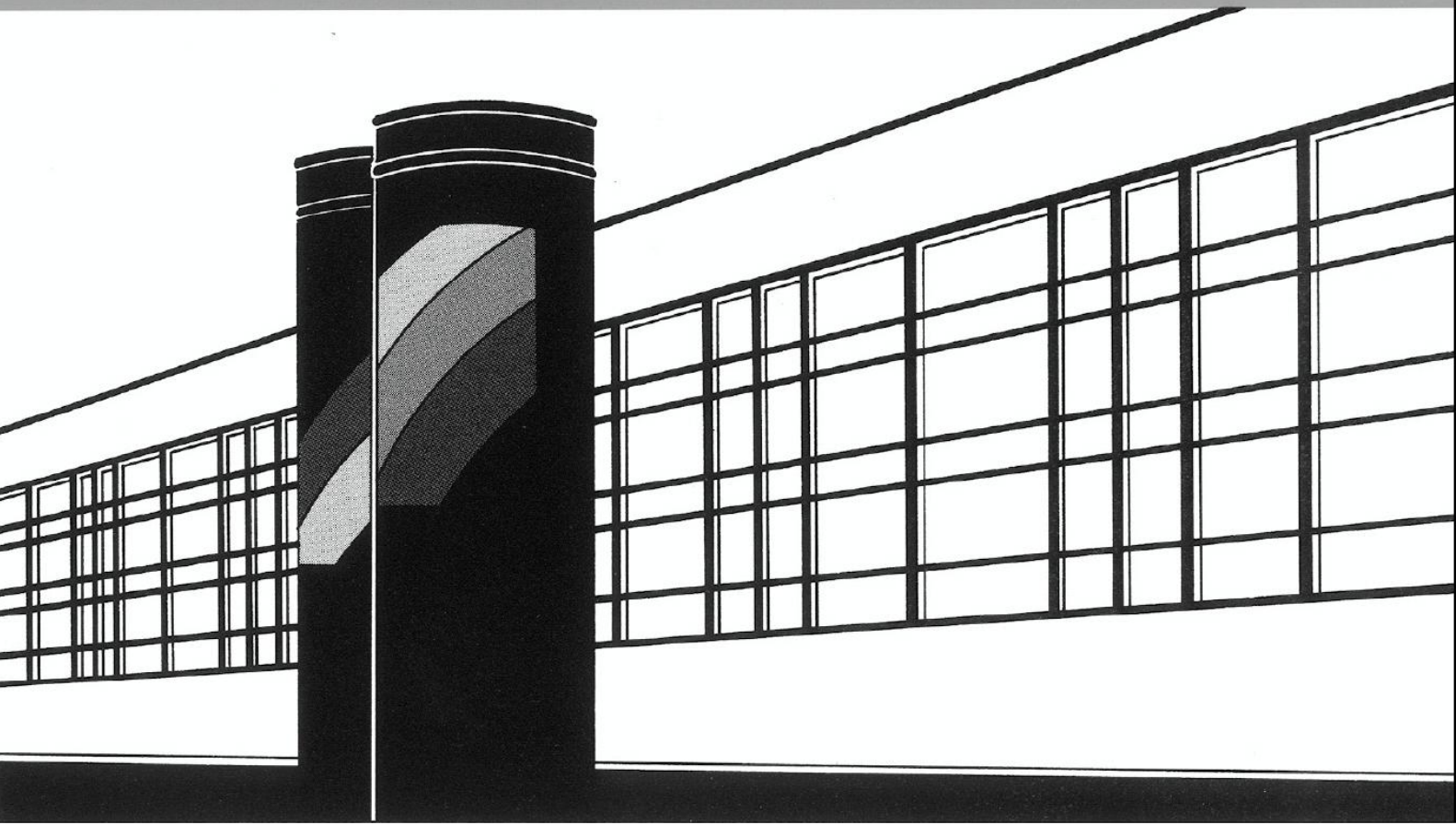


Universität Stuttgart



Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 240 Alejandro Chamorro Chávez

Stochastic and Hydrological Modelling
for Climate Change Prediction
in the Lima Region, Peru

Stochastic and Hydrological Modelling for Climate Change Prediction in the Lima Region, Peru

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

vorgelegt von
Alejandro Chamorro Chávez
aus Chile

Hauptberichter: Prof. Dr. rer. nat. Dr. -Ing. András Bárdossy
Mitberichter: Prof. Dr. -Ing Axel Bronstert
Prof. Dr. -Ing. Martin Kranert

Tag der mündlichen Prüfung: 08. 12. 2014

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2015

Heft 240 Stochastic and Hydrological
Modelling for Climate Change
Prediction in the Lima Region,
Peru

von
Dr.-Ing.
Alejandro Chamorro Chávez

D93 Stochastic and Hydrological Modelling for Climate Change Prediction in the Lima Region, Peru

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

Alejandro Chamorro Chávez:

Stochastic and Hydrological Modelling for Climate Change Prediction in the Lima Region, Peru, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2015

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 240)

Zugl.: Stuttgart, Univ., Diss., 2015

ISBN 978-3-942036-44-3

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

Gegen Vervielfältigung und Übersetzung bestehen keine Einwände, es wird lediglich um Quellenangabe gebeten.

Herausgegeben 2015 vom Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung

Acknowledgement

I am deeply grateful for having had the possibility of doing my PhD at the University of Stuttgart. Special thanks to Prof. Dr. rer. nat Dr. Ing. András Bárdossy for giving me the chance to learn a lot joining his research group. Thanks for valuable suggestions and ideas promoting my scientific development and for fostering my interest to and passion for the statistical and hydrological “world”. I really feel honored to have been his student during the time I performed my research. Many thanks also to Prof. Dr. Axel Bronstert for co-supervising my thesis.

I would like to thank many colleagues at IWS for their support in many issues, comments, feedbacks, friendship. I can mention, between others, Pawan Thapa, Shailesh Singh, Jhan Rodriguez, Henning Lebreuz, Thomas Pfaff.

Dr. rer. nat. Jochen Seidel was an invaluable help from the very beginning till the end at IWS. Your great attitude, disposition at any time and your ability to find solutions was a great support to my work. Many thanks Jochen.

My deepest gratitude to my family in their limitless support at any time. To my parents, encouraging me to continue learning in all aspects. To my brother for his wise support on theoretical and mathematical issues. To my sister, for always transmitting me energy and remembering me of my far dreams. To Sabine, who was always at my side.

Above all, to God, for giving me this opportunity in this present life.

Contents

List of Figures	vii
List of Tables	xi
Abstract	xiii
Kurzfassung	xvi
1 Introduction	1
1.1 Understanding Unpredictability	1
1.2 Some Physical Ideas	2
1.3 Objectives and PhD Thesis Structure	3
2 Study Area and Data	5
3 Geostatistics and Regionalization Methods	9
3.1 General Introduction	9
3.2 A Little Review in Geostatistics	9
Theoretical Variograms	12
3.3 Two Regionalization Methods	13
Ordinary Kriging	13
External Drift Kriging	16
Application: Comparison Between OK and EDK	18
Performing EDK in the Study Areas	18
3.4 Transformations	23
Basic Idea	23
Smoothing	23
Power Transformation	26
Rotation	30
3.5 Discussion of the Results	34
4 Downscaling	36
4.1 General	36
4.2 Quantil Quantil Transformation	36
4.3 Parametric and Non-Parametric Fit	38
Parametric Approach	38
Non-Parametric Approach	39
The kernel method	40
Orthogonal series density estimation	43

4.4	Application	47
	Downscaling Result on the Pacific Side of the Andes	47
	Precipitation pattern	48
	Downscaling result for precipitation	48
	Downscaling result for temperature	50
	Downscaling Result on the Atlantic Side	55
4.5	Summary and Final Comments	57
5	Hydrological Modeling	59
5.1	General	59
5.2	HBV and Hymod Model	60
	HBV Model	60
	Hymod	64
5.3	Parameter Estimation	64
	Objective Function	65
	Simulated Annealing	66
5.4	Application	68
5.5	Modeling in an Ungauged Catchment	71
5.6	Final Comments	75
6	Robust Estimation	76
6.1	General	76
6.2	Robust Parameter Estimation	76
	Depth Functions	77
	ROPE Algorithm	78
6.3	Application and results	79
7	Climate Change Impact in the Discharge: Future Assessment	83
7.1	General	83
7.2	Discharge Estimation	83
7.3	Future Assessment: Models Results	84
	Results from Simulated Annealing Optimization	84
	Results from the ROPE Algorithm	87
7.4	Summary and Conclusions	92
8	Time Series and Copula-Based Autorregressive Models	93
8.1	General	93
8.2	Trend Analysis	93
	General Idea of Trend Analysis	93
	Trend Estimate by Bootstrap Method	93
8.3	Discharge assessment: Autorregressive models and Copula based models	98
	Autorregressive models	98
	Copula-based autorregressive model	99
8.4	Summary and Final Comments	108
9	Conclusions and Outlook	110
9.1	Conclusions	110
9.2	Outlook	112

10 Bibliography

114

List of Figures

1.1	Wave Packet	2
1.2	Fourier transform (blue line) and inverse Fourier Transform (red line) representing a matter wave	3
2.1	Study area: Chillón, Rimac and Lurín (pacific side) , Marcas region (upper part of Mantaro catchment)	6
2.2	Rain stations used in the study	7
2.3	Natural discharge generated in the sub catchments Santa Eulalia (a), Atarjea (b) and Obrajillo (c)	7
3.1	Interpolation results for the catchment Chillon using both OK (figure on the top) and EDK (figure on the bottom) for the 2nd March 1999	19
3.2	Estimated monthly average precipitation for Chillón, Rimac and Lurín for the time period 1999-2008	21
3.3	Monthly average precipitation for the individual catchments Chillon Rimac and Lurin. Time period 1999-2008	22
3.4	Interpolation result for the catchment Rimac, considering interpolation with real digital elevation model data (figure on the top) and considering smoothed data (figure on the bottom)	24
3.5	Correlation between the vector error \vec{e}_v and elevation for the range $1 \leq R \leq 10$ km	26
3.6	Cumulative error (eq. 3.28) for different values of the parameter R	26
3.7	Correlation between the vector error \vec{e}_v and elevation \vec{z}_v for different exponents β (eq. 3.30)	27
3.8	Overall error (eq. 3.28) for a set of different exponents β	28
3.9	Correlation between the vector error \vec{e}_v and elevation \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents β and radius R	29
3.10	Correlation between the vector error \vec{e}_v and elevation \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents β and radius R , months December to April	30
3.11	Correlation between the vector error \vec{e}_v and vector elevation \vec{z}_v for a set of different K -parameters according to the set (3.35)	31
3.12	Overall error (eq. 3.28) for a set of different K -parameters according to the set (3.35)	31
3.13	Correlation between \vec{e}_v and \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents $beta$ and K -parameters	33
3.14	Correlation between \vec{e}_v and \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents $beta$ and K -parameters for the rainy season	34

4.1	Downscaling procedure. An input $\bar{v}_m^i(k)$ is evaluated into $G(\cdot)$, matched to $F(\cdot)$ and the downscaled value obtained from $F_i^{-1}(\cdot)$	37
4.2	Theoretical and empirical fitted distributions for the months December (upper left figure), march (upper right figure), June (lower left figure) and September (lower right figure) for the sub catchment Obrajillo, representing the begin of the seasons summer, autumn, winter and spring respectively	39
4.3	Yearly average precipitation relative frequency for the sub catchment Obrajillo	41
4.4	Naive estimate for the yearly average precipitation, sub catchment Obrajillo.	42
4.5	Kernel estimate for the yearly average precipitation, sub catchment Obrajillo using a Gauss Kernel function.	42
4.6	Kernel estimate of the yearly average precipitation in Obrajillo by using a log-transformation	43
4.7	Beta distributed sample	44
4.8	Density estimation based on Orthogonal Series with different limits l for the subindex j : $l = 30$ (fig. a), $l = 50$ (fig. b), $l = 100$ (fig. c) and $l = 500$ (fig. d)	45
4.9	Density estimation based on Orthogonal Series with optimal choice of the parameter l	47
4.10	Defined sub catchments for the purpose of downscaling and hydrological modeling: Obrajillo in catchment Chillón and Santa Eulalia (Sheque) and Atarjea in catchment Rimac	48
4.11	Monthly average precipitation pattern in the three analysed sub catchments: Santa Eulalia (upper figure) and Atarjea (middle figure) in the catchment Rimac, and Obrajillo (lower figure) in the catchment Chillón. Time span ranges from 1999 to 2008	49
4.12	Downscaling results for Santa Eulalia (upper figure) and Atarjea (Lower figure). Curves represent the monthly average precipitation for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod	51
4.13	Downscaling results for Obrajillo (upper figure) and Sub Lurin (Lower figure). Curves represent the monthly average precipitation for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod	52
4.14	Temperature downscaling results for Santa Eulalia (upper figure) and Atarjea (lower figure). Curves represent the monthly average temperature for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod	54
4.15	Temperature downscaling results for Obrajillo (upper figure) and Sub Lurín (lower figure). Curves represent the monthly average temperature for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod	55
4.16	Regions on the Atlantic side of the Andes comprising the 5 Marca project areas.	56
4.17	Present and downscaled precipitation for the period 1999-2008 and 2012-2050 respectively for the entire area involving the 5 Marca projects.	57
5.1	Scheme of HBV-IWS model used in this study	63

5.2	Schematic representation of Hymod model	64
5.3	Monthly average volume in lakes, Sub catchment Santa Eulalia	66
5.4	System in which a given configuration leaves a local minimum	67
5.5	Evolution of the objective function in Simulating Annealing when calibrating Santa Eulalia sub catchment	68
5.6	Calibration and validation of HBV model, subcatchment Santa Eulalia	69
5.7	Calibration and validation Hymod model, subcatchment Santa Eulalia	69
5.8	Calibration and validation result given by HBV model, subcatchment Atarjea	70
5.9	Calibration and validation result given by HBV model, subcatchment Obrajillo	71
5.10	Sub Lurin sub catchment and rain stations in Lurin	72
5.11	Discharge estimation in Atarjea with two different set of parameters: Atarjea and Santa Eulalia	72
5.12	Discharge estimation in Obrajillo with two different set of parameters: Obrajillo and Santa Eulalia	73
5.13	Histogram of the error (difference) between discharges estimated with real and transferred parameter set. left Atarjea, right Obrajillo	74
5.14	Monthly average generated discharge for the sub catchments Santa Eulalia, Atarjea, Obrajillo and Sub Lurin for the time period 1999-2008	74
6.1	Plot of reservoir parameters showing the outcomes of iteration 2 (run 3) and iteration 4 (run 5) of ROPE algorithm	80
6.2	Histograms for parameters K0 and K2. Top and down left, iteration 2. Top and down right, iteration 4.	81
6.3	Modeled discharge considering the parameter vector from the ROPE algorithm outcomes corresponding to the best and worst performance	81
6.4	Histogram model performances for different iterations	82
7.1	Subcatchments on the pacific side of the Andes for discharge time series estimation in the period 2012-2050	84
7.2	Monthly average generated discharge for the sub catchments Santa Eulalia (upper figure) and Atarjea (lower figure) for the different scenarios. Time period 1999-2008 (Present) and 2012-2050	85
7.3	Monthly average generated discharge for the sub catchments Obrajillo (upper figure) and Sub Lurín (lower figure) for the different scenarios. Time period 1999-2008 (Present) and 2012-2050 (future)	86
7.4	Monthly average generated (present) and estimated (future) discharge for Marca 1 (upper figure), Marca 2 (middle figure) and Marca 3 (lower figure). Time period 1999-2008 (present) and 2012-2050 (future)	88
7.5	Monthly average generated (present) and modeled (future) discharge for Marca 4 (upper figure) and Marca 5 (lower figure). Time period 1999-2008 (Present) and 2012-2050 (future)	89
7.6	Estimated variation of the discharge for the model Hadley and scenario A2 considering the ROPE algorithm outcomes	90
7.7	Histograms of the expected discharge variations for the sub catchment Santa Eulalia considering the vector sets form the ROPE algorithm. Time period 1999-2008 (present) and 2012-2050 (future)	91
8.1	Location of the Gauge stations Pariacancha and San José de Parac in the catchments Chillón and Rimac respectively	94

8.2	Yearly precipitation in the gauge stations Pariacancha and San José de Parac. The time period considered is from 1983 to 2008 for each station . . .	95
8.3	Bootstrap result for the stations Pariacancha and San José de Parac considering yearly precipitation	96
8.4	Bootstrap result for the stations Pariacancha and S.J. Parac for the rainy (left hand figures) and dry (right hand figures) season	97
8.5	Observed and 1-day ahead estimated discharge (AR process) for the year 2008, sub catchment Santa Eulalia	99
8.6	Empirical and Normal distribution functions for the time series TS	102
8.7	Empirical bivariate Copulas	103
8.8	Empirical 3D Copulas	104
8.9	Observed and 1-day ahead estimated discharge for the year 1999, sub catchment Santa Eulalia and considering a Gaussian Copula model	107
8.10	Individual daily errors between the Gaussian Copula model and the generated discharge for the year 1999	108
8.11	Modeled discharge, 10% quantil and 90% quantil for the last three months of the year 2008	108

List of Tables

2.1	Temperature stations used in this study. Period 1999-2008	8
3.1	Monthly average precipitation for the three main catchments Chillón, Rimac and Lurín	20
3.2	Calculated yearly average precipitation for Chillón, Rimac and Lurín for the time period 1999-2008	21
3.3	Dependence between the estimated error and height for the points with observations (stations), considering both a linear and squared relationship given by the equations (3.20) and (3.23)	22
3.4	Correlation between \vec{e}_v and \vec{z}_v , and overall error (eq. 3.28) representing the base case	26
4.1	Estimated yearly average precipitation according to the different scenarios analysed in the time period 2012 - 2050	50
4.2	Estimated variation of precipitation (%) for the period 2012-2050 with respect to the period 1999-2008	50
4.3	Yearly average temperature for the observation period 1999-2008 and scenarios period 2012-2050	56
4.4	Estimated yearly average variation of temperature ($^{\circ}C$) for the period 2012-2050 with respect to the period 1999-2008	56
4.5	Expected precipitation variation (%) for each region (Marca I - Marca V) for the period 2012-2050. Considered reference period 1999-2008	57
5.1	Parameters used in the model	64
5.2	Error associated to HBV and Hymod model for the subcatchments Santa Eulalia and Obrajillo [m^3/s]	69
5.3	Calibration and validation period for the sub catchments Santa Eulalia, Atarjea and Obrajillo	70
6.1	Error of the different vectors solution	79
7.1	Quantification of the climate change impact on the discharge (%) for each scenario. Sub catchments on the pacific side of the Andes	87
7.2	Quantification of the climate change impact on the discharge (%) according to each scenario. Areas correspond to Marca 1 to Marca 5	87
7.3	Average estimated variation of the discharge by applying the ROPE algorithm, sub catchment Santa Eulalia. Time period 2012-2050 with respect to 1999-2008	90

8.1	Observed trend (%) for the stations Pariacancha and S. J. Parac. Time period 1983-2008	94
8.2	Defined quantil of the observed trends on the sample distribution curve. Stations Pariacancha and San José de Parac	95
8.3	Comparison of the errors (RMSE) using copulas and using the model AR2	106

Abstract

Climate change has been an important field of research in the past years and certainly is a major concern in the present time. It involves a broad spectrum of subjects and significant different time scales, ranging from decades to thousands or millions of years. Generally speaking, in a climate change scenario a change in the pattern, average or extreme conditions of some variables is observed, and this can be due to many different causes as changing processes in the earth, human activities or extra terrestrial induced factors. This study concentrates on the influences on the climate due to human activities and focuses on the hydrological response to these influences or changes as a primarily goal, for the next few decades. The main motivation is the vulnerability and scarcity of the water availability in the Capital of Peru, Lima, and how the area under study will respond to a change in the climate. The extreme aridity in the city, with an average precipitation of about 10 mm/year makes it necessary to obtain the water resources from others regions and catchments. The inhomogeneous behavior and unfavorable distribution of precipitation has an important effect on the form of the water supply for this metropolitan area. An estimation of the future availability/variation of water especially in the upper regions is under these circumstances of major interest. Important questions are: Which effects are to be expected on precipitation, temperature and discharge, in a climate change scenario. An important focus of analysis in order to reduce the uncertainty in the predictions is the errors that appear when modeling a given variable or set of variables. Uncertainty is present in each step and in general it is present in the very act of observing (measuring) a quantity. Along the thesis this issue is addressed first in regionalization in which an unknown is desire to be calculated all along a discrete domain, given a set of observations in sampled points (locations), and Second in the calibration of hydrological models in which a robust parameter estimation in the process of optimization is performed. Here the final result is a set of different outcomes for the discharge variation for each area and scenario analysed. In the first issue concerning to regionalization, External Drift Kriging is applied to precipitation. Fundamentally, in EDK a description and modeling of the spatial variability of a random variable is performed through probabilistic and stochastic approaches. In this method important information can be summarized as the scale of fluctuations of the random variable which is subject to analysis, and also additional knowledge of an external variable is required to be incorporated into the model. In general terms the spatial structure is modeled and subsequently a linear estimation is carried out. In this part of the work the results of regionalization are analysed with focus on the errors and systematic errors which appear during the modeling. The main goal here is the reduction of these errors through some proposed transformations. Here, three approaches are suggested, namely smoothing of the digital elevation model (DEM) considering a symmetric area, power transformation and smoothing considering a non symmetric area. In the first case the z-component of the DEM is smoothed by means of a parabolic kernel function and then regionalization is performed in sampled locations. After changing the extent of the smoothing area the procedure is repeated until a clear tendency for the error to increase is observed. In the second case the attention is paid in the derivation of EDK. In the method, a linear relationship between the conditional expectation of the random variable to be regionalized and an external variable is supposed to exist. This condition is weakened and a non linear relationship is also considered to be a feasible assumption. The last case deals with smoothing the z-component of the DEM

as done in the first case, but here points inside a non symmetric area are considered. By defining a set of different angles this area is rotated and regionalization is performed each time. The systematic error expressed in terms of the correlation function is reduced to practically a half when a power law between the conditional expectation and the drift is assumed (second case). Although this constitutes an important result, a much more significant improvement is achieved when combining the previous approaches. In this way by considering smoothing and power transformation or smoothing with a non symmetric area and power transformation, the systematic error becomes considerably smaller and can be considered as negligible. The analysis was carried out in the whole time period, in the dry season and in the rainy season. Improvements were observed in all of the cases and the most significant reduction occurred in the rainy season. The second issue concerning uncertainty and related to parameter estimation was addressed two-fold, namely by optimizing the objective function by means of a heuristic optimization procedure based on Monte Carlo simulation, and by means of a robust parameter estimation (ROPE) algorithm developed quite recently by Bárdossy and Singh, which in general terms can be used as a general multivariate optimization procedure. This is a very important step in terms of the uncertainties and errors which may be introduced, especially when the main goal is to use the calibrated models for predictions. To this respect ROPE algorithm offers a way of finding a set of “good” parameter vectors which, among other characteristics, are transferable in time. The final result comprises an ensemble of estimations for expected discharge variations accounting for the uncertainty in parameterization and processes description in the model. In this way not a single estimation is constructed but a set of them. Two models were used for this purpose, namely the HBV and Hymod models. The assessment of the impact of climate change in the variables precipitation and temperature was carried out by a statistical downscaling procedure. More specifically, the estimations given by the Global Climate Models (GCMs) were statistically downscaled. This is a necessary step because of the scale mismatch between the GCMs outcomes and in general the scale involved for hydrological modeling purposes. A quantil-quantil downscaling which is based on the statistical characteristics of the sets of data was applied, in order to transform the large scale information to the local scale. The fitting of the cumulative distribution functions considered both parametric and non parametric approaches. In the first case a Weibull distribution which depends on two parameters was considered and fitted to precipitation data. The non parametric case was based on kernel functions and was applied to temperature information. The shape and scale parameter of the Weibull distribution as well as the smoothing parameter of the kernel based density estimator was found by the Likelihood function. As a result, monthly time series for precipitation and temperature for the estimation period 2012-2050 were obtained. This allowed us to get an insight into the possible changes which may occur in a climate change scenario for these two variables. The analysis done so far permitted us to have estimations for the temperature, precipitation and discharge for the next 40 years. The two GCMs models, namely Echem from Germany and Hadley from the UK and the different scenarios analysed permitted the definition of a range for the expected futures variations of these variables, accounting to same extent for the uncertainty in the human induced driven forces. These variations can be summarized as follows: regarding to precipitation, in almost all the cases a decrease is expected and in only one of the cases an increase may occur. They range between $\Delta P = -9.8\%$ and $\Delta P = 6.7\%$. For the discharge the same pattern is expected but the variations are more accentuated ranging from $\Delta Q = -13.7\%$ to $\Delta Q = 6.3\%$. Temperature is expected to increase regardless of the scenario consid-

ered, with a maximum expected variation of $\Delta T = 2.6^\circ\text{C}$. The last chapter of this thesis addresses the assessment of the discharge in the short term. The goal here is to “infer” the outcome of a random variable (discharge) in the next time step by taking information from past observations (previous steps). As we can regard the observations (time series) as a realization generated from a stochastic process, we can address this issue from a stochastic point of view. The task is addressed first by considering some of the existing autoregressive models (AR process), and second by considering a Copula-based autoregressive model. As a general concept, a joint distribution function can be expressed in terms of univariate marginal distributions by means of a Copula function or simply a Copula. In other words, by associating univariate marginal distributions by a Copula function a multivariate distribution function can be constructed. In order to perform the Copula-based autoregressive model, a given time series (modelled discharge) was transformed into three vectors representing the same original time series but shifted in time. A three dimensional Copula was then fitted to the univariate distributions. The empirical Copula suggested the Gaussian model to be more representative of the interdependence structure of the variables. In order to compare the results of the Copula model, a non parametric approach was also performed. Here a Beta kernel Copula was considered which is expressed in terms of the Beta function and depends on a single free parameter h . The time series used in the analysis was the discharge in one of the catchments involved in the study in a temporal resolution of 1 day. As the observed discharge was available only in a monthly time scale, the modelled discharge given by the HBV model was used instead. The results showed that the estimation given by both, the Gaussian Copula and the Beta kernel Copula model are very similar in magnitude and represent the real values well. The most significant differences with respect to the modeled discharge are observed in those periods where the discharge presents higher values and in those periods with higher discharge variability. By comparing these results with the AR process, a little improvement is observed by the Copula based autoregressive model, although the differences are not significant.

Kurzfassung

Klimawandel war in den vergangenen Jahren ein wichtiges Forschungsfeld und ist auch in der Gegenwart von großem Interesse. Die Thematik umfasst einerseits ein breit gefächertes Spektrum und andererseits signifikant unterschiedliche Zeitskalen, beginnend bei Jahreszeiten, über Jahrtausende bis hin zu Jahrmillionen. In einem Klimawandelszenario können im Allgemeinen Veränderungen einer oder mehrerer Variablen beobachtet werden, wie z. B. Veränderung in deren Struktur, in deren Durchschnitts- oder Extremwerten. Die Ursachen dafür können unterschiedlich sein, wie z. B. Veränderungen der Erde, menschenbedingte Veränderungen oder extraterrestrische Faktoren. Diese Forschungsarbeit untersucht den Einfluss von anthropogenen Aktivitäten auf das Klima, wobei der Fokus auf der hydrologischen Reaktion durch diese Einflüsse und Veränderungen in den nächsten Jahrzehnten liegt. Der zentrale Beweggrund für die Studie liegt in der Gefährdung und dem Mangel an Wasserverfügbarkeit in der Hauptstadt von Peru, Lima. Die extreme Trockenheit mit einem durchschnittlichen Jahresniederschlag von 10 mm in dieser Region erfordert den Bezug von Wasserressourcen aus anderen Regionen und Einzugsgebieten. Neben dem generellen Wassermangel hat die ungleichmäßige und ungünstige Verteilung des Niederschlags einen beachtenswerten Einfluss auf die Art und Weise der Wasserversorgung der Metropolregion. Eine Abschätzung des zukünftigen Vorhandenseins bzw. der zukünftigen Veränderungen der Wassernutzbarkeit, insbesondere in den höherliegenden Gegenden, ist unter diesen Gegebenheiten von großem Interesse. Wichtige Fragen sind: welche Auswirkungen des Klimawandels auf den Niederschlag, Temperatur und Durchfluss in Lima zu erwarten sind. Ein zentraler Aspekt bei der Abschätzung von Auswirkungen des Klimawandels in den oben erwähnten Kategorien liegt bei der Minimierung der Unsicherheit von Vorhersagen und von Fehlern, die bei der Modellierung einer oder mehrerer Variablen auftreten. Unsicherheiten treten in jedem Modellierungsansatz auf, so auch bei der Messung einer Menge, wie z.B. Niederschlag. Diese Thematik kann z. B. behandelt werden: Erstens bei der Regionalisierung, bei der unbekannte Größen in einem diskreten Definitionsbereich berechnet werden und zweitens bei hydrologischen Modellen, wo eine robuste Parameterschätzung durchgeführt wird. Das resultierende Ergebnis ist die Vorhersage der Abflussänderung in jeder untersuchten Region und für alle analysierten Szenarien. Für den ersten Aspekt, die Regionalisierung, wurde External Drift Kriging (EDK) für den Niederschlag angewendet. Grundlegend wird bei EDK die Beschreibung und Modellierung der räumlichen Variabilität einer Zufallsvariablen, mittels stochastischen Ansätzen durchgeführt. Mit Hilfe dieser Methode können wichtige Informationen zusammengefasst werden, wie z.B. die Bandbreite einer gegebenen Zufallsvariablen. Verrallgemeinert lässt sich sagen, dass die räumliche Struktur modelliert und anschließend eine lineare Schätzung durchgeführt wird. Dabei liegt ein besonderes Augenmerk auf den Fehlern und systematischen Fehlern, welche bei Regionalisierung auftreten können. Das Ziel hierbei ist es, diese Fehler durch Transformationen zu minimieren. Dafür wurden drei Ansätze gearbeitet: dem digitalen Höhenmodell (DHM) unter Betrachtung einer symmetrischen Fläche zu glätten, nicht-lineare Transformationen und der Glättung unter Berücksichtigung einer nicht symmetrischen Fläche. Beim ersten Ansatz wird die Höhe (z -Komponente) durch eine parabolische Kernel-Funktion geglättet. Anschließend wird für die Messpunkte (Niederschlagsstationen) Regionalisierung durchgeführt. Der Um-

fang der Fläche wird geändert und der Vorgang so lange wiederholt bis eine eindeutige Tendenz in der Fehlerzunahme zu beobachten ist. In einem zweiten Ansatz wird die Methode des EDK analysiert. Bei dieser Methode wird angenommen, dass ein linearer Zusammenhang zwischen dem Erwartungswert der regionalisierten Variablen und den externen Variablen besteht. Diese Annahme wurde im Folgenden nicht mehr vorausgesetzt und stattdessen ein nicht-linearer Zusammenhang angenommen. Für den letzten untersuchten Ansatz wird für die Glättung der Höheninformationen eine nicht symmetrische Fläche berücksichtigt. Es wurde ein Satz verschiedener Winkel definiert, um welche die Fläche jeweils gedreht und die Regionalisierung für jeden Zeitschritt durchgeführt wurde. Die Ergebnisse zeigen, dass bei einem angenommenen nicht-linearen Zusammenhang zwischen dem bedingten Erwartungswert und der Drift, der systematische Fehler ausgedrückt durch eine Korrelationsfunktion nahezu um die Hälfte reduziert wird. Obwohl dies ein wichtiges Ergebnis darstellt, wurden die Fehler durch die Kombination der vorher vorgestellten Transformationen weitaus reduziert. Hierdurch wurden die Fehler sehr klein und können vernachlässigt werden. Diese Analyse wurde für den gesamten Zeitraum, für die Regenzeit und für die Trockenzeit durchgeführt. Eine Verbesserung der Ergebnisse bzw. eine Reduktion der Fehler konnte in allen Fällen erzielt werden, wobei die signifikanteste Fehlerreduzierung während der Regenzeit zu beobachten war. Der zweite Aspekt in Bezug auf Fehlerminimierung wurde durch Parameterschätzung angegangen. Zum einen wurden das „Simulating Annealing“ und zum anderen eine robuste Parameterabschätzung (ROPE-Algorithmus) für die Optimierung der Zielfunktion angewandt. Der ROPE-Algorithmus wurde von Bárdossy und Singh entwickelt und kann als ein allgemeiner multivariater Optimierungsalgorithmus betrachtet werden. Dies war ein wichtiger Schritt zum Beschreiben der Unsicherheiten und Fehler, die bei der hydrologischen Modellierung auftreten. Mit dem ROPE-Algorithmus ist es möglich eine Menge „guter“ Parametervektoren zu finden, welche unter anderen Eigenschaften zeitlich übertragbar sind. Die zeitliche Übertragbarkeit ist insbesondere bei Modellen für Vorhersagen, die hier Thema sind, von zentraler Bedeutung. Als Ergebnis erhält man ein Ensemble möglicher Abflussänderungen. Die Auswirkungen des Klimawandels auf den Niederschlag und der Temperatur wurde ebenfalls analysiert. Da eine Diskrepanz zwischen den Maßstäben der verwendeten GCM's und der Untersuchungsgebiete besteht, wurde ein statistisches Downscaling-Verfahren durchgeführt. Es wurde eine Quantil-Quantil-Transformation angewandt, welche auf den statistischen Eigenschaften der Daten beruht. Auf diese Weise wird eine großflächige Information zu einer lokalen Information transformiert. Als Verteilungsfunktionen wurden parametrische und nicht-parametrische Ansätze verwendet. Für den Niederschlag wurde die Weibull-Verteilungsfunktion und für die Temperatur eine nicht-parametrische Kernfunktion angepasst. Die Skalen- und Formparameter der Weibull-Funktion sowie die Bandbreite der Kerndichtefunktion wurden mit Hilfe der Maximum-Likelihood-Methode optimiert. Damit konnten für den Vorhersagezeitraum von 2012-2050 monatliche Niederschlags- und Temperaturzeitreihen generiert werden. Diese Ergebnisse liefern eine Erkenntnis darüber, welche möglichen Veränderungen diese beiden Variablen auf ein Klimawandelszenario haben. Die Untersuchung stellt somit für den Niederschlag, die Temperatur und den Abfluss Abschätzungen für die nächsten 40 Jahre zur Verfügung. Zusammenfassend lassen sich die wichtigsten Ergebnisse wie folgt darstellen: Die Niederschlagsmenge hat eine klare Tendenz zur Abnahme. Bis auf ein Szenario zeigen alle analysierten Szenarien eine Abnahme der Niederschlagsmenge im Bereich von $\Delta P = -9.8\%$ bis $\Delta P = 6.7\%$. Die zu erwartenden Änderungen des Abflusses folgen einem ähnlichen Schema, wobei die Bandbreite zwischen $\Delta Q = -13.7\%$

und $\Delta Q = 6.3\%$ stärker variiert. Bei der Temperatur lässt sich mit einer maximale erwartete Zunahme von $\Delta T = 2.6^\circ\text{C}$ für alle untersuchten Szenarien eine deutliche Tendenz zur Temperaturzunahme erkennen. Der letzte Teil dieser Dissertation befasst sich mit Kurzzeitprognosen für den Abfluss. Das Ziel hierbei ist es, auf den Wert einer Zufallsvariablen (z.B. Abfluss) im nächsten Zeitschritt mit Hilfe von gemessenen bzw. beobachteten Daten aus vorherigen Zeitschritten zu schließen. Da man die Information (Zeitreihe) als eine aus einem stochastischen Prozess generierte Realisierung betrachten kann, kann man die Prognose aus einer stochastischen Perspektive heraus angehen. Zunächst wird ein existierendes autoregressives Modell (AR-Prozess) angepasst und in einem zweiten Schritt mit einem copula-basierenden autoregressiven Modell verglichen. Im Allgemeinen kann eine mehrdimensionale Verteilungsfunktion als Funktion eindimensionaler Randverteilungen mit Hilfe einer Copula ausgedrückt werden. Um ein copula-basiertes autoregressives Modell aufzubauen, wurde die analysierte Zeitreihe (modellierter Abfluss) in drei Vektoren transformiert, welche der ursprünglichen Zeitreihe entsprechen, aber zeitverschoben sind. Es wurde eine dreidimensionale Copula mit Hilfe ihrer Randverteilungen angepasst. Die empirische Copula hat gezeigt, dass ein Gauss-Modell geeignet ist, um die Interdependenz zwischen den Variablen abzubilden. Um die Ergebnisse des Copula-Modells vergleichen zu können, wurde ein nicht-parametrischer Ansatz zum Anpassen einer Copula-Funktion gewählt. Hierzu wurde eine „Beta-Kernel“-Copula angepasst, welche durch die Beta-Funktion und einen einzelnen freien Parameter ausgedrückt wird. Die untersuchte Abflusszeitreihe stammte aus einem in der Studie betrachteten Einzugsgebiet mit einer zeitlichen Auflösung von einem Tag. Da der gemessene Abfluss lediglich in einer monatlichen zeitlichen Auflösung vorlag, wurden für die Untersuchung die Ergebnisse aus dem HBV-Modell verwendet. Das Gauss-Modell und das „Beta-Kernel“-Modell zeigten sehr ähnliche Ergebnisse und stimmten mit dem modellierten Abfluss gut überein. Die größten Abweichungen waren bei hohen Abflusswerten oder Perioden mit einer hohen Abflussvariabilität zu erkennen. Im Vergleich zu den autoregressiven Prozessen hat das copula-basierte autoregressive Model bessere Ergebnisse gezeigt, wobei die Unterschiede nicht signifikant sind.

Chapter 1

Introduction

1.1 Understanding Unpredictability

Nature does not reveal itself in its full glory. This could be one of the first conclusions on a lecture about quantum physics at any University. And in fact, it is the message in one of the lectures introducing the uncertainty principle at the Punjab University National Centre for Physics (Modern Physics lectures). Suppose we have a system in which the initial conditions at a particular point in time are given and so are the equations governing the system. Classical Newtonian mechanics tells us that one is able to characterize the system at any instant of time. But in the last century this idea and the concept of determinism (which was previously strongly believed in) was replaced. This took place when the observation of some phenomena could not be understood by the knowledge of the time. New ideas were required, new laws were necessary, and ultimately a totally new picture of nature had to be developed in order to make an adequate description, a formal explanation from the mathematical point of view, of the unexplained phenomena. It was not until this that a deeper understanding was achieved when observation, computation and analysis of the consequences (Feynmann, about scientific method) was directed to the finer and finer components of matter. This advancement was made as many achievements in science are made, by taking in the world from a different point of view. This was the time when one of the two pillars of modern physics came into existence: Quantum mechanics, the physics of the very small. This totally new concept was destined to “just” change the perspective we understand the world.

What does this have to do with our study? It is directly connected to the core of the problem: Unpredictability. Knowing that the process of precipitation involves a range of different scales, departing from that of the very small, one can get closer to understanding this phenomenon through some ideas which arise from this new view of nature. Our understanding of the inherent uncertainty in the occurrence and internal variability of precipitation can be addressed from this point of view.

Uncertainty is inherent to the nature of the particles, and it is not *only* a consequence of the act of measuring (Heisenberg, from “Seis Genios en el Cielo”[73]). Heisenberg argued that the very act of observing a particle changes it, making precise knowledge impossible. From here one could face the fact that is is not the *consequence* (prediction), but the *premise* (present state) which fails. In the spirit of slightly clarifying the concept of unpredictability and to satisfy the perhaps already curious reader, a short example which relies on the Fourier Transform will be given in the next section. For more, [35], [85] and [59] can be reviewed.

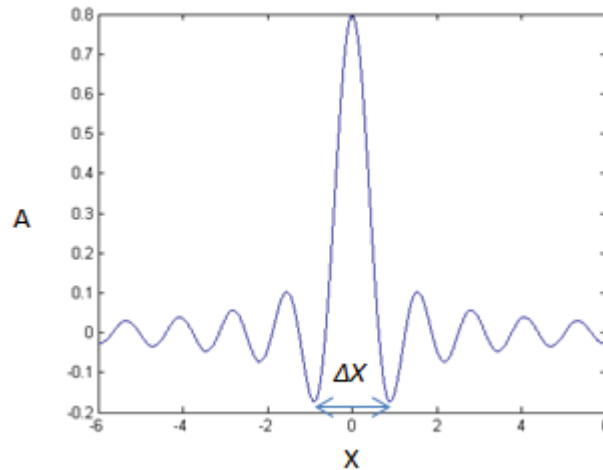


Figure 1.1: Wave Packet

1.2 Some Physical Ideas

In this section we will give a short but clarifying example of the uncertainty principle which is sure to help us to understand the problem introduced in the previous section. It is purely connected to modern physics, and of course with the deepest essence of our problem: Unpredictability.

