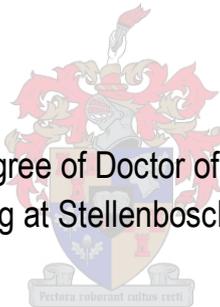


Modelling breakthrough curves and investigating the impact of models and numerical properties on parameter estimation.

By

Davy Danny Silawwe

Dissertation presented for the degree of Doctor of Civil Engineering in the Faculty of Engineering at Stellenbosch University



Supervisor: Dr I. C. Brink

April 2019

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained herein is my own original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining and qualification.

This dissertation includes 3 unpublished publications. The development and writing of the papers were the principle responsibility of myself and, for each of the cases where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contribution of co-authors.

Date: April 2019

Copyright © 2019 Stellenbosch University

All rights reserved

Abstract

The research investigated and applied several Eulerian numerical methods of the advection-dispersion model (AD-Model) for the analysis of concentration-time curves, also known as breakthrough curves (BTCs), to develop empirical models for predicting stream longitudinal dispersion coefficients. Typically, measured BTCs are analysed to estimate solute transport parameters which are then used to develop empirical equations by correlating optimised longitudinal dispersion coefficients with the bulk flow and channel properties. The investigation attempted to determine the impact of numerical methods and nondimensional numerical properties on optimised parameters and subsequently on constructed empirical models for predicting longitudinal dispersion coefficients.

Four concerns related to the construction of empirical models for estimating stream longitudinal dispersion coefficient based on estimates by numerical methods were addressed. (a) Dependence of estimated parameter values on the method used. (b) Influence of numerical properties on values of estimated parameters (c) Identification of model structure, and (d) Characterising model performance.

To address the concern (a), six optional numerical methods were assessed using a set of synthetic BTCs simulated for a hypothetical stream reach. This was followed by a selection of three numerical methods for the analysis of observed BTCs to determine parameter values for the development of empirical models. The selected numerical methods are well-known methods, namely, Backward-time/centred space (BTCS), Crank-Nicolson, Implicit QUICK, QUICKEST, MacCormack and third-order upstream-differencing methods. Shortlisted methods were Crank-Nicolson, MacCormack and QUICKEST methods. To address issue (b) parameter values for the development of empirical models were obtained over a range of numerical properties. To address issue (c) dimensional analysis and least-squares regression was used. To address issue (d) a combination of several model performance measures focusing on several features were used for a broad evaluation of models.

The study shows that optimal parameter values of the AD-Model determined by Eulerian numerical methods vary with numerical methods and model resolution, such that there is a possibility of overestimating or underestimating parameter values, especially the dispersion coefficient. Consequently, in this research, the Crank-Nicolson and the MacCormack methods were observed to overpredict the dispersion coefficient with an increase in Peclet number, while the the QUICKEST method was observed to underpredict dispersion coefficients with increase in Peclet number. Consequently, structures of developed empirical models and predictions varied with solution method used and nondimensional numerical parameters under which optimised parameters were determined. Based on performance analysis measures, adequate and comparable empirical models were developed for a range of 0.599 – 12.818 of the Peclet number. However, the quality of concentration predictions using predicted dispersion coefficients requires the use of numerical methods and model resolutions under which empirical models were developed. Therefore, empirical models may be well-founded within their calibrated conditions and channel characteristics.

Opsomming

Hierdie navorsing was gefokus op die ondersoek van verskeie Euleriese numeriese metodes wat tipies gebruik word in die adveksiespreidingsmodel (AD-Model). Die metodes was toegepas op die analise van konsentrasie-tydkrommes, ook bekend as deurbrekkrommes (BTCs), om empiriese modelle te ontwikkel vir die voorspelling van longitudinale verspreidingskoëffisiënte. Gemete BTC's word tipies ge-analiseer om opgeloste vervoerparameters te bereken, wat dan verder gebruik word om empiriese vergelykings te ontwikkel deur ge-optimaliseerde longitudinale verspreidingskoëffisiënte met massa vloeï en kanaal eienskappe te korreleer. Daar was gepoog in hierdie ondersoek om die impak van numeriese metodes en nie-dimensionele numeriese eienskappe op ge-optimaliseerde parameters te bepaal. Daar was ook voort gegaan om die verdere impak hiervan op empiriese modelle vir die voorspelling van longitudinale verspreidingskoëffisiënte te bepaal.

Vier kwessies wat verband hou met die konstruksie van empiriese modelle vir die bepaling van stroom longitudinale verspreidingskoëffisiënte, gebaseer op beraamings deur numeriese metodes, was aangespreek: (a) afhanklikheid van beraamde parameterwaardes op die metode wat gebruik word, (b) invloed van numeriese eienskappe op waardes van geskatte parameters, (c) identifikasie van modelstruktuur en (d) karakterisering van modelprestasie.

Om kwessie (a) aan te spreek, was ses opsionele numeriese metodes ge-assesseer met 'n stel sintetiese BTC's wat gesimuleer was vir 'n hipotetiese stroom lengte. Dit was gevolg deur 'n verdere ondersoek van drie gekose numeriese metodes gebruik vir die analise van die waargenome BTCs om parameterwaardes vir die ontwikkeling van empiriese modelle te bepaal. Die ondersoekte numeriese metodes het algemeen bekende metodes ingesluit, naamlik: die "Backward-time/centred space (BTCS)" metode, die "Crank-Nicolson" metode, die "Implicit QUICK" metode, die "QUICKEST" metode, die "MacCormack" metode en derde-orde stroomopwaartse differensieringsmetodes. Finaal geselekteerde metodes was die "Crank-Nicolson", "MacCormack" en "QUICKEST" metodes. Om kwessie (b) aan te spreek, was parameterwaardes vir die ontwikkeling van empiriese modelle oor 'n verskeidenheid numeriese eienskappe verkry. Dimensionele analise en kleinste-kwadrante regressie was gebruik om kwessie (c) aan te spreek. 'n Kombinasie van verskeie modelprestasiemaatreëls wat op verskeie funksies gefokus het was gebruik vir 'n breë evaluering van modelle onder kwessie (d).

Die studie het getoon dat optimale parameterwaardes van die AD-model, wat deur Euleriese numeriese metodes bepaal was, gewissel het volgens numeriese metodes en modelresolusie. Dit beteken dat daar 'n moontlikheid is om parameterwaardes, veral die verspreidingskoëffisiënt, te oorskakel of te onderskat. In hierdie ondersoek was die "Crank-Nicolson" en die "MacCormack" gebaseerde modelle waargeneem om die verspreidingskoëffisiënt te oorskakel met toename in Peclet-nommer, terwyl die "QUICKEST" gebaseerde modelle waargeneem was om verspreidingskoëffisiënte te onderskat met toename in Peclet-nommer. Gevolglik het die strukture van ontwikkelde empiriese modelle en voorspellings gewissel met oplossingsmetode wat gebruik was sowel as met die nie-dimensionele numeriese parameters waarvolgens ge-optimaliseerde parameters bepaal was. Op grond van prestasie analise maatreëls, was voldoende en vergelykbare empiriese modelle ontwikkel vir 'n reeks van 0.599 - 12.818 van die Peclet-nommer. Die kwaliteit van konsentrasie voorspellings wat gedoen word deur gebruik te maak van voorspelde verspreidingskoëffisiënte, vereis egter die gebruik van numeriese metodes en modelresolusies waaronder die empiriese modelle ontwikkel is. Dus mag empiriese modelle goed gegrond wees binne hul gekalibreerde toestande en kanaal eienskappe.

CONTENTS

Declaration	i
Abstract	ii
Opsomming	iii
List of figures	ix
List of tables	xii
Acknowledgements	xiv
Abbreviations	xv
Notation	xvi
1 INTRODUCTION	1
1.1 Introduction	1
1.2 Problem statement	4
1.3 Thesis statement	6
1.4 Aims and objectives of the research	6
1.5 Delineations and Limitations	7
1.5.1 Delineations	7
1.5.2 Limitations	9
1.6 Assumptions	10
1.7 The significance of the study	11
1.8 Dissertation Layout	12
2 LITERATURE REVIEW	14

2.1	Introduction	14
2.2	Physical processes of solute transport in streams	17
2.3	The one-dimensional advection dispersion model	20
2.4	Analytical solutions of the one-dimensional advection dispersion model	21
2.5	Numerical solutions of the advection dispersion model	24
2.5.1	Introduction	24
2.5.2	Grid discretisation approaches	25
2.5.3	Discretisation of the advection dispersion model	27
2.5.4	Non-dimensional form of the advection dispersion model and numerical properties	29
2.5.5	Eulerian Solutions of the advection dispersion model	31
2.6	Estimating parameters of the advection dispersion model	32
2.6.1	Global search parameter estimation	34
2.6.2	Local-search parameter estimation	35
2.6.3	Parameter estimation by gradient-based methods	36
2.6.4	Routing procedures	37
2.6.5	Concentration prediction by numerical methods	41
2.7	Empirical equations for estimating the dispersion coefficient.	43
2.7.1	Introduction	43
2.7.2	Previous works	43
2.7.3	Model construction process	49
2.7.4	Statistical analysis of observational data	49
2.7.5	Empirical model identification	50
2.7.6	Model Performance analysis	53
2.7.7	Model confirmation	54
2.8	Summary	55
3	APPRAISAL OF SELECTED NUMERICAL METHODS	57

3.1	Introduction	57
3.2	Selected numerical methods	58
3.2.1	The Backward-Time/Centred-Space (BTCS) method	58
3.2.2	The Crank-Nicolson (CN) method	59
3.2.3	The Implicit QUICK method	60
3.2.4	The QUICKEST Method	61
3.2.5	The MacCormack Method	62
3.2.6	Third-order upstream-differencing scheme	63
3.3	Generating synthetic data	64
3.4	Numerical Concentration prediction by explicit schemes	64
3.5	Concentration prediction by implicit schemes	65
3.6	Identification of parameters	66
3.6.1	Computational method	66
3.7	Performance assessment of the methods	69
3.8	Results and discussion	70
3.9	Conclusion	77
4	DATA AND ANALYSIS OF OBSERVED BREAKTHROUGH CURVES	79
4.1	Introduction	79
4.2	Data analysis	79
4.3	Estimating stream flow rates	83
4.4	Estimating stream parameter values.	84
4.4.1	Introduction	84
4.4.2	Testing the routing procedure and computation of initial parameter values.	86
4.4.3	Numerical analysis of observed breakthrough curves.	91

4.4.4	Results and discussion	92
4.4.5	Conclusion	102
5	EMPIRICAL MODEL DEVELOPMENT AND CALIBRATION	103
5.1	Introduction	103
5.2	Methods	104
5.2.1	Testing normality of computed dispersion coefficients	104
5.2.2	Results	105
5.2.3	Development of new models	107
5.2.4	Evaluating model parameters and model calibration	112
5.2.5	Results and discussion: regression parameters	115
5.2.6	Results and discussion: calibration of models	117
5.3	Conclusion	121
6	PERFORMANCE ANALYSIS OF EMPIRICAL MODELS	123
6.1	Introduction	123
6.2	Methods	124
6.2.1	Graphical method	124
6.2.2	Descriptive statistics	125
6.2.3	Error statistics	125
6.2.4	Residual error analysis	127
6.3	Results and discussion	129
6.4	Conclusion	135
7	CONFIRMATION OF NEW MODELS AND COMPARISON WITH EXISTING MODELS	137
7.1	Introduction	137

7.2	The method: confirmation of empirical models	138
7.2.1	Results and discussion: confirmation of empirical models	139
7.3	Comparison of new models with existing methods	151
7.3.1	Results and discussion: comparison of models	152
7.4	Conclusion	154
8	CONCLUSION AND RECOMMENDATIONS	156
8.1	Findings	156
8.2	Conclusion	159
8.3	Summary of contributions	160
8.4	Recommendations	161
	BIBLIOGRAPHY	163
	APPENDICES	168

List of figures

Figure 2-1: A structured computational grid for the finite volume method.....	26
Figure 2-2: A structured computational grid for the finite difference method	27
Figure 2-3: Fitting a simulated profile to a measured profile by optimising the velocity and dispersion coefficient (Wallis et al., 2013).....	32
Figure 3-1: General structure of the spreadsheet for the numerical solution domain.....	68
Figure 3-2: A flowchart for concentration prediction and parameter estimation, adopted from Chapra (2008).	69
Figure 3-3: Plot of computed dispersion coefficients versus the Peclet number	73
Figure 3-4: Plot of computed velocity versus the Peclet number.....	73
Figure 3-5: Plot of the residual sum of squares versus Peclet number	74
Figure 3-6: Plot of relative error of optimised dispersion coefficient versus the Peclet number.....	76
Figure 3-7: Plot of relative error of optimised velocity versus Peclet number.....	76
Figure 3-8: Simulation results of optional methods using space step of 20 m.....	77
Figure 4-1: Study site (Burke, 2002).....	80
Figure 4-2: Turner Designs 10 fluorometer (Heron, 2015).....	81
Figure 4-3: Fluorometer reading versus Time (experiment 2)	82
Figure 4-4: Concentration versus Time (experiment 2).....	83
Figure 4-5: Simulation result of the routing procedure for a time step of 40 seconds.	89
Figure 4-6: Simulation result of the routing procedure for a time step of 60 seconds.	89
Figure 4-7: Examples of simulation results during estimation of initial parameter values.	91
Figure 4-8: A Flowchart for analysis of observed BTCs, adapted from Chapra (2008).....	92
Figure 4-9: Computed dispersion coefficient versus Peclet number (experiment 10)	95
Figure 4-10: Computed dispersion coefficient versus Peclet number (experiment 11).....	95
Figure 4-11: Computed dispersion coefficient versus Peclet number (experiment 12).....	96
Figure 4-12: Computed velocity versus Peclet number for experiment 10	98
Figure 4-13: Optimised velocity versus Peclet number for experiment 11	98
Figure 4-14: Optimised velocity versus Peclet number for experiment 12	98
Figure 4-15: Residual sum of squares versus Peclet number for experiment 10	99
Figure 4-16: Residual sum of squares versus Peclet number for experiment 11	99
Figure 4-17: Residual sum of squares versus Peclet number for experiment 12	99

Figure 4-18: Fitting simulated and observed concentration-time profiles, space step = 12.266 m (experiment 10).....101

Figure 4-19: Fitting simulated and observed concentration-time profiles, space step = 12.266m (experiment 12).....102

Figure 5-1: Q-Q plots and normality test results of computed dispersion coefficients.....107

Figure 5-2: Computed dispersion coefficients versus independent model variables (Crank-Nicolson)110

Figure 5-3: Computed dispersion coefficient versus model independent variables (MacCormack) ..110

Figure 5-4: Computed dispersion coefficients versus independent model variables (QUICKEST)111

Figure 5-5: Computed dispersion coefficient versus the control variables (Singh and Beck routing procedure)111

Figure 5-6: Model coefficients, α , versus space step.....117

Figure 5-7: Plot of predicted dispersion coefficient versus the Peclet number (experiment 10)120

Figure 5-8: Plot of predicted dispersion coefficients versus the Peclet number (experiment 12)120

Figure 5-9: Plot of predicted dispersion coefficients against flow rate, $\Delta x = 2.00$ m for numerical method-based models.121

Figure 5-10: Plot of predicted dispersion coefficients against flow rate, $\Delta x = 18.40$ m for numerical method-based models.121

Figure 6-1: Predicted dispersion coefficients versus computed dispersion coefficients; $\Delta x = 18.40$ m (CN).....130

Figure 6-2: Predicted versus computed dispersion coefficients; $\Delta x = 18.40$ m (MacCormack)130

Figure 6-3: Predicted dispersion coefficient versus computed dispersion coefficient; $\Delta x = 18.40$ m (QUICKEST).....131

Figure 6-4: Predicted dispersion coefficient versus computed dispersion coefficient (routing method).....131

Figure 6-5: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; CN-based model.....134

Figure 6-6: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; MacCormack based model.....134

Figure 6-7: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; QUICKEST based model.....135

Figure 6-8: Normalised residuals versus predicted dispersion coefficient; routing method-based model.....	135
Figure 7-1: Predicted dispersion coefficient versus the Peclet number; experiment 4.....	140
Figure 7-2: Predicted dispersion coefficient versus Peclet number; experiment 9.....	140
Figure 7-3: Predicted dispersion coefficients versus Peclet number; experiment 13.....	140
Figure 7-4: Plot of the residual sum of squares versus Peclet number; experiment 4.....	141
Figure 7-5: Plot of the residual sum of squares versus Peclet number; experiment 9.....	141
Figure 7-6: Residual sum of squares (RSS) versus Peclet number; experiment 13.....	142
Figure 7-7: Coefficient of determination versus Peclet number; experiment 4.....	142
Figure 7-8: Coefficient of determination versus Peclet number; experiment 9.....	143
Figure 7-9: Coefficient of determination versus Peclet number; experiment 13.....	143
Figure 7-10: Fitting concentration profiles, dispersion coefficient predicted by CN-based model; experiment 13.....	144
Figure 7-11: Fitting concentration profiles, dispersion coefficient predicted by MacCormack-based model; experiment 13.....	145
Figure 7-12: Fitting concentration profiles, dispersion coefficient predicted by QUICKEST-based model; experiment 13.....	145
Figure 7-13: Fitting concentration profiles by Crank-Nicolson method, dispersion coefficient predicted by routing method-based model; experiment 13.....	146
Figure 7-14: Fitting concentration profiles by MacCormack method, dispersion coefficient predicted by routing method-based model; experiment 13.....	147
Figure 7-15: Fitting concentration profiles by QUICKEST method, dispersion coefficient predicted by routing method-based model, experiment 13.....	147
Figure 7-16: Residual sum of squares (RSS) versus Peclet number; experiment 13.....	148
Figure 7-17: Coefficient of determination versus Peclet number; experiment 13.....	150
Figure 7-18: Analogy of RSS obtained by numerical methods, predicted dispersion coefficient given by the routing-based model; experiment 13.....	150
Figure 7-19: Analogy of R^2 achieved by numerical methods, predicted dispersion coefficient given by the routing-based model; experiment 13.....	151
Figure 7-20: Plot of predicted dispersion coefficients obtained by several empirical methods versus flow rate.....	154

List of tables

Table 3-1: Dimensionless numerical properties and optimised dispersion coefficients (m^2/s) (UNS means unstable solution)	71
Table 3-2: Dimensionless numerical properties and optimised velocity (m/s).	71
Table 3-3: Dimensionless numerical properties and residual sum of squares ($\mu g^2/l^2$)	74
Table 3-4: Dimensionless numerical properties and relative error of optimized dispersion coefficient	75
Table 3-5: Dimensionless numerical properties and relative error of optimised velocity	75
Table 4-1: Physical characteristics of sub-reaches of the study site (Ani et al., 2009).....	81
Table 4-2: Summary of experiments (Heron, 2015) and flow rates.	84
Table 4-3: Analytical and optimised parameter values, and relative error statistics.....	88
Table 4-4: Optimized longitudinal dispersion coefficients and flow rates (Crank-Nicolson)	93
Table 4-5: Optimized longitudinal dispersion coefficients and flow rates (MacCormack method).....	94
Table 4-6: Optimized longitudinal dispersion coefficients and flow rates (QUICKEST method; UNS means unstable solution)	94
Table 4-7: Flow rates and optimised velocity (Crank-Nicolson method).....	96
Table 4-8: Flow rates and optimised velocity (MacCormack method).....	97
Table 4-9: Flow rates and optimised velocity (QUICKEST method; UNS means unstable solution)	97
Table 5-1: Peclet number and model parameters for the Crank-Nicolson based model	116
Table 5-2: Peclet number and model parameters for the MacCormack-based model	116
Table 5-3 Peclet numbers and model parameters for the QUICKEST-based model.....	117
Table 5-4: Regression parameters for the routing-based model	117
Table 5-5: Flow rates, space steps and predicted dispersion coefficients (CN-based model).....	118
Table 5-6: Flow rates, space steps and predicted dispersion coefficients (MacCormack-based model)	118
Table 5-7: Flow rates, space steps and predicted dispersion coefficients (QUICKEST-based model)	119
Table 5-8: Flow rates and predicted dispersion coefficients (routing method-based model).....	119
Table 6-1: Space step, minimum and maximum optimised and predicted dispersion coefficients (m^2/s) and performance metrics; CN-based model.....	132
Table 6-2: Space step, minimum and maximum optimised and predicted dispersion coefficients	

(m²/s) and performance metrics; MacCormack-based model133

Table 6-3: Space step, minimum and maximum optimised and predicted dispersion coefficients

(m²/s) and performance metrics; QUICKEST-based empirical model.....133

Table 6-4: Minimum and maximum optimised and predicted dispersion coefficients (m²/s) and performance metrics for routing method-based model.133

Table 6-5: Average values of performance metrics for numerical method-based models and for routing method-based model.134

Table 7-1: Parameters of new models at space step of 18.40 m.....153

Table 7-2: Flow rates and dispersion coefficients estimated by existing empirical models153

Acknowledgements

My sincere gratitude goes to the following people and organisations for the help provided during this research:

Supervisor

Dr I C Brink, Department of Civil Engineering, University of Stellenbosch.

Consultation

Dr S G Wallis, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University.

Professor D G Nel, Centre for Statistical Consultation, University of Stellenbosch.

Data and information

Dr S G Wallis, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University.

Funding

The Copperbelt University, Kitwe, Zambia

Software

Analyse-it statistical analysis and data visualisation software for Microsoft Excel, Version 5.10

(Analyse-it, Ltd)

Abbreviations

AD-Model	Advection dispersion model
BTC	Breakthrough curves
BTCS	Back-time/centred-space
CDS	Centred differencing scheme
CN	Crank-Nicolson
CV	Control volume
FD	Finite difference
FE	Finite element
FV	Finite volume
NSE	Nash-Sutcliffe efficiency
RMSE	Root mean square error
RSS	Residual sum of squares
SSE	Sum of squared errors
TS-Model	Transient storage model
TUDS	Third-order upstream differencing scheme

Notation

Symbol	Description	Dimension
A	Cross-sectional area of flow	L^2
A_p	Time step increment weighting function	T
C	Advection (or Courant) number	
\mathbf{C}	Parameter column vector	
d	Dispersion (or diffusion) number	
D	Longitudinal dispersion coefficient	L^2T^{-1}
e	Transformed error term	
E	Model efficiency	
F	Control variable	
g	Gravitational acceleration	LT^{-2}
h	Vertical space coordinate	L
H	Average flow depth	L
I	Fischer's expression for the triple integration	
I_j	Jain's expression for the triple integral	
I_L	Liu's expression for the triple integral	
l_t	Transverse mixing length	L
L	Characteristic length of velocity deviation	L
M	Mass of solute	M
N	Sample size	
$\mathbf{0}$	Null vector	
P_e	Peclet number	
Q	Flow rate	L^3T^{-1}
R	Hydraulic radius	L
\mathbf{R}	Error vector	
R^2	Coefficient of determination	
r	Correlation coefficient	
S	Channel bed slope	LL^{-1}
s_ω	Standard deviation of observational data	

t	Time	T
U_*	Shear velocity	LT ⁻¹
U'	Local deviation of longitudinal velocity from the cross-sectional mean longitudinal velocity.	LT ⁻¹
u'	Local depth-averaged flow velocity deviation from cross-sectional average longitudinal flow velocity	LT ⁻¹
u	Local deviation of depth-averaged longitudinal flow velocity from cross-sectional mean longitudinal flow velocity.	LT ⁻¹
V	Cross-sectional average longitudinal flow velocity	LT ⁻¹
w	Channel width at water surface	L
W	Shapiro-Wilk statistic	
x	Longitudinal spatial dimension	L
X	Stream reach length	L
X_i	Initial mixing zone	L
y	Transverse space coordinate	L
\mathbf{Y}	Input vector of a random variable	
$\hat{\mathbf{Y}}$	Response vector of a random variable	
Z	Transformed control variable	
Z_k	Numerical integration weighting function	

Greek notation

α	Empirical model coefficient	
$\beta, \gamma, \lambda, \gamma$	Empirical model exponents	
δ	Error term	
∂	Normalised error	
ϵ_t	Cross-sectional average transverse mixing coefficient	L ² T ⁻¹
ϵ_z	Local transverse mixing coefficient	L ² T ⁻¹
ϵ	Dummy integration variable denoting space	L
θ	Expected parameter value	
$\hat{\theta}$	Estimated parameter value	

ϕ	Time weighting factor for solution of the AD-Model	
ρ	Density of water	ML ⁻³
σ^2	Variance of observational data	
$\hat{\sigma}^2$	Variance of residuals	
ϕ	Observed stream solute concentration	ML ⁻³
$\hat{\phi}$	Predicted stream solute concentration	ML ⁻³
τ	Dummy integration variable denoting time	
μ	Pulse of unit concentration	
ν	Kinematic viscosity	L ² T ⁻¹
ξ	Response function for routing procedure	
ζ	Transformed model coefficient	
ζ	Local depth of flow	L
χ	Transformed dependent variable	
ω	Observed random variable	
ϖ	Predicted random variable	
ψ	Numerical convolution integral weighting function	

1 INTRODUCTION

1.1 Introduction

Many communities depend on surface water resources such as rivers and streams for several purposes, such as drinking water, irrigation, recreation, aquaculture and livestock watering. Management of streams for such a variety of uses requires an understanding of physical factors that affect its quality. Of concern is the continuous threat of high-level pollution of these freshwater resources because of discharge of effluents from treatment plants, accidental spillage or by the intentional disposal of pollutants (Jobson, 1997).

In many cases, accidental spills are the greatest economic danger to these water resources. Predicting the subsequent movement and longitudinal spreading of pollutants in streams is necessary for a timely response by water authorities in respect of downstream consumers and mitigation purposes. This can only be possible if characteristics of pollutants migration in such surface watercourses are reliably known. There are numerous processes that transport matter within streams. Recent studies of characterising pollutant transport in streams have focused on physical mechanisms affecting solute concentrations which include advection, dispersion and transient storage (Wallis et al., 1998; Chin, 2013).

Advection, the migration of pollutants with the flowing water, physically transports water and solutes from one location to another. Dispersion is responsible for the longitudinal mixing of pollutants within the stream and is mainly due to spatial differences in velocity of the flow (Rutherford, 1994; Chakra, 2008). The other form of spreading is caused by dispersion into transient storage zones created for instance by vegetation. Modelling of pollutant transport in streams may include time of arrival, peak concentration and duration of occurrence of pollution. Therefore, with good knowledge of the factors that influence pollutant transport and the parameters which control them, spatial and temporal concentration of pollutants can be predicted with reasonable accuracy (Wallis et al., 2013).

Mathematical models are often used by engineers and scientists to simulate the physical processes of mass transport in streams, and several of them exist. However, dispersive aspects are subject to

significant complexities (Ani et al., 2009). Common uses include evaluation of the impact of pollution events and issues relating to stream ecology. The methods range from theoretical approaches to interpreting tracer experiments and empirical methods (Wallis et al. 2013; Wallis & Manson 2004). However, modelling solute transport in natural streams often involves tracer experiments. Application of tracer experiments has the advantage of assessing the reliability of mathematical models of solute transport as well as simulating the fate of solutes (Rutherford, 1994).

Mass transport modelling in streams has commonly been undertaken using models based on appropriate one-dimensional governing transport equations of the advection-dispersion model (AD-Model) and the transient storage model (TS-Model) (Wallis et al. 2013). These methods combine solutions of the governing equations with some form of parameter estimation tool. The TS-Model was designed to include diffusion of material in the so-called dead zones. However, there is little evidence that dead zone effect plays an important role in the stream reach under this investigation (Ani et al., 2009). Therefore, due to this evidence and the increase in modelling the complexity of the TS-model, it was not considered in detail in this research.

Identification of relevant parameters involves the calibration of an AD-Model expression by optimising the parameters so that the best fit is attained between measured and simulated concentration-time profiles known as break-through curves (BTCs). The estimated parameter values can be used later in case of a pollution event. The AD-Model can be applied through solutions of its governing equation, and several solutions are available depending on the initial and boundary conditions that are specified. The available solutions include analytical solutions (Ogata and Banks, 1961; Kumar et al., 2009), Eulerian numerical solutions (Bencala and Walters, 1983), and semi-Lagrangian numerical solutions (Manson et al., 2001). Where tracer data is available, analytical solutions are applied using the so-called routing procedures (Singh and Beck 2003; Rutherford 1994; Fischer 1967).

Analytical solutions and routing procedures can be used in idealised situations where the model geometry is simple, and parameter values are constant. Therefore, most actual situations would require numerical approaches to the solution of the governing equation. Numerical solution

techniques involve converting the governing differential equation into algebraic difference equations that can be solved for values at incremental points in space and time (Chapra & Canale 2008; Wallis 2007). In applying numerical methods, the results are only physically realistic when the discretisation scheme has important properties (Ferziger and Peric, 2002; Versteeg and Malalasekera, 2007). The important properties of numerical solution methods are conservativeness, transportiveness, stability, boundedness and accuracy (Versteeg and Malalasekera, 1995; Ferziger and Perić, 1996). Numerical solutions of the models are conservation equations; therefore, the equations should obey the conservation laws on local and global basis. Boundedness entails that solutions of the numerical ought to be within proper limits. Stability connotes that errors are not magnified in the process of a numerical solution. Transportiveness is the directionality of influencing in the flow and, therefore the advective velocity, the direction of flow and the Peclet number should be considered in the discretisation scheme. (Abbott and Basco, 1989; Versteeg and Malalasekera, 1995; Ferziger and Perić, 1996). Generally, if a component of the solution method does not possess the desired properties, neither will the complete method.

The solution also requires a reference frame and one of the two basic reference frames are usually used: Eulerian and Lagrangian. The Eulerian reference frame is a spatially fixed reference frame and is employed by In Lagrangian approach the view is from the river and the computational grid is determined by the fluid motion. In practice, a semi-Lagrangian frame is particularly useful because it recognizes the consequences of fluid motion while the modeller retains control of the control grid (Manson et al., 2001; Wallis, 2007). Regardless of the frame of reference used, it is necessary that the mathematical description should represent the physical situation. The semi-Lagrangian approach is used effectively in advection-dominated mass transport problems (Sun and Sun, 2015). Previous studies show no evidence that mass transport is advection-dominated in the stream being investigated (Ani et al., 2009; Wallis et al., 2013). Therefore, the semi-Lagrangian approach was not considered further.

All numerical solutions suffer from a range of errors, which are determined by the details of each scheme used. Consequently, when such a model is optimized to observed concentration data, the estimated parameter values are likely to be influenced by the formulation of the numerical solution,

the choice of time marching and the choice of reference frame adopted. Hence, there is evidence to suggest that depending on the numerical method used different solute transport parameter values can be estimated with the same observed data (e.g. Wallis et al. 2013). Therefore, prediction and interpretation of mass transport in streams has been observed to depend on the approach used (Wallis et al., 2013). Therefore, it is apt to calibrate and evaluate such models to determine their reliability by using experimental data over a range of stream flow rates and model resolutions for a reach of interest. Additionally, estimating solute transport parameters using BTCs requires an inverse modelling tool. Inverse modelling (or calibration) is the process of determining values of model parameters which give the best fit between predicted and the measured data (Chin, 2013; Semuwemba, 2011).

Although experimental methods have the advantage of validating results, experimental data is not easy to obtain. It is for this reasons that many researchers have attempted to find simplified models in the form of empirical equations developed using parameters obtained by experimental methods and pertinent channel and hydraulic variables for the computation of solute transport parameters, especially the longitudinal dispersion coefficient (Seo and Cheong, 1998; Kashefipour and Falconer, 2002; Singh and Beck, 2003; Wallis and Manson, 2004; Martin et al., 2013).

1.2 Problem statement

While much research has been done to develop empirical models for estimating the longitudinal dispersion coefficient for streams based on experimental methods, little attention has been given to investigate the impact of numerical methods and numerical properties on the construction of empirical models. Commonly, routing methods and the method of moments have been used to obtain parameter estimates for development of empirical models (Kashefipour and Falconer, 2002; Wallis and Manson, 2004; Ani et al., 2009), and it is not known yet how numerical methods and numerical properties influence the development of empirical models.

Typically, observed breakthrough curves (BTCs) are analysed to obtain solute transport parameters. Estimated parameter values are then used to develop an empirical model, using regression analysis by correlating these parameters with the bulk channel and flow characteristics. Routing methods

used to estimate parameters for the development of empirical models may be inadequate as numerical methods would ultimately be used to predict concentration evolution in the case of a pollution event.

There are several issues associated with the construction of predictive models concerning the use of experimental methods of the AD-Model to obtain parameter estimates:

- a) It is commonly known that the scheme and numerical properties influence estimates given by numerical methods. This is dependent on the relative intensity of dispersion and advective transport such that if the solute transport is dispersion dominated, numerical difficulties are not expected (Szymkiewicz, 2010). In such a case to solve the AD-Model one could select any numerical method. However, it is not known a priori which process is dominating and which numerical methods would be appropriate for the application. Therefore, it would be necessary to appraise numerical methods using data in which solute transport parameters are known and are similar for the application.
- b) If several numerical methods are identified for an application, it is not known a priori how numerical properties impact such methods based on observed sampled data. Additionally, it is not known how the construction of an empirical model is impacted by different numerical methods and numerical properties.
- c) An empirical model based on parameter estimates obtained by a routing procedure would more likely predict parameter estimates appropriate for use with a routing method when predicting concentration distribution at a downstream site. In most actual situations, prediction of concentration distribution at a downstream site would be determined by a numerical method. It is not clear whether predicted dispersion coefficients obtained by an empirical model based on a routing method could be used reliably by numerical methods and without consideration of numerical properties.
- d) Construction of an empirical model involves among other things model performance analysis. It is not well known how the performance of an empirical model is affected by the numerical method

used to estimate parameters and the numerical properties under which parameters were determined for its development.

