UNESCO-IHE
INSTITUTE FOR WATER EDUCATION

A Hybrid Approach to Combine Physically Based and Data-Driven Models in Simulating Sediment Transportation

Sewagudde Sowed

MSc Thesis 1st Draft
March 2008

UNESCO-IHE
Institute for Water Education

UNESCO-IHE
Institute for Water Education

A Hybrid Approach to Combine Physically Based and Data-Driven Models in Simulating Sediment Transportation

Master of Science Proposal
by
Sowedi Sewagudde

Supervisor
Dr. B. Bhattacharya (UNESCO-IHE)

Examination committee

Prof. Dr. R.K. Price (UNESCO-IHE), Chairman
Dr. T. Van Kessel (Deltares)
Dr. B. Bhattacharya (UNESCO-IHE)

This research is done in partial fulfillment of requirements for the Master of Science degree at the UNESCO-IHE Institute for Water Education, Delft, The Netherlands

Delft
April 2008

To my mother

Abstract

Sediments are a problem in many aquatic systems. Their transportation and deposition has significant implication on morphology of water bodies, navigability in channels, and water quality. Understanding the dynamics of sediment transportation in time and space is therefore important in drawing interventions and making management decisions. In this thesis sediment resuspension, transportation and deposition is modelled by combining physically based and data-driven modelling approaches. This modelling approach is denoted as hybrid modelling.

An investigation on the possibility of approximating a physically based numerical model using a neural network data-driven model (DDM) is carried out. The study also examines the effect of specifying time varying sediment open boundary conditions instead of fixed boundary conditions in simulating suspended sediments along the Dutch coast. It further explores the applicability of knowledge gained in modelling sediment transportation along the Dutch coast to Lake Victoria, East Africa.

Results show that there is a strong possibility of approximating a numerical model using a data driven model. A combination of total bed shear stress and significant wave heights as input variables to the DDM produces satisfactory results. The study reveals that while applying a DDM (developed at one location) to another location, alongshore and cross-shore correction factors should be applied. It further reveals that errors in simulation results due to inaccuracies in sediment boundary conditions reduce exponentially inside the model domain.

The nature of the sediment open boundary (whether fixed or time varying) has minimal effect on the model outcome when a short simulation period is considered. Time varying sediment boundary conditions could be more suitable where simulation of large time-scale processes such as morphological changes is involved. Regarding Lake Victoria, application of hybrid modelling knowledge of the Dutch coast is most relevant when applied to the western coast of lake Victoria. Simulation accuracy of a point DDM is likely to be higher in Lake Victoria than in the Dutch coast because of the fewer complicated processes in the former system.

Keywords: suspended sediment, hybrid model, physically based model, data-driven model artificial neural networks.

Acknowledgements

I am deeply indebted to the Water Resources Planning and Management Project of the Nile Basin Initiative (NBI) who funded my MSc studies at UNESCO-IHE Institute for Water Education.

Sincere gratitude go to the stuff of UNESCO-IHE in particular Prof. S. Solomatine, Dr. I Popescu, Prof. A. Mynett, Emeritus Prof. R. Price, Dr. A Jonoski for the valuable guidance that enabled us to carry out this pre-feasibility study successfully.

Many thanks to hydroinformatics class of 2006/08 with whom we interacted during the bitter sweet times of the course.

Special thanks to Dr. B. Bhattacharya whose thoughtful guidance and stimulating discussions stirred this research to a successful completion.

I am grateful to Dr. Thijs van Kessel , Dr. Johan de Kok and Dr. Johannes Smits of WL| Delft hydraulics and Prof. Roevlink and Mick van Der Wegen for their warm hearted technical advice and support.

Finally I wish to thank my family for your endless support and love.

Table of Contents

1 Introduction

1.1 Background

Sedimentation is a problematic feature in many aquatic systems. Sediments impact on the use of water in many ways. Suspended sediment is crucial in determining the local morphology in coastal rivers, lakes, estuaries, and shelves environments. It always raises concern among coastal engineers and water managers about coastal protection, land reclamation, dredging of deepwater navigational channels and water quality management. Fluid mud, a high concentration aqueous suspension of fine-grained sediment (cohesive sediment) in which settling is substantially hindered also constitutes a significant management problem (McAnally, 2007). It impedes navigation, reduces water quality and damages equipment. Fluid mud accumulations have been reported in numerous locations worldwide, including the Maasmond area at the entrance of the port of Rotterdam, Savannah Harbor, U.S., and the Severn Estuary, U.K.(McAnally, 2007; Winterwerp, 1999).

Suspended solids reduce the available light to aquatic plants, clog the filtering capacity of filter feeders, cover fish spawning areas and food supplies, and clog and harm gills of fish. Turbidity lowers water transparency there by interrupting the feeding habits of fish and even primary production. Sediment deposition and resuspension also has considerable influence on eutrophication of lakes, streams, estuaries, and coastal waters. Chemicals favouring eutrophication conditions such as phosphorus and ammonia are transported with sediment in an adsorbed state. These effects combine to reduce the fish and plant populations as well as the productivity of the aquatic system (Blom, 1992; Ji *et al.*, 2002)

The process of sediment resuspension and transport in large water bodies particularly shallow lakes mainly starts with wind energy being delivered to the water surface and causing waves (Jin and Ji, 2004). The wind energy is transmitted from the surface to the bottom while being dissipated into wave motion in the vertical direction. This phenomenon creates orbital velocities at the sediment-water interface, which in combination with current velocities exert shear stresses leading to entrainment of sediment into the water column.

In coastal areas in addition to wind induced waves and currents, sediment resuspension/transport is also influenced by tide induced current. In areas where fresh river water is discharged into sea or ocean, density driven currents are created and they lead to a higher concentration of suspended sediments near the coast. A case in point is the Maasmond area along the Dutch where the river Rhine discharges into the North Sea (WL| Delft Hydraulics, 2001).

Models have been used to improve the understanding of the complex interaction between sediment, waves and currents in forecasting of siltation and suspended sediment. Physically based numerical models have been employed in the simulation of two dimension (2D) and three dimension (3D) sedimentation processes in aquatic systems (Ji *et al.*, 2002; Roelvink *et al.*, 2001; Vuurens, 2001; WL| Delft Hydraulics, 2001). The strength of numerical models lies in their ability to simulate physical processes by taking advantage of the deterministic characteristics embedded therein. This feature makes them quite handy in interpreting the evolution of physical processes of the studied system under different conditions. For example, it is possible to carry out worst or best case scenario assessments. In this way we can establish the boundaries within which an aquatic system can be managed.

Data-driven models have also been used in studying sedimentation processes (Bhattacharya *et al.*, 2007). These are built on the basis of collected data and they do not incorporate the underlying processes of the phenomenon. But this does not imply that the underlying processes are not always known. In most cases they are built when such knowledge is absent or disjointed. In many cases there is an understanding of the modelled processes, but not very detailed to facilitate development of accurate physically based models (Solomatine, 1999).

1.2 Problem description

Physically based and data driven models are used in simulating sediment transportation. However, there are several limitations with both approaches. One of the limitations of physically based models is that they require a lot computation time. This is largely caused by three factors namely: the long simulation time due to long residence time of fine sediments; large computation domain size and; high resolution grid as necessitated by a strong sediment concentration gradient and need to assess impacts of local morphological changes. To keep the computation time within reasonable limits usually meteorological variability is schematized over a short period and then assumed to occur in the same pattern over the rest of the simulation period. This is done at the expense of realism and consequently impacts on the accuracy of simulated results. Data-driven models require a lot of data for training to achieve satisfactory performance. Besides they cannot predict future changes in the environment that are not captured in the existing/training data e.g. changes in bathymetry. It is probable that more satisfactory sedimentation simulations can be achieved by combining the positive aspects of these two modelling approaches. This approach has been denoted as hybrid modelling.

1.3 Objectives

The objective of this study is to develop a methodology for hybrid modelling of sedimentation in a coastal basin or large shallow lake where physically based and data driven approaches are combined.

1.4 Research questions

1. To what extent can a physically based model be approximated by a data driven model. What are the most critical deterministic and stochastic variables for approximating a physically based model?

2. How can time varying boundary conditions be generated by a data driven model. What improvement is achieved by specifying time varying sediment boundary conditions instead of a fixed boundary? Is the physical description for sediment transport sufficient in the present physically based models? What improvements can be recommended based on the analysis of model results of suspended sediment along the Dutch coast?

3. How can a hybrid sediment modelling approach developed for the Dutch coast be applied to Lake Victoria? What similarities and differences would influence methodology/technological transfer?

1.5 Methodology

1.5.1 Approach

This research was broken down into three blocks. The first block explores the possibility of approximating a physically based model using a data driven model to predict suspended particulate matter (SPM) concentrations. It identifies the necessary input variables and data manipulation operations to expose maximum information to a data driven modelling tool.

The second block investigates the effect of specifying time varying open boundary condition instead of fixed boundary conditions on the simulation of (SPM) along the Dutch coast. First, a methodology for generating time varying open boundary conditions using a data driven model is developed. Local hydrodynamic and meteorological conditions are used as input variables. Secondly the improvement of the time varying open boundary conditions (OBC) on the simulation results is studied

The third block examines how knowledge gained in modelling sediment transport along the Dutch coast can be transfer to understand sediment transport in Lake Victoria. This part lays ground for more accurate and reliable modelling of sediment transport in Lake Victoria in future through transfer of methods developed in the preceding sections.

1.5.2 Modelling tools

Physically based modelling

The physically based modelling tool used in this study is Delft3D. It is a powerful software package developed by WL | Delft Hydraulics. It was selected based on the fact that it has been used extensively in the study of sedimentation processes in the Dutch coast and it is well documented.

Data-driven modelling

Data driven modelling is implemented using an Artificial Neuron Network (ANN) of a Multi layer perceptron topology. This preference is based on the fact that ANN are able to capture nonlinear dependencies that are inherent in the data as opposed to linear autoregressive data-driven models with exogenous inputs (Vojinovic *et al.*, 2003). It has also been used in previous modelling of sediment transport in the Dutch coast has produced quite satisfactory results (Bhattacharya *et al.*, 2007).

1.5.3 Outline of the report

The thesis is structured in such away that the first two chapters give general information of the study. Chapters 3, 4 and 5 address the three blocks of the study. They give the focus of the block, specific problem addressed, specific objectives and approach used. They have been written in such away that they are self sufficient in an attempt to reduce cross referencing with other chapters. Though this presents room for some repetition however it is intended to make the thesis easy to read. Chapter 6 presents a wrap up of the study. The outline of the chapters is as follows:

Chapter 1 introduces and justifies the need for employing hybrid modelling in the simulation of sediment transport. The objectives of the study are also introduced therein.

Chapter 2 reviews the basis for hybrid modelling, previous work done and the possible application of hybrid modelling to simulate suspended sediment dynamics. It also 3 discusses the schematisation of the physically based model. It highlights the nature of the computational grid, governing equations, boundary conditions and the transport processes implemented.

Chapter 3 presents an examination of the possibility of approximating a physically based model using a data driven model to predict suspended sediments.

Chapter 4 Impacts of specifying time varying open boundary condition as opposed to fixed boundary conditions on the simulation of suspended sediments is investigated. A methodology to generate time varying boundary conditions is discussed.

Chapter 5 explores how to apply hybrid expertise developed in the preceding chapters to Lake Victoria. The relevancy of hybrid modelling of sediment dynamics in Lake Victoria and the Nile basin in general is also demonstrated.

Chapter 6 presents conclusions drawn from the study and recommendations for further research.

2 Concepts of hybrid modelling concepts

2.1 Introduction

The previous chapter has emphasized the importance of understanding sediment transport dynamics. The tools that have been used to improve the understanding of the dynamics have been introduced and their strength and shortcomings highlighted. It established the need to combine these tools in a framework denoted as hybrid modelling. This chapter reviews the basis for hybrid modelling, previous work done and possible application of hybrid modelling in the context of suspended sediment modelling.

2.2 Sediment properties

Sediment refers to granular material that can settle in water by gravity. The size of the particles is one of the most important characteristics of sediment. The range of particle size of practical importance for sediment studies is quite diverse ranging from clay to breakaway armor stone blocks. Usually the particle size is defined in terms of its diameter which in turn is expressed as a fineness factor based on a phi unit as proposed by Krumbein (1936).

$$\varphi = -\log_2 D \tag{2.1}$$

Where D is the grain diameter in millimeters

Given the fact that even the best-sorted natural sediments have a range of grain sizes, normally there is a need to classify sediment as an aggregation of particles, rather than the diameter of a single particle. Thus sediment particles are classified based on their size, into six general categories: clay, silt, sand, gravel, cobbles, and boulders. Table 2.1 presents a classification of sediments according to American Geophysical Union These classifications are essentially arbitrary and as a result many grading systems are found in the engineering and geologic literature (van Rijn, 1993).

Cohesive sediments in a marine environment mainly consist of clay, silt, fine sand, organic material, water and sometimes gases(Ye, 2006). Its behavior is dominated by electrochemical forces and is primarily dependent on the particle size, water chemistry, and sediment mineralogy. The three most common minerals which have electrochemical forces causing individual particles to stick together are illite, kaolinite, and montmorillonite.

Table 2.1 Grain size model of American Geophysical Union (van Rijn, 1993)

Class Name	Millimeters	Phi Values
Boulders	>256	<-8
Cobbles	256-64	-8 to -6
Gravel	64-2	-6 to -1
Very Coarse sand	2.0-1.0	-1 to 0
Coarse sand	1.0-0.5	0 to +1
Medium sand	0.5-0.25	+1 to +2
Fine sand	0.25-0.125	+2 to +3
Very fine sand	0.125-0.062	+3 to +4
Coarse silt	0.062-0.031	+4 to +5
Medium silt	0.031-0.016	+5 to +6
Fine silt	0.016-0.008	+6 to +7
Very fine silt	0.008-0.004	+7 to +8
Coarse clay	0.004-0.002	+8 to +9
Medium clay	0.002-0.001	+9 to +10
Fine clay	0.001-0.0005	+10 to +11
Very fine clay	0.0005-0.00024	+11 to +12
Colloids	<0.00024	>+12

Cohesive sediment is characterized by the dispersed particle fall velocity, flocculated fall velocity of the suspension, the clay and nonclay mineralogy, organic content, and the cation exchange capacity. The fluid is characterized by the concentration of important cations, anions, salt, pH, and temperature. The boundary between cohesive and non-cohesive sediments is not clearly defined (Maggi *et al.*, 2007). Cohesion increases with decreasing particle size for the same type of material for instance, clays are much more cohesive than silts .

Sediments with size smaller than 62 microns are considered as the main constituent of cohesive sediment. Due to the complicated nature of cohesive sediments, particles in the range of 4 – 62 microns were the focus of this study. This classification helped in the quantification of sediments in the study.

2.3 Physically based models

Physically based models reproduce certain physical processes inherent in a process being modelled. For a given input data, a physically based model yields the same output for every model run provided that the model parameters, schematisation and boundary conditions remain unchanged. In other words a unique set of parameters will always yield the same solution. This feature makes them quite handy in interpreting the evolution of physical processes of the studied system under different conditions (Richardson *et al.*, 2007). For example, it is possible to carryout worst scenario or best scenario assessments. This way boundaries can be established within which a given aquatic system can be better managed. Likewise the behaviour of the system for various changes in system parameters can be implicitly and explicitly investigated.

Physically based models also present the possibility to predict the behaviour of the system over a long period, with the prospect of forecasting seasonal variability. With improved understanding of the system, intervention measures can be specifically directed towards controlling the most influential forcing parameters of the system basing on forecasted behaviour of the system under an altered physical environment.

Despite the fact that physically based models represent a range of physical processes an attribute (that favour their wide range of application over data-driven models) they have some shortcomings. Much as they incorporate the most important sub processes of the modelled process they remain a simplification of the physical system(Ye, 2006).

2.4 Data-driven models

Solomatine (1999) defines a data-driven model of a system as one that connects the system variables and is built on the basis of collected data. A number of data-driven models have been suggested to model various hydrological systems. These have included, autoregressive moving average (ARMA) models as suggested by Box and Jenkins (1970). Valencia and Schaake (1973) proposed disaggregation models whereas Panu and Unny (1980) evoked models based on the concept of pattern recognition. Stedinger and Taylor (1982) evaluated the performance of various data-driven models from which they established their strong points and shortcomings.

Some data-driven models such as linear regression models assume a linear behaviour of physical processes. In reality physical processes are non linear. Modelling of physical processes on the basis of linear relationships has often been found to be inadequate in comparison to non-linear relationships. It is therefore more logical to use data-driven models that encapsulate the non linear tendencies of the real world. Raman and Sunilkumar (1995) suggested artificial neural networks (ANN), which are suited to complex nonlinear problems, to be used for the analysis of real world processes.

Even in hybrid models where data-driven and physically based models tend to complement each other this notion still holds. Vojinovic et al, (2003) in their study of the hybrid approach to modeling of wet weather response in wastewater systems confirm that a hybrid model with nonlinear data-driven component has a superior performance to one with the linear data-driven component. It is on this basis that Artificial Neural Networks were selected to be used as the data-driven component in hybrid modeling.

However we should bear in mind that data-driven models are disadvantaged in the sense that their outcomes are difficult to interpret due to their 'black box' nature of operation. In addition they are unable to predict future performance of the system when subjected to physical change like changes in bathymetry say as a result of land reclamation.

2.5 Hybrid Modelling

2.5.1 Basis for hybrid modelling

Attempts have been made to combine physically based and data-driven modelling approaches. These have mainly concentrated on temporal modelling of processes at a single point in a system such as predicting flow in a river at a given point (Abebe, 2003; Vojinovic *et al.*, 2003). This approach can be extended to spatial modelling of complicated processes.

In the outgoing sections it has been disclosed that physically based models are based on the understanding of underlying mechanisms of the physical processes in a system. Conversely, data-driven models are built on the basis of collected data. This is does not imply that the underlying processes are not always known, but in most cases they are built when such knowledge is absent or disjointed. Solomatine, (1999) noted that in many cases there is an understanding of the modelled processes, but not very detailed to facilitate development of accurate behavioural models. Combining the two modelling approaches serves to compliment each other to produce more accurate results. This section further builds the case for hybrid modelling.

2.5.2 Similarities between physically based and data-driven views

Vogel (1992) considered a watershed model mathematically represented as

$$Q = f_d\left(P, PE/\Omega\right) + \varepsilon \qquad 2.2$$

Where Q is the river discharge, P is precipitation, PE is potential evapotranspiration, Ω represents model parameters and ε represents model errors. From the formulation it is noted that the stream flow is composed of both the physically based component represented by (P, PE/Ω) and a data-driven component represented by ε. In most cases such a model is regarded as physically based simply because the main driving forces have been relatively well represented and dominate the process. In the calibration of such a model one would aim at reducing model error term thereby reducing the calibration exercise to minimisation of the data-driven component. On this basis we can say that physically based models are not necessarily physically based since they contain empirical components. A data-driven model for the same problem can be formulated as

$$Q = f_d\left(P, PE/\Omega\right) + \psi \qquad 2.3$$

Where ψ represents the error term and the rest of the components are as previously defined. Still the physically based component is defined as (P, PE/Ω) and the data-driven component as the model error term. However as opposed to the physically based model the physically based component of the data-driven model is constructed with the intention of reproducing certain characteristics (such as variance, mean, skewness cross correlation between input and output ie P and Q etc) of the stream flow and it is not derived from the physical processes. With this kind of implementation data-driven models are unable to reproduce important physical processes. Therefore they can not be used for the same purpose as physically based models say for simulation of system performance under different conditions.

But since the general framework of data-driven and physically based models is the same, that is to say they both consist of physically based and data-driven terms several advantages can be derived from combining the two approaches. Data-driven views can help to develop physically based models which closely represent observed relationships between model input and outputs. Physically based view would help in building data-driven models which can realistically represent certain observed internal physical process (Vogel, 1999).

There are a couple of alternatives for the application of hybrid model to improve the reproduction of the modelled system (Abebe, 2003; Hur and Kim, 2008; Vojinovic *et al.*, 2003). The study concentrated on using a DDM to generate time varying boundary conditions for use in a physically based numerical model and secondly as an approximator for a physically based model.

2.6 Study area

2.6.1 The Maasmond area

The Maasmond is the approach channel to the port of Rotterdam one of the world's largest sea ports. The harbour provides passage to large ships. It is constantly dredged to maintain a depth of 24m. The area is one of the five major sediment entrapment areas along the French – Belgium- Dutch coastline. The other entrapment areas area: Wadden Sea along the Dutch and German coast, Haringvliet outlet, Eastern Scheldt mouth and Zeebrugge at the mouth of the Western Scheldt River. Figure 2.1 shows the location of the sediment traps in the North Sea. It is estimated that about 80% of the sediment in the Maasmond area originates from the British Channel and the French coast (Vuurens, 2001). The rest of the sediment is contributed by the discharges of the Rhine and Muese River.

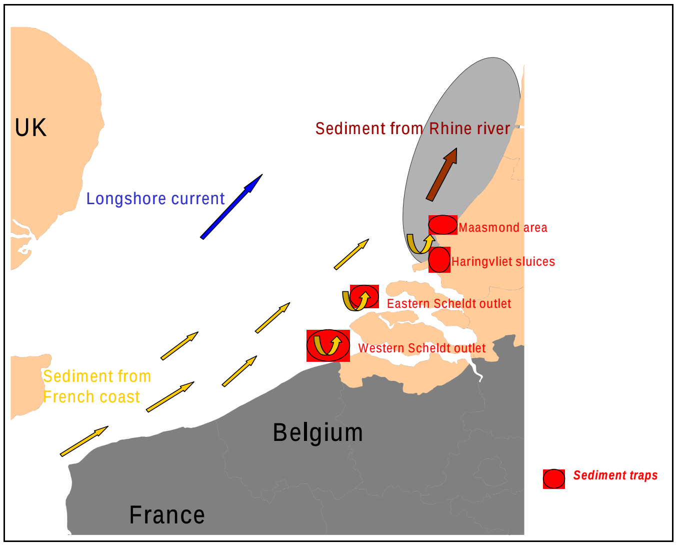


Figure 2.1 Location map showing the sedimentation traps in the North Sea adapted from Bhattacharya (2007)

The sediment in the Maasmond is mainly marine sediment originating from the down stream coastal areas. The rest is fluvial sediments originating from the Alps and the low-altitude mountain ranges in the Northern France and Belgium. The bed material of Maasmond is mainly silty-sand with D_{50} 0.055mm. The sediment size ranges from 0.1mm to 0.4mm whereas the organic component ranges from 8 – 23 % with increasing silt concentration in the landward direction (Chen and Eisma, 1995; Leussen, 1994)

Sediment in the North Sea has several sources the main ones being the Dover Straits and the Atlantic Ocean. A considerable amount is also obtained from the erosion of the French and British cliff coasts along the Dover Strait and rivers empting their waters into the North Sea. Dredging actives along the harbours in United Kingdom, France, German, Belgium, and The Netherlands also contribute a subtle amount. Estimate of the sediment flux from each of the sources is presented in table 2.

Table 2: Estimated sediment influx (in million ton/year) to the North Sea from different sources (Eisma D., 1988)

Dover Strait	Atlantic Ocean and Baltic Sea	Coastal erosion	Bed erosion	River	Total
20 to 30	10.5	2.2	9 to 13.5	4.8	46.5 to 61.0

During calm weather seasons, some of the sediment settles in the various sediment entrapment areas from which it is entrained during rough weather seasons. Frequent storms occur during winter and consequently the SPM concentration is an order of magnitude higher during winter as compared to summer (WL| Delft Hydraulics, 2001). In this context it can be concluded that the North Sea bed is not a net source of sediment material but rather a media that phases the availability of sediment.

The fresh water discharge of the Rhine River into the North Sea causes sediment – driven currents. This in addition to Coriolis force causes a freshwater *Coastal River* spanning a distance of 10-20km (WL| Delft Hydraulics, 2001). In the Coastal River, gravitational currents perpendicular to the coast are created. These lead to a higher concentration of SPM along the coast than offshore. As a result the density driven currents draw sediment to shore increasing the suspended sediment concentration even further.

As earlier alluded to, sedimentation in this area is mainly governed by the availability of sediment. The availability of sediment is a function of both stochastic meteorological forcing and river discharge and deterministic processes such as tidal and wave currents. From these complex forcing process modelling of sediment transport in the Dutch coast is quite complex. Yet it is necessary for the prediction of long term morphological changes so that proactive intervention measures are taken well in time.

2.6.2 Lake Victoria

Lake Victoria is the second largest freshwater body in the world. It is the most important shared natural resource of the East African Community (EAC). Located at 1,134m a.m.s.l, the lake has a mean surface area of 68,870 km^2, holds 2,760 Km^3 of water at an average depth of 40 m with a maximum depth of 85 m (LVEMP, 2002). Temperature range in the lake is narrow and warm with a mean annual temperature of 25 ^{o}C (Hecky, 1993). The shoreline is about 3,500 Km long shared between Uganda, Kenya and Tanzania. It encloses a number of small, shallow bays and inlets, many of which include swamps and wetlands that differ a great deal from one another and from the lake itself

The basin covers about 181,000 km2 with an estimated population of over 33 million. The gross economic product potential of the lake is estimated at US $ 5 billion (World Bank, 2005). The lake features the world's largest freshwater fishery with significant local consumption and exports to the European Union, and it is a global centre of aquatic biodiversity (LVEMP, 2007).

The lake has experienced serious decline in water quality since the 1960's. The decline is largely due to nutrient input from anthropogenic activities in the basin. Phosphorus concentrations and algal biomasses have increased significantly with cyanobacteria dominating the algal community. Total phosphorus concentration has risen by a factor of two. In short nutrient enrichment has stimulated increased algal primary production and has consequently led to massive eutrophication of the lake.

