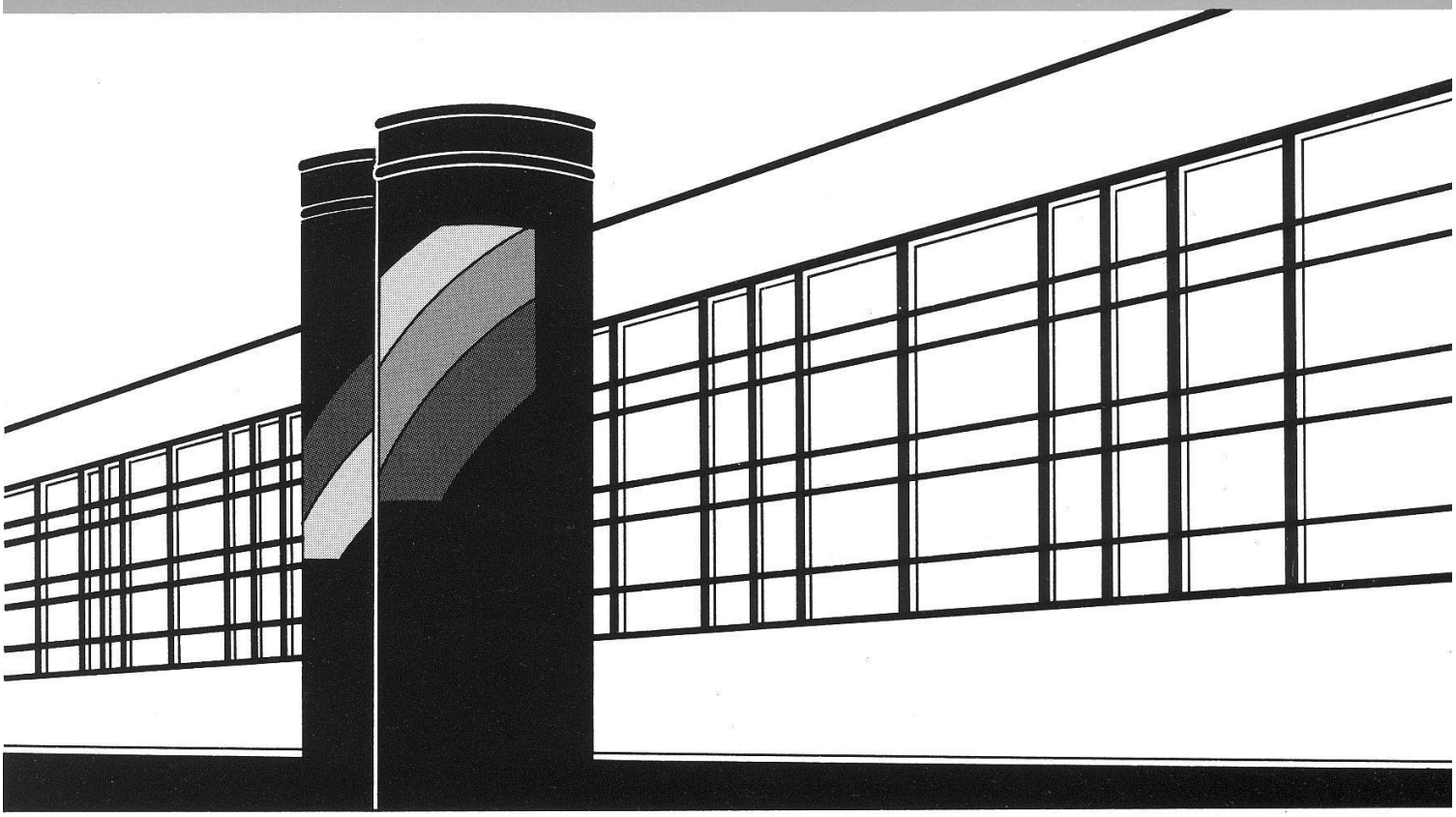
Institut für Wasserbau · Universität Stuttgart

# Mitteilungen



Heft 194 Jan Bliefernicht

## Probability Forecasts of Daily Areal Precipitation for Small River Basins

# Probability Forecasts of Daily Areal Precipitation for Small River Basins

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

Vorgelegt von

**Jan Bliefernicht**

aus Wesenstedt

| | |
|---|---|
| Hauptberichter: | Prof. András Bárdossy |
| Mitberichter: | Prof. Charles Obled |

Tag der mündlichen Prüfung: 22. Juni 2010

Institut für Wasserbau der Universität Stuttgart

2010

Heft 194

# Probability Forecasts of Daily Areal Precipitation for Small River Basins

von
Dr.-Ing.
Jan Bliefernicht

Eigenverlag des Instituts für Wasserbau der Universität Stuttgart

**D93 Probability Forecasts of Daily Areal Precipitation for Small River Basins**

**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

Bliefernicht, Jan:
Probability Forecasts of Daily Areal Precipitation for Small River Basins
von Jan Bliefernicht. Institut für Wasserbau, Universität Stuttgart.
Stuttgart: Inst. für Wasserbau, 2010

(Mitteilungen / Institut für Wasserbau, Universität Stuttgart: Heft 194)
Zugl.: Stuttgart, Univ., Diss., 2010
ISBN 978-3-933761-98-9
NE: Institut für Wasserbau <Stuttgart>: Mitteilungen



Herausgegeben 2010 vom Eigenverlag des Instituts für Wasserbau
Druck: Document Center S. Kästl, Ostfildern

# Acknowledgements

First of all, I want to thank Prof. Bárdossy who has given me the chance to work in his group, for his continuous support and for his patience with me. He also supported me with many ideas so that I have still a plenty of work to do in the future. Many thanks also to Prof. Obled for being my external reviewer, his detailed revision, for his help and for the fruitful discussions during his stay in Stuttgart.

The work was counter-checked by my colleagues, Ferdinand, Christian, Mahboob, Thomas P., Thomas J., Dirk, Shailesh and Claus. They gave me many valuable comments that improved the work a lot. The English spelling and grammar was proof-read by Leo Redcliffe in a very professional way.

Many further research colleagues and friends helped me directly or indirectly during the preparation of my work. Here, I would like to give a special mention to the following persons: My colleague Christian - we had more than five years a very good time in the same office; my current and former colleagues for the excellent working atmosphere in our group and the enjoyable coffee breaks; Norbert Demuth from the State Agency Rhineland-Palatinate for his patience during the project work so that I had enough time to complete my dissertation; my project colleagues in HORIX and in PREVIEW for the very good collaboration and for providing me with all necessary information; my roommates, Alice, Dagmar, Ralf and Nowo, for the perfect refuge when I had a bad working day.

Last but not least, I would like to thank my parents, Margret and Karl-Heinz, and the most important person in my life, my partner Rebekka. Due to the continuous support and the tolerance of my parents I was in the lucky position to study environmental science (geoecology) and to receive a PhD in hydrology. The help, the understanding and the ease of mind of Rebekka were extremely important during hard times. I am looking forward to our common future.

# Contents

**Acknowledgements** i

**List of Figures** vii

**List of Tables** viii

**List of Abbreviations** xi

**List of Symbols** xiii

**Zusammenfassung** xv

**1 Introduction** 1
- 1.1 Motivation . . . 1
- 1.2 Objective and Research Questions . . . 5
- 1.3 Outline of the Thesis . . . 6

**2 Basic Theory** 9
- 2.1 Weather Forecasting . . . 9
  - 2.1.1 Global Numerical Weather Prediction . . . 10
  - 2.1.2 Downscaling . . . 11
  - 2.1.3 Dynamical Downscaling and Deterministic Forecasting . . . 13
  - 2.1.4 Dynamical Downscaling and Ensemble Forecasting . . . 15
  - 2.1.5 Statistical Downscaling and Forecasting . . . 19
- 2.2 Forecast Verification . . . 23
  - 2.2.1 What is a Good Forecast? . . . 23
  - 2.2.2 Skill Scores . . . 25
  - 2.2.3 Probability Forecast . . . 26
  - 2.2.4 Binary Forecast . . . 29
  - 2.2.5 Cost-Loss Situation and Forecast Value . . . 31

**3 Data and Study Areas** 35
- 3.1 Description of Study Regions . . . 37
- 3.2 Description of Predictor Information . . . 44

**4 Analog Method** 49
- 4.1 Principle of the Analog Method . . . 49
- 4.2 Pattern Closeness and Pattern Similarity . . . 52
- 4.3 Maximizing the Forecast Value . . . 55
- 4.4 Probabilistic Precipitation Forecast . . . 58
- 4.5 Strategy of the Model Development . . . 60

4.6 Influence of User-Defined Criteria . . . 62
4.6.1 Predictor Selection and Predictor Configuration . . . 62
4.6.2 Predictor Domain . . . 65
4.6.3 Distance Parameters and Predictor Weights . . . 67
4.6.4 Number of Analogs and Probability Functions . . . 72
4.6.5 Selection Rule . . . 73
4.7 Forecast Skill and Value . . . 74
4.7.1 Optimal Use of a Probability Forecast . . . 74
4.7.2 Intra-annual Variability . . . 79
4.8 Summary and Conclusions . . . 80

**5 Classification** **83**
5.1 Subjective and Objective Classification . . . 83
5.2 Why Objective Classification in Weather Forecasting? . . . 84
5.3 Fuzzy Rule-Based Classification . . . 86
5.3.1 Basic Methodology . . . 86
5.3.2 Objective Functions . . . 87
5.3.3 Validation Procedure . . . 91
5.4 Wetness Index and Anomaly Maps . . . 92
5.5 Forecast Skill and Value . . . 96
5.5.1 Pure Classification . . . 96
5.5.2 Classification vs. Analog Forecasting . . . 97
5.5.3 Conditional Forecast Value . . . 100
5.6 Summary and Conclusions . . . 101

**6 Metric Optimization** **103**
6.1 Methodology . . . 104
6.1.1 Weighted Euclidean Distance . . . 104
6.1.2 Simulated Annealing . . . 105
6.1.3 Objective Function . . . 106
6.2 Model Development and Validation Strategy . . . 108
6.3 Results . . . 109
6.3.1 Air Flow Indices and Data Resolution . . . 109
6.3.2 Optimization Performance and Metric Weights . . . 111
6.3.3 Forecast Accuracy and Value . . . 114
6.3.4 Selection Rule . . . 118
6.4 Summary and Conclusions . . . 120

**7 Data Depth** **123**
7.1 How Can We Measure the Centrality of a Weather State? . . . 125
7.2 Applications . . . 127
7.3 Precipitation Probability and Risk Index . . . 128
7.4 Forecast Value . . . 133
7.5 Summary and Conclusions . . . 134

**8 Operational Application** **137**
8.1 Model History . . . 137
8.2 Global Forecast System . . . 138

8.3 Probabilistic Precipitation Forecast System . . . . . 138
8.3.1 Data Assimilation System and Postprocessor . . . . . 138
8.3.2 Statistical Precipitation Prediction Model . . . . . 141
8.4 Probabilistic Quantitative Precipitation Forecast . . . . . 143
8.5 Summary . . . . . 145

**9 Summary and Conclusion** **147**
9.1 Some Answers to the Research Questions . . . . . 148
9.2 Outlook for Future Investigations . . . . . 152

# List of Figures


1 Lage der Untersuchungsgebiete . . . . . xviii
2 Vorhersagefehler des COSMO-Modells . . . . . xx
3 Streudiagramm täglicher Geopotentialanomalien . . . . . xxii
4 Musterabstand und Musterähnlichkeit . . . . . xxiii
5 Wirkungsbereich der Geopotentialhöhe . . . . . xxiv

2.1 Illustration of the real world in comparison to the world of a general circulation model . . . . . 12
2.2 Differences between the precipitation forecast of the COSMO-model and observations . . . . . 14
2.3 Forecast uncertainty of the COSMO-LEPS . . . . . 18

3.1 Location and catchment size of the study areas . . . . . 36
3.2 Precipitation regimes for the four study regions . . . . . 39
3.3 Discharge regimes for the four study regions . . . . . 40
3.4 Seasonal regime of heavy precipitation and floods for the four study regions . . . . . 43
3.5 Missing value of the reanalysis information. . . . . 47

4.1 Pattern closeness and pattern similarity . . . . . 53
4.2 Variation of center, size and form of a predictor domain . . . . . 60
4.3 Influence of the predictor configuration on the model performance for the height field. . . . . 63
4.4 Influence of the predictor configuration on the model performance for the moisture flux. . . . . 64
4.5 Influence of the domain center on the model performance . . . . . 65
4.6 Influence of the domain form and size on the model performance . . . 67
4.7 Influence of the distance parameters on the model performance . . . . 69
4.8 Influence of the predictor combinations on the model performance . . 71
4.9 Influence of the number of most similar analogs on the model performance. . . . . 73
4.10 Influence of the selection rule on the model performance. . . . . 73
4.11 Mean forecast value of the probability forecast in dependency on the cost-loss ratio . . . . . 75
4.12 Maximum forecast value and user interval in dependence on the return frequency . . . . . 76
4.13 Intra-annual variability of analog forecasting . . . . . 80

5.1 Histogram of observed daily areal precipitation and observed daily discharge increments . . . . . 89
5.2 Time series of observed daily discharge differences . . . . . 90

5.3 Anomaly maps of daily mean sea level pressure . . . . . 94
5.4 Forecast value of the classifications based on daily precipitation in comparison to the classification with daily discharge differences . . . . 96
5.5 Forecast value of the analog method versus classification . . . . . 98
5.6 Forecast value of analog forecasting for wet and dry CPs. . . . . 101

6.1 Spatial distribution of metric weights in analog forecasting . . . . . 103
6.2 Model performance of the analog method for different pressure and moisture related variables . . . . . 110
6.3 Optimization performance for three parameterizations schemes . . . . 111
6.4 Calculated metric coefficients for four optimization strategies . . . . . 112
6.5 Contour map of metric coefficients for four optimization strategies . . 113
6.6 Forecast accuracy of analog forecasting . . . . . 115
6.7 Model performance of analog forecasting in comparison to a reference algorithm . . . . . 116
6.8 Relative value of analog forecasting . . . . . 117
6.9 Selection rules for analog forecasting . . . . . 118
6.10 Influence of the selection rule on the model performance . . . . . 119

7.1 Scatter plot of the daily height anomalies . . . . . 124
7.2 Functionality of a depth function . . . . . 126
7.3 Data depth of the height anomaly in relation to intensive daily areal precipitation . . . . . 128
7.4 Precipitation probability conditioned on data depth . . . . . 129
7.5 Scatter plot of daily height anomalies for wet and dry CPs . . . . . 130
7.6 Conditional precipitation probability and conditional risk index . . . 131
7.7 Forecast value of the conditional downscaling approach based on the data depth . . . . . 134

8.1 Data flow through the assimilation system . . . . . 139
8.2 Data flow through the postprocessor . . . . . 140
8.3 Forecast scheme of the resampling algorithm . . . . . 142
8.4 An example of a probabilistic quantitative precipitation forecast . . . 143
8.5 Probabilistic quantitative precipitation forecast for an extrem event . 146

9.1 Forecast uncertainty of analog forecasting with respect to the closeness of two resampling algorithms . . . . . 153

# List of Tables


1 Resultate einer binären Vorhersage . . . . . xx
2 Konsequenzen der Entscheidungsfindung . . . . . xxi

2.1 Model properties of several NWP models and the distributed hydrological model LARSIM. . . . . 19
2.2 Possible outcomes of a binary forecast . . . . . 30
2.3 Expense matrix of a binary warning system . . . . . 31

3.1 Climatogolical and hydrological characteristics of the four study regions 35
3.2 Properties of four representative catchments . . . . . 37

4.1 The predictor domain estimated for the height field and the moisture flux field . . . . . 68
4.2 Optimum distance parameters for analog forecasting . . . . . 69
4.3 Predictor configuration and predictor weights of predictor set A. . . . 70
4.4 Predictor configuration and predictor weights of predictor set B. . . . 72
4.5 Optimal number of analogs . . . . . 74
4.6 Binary outcomes for the prediction of extremes . . . . . 79

5.1 Statistical properties of the CPs classified for the Rhine basin with daily areal precipitation . . . . . 93
5.2 Statistical properties of the CPs classified for the Rhine basin and with daily discharge differences . . . . . 95
5.3 Binary outcomes of analog forecasting in comparison to the classification 99

7.1 Statistical properties of data depth categories conditioned on CPs . . 132

# List of Abbreviations

| Abbreviation | Meaning |
|---|---|
| m.a.s.l. | mean above sea level |
| CP | circulation pattern |
| COSMO | Consortium for Small-Scale Modeling |
| DWD | Deutscher Wetterdienst |
| ECMWF | European Centre of Medium Range Weather Forecast |
| EDs | algorithm with a small predictor domain - metric coefficient are uniform distributed |
| EDl | algorithm with a large predictor domain - metric coefficient are uniform distributed |
| EG | abbreviation for the study region located in Eastern Germany |
| EPS | Ensemble Prediction System |
| GCM | General Circulation Model |
| GFS | Global Forecast System |
| GME | Global-Modell |
| GPH | geopotential height |
| GPHa | anomalies of GPH |
| GPHn | geopotential height normalized to an interval [0;1] |
| HORIX | German abbreviation for Flood Risk Management of Extreme Events |
| IWS | Institute of Hydraulic Engineering at the University of Stuttgart |
| LAM | Limited Area Model |
| LARSIM | Large Area Runoff Simulation Model |
| LEPS | Limited Area Ensemble Prediction System |
| LM | Lokal-Modell |
| LUWG | German abbreviation for the Environment State Agency Rhineland-Palatinate |
| MG | abbreviation for the study region located in the middle of Germany |
| MM5 | Fifth-Generation Mesoscale Model |
| MOS | model output statistics |
| MSC | Meteorological Service of Canada |
| MSLP | mean sea level pressure |

| **Abbreviation** | **Meaning** |
|---|---|
| MSLPa | anomalies of MSLP |
| NCAR | National Center of Atmospheric Research |
| NCEP | National Center for Environmental Prediction |
| NMC | National Meteorological Center (today NCEP) |
| NOAA | National Oceanic and Atmospheric Administration |
| NWP | Numerical Weather Prediction |
| PP | perfect prognosis |
| PPFS | Probabilistic Precipitation Forecast System |
| PREC | precipitation |
| PREVIEW | Prevention, Information and Early Warning |
| RHUM | relative humidity |
| RIMAX | German abbreviation for Risk Management of Extreme Flood Events |
| SG | abbreviation for the study region located in Southern Germany |
| SHUM | specific humidity |
| TEMP | temperature |
| TGPH | resulting geostrophic flow strengths |
| UFLX | zonal moisture flux |
| UFLXn | zonal moisture flux normalized to an interval [0;1] |
| UGPH | zonal flow velocity |
| UTC | universal time coordinate |
| UWND | u-wind component |
| VFLX | meridional moisture flux |
| VFLXn | meridional moisture flux normalized to an interval [0;1] |
| VGPH | meridional flow velocity |
| VWND | v-wind component |
| WED | algorithm with metric coefficients which are symmetric distributed |
| WG | abbreviation for the study region located in Western Germany |
| WMO | World Meteorological Organization |
| WRF | Weather Research and Forecasting Model |

# List of Symbols

| Symbol | Definition | Unit |
|---|---|---|
| $\alpha$ | cost-loss ratio | $[-]$ |
| $\beta$ | weight balancing between closeness and similarity | $[-]$ |
| $\lambda$ | parameter of the exponential distribution | $[-]$ |
| $\mu$ | location parameter of the function for $g_k$ | $[-]$ |
| $\sigma$ | slope parameter of the function for $g_k$ | $[-]$ |
| | | |
| $a$ | number of hits | $[-]$ |
| $b$ | number of false alarms | $[-]$ |
| $c$ | number of miss | $[-]$ |
| $d$ | number of correct rejections | $[-]$ |
| $d_1$ | Euclidean distance | $[-]$ |
| $d_2$ | distance related to the Pearson correlation | $[-]$ |
| $d*$ | mixed distance | $[-]$ |
| $f_p$ | forecast probability | $[-]$ |
| $f_b$ | binary forecast | $[-]$ |
| $g_k$ | metric coefficient | $[-]$ |
| $i$ | indicator value of y | $[SI]$ |
| $k$ | number of k-nearest analogues | $[-]$ |
| $n$ | number of observations | $[-]$ |
| $n_{cp}$ | number of CPs | $[-]$ |
| $n_{st}$ | number of stations | $[-]$ |
| $nw$ | number of wet days | $[-]$ |
| $o_b$ | binary observation | $[-]$ |
| $\bar{o}$ | arithmetic mean of observation | $[SI]$ |
| $obj$ | objective function | $[-]$ |
| $p_t$ | decision threshold | $[-]$ |
| $s$ | standard deviation | $[SI]$ |
| $s$ | occurrence frequency | $[1/d]$ |
| $t$ | time | $[SI]$ |
| $x$ | predictor | $[SI]$ |
| $x_a$ | anomaly value of predictor | $[-]$ |
| $x_n$ | standardized predictor to an interval between zero and 1 | $[-]$ |
| $\bar{x}$ | arithmetic mean of x | $[SI]$ |
| $y$ | predictand | $[SI]$ |
| $\bar{y}$ | arithmetic mean of y | $[SI]$ |
| $y_t$ | threshold value for y | $[SI]$ |

| Symbol | Definition | Unit |
|---|---|---|
| $A_E$ | catchment area | $[km^2]$ |
| $B_{rel}$ | relative measure of reliability | [−] |
| $B_{res}$ | relative measure of resolution | [−] |
| $BS$ | Brier score | [−] |
| $BS_0$ | $BS$ of a reference forecast | [−] |
| $BS_1$ | $BS$ of the perfect forecast | [−] |
| $BSS$ | Brier skill score | [−] |
| $E_f$ | average expense of the forecast | [−] |
| $E_c$ | climate expense | [−] |
| $E_0$ | average expense of low-skill reference forecast | [−] |
| $E_1$ | average expense of the perfect forecast | [−] |
| F($y$) | cumulative distribution function of y | [−] |
| F'($y$) | inverse of cumulative distribution function of y | [−] |
| $FA$ | false alarm rate | [−] |
| $G$ | maximum forecast value | [−] |
| $HR$ | hit rate | [−] |
| $HSS$ | Heidke skill score | [−] |
| $J$ | number of predictors | [−] |
| $J$ | number of functions for $g_k$ | [−] |
| $K$ | number of grid points | [−] |
| $L$ | number of precipitation thresholds | [−] |
| $L_p$ | $L_p$-distance | [−] |
| $M$ | number of end users | [−] |
| $PC$ | proportion correct | [−] |
| $PC_r$ | PC of random forecast | [−] |
| $PI$ | precipitation index | [−] |
| $PSS$ | Peirce's skill score | [−] |
| $RPS$ | ranked probability score | [−] |
| $RPS_0$ | $RPS$ of a reference forecast | [−] |
| $RPS_1$ | $RPS$ of perfect forecast | [−] |
| $RPSS$ | ranked probability skill score | [−] |
| $S$ | value of a verification measure | [−] |
| $S_0$ | $S$ of a reference forecast | [−] |
| $S_1$ | $S$ of perfect forecast | [−] |
| $SS$ | skill score | [−] |
| $T_a$ | annealing temperature | [−] |
| $U_i$ | user interval | [−] |
| $V$ | forecast value | [−] |
| $V_{max}$ | maximum forecast value | [−] |
| $W$ | wetness index | [−] |

# Zusammenfassung

Lang anhaltende großflächige Starkniederschläge führen in zahlreichen Regionen der Erde immer wieder zu schweren Überschwemmungen und bedrohen so das Leben vieler Menschen. Zur Verringerung der negativen Auswirkungen solcher Ereignisse werden Hochwasservorhersagesysteme entwickelt. Auf Grundlage der Vorhersage dieser Systeme werden bei einem bevorstehenden Hochwasser Warnungen an die Bevölkerung ausgegeben, falls die Kapazitäten der Hochwasserschutzmaßnahmen in der betroffenen Region überschritten werden könnten. Den Kern eines Hochwasservorhersagesystems bilden mathematische Modelle, die die hydrologischen Prozesse in einem Flussgebiet mit mathematischen Gleichungen beschreiben, um den beobachteten und prognostizierten Niederschlag in Abfluss- oder Wasserstandsvorhersagen zu überführen.

Die Qualität einer kurzfristigen (1 bis 3 Tage) und mittelfristigen (4 bis 7 Tage) Hochwasserprognose wird im Wesentlichen von der Güte der Niederschlagsvorhersage bestimmt. Niederschlagsvorhersagen werden von einer Vielfalt mathematischer Prognosemodelle bereitgestellt. Die bewährtesten Modelle sind globale numerische Wetterprognosemodelle. Sie simulieren die großräumige atmosphärische Zirkulation und stellen so mehrmals täglich Vorhersagen für sämtliche Regionen der Welt bereit. Leider sind die Prognosen dieser Modelle für lokale oberflächennahe Prozesse, wie dem Niederschlagsgeschehen in kleinräumigen Flussgebieten, von geringer Güte. Der Hauptnachteil der globalen Modelle ist die grobe räumliche Auflösung, so dass eine adäquate Beschreibung dieser Prozesse nicht möglich ist. Allerdings ist gerade diese Information für die Beschreibung Hochwasser auslösender Prozesse von Bedeutung, da viele Hochwasservorhersagesysteme für kleinräumige und schnell reagierende Flussgebiete entwickelt werden.

Die Lücke zwischen den Vorhersagen globaler numerischer Wetterprognosemodelle und den Anforderungen lokaler Anwendungen wird durch Downscaling-Verfahren geschlossen (siehe Karl et al., 1990; Giorgi and Mearns, 1991). Diese Verfahren verknüpfen großräumige atmosphärische Variablen, wie die Geopotentialhöhe, mit lokalen Oberflächenvariablen, wie dem Niederschlag, um die globale Information mit dynamischen oder statistischen Verfahren auf die lokale Ebene zu überführen. Beim dynamischen Downscaling wird ein regionales numerisches Wetterprognosemodell in ein globales Modell integriert, um eine hochaufgelöste Wettervorhersage für das Zielgebiet zu erzeugen. Meistens werden beim dynamischen Downscaling deterministische Modellansätze verwendet, mit denen eine Realisierung einer Zielvariablen bestimmt werden kann, so dass eine Vorhersage von dem Modell bereitgestellt wird. Jedoch ist die Wettervorhersage (wie auch die Hochwasservorhersage) von Natur aus unsicher, egal ob die Vorhersage von einem mathematischen Modell oder von einem Experten stammt. Gerade die Niederschlagsvorhersage für kleinräumige Flussgebiete und für sehr selten eintretende Ereignisse ist mit einer großen Unsicherheit versehen. Für eine angemessene Beschreibung der Vorhersageunsicherheit sind da-

her Ansätze nötig, die die möglichen Realisierungen einer Zielgröße für verschiedene zeitliche Reichweiten eines Vorhersagemodells bestimmen können und deren Prognosen in Form von Wahrscheinlichkeiten ausgedrückt werden.

Zur rechtzeitigen Warnung der Bevölkerung werden schnelle und klare Entscheidungen benötigt. Experten weisen häufig darauf hin, dass Entscheidungen deutlich schwieriger zu fällen sind, wenn die Prognosen in Wahrscheinlichkeiten ausgedrückt werden. Dieses Argument ist richtig, aber es gibt mindestens zwei angemessene Gründe für die Wahl eines probabilistischen Ansatzes und für das Verwenden von Wahrscheinlichkeit bei der Vorhersage (Murphy, 1985):

1. Der Entwickler eines Prognosemodells und der Nutzer des Modells wissen aufgrund ihrer Erfahrung häufig, dass die Prognosen eines mathematischen Modells mit einer Unsicherheit behaftet sind. Diese Unsicherheit sollte sich in den Modellvorhersagen widerspiegeln und zu allen Empfängern einer Vorhersage kommuniziert werden.

2. Es kann einfach demonstriert werden, dass eine probabilistische Vorhersage verglichen mit anderen Vorhersageformen mindestens den gleichen oder einen höheren Informationswert besitzt.

Der letztgenannte Punkt ist wahrscheinlich das bedeutendste Argument für die Verwendung von Wahrscheinlichkeiten in der Vorhersage.

Zur Erfassung der Vorhersageunsicherheit wurden in den letzten beiden Jahrzehnten vermehrt Ensembleprognosesysteme entwickelt. Diese Systeme führen viele Läufe mit einem globalen numerischen Wetterprognosemodell durch, aber ausgehend von unterschiedlichen Beschreibungen des atmosphärischen Zustandes. Die Realisierungen des globalen Modells werden dann mit einem dynamischen Verfahren weiterverarbeitet, um ein Ensemble hochaufgelöster Wettervorhersagen für das Zielgebiet zu erstellen. In einer Reihe von Studien wurde veranschaulicht, dass Ensembleprognosen einen deutlich höheren Informationswert besitzen, wenn Wahrscheinlichkeiten zur Beschreibung der Vorhersageunsicherheit verwendet werden und der Entscheidungsträger weiss, wie er diese Information zu nutzen hat. Leider haben die Ensembleprognosesysteme den Nachteil, dass aufgrund des hohen Rechenaufwands nur eine geringe Anzahl von Vorhersagen erstellt werden können, so dass die Vorhersageunsicherheit für extreme Niederschlagsereignisse nicht angemessen bestimmt werden kann. Des Weiteren werden flexible Ausgaben von den Downscaling-Verfahren benötigt, da die optimale raum-zeitliche Auflösung eines Hochwasservorhersagesystems von den Gebietseigenschaften (z. B. der Reaktionszeit während eines Hochwassers) bestimmt wird.

Sehr gute Alternativen zum dynamischen Downscaling bieten die Verfahren des statistischen Downscalings. Bei dieser Art von Downscaling überführt eine Transferfunktion die globale Information auf die lokale Ebene. Nicht nur der geringe Aufwand und die hohe Effizienz dieser Techniken machen diese Verfahren attraktiv für den operationellen Einsatz, sondern auch der Fokus auf bestimmte Bereiche einer Vorhersagevariablen, z. B. der Extreme. Das wichtigste Argument für die Nutzung dieser Techniken ist, dass viele Realisierungen zur angemessenen Beschreibung der Vorhersageunsicherheit für eine Zielgröße erstellt werden können.

Ein besonderes Verfahren zur Vorhersage des Niederschlagsgeschehens in kleinräumigen Flussgebieten wurde von Obled et al. (2002) vorgestellt. Im Gegensatz zu den Verfahren nationaler Wetterdienste konzentriert sich die Methode auf die Vorhersage des täglichen Gebietsniederschlags, die zur Beurteilung der Hochwassersituation in einem Flussgebiet die bedeutendste Information ist. Den Hauptbestandteil des Verfahrens bildet die analoge Methode. Die Idee der analogen Methode ist vergleichbar mit einem erfahrenen Meteorologen, der die prognostizierte Wetterkarte mit Karten aus der Vergangenheit vergleicht, um die ähnlichsten (analogen) Wettersituationen zu identifizieren. Die lokalen Bedingungen dieser Wettersituationen (z. B. der tägliche Gebietsniederschlags eines Flussgebietes) werden dann für die Prognose verwendet. Allerdings wird bei diesem Verfahren der Kartenvergleich von einem mathematischen Suchalgorithmus übernommen, der die großräumigen Felder atmosphärischer Variablen, wie die des Luftdrucks oder des Feuchteflusses, über dem Untersuchungsgebiet vergleicht.

### Zielsetzung der Arbeit

In Rahmen dieser Untersuchung wurde ein Prognosemodell zur probabilistischen Vorhersage des Niederschlagsprozesses in kleinräumigen Flussgebieten vorgestellt. Die Komponenten des Modells sind statistische Verfahren des Downscalings, um die Anforderungen eines Hochwasservorhersagesystems für kleinräumige Flussgebiete möglichst zu erfüllen. Dabei kann der hier vorgestellte Downscaling-Algorithmus in zwei Schritte unterteilt werden: (i) Zunächst wird der tägliche Gebietsniederschlag für das Zielgebiet vorhergesagt; (ii) dann wird diese Information mit einem stochastischen Simulationsverfahren raum-zeitlich disaggregiert, um die Variabilität des Niederschlagsprozesses zu simulieren. Letztendlich stellt das Modell ein Niederschlagsensemble bereit, um die Unsicherheit der Niederschlagsprognose für kleinräumige Flussgebiete auf angemessene Weise für kurzfristige und mittelfristige Projektionen zu ermitteln.

Das Hauptziel dieser Untersuchung war die Entwicklung statistischer Verfahren zur Realisierung des ersten Schrittes des Algorithmus, der Vorhersage des täglichen Gebietsniederschlags. Die Umsetzung dieses Schrittes ist wohl die wichtigste Komponente des Modells, da eine angemessene Prognose des Gebietsniederschlags die notwendige Voraussetzung für die Entwicklung eines präzisen und somit möglichst qualitativ hochwertigen Hochwasservorhersagesystems ist. Dabei soll den Kern der statistischen Verfahren die analoge Methode bilden.

Neben der Entwicklung des Niederschlagsprognosemodells bestand ein weiteres Ziel dieser Arbeit darin, die folgenden Fragen näher zu untersuchen und zu beantworten:

- Wie präzise ist die Vorhersage des täglichen Gebietsniederschlags für kleinräumige Flussgebiete in Mitteleuropa, insbesondere in Warnsituationen?
- Wie kann ein Entscheidungsträger eine probabilistische Vorhersage in Warnsituationen optimal nutzen? Und wie wertvoll sind die Wahrscheinlichkeitsvorhersagen im Vergleich zu anderen Vorhersageformen?
- Welche Qualität besitzen die Prognosen der analogen Methode im Verhältnis zu anderen statistischen Downscaling-Verfahren?

- Wie kann die Vorhersage der analogen Methode verbessert werden?

Letztendlich war ein weiteres Ziel dieser Arbeit, die Anwendbarkeit des Niederschlagsmodells zu demonstrieren. Daher wurde ein Prototyp eines probabilistischen Niederschlagsvorhersagesystems entwickelt, der mehrmals täglich Vorhersagen für die Untersuchungsgebiete liefert.

### Untersuchungsgebiete und Modellentwicklung

Die Untersuchung wurde in mehreren kleinräumigen Flussgebieten der Elbe, der Oberen Donau und des Rheins durchgeführt, die typische montane und alpine Quellgebiete und Zuflüsse dieser Fließgewässer darstellen (Abb. 1). Die Einzugsgebiete zeichnen sich durch eine hohe Reliefenergie und eine geringe Reaktionszeit ($< 24$ h) während eines Hochwassers aus. Um angemessene Modellparameter zu bestimmen, wurden die Flussgebiete entsprechend ihrer Lage in vier Regionen (Süden, Westen, Mitte und Osten Deutschlands) unterteilt, für die dann die Modellentwicklungen ausgeführt wurden.

Die Untersuchung wurde hauptsächlich mit Reanalyse-Informationen durchgeführt. Auf diese Weise konnte die Güte der statistischen Verfahren für einen langen Zeitraum (mehr als 40 Jahre) mit Ausgabedaten von einem globalen numerischen Simulationsmodell bestimmt werden, das zur Rekonstruktion des vergangenen atmosphärischen Zustandes verwendet wurde und den globalen Modellen in der Wettervorhersage

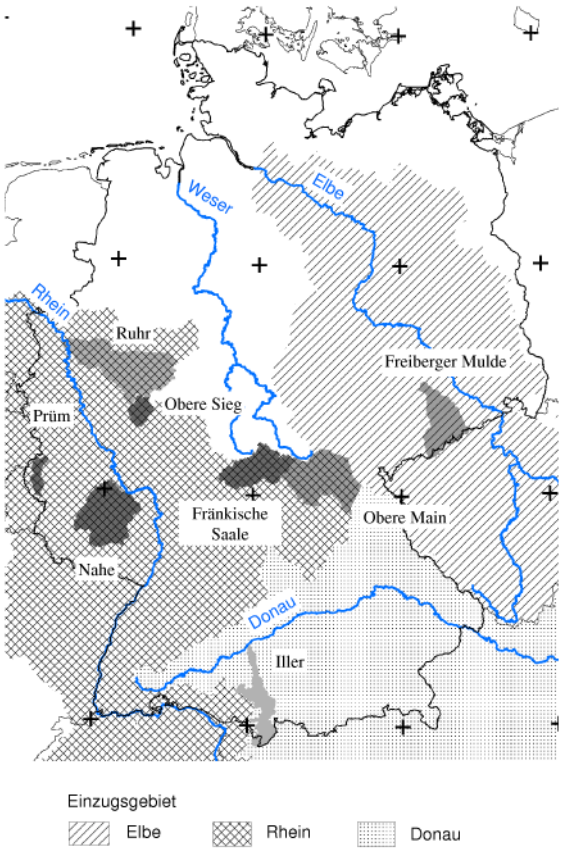



**Abbildung 1:** Lage der Untersuchungsgebiete

sehr ähnelt. Da das Niederschlagsmodell zukünftiger Bestandteil eines Hochwasserwarnsystems ist, wurden vorrangig die Prognosen des Modells für Warnsituationen bewertet. Der Vergleich mit einfachen Referenzvorhersagen zeigte, dass die analoge Methode angemessene Informationen für die Identifizierung extremer Niederschläge liefert.

Zusätzlich wurde in dieser Untersuchung gezeigt, dass die Prognosegüte der analogen Methode von mehreren Kriterien beeinflusst wird, die vom Entwickler subjektiv festgelegt werden müssen, bevor der Suchalgorithmus gestartet werden kann. Da eine Entscheidungshilfe für eine angemessene Auswahl der Kriterien nicht vorhanden ist und für das Downscaling von Niederschlägen in den Untersuchungsgebieten bisher die analoge Methode nur vereinzelt zur Anwendung kam, wurde eine umfassende Analyse zur Identifizierung der sensitivsten Modellgrößen durchgeführt. Die wichtigsten Einstellungen wurden dann für alle vier Regionen und für zwei verschiedene Sätze globaler Informationen ermittelt. Diese können dann für die Schätzung der prognostizierten Vorhersageunsicherheit für kurzfristige und mittelfristige Projektionen in den Gebieten eingesetzt werden. Des Weiteren können zukünftige Untersuchungen diese Information nutzen, um das Downscaling für die Regionen weiter zu verbessern.

### Beantwortung der Forschungsfragen

Am Anfang dieser Untersuchung besaß der Autor kaum Erfahrungen mit der Güte der Niederschlagsprognosen des dynamischen und statistischen Downscalings. Daher erhob sich zunächst folgende Frage:

***Wie präzise ist die Vorhersage des täglichen Gebietsniederschlags für kleinräumige Einzugsgebiete in Mitteleuropa, insbesondere in Warnsituationen?***

Zur Beantwortung dieser Frage wurden die Vorhersagen des COSMO-Modells (ehemals Lokal-Modell des Deutschen Wetterdienstes Schättler et al., 2008) analysiert, das zur Gruppe der dynamischen Downscaling-Verfahren gehört. Die Niederschlagsvorhersagen des Modells werden häufig als erstes von den regionalen Hochwasservorhersagezentralen in Deutschland zur Bewertung der Niederschlagssituation herangezogen.

Ein Vergleich mit beobachteten täglichen Gebietsniederschlägen zeigte deutlich die geringe Genauigkeit der Niederschlagsprognosen. Es wird häufig ein Fehlalarm ausgelöst oder das Ereignis sogar verfehlt (siehe Abb. 2). Dieses Resultat bestätigte die Vermutung vieler Experten, dass eine deterministische Niederschlagsvorhersage nicht ausreichend sei, um mit einem Hochwasservorhersagesystem eine wertvolle Warninformation zur Verfügung zu stellen.

Es wurden auch die Niederschlagsinformationen eines regionalen Ensembleprognosesystems (COSMO-LEPS; Montani et al., 2003) näher analysiert, das seit einigen Jahren für Mitteleuropa hochaufgelöste ensemble-basierte Wettervorhersagen liefert. Es konnte demonstriert werden, dass der Großteil der Beobachtungen innerhalb des Unsicherheitsintervalls der Ensembleprognosen liegt. Das Modell liefert somit wertvolle Informationen über die Vorhersageunsicherheit. Allerdings ist die Unsicherheit selbst für kurzfristige Projektionen (z. B. für den morgigen Tag) sehr

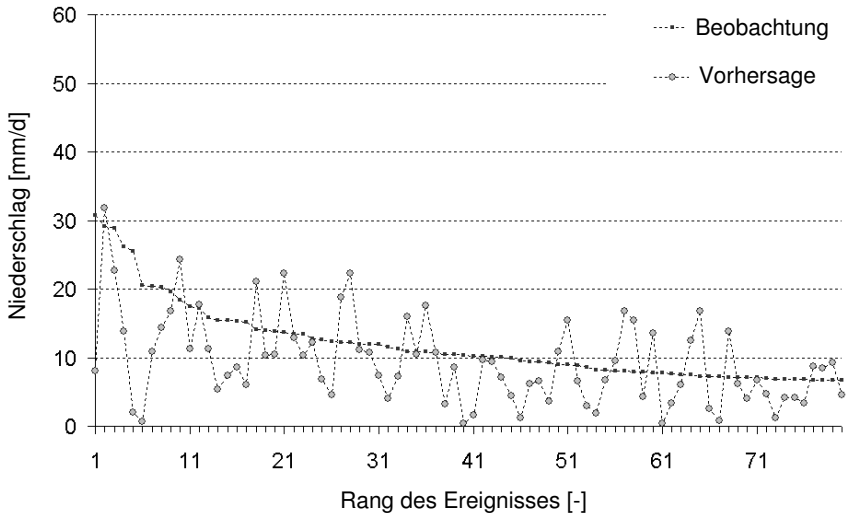



**Abbildung 2:** Vergleich zwischen der Niederschlagsvorhersage des COSMO-Modells und dem beobachteten Niederschlag für die 80 größten Ereignisse an der Freiberger Mulde, Elbe-Einzugsgebiet, Deutschland, täglicher Gebietsniederschlag (6 UTC bis 6 UTC des Folgetages) von April 2004 bis Dezember 2005; Vorhersagezyklus 0 UTC, zeitliche Reichweite: 6 fh bis 30 fh.

groß, obwohl nur wenige Realisierungen durch das regionale Modell transferiert werden. Da aufgrund der großen Vorhersageunsicherheit keine schnellen Entscheidungen mehr getroffen werden können, wird häufig von den Nutzern von Ensembleprognosen gefordert, die Anzahl der Realisierungen zu reduzieren. Leider bietet sich diese Strategie nicht an, da ansonsten vermehrt Niederschlagsereignisse außerhalb des Unsicherheitsintervalls liegen und eine Identifizierung von Extremereignissen zunehmend erschwert wird.

Den Problemen des dynamischen Downscalings wurden in dieser Arbeit mit statistischen Ansätzen begegnet, deren Prognosen in Form von Vorhersagewahrscheinlichkeiten ausgedrückt werden. Leider fallen klare Entscheidungen bei dieser Art der Vorhersageform weniger einfach. Daher war es notwendig, folgende Fragen näher zu beleuchten:

***Wie kann ein Entscheidungsträger eine probabilistische Vorhersage in Warnsituationen optimal nutzen? Und wie wertvoll sind die Wahrscheinlichkeitsvorhersagen im Vergleich zu anderen Vorhersageformen?***

Zur Lösung dieser Fragen wurde ein Kosten-Verlust-Ansatz verwendet (Thompson and Brier, 1955), um die Prozesse der Entscheidungsfindung zu beschreiben. Der Ansatz verknüpft die Resultate eines Warnsystems (Tab. 1) mit den Konsequenzen der Entscheidungsfindung (Tab. 2). Bei einer Alarmierung treten somit Kosten für Schutzmaßnahmen auf, während Verluste nur dann entstehen, wenn kein Alarm gegeben wird und das Ereignis dennoch eintritt. Die Untersuchung demonstrierte, dass die probabilistische Vorhersage verglichen mit anderen Vorhersageformen deutlich wertvollere Informationen liefert, da der Entscheidungsträger das opti-

**Tabelle 1:** Resultate einer binären Vorhersage nach Mason (2003)

| Vorhersage | Beobachtung | |
|---|---|---|
| | ja | nein |
| ja | Treffer | Fehleralarm |
| nein | Verfehlt | Treffer 2. Art |

male Quantil einer Vorhersage für einen individuellen Nutzer der Vorhersage wählen kann. Allerdings ist es notwendig, dass der Entscheider weiss, wie er Vorhersagewahrscheinlichkeiten zu interpretieren hat. Da dieses Wissen sehr bedeutsam für die Identifizierung intensiver Niederschläge ist, wurde die optimale Nutzung einer probabilistischen Vorhersage näher erläutert. Zusätzlich wurde nachgewiesen, dass mehrere Faktoren bei der Entscheidungsfindung berücksichtigt werden sollten, um letztendlich eine möglichst wertvolle Information an den Empfänger der Warnungen weiterzuleiten.

Der Kosten-Verlust-Ansatz wurde auch zur Bewertung der Prognosen der statistischen Verfahren verwendet, um ihren relativen Informationswert zu ermitteln. Die Niederschlagsprognosen der analogen Methode haben verglichen mit einfachen Referenzvorhersagen einen sehr hohen Informationswert, insbesondere für die Identifizierung lang anhaltender großflächiger Niederschläge. Letztendlich könnten viele individuelle Nutzer von den Modellvorhersagen profitieren. Allerdings gibt es bisher kaum Vergleiche mit anderen statistischen Vorhersagetechniken für die Untersuchungsgebiete. Daher war es notwendig, auch die folgende Frage näher zu untersuchen:

***Welche Qualität besitzen die Prognosen der analogen Methode im Verhältnis zu anderen statistischen Downscaling-Verfahren?***

Zur Beantwortung dieser Frage wurde eine auf Fuzzy-Regeln basierende Klassifikation benutzt (Bárdossy et al., 2002), um ein Katalog großräumiger täglicher Zirkulationsmuster für die Untersuchungsgebiete zu erstellen. Es konnte demonstriert werden, dass physikalisch plausible Zirkulationsmuster mit der Klassifikation erstellt werden können, mit denen angemessen zwischen trockenen und regenreichen Tagen unterschieden werden kann. Die Bewertung der Niederschlagsprognosen veranschaulichte, dass während nasser Zirkulationsmuster viele intensive Niederschlagsereignisse auftreten, so dass die Klassifikation einen angemessenen Informationswert für die Identifizierung dieser Ereignisse bereitstellen könnte. Allerdings werden im

**Tabelle 2:** Konsequenzen der Entscheidungsfindung

| Aktion | Beobachtung | |
|---|---|---|
| | Extrem | kein Extrem |
| Schutz | Kosten | Kosten |
| kein Schutz | Verluste | 0 |

Vergleich zur analogen Methode deutlich mehr Fehlalarme ausgelöst, so dass der Informationswert von der analogen Methode höher ist. Trotzdem sollte die Klassifikation zentraler Bestandteil des hier vorgestellten Niederschlagsmodells bleiben, da es eine Reihe von weiteren Anwendungsmöglichkeiten bietet, um das statistische Downscaling des Niederschlags weiter zu verbessern.

In dieser Untersuchung wurde zudem ein neuer Ansatz zum Downscaling des Niederschlags vorgestellt. Die Grundlage dieses Ansatzes bildet die Datentiefe mit der die Zentralität einer Wettersituation gemessen werden kann. Dabei wird angenommen, dass je größer die Datentiefe, desto zentraler und somit gewöhnlicher der atmosphärische Zustand über dem Untersuchungsgebiet. Da lokale Extremereignisse häufig durch Anomalien in der Atmosphäre, also durch ungewöhnliche Situationen, verursacht werden, scheint diese Idee ein vielsprechender Ansatz für die Identifizierung intensiver Niederschläge zu sein. Da ungewöhnliche Zustände der Atmosphäre sowohl zu sehr schlechtem wie auch zu sehr gutem Wetter führen können (Abb. 3), wurde zusätzlich zu diesem Verfahren die auf Fuzzy-Regeln basierende Klassifikation verwendet, um zwischen trockenen und regenreichen Situationen zu unterscheiden. Die Bewertung des neuen Verfahrens zeigte, dass deren Prognosen einen hohen Informationswert für die Identifizierung intensiver Niederschläge besitzen und mit dem Informationswert der analogen Methode vergleichbar ist. Daher ist es sehr zu empfehlen, diesen neuen Ansatz in zukünftigen Untersuchungen weiter zu verbessern.

Letztendlich lässt sich zusammenfassen, dass die analoge Methode ein qualitativ hochwertiger Ansatz zur Vorhersage des täglichen Gebietsniederschlags in kleinräumigen Einzugsgebieten darstellt und dass eine Weiterentwicklung dieses Ansatzes loh-

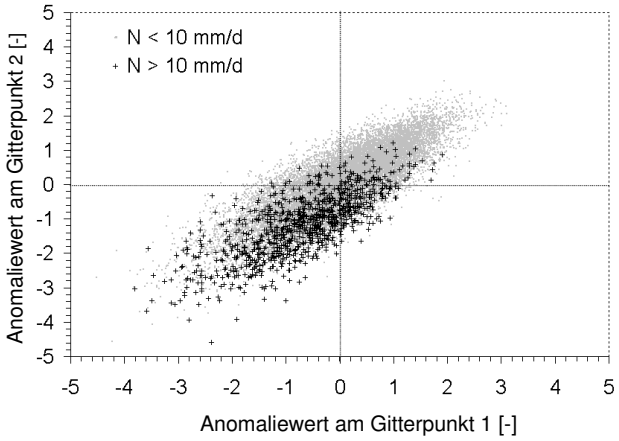



**Abbildung 3:** Streudiagramm täglicher Geopontialanomalien an zwei Gitterpunkten nahe dem Untersuchgsgebiet für zwei Niederschlagsklassen. Tage mit geringen Niederschlagsrisiken zeichnen sich durch positive Anomalien (hoher Lufdruck) aus, während bei negativen Anomalien (niedriger Luftdruck) die Anzahl der Tage mit intensiven Niederschlägen deutlich zunimmt.

nenswert erscheint.

### *Wie können die Vorhersagen der analogen Methode verbessert werden?*

In dieser Untersuchung wurden mehrere Ansätze zur Verbesserung der analogen Methode vorgestellt und näher untersucht. Hier werden zwei dieser Ansätze und deren Resultate näher erläutert:

1. **Abstand und Ähnlichkeit atmosphärischer Muster**: Die Modellgüte der analogen Vorhersage wird davon bestimmt, welches Maß zum Vergleich der atmosphärischen Felder (Muster) verwendet wird. Gewöhnlicherweise wird ein Entfernungsmaß ausgewählt, welches den Abstand zwischen zwei Feldern misst. Dabei wird angenommen, dass je geringer der Abstand der Felder, desto höher die Wahrscheinlichkeit, dass es sich um eine analoge Wettersituation handelt. Allerdings können sich zwei atmosphärische Muster auch hinsichtlich ihrer Form ähneln (Abb. 4). Jedoch werden der Abstand und die Ähnlichkeit zweier Muster bei der Ermittlung analoger Situationen nicht berücksichtigt. Aus diesem Grund wurde in dieser Untersuchung ein Maß vorgestellt, welches eine gewichtete Mischung zweier Entfernungsmaße ist, um je nach Bedarf zwischen Musterabstand und Musterähnlichkeit zu variieren. Es konnte gezeigt werden, dass die Wahl des neuen Vergleichsmaßes die Vorhersage intensiver Niederschläge leicht verbessern kann, wenn der Informationswert beider Entfernungsmaße ähnlich ist. Zukünftige Untersuchungen sollten überprüfen, ob das neue Vergleichsmaß auch die Vorhersage über den gesamten Bereich der Zielgröße verbessern kann.

2. **Auswahl einer angemessenen Entfernungsfunktion**: Bei der analogen Vorhersage muss der Entwickler neben dem Vergleichsmaß auch den Wirkungsbereich der großräumigen atmosphärischen Variablen festlegen, um den Abstand oder die Ähnlichkeit zwischen zwei atmosphärischen Feldern zu berechnen. Normalerweise erhalten die Gitterpunkte, die innerhalb des Wirkungsbereiches liegen, dasselbe Gewicht für die Entfernungsbestimmung. Allerdings ist diese Gewichtsverteilung eine unbefriedigende Wahl, weil gewöhnlicherweise der Einfluss einer Wirkungsvariablen zum Zielgebiet hin zunehmen sollte. Da die wahre Verteilung der Gewichte nicht bekannt und eine manuelle Festlegung

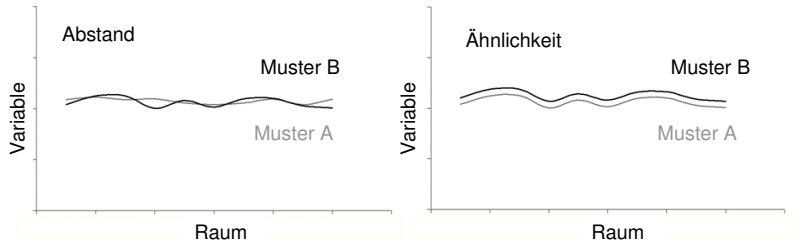



**Abbildung 4:** Zwei Arten von Ähnlichkeiten zwischen zwei atmosphärischen Feldern (Mustern): Musterabstand (linkes Bild) und Musterähnlichkeit (rechtes Bild)

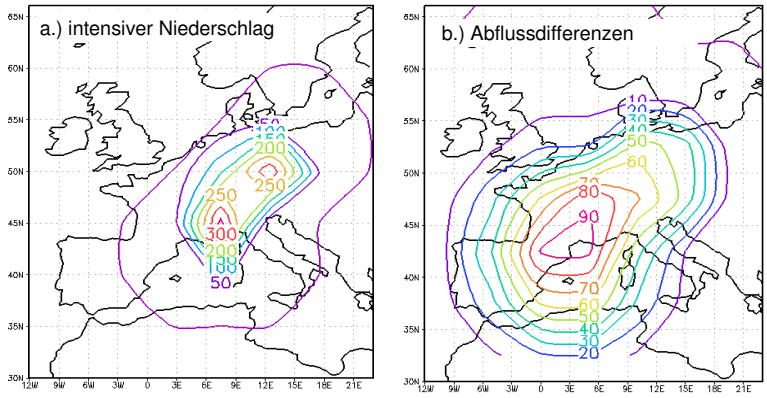



**Abbildung 5:** Wirkungsbereich der Geopotentialhöhe für zwei verschiedene Zielfunktionen: a.) „intensiver Niederschlag“ = Minimierung des Vorhersagefehlers für extreme Niederschlagsereignisse; b.) „Abflussdifferenzen“ = Minimierung des Vorhersagefehlers für kleine Abflussdifferenzen und hohe positive Abflussdifferenzen.

sehr zeitintensiv ist, wurde daher ein automatischer Algorithmus zur Optimierung der Gewichtsverteilung und somit zur Bestimmung einer geeigneten Entfernungsfunktion vorgestellt. Es wurde gezeigt, dass der Algorithmus einen angemessenen Wirkungsbereich für die globale atmosphärische Variable bestimmen kann. Zusätzlich konnten die Modellprognosen mit dieser Technik etwas verbessert werden. Ein Entwickler mit wenig Erfahrung könnte daher mit dem Optimierungsalgorithmus ein ähnliches Resultat erzielen wie ein Experte. Allerdings wird die Güte der Modellübertragbarkeit sehr von der Wahl der Zielfunktion beeinflusst. Ein Entwickler muss daher eine geeignete Zielfunktion für den Optimierungsalgorithmus wählen, ansonsten wird ein weniger glaubwürdiger Wirkungsbereich für die großräumige Information festgelegt (Abb. 5). Im Rahmen dieser Untersuchung stellte sich heraus, dass eine Zielfunktion, die nur den binären Vorhersagefehler für intensive Niederschlagsereignisse minimiert, weniger für eine Identifizierung robuster Parameter geeignet ist. Letztendlich lieferte ein Algorithmus, der den Vorhersagefehler für Abflussdifferenzen minimiert, die besten Ergebnisse hinsichtlich der Definition eines angemessenen Wirkungsbereiches und der Modellübertragbarkeit.

## Operationelle Anwendbarkeit

Um die Fähigkeiten des Niederschlagsprognosemodells für Warnsituationen der realen Welt zu testen, wurde in dieser Arbeit zusätzlich ein operationeller Prototyp des Vorhersagemodells entwickelt. Der Prototyp verwendet die Ausgaben des globalen numerischen Wetterprognosemodells des US-amerikanischen Wetterdienstes, um mehrmals täglich eine kurzfristige probabilistische Prognose des täglichen Gebietsniederschlags für die Einzugsgebiete dieser Untersuchung und für die nächsten drei Tage zu erstellen. In naher Zukunft wird der Projektionszeitraum erweitert und die zweite Komponente des Downscaling-Verfahrens, die stochastische

Simulation raum-zeitlich hochaufgelöster Niederschlagsfelder, in das operationelle System eingebettet. Das Vorhersagesystem wird dann für Testzwecke in einer der regionalen Hochwasservorhersagezentralen in Deutschland implementiert.

### Schlussfolgerung und Ausblick für zukünftige Untersuchungen

In dieser Untersuchung wurden vielversprechende statistische Techniken zur probabilistischen Vorhersage des täglichen Gebietsniederschlags vorgestellt und einige sehr interessante Forschungsfragen im Themengebiet der quantitativen Niederschlagsvorhersage und -warnung beantwortet. Dennoch gibt es eine Reihe von weiteren Techniken und Fragen, die in zukünftigen Studien näher untersucht werden sollten. Die für die Anwenderseite dringendste Frage ist, wie ein Vorhersagesystem erstellt werden kann, das möglichst vielen Empfängern nützliche Hochwasserwarninformationen liefert. Die Arbeit zeigte, dass die Erstellung probabilistischer Niederschlagsvorhersagen mit statistischen Downscaling-Verfahren und die Analyse des Entscheidungsprozesses mit dem Kosten-Verlust-Ansatz ein erster angemessener Schritt zur Erstellung eines solchen Warnsystems zu sein scheint. Die Weiterentwicklung dieser Ansätze sollte einen zentralen Bestandteil zukünftiger Untersuchungen bilden. Des Weiteren ist es wichtig, die Empfänger der Vorhersagen zu identifizieren und zu kategorisieren, wobei die Warninformation und deren Weiterleitung nach den Bedürfnissen der Empfänger ausgerichtet werden sollte. Dieser Schritt wurde in dieser Arbeit weniger betrachtet und bedarf einer näheren Untersuchung in künftigen Arbeiten.

Letztendlich hofft der Autor dieser Arbeit sehr, dass zukünftige Forschungsarbeiten von den hier vorgestellten Ideen und Methoden profitieren können und dass die eine oder andere Idee oder Methodik später in der hydrologischen Praxis ihren Platz findet.

# 1 Introduction

*(... In that summer, on the Saturday before St Mary Magdalene [19 July] it began raining at night it rained the whole night and on the following day on Sunday [20 July] the whole day, and on Monday [21 July] in the morning extreme flooding occurred and all the mills and baths around Prague were destroyed and swept away. And then the bridge of the Major Town of Prague collapsed at night before the vigils of St Mary Magdalene from Monday to Tuesday [21 July to 22 July] ...)*

Description of a Prague citizen during the tremendous Vltava flood in July 1432 (taken from Brazdil et al., 2006).

## 1.1 Motivation

The protection of people, property, and environment against river flooding is a major challenge in hydrology. In the temperate zones as well as in the subtropics and tropics flooding is mainly caused by long-lasting large-area precipitation. In these regions, many people live close to river beds, in particular in poor countries, whose habitats are threatened by floods. As a typical example the great flood of the Yangtze River serves in 1931 when more than four million people lost their life both from the flooding itself as well as from the subsequent famine (O'Connor and Costa, 2004). Fortunately, river flooding in Central Europe is less catastrophic since the potential of the atmosphere to store water and the speed of moisture fluxes is less in the temperate zones compared to subtropical and tropical climates. However, great floods can inundate large areas of a river basin affecting many people and can lead to high monetary losses for a certain region or country. For instance, the economic losses of the severe Danube and Elbe flood in August 2002 amounted to more than 20 Billion US Dollar (Toothill, 2002). During the last years several large floods occurred in Central Europe, e.g. the Rhine flood in 1993, the Odra flood in 1997, and the Northern Alp flood in 2005 which caused substantial damage in several countries.

There are several management strategies in water resources planning to reduce the flood risk of a certain region. Usually, the protection against river flooding begins in the head catchment with flood retention reservoirs and ends with a dike or a (removable) flood wall downstream. However, all flood protection structures are designed to give protection to a certain water level like a water level that occurs once every one hundred years. For water levels exceeding this threshold, there is no protection. It is merely a question of time and not a question of if, when an

intensive precipitation occurs to cause a flood that exceeds the capability of those structures.

For the protection against flooding there are warning systems (see Krzysztofowicz, 1993; Obled et al., 2004). The outcomes of these systems are taken by decision makers as a source of information for early warnings to save human lives and to reduce damages. The integral part of a flood warning system forms a hydrological prediction model which describes the runoff processes of the real world by mathematical equations to transfer local surface fluxes such as precipitation into discharge or water level forecast. To describe the actual state of the catchment and to estimate the future weather state, a data assimilation system collects observations from monitoring networks and forecasts from weather prediction models which are needed to force the hydrological prediction model. The quality of a flood warning system to trigger a warning early enough mainly depends on the performance of the weather forecast. In temperate climate zones characteristic for Central Europe, the key factor that determines the accuracy of a flood warning system is the quality of the precipitation forecast.

A precipitation forecast can be derived from a variety of prediction models. Probably, the most established models are global numerical weather prediction (NWP) models which provide forecasts for up to two weeks for different geographical regions around the world. The main component of a global NWP model is a general circulation model (GCM). This simulates the current situation of the large-scale atmospheric circulation and their prospective evolution. It is well accepted that GCMs have significant skill in simulating and predicting the atmospheric flow processes at continental and regional scale, because the main features of the atmospheric circulation are incorporated in a GCM (Zorita and von Storch, 1997). However, the GCMs predictions are poor for local surface variables like precipitation or temperature (see e.g. Grotch and MacCracken, 1991; von Storch et al., 1993). There are at least three reasons for the failure of GCMs (Zorita and von Storch, 1997). Probably, the most obvious shortcoming is the coarse spatial resolution of GCMs leading to an incorrect description of the earth surface. For example, the GCM of the global NWP model of the American Weather Service has a resolution of 0.5° (NCEP, 2003), which corresponds to a grid spacing of around 50 km in Central Europe. However, for an adequate description of flood producing processes at the local-scale, weather forecasts in a high spatiotemporal resolution are needed. For instance, the regional flood forecasting centers in Germany design prediction models for the description of the runoff processes in small river basins such as head catchments or sub-basins of the biggest rivers in Germany. In mountainous areas these catchments are often characterized by a short concentration time during a flood ($< 24$ h). Thus, a fine spatiotemporal resolution of the models is crucial to trigger an early warning. For instance, the prediction model of the regional forecasting centers in Bavaria operates in an hourly resolution (Hangen-Brodersen et al., 2008). However, the optimal resolution of a hydrological prediction model is not known (Obled et al., 2004). It depends on factors such as the concentration time of a catchment and the time range of the projections. Thus, there is a strong need to design techniques which can bridge the differences between global NWP models and local-scale applications.

