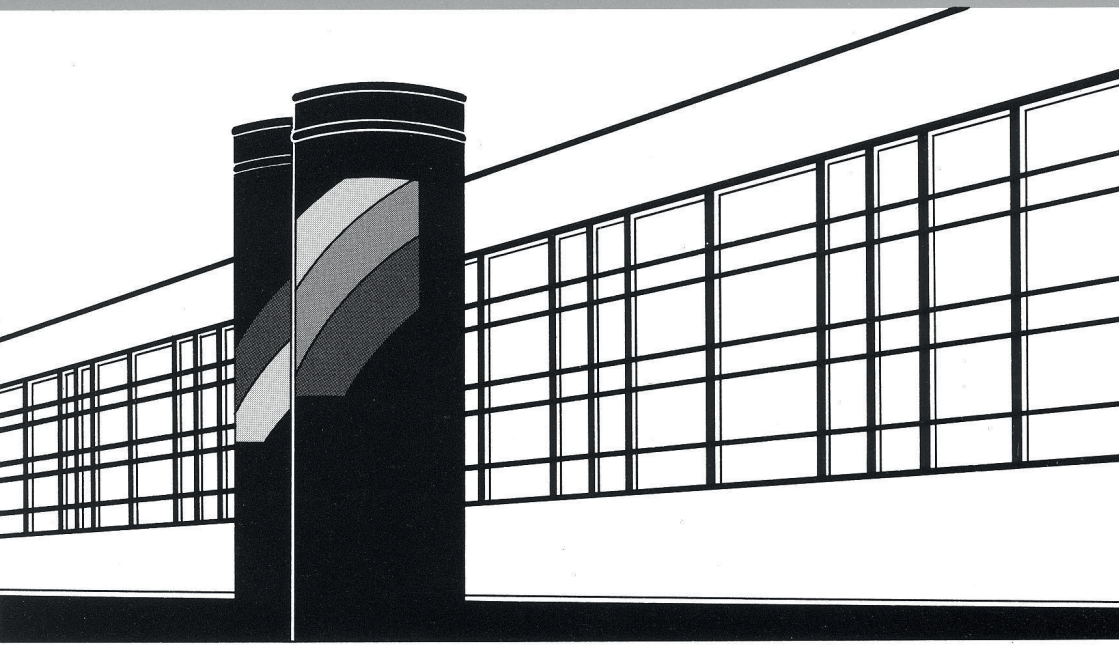
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 211 Habtamu Gezahegn Tolossa

Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach

Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach

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

Vorgelegt von

Habtamu Gezahegn Tolossa

aus Arba Minch, Äthiopien

Hauptberichter:	Prof. Dr.-Ing. Silke Wieprecht
Mitberichter:	Prof. Dr. Chih Ted Yang

Tag der mündlichen Prüfung: 11. Januar 2012

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2012

Heft 211

Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach

von
Dr.-Ing.
Habtamu Gezahegn Tolossa

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

D93 Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach

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

Tolossa, Habtamu Gezahegn:
Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach / von Habtamu Gezahegn Tolossa. Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2012

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: Heft 211)
Zugl.: Stuttgart, Univ., Diss., 2012
ISBN 978-3-942036-15-3
NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen

Herausgegeben 2012 vom Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung

Druck: Document Center S. Kästl, Ostfildern

Acknowledgements

It has been one of the most exiting challenges of my academic carrier to write and finalize this dissertation. Many people contributed to make this work successful. It is a pleasure to convey my gratitude and appreciation to them all.

First of all, I would like to express my deepest gratitude to my advisor Prof. Silke Wieprecht for giving me the opportunity to conduct this work under her supervision. Her guidance, support, and encouragement enabled me to finish the dissertation successfully. I would like to thank my co-advisor Prof. Chih Ted Yang for offering his invaluable advice and comments, and for providing a nice working environment during my research visit to Colorado State University. I am also grateful to Prof. Manfred Joswig, the chairman of my examination committee, for his time and effort.

I would like to take this opportunity to forward my earnest gratitude and appreciation to Dr. Matthias Schneider for his guidance, help and friendship throughout my stay in Germany. He provided several valuable comments that helped a lot to improve the quality of the work. Thanks also go out to Jeffrey Tuhtan for being a wonderful friend and colleague, and for his help in so many aspects. I enjoyed all our good-spirited discussions and debates concerning a variety of issues.

I am extremely grateful to Dr. Gabriele Hartmann, the course director of ENWAT program, for her consistent assistance and cooperation from the beginning till the end of my doctoral study. I would like to thank the German Federal Ministry of Education and Research (BMBF) and IPSWAT scholarship program for providing financial support to carry out this research.

I am thankful to Dr. Stefan Vollmer from the German Federal Institute of Hydrology (BfG) for providing the comprehensive field datasets used in the dissertation.

Many thanks go out to all my friends who made my stay in Stuttgart very pleasant and comfortable. I am very grateful for the friendship and support of all of the members of Institut für Wasser- und Umweltsystemmodellierung especially those of Abteilung 1.

A very special thank you to my wife Beti! Her encouragement, understanding and unwavering love were in the end what made this possible. Finally, I am deeply and forever indebted to my parents for their love and support throughout my entire life. I owe them all that I am right now. Thank you for everything!

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Definition
AARE	Average absolute relative error
ARE	Absolute relative error
B	Channel width
C_a	The reference concentration at depth =a (van Rijn,1984a)
C_b	Bed load concentration
C_t	Total sediment concentration by weight per volume in ppm
C_{tg}	Total gravel concentration in ppm by weight (Yang, 1984)
d	Sediment particle size
D	Water depth
d_{50}	Median sediment particle size
d_{90}	Size of the sediment for which 90% of the material is finer
d_{gr}	Dimensionless grain diameter
d_m	Mean particle diameter (Meyer-Peter and Müller (1948)
D_r	Discrepancy ratio
e_b	Efficiency coefficient for bed load transport (Bagnold,1966)
e_s	Suspended load transport efficiency (Bagnold,1966)
F	Shape coefficient
f'	Friction coefficient (Engelund and Hansen,1972)
F_{gr}	Sediment mobility number
g	Acceleration due t gravity
g_b	Bed load rate in weight per unit time and width
K_r	Coefficient of particle roughness
K_s	Sticklers' coefficient of bed roughness
N	Total number of observations
P	Probability for motion (Einstein, 1950)
p_i	Percent of sediment with particle diameter i
q_b	Bed load transport by weight per unit time and width

q_s	Suspended load transport by weight per unit width
q_t	Total bed-material sediment discharge by weight per unit width
r	Correlation coefficient
R	Hydraulic radius
$R_b^{'}$	Hydraulic radius due to bed form resistance
RE	Relative error in percentage
RMSE	Root mean squared error
s	Specific gravity of sediment
S	Bed or energy slope
$\bar{S}_c$	Mean computed value
S_{ci}	Computed sediment load for ith observation
S_m	Sediment transport capacity of flow
$\bar{S}_o$	Mean observed value
S_{oi}	Observed sediment load for ith data
T	Transport stage parameter
u_*	Bed shear velocity
$u_*^{'}$	Bed shear velocity related to grains
$u_{*,c}$	Critical bed shear velocity
u_{bs}	Particle velocity
V	Flow velocity
V_c	Critical velocity
$W_s^{'}$	The rate of energy available for transporting suspended load
W_i^*	Parker's dimensionless bed load transport function
W_s	The rate of work needed to transport suspended load
γ_s	Specific weight of sediment
γ_w	Specific weight of water
δ_b	Saltation height
θ	Dimensionless shear stress
$\mu_A(x)$	The membership function of the fuzzy set A
ν	Kinematic viscosity
ρ_s	Density of sediment
ρ_w	Density of water
τ	Bed shear stress
$\tau^{'}$	Bed shear stress caused by grain resistance

τ_c	Critical bed shear stress
ϕ	Dimensionless sediment discharge
Ψ	Dimensionless flow intensity
ω	Fall velocity of sediment particles
Φ	Dimensionless bed load intensity

List of Abbreviations

Abbreviation	Meaning
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
BfG	Bundesanstalt für Gewässerkunde
FCM	Fuzzy c-means clustering
FIS	Fuzzy inference system
MF	Membership function
TS	Takagi-Sugeno inference system

Abstract

Reliable approaches for the computation of sediment transport rates in natural rivers are vital for sustainable utilization and management of water resources. They are especially essential for prediction of erosion, sedimentation and related river morphological changes, design of flood control structures. During the last decades, many researchers have focused on analysing the process of sediment transport and put forward a plethora of sediment transport equations to choose from. Estimation of amount and mode of sediment under transport is challenging and is usually accomplished by utilizing empirical or semi-empirical equations. Prediction errors of the existing equations are usually very high for practical applications, and show significant discrepancy from observed transport rates. There is a very high degree of uncertainty and fuzziness associated with the results from different equations. Recently, the utilization of data-driven fuzzy modelling, which is especially attractive for modelling complex processes about which the physics governing them is too complex to be represented in mathematical equations, has become more popular. Sediment transport modelling is one of the focus areas in this regard.

This research focuses on assessing applicability of data-driven adaptive neuro-fuzzy based modelling approach for estimating sediment transport rates. A general fuzzy model has four components: fuzzifier, fuzzy rule base, fuzzy inference, and defuzzifier. Four dominant parameters affecting sediment transport capacity are selected and used for constructing the data-driven fuzzy model, using comprehensive sets of laboratory and field data. These parameters are depth, flow velocity, bed or energy slope and median size of sediment particles. The laboratory data is categorized into sand and gravel, and the river data includes measured input parameters and bed load and total bed-material load transport rates for the Rhine and Elbe Rivers.

The laboratory and river datasets are treated separately because the ranges of the input and output parameters are significantly different. A methodology for selecting training and test datasets is recommended. Two third of the available datasets are used for training and one third for testing. The Takagi-Sugeno fuzzy inference system is selected because it is the most suitable method for generating fuzzy rules from measured data, the output is crisp and adaptive techniques can be used for model optimization. The initial fuzzy model is obtained by grid partitioning of the input variables, and fuzzy clustering of input and output

variables. The optimisation of the model is performed by data-driven tuning of the fuzzy model parameters using the adaptive neuro-fuzzy inference system (ANFIS) so that the model output is able to reproduce the measured total bed-material and bed load transport rates. The antecedent and premise parameters of the fuzzy model are adjusted during model training with ANFIS. A sensitivity analysis for the combination of input parameters, and number and type of membership functions for each input parameter is also performed to determine the significance of the parameters on the accuracy of the developed model. The results of the sensitivity analysis are implemented to optimize the structure of the final rule base, which results in a compact and optimum rule system. Further refining of the developed fuzzy models is performed for the Rhine and Elbe by dividing the rivers into sections based on river morphology and relevant input parameters. A detailed comparison of the results of the fuzzy rule based model with the results of other commonly utilized sediment transport functions is also performed. Selection of the sediment transport equations is done based on the ranges of the input parameters and recommended applicability of the equations. Validation of the model developed for the Rhine river is done by implementing the ANFIS model for estimating erosion and deposition rates, and computing average annual bed level change for a section of the Upper Rhine.

The results of the investigation show that the data-driven adaptive neuro-fuzzy modelling approach can be a powerful alternative technique for correctly estimating both bed and total bed-material transport rates. Three generalized-bell shaped membership functions for each of the input parameters are found to be the most efficient with respect to accuracy and model complexity. The fuzzy model performed better than the selected explicit transport equations both for the laboratory and field datasets. The analysis of the model outputs shows that large computation errors are associated with low sediment transport rates which are not significant for river morphology changes. In the case of laboratory data, acceptable accuracy is obtained by using three input variables (velocity, slope, and median particle size) in the fuzzy model. However, water depth should be included for rivers like the Rhine and Elbe. Total load is estimated with better accuracy than bed load, this is also the case for most of sediment transport equations. Overtraining may occur during model optimization and care should be taken to avoid overtraining by checking the performance of the fuzzy model on the test data. Collection of sufficient and good quality training data which represent the properties of the process to be modelled is a necessary precondition for a successful application of data-driven modelling. The limitation of the approach is that its accuracy purely depends on the quality and quantity of input data. The generalization capacity and transferability of the developed models to other reaches with comparable hydraulic and morphologic characteristics should be assessed carefully.

Kurzfassung

Verlässliche Ansätze zur Berechnung der Sedimentbewegungen in Fließgewässern sind ein wichtiger Aspekt des nachhaltigen Gewässermanagements. Sie werden insbesondere benötigt zur Voraussage von Erosion, Sedimentation und damit verbundenen morphologischen Veränderungen, wie z. B. für die Planung von Hochwasserschutzmaßnahmen. In den letzten Jahrzehnten haben sich viele wissenschaftliche Arbeiten mit den Prozessen des Sedimenttransports beschäftigt und es existiert eine Vielzahl von Berechnungsformeln und -gleichungen. Dennoch – oder auch deshalb – ist die Berechnung von Sedimenttransportraten immer noch eine Herausforderung und es werden vor allem empirische oder semiempirische Ansätze verwendet. Die Fehler von Voraussagen auf Basis der existierenden Ansätze sind normalerweise groß und zeigen starke Abweichungen von den in der Natur beobachteten Transportraten, so dass ihre Anwendung in der Praxis für das Gewässermanagement zu unsicher ist. Die Differenzen zwischen den Ergebnissen der unterschiedlichen Berechnungsgleichungen sind groß, so dass ohne eine ausreichende Kalibrierung eine hohe Unsicherheit in Kauf genommen werden muss. In jüngerer Zeit werden datenbasierte, fuzzylogische Ansätze vermehrt eingesetzt, die besonders geeignet sind, Gesamtprozesse in Modellen zu abzubilden, deren physikalische Zusammenhänge zu komplex sind, um sie durch physikalisch basierte Gleichungen zu beschreiben. Der Sedimenttransport zählt zu dieser Art sehr komplexer Gesamtprozesse.

Die vorliegende Arbeit beschäftigt sich mit der Anwendbarkeit von datenbasierten, adaptiven Neuro-Fuzzy-Modellansätzen für die Ermittlung von Sedimenttransportraten. Prinzipiell umfasst ein Fuzzy-Modell vier Hauptkomponenten: die Fuzzifizierung, das Fuzzy-Regelwerk, das Fuzzy-Inferenz System und die Defuzzifizierung. In dieser Arbeit werden vier dominante Parameter identifiziert, die den Sedimenttransport maßgeblich bestimmen. Diese werden verwendet, um auf Grundlage umfangreicher Daten aus Labor- und Felduntersuchungen das datenbasierte Fuzzymodell zu entwickeln. Die Parameter sind Wassertiefe, Fließgeschwindigkeit, Sohlgefälle und der Medianwert der Sedimentkornverteilung. Die Labordatendatensätze werden unterteilt in die Kategorien Sand und Kies. Die vorliegenden Naturdaten beinhalten Messdaten für die genannten Parameter, sowie die Geschiebefracht und die Gesamtsedimentfracht für zahlreiche Messstationen an den Flüssen Rhein und Elbe.

Für die Untersuchungen werden sowohl Labor- als auch Naturmessdaten verwendet, aber getrennt behandelt, um den unterschiedlichen Parameterbereichen gerecht zu werden. Die Daten werden im Verhältnis zwei Drittel zu einem Drittel jeweils in Tranings- und Testdatensätze unterteilt. Das Takagi-Sugeno Fuzzy Inferenz System wird eingesetzt, da es sich in zahlreichen Untersuchungen als sehr geeignet für die Erzeugung von Fuzzy-Regelwerken auf der Basis von Messdaten erwiesen hat. Das Initialsystem wird erhalten durch die Methode des Grid Partitioning, das auf die Eingangsvariablen angewendet wird und durch Fuzzy-Clustering der Eingangs- und Ergebnisvariablen. Die Modelloptimierung erfolgt über eine datenbasierte Anpassung der fuzzy-logischen Modellparameter unter Verwendung eines adaptiven neuro-fuzzy Inference Systems (ANFIS), Ziel ist eine weitestgehende Annäherung an die gemessenen Transportraten für den Geschiebetransport und die Gesamtsedimentfracht.

Die in den Regelprämissen enthaltenen Parameter werden durch das ANFIS während des Modelltrainings angepasst. Gleichzeitig wird eine Sensitivitätsanalyse durchgeführt unter Berücksichtigung der Kombination, der Anzahl und des Typs der verwendeten Fuzzymengen für alle Eingangsparameter. Die Ergebnisse der Sensitivitätsanalyse dienen zur Optimierung der Regelbasis, mit dem Ziel ein vereinfachtes und bestangepasstes Regelwerk zu erstellen. Eine weitere Verfeinerung des Modells wird erreicht durch eine Aufteilung der untersuchten Gewässer in Abschnitte unter Berücksichtigung der unterschiedlichen Morphologie und der Eingangsparameter. Zusätzlich wird ein Vergleich der Ergebnisse aus dem auf Fuzzy-Regeln basierten Modell mit denen aus gängigen Transportgleichungen angestellt. Diese Transportgleichungen werden unter Beachtung der für ihre Anwendung empfohlenen Randbedingungen ausgewählt. Eine Validierung des Modells für den Rhein wird durchgeführt über den Vergleich der jahresgemittelten Modellergebnisse für Erosions- und Sedimentationsraten mit den jeweils gemessenen Sohlveränderungen.

Die Untersuchungsergebnisse belegen, dass der datenbasierte adaptive Neuro-Fuzzy-Ansatz eine geeignete Alternative sein kann für die Abschätzung des Geschiebetransports und der Gesamtsedimentfracht. Drei verallgemeinerte glockenförmige Zugehörigkeitsfunkionen für alle Eingangsparameter erweisen sich im Hinblick auf die Voraussagegenauigkeit und die Modellstruktur als am geeignetsten. Das Fuzzy Modell erzielt bessere Ergebnisse als die ausgewählten Transportgleichungen, sowohl für die Labor- als auch die Naturdaten. Die Betrachtung der Modellergebnisse zeigt, dass große Fehler vor allem bei geringen Sedimenttransportraten auftreten, welche jedoch nicht maßgebend für die Veränderungen der Flussmorphologie sind.

Für die Labordaten werden mit dem Fuzzy-Modell ausreichend gute Ergebnisse bei Verwendung von nur drei Eingangsvariablen erzielt (Fließgeschwindigkeit, Gefälle und

mittlere Korngröße). Bei Gewässern wie der Elbe und dem Rhein sollte allerdings die Wassertiefe las viertem Parameter mitberücksichtigt werden. Die Ergebnisse für die Gesamtsedimentfracht sind, wie auch bei den meisten gängigen Transportgleichungen, verlässlicher als die für die Geschiebefracht. Während der Modelloptimierung kann eine Überkalibrierung (Overtraining) stattfinden. Dies sollte durch die Überprüfung der Ergebnisse für den Testdatensatz vermieden werden. Voraussetzung für die erfolgreiche Anwendung des datenbasierten Ansatzes ist die Auswahl eines geeigneten und ausreichend umfangreichen Datensatzes, der die abzubildenden Prozesse repräsentativ widerspiegelt. Die Abhängigkeit des Ansatzes vom Umfang und der Qualität der Eingangsdaten ist gleichzeitig seine größte Einschränkung. Die Übertragbarkeit der entwickelten Modelle auf andere Gewässer mit vergleichbarer Hydraulik und Morphologie sollte deshalb äußerst vorsichtig beurteilt werden.

1 Introduction

1.1 Background

For many years investigators have tried to estimate the rate of sediment transport in natural rivers. Quantifying the amount and type of sediment particles transported by rivers is necessary for a successful implementation of sustainable projects in hydraulic engineering and water resources management. Prediction of long-term channel morphology changes, the useful operation life of reservoirs, maintaining efficient navigation channels, flood protection, fish habitat modelling, river aesthetics, etc. all require sediment transport modelling. There are many equations (over a hundred) which have been provided by various researchers to estimate the rate of sediment transport but the accuracy of these equations is not satisfactory and is mainly restricted to specific river characteristics in terms of slope, grain size or equilibrium state of the river section. The error ranges up to two to three orders of magnitude and even more. There is not any equation which is universally accepted and applicable under varying site conditions (Yang, 1996). Because of the very complex nature of the process of sediment transport and the different factors playing a role in determining the amount of transported material, sediment transport modelling is one of the challenging tasks facing engineers. Many researches are underway to improve the existing equations and get a better understanding of the process. There is a huge uncertainty and fuzziness associated with the prediction of sediment transport.

In recent times, fuzzy logic based modelling has gained significant attention for solving many engineering problems. The concept of fuzzy logic has been introduced by Zadeh in 1965, and as its name implies it is usually applied for modelling complex processes about which the physics governing them is not well understood. Fuzzy means uncertain, not clear, or vague. Based on this definition, sediment transport modelling is considered to be fuzzy so far, and many processes in hydraulic engineering are. Fuzzy models have four key components. These are fuzzification interface, fuzzy rule base, fuzzy inference, and defuzzification interface. In fuzzy models, a process is described using a natural language, and is close to the way human beings communicate. The process is usually represented qualitatively using IF-THEN statements.

There are two ways for generating the fuzzy rules describing the system to be modelled. If sufficient expert knowledge exists about the processes and their interaction, the rules can be defined by experts. But if the processes are too complicated and the rules governing the process are not known with sufficient quality, measured data can be used to generate the rules. Data-driven modelling is used to identify hidden relationships that exist in datasets by detailed analysis of the data and by utilizing different optimization algorithms. Data-driven modelling has been applied in many fields in water resources. The most popular methods of data-driven modelling are fuzzy logic, artificial neural networks, and evolutionary algorithms. In the case of sediment transport modelling, the process is too complicated and the rules cannot be expressed by experts with sufficient quality, therefore the option is to generate the fuzzy rules by using data-driven approaches.

A data-driven fuzzy logic based modelling can also play a key role in identifying dominant parameters deriving the process of sediment transport. It can provide insight toward understanding the importance and physical meaning of the parameters used in many sediment transport formulas by assessing the accuracy of different fuzzy models using the main variables in the sediment transport formulas as primary inputs to the fuzzy model. The corresponding changes in the parameters of membership functions of the resulting fuzzy model, which affect the shapes of final membership functions, can indicate the dominance of the parameters in influencing sediment transport rates.

The importance of carrying out a research program focusing on how to improve the predictive accuracy of sediment transport computation, analyzing the different factors playing a vital role in the process, and assessing ways to simplify the complex process using fuzzy logic based modelling approach is worth investing both time and resources and furnishes the main motivation of this work.

1.2 Research Objectives and Methodology

The main tasks of the research work are subdivided into four specific objectives. These goals and the methodologies implemented to achieve the objectives of this dissertation are summarized below.

1st **Objective**

Overview of existing sediment transport modelling approaches, selection of dominant parameters affecting sediment transport, and development of an initial fuzzy model by using a data-driven modelling approach.

- ***Methodology***: There are many approaches used for deriving sediment transport functions. The various available equations are analyzed and the main factors affecting

sediment transport rates are selected. The main criteria for the selection of the input parameters are they should be measurable as primary data and should have a clear physical meaning. A comprehensive set of laboratory and field data has been collected from various sources, and the data are carefully analyzed. Grid partitioning and fuzzy clustering of input and output parameters are used for the generation of initial membership functions of the fuzzy model. Additionally, the four basic variables are selected because they can be easily measured as primary data, and their physical meaning

2nd Objective

Optimize the developed data-driven fuzzy models and perform a sensitivity analysis for the model parameters and simplify the models accordingly.

- ***Methodology:*** An adaptive neuro-fuzzy inference system (ANFIS) is used for parameter optimization. Parameter optimization is achieved by membership function fine tuning and rule conclusion optimization. The sensitivity analysis is performed by changing the number and type of membership functions, and the combination of input parameters. A structure optimization of the fuzzy model is done to make it more simple and interpretable by identifying dominant input variables and rule based reduction. The results of the sensitivity analysis are implemented for the optimization of the model structure.

3rd Objective

Analyze the accuracy of the optimized fuzzy models in estimating transport rates and compare their accuracy with that of selected sediment transport functions.

- ***Methodology:*** Appropriate sediment transport functions are selected for comparison based on the ranges of input parameters within the datasets collected from different sources. A detailed comparison of the results of the developed fuzzy models with the results of other commonly utilized sediment transport equations is performed. Average absolute relative error, discrepancy ratio, correlation coefficient, and root mean squared error are selected as statistical model performance evaluation criteria for comparison of the accuracy of the different models.

4th Objective

Validation of the model with other case study reaches and further refining of the optimized fuzzy model, provide guidelines for further application and improvement of the developed model and the modelling approach.

- ***Methodology:*** The Rhine and Elbe river reaches are divided into different categories based on the input parameters like slope, width, and bed-material size, and separate

models are developed for the river reaches which have comparable morphologic characteristics. The model developed for river Rhine is validated by computing erosion and deposition in the upper Rhine and compared with measured data. The ANFIS model is utilized to identify sections of bed degradation and aggradation, estimate the bed level change by balancing the annual sediment load. Provide guidelines for future application of the developed model, for which specific cases the data-driven modelling approach is suitable and the quality and amount of datasets required. The transferability of the optimized fuzzy models to other river reaches should be carefully analyzed with additional datasets.

1.3 Dissertation Structure

This dissertation is compiled in nine chapters.

Chapter two briefly presents the process of sediment transport, the different approaches used for deriving sediment transport functions, a summary of selected transport equations, the dominant factors affecting sediment transport, and the factors to be taken into consideration while selecting sediment transport equations for different river reaches.

Chapter three focuses on describing fuzzy logic based modelling, the components of fuzzy models, types of fuzzy inference systems and techniques used to derive fuzzy rules. The adaptive neuro-fuzzy inference system which is utilized for optimizing the fuzzy models is presented, and also some of the recent applications of fuzzy logic in hydraulic engineering are summarized.

Chapter four provides a summary of the available laboratory and field measured sediment transport data and other hydraulic parameters obtained from different sources, and the methodology for training and test data selection is described.

Chapter five discusses the basic steps followed for deriving data-driven fuzzy models for computing bed load and total bed-material load transport rates using laboratory and field data, and the techniques implemented for model optimization.

Chapter six summarizes the results of various data-driven adaptive neuro-fuzzy models and the sensitivity analysis performed using the available data.

Chapter seven provides a comparison of the results of the models developed in the previous chapters with results of other commonly utilized sediment transport equations using statistical model performance evaluation criteria.

Chapter eight discusses the validation of the fuzzy model developed for river Rhine by computing the rates of erosion and deposition, and annual average bed level change by utilizing datasets obtained from fifteen stations in the Upper Rhine.

Chapter nine presents a summary of the work, the conclusions drawn from this research, recommendations for further application of the fuzzy models developed and the modelling approaches, and outlines interesting points to be considered for further research.

2 Sediment Transport Modelling Approaches

2.1 Introduction

In natural rivers, flow is able to transport sediment particles when it exceeds the critical condition required for the initiation of motion. This critical condition is called incipient motion and the criteria for the incipient motion can be expressed using critical bed shear stress (Shields, 1936), critical velocity, or flow power as characteristic parameters. After this condition is satisfied and the flow is able to exert enough force or energy, sediment particles start moving. The motion of sediment particles can be as bed load or suspended load. Different factors determine whether a sediment particle is transported as bed load or suspended load including particle size, flow velocity, flow depth, etc. Understanding the process of sediment transport in rivers and prediction of the amount of transported sediment is essential for the design of hydraulic structures and a sustainable management of water resources. Sediment transport plays a key role in channel morphology changes, reservoir sedimentation, maintaining navigation channels, design of intake structures for hydropower, habitat modelling, river aesthetic, and environmental impact assessment. One, two, or three dimensional numerical hydraulic models are usually used to simulate the process of sediment transport and the resulting short and long term river morphology changes so that appropriate mitigation measures can be implemented based on the results of the simulation.

Bed Load

If the movement of sediment particles is by rolling, sliding or saltation (regular jumping) along the river bed, it is called bed load transport. Bed load transport is a fundamental physical process in natural rivers (Barry et al., 2008), and plays a significant role for river morphology change and dynamic equilibrium (Goodwin, 2004). The bed load transport occurs until a certain height above the river bed. This thickness of bed load transport layer is proportional to the particle size and is usually determined from the saltation height. Van Rijn (1984a) suggested that the bed load transport layer is usually less than ten particle diameters. The amount of bed load transport depends on the flow condition, and it is usually difficult to measure accurately because defining the bed load transport layer clearly is not trivial (Wang et al., 1998).

Suspended Load

If the fall velocity of a sediment particle is less than the upward turbulent forces applied on the particle, the sediment particle stays in suspension. Suspended load is the portion of the sediment that is supported by the upward components of turbulent eddies and stays in suspension for a significant period of time (Yang, 1996). Detailed analysis of suspended load transport can be referred from (van Rjin, 1984b, 1993; Chien and Wan, 1999; Yang, 1996). The suspended load transport depends on the characteristics of the river basin upstream of the measurement station, and is relatively easy to measure (Syvitski et al., 2000).

Fine sediment particles ranging from clay to fine sand are usually transported in suspension and mostly have less impact on channel morphology changes. Larger sediment particles ranging from coarse sand to cobbles are mainly transported as bed load and affect river morphology change more significantly. There is a continuous exchange of materials between the suspended load and bed load layer, and it is usually difficult to separate the two (Xiaoqing, 2003). The source of sediment particles transported by rivers can be either the drainage basin or particles detached from the river bed (Alekseevskiy et al., 2008).

The sum of the amount of bed load and suspended load is called the *total load* that is under transport for given flow and boundary conditions. Sediment can also be classified as bed-material load and wash load. Bed-material load can move in the form of bed load and suspended load. Wash load is usually fine and mainly moves in suspension (Xiaoqing, 2003). The wash load portion is not dependent on the composition of the river bed-material rather it is a function of the sediment supply rate from upstream, and does not play an active role for river morphology change. The portion of wash load should be subtracted from the total sediment load to determine the morphologically active bed-material load.

Sediment transport modelling and river morphology change analysis are fields which are in the main focus of research in hydraulic engineering, and different approaches are utilized for deriving sediment transport equations.

2.2 Approaches for Deriving Sediment Transport Functions

The basic approaches used in the derivation of sediment transport functions are: regression, probabilistic, and deterministic approaches (Yang, 2006).

Regression Approaches

The regression approach uses a non-linear multiple regression analysis to derive the relationship between sediment transport, and input variables. Equations of Shen and Hung (1972), Karim and Kennedy (1990), Rottner (1959) are derived by regression. Sediment transport equations derived by regression analysis should be applied to conditions within

the range of datasets used for the formulation of the equations. Performing regression analysis requires the collection of a lot of data, and the resulting equation is usually descriptive and not predictive.

Probabilistic Approaches

Einstein (1950) introduced the derivation of sediment transport equations from a probabilistic approach. He observed the stochastic nature of sediment transport and combined probability and statistics with modern fluid mechanics to derive his sediment transport equation. Detailed analysis of Einstein's probabilistic approach can be found in Yang (1996), and Chien and Wan (1999). Equations of Colby (1964) and Toffaleti (1969) are based on Einstein's probabilistic approach.

Deterministic Approaches

The deterministic approach is based on the assumption that sediment transport can be estimated by using one or more dominant parameters (Yang, 2006). The most commonly used parameters are flow velocity, sediment particle diameter, slope, water depth, shear stress, stream power, unit stream power, etc.

Bagnold (1966) introduced the stream power approach for estimating sediment transport based on general physics. He assumed that the rate of dissipation of energy is proportional to the amount of material transported. Bagnold defined stream power as the power per unit bed area which can be used to transport sediment. Stream power is considered to be the product of shear stress (τ) and flow velocity (V).

Sediment transport equations of Engelund and Hansen (1972), and Ackers and White (1973) are derived based on Bagnold's concept of stream power. Yang (1973) derived his sediment transport equation based on the unit stream power approach. He defined unit stream power as the product of velocity and slope.

Equations of DuBoys (1879), Meyer-Peter and Müller (1948), and Laursen (1958) are based on the assumption that the amount of sediment transport is proportional to the excess of bed shear stress. Parker (1990) developed an empirical gravel transport function based on the equal mobility concept and field data. Van Rijn (1984a, b) introduced the computation of sediment transport as the product of the saltation height, particle velocity, and reference bed load concentration.

2.3 Sediment Transport Equations

There are many equations provided by different authors for computing the amount of sediment under transport. Some of the equations compute total bed-material load while

others are used for calculating bed load transport rate. Most of these equations are derived using data from laboratory flumes or limited field data, and the accuracies of the equations show considerable discrepancies from the measured transport rate (Yang, 2006). Estimation of the amount and composition of sediment transport is one of the challenging tasks facing engineers up to date and there is not any equation which is universally acceptable under varying site condition. Some of the commonly utilized equations are briefly summarized in this section. Interested readers can refer to the references mentioned for details of particular equations and approaches.

1. *DuBoys' Equation (1879)*

DuBoys (1879) derived his equation for computing sediment transport based on the assumption that the transport rate is proportional to the excess bed shear stress available to move sediment particles. DuBoys also assumes that sediment particles move in layers which slide along each other and comes up with the following simple equation to estimate bed load discharge by volume per unit width:

$$q_{b,v} = K\tau(\tau - \tau_c) \tag{2.1}$$

where:

$q_{b,v}$ = bed load discharge by volume per unit channel width (ft²/s),

τ = bed shear stress (lb/ft²),

τ_c = critical shear stress along the bed (lb/ft²), which can be obtained from the Shields diagram (Shields, 1936).

The value of K is given to be a function of the sediment particle size and can be computed using $K = \frac{0.173}{d^{3/4}}$ (Straub, 1935), where d = particle size (mm).

2. *Meyer-Peter and Müller's Equation (1948)*

Meyer-Peter and Müller (1948) propose an empirical formula for computing bed load transport in natural rivers. The formula is more accurate for estimating bed load transport rates for rivers carrying coarse sand and gravel (0.4 mm - 30 mm), and with large width-depth ratio (Julien, 1995). The bed load transport rate is assumed to be proportional to the difference between the average shear stress acting on the particles and the critical shear stress required to reach incipient motion, and is computed using:

$$q_b = 8\rho_w g^{1/2}\left[RS\left(\frac{K_s}{K_r}\right)^{3/2} - 0.047(s-1)d_m\right]^{3/2} * \left(\frac{s}{s-1}\right) \tag{2.2}$$

where:

ρ_w = density of water (kg/m^3),

g = the acceleration due to gravity (m/s^2),

R = hydraulic radius (m),

S = energy slope,

s = specific gravity of sediment,

d_m = mean particle diameter (m),

ρ_w = density of water (kg/m^3),

q_b = bed load transport by mass per unit time and width (kg/s.m),

K_s = Strickler's coefficient of bed roughness,

K_r = coefficient of particle roughness = $\dfrac{26}{d_{90}^{1/6}}$,

d_{90} = the size of the sediment for which 90% of the bed-material is finer (m).

3. *Einstein's Bed Load Equation (1950)*

Einstein (1950) suggests that the condition for initiation of motion of sediment particles is difficult to define and should be expressed with probability. According to Einstein, whether a particle moves or remains stationary on the bed, the probability depends on the diameter and the shape of the particle and the flow conditions (Chien and Wan, 1999). He uses a hiding correction factor and lifting correction factors to match results from his theoretical equation based on a probabilistic concept with measured laboratory data. Applying probability theory, Einstein derives a mathematical expression for the relationship between the bed load transport intensity ϕ and the flow parameter ψ (Xiaoqing, 2003):

$$P = 1 - \frac{1}{\sqrt{\pi}} \int_{-B_*\psi^{-1/\eta_0}}^{B_*\psi^{-1/\eta_0}} e^{-t^2} dt = \frac{A_*\phi}{1 + A_*\phi} \tag{2.3}$$

$$\phi = \frac{q_b}{\gamma_s} \left(\frac{\gamma_w}{g\left(\gamma_s - \gamma_w\right) d^3} \right)^{1/2} \tag{2.4}$$

$$\psi = \frac{\gamma_s - \gamma_w}{\gamma_w} \frac{d}{R_b^{'} S} \tag{2.5}$$

The constants A_*, B_* and η_0 which are determined through experiment are:

$$\eta_0 = 0.5\,,\ A_* = \frac{1}{0.023} = 43.5\ ,\ B_* = \frac{1}{7} = 0.143 \tag{2.6}$$

where:

P = the probability for motion,

ϕ = dimensionless bed load intensity,

ψ = dimensionless flow parameter,

q_b = the bed load rate in weight per unit of time and width (kg/s.m),

d = particle diameter (m),

R'_b = hydraulic radius due to bed form resistance (m),

γ_s = specific weight of sediment (kg/m³),

γ_w = specific weights of water (kg/m³),

S = bed slope.

A detailed analysis of Einstein's approach is presented in Chien and Wan (1999) and Yang (1996). His approach is the most comprehensive method that has been developed for sediment transport calculation from the theoretical point of view. But because of its complex computational procedure, it is not popular among engineers for practical application (Yang, 2006).

4. *Laursen's Equation (1958)*

Laursen's (1958) formula is used to compute total bed-material load by using mean hydraulic characteristics (velocity, depth and energy gradient) and sediment characteristics defined by grain size and fall velocity. The equation is expressed in dimensionally homogeneous form by American Society of Civil Engineers Task Committee (1971) as:

$$C_t = 0.01\gamma_w \sum_i p_i \left(\frac{d_i}{D}\right)^{7/6} \left(\frac{\tau'}{\tau_c} - 1\right) f\left(\frac{u_*}{\omega_i}\right) \tag{2.7}$$

where:

C_t = total bed-material concentration by weight per unit volume (ppm),

$u_* = \sqrt{gDS}$ = bed shear velocity,

p_i = percentage of material in size fraction i,

ω_i = fall velocity of particles of mean size d_i,

D = average water depth,

τ_c = critical shear stress for sediment size d_i from Shield's diagram,

τ' = Laursen's bed shear stress caused by grain resistance = $\frac{\rho V^2}{58}\left(\frac{d_{50}}{D}\right)^{1/3}$.

The function $f(\frac{u_*}{\omega_i})$ which is proportional to the ratio of shear velocity and particle fall velocity is given in a graphical form by Laursen (1958). The applicability of the equation is in the sand size range and with bed slope less than 0.025.

5. *Rottner's Bed Load Equation (1959)*

Rottner's bed load formula is one of the many equations based on regression analysis of laboratory data. Rottner (1959) formulates his equation for bed load transport in terms of the flow parameters and empirical coefficients (Yang, 2006). The dimensionally homogeneous Rottner's bed load equation is expressed as:

$$q_b = \gamma_s \left[(s-1) g D^3 \right]^{1/2} \left[\frac{V}{\sqrt{(s-1)gD}} \left[0.667 \left(\frac{d_{50}}{D} \right)^{2/3} + 0.14 \right] - 0.778 \left(\frac{d_{50}}{D} \right)^{2/3} \right]^3 \tag{2.8}$$

where:

q_b = bed load transport rate per width (lb/s.ft),

γ_s = specific weight of sediment (lb/ft^3),

s = specific gravity of sediment,

g = acceleration due to gravity in (ft/s^2),

D = mean depth in (ft),

V = mean velocity in (ft/s),

d_{50} = the median size of bed-material fraction by weight (ft).

6. *Colby's Equation (1974)*

Based on Einstein's (1950) probabilistic bed load function, Colby (1964) develops a graphical solution to determine the bed-material discharge in the sand size range. According to Yang (2006), Colby's equation is formulated as:

$$q_t = A(V - V_c)^B \; 0.672 \quad \text{in which } V_c = 0.4673 D^{0.1} d_{50}^{0.33} \tag{2.9}$$

where:

q_t = total bed-material discharge per width (lb/s.ft),

V = mean velocity (ft/s),

V_c = critical velocity (ft/s),

D = mean depth (ft),

d_{50} = the median bed-material by weight (mm),

A = a coefficient, and B = an exponent.

Yang (1996) recommends not applying Colby's method to rivers with median sediment size greater than 0.6 mm and depth more than 3 m. A detailed analysis of the equation and application examples can be referred from Yang (1996).

7. *Toffaleti's Equation (1969)*

Toffaleti (1969) presents a procedure to determine bed-material transport rate based on the probabilistic approach of Einstein (1950). The inputs required for Toffaleti's method consist of the average velocity, the hydraulic radius, water temperature, channel width, d_{65}, energy or water surface slope, bed-material fractions, settling velocity, and the kinematic viscosity of water-sediment mixture. Toffaleti's bed-material discharge formula is applicable for sediment particles size ranging from 0.062 to 16 mm (Yang and Huang, 2001). The complete procedure and a brief summary of Toffaleti's approach can be found in Yang (1996) and Wan and Chien (1999).

8. *Bagnold's Equation (1966)*

Bagnold (1966) assumes that the movement of bed load can be described by applying basic physics. He applies the theory of conservation of energy, and based on to the theory of conservation of energy the rate of sediment materials being transported is proportional to the rate of dissipation of energy.

The law of conservation of energy can be expressed as (Bagnold 1966):
Rate of energy consumed for transporting sediment particles = rate of energy supply of the flowing water * efficiency = rate of loss of potential energy of the flowing water * efficiency

Bagnold's equation for estimating bed load is:

$$q_b = \frac{\tau V e_b}{\tan\alpha} * \frac{\gamma_s}{\gamma_s - \gamma_w} \tag{2.10}$$

where:

q_b = bed load transport rate by weight per unit channel width,

$\tan\alpha$ = ratio of tangential to normal shear force,

τ = shear stress along the river bed,

V = average flow velocity,

e_b = efficiency coefficient for bed load transport.