Microsoft ® Encarta ® Premium Suite 2005. © 1993-2004 Microsoft Corporation. All rights reserved.

Figure 2.2 Location of Lake Victoria in East Africa

Recent studies have shown that fine grained sediments in the lake have accumulated large amounts of phosphorus (P) from these excessive external loads. Studies indicate that lake sediments are a major source of P to the water column (Reddy et al 1995). Sediments in Lake Victoria produce an internal P load, which is approximately equal to external P loads on an annual basis. This internal P load also significantly impacts algal growth in the lake (Moore *et al.*, 1998; Olia and Reddy, 1993). Therefore, understanding the dynamic mechanism of sediment resuspension becomes an important task in the lake research.

2.6.3 Previous modelling on the Maasmond area

Under the Delft Cluster, considerable effort has been undertaken to understand the processes involved in the transportation, deposition and resuspension of cohesive sediments along the Dutch coast. This has been achieved by undertaking case studies along the coast particularly on the areas in the vicinity of the port of Rotterdam. Notable has been the studies by Roelvink (2001),Winterwerp (2006), Ye (2006)and Wang (2007) who sought to improve the understanding of the aforementioned processes using numerical modelling. They used Delft3D an integrated 2D-3D modelling package. In their work it was noted that the computational time was large in relation to the simulation time. Ye reported that it took 13 days computation time for a simulation period of 83 days. Regarding data driven modelling Vuurens (2001) and Bhattacharya et al. (2007) carried out similar work on the Dutch coast.

Numerical modelling of the coastal process of the Dutch coast has taken place under mainly two stages. The first stage constituted developing a coarse grid model known as the ZUNO model. This was followed by a finer local model known as RIJMAMO (Wang, 2007). The grids of the ZUNO and RIJMAMO models are shown in figure 2.3

2.6.4 ZUNO Model

The ZUNO model is a coarse curvilinear, boundary grid covering an area of about 800,000 km^2 . The coarse nature of the model is a trade off between the computational time and the need to cover a wide area of the North Sea. In the vertical direction Delft3D uses the 'sigma grid'. Here the vertical direction is divided into equal number of layers irrespective of the depth. Roelvink, (2001) suggested a logarithmic scale in which we have closer layers at the surface and at the bottom than in the middle of the water column. This ensures that the effects of wind at the water surface and the computational requirements of sediment transport along the sea bed are taken care of well.

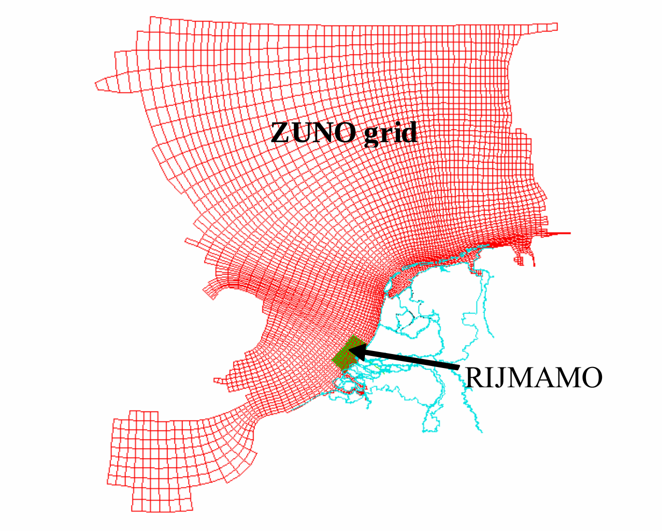


Figure 2.3 ZUNO coarse grid and RIJMAMO fine grid bathymetry (adapted from Wang 2007)

2.6.5 RIJMAMO Model

The RIJMAMO model is a finer grid model localised for the Dutch coast. The model was originally developed by WL | Hydraulic. Wang (2007) modified the model to facilitate modelling of sediments in the Maasmond area. It is the modified version by Wang that was used in this study. The model set up is in two phases. First is the FLOW model and then the WAVE model. The FLOW model is a 60 X 30 km finer grid and is used to simulate the hydrodynamics of the system. To facilitate a more accurate comparison of the simulated velocity and sediment concentration with the measured data, the model's vertical resolution near the sea bed has been amplified.

The WAVE model has the same resolution as the FLOW model but covers a larger area extending up to the Euro platform on the western side. It uses the same bathymetry data for the identical area covered both models. For the extended area of the WAVE model, bathymetry from the ZUNO model is used. Observed wave and wind data from the Euro Platform are used for the wave boundary. These data are obtainable at www.golfklimaat.nl/data and www.knmi.nl respectively.

Coupling of the FLOW model and the WAVE model provides the basis for the sediment model for the simulation of the SPM dynamics.

2.7 Implementation of the physically based Numerical model

Delft3D is the principal physically based modelling tool used in this study. The tool is an integrated flow and transport modelling system for the aquatic environment. The flow module implemented in Delft3D simulates the hydrodynamic conditions providing a backbone for other modules for modelling waves, water quality, morphology, and ecology.

In the subsequent sections the basic concepts of cohesive sediment transport processes as implemented in Delft3D are introduced. Only a limited number of processes whose understanding is presently clear are implemented.

2.7.1 Flow module

The flow module simulates two-dimensional depth averaged and three-dimensional unsteady flow and transport phenomena including density-driven flow. The depth averaged approach is appropriate for a fluid which is vertically homogenous (Herman, 2007). The three dimensional approach is important in transport problems in which the horizontal flow fields exhibit significant variation in vertical direction.

In the vertical direction the model provides two coordinate systems namely the Cartesian Z co-ordinate system and the sigma σ co-ordinate system. In the Z co-ordinate system the distance between the vertical layers is fixed. The number of vertical layers varies according to depth. In the σ co-ordinate system the number of vertical layers specified by the user is constant throughout the model domain but the distance between the layers varies with depth. The layers in the σ co-ordinate system are none-linear. They allow higher resolution to facilitate the study of a particular area in greater detail. It is for this reason that the σ-grid was adapted in this study.

The hydrodynamic model is used to solve the unsteady shallow water equations in 2D – depth averaged and 3D. In the horizontal direction the system of equations consist of the continuity equations and transport equations for conservative constituents (WL| Delft Hydraulics, 2001). On the basis of shallow water assumptions the momentum equation is reduced to a hydrostatic pressure equation in the vertical. The governing equations are solved with the finite difference scheme in curvilinear grid system. In this grid system the free water surface and the bathymetry are related to a flat horizontal plane.

The flow is forced by tide at the open boundaries, wind stress at the free surface, pressure gradients due to free surface gradients or density gradient. Source and sink terms are included in the equations to model the discharge and withdrawal of water.

Governing equations

The governing equations are based on the Navier Stokes equations for an incompressible fluid under shallow water, the Bousinesq and hydrostatic approximations are given in Equations 2.4 to 2.6

$$\frac{\partial \zeta}{\partial t}+\frac{\partial\left[h\bar{U}\right]}{\partial x}+\frac{\partial\left[h\bar{V}\right]}{\partial y}=S \tag{2.4}$$

$$\frac{\partial U}{\partial t}+U\frac{\partial U}{\partial x}+v\frac{\partial U}{\partial \eta}+\frac{\omega}{h}\frac{\partial U}{\partial \sigma}-fV=-\frac{1}{\rho_0}P_x+F_x+M_x+\frac{1}{h^2}\frac{\partial}{\partial \sigma}\left(\upsilon_v\frac{\partial u}{\partial \sigma}\right) \tag{2.5}$$

$$\frac{\partial V}{\partial t}+U\frac{\partial V}{\partial x}+V\frac{\partial V}{\partial y}+\frac{\omega}{h}\frac{\partial V}{\partial \sigma}-fU=-\frac{1}{\rho_0}P_y+F_y+M_y+\frac{1}{h^2}\frac{\partial}{\partial \sigma}\left(\upsilon_v\frac{\partial v}{\partial \sigma}\right) \tag{2.6}$$

Equation 2.4 is the depth averaged continuity equation where ζ is the water surface elevation, h is total water depth, U and V are depth averaged velocities in the X and Y directions. S is the contribution per unit area due to discharge or withdrawal of water, precipitation and evaporation.

Equations 2.5 and 2.6 are the horizontal momentum equations in the X and Y directions respectively. P_x and P_y, are the pressure gradients in the X and Y directions and for a fluid with non-uniform density, the local density is related to temperature and salinity by the equation of state. Thus the horizontal pressure gradients are given as

$$\frac{1}{\rho_0}P_x=g\frac{\partial \zeta}{\partial x}+g\frac{h}{\rho_0}\int_{\sigma}^{0}\left(\frac{\partial \rho}{\partial x}+\frac{\partial \sigma'}{\partial x}\frac{\partial \rho}{\partial \sigma'}\right)d\sigma' \tag{2.7}$$

$$\frac{1}{\rho_0}P_y=g\frac{\partial \zeta}{\partial y}+g\frac{h}{\rho_0}\int_{\sigma}^{0}\left(\frac{\partial \rho}{\partial y}+\frac{\partial \sigma'}{\partial y}\frac{\partial \rho}{\partial \sigma'}\right)d\sigma' \tag{2.8}$$

F_x and F_y represent the unbalance of horizontal Reynold's stresses determined by using eddy viscosity concept. M_x and M_y are contributions due to external sources or sinks of momentum (external forcing due to hydraulic structures, wave stresses, discharge or withdrawal of water, etc.). υ_v is the vertical eddy viscosity coefficient and ω is the vertical velocity in the adapting σ-grid system relative to the moving σ-plane.

With respect to the vertical, the momentum equation is reduced to the hydrostatic pressure equation on the basis of shallow water assumption. This is given by

$$\frac{\partial P}{\partial \sigma} = -g\rho h \qquad 2.9$$

Transport equations

The sediment transport model in Delft3D is expressed as an advection diffusion equation. A first order decay condition is also taken into account. It should be noted that this model is also applicable to the transportation of other dissolved substances, salinity and heat.

$$\begin{aligned}&\frac{\partial[hc]}{\partial t} + \frac{\partial[hUc]}{\partial x} + \frac{\partial[hVc]}{\partial y} + \frac{\partial[\omega c]}{\partial \sigma} \\ &= h\left[\frac{\partial}{\partial x}\left(D_H \frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_H \frac{\partial c}{\partial y}\right)\right] + \frac{1}{h}\frac{\partial}{\partial \sigma}\left[D_v \frac{\partial c}{\partial \sigma}\right] + hS\end{aligned} \qquad 2.10$$

S Is the source and sink terms per unit area due to discharge q_{in} or withdraw q_{out} of water and the exchange of heat through the free surface Q_{tot}.

$$S = h\left(q_{in}c - q_{out}c\right) + Q_{tot} \qquad 2.11$$

D_H is the horizontal vertical eddy diffusivity defined as

$$D_H = D_{2D} + D_{3D} + D_H^{back} \qquad 2.12$$

Where D_{2D} is 2D-turbulance associated with mixing due to horizontal motions and forcing, D_{3D} is the 3D-turbulance related to the turbulent eddy viscosity and D_H^{back} is the vertical background diffusion.

D_V is the vertical eddy diffusivities defined as

$$D_v = \frac{\upsilon_{mol}}{\sigma_{mol}} + \max\left(D_v^{back}, D_{3D}\right) \qquad 2.13$$

υ_{mol} is the kinematic viscosity coefficient and σ_{mol} is Prandtl-Schmidt number for constituent mixing (0.7).

Boundary conditions

To obtain a well posed mathematical problem with a unique solution a set of initial and boundary conditions for water levels and horizontal velocities have to be specified. Two types of boundaries exist in the 3D and 2D depth averaged model, namely open boundaries and closed boundaries. Open boundaries are artificial "water –water" boundaries and are normally introduced to restrict the computational area and consequently the computation effort. Closed boundaries are natural boundaries consisting of water-and any other natural environment such as land or air.

Vertical boundary conditions

The kinematic boundary conditions take the impermeability of the surface and the bottom into consideration in the following expressions.

$$\omega\big|_{\sigma=-1} = 0 \text{ and } \omega\big|_{\sigma=0} = 0 \qquad 2.14$$

Where σ= -1 is the bottom and σ = 0 is the surface.

Bed boundary conditions

At the bed a frictional boundary condition for the momentum equations is specified as

$$\frac{\upsilon_V}{h}\frac{\partial u}{\partial \sigma}\bigg|_{\sigma=-1} = \frac{1}{\rho}\tau_{bx} \; and \; \frac{\upsilon_V}{h}\frac{\partial v}{\partial \sigma}\bigg|_{\sigma=-1} = \frac{1}{\rho}\tau_{by} \qquad 2.15$$

τ_{bx} and τ_{by} represent the bed shear stress components and they combine the wave and current effects. Details of the computation of the bed shear stress are provided in (WL| Delft Hydraulics, 2006)

Surface boundary conditions

For the free surface the boundary conditions for the momentum equations are

$$\frac{\upsilon_v}{h}\frac{\partial u}{d\sigma}\bigg|_{\sigma=0} = \frac{1}{\rho_0}\left|\vec{\tau}_s\right|\cos(\theta),\; \frac{\upsilon_v}{h}\frac{\partial u}{d\sigma}\bigg|_{\sigma=0} = \frac{1}{\rho_0}\left|\vec{\tau}_s\right|\sin(\theta) \qquad 2.16$$

θ Is the angle between the wind.

Open boundary conditions.

At the open boundary either the water level or normal velocity component or both are specified. The boundary conditions are derived from the basic Riemann invariants for the linearised 1D equation assuming zero flow along the boundary. These are implemented to reduce reflections at the open boundary.

$$R = U \pm 2\sqrt{gh} \qquad 2.17$$

For the transport equations it is assumed that advection processes are dominant over diffusion processes. Thus the concentrations are specified at inflow and no concentrations are specified at the outflow.

2.7.2 The Wave model

In large open water bodies, wind induced waves have a big influence on flow and consequently sedimentations dynamics (Winterwerp, 2006) In this study a wave model has been incorporated to take into account the effects of waves on current and wave induced stress. This was implemented through online coupling of the SWAN model with the Delft3D-FLOW. SWAN model (Simulating WAves Nearshore) simulates the evolution of random, short-crested wind-generated waves in lakes, estuaries, tidal inlets etc (WL| Delft Hydraulics, 2006). The online coupling between the WAVE and FLOW modules implies dynamic interaction in which data is exchanged using the communication file. It contains the most recent data for the flow and wave computation.

The model is fully spectral described by a 2D density spectral wave action balance equation given as

$$\frac{\partial}{\partial t}N + \frac{\partial}{\partial x}c_x N + \frac{\partial}{\partial y}c_y N + \frac{\partial}{\partial \sigma}c_\sigma N + \frac{\partial}{\partial \theta}c_\theta N = \frac{S}{\sigma} \qquad 2.18$$

N represents the density spectrum with parameters σ and θ.

The first term in Equation 3.15 represents the local rate of change of action density in time. The second and third terms represent propagation of action in geographical space (with propagation velocities c_x and c_y in x - and y -space, respectively). The fourth term represents shifting of the relative frequency due to variations in depths and currents (with propagation velocity c_σ in σ -space). The fifth term represents depth-induced and current-induced refraction (with propagation velocity c_θ in θ -space). The term S (= S (σ, θ)) at the right-hand side of the action balance equation is the source term in terms of energy density representing the effects of generation, dissipation and non-linear wave-wave interactions (WL| Delft Hydraulics, 2006).

2.7.3 Sediment transport model

The sediment transport model is expressed as a three dimensional advection-diffusion mass conservation equation. i.e.

$$\frac{\partial c^{(\ell)}}{\partial t} + \frac{\partial uc^{(\ell)}}{\partial x} + \frac{\partial vc^{(\ell)}}{\partial y} \frac{\partial\left(w - w_s^{(\ell)}\right)c^{(\ell)}}{\partial z} - \frac{\partial}{\partial x}\left(\varepsilon_{s,x}^{\ell} \frac{\partial c^{(\ell)}}{\partial x}\right) - \frac{\partial}{\partial y}\left(\varepsilon_{s,y}^{\ell} \frac{\partial c^{(\ell)}}{\partial y}\right) - \frac{\partial}{\partial z}\left(\varepsilon_{s,z}^{\ell} \frac{\partial c^{(\ell)}}{\partial z}\right) = 0 \qquad 2.19$$

$c^{(\ell)}$ Mass concentration of sediment fraction (ℓ) Kg/m^3
u, v and w Flow velocity components m/s
$\varepsilon_{s,x}^{\ell}, \varepsilon_{s,y}^{\ell}$ and $\varepsilon_{s,z}^{\ell}$ Eddy diffusivities of sediment fraction (ℓ) m^2/s
$w_s^{(\ell)}$ (Hindered) sediment settling velocity of sediment fraction (ℓ) m/s

The equation is computed in the same manner as the transport equations of any other conservative constituent. However the main difference is that this particular one takes into account the exchange of sediment between the bed and flow and the settling velocity under the influence of gravity. The local flow velocity and eddy diffusivities are computed from the hydrodynamic module.

Computation of settling velocity takes into account the effect of salinity on cohesive sediments. In saline water, cohesive sediments flocculate, forming "flocs" that settle much faster than the individual sediment particles. The degree of flocculation depends on the salinity of the water. The salinity dependant settling velocity is computed as follows

$$w_{s,0}^{(\ell)} = \frac{w_{s,\max}^{(\ell)}}{2}\left(1-\cos(\frac{\pi S}{S_{\max}})\right)+\frac{w_{s,f}^{(\ell)}}{2}\left(1-\cos(\frac{\pi S}{S_{\max}})\right) \text{ when } S \leq SALMAX \qquad 2.20$$

$$w_{s,0}^{(\ell)} = w_{s,\max}^{(\ell)}, \text{ when } S > SALMAX$$

where:

$w_{s,0}^{(\ell)}$: The (non-hindered) settling velocity of sediment fraction (ℓ).

$w_{s,\max}^{(\ell)}$: WSM, settling velocity of sediment fraction (ℓ) at salinity concentration SALMAX.

$w_{s,f}^{(\ell)}$: WS0, fresh water settling velocity of sediment fraction (ℓ).

S : Salinity

$S_{\max}$: SALMAX, maximal salinity at which WSM is specified.

For the erosion and deposition of cohesive sediment at the bottom boundary the formulae of Parthniades and Krone are applied as in equations (3.18 through 3.20). The computed erosion and deposition is applied to the near bed computation cell through the appropriate sink and source terms. The advection, settling and diffusion through the bottom cell are set to zero to prevent double counting.

$$E^{(\ell)} = M^{(\ell)} S\left(\tau_{cw}, \tau_{cr,e}^{(\ell)}\right) \qquad 2.21$$

$$D^{(\ell)} = w_s^{(\ell)} c_b^{(\ell)} S\left(\tau_{cw}, \tau_{cr,d}^{(\ell)}\right) \qquad 2.22$$

$$c_b^{(\ell)} = c^{(\ell)}\left(z=\frac{\Delta z_b}{2}, t\right) \qquad 2.23$$

where

$E^{(\ell)}$: erosion flux [kg/m^2/s]

$M^{(\ell)}$: user specified erosion parameter [kg/m^2/s]

$S\left(\tau_{cw}, \tau_{cr,e}^{(\ell)}\right)$: erosion step function:

$$S\left(\tau_{cw}, \tau_{cr,e}^{(\ell)}\right) = \left(\frac{\tau_{cw}}{\tau_{cr,e}^{(\ell)}} - 1\right), \text{ when } \tau_{cw} > \tau_{cr,e}^{(\ell)},$$

$$= 0 \qquad \text{when}\, \tau_{cw} \le \tau_{cr,e}^{(\ell)}. \qquad 2.24$$

$D^{(\ell)}$: Deposition flux [kg/m^2/s]

$w_s^{(\ell)}$: Fall velocity (hindered) [m/s]

$c_b^{(\ell)}$: Averaged sediment concentration in the near bottom computational layer

$S\left(\tau_{cw}, \tau_{cr,d}^{(\ell)}\right)$: Deposition step function:

$$S\left(\tau_{cw}, \tau_{cr,d}^{(\ell)}\right) = \left(1 - \frac{\tau_{cw}}{\tau_{cr,d}^{(\ell)}}\right), \text{ when } \tau_{cw} < \tau_{cr,d}^{(\ell)},$$

$$= 0 \text{ when } \tau_{cw} \ge \tau_{cr,d}^{(\ell)}. \qquad 2.25$$

τ_{cw}: mean bed stress due to current and waves as calculated by the wave-current interaction model

$\tau_{cr,e}^{(\ell)}$: user specified critical erosion shear stress [N/m^2]

$\tau_{cr,d}^{(\ell)}$: user specified critical deposition shear stress [N/m^2]

(ℓ): Sediment fraction (ℓ)

2.8 Conclusion

Literature has shown that suspended sediment concentration and siltation can be modelled by both physically based and data-driven models. Although both modelling approaches have distinct advantages they have several drawbacks. If combined these two approaches could compliment each other to produce more accurate simulation results.

3 Approximating a Numerical model using a DDM

In this chapter we explore the possibility of approximating a physically based model using a data driven model to predict SPM concentrations. Necessary input variables and input data manipulation operations to expose maximum information to the data driven modelling tool were identified.

3.1 Introduction

Physically based models are undoubtedly powerful tools in predicting physical phenomena with the possibility of having numerous scenario assessments. However one of their drawbacks is the expensive computation time. Sometimes several simple approximations of the physical process are required to be developed in a short time - say during scenario development. But this may not be possible with fine grid physically based models given the fact that they may take several days of simulation time.

While modelling SPM concentration using Delft3D, two major modules are used that is, the SWAN module and the FLOW module. The former simulates wave propagation whereas the later simulates unsteady flow and transport phenomena in an aquatic environment. Running these modules concurrently consumes a lot of time, for example it would require 14 days computation time to simulate a period of 83 days at a time step of 1 min (Wang, 2007).

Alternatively hydrodynamic variables generated by the SWAN model can be used to build a data driven model (DDM) to approximate a physically based to predict suspended particulate matter (SPM). This approach could drastically reduce the computation time for predicting SPM. Consequently several scenarios could be analysed in a short period of time. This study therefore seeks to investigate the possibility of using a DDM to approximate a physically based model following the aforementioned approach.

3.2 Artificial Neural Network as data driven models

Artificial Neural Networks are computational systems comprising of several interconnected simple processing elements performing tasks in a manner that is analogous to the human brain. The processing elements are analogous to the biological neurons and process inputs from single and multiple sources producing outputs according to a predefined function. The objective of the ANN is to learn a function (*f)* given an input vector (*x*) that corresponds to a particular output vector (*y*).

ANN have several topologies the most commonly used being the Multi Layer Perceptron (MLP)(Bhattacharya *et al.*, 2007). The MLP network topology is constrained to feed forward in which connections are only possible from the input layer to the first hidden layer, from the first hidden layer to the second... and from the last hidden layer to the output layer. No communication is allowed between the processing units of a given layer. However the units can send their outputs to a single unit in a succeeding layer. Figure 3.1 shows a schematisation of an MLP network topology. The ANN used in this study was composed of 3 layers with one input layer, one hidden layer and an output layer.

3.2.1 Learning of ANN

The learning process of a MLP ANN is referred to as supervised learning. Tasks that are mainly performed within the paradigm of supervised learning are pattern recognition or classification and regression or function approximation. In supervised learning a set of example pairs say (x, y) $x \in X, y \in Y$ are supplied to the network and the aim is to find a function *f* that matches the instances. By this we try to infer the mapping implied by the data. This is achieved by implementing an algorithm known as the back propagation.

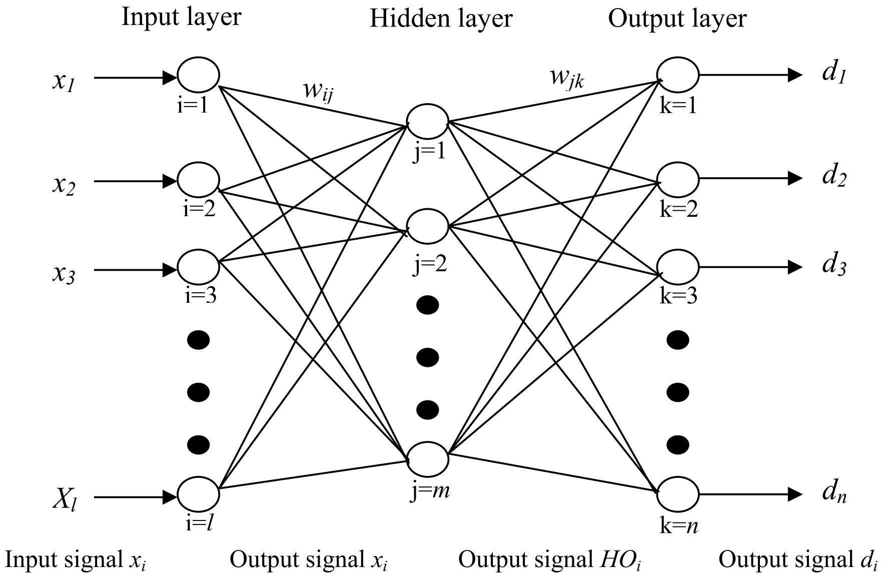


Figure 4.1 shows a schematisation of an MLP network topology

The learning process of the MLP network used in this study is described basing on the schematisation in Figure 3.1. The input layer consists of *p* data input variables each consisting of x_i $i = 1.....l$ and an output value O_k , $k = 1,\ldots,n$. At the beginning of the learning process the weights on the connections were assigned values randomly. Two phases are involved in the back propagation learning namely forward pass and backward pass.