The gap between GCMs and the requirements of hydrological applications is closed by downscaling techniques (see Karl et al., 1990; Giorgi and Mearns, 1991). Down-

scaling is based on the assumption that a functional relationship can be established between the large-scale atmospheric circulation and the local-scale surface fluxes such as precipitation (Bárdossy, 2000). In climate modeling, downscaling approaches are separated into two groups: dynamical (process-based) and statistical (empirical) downscaling (see e.g. von Storch et al., 1993; Hewitson and Crane, 1996; Zorita and von Storch, 1997). In weather forecasting the term downscaling is rarely used although the origin of downscaling comes from this research field (see Karl et al., 1990; Giorgi and Mearns, 1991). In weather forecasting the term downscaling is also called adaptation of NWP outputs (see Obled et al., 2002). Like global NWP, dynamical downscaling takes process-based equations for the simulation of the atmospheric flow processes. The common way of dynamical downscaling is to nest a limited-area model (LAM) into a GCM to simulate the atmospheric flow in a high spatiotemporal resolution for a limited geographical region (Giorgi and Mearns, 1991). This provides a high-resolved weather forecast for a region of interest. In dynamical downscaling it is quite common to apply deterministic approaches which provide a single realization of the surface variable of interest. These approaches are favored by many end users (and some researchers as well) since a single forecast is easy to handle and it allows for a straightforward decision in warning situations. Unfortunately, the accuracy and the value of deterministic forecasts is poor for the prediction of intensive precipitation (Atger, 2001; Richardson, 2003) as well as for other meteorological variables (Zhu et al., 2002; Zhu, 2004). The major shortcoming of deterministic approaches is that the uncertainty of the weather forecasts is ignored. Weather forecasts are inherently uncertain, irrespective of whether they are based on prediction models or the knowledge of human forecasters or decision makers (Murphy, 1985). The uncertainty of weather predictions (and hydrological predictions as well) is based on different sources such as the incomplete description of the atmosphere state due to the lacks in observational systems and the assumptions and the simplifications made during the model development (Murphy, 1985).

For an adequate description of the forecast, including uncertainty of local-surface variables, downscaling techniques are needed which can quantify all possible realizations of the surface variable of interest for different lead times (Bárdossy, 2000) and which can expresses the forecast uncertainty in probabilities. It is often argued by the end users that decision making in the face of uncertainty is more difficult compared to a deterministic forecast meaning that rapid decisions are not possible. It is true that decision making based on forecast probabilities is not straightforward. However, there are at least two reasons why the outcomes of weather prediction models should be formulated and expressed in forecast probabilities (Murphy, 1985):

1. Due to the different sources of uncertainty, a forecaster knows that the outcomes of weather prediction models are uncertain. This uncertainty should be reflected in model results and should be also communicated to all users within a forecasting chain.

2. It can be easily demonstrated that a probability forecast has at least the same or a higher value than a non-probability forecast.

The latter reason is probably the most convincing argument in favor of the use of probability forecasts in decision making situations. To illustrate this argument,

cost-loss approaches can be used (e.g. Thompson and Brier, 1955; Murphy, 1977). These approaches describe the decision making processes in the real world and their consequences. There are many examples which demonstrated the higher benefit of probabilistic techniques compared to deterministic approaches or other forecasting techniques (e.g. Richardson, 2000, 2003), especially for the prediction of local extremes with a high variability such as intensive precipitation.

To describe the forecast uncertainty, global ensemble prediction systems (EPS) have been developed in the last two decades and are nowadays well established in weather forecasting. Several years ago, a mesoscale EPS was proposed to transfer the outcomes of a global EPS through a LAM to provide a high-resolved ensemble forecast for Central Europe (Montani et al., 2003; Marsigli et al., 2005). The precipitation forecast of the mesoscale EPS are already tested by one regional flood forecasting center in Germany to run their hydrological prediction model (see Hangen-Brodersen et al., 2008). However, the spatiotemporal resolution of ensemble forecasts is usually coarser than the resolution of deterministic forecasts, due to the high computational demand of those prediction models. Furthermore, only a limited number of ensemble members are transferred through the LAM and it is questionable whether the limited number of realizations can approximate adequately the forecast uncertainty of local surface variables, in particular for longer lead times and for extremes.

An alternative technique of downscaling is called statistical downscaling. In statistical downscaling, a transfer function describes the observed relationship between the large-scale and the local-scale information (Hewitson and Crane, 1996). There are many reasons for the use of statistical downscaling (see e.g. Zorita and von Storch, 1997). Probably, the major advantage is that a number of realizations can be provided for the surface variable of interest to enable an adequate description of the forecast uncertainty. Like NWP, the techniques of statistical downscaling are also well established and they have a long history in weather forecasting (see Panofsky and Brier, 1958; Glahn, 1985). Today, several national weather forecast centers apply the techniques of statistical downscaling to provide probability forecasts for precipitation and other surface variables (Wilks, 2006). The predictand is mostly the precipitation amount at a certain location (point precipitation). However, hydrologists are mainly interested in the precipitation amount over a particular river basin (areal precipitation).

A methodology for the prediction of areal precipitation was proposed and tested by Krzysztofowicz et al. (1993). They noted that the forecaster can focus on the estimation of a single variable instead of taking into account the forecast uncertainty at certain points in a river basin, the high spatial variability and the spatial dependence of the predictand. Several years later, Obled et al. (2002) proposed a precipitation model which provides probabilistic precipitation forecasts of up to seven days for small river basins in France. This model forms a component of a flood forecasting system designed for early warnings and reservoir operation (Obled et al., 2003, 2004). The methodology was enhanced by Bontron and Obled (2003) who took reanalysis data for the model development. Reanalysis data are the simulation output of a GCM performed for the reconstruction of a past climate. Compared to the traditional approaches in weather forecasting (see Klein et al., 1959; Glahn and Lowry, 1972), this kind of model development might have the advantage that the

bias of a GCM can be incorporated into the model development. This case would be if the bias of the GCM used to produce the reanalysis information is similar to the bias of the GCMs applied in global NWP. A further advantage is that a long data set can be used for the model training and model validation. The latter reason is particularly useful for the predictions of rare events.

The core of the precipitation model developed by Obled et al. (2002) represents the analog method (e.g. Lorenz, 1969). The basic idea of the approach is quite similar to a human forecaster in that the current weather map is compared to an archive of older weather maps to devise empirical rules for the forecast. Here, the human forecaster is replaced by an automatic search algorithm. This algorithm is used to identify past weather states that are similar (analog) to the target one. Afterwards, the areal precipitation of the analogs is chosen to provide a probabilistic precipitation forecast for the basin of interest.

## 1.2 Objective and Research Questions

In this study the basic framework for a precipitation model for short-term (one to three days) and medium-term (four to seven days) projections is proposed. The integral parts of this model are the statistical downscaling techniques developed to meet the requirements of a flood warning system for small river basins. The general structure of the precipitation model can be divided into two elements: (i) The daily areal precipitation is predicted for the basin of interest and (ii) the disaggregation of the areal information into precipitation realizations with a high spatiotemporal resolution. Finally, the model provides a high-resolved probabilistic precipitation forecast for the river basin of interest to describe the forecast uncertainty of the precipitation process for different lead times.

The primary objective of this thesis is the development of statistical techniques to realize the first element of the downscaling algorithm, the prediction of daily areal precipitation. This is the most crucial step because an adequate estimation of the daily areal precipitation is the basic prerequisite for an accurate and valuable flood warning system.

The precipitation model for the prediction of daily areal precipitation follows the approach proposed by Obled et al. (2002). Thus, the integral part of this model represents the analog method. Furthermore, the following research questions are proposed and addressed in this thesis:

- How accurate is the forecast of daily areal precipitation for small river basins in Germany, in particular in flood producing situations?
- What is the optimal use of a probability forecast in decision making situations? And, how valuable are probability forecasts in comparison to other forecast techniques?
- How accurate and valuable are the predictions of the analog method in relation to other statistical forecasting techniques?
- How can the traditional way of the analog method be improved?

The statistical techniques presented in this study are evaluated with past data over a long investigation period. This is required, since long-lasting large-area precipitation is a rare event and a number of events are needed to evaluate the goodness and the weakness of model predictions. Furthermore, the precipitation model should also demonstrate its capability for operational purposes. Therefore, the basic framework of an operational system is also proposed in this work which has been developed during the last years.

The precipitation model is tested for head catchments and the basins of small tributaries of the Elbe, the Danube and the Rhine. Please note that the methods proposed here will be adapted for the catchments of this study. Since the operational model will be tested in the future by the regional flood forecasting center of the Environmental State Agency Rhineland-Palatinate, Germany, the development and the comparison of the statistical techniques focus on the catchments in their region of interest.

The second step of the algorithm is not presented in this thesis but a basic framework of a methodology for the disaggregation of daily areal precipitation was proposed in Bliefernicht et al. (2008). A brief description of the second step is also given in Obled et al. (2004).

## 1.3 Outline of the Thesis

The structure of the thesis is similar to the order of the aforementioned research goals and questions so that the reader can easily follow the investigations to find some solutions and answers to the proposed questions.

In **Chapter 2**, a brief overview of the variety of mathematical modeling techniques developed in weather forecasting is given. The strengths and the limitations of dynamical downscaling are discussed by means of a deterministic and an ensemble precipitation forecast, commonly taken at the regional flood forecasting centers in Germany. The motivation and some arguments for the use of statistical downscaling are presented. In the second part of this chapter, an introduction into forecast verification is given. It is analyzed which kind of performance measures is appropriate for the evaluation of the goodness of a probability forecast in warning situations. In **Chapter 3**, a brief description of the selected study regions and a short specification of the data used for the downscaling is given. Additionally, it is shown when intensive precipitation and floods occur in Germany and what the causes of the extremes are.

The integral part of this thesis concerns the Chapters 4 to 7. In **Chapter 4**, the basic principle of the analog method is presented. It is illustrated that the model performance of this approach depends on several criteria that must be defined a priori by the investigator before starting the search algorithm. Since the precipitation model represents an element of a flood warning system, the performance of the analog forecast is determined for warning situations. Furthermore, an objective function is proposed which minimizes the forecast error in terms of a binary warning system ("*no warning*" and "*warning*"). Finally, it is demonstrated how an end user should apply a probability forecast in decision making processes to maximize his benefit. In **Chapter 5**, a brief introduction into weather classification is given. An

automated fuzzy rule-based classification technique is selected to compile a catalog of daily large-scale circulation patterns for the regions of interest so that days with a high risk of intensive precipitation can be indicated. Furthermore, it is highlighted that the choice of the objective function can strongly affect the goodness of the model transferability. Finally, the forecast performance of this methodology is analyzed for warning situations and compared to predictions from the analog method. In **Chapter 6**, an optimization algorithm for the analog method is proposed to allow an optimal distance function for the selection of analogs to be defined. It is demonstrated that the new technique can improve the prediction for intensive precipitation compared to the traditional way of analog forecasting. Additionally, an example is given how the classification of circulation patterns can be used for the enhancement of the analog method. In **Chapter 7**, a new statistical downscaling approach is presented which is based on a data depth function. The basic idea of the approach is highlighted and it is illustrated how this methodology can be applied for the prediction of daily areal precipitation. The forecast performance of this approach is analyzed and compared to the analog method developed in this thesis.

Finally, the current status of the probabilistic precipitation forecast system is briefly described in **Chapter 8**. The performance of this system is illustrated by means of a heavy precipitation event that affected western Germany in December 2007. Some examples are presented how to communicate probabilistic precipitation forecast for warning levels in a suitable way. The thesis ends with a chapter summarizing the work and giving some answers to the proposed research questions and a brief outlook for future investigation.

Most of the work illustrated in this investigation was part of the project HORIX (a German abbreviation for Flood Risk Management of Extreme Events). This research project was funded in the framework of the national research program 'Risk Management of Extreme Flood Events' (RIMAX; www.rimax-hochwasser.de) that became initiated after the severe Elbe and Danube flood in August 2002. The main objective of HORIX was the development of a short-term flood prediction model for head catchments and small tributaries of the Rhine and Elbe basin in Germany. Parts of this work presented here were also realized in the international research project PREVIEW (Prevention, Information and Early Warning; www.preview-risk.com) and in cooperation with the regional flood forecasting center of the Environmental State Agency Rhineland-Palatinate.

# 2 Basic Theory

Since the introduction of computers in weather forecasting a number of prediction models have been developed. They are established as operational models at national weather forecast centers providing day by day predictions for different geographical regions around the world. In this chapter a short introduction about the various techniques applied in weather forecasting is given and their advantages and limitations are discussed. Additionally, a brief description of two numerical weather prediction models is presented, whose precipitation forecast are used by the regional flood forecasting centers in Germany to run their hydrological model. An example of the precipitation forecast of both models should illustrate the problems of predicting intensive daily areal precipitation for medium-sized river basins.

The outcomes of weather prediction models are applied by a variety of end users as source of information for decision making. The primary interest of the users is to gain a benefit from the weather forecast. To quantify their benefit, the correspondence between the model predictions and the observation must be quantified. This subfield of weather forecasting is called forecast verification. In the second part of this chapter an introduction to forecast verification is given and various performance measures are proposed to quantify the accuracy and the value of a forecast.

## 2.1 Weather Forecasting

Weather forecasting is the prediction of the state of the atmosphere for the future of a given location. In weather forecasting, the atmospheric state of the past is observed to understand the dynamics and the physical processes of the atmosphere. Based on this knowledge mathematical models are developed to predict the prospective evolution of the atmosphere.

The origin of modern weather forecasting goes back to the 19th century. In the early approaches a forecaster compared a weather map describing the current state of the atmosphere with past weather maps to devise empirical rules to forecast the variable of interest. This kind of technique belongs to the group of subjective (manual) forecasting. Since the advance of computer technology, objective (automated) techniques have mainly replaced the empirical rules made by human forecasters. These techniques are referred to as being objective since rerunning the model with the same input will not change the forecast.

The techniques of objective weather forecasting are mainly separated in two groups: statistical forecasting and numerical weather prediction (NWP, Glahn and Lowry, 1972). In NWP atmospheric processes are described by physically based equations. In statistical forecasting a transfer function is established linking a predictor or a predictor set to the forecast variable of interest. Both techniques are used at national weather forecast centers, but which kind of technique is preferred

depends on factors like the spatial scale, the time range of the forecast (lead time), the forecast variable and the scientist itself. The techniques of statistical forecasting are widely used in operational weather prediction. An introduction to statistical forecasting and the techniques used in the research field can be found in Murphy and Katz (1985) and Wilks (2006). However, since the introduction of global NWP models in the mid-70ties, the techniques of NWP are well established in weather forecasting, too. NWP models are mainly used to provide short-term (1 to 3 days) as well as the medium-term (4 to 10 days) forecasts. The application of statistical forecasting is usually restricted to very short lead times (e.g. 15 minutes to 6 hours) since the forecast of NWP models are due to the time-consuming data assimilation not available for this time range (Wilks, 2006). Statistical forecasting is also applied for long-term weather predictions (monthly or seasonal) because the lead time of NWP models is seldom longer than two weeks.

In this investigation a prediction model is developed to forecast the evolution of daily areal precipitation in meso-scale catchments for the next seven days. Numerical weather prediction is well established for this lead time. However, it is common to use statistical techniques since the precipitation forecasts of NWP models are poor under certain conditions which is explained in the next section.

### 2.1.1 Global Numerical Weather Prediction

Global NWP models have been established to forecast the evolution of the large-scale atmospheric circulation. Their development goes back to the work of von Neumann and his group which developed the first NWP model in 1950 (Simmons, 2004; Fraedrich et al., 2005). Around twenty years later, in 1974, the first global NWP model became operational at the US National Meteorological Center (NMC; Simmons, 2004). Today, several national weather forecast centers run their own global NWP models, so that various forecasts are available for a certain region. For Central Europe the most relevant global dynamical models are the Global Forecast System (GFS) of the US National Center of Environmental Predictions (NCEP; formerly NMC; NCEP, 2003), the global NWP model of the European Center of Medium Range Weather Forecast (ECMWF; Simmons et al., 1989) and the Global-Modell (GME; Majewski et al., 2002) of the German Weather Service (DWD).

The main integral part of a global NWP is a general circulation model (GCM) which simulates the evolution of the large-scale atmospheric circulation. A GCM consists of two components (Fraedrich et al., 2005): A core describing the dynamics of the atmosphere with primitive thermo-dynamic equation systems based on the Navier-Stokes equation. Most of the equations are partial differential equations which cannot be solved in an analytical way. To integrate the equations, numerical differentiation schemes are applied which give an approximate solutions of the equations. Usually, a three-dimensional GCM is taken for the description of the atmospheric flow processes often with an regular grid for the horizontal direction and a irregular grid for the vertical direction. For instance, the horizontal resolution of the GFS is 0.5 x 0.5° which corresponds to a grid spacing of 35 x 55 km in Central Europe. The vertical resolution of the simulation model consists of 64 layers describing the planetary boundary layer with a finer resolution than the free atmosphere. The internal time step of the GFS used for the integration of the differential

equations is 7.5 min (NCEP, 2003). Due to the coarse spatial resolution of a GCM sub grid-scale processes like cloud formation and surface fluxes e.g. precipitation are not described by the dynamical core. The sub grid-scale processes are simplified to empirical equations with parameters describing the net effect of these processes (von Storch et al., 1993). Originally, the GCMs have been developed in weather forecasting but they are also well established in climate modeling for simulation of the past climate and to perform projections of the future climate. The GCMs used in weather forecasting and in climate modeling are based on the same mathematical concept, but between the models there are often slight differences in model structure (e.g. global NWM models often contain no ocean model) or in model resolution.

A further component of a global NWP model is the global data assimilation system which collects a huge amount of observations from a number of meteorological variables which describes the current state of the atmosphere. Most of the measurements are from synoptic observations from land and sea stations, but also measurements from radiosondes and satellites as well as the results of the previous forecast are incorporated in the GCM (Fraedrich et al., 2005; Simmons, 2004). The meteorological observations are mostly point information. To define the initial conditions of a GCM, the data must be interpolated to each grid cell. Von Storch et al. (1999) mentioned that the increase of the forecast quality of global NWP mainly results from the improvement of the global data assimilation system.

### 2.1.2 Downscaling

GCMs are suited for the simulation of the current situation of the large-scale atmospheric circulation and to predict their future development (Zorita and von Storch, 1997). It is well accepted that GCMs have significant skill in predicting meteorological variables at continental and regional scale, because the main features of the atmospheric circulation are incorporated in a GCM. But in climate modeling as well as in weather forecasting it has been reported that the predictions of GCMs are poor for local surface variables like precipitation or temperature (see e.g. Grotch and MacCracken, 1991; von Storch et al., 1993). The performance of the GCMs are often at odds with the needs of hydrological applications (Hewitson and Crane, 1996) since the accuracy of the predictions of the GCMs grows with increasing scale (von Storch et al., 1993; Xu, 1999). Probably, the major drawback is the coarse spatial resolution of the GCMs leading to an inaccurate description of the earth surface (Zorita and von Storch, 1997). Figure 2.1 attempts to illustrate the crucial differences between the real world and the model world. A second shortcoming of the GCMs is the incomplete formulation of the sub-grid scale processes which are necessary to simulate the weather on the local-scale (Giorgi and Mearns, 1991). Zorita and von Storch (1997) emphasized that sub grid-scale processes have the greatest ecological and societal impact and there is a broad consensus in the research community for an adequate modeling of those processes.

To overcome the scale mismatch between GCMs and the requirements of local-scale applications, in weather forecasting as well as in climate modeling, downscaling approaches have been developed (see e.g. Karl et al., 1990; Giorgi and Mearns, 1991). A downscaling approach links large-scale atmospheric variables (predictor) such as the geopotential height to local-scale surface variables (predictand) e.g. precipita-

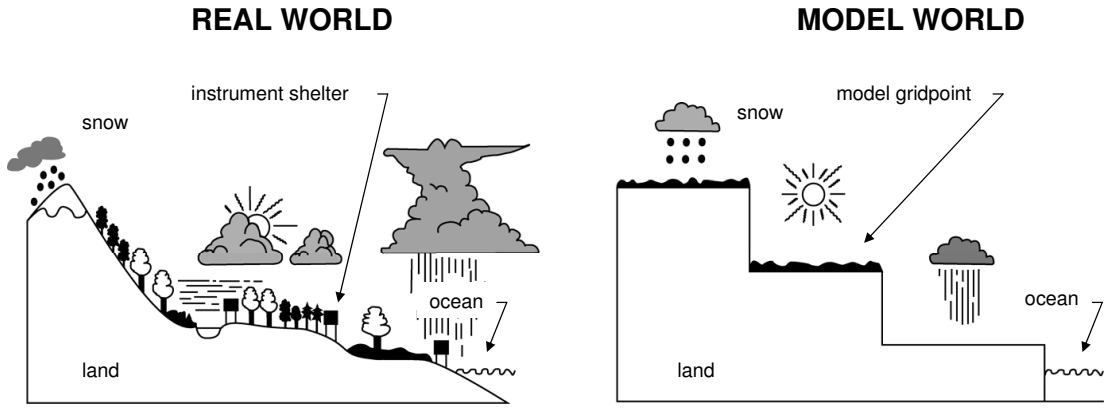



**Figure 2.1:** Illustration of the real world monitored by observations in comparison to the world of a general circulation model. Map from Karl et al. (1989) drawn from Wilks (2006).

tion. The downscaling is based on the assumption that a functional relationship can be established to describe the observed relationship between large-scale and local-scale information as follows (Bárdossy, 2000):

$$y\left(u,t\right) = F\left[x\left(U,T\right)\right] \tag{2.1}$$

where $x$ is the predictor over the domain $U$ at time $T$ and $y$ is the predictand at location $u$ and time $t$. The relationship between predictor and predictand is not unique and many different realizations are possible for the surface variable of interest (Bárdossy, 2000). Deterministic approaches are based only on a single realization and are therefore incomplete downscaling approaches, since the main task of downscaling is to determine the distribution of all possible realizations (Bárdossy, 2000). A full downscaling for hydrological applications should encompass all predictands (e.g. precipitation, temperature or wind) which are relevant for those applications (Bárdossy, 2000).

In the last two decades a variety of downscaling approaches have been developed. In climate modeling, downscaling approaches are separated in two groups: dynamical (process-based) and statistical (empirical) downscaling (see e.g. von Storch et al., 1993; Hewitson and Crane, 1996; Zorita and von Storch, 1997). In weather forecasting the term downscaling is seldom used, although the origin of downscaling is from this research field (see e.g. Karl et al., 1990; Giorgi and Mearns, 1991). Instead of speaking of downscaling the notion of „adaptation“ of NWP output is sometimes used (see Obled et al., 2002). The main difference between applications in climate modeling and weather forecasting is the purpose of downscaling. The primary objective in weather forecasting is a kind of maximization of the model performance on a short-term basis, e.g. a day-by-day basis, to describe the state of local surface variables (Karl et al., 1990). In climate modeling the downscaling applications focus on the long-term average and the long-term variability of the local surface variable to reproduce the climate characteristics at the local-scale as accurately as possible.

The techniques of dynamical downscaling applied in weather forecasting can be separated into deterministic and ensemble forecasting. In the next two sections a

brief description of both techniques is given. Afterwards, an overview about statistical forecasting is presented and a brief introduction of the different techniques used for the development of statistical models is given.

### 2.1.3 Dynamical Downscaling and Deterministic Forecasting

In dynamical downscaling the relationship between the large-scale and the local-scale information is described by process descriptions which are physically based. The typical application of dynamical downscaling is to nest a higher resolution circulation model within a GCM. The higher resolution circulation model is referred to as a limited area model (LAM) or regional circulation model, since the atmospheric circulation of a limited geographical area is simulated (Giorgi and Mearns, 1991). The GCM forecast is used as the driving force usually in form of time-varying boundary conditions to run the LAM.

Dynamical downscaling with a LAM allows for a better representation of the topography of the earth surface and enables a more sophisticated description of the sub-grid scale processes. The main advantage of dynamical downscaling is that consistent downscaled variables are provided (Bárdossy, 2000). Unfortunately, dynamic downscaling is computationally demanding and LAMs cannot be implemented on personal computers which hinder the wider use of this technique (Hewitson and Crane, 1996). The major shortcoming of dynamical downscaling is that often deterministic approaches are used providing only a single realization of the surface variable of interest with no estimation of the forecast uncertainty.

Today, several national meteorological services are running a LAM to provide spatiotemporal high-resolved weather forecasts for their region of interest. The MM5 (Fifth-Generation Meso-Scale Model) is probably the most-well known LAM, developed at the Pennsylvania State University and the US National Center for Atmospheric Research (NCAR, Grell et al., 1995). In the near future the MM5 will be replaced by the Weather Research and Forecasting Model (WRF, Skamarock et al., 2005). For Germany, the LAM of the Consortium for Small-Scale Modeling (COSMO) is the most relevant regional circulation model in weather forecasting, formerly known as the Lokal-Modell (LM(E)) of the DWD. A short description of the COSMO-model is given by Schättler et al. (2008) while a more extensive description can be found on the website of COSMO (www.cosmo-model.org). The model is well established in Germany and several regional flood forecasting centers uses the weather forecasts of COSMO to drive a hydrological model. The current version of the COSMO-model runs operational in a spatial resolution of 7 km. Since April 2007 a second version has became operational with a grid spacing of 2.8 km for shorter lead times (Schättler et al., 2008). However, even with the high resolution of the COSMO-model mismatches appear with the grid resolution of a distributed hydrological model. Unfortunately, the spatial resolution of the model is not the only limitation. The major drawback of the COSMO-model is that only a single realization of the predictand is forecast thereby ignoring the uncertainty of weather forecast. For example, in Figure 2.2 the precipitation forecast of the COSMO-model is compared to observed precipitation for several daily events which occurred in three medium-sized catchments located in the Rhine, Elbe and Danube basin in

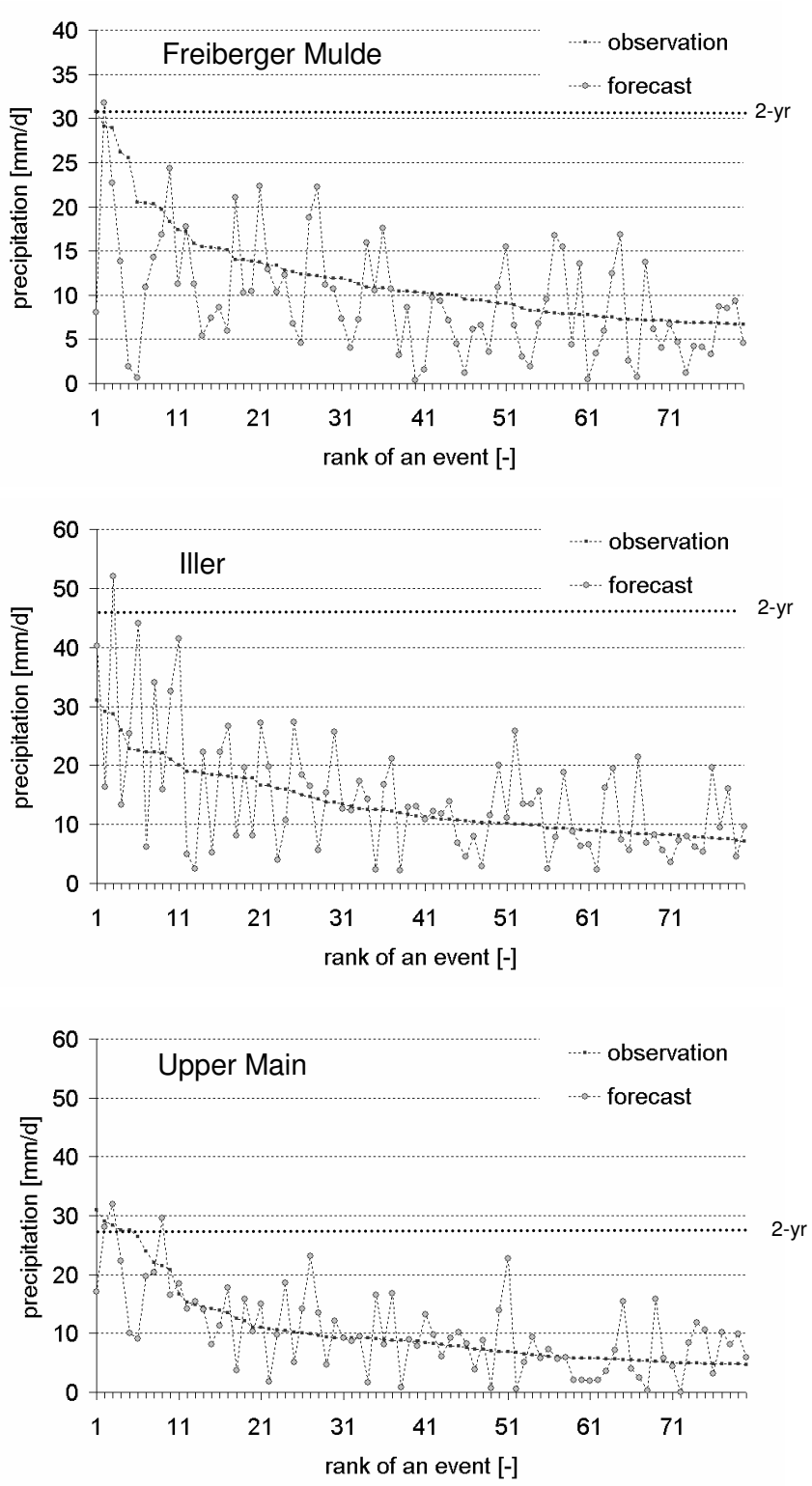



**Figure 2.2:** Comparison between precipitation forecast of the COSMO-model and observed precipitation for the 80 largest events (Freiberger Mulde - Elbe basin, Iller - Upper Danube basin and Upper Main - Rhine basin, Germany, $A_E$ = 2000 to 4500 km$^2$). Daily areal precipitation (6 UTC to 6 UTC of the next day) from April 2004 to December 2005, forecast cycle 0 UTC, lead time 6 fh to 30 fh. The thick dotted lines indicate the precipitation amount for an event with a return period of two years (2-yr).

Germany. For the comparison the daily areal precipitation of the forecast is taken which is the mean precipitation at the grid points of the COSMO-model situated in the catchment. It is compared to the observed daily areal precipitation. It is the average precipitation of a gridded data set (1 x 1 km) interpolated from more than 50 stations with daily measurements. Here, the daily areal precipitation is ordered according to the rank of an event where e.g. the largest event has the rank one. The comparison shows crucial differences between forecasts and observations which is a problem for all river basins. The COSMO-model often overestimates, but also underestimates the observed precipitation. A clear underestimation of the daily areal precipitation is for a hydrological warning system a big problem. In this case the warning system would give no alert and an important flood event would be missed. Please note that in this example the differences between forecasts and observations are presented for the largest daily observations during a short investigation period (April 2004 to December 2005). The thick dotted lines indicate the precipitation amount for an event with a return period of two years. Thus, most of the events have a low return period causing either no or only a moderate flood in the catchments. Unfortunately, the forecast uncertainty increases with the precipitation amount and the differences between model predictions and observations can be even higher for rare events than for those events illustrated in the example. The weakness of the deterministic precipitation forecast of the COSMO-model clearly illustrates that there is a need for the ability to quantify the forecast uncertainty of local-scale variables. To tackle this problem, ensemble prediction systems (EPS) have been developed in weather forecasting. They provide a number of realizations for the variable of interest.

### 2.1.4 Dynamical Downscaling and Ensemble Forecasting

The origin of ensemble forecasting goes back to the work of Edward N. Lorenz in the early sixties. He was probably the first meteorologist who studied the predictability of the atmospheric flow processes in more detail. Lorenz showed that small variations of the initial conditions of an ordinary differential equation system describing a convective process lead to substantially different states after a number of iterations (Lorenz, 1963). From his work it was concluded that the evolution of atmospheric flow processes depends not only on external factors like solar heating but also on internal factors. Slight variation of the internal factors can lead to different evolutions of the atmospheric flow processes. Due to the chaotic behavior of the atmosphere the forecast uncertainty of weather predictions systems grows with increasing lead time and the predictions are limited to a certain time range.

Based on the ideas of Lorenz, global EPSs have been developed for weather forecasting. Like deterministic systems the main components of a global EPS are a GCM and a global data assimilation system. Global ensemble forecasting is based on the assumption that a perfect GCM is used and that the state of the atmosphere cannot be exactly measured (Du et al., 1997). Due to the limitations in the observations, the initial conditions of a GCM are perturbed conditionally on the observational uncertainty. Multiple runs of a GCM are performed starting from the same time but from different initial conditions (Du et al., 1997). In contrast to a deterministic NWP model, a GCM of an EPS is usually run in higher spatiotemporal resolution

to reduce the computational demand.

There are a number of studies comparing global ensemble forecasting with deterministic approaches of the same model. Most of the studies showed that the skill of global EPS is significantly higher compared to a deterministic forecast (e.g. Richardson, 2000; Atger, 2001; Zhu et al., 2002). Zhu (2004) emphasized that the skill and value of global EPS is particularly high if the ensemble forecast is expressed in probabilities. They showed that probabilistic forecasts of the NCEP EPS for the geopotential height and precipitation performs better for all lead times compared to deterministic forecasts. Richardson (2000) and Atger (2001) did a similar work. Richardson (2000) determined the forecast value of the global EPS at ECMWF to predict precipitation. He illustrated that the economic value of a probabilistic EPS forecast for rare precipitation events is clearly higher than a deterministic forecast or the ensemble mean. Atger (2001) compared the precipitation forecast of the ECMWF EPS and NCEP EPS to the deterministic forecast of the same model. To verify the forecast he took 12-hourly observations for three winter months from the rain network in France. Even when the resolution of the global ensemble is lower, the ensemble forecast has a higher forecast value than the deterministic forecast. The work of Atger (2001) also indicated that the precipitation forecast of the global EPS at ECMWF are more skillful compared to the predictions of the NCEP EPS.

Probably, the major problem of ensemble forecasting is that it is computationally demanding. The researcher must find a good compromise between model resolution and ensemble size. This question was investigated by Mullen and Buizza (2002). They showed that the ensemble size is an important factor in predicting rare precipitation events. They concluded that a global EPS with a lower spatial resolution, though a larger ensemble size is more valuable than the same model with a higher resolution but a smaller ensemble.

In the past decade global EPSs have became well established in numerical weather prediction, providing useful information of the forecast uncertainty of meteorological variables for different geographical regions around the world. Nowadays, several national weather forecast centers run their own global EPS. A comparison of the predictions was performed by Buizza et al. (2005) for the global EPS at NCEP (Toth and Kalnay, 1993), ECMWF (Molteni et al., 1996) and at the Meteorological Service of Canada (MSC, Houtekamer et al., 1996). One of the main differences between the models is the ensemble size. ECMWF run their global EPS with 51 ensemble members which are more than twice as much compared to other ensemble models. Due to the different ensemble size, Buizza and co-workers selected a subset from each ensemble with the same size. They showed that the forecast skill differs only moderately between the models with the ECMWF EPS outperforming slightly the other models.

The aforementioned studies and techniques show that ensemble forecasting is well accepted in operational weather forecasting. Ensemble forecasting allows one to estimate the probability distribution of the predictand for future projections to identify weather states with low and high likelihoods (Zhu, 2004). The value and information content from a probabilistic forecast is clearly higher than from deterministic forecast, which is not surprising, because the user can select the optimal probability threshold for decision making (Atger, 2001). However, ensemble forecasts rarely contain more than 50 realizations and it is questionable if this small number can ap-

proximate the distribution of the forecast variable especially for longer lead times. A further drawback is that global ensemble predictions are biased like deterministic forecasts and the spatiotemporal resolution of their forecasts is often coarser compared to deterministic approaches (Zhu, 2004).

Due to these limitations, the techniques of dynamical downscaling are applied to translate the ensembles of large-scale variables to an ensemble of local surface variables. The common way of downscaling ensemble forecasts is to run a LAM which uses as boundary conditions the forecast of a global EPS. Each realization of the global EPS is transferred through the LAM to the local-scale. A second way of dynamical downscaling is to run a LAM with a deterministic forecast and to perturb the initial conditions of the regional circulation model to provide ensemble predictions. This approach is seldom applied, since most of the ensemble applications focus on medium-term predictions of large-scale atmospheric variables rather than on short-term applications for local surface variables (Du et al., 1997). A study pointing out this aspect is presented in Hamill and Colucci (1997).

An example of a regional EPS is the Limited Area Ensemble Prediction System of the COSMO (COSMO-LEPS, Montani et al., 2003) which takes the outcomes of the global ECMWF EPS to run the COSMO-model. In 2002, the COSMO-LEPS started operating to provide Central Europe with high-resolved ensemble weather forecasts (Marsigli et al., 2005). To save computational power, the ensemble size of the global ensemble is reduced to 16 members. Each realization of the ensemble subset is taken as initial and boundary conditions to run the LAM. COSMO-LEPS provides twice a day a forecast for the next five days. The temporal resolution of the forecast is three hours with a spatial resolution of around 10 km. A brief description and an application of COSMO-LEPS for the Elbe and Danube flood in August 2002 is presented in Marsigli et al. (2004). The skill of COSMO-LEPS in predicting intensive precipitation was investigated by Marsigli et al. (2005). They compared the predictions of the regional EPS with the ECMWF EPS and showed that the COSMO-LEPS can improve the detection of rare precipitation events.

An example of the precipitation forecast from the COSMO-LEPS is given in Figure 2.3. The forecast is compared to the largest observations which occurred from January 2007 to December 2007 in an alpine catchment located in the Upper Danube basin. The daily areal precipitation of the forecast is calculated for each ensemble realization. It is the mean precipitation at the grid points located in the river basin. To estimate the uncertainty of the precipitation forecast the maximum and the minimum value of the ensemble forecast is taken. The curves of the maximum and minimum value indicate a kind of a confidence interval of the forecast which encloses many observations. However, the interval also illustrates a large forecast uncertainty for daily areal precipitation. It is often argued that the forecast uncertainty of ensemble predictions must be reduced by decreasing the ensemble size. But even the selection of the maximum and minimum value does not prevent that some observations are outside the confidence interval, in particular for longer lead times and for the extremes. Furthermore, the example also illustrates that decision making based on the ensemble mean clearly underestimates the observations for longer lead times. In this case it is recommended to select a higher quantile for decision making to reduce the underestimation.

In Table 2.1 the model characteristics of the GFS, the ECMWF EPS, the COSMO-

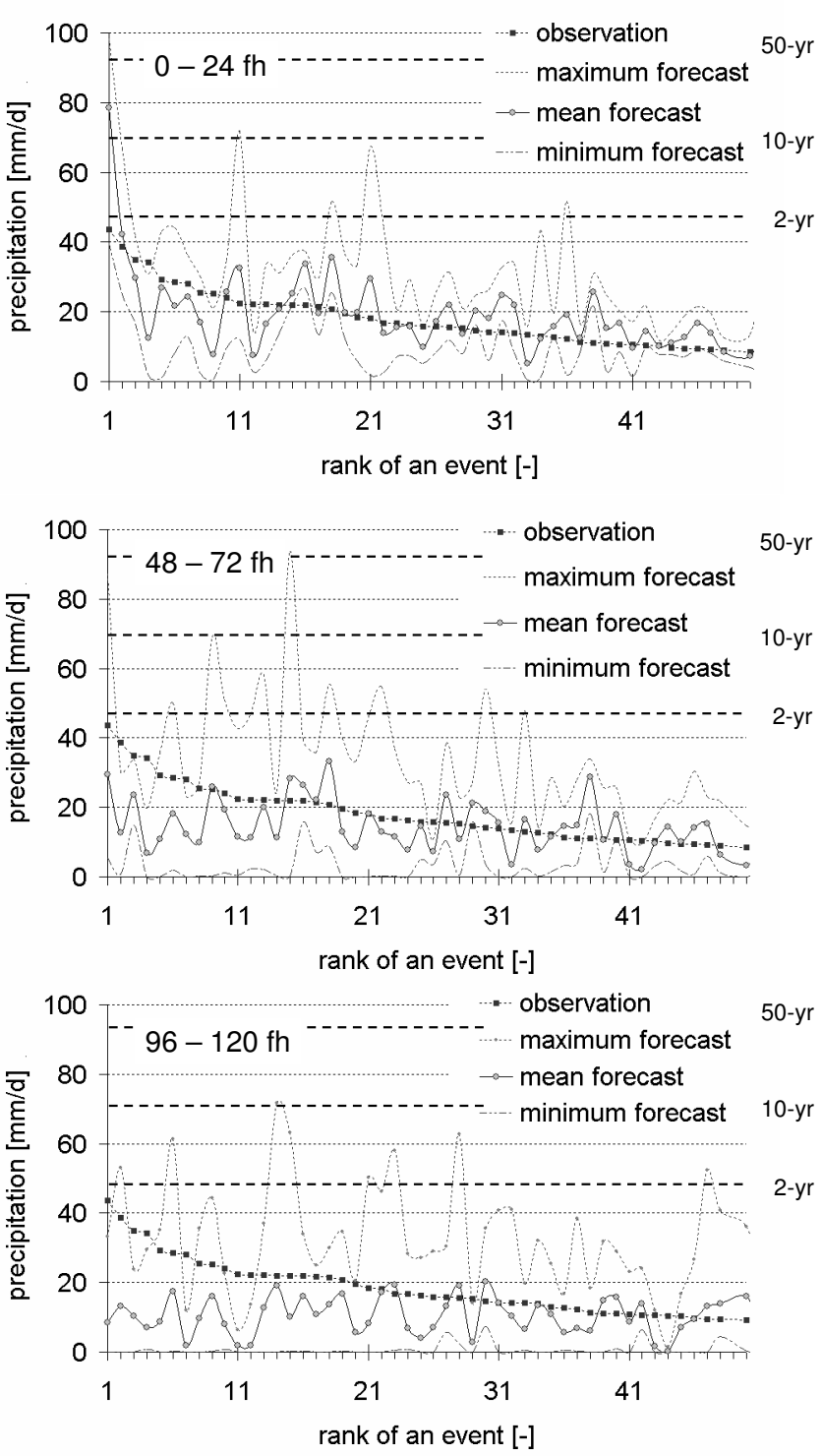



**Figure 2.3:** Forecast uncertainty of the COSMO-LEPS for the 50 largest observations and for three different lead times (0 - 24 fh, 48 - 72 fh and 90 fh - 120 fh), Iller catchment (Upper Danube basin, Germany, $A_E = 2177$ km$^2$). Daily areal precipitation (12 UTC to 12 UTC of the next day) from January 2007 to November 2007, forecast cycle 12 UTC. The thick dashed lines indicate the precipitation amount for event with a return period of two (2-yr), ten (10-yr) and fifty years (50-yr).

**Table 2.1:** Model properties of several NWP models and the distributed hydrological model LARSIM.

| property | GFS | EPS | COSMO -LM | COSMO -LEPS | LARSIM |
|---|---|---|---|---|---|
| institution | NCEP | ECMWF | COSMO | COSMO | LUWG |
| scale | global | global | regional | regional | local |
| spatial resol. [km] | 40 (80) | 40 | 7 (2.8) | 10 | 1 |
| temp. resol. [h] | 3 (12) | 6 | 3(1) | 3 | 1 |
| lead time [d] | 7 (16) | 10 | 3(1) | 5 | 3 (7) |
| forecast cycles [1/d] | 4 | 2 | 3 | 2 | 2 |
| realizations [-] | 1 | 51 | 1 | 16 | 1 |

LEPS and the COSMO-model are listed. Their characteristics are compared to the requirements of the distributed hydrological model LARSIM (Large Area Runoff Simulation Model; Bremicker, 2000) run at the LUWG. The operational mode of the LARSIM is commonly used at the regional flood forecasting centers in Germany (see e.g. Hangen-Brodersen et al., 2008). Even the high spatiotemporal resolution of the COSMO-model and the COSMO-LEPS does not fit the resolution of a distributed hydrological model. However, the spatiotemporal resolution of the LARSIM is usually an intuitive choice of the user which must be not necessarily the best choice. Obled et al. (2004) mentioned that the optimum spatiotemporal resolution of a flood forecasting model can vary and depends on catchment properties such as the response time. The same applies also for the number of realizations needed to describe the forecast uncertainty. The use of ensemble forecasts is usually restricted to a constant and to a small number of ensemble members. As the example in Figure 2.3 demonstrates, it is questionable whether the limited number of realizations can approximate the forecast uncertainty of the surface variable in particular for longer lead times.

The aforementioned examples illustrate a clear need to develop alternative downscaling techniques to describe the distribution of local surface variables with a variable number of realizations for different lead times. These downscaling techniques should be flexible and easy to handle in order to fit the requirements of flood forecasting models and to provide reliable and unbiased weather forecast with a high quality and value. In the next sections alternative forecasting techniques based on statistical approaches are proposed.

### 2.1.5 Statistical Downscaling and Forecasting

In statistical downscaling the observed relationship between the large-scale and a local-scale information is described through a transfer function (Hewitson and Crane, 1996). This transfer function is a statistical model which usually translates anomalies of the large-scale atmospheric flow to anomalies of the local-scale surface variable (Zorita and von Storch, 1997).

After Hewitson and Crane (1996) statistical downscaling is based on following

assumptions:

1. The local-scale surface variable is a function of the large-scale atmospheric flow.

2. The large-scale atmospheric flow is reliably simulated by a GCM.

3. The functional relationship remains valid for other time periods than the calibration period.

The assumptions made by the authors are not only valid for statistical downscaling. They must be also fulfilled in dynamical downscaling to produce reliable predictions for the local-scale surface variable. The latter assumption is the most difficult to reconcile for applications in climate change modeling, since the functional relationship established in the past must be also valid for a future climate with a doubled-CO2 concentration (Wetterhall, 2005). Fortunately, in weather forecasting the projections are only made for the next few days or weeks in advance which enables to test the third assumption directly.

During the last decades a variety of statistical downscaling approaches have been developed. A comprehensive list of case studies for downscaling precipitation and temperature is given in Wetterhall (2005). A classification for the techniques used in statistical downscaling has been proposed by several authors (see e.g. Zorita and von Storch, 1997; Wilby and Wigley, 1997). Bárdossy (2000) distinguishes three main techniques of downscaling:

- **Regression methods**: These approaches are the most frequently used techniques. Regression methods establish a linear or nonlinear relationship between the predictor and predictand (Wilby and Wigley, 1997). The choice of reliable predictors and the regression functions are crucial steps. Since the form of the function is often not known, a trial and error procedure is performed to identify a representative function. Afterwards, a screening procedure is used to scan through an archive of physical reasonable predictors and to integrate step by step the predictors into the equations system (see e.g. Wilks, 2006). Artificial neural networks as well as canonical correlation belong to this group (see Zorita and von Storch, 1997; Wilby and Wigley, 1997). Brunet et al. (1988) presents an example of a regression method used for the prediction of precipitation probabilities in weather forecasting.

- **Resampling**: Resampling techniques, like the analog method, identifies a subset of past weather states similar to the current one. Depending on the degree of similarity a probability can be assigned to each state of the subset (Bárdossy, 2000). In contrast to regression methods, only a portion of the data is selected for the forecast day to provide a prediction for the variable of interest. An example of analog forecasting for temperature is given in Kruizinga and Murphy (1983) and for daily areal precipitation in Obled et al. (2002).

- **Conditional (probability) approaches**: The techniques of conditional downscaling apply an intermediate step to classify the large-scale atmospheric flow

information into circulation (weather) patterns. The model parameters of regression methods and resampling techniques determined are then conditioned on the circulation pattern type. A number of studies use autoregressive techniques to perform a conditional simulation. In a temperate climate zone local-scale variables have a seasonal variation. To account for the intra-annual variability conditional approaches are also created by classifying the large-scale flow information into subcategories based on the season.

Each of the techniques used in statistical downscaling have advantages as well as limitations. The performance of the downscaling approaches depends on many factors so that the selection of a particular downscaling technique for a specific purpose is not straightforward. Therefore, it is quite common to work with several approaches to point out a suitable one for a certain downscaling objective.

There are several features of statistical downscaling which made their use in operational weather forecasting very attractive:

1. The techniques of statistical downscaling can be used to provide quick solutions. They are easy to implement on personal computers so consequently a greater number of researchers can use these techniques (Hewitson and Crane, 1996).

2. Statistical downscaling delivers by-products (Zorita and von Storch, 1997). It can be applied to verify and to reconstruct historical data, to fill data gaps and to identify erroneous measurements. It can be also used to validate the output of GCMs or LAMs. Statistical downscaling techniques are crucial for the development of reliable downscaling approaches in weather forecasting as well as in climate modeling.

3. Local surface variables are often influenced by local effects (e.g. topography, exposition). In contrast to dynamical approaches the local properties of the predictand are integrated to the equation system during the development.

4. Statistical downscaling enables the researcher to focus on certain areas of the distribution function of the predictand. An objective function can be created to maximize e.g. the forecast performance for extremes.

5. Probably, the major advantage of statistical downscaling is that the main task of the downscaling process can be fulfilled due to the computational efficiency. A large number of realizations can be generated to approximate the distribution of the predictand for different lead times.

In the literature it is sometimes mentioned that a detailed knowledge of physical processes is not needed for statistical downscaling. Nonetheless, an understanding of physical process is crucial and the process knowledge should be included in the model development to create reliable approaches (Wetterhall, 2005). Physical processes of the atmosphere are non-linear and a downscaling approach describing a process with a linear relationship should be carefully used. This model might reasonably predict the mean of the variable of interest, but it can fail for extremes. Thus, to set up a robust functional relationship long and homogeneous data for both the predictand

and the predictor are needed. Another shortcoming is the discrete character of the forecasting procedure, if conditional approaches are selected for the downscaling.

Before a statistical downscaling technique is implemented for the operational phase, the model predictions are tested for the past. To determine the model quality, the differences between the predicted and the observed value of the surface variable of interest is calculated. The validation with past data presents a crucial step in downscaling and there are usually different techniques how to perform a validation. The common way is split-sampling which divides the data of the investigation period in two subsamples. The first subsample is used to determine a reasonable predictor and to calibrate the model. The second subset is taken to test the model transferability with independent data by running the model with the same settings defined during the calibration. If the calibration and validation procedure is repeated several times, a more robust validation form can be created. This kind of techniques is called cross-validation. However, a validation form based on past data only determines the model skill for the past. The "true" performance of the methodology can only be quantified if the model has been implemented operationally.

Depending on the kind of data used for the model development, two different techniques are distinguished in weather forecasting (Wilks, 2006):

- **Perfect Prog**: In perfect prog(nosis) (PP) observed predictors are taken to develop the functional relationship (Klein et al., 1959) between predictor and predictand. The predictors are mostly derived from a gridded climatological data set describing the large-scale atmospheric flow processes over the area of interest. In PP long and homogeneous time series can be used for the model development which is an advantage for forecast models designed for extremes. However, in the implementation phase the forecast of global NWP model is used. PP relies on the assumption that the predictions of global NWP models have no systematic error.

- **Model Output Statistics**: A second technique is called Model Output Statistics (MOS, Glahn and Lowry (1972)). MOS takes the outcomes of a global NWP model to establish the statistical equations which allow to incorporate the bias of a NWP model. But for the development of the equation systems a long and homogeneous data set is rarely available, since operational NWP model are continuously enhanced and updated several times per year. After any significant change of the NWP model the statistical relationship must be redeveloped. Model changes are often too frequent so that a robust equation system cannot be established and development costs often become too high (Wilson and Vallée, 2002).

The techniques of MOS and PP are well established in operational weather forecasting and a number of national weather services perform statistical forecasts with MOS and PP. A list of national weather services using MOS approaches is given in Wilks (2006).

A third technique uses reanalysis data for the model development. Reanalysis data is the outcome of a GCM which has been used to simulate the large-scale atmospheric circulation for the past. These data are mainly taken for the development of downscaling approaches in climate modeling, but it can be also used in weather

forecasting (see Bontron and Obled, 2003). The use of reanalysis has the advantage that long and homogeneous time series for various large-scale predictor sets are available based on the information reproduced by a GCM. The model development with reanalysis data enables the incorporation of a number of past weather states leading to extremes at the local-scale. It allows also one to consider the bias of global NWP during the model calibration. Even so, the GCMs taken for the preparation of the reanalysis often use slightly different parameterization and discretization schemes compared to the GCMs used in operational applications. Downscaling with reanalysis data assumes that both GCMs produce the same model bias.

The techniques of statistical downscaling used at the national weather services are mainly designed to perform prediction for the local surface variable for a given point. Due to the stochastic character of precipitation, the performance of the forecast models predicting point precipitation at a daily time scale or even in a higher temporal resolution is often poor. To overcome that problem, the mean precipitation, for example, of an area of interest can be taken as predictand. Such a methodology has been proposed by Obled et al. (2002), who developed an analog forecasting algorithm for the prediction of daily areal precipitation for a river basin of interest. To meet finally the requirements of a hydrological prediction model, the daily areal precipitation is disaggregated into hourly precipitation realizations (see Obled et al., 2004) to reflect the mean precipitation characteristics of a catchment.

## 2.2 Forecast Verification

In forecast verification the correspondence between model predictions and observations is measured to assess the forecast performance of a prediction model for the variable of interest. It is based on the assumptions that the observations can represent the real world processes in an adequate manner. The closer the difference between model predictions and observations, the higher the performance of a prediction model is and the more appropriate the forecast of the real world processes are.

One of the earliest attempts was made by J. P. Finley in 1885 who measured the quality of an experimental tornado forecast (see Murphy, 1996; Jolliffe and Stephenson, 2003). The work well illustrates that the selection of an appropriate verification measure is a crucial factor in weather forecasting. The deficiencies of his method activated a broad discussion in this discipline and led to several researchers proposing more suitable performance measures (Murphy, 1996). Today, a large number of performance measures are used in forecast verification. A good overview is given in Wilks (2006). The book of Jolliffe and Stephenson (2003) probably presents the most comprehensive work, including also an outlook on present and future directions in forecast verification.

There is a common consensus in the research community that a single verification measure cannot be preferred since each measure has its strengths and also its weaknesses. To find an appropriate measure, the forecaster must select one or a combination of two or even more criteria depending on the purpose of the forecast model. In this study a verification measure must be selected which can quantify the performance of a probability forecast since the forecast uncertainty is expressed in

probabilities. Additionally, the presented forecast model is also a component of a warning system. If the forecast exceeds a certain warning level, a decision maker must uses model predictions as source of information to decide between an alarm for protective action or no alarm. Thus, in this investigation a verification measure is also needed which can quantify the performance of a warning system.

In the following sections some common verification measures are proposed which can be used to evaluate the performance of a probability forecast (Section 2.2.3) and the forecast of a warning system (Section 2.2.4). But at first, the following question will be addressed in more detail:

### 2.2.1 What is a Good Forecast?

This question was actually proposed and discussed by Murphy (1993). He defines three types of goodness in weather forecasting:

- *Consistency*: Type I goodness. It is defined as the correspondence between forecasters' judgment and their forecasts.
- *Quality*: Type II goodness. It is the correspondence between observation and forecast.
- *Value*: Type III goodness. A forecast has a value provided that the user gains a benefit from it.

All goodness types are essential in forecasting. The type I goodness, forecast consistency, is the most difficult type to measure, since the forecaster's (or decision maker's) judgments is based on a model forecast; their empirical knowledge and all further information are internal statements which are not available for evaluation. Only the external statements made by the forecaster in written and verbal form can be verified by the observations provided, of course, that these statements are listed consecutively (Murphy, 1993). Since forecast consistency is quite difficult to quantify, any verification technique to measure this type of goodness is not proposed in this chapter. For a more detailed discussion the reader is referred to Murphy (1993). In this study we assume that there is no inconsistency between the forecaster's judgments and their external statements.

The type II goodness, forecast quality, is the most well known goodness and has received a great deal of attention so that many performance measures have been developed to quantify this goodness. Murphy (1993) noted that several attributes can be used to describe this type of goodness. The following four are probably the most common attributes. Here, we follow the description of Wilks (2006):

- *Accuracy*: It defines the mean differences between observation and forecast. Needless to say, the higher the accuracy, the lower the differences between forecast and observations.
- *Bias*: It defines the differences between the average observation and the average forecast. If the average of the forecast is higher than the average observation, the forecast model overestimates the predictand. The bias is also called unconditional bias.

- *Resolution*: It is defined as the capability of the forecast model to distinguish between situations under which an event occurs in the future or not. For an illustration of this attribute we split up the rainfall forecast into five categories: *"no rain"*, *"light"*, *"moderate"*, *"intensive"* and *"very intensive rain"*. Yesterday, *"no rain"* and the day before yesterday *"very intensive rain"* was observed. If the same categories were predicted by the model, a forecast would have perfect resolution. The forecast has no resolution if the forecast is *"moderate rain"* for both days.

- *Reliability*: A reliable forecast provides unbiased estimates of the predictand with different forecast probabilities. To illustrate this attribute, we consider the daily forecast of rainfall probabilities for two categories *"no rain"* and *"rain"* for a given forecast period such as 10 years. Then, we determine e.g. the average observed frequencies of a rainy day for all days with a forecast probability of 10 %. A forecast is perfectly reliable for those days if the average observed frequency is also 10 %, which means that on 10 % of those days a rain event was observed.

- *Sharpness*: It is an important attribute for measuring the forecast uncertainty of probabilistic forecast. The smaller the sharpness, the higher the forecast uncertainty of a probability forecast. The probability forecasts of two prediction models can have the same accuracy but different sharpness. In the event of this, this prediction model should be selected that has the higher sharpness.

The aforementioned forecast attributes seem to be crucial for the development of a model. This is because the model predictions are related to the observations and each attribute presents particular information about the forecast quality which cannot be given by another attribute. The development of an overall measure that quantifies all attributes is certainly possible. However, it can be useful to determine the attributes step by step to investigate where the weakness and strength of the model prediction is. This is probably one of the main reasons why more than one performance measure is needed to describe the quality of a prediction model.

Note that the selection of the forecast attributes also depends on the purpose of the model. For example, bias is an attribute which may indicate a systematic error of the forecast. Usually, a forecast should be as accurate as possible and any systematic error should be avoided. However, a perfect forecast is not possible. There are always differences between forecast and observations leading to either an over- or an underestimation of the predictand. In the case of a warning system, a slight underestimation has different consequences than a slight overestimation of the predictand. Unfortunately, the traditional measures of bias cannot describe this asymmetry.

The forecast value, the type III goodness, is probably the most crucial type of goodness especially when the outcomes of the prediction model are used in warning situations for decision making which can affect the benefit of the end users. In Section 2.2.5 the idea of this type of goodness is highlighted in detail and a simple method is proposed for the description of a decision making process. Furthermore, it is illustrated that this approach can describe the aforementioned problem and the asymmetry between the negative outcomes of a warning system. To quantify the

accuracy of a probability and a binary forecast, suitable performance measures are proposed in Sections 2.2.3 and 2.2.4. Furthermore, in Section 2.2.3 a measure that determines resolution and reliability of a probability forecast is presented.

Before presenting appropriate performance measures for the evaluation of the prediction model developed in this study, the basic idea of a skill score and a skillful forecast is explained.

### 2.2.2 Skill Scores

A high degree of similarity between forecast and observation indicates also a high degree of forecast accuracy. But a high accuracy must not correspond necessarily to a skillful forecast. To decide whether a forecast has skill or not, skill scores have been developed in forecast verification. They are relative measures of a verification measure comparing the forecast against a low-skill reference forecast. A skill score $SS$ is defined in the following manner (Potts, 2003):

$$SS = \frac{S - S_0}{S_1 - S_0} \tag{2.2}$$

where $S$ is the value of the performance measure based on the forecast, $S_1$ is the value of the measure of the perfect forecast and $S_0$ is the value of the performance measure of the reference forecast. A positive skill score indicates an improvement compared to the reference forecast. If $S$ is equal to $S_1$, the forecast is perfect and the forecast skill of a model is 1. If the forecast skill is negative, the forecast is worse compared to the reference forecast.

At least three different types of low-skill reference forecast can be distinguished (see e.g. Potts, 2003):

- **Random forecast**: Forecasts are derived randomly from the marginal distribution of the observation.
- **Persistence forecast**: The observed value of the last measurement is taken for the forecast.
- **Climatological forecast**: The mean value of the observation is used for the forecast.

By using Equation 2.2 anyone can create their own skill score. However, it is important to select an appropriate baseline to create a useful skill score. The selection of a reference forecast depends on factors like the time scale of the predictand (e.g. daily or weekly), the predictand itself (e.g. precipitation or discharge) and the lead time of the prediction model (e.g. short-term or medium-term). The selection of the climatological forecast is not a good reference, e.g. if a short-term warning system of intensive rainfall is developed. In this case, the mean of the observations will always underestimate an extreme and no warning will be given by the forecast model. A random forecast gives sparse warnings but mostly to the wrong time. Predicting intensive rainfall with persistence is more suitable than a climatological or a random forecast, in particular for days with a similar weather pattern. Although, the longer

the lead time of the forecast model is, the poorer the persistence forecast is. For longer projections the climatological forecast is usually selected as reference.

The selection of a reference forecast is not only restricted to a low-skill reference forecast. The forecast of a model can be also used as baseline if the purpose of a study is the comparison of two or more models.

### 2.2.3 Probability Forecast

Forecast probabilities are mostly expressed in form of the exceedance probability $f_p(t)$ at a given forecast time $t$. The exceedance probability is the probability $\mathrm{P}(y(t))$ that the predictand $y(t)$ will exceed a certain event threshold $y_t$ defined by the forecaster. The exceedance probability is the complementary of the cumulative distribution function $\mathrm{F}(y(t))$ with $f_p(t) = 1{-}\mathrm{F}(y(t)) = 1{-}\mathrm{P}(y(t)) \leq y_t)$. For example, to define a wet day for a certain location in Germany, a forecaster selects a precipitation amount of 1 mm/d as threshold. The climatological occurrence frequency of this event is 30 %. A forecast probability of 90 % ($f_p(t) = 0.9$) for this event would indicate a high likelihood for the occurrence of a wet day since the forecast probabilities are three times larger than the climatological frequency. While a forecast probability of 10 % ($f_p(t) = 0.1$) would indicate the contrary, a low likelihood for a wet day.

To measure the accuracy of the probability forecast, the correspondence between a time series of the forecast probabilities and a time series of binary observations is calculated. The binary observations are derived by comparing the realizations of the predictand with the event threshold. If the predictand is larger than the threshold value, the event occurs and the observation $o_b\,(t)$ is 1:

$$o_b\,(t) = 1, \quad for\ y\,(t) > y_t \tag{2.3}$$

If the realization of the predictand is smaller or equal to the event threshold, no event is observed and the observation is zero:

$$o_b\,(t) = 0, \quad for\ y\,(t) \leq y_t \tag{2.4}$$

The accuracy of a probability forecast is measured by calculating the mean square difference between both time series. This performance measure is called Brier score (Wilks, 2006):

$$BS = \frac{1}{n}\sum_{t=1}^{n}\left[f_p\,(t) - o_b\,(t)\right]^2 \tag{2.5}$$

where $n$ is the number of joint pairs of observation and forecast used for the comparison. The values of the $BS$ ranges between zero and 1 since the observations and forecasts are standardized to an interval of zero and 1. The smaller the value of $BS$ is, the smaller are the differences between observations and forecasts. A $BS$ of zero indicates a perfect forecast. The original form of the BS proposed by Brier (1950) is different to the performance measure defined in Equation 2.5, though is commonly named Brier score in forecast verification (for a brief discussion see Wilks, 2006).

The $BS$ is a function of the event threshold. The variation of this threshold enables the forecaster to examine the skill of a forecast model over the entire range

of the predictand. For instance, if a precipitation threshold of 0.1 mm/d is selected, the probability forecast is evaluated whether it is able to separate among dry and wet days. To focus on extremes, a precipitation threshold which corresponds to a low return period of an event must be selected by the forecaster.

