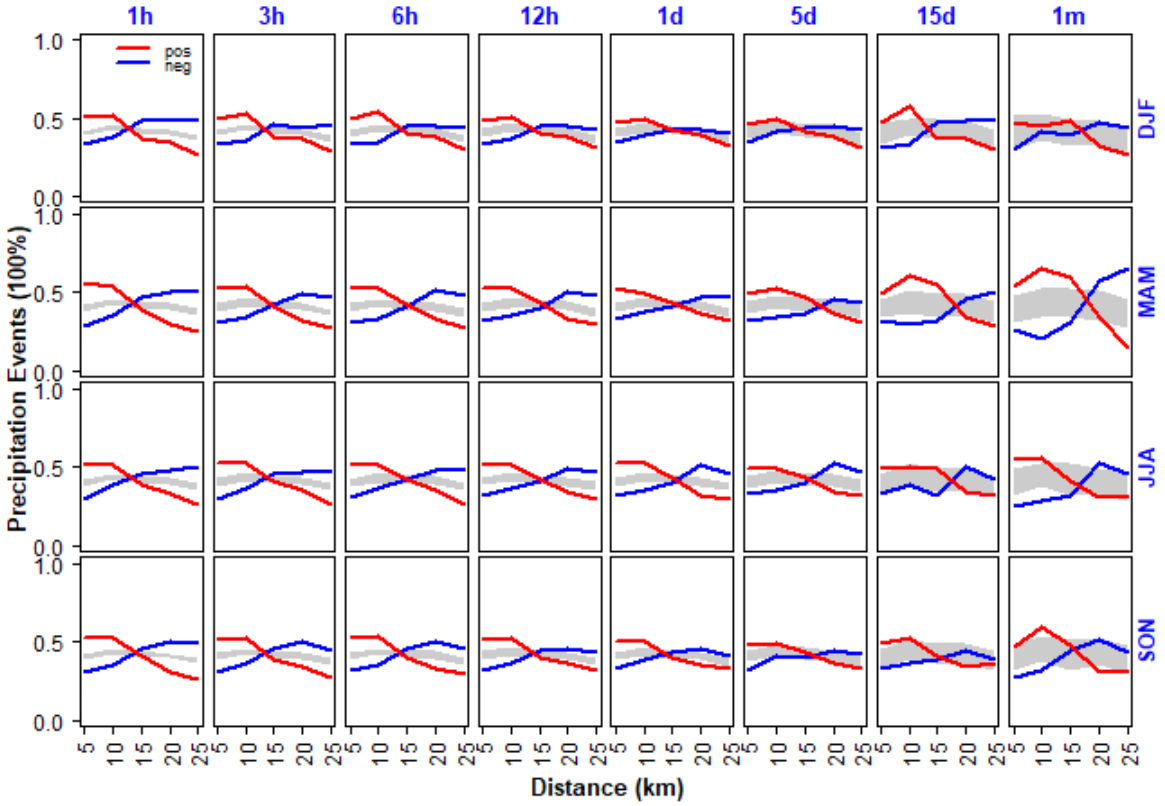



Figure 6.1: Asymmetry function using the geostatistical approach at different temporal scales in Singapore. The primary vertical axes represent the number of precipitation occurrences presented in 100% scale. The horizontal axes represent various distances (5 km to 25 km). The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The grey shadow areas denote the confidence interval of 90% of the symmetric Gaussian dependence. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent the time scales from hourly (1h) to monthly (1m).

words, the space scales play a crucial role to characterize the behavior of asymmetric spatial dependence. The positive asymmetric dependence exhibits extensively at neighborhood distances because precipitation amounts usually occur in a clustered manner. As a result, high values of precipitation at adjacent locations are more spatially correlated than low values.

**Effect of time scales** The effect of space scales on the characteristics of asymmetric spatial dependence, either positive or negative asymmetric dependence is restricted to a certain time scale. This means that temporal aggregation of precipitation amounts might generate an important influence on the characteristics of asymmetric dependence. In general, the number of events at which precipitation amounts exhibit the asymmetric spatial dependence, either positive or negative, is significantly dominant at the higher time resolutions. Precipitation values which follow either positive or negative asymmetric dependence then decrease with the increase of temporal scales. The positive asymmetric spatial dependence, which frequently occurs at nearby distances below 10 km and the negative asymmetric dependence, which often arises at far distances above 20 km, are more pronounced at the subdaily scales from hourly to daily, on the whole. The asymmetric dependence disappears at the higher temporal scales, for example, from 5-daily to monthly. In the other words, the symmetric Gaussian spatial dependence tend to exist significantly at the higher time scales, especially from 5-daily to monthly, no matter whether they are located at close distances or far distances.

**Effect of seasons** The characteristics of the asymmetric spatial dependence above generally exhibit similar patterns for all different seasons. However, there is still a small difference in the behavior of the asymmetry function, especially in the DJF season, which can be explained as follows. In the DJF season, the positive asymmetric spatial dependence occurs substantially at distances below 10 km from hourly to only 12-daily, and then it turns into the symmetric Gaussian spatial dependence from daily or 5-daily to monthly. Although at the daily scale, the positive asymmetric spatial dependence still presents slightly higher symmetrical dependence, it is a very few numbers and almost negligible. In contrast to the DJF season, for other seasons (for example, MAM, JJA, and SON), the positive asymmetric spatial dependence is significantly dominant at distances below 10 km from hourly to daily and then become the symmetric Gaussian dependence from 5-daily to monthly scale. Nevertheless, there is a strong empirical evidence that precipitation amounts generally exhibit the positive asymmetric spatial dependence at close distances and the high temporal resolutions for all different seasons no matter in which season of the year.

**Effect of precipitation intensities** The positive asymmetric spatial dependence of precipitation values describes that the high values of precipitation are more spatially correlated than the low values. This indicates that precipitation events with high intensities will yield the positive asymmetric dependence at adjacent gauges because the extreme precipitation events tend to occur in a clustering manner at nearby locations. In order to prove it, the

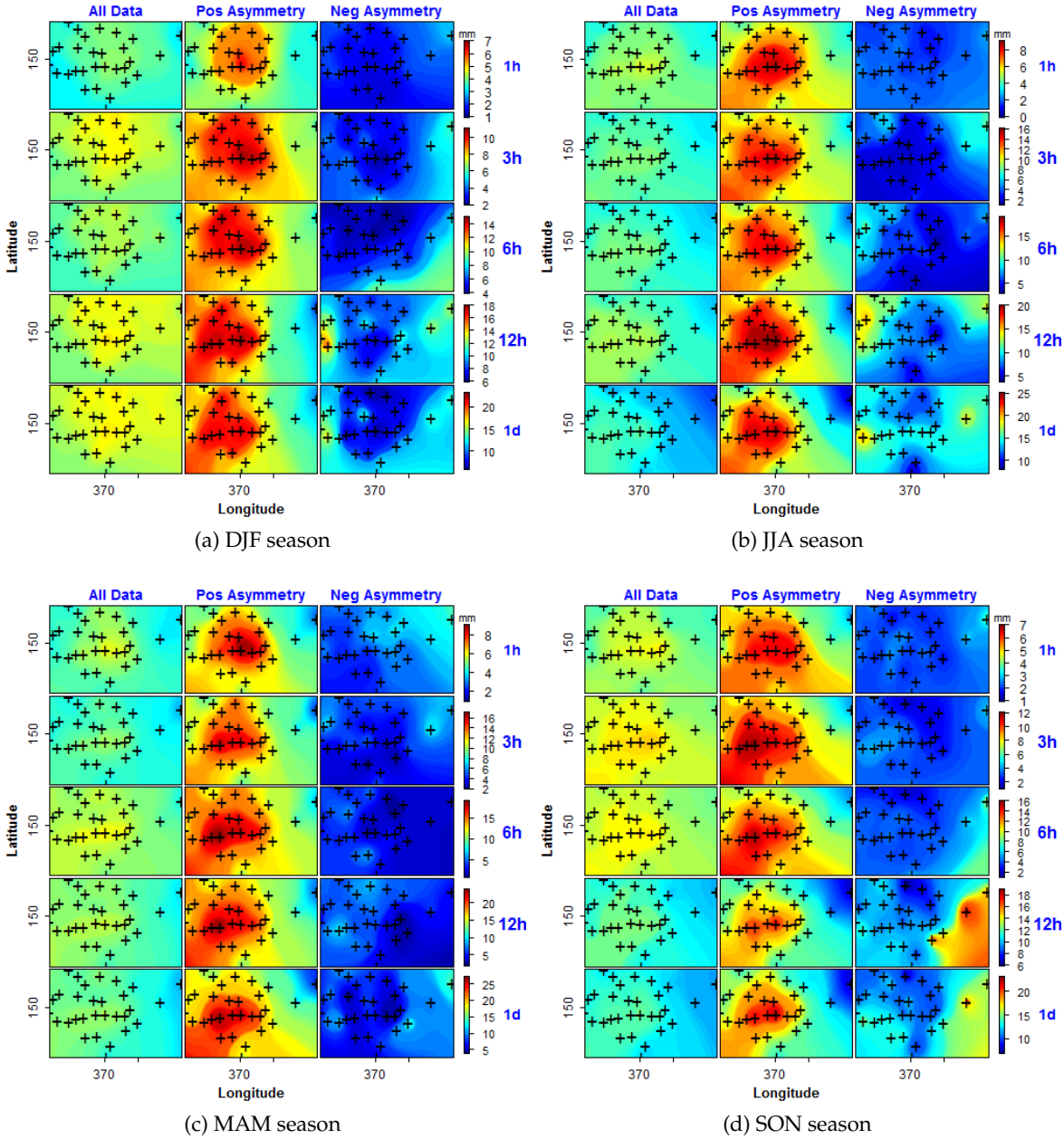
precipitation data sets are divided into three data sets, namely, all data sets, positive asymmetry, and negative asymmetry corresponding to the characteristics of asymmetric spatial dependence.

The first data set is from all precipitation events at a given aggregation time. The precipitation amounts for each station are then averaged over all time series yielding a mean value of precipitation at each gauge. This is also conducted for all gauges resulting in mean precipitation amounts over all stations. The interpolated precipitation amounts over the whole region of interest are created from those points resulting in a single realization of spatial precipitation amounts over a regular grid.

The second data set is from all selected events at which precipitation amounts follow the positive asymmetric spatial dependence at close distances below 10 km. The precipitation intensities at the observation points are also averaged over selected time periods above resulting in a single spatial realization of mean precipitation points. The spatial interpolations are carried out to estimate precipitation intensities at unsampled locations over the whole region of interest from those points resulting in a single realization of spatial precipitation field over a regular grid.

The third data set comes from the extractions of all events at which precipitation values exhibit the negative asymmetric spatial dependence at distances below 10 km. The mean value of precipitation for each station is calculated based on the available precipitation values which follow the negative asymmetric spatial dependence. This is carried out for all gauges resulting in mean precipitation amounts over all stations. The precipitation intensities at unsampled locations over the whole spatial domain of interest are estimated based on those points resulting in a single spatial realization of precipitation field over a regular grid.

Figure 6.2 shows spatial precipitation intensities using the different three data sets above (namely, all data sets, positive asymmetry, and negative asymmetry) at the different temporal scales from hourly to daily in the different seasons (DJF, JJA, MAM, and SON). It is seen that that heavy precipitation triggers more the positive spatial asymmetric dependence at close distances, and, on the contrary, the negative spatial asymmetric spatial dependence is commonly caused by the lighter precipitation intensities. In the context of urban water management, empirical investigation evidence reveals that instead of using the spatial Gaussian model or the symmetrical spatial dependence based model, the spatial precipitation models based on the positive asymmetric spatial dependence is recommended to be adopted for a better flood risks management.




(a) DJF season

(b) JJA season

(c) MAM season

(d) SON season

Figure 6.2: Precipitation intensities with different asymmetric dependence in Singapore. The primary vertical axes represent latitude and the horizontal axes define longitude (both in km). Panels from top to bottom represent temporal scales (hourly (1h) to daily (1d). Panels from left to right represent all data, positive and negative asymmetric dependence, respectively.

### 6.2.2 Implementation in Bavaria and Baden-Württemberg

Analysis of the asymmetric spatial dependence is also carried out in large scale regions in southern part of Germany, namely, the state of Baden-Württemberg and the state of Bavaria. The region of Baden-Württemberg has almost a fifty-fold greater area than Singapore and the area of Bavaria approximately doubles to the area of Baden-Württemberg. Figure 6.3 and Figure 6.4 present the asymmetric spatial dependence in these regions with a variety of different time scales from daily to monthly (1 d, 5 d, 15 d, 1 m) and various spatial distances from 5 km to 60 km with an increment of 5 km for the seasons of DJF, MAM, JJA, and SON.

**Case study in Bavaria** Empirical investigation results reveal that precipitation values in Bavaria, which exhibit the positive asymmetric spatial dependence, also frequently occur at the nearby gauge stations and the negative asymmetric dependence appears significantly at far distances. Precipitation values, which follow the positive asymmetric dependence decrease with the increase of distances, and, on the contrary, precipitation amounts which behave the negative asymmetric dependence increase with the increase of distances.

Unlike in the region of Singapore, this pattern applies to all time scales ranging from daily to monthly for all different seasons. This is an indication that increasing temporal scales from daily to monthly does not change the behavior of asymmetric spatial dependence. Instead, precipitation amounts that follow the positive asymmetric spatial dependence at close distances are still pronounced no matter which time scale is used.

The fact that precipitation amounts which exhibit the positive asymmetric spatial dependence at nearby gauges are influenced by seasons. For example, the domination of the positive asymmetric spatial dependence is discovered until the distance of 30 km for the DJF season, while in the JJA season, it is only until distance of 15 km. For the seasons of MAM and SON, precipitation values which follow the positive asymmetric spatial dependence occur at the same distances, which are below 25 km.

**Case study in Baden-Württemberg** Similar to Bavaria, precipitation values in Baden-Württemberg which following the positive asymmetric spatial dependence generally occur at close distances and precipitation amounts which exhibiting the negative asymmetric spatial dependence arise at larger distances, on the whole. However, the positive asymmetric spatial dependence of precipitation values does not arise at the closest distance, in this case, the distance of 5 km. Instead, the negative asymmetric spatial dependence is pronounced at the distance of 5 km.

The anomaly behavior of the asymmetric spatial dependence at the distance of 5 km might be due to the fact that the number of pairs in this distance is not sufficient to capture the phenomenon where the variability of topography in Baden-Württemberg exhibit a significant pronunciation. This might also be due to the local spatial extent, for instance, the effect of orographic precipitation which is a similar reason to the spatial rank correlation as discussed in section 5.3 of chapter 5.

The precipitation amounts which generally exhibit the positive asymmetric spatial dependence at nearby stations are also affected by seasons. The systematic pattern of domination of the positive asymmetric spatial dependence at the nearby gauges is detected pronounced in the DJF, MAM, and SON seasons for all various time scales from daily to monthly. The positive asymmetric spatial dependence occurs at a distance roughly between 10 km and 35 km. However, in the JJA season, the positive asymmetric spatial dependence tends to disappear, and precipitation values tend to follow the Gaussian symmetric spatial dependence, especially for daily scale. For the higher time scales, such as 5-daily, 15-daily, and monthly, the negative asymmetric spatial dependence even becomes more pronounced.

**Effect of precipitation intensities** The positive asymmetric spatial dependence of precipitation values describes that the high quantiles of precipitation amount are more spatially correlated than the low quantiles. This indicates that precipitation amounts with high intensities will yield the positive asymmetric dependence at nearby stations because the extreme precipitation values tend to occur in a clustered manner at nearby locations as mentioned.

Figure 6.5 and Figure 6.6 show interpolated spatial precipitation distribution for different type of data sets (namely, all data sets, positive asymmetry, and negative asymmetry) in different seasons (namely, DJF, JJA, MAM, and SON). Similar to Singapore, the first data set is all precipitation events and then averaged over all time periods for each gauge station separately in order to get a single realization of spatial precipitation field. The second data set is all precipitation events at which precipitation values follow the positive asymmetric dependence at close distances ($\leq$30 km) and it is also averaged over all time at each gauge station separately in order to obtain a single realization of spatial precipitation fields. The third data set is from the extraction of precipitation events with the negative asymmetric spatial dependence in neighboring distances ($\leq$30 km) and calculated mean value over time series at each gauge station to gain single realization of precipitation fields.

For the case of Bavaria (see Figure 6.5), it is seen that local heavy precipitation in the South of Bavaria triggers more positive asymmetry precipitation values at close distances. By contrast, precipitation amounts in Baden-Württemberg (see Figure 6.6) exhibit anomaly behavior because the asymmetric positive spatial dependence of precipitation amounts also occurs at far distances. This might be caused by the topography feature which is characterized a long series of mountains ranging from North-East to South-West over the region. There is also the mountainous region in the Southeast of Baden-Württemberg affecting this anomaly. High variabilities of topography features in Baden-Württemberg make the spatial stationary assumption questionable.

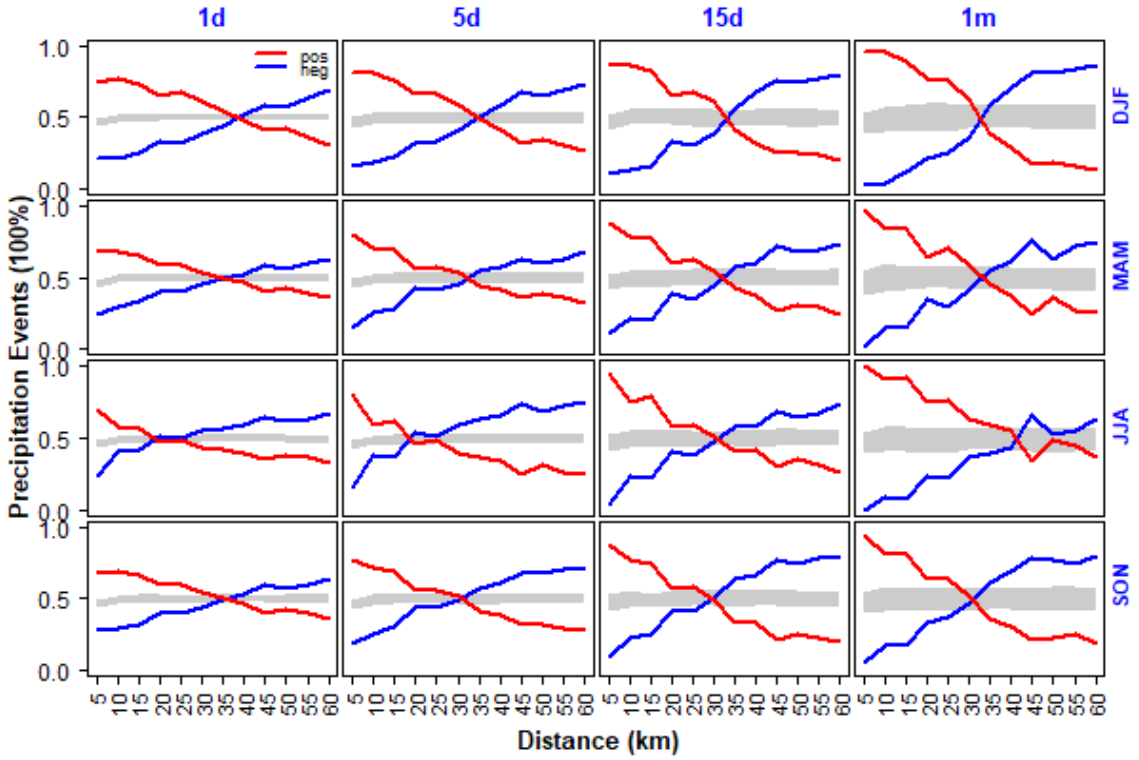



Figure 6.3: Asymmetry function using the geostatistical approach at different temporal scales in Bavaria. The primary vertical axes represent the number of precipitation occurrences presented in 100% scale. The horizontal axes represent various distances (5 km to 60 km). The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The grey shadow areas denote the confidence interval of 90% of the symmetric Gaussian dependence. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent the time scales from daily (1d) to monthly (1m).

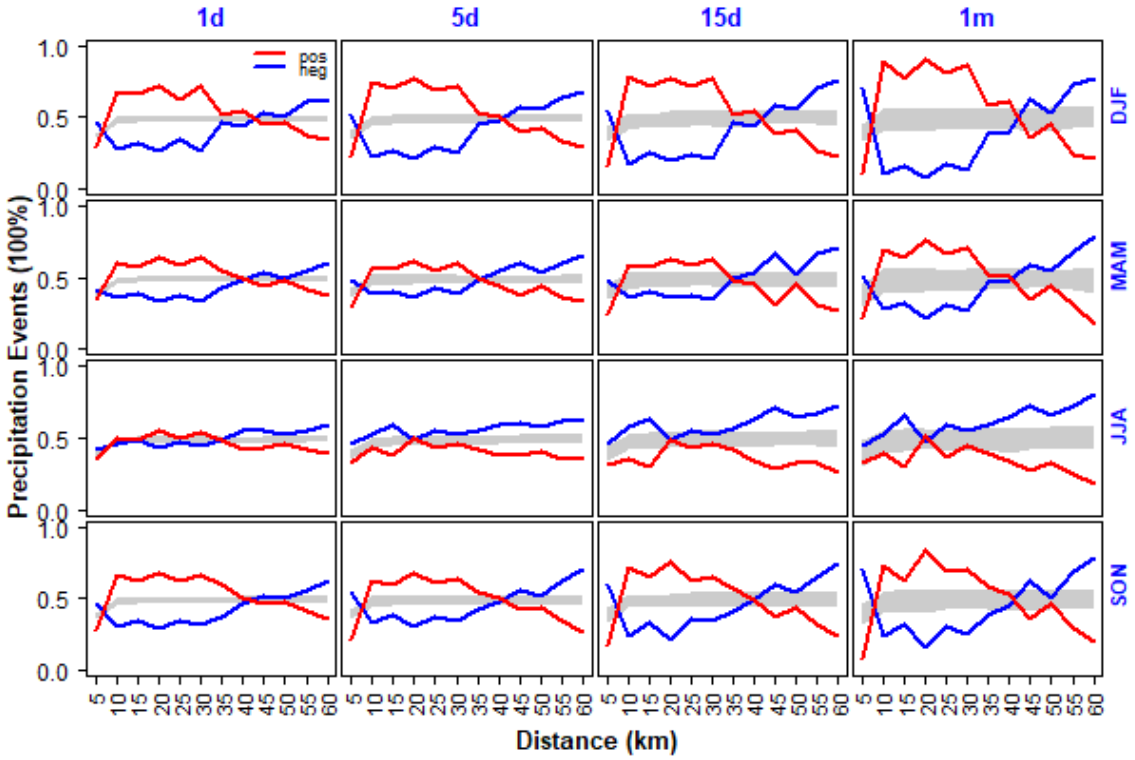



Figure 6.4: Asymmetry function using the geostatistical approach at different temporal scales in Baden-Württemberg. The primary vertical axes represent the number of precipitation occurrences presented in 100% scale. The horizontal axes represent various distances (5 km to 60 km). The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The grey shadow areas denote the confidence interval of 90% of the symmetric Gaussian dependence. Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent the time scales from daily (1d) to monthly (1m).

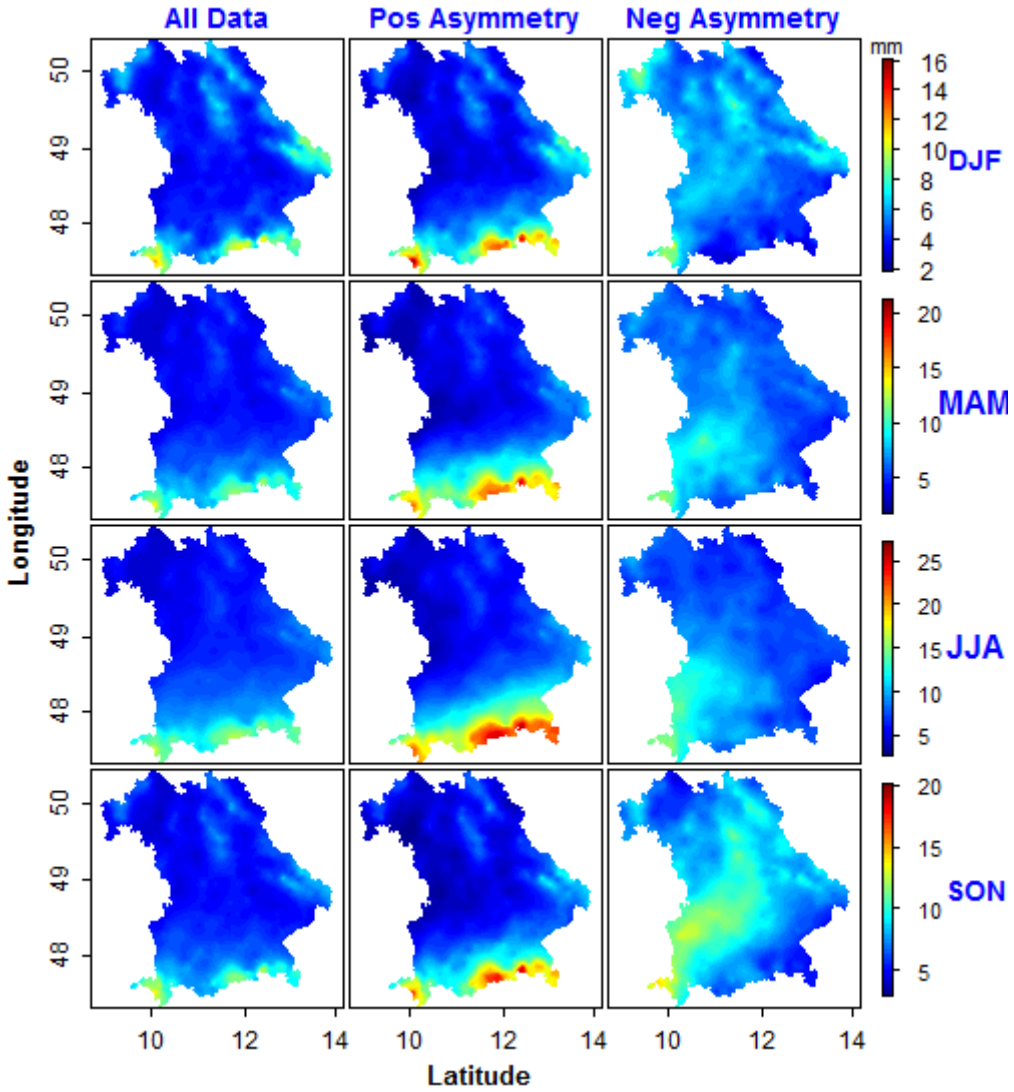



Figure 6.5: Daily precipitation precipitation intensities with different asymmetric dependence in Bavaria. The primary vertical axes represent latitude and the horizontal axes define longitude (both in km). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent all data, positive and negative asymmetric dependence, respectively.

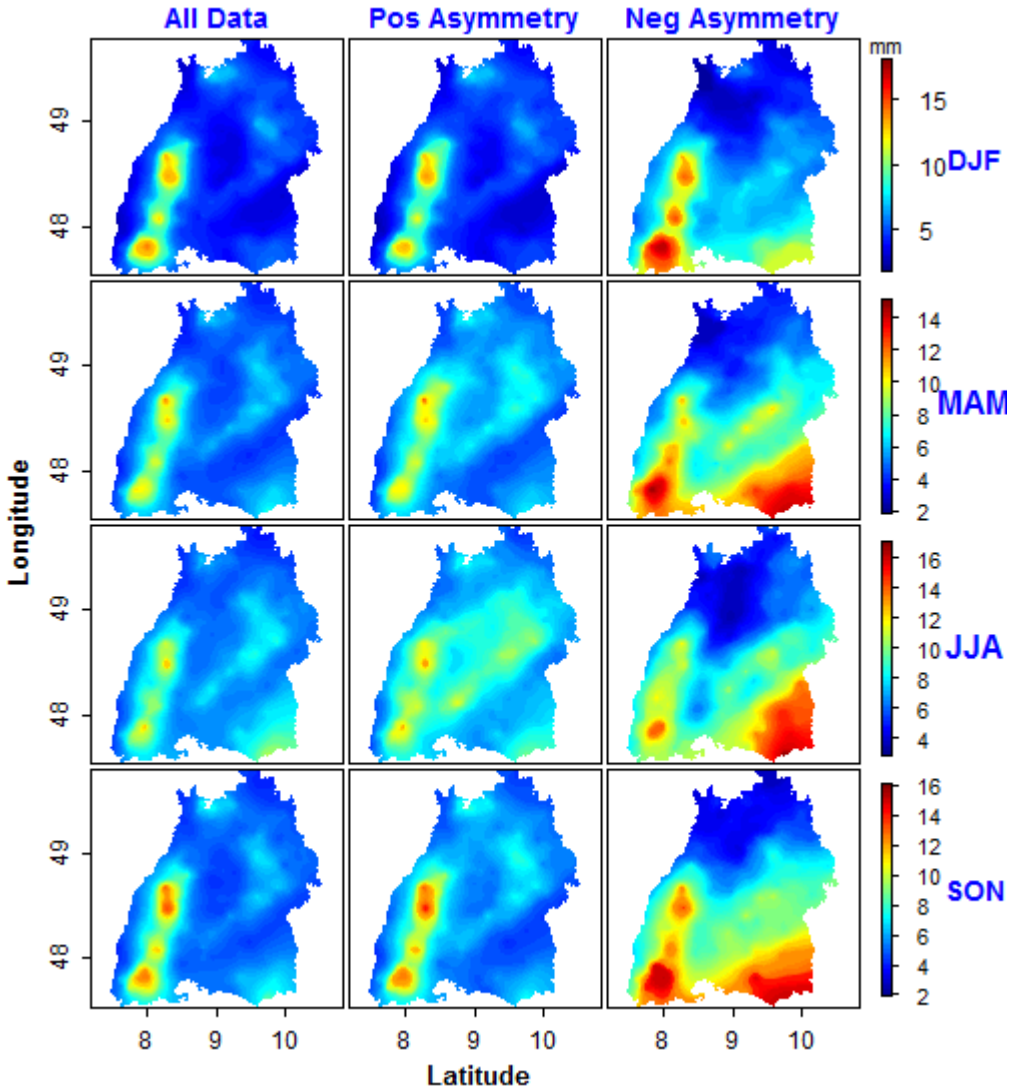



Figure 6.6: Daily precipitation precipitation intensities with different asymmetric dependence in Baden-Württemberg. The primary vertical axes represent latitude and the horizontal axes define longitude (both in km). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent all data, positive and negative asymmetric dependence, respectively.