We can model a particle (on the very small scale) as a wave packet, for example, a pulse. This matter wave has an associated amplitude and will occupy a certain region in space. This also means that we can associate a spread to the wave packet. An example can be seen in figure (1.1). Here, Δx holds for the associated spread of the matter wave. For our analysis we can consider the Fourier Transform (FT) and the inverse Fourier Transform of the Gaussian function (figure 1.2) where the abscissa in the FT is represented by the wave number k . It is known that this variable is related to the momentum of the particle according to the relation $p = \hbar k$ or $\Delta p = \hbar \Delta k$. We speak about the k – *space* or *momentum – space*. The constant \hbar is called the reduced Planck constant.

The two variables defined are actually conjugate variables and form the pair (k, x) . It is possible to demonstrate that if the spreading of one variable goes up, the other goes down, and vice versa. This is readily clear when constructing wave packs or matter waves from the Fourier Transform (FT). We can consider again figure (1.2) which constitutes an example where both the FT and the inverse FT are Gaussian waves.

Here the fact is clear that when the spread of the wave in the momentum or k – *space* decreases, the spread of the matter wave in the x – *space* increases, and vice versa. The two conjugate variables follow a reciprocal relationship of the form $\Delta k \Delta x \geq cte$, or equivalently $\Delta p \Delta x \geq cte$. Depending on the form of the matter wave, the relationship has a different limit. It can be demonstrated, however, that the ultimate limit is given by:

$$\Delta p \Delta x \geq \frac{\hbar}{2}$$

This means that the exact knowledge of a *state* of a particle is not possible and uncertainty is always present. Since the “origin” of precipitation involves several scales, between them the scale of the very small (photon interaction), the same concept is fully

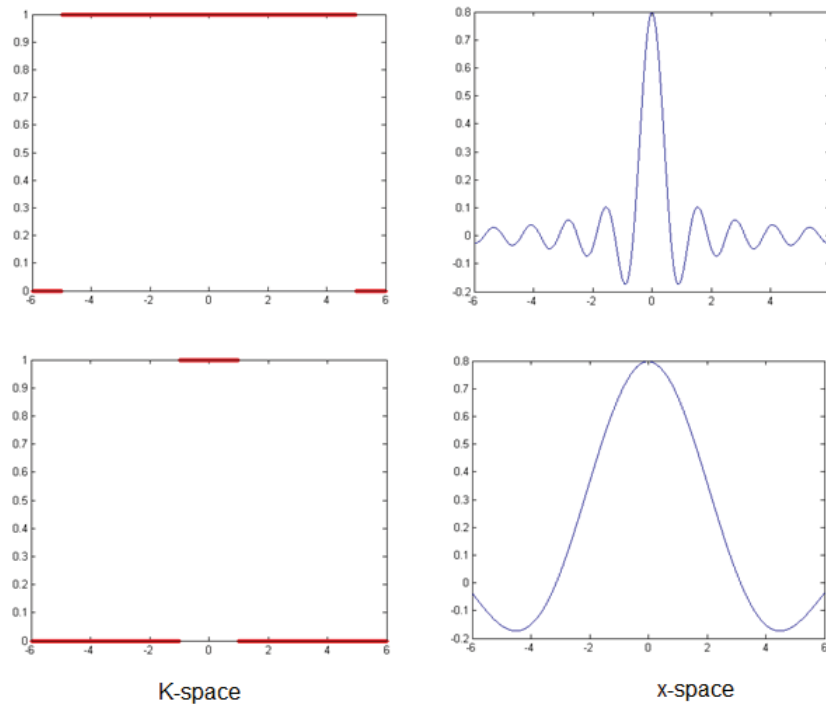


Figure 1.2: Fourier transform (blue line) and inverse Fourier Transform (red line) representing a matter wave

applied: inherent uncertainty, not only in the prediction (or future state) but also in the present description (present state) is present.

1.3 Objectives and PhD Thesis Structure

With a clearer picture of the difficulty when speaking about predictions, in the present research the issue of “knowing more” about some variables in the subsequent years is addressed. We investigate specially the future variations in precipitation, temperature and discharge in the area of research. The main goal is to estimate the impact of climate change in the hydrological variables which are important for Lima, especially the impact on the discharge.

It is known that, in general, the impact of climate change in the hydrological cycle at a regional scale may not be negligible (3). In this study, one of the most relevant foci is the quantification of this impact on the complex hydrological cycle of three catchments in Perú. A recent study (7) has estimated that at basin scale, Rimac, one of the catchments directly bound to Lima, present a downward significant trend for rainfall, considering 41 years of observations in the time interval 1964-2004.

To estimate the impact of climate change, different scenarios are considered, General Circulation Models (GCMs) outputs downscaled and two different models used to model the hydrological response of the areas under study in a climate-change scenario. The focus are the main catchments Chillón, Rimac and Lurin, and selected important sub catchments defined inside them as well as in the Mantro catchment (figures 2.1 and 4.16). Important from the view of water supply, water storage and energy generation.

The present PhD thesis is structured so that each chapter contains a theoretical introduction about the corresponding subject been treated therein. The general concepts

are explained in some detail followed by the necessary theoretical background about the methods and/or models used/performed in that particular chapter. Some applications and analysis are then done and conclusions drawn where it corresponds.

Chapter 2

Study Area and Data

The study area of this research involves all the areas which are “connected” to the capital of Perú, Lima. The term “connected” refers here to those areas from which the water resources of any particular region are drawn, partially or totally, used for irrigation, human consumption or energy generation in the city. It also refers to those projects which will contribute to water availability in the city once they are built. Separating the areas according to their location respect to the Andes range, they are:

- On the Pacific side: Chillón, Rimac and Lurín catchments
- On the Atlantic side: areas which comprises all 5 Marca projects: Marca I to Marca V, in Mantaro catchment

The highly variable topography (altitude) ranges from practically sea level in the city to near 5,000 m in the upper Andes. In the following figure (2.1) the regions on the Pacific and on the Atlantic side are depicted. A detailed delimitation of the areas corresponding to the Marca projects can be seen in figure (4.16) when dealing with downscaling.

The spatial precipitation pattern in the areas of Chillón, Rimac and Lurín is heterogeneous where rather higher precipitation is observed in the locations situated in the high altitudes than in those situated in the low altitude regions. An example of this heterogeneity is the rain stations of Aeropuerto and Milloc located in an altitude of $z = 13\text{ m}$ and $z = 4418\text{ m}$ respectively, which are depicted in figure (2.2). The separation between them is round $l = 115\text{ km}$ from each other. The yearly average precipitation in Aeropuerto is $p = 9\text{ mm/year}$, in contrast to Milloc where the same variable can reach almost $p = 900\text{ mm/year}$ for the period 1999-2008. This situation reflects the strong variation in the precipitation pattern.

The totality of the rain stations used in this study are presented in figure (2.2), where all of them are located inside the research area. The extent of the available information ranges from 1999 to 2008 in each of them with a time resolution of $\Delta t = 1\text{ day}$. This information was delivered from SENAMHI (National Service of Meteorology and Hydrology of Peru).

Regarding discharge information, observed time series exist in three different locations, whereby 3 subcatchments are defined and with which the subsequent analysis is performed. These three subcatchments are Santa Eulalia, Atarjea and Obrajillo (figure 4.10). The first two of them are located in Rimac catchment and the third one is located in Chillón catchment. The information was delivered from SENAMHI as well as from EDEGEL (EDEGEL is the largest electricity generating company of Peru). The time resolution

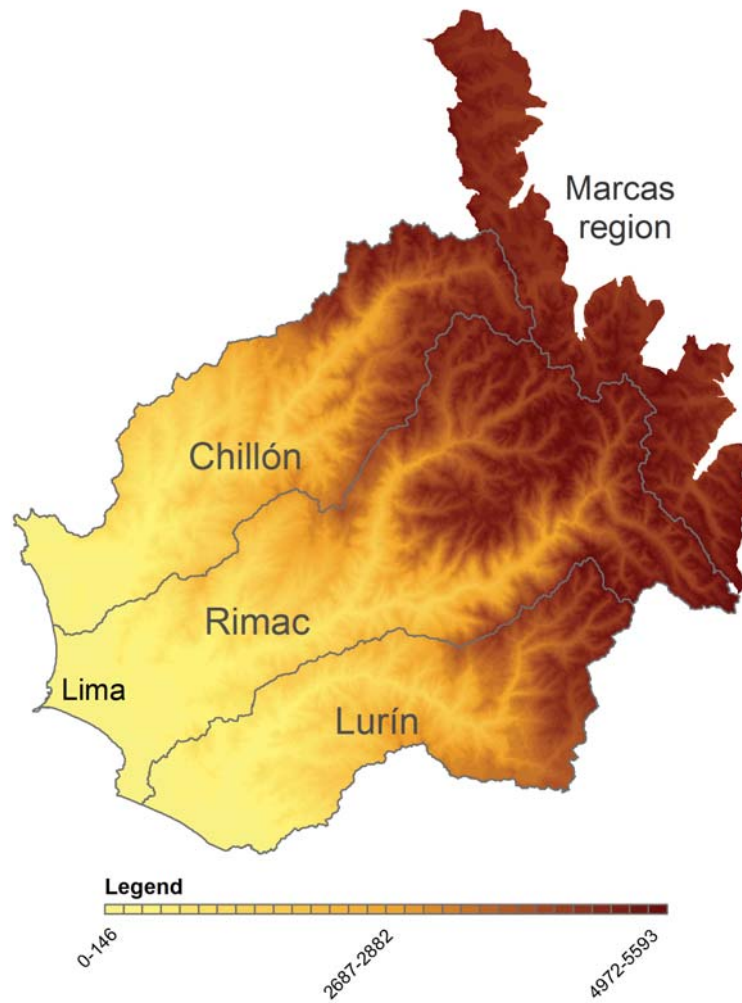


Figure 2.1: Study area: Chillón, Rimac and Lurín (pacific side) , Marcas region (upper part of Mantaro catchment)

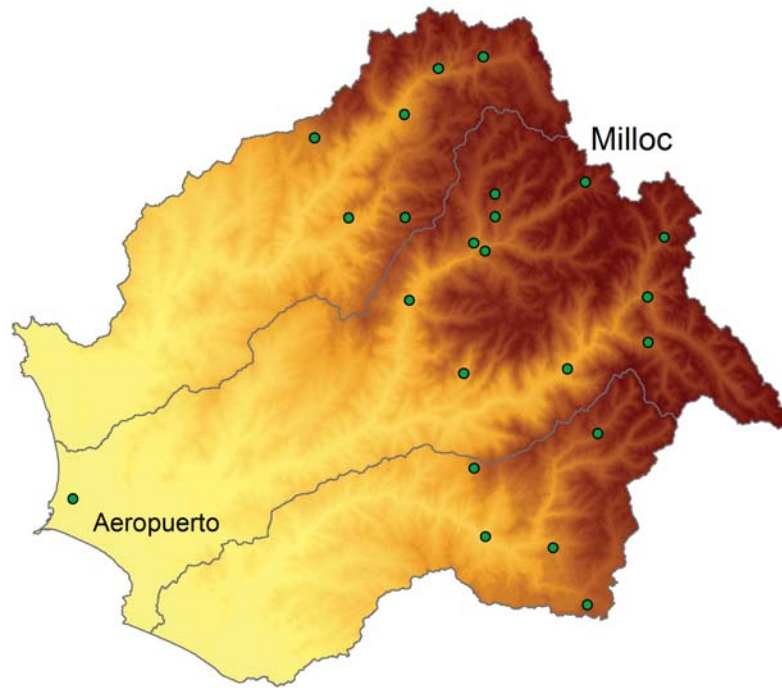


Figure 2.2: Rain stations used in the study

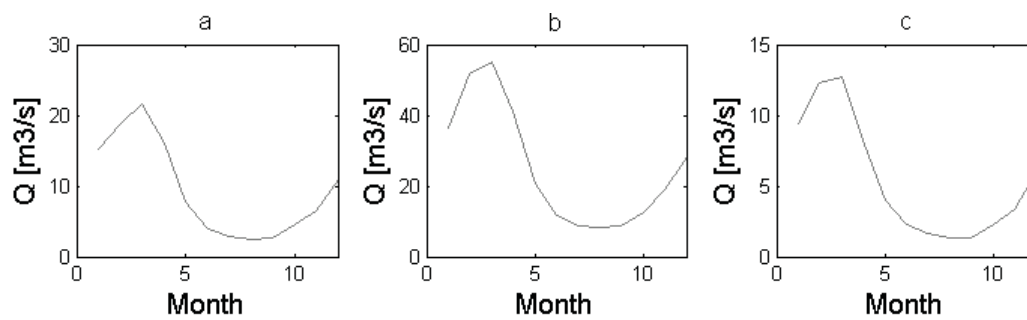


Figure 2.3: Natural discharge generated in the sub catchments Santa Eulalia (a), Atarjea (b) and Obrajillo (c)

for the observed discharge information is $\Delta t = 1$ month. The monthly average natural discharge in these three subcatchments are depicted in figure (2.3).

Temperature data are available for 6 different locations extending almost along the entire area and comprising both sides of the Andes. The annual cycle is more accentuated in the upper part of the area and the temperatures here are clearly lower than in the other areas. The stations together with their elevations and yearly averages are presented in table (2.1).

Table 2.1: Temperature stations used in this study. Period 1999-2008

Station	Elevation (masl)	Avg. Temperature (°C)
Humboldt	238	19.29
Nana	566	18.66
Matucana	2380	15.43
Yuracmayo	4300	7.66
Marca	4413	6.32
Milloc	4418	6.36

Chapter 3

Geostatistics and Regionalization Methods

Interpolation can be tackled in different ways. There are several methods of interpolation in the literature which have different degrees of difficulty, and their use depends on the requirement of each particular problem or application. In the present study external drift kriging is used and it will be explained in the next sections. First, a short review of geostatistics will be undertaken.

3.1 General Introduction

In this chapter we address the theory of regionalized variables. First of all, a review of geostatistics is done, in which some important concepts are presented. Subsequently theoretical considerations of the variogram and interpolation procedures, in which fundamentally some methods for the description of the spatial variability of some variable(s) with probabilistic and stochastic approaches are presented. Applications of the method of External Drift Kriging is also performed in different regions. This will permit us to have an estimation of some variable, in a delimited area and “arbitrary” resolution, and it will allow us to compute some of the basic necessary and required information for the subsequent rainfall-runoff modeling.

A very important issue is the error when applying some interpolation method, or in general, when performing some estimation or forecast. The associated error as well as the generated systematic error when estimating in unsampled locations is also analysed in this chapter. Finally, Some transformations are carried out and suggested for decreasing this two kind of errors.

3.2 A Little Review in Geostatistics

A historical development of geostatistics is not given here, but a description can be found for example in [47] and [9]. In geostatistics, one studies natural phenomena from the statistical point of view. In essence, geostatistics deals with the analysis of the spatial variability of a variable and how to model it. As natural phenomena are characterized by the distribution in space of one or more variables, an important definition is that of a regionalized variable, which can be viewed as a random variable (rv) distributed in space (one, two or three dimensional). It is a function of spatial coordinates

$$Z(\vec{X}) = Z(x, y, z)$$

where Z represents a random variable. Examples of a regionalized variable can be hydraulic head, transitivity, rainfall depths, etc. Another way of seeing a regionalized variable is as a realization of a random function. In our study, the considered variables are precipitation and temperature which we will deal with later in this chapter. When performing a geostatistical analysis, we can distinguish or divide the analysis into two different steps:

- Structural analysis of a regionalized variable: for example, how much correlation exists between the variable in two different points in the domain where the random field is defined
- Estimation problems: for example, estimation of the random variable at a point in the domain where observations are not available

An important concept which leads us to an essential hypothesis in geostatistics is that of stationary process. It is not only defined in this context but also in other areas. For example, a stationary process (or strictly stationary process) in time series is that process whose distribution remains invariant when a shift in time occurs. Analogously, this concept is defined in geostatistics in the same way, but considering not a displacement in time but in space. Mathematically this can be expressed as

$$P(Z(x_1) \leq d_1, \dots, Z(x_n) \leq d_n) = P(Z(x_1 + h) \leq d_1, \dots, Z(x_n + h) \leq d_n) \quad (3.1)$$

This can be thought of as a natural system which gives the same response under the same spatial configuration. This assumption of a strictly stationary process, or in other words the “strength” of this assumption, gives rise to some key definitions or hypotheses in geostatistics. One of them is the second order stationarity hypothesis, which is formulated through two conditions:

1. The expectation for the regionalized variable is constant over the domain of definition of the random variable
2. The covariance of two random variables is only dependent on the separation vector between their locations

Mathematically, if Z represent a random variable expressed as a function of location $Z = Z(x)$, then

$$E\{[Z(x+h) - m][Z(x) - m]\} = C(h)$$

where C is the covariance function.

An important consequence can be derived from the second condition regarding to the second moment of the random variable Z . If we consider a separation vector $h = 0$, the second condition says

$$E\{[Z(x) - m][Z(x) - m]\} = C(0) = Var[Z(x)]$$

This means the random variables have a constant second moment or variance independent of the location.

Another hypothesis in geostatistics is the intrinsic hypothesis, which leads to a fundamental function for the description of the variability of a regionalized variable: the variogram. This hypothesis reads

1. The expected value of the regionalized variable is constant over the entire domain
2. The variance of the difference in the value of the regionalized variable for two different locations depend only on the vector separating this two locations.

Mathematically we can write this hypothesis as follows:

$$E(Z(x)) = m \quad (3.2)$$

$$\frac{1}{2} \text{Var}[Z(x+h) - Z(x)] = \frac{1}{2} E\{[Z(x+h) - Z(x)]^2\} = \gamma(h) \quad (3.3)$$

with $x \in D$ (D , domain of definition of the variable x). The function $\gamma(h)$ is called the semivariogram or simply variogram. It is important to emphasize that the regionalized variable Z is defined in some domain in which each point is represented by a vector. This means that the variable h is also a vector. The equation (3.3) leads to the definition of the experimental variogram expressed as:

$$\gamma'(h) = \frac{1}{2n(h)} \sum_{x_i - x_j = h} [Z(x_i) - Z(x_j)]^2 \quad (3.4)$$

In the case of irregular-spaced observation points, the variogram can be constructed under the following considerations:

1. Two points x_i, x_j are considered to be separated a distance h if their separation is in the range $|h| - \epsilon \leq |x_i - x_j| \leq |h| + \epsilon$ for a given ϵ . This means, if

$$|x_i - x_j| \leq |h| \pm \epsilon \quad (3.5)$$

2. The angle formed by the vector $x_i - x_j$ is considered to be θ if this angle lies between $\theta - \delta$ and $\theta + \delta$ for a given value δ . Or equivalently, if

$$\text{Angle}(x_i - x_j) \leq \theta \pm \delta \quad (3.6)$$

with respect to the same reference.

As stated in these two considerations, both the value of separation ϵ and the angle δ has to be specified previously. The experimental variogram as well as the equations (3.5) and (3.6) require of course the knowledge of a minimum amount of information of the regionalized variable in some points of the domain. A less restrictive consideration for the calculation of the variogram is that which assumes isotropy, in which the restriction given by the equation (3.6) is not considered. This mean, one is interested only in the absolute value of the separation vector and not in its direction: For any $x_i, x_j \in D$,

$$\frac{1}{2} E\{[Z(x_i) - Z(x_j)]^2\} = \gamma(|x_i - x_j|) \quad (3.7)$$

irrespective of the direction of the vector $(x_i - x_j)$.

The previous paragraph and the equations (3.5), (3.6) and (3.7) require however a basic restriction or assumption as previously stated: they require information. To build a

variogram, a certain amount of observations corresponding to different points are necessary. What is more, for a selected distance h in the domain where a point of the variogram is desired to be calculated, the expected value in equation (3.3) has to be computed. This basic requirement for constructing a variogram is unfortunately not always met, and in fact, is not met in our case study and a theoretical variogram has to be used instead. However, in the case of having enough information, it is always advisable for theoretical reasons to fit a variogram [17], [19]. The usual way of doing this is by the method of least-squares and the method of maximum likelihood (David 1977, [44]). This leads us to the following point: in order to have a function suitable as a variogram for all points in the domain and not only, for example, for the points where observations are available, theoretical variograms have also been developed. In the present study a particular kind of variogram functions was considered, namely those which have the particularity that there exists a limit value L from which the function remains constant. This means

$$\gamma(|x_i - x_j|) = C \text{ if } |x_i - x_j| \geq L \quad (3.8)$$

if the second order stationarity is met and provided that after a certain distance L any pair of random variables are independent, we speak about variograms with a sill. In the case that only the intrinsic hypothesis is met we have variograms without a sill (See [4]). In the following, some typical theoretical variograms are presented.

Theoretical Variograms

Pure nugget effect variogram

$$\gamma(h) = C(1 - \delta(h))$$

$$\delta(h) = \begin{cases} 1 & h > 0 \\ 0 & h = 0 \end{cases}$$

Here $\delta(h)$ is the Kronecker delta function. This is the case where the random variables corresponding to different locations are independent. The variogram function is constant and equal to C everywhere except in the origin where take the value 0. In this point ($h = 0$), the function makes a jump and hence represent a point of discontinuity.

Exponential variogram

$$\gamma(h) = C \left[1 - \exp\left(-\frac{h}{l}\right) \right]$$

The exponential variogram depend on two parameters, the sill C and the parameter l . It can be seen that although the function (variogram) $\gamma(h)$ tend to C when $h \rightarrow \infty$ and therefore there is actually no particular distance from which the random variables are independent, an effective range of $3l$ can be defined from which the random variables separated by more than this distance can be considered as independent.

Spherical variogram

The spherical variogram depends also on two parameters, one of them the sill C and the other the range l , the latter being the distance from which two random variables separated by more than this distance are no longer correlated or dependent. The form of the spherical variogram is represented by the equation

$$\delta(h) = \begin{cases} C \left(\frac{3h}{2l} - \frac{1}{2} \frac{h^3}{l^3} \right) & h \leq l \\ C & \sim \end{cases} \quad (3.9)$$

The maximum value of the variogram corresponds to the point $h = l$ where $\gamma(h)$ takes the value C , the variance of the random variable.

Gaussian variogram

$$\gamma(h) = C \left[1 - \exp \left(-\frac{h^2}{l^2} \right) \right] \quad (3.10)$$

As we can see, the Gaussian variogram depends on two parameters, the variance of the random variable C and the parameter l . Analogous to the case of the exponential variogram, the function $\gamma(h)$ does not offer a particular limit from which two random variables separated by more than this distance can be considered as independent. It is defined however an effective range $h_l = l\sqrt{3}$.

If we draw the different theoretical variograms here presented we can see that the Gaussian variogram behave differently than the others in a vicinity of the origin. While the Gaussian variogram present a kind of quadratic curve in this vicinity, the others present a linear-like behavior. Also they present a different concavity, till a certain point is reached from where on all of them present the same concavity. If we consider the Gaussian variogram with parameters C and l (equation 3.10) and calculate the second derivative, it is possible to show by a simple analysis that the inflection point is given by

$$h' = l \frac{\sqrt{2}}{2}$$

from which the variograms present the same concavity.

3.3 Two Regionalization Methods

In the present section, an explanation and some derivations of two commonly used interpolation methods are given. A deeper insight into the methods can be gained for example in the textbooks [3], [4], [20] and [82].

Ordinary Kriging

Ordinary Kriging is one of the most used methods of interpolation. At first it can look rather difficult, but it can be easily implemented in a code. In the following, a derivation of the method is presented in order to understand its concept. An example will be given together with External Drift Kriging, which will be explained in the section 3.3.

Derivation of kriging

The general problem which we desire to be solved can be expressed as follows: Given n locations with coordinates (x_1, \dots, x_n) for which observation of a random variable have been made, we want to estimate the value of the random variable at a location x_0 without observation. We can then define:

$Z(x_i)$: Random variable with observation at location x_i
 $\hat{Z}(x_0)$: Random variable value to be estimated from the observations $Z(x_i)$, $\forall i = 1, \dots, n$

If we want to estimate our RV at a location x_0 as a linear combination of the observations we can then write

$$\hat{Z}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3.11)$$

The estimation of the random variable at a point without observation as a linear combination of the random variable at sampled locations (equation 3.11) simplified notoriously the estimation and further analysis. For estimating $\hat{Z}(x_0)$ we must then find the weights λ_i ($i = 1, \dots, n$) by some considerations. Here it is desired that the weights fulfill the following two properties:

1.- Unbiasedness

If we consider the error e as

$$e = Z(x_0) - \hat{Z}(x_0)$$

The unbiasedness condition impose that the expectation of the error between the random variable and the estimation is zero. This means,

$$E(e) = E[Z(x_0) - \hat{Z}(x_0)] = 0$$

Using either the second order stationarity or intrinsic hypothesis in this equation for the expectation we can then write

$$E[Z(x_0) - \hat{Z}(x_0)] = m \left(\sum_{i=1}^n \lambda_i - 1 \right) = 0$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (3.12)$$

Equation (3.12) already gives us a condition for the weights. This equation will also be a constraint in the minimization problem when minimizing the error variance in order to find the individual values of the weights.

2.- Minimum error variance

The error variance is defined as:

$$\sigma^2(x_0) = Var[Z(x_0) - \hat{Z}(x_0)] \quad (3.13)$$

For finding the variance we can express equation (3.13) differently. For this we can use either the covariance function or the variogram. This depends entirely on which kind of hypothesis is to be considered in our problem, namely the second order stationarity or intrinsic hypothesis. Using equation (3.12) and rearranging the terms in equation (3.13) we can write:

$$\sigma^2(x_0) = E \left\{ \left[Z(x_0) - \widehat{Z}(x_0) \right]^2 \right\}$$

$$\sigma^2(x_0) = E \left\{ Z^2(x_0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Z(x_i) Z(x_j) - 2 \sum_{i=1}^n \lambda_i Z(x_i) Z(x_0) \right\} \quad (3.14)$$

Now expressing the variance in terms of first the variogram and second in terms of the covariance function we can write:

- Intrinsic hypothesis:

$$\sigma^2(x_0) = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(|x_i - x_j|) + 2 \sum_{i=1}^n \lambda_i \gamma(|x_i - x_0|) \quad (3.15)$$

- Second order stationary hypothesis:

$$\sigma^2(x_0) = C(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(|x_i - x_j|) - 2 \sum_{i=1}^n \lambda_i C(|x_i - x_0|) \quad (3.16)$$

The goal is to have the smallest variance as a function of the values lambda. For this, the problem is transformed into an optimization problem where the function (3.13) is to be minimized. As the unbiasedness condition given by equation (3.12) is to be held for the error of the estimation to be zero (expected value), it represents a constraint of the optimization (minimization) problem. So the problem can be written in the following general form:

Given the function $f(\lambda) = \sigma^2(x_0)$ and the function $g(\lambda) = \sum \lambda_i$, the problem reads

$$\min f(\lambda_i)$$

$$\text{subject to } g(\lambda) - 1 = 0$$

The solution to this is given by introducing the Lagrange Multiplier and defining the new function

$$L(\lambda, \mu) = f(\lambda) + \mu [g(\lambda) - 1]$$

and the solution is found by solving

$$\nabla_{\lambda, \mu} L(\lambda, \mu) = 0 \quad (3.17)$$

where $\nabla_{\lambda, \mu}$ is the Nabla operator.

Replacing the function f by the error variance and solving (3.17), this means taking partial derivatives for each variable and equaling it to zero, we get the kriging system which depend either on the variogram or the covariance function:

- In terms of the variogram

$$\sum_{j=1}^n \lambda_j \gamma(x_i - x_j) + \mu = \gamma(x_i - x_0) \quad (3.18)$$

$$\sum_{i=1}^n \lambda_i = 1$$

- In terms of the covariance function

$$\sum_{j=1}^n \lambda_j C(x_i - x_j) - \mu = C(x_i - x_0)$$

$$\sum_{i=1}^n \lambda_i = 1$$

From which the coefficients lambda can be calculated and with this the estimation of the random variable (3.11) at the unsampled location. The associated variance can be calculated according to equation (3.13).

In order to simplify the calculation of the variance, let express equation (3.18) as follows:

$$\sum_{j=1}^n \lambda_j \gamma(x_i - x_j) = \gamma(x_i - x_0) - \mu \quad (3.19)$$

Now by replacing the right-hand side of equation (3.19) into equation (3.15), the error variance can be easier calculated as

$$\sigma^2(x_0) = \sum_{i=1}^n \lambda_i \gamma(x_i - x_0) + \mu$$

which is called the Kriging variance.

External Drift Kriging

In the conceptualization and derivation of the ordinary kriging method it is assumed that the random variable is stationary (or at least that the intrinsic hypothesis holds). However, in many cases it is possible to find situations where this condition is not satisfied and the random variable exhibits a trend or a systematic change. An example of this is the water table in an aquifer which can exhibit a global slope in the direction of the flow [64]. There are some methods which deal with this situation (see [19], [64], [21]), as for example:

- Universal Kriging
- Residual Kriging
- External Drift Kriging

In this section the method of External Drift Kriging will be explained which was introduced by Ahmed and deMarsily [1]. The condition for this method is the knowledge of an additional variable which is linearly related to the random variable under consideration. As a prerequisite, the new variable has to be known at any point in the domain where regionalization is to be performed.

Derivation of External Drift Kriging

As already pointed out in the previous section, the knowledge of an additional variable is necessary for the derivation and application of this method. Let's call this new variable $Y = Y(x)$ which depends upon the vector position x . Not any known variable can be a candidate for the purpose of the definition of the method, but only those which exhibit a linear relationship with the random variable. This means,

$$E[Z(x)/Y(x)] = a + bY(x) \quad (3.20)$$

Later we will see that this relationship may also contain a non linear relationship although in a non explicit way by introducing a new variable which is related to the variable Y . In the following this linear relationship given by equation (3.20) is supposed. A typical variable which is considered when interpolating precipitation is the topographic elevation, and it is used in this study.

The derivation of this method is quite similar to that of ordinary kriging from the previous section, but in the present case the new relation given by the equation (3.20) is incorporated in the analysis. In this way, given n sampled locations with coordinates (x_1, \dots, x_n) , the estimation of the outcome of the random variable at an unsampled location is a linear combination of the known random variables outcomes:

$$\hat{Z}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

The expectation of the error is again to be considered zero, as expressed by the unbiased condition. However, equation (3.20) implies a new relation and hence a new constraint for the optimization problem when minimizing the error variance.

By calculating the expectation of the random variable at the location x_0 without observation and using equation (3.20) we get the following relation:

$$\sum_{i=1}^n \lambda_i Y(x_i) = Y(x_0) \quad (3.21)$$

All in all, we have then two different constraints for the optimization problem where the estimation variance is to be minimized. Because we have two constraints, we define two Lagrange multipliers. The function to be minimized has then the form:

$$L(\lambda, \mu_1, \mu_2) = f(\lambda) + 2\mu_1 \left(\sum_{i=1}^n \lambda_i - 1 \right) + 2\mu_2 \left[\sum_{i=1}^n \lambda_i Y(x_i) - Y(x_0) \right]$$

where $f(\lambda)$ correspond to the error variance. The solution is found by solving

$$\nabla_{\lambda, \mu_1, \mu_2} L(\lambda, \mu_1, \mu_2) = 0 \quad (3.22)$$

This yield to $n+2$ equations for the $n+2$ unknowns. Expressing the solution as a function of the variogram we have:

$$\sum_{j=1}^n \lambda_j \gamma(x_i - x_j) + \mu_1 + \mu_2 Y(x_i) = \gamma(x_i - x_0)$$

$$\sum_{i=1}^n \lambda_i = 1$$

$$\sum_{i=1}^n \lambda_i Y(x_i) = Y(x_0)$$

From which we can find the weights Lambda which allows us to estimate the random variable at the location x_0 without observation.

Application: Comparison Between OK and EDK

So far we have derived two different methods of interpolation, namely Ordinary Kriging and External Drift Kriging. The equation systems for finding the vector parameter λ which optimize the estimation variance are very similar, but the restrictions and assumptions to be attained in each method have fundamental differences. In figure (3.1) an example of the implementation of Ordinary kriging and External drift Kriging is presented. Here the interpolated precipitation is depicted for the March 2nd, 1999, in the catchment Chillón. The selection of the date was arbitrary.

Both methods can reproduce some characteristics of the “real” precipitation, as is the case of the distribution of the maximum and minimum values. In fact, the highest precipitation is estimated to occur in the regions located in the upper parts of the catchment, and the lowest values are reached on the lowest regions. The evident difference is the distribution of precipitation throughout the entire area, in which the estimation using Ordinary Kriging seems to be fairly unrealistic.

The problem pointed out in the previous paragraph is a direct consequence of the applicability of the method. The assumptions and restrictions must then be checked carefully in order to avoid unwilling and wrong results. When considering a region where the assumption of constant expectation for the precipitation in the whole domain of definition is not true, the usage of ordinary kriging is not appropriate. In other words, when

$$E[Z(x)] \neq cte$$

An important issue concerns the additional information the method of External Drift Kriging requires for its derivation and application. An important question is how to find an optimal relationship in order to have a “better” estimation under some criterion. This issue will be addressed in the section (3.4) where numerical experiments will be performed using as basis the EDK method as interpolator.

Performing EDK in the Study Areas

From the last section it is clear that a crucial issue is to revise the different hypothesis and assumptions of a given method of interpolation before applying it in a certain region and

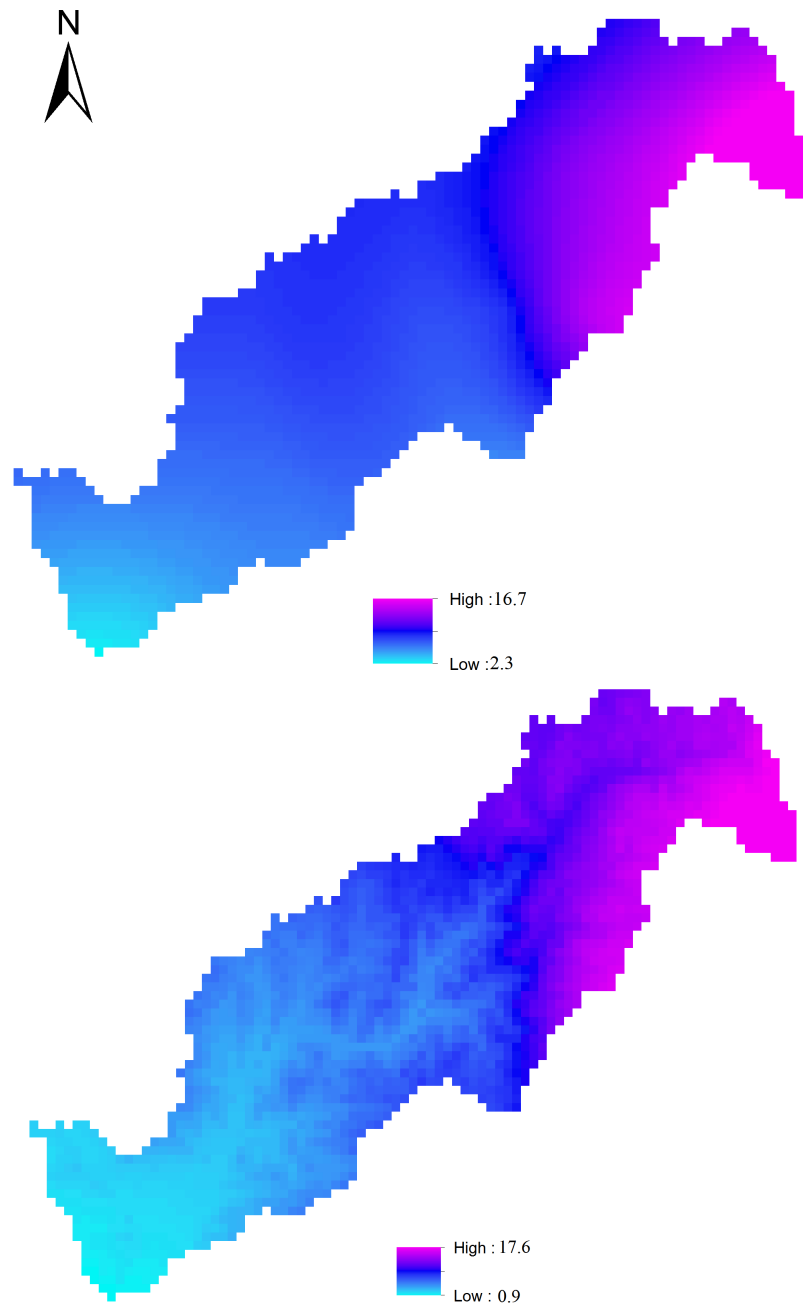


Figure 3.1: Interpolation results for the catchment Chillon using both OK (figure on the top) and EDK (figure on the bottom) for the 2nd March 1999

Table 3.1: Monthly average precipitation for the three main catchments Chillón, Rimac and Lurín

Month	Avg. Prec. [mm/yr]
Jan	61.94
Feb	82.44
Mar	84.62
Apr	29.13
May	5.36
June	2.05
July	2.11
Aug	2.87
Sept	7.26
Oct	17.60
Nov	20.49
Dec	48.94

with a data set at hand. As an example we implemented both OK and EDK in Chillón for a particular day (see figure 3.1), in which this conclusion is depicted quite clearly.

In this section we will show the implementation of External Drift Kriging in the areas of our study. The drift used here is the elevation. The kind of dependence introduced into the model (EDK) is at first a linear relationship.

By performing External Drift Kriging we calculate the spatial distribution of precipitation for the three main catchments. The time resolution for interpolation is one day. In order to also see the temporal variation in its pattern, the monthly average precipitation is also calculated. Figure (3.2) shows this monthly average precipitation for the entire area comprising the three catchments Chillón, Rimac and Lurín. In table (3.1) the averages values for each month are summarized.

If we take any particular month, the spatial distribution of precipitation can then be inferred clearly from figure (3.2). However, the degree of spatial variation can be rather different in dependence of the month in question. From the figures corresponding to the months December to April, the spatial distribution of precipitation is clearly recognizable. The precipitation generation is rather higher in the upper part of the region (north-east part of the area) been the rate of change much more accentuated in the month of March, where the maximum precipitation occurred. In contrast to this, the pattern is no more clear for the months May to September. For June and July, there is practically no difference in its distribution, where precipitation reaches its minimum value.

Another important analysis is the comparison of precipitation among the three main catchments, namely Chillón, Rimac and Lurín. Table (3.2) shows the calculated average precipitation for these three catchments, together with the individual percentage with respect to the total precipitation. The time period considered is 1999-2008. In figure (3.3) the monthly average precipitation for the three catchments and for the same period is depicted.