The Brier score can be decomposed in two terms describing reliability and resolution of a probability forecast. It illustrates a clear relationship between these forecast attributes (Murphy, 1973):

$$BS = \frac{1}{n}\sum_{i=1}^{k} n_i \left(f_{p,i} - \bar{o}_i\right)^2 - \frac{1}{n}\sum_{i=1}^{k} n_i \left(\bar{o}_i - \bar{o}\right)^2 + \bar{o}\left(1 - \bar{o}\right) \tag{2.6}$$

The first term of the equation describes reliability, the second term resolution and the third term uncertainty. In the case of the reliability term, forecasts and observations are sorted in dependence to the forecast probabilities $f_{p,i}$ into $k$ classes and for a class $i$ the arithmetic mean of the binary observation $\bar{o}_i$ is calculated:

$$\bar{o}_i = \frac{1}{n_i}\sum_{i=1}^{n_i} o_{b,i} \tag{2.7}$$

which is equal to the conditional relative frequency. $n_i$ is the number of joint pairs of observations and forecasts for a class. $\bar{o}$ is the mean of the binary observations of the entire sample:

$$\bar{o} = \frac{1}{n}\sum_{t=1}^{n} o_b\left(t\right) \tag{2.8}$$

which is equal to the unconditional relative frequency of the event selected for the evaluation. The reliability term presents a measure of the conditional bias. If the reliability is zero, the conditional observed frequencies are equal to the forecast probabilities and the forecast does not contain any conditional bias (perfect reliable). The resolution describes the ability of the forecast to separate between events with different frequencies. For each class the conditional relative frequency is compared with the climatological frequency. The larger the differences are, the better the resolution of the forecast is. The uncertainty term is not affected by the forecast since it depends only on the observed frequencies.

Forecast probabilities can be also issued for event categories (e.g. *"no rain"*, *"light rain"*, *"moderate rain"* and *"intensive rain"*) if the forecaster is interested in more than two outcomes. To measure the forecast accuracy for a categorical probability forecast, the ranked probability score *RPS* is usually selected (Epstein, 1969). This score is equal to the arithmetic mean of a Brier score for $k$ selected event thresholds if for both scores the same boundaries are taken to define the event categories (Toth et al., 2003):

$$RPS = \frac{1}{k}\sum_{i=k}^{k} BS_k \tag{2.9}$$

RPS summarizes the forecast accuracy of a probability forecast over the entire range

of the predictand if the corresponding event thresholds are selected. Thus, the value of RPS strongly depends on the selection of the thresholds.

To compare the quality of a probability forecast against a reference forecast the Brier skill score $BSS$ (Wilks, 2006):

$$BSS = \frac{BS - BS_0}{BS_1 - BS_0} = 1 - \frac{BS}{BS_0} \tag{2.10}$$

and the ranked probability skill score $RPSS$ are usually taken:

$$RPSS = \frac{RPS - RPS_0}{RPS_1 - RPS_0} = 1 - \frac{RPS}{RPS_0} \tag{2.11}$$

The scores of the perfect forecast $BS_1$ and $RPS_1$ are zero which enables a reduction of the skill score. $BS_0$ and $RPS_0$ represent the scores of a reference forecast which is the climatological average forecast (here the unconditional relative frequency of an event). $BSS$ and $RPSS$ range between minus infinity and 1. A model improvement is indicated if the score is larger than 1. The forecast is perfect if the skill score is 1.

It is also possible to create relative measures for reliability and resolution after dividing each term of Equation 2.6 by the uncertainty term (Toth et al., 2003):

$$B_{rel} = \frac{\frac{1}{n}\sum_{i=1}^{k} n_i \left(f_{p,i} - \bar{o}_i\right)^2}{\bar{o}\left(1 - \bar{o}\right)} \tag{2.12}$$

where $B_{rel}$ is the relative measure of reliability. To get a negatively oriented measure for the resolution, its relative measure $B_{res}$ is subtracted from one:

$$B_{res} = 1 - \frac{\frac{1}{n}\sum_{i=1}^{k} n_i \left(\bar{o}_i - \bar{o}\right)^2}{\bar{o}\left(1 - \bar{o}\right)} \tag{2.13}$$

Both measures range between 0 and 1. If both scores are zero, a probability forecast is perfect. They are related to the Brier skill score in the following way (Toth et al., 2003):

$$BSS = 1 - B_{rel} - B_{res} \tag{2.14}$$

The scores and skill scores presented in this section are the most common verification scores in weather forecasting used to quantify the quality (accuracy, reliability and resolution) of a probability forecast. However, in decision making situations a probability forecast is converted to a non-probability forecast (e.g. “no warning” and “warning”). In the next section this conversion is illustrated and some verification measures are introduced to estimate the forecast accuracy of a non-probability binary forecast.

### 2.2.4 Binary Forecast

A binary forecast presents the simplest type of a non-probability forecast. To create a time series of binary forecasts from a time series of probability forecast, the fore-

**Table 2.2:** Possible outcomes of a binary forecast. yes = event is forecast (observed), no = no event is forecast (observed). After Mason (2003).

| Forecast | Observations | |
|---|---|---|
| | Yes | No |
| Yes | Hit a | False alarm b |
| No | Miss c | Correct rejection d |

caster must select a probability (decision) threshold $p_t$. If the forecast probability exceeds this threshold value, the event is forecast and the binary forecast $f_b(t) = 1$. If the forecast probability is smaller or equal to the decision threshold, no event is forecast and $f_b(t) = 0$. Depending on the choice of the decision threshold the forecaster can choose between a warning system which gives alarms less or more frequently.

The following example may illustrate the conversion of the probability forecast and should demonstrate the importance of the decision threshold. A forecaster defines a precipitation event which is larger than 40 mm/d as the warning level. The event has a return interval of one year (one day in 365 days) which corresponds to a relative frequency of $s_l = 1/365 \approx 0.003$. The forecaster defines as decision threshold the exceedance probability $p_t = 0.3$ which is a hundred times larger than the climatological risk of the event. If the forecast probabilities exceed this threshold, 30 % of the precipitation realizations are larger than 40 mm/d which indicates a high likelihood for intensive rainfall. But it means also that a warning is only given if at least 30 % of the realizations are larger than the warning level. If the forecaster selects a smaller threshold e.g. $p_t = 0.03$ (ten times larger than the climatological risk), then only 3 % of the precipitation realizations are needed to trigger a warning.

To estimate the accuracy of a binary forecast, the time series of binary forecasts is compared to a time series of binary observations. The conversion of the observation to a binary time series is analog to the transformation shown in Equation 2.3 and Equation 2.4.

The outcomes of a binary forecast are two positive outcomes, a hit and a correct rejection, and two negative outcomes, a false alarm and a miss (see Table 2.2). The meaning of a hit and a false alarm should be familiar while the meaning of the other outcomes may not be so intuitive. A miss is an observed event which has not been forecast. If no event is forecast and also no event is observed, the result is called a correct rejection.

An accuracy measure of a binary forecast is the hit rate $HR$:

$$HR = \frac{a}{a+c} \tag{2.15}$$

where $a$ is the number of hits and $c$ is the number of misses. The hit rate is the ratio of events which are correctly forecast.

A forecaster is usually interested in maximizing the detection of events. If the warning system gives more alarms, the hit rate can rise but it is also possible that the number of false alarms $b$ increases. Thus, the accuracy of a binary forecast

cannot only be described by the hit rate and a second measure, the false alarm rate $FA$, must be taken into account:

$$FA = \frac{b}{b+d} \tag{2.16}$$

where $d$ is the number of correct rejection.

A simple skill score based on both accuracy measures is the Peirce's skill score $PSS = HR - FA$ (Peirce, 1884). If $HR$ is larger than $FA$ the binary forecast is skillful and $PSS$ is greater than zero. The binary forecast is perfect if PSS equal to 1 with $HR = 1$ and $FA = 0$. The hit rate can be also plotted in dependence of the false alarm rate for different event thresholds which is called relative operating characteristic curve (ROC-Curve) (see e.g. Mason, 2003). Zero skill is indicated by the diagonal of the diagram ($HR = FA$).

The skill of a binary forecast can be also expressed in terms of the Heidke skill score $HSS$ (Heidke, 1926):

$$HSS = \frac{PC - PC_r}{1 - PC_r} \tag{2.17}$$

It is based on the proportion correct $PC$ which describes the ratio of the correct forecasts:

$$PC = \frac{a+d}{n} \tag{2.18}$$

$PC_r$ is the proportion correct of a random forecast and $n$ is the number of joint pairs of observations and predictions. The Heidke skill score ranges from -1 to 1. A positive score indicates that the forecast has a higher accuracy predicting binary events than a random forecast. A score of 1 means a perfect forecast.

The $HSS$ and the $PSS$ represents common skill scores for the evaluation of a binary forecast. However, both measures are not appropriate for the verification of a binary warning system. In decision making situations both negative outcomes of a binary warning systems (a false alarm and a miss) lead to different consequences. Usually, the potential damages of a miss are higher compared to a false alarm. However, this asymmetric relationship between both negative outcomes is not accounted in the $HSS$ and the $PSS$. There are a number of further binary performance measures (see for an comprehensive overview Mason, 2003) and also modifications of the Heidke skill score (see von Storch and Zwiers, 1999). But none of these scores can consider the asymmetric relationship between a false alarm and a miss. To overcome that problem, performance measures based on a cost-loss approach must be selected. The basic idea of these approaches is presented in the next section.

### 2.2.5 Cost-Loss Situation and Forecast Value

The outcomes of weather prediction models are used as a source of information for decision making. In warning situations, the decision maker must choose between different courses of action such as "*protective action*" or "*no action*". The various

**Table 2.3:** Expense matrix of a binary warning system. The four outcomes (hit, false alarm, miss and correct rejection) of a binary forecast are related to a cost C and a loss L. After Murphy (1977).

| Action | Observations | |
|---|---|---|
| | Adverse weather | No adverse weather |
| Protect | Cost C | Cost C |
| Do not protect | Loss L | 0 |

ways can lead to different consequences but a decision maker does not know which way is the best, since the outcome of the event is not known. In dialog with decision makers therefore the following questions often arise:

- What is the best course of action?
- How can we quantify the different consequences of the decision making process?
- How valuable are weather forecasts?

A further problem is the forecast technique. A number of decision makers favor a deterministic forecast since the handling of a single forecast is straightforward in warning situations. If the realization of the variable of interest exceeds a given warning level, the decision maker can use this information to trigger an alarm or not. But many decision makers are aware that the outcomes of weather prediction models are uncertain and that this uncertainty should be reflected in the forecast. Unfortunately, decision making in the face of uncertainty is more difficult compared to a deterministic forecast because more information is provided due to the number of realizations describing the uncertainty for the variable of interest. Thus, the following questions should be also addressed:

- How to handle a probability forecast?
- What is the optimal use of a probability forecast, especially in warning situations?

A third problem in decision making are the various users which are located in the region of interest of a decision maker. The users are mainly interested in profiting from the weather predictions. The benefit to them does not only depend on the value of the decisions, but also on the user itself. For instance, if the warning system triggers an alarm, the user takes protective action which is associated with a cost. However, the amount of cost needed for protection depends on the user itself, since his needed level for protection might be higher or lower compared to that of the other users. Thus, the following questions should also be also taken into account when developing a warning system:

- How can we distinguish between users?
- And, can we quantify their benefit?

To address these questions, several cost-loss approaches have been introduced in weather forecasting. The simplest approach is presented in Thompson and Brier (1955) which is based on a method originally proposed by Anders Ångström in 1922 (see also Liljas and Murphy (1994). The approach describes the cost-loss situation of a two action, two state decision making problem (Murphy, 1977). It is based on the assumption that the decision maker is risk neutral and that a decision is only influenced by the weather forecast. The decision maker can decide between: "no action" and "protective action". Both decisions are associated with two possible states of weather: "adverse weather" and "no adverse weather". The four outcomes (hit, false alarm, miss and correct rejection) of a binary warning system have different consequences in decision making situations (see Table 2.3). If the decision maker gives an alarm, protective action is taken causing a cost $C$ but suffering eventually the loss $L$. If the decision maker gives no alarm, the expense depends on the actual weather state. There are no costs as long as no adverse weather occurs.

An individual user of a weather forecast can be defined by its cost-loss ratio $\alpha = C/L$. For an illustration of the cost-loss ratio, we consider the following example: If a flood warning system is designed for city protection, there are at least two individual users: (i) the local authority which placed the order for the design and (ii) the citizens. The local authority would like to use the warning system to save the life of the citizens and to reduce the monetary losses. To specify their cost-loss ratio, a decision maker must define all potential consequences of protection or flooding. The definition of all consequences is not straightforward, since monetary criteria and non-monetary criteria, such as the number of potential persons at risk or the socio-ecological impacts, must be summarized which are often expressed by imprecise statements such as "*high*" or "*low*". In this example we assume that the monetary losses and the persons at risk due to flooding are a hundred times larger than the expected costs of an alarm (e.g. 10 million euro compared to 100.000 euro or 10.000 person at risk compared to 100 person). The cost-loss ratio of the local authority is then 1:100 ($\alpha = 0.01$).

The cost-loss ratio of the citizens is different to that of the local authority. Each citizen can also determine his individual cost-loss ratio based on monetary values like in the previous example. However, the flood prediction model is mainly designed to raise warnings for the public. In this case, it is more important to specify the cost-loss ratio of the citizens in terms of their willingness for protection. If a warning system triggers too many false alarms, the citizens will be less inclined to take protective action again. It is probably true that not more than four or five false alarms are needed to reduce significantly their willingness for protection. In the event of this, we can assume that the cost-loss ratio of a citizen has an odds of 1:5 ($\alpha = 0.2$). Please note that the exact value of the cost-loss ratio (e.g. $\alpha = 0.2043$) is not the point of interest, since in real world situations the specification of all monetary and non-monetary criteria is quite difficult and only approximations can be made for a criterion which are associated with a high uncertainty. In this case, it is more interesting to define the range of possible realizations of the cost-loss ratio for a single user.

The cost-loss ratio describes also the asymmetric relationship between the consequences of a false alarm and a miss, as discussed in the previous section. For a warning system the value of $\alpha$ ranges between 0 and 1 with $C < L$. This means that

a false alarm is less relevant than a miss. However, it is often argued that a false alarm is as essential as a miss or even more important. However, this argument is not true for a warning system. If a false alarm is more relevant than a miss, the costs for protection are higher than the losses and no protection (and also no warning system) is needed!

Due to the relevance of the cost-loss approaches, they have been used in a number of studies to determine the expected value of weather prediction models (see e.g. Kolb and Rapp, 1962; Katz et al., 1982; Brown et al., 1986). A detailed overview about some early applications in weather forecasting is given in Winkler and Murphy (1985). Several years ago, Richardson (2000) determined the forecast value of the global EPS of ECMWF for the prediction of intensive precipitation by comparing the ensemble predictions with deterministic forecasts of the same model. He clearly demonstrated the superiority of probabilistic forecasts. Roulin and Vannitsem (2005) used the cost-loss approach to determine the value of the predictions of a hydrological model which provides an ensemble forecast for daily discharge.

A good introduction into cost-loss approaches and the determination of the forecast value is given by Murphy (1977) and Richardson (2003). In Section 4.3 of Chapter 4 a skill score is proposed for the simplest form of a cost-loss approach. Furthermore, an objective function is presented which maximizes the forecast value for a given set of end users. It is also demonstrated in this chapter how a decision maker should apply a probability forecast in cost-loss situations.

# 3 Data and Study Areas

The presented study is performed in eight small river basins located in Germany (Figure 3.1). The Fränkische Saale, the Upper Main, the Freiberger Mulde and the Iller are the study areas of the research projects HORIX and PREVIEW. The Nahe, the Prüm and the Upper Sieg are the test sites which has been provided by the regional flood forecasting center of Environmental State Agency Rhineland-Palatinate. The Ruhr is selected due to a strong research interest.

The study areas are sub-catchments of the Rhine, the Elbe and the Danube which are three of the four main rivers in Germany. They are typical mountainous and alpine tributaries with high relief energy and a short response time during floods. In the last two decades, several large floods occurred in this part of Germany such as the Neckar flood in 1993, the Elbe and Danube flood in 2002 and the Northern Alp flood in 2005. River flooding in Northern Germany is less problematic since the northern part is dominated by lowlands and rivers with a longer response time.

**Table 3.1:** Climatological and hydrological characteristics of the four study regions. Climate zones after Peel et al. (2007), mt = maritime temperate, wsc = warm summer continental. Additionally, the coordinates of the nearest grid point of the reanalysis information is given for each study region.

| | Western Germany | Central Germany | Eastern Germany | Southern Germany |
|---|---|---|---|---|
| abbreviation | WG | MG | EG | SG |
| project | HORIX,LUWG | HORIX | HORIX | PREVIEW |
| catchment | Prüm, Upper Sieg, Ruhr, Nahe | F. Saale Upper Main | F. Mulde | Iller |
| river basin | Rhine | Rhine | Elbe | Danube |
| landscape | mountainous | mountainous | mountainous | alpine |
| climate zone | mt | mt/wsc | wsc | wsc |
| latitude | 50.0° N | 50.0° N | 50.0° N | 47.5° N |
| longitude | 7.5° S | 10.0° S | 12.5° S | 7.5° S |

The crosses in Figure 3.1 mark the grid points of the reanalysis data used in this investigation. They illustrate the coarse resolution of the outcomes of a GCM used for the reconstruction of the past climate data. The figure also shows that the catchments have different grid points as neighbors which make a separate definition of the model parameters of the statistical approach necessary. Thus, the eight catchments are classified according to their location into four study regions (Table 3.1). The moderate differences between the climatological and hydrological characteristics of the catchments are a further reason for their subdivision. The study region located

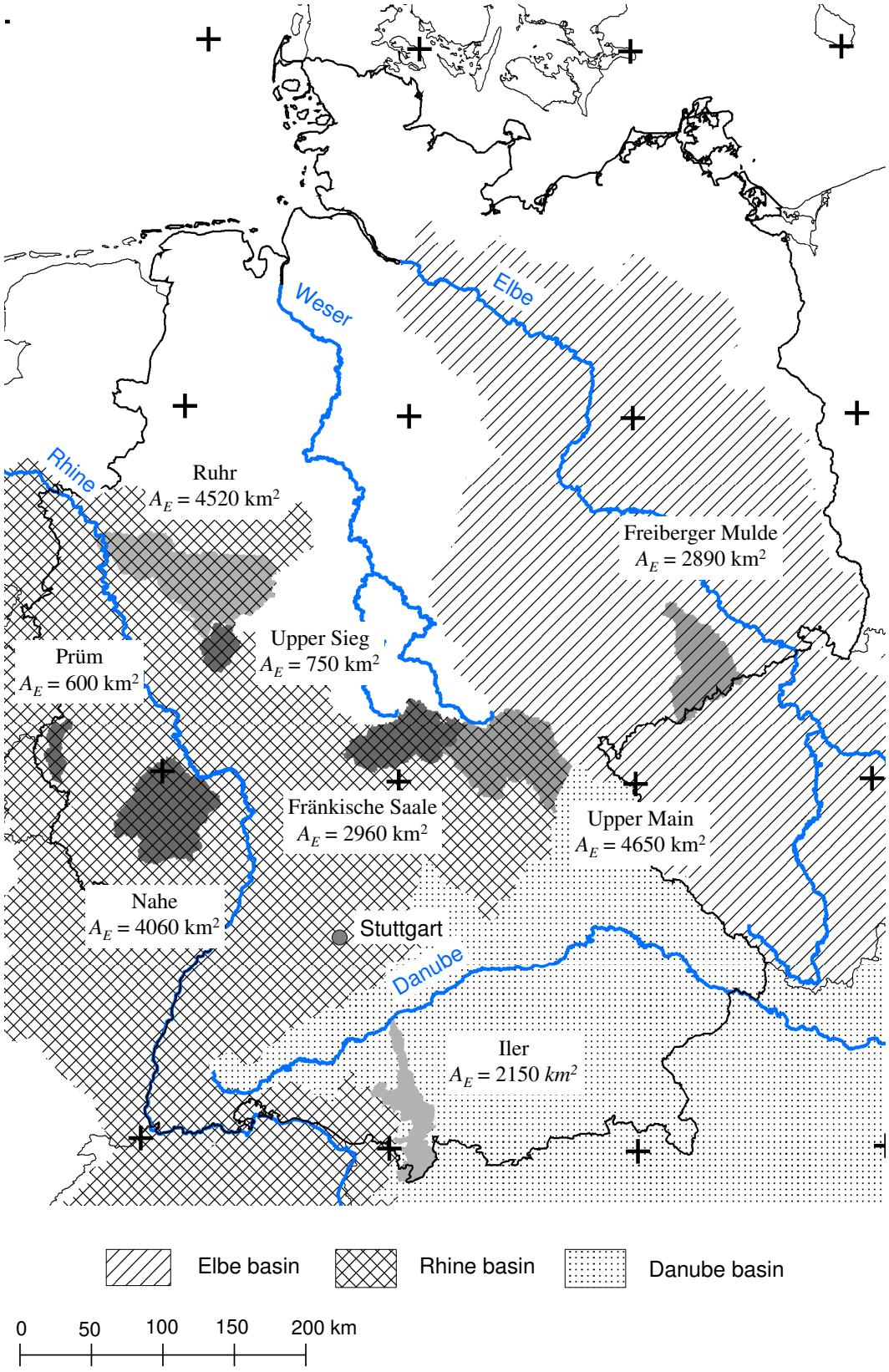



**Figure 3.1:** Location and catchment size $A_E$ of the study areas. The crosses mark the grid points of the reanalysis data used in this investigation.

**Table 3.2:** Properties of four representative catchments.

| catchment characteristic | Nahe | F. Saale | Iller | F. Mulde |
|---|---|---|---|---|
| region | WG | MG | SG | EG |
| gauging station | Grolsheim | Königshofen | Kempten | Erlln |
| catchment size $A_E$ [km$^2$] | 4060 | 2960 | 2150 | 2890 |
| river length [km] | 116 | 135 | 147 | 124 |
| lowest point [m.a.s.l] | 80 | 90 | 132 | 470 |
| highest point [m.a.s.l] | 816 | 950 | 2645 | 1214 |
| altitude difference [m.a.s.l] | 736 | 860 | 2513 | 744 |
| mean discharge [m$^3$/s] | 30.6 | 19.7 | 47.2 | 22.8 |
| mean ann. prec. [mm] | 858 | 788 | 1541 | 843 |
| 10-year areal prec. [mm/d] | 42 | 38 | 70 | 65 |
| ref. period discharge | 1958 - 2001 | 1958 - 2001 | 1958 - 2001 | 1958 - 2001 |
| ref. period precip. | 1958 - 2001 | 1979 - 2003 | 1958 - 2001 | 1963 - 2001 |

in Western Germany (WG) consists of four Rhine subcatchments (Nahe, Prüm, Upper Sieg and Ruhr). The region in the middle of Germany (MG) is composed of the Upper Main and its neighboring river, the Fränkische Saale, both located in the Rhine basin, too. The remaining two catchments, the Freiberger Mulde and the Iller, form their own class (EG and SG). They are located in the Elbe and Danube basin in Eastern and Southern Germany.

In the next section a brief description of the study regions is given. It is analyzed when intensive precipitation and floods occur and examined what the causes of these extremes are. Furthermore, a short description of the predictand, the daily areal precipitation, and the predictor information used in this investigation is given.

## 3.1 Description of Study Regions

### Catchment Characteristics

Four representative rivers (the Nahe, the Fränkische Saale, the Freiberger Mulde and the Iller) are selected to illustrate the climatological and hydrological characteristics of the four regions. An assortment of their catchment properties is listed in Table 3.2. The Nahe is situated in southwestern Germany. It is one of the smaller tributaries of the Rhine running north-eastwards among two low mountain ranges, the Hunsrück and the Pfälzer Bergland. The catchment of the Fränkische Saale is also situated in the Rhine basin around 200 km north-eastern to the Nahe. After the union of their headstreams, the Fränkische Saale flows south-westwards along a low mountain range, the Rhön, to meet the Main, one of the major tributaries of the Rhine. The catchment of the Freiberger Mulde is located in the Elbe basin in the East of Germany. The Freiberger Mulde springs in the Ore Mountains in Czech Republic and crosses the basin in a northward direction to meet the Zwickauer Mulde. The Zwickauer and the Freiberger Mulde represent the two headstreams of the Mulde

which is a tributary of the Elbe. The Iller is one of the smaller alpine tributaries of the Upper Danube located in the south. The source of the Iller is in the Bavarian Alps next to the Austria border.

The Nahe, the Fränkische Saale and the Freiberger Mulde are typical catchments of the low mountain ranges in Germany. Their altitude varies from 100 m.a.s.l at the catchment outlet to more than 800 m.a.s.l at the highest point. The higher altitudes of the catchments are covered by coniferous forest while the low lands are intensively used by agriculture. Large settlement areas are usually situated along the lower reaches as one would expect. The Iller catchment area is more heterogeneous compared to the mountainous rivers. The upper reaches of the Iller are characteristic for an alpine catchment with a strong change in vegetation zones. The mountain ranges of the upper reaches often exceed the tree line and their peaks are mostly covered by snow throughout the entire year. Between the tree line and the snow is a small transition zone of grassland which is extensively used for agriculture. The altitude of the upper reaches varies from 800 to 2645 m.a.s.l. Due to the large altitude difference the upper reaches have a high relief energy leading to a short response time ($<$ 6 h) during a flood event. The lower reaches of the Iller belong to the Alpine foothills. A change of plane and hilly areas with agricultural land and coniferous forest are characteristic for this region. The altitude of the Alpine foothills varies between 100 and 800 m.a.s.l though the response time of the catchments located in this region is similar to the mountainous catchments.

The size of all catchments range between 650 and 4500 km$^2$ (see also Figure 3.1) and thus correspond to the group of meso-scale catchments ($A_E$ = 100 - 10000 km$^2$).

### Precipitation and Runoff Regime

The study areas in the western part of Germany are situated in the maritime temperate climate zone which is characteristic for this region (Peel et al., 2007). The temperature regime of this climate zone has a clear seasonality with mild winters and warm summers. The precipitation regime of the catchments located in Western Germany is characterized by a weak seasonality with a slight maximum in December (Figure 3.2). The peak is the result of a sequence of storms which occur more frequently in autumn and at the beginning of winter due to the higher influence of the Prevailing Westerlies. In the east and the south of Germany, the maritime influence is weaker leading to a higher variability of the temperature regime with colder winters and warmer summers. The precipitation regime of the Freiberger Mulde and the Iller is marked by a summer maximum showing the higher continental character in these regions. The summer maximum is mainly caused by convective precipitation which is the dominant rainfall type in summer, especially in the continental part of Germany. The precipitation regimes of the Freiberger Mulde and the Iller are both characterized by a second moderate maximum in November and December indicating that the South and the East are also affected by autumn storms. Furthermore, the annual areal precipitation of the Iller is twice as large as the annual precipitation of the mountainous catchment (see Table 3.2) which is characteristic for the alpine catchments in Germany. The high areal precipitation in the alpine regions is mainly the result of orographic rainfall when warm and moist air masses are forced in the mountains. The precipitation regime is also characterized by a

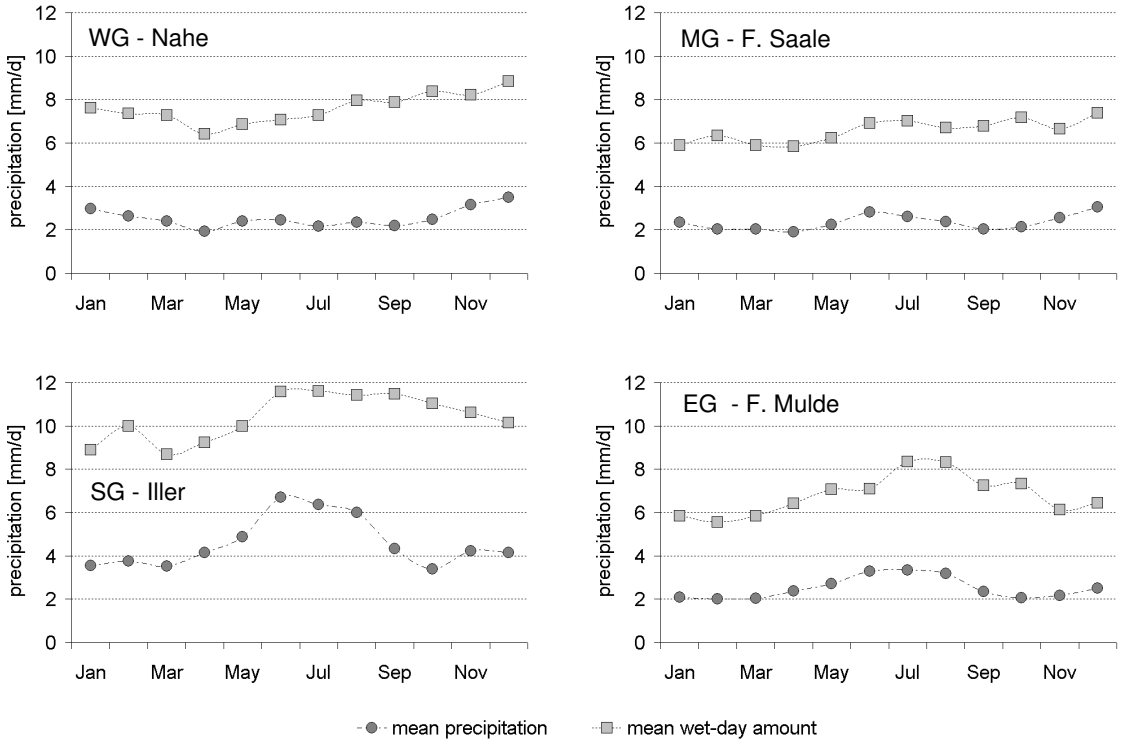



**Figure 3.2:** Precipitation regime of the representative catchments which are located in four different regions: WG = Western Germany, MG = Central Germany, EG = Eastern Germany and SG = Southern Germany. The precipitation regime is illustrated by the mean daily areal precipitation and the mean areal precipitation of a wet-day calculated for each month from daily observations of the reference period. In this investigation, all days with an areal precipitation of at least 2 mm are defined as a wet-day.

slight minimum in winter. This minimum is caused by high pressure systems which occur over Central Europe more frequently in winter and reducing the impact of the Westerlies. However, the winter minimum is not only a climatological effect. It can also be an artifact due to measurement errors. During winter, precipitation falls mostly as snow in the higher reaches of the mountainous and alpine catchments in Germany and ordinary pluviometers might underestimate the precipitation amount in those areas. The study region in the middle of Germany is located in the transition zone of both climate zones. Consequently, the precipitation regime of the catchments situated in this region is a mixture of the precipitation regimes of both climate zones.

The runoff regimes of the mountainous catchments are quite similar despite the differences in the precipitation regimes (Figure 3.3). They are characterized by a moderate seasonality with a maximum at the end of the winter which is mainly a by-product of snowmelt. The discharge regime of the alpine catchment has a clear maximum in May which is characteristic for an alpine flow regime with no glacial melt. The high discharges are caused by the gradual increase of the temperature at the end of spring and in the early summer so that the snow in the higher reaches of

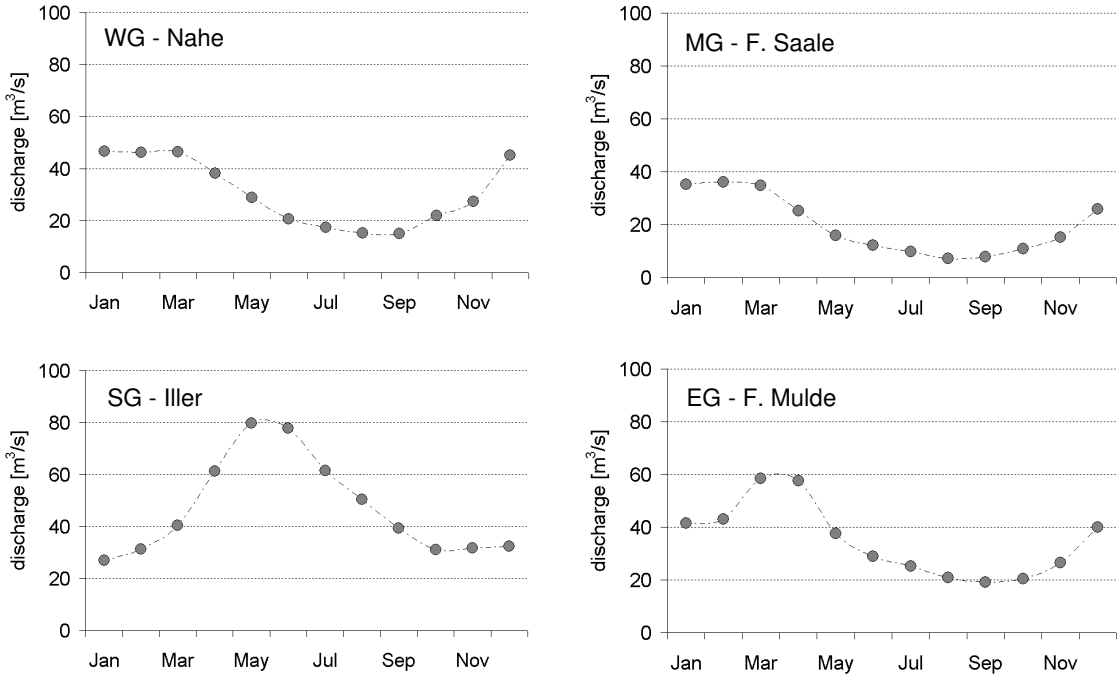



**Figure 3.3:** Discharge regimes of the representative catchments which are located in four study regions: WG = Western Germany, MG = Central Germany, EG = Eastern Germany and SG = Southern Germany. The discharge regime is illustrated by the mean daily discharge which has been calculated from daily observations for each month of the reference period. The observations are from a gauging station located at the catchment outlet. The reference period for the discharge regime is listed in Table 3.2.

the catchment is melted.

### Estimation of Daily Areal Precipitation

To ensure an appropriate downscaling, the time series of an ideal predictand should be long and homogeneous (Obled et al., 2002). Both properties are essential for the model development. However, the first property is in particular important for downscaling approaches designed for the prediction of extremes. Weather states leading to extremes at the local scale are unusual and occur seldom. If a short investigation period is selected, only a limited number of those situations are available for the training and validation. A long time series is also crucial for the development of a resampling algorithm. Resampling techniques selects a portion from all records to provide a forecast. If the subset of weather states cannot adequately approximate the predicted weather state, an accurate prediction of the local surface fluxes is less likely.

The second property, a homogeneous time series, means that in the case of daily areal precipitation the time series of the observed point precipitation as well as the techniques used for the estimation of the daily areal precipitation must be consis-

tent. Precipitation measurements are seldom homogeneous, since rain gauges are renewed from time to time by switching e.g. the measurement type (tipping bucket or weighing rain gauge) or by modifying the kind of registration (automatic, non-automatic). This type of information is usually not available making it more difficult to detect systematic shifts which are needed to remove inconsistencies in the time series.

The selection of a consistent interpolation technique for the estimation of the daily areal precipitation is actually straightforward. But the quality of the estimation is strongly affected by the goodness of the measurement technique. For instance, the measurement uncertainty of rain gauges is process-dependent. It is usually higher for snow than for rainfall. Thus, the quality of the estimation of daily areal precipitation is also process-dependent and can vary during the year. A further variable influencing the quality of the estimation of daily areal precipitation is the number of observations used for the estimation. Usually, the number of rain gauges varies during the observation period. It is often the case that fewer observations are available at the beginning of an investigation period.

In Germany, several institutions operate their own precipitation networks measuring precipitation at daily, hourly or even at higher temporal resolutions. The German Weather Service (DWD) has the densest network covering the entire country, providing daily observations for a number of stations for more than 60 years. Precipitation in a higher temporal resolution than daily measurements is also available, but usually the network density is clearly coarser and the time series are seldom longer than 15 years. This is one of the reasons why the daily time scale has been selected for this study. It is also relevant for the model development to know the exact observation interval of daily precipitation to avoid a temporal offset of the forecast. At the DWD the daily precipitation is usually measured from 6 to 6 UTC of the following day. However, the measurement interval of daily precipitation is not unique and it depends on the rain network as well as on measurement year. Before April 2001, daily precipitation had been observed from 6:30 to 6:30 UTC. Afterwards, the observation interval was changed to 6 to 6 UTC to fit the standard measurement interval of the World Meteorological Organization (WMO). But the change was only performed for automatic station. For non-automatic stations, daily precipitation is still measured from 6:30 to 6:30 UTC. Fortunately, the temporal offset of half an hour is less problematic for the prediction of daily precipitation.

Several techniques can be used for the estimation of daily areal precipitation for a river basin. The simplest technique selects for a given day the observed daily precipitation from the stations located in the catchment and calculates the mean value from this sample. This kind of technique is only recommended if many stations are available covering the entire catchment homogeneously. However, the precipitations stations are mostly irregularly distributed over the catchment. Therefore, an interpolation technique is chosen which transfers the point information to a regular grid. Afterwards, the mean precipitation at the grid points located in the catchment is calculated.

In this investigation external drift kriging (Ahmed and De Marsily, 1987) has been selected for the interpolation of daily areal precipitation. This technique belongs to the group of geostatistical techniques developed by G. Matheron and D. G. Krige around 50 years ago. Compared to simpler techniques like inverse distance weight-

ing, kriging has the advantage that the spatial dependence of the target variable is determined for the definition of the interpolation weights. The annex "external drift" comes from a second variable used as additional information. In this investigation the altitude is chosen as the external information since the precipitation amount in mountainous and alpine catchment usually rises with the elevation due to orographic rainfall. The rain network of the German Weather Service is denser in the mountainous regions than in the lowlands of Germany, so that more than 60 stations are available for the estimation throughout the entire observation period. The precipitation stations are mostly located in the catchment, but also stations outside the catchment are taken for the interpolation. Data gaps and erroneous data are marked as missing values. A quadratic grid with a resolution of one kilometer is selected and the precipitation is estimated for each grid point.

The interpolation by external drift kriging is only performed for the catchments located in EG and MG. The daily information for the catchments in WG is extracted from a gridded data set covering the Rhine basin (Hundecha Hirpa, 2005). This data set has been also interpolated by external drift kriging. The data set is available from 1958 to 2001 with a spatial resolution of 5 x 5 km. The daily areal precipitation for the catchment located in SG is calculated from REGNIE (Regionalisierung von Niederschlagshöhen), a gridded data set provided by the German Weather Service. REGNIE uses local factors like the height and the exposition as additional information for the interpolation. The spatial resolution of this data set is approximately 1 x 1 km. The precipitation data is available for entire Germany from 1951 to the current year. In this study the gridded data has been provided for a period from 1958 to 2001. REGNIE and the data set provided by Hundecha Hirpa (2005) are also based on daily measurements of the precipitation network operated by the DWD.

### Interannual Distribution of Precipitation Extremes and Floods

The interannual distribution of daily precipitation extremes is determined for the 80 largest daily peak values. The peak values are selected from a daily time series of areal precipitation which ranges from 1958 to 2001. To secure that the extremes are time-independent, only those daily peak values separated more than two weeks from another are selected. In Figure 3.4 the interannual distribution of the daily extremes is shown for the four catchments. Intensive precipitation can occur throughout the entire year, but there are certain seasons with a slightly higher risk. In Western Germany, intensive precipitation is more frequent in autumn and at the beginning of winter, when the influence of the Westerlies is strongest. In the east and south of Germany, intensive precipitation is more dominant in summer. Around 40 % of the extremes occur in the Freiberger Mulde from June to August. The summer extremes are either the result of local convective processes like thunderstorms due to an instable atmosphere or they are caused by low pressure systems coming from the North Atlantic or the Mediterranean Sea. Especially, low pressure systems coming from the Mediterranean Sea usually transport warm and moist air masses to Central Europe. They often caused severe floods in the past in southern and eastern Germany like the Elbe and Danube flood in August 2002 (for an event description see Philipp and Jacobeit, 2003) or the Northern Alp flood 2005.

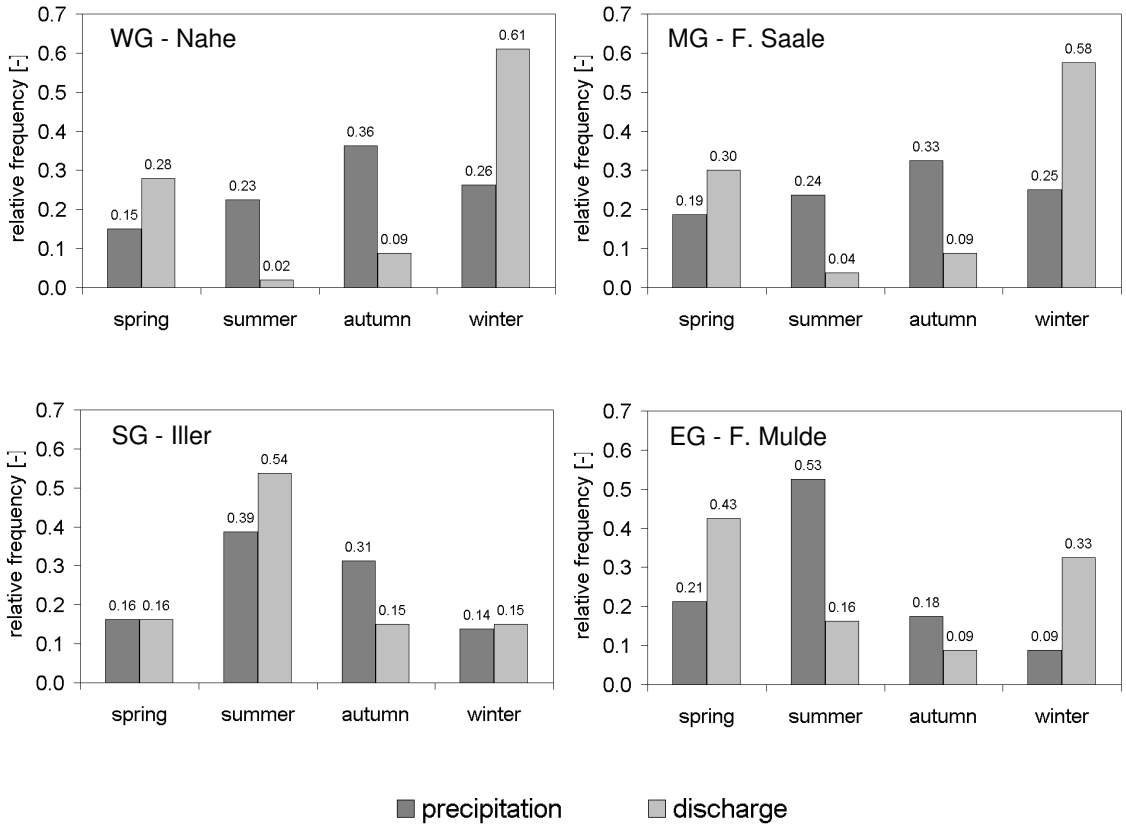



**Figure 3.4:** Seasonal regime of heavy preciptation and floods in the four study regions. Spring = March to May, summer = June to August, autumn = September to November and winter = December to February.

To define the flood regimes of the four catchments, the daily mean discharge of a gauging station at the catchment outlet is taken for the selection of the peak values. One exception is made for the Iller, since the runoff regime in the lower reaches of the Iller is strongly affected by hydropower production. For this catchment the extremes are selected from a time series of a station located in the upper reaches. Note that the investigation period for the flood analysis only partially overlaps with the investigation period of heavy precipitation and varies from catchment to catchment (see Table 3.2). The flood regime of the Freiberger Mulde and the Fränkische Saale indicates that spring floods are dominant in the mountainous regions of Germany. At the end of winter the runoff regime is higher in the mountainous catchment due to the snow melt. Intensive precipitation can activate the snow melt processes in the upper reaches, so that a combination of both factors can amplify the impact of a flood. Summer floods are not so frequent in the mountainous regions. In summer the soils are dry so that the effect of heavy precipitation is lessened. The flood risk of the Nahe and Fränkische Saale even tends to zero during June and August. This result is interesting since it could be said that a flood warning system is not important for the mountainous catchment in summer. However, we should be very careful with this conclusion since the flood regimes presented in this investigation are based on many extremes which cause little or no damages. The flood regimes of extremes with a higher damage potential (e.g. flood events with a return period

of 10 years or even longer) is not know. This regime can be different compared to the regimes presented here. For example, the flood regime of the Freiberger Mulde shows that the flood risk is highest in winter and in spring. However, the biggest floods of the last two decades (the Odra flood in July 1997 and the Elbe and Danube flood in August 2002) occurred in summer among many other big floods (see e.g. the description of the Vltava flood in the beginning of the Chapter 1). To determine the regime of those extremes, a clearly longer investigation period than 50 years is needed so that many extremes can be used for determining a suitable flood regime.

The flood regime of the alpine catchments has a peak in summer which is characteristic for the Upper Danube tributaries of the Northern Alps. More than 50 % of the extremes occur in the Iller catchment in summer. They are caused through intensive precipitation which is slightly amplified by the high base flow due to snow melt during this season. Throughout the other seasons, the risk of flooding is considerably lower since precipitation falls as snow in the higher reaches of the Alps and it is not available for runoff generation.

## 3.2 Description of Predictor Information

The selection of one or more reliable predictors is probably the most important step in downscaling. After Obled et al. (2002) and Wetterhall et al. (2005) an ideal predictor should fulfill the following properties:

- A predictor should be physical and conceptually reasonable and should be strongly correlated with the predictand.
- A time series of a predictor should be long, homogeneous and available for the development and the operational period.
- The predictor should be measured and simulated reliably.
- A long predictor time series is crucial for the development of analogue forecasting algorithm to incorporate a sufficient number of analogues for the model evaluation.

The selection of an appropriate predictor is probably the most important step in statistical downscaling. The choice of a predictor depends on factors like the predictand, the season and the purpose of downscaling. A list of predictors used for precipitation downscaling is given in Wilby and Wigley (2000). Wetterhall (2005) also provides a comprehensive list specifying downscaling techniques and predictors for precipitation and temperature. In many studies pressure related variables like the mean sea level pressure or the geopotential height are preferred. They can be also taken to calculate air flow indices like gradients or vorticity (e.g. Jones et al., 1993; Wilby and Wigley, 2000). Besides pressure related variables, humidity related variables like the specific humidity or the relative humidity have been also selected. For instance, Bontron and Obled (2003) used height gradients and relative humidity for analogue forecasting. Matulla et al. (2007) selected the specific humidity and the mean sea level pressure. Bontron and Obled (2003) mentioned that humidity related predictors are not as reliably forecast by a GCM compared to pressure related

variables in particular for longer lead times. They propose that multivariate predictors based on pressure and humidity related variables should be only selected for short-term forecasts rather than for longer lead times. Probably, the main problem of GCMs for an adequate description of humidity related variables is their coarse resolution. In particular, humidity related variables need a higher resolution due to their higher spatial variability than pressure related variables.

In this study the following predictors are mainly chosen:

- **Mean sea level pressure / geopotential height**: Both variables give suitable information of the actual state of the atmosphere. A pressure map shows the location of low and high pressure systems. Adverse weather is more likely when the pressure is low. High values are often linked to a high pressure system indicating good weather. The mean sea level pressure is an appropriate indicator for the separation of wet and dry days. The geopotential height has similar characteristics like the mean sea level pressure.

- **Moisture flux**: It describes the amount of moisture transported to the area of interest. In comparison to the geopotential height the moisture flux can give additional information of the precipitation amount of an event. Intensive precipitation is more likely when high moisture fluxes occur. The moisture flux can be also divided in its vector components. Due to the Westerlies humid westerly fluxes predominate in Central Europe. Zonal moisture fluxes should contain more information than meridional moisture fluxes.

- **Relative humidity**: It describes the actual saturation degree of the atmosphere. High moisture fluxes must not lead to intensive precipitation at all times, in particular, if the relative humidity is low over the area of interest. The relative humidity is a further indicator for wet or dry days especially in summer when the atmosphere can store more water due to the high temperatures.

The mean sea level pressure (MSLP) used in this investigation is from the Research Data Archive (RDA) which is maintained by the Computational and Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR). The original data are available from the RDA (http://dss.ucar.edu). It is a gridded daily data set (5.0 x 5.0°) which is available for the Northern Hemisphere from 1899 to the current date. The grid information is based on surface observations which were not interpolated directly from a single data source of daily point observations. It has been assembled from the grids of various meteorological chart digitization projects and operational analysis. For more detailed information see the information on the website of RDA.

The other predictors are derived from the reanalysis data set of the NCEP-NCAR archive (Kalnay et al., 1996). It provides a number of variables describing the state of the atmosphere from the earth surface to the tropospause for the entire globe. The reanalysis data is available as six hourly data at 0, 6, 12 and 18 UTC and as daily and monthly averages from 1948 to present. The spatial resolution of the data set is 2.5 x 2.5°. Atmospheric variables are provided for the surface and for 17 pressure levels. In this study the following six-hourly predictors are selected: the

geopotential height (GPH), the u-wind component (UWND), the relative (RHUM) and the specific humidity (SHUM). The information is taken from four pressure levels (1000, 850, 700 and 500 hPa). To derive the zonal moisture flux (UFLX), the specific humidity is multiplied by the u-wind component. Furthermore, the geopotential height is taken to calculate the zonal flow velocity (UGPH), the meridional flow velocity (VGPH) and the resulting geostrophic flow strengths (TGPH).

Besides NCEP-NCAR, ECMWF also provides a reanalysis data set called ERA 40 (Uppala et al., 2005) with a range from 1957 to 2002. A comparison of both archives has been recently performed by Ben Daoud et al. (2009). They selected the relative humidity field at the 850 hPa level and showed that ERA 40 contained a number of erroneous data over Europe.

The use of reanalysis data offers a number of advantages for the development of downscaling applications in weather forecasting (see for a short discussion also Section 2.1.5). However, the large-scale atmospheric flow simulations of GCMs can be characterized by systematic errors, so that the outcomes of a GCM can be marked by a spatial offset. This can lead to the problem that a predictor taken from the grid point next to the target area is not categorically the best choice. The location of the optimum grid point for the downscaling is usually not known. The same effect can also apply for the pressure level (and the observation time as well). Predictors taken from a pressure level of the free atmosphere should be preferred, since weather processes like convection mainly take place in the free atmosphere. However, many downscaling studies operate with measured and modeled predictor information describing the state of the atmosphere of the planetary boundary level or even at the earth surface, indicating that those predictors are probably more reliably measured and simulated than predictors taken from a higher pressure level.

### Inconsistency and Erroneous Data

The reanalysis data of NCEP used in this study is only partially homogeneous. For instance, NCEP provides a document on their webpage listing some problems and inconsistencies of the reanalysis data (see NCEP, 2009a). It is mentioned that the reanalysis data from 1948 to 1957 is forecast data, because the observations for the GCM were only available at 3, 9, 15 and 21 UTC instead of 0, 6, 12 and 18 UTC.

Beside this inhomogeneity there are further irregularities in the data which are not listed in the NCEP documentation. Although the reanalysis data is output from a simulation model, there are also time periods with missing values indicated by the black bars in Figure 3.5. The specific humidity and the u-wind component are both characterized by more than two years of data gaps which is around 4 % of the data, whereas the data gap of the geopotential height is short and only at the beginning of the reanalysis period. Unfortunately, the data gaps always occur at the same time for a single predictor, so that a replacement of gaps based on statistical techniques selecting the values from the same variable but from another observation time or pressure level is not possible. Furthermore, the application of multivariate predictor sets shortens the investigation period available for the model development since the data gaps of the different variables do not occur at the same time. On the other hand, the increase is not larger than 10 % and missing values do not appear in every time series. For instance, the relative humidity has no data

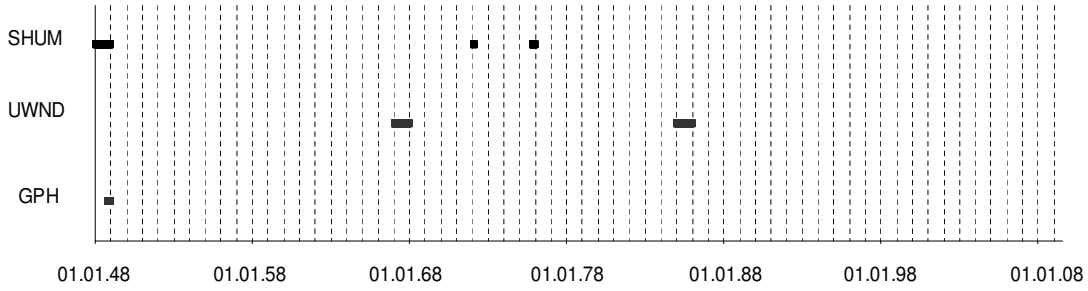



**Figure 3.5:** Black bars indicate periods with missing values for the geopotential height (GPH), the u-wind component (UWND) and the specific humidity (SHUM).

gaps, but compared to the other predictors this variable contains simulation data which seem to be physically not realistic. On 5 % of the days, the relative humidity at the grid points located next to the study regions is less than 10 %.

The examples of this section briefly demonstrate that reanalysis data is only partly homogeneous. There seems to be still many difficulties with the reanalysis information even with physically unreasonable data. A cross-check with other globally gridded data sets, like ERA-40, may help the detection of erroneous information (see Ben Daoud et al., 2009) to provide a more homogeneous and reliable predictor archive for downscaling applications.

### Standardization and Anomalies

The range of the predictor values is usually quite different which makes the definition of predictor weights more difficult if different predictors are used for the downscaling. For instance, the values of the 1000 hPa height field at 18 UTC ranges from -601 m to 449 m, whereas the value of the 500 hPa height field at the same observation time varies between 4642 m and 6026 m. To overcome this problem, a predictor is usually standardized. In this study two different standardization techniques are used.

The first technique is based on a linear transformation to transfer the predictor value $x(t)$ at a given grid point at time $t$ to a standard interval where the standardized values range between zero and 1:

$$x_n(t) = \frac{x(t) - x_{min}}{x_{max} - x_{min}} \tag{3.1}$$

where $x_n(t)$ is the value of the standardized predictor, $x_{min}$ is the minimum value and $x_{min}$ is the maximum value of the predictor within the given data set.

Beside this linear transformation the anomalies of the predictors are also calculated based on the following transformation:

$$x_a(t) = \frac{x(t) - \overline{x}}{s} \tag{3.2}$$

The arithmetic mean of the predictor $\overline{x}$ is subtracted from the predictor value divided by the standard deviation $s$ of the predictor. The calculation of the anomalies has the

advantage that the annual cycle of the predictand can be removed if a corresponding mean and standard deviation is selected. In this investigation the mean and the standard deviation are calculated by using a triangular kernel (weighting function) with a kernel length of 30 days for each grid point and for each day of a year.

# 4 Analog Method

It is well known that weather situations (patterns) of certain areas like Central Europe resemble each other from time to time (Radinović, 1975). This property of the atmosphere is exploited by the analog method. The basic idea of the approach is to search in an archive of past weather situations and to select those situations which are the most similar to the forecast (current) one. The selected situations are called as analogs and their local conditions are taken as a forecast for the surface variable of interest.

The analog method is not a new approach in weather forecasting. The methodology was probably firstly mentioned by Namias (1951) in medium-term weather forecasting. Several years later, Lorenz (1969) used the approach to study the predictability of the atmosphere. While more recently, Obled et al. (2002) has presented a detailed description of an analog forecasting algorithm developed for the prediction of daily areal precipitation. The methodology was enhanced by Bontron and Obled (2003) based on reanalysis data for the model development. The analog method was also selected to forecast other atmospheric variables. For example, Woodcock (1980), Kruizinga and Murphy (1983) and Radinović (1975) used the analog method for temperature forecasting and Sievers et al. (2000) and Fraedrich et al. (2003) selected the approach to forecast tropical cyclone tracks. The analog method was also applied in climate modeling for downscaling daily precipitation (see Zorita and von Storch, 1999; Wetterhall et al., 2005; Matulla et al., 2007). In this branch of atmospheric science the method is often called the benchmark approach since the methodology is easy to perform and the results are quite encouraging (Zorita and von Storch, 1999).

## 4.1 Principle of the Analog Method

The forecast performance of the analog method depends on several criteria which must be specified by the user. The first step is the definition of a search algorithm for the selection of analogs. To forecast daily areal precipitation for a specific river basin, the search algorithm can be as follows:

1. The target weather situation is compared with past situations to identify an analog situation.

2. The daily areal precipitation which occurred on the day of the analog is chosen for the specific river basin.

3. Step 1 and 2 are repeated to identify the most similar analogs and the corresponding precipitation sample to describe the forecast uncertainty.

The approach can be easily extended to forecast point precipitation and to other surface variables by replacing in the second step the daily areal precipitation with the surface variable of interest. It is also possible to create a more complex search algorithm if all steps of the algorithm are repeated. Bontron and Obled (2003) proposed such a two step search algorithm for the selection of analogs. They selected at first a large subsample of analog situations identified from all records. Afterwards, the search of analogs was repeated by selecting a smaller sample from the situations of the subsample.

Probably, the most important factor of the forecast approach is the choice of an appropriate predictor to describe the weather state over the basin of interest. A full description of the weather state is not possible, otherwise a three dimensional profile of all meteorological variables must be selected. Therefore, one or a set of two or three predictors are selected which must fulfill certain constraints (see Section 3.2). Screening procedures illustrate that mostly pressure related variables like the mean sea level pressure or the geopotential height taken from a certain pressure level have the highest potential. Their daily fields are usually selected as the first predictor (e.g. Zorita and von Storch, 1999; Obled et al., 2002). The introduction of a second and a third predictor can improve the performance of the approach. For example, Bontron and Obled (2003) used a combination of the geopotential height and humidity fluxes. Matulla et al. (2007) selected the daily mean sea level pressure and the daily specific humidity of the 700 hPa level. Instead of daily values, predictors with a higher temporal resolution (twice a day or six hourly) can be also selected (see Bontron and Obled, 2003) to avoid a temporal offset of the forecast.

After the selection of an appropriate predictor, a measure must be defined by the user for the comparison between two atmospheric states. The S1-Score is often selected (see Woodcock, 1980; Obled et al., 2002; Wetterhall et al., 2005) which compares the north-south and the west-east gradients of two atmospheric patterns. The S1-Score has been originally proposed in forecast verification by Teweles and Wobus (1954) to evaluate the height fields of the forecast with the observed height field. Beside the S1-Score, the Euclidean distance or a distance function related to the Pearson correlation has been also selected as measures (Bliefernicht and Bárdossy, 2007). A comparison of several functions was only performed in a few studies (see e.g. Toth, 1991). Recently, Matulla et al. (2007) have selected five measures for downscaling daily precipitation in California and the European Alps. They concluded that it is difficult to define a best measure and that a suitable measure depends on the purpose of downscaling. However, the Euclidean distance performs in most cases well and can be a reasonable first choice.

A crucial factor of the methodology is the definition of an objective function which measures the differences between forecasts and observations. The objective function is taken to maximize the forecast performance and to determine suitable model parameters. Like the other steps, the objective function also depends on the model purpose and should be carefully selected. In the case of a probability forecast, the performance measures defined in Section 2.2.3 are often taken as an objective function. For example, the ranked probability score was selected by Obled et al. (2002) and Wetterhall et al. (2005) and the continous version of the ranked probability score by Bontron (2004).

The last step of the methodology is the specification of the predictor settings and the estimation of the model parameters. There are many ways to find appropriate settings for the analog method and no general strategy is proposed in the literature. In the beginning of the model calibration a screening procedure can be taken which scans through an archive of possible predictors to find an optimal candidate. After the screening, the form or the size of the predictor domain can be varied to find a reasonable domain. For example, Obled et al. (2002) selected a rectangular domain covering Central Europe for their target region in France. Wetterhall et al. (2005) pointed out that the predictor settings also depends on the season because the link between the large-scale atmospheric circulation and the local-scale process changes during the year. To account for the intra-annual variability, they selected a smaller predictor domain for the summer than for the winter. Finally, in many applications the search of analogs is also restricted by a selection rule. For instance, Matulla et al. (2007) selected a moving window of 90 days for each year of the investigation period which was centered around the target day.

In this chapter the ideas of the aforementioned investigations are taken to develop an analog forecasting algorithm for the four study regions (see Chapter 3). The different ways of model calibration are illustrated and the influence of the user-defined criteria is determined and evaluated. Furthermore, the following new ideas and research topics are proposed and addressed in this section:

1. **Distance measure**: In a study of Bliefernicht and Bárdossy (2007) a mixed distance function is proposed which is a weighted mixture of the Euclidean distance and a distance function related to the Pearson correlation. The Euclidean distance measures the closeness between two patterns, but two atmospheric patterns can also be similar according to their shape. This property is measured by the second distance function. In this chapter a more general form of the mixed distance function is proposed which is based on the family of the $L_p$-distances.

2. **Parameter estimation and model validation**: There are only a few studies which used the analog method for downscaling daily precipitation in Germany (see e.g. Brandsma and Buishand, 1998; Buishand and Brandsma, 2001). In this investigation, the model parameters of the approach are specified for four different study regions in Germany and for two time periods (winter and summer term). To test the spatial transferability of the methodology, the model parameters are only specified for one catchment in each study region. Afterwards, the calibrated model is transferred to the neighboring catchment.

3. **Objective function**: The analog method developed in this section is designed for the prediction of intensive areal precipitation in mesoscale catchments. Since the precipitation forecast represents an element of a flood warning system, the performance of the analog forecast is determined for flood-producing situations based on a cost-loss approach (see Section 2.2.5). This approach is taken to formulate an objective function which enables the minimization of the negative outcomes ("false alarm" and "miss") of a binary warning system so that the value of the probability forecast in decision making situations is maximized.

The precipitation realizations provided by the search algorithm can be either taken as an ensemble forecast or as a probability forecast to describe the forecast uncertainty. It is reported in many studies that a probability forecast has a higher forecast value or at least the same value compared to other forecasting techniques (see Section 2.1.4) if the decision maker selects the optimal decision threshold for an individual user. However, the selection of an optimal threshold is not straightforward in decision making situations. In this section the basic framework of the cost-loss approaches is illustrated and it is demonstrated how a decision maker should apply a probability forecast in cost-loss situation to maximize the benefit for an individual user. Furthermore, it is shown that an accurate forecast must not necessarily correspond to a valuable forecast for an individual user and it is highlighted that even a low-skill reference forecast can outperform model predictions for certain end users.

The analog forecasting algorithm presented in this section is designed to predict the prospective evolution of the daily areal precipitation for the short (1-3 days) and the medium term (4 to 7 days). Since the forecast performance of the analog method strongly depends on the predictor set, the model parameters of the downscaling approach are defined for two different multivariate predictor sets. The first predictor set (A) contains three atmospheric variables: the geopotential height (GPH), the zonal moisture fluxes (UFLX) and the relative humidity (RHUM). In the operational mode this predictor set is taken to forecast the daily areal precipitation for lead times of two or three days. Beyond this lead time, humidity related variables are less reliably predicted by a global NWP model. Thus, a second predictor set (B) is chosen which consists of two geopotential height fields differing in observation time and pressure level. The maximal number of predictors in both predictor sets is not larger than three. The increase of the forecast performance is only minor and probably not statistically significant if more than three predictors are selected (Bontron and Obled, 2003).

In the next sections the distance functions used in this chapter are introduced and the objective function is described. The cumulative distribution functions chosen for the estimation of the daily precipitation are described in Section 4.4. In Section 4.5 the strategy of the model development is proposed. The results of the downscaling are presented and discussed in Section 4.6 and in Section 4.7. The chapter is closed by a short summary and a conclusion.