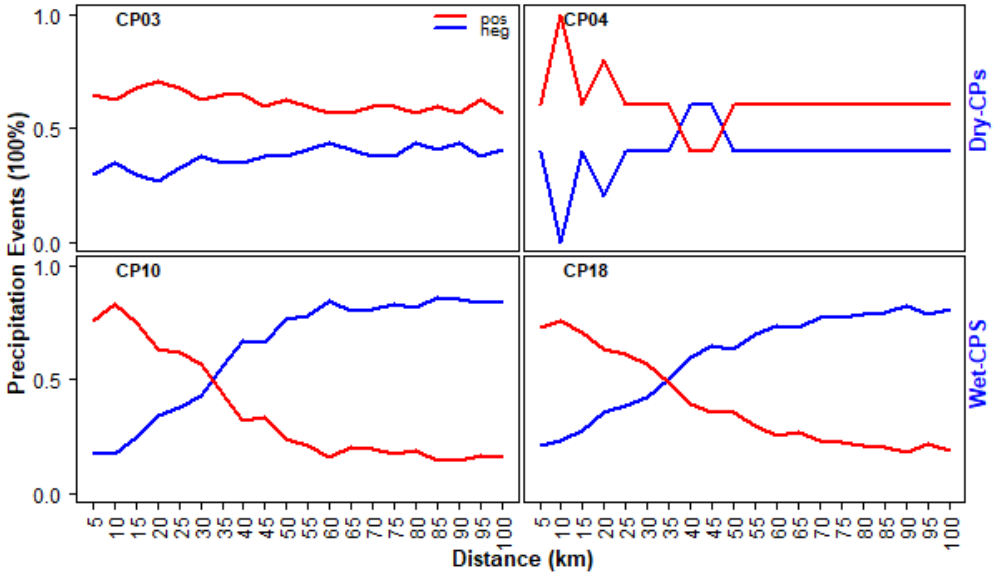
### 6.2.3 Implementation with conditional on circulation patterns

Precipitation events are high variable in space and time because the physical processes of precipitation events are strongly dependent on a very complex mechanism in the atmosphere which also depends on the space and time and reveals an extremely uncertainty of non-linear processes within large scales. Understanding the relationship between the atmosphere processes in huge scales on the one hand and the precipitation in local scales, on the other hand, is of importance to improve the estimation of the precipitation design.

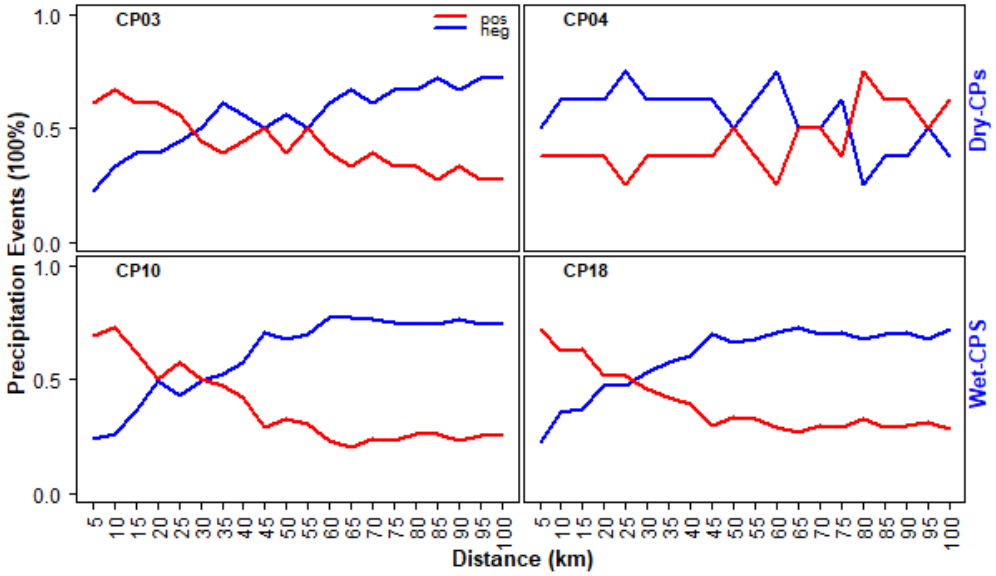
Among the atmospheric variables, the pressure is the driver for flow and transport. Air pressures at the land surface is a variable which can be measured simply with good accuracy. Selected observations and meteorological models provide information on high altitude pressure conditions. Air pressures and geopotential heights are among the best-modeled quantities. Therefore, it is reasonable to use them as a basis to describe atmospheric circulation (Bárdossy et al., 2015).

The classification of atmospheric circulation conditions is required to establish a relationship between large scale (mean air pressure distribution) and local scale variables (hydrological variables). The study of the atmospheric circulation patterns in the southern regions of Germany has been conducted by Bárdossy (2010); Beck and Bárdossy (2013); Liu et al. (2013) using the fuzzy rule-based approach developed by Bárdossy et al. (1995). A classification with 18 different circulation patterns (CPs) in southern Germany (Bárdossy, 2010; Beck and Bárdossy, 2013; Liu et al., 2013) is, therefore, used for this study. The most important remark about specific CPs is that the CP03 is characterized as the driest CP followed by the CP04 and the wettest CP is the CP10 and then followed by the CP18 as the second wettest CP.

Conditional investigation restricted to days with CPs is carried out. Analysis of asymmetric spatial dependence shows that the wettest CPs (CP10 and CP18) and the driest CPs (CP03 and CP04) present a different characteristic of the asymmetric spatial dependence. The positive asymmetric spatial dependence of precipitation amounts, which frequently occur at close distances, are pronounced for the wettest CPs (CP10 and CP18) in two different seasons (namely, November-April and May-October) and in both regions (Bavaria and Baden-Württemberg), as shown in Figure 6.7 and Figure 6.8. In addition, precipitation values which follow the positive asymmetric spatial dependence at nearby stations are higher intensities than the negative ones. In contrast, investigation of asymmetric spatial dependence shows that for the driest CPs (CP03 and CP04), the positive asymmetric spatial dependence is still pronounced at close distances in both seasons (November-April and May-October) for the CP04. For the CP03, the positive asymmetric spatial dependence is dominant in the period of November-April, but not in the period of May-October.

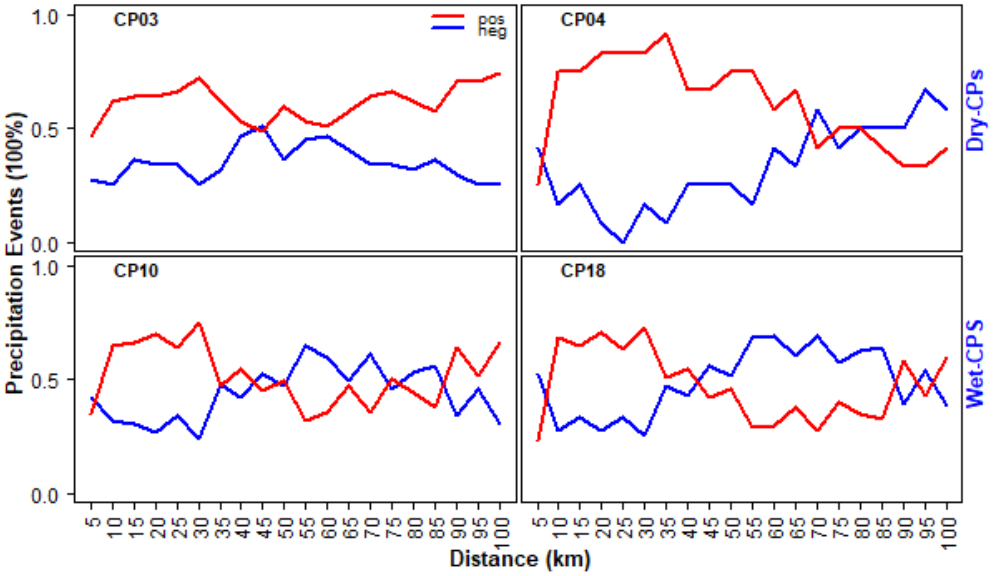



(a) Nov-April

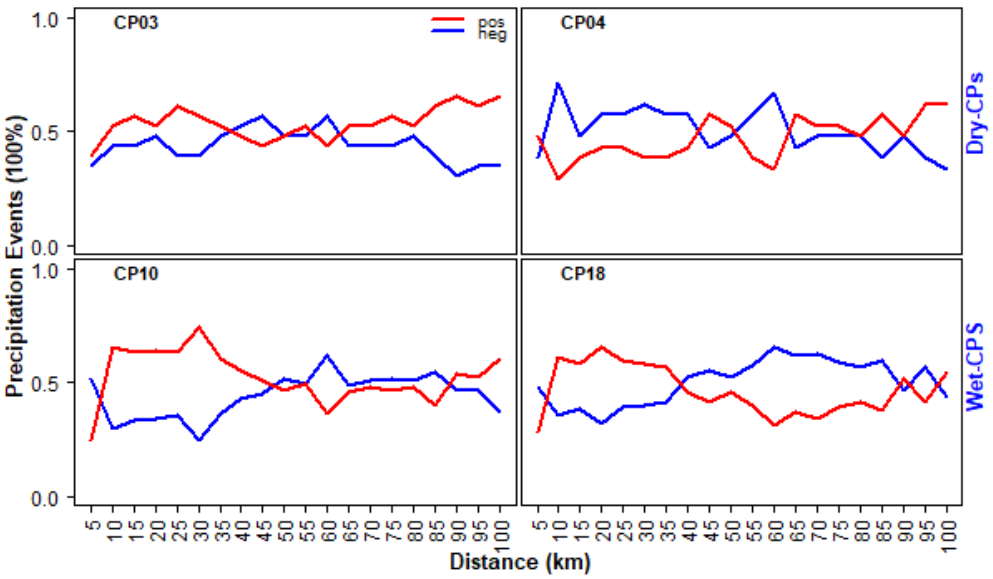



(b) May-Oct

Figure 6.7: Asymmetry function of daily precipitation with different circulation patterns (CPs) in Bavaria. The primary vertical axes represent the number of precipitation events presented in 100% scale. Panels from top to bottom represent the wet and dry CPs. The horizontal axes define various distances. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. Figure (a) is for Nov-April and Figure (b) for May-October.




(a) Nov-April




(b) May-Oct

Figure 6.8: Asymmetry function of daily precipitation with different circulation patterns (CPs) in Baden-Württemberg. The primary vertical axes represent the number of precipitation events presented in 100% scale. Panels from top to bottom represent the wet and dry CPs. The horizontal axes define various distances. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. Figure (a) is for Nov-April and Figure (b) for May-October.

### 6.2.4 Implementation with CRM precipitation data

A spatial precipitation model which is capable of deriving climate change scenarios at a finer spatial resolution has been playing a crucial role in many hydrological applications. Global circulation models (GCMs) provide scenarios for the possible future development of climate, but in a very coarse spatial resolution and thus they cannot be used directly for the assessment of regional consequences of climate change. Regional climate models (RCMs) are promising tools which can provide projected precipitation data with high spatial and temporal resolutions. RCMs are generally obtained by downscaling either statistically or dynamically from the output of the GCMs, that provide climate variables, including precipitation at a finer spatial resolution. Unfortunately, RCM models inherit some of the biases of the GCMs.

Large uncertainties in current global and regional climate models are related to the representation of clouds, moist convection, and complex topography. This is being solved by applying the convection-resolving modeling (CRM) by reducing the grid spacing of climate models down to some few kilometers and switching off the convection parameterization. In addition, CRM has some advantages such as much finer spatial resolutions to some few kilometers and a better representation of topography, surface fields, and boundary layer processes. Thus, CRM has become a promising climate model to generate more realistic precipitation patterns especially in the cases of moist convection and/or over mountainous regions (Ban et al., 2014).

However, an evaluation of performances on this model is necessary with different perspectives and different approaches. Most studies investigating RCMs biases are conducted on the basis of marginal distribution of precipitation at individual sites (for example, mean, variance, extreme events) and cross-correlation between stations. It is also interesting to investigate the biases of RCMs on the basis of asymmetric spatial dependence structures using empirical bivariate copulas. RCM-based precipitation simulations data used in this study, convection-resolving modeling (CRMs).

CRM Data used in this study is classified into three groups, namely historical observations (1951-2001), CRM simulations (1998-2007), and CRM projections (2081-2090). CRMs Data is here a type of CRM data with 2.2 km spatial resolution from a convection-resolving simulation integrated with the COSMO-CLM model (Consortium for Small-Scale Modeling in Climate Mode). Investigation regions are selected based on homogeneity areas with little topography variation. The regions of South Bavaria and South Baden-Württemberg are, therefore, only presented in this report.

**Case study in South Bavaria** Similarly, empirical evidence proves that the positive asymmetric spatial dependence is detected significantly at close distances for the three types of data sets (observations, CRM simulations, and CRM projections) for all different time scales (1d, 5d, 15d, 1m) and all seasons (DJF and JJA) as shown in Figure 6.9. It is clearly seen that in DJF, observed precipitation values exhibit frequently the positive asymmetric spatial dependence at close distances below 35 km for all different time scales. Identically, both CRM simulations and CRM projections also present similar characteristic of asymmetric spatial

dependence which are pronounced at the close distances below 35 km for all different temporal scales. In JJA, precipitation observation and both CRM precipitation simulations and CRM projections also present the positive asymmetric dependence at nearby distances below 20 km for all various temporal scales. Nevertheless, during both seasons (DJF and JJA), the positive asymmetric spatial dependence in Bavaria is pronounced due to the effect of convective precipitation which is strongly influenced by topography features of the high mountains over this region (namely, the Alps and the Swabian Jura).

**Case study in south Baden-Württemberg** By contrast, empirical investigations prove that the positive asymmetric spatial dependence of precipitation amounts in south Baden-Württemberg is detected significantly at distances between 15 and 30 km, and surprisingly, the negative asymmetric spatial dependence is pronounced at close distances below 10 km for DJF season. This applies for all three types of data set (observations, CRM simulations, and CRM projections) and for all different time scales as shown in Figure 6.10. In JJA season, precipitation observation and both CRM precipitation simulations and projections tend to present the symmetric Gaussian dependence at nearby distances for all various temporal scales and almost all distances. This is because typical precipitation processes dominating in this region are strongly influenced by orographic precipitation due to the mountainous region in the Black Forest extending from southwest to northeast resulting higher precipitation amounts on the windward side and lower amounts on the leeward side (Mosthaf and Bárdossy, 2017).

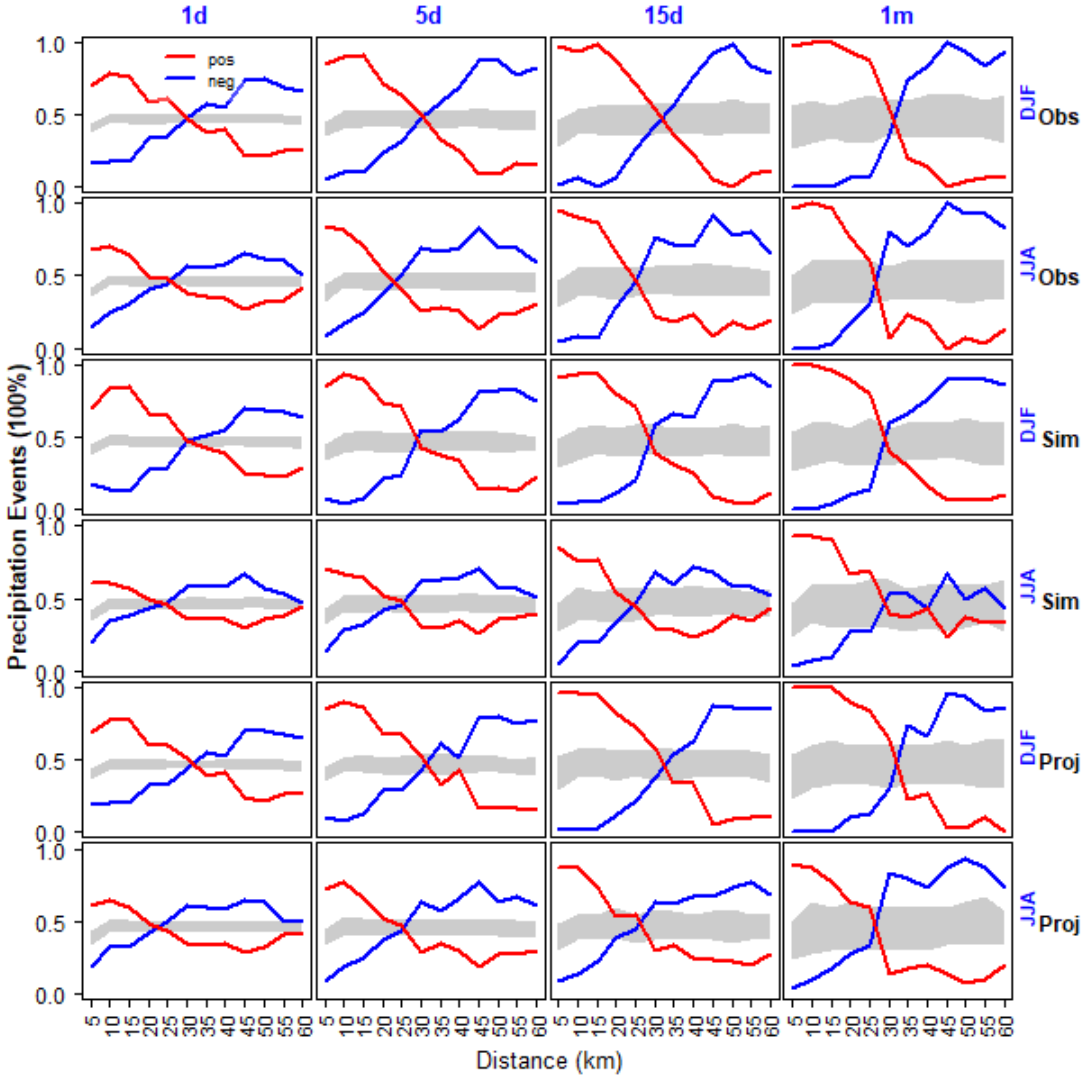



Figure 6.9: Asymmetry function of CRM precipitation data using the geostatistical approach in South Bavaria. The primary vertical axes represent the number of precipitation events presented in 100% scale. Panels from top bottom represent seasons (DJF and JJA) and type of data sets (Obs: observations, Sim: simulations, and Proj: projections). The horizontal axes define various distances. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The grey shadow areas denote the confidence interval of 90% of the symmetric Gaussian dependence.

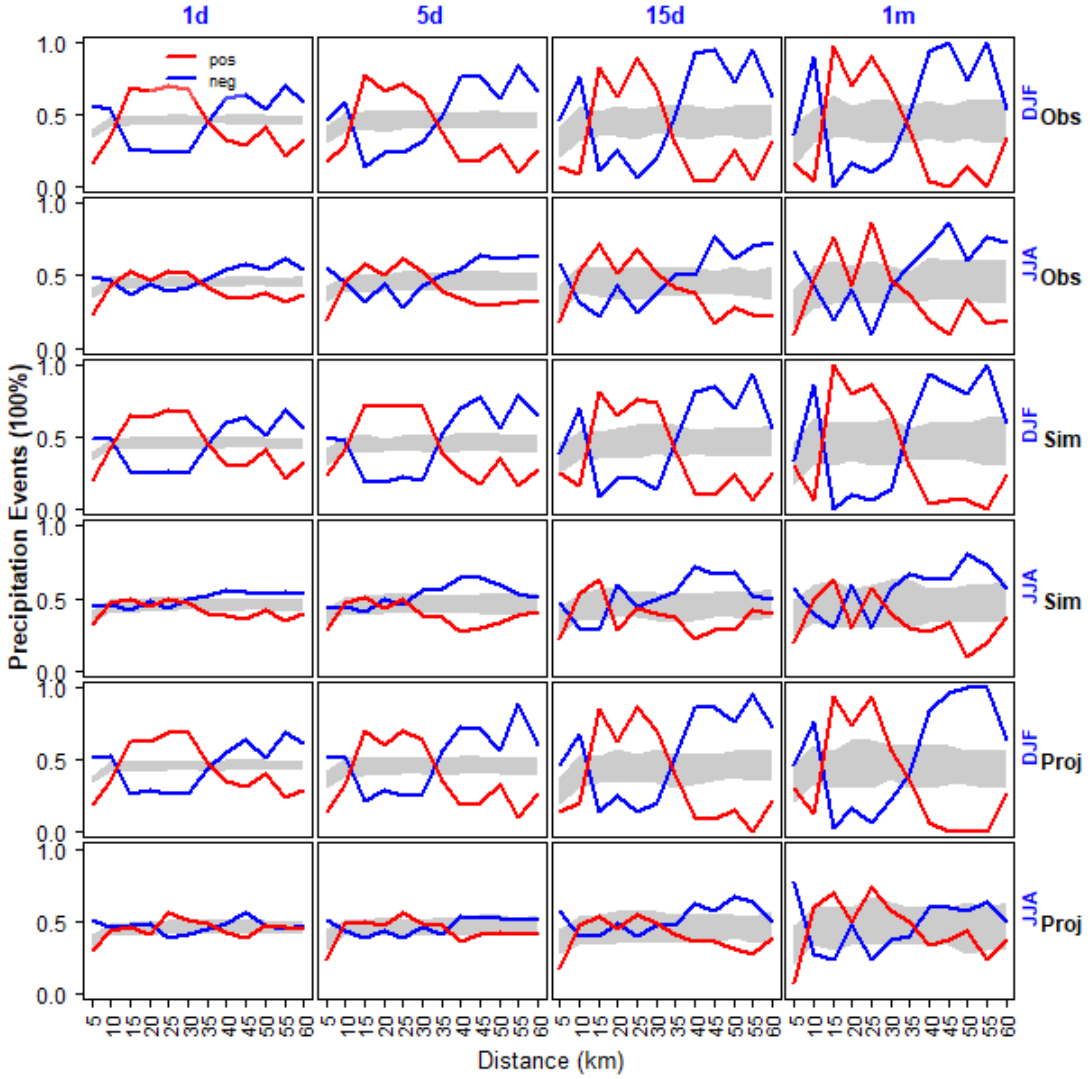



Figure 6.10: Asymmetry function of CRM precipitation data using the geostatistical approach in South Baden-Württemberg. The primary vertical axes represent the number of precipitation events presented in 100% scale. Panels from top bottom represent seasons (DJF and JJA) and type of data sets (Obs: observations, Sim: simulations, and Proj: projections). The horizontal axes define various distances. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The grey shadow areas denote the confidence interval of 90% of the symmetric Gaussian dependence.

## 6.3 Temporal investigation simultaneous occurrences of high and low precipitation

Unlike investigations of asymmetric spatial dependence using the geostatistical approach as mentioned in section 6.2, where precipitation amounts for each time step are treated separately using the assumption of a single realization of spatial random function, in this section, investigations of asymmetric spatial dependence are conducted on the basis of a pair of stations, which are located at the exact geographical location.

This study is carried out to investigate the simultaneous temporal occurrence of high and low precipitation amounts between two precipitation gauge stations of interest. This would determine whether those precipitation values between two locations present the symmetric Gaussian spatial dependence between high and low values and whether the Gaussian spatial dependence based model is a reasonable model for infilling missing data values. In this approach, empirical bivariate copulas are constructed for each pair of possible combination of two stations containing contemporaneous precipitation observations as same as approach described in section 5.4 of chapter 5. Thus, the total number of empirical bivariate distribution functions, which are investigated, is all possible pair combinations of stations located at different locations.

The asymmetry function $A_3$ for each pair of gauge stations is calculated on the basis of empirical bivariate copulas using Equation 4.20 presented in section 4.4 of chapter 4. If the asymmetry function $A_3$ is positive, then precipitation amounts between two stations of interest follow the positive asymmetric spatial dependence. If $A_3$ is negative, then it presents the negative asymmetric spatial dependence. If $A_3$ is zero, it shows the symmetrical spatial dependence.

Unlike in the geostatistical approach (see section 6.2), the Gaussian simulations for measuring the degree of uncertainty with 90% significance level is applied for each empirical bivariate copula of a pair of stations. The detail of investigation procedure and the statistical testing are described as follows.

1. For any pair of gauge stations, selected time scale, season, and quantile threshold $q_0$, the empirical asymmetric dependence is calculated using Equation 4.20 on the basis of the empirical bivariate copula of precipitation variables.
2. The precipitation variables at each the pair of gauge stations are transformed to the Gaussian domain using the normal score transformation.
3. Based on the statistical parameters obtained in step-1 above, 100 realizations of bivariate normal distributions are generated.
4. For each realization of 100 Gaussian simulations, the asymmetric dependence function is calculated using Equation 4.20.
5. For all 100 Gaussian realizations, the quantile 5% and 95% for lower and upper limit of confidence intervals of the asymmetric dependence function can be calculated.

6. If the empirical asymmetry function lies between the confidence limits then mark it as the symmetrical dependence.
7. If the empirical asymmetry function is bigger than the upper limit, then it is noted as the positive asymmetric dependence and conversely, if the empirical asymmetry function is less than the lower limit, then it is noted as the negative asymmetric dependence.
8. The steps above are conducted repeatedly again using the different quantile thresholds ($q_0$= 0.1, 0.2, 0.3, and 0.4) and the different seasons (the months of DJF, MAM, JJA, SON), and also different time scales ranging from hourly to monthly for Singapore and from daily to monthly for both Bavaria and Baden-Württemberg.
9. Finally, the number of pairs of gauge stations showing positive asymmetric, negative asymmetric, and symmetric dependence among all pairwise possible combinations are summarized for each group distance, such as 5 km, 10 km, other length of distances with an increment of 5 km.

### 6.3.1 Implementation in Singapore

In this part, investigation of the asymmetric spatial dependence of simultaneous temporal occurrence of high and low precipitation between two precipitation gauge stations of interest in the regions of Singapore are conducted. The empirical investigation results are shown in Figure 6.11. Figure 6.11 shows multiple plots presenting the number of pairs of stations which are classified as positive, negative, and symmetric Gaussian spatial dependence. The Gaussian simulations with confidence interval 90% are carried out for measuring the degree of uncertainty. This study is conducted using a variety of time scales (namely, 1 h, 3 h, 6 h, 1 d, 5 d, 15 d, and 1 m) and using in different seasons (namely, DJF, MAM, JJA, and SON).

The number of pairs of stations which significantly follow the positive asymmetric spatial dependence are significantly higher in comparison to the negative asymmetric spatial dependence as well as the symmetric Gaussian dependence. However, this applies for the lower time scales, for example from hourly to daily as shown in Figure 6.11. At these time scales (hourly to daily), not even any pair of stations can be classified as having a negative asymmetric dependence or a symmetric Gaussian dependence. This empirical evidence applies for all seasons (namely, DJF, MAM, JJA, and SON) with small differences in the magnitudes.

In the season of DJF for example, 100% of total pairs of stations exhibit significantly the positive asymmetric spatial dependence at the hourly scale (1 h) at the distance of 5 km. For 3 h, it presents substantially the positive asymmetric spatial dependence at the distances of 5 km and 10 km. For 6 h, it presents significantly the positive asymmetric spatial dependence at the distances of 5 km, 10 km, 15 km, and 20 km. For 12 h and 1 d scale, it shows substantially the positive asymmetric spatial dependence at the distances of 5 km, 10 km, 15 km, 20 km, and 25 km.

For the conditions in which not 100% of pairs of stations can be grouped as the positive

asymmetric dependence because the probability of zeros for all stations are higher than the quantile threshold 0.2. Consequently, the asymmetry function cannot be analyzed due to the lack of knowledge in that region of the quadrant. Note that, taking into zeros as latent variables are possible when the probability of zero precipitation amounts is less than the specific threshold $q_0$. The zero density is known, but the exact values are treated hidden as censored values.

On the higher time scales, such as 5-daily, 15-daily, and monthly, the number of pairs of stations which follow the positive asymmetric spatial dependence decrease systematically. For instance, on the 5 d scale, the number of pairs of stations for the negative asymmetric dependence is higher, roughly ranging from 55-65% of all pairs for distances between 5 km and 25 km, followed by the positive asymmetric dependence (31-42%) and the symmetric Gaussian dependence (0-7.4%). On the 10 d scale, for another example, the number of pairs of stations for the positive asymmetric dependence is slightly higher, roughly ranging from 50-62% of all pairs, followed by the negative asymmetric dependence (19-26%) and the symmetric Gaussian dependence (14-29%). For the monthly scale, for example, the number of pairs of stations for the negative asymmetric dependence is slightly higher, roughly ranging from 33-62% of all pairs, followed by the negative asymmetric dependence (18-26% ) and the symmetric Gaussian dependence (17-24%).