Therefore, the problem statement reads:

Generally, Eulerian numerical solutions to the governing AD-Model give a wide array of parameter results for solute movement in streams, depending on the method used and model resolution. This does not engender confidence in the reliability of the resultant empirical models and their application in real-world pollution events.

From this problem statement, it can be asked whether the reliability of a developed empirical model is dependent on the numerical model employed to estimate parameters for its development, a result of numerical properties under which parameters for its development were obtained or a result the numerical method and numerical properties adopted for application of its predicted values or the result of a combination of the above.

1.3 Thesis statement

To address the shortcoming highlighted in the problem statement above, research aimed at generating knowledge towards enhancing the development of empirical models was the thesis of this study.

The thesis statement reads:

The reliability of an empirical model for estimation the longitudinal dispersion coefficient for a stream developed based on observed parameters by Eulerian numerical methods of the AD-Model is influenced by the combination of the solution method of the AD-model and the numerical properties under which the parameters were obtained for its development, and the solution method and numerical properties for application of its predictions.

1.4 Aims and objectives of the research

The study aimed to analyse BTCs and investigate the impact of Eulerian numerical methods of the AD-Model and numerical properties on estimating stream solute transport parameters. Furthermore, the aim includes the construction of empirical models, based on parameters obtained by numerical

methods, for estimating stream longitudinal dispersion coefficients. Knowledge obtained from this process was used to support or negate the thesis statement.

The specific objectives are:

- a) To appraise Eulerian numerical solution methods and a routing procedure of the AD-Model for estimating stream transport parameters using data in which solute transport parameters are known and are appropriate for the Murray stream.
- b) To analyse observed BTCs using a routing procedure and selected numerical methods and investigate the impact of numerical methods and numerical properties on estimated solute transport parameters
- c) To develop and calibrate empirical mathematical models for estimating stream longitudinal dispersion coefficients based on a routing procedure and numerical methods and numerical properties. Furthermore, to investigate impacts of numerical methods and numerical properties on models developed based on numerical methods.
- d) To analyse performance and evaluate empirical models based on numerical solution methods used and numerical properties adopted.

1.5 Delineations and Limitations

The research scope and limitations arise from the experimental methods and modelling methods adopted.

1.5.1 Delineations

- a) For the AD-Model, the research only considered solutions for an instantaneous release of solute in a stream, namely, analytical solutions, routing procedures and Eulerian numerical solutions. Analytical solutions were used to generate synthetic data to assess numerical methods, parameter estimation tools and a routing procedure.
- b) Pollution events in streams are of many types depending on the spatial and temporal distribution. The present study focused on the gulp pollutant release, characterised by an

instantaneous injection of a known mass of a pollutant across a stream.

- c) The research used a set of existing raw data, collected using the tracer Rhodamine WT, from Scotland in the United Kingdom observed in a single stream reach. This was done for two reasons, namely, (1) Ensuring a stable data source whereby differences in modelling could be explained by model solutions rather than by influence of differences in physical stream parameters and (2) extreme difficulties in obtaining this type of data since these kinds of long-term tracer studies are rare. Notwithstanding the use of a single stream reach, the data set was adequately large due to an extended data capture period of twelve flow rates collected throughout fifteen months; from 2009 to 2011.
- d) The study only used one tracer, Rhodamine WT, which is considered conservative, i.e. non-reactive and non-biodegradable. Therefore, results may not apply to pollutants which are non-conservative.
- e) The study only used dimensional concentration-time tracer data. Therefore, the findings may not apply to non-dimensional forms of tracer data.
- f) The numerical work only considered second-order and third-order finite difference (FD) and finite volume (FV) discretization schemes and a limited range of numerical properties.
- g) Inverse modelling methods for analysis of BTCs have largely been local-search methods (e.g. Doherty, 2008). Therefore, inverse modelling techniques were developed in Microsoft Excel which uses the steepest descent method, a local-search numerical parameter estimation method.
- h) Parameters estimation only considered the classical inverse problem (CIP) approach, which assumes no model structure error (Sun and Sun, 2015).
- i) Empirical models were developed by correlating parameter values estimated by numerical methods to the site hydraulic characteristics of the Murray stream (flow rate, cross-sectional average velocity, shear velocity and depth of flow), and the structures of empirical models

were identified using a dimensional analysis approach and least-squares regression.

- j) Performance analyses of developed models only involved linear regression, descriptive statistics, error statistics, and residual error analysis.

1.5.2 Limitations

- a) Numerical methods refer only to second-order and third-order Eulerian numerical methods that are documented and published in well-known peer reviewed and accepted documentation.
- b) Discretisation methods refer to the finite difference, and the finite volume approaches for the structured numerical grid.
- c) The study only considered the one-dimensional AD-Model, with constant values of longitudinal dispersion and velocity, for an instantaneous release of a solute consistent with the observed tracer data.
- d) Solution methods refer to explicit and implicit time marching methods for the unsteady mass transport problem in a stream.
- e) Parameter estimation tool (solver) refers to Excel spreadsheets. Excel uses the steepest descent method for numerical parameter estimation. In this study, the Generalised Reduced Gradient method was used.
- f) Experimental data refer to concentration-time profiles, also known as breakthrough curves (BTCs) for an instantaneous release of Rhodamine WT as a tracer.
- g) Empirical model refers to a simplified model for predicting longitudinal dispersion coefficients in natural channels.
- h) Factors that influence dispersion considered for development of empirical equations refer to flow rate, cross-sectional average velocity, shear velocity and mean flow depth.
- i) Performance measures refer to graphical methods, descriptive statistics, error statistics and

residual error analysis.

1.6 Assumptions

The study considered a one-dimensional solute transport problem of a single stream reach of the Murray stream. The problem involved predicting observed upstream concentration-time profiles to simulate observed temporal concentration-time profiles a downstream site. The one-dimensional advection-dispersion stream mass transport problem using tracer data is based on several assumptions, most of which are presented in Novak et al. (2010), as follows:

- a. The ratio of longitudinal to cross-sectional stream dimensions is considered large enough for the flow and dispersion to be considered as one-dimensional. Therefore, the concentration of the solute was assumed to be uniform within a cross-section of the stream reach.
- b. The solute concentration is adequately low to affect the density of water. Consequently, stream hydrodynamics are not influenced by the solute transport processes.
- c. The main solute transport processes are advection, turbulent diffusion and differential advective dispersion and the solute is assumed to be conservative. Furthermore, turbulent diffusion and the differential advective dispersion are collectively modelled by the longitudinal dispersion coefficient, and advection is modelled by cross-sectional average flow velocity. Therefore, the advected dissolved tracer is transported at the same velocity as the cross-sectional velocity of flowing water.
- d. The channel cross-section is uniform, and the solute transport processes are constant over the reach length.
- e. Longitudinal dispersion is modelled using a Fickian diffusion-based law where the longitudinal dispersion coefficient is proportional to the cross-sectional mean flow velocity.
- f. The errors in simulations and estimated parameters are a result of the solution methods of the AD-Model used and the numerical properties, and a result of the numerical parameter estimation method.

1.7 The significance of the study

When appropriate stream in-situ data is unavailable and transport and dispersion characteristics are unknown, practising engineers and scientists requiring estimation of solute transport parameters, especially a longitudinal dispersion coefficient, have little choice but to employ an empirical model that aims to relate longitudinal dispersion coefficient to the bulk flow and channel characteristics (e.g. Wallis and Manson, 2004). The concern is that empirical equations have been investigated based on their formulations without considering the various methods and conditions under which data were observed for their development (Wallis and Manson, 2004).

Documented research shows that many workers have relied on using routing procedures and the method of moments to estimate parameter values for development of empirical models (Seo and Cheong, 1998; Falconer, Kashefipour and Falconer, 2002; Wallis and Manson, 2004). Routing procedures are typically applied for simple cases of initial and boundary conditions. In most practical situations a numerical method would be appropriate depending on its formulation, boundary conditions and initial conditions. However, numerical methods are influenced by the properties of the scheme and numerical properties (Ferziger and Peric, 2002; Versteeg and Malalasekera, 2007). Certainly, if parameter values obtained by numerical methods are to be used to derive predictive empirical models, the parameter values so determined must be reliable as predictive models are dependent on obtaining reliable estimates of parameter values by experimental methods, otherwise significant inaccuracies will result in predictions by empirical models (Wallis & Manson 2004; Falconer et al. 2002; Singh & Beck, 2003) and subsequently concentration predictions.

The study arose from a need to investigate the impact of numerical methods and numerical properties on parameter estimation and subsequently on the construction of empirical models. Previous studies have done similar but differed from the current research as parameter values have been estimated using routing procedures and the method of moments (Singh and Beck, 2003). However, where a numerical method has been used to estimate parameter values, the influence of the solution method and numerical properties has not been addressed (e.g. Ani et al., 2009).

Since parameters estimated by empirical models would ultimately be required to be used with a

numerical method for most practical situations, it would be necessary to investigate the impact of numerical methods and numerical properties on construction and application of empirical models. Additionally, parameter values for use with a numerical method must be appropriate for the type of numerical method, and the model resolution adopted. Therefore, the focus of this research was to investigate the impact of numerical methods and numerical properties on the construction and application of empirical models.

1.8 Dissertation Layout

Chapter 2 summarises relevant published literature and stresses issues relating to previous studies on parameter estimation of the AD-Model for development of predictive models. The emphasis is on estimating solute transport parameters using numerical methods, especially on the influence of properties of Eulerian numerical methods and non-dimensional numerical properties on estimated parameters. Central to this review is the theory on advection and dispersion. Secondly, the chapter discusses previous methods on the development of empirical models for estimating longitudinal dispersion coefficient in natural channels. The focus is on methods for estimating parameters for the development of empirical models. Finally, the chapter discusses the construction process of empirical models including performance characterisation of a constructed model.

Chapter 3 presents an assessment, using synthetic tracer BTCs, of six possible numerical methods of the AD-Model. The emphasis is on the influence of numerical properties on optimised parameter values. The assessment took into consideration the relative strengths of transport processes likely to be encountered when analysing observed BTCs for the Murray stream.

Chapter 4 presents an analysis of collected data and modelling of observed BTCs to estimate velocity and longitudinal dispersion coefficients using selected methods.

Chapter 5 presents data for the construction of models, identification of the general expression of the models, identification of unknown model parameters and calibration of developed models to estimate stream longitudinal dispersion coefficients.

Chapter 6 presents a performance analysis approach for models. In the context of model

development, several methods and criteria are considered to form a broad opinion on the performance of models.

Chapter 7 presents confirmation of developed and calibrated empirical models and comparison of developed models with existing empirical models testing of several existing models. Predicted dispersion coefficients are compared about stream flow rate.

Chapter 8 presents the conclusion from this research and recommendations for further research.

2 LITERATURE REVIEW

2.1 Introduction

A Pollutant released into a stream experience stages of mixing as the stream flow transports it downstream. The two main mixing zones in the movement of a cloud of solute along a channel after a slug are released are the initial mixing zone and the equilibrium zone. In the initial mixing zone, pollutants are spread longitudinally, transversely and vertically by transport processes. In the equilibrium zone, the cross-sectional spreading is complete, and the process of longitudinal shear dispersion is the dominant mechanism (Fischer et al., 1967). In the case that longitudinal dispersion is the dominant process, the one-dimensional AD-Model can be applied to make reasonable estimates of the rate of contaminant transport downstream. Thus, reasonable use of the AD-Model is restricted to the equilibrium zone, which is after the initial mixing zone (or advective zone) from the source of the pollutant (Wallis et al., 2013).

Application of the one-dimensional AD-Model to predict concentration evolution of dispersants in natural streams requires reliable parameter values. Therefore, the selection of proper parameter values, especially a dispersion coefficient, is the basic and the most demanding task (Seo and Cheong, 1998). Several methods have been developed for estimating solute transport parameters for streams. The methods range from theoretical equations to interpreting tracer experiments (e.g. Wallis, 2007; Wallis and Manson, 2004). Another method for estimating solute transport parameters is the use of empirical equations.

Theoretical equations include the flow integration method and the flux integration method. Both are based on knowing the cross-sectional distribution of certain parameters, which in practice is rarely achievable. However, the flow integration method is the basis of several simplified methods, including empirical equations, for estimating the dispersion coefficient (Wallis and Manson, 2004). Estimation of the solute transport parameters in natural streams often involves applying tracer experiments (Fischer et al., 1979). It is necessary to conduct tracer experiments in natural streams owing to the influence of many factors such as riparian vegetation, meanders, roughness etc., that influence transport processes. Application of tracer experiments has the advantage of assessing the

reliability of mathematical models of solute transport as well as simulating the fate of solutes (Wallis, 2005). Common experimental methods include the method of moments, the reduction of peak method, the Chatwin graphical method and calibration of the advection-dispersion model (AD-Model) and the transient storage model (TS-Model).

The method of moments requires capturing complete concentration-time profiles in estimating dispersion coefficients. In natural channels profiles have long tails, which require a long time to measure (Wallis, 2005). Difficulties in measuring profiles' leading and trailing edges result in errors in evaluating the dispersion coefficient (Rutherford, 1994). The enhanced method of moments is now being considered as an alternative to the traditional method of moments (Wallis and Manson, 1994).

The reduction of peak method involves plotting peak concentrations at several sites against the inverse of the square root of the peak time. The dispersion coefficient is estimated from the gradient of the plot. The method has the disadvantage that it requires knowledge of the cross-sectional flow area and the mass of the dispersant (Heron, 2015).

The Chatwin graphical method involves transforming Taylor's analytical solution of the ADM and rearranging the equation by taking natural logarithms of both sides which are then plotted. Like the reduction of peak method, the dispersion coefficient is estimated from the gradient of the plot. Experiments have shown that few concentration-time data plots a straight line (Rutherford, 1994). Therefore, parameter values obtained by this method are bound to be subjective.

It has been observed that best practices for estimating the parameters combine tracer experiments with mathematical models of the transport processes, namely the AD-Model (Fischer et al., 1979; Rutherford, 1994; Wallis, 2005) and the TS-Model (Chin, 2013; Wallis et al., 2013). Tracer data is obtained from tracer experiments in the form of concentration-time profiles (or breakthrough curves; BTCs). The models can be applied through calibration of solutions of the governing equations. There is a range of calibration methods for both the AD-Model and the TS-Model all of which aim to find those parameter values that obtain a good fit between the model output and the measured temporal concentration data. However, where transient storage is not significant, as in the Murray stream (Ani et al., 2009), only the AD-Model can be adequately applied.

The available solutions of the AD-Model include analytical solutions (Ogata and Banks, 1961; Kumar, Jaiswal and Kumar, 2009), Eulerian numerical solutions (Hoffman, 2001; Chapra, 2008) and semi-Lagrangian numerical solutions (Wallis et al. 1998; Manson et al. 2001). Additionally, analytical solutions can be applied with tracer data using the so-called routing procedures (Rutherford, 1994; Singh and Beck, 2003). When calibrating the AD-Model two parameters are estimated namely, the velocity and the longitudinal dispersion coefficient.

Most practical situations require the use of numerical solutions of the AD-Model, such as the Crank-Nicolson scheme, because ultimately the models are required for more complex problems than those that allow the use of analytical solutions or routing procedures (Martin and McCutcheon, 1998; Wallis, 2007b). There are various Eulerian discretisation schemes with different properties depending on the discretisation approach and the solution method. Also, the results are also influenced by numerical properties, namely, the advection (or Courant) number and the dispersion (or diffusion) number. These properties are characterised by the Peclet number which measures the significance of the advective flux against the diffusive flux (Wallis and Manson, 1997; Versteeg and Malalasekera, 2007; Wallis, 2007). Quantitatively, it is expressed as the ratio of advection number (or Courant number) to dispersion number (or diffusion number.)

Parameter estimation by methods based on the AD-Model using BTCs involves optimisation methods which involve fitting a parameterised function to a set of measured data points by minimising the errors between the data points and the function output. It may be a parameter-space-search (or manual) procedure, where the parameter estimates are manually adjusted during simulations to obtain the best fit between simulated and observed concentrations. Alternatively, a least-squares parameter estimation technique can be employed instead of the parameter-space-search approach where the principle of nonlinear least squares can be employed to find the best fit. The methods involve iterative improvements of parameter values to identify the optimal parameter set between the function and the measured data points. The steepest descent, the Gauss-Newton and the Levenberg-Marquardt methods are common techniques that can be used to estimate parameter values (Singh and Beck, 2003; Karahan, 2008). In this study, the Excel was used which employs the steepest descent method due to several advantages. The level of agreement between

the model output and the observations is used to assess the performance of a model. However, there are several quantitative measures of performance of simulations, depending on the relevance of the components of the model for the use (Chin, 2013).

Experimental methods give the best results in estimating mass transport parameters in streams (Wallis and Manson, 2004; Chin, 2006; Ani et al., 2009). However experimental data is not easy to obtain, and it is for this reason that many workers have attempted to find options in the form of simplified models (e.g. semi-empirical and empirical equations) for the computation of solute transport parameters (Seo & Cheong 1998; Fischer et al. 1979). Empirical equations are based on correlating the results from experimental methods against pertinent channel and hydraulic variables.

2.2 Physical processes of solute transport in streams

Solutes released into channels are translated downstream and are spread out along the channel mainly by shear flow dispersion with a small contribution from turbulent diffusion and a negligible contribution from molecular diffusion (Rutherford, 1994). A slug of solute released instantaneously at a location initially spreads in three dimensions. Complete vertical mixing occurs quickly, followed by complete transverse mixing at a much later time. Longitudinal mixing (or dispersion) takes place indefinitely.

Mixing because of the molecular-scale random velocity variations is called molecular diffusion and mixing because of random turbulent macroscopic velocity variations is known as turbulent diffusion. (Chapra, 2008; Chin, 2013). Mixing due to molecular and turbulent diffusion are related to the random movement of particles. Consequently, both molecular and turbulent diffusion are generally assumed to obey Fick's law (Chin, 2013). Additional mixing is caused by spatial variations of the local velocity and mixing by this process is known as dispersion (Wallis, 2005; Chin, 2013). In-stream mixing is dominated by dispersion due to the strong shears generated by the large flow rates and the confining channel. Generally, turbulent diffusion and dispersion can cause mixing individually or in combination (Chapra, 2008). An essential element of dispersion is the spreading of solute mass that occurs due to the interaction of velocity shear and diffusion (Chin, 2013). Unfortunately (and incorrectly), the terms dispersion and diffusion are often used interchangeably to describe the

macroscopic mixing of solutes in water (Rutherford, 1994). Consequently, as far as the physical mass transport is concerned in streams, dispersion is considered to obey Fick's law (e.g. Taylor, 1954).

Advection describes the downstream movement of solute mass at a local or cross-sectional mean flow velocity. Transient storage is a process where solutes are temporarily detained in eddies and stagnant pools of water. The main channel of the stream is the portion of the stream in which advection and dispersion are the dominant hydrologic processes, while the portion of the stream that contributes to transient storage is mainly confined to the stream bed and banks, including the hyporheic zone where connections with the surrounding saturated ground exist. Lateral inflow results from groundwater and overland flow during floods. Lateral inflow tends to dilute solute concentrations in the whole stream channel unless it is highly polluted.

Two main zones are recognised in the movement of a cloud of solute along a channel after a slug is released. The two zones are the initial mixing zone and the equilibrium zone (Rutherford, 1994). The initial mixing region or advective zone is where cross-sectional mixing of the solute takes place, and the equilibrium zone is where the solute is mixed well enough such that further evolution of the cloud takes place at a constant rate in the flow direction (Fischer and List, 1979; Rutherford, 1994; Wallis, 2005; Chin, 2013). The important parameters for estimating the initial mixing zone are the nature of the source (point or line source), a transverse characteristic length and the transverse mixing coefficient (Shucksmith et al., 2007). The distance for equilibrium to be established is estimated by the following expression (Rutherford, 1994):

$$X_i = \frac{\alpha v l_t^2}{\varepsilon_t}, \quad (2.1)$$

where X_i = the length of the initial mixing zone, v = cross-sectional mean velocity, l_t = a characteristic transverse length, ε_t = transverse mixing coefficient and α = a constant. It is suggested that the characteristic transverse length should be the lateral distance from the point of maximum velocity to the furthest bank. Published estimates of the constant α range from 0.2 to >

10 (Rutherford, 1994). The transverse mixing coefficient is estimated by the shear velocity and depth of flow. Shear velocity is an expression by which shear stress is rewritten in units of velocity. The turbulent velocity variations in the vertical and transverse directions are of the same order as the shear velocity (Chin, 2013). A reasonable estimation of the initial advective distance X from the source of the dispersant is given as (Magazine et al., 1988).

$$X_i = 1.8 \frac{l_t^2 v}{RU_*}, \quad (2.2)$$

where U_* = the shear velocity and R = the hydraulic radius. In the equilibrium zone, a balance exists between differential longitudinal advection, which stretches the cloud longitudinally, and cross-sectional mixing, which reduces the stretching potential of the differential longitudinal advection. The processes causing longitudinal dispersion in the equilibrium zone are transverse shear, turbulent longitudinal diffusion and cross-section mixing, consisting of transverse turbulent diffusion and secondary flows (Rutherford, 1994; Chanson, 2004; Wallis, 2005). The other processes that are often disregarded, but could be important, are processes of sorption and desorption of solutes on sediments on the bed of a stream (Chin, 2013).

Rutherford (1994), Chanson (2004) and Fischer et al. (1979) attribute the theory of longitudinal dispersion in rivers to pioneering work in pipe flow by Taylor (1954). Taylor's model has proved to give satisfactory predictions of solute transport from a point source in bounded channels (Rutherford, 1994). Taylor (1954) found that once the equilibrium between longitudinal advection and cross-sectional mixing is attained, the variance of the solute concentration distribution increases linearly with time, while its skewness diminishes.

Many experimental studies have shown that temporal concentration distributions are highly skewed with long tails because of tracers being trapped and released from zones of separated flow and boundary layers. The effect is generally known as storage zone effect or transient storage (Chapra, 2008; Chin, 2013). Transient storage is associated with dead zones, which are stagnant or eddy like flow structures located around the channel boundaries. However, similar, but larger and more

dynamic semi-isolated storage areas exist in river bends, at river junctions, and around obstructions (Wallis et al., 2013). The hyporheic zone, where there are interactions between the flow and the surrounding groundwater, also provides locations where solute can be dynamically stored.

2.3 The one-dimensional advection dispersion model

Mixing in streams is primarily caused by longitudinal dispersion which results from the stretching effect of velocity gradients. The longitudinal dispersion coefficient is used to measure the longitudinal mixing of a solute which is well mixed across a channel (Chin, 2013). Taylor (1954) argued that when a solute is well mixed in a stream longitudinal dispersion can be described by Fick's law (Wallis, 2009; Chin, 2013). Therefore, based on Taylor's (1953, 1954) analysis, in the equilibrium zone, concentration distribution can be modelled using an analogy of Fick's law (Fisher et al., 1979). Therefore, the diffusion coefficient in the AD-Model is replaced by a coefficient that encompasses all the processes that contribute to mixing in streams. The general one-dimensional equation describing longitudinal transport is expressed as follows (Rutherford, 1994; Chanson, 2004):

$$\frac{\partial(A\varphi)}{\partial t} + \frac{\partial(Av\varphi)}{\partial x} = \frac{\partial}{\partial x} \left[DA \frac{\partial \varphi}{\partial x} \right], \quad (2.3)$$

where A is the cross-sectional area of the channel (L^2), D is a longitudinal dispersion coefficient (L^2T^{-1}), v is cross-sectional average longitudinal velocity (LT^{-1}), φ is cross-sectional average concentration (ML^{-3}), x is longitudinal coordinate (L) and t is time (T). Equation (2.3) is the Taylor's (1953, 1954) and Fisher's one-dimensional advection-dispersion model (Fischer et al., 1979). Commonly constant mixing rates and cross-sectional average velocity are assumed (Chanson, 2002). The assumption of constant parameters is appropriate if the channel is straight and the discharge is uniform and steady. Therefore, for constant dispersion coefficient and velocity, the one-dimensional advection-dispersion model is written as (Rutherford, 1994),

$$\frac{\partial(\varphi)}{\partial t} + v \frac{\partial(\varphi)}{\partial x} = D \frac{\partial^2 \varphi}{\partial x^2}. \quad (2.4)$$

In the equation (2.4), the longitudinal dispersion coefficient quantifies the rate of longitudinal

stretching of a solute cloud, and the cross-sectional average velocity quantifies the rate of downstream movement of the whole cloud (Chanson, 2004; Wallis, 2009). An advection-dispersion problem requires a complete mathematical statement consisting of the AD-Model and the boundary and initial conditions specified. The AD-Model is applied through calibration of solutions of the governing equation, and several types of solution exist as stated above. Solutions to the AD-Model depend on the specified boundary and initial conditions (Fischer et al. 1979, Barnett 1983, Chanson 2004). When calibrating the AD-Model two parameters are estimated namely, the velocity and the dispersion coefficient.

The available solution types include analytical solutions (Ogata and Banks, 1961; Kumar, Jaiswal and Kumar, 2009), routing procedures (Rutherford, 1994; Singh and Beck, 2003), Eulerian numerical solutions (Hoffman, 2001; Chapra, 2008) and semi-Lagrangian numerical solutions (Wallis et al. 1998; Manson et al. 2001). When calibrating the AD-Model two parameters are estimated namely, the velocity and the longitudinal dispersion coefficient. Analytical approaches are derived as functions that are precise and continuous in time and space. The process of routing is a convolution operation by which a temporal tracer profile at a downstream site is predicted from an upstream temporal dispersant profile by a response function which is a solution of the AD-Model for unit concentration (McCuen, 1998; Wallis et al., 2013). The convolution process involves multiplication, translation with time and addition of the upstream concentration distribution. Usually, routing procedures are applied to temporal concentration profiles, and spatial resolution of the solution domain is not required (Singh and Beck 2003; Wallis et al. 2013). Most practical situations require the use of numerical solutions because ultimately the models are required for more complex problems than those that allow the use of analytical solutions or routing procedures (Martin and McCutcheon, 1998; Wallis, 2007). Numerical methods are applied by numerical convolution, where both time and space are discretised (Abbott and Basco, 1989).

2.4 Analytical solutions of the one-dimensional advection dispersion model

An advection-dispersion problem requires a complete mathematical statement consisting of an advection-dispersion model, and the boundary and initial conditions specified. There are essential

solutions of the AD-Model that can be considered as the bases from which other solutions can be developed. These essential solutions mostly correspond to slug releases of a solute in a uniform stream; with velocity field which is spatially uniform. The common solutions for a slug release of dispersant are Taylor's solution (Taylor, 1954), the Singh and Beck solution (Singh and Beck, 2003) and the Hayami solution (Barnett, 1983). The Taylor's solution of the one-dimensional AD-Model with constant coefficients is written as (Rutherford, 1994),

$$\varphi(x,t) = \frac{M}{A\sqrt{4\pi Dt}} \exp\left[-\frac{(x-vt)^2}{4Dt}\right], \quad (2.5)$$

where M is the mass of the solute, and other symbols are as previously defined. The solution satisfies the following initial condition and boundary conditions (Barnett, 1983):

$$\varphi(x,0) = \varphi_0(x) \quad -\infty < x < \infty, \quad (2.6)$$

$$\varphi(x,t) = \varphi_0 \quad x \rightarrow -\infty, \quad (2.7)$$

$$\varphi(x,t) = \varphi_1 \quad x \rightarrow \infty. \quad (2.8)$$

The Taylor solution is reliable at long distances downstream from the point where the slug is released (Rutherford, 1994). The solution is based on the situation where the concentration is known as a spatial distribution at an initial time. Taylor's expression shows that the concentration distribution at a time, along with the length of a stream, is Gaussian and that the temporal concentration distribution at a location is non-Gaussian (Rutherford, 1994; Singh and Beck, 2003). Observations of solute transport in streams are usually undertaken in the time domain and observed concentration distribution with time are not Gaussian (Chapra, 2008) and, therefore, do not necessarily obey Taylor's solution. There are several reasons why field data may not obey Taylor's solution, e.g. the solution may not fulfil the initial conditions in the field, and the assumption of uniform concentration across the stream section may not hold. Also, dead zones in the stream often create long tails in concentration profiles resulting in further deviations of temporal profiles from Gaussianity (Singh and Beck, 2003). Similarly, observed data may not obey other solutions of the AD-Model for similar reasons.

The Hayami solution for one-dimensional AD-Model is expressed as (Barnett, 1983; Rutherford, 1994):

$$\varphi(x,t) = \frac{Mx}{Avt\sqrt{4\pi Dt}} \exp\left[-\frac{(x-vt)^2}{4Dt}\right], \quad (2.9)$$

where, all the symbols are as previously defined. The solution satisfies the following initial condition and boundary conditions (Barnett, 1983):

$$\varphi(x,0) = 0 \quad -\infty < x < \infty, \quad (2.10)$$

$$\varphi(0,t) = \varphi_0 \quad 0 < t < \infty, \quad (2.11)$$

$$\frac{\partial \varphi}{\partial x}(x,t) = 0 \quad x \rightarrow \infty. \quad (2.12)$$

The solution applies to a situation where concentration is known as a temporal distribution at a boundary. Hayami solutions are more asymmetrical than the Taylor solutions (Barnett, 1983). Since experimental temporal distributions are asymmetrical, the Hayami solutions are more likely to fit observed BTCs more accurately (Barnett, 1983). Although the Taylor solution applies when an initial condition is known, and the Hayami solution applies when a boundary condition is known, they ultimately converge when applied to an impulse function representing an instantaneous loading on a channel (Barnett, 1983).

The Singh and Beck (2003) solution is applicable in the case where the initial concentration is equal to a fixed non-zero value; this initial concentration distribution is a step function and thus the initial and boundary conditions being as follows:

$$\varphi(x,0) = \begin{cases} \varphi_0, x < 0 \\ 0, x \geq 0 \end{cases}, \quad (2.13)$$

$$\varphi(0,t) = \begin{cases} \varphi_0, t = 0 \\ 0, t = \infty \end{cases}. \quad (2.14)$$

The solution of the governing equation, with constant velocity and dispersion coefficient, is (Singh and Beck, 2003; Barnett, 1983; Chapra, 2008):

$$\frac{\varphi}{\varphi_0} = \frac{1}{2} \operatorname{erfc}\left(\frac{x-vt}{2\sqrt{Dt}}\right) + \frac{1}{2} \exp\left(\frac{vx}{D}\right) \operatorname{erfc}\left(\frac{x+vt}{2\sqrt{Dt}}\right), \quad (2.15)$$

where $\operatorname{erfc}(\bullet)$ is a complementary error function which is expressed as

$$\operatorname{erfc}(\bullet) = 1 - \operatorname{erf}(\bullet) \quad (2.16)$$

where $\operatorname{erf}(\bullet)$ is the error function (e.g. Chin, 2013; Chakra, 2008). Equation (2.15) satisfies the following boundary and initial conditions (e.g. Singh and Beck, 2003):

$$\varphi(0, t) = \varphi_0, \quad (2.17)$$

$$\varphi(x, \infty) = 0 \quad (2.18)$$

Other solutions of the AD-Model for other pollution scenarios are developed based on the principle of superposition (McCuen, 1998). The AD-Model with constant parameter values is a linear partial differential equation. The principle of superposition states that ‘the sum of multiple solutions to a linear differential equation is also a solution to the differential equation and the corresponding boundary and initial condition is the sum of the multiple boundary and initial conditions respectively (Chin, 2013).

2.5 Numerical solutions of the advection dispersion model

2.5.1 Introduction

Numerical methods are derived by converting the AD-Model into algebraic difference equations that can be solved for values not known at incremental points in space and time (Wallis, 2007; Chapra, 2008). Traditionally, three discretisation methods have been popular for solving the AD-Model, namely, finite difference (FD), finite volume (FV) and finite element (FE) (Abbott and Basco, 1989; Ferziger and Peric, 2002; Versteeg and Malalasekera, 2007). In this study, the finite element method was not considered. In simple terms, differential terms at points in the space-time plane are approximated by functions of discrete concentration values at surrounding computational nodes. Usually, the main challenge is the formulation of an appropriate scheme for the values of the transported property when accounting for the advective contribution to the solution (Wallis and

Manson, 1997; Versteeg and Malalasekera, 2007; Wallis, 2007). This is dependent on a model resolution which is characterised by numerical properties (Wallis and Manson, 1997; Versteeg and Malalasekera, 2007; Wallis, 2007). The numerical properties that characterise model resolution are the advection number (also known as the Courant number), dispersion number (also known as the diffusion number) and the Peclet number (also known as the cell Reynolds number) (Abbott and Basco, 1989; Chapra, 2008; Chin, 2013).

A numerical method can be advanced in time in two ways, namely, explicit or implicit time marching approach. For explicit schemes, computation of the transported property at the advanced time step is determined by values of the transported property at the previous time step. In implicit methods, computation of the transported property at the advanced time can be computed simultaneously for all cells or nodes. The explicit schemes are computationally simple, while the implicit is more complex as generally a system of simultaneous equations must be solved (Wallis, 2007; Chapra, 2008). The calculations are also constrained by providing boundary conditions at the spatial edges of the computational domain.