Bagnold performed experiments and provides two graphs to determine the values of the efficiency coefficients e_b and $\tan\alpha$ (Bagnold, 1966; Yang, 1996).

The rate of work needed to transport suspended load is:

$$W_s = \frac{\gamma_s - \gamma_w}{\gamma_w} q_s \frac{\omega}{\overline{u}_s} \tag{2.11}$$

where:

q_s = suspended load discharge in dry weight per unit time and width,

$\overline{u}_s$ = mean transport velocity of suspended load, and

ω = fall velocity of suspended sediment.

The rate of energy available for transporting suspended load is:

$$W_s^{'} = \tau V (1 - e_b) \tag{2.12}$$

The rate of work done is the power available times the efficiency of the system:

$$\frac{\gamma_s - \gamma_w}{\gamma_w} q_s \frac{\omega}{\overline{u}_s} = W_s^{'} = \tau V (1 - e_b) e_s \tag{2.13}$$

where:

e_s = suspended load transport efficiency.

Bagnold (1966) assumes the mean transport velocity of suspended particles is approximately equal to the flow velocity, and from flume experiments he obtains $(1 - e_b)e_s$ = 0.01.

Then, the suspended load transport rate per width is calculated from:

$$q_s = 0.01 \tau \frac{V^2}{\omega} * \frac{\gamma_w}{\gamma_s - \gamma_w} \tag{2.14}$$

The final total bed-material transport rate equation given by Bagnold is:

$$q_t = q_b + q_s = \frac{\gamma_w}{\gamma_s - \gamma_w} \tau V \left(\frac{e_b}{\tan \alpha} + 0.01 \frac{V}{\omega}\right) \tag{2.15}$$

where:

q_t = total bed-material load in dry weight per unit time and unit width.

9. *Engelund and Hansen's Equation (1972)*

Engelund and Hansen (1972) apply Bagnold's stream power approach and derive the following relationships to predict total bed-material transport using the friction coefficient, dimensionless sediment discharge, and dimensionless shear stress as the main parameters:

$$f^{'}\phi = 0.1\theta^{5/2} \tag{2.16}$$

The friction coefficient $f^{'}$ is computed from:

$$f^{'} = \frac{2gSD}{V^2} \tag{2.17}$$

The dimensionless sediment discharge is calculated using:

$$\phi = \frac{q_t}{\gamma_s \sqrt{(s-1)gd_{50}^3}} \tag{2.18}$$

The dimensionless shear stress is computed by applying:

$$\theta = \frac{\tau}{(\gamma_s - \gamma_w)d_{50}} \tag{2.19}$$

where:

g = acceleration due to gravity,
S = energy slope,
V = average flow velocity,
γ_w = specific weight of water,
γ_s = specific weight of sediment,
d_{50} = median particle diameter
τ = shear stress along the bed,
D = mean water depth, and
q_t = total bed-material load in dry weight per unit time and unit width.

The equation is usually applied for estimating sediment transport rates for sandy rivers with considerable amount of suspended load.

10. *Ackers and White's Equation (1973)*

Following Bagnold's stream power approach and applying dimensional analysis, Ackers and White (1973) formulate a relationship between sediment transport and two dimensionless

parameters: sediment mobility number and dimensionless grain diameter. Their formula includes both bed load and suspended load.

The mobility number of sediment particles is calculated from (Ackers and White, 1973):

$$F_{gr} = \frac{u_*^n}{\sqrt{gd(s-1)}} \left(\frac{V}{\sqrt{32}\log(\frac{\alpha D}{d})} \right)^{1-n} \tag{2.20}$$

where:

u_* = shear velocity,

n = transition exponent which depends on sediment size,

α = coefficient in rough turbulent equation (α = 10),

d = sediment size,

D = water depth.

The sediment particle size is expressed in terms of a dimensionless grain diameter which is computed from:

$$d_{gr} = d\left(\frac{g(s-1)}{v^2} \right)^{1/3} \tag{2.21}$$

The dimensionless sediment transport is given as a function of the dimensionless mobility number and the dimensionless grain diameter. The Ackers and White's (1973) general dimensionless sediment transport function is (Yang, 1996):

$$\begin{aligned} G_{gr} &= f(F_{gr}, d_{gr}) \\ G_{gr} &= \frac{C_t D}{ds}\left(\frac{u_*}{V}\right)^n \end{aligned} \tag{2.22}$$

where:

C_t = total bed-material concentration by weight of flow (ppm).

By analysing 1000 sets of laboratory flume data, they formulate the final expression as:

$$G_{gr} = C\left(\frac{F_{gr}}{A} - 1 \right)^m \tag{2.23}$$

The values of the coefficients A, C, m, and n are determined from analysis of laboratory data with sediment size greater than 0.04 mm, Froude number less than 0.8, and energy gradient (slope) between 0.00006 and 0.037 (Yang, 1996). These coefficients depend on the value of the dimensionless grain diameter and are given as:

If $1 < d_{gr} \leq 60$:

$$n = 1.00 - 0.56 \log d_{gr} \tag{2.24}$$

$$A = \frac{0.23}{\sqrt{d_{gr}}} + 0.14 \tag{2.25}$$

$$m = \frac{9.66}{d_{gr}} + 1.34 \tag{2.26}$$

$$\log C = 2.86 \log d_{gr} - \left(\log d_{gr}\right)^2 - 3.53 \tag{2.27}$$

If $d_{gr} > 60$ (coarse material):

$$\text{n} = 0.00,\ \text{A} = 0.17,\ \text{m} = 1.5 \text{ and C} = 0.025 \tag{2.28}$$

A revised form of the coefficients was published in 1990 (HR Wallingford, 1990) and are given by:

If $1 < d_{gr} \leq 60$,

$$n = 1.00 - 0.56 \log d_{gr} \tag{2.29}$$

$$A = \frac{0.23}{\sqrt{d_{gr}}} + 0.14 \tag{2.30}$$

$$m = \frac{6.83}{d_{gr}} + 1.67 \tag{2.31}$$

$$\log C = 2.79 \log d_{gr} - 0.98\left(\log d_{gr}\right)^2 - 3.46 \tag{2.32}$$

If $d_{gr} > 60$,

$$\text{n} = 0.00,\ \text{A} = 0.17,\ \text{m} = 1.78 \text{ and C} = 0.025 \tag{2.33}$$

11. Yang's Equations (1973,1984)

Yang (1973) introduces the unit stream power approach for estimating sediment transport rate. He defines unit stream power as the product of velocity and slope and it is the power available to transport sediment particles per unit weight of water. He states that the total bed-material transport is directly related to unit stream power.

Yang (1973) considers the following functional relationship between the dominant variables and total sediment transport capacity:

$$\phi(C_t, VS, u_*, \nu, \omega, d) = 0 \tag{2.34}$$

where:

C_t = total bed-material concentration by weight (ppm),

VS = unit stream power,

u_* = shear velocity,

ν = kinematic viscosity,

ω = fall velocity of sediment, and

d = median particle diameter.

Yang (1973) performs dimensional analysis using Buckingham's π theorem, applies multiple regression analysis using sets of laboratory data and obtains the following dimensionless unit stream power formula for sand transport (0.15 mm - 1.7 mm):

$$\begin{aligned} \log C_t &= 5.453 - 0.286 \log \frac{\omega d}{\nu} - 0.457 \log \frac{u_*}{\omega} \\ &+ \left(1.799 - 0.409 \log \frac{\omega d}{\omega} - 0.314 \log \frac{u_*}{\omega}\right) \log\left(\frac{VS}{\omega} - \frac{V_c S}{\omega}\right) \end{aligned} \tag{2.35}$$

where:

C_t = total sand concentration in ppm by weight.

$\frac{V_c S}{\omega}$ is the critical dimensionless unit stream power and is the product of the dimensionless critical velocity ($\frac{V_c}{\omega}$) and the energy slope (S). The critical dimensionless unit stream power is a function of particle Reynolds number.

If $1.2 < \frac{u_* d}{\nu} < 70$,

$$\frac{V_c}{\omega} = \frac{2.5}{\log\left(\frac{u_* d}{\nu}\right) - 0.06} + 0.66 \tag{2.36}$$

If $70 \leq \frac{u_* d}{\nu}$,

$$\frac{V_c}{\omega} = 2.05 \tag{2.37}$$

Yang (1979) states that the incipient motion criterion is not that necessary for high sediment concentration (greater than 100 ppm) and introduces the following dimensionless unit stream power equation for computing sand transport with high concentration:

$$\begin{aligned} \log C_t &= 5.165 - 0.153 \log \frac{\omega d}{\nu} - 0.297 \log \frac{u_*}{\omega} \\ &+ \left(1.78 - 0.360 \log \frac{\omega d}{\omega} - 0.48 \log \frac{u_*}{\omega}\right) \log\left(\frac{VS}{\omega}\right) \end{aligned} \tag{2.38}$$

Keeping the basic form of the original equation, Yang (1984) modifies the coefficients in his equation and formulates a dimensionless unit stream power equation for gravel transport (d>2 mm).

$$\begin{aligned} \log C_{tg} &= 6.681 - 0.633 \log \frac{\omega d}{\nu} - 4.816 \log \frac{u_*}{\omega} \\ &+ \left(2.784 - 0.305 \log \frac{\omega d}{\omega} - 0.282 \log \frac{u_*}{\omega}\right) \log\left(\frac{VS}{\omega} - \frac{V_c S}{\omega}\right) \end{aligned} \tag{2.39}$$

where:

C_{tg} = total gravel concentration in ppm by weight.

Most sediment transport functions are developed for equilibrium sediment transport conditions without considering the effects of wash load. If a very high concentration of wash load exists, this can significantly affect the flow viscosity, sediment fall velocity, and relative specific weight of sediment particles (Yang, 1996). A non-equilibrium sediment transport of varying rates may occur because of high concentration of wash load. Thus Yang et al. (1996) modify Yang's 1979 formula for sediment-laden flow with high concentration of wash load (for Yellow river in China). This equation can be referred from (Yang, 1996).

12. *Van Rijn Equation (1984a, b)*

In van Rijn's equation (1984a), the bed load transport rate is calculated as the product of particle velocity, saltation height and the bed-load concentration.

$$q_b = u_{bs}\delta_b C_b \tag{2.40}$$

Van Rijn (1984a) determines the expressions for particle velocity and saltation height by solving the equations of motion for a saltating particle numerically. He then states that bed load transport can be sufficiently described by using two dimensionless parameters. These dimensionless parameters are the dimensionless particle diameter (D_*), and transport stage parameter (T) which are defined as:

$$D_* = d_{50}\left[\frac{(s-1)g}{\nu^2}\right]^{1/3} \tag{2.41}$$

$$T = \frac{(u_*^{'})^2 - (u_{*,c})^2}{(u_{*,c})^2} \tag{2.42}$$

where:

$u_*^{'}$ = bed shear velocity related to grains and is described in terms of the mean flow velocity and a Chezy coefficient related to the grains of the bed,

$u_{*,c}$ = Shields (1936) critical bed shear velocity.

The bed load concentration is computed from the following expression:

$$C_b = 0.18C_o\frac{T}{D_*} \tag{2.43}$$

where:

C_o is the maximum bed concentration = 0.65 which is a given constant.

The final expression given for computing the bed load transport is (van Rijn, 1984a):

$$q_b = 0.053\rho_s\frac{T^{2.1}}{D_*^3}\sqrt{(s-1)gd_{50}^3} \tag{2.44}$$

The suspended load is computed from (van Rijn, 1984b):

$$q_s = \rho_s FVDC_a \tag{2.45}$$

where:

C_a is the reference concentration at $D = a$ above the bed and is computed from:

$$C_a = 0.015 \frac{d_{50} T^{1.5}}{a D_*^{0.3}} \tag{2.46}$$

where:

F = the correction factor for suspended load,

D = flow depth,

V = average velocity.

The equation is reliable for estimating bed-load transport in the particle range of 0.2 mm - 2 mm, and transport of fine particles (suspended load) in the range of 0.1 mm - 0.5 mm. Other important parameters required for computing the suspended load and detailed derivation of the above formulations can be referred from van Rijn (1984a, b). Finally the total bed-material transport rate is computed as the sum of bed load and suspended load.

13. Parker's Equation (1990)

Parker (1990) develops an empirical gravel transport function based on the equal mobility concept and field data. He assumes that the bed load transport in gravel bed rivers is accomplished by means of the mobilization of grains exposed on the bed surface. Parker's dimensionless bed load transport function W_i^* and dimensionless shear stress parameter ϕ_i are defined as:

$$W_i^* = \frac{q_{bi} * (s-1)}{p_i \rho_s DS \sqrt{gDS}} \tag{2.47}$$

$$\phi_i = \frac{DS}{d_i (s-1) \tau_{ri}^*} \tag{2.48}$$

The value of τ_{ri}^* based on d_{50} is 0.875, i. e. $\tau_{ri}^* = 0.0875 \frac{d_{50}}{d_i}$.

where:

q_{bi} = bed load transport by weight per unit channel width in size fraction d_i,

D = water depth,

S = slope, and

p_i = fraction by weight in size d_i.

Because of the assumption of equal mobility of all sizes only d_{50} is used to characterize bed load transport as a function of dimensionless shear stress.

If $\phi_{50} < 1$,

$$W^* = 0.0025\phi_{50}^{14.2} \tag{2.49}$$

If $1.0 \leq \phi_{50} \leq 1.59$,

$$W^* = 0.0025\exp\left\{14.2\left(\phi_{50}-1\right)-9.28\left(\phi_{50}-1\right)^2\right\} \tag{2.50}$$

If $\phi_{50} > 1.59$,

$$W^* = 13.685\left(1-\frac{0.853}{\phi_{50}}\right)^{4.5} \tag{2.51}$$

The above equations are empirically fitted using field data with sediment size ranging from 18 to 28 mm.

The sediment transport functions mentioned above are only some of the many available ones. There are so many equations and none of them can be recommended for universal application under different and varying site conditions. The selection of the appropriate equation still requires the judgment and experience of the engineer. The accuracy of the equations is also under great debate.

2.4 Factors Affecting Sediment Transport

In order to be able to express the sediment transport capacity of a river, a formula should include all variables that play a significant role in determining the sediment transport process. Comprehensive analysis of variation in hydraulic resistance (Rooseboom and Le Grange, 1999) suggests that a number of variables play important roles in determining flow resistance which intern determines sediment transport capacity.

Factors affecting sediment transport include (Chien and Wan, 1999):

- Flow conditions: velocity V, flow depth D, slope S, gravitational acceleration g.
- Physical properties of water: specific weight γ_w and kinematic viscosity ν.
- Physical properties of sediment: specific weight γ_s, fall velocity ω, sediment diameter d.
- Boundary conditions: bed-material composition and channel width B.

Thus the sediment transport capacity of flow, represented by average sediment concentration, can be described by:

$$S_m = f(V, D, g, S, \gamma_w, \gamma_s - \gamma_w, \nu, \omega, d, B) \tag{2.52}$$

In addition to the above parameters, size and shape of bed material, morphology of bed forms, and availability of sediment from the source area affect the sediment transport capacity. Therefore the amount of materials under transport at a given boundary condition can be limited by sediment supply or transport capacity of the flow. The existence of so many governing variables creates uncertainty and makes the formulation of a single equation for accurately predicting sediment transport quite challenging.

2.5 Non-Uniform Sediment Transport

The composition of sediment particles in a natural river bed is usually non-uniform and the estimation of sediment transport rates in rivers is more difficult than prediction in laboratory flumes where bed composition is mostly uniform. Most sediment transport equations are valid only for uniform sediment distribution because they are derived under this assumption. The prediction of sediment transport for natural rivers is more difficult than for uniform sediment in laboratory flumes (Wu et al., 2004). Two methods are usually used to deal with non-uniform sediment. If the total transport rate is required, the formulas developed for uniform motion can be utilized by selecting a particle diameter which is representative of the river bed. But, in order to predict the transport rates of various diameters, the mutual effects of different particle sizes and their interaction should be taken into consideration (Xiaoqing, 2003).

The commonly used representative sediment particle diameters include (Wu et al., 2004):

- the median diameter of bed-material, d_{50}, for which 50% is finer by weight.
- the diameter of bed material, d_{35}, for which 35% is finer as proposed by Einstein (1944) and Ackers and White (1973).
- Meyer-Peter and Müller (1948) suggested mean diameter as the representative size and is computed from the following formula:

$$d_m = \frac{\sum d_i \Delta p_i}{100} \tag{2.53}$$

where:

Δp_i is percentage of particles of diameter d_i in the bed-material.

A single representative size may not represent the variation in the bed graduation correctly and some authors recommend the use of variable representative sizes for predicting the sediment transport rates (van Rijn, 1984; Hsu and Holly, 1992; Molinas and Wu, 1998, 2000). Because of the interaction between different particle sizes, the condition for initiation of motion for sediment mixtures is different from a uniform material (Samaga et al., 1986) and the gradation of sediment under transport is different from the gradation of the bed composition (Proffitt and Sutherland, 1983; Wu et al., 2000). All these factors add a huge uncertainty and complexity in modelling sediment transport accurately.

Transport Rates of Various Sediment Sizes

Many engineering problems require not only the calculation of total sediment transport rate but also the transport rate of various grain sizes of non-uniform sediment. For example, the fining process in the upstream reach of a dam and the armouring layer in the downstream reach require these computations. Very few researchers have studied the movement of various sizes of non-uniform sediment because the mutual interactions between the various sizes are complex. The methods are generally found to be unsatisfactory so far in predicting the transport rate for various sediment particles due to the complexity of transport of sediment mixtures and the lack of knowledge concerning the motion of individual size and its effect on other sizes (Chien and Wan, 1999).

2.6 Selection of Sediment Transport Equation

The selection of appropriate sediment transport formulas under different flow and sediment conditions is important to sediment transport and river morphologic studies. There are many equations derived by various researchers using different variables as the main factors influencing the amount of sediment transport. The results obtained from applying different equations usually have a very high prediction error and show significant discrepancy from each other and from the measured transport rates (Yang, 1996). There is not any equation which is universally applicable under varying site conditions. Great care should be taken in selecting appropriate equations based on site specific conditions and the problem required to be solved (bed load or total bed-material load). The range of datasets chosen for the use in developing the equations and the corresponding application guidelines recommended by the authors should be considered while selecting a particular equation for computing sediment transport.

The amount of transported material in natural rivers is influenced by many factors in addition to the main variables utilized in formulating many of the existing equations. Some of the additional factors include: the percentage of bed surface covered by coarse material, shape factor of sediment, sediment availability for transport, variations in the hydrologic cycle, rate of supply of wash load from upstream, water temperature, channel geometry,

strength of turbulence, etc. (Yang, 2006). All these factors make the development of a universally applicable equation quite challenging, and so far none of the sediment transport equations is proved to be generally applicable for accurately computing sediment transport rates in natural rivers (Yang, 1996). Therefore engineers have to analyse and compare the accuracies and ranges of application of different formulas before using any of the formulas for practical application.

A comprehensive procedure for selecting an appropriate sediment transport equation for natural rivers based on site specific conditions and data availability is provided by Yang (1996). While selecting equations for practical application, the characteristics of the river and the range of application recommended by the authors of the formulas have to be considered. Generally, many equations have to be tried and the best equation is selected based on comparison between measured sediment transport rates and the amount computed using a particular equation.

The equations selected for application using the laboratory and field datasets compiled in this dissertation are briefly presented in the previous section.

3 Fuzzy Logic Based Modelling and Its Application in Engineering

3.1 Introduction to Fuzzy Set Theory

Fuzzy set theory is used for modelling ambiguity and uncertainty in decision making. Fuzzy logic was first introduced by Zadeh (1965) and as its name implies it is usually applied for modelling complex processes about which the physics governing them is not well understood. A fuzzy model is a theoretical representation of a process in terms of fuzzy variables, rules and methods that define the input-output relationship (Garibaldi, 2005). Fuzzy set theory has made a vital contribution to the representation and processing of uncertainty. With increasing human knowledge, uncertainty has been expected to decrease and gradually even disappear. However, in the last decades it has been recognized that uncertainty, imprecision and ambiguity are inevitable and intrinsic parts of complex natural systems (Bardossy and Duckstein, 1995). Fuzzy logic theory provides a means for dealing with this uncertainty. The main advantage of fuzzy rules is their ability to express complex systems containing a great deal of uncertainty in a linguistic way which is easily interpretable. Fuzzy modelling is a description of a system using fuzzy quantities and provides a linguistic explanation of the behaviour of the system (Sugeno and Yasukawa, 1993). Fuzzy modelling is becoming popular because it is very difficult to express complex systems with a lot of variables by using mathematical equations. Fuzzy models are attractive alternatives in conditions where numerical models are computationally inefficient (Vernieuwe et al., 2005).

If X is a collection of objects generally represented by x, a fuzzy set A in X is defined as (Jang et al., 1997):

$$A = \{(x, \mu_A(x)) | x \in X; \mu_A(x) \in [0,1]\} \tag{3.1}$$

where $\mu_A(x)$ is called the membership function of the fuzzy set A, which assigns for each element of X a degree of membership between 0 and 1, inclusive.

Fuzzy logic is used to model uncertainty especially in cases where there is not sufficient knowledge about the system to be modelled and by using fuzzy logic it can be simplified and represented linguistically which is similar to the way human beings think. In fuzzy set theory, uncertainty is expressed by using membership functions. In the classical set theory an element either fully belongs to a set or it does not, i. e. the membership value of an element can only be either one or zero. But in fuzzy set theory an element can belong to more than one set with different degree of membership ranging from zero to one. Fuzzy set theory is a generalization of Boolean logic to conditions where data are modelled by variables whose elements have zones of gradual transition or overlapping boundaries, rather than sharp boundaries (Barreto-Neto and de Souza Filho, 2008). The fuzzy logic theory and applications in engineering are discussed in many books including: Dubois and Prade (1980), Ross (1995), and Klir et al. (1997).

Fuzzy model identification requires structure and parameter identification for the system to be modelled. Structure identification involves selection of input and output variables, type of inference, and determination of the number of membership functions for each of the variables which determine internally the number of rules (Garibaldi, 2005). Since fuzzy models have two parts, premise and consequent, the resulting rules have two structures called premise structure and consequent structure (Azar, 2010). Parameter identification is the process of determining the optimum parameters of input and output membership functions in the rule system and optimization of the rule base. Structure identification is usually associated with identifying the optimum number of fuzzy rules describing a system. The number of rules is proportional to the number of membership functions defined for each of the input variables. The structure of a fuzzy model can be optimized by merging similar fuzzy sets and eliminating rules which are not significant.

3.2 General Structure of a Fuzzy Model

In fuzzy modelling, a process is usually represented qualitatively using IF-THEN statements. If there are more than one input parameters, conjunction operators such as AND, OR can be used to relate the input variables to each other to define the result as a combination of the input variables. Fuzzy rule consists of a set of arguments $A_{i,k}$ in the form of fuzzy sets with membership functions $\mu_{Ai,k}$ and a consequence B_i which is also in the form of a fuzzy set (Bardossy and Duckstein, 1995). An example of a general fuzzy rule system is:

IF x_1 is $A_{i,1}$ AND x_2 is $A_{i,2}$ AND…AND x_k is $A_{i,k}$ THEN y_i is B_i

where $x_1, x_2, ..., x_k$ are input variables in the system and y_i is the corresponding output variable.

In general a fuzzy model has four basic components. These are a fuzzification interface, a fuzzy rule base, a fuzzy inference, and a defuzzification interface (Tayfur et al., 2003). Figure 3-1 shows these basic components and the flow chart for developing a fuzzy model.

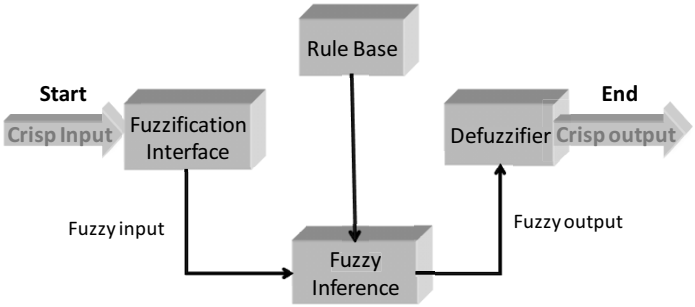


Figure 3-1: Basic components of a fuzzy model.

The steps followed to develop a fuzzy model and the basic components in Figure 3-1 are summarized below:

- First dominant input parameters affecting the problem to be modelled are selected, and the variables are divided into a number of fuzzy sets with linguistic values. Uncertainty and vagueness in the input parameters are described using fuzzy sets. A membership function expresses the degree to which an element belongs to a set. Fuzzification is a process by which crisp input variables are converted into fuzzy sets using membership functions.
- Next, a fuzzy inference system is selected and the rule is fired that maps the input variables to corresponding outputs using IF-THEN statements. For certain combinations of input variable values, the degree of fulfilment of a rule is evaluated using the membership degrees of the input variables for the corresponding fuzzy sets. Implication is a process that is used to evaluate the portion of the membership function that is representative for a particular rule, and compute the degree of fulfilment of a rule (Mahabir et al., 2003). The degree of fulfilment determines the contribution of a rule for the final output, or it is the truth value corresponding to the fulfilment of the conditions of a rule for given premises (Bardossy and Duckstein, 1995). Each rule results in one fuzzy output and the final output is obtained by combining the fuzzy outputs from all the rules. The fuzzy output sets of all the rules obtained from implication are combined into a single output fuzzy set and this process is called aggregation. There are different techniques of aggregation and detailed explanation can be found in many books (Dubois and Prade, 1980; Klir et al., 1997).

- Finally, the fuzzy output is defuzzified to obtain a crisp value.

Fuzzification of Input Variables

Input variable selection is a very important part of fuzzy modelling. The input parameter selection can be performed by using expert knowledge or by means of data-driven techniques such as artificial neural networks (Chen and Mynett, 2003), principal component analysis (Joliffe, 1986) or genetic algorithms (Holland, 1975; Goldberg, 1989). Data-driven methods are preferred whenever there is not enough prior knowledge giving explanation about relationships between variables. The key idea in fuzzy logic is the allowance of partial membership of an element to different subsets of a universal set, instead of completely belonging to a single set. Fuzzification of input variables is the first step in fuzzy modelling.

Membership Functions

In classical set theory, sets have non-overlapping boundaries, an element either belongs to a set or not. A fuzzy set is a set with overlapping boundaries making partial membership possible. Membership functions define how each element in the input parameter space is mapped to a membership value. Membership functions can have different shapes. The most commonly used shapes are piecewise linear (triangular or trapezoidal), Gaussian, generalized bell-shaped, and sigmoidal. Triangular and trapezoidal membership functions are usually preferred because of their simplicity and linear behaviour. Linguistic variables can be assigned to fuzzy sets. A linguistic variable is a variable with values which are words or sentences (Bardossy and Duckstein, 1995). The main advantage of fuzzy sets is the ability to assign linguistic values to the sets in a human way of thinking and this makes fuzzy models simple and easily interpretable. Examples of the most commonly used linguistic values are high, medium, low, etc. The linguistic variables can further be refined using linguistic modifiers such as very, mostly, not, and can be used together with the linguistic values to define additional fuzzy sets (Bardossy and Duckstein, 1995). There are many methods for defining membership functions for the parameters of a fuzzy model. Some of these techniques are: methods based on perception, intuition, neural network based methods, clustering-based methods, and genetic algorithms (Tayfur and Singh, 2006; Garibaldi, 2005). Defining membership functions in fuzzy expert systems is subjective with different shapes available to choose from, and this is considered to be one of the weaknesses of expert systems. The following guidelines are provided by Garibaldi (2005) for defining membership function which "...should be justifiable in number and should not overlap too much, maximum membership should be one, all terms should be convex and cover the whole range of input parameters". The membership function is considered to be both the strongest and the weakest point of fuzzy set theory (Cornelissen et al., 2001). The main advantage of membership functions is their ability to define overlapping regions which make expression of uncertainty and fuzziness possible (Bosserman and Ragade, 1982; Silvert, 1997).

The most commonly applied shapes of membership functions and their parameters are briefly presented next as discussed in Jang et al. (1997).

A triangular membership function has three parameters (a, b, c) which define the x-coordinates of the corners of the triangle and should satisfy the properties of a triangle.

$$\text{triangle}(x; a, b, c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ \dfrac{c-x}{c-b}, & b \le x \le c \\ 0, & c \le x \end{cases} \tag{3.2}$$

A trapezoidal membership function is defined by four parameters (a, b, c, d) as:

$$\text{trapezoid}(x; a, b, c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \dfrac{d-x}{d-c}, & c \le x \le d \\ 0, & d \le x \end{cases} \tag{3.3}$$

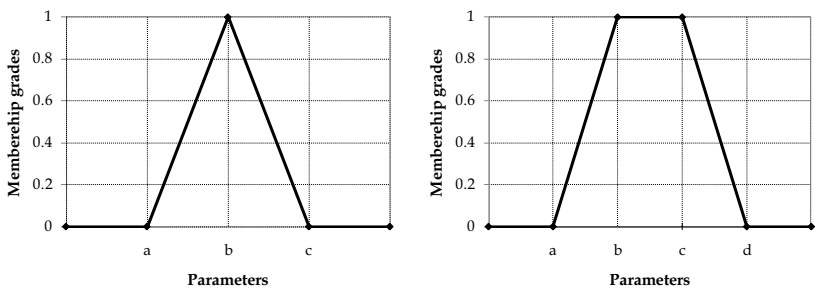


Figure 3-2: Triangular and trapezoidal membership functions and their parameters.

A graphical illustration of triangular and trapezoidal membership functions with their parameters is presented in Figure 3-2. This is a simple example case and the parameters of the membership functions do not necessarily have to be specified with equal intervals.

A Gaussian membership function is specified by two parameters (c, σ) where c represents the mean and σ the width:

$$\text{Gaussian}(x; c, \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \tag{3.4}$$

A generalized bell-shaped membership function is defined by three parameters (a, b, c) and will be discussed in detail later:

$$\text{bell}(x; a, b, c) = \frac{1}{1+\left(\frac{x-c}{a}\right)^{2b}} \tag{3.5}$$

The Gaussian and generalized bell-shaped membership functions have flexible shapes compared to triangular and trapezoidal membership functions which have sharp corners.

Figure 3-3 illustrates an example of trapezoidal membership functions representing flow velocity. Let's assume that flow velocity is described by using three trapezoidal membership functions with linguistic expressions. In this example, velocity values less than 0.5 m/s have full membership of 1 to the first fuzzy set (low flow velocity). Velocity values between 0.5 and 2 m/s belong to both the medium and low flow velocity sets with different membership grades. For example flow velocity of 1 m/s has a membership of 0.33 to medium and 0.67 to low flow velocity fuzzy set (see Figure 3-3). Velocities between 2 and 3 m/s belong only to the medium fuzzy set. Flow velocities greater than 4 m/s belong to the high flow velocity set.

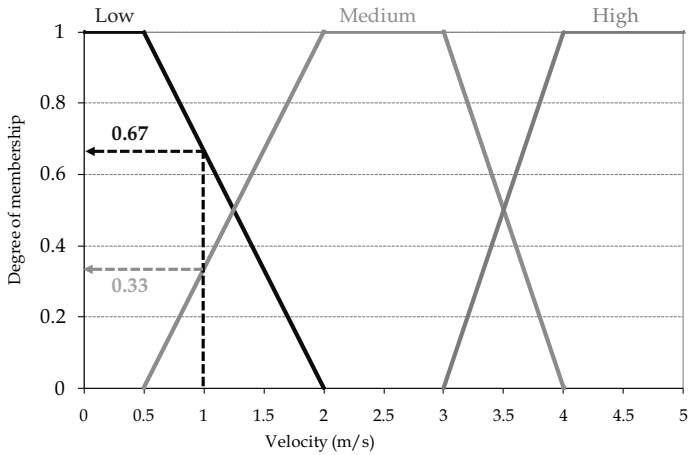


Figure 3-3: Example trapezoidal membership functions for flow velocity.

Defuzzification

Defuzzification is the process of converting the fuzzy output from the fuzzy inference engine into a crisp value. There are different methods of defuzzification such as centre of gravity, weighted average, bisector of area, mean of maxima, left-most maximum, right-most maximum, etc. (Jantzen, 1998; Bardossy and Duckstein, 1995).

3.3 Types of Fuzzy Inference Systems

There are two types of commonly utilized fuzzy inference systems. These are the Mamdani inference system and the Takagi-Sugeno fuzzy inference system. In Mamdani systems both the input and output variables are fuzzy (Mamdani, 1977) where as in Takagi-Sugeno inference systems the output is expressed as a linear function of the input variables (Takagi and Sugeno, 1985).

Mamdani Inference Systems

This inference methodology was proposed in 1975 by Ebrahim Mamdani. Fuzzy rules in Mamdani inference are defined in the following format:

Rule R_i

$$IF\ x_1\ is\ A_{i,1}\ AND\ IF\ x_2\ is\ A_{1,2}\ AND...IF\ x_k\ is\ A_{i,k}\ THEN\ y \text{ is } B_i \tag{3.6}$$

For $i = 1,2,...,n$

where n is the number of rules, $x_j (j = 1,2,...,k)$ are the input variables, y is the output variable, and A_{ij} and B_i are fuzzy sets that are characterized by membership functions $A_{ij}(x_j)$ and $B_i(y)$ respectively. The result of Mamdani inference is a fuzzy set which has to be defuzzified in order to get a crisp output value. Mamdani's system is the most commonly used system, especially for fuzzy systems based on expert knowledge.

Figure 3-4 demonstrates an example case where a system has two inputs (x and y) and one output (z). Input x has two membership functions (A_1 and A_2) and similarly, input y has two membership functions (B_1 and B_2). The output is expressed by two rules with two membership functions (C_1 and C_2). The figure illustrates how the Mamdani inference system derives the final output using minimum for implication (t-norm) and the maximum for aggregation (t-conorm) to obtain the final fuzzy output C′.

This fuzzy output is defuzzified to obtain the final crisp value using the centroid of the area (Z_{COA}) as a method of defuzzification.

Z_{COA} is computed by using:

$$Z_{COA} = \frac{\int_Z \mu_{C'}(z) z dz}{\int_Z \mu_{C'}(z) dz} \tag{3.7}$$

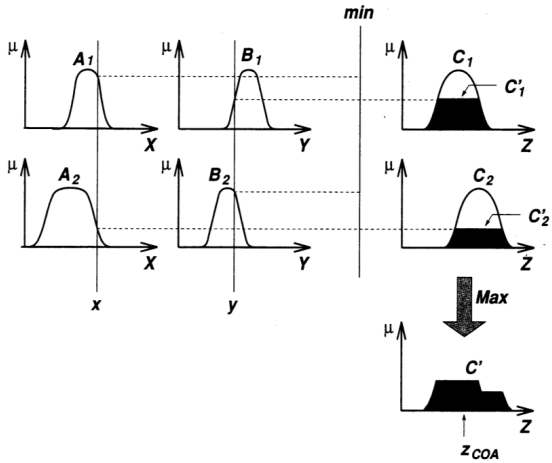


Figure 3-4: A Mamdani fuzzy inference system using min and max for t-norm and t-conorm operators respectively (Jang et al, 1997).

Takagi-Sugeno Inference Systems

A Takagi–Sugeno fuzzy model consists of a set of rules R_i for $i = 1,2,...,n$, and the rule is defined in the following format:

$$IF\ x_1\ is\ A_{i,1}\ AND\ IF\ x_2\ is\ A_{1,2}\ AND...IF\ x_k\ is\ A_{i,k}\ THEN\ y = f_i(x_1, x_2, ...x_k) \tag{3.8}$$

where $f_i(x_1, x_2, ..., x_k)$ is linear. If first order Takagi-Sugeno inference is used:

$$f_i\ (x_1, x_2,\ ...,\ x_k) = a_{i,1}x_1 + a_{i,2}x_2 + ... + a_{i,k}x_k + a_{i,0} \tag{3.9}$$

The output of the Takagi-Sugeno fuzzy model is computed by:

$$y = \sum_{i=1}^{n} f_i(x_1, x_2, ...x_n)\varphi_i(x) \tag{3.10}$$

where n is the number of fuzzy rules and $\varphi_i(x)$ is the basis function used to normalize the degree of rule fulfilment by using the product *t*-norm. Takagi-Sugeno's system is computationally more efficient because the output is crisp and does not require the time consuming defuzzification process.

The choice of the inference methodology depends on the problem required to be solved. If there is sufficient expert knowledge available, expression of uncertainty in the system is required, and non-numeric output is sufficient, Mamdani inference method is widely used. But, if the system to be modelled is too complicated, the fuzzy rules cannot be defined by experts, and if numerical crisp output is required then the data-driven rule generation with Takagi-Sugeno (TS) fuzzy inference system is applied (Garibaldi, 2005).

In this dissertation, the TS fuzzy inference system is applied because it is the most suitable for data-driven fuzzy modelling and adaptive techniques can be used for the optimization of the model.

3.4 Rule Induction

Fuzzy rules consist of two parts called antecedent and consequent. The antecedent part describes the conditions on one or more input variables; and the consequent part expresses the corresponding values of the output variable (Adriaenssens et al., 2004). There are two ways for generating the fuzzy rules describing the system to be modelled. These are approaches based on either expert knowledge or existing data.

3.4.1 Expert Knowledge Based Systems

If sufficient expert knowledge exists about the system to be modelled, rules and membership functions for the variables can be defined by experts. Membership functions with simple linear shapes are mostly implemented in expert systems. One example of a knowledge based system is the habitat model CASiMiR (Computer Aided Simulation System for Instream Flow Requirements) developed at the institute of hydraulic engineering at the University of Stuttgart (Jorde, 1996; Schneider, 2001). CASiMiR is used to estimate fish habitat requirements by using water depth, flow velocity, substrate, and cover type as input parameters. For different combinations of the input variables, the corresponding fish habitat preferences are defined by experts depending on fish species and life stage. Biologists who study fish behaviour express the rules based on their observation and expertise. The model is freely available and can be downloaded from:
http://www.casimir-software.de/download_eng.html

3.4.2 Data-Driven Rule Generation

But if the system under consideration is too complicated and the rules governing the processes are not known with sufficient quality, analysis of collected data can be used to generate the rules using data-driven techniques such as hybrid neuro-fuzzy approaches. Only a set of observations is available and the rule system has to be constructed to describe the interconnections between the input and output elements of the datasets. For complex systems, fuzzy rules based on expert knowledge may suffer from a loss of accuracy (Guillaume, 2001) and data-driven models are preferred.

There are combinations of the two techniques where the rules can be assessed by experts directly but available data can be used to update them, or the rules are not known explicitly but the variables required for the description of the system can be specified by experts (Bardossy and Duckstein, 1995).

In the case of sediment transport modelling, the hydraulic processes are complicated and the rules cannot be defined by experts with sufficient quality. Therefore the option is to generate the fuzzy rules by using data-driven approaches. In this dissertation, a neuro-fuzzy system is used for optimizing the model.

3.5 Adaptive Neuro-Fuzzy Inference System (ANFIS)

Adaptive Neuro-Fuzzy Inference System (ANFIS), first introduced by Jang (1993), is usually utilized for optimizing data-driven fuzzy models. ANFIS is a network structure consisting of a number of nodes and layers connected through directional links. Each node has a node function with adjustable or fixed parameters. ANFIS is widely used because it has the ability to combine the verbal power of a fuzzy system with the computational ability of a neural network. Neuro-fuzzy simulation refers to techniques that apply optimization algorithms utilized in neural networks to fuzzy modelling (Brown and Harris, 1994). ANFIS have the ability to be used as a universal approximator. Fuzzy logic and artificial neural networks are commonly used methodologies to extract relationships from observed data; they are able to describe the nonlinear behaviour of systems. Fuzzy logic and artificial neural networks have advantages and disadvantages. Neural networks have the ability to extract hidden relationship in a dataset and are vey suitable for optimization using different optimization techniques, but they tend to be black box models lacking interpretability (Haykin, 1999). Neural networks can handle nonlinearity and input-output mapping adaptively. Fuzzy logic has the ability to express a system in a human way using natural languages and can handle uncertainty, but lacks the flexibility of using adaptive techniques for optimization. The two powerful learning techniques are combined in adaptive neuro-fuzzy inference systems. The adaptive neuro-fuzzy systems are utilized to generate fuzzy models from data by using optimization algorithms that are commonly applied during the learning process of neural

networks. They are implemented to optimize all the parameters of a fuzzy inference system including membership function parameters (Jang et al., 1997).

There are a lot of similarities and differences between fuzzy logic and neural network models. The properties of fuzzy systems and artificial neural networks are summarized in Table 3.1 (Fuller, 2000).

Table 3.1: Properties of neural networks and fuzzy systems (Fuller, 2000).