For the regionalization of precipitation in the catchment Chillón (figure 3.1), a linear relationship between the precipitation and the elevation was assumed, according to equation (3.20). Now the model is performed again but with a little modification. Instead of

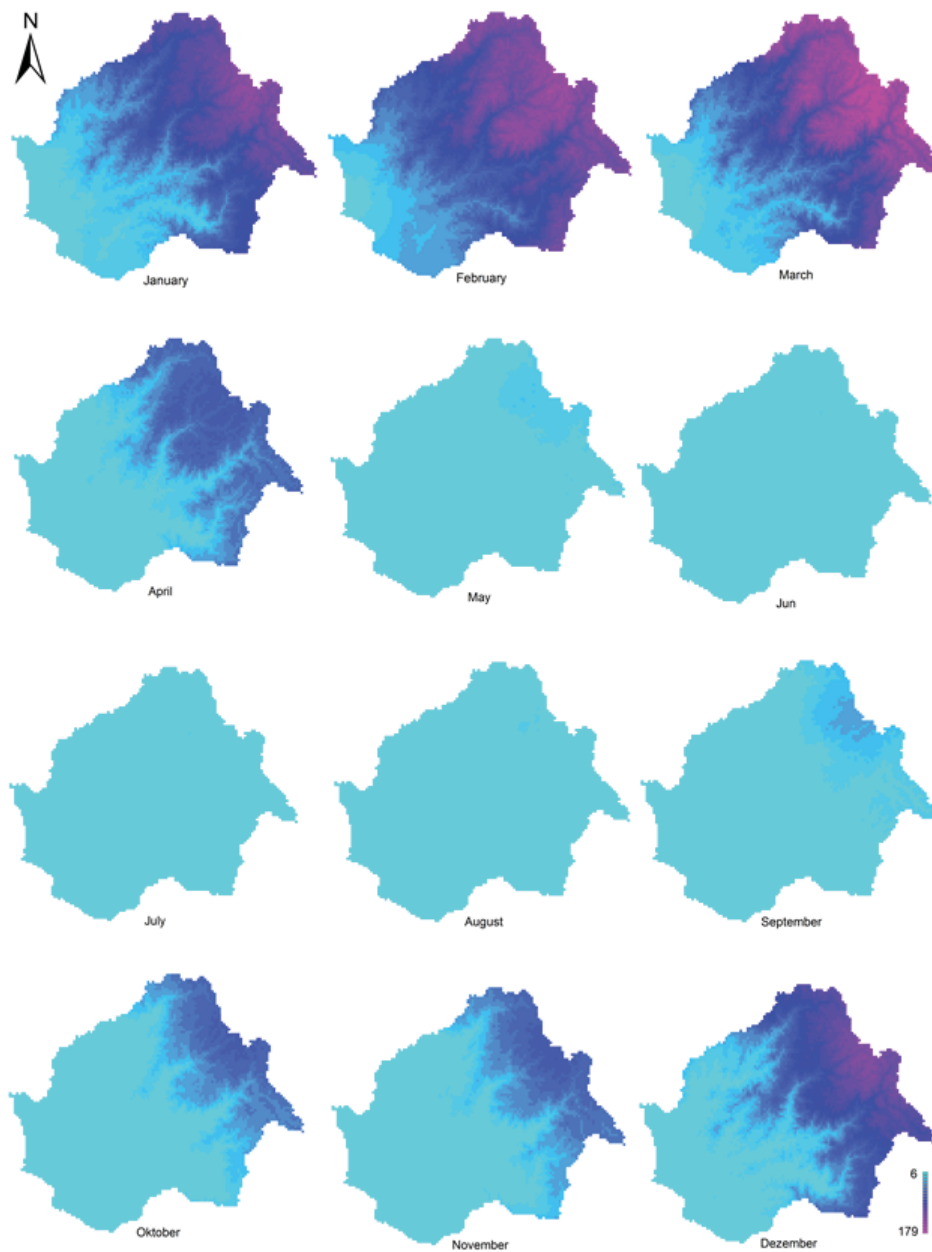


Figure 3.2: Estimated monthly average precipitation for Chillón, Rimac and Lurín for the time period 1999-2008

Table 3.2: Calculated yearly average precipitation for Chillón, Rimac and Lurín for the time period 1999-2008

Area	avg. Prec [mm]	%
Chillón	346.27	33.40
Rimac	430.79	41.55
Lurín	259.77	25.05

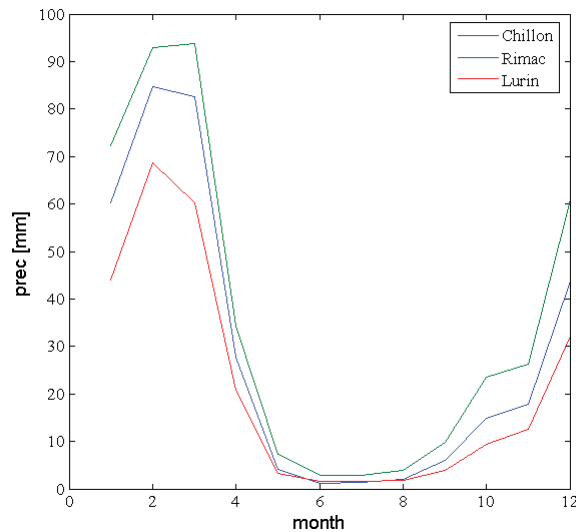


Figure 3.3: Monthly average precipitation for the individual catchments Chillon Rimac and Lurin. Time period 1999-2008

Table 3.3: Dependence between the estimated error and height for the points with observations (stations), considering both a linear and squared relationship given by the equations (3.20) and (3.23)

Relation	Correlation
Linear	-0.239
Power 1/2	-0.484

a linear relationship, a squared law between precipitation and elevation is adopted. This means

$$E[Z(x)/Y(x)] = a + b\sqrt{Y(x)} \quad (3.23)$$

In order to visualize the dependence between the cumulated error and elevation, the correlation between these quantities was computed considering all stations with observed data. The result is presented in the following table (3.3) where a comparison between these two different assumptions can be seen.

The result expressed in table (3.3) indicates the influence of the external drift on the structure of the error. If we view this as an indicator for the “goodness” or “appropriateness” of the relation *precipitation-external drift*, the linear relationship shows a better performance considering that the dependence error-elevation is less accentuated in comparison with the dependence given by the squared relationship.

The previous result leads us to the next section in which the efforts are concentrated in improving the interpolation by reducing the systematic error which may arise. The calculations will not only be based on modifying the kind of relationship for the drift which best estimates the distribution of the error with elevation, but also on modifying the characteristics of the region and/or the information (elevation). In the next section three different approaches are proposed.

3.4 Transformations

Basic Idea

So far we have described two different methods for regionalization of in general a random variable. The adequateness of each of them rely on the one hand on the spatial behavior of the random variable given by the conditions expressed in the last sections when defining a stationary process, either in its second order version or its intrinsic version, and on the other hand on a dependence between the random variable in question and an external variable or a drift. However, an analysis of the distribution of the error in space as for example in the *Z-coordinate* may become important when this error leaves the category of independent and a certain structure, let's say a trend, appears. This situation was tangentially outlined at the end of the last section when calculating the correlation of the cumulated error with the elevation as a measure of this dependency.

We can also write the relation between the random variable which has been analysed and the external variable or drift in the following way

$$E [Z(x) / f(x)] = f(x) + \varepsilon(x)$$

where the function f can be in general expressed as

$$f(x) = \sum_{i=0}^n c_k g_k(x)$$

and where the first component g_0 can be defined as 1.

The error ε is a random function which is supposed to have a zero mean and can be regarded as the stochastic part of the model. Hence, the function $f(x)$ can be regarded as the deterministic part of the model. This means we could eventually act on this function in order to improve the estimations. We can think of an undesired dependence between the error and some variable which arises when performing External drift Kriging. This issue will concern us and we will seek how to decrease the error in the estimation as well as the systematic error involved.

The analysis will be carried out by modifying the implementation of External drift Kriging done so far, consisting on the following three different approaches:

1. Smoothing
2. Power transformation
3. Rotation

Smoothing

Suppose we have a set of data which have a certain underlying structure and in general show an intrinsic non linearity. We want here to apply a model which represents the set of data adequately in some sense. We can estimate a smoothed component y_s (for example the z-component of a 3-dimensional vector) of a variable by defining different weights in terms of a kernel function and write (see for example [12])

$$\hat{s}(x) = \frac{\sum_{i=1}^n K_h(x - x_i) y_i}{\sum_{j=1}^n K_h(x - x_j)} \quad (3.24)$$

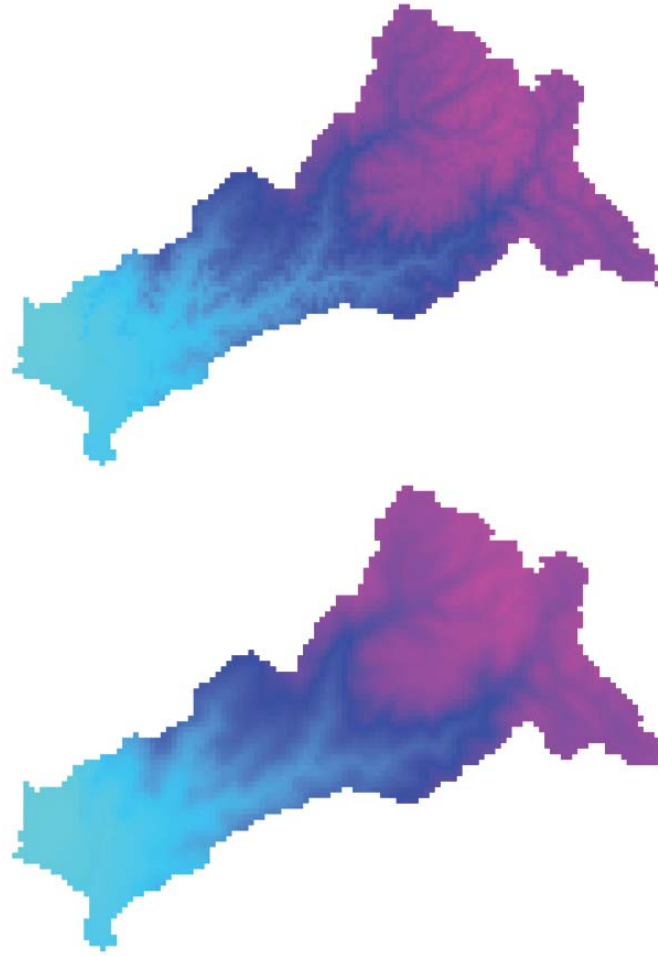


Figure 3.4: Interpolation result for the catchment Rimac, considering interpolation with real digital elevation model data (figure on the top) and considering smoothed data (figure on the bottom)

This estimator for y_s was first proposed by Nadaraya and Watson (1964) [56]. Here the Kernel function is to be chosen (see [81], [66]). In the present study, the Epanechnikov kernel was used. For the calculation, however, there is more than one possibility. Using a 1-dimensional kernel function in which the euclidean distance between two points is considered we can write

$$K(\nu) = c_d (1 - \nu^2) I_{\{\nu < 1\}}$$

with

$$\nu = \frac{\|x_0 - x_i\|}{R} \quad (3.25)$$

In figure (3.4) an example can be seen, in which regionalization is performed first without considering smoothing and second considering smoothing.

The effect of smoothing is clear from the previous picture. But as pointed out, an important issue is the dependence of the estimated error with elevation. Here we will formalize the idea in a little more detail. A cross validation is carried out with those points where information (measurements) are available. For any location with observation, let's say location k , the random variable is estimated at each time step, without considering

the specific information at that particular location. The associated error can be defined by the sum of the differences between the estimated and observed values. In this way, we can construct the vector error \vec{e}_v in which each argument correspond to the error in a specific location. We can then write

$$e_k = \sum_{j=1}^n [est_k(j) - obs_k(j)] \rightarrow \vec{e}_v = (e_1, \dots, e_m) \quad (3.26)$$

here

n : number of days

m : number of stations

e_k : error at station k

$est_k(j)$: Estimation at time step j and location k

$obs_k(j)$: Observation at time step j and location k

\vec{e}_v : Vector error

We can consider then the dependence between the error represented by the vector \vec{e}_v and the elevation by the correlation function. We write

$$\vec{z}_v = (z_1, \dots, z_m)$$

$$\rho = corr(\vec{e}_v, \vec{z}_v) \quad (3.27)$$

with \vec{z}_v vector elevation. For performing smoothing and cross validation we define a set $R_v = \{R_i\}_{i=1}^p$ which contains a given number of different radii in the range $1 \leq R \leq 10$ km. After performing cross-validation, figure (3.5) shows the correlation between the error and elevation. On the other hand Figure (3.6) shows the corresponding RMSE. We refer in this section to RMSE as the sum of the errors including all stations, ie.

$$e = \sum_{k=1}^m \left\{ \sum_{j=1}^n [est_k(j) - obs_k(j)]^2 \right\}^{\frac{1}{2}} \quad (3.28)$$

where n is the number of days and m the number of stations. The value of the correlation and error for the base case are presented in Table (3.4), which are the comparison points when performing our analysis, representing the case where no modification in both the information and the model have been performed. Analyzing the results, figure (3.5) shows clearly that the correlation between the error and elevation does not improve (decrease in absolute value) when performing smoothing disregarding the different values of R . On the contrary, the RMSE is reduced in all cases and a minimum value is attained for $R = 6$ with a value of $RMSE = 48.030$. These two facts exposed here make the results not wholly satisfactory since we are mainly interested in the reduction of the systematic error, though the decrease in the RMSE. This leads us also to the next points where other transformations are performed as well as some combination of them.

Table 3.4: Correlation between \vec{e}_v and \vec{z}_v , and overall error (eq. 3.28) representing the base case

Corr	RMSE
-0.239	48.297

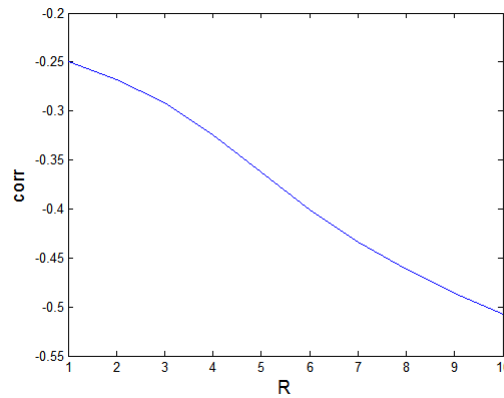


Figure 3.5: Correlation between the vector error \vec{e}_v and elevation for the range $1 \leq R \leq 10$ km

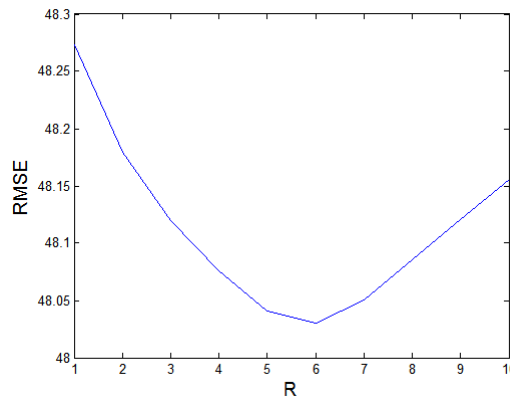


Figure 3.6: Cumulative error (eq. 3.28) for different values of the parameter R

Power Transformation

In the first point, when we smoothed the data in our attempt to decrease the systematic error when performing interpolation, the object of modification was the information itself. In this second approach, we will deal directly with the model. In the derivation of External Drift Kriging a linear relationship between the random variable to be regionalized and an external variable was supposed to exist. This means

$$E[Z(x)/y(x)] = a + by(x) \quad (3.29)$$

in which Z represents the random variable to be regionalized, y represents the drift and a and b are constants. Now we will suppose that the relationship between these two variables is not linear. Under this new assumption we can then write:

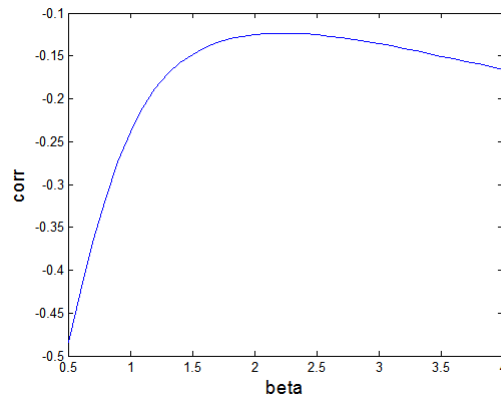


Figure 3.7: Correlation between the vector error \vec{e}_v and elevation \vec{z}_v for different exponents β (eq. 3.30)

$$E[Z(x)/y(x)] = a + by^\beta(x) \quad (3.30)$$

This means a power relationship is supposed to exist. For applying the model, a new variable is defined to hold linearity. We write,

$$y'(x) = y^\beta(x) \quad (3.31)$$

and the relationship (3.29) holds for the model, but this time between the conditional expectation of the random variable Z and y' . Since we do not know a priori which value of β fit best in the equation (3.30) in the sense of decreasing the systematic error (dependence error-elevation), we define a new set $\beta_v = \{\beta_i\}_{i=1}^n$ in which the exponent β is varied from a minimum of 1 and a maximum of 4. Here the sub index v holds again for designating a vector. Performing External Drift Kriging we calculate first the correlation between the error and elevation and second the error for each location. The results are presented in figure (3.7) which shows the exponent beta and the associated correlation and figure (3.8) which shows the error for different values of β .

In contrast to the case of smoothing, in figure (3.7) we do observe a decrease in the systematic error where the correlation changed importantly from a value $corr = -0.484$ corresponding to $\beta = 0.5$, to $corr = -0.166$ for $\beta = 4$. An optimum (minimum in absolute value) correlation coefficient is however attained for $\beta = 2.2$ with a value of $corr = -0.124$, which is much smaller in absolute value than the correlation given by the base case. The RMSE, on the other hand, shows always an increasing pattern which varies between $RMSE = 47.950$ and $RMSE = 49.306$. The improvement of the RMSE with respect to the base case is shown for the exponents in the interval $0.5 \leq \beta < 1$.

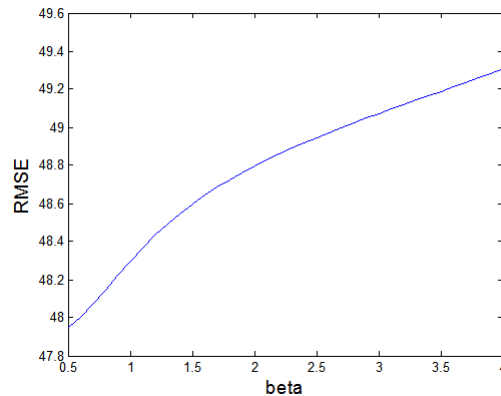


Figure 3.8: Overall error (eq. 3.28) for a set of different exponents β

Smoothing and power transformation

So far we have decreased the dependence error-elevation especially when defining a new variable and hence a different relationship between the random variable and the drift was shown to be more adequate for the purpose of reducing the systematic error. Now the analysis is turned into a combination of both smoothing and power transformation. For this we can define a set containing different ordered pairs as

$$T_v = \{(R_i, \beta_j) / i = 1, \dots, n_R; j = 1, \dots, n_\beta\} \quad (3.32)$$

in which R_i and β_j are varied between $1 \leq R_i \leq 10$ km and $1 \leq \beta_j \leq 4$ respectively. The spacing between consecutive values for each component is $\Delta R = 1$ km and $\Delta \beta = 0.1$. Previous analysis shown that for values of n_R and n_β larger than 10 and 4 respectively, the systematic error is much larger than that of the base case. With these constraints and definitions we first restrict the calculations to the defined boundaries and then perform cross validation for each ordered pair (R_i, β_j) in T_v . The results are presented in figure (3.9) namely the correlation error-elevation and RMSE.

By analyzing the correlation we observed that for each value of R there is a range of $beta$ values where the absolute value of the correlation is reduced. More precisely, up to a minimum value of $beta$ the correlation is nearer to zero than the corresponding correlation given by the base case, depending on R . In the maximum decrease of the absolute value, the correlation changes from $corr = -0.239$ (base case) to $corr = -0.106$ corresponding to the point $(R_{optimum}, \beta_{optimum}) = (3, 2.6)$ constituting a very important reduction.

Concentrating our attention now on the RMSE, figure (3.9) (bottom) shows that the larger the exponent β for a given R , the higher the estimated RMSE. There is a range for β between $1 \leq \beta \leq 1.5$ in which the RMSE is smaller than the corresponding value for the base case. The overall minimum for the error is reached in the point $(R_{optimum}, \beta_{optimum}) = (6, 1)$ with a value of $RMSE = 48.031$.

It is important to note that smoothing has some effect in the correlation in the sense of decreasing it respect to the base case in combination with a varying exponent, in contrast to the first case where only smoothing was performed and no improvement was achieved.

Finally, we perform the same analysis for the months December to April, corresponding to those months where precipitation is clearly higher than the rest of the year (rainy

season). The results are depicted in figure (3.10). In this case, the modulus of the correlation decreases considerably for most of the points (R_i, β_j) considered, where only for a small values of beta and high values of R the correlation does not improve. The minimum attains the value $corr = -0.029$ in the point $(R_m, \beta_m) = (3, 2.1)$.

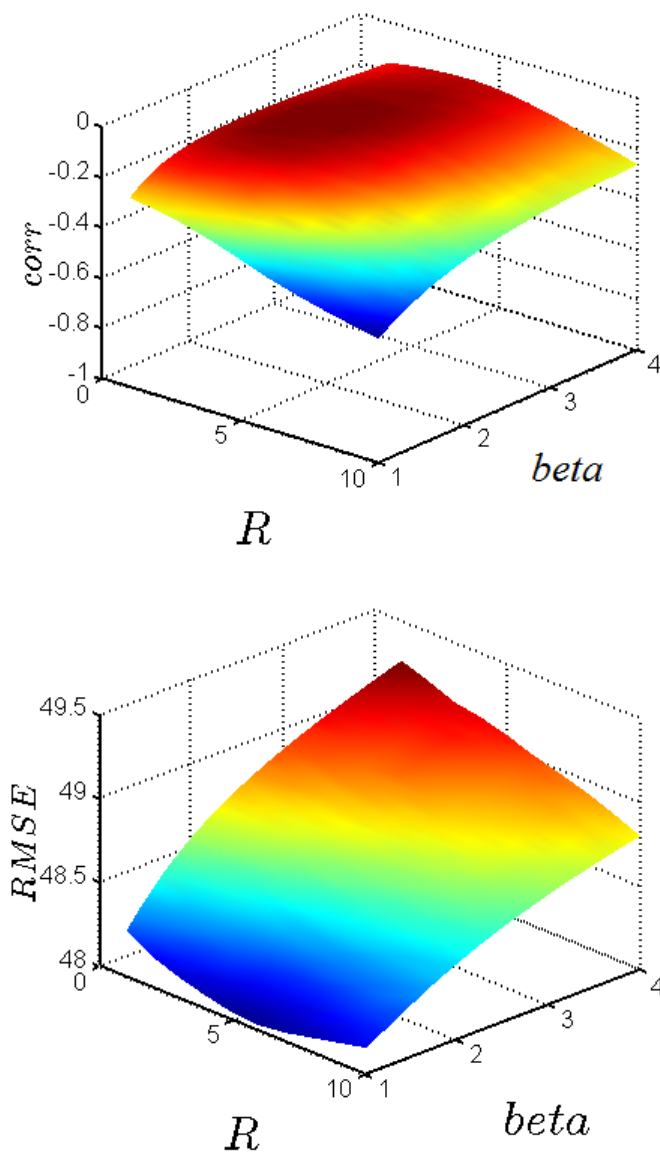


Figure 3.9: Correlation between the vector error \vec{e}_v and elevation \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents β and radius R

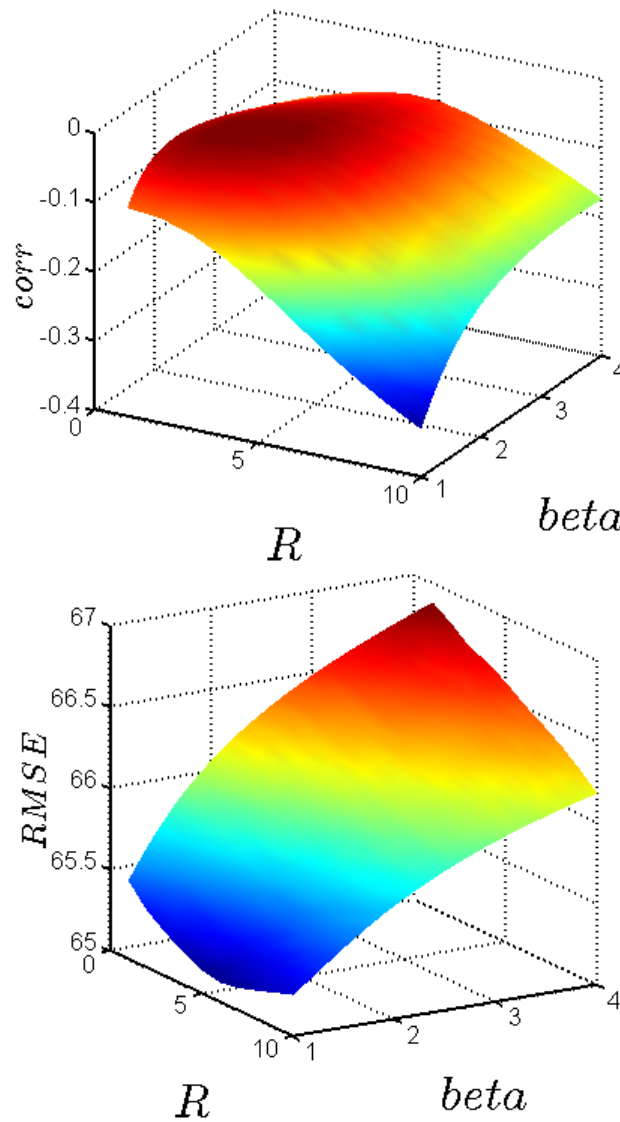


Figure 3.10: Correlation between the vector error \vec{e}_v and elevation \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents β and radius R , months December to April

Rotation

In the first attempt to improve regionalization in the sense of decreasing undesired systematic errors which could arise in the estimation of regionalization, we smoothed the Z-component of the DGM before applying External Drift Kriging. For this a circumference was defined with a defined radius. This means the only criterion for choosing a point was that it had to be inside this circumference. In this section our attention is turned into the definition of a non symmetric area for smoothing purposes, as for example a semicircle or a rectangle. The selection of the orientation of this area is the fundamental issue here. For each regionalization performance we position the area in a defined angle φ which is modified (rotated) in each iteration. By doing so we obtain a set of different values for the correlation and cum RMSE. An alternative way of seeing this is to think of the area as fixed in space and the DGM (digital elevation model) or the points to be analyzed is/are rotated. If we call the (x, y) the original point an (x', y') the

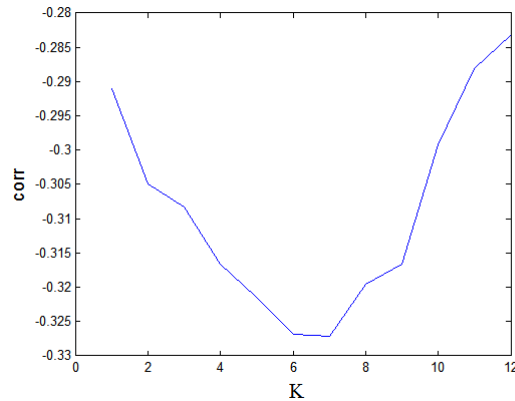


Figure 3.11: Correlation between the vector error \vec{e}_v and vector elevation \vec{z}_v for a set of different K -parameters according to the set (3.35)

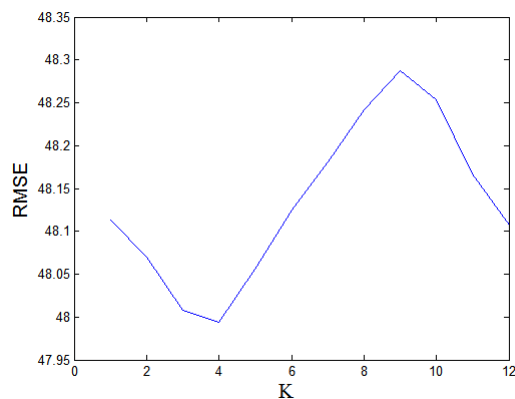


Figure 3.12: Overall error (eq. 3.28) for a set of different K -parameters according to the set (3.35)

transformed (rotated) point in an angle phi, we can relate them by the transformation matrix

$$M_T = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \quad (3.33)$$

so that

$$\begin{pmatrix} x' & y' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \quad (3.34)$$

In this way EDK is performed and for each iteration the angle of rotation is changed in $\Delta\varphi = \pi/6 \text{ rad}$ with respect to the previous step. In this way and in a similar fashion as in the case of smoothing and power transformation we defined a set

$$\Phi_v = \left\{ \varphi_i/\varphi_i = k \cdot \frac{\pi}{6}, k = 1, \dots, 12 \right\} \quad (3.35)$$

containing different angles or rotation ranging from $0 \leq \varphi \leq 2\pi$. In this way we calculate the correlation and RMSE considering each value in the set. In figure (3.11) and (3.12) the results are presented.

Comparing the results with the base case, no improvement can be observed in the correlation value for the defined set Φ_v . The RMSE improved however for each different

rotation, but the changes are quite slight with a maximum change from $RMSE = 48.297$ to $RMSE = 47.993$. In the next point, we will center our attention on the combination of power transformation and rotation.

Power Transformation and Rotation

In this point we will combine the modifications introduced in the last point, together with those introduced in the section dealing with power transformation. For this purpose we define a new set which has the form

$$S_v = \{(\varphi_i, \beta_j) / i = 1, \dots, n_\varphi; j = 1, \dots, n_\beta\} \quad (3.36)$$

in which we vary both the angle of rotation and the exponent. As in the last point, we express the angle φ_i as $\varphi_i = k \cdot \frac{\pi}{6}$, $k = 1, \dots, 12$. The difference between two consecutive values for the angle and the exponent is $\Delta\varphi = \frac{\pi}{6}$ and $\Delta\beta = 0.1$ respectively. The boundaries for the angle of rotation are clear since the results have a cycle 2π rad after which they repeat themselves. Regarding the boundaries for the exponent β , the same range is kept here as that used in the set T_v . We then perform cross validation for each point in the set S_v . The results are summarized in figure (3.13).

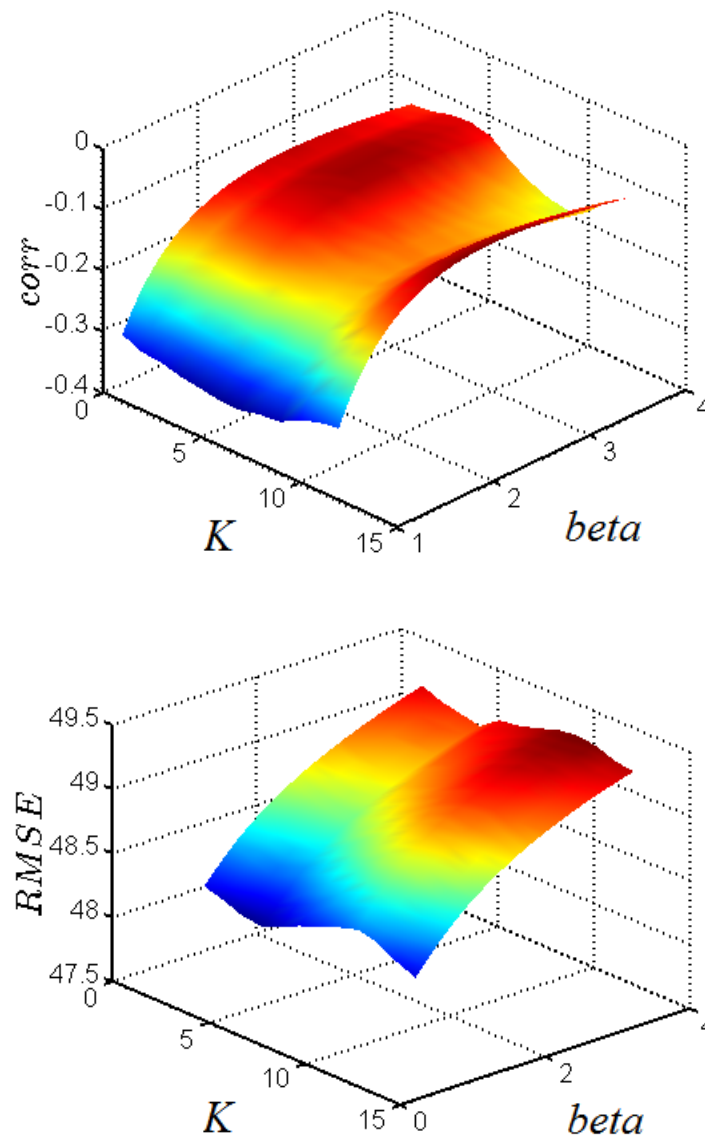


Figure 3.13: Correlation between \vec{e}_v and \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents $beta$ and K -parameters

From the figure we see that for each value of $beta$, the modulus of the correlation decreased with the exception of the lowest exponents $beta = 1$, $beta = 1.1$ and $beta = 1.2$. The maximum reduction is attained for the point $(\beta, \varphi) = (2.6, 11\pi/6)$ with a value of $corr = -0.10$. On the other hand, the RMSE (figure 3.13 bottom) shows a clear tendency of the error in dependence with β . For a given rotation, the error increases systematically and follows a curve with a slight curvature. The values vary from $RMSE = 47.994$ representing the minimum to $RMSE = 49.237$. Note that the minimum corresponds to an improvement with respect to the base case.

As performed in the case of smoothing and power transformation, the same analysis was carried out for the rainy season. The absolute value of the correlation also decreased notably for all the points considered with a minimum value of $corr = -0.023$. Results are presented in figure 3.14

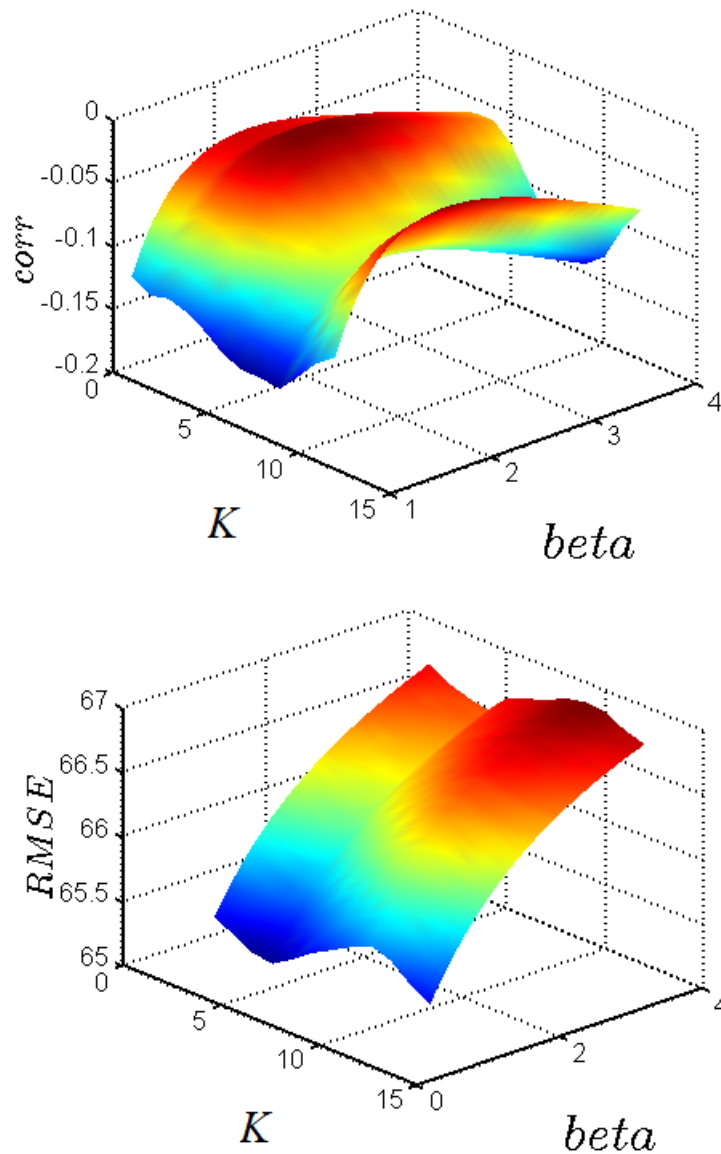


Figure 3.14: Correlation between \vec{e}_v and \vec{z}_v (upper figure), and overall error (lower figure) for a set of different exponents β and K -parameters for the rainy season

3.5 Discussion of the Results

In this chapter we gave some theoretical background about regionalized variables framed in the topic of geostatistics. For estimating precipitation in unsampled locations we performed EDK which conduced us to a model for the spatial pattern of this variable. The regionalized precipitation forms the basic information for the subsequent rainfall-runoff modeling. Together with this, the overall error as well as the systematic error was estimated, analyzed and reduced. For this purpose, three different approaches were implemented, namely the smoothing of the digital elevation model, a transformation in the drift where a power law was considered instead of a linear relationship, and the definition of a non symetric area which was allowed to rotate in a predefined manner. Here important results are obtained especially when combining the different methods. When considering a power transformation law, the systematic error was importantly reduced

from $corr = 0.239$ to $corr = 0.124$ in absolute value. A reduction in the total error was also observed which changed from $RMSE = 48.03$ to $RMSE = 47.95$, although they do not differ too much from the base case. By combining the methods, the systematic error was reduced further with optimum values $corr = 0.106$ and $corr = 0.100$ in the case of smoothing and power transformation, and a power transformation and rotation, respectively. This last optimum value constitutes a very important improvement where the $corr$ decreases more than a half of the original value. The $RMSE$ also improved (decreased) but in all cases the changes were small. A mentionable result came out when analyzing what happens with the correlation in a rainy-season scenario. Here it is observed that the systematic error drops to $corr = 0.023$ in absolute value, which can be considered practically zero.

Chapter 4

Downscaling

4.1 General

Downscaling is a term used when referring to a change of scale (from a coarser to a finer scale). When information or some model output is required in a finer scale, a downscaling procedure can be used for this purpose.

In order to assess the expected behavior of precipitation and temperature for the future, the use of the outcomes of General Circulation Models can not be made directly for hydrological climate change impact studies. The reason for this is the scale mismatch ([10]) and some adjustments have to be performed.

Downscaling techniques can be divided into two main groups, namely dynamic downscaling and statistical downscaling. The first of them is based on dynamic models in a scale which is finer than that of the GCM. The second approach is based on statistical relationships in which large-scale information is linked to local variables (Predictor/predictand) ([11]). In this study, the second approach is used and specifically a Quantil-Quantil downscaling is performed. In this way, information is transformed from GCM large-scale to local-scale given by each of the study areas.

4.2 Quantil Quantil Transformation

In the present study a Q-Q transformation is used. It can generally be applied to any meteorological parameter [76]. Q-Q transformation finds also application for example in bias correction for Regional Climate Models (RCM), because it is known that precipitation simulations by using RCM are biased [61].

For performing a Quantil-Quantil downscaling (Q-Q transformation) the following information is available:

- Time series from GCM 's outcomes for the variable to be downscaled, corresponding to the future assessment
- Time series from the GCM 's outcomes for the variable to be downscaled, corresponding to the control period
- Time series of observed data

If we consider the variable to be downscaled $v(t)$, lets call the forecast of the the GCM $v_m(t)$, the control period of the GCM $v_{cp}(t)$, the observation time series $v_{obs}(t)$ and the downscaled result $v_d(t)$.

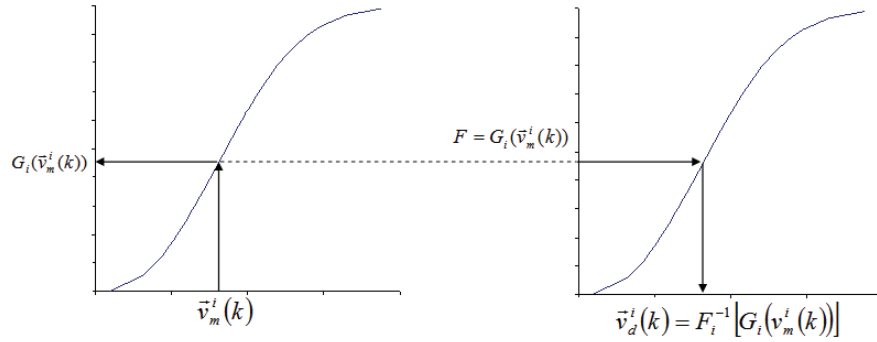


Figure 4.1: Downscaling procedure. An input $\vec{v}_m^i(k)$ is evaluated into $G(\cdot)$, matched to $F(\cdot)$ and the downscaled value obtained from $F_i^{-1}(\cdot)$

One important issue is that the time resolution for the downscaling performance or the time step to be considered is completely determined by the resolution of the GCM, corresponding to $\Delta t = 1$ month. Under this restriction, the observation time series $v_{obs}(t)$ as well as the downscaled time series $v_d(t)$ will have to have necessarily the same temporal resolution.

Under this condition, the Q-Q transformation requires vectors to be built for each month of the year. For this, each time series have to be split according to each month, resulting in 12 vectors of data from January to December. We can define:

- \vec{v}_m^i : Vector variable corresponding to the GCM forecast for the month i , $i = 1, \dots, 12$
- \vec{v}_{cp}^i : Vector variable corresponding to the GCM control period for the month i , $i = 1, \dots, 12$
- \vec{v}_{obs}^i : Vector variable corresponding to the observations for the month i , $i = 1, \dots, 12$
- \vec{v}_d^i : Vector variable corresponding to the downscaling results for the month i , $i = 1, \dots, 12$

Q-Q transformation is based on the distribution functions of the vectors \vec{v}_k^i . In the procedure, for each value in \vec{v}_m^i the cumulative distribution function (cdf) of \vec{v}_{cp}^i corresponding to the control period is evaluated and matched to the cdf of \vec{v}_{obs}^i . The transformed or downscaled value \vec{v}_d^i is then estimated. Mathematically we can write

$$\vec{v}_d^i(k) = F_i^{-1} \{ G_i [\vec{v}_m^i(k)] \} \quad (4.1)$$

in which

$G_i(\cdot)$: Cumulative distribution function for the control period vector \vec{v}_{cp}^i , $i=1, \dots, 12$

F_i^{-1} : Inverse cumulative distribution function for the observation vector \vec{v}_{obs}^i

$\vec{v}_d^i(\cdot)$: Downscaled value

Graphically, consider the following figure (4.1) where the downscaling procedure is explained. Here a distribution function for the control period and for the observations was drawn.

4.3 Parametric and Non-Parametric Fit

As already expressed in the last point, before performing downscaling a distribution function has to be fitted to each of the data sets. For this purpose we can mention here two different approaches which were used in the present study:

1. Parametric approach
2. Non parametric approach

In the next point we will introduce the theoretical distributions used, the method of fitting as well as the results for the parametric case. Subsequently we will present the nonparametric approach with the corresponding kernel function used, the selection of the optimum free parameter and results.