Forward pass

In the forward pass the external data at the input unit is propagated through the network to calculate the output signal at the output units. The inputs to the first hidden layer is obtained as the summation of the product of the input values x_i and the first interconnection weights, w_{ij}, $j = 1, ..., l$ at the hidden nodes over index *i*,

$$H_j = \sum_{i=1}^{l} w_{ij} x_i \quad j = 1,...,l \tag{3.1}$$

Where

H_j = input to the jth-hidden node,
w_{ij} = connection weight from the ith node to the jth node.

The input to the hidden node is transformed through a sigmoid function producing a hidden node output denoted as HO_j and is given by

$$HO_j = \frac{1}{1 + e^{-(H_j + \theta_j)}} \tag{3.2}$$

θ_j is a threshold value whose effect is to offset the center of the sigmoid function from zero. θ_j is learned in the same way as the connection weights. HO_j is subsequently fed into the succeeding hidden layer. This process is repeated until the last hidden layer is reached. The output O_n from the n^{th} output node is given by

$$O_k = f\left(\sum_{j=1}^{l} w_{ik} HO_j\right) \quad k = 1,...,n \tag{3.3}$$

Where the activation function is the same form as in (4.2) O_n is the neural network output however it is not the same as the desired value. Thus a cost function is invoked for the mismatch between the mapping function and the actual data. Normally the mean-squared error is used as the cost (Raman and Sunilkumar, 1995), so we try to minimise the average error between the networks' output, $f(x)$ and the desired value D over the entire example pairs. The error at the output units is computed from the difference between the desired out and the calculated network output. The error function e_p for the p^{th} input pattern is then given as;

$$e_p = \frac{1}{2}\sum_{k=2}^{n} (d_k - O_k)^2 \tag{3.4}$$

And consequently the entire system error E for all the input examples is given by

$$E = \frac{1}{2N}\sum_{p=1}^{N}\sum_{n=1}^{m} (d_{pn} - O_{pn})^2 \tag{3.5}$$

Where

d_{pn} = target value, d_n, for the pth pattern and
O_{pn} = network output value, O_n for the pth pattern.

The task then is to minimise this cost function (average squared error). This is achieved by adjusting the connection strength (weights) between the computational units. Weights are changed by an amount, which is proportion to the production of the error at the unit and the output of the unit feeding into the weights. In the back propagation-learning algorithm, the gradient descent method is used to optimise the cost. The method is accomplished by first computing the gradient (δ_n) for each node on the output layer given as;

$$\delta_n = O_n(1-O_n)(d_n - O_n) \qquad 3.6$$

Then the error gradient (δ_j) at the hidden nodes is calculated by propagating the error at the output units backwards through the weights for each intermediate layer.

$$\delta_j = HO_j(1-HO_j)\sum_{n=1}^{m}\delta_n w_{jn} \qquad 3.7$$

The associated adjustments to update the network weights are computed using expression 4.8. This process is repeated for each layer until the input layer is reached.

$$\begin{aligned} \Delta w_{ij}(r) &= \eta\delta_j x_i \\ \Delta w_{ij}(r+1) &= w_{ji}(r) + \Delta w_{ji}(r) \end{aligned} \qquad 3.8$$

Where

r = iteration number and

η = Learning rate that provides the step size during the gradient descent.

The learning rate determines the amount of the calculated sensitivity to weight change to be used for updating the weight. A large learning rate speeds up the convergence process, though it can result in non-convergence (Minns, 1998). Small learning rates produce more reliable results at the expense of increased training time.

Backward pass

The process of propagating the error signal backwards through the network is known as the backward pass and is repeated for each intermediate layer until the input layer is reached. The rate of learning can be improved by modifying the generalized delta rule by introducing a momentum term. The expression for updating the weights now becomes;

$$\Delta w_{ji}(r) = \eta\delta_j x_i + \alpha\Delta w_{ji}(r-1) \qquad 3.9$$

Where α is the momentum term. The momentum term determines the influence of the previous weight change on the current direction of movement in the weight space. $\Delta w_{ji}(r)$ is the change to be made to the weights for the current iteration r and $\Delta w_{ji}(r-1)$ is the weight change in the previous iteration. The momentum term speeds up convergence along shallow gradients by allowing the approach to the solution to pick up speed in the downhill direction (Raman and Sunilkumar, 1995)

The next input-output data pair is applied to the network and weights are adjusted to reduce the new error term. This process is repeated until all the training set has been applied. The whole process is then repeated starting with the first example in the training data set while modifying the weights in the process until convergence is achieved or the error is at an acceptable level. After convergence is attained the set of weights captures the knowledge and the information in the examples used in training process. When the network is presented with a test or production data set, the forward pass computation results in an output, which is a generalisation of the network learned and stored in the weights. Further details about ANNs can be found in Haykin (1999); Minns (1998) and Schalkoff (1997)

3.3 Understanding the process to be modelled

For successful modelling it is important to understand the relationship between the physical characteristics of the process being modelled. This helps in selecting the most appropriate input and output variables. This section therefore briefly reviews the interaction of variables involved in the sedimentation process.

SPM concentration and siltation is governed by the availability of sediment and transport processes. The availability is influenced by wind, waves, tides, river discharge, etc. Once sediments are available the transportation is then influenced by local hydro-meteorological conditions.

The process of sediment resuspension and transport in large shallow water bodies mainly starts with wind energy being delivered to the water surface and causing waves (Jin and Ji, 2004). The wind energy is transmitted from the surface to the bottom while being dissipated into wave motion in the vertical direction. This phenomenon creates orbital velocities at the sediment-water interface, which in combination with current velocities exert shear stresses leading to entrainment of sediment into the water column. In this respect the sediment transport flux is related to waves through the wave energy. We can then deduce that wind waves and current are the major driving forces to cause sediment resuspension in shallow water bodies.

In coastal areas in addition to waves and currents induced by wind, sediment resuspension/transport is also influenced by tidal induced current. In areas where fresh river water is discharged into the sea or ocean, density driven currents are created and they lead to a higher concentration of suspended sediments (WL| Delft Hydraulics, 2001).

3.4 Building the data-driven model

3.4.1 Identification of input variables

With the above understanding of the major processes influencing SPM concentration (wind waves and currents), we can identify a set of process parameters that can be used to successfully build a data driven model. In the case of wind waves the associated parameters can be identified as;

- Significant wave height $(Hsig)$ and,
- Bed shear stress due to waves, (τ_{bwav})

The current related variables are;

- Current velocity (v_{cur}) and
- Bed shear stress due to currents (τ_{bcur})

A more detailed justification for the choice of the variables is given in chapter 4. From the identified forcing functions the predictive sedimentation model can be conceptualised as;

$$S_s = f\left(H_{\text{sig}}, \tau_{bwav}, v_{cur}, \tau_{bcur} \right) \qquad 3.10$$

A desired situation would be to build a model with all the above-identified parameter. However, in this particular block of the study only the total bed shear stress could readily be extracted from the numerical model runs and thus it was the only input variable considered. In the subsequent blocks these variables were available and they were used effectively.

3.4.2 Building an Information Preparation Environment

Having identified the input parameters it follows that some manipulations should be employed on the input data to expose maximum information to the modelling tool. By expose maximum information we imply amplifying the information contained in the input data so that it becomes more conspicuous to the tool. This facilitates effective learning of the modelling tool and consequently more accurate predictions. The techniques involve in achieving this include, investigating variables to establish inter and cross correlation, redistributing distributions, normalising the range and reducing variance along the range among others. The outputs of this operation are the input variables to the DDM. The following section covers key data manipulation techniques that were found to increase information exposure to the DDM. The set of these successful operations in their chronological order is what has been referred to as the Preparation Information Environment (PIE)(Pyle, 1999)

3.4.3 Visual inspection of data

The importance of having an insight into the most suitable variables that form inputs to the DDM cannot be underestimated. It helps to determine the possibility or necessity of adjusting or transforming the data. Furthermore it establishes reasonable expectation of achieving a solution and reveals the relevance of the data to the task at hand. This was achieved by plotting the identified possible predictors together with the variable to be predicted.

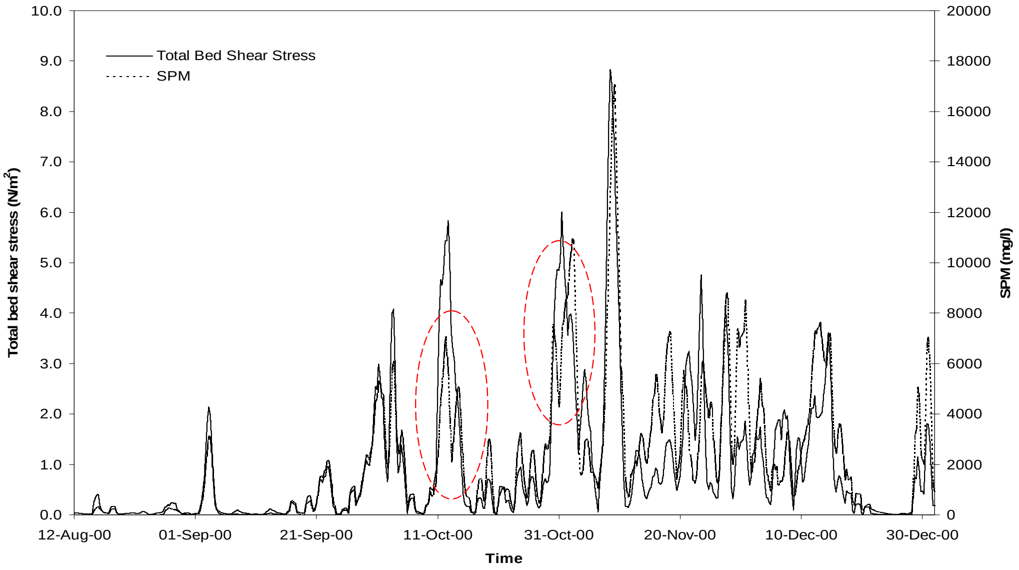


Figure 3.1 plot of total bed shear stress and SPM with a 12hr lag moving average

Figure 3.1 shows a plot of the total bed shear stress and the SPM concentration. The two series were smoothed using a 12 hours lag moving average so as to obtain a clear outlook. A 12 hours lag time was chosen to capture the long wave effect that is associated with the diurnal tidal movement and to smoothen out the shortwave effects that would otherwise appear as noise in the model.

The plot shows a good relationship between the total bed shear stress and the SPM. However, there are some discrepancies in some instances as indicated by the areas circled in the plot. In these instances the SPM abruptly drops tremendously and rises when the bed shear stress is continuously rising or falling. Secondly, in some instances the SPM does not rise in the same proportion as the total bed shear stress as seen in the plot. This suggests that the discrepancies could be due to other processes that could be fundamental in predicting SPM in addition to the total bed shear stress (TBSS).

3.4.4 Analysing the inter-dependencies between variables

As illustrated in section 3.4.1 the physical processes (such as bed shear stress, and SPM or Hsig and SPM) are related in one way or another however the degree of relation may vary from one variable to another. The measure of how values of one variable change as values of another variable change is known as correlation. By knowing one variable's values gives some idea of the value of the other variable though not a complete idea (Pyle, 1999). We should bear in mind that correlation does not tell much about the goodness of fit between two variables (because the two variables may not have a linear relationship) but rather, how one variable behaves in regard to the changing values of the other. Perhaps a better measure of the explanatory power of one variable has about the values of the other is the sample coefficient of determination (Pyle, 1999). But the correlation coefficient can give a big insight on the appropriate inputs to a data driven model.

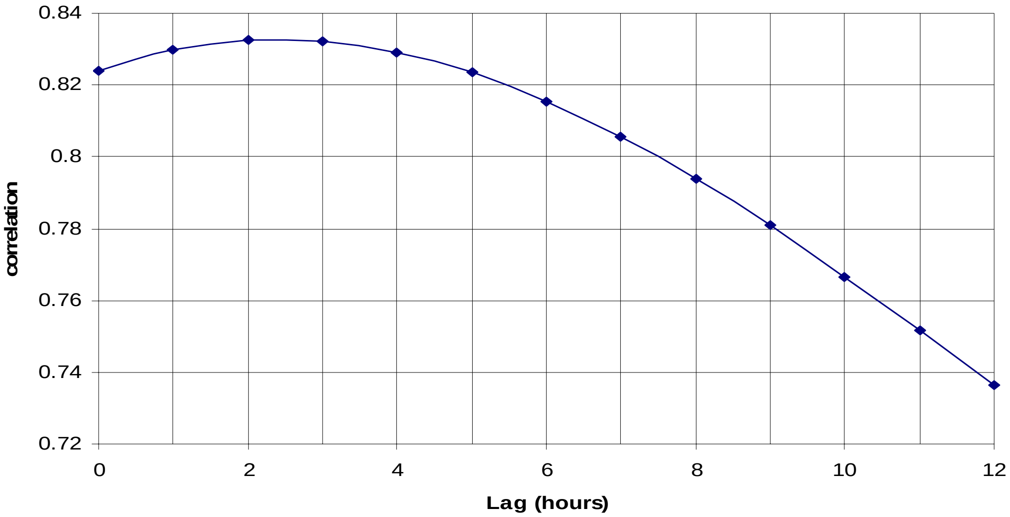


Figure 3.2 correlgram of SPM and Total bed shear stress up to lag 12 hours

Normally for physical processes, the state of the system at a point in time has influence on the state of the system at the subsequent step in time. Thus running a lagged correlation on the predictive time series gives a clue on the appropriate lag of the input variables. Figure 3.2 shows the correlgram of SPM and total bed shear stress lag up to 12 hours.

The bed shear stress with two hours lag gave the highest correlation with a correlation coefficient of 0.8235. The total bed shear stresses (tbss) with lag zero up to six were chosen as the input variables. Thus it follows that the predictive sedimentation model is as given in Equation 3.11

$$S_s = f\left(tbss, tbss-1, tbss-2, tbss-3, tbss-4, tbss-5, tbss-6, \right) \tag{3.11}$$

Where $tbss-1$, $tbss-2$,..........., $tbss-6$ are the total bed shear stress with lag 1 through 6

3.4.5 Data filtering

High variance in data is undesired for DDMs. The phenomenon considerably reduces the model's learning capability and consequently its performance. If a DDM is used as a regression model it is necessary to map the predictor to the desired values. Thus it follows that maintaining the pattern is the most important aspect in the modelled dataset. For better performance of the DDM, it calls for smoothing to reduce variance along the data range (Schalkoff, 1997). The intention is to reduce amplitude in the case of higher frequencies and to amplify in the case of lower frequencies. Several smoothing tools can be used including simple moving average, weighted moving average, peak-valley mean, Gaussian filter etc. In this study the simple moving average and the Gaussian filter were considered. To accommodate the tidal diurnal pattern a 12hours lag (smoothing kernel) has been considered in both cases.

The Gaussian filter

It is a type of filter that uses the Gaussian function as its filtering kernel. The Gaussian filter is expressed as

$$h_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-i^2 / 2\sigma^2} \tag{3.12}$$

h_i is the coefficient of the ith instance in the filtering kernel. σ is a parameter determining the amount of smoothing that is performed on the data. As sigma increases the amount of filtering also increases. The value of sigma can be determined if the desired size of the filter kernel is known. It is advisable that the sigma should be 70% of the size of the filter neighbourhood. That is to say the range to the left and right of the value we are looking at (http://www.cse.msu.edu/~cse471/gaussian.pdf). Thus σ=0.7n. Where n is the size of the neighbourhood to the left and right of the instance in consideration. The filter coefficients are then computed from Equation 3.13. If we a filtering kernel to be 12 hours then n will be 11. The Gaussian filtered values (g) is thus computed as

$$g_i = \sum_{i=-11}^{11} x_i h_i \tag{3.13}$$

Figure 3.3 shows a plot of the 12 hours SPM moving average and Figure 3.4 shows a plot of the smoothed SPM series using a Gaussian filter of 12 hours filtering kernel. From the plot it is noted that the simple moving average as a filtering kernel tends to shift the maxima and minima to the right as the kernel size increases. On the other hand, the Gaussian filter maintains the positions of the peaks and minimas. A draw back of the simple moving average is that each data point in the weighting kernel contributes equally. But it is logical that the more distant data points have less influence on the data point in focus.

The Gaussian filter has the advantaged of weighting in that the points close to the point in focus contribute more than the distant ones. The Gaussian filter gives a better correlation of 0.8942 between the SPM and total bed shear stress as compared to a correlation of 0.8127, when using the simple moving average. Since we are interested in have higher correlation between the predictor and the predicate the gaussian filter was choose as a smoothing tool for data.

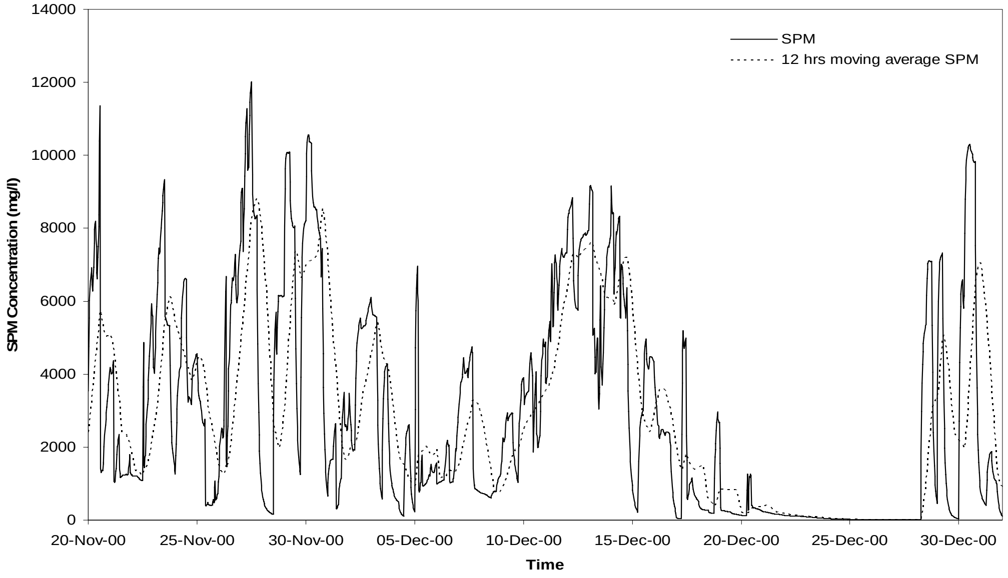


Figure 3.3 Plot of the SPM moving average

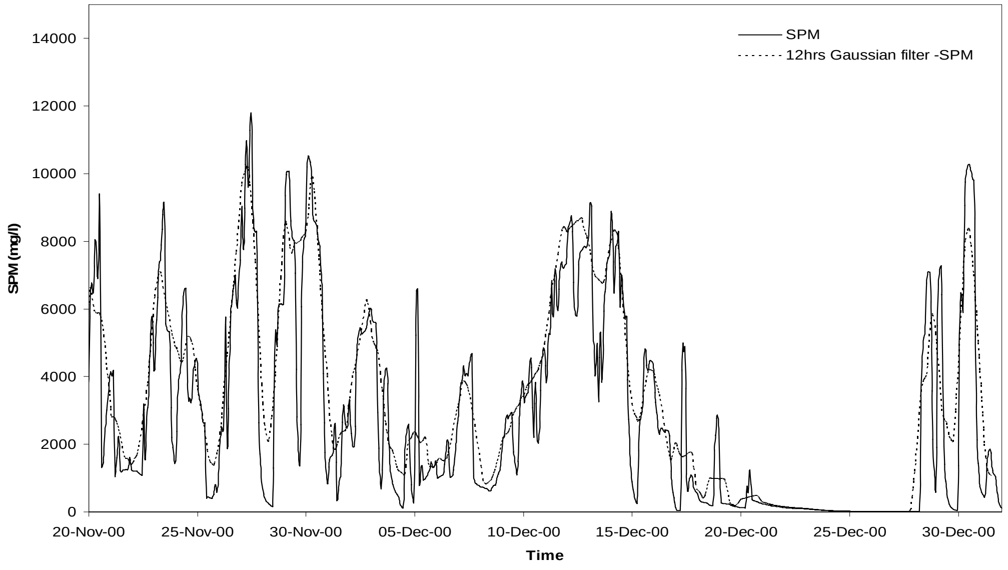


Figure 3.4 plot of SPM filter with a Gaussian filter

3.4.6 Choosing the training, validation and testing data sets

Developing a model mainly takes place in three phases. The first is training followed by testing and finely the execution phase Where the model is applied to production data. During training there is one data set, which is part of the data to build the model and a testing set, which could also be probably part of the same data set. It is important that the whole range of the building data set is captured during training, so that during the testing phase no values are seen as being out of range.

Depending on the distribution of the data, choosing the training, validation and testing data sets in the sequence of the time series is likely to train the network on a smaller (or larger) range of the data range. Thus when the network is presented with larger (or smaller) values which it did not see during the training phase the values are squashed within the tail and head sections of the logistic function (Solomatine, 1999). Figure 3.5 shows a plot of the observed and predicted values of SPM having chosen the training, cross validation and testing data sets in sequence. From the plot it is seen that there is a value above which the predicted SPM is suppressed to a straight line. This is probably due to the fact that the network was trained on smaller values of the series and that the testing set drawn from the end of the time series had larger fluctuations.

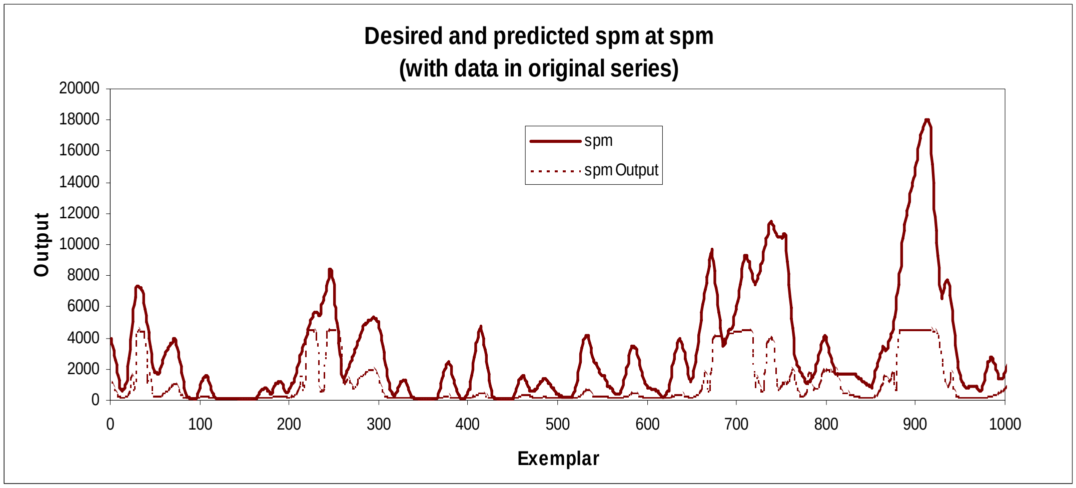


Figure 3.5: Plot of desired and predicted SPM with data in its original series

As a solution to this problem the data set was randomised. The training, cross-validation and testing data sets where then selected as percentages i.e. 60:15:25% of the randomised dataset respectively. This reduced the mean square error (MSE) by 78% as compared to that of a non-randomised dataset. Table 4.1 shows the statistics of the training, cross validation and testing datasets whereas table 4.2 shows the statistical performance parameters for the original data and the randomised data.

Table 3.1: Statics of training, cross validation and testing datasets

Statistics		Training	Cross validation	Testing
Bed shear stress (N/m^2)	Mean	0.735	0.412	0.939
	Std*	1.155	0.686	1.463
SPM (mg/l)	Mean	1278	534	2023
	Std	2097	809	2866

---

* Standard deviation

Table 3.2 performance measurements for original and randomised data set in predicting SPM units are mg/l

Performance	Original data	Re-arranged
MSE	10,784,906	2,206,838
NMSE	1.1487	0.2688
MAE	2289	916
Min Abs Error	0.013	0.320
Max Abs Error	13515.0	6414.5
r	0.70669	0.87070

3.4.7 Adjusting the distribution

Most modelling tools are designed to handle data on the assumption that its distribution is normal or regular. They may have difficulties to extract information from data with wide density distribution. Besides if the distribution of any two variables is dissimilar the problem is even intensified (Pyle, 1999). This means that not only the distributions of any two variables are different but also the irregularities are not shared by both variables. Transforming the data can eliminate this. Several transformation techniques are in existence but the Box-Cox method is widely used and therefore was preferred in this study.

The Box-Cox method is a non-linear data transformation method in which first a transformation is applied as shown in Equation 3.14.

$$x_i^i = \frac{x_i^\lambda - 1}{\lambda} \qquad 3.14$$

Where x_i is the original value

x_i^i is the transformed value

λ is a user-selected value taken as 0.6 in this case

A balancing of the distribution (Equation 3.15) follows this i.e.

$$x_i^{"} = \frac{x_i^{'} - E(x^{'})}{\delta_{x^{'}}} \qquad 3.15$$

Where $x_i^{"}$ is final standardized value

$E(x^{'})$ Is mean value of variable x’

$\delta_{x^{'}}$ Is standard deviation of variable x’

The inverse Box-Cox method is then applied to transform the outputs.

To test the effect of adjusting the distribution the model was run with both untransformed data and transformed data. Figures 3.6 and 3.7 show plots of observed and predicted SPM using untransformed and transformed data respectively.