## 4.2 Pattern Closeness and Pattern Similarity

The closeness between two atmospheric patterns is measured by a distance function. A function is called a distance if it fulfills the following conditions:

1. The distance is positive between two different points (positiveness).

2. The distance is zero for identical points (reflexivity).

3. The distance between two points is the same in either direction (symmetry).

4. The distance between two points is the shortest distance along any path (triangular inequality).

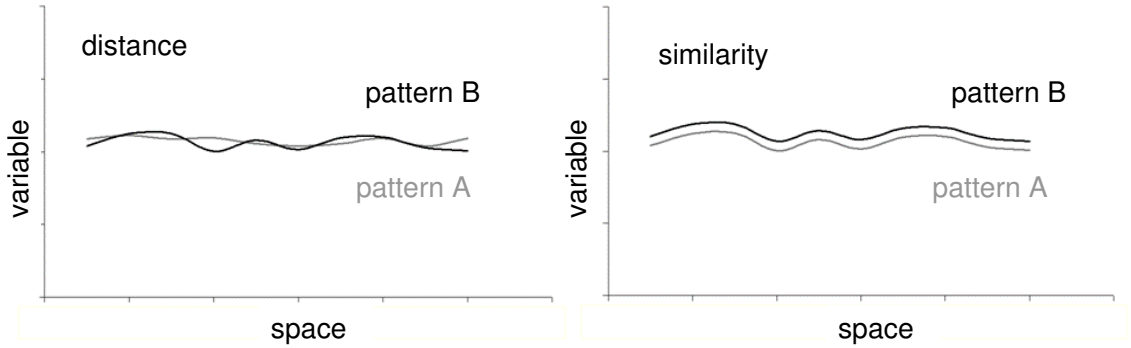



**Figure 4.1:** Two kinds of "similarity" between two atmospheric patterns illustrated for the one-dimensional case: pattern closeness (left figure) and pattern similarity (right figure).

In analog forecasting the distance between the values at the grid points of two atmospheric patterns is calculated. The closer the patterns are, the smaller is the value of the distance function. An example of a distance function is the Euclidean distance $d_1$. This distance is defined for the comparison of two atmospheric patterns $\mathbf{x}$ at time step $t_1$ and $t_2$ as follows:

$$d_1[\mathbf{x}(t_1), \mathbf{x}(t_2)] = \left( \sum_{k=1}^{K} [x_k(t_1) - x_k(t_2)]^2 \right)^{\frac{1}{2}} \tag{4.1}$$

where $x_k$ is the value of the atmospheric pattern at the grid point $k$ and $K$ is the number of grid points selected for the comparison.

The Euclidean distance is limited since it only measures the closeness between two spatial fields. But atmospheric patterns can also be similar according to their shape (see Figure 4.1). This property is measured by the Pearson correlation $r$:

$$r[\mathbf{x}(t_1), \mathbf{x}(t_2)] = \frac{\frac{1}{K} \sum_{k=1}^{K} [x_k(t_1) - \bar{x}(t_1)] [x_k(t_2) - \bar{x}(t_2)]}{s(t_1) s(t_2)} \tag{4.2}$$

where $\bar{x}(t_1)$ and $\bar{x}(t_2)$ are the arithmetic means and $s(t_1)$ and $s(t_2)$ are the corresponding standard deviations of both atmospheric patterns. The Pearson correlation varies between -1 and 1. If both atmospheric patterns are similar, the Pearson correlation tends to 1.

To construct a distance which measures pattern closeness and pattern similarity, the Euclidean distance and the Pearson correlation must be combined. However, the combination of both functions is not straightforward and a distance related to the Pearson correlation must be first introduced: In the first step of this procedure, each value of an atmospheric pattern for a given day is standardized by the corresponding mean and standard deviation of the atmospheric pattern:

$$x_k'(t) = \left( \frac{x_k(t) - \bar{x}(t)}{s(t)} \right) \tag{4.3}$$

The Euclidean distance $d_2$ between two atmospheric patterns in the normed space is:

$$d_2[\boldsymbol{x}^{'}(t_1), \boldsymbol{x}^{'}(t_2)] = \left( \sum_{k=1}^{K} \left[ x_k^{'}(t_1) - x_k^{'}(t_2) \right]^2 \right)^{\frac{1}{2}} \tag{4.4}$$

which can be decomposed to:

$$= \left( \sum_{k=1}^{K} \left[ x_k^{'}(t_1)^2 - 2x_k^{'}(t_1)x_k^{'}(t_2) + x_k^{'}(t_2)^2 \right] \right)^{\frac{1}{2}} = \tag{4.5}$$

$$= \left( 2K - 2\sum_{k=1}^{K} \left[ x_k^{'}(t_1)x_k^{'}(t_2) \right] \right)^{\frac{1}{2}} \tag{4.6}$$

Since the arithmetic mean of the standardized predictors tends to 0 and the corresponding standard deviation tends to 1, the Pearson correlation of the normed predictors can be simplified to:

$$r[\boldsymbol{x}^{'}(t_1), \boldsymbol{x}^{'}(t_2)] = \frac{1}{K} \sum_{k=1}^{K} x_k^{'}(t_1)x_k^{'}(t_2) \tag{4.7}$$

This formulation can be taken to replace the right term in Equation 4.6. Finally, the following relationship exists between the Euclidean distance in the normed space and the Pearson correlation:

$$d_2[\boldsymbol{x}^{'}(t_1), \boldsymbol{x}^{'}(t_2)] = 2K \left( 1 - r[\boldsymbol{x}^{'}(t_1)\boldsymbol{x}^{'}(t_2)] \right)^{\frac{1}{2}} \tag{4.8}$$

The two Euclidean distances $d_1$ and $d_2$ can be combined to a mixed distance $d$:

$$d = d_1[\boldsymbol{x}(t_1), \boldsymbol{x}(t_2)] + \beta d_2[\boldsymbol{x}^{'}(t_1), \boldsymbol{x}^{'}(t_2)] \tag{4.9}$$

where $\beta > 0$ balances between pattern closeness and pattern similarity.

The Euclidean distance belongs to the family of the $L_p$-distances:

$$L_p[\mathbf{x}(t_1), \mathbf{x}(t_2)] = \left( \sum_{k=1}^{K} |x_k(t_1) - x_k(t_2)|^p \right)^{\frac{1}{p}} \tag{4.10}$$

with a parameter p ranging between 1 and infinity. For $p = 2$ the $L_p$-distance is equal to the Euclidean distance. Another specific case of the $L_p$-distance is the Manhattan distance with $p = 1$.

The parameter $p$ can be used to change the weight of the grid points used for the distance calculation. The larger the value of $p$ is, the higher is the importance of

grid points with large differences. If $p$ tends to the infinite, the distance calculation only depends on the grid point with the maximum difference.

Replacing the Euclidean distance $d_1$ and $d_2$ in Equation 4.9 by the $L_p$-distance, a weighted mixture of two $L_p$-distances can be formulated for the mixed distance:

$$d^* = \left( \sum_{k=1}^{K} |x_k(t_1) - x_k(t_2)|^{p_1} \right)^{\frac{1}{p_1}} + \beta \left( \sum_{k=1}^{K} \left| x_k^{'}(t_1) - x_k^{'}(t_2) \right|^{p_2} \right)^{\frac{1}{p_2}} \tag{4.11}$$

This distance has three parameters: $p_1$ and $p_2$ of the corresponding $L_p$-distance and the weight $\beta$ balancing between both terms.

The mixed distance function proposed in Equation 4.11 compares the pattern of a single predictor. The distance between two patterns of a multivariate predictor set can be obtained by calculating the sum of the individual mixed distances $d_j^*$:

$$D = \sum_{j=1}^{J} a_j d_j^* \tag{4.12}$$

where the parameters $a_j$ defines the predictor's weight $j$ and $J$ represents the number of predictors. The range of the individual mixed distances is usually very different, especially for non-standardized predictors. The estimation of the predictor weight is easier if a single distance is standardized by the maximum distance. The maximum distance can be estimated e.g. by calculating the distances between all pairs of weather situations from a subsample of situations which is selected randomly from all records.

The S1-Score mentioned in the beginning of this chapter also belongs to the family of $L_p$-distances. Instead of comparing the absolute values of two atmospheric patterns, the gradients fields are compared. The S1-Score consists of two terms. The first term, the nominator, calculates the sum of the absolute differences between the gradients over all adjacent pairs of two atmospheric patterns. This term is identical to the Manhattan distance of two predictors that are equally weighted. The second term, the denominator, calculates the sum of the maximum values of the absolute gradients for all adjacent pairs. This term is used for a normalization to determine a kind of relative differences between the gradient fields. It is introduced to account for intra-annual variability of a predictor, because in summer the variability of the geopotential height is usually higher than in winter. For a more detailed explanation of the S1-Score the reader is referred to Wilks (2006).

Beside the family of the $L_p$-distances, there are also some other measures like the Mahalanobis distance which fulfill the conditions of a distance function and which takes the intercorrelation between the variables into account. This distance was already selected by Matulla et al. (2007) but the results were generally poorer compared to the Euclidean distance.

## 4.3 Maximizing the Forecast Value

A decision maker must decide between 'no action' and 'protective action' when an extreme event is forecast. Depending on the outcomes of the decision making

process, the user can profit from the weather forecast or he must sustain a loss. Cost-loss approaches have been introduced in weather forecasting to analyze the value of decision-making processes. An introduction into these approaches and some arguments for their use (in form of question) is presented in Section 2.2.5. In this section the basic idea of those approaches is briefly described and a skill score is proposed which determines the forecast value of a probability forecast. In the second part of this section this skill score is taken for the development of an objective function which maximizes the forecast performance of a binary probability forecast for a set of users.

In this study the simplest form of a cost-loss approach is taken which describes the cost-loss situation of a two action, two state decision making problem (Murphy, 1977). The decision maker has two alternatives: "*no action*" and "*protective action*". Both actions are associated to two possible states of the weather: "*adverse weather*" and "*no adverse weather*". The consequences of this particular decision making problem are four different outcomes ($C$,$C$,$L$,0) with three different realizations (see Table 2.3). If a decision maker gives a warning, protective action is taken preventing the damage but causing a cost $C$. If the decision maker gives no alarm, there are no costs as long as no adverse weather occurs. Otherwise, he must incur a loss.

To determine the average expense of the forecast for a particular user $\alpha_m$ and for a given investigation period, the decision maker must define a priori an event threshold which corresponds to a return (occurrence) frequency $s_l$ and a probability threshold $p_t$ used for decision making. Based on both thresholds, the outcomes of a binary forecast (see Table 2.2) are counted and related to the cost and the loss. Finally, the average expense of a forecast system $E_f$ represents the average cost of all warnings and the average loss of all misses which occurred throughout the entire investigation period:

$$E_f\left(\alpha_m, s_l, p_t\right) = \frac{a\left(s_l, p_t\right) + b\left(s_l, p_t\right)}{n} C + \frac{c\left(s_l, p_t\right)}{n} L \tag{4.13}$$

where $a$ is the number of hits, $b$ is the number of false alarms, $c$ is the number of misses and $n$ is the number of observation and forecast pairs. The value of the cost-loss ratio $\alpha_m = C/L$ depends on the individual user.

A skill score for the average expense can be derived by introducing the climate expense as reference forecast. The climate expense is the minimum average expense of two low-skill warning strategies: the user protects always ("*warning*") or he protects never ("*no warning*"). The average cost of the first strategy is the cost itself, since the cost must be spent for protection on every day of the investigation period:

$$E_a = C \tag{4.14}$$

If the decision maker selects the strategy "*no warning*", the average expense is the occurrence frequency of the event times the loss:

$$E_n\left(s_l\right) = s_l L \tag{4.15}$$

since the decision maker must incur the loss on those days when the event occurs. The optimum strategy between both low-skill warning strategies is the minimum of both expenses and is called average climate expense $E_c$:

$$E_c(s_l) = \min[E_a, E_n(s_l)] \tag{4.16}$$

A skill score can be specified with the help of the standard formulation of a skill score (Equation 2.2) and by determining the average expense of a perfect forecast $E_1$:

$$E_1(s_l) = s_l C \tag{4.17}$$

The average expense of a perfect forecast is the cost for protection needed when adverse weather occurs. The skill score is called relative value or economic value $V$ (Richardson, 2003):

$$V(\alpha_m, s_l, p_t) = \frac{E_f(\alpha_m, s_l, p_t) - E_c(s_l)}{E_1(s_l) - E_c(s_l)} \tag{4.18}$$

The score of the relative value ranges between minus infinity and one. If the relative value is greater than zero, the forecast of the warning system has a higher performance than the low-skill reference forecast. The forecast is perfect if the relative value is one.

The relative value proposed in the previous equation is a relative measure based on two low-skill warning strategies. A third straightforward protection strategy is the persistence forecast. In the event of this, the decision maker takes protective action if an adverse weather has been occurred the day before. This kind of protection strategy is an appropriate reference forecast for days with the same weather pattern. The mean expense of the persistence forecast $E_p$ is:

$$E_p(\alpha_m, s_l) = \frac{a_p(s_l) + b_p(s_l)}{n} C + \frac{c_p(s_l)}{n} L \tag{4.19}$$

where $a_p$ is the number of hits, $b_p$ is the number of false alarms and $c_p$ is the number of miss of a persistence forecast. If the expense of a persistence forecast and the climate expense are minimized, the average expense of the reference forecasts $E_0$ is the minimum expense of all three low-skill warning strategies:

$$E_0(\alpha_m, s_l) = \min[E_p(\alpha_m, s_l), E_a, E_n(s_l)] \tag{4.20}$$

To create a skill score on the climate expense and the persistence expense, $E_c$ must be replaced in Equation 4.18 by $E_0$:

$$V^*(\alpha_m, s_l, p_t) = \frac{E_f(\alpha_m, s_l) - E_0(\alpha_m, s_l)}{E_1(s_l) - E_0(\alpha_m, s_l)} \tag{4.21}$$

The relative value $V$ is a function of the cost-loss ratio, the probability threshold

and the occurrence frequency of an event. This dependency is used to create an objective function which maximizes the forecast value of a probability forecast for a set of users and for extremes. This function is derived in the following way:

If a decision maker selects the optimal threshold from a set of decision thresholds $\boldsymbol{p} = [p_1, p_2, ..., p_T]$ for an individual user, he can maximize the forecast value for this user:

$$G(\alpha_m, s_l) = max\left[V(\alpha, s_1, p_1), V(\alpha, s_2, p_2), ..., V(\alpha_M, s_L, p_N)\right] \tag{4.22}$$

where $T$ is the number of decision thresholds and $G$ represents the maximum forecast value for an individual user. The mean maximum value for a set of users $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_M]$ and for a set of events $\boldsymbol{s} = [s_1, s_2, ..., s_L]$ approximates the overall forecast value over the entire range of the predictand. This value is taken as the objective function and can be calculated in the following form:

$$obj = \frac{1}{L}\frac{1}{M}\sum_{l=1}^{L}\sum_{m=1}^{M} G(\alpha_m, s_l) \tag{4.23}$$

where $M$ is the number of users and $L$ is the number of events. To specify an objective function which calculates the forecast value over the entire range of the predictand with a focus on intensive precipitation, the following occurrence frequencies are selected $\mathbf{s} = [0.70, 0.80, 0.90, 0.95, 0.975, 0.99, 0.995]$. Additionally, a logarithmic distribution for the users is selected ranging between $\boldsymbol{\alpha} = [0.0001, ..., 1]$ to focus on users with low cost-loss ratios (small costs and high losses).

The objective function proposed in this section calculates the forecast value of the probability forecast for a specific set of users and over the entire range of the predictand. The aim of the model development is the maximization of the objective function so that finally the model performance in terms of the forecast value is maximized. Note that the increase of the objective function leads to an increase of the mean benefit for all users, but it can also decrease the forecast value of a single user (see Richardson, 2003). Furthermore, the objective function does not take into account the direction of a temporal offset of a warning. A warning issued one day too early is more valuable than a warning which is one day too late, evidently. In the first case, it is very likely that there is still protection due to the false alarm whereas in the latter case the warning comes definitely too late.

This section gives the theoretical framework of the objective function. In Section 4.7 the basic idea of the optimal use of a probability forecast in cost-loss situations is illustrated which will deliver more insight into the formulation of the objective function.

## 4.4 Probabilistic Precipitation Forecast

Three cumulative distribution functions are tested to provide a probability forecast for the daily areal precipitation: the empirical, the exponential distribution and the mixed exponential distribution. An illustrative overview about common probability

distribution functions used for the estimation of the precipitation process is given in Wilks and Wilby (1999).

The empirical distribution already represents an appropriate choice for the estimation of the daily areal precipitation. It can well describe the strongly skewed distribution of the predictand. However, forecasting local-surface variables with an empirical sample is restricted to the observations, and an extrapolation of the predictand larger than the maximum observed value is not possible. To overcome this problem a theoretical cumulative distribution function must be fitted to the empirical sample. However, the goodness of the extrapolation also depends strongly on the portion of analogs selected by the algorithm.

The cumulative probability function of the exponential distribution can be taken for the estimation of the precipitation amount for a given forecast day $t$:

$$\mathtt{F}\left[y(t)\right] = 1 - e^{-\lambda(t)y(t)} \;, y(t) \geq 0 \tag{4.24}$$

The parameter $\lambda(t)$ of the distribution is estimated by replacing the expected value of the exponential distribution:

$$\mathtt{E}\left[y(t)\right] = \frac{1}{\lambda(t)} \tag{4.25}$$

with the mean precipitation of the $k$-similar analogs: $\lambda(t) \approx 1/\bar{x}(t)$. Since the distribution can well describe the skewed character of daily precipitation, the exponential distribution is often selected for the estimation of daily precipitation amounts (Wilks and Wilby, 1999). However, the exponential distribution is not as flexible as probability distribution functions with two or more parameters. Those distribution functions are often preferred. They can more suitably approximate the discrete-continuous character of daily precipitation. For example, if the analog method is taken to forecast daily precipitation in Central Europe, the precipitation sample often contains a number of wet days and dry days ("zeros") in particular for summer days. If a distribution function is selected which cannot appropriately describe the discrete-continuous character of daily precipitation, it is very likely that high precipitation amounts are underestimated (see Wilks, 1999). To overcome this problem, the mixed exponential distribution is taken which is a weighted mixture of two ordinary exponential distributions. The cumulative distribution function of the mixed exponential distribution function is defined as:

$$\mathtt{F}\left[y(t)\right] = p_0(t)\left[1 - e^{-\lambda_1(t)y(t)}\right] + \left[1 - p_0(t)\right]\left[1 - e^{-\lambda_2(t)y(t)}\right], y(t) \geq 0 \tag{4.26}$$

The first term of this equation describes the left tail of the distribution function selected for the estimation of smaller precipitation amounts and to describe the portion of dry days. Higher precipitation amounts are modeled with the second term. The probability $p_0(t)$ defines the weight for each distribution function. The expected value of the mixed exponential distribution is the weighted average of the expected value of both ordinary exponential distributions:

$$\mathtt{E}\left[y(t)\right] = p_0(t)\frac{1}{\lambda_1(t)} + \left[1 - p_0(t)\right]\frac{1}{\lambda_2(t)} \tag{4.27}$$

Since small precipitation amounts cannot activate flood relevant processes, a precipitation threshold of 2 mm/d is selected to define all days with a precipitation amount lower than this threshold as dry days by replacing the actual value by a zero. A further reason for the conversion is that the precipitation probabilities are estimated for an area. The larger the area is, the higher is the likelihood for a precipitation event in this area. The selection of a precipitation threshold eliminates small events to provide finally a more appropriate description of the probabilities for low precipitation amounts.

The parameter $p_0(t)$ in Equation 4.26 is estimated in this study by the relative frequency of dry days of the analogs. In this case, the left term of this Equation can be simplified to:

$$\mathtt{F}\left[y(t)\right] = p_0(t) + \left[1 - p_0(t)\right]\left[1 - e^{-\lambda_2(t)y(t)}\right], y(t) \geq 0 \tag{4.28}$$

since the parameter $\lambda_1$ is infinite in the case of all days with zero observations and the exponential function of the left term of Equation tends 4.26 tends to 0. The expected value of this cumulative distribution function is:

$$\mathtt{E}\left[y(t)\right] = \left[1 - p_0(t)\right]\left(\frac{1}{\lambda_2(t)}\right) \tag{4.29}$$

The parameter $\lambda_2(t)$ of the mixed distribution function can be estimated by selecting the mean precipitation amount of wet days of the analogs.

## 4.5 Strategy of the Model Development

The model settings are specified for each study region and for the summer (April to September) and the winter half period (October to March). To test the spatial transferability of the approach, the catchments are separated into two groups. The daily areal precipitation of the Nahe, the Iller, the Freiberger Mulde and the Fränkische Saale is selected to specify the model parameters for the four study regions. Afterwards, the model is transferred to the neighboring catchments (Ruhr, Upper Sieg, Prüm and Upper Main) to predict the daily areal precipitation for these study sites.

The downscaling approach is validated by a cross-validation technique. This technique subdivides the data of the observation period into $k$ subsets. The data of $k-1$ subsets represent the calibration set taken to specify the model parameters. The remaining subset is used for the validation of the approach. Afterwards, the calibration and validation step is repeated $k-1$ times and in each repetition a new subset is selected as validation set. In this study an extreme form of the cross-validation technique, the leave-one-out cross-validation, is selected. This kind of cross-validation selects only the forecast day as validation set to evaluate the forecast quality. The remaining data is taken for the calibration. Afterwards, the validation is repeated

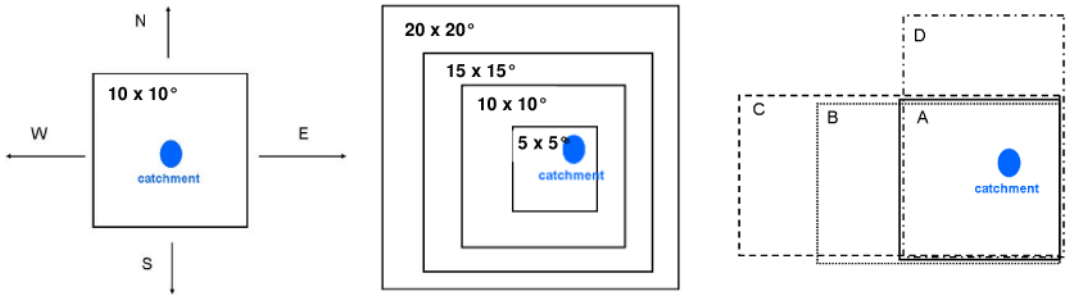



**Figure 4.2:** Variation of center, size and form of a predictor domain

for every day of the observation period. To ensure that the analogs are independent from the target day, the data of the neighboring days are excluded from the calibration set. The cross-validation period ranges from 1958 to 2001.

The cross-validation technique has several advantages compared to traditional split-sampling. In a split-sampling the data is also separated into a calibration and validation set, but the validation procedure is not repeated. This kind of technique has the shortcoming that only a subset of the observation data is used for model calibration as well as for model training. In contrary to the split-sampling, the cross-validation technique allows one to take the data of the entire observation period for model development. This kind of validation strategy is in particular helpful if the purpose of the model development is the prediction of extremes. In the event of this, the forecast quality of a model can be tested for the entire range of the investigation period and the model parameters can be estimated based on the entire data set. On the other hand, a cross-validation is more time consuming compared to split-sampling which makes e.g. the development of an automated optimization algorithm more difficult.

In this study the model parameters of the analog method are calibrated step-by-step which can be divided into five steps according to the following user-defined criteria:

1. **Predictor selection and configuration**: An appropriate configuration (observation time and pressure level) is determined for the predictors (GPH, UFLX and RHUM) by selecting the predictors from five observation times (6, 12, 18, 24 and 30 UTC) and four pressure levels (1000, 850, 700 and 500 hPa). The settings are specified for both the Euclidean distance and the Euclidean distance in the normed space to investigate whether the choice of the similarity measure also affects the predictor configuration.

2. **Predictor domain**: The domain configuration (form, size and center) is determined for the geopotential height and the moisture flux. At first, the location of the domain center is specified. Afterwards, a suitable size and form of the domain is determined by changing the size of four different forms (see Figure 4.2).

3. **Distance parameters and predictor weights**: The parameters ($p_1$, $p_2$, $\beta$) of the mixed distance functions are specified to find an appropriate $L_p$-

distance and to link pattern closeness to pattern similarity. Afterwards, the predictor coefficients ($a_j$) of the second and the third variable are stepwise determined. The predictor configuration is specified by screening through a predictor archive (GPH, MFLX and RHUM) of three pressure levels (1000, 850 and 700 hPa) and three observation times (12, 18 and 24 UTC).

4. **Selection rule**: The selection process is investigated by defining a moving window for every year of the investigation period centered on the target day. Then, the forecast value of the approach is determined for different window lengths (30 days to 365 days).

5. **Number of analogs and probability functions**: In step 1 to 4 an empirical sample of 60 analogs is taken to provide a probability forecast. To determine the objective function, the quantiles of the empirical distribution are selected as decision threshold $\boldsymbol{q_t} = [1/61, 2/61, ..., 60/61]$ with $\boldsymbol{p_t} = 1-\boldsymbol{q_t}$. However, the model performance also depends on the number of the most similar analogs and the cumulative distribution function. In the last step the forecast performance of the analog method, based on the exponential and the mixed exponential distribution, is calculated for different sets of the most similar analogs and compared to the reference algorithm based on the empirical distribution.

In every calibration step the score of the objective function is calculated and compared to the score of a reference algorithm based on standard settings to illustrate in a suitable manner the model improvement (or worsening) for the different user-defined criteria. However, the model improvements are in some particular cases small and probably not statistical significant or the 'best' model settings specified for a criterion seem to be not reliable. In both cases the model improvements are neglected and the new parameter set is replaced by a more plausible parameter set which is taken for the further steps of the model development.

In the following sections a more detailed description of the model development is given and the influence of the user-defined criteria on the forecast performance is illustrated.

## 4.6 Influence of User-Defined Criteria

### 4.6.1 Predictor Selection and Predictor Configuration

The predictor configuration (observation time and pressure level) is determined for the geopotential height, the moisture flux and the relative humidity. The predictors are taken from five observation times (6, 12, 18, 24 and 30 UTC) enclosing the measurement interval of daily precipitation (6 to 6 UTC of the following day) and from four pressure levels (1000, 850, 700 and 500 hPa). To hold the number of model runs as small as possible, the same predictor domain is selected for the four study regions. The domain is slightly smaller (15° x 15°) for the summer than for the winter (20° x 20°). The center of the predictor domain is located southwesterly to the study regions at 47.5° N and 5° E.

The model performance based on the Euclidean distance and the height field is given in Figures 4.3a and 4.3b. To illustrate the performance of the approach, the

mean value of the objective function is calculated for the four catchments selected for the calibration. The positive value of the objective function indicates that the analog forecast is superior compared to the low-skill reference forecasts for all observation times and pressure levels. The score of the objective function is also higher for winter than for summer indicating that the prediction of intensive precipitation is more skillful in winter. In summer, the relationship between the large-scale atmospheric circulation and precipitation is weaker. Intensive summer precipitation is often the result of local convective processes which have a high stochastic character. The prediction of those events is more difficult compared to events which are related to advective processes. But due to their local character the impact of convective precipitation is more relevant for head catchments.

The model performance is also strongly influenced by the predictor configuration. The forecast of intensive precipitation is improved if a predictor is selected from an observation time which is in the center of the measurement interval of daily precipitation. For example, the algorithm based on a height field at 18 UTC has the highest forecast value illustrating that a predictor taken from this observation time contains the most suitable information to explain intensive precipitation from 6 UTC to 6 UTC. However, the forecast performance is also influenced by the choice of the pressure level. Analog forecasting based on height fields of the boundary

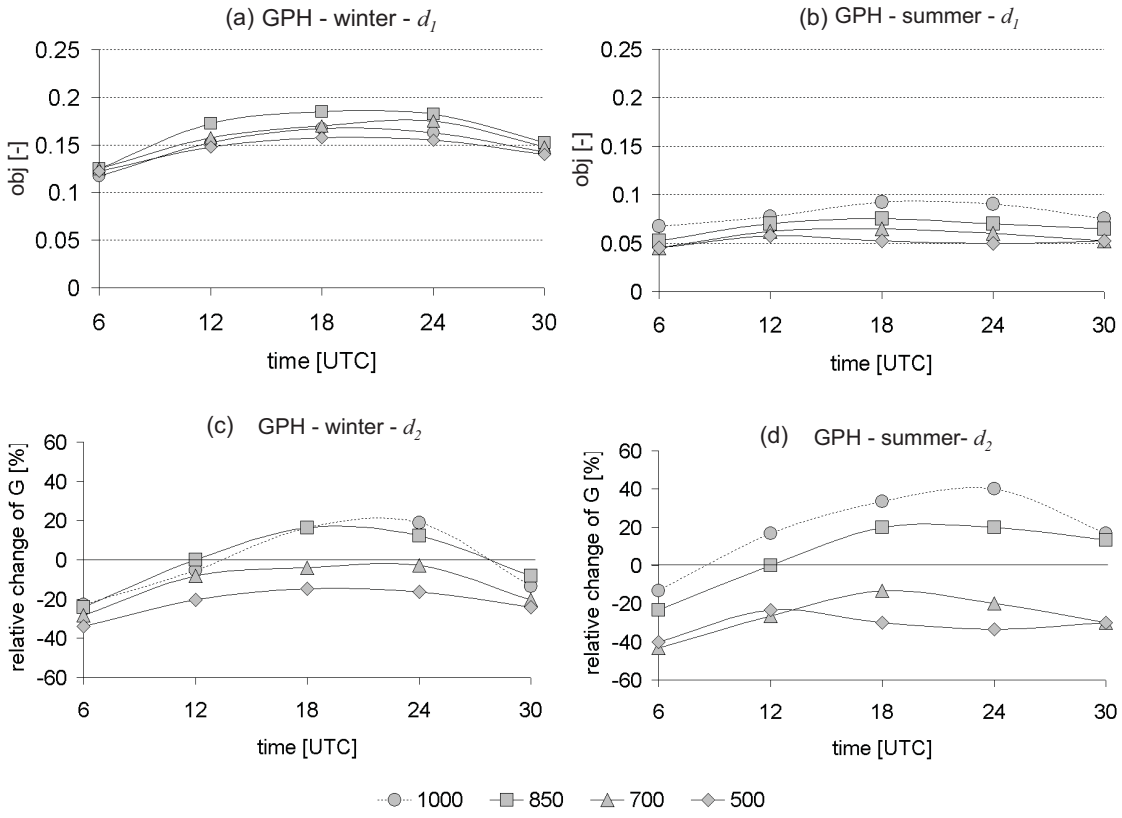



**Figure 4.3:** Influence of the predictor configuration on the model performance for the height field (GPH). $d_1$ = Euclidean distance, $d_2$ = Distance related to the Pearson correlation.

level is superior compared to height fields selected from the free atmosphere. For instance, the 1000 hPa height field performs best in summer. This result is a little bit surprising since the 1000 hPa field describes the state of the atmosphere on the ground while dominate precipitation processes like cloud formation take usually place in the free atmosphere. The result might indicate that the height fields of the boundary level are probably more reliably simulated by a GCM than the height fields of the upper pressure levels.

The optimum pressure level and observation time is also determined for an algorithm which measures the similarity between patterns to see whether the choice of the distance function also affects the predictor configuration. The Figures 4.3c and 4.3d show the model performance of this algorithm in comparison to a reference algorithm which measures pattern closeness and which performed suitable in the previous case (Euclidean distance and the 850 hPa height field at 18 UTC). On the ordinate the relative difference between both objective functions $\Delta obj^*$ is given. A positive value indicates a model improvement, whereas a negative difference represents a worsening of the model predictions. The selection of the algorithm comparing pattern similarity improves the forecast for summer and winter extremes. Further-

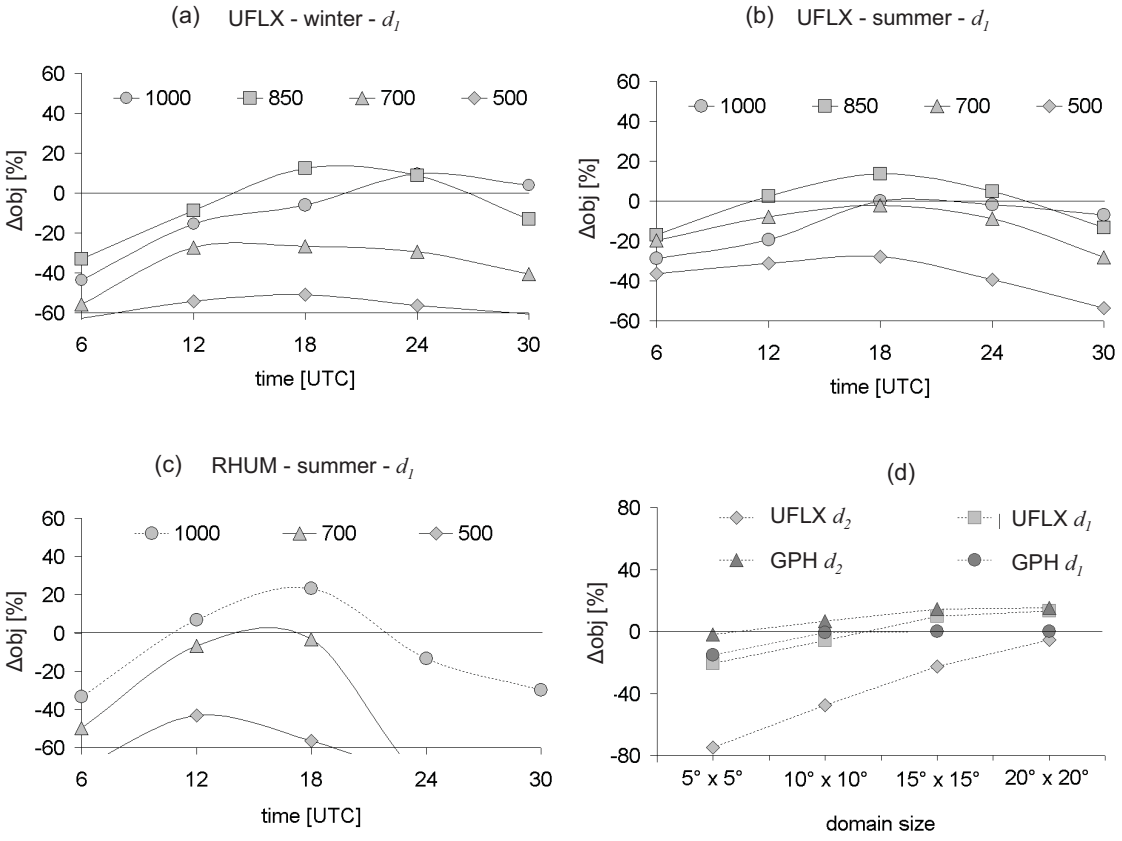



**Figure 4.4:** Influence of the predictor configuration on the model performance for the moisture flux (UFLX) and the relative humidity in comparison to a reference algorithm (Euclidean distance, 850 hPa height field at 18 UTC). $d_1$ = Euclidean distance, $d_2$ = distance related to the Pearson correlation.

more, the selection of this function also influences the predictor configuration. The optimum configuration for this distance is in summer the 1000 hPa height field at 24 UTC.

The results of the investigation for both humidity related variables are presented in Figure 4.4. The model performance is again compared to the reference algorithm selected in the previous case. Forecasting intensive precipitation based on the moisture flux field has a similar performance compared to the height fields (Figures 4.4a and 4.4b). An appropriate configuration for the moisture flux field is the 18 UTC field taken from a pressure level of the planetary boundary level such as the 850 hPa field. In summer, the relative humidity presents also a crucial variable for the prediction of intensive precipitation which can even outperform the prediction of the height fields (see Figure 4.4c). However, in contrast to the other two predictors, the selection of an appropriate configuration is more important, otherwise the forecast performance clearly decreases. In winter, the model performance based on the relative humidity is clearly lower when compared to the other predictors (not shown here). The investigation presented in this section is also repeated based on an algorithm which takes a smaller predictor domain (see Figure 4.4d). Forecasting intensive precipitation based on the moisture flux and the distance function related to the Pearson correlation is clearly poorer for smaller domains compared to other techniques. Thus, this kind of analog forecasting is neglected for further calibration steps.

The investigation in this section illustrates that the selection of an appropriate predictor configuration is an essential factor for the model performance. A reasonable first choice is a predictor field which describes the atmospheric state of the planetary boundary level at an observation time which is in the center of the measurement interval of daily precipitation. The investigation also pointed out that there are predictor configurations which outperform this choice. However, it is quite difficult to find an optimum configuration since the predictor configuration is influenced by other criteria like the distance function. In the following section it is demonstrated that the model performance is also influenced by the predictor domain.

### 4.6.2 Predictor Domain

The configuration of the predictor domain is specified for the 1000 hPa height field at 18 UTC and the 850 hPa zonal moisture field at 18 UTC. To point out a suitable location of the domain center, a quadratic domain (10° x 10°) is moved in each direction (see Figure 4.2). Afterwards, the optimum location of the domain center is taken to determine the size and form of the predictor domain. A quadratic predictor domain (A) and three rectangular forms (B, C, D) are selected (see Figure 4.2) which are changed from a small (e.g. 5°x 5°) to a large domain (e.g. 25° x 25°). The investigation is performed for all study regions to estimate the predictor domain for each region. The influence of the domain center on the model performance is given in Figure 4.5. The model performance is compared to a reference algorithm based on a predictor domain where the study region is located in the center of this domain. The figure shows the relative differences between both objective functions given in relation to the zonal distance. The zonal distance measures the distance between the longitude coordinate of the domain center and the longitude coordinate

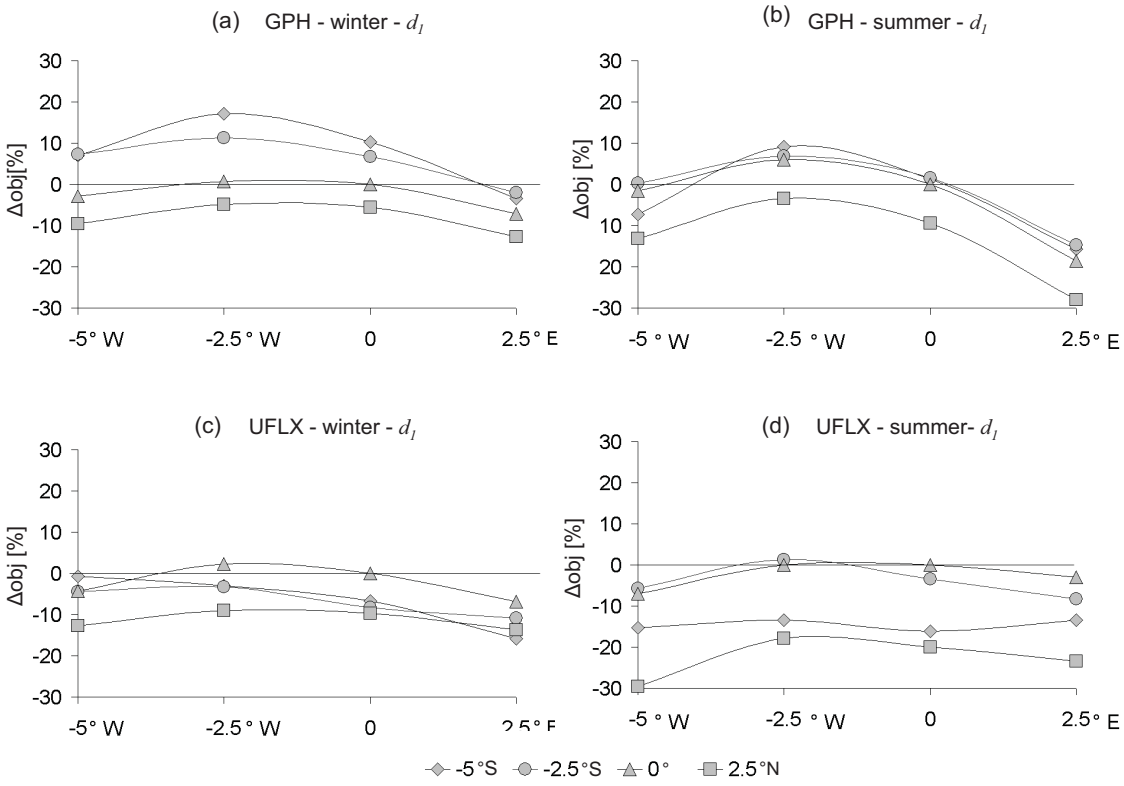



**Figure 4.5:** Influence of the domain center on the model performance for the geopotential height (GPH) and the moisture flux (UFLX) in comparison to a reference algorithm where the location of the domain center is identical to the location of the study region.

of the nearest grid point of a study region (see Table 3.1). The figure illustrates that the forecast performance increases if the domain center of the height field is located southwest of the catchments (Figure 4.5a and Figure 4.5b). The spatial offset of the predictor domain seems to be reliable since air masses from the southwest are usually warmer and can transport more moisture to the study region. In summer, the spatial offset of the height field is less due to the weaker link between the large-scale flow and intensive precipitation. A similar result is given in Figure 4.5c and Figure 4.5d for the moisture flux, but the spatial offset of the predictor domain is more dominant for the height field than for the moisture flux field.

The forecast performance of the analog method in relation to the domain size and the domain form is given in Figure 4.6 where the domain size is indicated by the number of grid points. For instance, if a quadratic domain with an extension of 10° x 10° is taken, the distance is calculated between the predictor values at 25 grid points. A reference algorithm based on a medium-sized quadratic domain (15° x 15° in summer and 20° x 20° in winter) is selected for the comparison.

The forecast is poor for a small height field (e.g. 5° x 5°). On the other hand, the choice of a large domain (e.g. 25° x 25°) can also worsen the forecast. This result illustrates that there is an optimum number of grid points needed for the distance calculation. However, this number is influenced by the season (compare

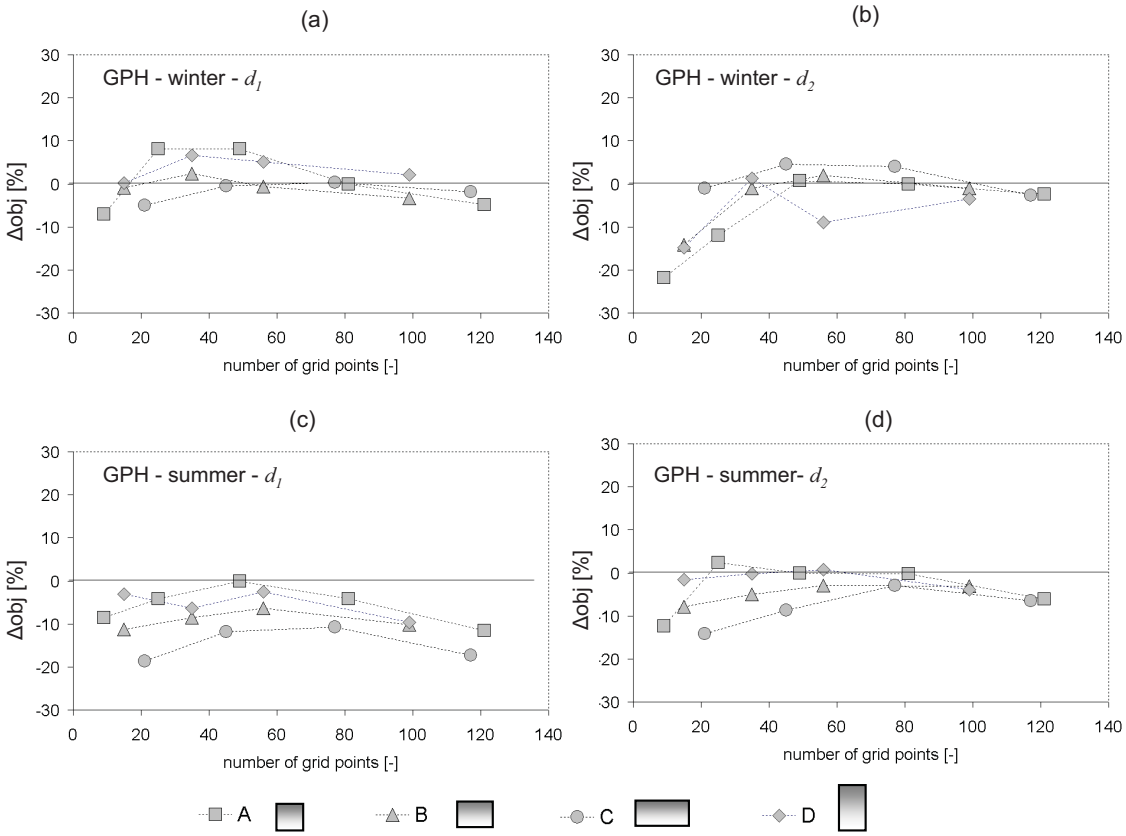



**Figure 4.6:** Influence of the domain form (A, B, C and D) and size (see for an illustration Figure 4.2) on the model performance for the geopotential height (GPH) in comparison to a reference algorithm based on a medium sized domain (winter: 20° x 20°; summer: 15° x 15°).

Figure 4.6a and Figure 4.6c) and the distance function (compare e.g. Figure 4.6a and Figure 4.6b). For instance, a smaller domain should be preferred for the predictions of summer extremes and a larger domain should be selected to determine pattern similarity. It is also illustrated in Figure 4.6a that the domain form influences the model performance as well. A quadratic domain seems to be a reasonable first guess, in particular for the predictions of summer extremes. But the selection of a rectangular domain (form C) can improve the forecast of winter extremes since humid westerly fluxes dominate in this season.

The investigations in this section demonstrate that the selection of an appropriate predictor domain is a further important step for the precipitation downscaling and can slightly improve the model performance. The domain size seems to be the most important factor for the prediction of intensive precipitation and should be determined before the form or the center is specified. In winter, a reasonable estimate for the height field and the moisture flux field is a rectangular domain based on more than 30 grid points (e.g. 15° x 10°) and which is slightly shifted to the southwest of the study region of interest. For the prediction of summer events a smaller domain can be selected which is quadratic (e.g. 10° x 10°) and more centered to the study region since local convective processes dominate during this season. Nevertheless, the investigations in this section demonstrate that a manual specification of the

**Table 4.1:** The predictor domain estimated for the height field (GPH) and the zonal moisture flux field (UFLX). $d_1$ = Euclidean distance, $d_2$ = distance function related to the Pearson correlation. The coordinates of the domain center are given in relative coordinates. The nearest grid point to the study region is located in the origin of the coordinate system.

| catchment | domain | period | GPH $d_1$ | GPH $d_2$ | UFLX $d_1$ |
|---|---|---|---|---|---|
| Mean | center | wi | (-5.0° ; -2.5°) | (-5.0° ; -2.5°) | (-2.5° ; 0.0°) |
| | | su | (-5.0° ; -2.5°) | (-2.5° ; -2.5°) | (-2.5° ; 0.0°) |
| | size | wi | 10° x 10° | 20° x 10° | 25° x 20° |
| | | su | 10° x 5° | 10° x 10° | 20° x 20° |
| Nahe | center | wi | (-5.0° ; -2.5°) | (-5.0° ; -2.5°) | (0.0° ; -2.5°) |
| | | su | (0.0° ; 0.0°) | (0.0° ; 0.0°) | (0.0° ; -2.5°) |
| | size | wi | 10° x 10° | 20° x 10° | 15° x 10° |
| | | su | 10° x 5° | 15° x 15° | 10° x 5° |
| F. Saale | center | wi | (-5.0° ; -2.5°) | (-5.0° ; -5.0°) | (-2.5° ; -5.0°) |
| | | su | (-5.0° ; -0.0°) | (-2.5° ; -2.5°) | (-2.5° ; -2.5°) |
| | size | wi | 10° x 10° | 15° x 15° | 20° x 10° |
| | | su | 20° x 15° | 5° x 5° | 15° x 15° |
| F. Mulde | center | wi | (-2.5° ; -5.0°) | (-2.5° ; 0.0°) | (0.0° ; 0.0°) |
| | | su | (-2.5° ; -2.5°) | (-5.0° ; -2.5°) | (0.0° ; 0.0°) |
| | size | wi | 10° x 15° | 25° x 15° | 20° x 20° |
| | | su | 5° x 10° | 10° x 10° | 25° x 25° |
| Iller | center | wi | (-2.5° ; -2.5°) | (-5.0° ; -5.0°) | (-2.5° ; -2.5°) |
| | | su | (-5.0° ; -2.5°) | (-5.0° ; -0.0°) | (-2.5° ; 0.0°) |
| | size | wi | 15° x 15° | 25° x 25° | 15° x 20° |
| | | su | 15° x 15° | 30° x 20° | 20° x 25° |

predictor domain is not straightforward and seems to be time consuming due to the dependence on several factors like the choice of the similarity measure, the season, the study region, the predictor and the predictand itself (see also the predictor domain determined for each study region in Table 4.1).

### 4.6.3 Distance Parameters and Predictor Weights

The parameters of the mixed distance ($p_1$, $p_2$ and $\beta$) and the predictor coefficients ($a_i$) are estimated one after another. At first, the parameters of the $L_p$-distances ($p_1$, $p_2$) are determined and afterwards the weight $\beta$ is specified to link pattern closeness to pattern similarity. All three parameters are only specified for the geopotential height. The parameter of the $L_p$-distance measuring pattern closeness is only determined for humidity related variables, since the model predictions based on the other distance measure are poor (see Section 4.6.1). The predictors are again the 1000 hPa height field at 18 UTC and the 850 hPa moisture flux field based on the predictor domain listed in Table 4.1.

After estimating the parameters for the distance function, this parameter set is

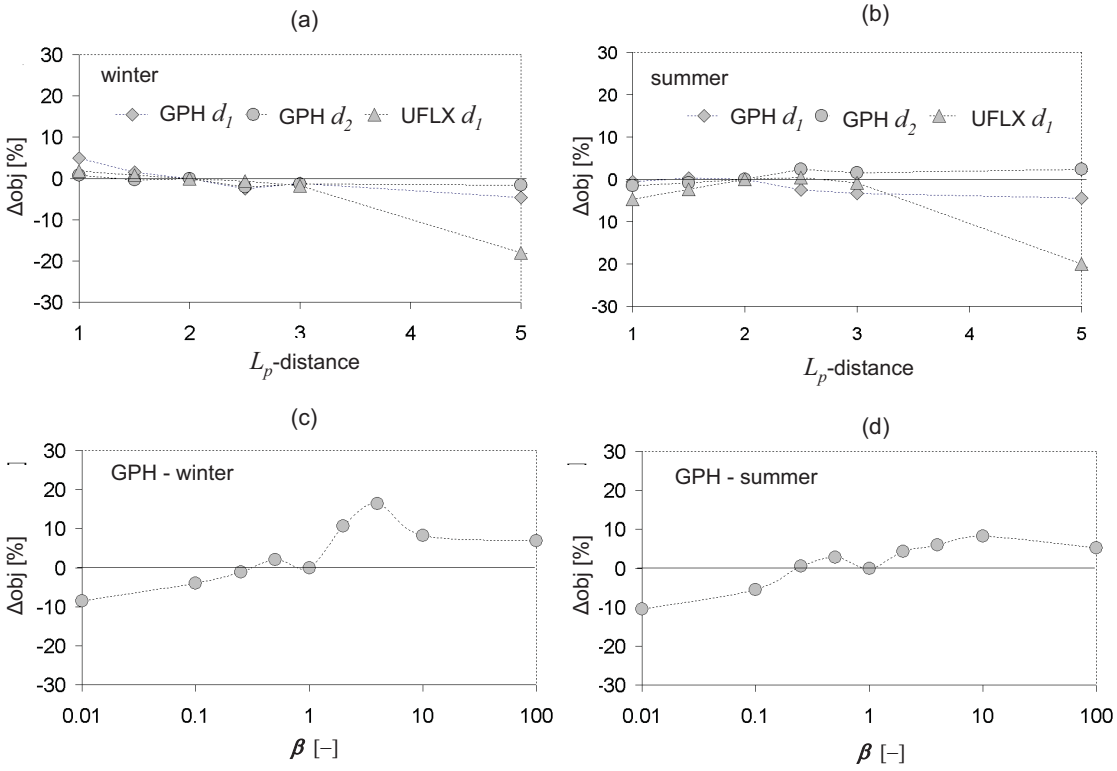



**Figure 4.7:** Influence of the distance parameters ($p_1$, $p_2$ and $\beta$) on the model performance in comparison to a reference algorithm.

taken to estimate the coefficient for the second and the third predictors. The variables are stepwise integrated to the distance formula. For predictor set A the first variable is the geopotential height, the second variable is the moisture flux and the third variable is the relative humidity. The order of the variable is chosen according to the general model performance based on a single predictor. The predictor configuration is determined by screening through an archive of nine candidates. The candidates (GPH, UFLX and RHUM) are selected from three pressure levels (1000, 850 and 700 hPa) of three observation times (12, 18 and 24 UTC). The predictor domains for the geopotential height and the moisture flux are listed in Table 4.1. A medium-sized domain (15° x 15°) is selected for the relative humidity and for all study regions. The location of the domain center is identical to the location of the study region. The predictor weight is determined by increasing the value of the coefficient in small equidistant steps. To simplify the estimation of the predictor coefficients, the single predictor distances are normalized to an interval between zero and one.

The Euclidean distance and the function related to the Pearson correlation seem to be an appropriate choice for a distance function (Figure 4.7a and Figure 4.7b). The model performance is again compared to a reference algorithm based on standard properties for the distance parameters ($p_1 = 2$ and $p_2 = 2$). The model predictions for intensive winter precipitations are slightly improved if the Manhattan distance ($p = 1$) is selected. However, this improvement is almost negligible compared to the

**Table 4.2:** Optimum distance parameters for analog forecasting derived with the 1000 hPa height field (GPH) and the 850 hPa zonal moisture flux field (UFLX).$d_1$ = Euclidean distance, $d_2$ = distance related to the Pearson correlation.

| catchment | season | GPH | | | UFLX |
|---|---|---|---|---|---|
| | | $p_1$ | $p_2$ | $\beta$ | $p_1$ |
| Mean | wi | 1 | 1 | 4 | 2.5 |
| | su | 1.5 | 2.5 | 10 | 1 |
| Nahe | wi | 2 | 1 | 2 | 1.5 |
| | su | 1 | 1 | 0.5 | 1.5 |
| F. Saale | wi | 1 | 1.5 | 4 | 2 |
| | su | 5 | 1.5 | 0.25 | 2 |
| F. Mulde | wi | 1 | 1 | 100 | 1 |
| | su | 1.5 | 2 | 10 | 3 |
| Iller | wi | 3 | 1 | 10 | 1 |
| | su | 5 | 1.5 | 0.25 | 2.5 |

results mentioned in the previous sections.

The introduction of the mixed distance can more enhance the forecasts in particular for the prediction of winter extremes (see Figure 4.7c and Figure 4.7d). In these figures the model predictions are compared to the predictions of an algorithm based on an equally weighted mixed distance ($\beta = 1$). However, it is crucial to find the optimal parameter value for $\beta$, otherwise the forecast with a single distance function outperforms the forecast based on the mixed distance. Furthermore, the model prediction can only be improved if the forecasts of two algorithms based on ordinary distance functions perform similar. For instance, for the Freiberger Mulde the algorithm based on an ordinary distance function performs as good as the algorithm based on the mixed distance which is indicated by the large value for the distance weight $\beta = 1$ given in Table 4.2.

The integration of the second and third predictor can clearly improve the prediction of intensive precipitation (see Figure 4.8). In particular the moisture flux has a similar information content than the geopotential height indicated by the same values of both predictor coefficients in Table 4.3. The predictor coefficient of the moisture flux determined for the Freiberger Mulde is even higher than the coefficient of the height field. The relative humidity is less important for the prediction of intensive precipitation. Nevertheless, the integration of the relative humidity slightly improves the forecast of intensive precipitation even for the winter. The predictor configuration determined for the second variable indicates that the information content of the second predictor is higher if a variable is not taken from the same pressure level and observation time like the first predictor (see Tables 4.3 and 4.4). For example, the best configuration for the Nahe is the 1000 hPa height field at 24 UTC and the 700 hPa moisture flux field at 18 UTC. Predictors of other pressure levels and observation times (like 500 hPa, 6 UTC) were also tested as candidates for the second variable, but their performance was generally lower. This result indicates

**Table 4.3:** Predictor configuration (time and level) and predictor weights ($a_1$, $a_2$ and $a_3$) determined for predictor set A.

| | | GPH | | UFLX | | RHUM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| site | seas. | time [UTC] | level [hPa] | time [UTC] | level [hPa] | time [UTC] | level [hPa] | $a_1$ | $a_2$ | $a_3$ |
| Mean | wi | 24(18) | 1000 | 18 | 700 | 18 | 700 | 1 | 1 | 0.5 |
| | su | 24(18) | 1000 | 12 | 700 | 18 | 700 | 1 | 1 | 0.5 |
| Nahe | wi | 24 | 1000 | 18 | 700 | 18 | 700 | 1 | 0.5 | 0.4 |
| | su | 24 | 1000 | 18 | 700 | 18 | 1000 | 1 | 1 | 0.4 |
| F. Saale | wi | 24 | 1000 | 12 | 700 | 18 | 700 | 1 | 1.2 | 0.5 |
| | su | 24 | 1000 | 12 | 700 | 18 | 700 | 1 | 1 | 0.5 |
| F. Mulde | wi | 18 | 1000 | 18 | 700 | 18 | 700 | 1 | 1.5 | 1 |
| | su | 18 | 1000 | 18 | 700 | 18 | 700 | 1 | 1.5 | 0.5 |
| Iller | wi | 18 | 1000 | 24 | 700 | 18 | 700 | 1 | 1 | 1 |
| | su | 18 | 1000 | 12 | 700 | 18 | 700 | 1 | 0.4 | 0.25 |

that there is an optimal distance in terms of the observation time and the pressure level between two predictors for the identification of analogs.

The investigation in this section demonstrates that the integration of humidity related variables clearly improves the prediction of intensive precipitation and that the selected multivariate predictor set seem to be appropriate for the description of the actual state of the atmosphere. However, there are still possibilities to test further predictors like air flow indices (see Section 6.3.4) which might improve the predictions. It is also very likely that intensive precipitation depends not only on

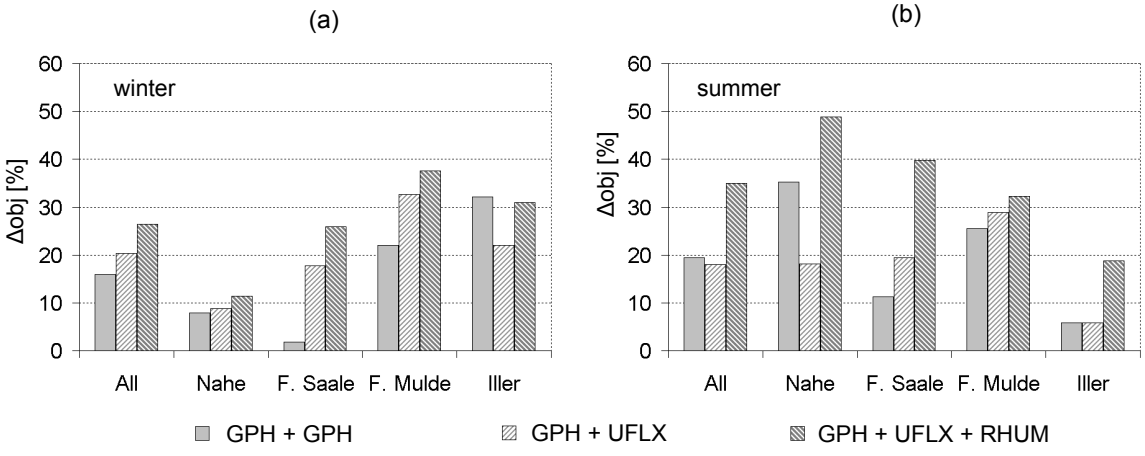



**Figure 4.8:** Influence of the predictor combinations on the model performance in comparison to a reference algorithm based on the geopotential height (GPH). UFLX = zonal moisture flux, RHUM = relative humidity.

**Table 4.4:** Predictor configuration (time and level) and predictor weights ($a_1$, $a_2$) of predictor set B.

| catchment | season | GPH time [UTC] | level [hPa] | GPH time [UTC] | level [hPa] | $a_1$ | $a_2$ |
|---|---|---|---|---|---|---|---|
| Mean | wi | 24(18) | 1000 | 12 | 700 | 1 | 0.5 |
| | su | 24(18) | 1000 | 12 | 700 | 1 | 0.6 |
| Nahe | wi | 24 | 1000 | 12 | 700 | 1 | 0.8 |
| | su | 24 | 1000 | 12 | 700 | 1 | 0.8 |
| F. Saale | wi | 24 | 1000 | 12 | 700 | 1 | 0.4 |
| | su | 24 | 1000 | 12 | 850 | 1 | 0.8 |
| F. Mulde | wi | 18 | 1000 | 12 | 700 | 1 | 0.6 |
| | su | 18 | 1000 | 18 | 700 | 1 | 0.6 |
| Iller | wi | 18 | 1000 | 12 | 700 | 1 | 0.4 |
| | su | 18 | 1000 | 12 | 700 | 1 | 0.4 |

the actual state of the atmosphere. Additional information which describe the preconditions e.g. the origin and the track direction of a low pressure system might be also relevant for the identification of suitable analogs. The analysis in this section also illustrate that the estimation of appropriate distance parameters and predictor coefficients is not straightforward and that it is influenced like the other criteria by a number of factors.

### 4.6.4 Number of Analogs and Probability Functions

The optimal number of the most similar analogs and the best cumulative distribution function for describing the forecast uncertainty is usually not known. To investigate this issue, various sets of the most similar analogs $k = [2, 4, ...., 100]$ are used for the estimation of the parameters of the exponential and the mixed exponential distribution. The percentiles $\boldsymbol{q_t} = [0.01, 0.02, ...., 0.99]$ of the cumulative distribution function are taken as the decision threshold to calculate the value of the objective function. It is compared to the score of a reference algorithm which is based on an empirical sample of the 60 most similar analogs. The value of the objective function for the reference algorithm is determined by selecting the following quantiles $\boldsymbol{q_t} = [1/60, 2/60, ..., 60/61]$. The investigation is performed for the predictor set A based on the model settings which are presented in the previous section.
The curves in Figures 4.9a and 4.9b illustrate that the forecast value of the analog method strongly depends on the most similar analogs. In winter, at least 30 analogs and in summer more than 40 analogs should be used for the parameter estimation to provide a valuable forecast for intensive precipitation. The mixed exponential distribution outperforms the empirical and the exponential distribution in particular for summer extremes, when the algorithm selects a number of dry days even for the prediction of extremes. Note that the model improvement due to the selection of the theoretical distributions is mainly the result of the higher number of decision thresholds selected for the theoretical distributions (99 compared to 60 for the empirical distribution), since the value of the objective function is sensitive to the number of thresholds in particular for extremes. The optimal number of analogs and optimal cumulative distribution function determined for the four study regions is listed in Table 4.5.