Parametric Approach

The parametric approach supposes the data are drawn from one of the existing parametric family of distributions. Here one of these theoretical cumulative distribution functions was fitted to the data, namely precipitation and temperature. Since the temporal resolution is given from the GCM resolution, monthly average precipitation was considered. The cdf used is the Weibull distribution which depends on two parameters, namely the shape and scale parameter. If we have a sample , the theoretical weibull cdf is given by the following equation:

$$f(x; k, \lambda) = \left(\frac{k}{\lambda}\right) \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \quad (4.2)$$

where

- k : Shape parameter
- λ : Scale parameter

Its corresponding cumulative distribution function is given by

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \quad (4.3)$$

As well known, fitting distributions, which means finding the optimum parameters in order the theoretical cdf to fit best the sample, can be accomplished by means of the method of moments or method of maximum likelihood. Here the latter approach was adopted. The likelihood function is defined as

$$L(\vec{\vartheta}/x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) \quad (4.4)$$

Where $\vec{\vartheta}$ is the parameter vector to be found. Equivalently and for convenience it is also sometimes worthwhile to work with the log-likelihood function, considering that several distribution functions follow an exponential law. In this way

$$l = \ln L(\vec{\vartheta}/x_1, \dots, x_n) = \ln \prod_{i=1}^n f(x_i) = \sum_{i=1}^n \ln [f(x_i)] \quad (4.5)$$

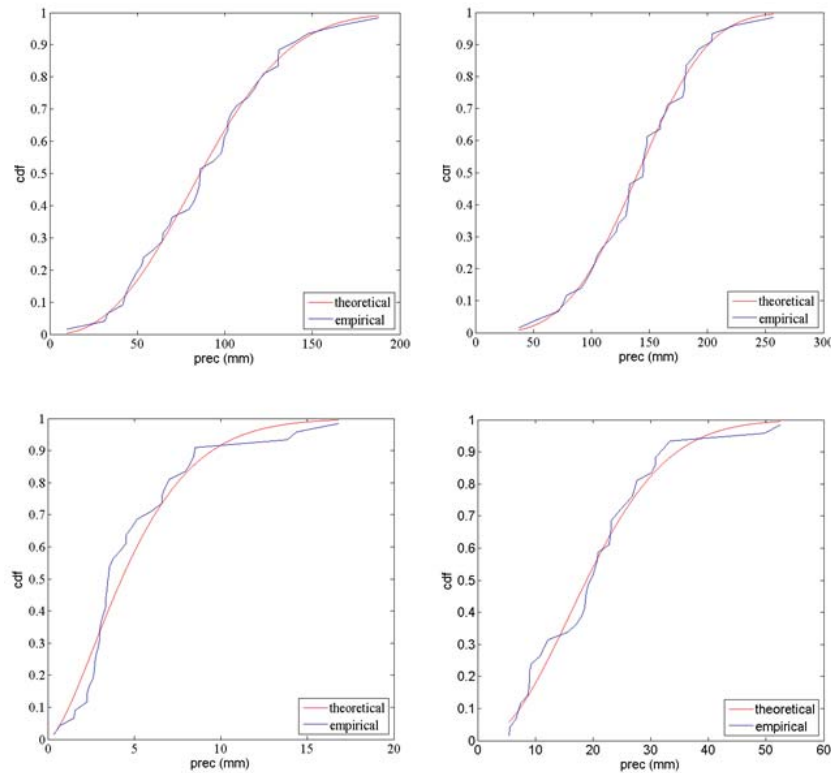


Figure 4.2: Theoretical and empirical fitted distributions for the months December (upper left figure), march (upper right figure), June (lower left figure) and September (lower right figure) for the sub catchment Obrajillo, representing the begin of the seasons summer, autumn, winter and spring respectively

it defined where the function f is the considered or selected density function and n the dimensionality of the sample vector \vec{x} . One can also compute the empirical cumulative distribution function of the sample \vec{x} , in order for example to compare it with the fitted distribution. It is defined as follows:

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{x_i \leq t\}}$$

By using the log-likelihood function (4.5), the Weibull distribution was fitted to the vectors \vec{v}_{cp}^i and \vec{v}_{obs}^i defined previously, for $i = 1, \dots, 12$. As an example, figure (4.2) shows the empirical distribution function together with the fitted theoretical distribution function for four months: December, March, June and September representing the beginning of the seasons (in the city of Lima) summer, autumn, winter and spring, respectively.

Non-Parametric Approach

In this approach, the construction of the density estimates relies solely on the data. This means that a particular function f is not assumed and with this no parameter o family of parameters has to be estimated from the observations in order to fit a theoretical distribution function. In the following, the kernel method as a non parametric methodology will be explained. As a complement to the method, special considerations will be adressed to deal with different kind of data, as for example when the random variable can take values

only in a constrained subset of the real line. A methodology based on Fourier coefficients (Orthogonal series density estimation) will also be addressed.

The kernel method

The second approach used in this study for fitting a distribution function to the data was performed by means of a Kernel function. Here no assumption for the data falling in a given family distribution is made.

If f is an unknown density function provided a sample $\vec{x} = (x_1, \dots, x_n)$, the Kernel density estimator of f is given by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (4.6)$$

or equivalently:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) \quad (4.7)$$

in which

$K(\cdot)$: Kernel function

h : Bandwidth or smoothing parameter

The kernel function K is non negative and satisfies the condition (in order to be a density function)

$$\int_{-\infty}^{\infty} K(x) dx = 1$$

The only parameter of the function, the free parameter h , plays an essential role when analyzing the shape of f . Depending on the value of h , important information carried by the data can be ignored as in the case when h is too large, or overemphasized when h is too small. In simple terms, the free parameter h controls the width of the kernel function used and the degree of smoothing which is applied to the data.

There are some kernel functions in the literature which can be used. The choice of which of them should be consider is more a consideration of computational efforts than in the error or the mean integrated square error MISE [66] obtained in the fitting procedure explained below. In the following an example of three different kernel functions is given, namely Epanechnikov, Gaussian and Triangular kernel.

- Epanechnikov

$$K(x) = c_d \max\left(1 - \frac{x^2}{d^2}, 0\right)$$

- Gaussian

$$K(x) = \frac{1}{\sqrt{2\pi}d} e^{-\frac{x^2}{2d^2}} \quad (4.8)$$

- Triangular

$$K(x) = c_d \max\left(1 - \frac{|x|}{d}, 0\right)$$

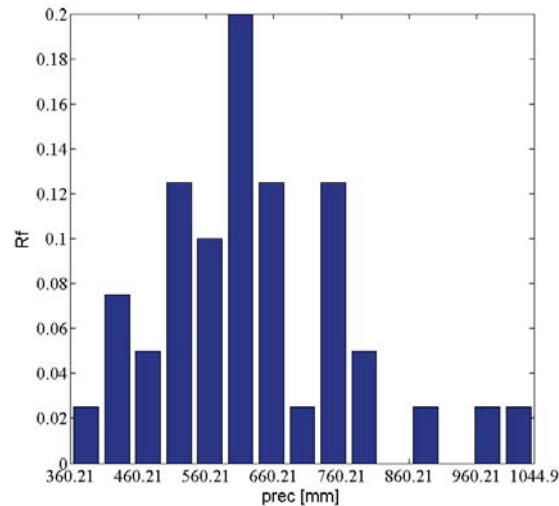


Figure 4.3: Yearly average precipitation relative frequency for the sub catchment Obrajillo

For the estimation of the free parameter h , one can have recourse to some existing methods. The method adopted in the present calculation is based on the likelihood function l defined in (4.5) with a little different consideration (see for example [72]). We can write the likelihood function

$$l = \sum_{i=1}^n \ln \hat{f}_{-i}(x_i) \quad (4.9)$$

Where the estimator $\hat{f}_{-i}(\cdot)$ has the form

$$\hat{f}_{-i}(x) = \frac{1}{(n-1)h} \sum_{i \neq j} K\left(\frac{x-x_j}{h}\right)$$

The likelihood function l (in this case the log likelihood function) depends clearly on the parameter h , which is found by maximizing l . By doing so a Kernel estimator for the unknown function f is obtained. The Kernel selected for the analysis in the present study is the Gaussian Kernel (4.8), although there is not any crucial reason as already said for this choice instead of another one.

An example can be seen in the following where a density function was fitted to the yearly average precipitation in the sub catchment Obrajillo. Figure (4.3) shows a histogram of the data and figure (4.4) and figure (4.5) shows the density estimation based on the well known naive estimator and the density estimation based on the kernel density estimator, respectively. The time period considered comprise 40 years, ranging from 1969 to 2008.

There are some important considerations when analyzing a density function. One of them is the extent of the domain of definition of the function. As we know, there are many cases in which the random variable can take merely non negative values. This means the random variable is bounded. Good examples are those variables referring to some natural phenomena as for instance precipitation, in which a negative outcome in measurement is simply not “allowed”. In general, such random variables can be left-bounded, right-bounded or bounded from each side (when considering a one dimensional space).

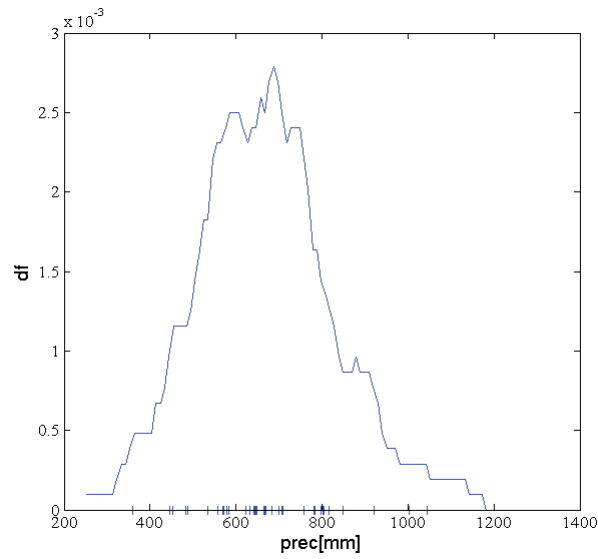


Figure 4.4: Naive estimate for the yearly average precipitation, sub catchment Obrajillo.

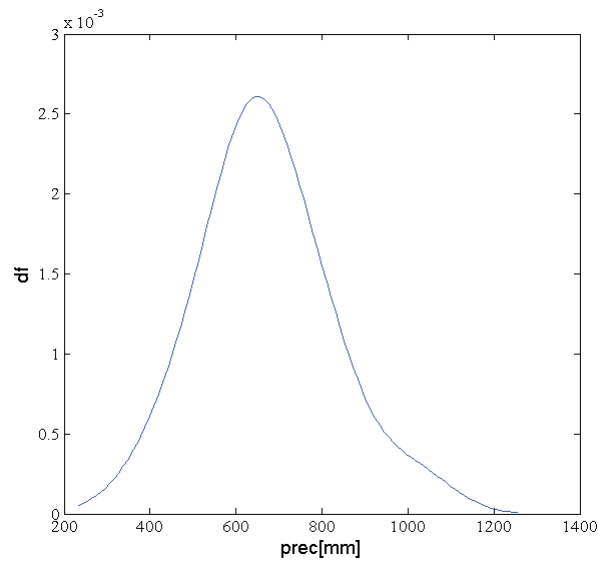


Figure 4.5: Kernel estimate for the yearly average precipitation, sub catchment Obrajillo using a Gauss Kernel function.

The point outlined above turns out to be important when the values have, for example, a natural bound in zero, and when the observations are near to this limit. More than one approach exists for treating this situation. One way is, for example, making a transformation of the random variable by defining a new function. If we have a random variable X with a density function $f(x)$ and if we first consider a transformation $t(x)$ and second its associated probability density function $g[t(x)]$, we can relate the two variables by (see [12])

$$f(x) = g[t(x)] t'(x)$$

The kernel estimate is then of the form:

$$\hat{f}(x) = \sum K_h [t(x) - t(x_i)] t'(x) \quad (4.10)$$

We can, for example, take the following transformation for building the kernel estimate of the yearly average for the precipitation in the sub catchment Obrajillo:

$$t(x) = \ln(x) \quad (4.11)$$

The kernel density estimate considering this transformation is depicted in figure (4.6)

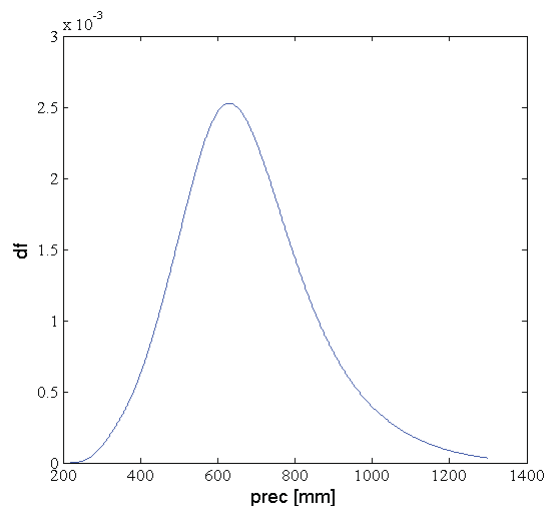


Figure 4.6: Kernel estimate of the yearly average precipitation in Obrajillo by using a log-transformation

The transformation may not make a difference when the values in the domain are not near the left bounded limit of the random variable. This is the case analyzed in figure (4.6) where the minimum precipitation registered in the period was quite a bit more than zero. However, the difference can become notable when the observations are near this limit imposed by nature, and a special treatment, for example the previous one, is recommended. Others adaptations can be made and reviewed in the bibliography given in this section.

Orthogonal series density estimation

Although the kernel method is of common use in nonparametric estimation, it possesses some drawbacks which are intended to be dealt with by some other modification methods

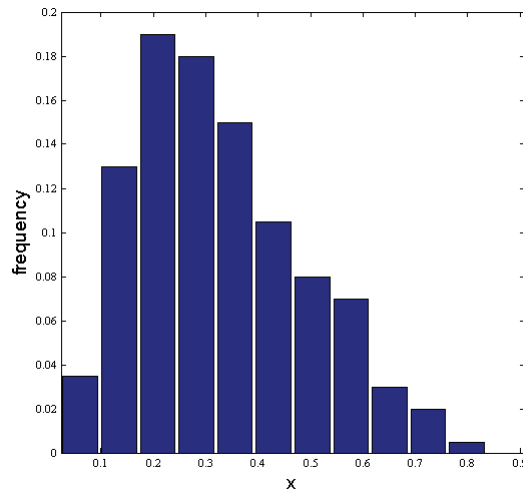


Figure 4.7: Beta distributed sample

which may not be easy to see directly or which can use a rather different mathematical approach for estimating a density function than the kernel method explained so far. We can mention for example the nearest neighbor method, variable bandwidths and orthogonal series method (see [12], [15], [46], and [48]). In general and for more information and more detail in this topic, some bibliography can be reviewed, for instance [66], [81] and [67]. In this section, a brief review of the orthogonal series density estimation will be discussed, been an interesting and different approach to density estimation. The Fourier Series Estimator is classified as one of those procedures known as general weight function estimators [84]. As the case of a non parametric estimation, this approach does not require information about the underlying density described by the data. As its name suggests, it is based on the Fourier transform.

Let's suppose that the domain of definition of the random variable lies in the subset $[0,1]$. If f denotes a function, it can be approximated with any desired accuracy by an infinite sum as expressed in equation (4.12). Suppose additionally that the function f is square integrable. We may write

$$f_l = \sum_{j=0}^{\infty} \hat{f}_j \varphi_j(x) \quad (4.12)$$

where the function \hat{f} given by

$$\hat{f}_j = \int_{-\infty}^{\infty} \varphi_j(x) f(x) dx \quad (4.13)$$

is the Fourier coefficient and the function φ is in general a complex exponential.

For the next analysis we will consider a sample following a Beta distribution depicted in figure (4.7)

The basis used for the estimation and which will be used in the analysis is based on the cosine function [25] and defined as

$$\varphi_k(x) = \begin{cases} 1 & k = 0 \\ \sqrt{2} \cos(\pi k x) & k = 1, 2, \dots \end{cases}$$

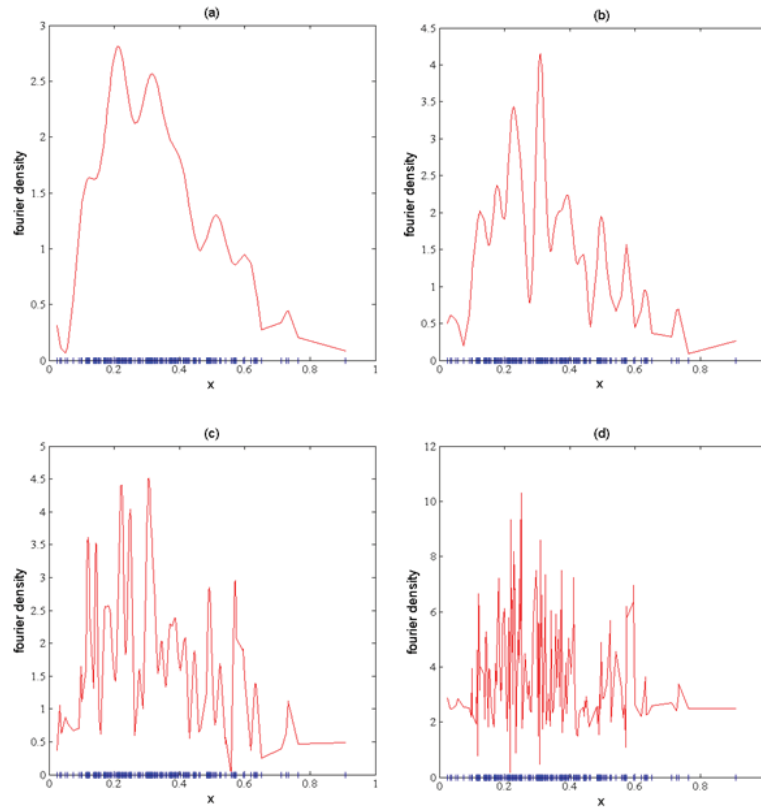


Figure 4.8: Density estimation based on Orthogonal Series with different limits l for the subindex j : $l = 30$ (fig. a), $l = 50$ (fig. b), $l = 100$ (fig. c) and $l = 500$ (fig. d)

if f denotes a density function, some important considerations can be made as we will see. The estimation of the density function according to equation (4.12) gives, however, a non desired shape of the function, far from been smoothed and well behaved from the point of view of integrability [66]. For seeing this, figure (4.8) depicts the estimated density f_l for the Beta distributed function (4.7) using the orthogonal series with different limits for the subindex j .

It is clear from the figure (4.8) that an arbitrary limit for the sum in equation (4.12) does not yield a good estimate of the density function. The greater the limit in the sum, the less smooth the estimate becomes and of course there is less resemblance to the real function that has been estimated. Indeed, the real shape can not be deduced or recognized any more beyond a certain value of l .

For solving this unwelcome outcome, an estimator can be defined by adding a new function to the equation (4.12):

$$f(x) = \sum_{j=0}^{\infty} \hat{\omega}_j \hat{f}_j \varphi_j(x) \quad (4.14)$$

where \hat{f}_j ($j=1,2,\dots$) are the Fourier coefficients and $\hat{\omega}_j$ is a shrinking coefficient. This can be defined as $\hat{\omega}_j = I_{\{j \leq l\}}$, with $I(\cdot)$ being the indicator function. This means

$$\hat{\omega}_j = \begin{cases} 1 & j \leq l \\ 0 & \sim \end{cases}$$

By defining this new function, equation (4.14) can be written in terms of a limited instead of an infinite sum. Rewriting (4.14) we have

$$f_l(x) = \sum_{j=0}^l \hat{f}_j \varphi_j(x) \quad (4.15)$$

in which the value l is to be chosen or estimated, acting here as a parameter of the density estimator. Before getting into the details, first we will discuss a key result which allows us to build this kind of orthogonal series density estimator.

As we are considering a special function f in (4.13), namely a density function, we can write

$$\hat{f}_j(x) = \int_0^1 \varphi_j(x) f(x) dx = E[\varphi_j(x)] \quad (4.16)$$

This is a very important result because it permits us to estimate the Fourier coefficients by means of the sample (observations) at hand. Using this result we can then write the Fourier coefficient estimator as

$$\hat{f}_j' = \frac{1}{n} \sum_{i=1}^n \varphi_j(x_i) \quad (4.17)$$

In this way, the density estimator is given by

$$\hat{f}_l(x) = \sum_{j=0}^l \hat{f}_j' \varphi_j(x) \quad (4.18)$$

Now coming back to the choice of the limit l , some investigations have been done in order to find an expression for a suitable value. We can mention for example one expression proposed by [38] which was used in our example. So, in order to find the parameter l we consider the following problem, provided we have a sample of data at hand

$$\sum_{k=1}^n \left(\frac{2d}{n} - \hat{f}_k' \right) \rightarrow \min$$

where the term d can be considered to be equal to 1. This can be derived by considering the variance of the Fourier coefficients and the Reyleigh's identity ([25], [38]). For more bibliographic data about this refer to [80], [48], [22], [74]. Another way of estimating the limit l is by considering the likelihood function. In our example it is found that parameter l is rather small when estimated by the expression given above. In this way, the density function for the beta distributed sample (figure 4.7) was estimated is depicted in figure (4.9)

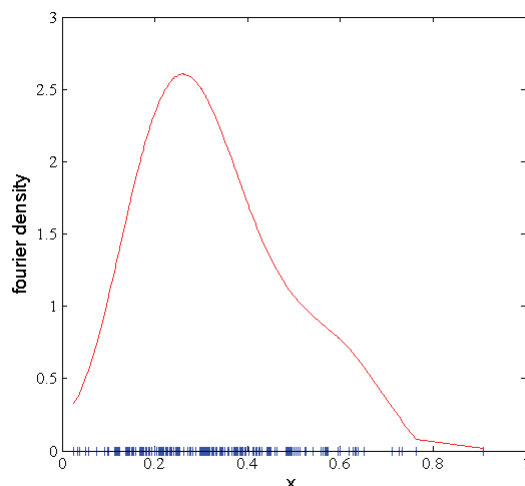


Figure 4.9: Density estimation based on Orthogonal Series with optimal choice of the parameter l

4.4 Application

The downscaling procedure explained at the beginning of this chapter was applied in some regions located in the catchments Chillon, Rimac (figure 4.10) Lurin and Mantaro (Marca projects, figure 4.16). The analysis will be done first in those areas located on the Pacific side of the Andes, and subsequently those located on the Atlantic side. The definition of the areas was based on the following criteria:

- Importance of the area defining the sub catchment, in the sense of the location of the point from which the sub catchment is drawn and the location of the main water storage sources
- Availability of information. Particularly relevant is the additional information needed for further analysis, after downscaling (discharge time series for hydrological modeling)

Downscaling Result on the Pacific Side of the Andes

In principle, it is possible to define here any number of areas for the purpose of downscaling, limited by the resolution of the digital elevation model. However, it is meaningful to define areas in view of the rainfall-runoff modeling step, where available discharge information can be used. This is especially important when calibrating the hydrological models in order to find a set of optimal parameters. Under this consideration, three sub catchments located in the catchments Chillon and Rimac are defined. One of these important areas is Santa Eulalia which contains the majority of lakes existing in the main catchment Rimac. Another important area is the sub catchment Atarjea from where water is taken out the main river Rimac for human consumption. In the following figure (4.10), the defined areas are shown. A fourth sub catchment (Sub Lurin) is also defined (5.10), but details will be given in a later chapter when the issue of transfer of parameters into unsampled locations is addressed.

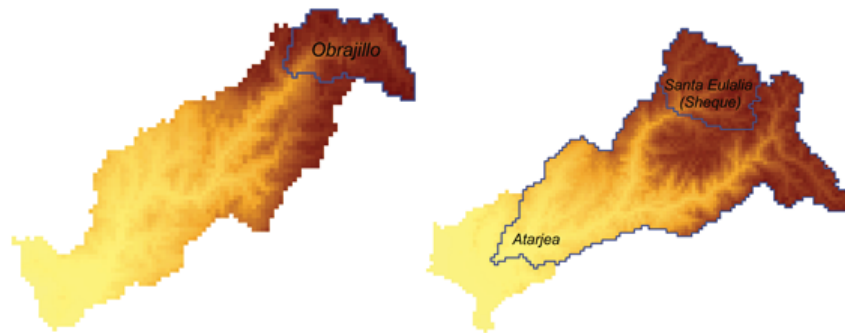


Figure 4.10: Defined sub catchments for the purpose of downscaling and hydrological modeling: Obrajillo in catchment Chillón and Santa Eulalia (Sheque) and Atarjea in catchment Rimac

Precipitation pattern

One of the characteristics of the precipitation in the regions linked to Lima is their spatial heterogeneity. This is rather evident when observing the times series of different stations along these regions or the precipitation gradient with respect to the elevation. The magnitude of this variation can for example be seen when comparing the information from a measurement station located very near the coast with another located in the upper part of the region. For this we consider the stations Aeropuerto and Milloc situated at 13 m and 4418 m above sea level respectively. The yearly annual precipitation for Aeropuerto is round 9 mm/year in contrast to Milloc with an annual precipitation of round 900 mm/year. The distance between them is about 115 km.

For the purpose of downscaling, the areal average precipitation in each sub catchment was considered. In the following figure (4.11) the monthly average precipitation for the defined sub catchments Santa Eulalia, Atarjea and Obrajillo are presented. The annual cycle is clearly recognizable in all of them ranging from a few mm in the dry season to round 200-230 mm in the rainy season, depending on the area, month and year.

The original data used for the regionalization of the precipitation have a temporal resolution of 1 day, but since the resolution of the GCM outcomes are monthly, the monthly average precipitation (regionalization output) was considered in the process of downscaling. The daily information plays an essential role in further steps, when the issue is to estimate the expected discharge variation in the next years, which will be discussed in detail in the next section where both the 1-day and 1-month temporal scales will be considered for the optimization procedure and further analysis.

Downscaling result for precipitation

Returning to our application case, the following information have been gathered and generated:

- Monthly average precipitation for the defined areas Santa Eulalia, Atarjea, Obrajillo and Sub Lurín figures (4.10) and (5.10)
- Monthly precipitation time series corresponding to the future forecast of the Global Circulation Models (Echam and Hadley)

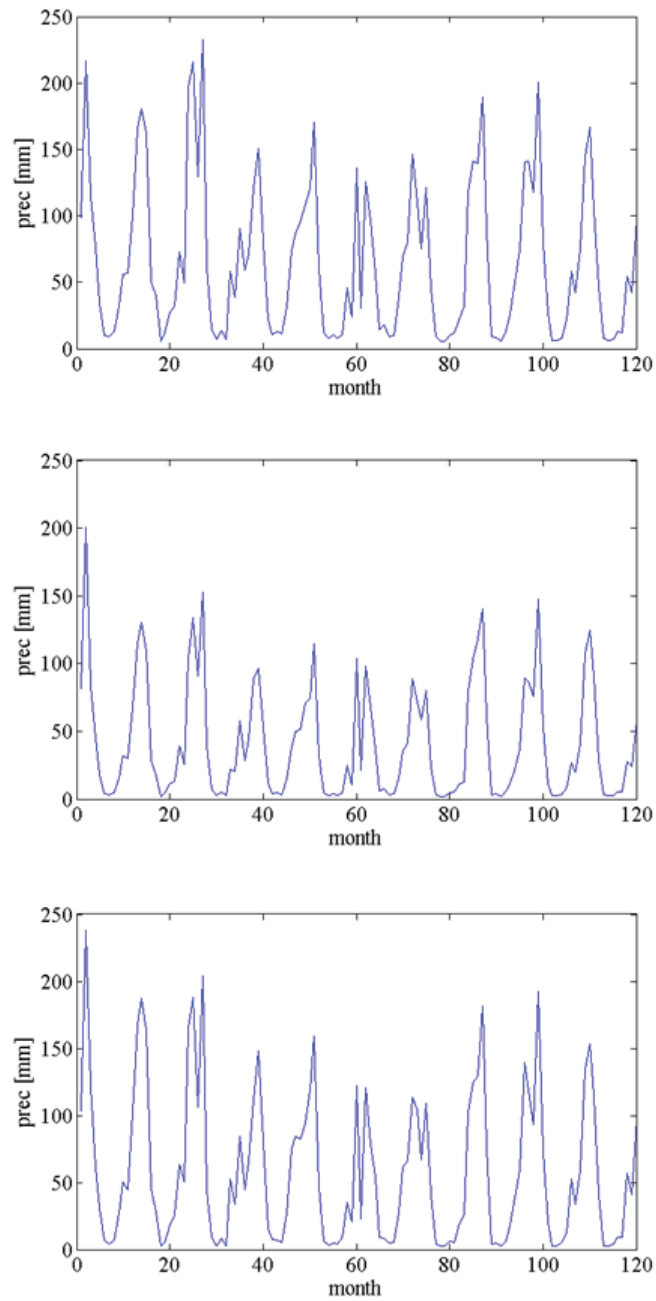


Figure 4.11: Monthly average precipitation pattern in the three analysed sub catchments: Santa Eulalia (upper figure) and Atarjea (middle figure) in the catchment Rimac, and Obrajillo (lower figure) in the catchment Chillón. Time span ranges from 1999 to 2008

Table 4.1: Estimated yearly average precipitation according to the different scenarios analysed in the time period 2012 - 2050

Region	Observed	EA2	EA1B	EB1	HA2	HA1B
Santa Eulalia	785.97	717.97	763.21	748.47	763.40	823.04
Atarjea	499.68	455.01	494.91	460.75	471.70	530.58
Obrajillo	700.98	639.34	694.53	658.73	670.53	751.53
Sub Lurín	548.11	500.68	541.39	516.73	521.44	580.81

Table 4.2: Estimated variation of precipitation (%) for the period 2012-2050 with respect to the period 1999-2008

Region	EA2	EA1B	EB1	HA2	HA1B
Santa Eulalia	-9.47	-2.98	-5.01	-2.96	4.50
Atarjea	-9.82	-0.96	-8.45	-5.93	5.82
Obrajillo	-9.64	-0.93	-6.41	-4.54	6.73
Sub Lurín	-8.65	-1.23	-5.72	-4.86	5.97

- Monthly precipitation time series corresponding to the control period of the Global Circulation Models outcomes
- Fitted distributions for the vectors \vec{v}_p^i and \vec{v}_{obs}^i as defined in section 4.2

With this necessary input data and fitted distributions we are able to estimate the transformed precipitation for each of the defined regions.

In the case of precipitation, the parametric approach was used where the Weibull distribution is considered (equation 4.2). The fitting is performed by means of the log-Likelihood function (4.5). Quantil-Quantil downscaling is then performed and new vectors and thus time series of future precipitation forecasts are obtained.

In the following, the results of downscaling are presented. In figures (4.12 and 4.13) the monthly average precipitation is shown where the observed monthly precipitation is drawn together with the models outcomes and the different scenarios analysed, namely Echam A2 (EA2), Echam A1B (EA1B), Echam B1 (EB1), Hadley A2 (HA2) and Hadley B1 (HB1).

Analysing the results, one important question is: what is the expected degree of change with respect to some pre defined time period? As the observations range from 1999-2008 and the GCM results time series used range from 2011 to 2050, the degree of change will be estimated regarding to these two periods. Tables (4.1 and 4.2) shows this expected variation together with the average annual precipitation for the observation and downscaled period.

Downscaling result for temperature

Similarly to the case of precipitation, the necessary information which was gathered and generated for temperature is

- Monthly average temperature for the sub catchments Santa Eulalia, Atarjea, Obrajillo and Sub Lurín

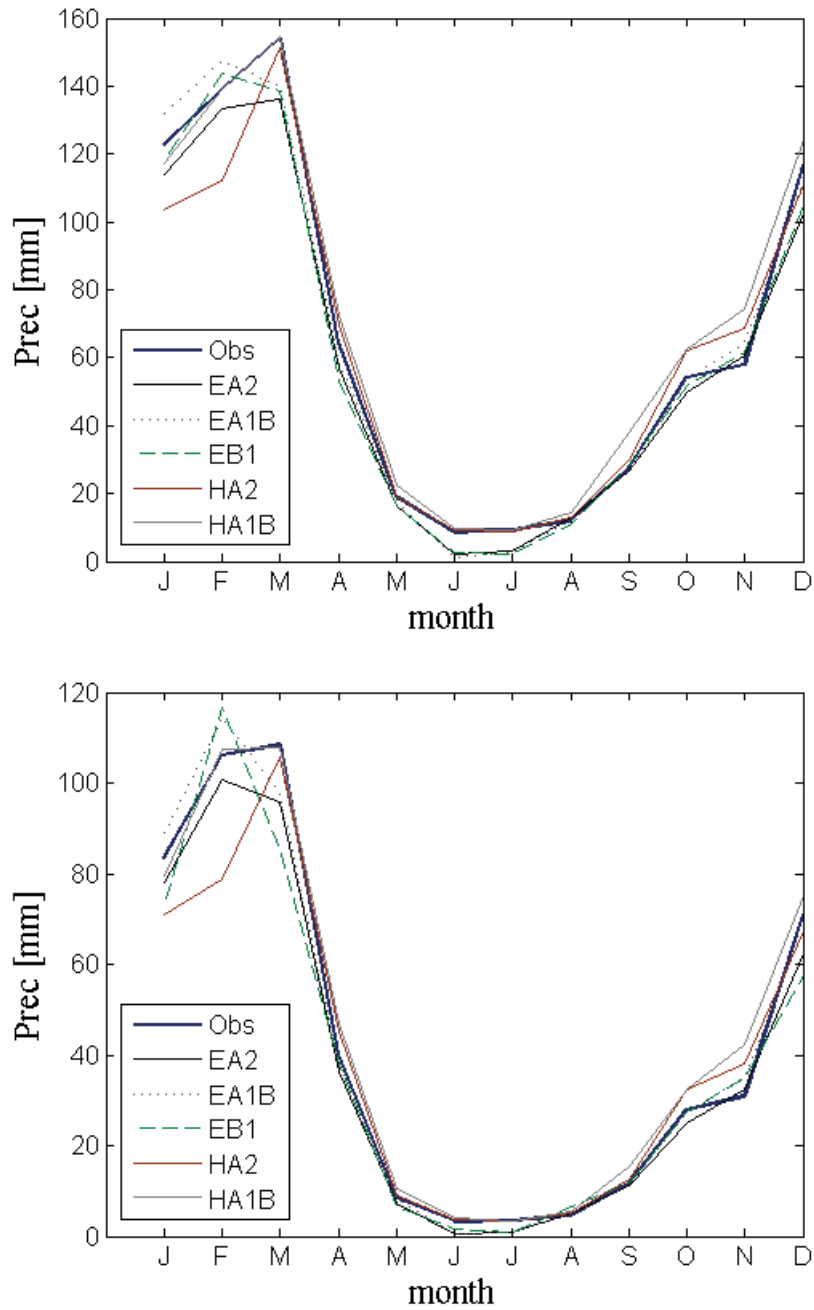


Figure 4.12: Downscaling results for Santa Eulalia (upper figure) and Atarjea (Lower figure). Curves represent the monthly average precipitation for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod

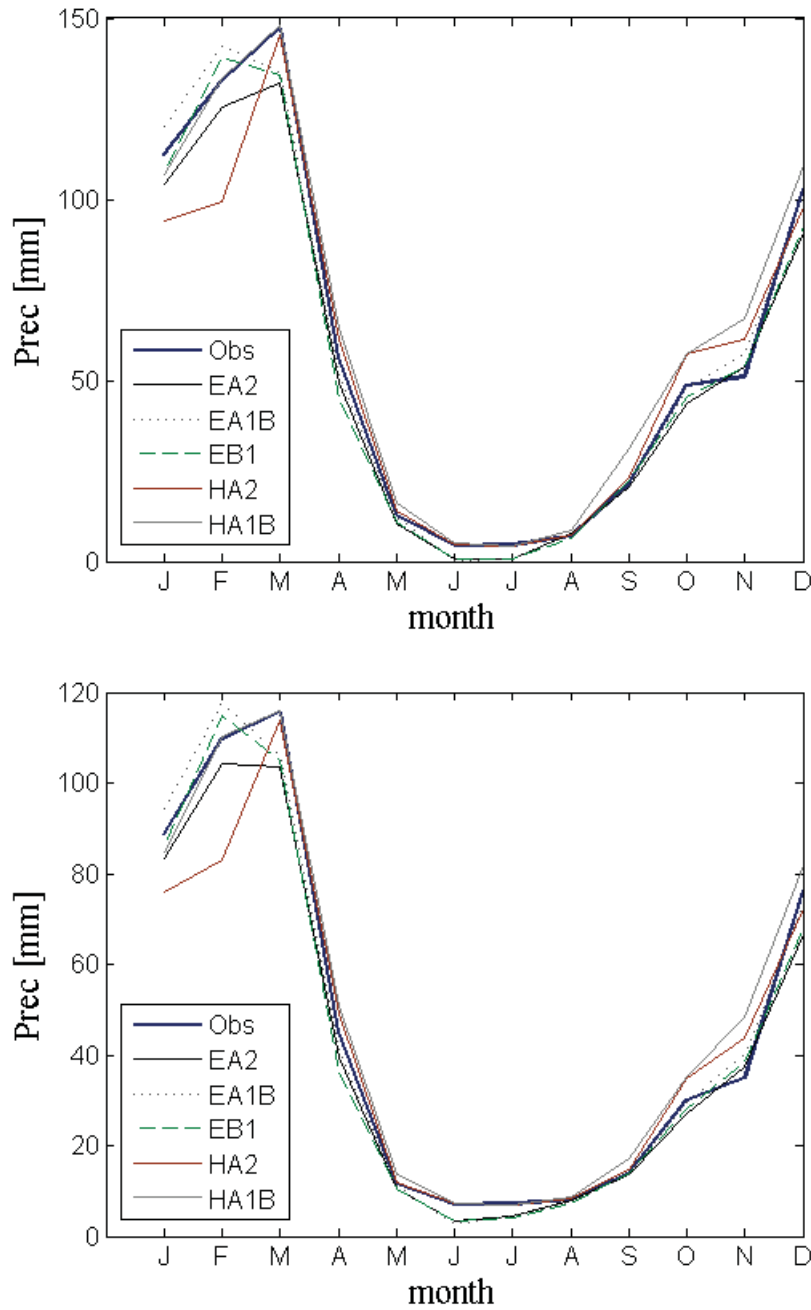


Figure 4.13: Downscaling results for Obrajillo (upper figure) and Sub Lurin (Lower figure). Curves represent the monthly average precipitation for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod

- Monthly precipitation times series corresponding to the forecast given by the global circulation models
- Monthly control period time series given by the global circulation models
- Fitted distributions for the vectors \vec{v}_p^i and \vec{v}_{obs}^i as defined in section 4.2

In contrast to precipitation where a Weibull distribution was fitted to the each of the vectors in the case of temperature a non parametric approach was used. Here no previous knowledge of a distribution function is required and the only parameter involved is the free parameter h , which is found by optimizing the likelihood function (equation 4.5).

In figures (4.14 and 4.15) the monthly average estimated temperature for the period 2012-2050 is presented for each different scenario and both models analysed, together with the corresponding monthly average temperature for the period 1999-2008. Here, E holds for the model Echam and H for the model Hadley.