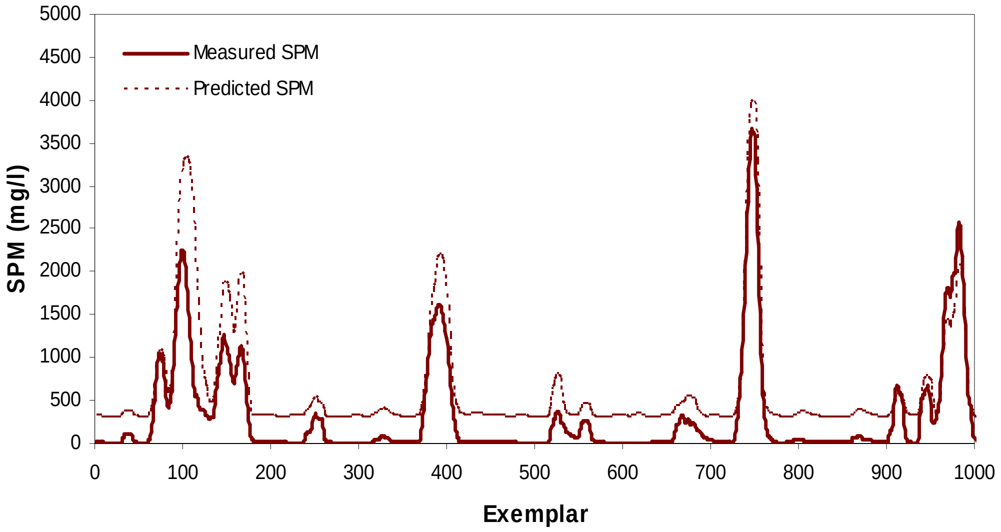


Figure 3.6 plot of desired and predicted SPM using untransformed data

It is observed that with untransformed data, the ANN consistently overestimates small values of SPM. This drawback is significantly alimented when transformed data is used. Transforming the data greatly enhances the models prediction of small values. The coefficient of correlation is improved from 0.8707 to 0.9177. By transforming the data max information is exposed to the network thereby improving its predicative capacity.

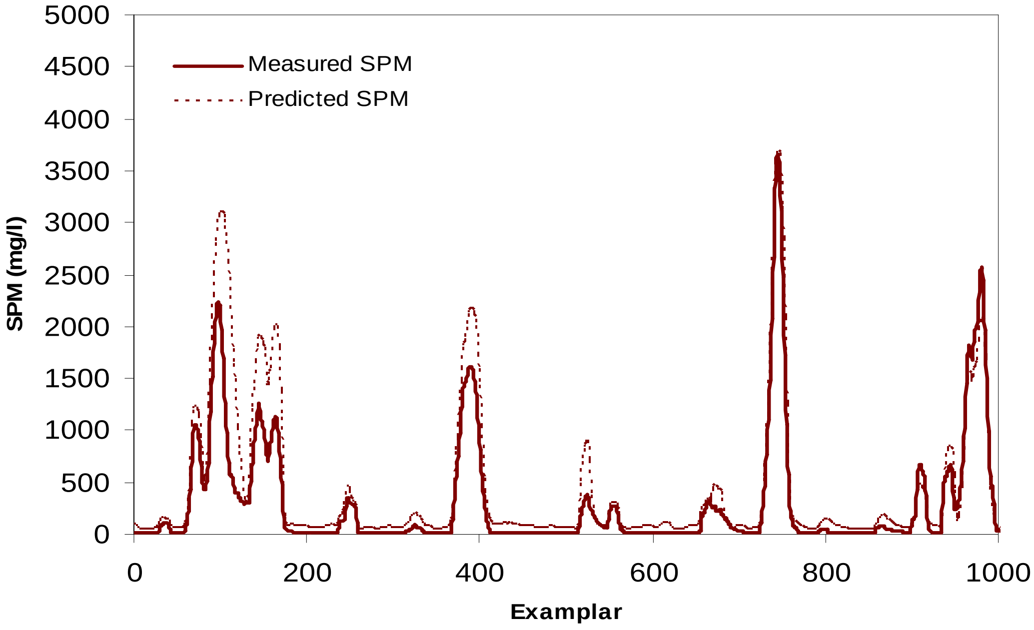


Figure 3.7 Plot of desired and predicted SPM using transformed data

3.5 Discussion

This block of the study has presented manipulation operations to maximally expose the information contained in the data to the modelling tool. A framework of these techniques to ensure that the training, testing and execution data sets are manipulated through different data preparation techniques before they are fed into the model has been referred to as the Information Prepared Environment (IPE) –Input. Given that the predicted variables are predictions of the manipulated data it is obvious that a reverse manipulation has to be employed to transform the predicted values back into the normal data range before data preparation. This is accomplished by the IPE-Output, which is the reverse of the IPE-Input. The IPE developed acts as a buffer to protect the modelling tool from distorted data.

When building a DDM the data selected for building the model is sampled from a large population. We cannot ascertain 100% that the whole data range (maximum and minimum values of the population) has been captured. Therefore the confidence level is always less than 100. Even with a confidence level of 95% there is a likelihood that a considerable number of out of range values could lay in the remaining 5% section for an execution dataset. The problem is even worsened by the fact that in such studies only a few years' data can be used to build the model and transferability of the model is further complicated.

One way to overcome this could be to apply a novelty measure to test the similarity of the execution dataset with the data that was used to train the model. If the statistical parameters of both data sets do not significantly differ, then the new data set should be used without any manipulation otherwise the new data set could be transformed to the same distribution as the training dataset. This could still be employed in the IPE modules.

3.6 Conclusions

- Comparison of the SPM computed by the numerical model and that predicted by the DDM showed a good fit with a correlation coefficient of 0.9177. However in a few cases the SPM was significantly underestimated.

- On the whole this chapter has indicted a strong possibility of approximating a numerical model for SPM using generated variables/parameters of the FLOW model.

- Using one parameter alone such as TBSS may not give satisfactory results but using more parameters could enhance the model's performance.

- In order to make the model more robust in regard to transferability a novelty measure of the execution data against the train data set should be performed. If significant difference is found transformation should be performed to conform to the new dataset to the same distribution as the training set.

4 Performance of time varying boundary conditions in the simulation of sediment transport

An investigation was conducted on the effect of specifying time varying open boundary condition for sediment concentration instead of fixed boundary conditions on the simulation of suspended particulate matter (SPM) along the Dutch coast. The aims were; a) to develop a methodology for generating time varying open boundary conditions with a data driven model using local hydrodynamic and meteorological variables as inputs, b) to study the performance of the time varying open boundary conditions on the simulation results. The performance of the boundary conditions was verified against observed data.

4.1 Introduction

In modelling, all problems with infinitely large region must be mapped to a finite domain for numerical computation. Since reducing the number of computational cells is an inexpensive way of saving simulation time, quite often a section of the interest area is cut out from the larger problem domain. Boundary conditions are then used to interface between an infinitely large problem space and a finite computation domain.

It is important that correct boundary conditions are specified and should be consistent with the physical description of the system. Marchesiello et al (2001) argues that it may not be necessary to know the solution of the numerical model within the interior of the numerical domain. But to implement physically realistic and sufficient boundary conditions requires a strong understating of the model interior and exterior of the phenomena and how they interact. Specifying conditions for the open boundary is usually complicated. Large-scale hydrodynamic simulations are normally run with a nested local grid. Simulation results from the large scale coarse model are then used to specify the open water boundary conditions for the nested grid (Schwab *et al.*, 2006).

In modelling of sediment transport in the Dutch coast, Wang (2007) implemented a RIJMAMO fine local grid model whose open boundary conditions were extracted from the ZUNO – coarse grid model of the North Sea. However the sediment concentration open boundary conditions could not be obtained from the larger model. Fixed values of sediment concentration were used as boundary conditions along the three open water boundaries namely the south, west and northern sides of the Rijmamo model. A schematization of the RIJMAMO grid is shown in Figure 4.1.

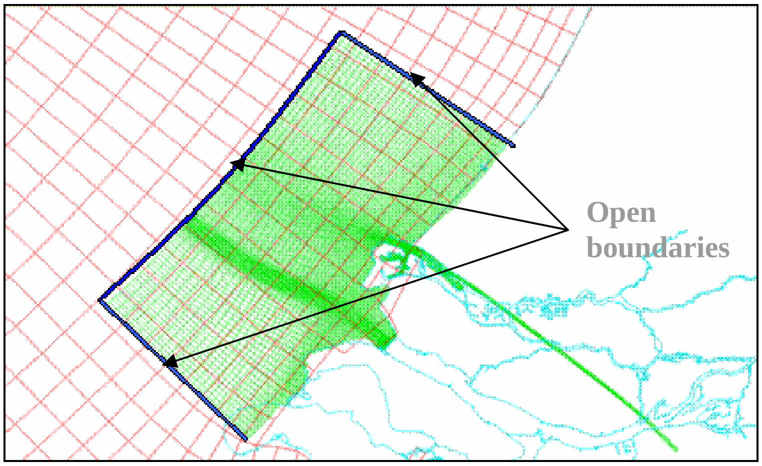


Figure 4.1: Schematization of the RIJMAMO g rid with its open boundaries

For the cross shore boundary the conditions were specified in such a way that two values were provided at both ends of the boundary line. A higher concentration was specified onshore and a lower value offshore. The concentration was linearly interpolated in between the two points. For the alongshore a constant sediment concentration was specified.

Considering that 80% of the sediments along the Dutch coast come from the English Channel (Chen and Eisma, 1995), It might be inappropriate to provide fixed SPM values at this boundary which does not take into account the time varying variability of sediment concentration. A fixed boundary condition is one with unchanging value in time irrespective of the changing hydro-meteorological conditions. Fixed boundary conditions do not seem to be realistic particularly for a domain with small length scales and therefore pose a potential source of low accuracy in simulation results.

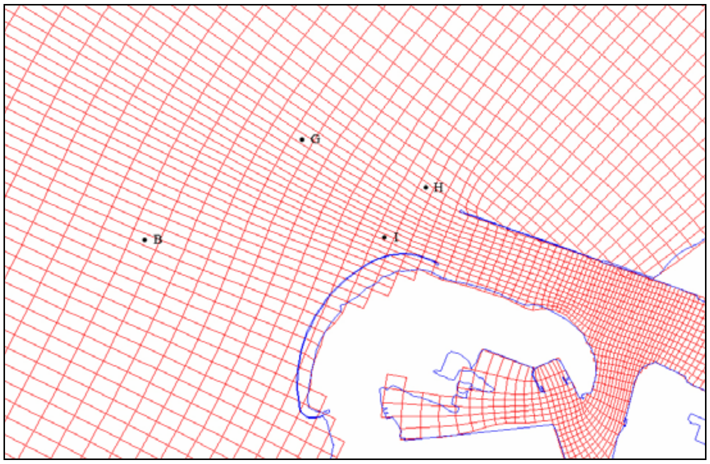


Figure 4.2 location of the SILTMAN observation points (adapted from Wang 2007)

The study therefore seeks to test the hypothesis that specifying time varying conditions at the south boundary of the RIJMAMO model would improve sediment prediction along the Dutch coast. The RIJMAMO model schematisation as built by Wang (2007) is adapted in this study. The study specifically develops a methodology for generating SPM time varying open boundary conditions the south boundary since measured data is not available. The time varying boundary conditions are generated using a data driven model with local hydrodynamic and meteorological conditions as input variables.

Secondly it tests the performance of the time varying open boundary condition (OBC) on the sediment simulation results. The performance of the boundary conditions is verified against observed data at the four monitoring points that were used for collecting the SILTMAN data. The location of the points are shown in Figure 4.2

4.2 Materials and Methods

4.2.1 Study area

The Maasmond is one of the five major entrapment areas for sediment along the French-Belgium – Dutch coastline. most of the sediment coming originates from the British Channel and the French coast (Vuurens, 2001). Marine sediments form the largest proportion of the sediment load followed by fluvial sediments from the Alps and mountain ranges of northern France and Belgium. The sediment size ranges from 0.1mm to 0.4mm. The organic component ranges from 8 – 23 % with increasing silt concentration in the landward direction (Chen and Eisma, 1995; Leussen, 1994).

At Maasmond, the Rhine River discharges into the North Sea. The plume of the fresh water extends several kilometres into the sea creating a coast river. During mixing (of fresh water with saline sea water) a density gradient is created which subsequently leads to the creation of density driven currents. These contribute to the process of sedimentation. A detailed description of the study area is given in section 2.6.1

4.2.2 Overview of approach

In this work an Artificial Neural Network (ANN) DDM built at Noordwijk 10[†] was adapted. The model was developed under the framework of the Delft Cluster II project. The input variables consisted of total bed shear stress, significant wave height (H_{sig}) and wind data and measured SPM. All variables had a time step of one hour. The simulated variable was SPM. The bed shear stress was derived from ZUNO run, whereas the H_{sig} and the wind data were measured data at Noordwijk 10km for the year 2000. The training dataset consisted of 715 records whereas the testing data set consisted of 271 records. Given the limited number of records no cross validation data set was specified to the model.

The DDM model was then used to predict SPM at the seven ZUNO grid cells that correspond to the south boundary of the Rijmamo grid. During model transfer a cross-shore and alongshore correction factors were applied to the simulated results. The simulation results were verified against measured SPM values at Schouwen 10km, a monitoring station quite close to the south boundary. The generated SPM time series were then supplied to Delft3D instead of the fixed boundary conditions. Four digital monitoring stations in series were introduced in the RIJMAMO model domain to investigate the influence of the boundary conditions within the model domain. Model simulation results were compared with both measured data and predicted results with fixed sediment boundary.

[†] 'Noordwijk 10' refers to a station say Noordwijk ten kilometres from the shore

4.2.3 Schematisation of the ANN DDM model at Noordwijk 10Km

The ANN used in the study was composed of 3 layers with one input layer, one hidden layer and an output layer. The input layer consisted of 3 data input parameters namely total bed shear stress, H_{sig}, and wind speed.

The process of sedimentation in a coastal environment is quite complicated. It involves several forcing functions including tidal forces, turbulent forces, radiation stress, bottom friction forces, etc. Given the fact that models should be kept simple (Roelvink and Walstra, 2005b) it is vital that the most important forcing functions should be represented in a model.

The following sections therefore discusse the justification of the choice of input variables to the DDM in relation to the longshore sediment transport in the Dutch coast. The input variables are total bed shear stress, south easterly winds and significant wave height. Alongshore transport is emphasised due to the importance of the north easterly movement of sediment in the area under review.

Total bed shear stress

Before any sediment is transported particles have to be moved from their initial position on the bottom. Particle movement can only occur when the water movement is strong enough to lift or drag the particle. The point of initiation of motion occurs when the velocity exceeds a critical value or critical bed shear stress. The grains then move, roll or come into suspension. Thus the bed shear stress is an important parameter in sediment resuspension and it on this basis that it was selected as one of the input variables.

The bed shear stress responsible for the particle movement is a combination of currents (due to wind, wave, tidal effect and density variation) and wave effects. To obtain the total bed shear stress, Bijker (1967) theoretically added the velocity due to currents and waves in a vector form. With several manipulations he established a relationship between the total bed shear stress, current shear stress and the wave shear stress. Equation 4.1 gives the above relationship mathematically.

$$\overline{\tau_{cw}} = \tau_c + \tfrac{1}{2}\hat{\tau}_w \qquad 4.1$$

where

$\overline{\tau_{cw}}$ = time - averaged value of total bottom shear stress

τ_c = Current shear stress

$\hat{\tau}_w$ = maximum wave shear stress

The individual shear stresses τ_c and $\hat{\tau}_w$ are denoted as

$$\tau_c = \rho\kappa^2 \mathrm{V_t}^2 \qquad 4.2$$

$$\hat{\tau}_w = \rho\kappa^2 (\mathrm{p\hat{u}_o})^2 \qquad 4.3$$

where

$\mathrm{V_t}$ = current velocity at elevetion $\mathrm{z_t}$
ρ = water density
κ = Von Karman coefficient (0.4)
$\mathrm{\hat{u}_o}$ = maximum velocity just outside the boundary layer
p = proportionality factor

To stir up sediments only the critical velocity need to be exceed regard less of the direction thus for the seek of stirring up sediments into the water column the time averaged value of total bottom shear stress is used (Velden, 1995). The total bed shear used in the DDM was computed as in equation 5.1. The wave induced and current induced bed shear stress where both locally extracted from ZUNO run results files. For the model at Noordwijk 10km, results for the 2000 model run were used whereas for the south boundary results for the 1996 model run were used.

Significant wave height

Sediment concentration distribution under wave action accounts for a great part of the total concentration in the water column (Booth *et al.*, 2000). In the presence of surfacial sediments, oscillating motion due to wave action causes ripples on the bed. Vortices are formed in the lee of these ripples. They are then thrown up with the reversing flow. The vortices cause a high concentration layer of a few vortices thick. With the ensuing orbital motion, the sediment is transported further into the vertical direction. These two mechanisms (vortices and orbital motion) lead to sediment concentration distribution due to waves to be of the order of 5 to 6 times that due to currents (Velden, 1995). For any sediment modelling it is important that the wave effect is well captured. Apparently one of the input variables for a DDM should be a wave characteristic that realistically represents the wave effect.

Since the concern is with sediment crossing the south boundary in a sense we are looking at the alongshore sediment transport. Conclusions regarding the appropriate wave characteristic can easily be drawn from the established alongshore sediment transport relationships. The Coastal Engineering Research Centre model (Velden, 1995) predicts alongshore sediment transport by use of the wave energy. It describes the relationship between the total transport through a unit width of the total breaker zone and the longshore component of energy flux as per equation 5.4.

$$S_x = A'U' \qquad 4.4$$

Where

S_x = longshore sand transport
$\mathrm{A'}$ = dimensional coeficient
$\mathrm{U'}$ = the component of energy flux entering the breaker zone

With manipulation equation 5.4 can be expressed as

$$S_x = AH_{sigb}{}^2 n_b c_b \cos(\varphi_b)\sin(\varphi_b) \qquad 4.5$$

where

H_{sigb} = where Significant wave height

c_b = wave progartions speed,

φ_b = angle between the wave creest and the coast at breaker line

nc gives the wave group velocity

An import aspect of this formulation is in the way how it relates to the residual shear stress. The radiation shear stress provides the driving force for the longshore current within the breaker zone. From this formulation we zero in on the significant wave height as one of the input variable to the DDM.

Bhattacharya et al.(2007)used the square of the significant wave height to represent the wave energy assuming a dimensional coefficient of proportionality equalling to unity. They argued that small waves did not have significant effect on sedimentation process and thus took the square of the difference between the instantaneous significant wave height (H_{sig_t}) the average significant wave height ($\overline{H}_{sig}$).This treatment of the significant wave height seem to be more appropriate. However the study did not use the square of significant wave height on the understanding that since the ANN DDM can pick up non linear relationship, then this manipulation is not important to the model. The absolute value of significant wave height was used in the DDM

For the model at Noordwijk 10 km observed significant wave height for the period 2000 was used. At the south boundary significant wave height were generated by interpolating results of the wave model (SWAN) in and observed measurements at 11 monitoring station for the year 1996. This interpolation was implemented in Matlab. The significant wave height was smoothed using a seven-day moving average

South westerly wind

Along the Dutch coast in general the water movement is in equilibrium, with a net transport in the north eastern direction (Postma, 1981). This effect is responsible for carrying sediment from the English Channel through the Dutch coast. The equilibrium of a water mass moving along the coast is determined by force acting parallel to the coastline(Velden, 1995). This force is as a result of changes in the radiation shear stress component acting in the same direction. The radiation shear stresses are due to waves approaching the coast line at an angle. The radiation shear stress is related to the waves through the wave energy expressed as

$$S_{yx} = En\sin(\varphi)\cos(\varphi) \tag{4.6}$$

where

E = wave energy,

φ = angle the wave crest makes with the depth contours parallel to the coast,

$$E = \tfrac{1}{16}\rho g H^2 \tag{4.7}$$

But the wave height is closely related to the wind speed. The Beach Erosion Board (CERC, 1984) established empirical model relating wind speed and wave height give as

$$\frac{gH}{w^2} = 0.0026\left(\frac{gF}{w^2}\right)^{0.47} \tag{4.8}$$

where

g = gravity due to acceleration,

w = wind speed,

F = effectiv fetch

H = water depth

From this relation we chose to use the wind component in the direction of the radiation shear stress was chosen as one of the DDM input variable to depict the effect of this stress in the alongshore sediment transport processes. The wind component was obtained by resolving the wind in the south westerly direction. Wind data for 2000 collected at Noordwijk was used to build the model at Noordwijk. At the south boundary wind data for the year 1996 collected at station 312 Oosterschelde (http://www.knmi.nl/samenw/hydra) was used as the production dataset during model application. The same wind speed was used for all the south boundary grid cells.

4.2.4 Applying the DDM and transfer correction factors

Besides meteorological and hydrodynamic conditions, local morphological conditions have an influence on the sediment processes. It is conceivable that a model built using local hydrodynamic conditions at one point in the system may require some adjustments to be applied at a different point within the same system. For the model developed at Noordwijk 10 km the adjustments entailed data manipulation and application of longshore and cross-shore correction factors. The manipulation and development of the correction factors are discussed in the following sections.

Data transformation to equalize the input data range

From the Atlas of the Near-Surface Total Suspended Mater Concentrations in the Dutch coastal zone (Rijkswaterstaat, 2002) it was noted that sediment concentration reduced gradually in the offshore direction. This would imply that the bed shear stress varies in the same manner. Indeed comparison of the average bed shear stress at Noordwijk 1km from the shore was 4 fold the average bed shear stress at Noordwijk 10 km. This meant that applicability of the model was impossible because the application dataset values would be far out of range particularly at the near shore grid cells.

To overcome this problem the input values for the model were all normalized to a zero mean and unit variance. The operation made the training, testing and production data set become comparable in magnitude

Along shore correction factor

Having run the model with normalized values it was noted that much as the production dataset at the grid cell along the south boundary had comparable statistical parameters as the training and testing dataset at Noordwijk 10 km (Figure 4.3), the predicted SPM at the two points was not comparable as seen in the Figure 4.4. Figure 4.5 shows the ZUNO grid cells that correspond to the south boundary of the RIJMAMO grid.

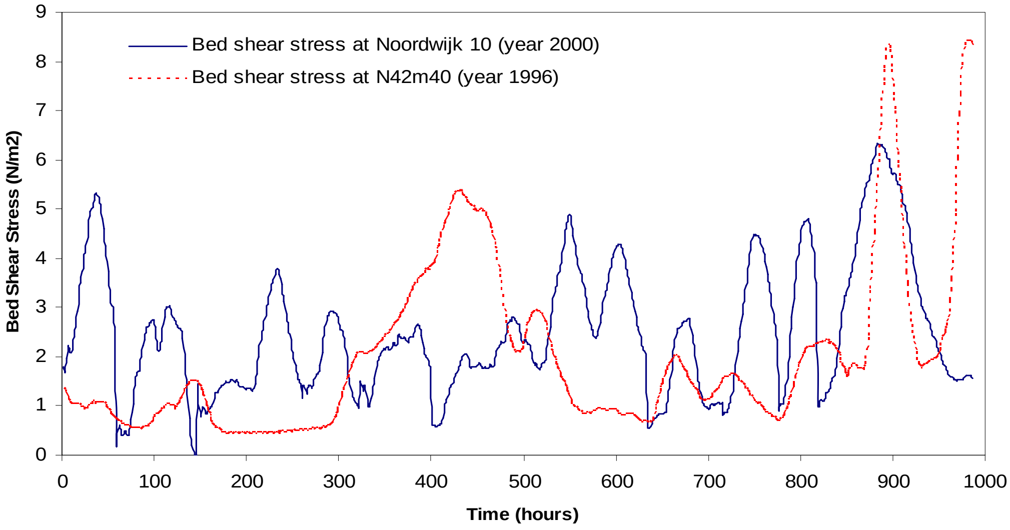


Figure 4.3Comparison of bed shear stress at Noordwijk and south boundary

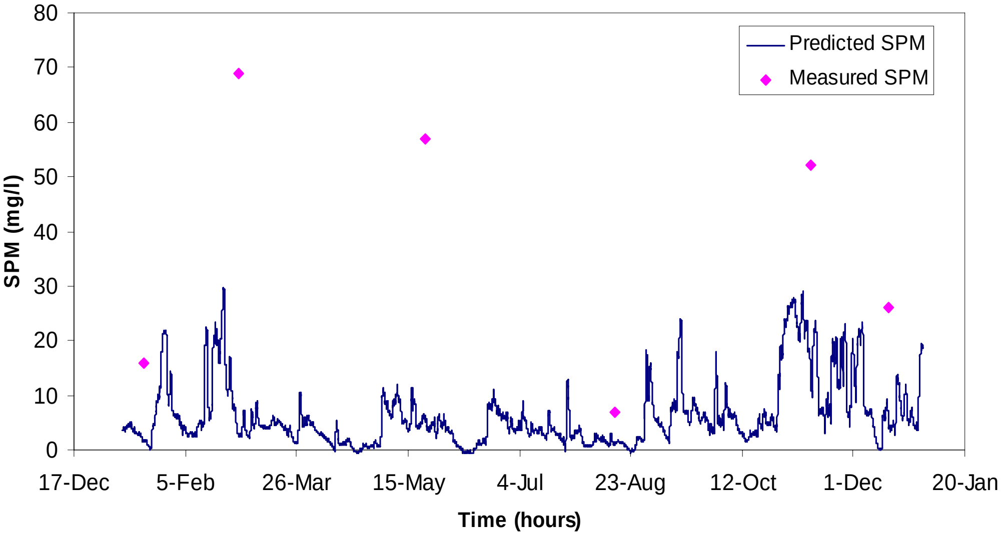


Figure 4.4 Predicted and observed SPM at the south boundary

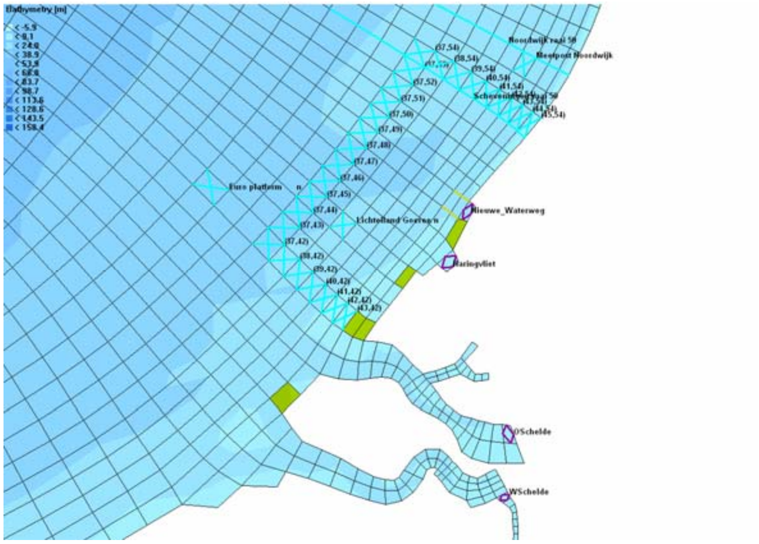


Figure 4.5 Location of the ZUNO grids coinciding with the RIJMAMO grid

Assessment of observed records and the Atlas of total suspended particulate matter (Rijkswaterstaat, 2002) revealed that there existed alongshore SPM gradient that necessitated applying a correction factor. The correction factor was computed as the ratio of average SPM values measured at Schouwen to the average SPM concentration at Noordwijk 10km for the period 1973 –1985 (http://www.waterbase.nl). Only days on which measurements were taken at either stations or time between measurements not exceeding 24hrs were considered. 160 Records in total were used in this operation.