### 4.6.5 Selection Rule

In analog forecasting a selection rule is often formulated separating the data into two or more subsets to speed up the algorithm and to account for the intra-annual variability of predictor and predictand. Usually, the selection process is restricted by a moving window which is defined for each year of the investigation period and

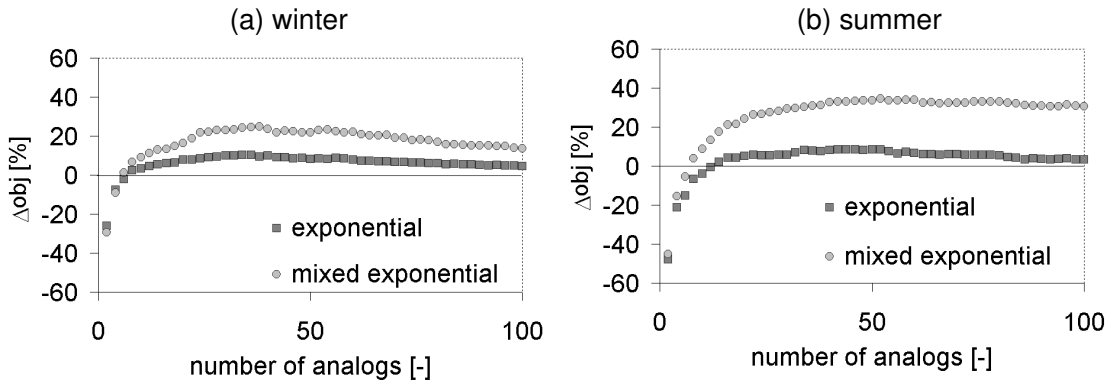



**Figure 4.9:** Influence of the number of most similar analogs on the model performance; Predictor set A.

**Table 4.5:** Optimal number of analogs n, cumulative distribution functions (EX = exponential, MEX = mixed exponential) and the length of the moving window, predictor set A.

| catchment / study region | season | n [-] | distrib. function | window lenght [days] |
|---|---|---|---|---|
| Mean | wi | 38 | MEX | 310 |
| | su | $\geq 40$ | MEX | 160 |
| Nahe | wi | 24 | MEX | 310 |
| | su | $\geq 40$ | MEX | 160 |
| F. Saale | wi | 52 | EX | 310 |
| | su | $\geq 40$ | MEX | 130 |
| F. Mulde | wi | 44 | MEX | 310 |
| | su | 60 | MEX | 220 |
| Iller | wi | 34 | MEX | 260 |
| | su | 48 | MEX | 160 |

centered on the target day. The window length defines the size of the subset used for the selection of analogs which is mostly not larger than two or three months (see e.g. Lorenz, 1969; Matulla et al., 2007). In this case, the algorithm selects e.g. for a winter day only analogs which occurred also in winter. However, weather states leading to intensive precipitation are rare and the selection of a small window can miss an important analog. To determine an appropriate size of the moving window, different lengths of the window are tested in this section. The investigation is performed for an algorithm based on the predictor set A and the model settings listed in the aforementioned sections. The model predictions are compared to a reference algorithm without any restrictions.

The influence of the selection rule is illustrated in Figures 4.10a and 4.10b. The optimum window length determined for the four study regions is listed in Table 4.5.

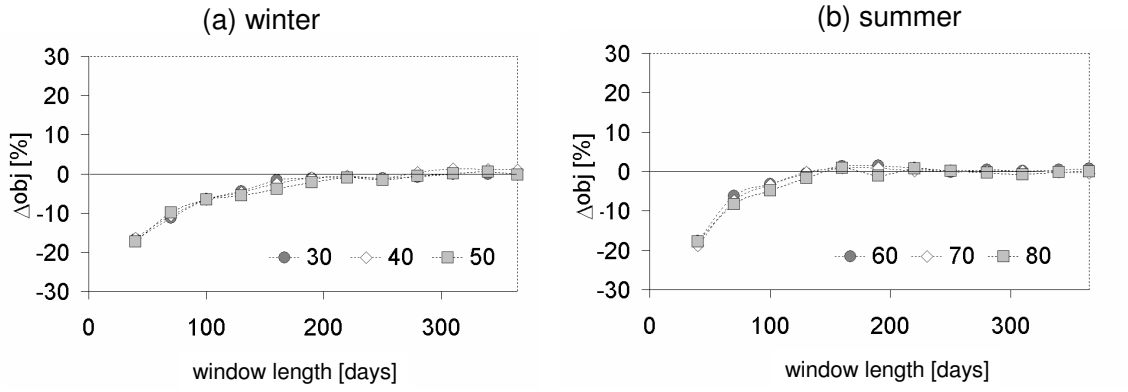



**Figure 4.10:** Influence of the selection rule on the model performance. The relative change of the objective function is given for different sets of the most similar analogs (30, 40, ..., 80) which performed best; Predictor set A.

The model predictions are marginally improved due to the restriction. The most important factor is that a window length is selected which is not too small; otherwise the forecast performance of the algorithm significantly decreases. It is recommended to choose a moving window of at least six months (> 180 days) for the prediction of winter extremes and a slightly smaller window (> 130 days) for summer events. This result is different to the aforementioned studies which selected a window of three or less months. However, if the selection process is analyzed for a forecast day, the algorithm without any restriction already distinguishes suitably between summer and winter situations (not shown here). A further influencing factor for the selection rule is the strength of the intra-annual variability of the predictand which is usually weaker for maritime influenced study regions compared to study regions located in a climate zone with a high continental or subhumid character.

## 4.7 Forecast Skill and Value

### 4.7.1 Optimal Use of a Probability Forecast

The model parameters specified in the aforementioned sections for the study region WG and MG and for the predictor set A are taken to test the model transferability of the approach for the neighboring catchments (Ruhr, Upper Sieg, Prüm and Upper Main). The value and the skill of the methodology is calculated for the relative value defined in Equation 4.21 for the following set of decision thresholds $\boldsymbol{p_t} = [0.01, 0.02, ..., 0.99]$. Note that for the study regions located in EG and MG the investigation period is slightly longer (1951 to 2005) than for the other study regions (1958 to 2001). Thus, the model performance is not only evaluated for the cross-validation period (1958 to 2001). In the following the mean performance of the downscaling approach is calculated for all study regions and for the entire investigation period.

The forecast value of the analog method in dependence on the cost-loss ratio is shown in Figure 4.11. In the upper left graph the relative value of the model predictions is illustrated for moderate winter precipitation if the forecaster takes the 50 %, the 30 % or the 10 % probability threshold for decision making corresponding to the median, the 70 %-quantile and the 90 %-quantile of the forecast. The selection of the median forecast is poor in particular for users with low cost-loss ratios (small cost and high loss). This warning strategy only produces a moderate number of alarms but the events are often missed increasing the losses for this user. On the other hand, the selection of the 30 % and the 10 % probability threshold is more suitable for them since more alarms are triggered by the warning system rising the costs for these users but minimizing their losses.

This example illustrates that the selection of an appropriate decision threshold strongly depends on the cost-loss ratio of an individual user. However, the optimal use of the probability forecast depends on the return frequency of an event, as well (see Figure 4.11b). The decisions are valuable if the forecaster selects a threshold which is clearly lower compared to the decision thresholds for moderate events. If the decision maker selects e.g. the 2 % probability threshold, his decisions are valuable for a number of users. In this case, the warning system triggers a number of alarms

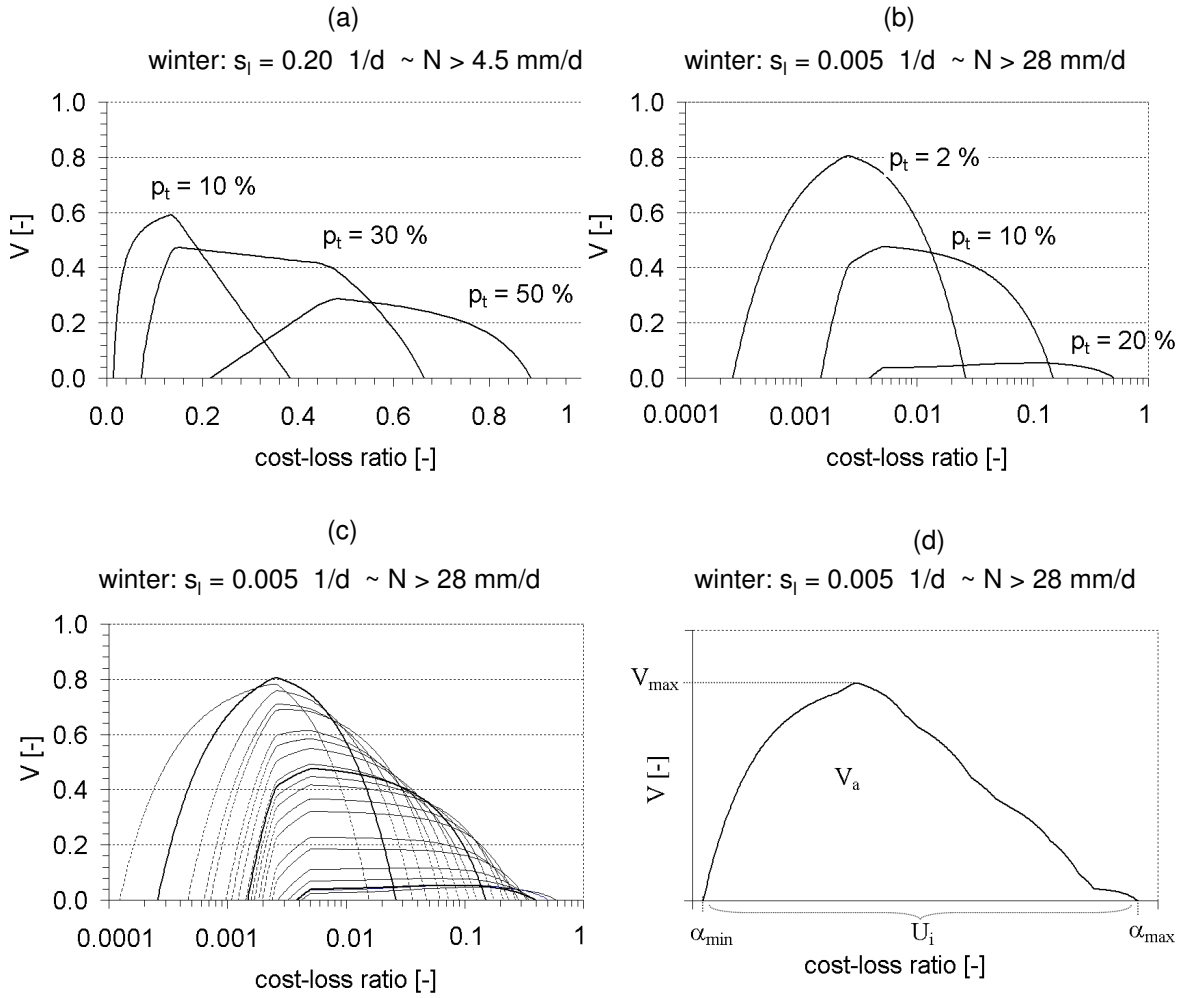



**Figure 4.11:** Mean forecast value of the probability forecast in dependence on the cost-loss ratio. a and b: forecast value for three probability thresholds for moderate and intensive winter precipitation. c: forecast value for the entire set of probability thresholds for intensive winter precipitation. d: Maximum forecast value of the probability forecast with three performance measures: $V_{max}$ = maximum value, $V_a$ = area of forecast value, $U_i$ = user interval $[\alpha_{min}; \alpha_{max}]$. Investigation period (1951 (1958) to (2001) 2005), predictor set A, all catchments

increasing the hit rate and minimizing the losses. However, this warning strategy is only suitable for users with low cost-loss ratios. For users with high cost-loss ratios the forecaster must select a higher decision threshold e.g. the 10 % probability threshold. In this example, a logarithmic scale is selected to focus more on users with low cost-loss ratios. In Figures 4.11a and 4.11b the forecast value is given for a limited number of decision thresholds. The forecast value for the entire set of decision thresholds is illustrated in Figure 4.11c. If the decision maker selects the optimal threshold for an individual user, he can maximize the forecast value for this user. If he selects the optimal decision thresholds for all users, the maximum value of the probability forecast can be determined which is described by the envelope curve

drawn in Figure 4.11d. The optimal use of the probability forecast ensures that the value of the probability forecast is always higher or at least the same compared to a non-probability forecast such as the ensemble mean for the same model. This example clearly demonstrates the advantage of a probability forecast compared to other forecasting techniques. The distribution of the maximum value shows that the highest forecast value is obtained for users with cost-loss ratios similar to the return frequency. For these users the low-skill reference forecasts are the weakest and they can gain a high benefit from the model predictions. The reference forecast is the strongest for cost-loss ratios tending to zero or 1. For example, the analog forecast has only a value for cost-loss ratios ranging between 0.0001 and 0.5 for the prediction of intensive winter precipitation. If a user is outside this interval, he gains no benefit from the model predictions. In this case, it is recommended that the low-skill reference forecast is selected for the decision making instead of the model predictions. Otherwise, the decisions based on the model outcomes are poorer than the low-skill warning strategy and the average expense of the model predictions are even higher than the average expense of the low-skill reference forecast. If only monetary costs are considered, the decision maker will even waste money due to his poor decisions. This is a further illustrative example demonstrating the importance

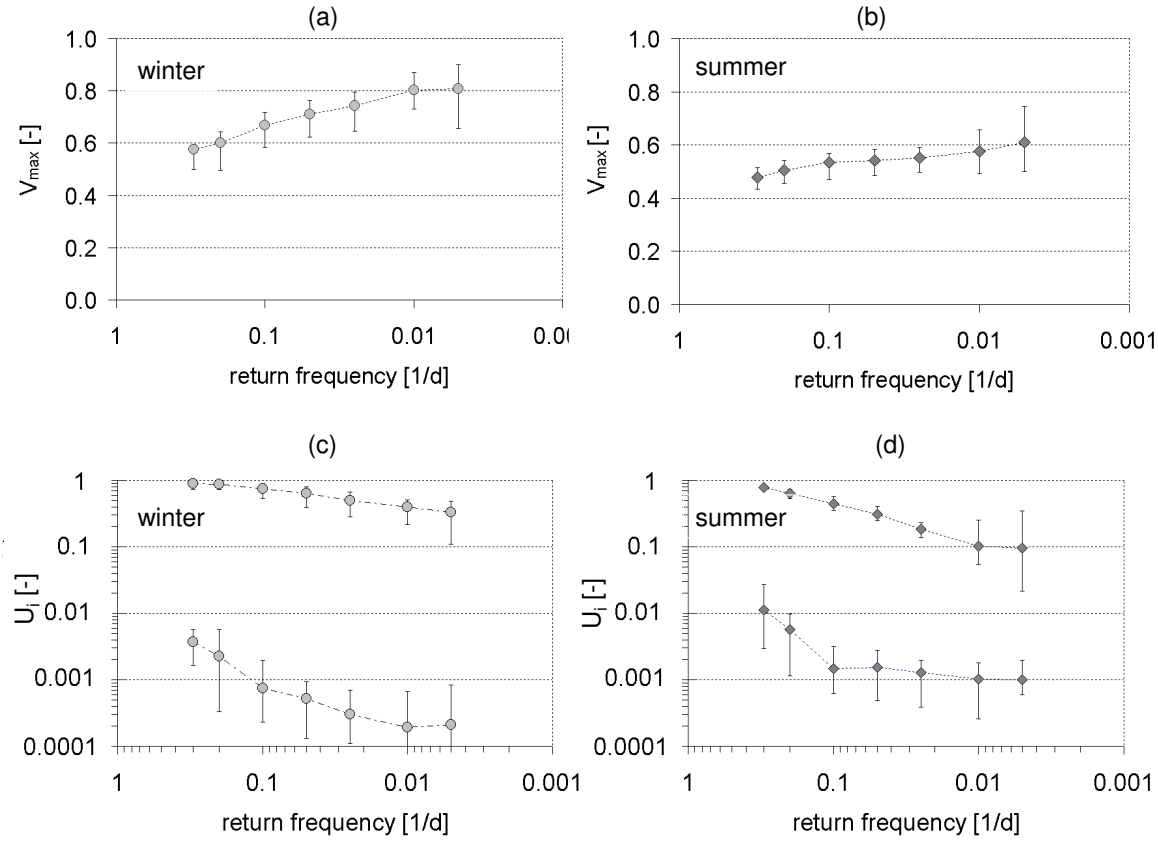



**Figure 4.12:** Maximum value (a and b) and user interval (c and d) in dependency on the return frequency. Investigation period (1951 (58)- (01) 2005), predictor set A, all catchments.

of cost-loss approaches for the analysis of decision-making processes and showing that a low-skill reference forecast can even outperform model predictions for certain users.

To determine the value of the model predictions over the entire range of the predictand, three performance measures are introduced: the maximum value $V_{max}$, the user interval $U_i$=[$\alpha_{min}$; $\alpha_{max}$] and the area of the forecast value $V_a$ (see Figure 4.11d). The maximum value represents the peak value of the envelope curve describing the maximum possible value of the model predictions. The user interval measures the user range gaining a benefit from the model predictions. The larger the interval is the more users can profit from the forecast. The area of the forecast value describes the area between the envelope curve and the line indicating zero value. It is approximated by the objective function proposed in Section 4.3. If the objective function is maximized, the area of the forecast value increases and at least one of the other two performance measures rises as well.

The maximum value in dependence on the return frequency of an event is given in Figures 4.12a and 4.12b illustrating a high model performance for intensive precipitation. However, the forecast value of analog forecasting also depends on the season. In summer, the value is lower than in winter indicating that the detection of summer extremes is weaker due to the higher stochastic character of those events. The curves in Figures 4.12a and 4.12b represent the mean maximum value of all catchments selected in this study. The error bars characterize the lowest and highest maximum value for the test sites illustrating a high forecast performance for all catchments and indicating the suitable model transferability to the neighboring catchments. However, the range of the maximum value increases with the precipitation amount. It is the largest for intensive precipitation indicating a high spatial variability of the forecast value, in particular for the extremes. Thus, the forecast performance depends also on the catchment. In future studies the spatial variability of the forecast performance must be also taken into account to investigate how large the influence is for decision making process.

The user interval determined for summer and winter is given in Figures 4.12c and 4.12d. The larger the area is between both curves, the more users can profit from the model predictions. The curves illustrate that the analog forecast is skillful for a number of users over the entire range of the predictand. Since the performance of the low-skill reference forecasts is the weakest for cost-loss ratios similar to the return frequency, the user range is shifted to lower cost-loss ratios with increasing precipitation amount. The forecast value in terms of the user interval is the largest for the prediction of winter extremes and a wide range of users can profit from the forecast for all catchments. This demonstrates again the suitable transferability of the approach.

### 4.7.2 Intra-annual Variability

Since the detection of extremes is clearly poorer in summer, the maximum value and the user interval is characterized by a strong intra-annual variability (Figure 4.13a and Figure 4.13b). In winter, the model predictions are strong for low and for intensive precipitation characterized with a high hit rate and a small number of false alarms (Table 4.6). In March the forecast performance clearly decreases. In summer the forecast performance is the lowest but the predictions are still skillful and have a moderate value for a number of users. But due to the variation of the user interval, the predictions are not skillful for all individual users throughout the entire year. For these specific users the decision maker must take into account the poorness of the model predictions in summer, otherwise the average expense of the analog forecast is higher compared to the low-skill reference forecast.

The different examples in this section demonstrate that the predictions of the analog method are skillful and have high forecast value for a wider range of users and over the entire range of the predictand. Nevertheless, a suitable forecast value is only obtained if the decision maker knows how to use a probability forecast in cost-loss ratio situations. The optimal use of a probability forecast depends on several factors like the cost-loss ratio and the return frequency. It is crucial for the design of a warning system that these factors are taken into account to provide finally valuable decisions for all users in a region of interest.

**Table 4.6:** Binary outcomes for the prediction of extremes ($a$ = number of hits, $b$ = number of false alarms, $c$ = number of misses, $d$ = number of correct rejections) in WG. $HR$ = hit rate, $FA$ = false alarm rate, $p_t$ = optimal probability threshold, predictor set A. Cross-validation period (1958 - 2001).

| catchment | season | $a$ | $b$ | $c$ | $d$ | $HR$ | $FA$ | $p_t$ |
|---|---|---|---|---|---|---|---|---|
| Mean | year | 72.0 | 2622.3 | 9.5 | 13367.3 | 0.88 | 0.16 | 0.02 |
| | wi | 39.0 | 577.0 | 2.0 | 7400.0 | 0.95 | 0.07 | 0.04 |
| | su | 33.5 | 1823.3 | 7.8 | 6188.5 | 0.81 | 0.23 | 0.01 |
| Upper | year | 70 | 1291 | 11 | 14699 | 0.86 | 0.08 | 0.03 |
| Sieg | wi | 40 | 504 | 1 | 7473 | 0.98 | 0.06 | 0.05 |
| | su | 32 | 1505 | 9 | 6507 | 0.78 | 0.19 | 0.01 |
| Nahe | year | 76 | 3199 | 5 | 12791 | 0.94 | 0.20 | 0.01 |
| | wi | 39 | 491 | 2 | 7486 | 0.95 | 0.06 | 0.03 |
| | su | 36 | 1934 | 6 | 6077 | 0.86 | 0.24 | 0.01 |
| Prüm | year | 72 | 3945 | 10 | 12044 | 0.88 | 0.25 | 0.01 |
| | wi | 40 | 709 | 1 | 7268 | 0.98 | 0.09 | 0.04 |
| | su | 31 | 2021 | 10 | 5991 | 0.76 | 0.25 | 0.01 |
| Ruhr | year | 70 | 2054 | 12 | 13935 | 0.85 | 0.13 | 0.04 |
| | wi | 37 | 604 | 4 | 7373 | 0.90 | 0.08 | 0.04 |
| | su | 35 | 1833 | 6 | 6179 | 0.85 | 0.23 | 0.02 |

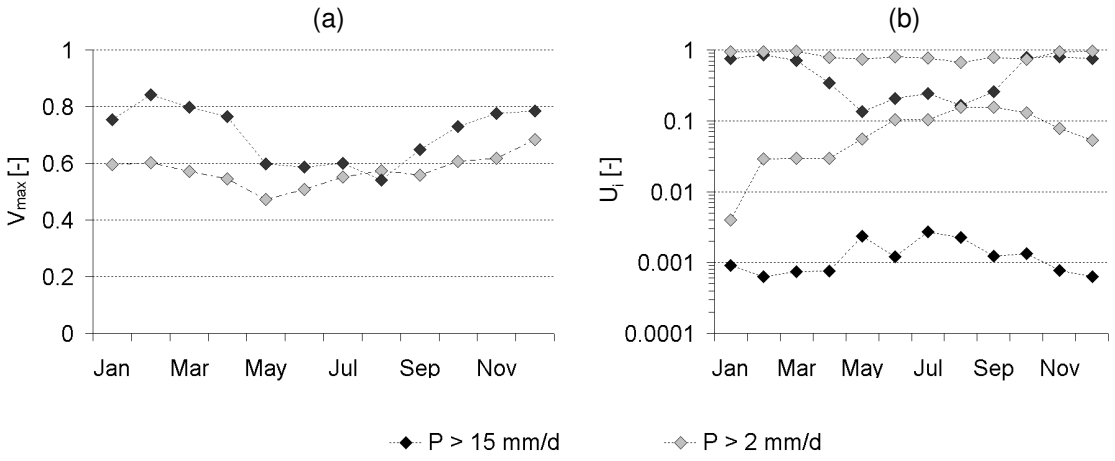



**Figure 4.13:** Intra-annual variability of the forecast value for two precipitation thresholds. Rhine basin (Ruhr, Sieg, Prüm and Nahe). Cross-validation period (1958 - 2001).

## 4.8 Summary and Conclusions

In this chapter an analog forecasting algorithm was proposed for the prediction of intensive daily areal precipitation. The methodology was performed for all catchments selected in this thesis which are located in four study regions in Germany. The parameters of the algorithm were determined for every study region and for two different parameter sets. The first predictor set is a combination of pressure and humidity related variables designed for short-term predictions. Since humidity related variables are less reliable for medium-term forecasts a second predictor set was proposed. This set consists of two geopotential height fields differing in observation time and pressure level.

Furthermore, a distance function was proposed for the comparison of two atmospheric patterns. This distance function represents a mixture of two $L_p$-distances to link two kinds of "similarity" between atmospheric patterns: pattern closeness and pattern similarity. It was demonstrated that the mixed distance function can slightly improve the predictions of intensive winter precipitation if model predictions based on the ordinary distance functions perform similarly. In future investigations the model performance must be also determined over the entire range of the predictand to see whether the selection of the analogs based on the mixed distance function can improve the model predictions also for light and moderate precipitation. The investigation also indicated that the selection of different $L_p$-distances like the Manhattan distance is of minor importance for the downscaling of intensive daily precipitation and that the Euclidean distance and the distance function related to the Pearson correlation already represents a suitable first choice for the description of pattern closeness and pattern similarity.

It was demonstrated that the model performance depends on a number of further criteria like the predictor domain or the number of the most similar analogs. These criteria have to be specified a priori by the user before starting the search

algorithm. Since a guideline for an appropriate selection of these criteria is not available, a comprehensive overview was given to investigate their influence on the model performance. Finally, some recommendations were made for the selection of appropriate criteria. However, due to the interaction of the model parameters a manual specification of the different criteria is not straightforward and a lot of user experience is needed. To tackle this problem, an automated optimization algorithm is proposed in Chapter 6 to specify an appropriate predictor domain and an optimum distance for the selection of analogs. The basic framework of a cost-loss approach was also presented in the chapter for a two action ("*protection*", "*no protection*"), two state ("*adverse weather*", "*no adverse weather*") decision making problem. The importance of the approach for the analysis of cost-loss situations was highlighted by means of several examples. The cost-loss approach was also taken to determine the forecast skill and value of the model predictions indicating the high potential of the analog method for the detection of intensive daily areal precipitation for a wide range of users. However, it is crucial for the goodness of the model predictions that the decision maker knows how to apply a probability forecast in cost-loss situations. Therefore, the optimal use of a probability forecasts was demonstrated so that a decision maker can maximize the benefit for the individual user. Additionally, it was shown that the optimal use of the probability forecast is influenced by several factors e.g. the cost-loss ratio and the occurrence frequency. These factors must be taken into account by a forecaster for the optimization of the decision making process to provide finally a high valuable forecast for all end users of the region of interest.

The cost-loss approach was also applied for the formulation of an objective function to maximize the model performance of a binary warning system ("*warning*", "*no warning*") based on a probability forecast. The spatial transferability of the approach was tested by transferring the calibrated model to the neighboring catchments to evaluate the robustness of the estimated model settings. This test was necessary since in the operational mode the precipitation model must be also transferable to further catchments which are not part of this thesis to provide a precipitation forecast for all catchments located in a region of interest. The transfer of the model parameter for the prediction of intensive precipitation in the neighboring catchment indicated a suitable model transferability. In the following Chapter a more detailed investigation of the transferability is presented. It is illustrated that the choice of the objective function strongly influences the model parameters and the model transferability.

Three performance measures based on the cost-loss approach were introduced to illustrate the forecast value over the entire range of the predictand. It was shown that the forecast performance in terms of the forecast value is characterized by an intra-annual variability in particular for the extremes. This temporal variability of the forecast performance must be accounted for by a forecaster to optimize his decision and to provide finally a valuable forecast for a number of users. The results also indicate that the forecast value varies from catchment to catchment. However, this spatial variability of the forecast performance was not investigated in this chapter in detail and should be accounted for in future work to investigate how the spatial variability of the forecast performance affects the decision making process.

Note that the analog method presented in this Chapter represents not only a

practical approach, but also a valuable methodology for the description of the forecast uncertainty of daily precipitation. The selection of this approach was necessary since the available forecasting approaches commonly applied by the regional flood forecasting centers in Germany are limited, as described in Chapter 2. Nevertheless, there are a several features of the presented approach which should be addressed in future works to develop a more robust model and to improve the model predictions. One limitation is the large dimensionality of the search algorithm due to the large predictor domains estimated for the multivariate predictors and the selection of the mixed distance function. To reduce the dimensionality of the search, a two-step search algorithm can be applied by selecting in the first step pressure related variables and in the second step humidity related variables (see e.g. Bontron and Obled, 2003). A further technique for the reduction of the dimensionality is the optimization algorithm presented in Chapter 6.

In the following chapters the methodology is compared to further statistical downscaling techniques to evaluate the strengths and the weakness of the forecasting approach. Furthermore, some ways are illustrated to improve the model development and the model performance.

# 5 Classification

The analysis of past floods in Central Europe clearly indicates a relationship between large-scale weather phenomena and long-lasting large-area precipitation. An example is the Vb-weather pattern which caused the Elbe and Danube flood in 2002 (for a description of the event see Philipp and Jacobeit, 2003). Actually, this weather pattern describes the track of a cyclone which moves from the Mediterranean Sea (the Gulf of Genoa) along the Eastern Alps over Austria and Czech Republic to Poland. In summer, the Vb-weather pattern and its relating patterns are probably the main flood producing patterns. They were also responsible for the Odra flood in 1997 (Malitz and Schmidt, 1997) and the Northern Alp flood in 2005 (Rudolf et al., 2005). However, in winter large river floods can also occur when zonal flows dominate so that moist air masses are transported from the Northern Atlantic to Central Europe. The Rhine floods in 1993 and in 1995 are two examples which were caused by such a pattern (Engel, 1997). Those weather patterns can be also characterized by a sharp temperature increase so that the impact of intensive rainfall can be amplified by snow melting in mountainous regions.

Due to the obvious link between the large-scale atmospheric circulation and local surface variables it is quite common to apply classification techniques. The main objective is to understand the relationship between the large-scale information and the local variable (Yarnal, 1992). This goal is attained by classifying the atmospheric circulation into groups and linking these groups to the local consequences to identify e.g. weather (circulation) patterns with a high risk of intensive precipitation or flooding. There are a lot of studies which related surface variables like precipitation or temperature to large-scale weather patterns (for recent application see Samaniego and Bárdossy, 2007). A detailed introduction into weather classification is given by Yarnal (1992) with good overview about many case studies for a number of surface variables. A brief survey of weather classification is also proposed by Wilby and Wigley (1997) and Zorita and von Storch (1997).

## 5.1 Subjective and Objective Classification

The classification approaches are usually separated in subjective (manual) and objective (automated) techniques (Yarnal, 1992). In subjective classification the knowledge and the judgment of an expert is incorporated into the classification procedure. A well-known example of a subjective classification is the catalog of Hess and Brezowsky (1969) which describes the Grosswetterlage of Central Europe. The Grosswetterlage is a sequence of days where the main features of the large-scale weather remain constant over Central Europe. The Lamb weather types are a further subjective classification which is a catalog of daily weather maps developed for the British Isles (Lamb, 1972).

There are many reasons for the use of subjective classification (see Yarnal, 1992). However, subjective classifications are restricted to certain geographical regions so that the classification is not transferable to other regions (Yarnal, 1992). A further limitation is that a subjective classification is not reproducible by a second investigator. Thus, the catalog of weather patterns must be reformulated again to avoid an inconsistent time series (Yarnal, 1992). Subjective classifications are also often characterized by a large number of weather patterns which make a robust estimation of statistical properties for a surface variable more difficult in particular for rare patterns. For example, the catalog of the Grosswetterlagen consists of 30 classes and the Lamb scheme divides the atmospheric flow processes into 29 weather types. To decrease the number of classes, weather patterns with similar large-scale behavior and local consequences must be merged together.

To overcome the problems in subjective classification, a variety of objective classification schemes have been developed (see e.g. Wilby and Wigley, 1997; Bárdossy et al., 2002) and they are applied for different geographical regions. Yarnal (1992) and Bárdossy et al. (2002) noticed that in subjective and objective classification the local surface variable is often not included in the classification procedure. The circulation patterns are independently derived from the surface variable of interest and the statistical properties of this variable are calculated for each pattern after compiling the classification. However, for downscaling applications it is desirable to develop classifications which incorporate the surface variable in the classification procedure to explain the variability of the predictand. Such a technique has been proposed by Bárdossy et al. (2002) for objective classification which specifies large-scale daily circulation patterns for the variable of interest. The methodology was applied for precipitation and temperature downscaling for a region in Germany and in Greece. Furthermore, the classification was also tested for downscaling daily discharge differences to determine flood producing large-scale circulation patterns for two study regions located in France and in Spain (Bárdossy and Filiz, 2005). The results of these studies indicated that the classification can distinguish between dry (cold) and wet (warm) circulation patterns and that the circulation patterns seem to be physically reasonable.

## 5.2 Why Objective Classification in Weather Forecasting?

The aforementioned studies illustrate that classification techniques are widely used in climate modeling. However, in operational weather forecasting regression-based approaches seem to be dominating (see e.g. Brunet et al., 1988; Lemcke and Kruizinga, 1988; Wilson and Vallée, 2002) and classification techniques are less represented. Nevertheless, there are many arguments which make their use for operational purposes very attractive. Some of these arguments are as follows:

1. The output of NWP models can be used to determine daily circulation patterns for different lead times to indicate days with a high or low risk of local extremes.

2. The conditional cumulative distribution function of the surface variable of interest can be specified for a given weather pattern. This function can be

used to describe the forecast uncertainty for the surface variable of interest.

3. The forecast performance of weather prediction models can be analyzed for a given weather pattern. This kind of technique is not often applied in forecast verification. The performance of the prediction model is usually determined for different seasons or locations. Furthermore, the model predictions of a classification technique can be also taken as a reference forecast to evaluate the skill of further weather prediction models.

4. In the case of the precipitation model proposed in Section 1.2, the classification is also essential for the second step of the downscaling algorithm, the disaggregation of the daily information into precipitation realizations with a high spatiotemporal resolution. The statistical properties of precipitation such as intensity or moving direction of precipitation fields are strongly influenced by the large-scale weather pattern. To allow a conditional simulation e.g. of precipitation, circulation patterns are needed. There are many studies which demonstrated the applicability of large-scale weather (circulation) patterns for the conditional simulation of surface variable (see for daily precipitation e.g. Bárdossy and Plate, 1992; Bárdossy, 1997; Stehlík and Bárdossy, 2002).

5. The model parameters of the analog method can be conditioned on a circulation pattern. For example, in Section 6.3.4 a selection rule is formulated which restricts the search of analogs by conditioning it on a circulation pattern. It is demonstrated that the model predictions can be slightly improved compared to a traditional approach and that a faster search algorithm can be created which is in particular helpful for the optimization of this algorithm.

Classification techniques are also characterized by some less pleasant features. Probably, their major disadvantage is that the continuous atmospheric flow processes and their relationship to the local variables are subdivided into discrete groups. If the classification is used for the prediction of surface variables, this can lead to the problem that there is an abrupt change of the forecast characteristics of a surface variable between two weather patterns. The discontinuity is in particular large, if the properties of the surface variables of two weather patterns are quite different.

### Model Purpose

The classification applied in this chapter should form a component of the precipitation model presented in Section 1.2. It transfers the output of a global NWP model to determine a time series of daily circulation patterns for the forecast horizon of interest. Before implementing the approach for operational purposes, we investigate in this chapter how effectively the approach can detect extremes. Thus, the main purpose is to shed light on the second and the third argument of the aforementioned list as follows:

1. The performance of the classification for the identification of circulation patterns with a high and low risk of intensive precipitation is analyzed.

2. The classification is taken to provide a probability forecast of daily areal precipitation for each day of the investigation period to determine the skill and the value of the classification for the detection of intensive precipitation.

3. The classification is also compared to the model predictions of the analog method to identify the strengths and the weaknesses of both downscaling techniques.

4. The conditional forecast performance of the analog method is determined for a particular circulation pattern.

In this chapter we would also analyze the transferability of the classification from a dependent data set to an independent data set. The model transferability of a statistical downscaling approach strongly depends on the choice of the objective function. Therefore, two objective functions are tested. The first function is based on a time series of daily areal precipitation. This optimization strategy should represent the intuitive choice of an investigator if the objective of the model development is the prediction of daily precipitation. The second strategy takes an indicator time series of daily discharge differences to allow for a more robust estimation of the model parameters. Both optimization strategies are used to compile a catalog of daily circulation patterns for the Rhine basin. Finally, the model performance of both classifications is analyzed for the prediction of daily areal precipitation for the same period of an independent data set.

To address these aims, the objective classification developed by Bárdossy et al. (2002) is applied to determine large-scale daily circulation patterns (CPs). The concept of this classification is based on fuzzy rules. In the following section the basic idea of the fuzzy rule-based classification is given and the objective functions used for classification are presented. The results of the investigation are shown in Section 5.4 and Section 5.5. The chapter is closed with a short summary and conclusions for further investigations.

## 5.3 Fuzzy Rule-Based Classification

### 5.3.1 Basic Methodology

In this section only a rough description of the fuzzy rule-based classification is proposed. For a comprehensive overview of the objective classification the reader is referred to Bárdossy et al. (2002). The development of the fuzzy rule-based technique is delineated in detail in Bárdossy et al. (1995). The basic principle of the automated classification can be described as follows:

Weather patterns of subjective classifications are often described by imprecise formulations. For example, the characterization of the Grosswetterlage TRM, the counterpart of the Vb-weather pattern, begins with the following sentence (Hess and Brezowsky, 1969): "*A trough over North and Central Europe is surrounded by higher pressure over the eastern part of the North Atlantic and western Russia ...*". The use of fuzzy-rules enables the investigator to incorporate such vague statements made by an expert. However, the definition of fuzzy rules by an expert is usually time-consuming and a lot of experience is needed. Instead of defining the fuzzy rules

in a subjective manner, the rules are specified by an optimization algorithm, so that even an investigator with less experience can obtain similar results compared to an expert.

The following steps must be performed to determine a catalog of daily CPs for a particular region:

1. The daily anomalies of a large-scale pressure related variable like the geopotential height or the mean sea level pressure is calculated (see Equation 3.2).
2. To describe the atmosphere state by imprecise statements, the daily anomalies are transformed into the following categories: "*high*", "*medium high*", "*medium low*", "*low*" and "*indifferent*".
3. For each CP a fuzzy rule is defined specifying the lows and the highs of the atmospheric circulation.
4. The degree of fulfillment is calculated for each rule.
5. The fuzzy rule with the highest degree of fulfillment is taken to define a CP for each day of the investigation period.
6. The objective function is calculated for the catalog of daily CPs.
7. Step 3 to 6 are repeated to find a suitable classification for the surface variable of interest.

The optimization algorithm for the definition of the fuzzy rules is based on simulated annealing (Aarts and Korst, 1989). A detailed description of the algorithm is given in Bárdossy et al. (2002). In Section 6.1.2 the basic idea of simulated annealing is presented for an optimization routine which has been developed for analog forecasting. In the following sections two objective functions are proposed for the classification.

### 5.3.2 Objective Functions

In Section 4.3 an objective function was introduced which measures the forecast performance of a binary warning system. This objective function was used to determine reasonable parameters for the analog method to maximize the forecast performance for the extremes. However, it was also mentioned in Section 4.5 that the manual optimization of the approach leads in some case to less plausible parameter settings. Thus, the parameters were replaced in a subjective manner by more reasonable parameters. This result already indicates that the determination of suitable parameters which are physically reasonable and which allow for high model transferability seems to be more complex. In this section we would like to address this topic in more detail.

A robust identification of model parameters strongly depends on the choice of the objective function. Therefore, two different objective functions are tested for the fuzzy rule-based classification. The first optimization strategy is based on daily areal precipitation. Unfortunately, this strategy has the disadvantage that the assumptions of robust parameter estimation can be violated. To enable robust parameter

estimation, the residuals which are the differences between the model predictions and the observations should fulfill several assumptions: (i) The mean value of the residuals is zero; (ii) the residuals are independently distributed (no serial correlation) and (iii) the variance of the residuals is constant and should not depend on the predictor. In the case of daily areal precipitation most of these assumptions are strongly violated (for a discussion see the following section).

To overcome this problem, an objective function is proposed which is based on an indicator time series of daily discharge differences. The selection of discharge differences has also a second reason. For hydrological applications like flood prediction, the main objective is the development of a classification which links the large-scale atmospheric flow to daily discharge for the detection of flood producing situations. If the concentration time of a catchment is taken into account, such a classification can be also used for the detection of precipitation extremes, since in the temperate climate zone, river flooding is mainly the result of long-lasting large-area heavy precipitation. Finally, a classification is designed for two purposes: (i) The detection of intensive daily precipitation and (ii) the detection of daily floods. In the following two sections a more detailed description of both optimization strategies is given.

### Optimization criterion based on daily areal precipitation

The objective function for the areal precipitation is based on a weighted mixture of two precipitation indices describing the wetness of a CP. It has been proposed in the investigation of Bárdossy et al. (2002).

The first precipitation index $PI_1$ is a measure which relates the conditional precipitation sum of a given CP $i$ to the average precipitation sum:

$$PI_1(i) = \frac{n_i}{n} \left( \frac{\sum_{k=1}^{n_i} y(k) \cdot n}{\sum_{t=1}^{n} y(t) \cdot n_i} - 1 \right)^2 \tag{5.1}$$

where $y(k)$ is the daily precipitation amount of day $k$ of a CP and $y(t)$ is the daily precipitation amount at day t. $n_i$ is the number of days with the same CP and $n$ is the total number of days of the investigation period. The ratio $n_i/n$ is the relative frequency of a CP defining the weight of a CP for the precipitation index.

The second precipitation index $PI_2$ is a measure relating the conditional occurrence of wet days of a CP to the average occurrence of wet days:

$$PI_2(i) = \frac{n_i}{n} \left( \frac{nw_i}{n_i} - \frac{nw}{n} \right)^2 \tag{5.2}$$

where $nw_i$ is the number of wet days of a CP and $nw$ is the total number of wet days. The value of both precipitation indices ranges between zero and infinite. The larger the indices are, the better the classification can differentiate between dry and wet days. An index equal to zero indicates a CP which is not different to the climatological average. The aim of the optimization is to maximize the differences between both functions to determine CPs whose precipitation properties deviate from the climatological average.

To account for the intra-annual variability of daily precipitation, the investigation period is divided into four periods according to the seasons. Then, the precipitation

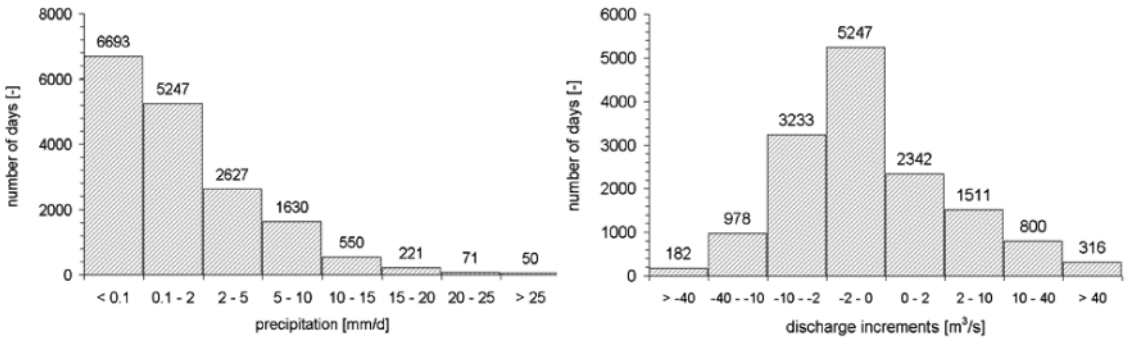



**Figure 5.1:** Histogram of observed daily areal precipitation (1951 - 2005, Upper Main) and observed daily discharge increments (1961 - 2000, gauging station Kemmern, Upper Main).

indices are determined for each season. Finally, the objective function $obj_1$ is a weighted mixture of both precipitation indices:

$$obj_1 = w_1 \sum_{i=1}^{n_{cp}} \sum_{j=1}^{4} PI_1(i,j) + w_2 \sum_{i=1}^{n_{cp}} \sum_{j=1}^{4} PI_2(i,j) \tag{5.3}$$

where $j$ is the index of the season and $n_{cp}$ is the number of CPs defined for the classification. The parameters $w_1$ and $w_2$ specify the weight for a precipitation index. In this investigation the same weight is chosen for both indices.

Unfortunately, daily precipitation is characterized by a highly skewed distribution function due to the discrete-continuous character of this variable (see left graph of Figure 5.1). It is very likely that the assumptions needed for robust parameter estimations are violated if a time series of daily precipitation is chosen to formulate an objective function. To reduce the highly skewed character of a variable, a suitable transformation function must be selected which transforms the skewed distribution function to more or less a Gaussian distribution. However, the transformation of daily precipitation is more difficult compared to other variables due to the number of dry days (zeros). An alternative way to represent the objective function is proposed in Equation 5.3 which is based on the mean seasonal properties for each CP and instead of on a daily time series.

### Optimization criterion based on daily discharge

There are several reasons why an establishment of a relationship between the large-scale atmospheric flow and the daily discharge is not straightforward (Bárdossy and Filiz, 2005): (i) Usually, a time lag has to be introduced due to the time-delayed reaction of a river basin, in particular if a gauging station is selected which is located at the catchment outlet. (ii) The catchment state which is predominantly influenced by the weather state of the previous days controls the time-delayed reaction of a river basin. (iii) A further problem is the high discharge which occurs in the raising and the falling limb of the curve during a flood event.

To overcome the latter problem, a subsample of independent discharge realizations can be selected. But this kind of technique usually leads to information loss which is also not desirable. An alternative way was proposed by Bárdossy and Filiz (2005) which is based on daily discharge differences $\Delta Q(t)$:

$$\Delta Q(t) = Q(t) - Q(t-1) \tag{5.4}$$

where $Q(t)$ is the observed discharge on day $t$ and $Q(t-1)$ is the observed discharge of the previous day. In Figure 5.2 a time series of daily discharge differences is shown for the gauging station Kemmern situated at the catchment outlet of the Upper Main. The positive discharge differences are mainly caused by moderate and intensive precipitation. The negative discharge differences are the response of the catchment due to the water excess after a precipitation event (Bárdossy and Filiz, 2005). Days with intensive precipitation are indicated by large positive discharge differences. They are usually followed by large negative responses when the discharge rapidly decreases after a flood event. The period between two floods is marked by small variations of discharge differences. The right graph of Figure 5.1 shows the histogram of the daily discharge differences. It indicates that the distribution func-

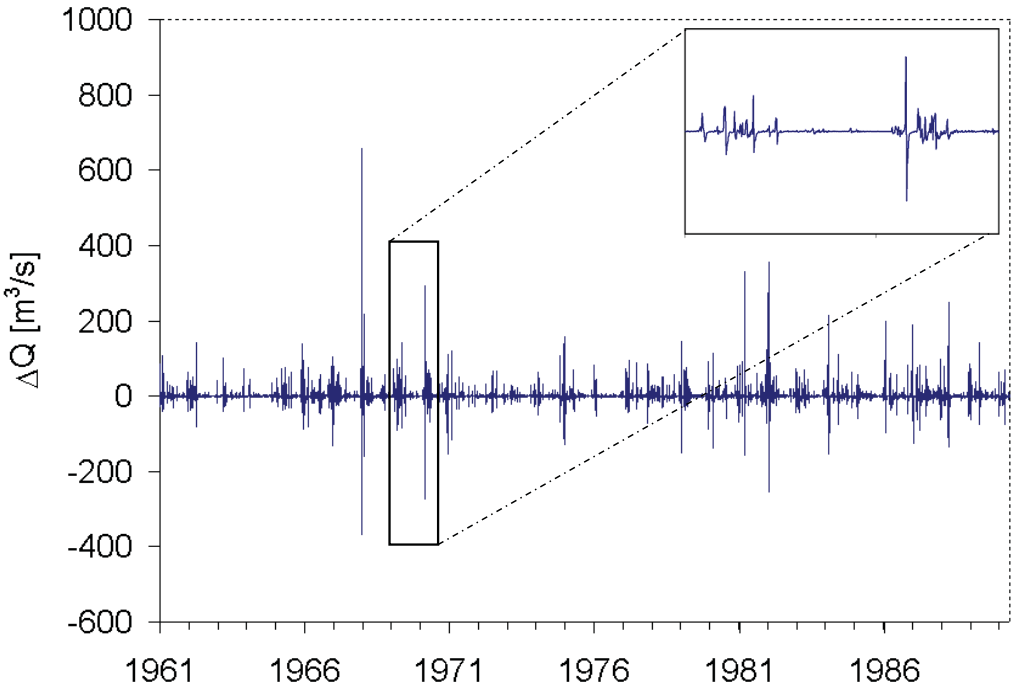



**Figure 5.2:** Time series of observed daily discharge differences measured at the gauging station Kemmern of the Upper Main.

tion is nearly symmetric. In this case, it is more likely that the assumptions needed for a robust estimation are fulfilled. Please note that the selection of daily discharge difference for the optimization has also disadvantages and some important information could be lost. For instance, positive discharge differences can be enhanced by snow melting or they can be the result it due to a sharp temperature increase.

Bárdossy and Filiz (2005) also demonstrated that an objective function based on positive discharge differences can be used for the optimization of the fuzzy rules. However, the positive discharge differences (as well as the negative discharge differences) of gauging stations which are located at the catchment outlet are often regulated due to reservoir operation, in particular during floods. To overcome this problem, Bárdossy and Filiz (2005) selected several discharge series from gauging stations which are situated in the head catchments and which are often less influenced by anthropogenic factors.

In this investigation a similar objective function is proposed which is based on small daily discharge differences. Those discharge differences mark periods between two flood events that is mainly dry days and days with light precipitation which produce either no or little runoff. If the classification can suitably predict days with small discharge differences, the classification can well distinguish those days from days with high discharge differences. Thus, for the inverse case, the differentiation of days with high discharge differences from dry days with small differences could be also enhanced so that finally the detection of flood producing situations might be improved.

The indicator time series of daily discharge differences $o_b\,(t)$ is calculated as follows. The indicator value is 1 if the value of the empirical distribution function of daily discharge differences $F_n[\Delta Q(t)]$ is in the interval between $[0.3; 0.8]$:

$$o_b\,(t) = 1, \quad 0.3 < F_n[\Delta Q(t)] < 0.8 \tag{5.5}$$

Otherwise the indicator value is zero:

$$o_b\,(t) = 0 \tag{5.6}$$

The objective function $obj_2$ is the weighted mean of the differences between the conditional mean indicator value of a given CP $i$ $\bar{o}_i$ and the mean of the indicator time series $\bar{o}_b$:

$$obj_2 = \sum_{i=1}^{n_{cp}} \frac{n_i}{n} [\bar{o}_i - \bar{o}_b]^2 \tag{5.7}$$

The value of the objective function can vary between minus and positive infinite. A value equal to zero indicates a classification which is not different to the climatological average. The aim of the optimization is to maximize the differences. In the event of this, a CP-catalog is compiled to distinguish between days with small discharge differences from days with high discharge differences. The intra-annual variability of the predictand can be accounted for by calculating the index for each season.

### 5.3.3 Validation Procedure

For the classification with daily precipitation, the predictor are the daily anomalies of the mean sea level pressure which has been derived from the Research Data Archive (RDA) of NCAR (see Section 3.2). The domain of the predictor field covers an area ranging from 30° N to 70° N and from 40° W to 40° E. The classification is performed for the study region WG by selecting the daily areal precipitation of the Nahe catchment. The validation approach is the split-sampling technique. The calibration period ranges from 1980 to 1999. The number of CPs is 13.

The classification with daily discharge differences has been performed by András Bárdossy who provided a catalog of daily CPs from 1900 to 2007 with 12 different CPs. The same predictor has been selected for this classification but the predictor domain (25° N to 75° N and from 40° W to 40° E) is slightly larger compared to the domain used for the other classification strategy. The calibration period ranges from 1980 to 1989. The time series of discharge differences are calculated from daily discharge series taken from six Rhine gauging stations. Note that the gauging stations are not part of the catchment used in this study. They are located around 200 kilometers south to the catchments of WG. Since the gauging stations are mainly located at the catchment outlet of the Rhine tributaries, a time lag of one day must be introduced to account for the concentration time of the tributaries. The comparison between both classifications is performed for the study region WG (see Section 3.2). The validation set used for the comparison covers a period from 1958 to 1979 and 2000 to 2001.

## 5.4 Wetness Index and Anomaly Maps

#### Classification with precipitation

The CP occurrence and three properties describing the precipitation characteristics of each CP are listed in Table 5.1 for the catchments of the study region WG. The mean wet day amount describes the average precipitation amount on a wet day. A wet day is defined in this study as a day with an areal precipitation larger than 2 mm/d. The precipitation index $W$ describes the wetness of a CP. It is the ratio between the conditional mean precipitation of a CP and the climatological mean precipitation. The wetness index ranges between zero and infinite. The larger the wetness index, the wetter the circulation pattern is. A wetness index tending to zero indicates a dry pattern, obviously.

The mean wet-day amount and the wetness index illustrates that the fuzzy rule-based classification can appropriately differentiate between dry and wet CPs. Three CPs (CP 2, CP 3 and CP 10) are characterized by a high wetness index and a high number of extremes compared to the other CPs. They occurred on 30 % of the days and were responsible for around 90 % of the 80 largest precipitation events. The seasonal properties indicate that the classification is slight weaker in summer. Some CPs with a low wetness index are also characterized by intensive precipitation during this season.

The investigation of the classification has been also performed for the study regions located in the Elbe and the Danube basin (EG and SG). To specify a CP-catalog for

**Table 5.1:** Statistical properties of the CPs classified for the Rhine basin with daily areal precipitation. The average properties are shown for three catchments (Nahe, Sieg and Prüm). Bold = Wet CPs with a higher risk of intensive precipitation, wi = winter half year (October to March), su = summer half year (April to September). Investigation period from 1958 to 2001.

| | CP occurence [%] | | | wet-day amount [mm/d] | | | wetness index [-] | | | number of extremes [-] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CP | wi | su | year | wi | su | year | wi | su | year | wi | su | year |
| 1 | 5.2 | 4.5 | 9.7 | 2.2 | 3.4 | 2.7 | 0.3 | 0.5 | 0.4 | 0 | 0.3 | 0.7 |
| **2** | **6.4** | **7.0** | **13.5** | **4.6** | **6.6** | **5.6** | **1.2** | **2.1** | **1.6** | **5.0** | **12.7** | **18.0** |
| **3** | **3.5** | **3.7** | **7.3** | **8.8** | **6.3** | **7.7** | **3.0** | **2.4** | **2.7** | **10.3** | **9.3** | **22.3** |
| 4 | 2.7 | 3.3 | 6.0 | 3.9 | 2.6 | 3.3 | 0.8 | 0.5 | 0.7 | 0 | 0 | 0 |
| 5 | 4.0 | 4.4 | 8.4 | 1.7 | 3.9 | 3.0 | 0.2 | 0.4 | 0.3 | 0 | 1.3 | 1.3 |
| 6 | 2.9 | 3.8 | 6.7 | 3.1 | 2.8 | 3.0 | 0.6 | 0.5 | 0.6 | 0.3 | 0.3 | 0.7 |
| 7 | 5.6 | 4.6 | 10.2 | 2.2 | 4.6 | 3.5 | 0.3 | 0.7 | 0.5 | 0 | 2.7 | 3.3 |
| 8 | 6.2 | 6.0 | 12.2 | 3.0 | 3.7 | 3.3 | 0.5 | 0.5 | 0.5 | 0.3 | 1.3 | 1.7 |
| 9 | 1.7 | 2.0 | 3.7 | 3.7 | 3.4 | 3.5 | 0.5 | 0.7 | 0.6 | 0 | 0.3 | 0.7 |
| **10** | **4.4** | **3.4** | **7.8** | **12.0** | **6.3** | **9.8** | **3.7** | **2.2** | **3.2** | **21.7** | **7.0** | **30.3** |
| 11 | 5.1 | 4.4 | 9.5 | 2.2 | 2.1 | 2.2 | 0.3 | 0.3 | 0.3 | 0.7 | 0 | 0.7 |
| 12 | 1.8 | 2.6 | 4.4 | 4.5 | 3.9 | 4.2 | 0.8 | 0.7 | 0.7 | 0.3 | 0.7 | 1.3 |
| 13 | 0.5 | 0.3 | 0.7 | 3.7 | 1.8 | 3.3 | 0.8 | 0.2 | 0.6 | 0 | 0 | 0 |

these basins, the optimized rules determined for the Rhine classification have been taken as initial rules for the new optimization. The anomaly maps of the daily mean sea level pressure of the Rhine and the Elbe classification are given in Figure 5.3 for the three wettest and the driest CP. The maps illustrate that physically reasonable CPs can be compiled by the fuzzy-rule based classification. For instance, CP 2 and CP 3 of the Rhine classification are characterized by low pressure over Central Europe leading to intensive precipitation over the study region. On the other hand, the driest CP (CP 5) of the Rhine classification is marked by high pressure north of the British Isles which reduces the influence of the Prevailing Westerlies. Thus, dry and cold air masses are transported from the north-east to the study region. On the other hand, CP 10 is another typical constellation for intensive precipitation in Western Germany which is extremely wet in winter. This CP is characterized by low pressure over the North Sea and high pressure over North Africa which causes a strong pressure gradient over West Europe so that wet air masses are transported from the Northern Atlantic to Central Europe. This large-scale circulation pattern usually dominates at the end of autumn and at the beginning of spring when the influences of the Westerlies is strongest for Central Europe.

The anomaly maps for the Elbe basin illustrate that reasonable CPs can be also determined for other basins in Germany as well. However, the maps of the wet CPs (CP 5, CP 6 and CP 12) are quite similar to the wet CPs of the Rhine classification.

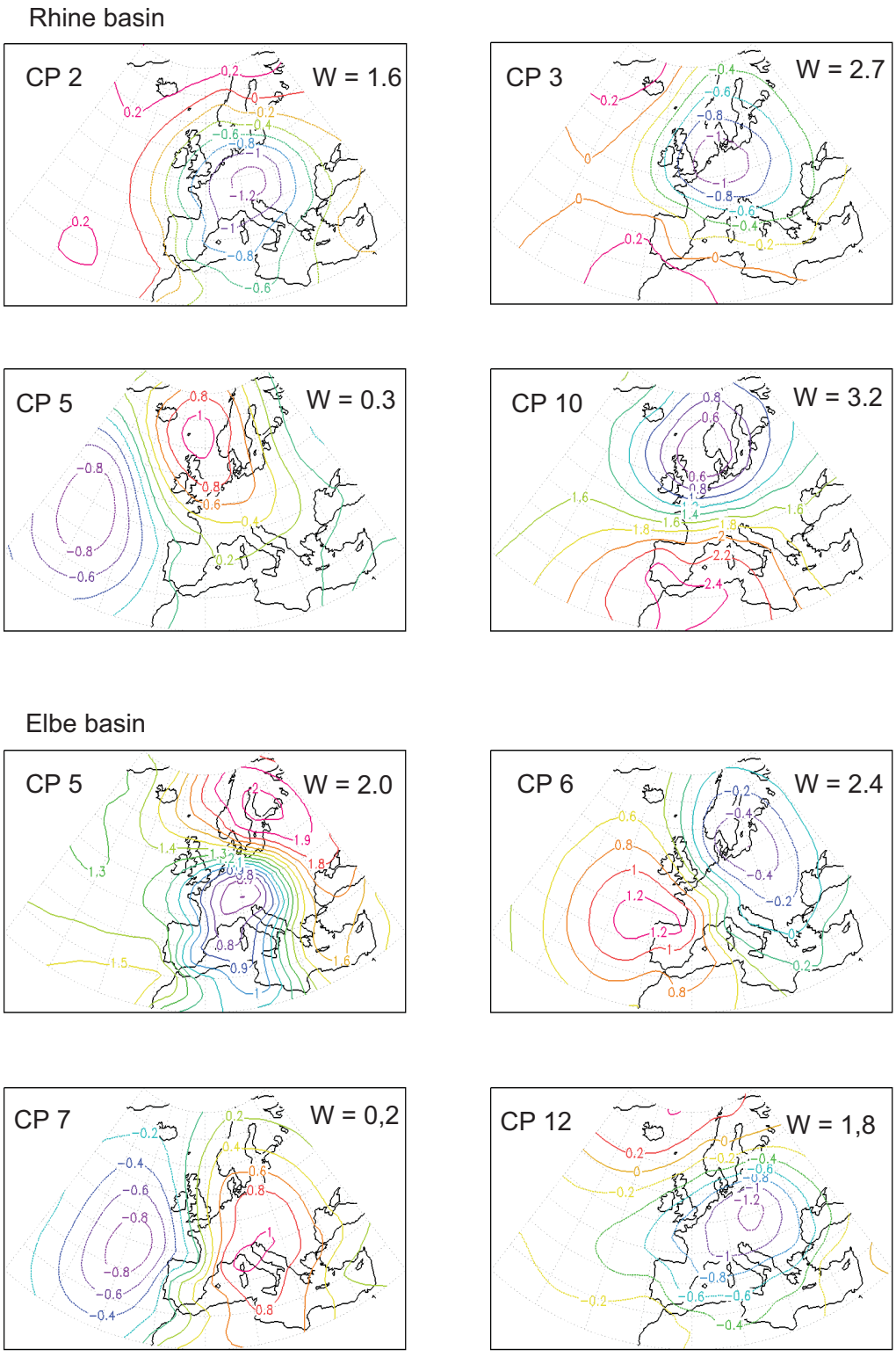



**Figure 5.3:** Anomaly maps of the daily mean sea level pressure for the three wettest and the driest CP determined for Rhine and Elbe basin. W = Wetness index. Classification with precipitation.

The classification for the Rhine catchments has been also tested for the Elbe and Danube catchment (study region EG and SG) but the precipitation properties of the CPs turn out to be not so appropriate for these catchments. The main reasons for the moderate transferability of the Rhine classification are the climatological differences between the study regions (see Section 3.1). Thus, the compilation of further catalogs for the other two basins seems to be a reasonable exercise which should also improve the detection of intensive precipitation.

### Classification with discharge differences

The classification with discharge differences can also appropriately distinguish between dry and wet CPs (see Table 5.2) although the objective of the optimization is the prediction of small discharge differences for gauging stations which are not situated in the study region. The comparison of the statistical properties with the previous CP-classification illustrates that the new classification can probably even slightly better separate between dry and wet CPs. This classification has the driest (CP 2) and the wettest circulation pattern (CP 5) in comparison to the other classification. The investigations also illustrate that the detection of intensive precipitation is also possible. Around 90% of the extremes can be explained if all wet CPs (CP 4, CP 5, CP 6, CP 7, CP 11 and CP 12) are used to trigger an alarm. However, the occurrence frequency of the wet CPs is very high (around 40%) so that

**Table 5.2:** Statistical properties of the CPs classified for the Rhine basin and with daily discharge differences. The average properties are shown for three catchments (Nahe, Sieg and Prüm). Bold = Wet CPs with a high risk of intensive precipitation, wi = winter half year (October to March), su = summer half year (April to September). Investigation period from 1958 to 2001.

| | CP occurrence [%] | | | wet-day amount [mm/d] | | | wetness index [-] | | | average number of extremes [-] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CP | wi | su | year | wi | su | year | wi | su | year | wi | su | year |
| 1 | 5.4 | 3.8 | 9.2 | 4.7 | 4.5 | 4.5 | 0.6 | 0.7 | 0.6 | 0.7 | 0.3 | 1.0 |
| 2 | 5.3 | 5.6 | 10.9 | 2.2 | 2.9 | 2.6 | 0.1 | 0.2 | 0.1 | 0 | 0.3 | 0.3 |
| 3 | 3.4 | 2.8 | 6.2 | 2.2 | 2.7 | 2.5 | 0.2 | 0.2 | 0.2 | 0 | 0 | 0 |
| **4** | **2.6** | **2.5** | **5.1** | **7.8** | **8.2** | **8.0** | **1.3** | **1.9** | **1.6** | **3.0** | **2.7** | **5.7** |
| **5** | **5.4** | **3.3** | **8.7** | **16.8** | **8.3** | **13.9** | **3.8** | **2.2** | **3.4** | **28.7** | **3.0** | **31.7** |
| **6** | **3** | **4.9** | **8** | **6.8** | **9.8** | **8.5** | **1.3** | **2.1** | **1.7** | **3.3** | **10.0** | **13.3** |
| **7** | **3.3** | **3.6** | **6.9** | **9.2** | **7.2** | **8.4** | **1.7** | **1.4** | **1.6** | **6.0** | **4.0** | **10.0** |
| 8 | 3.9 | 4.2 | 8 | 4.0 | 5.3 | 4.7 | 0.3 | 0.5 | 0.4 | 0.3 | 1.0 | 1.3 |
| 9 | 7.9 | 7.6 | 15.4 | 2.4 | 5.3 | 3.9 | 0.2 | 0.5 | 0.3 | 0.3 | 2.3 | 2.7 |
| 10 | 4.3 | 5.5 | 9.8 | 2.7 | 5.0 | 4.1 | 0.1 | 0.4 | 0.2 | 0 | 1.3 | 1.3 |
| **11** | **3.1** | **4.4** | **7.5** | **8.8** | **7.5** | **8.1** | **1.6** | **1.7** | **1.6** | **3.7** | **3.3** | **7.0** |
| **12** | **2.3** | **1.8** | **4.2** | **10.9** | **8.9** | **10.1** | **1.8** | **1.8** | **1.8** | **3.3** | **3.3** | **6.7** |

many alarms are needed to hit a number of extremes. It is obvious that this kind of warning strategy is not useful for users with high costs and low losses.