Skills	Type	Fuzzy Systems	Neural Networks
Knowledge acquisition	Inputs	Expert knowledge	Sample datasets
	Tools	Interaction	Algorithms
Uncertainty	Information	Qualitative and quantitative	Quantitative
Reasoning	Cognition	Heuristic approach	Perception
	Mechanism	Low	Parallel computation
	Speed	Low	High
Adaption	Fault-tolerance	Low	Very high
	Learning adjusting	Induction	Adjusting weights
Natural language	Implementation	Explicit	Implicit
	Flexibility	High	Low

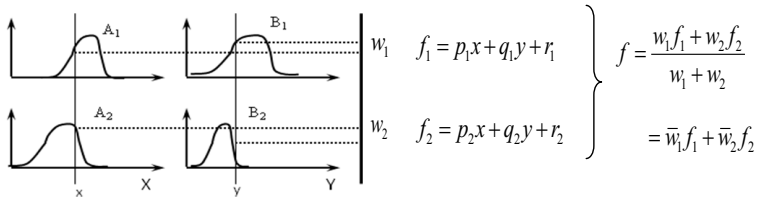
Fuzzy modelling is an effective approach to utilizing linguistic rules whereas neural control is suited for using numerical data pairs (Wang and Mendel, 1992). Deriving fuzzy rules from data involves computing the weight of each rule triggered, accumulation of weights, outputs from each rule, and computing the weighted output for each rule. Wang and Mendel (1992) proposed a five step procedure for generating fuzzy rules from numerical data pairs and how to apply the rules to obtain a mapping from inputs to outputs.

3.5.1 ANFIS Architecture

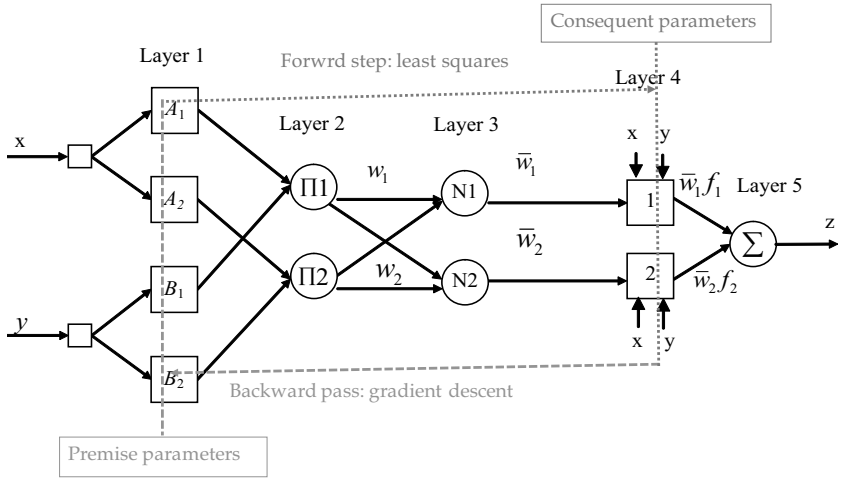
ANFIS has a total of five layers. These are fuzzification, rule, normalization, defuzzification, and summation layers. If first order Takagi-Sugeno fuzzy IF–THEN rules are implemented, the output of each rule is a linear combination of the input variables plus a constant term, and the final output is the weighted average of the model outputs from each fuzzy rule. The detailed algorithm and mathematical background of the hybrid-optimization algorithm in ANFIS can be found in Jang (1993).

Figure 3-5 demonstrates a simple example where the fuzzy inference system under consideration has two input variables, x and y and one output z and the rule base contains two fuzzy IF–THEN rules of Takagi –Sugeno type (Jang, 1993). The Takagi-Sugeno fuzzy

model is illustrated in Figure 3-5 (a), and the corresponding equivalent ANFIS architecture is shown in Figure 3-5 (b).



(a)



(b)

Figure 3-5: a) Two input first order TS model with two rules b) Equivalent ANFIS architecture (adopted from Jang, 1993).

The two rules are defined as:

$$Rule\ 1:\ If\ x\ is\ A_1\ AND\ y\ is\ B_1,\ THEN\ f_1 = p_1x + q_1y + r_1 \tag{3.11}$$

$$Rule\ 2:\ If\ x\ is\ A_2\ AND\ y\ is\ B_2,\ THEN\ f_2 = p_2x + q_2y + r_2 \tag{3.12}$$

where x and y are the crisp inputs to node i, A_i and B_i are the linguistic labels (such as: low, medium, high, etc.) characterized by convenient membership functions and p_i , q_i

and r_i are the consequence parameters (i = 1, 2). An adaptive network as indicated in Figure 3-5 is a multi-layered feed-forward network in which each node performs a computation on the incoming signals which are only propagating in the forward direction. The formulas for the node functions vary from node to node, and the choice of each node function depends on the overall input-output function which the adaptive network is required to carry out. In the network shown in Figure 3-5, a square node is adaptive (has parameters) while a circle node is fixed and does not have any parameters. The basic architecture of ANFIS and the functions and parameters of the layers are described as follows (Jang, 1993):

Layer 1: Fuzzification Layer

Each node in this layer evaluates membership values of the crisp input parameters by using membership functions. Every node i in this layer is an adaptive node, representing membership functions described by generalized bell shape functions. Parameters in this layer determine the final shape of the membership function and are called *premise parameters.*

$$O_i^1 = \mu_{A_i}(x) = \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^{2b_i}\right]} \tag{3.13}$$

where x is the input to node i, and A_i, is the linguistic label associated with this node function, $\{a_i, b_i, c_i\}$ are premise parameters of the membership function which are adjusted during model optimization. The parameters a and b of the generalized membership function vary the width of the curve, c locates the centre of the curve, and b should be always positive. The generalized bell-shaped membership function is continuously differentiable and is symmetric about the line $x = c$. These parameters are adjusted during model optimization. Other possible shapes like trapezoidal and Gaussian membership functions can also be defined and the most suitable is chosen by a sensitivity analysis.

Layer 2: Rule Layer

This layer consists of fixed nodes without parameters and these nodes multiply incoming signals from layer one, and the product represents *the firing strength* of a rule. This is similar to the degree of fulfilment of a fuzzy rule.

$$w_i = \mu_{A_i}(x) * \mu_{B_i}(x) \quad , i = 1, 2. \tag{3.14}$$

Layer 3: Normalization Layer

In this layer, the nodes calculate the ratio of the ith rule's firing strength to the sum of the firing strengths of all rules. Outputs of this layer are called *normalized firing strengths.* They

represent the weight of every rule in determining the final output which is computed as a combination of the outputs from each rule.

$$\overline{w}_i = \frac{w_i}{w_1 + w_2} \tag{3.15}$$

Layer 4: Defuzzification Layer

The nodes in this layer are adaptive with linear node functions. Parameters in this layer are called *consequent parameters* which are parameters of output membership functions. These parameters are adjustable during the training of the model with collected data.

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \tag{3.16}$$

where $\overline{w}_i$ is the output of layer 3 and $\{p_i,\ q_i, r_i\}$ is the parameter set.

Layer 5: Summation Layer

The node in this layer is a single fixed node and computes the final model output as the combination of all incoming signals from every rule fired. It is a weighted average combination.

$$O_i^5 = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{3.17}$$

3.5.2 Hybrid Learning Algorithm

ANFIS provides a fuzzy modelling procedure to learn information about a datasets, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input-output data. ANFIS uses back-propagation (gradient descent) learning to determine premise parameters and least mean squares estimation to determine the consequent parameters (Jang, 1993). This is referred to as a hybrid learning scheme. Premise parameters are parameters associated with input membership functions which determine the final shape of membership functions. The consequent parameters in the first order Takagi-Sugeno inference system are coefficients of the linear equations expressing output membership functions. Detailed analysis of ANFIS and the hybrid algorithm used for optimization can be found in (Jang, 1993; Jang and Sun, 1995; Jang et al., 1997).

The output in Figure 3-5 is computed by:

$$z = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 = \bar{w}_1 f_1 + \bar{w}_2 f_2$$
$$= (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + (\bar{w}_1) r_1 + (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 + (\bar{w}_2) r_2 \tag{3.18}$$

This is a linear equation which involves determination of the consequent parameters. In the hybrid learning algorithm of ANFIS these consequent parameters are identified first using the least squares algorithm.

If a fuzzy model has n input variables, and p number of fuzzy partitions (membership functions) for each variable then (Jang and Sun, 1995):

$$\text{Number of parameters in layer 1} = 3.(p.n) = |S_1| \tag{3.19}$$

$$\text{Number of parameter in layer 4} = (n+1).p^n = |S_2| \tag{3.20}$$

The parameter set $S_1 = \{(a_{11}, b_{11}, c_{11}), (a_{12}, b_{12}, c_{12}), ..., (a_{1p}, b_{1p}, c_{1p}), ..., (a_{np}, b_{np}, c_{np})\}$ represents the parameters of the fuzzy partitions (input membership functions) used in the rules, and is a set of premise parameters. The parameter set $S_2 = \{(a_{1,1}, ..., a_{1,n}, a_{1,0}), ..., (a_{p^n,1}, ..., a_{p^n,n} a_{p^n,0})\}$ represents the coefficients of the linear functions in the fuzzy rules, consequent parameters.

ANFIS uses a hybrid learning algorithm which has two steps (Jang, 1993, Jang et al., 1997):
In the forward pass S_1 is fixed and S_2 is computed by iterative least square error procedure. In the second or backward step, S_2 remains fixed and S_1 is computed with a gradient descent or back propagation algorithm (see Figure 3-5). This procedure is then repeated until the optimization criteria are satisfied.

The ANFIS technique is used for estimating the desired output parameters of a system when enough training data is provided. Because ANFIS allows the modelling of physical phenomena in complex systems without requiring explicit mathematical representations, this technique is particularly useful for engineering applications where classical approaches are too complicated to be used. In order to use this system, first the ANFIS model should be trained with a series of historical data. The training should be well performed, and the data for training should include enough measured data that are involved in the application of the model. The fuzzy models which are optimized using ANFIS can generate fuzzy sets with too much or no overlap, thereby making the interpretation of the model difficult (Espinosa and Vandewalle, 2000). Jang (1993) mentioned the importance of model interpretability and described how to constrain the optimization of ANFIS to preserve interpretability of the final

network. Espinosa and Vandewalle (2000) proposed an algorithm to improve the trade-off between numerical accuracy and interpretability while generating fuzzy rules from data.

The development of fuzzy models requires the selection of several parameters: position, shape and distribution of the membership functions, rule base construction, selection of the logical operations, consequences of the rules, etc. Because of the high number of degrees of freedom, there are many difficulties in developing accurate fuzzy models. Ali and Zhang (2001) outlined the following major difficulties in fuzzy modelling applications:

- How to define membership functions including their shapes and linguistic terms?
- How to obtain the fuzzy rule base? This requires sufficient expert knowledge or data.
- What are the best ways to perform different fuzzy operations like intersection, unions, and defuzzification?
- How to improve computational accuracy and reduce computational effort?

The above mentioned points should be carefully considered and tested while developing fuzzy models from expert knowledge or based on analysis of data.

3.6 Applications of Fuzzy Logic and ANFIS in Hydraulic Engineering

Fuzzy logic and ANFIS have been utilized to solve a number of problems in hydraulic engineering. Some of the recent application examples are summarized in this section.

3.6.1 Application of Fuzzy Logic

Recently, fuzzy logic based modelling techniques have been successfully applied in hydraulic engineering and water resources management. Some of these applications include: fuzzy logic for reservoir operation and management, river flow forecasting, modelling the relationship between stage and discharge, estimation of suspended sediment concentration, simulation of rainfall-runoff processes, and flood forecasting.

Fuzzy logic has been used successfully for estimation of suspended sediment during recent years (Kisi, 2004; Tayfur et al., 2003; Kisi et al., 2006, 2009). Barreto-Neto and de Souza Filho (2008) apply fuzzy logic for evaluating runoff in a tropical watershed. Mitra et al., (1998) use fuzzy logic to estimate soil erosion in large watershed using few input parameters and compare the results with that of the Universal Soil Loss Equation. Mahabir et al. (2003) use fuzzy logic to forecast seasonal runoff and show that the fuzzy logic results are better than other regression models. Tayfur et al. (2003) apply fuzzy logic algorithms and estimate runoff-induced sediment transport from bare soil surfaces using slope and rainfall as input variables.

Kisi et al. (2006) apply a fuzzy logic based model with triangular membership functions to estimate river suspended sediment from stream flow and compare the results with that of rating-curve models. The results in their analysis show the fuzzy model performs better than rating curve models in estimating suspended sediment. Tran et al. (2001) use multi-objective fuzzy regression and fuzzy rule-based modelling to improve the performance of the revised universal soil loss equation. Lohani et al. (2007) successfully apply fuzzy logic for modelling stage-discharge-sediment relationships for two gauging stations in India, and compare the results with the outputs from ANN, and sediment rating curves.

Oinam (2011) developed a simple and efficient fuzzy logic-based soil erosion risk model for monitoring soil erosion risk distribution from Awash river basin in Ethiopia. The fuzzy rules derived for qualitative estimation of erosion in his analysis are in good agreement with that of observed patterns and the results obtained from physically based models. He selects few dominant input parameters as slope, rainfall erosivity, and soil type, and fuzzified them by analysis of the ranges of the available data. The corresponding fuzzy rules are defined by detailed literature review and questionnaires filled by experts.

The fuzzy rule based control systems for reservoir operation are studied by Russel and Campbell (1996). Mohan and Prasad (2006) apply a fuzzy logic model for multi-reservoir operation.

Furthermore, the applicability of the fuzzy rule-based modelling to forecast river flow has been explored by various authors (Xiong et al., 2001; Chang et al., 2001; Lohani et al., 2006). Fuzzy logic modelling has also been applied for modelling karstic aquifer management (Bárdossy and Duckstein, 1992) and for simulating soil infiltration (Bárdossy and Disse, 1993). Abebe et al. (2000) apply fuzzy rule-based modelling for reconstructing missing precipitation events. Fuzzy knowledge-based models of annual production of skylarks are developed by Daunicht et al. (1996). Hundecha et al. (2001) demonstrated that a fuzzy logic approach can be used to simulate components of hydrologic processes (e.g. snowmelt, evaporation, runoff and basin response) in areas where sufficient data are not available to model these processes physically based.

3.6.2 Application of ANFIS

An adaptive neuro-fuzzy inference system (ANFIS), which combines fuzzy logic and ANN, has also been successfully applied in hydraulic engineering. Some of these application examples are briefly presented in this section.

Kisi (2005) uses an ANFIS model for daily suspended sediment estimation. Kisi et al. (2008) apply an adaptive neuro-fuzzy system to estimate suspended sediment concentration from stream flow data, and compare the results with that of artificial neural networks and

sediment rating curves. Their results show that the neuro-fuzzy model successfully estimates suspended sediment estimation and performs better than the artificial neural networks and sediment rating curves models.

Chang and Chang (2006) evaluate the capabilities of the ANFIS model with subtractive fuzzy clustering in estimating reservoir water level from hourly rainfall data. Rajaee et al. (2009) investigate the accuracy of ANN and neuro-fuzzy models in estimating daily suspended sediment concentration using discharge and sediment concentration as input parameters in two hydrometric stations, and show that the neuro-fuzzy models have a better accuracy than ANN models. They point out that the main reason for better model performance is embedded in its structure, because it uses the advantage of the simplified function of fuzzy reasoning and the self-learning ability of neural networks with the strong capability of eliminating pseudo signals (Rajaee et al. (2009).

Bakhtyar et al. (2008) apply ANFIS to estimate long shore sediment transport using breaking wave height, breaking angle, and wave period as input parameters and compare the ANFIS results with the results of other empirical equations. Their results indicate that the ANFIS model provides higher accuracy and reliability than the other techniques. Additionally, adaptive neuro-fuzzy inference system are successfully applied for hydrological time series prediction and river flow modelling (Nayak et al., 2004; Firat and Güngör, 2007; Zounemat-Kermani and Teshnehlab, 2008), estimation of scour depth near pile groups (Zounemat-Kermani et al., 2009), flood forecasting (Chen et al., 2006), stream flow reconstruction (Chang et al., 2001), etc.

Based on the literature review done so far, the application of ANFIS and data-driven fuzzy logic approach has not yet been investigated in detail to compute bed load and total bed-material load by utilizing comprehensive laboratory and river datasets. This research aims to accomplish that.

4 Data Available and Analysis

Data preparation and pre-processing which include selection of input variables, data collection, removal of extreme values, selection of training and test datasets are very important in data-driven modelling. Collection of sufficient data with good quality is necessary to derive reliable fuzzy rules from data. This chapter summarizes the different groups of data available and the analysis carried out to prepare the final datasets used for the data-driven fuzzy modelling.

In order to develop a fuzzy model which is capable of successfully estimating sediment transport rates, field measurement of bed load and suspended load is necessary. The resulting rules and equations describing the process of sediment transport can be implemented for analysis and estimation of river morphology changes. There are different methods and instruments for measuring sediment transport. A method for measurement of bed load is described in Emerson (1991). Methods of bed and suspended sediment load measurement can be referred from Xiaoqing (2003).

For this research, laboratory and field data are collected from different sources to develop the fuzzy models. The laboratory data is divided into sand (0.0625 mm < d_{50} < 2 mm) and gravel (d_{50} > 2mm). Field datasets are available for the Rhine and Elbe Rivers, and additional datasets for four rivers is obtained from Yang (1996).

4.1 Laboratory Data

The laboratory flume data is obtained from Prof. Chih Ted Yang from the engineering research centre of Colorado State University. The quality of the available datasets has been carefully analyzed by Prof. Yang in his previous studies and only datasets which are believed to have good measurement accuracy are included. As most sediment transport equations are applicable for a certain range of sediment particle sizes only, the datasets are divided into sand (0.0625 mm < d_{50} < 2 mm) and gravel (d_{50} > 2mm). The laboratory datasets contain total bed-sediment concentration and other hydraulic variables like flow velocity, depth, bed slope, width of the flume, water temperature, and median size of sediment particle (d_{50}).

4.1.1 Sand

The datasets in the sand size range used for this study are the same datasets used by Yang (1973) for developing his sediment transport formula based on the unit stream power approach. They are collected by Gilbert (1914), Nomicos (1956), Vanoni and Brooks (1957), Kennedy (1961), Stein (1956), Guy et al. (1966), Williams (1967), and Nordin (1976). Although these data are collected by different researchers, they are put together to obtain a more general model which is able of capturing the characteristics of the total laboratory data.

Further data analysis is conducted and very low and very high sediment concentration data are removed, because data-driven modelling results are sensitive to the ranges of input data used for the development of the model. Extreme events and outliers skew the final model results and filtering of these values from the datasets is of paramount importance for the successful development of any model based on data mining approaches. In order to ensure the accuracy of measured data, total sediment concentrations less than 100 parts per million (ppm) by weight are not included in the final analysis. To avoid the inclusion of wash load, datasets with median particle diameters of less than 0.0625 mm are excluded (Yang and Huang, 2001). Based on the above criteria a total of 1023 sets of laboratory flume data in the sand size range are selected for developing the data-driven fuzzy model to compute total sediment concentration. After pre-processing of the data and filtering, it is grouped into two categories. These are training and test datasets. The training data is used for optimizing the model so that model output is capable of reproducing observed sediment concentration. The test set is required to measure the performance of the system within the model optimization process to avoid overtraining. The number of training and test datasets from each author and the statistical distribution of the datasets in the sand size range are summarized in Tables 4.1 and 4.2.

Table 4.1: Number of datasets used for training and testing and the corresponding authors for laboratory data in the sand size range.

Author	Number of datasets		
	Training	Test	Total
Gilbert (1914)	418	210	628
Guy, et al. (1966)	155	79	234
Kennedy (1961)	27	13	40
Nomicos (1956)	8	4	12
Nordin (1976)	17	10	27
Stein (1956)	27	14	41
Vanoni and Brooks (1957)	8	4	12
Williams (1967)	19	10	29
Total data	669	344	1023

Table 4.2: Statistical distribution of input and output variables for the training and test datasets of laboratory data in the sand size range.

Data	Variables		Average	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	0.86	2.02	0.24	0.31
		depth (m)	0.10	0.84	0.01	0.10
		bed slope (mm/m)	7.84	25.50	0.58	5.49
		d_{50} (m)	0.54	1.71	0.14	0. 34
	Output	total concentration (ppm)	6668.20	36100.00	106.00	6805.46
Test	Input	mean velocity (m/s)	0.85	1.88	0.28	0.31
		depth (m)	0.11	0.86	0.01	0.12
		bed slope (mm/m)	7.89	29.60	0.51	5.84
		d_{50} (m)	0.54	1.71	0.14	0.34
	Output	total concentration (ppm)	6895.14	35900.00	140.19	7325.03

4.1.2 Gravel

Laboratory datasets with a median sediment particle size greater than 2 mm are considered to be gravel, and are treated separately from the sand data. The gravel datasets were collected by Casey (1935), Gilbert (1914), Graf and Suszka (1987), Meyer-Peter and Müller (1948), Sato et al. (1958), and Song et al. (1998). Here also, datasets values where the measured gravel concentration is less than 100 ppm are excluded from further consideration. After detailed data analysis and filtering 392 gravel transport datasets are selected for the final model development. The total number of datasets utilized for developing the final model from each of the authors for gravel laboratory data are summarized in Table 4.3.

Table 4.3: Number of datasets used for training and testing and the corresponding authors for laboratory data in the gravel size range.

Author	Number of datasets		
	Training	Test	Total
Casey (1935)	14	7	21
Gilbert (1914)	88	44	132
Graf and Suszka (1987)	53	26	89
Meyer-Peter and Müller (1948)	42	21	63
Sato et al. (1958)	44	21	65
Song et al. (1998)	22	10	32
Total data	263	129	392

The summary of statistical analysis such as the mean (average), maximum, minimum, and standard deviation (St. dev) of the input and output variables are presented in Table 4.4 for the training and test datasets.

Table 4.4: Statistics of input and output variables for training and test data of gravel.

Data	Variables		Average	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	1.08	2.88	0.53	0.42
		depth (m)	0.20	1.03	0.02	0.19
		bed slope (mm/m)	10.15	30.20	1.30	6.71
		d_{50} (mm)	8.69	28.65	2.04	7.56
	Output	total gravel concentration (ppm)	1965.85	10700.00	100.90	2349.41
Test	Input	mean velocity (m/s)	1.08	2.87	0.40	0.42
		depth (m)	0.19	0.96	0.02	0.18
		bed slope (mm/m)	10.65	31.00	1.00	7.23
		d_{50} (mm)	8.69	28.65	2.04	7.62
	Output	total gravel concentration (ppm)	1927.70	10700.00	100.60	2715.30

4.2 Field Data from Different Authors

To evaluate the applicability of the modelling approach to natural rivers, field datasets are also collected. The field datasets are obtained from two sources. Further categorization of the field datasets is performed and the different categories are analyzed separately because the general morphological characteristics of the rivers are quite different. This affects the ranges of the input and output data and the distributions which significantly affect the performance of models based on data analysis.

The first group is provided by Yang (1973) and consists of data considering only total bed-material load (bed load + suspended load – wash load). Here measured total bed-material concentrations and other hydraulic parameters are available for four rivers. They are the Niobrara River (Colby and Hembree, 1955), Mountain Creek (Einstein, 1944), Middle Loup River (Hubbell and Matejka, 1959), and Rio Grande (Nordin, 1964). The number of collected data from these four rivers is 166. Sediment discharges less than 100 ppm by weight are not included in the analysis. Observations with sediment particle sizes only in the sand size range (0.0625 mm – 2 mm) are used and the wash load portion is excluded. This results in 129 datasets in the sand size range for a final analysis.

Tables 4.5 and 4.6 summarize the number of training and test datasets from each author (river), and the statistical distribution of the input and output variables for field datasets of this category.

Table 4.5: Number of training and testing datasets and the corresponding authors for river data collected by different authors.

Author	Number of datasets		
	Training	Test	Total
Colby and Hembree (1955)	17	8	25
Hubbell and Matejka (1959)	10	5	15
Nordin (1964)	28	13	41
Einstein (1944)	33	15	48
Total data	88	41	129

Table 4.6: Statistical distribution of training and test sets for river data from different authors.

Data	Parameters		Average	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	0.94	2.35	0.47	0.48
		depth (m)	0.44	1.46	0.09	0.33
		bed slope (mm/m)	1.28	1.93	0.76	0.36
		d_{50} (mm)	0.55	1.00	0.16	0.36
	Output	total bed-material (ppm)	830.41	2440.00	101.70	691.28
Test	Input	mean velocity (m/s)	0.88	2.38	0.47	0.47
		depth (m)	0.43	1.25	0.1	0.31
		bed slope (mm/m)	1.26	1.80	0.74	0.35
		d_{50} (mm)	0.54	1.00	0.19	0.36
	Output	total bed-material (ppm)	697.11	2296.00	116.10	580.92

4.3 Field Data from the German Federal Institute of Hydrology

The second group of the field data was obtained for two of the largest rivers in Europe. A comprehensive dataset for the Rhine and Elbe rivers is provided by the German Federal Institute of Hydrology (BfG). Measured values of mean velocity, water surface slope, river bed width, median particle size (d_{50}), water depth, bed load, and total bed-material load are available along the two river reaches.

The German Federal Institute of Hydrology (BfG) is responsible for regularly collecting storing, and analyzing bed load and suspended load measurement data obtained by cross section measurements together with other morphological data. The BfG runs a dense measurement network with historical data back to the year 1974. The following section summarizes the method of bed and suspended load measurement used by the BfG as described in Spreafico and Lehmann (2009).

Bed Load Measurement

Bed load measurements are carried out with the bed load sampler "Arnheim-Koblenz" which is an in-house development of the BfG. It consists of a heavy frame, a sampling basket with a mobile rectangular mouth (16 to 8 cm), and a diffuser between the mouth and the basket. A video camera is also mounted above the mouth to control the correct hub of the sampler and the undisturbed inflow of sediment (Spreafico and Lehmann, 2009). The cross section is divided into vertical sections where the vessel is anchored successively for taking measurement samples. The number of vertical sections depends on the width of the river at the station, and usually five to eight sections are defined. The bed load sampler is lowered to the bed surface for five to fifteen minutes, at least three times at each vertical. After each measurement the trapped sediments are collected, then weighing and sieve analysis is performed in a laboratory. Figure 4-1 shows the "Arnheim-Koblenz" bed load sampler which is used by the BfG.

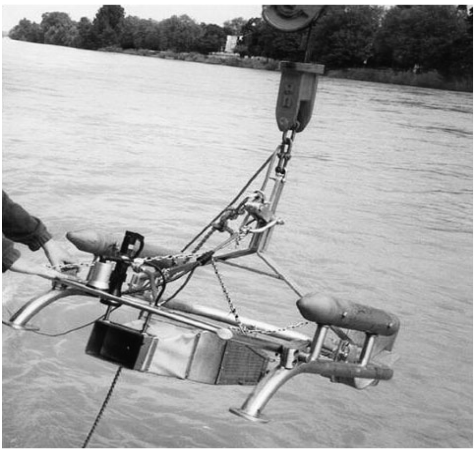


Figure 4-1: Bed load sampler "Arnheim-Koblenz" (BfG, 2008).

Suspended Load Measurement

For measuring suspended load at the permanent suspended load monitoring stations, a sample of five litres of water is taken every day at a point which is located in the middle of the river and approximately 50 cm below the water surface. Then the samples are filtered and sediment concentration is determined by drying and weighing the residuum.

At the cross-sections where bed load measurements are carried out, parallelly suspended load measurements are performed at the same verticals in four to five depth levels below the water surface. This gives detailed information about the distribution of sediment concentration over the whole cross section. A water sample of 50 litres is pumped at each

measuring point, and particles with size greater than 63 micro meters are separated by using an appropriate sieve. This is done in order to remove the portion of wash load from the suspended load. The collected sample is taken to a laboratory and sieve analysis is performed. The sum of bed load and suspended load (excluding wash load) is the total bed-material load used in this study. A detailed description of the measuring procedure and the equipments used can be referred from (Spreafico and Lehmann, 2009).

4.3.1 River Rhine

The Rhine river is the second largest river of Central Europe. It runs between the Swiss Alps and the North Sea crossing six countries (Switzerland, Austria, Liechtenstein, Germany, France and The Netherlands). It covers a length of 1320 km and a catchment area of 190,000 km^2. More than half of the catchment area lies in Germany. The Rhine is one of the most intensively used rivers on earth. The most important uses are navigation, hydropower, industry, drinking water, and recreation. It is one of the busiest shipping routes in the world. The Rhine river has an annual mean discharge of 2,290 m^3/s when it reaches the Dutch border. Different sediment management strategies are implemented along the Rhine including artificial sediment feeding and dredging operations in order to maintain the navigation efficiency of the river.

Bed load and suspended load are measured for the German part of the Rhine regularly at more than 40 cross sections between Iffezheim and the German-Dutch border. Sediment transport measurement is done at each station four to five times a year (Spreafico and Lehmann, 2009). Additional suspended sediment concentration measurement is done at eleven permanent stations between Lake Constance and Emmerich/Lobith. Particle size analysis is performed for sediment samples from the river bed and those collected by cross-sectional measurement of bed load and suspended load. The river cross-section is surveyed for the Rhine by echo sounding every two years from Iffezheim to the German-Dutch border.

The Rhine river with the different sections from its source to the North Sea is shown in Figure 4-2 (Alpine Rhine, High Rhine, Upper Rhine, Middle Rhine, Lower Rhine, Delta Rhine). The first and last measuring stations in the final datasets are marked with the pins on the figure. The datasets provided by the BfG contain measured bed load and suspended load data from stations of the Upper, Lower and Middle Rhine.

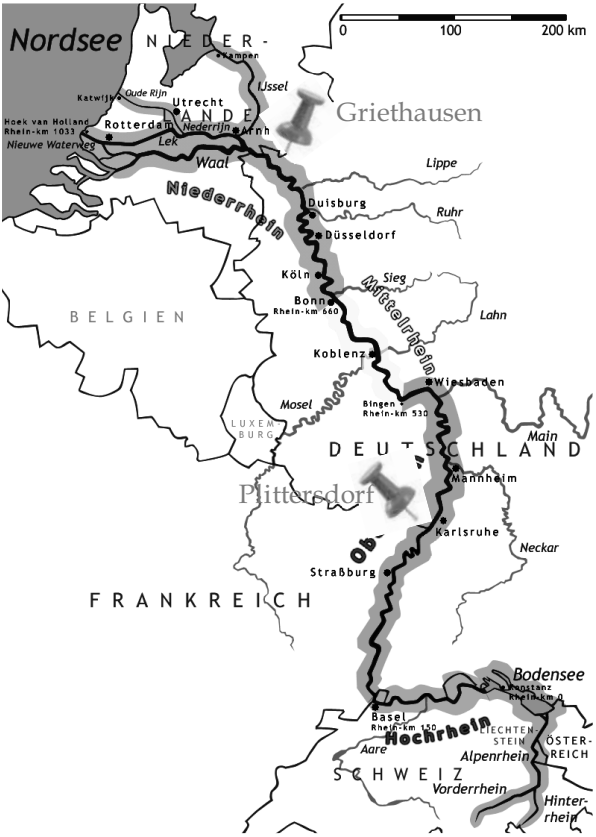


Figure 4-2: River Rhine (after Daniel Ullrich, 2005).

The available data is of 45 stations along the river from Plittersdorf to Griethausen covering a length of more than 500 km. The Rhine is a gravel bed river, specially the Upper Rhine, where a considerable portion of the sediment load is transported as bed load. The developed model will be applicable for computing sediment transport rates between the first and last stations. Detailed data analysis and filtering is performed to obtain the final datasets. There are two groups of data for the Rhine which are used to derive bed load and total bed-material load fuzzy models.

After serious data analysis and excluding those where one of the input variables is missing (see Table 4.7), a total of 560 datasets for bed load and 510 datasets for total bed-material load are selected for developing the models for the Rhine river. Stations with less than three measurements are not included in the analysis. From the total bed-material load datasets,

values with transport rates less than 20 g/s.m and greater than 1000 g/s.m are considered as extreme values with less representation and therefore are excluded from the input datasets. Similarly, from the bed load datasets, values with transport rates less than 5 g/s.m and greater than 100 g/s.m are not included. The statistics of the training and test datasets implemented for total bed-material load and bed load data-driven fuzzy models are summarized in Tables 4.7 and 4.8 respectively. The total bed-material and bed load transport rates are indicated in mass per unit time per width (g/s.m). The other input variables are: velocity (m/s), depth (m), water surface slope (m/m), and median sediment particle size (m). The technique used for selection of the training and test datasets is explained in section 4.4.

Table 4.7: Statistical distribution of input and output variables for the training and test data of total bed-material load model for the Rhine.

Data	Parameters		Mean	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	1.37	5.16	0.61	0.46
		depth (m)	3.96	6.98	0.94	0.98
		bed slope (m/m)	0.00019	0.00045	0.00000	0.00010
		d_{50} (m)	0.00749	0.01993	0.00057	0.00554
	Output	total bed-material (g/s.m)	154.68	729.63	20.51	122.80
Test	Input	mean velocity (m/s)	1.35	2.72	0.54	0.38
		depth (m)	3.96	7.18	2.08	1.00
		bed slope (m/m)	0.00018	0.00045	0.00000	0.00010
		d_{50} (m)	0.00748	0.01904	0.00055	0.00550
	Output	total bed-material (g/s.m)	155.96	658.30	24.17	110.54

Table 4.8: Statistical distribution of input and output variables for the training and test data of bed load model for the Rhine.

Data	Parameters		Mean	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	1.46	4.21	0.81	0.37
		depth (m)	4.16	8.18	0.94	1.15
		bed slope (m/m)	0.00020	0.00045	0.00000	0.00010
		d_{50} (m)	0.00862	0.02273	0.00200	0.00544
	Output	bed load (g/s.m)	20.08	99.89	5.01	16.66
Test	Input	mean velocity (m/s)	1.47	3.79	0.76	0.39
		depth (m)	4.03	7.66	2.12	1.12
		bed slope (m/m)	0.00020	0.00044	0.00000	0.00010
		d_{50} (m)	0.00901	0.02326	0.00203	0.00556
	Output	bed load (g/s.m)	20.30	86.59	5.08	17.75

4.3.2 River Elbe

The Elbe river is the third largest river of Central Europe. It flows from the Krkonose Mountains (Czech Republic) to the North Sea covering a total distance of 1091 km and a catchment area of 148,268 km^2. The Elbe river basin spans four countries: two third of the basin lies in Germany, one third in the Czech Republic, and less than 1% in Austria and Poland. The mean annual discharge into the North Sea is 877 m^3/s. The Elbe basin is a region with dense population, highly developed industry, and intensive agriculture.

The data obtained from the BfG covers the German part of the river. Measured parameters are available for 20 stations, with a total length of more than 500 km from the first to the last station. The first station in the final dataset is Dresden and the last measurement station is Neu Darchau. The Elbe is a sandy river where most sediment particles are transported in the form of suspended load. Here again, data analysis and removal of extreme values is done, wash load is excluded and training and test datasets are selected. This results in a total of 321 and 356 datasets for total bed-material load and bed load respectively for final analysis. The statistics of the training and test datasets for the two categories of data are indicated in Tables 4.9 and 4.10.

Figure 4-3 shows the Elbe River basin from its source to its mouth into the North Sea. The first and last measuring stations in the final dataset are also indicated in the figure.

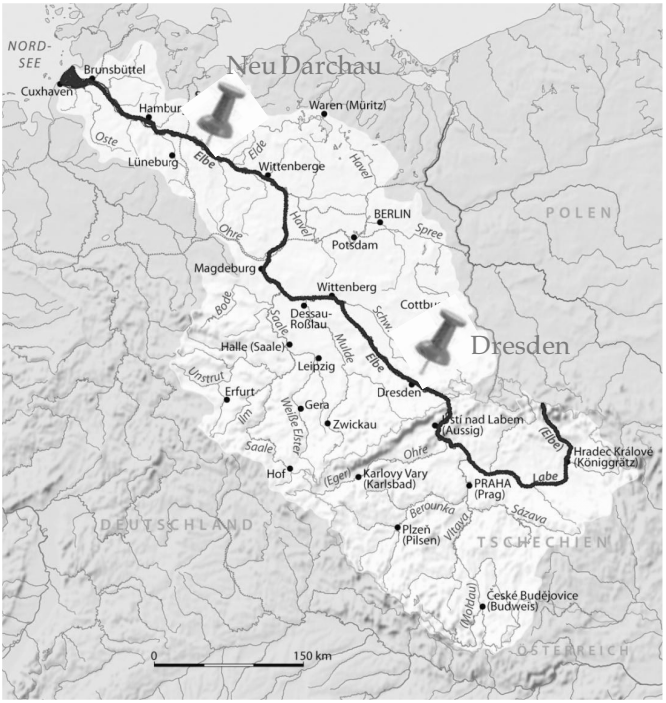


Figure 4-3: The Elbe river basin (after Titus Groan, 2008).

Table 4.9: Statistical distribution of input and output parameters for the training and test data of total bed-material load model for the Elbe.

Data	Parameters		Average	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	1.00	2.36	0.08	0.29
		depth (m)	2.77	5.07	0.69	0.71
		bed slope (m/m)	0.00019	0.00035	0.00013	0.00005
		d_{50} (m)	0.00125	0.00853	0.00050	0.00114
	Output	total bed-material (g/s.m)	123.54	595.98	20.15	93.85
Test	Input	mean velocity (m/s)	1.03	3.89	0.51	0.39
		depth (m)	2.80	4.72	0.38	0.73
		bed slope (m/m)	0.00019	0.00033	0.00013	0.00005
		d_{50} (m)	0.00129	0.00788	0.00051	0.00128
	Output	total bed-material (g/s.m)	130.77	464.75	28.46	88.67

Table 4.10: Statistical distribution of input and output parameters for the training and test data of bed load model for the Elbe.

Data	Parameters		Average	Maximum	Minimum	St. dev
Training	Input	mean velocity (m/s)	1.02	2.50	0.51	0.28
		depth (m)	2.87	5.53	1.07	0.72
		bed slope (m/m)	0.00018	0.00034	0.00013	0.00004
		d_{50} (m)	0.00122	0.00732	0.00053	0.00095
	Output	bed load (g/s.m)	39.79	96.64	15.02	20.37
Test	Input	mean velocity (m/s)	1.00	1.97	0.60	0.25
		depth (m)	2.88	4.89	1.63	0.63
		bed slope (m/m)	0.00018	0.00035	0.00012	0.00004
		d_{50} (m)	0.00116	0.00563	0.00051	0.00080
	Output	bed load (g/s.m)	40.74	98.23	15.08	22.22

4.4 Training and Test Datasets

The selection of training and test datasets is very important for any data-driven modelling. In data-driven modelling, the datasets should always be divided into training and test datasets. The reason for using test datasets for the evaluation of the model validity is that after some point in the training process over-fitting may occur on the training data. The system could be over-trained losing its general predictive ability and becomes very specific to the datasets used for model training. The test datasets are used to avoid overtraining by checking the performance of the optimized models at different iteration numbers. The number of training data should contain all possible combinations of events expected for the system to be modelled. For sediment transport, data should be collected in different seasons ranging from high to low flow, which mainly affects the amount of sediment under transportation. This is important to improve the generalization capacity of the model which refers to the ability of a data-driven model to produce a reasonable output for the combination inputs not encountered during training.

In this study, two third of the available datasets are used for training and one third for testing. In data-driven modelling, it is required that the training and testing datasets have more or less comparable statistical distributions to ensure that they come from similar populations. In order to maintain statistical homogeneity between the training and test sets, two consecutive data points for training from each dataset (author or station) followed by one data point for testing are chosen (Bhattacharya et al., 2005). If there are sufficient datasets available, creating a third data group called validation or evaluation dataset is helpful to further validate the developed model. For laboratory data in the sand size range, there are a

total number of 669 training sets and 344 test sets prepared for model development. The numbers for each category of data are presented in the previous sections.

Although there is a considerable difference between the hydraulic characteristics of the Rhine and Elbe rivers along the stations from the first to the last station, all the datasets from the stations are combined together for each river in order to develop general models which are applicable for the whole reach of each river. The different stations, the number of training and test data from each station, and the total number of final bed load and total bed-material load datasets for the Rhine and Elbe rivers are included in Appendix A.

5 Sediment Transport Modelling Using Data-Driven Fuzzy Logic Approach

5.1 Introduction

For the process of sediment transport it is complicated to derive mathematical equations which describe the physics accurately and data-driven modelling is a powerful alternative tool for such processes. Existing knowledge about the process can be used for selecting input and output parameters in data-driven modelling. Sufficient data is required to be collected to identify the relationship between input and output variables and different optimization techniques are used to optimize the model. Machine learning deals with the derivation of models from experience and collected data (Mitchell, 1997). It is especially attractive for modelling processes which are complicated and there is not sufficient knowledge to express them mathematically with acceptable accuracy. Modelling techniques such as artificial neural networks (ANN), fuzzy logic, decision trees, and hybrid neuro-fuzzy approaches are some examples of data mining methods. If a set of input and output parameters are given, a data mining algorithm generates a functional relationship between the input variable values and target outputs (Bhattachary et al., 2007). This chapter gives an introduction to data-driven modelling and presents the steps followed in the application of data-driven fuzzy logic modelling for computing sediment transport. Fuzzy rules derived from expert knowledge are called fuzzy expert systems. Fuzzy models can also be used to capture complex relationships between input and out parameters if sufficient data is available, and these fuzzy rules are called data-driven fuzzy systems.