Finally a correction factor of 2.65 was arrived at. Applying this factor to the DDM's prediction gave a satisfactory fit as seen in the Figure 4.6. It should be noted that Schouwen does not lie exactly on the south boundary but it is approximately 5kms from the south boundary. But its records were used since it is the only station with records close to the south boundary. It is assumed that the difference in concentration in relation to the distance is insignificant.

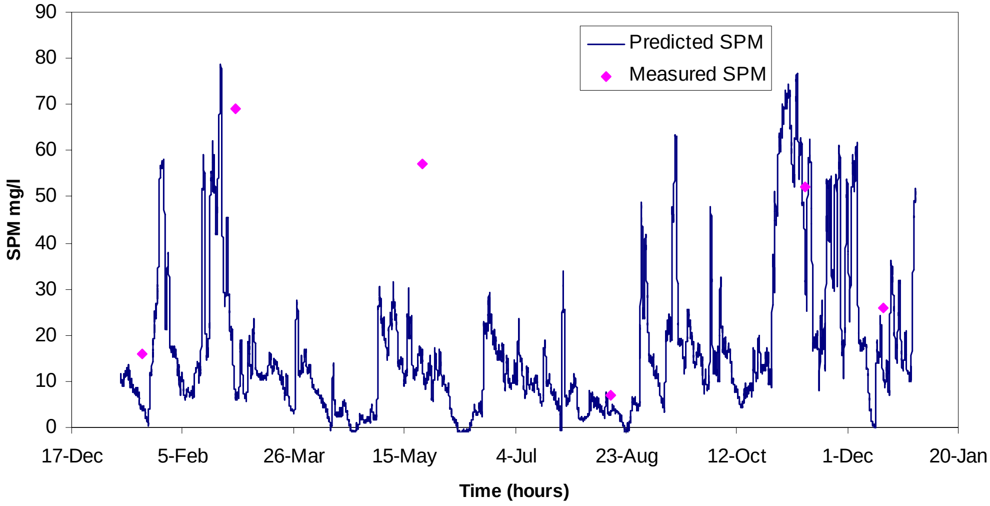


Figure 4.6 Comparison of measured and predicted SPM at the south boundary after applying the alongshore correction factor

:

Across shore correction factor

Similarly, across-shore gradient in SPM concentration exists in the Dutch coast. It is important that while applying the model to the other locations of the south boundary other than the point that is inline with Noordwijk 10 km, across-shore correction factor should be applied as well. This cross-shore correction factor was computed in a similar manner as the alongshore correction factor. However in this case 4 stations were used at the Schouwen transect. The results were related to depth in the cross-shore direction other than the across-shore distance. A second order polynomial function was fitted to the points to generate the cross-shore SPM-depth relationship expressed in equation 5.9.

$$\eta_{cs} = 0.0098h^2 - 0.0518h + 0.7263 \qquad 4.9$$

where

η_{cs} = cross shore correction factor

h = water depth

The resultant coefficient was then multiplied by the SPM concentration as predicted by the model using the production dataset. Figure 4.7 shows the across-shore correction factor with respect to depth.

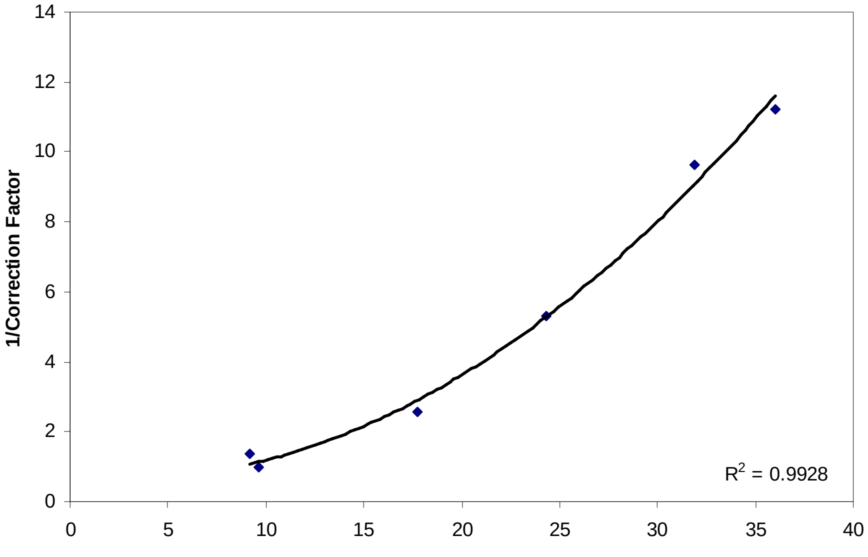


Figure 4.7Across-shore correction factor with respect to depth

By running the DDM and applying both the alongshore and across-shore correction factors to the DDM simulations, SPM time series were generated for the seven grid cells that correspond to the south boundary of the RIJMAMO model.. The time series where then supplied to the RIJMAMO local model instead of the fixed boundary conditions. The model was then run for the period 28th Jan 1996 to 16 March 1996 with the same model parameters as in the fixed boundary model. Figure 4.8 shows SPM time series generated for some of the grid cells.

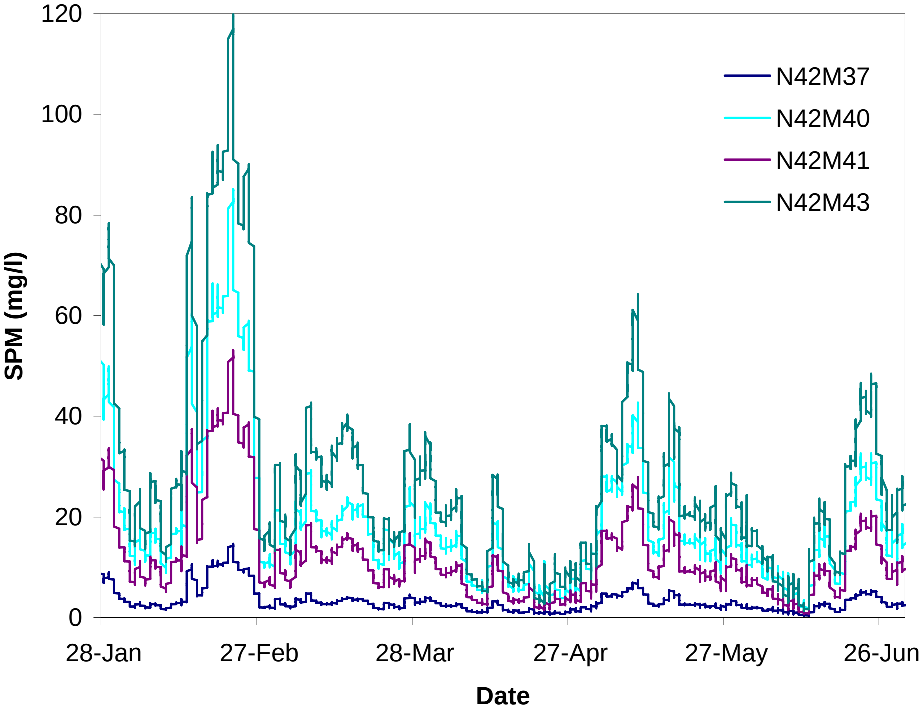


Figure 4.8: Generated time series for some of the ZUNO grid cells corresponding to the south boundary of RIJMAMO grid

Investigation of error propagation with respect to distance form south boundary

In numerical modelling the boundary data given at the point should not influence the results of the model computation under different scenarios (Verwey, 2005). But this principle may not hold in the case of water quality modelling in general and sediment transport modelling in particular. If the conditions specified at the boundary should have an effect on the outcome of the model results then effort should be made to specify as correct as possible the values at the boundary. Otherwise the location of the boundary should be located at a distance sufficient enough not to have effect on the model outcome.

In a river a good measure of the sufficient distance is the length of the back water curve. However there is no clear methodology in the case of the coastal system. In such a circumstance Verwey (2005) recommends a sensitivity analysis on the effect of an error in the chosen boundary conditions

To investigate the effects of an error in the boundary condition on the model outcome with respect to distance in the alongshore direction, four ‘history’ points were introduced in series in the RIJMAMO domain. History points refer to points within the model domain at which modelled flow variables can be inspected with respect to time. Figure 4.9 shows the location of the observation points whereas the Table 4.2 gives the distance of the observation point from the south boundary.

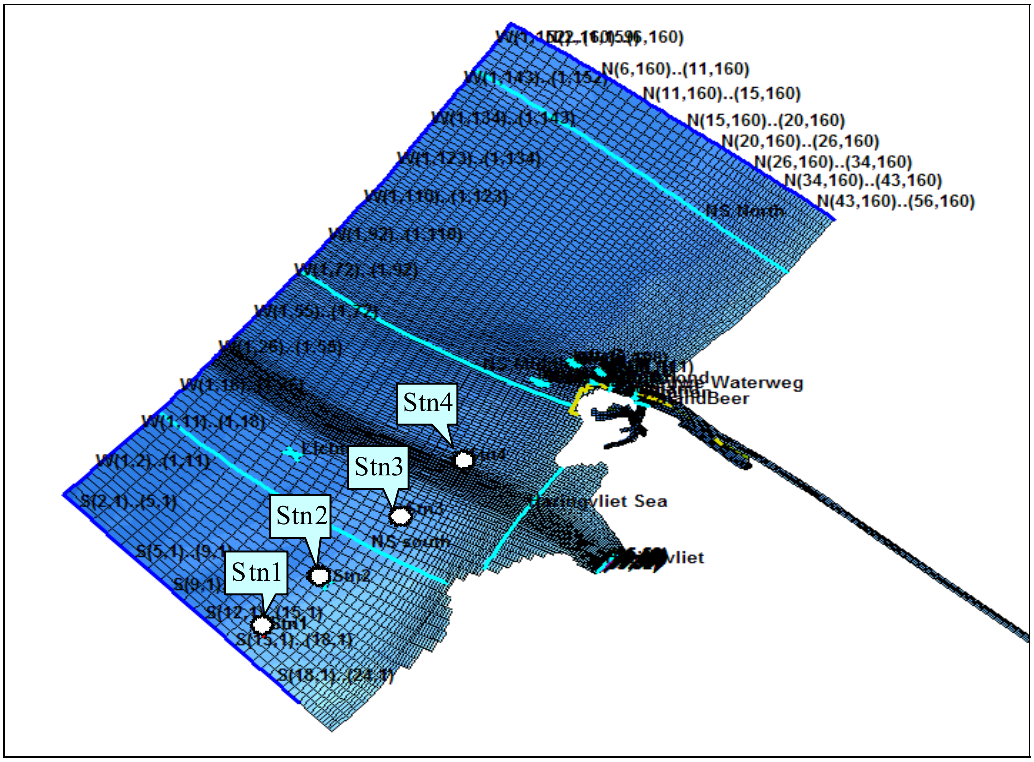


Figure 4.9: Location of observation points from the south boundary

Table 4.1: Distance of the observation points from the south boundary

	Observation points			
	Stn[‡] 1	Stn 2	Stn 3	Stn 4
Distance from south boundary (km)	2.8	9.4	18.2	26.0

For this operation the model was run for the period 28th Jan 1996 to 10th February 1996 with both time depended and fixed boundary values the same model parameters.

4.3 Results and discussion

In this investigation it was assumed that the time depended sediment concentration were the correct boundary conditions. Results of the fixed boundary values were compared with those of the time varying results and errors were computed accordingly.

4.3.1 Comparison of observed SPM and predicted SPM

Results show that using the time series for the SPM boundary conditions only affects the amplitude of predicted SPM but does not necessarily change the trend. Perhaps this is due to the fact that even when time varying boundary conditions are specified, the magnitudes of hydrodynamic variables such as current velocity, bed shear stress, etc still remain unchanged. Sediment resuspension and transport mechanism remain the same and consequently the prediction trend stays unchanged irrespective of the

---

[‡] Stn refers to the history points in the computation domain

variability in the values specified at the boundary. Figure 4.10 and 4.11 show plots of observed SPM and SPM as predicted using time varying boundary and fixed boundary at SILTMAN points B and H respectively.

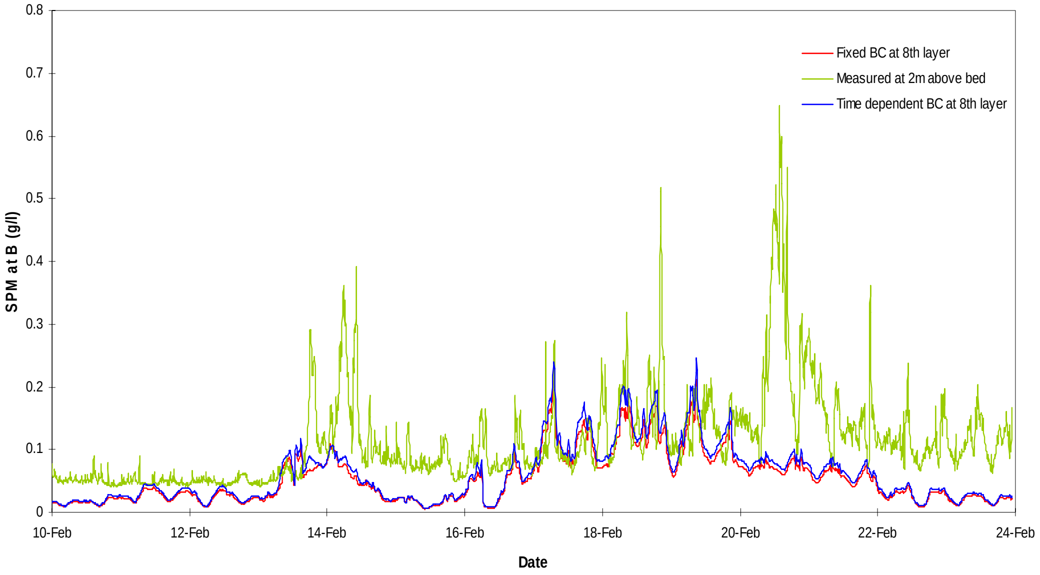


Figure 4.10: Plot of observed SPM and simulated SPM using time varying boundary and fixed boundary condition at SILTMAN point B

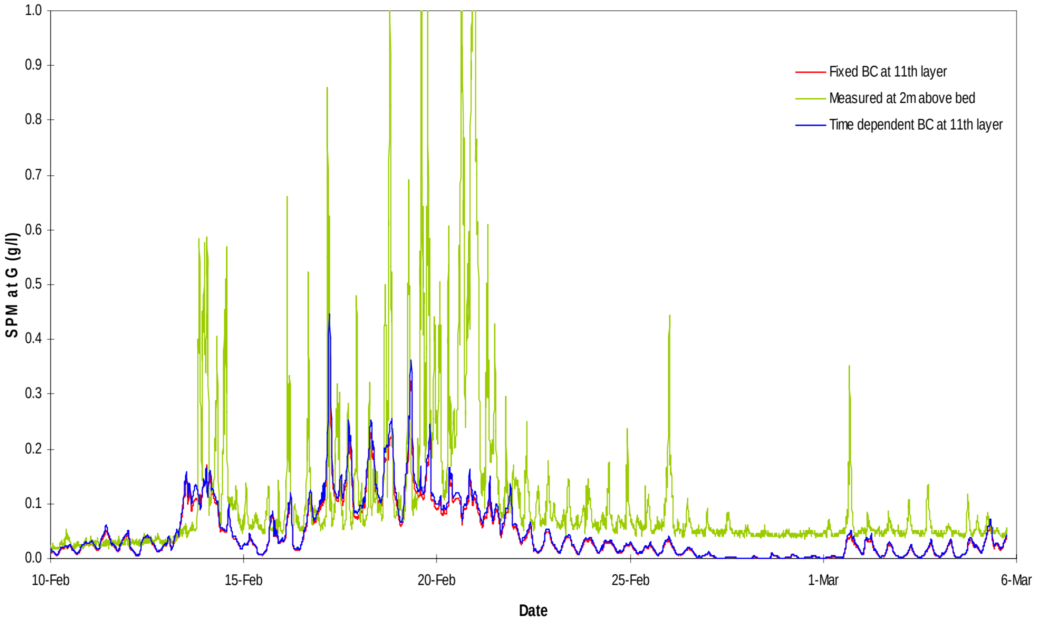


Figure 4.11: Plot of observed SPM and simulated SPM using time varying boundary and fixed boundary condition at SILTMAN point G

The fact that the difference between the model outcome of the two approaches (time varying and fixed sediment boundary conditions) is insignificant could imply that in the short term the local sediment supply is more important in computing the sediment concentration along the water column. The reference concentration or in some approaches referred to as the bottom concentration is greatly dependent on the local availability of surfacial sediments besides the fluid velocity and shear stress in the bottom layer. Thus the inference on the importance of local sediment supply is in agreement with sediment transport formulae that have emphasised the importance of the reference concentration in the computation of suspended sediments (Bijker, 1967; van Rijn, 1993).. The formulations stipulate that suspended and bed load sediment is directly proportional to reference concentration.

A time varying sediment boundary condition at the south boundary could be more useful if the objective is to predict long-term morphological changes. In this case an accurate sediment budget is needed to establish whether a net deposition or erosion would take place in the study area. It is important that the horizontal sediment transport is well schematised to give accurate predictions especially if a morphological factor is applied.

Considering that the equilibrium sediment transport is in the north eastern direction as a result of radiation shear stress. the sediments will always be moving out of the computation domain. In that respect the north boundary condition is rendered redundant thus a weak boundary condition can be considered at the north open boundary. No sediment concentrations need to be specified at this open boundary.

4.3.2 Change with distance from the boundary layer

The absence of a significant difference between the simulated results with fixed boundary conditions and time varying boundary conditions could also imply that the observation points of the SILTMAN data are far enough from the south boundary to be affected by the conditions supplied at this boundary. But this may not be the case with points that are close to the south boundary.

Investigation on the influence of a boundary condition on the model outcome with respect to distance in the alongshore direction yielded plots in figure 4.12 through 4.15. The plots show simulated sediment concentration using fixed and time varying sediment boundary condition at four points in series within the model domain as schematized in section 4.4.4.

From the plots it is notable that at history point Stn 1, the variation between the two plots is quite significant. However this variation gradually fades out at the subsequent stations. Of course this outcome is expected but the most important aspect of this observation is the rate of reduction in error. A relationship between the error propagation with respect to distance from the boundary can be established accordingly.

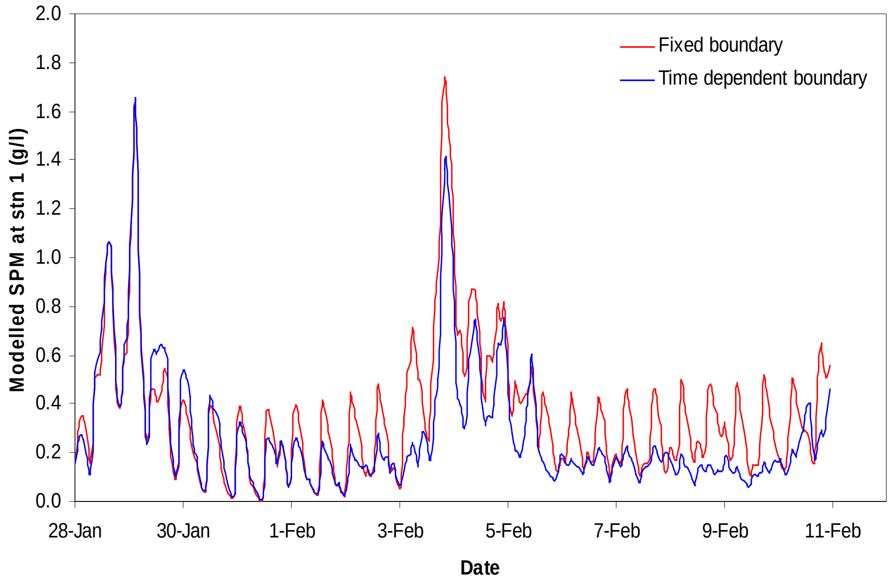


Figure 4.12: Plot of predicted SPM using fixed and time varying boundary conditions at observation point stn1

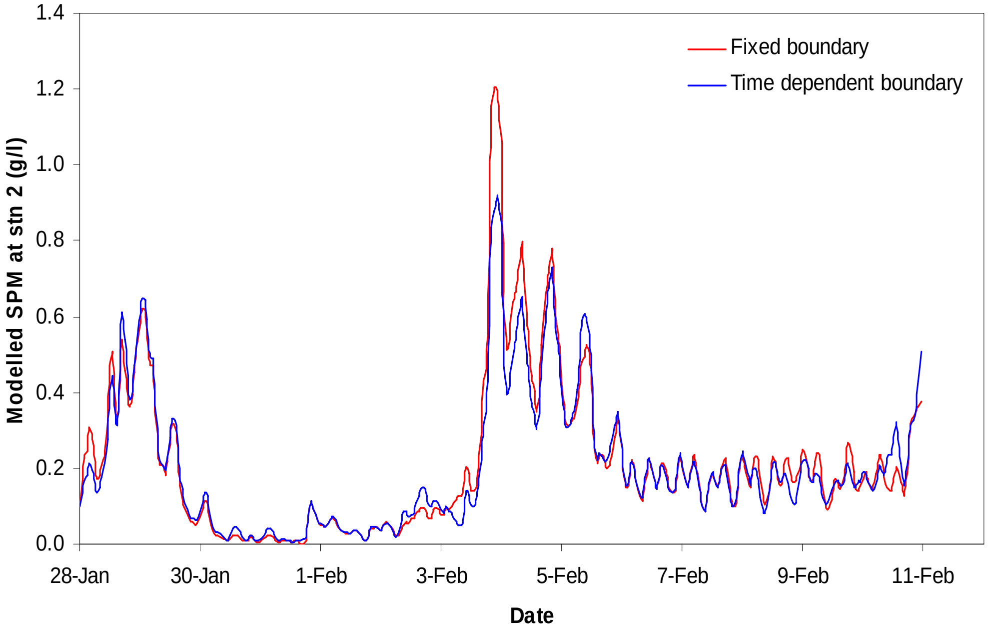


Figure 4.13: Plot of predicted SPM using fixed and time varying boundary conditions at observation point stn2

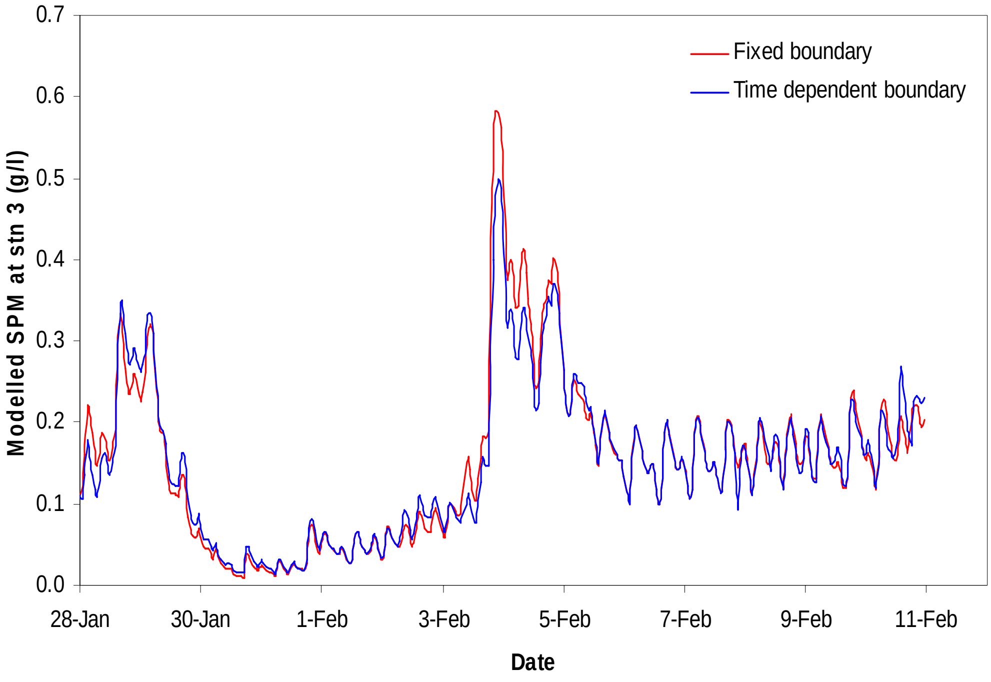


Figure 4.14: Plot of predicted SPM using fixed and time varying boundary conditions at observation point stn3

