

DELFT UNIVERSITY OF TECHNOLOGY

MSC THESIS

Final Report



Author:

Remko NIJZINK

Supervising Committee:

Prof. dr. ir. H.H.G. SAVENIJE

Prof. dr. Z.B. WANG

ir. K. KUIJPER

J.I.A. GISEN MSc.

October 3, 2013

IMPROVING THE PREDICTIVE EQUATION FOR DISPERSION IN ESTUARIES

By:
Remko NIJZINK

Supervising Committee:
Prof. dr. ir. H.H.G. SAVENIJE
Prof. dr. Z.B. WANG
ir. K. KUIJPER
J.I.A. GISEN MSc.

THESIS FOR THE DEGREE OF MASTER OF
SCIENCE

DELFT UNIVERSITY OF TECHNOLOGY



List of Symbols

A_0	=	Area at estuary mouth [L^2]
a	=	Convergence length of the cross-section [L]
a_1	=	Convergence length of the cross-section in the first part of the estuary [L]
a_2	=	Convergence length of the cross-section in the second part of the estuary [L]
a^*	=	Weighed convergence length [L]
B	=	Estuary width [L]
b	=	Convergence length of the width [L]
C	=	Chézy friction coefficient [$L^{0.5}/T$]
c	=	Wave celerity [L/T]
c_0	=	Wave celerity at estuary mouth [L/T]
D_0	=	Dispersion coefficient at estuary mouth [L^2/T]
D_x	=	Dispersion coefficient at location x [L^2/T]
E	=	Tidal excursion [L]
f	=	Dimensionless friction $[-]$
g	=	Gravitational acceleration [L/T^2]
H_0	=	Tidal range at estuary mouth [L]
h	=	Mean depth [L]
h_0	=	Depth at estuary mouth [L]
K	=	van der Burgh coefficient $[-]$
$K_{manning}$	=	Manning friction coefficient [$L^{1/3}/T$]
N_r	=	Richardson number $[-]$
P_t	=	Flood volume [L^3]
Q_r	=	River discharge [L^3/T]
r_s	=	Storage width to stream width ratio $[-]$
S_0	=	Steady state salinity at estuary mouth [M/M]
S_x	=	Steady state salinity [M/M]
s	=	Salinity [M/M]
T	=	Tidal period [T]
u_r	=	Velocity river water [L/T]
v_Δ	=	Velocity density current [L/T]
x_1	=	Inflection point [L]
α_0	=	Mixing coefficient [L^{-1}]
α_c	=	Calibration coefficient Kuijper & Van Rijn (2011) $[-]$
$\Delta\rho$	=	Density difference [ML^{-3}]
δ	=	Damping factor [L^{-1}]
ϵ	=	Phase lag $[-]$
η	=	Tidal amplitude [L]
ζ	=	Correction factor Kuijper & Van Rijn (2011) $[-]$
θ	=	Correction factor Kuijper & Van Rijn (2011) [L^{-1}]
γ	=	Inverse convergence length Kuijper & Van Rijn (2011) [L^{-1}]
λ	=	Tidal wave length [L]
ρ	=	Density [ML^{-3}]
v_0	=	Velocity amplitude in estuary mouth [L/T]
v_Δ	=	Velocity of a density current [L/T]
ω	=	Angular velocity [T^{-1}]

Abbreviations

TA	=	Tidal Average
HW	=	High Water
HWS	=	High Water Slack
LW	=	Low Water
LWS	=	Low Water Slack
NS	=	Nash-Sutcliff efficiency
NSlog	=	Log Nash-Sutcliff efficiency
RMSE	=	Root Mean Squared Error

Abstract

Dispersion is often hard to incorporate in analytical salt intrusion models. The analytical models of Savenije (1993) and Kuijper & Van Rijn (2011) are quite similar and use a predictive equation for the dispersion in the estuary mouth. The biggest difference between the two models is the Van der Burgh K for which Kuijper & Van Rijn (2011) stated that it should be equal to 0.5. The main goal of this research is to improve the applicability of the analytical salt models.

The two models were applied to 72 measurements in 27 estuaries. Both models gave reasonable results, but Savenije's model gave in general a slightly better fit. The differences were found in especially the tail of the curve, what indicates that K is probably not equal to 0.5.

Linear regressions were carried out in order to derive new possible predictive equations for the dispersion coefficient. Several existing dimensionless ratios were combined in different regressions. A different regression technique, genetic programming, confirmed that a linear combination of the log of the dimensionless ratios is correct. The linear regressions were carried out for both models and for the estuary mouth and the inflection point, where the shape of the estuary changes. Many of the derived equations showed however more or less comparable results. The significance of the different terms was tested to see if each term contributed significantly. In this way it was already possible to reduce the number of possible new predictive equations.

A selection of the predictive equations was applied locally to investigate if the predictive equations would show similar patterns as the dispersion used by the two models. Kuijper's model showed more similar patterns as it implicitly uses a dispersion relation of the same structure as the predictive equation. This dispersion relation was the start assumption in the derivation of this model.

The same selection of equations was applied to the salt models. Eventually a choice of a new predictive equation was made based on the regressions, the local applications and the applications in the salt models. The horizontal to vertical tidal range and a friction term should be added to the Richardson number to get an improved predictive equation. The applicability of the analytical models will also increase by starting the calculation from the more clearly defined inflection point.

A disadvantage of the new equation is especially the friction term, because friction is often not known a priori. The hydraulic model of Cai et al. (2012) was used to test if friction and depth could be estimated with just a minimum of information. It was possible to use these equations to make a "quick and dirty" estimate of these parameters. The new proposed approach is therefore as follows:

- Determine the location of the inflection point based on information about the geometry
- Estimate hydraulic parameters, also with help of the equations of Cai et al. (2012)
- Determine the dispersion coefficient with the new predictive equations
- Determine the salt distribution with the model of Savenije (1986) or Kuijper & Van Rijn (2011)

Preface

This master thesis was written as final part of the master Civil Engineering, track Water management for Delft University of Technology. The work on this thesis started in the last few months of 2012 and has been very challenging. I truly hope that my work adds something interesting to the topic of salt intrusion, which will stay an important topic in my view. I also think that the work is not finished yet and I hope new students keep on working on the challenges of salt intrusion.

An important challenge of this research was also the field survey that was conducted in Malaysia in February and March 2013. The assistance of the University Teknologi Malaysia (UTM) was extremely valuable. They helped with equipment and transport, but also provided personnel that joined us on our boats. Professor Khairi has a great team and I would like to thank him and his team for all the assistance during the data collection.

The link between TU Delft and UTM is actually one of my committee members Jacqueline Gisen. She did a great deal with arranging everything in Malaysia and also by taking care of the "Dutchies" in Malaysia. Next to that, as committee member she had valuable advices.

The two persons that are mentioned most in this thesis are ir. Kees Kuijper and prof. dr. ir. Hubert Savenije. Their two salt models were part of this research and I was therefore very happy to have them in my committee as well. They helped me a lot with comments and remarks to improve my work. Professor Savenije actually proposed this topic to me and I am very happy I accepted this as my thesis topic. In general I would like to thank my whole committee for their support, input and feedback. I think this master thesis improved a lot by them and I hope the results are interesting for anyone interested in salinity in estuaries.

Remko Nijzink
Delft, October 2013

Contents

1	Introduction	1
1.1	Mixing processes	1
1.2	Longitudinal dispersion	2
1.3	Hydraulic and geometric estuary characteristics	2
1.4	Savenije's and Kuijper's model	3
1.5	Scope and research questions	5
1.6	Methodology	6
1.7	Outline	7
2	Comparing the models	9
2.1	Performance of the models	9
2.2	Performance of the predictive equations	10
3	Linear regression	13
3.1	Dimensionless ratios	13
3.2	Derived equations	14
3.3	Performance of the equations	16
3.4	Significance testing	18
4	Linear regression D_1	23
4.1	Dimensionless ratios	23
4.2	Derived equations	23
4.3	Performance of the equations	25
4.4	Significance testing	25
5	Genetic Programming	31
5.1	Initialization	31
5.2	Objective functions	31
5.3	Optimization step	31
5.4	New generation	32
5.5	Derived equations	32
5.6	Performance of the equations	33
6	Simplifying the equations	39
6.1	Optimizing for D_0	39
6.2	Optimizing for D_1	41
7	Local Validity	43
7.1	Selected equations	43
7.2	Performance of the equations	43
8	Predictive Mode	47
8.1	Performance of the equations	47
8.2	Starting from the inflection point	47
9	Hydraulic calculations	51
9.1	Estimating damping and wave celerity	51
9.2	Equations for friction and depth	51
9.3	Estimation of the estuary parameters	55

10 Conclusions and Recommendations	61
10.1 Savenije's model	61
10.2 Kuijper's model	62
10.3 The new approach	63
10.4 The models compared	64
10.5 Further considerations	64
A Tables	69
B Salinity Graphs	73
C Graphs Local Validity Savenije	83
D Graphs Local Validity Kuijper	93
E Graphs Predictive Mode Savenije	103
F Graphs Predictive Mode Kuijper	113

List of Figures

1.1	Geometry characteristics on a semi-log scale	3
1.2	Schematization of a trumpet shaped estuary with x_0 the estuary mouth, x_1 the inflection point and E the tidal excursion	4
2.1	Salinity curve Bernam and Maputo estuary, with the root mean squared errors (RMSE)	9
2.2	Predictions and calibrations of D_0 for left Savenije ($R_{NS} = 0.56$) and right Kuijper and van Rijn ($R_{NS} = 0.52$)	10
2.3	Root mean squared error LWS	11
2.4	Root mean squared error HWS	12
3.1	R^2 -values for the derived equations in dimensionless form for the analytical model of Savenije	17
3.2	R^2 -values for the derived equations in dimensionless form for the analytical model of Kuijper	17
3.3	Efficiency for the derived equations in a dimensional form for the analytical model of Savenije	19
3.4	Efficiency for the derived equations in a dimensional form for the analytical model of Kuijper	19
3.5	P-values for the linear regressions with the model of Savenije	21
3.6	P-values for the linear regressions with the model of Kuijper	21
3.7	P-values for the linear regression with new ratios for the model of Savenije	22
3.8	P-values for the linear regression with new ratios for the model of Kuijper	22
4.1	R^2 -values for the derived equations for the analytical model of Savenije	26
4.2	R^2 -values for the derived equations for the analytical model of Kuijper	26
4.3	R_{NS} values for the derived equations in dimensional form for the analytical model of Savenije	27
4.4	R_{NS} values for the derived equations in dimensional form for the analytical model of Kuijper	27
4.5	P-values for the linear regressions with the model of Savenije	29
4.6	P-values for the linear regressions with the model of Kuijper	29
5.1	The crossover operation, from Babovic & Keijzer (2000)	32
5.2	Nash-Sutcliffe efficiency for the equations derived for Savenije's model	34
5.3	Nash-Sutcliffe efficiency for the equations derived for Kuijper's model	34
5.4	Nash-Sutcliffe log efficiency for the equations derived for Savenije's model	37
5.5	Nash-Sutcliffe efficiency for the equations derived for Kuijper's model	37
6.1	Performance for the optimized equations for Savenije's model	40
6.2	Performance for the optimized equations for Kuijper's model	40
6.3	Performance of the optimized equations for D_1 for Savenije's model	42
6.4	Performance of the optimized equations for D_1 for Kuijper's model	42
7.1	The predictive equations applied locally for Landak, Pungue and Edisto with left the model of Savenije and right the model of Kuijper	45

8.1	Log Nash-Sutcliff efficiency for the equations derived for Savenije's model used in a predictive manner	49
8.2	Log Nash-Sutcliff efficiency for the equations derived for Kuijper's model used in a predictive manner	49
8.3	The salt models applied for Kurau, Pungue and Sinnamary for left Savenije and right Kuijper and Van Rijn	50
9.1	Place-time diagram per high water, low water and (if observed), slack	52
9.2	Place-time diagram per high water, low water and (if observed), slack	53
9.3	Place-time diagram, with different celerities for the two parts in the estuaries	54
9.4	Estimated, modeled and observed depth	56
9.5	Estimated, modeled and observed friction	56
9.6	Estimated, modeled and observed slack times	58
9.7	Estimated, modeled and observed tidal excursion	59
10.1	The new predictive equation versus the calibrated values, $R_{NS} = 0.68$	62
10.2	The new predictive equation versus the calibrated values, $R_{NS} = 0.82$	64

List of Tables

3.1	Correlation coefficients between the dimensionless ratios	13
3.2	Correlation coefficients between the log of dimensionless ratios	14
3.3	Results linear regression	14
4.1	Results linear regression D_1	23
5.1	Settings in genetic programming	32
5.2	Results genetic programming	33
5.3	Results genetic programming with a Nash-Sutcliff log efficiency	35
A.1	Used data in the analyses	70

1 Introduction

Many people in the world live close to alluvial estuaries. Salt intrusion can be a problem for agriculture and drinking water in these estuaries. It is therefore important to have models to understand the processes that influence salt intrusion, which allows us to analyze the effect of environmental changes or human impacts. Numerical models are quite capable of calculating the salt intrusion in estuaries, but there will always be a need for verification of these models. Next to that, analytical models create in general a better understanding of the different processes that play a role in the salt intrusion process. A preferably simple analytical model can therefore be useful for verification and the understanding of the different processes. Several authors derived one-dimensional salt intrusion models for tidal averaged situations. Among them Prandle (1985), Savenije (1986) and Kuijper & Van Rijn (2011). The dispersion processes are however hard to incorporate in these models in a predictive manner. The definition of dispersion is in many cases already not clear. Dispersion is not a real physical process, and must not be confused with diffusion. It is actually just a mathematical artifact, counteracting advective mixing mechanisms. Several approaches exist to deal with dispersion in analytical salt models.

1.1 Mixing processes

An approach used by many authors is decomposing all the mixing processes in different small scale subprocesses. For example, Fischer (1972) distinguished transverse net circulation, vertical net circulation, transverse oscillatory shear and vertical oscillatory shear. In general, tide driven mixing and density-driven mixing can be distinguished.

When decomposing the mixing mechanisms a bit more, several mechanisms can be distinguished. One of them is tidal trapping. Salt water is trapped in small volumes, dead ends and irregularities along the estuary and will mix with water of a different density when the flow reverses. MacVean & Stacey (2010) worked recently on tidal trapping and concluded that advection was the main driving mechanism of trapping, instead of the common assumption that it's diffusion-driven. Another mixing mechanism in (more stratified) estuaries is tidal shear, see also Prandle (1985) and Simpson et al. (1990). Turbulent mixing due to bottom friction is another mixing mechanism taken into account by for example Bowden (1981).

Residual currents in an estuary also cause mixing, McCarthy (1993) dealt with this intensively. He used an exponentially varying width of the estuary. He concluded that tidal buoyancy transport, with his definition actually a landward tidal advection, is important in the estuary mouth. More landward, horizontal diffusive buoyancy transport, or mixing just because of the difference in density, takes over.

One other mixing mechanism related to residual currents is so-called tidal-pumping, as mentioned by Fischer (1976). Ebb and flood channels behave like tidal pumps and influence the mixing of salt and fresh water. Nguyen (2008) derived an analytical, but partly empirical, relation to deal with tidal pumping. He derived an equation for the dispersion coefficient due to tidal-pumping. The sum of the dispersion coefficients due to tidal-pumping and gravitational mixing should give the total dispersion coefficient.

Mixing due to density differences is often referred to as gravitational circulation or buoyancy driven mixing. Savenije (1993) showed that gravitational mixing can be described by a simple hydrostatic model, where the forces of the salt and fresh water front exercise a momentum that drives mixing.

1.2 Longitudinal dispersion

A different approach is by using a longitudinal dispersion coefficient that takes all the mixing processes into account. When decomposing all the mixing processes it is very hard to define which process is dispersion and which process is advection. The definition of dispersion is not well defined in this way. Dispersion is nothing else as averaging different advective mixing processes, and this is scale dependent. The water particle that changes his density actually always undergoes advective processes, but when zooming out more and more of these processes will be defined as dispersion. A more general definition used is that the longitudinal dispersion takes all the mixing processes into account that counterbalance the advective mixing caused by the incoming fresh water discharge of the river. The longitudinal dispersion coefficient can be a function of x in several ways, like Prandle (1981) stated:

$$D_x = D_0$$

$$D_x \propto \frac{\partial s}{\partial x}$$

$$D_x \propto \left(\frac{\partial s}{\partial x} \right)^2$$

This can be summarized as:

$$D_x \propto \left(\frac{\partial s}{\partial x} \right)^k$$

1.3 Hydraulic and geometric estuary characteristics

Authors like McArthy (1993) and Savenije (1986) made use of an exponentially varying width to describe the shape of an estuary. This approach will also be adopted in the current research. The depth h_0 is assumed to be constant, what leads to an exponentially varying cross-section. The so-called convergence length a is therefore an important parameter to describe this exponential shape. With A_0 the cross-section in the estuary mouth the geometry of an estuary can be defined:

$$A(x) = A_0 e^{-x/a} \tag{1.1}$$

In quite some cases two parts can be distinguished within the estuary and the estuary then has a kind of trumpet shape. The first part near the estuary mouth is mainly wave-dominated and has a short convergence length. The second part starts in general a few kilometers inland and has a longer convergence length.

The location where the geometry changes can be easily identified when the cross-sections are plotted on a semi-log scale, like is done in Figure 1.1. This point can also rather often already be identified from a top view, Figure 1.2 shows this. We define here this point as the inflection point or x_1 . In this report we will adopt a subscript 0 for the estuary mouth and a subscript 1 for the inflection point. In general, the x-axis will start from the estuary mouth and continue landward.

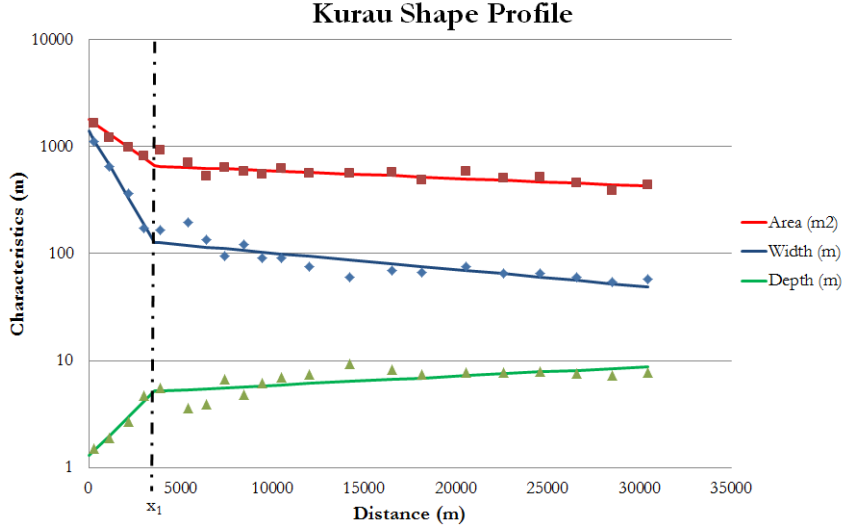


Figure 1.1: Geometry characteristics on a semi-log scale

The estuarine hydraulics are also not unimportant for salt intrusion and dispersion. Analytical hydraulic models are derived by for example Van Rijn (2011) and Cai et al. (2012), which all use more or less the same definitions of hydraulic parameters. The tidal range H is defined as the difference between the high water and low water levels. This tidal range can be damped or amplified more inland from the estuary mouth. The tidal damping is therefore defined as:

$$\delta = \frac{1}{H} \frac{dH}{dx}$$

The time it takes for the incoming or outgoing current to reverse after high water (HW) and low water (LW) is the so-called phase lag (ϵ). The moments the currents reverse are called high water slack (HWS) or low water slack (LWS). It can be observed at these moments that the water does not flow in a particular direction and the water looks stagnant (it slacks). Thus, a water particle actually travels a certain distance between HWS and LWS. Thus, the maximum salt intrusion also occurs during HWS and the minimum during LWS. This traveled distance is the tidal excursion E , or horizontal tidal range. This tidal excursion is in general fairly constant, but can also be damped or amplified.

1.4 Savenije's and Kuijper's model

Savenije (1986) used Van der Burg's equation to derive an analytical salt intrusion model, see Equation 1.2. Van der Burgh derived this equation based on

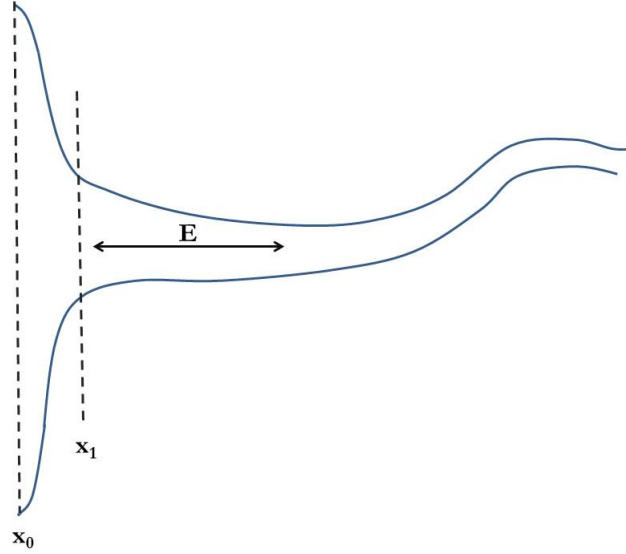


Figure 1.2: Schematization of a trumpet shaped estuary with x_0 the estuary mouth, x_1 the inflection point and E the tidal excursion

a number of measurements in the Rotterdam Waterway for a long period. He stated that the derivative of the dispersion was proportional to the flow velocity and a constant coefficient K (the Van der Burgh coefficient).

$$\frac{\partial D}{\partial x} = K u_r \quad (1.2)$$

The analytical salt model of Savenije (1986) is presented in Equation 1.3. It can be seen that it uses the convergence length a , the river discharge Q_r and cross-section in the estuary mouth A_0 to calculate the salinity S . The dispersion coefficient in the estuary mouth D_0 and the Van der Burg K are the calibration parameters in this model. A low value of K would indicate that tide-driven mixing is dominant, a high value of K would indicate that density-driven mixing is the more dominant process. Savenije (1993) also derived empirical relations for K (Equation 1.5) and D_0 (Equation 1.6) to make the model predictive. Shaha & Cho (2011) argued that K is longitudinally dependent, and proposed K as a function of x . Kuijper & Van Rijn (2011) stated however that K should have a fixed value of 0.5 and proposed some adjustments of the model, see Equation 1.4.

$$\frac{S_x}{S_0} = \left[1 - \frac{K a |Q_r|}{D_0 A_0} \left(e^{x/a} - 1 \right) \right]^{1/K} \quad (1.3)$$

$$\frac{S_x}{S_0} = \left[1 - \frac{1}{2} \frac{a |Q_r| \zeta}{D_0 A_0} \left(e^{x/a \zeta} - 1 \right) \right]^2 \quad (1.4)$$

The biggest difference between the two models is the use of the constant value of 0.5 for the Van der Burgh K and the introduction of the correction factor ζ , which is a function of damping and shape.

$$\zeta = \frac{1}{1 - \theta a}$$

$$\theta = \frac{\delta}{2} + \frac{3\gamma}{2} + \frac{1}{2a}$$

$$\gamma = \frac{1}{b} - \frac{1}{a}$$

The damping δ mentioned here should be determined with a hydraulic model, like Van Rijn (2011) or Cai et al. (2012), these however also use the Chézy roughness as input. Furthermore, γ is a shape parameter that takes the convergence of the cross-section (a) and the convergence of the width (b) into account.

Both models can be used for high water slack, low water slack and the tidally averaged (TA) situation. We will adopt the TA approach and the HWS and LWS curves can then be obtained by shifting the curve over the half of the tidal excursion length.

The models of Kuijper & Van Rijn (2011) and Savenije (1993) use in principle the same predictive equation for the dispersion coefficient at the estuary mouth D_0 , see Equations 1.6 and 1.7. Savenije choose to use the estuarine Richardson number, but Kuijper and Van Rijn decomposed this in a flow ratio, which is important for the level of stratification according to Prandle (1985), and a ratio of the propagation velocity of a density current (v_Δ) and the maximum tidal flow velocity (velocity amplitude) v_0 . A friction parameter was also added in the predictive equation of Kuijper & Van Rijn (2011). Both predictive equations also make use of a weighed convergence length a^* in case of two convergence lengths.

$$K = 0.2 * 10^{-3} \left(\frac{E}{H} \right)^{0.65} \left(\frac{E}{C^2} \right)^{0.39} (1 - \delta b)^{-2.0} \left(\frac{b}{a} \right)^{0.58} \left(\frac{Ea}{A_0} \right)^{0.14} \quad (1.5)$$

$$D_0 = 1400 N_r^{0.5} \frac{E}{a^*} v_0 h_0 \quad (1.6)$$

$$D_0 = \alpha_c 60 \sqrt{\pi} \left(\frac{v_\Delta}{v_0} \right) \left(\frac{C}{\sqrt{g}} \right) \left(\frac{u_r}{v_0} \right)^{0.5} v_0 h_0 \frac{E}{a^*} \quad (1.7)$$

1.5 Scope and research questions

The main goal of this research is to improve the predictive applicability of an analytical salt model. Predicting the dispersion is the most uncertain factor, and especially the predictive equations for the dispersion will need improvement. To achieve the goals, several questions will have to be answered:

- Can the Van der Burgh K be left out of the equations and have a constant value?
- Can we derive a new predictive equation for the dispersion coefficient?
- Can we apply local dispersion coefficients?

- Can the salt calculations be started from the inflection point of an estuary?
- Can we estimate hydraulic parameters in an analytical way?

1.6 Methodology

Data from many estuaries with different characteristics worldwide were needed to improve the predictive dispersion equation. A summary of the data can be found in Table A.1. Data was used for 27 estuaries worldwide with in total 72 measurements for high water slack (HWS) and low water slack (LWS). Much data was available from old measurement campaigns in for example Mozambique and Asia. New measurements were carried out from February till April 2013 in the Malaysian estuaries Kurau, Endau and Perak in addition to another Malaysian measurement campaign in 2012.

The two models of Savenije (1993) and Kuijper & Van Rijn (2011) were applied to the measurements to answer the question whether the Van der Burgh K is a constant or not. This would also lead to conclusions about the applicability, performance and ease of use of these two methods.

The existing equation was derived by a linear regression for the situation in the estuary mouth. The same approach was used to derive new predictive equations, assuming therefore that the structures of the old predictive equations of Savenije (1993) and Kuijper & Van Rijn (2011) were correct. The dimensionless ratios were first put in a log-format and different combinations of ratios were used in the linear regressions. A test statistic can be derived for each individual term in a linear regression to judge whether the contribution of this term is significant. This test statistic was derived for each ratio in each regression to see if the derived regression actually makes sense. The same analyses were carried out for the so-called inflection point of the estuary. Here the shape of the estuary changes, and it is a more clearly defined point than the estuary mouth. The practical applicability of a predictive equation might improve when all estuaries use the same starting point in the estuary.

A rather new technique that can be used to improve or derive equations is genetic programming or symbolic regression. The structure of the old predictive equation assumed a linear relation between the log of the ratios. Symbolic regression is a regression technique without a first assumption of the structure of the relation. A search algorithm that is based on the principles of natural selection tries to find the best suitable equation structure for the available data.

An important assumption made by Kuijper & Van Rijn (2011) is that the predictive equation for dispersion should be locally valid. The derived equations were therefore applied locally by plotting them along the length of the estuary. The outcomes of the derived predictive equations could then be compared with the dispersion that comes from the calibrated models. Eventually, conclusions about the local validity of the predictive equations could be drawn.

The new equations should also be tested. The new equations were applied to the measurements to see if the equations produce a reasonable fit of the salt curves. It was also tested if starting from the inflection point is possible and if it gives good results.

Some hydraulic parameters are hard to determine in advance. The practical applicability of a predictive equation is dependent on the fact if the used parameters are easy to estimate. The hydraulic model of Cai et al. (2012) was used

to estimate some hydraulic parameters like friction. The hydraulic parameter estimates were also compared with observations to draw conclusions about the applicability of this approach.

1.7 Outline

First a comparison of the two models will be presented in Chapter 2. The third chapter will deal with the linear regressions carried out for Kuijper and Van Rijn's model and Savenije's model. The fourth chapter will change the location of the regressions to the inflection point of the estuary, but the analysis will be the same. Chapter 5 will deal with genetic programming. The derived equations of Chapters 3, 4 and 5 will be simplified and optimized in Chapter 6. A selection of equations will be applied locally and also applied in the salt models of Kuijper and Savenije in the two subsequent Chapters. Eventually, Chapter 9 will deal with the hydraulic estimations of some parameters. To conclude, a new approach of calculating salt will be presented in Chapter 10.

The appendices contain a table with the parameters used in the calculations, the calibrated graphs of both salt models, graphs of the local dispersion and graphs of the two models used in a predictive manner.

2 Comparing the models

The two existing analytical salt models of Savenije (1986) and Kuijper & Van Rijn (2011) were compared to answer the question if K is indeed constant. Data from 27 estuaries worldwide with a total of 72 measurements for high water slack (HWS) and low water slack (LWS) were used in this comparison.

2.1 Performance of the models

Examples of the calculations for both models can be seen in Figure 2.1 for the Bernam and Maputo estuary. A summary of the used data and the calibration results can be found in Table A.1.

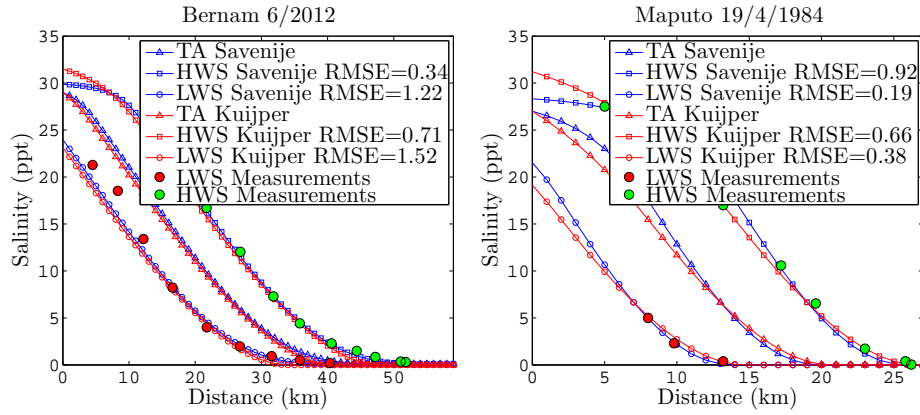


Figure 2.1: Salinity curve Bernam and Maputo estuary, with the root mean squared errors (RMSE)

It was very well possible to apply both models to all the available measurements. The root mean squared errors (RMSE) are summarized in Figures 2.3 and 2.4. It can be seen that just in a few cases these errors become large and the RMSE for Savenije's case are in general a bit lower. Also from just visual inspection of the individual graphs (see Appendix B) it can be noted that fitting the curves is in fact always possible. Applying the model of Savenije is however a bit easier. The model is simpler and doesnot use the correction factor ζ . To use Kuijper's model in a correct way first a hydraulic model must be used to calculate the correction factor ζ .

The Van der Burgh K is a factor that determines the shape of the salt curve, especially in the toe of the curve. The fact that Kuijper & Van Rijn (2011) fixes K to 0.5 leads therefore to differences between the two models in mainly the last part of the curves. The fit of the model was in general a bit better for Savenije's case, and the statement that K should have a fixed value of 0.5 is therefore very likely to be wrong.

However, differences can also be noted in the first part of the curve. This is probably caused by the correction factor ζ , which is multiplied with the convergence length in the exponential function. This correction factor and the

fixed Van der Burgh K make it therefore a bit harder to obtain equally good results with Kuijper's model.

The differences between both models are in most cases minor, it was generally possible to use both models with good results. Savenije's model had slight better fits. The total RMSE for HWS and LWS is also lower for Savenije's case than for Kuijper's case. In Savenije's case the RMSE is $1.49ppt$, for Kuijper's case $1.93ppt$. The difference in performance is therefore rather small. In some cases the model of Kuijper and Van Rijn needed different calibration values to give an equally good fit as the model of Savenije. The model of Kuijper still has the advantage of just one calibration parameter. Next to that, the physical meaning of K is still open for debate.

2.2 Performance of the predictive equations

The predictions for D_0 are plotted versus the calibrated values in Figure 2.2. Kuijper's predictive equation gives a higher scatter compared to Savenije's predictive equation. The two equations are quite similar, but the constant is different and a friction term is added in Kuijper's case. Kuijper's model uses different calibration values, because it has one degree of freedom less (K is fixed to 0.5). The two predictive equations show however that in general the same measurements are outliers.

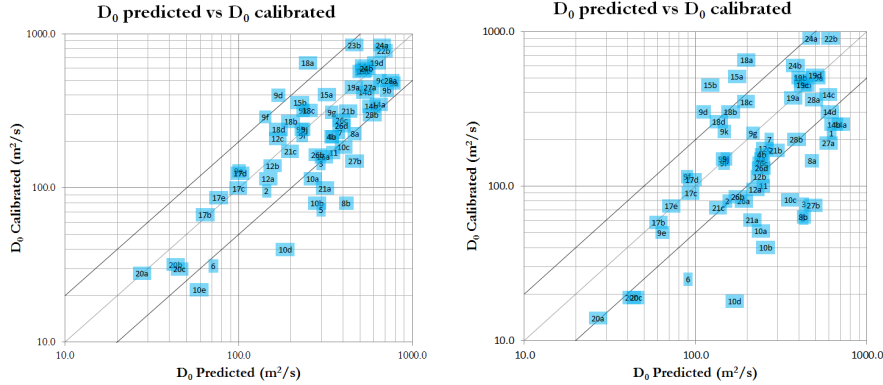


Figure 2.2: Predictions and calibrations of D_0 for left Savenije ($R_{NS} = 0.56$) and right Kuijper and van Rijn ($R_{NS} = 0.52$)

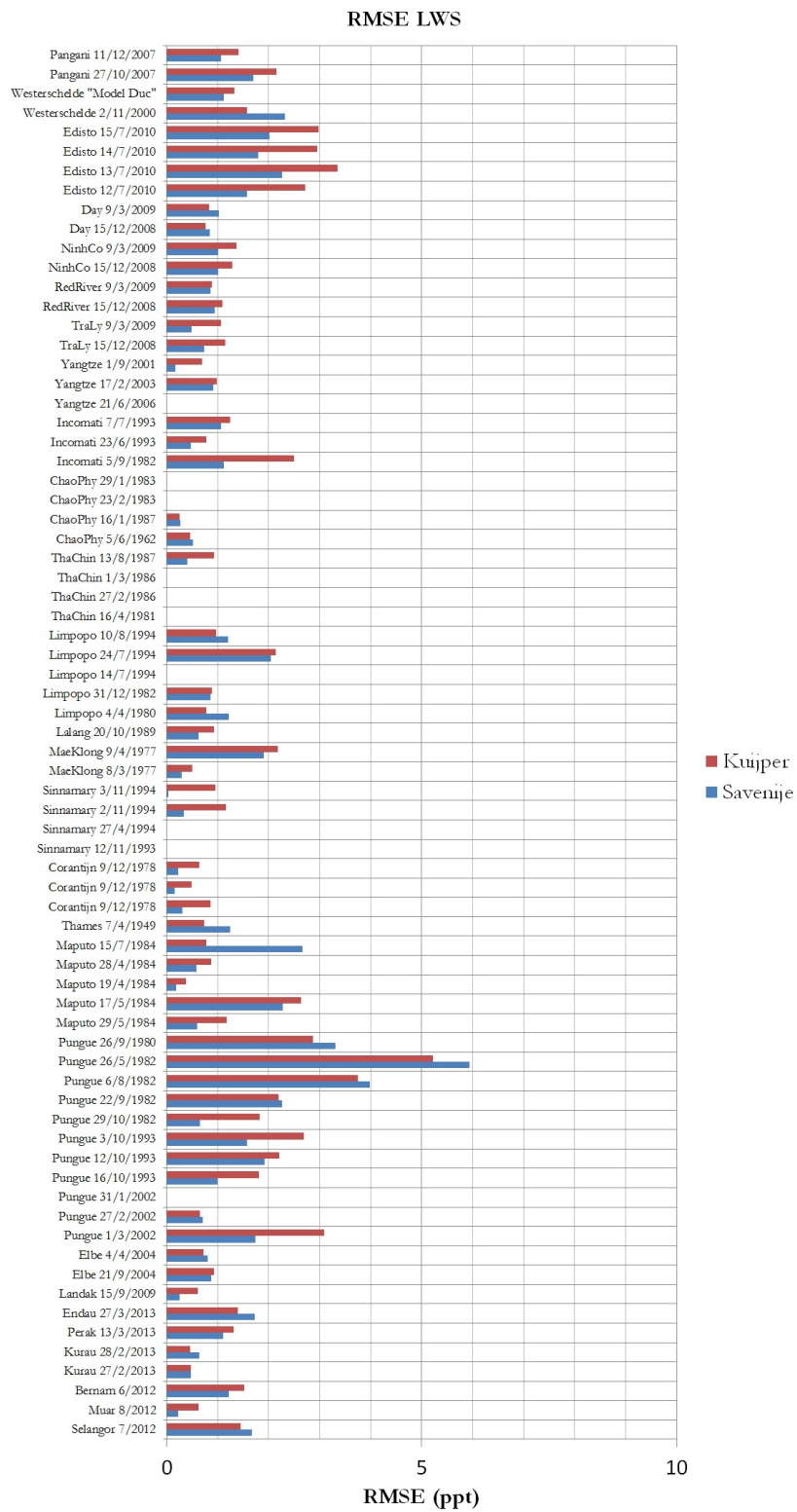


Figure 2.3: Root mean squared error LWS

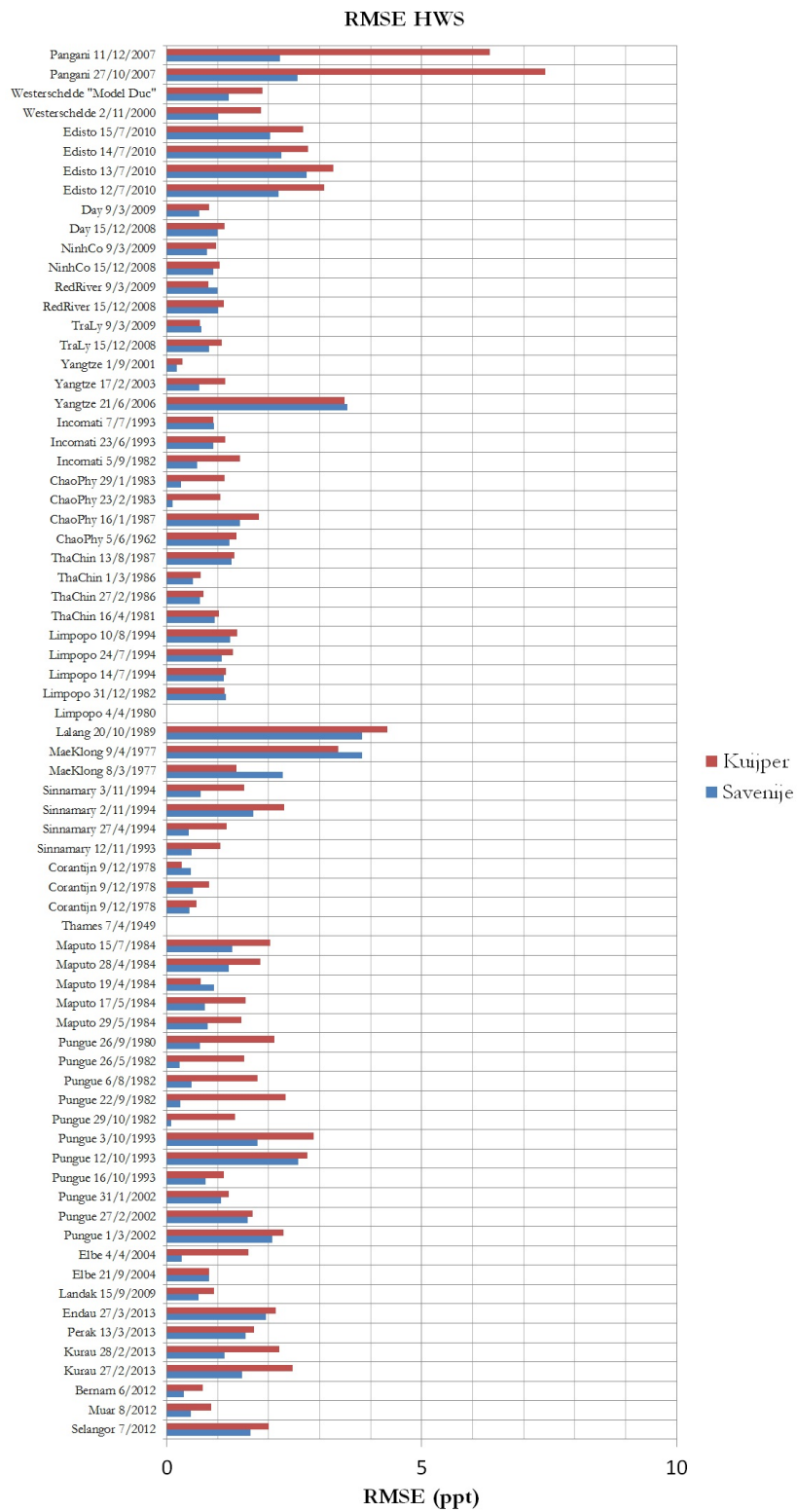


Figure 2.4: Root mean squared error HWS

3 Linear regression

A linear regression was carried out in order to derive a predictive dispersion equation based on a tidally averaged situation. To do so, the log values of several dimensionless ratios were used. The dependent variables were varied and the regression was done for the two available salt models, which both use different calibrated values for D_0 .