2.5.2 Grid discretisation approaches

The starting point of a numerical method is the mathematical model which is usually a differential equation followed by identification and selection of a suitable method of approximating the differential equation by discretisation. In the present context, the AD-Model is approximated by algebraic equations which are solved. The algebraic equations are applied to small spatial and temporal domains such that the numerical solution provides results at discrete points in space and time. The accuracy of the results of a numerical solution is partly influenced by the quality of discretisation used. Inaccuracies resulting from the use of discretised algorithms to produce solutions are a result of approximations made in the discretisation process, idealisations in the differential equations, and iterative methods used in solving the algebraic equations (Ferziger and Peric, 2002; Versteeg and Malalasekera, 2007). The most common numerical grid is the structured or regular grid. Once the choice of grid type has been made, an approximation method to be used in the discretisation process must be chosen.

The FV approach subdivides the solution domain into several control volumes (CVs), as illustrated in Figure 2-1. The numerical grid defines the discrete locations where the transported variable is to be determined. Thus, the domain of the solution is divided into control volumes with a computational node at the centroid of each control volume at which the variable values are calculated (Ferziger and Peric, 2002; Versteeg and Malalasekera, 2007).

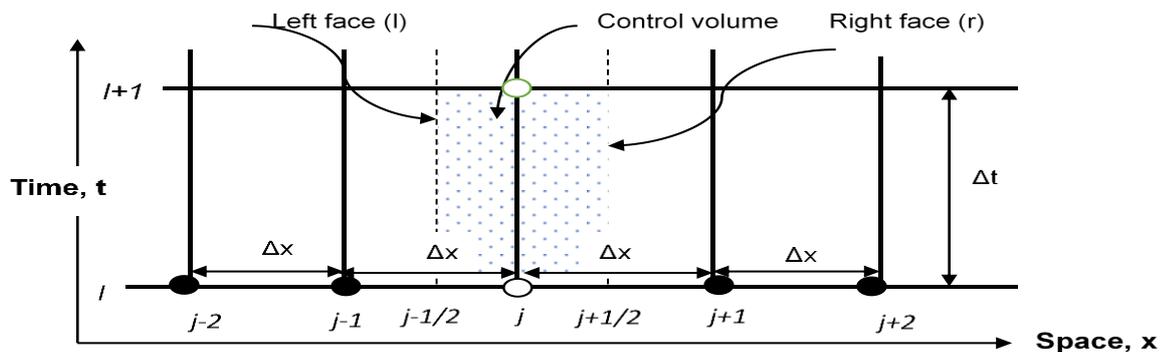


Figure 2-1: A structured computational grid for the finite volume method

The starting point for discretising the AD-Model using the FV approach is a conservation equation in integral form. The conservation equations are then applied for each control volume to determine the value of the dependent variable at the computational node. The values of the dependent variable at the surfaces of the control volume are expressed regarding nodal values by interpolation. The method results in an algebraic equation for each control volume, where several neighbouring nodal values appear. The approach is conservative if the integrals that represent advective and dispersive fluxes are the same for the control volumes with a common boundary (Ferziger and Peric, 2002).

The FD approach subdivides the solution domain into a mesh by grids, in which grid lines serve as local coordinates, as shown in Fig. 2-2. The starting point for discretising the AD-Model starts with a conservation equation in a differential form. Approximations for the derivatives at the grid points must be selected. The partial derivatives in the AD-Model are expressed regarding nodal quantities of both dependent and independent variables. The discretisation results in algebraic equation(s) with all unknowns prescribed at discrete mesh points of the solution domain. Approximation of the derivatives is achieved by Taylor series expansion or polynomial fitting of the dependent variable in

terms of the coordinates. If required values of the dependent variable can be determined similarly at other points other than the grid nodes. This results in an algebraic equation for each grid node, such that the value of the dependent variable at the node and several neighbour nodes are unknown (Ferziger and Peric, 2002; Wallis, 2006).

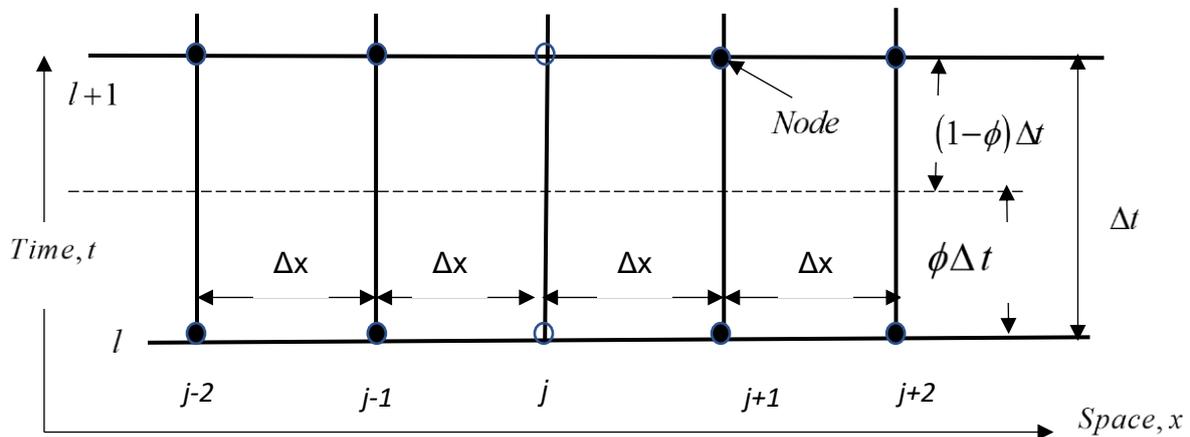


Figure 2-2: A structured computational grid for the finite difference method

In both FV and FD approximation methods, the choice of the approximation method influences the accuracy the approximations and the difficulty of developing the solution method. However, a compromise between ease of implementation, accuracy and computational efficiency should be made, depending on the problem. For unsteady transport problems, the methods based on marching in time are used, where the problem is solved at each time step (Wallis, 2007).

2.5.3 Discretisation of the advection dispersion model

There are various discretisation schemes with varying properties (van Leer, 1974; Leonard, 1979; Abbott and Basco, 1989; Versteeg and Malalasekera, 1995; Hoffman, 2001). The properties of the discretisation schemes are influenced by discretisation approach (FV or FD), discretisation scheme (e.g. upstream differencing, centred differencing, etc.), solution method (explicit or implicit) and order of error of the scheme.

The main problem in using numerical methods of the AD-Model is the formulation of an appropriate expression for the values of the dependent property when accounting for the advective contribution

of the transported property to the solution (Abbott and Basco, 1989; Versteeg and Malalasekera, 2007; Wallis, 2007b). The most important feature of the inaccurate formulation of expression for the advective term is that truncation errors of the term can result in artificial dispersion (also called numerical diffusion) that can subdue physical dispersion, and wiggles can occur even for discretisation schemes that are bounded (Ferziger and Perić, 1996). The problems associated with the discretisation of the advective term are also influenced by the numerical grid resolution, which is characterised by stream transport properties, namely, advection (or Courant) number and the dispersion (or diffusion) number (Wallis and Manson, 1997; Versteeg and Malalasekera, 2007; Wallis, 2007b). However, in many numerical computations of the AD-Model, the schemes are characterised in terms of the Peclet number (Sobey, 1984; Ferziger and Peric, 2002). The Peclet number indicates whether the advective flux dominates or the diffusive flux dominates (Chin, 2013). Accordingly, high values of the Peclet number indicate that transport due to advection dominates and low values of the number indicate that transport due to diffusion dominates; it is a measure of the relative strengths of advection to dispersion. It is necessary that the relationship between the value of the Peclet number and the transportiveness (or directionality of influencing) is considered in the discretization scheme of a numerical model (Versteeg and Malalasekera, 2007).

Numerous numerical solutions have been proposed to solve the AD-Model. They have various properties, and not all of them are equitably suitable (e.g. Wallis, 2007; Szymkiewicz, 2010). Basically, numerical methods are derived from the Taylor series expansion, and a truncation error is always present. Numerical schemes generally produce a numerical diffusion (artificial dispersion) and numerical dispersion. The numerical diffusion causes an unphysical amplitude attenuation and the numerical dispersion affects the advective velocity. The numerical dispersion frequently generates unphysical oscillations in the solution (Szymkiewicz, 2010). The numerical diffusion (or artificial dispersion) ordinarily induce effects like physical diffusion (or physical dispersion). It should be noted that in river mixing the main mixing mechanism is shear dispersion, and hence the advection-diffusion equation is called the advection-dispersion equation. Therefore, dispersion in the context of river mixing is completely different to dispersion in the context of numerical dispersion (e.g. Chapra, 2008)

Several numerical schemes have been proposed to solve the AD-Model. The schemes have different characteristics and not all them are of equal utility. Nevertheless, it can be said that practically, all numerical schemes work well for dispersion (or diffusion) dominated transport systems (Szymkiewicz, 2010). Therefore, it is essential to be aware of the kind of solution regarding the physical transport processes included in the considered equation. Generally, schemes that are first-order in space have been observed to be inaccurate and induce high values of artificial dispersion (Roache et al., 1986; Roache, 1997). Although the traditional centered-differencing scheme (CDS) is second-order in space, it is only stable when the Peclet number is not more than 2.0 and the advection number is not more than 1.0 (Versteeg and Malalasekera, 2007). This also applies to the Hybrid scheme, the power-law scheme and the exponential-differencing scheme when they are applied out of their limited conditions for which they are well founded such as steady transport problems (Roache, 1997; Versteeg and Malalasekera, 2007). However, very high-order schemes for unsteady transport problems are computationally expensive (Roache et al., 1986; Roache, 1997).

2.5.4 Non-dimensional form of the advection dispersion model and numerical properties

The AD-Model represents a combination of two dissimilar transport processes: advection and dispersion (or diffusion). Depending on the relative intensity of the transport processes, the solution can be influenced by either the diffusive transport or the advective transport. The one-dimension advection-dispersion model can be expressed in non-dimensional form by normalising the variables of the model relative to reference values. A reference length, x_r , which characterises the spatial dimension in which the pollutant is flowing, along with a reference concentration, φ_r , the concentration of the dispersant, a reference time, t_r , and with a reference velocity, v_r , can be defined. The concentration, length, velocity and time can be normalised relative to reference quantities to yield the following non-dimensional variables (Chin, 2013):

$$\varphi^* = \frac{\varphi}{\varphi_r}; x^* = \frac{x}{x_r}; t^* = \frac{t}{x_r/v_r}; v^* = \frac{v}{v_r}, \quad (2.19)$$

where the asterisk indicates that the variable is non-dimensional. Substituting these normalised

variables in the advection-dispersion model results in the following non-dimensional form of the model (Chin, 2013):

$$\frac{\partial \varphi^*}{\partial t^*} + v^* \frac{\partial \varphi^*}{\partial x^*} = \left(\frac{v_r x_r}{D_r} \right)^{-1} \frac{\partial^2 \varphi^*}{\partial x^{*2}}. \quad (2.20)$$

In this non-dimensional representation of the AD-Model, all non-dimensional variables are of the order of one, as they have been normalised by a reference value of the prevailing conditions. The only term whose magnitude is not necessarily of the order of one is the dispersion term, which is defined by the non-dimensional group, $v_r x_r / D_r$. The physical meaning of this non-dimensional group

is that it relates the advective flux, $v_r \varphi_r$, to the diffusive flux, $D_r \varphi_r / x_r$, and is called the Peclet number (Chin, 2013). These dimensionless numerical properties can be expressed based on grid discretisation. Thus, based on grid discretisation, the dispersion number (or diffusion number) is expressed as

$$d = \frac{D \Delta t}{(\Delta x)^2}. \quad (2.21)$$

The advection number (or Courant number) is expressed as

$$C = \frac{v \Delta t}{\Delta x}. \quad (2.22)$$

The Peclet number allows one to define the involvement of transport by advection and dispersion or diffusion. The diffusive concept has the dispersal effect; hence it poses smoothing properties inducing positive numerical results. For a numerical grid with the spatial discretization of Δx the Peclet number is expressed as

$$P_e = \frac{v \Delta x}{D}. \quad (2.23)$$

Numerical problems appear for high values of the Peclet number when the transport is dominated by advective flux. Numerical analysis shows that the dispersion (or diffusive) term of numerical schemes has positive effects on the numerical solution. The term has smoothing characteristics inducing positive numerical results. If the transport problem is diffusion dominated, minimal numerical difficulties are expected (Szymkiewicz, 2010). However, if in the numerical solution, the physical diffusion represented by the diffusion term is not strong it is unlikely to suppress oscillations or wiggles. In the case that numerical diffusion (or artificial dispersion) is introduced in the solution it can swamp the physical one, although it strengthens the smoothing effect resulting from the dispersion (or diffusion) term (Szymkiewicz, 2010). The difficulty with numerical diffusion (or artificial dispersion) is that it has the same effect as physical diffusion. Since the Peclet number is determined by model resolution, it is important that the properties of a scheme are considered when discretizing the computational domain.

2.5.5 Eulerian Solutions of the advection dispersion model

A general approach for advancing a numerical solution for the AD-Model over time employs a temporal weighting factor, ϕ , and is expressed as (Wallis, 2007; Versteeg and Malalasekera, 2007):

$$\frac{\phi^{l+1} - \phi^l}{\Delta t} + (1 - \phi) \frac{v \partial \phi}{\partial x} \Big|_l + \phi \frac{v \partial \phi}{\partial x} \Big|^{l+1} = (1 - \phi) \frac{D \partial^2 \phi}{\partial x^2} \Big|_l + \phi \frac{D \partial^2 \phi}{\partial x^2} \Big|^{l+1}, \quad (2.24)$$

where $0 \leq \phi \leq 1$, Δt is the time step, and the superscripts l and $l+1$ refer to the times at the start and end of the time step, respectively. Once finite difference approximations replace the spatial gradients, Equation (2.24) is used to evaluate the solute concentration at a time $l+1$ for all the nodes, assuming all nodal solute concentrations at the time l and all boundary conditions at the edges of the computational domain are known. When applying Equation (2.24) and $\phi = 0$ only transported variable values at the old time are used to evaluate one unknown concentration, resulting in an explicit calculation. If $\phi = 1$, transported variable values at the new time level are used; and if $\phi = 0.5$ transported variable values at both time levels are used. The latter two cases give a set

of simultaneous equations containing all the unknown concentrations, resulting in an implicit calculation. FV schemes can be expressed in a similar temporal weighting framework.

2.6 Estimating parameters of the advection dispersion model

The AD-Model can be applied through analytical solutions, numerical solutions and routing procedures. If the velocity and dispersion are known temporal or spatial concentration distribution at a downstream site can easily be predicted. However, the interest in this work lies in estimating these parameters from observed concentration profiles by applying parameter estimation tools and solutions of the model. Parameter estimation by methods based on the AD-Model involves optimisation methods which involve fitting a parameterised function to a set of measured data points by minimising the errors between the data points and the function. It may be a parameter-space-search (or manual) procedure, where the parameter estimates are manually adjusted during simulations to obtain the best fit between simulated and observed concentrations. Alternatively, a parameter estimation technique (automatic) can be used instead of the parameter-space-search approach (Chapra, 2008). The principle of least squares can be employed to find the best fit. The methods involve iterative improvements and use the derivative information to identify the optimal parameter set between the function and the measured data points (Chapra, 2008; Chin, 2013), as shown in Figure 2-3.

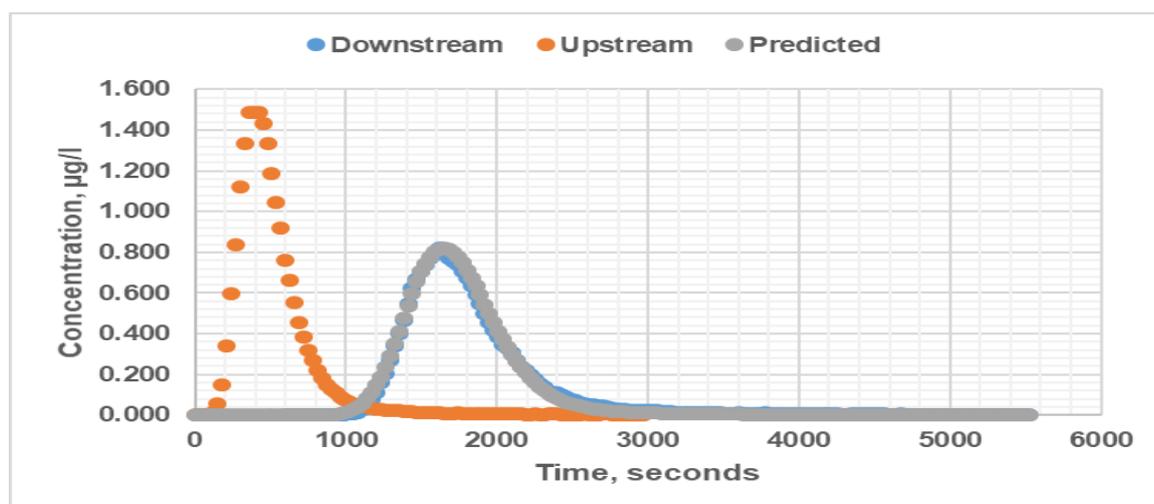


Figure 2-3: Fitting a simulated profile to a measured profile by optimising the velocity and dispersion coefficient (Wallis et al., 2013).

In the manual process, simulated and observed output responses are compared, and an incremental trial-and-error process of parameter adjustment is attempted to get the simulated response to approach more closely the observed response. In each trial, the values of the model parameters, chosen by the modeller, are substituted into the model, and a forward solution is carried out to give simulated values of the state variable. The fit between simulated and measured observations is then examined. If the fit is unacceptable, per modeller's judgement, parameters are adjusted manually repeatedly until a set is obtained which gives the best fit to the measured observations. The manual calibration process allows the modeller to develop a much deeper understanding of the data and models and their limitations. However, there are several drawbacks associated with manual inverse modelling. Firstly, apart from the expert or modeller's knowledge and judgement, there are no systematic rules followed in choosing which parameter to adjust. Secondly, for models with many parameters, the procedure can be difficult due to the multitudes of combinations of the parameters to be adjusted.

Thus, automatic calibration schemes have shown to produce improved calibrations. The advantages of automatic inverse modelling include the ability to get the optimal parameters quickly. Also, the process is systematic, and there is no subjectivity as in manual inverse modelling. Furthermore, there is a guarantee of finding optimal parameters, where a model is a true representation of the processes and observed data is accurate; and diagnostic statistics available at the end of the process are important for quantitatively judging the quality of the calibration. The major problem with traditional automatic calibration methods is their underlying assumption that the available model structure is correct, leading to the elusive goal of finding a unique optimal parameter set. Typically, the closeness of the model output to observations is measured by an objective function, as earlier stated, and there are typically large regions of feasible parameter space for which the objective function values are very similar.

A major weakness in automatic calibration is the dependence of the identification process on a single objective function that no matter how carefully chosen, is often inadequate to properly measure all the characteristics of the observed data deemed to be important. The dependence of automatic calibration methods on a single objective function contrasts with the manual calibration process,

which typically uses several complementary ways of evaluating model performance. In some cases, automatic calibration is used to fine tune a parameter set after manual calibration is complete. It should be recognised that any automatic calibration method, no matter how advanced, gives only the optimal solution concerning the objective function used in a specific case. There are several automatic inverse modelling algorithms in the literature, and it is important to understand the working of an inverse modelling algorithm to identify its capabilities (Gallagher and Doherty, 2007). Usually, inverse modelling methods are classified as either local or global methods (Blasone et al. 2007).

2.6.1 Global search parameter estimation

Global search algorithms initiate the search using many parameters (Blasone et al. 2007). The algorithm repeatedly evolves each parameter into a probable solution which gives the optimum objective function value. The process continues until a specified number of iterations have been attained or when the value of the relative change in optimum objective function between successive iterations is less than a user-specified value. An example of global search algorithms is the Shuffled Complex Evolution (Blasone et al., 2007).

Advantages of global search methods include (Semuwemba, 2011): (a) the final solution is often independent of initial values of the parameters; (b) the methods use a large set of parameters from which it can get the representative model parameter set; (c) they avoid the possibility of the objective function being trapped in areas of local minima.

However, global search methods have several disadvantages, which include (a) requirement of a considerable number of model's runs to give good estimates of the final model parameters (Skahill & Doherty 2006; Blasone et al. 2007). (b) incapable of detecting model insensitivity in the parameters and the observations as they do not compute derivatives (Matott & Rabideau 2008). (c) they do not give summary statistics that indicate the goodness of fit (Skahill and Doherty 2006), (d) the accuracy of the final estimates may be unreliable as the final value of the objective function may be false (Matott & Rabideau 2008).

2.6.2 Local-search parameter estimation

Local search methods begin from a single set of parameter values which they continue to adjust in subsequent runs. At the end of each run, the new parameter values are said to be an improvement of the previous values. The objective of the process is to find parameters in the parameter space which corresponds to the optimum objective function. The number of steps depends on among several factors, the model formulation, the initial parameter values, the calibration data and the local search algorithm used (Doherty 2008).

Depending on the method used to compute the next set of parameters, local search methods can either be classified as direct search or gradient search methods (Blasone 2007). Direct search methods derive the next parameter set by successively evaluating the value of the objective function near the initial parameter values. The main strength of direct search algorithms is that they do not require derivatives of model equations with respect to the parameters. Therefore, they are useful for models which are not differentiable or those that have discontinuous derivatives (Chapra and Canale, 2015). However, the initial estimate of the parameters should be as close to the true parameter values as possible. Examples of direct search methods are the random search method and the univariate search method (Chapra and Canale, 2008) among others.

Gradient-based methods use derivatives of the objective function and the derivatives of the model equations concerning the parameters to locate optima. Because gradient-based methods depend on the derivative computations of the objective function, must be differentiable with respect to the parameters and the derivatives must be continuous over the parameter space (Van den Bos, 2007).

The main disadvantage of all local search methods is that they can be trapped in areas of local optima, which may be of varying sizes. To circumvent this problem requires an expert assessment of the model so that the starting set of model parameters are close to the range expected for the modelled system (see Section 4.4.2).

However, the inverse modelling methodology used in interpreting tracer BTCs using the AD-Model in most studies has been based largely on the local-search approach (e.g. Wallis et al., 2013; Manson and Wallis, 2013, Doherty, 194).

2.6.3 Parameter estimation by gradient-based methods

A nonlinear least-square inverse optimisation problem can be stated as follows (Rao, 2009):

$$\text{Find } \mathbf{C} = \begin{Bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{Bmatrix} \text{ which minimises } f(\theta), \quad (2.25)$$

subject to constraints:

$$f_j(\theta) \leq 0 \quad j = 1, 2, \dots, k, \quad (2.26)$$

$$h_j(\theta) = 0 \quad j = 1, 2, \dots, m, \quad (2.27)$$

where θ is a parameter value, \mathbf{C} is an n dimensional parameter vector, $f(\theta)$ is the objective function, $f_j(\theta)$ and $h_j(\theta)$ are inequality and equality constraints respectively. Such an optimisation problem is termed as a constrained optimisation problem (Rao, 2009). The process of optimisation involves minimising the objective function and that the objective function is twice differentiable. The basic characteristic of an optimum point is that it is a stationary point of the objective function such that the gradient vector of the objective function is zero (van den Bos, 2007; Rao, 2009). Thus, in the case where $f(\theta)$ is a function of the elements of $\theta_j = (\theta_1, \theta_2, \dots, \theta_k)^T$, then θ_* is a stationary point if

$$\left. \frac{f}{\partial \theta} \right|_{\theta=\theta_*} = \begin{bmatrix} \frac{\partial f}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} \\ \vdots \\ \frac{\partial f}{\partial \theta_k} \end{bmatrix} = \mathbf{0}, \quad (2.28)$$

where, $\mathbf{0}$ is a $k \times 1$ null vector.

Stationarity is a necessary condition for a point to be optimum. An essential condition for a stationary point to be minimum is that the Hessian matrix must be positive definite at that point (van den Bos, 2007), that is,

$$\left. \frac{\partial f}{\partial \theta \partial \theta^T} \right|_{\theta=\theta_0} = \begin{bmatrix} \frac{\partial^2 f}{\partial \theta_1^2} & \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} & \dots & \frac{\partial^2 f}{\partial \theta_1 \partial \theta_k} \\ \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 f}{\partial \theta_2^2} & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial \theta_1 \partial \theta_k} & \dots & \dots & \frac{\partial^2 f}{\partial \theta_k^2} \end{bmatrix} \succ \mathbf{0}, \quad (2.29)$$

where, $\mathbf{0}$ is $\kappa \times \kappa$ null matrix.

Several methods can be used to solve the parameter identification problem including the steepest descent method, the Newton method, the Gauss-Newton method and the Levenberg-Marquardt method. The steepest descent and the Newton methods are general numerical minimisation methods of the objective function (van den Bos, 2007). The Newton method is obtained from the quadratic Taylor polynomial. The computational effort involves both the gradient and the Hessian matrix of the objective function. Also, the method needs an evaluation of eigenvalues and solution of a system of linear equations of the step vector. The Gauss-Newton and the Levenberg-Marquardt including other variants such as the generalised Gauss-Newton method are like the Newton method as they use an estimation of the Hessian matrix (van den Bos, 2007). The steepest descent is obtained from the linear Taylor polynomial of the objective function. The computational effort is moderate as it only comprises computation of the gradient only. Except for the steepest descent method, other methods require a routine for the minimisation solution, which must be supplied by the experimenter (van den Bos, 2007).

2.6.4 Routing procedures

Parameter estimation for the AD-Model using an analytical solution has the disadvantage that it requires knowledge of the area of the channel and the mass of the solute. When concentration-time data is available, analytical solutions cannot be used directly to estimate parameters. They are applied through routing procedures, and the common ones are the Fischer's routing procedure (Fischer et al., 1979), the Hayami solution (Rutherford, 1994) and the Singh and Beck routing procedure (Singh and Beck, 2003). The routing method is the process by which a temporal tracer profile at a downstream site is predicted from an upstream temporal tracer profile by a response

function which is a solution of the AD-Model governing equation for unit concentration (Rutherford, 1994; Chin, 2013; Wallis et al., 2013). The process can be used for either a discrete or a continuous function of upstream concentration. If the upstream source is transient, then the temporal concentrations at a downstream section are determined by the following expression:

$$\varphi(X, t) = \int_0^t \varphi(\tau) \xi(x, t - \tau) d\tau, \quad (2.30)$$

where $\varphi(\tau)$ is concentration upstream, $\varphi(X, t)$ is concentration downstream and $\xi(x, t)$ is the response function. As well as temporal routing, spatial routing can be employed in which a spatial concentration profile is routed downstream. Since only discrete data is available, numerical integration approximation is used in both approaches (Fischer, 1967; Singh and Beck, 2003; Wallis, 2013). In principle the spatial routing and the temporal routing approaches are similar, and the only difference is that one takes place in the space domain and the other takes place in the time domain.

2.6.4.1 Routing a spatially varying source

Fischer (1967) proposed the first routing procedure based on spatial routing, i.e. it starts with an initial condition of a spatial concentration profile. However, spatial profiles are practically difficult to observe. Owing to observations being made in the time domain, Fischer (1967) introduced the idea of the frozen cloud assumption to apply spatial routing to concentration-time data. The frozen-cloud assumption considers the centre of gravity of the cloud is moving with the velocity of the stream flow. The frozen cloud approach is used twice: firstly, the upstream temporal profile is converted into an equivalent spatial profile which is then routed downstream and, secondly the predicted spatial profile is converted to an equivalent downstream temporal profile. Since Fischer had discrete rather than continuous data he undertook the spatial routing using a numerical integration approximation. In Fischer's routing method, the errors introduced into estimates of dispersion coefficients may be a result of the frozen-cloud assumption and the numerical integration (Singh and Beck, 2003).

2.6.4.2 Routing a temporally varying source

As discussed above, it is convenient and customary to measure a temporal concentration distribution at a location. Predicting a temporally varying source uses a convolution integral and takes place

entirely in the time domain such as that proposed by Singh and Beck (2003) and in the Hayami solution presented by Barnett (1983). These methods do not require the frozen-cloud approximation and deal strictly with boundary conditions. Unlike Fischer's routing procedure, these methods directly predict a downstream temporal concentration distribution from an upstream temporal concentration distribution. The upstream concentration profile is discrete, rather than continuous, and transient. The advection-dispersion model is a linear equation, and the solutions of this equation can be added for transient sources. Thus, response temporal profiles from transient sources can be added together to obtain the combined effect (Rutherford, 1994; McCuen, 1998). By applying the process of convolution, the mean concentration in each time segment of the upstream concentration profile results in a temporal concentration profile at another point downstream. The process is repeatedly applied to every segment, and the separate concentration profiles are added by applying the principle of superposition to obtain the combined effect (Rutherford, 1994).

2.6.4.3 Frozen cloud approximation

As discussed above, temporal profiles are usually observed because spatial profiles are practically difficult to observe. The temporal profiles are always skewed as longitudinal dispersion continues to take place as a cloud of solute passes a monitoring site resulting in asymmetric profiles (Rutherford, 1994; Wallis, 2005). The frozen cloud approximation disregards the longitudinal dispersion that takes place during the time the solute passes a monitoring site (Rutherford, 1994). As explained earlier applying the frozen cloud approximation method involves converting temporal concentration profiles to spatial concentration profiles and then routing the spatial concentration distribution downstream. The predicted spatial profile is then converted back to a temporal profile. Where a temporal concentration profile has been observed at Site 1 downstream an injection point and requires predictions of a temporal concentration profile at Site 2 downstream, the procedure is summarised as follows (Rutherford, 1994; Fischer et al., 1979):

$$\varphi(x_2, t) = \int_{-\infty}^{\infty} \frac{\varphi(x_1, \tau) v}{\sqrt{4\pi D(t_2 - t_1)}} \exp\left(-\frac{v^2 (t_2 - t_1 - t + \tau)^2}{4D(t_2 - t_1)}\right) d\tau, \quad (2.31)$$

where $\varphi(x_1, t)$ is the observed temporal concentrations profile at Site 1, $\varphi(x_2, t)$ is the predicted

temporal concentrations profile at Site 2, t_1 and t_2 are travel times of centroid of profiles at Sites 1 and 2 respectively, v is the flow velocity, τ is a time increment and D is the longitudinal dispersion coefficient. Where a temporal concentration profile has been observed at Site 2, it can be compared with the predicted profile, and if necessary, the coefficients in the routing equation adjusted to obtain the best fit. Adjusting the values of the coefficients to obtain the best fit between the observed profile and the routed profile is a calibration (or parameter estimation) process. Velocity is estimated by matching the timing of the peak concentrations of the observed and routed profiles while dispersion coefficient is estimated by matching the peak concentrations and the breadth of the routed profile to that of the observed profile (Rutherford, 1994; Singh and Beck, 2003; Wallis, 2005).

2.6.4.4 Fischer's routing procedure

The method applies a frozen cloud approach, and the complete Fischer's procedure of routing an upstream temporal concentration profile to a temporal downstream profile is given by Equation (2.30). This equation is usually integrated numerically (Rutherford, 1994) as:

$$\varphi(x_2, t) = \sum_{j=1}^n \frac{\varphi(x_1, l\Delta t) v}{\sqrt{4\pi D(t_2 - t_1)}} \exp\left[-\frac{v^2(t_2 - t_1 - l\Delta t)^2}{4D(t_2 - t_1)}\right] \Delta t, \quad (2.32)$$

where $\varphi(x_1, l\Delta t)$ is the observed temporal concentration upstream, l is the index of time increment Δt , $\varphi(x_2, t)$ is predicted temporal concentration profile downstream, t_2 and t_1 are mean times of travel of centroids between the sites and v is cross-sectional average reach velocity. Rutherford (1994) discussed conditions that are required when applying this method. The location of Site 1 must be in the equilibrium zone. The entire temporal concentration profile must be captured at both Site 1 and Site 2 with a negligible loss between the sites otherwise a correction must be made before optimisation. The values of velocity and dispersion coefficient that result in the best fit of predicted and observed temporal concentrations at the downstream Site 2 are considered as the mean values of velocity and dispersion for the reach.

2.6.4.5 Hayami's routing procedure

The Hayami solution can be applied to route an observed upstream temporal concentration profile downstream directly without requiring the frozen cloud approximation (Rutherford, 1994; Chanson, 2004). The Hayami solution applied to a discretised upstream temporal profile results in a temporal profile at the downstream site is given by (Rutherford, 1994):

$$\varphi(x_2, t) = \int_{-\infty}^{\infty} \frac{\varphi(x_1, \tau)(x_2 - x_1)}{(t - \tau)\sqrt{4\pi D(t - \tau)}} \exp\left[-\frac{(x_2 - x_1 - v(t - \tau))^2}{4D(t - \tau)}\right] d\tau. \quad (2.33)$$

Similarly, this equation is integrated numerically. The values of advection and dispersion coefficient can be varied to get the best fit which gives the mean values of the reach. It also requires complete temporal profiles with a negligible loss between the sites (Rutherford, 1994).

2.6.4.6 Singh and Beck's routing procedure

The routing procedure derived by Singh and Beck (2003) applies the solution of the one-dimensional advective dispersion equation satisfying a spatially semi-infinite volume source. Assuming constant velocity and dispersion and satisfying the boundary conditions for the solution given by Equation (2.15) the procedure is (Singh and Beck, 2003; Chapra, 2008):

$$\varphi(x, t) = \frac{\varphi_0}{2} \left[\operatorname{erfc}\left(\frac{x - vt}{2\sqrt{Dt}}\right) + \exp\left(\frac{vx}{D}\right) \operatorname{erfc}\left(\frac{x + vt}{2\sqrt{Dt}}\right) \right]. \quad (2.34)$$

Routing procedures are formulated for different initial and boundary conditions. Consequently, it may not be correct to apply the procedures to the same data unless the conditions of data satisfy the basis on which these methods are formulated.