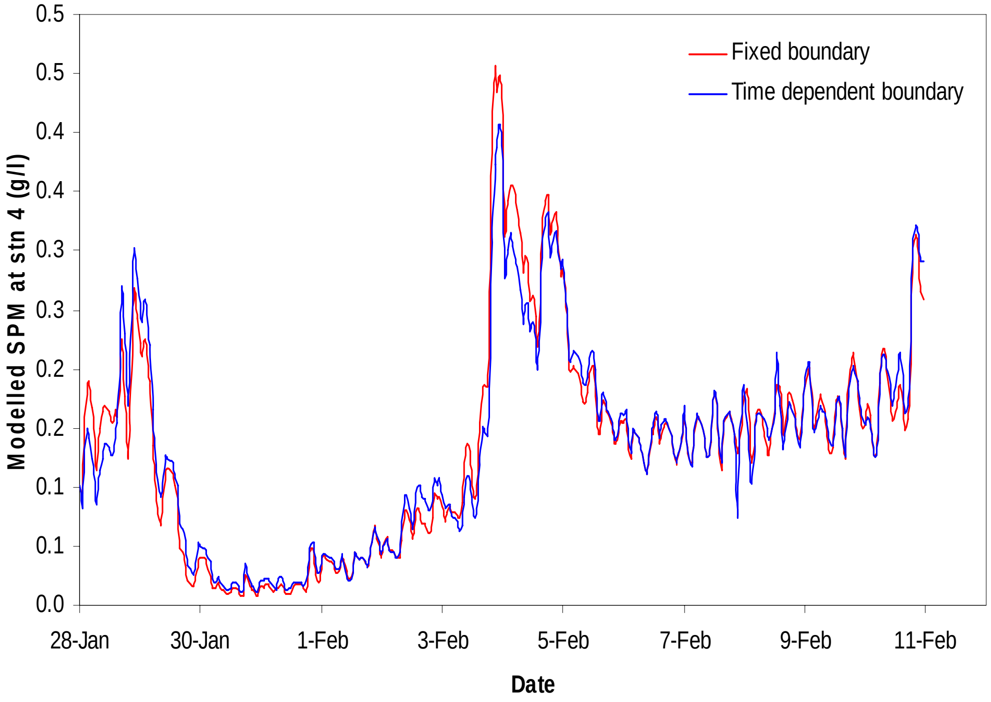


Figure 4.15: Plot of predicted SPM using fixed and time varying boundary conditions at observation point stn4

It was assumed that the time varying sediment concentrations were the correct boundary conditions. The error was computed as the difference in concentration by using fixed and time varying boundary conditions. Table 4.2 gives the computed Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE)

Table 4.2: Error analysis of time varying and fixed boundary predicted SPM at four monitoring points

	Observation points			
Error	Stn 1	Stn 2	Stn 3	Stn 4
Mean Absolute Error (g/l)	0.114	0.030	0.014	0.013
Root Mean Square Error (g/l)	0.156	0.054	0.023	0.019

A plot of both the MAE and RMSE verses distance (figure 4.16) shows that the error between results obtained by time varying boundary and the fixed boundary reduces exponentially with increasing distance from the boundary. Equations 4.10 and 4.11 give the MAE and RMSE error reduction function with respect to distance.

$$\varepsilon_{MAE} = 0.4247 d^{-0.9618} \qquad 4.10$$

$$\varepsilon_{RMSE} = \qquad 4.11$$

where

ε_{MAE} = Mean Absolute Error

ε_{RMSE} = Root Mean Square Error

d = distance from the south boundary

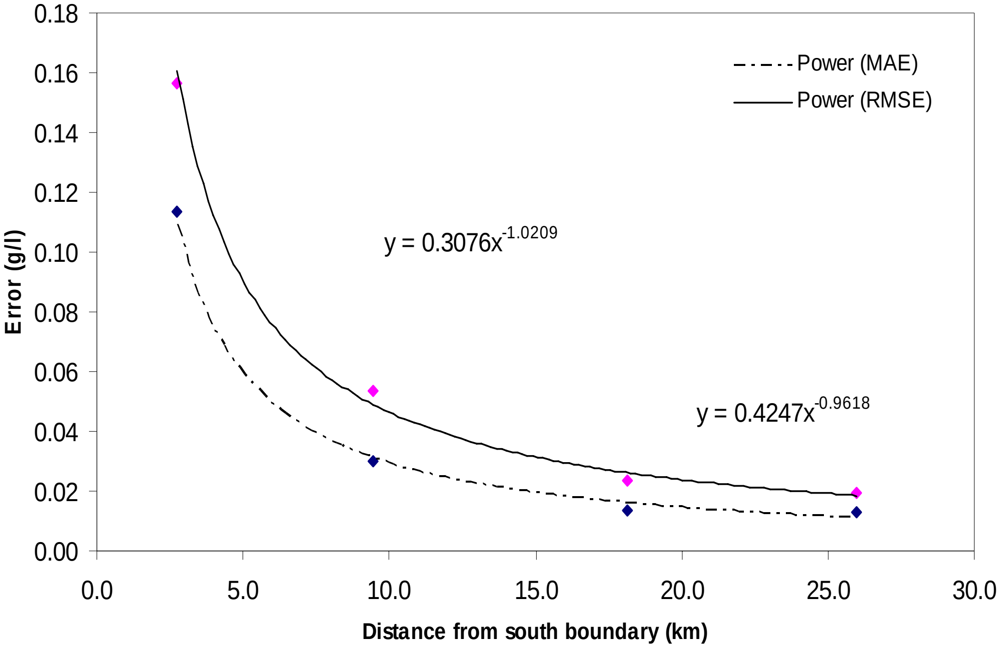


Figure 4.16: Variation of the Mean Absolute Error (MAE) and Root Mean Square Error with distance from the south boundary

4.3.3 Response of sediment to major forcing function

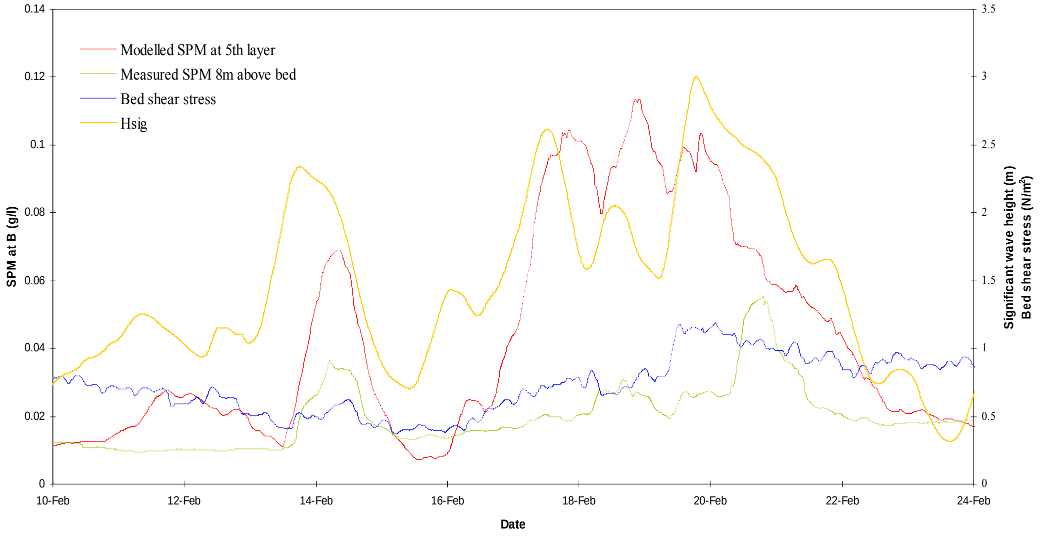
In general the model depicts the SPM variation though in some cases there are big discrepancies that can not be explained by even the forcing function. In the figure 4.17 and 4.18 the variables are averaged over a 12hr moving average for easy interpretation of the results12 hours are used to filter out the small variations while capturing the effect of the tidal movement. Result show that the observed SPM seems to respond more to the H_{sig} than the total bed shear stress. This is contrary to the widely known fact that bed shear stress is the most important parameter in sediment entrainment (Bhattacharya *et al.*, 2007; van Rijn, 1993; Velden, 1995; Winterwerp, 1999).

The discrepancy could be explained by the fact that regardless of the bed shear stress the mechanism of bringing sediments in suspension will influence the sediment concentration in the water column. For example for the same magnitude of bed shear stress the amount of sediment entrained due to wave induced bed shear stress will be several orders of magnitude compared to the current induced bed shear stress (Postma, 1981)

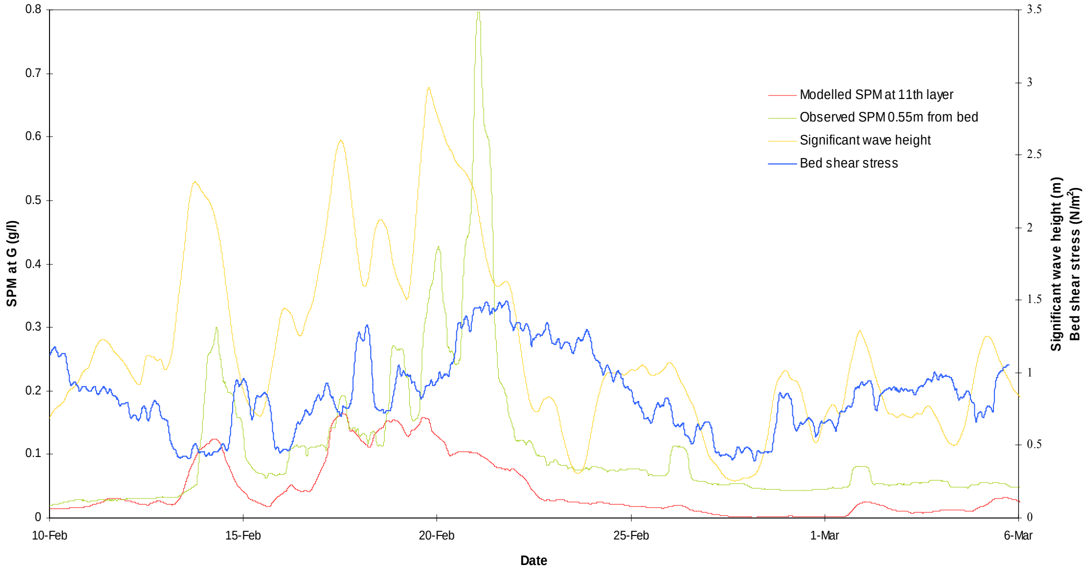
In case of wave action the oscillating motion causes ripples on the bed. Vortices are formed in the lee of these ripples. They are then thrown up with the reversing flow. The vortices cause a high concentration layer of a few vortices thick. With the ensuing orbital motion, the sediment is transported further into the vertical direction. These two mechanism lead to sediment concentration distribution due to waves to be of the order of 5 to 6 times that due to currents (Velden, 1995).

The above revelation could mean that using bed shear stress alone in DDM to predict SPM is quite less accurate. Secondly, that the significant wave height is an equally important variable in sediment prediction. Concurrent use of both variables (H_{sig} and total bed shear stress) could greatly any enhance the model's predictive capacity.

From the plots of observed SPM and forcing functions shown in figures 5.11 and 5.12, it is also noticed that in some cases the observed SPM does not to respond to any of the sediment transport forcing factors. But otherwise behaves in a stochastic manner. There is no clear explanation for this kind of behaviour. Perhaps the problem lies with the observed data. To establish whether the model is at fault or the data is in error the measured and observed significant wave height and the observed SPM should be compared in the same plot. The comparison not possible for the area under study for the reason that there were no H_{sig} data recorded at the SILTMAN data points.



4.17 comparison of SPM and hydrodynamic variables at SILTMAN point B



4.18 comparison of SPM and hydrodynamic variables at SILTMAN point G

4.4 Conclusion

In the preceding sections we have investigated the effect of specifying time varying open boundary condition instead of fixed boundary conditions on the simulation of suspended particulate matter (SPM) along the Dutch coast. The following are conclusions drawn from the results.

- The total bed shear stress, significant wave height and south-easterly winds as local hydrodynamic and meteorological input variables tot the DDM have shown satisfactory performance in the simulation of SPM.

- While applying the data-driven model it is necessary to apply both alongshore and cross-shore correction factors.

- Results have shown that nature of the sediment concentration boundary condition at the south boundary (either time varying or fixed) has no significant effect on the SILTMAN points that were used in the analysis. If the SILTMAN points are the focus of the study then they sufficiently located far enough not to be affected by the specified conditions.

- The error in the choice of the boundary reduces exponentially in the long shore direction.

- The choice of the south boundary i.e. whether fixed or time varying may have no effect on the model results when short simulation times are considered. However the choice might be important if the aim is to model large time scale processes such as morphological changes.

- Using the developed relationship of error reduction and propagation within the model domain, the size of the model domain may be reduced considerably depending on the location of the study area and the acceptable error in the model simulation. This would considerably reduce the computation time.

5 Modeling of sediment transport in Lake Victoria

This chapter examines how to transfer knowledge gained in modelling sediment transport along the Dutch coast to understand sediment transportation in Lake Victoria. The study lays ground for future more accurate and reliable modelling of sediment transport in Lake Victoria through transfer of methods developed in the preceding sections. This was achieved by identifying the most important forcing functions in sediment transport and resuspension in Lake Victoria. An assessment of the adequacy of existing data for sediment modelling of the lake has been carried out. The study also recommends parameters that should be collected and also suggests other data collection techniques other than the conventional methods

5.1 Introduction

Sediment deposition and resuspension has considerable influence on eutrophication of lakes, streams, estuaries, and coastal waters. Chemicals favouring eutrophication conditions such as phosphorus and ammonia are transported with sediment in an adsorbed state. Studies on chemical state and nutrient content of sediment as well as laboratory simulation of sediment release of nutrients have shown that sediment can be a source for release, or a sink for adsorption of nutrients under certain conditions (Krumbein, 1936).

Changes in the aquatic environment such as the development of anaerobic conditions in the bottom sediments can cause the release of the adsorbed nutrients from the sediment. Adsorbed phosphorus transported by the sediment may not be immediately available for aquatic plant growth but does serve as a long-term contributor to eutrophication (US Environmental Agency, 2007). From the above understanding it is apparent that there is a relationship between SPM and eutrophication. Understanding the dynamics of sediment transport can be used to predict eutrophication and to draw interventions accordingly.

A preliminary modelling framework for Lake Victoria was developed by WL| Delft Hydraulics in conjunction with HydroQual Inc. The frame work was meant to be used as a tool for water quality management of the lake to promote understanding of the interaction of the main lake processes as a basis for developing technological responses, and guiding water management decisions. However the data requirements for the water quality forcing are quite intensive. These include air temperature, atmospheric pressure, precipitation, evaporation, relative humidity, solar radiation, cloud cover, wind speed and wind direction in addition to (bio)chemical loading. With the challenges facing data collection in Lake Victoria satisfactory water quality modelling of Lake Victoria is and still likely to remain a big challenge in the near future (WL| Delft Hydraulics *et al.*, 1999).

In light of the aforementioned challenges it follows that the simplest way to predict eutrophication in Lake Victoria is by relating it to sediment transportation and resuspension. The objective of this section is therefore to explore how knowledge gained in modelling sediment transportation along Dutch Coast can be applied to Lake Victoria. The study examines similarities and differences between the two systems on the basis of which it recommends a modelling approach. The Dutch Coast was chosen because of the vast expertise gained from sediment modelling activities there over the years. This work was intended to lay ground for more accurate sediment modelling in Lake Victoria in Future.

5.2 Materials and methods

5.2.1 Study area

Lake Victoria is the world's second largest freshwater lake and the largest in Africa with a surface area of 68,800 Km^2. Temperature range in the lake is narrow and warm with a mean annual temperature of 25^{o} C. The shoreline is about 3,500 Km. It encloses a number of small, shallow bays and inlets, many of which include swamps and wetlands that differ a great deal from one another and from the lake itself.

Lake Victoria basin provides an immense economic potential to its estimated population of over 33 million people. Over the past four decades or so, the lake has come under increasing pressure from a variety of interlinked human activities such as industrial pollution, eutrophication, in addition to sediment loading (Hecky, R.E. 1993; Water Quality Synthesis Report, 2005).

Studies have shown that sediments in Lake Victoria produce an internal phosphorous load which is approximately equal to external loads on an annual basis. The Phosphorous load significantly influences algal growth in the lake (Moore *et al.*, 1998; Olia and Reddy, 1993). Understanding the dynamic mechanism of sediment resuspension is import into in drawing any intervention measures. A detailed description of the study area is given in section 2.6.2

5.2.2 Previous sedimentation modelling on Lake Victoria

The modelling framework for Lake Victoria implemented a sediment flux sub-model. The sub-model was intended to ensure that the sediment oxygen demand and nutrient fluxes bore a relationship for the delivery of organic materials from the water column to the surfacial sediments and vice versa (WL| Delft Hydraulics *et al.*, 1999). From the implementation it is noticeable that the sub model was not a dedicated model to accurately simulate sediment transportation, deposition and resuspension.

Azza (2006) conducted an investigation on the mechanisms of sediment distribution in Lake Victoria. He mainly examined the applicability of generalisations of sedimentation principles of small lakes and oceans to large shallow lakes. The work concentrated on getting the right sediment distribution patterns rather than the accurate sediment concentration in the water column on both the spatial and temporal scales.

If we are to predict eutrophication in Lake Victoria by inference on the sediments characteristics, it is paramount that the correct sediment concentration in both time and space is predicted likewise. This work therefore seeks to lay ground for more accurate modelling of transportation deposition and resuspension sediments in Lake Victoria.

5.2.3 Overview of the study

The hydrodynamic module of Lake Victoria developed WL| Delft Hydraulics under the Lake Victoria modelling framework was adapted. A wave module was built and coupled with the flow module to run online. Due to limited observed data the model was validated using knowledge on sediment resuspension and lake mixing patterns as described in published literature.

5.2.4 Model set up

The model set up was composed of the FLOW model and the WAVE model. These were schematised as follows;

The FLOW model

The FLOW model simulated the hydrodynamic phenomena of the system namely, residual currents, significant wave height, sediment transport etc. In space the model was schematized such that the mesh grid size was typically 4km in the horizontal and 20 equidistant layers in the vertical. It ran on a time step of 30minutes for a simulation period of one year (1996). Bathymetry used in the model was provided by WL| Delft Hydraulics as digitized from Admiralty charts 3252 and 3665. Figure 5.2 shows the computational grid and bathymetry of Lake Victoria.

The main rivers flowing to the Lake Victoria namely; Kagera, Sondu, Yala, Nzoia, Ruizi, Katonga, Nyando, Mara and the Nile were also schematized. It is likely that the bottom roughness varies in space due to difference in bed vegetation and sediment particle distributions however due to lack of information of this distribution; a uniform bottom resistance was used.

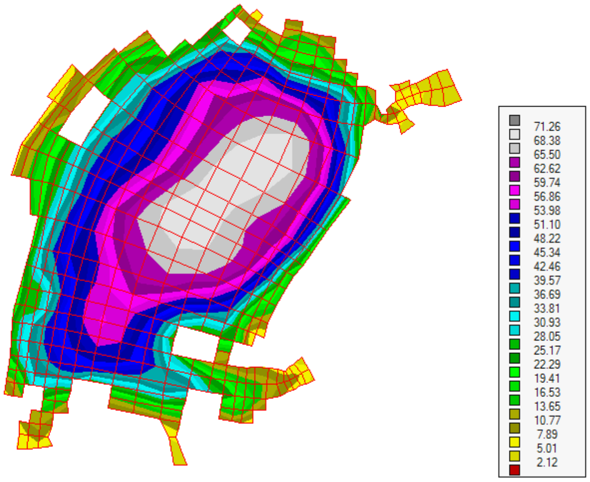


Figure 5.1 Computational grid and bathymetry of Lake Victoria

Wind-wave model

The process of sediment resuspension and transport in large water bodies particularly shallow lakes mainly starts with wind energy being delivered to the water surface and causing waves (Jin and Ji, 2004). The wind energy is transmitted from the surface to the bottom while being dissipated into wave motion in the vertical direction. This phenomenon creates orbital velocities at the sediment-water interface, which in combination with current velocities exert shear stresses leading to entrainment of sediment into the water column. Thus wind waves and current are the major driving forces to cause sediment resuspension in shallow lakes.

The WAVE model was used to simulate the generation, propagation and dissipation of waves. The WAVE model computed the significant wave height that was used by the FLOW model together with the FLOW simulated current velocity to compute the total bed shear stress. The wave model received wind speed from the FLOW model and used the same bathymetry data and computation grid.

Data input

1996 wind recorded at Entebbe international airport (WMO Center 750 station No. S89320660) was used. Although weather conditions vary from the south to the north of the lake, the same diurnal and seasonal trends occur across the basin (Newell, 1960). Hence wind data recorded at Entebbe international airport (located on the shores of Lake Victoria) can be taken as being representative of the general wind conditions over the lake. The wind data was collected on an averaged time step of one minute where as the direction was taken on an hourly basis. The data was averaged on a daily basis and implemented in the model.

Daily discharge for the nine rivers contributing the highest discharge where schematised in the model. Incidentally rivers contributing the highest discharge also contribute the highest sediment to the lake.

There is no dedicated sediment monitoring campaign on the rivers flowing into the lake no sediment time series were available for use in the model. However it is estimated that the average annual sediment loading from river Kagera is about 736 tones per year (Azza, 2006). This value was used to compute the average sediment concentration per unit volume of flow. The sediment concentration of the other rivers was computed in proportion to river Kagera's contribution.

5.2.5 Model calibration,

The validity of model simulation is to a large extent judged by its agreement with observed data. However there is no lake-wide monitoring campaign on Lake Victoria designed for the collection of data to facilitate systematic model calibration. With this limitation the primary aim of the calibration exercise was to correctly simulate the seasonal pattern of known hydrodynamic phenomena such as the periodical mixing of the lake. Information on the mixing pattern of the lake was sought from published literature. During calibration emphasis was placed on adjusting parameters (namely, sediment settling velocity and critical bed shear stress) that influence sediment transportation. The model was valid for the period 01 January 1996 to 31st December 1996.

5.3 Results and discussion

A good assessment of a model's performance would entail quantifying its prediction accuracy using standard performance measurements such as mean error, mean absolute error, root mean square error etc. Due to inadequate data particularly on hydraulic conditions and water quality parameters this kind of performance assessment could not be achieved. Conclusions on performance were drawn qualitatively on the basis of the model's simulation of documented physical processes/patterns of the lake system on both temporal and spatial scales.

Model simulations of hydrodynamic parameters seem to agree with wind patterns which are the major forcing for the system. The hydrodynamic parameter used in this evaluation is depth averaged velocity magnitude and direction. For a good comparison we need to understand the wind patterns over the lake.

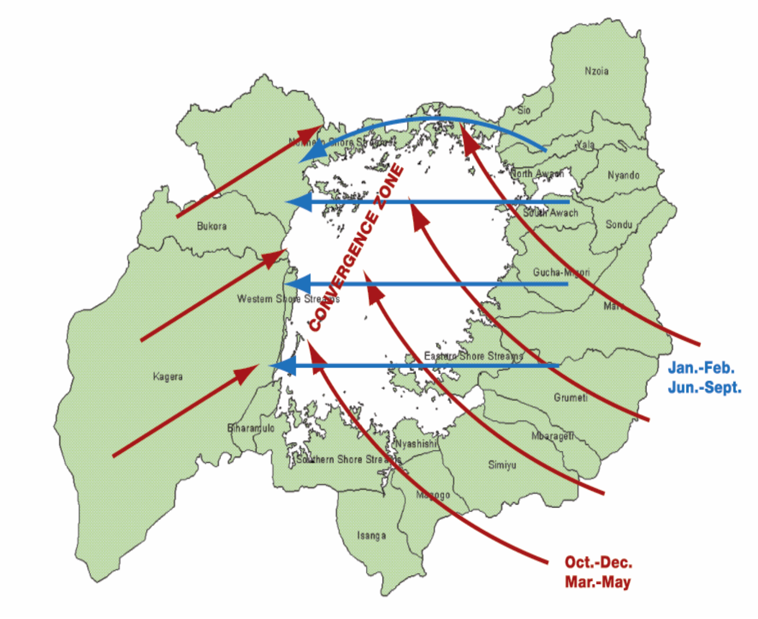


Figure 5.2 Seasonal wind patterns influencing the hydrological processes in Lake Victoria Basin, adopted from LVEMP(2002)

Wind blowing over Lake Victoria is the most important forcing for the hydraulic processes in Lake Victoria. The seasonal wind pattern is influenced by the movement of the Inter Tropical Convergence Zone (ITCZ) as a result of the apparent movement of the sun across the equator. Figure 5.2 shows a schematisation of the seasonal wind patterns In January-February and June-September, the wind pattern is predominantly East West, parallel to the equator, with origins from the Nandi hills in western Kenya. In the period March-May and October-December, the wind pattern changes towards the northern parts of the lake (Okonga and Sewagudde, 2006)

Comparison of the seasonal wind pattern and the depth averaged velocity shows that the model depicts the seasonal trends in the system. This comparison is useful in explaining the high sediment resuspension on the western coast of the lake and movement of sediment in the northern direction. Figures 5.3 through 5.6 show the depth averaged velocities in the lake, the overall direction of the velocity fields in Figure 5.3 and 5.4 is the East-west direction which is quite in agreement with the Jan-Feb and Jun – Sep, wind pattern respectively.