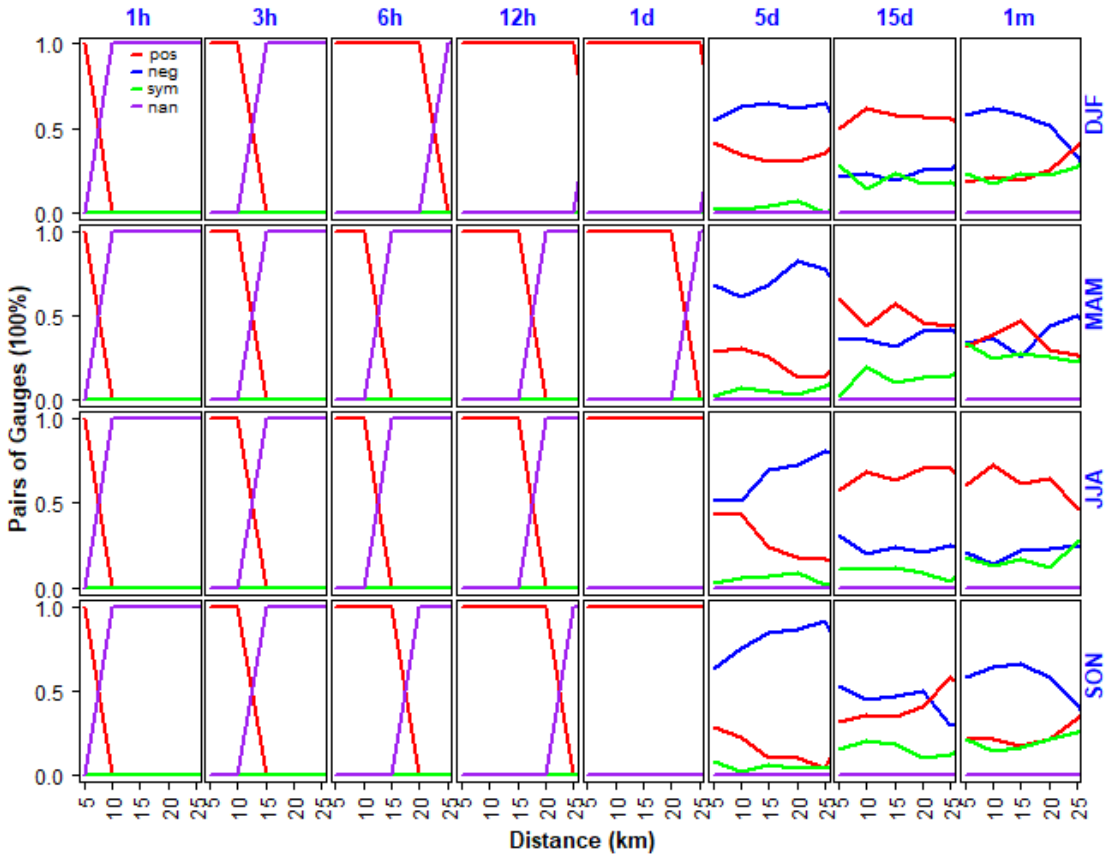



Figure 6.11: Asymmetry function using pairwise approach with the threshold quantile $q_0$ 0.2 in Singapore. The primary vertical axes are the portion of pairs of stations presented in 100% scale. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The symmetric Gaussian dependence is denoted by the green lines. The purple lines indicate the quantile threshold $q_0$ is less than the probability of zeros and marked as not a number (NaN). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent temporal scales (hourly (1h) to monthly (1m). The horizontal axes represent distances.

### 6.3.2 Implementation in Baden-Württemberg

The empirical investigation results of the asymmetric spatial dependence of simultaneous temporal occurrence of high and low precipitation between two precipitation gauge stations of interest in the region of Baden-Württemberg are shown in Figure 6.12. Figure 6.12 shows multiple plots presenting the number of pairs of stations which are classified significantly as positive, negative, and symmetric Gaussian spatial dependence. The study is conducted using a variety of time scales of precipitation amounts (namely, 1d, 5d, 15d, and 1m scale) and using different seasons (namely, DJF, MAM, JJA, and SON).

Similarly to the results for Singapore, the number of pairs of stations which follow the positive asymmetric spatial dependence are totally dominant in comparison to the negative asymmetric spatial dependence as well as the symmetric Gaussian dependence, especially for the daily scale. Not even any pair of gauge stations can be classified as negative asymmetric dependence or symmetric Gaussian dependence for the daily scale as shown in Figure 6.12. This empirical evidence applies for all seasons (DJF, MAM, JJA, and SON).

However, on the higher time scales, such as 5-daily, 15-daily, and monthly, the number of pairs of stations which exhibit the positive asymmetric spatial dependence of precipitation amounts decrease systematically. For instance, on the 5 d scale, the number of pairs of stations for the negative asymmetric dependence is higher, roughly ranging from 55-65% of all pairs for the distance of 5-80 km, followed by the positive asymmetric dependence (31-42%) and the symmetric Gaussian dependence (0-7.4%). On the 15 d scale, for another example, the number of pairs of stations for the positive asymmetric dependence is slightly higher, roughly ranging from 50-62% of all pairs, followed by the negative asymmetric dependence (19-26%) and the symmetric Gaussian dependence (14-29%). For the monthly scale, the number of pairs of stations for the negative asymmetric dependence is slightly higher, roughly ranging from 33-62% of all total pairs, followed by the negative asymmetric dependence (18-26%) and the symmetric Gaussian dependence (17-24%).

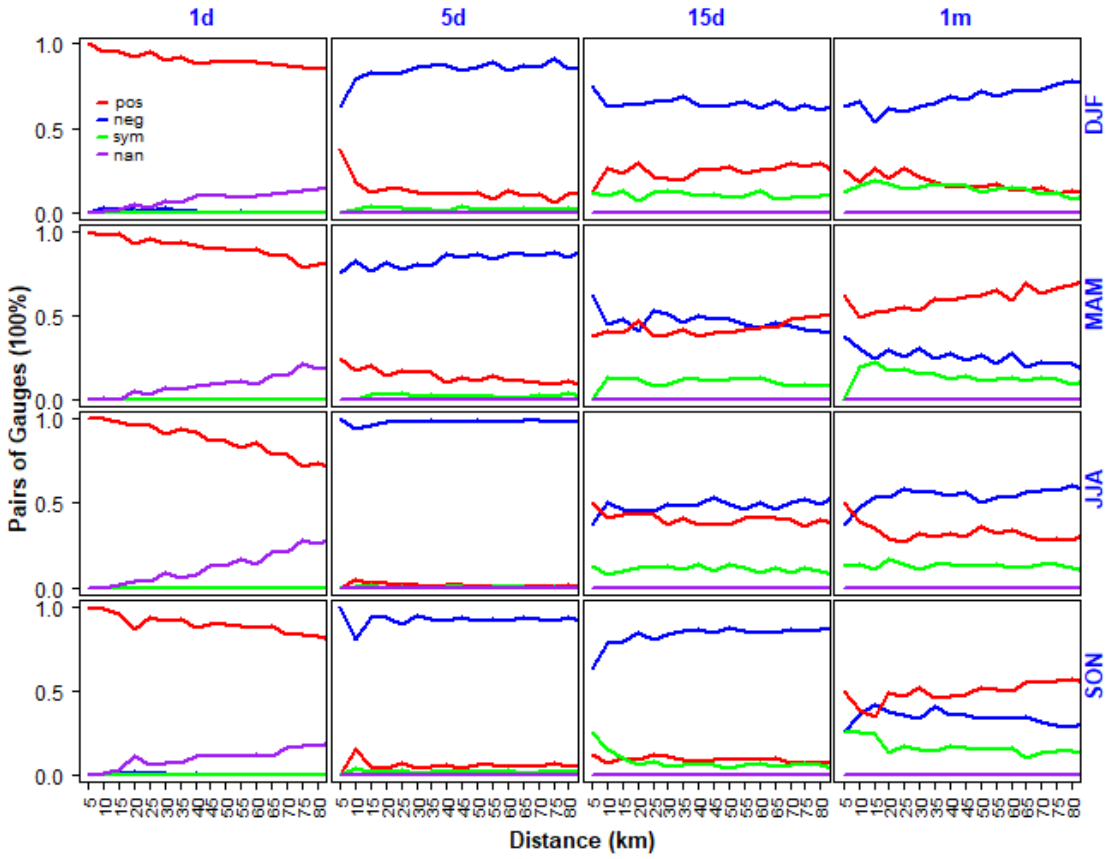



Figure 6.12: Asymmetry function using pairwise approach with the threshold quantile $q_0$ 0.2 in Baden-Württemberg. The primary vertical axes are the portion of pairs of stations presented in 100% scale. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The symmetric Gaussian dependence is denoted by the green lines. The purple lines indicate the quantile threshold $q_0$ is less than the probability of zeros and marked as not a number (NaN). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent temporal scales (daily (1d) to monthly (1m). The horizontal axes represent distances.

### 6.3.3 Implementation in Bavaria

The empirical investigation results of asymmetric spatial dependence of temporal simultaneous occurrence of high and low precipitation between two precipitation gauge stations of interest in the regions of Bavaria are shown in Figure 6.13, displaying multiple plots presenting the number of pairs of gauges which are classified as positive, negative, and symmetric Gaussian spatial dependences after it is significantly tested using the Gaussian simulations for measuring degree of uncertainty. The study is conducted using a variety of time scales of precipitation amount (namely, 1d, 5d, 15d, and 1m), and using different seasons (namely, DJF, MAM, JJA, and SON).

The number of pairs of gauges at different locations which substantially exhibit the positive asymmetric spatial dependence of precipitation amounts are totally dominant in comparison to the negative asymmetric dependence as well as the symmetric Gaussian dependence for daily scale. Not even one pair of gauges can be classified as negative asymmetric dependence or symmetric Gaussian dependence at the daily scale as shown in Figure 6.13. This empirical evidence applies for all seasons (DJF, MAM, JJA, and SON) with small differences in the magnitude.

However, on the higher time scales, such as 5-daily, 15-daily, and monthly, the negative asymmetric spatial dependence of precipitation amounts becomes more dominant regarding the number of pairs of stations. For instance, on the 5 d scale, the number of pairs of stations for the negative asymmetric dependence is higher, roughly ranging from 54-90% of all pairs for distances between 5 and 80 km, followed by the positive asymmetric dependence (6-39% ) and the symmetric Gaussian dependence (3-8%). On the 15 d scale, the number of pairs of stations for the positive asymmetric dependence is slightly higher, roughly ranging from 49-71% of all pairs, followed by the negative asymmetric dependence (14-33%) and the symmetric Gaussian dependence (12-19%). For the monthly scale, the number of pairs of stations for the negative asymmetric dependence is slightly higher, roughly ranging from 33-76% of all pairs, followed by the negative asymmetric dependence (7-34%) and the symmetric Gaussian dependence (15-37%).

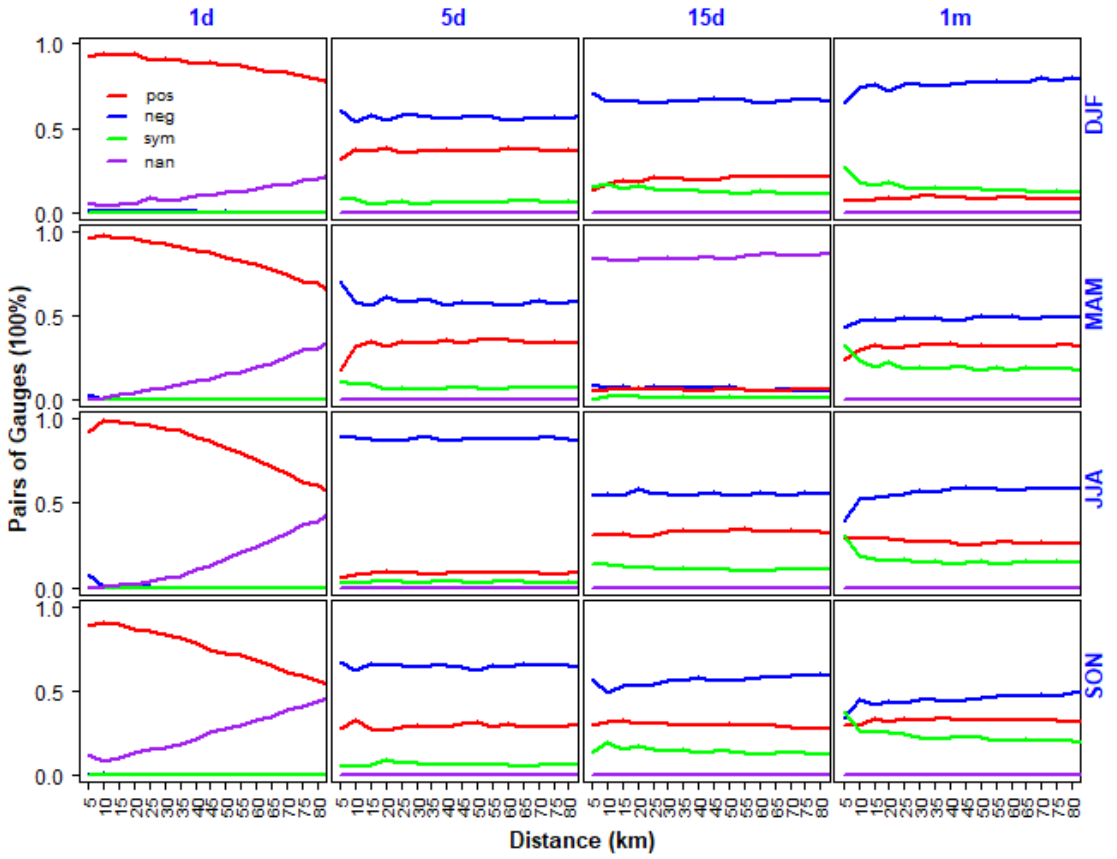



Figure 6.13: Asymmetry function using pairwise approach with the threshold quantile $q_0$ 0.2 in Bavaria. The primary vertical axes are the portion of pairs of stations presented in 100% scale. The red lines indicate the positive asymmetric dependence and the blue lines represent the negative asymmetric dependence. The symmetric Gaussian dependence is denoted by the green lines. The purple lines indicate the quantile threshold $q_0$ is less than the probability of zeros and marked as not a number (NaN). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). Panels from left to right represent temporal scales (daily (1d) to monthly (1m). The horizontal axes represent distances.

## 6.4 Summary

- Empirical investigation results reveal that, precipitation values, which exhibit the positive asymmetric spatial dependence frequently occur at nearby stations and the negative asymmetric dependence appears significantly at far distances. Precipitation values with positive asymmetric spatial dependence decrease with the increase of distances, and on the contrary, precipitation amounts which exhibit the negative asymmetric dependence increase with the increase of distances.
- In Singapore, the number of precipitation events at which precipitation values exhibit the asymmetric spatial dependence, is significantly dominant at the higher time resolutions and then decreases with the increase of temporal scales. In Bavaria and Baden-Württemberg, this pattern applies to all time scales ranging from daily to monthly.
- The positive asymmetric spatial dependence of precipitation values conditioned by the circulation patterns more frequently occur at close distances for the wettest CPs for both regions, Bavaria and Baden-Württemberg. In contrast, the positive asymmetric dependence for the driest CPs (CP03 and CP04) is pronounced at close distances for the CP04. For the CP03, the positive asymmetric spatial dependence is dominant only in the season of November-April.
- The positive asymmetric spatial dependences of the three type of data sets (observations, RCM simulations, and RCM projections in the Southern part of Bavaria) are also detected significantly at close distance for all different time scales. By contrast, the positive asymmetric spatial dependence in the Southern part of Baden-Württemberg is detected significantly at a distance of 15-30 km for DJF season.
- Investigation results of the simultaneous temporal occurrence of high and low precipitation between two precipitation gauge stations reveal that, the number of pairs of stations which significantly exhibit the positive asymmetric dependence are dominant in the regions of Singapore (from hourly to daily), Baden-Württemberg (daily) and Bavaria (daily).

# 7 Spatial copula models based on asymmetric dependence

## 7.1 Introduction

Multivariate analysis based models play a significant role in a variety of applications, as the dependence between two or more random variables is detected significantly. Thus, using a simple univariate approach is not capable of modelling it reasonably. As a result, it could yield to either an underestimation or overestimation of the risks. Multivariate normal models have been a widely applied in many fields for a long time. Multivariate normal distributions are attractive due to their simplicity because both the conditional and the marginal distributions are also normal (Favre et al., 2004).

However, a multivariate normal distribution is not suitable to model for most cases in a more realistic way because marginal distributions are often skewed and have a heavy tail. In addition, the dependence structure is generally different from the Gaussian case which is described only by Pearsons correlation coefficient as discussed in chapter 4. The drawbacks of the traditional multivariate normal model can be tackled by using a copula approach.

Multivariate analysis based on a copula approach has become very popular being adopted in various applications over the last decades. In the applications of spatial geostatistics, Bárdossy (2006) is one of the first who introduced and implemented the copula approach in a geostatistical context, then followed by Gräler (2014); Kazianka and Pilz (2011); Quessy et al. (2016); Wasko et al. (2013). Unlike traditional geostatistics which is restricted to the underlying Gaussian assumption, copulas are more flexible. Moreover, traditional spatial geostatistics assumes that variogram or a covariance function describing spatial variability, is just calculated as an integral over the whole distribution of the parameter values. This is in contradiction to the reality that different percentile values can have a different spatial dependence structure from the central values and the variograms are sensitive to the outliers (Bárdossy, 2006).

In this chapter, two different copula families, namely, Gaussian copulas and V-copulas are implemented in the context of spatial geostatistics. The V-copulas are constructed from the Gaussian copulas through a non-monotonic transformation. Both models are implemented in two different climate regions, Singapore and Bavaria. Characteristics of the properties of both models are discussed in the next sections. Random precipitation variables with different temporal scales, such as sub-daily and super-daily are used in this analysis.

## 7.2 Gaussian copulas

The advantage of the multivariate normal is that all higher order correlations are indirectly defined by the second order moments, namely the correlation coefficients. These models allow for a lot of flexibility in high dimensions since each pair is specifically parameterized (Quessy et al., 2016). In a simulation, generation of the multiple sites replicates of the field variables is readily achieved by generating properly associated multivariate normal variates which are back-transformed to synthetic field variables (Bárdossy and Pegram, 2009). This is another advantage adopting a Gaussian copula for multivariate modeling.

The Gaussian copula belongs to the elliptical copulas' family, for which the dependence structure is completely determined by the correlation matrix. This model is thus convenient and very simple to use to estimate the parameter by means of a symmetric and definite positive matrix (Favre et al., 2004; Renard and Lang, 2007). The fact that these dependence structures are indexed by a correlation matrix is helpful in applications such as spatial statistics, where distances between locations can easily be incorporated (Quessy et al., 2016).

Departure from Equation 4.13, the Gaussian copula can be formulated as follows:

$$C^n(u_1, ..., u_n) = \Phi^n\Big(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)\Big) \tag{7.1}$$

where $\Phi^{-1}$ is the inverse of the cdf of the standard normal distribution $N(0, 1)$ and $\Phi^n$ in this case is the cdf of a multivariate normal distribution with mean 0 and covariance matrix $\Sigma$.

The Gaussian copula density as a result of the differential of Gaussian copula can thus be written in the following formula.

$$\begin{aligned} c^n(u_{12}, ..., u_n) &= \frac{\partial^n C^n(u_1, ..., u_n)}{\partial u_1 ... \partial u_n} \\ &= \frac{\phi^n\Big(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)\Big)}{\prod_{i=0}^{n} \Phi^{-1}(u_i)} \end{aligned} \tag{7.2}$$

Where $\phi^{-1}$ is the pdf of a multivariate normal distribution with mean 0 and covariance matrix $\Sigma$ and $\phi$ is the pdf of the standard normal distribution $N(0, 1)$ given in the following formulas.

$$\phi^n(x_1, ..., x_n) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\Big[-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)\Big] \tag{7.3}$$

$$\phi(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\Big[-\frac{(X-\mu)^2}{2\sigma^2}\Big] \tag{7.4}$$

The Gaussian copula density can be written in another formula as follows

$$c^n(u_1, ..., u_n) = \frac{1}{\sqrt{|\Gamma|}} \exp\Big[-\frac{1}{2} X^T(\Gamma^{-1} - I)X\Big] \tag{7.5}$$

Where $I$ is identity matrix, $\Gamma$ is correlation matrix, $|\Gamma|$ is determinant of correlation matrix.

## 7.3 V-transformed Gaussian copulas

The Gaussian copula obviously cannot capture a tail dependence structure of multivariate distribution or a non-symmetric dependence structure, in particular between low and high values. If in a real world data set shows an asymmetric dependence structure, a non-Gaussian copula-based model is a more realistic way to mimic this nature.

The Gumbel copula is an alternative to model a positive non-symmetric dependence structure. The multivariate extension using the Gumbel copula to n dimensions, however, does not model the processes reasonably. This is because, instead of maintaining strong interdependence, the dependence structure deteriorates with increasing dimensions (Bárdossy and Pegram, 2009; Drouet-Mari, D. and S. Kotz, 2001). In contrast, meta-elliptical copulas do not suffer from this problem.

Bárdossy (2006) derived a non-Gaussian copula from the non-centered multivariate chi-square distribution. Bárdossy and Li (2008) introduced a quite similar approach, the so-called $V$-transformed multivariate normal copulas, in order to take into account the existence of non-symmetric dependence structure behaviour. The unique advantage of underlying the multivariate normal is that all higher order correlations are implicitly defined by only the second order moments.

The $V$-transformed normal copulas can simply be constructed by the non-monotonic transformations of the multivariate standard normal distributions (Bárdossy and Li, 2008). The $V$-transformed $n$-variate variable $X = (X_1, ..., X_n)$ is obtained from a normally distributed $n$-variate random variable $Y = (Y_1, ..., Y_n)$ and $Y \sim N(0, 1)$ by the following formulas.

$$X_j = \begin{cases} k(Y_j - m) & \text{if } Y_j \geq m \\ m - Y_j & \text{otherwise} \end{cases} \tag{7.6}$$

Where $k \in (0, \infty)$ and $m \in R$.

The cumulative distribution function of univariate $V$-copula thus can be expressed as a function of standard normal distribution $\Phi$ as follows.

$$\begin{aligned} H(x_i) &= P(X_i \leq x_i) \\ &= P(Y \leq Y \leq \frac{x_i}{k} + m) \\ &= P(Y \leq \frac{x_i}{k} + m) - P(Y \leq m - x_i) \\ &= \Phi(\frac{x_i}{k} + m) - \Phi(m - x_i) \end{aligned} \tag{7.7}$$

The density function of univariate $V$-copula can be written also as a function of density standard normal distribution as follows.

$$\begin{aligned} h(x_i) &= \frac{\partial H(x_i)}{\partial x} \\ &= \frac{1}{k}\phi(\frac{x_i}{k} + m) + \phi(m - x_i) \end{aligned} \tag{7.8}$$

The multivariate cumulative distribution function of $V$-transformed Gaussian copula with $n$ dimensions can be formulated as follows.

$$\begin{aligned} H^n(x_1, ..., x_n) &= P(X_i \leq x_i, ..., X_n \leq x_n) \\ &= \sum_{i=0}^{2^n-1} (-1)^i \Phi(\zeta_i + m) \end{aligned} \tag{7.9}$$

where

$$\zeta_i^T = \Big( b\big((-1)^{i1}\big) x_1, ..., b\big((-1)^{in}\big) x_n \Big) \tag{7.10}$$

with

$$b(t) = \begin{cases} -1 & \text{if } t \leq 0 \\ \frac{1}{k} & \text{otherwise} \end{cases} \tag{7.11}$$

and $i_j$=0 or 1 and

$$i = \sum_{j=0}^{n-1} i_j 2^j \tag{7.12}$$

The density of multivariate $V$-transformed Gaussian copula with $n$ dimensions can be described as follows.

$$h^n(x_1, ..., x_n) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \sum_{i=0}^{2^n-1} \frac{1}{k^n - \sum_{j=0}^{n-1} i_j} \exp\Big[ -\frac{1}{2} (\zeta_i + m)^T \Sigma^{-1} (\zeta_i + m) \Big] \tag{7.13}$$

Finally, the density of multivariate $V$-copulas can be calculated as follows.

$$\begin{aligned} c^n(u_1, ..., u_n) &= \frac{h^n(x_1, ..., x_n)}{\prod_{i=0}^n h(x_i)} \\ &= \frac{h^n(H^{-1}(u_1), ..., H^{-1}(u_n))}{\prod_{i=0}^n h(H^{-1}(u_i))} \end{aligned} \tag{7.14}$$

Where $c^n$ is density copulas with $n$ dimensions, $h^n$ is density of multivariate $V$-transformed multivariate normal with $n$ dimensions, $h$ is univariate $V$-transformed density of marginal distribution, $H^{-1}$ is the inverse of the cdf of the v-transformed standard normal distribution.

## 7.4 Parameter estimation

### 7.4.1 High dimensional issue

The parameter estimation of multivariate $V$-copulas in high dimensions requires a very large computational effort. In order to illustrate how huge computation is, two dimensions of a $V$-copula (bivariate copula) is taken just for example. From the Equation 7.9, the

bivariate $V$-transformed standard normal distribution can be written as follows.

$$\begin{aligned} H^2(x_1, x_2) = \Phi^2(\frac{x_1}{k} + m, \frac{x_2}{k} + m) - \Phi^2(\frac{x_1}{k} + m, x_2 + m) \\ - \Phi^2(x_1 + m, \frac{x_2}{k} + m) + \Phi^2(-x_1 + m, -x_2 + m) \end{aligned} \tag{7.15}$$

Then, the density of the bivariate $V$-transformed standard normal distribution can be derived easily as:

$$\begin{aligned} h^2(x_1, x_2) = \frac{1}{k^2}\phi^2(\frac{x_1}{k} + m, \frac{x_2}{k} + m) + \frac{1}{k}\phi^2(\frac{x_1}{k} + m, x_2 + m) \\ + \frac{1}{k}\phi^2(x_1 + m, \frac{x_2}{k} + m) + \phi^2(-x_1 + m, -x_2 + m) \end{aligned} \tag{7.16}$$

where $H^2$ is the bivariate $V$-transformed standard normal distribution, $h^2$ is the density of the bivariate $V$-transformed standard normal distribution,$\Phi^2$ is the cdf of the bivariate standard normal distribution, and $\phi^2$ is the density of the bivariate standard normal distribution, $m$ and $k$ are parameters of the bivariate$V$-transformed standard normal distribution.

From the Equation 7.15 and the Equation 7.16 above, one can see that the bivariate $V$-transformed distribution ($H^{d=2}$) consist of the summation over 4 terms from the $2^{d=2}$. Moreover, for the 30 dimensions ($H^{d=30}$) like in the Singapore (d=30 stations), it yields 1073741824 terms due to $2^{30}$ and the computational effort becomes significantly costly following the non linear relationship with the dimensions.

In order to handle this problem, the observation set is divided into several subsets $S_k$ of arbitrary sizes (Bárdossy and Li, 2008). The sets $S_k$ can be selected randomly in order to avoid possible biases caused by preferential sampling. However, each subset $S_k$ consists of $n(k)$ observation points ranging from 2-12 points in a clustered points (Bárdossy and Pegram, 2013). Each subset can be formulated as follows (Bárdossy and Li, 2008):

$$S_k = \{x_{k,1}, ..., x_{k,n(k)}\} \quad k = 1, ..., K \tag{7.17}$$

The clustered points in a subset make that the distance between pair points is reasonable enough since the spatial interpolation procedure is commonly carried out using nearby points. Each observation point is allowed to occur only once in a subset in such a way that all subsets are disjoint to each others, i.e.

$$S_k \cap S_j = 0 \quad \text{if } k \neq j \tag{7.18}$$

Further, the subsets should be created such that all subsets cover all observation points, i.e.

$$S = \bigcup_{k=1}^{K} S_k = S \tag{7.19}$$

For any of the subsets $S_k$, the likelihood of the parameter vector of the theoretical copula can be calculated which corresponds to the copula density of the points enclosed in the subset.

$$c(S_k, \theta) = c\Big(F_z\big(Z(x_k, 1)\big), ..., F_z\big(Z(x_k, n(k))\big); \theta\Big) \tag{7.20}$$

Then, as all the subsets are disjoint, under a weak assumption of independence, the overall likelihood of the observation points is the product of the copula densities of the individual subsets as follows.

$$L(\theta) = \prod_{k=1}^{K} c(S_k, \theta) \tag{7.21}$$

Parameter estimation can be made by maximizing the likelihood function above with respect to the parameters $\theta$, namely the correlation matrix, and the values of $m$, and $k$.

### 7.4.2 Treatment of zeroes precipitation

The special characteristic of daily precipitation is zero inflated data. This requires a specific treatment where zero precipitation is assumed to be a censored variable of a continuous distribution function. This special treatment of zero precipitation, however, would lead to an increase of the complexity of the calculations (Bárdossy, 2011). The likelihood of the parameter vector of the chosen theoretical copula is, therefore, calculated by using the multiple bivariate copula densities from all possible combinations of pair of stations enclosed in the subset.

The likelihood function used for estimating the parameter of the Gaussian copula is exactly the same as the Equation 5.4 in chapter 5. Notice that the likelihood function is calculated for each pair of gauge stations from possible pair-wise combinations in the region of interest after classified into several clusters (subsets). The parameters of the $V$-copula are estimated by maximizing numerically the likelihood function which is basically similar to the Equation 5.4. Here, $\theta$ represents the parameters of the $V$-copula, namely, the Spearman's rank correlation and the transformation parameters $m$ and $k$. The likelihood function of the censored bivariate $V$-copula is formulated as follows.

$$\begin{aligned} L(\theta) = & \prod_{(x_1,x_2)\in I_1} \frac{h^2\big(H^{-1}(u_1 = F_1(x_1)), H^{-1}(u_2 = F_2(x_2)); \theta\big)}{h\big(H^{-1}(u_1 = F_1(x_1))\big) h\big(H^{-1}(u_2 = F_2(x_2))\big)} \times \\ & \prod_{(x_1,x_2)\in I_2} \int_{u_1=0}^{u_1=P_{01}} \int_{u_2=0}^{u_2=P_{02}} \frac{h^2\big(H^{-1}(u_1), H^{-1}(u_2); \theta\big)}{h\big(H^{-1}(u_1)\big) h\big(H^{-1}(u_2)\big)} du_1 du_2 \times \\ & \prod_{(x_1,x_2)\in I_3} \int_{u_1=0}^{u_1=P_{01}} \frac{h^2\big(H^{-1}(u_1), H^{-1}(u_2 = F_2(x_2)); \theta\big)}{h\big(H^{-1}(u_2 = F_2(x_2))\big)} du_1 h\big(H^{-1}(u_2)\big) \end{aligned} \tag{7.22}$$

Where, $h^2$ is the density function of the bivariate $V$-transformed distribution, $H^{-1}$ is inverse of the univariate distribution function of the $V$-transformed standard normal distribution $N(0,1)$, $h$ is density function of the $V$-transformed standard normal distribution $N(0,1)$. The sets $I_1$, $I_3$, and $I_3$ are the same definition as the Equation 5.4.

### 7.4.3 Stationary assumption

The $n$-dimensional multivariate Gaussian copula is identified through a correlation symmetrical matrix $\Gamma_{n\times n}$, while the $n$-dimensional $V$-transformed normal copula is parameterized by the two transformation parameters $m$, $k$, and the correlation symmetrical matrix $\Gamma_{n\times n}$ of the underlying of multivariate normal $Y$. This study is conducted under the assumption of spatial stationary over the domain of interest which is a similar assumption to the geostatistical approach. Hence, the correlations between any two points can be written only as a function of the separating vector $h$ and do not depend on their exact geographical locations.