2.6.5 Concentration prediction by numerical methods

Concentration prediction using numerical methods of the AD-Model is analogous to the previous one except that a numerical solution to the AD-Model is used instead of a routing procedure. Starting with known initial concentration values, concentrations at later times are evaluated using either an implicit or an explicit time-marching approach (Wallis, 2007b). The calculations are also dependent

on boundary conditions at the spatial edges of the computational domain. All numerical solutions suffer from a range of errors, which are determined by the properties of the scheme used. Consequently, when such a model is optimised to observed concentration data, the estimated parameter values are likely to be influenced by the formulation of the numerical solution, the choice of time marching and the numerical grid (Wallis, 2007).

The procedure used for predicting temporal concentration distribution at a downstream site is a numerical convolution solution method (Abbott and Basco, 1989). The continuum form for the solution for the one-dimensional equation of fluid flow of the parabolic partial differential equation of linear operator form is expressed as (Abbott and Basco, 1989):

$$\varphi(x, t_0 + \tau) = \int_{\varepsilon=-\infty}^{\varepsilon=\infty} \psi(\varepsilon) \cdot \varphi(x + \varepsilon, t_0) d\varepsilon, \quad (2.35)$$

where:

$\psi(\varepsilon)$ is a weighting function depending on the integration method used; τ is a time increment; ε is a space increment, $\varphi(x, t_0)$ are specified initial conditions, and $\varphi(x, t_0 + \tau)$ is the required solution after the specified time step. With suitable boundary conditions, repeated application of Equation (2.33) is needed for the solution of marching type to find the solution function $\varphi(x, T)$ at a time $T = L\tau$. This equation applies directly to initial-value problems. However, it can be generalised to mixed initial-value and boundary-value problems by use of a suitable weighing function (Abbott and Basco, 1989).

Like the routing procedure, if a temporal concentration profile has been observed at the downstream point, it can be compared with the predicted profile, and if necessary, the parameters in the numerical model can be adjusted to obtain the best fit, thereby obtaining optimal parameter values for the stream reach.

2.7 Empirical equations for estimating the dispersion coefficient.

2.7.1 Introduction

A successful application of the AD-Model for prediction of contaminant transport in a stream depends on reliable parameter values. The important and difficult task is the determination of the longitudinal dispersion coefficient (Wallis and Manson, 2004). Since experimental data is not easy to obtain a common method for estimating solute transport parameters to be used with the AD-Model is by use of simplified models in the form of empirical equations. Empirical equations are based on correlating the results from other methods, especially experimental methods, against pertinent bulk hydraulic variables (Rutherford, 1994; Seo and Cheong, 1998; Wallis and Manson, 2004). Thus, using estimated solute transport parameters from experiments and measured parameters such as streamflow and channel characteristics an empirical equation could be developed for predicting solute transport parameters that can be used with the AD-Model (Fischer, 1979; Seo and Cheong, 1998; Ani et al., 2009). Several of the methods are based on the flow integration method (Fischer, 1967) and have been developed by regression analysis (Fischer et al. 1979; Seo & Cheong 1998; Falconer et al. 2002; Wallis & Manson 2004). Considerable reviews of these predictive models are presented in Seo and Cheong (1998), Kashefipour and Falconer (2002) and Wallis and Manson (2004).

Several empirical methods have been developed ranging from those that correlate one bulk characteristic such as discharge (e.g. Ani et al., 2009) to those correlating longitudinal dispersion to several bulk characteristics (Fischer, 1975; Seo and Cheong, 1998). Typically, bulk parameters considered include cross-sectional average velocity, channel width, shear velocity, flow rate and mean flow depth/hydraulic radius) (Fischer et al., 1979; Seo and Cheong, 1998; Kashefipour and Falconer, 2002). Majority of the methods have used the dimensional analysis approach to develop empirical models which have arrived at different combinations of bulk flow characteristics that lead to the same dimensions of the longitudinal dispersion (Kashefipour and Falconer, 2002).

2.7.2 Previous works

The majority of studies relating dispersion coefficient to bulk flow parameters are based on the work by Fischer (1967). The dependence of the dispersion coefficient on differential lengthwise advection

and cross-sectional mixing can be defined by the following expression (Wallis and Manson, 2004):

$$D = -\frac{1}{A} \int_A U' (y, h) f (y, h) dA, \quad (2.36)$$

where A is the cross-sectional of the channel, U' is the local spatial deviation of the lengthwise velocity from the cross-sectional mean lengthwise velocity, f is a function that defines the cross-sectional distribution of lateral and vertical mixing coefficients, y is the lateral space coordinate and h is the vertical space coordinate. Given data on the values and the distribution of the lengthwise velocity and the mixing coefficients can be used to determine D . Given only the transversal distribution of depth-averaged lengthwise velocity and depth-averaged transversal mixing coefficient the determination of D from Equation (2.36) is simplified (Wallis and Manson, 2004).

Following the work of Taylor (1954) and Fischer (1967), it has generally been believed that transversal differential lengthwise advection and transversal mixing govern longitudinal dispersion in natural channels for having large width-to-depth ratios, in comparison to the vertical variations of velocity. Based on that belief, Fischer (1967) developed a triple integral equation for estimating longitudinal dispersion coefficient in natural streams. The equation is expressed as follows (Seo and Cheong, 1998; Wallis and Manson, 2004; Kashefipour and Falconer):

$$D = -\frac{1}{A} \int_0^w \zeta u' \int_0^z \frac{1}{\varepsilon_z z} \int_0^z \zeta u' d\zeta d\zeta d\zeta, \quad (2.37)$$

where w is the top width of the channel, ε_z is the local transverse mixing coefficient, ζ is the local depth of flow, $u' = u(\zeta)$ is the local departure of the depth-averaged longitudinal velocity from the cross-sectional mean longitudinal velocity, u is the local depth-averaged lengthwise velocity, and the other symbols are as previously described. In practice, the integrals of Equation (2.37) are replaced by summations, and hence, substantial tracer data are required in the transverse and longitudinal directions of flow (Kashefipour and Falconer, 2002). Fischer (1967) later suggested a simpler theoretical form of the triple integral expressed as (Wallis and Manson, 2004)

$$D = \frac{\overline{Iu'^2}L^2}{\varepsilon_t}, \quad (2.38)$$

where I is the non-dimensional triple integral accounting for the interaction of the local cross-sectional mixing and the local flow structure, $\overline{u'^2}$ is the average squared value of the cross-sectional velocity deviation, L is the characteristic length upon which velocity deviation takes place and ε_t is the cross-sectional transverse mixing coefficient. Using this method for estimating the longitudinal dispersion coefficient to be used with the AD-Model depends on the availability of measured data. The triple integral approach has the strongest justification as it directly takes account of the physical processes that affect dispersion, but it is difficult to use the equation as the variables that the method requires cannot be measured easily (Seo and Cheong, 1998; Wallis and Manson, 2004). Fischer (1979) developed a simpler equation by introducing an approximation of the triple integration, the transverse dispersion coefficient and the velocity deviation. This yielded the following simplified equation:

$$D = 0.011 \frac{v^2 w^2}{HU_*}, \quad (2.39)$$

where H is the average flow depth and U_* is the shear velocity. Fischer's simplified equation has the benefit of simplicity as it can be used to predict longitudinal dispersion coefficient by using bulk cross-sectional flow and channel parameters. Expressing the simple form of Fischer's equation in the non-dimensional form using an approximation of the triple integral is written as (Wallis and Manson, 2004; Fischer et al., 1979),

$$\frac{D}{HU_*} = 0.011 \left(\frac{v}{U_*} \right)^2 \left(\frac{w}{H} \right)^2. \quad (2.40)$$

The non-dimensional form of the Equation (2.40) shows that the longitudinal dispersion coefficient is a function of measures of channel aspect ratio and channel friction. Many studies seem to approve the non-dimensional form of Fischer's equation, despite differences in some features (Wallis and

Manson, 2004). A look at later equations shows that some do not relate the non-dimension dispersion coefficient to aspect ratio and channel friction. Also, the exponents and numerical constants vary widely (Wallis and Manson, 2004). Most likely, the basic reason for the differences in equations is due to differences in relative strengths of various mechanisms that control dispersion. Thus, the observed dispersion coefficients used for calibrating the equations depend on channel characteristics. The other causes of differences could be variability of methods for estimating dispersion coefficients and differences in regression analyses used to develop the equations (Wallis and Manson, 2004). Following Fischer's work, several methods have been developed (Seo and Cheong, 1998; Seyed M Kashfipour and Falconer, 2002; Deng and Jung, 2009). Jain (1976) proposed an equation expressed as (Wallis and Manson, 2004)

$$D = \frac{I_j v^2 w^3}{\kappa_t A U_*}, \quad (2.41)$$

where I_j is a dimensionless triple integral, κ_t = cross-sectional average dimensionless transverse mixing coefficient (commonly expressed as $\kappa_t = \varepsilon_t / H U_*$). In non-dimensional form, the equation is expressed as

$$\frac{D}{H U_*} = \frac{I_j}{\kappa_t} \left(\frac{v}{U_*} \right)^2 \left(\frac{w}{H} \right)^2, \quad (2.42)$$

where, I_j is a dimensionless triple integral expressed in terms of local flow depth of the channel. The method only requires the shape of the channel to calculate the triple integral. Hence, the method applies to channels that are uniform and straight.

Liu (1977) obtained a similar theoretical equation which is expressed as

$$D = \frac{I_L v^2 w^2}{A U_*}, \quad (2.43)$$

where, I_L is a dimensionless triple integral, being a function of local depth of flow, local depth average velocity and a quasi-local transverse mixing coefficient. Observed longitudinal dispersion coefficients were used to calculate the triple integral which was correlated to the channel friction based on the idea that channel characteristic that enhances dispersion also increase resistance to

flow. Liu (1977) obtained the following expression for evaluation of the triple integral:

$$I_L = 0.18 \left(\frac{U_*}{v} \right)^{1.5}. \quad (2.44)$$

Combining Equations (2.41) and (2.42), the dimensionless expression obtained by Liu (1977) is written as

$$\frac{D}{HU_*} = 0.18 \left(\frac{v}{U_*} \right)^{0.5} \left(\frac{w}{H} \right)^2. \quad (2.45)$$

The equation was observed to be better than the Fischer's equation. However, the results were considered to be unreliable based on the method used for evaluating the triple integral (Wallis and Manson, 2004).

Like Equation (2.38) the use of Equation (2.41) and (2.43) have drawbacks because of challenges of evaluating detailed velocity and channel profiles and transverse dispersion coefficient (Martin et al., 2013). Consequently, in practical engineering studies, it has become preferable to estimate D using equations which are based on channel and flow characteristics, which can be easily obtained quantified, such as Equation (2.45) (Seo and Cheong, 1998).

Seo and Cheong (1998) used dimension analysis and the one-step Huber nonlinear regression method to correlate non-dimensional observed longitudinal dispersion coefficients to non-dimensional channel friction and aspect ratio. Longitudinal dispersion coefficients for developing the equation were estimated using a routing procedure (Seo and Cheong, 1998). In dimensionless form, their equation is expressed as

$$\frac{D}{HU_*} = 5.915 \left(\frac{v}{U_*} \right)^{1.428} \left(\frac{w}{H} \right)^{0.620}. \quad (2.46)$$

The model was developed based on 59 data sets measured in 26 streams in the United States (Seo and Cheong, 1998). The method was compared with three existing equations, namely, McQuivey and Keefer, Liu, and Iwasa and Aya (Seo and Cheong, 1998; Wallis and Manson, 2004). The

equation was observed to predict dispersion coefficient better than the three equations (Seo and Cheong, 1998).

Kashefipour and Falconer (2002) undertook an investigation based on dimension analysis and least-squares regression to relate observed longitudinal dispersion coefficients to non-dimensional groups using 81 data sets measured in 30 rivers in the United States. Initially, they inferred that the most appropriate equation is expressed as (Kashefipour and Falconer, 2002; Wallis and Manson, 2004)

$$\frac{D}{Hv} = 10.612 \left(\frac{v}{U_*} \right). \quad (2.47)$$

Using several inferential statistics and comparing with several existing empirical models, they found that the Equation (2.47) and Seo and Cheong's (1998) equations were more accurate than other predictors. However, Seo and Cheong's equation was found to overestimate while Equation (2.47) was observed to underpredict. Consequently, a linear combination was suggested (Kashefipour and Falconer, 2002) which resulted in the following equation:

$$\frac{D}{HU_*} = \left[7.428 + 1.775 \left(\frac{v}{U_*} \right)^{-0.572} \left(\frac{w}{H} \right)^{0.620} \right] \left(\frac{v}{U_*} \right)^2. \quad (2.48)$$

As a further clarification, they observed that Equation (2.48) was more appropriate than Equation (2.47) for stream flows with aspect ratio values greater than 50, Equation (2.47) was observed to predict more acutely than the other Equation (2.48) when aspect ratio values were less than 50 (Kashefipour and Falconer, 2002).

In summary theoretical methods have problems with evaluation of the triple integral (Martin et al. 2013). For example, the use of Fisher's theoretical method was inconsistent as the triple integral was constant instead of being flow dependent (Wallis and Manson, 2004). Liu's equation in correlating with channel friction resulted in values of the triple integral decreased with increased channel flow rate. Jain's equation showed uncertainties in the values of his triple integral as the order of the values was rather excessive (Wallis and Manson, 2004).

However, regressing even well-grounded empirical models to observed dispersion coefficients obtained from many channels requires careful thought due to a variety of channel types, amount of lengthwise nonuniformity and flow conditions under which observed dispersion coefficients were collected. Otherwise, there is a possibility of developing a single and general equation for use in all cases without recognising that several equations, each suited to particular conditions, may be better capable for a range of flow situations (Wallis and Manson, 2004; Ani et al., 2009).

2.7.3 Model construction process

The traditional process of constructing a model involves data collection, model conceptualisation and specification, model calibration and model evaluation (Sun and Sun, 2015). In this context data includes estimated parameter values of the AD-Model and flow characteristics that influence longitudinal dispersion in streams. Model conceptualisation and specification involved determining the functional form and the structure of the model. Model calibration involves applying the developed model to model building data to make predictions. Lastly, model confirmation or evaluation involves using the developed model to estimate dispersion coefficients to be used with the numerical methods of AD-Model to predict concentration evolution at a downstream site of the model evaluation data.

2.7.4 Statistical analysis of observational data

To make predictions based on the developed empirical equations requires an understanding of the principles of both probability and statistics (Chin, 2013). Estimated values of water quality parameters are stochastic variables that can only be characterised by probability distributions because of their natural variability. Probability distributions characterising the variables are the basis of the methods used for statistical analysis (Montgomery, 2003; Chin, 2013).

Since the true probability distribution of estimated dispersion coefficients is not known, the appropriate probability distribution can be estimated based on observed values. The use of observed data to estimate the stochastic properties of random variables is covered by the field of statistics (Chin, 2013). Several basic probability distributions that are used in the analysis of water quality observational data are derived probability distributions. Additionally, derived probability distributions are typically used in describing the probability distribution of statistics computed from observational

data. Most of these derived probability distributions describe the behaviour of random variables that are assumed to be normally distributed (Chin, 2013). Derived probability distributions are also used in hypothesis testing whether observed data support the assertion that they are drawn from a population distribution. Hypothesis testing methods identify a statistic that measures the difference between the sample distribution and the proposed population distribution and then determining the significance level of this statistic (Chin, 2013; Hanusz et al., 2016).

The common tests for normality involve normal scores (quantile-quantile; Q-Q) plot followed by a normality hypothesis test. Typical tests include the Shapiro-Wilk (Ryan and Joiner, 1976; Analyse-it Software Ltd., 2009), the Shapiro-Francia test and the Anderson-Darling test. The Shapiro-Francia test is a modification of the Shapiro-Wilk test that is recommended when observational data is more than 50. The Shapiro-Wilk test has been found to be a useful test that detects most departures from normality when the sample size is not more than 5000 (Ryan and Joiner, 1976; Razali and Wah, 2011; Shapiro and Wilk, 1965). The test is especially useful for detecting departures from normality in the tails of sample distributions (Chin, 2013). The test relates Shapiro-Wilk statistic, sample size and confidence level (Hanusz et al., 2016).

2.7.5 Empirical model identification

2.7.5.1 Selecting model bulk parameters

Primarily, longitudinal dispersion in streams is caused by differential longitudinal advection and turbulent diffusion. Differential longitudinal advection and turbulent diffusion are influenced by several factors which can be termed as bulk channel characteristics. Bulk characteristics can be classified as fluid properties, hydraulic characteristics and geometric forms (Rutherford, 1994; Falconer, Kashefipour and Falconer, 2002; Seo and Baek, 2004). These features may enhance turbulence, trap the dispersant and release it later. Hydraulic factors that influence dispersion include channel width, channel depth, shear velocity, flow velocity and channel flow rate. Geometric characteristics include bedforms such as dead zones, and channel sinuosity (Rutherford, 1994; Seo and Cheong, 1998; Kashefipour and Falconer, 2002).

Longitudinal advective dispersion is the lengthwise stretching of matter caused by velocity shear

(Chanson, 2004). The rate of longitudinal advective dispersion is a result of the balance between differential advection which tends to spread the dispersant longitudinally and transverse mixing which enhances uniform concentrations across the channel and consequently contracts the effects of differential advection (Rutherford, 1994). Channel width and depth influence longitudinal dispersion due to cross-sectional mixing which is caused by secondary currents and turbulent transversal diffusion. In channels where the rate of cross-sectional mixing is high, longitudinal dispersion is reduced (Rutherford, 1994). Additionally, an increase in shear velocity increases turbulent diffusion and cross-sectional mixing but reduces flow velocity thereby reduces longitudinal dispersion (Chanson, 2004; Chin, 2013).

The rate of longitudinal dispersion has been observed to increase with flow rate (Rutherford, 1994). The influence of flow on dispersion depends on the combined influence of flow on the rate of transverse mixing and differential longitudinal advection (Rutherford, 1994). An increase in flow rate results in more uniform flow and reduced bed friction, which may cause a reduction in longitudinal dispersion. However, if there is an increase in longitudinal differential advection and a reduced transverse mixing, the rate of longitudinal is expected to increase (Rutherford, 1994). Commonly, when flows and depth of flow are low, the rate of longitudinal dispersion tends to increase with flow rate. This could be a result of the effects of channel irregularities (Boxall and Guymer, 2007). The variation of longitudinal dispersion with features of a channel implies that it is possible to generalise the variation of longitudinal dispersion with discharge in a given reach of a stream (Rutherford, 1994).

Sinuosity increases the rate of transverse mixing as the flow constantly changes direction. Therefore, an increase in transverse mixing reduces the effect on shear flow dispersion processes (Rutherford, 1994). The channel form is characterised by the number of irregularities and riparian vegetation. These irregularities result in turbulence being inhomogeneous and reduce cross-sectional mixing. Therefore, the dispersion is expected to increase with an increase in irregularities. Flow complexities such as sinuosity and shape factor represent characteristics that are not easily measured in natural channels, and the impact of these characteristics can be included in the friction term (Seo and Cheong, 1998).

2.7.5.2 Identification of model structure

The common method of identification of model uses dimensional analysis and regression analysis (Seo and Cheong, 1998; Chapra and Canale, 2008). However, there are two typical types of engineering applications that are encountered in the regression analysis, namely. trend analysis and hypothesis testing.

Trend analysis is the process of using the pattern of the data to make predictions. If observed data is accurate, interpolation is commonly used, and if observed data is inaccurate, least-squares regression is commonly used. Least-squares requires additional information from the field of probability and statistics (Chapra and Canale, 2008).

Hypothesis testing is when an existing model is compared with observational data. If the parameters of an existing model are unknown, it may be required to determine the parameters that best fit the observational data. Alternatively, if the model parameters are known, it may be apt to assess the adequacy of the model. Commonly, optional models are assessed, and the most appropriate model is selected (Chapra and Canale, 2008).

A regression analysis aims to establish an equation, including its parameters, that adequately define the relationship between variables (Montgomery, 2003; Chapra and Canale, 2008). One approach to a regression analysis starts with the selection of the functional form of the equation to be matched to observed data (Chin, 2013). This is followed by adjusting the parameters until the residual sum of the differences between the observed data and those assumed by the equation is minimised. The drawback of this approach is that there could be an infinite number of functions that might fit the data. An alternative approach is to transform data until the relation between the variables is linear. After that linear regression can be used to determine the relationship between variables (Montgomery, 2003; Chapra and Canale, 2008; Chin, 2013). Model identification is followed by calibration which uses a set of observational data to identify or estimate the unknown model parameter (Sun and Sun, 2015).

2.7.6 Model Performance analysis

Calibration aims to adjust the parameters so that prediction optimally fits observations. Model performance is assessed by the quality of model calibration (Chapra, 2008). For effective use of models for management purposes, it is important to demonstrate a level of assurance in their performance (Bennett et al., 2013). Environmental models commonly comprise various definite characteristics. If performance analysis is based on one criterion, only certain features of performance are assessed (Bennett et al., 2013). Such a practice may lead to a preference of models that may not produce important properties of the system. Therefore, a system of measures addressing several characteristics may be required for a broad assessment of a model (Toprak and Cigizoglu, 2008). Generally, the performance of a model is assessed by the level of agreement between predictions and observations. Typically, a model structure will depend on its objectives. Therefore, the appropriate type of performance evaluation will depend on the model objectives (Bennett et al., 2013).

Several workers recommend that a combination of methods should be used to characterise the performance of models (e.g. Toprak and Cigizoglu, 2008; Chin, 2013). According to Toprak et al. (2004), models should be assessed using several descriptive statistics, such as maximum and minimum values of variables, standard deviation, variance, correlation coefficients (or coefficient of determination) etc., and several error statistics, such as the mean square error, the standard error, normalized error. Chin (2013) presents several methods for performance analyses, viz. graphical techniques, error statistics, residual error analysis, hypothesis testing, and linear regression. Statistical criteria are specifically appropriate if a model output variable is normally distributed, and the level of agreement between the observed and the model output can be assessed using the traditional measures such as the standard deviation and the variance (Chin, 2013). In hypothesis testing, the null hypothesis indicates similarity between observed and predicted values. Linear regression requires that the slope, the intercept and the correlation coefficient from the regression line be calculated together before making any inference (Chin, 2013). Residual error analyses include the coefficient of determination, model efficiency, residual analysis etc.

The analysis of residuals is usually useful in examining the assumption that residuals are normally

distributed with constant variance, and in establishing whether the model would require more terms. A plot of residuals against predicted values is the main diagnostic tool to check normality of residuals (Montgomery and Runger, 2003). It can be achieved by use of a statistical computer program such as Analyse-it for Microsoft Excel (2009).

2.7.7 Model confirmation

The interest in confirming models is to have reliable models for predicting mass transport in streams but also to use reliable parameters and to make the models to apply to other similar case studies (Ani et al., 2009). Once a developed model has been calibrated, all the modeller knows is that the model fits a single set of observations. However, before applying the model for management predictions with confidence, the model must be confirmed. This involves running the calibrated model for a new set of observations, with the stimuli changed to reflect the new conditions (Chapra, 2008). Contrary to calibration, the model parameters should now be kept unchanged at the values determined during calibration. If the model predictions fit the new set of observations, the model has been confirmed as an effective prediction model for the range of conditions determined by the calibration and the evaluated data sets. If there is no match, the model can always be investigated to establish possible reasons for discrepancies. This may lead to further mechanism characterisation and model refinement (Chapra, 2008; Chin, 2013). The simulation may not precisely match the observational confirmation dataset even if the model is considered evaluated. Thus, some workers fine-tune the model after evaluation so that the model optimally fits both the calibration and the evaluation datasets (Chapra, 2008).

However, model confirmation may require model performance assessment criteria, which stipulate an acceptable level of model performance (Bennett et al., 2013; Moriasi et al., 2015). Performance evaluation criteria may be predetermined values that an end-user has determined for decision making or based on comparison with other models (Arhonditsis et al., 2006). Some workers have attempted to relate expert judgement and quantitative performance metrics to model performance criteria (Bennett et al., 2013). Every modelling undertaking has distinctive aims and challenges. Therefore, even if it is possible to prescribe some general elements of performance measures that are worthwhile in modelling, there are no standard performance evaluation criteria that can apply to

all models (Bennett et al., 2013; Moriasi et al., 2015).

2.8 Summary

- a) The suitability of a numerical solution of the AD-Model depends on its properties, grid resolution and relative significance of the transport processes. The grid resolution and the relative significance of transport processes characterise the mass transport conditions for a numerical solution of the AD-Model. They are primarily described by two non-dimensional properties known as the advection (or Courant) and dispersion (or diffusion) numbers. The relative significance of the rate of advection of the solute by the flow (defined by the advection number) to the rate of dispersion of the solute driven by the dispersion coefficient (defined by the dispersion number) is known as the Peclet number.
- b) The inverse modelling methodology for analysing breakthrough curves (BTCs) using the AD-Model in most studies has largely been based on the local search approach.
- c) Parameter estimates for the construction of empirical models has largely been achieved using routing procedures. However, empirical models have been applied without considering the methods which were employed to estimate parameters for their development.
- d) Numerical models are influenced by details of the scheme and numerical properties. However, the significance of the impact of schemes and numerical properties on the construction (i.e. development and application) is not documented.
- e) Empirical equations are required to estimate parameter values that are required to be used with a numerical expression of the AD-Model. However, most researchers have based empirical methods on estimates by the classical and the routing methods. It is likely that predictions of these empirical methods would not necessarily be appropriate for application with numerical methods depending on numerical properties.
- f) Practically it has become preferable to base predictive models on channel and hydraulic characteristics which can be easily measured from experimental models.

- g) Most of the empirical methods have developed based on the channel aspect ratio and channel friction and average flow velocity. Therefore, the majority of methods do not consider the influence of stream flow rate, which as observed by Rutherford (1994) influences the rate of longitudinal dispersion.
- h) Most of the empirical equations have been developed based on many rivers/streams without considering the variety of channel types, amount of lengthwise non-uniformity and flow conditions under which observed dispersion coefficients were collected for development of empirical models.
- i) Model construction requires among other things quantifying model performance measures and criteria. Environmental models commonly comprise various definite characteristics and thus require a system of measures addressing several characteristics for a broad assessment of a model. Since every modelling endeavor has specific aims and challenges there are no standard performance evaluation criteria that can apply to all models.

3 APPRAISAL OF SELECTED NUMERICAL METHODS

3.1 Introduction

This chapter attempted to investigate how estimated parameter values of the AD-Model are influenced by Eulerian numerical methods and the numerical properties. Since numerical properties influence Eulerian numerical methods, it was necessary to appraise the methods using synthetic data generated by an analytical method. The other objective was to assess the parameter estimation tools. A set of synthetic BTCs was generated using Taylor's analytical solution (Equation (2.5)) of the AD-Model. The numerical methods were tested for a fixed time step as measured BTCs were collected at fixed time steps, while the space step was varied. The methods, when used with synthetic data, were required to predict concentrations-time curves given by the analytical method and reproduce the parameter values used with the analytical solutions when optimised. The use of synthetic data for testing models is common (Semuwemba, 2011; Vaghela and Vaghela, 2014) as conditions can be tailored to a particular situation. The advantage of using synthetic data is that the values of solute transport parameters are known and that analytical solutions provide results for a broader range of conditions which would not be practically possible with observed data (Semuwemba, 2011). Numerical methods were tested for a hypothetical reach of 200 m, comparable to the actual stream reach which is 184 m, for simplicity.

Parameter estimation used Excel (e.g. Billo, 2007; Wallis et al., 2013). Excel uses the steepest descent method which is a general optimisation package that can compute a minimum or user-specified value of the target cell (Billo, 2007). Testing of numerical methods was followed by a selection of numerical methods for analysis of observed BTCs to estimate stream flow velocity and longitudinal dispersion. The selection of numerical methods was based on the residual sum of squares (RSS) as the objective function for assessing simulations, and the relative error statistic for testing individual parameter estimates.

Numerical solution methods of the AD-Model must deal with several results that influence a solution. Primarily, these effects are a result of numerical approximation of the advective term (Sobey, 1984; Abbott and Basco, 1989), which may introduce artificial (or numerical) dispersion in the solution. It

is not commonly possible to differentiate in the calculated solution between numerical dispersion introduced by a numerical method and the physical process of longitudinal dispersion (Sobey, 1984). Therefore, it was reasonable to compare solutions determined by various numerical schemes and assess the accuracy of several numerical methods for estimating stream solute transport parameters in a small stream.

Six numerical methods were considered based on the order of the scheme (second-order and third-order), the solution method (explicit and implicit) and the discretisation method (finite difference (FD) and finite volume (FV)). The choice of the methods was based on the reasons outlined in Section 2.5.3 concerning the use of the first-order upstream differencing schemes, the central differencing scheme and very high-order schemes.

The FD schemes consisted of the Backward-Time/Centred-Space method (BTCS), the Crank-Nicolson method (CN), the MacCormack method (Versteeg and Malalasekera, 2007; Ferziger and Perić, 2002; Chapra, 2008) and the third-order upstream-differencing method (TUDS) (Kowalik and Murty, 1993). The FV schemes were the implicit scheme with QUICK (quadratic upstream interpolation for convective kinetics) differencing (Versteeg and Malalasekera, 2007) and the QUICKEST (quadratic upstream interpolation for convective kinetics with estimated streaming terms) method (Leonard, 1979).

3.2 Selected numerical methods

3.2.1 The Backward-Time/Centred-Space (BTCS) method

Using Equation (2.24) with $\phi = 1$, the method approximates the spatial derivatives at time level $l+1$ by the centred difference approach. Thus, the method is referred to as a backward-time/centred-space implicit scheme (Chapra, 2008). The method is an FD scheme and is expressed as:

$$\phi_j^{l+1} = \phi_j^l + \left[-v \frac{\phi_{j+1}^{l+1} - \phi_{j-1}^{l+1}}{2\Delta x} + D \frac{\phi_{j+1}^{l+1} - 2\phi_j^{l+1} + \phi_{j-1}^{l+1}}{(\Delta x)^2} \right] \Delta t, \quad (3.1)$$

where, \mathcal{V} is cross-sectional average velocity, ϕ is cross-sectional solute concentration, D is longitudinal dispersion coefficient, Δx is space step and Δt is time step. For computational purposes, the above equation can be expressed in terms of the non-dimensional numerical properties introduced above as

$$(1+2d)\phi_j^{l+1} + (C/2-d)\phi_{j+1}^{l+1} + (-C/2-d)\phi_{j-1}^{l+1} = \phi_j^l, \quad (3.2)$$

where C is advection (or courant) number and d is dispersion number (or diffusion) number.

In this method, values of the transported variable at the new time level are evaluated in terms of other unknown variable values at the new time level, requiring the solution of a set of simultaneous equations. This method is unconditionally stable and allows use of large time steps (Chapra, 2008), however, the method has disadvantages: it has first-order truncation errors in time, manifests artificial dispersion, which is time-step dependent, and oscillations are likely when $C > 1$ (Manson and Wallis, 1995; Ferziger and Peric, 2002; Chapra, 2008). Wallis (2007) warns that the increasing inaccuracy outweighs the promise of stability as C increases. The advantage of this scheme is the possibility of using large time steps, which may lead to a more productive procedure despite drawbacks. The need for solving the problem using matrix algebra can be avoided by writing the equation in a direct form as

$$(1+2d)\phi_j^{l+1} = \phi_j^l + (d-C/2)\phi_{j+1}^{l+1} + (C/2+d)\phi_{j-1}^{l+1}. \quad (3.3)$$

The equation can be solved using spreadsheets, by writing the equation to each cell which can be solved iteratively (Karahan, 2008).

3.2.2 The Crank-Nicolson (CN) method

Using Equation (2.24), with $\phi = 0.5$ the method employs a centred-time/centred-space approach in which estimates of the spatial derivatives are expressed using values of the transported property at time levels l and $l+1$. This is an implicit finite difference scheme and is expressed as:

$$\begin{aligned} \phi_j^{l+1} = \phi_j^l + \frac{\Delta t}{2} \left[-v \frac{\phi_{j+1}^{l+1} - \phi_{j-1}^{l+1}}{2\Delta x} + D \frac{\phi_{j+1}^{l+1} + \phi_{j-1}^{l+1} - 2\phi_j^{l+1}}{\Delta x^2} \right] \\ + \frac{\Delta t}{2} \left[-v \frac{\phi_{j+1}^l - \phi_{j-1}^l}{2\Delta x} + D \frac{\phi_{j+1}^l + \phi_{j-1}^l - 2\phi_j^l}{\Delta x^2} \right] \end{aligned} \quad (3.4)$$

For computational purposes, the above equation can be expressed in terms of dimensionless numerical properties as:

$$-\left(\frac{d}{2} + \frac{C}{4}\right)\phi_{j-1}^{l+1} + (d+1)\phi_j^{l+1} - \left(\frac{d}{2} - \frac{C}{4}\right)\phi_{j+1}^{l+1} = \left(\frac{d}{2} + \frac{C}{4}\right)\phi_{j-1}^l + (1-d)\phi_j^l + \left(\frac{d}{2} - \frac{C}{4}\right)\phi_{j+1}^l. \quad (3.5)$$

In this method, values of the transported variable at the new time level are evaluated in terms of variable values from both the old and the new time levels, requiring the solution of a set of simultaneous equations. The scheme is based on centred differencing in time and space and is thus second order accurate. Alternatively, Eq. 3.5 can be expressed in direct form as

$$(d+1)\phi_j^{l+1} = \left[\left(\frac{d}{2} + \frac{C}{4}\right)\phi_{j-1}^l + (1-d)\phi_j^l + \left(\frac{d}{2} - \frac{C}{4}\right)\phi_{j+1}^l + \left(\frac{d}{2} + \frac{C}{4}\right)\phi_{j-1}^{l+1} + \left(\frac{d}{2} - \frac{C}{4}\right)\phi_{j+1}^{l+1} \right]. \quad (3.6)$$

The scheme requires more computational effort than the BTCS. The scheme is unconditionally stable and does not induce numerical dispersion, and much larger time steps can be adopted without producing wiggles, and stability is improved when the implicit contribution is increased (Wallis, 2007). The fact that the domain of influence of a calculated concentration comes from the whole computational plane (indicated by characteristic lines), the calculated concentration is also affected by parts which physically are not necessary to affect it (Hoffman, 2001; Wallis, 2007). Consequently, the method becomes inaccurate as the advection number and Peclet number increases (Ferziger and Peric, 2002; Wallis, 2007; Chapra, 2008).