In Figures 5.5 to 5.6 the velocity fields are predominantly oriented in the south-direction a phenomenon that is consistent with the March-May and October-December wind patterns. In Figures 5.5 and 5.6 we notice that the powerful velocity currents mainly act on the western shore. This partly explains the enormous sediment resuspension/concentration along this shoreline. The SPM concentration is further amplified by enormous sediment loading from River Kagera. The sediment is sustained in suspension for a long time by the strong currents which translate into shear stresses that are above the critical bed shear stress for most of the time.

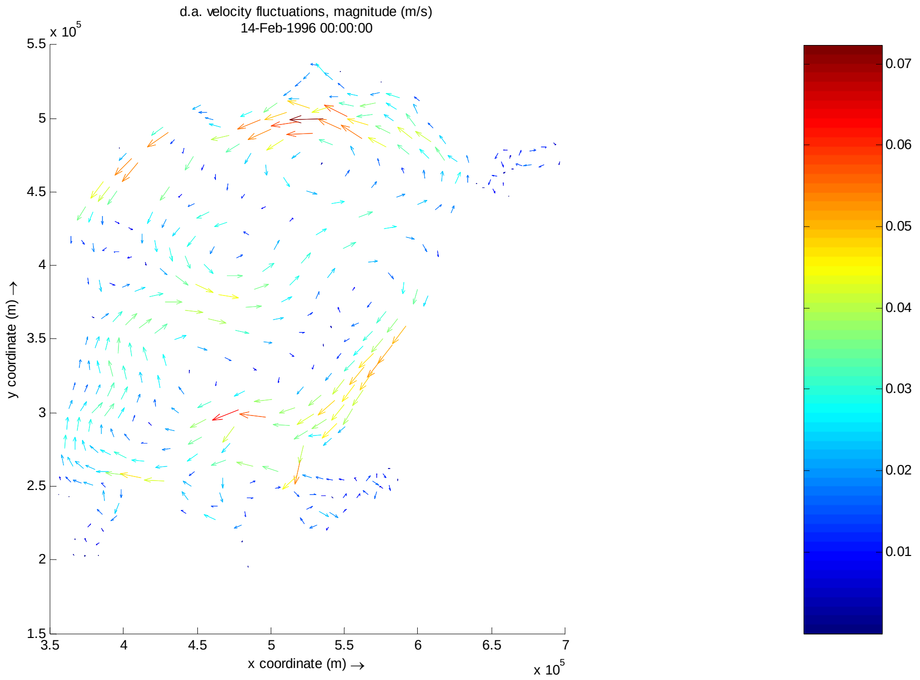


Figure 5.3. Hydrodynamic simulation under the January – February winds on 14 February 1996

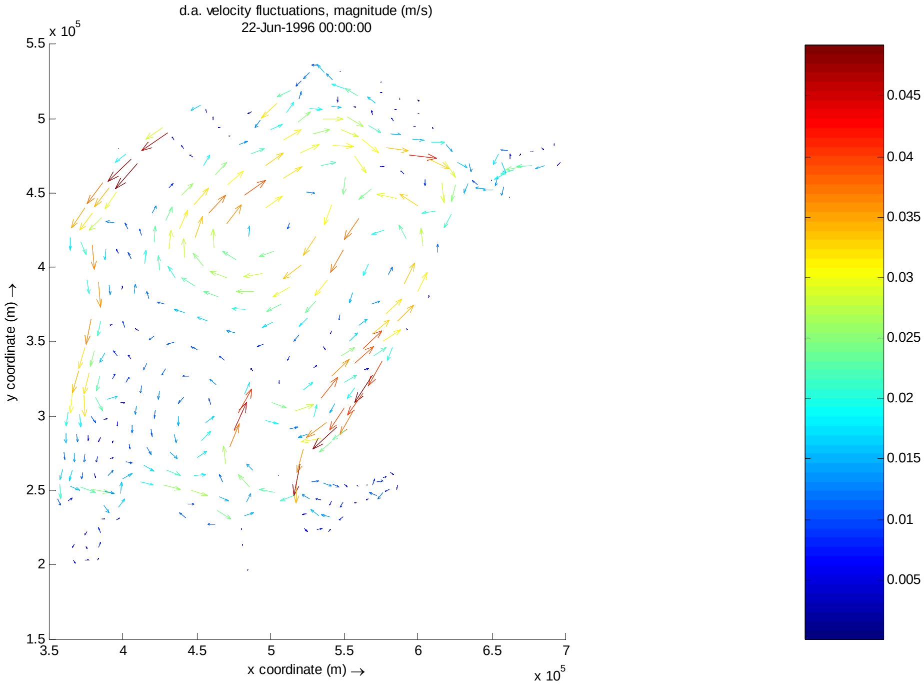


Figure 5.4: Hydrodynamic simulation under the June – September winds on 22 June 1996

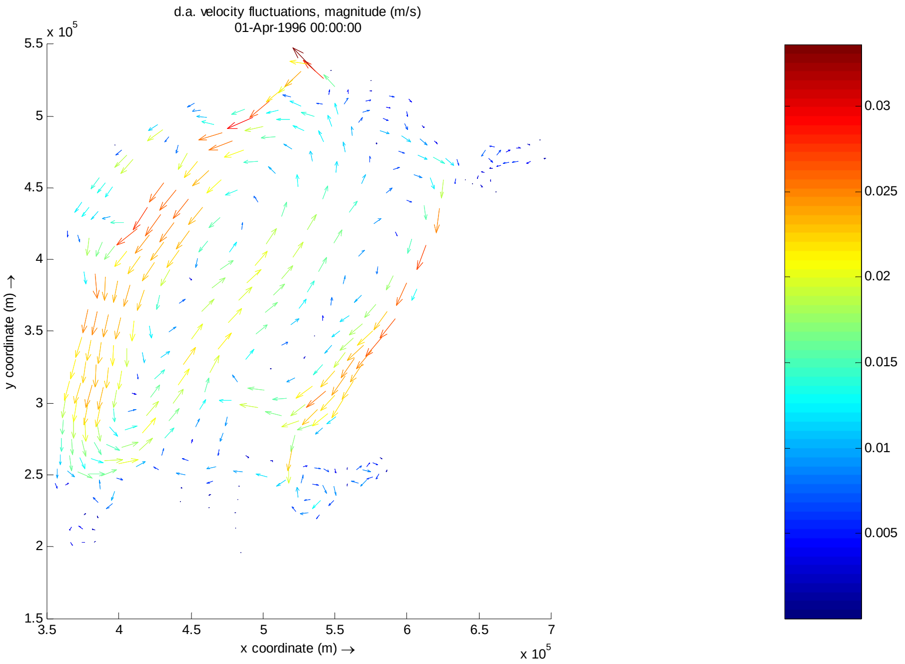


Figure 5.5: Hydrodynamic simulation under the March-May winds on 01 April 1996

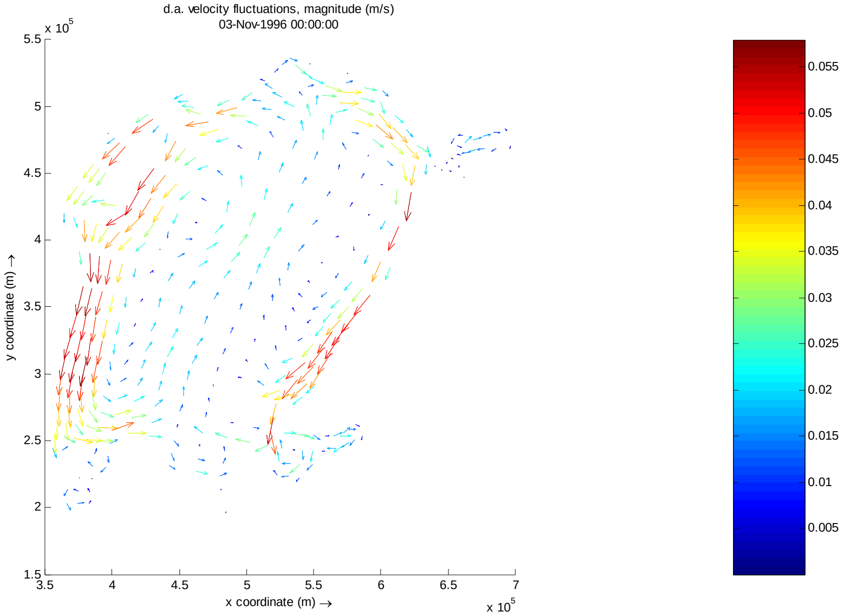


Figure 5.6: Hydrodynamic simulation under the October-December on 03 November 1996

From an examination of skewness, sorting and sedimentological zones, Azza (2006) suggested that a high frequency of resuspension and a northward movement of materials could be occurring at the western coast. He also suggested that some transportation at the mouth of river Kagera to be occurring due to the influence of the river. Model simulations that show a high sediment concentration in this part of the lake for the greatest part of the year confirmed his assertion. Results showed that transport due to the river plume does not persist for a long distance because of relentless wave attack on the west coast that brings about complete mixing in this area. In fact this relentless attacks is responsible for the fact that this area of the lake does not experience stratification all year round (LVEMP, 2002)

Model results without wave model showed that much as the model may not predict the sediment resuspension trend well, but the effect of current induced resuspension and transportation contributes a significant part in the sedimentation processes. Azza (2006) pointed out that predicting sediment resuspension and transport in large lakes from only the action of wind induced surface waves may leave out other processes that may not alter substantially the position of the mud deposition boundary but still lead to significant redistribution of fine-grained sediments and hence to nutrients and pollutants associated with this class of particles. It is further noted that generalizations of sedimentation processes from small lakes may give misleading results when applied to large lakes.

Results further show that suspended sediment along the western shower line moves along the coast especially during March-May and September-December, this is caused by the intensive oblique wave attacks due to the strong winds during the March-May and September-November period that cause enormous resuspension and consequently sediment transport parallel to the shoreline.

The model predicts well the annual complete mixing of the lake that occurs between June and September (Azza, 2006; Hecky, 1993). In the model results this complete mixing is indicated by almost an increased colour pattern and a random velocity field distribution throughout the lake as seen in Figure 5.7. An insignificant amount of sediment is brought into suspension on the east coast and centre of the lake as compared to the western coast. This could be explained by the fact that the west coast for almost its entire length is always the downwind side of the winds blowing over the lake and given that the winds reaching it have maximum fetch values this area of the lake actually experience maximum turbulence and scoring than any other part to the lake.

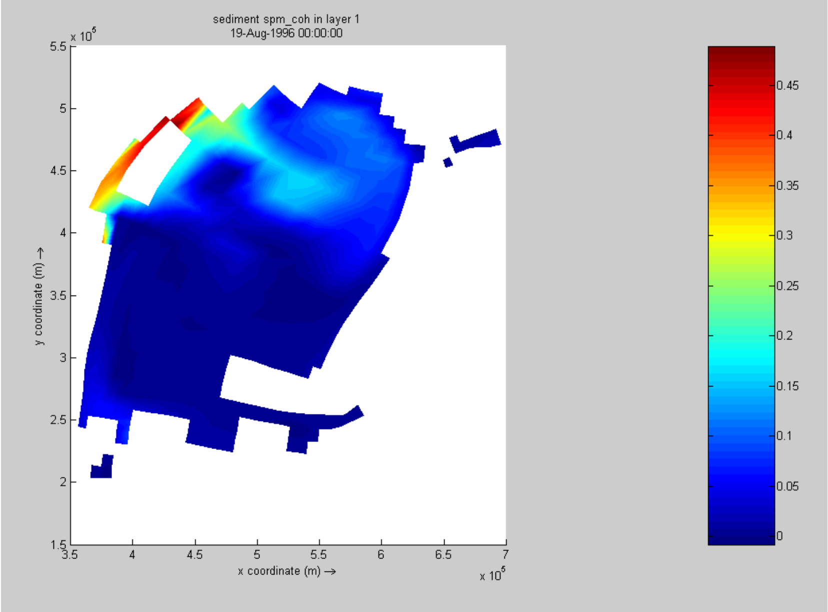


Figure 5.7: Simulation of sediment in Lake Victoria on 19 August 1996

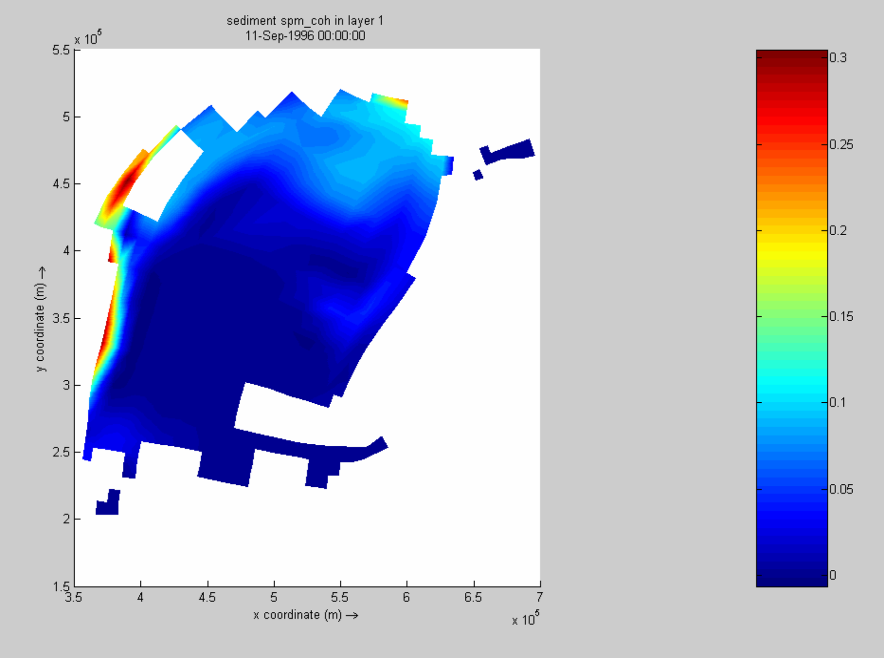


Figure 5.8Simulation of sediment in Lake Victoria on 11 September 1996

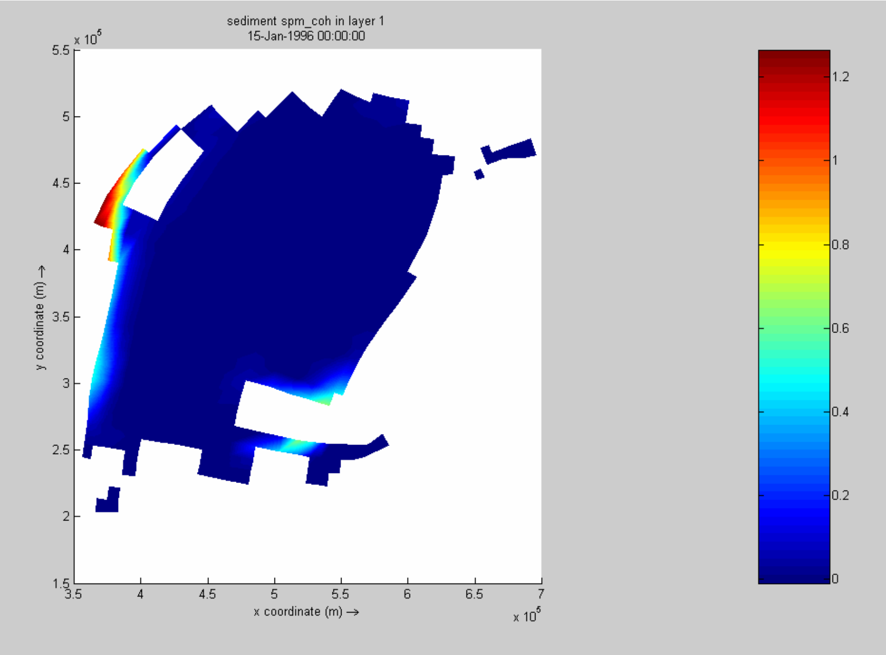


Figure 5.9: Simulation of sediment in Lake Victoria on 15 January 1996

The above discussion shows that the model calibration is good enough to draw general conclusions on sediment transport in Lake Victoria. Strategies for more accurate modelling and understanding of the transport in Lake Victoria can be drawn basing on this understanding. The following can be summarised from the model simulations;

1. That the western cost of Lake Victoria experiences intensive oblique wave attacks
2. It also experiences scoring and maximum sediment resuspension than any other part of the lake
3. Sediment travels along the coast due to wave induced alongshore currents
4. The river Kagera contributes a substantial amount of sediment that is transported along the coast and remains in suspense for a considerable amount of time.

5.4 Application of hybrid knowledge gained in the Dutch Coast

To transfer expertise in hybrid modelling of the Dutch coast to Lake Victoria, first we need to establish the similarities and differences between the two systems in as far as sediment resuspension and transport is concerned. From this understanding we can identify the feasibility, possible adjustment in the methodology and limitations.

5.4.1 Similarities between the Dutch costs and Lake Victoria

From the model simulations the following similarities in the Lake Victoria system and the Dutch coastal systems can be drawn.

1. Just like the Dutch coast, the west coast of Lake Victoria experiences relentless wave attacks that are responsible for sediment suspension and turbulent mixing of the Kagera river flow.

2. Sediment resuspension in both systems is strongly influenced by wave and velocity current interactions. This disproves earlier views that currents are of little significance in sediment transport in large lakes (Azza, 2006)

3. English Channel is the major sediment source of the Dutch coast and sediment continue to travel along the coast in the north eastern direction. In Lake Vitoria the major sediment source of the western coast is River Kagera. However, the major similarity is the transportation of the sediment along the coast by the alongshore currents.

5.4.2 Differences between the Dutch coast and Lake Victoria

1. Unlike the Dutch coast there is no effect of coriolis forces in Lake Victoria given that Lake Victoria lies on the equator. This phenomenon ensures adequate mixing of the Kagera river flow and the lake waters.

2. Density currents in Lake Victoria are as a result of temperature difference alone. In the Dutch coast density currents are as a result of salinity difference in addition to temperature. Salinity difference is an important parameter in the simulation of sediment transport due to the coastal river effect caused by fresh water flow from the river Rhine.

3. Given that Lake Victoria is generally a shallow lake (a bout 85m at the deepest point) (Mangeni *et al.*, 2006) the effect of flow transitions from very deep offshore to the shelf is not significant.

4. Unlike the Dutch coast Lake Victoria does not experience tidal effects.

5.4.3 Hybrid modelling framework in Lake Victoria

In the absence of necessary data for development a DDM, this section only highlights a framework within which hybrid modelling can be applied in Lake Victoria.

From findings in section 5.4.1 and 5.4.2 the most relevant area of application of the Dutch coast hybrid modelling principles is the western coast of lake Victoria. This side of the lake has several major towns that draw their domestic water supply from the lake including Kampala city and Jinja. As earlier illustrated, suspended sediments have a direct implication on the water quality through light attenuation and/or nutrient release. We can chose to build a fine local model to study in detail the sediment resuspension and transporting dynamics and their long term implication on the health of the lake along hot spots such as the Murchison bay, Napoleon gulf e.t.c. along the western coast.

In this case the model would need boundary conditions that correctly represent the sediment transport along the western coast. But just like in the Dutch coast the source of sediments could be out of the model domain. Accordingly the influence of the sediment source should be well captured in the boundary condition. As seen in chapter five, time varying boundary conditions would be handy in studying long term effects like in this case. These could be generated using the methodology developed in the hybrid modelling of sediment transport in the Dutch coast

From the identified similarities and differences we can conclude that if a DDM where to be applied to the western coast of lake Victoria there is a high likelihood that it would have a better simulation performance than the one built at Noordwijk. This is based on the understanding that there are fewer parameters that influence the resuspension and transport process in Lake Victoria as compared to the Dutch coast. There are only three variables used in the DDM compared to the parameters responsible for sediment resuspension. It is logical that if the modelled system has fewer parameters/processes then there is a high likelihood that the DDM will perform well on the predictions.

Unlike in the Dutch coast the importance of sediment loading from river Kagera can not be underestimated. River Kagera discharge would be a very important input variable in the DDM that would be used as the generator of time varying boundary. As a modification, it is suggested the Kagera discharge forms the fourth input variable in addition to significant wave height, southerly wind speed and total beds shear stress. Depending on the distance of the south boundary the discharge lag used in DDM would vary accordingly. The suggested DDM time boundary generator is given as in expression 6.1

$$S_s = f\left(H_{\text{sig}}, \tau_{cw}, w_{\text{southerly}}, q_{\text{Kagera}} \right) \qquad 5.1$$

Where

S_s = suspended sediment
H_{sig} = Significant wave height,
τ_{cw} = total bed sheare stress,
$w_{\text{southerly}}$ = southerly wind,
q_{Kagera} = River Kagera discharge

5.5 Relevancy to the development of the Nile Basin Decision Support System

A general framework in which sediment modelling in Lake Victoria could be used in drawing catchment based invention measures for sustainable management of the lake resources is illustrated in the following sections. Several auxiliary models are suggested whose output would form inputs to the sediment model and vice versa. The implementation and linkage of these models is briefly illustrated. It is suggested the models are integrated to form a Lake Victoria Decision Support Tool (LVDST). The LVDST could in turn form a component of the larger Nile Basin Decision Support System.

5.5.1 Approach

The deteriorating health of Lake Victoria is threatening to undermine its capacity to deliver benefits to the riparian communities(LVEMP, 2007). To realize the highly cherished economic potential of the lake there is need to salvage it from the deteriorating condition of water quality. This requires devising intervention measures that will reverse the trends. Measures should convince policy makers that the situation will worsen if the current status-quo is maintained.

Suggested interventions should bear proof of alleviating the current and predicted problems and that they are cost effective and/or sustainable. This calls for developing modelling tools that would be used to simulate the future state of the lake system under different scenarios. The tools should be integrated into a Decision Support System (DSS) that would assist policy makers in making informed decision.

Modelling tools to achieve this objective would include a hydrological model, a hydrodynamic model simulating the sediment resuspension and transportation, ecological model, and an economic model. The following sections illustrate how the models would be implemented and link to each other.

5.5.2 Hydrological/catchment model

A catchment model would simulate hydrological processes within the catchment. The tool serves to predict the impact of management changes on flow and sediment in rivers. A semi/distributed model would be preferred since it would facilitate making changes on a more representative scale during scenario development and assessment. Besides the hydro-meteorological data, the most important input (spatial) data for such a model would be a digital elevation model, land use data, soil data and crop management data. The management practices and land use data would be varied to generate scenarios that could represent suggested interventions. For example the model could be run for a scenario where minimum tillage and buffer stripes are recommended for a type of land use and soil type then see how much sediment reduction is achieved and its impact on the lake.

5.5.3 Hydrodynamic modelling

Dissolved substances are transported with water. Determining the transport of water through hydrodynamic modelling is therefore vital in sediment modelling and water quality modelling in general. We can study the effects of sediment loading, resuspension and transportation on the health of the lake system. For the hydrodynamic modelling two modules could be considered. One model to route sediment from the catchment to the lake, and other one for modelling the hydrodynamics of sediment transport, settlement and resuspension within the lake

The routing model would be a 1D model to route the water transporting the sediment up to the mouth of the river. This model is important in ensuring that the right amount of sediment is delivered to the lake. The loads from the routing model would then be used as boundary conditions to a 3D lake model. Ideally measured values would suffice however during scenario assessment the routing model becomes versatile in routing sediments through the rivers when different land use changes and land management options are considered.

To correctly simulate the Lake Victoria processes the minimum data required would include temperature, wind speed and direction, evaporation, rainfall, river discharge, solar radiation, cloud cover etc. These data are key in simulating the lake process particularly the stratification of the lake as well as complete mixing.

5.5.4 Ecological modelling

An ecological model is required to correlate the effect of SPM on the functioning of the lake biological system. The entire biological system can not be simulated by the ecological modelling technique. Therefore effort would be concentrated on simulating the development of species that affect the lake's health. In this case the easily quantifiable species is algae. The simulation of the algal bloom, in particular the toxic blue – green type would be indicative of the extent of the declining health of the lake.

The easiest way to link SPM to algal growth is through water transparency. Water transparency plays an important role in algal growth and growth of submerged aquatic vegetation. Water transparency in a water column is quantified by the light extinction coefficient. Cerco and Cole (1994) confirmed that the light extinction coefficient is closely related to the total suspended sediment in a water column.

Jin and Ji (2005) established a linear relationship between the light extinction coefficient and total suspended sediment by fitting linear equations to measured total suspended sediment and light extinction data. This relation can then be used to establish the Secchi depth by employing an empirical relation between Secchi depth and light extinction coefficient using one of the empirical relationships developed by Sverdrup and Munk (1947)

After simulating the above evolutions the next step would be to validate the model. Currently spatial monitoring of eutrophication or suspended sediments is being not carried out on Lake Victoria. It is not likely to take place even in the near future due to logistical impediments (WL| Delft Hydraulics *et al.*, 1999). This leaves us with one choice; that is using remote sensed data. The predicted algal growth should be compared with satellite images that have been processed for algal brooms. It is probable that this would even give a better spatial coverage than the ground monitoring measurements.

5.5.5 Economic model

Having achieved the quantification and proliferation of the algal bloom, economical implications can then be ascertained. This can be achieved by deriving the relationship between the extent of the area covered by algae and the number of fish kills plus all other benefits that are hindered by the presence of algae. Different land based intervention scenarios can now be generated and examined their impacts on the development of algal blooms in the lake and consequently the economic implications.

5.5.6 Climate change

An interesting aspect that could be considered is the effect of climate change on the health of the lake. This is an issue that runs across the four models considered in the approach. Climate change could be factored in say by assessing the implication of increased rainfall, flooding or droughts in the hydrology module. In the ecological module the effect of climate change can also be incorporated by considering the effect of increased average water temperature on the water quality and the development of algal blooms.