In the following section the relative value for both classifications is calculated to investigate how skillful the detection of intensive precipitation is by the fuzzy-rule based classification.

## 5.5 Forecast Skill and Value

### 5.5.1 Pure Classification

To perform a probability forecast of daily areal precipitation by a catalog of daily CPs, the conditional empirical cumulative distribution is specified for a given CP. Then, this function is used to provide a probability forecast of daily areal precipitation for each day of the investigation period. Finally, the forecast performance of the classification is evaluated by calculating the relative value proposed in Equation 4.18. The decision thresholds needed for the calculation of the relative value are the quantiles of the conditional empirical distribution function $q_t = [0.01, 0.02, ..., 0.99]$. In this section the score of the relative value is determined for both CP catalogs which have been compiled for the Rhine basin.

The forecast value for both classifications is given in Figure 5.4 for two performance measures, the maximum value and the user interval, in relation to the return frequency of an event. A detailed explanation of the performance measures is given in Section 4.7. The performance measures illustrate that the prediction of daily areal precipitation is skillful over the entire range of the predictand for both classifications. However, the maximum forecast value is only slightly higher compared to the low skill reference forecast for light precipitation (see Figure 5.4a and 5.4c). The figures also show that the maximum forecast value rises with increasing precipitation amount. Thus, the classification has a higher forecast performance for the detection of rare events than for small events which is quite encouraging. However, the user interval indicates that only users with low cost-loss ratios can profit from the model predictions (see Figure 5.4b and 5.4d). These specific users can compensate for the high number of false alarms due to their lower costs for protection.

The curve also illustrates that the model performance of both classifications only differs marginally. For the investigation period, the classification with daily precipitation slightly outperforms the other optimization strategy for moderate and for intensive precipitation (see Figures 5.4a and 5.4b). However, the performance of the classification with daily precipitation decreases if the model is transferred to an independent data set (see Figures 5.4c and 5.4d). On the other hand, the performance of the classification with discharge difference is more or less stable.

The duration of the calibration period might be a further factor which confirms the higher model transferability if discharge differences are used for optimization of the fuzzy rules. The calibration period for the classification with precipitation is 20 years while for the second classification only ten years has been selected for the model calibration. The design of downscaling approaches which need only a small portion of the entire data set for a robust calibration is highly desirable and might be a further argument for the formulation of an objective function based on

discharge differences.

The results also illustrate that the CP-catalog compiled with discharge differences slightly outperforms the other catalog for the validation period. Thus, the classification of discharge differences has a higher performance for the detection of intensive precipitation. Please note that the second catalog is not only suitable for the detection of intensive precipitation. This classification is also probably more valuable for the detection of floods since it has been primarily designed for the prediction of daily discharge differences.

Finally, we can summarize that there are several indications that the use of small discharge differences for the optimization of the fuzzy rules enhances the transferability of the fuzzy rule-based classification and its performance for the detection of intensive precipitation.

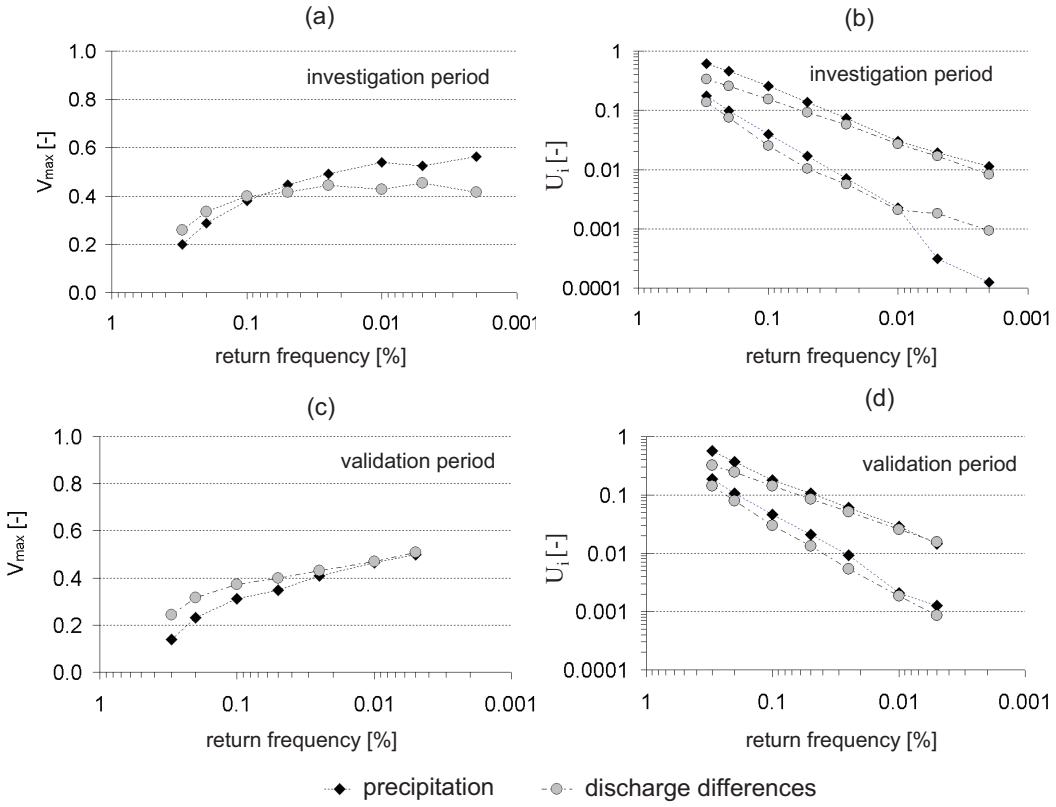



**Figure 5.4:** Forecast value of the classifications based on daily precipitation in comparison to the classification with daily discharge differences for the investigation period (1958 - 2001) and the validation period (1958 - 1979 and 2000 - 2001). Maximum forecast value $V_{max}$ (left) and user interval $U_i = [\alpha_{min}; \alpha_{max}]$ (right), study region WG, Rhine basin (Ruhr, Sieg, Prüm and Nahe).

### 5.5.2 Classification vs. Analog Forecasting

The precipitation forecast of the classification with daily discharge differences are compared to the predictions of the analog algorithm based on the predictor set A (GPH, UFLX and RHUM) and the model settings which are listed in the tables of the previous chapter. The investigation period ranges from 1958 to 2001. The comparison is performed for the catchments located in WG. Again, the relative value is calculated for both approaches based on the decision thresholds of the previous section.

The model performance of the classification in comparison to the analog method is illustrated on the Figures 5.5a and 5.5b. Both performance measures, the maximum value and the user interval, are clearly larger for the analog method than for the classification over the entire range of the predictand. Around 90% of the extremes can be detected with both downscaling techniques if an appropriate decision threshold is selected to trigger a warning (see Table 5.3). However, the false alarm rate of the analog method is clearly lower compared to the classification method which strongly

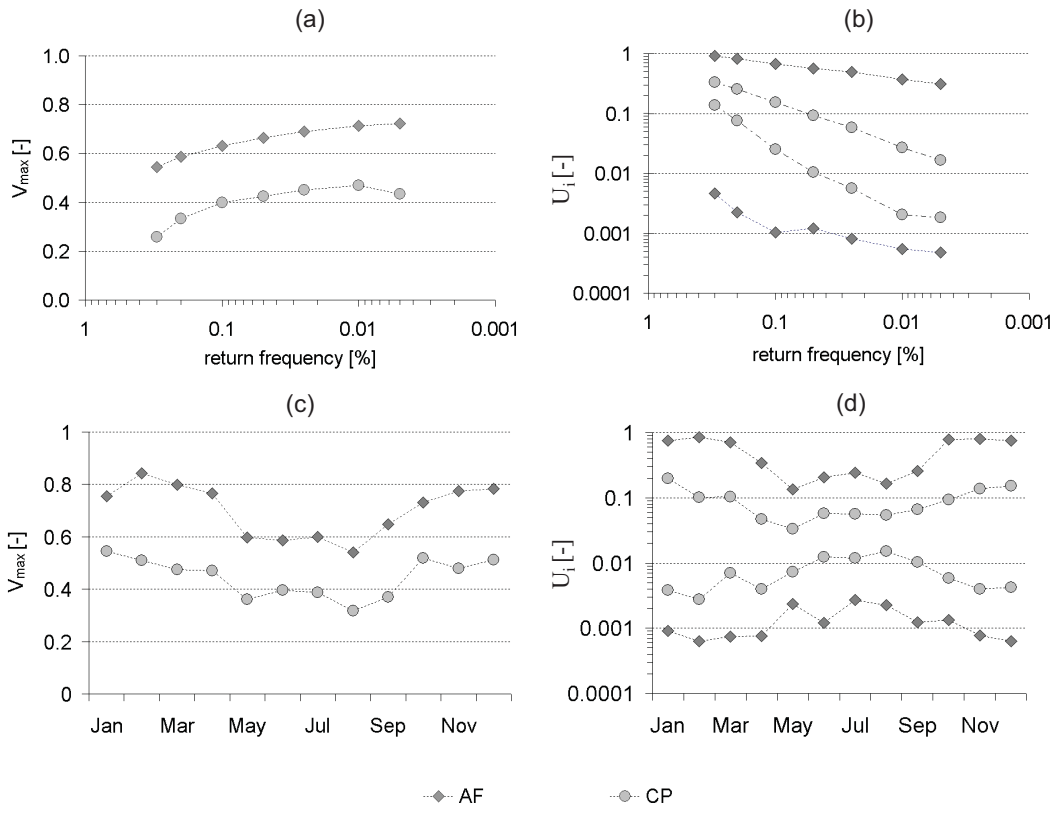



**Figure 5.5:** Forecast value of the analog method (AF) versus classification (CP). Maximum forecast value $V_{max}$ (left) and user interval $U_i = [\alpha_{min}; \alpha_{max}]$ (right). Figure a and b: Forecast performance over the range of the predictand. Figure c and d: Intra-annual variability of the forecast performance. Rhine basin (Ruhr, Sieg, Prüm and Nahe). 1958 to 2001.

**Table 5.3:** Binary outcomes of analog forecasting (AF) in comparison to the classification (CP). $a$ = hit, $b$ = false alarm, $c$ = miss, $d$ = correct rejection, $HR$ = hit rate, $FA$ = false alarm rate and $p_t$ = optimal decision threshold. 1958 to 2001, Rhine basin (Ruhr, Sieg, Prüm and Nahe). The average statistics are given for the three catchments.

| catchment | season | $a$ | $b$ | $c$ | $d$ | $HR$ | $FA$ | $p_t$ |
|---|---|---|---|---|---|---|---|---|
| AF | year | 72.0 | 2622.3 | 9.5 | 13367.2 | 0.88 | 0.16 | 0.02 |
| | wi | 39.0 | 577.0 | 2.0 | 7400.0 | 0.95 | 0.07 | 0.04 |
| | su | 33.0 | 2045.3 | 7.5 | 5967.2 | 0.81 | 0.26 | 0.01 |
| CP | year | 74.5 | 6432.5 | 7.0 | 9557 | 0.91 | 0.40 | 0.01 |
| | wi | 38.5 | 3125.3 | 2.0 | 4852.2 | 0.95 | 0.39 | 0.01 |
| | su | 36.0 | 3307.2 | 5.0 | 4704.8 | 0.88 | 0.41 | 0.01 |

increases the goodness of the model prediction of the analog algorithm. Furthermore, Figures 5.5c and 5.5d show that the analog method is also more suitable for the detection of intensive precipitation during the entire year.

The investigation in this section illustrates that the model predictions of the analog method are clearly superior compared to predictions of the fuzzy-rule based classification. However, this result can be expected since the analog algorithm uses more information for the downscaling. Beside a pressure related variable also two additional humidity related variables are selected for the detection of intensive precipitation. The analog method takes also a smaller predictor domain to focus more on the atmospheric state over the study region. This kind of technique should also improve the prediction of precipitation, since grid points located next to the study region usually contain more information compared to grid points which are situated far away from the region of interest.

Due to the lower model performance of the classification, the objective classification seems to be not needed for the detection of intensive precipitation. However, this statement is probably not true since there are still many reasons for the use of objective classification in weather forecasting (see e.g. the arguments presented in the beginning of this chapter).

A further reason for the use of the classification may be justified by considering the following problem. The model performance of both approaches presented in the chapter has been performed with reanalysis data. Thus, the investigation was performed under the assumption that the large-scale information of the GCMs are perfectly predicted. But it is obvious that this assumption doesn't hold for operational application when forecast data of a global NWP model are used for the downscaling. In this case, the forecast performance of the global NWP model usually decreases with increasing lead time and especially for the prediction of humidity related variables. Thus, for longer lead times only pressure related variables can be used for the downscaling. Furthermore, it is also likely that the predictor domain specified for the analog method in the reanalysis framework must be increased since the atmospheric circulation of a small predictor domain is probably less reliably predicted by a GCMs compared to the circulation of a larger predictor domain. Thus,

we can expect for longer lead times that the performance differences of both techniques may decrease so that the fuzzy-rule based classification performs similar to analog forecasting. To clarify this issue, the model performance of both downscaling techniques must be determined with forecast data of a global NWP and for different lead times in future investigations.

### 5.5.3 Conditional Forecast Value

We know from the previous section that the model performance strongly varies throughout the season since the prediction of summer precipitation is poorer due to high stochastic character of those events. However, the forecast performance of a downscaling approach can also vary for different circulation patterns which might effect the decision making process. To investigate this issue, the conditional forecast value of the analog method is calculated for a given CP. To perform this investigation, the Rhine classification with daily precipitation is grouped into four categories: the three wettest CPs (CP 2, CP 3 and CP 10) and all corresponding dry CPs which are merged into one group (CP dry). Then, the relative value of the analog algorithm is calculated for the four groups based on the same settings mentioned in the previous section.

The comparison of the maximum forecast value illustrates that the model performance of the analog method strongly varies for the four groups (see Figure 5.6a). The model predictions are the weakest for CP 2 and are the strongest for the dry CPs for light and for intensive precipitation. This result seems to be less promising since the model predictions of the analog method have a higher forecast value for dry CPs than for wet CPs. Usually, a warning system should trigger valuable warnings for wet patterns since those CPs have a high likelihood for intensive precipitation. Even so, it is also crucial for the development of a warning system that the detection of intensive precipitation is skillful for dry weather patterns.

We know from Section 4.7 that the predictions of the analog algorithm are valuable for users with cost-loss ratios similar to the return frequency of an event. In the event of this, the low-skill reference forecast ("*no warning*" and "*warning*") are the weakest and the model predictions of the analog method can be the strongest. However, if the user interval is estimated for a given CP, the model predictions of analog forecasting are the strongest for users with cost-loss ratios similar to the conditional return frequency of a precipitation event (see Figure 5.6a) . If a dry CP occurs, only those users can profit from the model predictions which are characterized by low cost-loss ratios. In this case, the reference forecast "*no warning*" represents for many users with high cost-loss ratios a suitable warning strategy due to the low return frequency of extremes during dry CPs. As expected for a wet CP, the user interval is shifted to higher cost-loss ratio due to the higher likelihood of intensive precipitation. In the event of this, even users with moderate cost-loss ratios can benefit from the model predictions. The investigation in this section demonstrates that the model performance of analog forecasting depends on the circulation pattern which also influences the decision making process. It means that the optimal use of a probability forecast is dependent on the various criteria determined in Section 4.7 as well as on the large-scale weather pattern. Since an optimal decision making is not straightforward due to a number of factors, future investigations should point

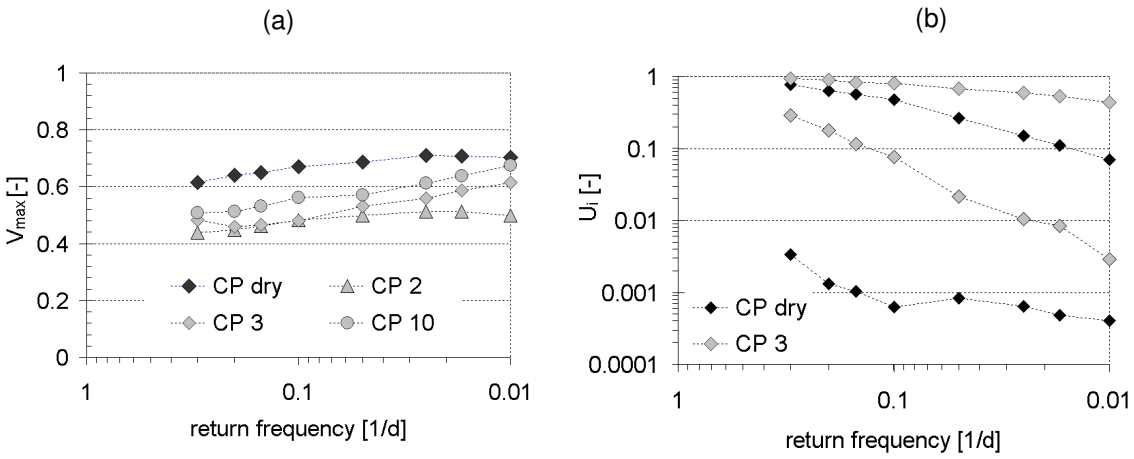



**Figure 5.6:** Forecast value of analog forecasting for wet and dry CPs over the entire range of the predictand. Maximum forecast value $V_{max}$ (left) and user interval $U_i = [\alpha_{min}; \alpha_{max}]$. 1958 to 2001, Rhine basin (Ruhr, Sieg, Prüm and Nahe). The average statistics are given for these catchments.

out the key factors to optimize the decision making process for the individual users.

The results presented in this section also illustrate a strategy of how to design a warning system which provides valuable information for the general public. As mentioned in Section 2.2.5, the members of the general public are usually characterized by a high cost-loss ratio since their willingness for protection is significantly reduced if too many false alarms are triggered. To overcome this problem, risk categories such as wet CPs must be specified which are marked by a high likelihood of intensive precipitation. In this case, there might be a chance to provide valuable warning information to users with high cost-loss ratios. Furthermore, the individual users must be also divided into different groups with similar cost-loss ratios. Then, the warning system should trigger only an alarm to these groups if there is a chance to provide a valuable warning. A "chance" means in the framework of the simple cost-loss approach that the mean cost-loss ratio of a certain user group should be in the range of the conditional return frequency of the extreme event of a given risk category.

The results in this section also illustrated that analog forecasting has the lowest skill for CP 2. At the moment, the reasons for the low model performance have been not investigated in detail. It is possible that the number of suitable analogs is less compared to the other CPs. However, this pattern occurs more frequently compared to other wet patterns so that this guess might be not true. The main reason for the low model performance could be that many precipitation events occurred in summer and that the forecasts for summer precipitation is usually poorer compared to winter precipitation. To improve the model prediction for CP 2, the identification of more suitable model settings of the analog approach might be worthwhile, since the model performance is clearly lower for this CP compared to the other CPs.

## 5.6 Summary and Conclusions

An objective classification method based on the concept of fuzzy rules was presented. This technique was used to compile a catalog of daily CPs for the study region WG located in the Rhine basin. It was demonstrated that physically reasonable CPs with a high likelihood of intensive precipitation can be specified by the classification. We also illustrated that reasonable CPs can be compiled for other basins in Germany.

Since the model transferability of the approach depends on the choice of the objective function, two different optimization strategies were compared for the prediction of intensive precipitation. The first strategy takes an objective function which focuses on the prediction of daily areal precipitation. This strategy should represent the intuitive choice of an investigator if the purpose of the classification is to link daily precipitation to the large-scale atmospheric flow processes. For the second optimization strategy an objective function was formulated for the prediction of daily discharge differences. In comparison to the first strategy, this way has the advantage that the assumptions for robust parameter estimation are probably less violated, since the distribution function of daily discharge differences is less skewed compared to daily precipitation. Both optimization strategies were used to compile a classification for the Rhine basin.

The results indicated that the selection of daily discharge differences can enhance the model transferability of the classification. We could demonstrate that the selections of daily discharge differences can even slightly improve the detection of precipitation over the entire range predictand. However, there are still further options for the formulation of an objective function based on daily areal precipitation which might produce a more suitable classification. In the following chapter, some optimization strategies for the downscaling of daily precipitation are proposed. It is analyzed how the choice of the objective function affects the model transferability for analog forecasting.

The evaluation of the forecast skill demonstrated that the predictions of the classification are valuable for the prediction of intensive areal precipitation. However, a warning system based on the classification triggers many alarms so that the classification is less valuable in comparison to the analog method. Nevertheless, there are still many reasons to apply a classification in operational weather forecasting. For example, it was highlighted that the forecast performance of the analog method varies for different large-scale circulation pattern and that the use of the circulation pattern for the evaluation of the forecast performance can improve the decision making process.

We can summarize that both, fuzzy-rule based classification and analog forecasting, are suitable approaches for the detection of intensive precipitation in small rivers basin in Germany. Finally, it seems to be promising to spend more time with both techniques. For example, future studies can increase the number of circulation patterns which might enhance the resolution of the precipitation predictions and the detection of extremes. In the following chapter two ways are presented to improve the analog method: (i) An optimization algorithm is presented to define a suitable distance function for the selection of analogs; (ii) It is highlighted how large-scale daily circulation patterns can be incorporated in the search algorithm to improve the selection process of analogs.

# 6 Metric Optimization

The performance of analog forecasting depends on several criteria which must be selected by the investigator before starting the search algorithm (see Chapter 4). Each criterion is related to a parameter set and the task of the investigator is to define suitable values for the parameters, usually with the objective to minimize the differences between model predictions and observations. Since a guideline for supporting a decision is not available, the selection of suitable model parameters is costly and a lot of experience is needed. The major problem for the definition of ap-

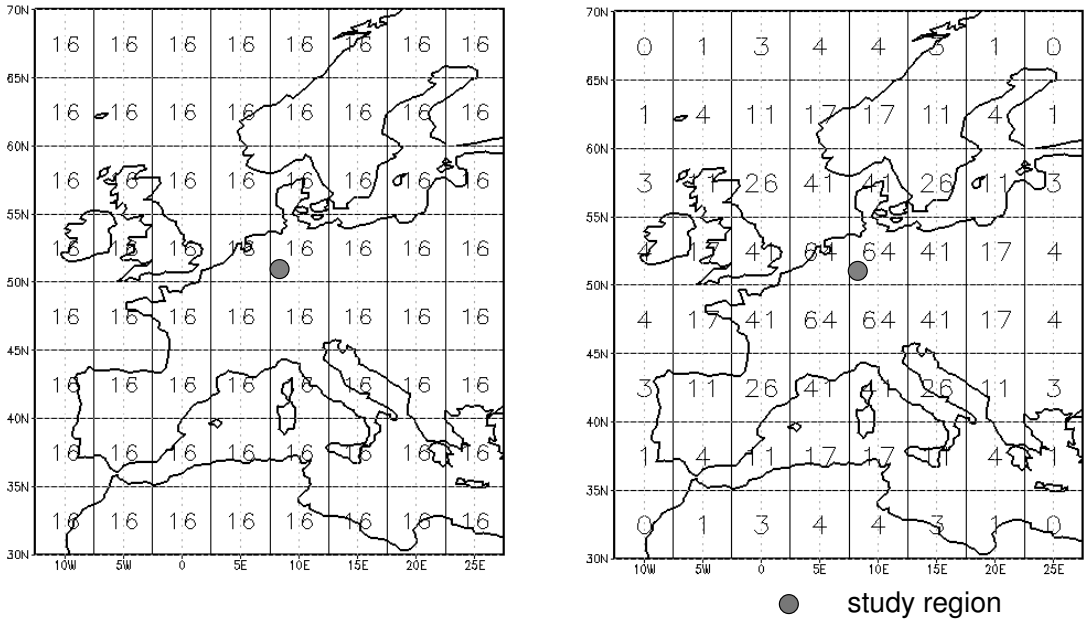



**Figure 6.1:** Metric weights at each grid point of a predictor domain. Left: Uniform selection of the weights representing the traditional way of analog forecasting. Right: Non-uniform selection of the weights with a symmetric distribution.

propriate parameters is their non-linear interaction which influences the parameter values of the previous choice if a stepwise calibration is performed. An investigator with much knowledge may point out the main sensitive parameters. However, due to the number of possible parameter combinations, a suitable definition of the parameter values is not straightforward.

Probably, one of the most time-consuming tasks in analog forecasting is the selection of an appropriate distance function. Usually, the investigator selects a distance function where all grid points of a predictor domain have the same weight for the comparison (see left graph of Figure 6.1). However, this guess is actually a poor

choice. The investigator knows that grid points located next to the study region should get a higher weight compared to grid points which are situated far away because the information content of the predictor values is higher at those points. A non-uniform selection of the weights is illustrated, for instance, in the right graph of Figure 6.1. The closer the grid point is to the location of the study area, the larger is the value of the weight. However, the best distribution of the weights is usually not known so that a specification is not intuitive. To tackle this problem, we need an optimization algorithm which calculates the values of the weights at each grid point. This technique has the advantage that two criteria, (i) the distance function and (ii) the predictor domain, can be specified without much effort and knowledge if the investigator selects a large predictor domain and the algorithm identifies suitable weights for the grid points.

There are only a few works which addressed the problem of determining an appropriate distance for the selection of analogs in the subfields of atmospheric sciences. Fraedrich and Rückert (1998) proposed such a methodology by selecting the weighted Euclidean distance for the optimization. This methodology was applied by Sievers et al. (2000) and Fraedrich et al. (2003) to forecast tracks of tropical cyclones. The selection of an appropriate metric function was also addressed in detail by Bárdossy et al. (2005). They applied the methodology for the estimation of annual mean discharge from catchment characteristics and for daily flood forecasting.

## 6.1 Methodology

For the definition of the metric weights, a suitable distance function, an optimization algorithm and an objective function is needed. In the following sections a brief description of these components is given.

### 6.1.1 Weighted Euclidean Distance

The weighted Euclidean distance is an appropriate distance function for the optimization problem because this measure consists of a coefficient which can be used to specify the weights. For the comparison of two atmospheric patterns $\boldsymbol{x}$ at time step $t_1$ and $t_2$, this measure is defined as follows:

$$d[\boldsymbol{x}(t_1), \boldsymbol{x}(t_2)] = \left( \sum_{k=1}^{K} g_k \left[ x_k(t_1) - x_k(t_2) \right]^2 \right)^{\frac{1}{2}} \tag{6.1}$$

where $g_k$ is the metric coefficient of the k-th grid point $P(i, j)$. $K$ is the number of grid points. The metric coefficient can be any positive value. If all coefficients are identical with $g_1 = g_2 = ... = g_K$, the weighted Euclidean distance is equal to the Euclidean distance (see Equation 4.1).

The function used for the calculation of the metric coefficient at a grid point $P(i, j)$ describes a symmetrical distribution of the weights similar to a Gaussian distribution. It is defined in the following way:

$$g_k = e^{-\frac{(i-\mu_i)^2+\left(j-\mu_j\right)^2}{2\sigma^2}} \tag{6.2}$$

The coordinates $i$ and $j$ of the grid point P are relative values to express the distance from the study region. If the coordinates are zero, the location of the grid point is identical to the location of the study region. The parameter $\mu_i$ and $\mu_j$ are the coordinates of the peak value $P_{max}(\mu_i, \mu_j)$ of the distribution. If $\mu_i$ and $\mu_j$ are zero, the location of the peak value is identical to the location of the study region. The parameter $\sigma$ specifies the slope of the function which can be any positive value. The larger $\sigma$ is, the more similar are the metric coefficients. If $\sigma$ tends to infinity, all metric coefficients are uniform and the distance function is equal to the Euclidean distance. If $\sigma$ tends to zero, the coefficient at the point $P(\mu_i, \mu_j)$ is one and the coefficients at the surrounding points are zero. In this specific case, the calculation of the distance function is only based on the predictor values at the peak location.

The distribution function proposed for the metric coefficients describes a two-dimensional isotropic distribution with a single peak. However, the best distribution of the metric coefficient is usually not known. It is very likely that the distribution of the metric coefficient is anisotropic with more than one single peak. To create a more flexible function, the metric coefficient of two or more functions, given in Equation 6.2 above, must be summarized:

$$g_k^* = \sum_{L}^{l=1} g_{k,l} \tag{6.3}$$

where $g_k^*$ is the sum of the metric coefficients at grid point $k$ and $g_{k,l}$ is the metric coefficient of the $l$-th function at the same grid point. $L$ is the number of functions selected for the description of the spatial distribution equivalent to the number of peaks. The number of peaks must be carefully selected since an increase of the peaks is associated with an increase of the model parameters needed for the optimization. For example, if a function with three peaks is selected, nine parameters (six location parameters and three slope parameters) must be specified by the optimization algorithm.

### 6.1.2 Simulated Annealing

An analytic solution of the mathematical problem needed for the estimation of the metric coefficient is not possible. To find an approximate solution of this problem, an optimization algorithm is needed. A trial and error procedure is not convenient due to the large number of parameter combinations.

In this study a simulated annealing algorithm (Aarts and Korst, 1989) is selected. The idea of this methodology comes from the annealing of metal. After the heating of a material, a slow controlled cooling enables the atoms to find a low-energy state

near their ground state. The cooling of the metal is controlled by the annealing temperature. The slower is the decrease of the temperature, the higher is the chance for finding a low-energy state though the process of annealing takes longer. In the case of a mathematical optimization problem, the model parameters represent the metal atoms and the model performance the energy state of the material. In contrast to other optimization techniques, like the gradient-descent or the hill climbing algorithm, simulated annealing allows an uphill movement of a parameter to leave a local minimum. Finally, a parameter set can be identified that is near to the optimal parameter set.

To determine suitable weights by the simulated annealing algorithm, the following steps have to be performed:

1. The number of distributions functions $J$, the initial values for the parameters $\boldsymbol{\mu}_i = \{\mu_{i,1}, ..., \mu_{i,J}\}$, $\boldsymbol{\mu}_j = \{\mu_{j,1}, ..., \mu_{j,J}\}$, $\boldsymbol{\sigma} = \{\sigma_1, ..., \sigma_J\}$ and the annealing temperature $T_a$ are randomly chosen.

2. A small change is randomly introduced to one element of $\boldsymbol{\mu}_i^*$, $\boldsymbol{\mu}_j^*$ or $\boldsymbol{\sigma}^*$.

3. The weighted Euclidean distance is computed to identify the most similar analogs for all forecast days of a given training period.

4. The objective function $obj^*$ is calculated.

5. If the objective function $obj^* < obj$, the change is accepted and $\boldsymbol{\mu}_i = \boldsymbol{\mu}_i^*$, $\boldsymbol{\mu}_j = \boldsymbol{\mu}_j^*$, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^*$ and $obj = obj^*$. Otherwise the change is accepted with a probability $p = e^{-(\frac{obj - obj^*}{T_a})}$ and $\boldsymbol{\mu}_i = \boldsymbol{\mu}_i^*$, $\boldsymbol{\mu}_j = \boldsymbol{\mu}_j^*$, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^*$ and $obj = obj^*$.

6. Steps 2 to 5 are repeated $k$-times.

7. The optimization is repeated $l$-times (steps 2 to 6) by decreasing the annealing temperature with $T_a = T_a * dt$ with a constant value $dt$ ranging between $0.9 < dt < 1$.

The search space is restricted for $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ to an interval between [-10°; 10°] in both directions. The values are given in relative coordinates to express the distance to the target region. The search interval for $\boldsymbol{\sigma}$ is between [0.2; 5]. As the initial guess, a symmetrical distribution is selected where the location of the peak value is identical to the location of the study region ($\boldsymbol{\mu}_i = 0°$, $\boldsymbol{\mu}_j = 0°$ and $\boldsymbol{\sigma} = 2$).

### 6.1.3 Objective Function

It was illustrated in Section 5.3.2 that the choice of the objective function affects the goodness of the model transferability to an independent data set. To analyze this issue for the search algorithm, four optimization strategies are proposed in this investigation:

- **Intensive precipitation**: The mean Brier skill score (Equation 2.10) is calculated for daily areal precipitation events with a low return frequency: $s_1 = 0.025$, $s_2 = 0.010$, $s_3 = 0.005$.

- **Light precipitation**: The mean Brier skill score is calculated for daily areal precipitation events with a high return frequency: $s_1 = 0.35$, $s_2 = 0.30$, $s_3 = 0.20$).
- **Small discharge differences**: An indicator time series of small discharge differences is taken to determine the objective function proposed in Equation 6.4.
- **Small and high positive discharge differences**: A weighted mixture of two indicator time series based on small and large positive discharge differences is taken for the optimization (see Equation 6.5).

The optimization strategy "intensive precipitation" should represent the intuitive choice of an investigator because he would like to maximize the forecast performance of the analog method for precipitation extremes. However, it is very likely that this optimization strategy is not an appropriate choice, since the distribution function of daily areal precipitation is highly skewed which can lead to a bias in the parameter estimation (for a more detailed discussion see Section 5.3.2). To reduce this negative effect, the time series of daily precipitation is converted to a binary time series. Finally, the mean Brier skill score is calculated for events with low occurrence frequencies to focus on extremes. Even so, a symmetric distribution of the indicator time series can be only derived if the relative frequency of the zeros (dry days) and the ones (wet days) is nearly the same. The selection of a high precipitation threshold for the transformation of the absolute values leads again to an indicator time series with a skewed distribution.

The optimization strategy "light precipitation" also uses an indicator time series of daily areal precipitation. This strategy focuses on days with no precipitation and low precipitation amounts. If the model predictions are accurate for dry days, the algorithm can well separate dry from wet situations. If a suitable discrimination for those days is possible, the inverse case, e.g. the separation of wet from dry situations, might be enhanced so that finally the model predictions would be improved for intensive precipitation.

The latter two optimization strategies use an objective function based on an indicator time series derived from daily discharge differences. The selection of discharge differences for the optimization of the model parameters has several advantages which are discussed in detail in Section 5.3.2. However, the third and fourth objective functions are slightly different to the functions proposed in Section 5.3.2. The value of these objective functions are determined by calculating the differences between predictions and observation based on seasonal properties for each CP. In this investigation the differences between model predictions and observations are calculated on a day-by-day basis.

In the case of the optimization strategy "small discharge differences", the objective function is defined in the following way:

$$obj_3 = \frac{1}{n_{st}} \frac{1}{n} \sum_{l=1}^{n_{st}} \sum_{t=1}^{n} [\bar{f}_b(t) - o_b(t)]^2 \tag{6.4}$$

where $\bar{f}_b(t)$ is the mean indicator value of the most similar analogs and $o_b(t)$ is the indicator value of the observation to the same time. $n_{st}$ is the number of gauging

stations selected for the optimization. The transformation of the observed time series to the indicator time series is analog to the procedure proposed in Equation 5.5. The objective function of the fourth optimization strategy $obj_4$ is a weighted mixture of two objective functions:

$$obj_4 = w_1 \frac{1}{n_{st}} \frac{1}{n} \sum_{l=1}^{n_{st}} \sum_{t=1}^{n} [\bar{f}_b(t) - o_b(t)]^2 + w_2 \frac{1}{n_{st}} \frac{1}{n} \sum_{l=1}^{n_{st}} \sum_{t=1}^{n} [\bar{f}_b^*(t) - o_b^*(t)]^2 \qquad (6.5)$$

where $\bar{f}_b^*(t)$ and $o_b^*(t)$ have the same meaning as the variables of the aforementioned equation, except they are the corresponding values for the indicator time series of the high positive discharge differences. The weights $w_1$ and $w_2$ can be used to balance between both terms. The conversion of the daily positive discharge differences to the indicator time series is performed for the entire time series. The discharge values above the value of the 90%-quantile are transformed to 1. The rest of the discharge information is converted to zero.

## 6.2 Model Development and Validation Strategy

The study is performed for the catchments located in the study region WG (see Chapter 3). The predictor is the 1000 hPa height field at 18 UTC derived from the NCEP/NCAR-reanalysis data set. The data resolution of the predictor is reduced from 2.5 x 2.5° to 5.0 x 5.0° to save time for the optimization. Usually, a data reduction should be avoided, but the loss of information is negligible for variables with a high spatial dependence like the geopotential height (see the following section).

The initial predictor domain of the height field for the optimization with precipitation covers an area which ranges from 35.0° N to 65.0° N and from 12.5° W to 22.5° E. To calculate the Brier skill score, the daily areal precipitation of the catchments located in the study region WG are selected. For the optimization with the daily discharge differences, a slightly larger height field was selected (32.5° N to 67.5° N and 10° W to 25° E). The indicator time series of the small discharge differences is calculated from the daily discharge series of six Rhine gauging stations. They are located around 200 kilometers south of the target region. The indicator time series of positive discharge differences is taken from a single gauging station located in the Nahe catchment.

Since all gauging stations are mainly situated at the catchment outlet, the concentration time of the catchments must be considered. This means that an appropriate observation time for a predictor must be selected. In the case of the geopotential height a suitable observation time is -30 UTC. Thus, a time lag of around two days is used to take into account the concentration time of the river basins.

The model transferability of the approach is tested by dividing the investigation period into a training (1961 to 1995) and a validation set (1958 to 1960 and 1996 to 2001). For the training period a leave-one-out cross-validation is performed. This approach is equivalent to the validation technique described in Section 4.5. However, the leave-one-out cross-validation technique is an optimistic validation strategy for the model development since every day a recalibration is performed. Nonetheless, if a model becomes operational, it usually runs for a certain forecast period (e.g.

one year) without any recalibration. Consequently, the predictions can be only performed with past weather situations. To describe this strategy of the model development in a more suitable way, the model settings determined for the training period by the cross-validation technique are transferred to an independent data set used for the validation.

The selection process of analogs is also restricted to ±3 months centered on the target day. The window length is 180 days. The number of analogs is 30. To evaluate the skill of the methodology, the model performance is compared to three reference algorithms:

- **EDs**: An algorithm with the Euclidean distance and a small predictor domain (10° x 10°) which is centered on the target region (50° x 7.5°). This algorithm has a uniform distribution of the weights.
- **EDl**: An algorithm with the Euclidean distance and a large predictor domain (20° x 20°) which is centered on the target region (50° x 7.5°). This algorithm has also a uniform distribution of the weights.
- **WEDs**: An algorithm with a weighted Euclidean distance which is based on a symmetric distribution with $\boldsymbol{\mu}_j = 0°$, $\boldsymbol{\mu}_j = 0°$ and $\boldsymbol{\sigma} = 2$ equal to the initial guess of the optimization.

The further settings (e.g. the selection rule and the number of the most similar neighbors) of the reference algorithms are equivalent to the new algorithm.

## 6.3 Results

### 6.3.1 Air Flow Indices and Data Resolution

Probably, the most important step of the model development is the selection of an appropriate predictor. The 1000 hPa height field at 18 UTC is a suitable predictor to explain daily precipitation from 6 UTC to 6 UTC of the following day (see Section 4.6.1). In this section the height field is taken to calculate the height anomaly (GPHa) and two air flow indices: the zonal flow velocity (UGPH) and the resulting geostrophic flow strengths (TGPH). To determine the model performance of each predictor over the entire range of the predictand, the mean Brier skill score of the reference algorithms (EDs, EDl and WED) is calculated for different precipitation thresholds.

The skill of the resampling algorithm is in general quite similar for the various predictors which is shown in Figure 6.2a but the geopotential height and the height anomaly slightly outperform the other predictors for light and for moderate precipitation. The relative differences of the Brier skill scores $\Delta BSS = (1 - BSS/BSS_{ref}) * 100$ between the new predictors (GPHa, UGPH and TGPH) and the height field are illustrated in Figure 6.2b. A positive value indicates a model improvement due to the selection of a new predictor and negative value a model worsening. The algorithms based on the height field slightly outperform the algorithm with the height anomaly and the air flow indices for intensive precipitation. This result indicates that the height field is probably one of the most suitable pressure related variables for the prediction of daily precipitation.

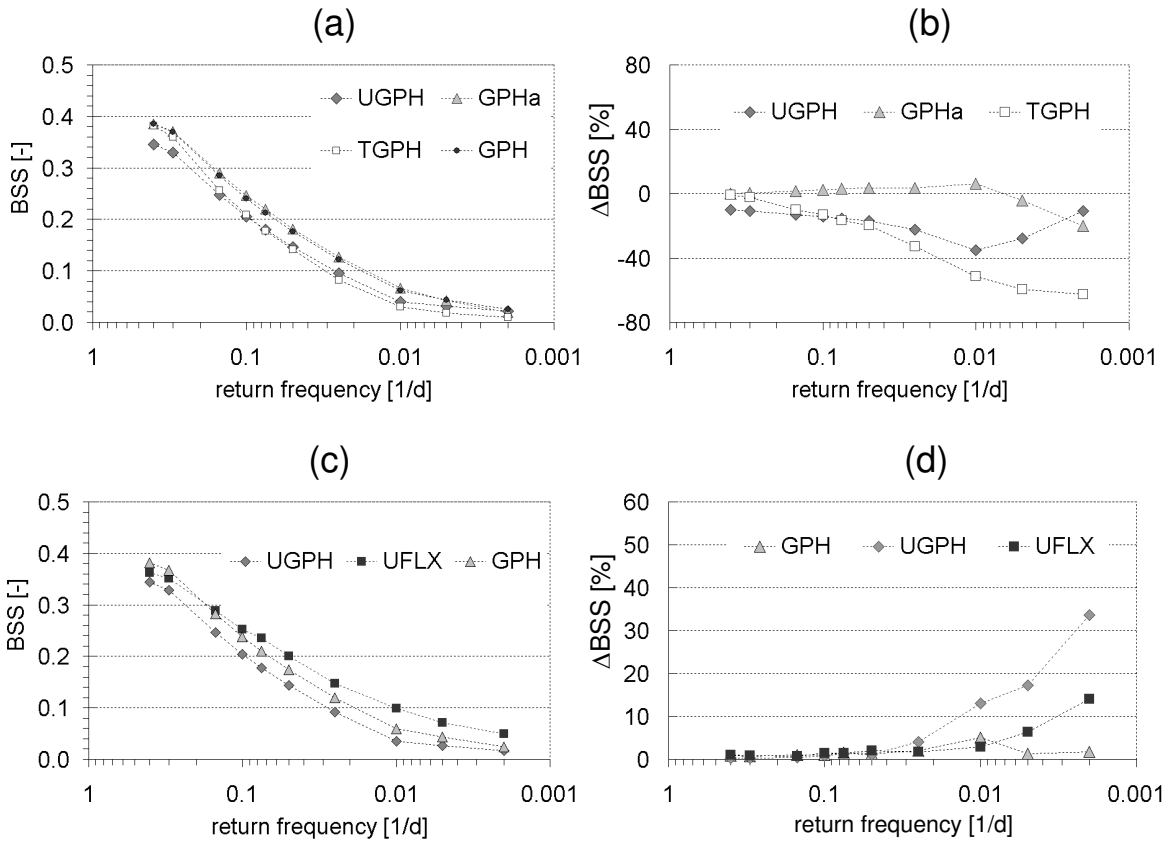



**Figure 6.2:** Figure a and b: Performance of the analog method for different pressure and moisture related variables. Figure c and d: Performance of analog forecasting for high-resolution and low-resolution reanalysis data. GPH = geopotential height, UFLX = zonal moisture flux, UGPH = zonal flow velocity, TGPH = geostrophic flow strengths, GPHa = height anomaly. Period from 1960 - 2000. Study region WG and MG.

The model performance of the analog method also depends on the spatial resolution of the predictor information. It is often assumed that the selection of a high-resolution data set outperforms low-resolution information in statistical forecasting. To investigate this question, the mean model performance of the reference algorithms is calculated for three predictors based on the low-resolution data (5° x 5°) illustrated in Figure 6.2c. The model performance of the resampling algorithms with the low resolution information indicates a suitable skill for the prediction of precipitation for all predictors.

In Figure 6.2d the performance of the algorithms is compared to the model performance of the same algorithms which is based on the original reanalysis information in 2.5°. The selection of the information with a higher resolution outperforms the other strategy for each predictor. However, it is questionable, whether the small improvement for the height field is significant. This result indicates that an increase of the data resolution has only a minor effect on the forecast performance of the algorithm based on a predictor with a low spatial variability. The influence on the

performance is clearly larger for variables with a higher variability such as the zonal moisture flux or the air flow indices, especially for extremes.

### 6.3.2 Optimization Performance and Metric Weights

The performance of the optimization algorithm is given for three parameterization schemes in Figure 6.3 for the strategy "intensive precipitation". The curves illustrate that the model performance can be increased for all three parameterizations compared to the initial guess. However, the selection of the highest parameterization scheme with five peaks is slightly inferior compared to a parameterization scheme with three peaks. This result shows that there are an optimal number of parameters needed for the description of the metric coefficients. In this example, a distribution function with three peaks and nine parameters is already a suitable parameterization scheme for the approximation of the metric weights. There is still a further possibility to reduce the number of parameters which might describe the non-uniform distribution of the metric coefficients in a suitable way. A direct consideration of the anisotropy in Equation 6.2 would only require two more parameters so that only five parameters must be estimated.

The metric coefficients of this parameterization scheme are given as standardized

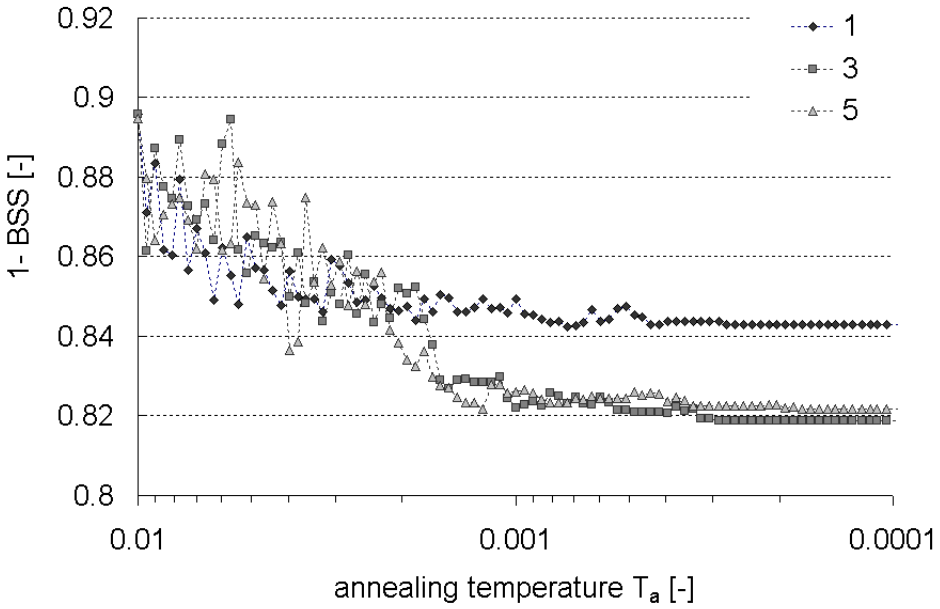



**Figure 6.3:** Optimization strategy for "intensive precipitation": The value of the objective function for three parameterizations schemes in relation to the annealing temperature. 1 = distribution function with a single peak, 3 = distribution functions with three peaks, 5 = distribution functions with five peaks.

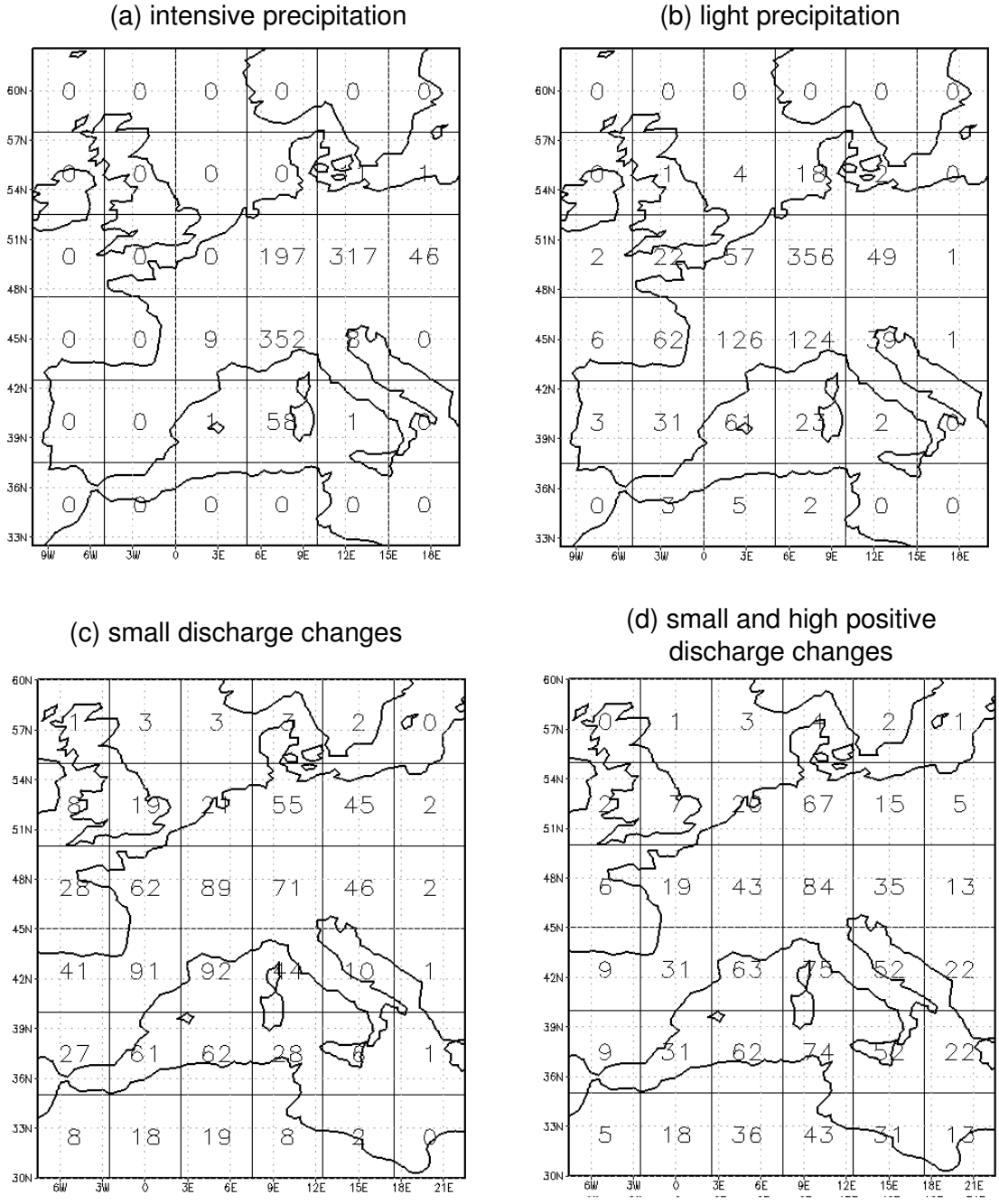



**Figure 6.4:** Calculated metric coefficients at each grid point for four optimization strategies.

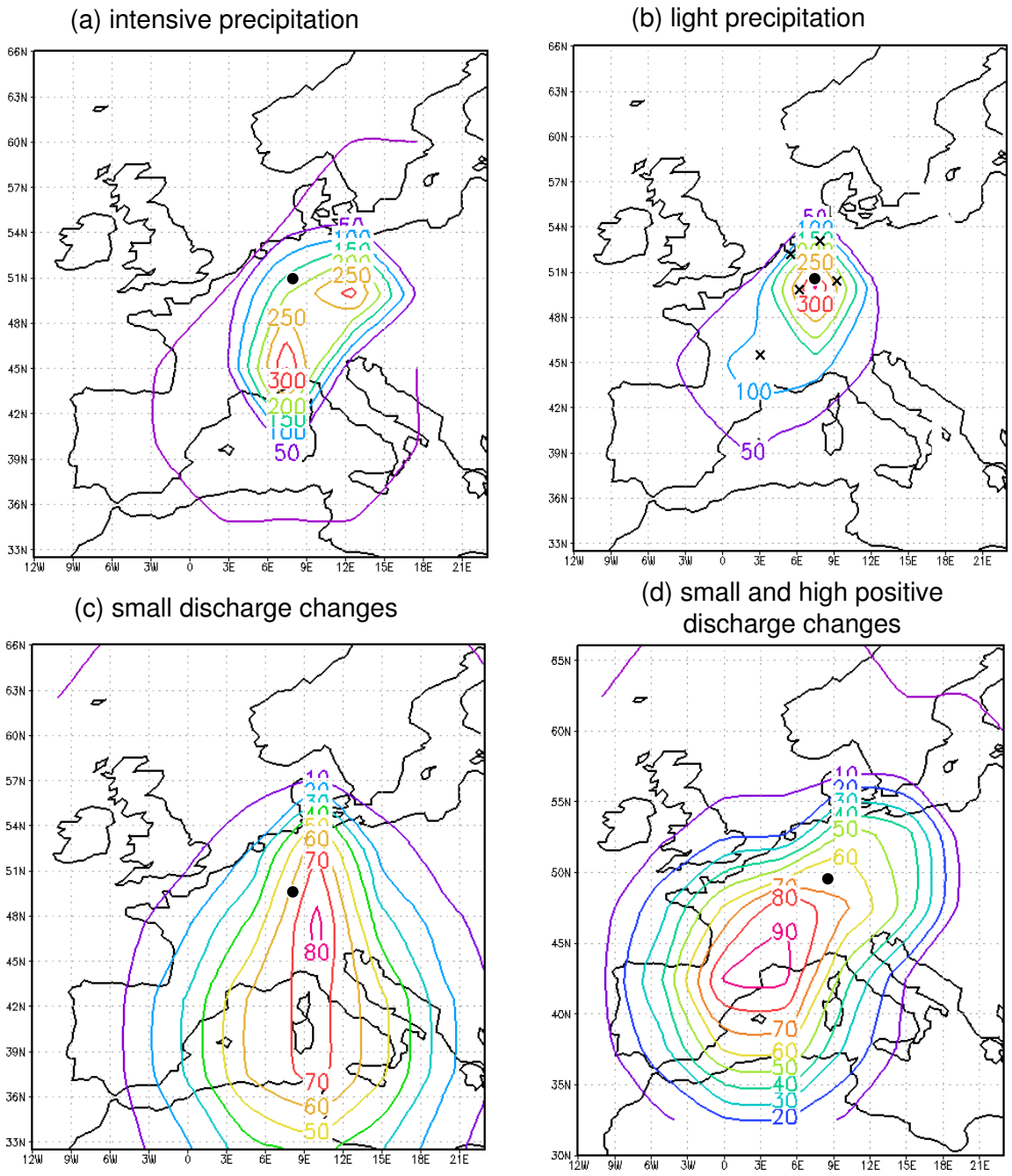



**Figure 6.5:** Contour map of metric coefficients for four optimization strategies interpolated from the values given in Figure 6.4. The dot marks the center of the study region. The crosses of the upper right map indicate the locations of the peaks for the selected parameterization scheme.

weights in Figure 6.4a for each grid point and in Figure 6.5a as contour plot. The contour plot indicates that the predictor domain can be clearly specified by the optimization algorithm. However, the predictor domain seems to be less reliable since the domain is quite small and the spatial distribution of the metric coefficient is characterized by a peak which is situated to the east of the catchments. The standardized metric coefficients illustrate that only a small number of grid points are needed to explain intensive daily precipitation for the study region.

The predictor domain of the height field specified by the second objective function ("light precipitation") seems to be more reliable (see Figure 6.4b and Figure 6.5b). The predictor domain is larger and grid points that are located to the southwest of the study region are marked by a high weight. The crosses on the contour plot illustrate the peak location of the five distribution functions selected for the approximation of the metric weights. Unfortunately, this result shows an over parameterization because four peaks are centered next to the study region. The selection of a parameterization scheme with fewer peaks (e.g. two or three) reduces the over-parameterization. Such a parameterization scheme is probably as valuable for the description of the metric coefficients as a scheme with more parameters.

The predictor domains determined by the discharge differences are illustrated in Figure 6.4c and Figure 6.5c. Both domains are clearly larger compared to the previous domains. Usually, small discharge differences mark dry periods which are often characterized by high pressure over the study region. Since the spatial extension of high pressure systems is larger compared to low pressure systems, a larger predictor domain should be also needed for an adequate description of the atmospheric state on those days. Probably, the most reliable predictor domain is the domain specified with small and high positive flow changes (Figure 6.4d and Figure 6.5d). This domain is clearly larger compared to the others and the highest weights are obtained at grid points which are located to southwest of the catchments. Usually, south-westerly fluxes are warmer and can transport more moisture to the study region than northerly or westerly fluxes. This result also coincides with the outcomes listed in Section 4.6.2 where the domain center of the height field was southwest of the catchments.

### 6.3.3 Forecast Accuracy and Value

The model performance of the algorithm based on the optimization strategy "intensive precipitation" is illustrated for the training period in Figure 6.6. The performance of the model predictions based on the optimized weights is not as good on days with light precipitation compared to the reference algorithms. However, for moderate and intensive precipitation, the optimized model outperforms the reference algorithms. In Figure 6.7a the model performance of the optimized algorithms is compared to the mean model performance of the reference algorithms by calculating the relative differences between the Brier skill scores. The lines characterize the model performance of the training period for the four optimization strategies. The model predictions based on the optimized weights outperform the predictions of the reference algorithms, particularly for the extremes. The performance of the model based on the optimization strategy "intensive precipitation" is the strongest. Interestingly, if the optimized models are transferred from the training period to the validation

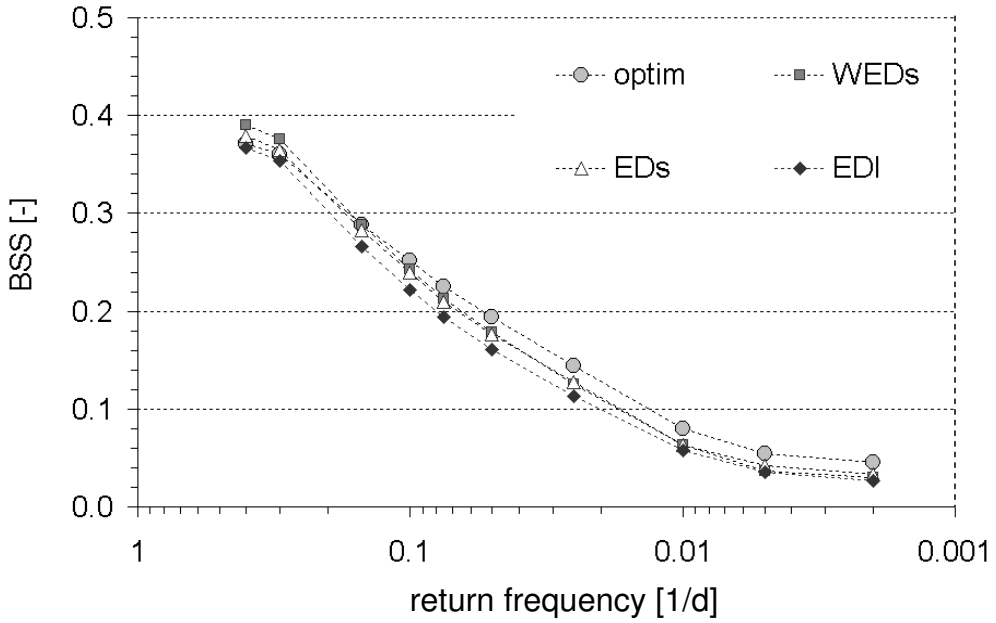



**Figure 6.6:** Forecast accuracy of analog forecasting based on the optimized coefficients (optim) in comparison to the reference algorithms (WEDs, EDs, EDl), training period.

period (see Figure 6.7b), the performance of the optimization strategy "intensive precipitation" becomes the poorest for the prediction of rare events. This strategy is even poorer as the baseline. On the other hand, the optimization strategy "small and high positive discharge differences" clearly outperforms the reference algorithms over the entire range of the predictand which indicates that the metric coefficients are suitably estimated. The second optimization strategy "light precipitation" also performs acceptably for the validation period. However, this optimization strategy seems to be less reliable for the prediction of intensive precipitation. The model performance is only known for smaller extremes due to the shorter validation period. The performance also decreases for the prediction of the largest extremes during the training period.

In Figure 6.6 the model performance of the analog method is compared to the low-skill reference forecast (which is the climatological average for the Brier skill score). The curves illustrate a high model performance for light precipitation but a poor performance for the extremes. However, a low accuracy of the model predictions must not correspond to a low forecast value. This problem is illustrated in Figure 6.8. Here, the maximum forecast value and the user interval (see Section 4.7) of the algorithm with optimized weights is compared to the mean scores of the reference algorithms. Both performance measures illustrate a high forecast value over the entire range of the predictand and in particular for the extremes. These data also illustrate that the algorithm based on the optimized weights outperforms the

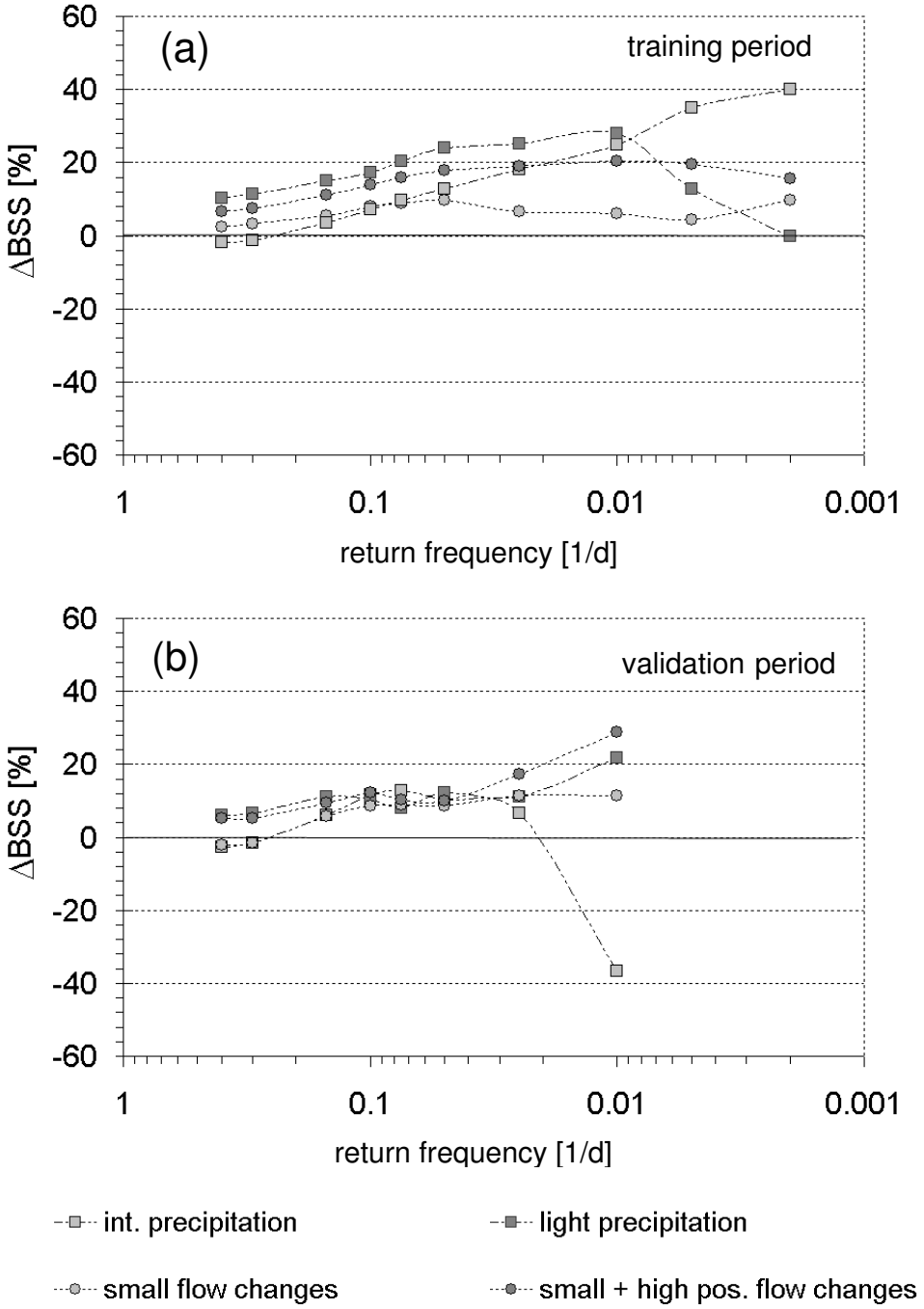



**Figure 6.7:** Model performance of analog forecasting in comparison to the mean performance of the reference algorithm for the training and validation period.