3.2.3 The Implicit QUICK method

This is a finite volume (FV) approach with control volume (CV) face values of the transported variable expressed in terms of an upstream weighted parabolic interpolation and spatial gradients of the transported variable expressed using linear interpolation (Leonard, 1979; Versteeg and

Malalasekera, 2007). Using $\phi = 1$ in the FV equivalent version of Eq. (2.24) gives:

$$\phi_j^{l+1} - \phi_j^l = -\frac{\Delta t v}{\Delta x} \left(\phi_{j+1/2}^{l+1} - \phi_{j-1/2}^{l+1} \right) + \frac{\Delta t D}{\Delta x} \left[\left(\frac{\partial \phi}{\partial x} \right)_{j+1/2}^{l+1} - \left(\frac{\partial \phi}{\partial x} \right)_{j-1/2}^{l+1} \right]. \quad (3.7)$$

Using Hayase et al.'s formulation (Hayase et al., 1992; Versteeg and Malalasekera, 2007) to express variable values at the CV surfaces gives:

$$\phi_{j+1/2}^{l+1} = \phi_j^{l+1} + \frac{1}{8} (3\phi_{j+1}^{l+1} - 2\phi_j^{l+1} - \phi_{j-1}^{l+1}). \quad (3.8)$$

$$\phi_{j-1/2}^{l+1} = \phi_{j-1}^{l+1} + \frac{1}{8} (3\phi_j^{l+1} - 2\phi_{j-1}^{l+1} - \phi_{j-2}^{l+1}). \quad (3.9)$$

The discretised equation for a general FV centred at the node, j , is expressed as:

$$\left(1 + 2d + \frac{3C}{8} \right) \phi_j^{l+1} = \phi_j^l + \left(d + \frac{7C}{8} \right) \phi_{j-1}^{l+1} + \left(d - \frac{3C}{8} \right) \phi_{j+1}^{l+1} - \frac{C}{8} \phi_{j-2}^{l+1}. \quad (3.10)$$

The accuracy of the scheme on a uniform grid is third-order in space and first-order in time, requiring small time steps for accurate results (Versteeg and Malalasekera, 2007). For constant time step and velocity, the QUICK scheme has artificial dispersion which increases with space-step. Additionally, the implicit solution method introduces artificial numerical dispersion which depends on the time-step (Chapra, 2008) and, like all implicit solution methods, calculated concentrations are also affected by parts which have no physical reason to influence the calculated values.

3.2.4 The QUICKEST Method

The QUICKEST (quadratic upstream interpolation of convective kinetics with estimated streaming terms) method (Leonard, 1979) is an FV approach similar to, but superior to, the QUICK method. As well as using the upstream weighted parabolic interpolation of the QUICK method (Abbott and Basco, 1989, Versteeg and Malalasekera, 2007) it is an explicit method, using the $\phi = 0$ version of the FV equivalent of Equation (2.24). Additionally, it includes 'estimated streaming terms' (EST) to account for advection and dispersion occurring during the time step (Leonard, 1979). Variable values at CV faces are given by the following expressions (Leonard, 1979; Abbott and Basco, 1989):

$$\phi_{j+1/2}^l = \frac{1}{2} (\phi_j^l + \phi_{j+1}^l) - \frac{1}{6} (\phi_{j-1}^l - 2\phi_j^l + \phi_{j+1}^l). \quad (3.11)$$

$$\varphi'_{j-1/2} = \frac{1}{2}(\varphi'_{j-1} - \varphi'_j) - \frac{1}{6}(\varphi'_{j-2} - 2\varphi'_{j-1} + \varphi'_j). \quad (3.12)$$

For calculation purposes the QUICKEST scheme can be expressed as (Leonard, 1979; Abbott and Basco, 1989):

$$\begin{aligned} \varphi_j^{l+1} = & \varphi_j^l + \left[d(1-C) - \frac{C}{6}(C^2 - 3C + 2) \right] \varphi_{j+1}^l - \left[d(2-3C) - \frac{C}{2}(C^2 - 2C - 1) \right] \varphi_j^l \\ & + \left[d(1-3C) - \frac{C}{2}(C^2 - C - 2) \right] \varphi_{j-1}^l + \left[d(C) + \frac{C}{6}(C^2 - 1) \right] \varphi_{j-2}^l \end{aligned} \quad (3.13)$$

The expression shows that the unknown values of the transported variable at the new time level can be computed one at a time (an explicit calculation). Ideally, the scheme is third order accurate, but its explicit nature leads to conditional stability and boundedness. Thus, it can be unstable at modest values of P_e ; Leonard (1979) provides some detail on this issue. The scheme manifests artificial dispersion which is both space step and time step dependent (Leonard, 1979).

3.2.5 The MacCormack Method

The MacCormack method is a predictor-corrector method. It is a two-step finite difference method, unlike the above-discussed techniques which are one-step methods. There are various formulations of the approach (e.g. Maccormack, 1982; Fürst and Furmánek, 2011). Here the implicit formulation by MacCormack (1969) described in (Chapra, 2008; p. 229) was followed. The first step (predictor) uses the following explicit estimator, which uses forward spatial differencing for the advective term and centred spatial differencing for the dispersion term:

$$\varphi^* = \left(\frac{\varphi_j^{l+1} - \varphi_j^l}{\Delta t} \right) = -v \frac{\varphi_{j+1}^l - \varphi_j^l}{\Delta x} + D \frac{\varphi_{j+1}^l - 2\varphi_j^l + \varphi_{j-1}^l}{\Delta x^2}. \quad (3.14)$$

The second step (corrector) uses the following implicit estimator, which uses backward spatial differencing for the advective term and centred spatial differencing for the dispersion term:

$$\varphi^{**} = \left(\frac{\varphi_j^{l+1} - \varphi_j^l}{\Delta t} \right) = -v \frac{\varphi_j^{l+1} - \varphi_{j-1}^{l+1}}{\Delta x} + D \frac{\varphi_{j+1}^{l+1} - 2\varphi_j^{l+1} + \varphi_{j-1}^{l+1}}{\Delta x^2}. \quad (3.15)$$

Finally, an average of the two estimators is used to obtain the result expressed as:

$$\phi_j^{l+1} = \phi_j^l + \left(\frac{\phi^* + \phi^{**}}{2} \right) \Delta t. \quad (3.16)$$

This method, like the Crank-Nicolson scheme, is an average of an explicit scheme and an implicit scheme. For computational purposes the equation can be expressed as:

$$-\left(\frac{d}{2} + \frac{C}{2} \right) \phi_{j-1}^{l+1} + \left(1 + d + \frac{C}{2} \right) \phi_j^{l+1} - \frac{d}{2} \phi_{j+1}^{l+1} = \frac{d}{2} \phi_{j-1}^l + \left(1 + \frac{C}{2} - d \right) \phi_j^l + \left(\frac{d}{2} - \frac{C}{2} \right) \phi_{j+1}^l. \quad (3.17)$$

This expression has the same form as other implicit methods described above and is similar to that presented by Furst and Furmanek (2011). The scheme is unconditionally stable (Furst and Furmánek, 2011) and it is second-order accurate in space and time and does not manifest artificial dispersion (Furst and Furmánek, 2011). Although Chapra (2008) claims that the scheme is conditionally stable, he has not provided details. Alternatively, Equation. (3.17) can be expressed in a direct form as

$$\left(1 + d + \frac{C}{2} \right) \phi_j^{l+1} = \frac{d}{2} \phi_{j-1}^l + \left(\frac{d}{2} + \frac{C}{2} \right) \phi_{j-1}^{l+1} + \left(1 + \frac{C}{2} - d \right) \phi_j^l + \left(\frac{d}{2} - \frac{C}{2} \right) \phi_{j+1}^l + \frac{d}{2} \phi_{j+1}^{l+1}. \quad (3.18)$$

3.2.6 Third-order upstream-differencing scheme

The first order upstream-differencing scheme is stable at high Peclet numbers and conserves positive definite property (Leonard, 1979; Roache, 1997; Versteeg and Malalasekera, 2007). However, the method induces an excessive numerical dispersion coefficient caused by the advective term. To reduce numerical dispersion a higher order of four-point formulation is used to approximate the advective term (Kowalik and Murty, 1993). It is based on the definition of the AD-Model by discretising the advection term using a third-order upstream-differencing scheme and second-order centred-differencing scheme for the dispersion term (Kowalik and Murty, 1993).

$$\frac{\partial \phi}{\partial x} = \frac{(\phi_{j-2}^l - 6\phi_{j-1}^l + 3\phi_j^l + 2\phi_{j+1}^l)}{6\Delta x}. \quad (3.19)$$

Thus, the general expression for interior nodes is expressed as:

$$\frac{\varphi_i^{l+1} - \varphi_i^l}{\Delta t} + \frac{v}{6\Delta x} (\varphi_{i-2}^l - 6\varphi_{i-1}^l + 3\varphi_i^l + 2\varphi_{i+1}^l) = \frac{D}{\Delta x^2} (\varphi_{i-1}^l - 2\varphi_i^l + \varphi_{i+1}^l). \quad (3.20)$$

For computational purposes, the algorithm can be expressed in terms of non-dimensional numerical properties as:

$$\varphi_i^{l+1} = -\frac{C}{6}\varphi_{i-2}^l + (C+d)\varphi_{i-1}^l + \left(1 - \frac{C}{2} - 2d\right)\varphi_i^l + \left(d - \frac{C}{3}\right)\varphi_{i+1}^l. \quad (3.21)$$

The expression shows that unknown variable values at the new time level can be calculated one at a time. The scheme is third-order accurate in space, but its explicit formulation leads to conditional stability.

3.3 Generating synthetic data

Synthetic BTCs, generated by parameter values similar to those observed in the Murray stream (e.g. Ani et al., 2009; Wallis et al., 2013; Heron, 2015), were used to appraise optional numerical methods. Synthetic BTCs were generated using Eq. 2.5 (Taylor, 1954; Rutherford, 1994). The synthetic BTCs were generated for one set of input values, namely. velocity, $v = 0.225$ m/s, longitudinal dispersion coefficient, $D = 0.750$ m²/s, and time step, $\Delta t = 20$ seconds. The parameter values were selected based on previous studies of the Murray stream. A hypothetical stream reach length of 200 m was used, which was comparable to the length of the actual study reach. Temporal concentration profiles were generated for upstream and downstream sites at 600 m and 800 m, respectively, from the solute source. The mass of solute was 100 g, and the cross-sectional area of the channel was 1.0 m². The numerical methods were applied to the synthetic BTCs over a range of space steps (5 m to 40 m), such that optimisation of the model parameters was observed under different values of the non-dimensional numerical properties.

3.4 Numerical Concentration prediction by explicit schemes

The procedure used for predicting temporal concentration distribution at a downstream site is a numerical convolution solution method as described by Abbott and Basco (1989). The continuum

form, Eq. 2.33, of the convolution, is replaced by the following summation:

$$\varphi_j^{l+1} = \sum_{\kappa=-K_m}^{\kappa=+K_m} Z_{\kappa} \cdot \varphi_{j+\kappa}^l \quad \kappa =, \pm 1, \pm 2, \dots, \pm K_m, \quad (3.22)$$

where Z_k is a suitably chosen function similar to $\psi(\varepsilon)$, of the continuum form (i.e. Equation (2.33)), and dependent on the translation method employed, with K_m such that at least one of Z_{K_m} and Z_{-K_m} is non-zero; k is a counter index, with maximum value K_m ; j is the spatial index of the space step Δx , with maximum value jj ; l is the time index of the time step Δt , with maximum value L ; φ_j^l is the given initial conditions over discrete points, j ; and φ_j^{l+1} is the desired solution after some time step Δt . Repetition of the above equation together with appropriate boundary conditions produces a numerical solution at a time $T = L \cdot \Delta t$.

3.5 Concentration prediction by implicit schemes

The continuum form and its translation into digital form are applicable for implicit-type solution procedures when K_m covers the whole solution domain (Abbott and Basco, 1989). The digital explicit-scheme concepts can be extended to the compact form of implicit schemes, where the convolution integral is considered in the way of summations for both sides of the equation, by introducing a second weighting function A_i , such that:

$$\sum_{-P_m}^{+P_m} A_p \cdot \varphi_{j+1}^{l+1} = \sum_{-K_m}^{+K_m} Z_{\kappa} \cdot \varphi_{j+\kappa}^l \quad \begin{array}{l} \kappa = 0, \pm 1, \pm 2, \dots, \pm K_m, \\ p = 0, \pm 1, \pm 2, \dots, \pm P_m, \end{array} \quad (3.23)$$

where A_p time-step-increment weighting function; p is a counter index. This results in a system of linear equations with variably-populated coefficient matrices for weighting functions which require the use of matrix algebra (Abbott and Basco, 1989). However, the formulation of implicit schemes can be expressed in a direct form such that the use of matrix algebra is not required. This approach can be achieved by making the variable value φ_j^{n+1} subject of the expression (such as in Equations (3.3), (3.60), (3.10) and (3.18)) and the solution can be obtained using Excel (Karahana, 2008).

3.6 Identification of parameters

In contrast to simulating a solute transport event, estimating solute transport parameters using BTCs is an inverse modelling problem. In inverse modelling, values of model parameters are determined which give the best fit between the simulated and the observed data (Chapra, 2008; Chin, 2013; Sun and Sun, 2015). The level of agreement between the model output and the observations is used to assess the performance of the model (Runkel & Broshears 1991; Wallis et al. 2013), which can be assessed using one or several performance measures (Bennett et al., 2013; Chin, 2013).

The numerical quadrature can be expressed in functional form as (Singh and Beck, 2003):

$$\varphi(X, j\Delta t) = f[X, \Delta t, D, v, \varphi_u(j\Delta t)] \quad j = 1, 2, \dots, n, \quad (3.24)$$

where $\varphi_u(\cdot)$ = observed upstream temporal concentration, $f(\cdot)$ = predicted concentration at the downstream site and n = number of simulated points. With a set of observed values for $\bar{\varphi}(X, j\Delta t)$, at the downstream site, and $\varphi_u(j\Delta t)$, at the upstream site, the v and D can be optimised at a minimum residual sum of squares (RSS). The RSS is the measure of goodness of fit between the observed and the predicted concentrations at the downstream site (Singh and Beck, 2003; Chin, 2013; Wallis et al., 2013). The optimisation procedure, in this case, may be achieved by the following formulation:

$$\text{Minimise } RSS = \sum_{i=1}^n (\bar{\varphi}(X, j\Delta t) - \varphi(X, j\Delta t))^2. \quad (3.25)$$

The minimisation of the RSS results in optimal values of velocity and dispersion coefficient. Estimation of the parameters can allow for the possibility, to some degree, the lack of fit of the AD-Model in representing the influence of the type of numerical method and numerical properties.

3.6.1 Computational method

Excel was used to apply numerical methods which followed the procedure similar to that described by Karahan (2006, 2007, 2008). Excel uses a general optimisation code that can compute a minimum or user-specified value of the target cell (van den Bos, 2007). The use of the code has been observed to be efficient and robust and compares well with other algorithms (Fylstra et al.,

2012). Spreadsheets can also be configured to solve implicit schemes without the need for matrix computation by expressing equations in direct form and using an inbuilt iterative technique. The equations can be written directly to each cell, which represents a computational node (Karahan, 2006). The general structure used for computations is shown in Fig. 3-1. The input variables were Δx , Δt , ν and D from which values of the three numerical properties were determined. Since time was fixed, the spatial resolution of the computational domain was varied to estimate parameters over a range of numerical properties. In the solution domain columns symbolise different spatial positions and rows symbolise different times. The solution for each method was carried out based on computational algorithms expressed for each cell. Before starting the model, it was necessary to enter the initial and boundary conditions. The initial conditions are shown in row 10, and the upstream boundary conditions are shown under the heading “upstream” in column E. Values under the heading “downstream” represent an observed concentration profile at the downstream site. A separate workbook was created for each numerical method, and each appropriate algorithm was written to spreadsheets. In each workbook, a separate spreadsheet was created for each space step. The upstream and downstream observed BTCs are presented in Appendix A1.

The equations were written to all grid points for all time steps and well beyond the downstream boundary, such that the last column had zero values. Values under the heading “downstream” represent an observed concentration profile at the downstream site. The column headed by the residual sum squares (RSS) represents the squared differences between the predicted concentration and the observed concentration values. The sum of residual sum of squares (Sum RSS) is the objective function.

The methods were then solved to predict the concentration distribution downstream using the upstream concentration profile. The optimisation code evaluates partial derivatives of the objective function (i.e. target cell) with respect to parameters (changing cells; ν and D) using the finite-difference approach. When estimating parameters, the lack of fit of a predicted concentration profile to the downstream observed concentration profile can account for, to some degree, the influence of the numerical method and numerical properties. The residual sum of squares was minimised to

achieve the best fit between the downstream concentration profile and the predicted profile. The flow chart for of the modelling process is given in Fig. 3-2. The process was repeated for several space steps ranging from 5 m to 40 m.

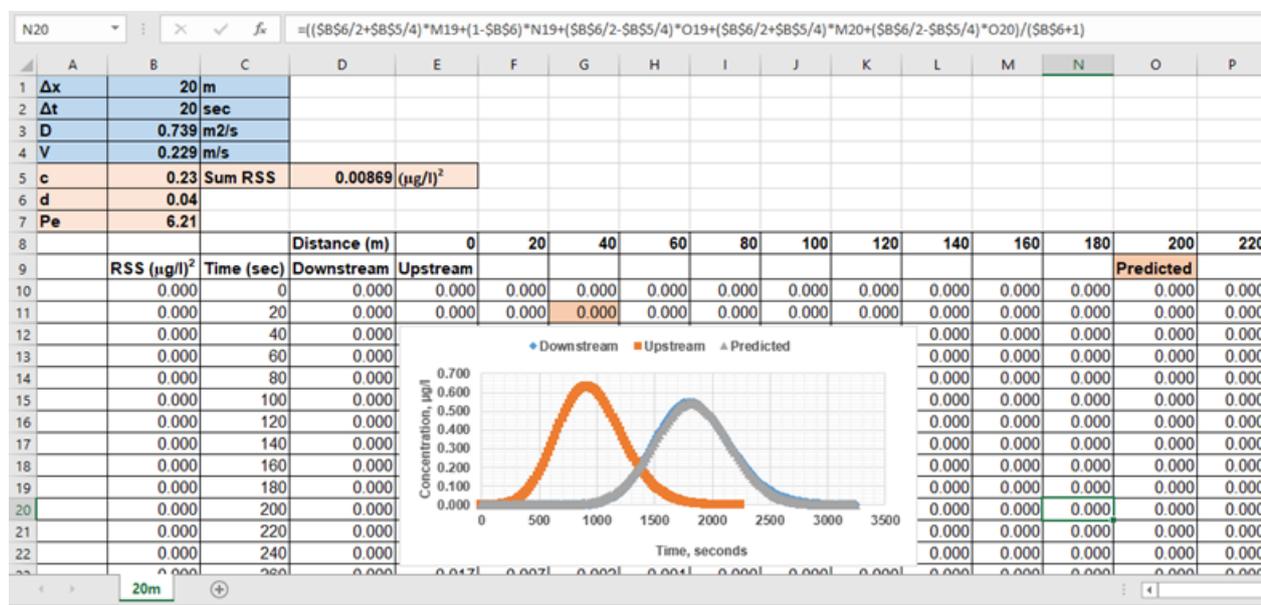


Figure 3-1: General structure of the spreadsheet for the numerical solution domain.

The QUICKEST and the third-order upstream-differencing methods yield the required concentration downstream explicitly, while the other numerical methods were implemented in a direct form. The implicit methods are exemplified using the Crank-Nicolson method below. In its direct form it is given by Equation (3.6), for which the computation written to the spreadsheet, for example for cell G11, is:

$$G11 = \frac{(\$B\$6/2 + \$B\$5/4) * F10 + (1 - \$B\$6) * G10 + (\$B\$6/2 - \$B\$5/4) * H10 + (\$B\$6/2 + \$B\$5/4) * F11 + (\$B\$6/2 - \$B\$5/4) * H11}{(\$B\$6 + 1)}, \quad (3.26)$$

where G11 is the cell at the intersection of column G and row 11, and contains the concentration at the time $l + 1$. The equation is written to all cells for all time steps and well beyond the downstream boundary such that the last column has zero values. Predicted concentrations at the downstream boundary (column O) are compared to the synthetic concentration profile which is given in column D. The disagreement between the two profiles is quantified using the (RSS). The RSS is minimised to attain the best fit between the predicted concentration profile and the downstream concentration

profile by tuning ν and D using Excel's solver function. Fylstra et al. (1998) claim that the use of the code is efficient and robust and compares well with other algorithms.

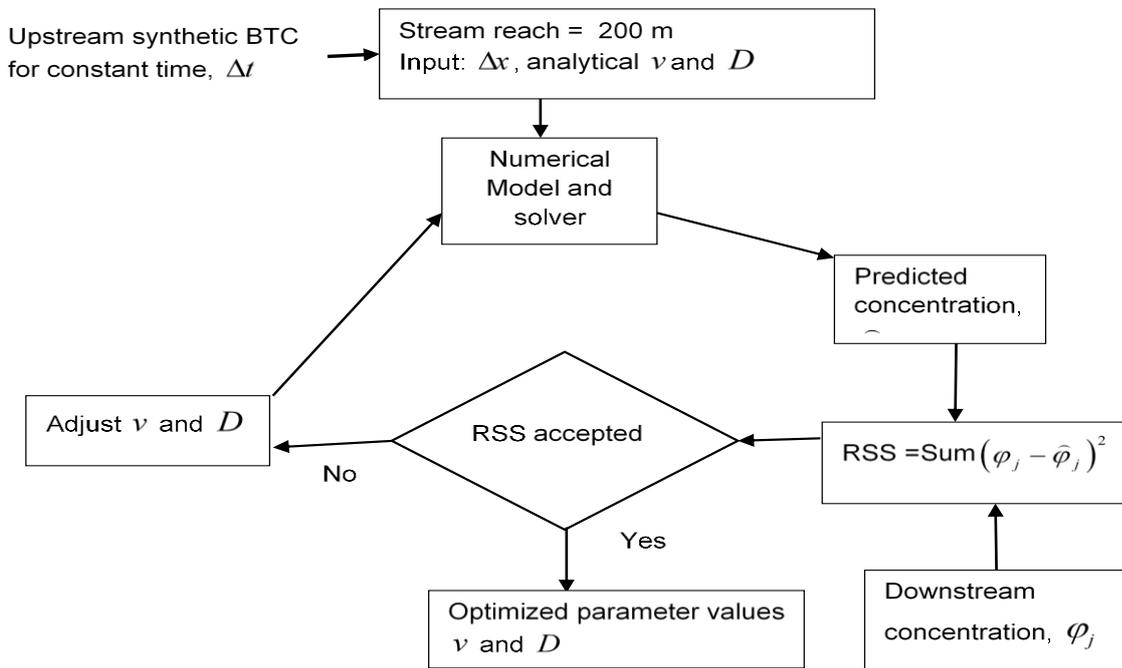


Figure 3-2: A flowchart for concentration prediction and parameter estimation, adopted from Chapra (2008).

3.7 Performance assessment of the methods

The values of the parameters were obtained by solving a numerical non-linear optimisation problem in Excel. There are two general approaches for assessing the quality of calibration; subjective and objective. Subjective approach is based on visual comparison of prediction and observed data (Chapra, 2008). In contrast, the objective approaches are based on developing some quantitative measure of the quality of fit between the model output and data. The method is based on minimising an error statistic known as the objective function, computed as the residual sum of squares (RSS) between the model output and the observed BTCs (Ani et al., 2009; Chin, 2013). This error statistic is used widely in water quality modelling (Chapra, 2008; Ani et al., 2009; Wallis et al. 2013).

The minimum of the error statistic is determined by adjusting values of velocity and dispersion coefficient so that prediction optimally fits observations. The initial search parameter values were those used in the analytical solution to generate synthetic data. Additionally, to assess the accuracy

of optimised parameter values the relative percentage error (RE) was used. The RE statistic is appropriate for assessing the accuracy of individual values (Chin, 2013).

The *residual sum of squares* is expressed as (Montgomery and Runger, 2003; Chin, 2013):

$$RSS = \sum_{i=1}^n (\varphi_j - \hat{\varphi}_j)^2, \quad (3.27)$$

where φ_j are the observed concentration values and $\hat{\varphi}_j$ are the predicted concentration values.

The relative percentage error (RE) statistic is expressed as (Chin, 2013)

$$RE = \frac{|\hat{\theta} - \theta|}{\theta} \times 100, \quad (3.28)$$

where θ is expected parameter value and $\hat{\theta}$ is the estimated parameter value.

3.8 Results and discussion

Six numerical methods were applied to synthetic data generated by Taylor's analytical solution of the of the AD-Model. Results of the numerical schemes behaved conservatively, and simulations were stable. There was very little evidence for any problems stemming from a lack of boundedness or stability often provided by the appearance of oscillations, except for third-order upstream-differencing scheme (TUDS) as it became unstable above advection numbers of 0.700. The results obtained were simulated concentrations, determined velocities and longitudinal dispersion coefficients. The statistics used to assess the methods were the residual sum of squares and the relative error. The residual sum of squares statistic was used to measure the differences between predicted concentration-time profiles and the downstream concentration profiles. The relative error was used to measure the error between the optimised parameter values and those used by the analytical solution to generate synthetic data.

Tables 3-1 lists space steps, dimensionless numerical properties and optimised dispersion coefficients for all optional numerical methods. Table 3-2 lists space steps, dimensionless numerical properties and optimised values of velocity for all optional numerical methods. Since the time step was constant, the space step was varied ranged from 5 m to 40 m. Below space steps of 5 m, the

QUICKEST method gave unstable and unbounded predictions, while above space step of 40 m both QUICKEST and Implicit QUICK failed to find positive values of dispersion coefficient because of the small value of dispersion coefficient used. However, the value of the dispersion coefficient used is in the range of values observed in the previous studies on the Murray stream.

It should be noted that the values of dimensionless numerical properties, i.e. advection number, dispersion number and Peclet number, listed in the tables were computed using the parameter values used for generating the synthetic data so that the results of all methods could be easily collated. It can be observed that there was a variation of optimised dispersion coefficients and velocities both with the numerical method used and with the numerical properties. However, the optimised velocity values show relatively little variation compared to the optimised dispersion coefficients.

Table 3-1: Dimensionless numerical properties and computed dispersion coefficients (m^2/s) (UNS means unstable solution)

Δx , (m)	Advection Number	Dispersion Number	Peclet Number	CN	BTCS	Implicit QUICK	MacCormack	QUICKEST	TUDS
5.00	0.900	0.600	1.500	0.749	0.397	0.234	0.749	0.708	UNS
6.06	0.743	0.408	1.818	0.749	0.438	0.253	0.748	0.764	UNS
7.14	0.630	0.294	2.143	0.748	0.492	0.269	0.747	0.762	0.885
8.00	0.563	0.234	2.400	0.747	0.520	0.253	0.746	0.759	1.163
10.00	0.450	0.150	3.000	0.746	0.545	0.234	0.745	0.751	1.229
12.50	0.360	0.096	3.750	0.743	0.429	0.197	0.744	0.736	1.208
14.29	0.315	0.073	4.286	0.742	0.361	0.176	0.744	0.721	1.190
16.67	0.270	0.054	5.000	0.741	0.224	0.148	0.745	0.695	1.160
18.18	0.248	0.045	5.454	0.750	0.384	0.136	0.747	0.674	1.137
20.00	0.225	0.038	6.000	0.739	0.392	0.109	0.751	0.646	1.107
25.00	0.180	0.024	7.500	0.752	0.242	0.025	0.770	0.544	1.002
28.57	0.158	0.018	8.571	0.766	0.255	0.000	0.790	0.450	0.909
40.00	0.113	0.009	12.000	0.827	0.316	0.000	0.854	0.041	0.517

Table 3-2: Dimensionless numerical properties and computed velocity (m/s).

Δx , (m)	Advection Number	Dispersion Number	Peclet Number	CN	BTCS	Implicit QUICK	MacCormack	QUICKEST	TUDS
5.00	0.900	0.600	1.500	0.225	0.231	0.225	0.226	0.227	UNS
6.06	0.743	0.408	1.818	0.225	0.226	0.225	0.226	0.227	UNS
7.14	0.630	0.294	2.143	0.226	0.226	0.229	0.226	0.226	0.223
8.00	0.563	0.234	2.400	0.226	0.226	0.224	0.226	0.226	0.223
10.00	0.450	0.150	3.000	0.226	0.227	0.224	0.227	0.225	0.222
12.50	0.360	0.096	3.750	0.227	0.227	0.224	0.228	0.224	0.222
14.29	0.315	0.073	4.286	0.227	0.228	0.224	0.228	0.224	0.221
16.67	0.270	0.054	5.000	0.228	0.229	0.223	0.229	0.223	0.221
18.18	0.248	0.045	5.454	0.229	0.229	0.223	0.230	0.223	0.221
20.00	0.225	0.038	6.000	0.229	0.230	0.223	0.231	0.222	0.220
25.00	0.180	0.024	7.500	0.232	0.233	0.223	0.234	0.221	0.220
28.57	0.158	0.018	8.571	0.234	0.235		0.236	0.220	0.219
40.00	0.113	0.009	12.000	0.242	0.243		0.244	0.219	0.219

Figures 3-3 and 3-4 show different features of the results plotted against Peclet number. As a fact, numerical discretisation can be expressed in terms of model parameters, space step and time step or in terms of dimensionless numerical properties, i.e. advection number and dispersion number. The results were plotted against Peclet number because it combines the effects of the other two dimensionless numerical properties and the plots are an effective way of showing the behaviour of the schemes when the space step is varied.

Figure 3-3 shows the determined longitudinal dispersion coefficients for all numerical methods plotted against the Peclet number. It can be observed that optimised dispersion coefficients vary significantly with the *Peclet* number and the numerical method used. The Crank-Nicolson and the MacCormack methods accurately reproduce dispersion coefficient values used in the analytical solution over a considerable range of values of the Peclet number. However, the methods tend to overestimate the dispersion coefficient for Peclet number greater than 8.0. The QUICKEST method only accurately reproduces the dispersion coefficient over a narrow range of low values of Peclet number (2 to 4). With the increase in the Peclet number, the QUICKEST method underestimates the dispersion coefficient. The Backward-Time/Centred-Space method shows fluctuating values of optimised dispersion coefficient over the full range of *Peclet* number, but they are all significantly smaller than the synthetic values. The Implicit QUICK method gives the lowest optimised dispersion coefficients at all values of the *Peclet* number, and they decrease with increasing value of the *Peclet* number: none of them is accurate. Indeed, the computation failed for the largest two values of the Peclet number. The third-order upstream differencing scheme (TUDS) gives highest values optimised dispersion coefficients between Peclet numbers 2 to 8. However, the method tends to underestimate the dispersion coefficient when the Peclet number is greater than 10. The tendency of these four methods to underestimate and overestimate the dispersion coefficient can be attributed to the presence of numerical dispersion, which is either dependent on time step or space step or both.

Fig. 3-4 shows a plot of computed velocity against Peclet number. Although there is a variation in the computed velocities, they are not highly affected by the method used, and the numerical properties since accurate values were obtained by all methods over the range of Peclet number

considered.