5.2 Data-Driven Modelling

Data-driven modelling, also called data mining or machine learning, has been attractive in many fields in water resources. Data-driven modelling is utilized to identify hidden relationships that exist in datasets by implementing different optimization algorithms. It is based on the investigation of collected data characterizing the process under consideration.

The unknown relationship between input and output parameters can be identified with limited assumptions about the physical behaviour of the system (Solomatine and Ostfeld, 2008). The most frequently used data-driven modelling techniques are fuzzy-rule based systems and artificial neural networks. The popularity of fuzzy models is due to their ability to roughly represent complex, imprecise, or approximate relationships which are difficult to describe using precise mathematical models (Hammell II and Sudkamp, 1998). Data-driven fuzzy models have two advantages. These are the ability to express a system with a natural language and the potential to approximate complex non-linear parameters by analysis of input-output data. The resulting linguistic expressions are more transparent if the fuzzy model is developed from expert knowledge. In general, data-driven modelling is more explanatory, and patterns discovered in datasets are usually descriptive in nature rather than being predictive (Hullermeier, 2005). The basic steps followed in developing data-driven models are: analyze the problem to be modelled, collect data, select model structure, build the model, optimize, and test the model (Solomatine and Ostfeld, 2008).

The flow chart in Figure 5-1 summarizes the basic steps followed in this dissertation to develop and validate data-driven adaptive neuro-fuzzy models for computing both bed load and total bed-material load transport rates. The modelling procedure can be summarized into four general steps. The first step is the identification of input and output variables and data preparation which includes: selection of input and output variables, collection of sufficient data, data analysis and filtering, and preparation of training and test datasets. The second step is the generation of initial membership functions; grid partitioning or fuzzy clustering can be used to define the initial membership functions and rules. The third step is the model optimization using ANFIS and performing a sensitivity analysis for the different parameters of the model. The final step is the model validation and testing using additional datasets. Model validation can also be performed by comparing the results of the ANFIS model with results of other commonly applied models.

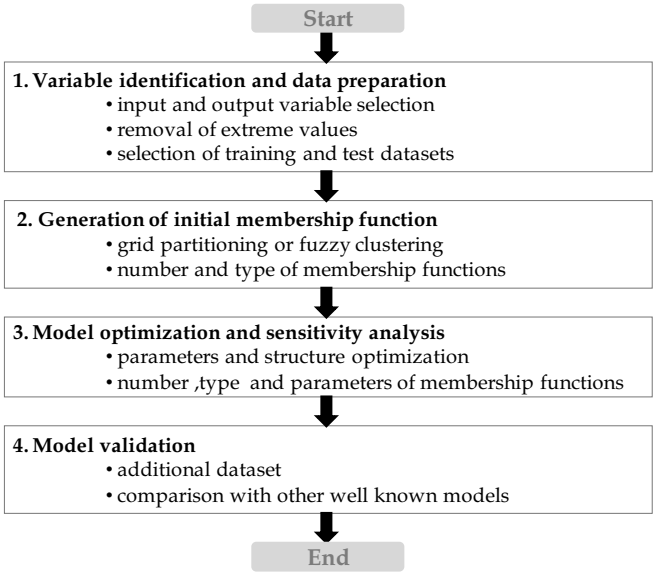


Figure 5-1: Steps followed for developing data-driven ANFIS models.

5.3 Selection of Input Variables

The first step for developing fuzzy rules from observed data is selecting input and output variables. In data-driven modelling selection of appropriate input variables is very important and requires the incorporation of the domain knowledge in the model. The most significant factors affecting sediment transport capacity are identified and used for constructing the fuzzy model. The general parameters governing the process of sediment transport can be described by (Yalin, 1977):

$$q_t = f(V, D, d, S, g, \rho_s, \rho_w, \mu) \tag{5.1}$$

where V= velocity, D = depth of flow, d = particle size (mainly d_{50}), S = energy slope, g = acceleration due to gravity, ρ_s = density of sediment, ρ_w = density of water, μ = viscosity of water and q_t = total bed-material transport rate.

Many equations are formulated by different researchers to compute the rate of sediment transport and these equations use different parameters as dominant factors influencing the process. The main parameters needed to estimate sediment transport rate are related to the

properties of flow conditions and sediment mixture. There are several flow variables which are relevant for the motion of sediment and the three important variables describing the flow condition are depth, velocity, and slope. These variables are important factors needed for the computation of shear stress, unit stream power or stream power which are used as the dominant parameters in many sediment transport equations. The amount of transported material depends on discharge which is a function of flow velocity and water depth. The water depth also determines the amount of materials in suspension. Sediment transport is influenced by flow velocity. Water flowing with high velocity exerts large force or energy on sediment particles on the river bed, and has higher tendency to transport sediment particles. The slope or energy gradient represents the loss of energy to overcome friction and move sediment particles. The flatter the slope, the lesser the rate of energy dissipation and this results in the transport of less material. Yang (1973) proved that the transport rate is closely related to unit stream power. The unit stream power is the product of velocity and slope. The most significant property of a sediment mixture is its particle size. For initiation of motion of sediment particles, the resistance force should be balanced by the drag force. The drag force is the shear stress which is exerted on the sediment particle and it is a function of slope and depth. The resistance to movement depends on the physical properties of the particle of which size is the most important. Particle size determines the criteria for incipient motion whether it is based on critical shear stress or critical stream power, or critical unit stream power.

As described above, most of the parameters used in many sediment transport equations are functions of the four fundamental parameters: bed or energy slope (S), mean flow velocity (V_m), water depth (D), and the median particle diameter (d_{50}). These parameters are selected as the main important variables governing the process of sediment transport. Additionally, the four basic variables are selected because they can be easily measured as primary data, and their physical meaning is obvious for practical applications of the models to be developed. Dimensionless variables which are used in many sediment transport equations could also be defined as alternative input parameters.

In order to successfully implement the approach, sediment transport data is collected from different sources and then the available data is categorized into two groups: training and testing. The training datasets are used to train the model and obtain best fitting parameters by implementing optimization algorithms. The test datasets are used to validate the model and avoid overtraining or over-fitting. The data available for the analysis and the training and test datasets are summarized in chapter four.

5.4 Generation of Initial Membership Functions

After analyzing the problem to be solved, collection of data and selection of input and output variables, the next step in fuzzy modelling is to define initial membership functions for the input and output variables. Data-driven fuzzy rule generation can be divided into two main steps. These are rule induction, and rule-base optimization. Fundamentally, there are two techniques for defining initial membership functions for rule induction (Guillaume, 2001). The first method uses grid partitioning of the multidimensional space. The grid partitioning can be generated from data or defined by experts. The number of fuzzy sets (partitions) for each variable has to be defined and linguistic labels can be associated with the fuzzy sets. A training procedure optimizes the grid structure as well as the rule consequences according to the available collected data. The second technique for generating initial membership functions for rule induction is fuzzy clustering. The training data pairs are categorized into homogeneous groups (clusters) and a rule is associated to each cluster. The fuzzy sets are not shared by the rules, and each cluster has its own characteristic and is described by one particular rule (Guillaume, 2001).

5.4.1 Grid Partitioning

The initial fuzzy model can be generated by using a fixed fuzzy partition for each of the input variables, and each cell of this grid is considered as a potential antecedent part of a rule. Equally shaped and equally spaced membership functions are defined for each of the input variables based on the ranges of measured data. This approach is advantageous from the interpretability point of view. An effective data partitioning in input–output space can lead to reducing the number of rules and thus improving the computational efficiency and interpretability of fuzzy models (Chen and Likens, 2004). In grid partitioning, the membership functions are defined independently for each variable. It does not consider any potential relationship which might exist between the input variables (Vernieuwe et al., 2005), and this is one of the drawbacks of using grid partitioning to generate initial fuzzy sets. Sensitivity analysis is usually helpful to obtain the optimum number of partitions (fuzzy sets) and type of membership functions for the input variables. Membership functions with different shapes are available. Depending on the range of values in the training datasets, the optimum number of fuzzy sets for each variable is obtained by sensitivity analysis. In general, defining 2 to 5 partitions for each variable is usually sufficient (Guillaume, 2001).

All possible combinations of antecedent fuzzy sets are covered in the grid partitioning approach. Three generalized bell-shaped membership functions are defined as initial membership functions for each of the input variables velocity, depth, slope, and median particle diameter. Since every combination of the antecedent fuzzy sets is covered, this results in 3x3x3x3 (81) initial rules (output equations) for the model to be optimized. The same procedure is followed and initial membership functions are generated for the

laboratory data in the sand size range, gravel laboratory data, field data of total bed-material transport obtained from different authors, bed load, and total bed-material load datasets for the Rhine and Elbe rivers obtained from the BfG.

Figures 5-2 and 5-3 show the initial membership functions generated for laboratory data in the sand range, and total bed-material sediment transport model for river Rhine. In the figures, the blue curves represent the first initial membership functions (mf_1) for each of the four input parameters, the green lines the second membership functions (mf_2), and the red graphs represent the third membership functions (mf_3) for each of the four input variables. Linguistic values can be defined for these membership functions. The parameters of these membership functions which determine the shapes will be changed during model optimization. From Figure 5-2, it can be observed that for laboratory data in the sand size range, the measured velocities are between 0 and 2 m/s, the depth between 0 and 0.8 m, the bed slope 0.0005 - 0.025, and the particle size 0.1 - 2 mm.

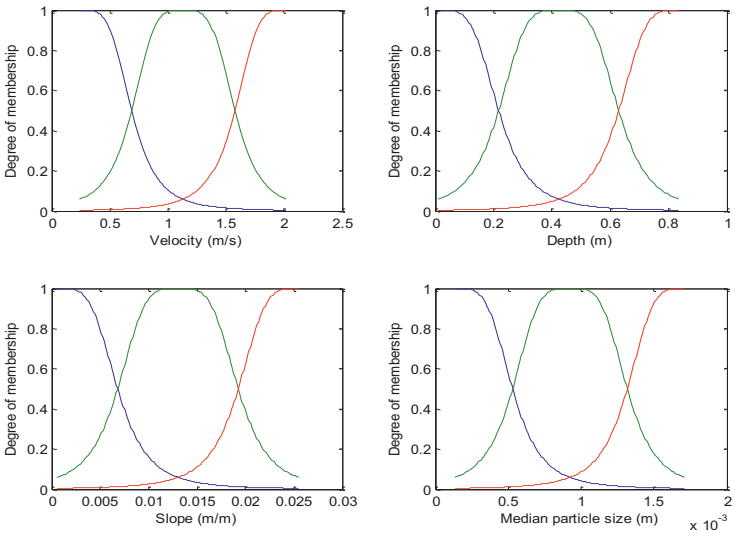


Figure 5-2: Example of generalized-bell shaped membership functions generated for laboratory data in the sand range.

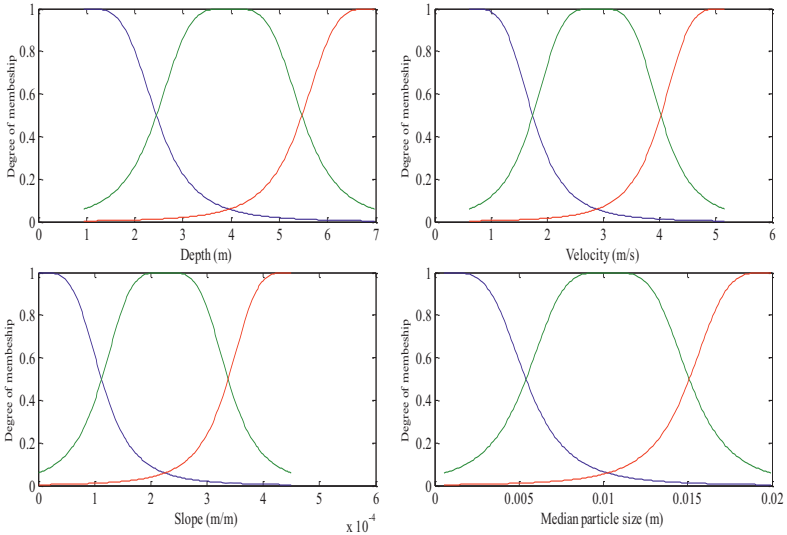


Figure 5-3: Initial generalized bell shaped membership function generated for total bed-material load ANFIS model of the Rhine river.

For the initial fuzzy system shown in Figure 5-3, there are 81 possible combinations of antecedent membership functions and the resulting rules for the first order TS fuzzy system are expressed as:

Rule 1:

IF (depth is in_1mf_1) AND (velocity is in_2mf_1) AND (slope is in_3mf_1) AND (d_{50} is in_4mf_1) THEN (total bed-material load is $out_1mf_1 = a_{1,1}$*depth+$a_{1,2}$* velocity +$a_{1,3}$* slope +$a_{1,4}$*d_{50}+$a_{1,0}$)

Rule 2:

IF (depth is in_1mf_1) AND (velocity is in_2mf_1) AND (slope is in_3mf_1) and (d_{50} is in_4mf_2) THEN (total bed-material load is out_1mf_2= $a_{2,1}$*depth+$a_{2,2}$* velocity +$a_{2,3}$* slope +$a_{2,4}$*d_{50}+$a_{2,0}$)

.

.

.

Rule 81:

If (depth is in_1mf_3) AND (velocity is in_2mf_3) and (slope is in_3mf_3) AND (d_{50} is in_4mf_3) THEN (total bed-material load is out_1mf_{81}= $a_{81\ 1}$*depth+$a_{81,2}$*slope+$a_{81,3}$*velocity+$a_{81,4}$*d_{50}+$a_{81,0}$)

In which in_1mf_1 represents input 1 (depth) and membership function 1, in_2mf_1 is the membership function 1 for input 2 (velocity), etc. There are three parameters for each

generalized bell-shaped membership function and there are four variables and three membership functions for each variable. This results in a total of 3*3*4=36 premise non-linear parameters for the input membership functions. The total number of rules is 81 and there are five ($a_{i,1}$, $a_{i,2}$, $a_{i,3}$, $a_{i,4}$, $a_{i,0}$) parameters for each rule or output membership functions resulting in a total of 81*5= 405 consequent linear parameters. These premise and consequent parameters are modified during the model optimization.

5.4.2 Fuzzy Clustering

Clustering is the process of identifying natural behaviour in a dataset and putting similar datasets into homogeneous groups or clusters (Bezdek et al., 1984; Hammouda and Karray, 2000). Fuzzy clustering is one recognized clustering technique to generate initial fuzzy models automatically from data. Fuzzy clustering algorithms are used to organize and categorize data by dividing the datasets into different fuzzy clusters. The clusters are fuzzy because the groups overlap. Each cluster represents a specific behaviour in the datasets collected for the system to be modelled. Clustering partitions a dataset into several groups such that the similarity within a group is larger than the similarity among groups (Chen and Likens, 2004). The number of clusters is the same as the number of rules. This method creates a very compact rule system because only one rule is associated with each cluster. The fuzzy rules are not shared between clusters. There are different types of available clustering techniques such as fuzzy c-means clustering (Bezdek, 1973), mountain clustering (Yager and Filev, 1994), Gustafson-Kessel fuzzy clustering (Gustafson and Kessel, 1979), and subtractive clustering (Chiu, 1994).

The general approach of all the clustering techniques is to find cluster centres that represent each cluster. A cluster centre indicates the central position of each cluster. If an input dataset is provided, the relationship of the variables to the different clusters is measured by its distance from the cluster centres. The shorter the distance between a cluster centre and a given data, the more related it is to this cluster and vice versa. In Fuzzy c-means clustering and K-means clustering the number of clusters should be known in advance to apply the algorithms (Hammouda and Karray, 2000). A trial and error procedure can be implemented to identify the optimum number of clusters in these methods. Subtractive clustering and mountain clustering algorithms can determine the number of clusters required to describe the behaviour of the dataset automatically.

The results from applying a clustering algorithm are: a set of clusters, cluster centres and a fuzzy partition matrix with membership degrees. After the clustering process, the membership degrees of a new data object to the computed clusters can be calculated using the cluster centres and the fuzzy partition matrix. The most widely used clustering algorithm is the fuzzy c-means clustering (FCM) due to its efficiency and simplicity (Chen and Likens,

2004), and is selected for application in this dissertation. The mathematical computation of FCM is described in the section below.

The Fuzzy C-Means Clustering

The fuzzy c-means clustering was proposed by Bezdek in 1973. The FCM algorithm partitions a collection of data points specified by n-dimensional vectors into c fuzzy clusters, and finds a cluster centre in each by minimizing an objective function (Jantzen, 1998). The cluster centres are determined by optimizing an objective function (minimizing a cost function). An element can belong to more than one cluster with different degrees of membership grades according to its distance from the cluster centres. The mathematical formulation of the FCM algorithm is described in detail in (Bezdek, 1980, 1981; Bezdek et al., 1984) and is summarized as follows.

A set of n vectors $x_j, j = 1,...,n$, are to be partitioned into c partitions. The membership matrix U has elements with values between 0 and 1. The sum of the degree of membership of an element to all clusters should be always 1.

$$\sum_{i=1}^{c} u_{ij} = 1, \forall j = 1..., n. \tag{5.2}$$

The objective function which is used to partition the dataset into different clusters is:

$$J(U, c_1, ..., c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2, \tag{5.3}$$

where:

n is the number of input vectors, u_{ij} is membership degree between 0 and 1, c_i is the centre of the cluster i, x_j are the input training data vectors ;

$d_{ij} = \|c_i - x_j\|$ is the Euclidean distance between the ith cluster centre and the jth data point; and m ε[1,∞) is a weighting exponent to adjust the membership degree weighting effect.

Parameter m controls the fuzziness of the clusters, and affects the degree of overlapping between the multidimensional fuzzy clusters; as m increases, the overlapping degree or fuzziness also increases.

The necessary conditions for equation 5.3 to be optimum (reach its minimum) are:

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{5.4}$$

and,

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \tag{5.5}$$

The following steps are followed by the algorithm of FCM to determine the cluster centres c_i and the membership matrix U (Hammouda and Karray, 2000):

1 Initialize the membership matrix U with random values between 0 and 1 such that the constraints in equation 5.2 are satisfied.
2 Calculate the cluster centres with equation 5.4.
3 Compute the objective function with equation 5.3, stoppage criteria can be reaching the tolerance value or if improvement with consequent iterations is less than a given threshold.
4 Calculate the new membership matrix with equation 5.5. Go to step 2.

The number of clusters is equal to the number of rules, because similarity within a dataset in a cluster is very strong compared to the similarity among clusters. Therefore clustering generates a very compact fuzzy rule system from data, this is especially advantageous where there are many parameters and grid partitioning creates too many rules. Clustering algorithms are popular due to their ability for sorting out and identifying complex interactions between variables in multidimensional data (Bezdek et al., 1984). The disadvantage of the clustering approach is that final clusters are usually difficult to interpret as the clustering is performed in a multi-dimensional space. The interpretability of the system significantly decreases and the approach tends to have more of a black-box nature. Application of clustering techniques to generate fuzzy rules can be referred in (Bruin and Stein, 1998; Priyono et al., 2005; Lohani et. al, 2007; Corduas, 2011). The default value of m is used in this research and the number of iterations for FCM clustering is set to 200.

After initial membership function generation using grid partitioning or fuzzy clustering, the next step is to optimize the data-driven fuzzy model.

5.5 Optimization of the Fuzzy Model

Model optimization is the process of fine tuning model parameters to increase the predictive accuracy of the model. Some of the available optimization methods include hill climbing,

Monte Carlo, simulated annealing, evolutionary algorithms, neuro-fuzzy or other hybrid approaches (Garibaldi, 2005). There are two kinds of optimization. These are parameter and structure optimization.

5.5.1 Parameter Optimization with ANFIS

Parameter optimisation adjusts aspects of the fuzzy model such as the shape and location of membership functions and the number and form of rules. Parameter optimization is the process of determining the parameters of the input and output membership functions. For the input membership functions, the parameters determine the shape and distribution of the fuzzy sets. In first order Takagi-Sugeno fuzzy models the parameters of the output membership functions are the coefficients of the linear equations describing the output membership functions. Garibaldi (2005) mentioned three properties which are usually required for an optimum rule base: continuity, consistency, and completeness. The continuity guarantees that small variations of the input do not induce big variations for the output. Consistency is the property that insures if two or more rules are simultaneously fired their conclusions are coherent. Completeness in the rule base is necessary to cover any possible combination of inputs and there is at least one rule describing it. In other words there shouldn't be any inference breaking.

In this dissertation, the adaptive neuro-fuzzy inference system which implements neural network learning algorithms to optimize fuzzy models is implemented for optimising the parameters of the fuzzy model. The hybrid learning algorithm in ANFIS (section 2.5), is the combination of the gradient descent and least-squares methods. The gradient descent method is employed to tune premise non-linear parameters, while the least-squares method is used to identify consequent linear parameters (Jang, 1993; Jang et al., 1997). The least mean square error is used as a model performance index in combination with the gradient descent technique to overcome the problem of getting trapped in local conditions in the parameter space which might happen if only the gradient descent scheme is implemented (Chen and Likens, 2004). To determine the final optimized model, different numbers of membership functions and iterations are carried out.

5.5.2 Avoiding Overtraining

During the optimization process, great care should be taken to ensure that the model will not be over-trained. An over-trained model loses its general predictive ability and becomes too specific to the training dataset rather than identifying the general behaviour of the system. The test dataset is used to avoid overtraining by checking the performance of the optimized models at different iteration numbers. Furthermore, the investigator should always keep in mind that there is a trade-off between accuracy and interpretability (Casillas et al., 2003) during model optimization. In fuzzy modelling, maintaining some degree of interpretability is essential. If only numerical accuracy is considered, the training algorithm can generate

final fuzzy sets which are very complicated and difficult to interpret. Casillas et al. (2003) stated that if interpretability is neglected, the final model will be nothing but a black-box description of input–output relationships and the advantage of fuzzy systems over neural networks might be completely lost.

Figure 5-4 represents an example where, after approximately 190 iterations, the model starts to be over-trained. The optimization should be stopped after that and the test dataset is necessary to identify this point. The final model should not be selected just based on the minimum training root mean squared error (RMSE), but the model which gives the minimum RMSE for the test set should be used for future applications.

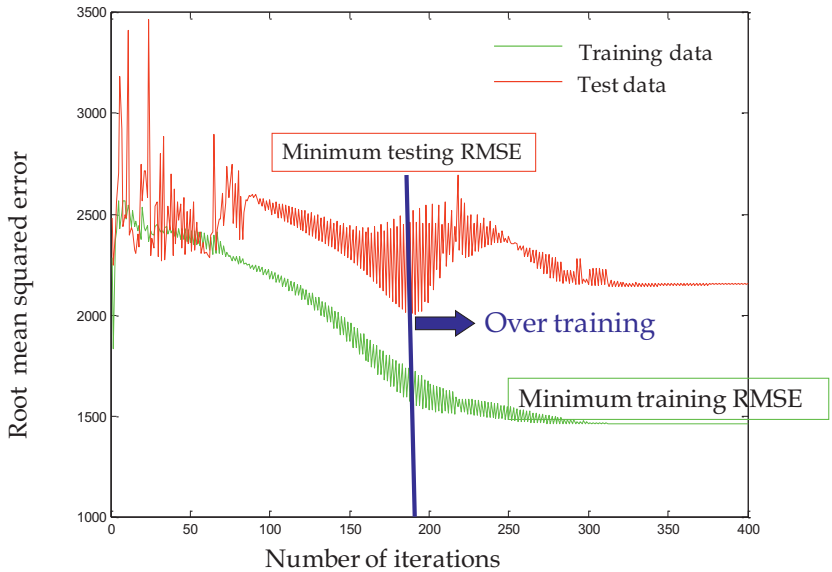


Figure 5-4: Graphical representation of overtraining.

5.5.3 Number of Iterations for Optimization

Selection of appropriate number of iterations is necessary during the model training process. The number of iterations determines the time required for model optimization. Time is a valuable consideration during model development and optimization.

As it is illustrated in Figure 5-4, the optimum number of iterations should be determined based on the performance of the model on the test dataset. If the training process is

continued without this constraint, it results in unrealistic optimized parameters just based on model performance on the training dataset.

Table 5.1 summarizes the general properties of an ANFIS model with four input and one output variables. If the number of membership functions for each of the input variables is three, this results in a total of 81 rules describing the system which are represented in the 81 output membership functions. In ANFIS, the product method is usually used for implication operation (to determine the degree of fulfilment of a rule) and the weighted average method is implemented for summation (combine outputs from each rule). For this first order TS fuzzy inference system, there are 36 premise (3*3*4) and 405 (81*5) consequent parameters.

Table 5.1: General properties of an ANFIS model with four input variables and 81 rules.

ANFIS parameter	Selected for application
TSK Type	first order
Number of inputs	4
Input labels	depth, velocity, slope, d_{50}
Number of outputs	1
Output variable	total bed-material sediment load
Number and type of input membership functions	3 generalized bell-shaped
Number and type of output membership functions	81 linear
Numbers of Rules	81
Number of training epochs	400
Number of nodes	193
Implication method	product
Defuzzification method	weighted average
Premise (nonlinear) parameters	36
Consequent (linear) parameters	405
Total fitting parameters	441

Optimizing the ANFIS model with the training data provides the final parameters of input and output membership functions. Based on these parameters, the output for any combination of input variables can be computed.

After parameter optimization, the optimal fuzzy rule-base is not yet finally constructed. The obtained fuzzy model may exhibit redundancy in terms of highly overlapping membership functions. To acquire an efficient and transparent fuzzy model, elimination of redundancy and making the fuzzy model as simple as possible is necessary by optimizing the structure of the model.

5.5.4 Structure Optimization

Structure optimization usually includes input variable selection, and rule base reduction. Simplification of fuzzy models is important to make the rule simple and interpretable. This can be achieved by reducing the number of fuzzy sets for each input variable and/or reducing the number of rules. Reducing the number of fuzzy sets for each input variable can be accomplished by removing redundant fuzzy sets and combining (merging) similar sets (Chen and Likens, 2004). The number of fuzzy rules can also be reduced by eliminating rules which are not very significant and combining two or more rules which are similar. There are different measures of similarity and can be referred from Chen and Likens (2004). There are two commonly used methods of rule base reduction. The first one consists of merging compatible elements: fuzzy sets, clusters, or input variables. The second family of methods is based upon statistic input domain transformation. Details can be referred from Guillaume (2001). After structural optimization, the obtained model is structurally simpler and interpretable.

Developing the fuzzy inference system using only the most important input variables and removing extra variables will increase interpretability and stability and leads to a more compact rule system, and this can significantly improve the transparency of the final model. Moreover, rule base and parameter optimization are easier to achieve once extra variables have been removed. Redundant variables have a tendency to bring more noise than useful information (Guillaume, 2001). Sensitivity analysis plays a key role in this regard.

In this dissertation, the results of sensitivity analysis (as presented in **Chapter 6**) are used for ranking the variables according to their influence on model performance, and the optimum number of membership functions (fuzzy sets) for each variable.

5.6 Model Performance Evaluation

The selection of an appropriate statistical model performance evaluation criterion is necessary during model optimization and for comparison of the accuracy of different models. Correlation coefficient (r), root mean squared error ($RMSE$), average absolute relative error ($AARE$) and discrepancy ratio (Dr) are selected as model performance evaluation criteria. The correlation coefficient measures the degree to which two variables are linearly related. The root mean squared error ($RMSE$) evaluates the difference between observed and computed sediment load. The $RMSE$ assumes that larger prediction errors have greater importance than smaller error values and measures the model performance with respect to high transport rates. The root mean squared error is more sensitive to occasional large errors because the squaring process gives disproportionate weight to very large errors. The lower the RMSE value, the better the model performance. A positive correlation coefficient indicates that the computed and observed values tend to go up and

down together. If the variables go in opposite directions it results in a negative correlation coefficient. The correlation coefficient provides information only about the linear dependence between measured values and corresponding computation results. It is not suitable for evaluating non-linear independence. The drawback of using it as model performance evaluation criteria is that it is insensitive to additive and proportional differences between model predictions and observed values. If a model consistently over or under predicts observed values, this can result in a high correlation coefficient which is misleading. Therefore, the correlation coefficient is not necessarily in agreement with other performance criteria such as the $RMSE$ and Dr.

The correlation coefficient (r), RMSE and AARE are computed as follows:

$$r = \frac{\sum_{i=1}^{N}(S_{oi} - \bar{S}_o) - (S_{ci} - \bar{S}_c)}{\sqrt{\sum_{i=1}^{N}(S_{oi} - \bar{S}_o)^2 * \sum_{i=1}^{N}(S_{ci} - \bar{S}_c)^2}} \tag{5.6}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_{ci} - S_{oi})^2} \tag{5.7}$$

$$AARE = \frac{1}{N}\sum_{i=1}^{N} Abs(RE) \tag{5.8}$$

$$RE = 100 * \frac{(S_{oi} - S_{ci})}{S_{oi}} \tag{5.9}$$

where RE is the relative error in prediction expressed as percentage, S_{oi} is the observed sediment load for the ith dataset, S_{ci} is the computed sediment load for the ith observation which is computed by ANFIS, $\bar{S}_c$ is the mean computed value, $\bar{S}_o$ is the mean observed value, N is the total number of observations. The $AARE$ not only gives the average performance index in terms of predicting sediment load but also the distribution of the computation errors.

Dr is computed by:

$$Dr = \frac{S_c}{S_o} \tag{5.10}$$

While evaluating the performance of the model for its predictive ability, it is important to analyze the distribution of the errors. The statistical performance evaluation criteria r and $RMSE$ do not provide any information on the distribution of errors. In order to test the robustness of the developed model, it is important to apply other performance evaluation criteria such as average absolute relative error and discrepancy ratio (Kisi, 2005). The RMSE is very sensitive to outliers (extreme errors), and also it is not scale free. To avoid the problem of scaling, it is recommended to implement Dr and AARE which are scale free error measurements. The mean absolute relative error calculates the error as a percentage of the measured value. The optimal value of AARE is zero, with low-magnitude values indicating better model performance. The limitation of ARE (Absolute Relative Error, see eq. 5.8 and 5.9) is that a small deviation in error can result in large changes in ARE, when calculating with small denominators. Few outliers can dominate the AARE. It is advantageous because it is scale free and is easily understandable. The Dr should be used only when computed values are positive. It cannot be used if the computed values are negative or contain zeros. If there are zero values in the observed dataset, there will be a division by zero. A Dr value of one indicates a perfect agreement between the computed and measured values. Dr is advantageous because it is not skewed towards extreme computational errors. Discrepancy ratio analysis of computed values is necessary to obtain the percentage of computed transport rates within certain ranges of discrepancies. For example the percentage of predicted values for which Dr is within 0.5-1.5 provides the number of model predictions which are between 0.5 times the observed values and 1.5 times the observed values.

To summarize this chapter, a modeller should always clearly understand what the main objective of the work is, just numerical accuracy or maintain some degree of interpretability. Interpretability is the main advantage of fuzzy systems over alternatives such as statistical models or neural networks (Mikut et al., 2005). Interpretability refers to the ability to understand the behaviour of the fuzzy system by examining the rule system. A trade-off between accuracy and interpretability is the main focus of recent research. Ishibuchi et al. (1997) describe a methodology to automatically generate compact fuzzy classification systems from numerical data by selecting a small number of important fuzzy rules using genetic algorithms. Mikut et al. (2005) illustrate a method for an automatic and complete design of fuzzy systems from data and improve the interpretability of the model. There are a number of algorithms which are available to automatically generate fuzzy rules from data. The different options should be carefully studied before selection for application.

6 Sensitivity Analysis and Model Results

This chapter summarizes the results of the sensitivity analysis of the different optimized ANFIS models developed using the datasets presented in chapter four.

6.1 Parameters for Sensitivity Analysis

Sensitivity analysis is a technique utilized to investigate effects of variations in model input parameters on the final accuracy of the model by systematically changing the parameters of the model. Performing sensitivity analysis is a key process in the development of successful and optimum mathematical models. In this dissertation, the effects of input variable combination, and number and type of membership functions on the performance of the ANFIS model are investigated.

6.1.1 Combination of Input Variables

A sensitivity analysis has been conducted to determine the relative importance of each of the four input parameters (depth D, slope S, velocity V, and median particle size d_{50}) on the accuracy of the model results, with one of the inputs removed at a time. Including the most dominant parameters significantly optimizes the structure of the fuzzy model. The different architectures of the models tried are: all four input parameters (D V S d_{50}), without depth (V S d_{50}), without flow velocity (D S d_{50}), excluding slope (D V d_{50}), and excluding particle size (D V S). Additionally, instead of using velocity and slope as separate parameters, the unit stream power (defined as the product of velocity and slope) is used as a single parameter. Unit stream power is the power available per unit weight of water and per unit width of the river reach. The analysis by Yang (1973) indicates that it has a very strong relationship to sediment transport rates compared to other conventional variables. The combination of unit stream power with water depth and grain size (VS D d_{50}) as well as with grain size only (VS d_{50}) resulted in a total of seven various combinations of input variables. Other possible combinations of the variables including dimensionless parameters as used in some sediment transport equations can be defined as alternative inputs. The main reason for using only the primary parameters is to construct the fuzzy model with simple hydraulic parameters which can be easily and directly measured and do not require further computation.

This is in agreement with the basic concept behind fuzzy logic based modelling which is to represent complex processes with simple fuzzy models. The output variable of the fuzzy model is either bed load or total bed-material load transport rate according to the measured data type.

6.1.2 Number and Type of Membership Functions

Number and type of membership functions defined in fuzzy models are important considerations for model optimization. There are different shapes of membership functions available to choose from. The number of fuzzy sets for each variable depends on the variation of the values within the dataset and is one of the key parameters for model optimization. Performing a sensitivity analysis and determining the optimum number of membership functions is important. Trapezoidal and triangular membership functions are usually used because of their simplicity and linear behaviour, especially for fuzzy expert systems. However, in this dissertation trapezoidal and triangular membership functions are not found to be suitable for optimization with ANFIS because the shapes are not flexible enough for changing parameters during optimization. The generalized bell-shaped and Gaussian membership functions show a comparable model performance. Finally generalized bell-shaped membership functions are chosen because of their simplicity and flexibility.

The number of membership functions (fuzzy sets) for each variable is varied from three to five. Defining too many membership functions makes the model too complicated because the number of fuzzy rules is proportional to the number of membership functions, whereas insufficient membership functions reduce model accuracy because the variation of the input parameters cannot be represented accurately. A large number of membership functions make the system more complex and difficult to interpret. Based on the results of the sensitivity analysis, three membership functions for each input variable are found to be the most efficient with respect to accuracy and model complexity. If four input variables are implemented and three membership functions are defined for each input variable, the total number of fuzzy rules (output linear equations in the case of Takagi-Sugeno fuzzy inference systems) is 3*3*3*3 (81) if every combination of antecedent fuzzy sets is considered (in the case of grid partitioning). However, if 5 fuzzy sets are defined for each of the input variables, the total possible combinations are 5*5*5*5 (625) rules which results in a very complicated fuzzy system. The number of premise and consequent parameters is proportional to the number of membership functions defined, and number of training datasets required is also proportional to the total number of adjustable parameters during model optimization. A model with more rules requires collection of more training data.

While applying fuzzy c-means clustering as an alternative technique for generating initial membership functions, the number of clusters is one of the parameters for sensitivity

analysis. The number of clusters is varied from three to five to determine the optimum number of clusters which is also equal to the number of fuzzy rules.

6.2 ANFIS Model Results for Laboratory Data

The model results obtained using different combinations of input variables in the ANFIS for the two categories of laboratory datasets are presented in this section. The procedure described in chapter five is followed to obtain the final optimized models.

6.2.1 Sand

For the laboratory datasets in the sand size range, three generalized bell-shaped membership functions are defined for each of the input variables and seven combinations of input variables are selected for sensitivity analysis. The different ANFIS architectures and the corresponding model accuracies for the laboratory datasets in the sand size range are shown in Tables 6.1 and 6.2. Scatter plots of measured versus computed sediment concentrations for ANFIS models with various combinations of input variables are shown in Figure 6-1.

In Table 6.1, the AARE (%), RMSE (ppm), r, and the percentage of data with ARE greater than 100% are summarized both for the training and test datasets. Discrepancy ratio analysis for the optimized ANFIS models showing the average discrepancy ratio, and percentage of computed values within discrepancy ratios between 0.75 - 1.25, 0.5 - 1.5, and 0.25 - 1.75 is presented in Table 6.2. The discrepancy ratio values within 0.75 - 1.25, 0.5 - 1.5, and 0.25 - 1.75 account for approximate model estimation errors of $\pm 25\%$, $\pm 50\%$, and $\pm 75\%$ respectively. The discrepancy ratio analysis is helpful to compare the accuracies of the various ANFIS models by identifying the distribution of the model estimation errors. The best results are those with average discrepancy ratio around 1.0 where the measured and computed values are comparable.

Table 6.1: Performances of ANFIS models with different combinations of input variables for laboratory data in the sand size range.

	Training data				Test data			
Input variables	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%
D V S d_{50}	22.4	1451.9	0.98	2.2	28.8	2233.8	0.95	4.7
V S d_{50}	22.6	1724.5	0.97	1.3	24.9	2010.6	0.96	2.6
D S d_{50}	31.3	1910.7	0.96	4.4	35.8	1936.9	0.96	7.3
D V d_{50}	46.5	3432.8	0.86	10.2	52.9	3898.4	0.85	10.2
D V S	37.1	2873.0	0.91	7.1	50.5	3172.2	0.90	8.7
VS D d_{50}	27.7	2310.9	0.94	3.8	29.6	2417.9	0.95	2.9
VS d_{50}	34.6	2291.8	0.94	5.9	34.2	2481.3	0.94	5.8

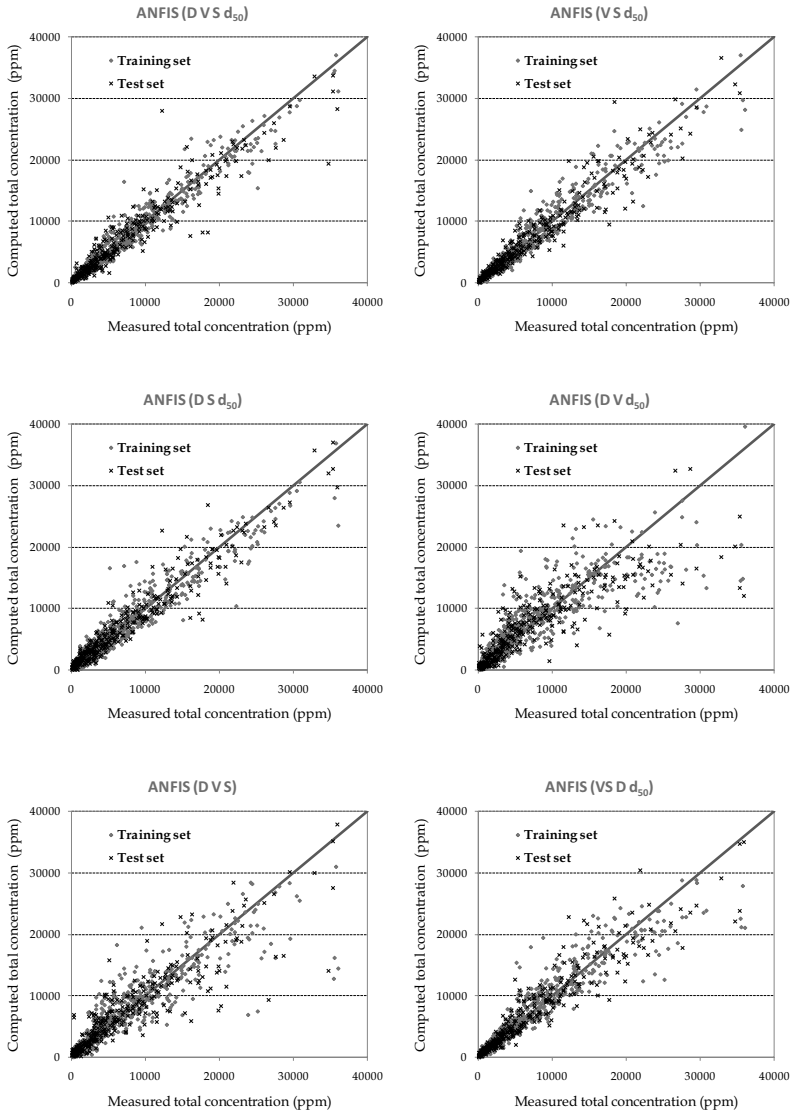


Figure 6-1: Scatter plots of measured versus computed total sediment concentrations for ANFIS models with various combinations of input variables for laboratory data in the sand size range.

Table 6.2: Discrepancy ratio analysis of different ANFIS models for laboratory data in the sand size range.

