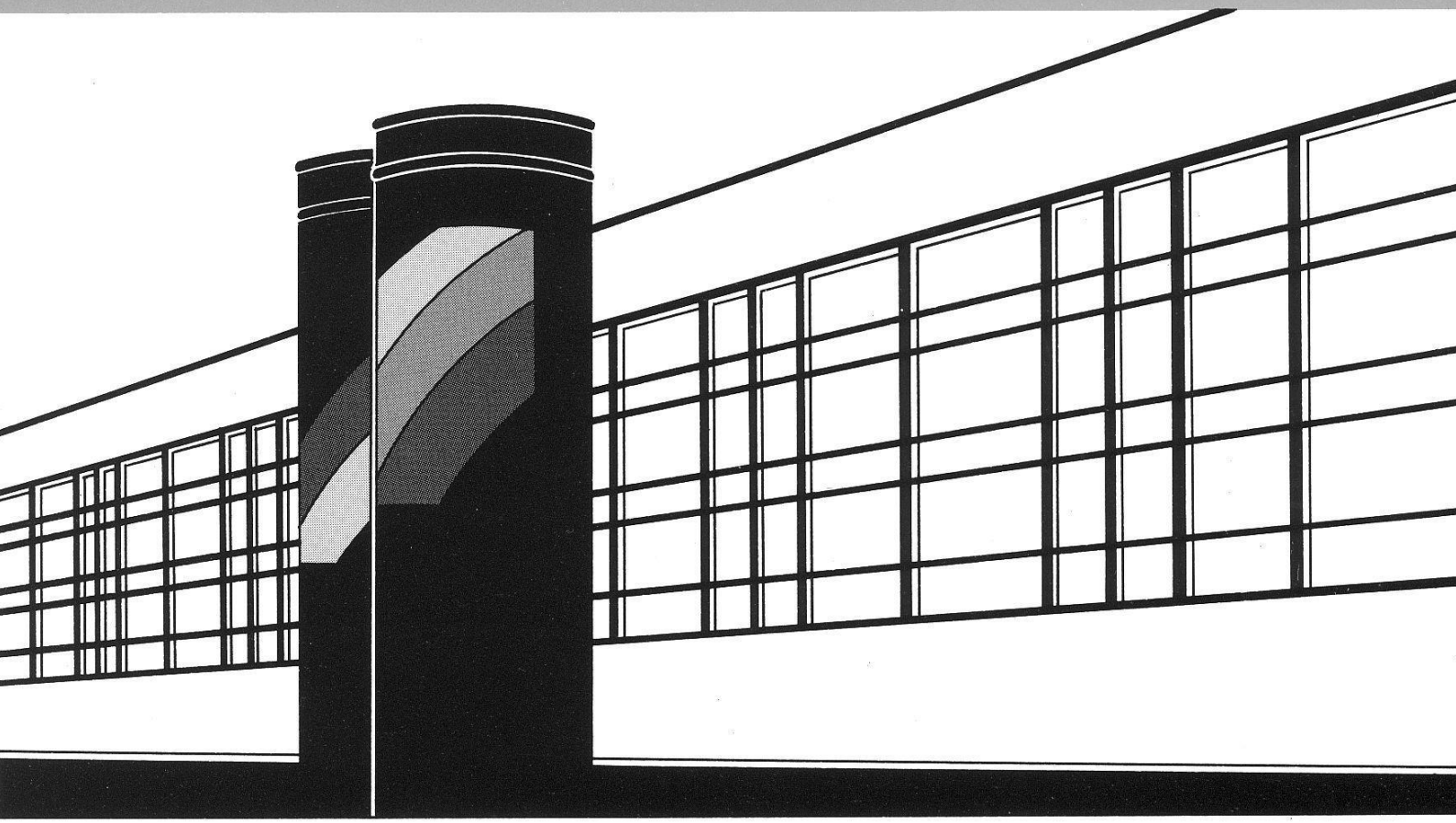
Institut für Wasserbau · Universität Stuttgart

Mitteilungen



Heft 207 Nguyen Viet Dung

Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta

Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta

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

Vorgelegt von

Nguyen Viet Dung

aus Thanh Hoa, Vietnam

Hauptberichter:	Prof. Dr. rer.nat. Dr.-Ing. habil. András Bárdossy
Mitberichter:	Prof. Dr. rer. nat. habil. Dr.-Ing. Bruno Merz
Tag der mündlichen Prüfung:	16. September 2011

Institut für Wasserbau
der Universität Stuttgart
2011

Heft 207

Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta

von
Dr.-Ing.
Nguyen Viet Dung

Eigenverlag des Instituts für Wasserbau
der Universität Stuttgart
2011

D93 Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta

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

Nguyen Viet Dung:
Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta / von Nguyen Viet Dung. Institut für Wasserbau,
Universität Stuttgart. - Stuttgart: Inst. für Wasserbau, 2011

(Mitteilungen / Institut für Wasserbau, Universität Stuttgart: H. 207).
Zugl.: Stuttgart, Univ., Diss., 2011
ISBN 978-3-942036-11-5
NE: Institut für Wasserbau <Stuttgart>: Mitteilungen

Herausgegeben 2011 vom Eigenverlag des Instituts für Wasserbau

Druck: Document Center S. Kästl, Ostfildern

Acknowledgements

I express my most sincere thanks to Prof. Dr. Dr. András Bárdossy, my main supervisor for his overall guidance and encouragement. I am also sincerely grateful to Prof. Dr. Dr. Bruno Merz, my co-supervisor for his warm-hearted support and invaluable suggestions during this study. I am honored to be one of their students. I owe my deepest gratitude to Dr. Heiko Apel, who provided solid support throughout my research. And I thank Jun.-Prof. Dr. Oliver Röhrle for his chairmanship of my defense.

I would like to acknowledge ENWAT International Doctoral Program of University of Stuttgart for offering me the academic program. A lot of thanks go to Dr. Gabriele Hartmann for her help. I thank my friends in ENWAT, University of Stuttgart.

This study was conducted within the framework of the WISDOM project (Water related Information System for a Sustainable Development of the Mekong Delta). I would like to thank the German Federal Ministry of Education and Research (BMBF), the Vietnamese Ministry of Science and Technology (MOST), the United Nations University, Bonn, Germany (UNU-EHS), and the German Research Centre for Geosciences (GFZ) in Potsdam for financial support and a convenient study environment in Germany. I would like to thank the Southern Institute of Water Resources Research (SIWRR) for offering me a chance to upgrade my study abroad. My thanks go to Prof. Dr. Le Manh Hung, Prof. Dr. Tang Duc Thang, Prof. Dr. Vo Khac Tri, Dr. Trinh Thi Long for their continuous encouragement.

I have had chance to work in the same group with Nguyen Nghia Hung and Jose Delgado in GFZ. We are the three musketeers. Our daily topic of discussion centered not only on science but also football! Many thanks go to them.

I do not have enough space here to list all of their names, but I sincerely thank all my colleagues in the Section 5.4 Hydrology, GFZ in Potsdam.

I am indebted to all the people who have helped me but I may have forgotten to say “thank you”.

Last but not least, as my greatest thanks to them, I dedicate this thesis to my family, to my parents, to my wife - Minh Tan and my two beloved little daughters - Luna Nguyet Minh and Sunni Nhat Minh for their understanding and for sharing my happiness and sadness during this long study in Germany. I will never be far from them again.

Potsdam, 2011

Nguyen Viet Dung

Author's declaration

I prepared this dissertation without illegal assistance. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. This dissertation has not been presented to any other University for examination, neither in Germany nor in another country.

Potsdam, 2011

Nguyen Viet Dung

Table of Contents

List of figures

List of tables

Abstract

Automatic and multi-objective calibration of hydrodynamic models is still underdeveloped, in particular, in comparison with other fields such as hydrological modeling. This is for several reasons: lack of appropriate data, the high degree of computational time demanded, and a suitable framework. These aspects are aggravated in large-scale applications. There are recent developments, however, that improve both the data and the computing constraints. Remote sensing, especially radar-based techniques, provide highly valuable information on flood extents, and in case high precision Digital Elevation Models (DEMs) are present, also on spatially distributed inundation depths. With regards to computation, the use of parallelization techniques brings significant performance gains. In the presented study, we build on these developments by calibrating a large-scale one-dimensional hydrodynamic model of the whole Mekong Delta downstream of Kratie in Cambodia: We combine in-situ data from a network of river gauging stations, i.e. data with high-temporal but low-spatial resolution, with a series of inundation maps derived from ENVISAT Advanced Synthetic Aperture Radar (ASAR) satellite images, i.e. data with low-temporal but high-spatial resolution, in a multi-objective automatic calibration process. It is shown that this kind of calibration of hydrodynamic models is possible, even in an area as large-scale and complex as the Mekong Delta. Furthermore, the calibration process reveals deficiencies in the model structure, i.e. the representation of the dike system in Vietnam, which would be difficult to detect by a standard manual calibration procedure.

In the last part of the dissertation the established hydrodynamic model is combined with flood frequency analysis in order to assess the flood hazard in the Mekong Delta. It is now common to state that climate change can lead to a change in flood hazard. Starting from this assumption, this study develops a novel approach for flood hazard mapping in the Mekong Delta. Typically, flood frequency analysis assumes stationarity and is limited to extreme value statistics of flood peaks. Both, the stationarity assumption and the limitation to univariate frequency analysis remain doubtful in the case of the Mekong Delta, because of changes in hydrologic variability and because of the large relevance of the flood volume for the impact of flooding. Thus, besides the use of the traditional approach for flood frequency analysis, this study takes non-stationarity and bivariate behavior into account. Copula-based bivariate analysis is used to model the dependence and to generate pairs of maximum discharge and volume, by coupling their marginal distributions to gain a bivariate distribution. In addition, based on cluster analysis, groups of characteristic hydrographs are identified and synthetic flood hydrographs are generated. These hydrographs are the input for the calibrated large-scale hydrodynamic model of the Mekong

Delta, resulting in flood hazard maps for the whole Mekong Delta. To account for uncertainty within the hazard assessment, a Monte Carlo framework is applied yielding probabilistic hazard maps.

Keywords: flood modeling, multi-objective optimization, automatic calibration, flood hazard, non-stationary, flood frequency analysis, Mekong Delta

Zusammenfassung

Hochwasser ist rezent eine der häufigsten Naturgefahren. Zudem herrscht weitgehend Einigkeit darüber, dass das Hochwasserrisiko generell durch Klimawandel, demographische und ökonomische Entwicklung in Zukunft weiter steigen wird, insbesondere in Küstenregionen und Ästuaren (Merz et al., 2008; Apel et al., 2009). Dadurch wird die Notwendigkeit der verlässlichen Abschätzung von Hochwassergefährdung und –risiken als Grundlage eines effektiven Hochwasserrisikomanagements weiter steigen.

Eine Kerntechnologie in der Abschätzung der Hochwassergefährdung und der Hochwasservorhersage ist die numerische Simulation von Überflutungsvorgängen (Apel et al., 2004; Vorogushyn et al., 2010). Hierbei muss für jeden Anwendungsfall basierend auf wissenschaftlichen und technischen Aspekten ein geeigneter Modellierungsansatz gewählt werden. Die vorliegende Studie beschäftigt sich mit den regelmäßigen Hochwässern im Mekong Delta, bei denen alljährlich große Flächen überflutet werden. Durch die niedrige Topographie und Meeresnähe ist das Mekong Delta durch die Auswirkungen des Klimawandels besonders gefährdet. Das Delta stellt durch die Vielzahl von künstlich angelegten Kanälen und Regulierungsbauten ein hydraulisch hochkomplexes System dar. Um dieser Komplexität und den großen Ausmaßen des Delta gerecht zu werden, wurden die Überflutungsvorgänge mittels eines ein-dimensionalen Modellansatzes quasi zwei-dimensional modelliert. Allerdings wurde das Modell bislang nur unzureichend manuell kalibriert und validiert, so dass die Modellergebnisse, insbesondere die Berechnung der Überflutungsflächen, mit großen Unsicherheiten belegt sind. An dieser Schwachstelle, die hydraulischen Modellen allgemein inhärent ist, setzt diese Studie an. Es wurde eine automatische und multikriterielle Kalibrierungsroutine entwickelt und implementiert und auf das Mekong Delta Modell angewandt.

Automatische Kalibrierung von hydraulischen Modellen ist generell kaum entwickelt, bzw. wir aufgrund fehlender Werkzeuge, geringer Kalibrierungs- und Validierungsdaten und hoher Rechenzeiten kaum angewandt, weder in Wissenschaft noch Praxis. In dieser Studie wurden aktuelle Entwicklungen in Kalibrierungsalgorithmen, beschleunigte Rechenzeiten durch Parallelisierung und verbesserter Dokumentation von Überflutungsvorgängen aufgegriffen und zu einem automatischen und multikriteriellen Kalibrierungsschema für hydraulische Modelle kombiniert. Zur Kalibrierung wurden räumlich explizite Überflutungskarten aus Satellitenbeobachtungen (ENVISAT ASAR) mit geringer zeitlicher Auflösung und zeitlich hochauflösende, aber räumlich begrenzt aussagekräftige Hydrographen verwendet. Daraus ergeben sich zwei Kalibrierungskriterien und die Möglichkeit, die Modellierungsergebnisse

hinsichtlich ihrer zeitlichen und räumlichen Dynamik zu bewerten. Die Gütefunktion für das erste Kriterium F1, die zeitliche Dynamik in den Gerinnen, wurde analog zu dem in der Hydrologie bekannten Nash-Sutcliffe Gütekriterium definiert. Die zweite Gütefunktion F2 bewertet die Simulation der Überflutungsausdehnung mit Hilfe des floodareaindex. In der Auswertung von F2 wurden zudem die Unsicherheiten in der Überflutungssimulation (bedingt durch Modellaufbau und vereinfachte räumliche Simulation) und auch in den kartierten Überflutungsflächen (durch Klassifizierungsfehler und räumlicher Referenzierung) in einem Monte-Carlo Verfahren berücksichtigt.

Das Ziel der Kalibrierung war die Bestimmung derjenigen Parametersätze, die beide Gütekriterien bestmöglich erfüllen. Die Kalibrierung erfolgte über die Anpassung der Oberflächenrauhigkeitsparameter. Diese wurden zur Reduzierung der Freiheitsgrade in verschiedene Gruppen klassifiziert, die auf dem vorherrschenden hydraulischen Regime und Charakteristik der verschiedenen Regionen im Delta beruhen. Die automatische Kalibrierung erfolgte mittels NSGA-II, einem globalen multikriteriellen genetischen Optimierungsalgorithmus. Der Algorithmus wurde durch eine Master-Slave Parallelisierung auf einem Multiprozessor-Rechenserver implementiert, um die erforderliche Anzahl der Modellläufe in akzeptabler Zeit zu realisieren. Das gesamte Kalibrierungsschema ist in Abbildung I illustriert.

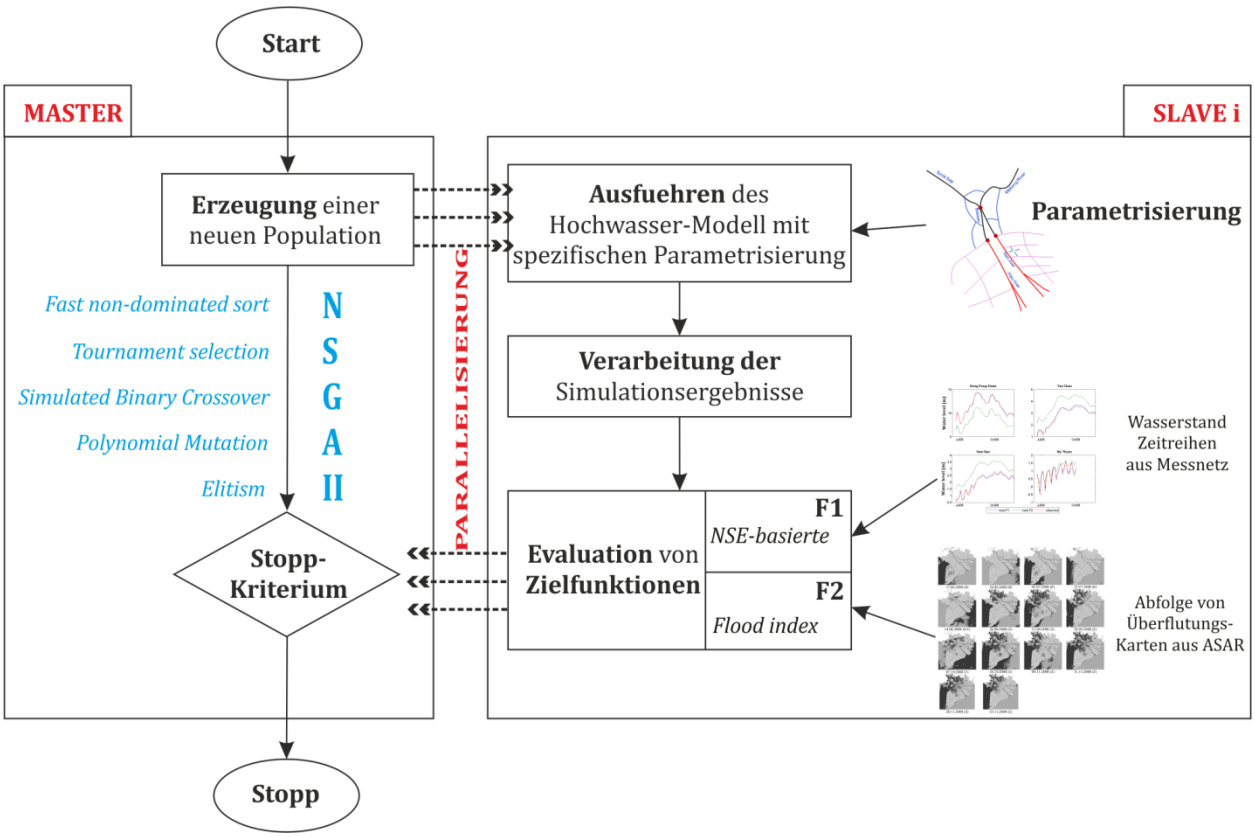


Abbildung I: Schema der automatischen multikriteriellen Kalibrierung

Folgende Schlüsse konnten aus der erfolgreichen Kalibrierung des Modells gezogen werden:

- **Erstens**: Die entwickelte Methode ermöglicht eine automatische und multikriterielle Kalibrierung von hydraulischen Modellen, auch von solcher Größe und Komplexität wie das vorgestellte Modell für das Mekong Delta. Dies ist ein wichtiger Schritt hin zu einer objektiveren Kalibrierung und Bewertung hydraulischer Modelle und deren Ergebnisse.
- **Zweitens:** Die Kombination von Überflutungskarten und Hydrographen ermöglicht eine gleichzeitige Auswertung der Simulationsergebnisse hinsichtlich ihrer räumlichen und zeitlichen Genauigkeit.
- **Drittens:** Die vorgestellte Kalibrierungsmethode identifiziert Parametersätze, die einen Kompromiss zwischen der Erfüllung der beiden Gütekriterien darstellen (Pareto-optimale Lösungen). Im vorgestellten Fallbeispiel heißt dass, das gute Ergebnisse in einem Kriterium nur auf Kosten des anderen Kriteriums zu erzielen sind. Daher konnte auch kein allgemein gültiger Parametersatz identifiziert werden, der beide Kriterien zufriedenstellend erfüllt. Das heißt in Konsequenz, dass die Parametersätze je nach Anwendungszweck des Modells mit Bedacht gewählt werden müssen. Der paretobasierte Ansatz erlaubt dem Nutzer unter gleichwertigen Parametersätzen zu entscheiden, aber auch entweder hinsichtlich der zeitlichen Dynamik in den Gerinnen oder der räumlichen Ausdehnung der Überflutungsflächen zu gewichten. Für den Fall, dass ausgeglichen Simlationsergebnisse gewünscht sind, wird die Wahl desjenigen Parametersatzes mit der geringsten euklidische Distanz zum theoretisch optimalen Modellempfohlen.
- **Viertens:** Durch den paretobasierten Ansatz konnten strukturelle Defizite im Modell identifiziert werden, was durch einen aggregierten oder gar manuellen Ansatz der Kalibrierung nicht möglich gewesen wäre. Durch die Analyse der besten Ergebnisse für die beiden Optimierungskriterien wurde klar, dass die Deichhöhen im vietnamesischen Teil des Deltas fehlerhaft sein müssen. Dieser durch die automatische Kalibrierung erzielte Schluss wurde weiter durch eine Simulation mit synthetisch erniedrigten Deichen bestätigt. Bei einer angenommenen Reduzierung der Deichhöhen um 20% konnte verbesserte Ergebnisse in beiden Kriterien erzielt werden. Die gleichwertigen Pareto-optimalen Lösungen sind auch ein Maß für die Unsicherheit in der Simulation. Diese kann z.B. durch Ensemblesimulationen oder durch Unsicherheitsbänder und Quantilkarten der Überflutungsflächen ausgedrückt werden.

Im zweiten Teil der Dissertation wird das kalibrierte Modell für eine Hochwassergefährdungsanalyse für das Mekong Delta angewendet. Grundlage für die Gefährdungsanalyse sind verschiedene extremwertstatistische Verfahren, aus denen die Randbedingungen für das hydraulische Modell abgeleitet werden. Hierbei werden verschiedene

Verfahren und Annahmen getestet, um den zu erwartenden Einflüssen des Klimawandels, beobachteten und zu erwartenden Instationaritäten in den Zeitreihen, und Unsicherheiten in der Ableitung der Randbedingung und Überflutungssimulation Rechnung zu tragen.

In Hochwassergefährdungsanalysen ist der Abfluss die wesentliche zu analysierende Komponente. Dies geschieht üblicherweise mittels stationärer Extremwertstatistiken, die fundamental die Stationarität und Unabhängigkeit der Zeitreihen voraussetzen. Diese Voraussetzungen werden oftmals von den zu analysierenden Zeitreihen nicht erfüllt, insbesondere mit Hinblick auf den Einfluss von Klimawandel auf die Abflussbildung (Strupczewski et al., 2001; Khaliq et al., 2006; Adlouni et al., 2007). Daher wird in der vorliegenden Arbeit neben dem üblichen stationären Ansatz auch nicht-stationäre Methoden angewandt und verglichen. Da im Mekong Delta nicht nur die jährliche Abflussspitze die Überflutungsintensität bestimmt, sondern auch die Volumen des Hochwassers und die Form des Hochwasserhydrographen, werden diese Faktoren ebenfalls mit in die Analyse einbezogen. Hieraus ergeben sich vier verschiedene Ansätze zur Hochwassergefährdungsanalyse, die miteinander verglichen werden:

- Stationär und univariat (Spitzenabfluss und Volumen unabhängig)
- Stationär und bivariat (Spitzenabfluss und Volumen in Abhängigkeit voneinander)
- Nicht-stationär und univariat (Spitzenabfluss und Volumen unabhängig)
- Nicht-stationär und bivariat (Spitzenabfluss und Volumen in Abhängigkeit voneinander)

Die bivariaten Analysen wurden mittels eines Copula-basierten Ansatzes realisiert. Die dritte Einflussgröße, der Abflusshydrograph, wurde mittels einer Clusteranalyse der beobachteten Abflussganglinien und der Identifikation charakteristischer Hydrographen mit in die Analyse einbezogen. Hierbei wurden die standardisierten typischen Hydrographen mit Eintrittswahrscheinlichkeiten belegt und durch die in der Extremwertstatistik ermittelten Abflussspitzen und Volumen zu synthetischen Hochwassern für definierte Wiederkehrintervalle skaliert. Diese synthetischen Hochwasser steuerten dann als Randbedingung das hydraulische Modell, mit dem Überflutungskarten generiert werden konnten. Unsicherheiten, die sich aus der Definition der Randbedingung (synthetische Hydrographen) ergeben, wurden mit Hilfe eines Monte-Carlo Ansatzes und Latin-Hypercubesamplings in probabilistische Überflutungskarten (Quantilkarten der Überflutungstiefen) überführt. Abbildung II illustriert das gesamte Vorgehen der Gefährdungsabschätzung und den Vergleich der verschiedenen Ansätze. Durch den Methodenvergleich konnten die Fehler, die aus vereinfachten und im Grunde nicht anwendbaren statistischen Ansätzen resultieren, quantifiziert und diskutiert werden.

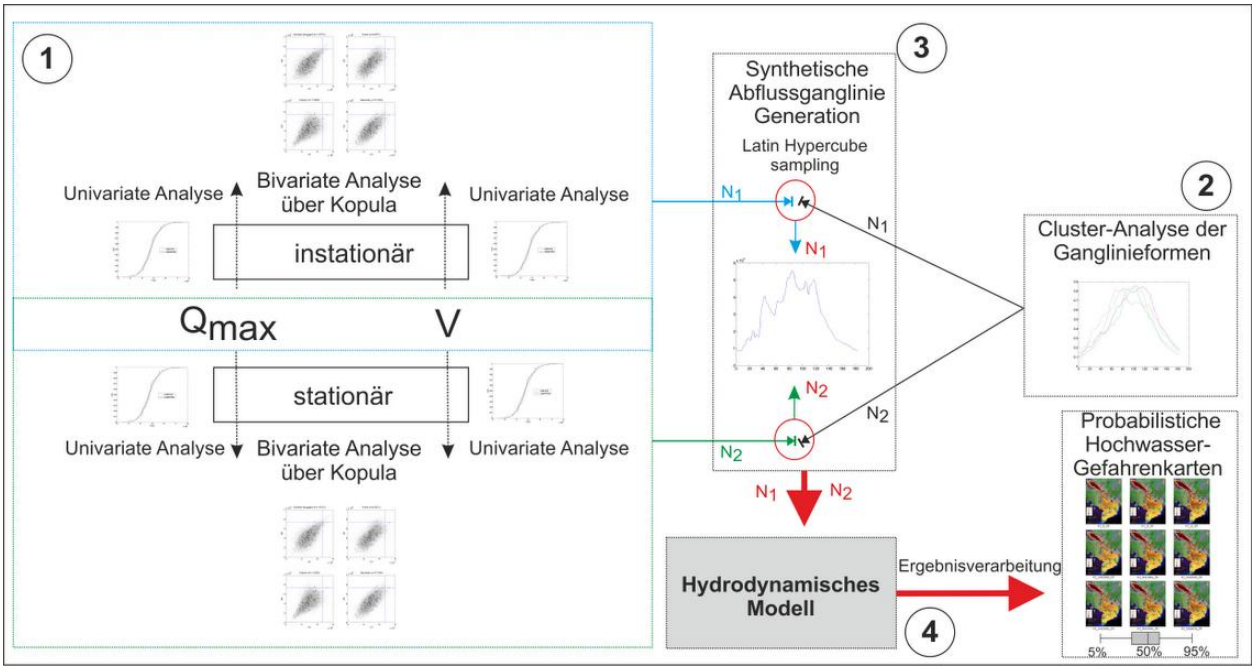


Abbildung II: Schema der Hochwassergefährdungsanalyse und Methodenvergleich für das Mekong Delta: (1) Extremwertstatistik der Spitzenabflüsse und Hochwasservolumen, (2) Clusteranalyse der Hochwasserganglinien, (3) Generierung der synthetischen Hochwasserganglinien und zugehörige Eintrittswahrscheinlichkeiten, (4) Hochwassergefährdungskarten

Statistische Tests ergaben, dass sowohl die Verteilung der jährlichen Abflussspitzen als auch des Hochwasservolumens am besten mit einer drei-parametrigen Log-Normalverteilung (LN3) modelliert werden können. Die univariate nicht-stationäre Häufigkeitsanalyse ergab, dass beide Zeitreihen einen negativen Trend im Lokationsparameter und einen positiven Trend im Skalenparameter aufweisen. Die Parameter der Copulas für die bivariaten Analysen wurden mit der zweischrittigen „inferencefrommargin“ (IFM) Methode geschätzt. Als geeignetes Copulamodell wurde die Gauß’scheCopula identifiziert, sowohl für die stationäre und nicht-stationäre Analyse. Anhand der Copulas wurden Abflussspitzen-Volumen-Paare für folgende Szenarien ermittelt:

(1)100-jähriges Ereignis, stationär, multivariat

(2)100-jähriges Ereignis, nicht-stationär für das Referenzjahr 2009, multivariat

(3)100-jähriges Ereignis, nicht-stationär für das Referenzjahr 2030, multivariat

Die Analyse der Hochwasserganglinien ergab vier charakteristische Hydrographen, dem jeweils eine Eintrittswahrscheinlichkeit zugeordnet wurde. Diese Hydrographen und Wahrscheinlichkeiten wurden mit den Spitzenabfluss und Volumenpaaren in einer Monte-Carlo Simulation kombiniert.

Jeder der so ermittelten synthetischen Hydrographen wurde durch das hydraulische Modell des Mekong Delta in Überflutungsflächen überführt. Hierbei wurden zwei Modellparametrisierungen verwendet: diejenige, die die besten Ergebnisse hinsichtlich der Überflutungsflächen ergab und diejenige mit der geringsten euklidischen Distanz zur optimalen Parametrisierung. Durch die Vielzahl der Monte-Carlo Simulationen konnten für die einzelnen Szenarien probabilistische Hochwassergefährdungskarten, die die maximalen Überflutungstiefen darstellen, generiert werden. Die Überflutungswahrscheinlichkeiten wurden hierbei durch 5%, 50% und 95% Quantilkarten dargestellt. Aus den Karten und dem Vergleich der verschiedenen Szenarien und statistischen Ansätze konnten folgende Schlüsse gezogen werden:

- **Erstens:** mit dem entwickelten innovativen Ansatz konnte die Hochwassergefährdung einschließlich der inhärenten Unsicherheiten analysiert werden.
- **Zweitens:** Änderungen in der oberen Randbedingung des Deltas – des Abflussgeschehens in Kratie – haben einen direkten Einfluss auf die Überflutungen im kambodschanischen Teil des Deltas, als auch auf den nördlichen Teil des vietnamesischen Deltas. Die Auswirkungen auf die Überflutung in den küstennahen Bereichen, insbesondere der östlichen, sind allerdings gering.
- **Drittens:** Die Unterschiede in den Überflutungstiefen zwischen den verschiedenen statistischen Ansätzen (stationär und nicht-stationär für zwei Referenzjahre) sind gering. Das heißt in Konsequenz, dass ein stationärer bivariater Ansatz für den präsentierten Fall und eine mäßige Zukunftsprojektion ausreichend ist, trotz der Verletzung der zu Grunde liegenden statistischen Annahmen und den identifizierten Trends. Dies gilt aber zunächst nur für Ereignisse bis zu einem Wiederkehrintervall von 100 Jahren und ist durch die geringe Ausprägung der Asymmetrie der Verteilungsfunktionen begründet. Das heißt, dass exteme Ereignisse im Mekong Delta sich nicht in Größenordnungen von den mittleren Ereignisen unterscheiden. Dies indiziert die "Normalität" der Hochwasser im Mekong Delta auch von einem statistischen Gesichtspunkt. Dennoch sollte für eine weiter reichende Zukunftsprojektion ein nicht-stationärer Ansatz verwendet werden, da die zu erwartenden Einflüsse des Klimawandels sich in weiterer Zukunft stärker auf die Hochwasserverteilung auswirken dürften.Eine Möglichkeit weiter reichender Zukunftsprojektionen besteht in der Einbeziehung von Simulationsensemblen globaler Zirkulationsmodelle und der Korrelation der Parameter der Extremwertstatistiken mit Klimaindizes (Coles, 2001; Delgado et al. 2011).
- **Viertens:** Im Gegensatz zu den verschiedenen statistischen Ansätzen ergaben die verschiedenen Parametrisierungen des hydraulischen Modells signifikante Unterschiede: Die Karten, die mit der besten F2-Parameztrisierung erzeugt wurden, zeigten tiefere überflutungstiefen und größere Ausdehnungen der Überflutungsflächen im Vergleich zur zwischen den Optimierungskriterien ausgleichenden Parametrisierung. Durch die

Verwendung mehrerer Pareto-optimaler Parametrisierungen kann die Unsicherheit, die durch das hydraulische Modell hervorgerufen wird, quantifiziert und illustriert werden.

- **Fünftens:** Ein Vergleich der Unsicherheiten in der Bestimmung der Randbedingungen für das Modell und der Modellparametrisierung ergab, dass die Unsicherheit in den Randbedingungen wesentlich größer ist als die des Modells. Dies wurde durch den Vergleich der Quantilkarten eines jeden Szenarios deutlich.
- **Sechstens:** Die abgeleiteten Quantilkarten der Überflutungstiefen (probabilistische Hochwassergefährdungskarten) bieten eine gute Grundlage für ein probabilistischen Hochwasserrisikomanagement, das die Unsicherheit in der Bestimmung extremer Hochwasserereignisse berücksichtigt.

Ableitend aus den dargestellten Schlüssen wird empfohlen, einen Copula-basierten Ansatz für bivariate Hochwassergefährdungsanalysen zu verwenden. Dies ist sowohl mit stationären und nicht-stationären statistischen Ansätzen möglich und bietet eine große Flexibilität hinsichtlich der verwendbaren extremwertstatistischen Modelle.

Schlüsselworte: Überflutungssimulation, multi-objektive automatische Kalibrierung, Hochwassergefährdung, nicht-stationär, Hochwasserhäufigkeitsanalyse, Mekong Delta

Chapter 1

INTRODUCTION

"Water is the driver of Nature."

(Leonardo da Vinci)

1.1 Background and problem identification

Flooding, often seen as a destructive abundance of water (Kundzewicz and Robson, 2004), is one of the most damaging catastrophes among natural disasters. However, flooding in some specific regions in the world (e.g. the Mekong Delta) should be perceived in a wider view. Floods can be not just damaging, they can also be beneficial by, for example, bringing sediments and nutrients to fields of flooded regions where agriculture is the main economic activity. In fact, floods have been one important factor in the development of modern civilization by enabling productive agricultural systems and the consequent organization and development of civilization, as documented e.g. in the early Persian and Egypt empires. However, recently, owning to the effects of climate change and other factors like demographic and economic developments, it is widely acknowledged that an increasing number of people are threatened by floods, especially in coastal and estuarine regions (Merz et al., 2008; Apel et al., 2009). Assessing flood risk, preparing effective flood mitigation measures and utilizing flood benefits at the same time have thus become an even more vital task in water resources engineering.

One of the key techniques in flood hazard analysis as well as in flood forecasting is the numerical simulation of inundation processes (Apel et al., 2004; Vorogushyn et al., 2010). For each application an appropriate approach for flood modeling should be carefully selected based on both scientific and technical aspects.

This study was, firstly, motivated by a very complex case study of flood modeling in the Mekong Delta, an area witnessing large scale floods annually and particularly susceptible to climate change impacts. This study used one-dimensional hydrodynamic approach to simulate

the inundation processes of the whole Mekong Delta, one of the largest estuaries in the world with a highly complex hydraulic system. Floods in the Mekong Delta occur annually. Average floods are perceived as beneficial to the Delta. In fact, the annual floods are the basis of the livelihoods of several million people in the Cambodian and Vietnamese part of the Delta. However, extreme floods (e.g. flood years 1961, 1978, 2000, 2001 and 2002) can cause huge damage (Hoa et al., 2007; MRC, 2009) and pose a serious threat to millions of people. This flood risk is likely to rise due to observed and expected sea level rise and changes in discharge caused by climate change (Wassmann et al., 2004; Hoa et al., 2007; IPCC, 2007; MONRE, 2009; Doyle et al., 2010).

In flood inundation modeling calibration is an essential, although not a trivial step. The usual procedure is a manual adjustment of the model parameters to fit the observations. Obviously, from a scientific point of view this procedure is not sufficient, due to the subjectivity and the known drawbacks (Gupta et al., 1998; Madsen, 2000) like e.g. local minima in the optimization space and equifinal model parameterizations. Therefore, automatic and multi-objective model calibration algorithms have been developed to tackle these problems. However, automatic and multi-objective calibration of hydrodynamic models is still underdeveloped compared to other disciplines such as e.g. hydrological modeling. The reasons for this are manifold, but the main obstacles were the high computational demands of hydraulic models, especially for two-dimensional applications, and the lack of suitable data against which the models could be calibrated. However, in recent years progress was made in both reduction of computational times and in the development of methods for collecting data sets that enable better identification of floodplain processes. For the latter aspect remote sensing applications, especially radar based derivation of inundation areas, made a significant impact in the understanding and mapping of spatial inundation dynamics (Schumann et al., 2009a)

The presented work was motivated by this observation and the idea to use the “new” data in an automatic multi-objective calibration framework for hydraulic models, trying to optimize model parameters with respect to both temporal and spatial inundation dynamics. Thus an automatic multi-objective calibration has been implemented for a large scale hydrodynamic model covering the whole Mekong Delta. In the calibration different data sources - inundation extent maps with high spatial information content and multiple gauge measurements with a high temporal resolution – were used in order to objectively calibrate the model for both temporal and spatial inundation dynamics and to identify likely equifinal model parameterizations.

Another aspect of the thesis is the application of the model calibrated with the methods outlined above for flood hazard analysis in the Mekong Delta, including climate change impacts. We applied a non-stationary extreme value statistic and a multivariate flood frequency analysis. An

approach like this does not exist for the region, not even a standard hazard analysis based on a univariate, stationary extreme value statistic. The scientific aspects of this part of the study are:

- Non-stationarity of discharge time series – recent studies have shown that there are significant trends in variability of flood magnitude in the Mekong calling for non-stationary flood frequency approaches
- Multivariate extreme value statistic-besides flood peak the flood volume is an important characteristic for the Mekong Delta calling for a bivariate analysis.

With the proposed methods a probabilistic flood hazard analysis was performed resulting in flood hazard maps for different return periods derived from the frequency analysis and the calibrated hydraulic model of the Mekong Delta. The flood hazard maps were associated with quantile maps expressing the inherent uncertainties, both from natural variability of flood events (aleatory uncertainty) and uncertainties in climate change predictions.

1.2 Objectives

In summary of the above, the objectives of this study are formulated as:

- Implementation of an automatic multi-objective calibration framework for a large scale hydrodynamic model taking the Mekong Delta as an example,
- Utilization of a series of spatially explicit inundation maps and of time series of a number of gauging stations with high temporal resolution for the calibration,
- Evaluation of model performance, detection of model deficiencies and model improvement through automatic calibration,
- Large scale flood hazard analysis for the Mekong Delta, considering non-stationarity, climate change impacts and uncertainties.

1.3 The structure of this thesis

The dissertation is divided into seven chapters.

Following the current “Introduction” chapter, chapter 2 gives a short description on the Mekong Delta. Chapter 3 reviews general approaches for flood inundation modeling and describes the large scale flood model used throughout this study, while chapter 4 presents an introduction and overview on multi-objective optimization techniques and algorithms with a special focus on the selected approaches for this study. The main contributions of this thesis are included in chapter 5. It presents the development and application of the automatic calibration framework and its

results. It also presents the discussion on model deficiencies detected by the calibration. Chapter 6 presents the methodology and results of the non-stationary and multivariate flood frequency analysis along with the derivation of flood hazard maps for the whole Mekong Delta using the calibrated model and the outcome of the frequency analysis. The final chapter 7 provides a summary and outlook of the whole thesis. In the appendix information and results additional to the ones presented in chapters 5 and 6 are collected for the interested reader.

Chapter 2

THE MEKONG DELTA

"God made only water, but man made wine."

(Victor Hugo)

2.1 Mekong River and its Basin

The Mekong River, originates in the Tibetan Plateau (where the large Yangtze, Salween, Red Rivers also rise) at the height of about 5000 m, spans from $8^030'N \rightarrow 30^0N$ and $95^0E \rightarrow 109^0E$, and runs through six countries (China, Burma, Laos, Thailand, Cambodia, Vietnam). It stretches to about 4800 km, forming the drained catchment area of 795,000 km with an annual runoff of over 475 billion cubic meters, and it is one of the greatest rivers in the world (Hoi, 2005; MRC, 2005a; Adamson et al., 2009). Because it shares its water resources by flowing through many countries, the Mekong River is seen as an international river (like Nile in Africa, Amazon in South America, Danube in Europe, etc…). The Mekong, from its source, flows 2,200 km to the Golden Triangle, where the borders of China, Thailand, Laos, and Burma intersect forming the so-called Upper Mekong Basin (UMB). Here, its altitude drops 4,500 m before entering the Lower Mekong Basin (LMB), or the remaining region of the river basin (Figure 2.1-1).

Table 2.1-1: Territory of the six countries within the Mekong River Basin (after MRC, 2005)

	China	**Burma**	**Laos**	**Thailand**	**Cambodia**	**Vietnam**	**Total**
Area (10^3 km^2)	165	24	202	184	155	65	795
Area portion	21	3	25	23	20	8	100
Flow portion	16	2	35	18	8	11	100

The UMB is characterized by its narrow and deep gorges, which make up 24% of the total basin area. However, the runoff generated from the melting snow on the Tibetan Plateau makes up only 10%-20% of that for the whole area, whereas the portion yielded in the LMB, generated from moisture brought in by the monsoon systems: Indian and Western North Pacific, takes up 80%-90% (Hoi, 2005; MRC, 2005a). Table 2.1-1 represents the area and flow contribution of the six countries sharing the Mekong River Basin.

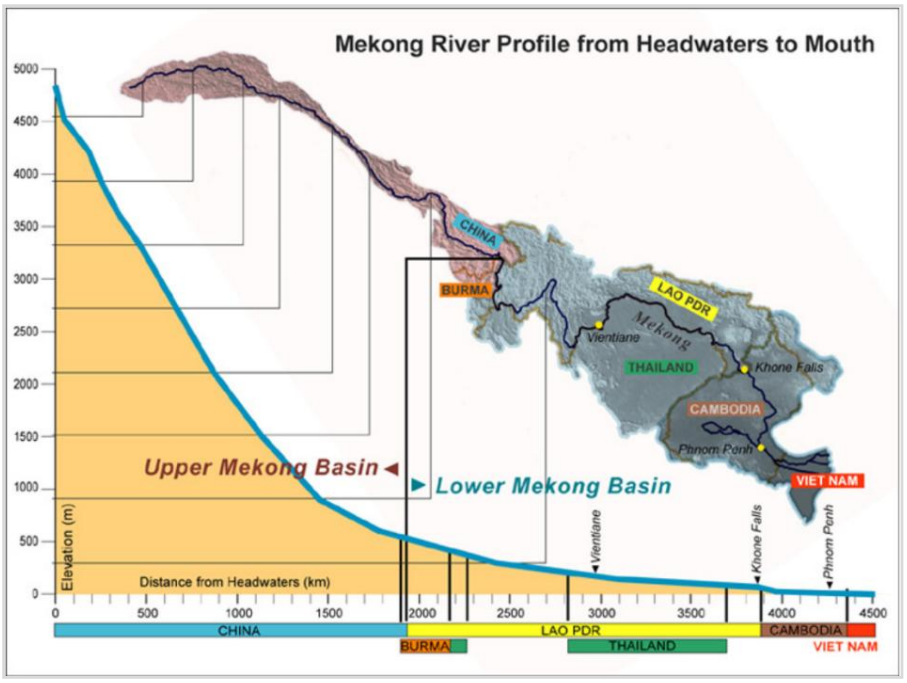


Figure 2.1-1: Longitudinal profile of the Mekong River from source to delta (MRC, 2004)

2.2 The Delta

The Mekong River Delta resides in the lowest part of Mekong River basin. In several published studies, the term Mekong Delta has been used differently. Some refer to the Delta as that in South Vietnam, while some include the region downstream of Phnom Penh (Cambodia). In this study, when referring to the study area, the Mekong Delta (Figure 2.2-1), we mean the area covering Kratie, Tonle Sap Lake (in Cambodia) and the delta in South Vietnam. As seen in Figure 2.2-1, this region is bordered by the East Sea and Gulf of Thailand.