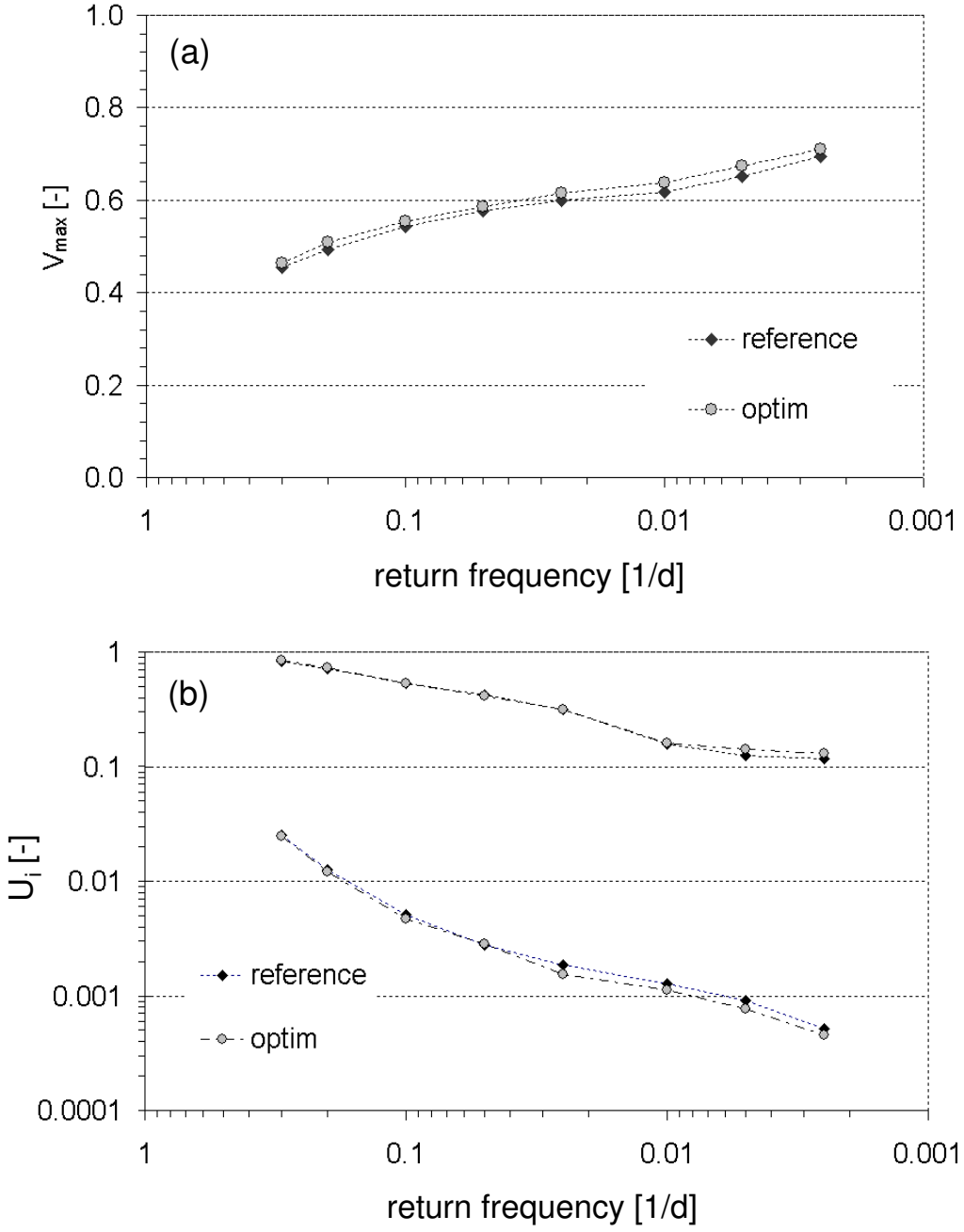



**Figure 6.8:** Forecast value of analog forecasting with optimized metric weights (optim) in comparison to an algorithm with a poor distribution of weights (reference). Figure a: Maximum forecast value $V_{max}$. Figure b: User interval $U_i$, training period and validation period.

reference algorithms for both performance measures. However, the increase of the forecast value is negligible so that negative outcomes (the number of false alarms and misses) of a warning system are only slightly decreased by the new algorithm.

### 6.3.4 Selection Rule

In Section 4.6.5 the selection process of the analogs is restricted by a moving window to account the intra-annual variability of the predictand and to accelerate the search process. However, it could be also useful to introduce CPs into the search process. In this case, the algorithm would only select analogs which occur in the same (similar) season and which belong to the same (similar) CP.

To investigate this issue, the data is separated into various subsets according to the rules given in Figure 6.9. At first, the entire information is divided into three different ways (no separation, the winter and summer half-year and the season) to take into account the intra-annual variability. Then, this information is again subdivided to incorporate the CP-variability. For the subdivision, the CP-catalog of the Rhine classification is selected which, itself, was determined by the daily discharge differences (see Table 5.2). This classification is then divided by the wetness index $W$ either in two groups (wet 2: "dry CPs" with $W < 1$; "wet CPs" with a $W > 1$) or into four groups (wet 4: "very dry": CP 2, CP 3 and CP 10; "dry": CP 1, CP 9 and CP 8; "wet": CP 4, CP 7 and CP 11; "very wet": CP 5, CP 6 and CP 12). The number of subsets for each rule is denoted by the values in the parenthesis in Figure 6.9. For example, the sixth rule divides the data into eight subsets.

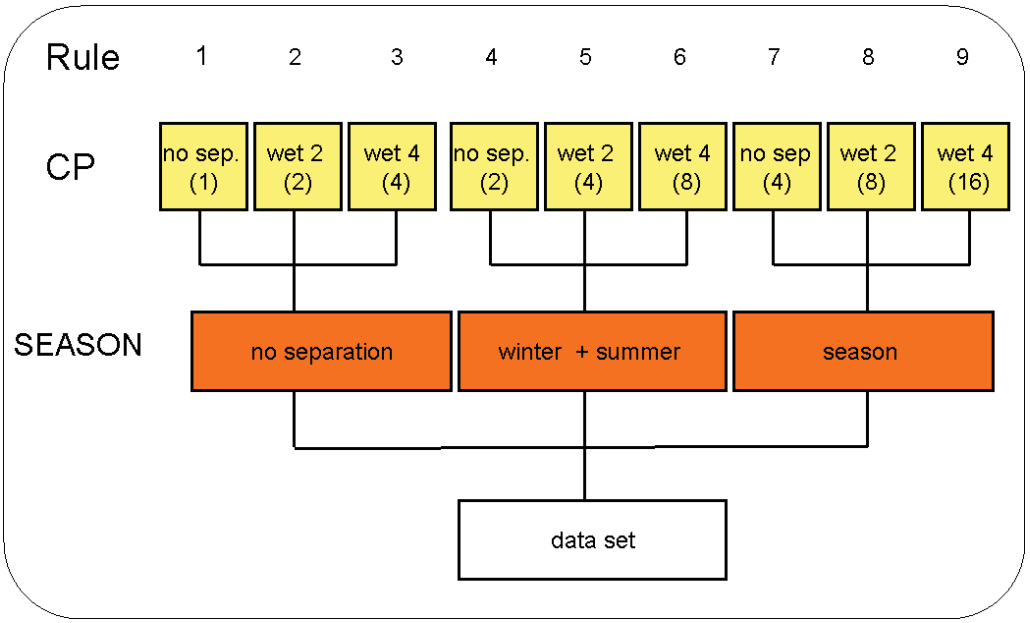



**Figure 6.9:** Subdivision of the data to account the intra-annual variability and the CP-variability in the selection process. The number of subsets of each rule is indicated by the number in the parenthesis.

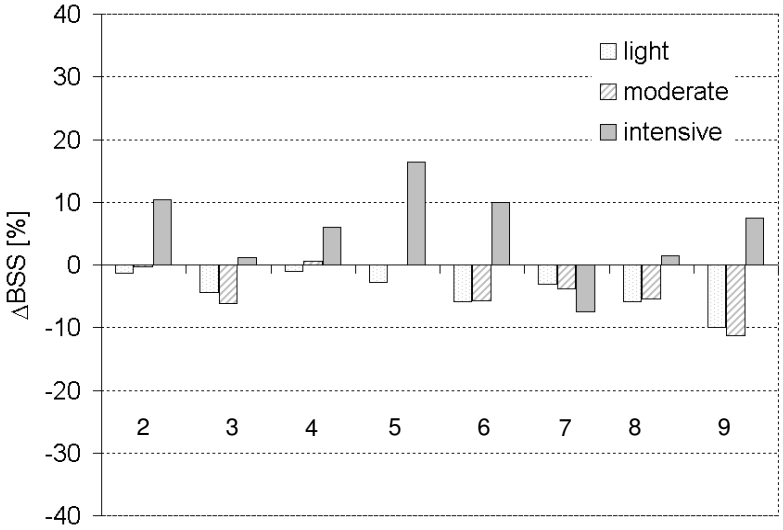



**Figure 6.10:** The influence of the selection rule on the model performance in comparison to an algorithm without any restriction (rule 1). The number indicates the identifier of the rule.

The investigation is performed with the analog algorithm based on the optimized metric weights determined by the optimization strategy "small and high positive discharge differences". The model performance of the algorithm is evaluated with the mean Brier skill score. It is calculated for an investigation period ranging from 1960 to 2001 and for catchments located in WG and MG. The Brier skill score is determined for three precipitation categories defined by the following event thresholds: light precipitation $\boldsymbol{s} = [0.60, 0.70, 0.80]$, moderate precipitation $\boldsymbol{s} = [0.85, 0.90, 0.95]$ and intensive precipitation $\boldsymbol{s} = [0.98, 0.99, 0.995]$. The model predictions are compared to the reference algorithm without any restriction (rule 1) by calculating the relative differences between the Brier skill scores.

The influence of the selection rule is illustrated in Figure 6.10 which shows that the second, the fourth and the fifth rule are probably the most suitable rules because the prediction of intensive precipitation can be increased without decreasing too much the skill for smaller events. A comparison of the rules indicates that the CP-variability seems to be as important as the intra-annual variability since the model performance slightly increases for intensive precipitation when CPs are introduced in the search process (see e.g. rule 2 and compare rule 4 to rule 5). However, the selection of more than eight categories (e.g. rule 6 or 8) clearly reduces the model performance for the prediction of light and moderate precipitation. Note that this investigation has been also repeated with all CPs for the subdivision but, for this configuration, the model performance is clearly poorer.

The investigation illustrates that the CP-variability should be taken into account to accelerate the search process and to improve the prediction for intensive precipitation. In addition, the results showed that the model performance is only slightly higher than the performance of an algorithm without any CP information. However,

there are still several ways to advance the selection process of the analogs. For instance, we can use a moving window to account for the intra-annual variability such as that shown in Section 4.6.5 and for the CP-variability. The selection of a moving window smoothes out the intersection between the different categories and should improve the prediction for those days which are in the transition zone between two categories.

## 6.4 Summary and Conclusions

The performance of analog forecasting depends on several criteria which must be specified by an investigator. The task of the investigator is to define suitable parameter values for the different criteria in order to allow for high model transferability and to maximize the prediction performance. However, a suitable selection of these criteria is time-consuming since a guideline for supporting a decision is not available. To tackle this problem, we presented an optimization algorithm in this chapter that can be used to define an appropriate distance function for the selection of analogs.

The methodology was tested for the prediction of daily areal precipitation for small river basins located in the Rhine basin (study region WG). We illustrated that the investigator must carefully select the objective function, otherwise the model transferability of the approach is poor. The intuitive choice of an objective function that focuses on the prediction of intensive precipitation is less promising. Therefore, an alternative way was proposed that is similar to the optimization strategy proposed in the previous chapter for the classification. This strategy focuses on the prediction of small discharge differences and high positive discharge differences. It was highlighted that this strategy is a suitable optimization strategy. It allowed for the determination of an appropriate distance function. Furthermore, a comparison of the new approach to the traditional way of analog forecasting illustrated that the prediction of intensive precipitation can be even improved by the new technique over the entire range of the predictand.

We can summarize that the optimization algorithm has a high potential to replace the costly manual calibration. A user with less knowledge can even obtain similar results as an investigator with much more experience. The optimization algorithm also presents a prototype of a methodology which enables an automated recalibration of the model parameters of an operational model. The methodology is not restricted to the prediction of daily areal precipitation. It can be also applied for the downscaling of other surface variables, such as daily temperature or daily wind speed.

There are still many aspects that need to be addressed in future studies. In this investigation the methodology was tested for the definition of an appropriate distance function for a pressure related variable, the geopotential height. We know, however, that the selection of further predictors can improve the prediction of intensive precipitation. In this respect, the optimization algorithm should be advanced so that further predictors, such as the moisture flux, can be used for the downscaling. This means also that future studies should find a way to define a suitable distance functions for large-scale information with a higher spatial resolution. The metric optimization performed in this chapter was done only on a 5° grid. Even though,

this could be carried out at a higher resolution using the readily available reanalysis information.

# 7 Data Depth

It is obvious that weather states which cause long-lasting large-area precipitation are unusual. They are often characterized by low pressure, a strong pressure gradient and a high moisture flux over the region of interest and by a slow movement of a low pressure system. In analog forecasting such weather states are identified by comparing the unusual weather state with all observed states to identify a subset of similar ones. In other words, this methodology performs a state-by-state comparison which is especially useful if many similar situations have been observed so that a suitable set of analogs can be identified for a particular forecast day. Unfortunately, unusual weather situations are rare events. In this case, it can be possible that only a limited number of appropriate analogs can be identified. For example, if a new weather situation develops over Central Europe in the future due to climate change, it could be even possible that this situation is not identified as an unusual state since no suitable analog has been observed so far.

In this chapter a further concept for the identification of rare weather states is proposed which might overcome the aforementioned problem. Instead of a direct comparison of two weather states, the centrality of a weather state using the concept of the data depth is measured. The closer a weather state is to the center, the higher the data depth becomes. If the actual state is identical to the center, the data depth is maximal and the state represents a usual weather state. Unusual weather states are indicated by low data depths.

The idea of measuring the centrality of a weather state seems to be a very reasonable concept for the downscaling since we know that extremes at the local scale are often caused by anomalies of the atmosphere. The data cloud of the daily height anomalies at two grid points which are located west and east to the study region WG (for a description of the study region see Chapter 3) are shown in Figure 7.1a. Pairs situated near the boundary of the data cloud are unusual atmospheric states. We know that those situations can cause local extremes. For instance, pairs with strong negative height anomalies indicate days with low pressure over the study region which are often characterized by a higher likelihood of intensive precipitation (Figure 7.1b). Data pairs with strong positive anomalies can be linked to high pressure which is usually a suitable indicator for good weather, while data pairs with strong positive and negative anomalies describe atmospheric states with high pressure gradient which can also correspond to a higher likelihood of intensive precipitation. This is especially for those days when the atmospheric state over the study region is characterized by strong westerly pressure gradients. In this event, wet air masses can be transported from the Atlantic Ocean to Europe.

The aforementioned example illustrates that the centrality of a weather state seems to be a suitable indicator for the identification of local extremes. Apparently, this kind of technique has been not used for the downscaling of local surface variables so far. Thus, the performance of this approach for the prediction of intensive pre-

cipitation is not known. To shed some light on this new downscaling technique, the following points are investigated in this chapter in more detail:

1. The functionality of a data depth function is illustrated for an independent and dependent standard normal distributed data set to identify unusual weather states.

2. The data depth is determined for the geopotential height and zonal moisture flux anomalies for each day of the investigation period to describe the actual state of the atmosphere over the study region. Then, a catalog of daily cir-

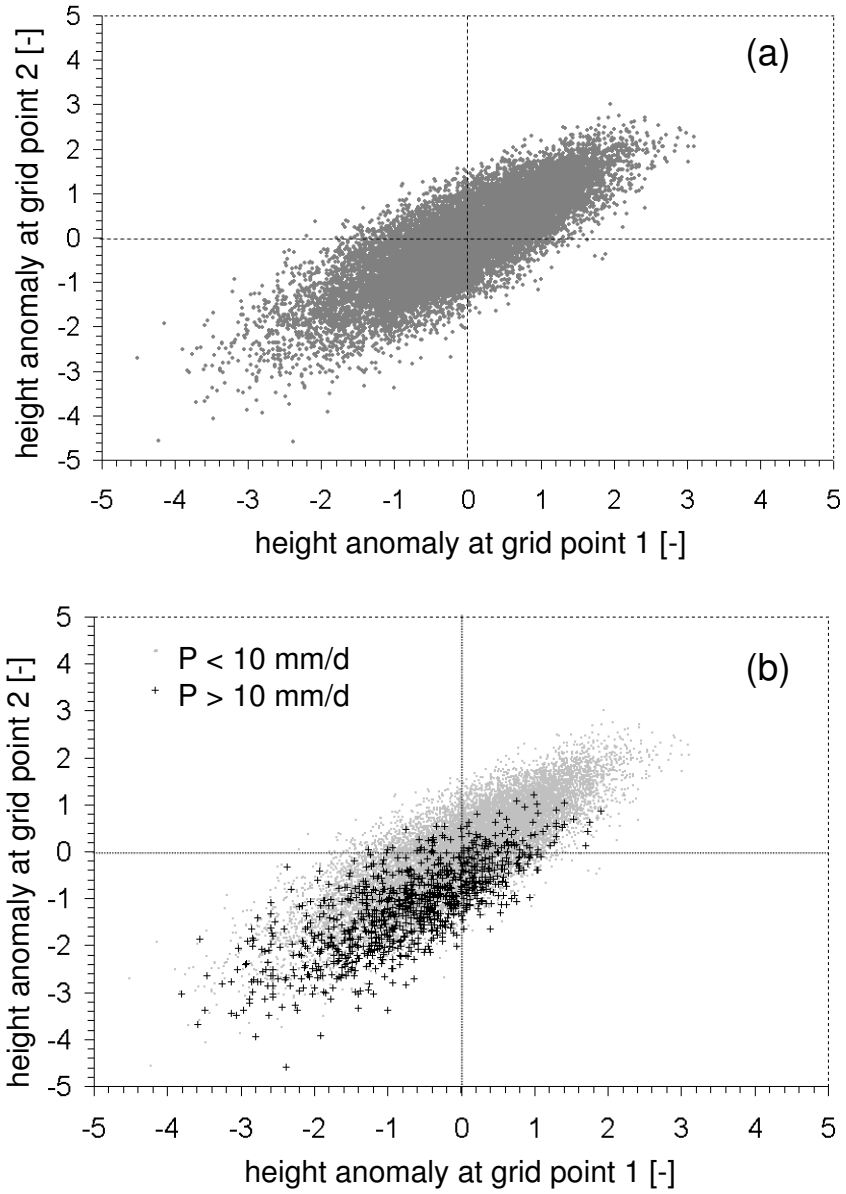



**Figure 7.1:** Scatter plot of the time series of daily height anomalies at two grid points close to the study region for two precipitation (P) categories. Grid point 1: 50.0° N and 5° E. Grid point 2: 50.0° N and 10° E. Investigation period from 1958 to 2001.

culation patterns (CP) is selected to distinguish between dry and wet days. Afterwards, the conditional precipitation probability for a given data depth is determined for each CP.

3. An index is proposed to describe in a suitable way the risk of a given weather state for intensive precipitation.

4. The relative value of the downscaling approach is calculated to determine the model performance for the detection of intensive daily areal precipitation. Finally, the model performance of this approach is compared to the analog method.

In the previous chapters the performance of the downscaling approaches has been analyzed for the detection of intensive precipitation in head catchments and small tributaries ($A_E \approx 1000$ to $4000\ km^2$) of the major rivers in Germany. In this chapter the catchment scale is increased to evaluate the model performance for larger catchments such as the main tributaries of the major rivers in Germany, in this case the Main or the Neckar. Their catchment size ranges between 10000 and 40000 $km^2$. The increase of the catchment size has the effect that the influence of small-scale convective precipitation events is reduced and the model performance of the approach is skewed towards precipitation events which are mainly caused by large-scale atmospheric anomalies. As illustrated briefly at the beginning of Chapter 5, the river floods in Central Europe were primarily caused by large-scale weather phenomena in the last two decades.

In the following section the depth function selected in this study is proposed and its functionality is tested with two examples.

## 7.1 How Can We Measure the Centrality of a Weather State?

There are many functions to measure the centrality for a data point of a multivariate distribution. All of these functions might be suitable depth functions for the selection of unusual weather states. An overview of data depth function can be found in Liu et al. (1999). The major reason for the variety of depth functions is that the definition of the centrality of a point is not intuitive for a multivariate distribution. For example, there are already several concepts which are used to measure the centrality of a data point for a univariate distribution such as arithmetic mean, median or $\alpha$-trimmed mean.

The most commonly used data depth function is probably the half-space depth proposed by Tukey (1975). Recently, Bárdossy and Singh (2008) selected this measure to develop a robust parameter estimation algorithm for hydrological modeling.

In this investigation the $L_1$-depth is selected, since the value of this function is easy to specify for a multivariate distribution in comparison to other depth functions. The $L_1$-depth function forms thereby an appropriate starting point to investigate the performance of the function for the detection of precipitation extremes.

The $L_1$-depth of an atmospheric state described by a predictor $\mathbf{x}(t_0) = (x_1, x_2, ..., x_K)$ with K dimensions at time $t_0$ of a data set $\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), ..., \mathbf{x}(t_N)]$ with $N$ time steps can be calculated as follows (Vardi and Zhang, 2000):

$$L_1[\mathbf{x}(t_0)] = 1 - |\sum_{i=1}^{N} f_i \frac{\mathbf{x}(t_0) - \mathbf{x}(t_i)}{|\mathbf{x}(t_0) - \mathbf{x}(t_i))|}| \tag{7.1}$$

$f_i$ is the weight of an atmospheric state. The sum of all weights is one. Like other depth functions, the values of the $L_1$-depth ranges between zero and 1. The closer is the atmospheric state to the center, the higher the $L_1$-depth.

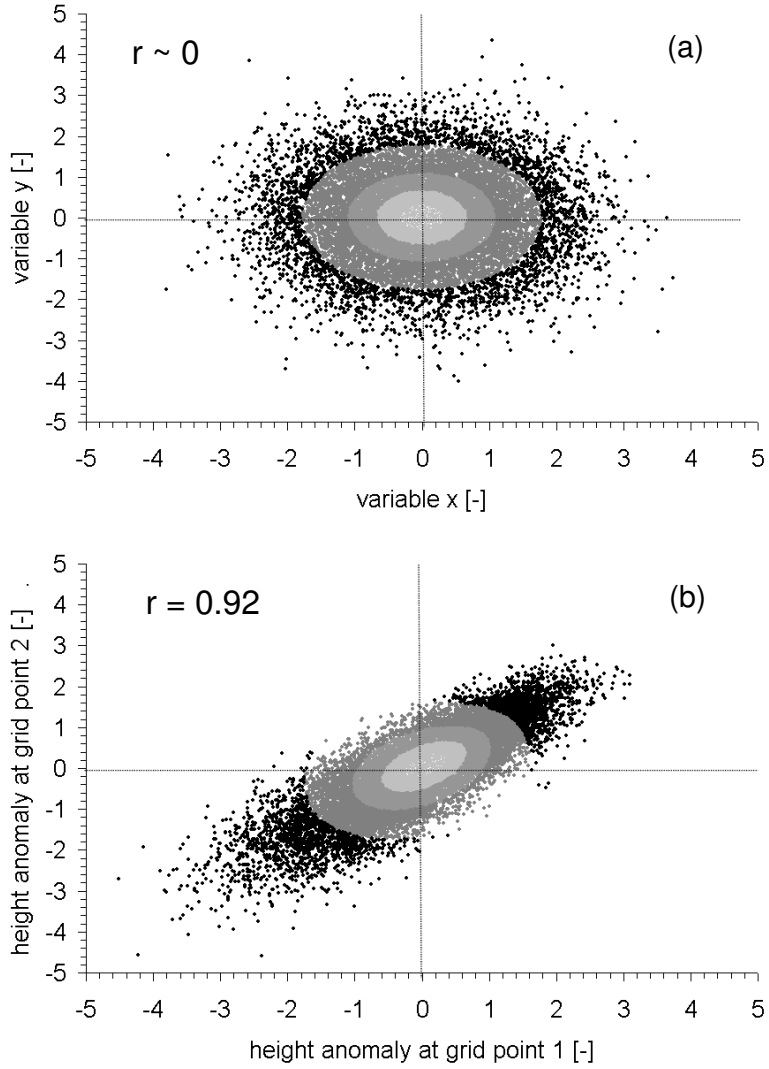



**Figure 7.2:** Figure a: Data cloud of two independent variables. Figure b: Data cloud of two dependent variables. Height anomalies. Grid point 1: 50.0° N and 5° E. Grid point 2: 50.0° N and 10° E. Investigation period from 1958 to 2001. r = Pearson correlation coefficient.

For the calculation of the data depth, the atmospheric states are divided into four subsets according to the seasons to take into account the intra-annual variability of the downscaling process. It should be taken into account that the data depth of a particular state is only calculated based on the states of the same subset. If we express this formulation in terms of the weights given in Equation 7.1, the weight of all atmospheric states which belong to the same subset of the state of interest is 1, while the weight is zero for all other states.

The functionality of the $L_1$-depth is illustrated in Figure 7.2a for two independent standard normal distributed variables. The scatter plot shows the depths of all data pairs for five data depth categories (*"very high"*, *"high"*, *"moderate"*, *"low"* and *"very low"*). The $L_1$-depth only designates those pairs as extremes which are located far away from the origin. Those pairs are indicated in Figure 7.2a by black dots. However, pressure related variables like height anomalies have a high spatial dependence. In this case, the time series of two pressure related variables which are located at grid points close to a study region are not independent and the resulting data cloud is elliptical (Figure 7.2b). In the event of this, the $L_1$-depth can only identify those atmospheric states as unusual situations when high negative or high positive anomalies occur to the same time at both grid points. Pairs with strong pressure gradients are not indicated as extremes. These pairs have only a low data depth although they are situated near the boundary of the data cloud. Actually, they represent unusual states which could also cause intensive precipitation over a study region (see Figure 7.1b). This examples illustrates that the $L_1$-depth is for certain cases not a suitable indicator for unusual situations. To overcome this problem, future studies should test further depth functions like the half space depth which might improve the selection of unusual atmospheric states.

## 7.2 Applications

The investigation is performed for the study region WG and MG (see Chapter 3) by selecting the anomalies of the 1000 hPa height field (GPHa) and the 700 hPa zonal moisture flux field (UFLXa) at 18 UTC. The absolute values of both variables are suitable predictors for intensive precipitation, already demonstrated for analog forecasting in Chapter 4.

To point out an adequate dimension for the calculation of the depth, a preliminary investigation has been performed. The results of this investigation indicated that only a few grid points ($< 15$) are needed. In this investigation a small quadratic predictor domain with nine grid points serves as first guess. The predictor domain is centered over the study region. It ranges from 47.5° S to 52.5° N and from 5.0° E to 10.0° E. The investigation period is from 1958 to 2001.

To analyze the model performance for precipitation extremes which may cause severe inundations in river basins which are an order of magnitude larger compared to the selected subbasins, the daily areal precipitation of the catchments located in MG and WG is averaged for each day of the investigation period. In this case, the influence of convective precipitation events is reduced and the verification of the downscaling approach focuses more on precipitation events which occur in both study regions at the same time.

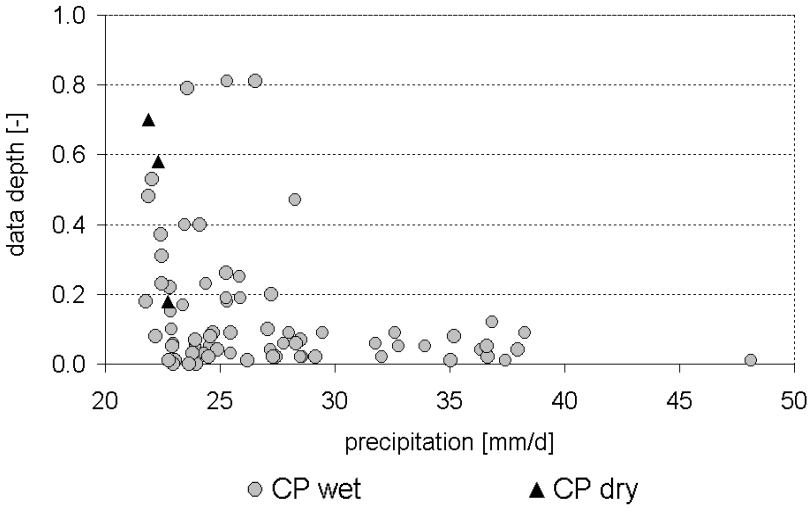



**Figure 7.3:** Data depth of the height anomaly in relation to daily areal precipitation extremes (circle = wet CPs, triangle = dry CPs). Mean daily areal precipitation of the catchments located in WG and MG. Investigation period from 1958 to 2001.

The data depth of the height anomalies for the 80 largest precipitation events in the study regions WG and MG is presented in Figure 7.3. The example illustrates that in many cases a low depth of the height anomalies can be measured if an intensive precipitation event occurs. Thus, there might be a relationship between data depth and intensive precipitation. The data pairs are also marked by triangles and circles to distinguish between precipitation extremes that have occurred either during a wet or dry CP. For the subdivision of the extremes, the Rhine classification presented in Table 5.2 of Chapter 5 was selected . It shows the high performance of the fuzzy-rule based classification for the detection of large-scale precipitation extremes since only three smaller extremes occur during a dry CP.

## 7.3 Precipitation Probability and Risk Index

The precipitation probability conditioned on the data depth of the height anomaly and the moisture flux anomaly are given in Figure 7.4 for an event which is larger than 2 mm/d. It describes that the conditional probability rises with decreasing data depth for both predictors; the relationship between data depth and precipitation probability is only moderate for both predictors. This result is actually not surprising since the $L_1$-depth identifies all weather states with low and high height anomalies over the study region as extremes, as illustrated in the example of Figure 7.2b. In this case, the extremes are a mixture of very dry and very wet situations. The relationship between data depth and precipitation probability is slightly stronger for the moisture flux since strong negative as well as strong positive anomalies denotes weather states with a high moisture flux either from west or from east.

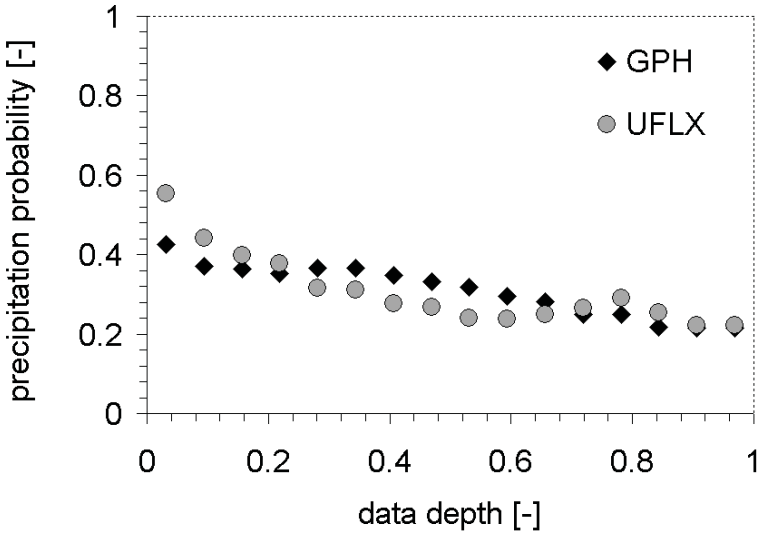



**Figure 7.4:** Precipitation probability conditioned on the data depth of the height anomalies (GPHa) and the zonal moisture flux anomalies (UFLXa).

To separate the mixture of unusual weather states, the CP-catalog compiled by the fuzzy-rule based classification is selected. For the classification the same CP-catalog is selected in the same manner as in the previous section. The influence of the classification is illustrated in Figure 7.5 for two circulation patterns: the wettest (CP 5) and the driest (CP 2). The gray dots in the background of the figure illustrate the pairs of all weather states. Due to the classification, many pairs with high positive anomalies belong to the driest CP while pairs with negative anomalies belong to the wettest pattern. If the conditional probabilities are calculated for the states of both patterns, the relationship between precipitation probability and data depth is stronger for both predictors compared to the relationship without any CP information. The precipitation probability rises with decreasing data depth for the wettest CP while for the driest pattern the effect is complimentary (Figure 7.6a). The precipitation probability of the wet pattern is approximately 90 % for very low depths, whereas for the dry CP the probability tends to zero for the same depth. Thus, a suitable identification of weather states with a high and low precipitation probability seems to be possible on the proviso that both classification and data depth is used for the downscaling.

In Chapter 5, the precipitation characteristics of a CP were described by the wetness index. It was illustrated that CPs with a high wetness can be used to identify many extremes so that this index seems to be a suitable indicator for CPs with a high risk of intensive precipitation. In this section a further precipitation index, the risk index $R_i$ , is proposed to describe the risk of a given atmospheric state for intensive precipitation. The idea of this index is similar to the wetness index. This risk index relates the conditional occurrence frequency $s_{i,l}$ for a given weather state $i$ to the climatological frequency of a precipitation extreme $s_l$ based on a corresponding threshold $l$:

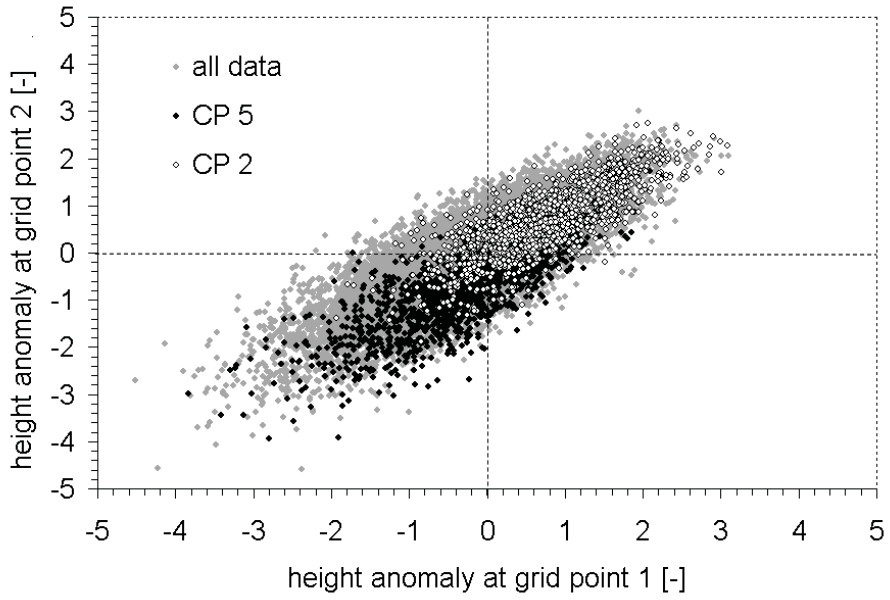



**Figure 7.5:** Scatter plot of the time series of daily height anomalies of two grid points close to the study region for the wettest (CP 5) and driest circulation pattern (CP 2). Grid point 1: 50.0° N and 5° E. Grid point 2: 50.0° N and 10° E. Investigation period from 1958 to 2001.

$$R_i = \frac{1/(1 - s_{i,l})}{1/(1 - s_l)} \tag{7.2}$$

Thus, the risk index is a measure that describes the risk of intensive precipitation of a certain weather state. This risk index can range between zero and infinity. The larger the index is the higher is the risk for intensive precipitation. If the index is equal to one, the risk of intensive precipitation is identical to the climatological mean. A risk index close to zero indicates a weather state with a very low risk of intensive precipitation.

Murphy (1991) noted that the risk index is also known as the odds ratio since the conditional odds $1/(1 - s_{i,l})$ are related to the climatological odds $1/(1 - s_l)$. Since precipitation probabilities are extremely small for rare events, it is very likely that the risk for extremes is underestimated by a decision maker. In particular in the case if the decision maker has less experience with forecast probabilities. To overcome this problem, the risk index is a suitable way to express the risk of a given weather state for intensive precipitation. A more detailed description of this problem is also given in Section 8.4.

In this investigation, the risk index is calculated for nine categories of the data depth conditioned again on CP 2 and CP 5 of the Rhine classification (Figure 7.6c and 7.6d). Thus, a certain weather state is described by its CP and its data depth information, calculated either from the height anomalies or the moisture flux anomalies. Here, the risk index is given for precipitation event which exceeds a threshold

of 15 mm/d. The corresponding occurrence frequency of the event is 2 %. The information in the Figure 7.6c and 7.6d illustrates that for the wet pattern the risk clearly rises if the data depth decreases. The risk is highest for the lowest category of the moisture flux, since the atmospheric state is characterized by a high zonal moisture flux over the study region. While for the dry pattern, the risk index is clearly smaller compared to the wet pattern. The risk index is even smaller compared to the climatological risk which indicates a low likelihood for intensive precipitation. The probability for an intensive precipitation event is smallest for very low data depth of the height anomaly. In this instance, the atmosphere is characterized by high positive height anomaly. This is usually linked to suitable weather conditions over the study region.

The risk index and further statistical properties are also listed in Table 7.1 for all CPs of the Rhine classification for the data depth of the moisture flux anomaly. The risk for intensive precipitation is very high on days with a wet CP and very low data depth. Thus, the selection of the classification and the data depth information allows for a clear identification of weather states with a high likelihood of intensive precipitation and for the determination of weather states with a low likelihood. If a decision maker takes all categories with a risk index being larger than one to trigger

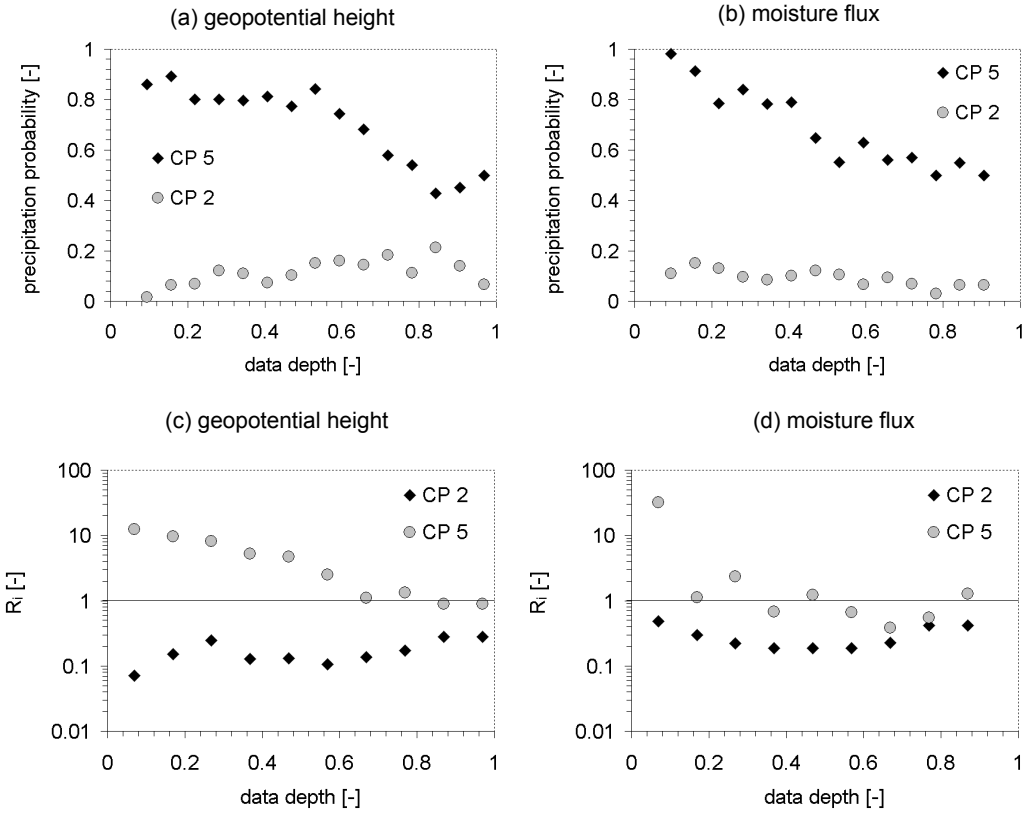



**Figure 7.6:** Conditional precipitation probability and risk index for the wettest (CP 5) and the driest circulation pattern (CP 2) of the Rhine classification.

**Table 7.1:** Risk index, number of extremes and relative frequency of nine data depth categories conditioned on 12 CPs. Rhine classification with 12 CPs. Investigation period from 1958 - 2001. WG and MG. 1000 hPa height field (GPHa) and the 700 hPa zonal moisture flux field (UFLXa) at 18 UTC.

risk index [-]

| depth | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.6 | 0.8 | 0.5 | 11.9 | 70.4 | 31.3 | 11.6 | 0.6 | 1.5 | 0.5 | 7.5 | 27.2 |
| 0.2 | 0.9 | 0.6 | 0.9 | 6.0 | 24.8 | 4.5 | 4.7 | 0.9 | 0.5 | 0.5 | 1.3 | 3.1 |
| 0.3 | 1.3 | 0.3 | 0.4 | 1.8 | 7.8 | 2.9 | 1.8 | 0.4 | 0.4 | 0.7 | 1.8 | 0.5 |
| 0.4 | 2.2 | 0.2 | 0.4 | 0.4 | 4.0 | 1.6 | 0.7 | 0.3 | 0.1 | 0.3 | 0.3 | 1.6 |
| 0.5 | 0.2 | 0.2 | 0.3 | 0.7 | 2.2 | 0.6 | 0.7 | 0.3 | 0.1 | 0.4 | 0.3 | 0.5 |
| 0.6 | 0.2 | 0.2 | 0.3 | 0.8 | 0.5 | 1.8 | 0.3 | 0.2 | 0.1 | 0.2 | 0.5 | 0.8 |
| 0.7 | 0.2 | 0.2 | 0.5 | 1.1 | 0.3 | 1.7 | 0.3 | 0.3 | 0.2 | 0.2 | 0.4 | 0.9 |
| 0.8 | 0.3 | 0.2 | 0.5 | 0.7 | 0.7 | 0.3 | 0.5 | 0.3 | 0.2 | 0.3 | 0.5 | 1.1 |
| 0.9 | 0.6 | 0.4 | 0.9 | 1.5 | 0.5 | 0.4 | 0.7 | 0.4 | 0.5 | 0.4 | 0.5 | 1.1 |

number of extremes [-]

| depth | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 0 | 0 | 0 | 3 | 32 | 6 | 4 | 0 | 1 | 0 | 5 | 11 |
| 0.2 | 0 | 0 | 0 | 4 | 18 | 2 | 3 | 0 | 0 | 0 | 1 | 2 |
| 0.3 | 1 | 0 | 0 | 2 | 11 | 3 | 2 | 0 | 1 | 1 | 3 | 0 |
| 0.4 | 4 | 0 | 0 | 0 | 6 | 2 | 1 | 0 | 0 | 0 | 0 | 2 |
| 0.5 | 0 | 0 | 0 | 1 | 4 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0.6 | 0 | 0 | 0 | 1 | 1 | 4 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0.7 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.8 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0.9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

relative frequency [%]

| depth | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.1 | 0.2 | 0.2 | 0.2 | 0.5 | 0.2 | 0.2 | 0.3 | 0.5 | 0.3 | 0.5 | 0.3 |
| 0.2 | 0.1 | 0.5 | 0.4 | 0.5 | 0.6 | 0.3 | 0.4 | 0.4 | 0.7 | 0.6 | 0.5 | 0.4 |
| 0.3 | 0.5 | 1.1 | 0.8 | 0.8 | 1.0 | 0.7 | 0.7 | 0.8 | 1.5 | 0.9 | 1.1 | 0.7 |
| 0.4 | 1.2 | 1.4 | 0.9 | 0.8 | 1.0 | 0.8 | 1.0 | 1.0 | 2.2 | 1.0 | 1.2 | 0.8 |
| 0.5 | 1.7 | 1.8 | 1.1 | 0.9 | 1.2 | 1.1 | 1.0 | 1.2 | 2.4 | 1.5 | 1.1 | 0.6 |
| 0.6 | 2.1 | 1.9 | 1.1 | 0.8 | 1.4 | 1.5 | 1.2 | 1.3 | 3.1 | 1.7 | 1.2 | 0.4 |
| 0.7 | 1.7 | 1.5 | 0.7 | 0.6 | 1.2 | 1.1 | 1.1 | 1.1 | 2.1 | 1.5 | 0.9 | 0.4 |
| 0.8 | 1.2 | 1.4 | 0.6 | 0.5 | 1.0 | 1.2 | 0.7 | 1.0 | 1.6 | 1.2 | 0.6 | 0.3 |
| 0.9 | 0.6 | 1.1 | 0.5 | 0.2 | 0.9 | 1.0 | 0.6 | 0.9 | 1.6 | 1.1 | 0.4 | 0.1 |

an alarm, he would be in the position to detect around 90 % of the extremes with an alarm rate of roughly 16 %. The maximum relative value of this warning strategy can be estimated by the Peirce's skill score which is the hit rate minus the false alarm rate (see Section 2.2.4). For the estimation of this score we can assume that the false alarm rate is equal to the alarm rate. In this case, the maximum value of the conditional downscaling approach is around 0.74. This forecast value indicates a high model performance for the detection of intensive precipitation events which affect both study regions in the Rhine basin to the same time. In the following section, the forecast value of the approach is determined over the entire range of the predictand and for the different users of the warning system. Furthermore, the approach is compared to the analog method.

## 7.4 Forecast Value

The downscaling skill of the approach is estimated by the relative value proposed in Equation 4.21 to evaluate the predictions for warning situations. The forecast probabilities for each day of the investigation period are calculated in the following way: At first, the 50-nearest data depths are determined from the entire subset of data depths of a CP identical to the pattern of the day of interest. Then, the corresponding precipitation values of this subset are taken to estimate the parameters of the mixed-exponential distribution (see Section 4.4). The decision thresholds needed for the calculation of the forecast value are the percentiles of the mixed-exponential distribution. To estimate the skill of the prediction over the entire range of the predictand, the forecast value is calculated for the following occurrence frequencies $\mathbf{s} = [0.70, 0.80, 0.90, 0.95, 0.975, 0.99, 0.995]$.

The maximum forecast value for the approach is illustrated in Figure 7.7a. It is compared to the forecast value of the analog method presented in Chapter 4. The maximum forecast value of the conditional downscaling approach is positive over the entire range of the predictand. The skill of the prediction is highest for intensive precipitation which indicates a suitable performance for the detection of intensive precipitation. However, in comparison to the analog method, the prediction of the conditional approach is less valuable which is indicated by both performance measures, the maximum forecast value and the user interval (Figure 7.7b).

Nevertheless, the results presented in this chapter are quite encouraging for the new downscaling technique since there are still many possibilities to improve this approach for the prediction of intensive precipitation. For example, the conditional approach only uses the moisture flux for the description of the atmospheric state over the study region as information, while the analog method selects the same information but also the geopotential height and the relative humidity. Thus, the integration of further large-scale predictors for the description of the atmospheric state over the study region could enhance the prediction for intensive precipitation. Furthermore, the identification of suitable parameters for the data depth approach could also improve the downscaling. In this investigation only a short preliminary investigation was performed to point out a suitable predictor domain for the height and the moisture flux anomalies.

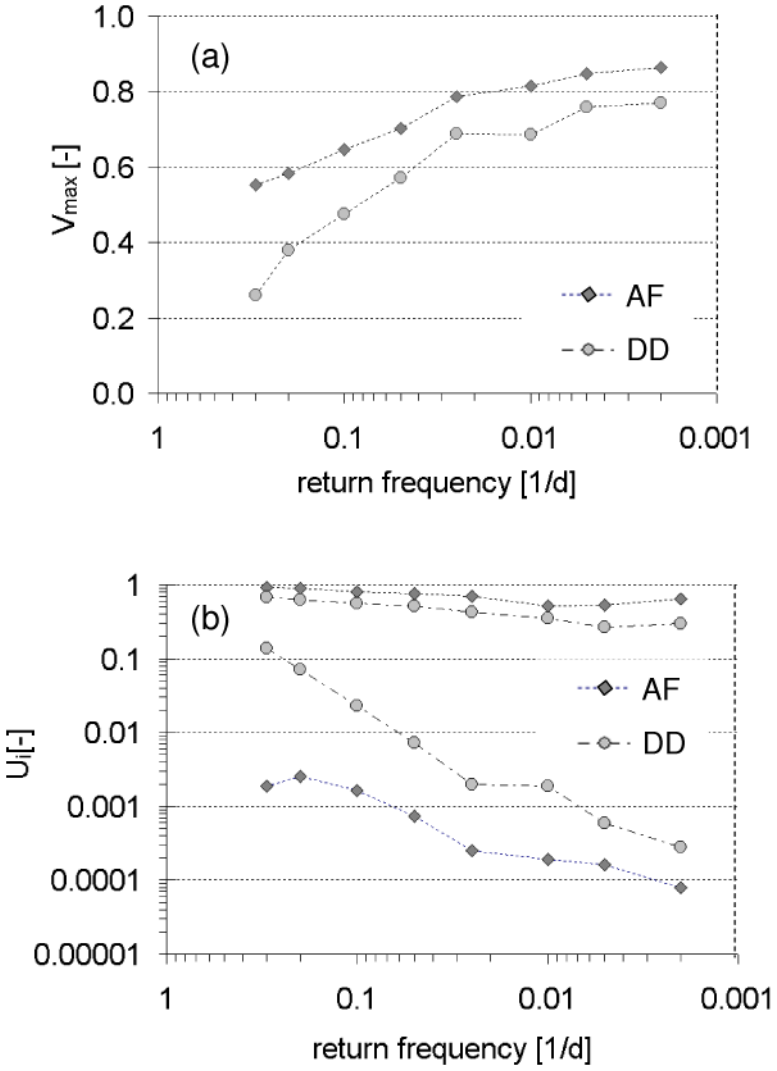



**Figure 7.7:** Forecast value of the conditional downscaling approach based on the data depth (DD) in comparison to analog forecasting (AF). Figure a: Maximum forecast value $V_{max}$. Figure b: User interval $U_i$.

## 7.5 Summary and Conclusions

In this chapter a downscaling approach based on a data depth function was presented for the prediction of intensive precipitation. A data depth function can be used to measure the centrality of a weather state. Since local extremes are often caused by atmospheric anomalies, this methodology seems to be a suitable technique for the identification of those events.

We illustrated that the selection of a data depth function based on daily height anomalies and without any additional information performs poorly for the detection of intensive daily precipitation events. This is because unusual height anomalies can correspond to intensive precipitation as well as to very dry conditions on the local scale. To overcome this problem, we selected a catalog of daily large-scale circulation patterns to classify the atmospheric states into dry and wet situations. Then, a strong relationship between the data depth and the precipitation probability could

be determined for a given circulation pattern. To describe the risk of a particular weather state, we used a risk index which relates the conditional occurrence frequency of an event to its climatological average. It was illustrated that the data depth conditioned on CPs allowed for the determination of weather states with high risk of intensive precipitation. The evaluation of the model performance for warning situations demonstrated that the predictions of the new downscaling technique have a high forecast value for the detection of intensive precipitation. The approach is nearly as efficient as the analog method.

Finally, we can summarize that the selection of the data depth conditioned on circulation patterns is also a valuable approach for the detection of intensive precipitation. An enhancement of this technique in future investigations would be worthwhile.

# 8 Operational Application

In the previous chapter the forecast performance of statistical downscaling techniques was evaluated with reanalysis information. Thus, the model performace was evaluated for past events under the assumptions that the large-scale atmospheric predictions of a global NWP model are perfect. To determine the performance of downscaling techniques with real time forecast information, a prototype of a probabilistic precipitation forecast system (PPFS) has been developed during this investigation. This forecast system transfers the outcomes of the Global Forecast System (GFS; NCEP, 2003) of the US National Center of Environmental Predictions (NCEP) to the local scale to estimate the forecast uncertainty of daily areal precipitation for small river basins in Germany.

The integral parts of the PPFS are a data assimilation system, a postprocessor and a statistical precipitation prediction model. The data assimilation system downloads the GFS predictions from the NCEP server and converts the forecasts into the format needed for the downscaling. The postprocessor visualizes the forecasts and collects all relevant observations and forecasts in a data archive. The key downscaling techniques of the statistical precipitation model are the fuzzy rule-based classification and the analog method. The classification is used to compile a time series of daily circulation patterns to identify days with a high risk of intensive precipitation for the forecast horizon of interest. The role of the search algorithm is to produce a probability forecast of daily areal precipitation for the basin of interest.

In this chapter the current status of the operational system is briefly described. At first, the model history of the PPFS is summarized to illustrate that a forecast system is continuously enhanced so that the structure of those systems must be flexible in respect to future advancements. Afterwards, the GFS and the main components of the PPFS are described in more detail. In the last section, an example of a probabilistic precipitation forecast is given for an event which caused partially severe inundation in southwest Germany in November 2007. The predictions for this event are also used to illustrate the case of a decision maker who, by using the forecast probabilities alone, might have underestimated that risk of this event.

## 8.1 Model History

The first prototype of the PPFS became operationally in February 2005. This system produced a probabilistic short-range forecast of daily precipitation for several precipitation stations around Stuttgart. The predictions were performed by an analog algorithm based on a weighted mixture of two $L_p$-distances (Equation 4.2) to describe the closeness and the similarity between two 1000 hPa height fields. In August 2006, a second predictor, the 700 hPa height field, was integrated into the downscaling process and the predictions were also performed for the catchments of

the study region WG (see Table 3.1 and Figure 3.1).

A major revision of the forecast system started in January 2008 when the output format of the GFS forecast was changed from GRIB1 to GRIB2. The existing conversion programs of the data assimilation system were not able to read the new output format so that these program components had to be replaced by some new tools. This provided the opportunity to upgrade the data archive of the assimilation system to include the relevant forecast information of the PPFS and the GFS in suitable manner for subsequent analysis. Besides the enhancement of the assimilation system, the prediction model was also refined by the fuzzy rule-based classification. The preparation of the classification for the operational use was performed by Christian Ebert. Additionally, the predictions of the analog method were extended for the catchments of the study region MG, SG and EG.

In Summer 2008, the latest version of the PPFS became operationally. Nowadays, the forecast system is running four times a day at 5, 11, 17 and 23 UTC to produce a short-range probabilistic precipitation forecast for all catchments proposed in this study and to compile daily circulation patterns for the Upper Danube basin. Please note that continuous time series of the PPFS predictions are not available due to the number of upgrades made during the development of the forecast system.

## 8.2 Global Forecast System

The Global Forecast System is probably the most established global NWP model. A detailed description of the current status of the system seems to be not available due to the continuous advancements. The version of the forecast system from 2003 is briefly proposed in NCEP (2003) which also gives a detailed overview about many references describing the subsequent development of the forecast system over the last two decades (see e.g. Kanamitsu, 1989). An overview about the latest model changes is given in (NCEP, 2009b).

The current version of the GFS is running every six hour at 0, 6, 12 and 18 UTC to produce forecasts for the next two weeks. Since the GFS predictions are performed by two model versions, the resolution of GFS outcomes decreases over the lead time. The first model component produces predictions for the first week with a temporal resolution of three hours and a spatial resolution of 0.5 x 0.5°, while the second part of the model is running in a slightly sparser resolution to provide 12-hourly forecasts in a spatial resolution of 2.5 x 2.5°.

## 8.3 Probabilistic Precipitation Forecast System

### 8.3.1 Data Assimilation System and Postprocessor

The data flow of the GFS predictions through the assimilation system developed during this investigation is illustrated in Figure 8.1.

The data assimilation begins with the download of the GFS forecasts via anonymous ftp from the NCEP server NCEP (2009c) where the forecast information is stored into GRIB2 files. An overview about the inventory of these files is listed in NCEP (2008). This documentation shows that all forecast variables are stored

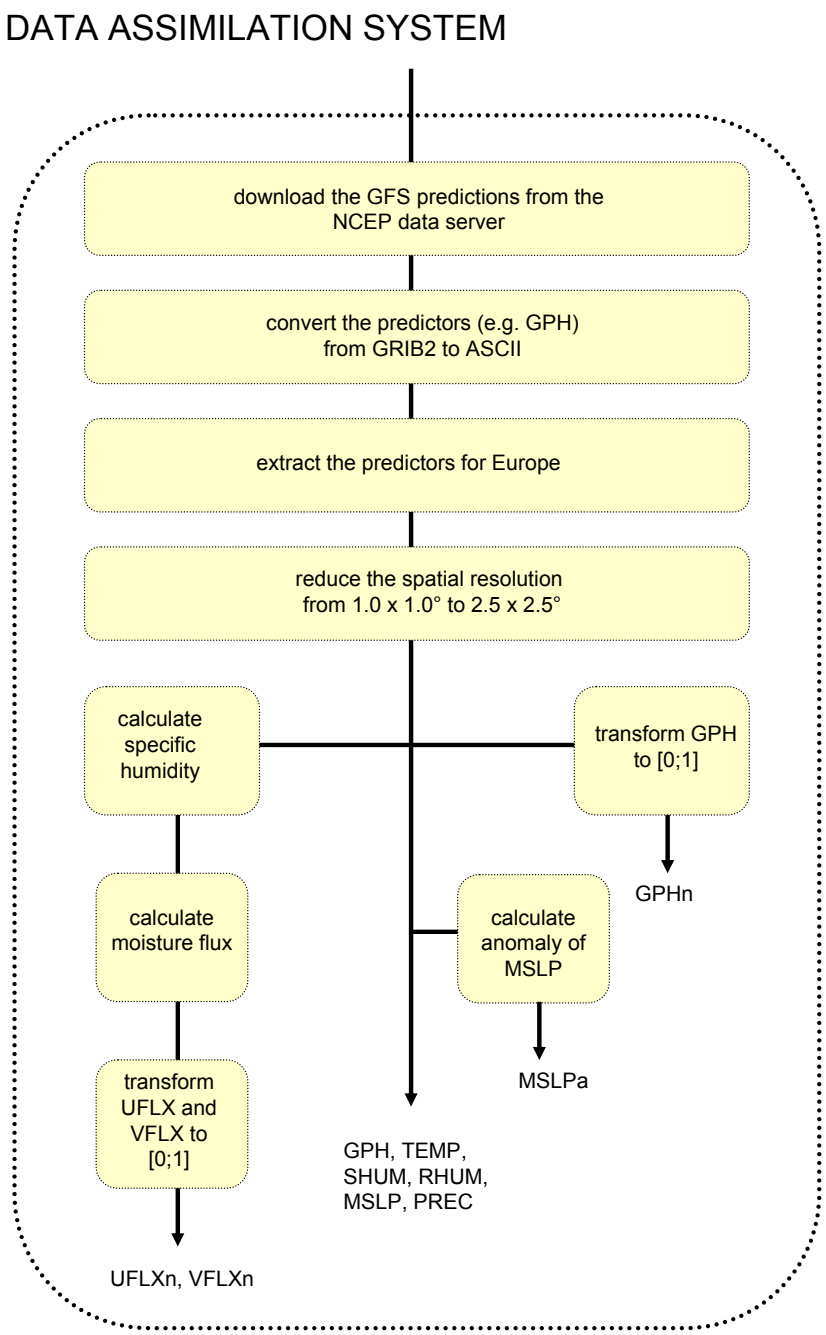



**Figure 8.1:** Data flow of the GFS forecasts through the assimilation system.

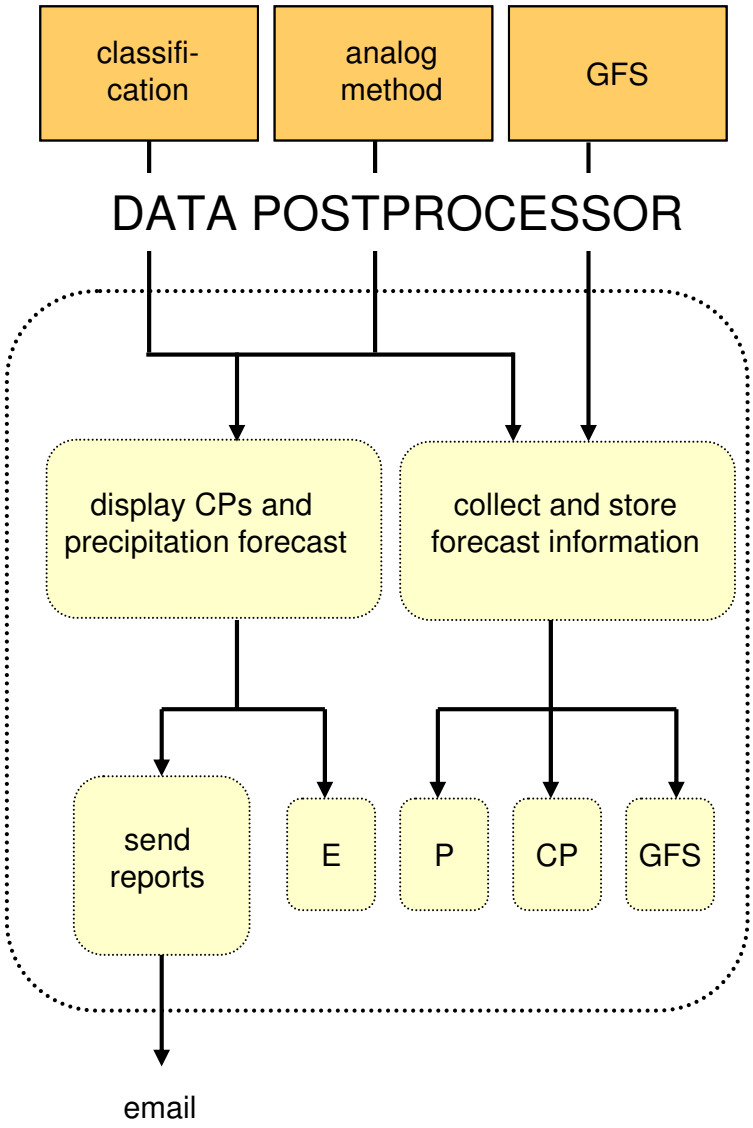



**Figure 8.2:** Data flow of the forecast information through the postprocessor, CP = prediction of the classification, GFS = GFS prediction, E = error reports, P = prediction of analog forecasting.

in a single data file for a particular forecast time step. At the moment, the entire forecast information of the GFS of a given forecast time step is downloaded from the server before the data preparation for the downscaling can begin. The three-hourly predictions of the first week are available in the standard resolution and in a slightly coarser resolution (1 x 1°). The assimilation system downloads the forecast information with the coarser resolution for twelve forecast time steps (6, 12, ..., 72 fh).