For any set of observations $x_1, ..., x_n$, the correlation symmetrical matrix $\Gamma_{n\times n}$ of the normal variable $Y$ can be written as follows (Bárdossy, 2011; Bárdossy and Li, 2008; Bárdossy and Pegram, 2013).

$$\Gamma_{n\times n} = \left[(\rho_{i,j})_{1,1}^{n,n}\right] \tag{7.23}$$

where $\rho_{i,j}$ is the correlation between two points $x_i$ and $x_j$. It only depends on the vector $h$ separating the points $x_i$ and $x_j$. Thus, the correlation $\rho_{i,j}$ can be written as follows.

$$\rho_{i,j} = R(x_i - x_j) = R(h) \tag{7.24}$$

Where $R(h)$ is a valid (positive and definite) spatial correlation function depending on the vector $h$ such as the negative exponential correlation function, the spherical correlation function, and other theoretical spatial correlation functions as discussed in the chapter 5.

For the $V$-transformed normal copula the parameters of the correlation function corresponding to $Y$ and the transformation parameters ($m$ and $k$) have to be assessed on the basis of the available observations. Notice that in the derivation of $V$-transformed normal copula, all the components of the $m$ are assumed to be equal.

## 7.5 Parameters interpretation

### 7.5.1 Case study in Singapore

The spatial copula models based on both the symmetric (Gaussian copulas) and the asymmetric spatial dependence (V-copulas) are implemented in the region of Singapore using the high-quality precipitation data sets with different temporal resolutions from hourly to monthly during the time period of 1980-2010. The parameter estimation of both families of copulas is conducted using the maximum likelihood method incorporating zero precipitation amounts, which are treated as censored variables as given in Equation 7.22 and Equation 5.4.

Notice that the modeling processes are carried out for each time step of precipitation values separately, without considering the temporal correlation behavior. The precipitation occurrences, which are analyzed in this study, are all selected events on a given time step at which more than 0.7 of all gauge stations in the region of interest are wet with precipitation depth

of more than 0.1 mm. This means that a precipitation event on a given time step with the probability of dry events bigger than 0.3 will be excluded from the further analysis. In Singapore for example, a precipitation occurrence on a given time event during the time period of 1980-2010 is considered in this analysis if at least 21 of 30 precipitation gauge stations receive precipitation of more than 0.1 mm. In order to detect the influences of different seasons on the parameters of interest, the precipitation events are classified into four different seasons, namely, the DJF, MAM, JJA and SON seasons.

**Parameter of range** The range parameter in this study is basically derived from the fitted negative exponential spatial correlation function and calculated based on the distance where the correlation coefficient of $\exp(-1) = 0.37$. Figure 7.1 depicts multiple box-plots of the parameters range at different time scales and different seasons for both the Gaussian and the V-copulas. Overall, the range parameter of the Gaussian copulas is systematically lower than the ones of the V-copulas nearly for all the variety of time scales and different seasons as portrayed in Figure 7.1. The red lines on Figure 7.1 represent the range parameter of 5 km to emphasize the different characteristics of both spatial models. The range parameter of the Gaussian copulas is obviously less than 5 km, on average, while the range parameter of the V-copulas is clearly greater than 5 km on the whole. This indicates that the V-copulas bring a stronger spatial dependence structure than the Gaussian copulas in the region of Singapore.

Furthermore, the kernel distribution functions of the range parameter of both models as shown in Figure 7.2a give an additional empirical evidence that the V-copula models present a larger spatial correlation range than the Gaussian copulas in all the quantile values. The kernel distribution functions of the V-copulas are represented by red lines on Figure 7.2a, while the blue lines represent the kernel distribution functions of the Gaussian copulas. This empirical evidence applies not only to the specific time scales but also to nearly all different temporal scales across the different seasons. Indeed, for several time scales, such as 15-daily and monthly, both models present a similar kernel distribution function of the range parameter, but those are only for the JJA season.

There is an interesting reason why the V-copulas have a stronger spatial dependence structure than the Gaussian copulas. Figure 7.2b shows multiple plots of the kernel density functions of the range parameter for both Gaussian and V-copulas at different temporal scales across the seasons. Again, the red lines represent the kernel density functions of the range parameter for the V-copulas, and the blue lines denote the kernel density functions of the parameters range for the Gaussian copulas. The kernel density functions of the range parameter for the Gaussian copulas increase quickly, but also decrease quickly. In contrast, the kernel density functions of the range parameter for the V-copulas rise slowly, but also decline slowly. As a result of these phenomena, the V-copulas can capture a larger spatial correlation structure, on the whole.

The mean values of the parameters range for the Gaussian copulas from all events during the time period of 1980-2010 are quite similar across seasons and different temporal scales. In the DJF season for instance, the range parameter roughly ranges from 4.9 km to 5.9 km. Similarly, the range parameter for the MAM season varies between 4.2 and 5.5 km. The

range parameter for the seasons of JJA and SON varies from 4.4 km to 5.5 km and from 4.3 km to 6 km, respectively.

Similarly to the Gaussian copula, the mean values of the range parameter for the V-copulas during the time period of 1980-2010 is similar to each other for different seasons and temporal scales. In the DJF season, the range parameter varies from 6.9 km to 7.9 km for sub-daily scales and ranges from 6.7 km to 7.4 km for the super-daily scales. The range parameter in the MAM season is lower varying from 5.9 to 6.9 km for the sub-daily scales and ranging from 5.4 km to 6.2 km for super-daily time scales. In the JJA season, the range parameter varies from 6.7 km to 7.1 km for sub-daily scales and ranges from 4.7 km to 6.3 km for the super-daily scales. The range parameter ranges from 7 km to 7.7 km for sub-daily scales and varies from 6 to 7.1 km for super-daily scales in the SON season.

In fact, the range parameter is generally higher in the DJF and SON seasons in comparison to the MAM and JJA seasons. However, in practical applications, the differences between values of the range parameter mentioned above is relatively very small seeing the common distance among gauge stations, and thus this can be concluded that there is no significant difference of the general correlation characteristics between different seasons. Furthermore, there is a fluctuation of the range parameter over different time scales, but it is quite difficult to conclude that there is an increase of range parameter as increase the time intervals because it is very small and almost negligible. This is a very similar result to the spatial correlations which are estimated by using the geostatistics approach on the basis of the censored bivariate Gaussian copulas as mentioned in section 5.3 of chapter 5.

**Parameter $m$ and $k$** The multivariate normal distributions are simply parameterized by the correlation matrix, which is described by the range parameter following the negative exponential function (for this study) due to the stationary assumption. In the V-copula models, all negative random variables underlying the multivariate normal distribution are inevitably transformed into the corresponding positive values using the non-monotonic transformations with the parameters $m$ and $k$ as written in Equation 7.6.

Unlike the monotonic transformation which converts one set of random variables into another set of the random variables so that the rank order of the original set of random variables is preserved, the non-monotonic transformations instead will destroy the rank order of random variables underlying normal distribution. This is one reason why the non-monotonic transformation is a suitable choice implemented in the copula model because copula works under the rank of the set of random variables and the dependence structures are strongly dependent on the rank order of random variables. If the ranks do not change after such a transformation, it means that the dependence structure does not change either. In the spatial context, the parameters $m$ and $k$ are commonly assumed to be the same value for any location.

The location parameter $m$ and scale $k$ play a fundamental role in the V-copula model which bring a change of the rank order of any set of random variables. If any random variable underlying the normal distribution is less than the parameter $m$, which is a small value in the normal domain (it can be either both negative and positive random variables or all

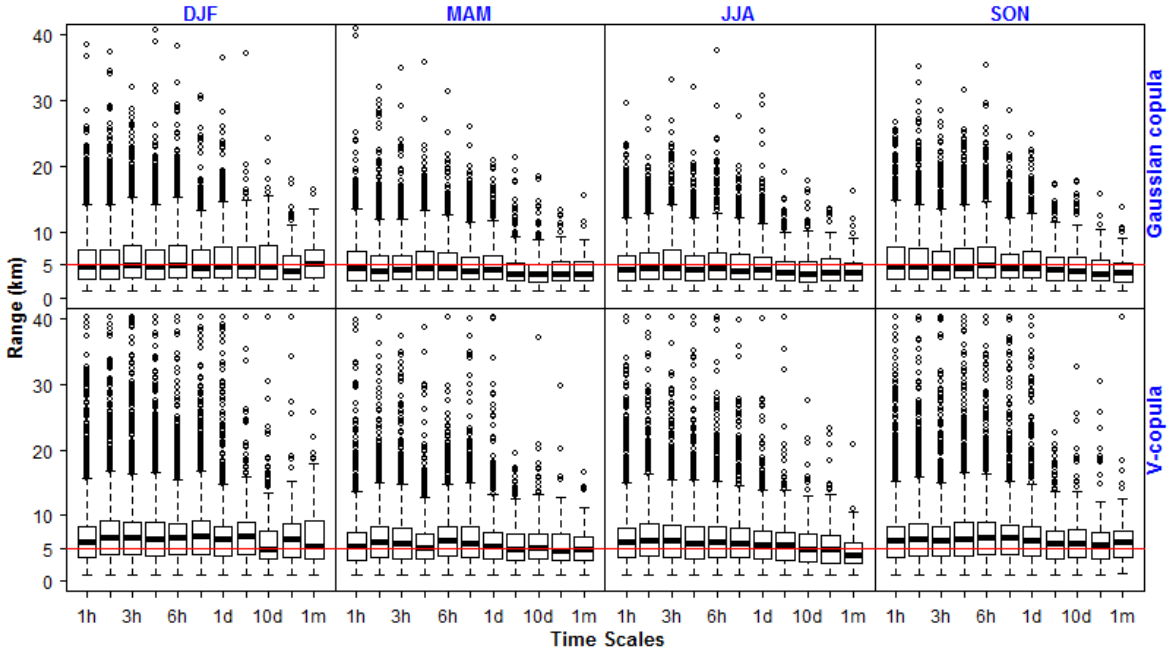
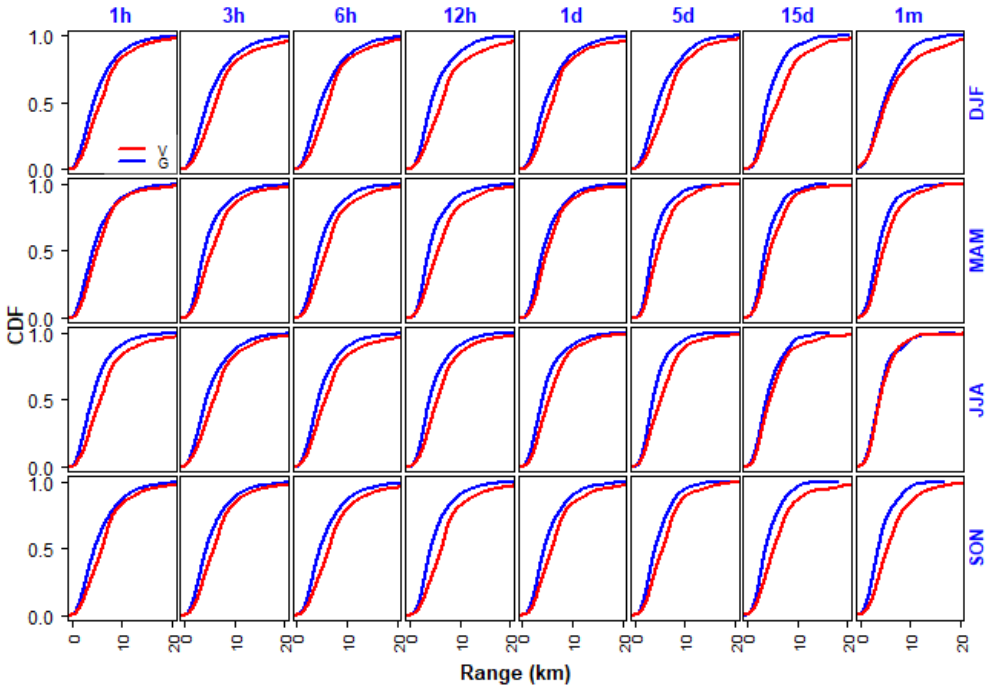
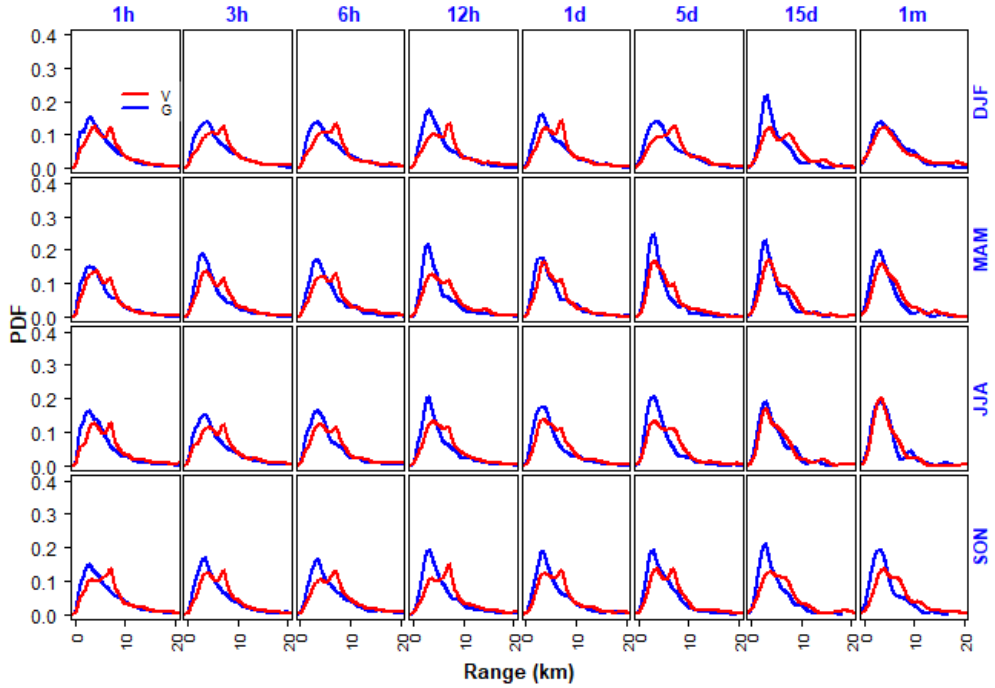



Figure 7.1: Parameters of range from the Gaussian and V-copulas in Singapore. The primary vertical axes represent the range parameter (km). Panels from top to bottom represent the Gaussian and V-copulas. The horizontal axes represent temporal scales from hourly to monthly. Panels from left to right represent seasons (DJF, MAM, JJA, and SON).




(a) CDF




(b) PDF

Figure 7.2: Distribution of the range parameter from the Gaussian and V-copulas in Singapore collected from the period 1980-2010. The primary vertical axes represent kernel distribution function (Fig-a) and kernel density function (Fig-b). Panels from top to bottom represent seasons (DJF, MAM, JJA, and SON). The horizontal axes define the range parameter (km). The blue lines indicate the Gaussian copulas and the red lines represent the V-copulas. Panel from left to right represent temporal scales (hourly (1h) to monthly (1m)).

negative variates), it turns into a high value in the new space due to the non-monotonic transformations of the subtraction from the normal variate to the parameter $m$.

Conversely, when a random variable is the high value in the previous space, it becomes the low value in a new domain. Furthermore, if any normal variate is greater than the parameter $m$, it is transformed into either more concentrated or more deviated from the value of the parameter $m$ depending on the parameter $k$. The higher the parameter $k$, the more spread it can yield in the distribution in the new domain causing the new distributions to be strongly right skewed. Consequently, the rank of the random variables in the new domain could be changed from the rank of the multivariate normal distributions by this transformation.

However, the non-monotonic transformation above does not guarantee a significant change of the rank order of any data set depending on the values of the parameter $m$. For example, for a greater parameter $m$, the ranks of the variates is less changed because the left branch of the transformation dominates the right one. As a result of this, marginal distribution remains more similar to the original one, and the dependence structure is more symmetric and more Gaussian. Moreover, if the parameter $m \geq 3$, then the rank order of the new domain remains the same as in the old space because all normal variates are transformed into the new domain using only the left arm of the transformation, and therefore the spatial dependence structure approaches Gaussianity which is the same as in the old domain.

Figure 7.3 presents dependence structures of simulations of bivariate Gaussian copulas with the correlation coefficient $\rho$=0.84 and bivariate V-copulas with a variety of combination parameters $m$ and $k$. One can see that the dependence structure exhibits the positive asymmetric dependence when the parameter $m$ is smaller or close to zero, and vice versa, the association structure tends to approach a Gaussian copula as the parameter $m$ is greater and approaching to infinitive. However, for any practical application, the dependence structure will be close to Gaussian when the parameter $m \geq 3$ as mentioned before. In addition, the V-transformed Gaussian copula will equal the non-centered square copula for the parameter $k$=1. If the parameter $m$ is small and the parameter $k$ is higher, then the dependence structure on the middle quantiles will increase.

**Consequences of the parameters $m$ and $k$ on mechanisms of precipitation** Precipitation mechanisms, which are generally characterized by convective and stratiform regimes (Anagnostou, 2004), can be mimicked by the $V$-copulas reasonably with the two branches of the non-monotonic transformation. This model can describe two rainfall generating processes separately; adjective/stratiform precipitation modeled by the lower arm of the $V$-transformation and convective precipitation mimicked by the upper arm of it (Bárdossy and Pegram, 2009).

The convective precipitation event is usually characterized by localized precipitation of high intensity on the short duration triggered by intense vertical movements as the product of atmospheric convection, from active cumulus and cumulonimbus clouds and mainly occurring in tropical countries (Houze, 1997, 2004). In contrast, the stratiform precipitation event is mainly described with relatively weak vertical velocity fields, greater horizontal homogeneity, and lower rainfall intensity. Although stratiform precipitation intensity is

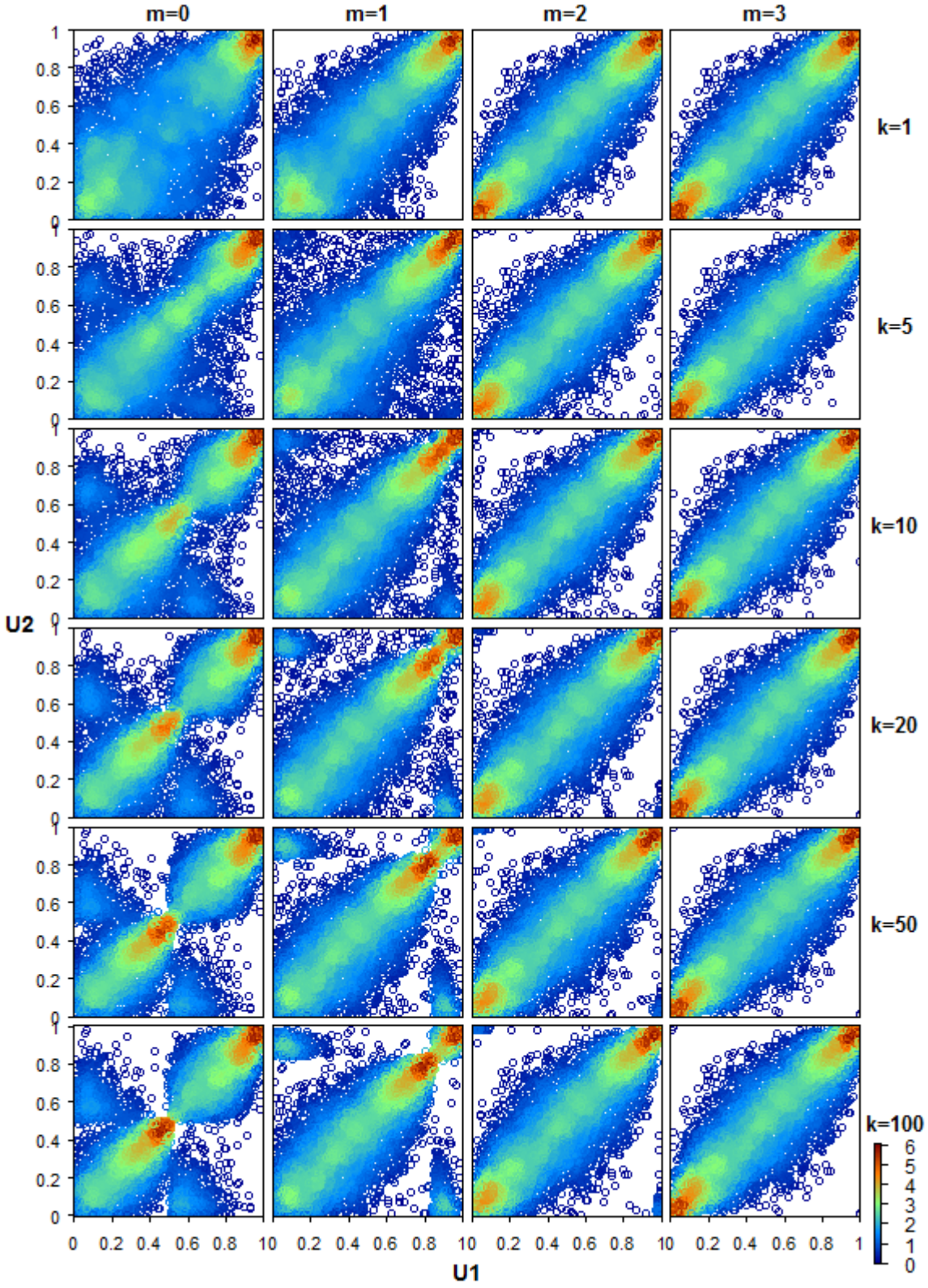
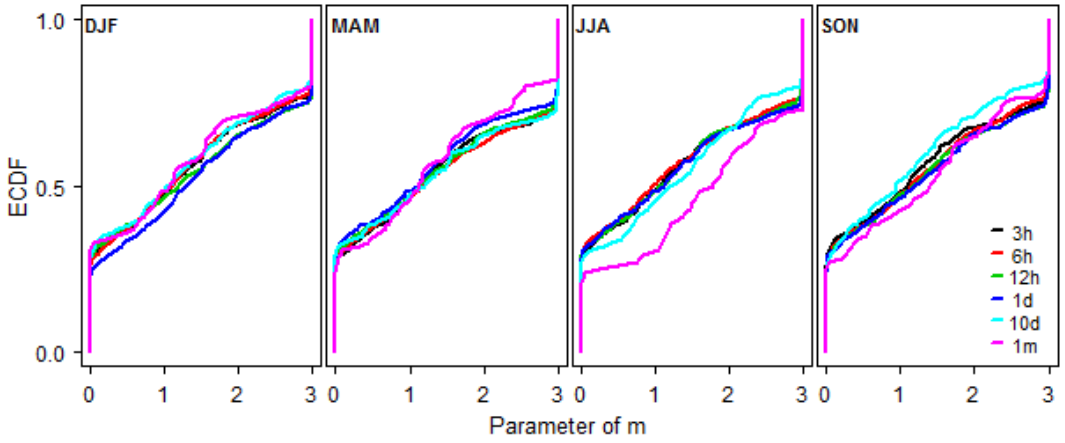
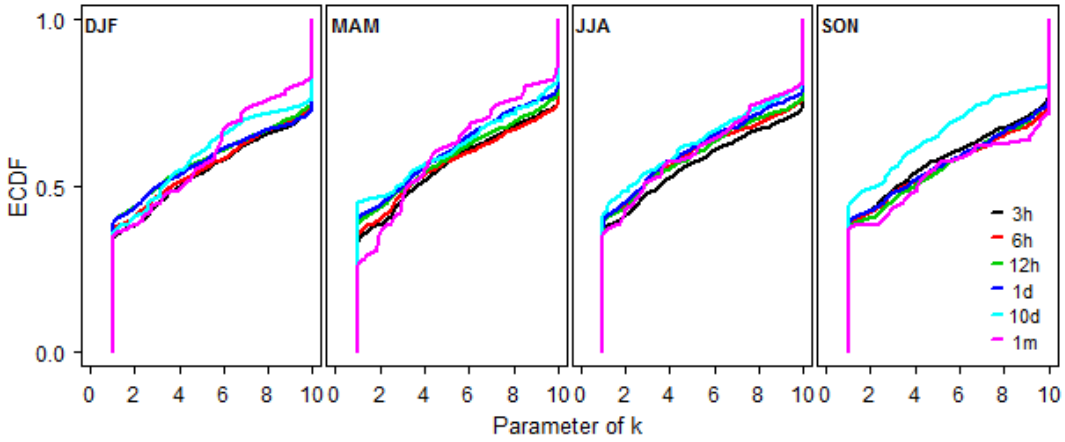



Figure 7.3: Bivariate V-copulas density with $\rho = 0.85$ and different parameter $m$ and $k$. Panels from left to right represent the parameter $m$. Panels from top to bottom represent the parameter $k$.

much lower than that from the convective precipitation event, the stratiform precipitation typically covers larger areas and always contributes a significant portion (40-50%) of the precipitation volume, even for major convective systems (Anagnostou, 2004; Houze, 2004).




(a)




(b)

Figure 7.4: Empirical distribution functions of the parameter $m$ (Fig-a) and the parameter $k$ (Fig-b) from the V-copulas on different temporal scales in Singapore collected from the period 1980-2010. Panels from left to right represent seasons (DJF, MAM, JJA, and SON).

Most precipitation events in Singapore exhibit an asymmetric spatial dependence structure. This is because those precipitation occurrences exhibit the smaller parameter $m$. Figure 7.4a presents empirical distribution functions of the parameter $m$ from the V-copula model which are estimated for each precipitation event at the various time scales and different seasons collected from the time period of 1980-2010. The averaged values of the parameter $m$ over realizations of precipitation events are relatively similar for all different time scales and different seasons. For instance, the parameter $m$ in the MAM season is very similar to the SON seasons, namely, ranging from 1.2-1.4 on average. Meanwhile, for the seasons of DJF and JJA, the parameter $m$ varies from 1.3 to 1.4 and from 1.2 to 1.6, respectively. For

daily scale, the parameter $m$ for the months of DJF, MAM, JJA, and SON is 1.4, 1.3, 1.3, and 1.4, respectively.

In general, the higher time scales yield less asymmetric dependence, but not very different. For example for the JJA season, monthly precipitation occurrences exhibit a higher values of the parameter $m$, namely 1.6. For other seasons, the effect of different time scales on the asymmetric behavior is not clearly seen. Furthermore, the most frequent values of the parameter $m$ roughly ranges from 0-0.5 for all different time scales and seasons indicating that the positive asymmetric spatial dependence of precipitation values occur significantly in this region.

In addition, Figure 7.4b presents empirical distribution functions of parameter $k$ for the V-copula estimated for each event at various time scales and different seasons from 1980-2010 in the region of Singapore. Overall, the parameter $k$ varies between 1 and 10 and tends to concentrate to smaller values or close to 1. In the case of the asymmetric dependence with the lower parameter $m$ near to 0, the smaller parameter $k$ results in a dependence structure more concentrated into the uppermost quantiles and conversely, larger value of the parameter $k$ causes the dependence structure on the uppermost quantiles to be more spread and tends to deviate from there to the middle quantiles. However, for the higher the parameter $m$ around 2, which there are two parts of the mass density concentrated on both uppermost and lowermost quantiles, the greater parameter $k$ reduces the mass density on the lowermost quantiles and increases to the uppermost quantiles or the middle quantiles (see Figure 7.3 for illustration).

The average value of the parameter $k$ is relatively small just ranging from 3.9 to 5.1 for all time scales and seasons. For example, the parameter $k$ in the DJF season roughly ranges from 4.7-5.1 (for hourly to daily) and 4.6-4.8 (for 5-daily to monthly). In the MAM season, the parameter $k$ varies from 4.7-5 (for sub-daily) and 4.4-4.6 (for super-daily). In the JJA season, the parameter $k$ ranges from 4.6-5 (for sub-daily) and 3.9-4.5 (for super-daily). In the SON season, the parameter $k$ ranges from 4.8-5.4 (for sub-daily) and 4-5.1 (for super-daily). The parameter $k$ is relatively slightly lower implying that local heavy precipitation occurs more intensively in this region due to the convective precipitation. Furthermore, precipitation values at the lower time scales exhibit more systematically a smaller value of the parameter $k$ than that of the higher time scales, describing that precipitation events with high intensity occur more clustered at the small time scales.

**Effect of precipitation intensities** The effect of precipitation intensity on the parameters of the spatial models of interest is very interesting and exhibits special pattern, in particular in the region of Singapore. In order to show it, the precipitation events at daily scale observed during the time period of 1980-2010 are investigated. For each precipitation event, the areal mean precipitation is calculated based on all precipitation points measured at all gauge stations. All precipitation events are then divided into two different parts depending on their intensity, namely the high extreme precipitation events and the low extreme precipitation events. Different precipitation intensities are required to distinguish the behavior on the parameters of interests. The precipitation events, that exhibit the areal mean precipitation intensity are higher the quantile 75% of the areal mean precipitation intensity,

are classified as high extremes. In contrast, the low extremes correspond to the precipitation events with the areal mean precipitation intensity lower than the quantile 25%. This is carried out because two sets of precipitation events with the same length of time events but different intensity are desired. These empirical investigation results are presented in Figure 7.5a and Figure 7.5b.

The effects of precipitation intensities on the parameters of V-copulas ($m$ and $K$) are shown in Figure 7.5a. This shows the empirical cumulative distribution function of the parameters of V-copulas, the parameter $m$ and the parameter $k$ at daily scale with different seasons. One can see clearly that precipitation events with high-intensity exhibit a lower parameter $m$, on the whole, across seasons except for the DJF season. This implies that the heavy precipitation triggers the asymmetric spatial dependence structure on the uppermost quantiles. In other words, precipitation with high intensity tends to occur in clustering manner due to the local convective precipitation arising in this region. In the DJF season, however, the distribution function of the parameters $m$ for the high extreme precipitation seems to be similar to the low extreme precipitation. Furthermore, the extreme precipitation amounts also generate the smaller parameter $k$. This indicates that asymmetric dependence structure concentrates more on the upper quantiles without split to other quantiles or resulting in more mass on the middle quantiles or even close to the middle quantiles. Nevertheless, both data sets of high extremes and low extremes present a high intensity of precipitation values in general.

The effects of precipitation intensities on the range parameter from V-copulas and Gaussian copulas are given in Figure 7.5b. Figure 7.5b shows the effect of extreme precipitation on the range parameter for both Gaussian and V-copula models. It can be seen that the extreme precipitation exhibits a slightly higher range parameter for both copula models than the precipitation with the low intensity. However, for Gaussian copula in the DJF seasons, the high extreme precipitation presents a significantly larger range parameter than the low extreme precipitation.