Input variables	Training data				Test data			
		Percent of data Dr with in				Percent of data Dr within		
	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.06	70.7	91.3	95.9	1.10	61.1	82.0	90.7
V S d_{50}	1.07	71.1	91.3	95.1	1.07	67.7	88.7	95.1
D S d_{50}	1.12	58.5	81.3	90.2	1.17	57.9	78.9	87.8
D V d_{50}	1.22	45.1	71.3	83.1	1.25	43.0	71.8	84.0
D V S	1.18	58.2	78.1	88.2	1.29	52.9	72.7	85.8
VS D d_{50}	1.11	64.7	85.7	91.8	1.09	59.3	84.3	91.0
VS d_{50}	1.18	62.0	85.7	91.0	1.15	56.4	85.2	91.6

The results of the sensitivity analysis (Tables 6.1 and 6.2, Figure 6-1) for the laboratory data in the sand range show that bed slope and particle size have the most significant effect on sediment transport rates. It can be observed from Figure 6-1 that the deviations from the measured values are high when these variables are not included. Depth is found to be having the least effect on model accuracy for the dataset used in the analysis and can be excluded from the input variables without reducing model accuracy considerably. This is because the depth variation in laboratory flumes is not significant. The model with three input parameters (V, S, and d_{50}) is found to be having comparable accuracy with that of the one with four input parameters (D, V, S, and d_{50}) by looking at the AARE (< 25%), RMSE, and r (r > 0.96, which indicates a very strong correlation) values both for the training and test datasets. Therefore it is chosen to be the most optimum model with an average discrepancy ratio of 1.07 for both datasets, and less than 3% of the total data have ARE greater than 100%. The percentage of data lying within discrepancy ratio of 0.5 - 1.5 is 91.3% and 88.7% for the training and test datasets respectively. This indicates a very good model performance. Figure 6-1 illustrates that the least accurate model is the ANFIS model without slope in the input variables with AARE greater than 45% both for the training and test datasets, and around 10% of the data have ARE greater than 100%. The analysis of the model outputs shows that large computation errors are associated with very high or low sediment transport rates. The quality of the input data for low sediment concentrations could also be low because of the difficulty in exact measurement of low sediment transport rates. This affects the accuracy of the developed model as data-driven models are highly sensitive to the quality of input data. The ANFIS model with the unit stream power (VS) as one parameter together with depth and particle size also shows a comparable accuracy with the one with four parameters resulting in AARE less than 30% and r greater than 0.94 for the total dataset. The ANFIS model with VS and d_{50} as input variables resulted in a very compact fuzzy inference system with nine rules and reasonably acceptable accuracy, but if three parameters are used it is better to use V and S separately. From the analysis of model results it can be concluded that

the data-driven fuzzy model successfully estimates observed total sediment concentrations. It is important to mention that the dataset contains measured sediment concentrations and other variables collected by eight different researchers (chapter 4, section 4.1.1) measured under different conditions. Model accuracy increases when the datasets from different authors are analyzed separately.

FCM clustering is also used to generate initial membership functions and a sensitivity analysis is performed for the combination of input variables and the number of clusters. The number of clusters is changed from three to five. The ANFIS model with five clusters performs better. The result of the sensitivity analysis for five combinations of input variables with 5 clusters is plotted in Figure 6-2. As it can be seen from Figure 6-2, the least accurate model is the model without slope in the input parameters with AARE greater than 50% and r less than 0.85 for the total dataset. Here also, slope and particle size are the dominant parameters and depth has the least effect on model accuracy by observing the AARE and r values from Figure 6-2. These results are in agreement with the results obtained by using grid partitioning for generating initial membership function as explained previously.

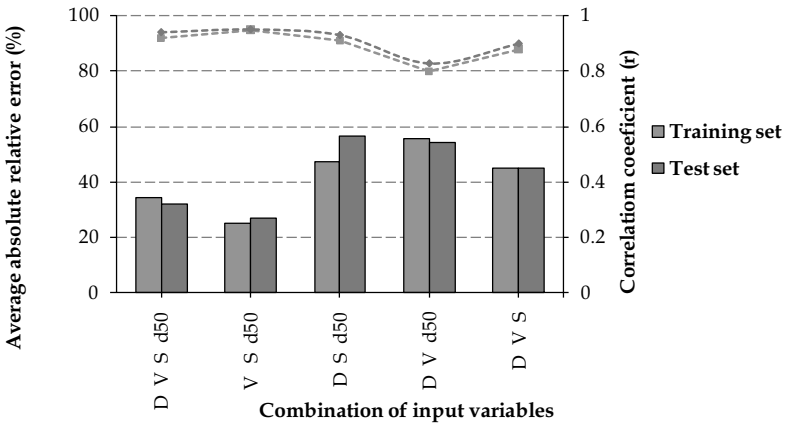


Figure 6-2: Performances of ANFIS models with various combinations of input variables and five clusters for laboratory data in the sand size range.

The model with five clusters (five rules) and three input parameters (V, S, and d_{50}) is the optimum model with AARE less than 28% and r = 0.95 both for the training and test datasets. The scatter plot of measured versus computed total sediment concentration shown in Figure 6-3 indicates that the model is capable of successfully estimating observed transport rates.

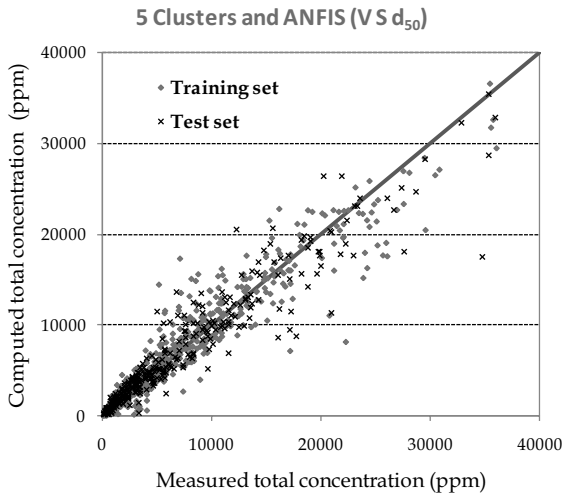


Figure 6-3: Scatter plots of measured versus computed total sand concentration for ANFIS model with three input variables (V S d_{50}) and five clusters.

The detailed model results using FCM for generation of initial membership functions are summarized in Appendix B and C for the model with four clusters and five clusters with different combinations of input variables.

6.2.2 Gravel

The gravel datasets contain 392 measured sediment concentrations collected by six researchers with median particle diameter greater than 2 mm and less than 30 mm (section 4.1.2). Grid partitioning and FCM clustering are implemented to generate initial membership functions.

Different ANFIS models with five combinations of input variables (see Table 6.3) and three, four and five clusters have been developed. The model with five clusters results in the highest accuracy and is therefore selected as the best model. The summary of model performance evaluation criteria for the ANFIS model with five clusters is presented in Tables 6.3 and 6.4. It can be seen that the model with four input parameters is selected to be the best model with AARE of 32.1% and 33.9%, correlation coefficients of 0.98 and 0.99, and average discrepancy ratios of 1.12 and 1.19 for the training and test datasets respectively. The RMSE values are 483.4 ppm for the training and 392.6 ppm for the test datasets, which are lower than the values from the ANFIS models with other input variable combinations. Less than

7% of the total dataset result in an absolute relative error > 100% and about 78% of the total data lie within a discrepancy ratio range 0.5 - 1.5. The ANFIS model, excluding depth from the input parameters, shows a comparable accuracy with 71.9% of the training and 64.3% of the test datasets lying within a discrepancy ratio range 0.5 - 1.5. The model without particle size in the input parameters (D V S) is the least accurate with AARE of 78.4% and 97.4% and an average discrepancy ratio of 1.51 and 1.71 for the training and test datasets respectively. The percentages of computed values with ARE greater than 100% are 18.2% for the training and 19.4% for the test sets. This result shows that the sediment particle size is a very important parameter for the computation of gravel transport where particle size varies considerably. The model results obtained by using four clusters and five combinations of input variables are included in Appendix D.

Table 6.3: Performances of various ANFIS models with five clusters for laboratory data in gravel size range.

Input variables	Training data				Test data			
	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%
D V S d_{50}	32.1	483.4	0.98	4.6	33.9	392.6	0.99	7.0
V S d_{50}	38.9	528.1	0.97	7.6	45.1	438.2	0.99	7.8
D S d_{50}	41.3	527.3	0.97	9.5	39.3	389.3	0.99	13.9
D V d_{50}	65.9	828.1	0.94	16.7	73.4	946.8	0.94	17.8
D V S	78.4	883.4	0.93	18.2	97.4	844.2	0.95	19.4

Table 6.4: Discrepancy ratio analysis of ANFIS models with five clusters for laboratory data in the gravel size range.

Input variables	Training data				Test data			
		Percent of data Dr with in				Percent of data Dr within		
	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.12	56.6	78.3	88.2	1.19	51.9	77.5	86.82
V S d_{50}	1.16	49.1	71.9	82.5	1.21	49.6	64.3	79.8
D S d_{50}	1.21	53.23	74.5	82.9	1.24	55.8	75.2	82.2
D V d_{50}	1.39	33.1	53.6	67.68	1.49	38.0	55.8	65.12
D V S	1.51	37.6	57.4	71.8	1.71	34.1	59.7	67.4

The computation of the results for the model with four input variables ANFIS (D V S d_{50}) and without depth in the inputs ANFIS (V S d_{50}) are shown in Figure 6-4. The figure illustrates that the model estimates the observed gravel concentration accurately with little deviation from the line of perfect agreement.

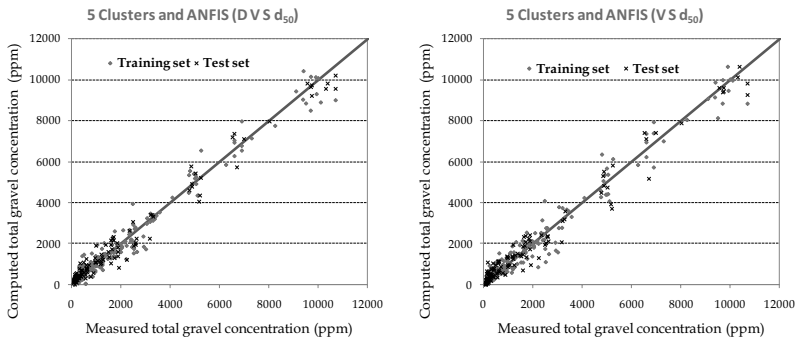


Figure 6-4: Scatter plots of measured versus computed gravel concentration for ANFIS models with five clusters.

The fuzzy clustering technique generates fuzzy rules which are capable of estimating the gravel transport rate. The drawback of this approach is that it is very difficult to interpret the final clusters and it tends to be more a black-box system. It is presented here as an alternative approach for a data-driven rule generation. Because the grid partitioning approach is better with respect to transparency and accuracy, it is selected as the better technique and the results from grid partitioning are presented in detail in the next sections.

The ANFIS model results for gravel dataset with four input variables and three fuzzy partitions for each input variable are indicated in Figure 6-5. From the scatter plots of observed and computed concentrations, it can be concluded that the ANFIS model performs quite well. The computed values are closer to the line of perfect agreement and show less scatter. The AARE values are 22.3 and 24.8% for the training and test datasets respectively, less than a 25% error for both datasets can be considered as very good. The correlation coefficient is 0.99 for the total dataset which indicates a very strong correlation between observed and computed values. The average discrepancy ratio is 1.08 for both the training and test datasets, and about 85% of the total dataset lie within a discrepancy ratio of 0.5 - 1.5. Only 4% of the training datasets and 2% of the test datasets show a computational error greater than 100%.

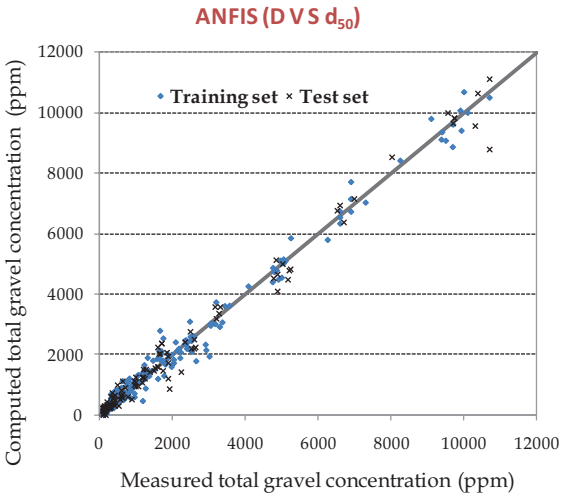


Figure 6-5: Scatter plots of measured versus computed total gravel concentration for ANFIS model with four input variables and 81 rules for the training and test datasets.

The results of the sensitivity analysis indicating the AARE and r values for the ANFIS models with five combinations of input variables and three membership functions for each variable are presented in Figure 6-6.

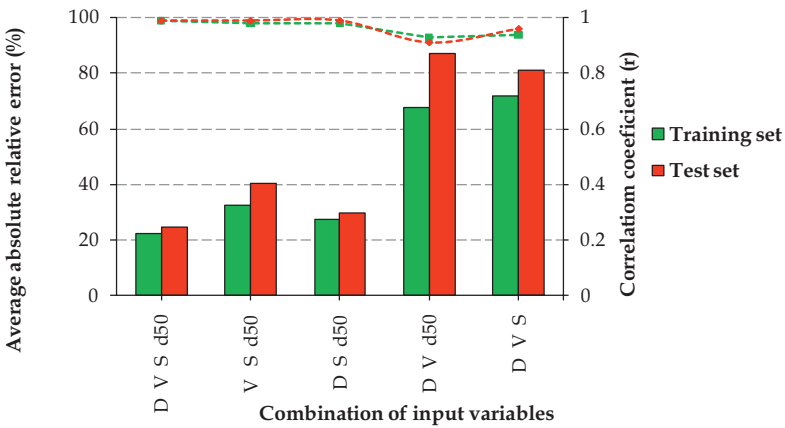


Figure 6-6: Performances of ANFIS models with various combinations of input variables and three generalized bell-shaped membership functions for each variable for gravel data.

From the results of the sensitivity analysis shown in Figure 6-6, slope and particle size are found to be the most dominant parameters. The ANFIS model derived by excluding particle size from the input parameters shows the least performance with AARE values of 67.5 % and 87% for the training and test datasets respectively. The detailed results for the model with five different combinations of input variables and three fuzzy sets (partitions) for each input variable can be found in Appendix E.

These results further prove the potential of the ANFIS model in successfully estimating measured gravel transport rates. Together with the results for laboratory data in the sand size range, this confirms that the data-driven fuzzy logic based modelling results are fairly accurate for both gravel and sand laboratory dataset. The laboratory flume sediment data typically have high measurement accuracy because it is easier to control the conditions. The data-driven fuzzy models results also reveal that with quite promising outputs.

The next sections of this chapter are devoted to testing whether the technique can also be applied to field datasets which have usually less measurement accuracy and high uncertainty.

6.3 ANFIS Model Results for Field Data from Different Authors

The first group of field dataset contains measured total bed-material concentration (ppm) and input variables for four sandy rivers (d_{50} < 2 mm) collected by different researchers. The accuracies of the ANFIS models developed by using seven combinations of the main input variables for field dataset obtained from different authors are presented in Tables 6.5 and 6.6 for the training and test datasets. The AARE, RMSE, r, percentage of data with ARE > 100%, average discrepancy ratio (Avg. Dr), and percent of data within discrepancy ratios between 0.75 - 1.25, 0.5 - 1.5, and 0.25 - 1.75 are summarized to compare the performance of the models on the training and test datasets.

Table 6.5: Performances of various ANFIS models with different combinations of input variables for field data from different authors.

Input variables	Training data				Test data			
	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%	AARE (%)	RMSE (ppm)	r	Percent of ARE>100%
D V S d_{50}	23.6	223.9	0.95	2.3	45.1	635.0	0.52	9.8
V S d_{50}	33.4	314.3	0.89	3.4	34.3	254.7	0.90	2.4
D S d_{50}	29.8	317.3	0.89	4.7	46.8	404.1	0.78	12.2
D V d_{50}	22.6	178.9	0.97	2.3	51.4	925.5	0.52	12.2
D V S	30.5	255.4	0.93	3.4	55.5	715.1	0.45	12.2
VS D d_{50}	29.7	315.3	0.89	1.1	27.2	261.9	0.89	2.4
VS d_{50}	28.4	317.6	0.89	2.3	32.9	307.8	0.85	7.3

Table 6.6: Discrepancy ratio analysis of different ANFIS models for field data from different authors.

Input variables	Training data				Test data			
		Percent of data Dr within				Percent of data Dr within		
	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg. Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.10	73.9	89.8	96.6	1.24	41.5	75.6	82.9
V S d_{50}	1.13	50.0	75.0	87.5	1.12	46.3	80.5	90.2
D S d_{50}	1.14	62.5	85.2	93.2	1.34	43.9	78.1	85.4
D V d_{50}	1.08	71.6	89.7	95.5	1.23	29.3	78.1	82.9
D V S	1.12	55.7	81.8	93.2	1.26	26.8	61.0	75.6
VS D d_{50}	1.14	55.7	77.3	93.2	1.13	68.3	90.2	90.2
VS d_{50}	1.12	56.8	83.0	96.6	1.19	58.5	82.9	87.8

The model with unit stream power, depth, and median particle size as input variables results in the best performance with AARE less than 30% both for the training and test datasets. Average discrepancy ratios of 1.14 and 1.13, and RMSE of 315.3 ppm and 216.9 ppm are obtained for the training and test datasets respectively. Table 6-6 shows that the percent of data within discrepancy range of 0.5 - 1.5 are 77.3% and 90.2% for the training and test datasets respectively. The ANFIS model with three input parameters (velocity, slope and d_{50}) shows good model accuracy with a correlation coefficient of 0.89 and AARE less than 35% for the total dataset, and average discrepancy ratios are 1.13 and 1.12 for the training and test datasets respectively. The correlation coefficient r = 0.89 for both datasets indicates a strong linear relationship between measured and computed sediment concentrations. The results also prove that having too many input variables does not necessarily result in a more accurate model, especially in cases where the number of available data is limited. Redundancy in input variables should be minimized as much as possible because during model optimization, redundant variables result in overtraining and complication. The ANFIS model with only VS and d_{50} as input parameters also shows an acceptable accuracy with only nine rules, the percentage of data with model computation error greater than 100% are 2.3% and 7.3 % for the training and test datasets respectively (Table 6-5). The analysis of the accuracy of the ANFIS models with different combinations of input variables demonstrates that, slope and particle size are the most important parameters whose exclusion or inclusion are significantly affecting the model performance. This is in agreement with many sediment transport equations based on stream power, unit stream power or shear stress approaches. Particle size determines the critical shear stress required for assessing incipient motion, and slope is a key parameter for computing stream power or unit stream power.

In combination with the results in sections 6.2, it can be concluded that data-driven fuzzy modelling can also play an important role in identifying the dominant parameters for developing sediment transport models with basic physics as their fundamental background.

6.4 Model Results for the Rhine River

Separate ANFIS models are developed for estimating total bed-material load and bed load transport rates for the Rhine, the datasets used for training and testing are summarized in chapter 4.

6.4.1 Total Bed-Material Load ANFIS Model Outputs for the Rhine

The accuracies of ANFIS models with different combinations of input variables using the total bed-material load data for the Rhine river are presented in Tables 6.7 and 6.8. Table 6.7 summarizes the AARE, RMSE, r and percentages of data with ARE greater than 100% for training and test datasets. In Table 6.6 the discrepancy ratio analysis of the different total bed-material load ANFIS models developed for the Rhine is presented.

Table 6.7: Performances of different total bed-material load ANFIS models for the Rhine river.

	Training data				Test data			
Input Variables	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%
D V S d_{50}	39.9	66.2	0.84	9.3	44.4	72.7	0.76	7.1
V S d_{50}	57.1	91.3	0.67	14.4	57.0	86.8	0.63	14.2
D S d_{50}	48.6	80.2	0.76	12.4	53.4	83.5	0.66	10.3
D V d_{50}	39.7	61.5	0.87	7.3	42.9	75.1	0.74	6.5
D V S	41.1	67.3	0.84	10.7	44.5	73.0	0.76	7.7

Table 6.8: Discrepancy ratio analysis of different total bed-material load ANFIS models for the Rhine river.

	Training data				Test data			
		Percent of data Dr within				Percent of data Dr within		
Inputs	Avg. Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg. Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.20	46.8	76.6	86.8	1.21	42.6	74.2	89.7
V S d_{50}	1.35	38.9	67.0	79.2	1.35	37.4	73.6	81.3
D S d_{50}	1.27	42.3	70.1	80.3	1.27	38.7	69.7	81.3
D V d_{50}	1.19	48.2	77.2	86.8	1.15	41.9	72.9	90.3
D V S	1.21	45.6	76.6	86.5	1.19	43.2	74.2	89.0

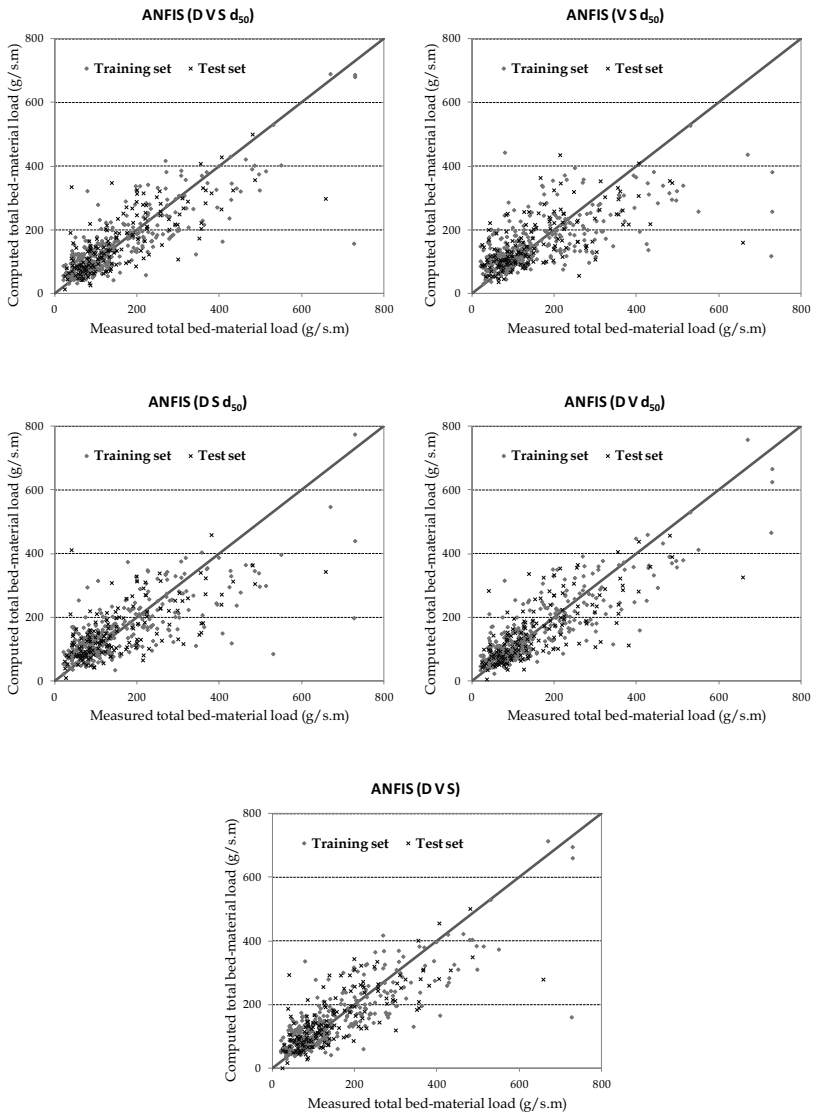


Figure 6-7: Scatter plots of measured versus computed total bed-material load transport rates per width for ANFIS models with different combinations of input variables for the Rhine river.

Figure 6-7 shows the plots of measured versus computed values of total bed-material load transport rates for the Rhine by using ANFIS models with five combinations of input

variables. From the results summarized in Tables 6.7 and 6.8, and the scatter plot in Figure 6-7 , the ANFIS model with four input parameters (D V S d_{50}) is found to be showing the best performance with AARE of 39.9% and 44.4%, average discrepancy ratio of 1.20 and 1.21 for the training and test datasets respectively. Better correlation coefficients of 0.84 and 0.76 are obtained for the training and test datasets respectively. Less than 10% of the total datasets have an absolute relative error (ARE) greater than 100%, and more than 74% of the total data lie within the discrepancy ratio range of 0.5 - 1.5, which is a good performance for estimating sediment transport in large rivers like the Rhine. From the results of the sensitivity analysis, depth and velocity are found to be the most important parameters influencing the model performance. From Figure 6-7 it can be seen that the computation results of ANFIS (V S d_{50}) and ANFIS (D S d_{50}) models show large deviations from the measured transport rates. This is because the variations in depth and velocity are significant for rivers like the Rhine. These variations affect many of the parameters influencing the sediment transport rate, like shear stress, stream power, or unit stream power. The ANFIS model with three input parameters D, V, and d_{50} shows comparable accuracy with that of four input parameters (D, V, S, d_{50}) with less than 7% of the total data having a model computational error >100%. The average discrepancy ratios are 1.19 and 1.15, and the AARE values are 39.7% and 42.9% for the training and test datasets respectively. The model without depth in the input variables, ANFIS (V S d_{50}), is the least accurate with an average discrepancy ratio of 1.35, $r < 0.7$ and AARE > 57% for both datasets. This proves that for large rivers like the Rhine, water depth should be included for a successful estimation of transport rates. The analysis of the model results show that large computational errors are associated with very low or very high sediment transport rates. This can also be clearly observed from Figure 6-7. The measurement accuracy for those events is usually relatively low which is also represented on the performance of the models. The distance between the first and last station in this analysis is more than 500 km, and the river has different morphological and hydraulic characteristics along the total section included in the analysis.

Based on the above analysis it can be concluded that the ANFIS model results are quite promising and can be used as a basis for estimating total bed-material sediment transport rate for the Rhine. As we are dealing with computing sediment transport with very high uncertainty, an average computation error of 40% is quite acceptable. If additional data is collected in the future, it can be included in the training datasets and the fuzzy model and the final rules (output equations) can be updated accordingly to improve model accuracy. In order to show a better performance of the fuzzy model, comparison of the model results with the results from other sediment transport equations is done and is presented in chapter 7.

6.4.2 Bed Load ANFIS Model Outputs for the Rhine

The results of the optimized bed load ANFIS models for the Rhine are given in Tables 6.9 and 6.10, and from the model performance summarized in the tables it can be concluded that

the ANFIS model estimates the observed bed load with reasonably acceptable accuracy. The ANFIS model with four input parameters is selected to be the best model with around 60% of the total data lying in the discrepancy range of 0.5 - 1.5. The AARE is 50.4% for the training datasets and 67% for the test sets, and the correlation coefficient is greater than 0.7 for both datasets. The RMSE values are 11.0 and 13.2 g/s.m for the model with four input parameters which are lower than those from other models. Slope is found to be having the least influence on the accuracy of the model results from the results of the sensitivity analysis. The slope difference from section to section is not very significant. The slope is also considered in the flow velocity if we consider the Manning's or Chezy's equation where V is computed as a function of slope. The percentage of data with absolute relative error greater than 100% is 13.5% for the training and 22.4% for the test dataset, which are less than the values obtained for the other models. The discrepancy ratio analysis indicates 80.3% of the training data and 71.3% of the test data lie within discrepancy ratios 0.25 - 1.75. The fuzzy model with three input parameters (without slope) results in acceptable accuracy with AARE values of 51.5% and 70.1% for the training and test datasets respectively. As stated in the previous section, these fuzzy models for estimating bed load transport in the Rhine can be updated in the future based on the availability of additional data.

Table 6.9: Performances of different bed load ANFIS models for the Rhine river.

	Training data				Test data			
Inputs	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%
D V S d_{50}	50.4	11.0	0.75	13.5	67.0	13.2	0.70	22.4
V S d_{50}	61.2	13.3	0.60	19.2	73.5	14.8	0.57	22.4
D S d_{50}	67.2	13.9	0.55	19.7	70.9	13.7	0.65	24.1
D V d_{50}	51.5	11.4	0.73	13.0	70.1	20.0	0.45	20.7
D V S	56.6	12.3	0.67	14.5	72.7	21.8	0.61	21.8

Table 6.10: Discrepancy ratio analysis of different bed load ANFIS models for the Rhine river.

	Training data				Test data			
		Percent of data Dr within				Percent of data Dr within		
Inputs	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.27	40.16	65.3	80.3	1.45	30.5	58.6	71.3
V S d_{50}	1.37	35.2	58.0	72.8	1.50	30.5	56.3	67.8
D S d_{50}	1.41	26.9	54.7	73.1	1.46	30.5	53.5	69.5
D V d_{50}	1.29	36.3	65.5	80.6	1.48	32.2	54.0	67.8
D V S	1.34	35.2	61.4	75.1	1.51	34.5	54.6	70.1

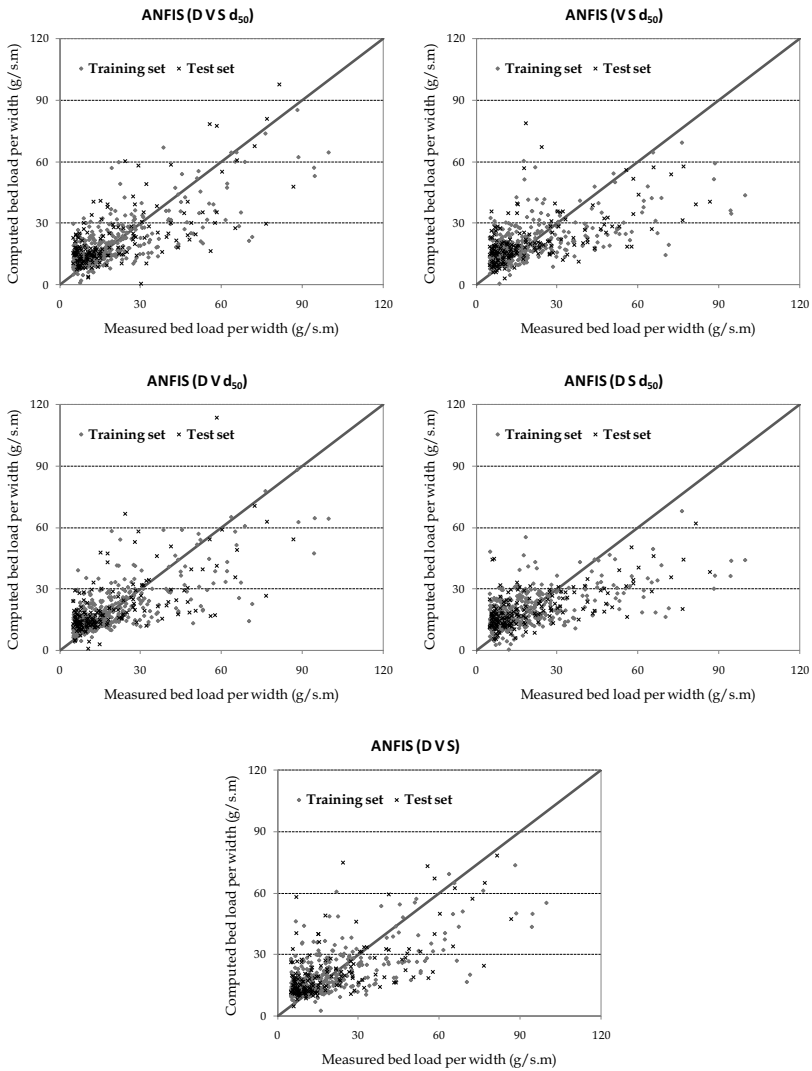


Figure 6-8: Scatter plots of measured versus computed bed load transport rates per width for ANFIS models with different combinations of input variables for the Rhine river.

The computation results plotted in Figure 6-8 for the ANFIS models with five combinations of input variables further strengthen the above conclusions. Here again, large computational errors with significant deviations are associated with high transport rates.

Compared to the results of total bed-material load ANFIS models (section 6.4.1), the bed load ANFIS model results are less accurate. This can be attributed to the complexity of bed load transport, and is usually the case for other sediment transport equations as well. Measuring the amount of bed load is difficult and varies from section to section in a cross section. The measurement accuracy is low most of the time and it is difficult to separate bed load and suspended load transport layers accurately. Considering the uncertainty and fuzziness associated with the process of bed load transport, and its measurement accuracy, the data-driven ANFIS model results are considered to be quite good. Usually prediction error of most of the existing bed load transport equations is greater than 100%, and thus the fuzzy model is performing relatively very well.

6.5 Model Results for the Elbe River

Using the dataset obtained from the BfG for the section of the Elbe river from Dresden to Neu Darchau, data-driven ANFIS models are developed for computing the amount of bed load and total bed-material load transport rates. The total section of the river is used in the analysis to build a data-driven fuzzy model which can give a fairly accurate estimation for the total section of the Elbe where measured sediment transport values are available.

6.5.1 Total Bed-Material Load ANFIS Model Outputs for the Elbe

Tables 6.11 and 6.12 show the accuracy of different total bed-material load ANFIS models developed for the Elbe river. The ANFIS model with four input parameters (D, V, S, d_{50}) is found to be having the best performance with AARE values of 52.4% and 50.4%, average discrepancy of 1.29 and 1.22 for the training and test datasets respectively. Around 70% of the total data lie in the discrepancy ratio range of 0.5 - 1.5 and less than 15% of the total data have a computational error greater than 100%. This is a good performance for estimating sediment transport in large rivers like the Elbe. The scatter plot of the results from the ANFIS models with different combinations of input variables is given in Figure 6-9. The figure also proves that the model results are sensitive to water depth unlike in the case of laboratory data and it is important to include depth in the input parameters. As most of the sediment load in the Elbe river is transported as suspended load (the Elbe is a sandy bed river), the effect of excluding the sediment particle size from the input parameters is not quite significant. The fuzzy model with three input parameters D, V, and S shows a comparable accuracy as that with four input parameters, with AARE of 54.1% for the training datasets and 45.5% for the test datasets. Additionally, average discrepancy ratios of 1.30 and 1.20, correlation coefficient of 0.8 and 0.6 are obtained for the training and test datasets respectively. Around 80% of the total data lie within discrepancy ratio of 0.25 - 1.75. The RMSE are 57.4 and 82.9 g/s.m, and the percent of data with ARE > 100% are 13.6% and 15.2% for the training and test datasets respectively.

Table 6.11: Performances of different total bed-material load ANFIS models for the Elbe river.

	Training data				Test data			
Inputs	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%	AARE (%)	RMSE (g/s.m)	r	Percent of ARE>100%
D V S d_{50}	52.3	56.7	0.80	14.1	50.4	80.7	0.56	13.9
V S d_{50}	62.5	67.8	0.69	18.6	53.2	76.2	0.53	13.9
D S d_{50}	61.2	67.5	0.69	19.5	48.6	82.5	0.62	10.9
D V d_{50}	55.8	60.0	0.78	17.7	54.8	84.3	0.54	13.9
D V S	54.1	57.4	0.79	13.6	45.5	82.9	0.59	15.2

Table 6.12: Discrepancy ratio analysis of different total bed-material load ANFIS models for the Elbe river.

	Training data				Test data			
		Percent of data Dr within				Percent of data Dr within		
Inputs	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.29	46.4	70.9	80.5	1.22	37.6	67.3	78.2
V S d_{50}	1.37	35.5	63.6	75.9	1.19	38.6	60.4	77.2
D S d_{50}	1.38	42.7	66.4	77.7	1.22	34.7	68.3	85.2
D V d_{50}	1.33	41.8	70.9	79.1	1.29	40.6	69.3	82.2
D V S	1.30	43.6	69.1	79.6	1.20	39.6	68.3	79.2

Graphical illustration of the different total bed-material load ANFIS model outputs for the Elbe is given in Figure 6-9. This figure shows that computed values by using the model with four input variables ANFIS (D V S d_{50}) and three input variables without particle size ANFIS (D V S) are closer to the line of perfect agreement and result in less scatter. It can be concluded that with three input variables and three membership functions for each of the variables, which results in twenty seven fuzzy rules, acceptable model accuracy is obtained. Hence, the model ANFIS (D V S) can be used as alternative fuzzy model for estimating total bed-material load transport in the Elbe for cases where particle size is not available.

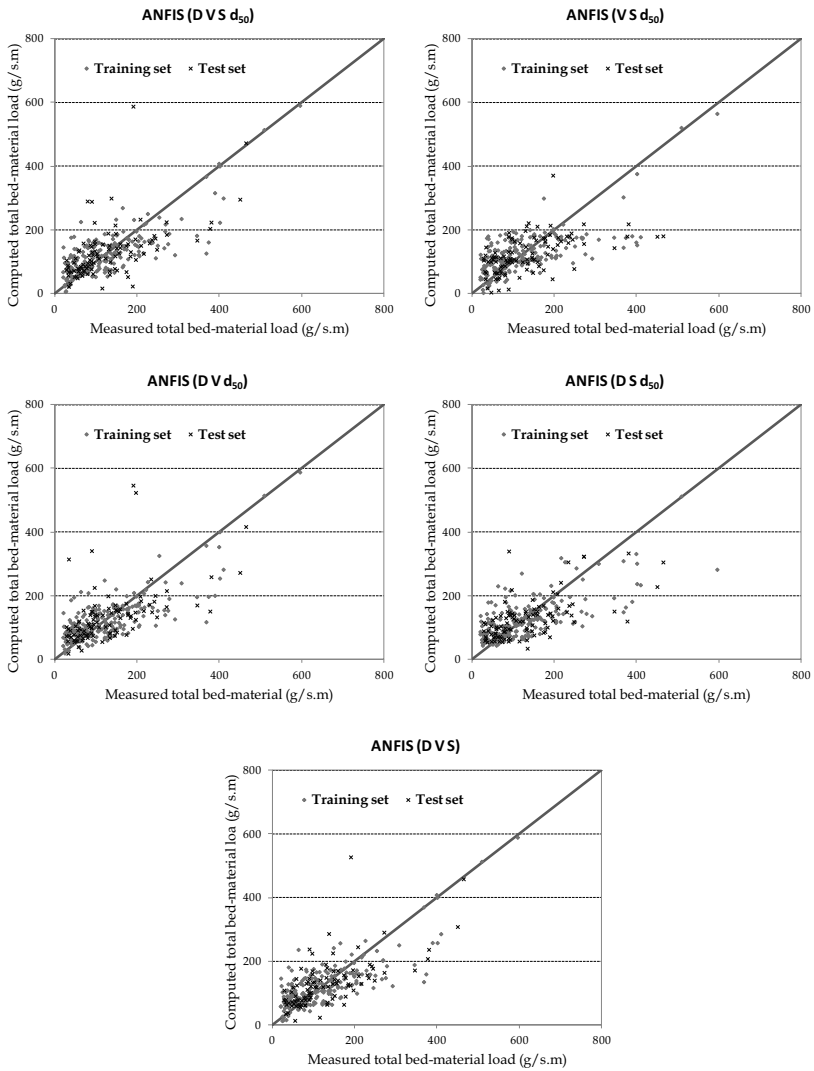


Figure 6-9: Scatter plots of measured versus computed total bed-material load transport rates per width for ANFIS models with different combinations of input variables for the Elbe river.

Those who are not familiar with the accuracies of the different sediment transport equations might be surprised that a 50% computational error using an uncalibrated equation is acceptable, but most sediment transport equations are usually less accurate when applied to

natural rivers like the Elbe. Until a breakthrough is achieved in the field of sediment transport calculation, engineers have to work with this rough estimate.

6.5.2 Bed Load ANFIS Model Outputs for the Elbe

Similar steps are followed and ANFIS models are developed for estimating the bed load transport rate in the Elbe river. The results are not as good as those for the Rhine river. Further data analysis and filtering is performed aiming at achieving better accuracy; more and more extreme values are removed to exclude high and low transport rates. In Figure 6-10, the average absolute relative error values and the correlation coefficients are plotted for five combinations of the primary input variables. Table 6.13 summarizes the discrepancy ratio analysis for the training and test datasets of the final of bed load ANFIS models for the Elbe.

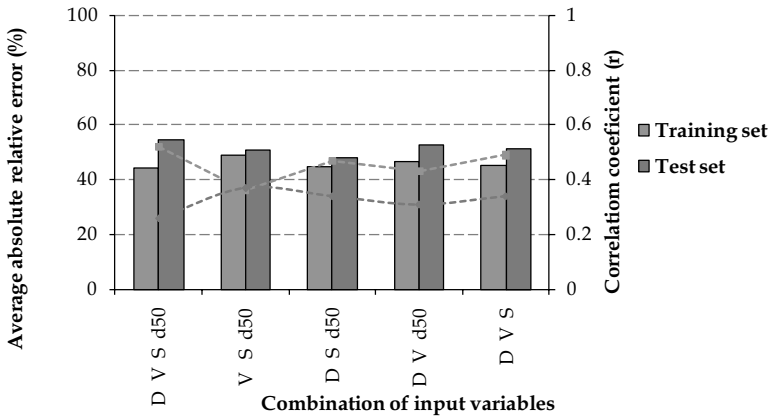


Figure 6-10: Performance of different bed load ANFIS models for the Elbe river.

Table 6.13 Discrepancy ratio analysis of different bed load ANFIS models for the Elbe river.