3.1 Dimensionless ratios

A linear regression should preferably be carried out with an independent set of input variables to avoid spurious relations. The correlation between some already existing dimensionless ratios was therefore calculated. The uncorrelated ratios (so with a correlation coefficient of approximately 0) can be combined in a linear regression. This led to the following possible combinations, based on Table 3.2:

- $N_r, \frac{C^2}{g}$ and $\frac{H}{E}$
- $\frac{B_1}{h_0}$ and $\frac{C^2}{g}$
- $N_r, \frac{C^2}{g}$ and $\frac{\lambda}{a}$
- $\frac{h_0}{a}, \frac{C^2}{g}$ and $\frac{H}{E}$

Table 3.1: Correlation coefficients between the dimensionless ratios

	$\frac{D_0}{v_0 E_0}$	$\frac{C^2}{g}$	N_r	$\frac{h}{a}$	$\frac{H}{E}$	$\frac{\lambda}{a}$	$\frac{B_1}{h_0}$	$\frac{\lambda}{E}$
$\frac{D_0}{v_0 E_0}$		-0.07	0.96	-0.34	0.02	-0.25	-0.37	0.60
$\frac{C^2}{g}$	-0.07		-0.11	-0.07	-0.13	-0.32	-0.23	-0.15
N_r	0.96	-0.11		-0.37	-0.02	-0.28	-0.29	0.59
$\frac{h}{a}$	-0.34	-0.07	-0.37		0.43	0.69	0.42	-0.27
$\frac{H}{E}$	0.02	-0.13	-0.02	0.43		0.64	0.49	0.42
$\frac{\lambda}{a}$	-0.25	-0.32	-0.28	0.69	0.64		0.66	-0.04
$\frac{B_1}{h_0}$	-0.30	-0.23	-0.29	0.42	0.49	0.66		-0.24
$\frac{\lambda}{E}$	0.60	-0.15	0.59	-0.27	0.42	-0.04	-0.24	

It is also possible to derive new dimensionless ratios with the Buckingham PI-theorem. The dispersion coefficient D_0 is the dependent variable, and the following 12 variables are probably influencing this dependent variable:

$$v_0, E_0, C, g, \rho, \Delta\rho, h, B, Q_f, T, a, H$$

These variables use three dimensions: [L], [T] and [M]. This means that $13 - 3 = 10$ dimensionless ratios can be formed. The variables v_0 , ρ and E_0 were chosen as base variables. A base variable represents one of the dimensions and the dimensions of these variables cannot be expressed into each other. Thus, E_0 represents length [L], v_0 represents time [T] and ρ represents mass [M]. The following dimensionless ratios were derived:

Table 3.2: Correlation coefficients between the log of dimensionless ratios

	$\frac{D_0}{v_0 E_0}$	$\frac{C^2}{g}$	N_r	$\frac{h}{a}$	$\frac{H}{E}$	$\frac{\lambda}{a}$	$\frac{B_1}{h_0}$	$\frac{\lambda}{E}$
$\frac{D_0}{v_0 E_0}$		-0.16	0.88	-0.45	0.06	-0.18	-0.46	0.69
$\frac{C^2}{g}$	-0.16		-0.01	-0.00	-0.19	-0.30	-0.14	-0.15
N_r	0.88	-0.01		-0.69	-0.19	-0.48	-0.72	0.61
$\frac{h}{a}$	-0.45	-0.00	-0.69		0.50	0.77	0.61	-0.26
$\frac{H}{E}$	0.06	-0.19	-0.19	0.50		0.67	0.41	0.46
$\frac{\lambda}{a}$	-0.18	-0.30	-0.48	0.77	0.67		0.69	0.10
$\frac{B_1}{h_0}$	-0.46	-0.14	-0.72	0.61	0.69	0.69		-0.26
$\frac{\lambda}{E}$	0.69	-0.15	0.61	-0.26	0.46	0.10	-0.26	

$$\frac{D_0}{v_0 v_0}, \frac{C\sqrt{E_0}}{v_0}, \frac{gE_0}{v_0^2}, \frac{\Delta\rho}{\rho}, \frac{h}{E_0}, \frac{A_0}{E_0^2}, \frac{Q_f}{v_0 E_0^2}, \frac{Tv_0}{E_0}, \frac{a}{E_0}, \frac{H}{E_0}$$

Combining ratio three till eighth gives again the estuarine Richardson number:

$$\frac{\Delta\rho}{\rho} * \frac{gE_0}{v_0^2} * \frac{h}{E_0} * \frac{Q_f}{v_0 E_0^2} * \frac{Tv_0}{E_0} * \frac{E_0^2}{A_0} = \frac{\Delta\rho}{\rho} \frac{ghQ_f T}{v_0^2 E_0 A_0} = N_r$$

Another important aspect is that one of the ratios is constant:

$$\frac{Tv_0}{E_0} = \pi$$

3.2 Derived equations

The derived formulas are presented in Table 3.3. The equations were derived for the same estuaries as used in Savenije (2005) and validated with new measurements from more recent years. The derivation dataset had 11 estuaries with 43 measurements, the validation dataset had 16 estuaries with 29 measurements.

Table 3.3: Results linear regression

Savenije	Kuijper and Van Rijn
$\frac{D_0}{v_0 E_0} = 1569.2 N_r^{0.56} \frac{h_0}{a} \quad (3.1)$	$\frac{D_0}{v_0 E_0} = 1469.6 N_r^{0.62} \frac{h_0}{a} \quad (3.2)$
$\frac{D_0}{v_0 E_0} = 8.0 N_r^{0.48} \left(\frac{h_0}{a} \right)^{0.46} \quad (3.3)$	$\frac{D_0}{v_0 E_0} = 0.71 N_r^{0.49} \left(\frac{h_0}{a} \right)^{0.21} \quad (3.4)$

$\frac{D_0}{v_0 E_0} = 42.2 N_r^{0.50} \left(\frac{h_0}{a}\right)^{0.48} \left(\frac{g}{C^2}\right)^{0.24} \quad (3.5)$	$\frac{D_0}{v_0 E_0} = 22.1 N_r^{0.53} \left(\frac{h_0}{a}\right)^{0.31} \left(\frac{g}{C^2}\right)^{0.42} \quad (3.6)$
$\frac{D_0}{v_0 E_0} = 90.4 N_r^{0.49} \left(\frac{h_0}{a}\right)^{0.24} \left(\frac{g}{C^2}\right)^{0.18} \left(\frac{E}{H}\right)^{0.39} \quad (3.7)$	$\frac{D_0}{v_0 E_0} = 28.5 N_r^{0.52} \left(\frac{h_0}{a}\right)^{0.11} \left(\frac{g}{C^2}\right)^{0.39} \left(\frac{H}{E}\right)^{0.28} \quad (3.8)$
$\frac{D_0}{v_0 E_0} = 0.25 N_r^{0.42} \left(\frac{g}{C^2}\right)^{0.16} \quad (3.9)$	$\frac{D_0}{v_0 E_0} = 0.88 N_r^{0.48} \left(\frac{g}{C^2}\right)^{0.38} \quad (3.10)$
$\frac{D_0}{v_0 E_0} = 6.44 N_r^{0.44} \left(\frac{H}{E}\right)^{0.48} \quad (3.11)$	$\frac{D_0}{v_0 E_0} = 2.64 N_r^{0.49} \left(\frac{H}{E}\right)^{0.38} \quad (3.12)$
$\frac{D_0}{v_0 E_0} = 0.000017 \left(\frac{E}{H}\right)^{0.81} \left(\frac{C^2}{g}\right)^{0.005} \quad (3.13)$	$\frac{D_0}{v_0 E_0} = 0.0004 \left(\frac{E}{H}\right)^{1.33} \left(\frac{g}{C^2}\right)^{1.40} \quad (3.14)$
$\frac{D_0}{v_0 E_0} = 25.3 N_r^{0.46} \left(\frac{H}{E}\right)^{0.54} \left(\frac{g}{C^2}\right)^{0.13} \quad (3.15)$	$\frac{D_0}{v_0 E_0} = 19.2 N_r^{0.51} \left(\frac{H}{E}\right)^{0.36} \left(\frac{g}{C^2}\right)^{0.37} \quad (3.16)$
$\frac{D_0}{v_0 E_0} = 0.14 \left(\frac{C^2}{g}\right)^{0.15} \left(\frac{h_0}{B_1}\right)^{0.56} \quad (3.17)$	$\frac{D_0}{v_0 E_0} = 0.46 \left(\frac{g}{C^2}\right)^{0.0003} \left(\frac{h_0}{B_1}\right)^{0.66} \quad (3.18)$
$\frac{D_0}{v_0 E_0} = 0.08 N_r^{0.48} \left(\frac{\lambda}{a}\right)^{0.29} \quad (3.19)$	$\frac{D_0}{v_0 E_0} = 0.07 N_r^{0.48} \left(\frac{\lambda}{a}\right)^{0.15} \quad (3.20)$
$\frac{D_0}{v_0 E_0} = 0.14 N_r^{0.50} \left(\frac{\lambda}{a}\right)^{0.31} \left(\frac{g}{C^2}\right)^{0.11} \quad (3.21)$	$\frac{D_0}{v_0 E_0} = 0.71 N_r^{0.50} \left(\frac{\lambda}{a}\right)^{0.14} \left(\frac{g}{C^2}\right)^{0.38} \quad (3.22)$
$\frac{D_0}{v_0 E_0} = 1.03 \left(N_r \frac{g}{C^2}\right)^{0.41} \quad (3.23)$	$\frac{D_0}{v_0 E_0} = 1.49 \left(N_r \frac{g}{C^2}\right)^{0.48} \quad (3.24)$

$\frac{D_0}{v_0 E_0} = 0.097 N_r^{0.40} \quad (3.25)$	$\frac{D_0}{v_0 E_0} = 0.092 N_r^{0.45} \quad (3.26)$
$\frac{D_0}{v_0 E_0} = 0.017 N_r^{0.36} \left(\frac{\lambda}{E} \right)^{0.54} \quad (3.27)$	$\frac{D_0}{v_0 E_0} = 0.068 N_r^{0.49} \left(\frac{\lambda}{E} \right)^{0.21} \quad (3.28)$
$\frac{D_0}{v_0 E_0} = 50.0 \left(N_r \frac{g}{C^2} \frac{H}{E} \right)^{0.43} \quad (3.29)$	$\frac{D_0}{v_0 E_0} = 185.9 \left(N_r \frac{g}{C^2} \frac{H}{E} \right)^{0.52} \quad (3.30)$

The results of the linear regression for the newly derived ratios are presented in Equations 3.31 for Savenije and 3.32 for Kuijper and Van Rijn.

$$\frac{D_0}{E_0 v_0} = 287.6 \left(\frac{C \sqrt{E_0}}{v_0} \right)^{-0.45} \left(\frac{g E_0}{v_0^2} \right)^{0.66} \left(\frac{\Delta \rho}{\rho} \right)^{-0.02} \left(\frac{h}{E_0} \right)^{0.99} \left(\frac{A_0}{E_0^2} \right)^{-0.50} \left(\frac{Q_f}{v_0 E_0^2} \right)^{0.42} \left(\frac{a}{E_0} \right)^{-0.33} \left(\frac{H}{E_0} \right)^{0.46} \quad (3.31)$$

$$\frac{D_0}{E_0 v_0} = 4023.5 \left(\frac{C \sqrt{E_0}}{v_0} \right)^{-0.92} \left(\frac{g E_0}{v_0^2} \right)^{0.54} \left(\frac{\Delta \rho}{\rho} \right)^{-0.27} \left(\frac{h}{E_0} \right)^{0.86} \left(\frac{A_0}{E_0^2} \right)^{-0.55} \left(\frac{Q_f}{v_0 E_0^2} \right)^{0.51} \left(\frac{a}{E_0} \right)^{-0.07} \left(\frac{H}{E_0} \right)^{0.32} \quad (3.32)$$

3.3 Performance of the equations

In Figure 3.1 and Figure 3.2 the R^2 -values for the derived equations are summarized. These R^2 -values are the values that were obtained during the regression, so with the equations written in a log-format and as dimensionless ratios. The adjusted R^2 -values are also presented in Figures 3.1 and Figure 3.2. The adjusted R^2 -values take also the number of terms into account to make a more fair comparison. Almost all derived equations have high R^2 -values, only Equations 3.13, 3.14, 3.17 and 3.18 have low values. These equations do not use the Richardson number and the Richardson number is apparently necessary to reach high R^2 -values. The other equations do not differ much in the achieved R^2 -values. Still Equation 3.1, which is the original format of Savenije (1986), gives a bit higher values compared to the other equations, but this is minor.

The equations were converted back to a dimensional form to see if the the derived equations were useful as predictive equation:

$$D_0 = E_0 v_0 f(x_1, \dots, x_n)$$

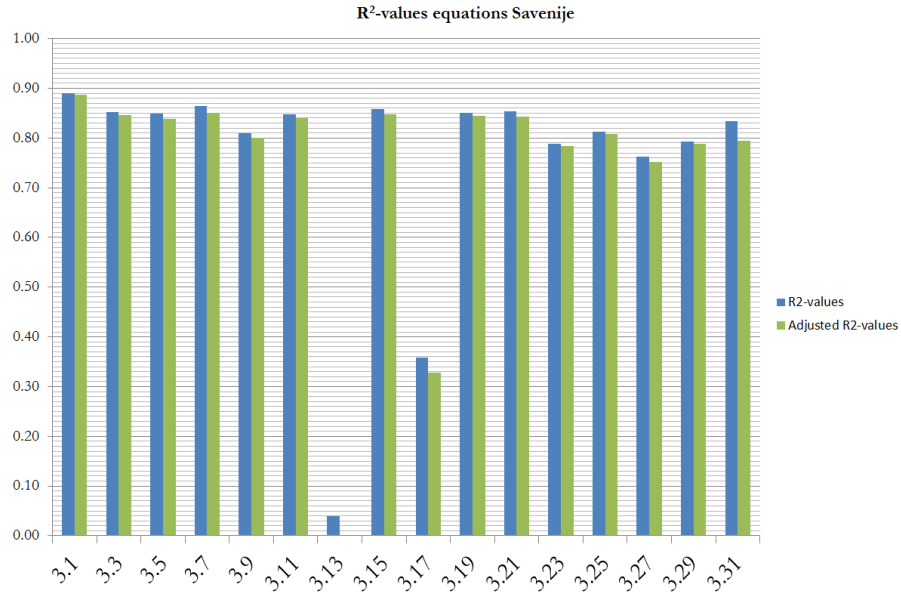


Figure 3.1: R^2 -values for the derived equations in dimensionless form for the analytical model of Savenije

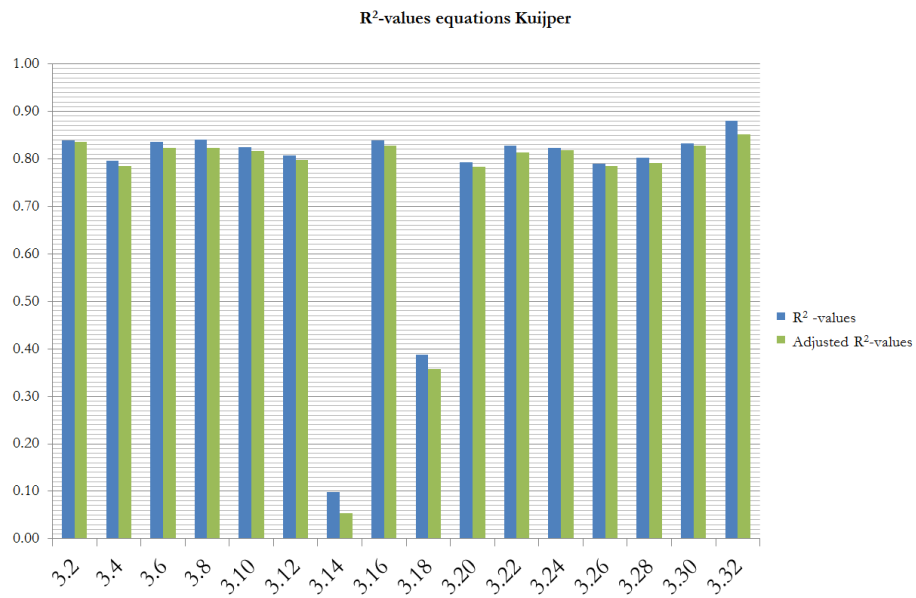


Figure 3.2: R^2 -values for the derived equations in dimensionless form for the analytical model of Kuiper

The relation is now not completely linear anymore and the several definitions of the R^2 -efficiency can not be expressed into each other anymore. Now, the most general definition was used, which is actually the same as the Nash-Sutcliffe efficiency:

$$R_{NS} = 1 - \frac{\sum (y_{modeled} - y_{observed})^2}{\sum (y_{observed} - \bar{y}_{observed})^2}$$

Figures 3.3 and 3.4 show the results for the calibration (11 estuaries with 43 measurements, see Table A.1 in the appendix) and validation (16 estuaries with 29 measurements). The differences between the equations become bigger and it can be clearly seen that Equations 3.7, 3.15, 3.19 and 3.21 and 3.29 have the best performance. The regression for Kuijper's model show a similar pattern. The equivalent equations 3.8, 3.16, 3.22 and 3.30 are also here the best ones.

Both models also show an increase in performance during the validation, it is however expected that the performances would decrease. The reason that this happens is not very clear, but it could indicate that the calibration and validation datasets are still too small.

3.4 Significance testing

As the performance of the equations dramatically drops by leaving out the Richardson number it might be interesting to look at the importance of the different terms in the equations. The errors of a linear regression are assumed to have a Student-T distribution, so now we can define the test statistic:

$$t = \frac{\beta_i}{SE_i}$$

With β_i the regression coefficient for the i -th term and SE_i the standard error for the i -th term.

The two hypotheses are defined as follows:

$$\begin{aligned} H_0 : \beta_i &= 0 \\ H_1 : \beta_i &\neq 0 \end{aligned}$$

A confidence level of $\alpha = 0.05$ was chosen, so for $P(T \leq t) < 0.025$ the null hypothesis was rejected. When the null hypothesis is not rejected the regression coefficient is more likely to be equal to zero, and the term in the equation does not really contribute to the result. The outcomes are summarized in Figures 3.5 and 3.6.

For the equations of Savenije the P-value of the friction term always exceeds the confidence level. Apparently, the friction term doesnot play an important role in the derived equations for Savenije's model, this can be caused by the Van der Burg K , which is friction dependent. Another reason for this might be the place in the estuary for which the regression is carried out. More seaward, the influence of the friction will probably be less. For Equation 3.7 the Richardson number alone contributes to the result, while the other terms exceed the confidence level. The same applies actually to 3.27, also here just the Richardson

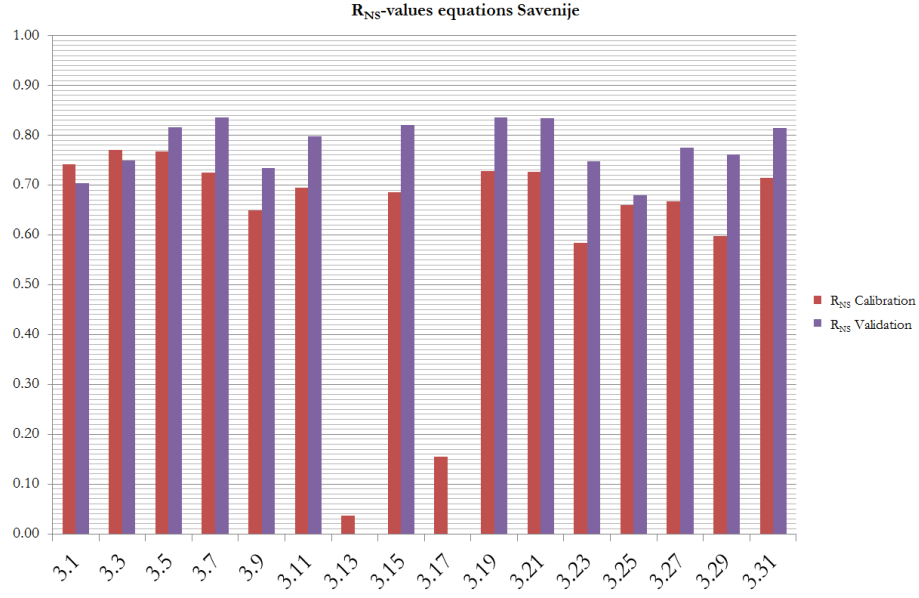


Figure 3.3: Efficiency for the derived equations in a dimensional form for the analytical model of Savenije

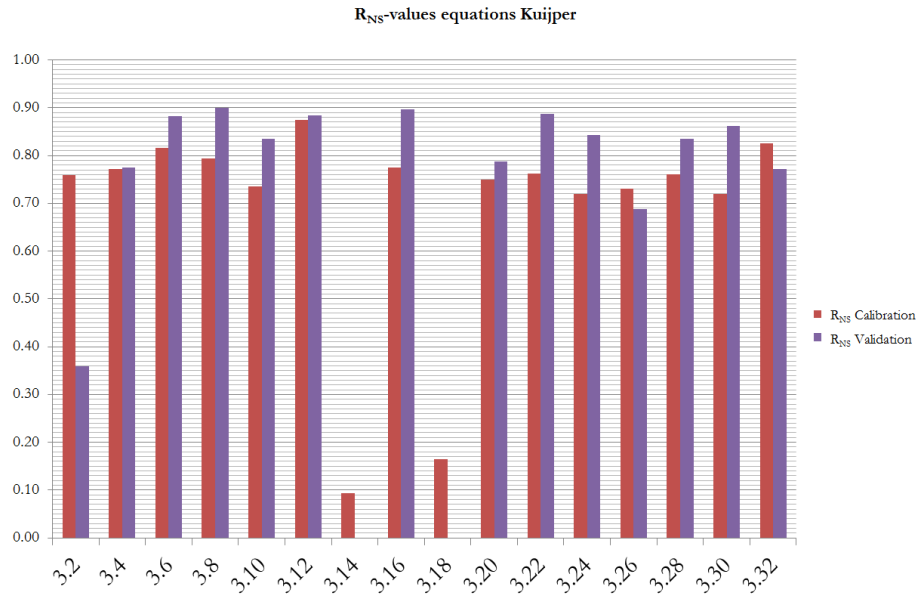


Figure 3.4: Efficiency for the derived equations in a dimensional form for the analytical model of Kuiper

number contributes. Equations 3.1, 3.3, 3.11 and 3.19 are now left as possible new predictive equations, 3.1 and 3.3 show however a strong decrease in performance during validation. It is also interesting to see that for 3.15 the friction term exceeds the confidence level, and neglecting this term here leads again to the same format as Equation 3.11. The same holds for Equation 3.21, when here the friction term is neglected 3.19 is found back. It is therefore very likely that the Richardson number should be combined with either $\frac{H}{E}$ or $\frac{\lambda}{a}$. Equations 3.5 and 3.15 however showed very good results during validation, these equations are also not exceeding the confidence level very much. Also Equations 3.23, 3.25 and 3.29 are still an option, the confidence level is not exceeded. However, these equations just use one term, and this term always contributes significantly. If we require that the new predictive equation should have no terms that exceed the confidence level and should have good results during the calibration and validation Equations 3.3, 3.11, 3.19 and 3.29 look the most promising.

One would expect to see the same patterns back for the model of Kuijper. However, the null hypothesis is now rejected for the friction term in Equation 3.6, 3.10, 3.16 and 3.22. For the model of Kuijper only 3.2 and 3.10 show good results, the results of Equation 3.12 are however still acceptable. Now the opposite holds for Equations 3.16 and 3.12 compared to Savenije's case. In Equation 3.16 $\frac{H}{E}$ drops out and we find Equation 3.12 back. So for Kuijper's model it seems that friction must be included instead of the horizontal to vertical tidal range ratio. Also here the one-term equations never exceed the confidence level, and Equations 3.24, 3.26 and 3.30 are still suitable to become the new predictive equation. If we set the requirements the same as for Savenije's model, we can see that Equations 3.10 and 3.30 are for Kuijper's model the most promising.

The P-values for the newly derived ratios can be seen in Figures 3.7 and 3.8. The biggest difference between both models is again that for Kuijper's model the null hypothesis for the friction term can be rejected, thus should be included, but this is again not the case for Savenije's model. For some ratios that are part of the Richardson number the null hypothesis cannot be rejected, what is rather surprising. For example the density ratio does not contribute significantly to the result. The derived coefficients in these equations and the P-values in Figures 3.7 and 3.8 indicate that the most important terms are the discharge ratio and the ratio with the cross-section in the mouth for both models.

The objective function values (as shown also in Figures 3.3 and 3.4) are however not as good for the equations with the new ratios as for the other derived equations.

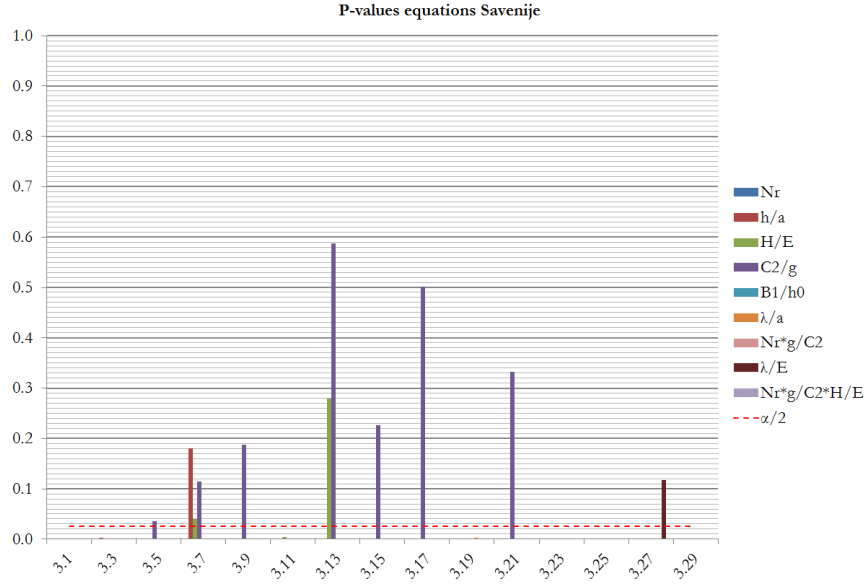


Figure 3.5: P-values for the linear regressions with the model of Savenije

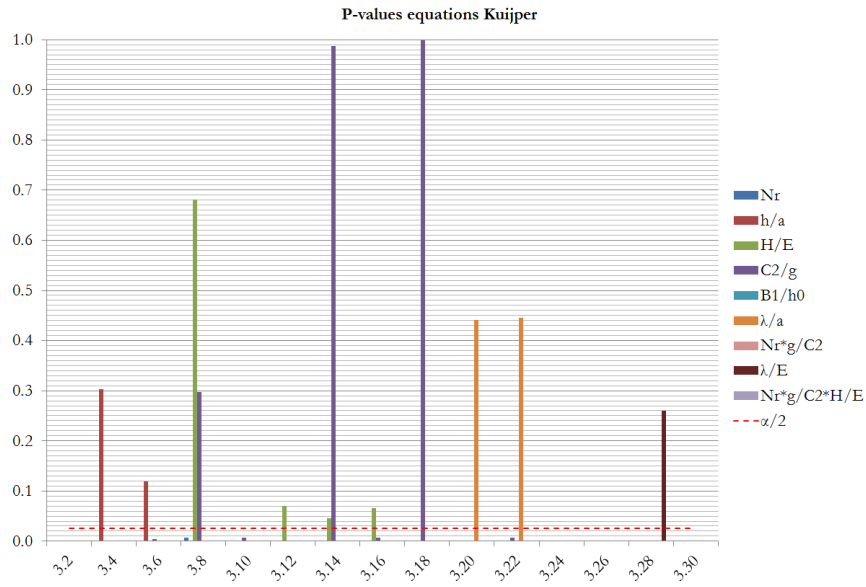


Figure 3.6: P-values for the linear regressions with the model of Kuijper

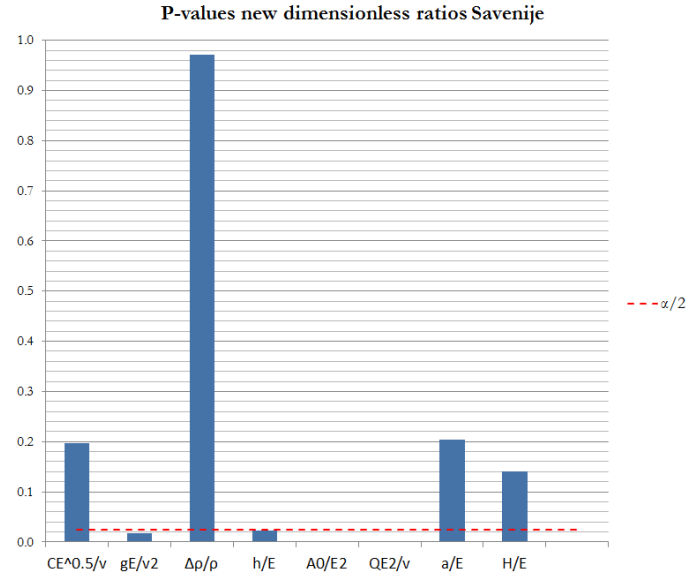


Figure 3.7: P-values for the linear regression with new ratios for the model of Savenije

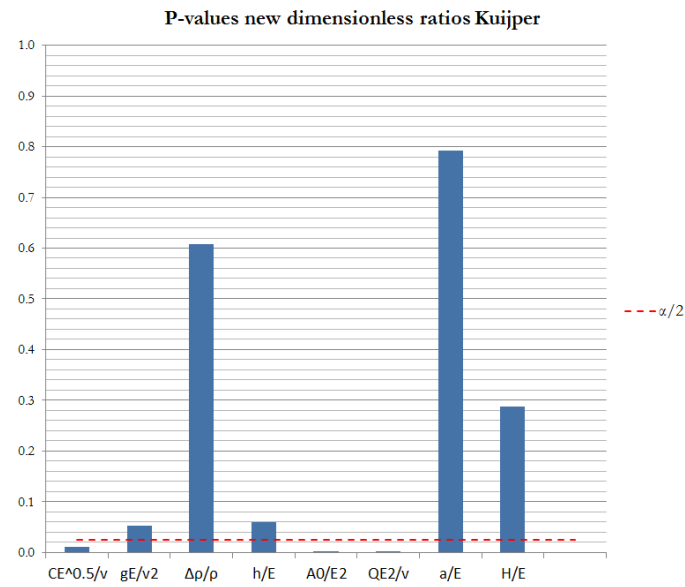


Figure 3.8: P-values for the linear regression with new ratios for the model of Kuijper

4 Linear regression D_1

A clearly defined point in estuaries is the inflection point where the convergence length of the estuary changes. It may be convenient to use this point as a starting point for the salt calculations and therefore linear regressions were carried out for the dispersion coefficient D_1 at this inflection point.

4.1 Dimensionless ratios

The same combinations of parameters were used in the analyses, but some of these parameters had to be calculated for x_1 :

$$H_1 = H_0 e^{\delta x_1}$$

$$E_1 = E_0 e^{\delta x_1}$$

$$v_1 = v_0 e^{\delta x_1}$$

Also the Richardson number need to be adjusted:

$$N_{r1} = \frac{\Delta \rho_1 g \bar{h} Q_f T}{\rho A_1 E_1 v_1^2}$$

So now we have the following combinations:

- $N_{r1}, \frac{C^2}{g}$ and $\frac{H_1}{E_1}$
- $\frac{B_1}{h_0}$ and $\frac{C^2}{g}$
- $N_{r1}, \frac{C^2}{g}$ and $\frac{\lambda}{a_2}$
- $\frac{h_0}{a_2}, \frac{C^2}{g}$ and $\frac{H_1}{E_1}$

4.2 Derived equations

The derived formulas are presented in Table 4.1. As stated before, the same combinations were used as for the regressions at the estuary mouth. The regressions were also carried out with the dimensionless ratios in a log-format.

Table 4.1: Results linear regression D_1

Savenije	Kuijper and Van Rijn
$\frac{D_1}{v_1 E_1} = 843.2 N_r^{0.50} \frac{h_0}{a} \quad (4.1)$	$\frac{D_1}{v_1 E_1} = 1585.3 N_r^{0.62} \frac{h_0}{a} \quad (4.2)$
$\frac{D_1}{v_1 E_1} = 0.00001 N_r^{0.44} \left(\frac{a}{h_0} \right)^{0.44} \quad (4.3)$	$\frac{D_1}{v_1 E_1} = 0.14 N_r^{0.41} \left(\frac{h_0}{a} \right)^{0.07} \quad (4.4)$

$\frac{D_1}{v_1 E_1} = 0.0018 N_r^{0.23} \left(\frac{a}{h_0}\right)^{0.41} \left(\frac{g}{C^2}\right)^{0.05} \quad (4.5)$	$\frac{D_1}{v_1 E_1} = 0.39 N_r^{0.42} \left(\frac{h_0}{a}\right)^{0.09} \left(\frac{g}{C^2}\right)^{0.12} \quad (4.6)$
$\frac{D_1}{v_1 E_1} = 0.009 N_r^{0.24} \left(\frac{a}{h_0}\right)^{0.56} \left(\frac{C^2}{g}\right)^{0.007} \left(\frac{H}{E}\right)^{0.40} \quad (4.7)$	$\frac{D_1}{v_1 E_1} = 1.12 N_r^{0.39} \left(\frac{a}{h_0}\right)^{0.18} \left(\frac{g}{C^2}\right)^{0.11} \left(\frac{H}{E}\right)^{0.44} \quad (4.8)$
$\frac{D_1}{v_1 E_1} = 0.41 N_r^{0.32} \left(\frac{g}{C^2}\right)^{0.32} \quad (4.9)$	$\frac{D_1}{v_1 E_1} = 0.14 N_r^{0.40} \left(\frac{g}{C^2}\right)^{0.10} \quad (4.10)$
$\frac{D_1}{v_1 E_1} = 0.20 N_r^{0.32} \left(\frac{H}{E}\right)^{0.13} \quad (4.11)$	$\frac{D_1}{v_1 E_1} = 0.77 N_r^{0.42} \left(\frac{H}{E}\right)^{0.27} \quad (4.12)$
$\frac{D_1}{v_1 E_1} = 0.0006 \left(\frac{E}{H}\right)^{0.49} \left(\frac{g}{C^2}\right)^{0.09} \quad (4.13)$	$\frac{D_1}{v_1 E_1} = 0.00006 \left(\frac{E}{H}\right)^{0.57} \left(\frac{C^2}{g}\right)^{0.16} \quad (4.14)$
$\frac{D_1}{v_1 E_1} = 2.28 N_r^{0.34} \left(\frac{H}{E}\right)^{0.18} \left(\frac{g}{C^2}\right)^{0.34} \quad (4.15)$	$\frac{D_1}{v_1 E_1} = 2.08 N_r^{0.42} \left(\frac{H}{E}\right)^{0.29} \left(\frac{g}{C^2}\right)^{0.14} \quad (4.16)$
$\frac{D_1}{v_1 E_1} = 0.33 \left(\frac{g}{C^2}\right)^{0.19} \left(\frac{h_0}{B_1}\right)^{0.30} \quad (4.17)$	$\frac{D_1}{v_1 E_1} = 0.09 \left(\frac{C^2}{g}\right)^{0.07} \left(\frac{h_0}{B_1}\right)^{0.36} \quad (4.18)$
$\frac{D_1}{v_1 E_1} = 0.14 N_r^{0.23} \left(\frac{a}{\lambda}\right)^{0.45} \quad (4.19)$	$\frac{D_1}{v_1 E_1} = 0.072 N_r^{0.40} \left(\frac{\lambda}{a}\right)^{0.04} \quad (4.20)$
$\frac{D_1}{v_1 E_1} = 0.25 N_r^{0.25} \left(\frac{a}{\lambda}\right)^{0.40} \left(\frac{g}{C^2}\right)^{0.11} \quad (4.21)$	$\frac{D_1}{v_1 E_1} = 0.13 N_r^{0.41} \left(\frac{\lambda}{a}\right)^{0.05} \left(\frac{g}{C^2}\right)^{0.11} \quad (4.22)$
$\frac{D_1}{v_1 E_1} = 0.42 \left(N_r \frac{g}{C^2}\right)^{0.32} \quad (4.23)$	$\frac{D_1}{v_1 E_1} = 0.70 \left(N_r \frac{g}{C^2}\right)^{0.39} \quad (4.24)$

$\frac{D_1}{v_1 E_1} = 0.06 N_r^{0.31} \quad (4.25)$	$\frac{D_1}{v_1 E_1} = 0.08 N_r^{0.39} \quad (4.26)$
$\frac{D_1}{v_1 E_1} = 0.06 N_r^{0.31} \left(\frac{\lambda}{E} \right)^{0.02} \quad (4.27)$	$\frac{D_1}{v_1 E_1} = 0.35 N_r^{0.46} \left(\frac{E}{\lambda} \right)^{0.39} \quad (4.28)$
$\frac{D_1}{v_1 E_1} = 10.3 \left(N_r \frac{g}{C^2} \frac{H_1}{E_1} \right)^{0.35} \quad (4.29)$	$\frac{D_1}{v_1 E_1} = 35.6 \left(N_r \frac{g}{C^2} \frac{H_1}{E_1} \right)^{0.43} \quad (4.30)$

4.3 Performance of the equations

In Figure 4.1 and Figure 4.2 the R^2 -values for the derived equations are again summarized. The values for R^2 were in general lower compared to the equations for D_0 . These lower values could be caused by the usage of derived values for x_1 , while for x_0 the parameters were mainly always obtained by direct measurement. The differences between the equations become now already a bit more distinct for the R^2 -values obtained with the regression. The equations that performed well for x_0 , had in general a good performance for x_1 as well.

The performance of the equations for the model of Savenije dropped rather strong after validation in dimensional form, while in Kuijper's case it stayed rather constant. Only Equation 4.29 gave also good results after validation. This can be seen in Figures 4.3 and 4.4. Kuijper's model gave better results for x_1 , and looked more consistent as the validation also kept on giving high R^2 -values. For Kuijper's model more equations kept on giving good results, like Equations 4.6, 4.10, 4.16, 4.22 and 4.30.

4.4 Significance testing

The difference between the performances of the equations for D_1 were still very small, thus the importance of the different terms was also tested again in the same manner as for x_0 .

The test statistic can be defined again:

$$t = \frac{\beta_i}{SE_i} \quad (4.31)$$

With β_i the regression coefficient for the i-th term and SE_i the standard error for the i-th term. The hypotheses are defined as:

$$\begin{aligned} H_0 : \beta_i &= 0 \\ H_1 : \beta_i &\neq 0 \end{aligned}$$

The outcomes are summarized in Figure 4.5 and 4.6. The friction terms for Savenije's equations exceeded the confidence level again always, just as for the regressions with D_0 . Actually none of the equations was fully accepted and

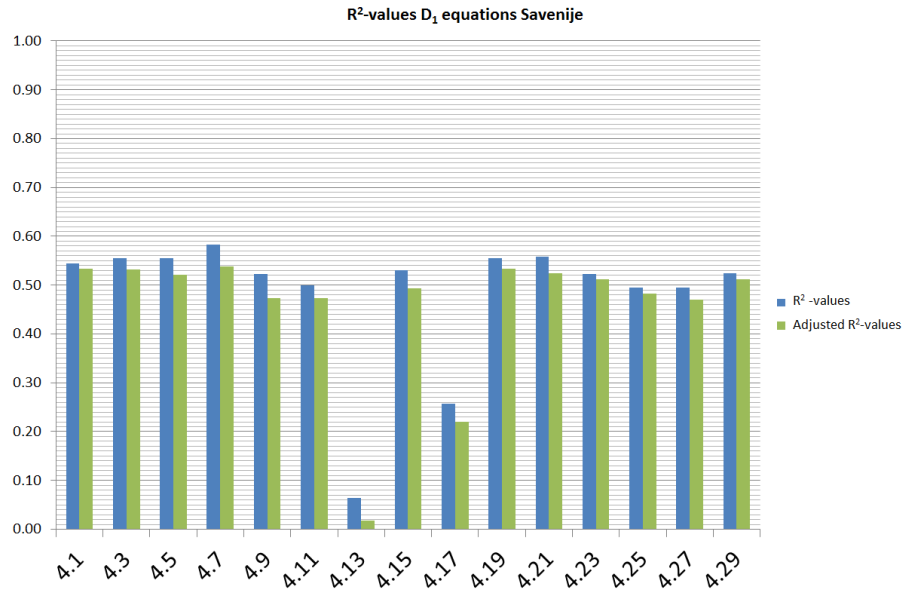


Figure 4.1: R^2 -values for the derived equations for the analytical model of Savenije

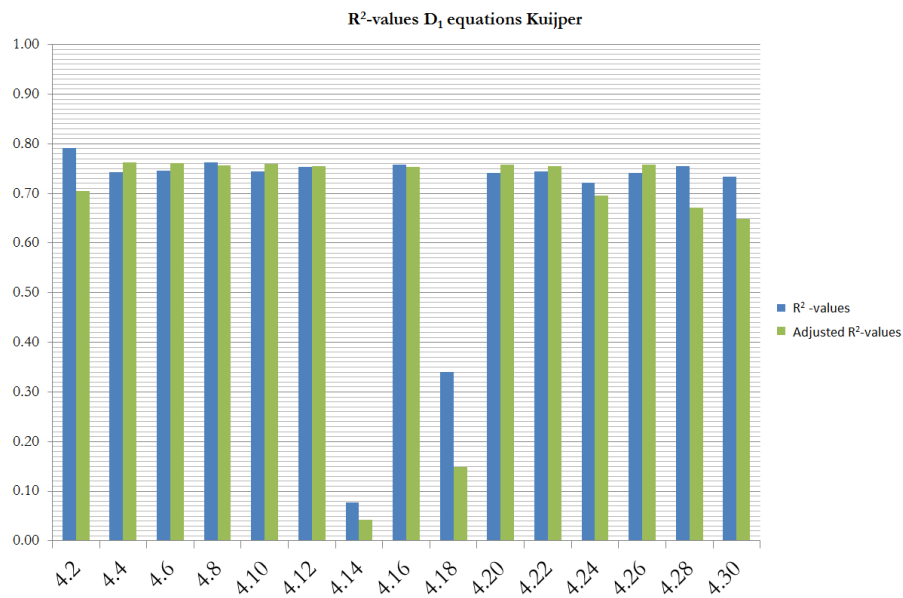


Figure 4.2: R^2 -values for the derived equations for the analytical model of Kuijper

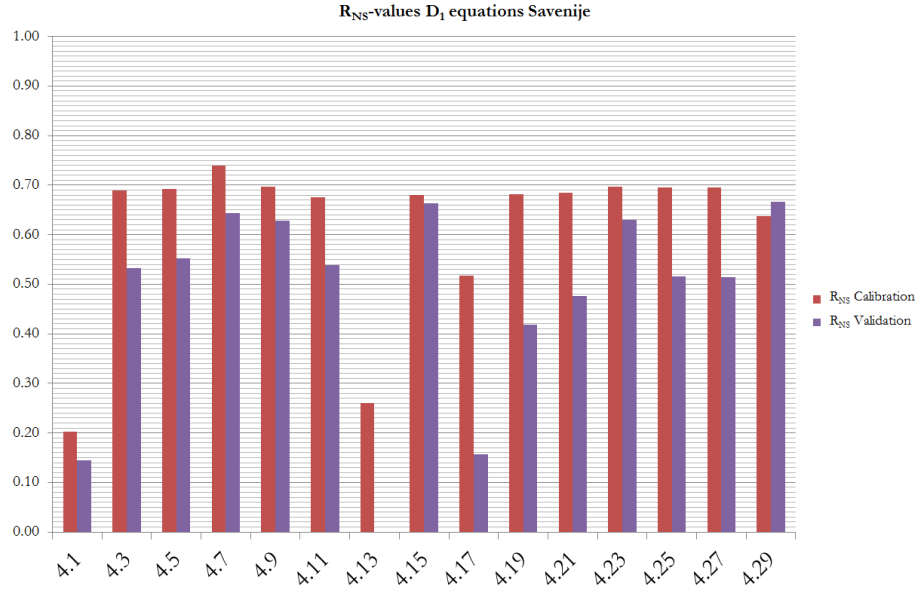


Figure 4.3: R_{NS} values for the derived equations in dimensional form for the analytical model of Savenije

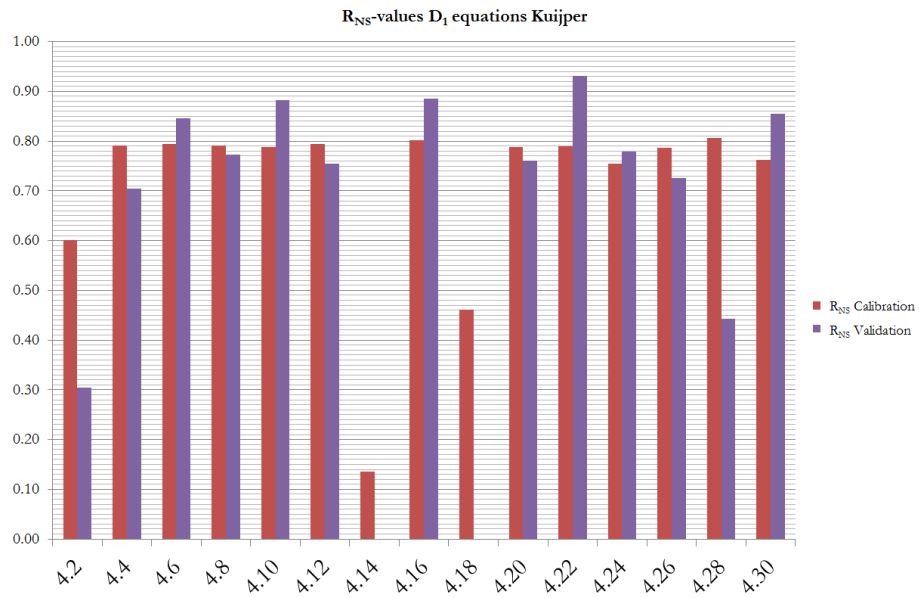


Figure 4.4: R_{NS} values for the derived equations in dimensional form for the analytical model of Kuiper

the most equations had at least one term that didnot contribute significantly. Furthermore only the one-term equations did not exceed the confidence level. The exceedence of the different terms looked in general much higher compared to the regressions carried out for the estuary mouth. Also for Kuijper's case none of the equations were fully accepted, except again for the one-term equations.