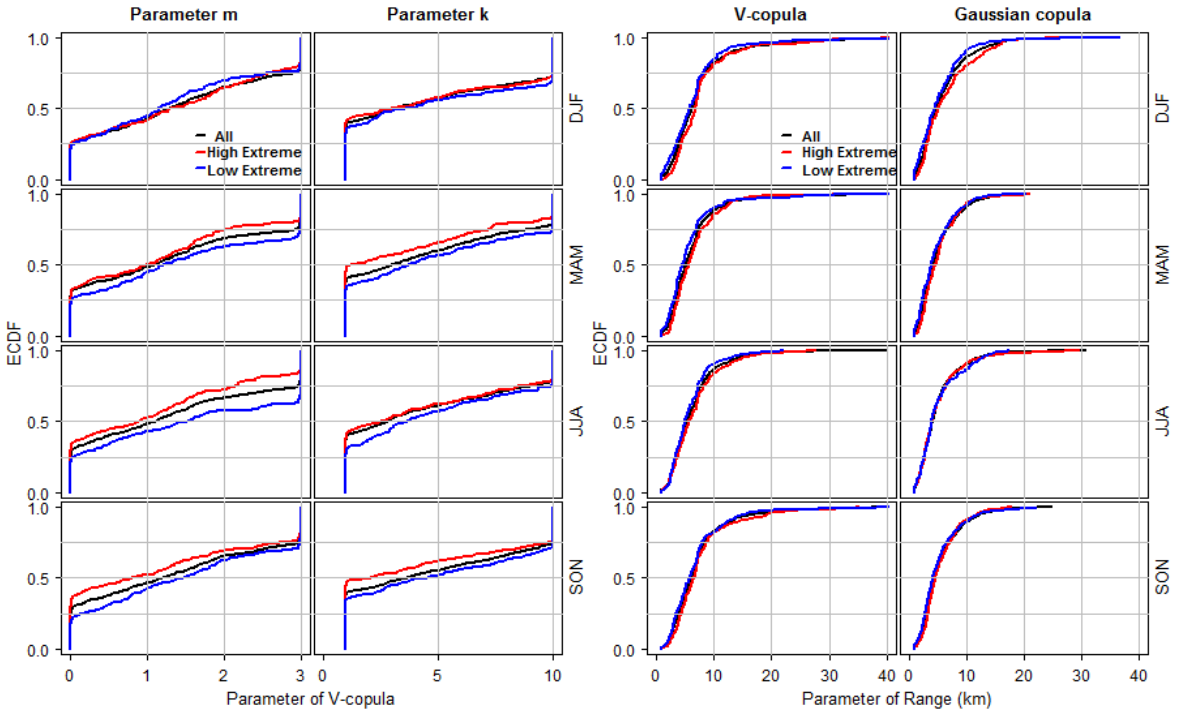



Figure 7.5: Empirical distribution functions of the parameters of the Gaussian copulas (range) and V-copulas (range, $m$, and $k$) on daily scale in Singapore with different intensity of precipitation amounts. Panels from top bottom represent seasons (DJF, MAM, JJA, and SON). Black lines represent all precipitation data collected from the period 1980-2010. Red lines represent precipitation data with high intensity. Blue lines represent precipitation data with low intensity. Panels from left to right (Fig-a) represent parameters $m$ and $k$. Panels from left to right (Fig-b) represent range parameter (for V-copulas) and range parameter (for Gaussian copulas).

**Effect of skewness of precipitation** It is widely known that precipitation amounts, especially for sub-daily timescales, exhibit a strong right skewness or positive skewness in time series observations because it has a long tail which extends to the right direction or to the positive values containing extremely large values resulting in the mean of the distribution toward the tail. It might also be interesting to investigate the effect of the right skewness of precipitation values on the asymmetric spatial dependence structure. This is because precipitation values generally fall in a clustering way at a location.

The degree of right skewness can be approximated by calculating its parameter of shape from the Gamma distribution. As pointed out in chapter 3, the parametric Gamma distribution has two parameters; the shape and scale. The Gamma distribution is used in this investigation because it has unique characteristics. The expected value of the Gamma distribution is simply calculated by the product of the two parameters, shape parameter and scale parameter. Hence, it is easier to interpret the characteristics of precipitation values based on those parameters.

A similar investigation as previously done the effect of the extreme precipitation on the spatial models can be performed. Instead, the shape parameter of the Gamma distribution is used. All precipitation events, here again at daily scale measured during the time period of 1980-2010, are classified into the two different groups with the same time length of events. The groups are events at which precipitation values exhibit the strong right skewness on one hand, and on the other hand, events at which precipitation values present the weak right skewness. Precipitation values that show a greater degree of right skewness can be determined from the precipitation amounts with a shape parameter bigger than the quantile 75% of it. Meanwhile, precipitation values that exhibit the shape parameter lower than the quantile 75% are categorized as the lower degree of the positive of skewness. These empirical investigation results are presented in Figure 7.6a and Figure 7.6b.

Figure 7.6a depicts the empirical distribution function of the parameters of V-copulas, the parameters $m$ and the parameters $k$ at daily scale with different seasons. One can see clearly that precipitation values with low shape exhibit the higher parameter $m$, on the whole. This implies that the skewed precipitation amounts do not automatically trigger the asymmetric spatial dependence structure on the uppermost quantiles. Instead, the skewed precipitation values can also result in the symmetric spatial dependence. Furthermore, the skewed precipitation values also generate the higher parameter $k$. This indicates that the asymmetric dependence structures of precipitation values in the region of attention do not concentrates on the upper quantiles only if the parameter $k$ higher, but it also result in more mass density on the middle quantiles or even close to the middle quantiles.

Figure 7.6b shows the effect of the skewed precipitation on the range parameter of both Gaussian copula and V-copula model. It can be seen that the low skewness of precipitation amounts exhibit a slightly higher range parameter for both copula models than the precipitation values with higher skewness. However, for the Gaussian copula in the DJF seasons, the low skewness of precipitation values present significantly greater correlations length than the precipitation amounts with high skewness.

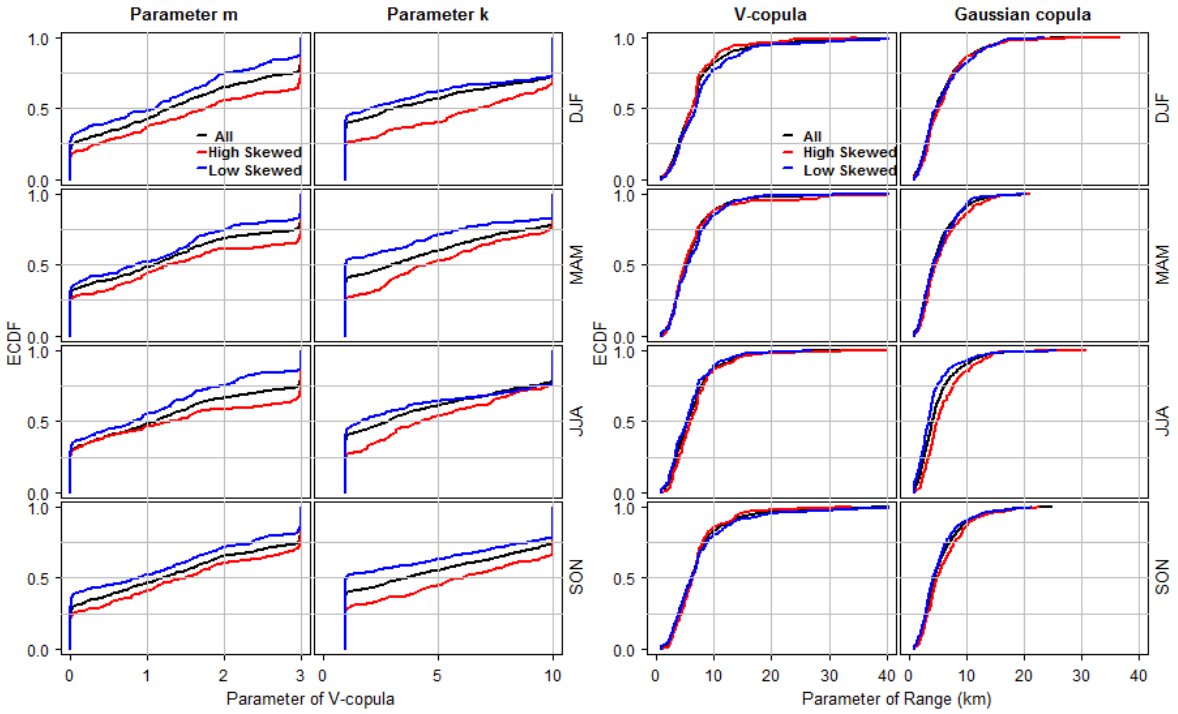



Figure 7.6: Empirical distribution functions of the parameters of the Gaussian copulas (range) and V-copulas (range, $m$, and $k$) on daily scale in Singapore with different right-skewed of precipitation amounts. Panels from top bottom represent seasons (DJF, MAM, JJA, and SON). Black lines represent all precipitation data collected from the period 1980-2010. Red lines represent precipitation data with high right-skewed. Blue lines represent precipitation data with low right-skewed. Panels from left to right (Fig-a) represent parameters $m$ and $k$. Panels from left to right (Fig-b) represent range parameter (for V-copulas) and range parameter (for Gaussian copulas).

### 7.5.2 Case study in Bavaria

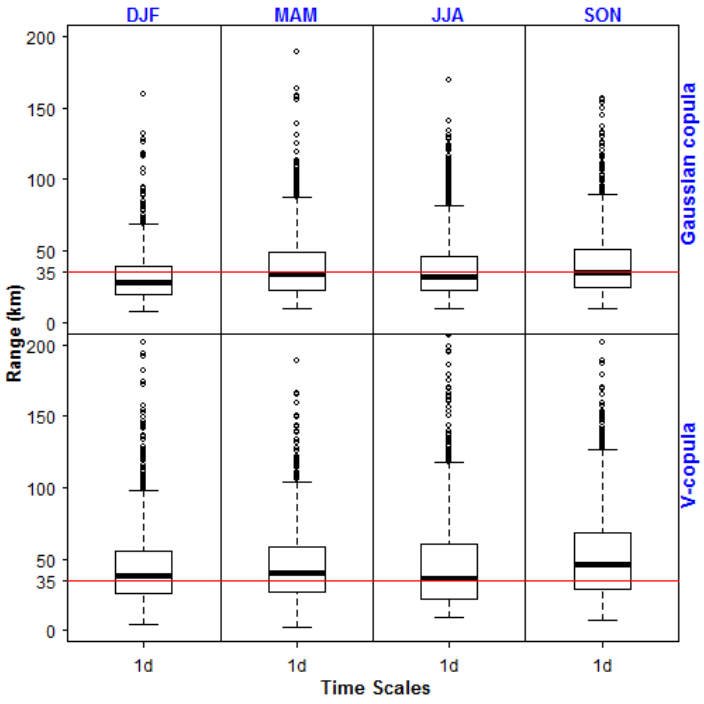
Similarly to the region of Singapore, the spatial copula models based on both the symmetric (Gaussian copulas) and the non-symmetric spatial dependence ($V$-copulas) have also been implemented in the region of Bavaria. Parameters of both models are estimated using the maximum likelihood method taking into account zero precipitation which is treated as latent variables. However, unlike in the study of Singapore, which employs precipitation data sets at the different temporal scales, ranging from hourly to monthly scale, this study in Bavaria uses only daily precipitation, which was observed during the time period of 1951-2001 from 573 gauge stations. This is because it requires a large effort of the calculations due to limitation time consuming on the computation.

In addition, higher than daily time scales are also not of importance in many hydrological applications especially dealing with risk assessment. Investigations are conducted at every time step independently due to the assumption of no temporal correlations. Precipitation events which are examined in this study should have the probability of dry events smaller than 0.3. Thus, precipitation events, for which the probability of a dry event is bigger than 0.3, will be excluded from the further analysis and modeling. In order to detect the influences of different seasons on the spatial models, the precipitation events are classified into 4 different seasons, namely, the DJF, MAM, JJA and SON seasons.

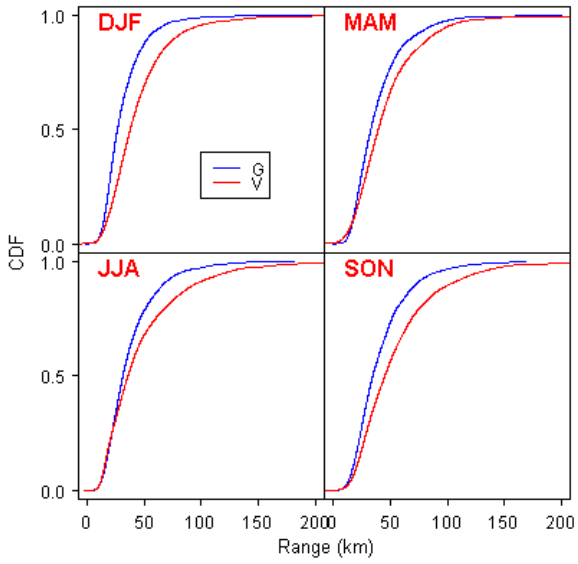
**Parameter of range** Figure 7.7 shows the range parameter of both Gaussian and V-copula models comprehensively. Similar result as the ones for the region of Singapore, the V-copula models exhibit a systematically greater range parameter than the Gaussian copulas for all different seasons. In general, the range parameter of the Gaussian copula are smaller than 35 km, while the V-copulas present values of the range parameter higher than 35 km as shown in Figure 7.7a.

Based on the kernel distribution functions of the range parameter, the V-copulas also present higher the range parameter than the Gaussian copulas for all seasons as shown in Figure 7.7b. Similarly, the red lines on figure represent the kernel distribution functions of the range parameter of the V-copula and the blue lines on it describe the kernel distribution function of the range parameter of the Gaussian copulas. The range parameter of the V-copula in the seasons of DJF and SON is clearly higher than the ones of the Gaussian copula in comparison to other seasons. Even, for the JJA and SON seasons, the kernel distribution functions of both models present similar curves for small range parameter, especially for the JJA season. This is more obviously seen on Figure 7.7c.

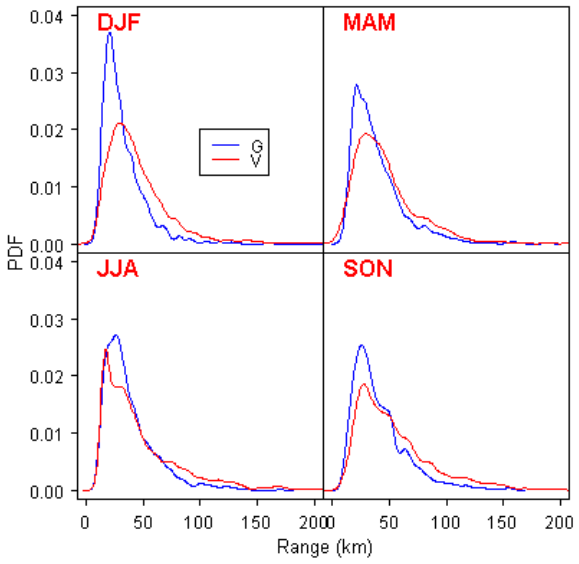
Figure 7.7c depicts the density of the kernel distribution functions of the range parameter of both Gaussian and V-copula. Similarly, the red lines on figure represent the kernel distribution functions of the range parameter of the V-copula and the blue lines on it describe the kernel distribution function of the range parameter of the Gaussian copulas. Again, the density function of range parameter of the Gaussian copulas increase quickly but also decrease rapidly. In contrast, the density function of range parameter of the V-copulas rises slowly, but also declines slowly. As a result of this phenomenon, the V-copulas can capture more significantly a larger spatial correlation structure.




(a)




(b)




(c)

Figure 7.7: Distribution of the range parameter from the Gaussian and V-copulas in Bavaria on daily scale. Fig-a: Box-plot of range parameter in different seasons (panels from left to right: DJF, MAM, JJA, and SON), and panels from top to bottom: the Gaussian and V-copulas. Fig-b: Kernel distribution function of range parameter for the Gaussian copulas (blue lines) and V-copulas (red lines) in different seasons (DJF, MAM, JJA, and SON). Fig-c: Kernel density function of range parameter for the Gaussian copulas (blue lines) and V-copulas (red lines) in different seasons (DJF, MAM, JJA, and SON).

The mean values of the parameters range of the Gaussian copulas from all daily precipitation events during the time period of 1951-2001 are slightly different across seasons, namely 31.6 km, 39 km, 37.4 km, and 41 km, for the seasons of DJF, MAM, JJA, and SON, respectively, while the average values of the parameters range of the V-copula yield higher range parameter than Gaussian copulas, namely 44.6 km, 46.7 km, 47.2 km, 54.6 km for seasons of DJF, MAM, JJA, and SON, respectively.

**Parameters $m$ and $k$** The average values of the parameters $m$ in Bavaria are around 2 or more for all different seasons, namely, 2, 2.4, 2.6, and 2.1 for the seasons of DJF, MAM, JJA, and SON, respectively. This indicates the existence of the weak positive asymmetric spatial dependence structure for most precipitation events. The average values of the parameters $k$ for all different seasons, namely, 13.9, 25.7, 6.6, and 17, for the seasons of DJF, MAM, JJA, and SON, respectively. This implies that mass density from both uppermost and lowermost quantiles move slightly to the middle quantiles. However, overall, the uppermost quantiles exhibit a higher density value than the lowermost quantiles. This is an indication of a very strong dependence of the large values

**Effect of precipitation intensities** Figure 7.8a shows the empirical distribution function of the parameters $m$ and $k$ on a daily scale with different seasons using the different precipitation intensities. One can see that precipitation with high-intensity presents slightly greater parameter $m$, on the whole, for all seasons except for DJF. This implies that the heavy precipitation brings a weak positive asymmetric spatial dependence structure on the uppermost quantiles. This is because the parameter $m$ is close to 3 which is the symmetric dependence. In other words, precipitation with a higher intensity in the region of interest tends to cover larger areas due to the frontal stratiform precipitation arising in this region. In the DJF season, however, the distribution function of the parameters $m$ from the high extreme precipitation seems to be higher than the low extreme precipitation, in particular for large values. Furthermore, the extreme precipitation also generates smaller value of parameter $k$ except for the DJF seasons.

However, overall, the parameters $k$ in Bavaria are larger. This indicates that the asymmetric dependence structure concentrates more on the upper quantiles and splits to other quantiles close to the middle quantiles. Nevertheless, both data sets of high extremes and low extremes present a modest precipitation intensity in general.

Figure 7.8b shows the effect of extreme precipitation on the range parameter of both Gaussian copula and V-copula model. It can be seen that the extreme precipitation exhibits a significantly higher range parameter for both copula models than the precipitation with low intensity. However, for the Gaussian copula in the DJF seasons, the high extreme precipitation presents only a slightly higher correlation length than the low extreme precipitation.

**Effect of skewness of precipitation** Figure 7.9a depicts the empirical distribution function of the parameters of V-copulas, the parameters $m$ and $k$ at the daily scale with different seasons corresponding to a different degree of the right skewed precipitation values. One

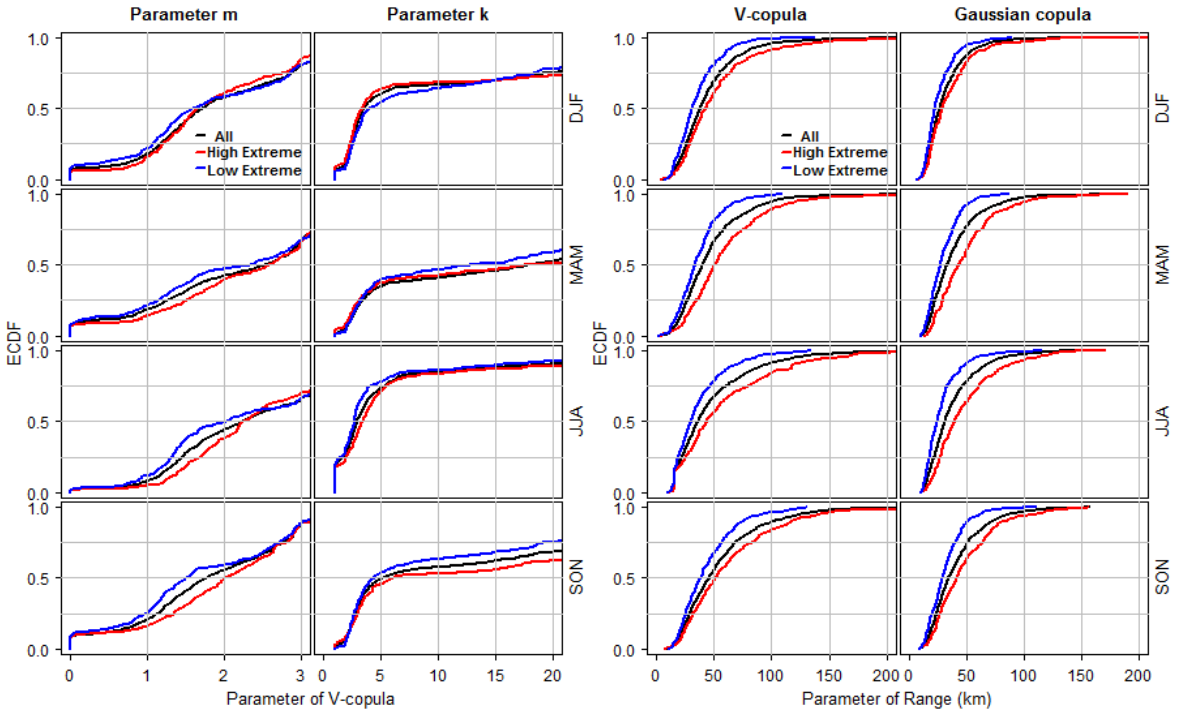



Figure 7.8: Empirical distribution functions of the parameters of the Gaussian copulas (range) and V-copulas (range, $m$, and $k$) on daily scale in Bavaria with different intensity of precipitation amounts. Panels from top bottom represent seasons (DJF, MAM, JJA, and SON). Black lines represent all precipitation data collected from the period 1980-2010. Red lines represent precipitation data with high intensity. Blue lines represent precipitation data with low intensity. Panels from left to right (Fig-a) represent parameters $m$ and $k$. Panels from left to right (Fig-b) represent range parameter (for V-copulas) and range parameter (for Gaussian copulas).

can see that precipitation amounts with low shape (a higher degree of right skewness) exhibit more frequent smaller values of parameter $m$ (0-1.5), but then bring more often a slight greater one of the parameter $m$ (2-3). This implies that the skewed precipitation amounts in Bavaria could result in more frequent positive asymmetric dependence.

Furthermore, the skewed precipitation values also generate a significant higher parameter $k$. This indicates that the asymmetric dependence structures of most precipitation events in the region of attention do not only concentrate on the upper quantiles, but they also result in more mass on the middle quantiles or even close to the middle quantiles. For the cases of weak positive asymmetric dependence structures that exhibit a parameter $m$ ranging from 2-3, densities on the lowermost quantile also decline and move into the middle quantiles close to the uppermost quantiles. Consequently, the precipitation shows a strong dependence not only on the uppermost quantiles but also on the middle quantiles or near the upper most quantiles.

Figure 7.9b shows the effect of the skewed precipitation on the range parameter of both Gaussian copula and V-copula model. It can be seen that the low skewness precipitation amounts exhibit a significantly higher range parameter for both copula models than the precipitation values with higher right skewness.

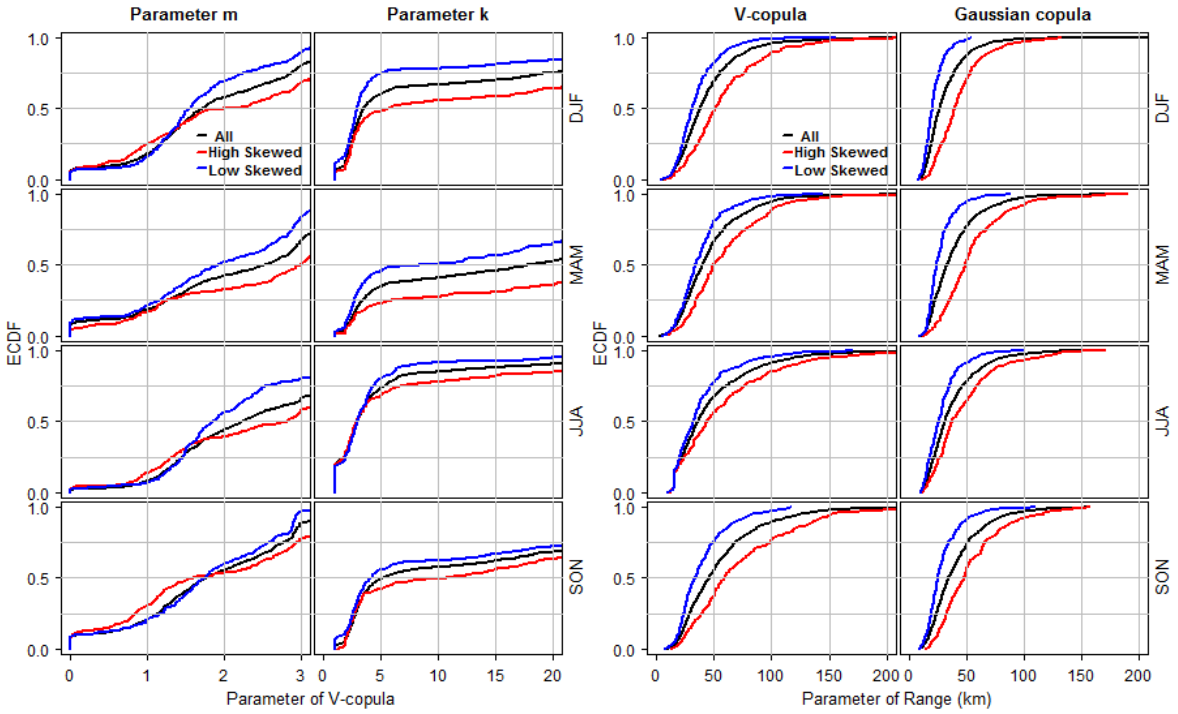



Figure 7.9: Empirical distribution functions of the parameters of the Gaussian copulas (range) and V-copulas (range, $m$, and $k$) on daily scale in Bavaria with different right-skewed of precipitation amounts. Panels from top bottom represent seasons (DJF, MAM, JJA, and SON). Black lines represent all precipitation data collected from the period 1980-2010. Red lines represent precipitation data with high right-skewed. Blue lines represent precipitation data with low right-skewed. Panels from left to right (Fig-a) represent parameters $m$ and $k$. Panels from left to right (Fig-b) represent range parameter (for V-copulas) and range parameter (for Gaussian copulas).

## 7.6 Summary

- The V-copulas or the V-transformed normal copulas are constructed from the Gaussian copulas through a non-monotonic transformation using two additional parameters, namely $m$ and $k$. The Gaussian copulas are the symmetric Gaussian dependence based models, and the V-copulas represent the asymmetric dependence based models. Both copulas are applied to describe precipitation fields located in the regions of Singapore and Bavaria.
- Empirical evidence proves that the range parameter for the V-copulas are systematically higher than Gaussian copulas for both regions, but more significant results are discovered in the region of Bavaria.
- Precipitation amounts for most events exhibit an asymmetric dependence structure for both regions, but more significant results are found in the region of Singapore.
- Precipitation with high-intensity over the whole region of Singapore tend to exhibit a stronger positive asymmetric dependence structure than low-intensity ones. This implies that precipitation with high intensity tends to occur in a clustering manner due to the local convective precipitation. In contrast, precipitation with high-intensity over the whole region of Bavaria tend to exhibit a weaker positive asymmetric dependence structure than low-intensity ones. This might be due to the fact that the size of the area of Bavaria is too large.
- Positive-skewed precipitation values over the whole region of Singapore do not generate the strong positive asymmetric dependence structure. In contrast, right-skewed precipitation amounts over the entire region of Bavaria trigger the strong positive asymmetric dependence structure due to the local convective precipitation occurring in a small part of the region.

# 8 Consequences of asymmetric dependence based copulas on spatial extremes

## 8.1 Introduction

Recently, spatial distributed hydrological modeling has attracted more attention for many hydrologists to develop and implement it into various catchment areas from small into large sizes at the finer spatial scales. Unlike the lumped hydrological models whose inputs and parameters are assumed to be spatially homogeneous using averaged inputs, parameters, and processes over a catchment area, the distributed models consider the spatial variability of inputs, processes, and parameters.

Consequently, the distributed hydrological models require reliable gridded precipitation as main input (Bárdossy and Pegram, 2013; Gyasi-Agyei and Pegram, 2014). Availability of daily precipitation data at the desired grid spacing and a record length enhances the capabilities of hydrological models used for water balance studies, sediment yield, sensitivity analysis of reservoirs, urban water management, flood risk management and assessment of climate change (Gyasi-Agyei, 2016).

Unfortunately, the main input of daily precipitation is frequently not available in terms of continuous temporal observation and spatial coverage at the desired grid size and sufficient record length due to a very limited number of stations or error of measurement instruments (Gyasi-Agyei, 2016). Precipitation spatial observation points represented by gauge stations just provide limited information scattered over the catchments with low gauge densities (Gyasi-Agyei and Pegram, 2014). A good quality spatial simulation of gridded precipitation is, therefore, required for a better spatial distributed hydrological model, as the model outputs are dependent on the quality of the inputs.

In fact, gridded precipitation models have been developed with a variate of different approaches. However, most recent models were developed on the basis of the concept of underlying Gaussian assumptions such as Gaussian copula and Generalized Linear Model (GLM) with uniform marginals (Kleiber et al., 2012; Serinaldi, 2009; Serinaldi and Kilsby, 2014; Wilks, 1998), even though zero precipitation amounts were also treated as latent spatial processes (Serinaldi and Kilsby, 2014).

In addition, most past studies implemented gridded model with coarse resolutions (Kleiber et al., 2012). Hence, a gridded precipitation model based on the assumption of symmetric

Gaussian dependence might yield an unrealistic estimation since precipitation amounts are characterized behaving the asymmetric spatial dependence.