Topography

Topography of the Mekong Delta is very low compared to the Kratie upstream where it is generally regarded as the point in the Mekong system where the hydrology and hydrodynamics of the river change significantly. This is illustrated clearly in the standard Shuttle Radar

Topography Mission (SRTM) DEM shown in Figure 2.2-2 (left). Most parts of the delta are lower than 5 m above sea level. Especially, the delta in Vietnam is a very wide and flat area. It occupies approximately 39000 km^2 (a bit larger than the state of Baden-Württemberg in Germany) with the average elevation from 0.8 – 1.2 m and decreases gradually from the Vietnamese-Cambodian border to the Vietnamese sea coast. The ground base of the delta in Cambodia is higher.

Climate

The Mekong Delta is located in the tropical zone and is under the influence of both the Indian monsoon and the Western North Pacific monsoon (Delgado et al., 2010). The climate in this area is mainly dependent on the monsoon wind system, which causes the two clear seasons: the dry season (enduring about 5 to 6 months from December to April/May) and the rainy season (enduring about 6 to 7 months from April/May to November) (Hoi, 2005;MRC, 2005a). According to the Koppen classification, this makes the climate of the Mekong Delta fall into the savanna category (Sakamoto et al., 2006).

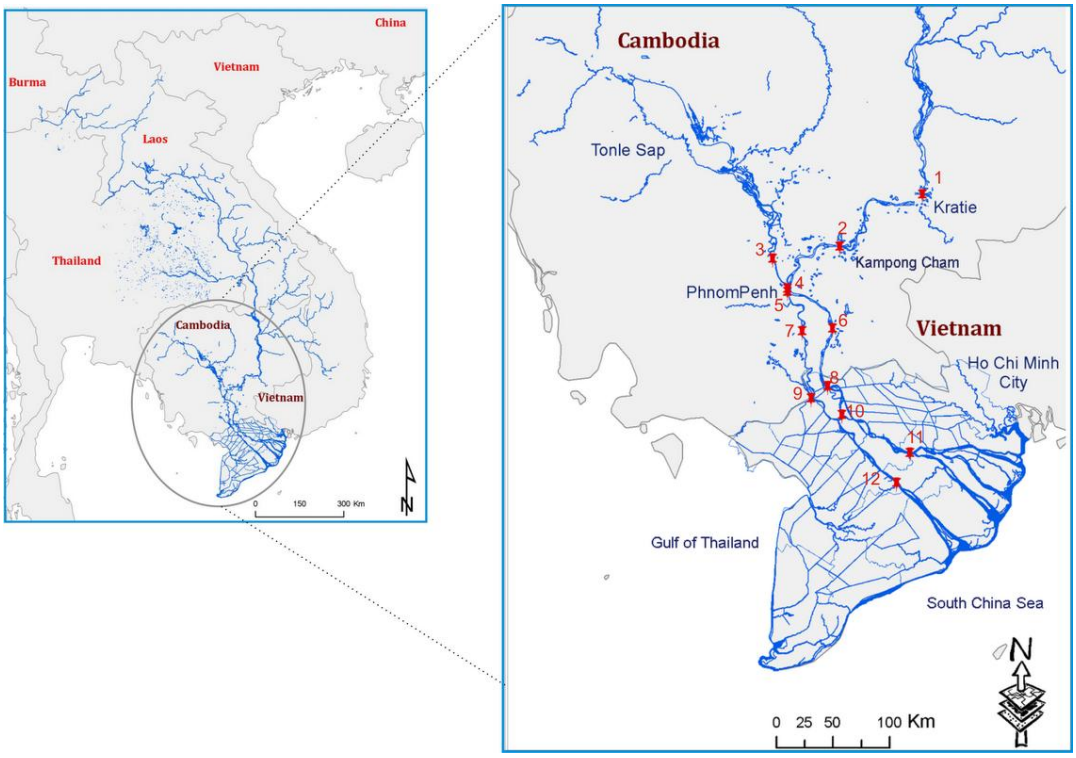


Figure 2.2-1: The Mekong basin and delta and gauging stations in Cambodia and Vietnam used for the calibration (red points); station names are given in Table 5.1-1.

River system

From Kratie, the Mekong mainstream flows down to Kampong Cham and Phnom Penh. At Phnom Penh, it divides into the Bassac, the other branch which retains the name Mekong, and

the Tonle Sap River (see Figure 2.2-1). The Bassac River and the Mekong River continue to run into the delta in Vietnam. In Vietnam, they are called the Hau and Tien Rivers, respectively. The Tonle Sap River connects the Mekong to the Tonle Sap Lake. While the river system in Cambodian territory is quite natural and not complex, that in the delta of Vietnam is extremely dense, including many natural rivers, waterways and a huge number of artificial canals. Here, there is a complex multi-layered circle system of rivers and canals (see Figure 2.2-2).

Tide

The Mekong Delta is influenced by two tidal modes: (i) semidiurnal tide in the East Sea, where amplitudes increase gradually and may reach 4 m from the Vung Tau to Ganh Hao; (ii) the diurnal tide in the Gulf of Thailand, whose amplitudes are low (a maximum of only 1 m). The tidal impact could be found almost everywhere in the Mekong Delta. For example, in the wet season, or flood season, the tidal signal is found in Tien River at Tan Chau, 200 km from the East Sea, 80 km from the Gulf of Thailand. In the dry season, the tide signal may be seen even further from the coast at Phnom Penh, Cambodia. The mixed impact of the two tidal modes is strongest in the Camau Peninsula of the delta in Vietnam.

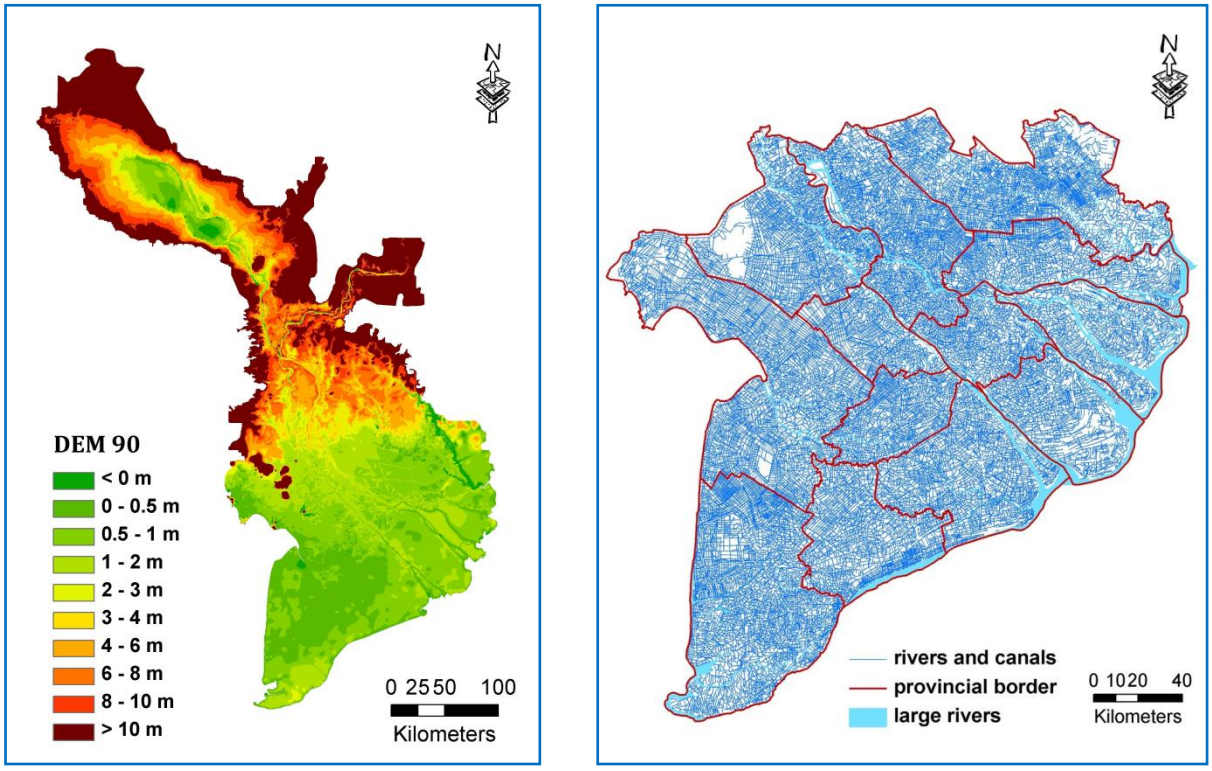


Figure 2.2-2: (left) – SRTM DEM with a horizontal resolution of 90 m covering the Mekong Delta (source: MRC); (right) – system of rivers and canals in the delta of Vietnam (updated until 2006, Dung and Thang, 2006)

2.2.1 Flows and floods in the Mekong Delta

Floods occur only in the rainy season at Kratie, when the flow rises quickly from receiving large amounts of water from upstream. The flood flow regime is highly complex. This is due to a combination of several components: (i) the very low topography of the Mekong Delta, especially in the Vietnamese division; (ii) the presence of a complex multiple-layered circle channel system; (iii) the impact of two different tidal modes; (iv) the peculiarity of the upstream boundary in Cambodia with the Tonle Sap Great Lake; (v) the numerous man-made hydraulic structures in the Vietnamese Delta.

The presence of low topography in the Delta may not be the lone cause of water movement complexity. When combined with other components however, the hydrodynamic complexity of the flow increases significantly.

When the water level in Kratie, the upper boundary of the Delta in east Cambodia, exceeds 17m above sea level, the river flows downstream and west towards Kampong Cham and there an overbank flow begins (Hoi, 2005). This flow is complex and unclear but basically separates into two different flow-paths. The right overbank flow-path follows the main stream west but divides again near Phnom Penh, flowing north towards the Tonle Sap Lake, and south towards the Mekong Delta in Vietnam. A large proportion of the remaining overbank flow flows directly south and reaches the Plain of Reeds, the north-eastern part of the Delta in Vietnam, thus causing a second flood wave in these parts, besides the floods from the Mekong main stem. The most striking natural feature of this area is the Tonle Sap Lake in Cambodia, As a direct result of the overbank flow into the Tonle Sap, the surface area of the lake increases roughly fivefold from 2500 km^2 to 15000 km^2,by discharging around 50 x 10^9 m^3 of water on the annual average. This lake serves as a huge buffer for flood discharges during the monsoon season. The Tonle Sap Lake restores water during the flood season and becomes the main source of water supply (about 80 x 10^9 m^3) to the delta during the dry season.

Floods are a regular event in the Mekong Delta. Flooding of the main stream and tributaries of the Mekong River crucially contributes to the wealth of biodiversity, the abundance of fish, and soil fertility for agricultural activities. Floods in the Mekong Delta, however, often cause a large extent of inundation in both Cambodia and Vietnam. Moreover, several areas are deeply inundated for a very long time. As a result each year, the flood causes loss of life and loss of rural infrastructure, disrupting the social and economic activities of the people living in the delta (MRC, 2005b). In particular, some “disastrous” flood events, such as the floods of 1961, 1978 the three consecutive flood events in 2000, 2001 and 2002, caused severe damage in the Mekong Delta.

2.2.2 Dike system

Unlike the mostly natural status of the Cambodian division of the Mekong Delta, a huge dike system has been constructed in the Vietnamese division, to control floods for several sub-regions. In order to control floods effectively, different types of dike systems were designed to meet the requirements in each flooding zone of the Mekong Delta, Vietnam, such as the Long Xuyen Quadrangle, the Plain of Reeds, the region between Hau River and Tien River, etc. (see Hoi, 2005). A national master plan on the construction of the dike system was officially drawn up in 1999 and adjusted after the most serious flood of 2000. The design of the dike system was mainly based upon the hydrographic shape and the magnitude of historical flood events that significantly damaged the Mekong Delta (e.g. the floods 1961, 1978, 2000). In fact, construction of the dike system was implemented a long time ago to mainly protect the agricultural activities in this region. One single crop was the typical crop pattern in the past (before the disastrous flood in 1978). In the last decades, two-crop or even three-crop patterns have become popular because of the increasing demand for food, as the population has quickly grown in this region (Hoi, 2005). Numerous small dikes, ranging from 3-10m in width, have therefore been raised. A floodplain compartment is an area closed by three, four, or even five dikes at different heights. There are a large number of small compartments in the Mekong Delta, Vietnam. This further explains why the inundation process for the floodplains in the delta of Vietnam is very sophisticated. Figure 2.2-3 shows an example of a dike system in a small area upstream of Mekong Delta, Vietnam (Tam Nong district, Dong Thap province in the Plain of Reeds) where different heights of dike lines can be seen. Water from the main stream starts moving into the floodplains. Some of the floodplains are fully protected by high dikes, while the others are semi-protected by low dikes. The figure also shows the complexity of the inundation process by overbank flow.

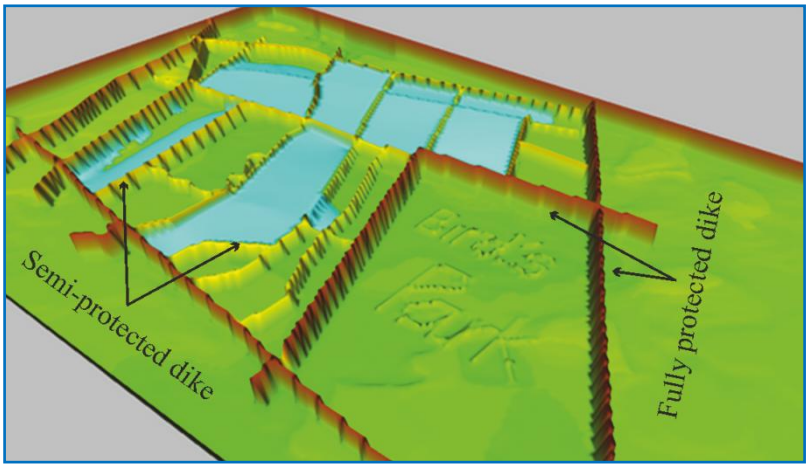


Figure 2.2-3: Illustration of overbank flow in a diked system in floodplains (from the small scale coupled one-dimensional/two-dimensional model example) (Dung, 2008).

Chapter 3

NUMERICAL FLOOD MODELING

"Everything should be made as simple as possible, but not simpler."

(Albert Einstein)

In this chapter, firstly, a short review on the approaches used in flood modeling is presented. Secondly, the large scale flood modeling in the Mekong Delta is described.

3.1 Review of flood simulation approaches

The numerical simulation of the hydrodynamic processes in the river and floodplains is essential for flood hazard analysis as well as for flood forecasting. Numerical models for flood simulation have been developed and applied since several decades for many engineering, planning and risk assessment studies (Chow, 1973; Cunge, 1975; Aronica et al., 1998a, b; Bates and De Roo, 2000; Cunge, 2003; Horritt, 2004). Several methods have been used varying from zero-dimensional models to three-dimensional models, however, with a focus on one-and two-dimensional models (Bates and De Roo, 2000; Sanders, 2007; Apel et al., 2009). The choice of the modeling approach for a certain application depends on both scientific and technical aspects as well as on the resources available. These aspects include, among others, scientific question to be addressed, the time and spatial scale of the simulation, topography, topographical data available, the complexity of the hydraulic regime and computational costs. Table 3.1-1 summarizes several common approaches used in flood inundation modeling (three-dimensional numerical approaches which are rarely applied in practice due to its complexity will not be described further).

Non-hydrodynamic approaches

Non-hydrodynamic approaches are not based on the study of water motion. These methods are adapted to simplify the prediction of inundation processes by combining the information about the water level at gauge stations with an available Digital Elevation Model (DEM) to produce the flood depth and the extents. They are often called planar surface method (Priestnall et al., 2000; Horritt and Bates, 2001; Werner, 2004). Alternative approach introduces an empirical

relationship between the flooded area and the magnitude of the flood discharge via synthetic aperture radar (SAR) imagery for estimating flood extent and duration e.g. LR-FIM model (Townsend and R. Foster, 2002). The benefits of these methods are their simplicity and cheap computational cost. Hence, they can be applied in cases where the data needed for developing more complex models are not available. However, the planar surface approach may show significant overestimation of floodplain extent, and these methods may not be able to capture the inundation extent correctly since they do not consider the hydraulic processes due to local inputs.

Table 3.1-1: Common approaches of flood inundation modelling. Model complexity increases down the table (modified from Bates and De Roo, 2000;Hunter, 2005)

Method	Example codes	Short Description	Example applications
Non-hydrodynamic (planar surface)		DEM and water level information at gauge stations are used to simply interpolate flood extent and depth	(Priestnall et al., 2000; Horritt and Bates, 2001)
Non-hydrodynamic (statistical approach)	LR-FIM	Area – discharge relationship (power law) is built up via synthetic aperture radar (SAR) to estimate the flood extent and duration	(Townsend and R. Foster, 2002)
Storage cell	LISFLOOD-FP, KOD	Different treatment is applied for main channel and floodplain. Floodplain is represented by a series of storages (irregular or raster based). Uniform, kinematic or diffusive wave flow equations are used to treat the flows in mainstream and that between cells.	(Cunge, 1975) (Bates and De Roo, 2000) (Nien, 1996)
One-dimensional hydrodynamic	HEC-RAS, ISIS, MIKE11, SOBEK, VRSAP, HYDROGIS,	The fully hydrodynamic one-dimensional Saint-Venant equations are adopted. Floodplain and channel are together discritized by a series of wide cross-sections perpendicular to the main flow direction. This approach is often combined with an irregular storage cell module to represent floodplain storage and conveyance effects as an alternative to specifying extended cross-sections.	(Werner, 2001) (DHI, 2004) (Horritt and Bates, 2002) (Pappenberger et al., 2007b) (Dac, 2005) (Hoa et al., 2007)
Two-dimensional hydrodynamic	MIKE21 DELFT-FLS (SOBEK 2D) TELEMAC-2D	Solution of the full or simplified two-dimensional depth-averaged shallow water equations. Discretization of floodplain and main channel using structured, unstructured or curvilinear model grids.	(Aronica et al., 1998b) (DHI, 2007a) (Hervouet et al., 1994) (Carrivick, 2006)
Coupled hydrodynamic	MIKE FLOOD SOBEK 1D-2D	Floodplains are represented using two-dimensional hydrodynamic approach. Main channels are described using 1D hydrodynamic approach. Using special links to describe the connection between the main channels and floodplains.	(DHI, 2007b; Dung, 2008) (Frank et al., 2001) (Vorogushyn et al., 2010)
Three-dimensional hydrodynamic		A 3-dimensional Navier-Stokes equation system is solved to simulate flows in main channel and floodplains	(Wilson et al., 2006)

One-dimensional hydrodynamic approaches

In one-dimensional modeling approaches, the floodplain, normally wide and shallow, and the channel, narrow and deep, are jointly discretized as a modeled compound channel. That is, such approaches assume that the flow in the floodplain is parallel to the main channel. The motion of water is governed by the full one-dimensional Saint-Venant equation, consisting mass and momentum conservation equations. Applying these equations requires making a series of assumptions, e.g. that water depth and velocity vary only in the longitudinal direction of the channel, the bottom slope is small, vertical acceleration can be neglected, and a hydrostatic pressure variation along the vertical axis prevails. In one-dimensional hydrodynamic models, water levels and discharge in channels are computed relatively quickly. These models are capable of predicting the bulk of river flow properties, such as propagation and attenuation of the flood wave and backwater effects, provided they are correctly parameterized (Hunter, 2005). Moreover, these approaches are often combined with an irregular storage cell module to represent floodplain storage and conveyance effects as an alternative to specifying extended cross-sections.

Deriving analytical solutions from the nonlinear partial differential Saint-Venant equations is not possible, except for very simplified cases. Hence, a number of finite-difference methods, which are implicit or explicit, have been developed for obtaining the numerical solutions to these equations (Chaudhry, 2008). Many available codes developed using one-dimensional methods have been applied successfully for flood modeling world-wide (see Table 3.1-1) because they are less complex than two-dimensional approach, while still describing the flow regimes well, when a two-dimensional approach is not necessarily required.

Two-dimensional hydrodynamic approaches

Two-dimensional approaches typically integrate the three-dimensional Navier-Stokes equations over the flow depth to produce depth-averaged values of velocity. In case hydrostatic pressure distribution is assumed, the Saint-Venant equations are derived. Otherwise, the Boussinesq equations are adopted. In flood modeling, the former is of much use. Two-dimensional approaches are used where lateral variations in velocity field or the flow direction are of importance. Numerous classes of two-dimensional equations have been developed (see Table 3.1-1). These can be broadly classified into two categories: full and simplified. In the former class, the two-dimensional Saint-Venant shallow water equation system is used in their full form. In the later class, the convective acceleration term are excluded from the governing equations. Solving these equations analytically is almost impossible, hence they are solved numerically via finite difference methods (Abbott, 1979; Lai, 1986), finite element methods (Katopodes, 1984) or finite volume methods (Hirsch, 1990; Yoon and Kang, 2004).

Two dimensional means more accuracy and better resolution, however requires more computational effort. Therefore, applications using two dimensional models have been mainly found in relative small scale reaches.

Coupled one-dimensional and two-dimensional approaches

Coupled approaches combine one-dimensional and two-dimensional systems, utilizing the advantageous features and avoiding several limitations of each. A one-dimensional model is applicable for very long or complicated river systems with little computational effort, but is difficult to model flow paths over complex floodplains or coastal areas. Inversely, a two-dimensional model is capable of simulating flows, esp. overland flow, more accurately but requires high computational costs.

In these approaches, floodplains are represented using a two-dimensional approach and water flow in channel network are modeled using one-dimensional approach. Hydrodynamic links to describe the connection between the main channels and floodplains are required. Many available codes (see Table 3.1-1) have been developed for such hybrid approaches to simulate flood inundation process. Depending on application requirements, coupled approaches could be simplified by relaxing the governing equations of each component. Most of practical applications are in the simplified form (e.g. Aronica et al., 1998b; Bates and De Roo, 2000; Frank et al., 2001; Dutta et al., 2007; Vorogushyn et al., 2010)

3.2 Flood modeling in the Mekong Delta

Some attempts have been made into the establishment of hydraulic models for simulating floods in the Mekong Delta (or its neighbor regions) using two-dimensional and coupled approaches (e.g. Dutta et al., 2007; Dung, 2008).

Using MIKE FLOOD modeling software, Dung (2008) tested the simulation of floods for an area (15 x 15 km) prone to deep inundation in the Plain of Reeds in Vietnam. The study illustrated that, even for such a small area, the simulation of the flood processes is very difficult due to the high complexity of the flow regime, different size of dike systems and the regulation by means of flood control structures. Dutta et al. (2007) developed their own numerical model which combined one-dimensional river flow model and two-dimensional surface flow to simulate flood inundation for the Cambodia floodplain of the Mekong River including the Tonle Sap channel and the Bassac River. They reported that the model was able to simulate the flow characteristics along the main river channels and in the floodplain with a reasonable level of accuracy.

However, all available flood models developed for the whole Mekong Delta and covering the area downstream of Kratie in Cambodia are in the core one-dimensional (Nien, 1996; Thuy and Dac, 2000; Dac, 2005; Dung and Thang, 2007; Hoa et al., 2007; Hoa et al., 2008), except the high cost application utilizing the storage cell approach documented in Cunge (1975). The reasons are the vast extent of the simulation domain and the peculiarities of the hydraulic regime in the Cambodian part of the Delta including the Tonle Sap lake, the large number of man-made structures in the Vietnamese part of the Delta, and the complex bi-modal tidal influence of the East Sea and the Gulf of Thailand which was illustrated in Chapter 2. The differences between these models reside in the different representation of the floodplains, of the numerical scheme used for solving the governing equations, and – most important – the database for the hydraulic control structures in the Vietnamese part of the Delta.

The flood model used for this study was created by the Southern Institute of Water Resources Research (SIWRR) using a huge updated database for the infrastructure of the Mekong Delta (Dung and Thang, 2007). It is based on the software package MIKE 11 developed by the Danish Hydraulic Institute (DHI).

3.2.1 Development of the model used in this study

The establishment of the model has a long history. It was an ambitious project developed by the Southern Institute of Water Resources Research (SIWRR), Vietnam by a modeling group in which the author was one of the chief developers responsible for the schematization of the river network in most regions of the delta. The need for developing the model for the whole Mekong Delta was not only for flood modeling but also for water quality simulation. Several smaller models were set up for training purposes during the training period provided by DHI in Vietnam.

The early version of the model was derived by merging all appropriate smaller models. Different small tools (about 20) were developed by the author. Component models are the model for Cambodia (Fujii et al., 2003), the model for Long Xuyen Quadrangle, the model for Plain of Reeds, the model for Camau peninsula, the model for the South of Mang Thit, and the model for SaiGon – Dong Nai system.

Further useful tools are, among others: the merging tool (to merge different river networks, cross-section profiles, etc.), the editing tool (to edit branch connections, cross section profiles, etc.), the error checking tool (to check simple errors in the model, etc.), the control structure tool, the tool for controlling the initial and boundary conditions (hydraulic and water quality), the tool for supporting the rainfall-runoff model (NAM) used in the original model, and the result processing tool.

3.2.2 Data source

Data sources used to establish and update the model (calibration and validation data will be described in detail in chapter 5) for the Mekong Delta are:

- 25 detailed river network maps (scale 1:100000) covering all the Mekong Delta in Vietnam,
- A 90 m resolution Shuttle Radar Topography Mission (SRTM) DEM covering the whole study area (model domain) and a 5 m resolution DEM provided by Danish Hydraulics Institute (DHI) covering a small area in Tam Nong province (see Section 2.2.2),
- Data on the dike system raised so far in some flood prone areas in the Mekong Delta in Vietnam, such as: Long Xuyen Quadragle, Plains of Reeds, the region between Tien river and Hau river belonging to An Giang and Dong Thap province, some parts of Can Tho city,
- Hydrology data used to define the boundary conditions for the hydrodynamic model,
- Hydrological data for a lumped rainfall-runoff model (NAM),
- River cross-section profiles obtained for most large rivers (e.g. the Mekong River, Hau River, etc) and main channels. Profiles for small channels were assumed using typical design profiles in each region.

This huge data set was collected from various sources or officially inherited from the dataset of the previous models and early version of this model used in the Mekong Delta. Some of the main institutional sources are Mekong River Commission (MRC), Southern Institute of Water Resources Research (SIWRR), Southern Region Hydro-Meteorological Center (SHRMC), and Southern Institute of Water Resources Planning (SIWRP). Other sources were local authorities and several own surveys.

It should be noted here that the above data are of different levels of accuracy, especially those of the dike system data.

3.2.3 MIKE11 HD

The hydrodynamic module of MIKE 11 software package was used to simulate floods in the Mekong Delta. A short description of MIKE11 HD is given below. More details can be found in (DHI, 2004).

MIKE11 HD applied with the dynamic wave description solves the vertically integrated Saint-Venant equations of conservation of mass and momentum. The derivation of the equations of continuity and momentum, as used by MIKE11, is given in the box below:

The Saint-Venant system of equations including the equation of continuity and the equation of momentum used in MIKE 11:

Continuity equation:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{3.1}$$

Momentum equation:

$$\frac{\partial Q}{\partial t} + \frac{\partial \left(\alpha \frac{Q^2}{A}\right)}{\partial x} + gA\frac{\partial h}{\partial x} + \frac{gQ|Q|}{C^2 AR} = q \tag{3.2}$$

where:

- Q : discharge
- A : flow area
- q : lateral flow
- h : stage over datum
- C : Chezy coefficient
- R : hydraulic or resistance radius
- α : momentum distribution coefficient

The system of equations is numerically solved based on an implicit finite difference scheme developed by Abbott and Ionescu (DHI, 2004). This 6-point scheme, often called Abbott-Ionescu scheme (see Figure 3.2-1), is structured in such a way that it is independent of the wave description specified (i.e. kinematic, diffusive or dynamic). The computational grid is automatically generated using switching discharge points Q-points and water level points H-points based on the requirements of the modeler (see Figure 3.2-2).

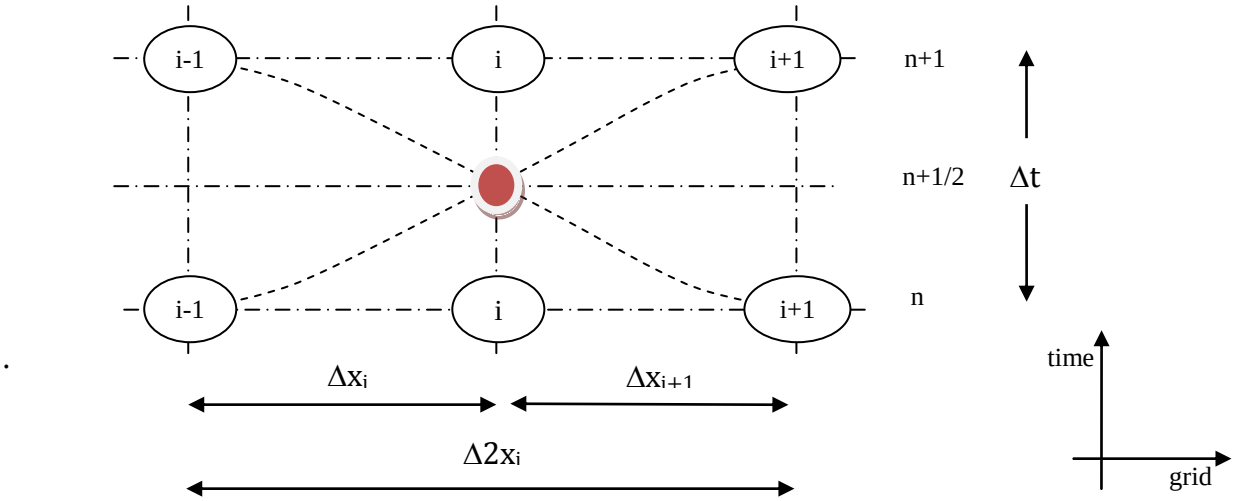


Figure 3.2-1: Abbott and Ionescu 6-point scheme for solving the Saint-Venant equation system.

H-points are placed at the node i and Q-points are placed at node i-1 and i+1 as applying the Abbott and Ionescu scheme to the continuity equation. H-points are placed at the nodes i-1 and i+1 and Q-points are placed at node i as applying the Abbott and Ionescu scheme to the momentum equation. Derivatives in Equations 3.1 and 3.2 at nodes i are approximated by a forward finite difference scheme with respect to time step and by a "modified" central finite difference scheme with respect to grid size (DHI, 2004).

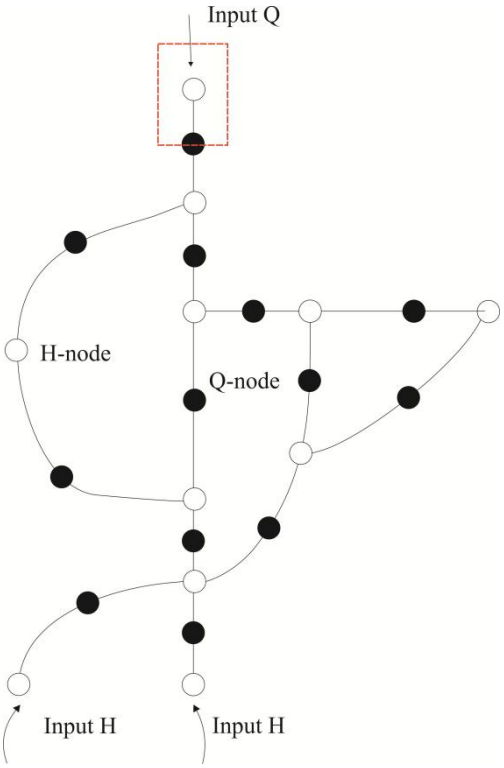


Figure 3.2-2: Computational grid with alternating H-point (in white) and Q-point (in black). Red dot box notates the local treatment in case the discharge is used for the boundary condition.

Nodes at connections or boundaries are always H-points although discharge Q is sometimes given as boundary condition. In such cases, a local treatment is needed (red dot box). To model floodplains in MIKE11 HD, various approaches could be taken. To the case study in the Mekong Delta, full one-dimensional approach is adequately suited (quasi-two-dimensional). This will be explained and justified in details in the next subsections. To discretize the floodplains, structures are used to represent the dikes. MIKE11 includes descriptions for a wide range of structures which act as control points. The formulation of these features permits great flexibility since they can be arranged both in their degree of user-intervention and in their level of complexity.

The selection of an appropriate time step Δt (fixed, not adaptive) and grid size Δx (maximum distance of two computational nodes) is crucial in order to obtain a stable and adequate solution for the finite difference scheme. The basic criterion for Δt and Δx is that they should be

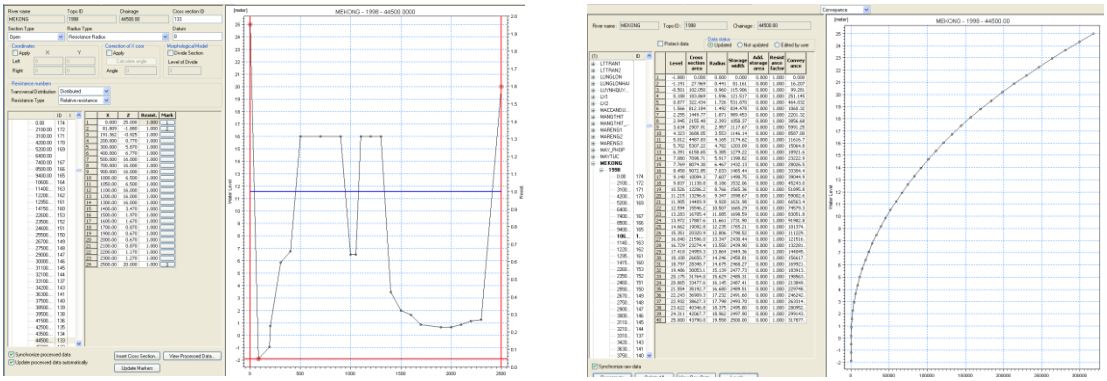
sufficiently fine to work well with the nonlinear behavior in terms of temporal and spatial variations. The Courant condition is an important hint to guide the selection (DHI, 2004).

In general, given a defined grid size, the smaller the time step the more stable the model is. However, for the large scale modeling application of this study and, especially in the context of multi-objective automatic calibration, where a very large number of model simulations may be needed, the time step value should be chosen accordingly.

3.2.4 Model set-up

The flood model was set up to represent the river network and floodplains in the Mekong Delta (Dung and Thang, 2006, 2007; Dung et al., 2009, 2010). The model domain embraces the complete Delta from Kratie including the Tonle Sap Great Lake in Cambodia to the river mouths in Vietnam (see Figure 3.2-5). For such a large-scale model only a one-dimensional approach is feasible from a computational point of view. Figure 3.2-3 depicts a cross section profile in a main channel in the study area. However, the model needs to represent the floodplains in order to be hydraulically meaningful and to enable prediction of flood extends. Two different ways of representing the floodplains were followed. The comparatively natural floodplains in Cambodia including the Tonle Sap, where hardly any channels and dikes exist, are simulated by wide cross sections including the floodplains, which is the usual approach in one-dimensional hydrodynamic flood models. The floodplains in Vietnam, which are separated into a multitude of compartments enclosed by high dikes, which are comparable to dike rings in the Netherlands, were treated differently. Because most of the compartments represent a closed system surrounded by dikes and channels, flood cells are modeled by artificial branches with low and wide cross sections extracted directly from the DEM being the standard Shuttle Radar Topography Mission (SRTM) DEM with a horizontal resolution of 90m (Figure 2.2-2). Those branches are linked to the channel by control structures. Here weirs were used to represent dikes and dike overflow and sluice gates were used whenever information on existing sluice gates was present. Figure 3.2-4 illustrates the approach, by which a quasi-two-dimensional model is established. This approach resembles the first hydrodynamic model developed for the Mekong Delta by Cunge (1975) but adopted to the present channels and dikes, which were hardly present at that time. The model consists of 4235 branches, of which more than 2134 are used to represent the floodplains of the Mekong Delta in Vietnam in 564 compartments, which is equivalent to about 26376 computational nodes (see Table 3.2-1). The length of the simulated channel system is about 25000 km. The topographical data for the model were collected from various sources, thus having different levels of accuracy. The most accurate data could be collected for the main stems of the Mekong and for some larger channels. For the smaller channels, where usually no bathymetric data were available, bed elevation had to be assumed. The same holds true for the

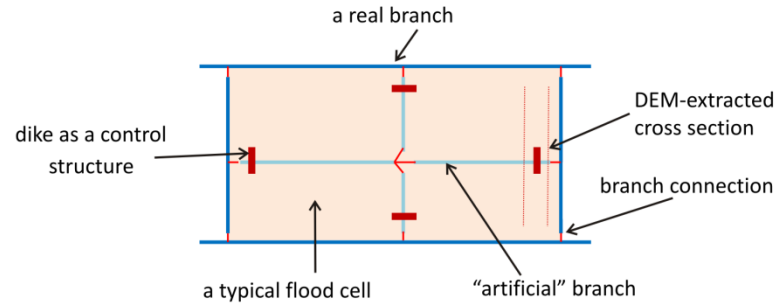
dike elevations, with the additional problem of different datum used by the different agencies and districts for geo-referencing of elevations.



(a) (b)

Figure 3.2-3: An example of a modelled channel cross-section profile ((a): raw; (b): processed) illustrated in the cross-section editor (the file size the cross-section is roughly 30Mb)

(a)



(b)

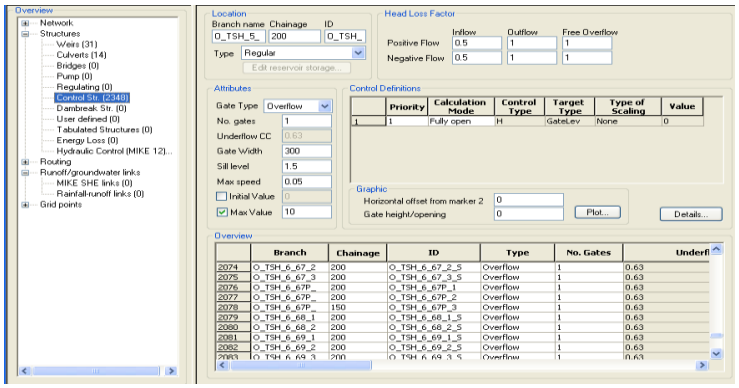


Figure 3.2-4: (a): Representation of a typical floodplain compartment in the Vietnamese part of the Delta, (b): dikes are modelled as flexible control structures in the structure editor.

Table 3.2-1: Statistics of the hydrodynamic model of the Mekong Delta

Item	value
Number of real branches	4235
Length of simulated channel system (km)	~25000
Number of "artificial" branches	2134
Number of computing nodes	26376
Downstream boundary conditions	84
Upstream boundary conditions	4
Number of flood plain compartments	564

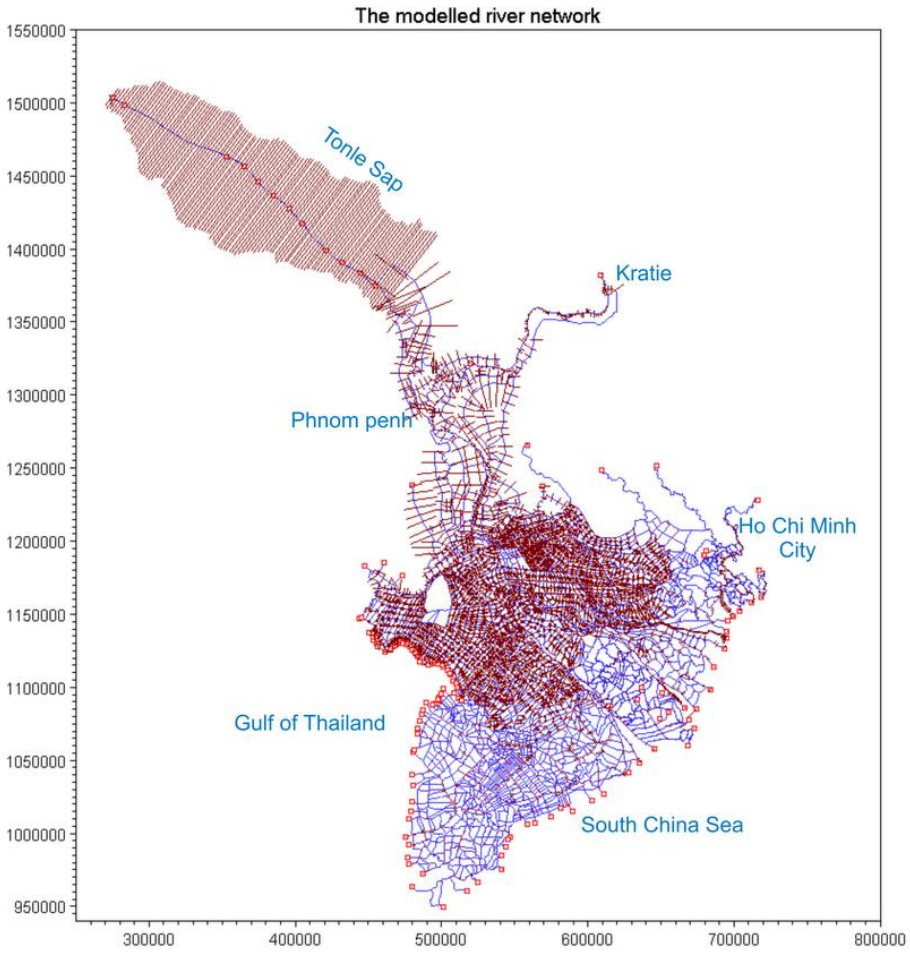


Figure 3.2-5: The model river network including Tonlé Sap – Great Lake

3.2.5 Boundary conditions

The upstream flow of the delta in Kratie is the most important input for the hydrodynamic model. Besides, the Sai Gon – Dong Nai river basin was also integrated in the model, hence the upstream boundary conditions of this basin were also included. Tidal stage measurements of the 10 stations in this area are used to define the downstream boundary conditions for 84 open nodes along the sea.

(a)

(b)

(c)

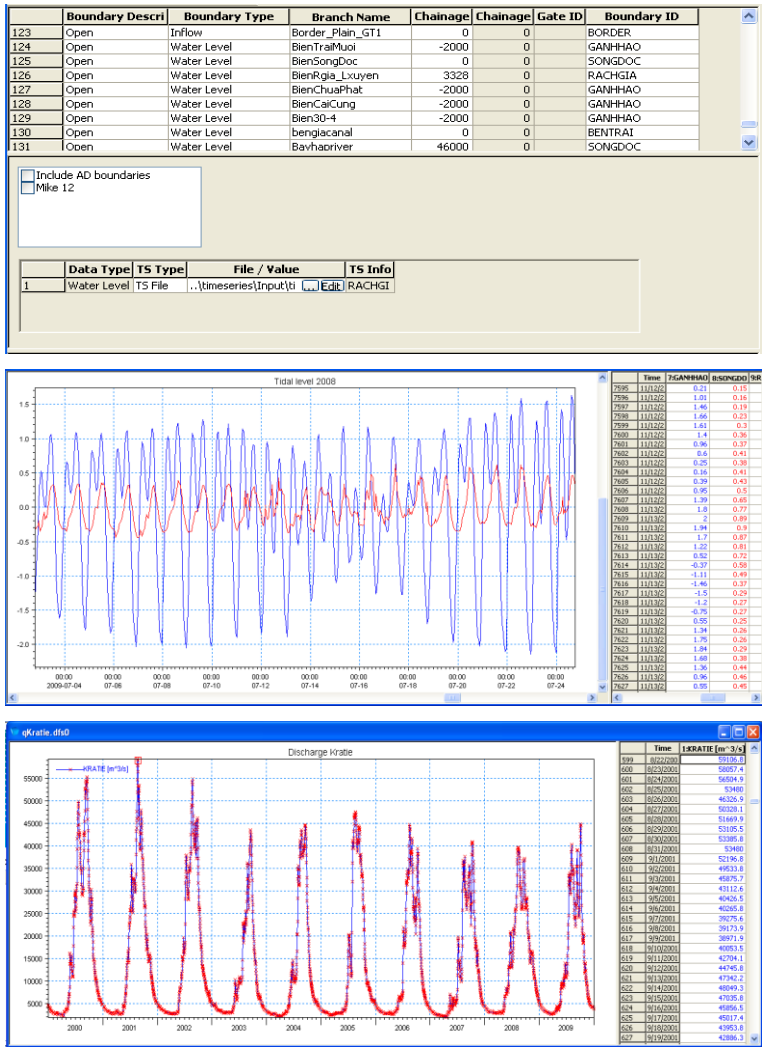


Figure 3.2-6: Boundary conditions: (a) snapshot of the boundary condition editor, (b) tidal stage data (red for level in the Thailand Gulf, blue for level in the East Sea), (c) upstream boundary discharge series at Kratie

3.2.6 Justification of the selected one-dimensional approach

The main justification for using a one-dimensional modeling approach for the simulation is the sheer dimension of the simulation domain, which has a size of more than 55000 km^2. For this dimension, two-dimensional approaches are still very CPU-demanding, even with modern

computational facilities and parallelization techniques. The only study of such large scale dimension was published by Wilson et al., (2007) for a reach of the Amazon. However, the simulation domain was about 13.000 km^2, less than a quarter of the size of the presented study, while the size of the computational grid cells was 270 m. The simulation time in this study for a 22 month dynamic simulation period was reported to be 14 days. Projecting this on the Mekong study, a single simulation for the Mekong would last about 360 hours or about 15 days. Clearly, with this computational time the presented automatic calibration would be impossible. Moreover, a grid cell size of 270 m would result in badly represented inundation areas and inundation processes. While this spatial discretization may be acceptable for natural large scale floodplains as the Amazon, it is insufficient in a heavily controlled environment such as the Mekong Delta, where the inundation is mainly controlled by dikes and dike overflow, respectively. In order to capture this feature, the information of the dike lines and elevations would have to be mapped into the DEM, which itself would be a major effort, and then represented as either 270 m wide dikes resulting in unrealistic inundation areas and processes, or with a higher spatial resolution of the DEM causing even longer computational times.