For the regression at the inflection point x_1 the exceedence of the friction terms was less compared to the situation for x_0 , but now $\frac{H}{E}$ exceeds the confidence level rather strong. This could indicate that the friction term gains influence when moving more inland, while the influence of the tidal range decreases. This statement is also supported by the fact that Equation 4.9 (with friction) has a better performance than Equation 4.11 (with tidal range) for x_1 , while the opposite was true for x_0 . The same holds for the equations derived for Kuijper's case.

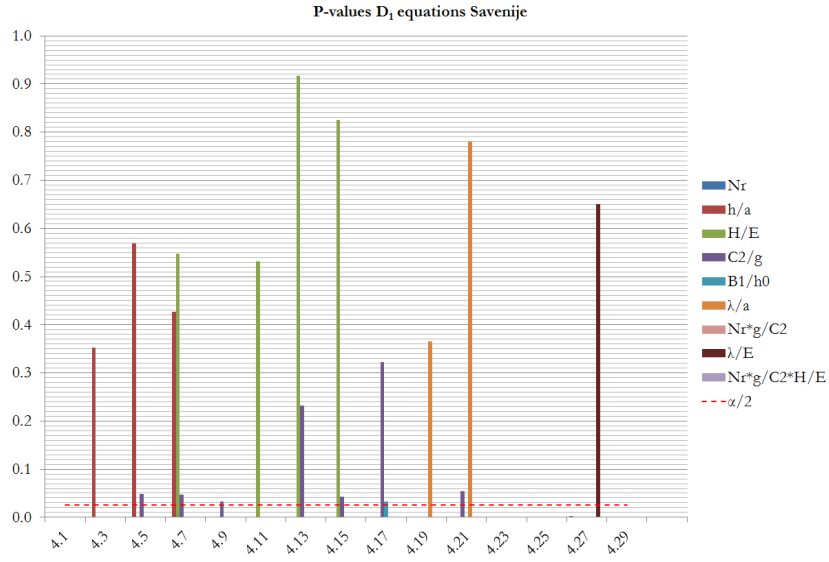


Figure 4.5: P-values for the linear regressions with the model of Savenije

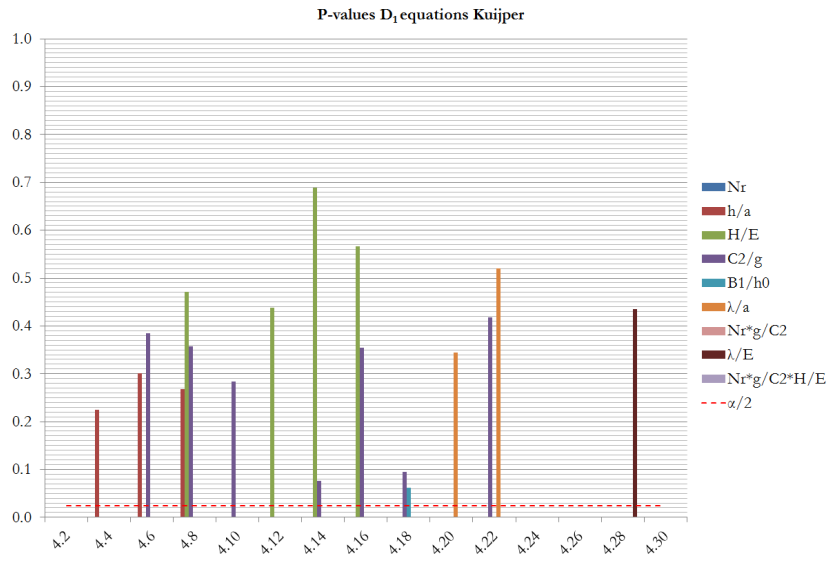


Figure 4.6: P-values for the linear regressions with the model of Kuijper

5 Genetic Programming

Genetic programming (GP) was introduced by Koza (1994) and afterwards it was used by several authors, like Babovic & Keijzer (2000). Genetic programming uses the principles of natural selection to derive new relations between variables. This technique is also referred to as symbolic regression. Functions are represented as trees, as can be seen in Figure 5.1. In this research GPTIPS from Searson et al. (2010) was used to perform symbolic regression analyses. The Matlab package was slightly modified to serve the goals of this research.

5.1 Initialization

First a so-called terminal set is created that exists out of the input variables and assumed constants. A function set defines all the possible functions that can be used, in our case plus, minus, times, log, exp and divide. The terminal set is composed out of (combinations of) the earlier defined dimensionless ratios. Three options exist to make a first generation:

- full
- grow
- ramped 1/2 and 1/2.

The full method fills the whole tree to the maximum depth, while the grow method creates trees of different sizes. The last option is a combination of both methods. This last method is widely used and also adopted in the current research.

5.2 Objective functions

A selection is made based on an objective function. A percentage of the best relations will be directly copied to the next generation, this is called elitism. When the existing relations are ranked to determine the elite group, the length of the relation will be taken into account. When two relations have the same length, the shortest relation will be ranked higher.

To select relations that will be used to create a new generation several methods exist. One often applied method is tournament selection. A number of relations will be randomly selected from the existing population. These selected relations are in a tournament and the relation with the best objective function value will win this tournament and will be used to create a new generation.

5.3 Optimization step

To make the algorithm a bit more efficient, an extra optimization step was added. In the original code only fixed numbers were added. This would mean that the algorithm could have found the correct format of the equation, but it would give it the wrong, fixed constant, what could lead to a bad performance of this equation. The constant should therefore be optimized before the algorithm starts looking for other equation structures. The method used to optimize the constants was the Nelder-Mead Simplex Method, as this is the standard

Table 5.1: Settings in genetic programming

Parameters GP	
Population size	100
Number of generations	500
Tournament size	50
Elite fraction	0.05
Crossover probability	0.9
Mutation probability	0.05
Reproduction probability	0.05
Tree depth	10
Build method	Ramped 1/21/2

optimization method for Matlab. It uses the simplex algorithm as described by Lagarias et al. (1998).

5.4 New generation

The new generation can be created by reproduction (transferring a complete relation to the new generation), mutation (changing a relation) or crossover (combining two relations). The choice of operation depends on a probability that is assigned to each operation. The new generation will replace the existing one and after a few generations the best relations will have survived.

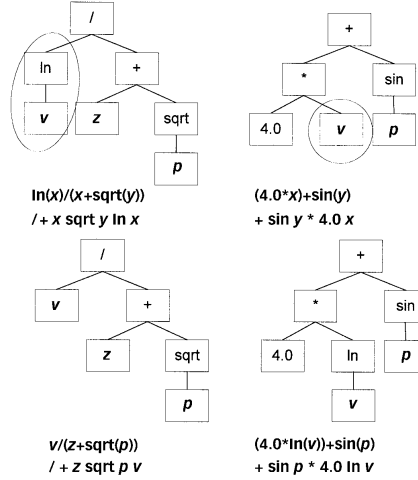


Figure 5.1: The crossover operation, from Babovic & Keijzer (2000)

5.5 Derived equations

A selection of the best results can be found in Table 5.2. The R_{NS} -values are summarized in Figure 5.2 and 5.3 .

Table 5.2: Results genetic programming

Savenije	Kuijper
$\frac{D_0}{v_0 E_0} = 0.14 N_r + 0.014$ (5.1)	$\frac{D_0}{v_0 E_0} = 0.13 N_r + 0.011$ (5.2)
$\frac{D_0}{v_0 E_0} = 0.14 N_r + \frac{h}{a} + 0.014$ (5.3)	$\frac{D_0}{v_0 E_0} = (-0.088 - N_r) \left(\frac{h}{a} - 0.13 \right)$ (5.4)
$\frac{D_0}{v_0 E_0} = 0.15 N_r + 77.8 \frac{h}{a}$ (5.5)	$\frac{D_0}{v_0 E_0} = 0.14 \left(\frac{h}{a} + N_r \right)$ (5.6)
$\frac{D_0}{v_0 E_0} = \frac{h}{a} + 1.13^{N_r} - 0.98$ (5.7)	$\frac{D_0}{v_0 E_0} = 1.01 - 0.85^{N_r}$ (5.8)
$\frac{D_0}{v_0 E_0} = 0.15 N_r + 0.0009 \frac{\lambda}{a}$ (5.9)	$\frac{D_0}{v_0 E_0} = N_r \left(\frac{0.21}{\frac{\lambda}{a}} + 0.02 \frac{\lambda}{a} \right)$ (5.10)
$\frac{D_0}{v_0 E_0} = 0.15 N_r$ (5.11)	$\frac{D_0}{v_0 E_0} = -0.015 \ln \left(\frac{h}{a} \right) N_r$ (5.12)
$\frac{D_0}{v_0 E_0} = (N_r + 0.06) * \frac{\frac{C^2}{g}}{2428.1}$ (5.13)	$\frac{D_0}{v_0 E_0} = 0.13 N_r + 39.2 \frac{H}{E}$ (5.14)
$\frac{D_0}{v_0 E_0} = \frac{N_r - 6.4}{\frac{H}{E}}$ (5.15)	$\frac{D_0}{v_0 E_0} = 0.14 N_r e^{\frac{H}{E}}$ (5.16)
$\frac{D_0}{v_0 E_0} = \left(0.16 + \frac{H}{E} \right) N_r$ (5.17)	$\frac{D_0}{v_0 E_0} = \frac{H}{E} + 0.14 N_r$ (5.18)
$\frac{D_0}{v_0 E_0} = 0.15 \left(N_r + \frac{0.17}{\frac{\lambda}{a}} \right)$ (5.19)	$\frac{D_0}{v_0 E_0} = 0.15 \left(e^{\frac{H}{E}} N_r \right)^{0.79}$ (5.20)

5.6 Performance of the equations

Equation 5.1 was found in more runs with the algorithm. This is not surprising as it is actually the result that also should be obtained when a linear regres-

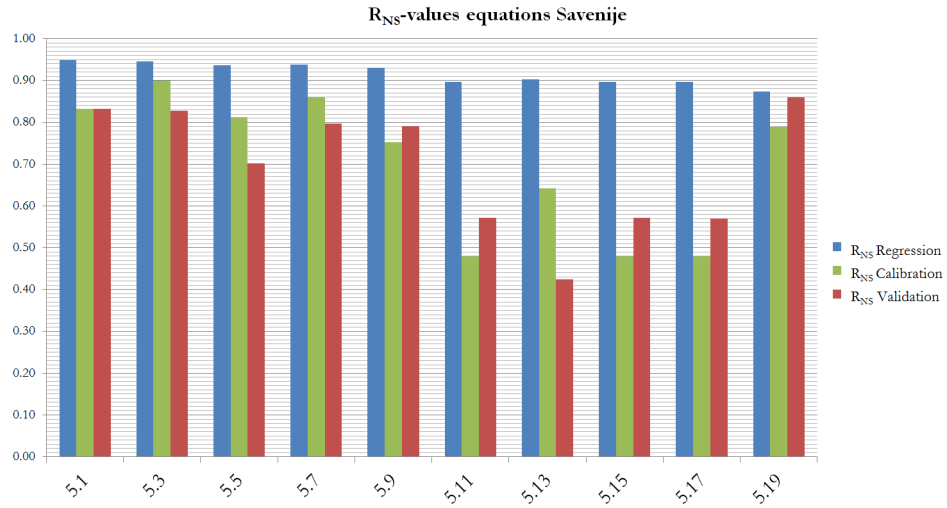


Figure 5.2: Nash-Sutcliff efficiency for the equations derived for Savenije's model

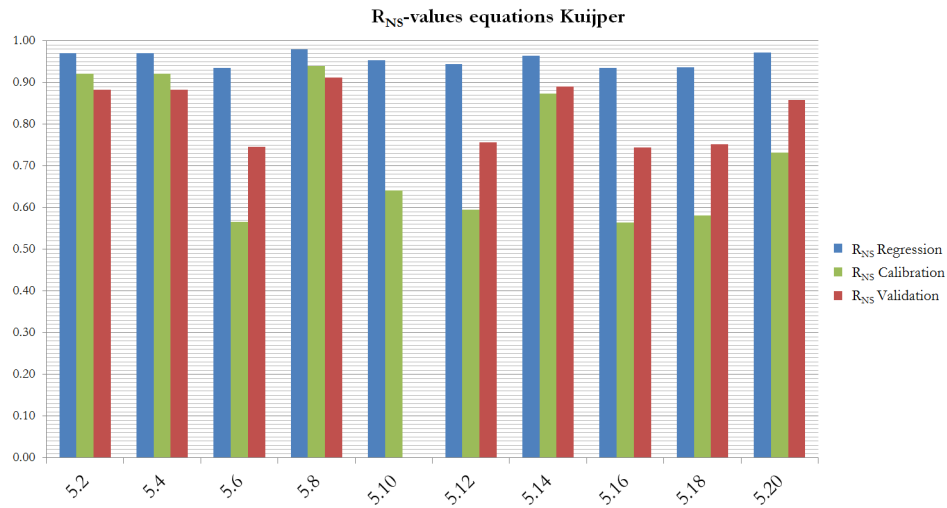


Figure 5.3: Nash-Sutcliff efficiency for the equations derived for Kuijper's model

sion without log-format should be carried out. The most interesting is to see if the derived equations still give reasonable outcomes during the validation in a dimensional form. Equations 5.5, 5.7 and 5.13 are in that respect not very promising equations. Equations 5.1, 5.3 and 5.19 are more interesting, but these are again just linear combinations of the Richardson number. For Kuijper's model Equations 5.2, 5.4, 5.8, 5.14 and 5.20 gave good results, but also these equations are mainly just simple linear combinations. Only Equations 5.8 and 5.20 are of a different structure. From these equations, Equation 5.20 is very similar to the format used in the old predictive equation and the forgoing chapters.

It is very well possible that the predictive equation is a simple linear relationship, but it must be said that deviations from the lower values are a bit neglected by the choice of the objective function. For the lower values of the dispersion coefficient these linear approximations show relatively large deviations. A different objective function should be used when focusing more on these low values. The Nash-Sutcliff log efficiency was used in some more runs of the algorithm:

$$R_{NS_{log}} = 1 - \frac{\Sigma (\log y_{modeled} - \log y_{observed})^2}{\Sigma (\log y_{observed} - \log \bar{y}_{observed})^2}$$

The results of these runs can be found in Table 5.3.

Table 5.3: Results genetic programming with a Nash-Sutcliff log efficiency

Savenije	Kuijper
$\frac{D_0}{v_0 E_0} = -0.99 + 1.19^{N_r} \quad (5.21)$	$\frac{D_0}{v_0 E_0} = \ln \left(\ln \left(\ln \left(15.4 + \frac{N_r}{1.88} \right) \right) \right) \quad (5.22)$
$\frac{D_0}{v_0 E_0} = \left(\frac{N_r}{\frac{C^2}{g}} \right)^{\ln(1.503)} \quad (5.23)$	$\frac{D_0}{v_0 E_0} = 0.92 + 25.9 N_r + \frac{H}{E} \quad (5.24)$
$\frac{D_0}{v_0 E_0} = \frac{-1631.9 \frac{H}{E} \frac{h}{a} + N_r + 0.05}{e^{\frac{H}{E} N_r} - 10748} \quad (5.25)$	$\frac{D_0}{v_0 E_0} = \frac{\left(\frac{\lambda}{a} - 567.2 - N_r^{\frac{h}{a}} \right) (0.03 + N_r)}{\frac{h}{a} + \frac{H}{E} - 3251.3} \quad (5.26)$
$\frac{D_0}{v_0 E_0} = \frac{N_r}{353.46}^{\ln(1.49)} \quad (5.27)$	$\frac{D_0}{v_0 E_0} = 0.092 N_r^{0.45} \quad (5.28)$
$\frac{D_0}{v_0 E_0} = e^{\frac{\frac{h}{a} - 2.11}{0.46 + N_r}} \quad (5.29)$	$\frac{D_0}{v_0 E_0} = \frac{e^{N_r}}{610049.8 \frac{h}{a}} \quad (5.30)$

$\frac{D_0}{v_0 E_0} = 0.009 + 0.18 N_r \quad (5.31)$	$\frac{D_0}{v_0 E_0} = 0.006 + 0.175 N_r \quad (5.32)$
---	--

Equation 5.23 gave the highest values of the objective function during the validation, the structure of this equation is the actually the same when a linear regression with log-format should be carried out. The same holds for Equation 5.27. This confirms that the structure that is obtained after the linear regressions with log-format is probably the right one. For Kuijper's model Equation 5.28 also confirms that the assumed equation structure is correct. The other presented equations here are actually again just linear combinations. Equation 5.26 gave also good results, but is very complicated. To conclude, the assumed structure of the predictive equation in Chapters 3 and 4 is probably correct as it is found back when the right objective function is applied. The remark must however be made that the weakness of this conclusion lies in the fact that the choice of the objective function stays a subjective one.

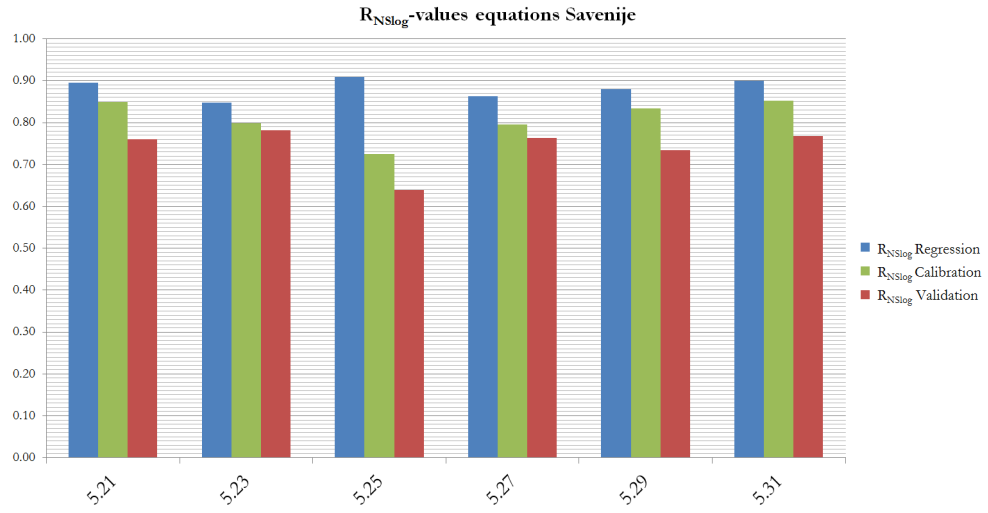


Figure 5.4: Nash-Sutcliff log efficiency for the equations derived for Savenije's model

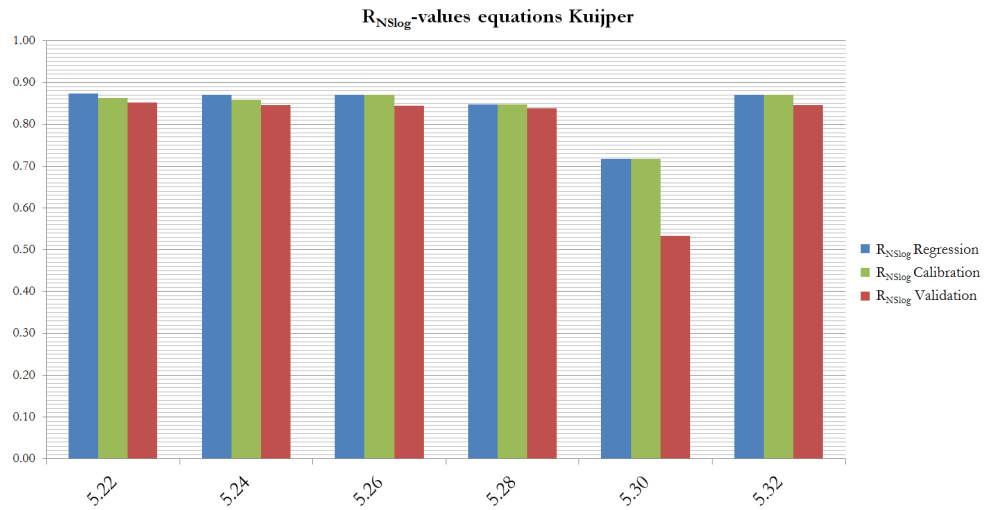


Figure 5.5: Nash-Sutcliff efficiency for the equations derived for Kuijper's model

6 Simplifying the equations

The total dataset was used to simplify and optimize the equations, this would lead to bigger differences between the performances of the equations. This could help in identifying the best equation structure. The power exponents of the equations were rounded to one decimal and the remaining constant was optimized. This was done by a simple Monte Carlo script.

One single objective function would probably not be sufficient to determine the best equation. A reliable predictive equation should be able to predict high and low values of the dispersion coefficient and therefore the Nash-Sutcliff (R_{NS}) and the log Nash-Sutcliff ($R_{NS_{log}}$) as defined in Chapter 5 are both used to judge the performance of the equations. As the applicability of the equation is the most important, the comparison was mainly based on the dimensional equation format:

$$D_x = E_x v_x f(x_1 \dots x_n)$$

6.1 Optimizing for D_0

The performances after optimizing the equations can be seen in Figure 6.1 and 6.2. The differences between the performances became bigger. After optimization Equation 3.3 showed the best results, it is however more interesting to look at the applicability of the equation and the dimensional form. Equation 3.15 shows for high and low values a good performance in dimensional form. Equation 3.29 is the same as Equation 3.15, but now with the same exponent for all the terms. This equation also gives reasonable results. Next to these, also Equations 3.19 and 3.21 still perform well when applied.

The same holds for these equations for Kuijper's case, so Equations 3.16, 3.20 and 3.22. Equation 3.30 gives also good results, this is again the same Equation as 3.16, but with one exponent for all the individual terms. These two equations also look the most consistent for high values and low values. Equation 3.8 looks also promising now.

Equations 3.19 and 3.21 (or for Kuijper's case Equations 3.20 and 3.21) stay among the best equations. In Chapter 3 and Table 3.1 it can already be seen that $\frac{H}{E}$ and $\frac{\lambda}{a}$ have a strong mutual correlation. It is also possible to derive this analytically.

The geometry-tide relation is defined by Savenije (2005) as follows:

$$\frac{H}{E} = \frac{h}{b} \frac{1 - \delta b}{\cos(\epsilon)}$$

With the second term more or less equal to one, and $a \approx b$:

$$\frac{H}{E} \approx \frac{h}{a}$$

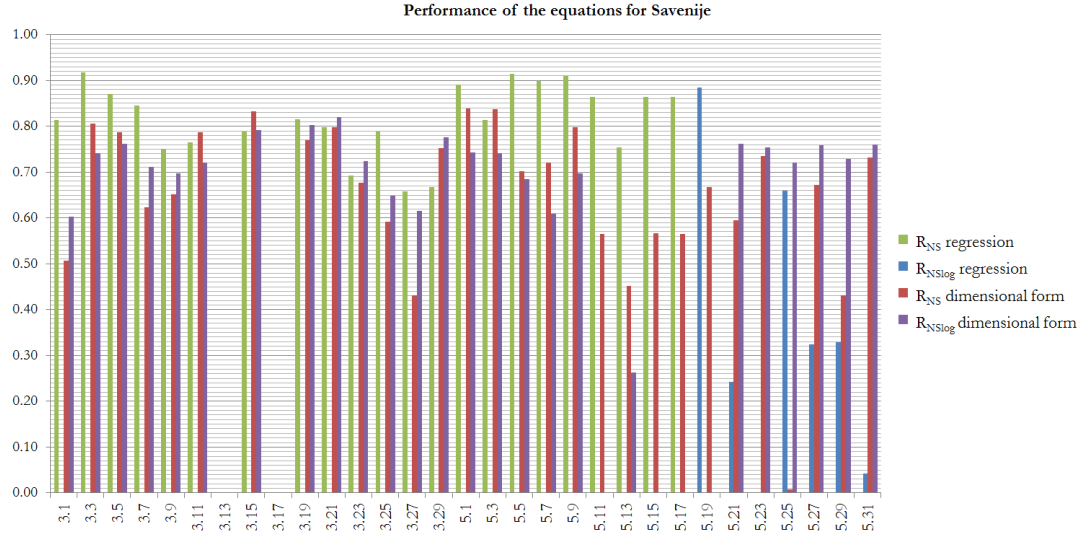


Figure 6.1: Performance for the optimized equations for Savenije's model

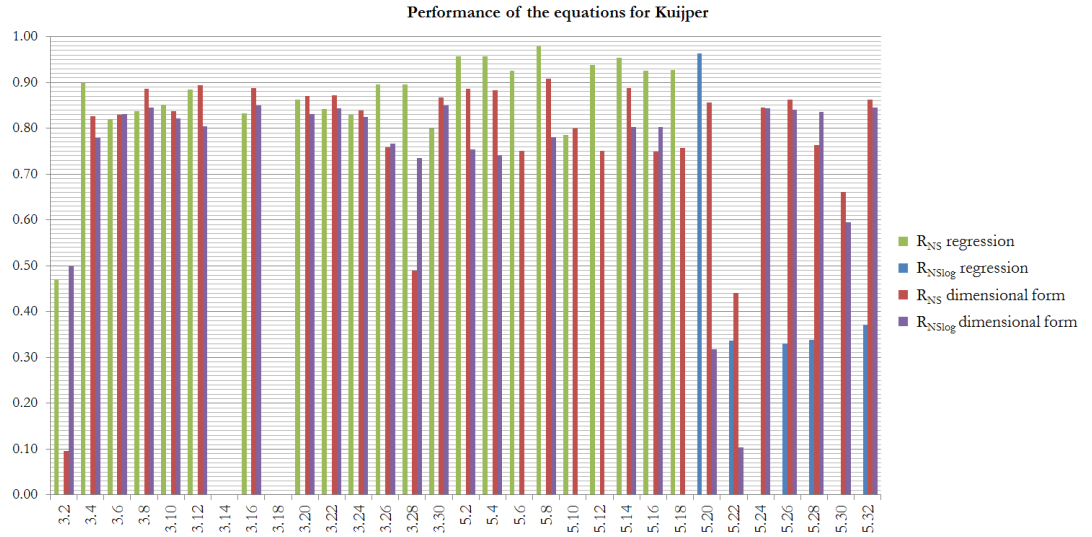


Figure 6.2: Performance for the optimized equations for Kuijper's model

Now we look at the other ratio:

$$\frac{\lambda}{a} = \frac{c_0 T}{a} = \frac{\sqrt{gh} T}{a} = \text{constant} * \frac{\sqrt{h}}{a}$$

From this we can actually see that the question is whether the square root of the depth, or just the depth should be included in the equation. From this reasoning it is also not surprising that also Equation 3.3 has a rather good performance. As Equations 3.15 and 3.16 both perform better for the optimized equations, it is more likely that the format of Equations 3.15 and 3.16 and so 3.29 and 3.30 is the correct one.

6.2 Optimizing for D_1

The same analysis was carried out for the derived equations at the inflection point. The results after optimizing these equations can be seen from Figure 6.3 and 6.4. The same patterns come up from these figures. Also here the equations with friction and horizontal to tidal range ration show good results for high values and low values. It is very interesting to note that now Equation 4.29 and 4.30 show better results than than Equation 4.15 and 4.16. It is very attractive to use the simple format of Equation 4.29 and 4.30.

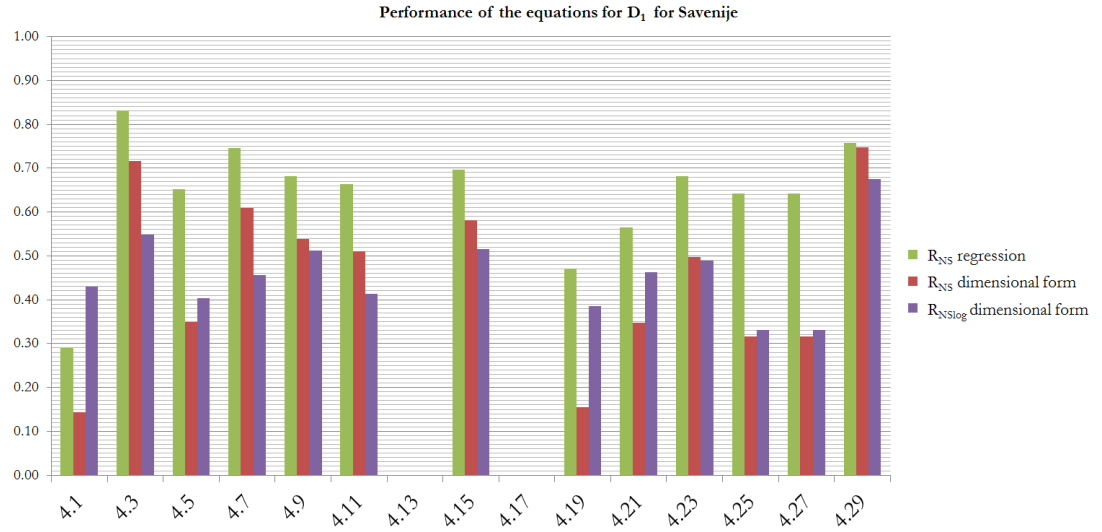


Figure 6.3: Performance of the optimized equations for D_1 for Savenije's model

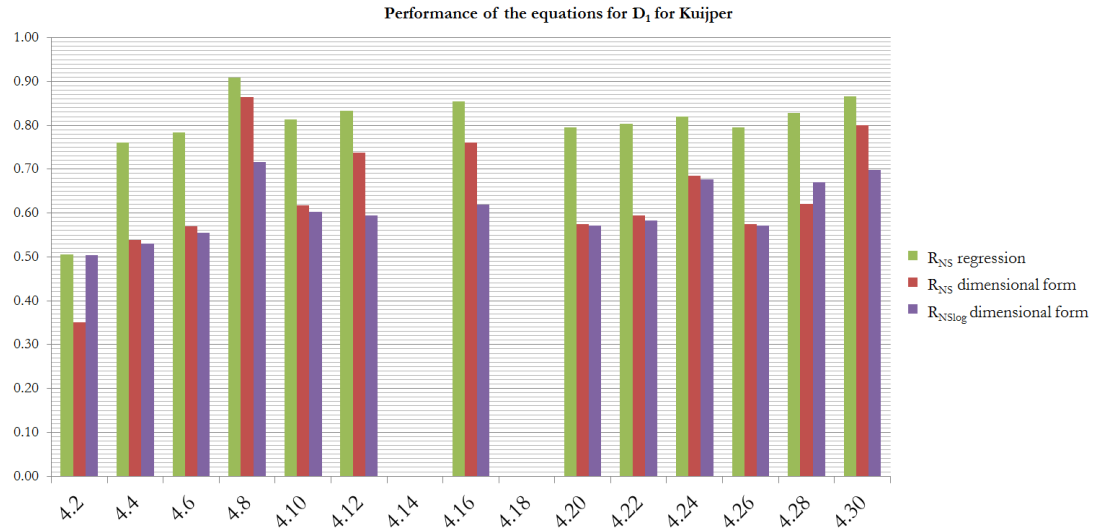


Figure 6.4: Performance of the optimized equations for D_1 for Kuijper's model

7 Local Validity

An important assumption made by Kuijper & Van Rijn (2011) is that the predictive equation is locally valid, and it should not matter at what place it is used in the estuary. To see if this statement is correct a selection of the best new predictive equations was applied locally.

7.1 Selected equations

The selection of equations that was used was based on performance and on the significance of the individual terms. For Savenije's case only 3.3, 3.11, 3.19 and 3.29 were fully accepted and gave good results. Thus, these equations were tested, just as the equivalent equations for Kuijper's case. For Kuijper's case furthermore Equations 3.10 and 3.30 were completely accepted. Also 3.9 was therefore added to the selection of Savenije's equations. Applying the same reasoning for the equations derived for x_1 led to a selection of Equations 4.3, 4.19 and 4.29 for Savenije's case and 4.4, 4.20 and 4.30 for Kuijper's case. The selection looked therefore as follows:

Savenije: Equations 3.3, 3.9, 3.11, 3.19, 3.29, 4.3, 4.19, 4.29

Kuijper: Equations 3.4, 3.10, 3.12, 3.20, 3.30, 4.4, 4.20, 4.30

To be able to apply these equations locally, some parameters had to be made x-dependent. Next to that, the Richardson number uses implicitly the salinity (in the density term). The tidally averaged salt curves were therefore used to make an x-dependent Richardson number, these salinity curves were assumed to be a good approximation of the true tidally averaged salinity. This approximation of the tidally averaged salinity was thus input for the Richardson number. The following equations apply:

$$N_{r,x} = \frac{\Delta \rho g h_x Q_f T}{\rho A_x E_x v_x^2}$$

With:

$$E_x = E(x) = E_0 e^{\delta x}$$

$$v_x = v(x) = v_0 e^{\delta x}$$

7.2 Performance of the equations

Some results can be found in Figure 7.1, but more graphs can be found in Appendices C and D. In cases with two convergence lengths it can be seen that in the first part the predictive dispersion equations almost always increased strongly. This is caused by the strong convergence in the first part of the estuary. When the cross-section decreases strongly, the Richardson number, and therefore the predicted dispersion, increases strongly. In the second part of the estuaries the predictive equations approach the dispersion generally better.

The predictive equations were generally not in agreement with the dispersion derived from the salinity curve. Because of this, it is quite reasonable to

state that there is still something missing in the predictive equations, when applied locally. The missing term should be able to compensate for the strong convergence in the first part of the estuary, and it can be hypothesized that here different depth or friction values may be applied. The first part in the estuary is mainly a wider part where sand will be deposited because of a reduction in velocity. A decrease in depth and friction (the sediment will be less coarse) could maybe compensate for the strong convergence in this first part of the estuary.

It is however important that in Savenije's case the predictive equation should provide a boundary condition for the model. The used predictive equations can still be correct, but only as the boundary condition for the dispersion at a seaward point (estuary mouth or inflection point).

The dispersion pattern for Kuijper's model was already much better. The dispersion derived from Kuijper's model mainly followed the same pattern as the locally applied predictive equations. In one way this is logical as this model uses the start assumption that the dispersion is indeed locally valid and has a form similar to the predictive equations. The salt model of Kuijper was derived with this predictive equation as a start. Kuijper's model is in that sense mathematically smart and becomes more predictive by this start assumption.

The best local approaches were in general Equations 3.3, 4.29 and 4.3 , however there is still room for improvement. Equation 4.3 uses the two different convergence lengths, which led to a jump in the dispersion curve. Equation 4.29 generated a continuous curve and showed also rather good results. In Kuijper's case Equations 3.10, 3.30 performed quite reasonable, but also Equation 4.30 was not bad. It is rather interesting that again for Kuijper's case the equation with a friction term (3.10) showed better results than for Savenije's case, Equation 3.9 is not very bad locally, but is also not among the best.

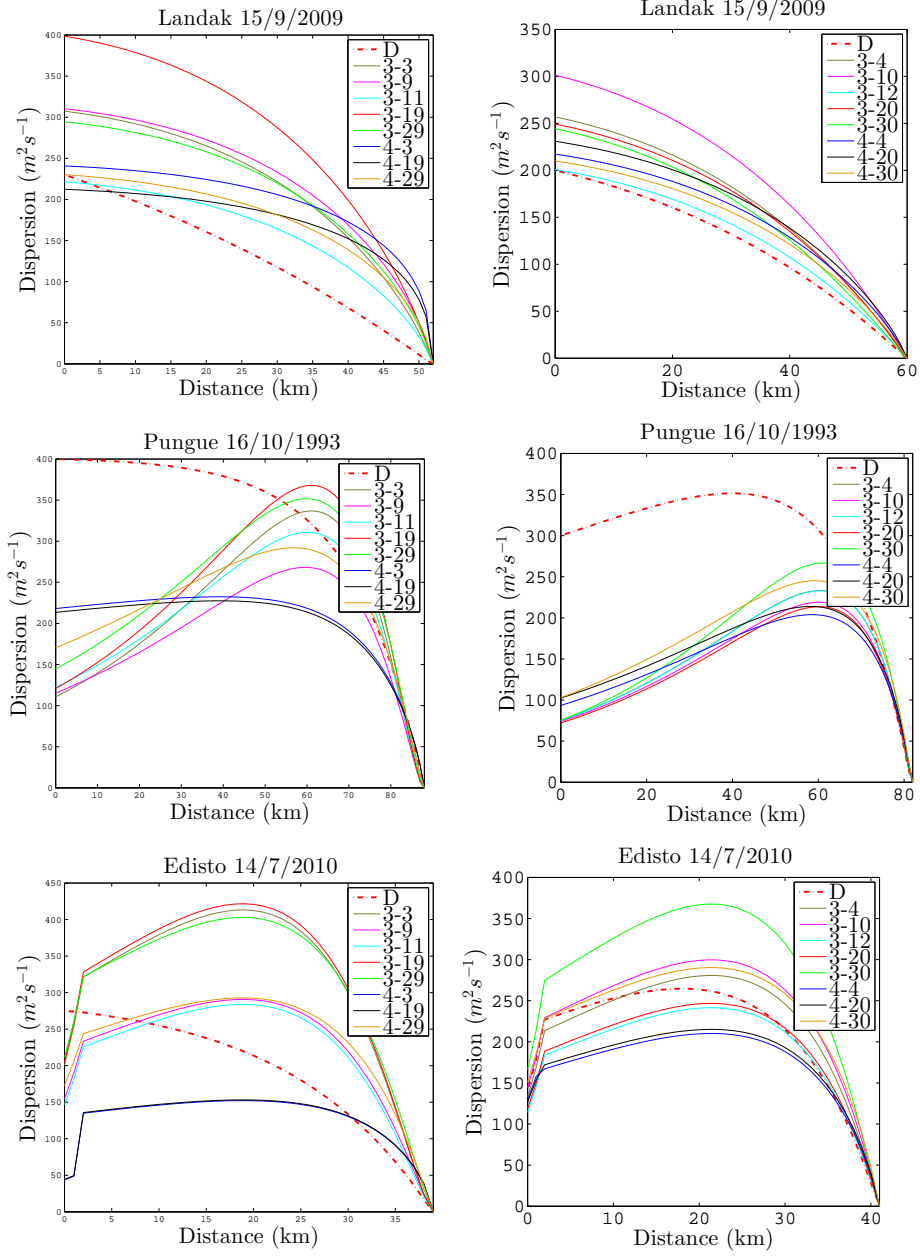


Figure 7.1: The predictive equations applied locally for Landak, Pungue and Edisto with left the model of Savenije and right the model of Kuijper

8 Predictive Mode

The same selection of equations as in Chapter 7 was applied to see if these equations were actually capable of producing well fitted salinity curves. All the curves can be found in Appendices E and F.

8.1 Performance of the equations

Looking qualitatively at the graphs shows that all the predictive equations were able to produce reasonable results. Some of the graphs can be found in Figure 8.3. The different graphs did not differ much and it looks hard to draw conclusions from it. A closer visual inspection however shows that especially Equations 3.29 and 4.29 approach the calibrated lines in most cases very well. These equations are also quite similar, they only differ in the constant and the location where the calculation starts (estuary mouth versus the inflection point). Exactly the same can be observed from the graphs for Kuijper's case. Also here the same equations for Kuijper's case performed the best. Equations 3.30 and 4.30 show the best approach of the calibrated curves.

The performances of the different equations were quantified with the log Nash-Sutcliffe efficiency as defined in Chapter 5. This was done for each measurement individually, but also for all the measurements together. The total performance over all the measurements for HWS and LWS can be found in Figures 8.1 and 8.2. It can be noted that for Savenije's case Equation 3.29 shows the best results and for Kuijper's case Equation 4.30. The second best equation for Kuijper's case is Equation 3.30, which is actually the same equation as 4.30 only now derived for the estuary mouth. For Savenije's case Equations 3.3 and 3.19 show actually better results than 4.29, the differences are however small.

Looking at the individual performances gives a more or less similar impression. For Savenije's case and Kuijper's case Equations 3.29, 3.30, 4.29 and 4.30 give in general the highest objective function values. In 30 measurements out of 72 measurements Equations 3.29 or 4.29 showed the highest log Nash-Sutcliffe values for Savenije's model. In Kuijper's case this was similar, here 38 cases had 3.30 or 4.30 as the best.

It can be questioned of course if the deviations from the calibrated curves are big enough to draw conclusions from. The error introduced by the measurements and the calibration might even be bigger than the deviation from the furthest calculated line. The conclusions must therefore be drawn with caution. However, when for Kuijper's case and Savenije's case the same type of equations perform well in by far the most cases, something is probably done right by applying these equations.

8.2 Starting from the inflection point

From applying the predictive equations it can also be concluded that it is very well possible to start the calculations from the inflection point. The results for starting at the inflection do not differ much from the results for the calculations that started at the estuary mouth. One requirement that should be met to start the calculations from the inflection point is however that the salt intrusion should of course come to this point. Otherwise the dispersion becomes 0 and

the calculation doesnot make sense. The inflection point should therefore not be too far away from the estuary mouth.