In this chapter, spatial stochastic simulations of gridded daily precipitation are performed using both Gaussian and V-copula models. The Gaussian copulas represent the spatial precipitation models based on the symmetric spatial dependence whereas the asymmetric spatial dependence based models are represented by the V-copulas. The consequences of both models on the spatial extremes of daily precipitation can, therefore, be quantified. The concept of regionalization approach is adopted which assumes that the precipitation occurrences at a time step are treated as a single realization of spatial random variables over the whole spatial domain of interest. The spatial stochastic simulations are implemented in the region of Singapore with the spatial domain size 100 km $\times$ 100 km and in the region of Bavaria with the spatial domain size 500 km $\times$ 500 km. The spatial resolution of 1 km $\times$ 1 km is used for both regions.

## 8.2 Marginal distributions

A copula function describes the dependence structure of multivariate distribution functions independently from the marginal distributions in the unit hypercube as mentioned in section 4.1. A joint distribution can be carried out separately from the analysis of the marginal distributions. However, an adequate understanding of the marginal distributions is necessary in order to be able to estimate local values at the marginal domain from the dependence structure of the joint distribution in a more precise and reasonable way.

In this case, precipitation values at a specific time event independently are treated as the univariate marginal distribution without considering temporal correlations. This is a similar treatment to the univariate spatial distribution as discussed in chapter 3. This is because the spatial stationary assumption is taken over the whole spatial domain of interest. As a consequence, for any given time event of precipitation $t_j$, the marginal distributions at each site $x_i$ are assumed to be the same for all locations as illustrated follows (Bárdossy and Pegram, 2013).

$$F_{x_i,t_j}(z) = F_{t_j}(z) \tag{8.1}$$

Due to intermittent processes, the marginal distributions of precipitation values might consist of zero values. These are modelled by using a mixed discrete and continuous distribution as introduced by Bárdossy and Pegram (2009, 2013) and Serinaldi (2009). Here, the zero precipitation amounts at the marginal distribution are treated as latent variables underlying a continuous distribution function corresponding to the probability of dry stations ($P_0(t_j)$) which is calculated by the total number of stations with zero precipitation ($n_0(t_j)$) divided by the total number of stations $n$. Incorporating the dry and the wet precipitation amounts in the spatial copula model could improve the correlation estimation (Bárdossy and Pegram, 2013).

Consequently, the marginal distribution function of precipitation amounts $F_{t_j}(z)$ is composed of the two parts, namely, the zero precipitation amounts (modeled by a discrete dis-

tribution) and non-zero precipitation amounts (modeled by a continuous distribution) as described in the following equation.

$$\begin{aligned} F_{t_j}(z) &= P(Z(x_i, tj) \leq z) \\ &= \begin{cases} p_0(t_j) & \text{if } z = 0 \\ p_0(t_j) + G_{t_j}(z)(1 - p_0(t_j)) & \text{if } z \geq 0 \end{cases} \end{aligned} \tag{8.2}$$

Where $Z(x_i, t_j)$ denotes the precipitation amounts measured at location $x_i$ and at time $t_j$, $G_{t_j}$ represents the distribution of the positive precipitation amounts, and $p_0(t_j)$ is the probability that a station is dry or $z(x, tj) = 0$.

The distribution function $G$ is fitted from the observed precipitation amounts using an univariate distribution. Empirical evidence reveals that the Weibull distribution generally fits quite well to the univariate spatial distribution of precipitation amounts in the region of Singapore, whereas the non-parametric Gamma kernel function underperforms as pointed out in chapter 3. In contrast, the non-parametric Gamma kernel function exhibits the best performances of fitting the univariate spatial distribution of precipitation amounts in the region of Bavaria. Since spatial stochastic simulations using the non-parametric Gamma kernel function for fitting the marginal distribution requires an extremely time-consuming computation on the computer calculation, the Weibull parametric distribution is, therefore, adopted for the further simulation experiments.

## 8.3 Simulation experiments

Gaussian copulas, which are characterized by symmetric dependence structures, are not capable of modelling the asymmetric spatial dependence structure. Precipitation amounts at daily scale often exhibit strong asymmetric behavior, not only of the skewed distributions on the univariate marginal distribution but also of the asymmetric spatial dependence structures in the multivariate senses as mentioned. Evaluation of Gaussian copula-based model in precipitation modelling is necessary to conduct, since a lot of spatial precipitation models are implemented using the Gaussian dependence assumption. The evaluations of the performances of the model are carried out based on the spatial extremes through spatial stochastic simulations. These simulation experiments under the Gaussian copulas are evaluated in comparison to the V-copulas. The V-copula models are capable of modelling asymmetric dependence structures.

The performances of both Gaussian and V-copula simulations are evaluated through the spatial extremes using the same marginal distribution. When the same marginal distribution is used for the spatial simulation experiments, by definition, both copula models should yield the same basic statistical values of the simulated precipitation, such as the average value, the sum value, the minimum value and the maximum value over the whole spatial domain of interest. On the other hand, however, both Gaussian and V-copulas have a different structure of spatial dependence. As a result, the spatial simulations using the same

marginals could give a different distribution of spatial extremes between both Gaussian and V-copulas.

Evaluation of both Gaussian and V-copulas are carried out using three different spatial simulation experiments. Each simulation experiment consists of sufficient simulations over the whole spatial domain, which are randomly selected from the historical precipitation occurrences using bootstrap random sampling. Each spatial simulation is conducted for each time step separately. These simulation experiments are implemented in different climate regions, in Singapore and Bavaria. The procedures of the simulation experiments are explained in details in the next subsections.

In general, for each spatial stochastic simulation, simulated precipitation at a regular grid size of 1 km $\times$ 1 km are generated within a spatial domain size of 100 km $\times$ 100 km for the region of Singapore and 500 km $\times$ 500 km for Bavaria. Simulations at large scales with a lot of repetition procedures require a huge computation cost. Hence, the Fast Fourier Transform (FFT) based Gaussian simulation is implemented for both Gaussian and V-copulas. The FFT based Gaussian simulation basically depends only on the range parameter. For the Gaussian copulas, precipitation simulations in the normal domain obtained from the FFT simulations are transformed into uniform distribution and then transformed back using the inverse of the marginal distribution to achieve simulated precipitation in the original domain. For the V-copula models, the precipitation simulations under normal domain are transformed into the new non-Gaussian domain using the parameters $m$ and $k$, and then transformed back into original domain using inverse of the marginal distribution. Notice that, the spatial dependence structures of both Gaussian and V-copulas for each time step are taken from the parameters which are estimated using the method explained in chapter 7.

### 8.3.1 Experiment-1: Spatial clustering of maximum gridded precipitation

Precipitation events physically tend to occur in a clustering manner at a location (Kilsby et al., 2007). This is because there is a tendency for the clouds to cluster as well. From the point of view of the precipitation mechanism, the high extreme of precipitation also has a tendency to occur in a clustering way. The simulation experiment here is, therefore, conducted in order to evaluate whether both Gaussian and V-copulas could mimic the natural precipitation processes reasonably, especially in terms of clustering spatial extremes. The specific goal of this experiment is to calculate the areal mean precipitation for which the grids are near to the maximum gridded precipitation in the whole spatial domain for both Gaussian and V-copulas. The areal mean precipitation are calculated for any given spatial size which is the closest radius to the maximum precipitation at a grid point. This experiment is carried out for each time step independently and for any given season (DJF, MAM, JJA, or SON) implemented in the regions of Singapore and Bavaria. The details of the simulation procedure are described in the following steps:

1. A precipitation event for given aggregation time is drawn in a bootstrap random sampling with replacement resulting $\hat{X}_{1,t}, ..., \hat{X}_{n,t}$ where $n$ is the total number of precipitation gauge stations at the region of interest and $t$ is a sequence of bootstrap random samplings.

2. The precipitation event drawn above is modelled by the Weibull distribution and the parameters are estimated by the maximum likelihood method as presented in chapter 3.

3. The precipitation variables from the precipitation event drawn above are transformed into the uniform distribution using the rank transformation.

4. Parameters of Gaussian copula $\Gamma_{g,t}$ and V-copula $\Gamma_{v,t}, m_t, k_t$ are estimated for each precipitation event drawn above using the method described in chapter 7.

5. Gaussian random field precipitation simulations are conducted using FFT on the regular grid size $s \times s$ = 1 km $\times$ 1 km on the domain size $d \times d$ = 100 km $\times$ 100 km for the region of Singapore and $d \times d$ = 500 km $\times$ 500 km km for the region of Bavaria using the range parameter from both models obtained above.

6. Gaussian random fields of precipitation are transformed into the uniform distribution for the Gaussian copula or are transformed first to the new non-Gaussian domain using the parameters $m$ and $k$ and then transformed into the uniform distribution for the V-copula model.

7. Gaussian random field precipitation simulations in the uniform domain are transformed back directly into the original domain using the inverse of the marginal of the precipitation values for the Gaussian copula. For the V-copula, the new non-Gaussian random field precipitation simulations in the uniform distribution are transformed to the original domain using the inverse of the marginal.

8. Simulations of the gridded precipitation random fields generated by using the Gaussian copula would yield $Y_{1,t}, ..., Y_{S\times S,t}$ and using the V-copula model would produce $Z_{1,t}, ..., Z_{S\times S,t}$ with $S = \frac{d}{s}$.

9. The maximum gridded precipitation $Y_{max,t}$ for the Gaussian copula and $Z_{max,t}$ (both in mm) for the V-copula at the original grid size $s \times s$ = 1 km $\times$ 1 km from both models are determined as shown in Figure 8.1 which the maximum value is marked with the white circle.

10. The mean gridded precipitation $\bar{Y}_{max,t}$ (in mm) for the Gaussian copula and $\bar{Z}_{max,t}$ (in mm) for the V-copula are calculated among other grid points which are closest to the maximum grid point. The following areal sizes are chosen: 9, 16, 25, 36, 47, 49, 64, 81, 100, 121, 144, 169, 196, 225, 256, 289, 324, 361, and 400 km$^2$.

11. Steps 1-10 are carried out repeatedly until the total number of precipitation events drawn $t$ equals to the total number of precipitation observations $T$, and should be at least 500 times.

12. Density and distribution function of the mean gridded precipitation from both models above are plotted using kernel density estimation for each area size separately in order to investigate visually whether areal mean precipitation from V-copula is higher than the Gaussian copula. The mean values of both models are also calculated.

13. Statistical testing whether the distribution function of the mean gridded precipita-

tion at a given area size from the V-copula simulation is significantly higher than the Gaussian copula are calculated using 1000 bootstrap simulations with the confidence interval of 90%. The statistical tests are conducted using 3 different statistical tests, namely, Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises as presented in chapter 3.

14. Steps 1-13 are conducted repeatedly with $N$=100 times. Thus, rejection or acceptance rate of the null hypothesis that simulated values of the V-copulas are significantly higher than the Gaussian copulas can be determined.

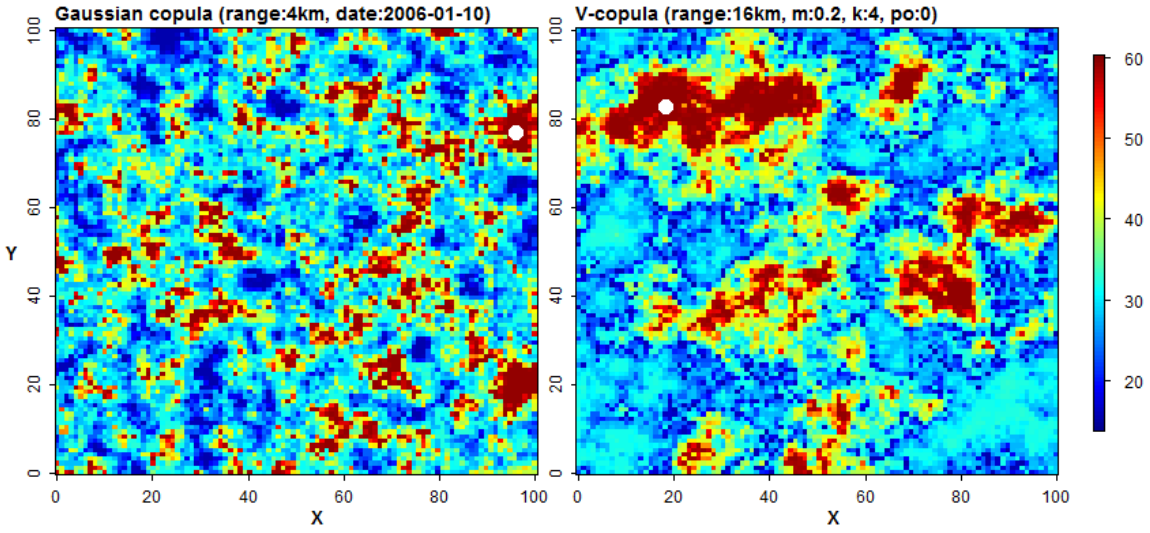



Figure 8.1: Gridded precipitation simulation using the Gaussian copula (left panel) and the V-copula (right panel) for Experiment-1. White circle represents the maximum gridded precipitation.

**Effect of spatial scales** Based on the stochastic simulation experiment results, overall, the V-copulas could capture the behavior of clustering spatial extreme precipitation occurrences more properly than the Gaussian copulas. This judgment is made based on the fact that the areal means of gridded simulated precipitation surrounding the maximum grid point of simulated precipitation, which are generated by the V-copulas, are statistically significantly greater than those areal means which are generated by the Gaussian copulas. Nevertheless, this pattern is dependent on spatial scales and the climate conditions in the region of attention as depicted in Figure 8.2.

Figure 8.2 shows the numbers of acceptance rates for which the V-copulas show significant higher areal mean precipitation which surround the maximum precipitation value at a given spatial scale and in different regions. The investigation results of the region of Singapore are shown in Figure 8.2a, and the results in the region of Bavaria are presented in Figure 8.2b. In order to detect the effect of different seasons, precipitation events for given time aggregation are classified into 4 different seasons, namely, DJF, MAM, JJA, and SON for Singapore, whereas investigations in Bavaria are limited to 2 different seasons, namely, DJF and JJA

due to a very high computational time.

The V-copula simulations exhibit larger areal mean precipitation adjacent to the maximum precipitation amount than the Gaussian copulas affected by the grids sizes. At the size of the original grids (here, 1 km $\times$ 1 km), by definition, the areal mean precipitation should be the same for Gaussian and V-copulas. When the grids size is aggregated to higher scales, such as 3 km $\times$ 3 km, 4 km $\times$ 4 km, or even higher spatial scale with respect to the grids which are the closest radius to the maximum precipitation, then the areal mean precipitation within the areas containing the gridded precipitation can be calculated. The areal mean precipitation of the grids inside the areas represents the clustering of spatial extreme precipitation where precipitation events frequently occur.

The areal mean of gridded precipitation adjacent to the maximum precipitation where is nothing else but the center of the spatial extremes that generated by the V-copulas, is significantly higher than the Gaussian copulas if the original grid size is aggregated to higher spatial scales. However, the fact, that the V-copulas are superior to the Gaussian copulas in the areal mean precipitation surrounding to the maximum precipitation applies only until a certain spatial scale is reached. Beyond this spatial scale, the areal mean precipitation generated by the V-copulas decreases or has a tendency to be similar to the values generated by the Gaussian copulas.

**Case study in Singapore** Overall, the V-copulas exhibit higher areal mean precipitation than the Gaussian copulas at certain spatial scales with different acceptance rates based on the different statistical testing. The acceptance rates tend to increase quickly with increasing spatial scale until a certain spatial scale is reached. Then, the acceptances levels decrease gradually with increasing spatial scale approaching to a zero acceptance rate. In the case of Singapore, the acceptance rates for which the V-copulas yield significantly higher areal mean precipitation than the Gaussian copulas are generally above 0.5 except for the DJF season for area sizes roughly ranging from 4 km $\times$ 4 km to 15 km $\times$ 15 km.

In the JJA season, for example, the V-copulas yield significantly larger areal mean precipitation with acceptance rates above 0.5 for area sizes roughly ranging from 4 km $\times$ 4 km to 15 km $\times$ 15 km for the Kolmogorov-Smirnov test, and from 5 km $\times$ 5 km to 13 km $\times$ 13 km for both Anderson-Darling and Cramer-Von Mises test. In the SON season, for another example, the V-copulas yield significantly greater areal mean precipitation for area sizes from 4 km $\times$ 4 km to 19 km $\times$ 19 km with acceptance rates above 0.5 for all three different statistical tests. In the MAM season, the V-copulas yield significantly greater areal mean precipitation with acceptance rates above 0.5 for cell sizes from 4 km $\times$ 4 km to 15 km $\times$ 15 km for both AD and KS tests and roughly ranging from 4 km $\times$ 4 km to 17 km $\times$ 17 km for CVM test.

Surprisingly, for all three seasons, MAM, SON, and JJA, the area sizes, for which the V-copulas outperform the Gaussian copulas, roughly range from 3 km $\times$ 3 km to 20 km $\times$ 20 km with acceptance rates above 0.25 for all different statistical tests, except for AD and CVM tests in the JJA season. This indicates that the V-copula models could mimic natural processes of precipitation reasonably in terms of spatial extremes.

However, in the DJF season, the acceptance rates for which the V-copulas significantly have

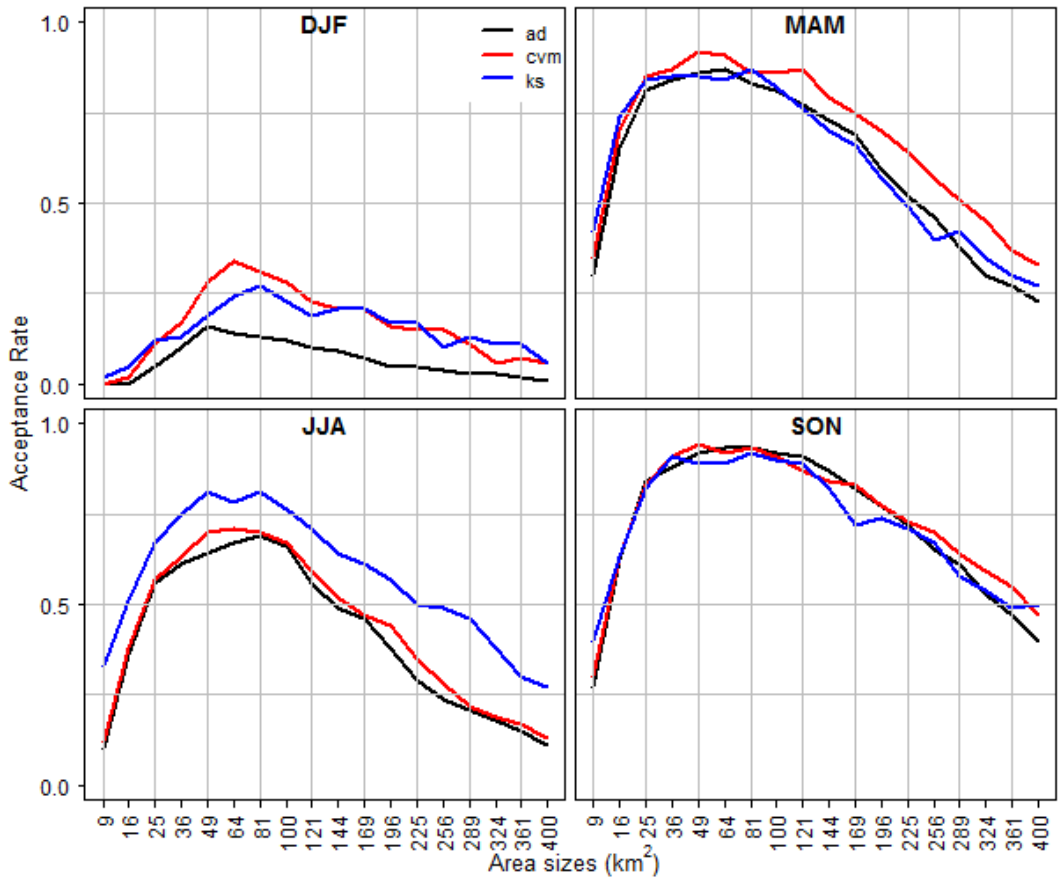
higher simulated values than the Gaussian copulas, are below 0.5 and just around 0.25 for all statistical tests. This might be due to the fact that the number of daily precipitation occurrences, which follow the positive asymmetric spatial dependencies during that season, is also low as found in the previous empirical investigation shown in Figure 6.1.

**Case study in Bavaria** Similar to the region of Singapore, in Bavaria, the V-copulas also exhibit greater areal mean precipitation than the Gaussian copulas at certain spatial scales with different acceptance rates. However, there is a different behavior in spatial scales. In contrast to Singapore, where the V-copulas start to exhibit significantly greater areal mean precipitation than the Gaussian copulas in acceptance levels above 0.5 at a grid size of 4 km $\times$ 4 km, a different finding in Bavaria is that the V-copulas present significantly higher areal mean precipitation in the level of acceptance above 0.5 at the larger spatial scale roughly ranging from 10 km $\times$ 10 km to 13 km $\times$ 13 km.

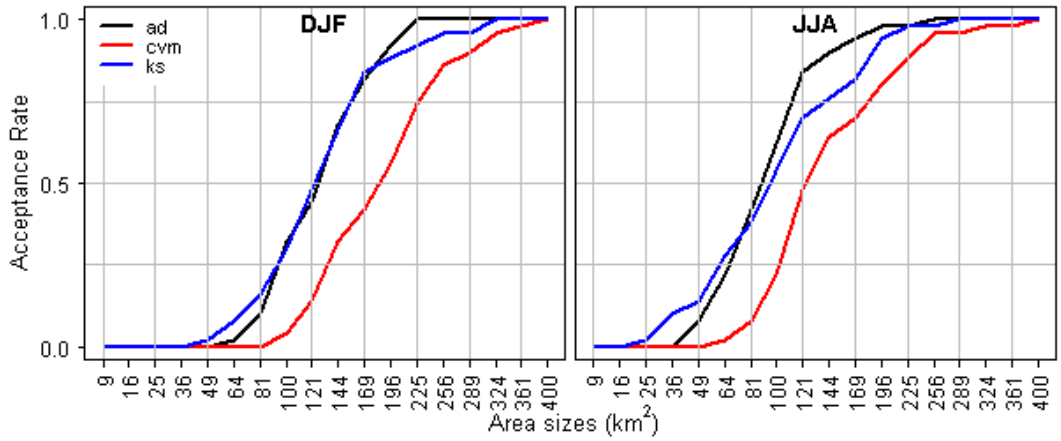
In the DJF season, for instance, the V-copulas exhibit significantly greater areal mean precipitation than the Gaussian copulas with acceptance rates above 0.5 starting from spatial scales of 11 km $\times$ 11 km for both AD and KS statistical tests and 13 km $\times$ 13 km for the CVM statistical test. In the JJA season for another example, the V-copulas exhibit significantly greater areal mean precipitation than the Gaussian copulas with acceptance rates above 0.5 starting from spatial scales of 9 km $\times$ 9 km for both AD and KS statistical tests and 11 km $\times$ 11 km for the CVM statistical test.

When the acceptance rates below 0.5 are considered, the acceptance rates of 0.25 for instance, the V-copulas exhibit significantly greater areal mean precipitation than the Gaussian copulas during the DJF season starting from spatial scales of 10 km $\times$ 10 km for both AD and KS tests, and 12 km $\times$ 12 km for CVM test. Similarly, in the JJA season, the V-copulas exhibit significantly greater areal mean precipitation than the Gaussian copulas in the acceptance rate of 0.25 for area sizes around 8 km $\times$ 8 km for both AD and KS tests, and around 10 km $\times$ 10 km for CVM test.

The areal mean precipitation generated by the V-copulas presents similar results to those simulated by the Gaussian copulas although the original grid size is aggregated to higher spatial scales such as 3 km $\times$ 3 km to 7 km $\times$ 7 km. This means that the V-copula models would yield significantly higher areal mean precipitation for larger areas with respect to the centre of the spatial extremes in comparison to the region of Singapore. This might be because the size of the spatial domain in Bavaria is very large compared to Singapore.




(a) Singapore




(b) Bavaria

Figure 8.2: Acceptance rates for which V-copulas are significantly higher than the Gaussian copulas for the experiment-1 in Singapore (Fig-a) and Bavaria (Fig-b). The vertical axes represent the acceptance rates. Blue lines represent Kolmogorov-Smirnov test. Black lines represent Anderson-Darling test. Red lines represent Cramer-von Mises test. The horizontal axes denote the area size in km$^2$.

### 8.3.2 Experiment-2: Maximum gridded precipitation at different spatial scales

In experiment-1, the investigation just focused on the maximum precipitation, by calculating the areal mean precipitation surrounding to the maximum precipitation. This experiment, however, aims to investigate the areal mean of the maximum gridded precipitation at the different spatial scales over the whole spatial domain. The maximum gridded precipitation here are obtained from all realizations for each original grid point separately. The various spatial scales are created by aggregating from the original grid size (1 km $\times$ 1 km) into higher sizes, for example, 2 km $\times$ 2 km, 4 km $\times$ 4 km, or even larger spatial scale. The areal mean precipitation can then be calculated from the subset with the higher spatial scales consisting of the maximum gridded precipitation at each grid cell from all realizations simulated by the Gaussian and V-copulas. The details of the simulation procedure are described in the following steps.

1. A precipitation event for given aggregation time is drawn in a bootstrap random sampling with replacement resulting $\hat{X}_{1,t}, ..., \hat{X}_{n,t}$ where $n$ is the total number of precipitation gauge station at the region of interest and $t$ is a sequence of bootstrap random samplings .
2. The precipitation event drawn above is modelled by the Weibull distribution.
3. The precipitation variables from the precipitation event drawn above are transformed into the uniform distribution using the rank transformation.
4. Parameters of Gaussian copula $\Gamma_{g,t}$ and V-copula $\Gamma_{v,t}, m_t, k_t$ are estimated for each precipitation event drawn above.
5. Gaussian random field precipitation simulations using FFT are conducted on the regular grid size $s \times s$= 1 km $\times$ 1 km on the domain size $d \times d$= 100 km $\times$ 100 km for the region of Singapore and $d \times d$= 500 km $\times$ 500 km for the region of Bavaria using the range parameter from both models obtained above.
6. Gaussian random field precipitation simulations are transformed into the uniform distribution for the Gaussian copula or are transformed first to the new non-Gaussian domain using the parameters $m$ and $k$ and then transformed into the uniform distribution for the V-copula model.
7. Gaussian random field precipitation simulations in the uniform domain are transformed back directly to the original domain using the inverse of the marginal of the precipitation values for the Gaussian copula. For the V-copula, the new non-Gaussian random field precipitation simulations in the uniform distribution are transformed to the original domain using the inverse of the marginal.
8. Simulations of the gridded precipitation random field generated by using the Gaussian copula would yield $Y_{1,t}, ..., Y_{S\times S,t}$ and using the V-copula model would produce $Z_{1,t}, ..., Z_{S\times S,t}$ with $S = \frac{d}{s}$.
9. Steps 1-8 are conducted repeatedly with the next sample of precipitation events until the number of precipitation events drawn $t$ equals to the total number of precipitation

events $T$ or is at least 500. Hence, the simulated gridded precipitation fields with number of $T$ realizations can be made, $\Big(\{Y_{1,t},...,Y_{S\times S,t}\},...,\{Y_{1,T},...,Y_{S\times S,T}\}\Big)$ from Gaussian copulas and $\Big(\{Z_{1,t},...,Z_{S\times S,t}\},...,\{Z_{1,T},...,Z_{S\times S,T}\}\Big)$ from V-copulas.

10. On each grid point at the regular grid size $s \times s$= 1 km $\times$ 1 km, the maximum simulated gridded precipitation is selected from T realizations in order to get a single realization of spatial random field having a maximum value on each grid point, namely, $Y_{1,max},...,Y_{S\times S,max}$ for Gaussian copulas and $Z_{1,max},...,Z_{S\times S,max}$ for V-copulas.

11. For the case of Singapore, the domain size $d \times d$= 100 km $\times$ 100 km is divided equally into several subsets with regular dimensions, such as 2 x 2 km$^2$, 4 x 4 km$^2$, 10 x 10 km$^2$, 20 x 20 km$^2$, and 25 x 25 km$^2$. The maximum simulated gridded precipitation are averaged over grids on each subsets. Every subset with regular dimension would yield $\bar{Y}^j_{1,max},...,\bar{Y}^j_{S\times S,max}$ for Gaussian copulas and $\bar{Z}^j_{1,max},...,\bar{Z}^j_{S\times S,max}$ for V-copulas, where $j = (2\times 2),(4\times 4),(10\times 10),(20\times 20),$ and$(25\times 25)$km (See the Figure 8.3).

12. For the case of Bavaria, the domain size $d \times d$= 500 km $\times$ 500 km is divided equally into several subsets with regular dimensions, such as 2 x 2 km$^2$, 4 x 4 km$^2$, 10 x 10 km$^2$, 20 x 20 km$^2$, 25 x 25 km$^2$, 50 x 50 km$^2$, and 100 x 100 km$^2$. The maximum simulated gridded precipitation are averaged over grids on each subsets.

13. The density and distribution functions of the mean simulated gridded precipitation at a given subsets over the whole spatial domain are estimated using kernel density estimation. Notice that, vector lengths at each subset from both models are the same number, namely, $\frac{d}{j}\times\frac{d}{j}$.

14. Statistical testing, whether the distribution function from V-copula simulation is a significantly higher than Gaussian copula, are calculated using 1000 bootstrap simulations with the confidence interval of 90%. The statistical tests are conducted using 3 different statistical tests, namely, Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises.

15. Steps 1-14 are conducted repeatedly with $N$ =100 times. Thus, rejection or acceptance rate of the null hypothesis that simulated values of V-copulas are significantly higher than the Gaussian copulas can be carried out.

**Effect of spatial scales** Based on the stochastic simulation experiment results, there is a pattern for which the V-copulas yield significantly greater areal mean extreme precipitation than the Gaussian copulas when the original grid size (in this study $1\times 1$ km) is aggregated into greater sizes until reaching the specific aggregation scale where it will decrease and then disappear. The extreme precipitation for each grid is calculated based on the maximum precipitation among all realizations. This pattern is also really dependent on spatial scales and the climate condition in the regions of attention as depicted in Figure 8.4.