After the download of the GFS predictions, the binary forecast information is extracted and converted to an ASCII format. The conversion is performed by a program routine based on the python module pygrib2 which is freely available. A

brief user documentation of the python program and some references about GRIB2 is given by Whitaker (2009). The conversion of the GFS forecasts are performed for five pressure level variables, the geopotential height (GPH), the u-wind and v-wind component (UWND, VWND), the relative humidity (RHUM) and the air temperature (TEMP), taken from four atmospheric pressure levels (1000, 850, 700 and 500 hPa). Apart from the conversion of the pressure level variables, two surface variables, the mean sea level pressure (MSLP) and the precipitation (PREC) are also extracted from the original GRIB2 information. The domain of the window used for the conversion covers the eastern part of the North Atlantic Ocean and Europe. The domain ranges between 40° W to 40° E and 25° N to 75° N.

The last step of the data processing is the preparation of the predictors needed for the downscaling. At first, the resolution of the large-scale information is reduced from 1 x 1° to 2.5 x 2.5° to make the forecasts consistent with the reanalysis information. Please note that in contrast to the NCEP/NCAR reanalysis data, the GFS predictions contain no information about the specific humidity. To overcome this problem, the relative humidity and the air temperature are used for the calculation of this variable. The meridional and the zonal moisture flux (VFLX, UFLX) are determined by multiplying the specific humidity with the corresponding wind component. Finally, the anomalies of the mean sea level pressure (MSLPa) are calculated (Equation 3.2) and the geopotential height and the moisture fluxes are standardized (GPHn, UFLXn, VFLXn) to an interval between zero and one (Equation 3.1).

The postprocessor collects and saves the relevant forecast information of the PPFS and GFS into an data archive (see Figure 8.2). Due to the large data amount which is downloaded four times a day (180 MB) from the NCEP server, only the large-scale information needed for the downscaling is collected. This amount of data is around 90 % smaller compared to the original amount. Apart from the forecast archive, the postprocessor consists of an observation archive with the NCEP/NCAR reanalysis and the precipitation information. A further task of the prostprocessor is to visualize the forecasts and to send this information to end users by email.

### 8.3.2 Statistical Precipitation Prediction Model

The statistical techniques for the precipitation prediction model are the analog method (Chapter 4) and fuzzy rule-based classification (Chapter 5). The predictions of the search algorithm are performed for all catchments selected in this investigation and for several precipitation stations around Stuttgart. The predictors are the height field and the zonal moisture flux field based on the model settings presented in Chapter 4 in Tab. 4.1 and 4.3. The fuzzy rule-based classification compiles a time series of daily circulation patterns for the Upper Danube basin in Germany. Both downscaling techniques are used to perform a forecast for three days.

The forecast cycle of the PPFS is illustrated for the resampling algorithm in Figure 8.3. To perform the predictions, the forecast information of the latest GFS model run is selected by the PPFS. For example, the model run at 5 UTC uses the forecast information of the GFS model run initialized at 0 UTC. It is important to mention that the GFS needs at least three hours to produce the first forecast information and to store it on the data server. This means that the entire information needed for the downscaling by the PPFS is soonest available four hours after the

intializing model run e.g. at around 4 UTC for the model run at 0 UTC. Thus, there is still a time delay of one hour before the assimilation system of the PPFS begins to download the GFS predictions.

The PPFS model run at 5 UTC produces probability forecasts for the time interval from 6 to 6 UTC of the next day for all forecast days. For the first forecast day, the predictors are the 700 hPa zonal moisture flux field at 18 fh and the 1000 hPa height field at 24 fh. They are compared to the corresponding reanalysis information at 18 and 24 UTC. The time information of the GFS predictions is given in forecast time measured in hours (fh) since the intialization. To predict the daily areal precipitation for the second and the third forecast day, the same forecast information is taken except that it has a temporal offset of 24 and 48 fh, respectively.

The further model runs of the PPFS at 11, 17 and 23 UTC use the GFS predictions

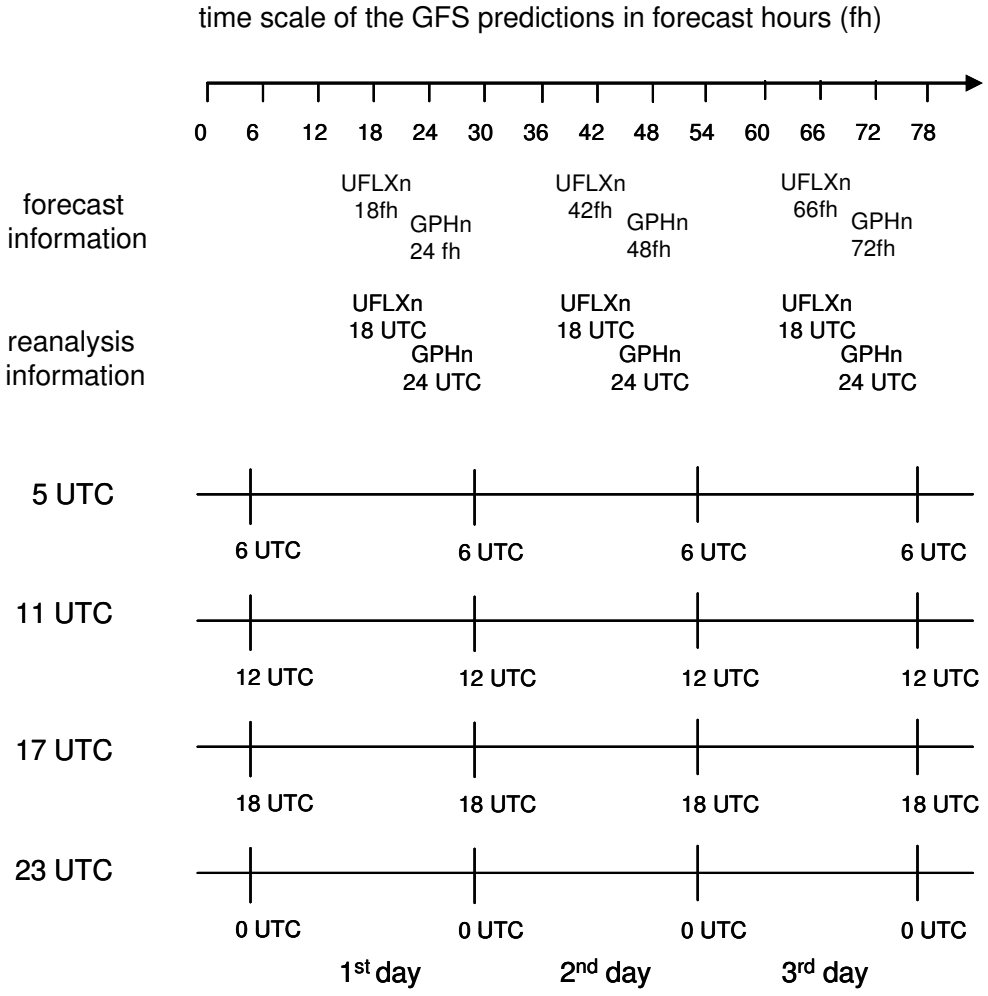



**Figure 8.3:** Predictor information needed to perform a three day forecast by the resampling algorithm for the model runs of the PPFS at 5, 11, 17 and 23 UTC.

Probabilistic quantitative precipitation forecast for the Upper Sieg

| | 0 - 24 fh | 24 - 48 fh | 48 - 72 fh |
|---|---|---|---|
| Precipitation amount [mm/d]: | | | |
| 10%-quantil: | 1.6 | 0.0 | 0.0 |
| median: | 5.0 | 0.2 | 0.0 |
| 90%-quantil: | 19.3 | 4.2 | 2.8 |
| Forecast probabilities [%]: | | | |
| <2 mm/d: | 18 | 76 | 88 |
| 2-10 mm/d: | 54 | 24 | 12 |
| 10-20 mm/d: | 18 | 0 | 0 |
| >20 mm/d: | 10 | 0 | 0 |

**Figure 8.4:** An example of a probabilistic precipitation forecast sent to some IWS members on each day by email, 6th - 8th October 2006. The time interval of the daily precipition is here from 12 to 12 UTC.

from the corresponding second, third or fourth model run initialized at 6, 12 or 18 UTC. To perform a three day forecast, the same predictor information is used for these model runs. Only the time interval of the predicted daily areal precipitation is changed for each model run. The second model run at 11 UTC uses a time interval from 12 to 12 UTC while predictions of the third and the fourth model runs are based on a time interval from 18 to 18 UTC and 0 to 0 UTC, respectively.

An entire run of the forecast system takes around 25 minutes with a standard personal computer. The predictions are therefore available at 5:25, 11:25, 17:25 and 23:25 UTC. The total time needed for the download of the GFS predictions and the predictor preparation is at least 20 minutes, while the prediction by the analog method and the classification only take a few minutes.

Finally, the forecast information of the prediction model is visualized and can be sent to an end user via email. A sketch of the probabilistic precipitation forecast of the analog method sent daily to several members of the IWS is presented in Figure 8.4.

## 8.4 Probabilistic Quantitative Precipitation Forecast

The PPFS predictions have been performed on around 60 % of the days since the first model version became operational in February 2005. One of the largest precipitation events that occurred during this test period was the event caused by the Kyrill storm on 18th January 2007. The uncertainty of the precipitation forecast for this event is indicated by the 10 % and 90 %-quantile in Figure 8.5 for the Blies catchment. The forecasts are compared to the mean observed daily areal precipitation of this catchment. Here, the forecast uncertainty of the analog method is given for the

third forecast day (lead time from 48 to 72 fh).

The figure illustrates that the search algorithm can suitably distinguish between dry and wet days even for the third forecast day since the predicted precipitation probabilitiy is clearly smaller on days with dry or light precipitation than the precipiation probabilitiy on days with moderate or intensive precipitation. For example, for days with no rain or light precipitation, around 80 % of the precipitation realizations are below 2 mm/d. While for the extreme, more than 90 % of the forecast realizations exceed this threshold. Furthermore, the visualization of the probability forecast illustrates that the confidence interval of the forecast can enclose all observations. Thus, the PPFS seems to provide a suitable estimate of the predictive uncertainty of daily areal precipitation for the selected forecast period.

The forecast example also highlights that there is a tremendous degree of uncertainty regarding the forecast of extreme precipitation. Approximately 80 % of the precipitation realizations are between 5 and 35 mm/d. In this case, it is obvious that many decision makers argue that a fast and valuable decision cannot be performed due to the high forecast uncertainty. To overcome this limitation, the cost-loss approach proposed in Section 2.2.5 must be selected to point out a suitable quantile and to optimize the decision making process for those situations. This example also illustrates that the ensemble mean is not a suitable indicator to estimate the precipitation amount for the storm event and the selection of a higher quantile provides more suitable warning information for this event.

The PPFS has been also designed to be an element of a warning system. This system should give a suitable forecast information for various warning levels so that the decision maker can use this information to trigger an alarm. This means also that the forecast probabilities must be predicted for a given warning level. Unfortunately, the probabilities of these thresholds are small, in particular in this case that the predictions are performed for levels of the highest warning category. In this case, it is possible that the risk is underestimated by a decision maker (for further discussion see Murphy, 1991). This problem will be illustrated in more detail in the following example.

The forecast probabilities for a precipitation event exceeding a warning level of 20 mm/d are illustrated for the storm event in Figure 8.5b. The forecast probabilities are small for the day even when intensive precipitation occurred. Only 25 % of the precipitation realizations are above this warning level. However, the forecast probabilities determined for this warning level correspond to an event with a high return period (around six times a year) but the forecast probabilities for higher warning levels are even smaller. Thus, the forecast probabilities might not be high enough for a decision maker with less experience to recognize that an extreme event is indicated by the forecast system. To overcome this problem, the decision maker must compare the forecast probabilities to the climatological occurrence frequency of the event to which the warning level corresponds to. This comparison can be performed by the risk index described in Equation 7.2. For the storm event, the risk index for the given warning level is around 15 times larger in comparison to the mean climatology (Figure 8.5c). This information should alert a decision maker to observe this event in more detail in order to decide finally between protective action or no action.

This example demonstrates that the risk index can simplify the interpretation of

the forecast probabilities for the prediction of rare events. Besides the formulation of forecast probabilities, the risk index should also play a major role in the interpretation of the output information of warning system.

## 8.5 Summary

In this chapter a prototype of a probabilistic precipitation forecast system (PPFS) was proposed which was developed during this investigation. This forecast system uses the large-scale information of the GFS to perform a probabilistic forecast of daily areal precipitation for the river basin selected in this investigation for up to three days. The PPFS also produces a times series of daily circulation patterns for the identification of days with a high risk of intensive precipitation for the same forecast horizon.

The integral parts of the forecast system are the data assimilation system, the postprocessor and the statistical precipitation prediction model. The data assimilation system collects and prepares the large-scale information of the GFS needed for the downscaling. Afterwards, the large-scale information is transferred by the statistical precipitation prediction model into forecasts for the local scale. The key downscaling techniques of the statistical prediction model are the analog method and the fuzzy rule-based classification. After compiling the predictions, the postprocessor visualize the forecast information which can be sent *via* email to a potential user.

The PPFS was tested with forecast information over the last few years. In this investigation the precipitation forecast of the operational system was illustrated for one extreme event, the storm event Kyrill which caused partially severe inundation in southwest Germany in January 2007. The visualization of the precipitation forecast for the third day demonstrated that the forecast system can provide suitable information about the forecast uncertainty of daily areal precipitation. The selection of the risk index illustrated that days with a high risk of intensive precipitation can be indicated in a suitable way so that the prediction model is in a position to provide adequate warning information for rare precipitation events. Since the forecast probabilities for those events are small, the calculation of the risk index is a suitable technique to illustrate the risk for the prediction of rare event. This technique is in particular useful if the decision maker has less experience with probabilistic precipitation forecasts and should play a crucial role alongside the formulation of forecast probabilites.

In the near future, the forecast horizon of the PPFS will be extended so that the uncertainty of the precipitation forecast can be described for up to seven days. Additionally, a simulation model will be installed to disaggregate the daily information into precipitation fields with a high spatiotemporal resolution. Afterwards, the precipitation information will be equivalent to the resolution of the distributed hydrological model which is running at the flood forecasting center of the LUWG. Finally, the PPFS will be implemented and tested by the members of the forecasting center for real world situations.

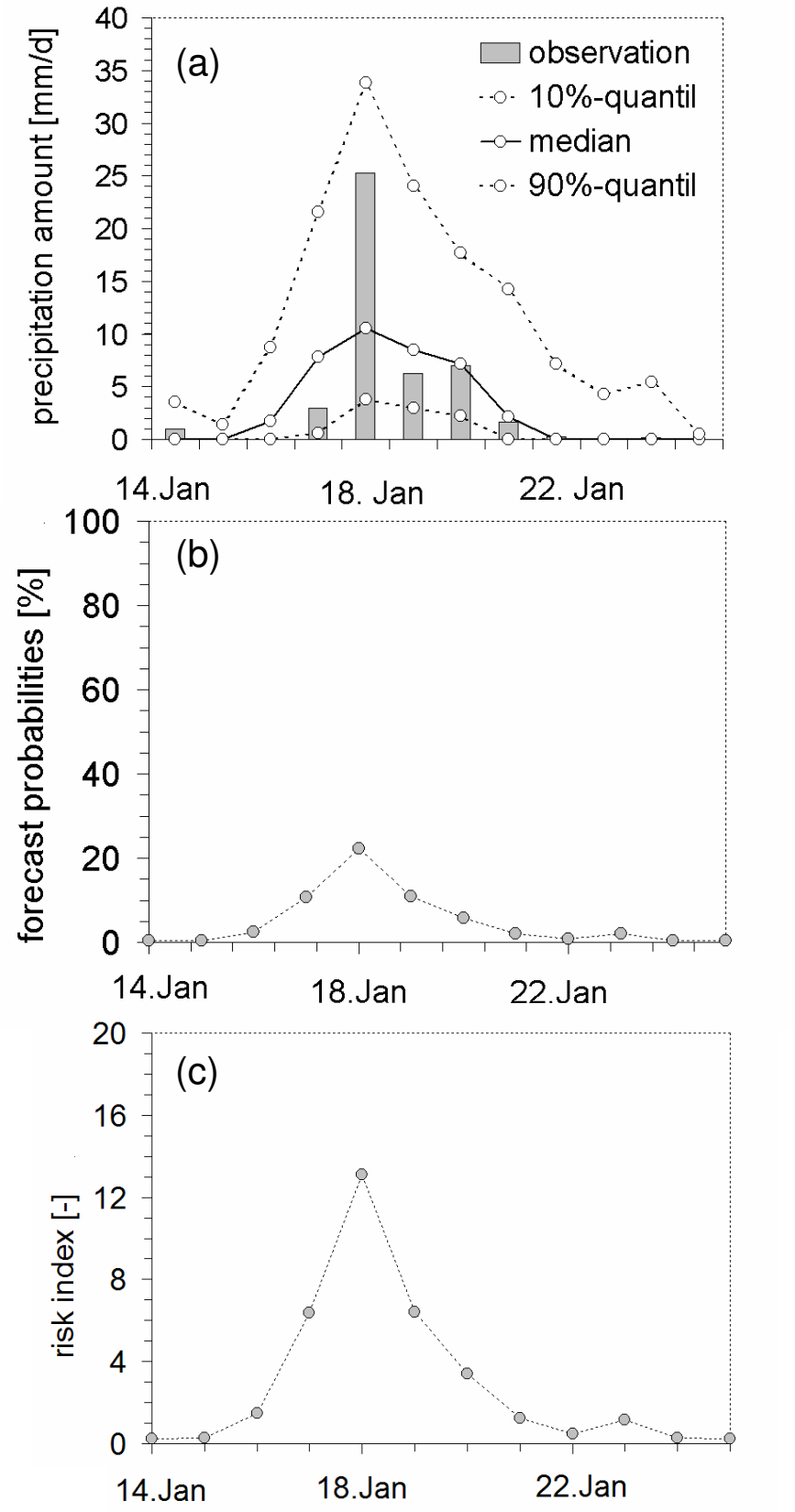



**Figure 8.5:** Probabilistic quantitative forecast of daily areal precipitation for the Blies by the analog method. Figure a: Forecast uncertainty. Figure b: Forecast probabilities for an event exeeding a warning level of 20 mm/d. Figure c: Risk index for the same event. Lead time from 48 to 72 fh. Kyrill storm event in January 2007.

# 9 Summary and Conclusion

An adequate estimation of the predictive uncertainty for precipitation is the basic prerequisite for the development of an accurate and valuable flood forecasting system for small river basins in temperate zone. The primary goal of this investigation was the development of a prediction model that can be used to produce a probabilistic forecast of daily areal precipitation to estimate the uncertainty of the precipitation forecast for small river basins in Germany.

The integral parts of this model are the techniques of statistical downscaling. These techniques establish a transfer function between the outcomes of a global NWP model and the surface variable of interest to translate the large-scale information to the local scale. The key downscaling technique of this investigation was the analog method. It is a straightforward and valuable approach for the estimation of the predictive uncertainty for local surface variables. The main component of the analog method is an automatic search algorithm to identify past weather states that are similar (analog) to the forecast one.

The study was performed in several medium-sized catchments located in the Elbe, Danube and Rhine basins in Germany. These catchments are typical mountainous and alpine subbasins of the major rivers in Germany with large height differences and short concentration time ($< 24$ h) during a flood event. To define suitable model parameters for the analog method, the catchments were divided into four study areas according to their catchment characteristics. The parameter estimation was performed with reanalysis information so that the goodness of a forecasting technique can be evaluated over a long investigation period (more than 40 years) with simulation output from a general circulation model. Since the precipitation model is designed to be an element of a flood warning system, the performance of the downscaling process was determined for warning situations. The comparison with low-skill reference forecasts demonstrated that the analog method can provide suitable warning information for the detection of intensive precipitation.

It was also demonstrated that the skill of the analog method is affected by several criteria which must be defined by an investigator before starting the search algorithm. Since a guideline for an appropriate selection of these criteria is not available and only a few studies have selected so far the analog method for downscaling daily precipitation in Germany, comprehensive model sensitivity was performed to point out the most sensitive criteria. The model parameters were determined for two suitable predictor sets which were chosen to estimate the uncertainty of the precipitation forecast for short-term and medium-term projections. Finally, the model settings were listed for each study region in detail. Future investigations can use this information to advance the downscaling process for these regions in Germany.

To demonstrate the capability of the statistical downscaling technique for real situations, a prototype of a probabilistic precipitation forecast system was developed. This forecast system uses the outcomes of the global NWP model of the Amer-

ican weather service to provide a short-term probabilistic forecast of daily areal precipitation for the catchments selected in this investigation. In the near future, the spatiotemporal resolution of the forecast system will be extended to meet the requirements of the distributed hydrological prediction model of a regional flood forecasting centre. Finally, the precipitation model will be implemented at the flood forecasting centre to test the model for decision making in real world situations.

## 9.1 Some Answers to the Research Questions

Apart from the design and the development of the precipitation model, four research questions were proposed and addressed in this study. Since we had little experience with the precipitation forecast of dynamical and statistical downscaling, the following question arose at the beginning of this investigation:

***How accurate is the forecast of daily areal precipitation for small river basins in Germany, in particular for flood producing situations?***

To answer this question, an example of a deterministic precipitation forecast from a limited area model was presented. This is the most commonly used at the regional flood forecasting centers in Germany. The comparison with observations illustrated the low accuracy of the deterministic forecast so that in many cases either a false alarm is given or the event is missed. This result coincides with the experiences of many decision makers. They often have mentioned that a deterministic precipitation forecast is not accurate enough to develop a reliable flood prediction model.

We also analyzed the precipitation forecast of a mesoscale ensemble prediction model which transfers a selected number of global ensembles through a limited area model. This model provides a high-resolved ensemble forecast for Central Europe which can be used to estimate the forecast uncertainty of local surface variables for up to five days. It was illustrated that the confidence interval of the forecast can enclose many observations so that suitable information about the forecast uncertainty can be provided by the ensemble prediction model. The analysis of the ensemble forecasts also demonstrated that the precipitation uncertainty is tremendous even for short-term forecasts and although only a small ensemble size is used to estimate the uncertainty. We would like to stress that the selection of less realizations is not a reasonable strategy to reduce the forecast uncertainty. Otherwise, the model performance of the ensemble prediction system clearly decreases, in particular for longer lead times, so that suitable warning information can be not provided for the end users.

To overcome the problems of dynamical downscaling, we proposed in this investigation statistical downscaling techniques to estimate in an adequate manner the forecast uncertainty for the surface variable of interest. Moreover, many reasons for the use of statistical forecasting were given and suitable arguments were presented to formulate and to express the outcomes of prediction models in forecast probabilities. Unfortunately, decision making in the face of uncertainty is more complex compared to deterministic forecasts. Therefore, the following questions were also

discussed in more detail:

***What is the optimal use of a probability forecast in decision making situations? And, how valuable are probability forecasts in comparison to other forecasting techniques?***

To answer this question, we selected a cost-loss approach to approximate the decision making process of warning situations. This approach relates the negative outcomes of a binary warning system (hit, false alarm, miss and inverse hit) to the consequences of the decision making process (cost for protective action and loss if the event is missed). We demonstrated that a probabilistic forecast is more valuable compared to other forecasting techniques since the decision maker can select the optimal decision threshold for an individual user. However, it is crucial for the goodness of the model predictions that the decision maker knows how to apply a probability forecast in cost-loss situations. Therefore, the optimal use of a probability forecasts was demonstrated so that the decision maker can use this information to maximize the benefit for an individual user. Additionally, it was shown that the optimal use of the probability forecast is influenced by several factors e.g. the cost-loss ratio and the occurrence frequency. These factors should be taken into account by a forecaster for the optimization of the decision making process to provide finally a highly valuable forecast for all users for their region of interest.

The approach was also chosen to estimate the forecast value of the analog method in warning situations. In comparison to the low-skill reference forecasts, the prediction of the analog method is characterized by a high forecast value for the detection of rare events so that many users can profit from the model predictions. Nonetheless, we have the opportunity to make any comparison with other statistical forecasting techniques. Thus, it was also necessary to clarify the following question:

***How accurate are the predictions of the analog method in relation to other statistical forecasting techniques?***

To shed light on this question, an objective classification based on the concept of fuzzy rules was selected to compile a catalog of daily large-scale circulation patterns for the river basin of interest. We illustrated that this technique can determine physically reasonable patterns with a high risk of intensive precipitation. The evaluation of the predictions by the cost-loss approach demonstrated that the classification can provide valuable information for the detection of intensive precipitation. If the decision maker selects an appropriate decision threshold, the hit rate of the classification is similar to analog forecasting. Unfortunately, many alarms are triggered by the classification so that the warning information provided by this technique seems to be less valuable in comparison to the analog method. This result would be expected since the search algorithm uses more information for the downscaling and it focuses more on the description of the weather state over the region of interest in comparison to the classification.

To give a more precise answer to the aforementioned question, future investigations should also select forecast information for the model evaluation. In this investigation

the model performance of each downscaling approaches was mostly determined with reanalysis data. Thus, under assumptions that the GCM predictions are perfect. However, it is obvious that this assumption is violated when forecast data from a global NWP model is selected.

Please note that the classification was not only selected for the model comparison. The large-scale daily circulation patterns can be also used to introduce additional information in the downscaling process. Besides this argument, there are many further reasons for the use of a classification technique in objective weather forecasting. The compilation of large-scale circulation patterns should actually form an integral part when a prediction model for local surface variable is developed. Future investigations can highlight this aspect in more detail to join the strengths of both downscaling approaches and to reduce finally the limitations of a prediction model for local surface variables.

We also presented in this investigation a new technique for the downscaling of surface variables. This technique is based on the data depth which measures the centrality of a weather state. Since local extremes are often caused by anomalies of the atmosphere, this idea seems to be a promising technique for the identification of weather states with a high likelihood of intensive precipitation at the local scale. However, the selection of this approach for precipitation downscaling is not straightforward since unusual weather situations can be linked to adverse weather and good weather over a region of interest. To overcome this limitation, the large-scale circulation patterns compiled by the fuzzy rule-based classification were taken to distinguish between dry and wet weather states. The model performance of this conditional downscaling approach indicated a high potential for the detection of intensive precipitation similar to the performance of analog forecasting. This result was quite encouraging and therefore an advancement of this technique should be continued.

We can summarize that the analog method is a suitable technique for the estimation of the forecast uncertainty of daily precipitation and that this approach was able to provide valuable information for the detection of intensive precipitation for small river basins in Germany. Thus, the enhancement of the methodology should be worthwhile. Even so, the next pertinent question that arises is ...

***How can the traditional way of the analog method be improved?***

To address this issue, four different ways were proposed and investigated in detail:

1. **Pattern closeness and pattern similarity**: The model performance of analog forecasting strongly depends on the measure which is taken for the comparison of two weather patterns to identify suitable analogs. Usually, a distance function is selected which describes the closeness between two weather patterns. However, we highlighted that two patterns can be similar according to their form, or so called pattern similarity. Unfortunately, pattern closeness and pattern similarity are not taken into account by the traditional distance functions selected in analog forecasting. To create a measure balancing between similarity and closeness, a weighted mixture of two $L_p$-distances was proposed. We demonstrated that the selection of this measure can slightly improve the

prediction of intensive precipitation if both ordinary distance functions posses a similar downscaling performance. Future work should also be able to point out whether the new distance measure can enhance the predictions over the entire range of the predictand or not.

2. **Selection of an appropriate distance function**: Usually, in analog forecasting a measure for the comparison of two patterns is selected where all grid points of the predictor domain have the same weight for the comparison. However, such a distribution of the weight is actual a poor choice and the selection of more sophisticated distribution is costly. To overcome this problem, we presented in this investigation an optimization algorithm. It was highlighted that the algorithm can define weights for a distance function so that the time-consuming manual specification of the predictor domain and the selection of a measure for the comparison can be skipped. Furthermore, we illustrated that the model prediction can be improved by this technique so that an investigator with less experience can even obtain similar results compared to an expert. However, the goodness of the model transferability is strongly influenced by the choice of the objective function. An investigator must select a suitable objective function, otherwise the optimization algorithm determines a less reliable predictor domain and the model transferability of the approach is poor.

3. **Model transferability**: To improve the transferability of analog forecasting, several objective functions were tested which focused either on the prediction of daily precipitation or the prediction of daily discharge differences. The selection of daily discharge differences is an uncommon optimization strategy, though it has the advantage that robust identification of model parameters is more likely. Since river flooding is mainly the result of intensive precipitation in the temperate zone, a further benefit of this strategy is that a prediction model can be developed for two purposes, namely the prediction of intensive precipitation and river floods. The model optimization with discharge differences was also tested to identify suitable parameters for the fuzzy rule-based classification. The calculation of the model performance demonstrated that the selection of discharge differences can enhance the model transferability for both, analog forecasting and classification. This optimization strategy can even improve the model predictions for intensive precipitation in comparison to the optimization with daily precipitation. However, there are still further options available for a more suitable formulation of an objective function based on daily areal precipitation which could outperform the optimization strategy presented in this investigation.

4. **Selection process**: The predictions of the analog method are also influenced by the choice of the selection rule used to restrict the search process. Usually, a selection rule is used to take into account the interannual variability of the downscaling process. However, the investigation highlighted that the selection of circulation patterns can be also useful in restricting the search process so that e.g. for a summer day with a wet circulation pattern only analogs are selected that occurred in the same (similar) season and that belong to the same

(similar) pattern. The calculation of the model performance demonstrated that the selection of the new search algorithm was only able to improve the prediction for moderate and intensive precipitation in comparison to standard algorithms. The use of circulation patterns has also the advantage that the search process can be accelerated to allow for quicker estimation of the model parameters. Future investigation can focus on the development of more sophisticated rules compared to those rules selected in this study so that finally the prediction of the algorithm can be improved over the entire range of the predictand.

In this section we presented some answers for the research questions proposed in this investigation. However, there are still many open questions which need to be addressed in future investigations. Some of these questions are presented in the following section.

## 9.2 Outlook for Future Investigations

***Are there any ways to reduce the forecast uncertainty of the analog method for daily areal precipitation?***

This question often came up when the uncertainty of the precipitation forecast was presented for specific events. Then, decision makers and many researchers as well were often surprised about the high uncertainty of the precipitation forecast.

There are still some ways and techniques to improve the methodology. One possible way is the identification of systematic differences (bias) between the predictor information used for the model development and the predictor information used for operational real-time forecast. In this investigation the precipitation model was developed with reanalysis information. But in the operational mode the predictor information is provided by a GCM with a different model structure (e.g. resolution and parameterization) and probably with a different data source. In particular for variables with a high spatial variability like the moisture flux, it can be expected that there is a bias in the prediction information which influences the performance of the precipitation model. To investigative this question, the predictor information of the reanalysis archives provided by NCEP/NCAR and ECMWF can be used for the comparison.

Another way could be the advancement of the predictor sets used to describe the downscaling process. For example, additional information can be introduced in the downscaling process to describe more accurately the preconditions of an event. It is obvious that the origin and the tracks of cyclones should play a role for long-lasting large-area heavy precipitation. A further way of introducing additional information is the selection of predictors which describe the long-term variability of the weather over the region of interest. We know that the evolution of the weather in Europe is influenced by the state of the ocean circulation of the North Atlantic. Wilby (1998) illustrated that anomalies of the mean sea temperature of the North Atlantic have a small effect on the long-term variability of the precipitation process in Europe and can slightly improve the downscaling.

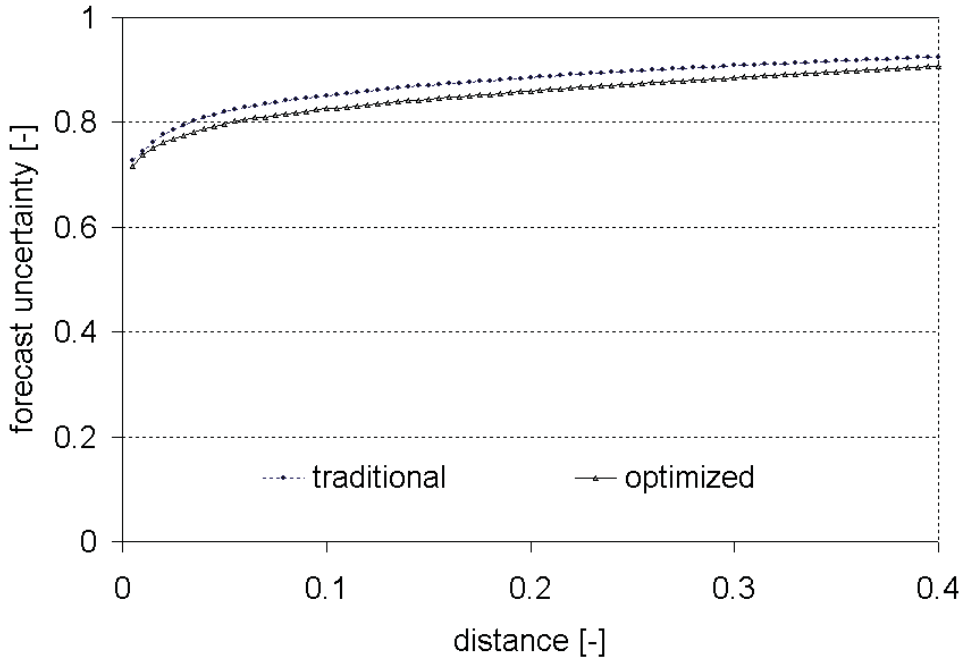



**Figure 9.1:** Forecast uncertainty of daily areal precipitation with respect to the closeness of two resampling algorithms. Traditional = traditional approach of analog forecasting (Chapter 4). Optimized = analog forecasting based on an appropriate distance function (Chapter 6).

Apart from the introduction of additional information in the downscaling process, the enhancement of the search algorithm should also play an important role in future studies. Usually, we assume in analog forecasting that the closer two weather situations are, the more similar is their outcome and the smaller should be the forecast uncertainty. However, in the case of forecasting daily areal precipitation, this assumption seems to be not satisfactorily fulfilled. Figure 9.1 illustrates the forecast uncertainty of daily areal precipitation in dependence on the closeness of two weather patterns for two search algorithms proposed in this investigation. Even for very close patterns, the forecast uncertainty is large for both algorithms. Only 30 % of the variability of the predictand can be explained even for the algorithm with an optimized distance function (see Chapter 6). This internal validation illustrates that there still might be potential to improve the predictions of the analog method. A methodology to tackle the aforementioned problem was proposed in Bárdossy et al. (2005). They tested this idea for a resampling algorithm used for daily flood forecasting. While the same idea was selected by Samaniego et al. (2008) for the classification of land use data, that was derived from remote sensing. In this study we also developed a methodology based on the approach presented in Bárdossy et al. (2005). Unfortunately, the transferability of this approach to an independent data set was poor so that an alternative way of distance optimization was proposed. In the near future, we would like to continue this work to find a more satisfying solution

for the aforementioned problem.

We are sure that all ways (i) the identification and the correction of the bias in the predictor information, (ii) the introduction of additional information in the downscaling process and (iii) the advancement of the search algorithm can reduce the forecast uncertainty. However, it should be remembered that the model predictions presented in this investigation were performed with reanalysis information (one exception was made in Chapter 8 where the model predictions was based on forecast information). Thus, the precipitation model was tested under the assumption that the large-scale predictions of a global NWP model are perfect. Unfortunately, this assumption is not true for future projections. Thus, there will be still the dilemma that the forecast uncertainty of daily precipitation is high, in particular for longer lead times and for the prediction of rare events. To describe the forecast uncertainty for daily areal precipitation, the statistical forecasting techniques proposed here seem to be appropriate downscaling techniques. Moreover, the selection of cost-loss approaches presented in this investigation enables the investigator to identify suitable strategies for decision making even under high uncertainty of the precipitation forecast.

***What about extrapolation?***

Analog forecasting is restricted to observations. Since the archive of daily precipitation observations are seldom longer than fifty years, the estimated precipitation amount is poor for events outside the observation range. However, Bontron and Obled (2003) mentioned that the methodology provides a suitable warning information already for decision making due to the selection of past weather situations. For instance, if the resampling algorithm identifies a weather state with the largest ever recorded precipitation amount in a river basin, this information should be enough for a decision maker to trigger an alarm. Nevertheless, we must also develop techniques which can adequate forecast the precipitation amount of rare events. This means that the extrapolation capabilities of global NWP model must be analyzed for rare weather phenomena. If they can reliably simulate the atmospheric flow processes for these situations, then this information can be used for the development of downscaling approaches with a suitable extrapolation technique. In the case of analog forecasting, the selection of a local regression might be a possible strategy to provide an appropriate estimate of the precipitation amount for rare events. Please note that even in the case when a suitable extrapolation technique can be developed, there will be still the problem that only a small number of observed extremes are available for the model validation. Then we can only assume that the transfer function developed with past observations hopefully holds for future extremes which are outside the observation range and which occur in a changing climate.

***How can we develop a reliable warning system for the public?***

This question usually came up when the warning strategy for the detection of intensive precipitation was presented. To obtain a suitable hit rate for rare events, the warning system must trigger a number of alarms, since the prediction of intensive precipitation is associated with a high forecast uncertainty. Unfortunately, this pessimistic warning strategy is not useful for the members of the general public. It is

obvious that their willingness for protection will be significantly reduced if this kind of warning strategy is used. To overcome this dilemma, the following points should be taken into account:

- Firstly, the decision makers of the flood forecasting centers must be trained so that they can optimally use a probability forecast to provide valuable warning information for all users of their region of interest. Many decision makers are aware that the outcomes of predictions models are uncertain. This step could be realized with less effort.

- Secondly, the users must be distinguished into classes according to their willingness for protection. Then, a warning system must be developed which mobilizes the user groups step-by-step during an alarm situation. For example, if there is any indication of an extreme, the members of a flood forecasting centers are alarmed to observe the weather situation in more detail. If they obtain new information and they are more certain that the event will occur, they can inform users with a higher willingness for protection. Only in the last step of the warning chain, if a suitable warning can be provided to users with a low willingness for protection, the members of the general public should be informed so that they can take protective action.

- Thirdly, the uncertainty of the weather forecast must be communicated to **all** members of a forecasting chain. It is obvious that many people know from their experience that weather forecasts are uncertain. Thus, this issue is not new for them and the introduction of uncertainties in the form of forecast probabilities might be less difficult. However, the use of forecast probabilities also means that all members of a warning chain have to be trained since misinterpretations of probability forecast are quite common (see Murphy et al., 1980; Gigerenzer et al., 2005; Morss et al., 2008). Probably, a further problem is that many people do not understand why the predictions are uncertain. Most of the people know that weather forecasts are performed with computers. Unfortunately, there seems to be the belief that due to the increase of computer power over the last decades, the performance of a prediction model must rise in a similar way. However, it is obvious that the model performance of a prediction model is not only a condition of computer power. Thus, there seems to be the need to communicate in a suitable way the main reasons for the forecast uncertainty. If we do not engage in any communication, we pretend that a high model precision is present and users might overestimate the capabilities of the prediction model.

  We are convinced that both, an adequate communication of the forecast uncertainty and their reasons behind it, will increase the motivation for protection so that finally a more valuable warning system can be developed for all members of a forecasting chain.

Finally, we can summarize that there are still many open questions and issues which should be addressed in future studies. The ideas and methodologies presented in this investigation are only a small portion of the current state of those that represent the art in the field of probabilistic precipitation forecast. The author hopes that the

investigation was able to highlight the basic principles and techniques in a suitable way so that future research activities can benefit from this investigation and that some ideas or techniques can also find their way into the hydrological practice.

# Bibliography


Aarts, E. and Korst, J.: Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing, 1989.

Ahmed, S. and De Marsily, G.: Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, Water Resources Research, 23, 1987.

Atger, F.: Verification of intense precipitation forecasts from single models and ensemble prediction systems, Nonlinear Processes in Geophysics, 8, 401–418, 2001.

Bárdossy, A.: Downscaling from GCMs to local climate through stochastic linkages, Journal of Environmental Management, 49, 7–17, 1997.

Bárdossy, A.: Stochastic downscaling methods to assess the hydrological impacts of climate change on river basin hydrology - Keynote Paper 1. In: Beersma J., Agnew M.D., Viner D., and Hulme M.: Climate Scenarios for Water-Related and Coastal Impacts. Workshop Report. Proceedings of the EU Concerted Action Initiative ECLAT-2 Workshop 3. KNMI, Netherlands, May 10-12th 2000, Climatic Research Unit, Nor-wich 140pp., 2000.

Bárdossy, A. and Filiz, F.: Identification of flood producing atmospheric circulation patterns, Journal of Hydrology, 313, 48–57, 2005.

Bárdossy, A. and Plate, E.: Space-time model for daily rainfall using atmospheric circulation patterns, Water Resour. Res, 28, 1247–1259, 1992.

Bárdossy, A. and Singh, S.: Robust estimation of hydrological model parameters, Hydrology and Earth System Sciences, 12, 1273–1283, 2008.

Bárdossy, A., L., D., and Borgardi, I.: Fuzzy rule-based classification of atmospheric circulation patterns, International journal of climatology, 15, 1087–1097, 1995.

Bárdossy, A., Stehlík, J., and Caspary, H.: Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules, Climate Research, 23, 11–22, 2002.

Bárdossy, A., Pegram, G., and Samaniego, L.: Modeling data relationships with a local variance reducing technique: Applications in hydrology, Water Resources Research, 41, W08 404, 2005.

Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses: detection of suspicious data in ERA-40, Atmospheric Science Letters, 10, 2009.

Bliefernicht, J. and Bárdossy, A.: Probabilistic forecast of daily areal precipitation focusing on extreme events, Natural Hazards and Earth System Sciences, 7, 263–269, 2007.

Bliefernicht, J., Bardossy, A., and Ebert, C.: Stochastische Simulation stündlicher Niederschlagsfelder für Extremereignisse an der Freiberger Mulde, dem Oberen Main und der Fränkischen Saale, Hydrologie und Wasserbewirtschaftung/Hydrology and Water Resources Management-Germany, 52, 2008.

Bontron, G.: Prévision quantitative des précipitations: adaptation probabiliste par recherche d'analogues. Utilisation des ré-analyses NCEP/NCAR et application aux précipitations du Sud-Est de la France, thèse de doctorat, INPG, Grenoble, 276 p, 2004.

Bontron, G. and Obled, C.: New developments in quantitative precipitation forecasts by analog sorting techniques, in: Mediterranean Storms. Proc. Of the 5th EGS Plinius Conference held at Ajaccio, Corsica, France, October, 2003.

Brandsma, T. and Buishand, T.: Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling, Hydrology and Earth System Sciences, 2, 195–210, 1998.

Brazdil, R., Kotyza, O., and Dobrovolny, P.: July 1432 and August 2002-two millennial floods in Bohemia?/Juillet 1432 et Aout 2002-deux crues millenaires en Boheme?, Hydrological Sciences Journal/Journal des Sciences Hydrologiques, 51, 848–863, 2006.

Bremicker, M.: Das Wasserhaushaltsmodell LARSIM-Modellgrundlagen und Anwendungsbeispiele, 2000.

Brier, G.: Verification of forecasts expressed in terms of probability, Monthly Weather Review, 78, 1–3, 1950.

Brown, B., Katz, R., and Murphy, A.: On the economic value of seasonal-precipitation forecasts: The fallowing/planting problem, Bulletin of the American Meteorological Society, 67, 833–841, 1986.

Brunet, N., Verret, R., and Yacowar, N.: An objective comparison of Model Output Statistics and Perfect Prog systems in producing numerical Weather element forecasts, Weather and Forecasting, 3, 273–283, 1988.

Buishand, T. and Brandsma, T.: Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling, Water Resources Research, 37, 2761–2776, 2001.

Buizza, R., Houtekamer, P., Toth, Z., Pellerin, G., Wei, M., and Zhu, Y.: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems, Monthly Weather Review, 133, 1076–1097, 2005.

Du, J., Mullen, S., and Sanders, F.: Short-range ensemble forecasting of quantitative precipitation, Monthly Weather Review, 125, 2427–2459, 1997.

Engel, H.: The flood events of 1993/1994 and 1995 in the Rhine River basin, IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 239, 21–32, 1997.

Epstein, E.: A scoring system for probability forecasts of ranked categories, Journal of Applied Meteorology, 8, 985–987, 1969.

Fraedrich, K. and Rückert, B.: Metric adaption for analog forecasting, Physica A, 253, 379–393, 1998.

Fraedrich, K., Raible, C., and Sielmann, F.: Analog ensemble forecasts of tropical cyclone tracks in the Australian region, Weather and Forecasting, 18, 3–11, 2003.

Fraedrich, K., Aigner, A. A., Kirk, E., and Lunkeit, F.: General Circulation Models of the atmosphere. In Encyclopedia of Nonlinear Science, Routledge, New York, p 359-361, 2005.

Gigerenzer, G., Hertwig, R., Van den Broek, E., Fasolo, B., Katsikopoulos, K., and Street, H.: A 30% chance of rain tomorrow: How does the public understand probabilistic weather forecasts?, Risk Analysis, 25, 623–629, 2005.

Giorgi, F. and Mearns, L.: Approaches to the simulation of regional climate change: A review, Reviews of Geophysics, 29, 1991.

Glahn, H.: Statistical weather forecasting, Probability, Statistics, and Decision Making in the Atmospheric Sciences, pp. 289–335, 1985.

Glahn, H. and Lowry, D.: The use of Model Output Statistics (MOS) in objective weather forecasting, Journal of Applied Meteorology, 11, 1203–1211, 1972.

Grell, G., Dudhia, J., Stauffer, D., for Atmospheric Research Mesoscale, N. C., and Dicision, M. M.: A description of the fifth-generation Penn State/NCAR mesoscale model: MM5, National Center for Atmospheric Research. Mesoscale and Microscale Meteorology Division, 1995.

Grotch, S. and MacCracken, M.: The use of General Circulation Models to predict regional climatic change, Journal of Climate, 4, 286–303, 1991.

Hamill, T. and Colucci, S.: Verification of Eta-RSM short-range ensemble forecasts, Monthly Weather Review, 125, 1312–1327, 1997.

Hangen-Brodersen, C., Vogelbacher, A., and Holle, F.: Operational flood forecast in Bavaria, in: XXIVth Conference of the Danubian Countries, p. 9, 2008.

Heidke, P.: Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst, Geogr. Ann, 8, 301–349, 1926.

Hess, P. and Brezowsky, H.: Katalog der Grosswetterlagen Europas, Dt. Wetterdienst, 1969.

Hewitson, B. and Crane, R.: Climate downscaling: techniques and application, Climate Research, 7, 85–95, 1996.

Houtekamer, P., Lefaivre, L., Derome, J., Ritchie, H., and Mitchell, H.: A system simulation approach to ensemble prediction, Monthly Weather Review, 124, 1225–1242, 1996.

Hundecha Hirpa, Y.: Regionalization of parameters of a conceptual rainfall-runoff model, Ph.D. thesis, 2005.

Jolliffe, I. and Stephenson, D.: Forecast Verification: A practitioner's guide in atmospheric science, Wiley, 2003.

Jones, P., Hulme, M., and Briffa, K.: A comparison of Lamb circulation types with an objective classification scheme, International Journal of Climatology, 13, 1993.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American Meteorological Society, 77, 437–471, 1996.

Kanamitsu, M.: Description of the NMC Global Data Assimilation and Forecast System, Weather and Forecasting, 4, 335–342, 1989.

Karl, T., Schlesinger, M., and Wang, W.: A method of relating general circulation model simulated climate to the observed local climate. I- Central tendencies and dispersion, in: Conference on Applied Climatology, 6 th, Charleston, SC, pp. 188–196, 1989.

Karl, T., Wang, W., Schlesinger, M., Knight, R., and Portman, D.: A method of relating General Circulation Model simulated climate to the observed local climate. Part I: Seasonal statistics, Journal of Climate, 3, 1053–1079, 1990.

Katz, R., Murphy, A., and Winkler, R.: Assessing the value of frost forecasts to orchardists: A dynamic decision-making approach, Journal of Applied Meteorology, 21, 518–531, 1982.

Klein, W., Lewis, B., and Enger, I.: Objective prediction of five-day mean temperatures during winter, Journal of the Atmospheric Sciences, 16, 672–682, 1959.

Kolb, L. and Rapp, R.: The utility of weather forecasts to the raisin industry, Journal of Applied Meteorology, 1, 8–12, 1962.

Kruizinga, S. and Murphy, A.: Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands, Monthly Weather Review, 111, 2244–2254, 1983.

Krzysztofowicz, R.: A theory of flood warning systems, Water Resources Research, 29, 1993.

Krzysztofowicz, R., Drzal, W., Rossi Drake, T., Weyman, J., and Giordano, L.: Probabilistic quantitative precipitation forecasts for river basins, Weather and forecasting, 8, 424–439, 1993.

Lamb, H.: British Isles weather types and a register of the daily sequence of circulation patterns, 1861-1971, HMSO, 1972.

Lemcke, C. and Kruizinga, S.: Model Output Statistics forecasts: Three years of operational experience in the Netherlands, Monthly Weather Review, 116, 1077–1090, 1988.

Liljas, E. and Murphy, A.: Anders Ångström and his early papers on probability forecasting and the use/value of weather forecasts, Bulletin of the American Meteorological Society, 75, 1227–1236, 1994.

Liu, R., Parelius, J., and Singh, K.: Multivariate analysis by data depth: descriptive statistics, graphics and inference, Annals of statistics, pp. 783–840, 1999.

Lorenz, E.: Atmospheric predictability as revealed by naturally occurring analogues, Journal of the Atmospheric Sciences, 26, 636–646, 1969.

Lorenz, E. N.: Deterministic nonperiodic flow, Journal of the Atmospheric Sciences, 20, 130–141, 1963.

Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., and Baumgardner, J.: The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests, Monthly Weather Review, 130, 319–338, 2002.

Malitz, G. and Schmidt, T.: Hydrometeorologische Aspekte des Sommerhochwassers der Oder 1997.Klimastatusbericht 1997. DWD, p. 28, 1997.

Marsigli, C., Montani, A., Nerozzi, F., and Paccagnella, T.: Probabilistic high-resolution forecast of heavy precipitation over Central Europe, Natural Hazards and Earth System Sciences, 4, 315–322, 2004.

Marsigli, C., Boccanera, F., Montani, A., and Paccagnella, T.: The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification, Nonlinear Processes in Geophysics, 12, 527–536, 2005.

Mason, I.: Binary Events. Forecast verification. A practitioners guide in atmospheric sciences, 137-163. Eds I. Jolliffe and DB Stephenson, pp. 37–76, 2003.

Matulla, C., Zhang, X., Wang, X., Wang, J., Zorita, E., Wagner, S., and Storch, H.: Influence of similarity measures on the performance of the analog method for downscaling daily precipitation, Climate Dynamics, 30, 133–144, 2007.

Molteni, F., Buizza, R., Palmer, T., and Petroliagis, T.: The ECMWF ensemble prediction system: methodology and validation, Quarterly Journal of the Royal Meteorological Society, 122, 73–119, 1996.

Montani, A., Capaldo, M., Cesari, D., Marsigli, C., Modigliani, U., Nerozzi, F., Paccagnella, T., Patruno, P., and Tibaldi, S.: Operational limited-area ensemble forecasts based on the Lokal Modell, ECMWF Newsletter, 98, 2–7, 2003.

Morss, R., Demuth, J., and Lazo, J.: Communicating uncertainty in weather forecasts: A survey of the US public, Weather and Forecasting, 23, 974–991, 2008.

Mullen, S. and Buizza, R.: The impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF Ensemble Prediction System, Weather and Forecasting, 17, 173–191, 2002.

Murphy, A.: A new vector partition of the probability score, Journal of Applied Meteorology, 12, 595–600, 1973.

Murphy, A.: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation, Monthly Weather Review, 105, 803–816, 1977.

Murphy, A.: Probabilistic weather forecasting, Probability, statistics, and decision making in the atmospheric sciences, pp. 337–377, 1985.

Murphy, A.: Probabilities, odds, and forecasts of rare events, Weather and Forecasting, 6, 302–307, 1991.

Murphy, A.: What Is a good forecast? An essay on the nature of goodness in weather forecasting, Weather and Forecasting, 8, 281–293, 1993.

Murphy, A.: The Finley affair: A signal event in the history of forecast verification, Weather and Forecasting, 11, 3–20, 1996.

Murphy, A. and Katz, R.: Probability, statistics, and decision making in the atmospheric sciences, Westview Press, 1985.

Murphy, A., Lichtenstein, S., Fischhoff, B., and Winkler, R.: Misinterpretations of precipitation probability forecasts, Bulletin of the American Meteorological Society, 61, 695–701, 1980.

Namias, J.: General aspects of extended range forecasting, Compendium of Meteorology, pp. 802–813, 1951.

NCEP: The GFS atmospheric model, NCEP Office Note, 442, 14, 2003.

NCEP: NCEP product inventory, http://www.nco.ncep.noaa.gov/pmb/products/gfs/, 22. December 2008, 2008.

NCEP: NCEP/NCAR reanalysis problems list, http://www.cdc.noaa.gov/data/reanalysis/problems.shtml, 27. March 2009, 2009a.

NCEP: GFS/GDAS changes since 1991, http://wwwt.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html, 12. November 2009, 2009b.

NCEP: NCEP ftp server with the GFS forecasts, ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/gfs/prod/, 12. November 2009, 2009c.

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmospheric Research, 63, 303–324, 2002.

Obled, C., Djebboua, A., and Bontron, G.: Toward a simple probabilistic hydrometeorological forecasting chaing, in: Proc. 5th EGU Plinius Conf., Ajaccio, Corsica, France, 1–3 October, 2003, pp. 168–185, 2003.

Obled, C., Djerboua, A., Zin, I., and Garçon, R.: A simple probabilistic flood forecasting chain with focus on the use of QPFs, in: Proc. of the ESF LESC Exploration Workshop, Bologna, Italy, in: Hydrological risk recent advances in peak river flow modelling, prediction and real-time forecasting-assessment of the impacts of land-use and climate changes, edited by: Brath, A., Montanari, A., and Toth, E., Editoriale Bios, Castrolibero, CS, Italy, pp. 168–185, 2004.

O'Connor, J. E. and Costa, J. E.: The world's largest floods, past and present: Their causes and magnitudes: US Geological Survey Circular 1254, 13p., 2004.

Panofsky, H. and Brier, G.: Some applications of statistics to meteorology, Mineral Industries Extension Services, College of Mineral Industries, Pennsylvania State University, 1958.

Peel, M., Finlayson, B., and McMahon, T.: Updated world map of the Koppen-Geiger climate classification, Hydrology and Earth System Sciences, 11, 1633, 2007.

Peirce, C.: The numerical measure of the success of predictions., Science, 4, 453–454, 1884.

Philipp, A. and Jacobeit, J.: Das Hochwasserereignis in Mitteleuropa im August 2002 aus klimatologischer Perspektive, Petermanns geographische Mitteilungen, 147, 50, 2003.

Potts, J.: Basic concepts. Forecast Verification. A practitioners guide in atmospheric sciences, 137-163. Eds I. Jolliffe and DB Stephenson, 2003.

Radinović, D.: An analogue method for weather forecasting using the 500/1000 mb Relative Topography, Monthly Weather Review, 103, 639–649, 1975.

Richardson, D.: Skill and relative economic value of the ECMWF ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 126, 649–667, 2000.

Richardson, D.: Economic value and skill. Forecast verification. A practitioners guide in atmospheric sciences, 137-163. Eds I. Jolliffe and DB Stephenson, 2003.

Roulin, E. and Vannitsem, S.: Skill of medium-range hydrological ensemble predictions, Journal of Hydrometeorology, 6, 729–744, 2005.

Rudolf, B., Frank, H., Grieser, J., Müller-Westermeier, G., Rapp, J., and Trampf, W.: Hydrometeorologische Aspekte des Hochwassers in Südbayern im August, Statusbericht. Deutscher Wetterdienst., 2005.

Samaniego, L. and Bárdossy, A.: Relating macroclimatic circulation patterns with characteristics of floods and droughts at the mesoscale, Journal of Hydrology, 335, 109–123, 2007.

Samaniego, L., Bardossy, A., and Schulz, K.: Supervised classification of remotely sensed imagery using a modified$ k$-NN technique, IEEE Transactions on Geoscience and Remote Sensing, 46, 2112–2125, 2008.

Schättler, U., Doms, G., and Schraff, C.: Description of the nonhydrostatic regional model (LM). Part VII: users guide, COSMO-Consortium for Small Scale Modelling, 2008.

Sievers, O., Fraedrich, K., and Raible, C.: Self-adapting analog ensemble predictions of tropical cyclone tracks, Weather and Forecasting, 15, 623–629, 2000.

Simmons, A.: Global Numerical Weather Prediction: The past. Workshope note: The past, present and future of numerical modelling. 16th BMRC Modelling Workshop. 6-9 December. 2004. Melbourne, 2004.

Simmons, A., Burridge, D., Jarraud, M., Girard, C., and Wergen, W.: The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution, Meteorology and Atmospheric Physics, 40, 28–60, 1989.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., and Powers, J.: A description of the advanced research WRF Version 2, NCAR Tech Notes-468+ STR, 2005.

Stehlík, J. and Bárdossy, A.: Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation, Journal of Hydrology, 256, 120–141, 2002.

Teweles, S. and Wobus, H.: Verification of prognostic charts, Bull. Amer. Meteor. Soc, 35, 455–463, 1954.

Thompson, J. and Brier, G.: The economic utility of weather forecasts, Monthly Weather Review, 83, 249–253, 1955.

Toothill, J.: Central European Flooding August 2002 - An EQECAT Technical Report, 2002.

Toth, Z.: Intercomparison of circulation similarity measures, Monthly Weather Review, 119, 55–64, 1991.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: The generation of perturbations, Bulletin of the American Meteorological Society, 74, 2317–2330, 1993.

Toth, Z., Talagrand, O., Candille, G., and Zhu, Y.: Probability and ensemble forecasts. A practitioners guide in atmospheric sciences, 137-163. Eds I. Jolliffe and DB Stephenson, 2003.

Tukey, J.: Mathematics and the picturing of data, in: Proceedings of the International Congress of Mathematicians, vol. 2, pp. 523–531, 1975.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2005.

Vardi, Y. and Zhang, C.: The multivariate L1-median and associated data depth, 2000.

von Storch, H. and Zwiers, F.: Statistical analysis in climate research, Cambridge University Press, 1999.

von Storch, H., Zorita, E., and Cubasch, U.: Downscaling of global climate change estimates to regional scales: An application to iberian rainfall in wintertime, Journal of Climate, 6, 1161–1171, 1993.

Von Storch, H., Güss, S., and Heimann, M.: Das Klimasystem und seine Modellierung: Eine Einführung, Springer, 1999.

Wetterhall, F.: Statistical downscaling of precipitation from large-scale atmospheric circulation: Comparison of methods and climate regions [Statistisk nedskalning av nederbörd från storskalig atmosfärscirkulation: Jämförelse mellan metoder och klimatregioner], Uppsala universitet, Teknisk-naturvetenskapliga vetenskapsområdet, Geovetenskapliga sektionen, Institutionen för geovetenskaper, 2005.

Wetterhall, F., Halldin, S., and Xu, C.: Statistical precipitation downscaling in central Sweden with the analogue method, Journal of Hydrology, 306, 174–190, 2005.

Whitaker, J.: pygrib2 - python module for reading and writing GRIB2 files, http://code.google.com/p/pygrib2, 12. June 2009, 2009.

Wilby, R.: Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices, Climate Research, 10, 163–178, 1998.

Wilby, R. and Wigley, T.: Downscaling General Circulation Model output: A review of methods and limitations, Progress in Physical Geography, 21, 530, 1997.

Wilby, R. and Wigley, T.: Precipitation predictors for downscaling: Observed and General Circulation Model relationships, Int. J. Climatol, 20, 641–661, 2000.

Wilks, D.: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models, Agricultural and Forest Meteorology, 93, 153–169, 1999.

Wilks, D.: Statistical methods in the atmospheric sciences, Academic Press, 2nd edition, 2006.

Wilks, D. and Wilby, R.: The weather generation game: a review of stochastic weather models, Progress in Physical Geography, 23, 329, 1999.

Wilson, L. and Vallée, M.: The Canadian updateable model output statistics (UMOS) system: Design and development tests, Weather and Forecasting, 17, 206–222, 2002.

Winkler, R. and Murphy, A.: Decision analysis, Probability, Statistics, and Decision Making in the Atmospheric Sciences, pp. 493–524, 1985.

Woodcock, F.: On the use of analogues to improve regression forecasts, Monthly Weather Review, 108, 292–297, 1980.

Xu, C.: From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches, Progress in Physical Geography, 23, 229, 1999.

Yarnal, B.: Synoptic climatology in environmental analysis, Belhaven Press, 1992.

Zhu, Y.: Probabilistic forecasts and evaluations based on a global ensemble prediction system, Observation, Theory and Modeling of Atmospheric Variability: Selected Papers of Nanjing Institute of Meteorology Alumni in Commemoration of Professor Jijia Zhang, 2004.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., and Mylne, K.: The economic value of ensemble-based weather forecasts, Bulletin of the American Meteorological Society, 83, 73–83, 2002.

Zorita, E. and von Storch, H.: A survey of statistical downscaling techniques, NASA, 1997.

Zorita, E. and von Storch, H.: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods, Journal of Climate, 12, 2474–2489, 1999.

# Curriculum Vitae

| | |
|---|---|
| 22 September 1976 | born in Sulingen, Germany |

**Experiences**

| | |
|---|---|
| since 06/2010 | Post-Doc at the Institute for Geography of the University of Augsburg, Chair of Regional Climate and Hydrology, Prof. Kunstmann |
| 05/2004 - 05/2010 | Research assistant at the Institute of Hydraulic Engineering of the University of Stuttgart, Chair of Hydrology and Geohydrology, Prof. Bárdossy |
| 2003 - 2004 | Internship at the Gen Re (Reinsurance Company) in Cologne, Natural Hazard Division |
| 2002 | Internship at the German Research Center For Geoscience (GFZ) in Potsdam, Section Hydrology |

**Research projects**

| | |
|---|---|
| soon | WASCAL: Regional weather and climate predictions for West Africa |
| current | Implementation of a probabilistic precipitation forecast system for a regional flood forecasting center in Germany |
| 2005 - 2008 | HORIX: Development of an operational flood forecasting system for mesoscale river basins in Germany |
| 2005 - 2008 | PREVIEW: Enhancement of the Bavarian flood forecasting system for the Upper Danube basin |

**Education**

| | |
|---|---|
| 2004 - 2009 | Assistant of the following lectures: Hydrology, Statistics, Stochastic Hydrology and Multi-Objective Decision Making in Water Resources Management |
| 1997 - 2003 | Study of Environmental Science (Geoecology) at the Technical University of Brunswick, Germany |
| 1996 - 1997 | Alternative civilian serve in a German organization of physically and mentally disabled children |

**Institut für Wasserbau**
**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

1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und
Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963

2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964

3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächentragwerke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Randwert- und Lastwertmatrizen*, 1965

4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

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: *Versucheinrichtung 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, ISBN 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

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: Upscaling, 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, 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 Hochwasserrückhaltebecken-systemen,* 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 3-933761-57-3

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*, 2010, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2010, 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,* 2010, ISBN 978-3-933761-98-9

*Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.*