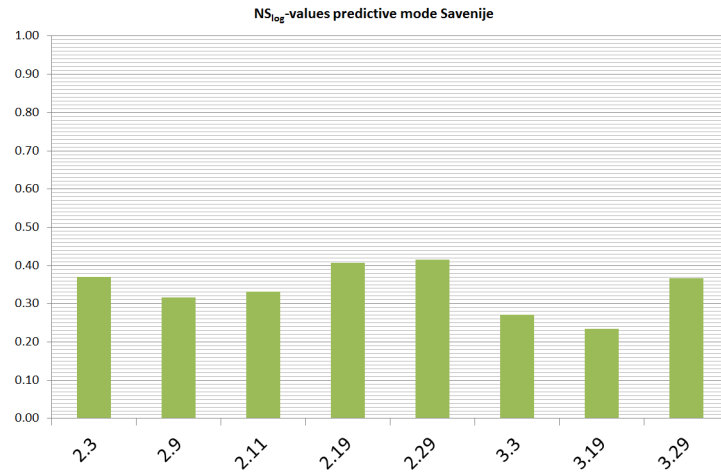


Figure 8.1: Log Nash-Sutcliffe efficiency for the equations derived for Savenije's model used in a predictive manner

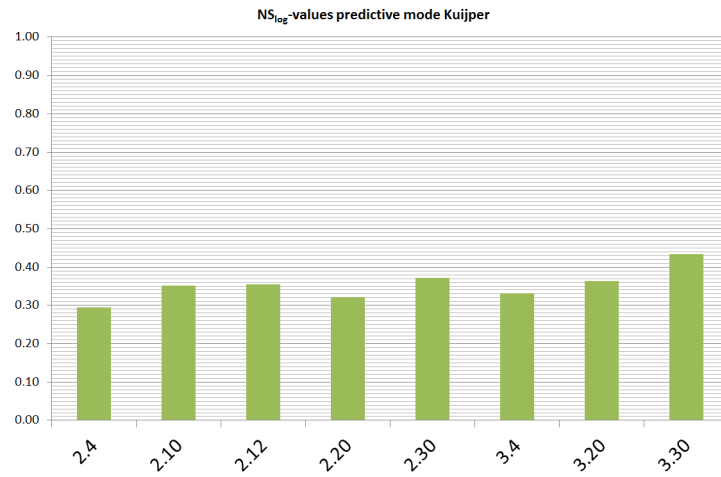


Figure 8.2: Log Nash-Sutcliffe efficiency for the equations derived for Kuijper's model used in a predictive manner

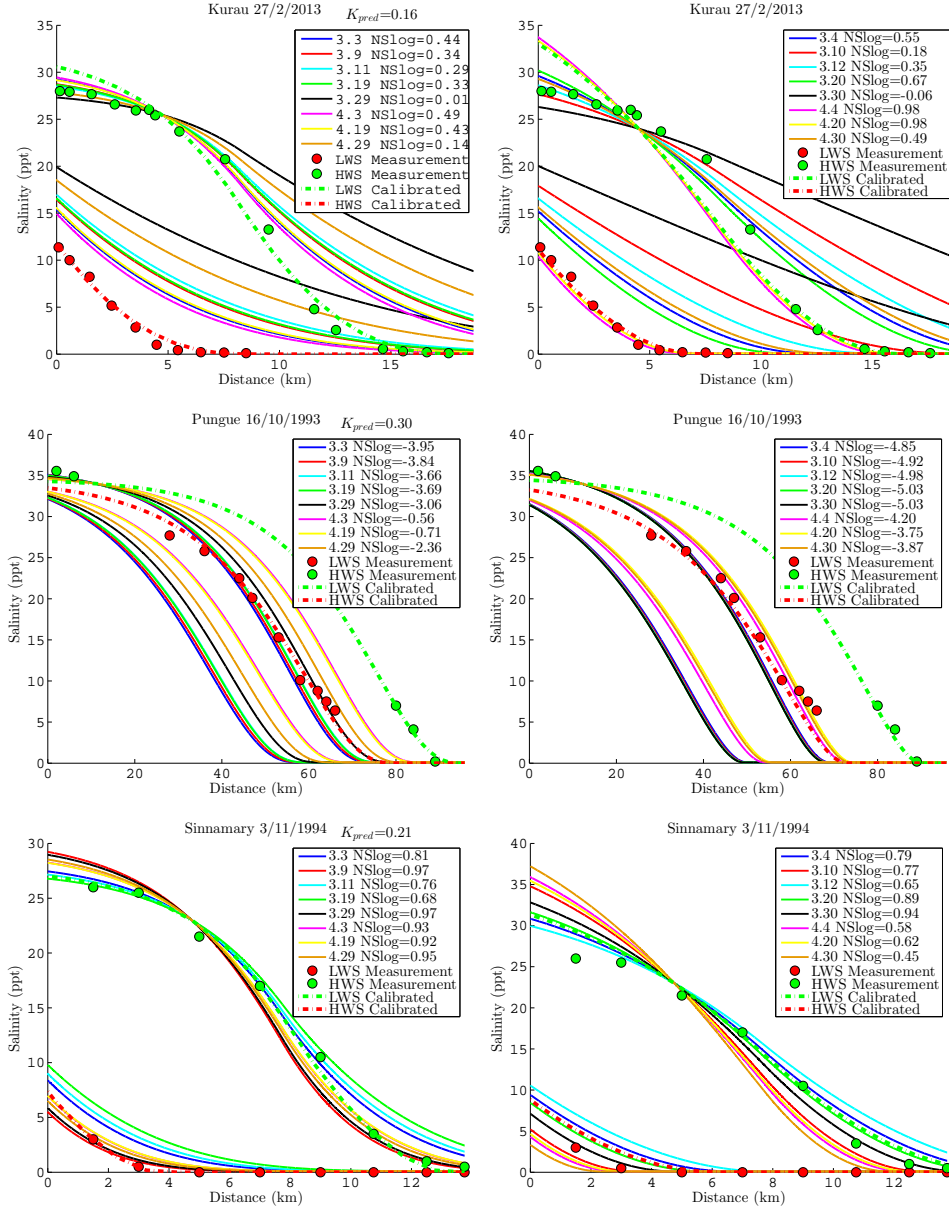


Figure 8.3: The salt models applied for Kurau, Pungue and Sinnamary for left Savenije and right Kuijper and Van Rijn

9 Hydraulic calculations

In order to make the analytical salt model completely predictive, estimations should be made of depth and friction. It was therefore tested if the equations used by Cai et al. (2012) could be used with observations of water levels and a known convergence length to predict these parameters .

9.1 Estimating damping and wave celerity

The first step consists out of estimating the damping and wave celerity for an estuary with available water level recordings. This was carried out for the estuaries Kurau, Perak, Endau, Bernam, Muar and Selangor.

The damping is defined as:

$$\delta_H = \frac{1}{\eta} \frac{d\eta}{dx} \quad (9.1)$$

With η the tidal amplitude, which is the half of the tidal range H . The damping can be estimated for a reach with divers at location 1 and 2:

$$\delta_H = \frac{1}{(\eta_1 + \eta_2)/2} \frac{\Delta\eta_{1-2}}{\Delta x} \quad (9.2)$$

This leads to estimations of the damping for a certain reach in the estuary for a certain time during the tidal cycle.

The wave celerity can be estimated by plotting the high water and low water times against the distance, as can be seen in Figure 9.1 and 9.2. This can be done for each tidal cycle, leading to an estimate of the wave celerities c_{HW} and c_{LW} for each high water and each low water. The estuaries under consideration have an inflection point in the first part of the estuary, what led to different wave celerities in the two parts of the estuary.

The average wave celerities over time were obtained by plotting the travel time of the wave against traveled distance of the wave, see Figure 9.3. Now a linear approximation was made by drawing a line through the high water and low water points. This was done for all high and low water points, but it was also done separately for the points before and after the inflection point. In the last way two time-averaged celerities were obtained.

9.2 Equations for friction and depth

The average depth of the estuary can be estimated with the wave celerity (c), damping (δ_h), angular velocity (ω), convergence length (a) and the storage width to stream width ratio (r_s) with the following equation based on Cai et al. (2012):

$$\bar{h} = \frac{r_s}{g \left[\frac{1}{c^2} + \frac{\delta_H}{\omega^2} \left(\frac{1}{a} - \delta_H \right) \right]} \quad (9.3)$$

The phase lag ϵ and velocity v can be calculated with the phase lag equation and scaling equation of Savenije (2005), which are also used by Cai et al. (2012):

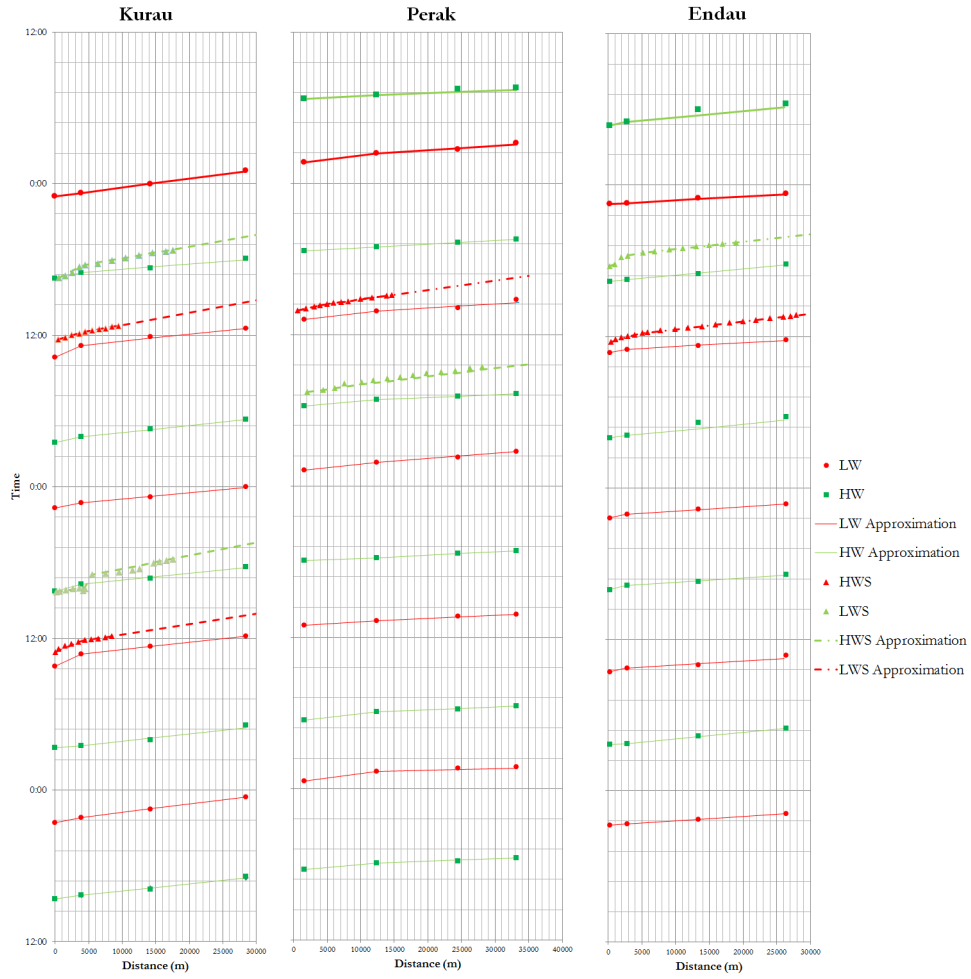


Figure 9.1: Place-time diagram per high water, low water and (if observed), slack

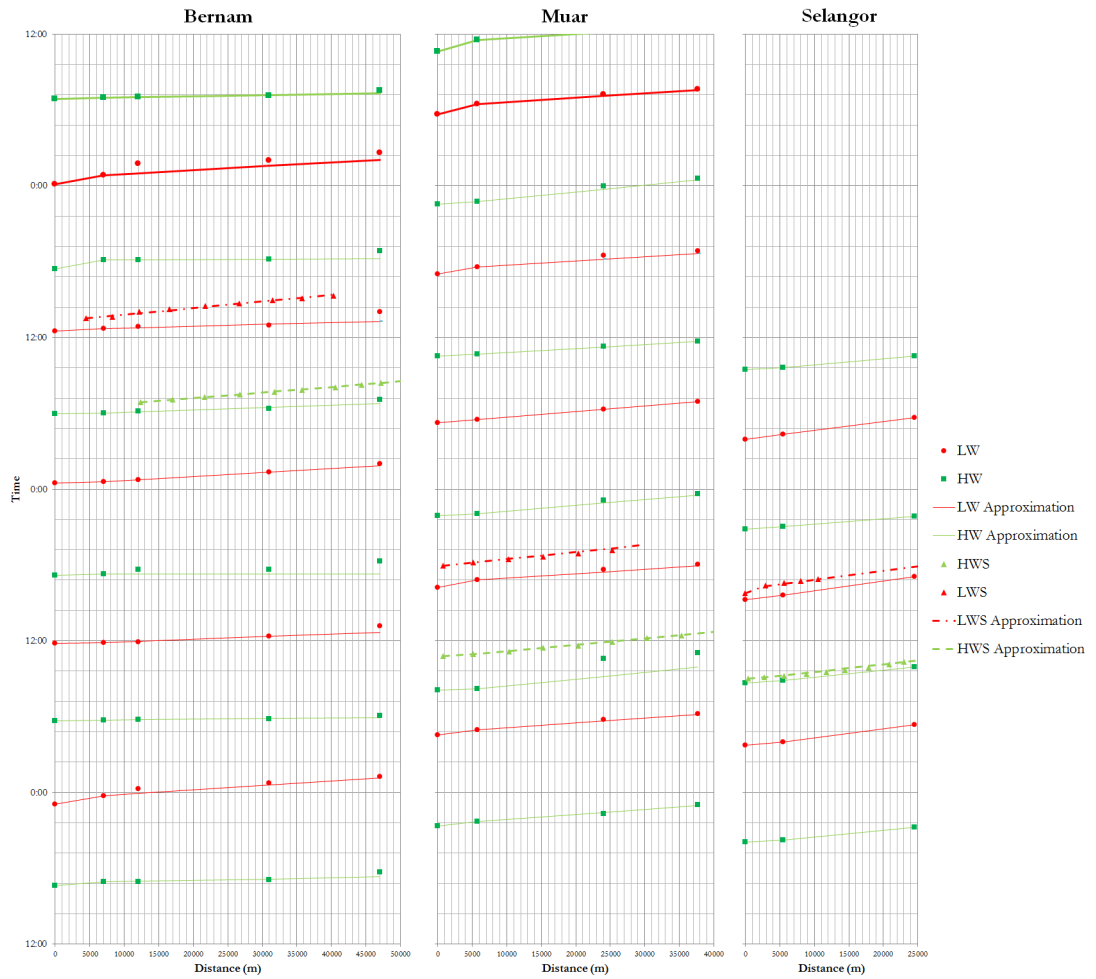


Figure 9.2: Place-time diagram per high water, low water and (if observed), slack

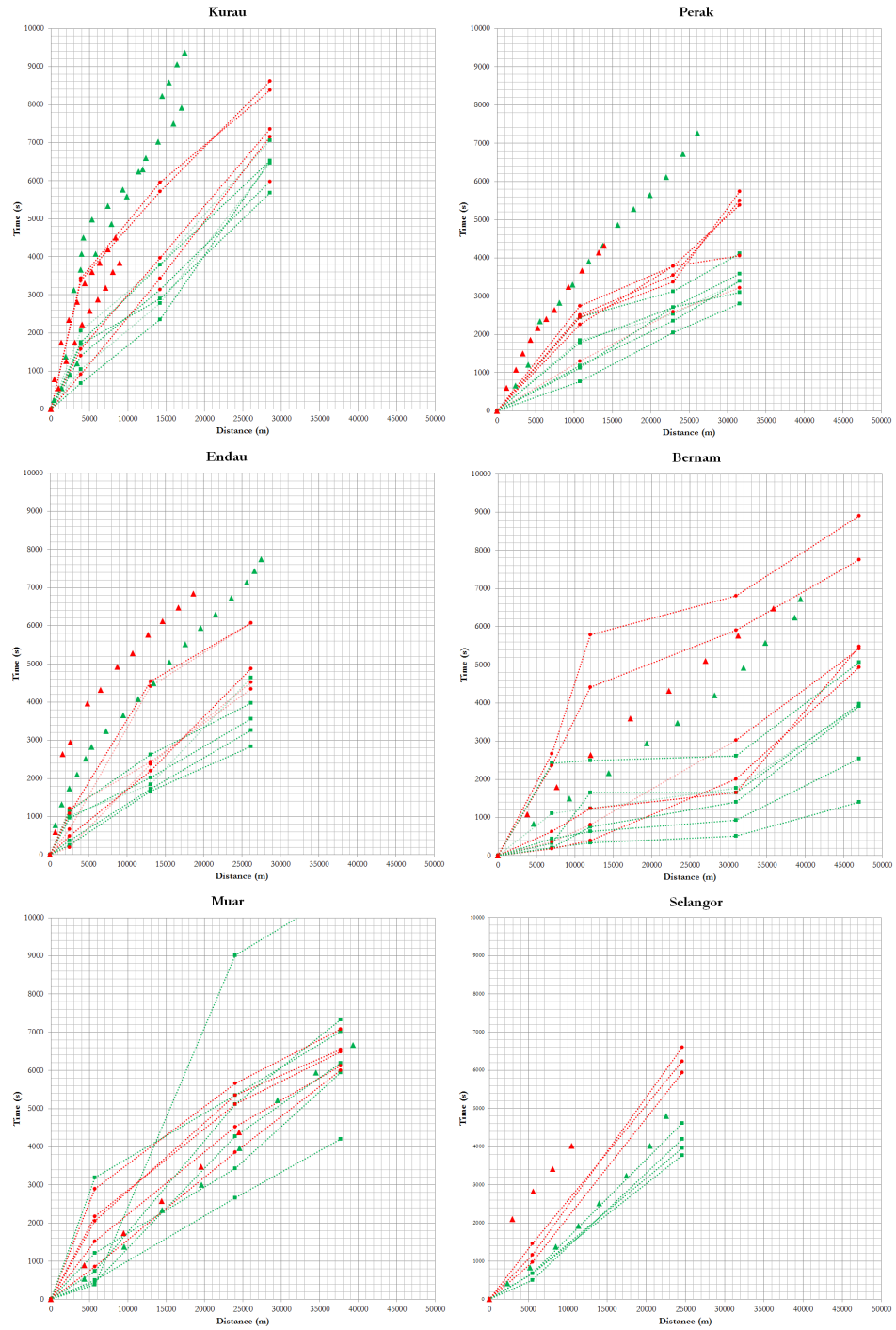


Figure 9.3: Place-time diagram, with different celerities for the two parts in the estuaries

$$\tan \epsilon = \frac{\omega a / c}{1 - a \delta_H} \quad (9.4)$$

$$v = r_s \frac{\eta}{\bar{h}} c \sin \epsilon \quad (9.5)$$

The damping equations according to the hybrid model of Cai et al. (2012) look as follows:

$$\delta_H \left(1 + \frac{g\eta}{cv \sin(\epsilon)} \right) = \frac{1}{a} - f \frac{v}{\bar{h}c} \left(\frac{2}{3} \sin(\epsilon) + \frac{8\pi}{9} \right) \quad (9.6)$$

$$f = \frac{g}{K_{Manning}^2 \bar{h}^{1/3}} \left[1 - \left(\frac{4}{3} \frac{\eta}{\bar{h}} \right)^2 \right]^{-1} \quad (9.7)$$

This can be rewritten to find a formula for the Manning-Strickler friction factor:

$$K_{Manning} = \sqrt{\frac{gv (6 \sin(\epsilon) + 8\pi)}{\bar{h}^{4/3} c \left[9 + 16 \left(\eta / \bar{h} \right)^2 \right] \left[\frac{1}{a} - \delta_H \left(1 + \frac{g\eta}{cv \sin \epsilon} \right) \right]}} \quad (9.8)$$

9.3 Estimation of the estuary parameters

For c , v and δ_H estimates for a certain time in the tidal cycle and a certain reach of the estuary can be used, leading to local estimates for a certain time and place of ϵ , v and $K_{Manning}$, but it is also possible to use the averaged values for c , η and δ_H over time, and, with respect to space, over the whole estuary or over the two parts with different convergence lengths. It was tested which of the approaches gave the best estimates. The analytical model of Cai et al. (2012) was also used to compare the results. The main difference between these approaches is that the analytical model uses the depth as input. If observations were available these were also used to compare the results of estimates and analytical model.

The results for the depth estimations can be found in Figure 9.4. The local estimations showed some deviations from observed depth, sometimes even negative values occurred. This could be caused by noise in the water level recordings, what makes it more difficult to distinguish the high water and low water times and therefore celerities.

Errors are also introduced by the location of the water level recordings. A wrong estimate of celerity and maybe even damping for the first part of the estuary is obtained when the inflection point is for example at $x=4900$ and the water level recording is located at $x=10000$.

Averaging over more tidal cycles, using therefore Figure 9.3, gave better estimations of the depth in the estuary. Using the average values over the first and second part of the estuary also gave reasonable results. It should however still be noted that this will be harder to do when the water level recording is rather far from the inflection point, as is the case for Perak estuary.

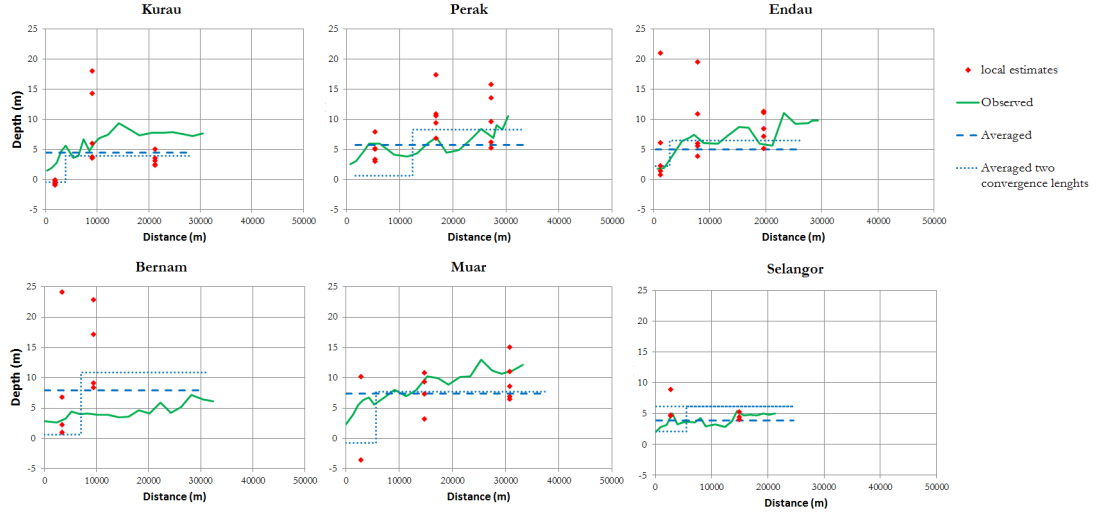


Figure 9.4: Estimated, modeled and observed depth

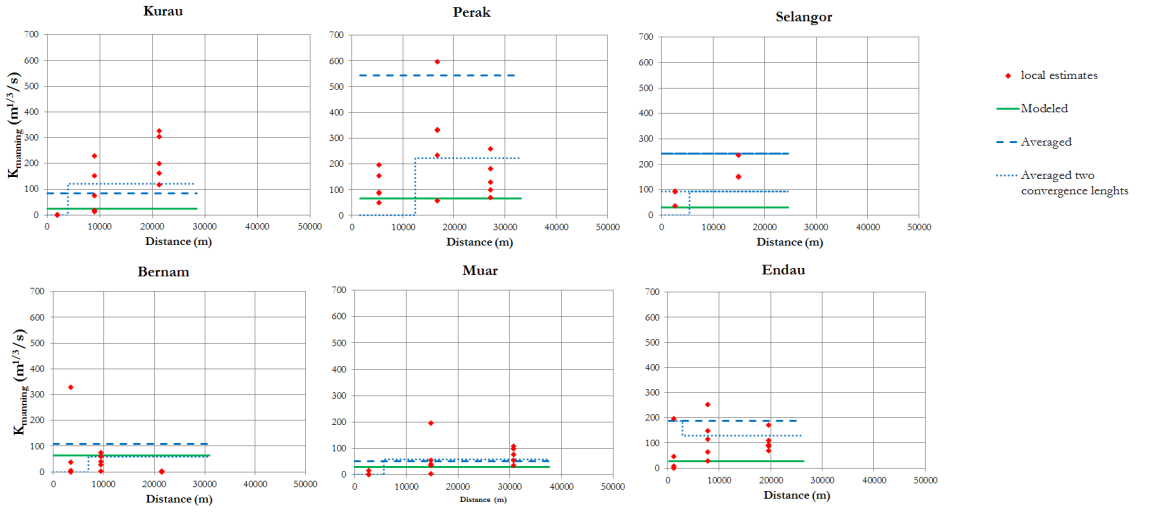


Figure 9.5: Estimated, modeled and observed friction

The slack times can be determined with the estimation of the phase lag. The slack times obtained by averaging over more tidal cycles were compared with observations and the analytical model in Figure 9.6. The slack times were determined by adding the phase lag to the observed high water times. The model only used the first location for an initial high or low water time, with the modeled celerity and phase lag the slack times were determined. The model showed discontinuities at the inflection point. After the jump the estimated phase lag and modeled results show reasonable agreement, also with the observed values.

The local estimates for the friction did not give satisfactory results for the first part of the estuary, especially in the cases where the depth was estimated to be negative. This can also be seen in Figure 9.5. For Kurau and Endau the estimates with the averaged approach gave reasonable results compared to the analytical model. Perak however still showed a rather large deviation.

The estimated tidal excursion in Figure 9.7 also looks rather reasonable compared with modeled and observed results. Again the averaged values look better than the local estimates.

To conclude, the suggested approach can be useful in a first estimation of depth and friction, it should however be kept in mind that the uncertainties are large. The best way to use the approach is by using time and place averaged values for celerity and damping.

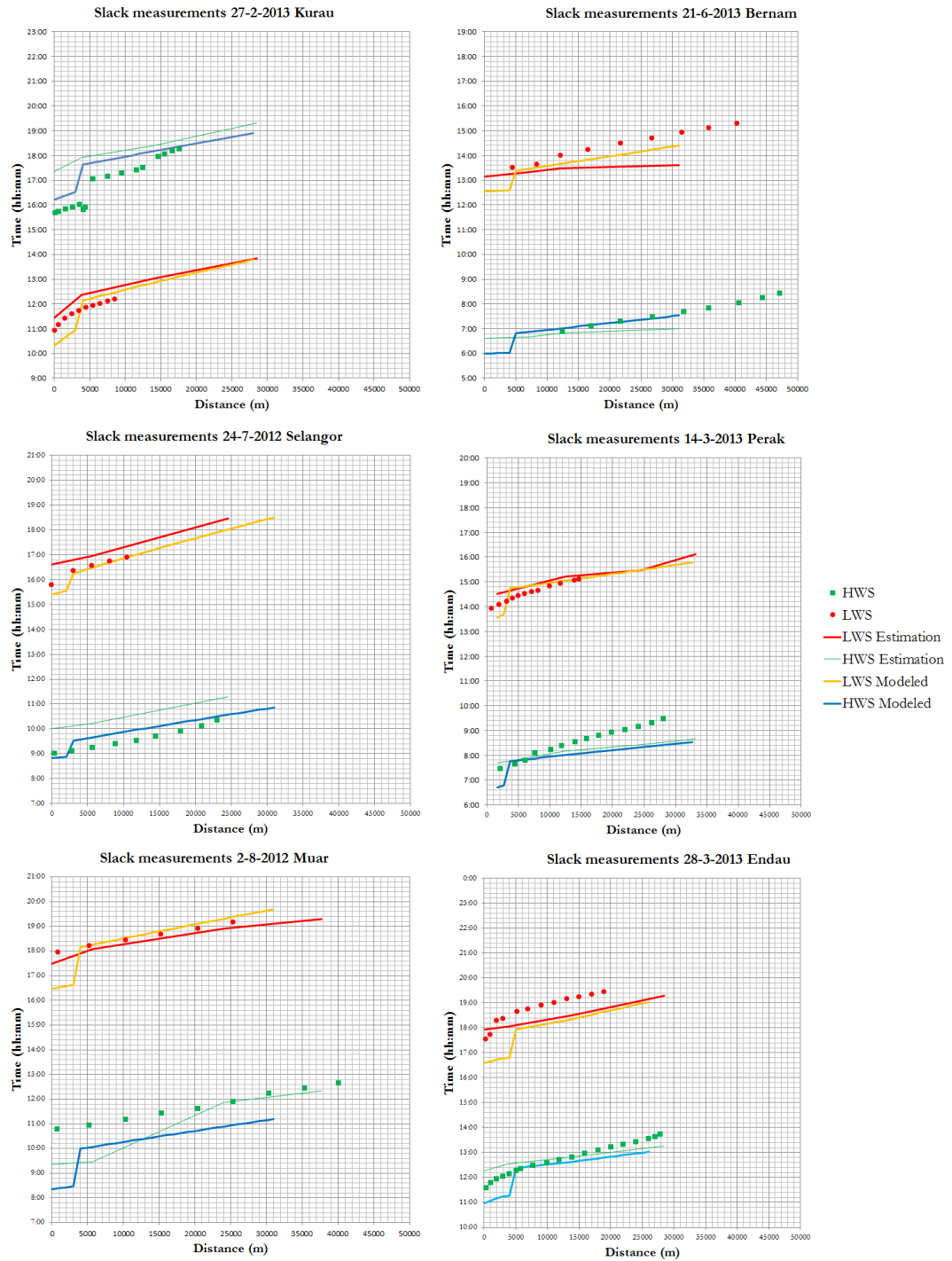


Figure 9.6: Estimated, modeled and observed slack times

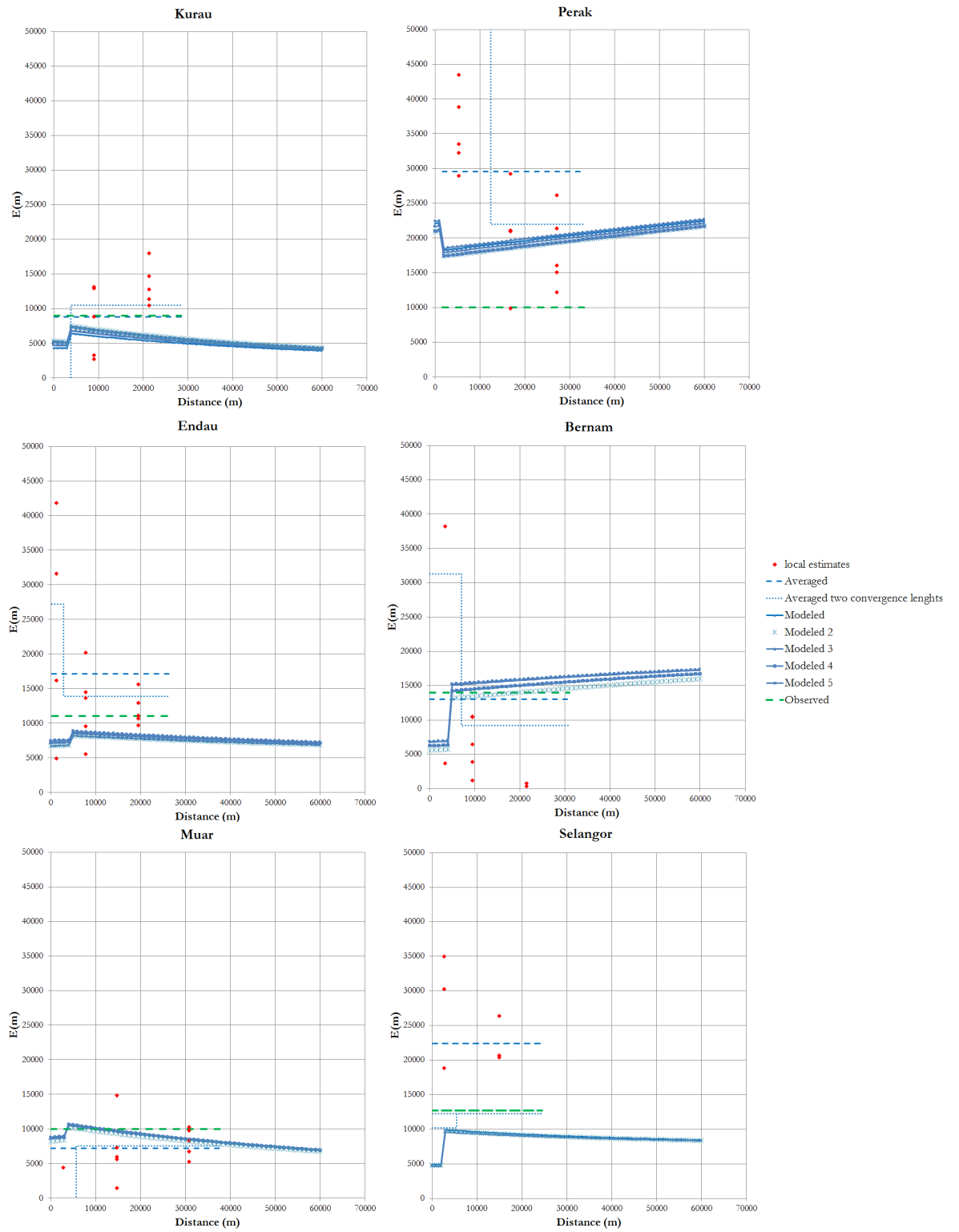


Figure 9.7: Estimated, modeled and observed tidal excursion

10 Conclusions and Recommendations

10.1 Savenije's model

The equations derived by the linear regressions gave more or less equally good results. After looking at the significance of the different terms the number of possible equations could already be brought back. For Savenije's model Equations 3.3, 3.11, 3.19 and 3.29 showed good results and all the terms of these equations were proven to be of significant importance. All the equations were simplified and the coefficients were rounded, this mainly led to more differences between the performance of the equations. It was shown in Chapter 6 that from these equations Equations 3.3, 3.19 and 3.29 also performed well when optimized.

The same analyses were carried out for the inflection point and here Equations 4.3, 4.19, 4.29 showed good results and had terms that were all significant. Equation 4.29 was the only equation that also kept performing well during the validation, the other equations all showed a strong decrease in performance.

Genetic programming was used to check if the used structure for the predictive equations was actually the correct one. The algorithm confirmed that the structure is correct, however the remark is made that this confirmation depends on the subjective choice of objective function that was used. It is also interesting to note that the same power-exponent (0.4) was found back for the Richardson number.

The total selection of possible equations consisted till now out of Equations 3.3, 3.11, 3.19, 3.29, 4.3, 4.19 and 4.29. These equations were applied locally and it was shown that the curves didnot approach the modeled dispersion very well. The best ones were however again Equations 3.3, 4.3 and 4.29. From these Equation 4.29 was the only one that did not show a jump in the curve because of the change in convergence lengths. The curves didnot correspond well with the dispersion curve from the salt model, but this doesnot mean that the predictive equations are not suitable. In Savenije's case, the predictive equation provides a boundary condition at the seaward side of the estuary, and is not meant as a local equation.

The same selection of equations was finally applied to 72 measurements in 27 estuaries. The best equations were now Equations 3.29 and 4.29. These equations were always among the best performing equations. The structure of these equations is also interestingly simple. The used terms can also be explained by reasoning. The equation accounts for gravitational mixing by the Richardson number, turbulent mixing by the friction term and tidal mixing by the vertical to horizontal tidal range ratio. The only question that remains is whether the predictive equation should be applied at the inflection point or in the estuary mouth. The overall fit for the regressions carried out for the inflection point were not as good as for the other regressions. The inflection point is however more clearly defined and it was very well possible to start the calculation from here. The recommendation is therefore to start the calculation from the inflection point.

The new proposed predictive equation is therefore:

$$\frac{D_1}{E_1 v_1} = 10.29 \left(N_r \frac{H}{E} \frac{g}{C^2} \right)^{0.35} \quad (10.1)$$

D₁ predicted vs D₁ calibrated

Y-axis: D₁ Calibrated (m²/s)

X-axis: D₁ Predicted (m²/s)

Legend: \square D₁ predicted vs D₁ calibrated

Species and D₁ values (m²/s):

- 16a, 25a: ~1500
- 23b, 25b: ~1200
- 22a, 24a: ~800
- 18a, 19a, 23b, 24b: ~600
- 21b, 28a, 24b: ~400
- 15b, 18a, 24b: ~300
- 1, 28b: ~200
- 9k, 21c, 21b, 18b, 10c: ~150
- 26a, 13d, 10b: ~100
- 12c, 9f, 9j, 9h, 27b, 3, 10b, 10d, 6: ~50
- 9e, 12b, 12a, 17d, 17c, 17b: ~30
- 20a, 20b, 20c, 10e: ~20

10.2 Kuiper's model

62

3.30. Again, all the equations were symplified and optimized, leading to more differences between the performance of the equations. The same equations as mentioned before kept on giving good results, Equation 3.30 looked also the most consistent for high flows and low flows.

The regressions for the inflection point in the estuary were also carried out for Kuijper's model. The regressions showed good results, but none of the equations were fully accepted. Only the one-term equations, like Equation 4.30, were accepted.

Symbolic regression showed also for Kuijper's case that the original equation structure is probably the correct one. The fact that also for Kuijper's model and Savenije's model this confirmation is found only strengthens the idea that the correct format is a linear relation of the log of the ratios.

Equations 3.4, 3.10, 3.12, 3.20, 3.30, 4.4, 4.20 and 4.30 were again applied locally. These dispersion curves showed to correspond more with the modeled pattern of Kuijper's model, in contrast to the situation for Savenije's model. Kuijper's model implicitly uses a dispersion relation similar to the format for the predictive equations, what is mathematically smart and makes the dispersion relations more locally valid. The best equations locally were Equations 3.10 and 3.30.

The selection of equations was applied to the measurements for Kuijper's case as well. The best equations were now Equations 3.30 and 4.30, so actually the same equations as for Savenije's model. The same argumentation can be therefore used as for Savenije's model. These equations are simple, and performed always well. They also showed to be consistent for high values and low values, and also the estuary mouth and the inflection point (especially for Kuijper's model).

The new proposed predictive equation for Kuijper's model is therefore:

$$\frac{D_1}{E_1 v_1} = 35.62 \left(N_r \frac{H}{E} \frac{g}{C^2} \right)^{0.43} \quad (10.2)$$

As the model of Kuijper uses the same power function in the model as in the predictive equation, this means that the value of 0.5 for the Van der Burgh K should also be changed to 0.4. This was not tested in the current research, but is worth testing in the future.

Also here the regression results are repeated in Figure 10.2. The same measurements show in general up as outliers, like 4, 5, 6, 10b, 10d and 10e. The same argumentation for Savenije's case holds and the outliers are probably caused by an unreliable (low) discharge and the estuary in a non-steady state.

10.3 The new approach

As the new predictive equation now uses the friction it is important to have a way to estimate friction. It seemed to be possible to do this with help of the analytical model of Cai et al. (2012). Summarizing, the approach to calculate salt intrusion is now recommended to be as follows:

- Determine the location of the inflection point based on information about the geometry
- Determine average wave celerity and average damping based on water level readings at several places along the estuary

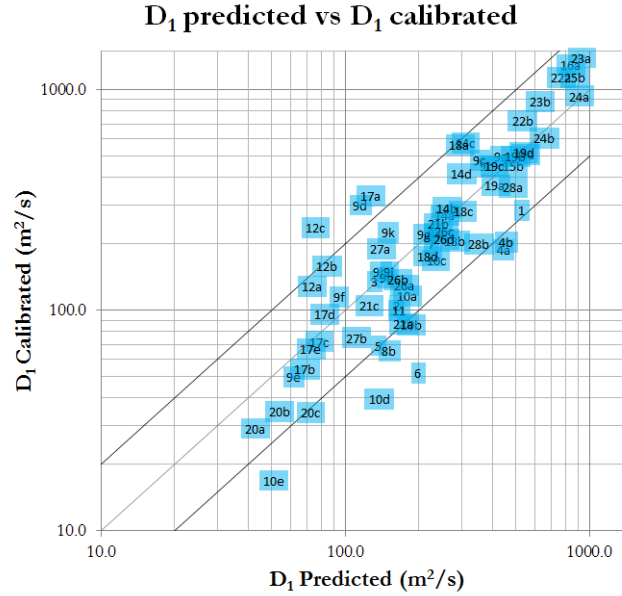


Figure 10.2: The new predictive equation versus the calibrated values, $R_{NS} = 0.82$

- Determine average depth and friction with the equations of Cai et al. (2012)
- Determine the dispersion coefficient with help of Equations 10.1 or 10.2
- Determine the salt distribution with the model of Savenije (1986) or Kuiper & Van Rijn (2011)

10.4 The models compared

Both models showed acceptable results, with mainly differences in tails of the salinity curve. This therefore indicates that a fixed value of 0.5 is not correct. It is however still convenient to use a fixed value, and, next to that, the physical meaning of K is still open for debate. As presented in the new approach in the forgoing section a preferred model is not chosen, as both models have their advantages and drawbacks. Savenije's model is easier to apply, but has more calibration parameters. Kuiper's model looks less accurate in especially the tails of the salt curves, but the connection between predictive equation and salt model is mathematically smart.

10.5 Further considerations

Every research has his own disadvantages and weaknesses. To improve this research more data should actually be obtained. It can still be questioned if the number of 72 measurements is sufficient to perform consistent regressions with. From the local application of the predictive equations it is very likely that there is still something missing in the equations. The convergence of the

cross-section is the main reason that the curves of predictions and model do not correspond, therefore an element should be included that counteracts this effect. Local values for example friction or depth may be an option, these are both assumed to be constant throughout the whole estuary in this research.

References

- Babovic, V. & Keijzer, M. (2000). Genetic programming as a model induction engine. *Journal of Hydroinformatics*, 35–60.
- Bowden, K. F. (1981). Turbulent mixing in estuaries. *Ocean Management*, 6, 117–135.
- Cai, H., Savenije, H. H. G., & Toffolon, M. (2012). A new analytical framework for assessing the effect of sea-level rise and dredging on tidal damping in estuaries. *Journal of Geophysical Research*, 117.
- Fischer, H. B. (1972). Mass transport mechanisms in partially stratified estuaries. *Journal of Fluid Mechanics*, 53(4), 671–687.
- Fischer, H. B. (1976). Mixing and dispersion in estuaries. *Ann. Rev. Fluid Mech.*, 8, 107–133.
- Koza, J. R. (1994). Genetic programming as a means for programming computers by natural selection. *Statistics and computing*, 4, 87–112.
- Kuijper, K. & Van Rijn, L. C. (2011). Analytical and numerical analysis of tides and salinities in estuaries; part ii: salinity distributions in prismatic and convergent tidal channels. *Ocean Dynamics*, 61(11), 1743–1765.
- Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1), 112–147.
- MacVean, L. J. & Stacey, M. T. (2010). Estuarine dispersion from tidal trapping: A new analytical framework. *Estuaries and Coasts*, 59, 404–406.
- McCarthy, R. K. (1993). Residual currents in tidally dominated, well-mixed estuaries. *Tellus*, 45, 325–340.
- Nguyen, A. D. (2008). *Salt Intrusion, Tides and Mixing in Multi-channel Estuaries*. PhD thesis.
- Prandle, D. (1981). Salinity intrusion in estuaries. *Journal of Physical Oceanography*, 113.
- Prandle, D. (1985). On salinity regimes and the vertical structure of residual flows in narrow tidal estuaries. *Estuarine, Coastal and Shelf Science*, 20, 615–635.
- Savenije, H. H. G. (1986). A one-dimensional model for salinity intrusion in alluvial estuaries. *Journal of Hydrology*, 85(1-2), 87–109.
- Savenije, H. H. G. (1993). Composition and driving mechanisms of longitudinal tidal average salinity dispersion in estuaries. *Journal of Hydrology*, 144, 127–141.
- Savenije, H. H. G. (2005). *Salinity and Tides in Alluvial Estuaries*. Elsevier.