Another important aspect is the high complexity of the channel network. Simplified two-dimensional inundation models like LISFLOOD-FP (Bates and De Roo, 2000) are mainly designed for tree-like channel topologies and hence cannot cope with the multiple layered circle systems as present in the Mekong Delta. Considering these computational cost implications and structural model demands, the choice of a sophisticated one-dimensional approach (i.e. quasi-two-dimensional) is therefore without alternatives. In order to represent floodplain inundation within a one-dimensional model, we developed the presented approach of shallow cross sections for the floodplains, connected via weir structures to the channels, representing the dikes. This simplification is acceptable, because of the low topography and unstructured morphological characteristics of the floodplains. The main floodplain inundation features, the dike overflow and the flood propagation over an almost flat surface can thus be represented in sufficient detail and resolution.

3.2.7 Some previous applications of the model

The model has been updated several times, mainly due to new information on topography and the dike system. There have been a few preliminary attempts for calibration and validation. A manual calibration was implemented for the flood season of the year 2000. Validation was done for the flood year 2001. As observation data time series information at some stations in the Mekong Delta was used.

However, because of the missing calibration against spatial inundation dynamics and the inherent subjectivity in the manual calibration, the results, especially in terms of inundation extents, are to be treated with caution.

Chapter 4

OPTIMIZATION TECHNIQUES

"Things (or foods) that are delicious, healthy and cheap do not exist."

(Vietnamese proverb)

The quote above is somehow translated directly from a Vietnamese proverb, commonly used in daily life. The real meaning, from the author's understanding, implies that it is very unlikely to find out the "best" things gaining the most deliciousness, the most healthiness and the best price, concurrently. The quote, in some sense, is very similar to the "trade-off" concept often found in multi-objective optimization where the deliciousness, the healthiness and the price could be seen as three mutual conflicting objectives. Generally, the "best" things satisfying all criteria may or may not exist. But, one still tends to search for "as good as possible" things. The searching process is somehow very alike to what is called "optimization technique". That is just one example in our real world where one of the most fundamental principles is the search for an optimal state.

Not all problems can be represented well in a mathematical form (e.g. in the example above, it is not easy to measure the deliciousness because deliciousness is a very "fuzzy" or subjective concept). However, when they do, the branch of applied mathematics called mathematical optimization is likely to be the technique to address them.

In this chapter, a general introduction and the theoretical background of optimization problems and algorithms will be presented. The two optimization algorithms used in this study, one single-objective and one multi-objective, will be introduced and explained in more detail.

4.1 Optimization problem

Optimization is a very wide field. Hence, this section does not aim at covering all its theory. To pose a very precise and "formal" definition of an optimization problem is out of the scope of this

thesis. Instead, a few background definitions are mentioned, mainly to put optimization in the context of water resources research. More details can be found in text books (e.g. Goldberg, 1989; Weise, 2009; Hendrik and Toth, 2010).

(Optimization problem)

Mathematically, an optimization problem, in a simple way, can be stated as follows:

$$opt(F), F = \{F_i, F_i: D \to U_i \subset \mathbb{R}, i = 1..n\} \quad (4.1)$$

where:

$opt(F)$	: minimization/maximization operator
$F = (F_i)_1^n$	: objective functions
D	: feasible space (search space, parameter space)
$U = \prod_1^n U_i$	: objective space

A vector x in D is called a decision vector (or parameter vector). $|D|$ denotes the dimension of D.

$F(x) = (F_1(x), F_2(x), \dots F_n(x)) \in U$ is called the objective vector (or the fitness vector). $F(D)$ forms what is called response surface.

When n equals 1, the problem is called single-objective optimization problem. Otherwise, it is called multi-objective optimization problem.

In most cases, the problem defined above is impossible to be solved analytically. There are many effective algorithms expressed as a list of well-defined rules. They have been proposed to deal with a specific class of problems. However, it should be noted here that there is no algorithm being superior to the others. If an algorithm can be used to solve a class of problems, it is always possible to find a problem lying outside that class where the algorithm cannot be applied. This is confirmed and proved in the "No free lunch theorems for optimization" paper of Wolpert and Macready (1997). Under the light of this theorem, two feelings are possible. The first, unfortunately, is an unhappy feeling because it is possible that the problem may not be solved by any available algorithm. Inversely, the second brings happiness to the one who wants to invent a new algorithm! This study uses the available optimization algorithms to solve a very complex problem rather than developing new ones.

Generally, the algorithms for the optimization problems having only one objective function are different from those having two or more objective functions. Therefore, it is reasonable to categorize the optimization problems into two classes: single-objective optimization and multi-objective optimization.

4.2 Single objective optimization

Single objective optimization problems are optimization problems having only one objective function to be optimized. To solve these problems, several algorithms can be used. It is possible to divide these algorithms into two groups: local and global algorithms.

Local methods

Local algorithms start with an initial guess located in the feasible space, and then search for the solution locally around that guess. They are specifically designed to efficiently find the optimum of a response surface $F(D)$ with one well-defined optimum, since in this case the search will eventually reach the optimum irrespective of the starting point (Madsen, 2007). Local search methods can be divided into two groups, direct and gradient-based methods. Direct search methods use information only about the objective function value to guide the search (and therefore also called derivative-free methods). One of the most used methods is the Simplex search algorithm developed by Nelder and Mead (1965). This algorithm gains popularity due to its simplicity and low temporary storage requirements. The gradient-based methods differ in that they require information about the gradient of the response surface. Most of these methods are based on the Gauss-Newton method. A good example is the Marquardt-Levenberg scheme which is implemented in the generic optimization packages PEST (Skahill and Doherty, 2006). Although a few models include explicit calculations of the gradient, most implementations of gradient-based search methods use finite difference approximations (Madsen, 2007).

Global methods

If multimodality exists in the response surface, it could make local search procedures less effective in locating the global optimum. In such cases convergence of local search procedures will depend on the starting point of the search (see Figure 4.2-1). Global optimization procedures are developed to locate the global optimum of a multimodal response surface, thus reducing the probability of being trapped in local optima. A natural way which is able to increase the effectiveness of the local search is to combine several local searches with different starting points, known as multi-start local search. Typically, a set of initial starting points for the local searches are randomly selected within the feasible parameter region. Individual local searches are performed, and the best optimized objective function obtained for the optimization is taken as the global optimum. The probability of the search being trapped in a local optimum is therefore reduced. While multi-start optimization can significantly increase the chance of converging to the global optimum, the procedure is not very efficient because individual searches are performed without sharing of information between them. More efficient global optimization can be obtained if information from previous searches is used when a new local search is carried out.

Skahill and Doherty (2006) presented a multi-start procedure where new local searches are started from points in the parameter space that are most distant from points that have been explored in the previous local searches. There are many methods which could be used for global optimization problems. Weise (2009) listed many global optimization methods which can be divided in two basic classes: deterministic and probabilistic. Among these techniques, evolutionary algorithms (EAs) have found widespread application to various optimization problems and have also been widely applied in parameter estimation problem of river basin models. EAs deal with the evolution of the population based on the natural selection process in nature (Darwin's theory of evolution). Well known algorithms belonging to EAs are genetic algorithm (GA) and the shuffled complex evolution algorithm (SCE-UA). While GA (see (Holland, 1975) uses three basic operators: selection, crossover, and mutation, the SCE-UA (Duan et al., 1992, 1993) method combines various search strategies. These two methods will be presented in different ways in the following subsections. Other global optimization techniques for single objective problems frequently used in water resources studies are, for example, simulated annealing (Kirkpatrick et al., 1983), swarm intelligent (Eberhart and Kennedy, 1995), tabu search (Glover, 1986), robust parameter estimation – ROPE (Bárdossy and Singh, 2008), dynamically dimensioned search–DDS (Tolson and Shoemaker, 2007)…

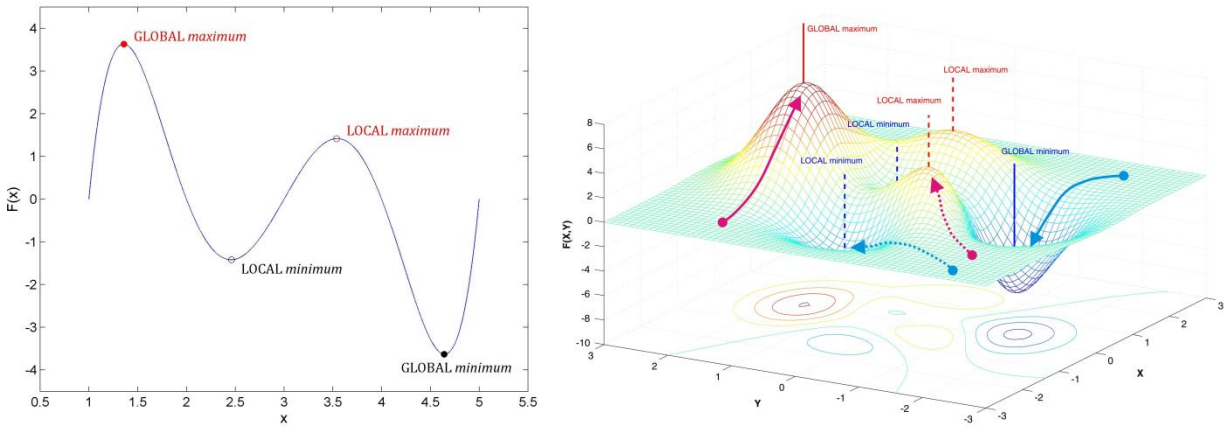


Figure 4.2-1: Illustrations of the multi-modal response surface (left: |D|=1; right: |D|=2) where global algorithms may be more suitable

4.2.1 The Shuffled Complex Evolution (SCE-UA)

The SCE-UA method (UA standing for University of Arizona) (Duan et al., 1992) is a global optimization algorithm which is designed to improve on the best features of the Control Random Search (CRS) method (Price, 1983), by incorporating the powerful concepts of competitive evolution (Holland, 1975) and complex shuffling (new developed concept). Both competitive evolution and complex shuffling help to ensure that the information contained in the sample is efficiently and thoroughly exploited. They also help to ensure that the

information set does not become degenerate. These properties lead us to expect that the SCE method will have better global convergence properties over a broader range of problems. In other words, given a pre-specified number of function evaluations (fixed level of efficiency), the SCE method should have a higher probability of succeeding in its objective of finding the global optimum (Duan, 1994).

Table 4.2-1: SCE-UA algorithm - searching for a global solution

Step	Descriptions
1	Predefine: - the number of complexes $nC \geq 1$ - the number of elements in each complex $mC \geq 2$ (recommended value: $2 * \|D\| + 1$) - the size of a sub-complex $2 \leq mS \leq mC$ (recommended value: $\|D\| + 1$) - the number of evolutions in a sub-complex $1 \leq mI$ (recommended: $2 * \|D\| + 1$) Initialize population P of size $n = nC * mC$ Evaluate the objective value (fitness) of each element in P
2	Rank P in the order of increasing fitness (minization case) $F_1 \leq \cdots \leq F_{nC} \leq F_{nC+1} \leq \cdots \leq F_{nC*2} \leq F_{nC*2+1} \leq \cdots \leq F_{nC*(mC-1)+1} \cdots \leq F_{mC*nC}$ Check for the termination criterion - If it is met, STOP the process; otherwise move to NEXT STEP
3	Partition n elements into nC complexes by $C_i = \{F_i, F_{i+mC}, \ldots, F_{i+nC*(mC-1)}\}$ for $i = 1..nC$
4	For each complex - form a sub-complex by randomly choosing mS elements in the working complex - assign probability to each element in the sub-complex using a triangular probability distribution - evolve the sub-complex using the Simplex algorithm (Nelder and Mead, 1965) mI times
5	Shuffle (combine) all the complexes, back to step 2.

The initial random sampling of the parameter space provides the potential for locating the global optimum without being biased by pre-specified starting points. However, an initial guess of the population could help finding the optimum solution faster for specific problems. SCE-UA and its variant Shuffled Complex Evolution Metropolis (SCEM-UA) developed by Vrugt et al. (2003b) have been applied successfully in many studies, especially in hydrology (Madsen, 2000; Vrugt et al., 2003a; Ngo et al., 2007). The SCE-UA will be used in this study for

all the parameter estimation of statistical models fitted to observed discharge series data (see chapter 6).

4.3 Multi-objective optimization

Real-world problems usually include several criteria that should be fulfilled at the same time. In many cases, the criteria may be conflicting, i.e. trying to improve one of them will result in worse values for some other. Therefore, a perfect multi-objective solution that simultaneously optimizes each objective function is almost impossible or happens only in some special cases. Instead, a reasonable solution to a multi-objective problem is to investigate a set of solutions, each of which satisfies the objectives at an acceptable level without being dominated by any other solution. The mathematical foundations for multi-objective optimization which considered conflicting criteria in a fair way was laid by Vilfredo Pareto more than a hundred years ago (Weise, 2009). The following box shows some basic terms:

(Domination):

Given two elements x_1, x_2 in decision space D, x_1 is said to dominate (or is preferred to) x_2 if x_1 is better than x_2 in at least one objective function and not worse with respect to all other objectives. In this case we denote $x_2 \angle x_1$.

(Non-dominated set):

A set NS being a subset of the feasible space U is a non-dominated set if every element in PS is not dominated by any other element in PS.

(Pareto set):
A set PS being a subset of $U_0 \subset U$ is called Pareto set of U_0 if its elements are not dominated by any elements in U_0.

(Pareto optimal set):
A set Pareto set of U is called Pareto optimal set.

(Non-dominated front):
the image of a non-dominated set is called a non-dominated front.

(Pareto front):
the image of a Pareto set of $U_0 \subset U$ is called a Pareto front of $F(U_0)$.

(Pareto optimal front):
the image of a Pareto optimal set is called a Pareto optimal front.

Figure 4.3-1 illustrates some of the terms defined above in both parameter and objective spaces. As shown, the parameter space can be divided into “good” (Pareto optimal) and “bad” (dominated) solutions, and a member of the Pareto optimal set is better than other members with respect to some objective functions but not with respect to all.

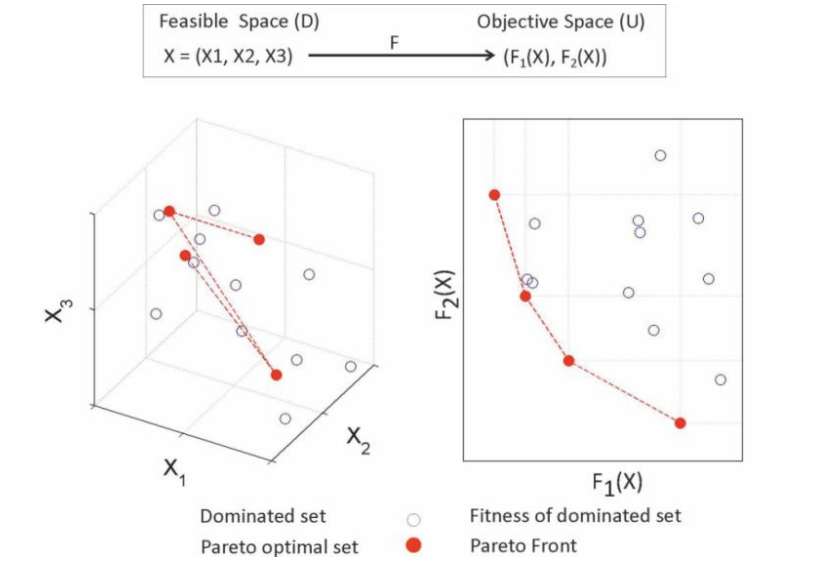


Figure 4.3-1: Illustration of the Pareto-optimality concept: an example of 3-dimensional parameter space (feasible space) and two-dimensional objective space

Identifying solutions located in Pareto optimal is the crucial goal of a multi-objective optimization algorithm. However, in practical problems, estimation of that "real Pareto optimal" set is computationally infeasible. Instead, a more pragmatic approach is to look for a set of solutions (the best known Pareto set) that represent the Pareto optimal set as well as possible. In general, approaches to solve a multi-objective optimization problem can be classified into two main groups: aggregation approaches and Pareto ranking approaches.

4.3.1 Weighted sums (linear aggregation) approaches

In this approach, each of the examined objective functions is normalized (to be comparable) and assigned a weight and then aggregated into a unique measure defining the new objective function in the optimization problem. Then it makes use of algorithms well designed for single objective optimization problems (e.g. SCE-UA, GA). This classic approach is to solve a multi-objective optimization by assigning different set of weights and optimizing the corresponding aggregated objective functions.

In Equation 4.2, weight w_i is assigned to each normalized objective F_i^{norm} so that the problem is transformed to a new single objective problem $G(w)$ with a scalar objective function as follows:

$$opt(G(w)), G(w) = \sum_{1}^{n} w_i F_i^{norm} \ , \qquad \sum w_i = 1 \tag{4.2}$$

The notation $G(w)$ means that the objective function G depends on the set of weights $w = \{w_i\}$. Optimizing $G(w)$ for a given vector w yields a single solution. Therefore, multi-objective

problems must need to specify many vectors w in advance (that's why it's also called the prior approach). Specifying the weight vector may cause difficulty. However, the aggregation of the different objective functions is not trivial since they reflect different information contents, different scales of magnitudes, etc. Applications and discussions of the use of this approach were documented in Madsen (2000, 2003), Blasone et al. (2006), and Konak et al. (2006).

4.3.2 Pareto ranking approaches

Unlike the aggregation approaches, the Pareto ranking approach does not rely on a single comparative measure but on whether one solution is dominated by another for the different calibration objectives considered. Pareto ranking approaches use the concept of Pareto dominance in evaluating fitness or in assigning selection probabilities to solutions. In this approach, the population of estimated parameters is ranked according to a dominance rule, and then each solution is assigned a fitness value based on its rank in the population. Among several ranking methods, the method first proposed by Goldberg (1989) is one of the most common used (the others are Fleming's ranking or Belegundu's ranking) and is still a framework of several algorithms for multi-objective optimization problem. This method assigns uniform probability of reproduction to all non dominated individuals in the population (see Figure 4.3-2). The "onion peeling" algorithm of this method is introduced in Table 4.3-1 below.

Table 4.3-1: Pareto ranking (Goldberg, 1989)

Step	Descriptions
	Input: A finite set P in U **Goal:** Rank all elements in P
1	Initialize the counter $i = 1$
2	Find the Pareto set of P: $PS_i = PS(P)$, increase the counter by 1: $i = i + 1$
3	Set $P = P \backslash PS_i$ (remove the current Pareto set) If $P \neq \emptyset$ then increase the counter $i = i + 1$ back to step 2 Else STOP **Output**: $PS_k, \bigcup_{k=1}^{i-1} PS_k = P$ (original set), all elements in PS_k having the same rank k.

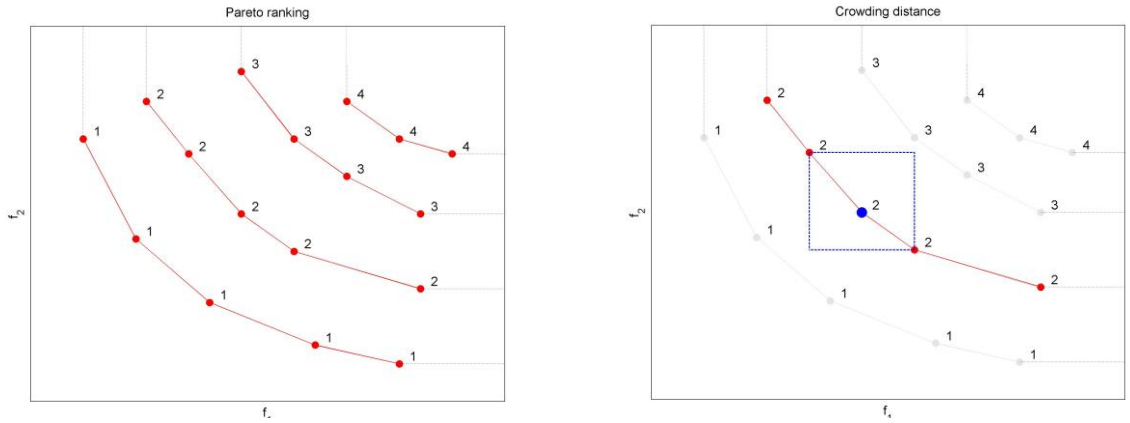


Figure 4.3-2: (left) - representation of the Pareto ranking concept first introduced in Goldberg (1989); (right) - representation of the "crowding distance" measure concept first introduced in Deb et al. (2002)

In the Pareto domination approach, most algorithms suited for the multi-objective optimization belong to evolutionary algorithms. This was motivated by the concept of genetic algorithm developed by Holland and his colleges (Holland, 1975). Because GA is a population-based approach, it is well suited to solve multi-objective optimization problems. From the first multi-objective GA proposed by Schaffer (1985) as vector evaluated genetic algorithm (VEGA), several multi-objective GA were developed including Niched Pareto Genetic Algorithm (NPGA, Horn et al.,1994), Strength Pareto Evolutionary Algorithm (SPEA) and its variants, Pareto-Archived Evolution Strategy (PAES) and especially Non-dominated Sorted Genetic Algorithm II (NSGA-II) and its variants. Among these, NSGA-II has been widely used to different multi-objective optimization problems in water resources (Shafii and De Smedt, 2009). This study also adopts NGSA-II for multi-objective calibration of the large scale hydrodynamic model. Hence, NSGA-II will be presented in details in the next chapter (see Section 4.3.3). Other commonly used multi-objective optimization algorithms that have been developed to deal with multi-criteria calibration problems in water related fields are multi-objective complex evolution (MOCOM, Yapo et al., 1998) and multi-objective complex evolution (MOSCEM, Vrugt et al., 2003a). Tang et al. (2006) and Kollat and Reed (2006) compared the performance of different considered multi-objective algorithms.

4.3.3 The NSGA-II

NSGA-II introduced by Deb et al. (2002) is a new variant of the original non-dominated sorting genetic algorithm (NSGA) which was first proposed by Srinivas and Deb (1995). It is a multi-objective optimization algorithm, and by its name, belongs to the family of genetic algorithm (and larger, evolutionary algorithm family). Compared to the original version, in NSGA-II, three new developed features were presented:

- The fast sorting algorithm for Pareto ranking for a given population.

- The "crowding distance" for measuring the density of the individuals in a population. In the old version NSGA, the sharing parameter is used instead.
- The elitism for making sure that all the best solutions which have been so far evaluated will be kept for the evolution in the next step.

The algorithm and it variants (e.g. ε-NSGA-II) have been applied in many studies in the water resource sector (Gupta et al., 1998; Kollat and Reed, 2007; Tang et al., 2007; Shafii and De Smedt, 2009).

In genetic algorithms, an element x in D is often called an individual, whose components$x_i, i = 1..|D|$ are called single "genes" (parameters). Those genes form a "chromosome" (parameter set). Therefore, it may be said that "chromosome" is an intrinsic property of an individual because it defines the individual ("chromosome to individual" is the same as "coordinates to point"). In a specific context, when the individual x is mentioned, it may refer to its chromosome x! That's because genetic operators are actually applied on the chromosome of individuals. A group of individuals, P (a subset of D) is called a population. $F(x)$ is called the fitness vector (or fitness) of the individual x. One may say that x has an extrinsic property $F(x)$.

In Chapter 5, the descriptions of the core of NSGA-II are introduced. Below are the main features mentioned above and covered by NSGA-II.

4.3.3.1 Fast non-dominated sorting algorithm

Fast non-dominated sorting algorithm is a way to implement the Pareto ranking proposed by Deb et al. (2002). This algorithm helps decrease significantly the complexity of the sorting algorithm which was introduced in the older version NSGA. The algorithm is presented in Table 4.3-2.

Table 4.3-2: The fast non-dominated sorting algorithm in NSGA-II

Step	Descriptions
	Input: A population P with size of n, each individual $p \in P$ are given fitness **Goal:** implement a fast non-dominated sorting over a given population P of the size forming the several Pareto sets at different ranking.
1	For every individual $p \in P$, initialize: - A set containing all individuals dominated p: $S_p = \emptyset$ (or $p.S$) (*) - A counter for the number of individuals dominating p: $ND_p = 0$ ($p.ND$) - The Pareto set of P: $F_1 = \emptyset$
2	For every $p \in P$ For every $q \in P$ If p dominates q then $S_p = S_p \cup \{q\}$ else if q dominates p then $ND_p = ND_p + 1$ If $ND_p = 0$, $F_1 = F_1 \cup \{p\}$ **Output step 2**: $F_1 \neq \emptyset, S_p, ND_p, \forall p \in P$
3	Initialize counter $i = 1$ Do until $F_i = \emptyset$ $H = \emptyset$ For every $p \in F_i$ For every $q \in S_p$ $ND_q = ND_q - 1$ If $ND_p = 0$, $H = H \cup \{q\}$ $i = i + 1$ $F_i = H$ **Output step 3:** $F_k, \bigcup_{k=1}^{i-1} F_k = P$. (each individual has another property which is its rank in the considered population.)

(*) Note:

- Both notations S_p and $p.S$ have the same meaning the "property" or "attribute" S of an "object" p. ($p.S$ is more common used in object-oriented programming languages)
- Analogous rule is applied to notation $p.ND$ and ND_p.

4.3.3.2 Density estimation

Crowding distance approaches aim at obtaining a uniform spread of solutions along the best known Pareto front without using a fitness sharing parameter. NSGA-II uses a crowding distance method as follows (Table 4.3-3):

Table 4.3-3: The algorithm for "crowding distance" assignment in NSGA-II

Step	Descriptions
	Input: a Non-dominated set NS containing n individuals **Goal:** Compute "crowding distance" of all the elements. This algorithm stays the same for both maximization of minimization problem.
1	Assign zero to the "crowding distance" of all the element of NS: ○ $NS[i].cd = 0, i = 1..n$
2	Set k = 1.
3	Sort PS based on the objective k ○ $NS = sort(NS)$ Assign "crowding distance" ∞ to both the first and last element of the PS as: ○ $NS[1].cd = \infty; NS[n].cd = \infty$ And assign the "crowding distance" to other elements of PS as: for i =2 to $n-1$ compute: ○ $A = \lvert NS[i+1].f_k - NS[i-1].f_k \rvert$ ○ $B = \lvert NS[1].f_k - NS[n].f_k \rvert$ ○ $NS[i].cd = NS[i].cd + A/B$ [1]
4	Increase k by 1, back to Step 2 until k exceeds the number of objectives

Note:

- $NS[i].cd$ denotes crowding distance value ("attribute") of the individual number i^{th} in NS ("object" $NS[i]$). (see also Table 4.3-2)
- $NS[i].f_k$ denotes the k^{th}-objective value of the individual number i^{th} in NS.

4.3.3.3 Genetic operator

Tournament selection

In GAs, selection is the stage of choosing "appropriate" individuals in a population to form a mating pool for later breeding. Many different selection schemes for GAs are available. Among them, tournament selection has gained more preference because it satisfies various criteria such as simplicity, efficiency for parallel and non-parallel architectures, and flexibility in changing the selection pressure to be adaptive in different problems. Table 4.3-4 illustrates the tournament selection adopted in the NSGA-II.

[1] Normalization step (not introduced in the original version of NSGA II)

Table 4.3-4: Tournament selection algorithm in NSGA-II

Step	Descriptions
	Input: A population P of size n of which all individuals are given fitness and "density" **Goal:** A mating pool MP with a predefined size of $m < n$
1	Predefine the selection pressure factor: k (*) Initialize $MP = \emptyset$
2	Pick k individuals of P forming a temporary set R: (with or without replacement) Set $bestR = \{p \in R, p$ has smallest rank value$\}$ ($bestR$ will never be empty) **If** $bestR$ contains only one individual add that individual to MP **else** set $bestCD = \{$p $\in bestR, p$ has largest "density" value$\}$ **if** $bestCD$ contains only one individual add that individual to MP **else** select randomly an individual in $bestCD$ and add it to MP
3	Back to step 2 until the size of MP equals m
(*) Notes: the selection of the pressure factor k is dependent on the specific application.	

Simulated Binary Crossover

In genetic algorithms, the crossover operator is a very important operator. It is a binary operation (because two individuals are needed for the mapping). Simulated binary crossover (SBX) was proposed by Deb and Agrawal (1995). It was designed with respect to the one-point crossover properties in binary-code GA. In this operator, the search power is defined by "similarity" probability of a pair of parent solutions to their offspring. This is illustrated in Table 4.3.5.

Table 4.3-5: Binary Simulated Crossover (SBX) operator

Step	Descriptions
	Input: Two individuals x_1, x_2 (parents) in D **Goal:** Their offspring c_1, c_2 (children)
1	Predefine: the "Spread factor" (defining the "gene similarity" probability of the offspring to their parents): $\beta \geq 0$ (*)
2	Compute: $\bar{x} = x_1 + x_2$ $c_{1,2} = \bar{x} \pm 0.5\beta(x_1 - x_2)$

(*) Note:

- If $\beta > 1$ then the crossover is expanding.
- If $\beta = 1$ then the crossover is stationary.
- If $\beta < 1$ then the crossover is contracting.
- The probability density distribution of β is illustrated in Figure 4.3-3. (Probability of occurrence $\beta \approx 1$ is higher than any other β value.)

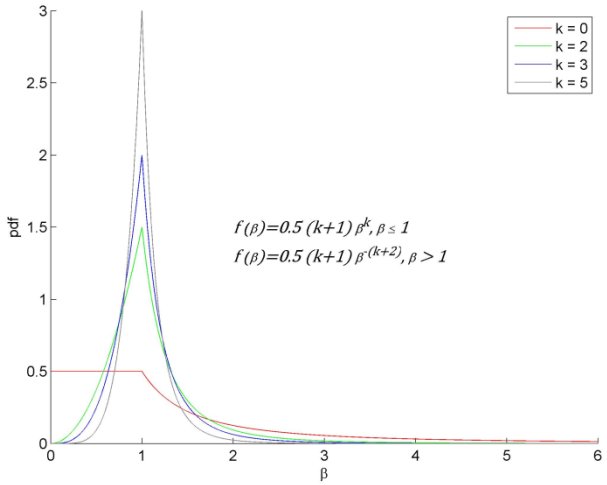


Figure 4.3-3: Probability distribution of the "Spread factor" β

Polynomial Mutation

In GA, mutation used to maintain genetic diversity from one generation of a population to the next. It is analogous to biological mutation. Because it needs one individual to implement the operation, it is a unary operation.

Table 4.3-6: Polynomial Mutation operator

Step	Descriptions
	Input: An individual x in D (assuming that $x_i^L \leq x_i \leq x_i^U, \Delta_i = x_i^U - x_i^L, i = 1..|D|$) **Goal:** A new mutated individual c in D
1	Predefine the factor $\delta \in (-1,1)$ (*)
2	Compute: $c_i = x_i + \delta\Delta_i, i = 1..|D|$

(*) Note:

- The probability density distribution of δ is illustrated in Figure 4.3-4.

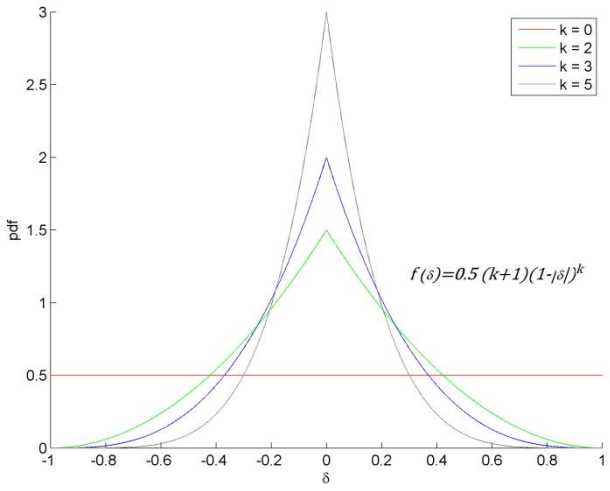


Figure 4.3-4: Probability distribution of the factor δ

4.3.3.4 Elitism

In evolutionary algorithms, the concept of "elitism" is often used. It means that the best solutions found hitherto during the search of the current generation survive to the next generation. In the context of multi-objective optimization problem, the best solutions mean the non-dominated solutions in the generation under consideration. Not all GAs algorithms use this property, especially in the early versions. For example, the vector evaluated genetic algorithm (VEGA,

Schaffer, 1985), being considered as the first multi-objective GA, did not use this concept. Particularly, NSGA (the early version of NSGA-II) did not use this concept either. Elitism is important for the convergence properties of evolutionary algorithms.

Recent GAs, for example NSGA-II, often use elitism concept. NSGA-II uses a fixed population size (N). When the current population is coupled with its offspring having the same size (N), it doubles the size of this temporary population of which each individual will be ranked and assigned the density again. After this step, the best N solutions will be selected to form the new population (see chapter 5).

4.4 Model calibration

Numerical modelling helps enhance knowledge of natural phenomena by transforming them into digital forms that are amenable to numerical analysis. However, there is always a gap between the model and reality since uncertainty is an inescapable and inherent element of any numerical model. Many attempts on reducing model uncertainty have focussed on the reduction of parameter uncertainty. The process of tuning model parameters to improve the goodness-of- fit between model results and the observations is defined as model calibration. Model calibration can be implemented either manually or automatically (or mixed – expert system).

Manual calibration

A process of trial and error parameter alteration is made in manual calibration. Visual inspection is frequently used to evaluate the goodness-of-fit between the modeled and observed behavior. Manual calibration may, sometimes, give acceptable result when employed by experienced modelers. However, it is usually criticized as it has several drawbacks: inherent subjectivity, tediousness, and labor intensiveness, and inability to ensure the best parameter sets (Duan et al., 1993; Madsen, 2000; Fabio et al., 2010).

Automatic calibration

In the last decades, motivated by the development in the computer sciences, a large number of numerical calibration procedures have been developed to reduce the shortcomings of manual calibration. In automatic calibration, the parameter adjustment is guided by optimization algorithms thanks to the power of computers. Mathematical forms are used for defining objective functions (see Section 4.1) which measure the goodness-of-fit of the modeled responses against the observations. The subjectivity is hence eliminated. And the probability to obtain the optima (or solutions closed to the optima) is higher.

4.4.1 Calibration framework

Model calibration may be different from application to application. However, in general, to implement a model calibration, the following steps are followed:

- Setting up a model
- Selecting model parameters to be calibrated
- Defining calibration objectives
- Selecting a suitable optimization algorithm

The model set-up step is an important part that conditions the outcome of the model calibration. It is dependent on applications considered (e.g. the flood model described in Chapter 3), hence not further discussed in this chapter. The optimization techniques were also mentioned as a guide to select suitable algorithm for certain calibration problem. The following text will shortly discuss on the model parameterization and the formulation of objective functions.

4.4.2 Choice of calibration parameters

The main goal of model calibration is to estimate model parameters. Hence, based on the model setup of a specific application, model parameters to be estimated should be clearly defined. Models may contain a large number of parameters which can cause, in many cases, a problem to the calibration process. Reducing the dimensionality of the problem may be necessary. Some of the model parameters may be much sensitive to the modeled responses (related to the objective functions) than the others. Furthermore, they may possibly be correlated. Using prior knowledge (or expert experience) or implementing sensitivity analysis is hence useful, and in some cases, mandatory.

4.4.3 Formulation of objective functions

The appropriate choice of calibration data, which are observations of modeled responses, and formulation of objective functions are essential for the success of the model calibration. In water resources studies, data requirements for model calibration depend on the type of river basin model considered. For a lumped rainfall run-off model, catchment run-off data are needed (Madsen, 2000). For highly distributed hydrodynamic models (e.g. the model considered in this study), spatially distributed data, such as satellite-derived flood extents or water-stage maps, are strongly recommended as calibration datasets (Bates, 2004). When the calibration data is properly chosen, the objective functions are able to be defined using proper numerical performance measures (also goodness-of-fit measures) serving different purposes. There exist numerous numerical performance measures (NPS). However, to follow the context of this study,

three types of them: NPS based on time series, NPS based on binary pattern, NPS based on likelihood are shortly presented.

Performance measure based on time series

Hydrological model outcome is usually presented in time series format. Goodness-of-fit measures based on time series are used to formulate objective functions for model calibration. They mainly measure the model skill with respect to the temporal aspect. Some of them are dimensional measures, e.g. bias, mean square error, standard deviation, some are dimensionless, e.g. correlation coefficient, coefficient of efficiency. Table 4.4-1 lists some measures that are commonly referenced in literature. In the table, X is the observed data series with the size of n and Y is the modeled data series having the same size as X.

Table 4.4-1: Some common "goodness-of-fit" measures (after Gupta, 1998; Madsen, 2007)

Performance measure	Equation	Notes
Bias (ME)	$Bias = \frac{1}{n}\sum_{1}^{n}(X_i - Y_i)$	dimensional
Mean absolute error (MAE)	$MAE = \frac{1}{n}\sum_{1}^{n}\lvert X_i - Y_i\rvert$	dimensional
Mean squared error (MSE)	$MSE = \frac{1}{n}\sum_{1}^{n}(X_i - Y_i)^2$	dimensional
Root mean squared error (RMSE)	$RMSE = \sqrt{MSE}$	dimensional
Standard deviation of residuals (STD)	$STD = \sqrt{RMSE^2 - MAE^2}$	dimensional
Max difference (MD)	$MD = \lvert\max(X) - \max(Y)\rvert$	dimensional
Coefficient of determination (R^2) (square of correlation coefficient)	$R^2 = \frac{\left[\sum_1^n (X_i - \overline{X})(Y_i - \overline{Y})\right]^2}{\sum_1^n (X_i - \overline{X})^2 \sum_1^n (Y_i - \overline{Y})^2}$	dimensionless
Coefficient of efficiency (NSE)[2]	$NSE = 1 - \frac{\sum_1^n (X_i - Y_i)^2}{\sum_1^n (X_i - \overline{X})^2}$	dimensionless
Index of agreement (d)	$d = 1 - \frac{MSE}{\sum_1^n \left(\lvert X_i - \overline{X}\rvert + \lvert Y_i - \overline{Y}\rvert\right)^2}$	dimensionless

[2] To be used in this study

This study uses the coefficient of efficiency to formulate the objective function based on the observed and simulated water level series. The Nash-Sutcliffe efficiency (NSE, defined by Nash and Sutcliffe, 1970) is the statistic which is most widely used in calibration of hydrological models (Gupta et al., 2009). It is dimensionless and ranges in the interval $(-\infty, 1]$. The perfect agreement between the model simulation and the observation derives when NSE is equal to 1. However, one should note that while NSE is frequently used when comparing model performance related to different observed time series, one of the main concerns about the NSE is its use of the observed mean as a reference, which can lead to overestimation of model skill for strongly seasonal time series (Schaefli and Gupta, 2007; Gupta et al., 2009).

Detailed discussions on the use of each goodness-of-fit measure in Table 4.4-1 can be found in Nash and Sutcliffe (1970); Gupta et al. (1998,2009); Legates and McCabe (1999) and Madsen and Khu (2006).

Performance measure based on binary pattern

When observations exist in form of binary patterns (e.g. flood extent maps, snow cover maps, etc.) and the result of the model can also be represented in this format (see Figure 4.4-1), both can be directly compared. Binary patterns could be spatial or temporal. Each cell of the pattern has either the state "1" (e.g. present, wet, snowy, rainy) or the state "0" (e.g. absent, dry, no-snow, no-rain). When comparing the two quantities, a contingency table is obtained as shown in Table 4.4-2. In this table, P^{11} is the number of cells for which both model and observation indicate the state "1" (present state), P^{01} the number of cells for which observation indicates "0" and simulation indicates "1", P^{10} the number of cells for which observation indicates "1" and simulation indicates "0" and P^{00} the number of cells for which both model and observation indicate "0". Table 4.4-3 lists several candidate goodness-of-fit measures for binary pattern comparison.

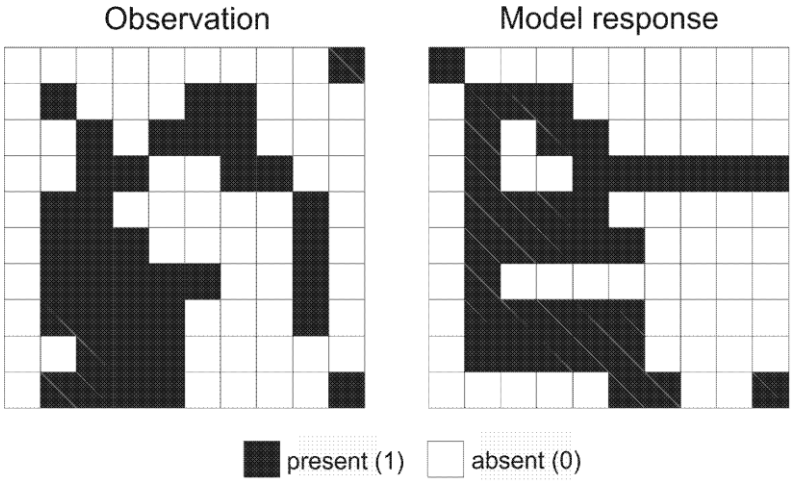


Figure 4.4-1: Binary pattern in both observation and model response

Table 4.4-2: Contingency Table

		Observation		
		present	*absent*	
Model	*present*	P^{11}	P^{01}	$P^{11} + P^{01}$
	absent	P^{10}	P^{00}	$P^{10} + P^{00}$
		$P^{11} + P^{10}$	$P^{01} + P^{00}$	n

Table 4.4-3: Measures for binary pattern comparison (after Hunter et al., 2005a; Schumann et al., 2009a)

Name	**Form**	**Notes**
Bias	$\frac{P^{11} + P^{01}}{P^{11} + P^{10}}$	Recommended for summarizing aggregate model performance
PC	$\frac{P^{11} + P^{00}}{n}$	Not recommended for either deterministic and uncertain calibration
ROC	$ROC_F = \frac{P^{11}}{P^{11} + P^{10}}$ $ROC_H = \frac{P^{01}}{P^{01} + P^{00}}$	(Receiver operating characteristic) Summarizes two types of model error that can occur and is potentially a useful tool for exploring their relative consequences and weighting in any subsequent risk analyses.
PSS	$ROC_F - ROC_H$	(Peirce skill score) Not recommended for either deterministic and uncertain calibration
$F^{<1>}$	$\frac{P^{11}}{n - P^{00}} = \frac{P^{11}}{P^{11} + P^{10+}P^{01}}$	Recommended for both deterministic and uncertain calibration
$F^{<2>}$	$\frac{P^{11} - P^{10}}{n - P^{00}} = \frac{P^{11} - P^{10}}{P^{11} + P^{10+}P^{01}}$	Recommended for deterministic calibration when under-prediction is preferable
$F^{<3>}$	$\frac{P^{11} - P^{01}}{n - P^{00}} = \frac{P^{11} - P^{01}}{P^{11} + P^{10+}P^{01}}$	Recommended for deterministic calibration when over-prediction is preferable

The discussion on the use of each measure can be found in Hunter (2005a), Schumann et. al (2009a) and Aronica et. al (2002). In general, there is no perfect measure for all cases. However, in flood inundation modeling, the measure $F^{<1>}$ is often used to compute flood indexes, although it may pose some deficiencies, e.g. bias towards large inundation extent which are

known and reported (Schumann, 2009a). Nonetheless, due to the lack of better alternatives up to date, it is still the basic measure used and recommended for deterministic calibration. Hence, it will be used in this study (see Chapter 5).