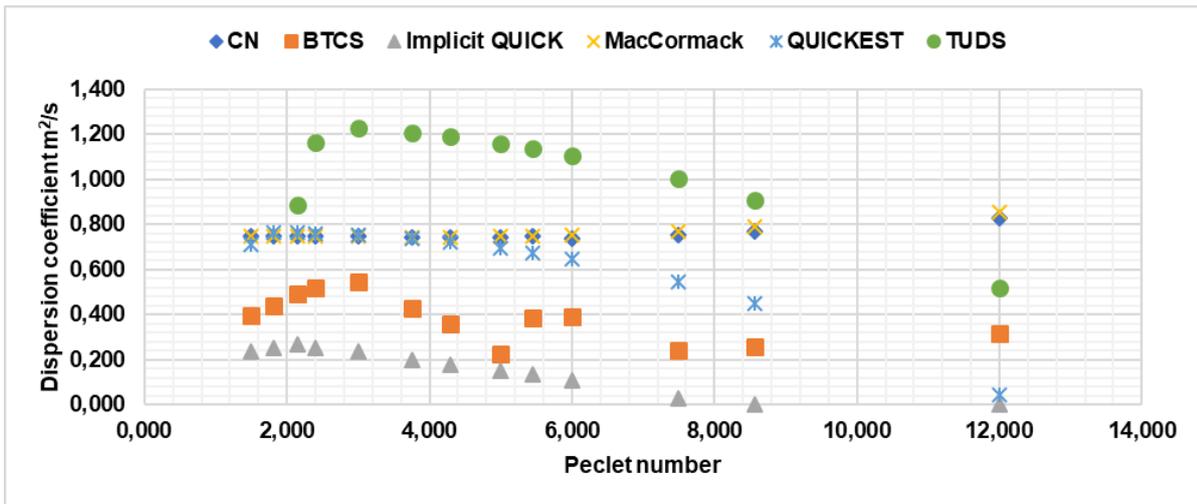


Figure 3-3: Plot of computed dispersion coefficients versus the Peclet number

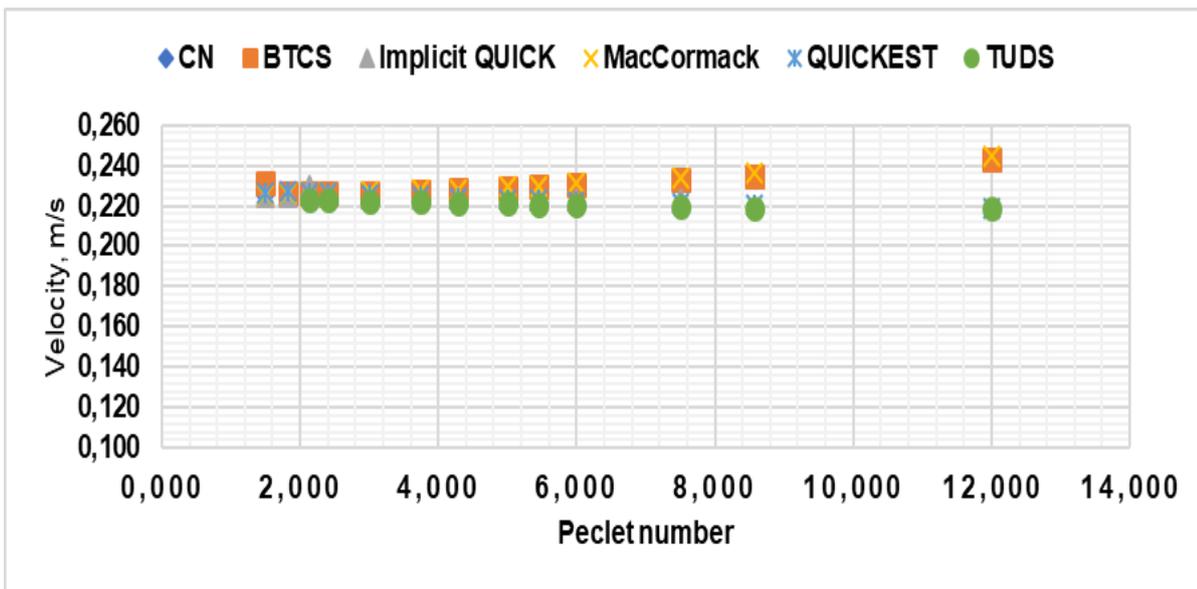


Figure 3-4: Plot of computed velocity versus the Peclet number

Table 3-3 lists values of dimensionless numerical properties and residual sum of squares for all the methods which were plotted as shown in Fig. 3-5. The results of the residual sum of squares obtained over the range of Peclet number were quite low for all the methods used but varied between them and with values of the Peclet number. Generally, the CN, BTCS, and MacCormack, which are finite-difference methods gave higher simulation errors than the Implicit QUICK and the QUICKEST, which are finite-volume methods, particularly at higher values of the *Peclet* number.

Table 3-3: Dimensionless numerical properties and residual sum of squares ($\mu\text{g}^2/\text{l}^2$)

Δx , (m)	Advection Number	Dispersion Number	Peclet Number	CN	BTCS	Implicit QUICK	MacCormack	QUICKEST	TUDS
5.00	0.900	0.600	1.500	0.00006	0.00303	0.00028	0.00023	0.00001	UNS
6.06	0.743	0.408	1.818	0.00011	0.00732	0.00041	0.00038	0.00000	UNS
7.14	0.630	0.294	2.143	0.00018	0.01138	0.00047	0.00059	0.00000	0.02823
8.00	0.563	0.234	2.400	0.00027	0.01388	0.00056	0.00081	0.00000	0.00188
10.00	0.450	0.150	3.000	0.00060	0.01699	0.00058	0.00157	0.00000	0.00051
12.50	0.360	0.096	3.750	0.00138	0.00938	0.00067	0.00318	0.00000	0.00046
14.29	0.315	0.073	4.286	0.00231	0.00703	0.00081	0.00493	0.00001	0.00042
16.67	0.270	0.054	5.000	0.00424	0.00644	0.00117	0.00831	0.00003	0.00035
18.18	0.248	0.045	5.454	0.00628	0.01543	0.00173	0.01121	0.00006	0.00032
20.00	0.225	0.038	6.000	0.00869	0.01938	0.00225	0.01561	0.00012	0.00030
25.00	0.180	0.024	7.500	0.02152	0.02751	0.00516	0.03400	0.00067	0.00057
28.57	0.158	0.018	8.571	0.03631	0.04490		0.05362	0.00170	0.00140
40.00	0.113	0.009	12.000	0.11996	0.13816		0.15241	0.01403	0.01365

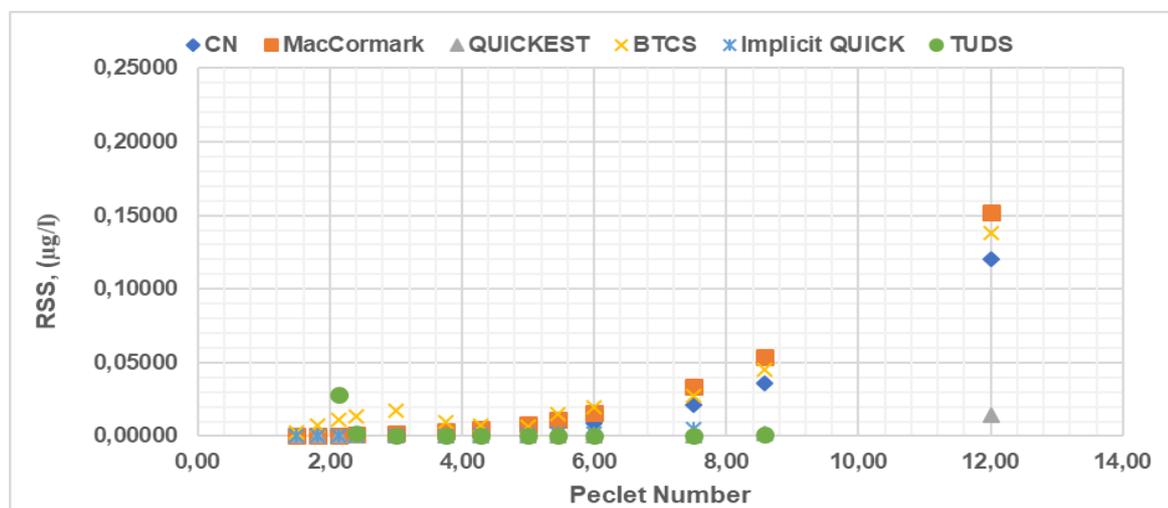


Figure 3-5: Plot of the residual sum of squares versus Peclet number

Tables 3-4 and 3-5 list relative percentage errors of optimised longitudinal dispersion coefficients and velocity, respectively. It can be observed that optimised dispersion values show high variations while optimised velocity values show little variation. For longitudinal dispersion coefficient, Table 3.4 shows that there are considerable differences between optimised and synthetic values for the Backward-Time/Centred-Space, Implicit Quick, QUICKEST and TUDS methods, but much smaller differences for the Crank-Nicolson and the MacCormack methods. As might be expected from the velocity results shown earlier, the relative errors of optimised to analytical velocity values were much lower than for the dispersion coefficients for all the schemes over the range of Peclet number used, as shown in Figures 3-6 and 3-7.

Table 3-4: Dimensionless numerical properties and relative percentage errors of estimated dispersion coefficient

Δx , (m)	Advection Number	Dispersion Number	Peclet Number	CN	BTCS	Implicit QUICK	MacCormack	QUICKEST	TUDS
5.00	0.900	0.600	1.500	0.08	47.07	68.86	0.19	5.58	UNS
6.06	0.743	0.408	1.818	0.19	41.58	66.33	0.31	1.87	UNS
7.14	0.630	0.294	2.143	0.27	34.38	64.08	0.39	1.59	18.05
8.00	0.563	0.234	2.400	0.36	30.73	66.25	0.48	1.27	55.10
10.00	0.450	0.150	3.000	0.60	27.28	68.84	0.66	0.19	63.81
12.50	0.360	0.096	3.750	0.89	42.78	73.74	0.82	1.88	61.08
14.29	0.315	0.073	4.286	1.07	51.93	76.58	0.83	3.90	58.64
16.67	0.270	0.054	5.000	1.21	70.17	80.30	0.66	7.38	54.65
18.18	0.248	0.045	5.454	0.04	48.86	81.88	0.42	10.12	51.63
20.00	0.225	0.038	6.000	1.46	47.72	85.41	0.09	13.89	47.59
25.00	0.180	0.024	7.500	0.23	67.70	96.66	2.64	27.44	33.64
28.57	0.158	0.018	8.571	2.17	65.99		5.39	39.97	21.18
40.00	0.113	0.009	12.000	10.25	57.89		13.91	94.57	31.12

Table 3-5: Dimensionless numerical properties and relative percentage errors of estimated velocity

Δx , (m)	Advection Number	Dispersion Number	Peclet Number	CN	BTCS	Implicit QUICK	MacCormack	QUICKEST	TUDS
5.00	0.900	0.600	1.500	0.17	2.83	0.02	0.33	0.68	UNS
6.06	0.743	0.408	1.818	0.21	0.41	0.10	0.41	0.91	UNS
7.14	0.630	0.294	2.143	0.29	0.47	1.93	0.52	0.64	0.93
8.00	0.563	0.234	2.400	0.35	0.57	0.25	0.61	0.45	1.01
10.00	0.450	0.150	3.000	0.52	0.72	0.43	0.85	0.08	1.18
12.50	0.360	0.096	3.750	0.79	1.03	0.52	1.20	0.32	1.41
14.29	0.315	0.073	4.286	1.03	1.29	0.59	1.50	0.57	1.58
16.67	0.270	0.054	5.000	1.39	1.75	0.69	1.94	0.89	1.79
18.18	0.248	0.045	5.454	1.61	1.90	0.79	2.24	1.10	1.94
20.00	0.225	0.038	6.000	2.00	2.29	0.83	2.65	1.31	2.08
25.00	0.180	0.024	7.500	3.10	3.48	0.90	3.91	1.84	2.44
28.57	0.158	0.018	8.571	4.03	4.44		4.91	2.16	2.63
40.00	0.113	0.009	12.000	7.49	7.99		8.56	2.70	2.83

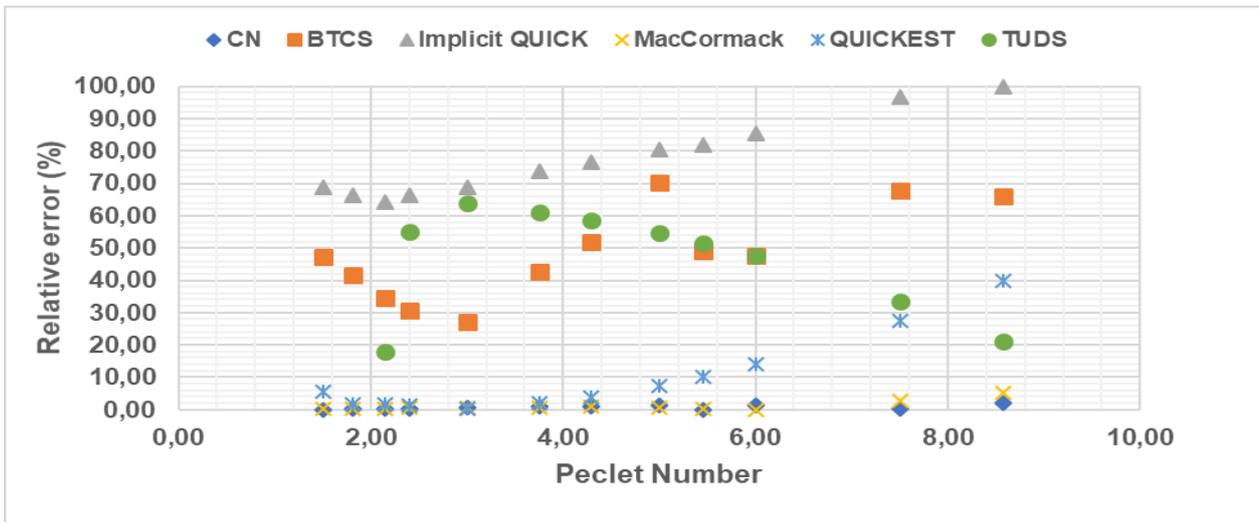


Figure 3-6: Plot of relative error of optimised dispersion coefficient versus the Peclet number

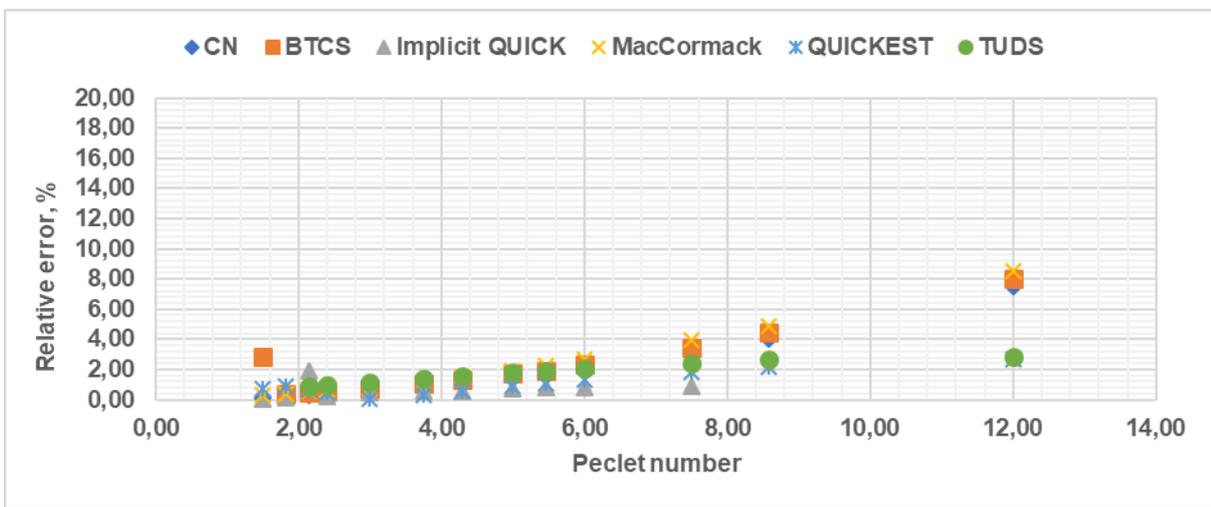


Figure 3-7: Plot of relative error of optimised velocity versus Peclet number

Fig. 3-8 shows simulated temporal concentration profiles obtained using a space step of 20 m of the reach with all six numerical methods. The results show that accurate simulations can be achieved by all methods, using the same spatial and temporal resolution, but with different parameter values to those used to generate the synthetic data (especially the dispersion coefficient, as discussed earlier). The values of dispersion coefficients and velocity used in simulations are listed in Tables 3-1 and 3-2, respectively, for space step of 20 m.

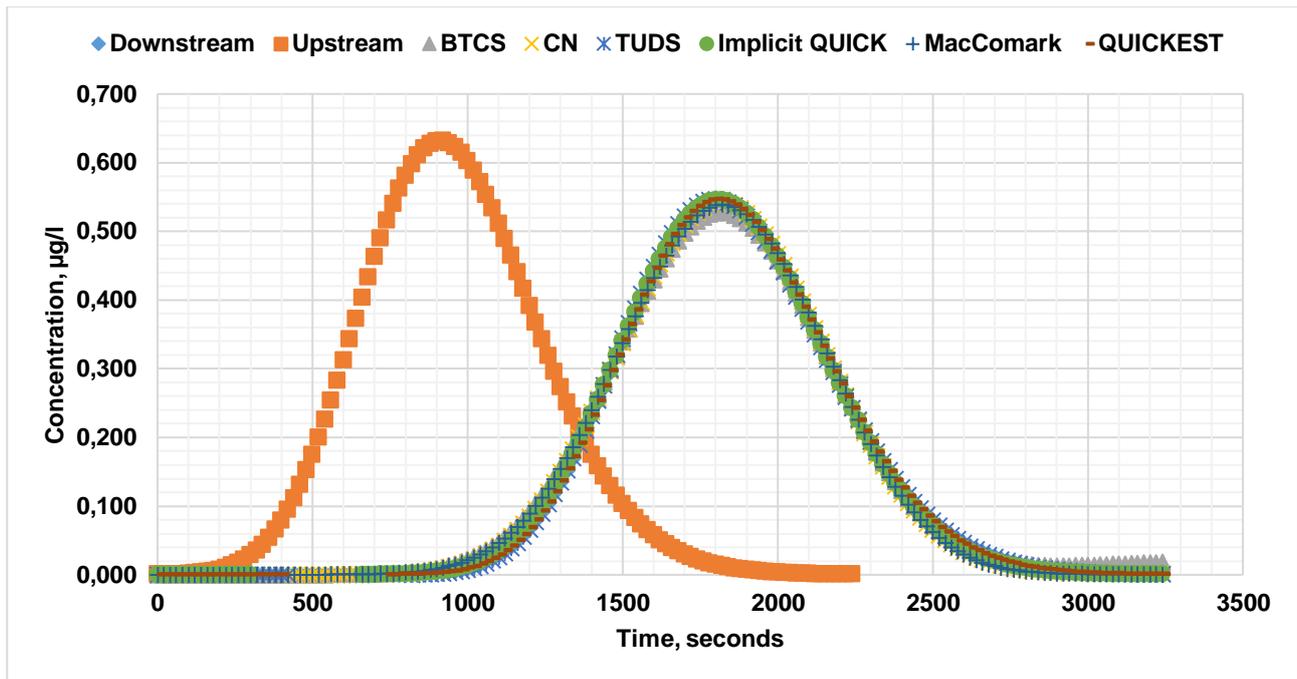


Figure 3-8: Simulation results of optional methods using space step of 20 m

3.9 Conclusion

Six different numerical methods were applied to synthetic data under the same numerical discretisation and non-dimensional numerical properties and optimised longitudinal dispersion coefficients and velocities were obtained, listed in Tables 3-1 and 3-2, respectively. The ranges of advection number, dispersion number and Peclet number covered were 0.900 – 0.113, 0.600 – 0.009 and 1.5 – 12.0. The methods yielded a range of values of longitudinal dispersion coefficients. Some of which were notably different to the value used to generate the synthetic data. However, similar velocity values to the analytical ones were reproduced by all the methods over the whole range of grid resolutions and dimensionless numerical properties. The behaviour of the methods was consistent with the known presence of artificial numerical dispersion in the BTCS, Implicit QUICK, QUICKEST and TUDS methods. Up to a Peclet number of 6.00, all the numerical methods were able to simulate the generated data well, although simulations were increasingly inaccurate for higher Peclet numbers. However, this was accomplished by adjusting the dispersion coefficient to allow the presence of numerical dispersion. Therefore, if dispersion coefficients are to be estimated by optimising numerical methods, there is a risk of underestimating or overestimating the dispersion

coefficient if the numerical method induces considerable amounts of numerical dispersion. Overall, the best performing methods were the Crank-Nicolson and MacCormack methods. The QUICKEST method was quite good, but only for a narrow range of Peclet number (2 to 4). The Implicit Quick and Backward-Time/Centred-Space (BTCS) and third-order upstream-differencing schemes (TUDS) were much less accurate.

Additionally, methods such as the Implicit QUICK would not be appropriate for application to a small stream such as the Murray stream. In respect to the residual sum of squares between predicted and synthetic concentration profiles at the downstream boundary, simulation errors over the range of the Peclet number were quite low for all the methods. This shows that, for accurate predictions, different methods would require different parameter values, especially the dispersion coefficient. Based on accuracy in predictions and optimised values of dispersion coefficients the CN, MacCormack and QUICKEST were selected for analysis of observed BTCs.

4 DATA AND ANALYSIS OF OBSERVED BREAKTHROUGH CURVES

4.1 Introduction

The basic components this chapter are analysing data by converting instrument readings to concentration-time profiles (BTCs) and analysing BTCs to estimate stream velocity and longitudinal dispersion coefficient values using shortlisted numerical methods over a range of numerical properties. The data was obtained in a raw form consisting of instrument readings of Rhodamine tracer versus time collected from experiments conducted on Murray stream that passes through Heriot-Watt University, Edinburgh campus, Scotland (Heron, 2015).

Parameter estimation using concentration-time profiles using numerical methods required initial search values of parameters. Initial search values were obtained by the Singh and Beck (2003) routing procedure. The method has been observed to be comparatively more accurate than the Fischer's method and the Hayami solution (Wallis et al., 2013). After testing, the routing method was used to estimate the values of velocity and dispersion coefficient from observed BTCs. The estimated parameter values were later used as initial search values for numerical methods.

To investigate the impact of numerical methods on estimated parameter values three numerical, namely, Crank-Nicolson, MacCormack and QUICKEST methods were used to estimate the velocity and dispersion coefficient. To investigate the impact of numerical properties on estimated parameter values, parameters were estimated over a range of numerical properties by varying space step. Analysis of BTCs was achieved by simulating observed downstream BTCs by adjusting both velocity and longitudinal dispersion coefficient.

4.2 Data analysis

This study used data that were collected by Heron (2015). As stated above, experimental work was conducted on the Murray stream which flows through the Riccarton campus of Heriot-Watt University in Edinburgh. This work followed the previous one undertaken by Burke (2002) which investigated travel times of solutes in the stream. The new experiments were required to obtain more data from a range of flow rates to estimate dispersion coefficients over several flow conditions on one of the

reaches used previously which extends between Site 3 and Site 4 of the sitemap shown in Figure 4-1. Table 4-1 shows the characteristics of the sub-reach together with other sub-reaches of the site.

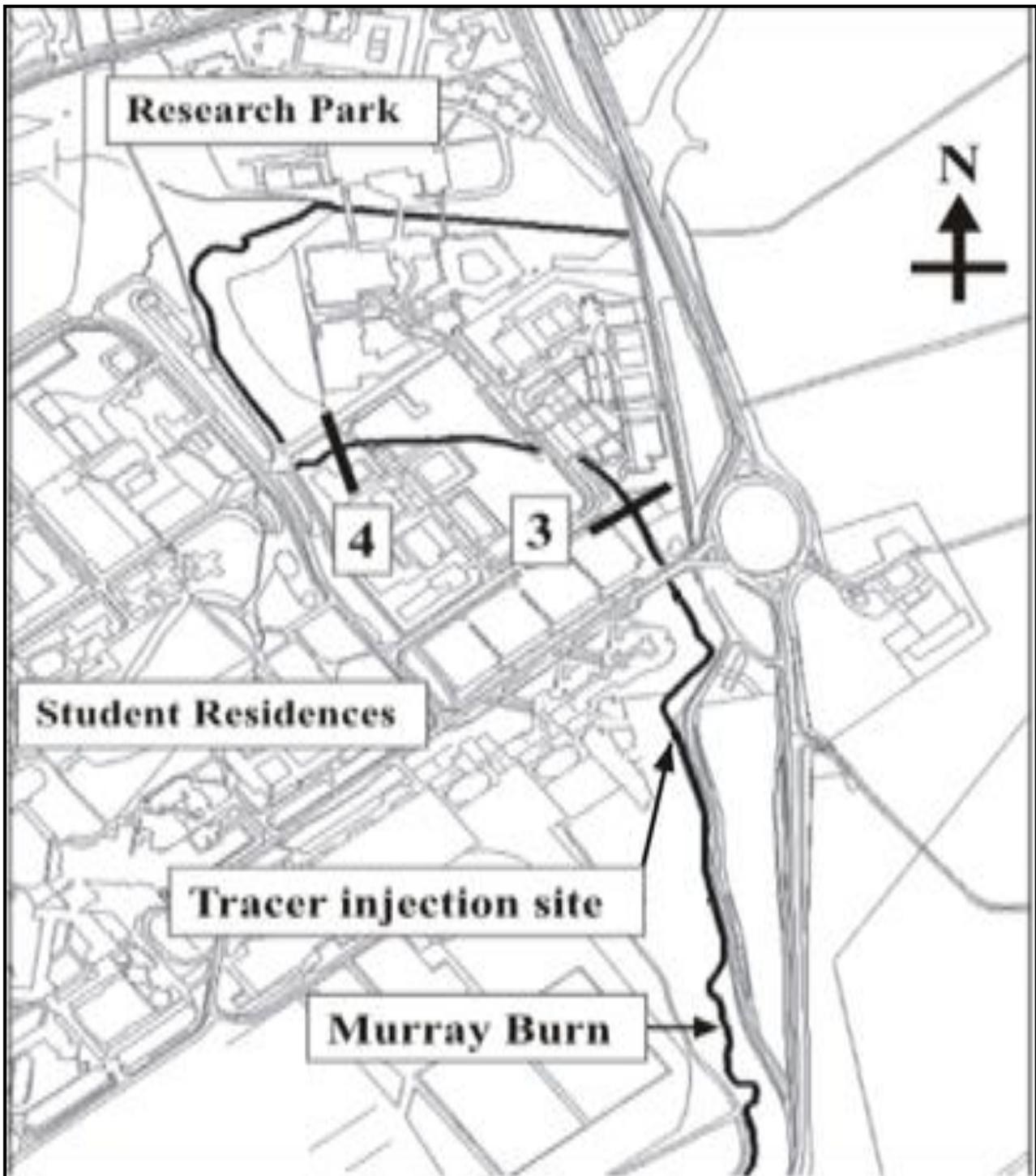


Figure 4-1: Study site (Burke, 2002)

Table 4-1: Physical characteristics of sub-reaches of the study site (Ani et al., 2009)

Reach number	Boundaries(upstream-downstream)	Reach length	Mean width	Mean slope (m/km)	Other characteristics
1	Source - Site 1	120 m	3.6 m	3.9	Boulders
2	Site 1-Site 2	137 m	3.7 m	24.0	Boulders; Has two bends
3	Site 2 - Site 3	99 m	3.3 m	16.4	Straight; The upper fifth is wider and shallower than the lower four-fifths
4	Site 3 - Site 4	184 m	2.4 m	9.3	Cobbles; Has a long curve; two upper thirds are wider and shallower than lower third

The data was collected using Rhodamine WT as a tracer and analysed using the Turner Designs model 10 fluorometer (Heron, 2015), shown in Fig. 4-2.



Figure 4-2: Turner Designs 10 fluorometer (Heron, 2015).

Thirteen experiments, listed in Table 4-2, were conducted over a period of 15 months. All experiments consisted of a gulp release of a known mass of tracer at an injection site 326 m upstream of site three shown in Figure 4-1. This was followed by a sampling of water for analysis from Sites 3 and 4. Rhodamine WT is a soluble, non-toxic, fluorescent dye tracer that was used in

all the experiments. Rhodamine WT is a well-known tracer in both laboratory and field tracer studies (Kilpatrick et al., 1989; Wallis, 2005). The first experiment was not used in this study as it was a trial experiment and it was not successful.

Processing of the data involved converting instrument readings into concentration, smoothing concentration profiles, removing background concentration and adjusting for any real or apparent non-conservative behaviour of the tracer. The background fluorescence signal of the stream is attributed to the naturally occurring material (e.g. Kirkpatrick, 1986), and it needs to be removed during analysis. Since the modelling process assumed that the tracer behaved conservatively, the concentration profiles were scaled such that the area under the downstream BTC was the same as the area under the upstream BTC in all the experiments. The instrument readings were then converted to concentration values by multiplying with the calibration factor which was determined earlier (Heron, 2015). Figure 4-3 and 4-4 are plots of instrument readings versus time and concentration versus time for experiment 2, respectively. Plots of analysed concentration-time profiles are given in Appendix A2.

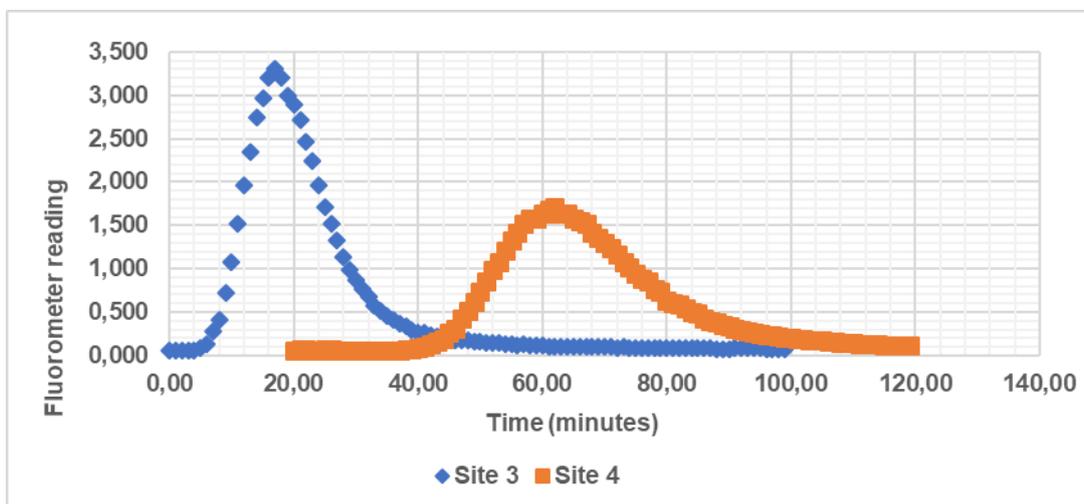


Figure 4-3: Fluorometer reading versus Time (experiment 2)

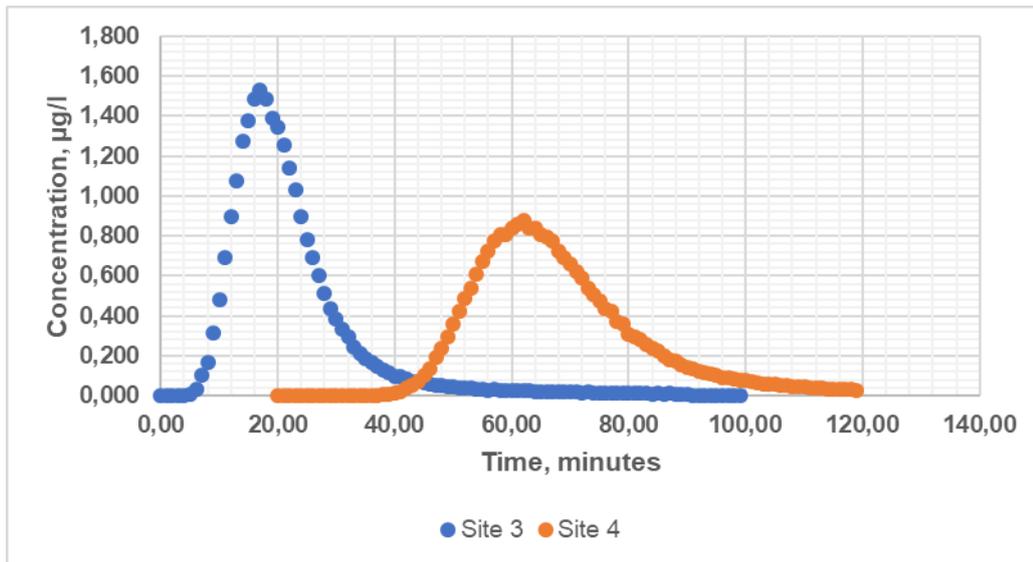


Figure 4-4: Concentration versus Time (experiment 2)

4.3 Estimating stream flow rates

Streamflow rate is one of the factors that influence the advection and dispersion of solutes (Rutherford, 1994). Flow rates could either be obtained from a calibrated weir or estimated using dilution gauging, as the stream is too small for the use of other methods such as the current meter. Ani et al. (2009) observed that the weir was not wide enough to cater for high flow rates. Consequently, Ani et al. (2009) used dilution gauging to estimate flow rates. Generally, the stream has favourable conditions for the use of dilution gauging to estimate flow rates. The conditions favourable for the use of dilution gauging to estimate stream flow rate are documented in several publications, e.g. the United States Geological Survey (Kilpatrick and Cobb, 1985). Hence, this study used the dilution gauging technique to estimate flow rates. The method is based on mass conservation of a released dispersant that is cross-sectionally well mixed. The basic principle of this method is that when a known mass of a conservative tracer is injected upstream of a sampling point and the concentration profile observed at a sampling point downstream is used to determine the stream flow rate. The mass of tracer at the sampling site is calculated using the following equation (Wallis, 2005):

$$M = Q \int_{t_1}^{t_2} \varphi(t) dt, \quad (4.1)$$

where M is a mass of tracer, Q is the stream flow rate, φ is the observed concentration at sampling site and t is time. The expression $\int_{t_1}^{t_2} \varphi(t) dt$ is the area under the concentration-time profile, and it represents the total mass of tracer passing through the sampling point. The flow rate is then computed using the expression,

$$Q = \frac{M}{\int_{t_1}^{t_2} \varphi(t) dt}. \quad (4.2)$$

The results of the dilution gauging technique depend on the tracer being completely mixed cross-sectionally at the sampling point with minor loss of the tracer (Wallis, 2005). The flow rates were computed using data at site 4. One of the advantages of using data from site 4 is that there is a possibility of the tracer not being fully mixed at site 3, especially at high flow rates.