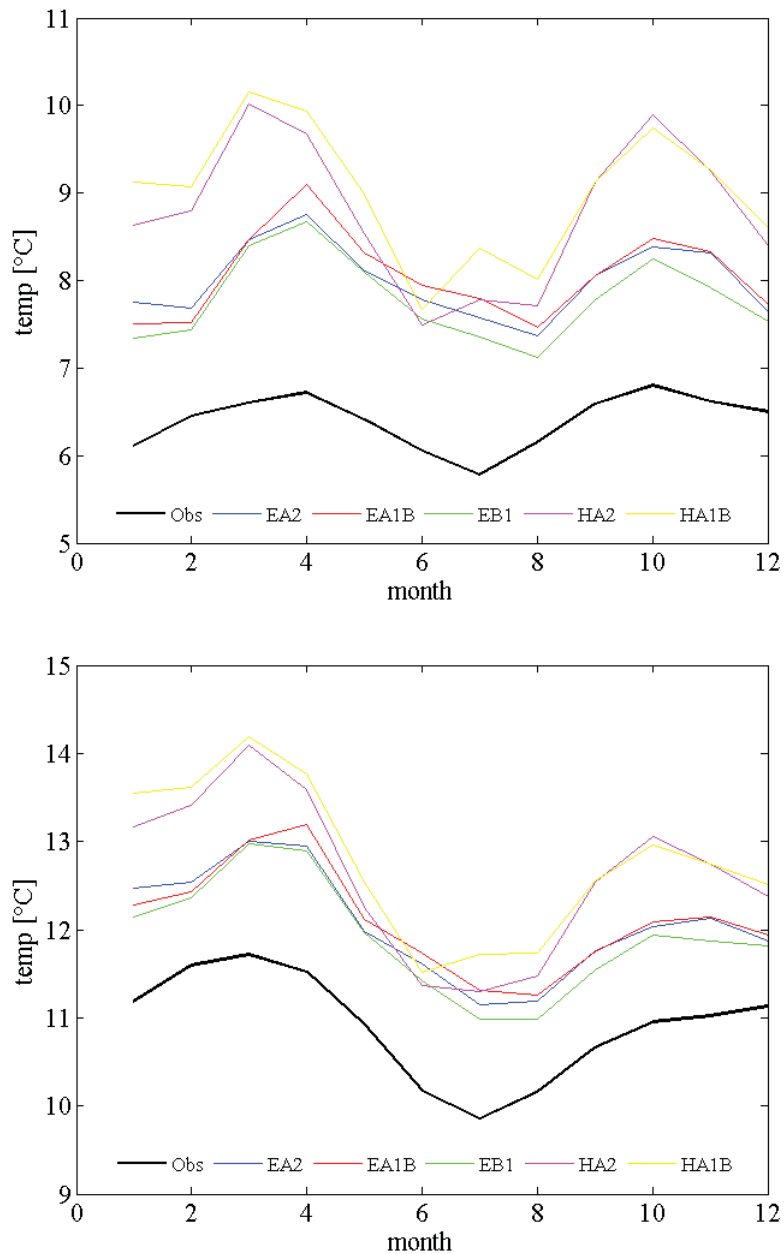


Figure 4.14: Temperature downscaling results for Santa Eulalia (upper figure) and Atarjea (lower figure). Curves represent the monthly average temperature for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod

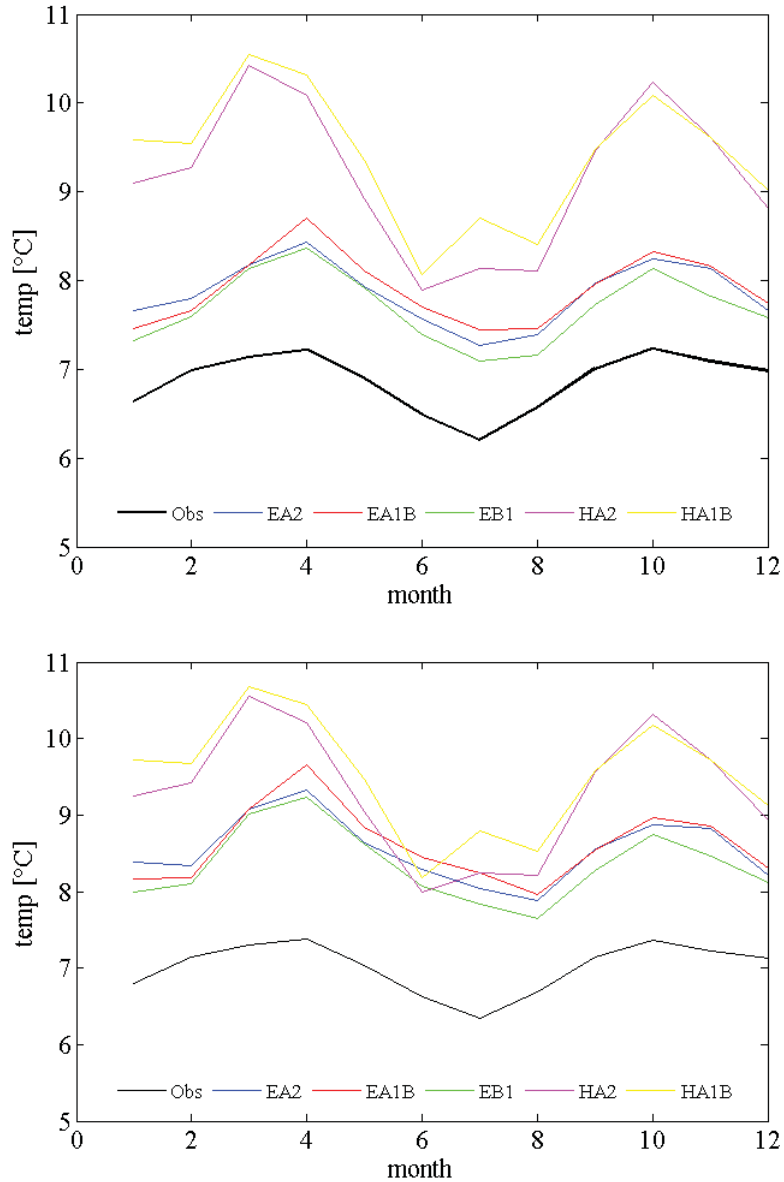


Figure 4.15: Temperature downscaling results for Obrajillo (upper figure) and Sub Lurín (lower figure). Curves represent the monthly average temperature for the observed as well as for the downscaled period. E holds for the model ECHAM and H for the model Hymod

The expected variation depends on the scenario and model considered and can be expressed in terms of the variation in temperature. In table (4.3 the average estimated temperature is presented and table 4.4) shows the expected variation for the time period 2012-2050 in comparison with 1999-2008.

Downscaling Result on the Atlantic Side

The downscaling was also performed in the regions comprising the 5 Marca projects on the Mantaro catchment. This was done for each individual area as well as for all of them together. The following picture (4.16) depicts them.

Table 4.3: Yearly average temperature for the observation period 1999-2008 and scenarios period 2012-2050

Region	Obs	EchamA2	EchamA1B	EchamB1	HadA2	HadA1B
Santa Eulalia	6.40	7.99	8.06	7.79	8.78	9.01
Atarjea	10.91	12.06	12.10	11.91	12.61	12.78
Obrajillo	6.87	7.85	7.91	7.69	9.17	9.39
Sub. Lurin	7.01	8.54	8.60	8.34	9.29	9.51

Table 4.4: Estimated yearly average variation of temperature ($^{\circ}C$) for the period 2012-2050 with respect to the period 1999-2008

Region	EchamA2	EchamA1B	EchamB1	HadA2	HadA1B
Santa Eulalia	1.59	1.66	1.39	2.38	2.61
Atarjea	1.15	1.20	1.00	1.71	1.88
Obrajillo	0.98	1.04	0.81	2.30	2.52
Sub. Lurin	1.52	1.59	1.33	2.28	2.49

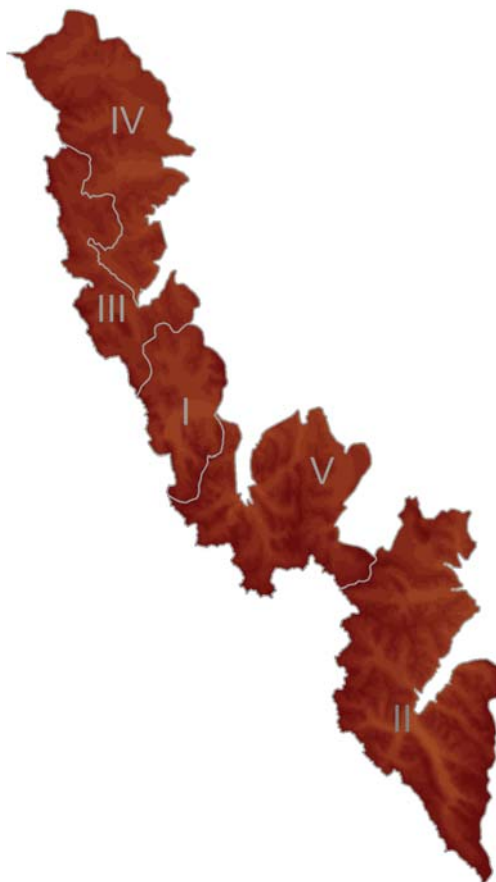


Figure 4.16: Regions on the Atlantic side of the Andes comprising the 5 Marca project areas.

In the same way as in the case of the subcatchments on the Pacific side of the Andes, a Weibull distribution was fitted (equation 4.2) by means of the Likelihood function. the results are monthly precipitation time series for each area. A summary can be seen in figure (4.17) where the average precipitation for the period 1999-2008 as well as the downscaled precipitation for the period 2012-2050 for each scenario is presented. Here the entire region is considered. Finally, the expected variation of precipitation is estimated and is summarized in table 4.5.

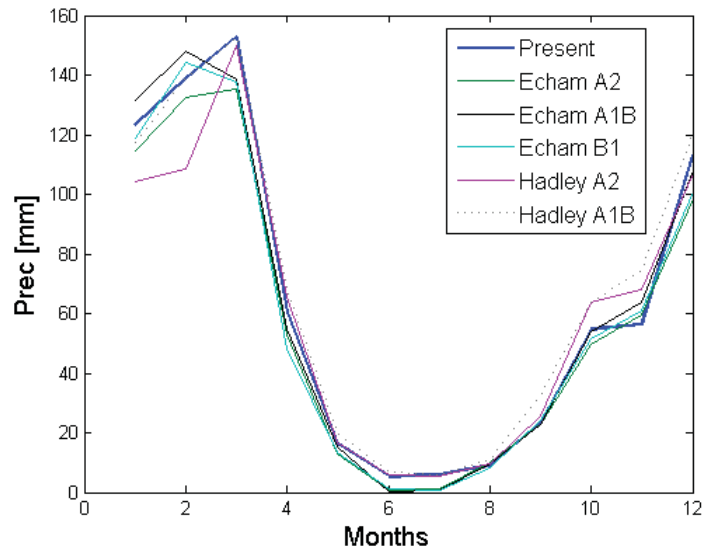


Figure 4.17: Present and downscaled precipitation for the period 1999-2008 and 2012-2050 respectively for the entire area involving the 5 Marca projects.

Table 4.5: Expected precipitation variation (%) for each region (Marca I - Marca V) for the period 2012-2050. Considered reference period 1999-2008

	EchamA2	EchamA1B	Echam B1	HadA2	HadA1B
Marca1	-9.42	-1.89	-6.90	-3.84	7.20
Marca2	-8.43	-1.22	-6.01	-3.04	7.11
Marca3	-8.64	-1.26	-5.92	-3.87	7.01
Marca4	-8.63	-1.23	-5.78	-4.39	6.58
Marca5	-9.78	-2.07	-7.33	-3.80	7.23

4.5 Summary and Final Comments

GCM outputs and observation time series allowed us to have an estimation of precipitation and temperature for a time period into the future by means of a downscaling procedure. As the time series form the GCMs, the transformed time series also contain an uncertainty which is inherited form the GCMs. The main results of this chapter can be summarized as follows:

Variation in precipitation is expected, but the assessment depends on the model as well as on the scenario considered. Positive and negative changes are observed, but clearly

there is a tendency to estimate a decrease in this variable. In fact, from all the scenarios, 4 of them estimate a decrease, and this is valid for both GCM analysed. Paying attention to the maximum expected variations, the results range from $\Delta Prec = -9.82\%$ in an EA2 scenario to $\Delta Prec = 6.73\%$ in a HA1B scenario. However, the differences between the values for each individual scenario do not vary importantly with the different regions. Regarding the temperature, all of the cases analysed assess an increase and the amount of change depend on the scenario and model. The maximum impact is observed in the model and scenario Hadley A1B with a $\Delta Temp = 2.61^\circ\text{C}$. The smallest variation is given by the model and scenario Echem B1 with a value of $\Delta Temp = 0.81^\circ\text{C}$. Finally, the generated time series will serve as an input for the hydrological modeling and hence the discharge assessment in the corresponding chapter “7”.

Chapter 5

Hydrological Modeling

5.1 General

As we know, models are a kind of representation of some reality. Hydrological modeling is normally applied to a part of the entire hydrological cycle, consisting of the transitions and interactions between the soil, in its surface and subsurface components, and the atmosphere, where important processes take place such as evapotranspiration and precipitation. This does not mean necessarily that all of these processes are represented in full detail and described with rigorous theoretical formulations, but simplifications and/or conceptual representations can instead be used. Purposes or applications of hydrological models are numerous, some of them been for example tools for the design of flood protection, irrigation, tools for hydropower design, water resources management and forecast. This last application is what concerns us primarily. For an understanding of what kind of models are available or can be developed, there is not a unique classification. One of these classifications is that which divides them into three categories: empirical, conceptual and physically based models (see [78], [83] and [33]). The first of them is mathematically simpler and requires observations of an input and an output (rainfall and discharge). It lacks an understanding of the processes in any detail because no insight into these processes is gained. Under variation or changes in the system been modeled, they are no longer applicable. The second category makes a conceptual description of some of the processes involved. This means that these processes are not expressed by the governing equations but by a simplified representation of them, giving rise with this to the definition of some parameters. In this way, the parameters do not have to necessarily have a physical meaning, but can have an abstract meaning. A consequence or an advantage of this is that the mathematics involved is not too complicated which have an effect, for example, in an easier implementation in a program or code. A greater effort, however, has to be put into finding an appropriate set of parameters by the process of calibration where an optimization procedure can be used. Finally, the third type of model is based on first principles or governing equations of the processes. This makes it rather difficult to find a solution since they require complicated numerical techniques for solving the governing equations. Nevertheless they can achieve a much more advanced process description because of the conception itself of this kind of models and it is expected a good performance especially at small scales. The applications of these kinds of models has been however restricted or limited due to the heterogeneity of process responses and unknown scale-dependence of parameters [7].

As an idea for the extent in the application of the previous classification, we can

say that conceptual models are the most commonly used kind of models [83]. Here the structure of the model is defined a priori depending on the hydrologist as well as the way the different processes can be considered, and the defined parameters are calibrated on the basis of observations.

An important issue to be aware of is that no matter whether we are talking about empirical, conceptual or physically based models, the precise nature of rainfall- runoff processes cannot yet be fully described by the models. The reason for this is mainly an unclear understanding of the physics involved and an imprecise representation of the processes (Shailesh 2010, [60]). Another important matter regarding the errors and hence the precision of the outcomes are the uncertainties, especially those which are unavoidable. We can mention here, for example [79] and [69]:

1. Data uncertainty
2. Model specification uncertainty
3. Model structural uncertainty

Data uncertainty refers clearly to the errors in measurements, since a precise value can never be achieved. Model specification uncertainty refers to the absence or the inability of finding a single best set of parameters. The model structure uncertainty is due to the simplifications made by representing the different processes, or in other words, due to the representation of the analyzed phenomena in a not-fully-detailed manner.

In our study we used two conceptual hydrological models, namely HBV and Hymod model. The objective of this section is to adapt, calibrate and validate these two models in different areas. A robust parameter estimation will be carried out and the results will be used in later sections to analyse the impact of climate change in the discharge. More detail about Rainfall-runoff modeling can be found for example in Beven [8].

5.2 HBV and Hymod Model

In this section a little introduction to HBV and Hymod model will be given. As already pointed out, both of them will be used and thereafter compared. A more extensively description of them can be found for example in [41], [6] for HBV and [55], [14] and [78] for Hymod.

HBV Model

The HVB model appeared in the early 1970s at the Swedish Hydrological and Meteorological Institute (IMHI) and was introduced by Bergström and Forsman in 1973 [71]. It falls into the category of conceptual models and in the present study will be used in its distributed version. It can be divided into three main parts or routines, namely snow routine, soil moisture routine, evapotranspiration routine and runoff-response routine. In the following, a short explanation of each of these processes will be given.

Snow routine

This routine is modeled by the degree-day approach, which considers a linear relationship between the snow melting and the temperature, given that the temperature is higher than a threshold temperature t_{crit} . The relation is given by

$$Melt = DD (t - t_{crit}) \quad (5.1)$$

where

DD : degree day factor [$mm / (day^\circ C)$]

t, t_{crit} : Temperature and threshold temperature respectively

In the case that the temperature is lower than the threshold temperature, the precipitation is accumulated as a snow. In the original version of HBV the degree-factor DD is considered constant. A modification of this setting consider that the degree-day factor varies proportional to the depth of precipitation [42].

$$DD = \begin{cases} DD_0 + kp & p \leq \frac{DD_{max} - DD_0}{k} \\ DD_{max} & \sim \end{cases} \quad (5.2)$$

where

DD_0 : Degree-day factor in the absence of precipitation

P : daily depth of precipitation

This modification (note the linearity of it) is introduced to take into account the energy available in the rain in a temperature greater than zero, in which this energy interacting with the snow would induce the melting, corresponding to a direct consequence of a basic law of thermodynamics in which heat always flows from the system where the temperature is higher to the one where the temperature is lower. The maximum value it is imposed to be DD_{max} , which is introduced in order to prevent that the DD takes high values, for example, in a case of a heavy rainfall event.

Evapotranspiration routine

Evapotranspiration is a very complex phenomenon and many different processes/variables may influence it. We can mention for example weather variables such as temperature, radiation, wind, air humidity, soil and vegetation conditions, type and texture of the soil and density of vegetation, and water availability. Given a potential evapotranspiration, the routine estimate the actual evapotranspiration as a function of the soil moisture as

$$E_a(t) = \begin{cases} \left(\frac{SM(t)}{PWP} \right) ET_o(t), & SM(t) < PWP \\ ET_0(t), & SM \geq PWP \end{cases} \quad (5.3)$$

Here

SM : Soil moisture

PWP : Permanent Wilking point

The potential evapotranspiration ET_0 can be given or can be estimated from some empirical relationships. One of them is for example the Hargreaves method ([37], [36]) which computes it as

$$ET_0(t) = 0.0023R_a(t + 17.8) \sqrt{t_{max} - t_{min}} \quad (5.4)$$

in which t_{max} and t_{min} are the maximum and minimum daily temperatures respectively and $R_a [MJ/(m^2d)]$ refers to the extraterrestrial radiation. An adjustment of ET_0 as a function of the temperature and the long term mean monthly air temperature can also be included and expressed as following:

$$ET'_0(t) = (1 + C_e(t - t_m)) PE_m$$

Where a new parameter C_e is added to the model. t_m and ET_0 represent the long term mean monthly temperature and potential evapotranspiration respectively. In this way the corrected value equals the potential evapotranspiration plus a term which changes linearly with the variation in temperature.

In our study measurements of ET_0 are available in some locations and hence are a direct input to the routine.

Soil moisture routine

The soil moisture routine takes as input the total water available $P(t)$ at each time step comprising rainfall and snow melt. The routine transforms this incoming water into two parts, namely as input to the soil moisture SM , in which it is simply added to the available SM from the previous step, and as a direct runoff according to the following equation:

$$P_{eff}(t) = \left(\frac{SM(t)}{FC} \right)^\beta P(t) \quad (5.5)$$

where

P_{eff} : Effective precipitation directly diverted to runoff

FC : Field capacity

β : Parameter of the model

Field capacity FC and the exponent β are the new parameters incorporated into the model. The later defines the shape of the curve given in the soil moisture routine of fig (5.1) and the higher this value the less the contribution to direct runoff, given that $\beta > 1$. When the soil moisture reaches the field capacity, its ratio remains constant e equal to 1 until the soil moisture decreases again below FC .

Runoff-response routine

This routine estimates the discharge generation using the effective precipitation from the soil moisture routine. It conceptualizes the complex runoff generation by means of two reservoirs set in parallel, accounting for the delay in time of the response. The first reservoir gives rise to the near surface flow Q_0 and the interflow Q_1 in the sub surface, and the second reservoir simulates the base flow Q_2 . The input for this second reservoir is calculated as the percolation Q_{per} produced in the first of them. The conceptualization of the process is summarized in the following relations:

$$Q_0(t) = \begin{cases} \frac{1}{k_0} (S_1(t) - L) A_s & S_1 > L \\ 0 & \sim \end{cases}$$

$$Q_1(t) = \frac{1}{k_1} S_1(t) A_s$$

$$Q_{per}(t) = \frac{1}{k_{per}} S_1(t) A_s$$

$$Q_2(t) = \frac{1}{k_2} S_2(t) A_s$$

Here S_1 and S_2 represent the levels in the upper and lower reservoir respectively, L a parameter which marks the limit from which Q_0 takes place, A_s is the area of the sub-catchment being modeled and the parameter k_0 , k_1 , k_{per} and k_2 are the storage constants for the near surface flow, interflow, percolation and baseflow respectively. Finally, the modeled discharge is the sum of the outputs from the first and second reservoir. This means

$$Q_M(t) = Q_0(t) + Q_1(t) + Q_2(t) \tag{5.6}$$

The structure of the model used in this study is depicted in figure 5.1, where each of the components explained above are presented. A summary of the defined parameters can be seen in table (5.1).

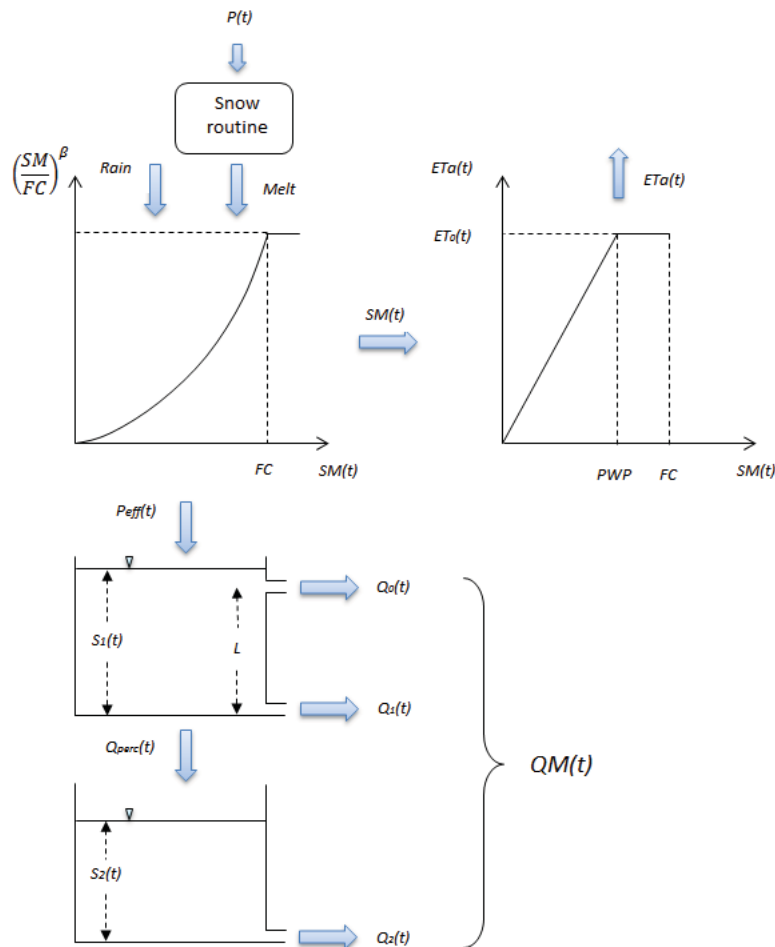


Figure 5.1: Scheme of HBV-IWS model used in this study

Table 5.1: Parameters used in the model

Parameter	Description
DD	Degree-day factor
T_{crit}	Threshold temperature in snow routine
PWP	Permanent wilting point, ET routine
FC	Field capacity, soil moisture routine
β	Exponent in soil moisture routine
k_0	Near surface flow storage constant
k_1	Interflow storage constant
k_{per}	Percolation storage constant
k_2	Baseflow storage constant

Hymod

As the HBV model, Hymod is also a conceptual model and it can be understood as being composed of two main components. The first of them consists of a rainfall excess routine described by two parameters. The second component comprises two series of reservoirs in which the first series is built up with three identical linear reservoirs and the second is built up with a single linear reservoir. One of the characteristics of the model is that the soil moisture storage capacity varies along the catchment. In this way a distribution function for the variability in space can be defined:

$$F(c) = 1 - \left(\frac{c}{c_{max}} \right)^\beta, \quad 0 \leq c \leq c_{max}$$

In this description, five parameters are defined, namely the maximum storage capacity c_{max} , the degree of spatial variability of the soil moisture capacity within the catchment β , the distributing factor between the two series of reservoirs α , and the storage constants k_1 for the first series of reservoirs and k_2 for the second reservoir. Figure (5.2) shows the structure of the model. More detailed information about the model can be found in [55], [14] and [78].

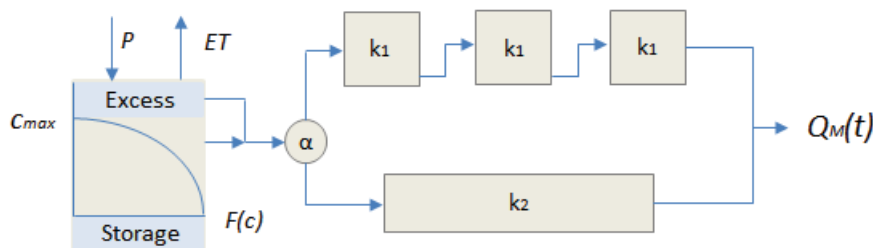


Figure 5.2: Schematic representation of Hymod model

5.3 Parameter Estimation

So far we have discussed how the different processes are modeled by the models HBV and Hymod, and introduced the basic equations that describe them. As already pointed out

at the beginning of this section, conceptual models describes the important components of the hydrological processes as a simplified conceptualization [45].

We should not forget that when dealing with these kind of models, the equations are not based or derived from the real physics that take place in the catchment. For this reason, several parameters are defined which not necessarily represent a feature of the system. These parameters should then be estimated taking into account the available information from the system. In other words, the model parameter set should be estimated through calibration as a function of the observed data, where the parameters are adjusted stepwise in the attempt to have modeled discharge and observed discharge as closely and consistently as possible (see [33]). To accomplish this task two methodologies can be used: manual calibration and automatic calibration. With manual calibration, the adjustment is done by defining a set of parameters and observing its performance according to a defined criterion. This procedure can be time consuming depending on the number of parameters involved and its interdependence, and it does not guarantees that a suitable set can be found. The automatic estimation needs less effort, and is based on the application of some algorithm. One important issue worth mentioning is that the algorithm should not get stuck in a local minimum or in its neighborhood as traditional methods do depending on the initial condition. To avoid this, some alternative methods have been proposed where this dependence and unwilling characteristic is overcome (Kavetsky et al. 2006). In the present study, simulated annealing algorithm and ROPE algorithm are used which will be explained in section 5.3 and 6.2 respectively.

Objective Function

The objective function measures the performance of a model (model structure) in a certain state or parameter configuration. Varying the parameter set in any or all the components, permits the objective function to vary in a certain degree. As we do not know a priori which configuration perform best, an iterative optimization procedure must be applied where the objective function is to be optimized (see next section). As an example we can mention the Nash Sutcliffe coefficient NS [57] and the MSE (or RMSE) function

$$NS = 1 - \frac{\sum_{i=1}^n [Q_0(t) - Q_m(t)]}{\sum_{t=1}^n [Q_0(t) - \bar{Q}_0]^2}$$

where

$Q_0(t)$: Observed discharge at time t

\bar{Q}_0 : Average observed discharge

$Q_m(t)$: Modeled discharge

$$MSE = \frac{1}{n} \sum_{t=1}^n [Q_0(t) - Q_m(t)]^2$$

In this study, the later approach is used.

In our study additional information was also incorporated since the main catchment, Rimac, is largely regulated. In this catchment, water is gathered mainly in the rain season, the lakes operated and the water distributed for both human consume and hydropower generation. This control causes that rainfall is not directly converted to discharge and is added to the lakes as water accumulation, causing an alteration of the natural hydrological cycle and hereby in the time discharge series. Another additional information which has

to be considered is the bypass of water from other regions into Rimac as input discharge. Figure 5.3 shows the monthly average mean, maximum and minimum volume involving the lakes in Santa Eulalia for the time period 1999-2008.

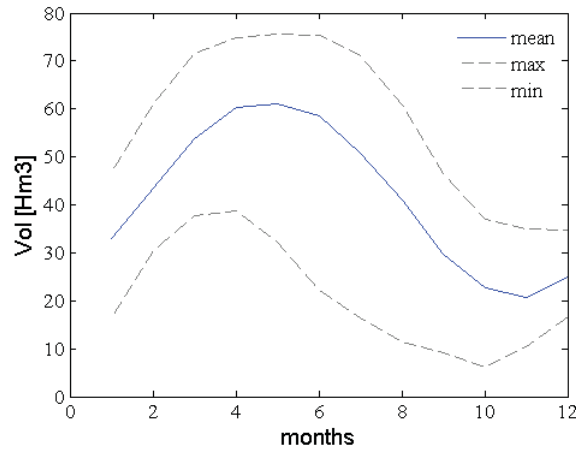


Figure 5.3: Monthly average volume in lakes, Sub catchment Santa Eulalia

Simulated Annealing

Simulated annealing was first proposed by Kirkpatrick (1983) [43]. It is an optimization algorithm which imitates a system in which its constituent atoms undergo a heating phase and a consequent slow, cooling phase. In other words, an analogy between a physical process and an optimization problem is made. The central idea is to take the concept of the Metropolis algorithm from statistical mechanics, in which the behavior of a system of particles at a finite temperature is simulated and applied into an optimization problem (Metropolis, 1953). This procedure generalizes the idea of iterative improvements of some function in which only better values of the function is permitted, by allowing a controlled “deterioration” of the performance as a part of the search of a better solution.

In the original problem one speak about Energy, temperature of the system and configuration, among others variables. In an optimization problem or more specifically in our problem where we want to find a set of parameters with a good performance, we speak about objective function instead of energy and and parameter vector configuration instead of spatial configuration of the particles (atoms). The temperature is replaced by a parameter which varies according to a predefined rule. One of the advantages of Simulated Annealing over traditional optimization methods is that it does not necessarily converge to a local optimum even though this optimum may be reached at some intermediate step. Figure (5.4) shows for example this concept using a given physical system in which a minimum is to be reached. A state E_i corresponding to the vector \vec{r}_i changes to the state E_{i+1} corresponding to the vector \vec{r}_{i+1} , where $E_{i+1} \geq E_i$.

For finding an optimum using simulating annealing we can define a series of consecutive steps as follows:

1. Define an initial T value and an initial parameter vector $\vec{\vartheta}_0$
2. Generate a random displacement $\vec{\delta}$. In this case the new configuration is given by $\vec{\vartheta}_{i+1} = \vec{\vartheta}_i + \vec{\delta}$. For the first generation of $\vec{\delta}$, $\vec{\vartheta}_i$ correspond to $\vec{\vartheta}_0$ defined in step 1

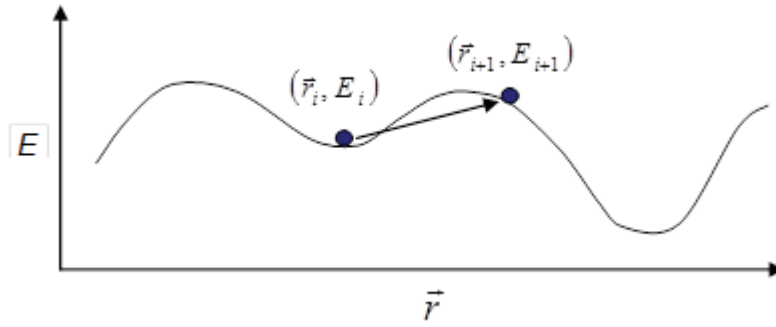


Figure 5.4: System in which a given configuration leaves a local minimum

3. Given the state $E_i = f(\vec{\vartheta}_i)$ and the new state $E_{i+1} = f(\vec{\vartheta}_{i+1})$, we can calculate

$$\Delta E = E_{i+1} - E_i$$

or equivalently

$$d = f(\vec{\vartheta}_{i+1}) - f(\vec{\vartheta}_i)$$

where f represents the objective function

4. if $d \leq 0$ the new configuration is accepted, and $\vec{\vartheta}_{i+1} = \vec{\vartheta}_i + \delta$. If $d > 0$, a pseudo random number r is generated so that $r \sim U[0, 1]$ and the configuration $\vec{\vartheta}_{i+1}$ for the next step is defined by

$$\vec{\vartheta}_{i+1} = \begin{cases} \vec{\vartheta}_i + \delta, & r \leq \exp\left(-\frac{f(\vec{\vartheta}_{i+1}) - f(\vec{\vartheta}_i)}{T}\right) \\ \vec{\vartheta}_i & \sim \end{cases} \quad (5.7)$$

The exponential in equation 5.7 can also be written as $\exp(-\Delta E/T)$ and represents the Boltzmann factor. In general and apart from normalization, it represents the probability that a particle in a system be in a defined state, at a defined temperature and in equilibrium (refer to [31], [65]).

The iterative process comprises the steps 2 to 4. A definition for the change of the parameter T (decreasing) after a number of iterations has to be defined. Also a criterion to stop the algorithm must be given, criterion which is based on the performance of the model given by the objective function, on the change of the parameter vector $\vec{\vartheta}$ and on the parameter T (for example when it reaches a predefined small value). Figure 5.5 shows an example of how simulated annealing procedure evolves with the iterations. The optimization problem depicted here is that which find a set of parameters for fitting the HBV model for the sub catchment Santa Eulalia (figure 4.10). The x-axis represents the number of iterations and the y-axis represents the objective function.

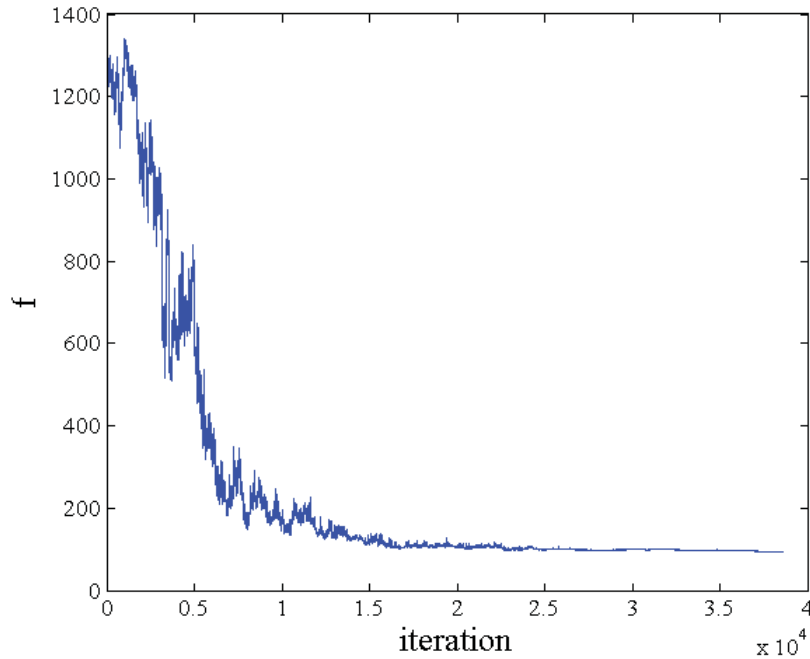


Figure 5.5: Evolution of the objective function in Simulating Annealing when calibrating Santa Eulalia sub catchment

From the figure 5.5 we can see how the objective function changes importantly at the beginning in comparison to the little changes in the last part of the curve, where the acceptance of a new configuration or parameter vector $\vec{\delta}$ is much more restrictive (Eq. 5.7).

5.4 Application

So far we have defined which models we will apply, and regarding them which method of optimization we will use for finding an appropriate set of parameters with which the model can characterize the input-response of each analysed area. In the following, we will first do a comparison between both models HBV and Hymod and subsequently we will use one of them in each area we want to study as a basis for further analysis when estimating the expected variation of the future discharge. Figure (5.6) and figure (5.7) shows the results of calibration and validation for HBV and Hymod models respectively in the subcatchment Santa Eulalia (figure 4.10).

It is clear from the figures that both models perform similarly in the process of calibration as well as in the process of validation. A closer analysis can be done by observing the error (RMSE) associated with each model performance. In order to compare the models in different situations, two sub catchments are analysed with the particularity that they present completely different characteristics from each other (regulation, input discharge), as is the case of Santa Eulalia and Obrajillo (see figure 4.10). After performing the models and calculating the error in both sub catchments, the results can be seen in table 5.2. A comparison of each model performance can then be made.

The sub catchment Santa Eulalia represents a very regulated region with transfer of water from other regions and several operated lakes, while the sub catchment Obrajillo represent a non regulated area. From these two regions where the models were applied it

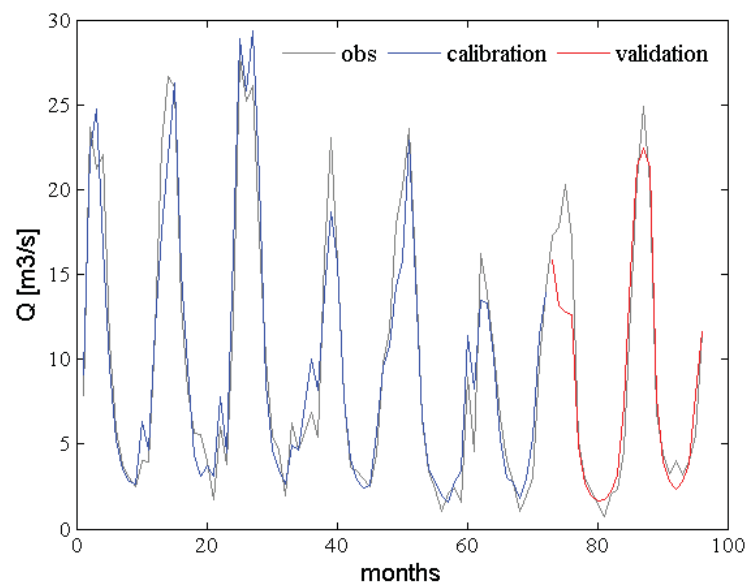


Figure 5.6: Calibration and validation of HBV model, subcatchment Santa Eulalia

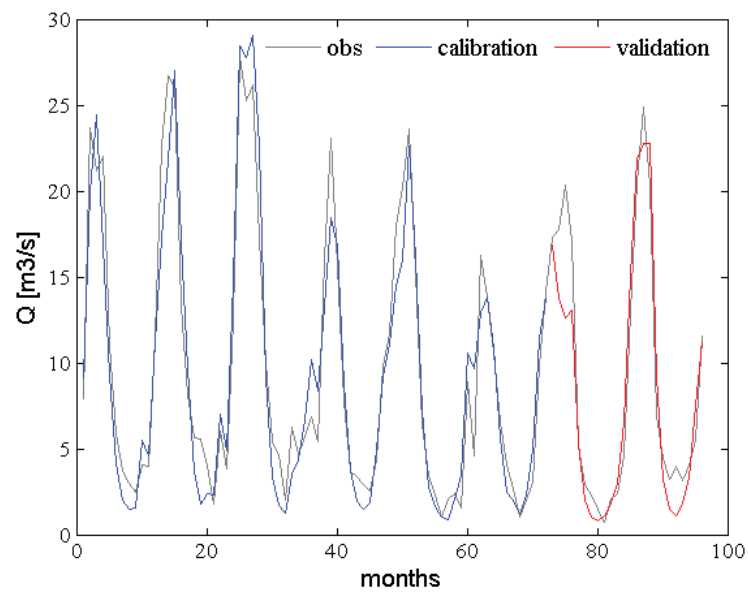


Figure 5.7: Calibration and validation Hymod model, subcatchment Santa Eulalia

Table 5.2: Error associated to HBV and Hymod model for the subcatchments Santa Eulalia and Obrajillo [m^3/s]

model	Santa Eulalia	Obrajillo
HBV	2.00	0.95
Hymod	2.29	1.35

can be inferred that HBV performs a little better showing a smaller error in the discharge estimation. Regarding the performance of the models, the smallest error occurs in sub catchment Obrajillo for HBV and Hymnod .

The models were also implemented in other regions with available discharge information, specifically in the sub catchments Atarjea and Obrajillo. In the best we should choose data which presents a high hydrological variability when calibrating the models, since much information is carried in it which is required for parameter identification [34]. Here, the availability of information is different for each analysed area, especially the information concerning to discharge time series, involving the calibration as well as the validation time period. The complete period ranges from 6 years in the case of Atarjea to 8 years in the case of Santa Eulalia and Obrajillo. Table 5.3 shows the calibration as well as the validation period for each sub catchment.

Table 5.3: Calibration and validation period for the sub catchments Santa Eulalia, Atarjea and Obrajillo

Sub catchment	Calibration period	Validation period
Santa Eulalia	1999-2004	2005-2006
Atarjea	2001-2004	2005-2006
Obrajillo	2001-2006	2007-2008

Considering the better outcomes applying the HBV model, in the following figures (5.8) and (5.9) the results of the process of calibration and validation for the sub catchments Atarjea and Obrajillo are presented. In the subsequent analysis and sections, the HBV model will be used in the estimation of the expected future discharge variation for the different scenarios.