Figure 8.4 presents the acceptance rates for which the V-copulas perform significantly higher than the Gaussian copula with respect to the areal mean extreme precipitation for each grid

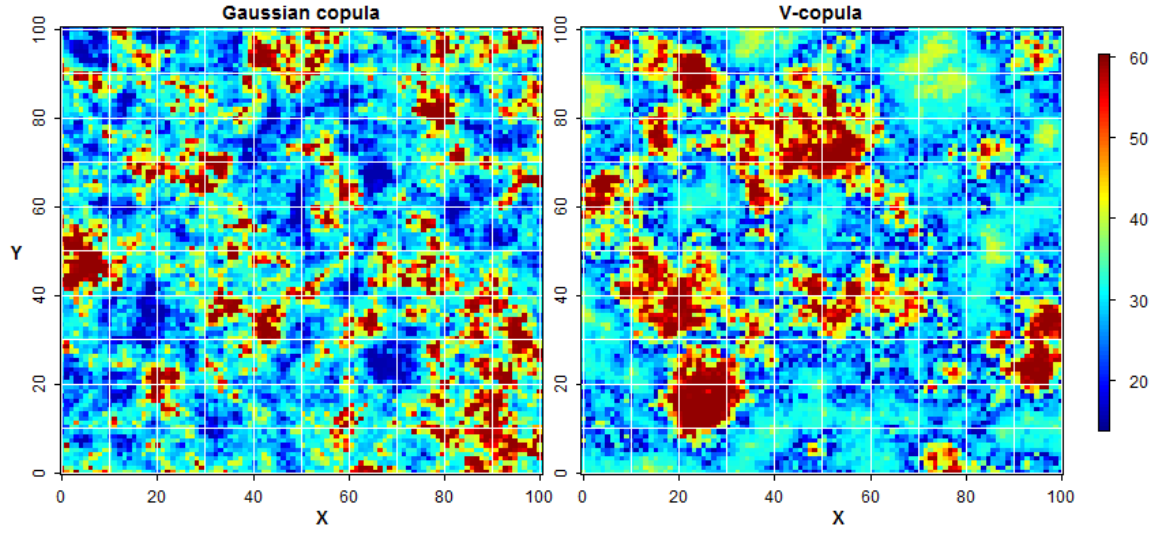



Figure 8.3: Gridded precipitation simulation using the Gaussian copula (left panel) and the V-copula (right panel) for Experiment-2.

at a given spatial scale and in different regions, in Singapore and Bavaria. The investigation results of the region of Singapore are shown in Figure 8.4a and empirical findings of the region of Bavaria are presented in Figure 8.4b. In order to detect the effect of different seasons on both models, precipitation events are classified into 4 different seasons (namely, DJF, MAM, JJA, and SON) for Singapore, whereas investigations in Bavaria are restricted to 2 different seasons (namely, DJF and JJA) due to a very expensive computational cost.

The V-copula simulations exhibit higher areal mean extreme gridded simulated precipitation over all spatial domain of interest. At the original grids size (here 1 km $\times$ 1 km),by definition, the areal mean extreme precipitation over all realizations should be the same for both Gaussian and V-copulas. When the size of the grids is aggregated into higher scales, such as 2 km $\times$ 2 km, 4 km $\times$ 4 km, or even higher spatial scales with respect to an equally regular grid over the spatial domain, then the areal mean extreme precipitation over the spatial domain can be calculated. The V-copulas generate higher simulated values of the areal mean extreme precipitation than the Gaussian copula. This applies to a certain spatial scale. Beyond this spatial scale, the areal mean extreme precipitation which are generated by the V-copulas would decrease or has a tendency to be similar with the values which are generated by the Gaussian copulas.

**Case study in Singapore** In the region of Singapore, the investigation results show that the acceptance rates for which the V-copulas exhibit higher values of the averages of the maximum precipitation than the Gaussian copula, are roughly below 0.5. The highest acceptance rate is reached at a grid size of 2 km $\times$ 2 km for all seasons. In this case, the V-copulas, that exhibit higher simulated values of the areal mean maximum precipitation than the Gaussian copulas, are more pronounced during the DJF season and followed by the MAM and SON seasons in the similar significant levels, and finally followed by the

season of JJA. This pattern is contradictory to the empirical results from the experiment-1 mentioned in section 8.3.1 where the V-copulas exhibit significantly higher simulated values than the Gaussian copulas for both MAM and SON seasons and followed by the seasons of JJA and DJF, respectively.

In the DJF season for example, the acceptance rates for which the V-copulas present significantly higher areal mean maximum gridded precipitation than the Gaussian copulas at the aggregation size of 2 km $\times$ 2 km are roughly 0.57 for AD test, 0.42 for CVM test, and 0.45 for KS test. Then, the acceptance rates decline to the levels of 0.39, 0.29, and 0.35 at the aggregation scale of 4 km $\times$ 4 km for the AD test, CVM test, and KS test, respectively. For greater aggregation sizes, the acceptance rates decrease dramatically close to zero. This implies that the V-copulas tend to yield similar simulated values of areal mean precipitation compared to the Gaussian copulas for aggregation sizes of greater than 4 km $\times$ 4 km as the case of the original grid size of 1 km $\times$ 1 km.

A similar behaviour occurs in both MAM and SON seasons. During those seasons, the V-copulas exhibit higher values of areal mean maximum gridded precipitation than the Gaussian copulas with similar acceptance rates, around 0.25 for both AD and KS tests and roughly 0.17 for CVM test at a grid size of 2 km $\times$ 2 km. At an aggregation scale of 4 km $\times$ 4 km, the acceptance rates during both seasons decrease to the levels roughly ranging from 0.06 to 0.15 for all statistical tests.

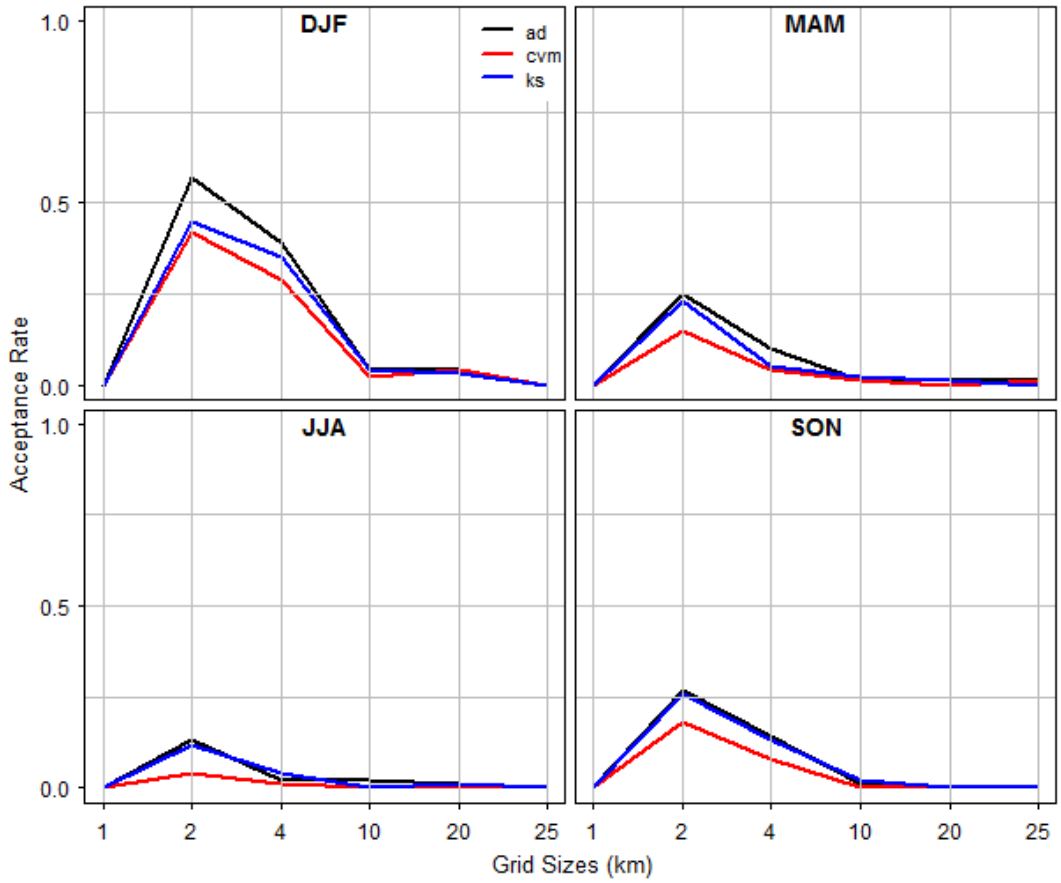
In the JJA season in contrast, the highest acceptance rate for which the V-copulas exhibit significantly higher areal mean maximum gridded precipitation than the Gaussian copula occurred at a grid size of 2 km $\times$ 2 km, is very low compared to other seasons, just around 0.15 for both AD and KS tests and about 0.05 for CVM test as depicted in Figure 8.4a. At an aggregation scale of 4 km $\times$ 4 km, the acceptance rates during the JJA season decrease dramatically close to zero for all statistical tests.

Overall, the acceptance rates are below 0.5 but greater than zero at the spatial scales of 2 km $\times$ 2 km and 4 km $\times$ 4 km. This indicates that most of V-copula simulations exhibit similar estimated values of areal mean maximum gridded precipitation compared to the Gaussian copulas. Some of V-copula simulations, however, still present higher simulated values of areal mean maximum precipitation than the Gaussian copulas. Nevertheless, for a conservative approach in the flood risk assessment, the V-copula should be considered for the planning and design.

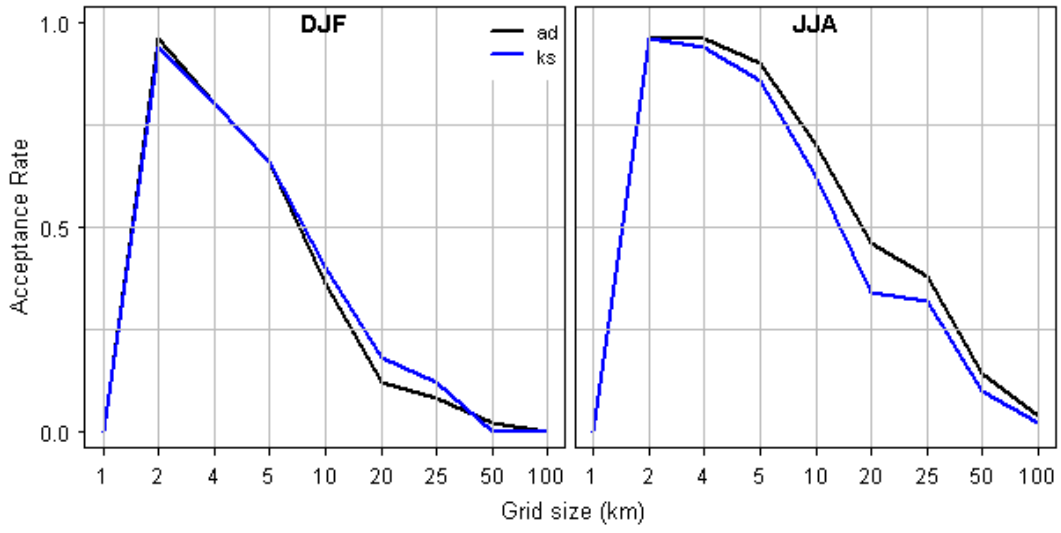
**Case study in Bavaria** In contrast to the region of Singapore, where the acceptance rates for which the V-copulas exhibit higher values of the averages of the maximum gridded precipitation than the Gaussian copula, is lower roughly below 0.5, the acceptance rates in Bavaria are higher in general. Based on the statistical tests from simulation experiments reveal that the acceptance rates are above 0.5 for the grid sizes from 2 km $\times$ 2 km to 5 km $\times$ 5 km for the DJF season and for the grid sizes from 2 km $\times$ 2 km to 10 km $\times$ 10 km for the JJA season as shown in Figure 8.4b. In this case, the V-copulas, which exhibit a higher estimation of the areal mean maximum gridded precipitation than the Gaussian copulas, are more pronounced during the JJA season and followed by the DJF season.

In the JJA season for example, the acceptance rates at a grid size of 2 km $\times$ 2 km is around 0.96 for the AD test and remain the same until at a grid size of 4 km $\times$ 4 km. The acceptance rates decline smoothly into the levels of 0.9, 0.7, and 0.4 for the grid sizes of 5 km $\times$ 5 km, 10 km $\times$ 10 km, and 20 km $\times$ 20 km, respectively. Further greater grid sizes would yield acceptance rates below 0.5. Similar to the AD test, the KS test presents the acceptance rates of around 0.96 at a grid size of 2 km $\times$ 2 km but it then decreases gradually into 0.94, 0.86, 0.62, and 0.34 for grid sizes of 4 km $\times$ 4 km, 5 km $\times$ 5 km, 10 km $\times$ 10 km, and 20 km $\times$ 20 km, respectively. Further higher grid sizes would also yield acceptance rates below 0.5.

In the DJF season for another example, the acceptance rates at a grid size of 2 km $\times$ 2 km are around 0.95 for both AD and KS test and decrease gradually to the levels of 0.8 and 0.66 for grid sizes of 4 km $\times$ 4 km and 5 km $\times$ 5 km, respectively. Investigations at larger grid sizes result in acceptance rates below 0.5.




(a) Singapore




(b) Bavaria

Figure 8.4: Acceptance rates for which V-copulas are significantly higher than the Gaussian copulas for the experiment-2 in Singapore (Fig-a) and Bavaria (Fig-b). The vertical axes represent the acceptance rates . Blue lines represent Kolmogorov-Smirnov test. Black lines represent Anderson-Darling test. Red lines represent Cramer-von Mises test. The horizontal axes denote the grid sizes in km.

### 8.3.3 Experiment-3: Spatial clustering of mean gridded precipitation at the specific location

Unlike the experiment-1 where the areal mean precipitation are calculated from the grids surrounding to the maximum precipitation for each realization, this experiment aims to investigate the areal mean precipitation that closes to a fixed location for the both Gaussian and V-copula models. In this study, the fixed location is selected in the middle of the spatial domain of interest. This experiment is also carried out for each precipitation events at a time step independently and for any given season (DJF, MAM, JJA, or SON) implemented in the regions of Singapore and Bavaria. The details of the simulation procedure are described in the following steps.

1. Precipitation event for given aggregation time is drawn in a bootstrap random sampling with replacement resulting $\hat{X}_{1,t}, ..., \hat{X}_{n,t}$ where $n$ is the total number of precipitation gauge stations at the region of interest and $t$ is a sequence of bootstrap random samplings.
2. The precipitation event drawn above is modelled by the Weibull distribution and the parameters are estimated by the maximum likelihood method.
3. The precipitation variables from the precipitation event drawn above are transformed into the uniform distribution using the rank transformation.
4. Parameters of Gaussian copula $\Gamma_{g,t}$ and V-copula $\Gamma_{v,t}, m_t, k_t$ are estimated for each precipitation event drawn above.
5. Gaussian random field precipitation simulations are conducted using FFT on the regular grid size $s \times s$ = 1 km $\times$ 1 km on the domain size $d \times d$ = 100 km $\times$ 100 km for the region of Singapore and $d \times d$ = 500 km $\times$ 500 km for the region of Bavaria using the range parameter from both models obtained above.
6. Gaussian random fields of precipitation are transformed into the uniform distribution for the Gaussian copula or are transformed first to the new non-Gaussian domain using the parameters $m$ and $k$ and then transformed into the uniform distribution for the V-copula model.
7. Gaussian random field precipitation simulations in the uniform domain are transformed back directly into the original domain using the inverse of the marginal of the precipitation amounts for the Gaussian copula. For the V-copula, the new non-Gaussian random field precipitation simulations in the uniform distribution are transformed to the original domain using the inverse of the marginal.
8. Simulations of the gridded precipitation random fields generated by using the Gaussian copula would yield $Y_{1,t}, ..., Y_{S\times S,t}$ and using the V-copula model would produce $Z_{1,t}, ..., Z_{S\times S,t}$ with $S = \frac{d}{s}$.
9. A grid point reference is selected, for example, in the middle of the spatial domain $c$ as shown in Figure 8.5 which are marked with the white circle.

10. The mean gridded precipitation $\bar{Y}_{c,t}$ (in mm) for the Gaussian copula and $\bar{Z}_{c,t}$ (in mm) for the V-copula are calculated among other grid points which are closest to the grid point reference. The areal sizes are chosen such as 9, 16, 25, 36, 47,49, 64, 81, 100, 121, 144, 169, 196, 225, 256, 289, 324, 361, and 400 km$^2$.

11. Steps 1-10 are carried out repeatedly until the total number of precipitation events drawn $t$ equals to the total number of precipitation observations $T$, or at least T=500 times.

12. Density and distribution function of the mean gridded precipitation from the both models above are plotted using kernel density estimation for each area size separately in order to investigate visually whether V-copula is higher than the Gaussian copula. Notice that, vector lengths from those models are the same number, namely, $T$.

13. Statistical testing whether the distribution function of the mean gridded precipitation at a given area size from the V-copula simulation is significantly higher than the Gaussian copula are calculated using 1000 bootstrap simulations with the confidence interval of 90%. The statistical tests are conducted using 3 different statistical tests, namely, Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises.

14. Steps 1-13 are conducted repeatedly with $N$=100 times. Thus, rejection or acceptance rate of the null hypothesis that simulated values of V-copulas are statistically significantly higher than the Gaussian copulas can be determined.

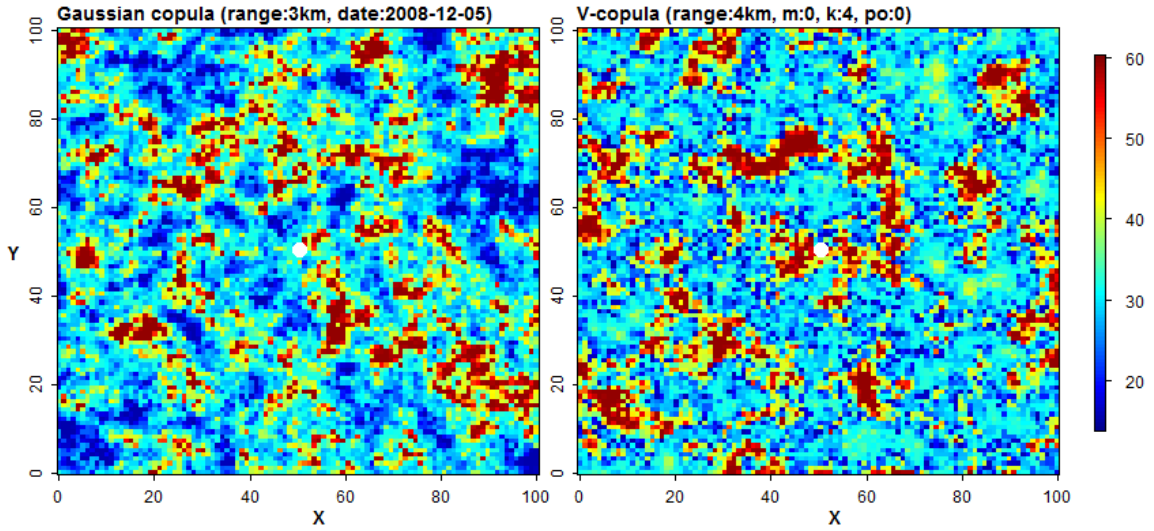



Figure 8.5: Gridded precipitation simulation using the Gaussian copula (left panel) and the V-copula (right panel) for Experiment-3. White circle represents a grid point reference, for example in the middle of domain.

**Effect of spatial scales** Overall, the results of the spatial stochastic simulations reveal that the V-copulas exhibit significantly higher areal mean precipitation than the Gaussian copulas at the smallest area size (3 km $\times$ 3 km) surrounding to a fixed location in the middle

of the spatial domain. However, the significant level for which the V-copulas perform higher simulated values than the Gaussian copulas at the smallest area size is low with acceptance rates below 0.5. The acceptance rates for which the V-copulas perform higher simulated values than the Gaussian copulas generally decrease systematically with increasing area sizes as shown in Figure 8.6.

Figure 8.6 presents the acceptance rates for which the V-copulas perform higher simulated values than the Gaussian copulas with respect to the areal mean precipitation surrounding to a fixed location in the middle of the spatial domain in the region of Singapore (Figure 8.6a) and the region of Bavaria (Figure 8.6b). Investigations in Singapore are carried out using 4 different seasons, namely, DJF, MAM, JJA, and SON, whereas in Bavaria, two different seasons, namely, DJF and JJA are considered in this study.

**Case study in Singapore** In Singapore, there is a clear pattern that the acceptance rates decrease gradually with increasing the area sizes as shown in Figure 8.6a. Overall, the acceptance rates for which the V-copulas exhibit significantly higher areal mean precipitation than the Gaussian copulas during the MAM season are higher than all other three seasons and followed by the SON, JJA, and DJF seasons, respectively. This general trend is similar to the result of the experiment-1, but the acceptance rates of this experiment-3 are lower than the experiment-1.

In the MAM season for example, the acceptance rates at the smallest area of 3 km $\times$ 3 km are above 0.25 for all statistical tests until the grid size of 7 km $\times$ 7 km. The acceptance rates at the smallest area of 3 km $\times$ 3 km are roughly 0.4, 0.3, 0.32 for AD, CVM and KS test, respectively, and then they decline gradually with the increase of grid sizes until the acceptance rates are close to zero at an area sizes approximately 18 km $\times$ 18 km for the AD test and 20 km $\times$ 20 km for both CVM and KS test.

In the SON season for another example, the acceptance rates at the smallest area of 3 km $\times$ 3 km are roughly 0.28, 0.19, and 0.25 for AD, CVM and KS test, respectively, and then decline gradually with the increase of grid sizes until the acceptance rates are close to zero at an area sizes approximately 20 km $\times$ 20 km, namely, 0.01, 0.03, and 0.05 for the AD, CVM and KS test, respectively.

In the JJA season, the acceptance rates at the smallest area of 3 km $\times$ 3 km are below 0.25 for all statistical tests. The acceptance rates at the smallest area of 3 km $\times$ 3 km are roughly 0.22, 0.16, and 0.16 for AD, CVM and KS test, respectively. Then, they decline gradually with the increase of grid sizes until the acceptance rates are close to zero at an area size approximately 15 km $\times$ 15 km, and 16 km $\times$ 16 km for the AD and CVM, respectively. For the KS test, it increases slightly to an acceptance level of 0.24 at a grid size of 6 km $\times$ 6 km before it goes down gradually to a level 0.02 at a grid size of 20 km $\times$ 20 km.

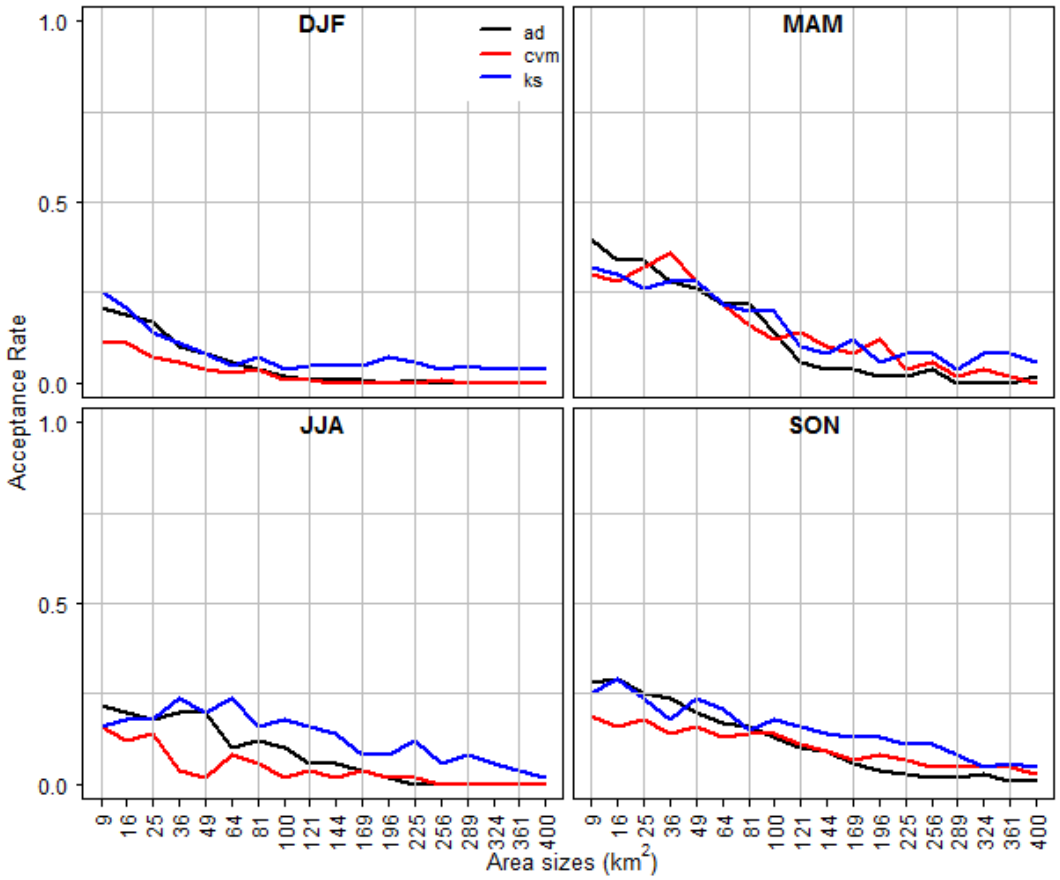
In the DJF season, the acceptance rates are below 0.25 for all statistical tests even at the smallest grid size from the reference point. The acceptance rate at the smallest area of 3 km $\times$ 3 km is roughly 0.21, and then declines gradually with the increase of grid sizes until an acceptance rate of zero for area sizes approximately 14 km $\times$ 14 km for the AD test. The CVM test yields slightly smaller values of the acceptance rate of around 0.11 at a grid size of

3 km $\times$ 3 km, and then it deteriorates until an acceptance rate of zero at the grid sizes of 12 km $\times$ 12 km. In contrast, the KS test results slightly greater acceptance rate at the level 0.25 at a grid size of 3 km $\times$ 3 km and it continuously decreases to the acceptance level of 0.04 at a grid size of 20 km $\times$ 20 km.

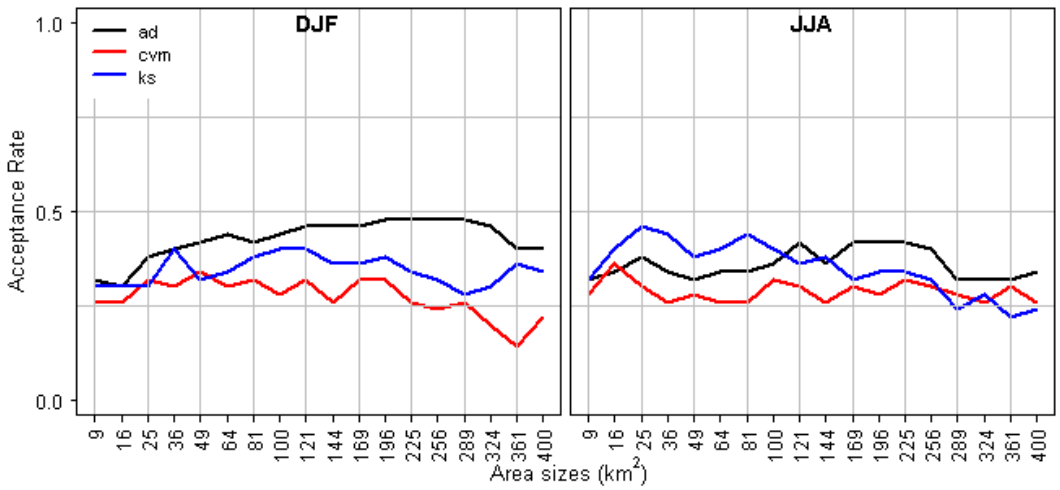
**Case study in Bavaria** In the region of Bavaria, the acceptance rates generally fluctuate and tend to be stationary at the levels from 0.25 to 0.5 over the grid sizes from 3 km $\times$ 3 km to 20 km $\times$ 20 km as presented in Figure 8.6b. This might be because the spatial domain for Bavaria is too large (500 km $\times$ 500 km) in comparison to the spatial resolution grid (1 km $\times$ 1 km).

For instance, in the DJF season, the acceptance rates at grid size of 3 km $\times$ 3 km are about 0.32, 0.26, and 0.3 for the AD, CVM, and KS test, respectively. The acceptance rates then slightly increase to the levels of 0.46, 0.26, and 0.36 at a grid size of 12 km $\times$ 12 km and change slightly into the levels of 0.4, 0.26, and 0.34 at a grid size of 20 km $\times$ 20 km for the AD, CVM, and KS test, respectively.

Similarly, for the grid sizes of 3 km $\times$ 3 km, 12 km $\times$ 12 km, and 20 km $\times$ 20 km in the JJA season, the acceptance rates based on the AD test are about 0.32, 0.36, and 0.34, respectively. Using the CVM test, the acceptance levels are about 0.28, 0.26, and 0.26, respectively. The acceptance rates based on the KS test are about 0.32, 0.38, and 0.24 for the grid sizes of 3 km $\times$ 3 km, 12 km $\times$ 12 km, and 20 km $\times$ 20 km, respectively




(a) Singapore




(b) Bavaria

Figure 8.6: Acceptance rates which V-copula is significantly higher than the Gaussian copula for the experiment-3 in Singapore (Fig-a) and Bavaria (Fig-b). The vertical axes represent the acceptance rates . Blue lines represent Kolmogorov-Smirnov test. Black lines represent Anderson-Darling test. Red lines represent Cramer-von Mises test. The horizontal axes denote the area sizes in km$^2$.

## 8.4 Consequences of the asymmetric dependence based spatial models on the floods risk assessment

Several severe flooding events over the last few years have caused terrible socio-economic impacts including extensive property damage, disrupted businesses and lives, and loss of life. Of all natural disasters worldwide, floods are not only the most costly natural disaster in property damage, but have also affected the most people in businesses and lives (Kousky and Walls, 2014; Strömberg, 2007). Flood risk assessment analyses is, therefore, required to mitigate this natural disasters. The relationship between the flood flow and the cost of damage to infrastructure is generally non-linear. Consequently, an underestimate of streamflow would result the cost of damage which may outweigh the alternative cost of adopting a larger design flood (Bennett et al., 2015).

An essential input of flood risk assessment analyses is the spatial intensity of extreme precipitation events over a catchment. Underestimating extreme spatial precipitation will typically lead to an underestimation of streamflow resulting in costly damage. It is common practice for estimating extreme spatial precipitation to ignore the spatial variation in precipitation intensity and just convert intensity estimates of extreme point to extreme spatial precipitation. Improving flood design of spatial precipitation extreme through understanding of the natural process of precipitation is required to reduce the economic impact and cost to society significantly (Bennett et al., 2015).