Table 4-2: Summary of experiments (Heron, 2015) and flow rates.

Experiment number	Date of experiment	Mass of tracer (mg)	Flow rate at site 4 (m ³ /s)
2	2009/10/28	50	0.036
3	2009/11/04	100	0.147
4	2009/11/11	50	0.084
5	2009/11/18	100	0.097
6	2009/11/26	200	0.385
7	2010/05/14	100	0.041
8	2010/05/27	75	0.044
9	2010/06/17	50	0.037
10	2010/07/08	50	0.017
11	2010/11/03	150	0.150
12	2011/02/08	300	0.456
13	2011/02/15	150	0.148

4.4 Estimating stream parameter values.

4.4.1 Introduction

The Analysis of observed BTCs to determine solute transport parameters was achieved using the Singh and Beck routing procedure and shortlisted numerical methods. The Analysis of observed BTCs using the AD-Model is an inverse modelling process to estimate cross-sectional average velocity and longitudinal dispersion coefficient (Wallis et al., 2013). Observed BTCs were used for model development, calibration and confirmation of developed models.

By the common practice (e.g. Ani et al., 2009), the observed BTCs were divided into two sets: predictive model development BTCs and predictive model confirmation BTCs. However, all BTCs were analysed for parameter estimation by shortlisted numerical methods for purposes of statistical analysis. There are several approaches to data division for model development and testing such as the cross-validation, re-substitution, bootstrapping etc. The re-substitution approach has the disadvantage that the same data are used for model development, calibration and model confirmation (Bennett et al., 2013). Bootstrapping involves repeated random re-sampling with the replacement of the original measurements. This process is repeated for all data points each one in turn being left out. The method has been observed to be extremely biased in some problems (Kohavi, 1995). In the cross-validation, method data are split into specific groups for model development and performance analysis, and the common method being the holdout method. In the holdout method, data are divided into two groups, one for model development and the other for model evaluation. Reliability of model testing is influenced by both the size and position of group division of the data (Kohavi, 1995; Bennett et al., 2013). Therefore, it is necessary to consider the extent to which the model evaluation data are independent of the model development and calibration data (Bennett et al., 2013). The holdout method was used in this study to divide the data for model development and model confirmation.

Generally, a quantitative performance measure normally gives a single value for the whole data set, which can disguise noteworthy unknown behaviour over the intervals of data sampling. Therefore, data may be divided deductively or divided into low, medium and high events (Bennett et al., 2013). In this study, the holdout method of data division was used. This method was also used by Ani et al. (2009), where evaluation data was obtained from low, medium and high flow rates. Subsequently, observed BTCs were divided into two sets: models' development and calibration data (comprising nine experiments) and models' confirmation data (comprising three experiments). The model development data was used for identifying models' structure and calibration.

When estimating parameter values using numerical methods, especially implicit method, unless the optimisation process starts with a good initial estimate (or guess), the estimation may take an unnecessary amount of time to converge (Dattner, 2015). The Singh and Beck routing procedure

was used to determine the initial values of the parameters. These parameter values were used in numerical models to make the first concentration predictions, which were the starting point for parameter estimation. Values of velocity and dispersion were determined by solving a non-linear least-squares optimisation problem. The algorithm is based on the minimisation of the objective function computed as the residual sum of squares (RSS) between the predicted and observed concentration-time profiles. The nonlinear least-squares problem was solved using Excel. However, before estimating initial parameter values, it was prudent to assess how well the routing procedure could estimate the solute transport parameters.

4.4.2 Testing the routing procedure and computation of initial parameter values.

The Singh and Beck routing procedure was tested using synthetic data. Synthetic data was generated using Equation (2.5) which is Taylor's analytical solution of the AD-Model. The data were generated using tracer mass of 100 mg, the channel width of 1.0 m² and a hypothetical reach length of 200 m for time steps 20 s, 30 s, 40 s, 60 s, and 90 s, respectively. The time steps used to assess the routing procedure were those used for sampling observed BTCs. Parameter values used to generate data were $v = 0.40$ m/s and $D = 0.8$ m²/s. These values are in the range of values observed in previous studies. Like the numerical methods in Chapter 3, the routing procedure was tested if it could accurately predict concentration distribution and reproduce the parameter values used with the analytical solutions. The performance of the routing procedure was also assessed by the RSS and the relative percentage error statistics. The RSS statistic was used to assess the performance of concentration prediction while the relative percentage error statistic was used to assess the accuracy of optimised parameter values (Chin, 2013).

4.4.2.1 Concentration prediction by routing method

Temporal evolution of the concentration at the downstream site because of time-varying concentration at the upstream site can be predicted by use of a solution of the AD-Model if parameter values are known. The temporal variation of concentration at the downstream section, $\varphi(X, t)$ as a result of time-varying concentration input at the upstream section, $\varphi_u(t)$ can be expressed by the following quadrature procedure (Singh and Beck, 2003):

$$\varphi(X, t) = \int_0^t \varphi_u(\tau) \xi(X, t - \tau) d\tau, \quad (4.3)$$

where X is the distance between the upstream and downstream points, $\varphi_u(t)$ is the upstream concentration at the time t and $\xi(X, \tau)$ is the downstream response as a result of the upstream unit concentration. Typically, the period of concentration evolution is discretised in time steps. Thus, Eq. (4.3) can be expressed in a discretised form as (Singh and Beck, 2003),

$$\varphi(X, l\Delta t) = \sum_{j=1}^l \bar{\varphi}_u(\lambda\Delta t) \mu\{(l - \lambda + 1)\Delta t\}, \quad (4.4)$$

where $\bar{\varphi}_u(\lambda\Delta t)$ is the mean concentration over the l^{th} time step and $\mu(*)$ is a pulse of unit concentration for the response, which may be expressed as

$$\mu(\kappa) = \psi(\kappa) - \psi(\kappa - 1), \quad (4.5)$$

where k , l and j symbolise time steps. The unit response at the end of the κ^{th} time step $\psi(k)$ is determined by Equation (2.15) for $\varphi_0 = 1$ and $x = X$ as (Singh and Beck, 2003; Wallis et al., 2013):

$$\psi(\kappa) = \frac{1}{2} \left[\operatorname{erf} \left(\frac{X - v\kappa\Delta t}{2\sqrt{D\kappa\Delta t}} \right) + \exp \left(\frac{vX}{D} \right) \operatorname{erfc} \left(\frac{X + v\kappa\Delta t}{2\sqrt{D\kappa\Delta t}} \right) \right]. \quad (4.6)$$

If values of velocity and longitudinal dispersion are known, the temporal evolution of concentration at the downstream site resulting from the concentration-time profile at the upstream site can be predicted using Equation (4.4).

4.4.2.2 Identification of parameters

Parameter identification by the routing method is achieved using Eq. 2.15 and Eq. 4.4. Thus Eq. 2.15 and Eq. 4.4 can be expressed as (Singh and Beck, 2003):

$$\varphi(X, l\Delta t) = g \left[X, \Delta t, D, v, \varphi_u(l\Delta t) \right], \quad j=1, 2, \dots, l, \quad (4.7)$$

where $\varphi_u(\bullet)$ is upstream concentration and $\varphi(\bullet)$ is predicted temporal concentration distribution

at the downstream section and l are measurement points. The optimisation can be achieved according to Eq. 3.26. After testing, the routing procedure was used to analyse observed BTCs to estimate velocity and longitudinal dispersion coefficients. The estimated parameters were also used in numerical methods as initial parameter values.

4.4.2.3 Results and discussion

In the assessment of the routing procedure, the optimised parameters were statistically examined, using relative error, for different time steps ranging from 20 s to 90 s. The routing procedure predicts the downstream profile by numerical integration. The time step is expected to influence the quality of simulations and computed values of velocity and dispersion coefficient (Wallis et al., 2013). This shows that depending on sampling time step different values would be obtained. Table 4.3 lists the results of calculated velocity and dispersion coefficient values and relative percentage errors obtained for several time steps by the routing procedure. As expected the calculated parameter values vary with time step as less accurate results were obtained with an increase in time steps. However, calculated dispersion coefficients show a higher variation of calculated values than calculated velocity values.

Figure 4-2 and 4-3 show simulated temporal concentration profiles for 40 s and 60 s time steps obtained by the routing procedure. The results show that accurate simulations could be achieved at different time steps, with very small values of the residual sum of squares.

Table 4-3: Analytical and optimised parameter values, and relative percentage error (R.E.) statistics.

Time step, seconds	Analytical D, (m ² /s)	Analytical V, (m/s)	Optimized D, (m ² /s)	Optimized V (m/s)	R.E. D, (%)	R.E. V, (%)
20	0.800	0.400	0.749	0.392	6.42	1.95
30	0.800	0.400	0.721	0.388	9.89	2.89
40	0.800	0.400	0.692	0.385	13.52	3.82
60	0.800	0.400	0.631	0.378	21.17	5.60
90	0.800	0.400	0.533	0.367	33.38	8.16

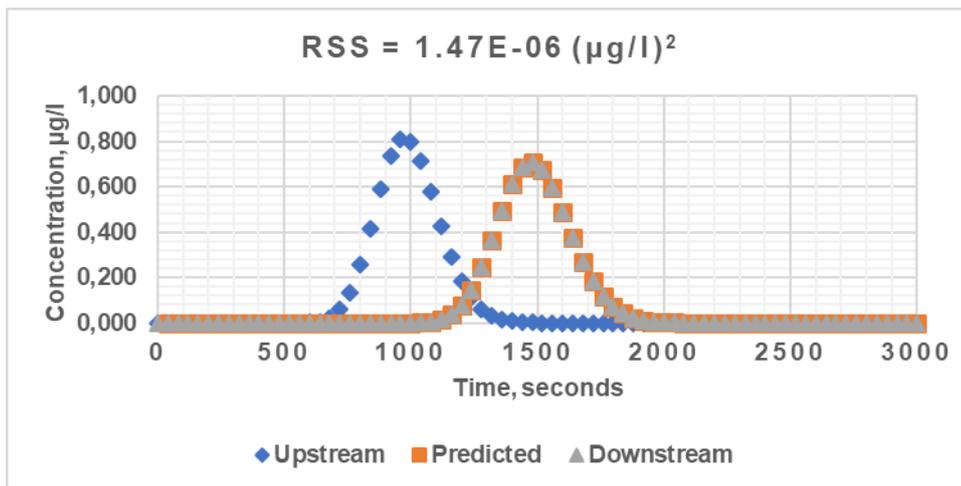


Figure 4-5: Simulation result of the routing procedure for a time step of 40 seconds.

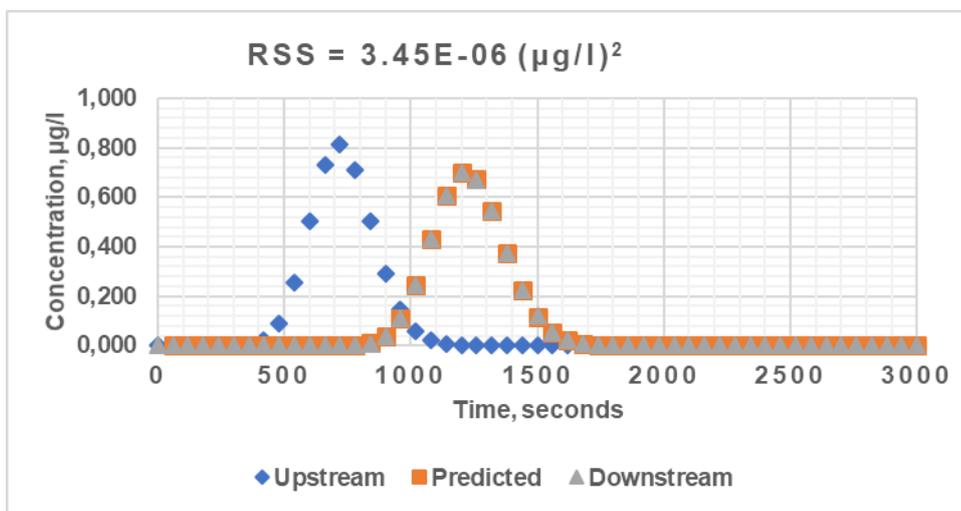


Figure 4-6: Simulation result of the routing procedure for a time step of 60 seconds.

Table 4-5 lists the results of optimised values of velocity and longitudinal dispersion coefficient obtained from the analysis of measured BTCs using the routing procedure. These were also used later as the initial search values for numerical methods. As expected the optimised parameter values increase with flow rate (Rutherford, 1994). The computed values were also compared with results obtained earlier by the Fischer's routing procedure, given in Table 4-4 (Heron, 2015). Fischer's routing procedure tends to estimate higher values than the Singh and Beck routing procedure (Wallis et al., 2013) However, the values obtained by Singh and Beck were in the same range as those obtained by the Fischer's routing method.

Figures 4-5 and 4-6 show examples of simulations of downstream observed BTCs by the routing procedure. Similar results were observed for all experiments used in this study. The results show

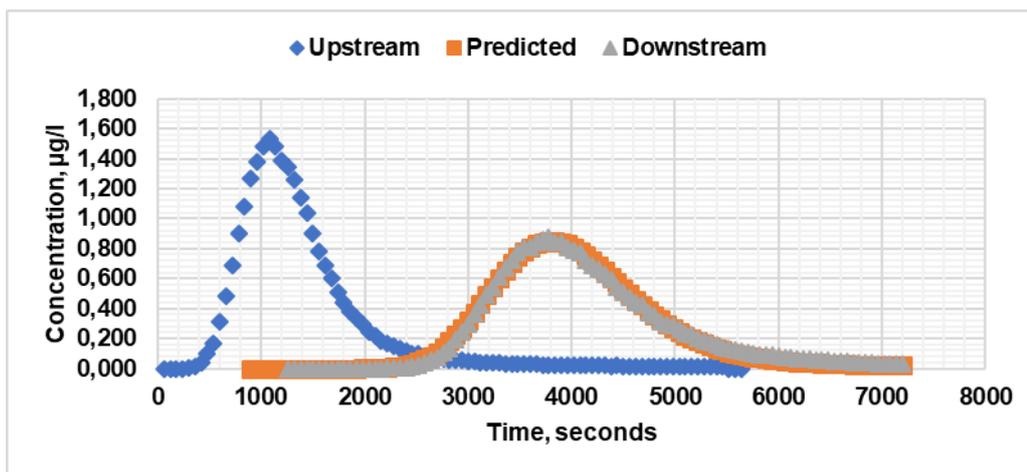
that accurate simulations were achieved at all sampling time steps which ranged from 20 seconds to 90 seconds. However, optimised parameter values obtained by routing methods are dependent on time step; they are less accurate with an increase in time step as listed in Table 4-3.

Table 4-4: Computed parameter values from a previous study using Fischer's routing procedure (Heron, 2015).

Experiment number	Velocity (m/s)	Dispersion coefficient (m ² /s)
3	0.197	0.538
4	0.148	0.714
5	0.162	0.805
6	0.347	0.709
7	0.091	0.426
8	0.097	0.445
9	0.083	0.380
10	0.041	0.202
11	0.191	0.868
12	0.389	0.920
13	0.214	0.714

Table 4-5: Optimized parameter values obtained by Singh and Beck routing procedure.

Experiment number	Time step (sec.)	Velocity (m/s)	Dispersion coefficient (m ² /s)
2	60	0.067	0.232
3	60	0.208	0.566
4	30	0.147	0.497
5	30	0.160	0.561
6	20	0.343	0.773
7	60	0.089	0.279
8	90	0.095	0.275
9	60	0.081	0.249
10	60	0.041	0.117
11	30	0.190	0.630
12	30	0.381	0.773
13	40	0.212	0.525



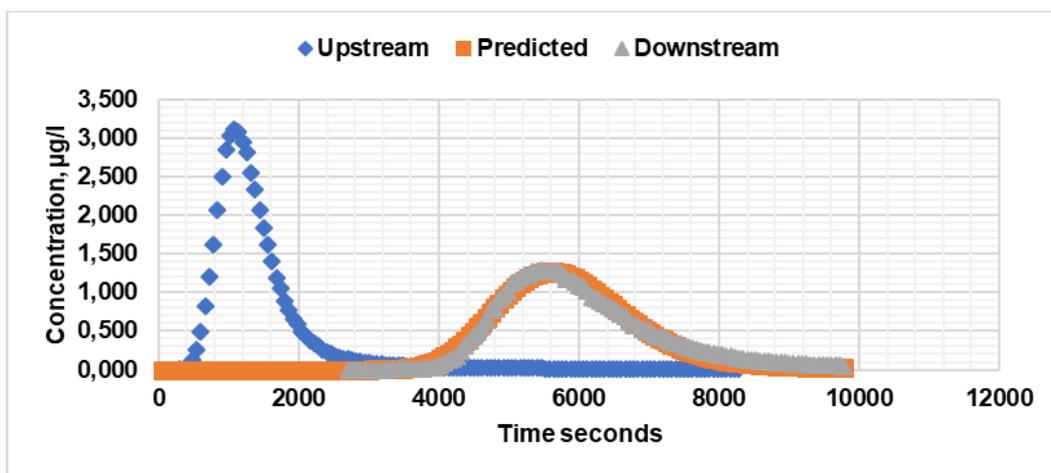
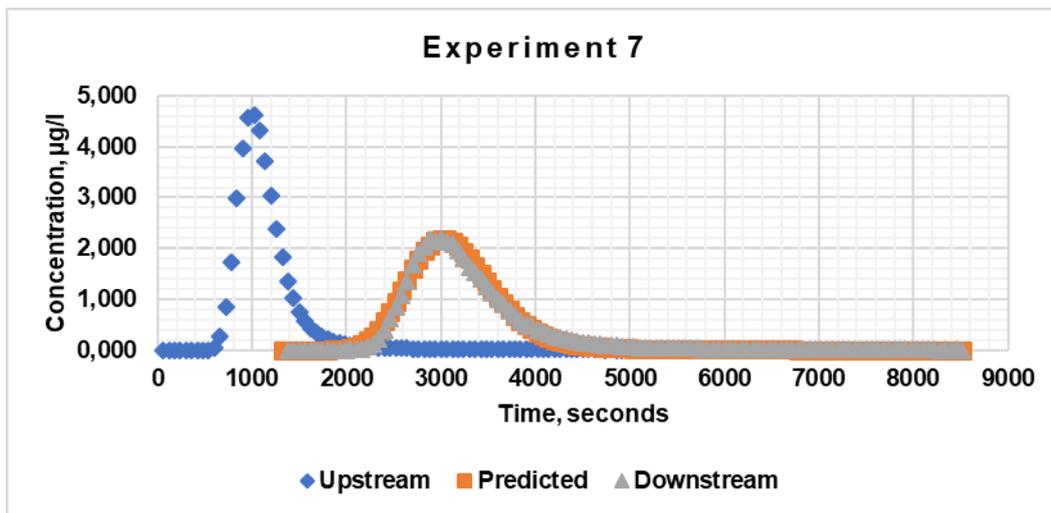


Figure 4-7: Examples of simulation results during estimation of initial parameter values.

4.4.3 Numerical analysis of observed breakthrough curves.

Based on the assessment of methods in Chapter 3, three methods were shortlisted for analysis of observed BTCs, namely, the Crank-Nicolson, the MacCormack and the QUICKEST method. The numerical methods were applied to start with the initial values obtained by the Singh and Beck routing method. Computed values of velocity and dispersion coefficients are listed in Table 4-5.

Like Chapter 3, Excel was used to analyse observed BTCs. The input variables were Δx , Δt , V and D as shown in Fig. 3-1. The input values determined the values of non-dimensional numerical properties. The equations were written to the cells depending on the method used. The solution of each model was carried out based on computational algorithms expressed for each cell. Before

starting the modelling, initial conditions, boundary conditions and initial parameter values were specified. The results obtained were simulated temporal concentration-time profiles at the downstream site, computed velocities and computed dispersion coefficients. The RSS (i.e. the objective function) was minimised to achieve the best fit between the downstream concentration profile and the predicted profile. The process was repeated for several space steps. The inverse modelling process is shown in the flowchart in Figure 4-9. Inverse modelling, also known as calibration, involves adjusting parameter values to achieve an ideal agreement between the model computations and the downstream observed BTC (Chapra, 2008; Chin, 2013).

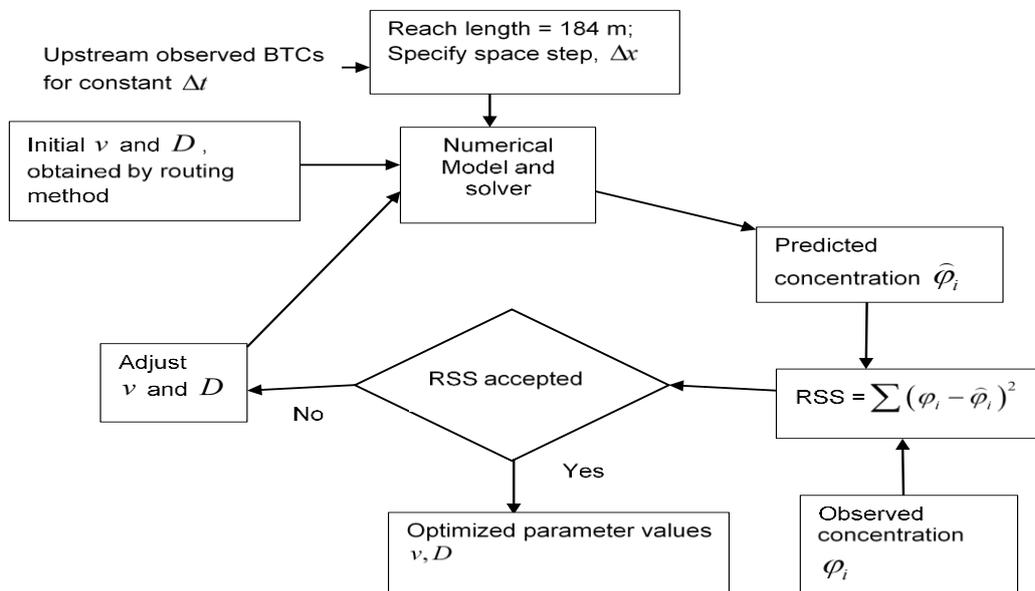


Figure 4-8: A Flowchart for analysis of observed BTCs, adapted from Chapra (2008).

To observe the influence of numerical properties, observed BTCs were analysed over a range of numerical properties, by varying space step. Since the sampling frequency of each experiment fixed by the time step, it was only possible to vary the space steps. The use of several methods and space steps enabled the investigation of the influence of numerical properties and numerical methods on estimated parameter values.

4.4.4 Results and discussion

There were problems of lack of stability and boundedness with the QUICKEST method at small

space steps as the scheme is conditionally stable depending on the combination of the advection number and the dispersion number (Leonard, 1979). The results obtained from the analysis were concentration simulations, optimised longitudinal dispersion coefficients and velocities. Concentration simulations were measured by the residual sum of squares (RSS).

Table 4-5 to 4-7 list computed dispersion coefficients obtained over a range of space steps and flow rates obtained by the Crank-Nicolson, the MacCormack and the QUICKEST schemes respectively. Optimised parameter values increase with discharge as expected (Rutherford, 1994). It can be observed that there is a variation of the computed dispersion coefficients with the numerical method and the space step. Since for each experiment values of velocity and dispersion coefficient are constant, increase in space step results in an increase in Peclet number.

Table 4-4: Computed longitudinal dispersion coefficients and flow rates (Crank-Nicolson)

Discharge, m ³ /s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.15	0.436	0.148
Δt , seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment No.	2	3	4	5	6	7	8	9	10	11	12	13
Δx , meters												
2.00	0.232	0.688	0.497	0.561	0.839	0.297	0.309	0.263	0.179	0.669	0.905	0.581
4.00	0.241	0.690	0.520	0.590	0.841	0.297	0.309	0.264	0.184	0.669	0.907	0.582
4.60	0.241	0.691	0.520	0.590	0.842	0.298	0.310	0.264	0.184	0.670	0.909	0.582
5.75	0.241	0.692	0.520	0.591	0.844	0.298	0.310	0.265	0.184	0.670	0.911	0.583
7.36	0.242	0.695	0.521	0.592	0.849	0.300	0.312	0.267	0.185	0.671	0.917	0.584
8.00	0.242	0.697	0.521	0.593	0.852	0.300	0.312	0.267	0.185	0.671	0.920	0.585
8.36	0.242	0.697	0.521	0.594	0.853	0.301	0.313	0.268	0.186	0.672	0.922	0.586
9.20	0.242	0.700	0.522	0.595	0.858	0.302	0.314	0.269	0.186	0.672	0.927	0.587
11.50	0.244	0.707	0.525	0.599	0.872	0.305	0.317	0.273	0.189	0.676	0.773	0.572
12.27	0.245	0.709	0.527	0.601	0.878	0.307	0.318	0.274	0.190	0.677	0.953	0.595
13.14	0.245	0.713	0.529	0.603	0.886	0.308	0.320	0.276	0.191	0.679	0.962	0.598
14.15	0.247	0.717	0.532	0.605	0.895	0.310	0.322	0.278	0.193	0.681	0.974	0.602
16.73	0.250	0.728	0.539	0.613	0.921	0.316	0.327	0.283	0.197	0.688	1.006	0.614
18.40	0.252	0.736	0.545	0.618	0.940	0.320	0.331	0.287	0.200	0.694	1.029	0.623
Average D, m²/s	0.243	0.704	0.524	0.596	0.869	0.304	0.316	0.271	0.188	0.676	0.930	0.591

Computed dispersion coefficients obtained by the Crank-Nicolson and the MacCormack methods increase with space step, while those obtained by the QUICKEST method tend to decrease with the increase in space step. The trends of computed dispersion coefficients with an increase in Peclet number and space step are like those observed in Chapter 3. However, the optimised dispersion coefficients are the optimal values for each numerical method and numerical properties.

Table 4-5: Computed longitudinal dispersion coefficients and flow rates (MacCormack method)

Discharge, m ³ /s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.15	0.436	0.148
Δt , seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment No.	2	3	4	5	6	7	8	9	10	11	12	13
Δx , meters												
2.00	0.232	0.698	0.497	0.561	0.846	0.279	0.311	0.249	0.182	0.670	0.917	0.585
4.00	0.242	0.709	0.522	0.594	0.855	0.301	0.315	0.267	0.184	0.673	0.937	0.589
4.60	0.242	0.713	0.523	0.595	0.859	0.302	0.316	0.268	0.185	0.674	0.944	0.591
5.75	0.243	0.720	0.524	0.597	0.867	0.304	0.319	0.270	0.185	0.676	0.960	0.595
7.36	0.245	0.736	0.497	0.602	0.887	0.308	0.325	0.274	0.187	0.681	0.995	0.605
8.00	0.245	0.736	0.497	0.602	0.887	0.308	0.325	0.274	0.187	0.681	0.995	0.605
8.36	0.245	0.739	0.529	0.603	0.891	0.309	0.326	0.275	0.188	0.682	1.001	0.607
9.20	0.246	0.745	0.531	0.606	0.900	0.311	0.328	0.277	0.189	0.685	1.015	0.612
11.50	0.249	0.763	0.537	0.613	0.928	0.317	0.335	0.282	0.192	0.692	1.057	0.626
12.27	0.261	0.768	0.540	0.616	0.937	0.319	0.337	0.284	0.193	0.695	1.071	0.631
13.14	0.251	0.775	0.543	0.619	0.949	0.321	0.340	0.287	0.195	0.698	1.088	0.637
14.15	0.250	0.783	0.547	0.623	0.963	0.324	0.344	0.289	0.196	0.703	1.107	0.645
16.73	0.257	0.803	0.557	0.634	0.999	0.332	0.352	0.296	0.201	0.714	1.156	0.665
18.40	0.261	0.815	0.565	0.641	1.024	0.337	0.357	0.301	0.205	0.723	1.187	0.678
Average D, m ² /s	0.248	0.750	0.529	0.608	0.914	0.312	0.331	0.278	0.191	0.689	1.031	0.619

Table 4-6: Computed longitudinal dispersion coefficients and flow rates (QUICKEST method; UNS means unstable solution)

Discharge, m ³ /s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.15	0.436	0.148
Δt , seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment	2	3	4	5	6	7	8	9	10	11	12	13
Δx , meters												
8.00	0.243	0.649	0.523	0.594	0.840	0.305	0.299	0.263	0.183	0.676	UNS	0.577
8.36	0.242	0.649	0.523	0.594	0.840	0.305	0.302	0.263	0.200	0.676	UNS	0.579
9.20	0.241	0.649	0.520	0.591	0.838	0.304	0.304	0.262	0.182	0.674	0.858	0.582
11.50	0.236	0.649	0.510	0.581	0.821	0.291	0.305	0.255	0.179	0.663	0.894	0.580
12.27	0.234	0.656	0.504	0.576	0.811	0.295	0.304	0.252	0.178	0.658	0.896	0.576
13.14	0.231	0.660	0.498	0.569	0.797	0.291	0.301	0.248	0.177	0.650	0.894	0.570
14.15	0.227	0.662	0.489	0.560	0.777	0.286	0.298	0.243	0.175	0.640	0.888	0.562
15.33	0.221	0.660	0.476	0.547	0.750	0.279	0.291	0.235	0.172	0.625	0.874	0.549
16.73	0.213	0.651	0.459	0.529	0.710	0.268	0.282	0.225	0.169	0.604	0.848	0.529
18.40	0.203	0.633	0.436	0.504	0.655	0.245	0.267	0.211	0.164	0.574	0.803	0.499
20.44	0.187	0.599	0.401	0.467	0.572	0.232	0.244	0.190	0.157	0.530	0.729	0.454
23.00	0.165	0.538	0.352	0.414	0.450	0.202	0.210	0.161	0.147	0.465	0.608	0.384
Average D, m ² /s	0.220	0.638	0.474	0.544	0.739	0.275	0.284	0.234	0.174	0.620	0.829	0.537

Figures 4-10 to 4-12 show plots of optimised dispersion coefficient, obtained by the three methods, against Peclet number for experiments 10, 11 and 12, respectively. These plots are for experiments with the lowest, medium and highest flow rates. It should be noted that the values of the Peclet number were calculated using initial values determined by the routing procedure so that the results could be easily compared. The results were plotted against Peclet number as it combines the effects

of the advective flux and the dispersive flux. The plots show that computed dispersion coefficients vary with Peclet number and the numerical method used. Beginning with the initial values of the dispersion coefficient the Crank-Nicolson and the MacCormack methods give higher values with increase in Peclet number. The QUICKEST method gives values which decrease with the increase of the Peclet number. The tendency of the QUICKEST method to produce decreasing values of longitudinal dispersion can be ascribed to the presence of numerical dispersion, which varies with space step and advection number (Leonard, 1979). The QUICKEST method induces artificial dispersion which is dependent on space step and advection number. The induced artificial dispersion increases with space step and when advection number is decreasing (Leonard, 1979). Variation of optimised dispersion coefficient means that for a constant value of velocity and time step numerical dispersion increases with space step.