Performance measure based on likelihood

In statistics, a likelihood function (or simply the likelihood) is a function of the parameters of a statistical model, defined as follows: the likelihood of a set of parameter values given some observed outcomes is equal to the probability of those observed outcomes given those parameter values. Likelihood functions play a key role in statistical inference, especially for estimating a parameter of the unknown population from a sample of observation (Wiki, 2011). To fit a sample to a distribution, the method of maximum likelihood is often preferred. In most cases, the log-likelihood function is used instead of the likelihood function.

Chapter 5

MULTI-OBJECTIVE CALIBRATION OF LARGE SCALE HYDRODYNAMIC MODELS UTILIZING FLOOD MAPS AND GAUGING DATA[1]

"Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom."

(Anonymous)

The large scale hydrodynamic model was developed to simulate floods in the Mekong Delta. However, even though large effort was put into obtaining the best possible data for setting up the model, calibration is still required, which both searches for the best possible representation of natural flow-resistance within the simplified flow representation in the model, and compensates for model insufficiencies and errors, which are unavoidable In fact, most models must be calibrated to some degree to be useful for any practical application (Gupta et al., 1998). The calibration can be performed either manually or automatically (see Chapter 4).

For hydrodynamic models, manual calibration is the standard - in contrast to hydrological modeling, where sophisticated automatic calibration dominates nowadays, both in research and application. In fact, just a few studies have been published dealing with the automatic calibration of hydrodynamic flood models (two-dimensional: Fabio et al., 2010, one-dimensional: Madsen and Vinter, 2006). Automatic calibration adjusts parameters automatically according to a specified search scheme optimizing numerical measures of goodness-of-fit of the model results to the data. Automatic calibration has gained increasing popularity in the past decades, because it

[1] Parts of this chapter (modified and improved to suit the structure of a monograph format) was published as:

Dung, N. V., Merz, B., Bárdossy, A., Thang, T. D., and Apel, H.: Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data, Hydrology and Earth System Sciences, 15, 1339-1354, 10.5194/hess-15-1339-2011, 2011.

alleviates the chief drawbacks of manual approaches: subjectivity, tediousness, over-dependence on the expertise of modelers, and the need for a huge amount of labor (Duan et al., 1993; Madsen, 2000; Fabio et al., 2010). A lot of research has been undertaken into developing automated calibration routines or applying them to a large number of water-related applications (Duan et al., 1992; Solomatine et al., 1999; Vrugt et al., 2003b; Skahill and Doherty, 2006; Bárdossy and Singh, 2008). Depending on the specified technique and the dimension of the parameter space, calibration algorithms have to evaluate many simulations, up to hundreds or thousands or even much more. Hydrological models usually take only minutes or even seconds per simulation run. Most hydrodynamic models, particularly flood inundation models, require much longer computation times, typically some hours or even days. Consequently, work on the automatic calibration of hydrodynamic models, especially of large-scale flood models, is rare. In the Mekong Delta, all the mentioned existing flood models have been calibrated manually. However, due to the increase in computational power, and parallelization techniques, automatic calibration of such large-scale hydrodynamic model as the one presented here have become feasible from a computational point of view.

The lack of appropriate data is another handicap to model calibration. Data requirements used for a calibration process depend on the model. For example, for a lumped rainfall run-off model, catchment run-off data are needed (Madsen, 2000). For highly distributed hydrodynamic models, spatially distributed data, such as satellite-derived flood extents or water-stage maps, are strongly recommended as calibration datasets (Bates, 2004). In the last few decades, along with growth in the area of flood modeling, major advances have been made in the field of remote sensing. Reinforcing the connection between increased computation power and increased data availability could improve the model performance significantly (e.g. Aronica et al., 2002; Di Baldassarre et al., 2009; Mason et al., 2009; Schumann et al., 2009a). The dataset used for this calibration study contains in-situ data from a network of river gauging stations and a series of flood-extent maps, derived from the ENVISAT Advanced Synthetic Aperture Radar (ASAR) satellite platform (http://envisat.esa.int/instruments/asar/). As a result of the different type, spatial and temporal coverage of the calibration data, a multi-objective optimization framework had to be chosen.

Several multi-objective optimization techniques have been developed in the past decade and have been applied. In light of the "No free lunch theorem for optimization" (Wolpert and Macready, 1997), as discussed in Chapter 4, and from recent studies comparing different multi-objective automatic calibration algorithms, it can be concluded that there is no algorithm that is superior in all cases (Zitzler and Deb, 2000; Kollat and Reed, 2006; Tang et al., 2006; Wöhling et al., 2008). Given the comparable performance of the available algorithms, the NSGA-II proposed by Deb et al. (2002) and mentioned in Chapter 4, was selected because of its ease-of-use and suitability for a parallelization scheme.

In the following sections the calibration data is introduced, and followed by the actual model calibration framework. This includes the definition of objective functions, the parameter classification, the implementation of the optimization algorithm, and computational aspects of the calibration process. Section 5.3 presents the results and related discussions.

5.1 Calibration data

The database used for this study consists of data used for building and modifying the model and for evaluating the performance of the model. A large amount of topographic data obtained from numerous administrative and scientific providers was used for model development, including data acquired by both ground survey and remote sensing. However, in the scope of this chapter, the data used for the calibration will be mainly presented.

Two types of data were used for this purpose. The first type consists of water level time series from a network of gauging stations. Further, a series of flood extent maps derived from ENVISAT ASAR satellite images is used. While the former data are point data with high temporal resolution, the latter have a high spatial, but a low temporal resolution. Using these different data sets the model performance with respect to temporal and spatial dynamics can be evaluated through the calibration process.

5.1.1 Gauge data

The typical source of hydrodynamic model calibration data consists of main stream measurements at gauging stations at the boundaries of and sometimes also inside the model domain. In the present study, water level time series at 12 gauging stations of the measurement network (see Figure 3.2-5) along the Mekong River and Bassac were used, of which 7 stations are located in Cambodia and 5 stations in Vietnam. The time series used in this study cover the flood season 2008, e.g. from beginning of June to the end of December. These gauge data were provided by Mekong River Commission (MRC). Figure 5.1-1 gives details about the location of the gauging stations used in this study.

Table 5.1-1: Gauging stations used for calibrating the hydrodynamic model (data source: MRC); note: the Mekong in Cambodia is named Tien river in Vietnam, the Bassac in Cambodia is the Hau river in Vietnam (*: those stations are shown in the hydrograph comparison in Section 5.4), CAM in the bracket indicates station in Cambodia, VN indicates station in Vietnam. The numbers in the last column indicate the weights assigned to the single gauging stations in the calibration process.

no.	**name of station**	**geographical location**		**location at**	**weight assigned**
		long	lat	river	to the station
1	Kratie (CAM)	105.987	12.240	Mekong	1
2	Kampong Cham (CAM)	105.338	11.909	Mekong	2*
3	Prekdam (CAM)	104.804	11.813	Tonle Sap	1
4	Phnom Penh port (CAM)	104.923	11.575	Tonle Sap	1
5	Chaktomouk (CAM)	104.923	11.552	Mekong	1
7	Neakluong (CAM)	105.284	11.261	Mekong	1.5
6	Khokhel (CAM)	105.040	11.240	Bassac	1.5
8	Tan Chau (VN)	105.243	10.803	Tienriver	2*
9	Chau Doc (VN)	105.113	10.707	Hauriver	2
10	Vam Nao (VN)	105.357	10.575	VamNao	2*
11	My Thuan (VN)	105.900	10.273	Tienriver	1.5*
12	Can Tho (VN)	105.790	10.033	Hauriver	1.5

5.1.2 Remote sensing data

The availability of remotely sensed flood extents has made a significant change from a data poor to a data rich environment for flood modeling (Schumann et al., 2009c). By combining these data with gauge data the calibration of hydrodynamic models can be better constrained, because the ability of the model to reproduce the temporal and spatial inundation dynamics can be evaluated at the same time.

Most recently the mapping of inundation areas by Synthetic Aperture Radar (SAR) sensors has gained large popularity because of the insensitivity to cloud coverage, which is the main drawback of optical sensors in inundation mapping. The value of such data for model calibration has been documented in a number of studies, which all used one or sometimes two flood extent maps in manual calibration mode (Horritt, 2006;Di Baldassarre et al., 2009). In this study a series of flood extent maps (Figure 5.1-1) covering the period of 17th June 2008 to 30th November 2008 were used.

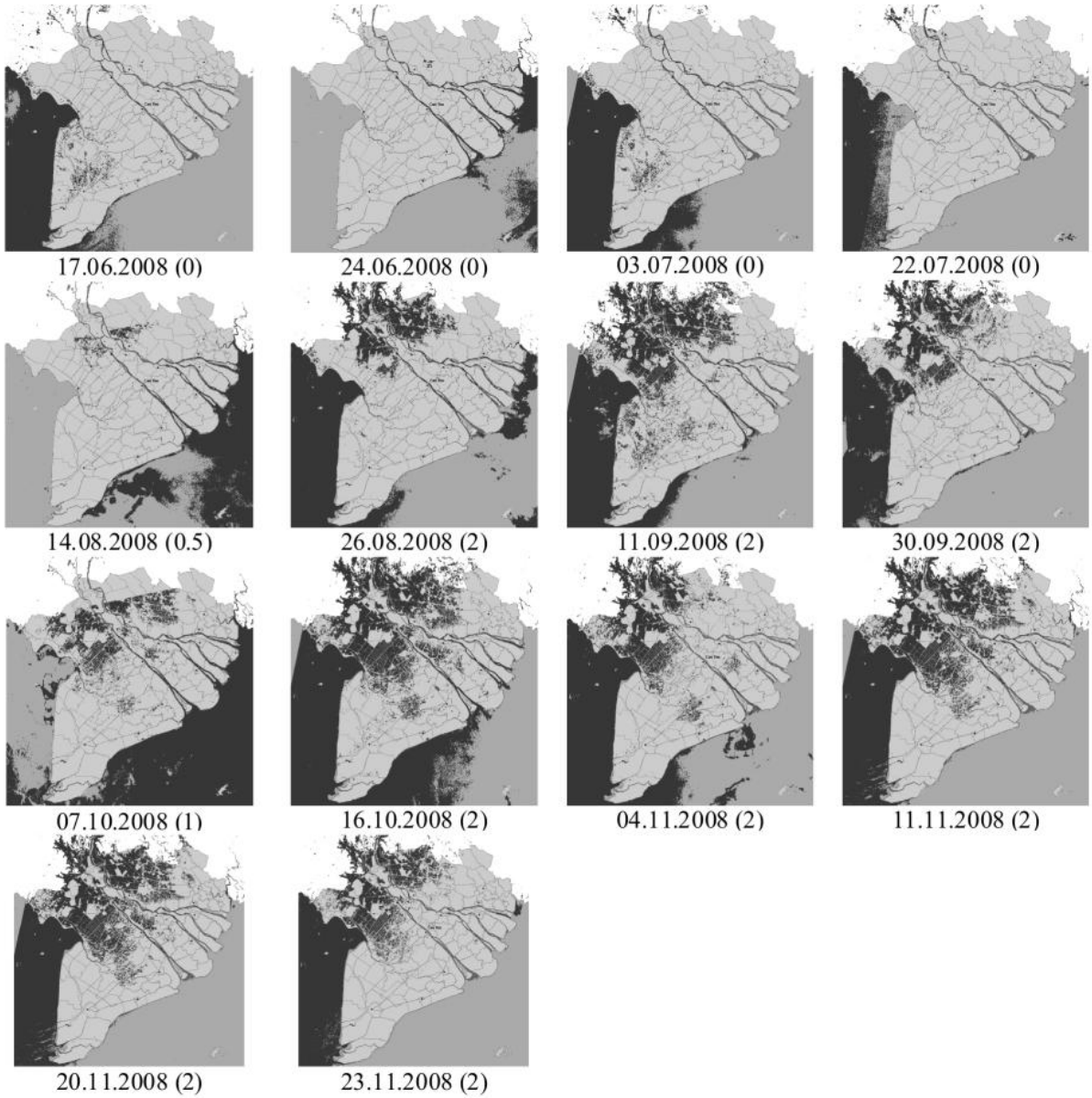


Figure 5.1-1: Flood extent maps derived from ASAR used for model calibration. The numbers in the brackets indicate the weights assigned for the individual map in the calibration process (cf. Subsection 5.2.2).

These maps were provided by German Aerospace Center (DLR) utilizing ENVISAT ASAR radar imagery. The Advanced Synthetic Aperture Radar (ASAR) on the platform uses the C-Band. Data used for generation of the inundation maps are acquired in wide swath mode (image mode) with a geometric resolution of 90m to ensure the coverage of the whole Mekong Delta in one dataset for one point in time. The derivation of the inundation maps is performed by a histogram threshold based approach similar to the one described by (Schumann et al., 2009b). The implementation of the method used for inundation map generation (Huth et al., 2009) was integrated in an automated processing chain for standardized and repeatable processing of

inundation time series (Gstaiger and Gebhardt, 2011). The histogram threshold based method is based on the assumption that water surfaces are forward scattering the radar signal resulting in low backscatter signals to the sensor. It uses multiple grey level thresholds and image morphological operations. The derived inundation maps were validated by several field surveys in the Mekong Delta. The accuracy at the edges of the inundation maps is estimated as 1–2 pixels, i.e. 90–180m for the ASAR derived inundation maps. Further details can be found in (Gstaiger and Gebhardt, 2011).

5.2 Multi-objective optimization

For the automatic calibration of such a large-scale hydrodynamic model the framework shown in Figure 5.2-1 has been developed. It consists of the following four elements:

- parameter classification
- multi-objective calibration algorithm
- formulation of objective functions
- parallelization scheme

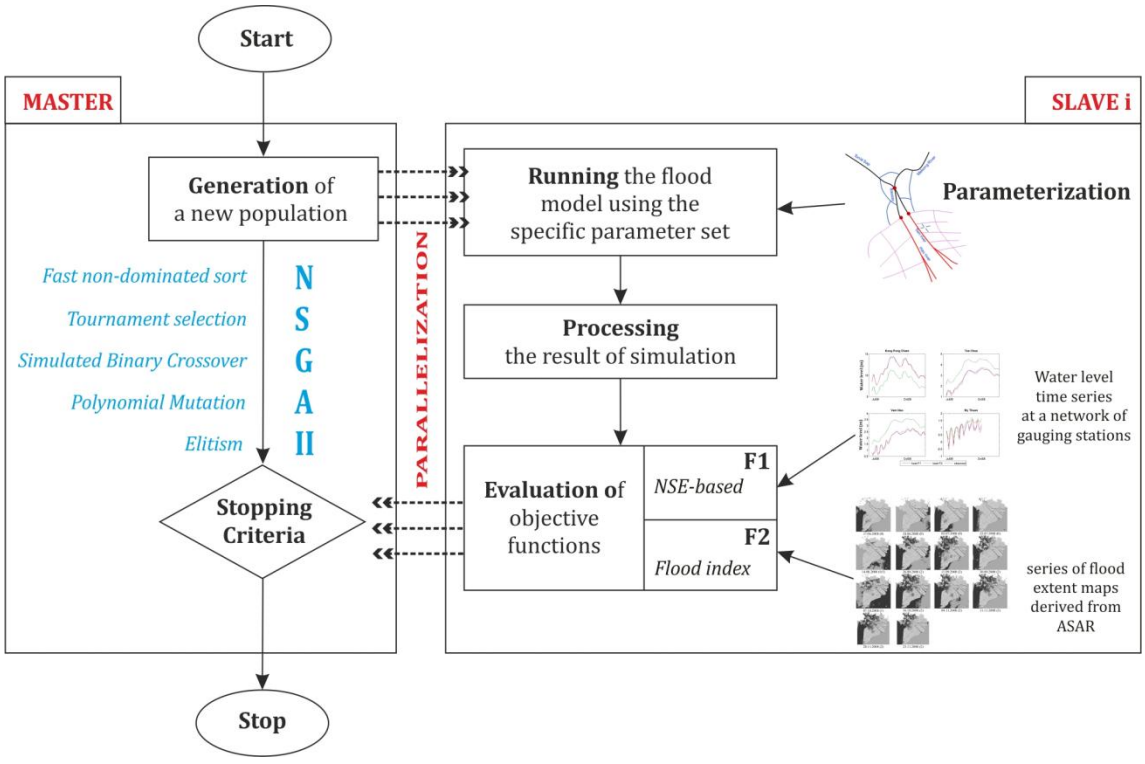


Figure 5.2-1: The calibration framework for the hydrodynamic model

5.2.1 Parameter classification

The calibration of the model is performed by adjusting the roughness parameters, which is the standard procedure in the calibration of hydraulic models. However, with a model of more than 26000 computational nodes, resp. possible different roughness parameters, is it obvious that the number of adjustable parameters has to be reduced. Otherwise the optimization problem would have too many degrees of freedom to be solved unambiguously. This is achieved by classifying the channel and floodplain elements. Five classes were defined: The main rivers in the Cambodian model domain (Mekong and Bassac), the main rivers (Tien and Hau) and major channels in the Vietnamese part of the Delta, the floodplains in Cambodia and in Vietnam respectively, due to their different characteristics, and the remaining medium and small channels in Vietnam. Figure 5.2-2 illustrates the spatial distribution of the different classes.

For each class feasible ranges of roughness parameters were defined. The ranges are comparable to those published in hydraulic text books and other publications (e.g. Chow et al., 1988; Werner et al., 2005; Pappenberger et al., 2007b). The ranges chosen enclose the values typically expected for the flow conditions, but are also large enough to enable compensation of model errors – most likely geometric errors – through the calibration. However, the ranges were limited to ranges preventing model instabilities. Table 5.2-1 summarizes the roughness parameterization.

Table 5.2-1: Assignment of roughness values to five classes of channels and floodplains

No.	Group's Name	Strickler's *[Manning's]* coefficient range	Descriptions
1	Mekong_Bassac	[20,60] *[0.016,0.05]*	Branches used to model Mekong River in Cambodia, Bassac River in Cambodia, Tonle Sap Great Lake
2	Tien-Hau	[20,60] *[0.016,0.05]*	Mekong River in Vietnam (Tien River), Bassac River in Vietnam (Hau River), and major branches of these rivers
3	CambodiaFP	[10,50] *[0.02,0.1]*	Branches for modeling Cambodia floodplains
4	VietnamFP	[10,50] *[0.02,0.1]*	Artificial branches for modeling Vietnam floodplains
5	Global	[20,60] *[0.016,0.5]*	Other from above (remaining branches)

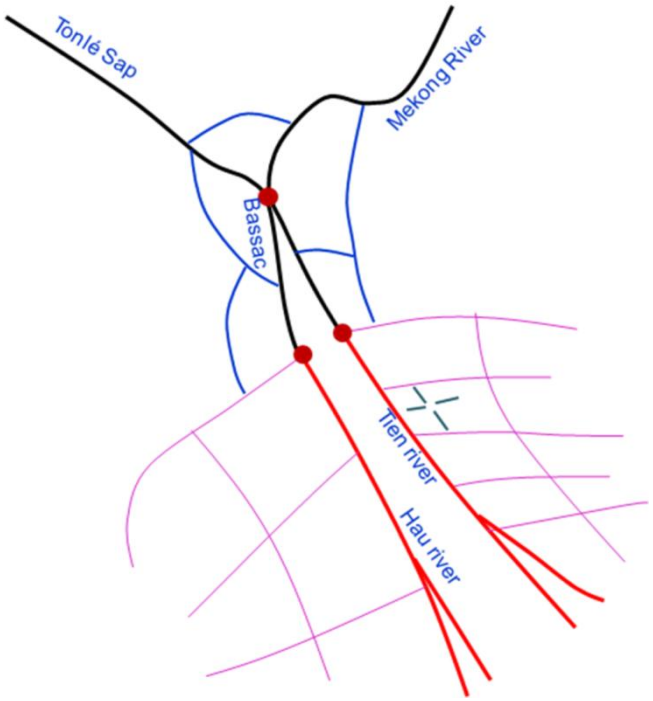


Figure 5.2-2: Scheme of the roughness classification of channels and floodplains: black for Mekong and Bassac in Cambodia, red for Tien and Hau river and major channels in Vietnam, blue for Cambodian floodplains, green for Vietnamese floodplain compartments and magenta for the remaining channels in the Vietnamese delta (cf. Table 5.2-1).

5.2.2 Definition of objective functions

5.2.2.1 The first objective function

Using the stage hydrographs recorded at 12 stations along the main stream of the Mekong and Bassac rivers, the first objective function evaluates the temporal performance in simulating water levels in the main channels and is formulated based on the Nash-Sutcliffe model efficiency coefficient (see Table 4.4-1):

$$F_1 = \sum_{i=1}^{n_S} \omega_i^S F_i^S \tag{5.1}$$

where:

$\sum \omega_i^S = 1$ and

$$F_i^S = 1 - \frac{\sum_{t=1}^{n} h_t^2 \, (W_{i,t}^{OBS} - W_{i,t}^{SIM})^2}{\sum_{t=1}^{n} h_t^2 \, (W_{i,t}^{OBS} - \bar{W}_{i,t}^{OBS})^2}$$

In the above equations, $W_{i,t}^{OBS}$, $W_{i,t}^{SIM}$ are the observed and simulated water level, respectively, at station number i and at time t, $\bar{W}_{i,t}^{OBS}$ the average observed water level at station number i, n_S the number of stations, n the number of time steps in the calibration period. h_t and ω_i^S are weighing coefficients. h_t indicates the importance given to particular portions of the hydrograph. This reflects the idea that it is difficult to obtain a model which perform equally well for high and low flows (Madsen, 2000), and that, in our flood study, high flows are of higher importance. Therefore, the flood peak period of the hydrograph is given a higher weight. ω_i^S indicates the importance given to a certain location of the network of gauging stations. F_i^S is the weighted form of the Nash-Sutcliffe coefficient given to station number i. F_1 denotes the first objective function which is the maximization form of the weighted average of the station coefficients mentioned. The optimal solution is $F_1 = 1$.

The weights were assigned to the single stations according to the following rationale: Gauging stations located closer to the sea and showing a large impact of the ocean tides even during high flows are given a lower weight. This reflects the lower impact of the flood wave on the inundation compared to the tidal influence. Furthermore, gauging stations in Cambodia were also assigned with a lower weight because of their relatively low impact on the inundation in Vietnam, which is the main focus of this study. Table 5.1-1 gives the weights associated to the different stations.

5.2.2.2 The second objective function

The second objective function evaluates the spatial performance of the model in predicting inundation extent utilizing the series of ASAR derived flood extent maps. As shown in chapter 4 and Table 4.4-3, several approaches to compare simulated and observed inundation extents have been proposed and discussed (Aronica et al., 2002; Hunter et al., 2005b; Pappenberger et al., 2007a; Schumann et al., 2009a). The most recommended measure is the flood area index (see Section 4.4), which is a binary pixel-wise comparison of observed and simulated flood extent maps and which is formulated for a single flood extent map as:

$$F_i^M = \frac{P_i^{11}}{n - P_i^{00}} = \frac{P_i^{11}}{P_i^{11} + P_i^{10} + P_i^{01}} \tag{5.2}$$

where:

- P_i^{11} is the number of pixels for which simulation and observation indicate “wet”.
- P_i^{10} is the number of pixels for which observation indicates “wet” and simulation indicates “dry”.

- P_i^{01} is the number of pixels for which simulation indicates "wet" and observation indicates "dry".
- F_i^M is the flood area index to the flood map number i.
- n is the number of cells taken into account.

The deficiencies of this measure, for example bias towards large inundation extent, are known and reported. Nonetheless, due to the lack of better alternatives up to date, it is still the basic measure used and recommended for deterministic calibration (Schumann et al., 2009a). In this study, we accept this limitation and put the focus more on the development and testing of automatic calibration routines instead of improving the goodness-of-fit measure for the inundation extent. However, since the hydrodynamic model is basically one-dimensional and does not deliver inundation maps directly, the method for deriving the flood area index had to be revised. Interpolating a two-dimensional flood extent map for comparison with the observed inundation extent from the nodes of the one-dimensional model, which is a quite error-prone procedure especially in the complex and heavily dike protected floodplains in Vietnam, was an inappropriate option.

Therefore, the following method was developed, which also considers uncertainties of the simulation (by the model setup and imperfect spatial representation) and flood maps (by classification errors and geo-referencing). Figure 5.2-3(a) shows the overlay of a flood extent map and a typical floodplain in Vietnam as represented in the model. The red dot at the junction of the four floodplain branches represents the inundation state and depth of an enclosed floodplain compartment. This is overlain by a single pixel of the flood extent map (yellow in Figure 5.2-3(a)). The probability of being flooded of the simulated node P_{SIM} representing the floodplain compartment is defined by a fuzzy set, i.e. a membership function is assigned to each floodplain node as shown in Figure 5.2-3(b). Just one computational node at the centre of the compartment (red node in Figure 5.2-3(a)) is taken into account in the comparison. However, this is an imperfect representation of the inundation state of the area around the node in the compartment. We therefore assume that the higher the water depth at the node, the higher is probability of the area around the node being flooded completely. If the water depth is lower than or equal to 5cm, the probability is defined as 0 given the uncertainties of the DEM and the actual micro-topography. In other words, with simulated water levels below 5 cm the probability of the major parts around the node of the compartment being flooded is zero. On the other hand, if the water depth is higher than 30cm, it is assumed that the area surrounding the node is inundated completely. This assumption is based on the typical micro-topography, especially the height of the low dikes surrounding paddy fields. The probability of inundation for water levels between 5 and 25 cm rises linearly. And, in order to reduce the spatial error in comparing just a

single pixel of the flood extent map with the state of the node for the floodplain compartment, also the neighboring eight pixels are included in the performance evaluation. Here, the probability of being flooded P_{SAR} is determined by the proportion of the nine cells identified as flooded in the extent map.

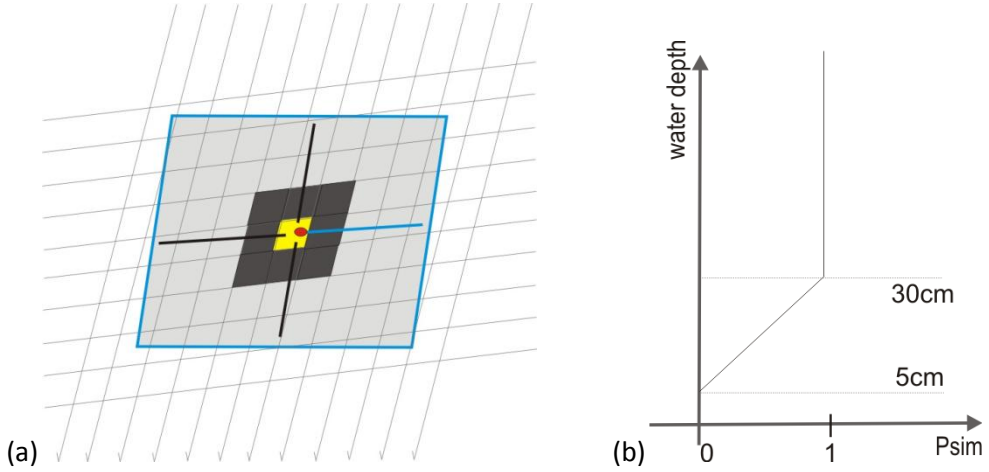


Figure 5.2-3: Illustration of the evaluation of the spatial performance of the model using flood extent maps. a) Representation of the diked floodplains in the Vietnamese part of the Delta in the model overlain by the flood extent map. Red dot: node that represents the inundation state of the area in the floodplain compartment surrounding the node in the model; yellow pixel: pixel of the extent map matching the node; gray pixels: neighbouring pixels of the extent map used in the performance evaluation. b) Fuzzy membership function used for the determination of the floodplain compartment nodes as being flooded

By the assignment of probabilities of a floodplain compartment of being flooded both in the simulation and the mapping, the performance of the model can be evaluated probabilistically in a Monte-Carlo procedure comprised of two steps:

Step 1:

- generate a random number r_{SIM} in (0,1) for every simulated node; if r_{SIM} is smaller P_{sim}, then this node is considered being wet, otherwise it is dry.
- generate a random number r_{SAR} in (0,1) for every 9-pixel cell; if r_{SAR} is smaller P_{SAR}, then this cell is considered being wet, otherwise it is dry.
- repeat the actions above for all flood cells

Step 2:

- calculate the measure F_i^M using Equation 5.2

The above two steps are repeated 1000 times (Monte Carlo sampling). The median (50% percentile of the distribution function of F_i^M) is considered as the goodness-of-fit measure based on a single map i.

To calculate the second objective function, F_i^M of the individual extent maps are combined as a weighted sum:

$$F_2 = \sum_{i=1}^{n_M} \omega_i^M F_i^M \tag{5.3}$$

where: $\sum \omega_i^M = 1$

F_2 is the second objective function to be maximized, with ω_i^M as the weighing coefficient which indicates the importance given to flood extent map i. If all flood extent maps would match perfectly, $F_2 = 1$. The weighting coefficients used are shown in Figure 5.1-1. Emphasis is given on maps covering the whole floodplain in Vietnam and mostly acquired during the flood season. To four early "almost dry" maps the value 0 was assigned, because our study is mainly focused on flood inundation modeling. Maps which do not cover the whole area of interest were assigned a lower weight.

5.2.3 Optimization algorithm

A general introduction to the optimization algorithm has been given in Chapter 4. A multi-objective calibration problem can be stated as a minimization or maximization problem over several objective functions (Madsen, 2000). In this study, both objective functions are based on maximum forms (Nash-Sutcliffe coefficient and flood area index). Hence, the optimization of the maximum is suitable in this case. It can be formalized as:

$$Maximize(F), F = \{F_1(\theta), F_2(\theta)\} \tag{5.4}$$

Where $\theta = (\theta_1, \theta_2, \dots, \theta_{|D|})$ is a parameter vector located within the parameter space D. From the proposed parameter classification in Table 5.2-1, D could be defined as $[20, 60] \times [20, 60] \times [10,50] \times [10, 50] \times [20, 60]$ and $|D|$ equals five.

In this case, two objectives are involved. They reflect the model performance with respect to the selected calibration objectives. Higher values in $F_i(\theta), i = 1,2$ indicate better model performance in the calibration objective i. In the ideal case with perfection in the model structure, model conceptualization and data measurement, the solution to Equation 5.4 would be a single unique parameter set. However, practically, they rather consist of a set of Pareto optimal solutions, which are best solutions from a multi-objective point of view (see Chapter 4).

As described in Chapter 4, genetic algorithms (GA) are well suited heuristic methods for multi-objective optimization problems. They encompass selection, mating, crossover, and mutation principles. Since the pioneering work by Schaffer (1985), a number of studies on multi-objective genetic algorithms (MOGA) have emerged (Konak et al., 2006). Most of these studies were motivated by a suggestion of a non-dominated GA outlined in (Goldberg, 1989). The most notable one is the Non-sorting Genetic Algorithm (NSGA) II developed by Deb (2002). Compared to its previous version NSGA, first proposed by Srinivas and Deb (1995), NSGA-II shows a significant improvement by using a fast non-dominated sorting algorithm and the elitism concept. Several calibration studies of hydrological models based on NSGA-II have been published (Khu and Madsen, 2005; Kollat and Reed, 2006; Madsen and Khu, 2006; Bekele and Nicklow, 2007; Fenicia et al., 2007; Reed et al., 2007; Tang et al., 2007). Table 5.2-2 illustrates the sequence of computational steps in the application of the NSGA-II.

Table 5.2-2: Sequence of computational steps of the NSGA-II algorithm

Step	Descriptions
1	**Population Initialization** - Generate an initial population P of size N. (Genes of each p in P are assigned by random sampling in the feasible space D defined above) - Evaluate the fitness of each individual in P.
2	**Ranking: fast sorting and density assignment** - Sort P based on domination to form fronts. (Fast non-dominated sorting which reduces significantly the complexity of the algorithm compared to the original NSGA) - Compute "density" or "crowding distance" for each individual in P (see Chapter 4) - Make a complete ordered sorting using both above concepts.
3	**Offspring Generation: applying genetic operators** - Form a pooling group of a size less than N of the current population using binary tournament selection - Implement crossover using Simulated Binary Crossover (SBX) to create a temporary offspring set of size of N. - With a predefined probability for the mutation (recommended to be $1/N$), mutate each individual of the above set using Polynomial Mutation to create the offspring population Q of size N.
4	**Forming Next Generation Population** - Combine P and Q to make R (size of $2 \times N$), then rank R based on domination and diversity ("crowding distance"). - Select the best N individual in R to form next generation (elitism –see Chapter 4) - Back to step 3. (The loop consisting of step 3 and step 4 is called main loop)

As proved in (Deb et al., 2002), the overall computational complexity[2] of the algorithm is $O(mN^2)$ where m is the number of the objectives and N is the size of the initial population. This is a significant improvement compared to its previous version.

Table 5.2-3 gives the values of the main parameters used for the NSGA-II in this study.

The stopping criteria of the process

In multi-objective optimization problems, several terminating conditions can be applied to stop the calibration process, e.g. the limit of evaluation time, a fixed number of generations. The latter can be defined at the beginning of the calibration process. Another way to determine the termination of the process is to measure the improvement over the iterations by comparing a function based on the two Pareto parameter sets which belong to two successive populations (Shafii and De Smedt, 2009). In this implementation of the calibration framework a fixed number of loops MI (seeTable 5.2-3) was used. Due to the computational demand for the whole process, MI should be chosen so that the initial population N multiplied by MI is not exceedingly large. The choice of N is elaborated in detail in Section 4.4. Through a series of implementing tests with N ranging from 36 to 60 and MI from 30 to 50, MI was fixed to 30.

Table 5.2-3: Summary of parameter settings for the multi-objective optimization in the master-slave parallelization scheme.

Parameters	**value**
Initial population (N)	52
Mating pool size	$N/2$
Tournament size	2 (binary)
Crossover probability	1
Mutation probability	$1/N$
Number of processors (nP)	12 – 14
Maximum iterations (MI)	30

5.2.4 On the normalization of objective functions

Many Multi-Objective Evolutionary Algorithms (MOEAs) use a distance metric to ensure a well-spread distribution of individuals along the Pareto front. In NSGA-II, it is the crowding distance d. However, the individual fitness values may not operate over a comparable scale. It is,

[2] Complexity measure of an algorithm which is commonly used in computer science

therefore, important to consider and adapt to widely disparate scaling among different objectives (Pedersen and Goldberg, 2004).

In the calibration, two objective functions were formulated using different performance measures. The first objective function is based on the Nash-Sutcliffe coefficient, hence, it ranges from $-\infty$ to 1 (not fully bounded). The second objective function is founded on flood indexes, which vary between 0 and 1 (bounded). In the published version of NSGA-II proposed by Deb et al. (2002), the crowding distance method did not discuss on the scaling factor which may cause bad distribution on the Pareto front (Weise, 2009). Or, the crowding distance only took account for the case of known bounds (Pedersen and Goldberg, 2004). The normalization is formalized as: (see 4.3-3)

$$NS[j].cd = NS[j].cd + \frac{|NS[j+1].f_i - NS[j-1].f_i|}{f_i^{max} - f_i^{min}} \tag{5.5}$$

where $NS[j].cd$ is the crowding distance for an individual j in a non-dominated setNS, $NS[j].f_i$is the value of the i^{th} objective of the individual j. In the literature, f_i^{max} and f_i^{min} are recommended to be fixed or equal to the two bound values of the objective i. Without the normalization by f_i^{max} and f_i^{min} the crowding distance would be dominated by the objective functions with larger values. Hence the distribution of solutions on the Pareto front would be biased towards those objectives.

For the case of unbounded or not fully bounded range of parameters, it is necessary to find another way for the normalization. Therefore f_i^{max} and f_i^{min} will be selected as local maximum and minimum values of F_1 and F_2 for each Pareto front and the normalization is consequently performed on these varying bounds.

5.2.5 The master slave parallelization scheme

In order to facilitate the automatic calibration of the model, a parallelization scheme for the optimization process was implemented. Computational time is still the bottleneck of automatic calibration on a single processor unit.

In this study, a typical model run for the simulated time period of five months took roughly 150 minutes, which is not extremely long, but still too long for automatic calibration on a single processor requiring several hundreds or even thousands of model runs. Therefore a master-slave parallelization routine was designed and implemented on a Windows-based computational server with 16 processors. In a master-slave scheme the master processor controls the communication and work load of subordinated slave processors (Tang et al., 2007). Applying this scheme to the calibration framework (see Figure 5.2-1), the master processor has a fully functional version of

the NSGA-II that uses slave processors to evaluate solution and return objective values to perform all of the required evolutionary search operations.

The number of processors nP used for the computation was selected between 12 and 14. The reason is that nP should be big enough to maximize the utilization of multiprocessors units but should not be too big to cause the unit to stall by processor communication overload. And, in order to optimize the performance of the computation, the size of the initial population N and the number of CPUs nP should suit (see Table 5.3-2). In general, nP should be an integer divisor of N in order to achieve optimal synchronization of the model runs of the slaves. In the present study, for example, we set N to 52 and nP to 13.

5.2.6 Computational framework

As mentioned before, MIKE 11 was used to create the flood model in this study. The MIKE software package contains already a generic tool for automatic calibration AUTOCAL. The module includes two global optimization algorithms (Shuffled Complex Evolution SCE and Population Simplex Evolution PSE), which were applied in some studies on hydrological model calibration (Madsen, 2000; Madsen, 2003; Blasone et al., 2006; Ngo et al., 2007). However, the current version of AUTOCAL is not appropriate for our purpose for the following reasons:

- Both algorithms are designed for single objective calibration problem (although they can be suited for multi-objective optimization by aggregating single objective calibration (see Section 4.3.1).
- The format of the input and output file should follow the format supported by AUTOCAL, which at the moment does not support the use of flood extent maps or spatial information in general.
- It requires the installation of the additional commercial OfficeGrid module for parallel computing facilities. Possibly, this feature will be integrated into the future version (Dr. Hendrik Madsen, DHI modeling expert, personal communication).

DHI MIKE is a commercial software package, therefore access to the source code in order to implement the required changes is not possible. In order to meet the requirements of this part of the study, wrapper codes around the actual hydrodynamic model were developed. A parallel version of NSGA-II has been developed to control the calibration process and serve as a wrapper for all components. The scripting language Python was selected due to available packages for parallelization (http://wiki.python.org/moin/ParallelProcessing). Furthermore, Python is the main scripting language in ArcGIS (http://resources.arcgis.com), which could be included in the

evaluation of the second objective functions in future applications, e.g. for two-dimensional hydrodynamic model calibration, where a direct comparison of inundation extents is possible. Others issues to deal with were reading the ASCII output files of MIKE11 to retrieve the simulation results and comparison with the flood extent maps.

5.3 Results and Discussions

5.3.1 Calibration

The flood model was used to simulate the flood season of 2008. A time step of 30 minutes was chosen to maintain model stability. For the optimization algorithm N was set to 52, the maximum iterations to 30, and hence the total number of model runs and objective function evaluations to 1560. The number of processors used was 13. The evaluation of each population took about 10 hours, i.e. the whole calibration process took about 300 hours or 12.5 days. Figure 5.3-1 shows the Pareto front of the final population, which consist of 52 Pareto-optimal solutions. Table 5.3-1 lists the parameter sets and objective function values for the optimal solutions for both F_1 and F_2.

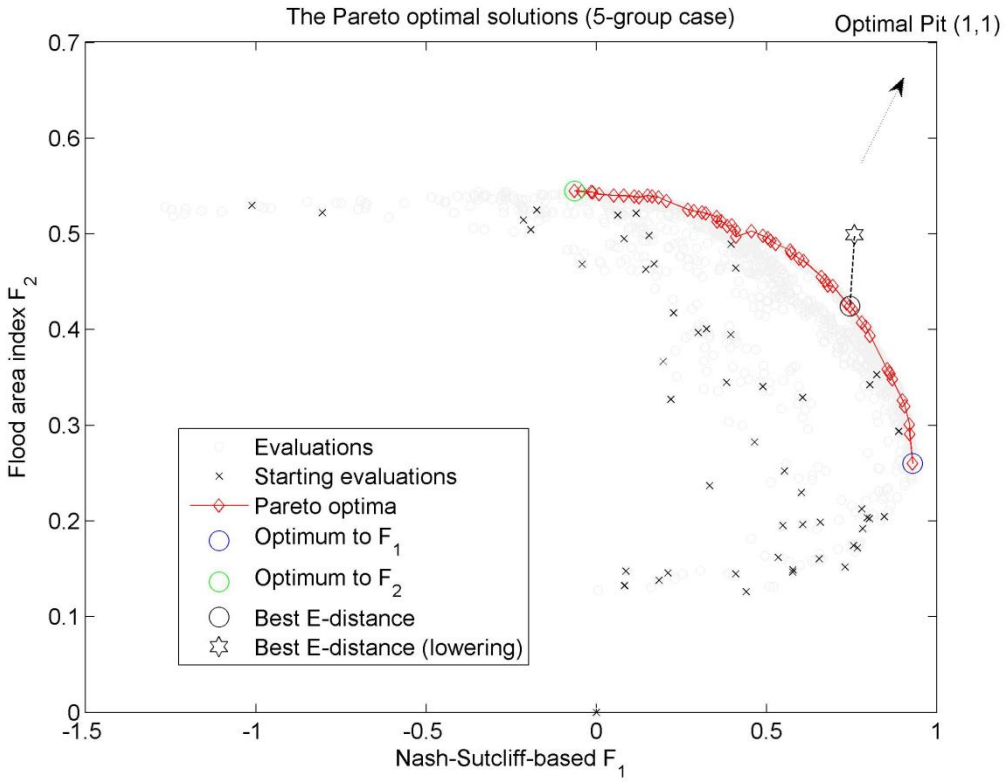


Figure 5.3-1: Pareto-optimal solutions of the final population maximizing the objective functions.

Table 5.3-1: Parameter sets (for 5-group case) and objective function values for best solutions for F_1 (first line) and F_2 (second line) and for the best solution with respect to the Euclidean distance (third line) (see also Subsection 5.3.4).

Strickler roughness coefficient					**Objective functions**	
Global	Mekong	Tien-Hau	CamFP	VietFP	F1	F2
21.502	25.084	43.471	10.315	19.153	**0.930**	**0.242**
25.281	42.294	29.483	34.047	10.014	**-0.064**	**0.531**
20.82	24.81	34.51	10.00	14.86	**0.745**	**0.405**

Figure 5.3-1 and Table 5.3-1 indicate that the objective function values for F1 and F2 exhibit a significant spread over the final population and also on the Pareto front. The best solution for either objective can only be reached in combination with a rather poor performance in the other objective. However, F1 is more sensitive than F2 as the wider spread on the horizontal axis in Figure 5.3-1 illustrates. Because the objective function values were normalized (cf. Section 5.2.4), a direct comparison of the sensitivity by value range is valid.

Figure 5.3-2 shows the parameter distribution for the final population over the five parameter classes. The parameter distributions on the Pareto optimal set exhibit two distinct features: 1) As already indicated in Table 5.3-1, the best solutions for F1 and F2 show an almost contrary behavior, and 2) the largest spread in parameter range from the Pareto optimal solutions is observed in the Cambodian floodplain.