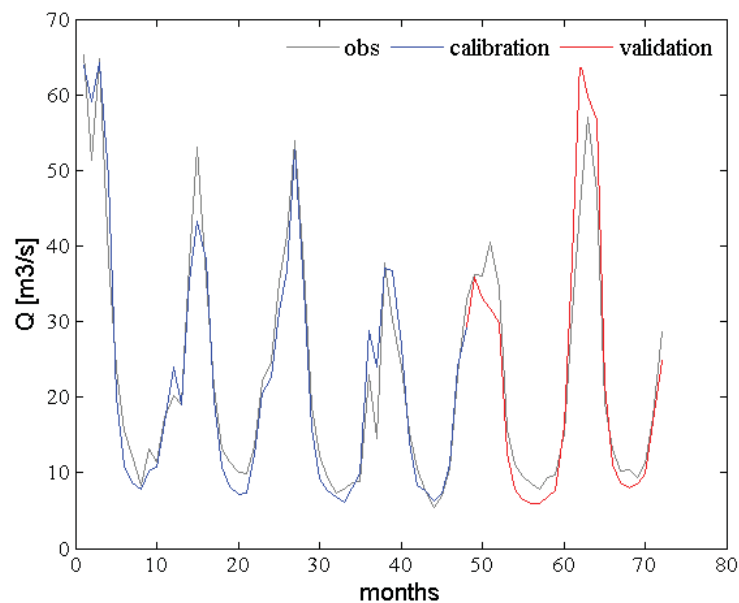


Figure 5.8: Calibration and validation result given by HBV model, subcatchment Atarjea

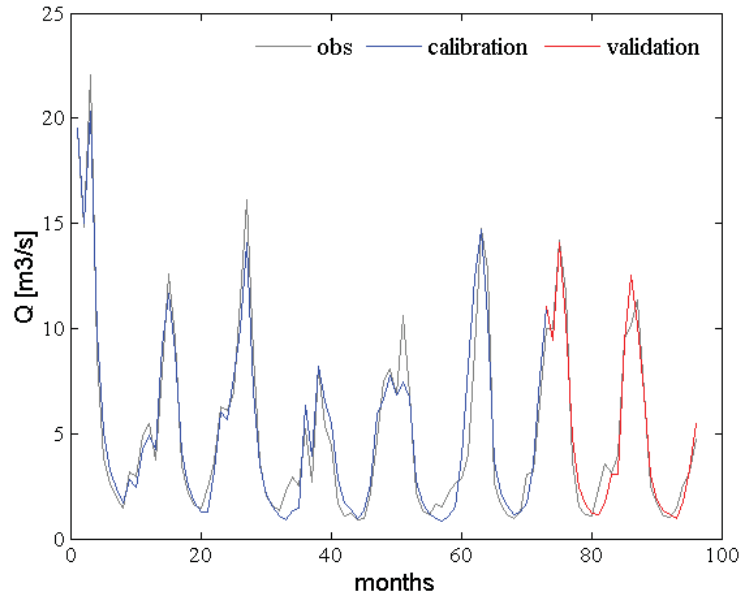


Figure 5.9: Calibration and validation result given by HBV model, subcatchment Obrajillo

5.5 Modeling in an Ungauged Catchment

In the three analysed catchments, the model was calibrated and validated and hence a set of parameters for each of them was found. This process was carried out against monthly discharge time series, with an extent according to the table 5.3. These areas are located in two different catchments, namely Rimac where Santa Eulalia and Atarjea are located, and Chillón where Obrajillo is located. However, the analysis is not only concentrated on these two catchments, but also in a third one, Lurín, located below the other two (figure 2.1). Unfortunately, no discharge information is here available with the consequence that the processes of calibration and validation are not possible, and therefore additional steps or a different strategy has to be made in order to apply the models. The analysis will be carried out in a subcatchment located in the upper part of Lurín, denominated Sub Lurín. In contrast to Rimac, Lurín catchment is not regulated at all with no input discharge from others region and no hydropower plant projects laying on it. Despite no records of discharge time series are available, 5 of the 23 rain stations used for regionalization lay in this catchment. Sub Lurín and the rain stations in Lurín are depicted in figure 5.10.

For the estimation of the discharge, a set of parameters which fit a different region may be used. As we already have performed our models in Santa Eulalia, Atarjea and Obrajillo, we can choose a set from these specific sub catchments. But before arbitrarily choosing a set and using it in Sub Lurín, we have to perform some tests for checking its appropriateness, in the sense that the estimated discharge in Sub Lurín should correspond or reproduce very closely the discharge data in Lurín, supposing they were measured. This implies that similarity from the hydrological point of view has to exist and be verified between Sub Lurín and another region. This will be done by first estimating the discharge in the sub catchment Atarjea, and second by estimating the discharge in the sub catchment Obrajillo, in both cases using the parameters calibrated for Santa Eulalia. These two results will be compared with those given by the corresponding estimated parameters for

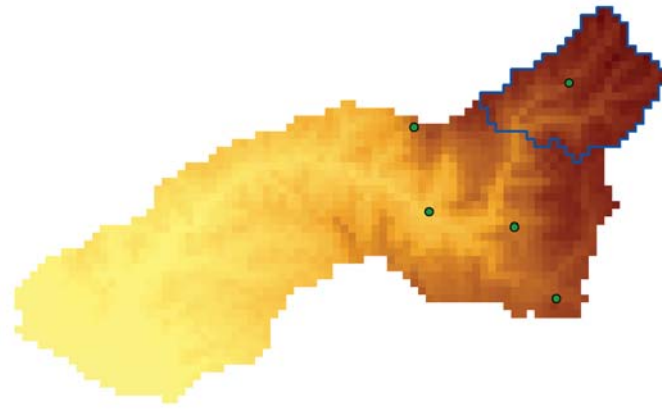


Figure 5.10: Sub Lurin sub catchment and rain stations in Lurin

Atarjea and Obrajillo. The first case involves the calculation of the response of a region in which part of it lays in a relatively low altitude (low part of Atarjea) considering a parameter set fitted in a high located region (Santa Eulalia). Contrarily to this, the second case estimates the response of a region located in a high-altitude area (Obrajillo) using a parameter set fitted in a high located area (Santa Eulalia). The results are depicted in figure (5.11) and figure (5.12)

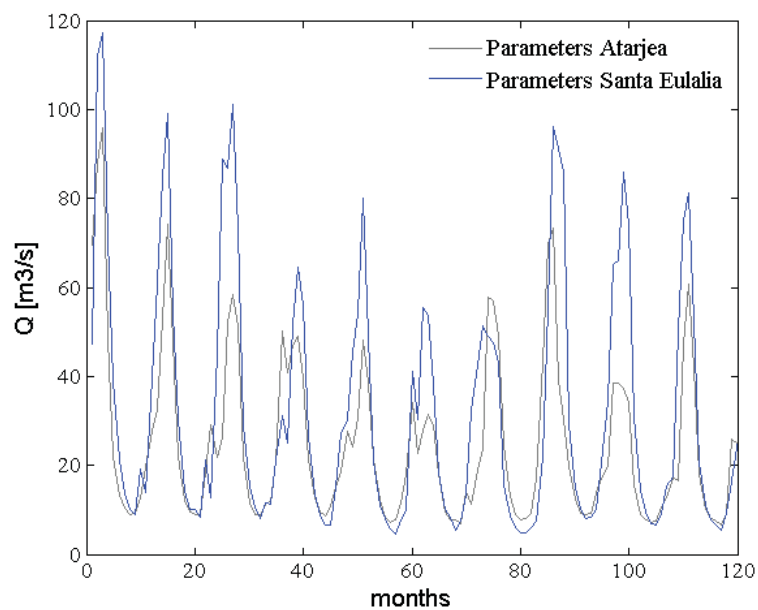


Figure 5.11: Discharge estimation in Atarjea with two different set of parameters: Atarjea and Santa Eulalia

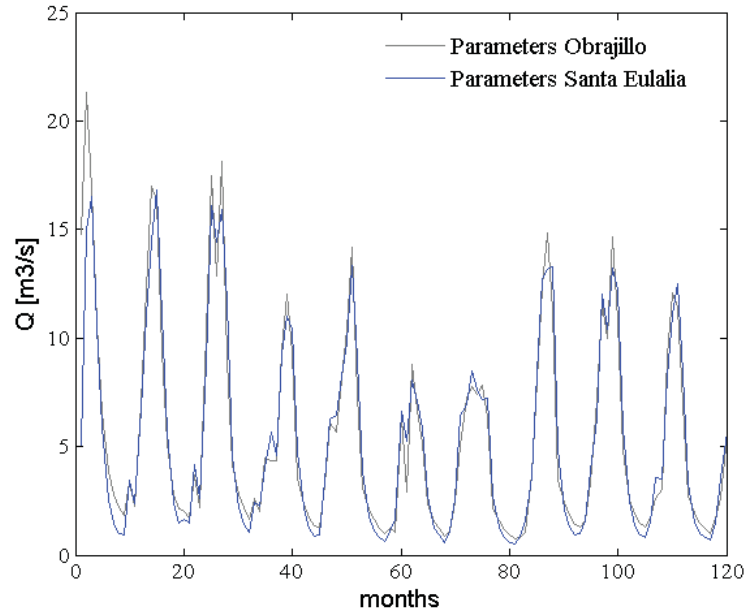


Figure 5.12: Discharge estimation in Obrajillo with two different set of parameters: Obrajillo and Santa Eulalia

Clearly by estimating the discharge in a low located region delivers poor results when using a set of parameters calibrated for in high located region (fig 5.11). On the other hand, if the two regions lay in a similar altitude, the transfer of parameters results in a much more accurate estimation. By quantifying the error (RMSE) in the discharge estimation using the transferred parameters against the real parameter set we obtain

$$e_{Atarjea} = 17.73$$

$$e_{Obrajillo} = 0.84$$

corresponding to the error in the discharge estimation in Atarjea and Obrajillo respectively. If we calculate the individuals errors as the difference between the discharge calculated with the real and transferred parameter set, we can depict the resulting histogram (figure 5.13) and calculate the spreading of both time series.

The calculated standard deviations are given by

$$s_{Atarjea} = 24.5$$

$$s_{Obrajillo} = 4.6$$

Figures 5.11, 5.12 and 5.13 show us that the transfer of parameter sets is far more accurate in the high regions or regions with similar elevations, where the model presents an estimated error of $e_{Obrajillo} = 0.84$ and a much smaller standard deviation $s_{Obrajillo} = 4.6$. As a consequence, the estimation in Sub Lurin located in a high located region is carried out with the parameters fitted in Santa Eulalia. Figure 5.14 shows the monthly average generated discharge for Sub Lurin as well as for the others sub catchments analysed previously. These will be the basis for the estimation of the future discharge variation (chapter 7). The time period ranges from 1999 to 2008.

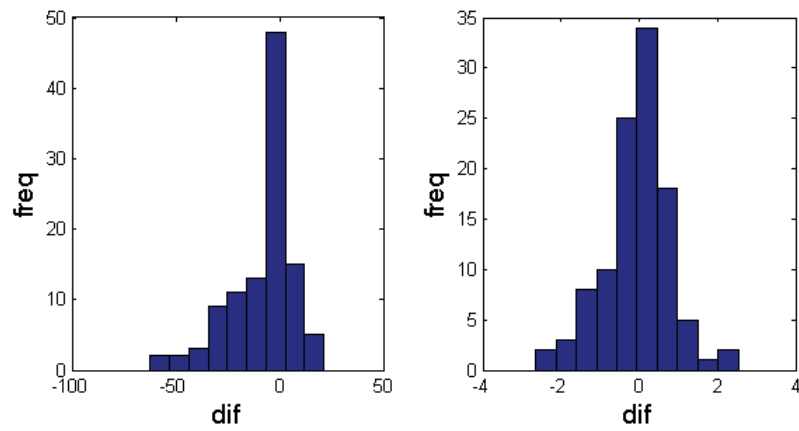


Figure 5.13: Histogram of the error (difference) between discharges estimated with real and transferred parameter set. left Atarjea, right Obrajillo

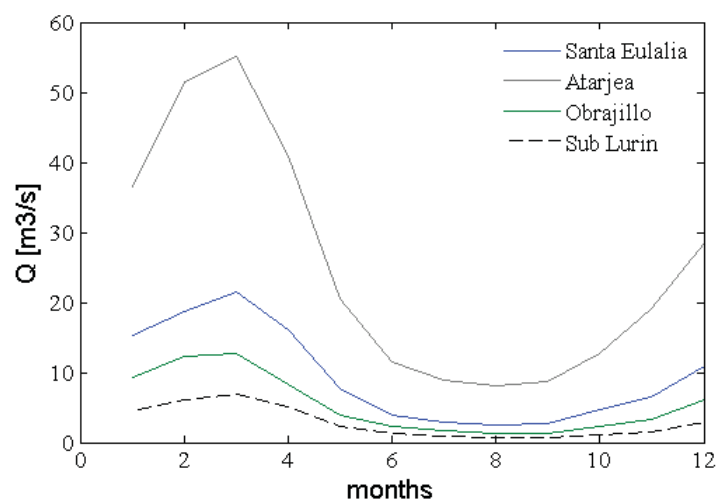


Figure 5.14: Monthly average generated discharge for the sub catchments Santa Eulalia, Atarjea, Obrajillo and Sub Lurin for the time period 1999-2008

The two most important sub catchments are Santa Eulalia and Atarjea. The first of them possess several lakes for water regulation and hydro power generation, and the second correspond to the region of water generation for human consumption. In our study, this two sub catchments together with Obrajillo and Sub Lurin will be the regions for which the impact of the climate change will be estimated.

5.6 Final Comments

In this chapter we have worked out the third important step in a discharge forecast, after regionalization and downscaling. Before using the information obtained (computed) from the previous chapters, the models for discharge estimation had to be calibrated in order to get an appropriate description (response) of the system been analysed. Special considerations were put into the optimization algorithm for the selection of a set of parameters and also a transfer of these parameters was carried out by considering the error associated in the estimations.

The next step is the generation of discharge time series with the generated precipitation and temperature time series as input information. But before getting into this step, a new algorithm ([68]) for estimating a set of parameters will be introduced and applied, which defines a robust algorithm for parameter estimation.

Chapter 6

Robust Estimation

6.1 General

Many factors can influence the estimation of an appropriate parameter set for hydrological modeling purposes. An example is the input data where measurement errors can have important impacts on the estimation of the best performing parameter vector [68]. In the previous chapter we have modeled the discharge in some previously defined points, some of them having enough information for the process of calibration and validation, and some of them not having enough information. For the first group a set of parameters was found whose performance was based on a selected objective function. Though showing a good performance, it may not be an optimal set. Added to this, different parameter vectors can perform equally well or very similar. It would be desirable to have an optimization procedure which on the one hand has a low sensitivity and on the other hand has a well performance when transferred to another time period. This last point is here of particular importance since we want to “say something” about the future by means of a calibrated model and by using the very same parameter vector in different points of time. We will find in this chapter a set of robust parameter vectors based on the ROPE algorithm (Robust parameter estimation algorithm) developed and introduced by Bardossy and Singh [68].

6.2 Robust Parameter Estimation

As already indicated in the general introduction, a new algorithm was proposed by Bardossy and Singh which was elaborated in detail by Singh in his Ph.D thesis. As a basis, a set of parameters should have following properties:

- Good performance of the model in the period of selection
- A hydrologically reasonable representation of the different processes should be achieved
- Sensitivity: Performance of the model should not be affected importantly by small changes in the parameter vector
- Transferability: Well performance of the model for different time periods

The robust parameter estimation is based on depth functions, concept which deserves to be explained in some detail and it will be addressed in the following point.

Depth Functions

Let's x be a vector from a set S so that $x \in S$ and $S \subseteq R^p$. A depth function is a function

$$\begin{aligned} D : R^p &\rightarrow R \\ \{x \in R^p\} &\rightarrow y \end{aligned}$$

in which to each vector $x \in S$ a number (depth) y is associated so that an ordering of them in the center-outward direction is defined. It can be seen as a quantitative measure of how central a vector is located when compared with a given vector set. With this, several definitions of depth functions have arisen disregarding any kind of criteria that these functions should satisfy. But we will consider depths function which ideally meet some conditions, as those introduced by [51] ([89]). These conditions suppose that the functions are nonnegative and bounded. An informal definition is given in the following (see [89]):

1. Affine invariance: The depth of a point $x \in R^d$ should not depend on the underlying coordinate system or, in particular, on the scales of the underlying measurements
2. Maximality at center: For a distribution having a uniquely defined "center" (e.g., the point of symmetry with respect to some notion of symmetry), the depth function should attain maximum value at this center
3. Monotonicity relative to deepest point. As a point $x \in R^d$ moves away from the "deepest point" (for symmetric distributions the center) along any fixed line through the center, the depth at x should decrease monotonically
4. Vanishing at infinity: The depth of a point x should approach zero as $\|x\|$ approaches infinity

A formal definition of these "desire properties" that should be possessed by a defined depth function can be found in [89]. From the previous listed properties it can immediately be seen that the data lying in the interior of the data cloud have a high depth value while those located far from the center have a low depth value. More about this can be found for example in [63], [62], [52], [23], [54], [49].

Halfspace depth function

Tukey in 1975 [77] proposed the halfspace depth function as a kind of generalization in the multivariate space of the univariate rank (order statistics). Let's define a set X in the d -dimensional space, $X \subset R^d$. If we consider a point p so that $p \in X$, the halfspace depth of this point with respect to X is given by

$$D_X(p) = \min_{n_h} (\min(|\{x \in X, \langle n_h, x - p \rangle \succ 0\}|), (|\{x \in X, \langle n_h, x - p \rangle \prec 0\}|))$$

where,

$\langle \alpha, \beta \rangle$: Scalar product between α and β

n_h : Unit vector representing a hyperplane, $n_h \in R^d$

The scalar product represent here just the scalar projection of the vector $(x - p)$ onto the unit vector n_h . So given an unit vector, the minimum set of points lying in "one or the other side" of the hyperplane is chosen. This is done for all possible hyperplanes, each

one defined by its unit vector n_h , orthonormal to it. It can be shown ([89], Theorem 2.1) that this depth function satisfies all properties (1 to 4) listed previously.

An important consideration related to depth functions is that they are computationally very demanding. The efforts can increase exponentially with increase of dimension of the vectors [68], and even the calculations can become impractical [40].

ROPE Algorithm

It is already clear from the last sections that when optimizing our objective function we do not necessarily find an optimum solution, but a near-optimum one, in general. Also this near-optimum solution is not unique, existing several vectors which perform similarly for the same problem. Important properties have been presented in the previous section which are desired for a vector solution to have, as for example that the model present a good performance not only in the selected period, and that the performance does not suffer an important decrement by small changes in the vector solution, or equivalently, the performance of the model is not sensitive to small variations in the vector solution. In the spirit of finding a robust parameter set which possess these good properties, the ROPE algorithm was developed. In the following, the ROPE algorithm is presented as introduced by Singh [68]:

Given the dimension of the parameter vector, let's say d ,

1. The limits for the d selected parameters are identified
2. n random parameter vectors are generated conforming the set $X_n = \{\vartheta_1, \dots, \vartheta_n\}, \vartheta_j \in R^d, j = 1, \dots, n$. The limits are those defined in 1
3. For each parameter vector $\vartheta_i \in X_n$ the hydrological model is run and the performance of the model is then calculated
4. A new set is defined, let's say X_m , which contains the m parameter vectors with the best performance. The number of selected vectors m can be for example the 10% of best performing vectors from X_n
5. A set of p new random vectors Y_p is generated such that

$$Y_p = \{\vartheta_1, \dots, \vartheta_p / D(\vartheta_i) \geq L, L \geq 1, i = 1, \dots, p\}$$

In this step, the depth $D(\vartheta)$ is calculated with respect to the vector set X_m

6. The new constructed vector set Y_p is relabeled as X_n and the procedure repeats from point 3. The condition for stopping the iteration is given by the performances of the sets Y_p and X_n , which should not differ more than the expected observational errors.

From the definition of the vector X_m , the $m = 1000$ best performing vectors parameters are selected in each iteration which was the base for the generation of the new vector set Y_p in the next step, and hence, the new set X_n for the next iteration. In the next section an application of this algorithm will be shown.

Table 6.1: Error of the different vectors solution

Location	error
25%	0.15
50%	0.29
75%	0.49
100%	0.91

6.3 Application and results

The aforementioned algorithm was applied to the sub catchment Santa Eulalia, in Rimac. The initial set X_n was chosen to have $n = 10,000$ vectors. The number of iterations necessary for observing a very important improvement in the objective function resulted to be $n_{iter} = 5$, where the algorithm stopped and the final set of robust parameter vectors was found. As an example of the geometry of the vector set in some iteration levels, figure 6.1 shows the parameters K_0 and K_1 , and K_1 and K_2 respectively in a two dimensional scatter plot. Here two different steps are depicted, namely iteration 3 (blue dots) and iteration 4 (red dots). On the other hand, figure 6.2 shows the histograms for the parameters K_0 and K_1 in the iteration 3 and iteration 5.

The parameters K_0 , K_1 and K_2 are the parameters of the reservoirs according to the definition given in HBV model (section 5.2). Since we are actually dealing with vectors in a dimension higher than 2, figure 6.1 represent a projection on the plane K_0 - K_1 and K_1 - K_2 respectively. An important characteristic we can see is that the points are getting deeper and deeper into the point cloud as the number of iterations increase. From figure 6.2 we can see that the parameters show a less skewness when performing further steps (iterations). It can also be seen from the figure that the range of the parameters is more restricted. This result does not only apply to the parameters presented here, but it is a general characteristic. We say that the volume of the hyper figure containing all the points decreases as the iteration steps increase.

Another interesting analysis is the comparison between the different outcomes of the HBV model when run with different parameter vectors given by the solution of the ROPE algorithm, in order to see how much the estimations differ from each other. In figure 6.3 the model discharge is depicted for the parameter vector with the best and the worst performance. Here we can see that the model performs similar when different vector solutions are used.

In order to quantify the differences in the results, the performance of each parameter vector and the best performing parameter vector are compared by means of the RMSE. Table 6.1 shows this comparison in which the vectors are sorted in ascending order according to this error.

Another interesting analysis is the comparison of the histograms of the model performance. figure 6.4 shows these histograms for different steps (runs), corresponding to different iterations of the ROPE algorithm

We observe that the model results present a lesser spreading in a new iteration in comparison with the previous one. This shorter range involves the lowest values of the objective function (error).

The model results obtained here will also be used to estimate the change in the discharge for the period 2012-2050, as we will do using the results from 5.4. Later, the two estimations will be compared.

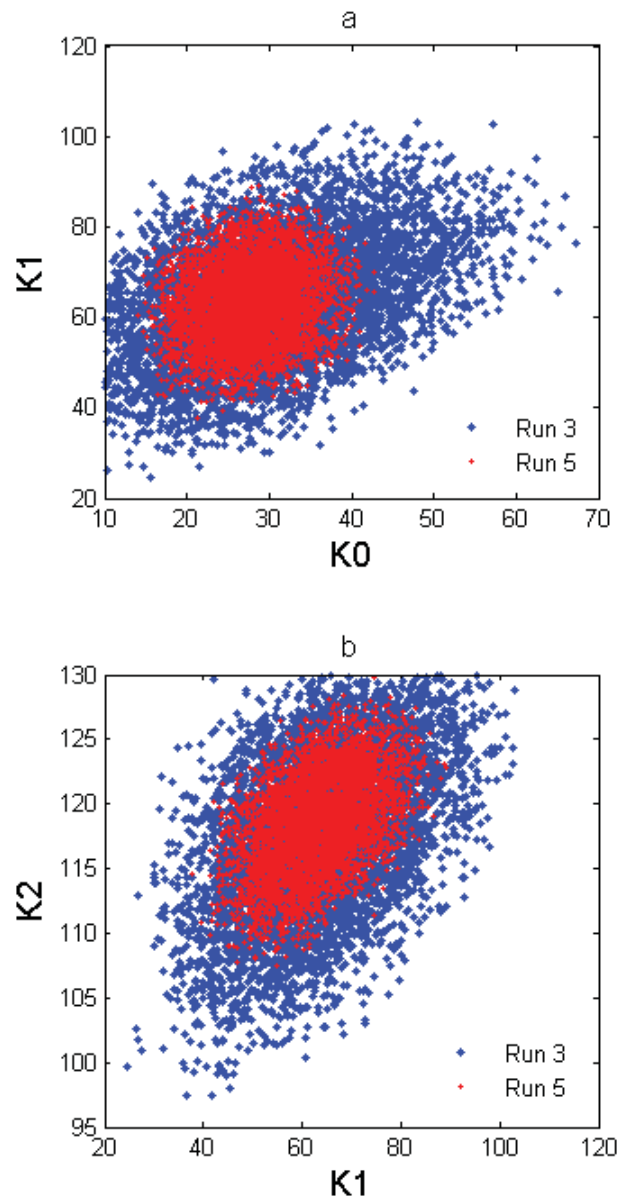


Figure 6.1: Plot of reservoir parameters showing the outcomes of iteration 2 (run 3) and iteration 4 (run 5) of ROPE algorithm

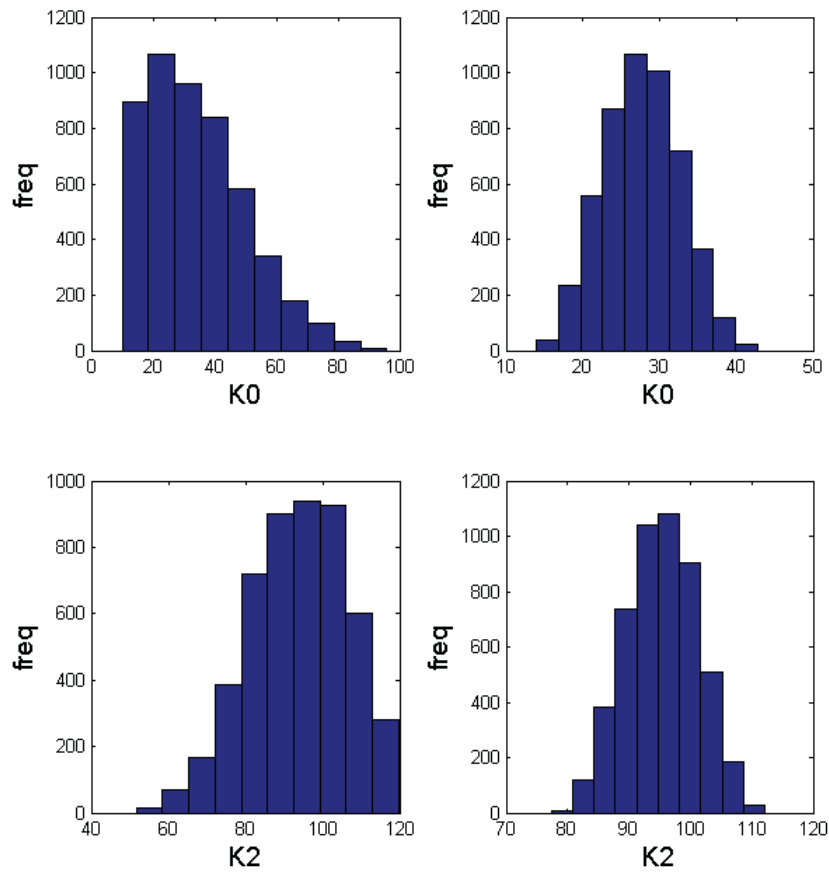


Figure 6.2: Histograms for parameters K_0 and K_2 . Top and down left, iteration 2. Top and down right, iteration 4.

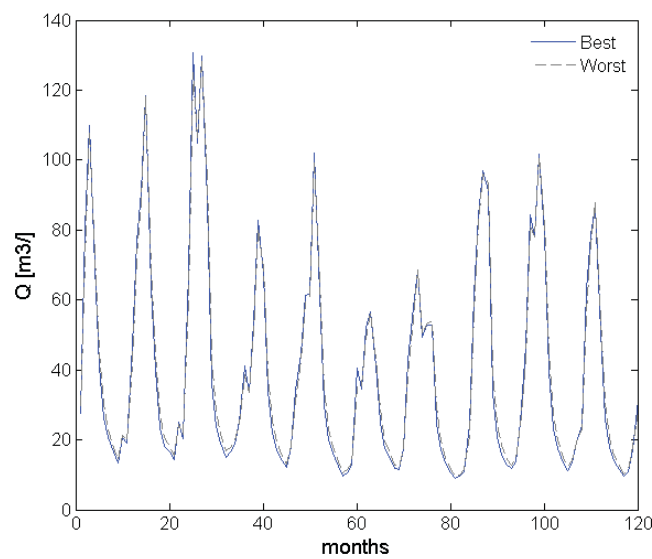


Figure 6.3: Modeled discharge considering the parameter vector from the ROPE algorithm outcomes corresponding to the best and worst performance

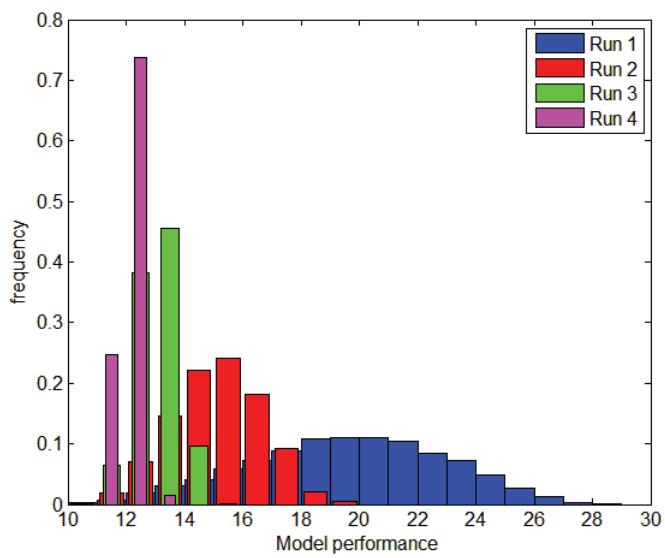


Figure 6.4: Histogram model performances for different iterations

Chapter 7

Climate Change Impact in the Discharge: Future Assessment

7.1 General

In this chapter we will present the final results of the estimation of the climate change impact in the discharge. As introduced in the first chapter, it is absolutely *meaningless* to try to make some reliable predictions at some specific point of time, especially when this point in time lies years ahead. We instead consider some different possible projections which depend on several factors as expressed in the IPCC scenarios. In this way we chose the developments given by the scenarios A2, A1B and B1. The final estimation for discharge is performed for the time period 2012-2050, which can serve for example as a tool for future planning of the water availability for the city of Lima.

7.2 Discharge Estimation

For the estimation of the future evolution of the discharge we have obtained important information in the previous steps. This is the case of precipitation and temperature time series for the period 2012-2050 and the calibrated model parameter sets. In the first case the time series are the results of a downscaling procedure where the spatial scale given by the global climate models was transformed to the basin scale. In this way five different time series for precipitation and temperature were generated, corresponding to the scenarios and models adopted. In the second case the parameter sets refer to vectors arose from both the calibration of the hydrological model by simulated annealing and the application of the ROPE algorithm. Our task now is to use the obtained time series as input data for the HBV model runs. As we have several parameter vectors for the model, we will have several discharge estimated time series (case Santa Eulalia). The different results will be at the end compared and a range of discharge variation will be found. The different sub catchments, where the future discharge is estimated, are summarized in figure 7.1, corresponding to the areas on the Pacific side of the Andes.

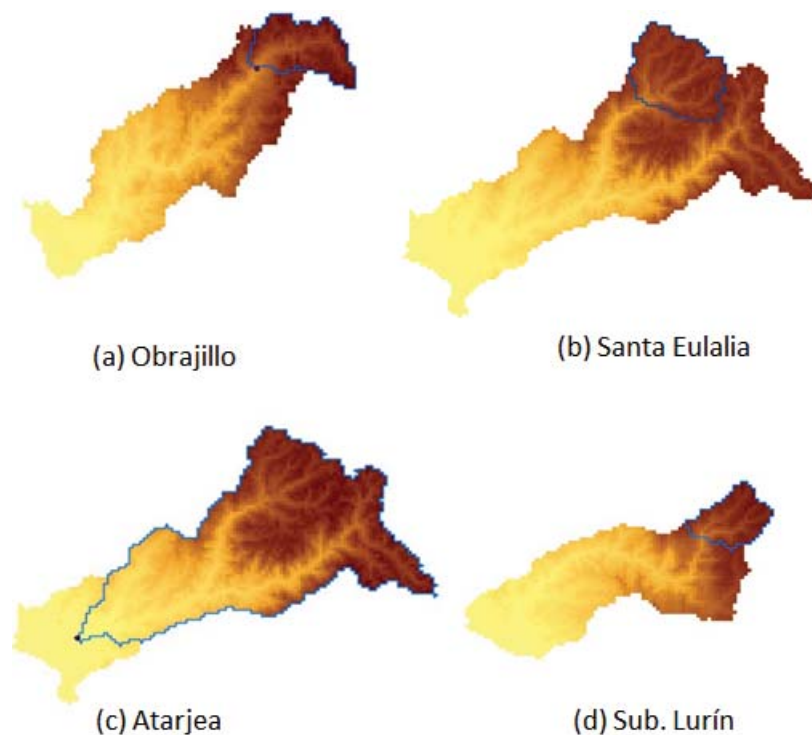


Figure 7.1: Subcatchments on the Pacific side of the Andes for discharge time series estimation in the period 2012-2050

7.3 Future Assessment: Models Results

In this section, we will present the result of the future discharge assessment for each studied sub catchment. Together with this, the estimated expected discharge variation will be calculated. Regarding this last point, the periods of comparison are the present period ranging from 1999 to 2008 and future period ranging from 2012 to 2050. As some catchments are strong regulated (Santa Eulalia and Atarjea) and the time interval corresponding to the observed period for the different sub catchment do not match, the generated time series for the present discharge will be considered, instead of the observed discharge. With this, the estimated discharge variation will be referred to at the same period of time, independent of the area analysed. In the next two sections, the results will be divided into those which arose from the usage of simulated annealing optimization and those which arose from the application of the ROPE algorithm.

Results from Simulated Annealing Optimization

Here we perform the calculations from a single parameter vector found with simulating annealing optimization (section 5.3). Using the different downscaled time series for precipitation and temperature, figures 7.2 and 7.3 show on the one hand the monthly average generated discharge for the period 1999-2008 and on the other hand the monthly average modeled discharge for the period 2012-1050.

From figures (7.2) and (7.3) we see that the average discharge for the present and the average modeled discharge for the future behave in a similar manner during the year. However, the individual months present differences which vary significantly depending on

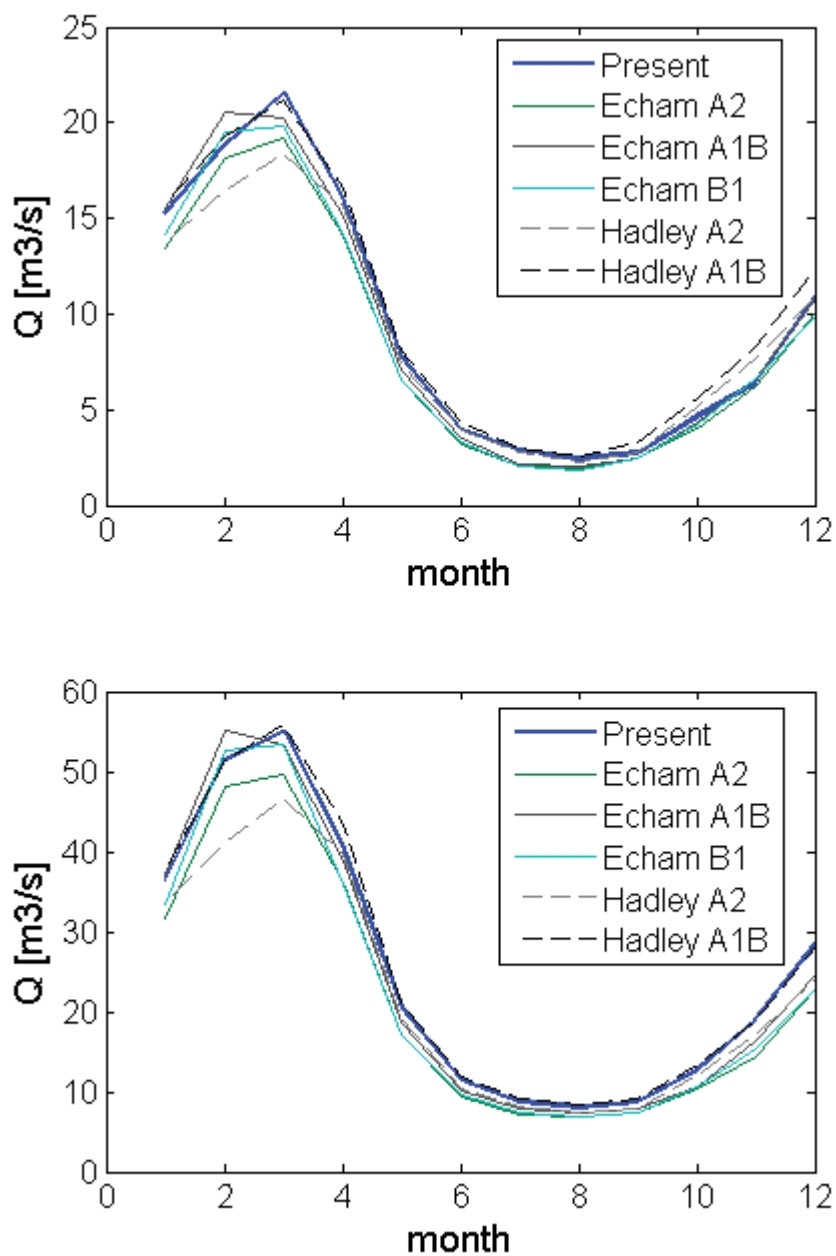


Figure 7.2: Monthly average generated discharge for the sub catchments Santa Eulalia (upper figure) and Atarjea (lower figure) for the different scenarios. Time period 1999-2008 (Present) and 2012-2050

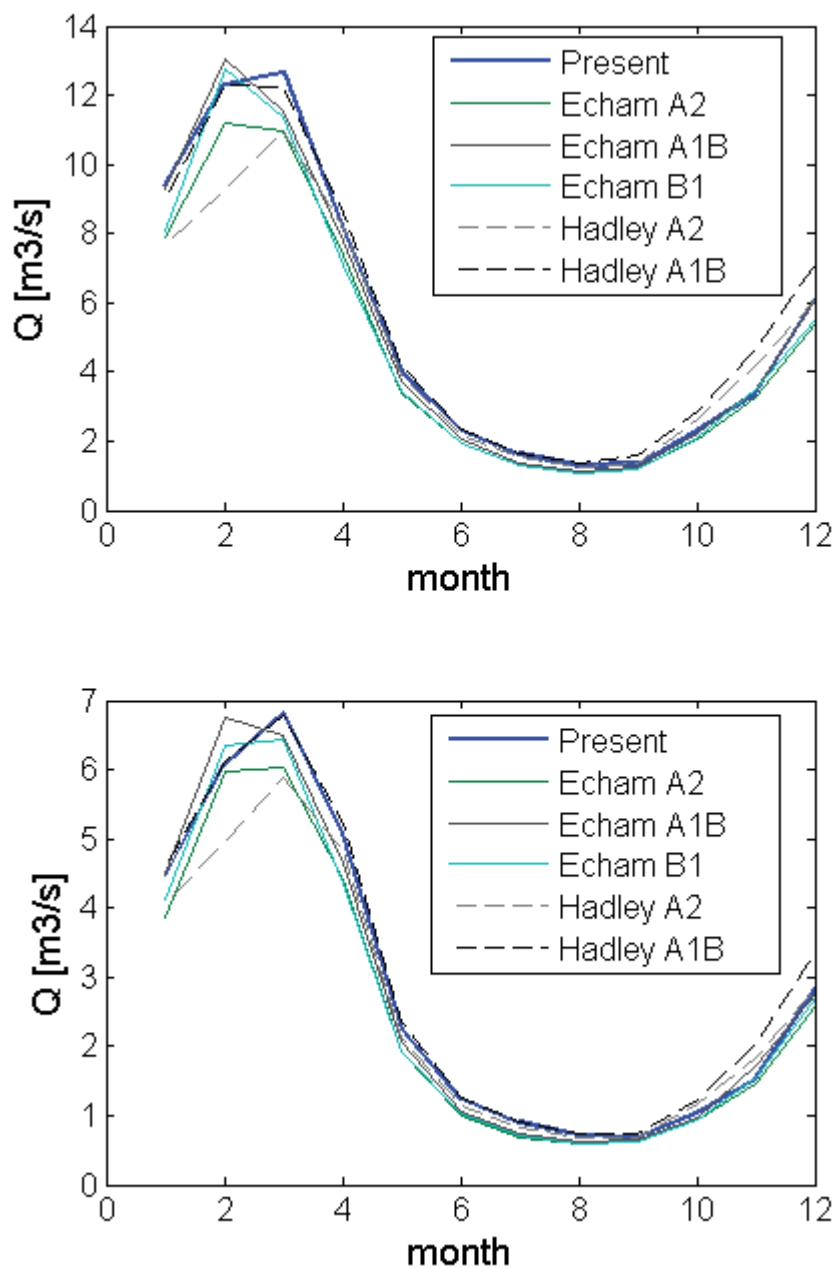


Figure 7.3: Monthly average generated discharge for the sub catchments Obrajillo (upper figure) and Sub Lurín (lower figure) for the different scenarios. Time period 1999-2008 (Present) and 2012-2050 (future)

Table 7.1: Quantification of the climate change impact on the discharge (%) for each scenario. Sub catchments on the pacific side of the Andes

Sub catchment	Echam A2	Echam A1B	Echam B1	Hadley A2	Hadley A1B
Santa Eulalia	-10.52	-2.61	-7.83	-5.44	6.28
Atarjea	-13.72	-4.80	-9.83	-11.50	2.15
Obrajillo	-12.08	-3.05	-8.63	-8.72	4.69
Sub Lurín	-10.59	-2.05	-6.97	-7.48	5.30

Table 7.2: Quantification of the climate change impact on the discharge (%) according to each scenario. Areas correspond to Marca 1 to Marca 5

Area	Echam A2	Echam A1B	Echam B1	Hadley A2	Hadley A1B
Marca 1	-10.93	-2.18	-7.81	-5.65	7.05
Marca 2	-9.77	-1.50	-6.82	-4.72	6.87
Marca 3	-10.13	-1.63	-6.81	-5.88	6.73
Marca 4	-10.32	-1.80	-6.78	-6.62	6.09
Marca 5	-11.33	-2.49	-8.35	-5.56	7.02

the month considered. So, for example, at the beginning of the year (months of January, February and March) the differences in the discharge are clearly higher than in the other months. The minimum variations are observed in those months with a low generation of discharge, namely from June to September.