- Searson, D. P., Leahy, D. E., & Willis, M. J. (2010). Gptips: an open source genetic programming toolbox for multigene symbolic regression. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1. Citeseer.
- Shaha, D. C. & Cho, Y. K. (2011). Determination of spatially varying van der burgh’s coefficient from estuarine parameter to describe salt transport in an estuary. *Hydrology and Earth System Sciences*, *15*, 1369–1377.
- Simpson, J. H., Brown, J., Matthews, J., & Allen, G. (1990). Tidal straining, density currents, and stirring in the control of estuarine stratification. *Estuaries*, *13*(2), 125–132.
- Van Rijn, L. C. (2011). Analytical and numerical analysis of tides and salinities in estuaries; part i: tidal wave propagation in convergent estuaries. *Ocean Dynamics*, *61*, 1719–1741.

A Tables

Table A.1: Used data in the analyses

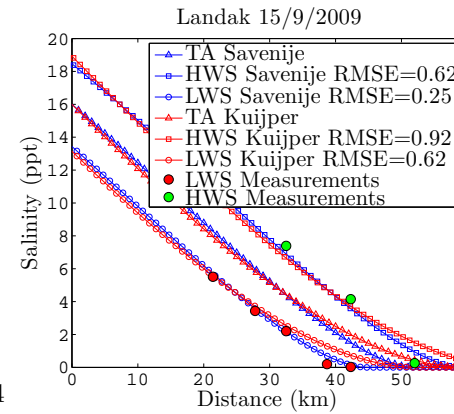
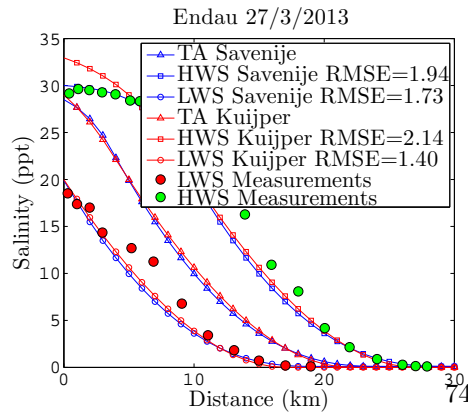
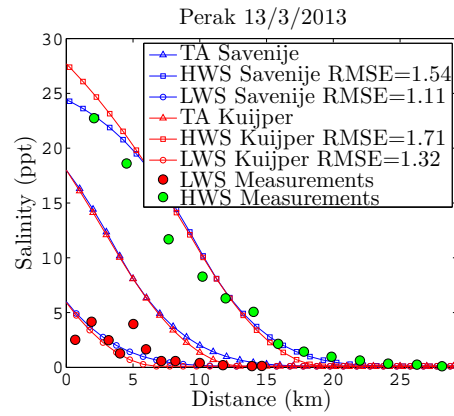
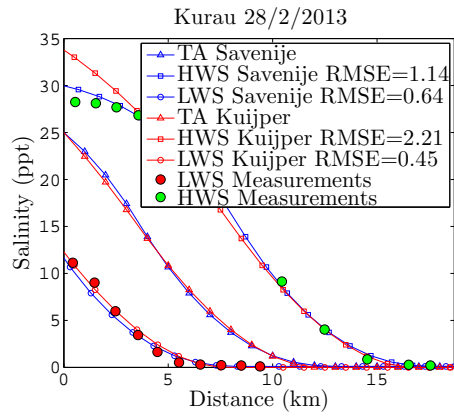
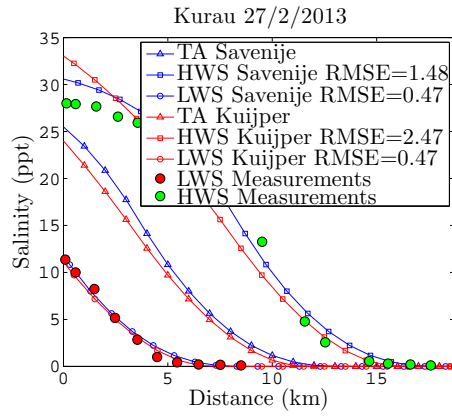
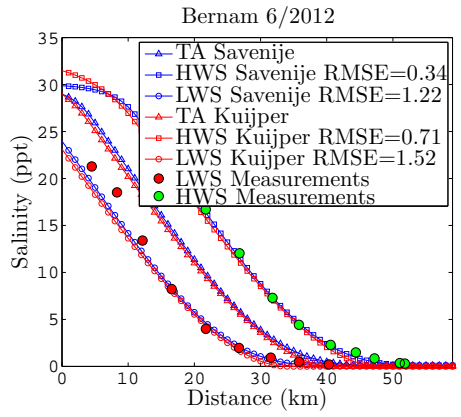
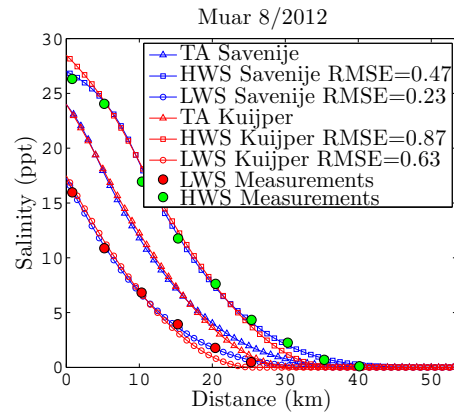
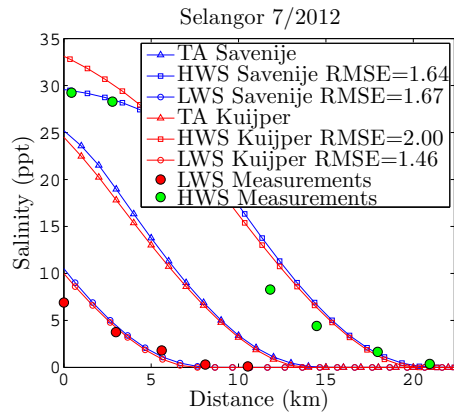
Estuaries	Date	A_o [m^2]	A_1 [m^2]	a_1 [m]	a_2 [m]	x_1 [m]	S_0 [m][ppt]	H_0 [m]	E [m]	h_0 [m]	T [s][m^3/s]	Q [$-$]	K [$-$]	a_0 [m^{-1}]	D_0 [m^2/s]
Selangor	1/7/2012	2200	1004	3440	13000	2700	25.2	4.0	12700	3.6	44400	38.0	0.34	8.00	304
Muar	1/8/2012	3300	1600	5300	118000	3900	24.0	2.0	11000	7.9	44400	11.0	0.25	8.70	95.7
Bernam	1/6/2012	15800	4500	3400	25000	4300	29.0	2.9	14000	5.2	44400	23.0	0.20	6.30	144.9
Kurau	27/2/2013	1800	660	3600	62000	3610	25.5	2.9	9400	5.2	44400	28.0	0.40	7.70	215.6
	28/2/2013	1800	660	3600	62000	3610	25.0	2.9	9400	5.2	44400	28.0	0.40	7.70	215.6
Perak	13/3/2013	20500	9200	5000	49000	4000	18.0	2.5	12500	6.0	44400	132.0	0.20	0.55	72.6
Endau	27/3/2013	8000	1900	2500	100000	3600	28.5	1.9	10000	7.0	44400	6.0	0.40	5.20	31.2
Landak	15/9/2009	2000		70000		16.0		1.6	14500	8.1	86400	10.0	0.60	23.00	230
Elbe *	21/9/2004	110435	32531	18000	66000	22000	24.0	3.0	18000	11.5	44400	200.0	0.30	1.13	225
	4/4/2004	110435	32531	18000	66000	22000	16.0	2.0	20000	11.5	44400	211.0	0.30	0.38	80
Pungue *	1/3/2002	23700		16750		22.0		6.7	24000	3.5	44440	300.0	0.30	1.60	480
	27/2/2002	23700		16750		22.0		6.1	20000	3.5	44440	300.0	0.30	1.43	430
	31/1/2002	23700		16750		21.0		6.2	20000	3.5	44440	262.0	0.30	1.91	500
	16/10/1993	23700		16750		34.0		6.4	22000	3.5	44440	10.0	0.30	40.00	400
	12/10/1993	23700		16750		35.0		3.8	15000	3.5	44440	5.0	0.30	26.00	130
	3/10/1993	23700		16750		35.0		5.3	15000	3.5	44440	10.0	0.30	34.00	340
	29/10/1982	23700		16750		32.0		6.0	16000	3.5	44440	60.0	0.30	5.17	310
	22/9/1982	23700		16750		33.0		5.2	16000	3.5	44440	26.0	0.30	9.23	240
	6/8/1982	23700		16750		32.5		5.2	12000	3.5	44440	36.0	0.30	6.11	220
	26/5/1982	23700		16750		30.0		5.0	10000	3.5	44440	50.0	0.30	4.80	240
	26/9/1980	23700		16750		32.5		6.3	20000	3.5	44440	22.0	0.30	14.55	320
Maputo *	29/5/1984	40000	5413	4000	16000	8000	30.0	3.0	12000	3.6	44440	50.0	0.40	2.30	115
	17/5/1984	40000	5413	4000	16000	8000	29.0	3.4	14000	3.6	44440	50.0	0.40	1.60	80
	19/4/1984	40000	5413	4000	16000	8000	27.0	3.3	12000	3.6	44440	115.0	0.40	1.61	185
	28/4/1984	40000	5413	4000	16000	8000	34.0	3.5	13000	3.6	44440	21.0	0.40	1.90	40

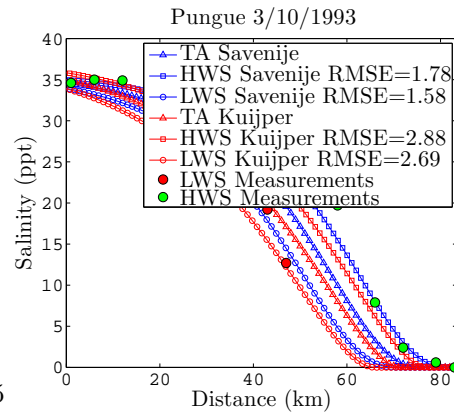
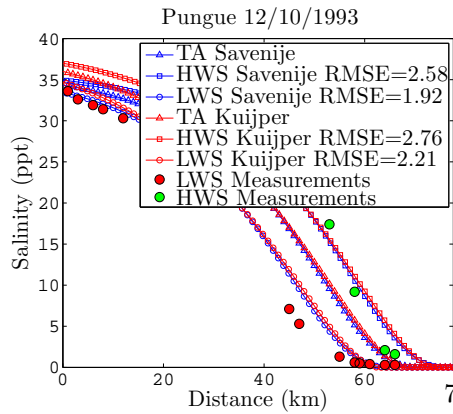
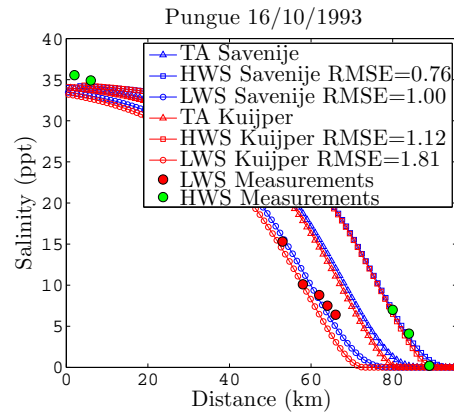
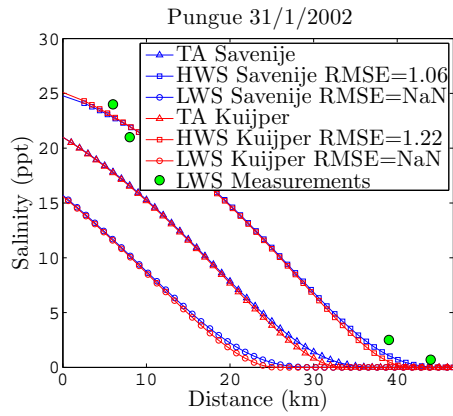
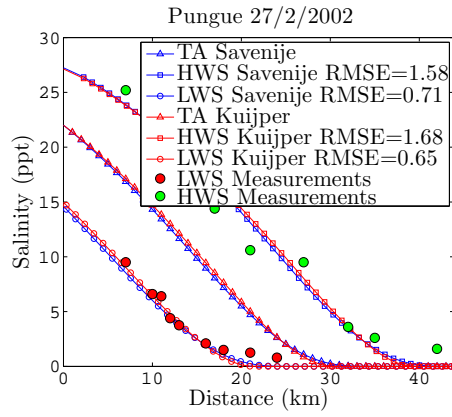
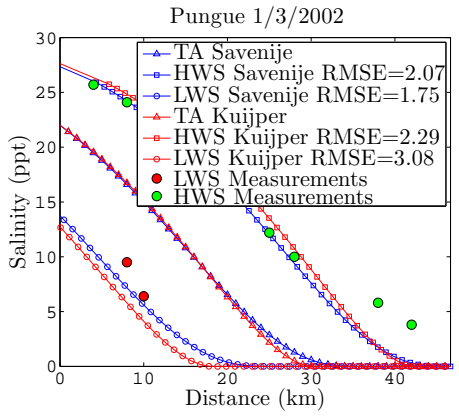
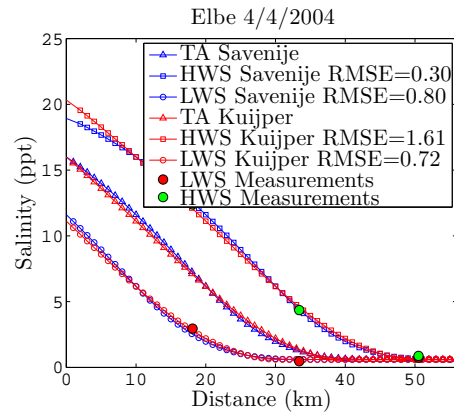
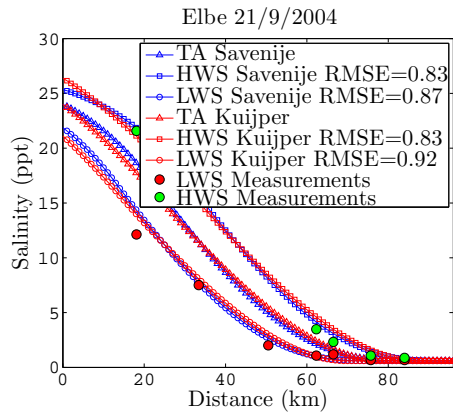
Thames *	15/7/1984	40000	5413	4000	16000	8000	34.0	2.1	6000	3.6	44440	4.8	0.40	4.58	22
Corantijn *	7/4/1949	58500		23000		32.0		5.3	14000	7.1	44400	40.0	0.20	0.20	170
	9/12/1978	69000	26755	19000	64000	18000	23.5	1.8	10000	6.5	44440	120.0	0.21	0.96	115
	14/12/1978	69000	26755	19000	64000	18000	18.5	2.3	13000	6.5	44440	130.0	0.21	1.08	140
	20/12/1978	69000	26755	19000	64000	18000	16.5	1.6	9000	6.5	44440	220.0	0.21	0.95	210
Sinnamary *	12/11/1993	3500	1123	2640	39000	3000	20.0	2.6	9000	3.8	44440	168.0	0.55	2.08	350
	27/4/1994	3500	1123	2640	39000	3000	17.0	2.9	10000	3.8	44440	148.0	0.55	2.30	340
	2/11/1994	3500	1123	2640	39000	3000	28.0	2.7	7500	3.8	44440	112.0	0.55	5.54	620
	3/11/1994	3500	1123	2640	39000	3000	23.0	2.9	9500	3.8	44440	112.0	0.55	3.75	420
MaeKlong *	8/3/1977	1400		102000		30.0		1.5	10000	5.2	44400	60.0	0.30	6.75	405
	9/4/1977	1400		102000		35.0		2.1	7000	5.2	44400	36.0	0.30	10.00	360
Lalang *	20/10/1989	2070		217000		18.0		2.6	27000	10.6	86400	120.0	0.70	12.50	1500
Limpopo *	4/4/1980	1710	1196	50400	130000	18000	24.0	1.1	7000	6.3	44440	150.0	0.50	9.00	1350
	31/12/1982	1710	1196	50400	130000	18000	35.0	1.1	8000	6.3	44440	2.0	0.50	33.50	67
	14/7/1994	1710	1196	50400	130000	18000	32.0	1.0	7000	6.3	44440	5.0	0.50	20.00	100
	24/7/1994	1710	1196	50400	130000	18000	35.0	0.9	6800	6.3	44440	5.0	0.50	25.00	125
	10/8/1994	1710	1196	50400	130000	18000	35.0	1.0	7100	6.3	44440	3.0	0.50	29.00	87
Tha Chin *	16/4/1981	3000	2674	87000	135000	10000	23.0	1.6	12000	5.3	86400	55.0	0.40	11.82	650
	27/2/1986	3000	2674	87000	135000	10000	24.0	2.6	20000	5.3	44400	40.0	0.40	6.75	270
	1/3/1986	3000	2674	87000	135000	10000	30.0	1.9	14000	5.3	86400	40.0	0.40	8.00	320
	13/8/1987	3000	2674	87000	135000	10000	23.0	2.0	15000	5.3	44400	39.0	0.40	6.15	240
ChaoPhya *	5/6/1962	5300		109000		22.0		2.2	22000	7.2	86400	63.0	0.80	7.14	450
	16/1/1987	5300		109000		14.0		2.5	15000	7.2	86400	180.0	0.80	3.17	570
	23/2/1983	5300		109000		21.0		1.6	19000	7.2	86400	100.0	0.80	5.80	580
	29/1/1983	5300		109000		24.0		2.4	26000	7.2	86400	90.0	0.80	7.22	650
Incomati *	5/9/1982	8100	1253	7500	42000	14000	35.0	1.4	7000	2.9	44440	2.0	0.20	14.00	28
	23/6/1993	8100	1253	7500	42000	14000	35.0	1.4	8000	2.9	44440	4.0	0.20	8.00	32
	7/7/1993	8100	1253	7500	42000	14000	35.0	2.6	9000	2.9	44440	4.0	0.20	7.50	30

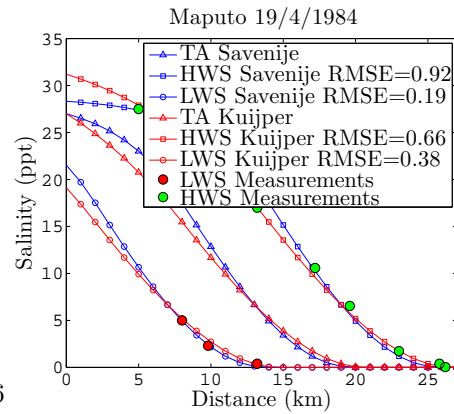
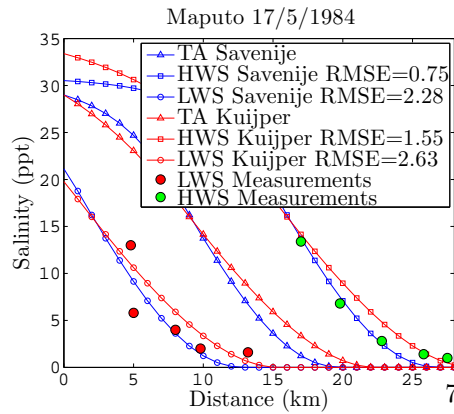
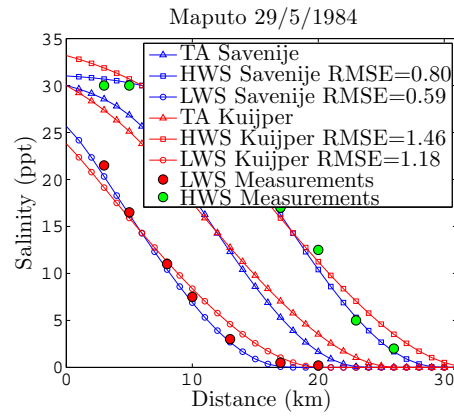
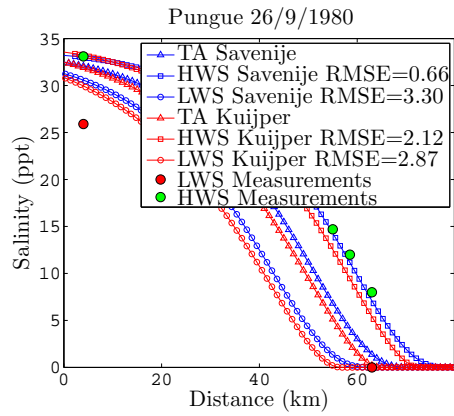
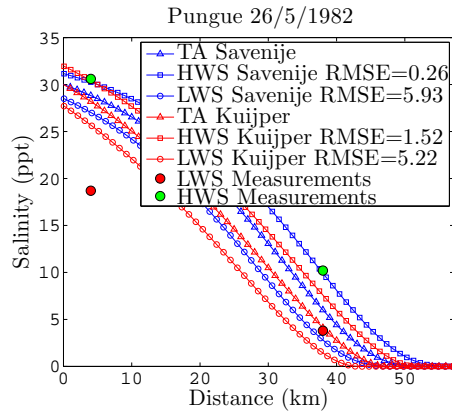
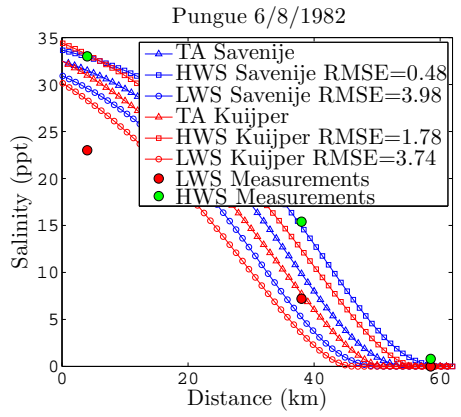
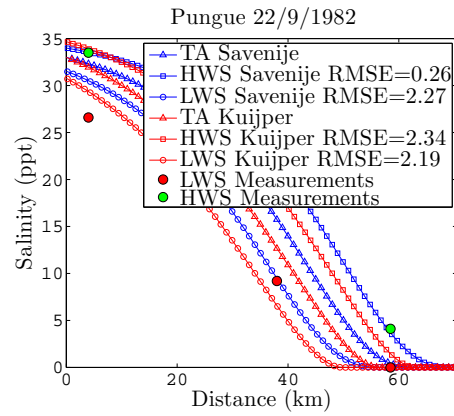
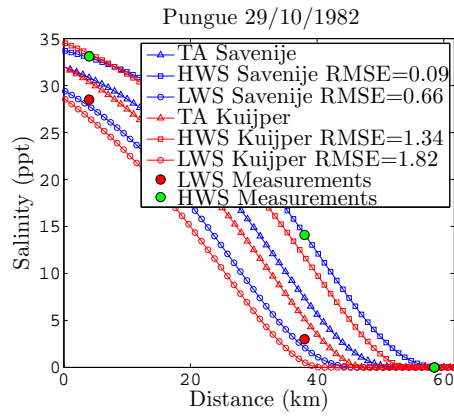
Yangtze (north)	21/6/2006	140000	64872	6500	24000	5000	33.0	3.1	24000	7.0	44400	46.2	0.10	2.16	100
	17/2/2003	140000	64872	6500	24000	5000	30.0	3.4	20500	7.0	44400	109.9	0.10	2.88	316
	1/9/2001	140000	64872	6500	24000	5000	30.0	3.4	18400	7.0	44400	26.4	0.10	6.55	172.8
Tra Ly	15/12/2008	1940	975	16000	68000	11000	20.0	3.9	14000	4.1	86400	140.0	0.21	8.39	1175
	9/3/2009	1940	975	16000	68000	11000	17.0	2.8	14000	4.1	86400	110.0	0.21	7.09	780
Red River	15/12/2008	4350	2055	12000	135000	9000	22.0	4.0	13000	4.5	86400	350.0	0.18	3.71	1300
	9/3/2009	4350	2055	12000	135000	9000	20.0	2.8	11000	4.5	86400	275.0	0.18	3.09	850
Ninh Co	15/12/2008	2950	1579	16000	45000	10000	24.0	4.2	20000	3.4	86400	84.0	0.27	10.12	850
	9/3/2009	2950	1579	16000	45000	10000	24.0	2.9	17000	3.4	86400	66.0	0.27	9.09	600
Day	15/12/2008	2950	1945	12000	95000	5000	20.0	4.2	15000	5.1	86400	308.0	0.24	4.87	1500
	9/3/2009	2950	1945	12000	95000	5000	15.0	2.9	15000	5.1	86400	242.0	0.24	4.75	1150
Edisto	12/7/2010	14000	5150	2000	15000	2000	49.5	2.3	10000	4.0	44400	15.0	0.35	10.67	160
	13/7/2010	14000	5150	2000	15000	2000	45.0	2.3	10000	4.0	44400	14.0	0.35	11.79	165
	14/7/2010	14000	5150	2000	15000	2000	47.0	2.3	10000	4.0	44400	25.0	0.35	11.00	275
	15/7/2010	14000	5150	2000	15000	2000	47.0	2.3	10000	4.0	44400	25.0	0.35	10.20	255
Westerschelde	2/11/2000	273000	4643	27000	18000	110000	33.0	4.0	12000	10.5	44400	220.0	0.25	2.05	450
	Model Duc	273000	4643	27000	18000	110000	33.0	3.5	11500	10.5	44400	150.0	0.25	1.00	150
Pangani	27/10/2007	1070		14300			27.0	4.2	19000	3.0	44440	14.8	0.90	33.78	500
	11/12/2007	1070		14300			28.0	3.0	15000	3.0	44440	11.1	0.90	27.03	300

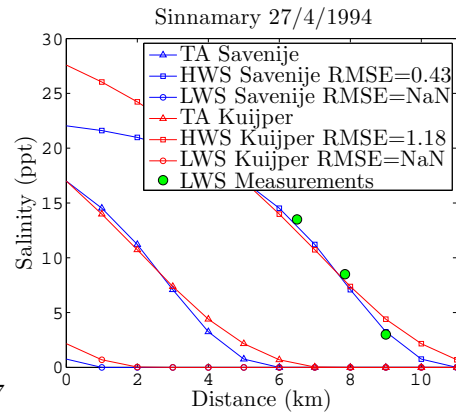
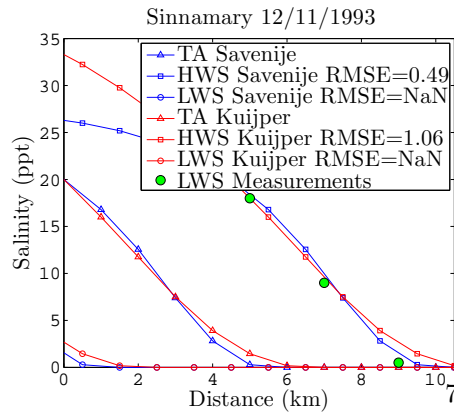
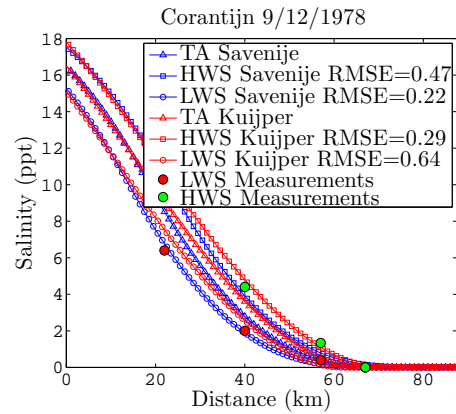
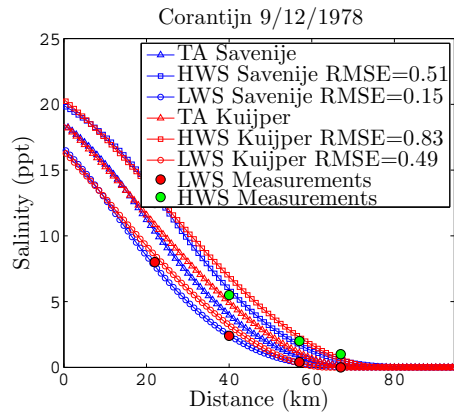
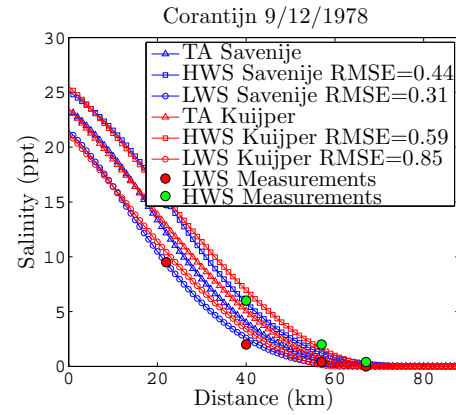
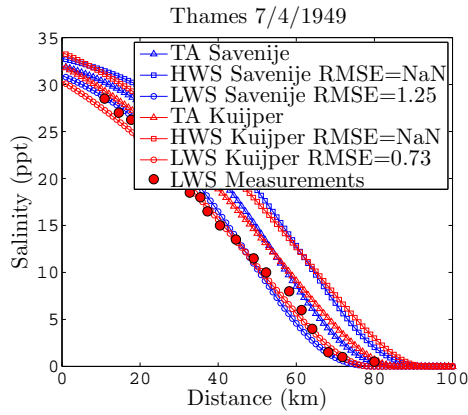
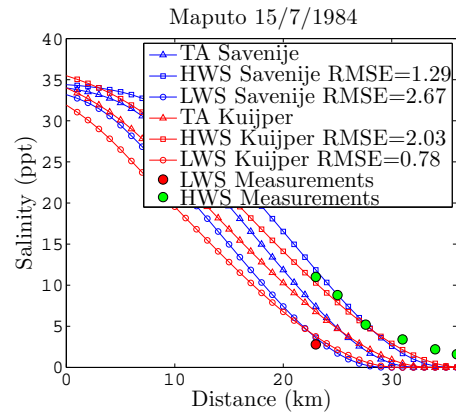
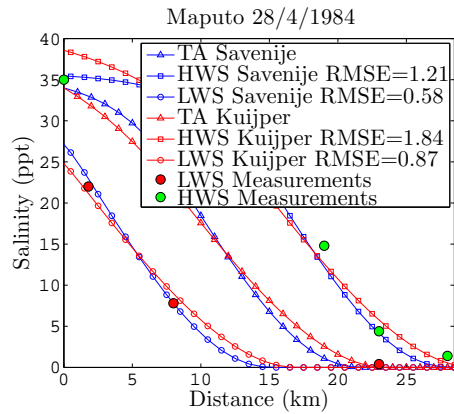
* Calibration set

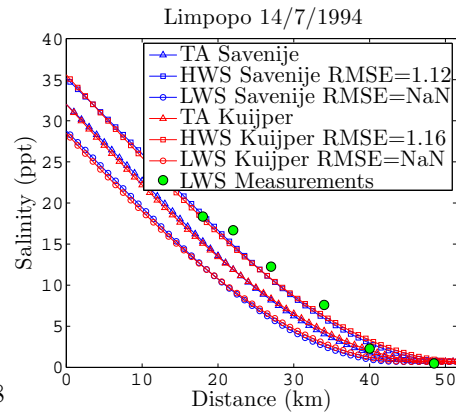
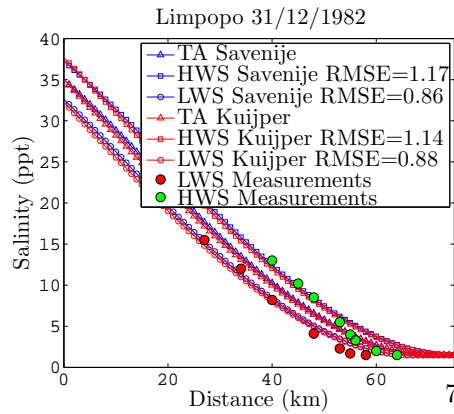
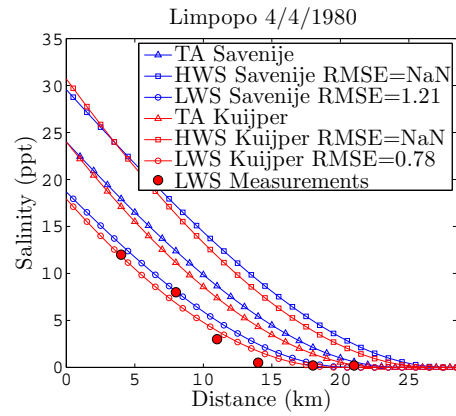
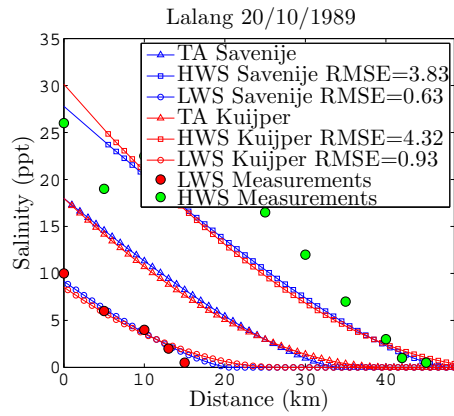
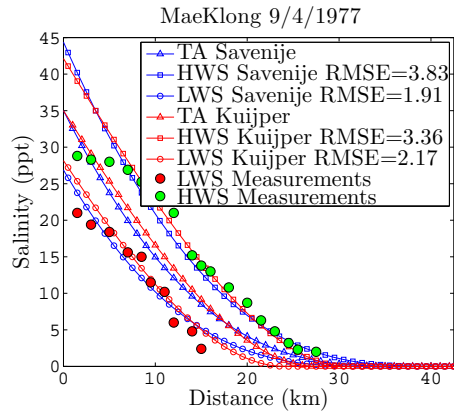
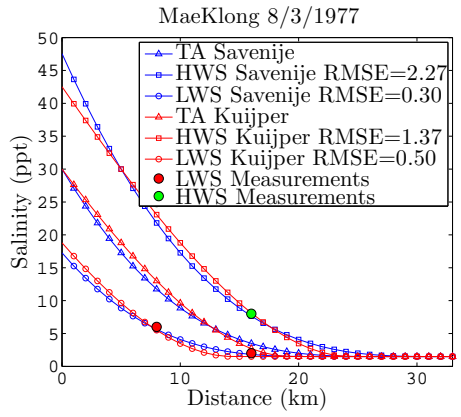
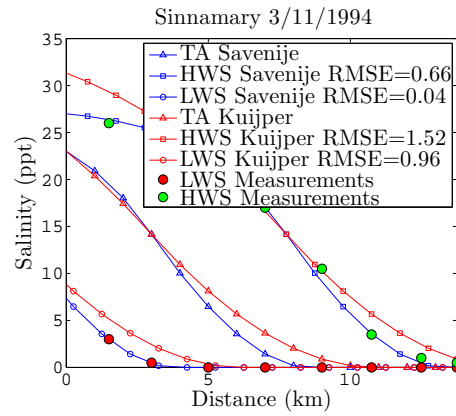
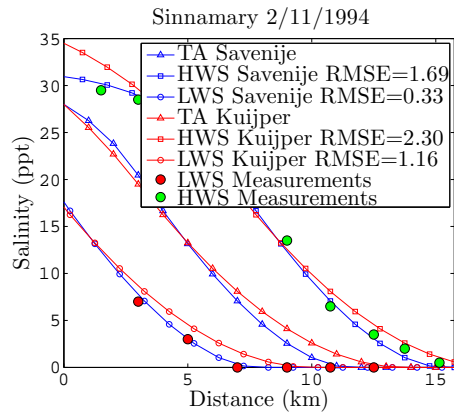
B Salinity Graphs

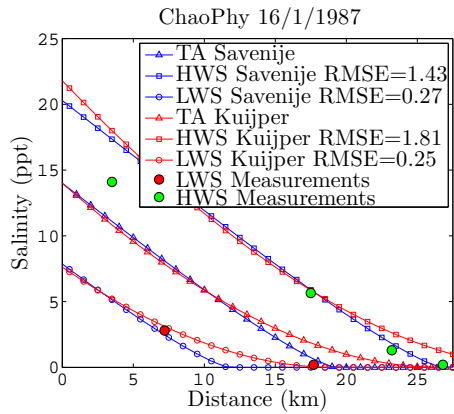
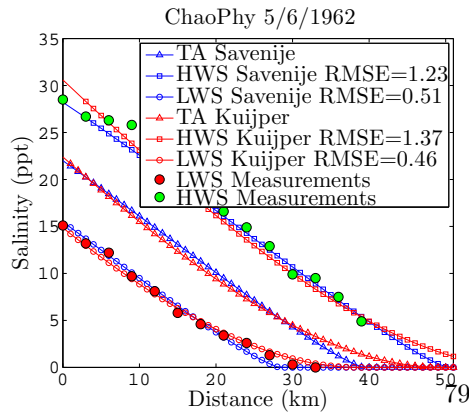
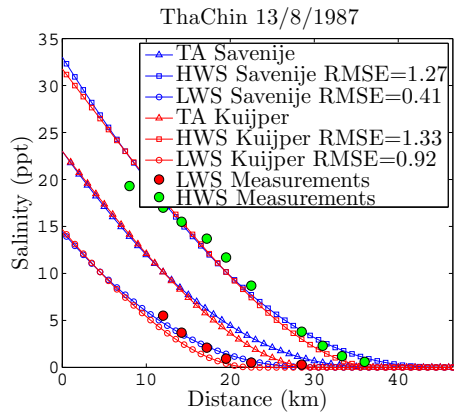
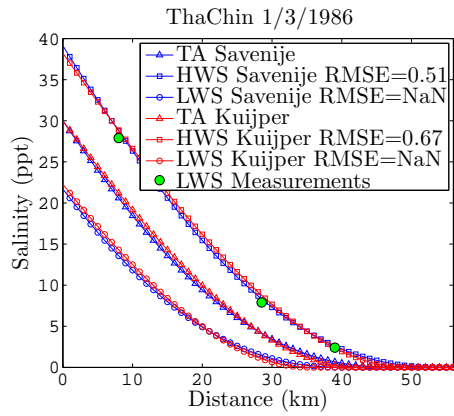
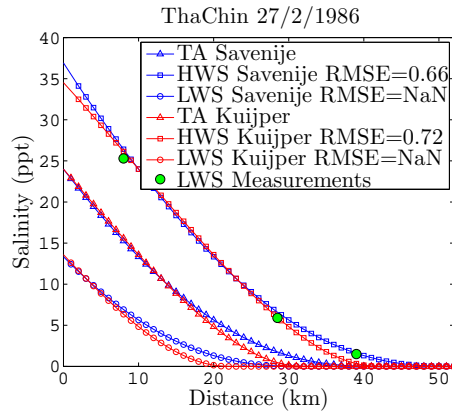
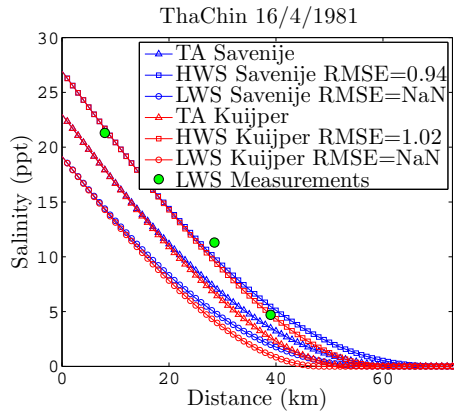
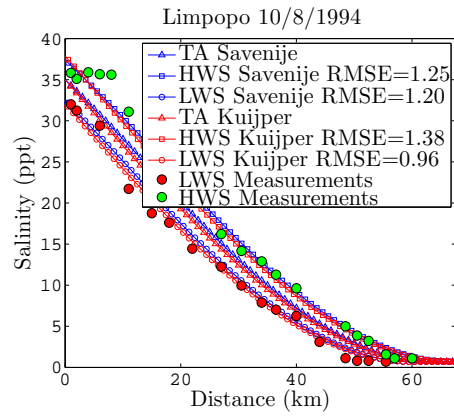
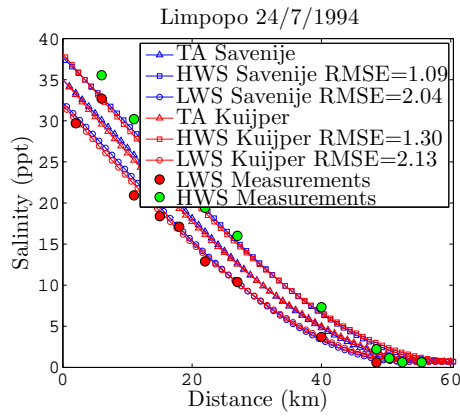


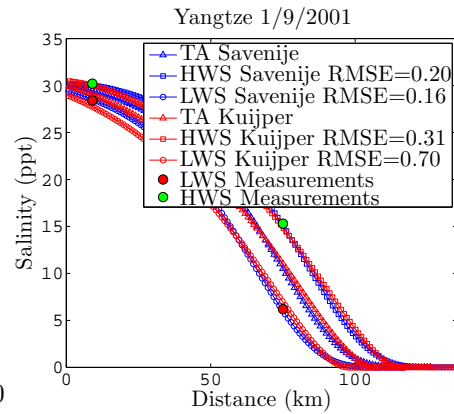
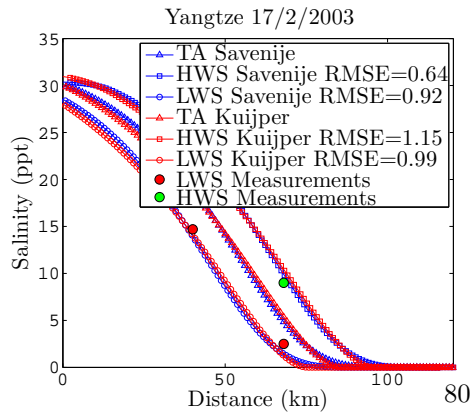
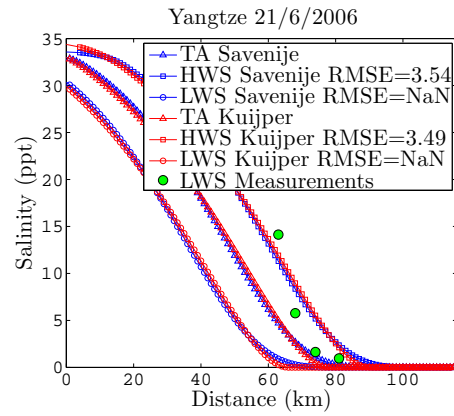
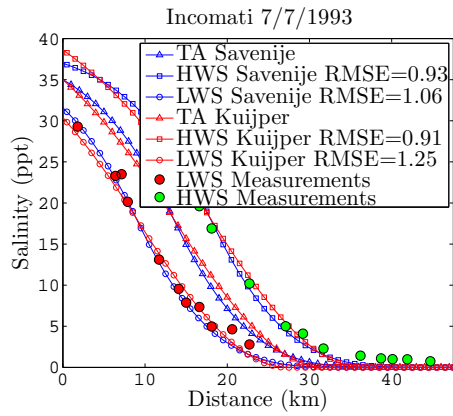
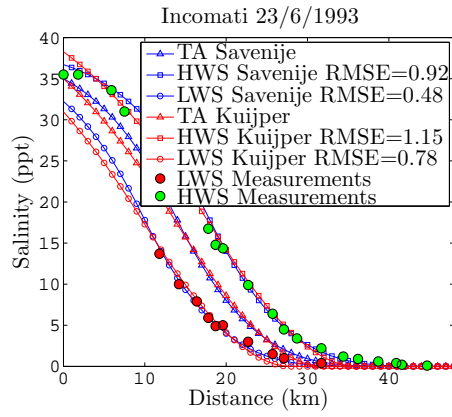
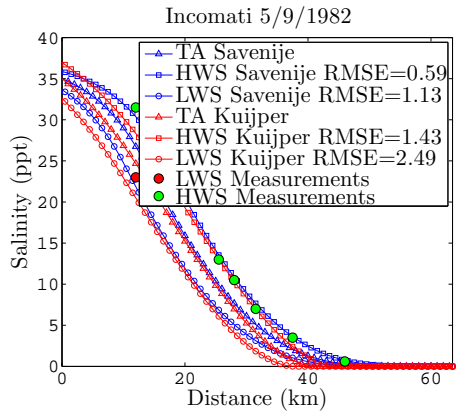
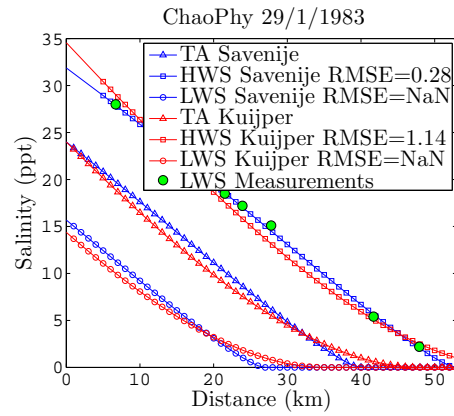
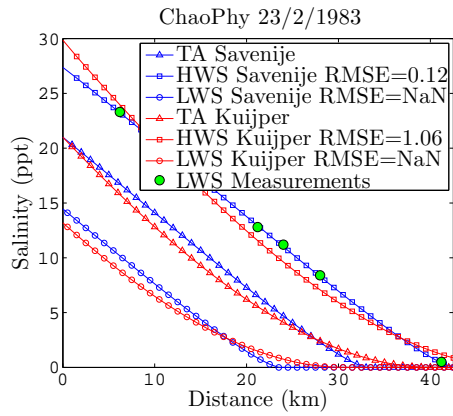