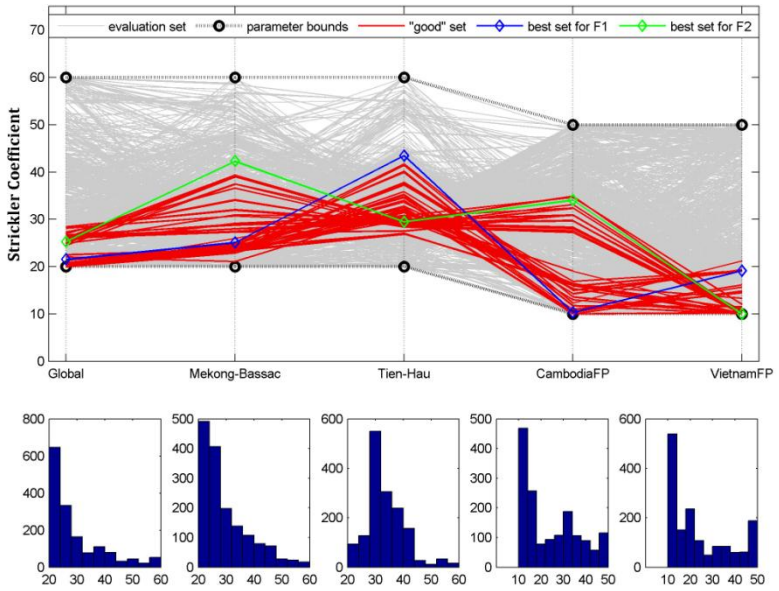


Figure 5.3-2: Parameter distribution of the simulations of the final population.

While these results are hard to explain by the parameter values alone, the comparison of the simulated hydrographs as well as the simulated inundation areas for the best solutions for F1 and F2 with the observed values provides more specific insights. Figure 5.3-3 shows the observed and simulated inundation areas for two different dates in the high flood period. Here it becomes obvious, that the best solution for F1 underestimates the inundation areas, especially in the Long Xuyen Quadrangle west of the Bassac/Hau river. Also in the Plain of Reeds the simulation exhibits some problems in simulating the observed inundation pattern. Some compartments are simulated as flooded while being observed as dry and vice versa. This situation is improved in the best simulation for F2, as expected. Here the Long Xuyen Quadrangle is inundated to a large extent and also the inundation pattern in the Plain of Reeds is matched better. However, this improvement comes at the cost of large errors in the simulation of the hydrographs as shown in Figure 5.3-4. In order to improve the spatial performance of the model the calibration routine tries to increase the water levels in the Vietnamese part of the Delta, for which the comparison of the inundation maps is performed. This is achieved by lowering the roughness (higher Strickler coefficients) in the Mekong/Bassac and in the Cambodian floodplains, causing less inundation and inflow in the Tonle Sap. This is illustrated by the simulated water levels in Kampong Cham, which are about 3 meters below the observed water levels. By this the buffering capabilities of the Tonle Sap in the model is reduced and more water is conveyed to the Vietnamese Delta, both through the main channels and the floodplains. In addition, the roughness was enlarged in the Tien and Hau rivers (lower Strickler coefficients) resulting in simulated water levels about 1 m above the recorded values (cf. stations Tan Chau and Vam Nao in Figure 5.3-4). This in turn causes larger inundation areas in the Long Xuyen Quadrangle and the Plain of Reeds mainly due to overflow of dikes. The overflow of dikes, respectively the dike elevations, controls the inundation of closed floodplain compartments in Vietnam, while the actual floodplain flow has only little influence on the inundation extent. This can be derived from the comparatively small changes in the roughness of the Pareto optimal solutions for the Vietnamese floodplains in Figure 5.3-2. These findings indicate that the dike elevations as implemented in the model are erroneous, despite the efforts taken in gathering the best possible information.

Possible error sources are the different datum and projections used by the different districts and provinces in the Delta when surveying the dikes. These often do not conform and can cause inconsistencies in the model. In general it can be noted that the spatial performance of the model is average at best and that the main reason is the representation of the dike elevations in the model. Another fraction of the error has to be attributed to the coarser representation of the floodplain compartments in the model compared to reality.

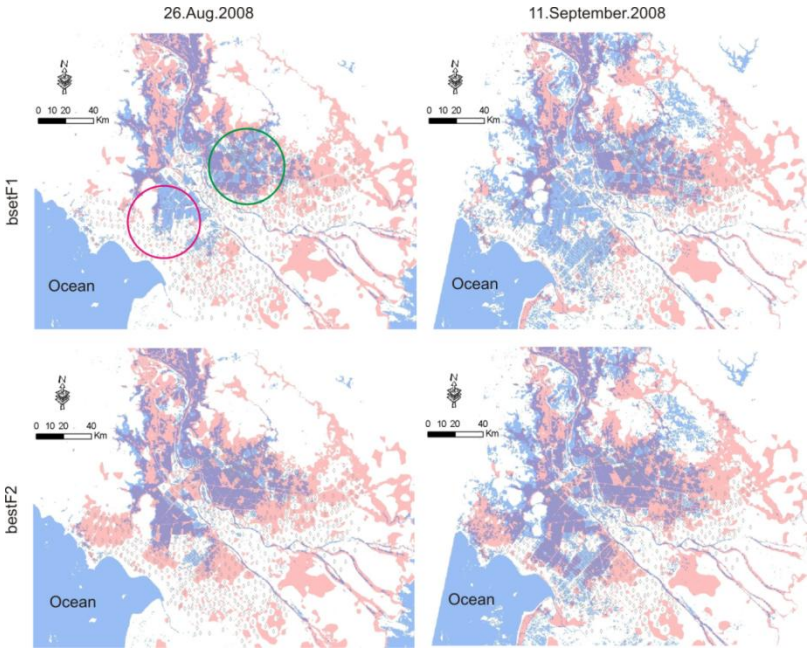


Figure 5.3-3: Inundation areas from ENVISAT ASAR (blue) and simulated inundation areas (light red) for best solutions for F1 and F2 for two satellite overpasses during the high flood period. Areas which are both observed and simulated wet appear in purple, both simulated and observed dry appear white. The red circle indicates the Long Xuyen Quadrangle, the green circle the Plain of Reeds. The small diamonds represent the floodplain compartments for which the flood area index was calculated.

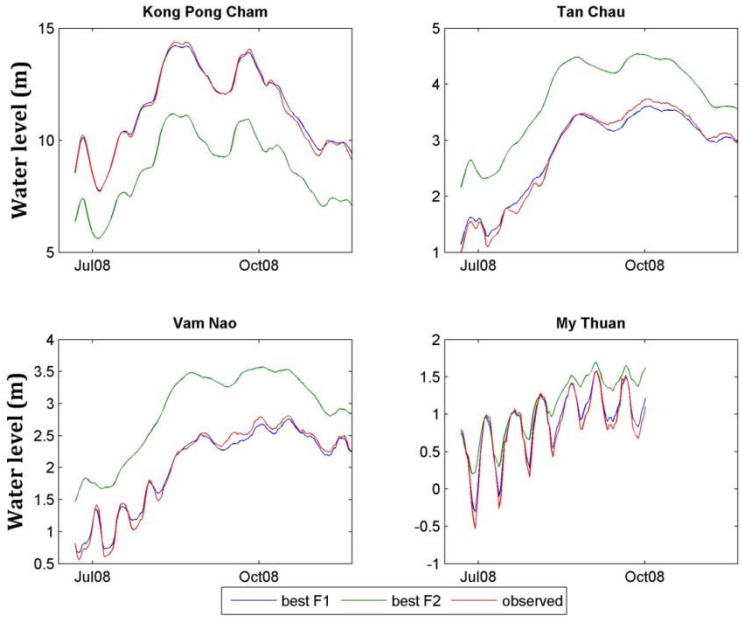


Figure 5.3-4: Observed and simulated hydrographs for best solutions for F1 and F2 for selected gauging stations.

5.3.2 Validation

Model validation is essential in order to document the predictive capabilities and credibility of the calibrated model. Validation must be carried out on independent data that have not been used in the calibration, and must assess the expected accuracy of the model predictions with respect to the model application being considered. For validation of the calibrated model, the same two types of data were used. Water level time series derived at the same stations used in the calibration process covering the flood season in the year 2009 are used. Two ENVISAT ASAR images with a spatial resolution of 90 m acquired on 27th August 2009 and 1st October 2009 are exploited. The flood extent maps utilized for the validation are derived by German Aerospace Center (DLR) (Gstaiger and Gebhardt, 2011).

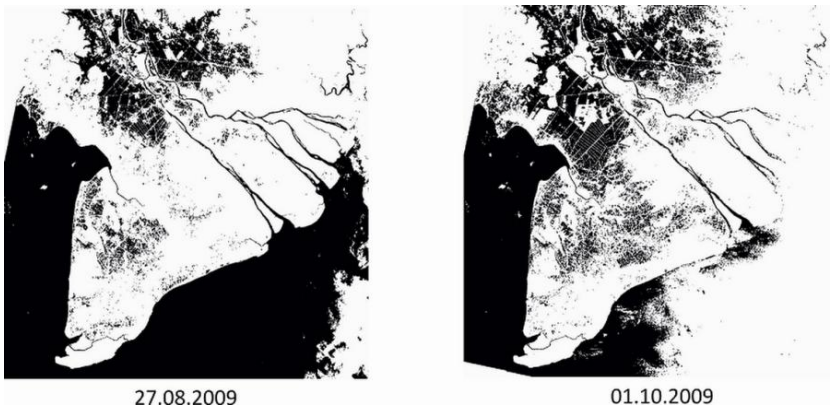


Figure 5.3-5: Flood extent maps derived from ASAR used for model validation.

The validation process is carried out for the calibrated model using the paretooptimal parameter sets. Table 5.3-2 shows the goodness-of-fit given for both the calibration and validation stages. As can be seen, the ranges of the statistics for the validation are comparable to those of the calibration. An interesting and encouraging aspect is that the best F1 parameters set from the calibration gives also the best performance in F1 in the validation.The same holds true for the second objective function: the same parameter set produces the best results in F2 in both calibration and validation. However, in case of the best Euclidean calibration parameter set, the validation simulation does not produce comparable results in F1, while for F2 comparable performance is achieved. Overall, the validation yielded results comparable to the calibration, as Figure 5.3-6 illustrates. Here the performance measures of calibration and validation are plotted against each other. For both F1 and F2 the scatter plots show an acceptable performance over all pareto optimal parameter sets (Fig. 5.3-6 lower row), with a slightly lower performance in the intermediate range between the optimal parameters sets for F1 and F2, as already mentioned above and shown in Table 5.3-2. This evaluation increases the confidence in the validity of the calibration process and the derived Pareto optimal parameter sets.

To further corroborate the validation, a spatially explicit evaluation of the simulated inundation areas was performed. Inundation areas were interpolated from the best F1 and F2 simulations of the validation and compared to the observed inundation areas for the two dates on 27th August 2009 and 1st October 2009. Figure 5.3-7 conveys comparable implications as for the calibration period, i.e. maps corresponding to best F1 do not map the inundation areas sufficiently, while the best F2 simulation reproduce the inundation pattern reasonable well. However, just as in the calibration, the good spatial performance of the best F2 parameter set comes at the cost of deteriorating performance in the hydrographs of the main channels (see Figure 5.3-8). These graphs show the same behavior as those in Figure 5.3-4, i.e. the water levels in the Vietnamese part of the Delta are simulated with a large bias. These findings again consolidate the hypothesis of an erroneous representation of the dike system in the model. Figure 5.3-8 also shows that the surveyed hydrographs used in the validation have some significant data gaps, which are at least partly responsible for the inferior performance in F1 of the validation simulations.

Table 5.3-2: Performance statistics of the calibration and validation simulations

	Calibration		**Validation**	
	F1	F2	F1	F2
Pareto optimal calibration range	[-0.064:0.930]	[0.240:0.531]	[-0.229:0.875]	[0.254:0.553]
Best F1 calibration point	0.930	0.242	0.875	0.254
Best F2 calibration point	-0.064	0.531	-0.189	0.553
Best Euclidean calibration point	0.745	0.405	0.592	0.399

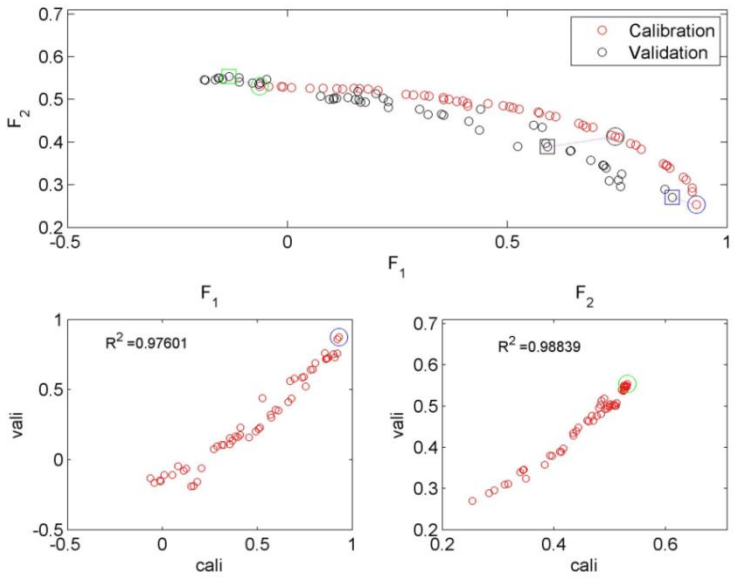


Figure 5.3-6: Checking the performance of the calibration – validation (5 group case)

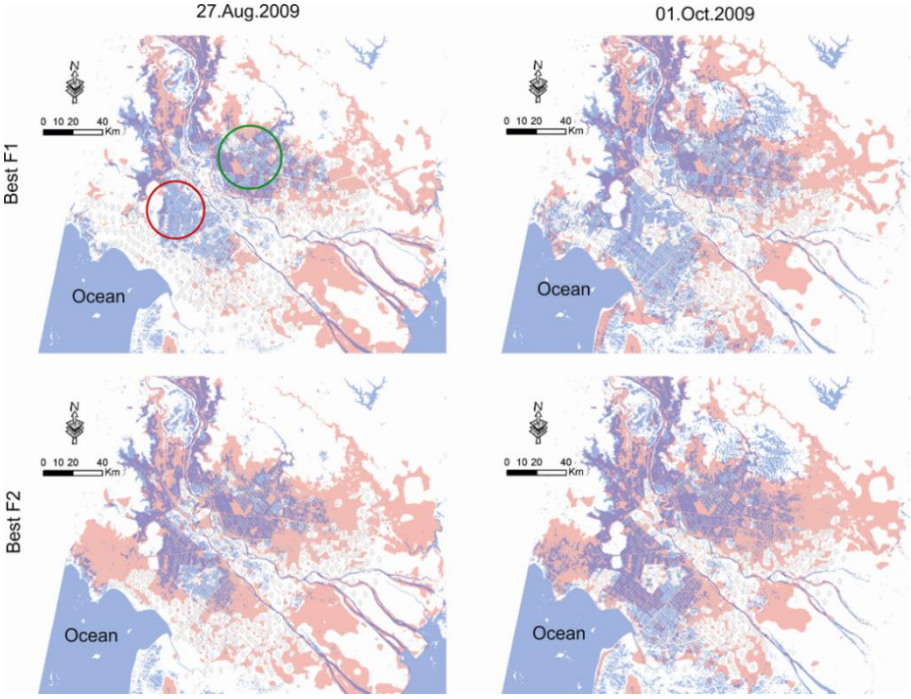


Figure 5.3-7: (Validation) Inundation areas from ENVISAT ASAR (blue) and simulated inundation areas (light red) for best solutions for F1 and F2 for two satellite overpasses during the high flood period. (legend translated the same as Figure 5.3-3 for the calibration)

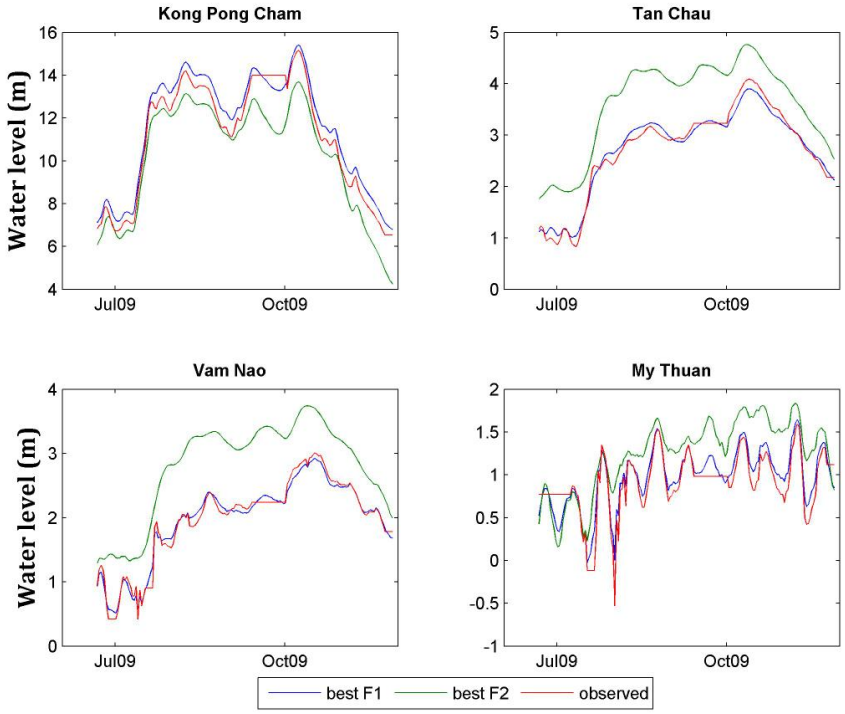


Figure 5.3-8: Observed and simulated hydrographs for best solutions for F1 and F2 for selected gauging stations (validation stage).

5.3.3 Re-calibration with 6 groups of parameters

A slight change was made on the parameter classification which was introduced in the Subsection 5.2.1. The Mekong-Bassac group is now divided into two smaller groups: Great Lake group consisting of only the branch modeling the Great Lake (>200 km long) and the Mekong-Bassac group containing the remaining. The reason of this grouping is the Great Lake (Tonle Sap Lake) is a huge lake which may have different resistance behavior compared to the other branches. Hence, 6 new groups of parameters are considered and illustrated in Table 5.3-3.The optimization problem described in Section 5.2.3 has been changed in term of the feasible space (searching space). The feasible space is now comprised of six dimensions, i.e. $D = [20, 60\,] \times [20, 60\,] \times [20, 60\,] \times [10{,}50\,] \times [10, 50\,] \times [20, 60\,]$. In general, the number of evaluations may be higher for a better convergence. However, due to the limited CPU-time, the same parameters for NSGA-II given in Table 5.2-3 were applied for this re-calibration, except for the number of the maximum iterations (now 40 instead of 30 as in the previous launch). The run time for this re-calibration took approximately 17 days.

Table 5.3-3: Assignment of roughness values to six classes of channels and floodplains (shaded region indicates the change compared to the five group case introduced in Table 5.2-1).

No.	Group's Name	Strickler's *[Manning's]* coefficient range	Descriptions
1	Mekong-Bassac	[20,60] *[0.016,0.05]*	Branches used to model Mekong River in Cambodia, Bassac River in Cambodia
2	GreatLake	[20,60] *[0.016,0.05]*	Tonle Sap Great Lake
3	Tien-Hau	[20,60] *[0.016,0.05]*	Mekong River in Vietnam (Tien River), Bassac River in Vietnam (Hau River), and major branches of these rivers
4	CambodiaFP	[10,50] *[0.02,0.1]*	Branches for modeling Cambodia floodplains
5	VietnamFP	[10,50] *[0.02,0.1]*	Artificial branches for modeling Vietnam floodplains
6	Global	[20,60] *[0.016,0.5]*	Other from above (remaining branches)

The results of the re-calibration process are shown in Figure 5.3-9, Figure 5.3-10 and Table 5.3-4. There is a similarity of the Pareto optimal solutions of the final population (after 40 generations of the population of size 52) between the two classification cases. The range of the

best F1 and best F2 objective functions is very similar. In the 6-group case, one additional group is added, hence, there is more flexibility in the search. As can be seen, the parameter value of the Great Lake freely moved from the group of the Mekong-Bassac. Therefore, it may help to find better solutions than those in the 5-group case. However, it this case, comparing the goodness-of-fit values of the 6-group case with the 5-group case, it turns out that the fit is quite similar. The improvement is not significant. It means that the 5-group of parameters is sufficient for the representation of the spatial roughness of modeled channels.

Table 5.3-4: Parameter sets (for 6-group case) and objective function values for best solutions for F_1 (first line), F_2 (second line) and the best solution in term of Euclidean distance to the optimal pit (third line).

Strickler roughness coefficient						**Objective functions**	
Global	GLake	Mekong	Tien-Hau	CamFP	VietFP	F1	F2
23.03	45.95	25.24	43.27	10.22	13.23	**0.930**	**0.252**
23.59	20.00	26.51	25.69	37.24	10.35	**-0.048**	**0.537**
20.00	55.99	24.66	34.76	10.00	18.10	**0.750**	**0.421**

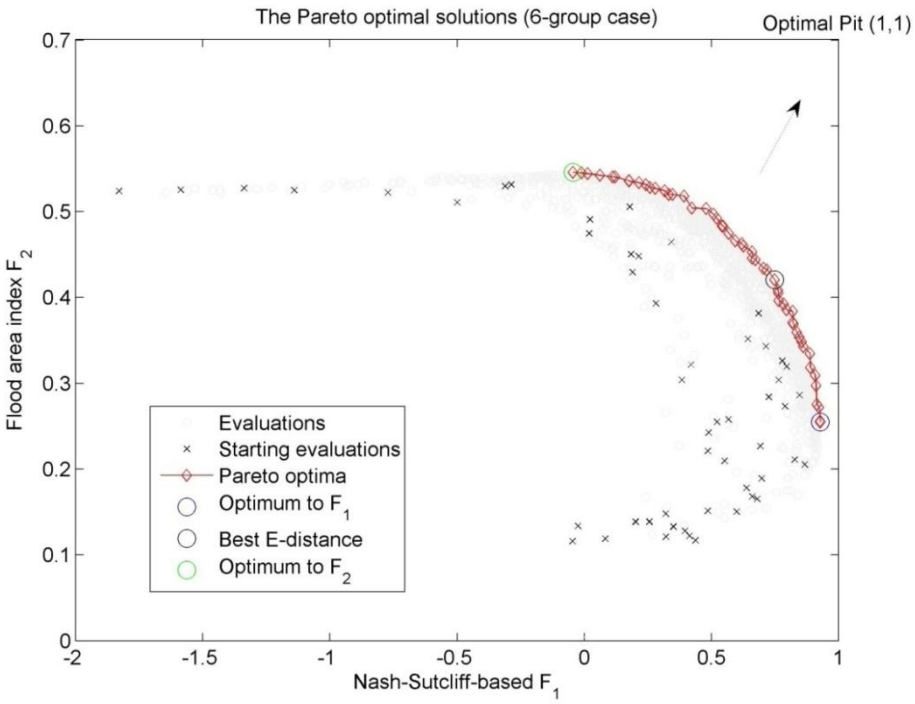


Figure 5.3-9: Pareto-optimal solutions of the final population maximizing the objective functions for the case of 6-group parameter set.

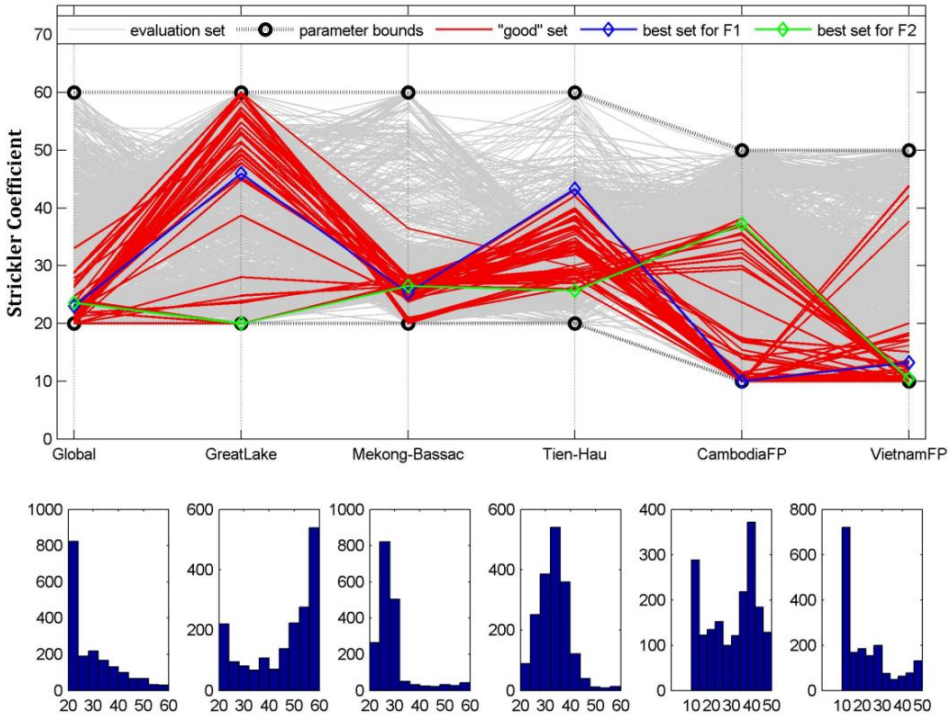


Figure 5.3-10: Parameter distribution of the simulations of final population (6 group case).

5.3.4 Simulation with best Euclidean parameter set and dike lowering

In order to test the hypothesis of wrong dike elevations, a simulation with lowered dike heights was performed. For this simulation the parameter set from the final Pareto optimal solution with the "smallest Euclidian distance" to the optimal pit (1,1) was selected:

$$min\left(\sqrt{(1-F_1)^2+(1-F_2)^2}\right) \tag{5.6}$$

Table 5.3-5 lists the parameter set fulfilling this criterion. This set was used simulating the inundation with dike heights generally lowered by 20% for the case of the 5-group parameter classification. The result in terms of model performance is given in Table 5.3-5 and in Figure 5.3-1. Both objectives show increased performance when lowering the dike heights by 20%. This corroborates the hypothesis that the dike representation in the model is responsible for the errors in predicting inundation extents. Further investigations on the real dike heights are therefore advised. However, this is not within the scope of this study.

Table 5.3-5: Best Euclidian distance parameter set of the final Pareto optimal population and performance values of F1 and F2 with original (top row) and 20%-lowered dike heights (bottom row)

Strickler coefficient					**Objective functions**		**Euclidian Distance to (1,1)**
Global	Tien-Hau	Meko-Bass	CamFP	VietFP	F1	F2	
20.82	24.81	34.51	10.00	14.86	0.745	0.405	**0.647**
					0.759	0.499	**0.556**

Chapter 6

FLOOD HAZARD MAPPING IN THE MEKONG DELTA

"Essentially, all models are wrong, but some are useful"

(George E. P. Box)

Current or future flood hazard is an indispensible input for flood risk analysis and assessment, in particular for an area as the Mekong Delta with natural high flood hazard that is expected to rise even higher due to climate change. Under these circumstances it is quite surprising that a flood hazard analysis for the Mekong Delta does not exist at present. The published reports present scenario simulations considering expected sea level rise (Hoa et al., 2007; Dung et al., 2009), but none of the studies considers changes in discharge triggered by climate change, nor do they quantify the probabilities of occurrences of flood events, which is a core aspect of a hazard analysis. Therefore this part of the study utilizes the calibrated hydrodynamic model for flood simulation as a core module for generating flood hazard maps based for the whole Mekong Delta. Moreover, it builds on recent publications on detected non-stationarities in the discharge variability of the Mekong and establishes a multivariate and non-stationary hazard analysis, which finally yields probabilistic flood hazard maps.

6.1 Motivation

Hydrological data, e.g. precipitation or flow discharges, being the main input for hydrological and hydraulic models, are usually examined using conventional frequency analysis in flood hazard analysis. Traditionally, stationarity and independence are the fundamental underlying assumptions. However, these assumptions remain doubtful because of the change in natural variability, especially in the context of climate change (Strupczewski et al., 2001; Khaliq et al., 2006; Adlouni et al., 2007). Thus, in this study, beside the use of the traditional approach of flood frequency analysis using the stationary assumptions, a new approach taking the non-stationarity into account is developed. In order to derive discharge hydrographs for the flood hazard mapping at Kratie, the upper boundary of the Mekong Delta in Cambodia, the three

essential variables flood peak, flood volume and hydrograph shape are analyzed. This is a necessary extension to the standard flood hazard analysis procedures based on annual maximum flows, where usually only the peak discharge is considered. In the Mekong Delta not only flood peak discharge is determining the flood hazard, but also the flood volume, as the devastating flood in the year 2000 illustrates. In this year the peak discharge was not extraordinarily high, but the flood volume was the largest recorded causing deep inundation through prolonged water logging and ponding in the flat floodplains of the Delta (Hoc, 2000; Hoi, 2005; MRC, 2005a). Univariate flood frequency analysis (stationary and non-stationary) is used for flood peak and flood volume. Therefore, a copula-based analysis is developed to deal with the bivariate frequency analysis of both flood peak flow and volume. Furthermore, a cluster analysis is applied to identify characteristic hydrograph shapes along with there probability of occurence. In the hazard analysis the characteristic flood hydrographs are scaled by discharge and volume with probabilities of occurrence, resp. return interval taken from the bivariate frequency distribution. Finally, flood hazard maps for the whole Mekong Delta are derived by the calibrated large scale hydrodynamic model of the Mekong Delta driven by the hydrographs derived for Kratie.

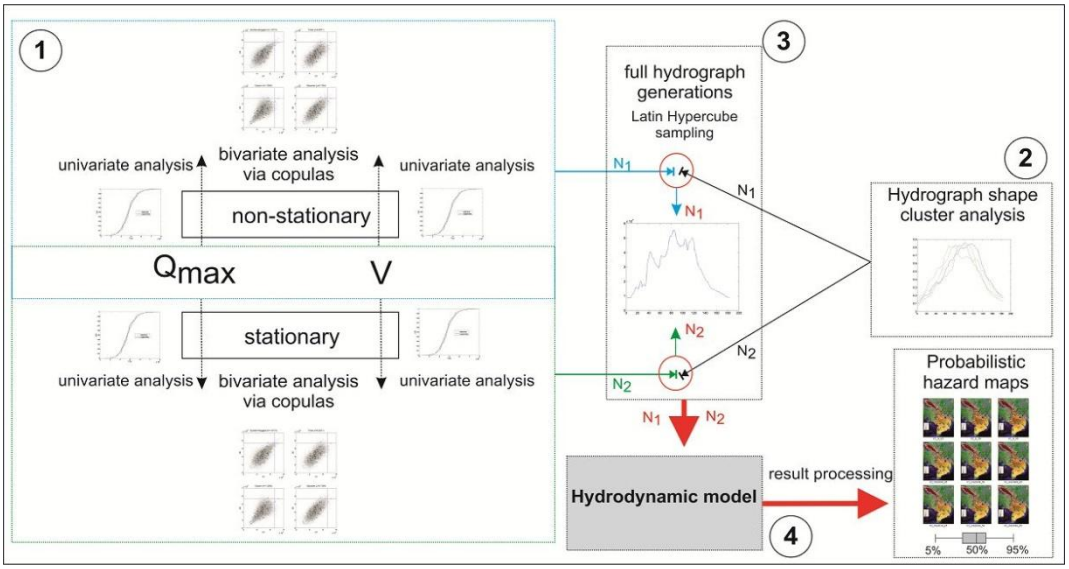


Figure 6.1-1: The strategic framework: (1) – frequency analysis on the peak and volume series, (2) – hydrograph shape cluster analysis, (3) – the generation of flood hydrographs and associated probability of occurrence and (4) – flood hazard mapping

To account for uncertainties in the hazard assessment, a Monte Carlo (Latin Hypercube sampling) framework sampling from combinations of flood peak and volume with identical probabilities of occurrences and different hydrograph shapes is applied. Figure 5 illustrates the general procedure of the synthetic hydrograph generation for the hazard analysis. By this method probabilistic hazard maps are produced for selected probabilities of occurrences. Additionally we applied both stationary and non-stationary frequency distributions and compare the results in

terms of hazard maps derived by the different approaches. By this comparison it is also possible to discuss the errors in the hazard analysis introduced by reported inappropriate statistical methods, in this case a stationary approach (Delgado et al., 2010).

6.2 Discharge series at Kratie

The Kratie station is located in Cambodia at the longitude 105.987 and latitude 12.240 (see Table 6.2-1). Kratie represents the upper boundary of the Mekong Delta, respectively the transition from the Lower Mekong Basin to the Delta. A daily discharge series from 1st January1924 to 31st December 2009 is used for the flood hazard analysis. Discharge at Kratie is considered as the drainage of the Mekong basin area of about 646000 km^2. The maximum specific runoff (unit discharge) is reported as one of the biggest in the world. Its value is about 0.12 $m^3/s/km^2$ (O'Connor and Costa, 2004; MRC, 2005a). For each year a fixed flood period from the beginning of June to the end of November is selected for the hazard analysis. This period covers the stable annual flood season of the Mekong (MRC, 2005a), i.e. the peak of each flood season always falls in this period. The flood volume is consequently calculated as the flood volume within this fixed period. The volume of flow may be expressed in cubic meters but this leads to very large numbers (e.g. the volume in Kratie) and therefore large volume units are commonly used (Reddy, 2005). In this study, the volume is calculated using the unit of day.m^3/s (or cumec-day) instead of m^3 as usual. Figure 6.2-1 shows the annual peak flow, the average flow and the minimum flow series of the flood period and the corresponding annual flood volume series. Figure 6.2-2 presents the peak flow and volume as scatter plots along with the histograms of the series. It can be seen that peak and volume are roughly linearly correlated. Table 6.2-1 lists some basic statistic parameters of the peak and volume series. Both series are slightly skewed to the right. However, the peak series shows stronger skewness, while the skewness in the volume series is very small. The Pearson correlation coefficient is 0.73, which implies a quite large linear dependence between the two variables.

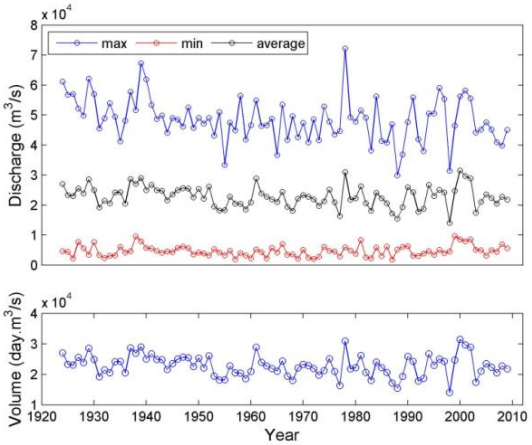


Figure 6.2-1: (top)-maximum, mean, minimum discharge time series; (bottom) - volume timeseries

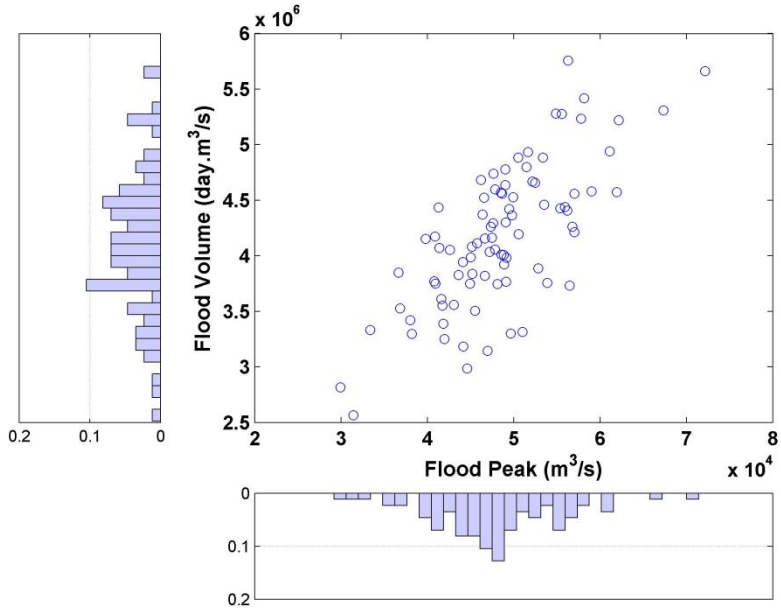


Figure 6.2-2 Observed flood peak versus flood season volume and histograms for 1924-2009 at Kratie

Table 6.2-1: Statistics of flood peak and volume series at Kratie

	Flood peak (m^3/s)	**Flood volume (10^3 day.m^3/s)**
Mean	48529	4178
Standard deviation	7366	642
Coefficent of skewness	*0.2782*	*0.0868*
Coefficent of kurtosis	*3.8344*	*2.9008*
Minimum	29920	2565
Maximum	72141	5757
Median	48309	4160
Interquartile Range	8684	818

6.3 Univarite flood frequency analysis

The traditional approach for flood frequency analysis is based on univariate extreme value statistics assuming independence and stationarity (Khaliq et al., 2006; Merz and Thieken, 2009; Delgado et al., 2010). Following these, the frequency analysis will first be carried out for both peak series and volume series.

6.3.1 Stationary case

For the univariate and stationary frequency analysis several probability distributions are used in research and practice. In this study we do not select a single distribution a priori, because no recommendations or previous studies exist for the Mekong Delta. Different distributions, which are often used in extreme value statistics, and the Gaussian normal distribution as reference, are used for fitting the observed peak and volume data series. The detailed descriptions of these distributions are presented in many publications (e.g. Stedinger et al., 1993; Hosking and Wallis, 1997; Robson and Reed, 1999; Rao and Hamed, 2000; Katz et al., 2002; Wilks, 2006). As candidate distributions were selected:

- Normal distribution 2 parameters (N2)
- Log-normal distribution (LN3)
- Gamma distribution (GAM2)
- Pearson type III distribution (P3)
- Log-Pearson type III distribution (LP3)
- Gumbel max distribution (GUM2)
- Generalized Extreme Value (GEV3)

6.3.1.1 Parameter estimation method

To estimate the parameters of the candidate distributions, we first use the L-Moment method (Hosking and Wallis, 1997) to make initial guesses for the parameters, then we apply the maximum likelihood method. The reason for using two steps for the estimation is, that although the L-Moment method has been reported to be powerful for small size data sets ($n<30$), it is not suitable for the parameter estimation of non-stationary series, resp. distributions, which will be considered later. However, using the L-moment parameter estimates as initial guesses for the maximum likelihood method increases the efficiency of the maximum likelihood method. The optimization algorithm used for the maximum likelihood estimation is the well-known single

objective evolutionary optimization algorithm Shuffled Complex Evolution (SCE-UA) which was shortly presented in Chapter 4.

6.3.1.2 Goodness-of-fit

To compare the goodness-of-fit among the different distribution in the fitting, the Akaike's Information Criterion (AIC) (Akaike, 1974) is used. The AIC is a way of selecting a model from a set of models under consideration. The AIC criterion adopted in this case is defined as:

$$AIC = -2 \times loglikelihoood + 2 \times number\ of\ parameters\ to\ be\ estimated \quad (6.1)$$

In general, the best model has the minimum AIC value. However, in practice, when examining the performance of a model in a set of candidate models, it is necessary to examine how much its AIC is larger than the minimum AIC of the model set. Guideline for this issue is clearly shown in Brunham and Anderson (2002).

Furthermore, the Kolmogorov-Smirnov (K-S) test (Wilks, 2006) is used to evaluate the goodness-of-fit of the individual distributions. The K-S test compares the empirical and theoretical cumulative distribution functions. For continuous distributions the K-S test is usually more powerful than the Chi-square test, and is thus preferred in this study.

6.3.1.3 The result of the fitting

Peak series

Table 6.3-1 shows the fitting parameters for all the candidate distributions. The Kolmogorov-Smirnov static values shown in Table A-1 (Appendix A) imply that there is insufficient evidence to reject the null hypothesis that the observed flood peaks are drawn from any of the considered distributions at the significant level of 10%. However, from Figure 6.3-1, it can be seen that the Gumbel distribution differs significantly from the remaining distributions. Also its AIC is the highest of all distributions (Fig. 6.3-2). It is, therefore, excluded from the selection of distributions. The other distributions give very similar AIC values. However, there are reasons to exclude some of them from the analysis. The Gaussian distribution is usually not appropriate to describe extreme value behavior. The Gamma distribution is not regularly used in flood frequency analysis. Although Pearson type III and Log-Pearson type III were recommended for flood frequency analysis (especially in USA), LN3 fits the data series better and has the simpler mathematical form as an additional bonus. GEV3 is likely the most frequently used distribution in flood frequency analysis. Therefore, it will be selected for the non-stationary case. In

summary the LN3 and GEV3 will be used for the stationary and univariate frequency analysis of the annual peak flows.

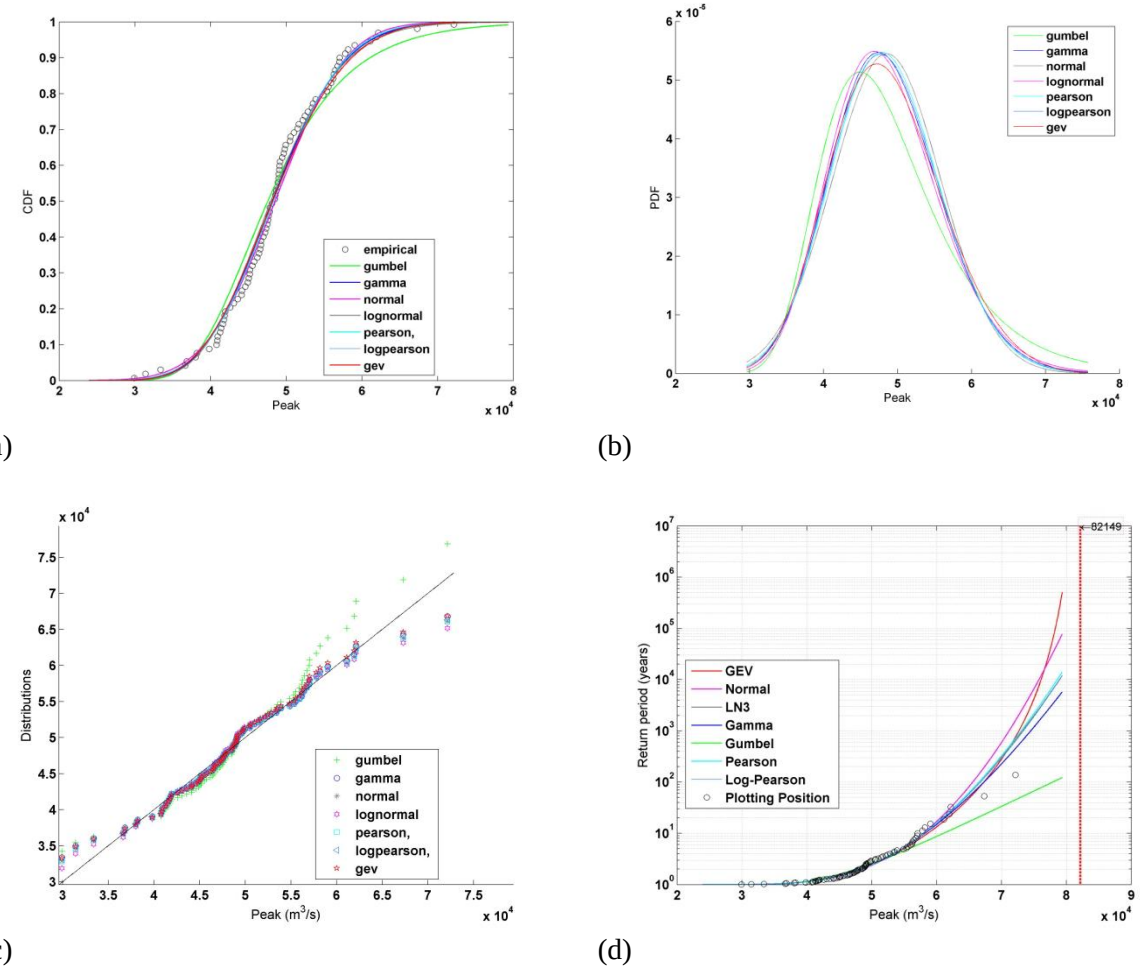


Figure 6.3-1: Fitting the peak series by the candidate distributions: (a) CDFs,(b) PDFs, (c) Q-Q-plot, (d) and flood frequency curves.