	Training data				Test data			
		Percent of data Dr within				Percent of data Dr within		
Inputs	Avg. Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg. Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.21	36.2	70.8	82.7	1.29	34.5	62.0	77.9
V S d_{50}	1.25	31.7	64.6	77.4	1.29	32.7	67.3	77.0
D S d_{50}	1.23	37.9	67.1	82.3	1.25	35.4	67.3	80.5
D V d_{50}	1.24	33.3	69.1	79.4	1.31	36.3	62.8	77.0
D V S	1.22	35.0	72.0	82.7	1.26	34.5	61.9	77.9

From Figure 6-10 it can be clearly observed that although the AAREs of all of the various models are less than 50%, all of them have poor correlation coefficients (r < 0.5). Irrespective of the combination of input variables, model performance is not varying. The AARE, average Dr, and r are almost the same for various ANFIS architectures. The summary of discrepancy ratio analysis provided in Table 6.13 shows that the average discrepancy ratio and the percentage of data with discrepancy ratios between 0.75 - 1.25, 0.5 - 1.5, 0.75 - 1.25 are not changing with the different combinations of input variables. The model with three input parameters D, S, d_{50} shows a slightly better accuracy with an average discrepancy ratio of 1.23 for the training and 1.25 for the test datasets. The Elbe is a sandy river where most of the sediments are transported as suspended load; the quality of the bed load measurement can be the main reason for less accuracy of the results. It is very difficult to separate the bed load transport layer from the suspended load transport layer for sandy rivers with fine bed materials. In Figure 6-11, the discharge versus the measured bed load transport rate (g/s.m) is plotted to demonstrate that there is not any clear trend in the measured bed load transport dataset for the Elbe. This is also observed in the computed transport rates using the ANFIS models. This figure clearly shows that discharge should not be used to estimate sediment load for a transport formula or rating curve.

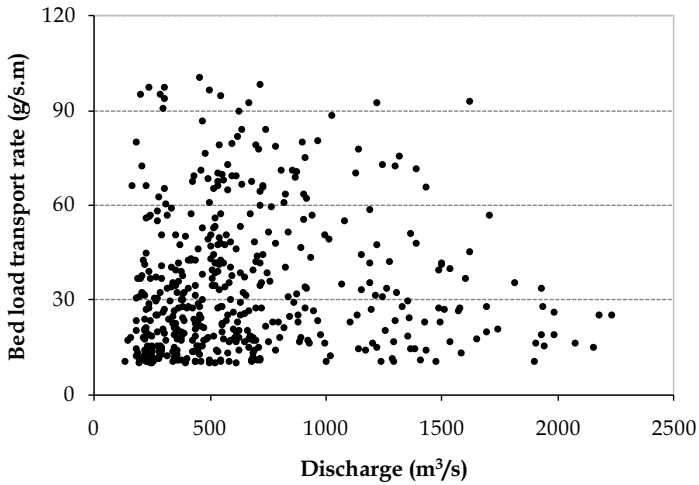


Figure 6-11: Discharge versus bed load per unit width and time for the Elbe river.

The performance of data-driven models depends primarily on the quality of the input data utilized to develop the models. The results of the sensitivity analysis show that the data utilized for training and testing are not representative for the bed load transport behaviour in the river. The point of measurement in a cross-section is usually a factor in the

measurement error because the bed load transport is not uniform throughout the cross-section. Computational error is very high for bed load, which is in agreement with the complexity of measuring it, and separating the bed load from the suspended load. In order to improve the performance of the model, collection of additional data is necessary. The results in this section prove the importance of data quality for successful development of models based on data-driven approaches. The final model performance depends on the quality of the measured data and it is very necessary to verify the data before using it for the modelling. The ANFIS model detects this and is a further indicator for the validity of the modelling approach.

6.6 Further Refining of the Developed ANFIS Models

The ANFIS models developed in the previous sections for the Rhine and Elbe rivers are considering the total sections of the rivers as a whole. The rivers are not divided into sections with similar morphological characteristics although there is a significant difference between the various sections of the rivers. The measured transport rates and the input parameters are put together in the datasets. Assuming that model accuracy might increase if the rivers are categorized into different reaches, the Rhine and Elbe rivers are divided into sections based on the input parameters slope, width, and bed-material size, and separate models are developed for parts of the river reaches which have comparable morphologic characteristics.

6.6.1 The Rhine River

The Rhine river is divided into three sections: Upper, Middle and Lower Rhine based on the slope and river morphology and separate models are developed for each section. The Upper Rhine is from Iffezheim to Bingen, Middle Rhine between stations Bingen and Porz, and the Lower Rhine is downstream of Porz till the German Dutch border (station Griethausen). Additional classification of the Rhine river is done based on the input parameters slope, median particle size, and width. The Rhine has been categorized into three classes for each of these parameters. For example on the basis of the slope, the river is divided into three groups. Category one is $0.10 < S < 0.13$, category two is selected to be $0.13 < S < 0.2$, and category three is slope between 0.20 and 0.48 mm/m.

Table 6.14 shows the ranges of the input parameters used for the categorization of the Rhine and Elbe rivers into sections based on various hydraulic variables. During the classification of the rivers it has been taken into consideration to get approximately equal amount of datasets for each section. It should be mentioned that the ranges of the input parameters and the descriptions provided are just for classification in this study and are not necessarily in agreement with other official classifications and nomenclatures.

6.6.2 The Elbe River

The Elbe river is divided into sections upstream and downstream of the station Magdeburg because the characteristics of the river are different for these two reaches. Additional classifications of the Elbe based on slope, particle size, and width are provided in Table 6.14.

Table 6.14: Classification of the Rhine and Elbe rivers based on hydraulic input parameters and morphology.

Parameter	Rhine		Elbe	
	Range	Description	Range	Description
Slope (mm/m)	0.10-0.12 0.13-0.19 0.20-0.48	mild medium steep	0.12-0.19 0.2-0.21 0.22 -0.35	low medium high
Particle size (mm)	0.5-2 2.1-10 10.1-40	sand gravel coarse gravel	0.4-1 1.1-2 2.1-10	fine sand coarse sand gravel
Width (m)	150-300 301-350 351-800	narrow medium wide	70-150 151-200 201-1000	narrow medium wide
Morphology	Iffezheim-Bingen Bingen-Porz Porz-Griethausen	upper middle lower	Dresden-Barby Magd-Neu Darchau	Upstream Magd. Downstream Magd.

The results obtained after dividing the river reaches into sections indicate that model accuracy is not improved compared to the ANFIS models developed for the whole sections for both rivers. Even less accurate model results are obtained, especially for the test datasets. The number of available data is one of the main reasons for this because there is not sufficient data describing the sections. The available number of datasets is limited and gets even smaller when the rivers are divided into sections. The variation in the distribution of input and output parameters in the datasets of the sections grouped is still high. When a river reach is divided into three sections, the total number of datasets becomes one third of the datasets for the whole river reach. This results in less datasets to represent the non-linear behaviour of sediment transport. The performance over the test datasets proves that the quality and quantity of training datasets is not good enough. Additional data collection is required in order to develop separate models for different sections.

6.7 Discussion

This chapter summarizes the results of the optimized data-driven ANFIS models for the different groups of data available in this research. The ANFIS models developed estimate the observed transport rates with acceptable accuracy for both the laboratory and field datasets. The model results for laboratory data are better than those for field datasets where the measurement accuracy is less and the uncertainty is high. The ANFIS model results of the sensitivity analysis indicate that different parameters play dominant roles in governing the process of sediment transport. For example, the depth is a more important input variable for rivers such as the Rhine and Elbe, but not that important for laboratory flumes and small rivers. This is consistent with basic knowledge in hydraulics. Additionally, total bed-material load is estimated with better accuracy than bed load for both the Rhine and Elbe rivers. The measuring accuracy of the bed load data is usually less, and it is usually difficult to separate the bed load from the suspended load transport layer. The bed transport is a more complicated process and the accuracy of most other existing sediment transport equations is not good either.

The model performance is data dependent and the collection of sufficient amount of data is necessary and the quality of data should be carefully analyzed before application. In general, the number of training datasets should be as much as the number of adjustable parameters in the fuzzy models. Data-driven models are generally descriptive not predictive. Therefore, if the river characteristic is artificially changed by construction of hydraulic structures, this alters the hydraulics and sediment transport behaviour in the rivers and the ANFIS models should be updated with new data. Like most regression models, data-driven models are specific to the datasets used during model optimization. The generalization capacity of the models and transferability to other rivers is an issue that has to be addressed. If additional data is obtained for other river reaches and implemented in the model development, it might be possible to develop a fuzzy model that is fairly accurate for many rivers. Introducing dimensionless parameters is an alternative that should be explored in order to remove the problems associated with having significantly different ranges of variables because of different morphological characteristics of the various rivers.

From the model results for the Rhine and Elbe rivers, it can be concluded that the ANFIS technique can be a very powerful alternative modelling approach for river reaches like the Rhine and Elbe where the BfG currently invests a significant effort in measuring and monitoring sediment transport. The models developed here can be implemented for future estimation of transport rates in the rivers.

7 Comparison of Sediment Transport Equations and ANFIS Models

Comparison of the accuracy of the developed data-driven ANFIS models and the results from commonly utilized sediment transport equations is an indicator of the validity of the approach and its performance. Many authors have analysed and compared the accuracy of available sediment transport equations. These comparisons can be done directly or indirectly. Direct comparison is possible using measured and computed transport rates and indirect comparison is usually based on analysis of results from computer model simulations (Yang, 1996). Some of the results of the investigations regarding the accuracy and ranking of sediment transport formulas include (White et al., 1975; Yang, 1979, 1984, 1996; Alonso et al., 1980; Brownlie, 1981; Yang and Molinas, 1982; Vetter, 1989; Yang and Wan, 1991; van Rijn, 1993; Yang and Huang, 2001). Most of the comparisons are performed using laboratory data and the ranking of the different equations is not consistent.

Yang and Huang (2001) published a comprehensive comparison of 13 sediment transport formulas to determine their limits of application. Wu et al. (2008) compare the accuracy of eight different equations in predicting sediment transport for the Yellow river in China which has a high concentration of fine sediments. Some of the commonly utilized bed-load equations had been evaluated by different researchers (White et al., 1975; van den Berg, 1987) and the predictive accuracy of most of the equations is found to be poor. Barry et al. (2008) analyze the performance of five bed load transport equations with regard to different geomorphology. The ranking of the accuracy of formulas in the comparisons is not consistent, mainly because they are based on different sets of data.

This chapter focuses on comparing the performances of the ANFIS models developed for each group of data with that of the computation results obtained by applying selected sediment transport equations. The guidelines summarized in section 2.6 are used for selecting appropriate transport equations based on the characteristics of input parameters in the datasets.

7.1 Laboratory Data Model Comparison

7.1.1 Sand

The performance of the ANFIS models developed for computing total sand transport in laboratory flumes is compared with the sediment transport equations of Yang (1973), Ackers and White (1973), Bagnold (1966), Engelund and Hansen (1972), and Laursen (1958) on the basis of the calculated RMSE, AARE, Dr, and r values. Additionally, Colby's (1964) and Toffaleti's (1969) equations which are derived based on the probabilistic approach of Einstein (1950) are implemented. These total bed-material load transport functions are selected for comparison because they are commonly applied. The median particle diameter (d_{50}) is used for all sediment transport formulas during the computation. Stevens and Yang (1989) published FORTRAN and BASIC computer programs for thirteen commonly used sediment transport formulas in river engineering and the source code can be found in Yang (1996). The FORTRAN program is here used for computing transport rates using the selected total bed-material transport equations.

Table 7.1 summarizes the comparison of the statistical model performance criteria AARE, RMSE, r, average discrepancy ratio, and the percentage of the computed values within discrepancy ratios of 0.75 - 1.25, 0.5 - 1.5, and 0.25 - 1.75 for the different models.

The scatter plots of computed versus measured total bed-material sediment concentration (ppm) for the ANFIS models with three input variables (V S d_{50}) and four input variables (V S D d_{50}), and three generalized bell-shaped membership functions for each input variable are shown in Figure 7-1 (section 6.2.1).The computation results from six selected sediment transport equations of Yang (1973), Ackers and White (1973), Bagnold (1966), Engelund and Hansen (1972), Laursen (1958), and Colby (1964) are presented in Figure 7-2. The results from the Toffaleti (1964) equation are not included because they are worse than the other equations. As it can be seen from Table 7.1, the discrepancy ratio of the different models does not vary significantly. It is desirable to have a graphical representation and analysis of the results. Using only the statistical model performance parameters might result in misleading conclusions. This is why in the following sections, the graphs of measured and computed transport rates are presented together with the tabular summary of the statistical parameters.

Table 7.1: Comparison of performances of different sediment transport models for laboratory data in the sand size range.

Model	AARE (ppm)	RMSE (ppm)	r	Avg.Dr	Percent of data Dr within 0.75-1.25	0.5-1.5	0.25-1.75	ARE >100%
ANFIS (V S d_{50})	23.4	1825.8	0.97	1.07	70.0	90.4	95.1	1.8
ANFIS (D V S d_{50})	24.6	1751.8	0.97	1.07	67.5	88.2	94.1	3.0
Yang	27.2	2953.4	0.91	1.03	55.8	88.0	96.1	2.0
Engelund and Hansen	35.6	3948.5	0.84	0.97	42.3	79.2	93.4	4.3
Ackers and White	43.8	6938.8	0.81	1.25	39.2	69.9	86.4	6.9
Laursen	70.4	10533.2	0.75	1.44	20.8	45.5	69.0	19.4
Bagnold	48.9	5019.6	0.92	1.03	29.8	72.6	88.9	7.9
Colby	45.2	6223.95	0.72	0.63	21.1	50.5	77.5	11.5

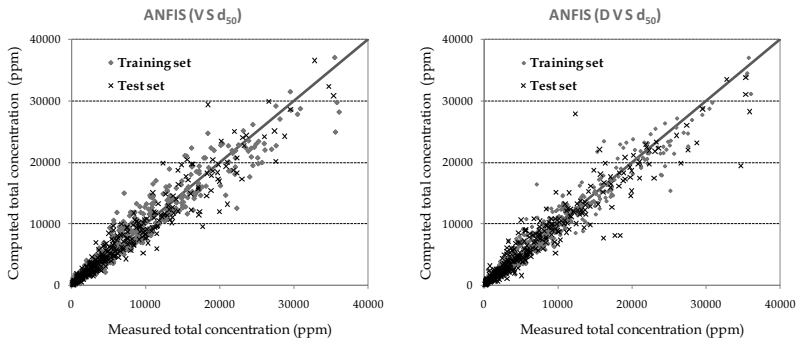


Figure 7-1: Scatter plots of measured versus computed total bed-material sediment concentration by using ANFIS models for laboratory data in the sand size range.

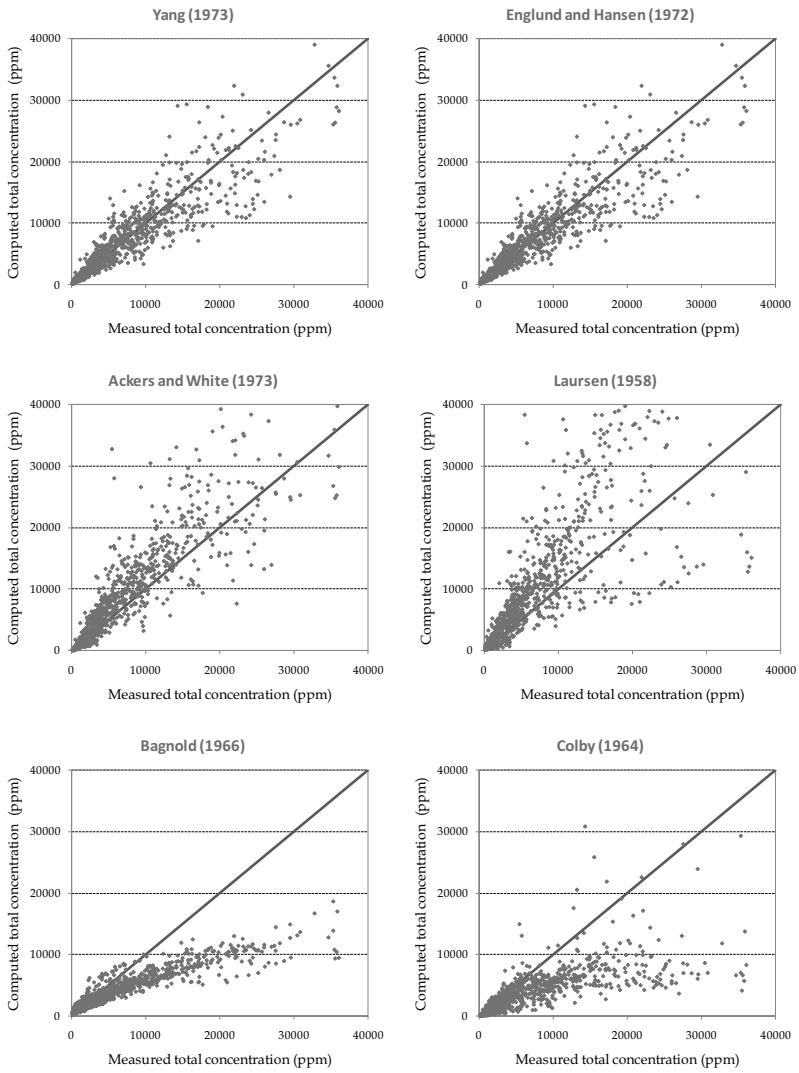


Figure 7-2: Scatter plots of measured versus computed total bed-material concentration by using selected sediment transport equations for laboratory data in the sand size range.

From Figures 7-1 and 7-2, it can be observed that the computation results of the ANFIS models show less scatter and are closer to the line of perfect agreement than the classical physical based equations. Together with the model performance evaluation results indicated in Table 7.1, the ANFIS model with three input parameters without depth (V S d_{50}) is found

to be the best with AARE of 23.4% and a very high correlation coefficient of 0.97. Less than 2% of the computed values have ARE greater than 100% and 90.4% lie within a discrepancy ratio of 0.5 - 1.5. Additionally, the ANFIS model with four input parameters also estimates the observed sediment transport concentrations successfully with AARE of 24.6% and 95.1% of the computed values lie within the discrepancy range of 0.5 - 1.5. The RMSE values provided in Table 7.1 indicate the two ANFIS models have the least values showing better estimation accuracy.

The values computed by using transport equations show a considerable deviation from the line of perfect agreement. The accuracies of formulas in descending order are those of Yang (1973), Engelund and Hansen (1972), Ackers and White (1973), Bagnold (1966), Colby (1964) and Laursen (1958). The AARE of Yang's equation is 27.2% with 88% of the computed values lying within discrepancy ratio range 0.5 - 1.5. The Laursen equation is found to be the least accurate with low correlation coefficient and very high AARE of 70.4% and RMSE of 10533.2 ppm, and 19.4%of the computed values show absolute relative computation error greater than 100%. The Bagnold (1966) and Colby (1964) equations consistently under predicted especially high sediment concentrations as can be observed from Figure 7-2. The average discrepancy ratios are 0.63 and 1.44 for Colby's and Laursen's equations respectively, indicating a very poor prediction quality. The above ranking is only based on the datasets used in this analysis. The data-driven fuzzy model significantly out-performed all the equations and successfully estimates total sand concentration. However, it should be noted that data-driven models could give accurate results only within the range of datasets used in the analysis. The ANFIS model optimized for laboratory datasets cannot be used for estimating transport rates for field conditions in natural rivers with different ranges of input and output parameters.

7.1.2 Gravel

There are some equations which are formulated for prediction of the transport of coarse material or gravel. Gravel is mainly transported as bed load. The equations selected for comparison are Yang (1984), Meyer-Peter and Müller (1948), Engelund and Hansen (1972) Ackers and White (1973), and Parker (1990). The accuracy of ANFIS model is compared with these equations in successfully estimating measured gravel concentrations. The gravel laboratory datasets collected from different sources and summarized in chapter 4 is used to evaluate the performance of the models.

The statistical model performance evaluation criteria for the equations of Yang, Engelund and Hansen, Ackers and White, and Meyer-Peter and Müller are summarized in Table 7.2. Figure 7-3 illustrates the comparison between measured and computed gravel concentration computed by four different equations and ANFIS models developed in section 6.2.2. As can

be seen from Figure 7-3, the computation results of the ANFIS models are closer to the measured gravel concentration and show less scatter than the selected transport equations.

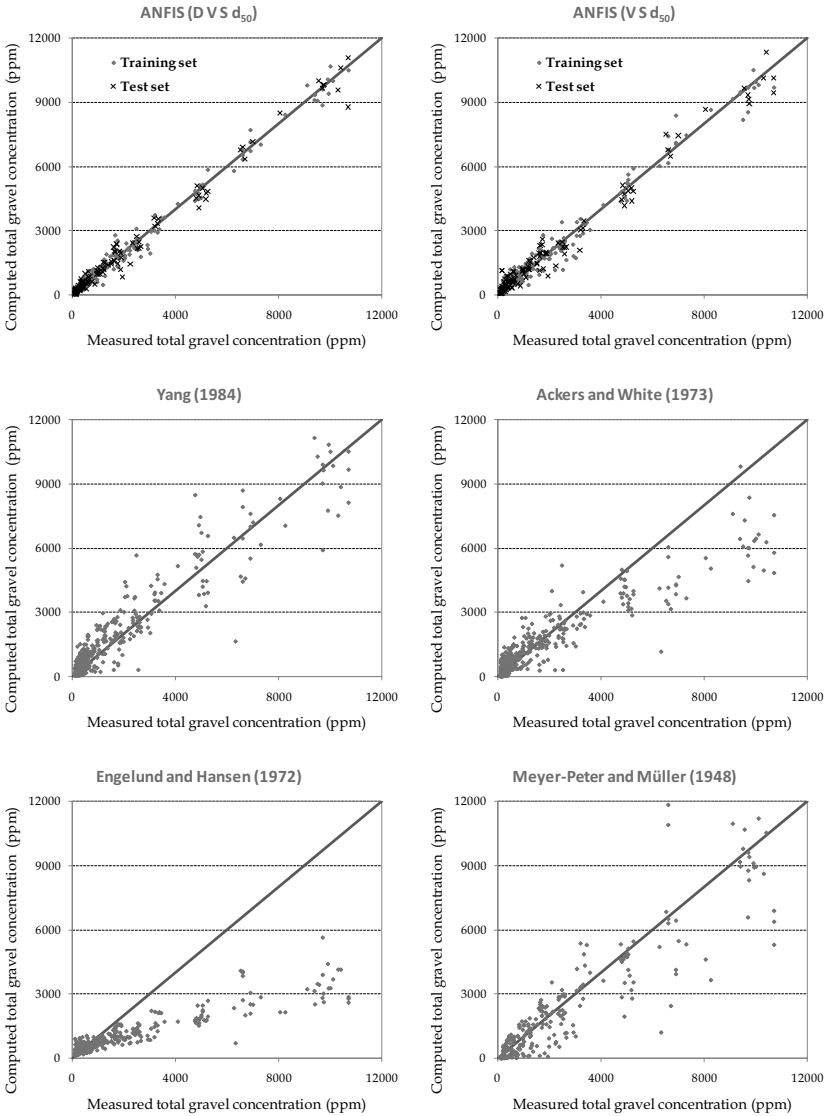


Figure 7-3: Scatter plots of measured versus computed total gravel concentration by using ANFIS models and selected equations.

Table 7.2: Comparison of performances of selected sediment transport equations and ANFIS models for laboratory data in the gravel size range.

Model	AARE (ppm)	RMSE (ppm)	r	Avg.Dr	Percent of data Dr within 0.75-1.25	0.5-1.5	0.25-1.75	ARE >100%
ANFIS (V S D d_{50})	22.8	371.6	0.99	1.08	68.4	86.2	92.4	2.3
ANFIS (V S d_{50})	35.1	440.3	0.98	1.18	57.6	78.6	89.3	6.4
Yang	64.4	953.7	0.93	1.25	29.6	47.7	67.1	20.9
Ackers and White	68.3	1169.7	0.92	1.25	31.1	58.2	72.5	17.4
Engelund and Hansen	56.4	1911.1	0.92	1.05	25.0	55.1	84.9	10.0
Meyer-Peter and Müller	50.9	1128.8	0.92	0.84	23.5	39.03	52.3	32.1

From the summary of the model evaluation criteria in Table 7.2, the ANFIS model with four input variables is selected to be the best model with AARE less than 25% and 86.2% of the computed values lie within a discrepancy ratio range 0.5 - 1.5. The model with three input variables also shows an acceptable accuracy with AARE of 35.1% and r = 0.98. All of the equations have AARE greater than 50%. The Engelund and Hansen equation consistently under predicts the observed gravel concentration. It is usually applied for computing sand transport. If only the correlation coefficient is used for model comparison, it leads to false conclusions as almost all the models show a high correlation coefficient. The Meyer-Peter and Müller equation which is formulated for computing gravel transport also underperforms with 32.1% of the computed values giving ARE > 100%. The mean discrepancy ratio for the equations of Yang (1984), Ackers and White (1973), Engelund and Hansen (1972), and Meyer-Peter and Müller (1948) are 1.25, 1.25, 1.05, and 0.84 respectively. All these values are higher than the values computed for the two ANFIS models. The computation results obtained by using Parker's (1990) equation show even larger deviations from observed transport rates and therefore are not included here.

7.2 Field Data Model Comparison

Similar to the analysis done for laboratory data, a detailed comparison of the results from the data-driven ANFIS models and the computation results from selected transport equations is performed for the different categories of river datasets.

7.2.1 Field Data from Different Authors

Table 7.3 presents the results of two ANFIS models (section 6.3) and the results from the sediment transport equations of Yang (1973), Engelund and Hansen (1972), Ackers and White (1973), Laursen (1958), and Bagnold (1966) for the field datasets in the sand size range collected by different authors. The plots of computed versus observed sediment

concentration are shown in Figures 7-4 and 7-5. The computation results from the ANFIS models with three input parameters (V S d_{50}), and the model with unit stream power, depth and median particle size (VS D d_{50}) are included for comparison.

Table 7.3: Comparison of performances of selected sediment transport equations and ANFIS models for the field data from different authors.

Model	AARE (ppm)	RMSE (ppm)	r	Avg.Dr	Percent of data Dr within 0.75-1.25	0.5-1.5	0.25-1.75	ARE >100%
ANFIS (VS D d_{50})	28.9	299.4	0.89	1.14	59.7	92.3	94.6	1.6
ANFIS (V S d_{50})	33.7	296.7	0.89	1.12	44.8	76.7	88.4	3.1
Yang	43.5	540.9	0.80	1.34	44.2	69.0	85.3	7.0
Engelund and Hansen	59.1	635.3	0.85	1.51	32.6	58.1	75.2	17.8
Ackers and White	66.0	1007.9	0.77	1.55	33.3	63.6	72.9	20.2
Laursen	136.2	672.1	0.32	1.72	11.6	29.5	58.9	31.8
Bagnold	118.1	476.4	0.71	1.96	25.6	49.6	56.6	38.0

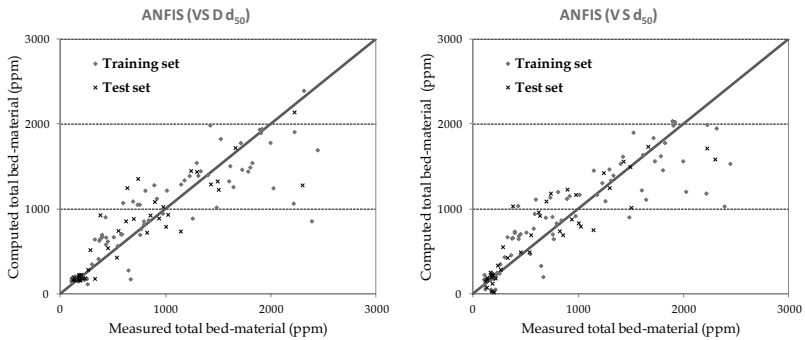


Figure 7-4: Measured versus computed total bed-material concentration by using ANFIS models for field data from different authors.

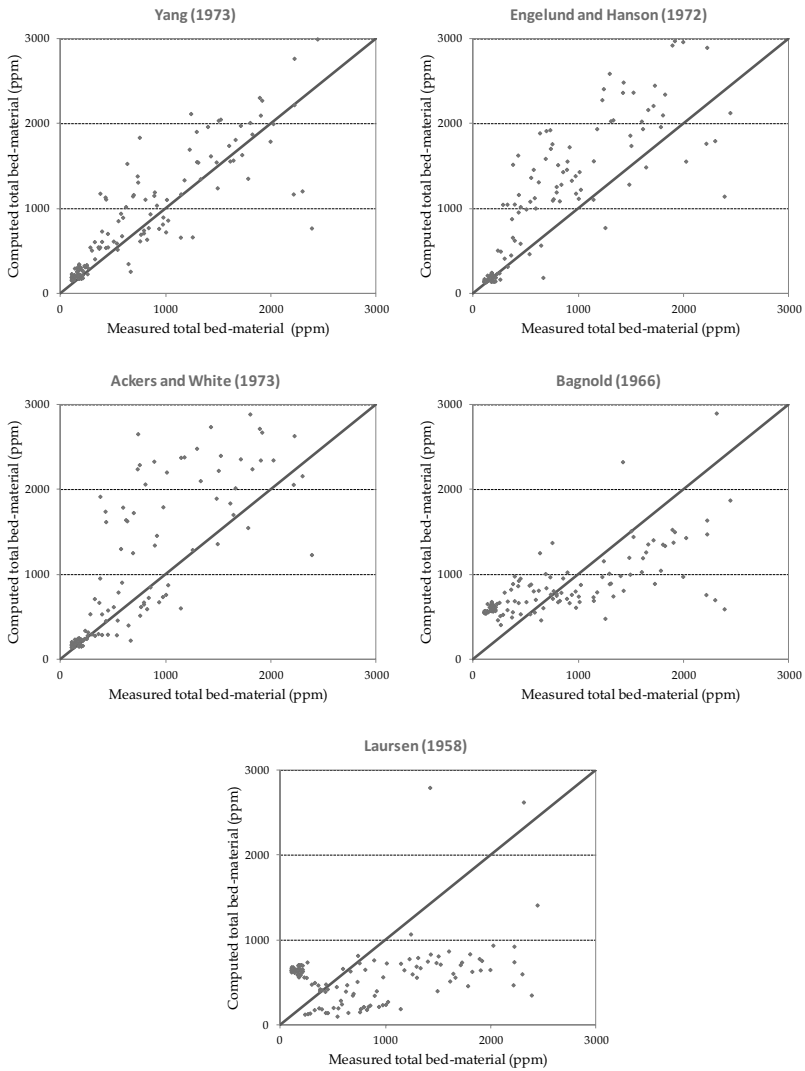


Figure 7-5: Measured versus computed total bed-material concentration by using selected sediment transport equations for field data from different authors.

From the results in Table 7.3, and Figures 7-4 and 7-5, it can be clearly observed that the ANFIS models perform better than the selected transport equations with lower AARE and RMSE, and higher r values. The equations show significant variations in model performance evaluation criteria given in Table 7.3. The ANFIS model with VS, D and d_{50} as input variables

has AARE 28.9% and 92.3% of the total data lie within the discrepancy ratio range of 0.5 - 1.5, and less than 2% of the computed values show an estimation error greater than 100%. This is a very good result for estimating sediment transport rates. The ANFIS model with V, S, and d_{50} as input variables also performs better than the selected equations as can be seen from Figure 7.2 and Table 7.3. Yang's equation performs better than the rest with lower AARE of 43.5%, RMSE of 540.9 ppm, and a higher correlation coefficient of 0.80. Laursen's and Bagnold's equations are found to be the least accurate with higher AARE values of 136.2% and 118.1% respectively. The ranking is similar to the ranking for laboratory datasets in the sand size range. The advantage of presenting computation results graphically is evident here as well. Although the statistical model performance criteria AARE and Dr are comparable for the Engelund and Hansen (1972) and Ackers and White (1973) formulas, Figure 7-5 clearly illustrates there is a better match between observed values and the results of the Engelund and Hansen equation.

It should be mentioned that this ranking is applicable for the datasets used in this analysis and if a different group of datasets is used, the ranking can be different. As expected the accuracy for field data is lower than the laboratory data because of the complicated nature of sediment transport in the field. These results indicate that the ANFIS based approach can be applied successfully and provides high accuracy and reliability for computing sediment transport in natural rivers as well if sufficient data is available.

7.2.2 Total Bed-Material Load for the Elbe

The accuracy of the ANFIS model developed to compute total bed-material transport rate for the Elbe (section 6.5.1) is also investigated by comparing it with the results obtained from five sediment transport equations. The van Rijn (1984) equation is the most commonly applied for predicting sediment transport in sandy rivers as the Elbe river. The Comparison of the total bed-material load ANFIS model with four input parameters on the training and test datasets with the results computed by the total bed-material load equations of Yang (1973), Ackers and White (1973), Engelund and Hansen (1972), Bagnold (1966), and Van Rijn (1984a,b) is shown in Figure 7-6 below. The statistical model performance evaluation criteria for the different models are summarized in Table 7. 4.

From the scatter plots of measured total bed-material load transport rate per width (g/s.m) versus computed total bed-material load transport rates (Figure 7-6), it can be clearly observed that the data-driven adaptive neuro-fuzzy model performs significantly better than the selected sediment transport equations. The transport equations result in large scatter for the computed results. The ANFIS model is selected as the best model with the lowest AARE of 51.7%, RMSE of 65 g/s.m, and the highest correlation coefficient of r = 0.72. The average discrepancy ratio of the data-driven model is 1.27 and 68.7% of the computed values lie within the discrepancy range of 0.5 - 1.5.

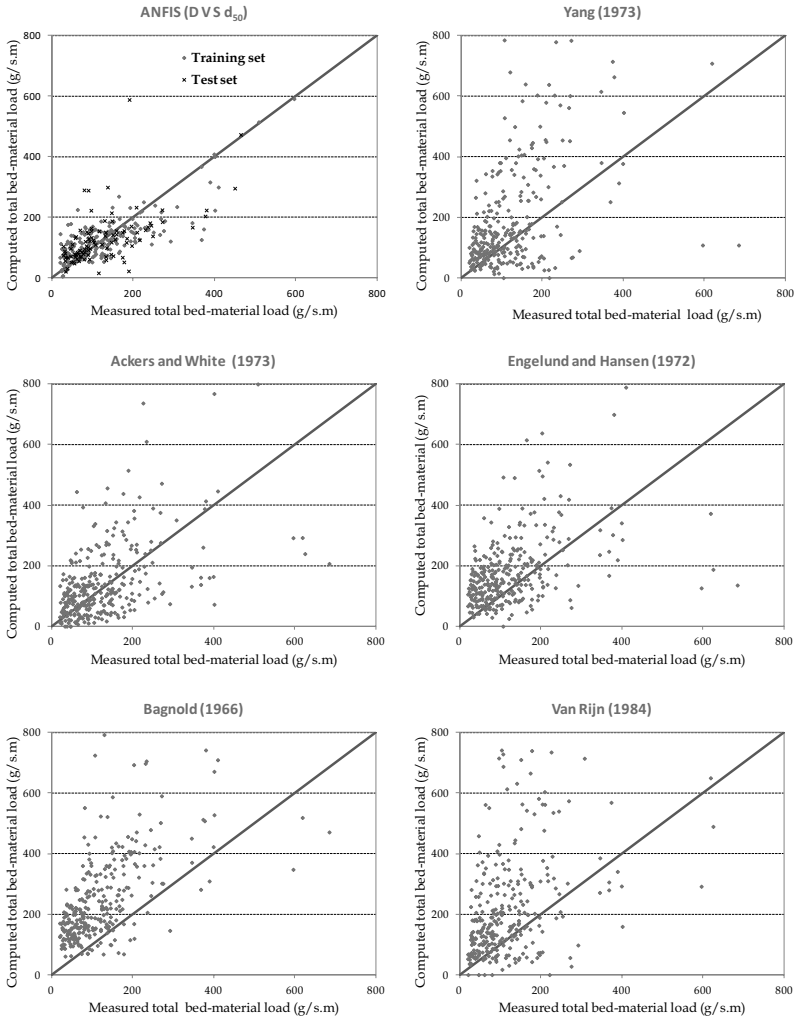


Figure 7-6: Measured versus computed total bed-material transport rate per width by using ANFIS model and selected transport equations for the Elbe.

Table 7.4: Comparison of performances of selected sediment transport equations and ANFIS model for the Elbe.

Model	AARE (g/s.m)	RMSE (g/s.m)	r	Avg. Dr	Percent of data Dr within			ARE >100 %
					0.75-1.25	0.5-1.5	0.25-1.75	
ANFIS (D V S d_{50})	51.7	65.0	0.72	1.27	42.9	68.7	78.5	13.4
Yang	121.6	284.5	0.62	1.99	18.4	34.4	47.2	36.2
Engelund and Hansen	116.8	322.8	0.52	2.00	22.1	40.2	52.2	35.9
Ackers and White	85.9	177.8	0.53	1.46	22.4	40.8	63.2	20.6
Bagnold	167.2	207.1	0.67	2.63	9.8	20.6	29.1	56.1
van Rijn	165.1	344.8	0.58	2.54	15.6	27.9	39.6	46.9

The AARE values for Yang (1973), Engelund and Hanson (1972), Ackers and White (1973), Bagnold (1966) and van Rijn (1984) are 121.6, 116.8, 85.9, 167.2, 165.1 % respectively. The AARE of all equations except the Ackers and White is greater than 100% and the deviations from the line of perfect agreement are significant. Bagnold's equation is the least accurate with 56.1% of the computed values having absolute relative error greater than 100%. The Ackers and White equation performs relatively better than the other equations with average discrepancy ratio of 1.46 and 40.8% of the values lying between 0.5 - 1.5 times the observed transport rate. These results further strengthen the complexity of computing sediment transport rates in large rivers like the Elbe where a great deal of factors plays a significant role and the uncertainty is very high. The results from the data-driven model are fairly accurate and provide better estimation of total bed-material load transport rate. The ANFIS model is quite simple to apply and results are promising. But like any regression and data-driven model, the ANFIS model developed for estimating total bed-material load rate in the Elbe captures the hidden relationship between the input variables and the total bed-material load transport rate for the river. This relationship can be different for other rivers and the model developed for the Elbe cannot be directly applied to river reaches with different characteristics. But the sediment transport equations are supposed to be generally applicable.

7.2.3 Total Bed-Material Load for the Rhine

Comparison of computed and observed total bed-material load transport rates has also been done for the Rhine river. The results further strengthen the better performance of the ANFIS model. Since most transported particle sizes in the Rhine River are quite coarse, the transport functions fail to predict the observed sediment load with reasonable accuracy and perform worse than for the Elbe River. The scatter plots of measured versus computed total bed-material transport rate by using the ANFIS model with four variables (section 6.4.1) and the Ackers and White's equation are shown in Figure 7-7. The results in Figure 7-7 clearly show

that the ANFIS model is more accurate because the deviations (scatters) from the line of perfect agreement are less.

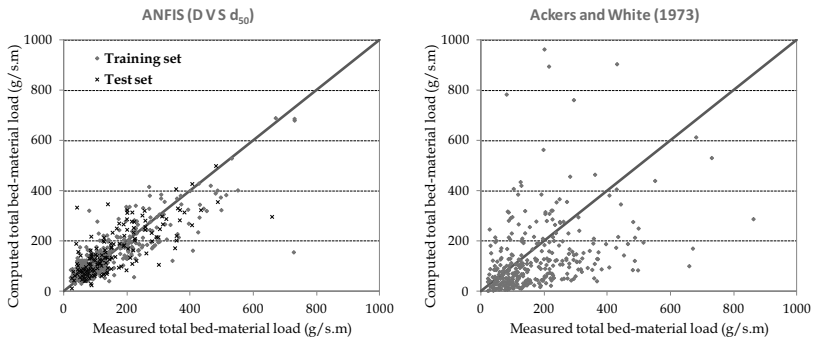


Figure 7-7: Scatter plots of measured versus computed total bed-material load transport rate per width by ANFIS model with four parameters and the Ackers and White (1973) equation for the Rhine.

7.2.4 Bed Load for the Rhine

The Meyer-Peter and Muller's (1948) equation is selected for comparison because it is the most widely used equation for predicting the amount of bed load transport. Additionally, Parker's (1990) equation which is developed for calculating gravel transport in natural rivers is implemented. The scatter plots of computed versus measured bed load transport rates for the ANFIS model with four input variables (section 6.4.2). and the results from the Meyer-Peter and Müller (1948) equation are shown in Figure 7-8.

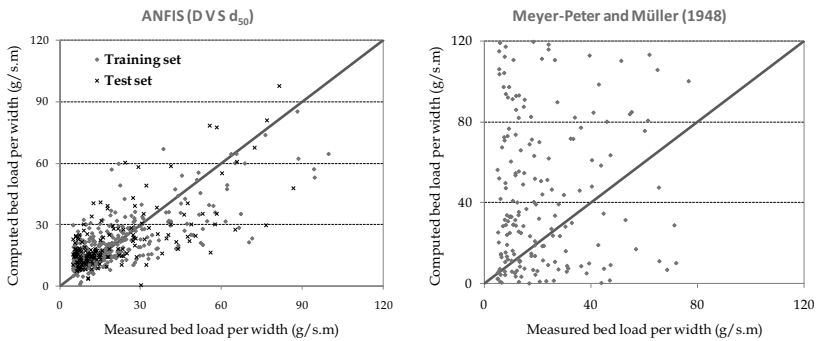


Figure 7-8: Scatter plots of measured versus computed bed load transport rates per width by ANFIS model with four input parameters and the Meyer-Peter Müller (1948) equation.