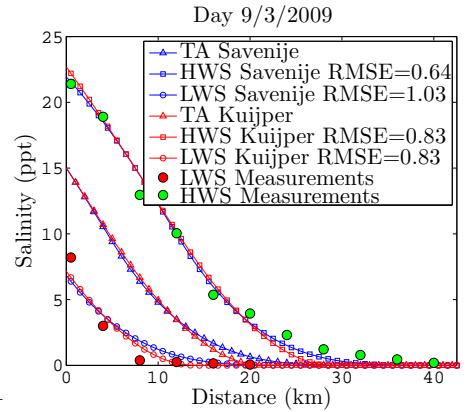
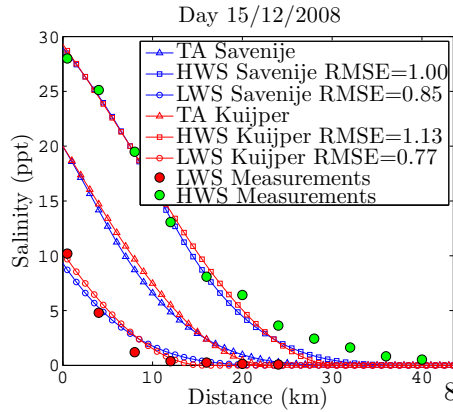
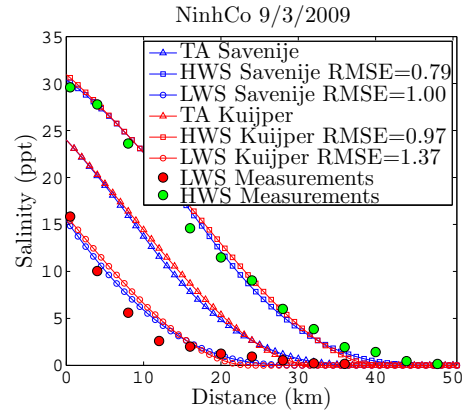
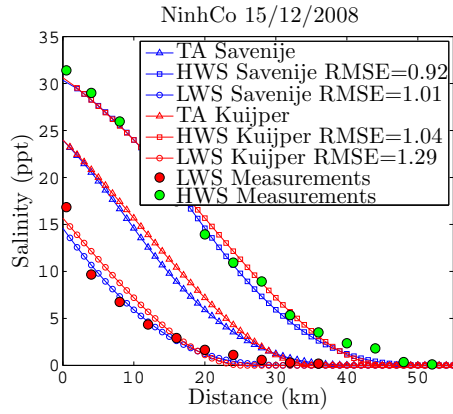
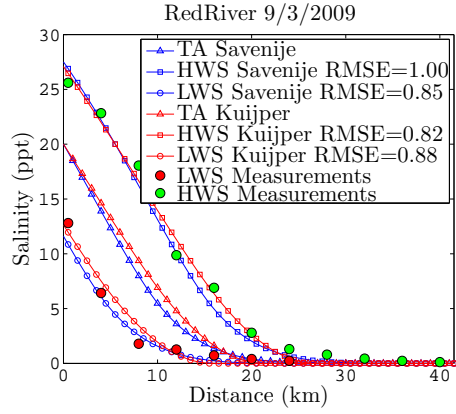
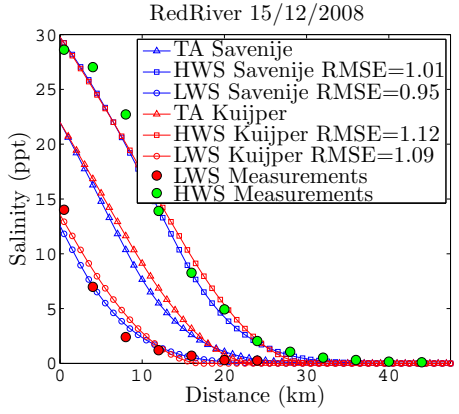
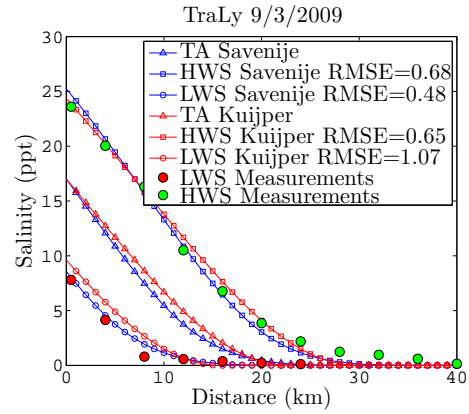
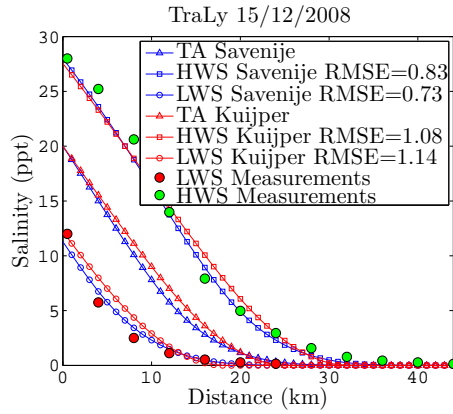


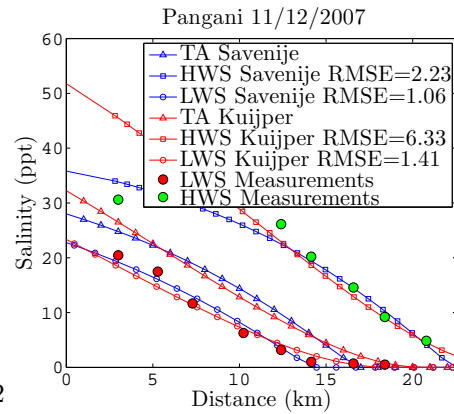
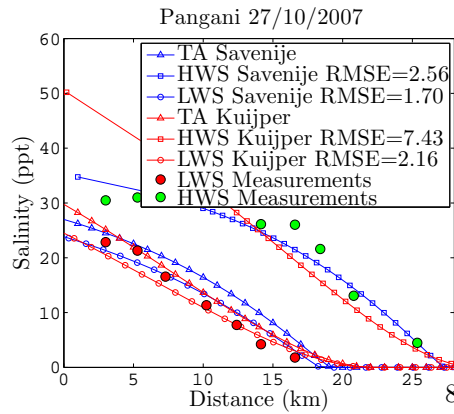
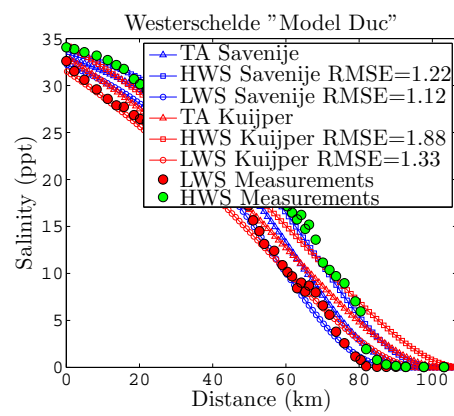
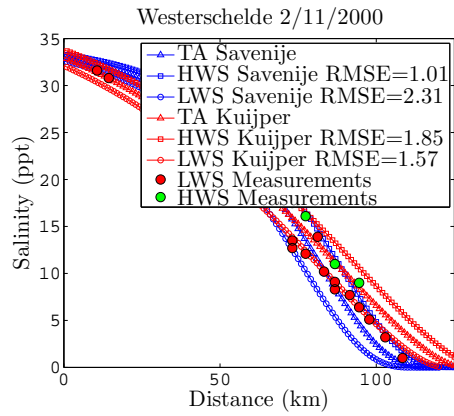
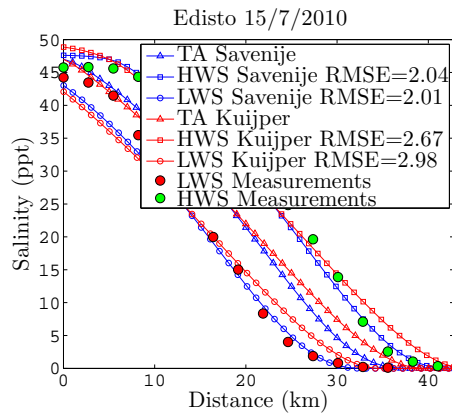
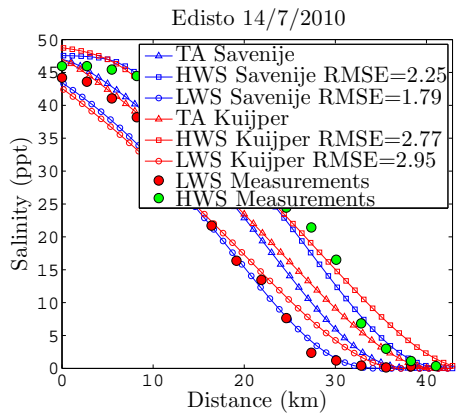
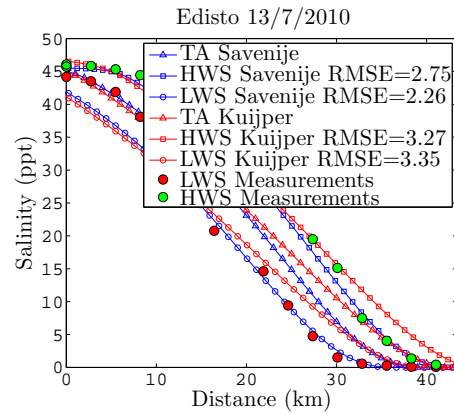
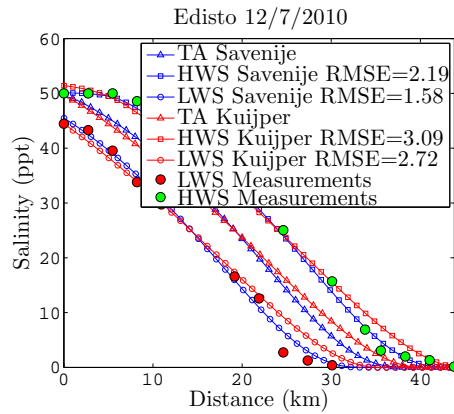




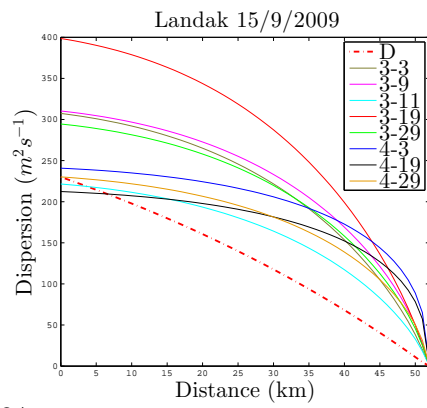
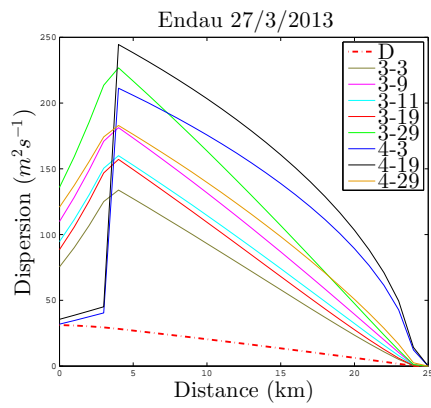
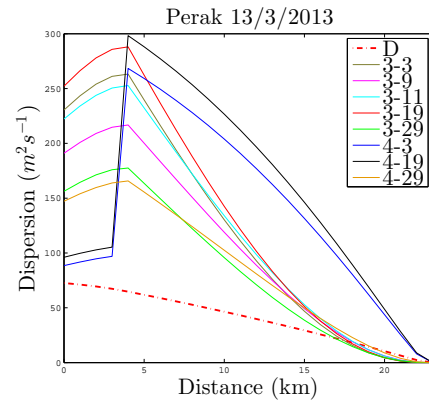
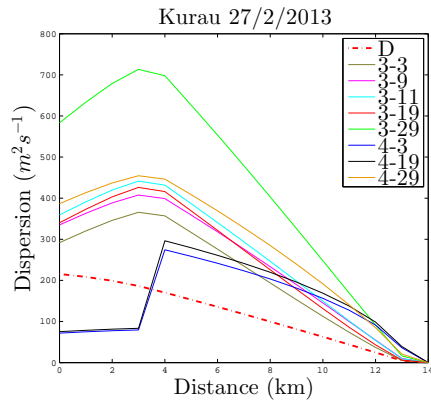
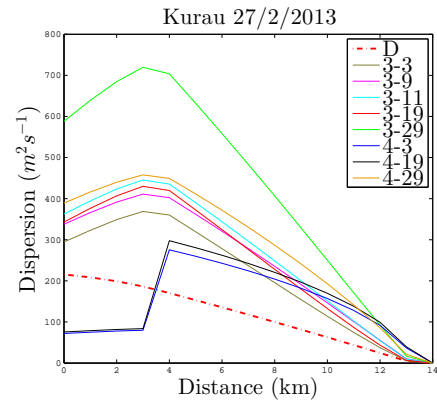
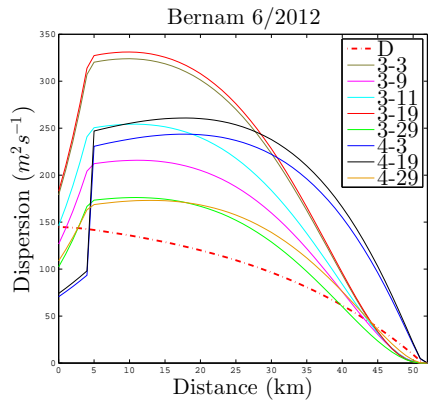
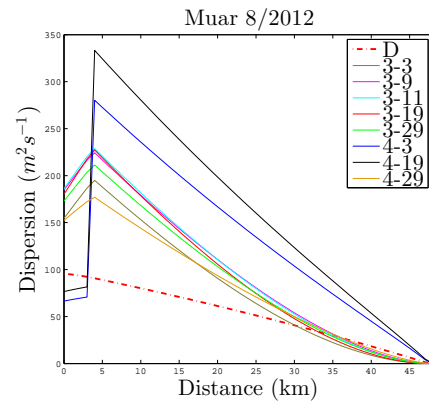
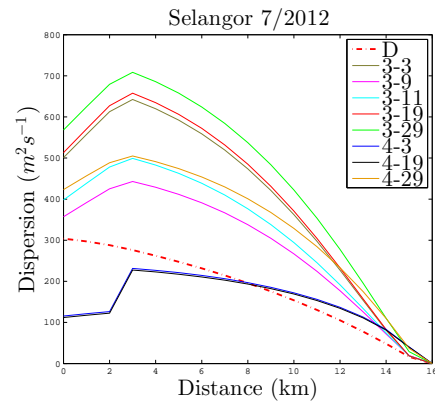


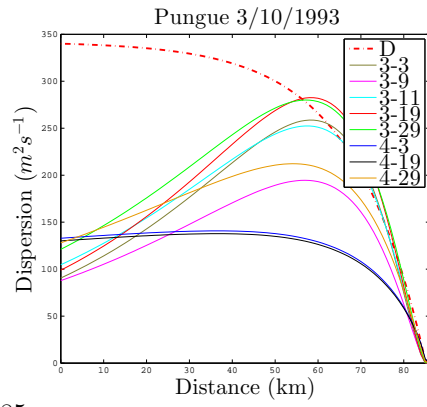
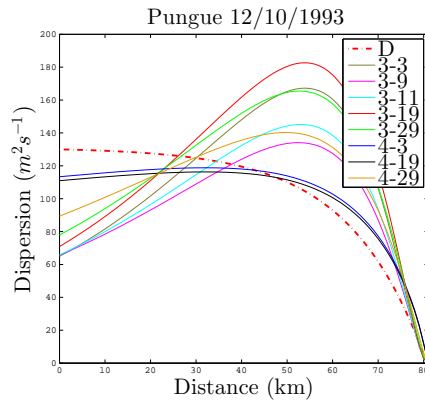
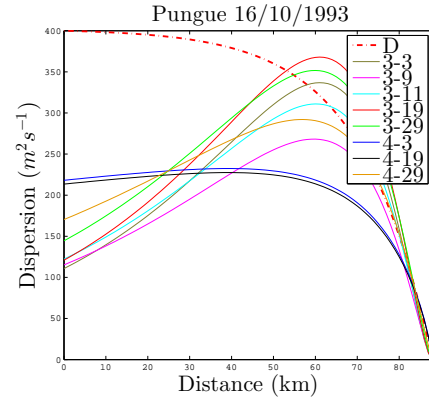
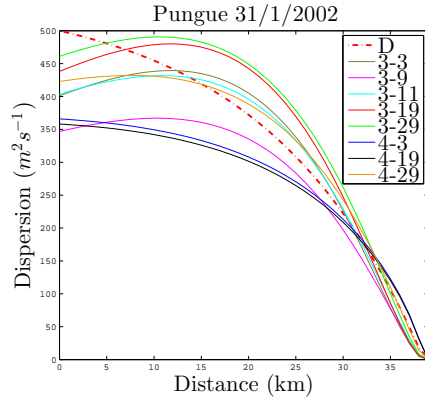
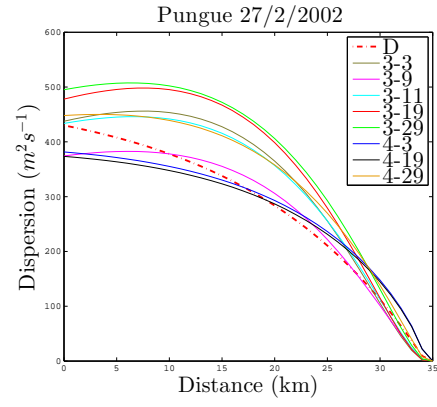
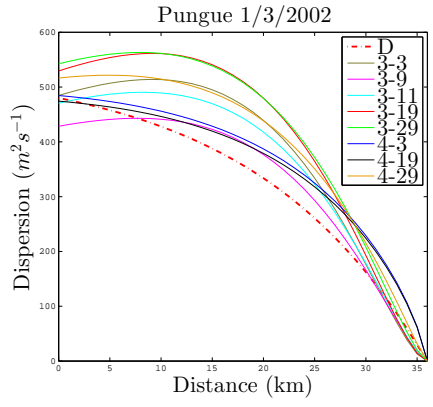
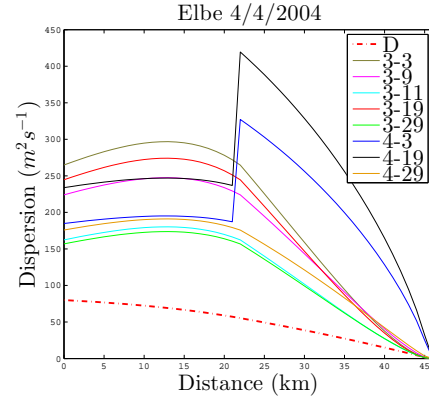
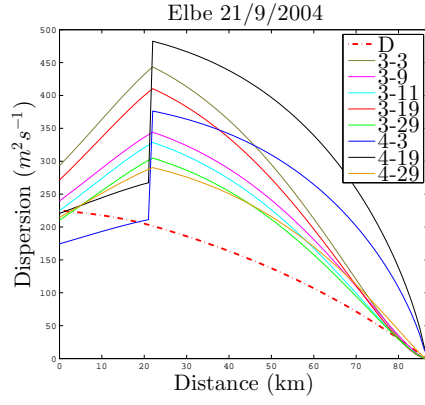


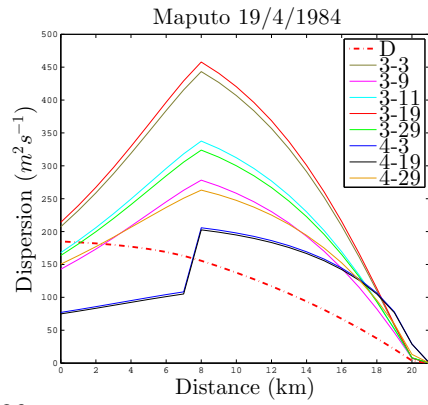
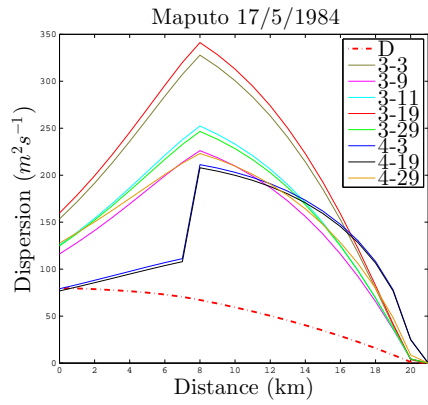
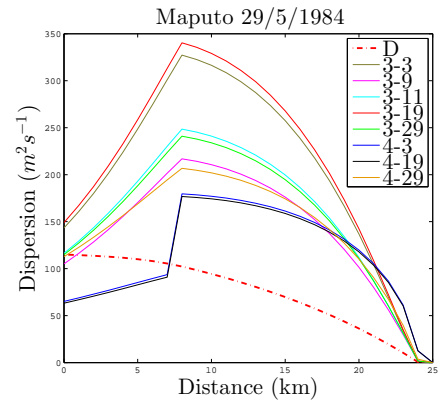
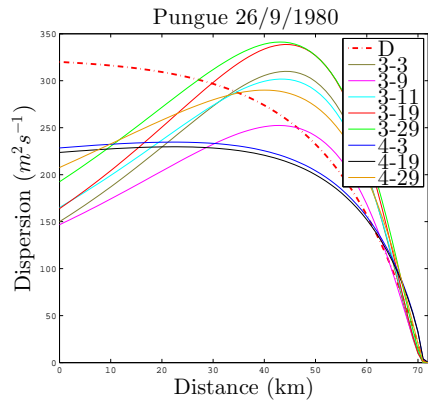
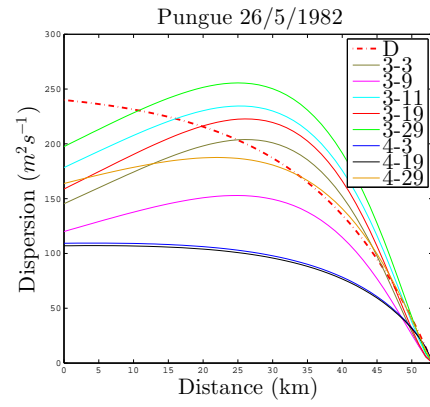
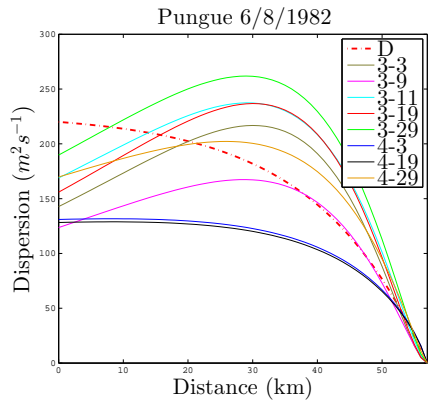
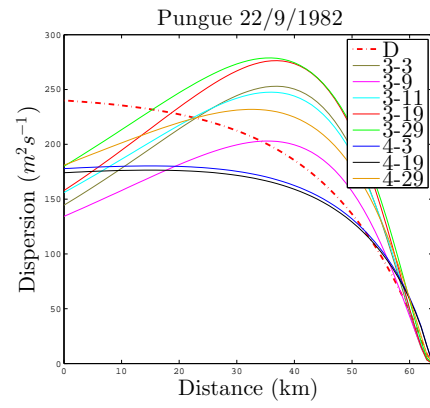
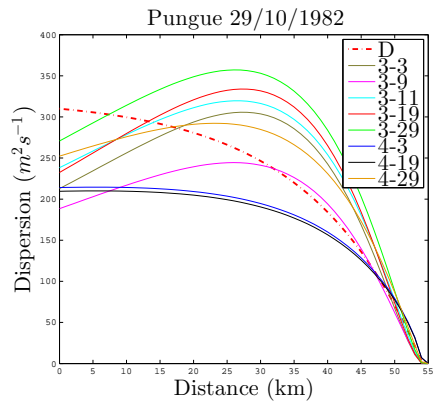


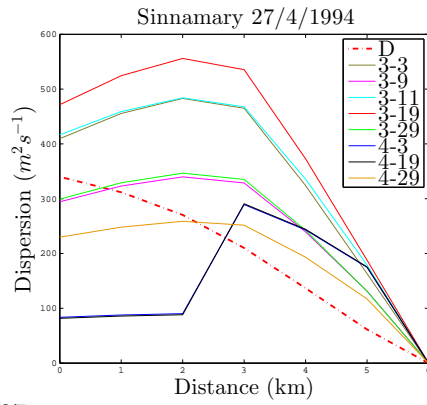
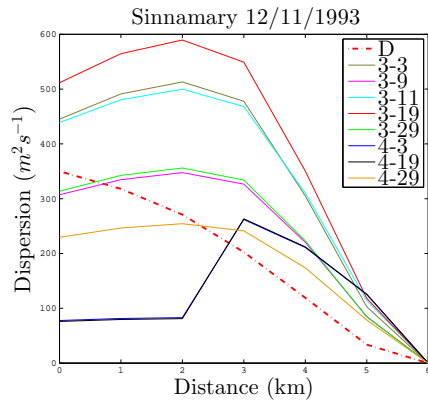
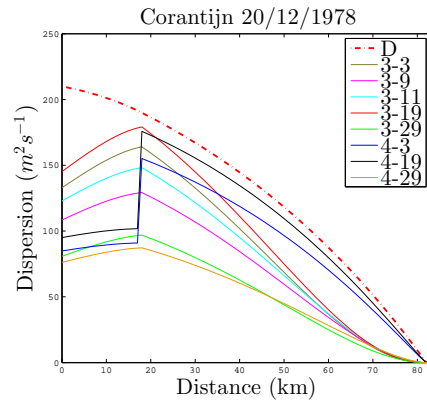
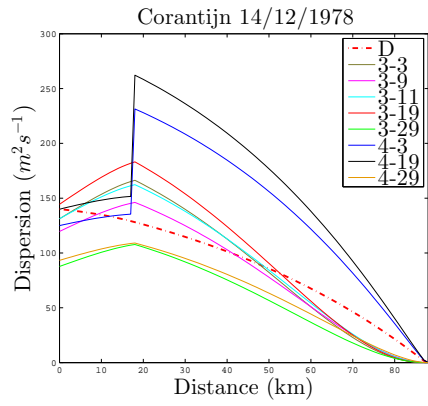
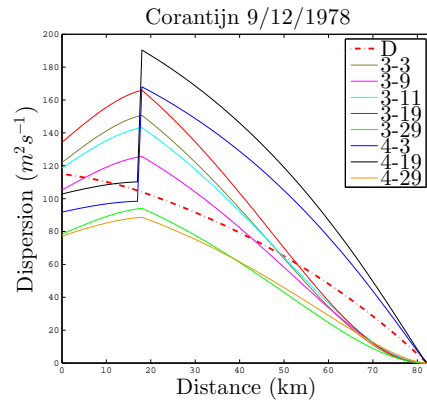
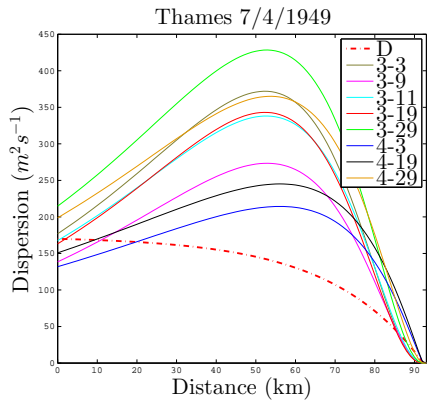
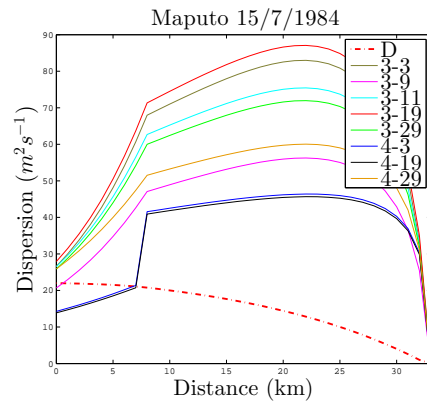
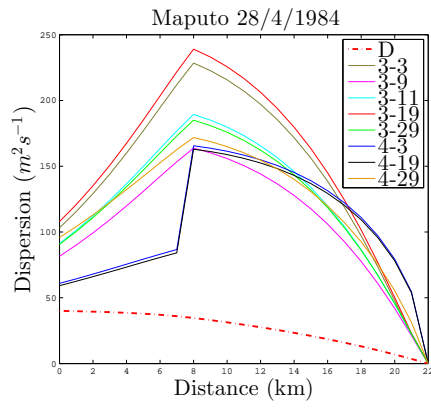


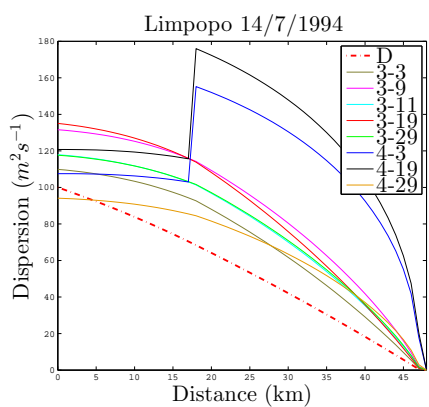
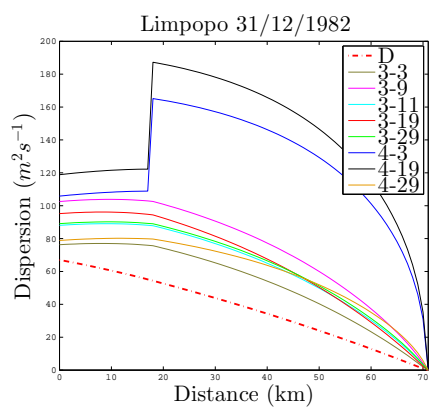
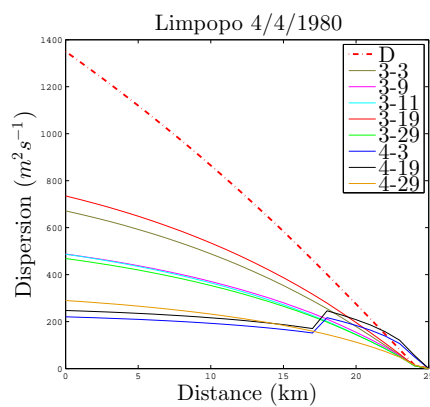
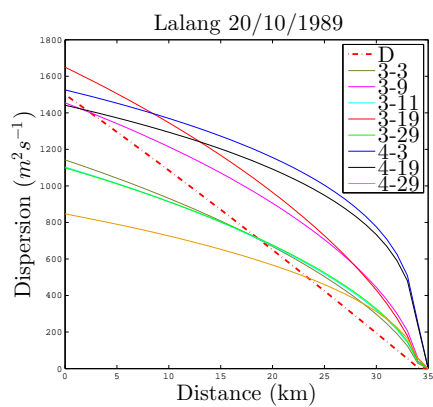
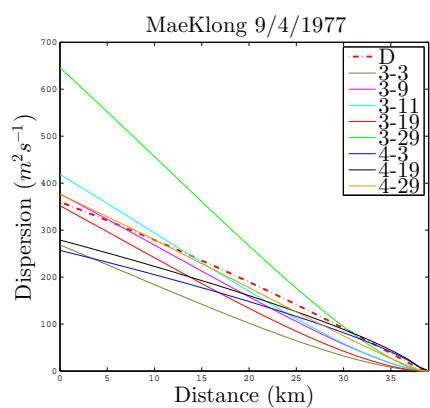
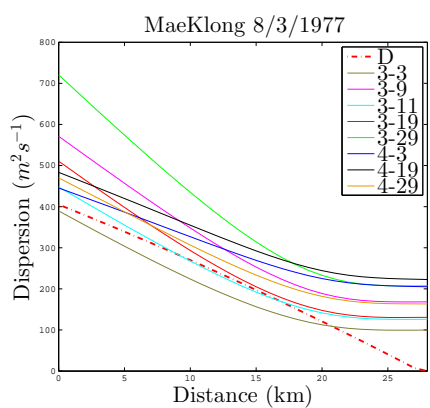
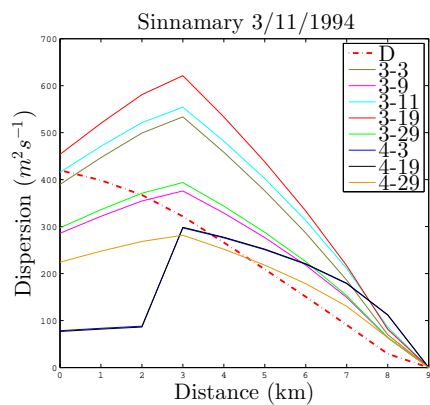
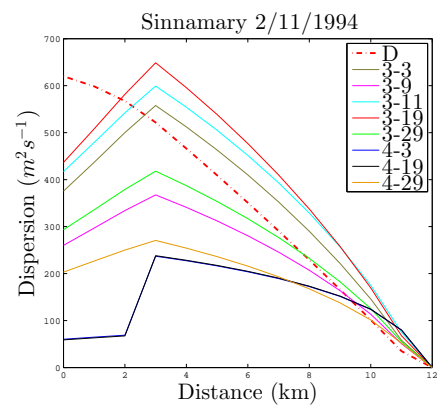
C Graphs Local Validity Savenije

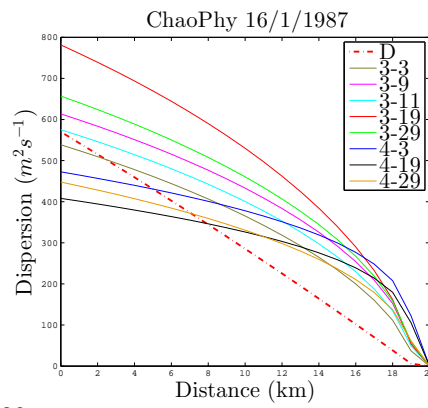
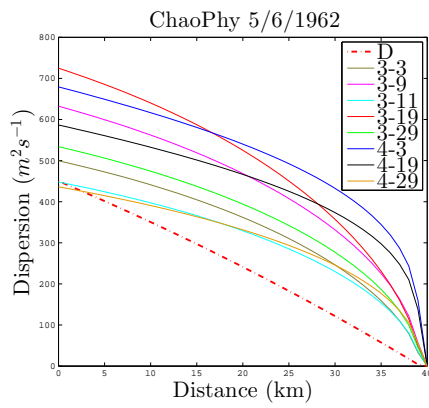
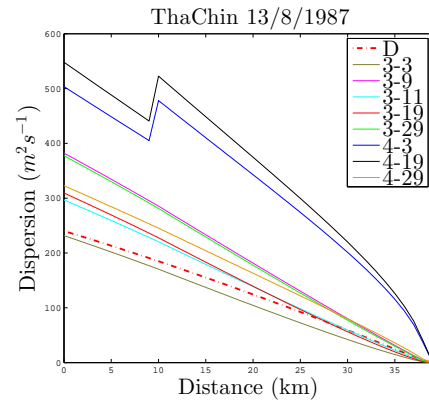
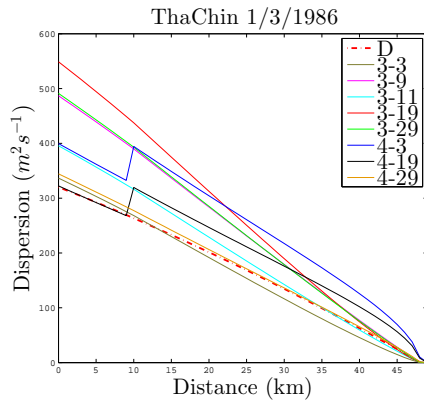
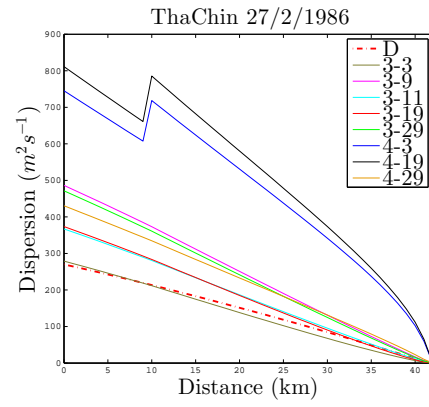
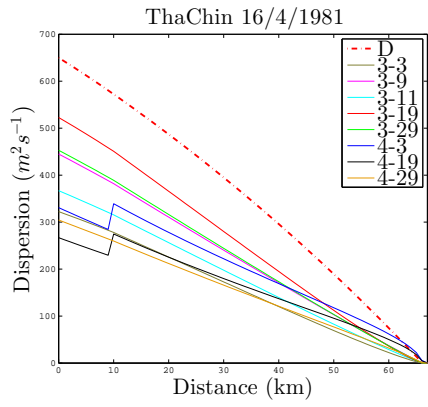
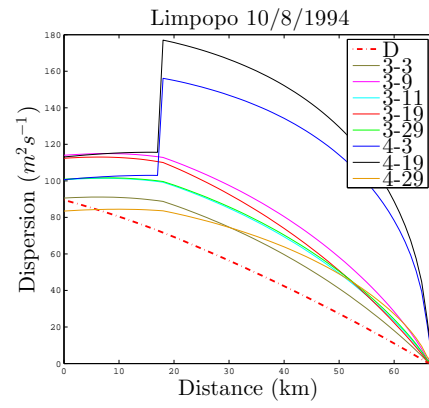
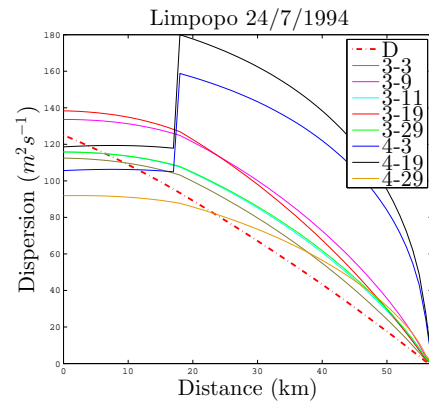


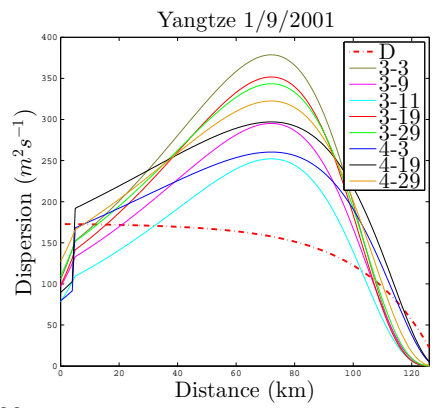
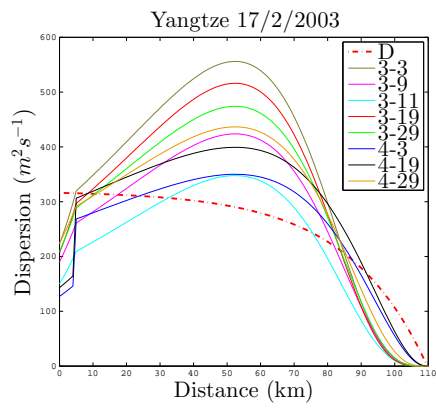
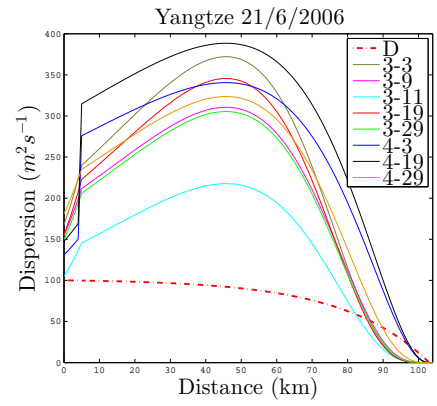
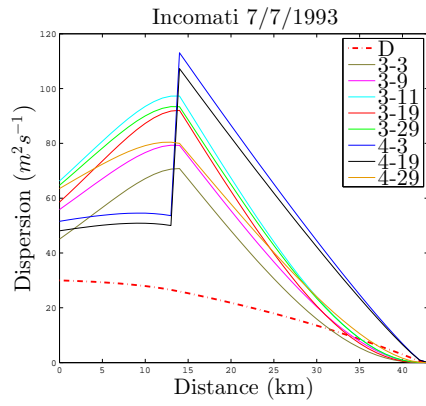
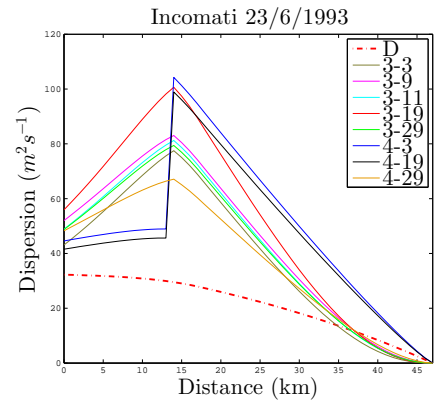
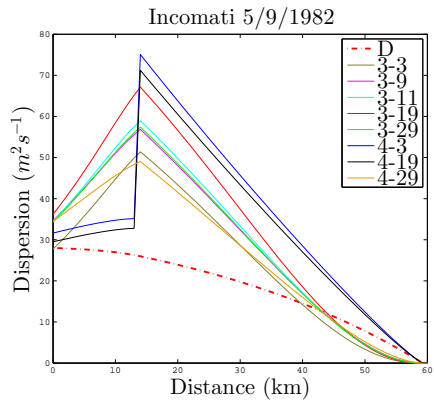
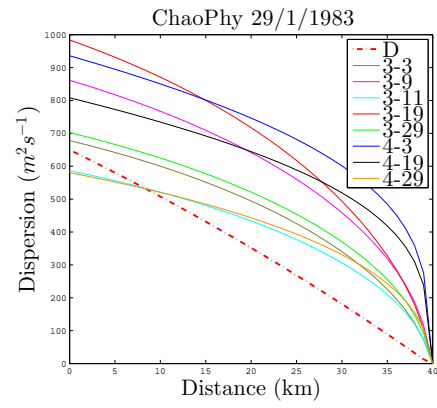
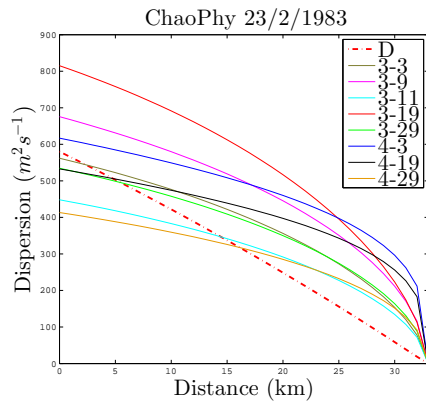