5.6 Conclusion

- The similarities and differences between Lake Victoria and the Dutch coast suggest that the application of hybrid modelling knowledge of the Dutch coast is most relevant when applied to the western coast of lake Victoria.

- Localised models to study sediment dynamics on the long-term impacts on water quality would benefit the most from a hybrid model approach.

- The time varying hybrid modeling concept has little relevancy when the entire lake is considered. However the concept of approximating a numerical model using a DDM would be handy if the entire lake is considered.

- The accuracy of simulation using the hybrid modeling in Lake Victoria is likely to be higher than in the case of the Dutch coast. This is due to the fact that there are fewer complicated processes in Lake Victoria than in the Dutch coast.

- An approach on how modelling of suspended sediment can be applied in a broad perspective of managing Lake Victoria water resources has been highlighted. The approach described could be integrated with previous lake Victoria decision support tools namely the lake Victoria model framework and the lake Victoria decision support tool to form a lake Victoria DDS. Integrating these tools will enable informed decision making and management of the water resources in a more sustainable manner.

6 Conclusion and recommendations

6.1 Introduction

This study was aimed at exploring the applicability of hybrid modelling to improve the simulation accuracy and understanding of sediment transportation processes in the Dutch coast and Lake Victoria. First it explored the possibility of approximating a physically based model using a data driven model to predict SPM concentrations. Relevant input variables and input data manipulation operations to expose maximum information to a data driven modelling tool with the purpose of improving the model's predictive power were identified

In another case the effect of specifying time varying open boundary condition instead of fixed boundary conditions on the simulation of suspended particulate matter (SPM) along the Dutch coast was investigated. The aim was to develop a methodology for generating time varying open boundary conditions using a data driven model with local hydrodynamic and meteorological conditions as input variables. The generated time varying boundaries were then used in a numerical model to study the improvement on the sediment simulation results as compared to fixed boundary.

The study also examined the applicability of the knowledge gained in modelling sediment transport along the Dutch coast to understand sediment transport in Lake Victoria. The aim was to lay ground for future more accurate and reliable modelling of sediment transport in Lake Victoria. This was achieved by identifying the most important forcing functions in sediment transport and resuspension in Lake Victoria. The subsequent sections give conclusions and recommendations drawn form these three broad activities.

6.2 Conclusion

6.2.1 Approximating a numerical model using a DDM

Results have shown a strong possibility of approximating a numerical model using a point data driven model. Comparison of the SPM computed by the numerical model and that predicted by the DDM showed a good fit with a correlation coefficient of 0.9177. However in some instances the SPM was significantly underestimated by the DDM. The discrepancy has been likened to the use of one input variable (total bed shear stress) that may not capture well the major processes in the sediment resuspension and transport process.

Indeed this assertion was been verified by the comparison of SPM and significant wave height. Results showed that SPM responds more to the significant wave height then to the total bed shear stress. It is probable that besides the magnitude of the bed shear Stress the mechanism by which the sediments are brought into suspension is equally important. For the same magnitude of bed shear stress more sediment is bought into the water column under wave conditions as compared to currents alone. Eddies and orbital motion associated with wave action are responsible for this phenomenon. Thus significant wave height is an important input variable to any DDM that is seeking to predict sediment resuspension.

6.2.2 Performance of time varying sediment boundary conditions

The total bed shear stress, significant wave height and wind as local hydrodynamic and meteorological input variables to a DDM have shown satisfactory performance in the simulation of SPM. For a system in which the overall direction of sediment transport is known such as the Dutch coast, it is likely that using the wind component in that direction could improve the results further.

The study revealed that while transferring a point DDM it is necessary to apply both alongshore and cross-shore correction factors. For two points in the modelled system having comparable input variables (with respect to magnitude) can have significantly different sediment concentrations. This discrepancy has been partly associated to the difference in the local morphological conditions, sediment availability and the mechanism that bring sediments into suspension.

In the case of the RIJMAMO grid of the Dutch coast, the time varying boundary has negligible effect on the SILTMAN points. If the SILTMAN points are the object of a given study then the location of the south boundary is far enough not to affect the model results in the study area significantly. We can comfortably conclude that the choice of the south boundary i.e. whether fixed or time varying may have no effect on the model results for short simulation times. However the choice might be important if the simulations involve large time scale processes such as morphological changes. Here the accuracy of the horizontal sediment flow is important as far as computing the sediment budget is concerned.

Investigation of the effect of an error in the choice of the boundary value indicted that the error in model domain reduces exponentially in the alongshore direction. Basing on the developed relationship of error reduction and propagation within the model domain, the size of the RIJMAMO model domain can be reduced considerably and accordingly the computation time. The location of the boundary would depend on the location of the study area and the acceptable error in the simulation results. If say the SILTMAN points or Rotterdam inlet channel are the intent of the study, the RIJMAMO grid/model could be reduced by half the size. The computation time would consequently be reduced in the same proportion but still achieving an acceptable degree of accuracy in the predicted SPM.

6.2.3 Modelling of sediment in Lake Victoria

This is the first study to make an attempt at modelling sediment transportation, resuspension and deposition in a 3-dimensional numerical model. Model simulation results though based on limited data, they have shaded more light on the transportation and resuspension of sediments in Lake Victoria.

Simulations revealed some process similarities and differences between Lake Victoria and the Dutch coast. The major similarities are;

- That sediment resuspension in both systems is mainly as a result of wave and velocity current interactions.
- Just like the Dutch coast, the western coast of Lake Victoria experiences relentless wave attacks that are responsible for sediment suspension and turbulent mixing.

- In general sediment in the Dutch coast is transported along the coast in the northern direction. This phenomenon is also observed in the western coast of Lake Victoria in which sediments mainly fed by river Kagera are transported along the cost in the northern direction.

The similarities lead us to conclude that the application of hybrid modelling knowledge of the Dutch coast is most relevant when applied to the western coast of lake Victoria. The time varying hybrid modeling concept has little relevancy when modeling the entire lake as a single unit. However the concept of approximating a numerical model using a DDM would be practical if the entire like is considered.

It is probable that the accuracy of simulation using the hybrid modeling in Lake Victoria would be higher than in the Dutch coast since there are fewer complicated processes in Lake Victoria than in the Dutch coast. For example there is no coriolis, tidal, salinity, breaker zones effects in Lake Victoria. Given that the same input variables for the data driven models would be considered for both the Dutch cost and the Lake Victoria and yet the latter processes are not represented in the input data it is expected that the simulation results in Lake Victoria's case would be much better.

6.2.4 Relevancy to NBI in the context of sustainable development

A framework has been proposed in which the suggested modeling methodology can be incorporated into the larger perspective of managing Lake Victoria water resources in general. The following interconnected modules have been suggested namely; Hydrological model, flow and sediment routing model, 3D hydrodynamic model, ecological model and an economic model. The economic model would be used for cost benefit analysis to assess land based interventions suggested under different scenarios. The framework could be implemented in a Lake Victoria DSS unit that could eventually form a component of the Nile basin Decision Support System.

6.3 Recommendations

In this study a simple average ratio approach was employed to generate along shore and cross-shore correction factors that were used in transferring the data driven model. It is probable that there are physical characteristics of the system responsible for the variation in SPM between say two points with similarly hydrodynamic and meteorological parameters. Further research is recommended to explore the possibility of empirically relating these physical characteristics to formulate a mathematical expression for the computing the two correction factors.

The hybrid approach used in the study has shown that a small improvement in the prediction of SPM resuspension pattern is achieved by supplying time varying variables. Another hydride approach involving error prediction is being proposed. In this approach a data-driven model could be used to model the error pattern of the physically based model. The DDM predicted error would then be added to the outcome of the physically based model in the subsequent time step to generate the final prediction. The approach may require building several point DDMs within the numerical model domain. A similar approach was used by Abebe and Price (2003) to predict runoff, Vojinovi *et al* (2003) used the approach to model the wet weather response in a wastewater system whereas Hur and Kim (2008) used the method to improve classification accuracy . The main difference between the proposed approach and the previous ones is that it will be implemented with a 3D model whereas previous work was implemented on 1D models.

In Lake Victoria the hybrid modeling is more relevant to the western part of the lake. Further research could be directed towards developing a methodology to upscale results obtained in this part to the entire lake.

In-lake monitoring of water quality particularly suspended sediments faces serious impediments (WL| Delft Hydraulics *et al.*, 1999). The fact that the lake lies within three countries makes it difficult to carryout a systematic lake wide monitoring programme. Any immediate future modeling work will heavily relay upon remote sensed/ satellite data. It is recommended that an inventory of satellite imagery over Lake Victoria is made. The imageries should be processed and validated against the limited ground measurement. The following are the crucial parameters that should be reflected in the images; air temperature, atmospheric pressure, precipitation, evaporation, relative humidity, solar radiation, cloud cover and suspended sediments.

References

Abebe, A. J., and Price, R.K., (2003) Managing uncertainty in hydrological models using complementary models. *Hydrological Sciences*, **48**(5), 679-692.

Azza, N. (2006) *Pattern and mechanisms of sediment distribution in Lake Victoria*, Technical University of Delft, Delft.

Bhattacharya, B., Deibel, I. K., Karstens, S. A. M., and Solomatine, D. P. (2007) Neural networks in sedimentation modelling for the approach channel of the port area of Rotterdam. Delft Cluster research programme.

Bijker, E. W. (1967) *Some considerations about Scales for Coastal Models with Movable bed*, Delft University of Technology, Delft.

Blom, E., Van Duin, E. JH. S., Aalderink, R.H., Liklema, L., and Toet, c. (1992) Modeling sediment transport in shallow lakes - Interactions between sediment transport and sediment composition. . *Hydrobiologia*, **235**(236), 153-166.

Booth, J. G., Miller, R. L., McKee, B. A., and Leathers, R. A. (2000) Wind-induced bottom sediment resuspension in a microtidal coastal environment. *Continental Shelf Research*, **20**(7), 785-806.

Box, G. E. P., Jenkins, G. M. (1970) *Time Series Analysis Forecasting and Control*, Holden-Day Press, San Francisco, California, USA.

CERC (1984) *Shore protection manual*, U.S. Army Corps of Engineers, Fort Belvoir, Virgnia.

Cerco, C. F., and Cole, T. (1994) *Three-dimensional eutrophication model of Chesapeake Bay*. *El-94-4*, US Army Corps of Engineers, vicksburg Mississipi.

Chen, S., and Eisma, D. (1995) Fractal geometry of in-situ flocks in the estuarine and coastal environments. *Netherlands Journal of Sea Research*, **32**(2), 173-182.

Eisma D., I. G. (1988) Suspended matter and sediment transport. In: *Pollution of the North Sea: an assessment*, W. S. e. al, ed., Springer, Berlin, German, 20–33.

Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*, Tom Robbins, 2nd.

Hecky, R. E. (1993) The eutrophication of Lake Victoria. *Verh. Limnology*, **25**, 39-48.

Herman, A. (2007) Numerical modelling of water transport processes in partially-connected tidal basins. *Coastal Engineering*, **54**(4), 297-320.

http://www.cse.msu.edu/~cse471/gaussian.pdf.

http://www.knmi.nl/samenw/hydra Wind Climate assessment of the Netherlands.

http://www.waterbase.nl [Online] [Access date:

Hur, J., and Kim, J. W. (2008) A hybrid classification method using error pattern modeling. *Expert Systems with Applications*, **34**(1), 231-241.

Ji, Z. G., Hamrick, J. M., and Pagenkopt, F. (2002) Sediment and metals modeling in shallow river *Journal of Environmental Engineering*, **128**(2), 105-119.

Jin, K. K., and Ji, A. G. (2004) Case study: Modeling of sediment transport and wind-wave impact in lake Okeechobee. *Journal of Hyadraulic Engineer*, **130**(11), 1055-10067.

Jin, k. R., and Ji, Z. G. (2005) Application and vildation of Three Dimensional model in a shallow lake. *Jouurnal of waterway, port coastaa and oceen engineering*.

Krumbein, W. C. (1936) Application of Logarithmic Moments to Size Frequency Distribution of Sediments. *Journal of Sedimentary Petrology,*, **6**(1), 35 - 47.

Leussen, W. v. (1994) *Estuarine macro-flocs and their role in fine-grained sediment transport*, University of Utrecht Utrecht, The Netherlands.

LVEMP (2002) *Integrated Water Quality and Limnology Report of Lake Victoria*. LVEMP.

LVEMP (2007) Project information document. Lake Victoria Environmental Management Programme.

Maggi, F., Mietta, F., and Winterwerp, J. C. (2007) Effect of variable fractal dimension on the floc size distribution of suspended cohesive sediment. *Journal of Hydrology*, **343**(1-2), 43-55.

Mangeni, B., Ngirane-Katashaya, G., Mwakali, J. A., and Taban-Wani, G. (2006) Preliminary investigation of Lake Victoria groundwater situation from advanced very high resolution radiometer data. In: *Proceedings from the International Conference on Advances in Engineering and Technology*, Elsevier Science Ltd, Oxford, 195-202.

Marchesiello, P., McWilliams, J. C., and Shchepetkin, A. (2001) Open boundary conditions for long-term integration of regional oceanic models. *Ocean Modelling*, **3**(1), 20.

McAnally, W. H., ; Friedrichs, C.; Hamilton, D., ; Hayter, E.,; Shrestha, P.,; Rodriguez, H.,; Sheremet, A.; and Teeter, A., (2007) Management of Fluid Mud in Estuaries, Bays, and Lakes. part I: Present State of Understanding on Character and Behavior. *journal of Hyadraulic Engineer*, **133**(1), 9-22.

Minns, A. W. (1998) *Artificial Neural Networks as subsymbolic process descriptors*, PhD, Delft University of Technology, Delft.

Moore, R. A., Reddy, D. R., and Fisher, M. M. (1998) Phosphorus flux between sediment and overlying water in Lake Okeechobee, Florida: Spatial and temporal variations. *Journal of Environmental Quality*, **27**, 1428-1439.

MVogel, R., and Kroll, C. N. (1992) Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research*, **28**(9), 2451-2458.

Newell, B. S. (1960) The hydrology of Lake Victoria. *Hydrobiologia*, **15**, 363-383.

Okonga, J. R., and Sewagudde, S. M. (2006) *Hydro-Meteorological Observations Over The Ugandan Side of Lake Victoria. under publication*. LVEMP.

Olia, O. G., and Reddy, K. R. (1993) Phousphours sorption characteristics of sediments in shallow eutrophic lakes of Florida. *Archiev Hydrobiol*, **129**, 45-65.

Panu, U. S., Unny, T. E. (1980) Extension and application of feature prediction model for synthesis of
hydrologie records. *Water Resources Research*, **16**(1), 77-96.

Postma, H. (1981) Exchange of materials between the North Sea and the Wadden Sea. *Marine Geology*, **40**(1-2), 199-213.

Pyle, D. (1999) *Data preparation for data mining*, Academic press.

Raman, H., and Sunilkumar, N. (1995) Multivariate modelling of water resources time series using artificial neural networks. *Hydrological Sciences*, **40**(2), 145 -163.

Richardson, M. C., Branfireun, B. A., Robinson, V. B., and Graniero, P. A. (2007) Towards simulating biogeochemical hot spots in the landscape: A geographic object-based approach. *Journal of Hydrology*, **342**(1-2), 97-109.

Rijkswaterstaat (2002) *Atlas of near-surface total suspended matter concentratios in the Dutch coastal zone of the North Sea*. *RIKZ/20002.059*, Ministry of Transport, Public Works and Water Management.

Roelvink, J. A., Kaaij, T. v. d., and Ruessink, B. G. (2001) *Calibration and verification of large scale 2D/3D flow models*

Roelvink, J. A., and Walstra, D. J. R. (2005b) Keeping it simple by using complex models. *Advances in Hydro-Science and Engineering* **VI**.

Schalkoff, R. J. (1997) *Artificial Neural Networks*, McGraw-Hill companies. Inc. .

Schwab, D. J., Beletsky, D., Frick, W., Ge, Z., Mccormick, M., Winkelman, A., and Foley, D. (2006) Development of Near-shore Hydrodynamic Models for Beach Closure Forecasting in the Great Lakes. *Annual Conference on Great Lakes Research*.

Solomatine, D. P. (1999) Neural network approximatiion of a hydrodynamic model in optimizing reservoir operation. 2nd Internation Conference on Hydroinformatics, Zurich.

Stedinger, J. R., and Taylor, M. R. (1982) Synthetic streamflow generation - model verification and validation. *Water Resource Research*, **18**(4), 909-918.

Sverdrup, H. V., and Munk, W. H. (1947) Wind, sea and swell: Theory of relations for forecasting. *US Navy hydrographic Office*, **601**.

US Environmental Agency (2007) *Management Measures for Agricultural Sources* http://www.epa.gov/owow/nps/MMGI/Chapter2/ch2-1.html. [Online] [Access date: 1 September]

Valencia, R. D., Schaake, J. C. (1973) Disaggregation processes in stochastic hydrology *Water Resources Research*, **9**(3), 580-585.

van Rijn (1993) *Principle of Sediment Transpor in Rivers, Estuaries, Coastal seas and Oceans*, Aqua Publications, Amsterdam.

Velden, E. T. J. M. v. d. (1995) *Coastal Engineering*, Delft University of Technology.

Verwey, A. (2005) Latest Development in floodplain modelling -1D/2D integration, . *6th conference on Hydraulics in Civel Engineering*, Hobart, Autralia.

Vogel, R. M. (1999) Stochastic and deterministic world views. *Journal of Water Resources Planning and Management*, **125**(6), 311-313.

Vojinovic, Z., Kecman, V., and Babovic, V. (2003) Hybrid approach for modeling wet weather response in wastewater systems. *Journal of Water Resources Planning and Management*, **129**(6), 511-521.

Vuurens R. S. (2001) *Data-driven modelling of sedimentation in the Rotterdam port area*, Delft University of Technology, Delft, The Netherlands.

Vuurens, R. S. (2001) *Data-driven modelling of sedimentation in the Rotterdam port area*, Delft University of Technology, Delft, The Netherlands.

Wang, L. (2007) *Modelling of cohesive sediment transport in the Maasmond area*, UNESCO-IHE Institute for Water Education, Delft, The Netherlands.

Winterwerp, J. C. (1999) *On the dynamics of high-concentration mud suspensions*, Deflt University of Technology, Deflt , The Netherlands.

Winterwerp, J. C. (2006) *Fluxes of fine sediment along the Dutch coast and the impact of Maasvlakte 2: a system description, .*

WL| Delft Hydraulics (2001) *Description and Model Representation T0 Situation WL2001003 Z3030.10*, WL | Delft Hydraulics,, Delft, The Netherlands.

WL| Delft Hydraulics (2006) *Delft3D user manual*, Delft, The Netherlands.

WL| Delft Hydraulics, HydroQual Inc, and Delft, I. (1999) Prepaation of a preliminary Lake Victoria Physical Processes and Water Quality Model. The World Bank.

World Bank (2005) *The African 10 year framework programme (10YFP) on sustainable consumption and production*.

Ye, Q. (2006) *Modelling of cohesive sediment trasprotation, deposition and resuspension in the Haringvliet Mouth*, UNESCO-IHE Institute for Water Education, Delft.

List of acronyms

a : reference height for suspended sediment concentration (m)
A : surface area (m^2)
c : suspended sediment concentration (kg/m^3)
$c_b^{(\ell)}$: averaged sediment concentration in the near bottom computational layer
$\bar{c}$: depth-averaged concentration (kg/m^3)
$c_{a,e}$: equilibrium concentration (kg/m^3)
c_b : near-bed concentration (kg/m^3)
c_{gel} : gelling concentration (kg/m^3)
c_v : consolidation coefficient of the soil
$c_x, c_y, c_\sigma, c_\theta$: propagation velocities in x, y and σ, θ space (m/s)
C_{x_0} : concentration at $x = x_0$ (g/m^3)
D : diameter of the sediment (mm); deposition rate (kg/m^{2}s)
$D^{(\ell)}$: deposition flux (kg/m^2/s)
D_* : dimensionless particle diameter
D_{50} : median grain size (m)
D_H, D_V : horizontal and vertical eddy diffusivities (m^2/s)
D_V^{back} : background turbulent eddy viscosity in vertical direction (m^2/s)
$E(\sigma,\theta)$: energy density spectrum
$E^{(\ell)}$: erosion flux (kg/m^2/s)
F_ζ, F_η : horizontal Reynold's stresses determined by using eddy viscosity concept (m/s^2)
h : water depth (m)
k : diffusion coefficient (m^2/s)
ℓ : length scale of the deformation process
M : erosion rate parameter (kg/m^{2}s)
$M^{(\ell)}$: user specified erosion parameter for sediment fraction (ℓ) (kg/m^{2}s)
M_i^t : mass in volume i at time t (g)
M_ξ, M_η : contributions due to external sources or sinks of momentum (m/s^2)
$N(\sigma,\theta)$: the action density spectrum
p : porosity of bed layer
p^w : total pore water pressure (kg/ms^2)
p_e : hydrostatic pore water pressure (kg/ms^2)
Pe : Peclet number
P_x, P_y : horizontal pressure terms, which is given by Boussineq approximation (kg/m^{2}s^2)

Ri : Richardson number
Re : Reynolds
Re_e : effective Reynolds number
S : source or sink term per unit area; salinity (ppt)
$S_{\max}$: maximal salinity at which $w_{s,\max}^{(\ell)}$ is specified (ppt)
$S\left(\tau_{cw}, \tau_{cr,e}^{(\ell)}\right)$: erosion step function:

77

$$S\left(\tau_{cw}, \tau_{cr,e}^{(\ell)}\right) = \left(\frac{\tau_{cw}}{\tau_{cr,e}^{(\ell)}} - 1\right), \text{ when } \tau_{cw} > \tau_{cr,e}^{(\ell)},$$

$$= 0 \qquad \text{when } \tau_{cw} \le \tau_{cr,e}^{(\ell)}.$$

$S\left(\tau_{cw}, \tau_{cr,d}^{(\ell)}\right)$: deposition step function:

$$S\left(\tau_{cw}, \tau_{cr,d}^{(\ell)}\right) = \left(1 - \frac{\tau_{cw}}{\tau_{cr,d}^{(\ell)}}\right), \text{ when } \tau_{cw} < \tau_{cr,d}^{(\ell)},$$

$$= 0 \qquad \text{when } \tau_{cw} \ge \tau_{cr,d}^{(\ell)}$$

W_s: sediment settling velocity (m/s)
Q_{x_0} : flow at x = x0 (m^3/s)
U, V : Generalized Lagrangian Mean (GLM) velocity components (m/s)
$\bar{U}, \bar{V}$: Depth-averaged GLM velocity components (m/s)
U_m: mean velocity in the fluid mud layer (m/s)
u_*: shear velocity (m/s)
V : velocity of the deformation process
v: kinematic viscosity (m^2/s)
v_m: viscosity of fluid mud (m^2/s)
υ_V : vertical eddy viscosity (m^2/s)
v_{x_0} : velocity at x = x0 (m/s)
W_s: settling velocity (m/s)
$W_{s,b}$: settling velocity of the sediment at the bed (m)
$w_s^{(\ell)}$: fall velocity (hindered) (m/s)
$w_{s,0}^{(\ell)}$: non-hindered settling velocity of sediment fraction (ℓ)
$w_{s,f}^{(\ell)}$: fresh water settling velocity of sediment fraction (ℓ) (m/s)
$w_{s,\max}^{(\ell)}$: settling velocity of sediment fraction (ℓ) at salinity concentration SALMAX (m/s)
β : Rouse number
δ_b: thickness of bed layer (m)
δ_m: thickness of fluid mud layer (m)
δ_{nb}: thickness of near-bed layer (m)
ε_s^{ℓ} : vertical sediment mixing coefficient for sediment fraction (ℓ) (m^2/s)

ε_f : vertical fluid mixing coefficient calculated by turbulence closure model (m^2/s)
κ : von Kármàn constant
ρ : local fluid density (kg/m^3)
ρ_0 : reference density of water (kg/m^3)
ρ_m : density of fluid mud (kg/ms^2)
ρ_s : density of solid sediment particles (kg/m^3)
ρ_s^ℓ : specific density of sediment fraction (ℓ) (kg/m^3)
ρ_w : specific density of water with salinity concentration S (kg/m^3)
ζ : water surface elevation above reference datum (m)
σ : externally applied stress; vertical “sigma” coordinate
σ_{mol} : Prandtl-Schmidt number for molecular mixing
σ_T : turbulent Prandtl-Schmidt number
τ_B : Bingham strength of fluid mud (N/m^2)
τ_b : bed shear stress (N/m^2)
$\tau_{b,w}$: wave-induced bottom shear stress (N/m^2)
$\tau_{b,f}$: tidal flow-induced bottom shear stress (N/m^2)
$\tau_{c,w}$: critical shear stress for erosion of bed layer (N/m^2)
$\tau_{c,f}$: critical shear stress for erosion of near-bed layer (N/m^2)
τ_{cw} : mean bed stress due to current and waves as calculated by the wave-current interaction model (N/m^2)
$\tau_{cr,e}^{(\ell)}$: user specified critical erosion shear stress (N/m^2)
$\tau_{cr,d}^{(\ell)}$: user specified critical deposition shear stress (N/m^2)
τ_d : critical shear stress for deposition (N/m^2)
τ_e : critical shear stress for erosion (N/m^2)
τ_v : yield strength (N/m^2)
υ_{mol} : kinematic viscosity (molecular) coefficient (m^2/s)
ω : vertical velocity component in sigma coordinate system (s^{-1})
ϕ : fineness factor of sediment
$\left.\frac{\partial C}{\partial x}\right|_{x_0}$: concentration gradient at $x = x_0$ (g/m^4)
(ℓ) : sediment fraction (ℓ)
Δt : time step (s)