Recently, numerous spatial precipitation models incorporating spatial variability on the basis of grid cells have been extensively developed but mostly based on underlying symmetric Gaussian dependence. Studies on the effect of the symmetric Gaussian dependence based model have been conducted. Spatial precipitation models based on symmetric spatial dependence are represented here by Gaussian copulas, and asymmetric spatial dependence are represented by V-copulas. Both models are demonstrated using daily precipitation on the basis of grid cells.

Empirical investigations focusing on spatial extremes reveal that the Gaussian copulas frequently exhibit lower spatial estimations than the V-copulas. Precipitation occurrences, which typically behave in clustering manner, could not be modelled reasonably by Gaussian copulas as expected. This is because the spatial extreme precipitation estimated by Gaussian copulas are systematically lower than V-copulas.

# 9 Conclusions

Pearson's correlation has been broadly used in a variety of applications in order to quantify the strength of linear dependence between two variables. Unfortunately, the correlation as measure has some drawbacks, such as restriction to the assumption of normality, requiring a linear relationship, and being sensitive to outliers. Spearman's rank correlation is one of the alternatives.

Spatial Spearman's rank correlations of precipitation fields are computed using censored bivariate Gaussian copulas on the basis of empirical bivariate copulas. These are constructed using two different approaches (geostatistical and pairwise). Empirical investigations located in different climate regions (namely, Singapore, Bavaria, and Baden-Württemberg) reveal that the spatial Spearman's rank correlation in the region of Singapore is significantly lower compared to Baden-Württemberg and Bavaria, respectively. The spatial correlation in Singapore is very low, indicating that high spatial variability exists caused by local convective precipitation being driven mainly by vertical processes. Southern Germany (Bavaria and Baden-Württemberg), on the other hand, is characterized rather by large scale frontal systems driven by atmospheric variables influencing local variability of precipitation mainly through horizontal processes.

The spatial rank correlations estimated using the pairwise approach, with simultaneous temporal occurrences; exhibit systematically higher values than using the geostatistical approach. This empirical evidence is consistent with different climate regions, that of Singapore and the South of Germany. The reason for this is mainly caused by driving processes with large spatial extent, where the geostatistical approach assumes spatial stationarity over the whole spatial domain of interest which the influence of the mean value is neglected. Spearman's rank correlations increase with the increase of temporal scales. This tendency is more decisive for the pairwise approach than the geostatistical approach.

The correlation coefficient is a fundamental parameter for spatial modelling, especially for the Gaussian copulas where the dependence structure is completely determined by the correlation coefficient matrix. The correlation coefficient is, however, not enough to describe the structure of dependency comprehensively since this measure just quantifies the linear association based on the entire dependence structure of the two random variables. This measure is not capable of detecting dependency structure between high and low values, for example. A measure of asymmetry which can incorporate zero precipitation amounts are introduced on the basis of empirical bivariate copulas. Asymmetric models offer a solution to quantify asymmetric dependence structure between high and low values.

Empirical investigations using the high-quality precipitation datasets prove that most of the precipitation events (case study in Singapore, Baden-Württemberg and Bavaria) exhibit

positive asymmetric spatial dependence at close distances between two points at different locations. This implies that precipitation with higher amounts tends to be more spatially correlated than lower values between nearby points. Consequently, spatial precipitation models based on the symmetric Gaussian dependence could yield a biased estimation. A V-transformed normal copula provides a possible solution to model the asymmetric dependence structure reasonably with high multidimensional problems.

Numerous spatial precipitation models incorporating spatial variability on the basis of grid cells have been extensively developed for input to distributed hydrological modelling, with most relying on an underlying symmetric Gaussian dependence. Studies of the effect of the symmetric Gaussian dependence based model were conducted in the regions of Singapore and Bavaria. Spatial precipitation models based on symmetric spatial dependence are represented by Gaussian copulas, and the asymmetrical spatial dependence based-models are represented by V-copulas. Both models are applied using daily precipitation on the basis of grid cells.

Investigations based on simulation examples focusing on the spatial extremes of areal gridded precipitation amounts reveal that the Gaussian copulas frequently exhibit lower areal mean gridded precipitation than V-copulas. Precipitation occurrences, which typically behave in a clustering manner, could not be modelled reasonably by Gaussian copulas, opposite to as expected. This is because the extreme spatial precipitation amounts estimated by the Gaussian copulas are systematically lower than V-copulas. As a result, Gaussian copulas would yield an underestimation of flood risks and should therefore be implemented with care in the wider practice of flood designs. In the financial sector, Salmon (2012) has mentioned that the Gaussian copula is the spectacular recipe for disaster which killed Wall Street.

This study focused on the possible effects of asymmetry on extreme precipitation. The consequences including area reduction factors and the estimation of extremes should be subject of future research.

# Bibliography


Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American statistical Association*, 97(457):284–292.

Anagnostou, E. N. (2004). A convective/stratiform precipitation classification algorithm for volume scanning weather radar observations. *Meteorological Applications*, 11(4):291–300.

Anderson, T. W. (1962). On the distribution of the two-sample cramer-von mises criterion. *The Annals of Mathematical Statistics*, pages 1148–1159.

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, 27(1):17–21.

Arnold, T. B. and Emerson, J. W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *The R Journal*, 3(2):34–39.

Ban, N., Schmidli, J., and Schär, C. (2014). Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres*, 119(13):7889–7907.

Bardossy, A. (1998). Generating precipitation time series using simulated annealing. *Water Resources Research*, 34(7):1737–1744.

Bárdossy, A. (2006). Copula-based geostatistical models for groundwater quality parameters. *Water Resources Research*, 42(11):1–12.

Bárdossy, A. (2010). Atmospheric circulation pattern classification for south-west germany using hydrological variables. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9-12):498–506.

Bárdossy, A. (2011). Interpolation of groundwater quality parameters with some values below the detection limit. *Hydrology and Earth System Sciences*, 15(9):2763–2775.

Bárdossy, A., DUCKSTEIN, L., and BOGARDI, I. (1995). Fuzzy rule-based classification of atmospheric circulation patterns. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, 15:1087–1097.

Bárdossy, A. and Li, J. (2008). Geostatistical interpolation using copulas. *Water Resources Research*, 44(7):1–15.

Bárdossy, A. and Pegram, G. (2009). Copula based multisite model for daily precipitation simulation. *Hydrology and Earth System Sciences*, 13:2299–2314.

Bárdossy, A. and Pegram, G. (2013). Interpolation of precipitation under topographic influence at different time scales. *Water Resources Research*, 49(8):4545–4565.

Bárdossy, A. and Pegram, G. (2014). Infilling missing precipitation records – a comparison of a new copula-based method with other techniques. *Journal of Hydrology*, 519:1162–1170.

Bárdossy, A., Pegram, G., Sinclair, S., Pringle, J., and Stretch, D. (2015). Circulation patterns identified by spatial rainfall and ocean wave fields in southern africa. *Frontiers in Environmental Science*, 3.

Bárdossy, A. and Plate, E. J. (1992). Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research*, 28(5).

Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *Journal of multivariate analysis*, 88(1):190–206.

Beck, F. and Bárdossy, A. (2013). Indirect downscaling of hourly precipitation based on atmospheric circulation and temperature. *Hydrology and Earth System Sciences*, 17(12):4851–4863.

Beck, F., Bárdossy, A., Seidel, J., Müller, T., Fernandez Sanchis, E., and Hauser, A. (2015). Statistical analysis of sub-daily precipitation extremes in singapore. *Journal of Hydrology: Regional Studies*, 3:337–358.

Bennett, B., Lambert, M., Thyer, M., Bates, B. C., and Leonard, M. (2015). Estimating extreme spatial rainfall intensities. *Journal of Hydrologic Engineering*, 21(3):04015074.

Cameron, A. C., Li, T., Trivedi, P. K., and Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, 7(2):566–584.

Casas, M. C., Codina, B., Redaño, A., and Lorente, J. (2004). A methodology to classify extreme rainfall events in the western mediterranean area. *Theoretical and Applied Climatology*, 77(3):139–150.

Chatterjea, K. (1998). The impact of tropical rainstorms on sediment and runoff generation from bare and grass-covered surfaces: a plot study from singapore. *Land Degradation & Development*, 9(2):143–157.

Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied Hydrology*. McGrawHill, New York.

Clarke, R. T. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, 19(1):1–20.

Dai, Q., Han, D., Rico-Ramirez, M. A., and Islam, T. (2014). Modelling radar-rainfall estimation uncertainties using elliptical and archimedean copulas with different marginal distributions. *Hydrological Sciences Journal*, 59(11):1992–2008.

Davison, A. C., Padoan, S., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science*, pages 161–186.

Dooge, J. C. (1968). The hydrologic cycle as a closed system. *International Association of Scientific Hydrology. Bulletin*, 13(1):58–68.

Drouet-Mari, D. and S. Kotz (2001). *Correlation and Dependence*. Imperial College Press, London.

Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., and Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40(1):1–12.

Fong, M. and Ng, L. K. (2012). *The weather and climate of Singapore*. Meteorological Service Singapore.

Franz, C. (2006). cramer: Multivariate nonparametric cramer-test for the two-sample-problem. *R package version 0.8-1*.

Frei, C. and Schär, C. (1998). A precipitation climatology of the alps from high-resolution rain-gauge observations. *International Journal of Climatology*, 18(8):873–900.

Gautheir, T. D. (2001). Detecting trends using spearman's rank correlation coefficient. *Environmental forensics*, 2(4):359–362.

Gräler, B. (2014). Modelling skewed spatial random fields through the spatial vine copula. *Spatial Statistics*, 10:87–102.

Guthke, P. (2013). Non-multi-gaussian spatial structures:process-driven natural genesis,manifestation, modeling approaches,and influences on dependent processes. *Institut für Wasserbau-Mitteilungen*.

Gyasi-Agyei, Y. (2016). Assessment of radar-based locally varying anisotropy on daily rainfall interpolation. *Hydrological Sciences Journal*, pages 1–13.

Gyasi-Agyei, Y. and Pegram, G. (2014). Interpolation of daily rainfall networks using simulated radar fields for realistic hydrological modelling of spatial rain field ensembles. *Journal of Hydrology*, 519:777–791.

Ha, E. and Yoo, C. (2007). Use of mixed bivariate distributions for deriving inter-station correlation coefficients of rain rate. *Hydrological Processes*, 21(22):3078–3086.

Habib, E. and Krajewski, W. F. (2001). Estimation of rainfall interstation correlation. *Journal of Hydrometeorolgy*, 2(6):621–629.

Haddad, K. and Rahman, A. (2011). Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for tasmania in australia. *Stochastic Environmental Research and Risk Assessment*, 25(3):415–428.

Hesse, R. (2010). Lidar-derived local relief models–a new tool for archaeological prospection. *Archaeological prospection*, 17(2):67–72.

Hofert, M., Kojadinovic, I., Maechler, M., and Yan, J. (2017). *copula: Multivariate Dependence with Copulas*. R package version 0.999-16.

Houze, R. (1997). Stratiform precipitation in regions of convection: A meteorological paradox? *Bulletin of the American Meteorological Society*, 78(10):2179–2196.

Houze, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, 42(4).

Husak, G. J., Michaelsen, J., and Funk, C. (2007). Use of the gamma distribution to represent monthly rainfall in africa for drought monitoring applications. *International Journal of Climatology*, 27(7):935–944.

Ison, N., Feyerherm, A., and Bark, L. D. (1971). Wet period precipitation and the gamma distribution. *Journal of Applied Meteorology*, 10(4):658–665.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. London: Chapman&Hall.

Kazianka, H. and Pilz, J. (2011). Bayesian spatial modeling and interpolation using copulas. *Computers & Geosciences*, 37(3):310–319.

Kilsby, C., Jones, P., Burton, A., Ford, A., Fowler, H., Harpham, C., James, P., Smith, A., and Wilby, R. (2007). A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, 22(12):1705–1719.

Kleiber, W., Katz, R. W., and Rajagopalan, B. (2012). Daily spatiotemporal precipitation simulation using latent and transformed gaussian processes. *Water Resources Research*, 48(1).

Kousky, C. and Walls, M. (2014). Floodplain conservation as a flood mitigation strategy: Examining costs and benefits. *Ecological Economics*, 104:119–128.

Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *Applied Statistics*, pages 1–12.

Laio, F. (2004). Cramer–von mises and anderson-darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*, 40(9).

Lee Rodgers, J. and Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1):59–66.

Li, J. (2010). Application of copulas as a newgeostatistical tool. *Institut für Wasserbau-Mitteilungen*.

Liu, M., Bárdossy, A., and Zehe, E. (2013). Interaction of valleys and circulation patterns (cps) on spatial precipitation patterns in southern germany. *Hydrology and Earth System Sciences*, 17(11):4685–4699.

Llasat, M.-C. (2001). An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of spain. *International Journal of Climatology*, 21(11):1385–1400.

Maity, A. and Sherman, M. (2012). Testing for spatial isotropy under general designs. *Journal of statistical planning and inference*, 142(5):1081–1091.

Mandapaka, P. V. and Qin, X. (2013). Analysis and characterization of probability distribution and small-scale spatial variability of rainfall in singapore using a dense gauge network*. *Journal of Applied Meteorology and Climatology*, 52(12):2781–2796.

Mao, G., Vogl, S., Laux, P., Wagner, S., and Kunstmann, H. (2015). Stochastic bias correction of dynamically downscaled precipitation fields for germany through copula-based integration of gridded observation data. *Hydrology and Earth System Sciences*, 19(4):1787–1806.

Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.

Matheron, G. (1971). *The theory of regionalized variables and its applications*, volume 5. École national supérieure des mines.

Mosthaf, T. and Bárdossy, A. (2017). Regionalizing nonparametric models of precipitation amounts on different temporal scales.

Nelsen, R. (2006). *An Introduction to Copulas*. New York: Springer.

Papalexiou, S., Koutsoyiannis, D., and Makropoulos, C. (2013). How extreme is extreme? an assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 17(2):851–862.

Pettitt, A. N. and Stephens, M. A. (1976). Modified cramer-von mises statistics for censored data. *Biometrika*, 63(2):291–298.

Pinto, J. G., Ulbrich, S., Parodi, A., Rudari, R., Boni, G., and Ulbrich, U. (2013). Identification and ranking of extraordinary rainfall events over northwest italy: The role of atlantic moisture. *Journal of Geophysical Research: Atmospheres*, 118(5):2085–2097.

Quessy, J.-F., Rivest, L.-P., and Toupin, M.-H. (2016). On the family of multivariate chi-square copulas. *Journal of Multivariate Analysis*, 152:40–60.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Renard, B. and Lang, M. (2007). Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. *Advances in Water Resources*, 30(4):897–912.

Salmon, F. (2012). The formula that killed wall street. *Significance*, 9(1):16–20.

Scholz, F., Zhu, A., et al. (2012). ksamples: K-sample rank tests and their combinations. *R package version*, 1.

Scholz, F. W. and Stephens, M. A. (1987). K-sample andersondarling tests. *Journal of the American Statistical Association*, 82(399):918–924.

Sekhon, J. S. (2011). Multivariate and propensity score matching software with automated balance optimization: the matching package for r.

Sen, Z. (2009). *Spatial modeling principles in earth sciences*. Springer.

Serinaldi, F. (2008). Analysis of inter-gauge dependence by kendall, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stochastic Environmental Research and Risk Assessment*, 22(6):671–688.

Serinaldi, F. (2009). A multisite daily rainfall generator driven by bivariate copula-based mixed distributions. *Journal of Geophysical Research*, 114(D10).

Serinaldi, F., Bárdossy, A., and Kilsby, C. G. (2014). Upper tail dependence in rainfall extremes: would we know it if we saw it? *Stochastic Environmental Research and Risk Assessment*, 29(4):1211–1233.

Serinaldi, F. and Kilsby, C. G. (2014). Simulating daily rainfall fields over large areas for collective risk estimation. *Journal of Hydrology*, 512:285–302.

Sklar, A. (1959). Fonctions de rpartition n dimensions et leurs marges. *Publications de lInstitut de statistique de l'Universit de Paris*.

Soláková, T., De Michele, C., and Vezzoli, R. (2013). Comparison between parametric and nonparametric approaches for the calculation of two drought indices: Spi and ssi. *Journal of Hydrologic Engineering*, 19(9):04014010.

Spearman, C. (1904). The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101.

Strömberg, D. (2007). Natural disasters, economic development, and humanitarian aid. *The Journal of Economic Perspectives*, 21(3):199–222.

Vernieuwe, H., Vandenberghe, S., De Baets, B., and Verhoest, N. (2015). A continuous rainfall model based on vine copulas. *Hydrology and Earth System Sciences*, 19(6):2685–2699.

Viglione, A., Laio, F., and Claps, P. (2007). A comparison of homogeneity tests for regional frequency analysis. *Water Resources Research*, 43(3).

Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H. (2012). Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. *Hydrology and Earth System Sciences*, 16(7):2311–2328.

Vrac, M., Naveau, P., and Drobinski, P. (2007). Modeling pairwise dependencies in precipitation intensities. *Nonlinear Processes in Geophysics*, 14(6):789–797.

Warrach-Sagi, K., Schwitalla, T., Wulfmeyer, V., and Bauer, H.-S. (2013). Evaluation of a climate simulation in europe based on the wrf–noah model system: Precipitation in germany. *Climate Dynamics*, 41(3-4):755–774.

Wasko, C., Sharma, A., and Rasmussen, P. (2013). Improved spatial prediction: A combinatorial approach. *Water Resources Research*, 49(7):3927–3935.

Wilcox, R. (2005). Kolmogorov–smirnov test. *Encyclopedia of biostatistics*.

Wilks, D. S. (1989). Rainfall intensity, the weibull distribution, and estimation of daily surface runoff. *Journal of Applied Meteorology*, 28(1):52–58.

Wilks, D. S. (1990). Maximum likelihood estimation for the gamma distribution using data containing zeros. *Journal of Climate*, 3(12):1495–1501.

Wilks, D. S. (1998). Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology*, 210(1-4):178–191.

Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G. C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C. N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G., Richard, E., Rotach, M. W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D. D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A., and Wirth, M. (2011). The convective and orographically-induced precipitation study (cops): The scientific strategy, the field phase, and research highlights. *Quarterly Journal of the Royal Meteorological Society*, 137(S1):3–30.

# Curriculum Vitae

Personal

| | |
|---|---|
| Name: | Suroso |
| Date of Birth: | 01.12.1979 |
| City of Birth: | Wonogiri, Indonesia |
| Nationality: | Indonesian |

Education

| | |
|---|---|
| 1998 - 2002: | Bachelor in Civil Engineering, Gadjah Mada University, Indonesia |
| 2010 - 2012: | M.Sc in Water Resources Engineering and Management (WAREM), University of Stuttgart, Germany |
| 2013 - 2017: | Dr.-Ing, International Doctoral Program Environment Water (ENWAT), University of Stuttgart, Germany |

Work Experiences

| | |
|---|---|
| 2000 - 2002 | Teaching assistant at the Department of Civil Engineering, Gadjah Mada University, Indonesia |
| 2003 - present | Lecturer/Researcher at the Department of Civil Engineering, Jenderal Soedirman University, Indonesia |
| 2013 - 2017 | Teaching assistant at the Department of Hydrology and Geohydrology, Institute of Hydraulic Engineering, University of Stuttgart, Germany. |
| Nov 2016 - Jan 2017 | Internship work (Praktikum) at RSS-Remote Sensing Solutions GmbH, Munich, Germany. |

Awards

1. DAAD-BMBF Doctoral Scholarship Award from the Ministry of Education and Research, the German Government (2013-2017)
2. IPSWAT-BMBF Master Scholarship Award from the Ministry of Education and Research, the German Government (2010-2012)
3. Achievement Graduation with Cumlaude predicate from Faculty of Engineering, Gadjah Mada University, November 2002.

**Institut für Wasser- und Umweltsystemmodellierung**
**Universität Stuttgart**

Pfaffenwaldring 61
70569 Stuttgart (Vaihingen)
Telefon (0711) 685 - 64717/64749/64752/64679
Telefax (0711) 685 - 67020 o. 64746 o. 64681
E-Mail: iws@iws.uni-stuttgart.de
http://www.iws.uni-stuttgart.de



## Verzeichnis der Mitteilungshefte

5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht*, und
Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gestehungskosten und jeweilige Vorteile*, sowie
Unny, T.E.: *Schwingungsuntersuchungen am Kegelstrahlschieber*, 1967

6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschiffahrt an diesem Wert*, 1967

7 *Sonderheft anläßlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von* Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967

8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968

9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968

10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968

11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969

12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969

13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969

14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970

15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970

16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit*, 1970

17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970

18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, *vergriffen*

19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, *vergriffen*

20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971

21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972

22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972

23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972

24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, *vergriffen*

25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, *vergriffen*

26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972

28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschiffahrt durch Schwimmerhebewerke*, 1973

29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973

30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973

31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973

32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974

34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975

35 *Sonderheft anläßlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von*: Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975

36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975

37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976

38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976

39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6

40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X

41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, *vergriffen*

42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6

43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4

44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2

45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0

46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9

47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7

48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5

49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, *vergriffen*

50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, *vergriffen*

51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5

52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, *vergriffen*

53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, *vergriffen*

54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X

55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, *vergriffen*

56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6

57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, *vergriffen*

58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2

59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, *vergriffen*

60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, *vergriffen*

61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebenen Grundwassermodellen*, 1986, ISBN 3-921694-62-0

63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9

64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7

65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5

66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3

67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1

68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, *vergriffen*

69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, *vergriffen*

70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1

71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*,1991, ISBN 3-921694-72-8

73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6

74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, *vergriffen*

75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2

76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, *vergriffen*

78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7

79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, *vergriffen*

80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, *vergriffen*

81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern,* 1993, ISBN 3-921694-81-7

82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versuchseinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5

83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, *vergriffen*

84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1

85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, *vergriffen*

86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, *vergriffen*

87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation -*, 1996, ISBN 3-921694-87-6

88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4

89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, *vergriffen*

90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, *vergriffen*

91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, *vergriffen*

92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien,* 1997, ISBN 3-921694-92-2

93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, *vergriffen*

94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen 1997*, ISBN 3-921694-94-9

95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, *vergriffen*

96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5

97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, SBN 3-921694-97-3

98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, *vergriffen*

99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden,* 2000, ISBN 3-933761-02-2

100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4

101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, *vergriffen*

102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0

103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9

104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7

105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien,* 2001, ISBN 3-933761-08-5

106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3

107 Josef, Reinhold:, *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5

109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3

110 Lang, Stefan: *Parallele numerische Simulation instätionärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1

111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X

112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten,* 2002, ISBN 3-933761-15-8, *vergriffen*

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002, ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Up-scaling, Mass-Conservative Discretization and Efficient Multilevel Solution,* 2004, ISBN 3-933761-33-6

131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004,* Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4

132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2

133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers,* 2004, ISBN 3-933761-36-0

134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface,* 2005, ISBN 3-933761-37-9

135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7

136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers,* 2005, ISBN: 3-933761-39-5

137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments,* 2005, ISBN 3-933761-40-9

138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7

139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5

140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media,* 2005, ISBN 3-933761-43-3

141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen,* 2005, ISBN 3-933761-44-1

142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X

143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser,* 2005, ISBN 3-933761-46-8

144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, ISBN 3-933761-47-6

145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden,* Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4

146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes,* 2006, ISBN 3-933761-49-2

147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebeckensystemen,* 2006, ISBN 3-933761-50-6

148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4

149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2

150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS-Statuskolloquium 2006,* Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0

151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9

152 Fischer, Markus: *Beanspruchung eingeerdeter Rohrleitungen infolge Austrocknung bindiger Böden,* 2006, ISBN 3-933761-55-7

153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5

154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models,* 2006, ISBN 3-33761-57-3

155 Bielinski, Andreas: *Numerical Simulation of $CO_2$ sequestration in geological formations*, 2007, ISBN 3-933761-58-1

156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3

157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation,* 2007, ISBN 3-933761-61-1

158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines,* 2007, ISBN 3-933761-62-X

159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8

160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, 2007,* ISBN 3-933761-64-6

161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4

162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2

163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0

164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9

165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS – Kolloquium 2007,* Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7

166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9

168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7

169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling,* 2008*,* ISBN 3-933761-73-5

170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3

171 Papafotiou, *Alexandros: Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1

172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7

173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4

174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1

175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8

176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach,* 2008, ISBN 978-3-933761-80-4

177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1

178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen,* 2009, ISBN 978-3-933761-82-8

179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa,* 2009, ISBN 978-3-933761-83-5

180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding,* 2009, ISBN 978-3-933761-84-2

181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers,* 2009, ISBN 978-3-933761-85-9

182 Kopp, Andreas: *Evaluation of $CO_2$ Injection Processes in Geological Formations for Site Screening,* 2009, ISBN 978-3-933761-86-6

183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on $CO_2$ and water (two-phase) flow in porous media,* 2009, ISBN 978-3-933761-87-3

184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0

185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7

186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3

187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0

188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7

189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4

190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8

192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5

193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2

194 Bliefernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins,* 2011, ISBN 978-3-933761-98-9

195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -,* Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen*,* 2010*,* ISBN 978-3-933761-99-6

196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia,* 2010, ISBN 978-3-942036-00-9

197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen,* 2010, ISBN 978-3-942036-01-6

198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins,* 2010, ISBN 978-3-942036-02-3

199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow,* 2011, ISBN 978-3-942036-03-0

200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation,* 2011, ISBN 978-3-942036-04-7

201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas,* 2011, ISBN 978-3-942036-05-4

202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen,* 2011, ISBN 978-3-942036-06-1

203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks,* 2011, ISBN 978-3-942036-07-8

204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology ,* 2011, ISBN 978-3-942036-08-5

205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken,* Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen*,* 2011*,* ISBN 978-3-933761-09-2

206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour ,* 2011, ISBN 978-3-942036-10-8

207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta,* 2011, ISBN 978-3-942036-11-5

208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam,* 2011, ISBN 978-3-942036-12-2

209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone,* 2012, ISBN 978-3-942036-13-9

210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking,* 2012, ISBN 978-3-942036-14-6

211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach,* 2012, ISBN 978-3-942036-15-3

212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media,* 2012, ISBN 978-3-942036-16-0

213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung,* 2012, ISBN 978-3-942036-17-7

214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish,* 2012, ISBN 978-3-942036-18-4

215 De Boer, Cjestmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1

216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7

217 Lebrenz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4

218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of $CO_2$ Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1

219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8

220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5

221 Walter, Lena: *Uncertainty studies and risk assessment for $CO_2$ storage in geological formations*, 2013, ISBN 978-3-942036-25-2

222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects,* 2013, ISBN 978-3-942036-26-9

223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes,* 2014, ISBN 978-3-942036-27-6

224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3

225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0

226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6

227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3

228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0

229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7

230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4

231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1

232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8

233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5

234 Zhang, Jin: *Investigations on Urban River Regulation and Ecological Rehabilitation Measures, Case of Shenzhen in China*, 2014, ISBN 978-3-942036-38-2

235 Siebel, Rüdiger: *Experimentelle Untersuchungen zur hydrodynamischen Belastung und Standsicherheit von Deckwerken an überströmbaren Erddämmen,* 2014, ISBN 978-3-942036-39-9

236 Baber, Katherina: *Coupling free flow and flow in porous media in biological and technical applications: From a simple to a complex interface description,* 2014, ISBN 978-3-942036-40-5

237 Nuske, Klaus Philipp: *Beyond Local Equilibrium — Relaxing local equilibrium assumptions in multiphase flow in porous media,* 2014, ISBN 978-3-942036-41-2

238 Geiges, Andreas: *Efficient concepts for optimal experimental design in nonlinear environmental systems*, 2014, ISBN 978-3-942036-42-9

239 Schwenck, Nicolas: *An XFEM-Based Model for Fluid Flow in Fractured Porous Media*, 2014, ISBN 978-3-942036-43-6

240 Chamorro Chávez, Alejandro: *Stochastic and hydrological modelling for climate change prediction in the Lima region, Peru*, 2015, ISBN 978-3-942036-44-3

241 Yulizar: *Investigation of Changes in Hydro-Meteorological Time Series Using a Depth-Based Approach*, 2015, ISBN 978-3-942036-45-0

242 Kretschmer, Nicole: *Impacts of the existing water allocation scheme on the Limarí watershed – Chile, an integrative approach*, 2015, ISBN 978-3-942036-46-7

243 Kramer, Matthias: *Luftbedarf von Freistrahlturbinen im Gegendruckbetrieb*, 2015, ISBN 978-3-942036-47-4

244 Hommel, Johannes: *Modeling biogeochemical and mass transport processes in the subsurface: Investigation of microbially induced calcite precipitation*, 2016, ISBN 978-3-942036-48-1

245 Germer, Kai: *Wasserinfiltration in die ungesättigte Zone eines makroporösen Hanges und deren Einfluss auf die Hangstabilität*, 2016, ISBN 978-3-942036-49-8

246 Hörning, Sebastian: *Process-oriented modeling of spatial random fields using copulas*, 2016, ISBN 978-3-942036-50-4

247 Jambhekar, Vishal: *Numerical modeling and analysis of evaporative salinization in a coupled free-flow porous-media system,* 2016, ISBN 978-3-942036-51-1

248 Huang, Yingchun: *Study on the spatial and temporal transferability of conceptual hydrological models*, 2016, ISBN 978-3-942036-52-8

249 Kleinknecht, Simon Matthias: *Migration and retention of a heavy NAPL vapor and remediation of the unsaturated zone*, 2016, ISBN 978-3-942036-53-5

250 Kwakye, Stephen Oppong: *Study on the effects of climate change on the hydrology of the West African sub-region*, 2016, ISBN 978-3-942036-54-2

251 Kissinger, Alexander: *Basin-Scale Site Screening and Investigation of Possible Impacts of $CO_2$ Storage on Subsurface Hydrosystems*, 2016, ISBN 978-3-942036-55-9

252 Michael Sinsbeck: *Uncertainty Quantification for Expensive Simulations - Optimal Surrogate Modeling under Time Constraints*, 2017, ISBN 978-3-942036-56-6

253 Grüninger, Christoph: *Numerical Coupling of Navier-Stokes and Darcy Flow for Soil-Water Evaporation*, 2017, ISBN 978-3-942036-57-3

254 Suroso: Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields, 2017, ISBN 978-3-942036-58-0

*Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.*