Table 6.3-1: Estimated parameters for fitting different distributions to the peak series

No.	Distribution	Estimated Parameters	Domain
1	N2	σ =7322; μ =48529	$(-\infty,+\infty)$
2	LN3	k =-0.065; α =7291; ξ = 48293	(-6463,+∞)
3	GAM2	α =43.43; β =1117	(0, +∞)
4	P3	α =112.0; β =662; γ =-32268.0	(-32268.0, +∞)
5	LP3	α =82.2 ; β =-0.017; γ =12.16	$(-\infty,190994)$
6	GUM2	σ =7163; μ =44931.0	$(-\infty,+\infty)$
7	GEV3	k =-0.195; σ =7116; μ =45675.0	$(-\infty,82170)$

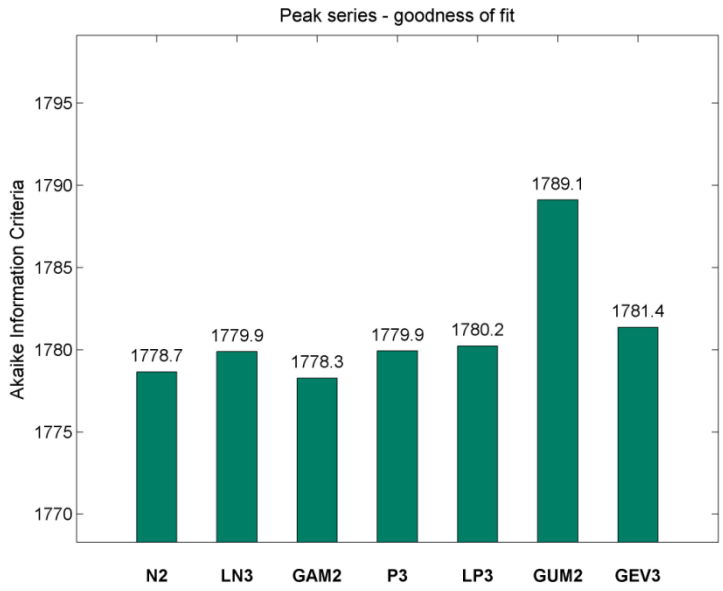


Figure 6.3-2: Akaike Information Criteria value (AIC) based on the maximum likelihood method for fitting the peak series to the 7 candidate distributions

Volume series

Figure 6.3-3 shows the results when fitting the volume data to the 7 candidate distributions. Table 6.3-2 shows the estimated parameters of the selected distributions. The Kolmogorov-Smirnov static values shown in Table A-2 (Appendix A) imply that there is insufficient evidence to reject the null hypothesis that the observed flood volume data are drawn from any of the considered distributions at the significant level of 10%. However, the Gumbel distribution perform worst and has a significantly larger AIC value compared to the others (see Figure 6.3-4). It is, therefore, excluded from the analysis. The Gaussian distribution fits well to the data. And different to the peak series, the volume series is not generated by an extreme process, since it is the volume integrated over the complete flood season.

Thus, the Normal distribution will be selected. The other distributions give very similar AIC values. The GEV distribution was removed from the selection because its AIC value (1781.4) is larger than that of LN3, GAM2, P3 and LP3 and because it is often applied for data which is generated from extreme mechanisms. GAM2, P3 and LP3 include rather complex mathematical terms, whereas LN3 is easier to handle, especially for the non-stationary analysis, when covariates are taken in account. For this reason, LN3 will be preferred and selected.

In summary, N2 and LN3 are further chosen for the non-stationary analysis of the volume series in chapter 6.3.2.

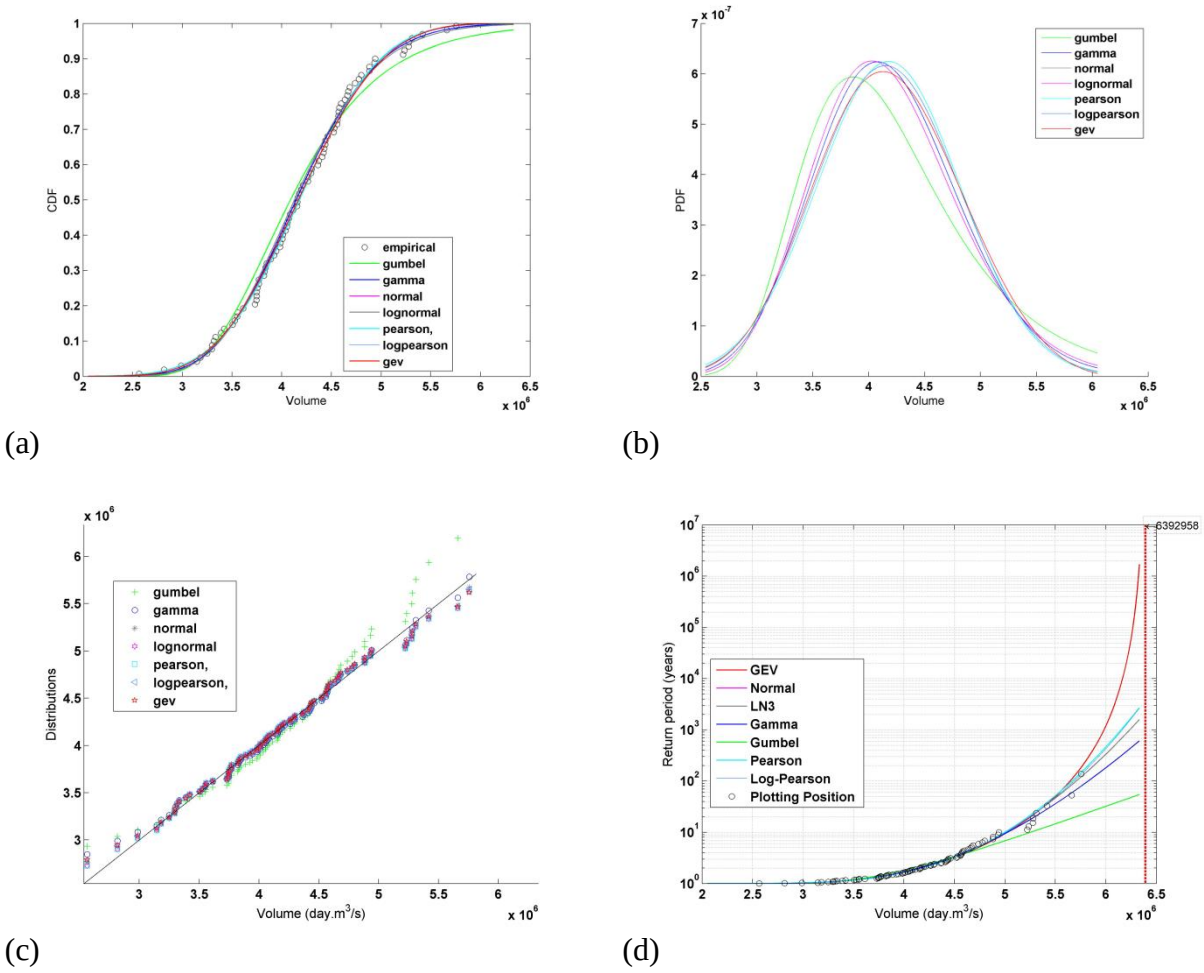


Figure 6.3-3: Fitting the volume series by the candidate distributions: (a) CDFs, (b) PDFs, (c) Q-Q plot, (d) and frequency curves.

Table 6.3-2: Estimated parameters for fitting different distributions to the volume series

No.	Distribution	Estimated Parameters – MLM	Domain
1	N2	σ = 638713; μ = 4177927	$(-\infty,+\infty)$
2	LN3	k = -0.031; α = 638324; ξ = 4167937	(-1654794, $+\infty$)
3	GAM2	α = 41.84; β = 9958.0	(0, $+\infty$)
4	P3	α = 638726; μ = 4177807 (special case)	$(-\infty,+\infty)$
5	LP3	α = 27.8; β = -0.03; γ =16.06	(0,9423575)
6	GUM2	σ = 619748; μ =3860463	$(-\infty,+\infty)$
7	GEV3	k = -0.258; σ = 631342; μ = 3945561	$(-\infty,6392957)$

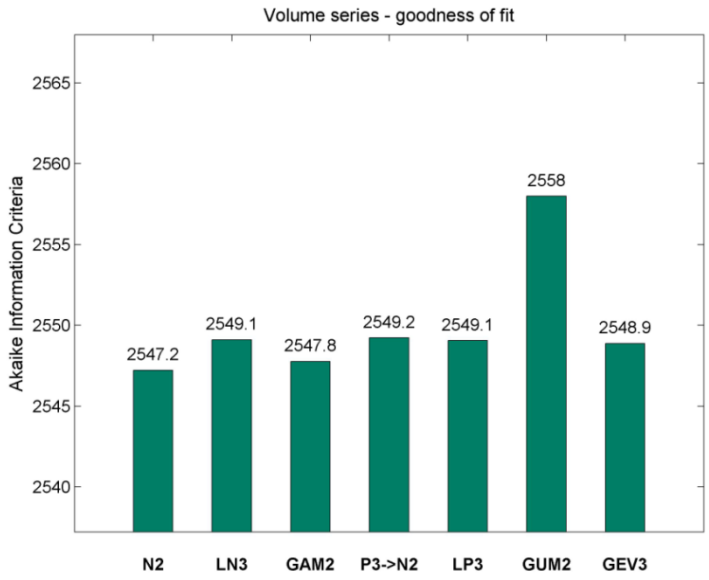


Figure 6.3-4: Akaike Information Criteria (AIC) based on the maximum likelihood of the fitting volume series to 7 distributions

6.3.2 Non-stationary frequency analysis

A data series is non-stationary if some of the underlying properties change over time. A series with trend is one example of non-stationary data. Non-stationarity also arises if there is a sudden jump or step change in the data. A data series is said to show a trend if, on average, the series is progressively increasing or decreasing (Robson and Reed, 1999; Kundzewicz and Robson, 2004).

Recently, there have been an increasing number of studies on non-stationarity in flood frequency. Delgado et al, (2010) analyzed the annual maximum discharge of the Mekong and detected non-stationarity and trends, and advocated the use of non-stationary methods in flood frequency analysis for the Mekong. Therefore, it seems appropriate to use non-stationary approaches and develop a methodology for the combined frequency analysis of flood peak and volume. This will be explored in a first step for a univariate analysis analogously to the previous chapter.

6.3.2.1 Mann-Kendall test

There exist several methods which can be used for detecting trends. This study uses a simple but robust non-parametric test, the Mann-Kendall test (Mann, 1945), to test for trends in the peak and volume series. The null hypothesis is that there is no trend in the peak, resp. volume series at the significance level of 10%. The Mann-Kendall test shows that trends are present in both peak

and volume series. It means that it is reasonable to apply the non-stationary flood frequency analysis for both series.

6.3.2.2 Non-stationary models

For the application of non-stationary frequency analysis it is assumed that the trend is only caused by the time covariate. As outlined in chapter 6.3.1, LN3 and GEV3 will be analyzed for the non-stationary case of the peak series, and N2 and LN3 for the volume series. These distributions are briefly described in the Appendix A.

Based on the trend analysis and the selection of the original (parent) candidate distribution, the following non-stationary models for flood peak were proposed. Candidate "parent" models and their parameters for the peak series are:

- Log-normal distribution (LN3): k (shape), α (scale), ξ (location)
- Generalized extreme value (GEV3): k (shape), σ (scale), μ (location)

Only trend in location and trend in scale (variance) are examined in the non-stationary analysis for this series. Furthermore, trend in location is assumed to be linear. Trend in scale follows an exponential behavior to ensure that the scale parameter is always positive. Table 6.3-3 illustrates the "child" models constructed for examining the non-stationarity in the peak flow series.

Table 6.3-3: Non-stationary models for peak series

Parameters	**Log-normal based**		**GEV based**	
	q_M10	q_M11	q_M20	q_M21
Shape	k	k	k	k
Scale	$\alpha > 0$	$\alpha(t) = \exp(\alpha_0 + \alpha_1 t)$	$\sigma > 0$	$\sigma(t) = \exp(\sigma_0 + \sigma_1 t)$
Location	ξ	$\xi(t) = \xi_0 + \xi_1 t$	μ	$\mu = \mu_0 + \mu_1 t$

The q_M10 and q_M20 are stationary models identical to the distributions used in 6.3.1. They are listed here as reference for the evaluation of the non-stationary models q_M11 and q_M21, respectively.

Candidate "parent" models for the volume series are:

- Normal distribution (N2): σ (scale), μ (location)
- Log-normal distribution (LN3): k(shape), α (scale), ξ (location)

Table 6.3-4 illustrates the "child" models (but more general models) constructed for examining the non-stationarity in the volume data.

Table 6.3-4: Non-stationary models for volume series

Parameters	**Log-normal based**		**Normal based**	
	v_M10	v_M11	v_M20	v_M21
Shape	k	k	$-$	$-$
Scale	$\alpha > 0$	$\alpha(\mathrm{t}) = \exp(\alpha_0 + \alpha_1 t)$	$\sigma > 0$	$\sigma(\mathrm{t}) = \exp(\sigma_0 + \sigma_1 t)$
Location	ξ	$\xi(t) = \xi_0 + \xi_1 t$	μ	$\mu = \mu_0 + \mu_1 t$

The v_M10 and v_M20 are again stationary models identical to the distributions in chapter 6.3.1 and are used for reference in this section. It is necessary to note here that the time varying scale function should be selected with care, especially for the case of the normal distribution, because the scale parameter has a very large influence on the flood probabilities and the exponential function is not upper-bounded. In some cases, the logistic function may be more appropriate. In this study, to be consistent in the choice of the function to represent the time-varying scale, exponential function is used in all cases.

6.3.2.3 Parameter estimation

The parameters of the non-stationary models were estimated by the maximum likelihood method (Coles, 2001), with some modifications from its original form (Fisher, 1932). Time was used in the log-likelihood form as a covariate. The parameter sets estimated for the stationary cases were utilized as starting points. The optimization algorithm is the Shuffled Complex Evolution (SCE-UE) (Duan et al., 1993). For q_M11, q_21, v_M11 there are 5 parameters to be estimated for each model. For v_M21 there are 4 parameters to be estimated.

A criterion for choosing the model:

Coles (2001) described the deviance statistic for determining the suitable models when covariates are taken into account in the non-stationary case. Given M_0, M_1 are the two models considered and M_0 is a sub-model of M_1, the deviance statistic is defined as:

$$D = 2 \times \{L_1(M_1) - L_0(M_0)\} \quad (6.2)$$

Where $L_0(M_0)$, $L_1(M_1)$ are the maximized log-likelihoods under models M_0, M_1 respectively. Large values of D imply that model M_1 explains substantially more of the variation in the data than M_0. Small values of D suggest that the increase in model size does not bring worth-while improvements in the model's capacity to explain the data. A Chi-square based test is used to help

defining how large D should be before the model M_1 is preferred to model M_0 (based on the asymptotic distribution of the deviance function).

Fitting results

Table 6.3-5: Summary of the fitted parameters for the proposed stationary (S) and non-stationary (NS) models

		Location parameter			Scale parameter			Shape parameter	
		S	NS		S	NS		S	NS
		$\xi(\mu)$	$\xi_0(\mu_0)$	$\xi_1(\mu_1)$	α	α_0	α_1	k	k_0
Peak	GEV3	45675	50565	-113	7118	8.571	0.0049	-0.1950	-0.155
	LN3	48293	52290	-93.1	7291	8.667	0.0029	-0.0646	-0.1014
Volume	N	4177927	4210009	-199	638713	13.08	0.0059	-	-
	LN3	4167936	4449238	-6528	638323	12.92	0.0081	-0.0308	-0.06

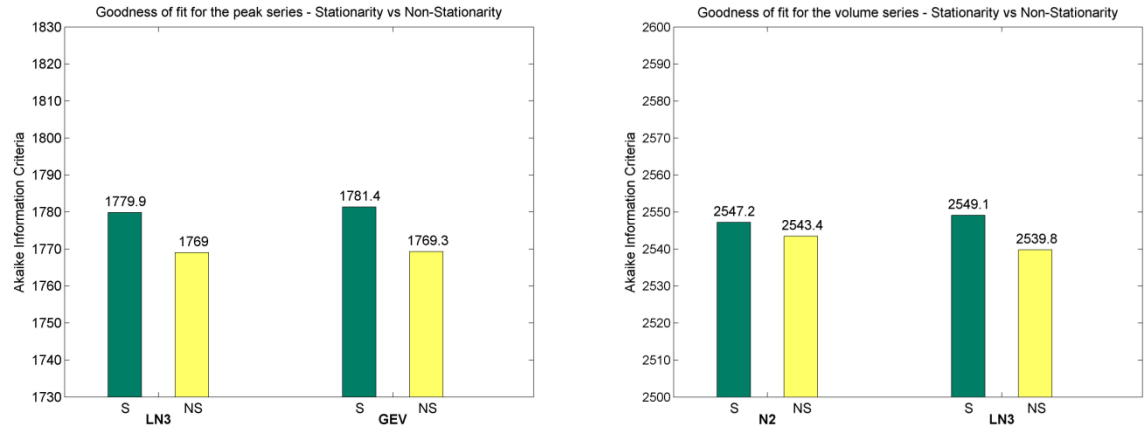


Figure 6.3-5: Akaike Information Criteria (AIC) based on the maximum likelihood of the fitting the non-stationary series to the proposed models: (left) - peak series, (right) - volume series; green denotes the stationary case, yellow is for the non-stationary case.

Peak series

Figure 6.3-5shows that both distributions are equally well suited for describing the time-varying peak series. The deviance statistic values are 10.850 and 12.076 for LN3 and GEV, respectively. In both cases, the non-stationary models are superior to the stationary models at the significance level of 10%. The non-stationary GEV3 model can describe well the time-trend series. It is an upper-bound distribution in this case. As mentioned in Section 6.2, the maximum unit discharge at Kratie was reported as one of the biggest in the world (O'Connor and Costa, 2004; MRC,

2005a). However, it is difficult to prove the physical limit of the runoff here. Therefore, in flood frequency analysis, upper-bounded distributions should be handled with care. The LN3 is hence chosen.

Volume series

Figure 6.3-5 shows that the LN3 distribution describes better the time-varying behavior of the volume series than the normal distribution. The deviance statistic is 3.794 and 9.352 for N2 and LN3, respectively. Applying the Chi square test leads to the conclusion that there is no improvement if applying v_M21 but there is a significant improvement in fitting the volume data to v_M22. Hence, LN3 is selected to model the volume series.

Figure 6.3-6 illustrates the time-varying probability density function which has been fitted to the peak and volume series using the non-stationary Log-Normal 3 parameter distribution.

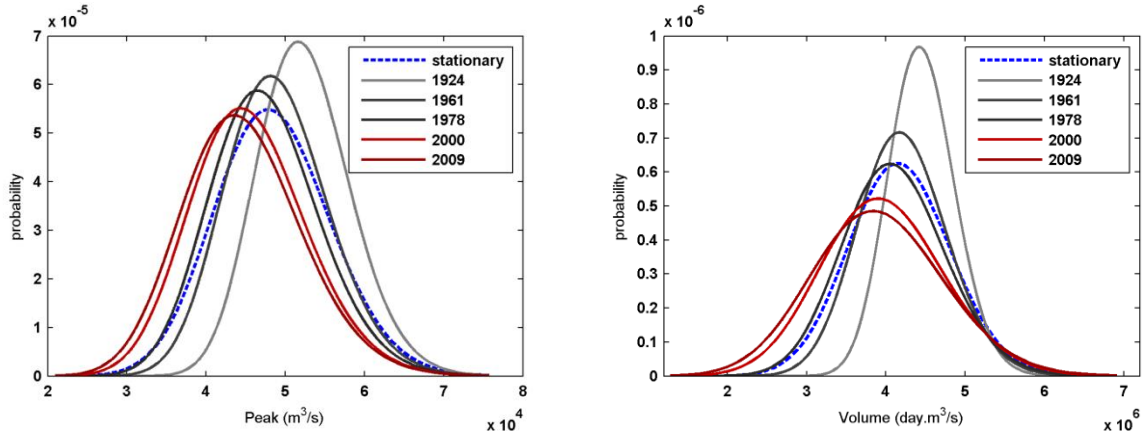


Figure 6.3-6: estimated PDF of peak (left) and volume (right) for different year according to the non-stationary LN3

6.3.2.4 Standardization of the non-stationary series

In the non-stationary case, the homogeneity assumption in the distribution of the observations does not hold. Therefore, the model-testing procedure has to be modified. The idea is that the "time-dependent" series is transformed to a "time-independent" series. Coles (2001) suggested a way to standardize the original data which can be applied to homogenous or non-homogenous series. Log-Normal distribution was suited to fit to the series in both stationary and non-stationary analysis. Thus the standardization procedure for this distribution is introduced below.

When random variables X_t follow a LN3 distribution with three parameters $k(t)$ (shape), α(t) (scale), $\xi(t)$ (location) (see Appendix A), we say:

$$X_t \sim LN3(k(t), \alpha(t), \xi(t)) \tag{6.3}$$

Standardized variables Z_t are then defined by:

$$Z_t = \begin{cases} -k(t)^{-1}\log\left\{1 - \dfrac{k(t)(X_t - \xi(t))}{\alpha(t)}\right\}, & k(t) \neq 0 \\ \dfrac{(X_t - \xi(t))}{\alpha(t)}, & k(t) = 0 \end{cases} \tag{6.4}$$

The standardized variables Z_t, hence, follow standard Normal distribution. Therefore, the probability and quantile plots of the observed z_t can be made. The above standardization form is also applied for the stationary case. Figure 6.3-7 shows the probability plots and quantile plots in both cases – stationary and non-stationary. In general, it illustrates good fits of the models to the series, although the fit in the tails of the peak series is not at high satisfaction. Furthermore, Figure 6.3-8 shows that the standardization does not change the rank of the data, while it changes the rank of the data in the non-stationary case. This implies a change in the dependence structure of the data (rank correlation) leading to a possible change in the bivariate model which will be applied later.

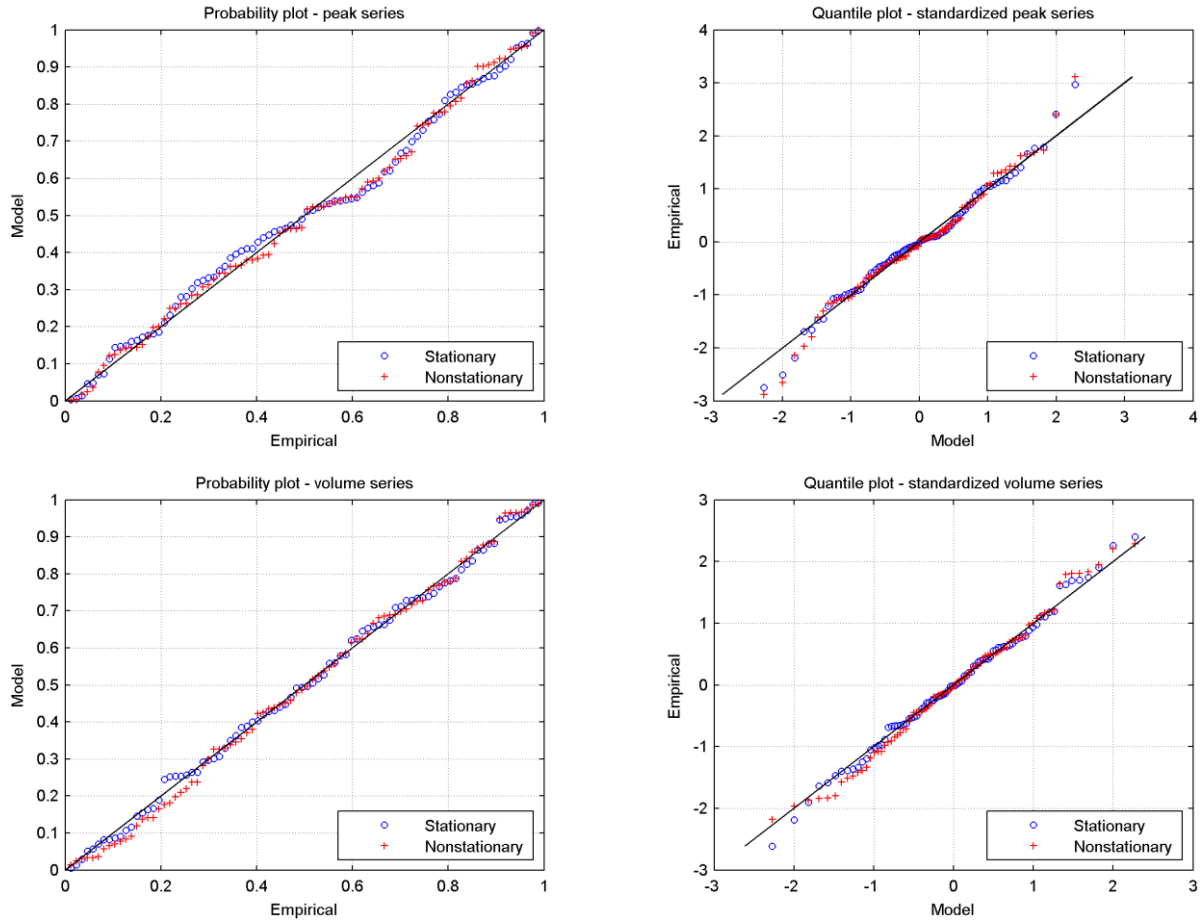


Figure 6.3-7: Residual diagnostic plots: probability plots (first column) and quantile plots (second columns) for the peak series (first line) and the volume series (second line) for stationary and non-stationary analysis

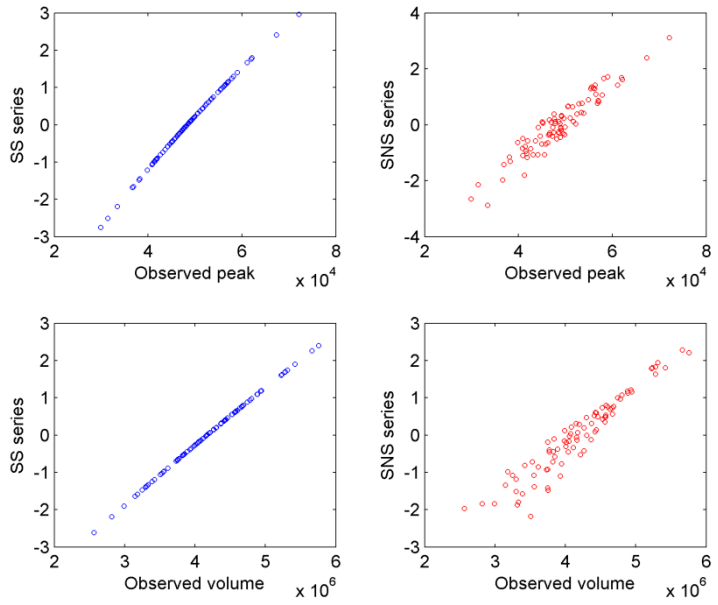


Figure 6.3-8: Standardized peak (first line) and volume series (second line) for the stationary case (left column) and the non-stationary case (right column); the label in vertical axis indicates the standardized value for the stationary case (first column) and for the non-stationary case (second column)

6.3.2.5 Time-trend in peak and volume series

Figure 6.3-9 illustrates the trend found in both location and scale parameter. Both series show a negative trend in the location parameter and a positive trend in the scale parameter. These trends will be used for extrapolating the flood hazard to "a near" future (see Section 6.6).

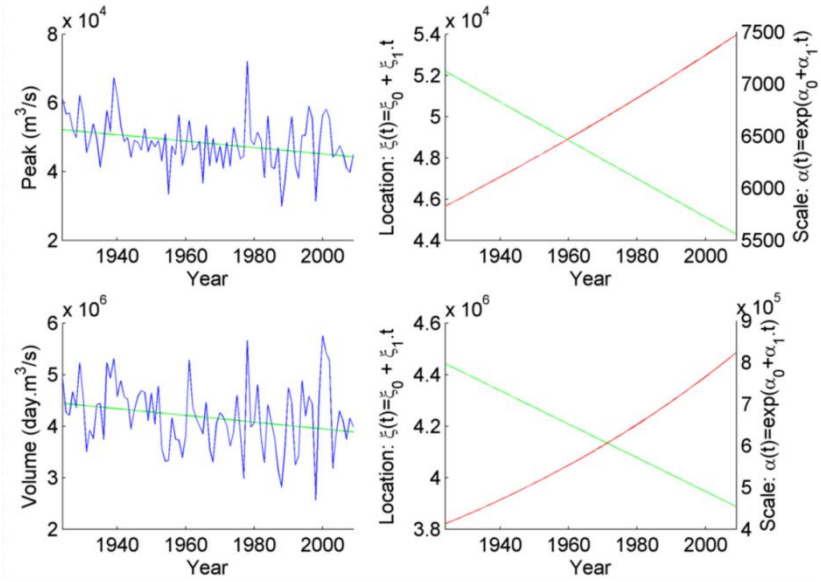


Figure 6.3-9: Time-varying parameters of non-stationary models for flood peak (first line) and volume (second line); blue for the time series, green: location parameter; red: scale parameter

6.3.3 Return period via univariate frequency analysis

For a random variable, probability of exceedance is defined for any given value. The return period (also known as recurrence interval) of a given value is the inverse of that probability. That mapping is a one-to-one relationship. The return period $T_X(x)$ is defined as follows (in this case, the unit of the return period is year):

$$T_X(x) = \frac{1}{Pr(X > x)} = \frac{1}{1 - Pr(X \leq x)} = \frac{1}{1 - F_X(x)} \tag{6.5}$$

Where $Pr(-)$ denotes the probability of an outcome, $F_X(x)$ the cumulative distribution functionof the random variable X. The return period is dependent on the type of the cumulative distribution function considered. The Log-normal distribution with 3 parameters LN3 is used for fitting both observed time series. Figure 6.3-10illustrates the estimation of return periods for flood peak and volume in both stationary and non-stationary cases, along with examples of changes in return period when the non-stationary model is applied. It can be seen that the changes can be significant, which in turn has consequences for the eventual use of the return periods in flood risk management or mitigation planning.

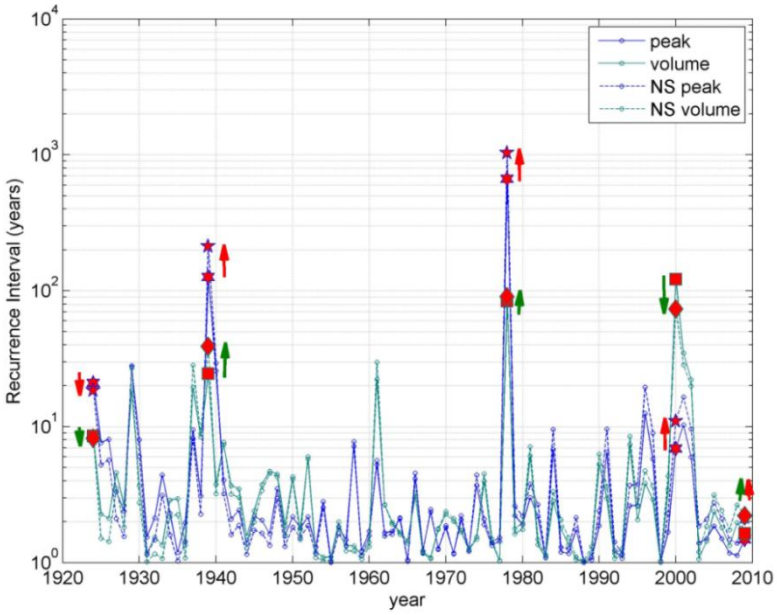


Figure 6.3-10: The return period estimations for the observed data using the LN3-based stationary and non-stationary distribution functions: five points are selected to represent the different pattern of the change in the recurrence intervals. Arrow denotes the comparison at the points taken into account (red for peak, green for volume). The direction of the arrows illustrates increase or decrease of the recurrence interval for the non-stationary case compared to the stationary approach.

6.4 Bivariate flood frequency analysis via copulas

In the previous chapters the flood peak and volume were analyzed separately. However, peak and volume correlate as shown in chapter 6.2. Therefore the frequency analysis should take account of this correlation by expanding it to a multivariate, resp. bivariate analysis. Classical bivariate approaches are only applicable when the two margins of the two random variables (or their transformations) follow the same family of distribution (Adamson et al., 1999;Yue, 1999, 2001; Yue et al., 1999). Extending the uncorrelated approach and the LN3 distribution, a classical bivariate log-normal distribution could be theoretically applied (Hiemstra et al., 1976; Yue, 2000). However, a more flexible approach based on copulas is adopted for this study. In case different distribution functions have to be selected for representing the bivariate data (e.g. Generalized extreme value distribution for peak series and Gamma distribution for volume series), the applicability of the traditional approach is no longer given (Salvadori and De Michele, 2007; Durante and Salvadori, 2010).

The copula concept emerged a long time ago when Sklar's theorem was first stated in 1959 (Joe, 1997; Nelsen, 2005). Recently, copula-based techniques are frequently applied to a wide range of real-world applications (Genest and Favre, 2007), because copula-based models are able to lift the limitations found in classical multivariate analysis. While copulas have been applied very often in "high-risk" financial and accrual sectors, its use in water resources is quite recent. Singh et al. (2005) applied the Archimedean copula family to analyze multivariate rainfall data. Salvadori and De Michele (2004) employed copulas to examine flood peak and volume when studying dam protection. Klein et al. (2010) applied the concept for flood design. Bárdossy, (2006) proposed copula-based geo-statistical models for groundwater quality parameters.

Copula is not a magic tool which can be applied to all applications where a multivariate analysis is needed. There is no consensus on copulas being the optimal solution to model dependencies, as illustrated by the recent and heated discussion initiated by (Mikosch, 2006) and followed by numerous authors (see Genest and Rémillard, 2006; Joe, 2006)

6.4.1 Basic concept of copulas

A very short introduction to the concepts of copulas is given. Details can be found in several text books (e.g. Joe, 1997; Cherubini et al., 2004; Nelsen, 2005; Salvadori et al., 2005).

Any multivariate distribution function with support on the n-dimensional unit hypercube and uniform margins is called a copula. Copulas are bounded by the Frechet-Hoeffding bounds. Given marginals, it is always possible to earn an infinite number of multivariate joint distributions. The main advantage of copulas is their 'scale invariant' property, meaning that they remain the same by any scaling transformation. The nature of the bivariate copula is the

ability to seize the dependence of the two random variables without being necessarily influenced by the marginals.

In the case of flood peak and volume at Kratie, the Gaussian copula is selected because both marginals were fitted well to the log-normal distribution which in turn originates from the normal distribution family. Additionally, the one-parameter Archimedean copula families are applied. Archimedean copulas are among the most frequently used, especially in hydrology (Genest and Favre, 2007). From the Archimedean copula families, the Ali-Mikhail-Haq copula is not considered because its application is based on the cases where the rank dependence is rather small (<0.4). Hence, the four candidate bivariate copulas Gumbel-Houggard, Frank, Clayton and Gaussian described in Table 6.4-1will be further investigated.

Table 6.4-1: Summary of the four candidate bivariate copulas (Nelsen, 2005;Bárdossy, 2006;Renard and Lang, 2007)

Name	**Descriptions**	
Gumbel-Hougaard	$C_\theta(u,v) = \exp\left(-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{1/\theta}\right), \theta \in [1,+\infty)$	(6.6)
Frank	$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + (e^{-\theta u} - 1)(e^{-\theta v} - 1)(e^{-\theta} - 1)\right),\ \theta \in R\backslash\{0\}$	(6.7)
Clayton	$C_\theta(u,v) = \left(u^\theta + u^\theta - 1\right)^{-1/\theta}, \theta \in (0,+\infty)$	(6.8)
Gaussian	$C_\theta(u,v) = \Phi_\theta(\phi^{-1}(u) + \phi^{-1}(v)), \theta \in [-1,1]$ ϕ^{-1} : the inverse function of the standard normal distribution cdf Φ_θ : the bivariate normal distribution cdf with correlation θ	(6.9)

When applying copulas for extreme value statistics, tail dependence is of high importance. Tail dependence is defined as the "amount of dependence" in the upper right quadrant tail (or lower left quadrant) of a multivariate distribution (Joe, 1997; Salvadori et al., 2005). A copula can have upper tail dependence or lower tail dependence or both. Among the candidate copulas, Gaussian is a symmetric copula but it has no tail dependency (in both upper and lower parts). The Frank copula does not have tail dependency either. Gumbel-Houggard has upper tail dependency and the Clayton copula has lower tail dependency.

6.4.2 Parameter estimation methods

When a parametric copula is selected to model the dependence between two random variables, parameters related to that copula family must be estimated. To this end, different methods exist (Cherubini et al., 2004; Joe, 2005; Genest and Favre, 2007):

- Direct method (Genest and Favre, 2007): copula parameters are estimated via one of the two common used rank correlation measures Kendall's tau (Kendall, 1970) and Spearman's rho.
- Pseudo maximum likelihood method (PML) (Genest et al., 1995; Genest and Favre, 2007): the sample data are transformed to uniform variates (empirical margins) and then coupled into the likelihood (or log-likelihood) function of the candidate copula; the copula parameters are estimated by maximizing the likelihood function. This method is also called the canonical maximum likelihood (CML) because it can be derived from the canonical representation form introduced in Cherubini et al., (2004).
- Inference from margin method (IFM) (Cherubini et al., 2004; Joe, 2005): instead of using the empirical margins in the likelihood function in the CML method, the theoretical margins are used. Parameters of the theoretical margins are estimated in advance, and then the copula parameters are estimated. This is why this method is called the two stage estimation method.
- Exact maximum likelihood (EML) (Cherubini et al., 2004): instead of using the two stage estimation of the IFM method, parameters of the theoretical margins and the copula are estimated at the same time.

Some of these methods are shortly presented in the Appendix B.

6.4.3 Results of parameter estimation and copula identification

For the stationary case, the first three estimation methods mentioned are used. For the non-stationary case, only the third estimation method was applied. Analogous to the univariate fitting, the AIC static value is used for evaluating the fitting performance.

Stationary analysis

Figure 6.4-1 shows the AIC values for the candidate copulas. The Gaussian copula outperforms the other candidates. Figure 6.4-2 provides a graphical test of the goodness-of-fit. A random sample of size 20,000 (small dots) from each of the four copulas with parameters estimated by the IFM method was simulated. The 86 pairs of peak and volume observations are represented by black circles. The two blue lines represent the threshold value corresponding to the 100-year

return period of the univariate analysis of peak and volume. The plot allows judging the viability of the complete model for frequency analysis. The two copulas Gumbel and Frank fit well to the observed data. However, the Gaussian copula seems to be the best (within the range of the observed data), while the Clayton copula is not a suitable choice because it spreads out significantly in the upper part. Furthermore, most of the points are not in the upper part formed by the two blue lines. And we do not want to go very far from what we observed. Or it means the "very extreme" values are not considered. Together with the fact that Gaussian copula has the best AIC value, it is selected.

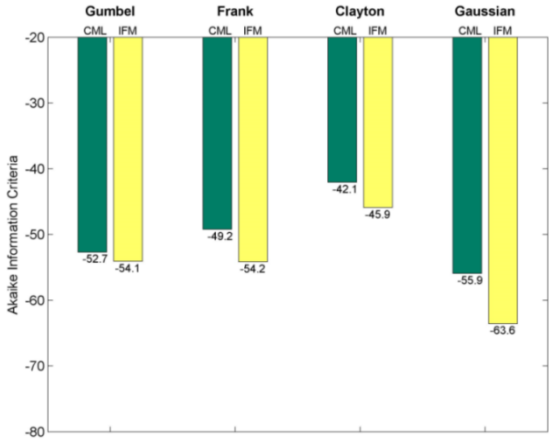


Figure 6.4-1: The AIC values of the estimation for the stationary analysis

Table 6.4-2: Copula parameter estimation for the stationary case

Copula	Method	Parameter	ML	AIC
Gumbel-Hougard	Direct	1.961		
	CML	1.917	27.4	-52.7
	IFM	1.872	28.1	-54.1
Frank	Direct	5.558		
	CML	5.578	25.6	-49.2
	IFM	6.007	28.1	-54.2
Clayton	Direct	1.922		
	CML	1.349	22.0	-42.1
	IFM	1.338	24.0	-45.9
Gaussian	Direct	0.730		
	CML	0.721	29.0	-55.9
	IFM	0.730	32.8	-63.6

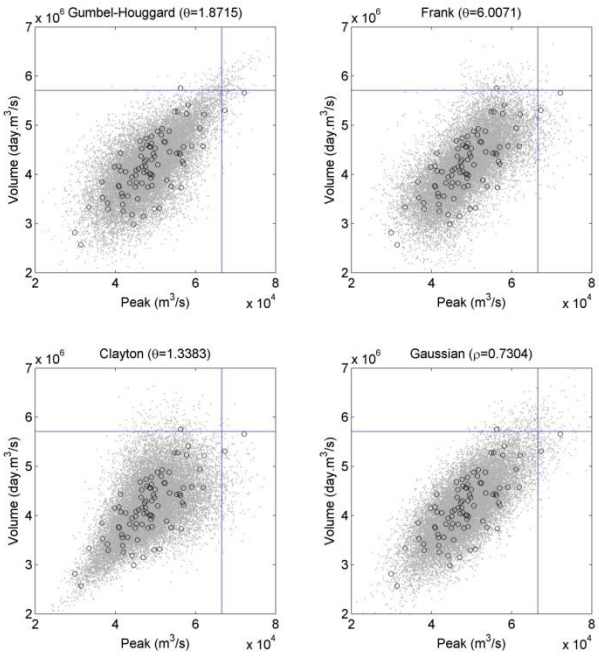


Figure 6.4-2: A graphical test for goodness-of-fit in the stationary analysis.

Non-stationary analysis

For the non-stationary analysis, the estimated parameters and AIC values based on the IFM method are presented in Table 6.4-3 and Figure 6.4-3. Again, the Gaussian copula shows the best fit to the data. The Clayton and the Frank copula are eliminated because of their significantly higher AIC.

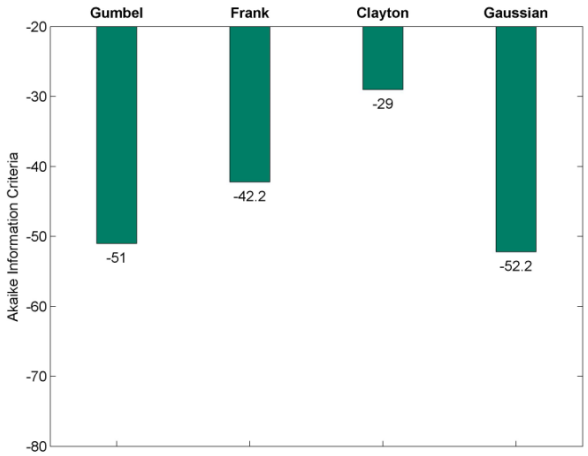


Figure 6.4-3: The AIC values for the candidate copulas using the inference from margins (IFM) method for the non-stationary case

Table 6.4-3: Copula parameter estimation for the non-stationary case

Copula	Method	Parameter	ML	AIC
Gumbel-Hougard	IFM	1.765	26.5	-51.0
Frank	IFM	3.960	22.1	-42.2
Clayton	IFM	0.890	15.5	-29.0
Gaussian	IFM	0.665	27.1	-52.2

For the graphical check of the suitability a small modification was made. In the non-stationary analysis, data is time-dependent. Therefore, standardization by "removing time" from the data is needed. The procedure for the standardization of the series was described in Subsection 6.3.2. We assume that the copula for the original series and the copula for the transformed series remain the same. Based on this assumption, the graphical check is given in Figure 6.4-4. The same argument as in the stationary analysis holds. Again, we decided not to go very far from what we observed. Hence, the Gaussian copula is the best choice for the non-stationary case. If we had indication of tail dependency, the Gumbel copula would be a more appropriate choice for flood frequency analysis, because it can represent upper tail dependency and the degree of dependence depends on the parameter value of the copula under consideration.