Quantifying the overall differences between the present and future modeled discharge, we can have an estimate of the impact of the climate change on the discharge for each scenario. The extrem variations can then be drawn. The results can be seen in Table 7.1 for all areas

Table 7.1 makes it clear that although the impacts of the different scenarios show a similar structure for all sub catchments, they vary in magnitude and range. The maximum variation is always expected from the model Echam and scenario A2, corresponding to a decrease in the discharge in more than 10% in all cases. The only case in which an increase is expected is given by the model Hadley and scenario A1B.

The same calculations can be performed for the areas located on the Atlantic side of the Andes, namely Marca 1 to Marca 5 (see figure 4.16). The corresponding monthly average generated (present) discharge as well as the monthly average estimated (future) discharge are presented in figures (7.4) and (7.5), and the expected variation in the discharge according to each scenario are presented in table 7.2.

As in the case of the areas on the Pacific side of the Andes, the results of the different scenarios retain the same structure for Marca 1 to Marca 5, differing in magnitude and range. Again, the maximum decrease is observed to occur in the scenario Echam A2, and the single increase is expected from the scenario Hadley A1B.

Results from the ROPE Algorithm

In the last section we addressed the issue of estimating the expected variation of the discharge for each scenario. For this we applied the calibrated HBV model considering a single parameter vector. This meant that for each projection and final assessment we

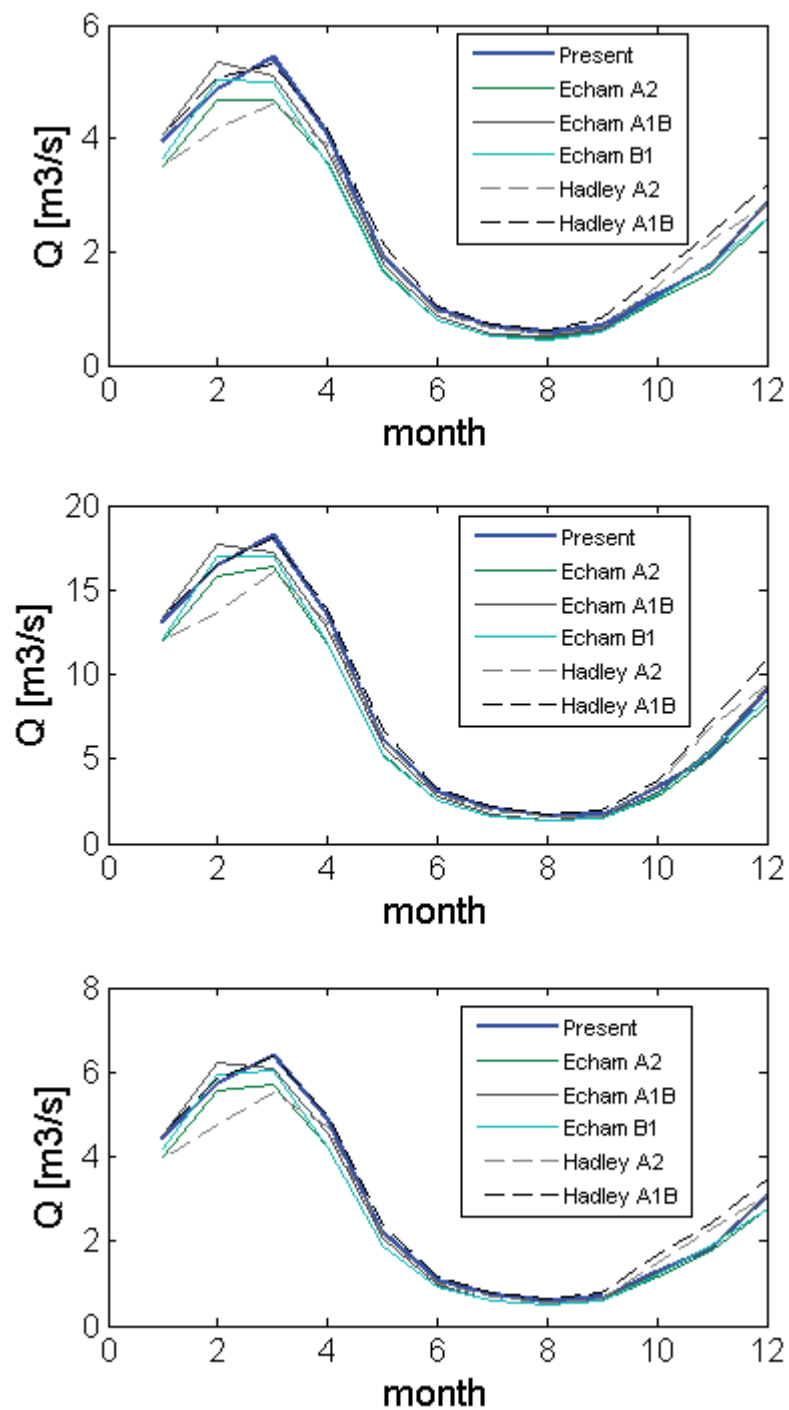


Figure 7.4: Monthly average generated (present) and estimated (future) discharge for Marca 1 (upper figure), Marca 2 (middle figure) and Marca 3 (lower figure). Time period 1999-2008 (present) and 2012-2050 (future)

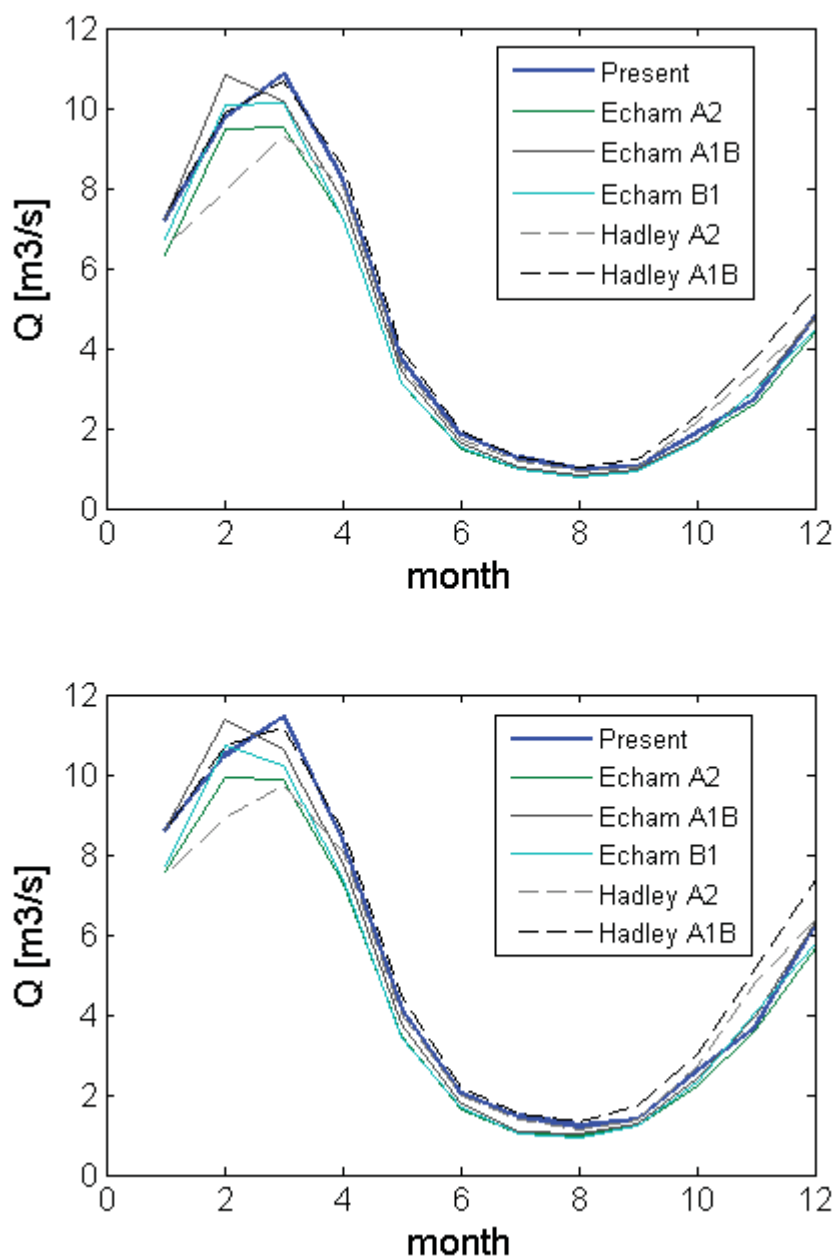


Figure 7.5: Monthly average generated (present) and modeled (future) discharge for Marca 4 (upper figure) and Marca 5 (lower figure). Time period 1999-2008 (Present) and 2012-2050 (future)

Table 7.3: Average estimated variation of the discharge by applying the ROPE algorithm, sub catchment Santa Eulalia. Time period 2012-2050 with respect to 1999-2008

Sub catchment	Echam A2	Echam A1B	Echam B1	Hadley A2	Hadley A1B
Santa Eulalia	-11.32	-3.02	-8.38	-7.09	5.00

found a single estimation. In this section we address the issue of estimating the variation of the discharge for a set of parameter vectors which perform similarly (see section 6.3). As we have a set of parameter vectors, we will have a set of estimations. In this way, for each scenario we can build a set of different time series for the discharge and its expected change with respect to a known time period. As an example, figure 7.6 shows the estimation of the variation of the discharge by performing the HBV model with all the individuals parameter sets obtained from the application of the ROPE algorithm. The case shown is the model Hadley and scenario A2.

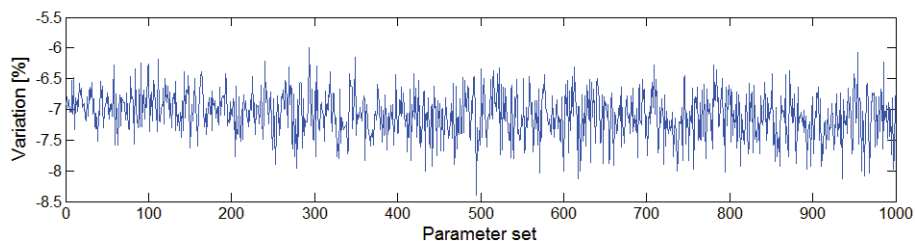


Figure 7.6: Estimated variation of the discharge for the model Hadley and scenario A2 considering the ROPE algorithm outcomes

The number of vectors used here are 1000. The minimum value which correspond to the maximum impact is $v_{max} = -8.4\%$ and the maximum value which correspond to the minimum impact is $v_{min} = -6.0\%$. Similarly to the last section, we compare the time period 2012-2050 with the time period 1999-2008. A summary of all the estimates (expected differences) for each scenario is presented in figure (7.7) in the form of their corresponding histograms. Here the range of the estimates as well as the extreme differences can be observed. The expected variation for each scenario can then be calculated (figure 7.3) and a comparison with the outcomes considering the parameters from simulationg annealing optimization can be carried out.

Similarly to the scenario A2, a set of 1000 parameter vectors were used for the discharge estimation for the remaining scenarios. Comparing the results given by the optimization performed on the one hand by Simulated Annealing algorithm and on the other hand by the ROPE algorithm, we see that the results are very similar showing small differences. The ROPE algorithm outcomes assess a little higher variation when a decrease in the discharge is estimated, and conversely a lower variation when an increase in the discharge is calculated. As an example, scenario A2 assesses a decrease of $\Delta Q = -10.52\%$ with S. Annealing optimization and a decrease of $\Delta Q = -11.32\%$ with ROPE algorithm optimization. When an increase is assessed, the results show a variation of $\Delta Q = 6.29\%$ from Simulated annealing against $\Delta Q = 5.00\%$ from the ROPE algorithm. Finally the maximum difference in the estimate between both simulated annealing and ROPE algorithm occurs in the model Hadley and scenario A2 with a value of $|d_{max}| = 1.65\%$, and the minimum difference occurs in the scenario Echam A1B with a value of $|d_{min}| = 0.41\%$.

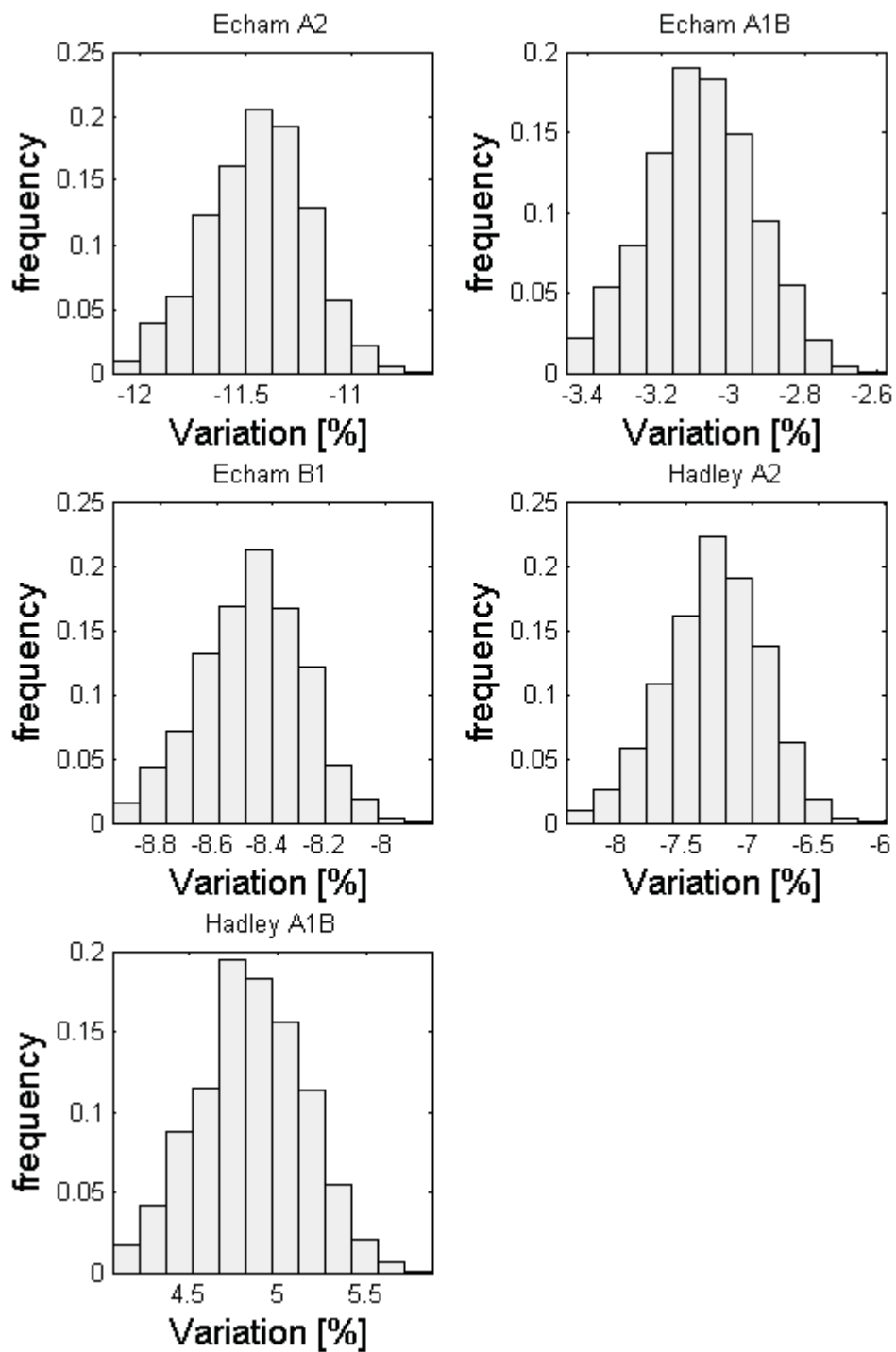


Figure 7.7: Histograms of the expected discharge variations for the sub catchment Santa Eulalia considering the vector sets from the ROPE algorithm. Time period 1999-2008 (present) and 2012-2050 (future)

7.4 Summary and Conclusions

In this chapter we have carried out the calculations for the estimation of the expected variations in the discharge until the year 2050. It is very important to be aware that there is an associated uncertainty which can not be removed from the analysis. We already witnessed that the two GCM outcomes led to different rainfall-runoff model results. So from the “list” of uncertainties such as, for example, those given by measurements, regionalization, model structure, etc., those arisen from each GCM should also be considered.

Some conclusions regarding this chapter about the final results and assessments are as follow:

1. The forecasts show the same pattern for the discharge as that observed in the present. This includes all the studied areas from the Pacific and from the Atlantic side of the Andes. The difference lies in the amount of generated discharge
2. The estimated variations in the discharge are in general similar for all the areas analysed for each particular model and scenario, and hence the ranges are comparable with each other. As in the first point, this conclusion includes the areas from both side of the Andes
3. From the 5 scenarios analyzed, 4 of them assess a decrease in the discharge and hence only one of them an increase. The assessed increase occurs in all cases considering the model and scenario Hadley A1B respectively.
4. The maximum negative variation (decrease) is estimated to occur in the sub catchment Atarjea considering the model and scenario Ecam A2 with $\Delta Q = -13.72\%$. The maximum positive variation (increase) is calculated with the model and scenario Hadley A1B in the sub catchment Santa Eulalia with $\Delta Q = 6.28\%$. These sub catchments represent the 2 most important areas from the point of view of water management, energy generation and water distribution.
5. The results obtained from the S. Annealing optimization are very similar to those obtained from the application of ROPE algorithm.

Chapter 8

Time Series and Copula-Based Autorregressive Models

8.1 General

In this chapter an analysis of precipitation time series as well as a 1-day ahead forecast for the modeled discharge in one of the studied sub catchments will be performed. In the first case the time series will be subject to a trend analysis based on a Montecarlo simulation technique. For the 1-day ahead discharge forecast, 2 different methods will be compared, namely a traditional autorregressive model and a Copula based autorregressive model. A little review of the theory will be first given.

8.2 Trend Analysis

General Idea of Trend Analysis

The question we want to answer here is if a series of data is given, for example yearly precipitation data from a gauge station, whether or not we can conclude or assert that a trend is observed which could be the result of a change in the pattern of the data generation process, or it is more likely for the trend to have occurred by chance. For this we will use a special re-sampling technique known as Bootstrapping. We will apply this procedure to precipitation data and to two different locations as we will see in the next section

Trend Estimate by Bootstrap Method

The main idea of bootstrap is to construct an empirical estimate of a sampling distribution for a predefined random variable using Monte Carlo sampling. We depart from the assumption that we have a random sample which is observed in a given period of time. In our case, we are interested in analyzing a possible trend in the yearly precipitation in two different locations. We define then the sample $x_i, i = 1, \dots, n$ from an unknown distribution of a random variable X , with n the number of years, and we consider three different cases, namely

1. All months of the year are considered for the specified period of time
2. Only the rainy season is considered

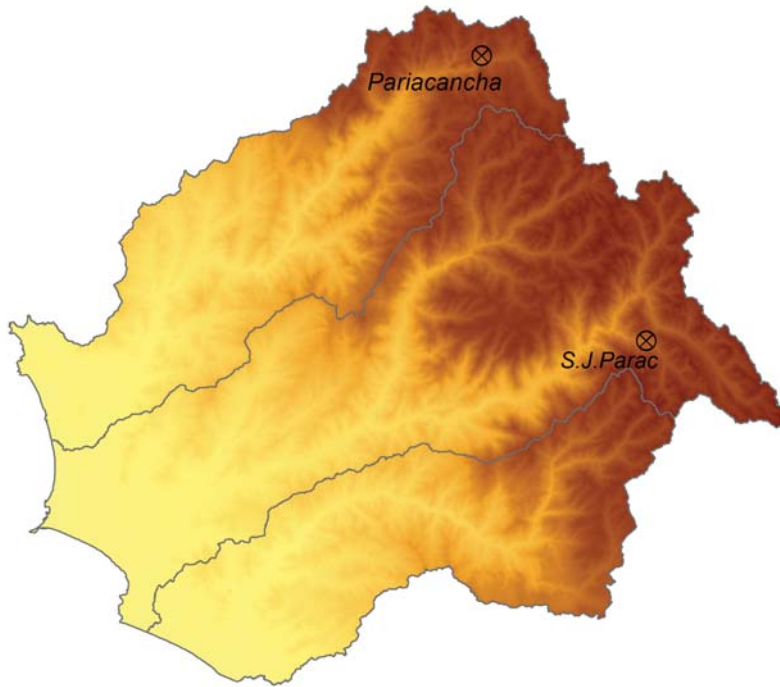


Figure 8.1: Location of the Gauge stations Pariacancha and San José de Parac in the catchments Chillón and Rimac respectively

Station	All months	Rainy season	Dry season
Pariacancha	-1.48	-0.37	-1.09
S. J. Parac	-1.02	0.60	-1.62

Table 8.1: Observed trend (%) for the stations Pariacancha and S. J. Parac. Time period 1983-2008

3. Only the dry season is considered

To analyse the trend of the observations we consider the slope of the lineal trend of consecutively yearly precipitation, that is, we define the random variable $S_l = S_l(X)$ representing this slope. For the subsequent simulations and analysis, the basis sample can be defined as the sample representing the original set $V_0 = \{x_i^0\}_{i=1}^n$.

In the Bootstrap method a large number of samples are randomly generated from the basis sample, this is, $V_j = \{x_i^j\}_{i=1,n}$ with j large. The procedure is built so that a value in the basis sample has a probability non zero to be generated more than once when constructing a new set. The random variable S_l is calculated then for each set. In this manner a sampling distribution for S_l is built. Figure 8.1 shows the gauge stations Pariacancha and San José de Parac located in the catchments Chillón and Rimac respectively where the analysis is performed. The yearly average precipitation for the two stations is presented in figure 8.2. The time period considered for the analysis ranges from 1983 to 2008.

Analyzing the yearly average precipitation considering first all months, second the rainy season and third the dry season we can notice that a little trend is observed, where both a negative as well as a positive trend occurs. Table 8.1 shows the original trend for the three cases.

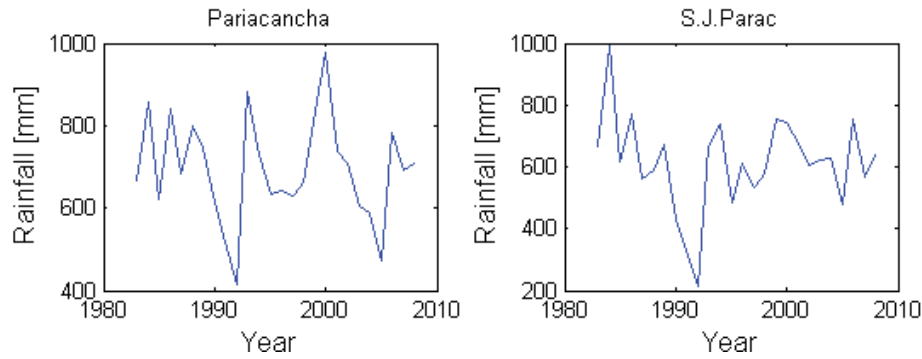


Figure 8.2: Yearly precipitation in the gauge stations Pariacancha and San José de Parac. The time period considered is from 1983 to 2008 for each station

Table 8.2: Defined quantil of the observed trends on the sample distribution curve. Stations Pariacancha and San José de Parac

Station	Yearly	Rainy season	Dry season
Pariacancha	0.33	0.44	0.13
S.J. Parac	0.39	0.57	0.45

The only positive trend is observed (table 8.1) when we consider the yearly precipitation values for the months belonging to the rainy season and for the station San José de Parac.

For the Bootstrap performance a number of $n_{iter} = 10,000$ Monte Carlo simulations were carried out and a sampling distribution $S_l = S_l(X)$ for each of the three different situations (points 1, 2 and 3) was found. Figures 8.3 and 8.4 show the result of the simulations. Figure 8.3 consider all months of the year without dividing it into rainy and dry months. Figure 8.4 consider on one hand the months of the year belonging to the rainy season (left hand figures) and on the other hand the months corresponding to the dry season (left hand figures). The red circle depicts the slope of the trend for the observations (x-axis) and the quantil it defines for the original sampling distribution $S_l = S_l(x_i^0)$.

A summary of the location (defined quantils) of the observed trends are presented in table 8.2. From there we can see that the minimum value of the basis quantil correspond to the station Pariacancha for the dry season with $q = 0.13$. However, in order it to be considered to have arisen from a change in the precipitation pattern through the years, it should fall into an extreme part of the curve, let's say the lowest 5% quantil or highest 5% of the curve, which is not the case. Regarding the other points, all of them are located quite centrally in the sample distribution curve.

As a summary of this section, what we have done here is to construct a sampling distribution of the random variable S_l by drawing randomly new samples several times from a set of data points (original sample). The results have shown that the original trends for both stations (8.1) may not conclusively have occurred due to a change in the pattern of precipitation, but may have occurred “by chance”. Bibliography can be consulted for example in [26] and [27].

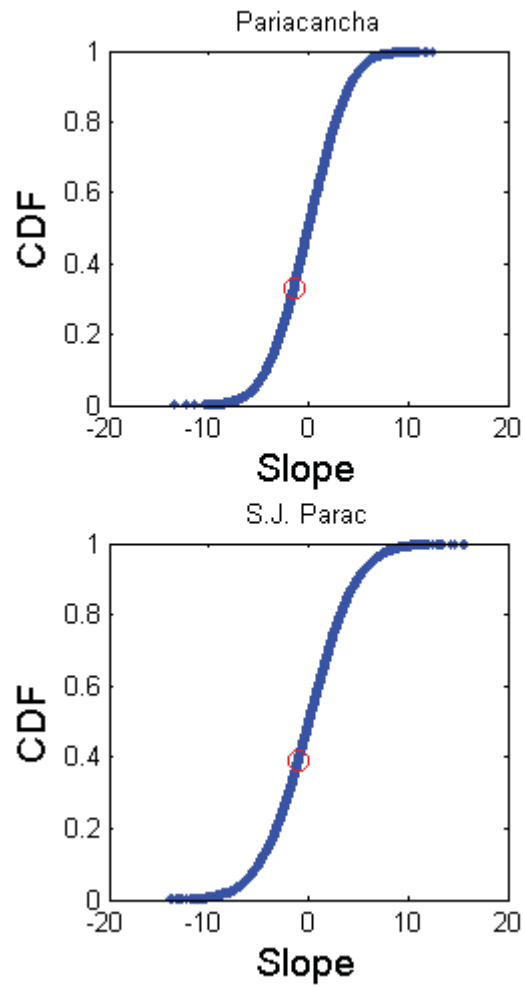


Figure 8.3: Bootstrap result for the stations Pariacancha and San José de Parac considering yearly precipitation

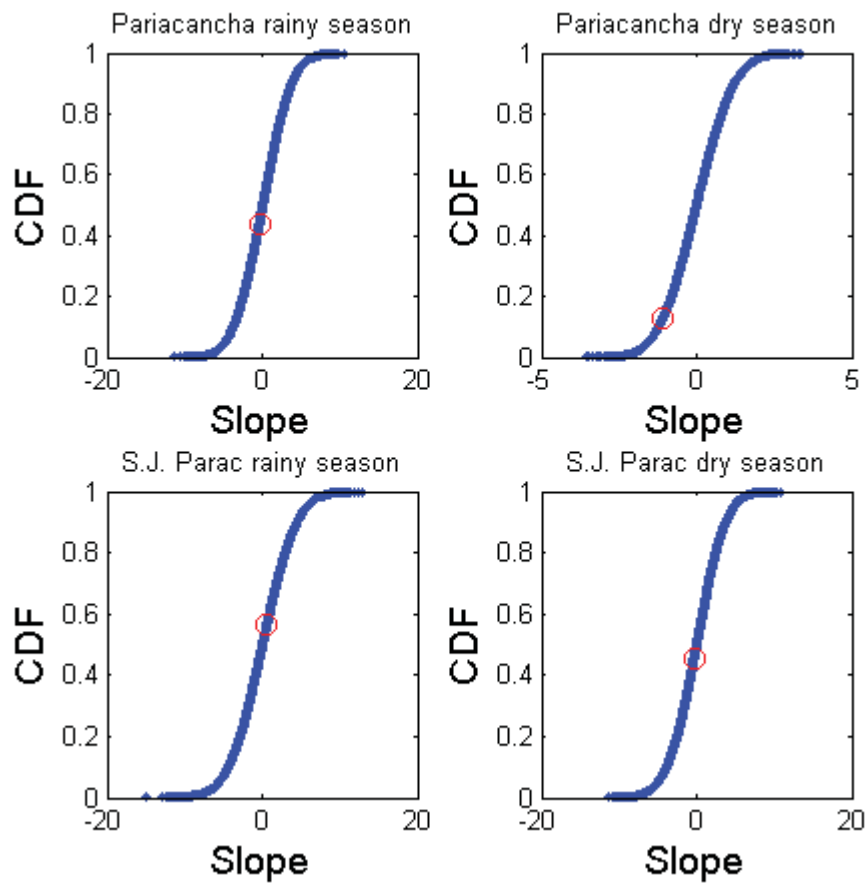


Figure 8.4: Bootstrap result for the stations Pariacancha and S.J. Parac for the rainy (left hand figures) and dry (right hand figures) season

8.3 Discharge assessment: Autorregressive models and Copula based models

In this section a short term estimation for the discharge is performed. This can be significantly important for example by reservoir operation purposes and depending on for example the present condition of these reservoirs (reservoir water level) or on the operational strategies, a reliable statement on the discharge is desirable to have at hand. Here we perform a 1-day ahead estimation of the discharge in the sub catchment Santa Eulalia, which is very important from the operational point of view for Edegel, the largest electric power generation company in Peru.

Firstly a short introduction to some of the existing autorregressive models will be given, namely the AR and MA processes. For more detail we can refer to [16] and [13]. A Copula-based autorregressive model is then introduced and results are analyzed and compared.

A key concept here is a time series. It can be regarded as a realization which was generated from a stochastic process, and from a population of possible outcomes. What we want to do here is to infer the outcome of a variable (discharge) in the next time step, by taking information from the past observations. The next stochastic processes are defined for modeling time series.

Autorregressive models

Autorregressive process

An autorregressive process is expressed as following

$$z_t = \sum_{i=1}^p \alpha_i z_{t-i} + r_t \quad (8.1)$$

here

r_t : Random process with mean zero and variance σ^2

α_i : Parameters of the model

The definition of the order of the model depends on how many past values the variable z_t is regressed on. So equation 8.1 refers to an Autorregressive process of order p , abbreviated $AR(p)$. It can be shown that the Autorregressive process is a particular case of a linear filter.

There is more than one approach to estimate the parameters of the model. One approach is by using the Yule-Walker equations which is based in the sample auto correlation coefficients. By multiplying equation (8.1) by z_{t-k} , taking expectation and multiplying by $\frac{1}{\gamma_0}$ we obtain the following set of equations:

$$\rho_k = \alpha_1 \rho_{k-1} + \dots + \alpha_p \rho_{k-p} \text{ for } k = 1, \dots, p \quad (8.2)$$

which are the Yule-Walker equations. For the estimations of the parameters $\{\alpha_i\}_{i=1}^p$, the sample auto correlation coefficients are used instead the theoretical auto correlations.

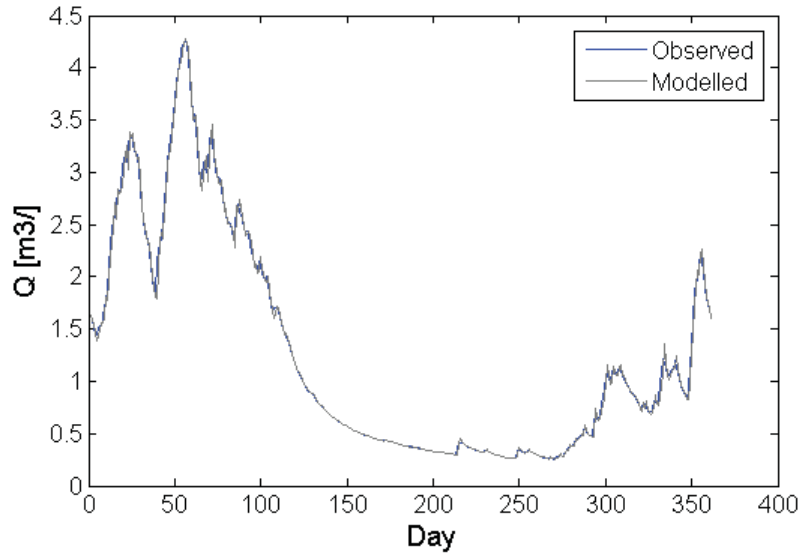


Figure 8.5: Observed and 1-day ahead estimated discharge (AR process) for the year 2008, sub catchment Santa Eulalia

Moving Average process

The general form of a moving average process is

$$z_t = r_t + \sum_{i=1}^q \beta_i r_{t-i} \quad (8.3)$$

where r_j is as defined for the AR process and β_j are parameters of the model, $j = 1, \dots, q$. The model 8.3 is said to be a Moving average process of order q which is abbreviated $MA(q)$. As for the AR process, the MA process can also be seen as a particular case of a Linear Filter.

There are other existing processes as ARMA and ARIMA, which will not be presented here, but they can be reviewed in the bibliography given at the beginning of this section.

Application

An $AR(2)$ was applied to the generated discharge in the sub catchment Santa Eulalia. The existing seasonality was “removed” by taking from each value the corresponding monthly average discharge. The resulting time series was scaled by dividing it by the standard deviation of the corresponding month (see [86]). In figure 8.5 the generated as well as the AR-modelled discharge is depicted.

Copula-based autorregressive model

A short review about Copulas

Copula theory is a new tool which arose only some decades ago and which has garnered increased interest in the statistical and non statistical community. A definition can be

found for example in [58] where copulas are defined as “functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions”, or as “multivariate distribution functions whose one-dimensional margins are uniform on the interval $[0,1]$ ”. In field hydrology Copula has been introduced not long ago, see for example [28], [32], [88]. In spatial statistics we can mention Bardossy [?], Bardossy and Li [5], Haslauer [39]. A more formal definition of copulas will be given in the following.

In the n -dimensional space copula is a multivariate distribution defined on the Cartesian product I^n as

$$C : I^n \rightarrow I$$

$$(u_1, \dots, u_n) \rightarrow v$$

where

$$I: [0,1]$$

$$u = (u_1, \dots, u_n): \text{vector so that } u_i \in [0, 1]$$

$$v \in [0, 1]$$

In this way the domain is the unit n -cube I^n and the codomain I . For example, in 2 dimensions we can write

$$C : [0, 1]^2 \rightarrow [0, 1]$$

$$(u_1, u_2) \rightarrow v$$

where $u_1, u_2, v \in [0, 1]$. Additionally, some conditions have to be fulfilled, as to have uniform marginals

$$C(u) = u_i \text{ if } u = (1, 1, \dots, u_i, 1, 1, \dots, 1)$$

and if any of the arguments is zero, then it has to be zero

$$C(u) = 0, \text{ if } \exists u_i \in u \mid u_i = 0$$

Also for any n -dimensional cube in the unit n -cube, the corresponding probability has to be nonnegative.

One very important result is the Sklar’s theorem [70] in which we find the clear dependence or relationship between a multivariate distribution function and its univariate margins. Sklar’s theorem states that any multivariate distribution function F can be represented with a Copula C .

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \tag{8.4}$$

Here

C : Copula function

F_i : 1-D marginal distribution of the multivariate distribution

If the distributions F_i are continuous, then the Copula defined in equation 8.4 is unique. Copulas show important properties, as for example that are invariant to monotonic transformations of the marginals variables. In this way, by taking for example the logarithm on one of the variables, it does not have an influence in the Copula.

There exists a variety of Copulas, as for example Archimidean Copulas [29], [30], Elliptical Copulas [24], and also empirical Copulas. More detail can be found for example in [58].

Two different bound can be defined on the lower and upper part of the Copulas. So if we consider a Copula C , for every vector $(u_1, u_2) \in Dom C$

$$\max(u_1 + u_2 - 1, 0) \leq C(u_1, u_2) \leq \min(u_1, u_2) \tag{8.5}$$

in which the proof is direct from the considerations of the marginals and the C -volume (as defined in [58]). It is used to denote $\max(u_1 + u_2 - 1, 0) = W(u_1, u_2)$ and $\min(u_1, u_2) = M(u_1, u_2)$ and the expression 8.5 can be written as

$$W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2)$$

M is referred to the Fréchet-Hoeffding upper bound and W as the Fréchet-Hoeffding lower bound. This two bounds W and M are not only bounds but are themselves copulas. In the case that the marginals distributions are fully dependent, then the copula is represented by

$$C(u_1, u_2) = M(u_1, u_2)$$

On the contrary, if the variables are independent, the copula is given by the product copula expressed as

$$\Pi(u_1, u_2) = u_1 u_2$$

If we assume that the Copula is continuous, we can then write the copula density as follows

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \tag{8.6}$$

The conditional copula is given by

$$C(u | U_1 = u_1, \dots, U_n = u_n) = \frac{\partial^n C(u, u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \cdot \frac{1}{c(u_1, \dots, u_n)}$$

Fitting a model

As in the case of the autorregressive models in section 8.3, the generated time series of discharge in Santa Eulalia was here analysed in the temporal resolution of $\Delta t = 1 \text{ day}$. For fitting a copula and the subsequent 1-day ahead estimation, 3 vectors are defined. The first of them, let's say TS1 (Time series 1), consider the original time series. The second vector, let's say TS2, consider the original time series but shifted in 1 day. The third vector, let's say TS3, consider the original time series but shifted in 2 days. In this way a 3-dimensional copula can be built.

Before getting into fitting a model, we have to transform the data considering the distribution functions they most likely represent. For this purpose the Empirical distribution function as well as the Normal distribution function were fitted. The (stationary) time series (TS) and the corresponding empirical as well as the fitted normal distribution is shown in figure 8.6.

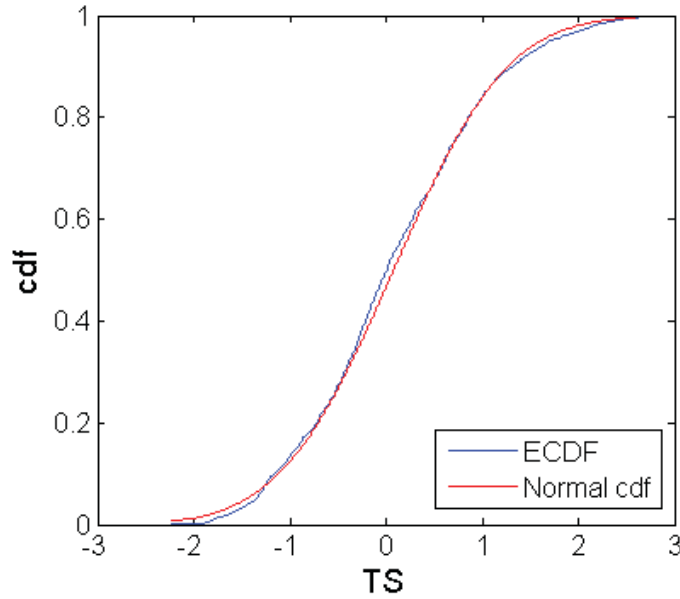


Figure 8.6: Empirical and Normal distribution functions for the time series TS

Here the ordinate (cdf-axis) becomes the independent variable when considering a copula function. In this way, for the 3 vectors TS1, TS2 and TS3, we consider the 3 transformed vectors through the distributions functions. We call them U1, U2 and U3 (as they are normally defined in the literature), which verify:

$$\forall u \in U_i, 0 \leq u \leq 1, i = 1, \dots, 3$$

The relation between the vectors can be appreciated in the empirical bivariate copulas, shown in the figure 8.7. Here the characteristics of the dependence between U1 and U2, U2 and U3, and U1 and U3 can be seen. A 3-dimensional representation involving the 3 vectors are depicted in figure 8.8.

Two different kind of copulas are here considered, namely the Gaussian Copula and Beta Kernel Copula. A definition of both of them are presented in the following.