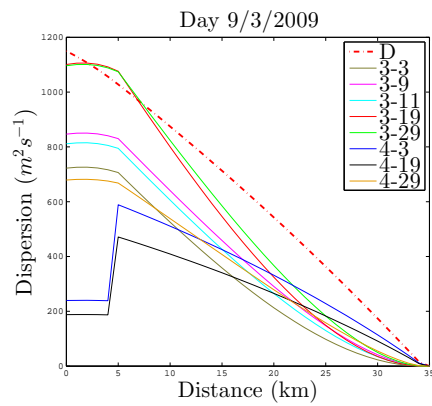
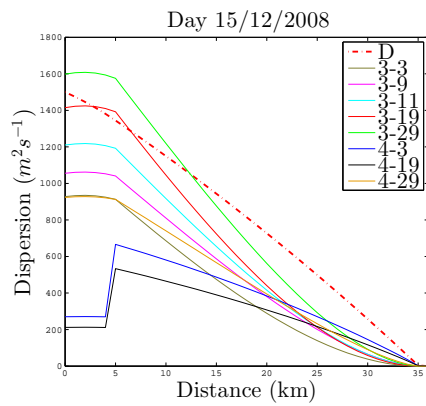
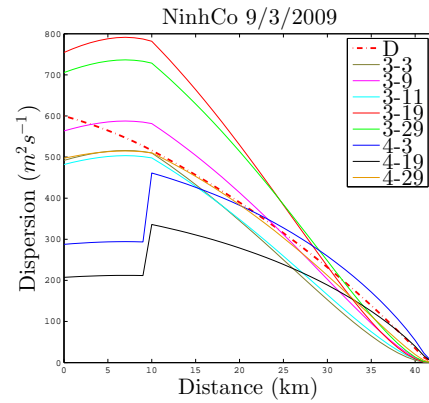
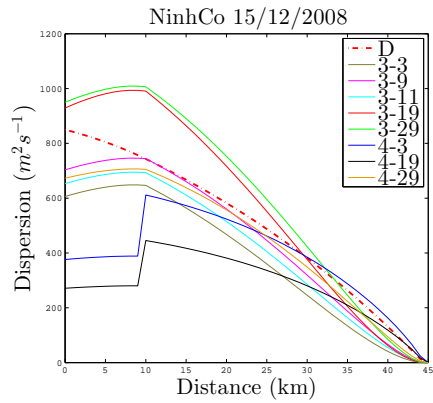
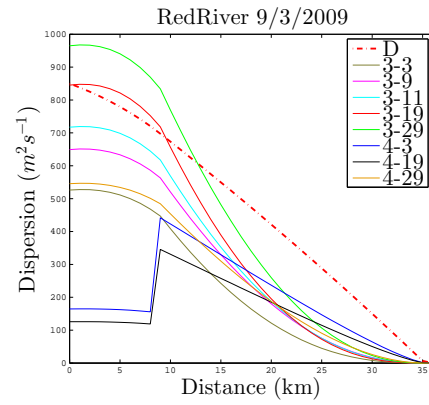
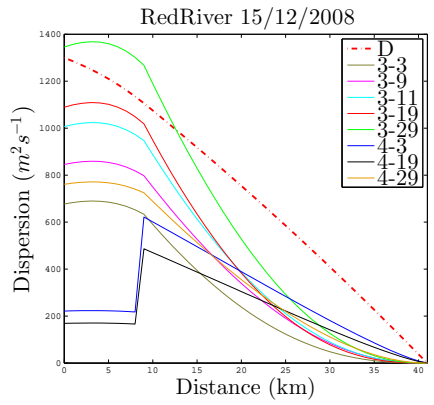
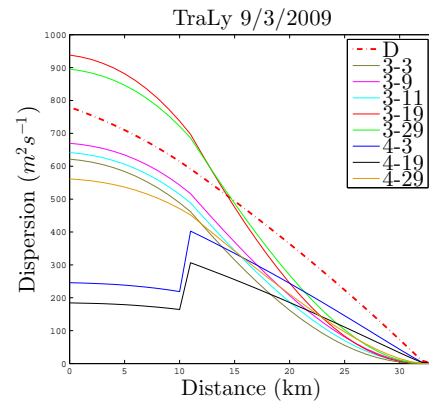
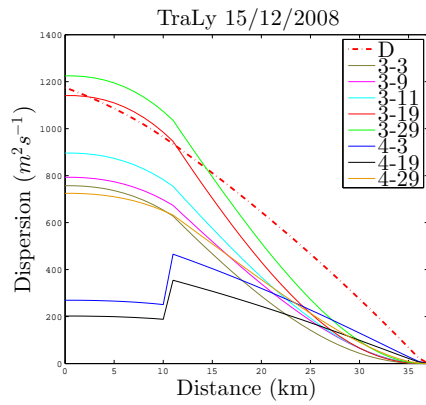


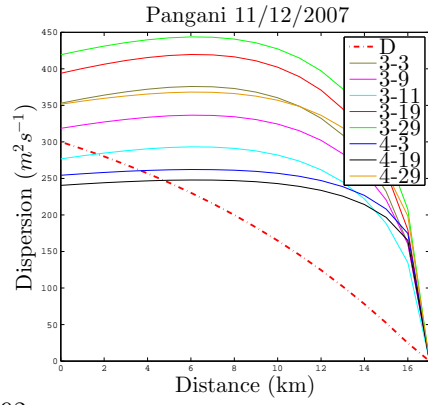
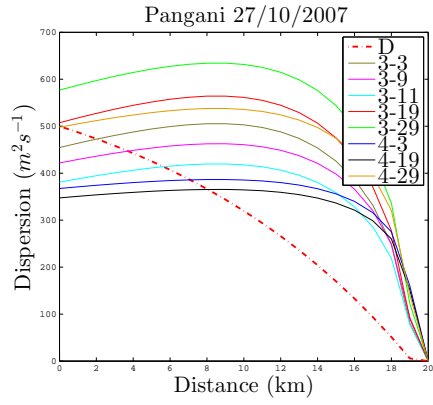
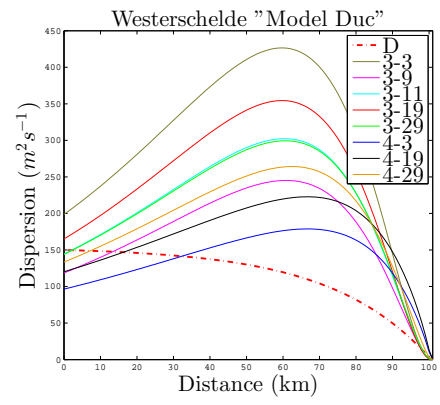
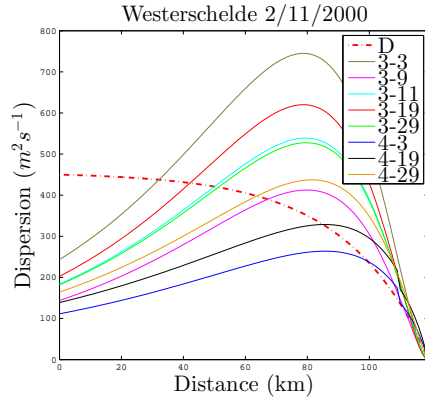
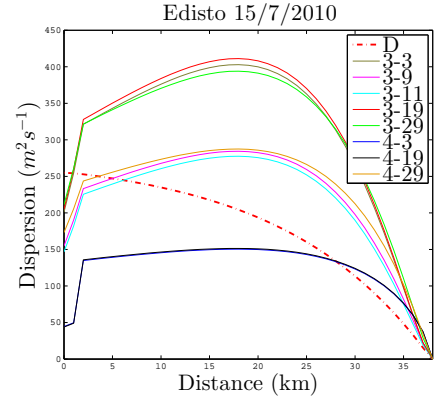
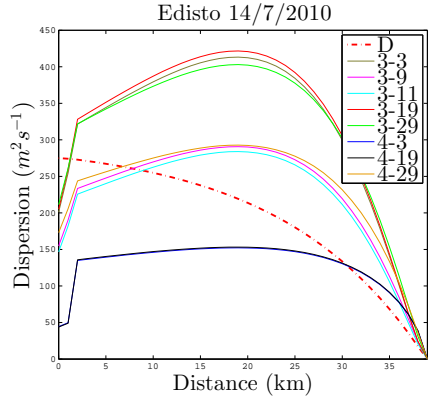
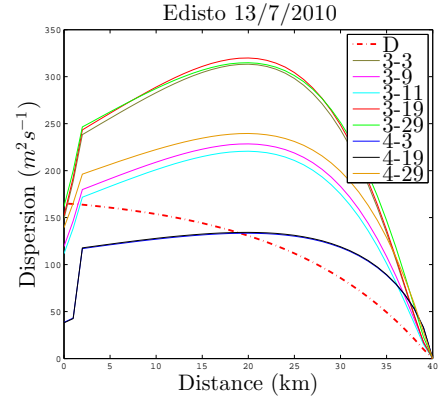
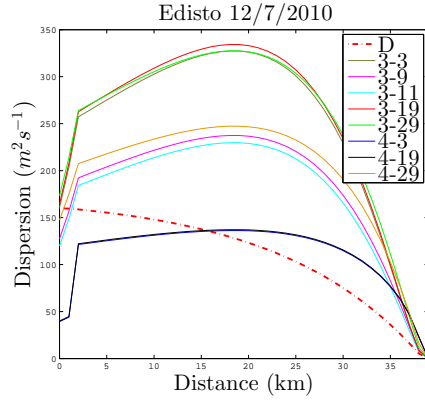




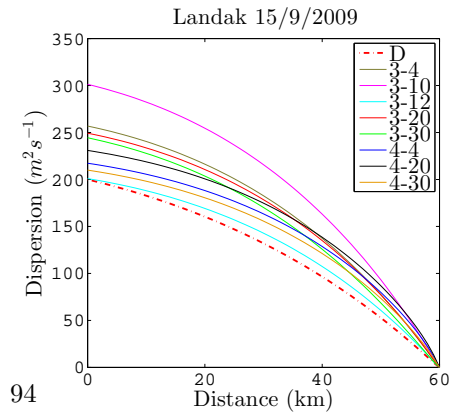
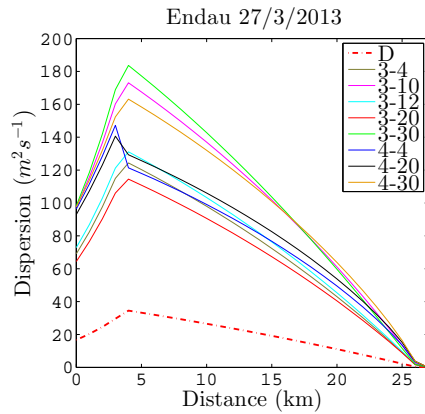
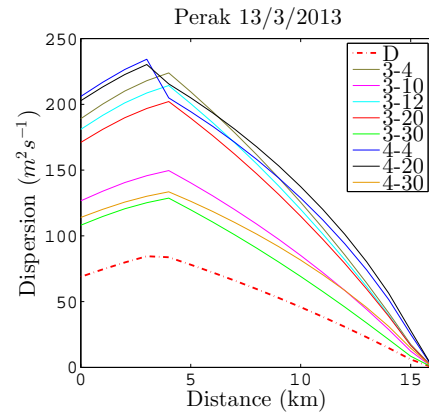
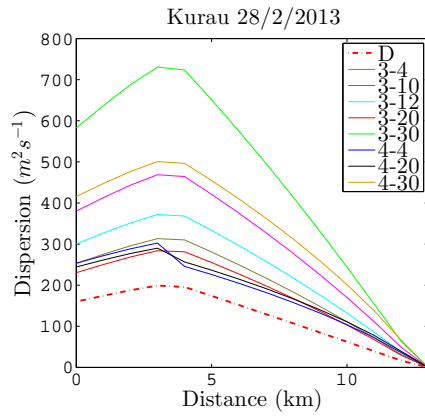
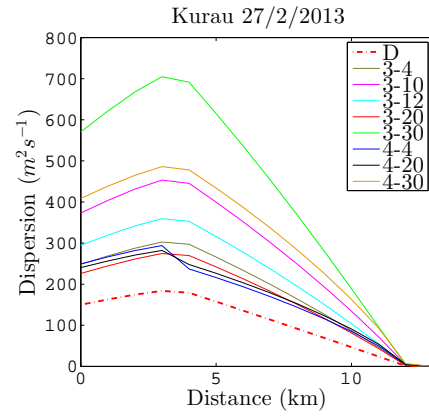
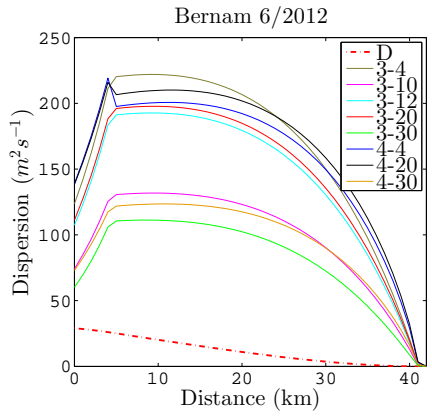
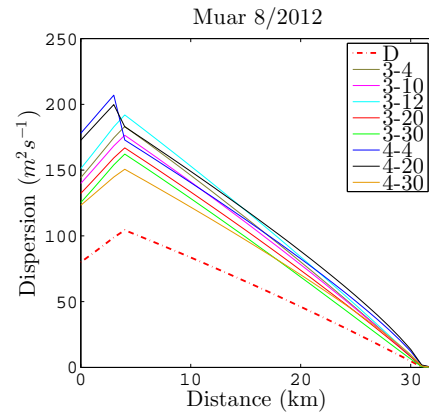
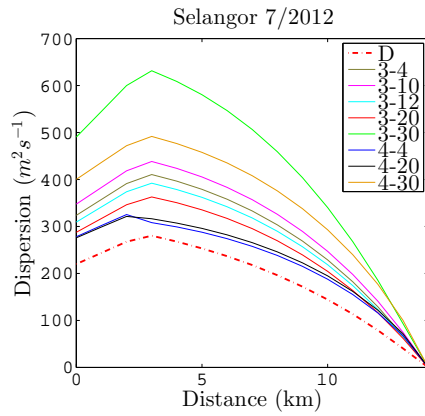


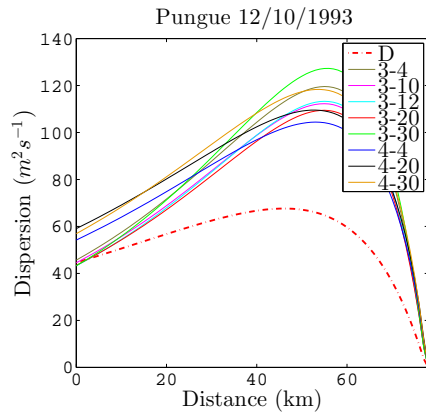
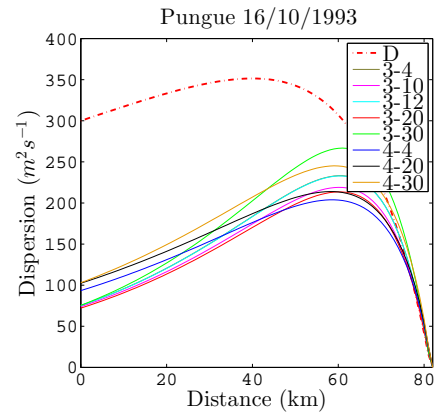
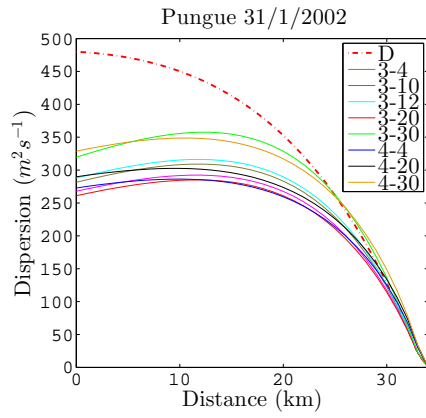
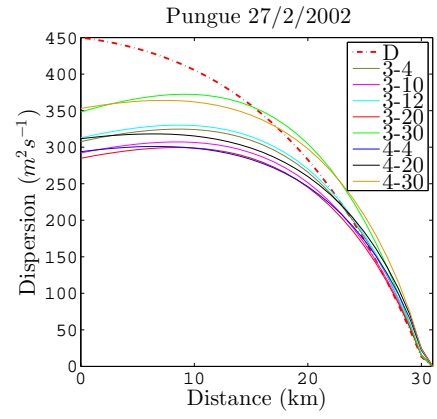
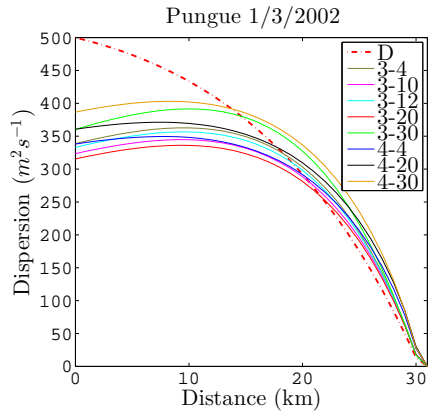
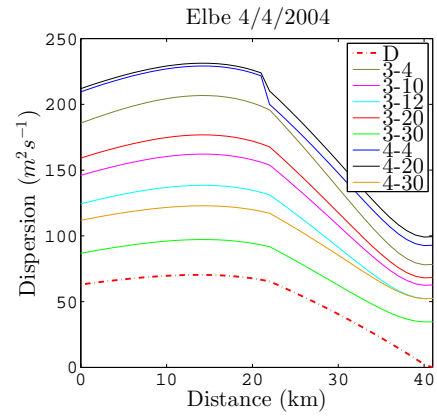
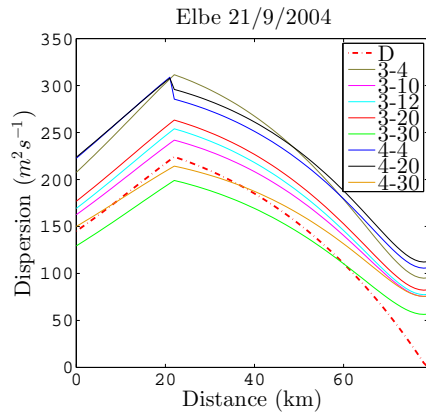




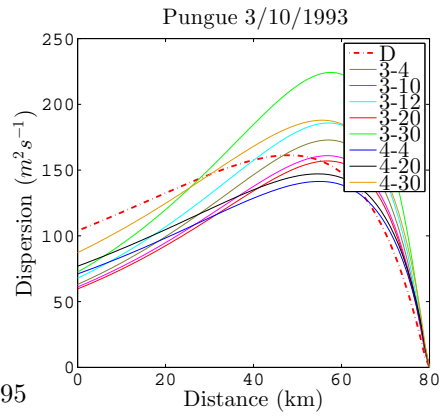


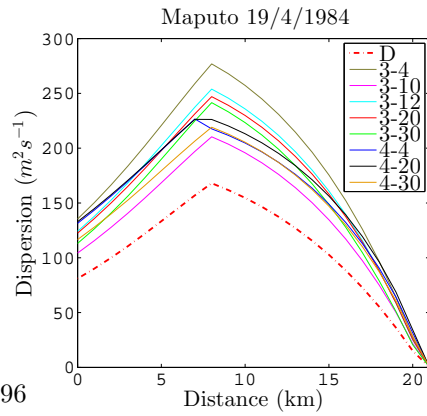
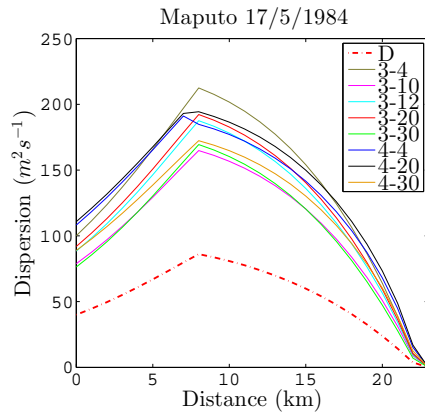
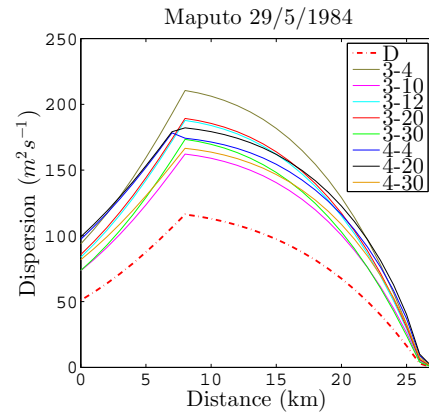
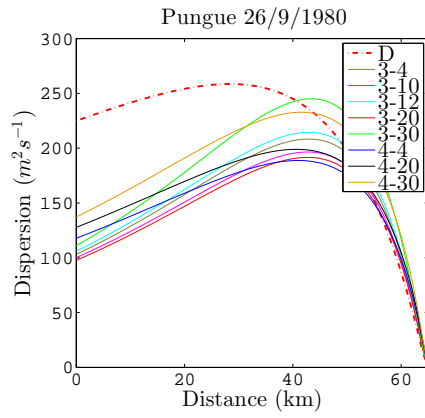
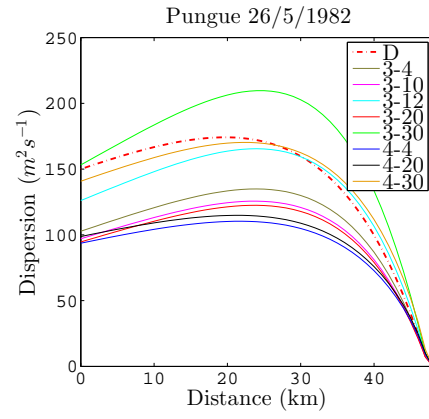
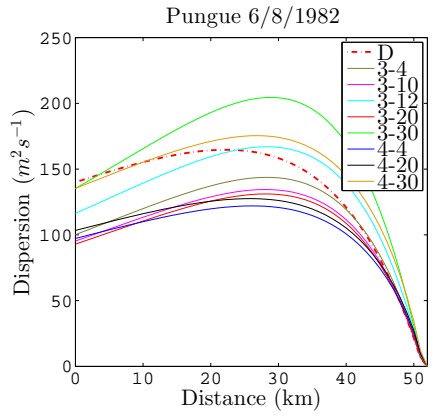
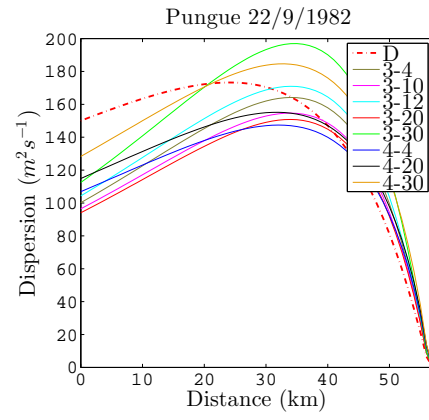
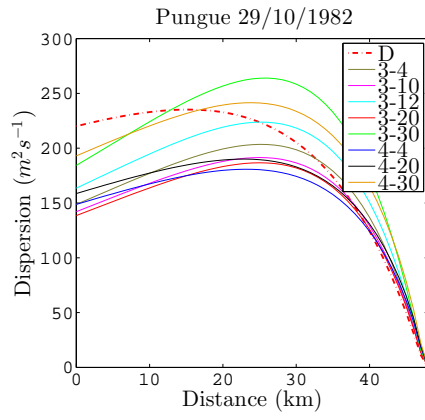
D Graphs Local Validity Kuijper

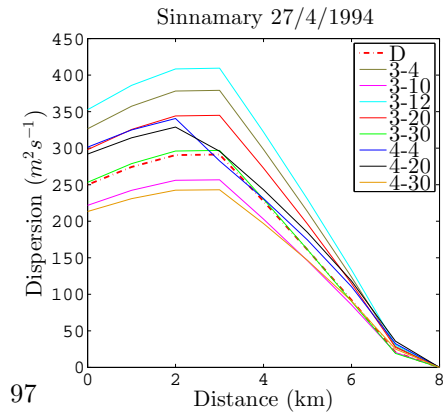
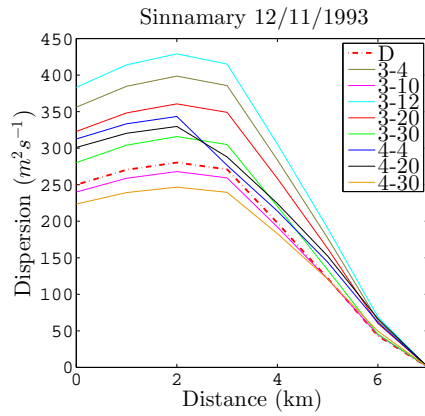
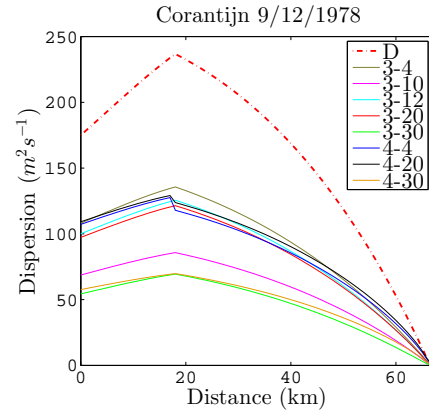
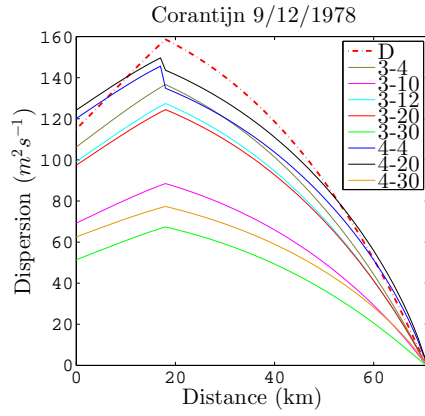
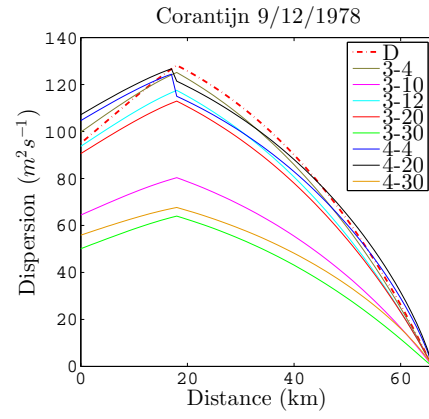
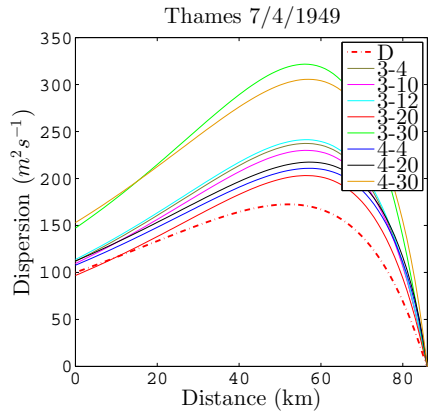
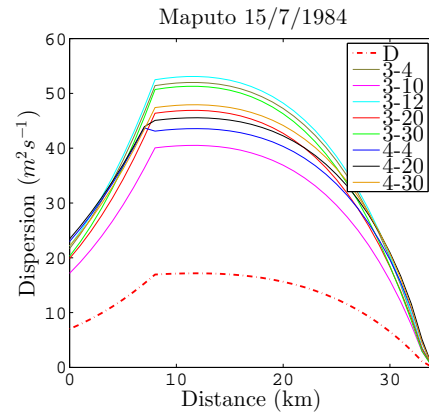
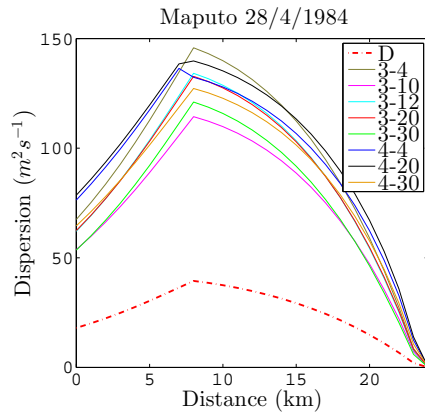


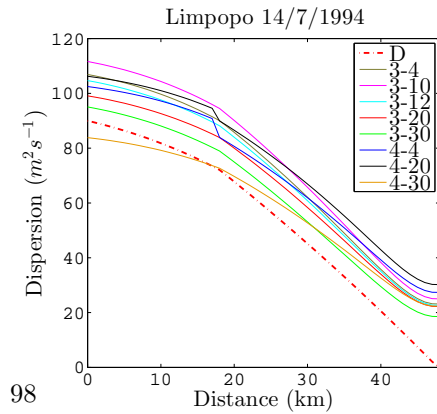
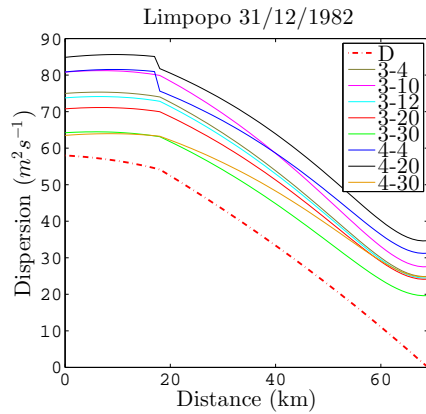
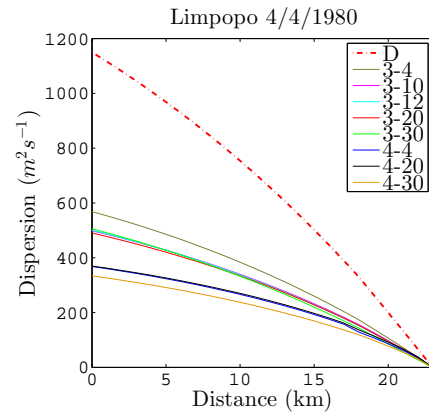
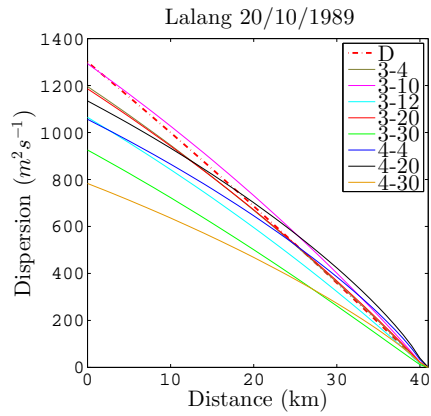
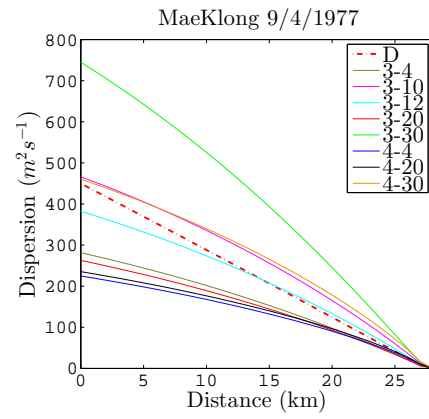
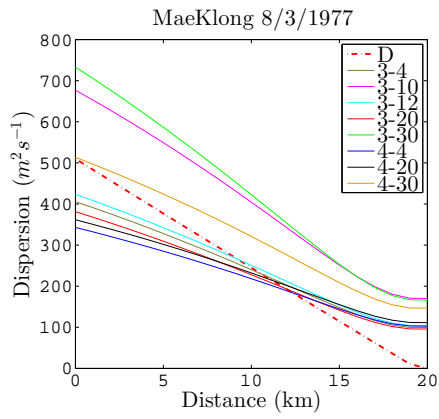
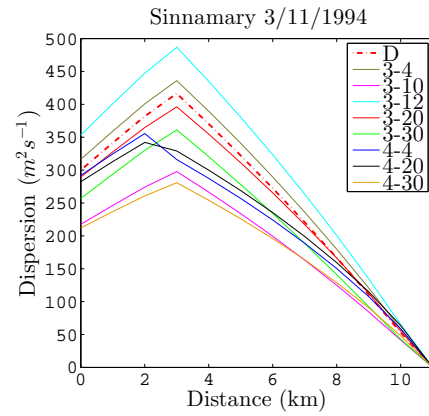
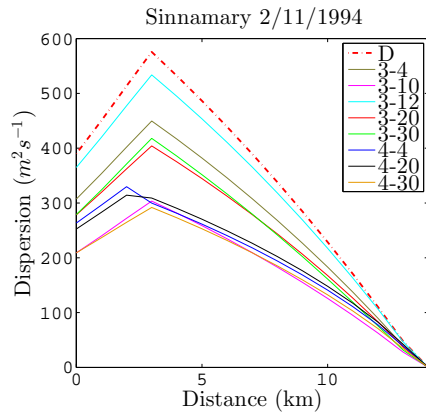


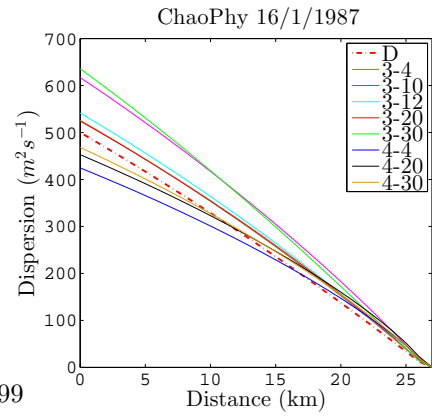
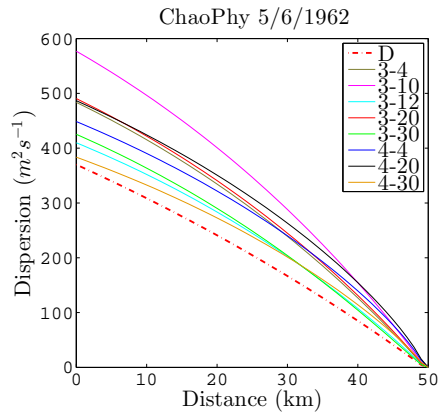
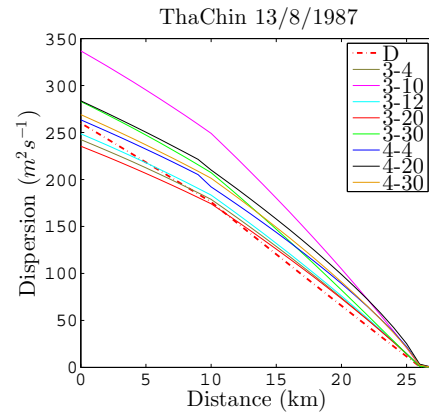
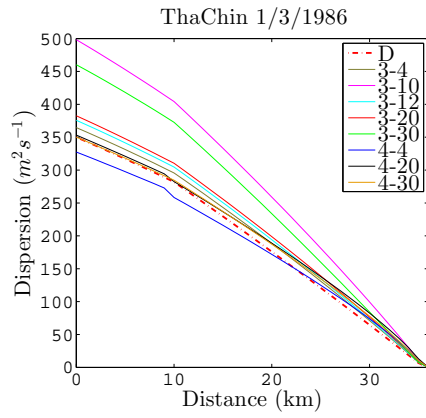
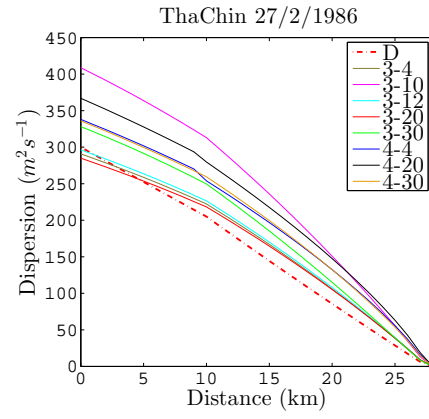
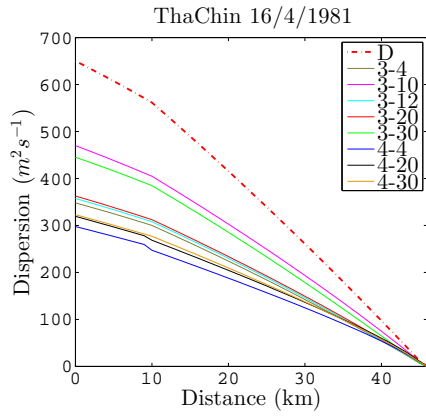
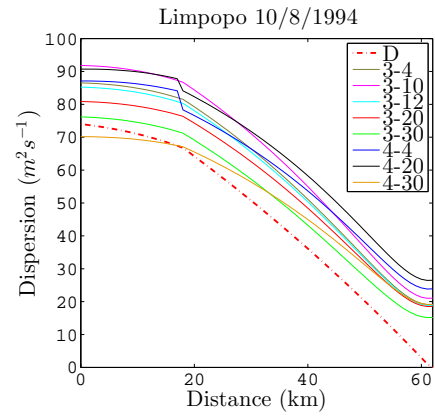
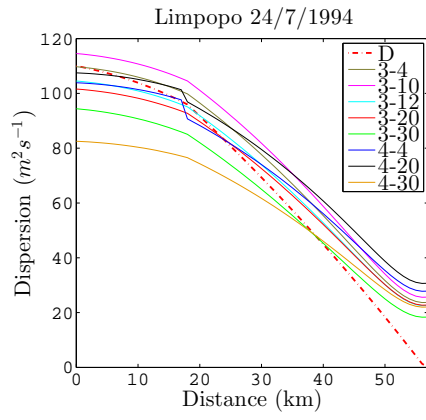
95

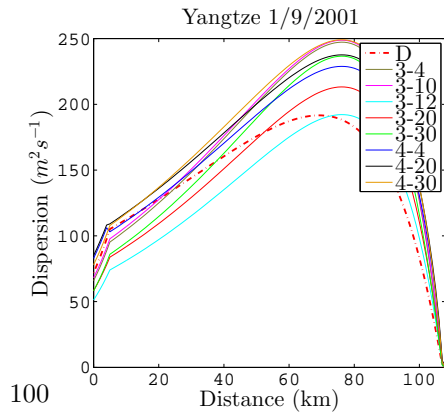
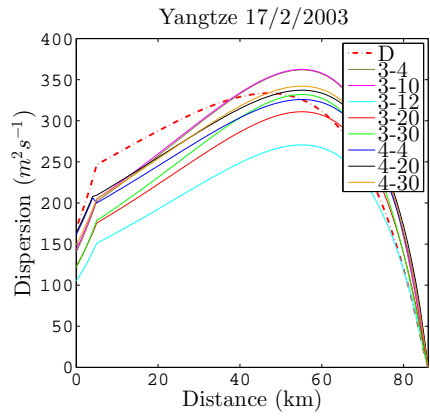
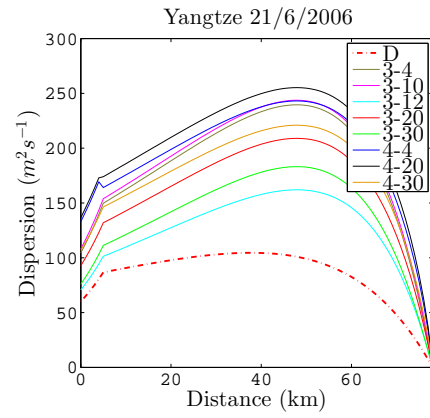
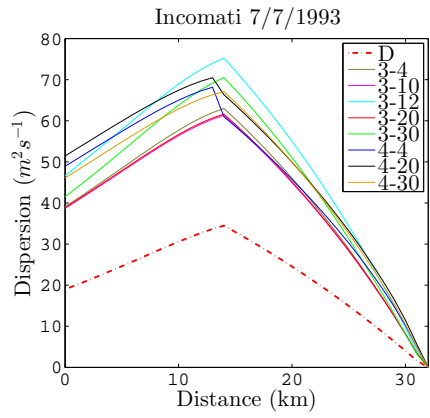
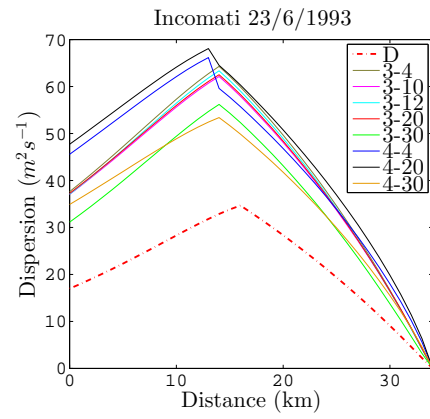
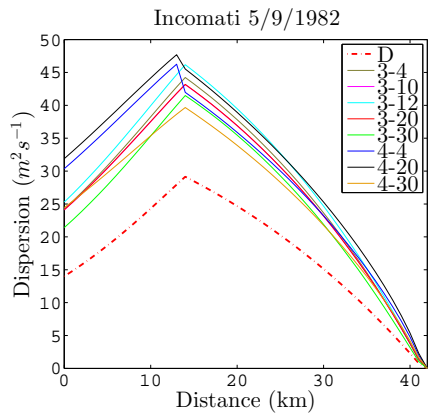
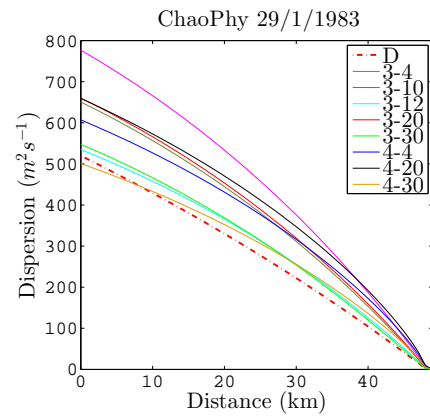
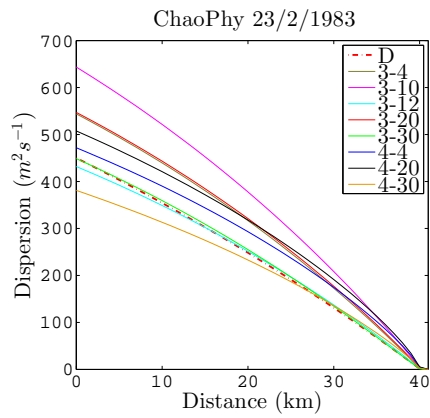