From Figure 7-8, it can be observed that the bed load ANFIS model estimates the observed transport rates better than the Meyer-Peter and Müller (MPM) equation. The AARE is 55.4% for the ANFIS model, whereas the AARE of MPM equation is greater than 100%. The correlation coefficient of the ANFIS model is 0.73 which is better than that of the equation, and 77.5% of the computed values are in the discrepancy ratio range 0.25 - 1.75. Only 16.3% of the computation results have an absolute relative estimation error of 100%. Figure 7-8 reveals that the ANFIS model performs better with less scatter. As it is pointed out in the previous section, this model is specific for estimating bed load transport in the Rhine.

In general the measured amount of bed load transport rate is sensitive to the location of the sampling device and the duration of measurement and this creates a significant challenge in accurately measuring it in field conditions. Transverse variability of hydraulic parameters significantly affect bed load transport and usually utilizing cross-section averaged values of hydraulic parameters result in high error (Bertoldi et al., 2008). Considering the uncertainty in measurement and the associated error, the ANFIS model is considered to be sufficiently accurate. It is usually required to get reasonable estimates not accurate values of transport rates in large rivers like the Rhine. This result can be used for implementing sediment management strategies for the Rhine, which is one of the main navigation routes in Europe.

7.3 Summary

The estimation accuracy of the adaptive neuro-fuzzy model is compared to that of other well known sediment transport equations. The results of the analysis in the previous sections prove the potential of the data-driven adaptive neuro-fuzzy modelling for estimation of both bed load and total bed-material transport rates. The model results show that the data-driven adaptive neuro-fuzzy modelling approach can be used for reasonably accurate estimation of sediment transport rates. The ANFIS models are found to be performing better than selected sediment transport equations for all the datasets used in this research. The ANFIS model results are quite promising for the laboratory datasets both in the sand and gravel size ranges. The ANFIS model outputs of the field data are also acceptable and more accurate than available equations. The approach is implemented to compute bed load and total bed-material load for the Rhine and Elbe rivers which are quite large and none of the existing equations are found to be reliable. In general, total bed-material load transport rates are estimated with better accuracy than bed load transport rates, this is also the case when other sediment transport equations are used. The measurement of bed load in most natural rivers is difficult and the accuracy and reliability of the measuring instruments is generally less. In general, most sediment transport equations are originally developed using laboratory data and when they are applied to natural rivers where bed composition is usually non-uniform, their accuracy decreases dramatically.

The plots of measured versus computed transport rates for the different groups of data presented in the previous sections demonstrate that the ANFIS model estimates show less scatter compared to the other equations. Additionally, the fuzzy models show the best performance with respect to model performance evaluation criteria summarized in the tables presented in this chapter. From the results, it can be concluded that data-driven models are capable of giving fairly accurate results and can be used as alternative tools in certain applications where the physically based equations are not as accurate. Sediment transport is definitely one of the focus areas in this regard as proved in this dissertation. Therefore, the approach presented in this research is believed to be worthy of further investigation. In general, computing the rate of sediment transport using existing equations requires: selecting an appropriate equation, comparing values computed by the equation with actual values measured in the field and calibrating the coefficients in the equation to attain a better matchup with the measured data, and finally validating the calibrated equation. The results of the conventional equations might improve if careful calibration is done and this should be considered while comparing the results from the equations with the results of the ANFIS models.

8 Model Validation

After developing any model the model validation is a very important step. It is an essential part of a model development process if the developed model is going to be accepted and implemented to solve problems and support decision making. Validation is the process of determining the degree to which a model represents the process accurately and is capable of giving acceptable results for the intended uses of the model. The validation data should be different from the data used for model optimization or calibration.

In this chapter, the ANFIS model developed for the Rhine River is validated by using input hydraulic parameters and computing the rate of sediment transport and then determining the sections of erosion and deposition in the Upper Rhine.

8.1 Input Hydraulic Data

The bed load and total bed-material load ANFIS models developed for the Rhine require four primary input parameters: flow velocity, water surface slope, water depth, and median particle size (d_{50}). These input variables are used to compute the total bed-material load and bed load transport rates. It is required to validate the ANFIS model of the Rhine using input hydraulic data from fifteen stations on the Upper Rhine from Iffezheim to Speyer with a spatial resolution of about 5 km between the stations. Average daily values of hydraulic variables depth, velocity, and the Manning's roughness coefficient are provided by the BfG for a period of 13 years from 1993-2006. The slope is computed using the Manning's equation. Additionally, the daily discharge and the corresponding flow width are provided for the stations. The median particle size for each of the stations is estimated from the particle sizes values provided previously for the development of the ANFIS model and for stations where particle size is not available, the value of a nearby station is used. In order to determine the median particle size for each year, interpolation is used between the available measured values. The data has been analyzed and filtering has been performed and the values where the unit discharge (V*D) and unit stream power (V*S) are less than those in the range of the training and test datasets during the development of the models are discarded.

Since measured bed elevation values of the 15 stations are available for 1993 and 2004, the analysis is done for years 1994-2004. From the available daily values, the mean discharge is 1263.4 m^3/s and the maximum discharge is 4511.2 m^3/s which is observed in May 1999. Since there is not any significant tributary coming into the reach, the discharge passing through the different sections is assumed to be the same.

Figure 8-1 shows these 15 stations and the corresponding station ID in river-km. Graphical representations of the statistical mean, maximum and minimum values of the input variables for the fifteen stations are shown in Figures 8-2, 8-3, and 8-4. The plots illustrate how the main input variables are varying along the different sections and are helpful to identify areas where there is a significant change in the hydraulic input variables. The mean, maximum, and minimum values for the flow width along the different stations are plotted in Figure 8-5. The detailed statistical distributions of the daily values of the input hydraulic variables for all stations are included in Appendix F.

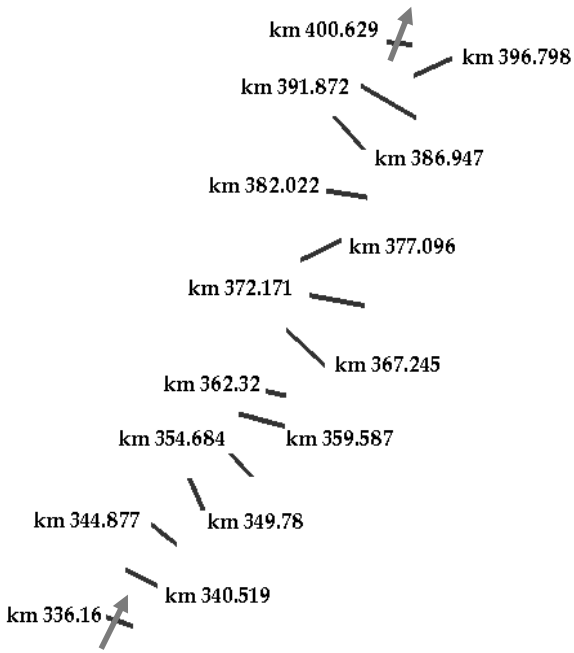


Figure 8-1: Schematic representation of the fifteen cross-sections in the Upper Rhine for which hydraulic input data is available.

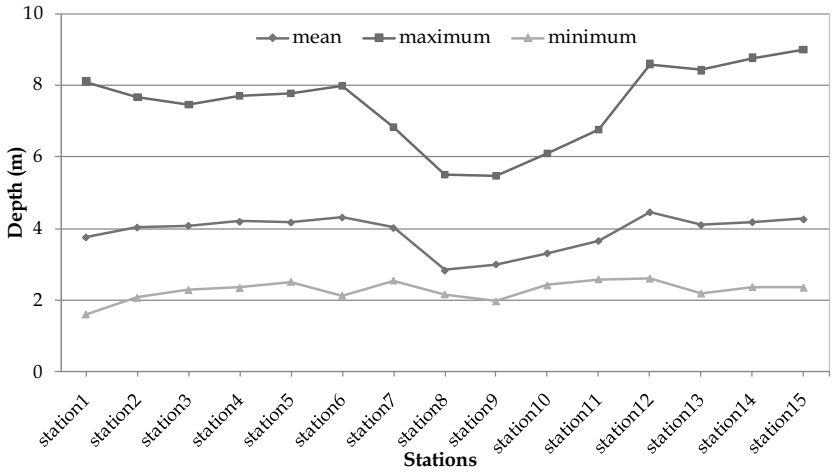


Figure 8-2: Mean, maximum, and minimum water depth for different stations in the Upper Rhine from 1994-2004.

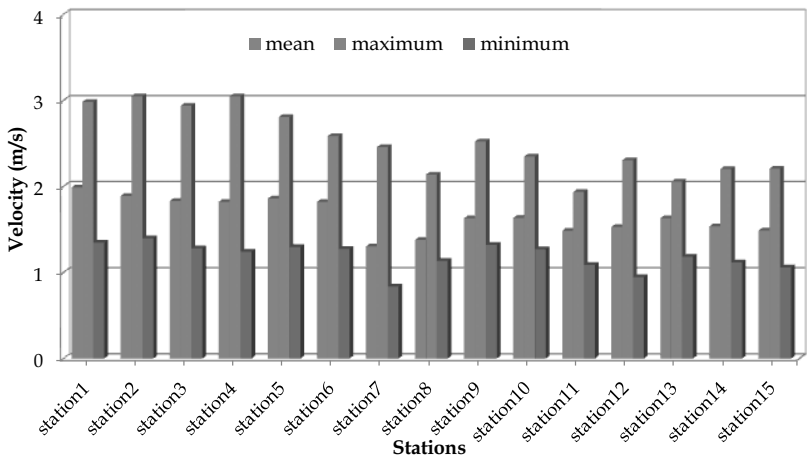


Figure 8-3: Mean, maximum, and minimum flow velocity for different stations in the Upper Rhine from 1994-2004.

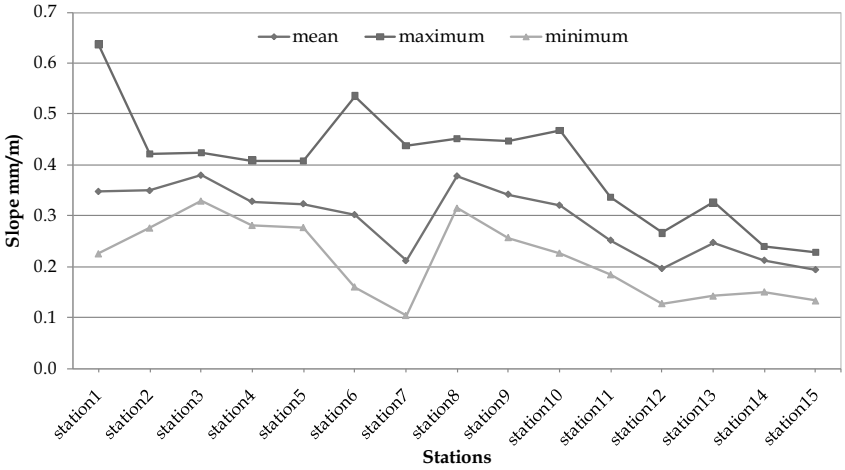


Figure 8-4: Mean, maximum, and minimum water surface slope for different stations in the Upper Rhine from 1994-2004.

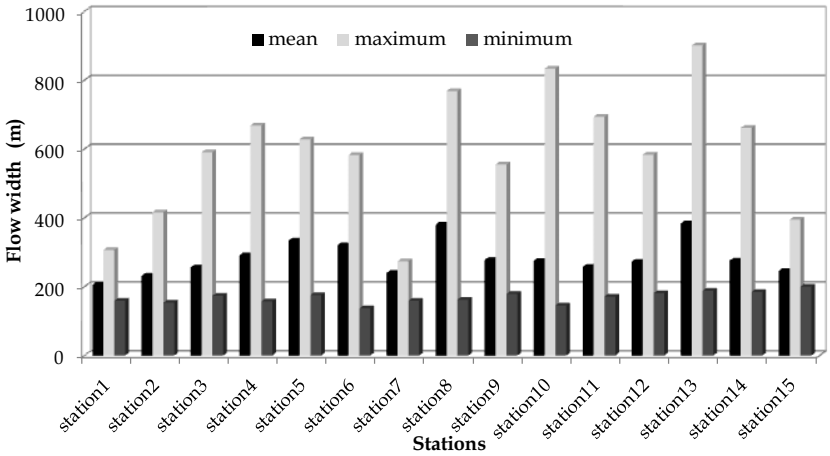


Figure 8-5: Mean, maximum, and minimum flow width for different stations in the Upper Rhine from 1994-2004.

8.2 Computation of Sediment Transport Rate

With the three hydraulic input variables for each of the stations and the median sediment particle size, the rate of total bed-material transport is computed using the ANFIS model developed for the Rhine in chapter 5. Then the sediment mass balance is computed from section to section. If the sediment load increases, there is erosion and if the sediment load decreases from section to section, there must have been deposition. The amount of erosion or deposition is the difference between the computed sediment loads for the consecutive sections. The analysis is performed for the 15 station which make up 13 sections between the stations. The bed level elevation for the stations is provided by the BfG for 1994 and 2004. In order to evaluate the difference in the computed values of sediment load, the datasets from 1994 to 2004 are used.

The total bed-material load ANFIS model with four input variables (V, S, D, and d_{50}) and the ANFIS model with three input variables (V, S, D) are used for computing the total bed-material transport in g/s.m. The ANFIS model with three input parameters has a comparable accuracy as the one with four input parameters (including particle size) for the training and test datasets during the development of the data-driven fuzzy model (refer section 6.4). Particle size is not available for the whole range of data, and values for nearby stations are assigned, and for the corresponding years interpolation is performed to get the particle size. The main aim of this chapter is not to get exact values but to show how the model can be implemented in the future and do a rough validation. The computed values from the ANFIS model are sediment transport rates per width (g/s.m). These are multiplied by width in order to obtain the mass transport rate. The total bed-material load transport rate is computed from:

$$M_{t,i} = \frac{q_{t,i} * W_i}{1000} \tag{8.1}$$

where:

$q_{t,i}$ = total bed-material transport rate for each day computed using the ANFIS model (g/s.m),

$M_{t,i}$ = the mass total bed-material transport rate (kg/s),

W_i = top width of the water surface elevation for the given date and station (m).

This total mass sediment transport rate can be converted into volume using the bulk density of gravel (ρ_g). The volumetric total bed-material load for each time step (day) can be estimated from:

$$V_{t,i} = \frac{M_{t,i} * 24 * 3600}{\rho_g} \tag{8.2}$$

Where $V_{t,i}$ = volumetric sediment load for each day (m^3).

The total volumetric bed-material load transported in a station is the summation of the transport values computed for each combination of input variables and this can be used for the analysis of erosion and deposition section in the reach.

$$V_t = \sum_{i=1}^{N} V_{t,i} \tag{8.3}$$

where:

V_t = total volumetric bed-material transport for the total period 1994-2004,

N = the total number of days for the ten years.

Table 8.1 summarizes the total volume of bed-material computed for each station using two ANFIS models, (V S D d_{50}) and (V S D), for the period 1994 - 2004. The nearby sediment measuring stations are also indicated for the cross-sections. Datasets from the nearby stations are used in chapter 5 to develop the data-driven ANFIS models from measured transport rates and input hydraulic variables.

Table 8.1: Total volumetric bed-material transported for the period of 1994 - 2004 computed by using two ANFIS models.

Station (Rhine km)	Nearby measuring station	V_t (Mil. m^3) ANFIS (V S D d_{50})	V_t (Mil. m^3) ANFIS (V S D)
336.16	Iffezheim @ km 337.39	12757.4	12534.3
340.519	Plittersdorf (alt) @ km 340	17752.5	16247.4
344.877	Plittersdorf@ km 342.7	17835.4	17328.5
349.78	Illingen @ km 347	23607.8	22522.1
354.684		25469.6	25321.8
359.587	Neuburgweier @ km 356	21276.9	21614.1
362.32	Maxau @ km 362.7	6304.4	6783.13
367.245		7413.0	7365.06
372.171	Leimersheim @ km 371.8	7143.6	7245.44
377.096		9574.4	9587.87
382.022		8485.8	8085.57
386.947		11711.1	11976.1
391.872	Philippsburg @ km 390	20289.2	20052.4
396.798		11542.9	11798.9
400.629	Speyer @ km 403.5	9521.0	9826.96

The total amount of volumetric bed-material load can be used to identify sections of erosion and deposition in the reach. This can be easily converted into mass transport rate and Figure 8-6 illustrates a graphical representation of the mean total bed-material transport rate for the 15 stations.

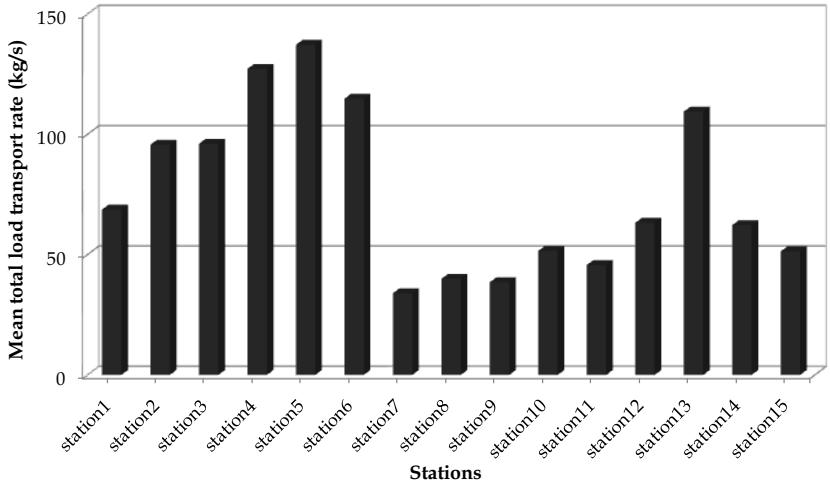


Figure 8-6: Mean total bed-material transport rates for the 15 stations between 1994 and 2004.

8.3 Erosion and Deposition Processes

Natural rivers continuously change the shape of their cross-sections because of the process of erosion and deposition. River bed erosion and deposition are challenges for implementation of sustainable projects in hydraulic engineering. They are also major causes of failure of projects resulting in huge economic loss and sometimes human lives. It is important to adequately assess and predict channel aggradation and degradation patterns and implement appropriate mitigation measures. Aggradation refers to an increase in bed elevation due to deposition and degradation is a decrease in bed elevation through the process of erosion. When sediment transport capacity exceeds sediment supply, channel bed degradation occurs and when the sediment supply is greater than the sediment transport capacity, deposition occurs. In order to determine the amount of erosion and deposition, sediment continuity analysis is necessary.

Sediment Continuity Analysis

Sediment continuity analysis is the application of the law of conservation of sediment mass or sediment volume respectively, in a river reach. According to continuity, the volume of

sediment deposited in a section or eroded from a section during a given period of time is computed as the difference between the volumes of sediment entering and leaving the reach (Shields, 2007):

$$\Delta V_s = V_{s_{out}} - V_{s_{in}} \tag{8.4}$$

where:

ΔV_s = volume of sediment deposited or eroded,

$V_{s_{out}}$ = volume of sediment transported out of the reach or leaving the section,

$V_{s_{in}}$ = volume of sediment entering the section.

The resulting average amount of bed level change in the section is computed by dividing the change in sediment volume (ΔV_s) by the area of the section.

The annual rate of bed degradation or aggradation can be computed by dividing the total volumetric rate by the total area of the section assuming the section is rectangular and the change in geometry is uniformly distributed along the section. This is a rough assumption and is similar to what is used in one-dimensional hydraulic models. In reality, both erosion and deposition can occur at a given cross-section depending on the cross-section geometry and bed substrate. In this simplified assumption, only either erosion or deposition is allowed at a section. To compute the bed level change at a section per year, the volumetric total annual bed-material load is divided by the bed area of the section. The area is computed by multiplying the average width of the cross-section by the length of the section.

$$\Delta z = \frac{V_{t,i} - V_{t,i-1}}{W_i * 0.5 * (L_{i-1,i} + L_{i,i+1})} \tag{8.5}$$

where:

Δz = the annual bed level change (m),

W_i = the width of station i,

$L_{i-1,i}$ = the length between the current station and the upstream station,

$L_{i,i+1}$ = the length between the current station and the downstream station.

The length of a section is taken to be half of the sum of the lengths between the current station and the upstream and downstream stations. The results of the computation of erosion and deposition processes identify sections of bed degradation and aggradation. This can be used for applying proper sediment management strategies and maintain the stability of a section. The average lengths of the sections are provided in Table 8.2.

In order to compare the sections of erosion and deposition identified using the ANFIS model with that of the actual observed changes in bed elevation, measured values of cross-section geometry are analyzed. Measured values of bed elevation data are available for 1993 and 2004. The difference of the bed elevations between these values is used to identify the sections of erosion and sedimentation in the reach. The plot of the measured cross-section elevations for station 2 are shown in Figure 8-7 for 1993 and 2004, as an example.

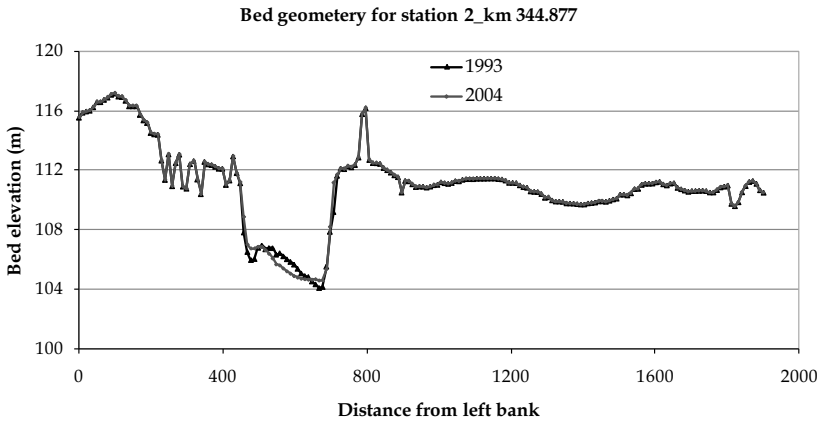


Figure 8-7: Measured cross-section geometry of station 2 for the years 1993 and 2004.

Figure 8-7 illustrates that there are both erosion and deposition in this station in the lateral direction. To determine whether the section is a net erosion or deposition section, the areas of erosion (negative) and deposition (positive) are summed up. Figure 8-8 illustrates how this calculation is performed by taking the points with change in bed elevation. The net area is computed by using the trapezoidal rule. Alternatively, the summation of the bed level changes can be used to identify whether there is a net erosion or deposition in the station because the points are equally distributed.

For example in Figure 8-8, the computation result shows that there is net erosion in this station. The calculated area is -2.53 m^2. Similar analysis is performed for all remaining stations and sections of erosion and deposition is identified. These results are compared with the results from the ANFIS model.

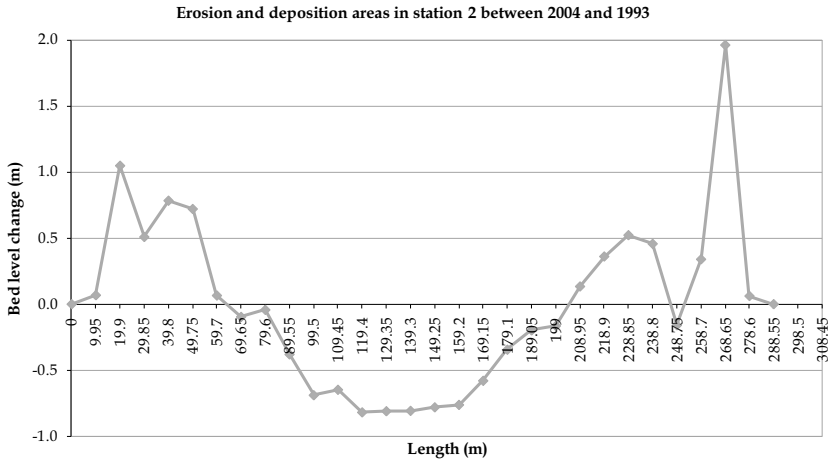


Figure 8-8: Computation of net erosion or deposition in a station by using the trapezoidal rule for station 2.

The summary of the analysis is presented in Table 8.2.

Table 8.2: Sections of erosion and deposition identified by using measured bed level elevations, and sediment balance analysis of the outputs from ANFIS model.

Station (Rhine km)	Analysis of the measured bed elevation values	Section	Section length	Computed by using ANFIS (V S D d_{50}) ΔV_t (Mil. m^3)	Erosion or deposition
336.16	Deposition				
340.519	Erosion	Section1	4358.5	3713.10	
344.877	Erosion	Section2	4630.5	1081.12	Erosion
349.78	Erosion	Section3	4903.5	5193.54	Erosion
354.684	Erosion	Section4	4903.5	2799.73	Erosion
359.587	Erosion	Section5	3818.0	-3707.77	Deposition
362.32	Deposition	Section6	3829.0	-14830.91	Deposition
367.245	Deposition	Section7	4925.5	581.93	Erosion
372.171	Erosion	Section8	4925.5	-119.62	Deposition
377.096	Deposition	Section9	4925.5	2342.43	Erosion
382.022	Deposition	Section10	4925.5	-1502.30	Deposition
386.947	Deposition	Section11	4925.0	3890.48	Erosion
391.872	Deposition	Section12	4925.5	8076.31	Erosion
396.798	Deposition	Section13	4378.5	-8253.47	Deposition
400.629	Erosion				

The first section (Iffezheim) is neglected because it is not used in the development of the ANFIS model. It is a station where there is an artificial sediment supply in the Rhine. From the computation results summarized in Table 8.2, it can be seen that the ANFIS model estimates accurately three out of five erosion sections and three out of seven deposition sections observed. It should be noted that the measured bed elevation values are with a spatial resolution of about 5 km. There can be other sections of erosion or deposition which are between any two stations. This is a very crude approximation of the process. The results might match better if there is finer spatial resolution.

The average annual bed level change in the section between the periods of analysis can be easily computed by dividing the difference in sediment volumes between two stations by the area of the section which can be approximated by multiplying the length of the section by the average width. Figure 8-9 shows the average annual bed level change computed for the 13 sections in the Upper Rhine from 1994 - 2004. In the figure, sections of positive bed level change are sections of deposition where as negative bed level changes are sections of erosion.

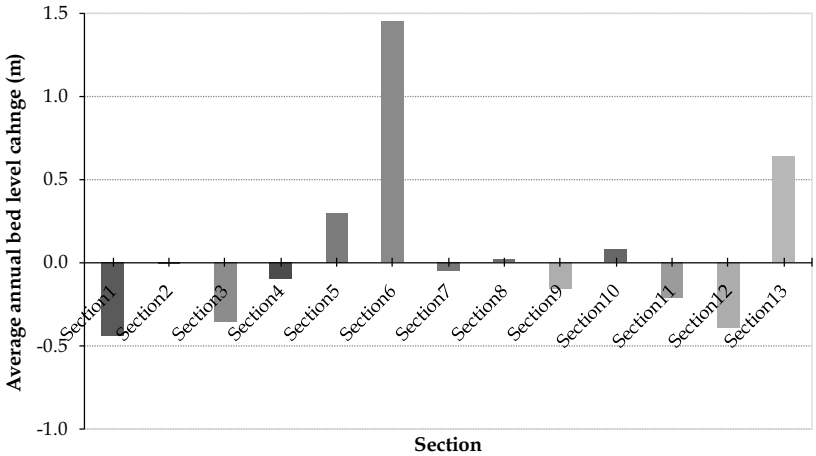


Figure 8-9: Average annual bed level change computed by analysing the ANFIS model outputs in the Upper Rhine between 1994 and 2004.

As it is mentioned previously, this is a very rough computation. Detailed assessment of erosion and deposition, and computation of long term morphology change in a river reach for given inputs of discharge and sediment require a long term computer model simulation. One and two dimensional hydraulic models are usually used to perform the complete analysis. One dimensional models usually simulate changes in bed elevation in stream wise distance, but ignore variations in the transverse direction. Two dimensional models are able

to compute transverse variations, but do not simulate variations in velocity in the vertical direction (Shields, 2007). One dimensional models like Mike 11 (Danish Hydraulic Institute, 2002), SRH-1D (Greimann and Huang, 2007), HEC–6 (USACE, 1993) are examples of sediment transport models that can be used for single event or long-term degradation and aggradation computation. GSTARS3 (Yang and Simoes, 2002) is a semi-2D model which uses the concept of stream tubes and is also capable of computing width adjustments as well. Two dimensional models like SRH-2D (Lai, 2008) and MIKE 21C (Danish Hydraulic Institute, 2003) can be used to get better results, because erosion and deposition can be determined at a higher resolution. The models route sediment along the channel and adjust the channel geometry (usually bed elevation, and some also lateral bank position) according to the computed sediment transport capacity and sediment supply rate.

8.4 Discussion

The amount of erosion and deposition for fifteen cross-sections is computed using the ANFIS model developed for the Rhine. The computation results identify the sections of erosion and deposition in the Upper Rhine. Most of the assumptions made here are very simplified and focus on how the model can be applied to implement proper sediment management strategies and assess long term river morphology changes. In order to accurately simulate the annual level of bed aggradation and degradation, proper hydraulic computation should be done. Different effects of bed sorting and armouring and other factors which are implemented in many morphologic models have to be accounted for. River beds are composed of non-uniform sediment material containing a wide range of sediment sizes. Usually selective transport where fine particles are easily transported leaving most coarse size for a given hydraulic condition. The layer of coarser materials left behind forms an armour layer that limits further erosion unless and until higher levels of shear stress destroy the armour layer (Shields, 2007). For example, armour layer formation is often observed downstream of dams. The complete sediment balance requires the consideration of further impacts like the sediment supply from tributaries, dredging operation, artificial supply of sediments, etc. For a successful model simulation, most of the numerical models require additional information like geologic or structural barriers, tributary inflow hydrographs for flow and sediment concentration. These points and how to integrate these effects should be taken into consideration while applying the ANFIS model. It would be interesting to implement the developed ANFIS model as one of the equations for sediment transport and calculate long term morphological changes in the sections

9 Summary and Outlook

9.1 Summary and Conclusions

The objective of this dissertation is to investigate the potential application of a data-driven fuzzy logic based modelling approach for computing the rate of sediment transport using measured laboratory and field datasets. Computation of sediment transport is a complicated process and is usually accomplished by using empirical or semi-empirical equations. There are many equations which are derived based on three basic approaches: regression, probabilistic, and deterministic. These approaches are briefly discussed and some of the commonly applied equations are presented. None of the existing equations has gained universal acceptance and the ranges of applications of the equations is quite different. For a given type of river with specific characteristics, the selection of an appropriate sediment transport equation is a trial and error procedure in which different equations have to be tried and the better one is selected and calibrated. The accuracy of existing equations is not satisfactory for practical application, especially for natural rivers. There are many factors which influence the amount of sediment under transport and because of this there is quite an uncertainty and fuzziness in the process. For such complicated processes, fuzzy logic is an alternative modelling approach. Fuzzy logic which is introduced by Zadeh (1965) is capable of incorporating uncertainty and imprecision in the modelling process. It is usually applied for qualitative analysis of processes. The four basic components of a fuzzy model are: fuzzification interface, fuzzy rule base, fuzzy inference, and defuzzifier. These are briefly summarized in chapter two. The fuzzy rules describing a system can be defined by experts or derived from analysis of collected data. Since the process of sediment transport is complicated and the rule system cannot be defined by experts accurately, a data-driven rule generation with the Takagi-Surgeno fuzzy inference system is implemented.

In this study, the data-driven fuzzy logic based modelling is accomplished by following four general steps. These four basic steps are: identification of input and output variables and data preparation, generation of initial membership functions using grid partitioning or fuzzy clustering, model optimization using ANFIS and performing a sensitivity analysis for the different parameters of the model, and the final step is model validation and testing.

In order to examine the reliability and performance of the presented approach, the investigation is done by using various categories of datasets obtained from different sources.

The most significant parameters influencing the rate of sediment transport are identified and used for generating the fuzzy model. These variables are: bed or energy slope, mean velocity, water depth, and the median particle diameter (d_{50}). The four basic variables are selected because they can be easily measured as primary input data, and their physical meaning is quite obvious for the future application of the model. For a successful implementation of any data-driven modelling technique, collection of sufficient data with good quality is crucial and should be carefully done. In this study, laboratory and field datasets are collected from different sources. The laboratory datasets are categorized into sand and gravel and separate analysis is performed. Measured values of the input hydraulic variables and total bed-material and bed load transport rates are available for the Rhine and Elbe rivers along different sections of the rivers, and these datasets are provided by the BfG. Detailed data analysis is done and extreme values are excluded from the datasets. Two third of the available data are used for model training and the remaining one third are used for testing. The test dataset is required to avoid overtraining during optimization. In general, it is recommended that the number of training data should be at least as big as the total number of parameters in the model to avoid overtraining.

Grid partitioning and fuzzy clustering are presented as alternative techniques for the generation of the initial fuzzy systems based on the ranges of the input variables in the different categories of datasets. The grid partitioning approach is selected for further application because it is better with regard to accuracy, and interpretability of the resulting final rule base. In this study, the adaptive neuro-fuzzy inference system (ANFIS) which is introduced by Jang (1993) is utilized for optimizing the initial fuzzy system. ANFIS is a hybrid optimization scheme which implements gradient descent to determine premise parameters and least squares technique to identify optimum consequent parameters. The basic architecture of ANFIS and the functions and parameters of the layers are described in chapter 3. Several architectures of ANFIS are tried to obtain the optimum model. In order to determine the relative importance of each of the four primary input parameters (D, S, V, and d_{50}) on the accuracy of the model results, a sensitivity analysis is performed with one of the inputs removed at a time. The different architectures of the models tried are ANFIS (D V S d_{50}), ANFIS (V S d_{50}), ANFIS (D S d_{50}), ANFIS (D V d_{50}), and ANFIS (D V S). Additionally, instead of using velocity and slope as separate parameters, unit stream power is used as a single parameter. The combination of unit stream power with water depth and grain size, ANFIS (VS D d_{50}), as well as with grain size only, ANFIS (VS d_{50}), resulted in a total of seven various combinations of input parameters. Furthermore, a sensitivity analysis is performed for the number (2 to 5) and type (triangular, trapezoidal, generalized bell-shaped, Gaussian)

of membership functions to achieve an optimized model. Finally three generalized bell-shaped membership functions are selected for each input variable based on the results of the sensitivity analysis.

The study also investigates the performance of the ANFIS models developed for computing sediment transport by comparing their results with the results of other commonly utilized sediment transport equations. For Laboratory data in the sand size range, equations of Yang (1973), Ackers and White (1973), Bagnold (1966), Engelund and Hansen (1972), Laursen (1958), Colby (1964), and Toffaleti (1969) are selected for comparison. The ANFIS model developed for computing gravel transport in laboratory flumes is compared with the equations of Yang (1984), Meyer-Peter and Müller (1948), Ackers and White (1973), and Parker (1990). Similarly, model comparison and performance analysis is done for the ANFIS models developed using the datasets for the Rhine and Elbe rivers. Correlation coefficient, root mean squared error, average absolute relative error, and discrepancy ratio are selected as statistical model performance evaluation criteria.

The results of this study prove that the data-driven fuzzy logic modelling approach can be successfully implemented for computing sediment transport both in laboratory flumes and natural rivers. The main idea behind the fuzzy logic is to obtain a qualitative analysis of complex processes. But if sufficient data with good quality is available, it is possible to obtain an accurate quantitative estimation with data-driven fuzzy logic modelling. From the results of the investigation it can be concluded that the ANFIS model, which does not consider the complicated physical process, estimates measured transport rates successfully. With respect to the statistical model performance evaluation criteria, the ANFIS model performs better than the selected transport equations for all categories of data analyzed (laboratory and field, sand and gravel, bed load and total bed-material load). Better agreement is found between the values computed by using the ANFIS model and measured sediment transport rates. The model comparison assessments show that the existing equations are unable to predict the measured transport rates accurately especially for rivers. The scatter plots comparing observed and computed transport rates (chapter 7) illustrate that the ANFIS models show less deviation. As expected, the estimation accuracy of the model for laboratory datasets is found to be much better than for field datasets. This is because the quality of measurements in laboratory flumes is better than in field conditions where there is a lot of uncertainty. Furthermore, the sensitivity analysis shows that different parameters are more important in influencing the amount of sediment transport depending on the ranges of input data, and river characteristics. For example, for laboratory flumes, the depth variation is not that important and therefore it can be excluded from the input variables. But for large rivers like the Rhine and Elbe, flow depth and velocity are found to be influential parameters. In the case of the Rhine river which is a gravel bed river, particle size affects the model

performance more than slope whereas for the Elbe river which is a sandy river the slope has more influence. This further proves how the fuzzy model can also provide some insight about the characteristics of the rivers and the governing variables for sediment transport processes in different morphologic regimes. The ANFIS model developed for the Rhine river is further validated by using daily hydraulic input parameters from fifteen stations on the Upper Rhine from Iffezheim to Speyer for a period of ten years (1994-2004). This is accomplished by computing the rate of sediment transport as well as identifying sections of erosion and deposition by applying sediment continuity analysis.

In a nutshell, this dissertation demonstrates the ability of simple neuro-fuzzy models to successfully capture complex relationships in datasets, and the implementation to computing sediment transport rates where the existing sediment transport equations fail to give reasonable results. The results of this research are highly promising and prove the potential applicability of the data-mining approach to modelling sediment transport. It is concluded that the proposed new data-driven fuzzy logic based modelling approach can be applied to compute sediment transport. The relatively simple ANFIS model is a powerful tool where the physical processes are too complicated to be expressed mathematically. The results from the present study indicate that employing measured sediment transport values, and selected hydraulic parameters, it is possible to obtain a fuzzy model that can successfully estimate measured bed load and total bed-material load transport rates.

However, it needs to be pointed out that data-driven models do not consider mathematical expressions describing the physical process and they are developed and optimized based on the analysis of available measured data. This is a major drawback in the case of extrapolating transport rates using variable input values that are not within the range of the training datasets. The models developed are usually data specific. They require collection of sufficient data with good quality describing the process under consideration. Data analysis and removal of extreme values is one of the important steps and should be seriously done. To assess the general ability of the ANFIS model in successfully estimating sediment transport for different rivers under varying conditions accurately, more training datasets from additional rivers should be included in the analysis. Further research with additional data from other river reaches is required to generalize the model.

9.2 Perspective of Future Work

The results of the data-driven ANFIS approach prove its potential as alternative modelling approach for estimating sediment transport. Additional works should be done in the future to better understand the role of other morphological parameters and their effect on model results, and to come up with a generalized model. There is always room for improvement

and based on the results obtained and the experience gained in this study, the following recommendations and suggestions are provided for further improvement and future research:

- **Model transferability** should be tested with datasets from other river reaches. Further research should be done to **generalize the models** by including more datasets from additional river reaches with different morphological and hydraulic characteristics to cover various ranges of input and output variables. Introducing additional parameters to describe river morphology, and other important site characteristics should be explored.

- In this dissertation, analysis of **interpretability** of the final parameters of the membership functions has not been done in detail. It is necessary to investigate the output membership function parameters further and see if there is a relationship which can be used to generalize their properties. Furthermore, detailed analysis of final input membership functions is required to examine if there is a pattern that can be identified and be used to come up with membership functions that are transferable. Implementing **dimensionless variables** and their ability to improve model accuracy and interpretability of final membership functions should be investigated. The dimensionless variables remove the problem of having different dimensions and might be helpful in developing a general ANFIS model.

- **Further validation** of the fuzzy models should be carried out by computing erosion and deposition processes. A method should be investigated to incorporate the fuzzy model into hydraulic models in order to simulate the long-term morphological development of the rivers. Whenever available, the inclusion of newly collected data to update the parameters of the input and output membership functions is necessary.

- The ANFIS models developed here are all one-dimensional obtained by using average hydraulic variables. However, transport rates and hydraulic variables in natural rivers show spatial variation in the transverse direction as well and a method to obtain **two-dimensional** results should be explored.

- The implementation of fuzzy **expert systems** by using the **Mamdani inference** system should also be examined. Implementation of additional simple parameters to take into account river morphology, bed forms, etc. is recommended in the expert system. It should be attempted to analyze and describe the results in terms of hydraulic processes. Further works aimed at improving the modelling approach and accuracy should be done.

References

Abebe, A. J., Solomatine, D. P., and Venneker, R. G. W. (2000). "Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events." *Hydrological Sciences Journal*, 45(3), 425–436.

Ackers, P., and White, W. R. (1973). "Sediment transport: a new approach and analysis." *Journal of the Hydraulics Division, ASCE*, 99(11), 2041–2060.

Adriaenssens, V., Baets, B. D., Goethals, P. L. M., and Pauw, N. D. (2004). "Fuzzy rule-based models for decision support in ecosystem management." *Science of The Total Environment*, 319(1-3), 1–12.