Gaussian Copula

The Gaussian family of copulas, which are often used in hydrology [?], have a symmetrical density function which can be expressed through the correlation matrix Γ as

$$C_{\Gamma}(u_1, \dots, u_n) = \Phi_{\Gamma, n}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)) \tag{8.7}$$

Here

- $C_n(\vec{u})$: copula defined on the n-dimensional space
- $\Phi(x_i)$: Standard normal distribution function
- Γ : Covariance matrix

Beta Kernel Copula

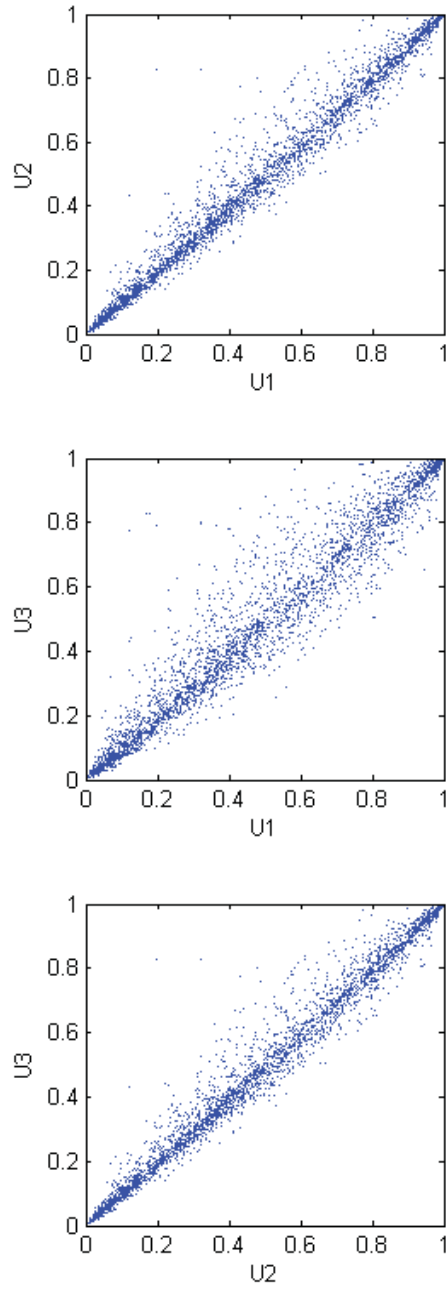


Figure 8.7: Empirical bivariate Copulas

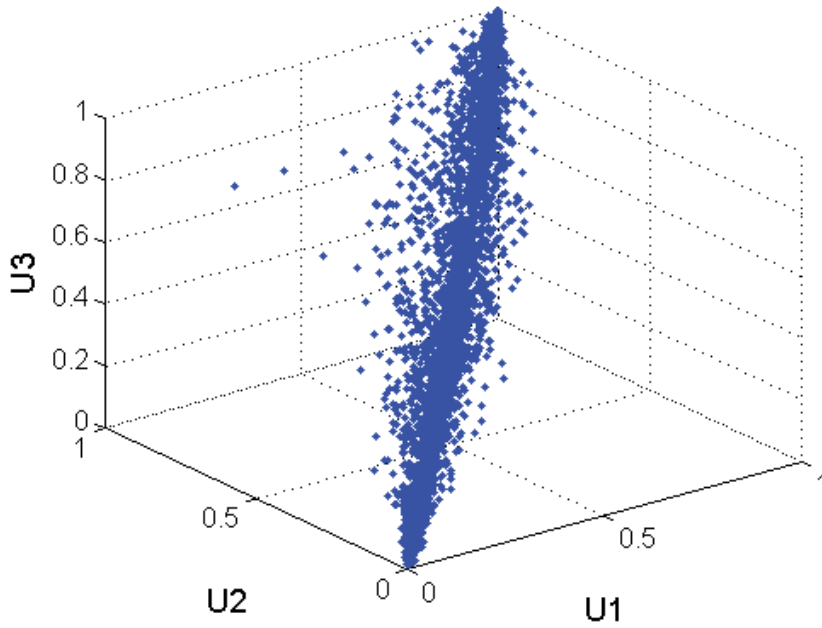


Figure 8.8: Empirical 3D Copulas

The beta kernel function can be written in terms of the beta function as

$$K(x, x_i; h) = \frac{x^{\frac{x_i}{h}} (1-x)^{\frac{1-x_i}{h}}}{B(\frac{x_i}{h} + 1, \frac{1-x_i}{h} + 1)} \quad (8.8)$$

By calling $\alpha = \frac{x_i}{h} + 1$ and $\beta = \frac{1-x_i}{h} + 1$, equation 8.8 can be expressed in a more compact way as

$$K(x, x_i; h) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}$$

where $B(\alpha, \beta)$ represents the Beta function. It is defined as

$$B(z, w) = \int_0^1 t^{z-1} (1-t)^{w-1} dt$$

The Beta function is closely connected to the Gamma function. Writing the former in terms of the latter the following relationship holds

$$B(z, w) = \frac{\Gamma(z) \Gamma(w)}{\Gamma(z+w)} \quad (8.9)$$

where the Gamma function is given by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, notation and name given by Legendre and which represent (in its original formulation) a generalization of the factorial. A prove of this equality (8.9) can be found for example in [75]. Alternatively the Beta function is termed the Euler integral of the first kind, and the Gamma function termed the Euler integral of the second kind.

By taking logarithms and thereafter applying the exponential function, equation 8.9 can be written in terms of a sum of individual Gamma functions as

$$B(z, w) = \exp(\ln\Gamma(z) + \ln\Gamma(w) - \ln\Gamma(z + w)) \tag{8.10}$$

In the special case that the argument of the Gamma function is a positive integer, it can be easily calculated as a factorial

$$\Gamma(m) = (m - 1)!,$$

result one can arrive at by integrating the Gamma function and expressing it in the recursive form $\Gamma(m + 1) = m\Gamma(m)$. The calculation of the Gamma function and hence the Beta function can be made via some algorithms as for example those outlined in [18] or directly by using some program languages as for example R or Matlab. For more detail refer to [75], [2] and [53].

In this way, calculating the Beta function, the Beta density estimator can be obtained by 8.11

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K(x, x_i; h) = \frac{1}{n} \sum_{i=1}^n \frac{x^{\frac{x_i}{h}} (1 - x)^{\frac{1-x_i}{h}}}{B\left(\frac{x_i}{h} + 1, \frac{1-x_i}{h} + 1\right)} = \hat{c}(x) \tag{8.11}$$

In the case of a multivariate kernel density estimation we can apply a multiplicative Kernel, which in the q-dimensional space is of the form

$$K(x) = K(x_1) K(x_2) \cdots K(x_k)$$

and the density estimator is given by

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \left\{ \prod_{j=1}^q K(x, x_i, h) \right\}$$

For example, in the case we want to fit a 3-dimentional Copula we have for $q = 3$,

$$\hat{f}(u) = c(u) = \frac{1}{n} \sum_{i=1}^n K(u_1, h) K(u_2, h) K(u_3, h)$$

where $K(\cdot)$ denotes the univariate Kernel function.

Correlation coefficient

For the estimation of the correlation coefficient, the well known relationship between the Kendall's tau and correlation coefficient is applied (equation 8.12) for the case of the normal distribution, although the result also holds for all non-degenerate elliptical distributions but with slightly modifications [50]. We have

$$\tau = \frac{2}{\pi} \arcsin(\rho) \tag{8.12}$$

from where the standard correlation coefficient ρ is calculated as

$$\rho = \sin\left(\frac{\pi}{2}\tau\right)$$

Kendal's tau correlation is a measure of concordance and in contrast to the Pearson correlation coefficient, is invariant on monotone transformations [87].

Table 8.3: Comparison of the errors (RMSE) using copulas and using the model AR2

Year	Error Copulas	Error AR2	e_Copula - e_AR2
1999	0.0640	0.0735	-0.0095
2000	0.0657	0.0803	-0.0146
2001	0.0724	0.0975	-0.0251
2002	0.0640	0.0739	-0.0099
2003	0.0601	0.0693	-0.0092
2004	0.0711	0.0821	-0.0110
2005	0.0622	0.0637	-0.0015
2006	0.0834	0.0915	-0.0081
2007	0.0801	0.0940	-0.0139
2008	0.0557	0.0611	-0.0054

For the estimation of the value (discharge) 1-day ahead, the following relationships was applied

$$\begin{aligned}
 P(X \prec x \mid X_i = x_i, i = 1, \dots, n) &= \int_0^u c(u \mid u_1, \dots, u_n) du \\
 &= \frac{\int_0^u c(u, u_1, \dots, u_n) du}{c(u_1, \dots, u_n)}
 \end{aligned} \tag{8.13}$$

In our case where a 3-dimensional vector is consider the previous expression reduces to

$$P(X \prec x \mid X_1 = x_1, X_2 = x_2) = \frac{\int_0^u c(u, u_2, u_3) du}{c(u_2, u_3)} \tag{8.14}$$

from where a cumulative distribution and hence a density was built. In this way, for each day and from its corresponding generated density function, the expected value of the variable for the next time step (day) was estimated considering the two models Gaussian Copula and Beta Kernel Copula. First the estimation was done for the year 2008 and subsequently this estimation was compared with the AR2 model outcomes in Santa Eulalia for the same year. For this comparison, the RMSE was calculated. The results showed a little improvement when considering the Gaussian Copula, where the associated errors (RMSE) equal $e_{Gaussian-Copula} = 0.0557$ and $e_{AR2-model} = 0.0611$. The situation did not repeat when comparing the $AR(2)$ model with the Beta kernel Copula, where a small deterioration in the accuracy of the estimation was observed with $e_{Beta-Kernel-Copula} = 0.0641$.

An example of the estimate together with the observed values can be seen in figure 8.9. Here the observed discharge and estimated discharge from the Gaussian Copulas for the year 1999 is depicted. In order to see how this estimation behaves in others periods and to compare it with the estimation given by the $AR(2)$ model, the same analysis and calculations were done for the others years with observations, this means, from 1999 to 2007. The estimations led to the same results in which the Gaussian Copula model delivered a little improvement when summing up the errors of the entire year. A summary of the results for the time period 1999 to 2008 are presented in table 8.3.

One of the results one would expect is that the errors were higher in the period of more discharge generation (beginning and end of the year) than the period of less discharge

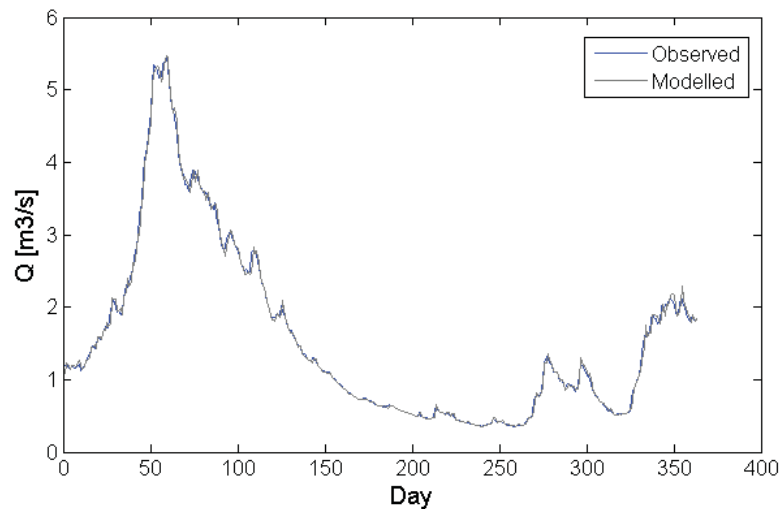


Figure 8.9: Observed and 1-day ahead estimated discharge for the year 1999, sub catchment Santa Eulalia and considering a Gaussian Copula model

generation (months in the middle part of the year), which is confirmed by analyzing the time series of the daily errors. Figure 8.10 shows these errors obtained from the applied Gaussian Copula model and for the year 1999.

Another result worthy of being analysed is the “position” of the modeled discharge (HBV) with respect to the generated discharge density function of the copula model, or how central are the modeled discharge values with respect to the calculated expected value of the discharge given by the copula-based autorregressive model. For this purpose two limits were defined, namely the 10% and the 90% quantil of the obtained density function of the copula model. In this way, the position of the modeled discharge was determined and whether or not it was between these limited. An example is shown in figure 8.11 where the 10% quantil and 90% quantil together with the modeled discharge are shown. The percentage of days where the modeled discharge lies between the two bounds resulted to be $P = 85.7\%$.

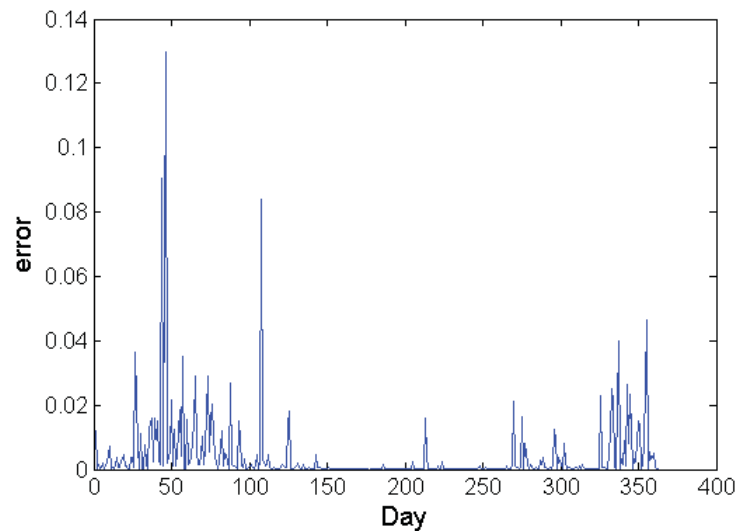


Figure 8.10: Individual daily errors between the Gaussian Copula model and the generated discharge for the year 1999

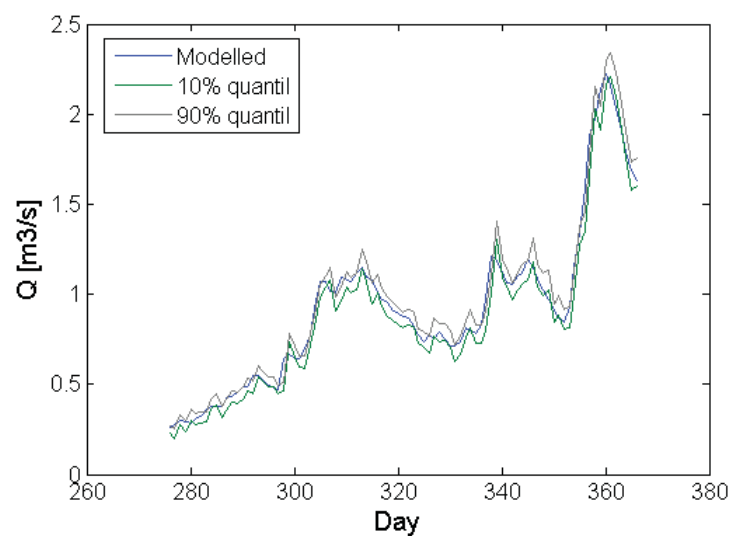


Figure 8.11: Modeled discharge, 10% quantil and 90% quantil for the last three months of the year 2008

8.4 Summary and Final Comments

In this chapter we have reviewed in a bit of detail some of the existing autorregressive models for calculating a variable 1-day ahead. The variable selected was the discharge in the subcatchment Santa Eulalia. Because discharge information consisted of monthly data, 1-day resolution time series were generated with the calibrated HBV model given rise to the generated discharge time series for the sub catchment.

We observed that both the estimation given by an AR(2) model and the estimation given by a copula-based model are similar in magnitude and represent the real values relatively well. The most significant differences are observed in the periods where the discharge is higher and where consecutive values do not follow a clear tendency and

bigger jumps are observed. Regarding the errors, the copula-based model shows a little more accuracy in the estimations, although the difference is not important.

Finally, it was shown that in most of the cases, the real value lies between the 10% and 90% quantile defined from the estimated density of the transformed estimation.

Chapter 9

Conclusions and Outlook

In this last chapter a summary of the overall conclusions of this thesis will be given, and a brief outlook for possible improvements of the work done here is presented.

9.1 Conclusions

This research was divided into several tasks which led us to the main goal: an estimate of the climate change impact in the areas analysed. So the main question to be answered was: What is to be expected for the main hydrological variables in the next 40 years? This was done for the variable temperature, precipitation and discharge where different tools were applied along this research.

The downscaling procedure based on Quantil-Quantil transformation offered a way of going from a coarse spatial scale to a finer spatial scale. The variables involved here were precipitation and temperature. The areas considered are subcatchments on the Pacific side as well as on the Atlantic side of the Andes range. As uncertainties are always present in many forms when speaking about the future, especially in the long term prediction, some different scenarios permitted us to estimate different possible outcomes. After performing the estimations, the future calculated variations are to be similar in all the analysed areas. The outcomes showed us that in most of the cases, a decrease in precipitation is expected. The only exception is the model Hadley in the possible scenario A1B (*HA1B*). In general, the differences depend on the model and scenario, but roughly they vary between $\Delta p = -10\%$ and $\Delta p = 7\%$.

Regarding the downscaling results for temperature, in all cases an increase in temperature is predicted and the magnitude of the variation depends on the model and scenario considered. To summarise, the models tell us that the future expected increase lies between $\Delta T = 1^\circ C$ to around $\Delta T = 2.6^\circ C$, where the minimum variation is predicted by the model Echem and scenario A2, and the maximum variation by the model Hadley and scenario A1B.

An important detail to be noted is the annual cycle presented for the time series of both precipitation and temperature. Independent of the model or scenario, the annual cycle is clearly recognizable and presents the same behaviour. This means that what is subject to variation is the amount but not the pattern (considering the scope of this thesis).

What is of major importance for the residents is what the models can say about the discharge and the expected variation this variable can undergo. A crucial issue here was the availability of discharge information for the process of calibration and validation.

Certainly, only in those points with enough information could these tasks be performed. Adaptations of the hydrological models HBV and Hymod were addressed and permitted the generation of the future monthly discharge. The areas considered are Santa Eulalia and Atarjea in the catchment Rimac, and Obrajillo in the catchment Chillón.

By calibrating and validating the model and the subsequent estimate of the future discharge, the results were not quite conclusive, which is analogous to the estimated precipitation for the same period. It is not conclusive in the sense that both an increase and a decrease resulted from the modeling. The most pessimist prediction is given by the model Echam and for the scenario A2 (*EA2*). The only positive scenario is given by the model Hadley and scenario A1B (*HA1B*). The predicted variations are a little more accentuated than those estimated for precipitation. The range fluctuates between a maximum negative variation $\Delta Q = -13.7\%$ in Atarjea and a maximum positive variation of $\Delta Q = 6.3\%$ in Santa Eulalia, on the Pacific side of the Andes. In Mantaro catchment (Marca Projects) the fluctuations lie between $\Delta Q = -11.3\%$ and $\Delta Q = 7.1\%$.

In order to have an order of magnitude of the variations in the discharge, the maximum negative change in Atarjea corresponds to a decrease in round $Q = 3.4 \text{ m}^3/\text{s}$ in the mean average discharge, with a corresponding expected maximum increase of $Q = 0.5 \text{ m}^3/\text{s}$. On the other hand, the estimated maximum decrease in Santa Eulalia corresponds to a discharge variation of $Q = -1.0 \text{ m}^3/\text{s}$ in the mean average discharge, with a corresponding maximum increase of $Q = 0.6 \text{ m}^3/\text{s}$.

As discharge observations were available only in the catchments Chillón and Rimac, a transfer of parameters from these two regions had to be attempted in order to estimate the discharge variation in Lurín and the areas in Mantaro defining the Marca projects. This parameter transfer was made between areas located in similar elevation on the one hand (case Santa Eulalia - Obrajillo) and in areas located in a clearly different elevation (case Santa Eulalia - Atarjea). The results showed that the performance of the model HBV is better in those areas of similar altitudes. This was clear when quantifying the RMSE, obtaining $e = 17.73 \text{ m}^3/\text{s}$ against $e = 0.84 \text{ m}^3/\text{s}$ for different and similar elevations respectively. With this result a sub catchment was defined in the upper part of Lurín and the HBV model applied. The results showed that in this new defined sub catchment the estimated expected variation in the discharge fluctuates between $\Delta Q = -10.6\%$ and $\Delta Q = 5.3\%$.

In summary, we can say that most of the scenarios analysed estimate a decrease in precipitation and discharge of about 10.0% and 13.7% respectively and all scenarios estimate an increase in temperature lying between 1 and 2.6 degrees. This conclusion was also attended by using a different method of optimization when calibrating the model HBV, namely the ROPE algorithm, in order to find a set of robust parameters in the hydrological modeling step.

Another point treated in this study was the 1-day ahead forecast. This was specially important for the largest electric power company of Peru, EDEGEL. When the store lakes, particularly those located in the upper part of Rimac are at full capacity, for example in the rain season, it is very important to have an estimate of the magnitude of the expected discharge for the next day. For this, a traditional autorregressive model was fit and compared with a copula based autorregressive model. Two different copula functions were used, namely the non parametric Beta Kernel function and the parametric Gaussian Copula function. The comparison between the traditional and copula based models showed a similar accuracy. The estimate was done for each year from 1999 to 2008 and each day. For any particular year, the daily estimation was sometimes better with

the AR process and sometimes with the copula-based model. However, when summing up the results of the year, the latter showed a little more accuracy than the former. This result repeated in each year analysed.

In this research the lack of appropriate information was pointed out. This refers not only to the absence of observations in locations where it would have been quite desirable to have them, but also to the mismatch between the different existing time series resolutions. This obliged, on the one hand, looking for alternative ways of estimating discharge by means of transfer of parameters between different areas, and on the other hand, generating discharges and calibrating them in different time scales. It would be good practice to obtain a better set of measurements and combine them with other sources of information, for example with satellite information in order to compare and especially improve the estimates done so far.

9.2 Outlook

There are important issues and questions which can be addressed after having performed the different methods, models and analyses in this study which could help us to understand and to improve the results obtained in this research:

1. Can the performance of the methods used in this study, the hydrological models outputs, and the predictions be improved?
2. Can we get a better understanding of the different hydrological processes present in the areas analysed?
3. Should we expect that the future will develop in a similar way as the predictions tell us it will?

To answer the first question we can refer to the availability of information in the regions. One of the first tasks performed in the study was the regionalization of precipitation where not only External Drift Kriging was applied but also some transformations were suggested. The results of regionalization were used in further steps, for example, in the application of the calibrated hydrological models for forecasting purposes. Hence the availability of information was essential for a good performance of the methods and estimates throughout the work. It would be very advisable and strongly suggested to perform the estimates with a better set of measurements, both in the spatial and temporal scale. It is also worth mentioning the poor time resolution in both the discharge and the water storage in lakes which differ significantly from that of precipitation and temperature. Addressing the second question, a better understanding can be achieved, for example, by regional simulation in a scale much finer than the offered by the GCM's. The higher resolution translates into a more detailed description of the processes and with this local features can be better captured, especially in those zones which are rather mountainous with a high variability in their characteristics. This also offers an alternative way for estimating climate changes over a desired region. This is clear by considering that when a certain pattern change, for example due to greenhouse effects, the way local features such as mountains interact with this also change. The third question was already addressed along the way in this thesis, especially when speaking about unpredictability and in the estimates of the future discharge. Several sources of uncertainties, in all the steps, cannot be removed from the analyses and the predictions are far from precise. This

is apparent when observing the different outcomes given by the GCM's used in this work. Uncertainty is already present in the input information. On the other hand, the consideration of different scenarios makes even more difficult to get an appropriate understanding and description of the future. All together, this obliges us to speak about possible future changes, understanding that other possible scenarios are also feasible.

Chapter 10

Bibliography

- [1] S. Ahmed and G. De Marsily. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resources Research*, 23(9):1717–1737, 1987.
- [2] G. Andrews, R. Askey, and R. Roy. *Special Functions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1999.
- [3] M. Armstrong. *Basic linear geostatistics*. Springer, 1998.
- [4] A. Bardossy. Introduction to geostatistics. stuttgart university, lecture book. 2010.
- [5] A. Bardossy and J. Li. Geostatistical interpolation using copulas. *Water Resources Research*, 44(7), 2008.
- [6] S. Bergstroem, V. Singh, et al. The hbv model. *Computer models of watershed hydrology.*, pages 443–476, 1995.
- [7] K. Beven. Changing ideas in hydrology. the case of physically based models. *Journal of hydrology*, 105(1):157–172, 1989.
- [8] K. Beven. *Rainfall runoff modelling: the primer*, volume 15. Wiley Chichester, 2001.
- [9] M. Bilodeau, F. Meyer, and M. Schmitt. *Space, structure and randomness: contributions in honor of Georges Matheron in the fields of geostatistics, random sets, and mathematical morphology*, volume 183. Springer, 2005.
- [10] G. Blöschl. *Statistical upscaling and downscaling in hydrology*. Wiley Online Library, 2005.
- [11] J. Boé, L. Terray, F. Habets, and E. Martin. Statistical and dynamical downscaling of the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, 27(12):1643–1655, 2007.
- [12] A. Bowman and A. Azzalini. Applied smoothing techniques for data analysis. 1997.
- [13] G. Box, G. Jenkins, and G. Reinsel. *Time series analysis: forecasting and control*. Wiley. com, 2013.
- [14] D. Boyle, H. Gupta, S. Sorooshian, V. Koren, Z. Zhang, and M. Smith. Toward improved streamflow forecasts: Value of semidistributed modeling. *Water Resources Research*, 37(11):2749–2759, 2001.

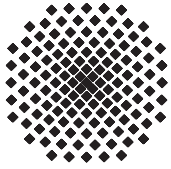
- [15] N. Cencov. Evaluation of an unknown distribution density from observations. *Soviet Mathematics*, 3:1559–1562, 1962.
- [16] C. Chatfield. *The analysis of time series: an introduction*. CRC press, 2003.
- [17] G. Christakos. On the problem of permissible covariance and variogram models. *Water Resources Research*, 20(2):251–265, 1984.
- [18] W. Cody. An overview of software development for special functions. In *Numerical Analysis*, pages 38–48. Springer, 1976.
- [19] N. Cressie. *Statistics for spatial data*. 1993.
- [20] P. Delfiner et al. *Geostatistics: modeling spatial uncertainty*, volume 497. Wiley.com, 2009.
- [21] C. Deutsch and A. Journel. *Geostatistical software library and user's guide*, volume 1996. Oxford university press New York, 1992.
- [22] P. Diggle and P. Hall. The selection of terms in an orthogonal series density estimator. *Journal of the American Statistical Association*, 81(393):230–233, 1986.
- [23] D. Donoho and M. Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, pages 1803–1827, 1992.
- [24] F. Durante and C. Sempi. Copula theory: an introduction. In *Copula theory and its applications*, pages 3–31. Springer, 2010.
- [25] S. Efromovich. Orthogonal series density estimation. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):467–476, 2010.
- [26] B. Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, pages 1–26, 1979.
- [27] B. Efron and R. Tibshirani. *An introduction to the bootstrap* (chapman & hall/crc monographs on statistics & applied probability). 1994.
- [28] G. Evin and A. Favre. A new rainfall model based on the neyman-scott process using cubic copulas. *Water Resources Research*, 44(3), 2008.
- [29] E. Frees and E. Valdez. Understanding relationships using copulas. *North American actuarial journal*, 2(1):1–25, 1998.
- [30] C. Genest and R. MacKay. Archimedean copulas and families of bidimensional laws for which the marginals are given. *Can. J. Stat*, 14:145–159, 1986.
- [31] W. Greiner, L. Neise, and H. Stroecker. *Thermodynamics and statistical mechanics*. Springer, 1999.
- [32] S. Grimaldi and F. Serinaldi. Design hyetograph analysis with 3-copula function. *Hydrological sciences journal*, 51(2):223–238, 2006.
- [33] H. Gupta, S. Sorooshian, T. Hogue, and D. Boyle. Advances in automatic calibration of watershed models. *Water Science and Application*, 6:9–28, 2003.

- [34] V. Gupta and S. Sorooshian. The relationship between data and the precision of parameter estimates of hydrologic models. *Journal of Hydrology*, 81(1):57–77, 1985.
- [35] H. Haken and H. C. Wolf. *Atom-und Quantenphysik: eine einfuehrung in die experimentellen und theoretischen Grundlagen*. Springer, 1980.
- [36] G. Hargreaves. Defining and using reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 120(6):1132–1139, 1994.
- [37] G. Hargreaves and Z. Samani. Reference crop evapotranspiration from ambient air temperature. *American Society of Agricultural Engineers*, 1985.
- [38] J. Hart. On the choice of a truncation point in fourier series density estimation. *Journal of Statistical Computation and Simulation*, 21(2):95–116, 1985.
- [39] C. Haslauer. Analysis of real-world spatial dependence of subsurface hydraulic properties using copulas with a focus on solute transport behaviour. 2011.
- [40] J. Hugg, E. Rafalin, K. Seyboth, and D. Souvaine. An experimental study of old and new depth measures. In *ALENEX*, pages 51–64, 2006.
- [41] Y. Hundecha and A. Bardossy. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of Hydrology*, 292(1):281–295, 2004.
- [42] Y. Hundecha Hirpa. Regionalization of parameters of a conceptual rainfall runoff model. 2005.
- [43] S. Kirkpatrick, D. G. Jr., and M. P. Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.
- [44] P. Kitanidis and R. Lane. Maximum likelihood parameter estimation of hydrologic spatial processes by the gauss-newton method. *Journal of Hydrology*, 79(1):53–71, 1985.
- [45] T. Kokkonen, A. Jakeman, P. Young, and H. Koivusalo. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the coveeta hydrologic laboratory, north carolina. *Hydrological Processes*, 17(11):2219–2238, 2003.
- [46] D. Kreider. *An introduction to linear analysis*. Addison-Wesley Pub. Co., 1966.
- [47] D. Krige and W. Kleingeld. The genesis of geostatistics in gold and diamond industries. In *Space, Structure and Randomness*, pages 5–16. Springer, 2005.
- [48] R. Kronmal and M. Tarter. The estimation of probability densities and cumulatives by fourier series methods. *Journal of the American Statistical Association*, pages 925–952, 1968.
- [49] L. Lin and M. Chen. Robust estimating equation based on statistical depth. *Statistical Papers*, 47(2):263–278, 2006.
- [50] F. Lindskog, A. McNeil, and U. Schmock. *Kendall tau for elliptical distributions*. Springer, 2003.

- [51] R. Liu. On a notion of data depth based on random simplices. *The Annals of Statistics*, 18(1):405–414, 1990.
- [52] R. Liu, J. Parelius, and K. Singh. Multivariate analysis by data depth: descriptive statistics, graphics and inference,(with discussion and a rejoinder by liu and singh). *The Annals of Statistics*, 27(3):783–858, 1999.
- [53] A. Mathai and H. Haubold. *Special functions for applied scientists*. Springer, 2008.
- [54] K. Miller, S. Ramaswami, P. Rousseeuw, J. A. Sellarès, D. Souvaine, I. Streinu, and A. Struyf. Efficient computation of location depth contours by methods of computational geometry. *Statistics and Computing*, 13(2):153–162, 2003.
- [55] R. Moore. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30(2):273–297, 1985.
- [56] E. Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
- [57] J. Nash and J. Sutcliffe. River flow forecasting through conceptual models part i. a discussion of principles. *Journal of hydrology*, 10(3):282–290, 1970.
- [58] R. Nelsen. *An introduction to copulas*. Springer, 1999.
- [59] H. Ohanian. *Modern Physics*. Prentice-Hall International, 1995.
- [60] S. Patil. *Regionalization of an event based Nash cascade model for flood predictions in ungauged basins*. PhD thesis, Zugl.: Stuttgart, Univ., Diss., 2008, 2008.
- [61] S. Rauscher, E. Coppola, C. Piani, and F. Giorgi. Resolution effects on regional climate model simulations of seasonal precipitation over europe. *Climate Dynamics*, 35(4):685–711, 2010.
- [62] P. Rousseeuw and A. Struyf. Computing location depth and regression depth in higher dimensions. *Statistics and Computing*, 8(3):193–203, 1998.
- [63] I. Ruts and P. Rousseeuw. Computing depth contours of bivariate point clouds. *Computational Statistics & Data Analysis*, 23(1):153–168, 1996.
- [64] F. Samper and J. Carrera. *Geoestadística: aplicaciones a la hidrología subterránea*. CIMNE. Barcelona. Spain, 1990.
- [65] F. Schwabl. *Statistical mechanics*, by franz schwabl. Technical report, ISBN 3-540-32343-0, Berlin. Springer, 2006.
- [66] B. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.
- [67] J. Simonoff. *Smoothing methods in statistics*. Springer, 1996.
- [68] S. Singh. *Robust parameter estimation in gauged and ungauged basins*. 2010.

- [69] M. Sivapalan, K. Takeuchi, S. Franks, V. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. McDonnell, E. Mendiondo, P. O'connell, et al. Iahs decade on predictions in ungauged basins (pub), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, 48(6):857–880, 2003.
- [70] M. Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959.
- [71] B. Sten and F. Arne. Development of a conceptual deterministic rainfall-runoff model. *Nordic hydrology*, 4(3):147–170, 1973.
- [72] M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):pp. 111–147, 1974.
- [73] S. Sverdlik. *Seis genios en el cielo*. Lugar Editorial, 2004.
- [74] M. Tarter and M. Lock. *Model-free curve estimation*, volume 56. CRC Press, 1993.
- [75] N. Temme. *Special functions: An introduction to the classical functions of mathematical physics*. John Wiley & Sons, 2011.
- [76] M. Themeßl and A. Gobiet. Empirical-statistical downscaling and model error correction of daily temperature and precipitation from regional climate simulations and the effects on climate change signals. In *EGU General Assembly Conference Abstracts*, volume 12, page 9664, 2010.
- [77] J. Tukey. Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians*, volume 2, pages 523–531, 1975.
- [78] T. Wagener, D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian. A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, 5(1):13–26, 1999.
- [79] T. Wagener and H. Wheater. *Rainfall-runoff modelling in gauged and ungauged catchments*. Imperial College Press, 2004.
- [80] G. Wahba. Data-based optimal smoothing of orthogonal series density estimates. *The Annals of Statistics*, pages 146–156, 1981.
- [81] M. M. P. Wand and M. C. Jones. *Kernel smoothing*, volume 60. Crc Press, 1995.
- [82] R. Webster and M. Oliver. *Geostatistics for environmental scientists*. John Wiley & Sons, 2007.
- [83] H. Wheater, A. Jakeman, and K. Beven. Progress and directions in rainfall-runoff modelling. 1993.
- [84] P. Whittle. On the smoothing of probability density functions. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 334–343, 1958.
- [85] D. Wick. *The infamous boundary: Seven decades of controversy in quantum physics*. Springer, 1995.

- [86] D. Wilks. Statistical methods in the atmospheric sciences, vol. 59 of international geophysics series, 1995.
- [87] I. Žežula. On multivariate gaussian copulas. *Journal of Statistical Planning and Inference*, 139(11):3942–3946, 2009.
- [88] L. Zhang and V. Singh. Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering*, 11(2):150–164, 2006.
- [89] Y. Zuo and R. Serfling. General notions of statistical depth function. *Annals of Statistics*, pages 461–482, 2000.



Institut für Wasser- und Umweltsystemmodellierung Universität Stuttgart

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

Direktoren

Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Prof. Dr.-Ing. Rainer Helmig
Prof. Dr.-Ing. Silke Wieprecht

Vorstand (Stand 03.11.2014)

Prof. Dr. rer. nat. Dr.-Ing. A. Bárdossy
Prof. Dr.-Ing. R. Helmig
Prof. Dr.-Ing. S. Wieprecht
Prof. Dr. J.A. Sander Huisman
Jürgen Braun, PhD
apl. Prof. Dr.-Ing. H. Class
Dr.-Ing. H.-P. Koschitzky
Dr.-Ing. M. Noack
Prof. Dr.-Ing. W. Nowak
Dr. rer. nat. J. Seidel
Dr.-Ing. K. Terheiden
Dr.-Ing. habil. Sergey Oladyshkin

Emeriti

Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Jürgen Giesecke
Prof. Dr.h.c. Dr.-Ing. E.h. Helmut Kobus, PhD

Lehrstuhl für Wasserbau und Wassermengenwirtschaft

Leiter: Prof. Dr.-Ing. Silke Wieprecht
Stellv.: Dr.-Ing. Kristina Terheiden
Versuchsanstalt für Wasserbau
Leiter: Dr.-Ing. Markus Noack

Lehrstuhl für Hydromechanik und Hydrosystemmodellierung

Leiter: Prof. Dr.-Ing. Rainer Helmig
Stellv.: apl. Prof. Dr.-Ing. Holger Class

Lehrstuhl für Hydrologie und Geohydrologie

Leiter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Stellv.: Dr. rer. nat. Jochen Seidel
Hydrogeophysik der Vadosen Zone
(mit Forschungszentrum Jülich)
Leiter: Prof. Dr. J.A. Sander Huisman

Lehrstuhl für Stochastische Simulation und Sicherheitsforschung für Hydrosysteme

Leiter: Prof. Dr.-Ing. Wolfgang Nowak
Stellv.: Dr.-Ing. habil. Sergey Oladyshkin

VEGAS, Versuchseinrichtung zur Grundwasser- und Altlastensanierung

Leitung: Jürgen Braun, PhD, AD
Dr.-Ing. Hans-Peter Koschitzky, AD

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 Flächen-tragwerke 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; Gesteigungskosten 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 Binnenschifffahrt durch Schwimmerhebwerke*, 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 rotations-symmetrisch 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 Axialwirbelkammverstärker 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 horizontalebene 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 Karst-aquiferen 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 instationärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1
- 111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X
- 112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten*, 2002, ISBN 3-933761-15-8, vergriffen
- 113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation*, 2002, ISBN 3-933761-16-6
- 114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung*, 2002, ISBN 3-933761-17-4
- 115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen*, 2003, ISBN 3-933761-18-2
- 116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung*, 2003, ISBN 3-933761-19-0
- 117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water*, 2003, ISBN 3-933761-20-4
- 118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2
- 119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0
- 120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9
- 121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7
- 122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -*, 2003, ISBN 3-933761-25-5
- 123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance*, 2003, ISBN 3-933761-26-3

- 124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003*, 2003, ISBN 3-933761-27-1
- 125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip*, 2003, ISBN 3-933761-28-0
- 126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8
- 127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1
- 128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X
- 129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flusstauhaltungen*, 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 Sub-surface - 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; Stuhmann, 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 eingeeerdeter 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 CO₂ 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 Götzing, 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; Stuhmann, 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; Stuhmann, 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 CO₂ 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 CO₂ 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 modeling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4
- 190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1
- 191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8
- 192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5
- 193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2
- 194 Bliedernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins*, 2011, ISBN 978-3-933761-98-9
- 195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -*, Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2010, ISBN 978-3-933761-99-6

- 196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia*, 2010, ISBN 978-3-942036-00-9
- 197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen*, 2010, ISBN 978-3-942036-01-6
- 198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins*, 2010, ISBN 978-3-942036-02-3
- 199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow*, 2011, ISBN 978-3-942036-03-0
- 200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation*, 2011, ISBN 978-3-942036-04-7
- 201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas*, 2011, ISBN 978-3-942036-05-4
- 202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen*, 2011, ISBN 978-3-942036-06-1
- 203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks*, 2011, ISBN 978-3-942036-07-8
- 204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology*, 2011, ISBN 978-3-942036-08-5
- 205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken*, Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2
- 206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour*, 2011, ISBN 978-3-942036-10-8
- 207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta*, 2011, ISBN 978-3-942036-11-5
- 208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam*, 2011, ISBN 978-3-942036-12-2
- 209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone*, 2012, ISBN 978-3-942036-13-9

- 210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking*, 2012, ISBN 978-3-942036-14-6
- 211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach*, 2012, ISBN 978-3-942036-15-3
- 212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media*, 2012, ISBN 978-3-942036-16-0
- 213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung*, 2012, ISBN 978-3-942036-17-7
- 214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish*, 2012, ISBN 978-3-942036-18-4
- 215 De Boer, Cjstmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1
- 216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7
- 217 Lebreuz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4
- 218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of CO₂ Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1
- 219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8
- 220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5
- 221 Walter, Lena: *Uncertainty studies and risk assessment for CO₂ storage in geological formations*, 2013, ISBN 978-3-942036-25-2
- 222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects*, 2013, ISBN 978-3-942036-26-9
- 223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes*, 2014, ISBN 978-3-942036-27-6
- 224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3
- 225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0

- 226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6
- 227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3
- 228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0
- 229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7
- 230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4
- 231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1
- 232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8
- 233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5
- 234 Zhang, Jin: *Investigations on Urban River Regulation and Ecological Rehabilitation Measures, Case of Shenzhen in China*, 2014, ISBN 978-3-942036-38-2
- 235 Siebel, Rüdiger: *Experimentelle Untersuchungen zur hydrodynamischen Belastung und Standsicherheit von Deckwerken an überströmbaren Erddämmen*, 2014, ISBN 978-3-942036-39-9
- 236 Baber, Katherina: *Coupling free flow and flow in porous media in biological and technical applications: From a simple to a complex interface description*, 2014, ISBN 978-3-942036-40-5
- 237 Nuske, Klaus Philipp: *Beyond Local Equilibrium — Relaxing local equilibrium assumptions in multiphase flow in porous media*, 2014, ISBN 978-3-942036-41-2
- 238 Geiges, Andreas: *Efficient concepts for optimal experimental design in nonlinear environmental systems*, 2014, ISBN 978-3-942036-42-9
- 239 Schwenck, Nicolas: *An XFEM-Based Model for Fluid Flow in Fractured Porous Media*, 2014, ISBN 978-3-942036-43-6
- 240 Chamorro Chávez, Alejandro: *Stochastic and hydrological modelling for climate change prediction in the Lima region, Peru*, 2015, ISBN 978-3-942036-44-3

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.