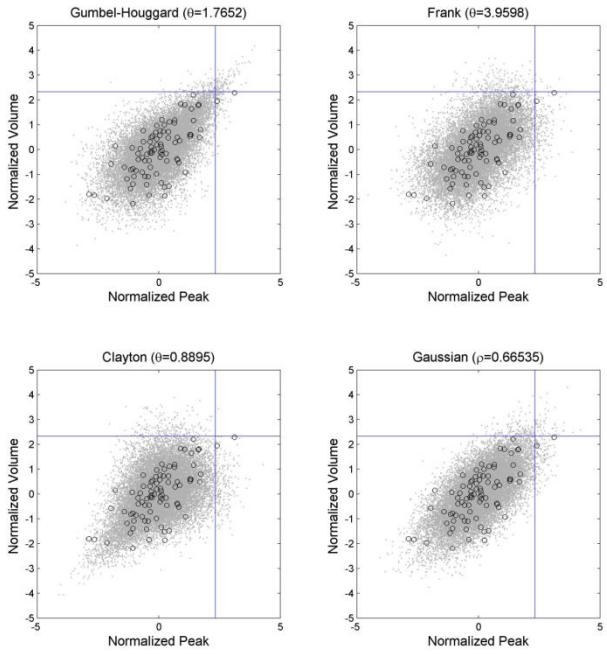


Figure 6.4-4: A graphical test for goodness-of-fit in the non-stationary analysis.

From the bivariate analysis via copulas implemented above for both cases: stationary and non-stationary, the Gaussian copula is thus selected for bivariate flood frequency analysis.

6.4.4 Return period via copula

Following (Salvadori et al., 2005), in the bivariate frequency analysis the return period can be defined by the average lag years between two flood events of which:

- either volume or peak exceeds certain values. ("OR" case return period)
- both volume and peak exceed certain values. ("AND" case return period)

Given a pair (x, y) of a random bivariate, the "OR" case return period is defined as:

$$T_{XY}(x,y) = \frac{1}{Pr(X \geq x \; OR \; Y \geq y)} \tag{6.10}$$

And the "AND" case return period is defined as:

$$T^{XY}(x,y) = \frac{1}{Pr(X \geq x \; AND \; Y \geq y)} \tag{6.11}$$

After some simple transformation steps, the above formulas read:

"OR" case:

$$T_{XY}(x,y) = \frac{1}{1 - F_{XY}(x,y)} = \frac{1}{1 - C(1 - \frac{1}{T_X(x)}, 1 - \frac{1}{T_Y(y)})} \tag{6.12}$$

"AND" case:

$$T^{XY}(x,y) = \frac{1}{\frac{1}{T_X(x)} + \frac{1}{T_Y(y)} - \frac{1}{T_{XY}(x,y)}} \tag{6.13}$$

From Equations 6.12 and 6.13, it is obvious to see that if the return periods of the two marginal distributions T_X, T_Y are known and the copula function is clearly defined, the corresponding return period of the bivariate can be derived. However, when the return period of the bivariate and the copula function is known, infinite pairs of (T_X, T_Y) can be derived from the distribution functions for the given return period. Furthermore,

$$T_{XY}(x,y) \leq T_X(x), T_Y(y) \leq T^{XY}(x,y) \tag{6.14}$$

holds true. The unit of T_{XY} and T^{XY} is "year" in this study (because annual series were considered in frequency analysis).

6.5 Hydrograph shape analysis

As described above, a period from the beginning of June to the end of November is fixed to define flood events. The flood frequency analysis examined flood peak and flood volume. To generate a discharge hydrograph, the shapes of the flood hydrographs are analyzed to derive characteristic flood hydrographs. The applied method follows the procedure of Apel et al. (2004, 2006). First, the series of hydrographs are normalized to the peak flow, i.e. dividing the series by the peak value corresponding to each event. We denote annual hydrograph (6 months) series as $H_t, t = 1..86.$ (86 means the number of years)

$$NH_t.q[i] = H_t.q[i]/H_t.q^{max} \tag{6.15}$$

Where NH_t is the normalized hydrographs, $q[i]$ denote the discharge value corresponding to day $i = 1..183$. The notation "." should be read as "of which" (not multiplication). Then, we derive:

$$\begin{aligned} &NH_t.q_{max} = 1 \\ &NH_t.v = H_t.v/H_t.q^{max} \end{aligned} \tag{6.16}$$

Next, a cluster analysis is implemented to classify the normalized hydrographs into groups based on their "similarity".

Cluster analysis

Cluster analysis using a simple hierarchical approach, the "ward" method, was applied to the normalized hydrograph series. The similarity between hydrographs is determined by the squared Euclidean distance. Given the number of clusters, the selected number of clusters are formed satisfying the criteria in which the total variance in each cluster is minimized.

For the hydrographs at Kratie, four clusters are proposed. Actually, this choice is – to a certain extent - subjective. In general, the number of hydrograph clusters should be large enough to depict the variability of the hydrograph shape but small enough in order to identify characteristic hydrographs and to overcome the problem of computational demand when dealing with a Monte Carlo (or Latin Hypercube) simulation.

The dendrogram shown in Figure 6.5-1 describes the process of the cluster analysis of the normalized hydrograph shape series. The values in x-axis show the identity of each hydrograph, the values in y-axis show the "distance" assigned to all intermediate groups. The shaded yellow region demonstrates the threshold for the four cluster selection. To each of the four groups an occurrence probability is assigned, given as the number of elements in each group divided by the total number of hydrographs (in this case 86). Figure 6.5-1 shows the cumulative probability assigned to the clusters.

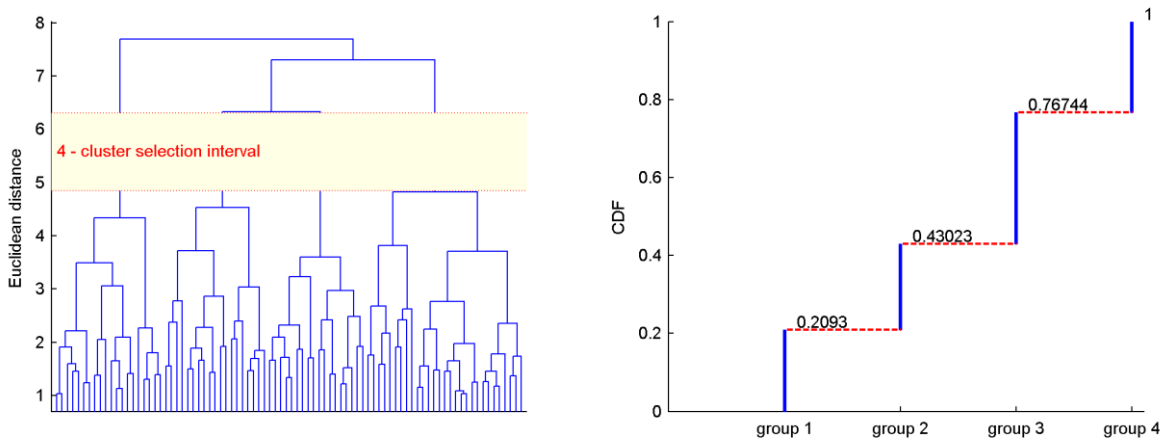


Figure 6.5-1: (left) - Dendrogram illustrating the cluster analysis for grouping hydrographs using the "ward" algorithm; (right) - and cumulative frequency of the 4 clusters

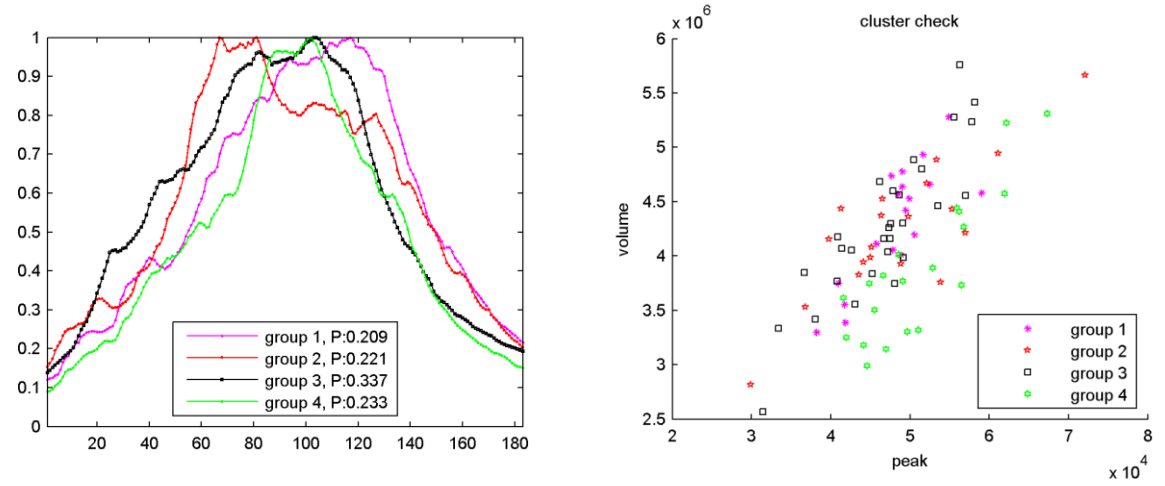


Figure 6.5-2: (left) – characteristic hydrographs; (right) –distribution of all clustered hydrographs

Figure 6.5-2 (left) shows 4 non-dimensional characteristic hydrographs computed as the mean hydrograph from the hydrographs in the clusters along with their probabilities. The shape of the hydrographs in group 2 and group 4 is more clearly defined compared to group 1 and 3. In group 2, the largest peak appears early in the flood season (e.g. the record flood year 1978), and its rising slope is much higher than the recession slope (many small peaks may occur). The hydrograph in group 4 is characterized by its small "head" with very steep slope in both rising and falling stages (e.g. the flood year 1984). The most frequent hydrograph shape is group 3. In this group, the shape shows 2 high peaks occurring closely in time; the lower peak appears first and higher peak appears with a delay of about 30 days (e.g. the flood year 2000, 2001 and 2002). Group 1 has the lowest probability of occurrence; it looks similar to the shape in group 3 but the largest peak appears later and there is no clear smaller peak (e.g. the flood year 1961).

The flood events given in brackets are big floods in the Mekong Delta. It means that the cluster analysis is valid in the sense that there is no big difference between the shape of the high peak and low peak events. This is checked in the right plot in Figure 6.5-2. The plot indicates that there is no relation between the hydrograph shape and the severity of the flood event, although group 4 tends to cover smaller volume events (green points). Therefore, four characteristic hydrographs could be used to build up the shape of hydrograph.

6.6 Synthetic hydrograph generations

Synthetic hydrographs are generated using the identified characteristic hydrographs scaled by peak discharge and flood volume estimated from both the stationary and non-stationary bivariate distributions. We do not move "very far" from what we observed, meaning that the flood scenarios are not derived for very extreme peak and volume pairs, resp. return periods. We therefore select the 100-year return period as a reference for building the scenarios. Furthermore, the "AND" case return period is chosen (see Section 6.4). From the Equation 6.14, both T_Q and T_V range in [1 100]. It means that some "unrealistic" pairs (q_{max}, v) could be possibly generated from the chosen Gaussian copula. Therefore, the above range is limited to $64.99 \leq v/q_{max} \leq 107.45$ (of the observed data) with a so-called expansion scale parameter λ representing the "prediction factor". This factor enables the selection of unobserved pairs of q_{max} and v in the hazard analysis. We choose 1.2 for λ. The lower bound is hence equal to $64.99\lambda - (107.45 + 64.99)(\lambda - 1)/2 = 60.74$ and the upper bound is $107.45\lambda - (107.45 + 64.99)(\lambda - 1)/2 = 111.69$. Therefore, pairs of (q_{max}, v) will be chosen satisfying the conditions:

$$\begin{aligned} &60.74 \leq v/q_{max} \leq 111.69. \\ &T^{QV}(q_{max}, v) = 100 \end{aligned} \tag{6.17}$$

Figure 6.6-1 illustrates the range for the sampling of pairs (q_{max}, v) meeting above conditions. The red and blue curve indicate the 100-year return period curve (for the stationary and non-stationary analysis, respectively) conditioned by the observed range, the yellow curve and the green curve indicate the 100-year return period curve (for the stationary and non-stationary analysis, respectively) conditioned by the observed range and expanded by the "prediction factor". Only the latter curves will be used for the sampling process. The Latin Hypercube sampling technique was used to sample 100 pairs of (q_{max}, v) for each stationary and non-stationary analysis. By this technique the 100-year event is simulated in a probabilistic framework considering the uncertainty in the prediction of the 100-year event from the sampling of distribution functions and the dependency between flood peak and flood volume. The sampled pairs of q_{max} and v were associated with a randomly selected characteristic hydrograph

conditioned by their probability of occurrences. Finally, synthetic hydrographs are generated by rescaling the normalized characteristic hydrographs to the sampled q_{max} and v.

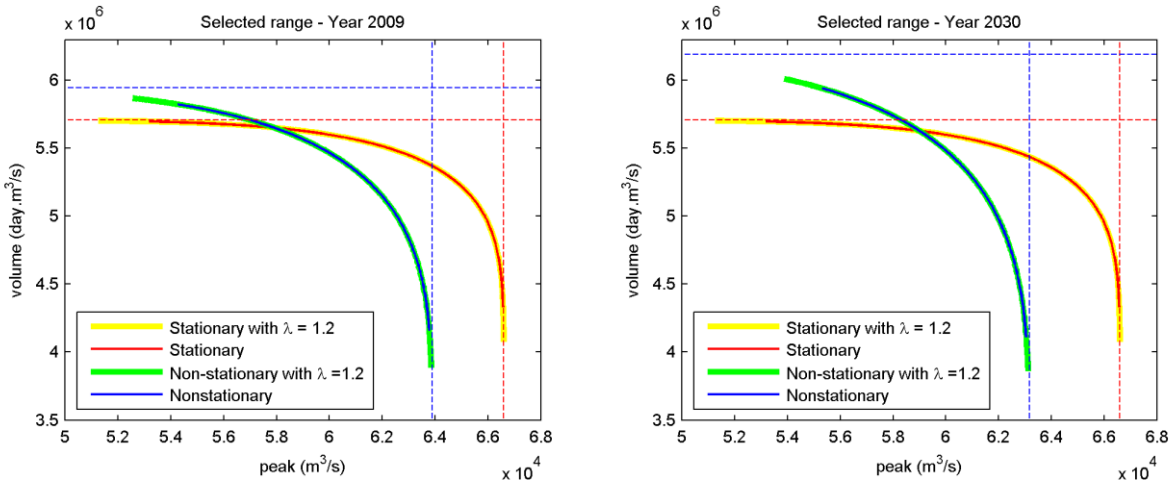


Figure 6.6-1: Range of pairs (q_{max}, v) for a 100-year flood for the years 2009 (left) and 2030 (right) for stationary and non-stationary analysis with and without expansion factor λ.

Two reference years are selected for mapping the probabilistic flood hazard. The year 2009 is used as baseline scenario or current status, 2030 is used as future scenario. In order to estimate the flood hazard in 2030, the detected trends in the time varying location and shape parameters of the non-stationary distributions are extrapolated to the reference year 2030. We assume hereby that the detected trends will continue in the near future, which is justifiable from our point of view given the observed monotonous trends. However, further extrapolation in future should be avoided due to the unquantifiable uncertainty associated. The estimations of the 100-year event derived by the stationary approach are compared to the non-stationary approach for these reference years. The synthetic hydrographs are transferred into inundation areas utilizing the calibrated and validated hydrodynamic model described in chapter 5. In order to account for model parameter uncertainty, roughness parameter sets for the best F2 and the best Euclidean distance are used for the modelling. We neglect the parameter set for the best F1 performance in F1, because in the derivation of the hazard maps the spatial performance has superior priority compared to temporal performance. Hence, in total, 600 runs (for the stationary case, the generation of peak and volume is the same for the year 2009 and year 2030) of the hydrodynamic model are launched and evaluated in the next section.

6.7 Simulations and flood hazard mapping

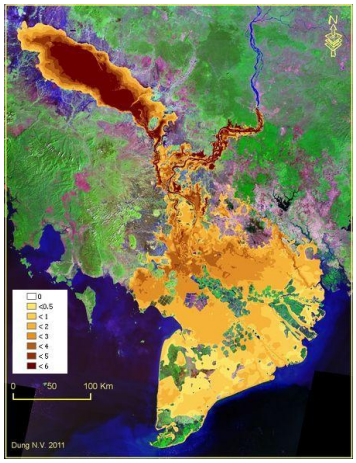
600 simulations of the hydrodynamic model were launched (taking 5 days) and results of the simulation were processed. Maximum water level at more than 26000 computational nodes was interpolated to two-dimensional representation with horizontal resolution of 90 m. Water depth

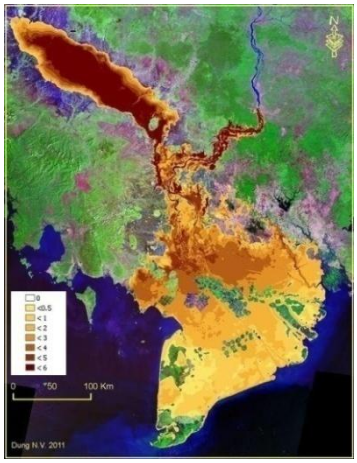
maps were then derived using the SRTM DEM for the Mekong Delta. These maps are then interpreted in a probabilistic way by calculating inundation statistics for every grid cell. For each case, the 3 commonly used percentiles 5%, 50% (mean), 95% were utilized to illustrate the uncertainty in flood hazard. Table 6.7-1 lists the codes used in the following hazard maps.

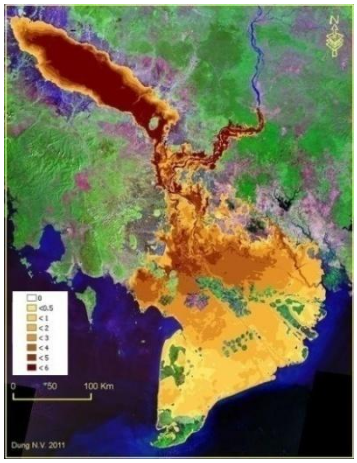
Figure 6.7-1 shows the mapping results using the best F2 parameter set. The first row of figure indicates the maximum depth over the Mekong Delta when stationarity in discharge series at Kratie is assumed. The second and the third rows illustrate the maximum depth at the same domain when non-stationarity is taken into account. The second row shows the 100 year flood hazard for the baseline year 2009, while the third row is for the future scenario year 2030. Examining the maps corresponding to the 3 percentiles in the columns, one can see that uncertainty in the upstream discharge at Kratie affects mainly the area in the Cambodian part of the Delta and in the upper part of the Delta in Vietnam. These regions are deeply inundated and the inundation depth ranges from 3 m to more than 6 m. The regions that are closed to the Vietnamese coast, especially the Eastern coast are less influenced. This could be explained that the flood flow from the upstream stretches more widely on the direction to the coast. The two first plots in Figure 6.7-3 show large differences between the 95% and 5% percentile maps (the inter-percentile map) for the stationary case and non-stationary case for the baseline year 2009. Comparison of the inundation depths derived by the stationary case and the non-stationary cases for the year 2009 and 2030 does not show significant difference (see Figure 6.7-3). This holds true for all the quantile maps. The maximum difference is shown in the Cambodian floodplains and the upstream of the delta Vietnam ranging from 20 cm to 50 cm (see the third map in Figure 6.7-3). Figure 6.7-2 shows the results for the case in which the best Euclidean parameter set is used. The interpretation of this inundation depth information is similar to that in the best F2 case described, except that the inundation depth is lower.

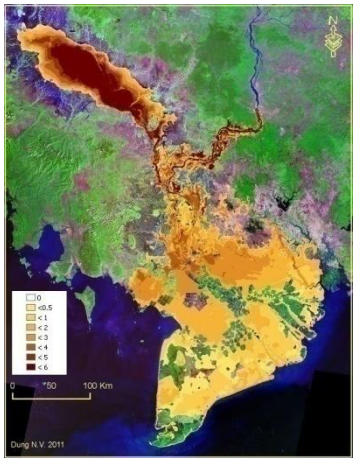
Table 6.7-1: Explanation for the text labels in Figure 6.7-1 and Figure 6.7-2

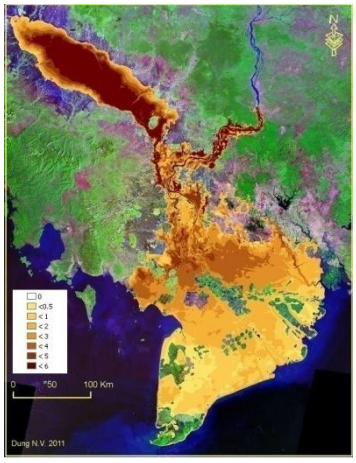
Notation	Descriptions	Notes
F2	refers to the best parameter set of roughness representing the spatial performance (best F2)	Parameter uncertainty
EU	refers to the best parameter set of roughness balancing the temporal and spatial performance (best Euclidean distance)	
S	refers to the stationary analysis	Uncertainty caused by the selection of statistical model
NS2009	refers to the non-stationary analysis for the current baseline year 2009	
NS2030	refers to the non-stationary analysis for the future year 2030	
05,50,90	refers to the percentiles (%)	Probabilistic uncertainty

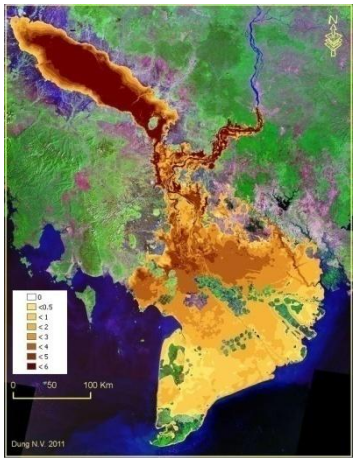

F2_S_05

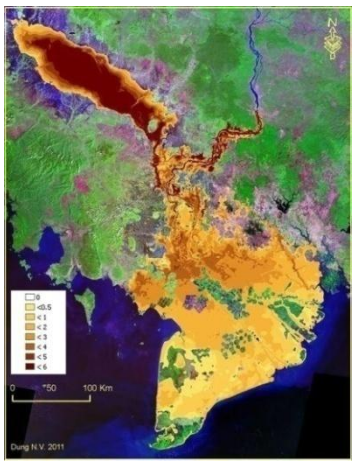

F2_S_50

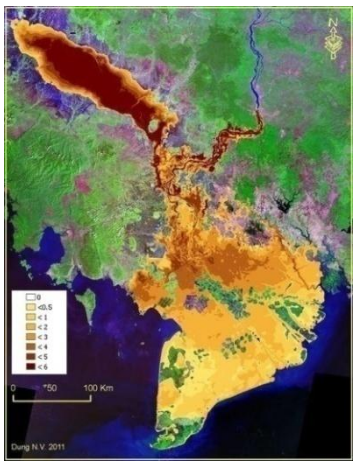

F2_S_95


F2_NS2009_05


F2_NS2009_50


F2_NS2009_95


F2_NS2030_05


F2_NS2030_50

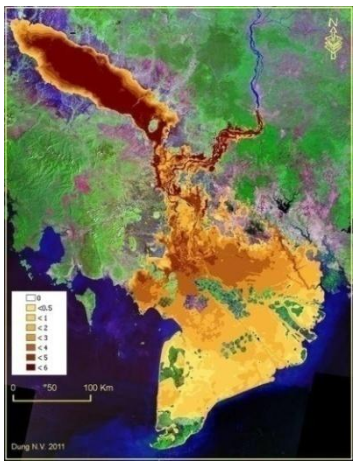
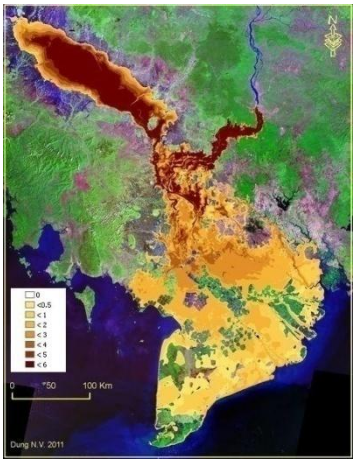
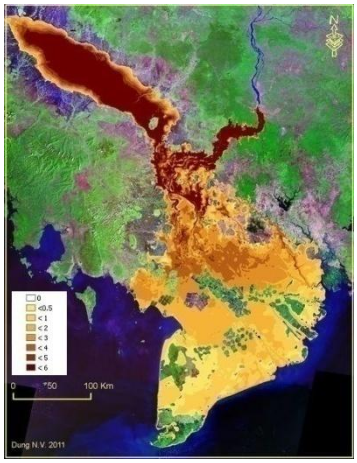

F2_NS2030_95

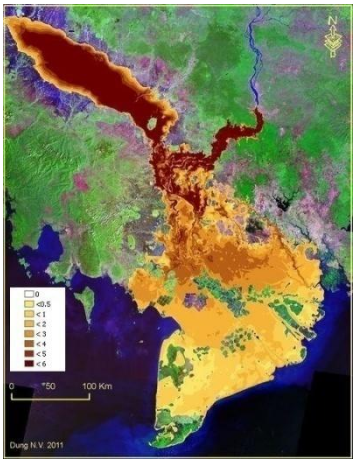
Figure 6.7-1: Flood hazard maps for maximum inundation depth over the study domain derived with the roughness parameter set best F2 for the hydraulic model (Table 6.7.1 explains the codes of the individual maps)



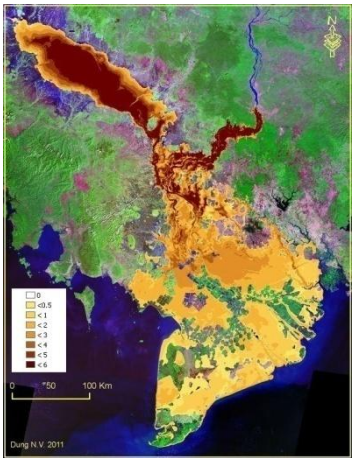
EU_S_05



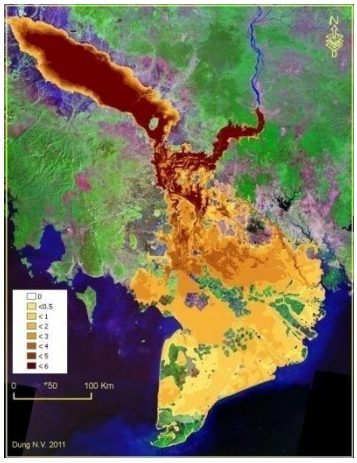
EU_S_50



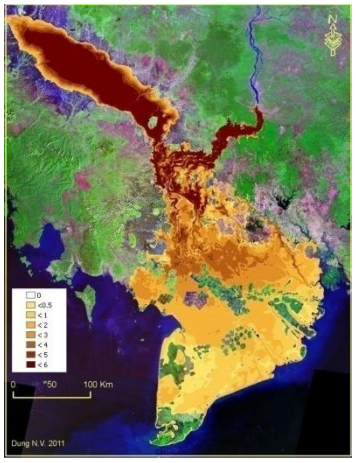
EU_S_95



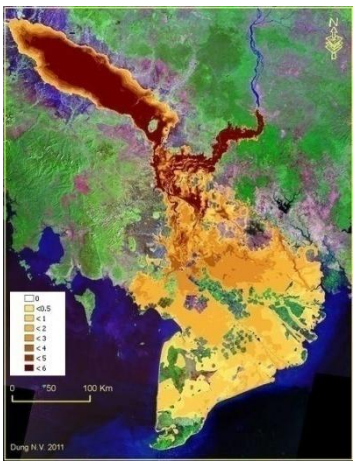
EU_NS2009_05



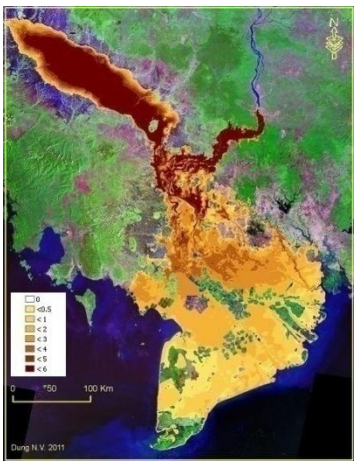
EU_NS2009_50



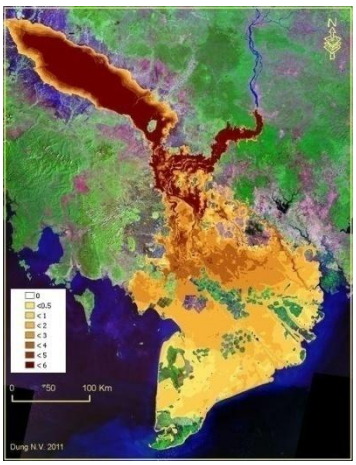
EU_NS2009_95



EU_NS2030_05

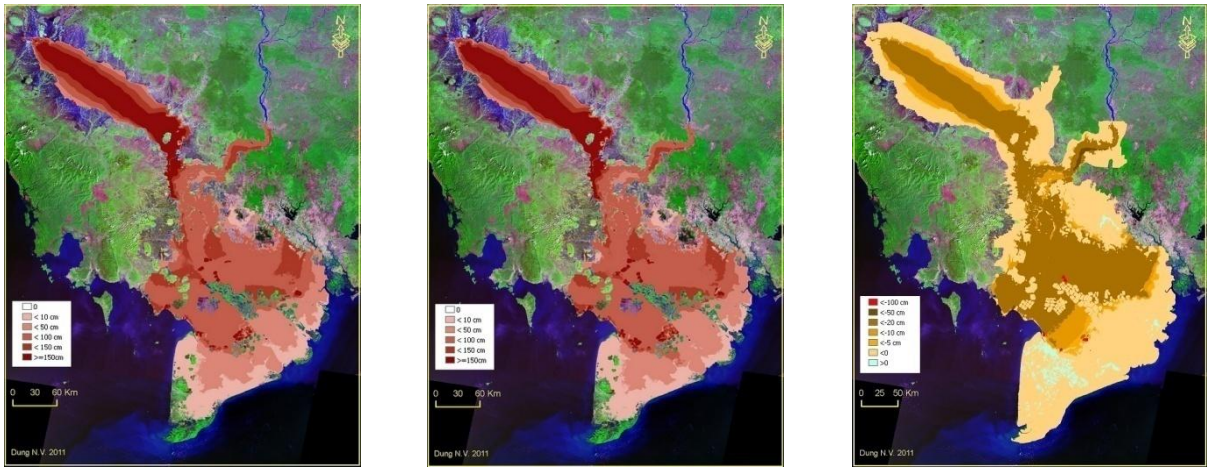


EU_NS2030_50



EU_NS2030_95

Figure 6.7-2: : Flood hazard maps for maximum inundation depth over the study domain derived with the roughness parameter set “shortest Euclidean distance” for the hydraulic model (Table 6.7.1 explains the codes of the individual maps)



D1 D2 D3

Figure 6.7-3: Comparison maps. Left: the difference map derived by subtracting F2_S_05 from F2_S_95; middle: the difference map derived by subtracting F2_NS2009_05 from F2_NS2009_95; right: the difference map derived by subtracting F2_S_50 from F2_NS2009_50 (explanation to codes in Table 6.7.1)

Chapter 7

CONCLUSIONS AND OUTLOOK

7.1 Conclusions

This thesis focused on two main aspects:

(i) the problem of calibration and validation of large-scale hydrodynamic models

(ii) flood hazard mapping using a probabilistic approach

To deal with the first aspect, we developed a suitable concept for the automatic calibration of hydrodynamic models, and applied it to the complex case study of the Mekong Delta.

This study used one-dimensional (quasi-two-dimensional) hydrodynamic approach to simulate the inundation processes of the whole Mekong Delta, one of the largest estuaries in the world with a highly complex hydraulic system. There had been a few preliminary attempts for calibration and validation. A manual calibration against gauging data had been implemented for the flood season of the year 2000. Validation had been done for the flood year 2001. However, due to the missing calibration against spatial inundation dynamics, and the inherent subjectivity within manual calibration, the results, especially in terms of inundation extents, needed to be treated with caution. These limitations suggested to constrain the calibration by more than one objective function, and directed me towards a multi-objective view of the calibration problem.

We started from the point that automatic calibration techniques are hardly ever found in flood inundation modeling due to lack of appropriate data, large computational demand, and an inappropriate framework. This thesis studied recent developments that could be used to overcome such limitations. The calibration was seen from a multi-objective viewpoint using in parallel high-temporal, low-spatial resolution data and low-temporal, high-spatial data. The objective functions for the calibration were defined based on two data sets: six months of daily stage measurements at gauging stations along the Mekong main streams representing mainly the temporal flow variation, and the series of 14 flood extent maps derived from ENVISAT ASAR products representing mainly the spatial variation. The first objective function was formulated

using the most commonly used measure in the calibration of hydrological models: the Nash-Sutcliffe coefficient of efficiency. The second objective function was defined utilizing the most recommended flood-area index, which is based on the binary pattern measure. Furthermore, we developed a new method making use of the Monte Carlo procedure. This procedure considers uncertainties residing in both simulation (by the model set-up and imperfect spatial representation) and flood maps (by classification errors and geo-referencing).

The crucial goal of the calibration process was to find the best parameter sets to represent both objectives. The calibration parameter is the roughness, or the resistance to flow coefficient. To reduce the degree of freedom that could cause a handicap to the automatic calibration, the huge number of river and floodplain elements (for a model with more than 26000 computational nodes) was grouped into a reasonable number of classes (five for the calibration and six for the re-calibration), based on prior knowledge of the hydraulic regime of the flow in the study area. An automatic calibration process based on the multi-objective genetic algorithm, NSGA-II, was developed in order to optimize both objectives simultaneously. A large number of model evaluations were launched (1560 simulation runs for the calibration with five groups of parameters, 2080 simulation runs for the calibration with six groups of parameters, and 52 simulation runs for the validation). It would have taken more than a year if these model runs were implemented on a fast single processor computer (excluding the time for input and output processing). Therefore, in order to cut down on computation time, we implemented a master-slave parallelization scheme on a multi-processor CPU. We also developed the wrapper code, the parallel version of NSGA-II, controlling the whole calibration process.

The calibration was successfully achieved, resulting in several findings:

- **Firstly**, we conclude that an automatic, multi-objective calibration of hydrodynamic models is possible, even at such a large-scale with such complex applications as in the Mekong Delta. This is an important step towards delivering more objectivity in hydrodynamic model calibration.
- **Secondly**, a combination of inundation maps and stage hydrographs enables identification of temporal and spatial performance of the hydrodynamic models.
- **Thirdly**, the calibration results provide trade-off solutions, known as Pareto optimal solutions, which are the best model parameter sets taking into account both objectives. It means that an improved performance in one of the two objectives, which measure the model skill with respect to the temporal or spatial aspects, can only be achieved at the expense of the other objective. This means in consequence that model parameterizations providing satisfying performance in both temporal and spatial dynamics cannot be found with the given model setup. Thus no general recommendable parameter set exists

and an appropriate parameterization has to be chosen according to the specific purpose of different model applications.

- **Fourthly**, in light of the Pareto-based approach, all Pareto optimal solutions are equally important, as it is difficult to prefer one solution over another without any further information about the problem (Fenicia et al., 2007). Thus, the Pareto-based approach allows the user to choose a particular solution based on the importance of the calibration criteria involved. For example, if the spatial pattern of flooding is of higher relevance, the parameter set obtaining the best objective function F2 is the one to choose. If the temporal performance in the main stream of the simulation area is highly important, the solution for the best objective function F1 should be selected. Decision makers could also use subjective, pre-defined criteria. In this research work, we proposed a criterion based on the Euclidean distance concept, which can harmonize the performance of the two objectives.
- **Fifthly**, there was no significant difference between the calibration with five groups of parameters and re-calibration with six groups of parameters. This was illustrated by the similarity of their Pareto-optimal solutions. It means that five groups of parameters can be used to represent the spatial roughness in the network of the hydrodynamic model of the Mekong Delta. Furthermore, the validation results showed consistency with those in the calibration, which means the parameter sets obtained are acceptable.
- **Sixthly,** using a Pareto-based approach for multi-objective calibration highlighted a possible model structural deficiency that would not have been detected when the aggregation method had been applied. When exploring the best solutions for the single objective functions through calibration and re-calibration, it became clear that the model contains deficiencies in representing the dike system in Vietnam. Thus, we can conclude that the automatic, multi-objective calibration is not only able to parameterize a hydrodynamic model properly, but it is also able to identify model deficiencies on an objective basis. This conclusion was corroborated by a sensitivity simulation using the best Euclidean distance parameter set and lowering dike heights by generally 20%, which improved the performance of the model. By obtaining a set of Pareto-optimal solutions, it is also possible to provide an ensemble of simulation runs, by evaluating the Pareto-optimal parameterizations.

As a final remark or recommendation with respect to this first aspect of the study, we suggest that automatic, multi-objective calibration of hydrodynamic models can offer several benefits. Hence, we recommend to use this approach more often in hydrodynamic modeling.

In the last part of the dissertation, we developed a novel approach for mapping flood hazard in the Mekong Delta, making use of the large scale model with calibrated parameter sets derived by an automatic multi-objective optimization method. We also took the uncertainties in the main hydrological input of the Mekong Delta as well as in the hydraulic model parameterization into account.

Flood frequency analysis and cluster analysis were applied to study the three aspects of the complete discharge hydrograph: peak, volume and shape. A method based on copulas was used to construct a bivariate flood peak and volume distribution and to analyze the bivariate frequency. The hydrograph shape was considered by the extraction of dimensionless characteristic hydrograph from the observed time series by a cluster analysis. Furthermore, alongside the traditional approach of flood frequency analysis using stationary assumptions, an approach taking the non-stationarity into account was also examined. The analysis has shown that both flood-peak and flood-volume can be modeled as log-normal distributions for in a stationary and non-stationary analysis. Negative trends in location-parameter and a positive trend in scale-parameter were detected in both peak discharge and volume time series. The two-step inference from margin (IFM) method was adopted to estimate the parameters of all the candidate copulas in both stationary and non-stationary analysis. The Gaussian copula was selected to model the bivariate flood-peak and flood-volume, as it provided the best fit for the relationship between them. Based on this copula, pairs of peak and volume were generated, based on pre-defined hydrological scenarios: (1) stationary, 100-year return period; (2) non-stationary, for the baseline year 2009, 100-year return period; (3) non-stationary for the future year 2030, 100-year return period. The analysis of the hydrograph shape yielded four characteristic hydrograph classes, each of which was assigned a probability of occurrence. By combining the different analyses, numerous synthetic hydrographs were generated corresponding to the three 100-year scenarios.

To map the flood hazard, the calibrated large-scale hydrodynamic model of the Mekong Delta was applied using the generated synthetic flood hydrographs as the boundary conditions, and the two optimal parameter sets derived from the calibration stage of the model (best F2 and best Euclidean sets, see Chapter 5). The Monte Carlo/Hypercube sampling framework was used to deal with the uncertainties in the hazard assessment. 600 model simulations were evaluated (taking 5 days in parallel mode) and interpolated to inundation maps. For each scenario, they illustrated different flood depth percentile maps (5%, 50%, 95%). The maps show the large range of potential maximum inundation depths associated with the probability of the occurrence of a flood event. The second aspect of this study was successfully achieved, resulting in the following main findings:

- **Firstly**, changes in the upstream flow (at Kratie) directly influence the Cambodia floodplains and the Northern areas in the delta of Vietnam. The impact of this change over the coastal regions in Vietnam, especially the Eastern area, is small.
- **Secondly**, the difference in the flood depths among the various scenarios is not significant. This was illustrated by the similarity in the flood maps for the stationary case, and the two non-stationary cases, for the two reference years of 2009 and 2030. This means that a bivariate stationary approach would be appropriate for the Mekong Delta, even for a moderate projection into the future despite detected trends in flood variability. This finding has to be attributed to a) the selected moderate probability of occurrence (100 year event) in the hazard analysis and b) to the fact, that the fitted stationary and non-stationary probability density distributions do not exhibit a heavy tail, i.e. extreme events do not differ to normal events in magnitudes. This indicates the "normality" of floods in the Mekong Delta also from a statistical point of view.
- **Thirdly**, there was a significant difference among the flood maps when different parameterizations were applied to the hydrodynamic model. The maps for the best performance in simulating inundation extent show deeper inundation and larger extent compared to those for the case best Euclidean Parameter set harmonizing spatial and temporal model performance. The use of different model parameterizations by taking into account all Pareto optimal solutions, allows considering uncertainties in the hydraulic model parameterization.
- **Fourthly**, the derived quantile maps of inundation depths (hazard maps) may serve as a valuable basis for probabilistic flood-risk management plans, allowing for inclusion of uncertainties in the estimation of extreme events.

As a final remark or recommendation with respect to the second aspect of this study, copula-based analysis is a useful tool for stationary and non-stationary multivariate flood-frequency analysis. The flood model in the Mekong Delta has proven its usefulness in flood-hazard assessment including different types of uncertainties.

7.2 Outlook

Based on the observations made during the research work and the scope of this study presented here, it is also worthwhile to describe the following outlooks on the direction that future work could take:

- **Towards improvement of the flood model representation:** the main goal of the calibration process of the hydrodynamic flood model was to cover the uncertainty caused by the parameterization. Our primary aim was not to improve the model

structure through calibration, although the multi-objective calibration gave insights into model deficiencies. Although great efforts have been spent in implementing the model, its representation of reality may not be sufficiently correct for many flood-related questions, e.g. the size of the compartments in some areas is still too large. Hence, research on how to improve the model structure could be a way for future work.

- **Calibration using water quality data**: As mentioned in Chapter 3, the hydrodynamic model has the potential to do more than just model floods. It is also capable of simulating water quality, e.g. salinity intrusion, sediment transportation, etc. To be useful for this purpose, the model must be calibrated using appropriate data. It would be interesting to investigate, to which extent a multi-objective calibration of higher order than two would deliver additional insight in model deficiencies.
- **Inundation duration map generation:** Due to the constraint caused by computation time, in this study only maximum flood depth maps were produced. For future work, duration maps should be generated as the inundation duration information could also give valuable data for risk assessment in the study area.
- **The issue of probability of occurrence**: Several probabilistic hazard maps were derived from the simulations. These maps, however, were formed by selecting the return period "AND" case. Selecting this return period may lead to an over-estimation of the risk associated with a given flood. For future mapping scenarios, beside the use of the "AND" return period case, the "OR" case scenarios should be considered.
- **Bivariate analysis of highly extreme values:** In this study, the hydrological scenarios were built on a "not very high" return period (100 years).Therefore, it was reasonable to select the bivariate Gaussian copula, although it does not take into account tail dependency. To work with a "high return period", there would be need to study tail dependence in detail. The use of the Gumbel copula or the t-copula, both of which introduce tail-dependence, may be more appropriate.
- **Correlation between climate signals and the frequency of floods in the Mekong Delta:** Although the bivariate stationary analysis seems to be appropriate for moderate future projection (2030), there is a tendency that the difference between the stationary and non-stationary approaches increases in the future which should be checked and corroborated. One possibility would be to derive scenarios of future flood hazard from Global Circulation Model (GCM) ensembles, e.g. by using the correlation between discharge PDF parameters and atmospheric circulation indexes (Coles, 2001; Delgado et al. 2011).