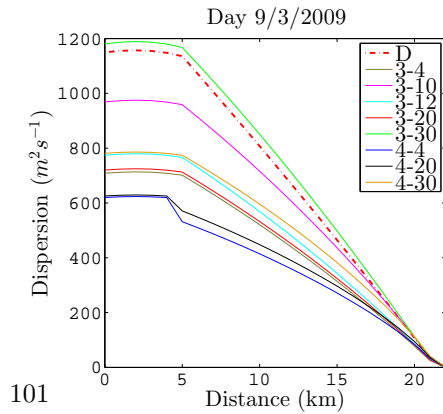
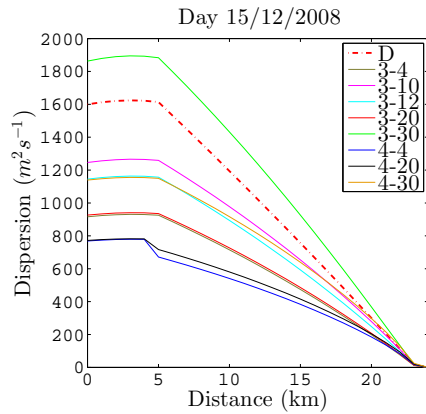
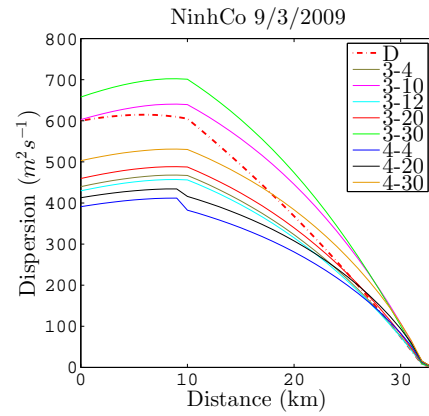
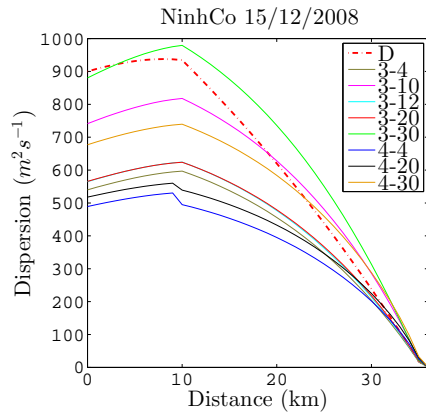
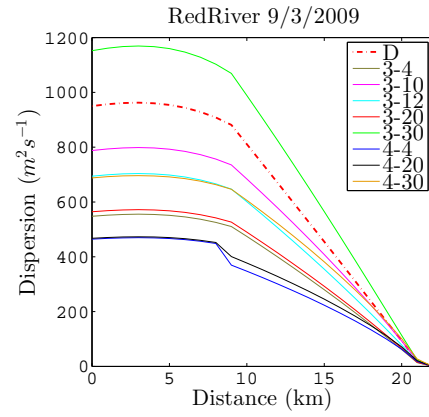
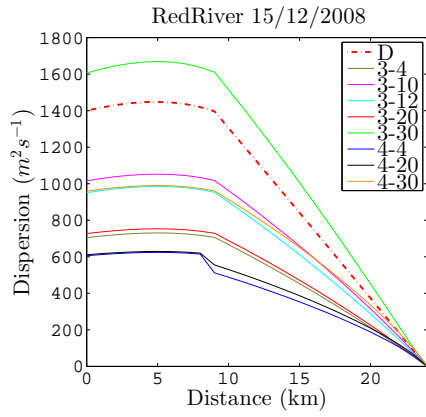
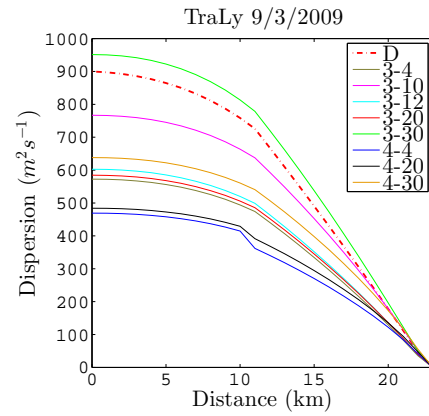
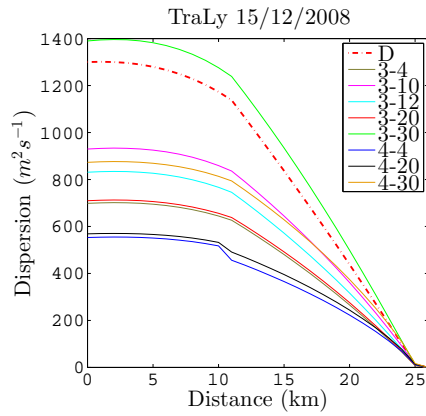


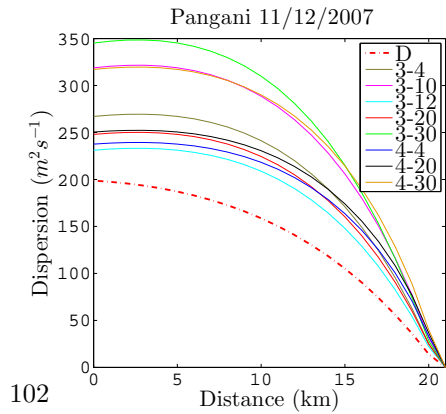
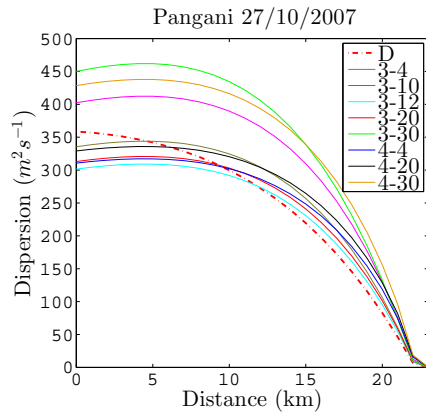
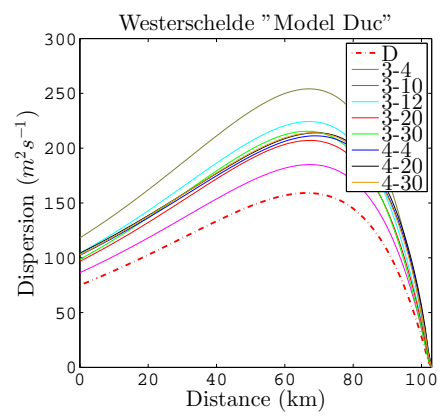
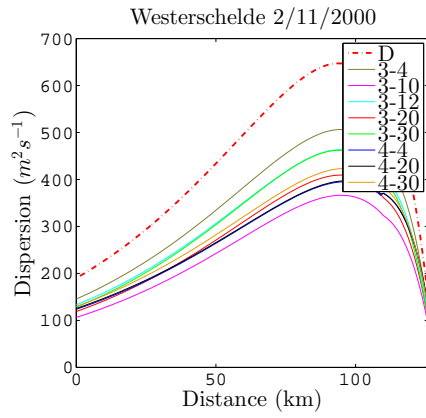
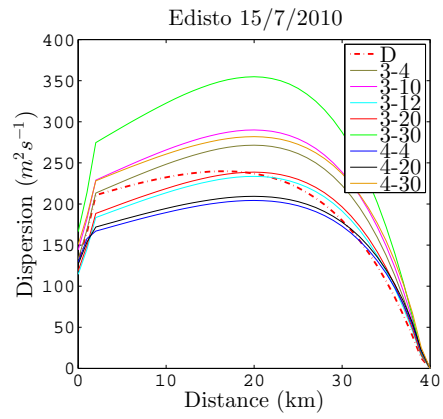
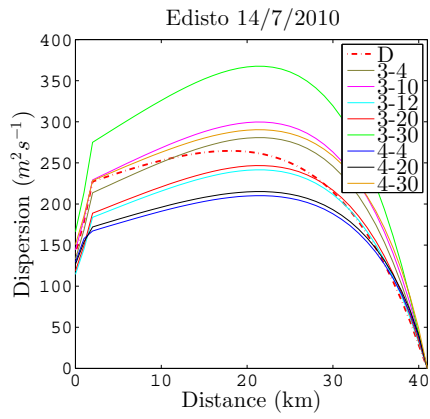
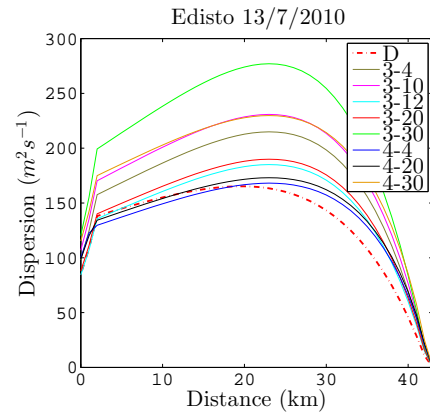
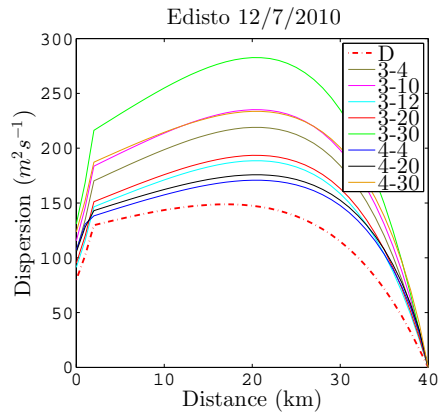




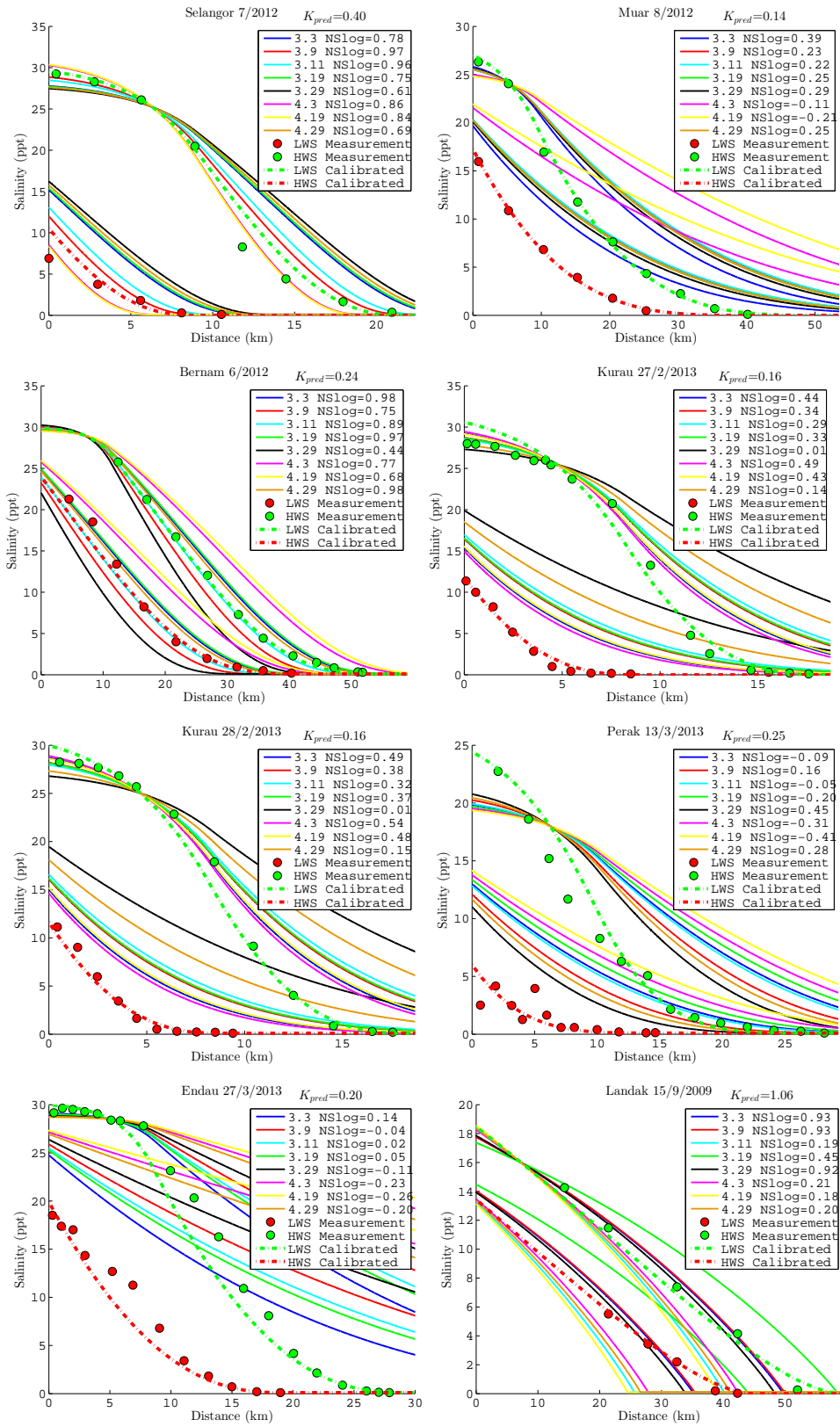


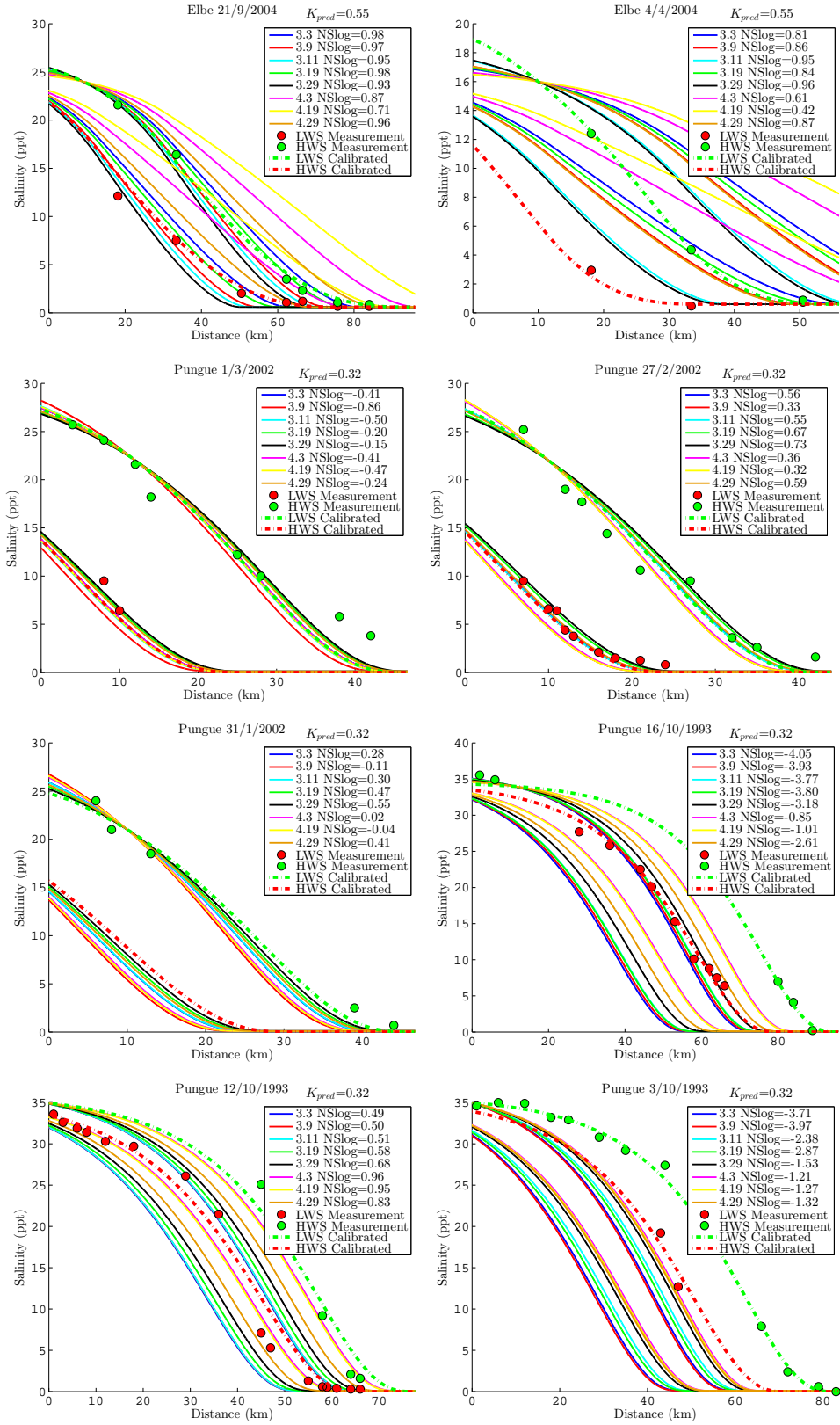


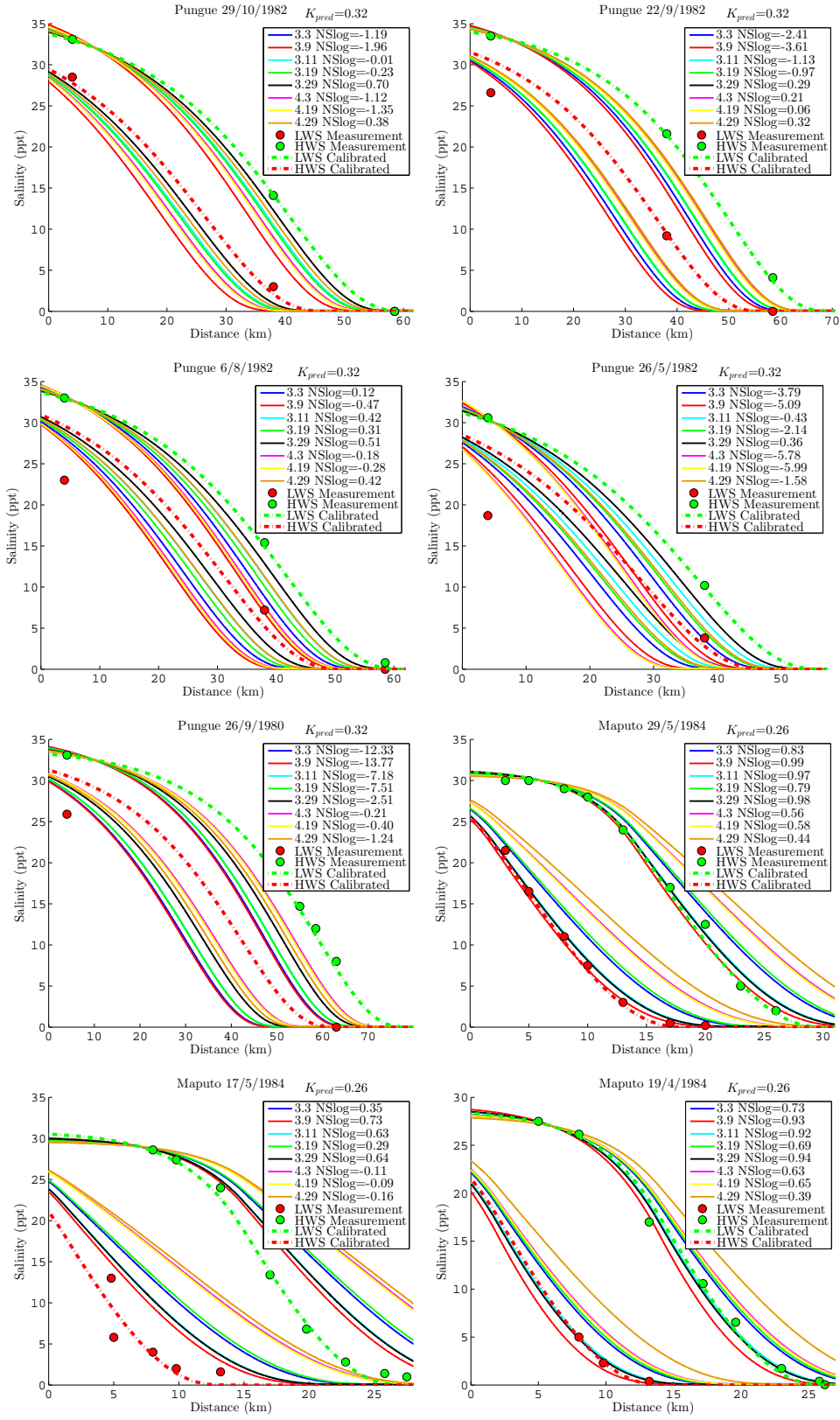


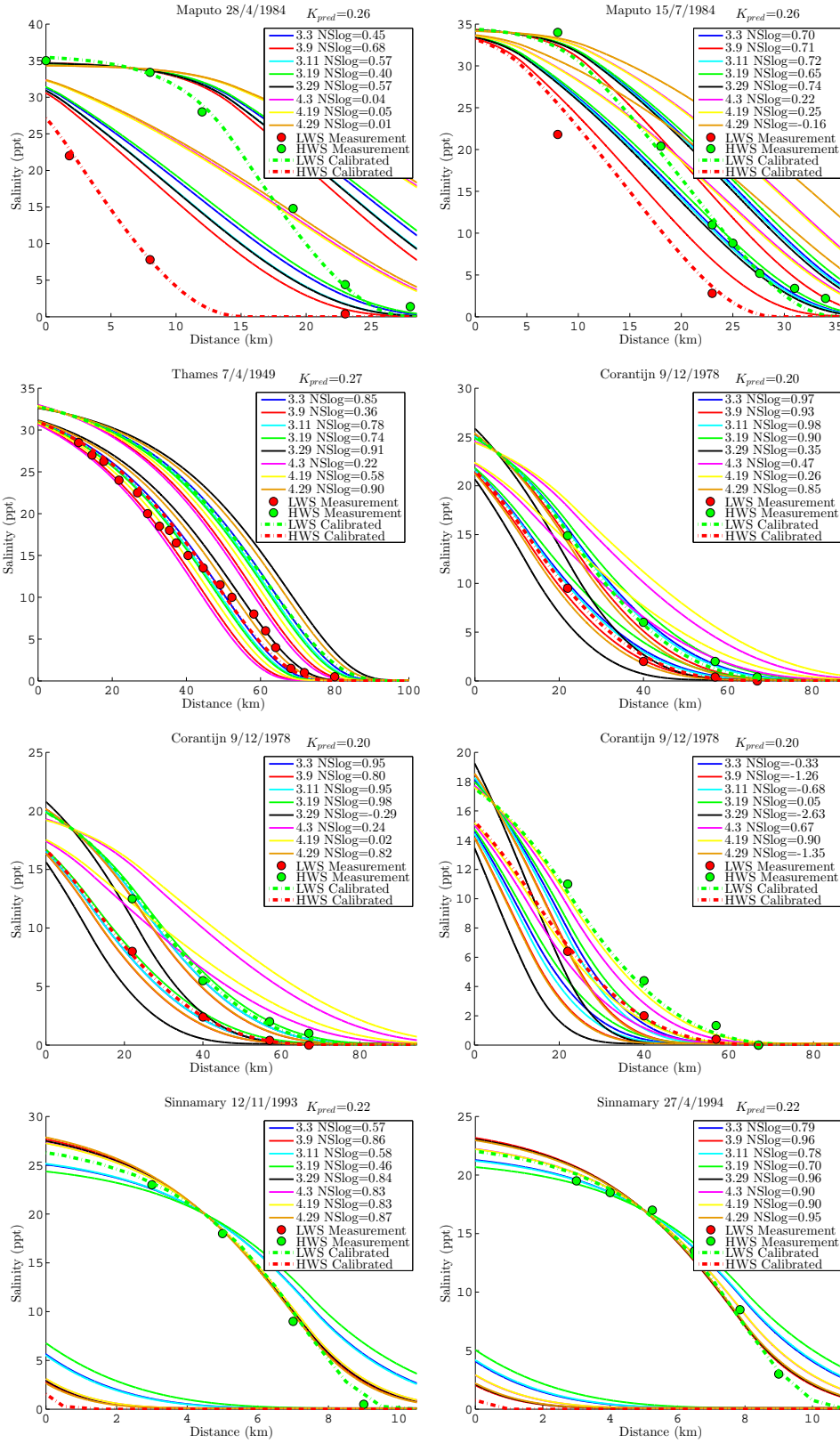


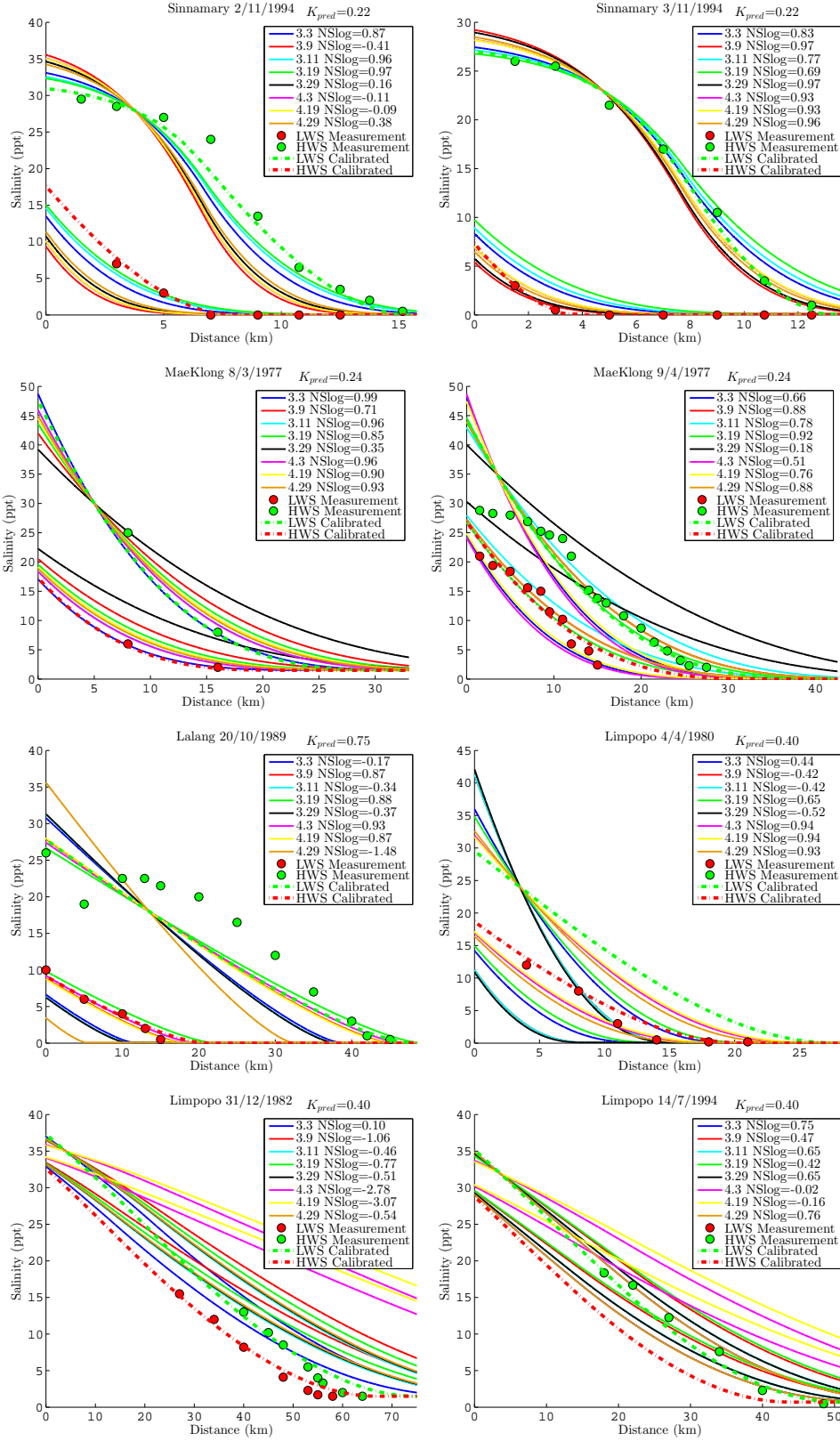
E Graphs Predictive Mode Savenije

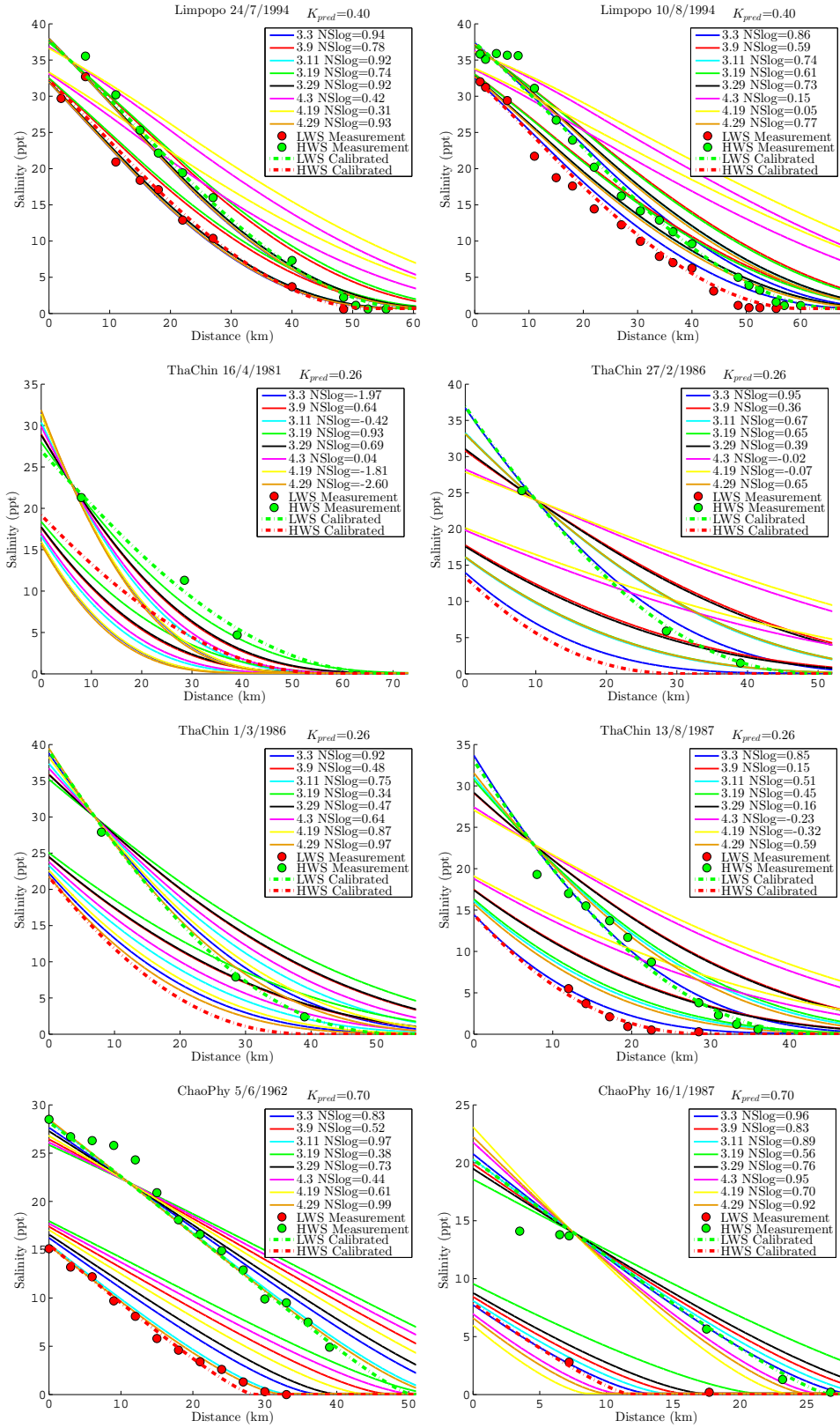


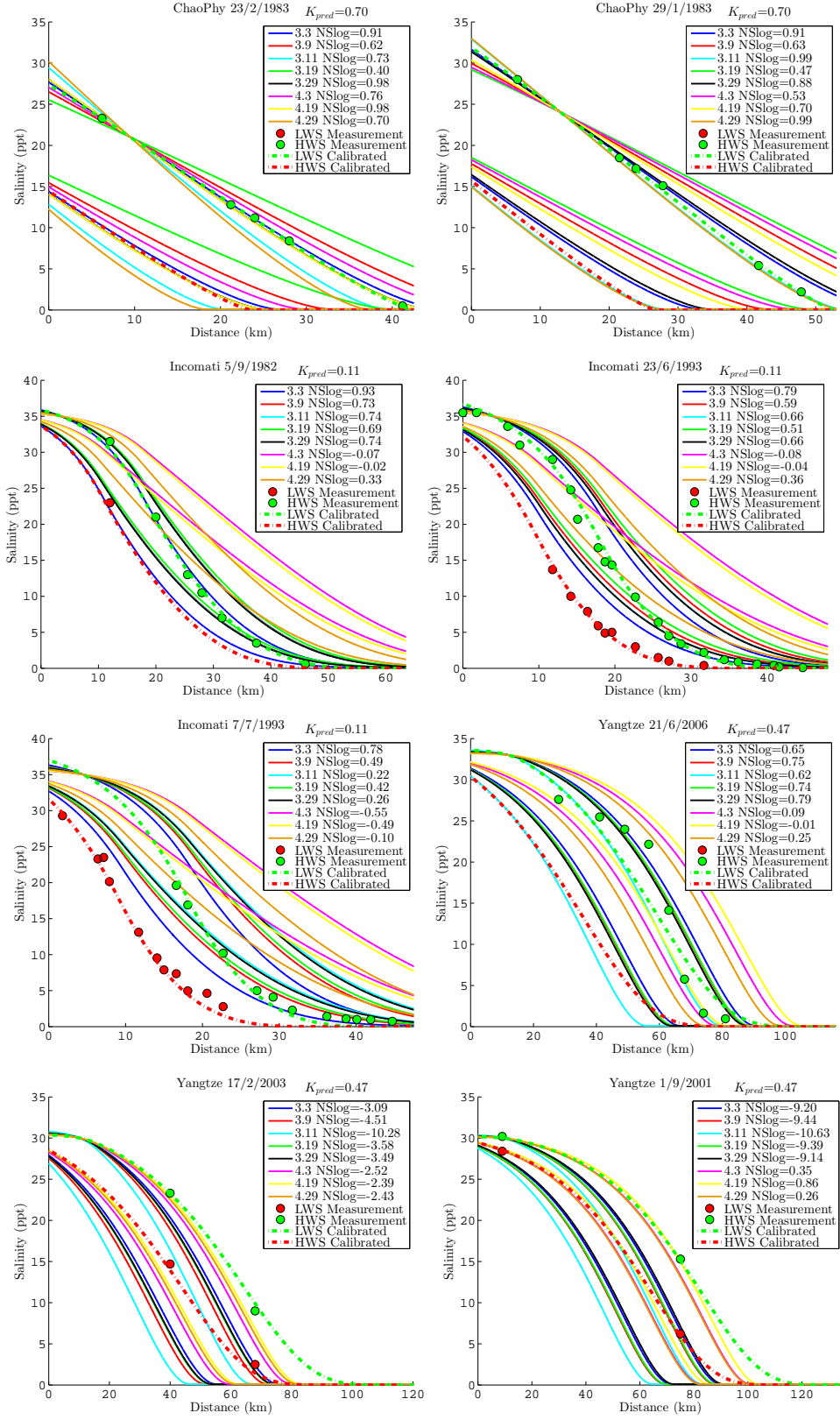


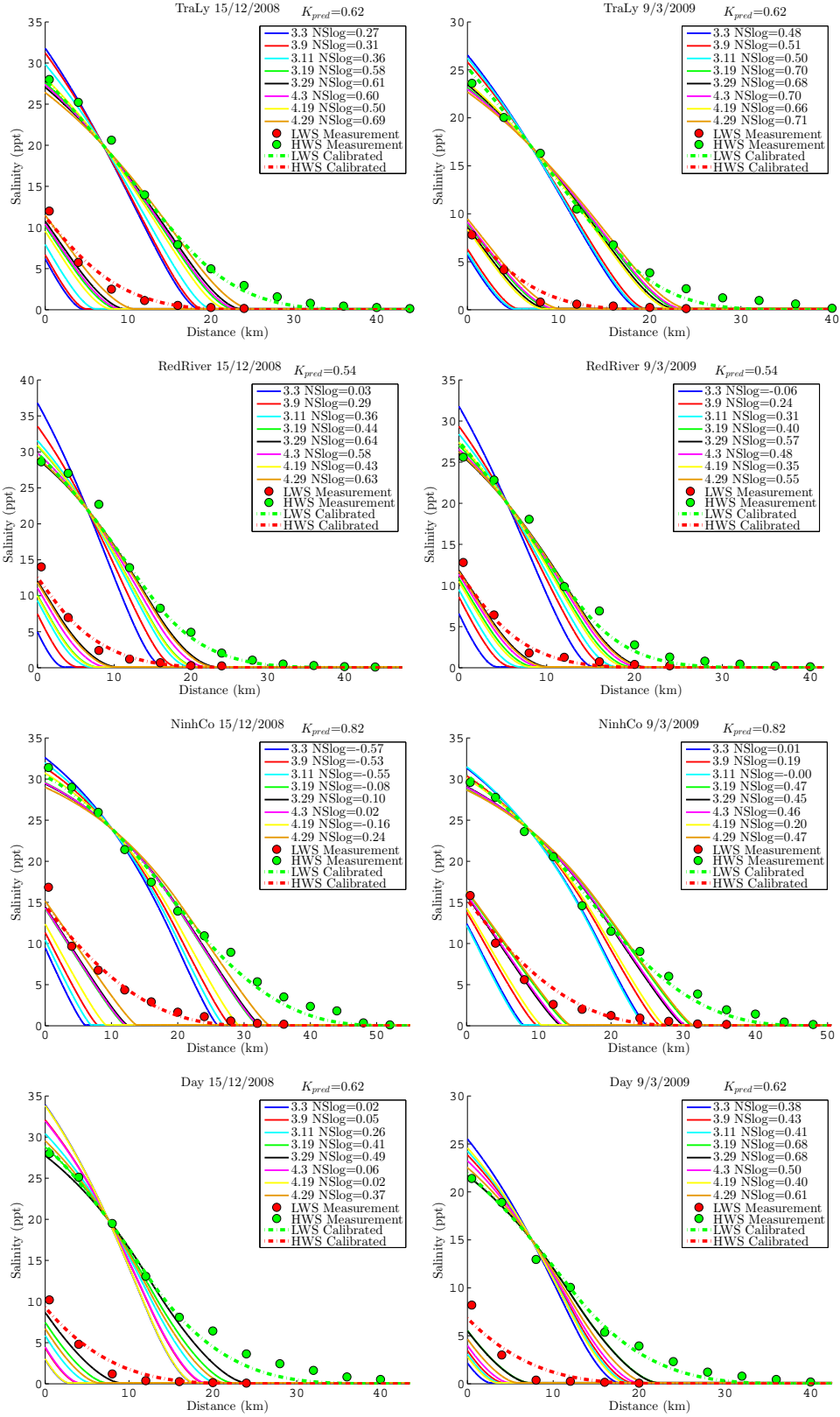


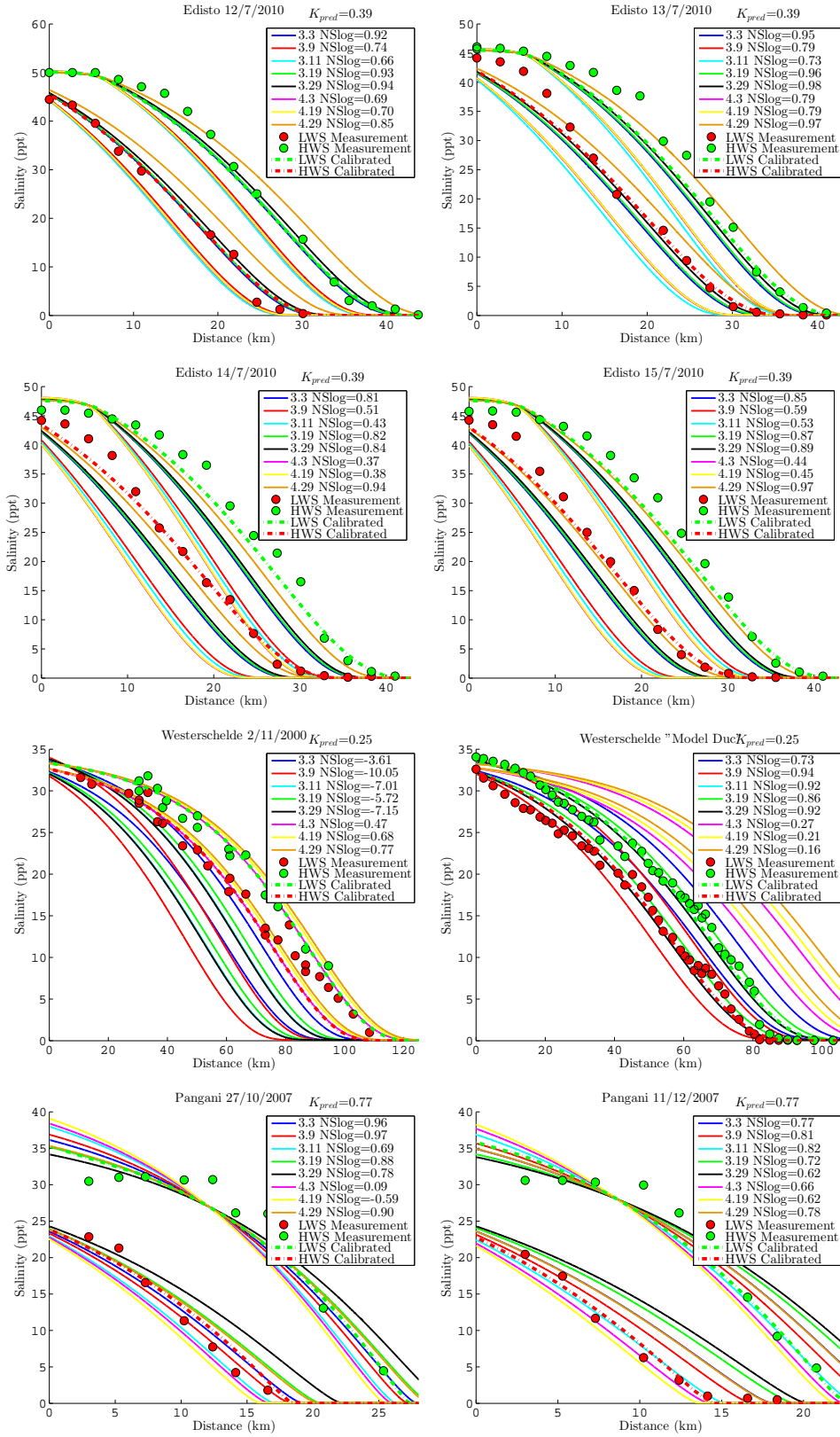












F Graphs Predictive Mode Kuijper