Alekseevskiy, N. I., Berkovich, K. M., and Chalov, R. S. (2008). "Erosion, sediment transportation and accumulation in rivers." *International Journal of Sediment Research*, 23(2), 93–105.

Ali, Y. M., and Zhang, L. (2001). "A methodology for fuzzy modeling of engineering systems." *Fuzzy Sets and Systems*, 118(2), 181–197.

Alonso, C. V., Neibling, W. H., and Foster, G. R. (1981). "Estimating sediment transport capacity in watershed modeling." *Transactions of the ASAE*, 24(5), 1211–1220.

Azar, A. T. (2010). "Adaptive Neuro-Fuzzy Systems." *Fuzzy Systems*, A. T. Azar, ed., InTech.

Bagnold, R. A. (1966). *An approach to the sediment transport problem from general physics*. U.S. Geological Survey Professional Paper 422-I.

Bakhtyar, R., Ghaheri, A., Yeganeh-Bakhtiary, A., and Baldock, T. E. (2008). "Longshore sediment transport estimation using a fuzzy inference system." *Applied Ocean Research*, 30(4), 273–286.

Bárdossy, A., and Disse, M. (1993). "Fuzzy rule-based models for infiltration." *Water Resources Research*, 29(2), 373–382.

Bárdossy, A., and Duckstein, L. (1992). "Analysis of a karstic aquifer management problem by fuzzy composite programming." *Water Resources Bulletin*, 28(1), 63–74.

Bárdossy, A., and Duckstein, L. (1995). *Fuzzy rule-based modeling with applications to geophysical, biological, and engineering systems*. CRC Press.

Barreto-Neto, A. A., and de Souza Filho, C. R. (2008). "Application of fuzzy logic to the evaluation of runoff in a tropical watershed." *Environmental Modelling and Software*, 23(2), 244–253.

Barry, J. J., Buffington, J. M., Goodwin, P., King, J. G., and Emmett, W. W. (2008). "Performance of bed-load transport equations relative to geomorphic significance: predicting effective discharge and its transport rate." *Journal of Hydraulic Engineering*, 134(5), 601–615.

van den Berg, J. H. (1987). "Bedform migration and bed-load transport in some rivers and tidal environments." *Sedimentology*, 34(4), 681–698.

Bertoldi, W., Ashmore, P., and Tubino, M. (2009). "A method for estimating the mean bed load flux in braided rivers." *Geomorphology*, 103(3), 330–340.

Bezdek, J. C. (1973). "Fuzzy mathematics in pattern classification." PhD Thesis, Applied Math. Center, Cornell University, Ithaca, New York, USA.

Bezdek, J. C. (1980). "A convergence theorem for the fuzzy clustering algorithms." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(1), 1–8.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Plenum Press.

Bezdek, J. C., Ehrlich, R., and Full, W. (1984). "FCM: The fuzzy c-means clustering algorithm." *Computers and Geosciences*, 10(2-3), 191–203.

Bhattacharya, B., Price, R. K., and Solomatine, D. P. (2005). "Data-driven modelling in the context of sediment transport." *Physics and Chemistry of the Earth*, 30(4-5), 297–302.

Bhattacharya, B., Price, R. K., and Solomatine, D. P. (2007). "Machine learning approach to modeling sediment transport." *Journal of Hydraulic Engineering*, 133(4), 440–450.

Bosserman, R. W., and Ragade, R. K. (1982). "Ecosystem analysis using fuzzy set theory." *Ecological Modelling*, 16(2-4), 191–208.

Brown, M., and Harris, C. (1994). *Neurofuzzy adaptive modelling and control*. Prentice Hall.

Brownlie, W. R. (1981). *Prediction of Flow Depth and Sediment Discharge in Open Channels*. W.M. Keck Laboratory of Hydraulics and Water Resources, California Institute of Technology, Pasadena.

Casey, H. J. (1935). *Uber Geschiebebewegung*. Preuss. Versuchsanst. fur Wasserbau und Schifibau (Translation on file at U.S. Soil Conservation Service, Washington, DC.), .

Casillas, J., Cordón, O., Herrera, F., and Magdalena, L. (2003). "Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview." *Interpretability Issues in Fuzzy Modeling*, 3–22.

Chang, F.-J., and Chang, Y.-T. (2006). "Adaptive neuro-fuzzy inference system for prediction of water level in reservoir." *Advances in Water Resources*, 29(1), 1–10.

Chang, F.-J., Hu, H.-F., and Chen, Y.-C. (2001). "Counterpropagation fuzzy-neural network for streamflow reconstruction." *Hydrological Processes*, 15(2), 219–232.

Chen, M.-Y., and Linkens, D. A. (2004). "Rule-base self-generation and simplification for data-driven fuzzy models." *Fuzzy Sets and Systems*, 142(2), 243–265.

Chen, Q., and Mynett, A. E. (2003). "Integration of data mining techniques and heuristic knowledge in fuzzy logic modelling of eutrophication in Taihu Lake." *Ecological Modelling*, 162(1-2), 55–67.

Chen, S.-H., Lin, Y.-H., Chang, L.-C., and Chang, F.-J. (2006). "The strategy of building a flood forecast model by neuro-fuzzy network." *Hydrological Processes*, 20(7), 1525–1540.

Chien, N., and Wan, Z. (1999). *Mechanics of sediment transport*. American Society of Civil Engineers, Reston, VA.

Chiu, S. L. (1994). "Fuzzy model identification based on cluster estimation." *Journal of Intelligent and Fuzzy systems*, 2(3), 267–278.

Colby, B. R. (1964). "Practical computations of bed-material discharge." *Journal of the Hydraulics Division, ASCE*, 90(2), 217–246.

Colby, B. R., and Hembree, C. H. (1955). *Computation of total sediment discharge, Niobrara River near Cody, Nebraska*. U.S. Geological Survey Water Supply Paper 1357.

Cornelissen, A. M. G., van den Berg, J., Koops, W. J., Grossman, M., and Udo, H. M. J. (2001). "Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory." *Agriculture, Ecosystems and Environment*, 86(2), 173–185.

Danish Hydraulic Institute. (2002). *MIKE11, a modelling system for rivers and channels, reference manual*. DHI Water and Environment, Hørsholm, Denmark.

Danish Hydraulic Institute. (2003). *MIKE 21C-2D river hydraulics and morphology, user's manual*. DHI Water and Environment, Hørsholm, Denmark.

Daunicht, W., Salski, A., Nohr, P., and Neubert, C. (1996). "A fuzzy knowledge-based model of annual production of skylarks." *Ecological modelling*, 85(1), 67–73.

Doğan, E., Yüksel, İ., and Kişi, Ö. (2007). "Estimation of total sediment load concentration obtained by experimental study using artificial neural networks." *Environmental Fluid Mechanics*, 7(4), 271–288.

Dubois, D., and Prade, H. M. (1980). *Fuzzy sets and systems: theory and applications*. Academic Press.

Einstein, H. A. (1944). *Bed-load transportation in Mountain Creek*. U.S. Department of Agriculture, Soil Conservation Service Technical Paper 55.

Einstein, H. A. (1950). *The bed-load function for sediment transportation in open channel flows*. U.S. Department of Agricultur, Agriculture, Soil Conservation Service, Technical Bulletin 1026.

Emerson, C. W. (1991). "A method for the measurement of bedload sediment transport and passive faunal transport on intertidal sandflats." *Estuaries and Coasts*, 14(4), 361–371.

Engelund, F., and Hansen, E. (1972). *A monograph on sediment transport in alluvial streams*. Teknisk Forlag.

Espinosa, J., and Vandewalle, J. (2000). "Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm." *IEEE Transactions on Fuzzy Systems*, 8(5), 591–600.

Firat, M., and Güngör, M. (2007). "River flow estimation using adaptive neuro fuzzy inference system." *Mathematics and Computers in Simulation*, 75(3-4), 87–96.

Fullér, R. (2000). *Introduction to neuro-fuzzy systems*. Advances in Soft Computing, Springer-Verlag.

Garibaldi, J. M. (2005). "Fuzzy expert systems." *Do smart adaptive systems exist?* Studies in Fuzziness and Soft Computing, B. Gabrys, K. Leiviskä, and J. Strackeljan, eds., Springer-Verlag.

Gilbert, K. G. (1914). *The transportation of debris by running waters,*. U.S. Geological Survey Professional Paper 86.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Pub. Co.

Goodwin, P. (2004). "Analytical solutions for estimating effective discharge." *Journal of Hydraulic Engineering*, 130(8), 729–738.

Graf, W. H., and Suszka, L. (1987). "Sediment transport in steep channels." *Journal of Hydroscience and Hydraulic Engineering*, 5(1), 11–26.

Guillaume, S. (2001). "Designing fuzzy inference systems from data: an interpretability-oriented review." *IEEE Transactions on Fuzzy Systems*, 9(3), 426–443.

Gustafson, D. E., and Kessel, W. C. (1979). "Fuzzy clustering with a fuzzy covariance matrix." *Proc. IEEE Transactions on Fuzzy Systems*, 195–204.

Guy, H. P., Simons, D. B., and Richardson, E. V. (1966). *Summary of alluvial channel data from flume experiments, 1951-61*. U.S. Geological Survey Professional Paper 462-I.

Hammell II, R. J., and Sudkamp, T. (1998). "Learning fuzzy rules from data." *The application of information technologies (computer science) to mission systems*, California, USA, 1–10.

Hammouda, K., and Karray, F. (2000). "A comparative study of data clustering techniques." *Tools of Intelligent Systems Design. In Course Project, SYDE*, 625.

Haykin, S. S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.

HR Wallingford. (1990). *Sediment transport, the Ackers and White theory revised*. Report SR237, England.

Hsu, S. M., and Holly, F. M. (1992). "Conceptual bed-load transport model and verification for sediment mixtures." *Journal of Hydraulic Engineering*, 118(8), 1135–1152.

Huang, J. V., and Greimann, B. P. (2007). *User's manual for SRH-1DV2.0*. Sedimentation and River Hydraulics Group, Technical Service Center, Bureau of Reclamation, Denver, Colorado.

Hubbell, D. W., and Matejka, D. Q. (1959). *Investigations of sediment transportation, Middle Loup River at Dunning, Nebraska.* U.S. Geological Survey Water-Supply Paper 1476.

Hüllermeier, E. (2005). "Fuzzy methods in machine learning and data mining: Status and prospects." *Fuzzy Sets and Systems*, 156(3), 387–406.

Hundecha, Y., Bardossy, A., and Theisen, H. W. (2001). "Development of a fuzzy logic-based rainfall-runoff model." *Hydrological Sciences Journal*, 46(3), 363–376.

Ishibuchi, H., Nozaki, K., Yamamoto, N., and Tanaka, H. (1994). "Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms." *Fuzzy Sets and Systems*, 65(2-3), 237–253.

Jang, J.-S. R. (1993). "ANFIS: Adaptive-network-based fuzzy inference system." *IEEE Transactions on Systems, Man and Cybernetics*, 23(3), 665–685.

Jang, J.-S. R., and Sun, C.-T. (1995). "Neuro-fuzzy modeling and control." *Proceedings of the IEEE*, 83(3), 378–406.

Jang, J.-S. R., and Gulley, N. (1998). *Fuzzy logic toolbox: for use with MATLAB: user's guide*. MathWorks, Inc.

Jang, J.-S. R., Sun, C.-T., and Mizutani, E. (1997). *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice Hall.

Jantzen, J. (1998). *Neurofuzzy modelling*. Technical Report 98-H-874, Technical University of Denmark, Department of Automation.

Jolliffe, I. T. (1986). *Principal component analysis*. Springer-Verlag.

Jorde, K. (1996). "Ecological evaluation of instream flow regulations based on temporal and spatial variability of bottom shear stress and hydraulic habitat Quality." *Proceedings of the Second International Symposium on Habitat Hydraulics*, Quebec, 163–174.

Julien, P. Y. (1995). *Erosion and Sedimentation*. Cambridge University Press.

Karim, M. F., and Kennedy, J. F. (1990). "Menu of coupled velocity and sediment - discharge relations for rivers." *Journal of Hydraulic Engineering*, 16(8), 978-996.

Kennedy, J. F. (1961). *Stationary waves and antidunes in alluvial channels*. Water Resources Number KH-R-2, California Institute of Technology, Pasadena.

Kisi, Ö. (2004). "Daily suspended sediment modelling using a fuzzy differential evolution approach." *Hydrological Sciences Journal*, 49(1), 183–197.

Kisi, Ö. (2005). "Suspended sediment estimation using neuro-fuzzy and neural network approaches." *Hydrological Sciences Journal*, 50(4), 683–696.

Kisi, Ö. (2008). "Constructing neural network sediment estimation models using a data-driven algorithm." *Mathematics and Computers in Simulation*, 79(1), 94–103.

Kisi, Ö., Haktanir, T., Ardiclioglu, M., Ozturk, O., Yalcin, E., and Uludag, S. (2009). "Adaptive neuro-fuzzy computing technique for suspended sediment estimation." *Advances in Engineering Software*, 40(6), 438–444.

Kisi, Ö., Karahan, M. E., and Şen, Z. (2006). "River suspended sediment modelling using a fuzzy logic approach." *Hydrological Processes*, 20(20), 4351–4362.

Kisi, Ö., Yuksel, I., and Dogan, E. (2008). "Modelling daily suspended sediment of rivers in Turkey using several data-driven techniques." *Hydrological Sciences Journal*, 53(6), 1270–1285.

Klir, G. J., Clair, U. H. S., and Yuan, B. (1997). *Fuzzy set theory: foundations and applications*. Prentice Hall.

Lai, Y. G. (2008). *SRH-2D version 2: Theory and user's manual*. Sedimentation and River Hydraulics Group, Technical Service Center, Bureau of Reclamation, Denver, Colorado.

Laursen, E. M. (1958). "The total sediment load of streams." *Journal of the Hydraulics Division, ASCE*, 84(1), 1531–1536.

Lohani, A. K., Goel, N. K., and Bhatia, K. K. S. (2006). "Takagi-Sugeno fuzzy inference system for modeling stage-discharge relationship." *Journal of Hydrology*, 331(1-2), 146–160.

Lohani, A. K., Goel, N. K., and Bhatia, K. K. S. (2007). "Deriving stage–discharge–sediment concentration relationships using fuzzy logic." *Hydrological Sciences Journal*, 52(4), 793–807.

Mahabir, C., Hicks, F. E., and Fayek, A. R. (2003). "Application of fuzzy logic to forecast seasonal runoff." *Hydrological Processes*, 17(18), 3749–3762.

Mamdani, E. H. (1977). "Application of fuzzy logic to approximate reasoning using linguistic synthesis." *IEEE Transactions on Computers*, 100(12), 1182–1191.

Meyer-Peter, E., and Müller, R. (1948). "Formulas for bed-load transport." *Proceedings of the 2nd Meeting of the International Association for Hydraulic Structures Research, Stockholm*, 39–64.

Mikut, R., Jäkel, J., and Gröll, L. (2005). "Interpretability issues in data-based learning of fuzzy systems." *Fuzzy Sets and Systems*, 150(2), 179–197.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Mitra, B., Scott, H. D., Dixon, J. C., and McKimmey, J. M. (1998). "Applications of fuzzy logic to the prediction of soil erosion in a large watershed." *Geoderma*, 86(3-4), 183–209.

Mohan, S., and Prasad, M. A. (2006). "Fuzzy logic model for multi reservoir operation." *Proceedings of the 2nd Conference of Asia Oceanic Sciences Society*, Singapore, 20–24.

Molinas, A., and Wu, B. (1998). "Effect of size dradation on transport of sediment mixtures." *Journal of Hydraulic Engineering*, 124(8), 786–793.

Molinas, A., and Wu, B. (2000). "Comparison of fractional bed-material load computation methods in sand-bed channels." *Earth Surface Processes and Landforms*, 25(10), 1045–1068.

Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S. (2004). "A neuro-fuzzy computing technique for modeling hydrological time series." *Journal of Hydrology*, 291(1-2), 52–66.

Nomicos, G. N. (1956). "Effects of sediment load on the velocity field and friction factor of turbulent flow in an open channel." PhD Thesis, California Institute of Technology, Pasadena, California.

Nordin, C. F. (1976). *Flume studies with fine and coarse sands*. U.S. Geological Survey Open-File Report 76-762.

Nordin, C. F., Jr. (1964). *Aspects of flow resistance and sediment transport, Rio Grande River near Bernalillo, New Mexico*. U.S. Geological Survey Water Supply Paper 1498.

Oinam, B. (2011). "Integrated fuzzy-GIS approach for assessing regional soil erosion risks." PhD Thesis, Institute of Hydraulic Engineering, University of Stuttgart.

Parker, G. (1990). "Surface-based bedload transport relation for gravel rivers." *Journal of Hydraulic Research*, 28(4), 501–518.

Proffitt, G. T., and Sutherland, A. J. (1983). "Transport of non-uniform sediments." *Journal of Hydraulic Research*, 21(1), 33–43.

Rajaee, T., Mirbagheri, S. A., Zounemat-Kermani, M., and Nourani, V. (2009). "Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models." *Science of The Total Environment*, 407(17), 4916–4927.

van Rijn, L. C. (1984a). "Sediment transport, part I: bed load transport." *Journal of Hydraulic Engineering*, 110(10), 1431–1456.

van Rijn, L. C. (1984b). "Sediment transport, Part II: suspended load transport." *Journal of Hydraulic Engineering*, 110(11), 1613–1641.

van Rijn, L. C. (1993). *Principles of sediment transport in rivers, estuaries, and coastal seas*. Aqua Publications.

Rooseboom, A., and Le Grange, A. (2000). "The hydraulic resistance of sand streambeds under steady flow conditions." *Journal of Hydraulic Research*, 38(1), 27–35.

Ross, T. J. (1995). *Fuzzy logic with engineering applications*. John Wiley and Sons.

Rottner, J. (1959). "A formula for bed-load transportation." *La Houille Blanche*, 14(3), 285–307.

Russell, S. O., and Campbell, P. F. (1996). "Reservoir operating rules with fuzzy programming." *Journal of Water Resources Planning and Management*, 122(3), 165–170.

Samaga, B. R., Raju, R., Kuttur, G., and Garde, R. J. (1986). "Bed load transport of sediment mixtures." *Journal of Hydraulic Engineering*, 112(11), 1003–1018.

Sato, S., Kikkawa, H., and Ashida, K. (1958). "Research on the bed load transportation." *Journal of Research, Public Works Research Institute*, 3, 21.

Schneider, M. (2001). "Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen." PhD Thesis, Institute of Hydraulic Engineering, University of Stuttgart.

Shen, H. W., and Hung, C. S. (1972). "An engineering approach to total bed-material load by regression analysis." *Proceedings of the Sedimentation Symposium*, Fort Collins, Colorado, 1–17.

Shields, A. (1936). *Application of similarity principles and turbulence research to bed-load movement*. California Institute of Technology, Pasadena.

Shields, F. D., Jr. (2007). *Scour calculations*. Technical Supplement 14B in Stream Restoration Design, National Engineering Handbook Part 654, USDA-NRCS Washington, D. C.

Silvert, W. (1997). "Ecological impact classification with fuzzy sets." *Ecological Modelling*, 96(1-3), 1–10.

Solomatine, D. P., and Ostfeld, A. (2008). "Data-driven modelling: some past experiences and new approaches." *Journal of Hydroinformatics*, 10(1), 3–22.

Song, T., Chiew, Y. M., and Chin, C. O. (1998). "Effect of bed-load movement on flow friction factor." *Journal of Hydraulic Engineering*, 124(2), 165–175.

Spreafico, M., and Lehmann, C. (2009). *Erosion, transport and deposition of sediment - case study Rhine*. Contribution to the International Sediment Initiative of UNESCO/IHP, Report NumberII-20.

Stein, R. A. (1965). "Laboratory studies of total load and apparent bed load." *Journal of Geophysical Research*, 70(8), 1831–1842.

Sugeno, M., and Yasukawa, T. (1993). "A fuzzy-logic-based approach to qualitative modeling." *IEEE Transactions on Fuzzy Systems*, 1(1), 7–31.

Syvitski, J. P., Morehead, M. D., Bahr, D. B., and Mulder, T. (2000). "Estimating fluvial sediment transport: The rating parameters." *Water Resources Research*, 36(9), 2747–2760.

Takagi, T., and Sugeno, M. (1985). "Fuzzy identification of systems and its applications to modeling and control." *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1), 116–132.

Tayfur, G., and Singh, V. P. (2006). "ANN and fuzzy logic models for simulating event-based rainfall-runoff." *Journal of Hydraulic Engineering*, 132(12), 1321–1330.

Tayfur, G., Ozdemir, S., and Singh, V. P. (2003). "Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces." *Advances in Water Resources*, 26(12), 1249–1256.

Toffaleti, F. B. (1969). "Definitive computations of sand discharge in rivers." *Journal of the Hydraulics Division, ASCE*, 95(1), 225–248.

Tran, L. T., Ridgley, M. A., Duckstein, L., and Sutherland, R. (2002). "Application of fuzzy logic-based modeling to improve the performance of the Revised Universal Soil Loss Equation." *CATENA*, 47(3), 203–226.

U.S. Army Corps of Engineers. (1993). *Scour and deposition in rivers and reservoirs: HEC-6user's manual*. Hydrologic Engineering Center, Davis, CA.

Vanoni, V. A., and Brooks, N. H. (1957). *Laboratory studies of the roughness and suspended load of alluvial streams*. MRD Sediment Series 11, California Institute of Technology, Pasadena.

Vernieuwe, H., Georgieva, O., De Baets, B., Pauwels, V. R. N., Verhoest, N. E. C., and De Troch, F. P. (2005). "Comparison of data-driven Takagi-Sugeno models of rainfall-discharge dynamics." *Journal of Hydrology*, 302(1-4), 173–186.

Vetter, M. (1989). *Total sediment transport in open channels*. Institute of Hydrology, University of the German Federal Army, Munich, Germany.

Wang, L. X., and Mendel, J. M. (1992). "Generating fuzzy rules by learning from examples." *IEEE Transactions on Systems, Man, and Cybernetics*, 22(6), 1414–1427.

Wang, Y., Ren, M., and Syvitski, J. (1998). *Sediment transport and terrigenous fluxes*. The Global Coastal Ocean: Processes and Methods, 253–292.

White, W. R., Milli, H., and Crabbe, A. D. (1975). "Sediment transport theories: a review." *Proceedings of the Institute of Civil Engineers*, London, 265–292.

Williams, G. P. (1967). *Flume experiments on the transport of a coarse sand*. U.S. Geological Survey Professional Paper 562-B.

Wu, B., van Maren, D. S., and Li, L. (2008). "Predictability of sediment transport in the Yellow River using selected transport formulas." *International Journal of Sediment Research*, 23(4), 283–298.

Wu, B., Molinas, A., and Julien, P. Y. (2004). "Bed-material load computations for nonuniform sediments." *Journal of Hydraulic Engineering*, 130(10), 1002–1012.

Wu, W., Wang, S. S. Y., and Jia, Y. (1998). "Incipient motion and bed load transport for nonuniform sediment mixtures." *Proceedings of the 3rd International Conference on Hydroscience and Engineering*, Germany.

Wu, W., Wang, S. S. Y., and Jia, Y. (2000). "Nonuniform sediment transport in alluvial rivers." *Journal of Hydraulic Research*, 38(6), 427-434.

Xiaoqing, Y. (2003). *Manual on sediment management and measurement*. Operational Hydrology Report, World Meteorological Organization, Geneva.

Xiong, L., Shamseldin, A. Y., and O'Connor, K. M. (2001). "A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system." *Journal of Hydrology*, 245(1-4), 196–217.

Yager, R., and Filev, D. (1994). "Generation of fuzzy rules by mountain clustering." *Journal of Intelligent and Fuzzy Systems*, 2(3), 209–219.

Yalin, M. S. (1977). *Mechanics of sediment transport*. Pergamon Press.

Yang, C. T. (1973). "Incipient motion and sediment transport." *Journal of the Hydraulics Division, ASCE*, 99(10), 1679–1704.

Yang, C. T. (1979). "Unit stream power equations for total load." *Journal of Hydrology*, 40(1-2), 123–138.

Yang, C. T. (1984). "Unit stream power equation for gravel." *Journal of Hydraulic Engineering*, 110(12), 1783–1797.

Yang, C. T. (1996). *Sediment transport: theory and practice*. McGraw-Hill, New York.

Yang, C. T. (2006). "Chapter 3: Non-cohesive sediment transport." *USBR Erosion and Sedimentation Manual*, 3-1–3-111.

Yang, C. T, and Huang, C. (2001). "Applicability of sediment transport formulas." *International Journal of Sediment Research*, 16(3), 335–353.

Yang, C. T, and Molinas, A. (1982). "Sediment transport and unit stream power function." *Journal of the Hydraulics Division, ASCE*, 108(6), 774–793.

Yang, C. T, and Wan, S. (1991). "Comparisons of selected bed-material load formulas." *Journal of Hydraulic Engineering*, 117(8), 973–989.

Yang, C. T, Molinas, A., and Wu, B. (1996). "Sediment transport in the Yellow river." *Journal of Hydraulic Engineering*, 122(5), 237–244.

Yang, C. T., and Simoes, F. J. M. (2002). *User's manual for GSTARS3 (generalized sediment transport model for alluvial river simulation version 3.0)*. Technical Service Center, U.S. Bureau of Reclamation, Denver, Colorado.

Zadeh, L. A. (1965). "Fuzzy sets." *Information and Control*, 8(3), 338–353.

Zounemat-Kermani, M., and Teshnehlab, M. (2008). "Using adaptive neuro-fuzzy inference system for hydrological time series prediction." *Applied Soft Computing*, 8(2), 928–936.

Zounemat-Kermani, M., Beheshti, A.-A., Ataie-Ashtiani, B., and Sabbagh-Yazdi, S.-R. (2009). "Estimation of current-induced scour depth around pile groups using neural network and adaptive neuro-fuzzy inference system." *Applied Soft Computing*, 9(2), 746–755.

Appendices

Appendix A: The number of training and test datasets from each station and the total number of final bed load and total bed-material load transport datasets used for the Rhine and Elbe rivers.

Table A.1: Sediment measuring stations and the number of training and test datasets from each station for the Rhine river.

Bed Load			Total Load		
Station	Training	Test	Station	Training	Test
Plittersdorf (alt)	8	4	Plittersdorf (alt)	7	3
Plittersdorf (akt.)	13	6	Plittersdorf (akt.)	7	3
Illingen	16	7	Illingen	11	5
Neuburgweier	19	9	Neuburgweier	14	6
Maxau	16	8	Maxau	12	5
Leimersheim	13	6	Leimersheim	14	6
Philippsburg	3	2	Philippsburg	5	2
Speyer	12	5	Speyer	8	4
Mannheim	8	4	Mannheim	9	4
Worms	12	6	Worms	11	5
Gernsheim	5	2	Gernsheim	5	2
Nierstein	3	1	Nierstein	14	7
Oestrich	2	1	Weisenau 1	4	1
Bingen	2	1	Weisenau 2	1	1
Oberwesel	3	1	Mainz	10	4
Braubach	4	2	Niederwalluf	11	5
Neuwied	10	4	Oestrich	9	4
Brohl	6	2	Bingen	12	6
Koenigswinter	6	2	Braubach	7	3
Bonn	2	1	Neuwied	8	3
Porz	9	4	Brohl	6	2
Rheindorf	14	6	Koenigswinter	7	3
Urdenbach	9	4	Mondorf	2	1
Uedesheim	12	5	Porz	6	2
Düsseldorf V	7	3	Rheindorf	11	5
Buederich	3	1	Urdenbach	6	2

Kaiserwerth	10	5
Gellep	10	5
Krefeld	17	8
Duisburg	12	6
Orsoy	18	8
Goetterswickerhamm	4	2
Ork	11	5
Wesel I	10	4
WeselPerrich	22	10
Xanten	9	4
Rees	14	6
Grieth	10	4
Emmerich 1	4	2
Emmerich 2	2	1
Emmerich 4	2	1
Emmerich 6	3	1
Emmericher Ward	2	1
Griethausen	9	4
Total	386	174
Percentage (%)	68.93	41.07

Uedesheim	11	5
Düsseldorf V	6	3
Buederich	5	2
Kaiserwerth	6	2
Gellep	6	3
Krefeld	14	6
Duisburg	12	6
Orsoy	10	4
Goetterswickerhamm	10	5
Ork	10	5
Wesel I	4	2
WeselPerrich	12	6
Xanten	4	1
Rees	12	5
Grieth	6	2
Griethausen	10	4
Total	355	155
Percentage (%)	69.6	30.4

Table A.2: Sediment measuring stations and the number of training and test datasets from each station for the Elbe river.

Bed Load			Total Load		
Station	Training	Test	Station	Training	Test
Torgau	3	1	Dresden	6	2
Mockritz	8	3	Mühlberg	12	6
Pretsch/Mauken	10	4	Torgau	12	5
Wittenberg	12	6	Mockritz	10	5
Vockerode	8	3	Pretsch/Mauken	16	8
Dessau 259	5	2	Wittenberg	21	10
Dessau 260	6	3	Vockerode	8	3
Aken	17	8	Aken	21	10
Barby	20	9	Barby	13	6
Magdeburg_321	15	7	Magdeburg_321	4	2
Magdeburg_324	16	7	Magdeburg_324	7	2
Magdeburg_332	22	11	Magdeburg_332	14	7
Magdeburg_338	5	2	Niegripp	14	6
Niegripp	14	7	Tangermünde	5	2
Tangermünde	10	5	Sandau	4	2
Sandau	8	3	Wittenberge	11	5
Wittenberge	13	7	Langendorf	11	5
Langendorf	15	7	Wilkenstorf	17	8
Wilkenstorf	18	9	Neu Darchau	14	7
Neu Darchau	18	9			
Total	243	113	Total	220	101
Percentage of total data	68.3	31.7	Percentage of total data	68.5	31.5

Appendix B: ANFIS model results for laboratory datasets in the sand size range using four clusters and five combinations of input variables.

Table B.1: Performances of various ANFIS models with four clusters for laboratory data in sand size range.

Input variables	Training data				Test data			
	AARE	RMSE	r	Percent of ARE>100%	AARE	RMSE	r	Percent of ARE>100%
D V S d_{50}	33.3	2881.0	0.91	4.3	34.1	2669.9	0.94	4.4
V S d_{50}	41.4	2692.9	0.92	3.4	41.8	2504.8	0.94	3.2
D S d_{50}	41.7	2641.7	0.92	8.0	41.4	2487.0	0.94	8.4
D V d_{50}	51.7	4235.4	0.80	11.5	54.6	4354.5	0.83	12.2
D V S	39.0	3475.8	0.87	8.1	40.2	3283.7	0.90	7.3

Table B.2: Discrepancy ratio analysis of various ANFIS models with four clusters for laboratory data in the sand size range.

Input variables	Training data				Test data			
	Avg.Dr	Percent of data Dr with in			Avg.Dr	Percent of data Dr within		
		0.75-1.25	0.5-1.5	0.25-1.75		0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.13	51.0	71.4	82.6	1.13	50.6	71.8	83.7
V S d_{50}	1.17	54.1	76.9	86.2	1.17	54.9	77.0	87.5
D S d_{50}	1.21	53.2	76.0	84.8	1.20	53.8	75.9	84.3
D V d_{50}	1.28	39.5	61.9	74.8	1.24	39.8	61.9	74.1
D V S	1.21	48.9	67.8	77.8	1.19	46.5	62.5	75

Appendix C: ANFIS model results for laboratory datasets in the sand size range using five clusters and five combinations of input variables.

Table C.1: Performances of various ANFIS models with five clusters for laboratory data in sand size range.

Input variables	Training data				Test data			
	AARE	RMSE	r	Percent of ARE>100%	AARE	RMSE	r	Percent of ARE>100%
D V S d_{50}	34.5	2775.5	0.92	4.3	31.9	2572.8	0.94	4.7
V S d_{50}	25.12	2091.4	0.95	1.5	27.2	2311.34	0.95	2.0
D S d_{50}	47.2	2863.9	0.91	10.3	56.8	2737.8	0.93	11.0
D V d_{50}	55.8	4136.9	0.80	13.0	54.1	4237.6	0.83	12.8
D V S	45.0	3264.5	0.88	10.3	45.2	3205.9	0.9	9.6

Table C.2: Discrepancy ratio analysis of various ANFIS models with five clusters for laboratory data in the sand size range.

Input variables	Training data				Test data			
	Avg.Dr	Percent of data Dr with in			Avg.Dr	Percent of data Dr within		
		0.75-1.25	0.5-1.5	0.25-1.75		0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.12	48.3	70.3	83.8	1.12	51.2	73.3	82.9
V S d_{50}	1.09	61.6	82.0	89.3	1.10	59.6	79.7	88.1
D S d_{50}	1.31	48.6	68.9	77.9	1.38	47.7	65.7	74.1
D V d_{50}	1.33	39.6	61.6	76.6	1.27	40.4	60.8	71.8
D V S	1.25	51.7	70.0	79.7	1.25	50.0	71.2	81.1

Appendix D: ANFIS model results for laboratory datasets in the gravel size range using four clusters and five combinations of input variables.

Table D.1: Performances of various ANFIS models with four clusters for laboratory data in gravel size range.

Input variables	Training data				Test data			
	AARE	RMSE	r	Percent of ARE>100%	AARE	RMSE	r	Percent of ARE>100%
D V S d_{50}	56.0	606.0	0.97	12.2	48.5	496.9	0.98	14.0
V S d_{50}	35.5	525.6	0.97	4.6	37.6	414.5	0.99	7.0
D S d_{50}	37.1	507.3	0.98	7.9	39.0	368.7	0.99	10.9
D V d_{50}	62.7	794.8	0.94	15.2	56.9	821.8	0.96	12.4
D V S	84.5	898. 0	0.92	20.5	96.9	896.2	0.95	21.7

Table D.2: Discrepancy ratio analysis of various ANFIS models with four clusters for laboratory data in the gravel size range.

Input variables	Training data				Test data			
		Percent of data Dr with in				Percent of data Dr within		
	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.31	39.1	55.9	73.6	1.3	47.3	65.9	75.2
V S d_{50}	1.16	51.3	73.0	86.3	1.2	48.8	74.4	85.3
D S d_{50}	1.16	54.0	71.1	83.7	1.18	54.3	68.2	79.1
D V d_{50}	1.35	34.6	57.4	70.7	1.27	41.1	59.7	75.2
D V S	1.51	35.5	57.4	69.6	1.69	33.3	56.6	72.1

Appendix E: ANFIS Model results with five different combinations of input variables and three fuzzy sets for each input variable for laboratory datasets in the gravel size range.

Table E.1: Performances of various ANFIS models with three fuzzy partitions for each input variable for laboratory data in the gravel range.

Input variables	Training data				Test data			
		Percent of data Dr with in				Percent of data Dr within		
	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75	Avg.Dr	0.75-1.25	0.5-1.5	0.25-1.75
D V S d_{50}	1.08	69.6	87.5	93.5	1.08	65.9	83.7	89.9
V S d_{50}	1.15	59.3	79.5	89.4	1.24	54.3	76.7	89.2
D S d_{50}	1.12	66.5	83.3	90.9	1.11	65.1	80.6	89.2
D V d_{50}	1.40	36.1	53.2	65.8	1.59	38.0	52.7	65.12
D V S	1.45	41.1	67.7	77.2	1.57	31.0	63.6	76.7

Table E.2: Discrepancy ratio analysis of various ANFIS models with three fuzzy partitions for each input variable for laboratory data in the gravel range.

Input variables	Training data				Test data			
	AARE	RMSE	r	Percent of ARE>100%	AARE	RMSE	r	Percent of ARE>100%
D V S d_{50}	22.3	3908	0.99	4.0	24.8	326.8	0.99	1.5
V S d_{50}	32.6	465.1	0.98	4.9	40.2	383.9	0.99	6.2
D S d_{50}	27.3	461.5	0.98	6.1	29.8	367.1	0.99	6.2
D V d_{50}	67.5	896.5	0.93	17.9	87.0	1166.9	0.91	18.6
D V S	71.9	806.8	0.94	18.3	81.3	780.6	0.96	17.1

Appendix F: Statistical analysis of daily values of hydraulic input variables and their combinations for all thirteen stations on the Upper Rhine used for model validation.

Table F.1: Statistical distributions of the daily values of input hydraulic variables and their combinations for the stations on the Upper Rhine.

Parameters	Mean	Maximum	Minimum	St. dev
Mean velocity (m/s)	1.67	3.05	0.84	0.28
Depth (m)	3.87	8.76	1.61	0.94
Energy slope (m/m)	0.00030	0.00064	0.00010	0.00007
Width (m)	285.82	900.83	136.90	115.14
Unit stream power (m/s)	0.00051	0.00126	0.00009	0.00017
Discharge per unit width (m^3/s.m)	6.61	24.15	2.52	2.58
Depth times slope (m)	0.001141	0.003171	0.000368	0.000343
Discharge (m^3/s)	1263.46	4511.20	441.00	520.84

Curriculum Vitae

PERSONAL DETAILS

Name:	Habtamu Gezahegn Tolossa
Date of Birth:	March 29, 1982
Nationality:	Ethiopian
Sex:	Male

EDUCATION

02/2009 – 02/2012	Ph.D. - Civil and Environmental Engineering University of Stuttgart, Germany Ph.D. Research Title: "Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach"
10/2006 – 10/2008	Master of Science (M.Sc.) - Water Resource Engineering & Management University of Stuttgart, Germany Master's Thesis Title: "Comparsion of 2D Hydrodynamic Models in River Reaches of Ecological Importance: Hydro_AS-2D and SRH-2D"
09/2000 – 07/2005	Bachelor of Science (B.Sc.) - Hydraulic Engineering Arba Minch University, Ethiopia B.Sc. Thesis Title: "Hydropower Potential Studies in Gilgel-Abay and Koga Rivers"

PROFESSIONAL EXPERIENCE

02/2009 – 02/2012	**Ph.D. Researcher** Institute for Modelling Hydraulic and Environmental Systems University of Stuttgart, Germany
05/2010 – 03/2012	**Civil Engineer – part-time** sje Schneider and Jorde Ecological Engineering GmbH Stuttgart, Germany
08/2005 – 12/2006	**Assistant Lecturer** Department of Hydraulic Engineering, Arba Minch University Arba Minch, Ethiopia

COMPUTER SKILLS

Hydraulic Modelling:	SMS (Surface Water Modelling System), Hydro_As-2D, SRH-2D, HEC-RAS, EPANET, CROPWAT

Morphodynamic Modelling:	SRH-2D, GSTARS3, MIKE 11, HEC-6
Habitat Modelling:	CASiMiR, HABITAT
Surface Mapping:	Surfer8, contouring and 3D surface mapping program
GIS:	ArcView, ArcGIS
MS Office:	Word, Excel, Access, PowerPoint
Programming Language:	MATLAB, FORTRAN
Document & typesetting:	LaTeX, JabRef, Zotero

LANGUAGE SKILLS

English	Fluent spoken and written
German	Good knowledge
Amharic	Native language, Ethiopia

PUBLICATIONS

Tolossa, H. G., Wieprecht, S., and Schneider, M. (2012). Evaluation of sediment transport equations and ANFIS-based models for computing total in the Elbe river. *Proceedings of the Second European IAHR Congress, Munich, Germany* (accepted).

Wieprecht, S., Tolossa, H. G., and Yang, C. T. (2011). A neuro-fuzzy based modelling approach for sediment transport computation. *Hydrological Sciences Journal* (in review).

Tolossa, H. G., Wieprecht, S., and Schneider, M. (2010). Application of data-driven fuzzy logic approach for modelling sediment transport. *Proceedings of the First European IAHR Congress, Edinburgh, Scotland.*

Tolossa, H. G., Tuhtan, J., Schneider, M., and Wieprecht, S. (2009). Comparison of 2D hydrodynamic models in river reaches of ecological importance: Hydro_AS-2D and SRH-W. *Proceedings of the 33rd IAHR Congress Water Engineering for a Sustainable Environment, Vancouver, Canada.*

AWARDS

- 02/2009 – 02/2012: IPSWaT scholarship from the German Federal Ministry of Education and Research (BMBF) for Ph.D. studies at the international doctoral program environment and water (ENWAT) at the University of Stuttgart.

- 01/2009 – 10/2008: IPSWaT scholarship from the German Federal Ministry of Education and Research (BMBF) to study M.Sc. in Water Resources Engineering and Management (WAREM) at the University of Stuttgart.

- 07/2005: Graduation Award from Arba Minch University for graduating with great distinction ranking first from all students in the water technology institute of Arba Minch University.

PROFESSIONAL MEMBERSHIP

Since 2007	International Association of Hydro-Environment Engineering and Research (IAHR)
Since 2008	German Association for Water, Waste-Water and Waste (DWA)

Institut für Wasser- und Umweltsystemmodellierung Universität Stuttgart

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

Verzeichnis der Mitteilungshefte

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, vergriffen

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002, ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: 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

208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam,* 2011, ISBN 978-3-942036-12-2

209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone,* 2012, ISBN 978-3-942036-13-9

210 Tuthan, Jeff: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking,* 2012, ISBN 978-3-942036-14-6

211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach,* 2012, ISBN 978-3-942036-15-3

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