APPENDIX A

SOME STATISTICAL DISTRIBUTIONS

(Hosking and Wallis, 1997; Robson and Reed, 1999; Coles, 2001)

1. Normal (Gaussian) distribution 2 parameters – N2

Probability density function (PDF):

$$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}{\sigma\sqrt{2}\pi} \quad \text{(A.1)}$$

Cumulative distribution function (CDF):

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad \text{(A.2)}$$

where Φ(x) is the Laplace Integral:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x \exp\left(-\frac{t^2}{2}\right)dt \quad \text{(A.3)}$$

Parameters:

σ – scale parameter (σ >0)

μ – location parameter

Domain:

D = (-∞,+∞) - unbounded

2. Log-Normal distribution (3 parameters) – LN3

Probability density function (PDF):

$$f(x) = \frac{\exp(ky - y^2/2)}{\sigma\sqrt{2}\pi} \quad \text{(A.4)}$$

Where:

$$y = \begin{cases} -k^{-1}\log\left\{1 - \dfrac{k(x-\xi)}{\alpha}\right\}, & k \neq 0 \\ \dfrac{(x-\xi)}{\alpha}, & k = 0 \end{cases} \quad \text{(A.5)}$$

Cumulative distribution function (CDF):

$$F(x) = \Phi(y) \quad \text{(A.6)}$$

Parameters:

α – scale parameter ($\alpha > 0$)

ξ – location parameter

k – shape parameter

Domain:

- If $k > 0$ then D = $(-\infty, \xi + \alpha/k]$ (left skew)
- If $k = 0$ then D = $(-\infty, +\infty)$ (special case, standard Normal distribution)
- If $k < 0$ then D = $[\xi + \alpha/k, +\infty)$ (right skew)

3. Generalized Extreme Value (3 parameters) – GEV3

Probability density function (PDF):

$$f(x) = \begin{cases} \sigma^{-1}(1+ky)^{-1-1/k}\exp(-(1+ky)^{-1/k}) & k \neq 0 \\ \sigma^{-1}\exp(-y-\exp(-y)), & k = 0 \end{cases} \quad \text{(A.7)}$$

Where:

$$y = \frac{x-\mu}{\sigma}$$

Cumulative distribution function (CDF):

$$F(x) = \begin{cases} \exp(-(1+ky)^{-1/k}) & k \neq 0 \\ \exp(-\exp(-\mathrm{y})), & k = 0 \end{cases} \quad \text{(A.8)}$$

Parameters:

σ – scale parameter ($\sigma > 0$)

μ – location parameter

k – shape parameter

Domain:

- If $k < 0$ then D = $(-\infty, \mu - \sigma/k]$ (inverse Weillbull distribution)
- If $k = 0$ then D = $(-\infty, +\infty)$ (Gumbel distribution)
- If $k > 0$ then D = $[\mu - \sigma/k, +\infty)$ (Fréchet distribution)

KOLMOGOROV-SMIRNOV TEST RESULTS

Table A-1: K-S test result for peak series

No.	Distribution	statistic	p-value	CV			Reject at		
				95%	*90%*	*80%*	*95%*	*90%*	*80%*
1	N2	0.087	0.500	0.144	0.130	0.114	No	No	No
2	LN3	0.075	0.690	0.144	0.130	0.114	No	No	No
3	GAM2	0.069	0.785	0.144	0.130	0.114	No	No	No
4	P3	0.076	0.676	0.144	0.130	0.114	No	No	No
5	LP3	0.073	0.717	0.144	0.130	0.114	No	No	No
6	GUM2	0.100	0.339	0.144	0.130	0.114	No	No	No
7	GEV3	0.074	0.701	0.144	0.130	0.114	No	No	No

Table A-2: K-S test result for volume series

No.	Distribution	KS test	p-value	CV			Reject at		
				95%	*90%*	*80%*	*95%*	*90%*	*80%*
1	N2	0.044	0.993	0.144	0.130	0.114	No	No	No
2	LN3	0.047	0.988	0.144	0.130	0.114	No	No	No
3	GAM2	0.055	0.945	0.144	0.130	0.114	No	No	No
4	P3	0.044	0.993	0.144	0.130	0.114	No	No	No
5	LP3	0.049	0.979	0.144	0.130	0.114	No	No	No
6	GUM2	0.094	0.411	0.144	0.130	0.114	No	No	No
7	GEV3	0.052	0.962	0.144	0.130	0.114	No	No	No

APPENDIX B

COPULA PARAMETER ESTIMATION METHODS

1. Canonical representation

For continuous random variables (limiting to bivariate cases in this study), the copula density c of the copula C is related to the density of the joint distribution H, denoted as h by the canonical representation (for the case of multivariate representation, see Cherubini et al. (2004)):

$$h(x,y) = c\big(F(x), G(y)\big) f(x) g(y) \tag{B.1}$$

Where:

$$c(F(x), G(y)) = \frac{\partial^2 (C(F(x), G(y))}{\partial F(x) \partial G(y)}$$

And f and g are the densities of the marginals having the distribution function F and G, respectively:

$$f(x) = \frac{dF(x)}{dx} \text{ and } g(y) = \frac{dG(y)}{dy} \tag{B.2}$$

The log-likelihood for a sample of N observations is given by

$$L(\theta) = \sum_{t=1}^{N} L_t(\theta) = \sum_{t=1}^{N} ln(c(F(x_t), G(y_t)) + \sum_{t=1}^{N} (\ln(f(x_t) + g(y_t)) \tag{B.3}$$

Where: $\theta = (\theta_C, \theta_M) \in \Omega_C \times \Omega_M$ is the set of the all parameters of both the copula $\theta_C \in \Omega_C$ and the marginals $\theta_M \in \Omega_M$. The log-likelihood form above could be represented as belows:

$$L(\theta) = L_C(\theta_C, \theta_M) + L_M(\theta_M) \tag{B.4}$$

2. Exact maximum likelihood method

In this method, all the parameters appeared in the log-likelihood in Equation (B.3), $L(\theta)$, will be estimated by maximizing the likelihood.

$$\hat{\theta} = (\hat{\theta}_C, \hat{\theta}_M) = \underset{\theta \in \Omega}{\operatorname{argmax}}\, L(\theta) \tag{B.5}$$

Under regularity conditions mentioned in Cherubini et al.(2004), the maximum likelihood estimator exists and it is consistent and asymptotically efficient. However, this method may be very computationally intensive, especially in the case of a high dimension, because it is necessary to estimate jointly the parameters of the marginal distributions and the parameters of the dependence structure represented by the copula.

3. The Inference From Margins method (IFM)

This method was first mentioned in Joe (1997) and later introduced in Cherubini (2004) and Genest and Favre (2007). Looking more closely at the log-likelihood function, one can find that the log-likehood function above is composed by two terms: one term involving the copula density and its parameters, and one term involving the margins and all parameters of the copula density. For that reason, Joe (1997) proposed that these set of parameters should be estimated in two steps:

Step 1: Estimate the margins' parameters θ_M by performing the estimation of the univariate marginal distributions:

$$\hat{\theta}_M = (\hat{\theta}_{Mf}, \hat{\theta}_{Mg}) = \underset{\theta_M \in \Omega_M}{\operatorname{argmax}}\, L_M(\theta_M) \tag{B.6}$$

where $\hat{\theta}_{Mf}$, $\hat{\theta}_{Mg}$ are the estimated parameter vectors of the marginals with density function f and g, respectively. $\hat{\theta}_M$ is the combined vector of both.

Step 2: given $\hat{\theta}_M$, the estimation of the copula parameter θ_C is carried out:

$$\hat{\theta} = \left(\hat{\theta}_C, \hat{\theta}_M\right) = \underset{\theta_C \in \Omega_C}{\operatorname{argmax}}\, L(\theta_C, \hat{\theta}_M) \tag{B.7}$$

This two-step method is called inference from the margins or IFM. The equivalence of the two estimators, in general, does not hold. It is simple to see that the IFM estimator provides a good starting point for obtaining an EML. Joe (1997) points out that the IFM method is highly efficient compared with the EML method. It is worth noting that the IFM method may be viewed as a special case of the generalized method of moments (GMM) with an identity weight matrix.

Although IFM method often does well for the parameter estimation, the estimates of the association parameters derived by the IFM technique clearly depend on the choice of F and G, and thus always run the risk of being unduly affected if the models selected for the margins turn out to be inappropriate.

4. Canonical maximum likelihood (CML)

In this method, the copula parameters may be estimated without specifying the marginals. In fact, another estimation method consists in transforming the sample data into uniform variates and then estimating the copula parameters. This method is called the Canonical maximum likelihood (CML) method. In this case, the CML estimator could be viewed as an EML, given the observed margins. This method is also called maximum pseudo-likelihood (MPL) method (Genest and Favre, 2007). This method could also be seen as a rank-based method. The others are the rank-based methods base on Kendall's tau and Spearman's Rho. It is worth being noted that the exclusive reliance on ranks for copula parameter estimation does not make complete consensus in the statistical community.

BIBLIOGRAPHY

Abbott, M. B.: Computational Hydraulics: Elements of the Theory of Free Surface Flows, Pitman, London, 1979.

Adamson, P. T., Metcalfe, A. V., and Parmentier, B.: Bivariate extreme value distributions: An application of the Gibbs sampler to the analysis of floods, Water Resources Research, 35, 2825-2832, 1999.

Adamson, P. T., Rutherfurd, I. D., and M.C., P.: The hydrology of the Mekong river, in: Biophysical Environment of an International River Basin, edited by: Campbell, I. C., Elservier, New York, USA, 432pp, 2009.

Akaike, H.: A New Look at the Statistical Model Identification, IEEE transactions on automatic control, 19, 1974.

Apel, H., Thieken, A. H., Merz, B., and Blöschl, G.: Flood risk assessment and associated uncertainty, Natural Hazards and Earth System Science, 4, 295-308, 2004.

Apel, H., Thieken, A. H., Merz, B., and Blöschl, G.: A probabilistic modelling system for assessing flood risks, Natural Hazards, 38, 79-100, 2006.

Apel, H., Aronica, G., Kreibich, H., and Thieken, A.: Flood risk analyses—how detailed do we need to be?, Natural Hazards, 49, 79-98, 2009.

Aronica, G., Hankin, B., and Beven, K.: Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data, Advances in Water Resources, 22, 349-365, 1998a.

Aronica, G., Tucciarelli, T., and Nasello, C.: 2D multilevel model for flood wave propagation in flood-affected areas, Journal of Water Resources Planning and Management, 124, 210-217, 1998b.

Aronica, G., Bates, P. D., and Horritt, M. S.: Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE, Hydrological Processes, 16, 2001-2016, 2002.

Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters, Water Resources Research, 42, 1-12, 10.1029/2005wr004754, 2006.

Bárdossy, A., and Singh, S. K.: Robust estimation of hydrological model parameters, Hydrol. Earth Syst. Sci., 12, 1273-1283, 2008.

Bates, G., Montana Constitutional Society of 1972., and Montana Centennial Commission.: 100 delegates, Montana Constitutional Convention of 1972, Montana Centennial Commission, Manhattan, Mont., 110 p. pp., 1989.

Bates, P. D., and De Roo, A. P. J.: A simple raster-based model for flood inundation simulation, Journal of Hydrology, 236, 54-77, 2000.

Bates, P. D.: Remote sensing and flood inundation modelling, Hydrological Processes, 18, 2593-2597, 2004.

Bekele, E. G., and Nicklow, J. W.: Multi-objective automatic calibration of SWAT using NSGA-II Journal of Hydrology, 341, 165-176 doi:10.1016/j.jhydrol.2007.05.014 2007.

Blasone, R. S., Madsen, H., and Rosbjerg, D.: Comparison of parameter estimation algorithms in hydrological modelling, IAHS-AISH Publication, 304, 67-72, 2006.

Brunham, K. P., and Anderson, D. R.: Model Selection and Inference: A Practical Information-Theoretic Approach, Second Edition ed., Springer-Verlag, New York, USA, 488 pp., 2002.

Carrivick, J. L.: Application of 2D hydrodynamic modelling to high-magnitude outburst floods: An example from Kverkfjöll, Iceland, Journal of Hydrology, 321, 187-199, 10.1016/j.jhydrol.2005.07.042, 2006.

Chaudhry, M. H.: Open-Channel Flow Second Edition Springer, New York, USA, 540 pp., 2008.

Cherubini, U., Luciano, E., and Vecchiato, W.: Copula Methods in Finance, John Wiley & Sons, Ltd, Chichester, 308 pp., 2004.

Chow, V. T.: Open-Channel Hydraulics, McGraw-Hill, New York, 680 pp., 1973.

Chow, V. T., Maidment, D. R., and Mays, L. W.: Applied Hydrology, McGraw-Hill, New York, 572 pp., 1988.

Coles, S.: An introduction to Statistical Modeling of Extreme Values, Springer, 2001.

Cunge, J. A.: Two-dimensional modeling of floodplains, Water Resources Publications, 2, 705-762, 1975.

Cunge, J. A.: Of data and models, Journal of Hydroinformatics, 5, 75-98, 2003.

Dac, N. T.: Mathematical method for modelling flows and water quality in channels, Agriculture Publishing House, Ho Chi Minh city, 2005.

Deb, K., and Agrawal, R. B.: Simulated Binary Crossover for Continuous Search Space, Complex Systems, 9, 115-148, 1995.

Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T.: A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II, Kanpur Genetic Algorithms Laboratory (KanGAL), Indian Institute of Technology Kanpur, Kanpur, 11, 2002.

Delgado, J., Merz, B., and Heiko, A.: Flood-frequency change in the Mekong river: scenarios for the 21st century, Geophysical Research Abstracts, EGU General Assembly, Vienna, Austria, 2011, EGU2011-11147,

Delgado, J. M., Apel, H., and Merz, B.: Flood trends and variability in the Mekong river, Hydrology and Earth System Sciences, 14, 407-418, 10.5194/hess-14-407-2010, 2010.

DHI: MIKE 11 - A Modelling System for Rivers and Channels, DHI Water & Environment, 516, 2004.

DHI: MIKE 21 Flow Model - Hydrodynami module scientific document, 58, 2007a.

DHI: MIKE FLOOD: 1D-2D Modelling user manual, 108, 2007b.

Di Baldassarre, G., Schumann, G., and Bates, P. D.: A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent, Journal of Hydrology, 367, 276-282, 2009.

Doyle, T. W., Day, R. H., and Michot, T. C.: Development of Sea Level Rise Scenarios for Climate Change Assessments of the Mekong Delta, Vietnam, U.S. Geological Survey Open-File Report 2010–1165, 110, 2010.

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and efficient gloabl optimization for conceptual rainfall-runoff models, Water Resources Research, 28, 1015-1031, 1992.

Duan, Q., Gupta, V., and Sorooshian, S.: Shuffled Complex Evolution Approach for effective and efficient global minimization, Journal of Optimization Theory and Applications, 76, 501-521, 1993.

Duan, Q., Sorooshian, S., and Gupta, V. K.: Optimal use of the SCE-UA global optimization method for calibrating watershed models Journal of Hydrology, 158, 265-284, 1994.

Dung, N. V., and Thang, D. T.: Updating information and analysing national and international impacts to the dike system in the Mekong delta, Southern Institute of Water Resources Research (SIWRR), Ho Chi Minh city, Technical report, 2006.

Dung, N. V., and Thang, D. T.: Establishment of a large scale hydrodynamic model for modelling floods in the Mekong Delta using MIKE11, SIWRR - Southern Instute of Water Resources Research Ho Chi Minh City, Technical report, 2007.

Dung, N. V.: Hydrodynamic modeling of floods in the Mekong Delta on different scales, IWRM / WISDOM Modeling-Workshop, Bochum, 2008.

Dung, N. V., Apel, H., and Merz, B.: Floods in the Mekong Delta: a future perspective, European Geosciences Union, Austria, 19 – 24. April 2009 (poster).

Dung, N. V., Apel, H., and Merz, B.: Integration of remote sensing and in-su data in automatic calibration of hydraulic models. A case study in the Mekong Delta, European Geosciences Union, Vienna, Austria, 2010 (talk).

Dung, N. V., Merz, B., Bárdossy, A., Thang, T. D., and Apel, H.: Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data, Hydrol. Earth Syst. Sci., 15, 1339-1354, doi:10.5194/hess-15-1339-2011, 2011.

Durante, F., and Salvadori, G.: On the construction of multivariate extreme value models via copulas, Environmetrics, 143-161, 10.1002/env, 2010.

Dutta, D., Alam, J., Umeda, K., Hayashi, M., and Hironaka, S.: A two-dimensional hydrodynamic model for flood inundation simulation: a case study in the lower Mekong river basin, Hydrological Processes, 21, 1223 - 1237, 2007.

Eberhart, R. C., and Kennedy, J.: A new optimizer using particle swarm theory, Proceedings of the Sixth International Symposium on Micro Machine and Human Science MHS'95, 1995, 39-43,

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R., & Bobée, B.: Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. Water Resources Research, *43*(3), 1-13. doi: 10.1029/2005WR004545, 2007.

Fabio, P., Aronica, G. T., and Apel, H.: Towards automatic calibration of 2-D flood propagation models, Hydrol. Earth Syst. Sci., 14, 911–924, doi:10.5194/hess-14-911-2010, 2010.

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: A comparison of alternative multiobjective calibration strategies for hydrological modeling, Water Resour. Res., 43, W03434, doi:10.1029/2006WR005098, 2007.

Frank, E., Ostan, A., Coccato, M., and Stelling, G. S.: Use of an integrated one dimensional modelling approach for flood hazard and risk mapping, In: Proceedings of the 1st Conference on River Basin Management, Southampton,UK, 2001, 99-108,

Fujii, H., Garsdal, H., Peterward, Ishii, M., Morishita, K., and Boivin, T.: Hydrological roles of the Cambodian floodplain of the Mekong River, Intl. J. River Basin Management, 1, 2003.

Genest, C., Ghoudi, K., and Rivest, L.-P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distribution., Biometrika, 82, 543-552, 1995.

Genest, C., and Rémillard, B.: Discussion of “Copulas: Tales and facts" by Thomas Mikosch, Extremes, 9, 27-36, 2006.

Genest, C., and Favre, A.-C.: Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask, Journal of Hydrologic Engineering, 12, 347-347, 10.1061/(asce)1084-0699(2007)12:4(347), 2007.

Glover, F.: Future paths for integer programming and links to artificial intelligence, Computers & Operations Research, 13, 533-549, 1986.

Goldberg, D.: Genetic algorithms in search, optimization, and machine learning, ADDISON-WESLEY Publ. Co., Inc., Massachusetts, 1989.

Gstaiger, V., and Gebhardt, S.: Multi-sensoral derivation of inundated areas using TerraSAR-X and ENVISAT ASAR data., International Journal of Remote Sensing, summited, under final revision, 2011.

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models; multiple and noncommensurable measures of information, Water Resources Research, 34, 751-763, 1998.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80-91, 10.1016/j.jhydrol.2009.08.003, 2009.

Hendrik, E. M. T., and Toth, B. G.: Introduction to nonlinear and global optimization, Optimization and Its Applications, edited by: Pardalos, P., and Du, D.-Z., Springer, New York, 220 pp., 2010.

Hervouet, J.-M., Hubert, J.-L., Janin, J.-M., Lepeintre, F., and Peltier, E.: The computation of free surface flows with TELEMAC. An example of evolution towards hydroinformatics, Journal of Hydraulic Research, 39, 45-64, 1994.

Hiemstra, L. A. V., Zuchhini, W.S., and Pegram, G.G.S.: A method of finding the familiy of runhydrographs for given return periods, Journal of Hydrology, 30, 95-103, 1976.

Hirsch, H.: Numerical Computation of Internal and External Flows. Vol.2:Computational Methods for Inviscid and Viscous Flows, Wiley, New York, 714 pp., 1990.

Hoa, L. T. V., Nhan, N. H., Wolanski, E., Tran, T. C., and Haruyama, S.: The combined impact on the flooding in Vietnam's Mekong River delta of local man-made structures, sea level rise, and dams upstream in the river catchment, Estuarine, Coastal and Shelf Science, 71, 110-116, 2007.

Hoa, L. T. V., Shigeko, H., Nhan, N. H., and Cong, T. T.: Infrastructure effects on floods in the Mekong River Delta in Vietnam, Hydrological Processes, 22, 1359-1372, 2008.

Hoc, D. X.: Assessment on the process of the flood year 2000 and its impact on the Plain of Reeds, Vietnam, University of Water Resources, Ho Chi Minh City, 67, 2000.

Hoi, T. N.: Research on scientific and technological solutions on building dike system for sustainable development in the Mekong Delta, Southern Institute of Water Resources Research, Ho Chi Minh City, 2005.

Holland, J. H.: Adaptation in Natural and Artificial Systems, University of Michigan Press, An Arbor, Michigan, 1975.

Horn J., Nafpliotis N., and Goldberg D.E.: A niched Pareto genetic algorithm for multiobjective optimization., In: Proceedings of the first IEEE conference on evolutionary computation., Orlando, USA, 1994,

Horritt, M.: Development and testing of simple 2D finite volume model of sub-critical shallow water flow, International Journal for Numerical Methods in Fluids, 44, 1231-1255, 2004.

Horritt, M. S., and Bates, P. D.: Predicting floodplain inundation: Raster-based modelling versus the finite-element approach, Hydrological Processes, 15, 825-842, 2001.

Horritt, M. S., and Bates, P. D.: Evaluation of 1D and 2D numerical models for predicting river flood inundation, Journal of Hydrology, 268, 87-99, 2002.

Horritt, M. S.: A methodology for the validation of uncertain flood inundation models, Journal of Hydrology, 326, 153-165, 2006.

Hosking, J. R. M., and Wallis, J.R.: Regional frequency analysis, Cambridge University Press, New York, 238-238 pp., 1997.

Hunter, N. M.: Development and assessment of dynamic storage cell codes for flood inundation modelling, School of Geographical Sciences, University of Bristol, Bristol, 360 pp., 2005.

Hunter, N. M., Bates, P. D., Horritt, M. S., De Roo, A. P. J., and Werner, M. G. F.: Utility of different data types for calibrating flood inundation models within a GLUE framework, Hydrol. Earth Syst. Sci., 9, 412-430, 2005a.

Hunter, N. M., Bates, P. D., Horritt, M. S., de Roo, P. J., and Werner, M. G. F.: Utility of different data types for calibrating flood inundation models within a GLUE framework, Hydrology and Earth System Sciences, 9, 412-430, 2005b.

Huth, J., Gebhardt, S., W., T., S., I. , , Kunzer, C., Schmidt, M., and Dech, S.: Automated inundation monitoring using TerraSAR-X multi-temporalimagery,, European Geoscience Union Annual Assembly 2009, Vienna, Austria, 2009.

IPCC: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A.(eds.)], IPCC, Geneva, Switzerland, 104pp, 2007.

Joe, H.: Multivariate models and dependence concepts, Chapman & Hall, USA, 418-418 pp., 1997.

Joe, H.: Asymptotic efficiency of the two-stage estimation method for copula-based models, Journal of Multivariate Analysis, 94, 401-419, 10.1016/j.jmva.2004.06.003, 2005.

Joe, H.: Discussion of "Copulas: Tales and facts", by Thomas Mikosch, Extremes, 9, 37-41, 10.1007/s10687-006-0019-6, 2006.

Katopodes, N. D.: A Dissipative Galerkin Scheme for Open-Channel Flow, Journal of the Hydraulics Division, ASCE, 110, 450-466, 1984.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Advances in Water Resources, 25, 1287-1304, 10.1016/s0309-1708(02)00056-8, 2002.

Kendall, M. G.: Rank correletion methods, Fourth Edition ed., J.W.Arrowsmith Ltd, Bristol, 210 pp., 1970.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P., and Bobee, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, Journal of Hydrology, 329, 534-552, 2006.

Khu, S. T., and Madsen, H.: Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration, Water Resour. Res., 41, W03004, doi:10.1029/2004WR003041, 2005.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P.: Optimization by simulated annealing, Science, 220, 671-680, doi:10.1126/science.220.4598.671, 1983.

Klein, B., Pahlow, M., Hundecha, Y., and Schumann, A.: Probability Analysis of Hydrological Loads for the Design of Flood Control Systems Using Copulas, Journal of Hydrologic Engineering, 15, 360-360, 10.1061/(asce)he.1943-5584.0000204, 2010.

Kollat, J. B., and Reed, P. M.: Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design, Advances in Water Resources, 29, 792-807, DOI 10.1016/j.advwatres.2005.07.010, 2006.

Kollat, J. B., and Reed, P. M.: A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications, Advances in Water Resources, 30, 408-419, DOI 10.1016/j.advwatres.2006.05.009, 2007.

Konak, A., Coit, D. W., and Smith, A. E.: Multi-objective optimization using genetic algorithms: A tutorial, Reability engineering and system safety, 91, 16, 2006.

Lai, C.: Numerical Modeling of Unsteady Open-Channel Flows, in: Advances in Hydrosciences, Academic Press, New York, 161-333, 1986.

Legates, D. R., and McCabe, G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, 233-233, 10.1029/1998wr900018, 1999.

Madsen, H.: Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, Journal of Hydrology, 235, 276-288, 2000.

Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, Advances in Water Resources, 26, 205-216, 2003.

Madsen, H., and Khu, S. T.: On the use of Pareto optimization for multicriteria calibration of hydrological models, Calibration and Reliability in Groundwater Modelling: From Uncertainty to Decision Making, 304, 93-99, 2006.

Madsen, H., and Vinter, B.: Parameter optimisation in complex hydrodynamic and hydrological modelling systems using distributed computing, 7th International Conference on Hydroinformatics HIC 2006, Nice, France, 2006, 2489-2496,

Madsen, H.: Model calibration guideline, Report, Harmoni-CA project, 2007

Mann, H. B.: Nonparametric tests against trend, Econometrica, 13, 245-259, 1945.

Mason, D. C., Bates, P. D., and Dall' Amico, J. T.: Calibration of uncertain flood inundation models using remotely sensed water levels, Journal of Hydrology, 368, 224-236, 2009.

Merz, B., Kreibich, H., and Apel, H.: Flood risk analysis: uncertainties and validation, Östereichische Wasserwirtschaft (ÖWAWI), 05-06, 1-6, 2008.

Merz, B., and Thieken, A. H.: Flood risk curves and uncertainty bounds, Nat Hazards, 51, 437–458, 2009.

Mikosch, T.: Copulas: Tales and Facts, Extremes, 9, 3-20, 2006.

MONRE: Climate Change, sea level rise scenarios for Vietnam, Ministry of Natural Resources and Environment, Ha Noi, 34, 2009.

Most, H. V. D., and Wehrung, M.: Dealing with Uncertainty in Flood Risk Assessment of Dike Rings in the Netherlands, Natural Hazards, 36, 191-206, 2005.

MRC: Review of hydro-meteorological database of the MRC Secretariat, Mekong River Comission, Progress Report (DRAFT) 58, 2004.

MRC: Overview of the Hydrology of the Mekong Basin, Mekong River Commission, Vientiane, 73pp, 2005a.

MRC: Annual Flood Report 2005 Flood Management and Mitigation Progamme, Mekong River Commission, 2005b.

MRC: Annual Mekong Flood Report 2008, Mekong River Commission, Vientiane, 2009.

Nash, J. E., and Sutcliffe, I. V.: River flow forecasting through conceptual models. Part 1 - A discussion on principles, Journal of Hydrology, 10, 282-290, 1970.

Nelder, J., and Mead, R.: A simplex method for function minimization, Computer Journal, 7, 308-313, 1965.

Nelsen, R. B.: Introduction to copulas, 2th ed., Springer, New York, 2005.

Ngo, L. L., Madsen, H., and Rosbjerg, D.: Simulation and optimisation modelling approach for operation of the Hoa Binh reservoir, Vietnam, Journal of Hydrology, 336, 269-281, 2007.

Nien, N. A.: KOD model for flood computation in Cuu Long river delta., Proceedings of Conference on Flood Modeling in Cuu Long river delta, Ho Chi Minh city, 1996.

O'Connor, J. E., and Costa, J. E.: The world’s largest floods, past and present—Their causes and magnitudes, U.S. Geological Survey Circular 1254, 13pp, 2004.

Pappenberger, F., Beven, K., Frodsham, K., Romanowicz, R., and Matgen, P.: Grasping the unavoidable subjectivity in calibration of flood inundation models: A vulnerability weighted approach, Journal of Hydrology, 333, 275-287, 2007a.

Pappenberger, F., Frodsham, K., Beven, K., Romanowicz, R., and Matgen, P.: Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations, Hydrology and Earth System Sciences, 11, 739-752, 2007b.

Pedersen, G. K. M., and Goldberg, D. E.: Dynamic Uniform Scaling for Multiobjective Genetic Algorithms, in: Proceedings 3103, Genetic and Evolutionary Computation Gecco, 2004, 11-23, 2004.

Price, W. L.: Global optimization by controlled random search, Journal of Optimization Theory Application, 40, 333-348, 1983.

Priestnall, G., Jaafar, J., and Duncan, a.: Extracting urban features from LiDAR digital surface models, Computers, Environment and Urban Systems, 24, 65-78, 10.1016/s0198-9715(99)00047-2, 2000.

Rao, A. R., and H.Hamed, K.: Flood frequency analysis, CRC Press, USA, 356 pp., 2000.

Reddy, P. J. R.: A Textbook of Hydrology, Firewall Media, Delhi, India, 530 pp., 2005.

Reed, P., Kollat, J. B., and Devireddy, V. K.: Using interactive archives in evolutionary multiobjective optimization: A case study for long-term groundwater monitoring design, Environmental Modelling & Software, 22, 683-692, DOI 10.1016/j.envsoft.2005.12.021, 2007.

Renard, B., and Lang, M.: Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, Advances in Water Resources, 30, 897-912, 10.1016/j.advwatres.2006.08.001, 2007.

Robson, A., and Reed, D.: Flood Estimation Handbook, Volume 3: Statistical procedures for flood frequency estimation, edited by: Reed, D., Institute of Hydrology, Wallingford, UK, 338 pp., 1999.

Sakamoto, T., Van Nguyen, N., Ohno, H., Ishitsuka, N., and Yokozawa, M.: Spatio–temporal distribution of rice phenology and cropping systems in the Mekong Delta with special reference to the seasonal water flow of the Mekong and Bassac rivers, Remote Sensing of Environment, 100, 1-16, 10.1016/j.rse.2005.09.007, 2006.

Salvadori, G., and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, Water Resources Research, 40, 1-17, 2004.

Salvadori, G., Michele, C. D., Kottegoda, N. T., and Rosso, R.: Extremes in Nature: An approach using Copulas, Springer, Dordrecht, The Netherlands, 303-303 pp., 2005.

Salvadori, G., and De Michele C.: On the Use of Copulas in Hydrology: Theory and Practice, Journal of Hydrologic Engineering, 12, 10.1061/(ASCE)1084-0699(2007)12:4(369), 2007.

Sanders, B. F.: Evaluation of on-line DEMs for flood inundation modeling, Advances in Water Resources, 30, 1831-1843, 2007.

Schaffer, J. D.: Multiple Objective Optimization with Vector Evaluated Genetic Algorithms, In: Proceedings of the First International Conference on Genetic Algorithms, 1985,

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., and Pappenberger, F.: Progress in integration of remote sensing derived flood extent and stage data and hydraulic models, Rev. Geophys., 47, RG4001, doi:10.1029/2008RG000274, 2009a.

Schumann, G., Di Baldassarre, and G., B., P.D: The utility of Spaceborne Radar to render flood inundation maps based on multialgorithm ensembles, IEEE T. Geosci. Remotes, 47, 2801–2807, 2009b.

Schumann, G., Lunt, D. J., Valdes, P. J., de Jeu, R. A. M., Scipal, K., and Bates, P. D.: Assessment of soil moisture fields from imperfect climate models with uncertain satellite observations, Hydrol. Earth Syst. Sci., 13, 1545-1553, 2009c.

Shafii, M., and De Smedt, F.: Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm, Hydrol. Earth Syst. Sci., 13, 2137-2149, 2009.

Schaefli, B. Gupta H.: Do Nash values have value? Hydrological Processes 21*(*15*):* 2075-2080, 2007

Singh, V. P., Wang, S. X., and Zhang, L.: Frequency analysis of nonidentically distributed hydrologic flood data, Journal of Hydrology, 307, 175-195, 2005.

Skahill, B. E., and Doherty, J.: Efficient accommodation of local minima in watershed model calibration Journal of Hydrology, 329, 122-139, 2006.

Solomatine, D. P., Dibike, Y. B., and Kukuric, N.: Automatic calibration of groundwater models using global optimization techniques, Hydrological Sciences Journal - Journal des Sciences Hydrologiques, 44, 879-894, 1999.

Srinivas, N., and Deb, K.: Multi-Objective function optimization using non-dominated sorting genetic algorithms, Evolutionary Computation, 2(3), 221–248, 1995.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: Handbook of Hydrology, edited by: Maidment, D. R., McGraw-Hill, New York, 18.11 - 18.66, 1993.

Strupczewski, W., Singh, V., and Feluch, W.: Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation, Journal of Hydrology, 248, 123-142, 10.1016/s0022-1694(01)00397-3, 2001.

Tang, Y., Reed, P., and Wagener, T.: How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?, Hydrology and Earth System Sciences, 10, 289-307, 2006.

Tang, Y., Reed, P. M., and Kollat, J. B.: Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications, Advances in Water Resources, 30, 335-353, DOI 10.1016/j.advwatres.2006.06.006, 2007.

Thuy, D. T., and Dac, N. T.: Modelling of flow and salinity in the Mekong Delta by SAL99 Model., Mekong River Commission, Phnom Penh, 225–235, 2000.

Tolson, B. A., and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, 2007.

Townsend, P. A., and Foster, J.R.: A synthetic aperture radar–based model to assess historical changes in lowland floodplain hydroperiod, Water Resour. Res., 38, doi:10.1029/2001WR001046. , 2002.

Vorogushyn, S., Merz, B., Lindenschmidt, K.-E., and Apel, H.: A new methodology for flood hazard assessment considering dike breaches, Water Resources Research, in press, doi:10.1029/2009WR008475, 2010.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resources Research, 39, 1214, doi:1210.1029/2002WR001746, 2003a.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research, 39, 1201, doi:1210.1029/2002WR001642, 2003b.

Wassmann, R., Hien, N. X., Hoanh, C. T., and Tuong, T. P.: Sea Level Rise Affecting the Vietnamese Mekong Delta: Water Elevation in the Flood Season and Implications For Rice Production, in: Climate Change, Kluwer Academic Publishers, 89-107, 2004.

Weise, T.: Global Optimization Algorithms – Theory and Application, 2009.

Werner, M. G. F.: Impact of grid size in GIS based flood extent mapping using a 1D flow model, Physics and Chemistry of the Earth Part B: Hydrology Oceans and Atmosphere, 26, 517-522, 2001.

Werner, M. G. F.: A comparison of flood extent modelling approaches through constraining uncertainties on gauge data, Hydrology and Earth System Sciences, 8, 1141-1152, 2004.

Werner, M. G. F., Hunter, N. M., and Bates, P. D.: Identifiability of distributed floodplain roughness values in flood extent estimation, Journal of Hydrology, 314, 139-157, 2005.

Wilks, D. S.: Statistical methods in the atmospheric sciences, Elsevier Inc., USA, 649-649 pp., 2006.

Wilson, C., Yagci, O., Rauch, H., and Olsen, N.: 3D numerical modelling of a willow vegetated river/floodplain system, Journal of Hydrology, 327, 13-21, 10.1016/j.jhydrol.2005.11.027, 2006.

Wilson, M., P. Bates, D. Alsdorf, B. Forsberg, M. Horritt, J. Melack, F. Frappart, and Famiglietti, J.: Modeling large-scale inundation of Amazonian seasonally flooded wetlands, GEOPHYSICAL RESEARCH LETTERS, 34, doi:10.1029/2007GL030156. , 2007.

Wikipeadia: Likelihood function, online available at http://en.wikipedia.org, 2011

Wolpert, D., and Macready, W. G.: No Free Lunch Theorems for Optimization, IEEE Transactions on Evolutionary Computation, 1, 67-82, 1997.

Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, Journal of Hydrology, 204, 83-97, 1998.

Yoon, T. H., and Kang, S. K.: Finite Volume Model for Two-Dimensional Shallow Water Flows on Unstructured Grids, Journal of Hydraulic Engineering, ASCE, 130, 678-688, 2004.

Yue, S.: Applying the bivariate normal distribution to flood frequency analysis, Water International, 24, 248-252, 1999.

Yue, S., Ouarda, T. B. M. J., Bobee, B., Legendre, P., and Bruneau, P.: The Gumbel mixed model for flood frequency analysis, Journal of Hydrology, 226, 88-100, 1999.

Yue, S.: The bivariate lognormal distribution to model a multivariate flood episode, Hydrological Processes, 14, 2575-2588, 2000.

Yue, S.: A bivariate gamma distribution for use in multivariate flood frequency analysis, Hydrological Processes, 15, 1033-1045, DOI: 10.1002/hyp.259, 2001.

Zitzler E., and Deb K.: Comparison of multiobjective evolutionary algorithms: empirical results, Evolutionary Computation, 8, 125-148, 2000.

Curriculum Vitae

Name:	Nguyen Viet Dung
Day of birth:	17.03.1977
Place of birth:	Thanh Hoa, Vietnam
Marital Status:	Married

1991-1994	Pupil at the Bac Lieu Senior High School, Bac Lieu province, Vietnam.
1994-1998	Bachelor of Mathematics at the University of Natural Sciences (UNS), Ho Chi Minh city, Vietnam
1999-2002	Master of Mathematics at the University of Natural Sciences (UNS), Ho Chi Minh city, Vietnam
1998-2003	Assistant Teacher of Mathematics at the University of Natural Sciences (UNS), Ho Chi Minh city, Vietnam
2003-present	Researcher at Southern Institute of Water Resources Research (SIWRR), Vietnam
2007-present	Researcher at German Research Center for Geosciences (GFZ), Potsdam, Germany PhD student of ENWAT International Doctoral Program provided by the University of Stuttgart

Publications: (within the Doctoral study period)

Dung, N. V., Merz, B., Bárdossy, A., Thang, T. D., and Apel, H.: Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data, Hydrol. Earth Syst. Sci., 15, 1339-1354, doi:10.5194/hess-15-1339-2011, 2011.

Dung, N.V., Delgado, J. M., Merz, B., Apel, H.: Changing flood hazard in the Mekong Delta**.** European Geosciences Union, Vienna, Austria, 2011 (poster)

Dung, N. V., Apel, H., and Merz, B.: Integration of remote sensing and in-su data in automatic calibration of hydraulic models. A case study in the Mekong Delta, European Geosciences Union, Vienna, Austria, 2010. (talk)

Apel H., Garschagen, M., Dung, N. V., Delgado, J. M., Zimmer J.: Climate change impact on flood hydrology and vulnerabilities in the Mekong Delta. Hydrology Conference, San Diego, California, USA, 2010 (poster)

Dung, N. V., Apel, H., and Merz, B.: Floods in the Mekong Delta: a future perspective, European Geosciences Union, Austria, 19 – 24. April 2009. (poster)

Dung, N. V.: Hydrodynamic modeling of floods in the Mekong Delta on different scales, IWRM/WISDOM, Modeling-Workshop, Bochum, 2008. (talk)

Institut für Wasserbau
Universität Stuttgart

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

Verzeichnis der Mitteilungshefte

1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und
Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963

2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964

3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächentragwerke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Randwert- und Lastwertmatrizen*, 1965

4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

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

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

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

8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968

9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968

10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968

11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969

12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969

13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969

14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970

15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970

16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit*, 1970

17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970

18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, vergriffen

19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, vergriffen

20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971

21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972

22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972

23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972

24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, vergriffen

25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, vergriffen

26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972

28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschiffahrt durch Schwimmerhebewerke*, 1973

29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973

30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973

31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973

32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974

34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975

35 *Sonderheft anläßlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von*: Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975

36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975

37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976

38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976

39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6

40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X

41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, vergriffen

42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6

43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4

44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2

45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0

46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9

47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7

48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5

49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, vergriffen

50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, vergriffen

51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5

52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, vergriffen

53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, vergriffen

54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X

55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, vergriffen

56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6

57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, vergriffen

58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2

59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, vergriffen

60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, vergriffen

61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebenen Grundwassermodellen*, 1986, ISBN 3-921694-62-0

63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9

64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7

65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5

66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3

67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1

68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, vergriffen

69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, vergriffen

70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1

71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*,1991, ISBN 3-921694-72-8

73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6

74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, vergriffen

75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2

76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, vergriffen

78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7

79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen

80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen

81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern,* 1993, ISBN 3-921694-81-7

82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versucheinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5

83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen

84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1

85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, vergriffen

86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, vergriffen

87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation* -, 1996, ISBN 3-921694-87-6

88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4

89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, vergriffen

90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, vergriffen

91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung*-, 1998, ISBN 3-921694-91-4, vergriffen

92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien,* 1997, ISBN 3-921694-92-2

93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, vergriffen

94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen 1997*, ISBN 3-921694-94-9

95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, vergriffen

96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5

97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, ISBN 3-921694-97-3

98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, vergriffen

99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden,* 2000, ISBN 3-933761-02-2

100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4

101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, vergriffen

102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0

103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9

104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7

105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien,* 2001, ISBN 3-933761-08-5

106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3

107 Josef, Reinhold:, *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5

109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3

110 Lang, Stefan: *Parallele numerische Simulation instätionärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1

111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X

112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten,* 2002, ISBN 3-933761-15-8

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002, ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Upscaling, Mass-Conservative Discretization and Efficient Multilevel Solution,* 2004, ISBN 3-933761-33-6

131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004,* Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4

132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2

133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers,* 2004, ISBN 3-933761-36-0

134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface,* 2005, ISBN 3-933761-37-9

135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7

136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers,* 2005, ISBN: 3-933761-39-5

137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments,* 2005, ISBN 3-933761-40-9

138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7

139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5

140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media,* 2005, ISBN 3-933761-43-3

141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen,* 2005, ISBN 3-933761-44-1

142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X

143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser,* 2005, ISBN 3-933761-46-8

144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, 3-933761-47-6

145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden,* Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4

146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes,* 2006, ISBN 3-933761-49-2

147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebecken-systemen,* 2006, ISBN 3-933761-50-6

148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4

149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2

150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS-Statuskolloquium 2006,* Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0

151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9

152 Fischer, Markus: *Beanspruchung eingeerdeter Rohrleitungen infolge Austrocknung bindiger Böden,* 2006, ISBN 3-933761-55-7

153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5

154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models,* 2006, ISBN 3-933761-57-3

155 Bielinski, Andreas: *Numerical Simulation of CO_2 sequestration in geological formations*, 2007, ISBN 3-933761-58-1

156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3

157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation,* 2007, ISBN 3-933761-61-1

158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines,* 2007, ISBN 3-933761-62-X

159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8

160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, 2007,* ISBN 3-933761-64-6

161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4

162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2

163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0

164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9

165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS – Kolloquium 2007,* Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7

166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9

168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7

169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling,* 2008*,* ISBN 3-933761-73-5

170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3

171 Papafotiou, *Alexandros: Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1

172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7

173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4

174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1

175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8

176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach,* 2008, ISBN 978-3-933761-80-4

177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1

178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen,* 2009, ISBN 978-3-933761-82-8

179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa,* 2009, ISBN 978-3-933761-83-5

180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding,* 2009, ISBN 978-3-933761-84-2

181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers,* 2009, ISBN 978-3-933761-85-9

182 Kopp, Andreas: *Evaluation of CO_2 Injection Processes in Geological Formations for Site Screening,* 2009, ISBN 978-3-933761-86-6

183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on CO_2 and water (two-phase) flow in porous media,* 2009, ISBN 978-3-933761-87-3

184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0

185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7

186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3

187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0

188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7

189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4

190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8

192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5

193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2

194 Bliefernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins,* 2011, ISBN 978-3-933761-98-9

195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -,* Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2010, ISBN 978-3-933761-99-6

196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia,* 2010, ISBN 978-3-942036-00-9

197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen,* 2010, ISBN 978-3-942036-01-6

198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins,* 2010, ISBN 978-3-942036-02-3

199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow,* 2011, ISBN 978-3-942036-03-0

200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation,* 2011, ISBN 978-3-942036-04-7

201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas,* 2011, ISBN 978-3-942036-05-4

202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen,* 2011, ISBN 978-3-942036-06-1

203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks,* 2011, ISBN 978-3-942036-07-8

204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology,* 2011, ISBN 978-3-942036-08-5

205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken,* Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2

206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour,* 2011, ISBN 978-3-942036-10-8

207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta,* 2011, ISBN 978-3-942036-11-5

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