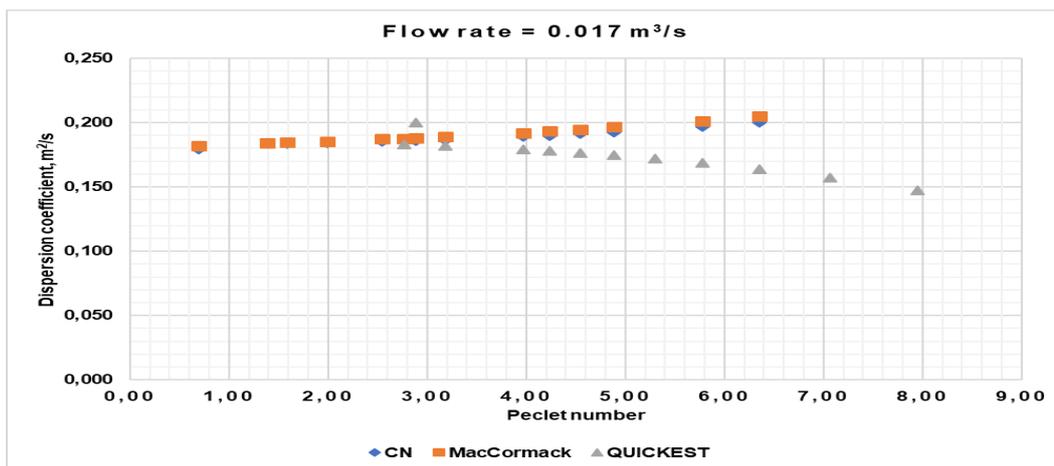


Figure 4-9: Computed dispersion coefficient versus Peclet number (experiment 10)

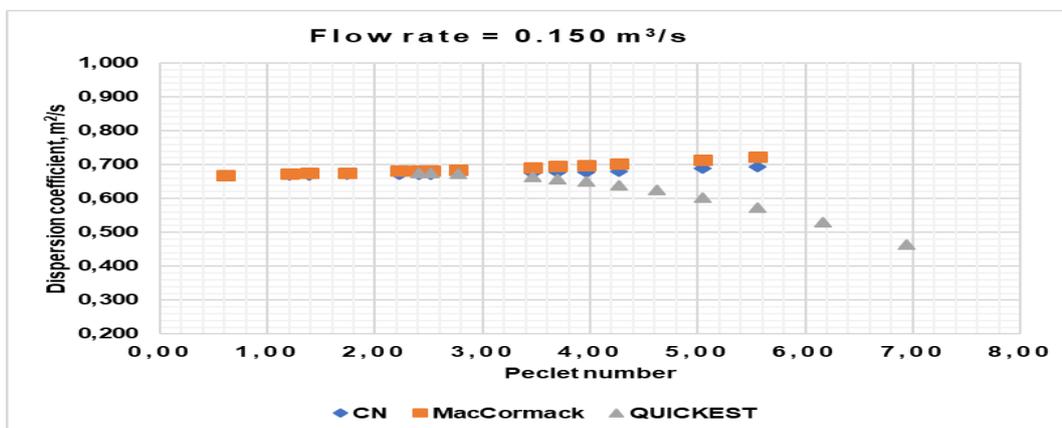


Figure 4-10: Computed dispersion coefficient versus Peclet number (experiment 11)

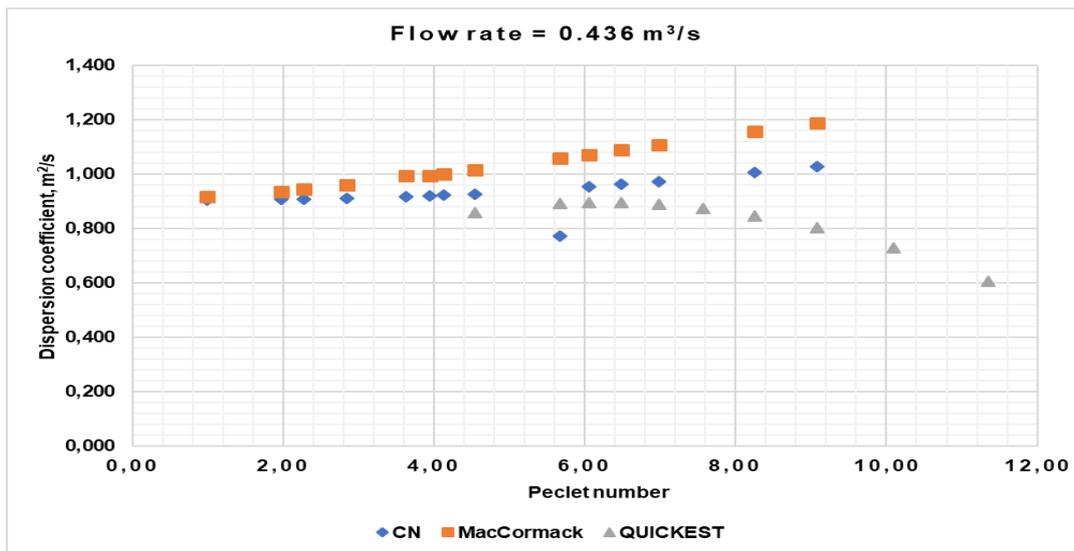


Figure 4-11: Computed dispersion coefficient versus Peclet number (experiment 12)

Tables 4-8 to 4-10 list computed values of velocity obtained over a range of space steps and flow rates. Computed values of velocity increase with the flow as expected. It can be observed that computed values of velocity vary with both numerical methods used and with the space steps (or Peclet number). However, the estimated velocities show comparatively small variations. These results are like those obtained in Chapter 3.

Table 4-7: Flow rates and computed velocity (Crank-Nicolson method)

Discharge, m³/s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.150	0.436	0.148
Δt, seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment No.	2	3	4	5	6	7	8	9	10	11	12	13
Δx, meters												
2.00	0.068	0.218	0.149	0.163	0.351	0.091	0.098	0.083	0.040	0.193	0.398	0.218
4.00	0.068	0.218	0.149	0.163	0.351	0.091	0.098	0.083	0.040	0.194	0.399	0.219
4.60	0.068	0.218	0.149	0.163	0.352	0.091	0.098	0.083	0.040	0.194	0.399	0.219
5.75	0.068	0.219	0.149	0.164	0.352	0.091	0.098	0.083	0.040	0.194	0.399	0.219
7.36	0.069	0.219	0.150	0.164	0.353	0.092	0.099	0.083	0.040	0.195	0.401	0.220
8.00	0.069	0.219	0.150	0.164	0.354	0.092	0.099	0.083	0.040	0.195	0.401	0.220
8.36	0.069	0.220	0.150	0.164	0.354	0.092	0.099	0.083	0.040	0.195	0.401	0.220
9.20	0.069	0.220	0.150	0.165	0.355	0.092	0.099	0.083	0.040	0.195	0.402	0.221
11.50	0.069	0.221	0.151	0.165	0.357	0.093	0.100	0.084	0.040	0.196	0.381	0.222
12.27	0.069	0.222	0.152	0.166	0.358	0.093	0.100	0.084	0.041	0.197	0.405	0.222
13.14	0.069	0.222	0.152	0.166	0.359	0.093	0.100	0.084	0.041	0.197	0.406	0.223
14.15	0.070	0.223	0.152	0.167	0.360	0.093	0.100	0.085	0.041	0.198	0.408	0.223
16.73	0.070	0.224	0.154	0.168	0.363	0.094	0.101	0.085	0.041	0.199	0.411	0.225
18.40	0.071	0.226	0.154	0.169	0.365	0.095	0.102	0.086	0.041	0.200	0.413	0.226
Average V, m/s	0.069	0.221	0.151	0.165	0.356	0.092	0.099	0.084	0.040	0.196	0.402	0.221

Table 4-8: Flow rates and computed velocity (MacCormack method)

Discharge, m ³ /s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.150	0.436	0.148
Δt , seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment No.	2	3	4	5	6	7	8	9	10	11	12	13
Δx , meters												
2.00	0.068	0.218	0.149	0.163	0.352	0.091	0.098	0.083	0.040	0.194	0.400	0.219
4.00	0.068	0.220	0.150	0.164	0.354	0.092	0.099	0.083	0.040	0.195	0.402	0.220
4.60	0.068	0.221	0.150	0.164	0.354	0.092	0.099	0.083	0.040	0.195	0.403	0.220
5.75	0.069	0.221	0.150	0.164	0.355	0.092	0.099	0.083	0.040	0.195	0.405	0.221
7.36	0.069	0.223	0.151	0.165	0.357	0.092	0.100	0.084	0.040	0.196	0.407	0.222
8.00	0.069	0.223	0.147	0.165	0.358	0.092	0.100	0.084	0.040	0.196	0.408	0.223
8.36	0.069	0.223	0.151	0.165	0.358	0.093	0.100	0.084	0.040	0.197	0.408	0.223
9.20	0.069	0.224	0.151	0.166	0.359	0.093	0.101	0.084	0.040	0.197	0.410	0.223
11.50	0.070	0.226	0.153	0.167	0.362	0.093	0.101	0.085	0.041	0.198	0.413	0.225
12.27	0.070	0.226	0.153	0.167	0.363	0.094	0.102	0.085	0.041	0.199	0.414	0.226
13.14	0.070	0.227	0.153	0.168	0.364	0.094	0.102	0.085	0.041	0.199	0.416	0.226
14.15	0.070	0.228	0.154	0.168	0.365	0.094	0.102	0.086	0.041	0.200	0.417	0.227
16.73	0.071	0.230	0.155	0.170	0.369	0.095	0.103	0.086	0.041	0.202	0.421	0.229
18.40	0.071	0.232	0.156	0.171	0.371	0.096	0.104	0.087	0.041	0.203	0.424	0.231
Average V, m/s	0.069	0.224	0.152	0.166	0.360	0.093	0.101	0.084	0.040	0.198	0.411	0.224

Table 4-9: Flow rates and computed velocity (QUICKEST method; UNS means unstable solution)

Discharge, m ³ /s	0.036	0.147	0.084	0.097	0.385	0.041	0.044	0.037	0.017	0.150	0.436	0.148
Δt , seconds	60	60	30	30	20	60	90	60	60	30	30	40
Experiment No.	2	3	4	5	6	7	8	9	10	11	12	13
Δx , meters												
8.00	0.068	0.220	0.149	0.164	0.353	0.091	0.101	0.083	0.040	0.195	UNS	0.221
8.36	0.068	0.220	0.149	0.164	0.352	0.091	0.099	0.083	0.042	0.195	UNS	0.221
9.20	0.068	0.220	0.149	0.163	0.352	0.091	0.099	0.083	0.040	0.194	0.402	0.220
11.50	0.068	0.220	0.148	0.162	0.350	0.091	0.098	0.082	0.040	0.193	0.399	0.219
12.27	0.068	0.219	0.148	0.162	0.349	0.091	0.098	0.082	0.040	0.193	0.398	0.218
13.14	0.068	0.218	0.148	0.162	0.349	0.090	0.098	0.082	0.039	0.193	0.397	0.218
14.15	0.068	0.218	0.148	0.162	0.348	0.090	0.097	0.082	0.039	0.192	0.396	0.217
15.33	0.067	0.217	0.147	0.161	0.347	0.090	0.097	0.082	0.039	0.192	0.395	0.217
16.73	0.067	0.216	0.147	0.161	0.346	0.090	0.097	0.081	0.039	0.191	0.394	0.216
18.40	0.067	0.215	0.147	0.161	0.345	0.090	0.097	0.081	0.039	0.191	0.392	0.216
20.44	0.067	0.214	0.146	0.160	0.344	0.089	0.096	0.081	0.039	0.190	0.390	0.215
23.00	0.067	0.213	0.146	0.160	0.343	0.089	0.096	0.081	0.039	0.190	0.389	0.214
Average V, m/s	0.068	0.217	0.148	0.162	0.348	0.090	0.098	0.082	0.040	0.192	0.395	0.218

Figures 4-13 to 4-15 show velocities obtained by the three numerical methods plotted against the Peclet number for experiments 10, 11 and 12, respectively. With the increase in Peclet number, the Crank-Nicolson and the MacCormack methods overestimate values of velocity while the QUICKEST method underestimates velocities. Although there is a variation in optimised velocity values, they are not markedly influenced by the method used and numerical properties.

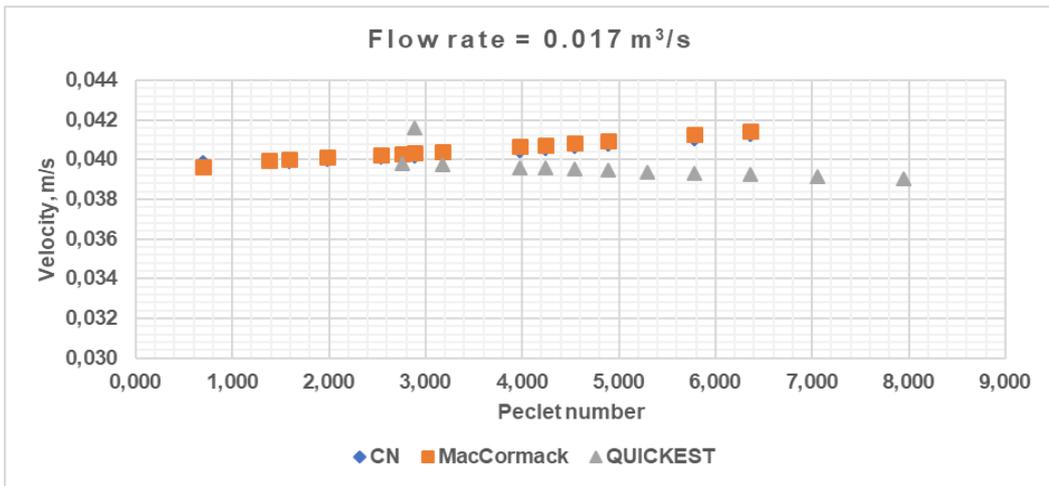


Figure 4-12: Computed velocity versus Peclet number for experiment 10

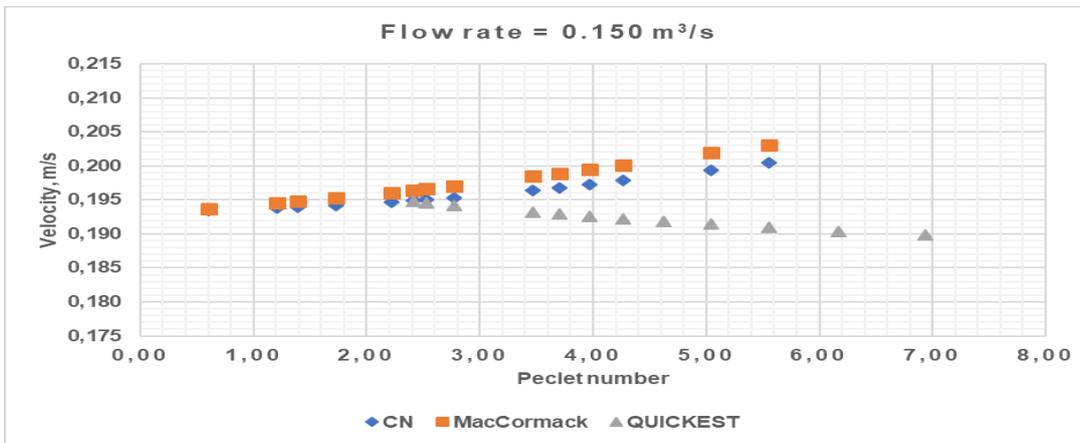


Figure 4-13: Optimised velocity versus Peclet number for experiment 11

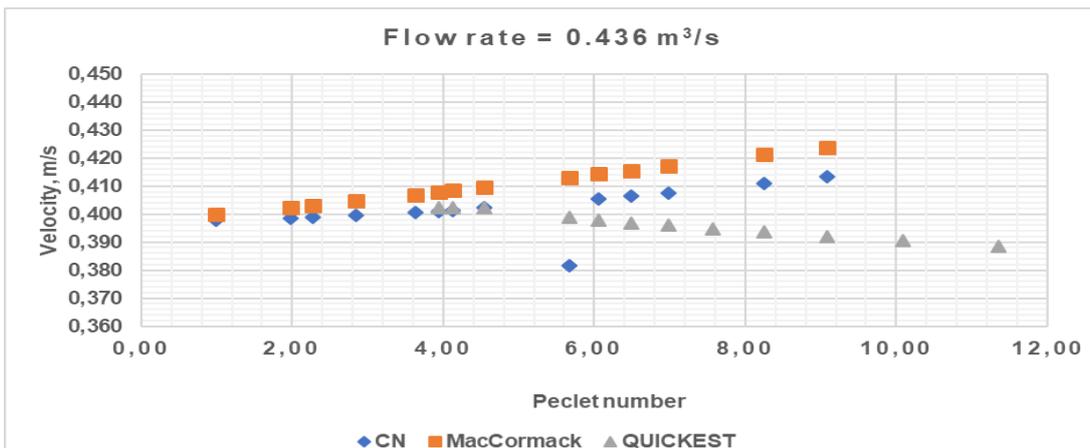


Figure 4-14: Optimised velocity versus Peclet number for experiment 12

Figure 4-16 to 4-18 show plots of the residual sum of squares versus Peclet number for experiments 10, 11, and 12, respectively. The values of the residual sum of squares varied consistently with the

optimised velocities and Peclet number.

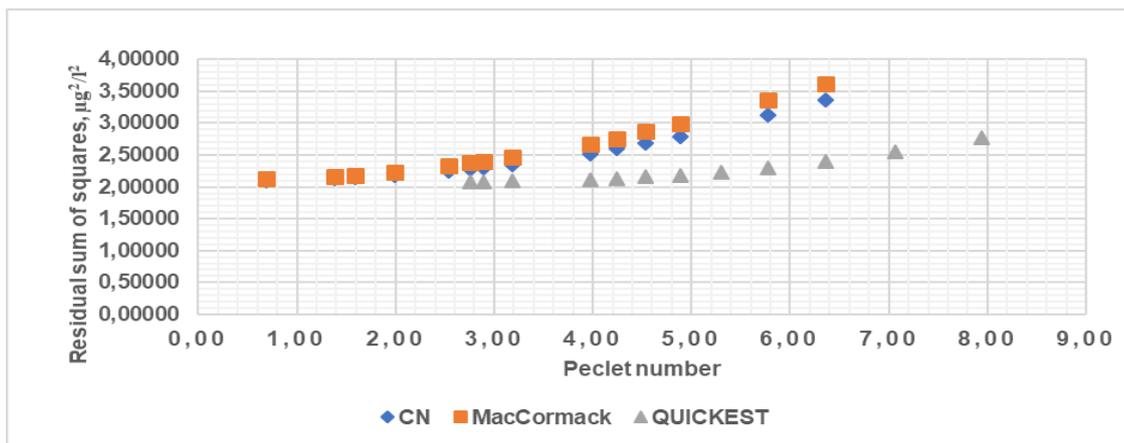


Figure 4-15: Residual sum of squares versus Peclet number for experiment 10

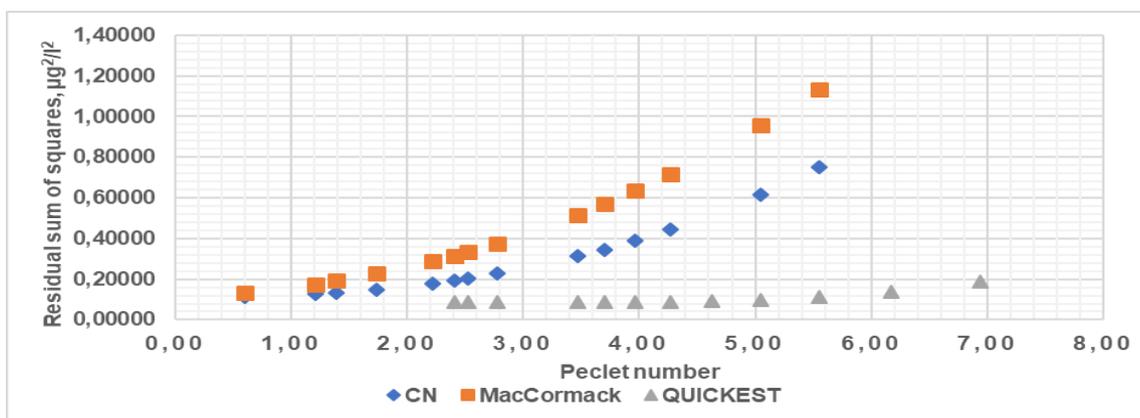


Figure 4-16: Residual sum of squares versus Peclet number for experiment 11

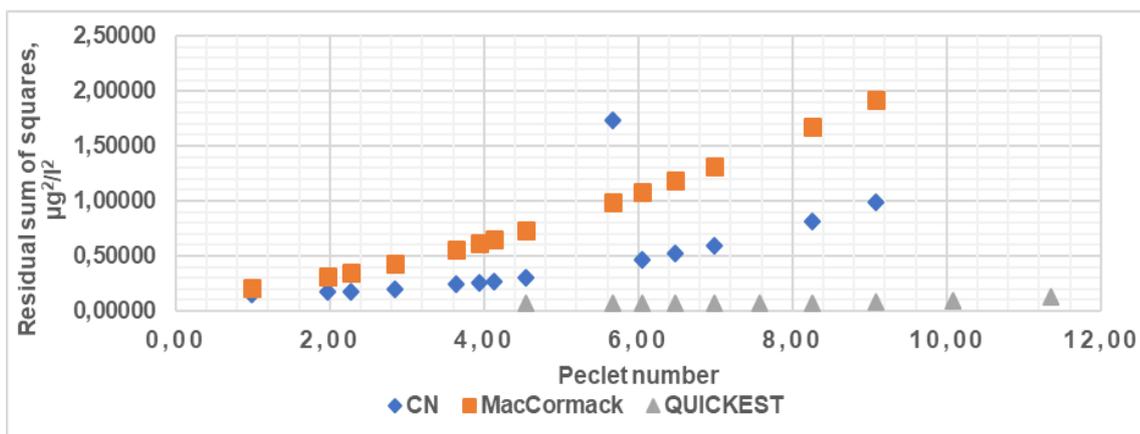
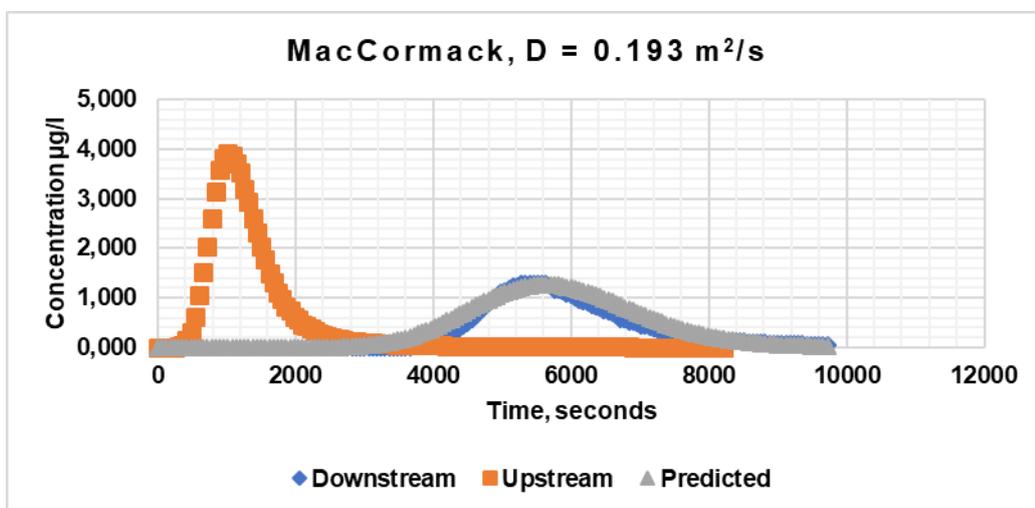
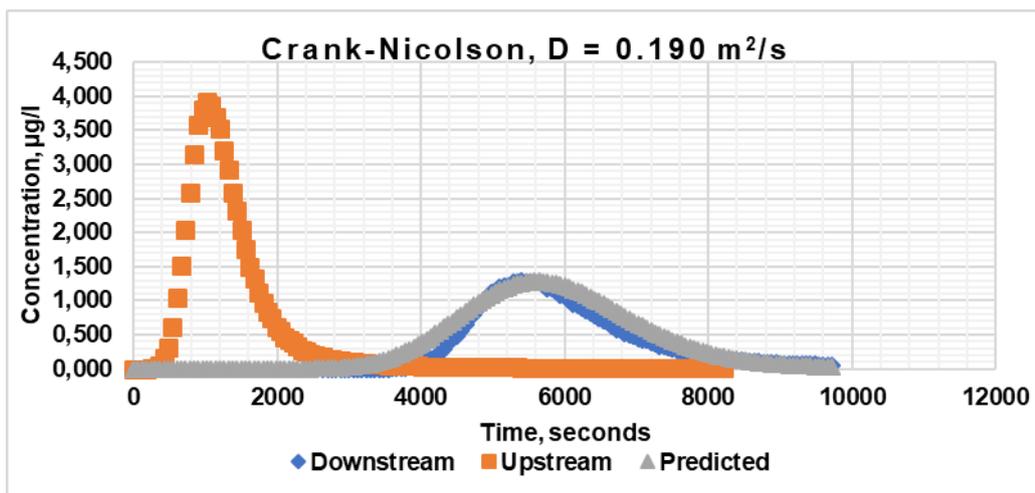


Figure 4-17: Residual sum of squares versus Peclet number for experiment 12

The MacCormack method gave the highest values of the residual sum of squares while the

QUICKEST gave the lowest values. The results also show that values of the residual sum of squares varied with flow rate, with the lowest flow rate resulting in highest values. The differences in accumulated residuals could be a result of differences in the number of time steps executed and on the resolution at which boundary conditions were defined. Comparatively, simulations by the QUICKEST method resulted in lowest accumulated residuals.

Figures 4-18 and 4-19 show simulated concentration-time profiles for experiments 10 and 12, which are for the lowest and highest flow rates, respectively, at space step 12.266 m. The results show variations in simulations and accumulated errors for the methods and experiments. The results show that accurate fits between observed and simulated profiles were achieved at low and high flow rates. Also, the figures demonstrate that accurate fits can be achieved with different parameter values, depending on the method used.



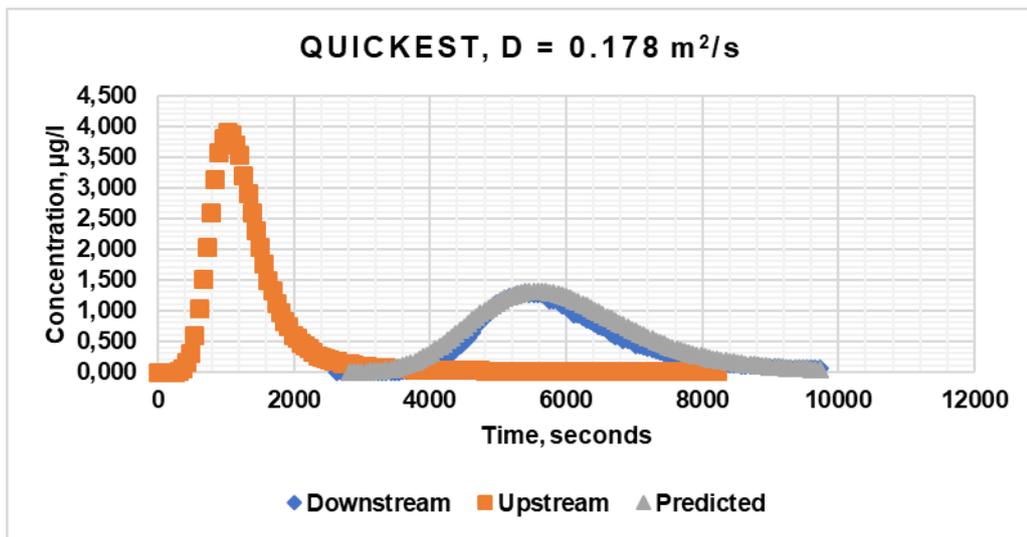
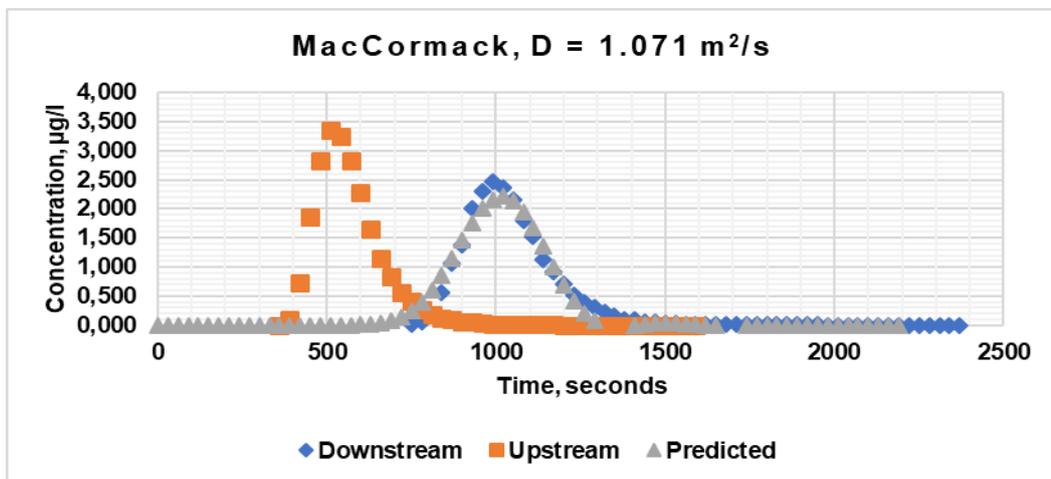
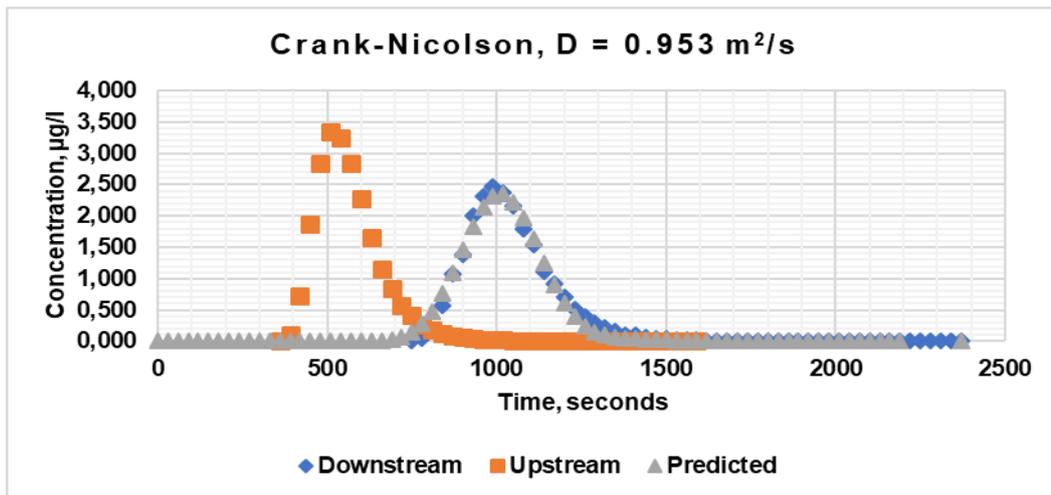


Figure 4-18: Fitting simulated and observed concentration-time profiles, space step = 12.266 m (experiment 10)



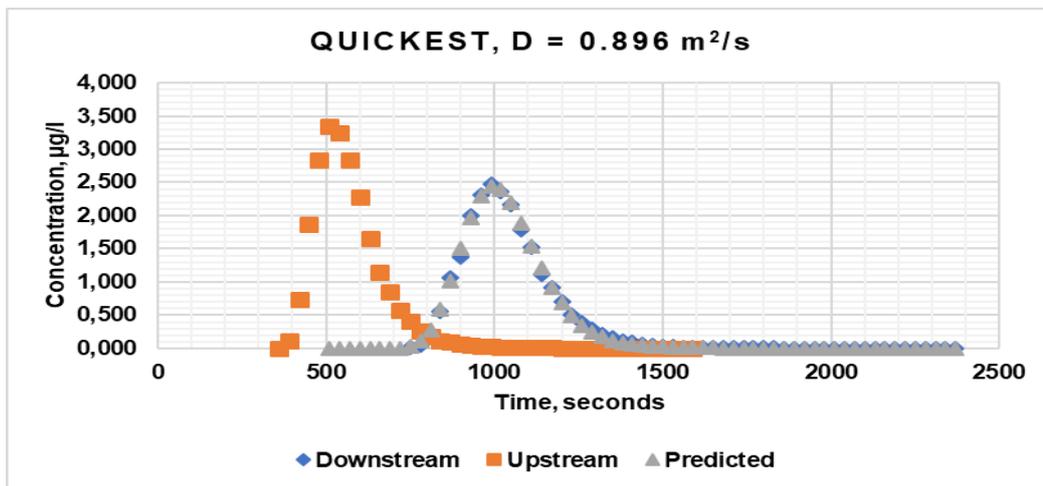


Figure 4-19: Fitting simulated and observed concentration-time profiles, space step = 12.266m (experiment 12)

4.4.5 Conclusion

Three methods were applied to observed temporal concentration profiles, namely, Crank-Nicolson, MacCormack and QUICKEST schemes. The methods were applied under similar space and time steps and calculated longitudinal dispersion coefficients and velocities were obtained. Time steps ranged from 20 seconds to 90 seconds and space steps ranged from 2.00 m to 23.00 m. The QUICKEST method gave unstable and unbounded solutions at space steps below 8.00 m in some experiments. In each experiment, the time step was fixed, and the space was varied. This was deliberately done to estimate parameters under different numerical properties. All the methods were able to simulate the observed BTCs, but the quality of simulations varied with space step such that poorer results were obtained with an increase in space step. The methods yielded a range of optimal values of dispersion coefficients which varied with the method and grid resolution. Computed dispersion coefficients showed high variations with the method used and Peclet number, while optimised velocity values showed much lower variations with the method used and Peclet number. Therefore, variations of estimated parameter values with the numerical method used and numerical properties demonstrate that estimated parameter values are influenced by the numerical method used and the model resolution. Additionally, high variations of calculated dispersion coefficients show that the dispersion coefficient is more influenced by the numerical method used and numerical properties than velocity.

5 EMPIRICAL MODEL DEVELOPMENT AND CALIBRATION

5.1 Introduction

Prediction of concentration evolution of pollutants in natural streams requires reliable pollutant transport parameter values. The selection of proper parameter values, especially a dispersion coefficient is the basic and the most demanding task (Seo and Cheong, 1998). The AD-Model is often used to model pollutant transport through its solutions. The common method for estimating pollutant transport parameters in natural streams is by simplified models in the form of empirical equations (Wallis and Manson, 2004). The parameter values determined by empirical equations are dependent on estimates of parameter values by experimental methods. Estimated parameter values obtained from measured concentration-time profiles by experimental methods together with the channel and hydraulic characteristics are commonly used to develop empirical models (Seo and Cheong, 1998; Kashefipour and Falconer, 2002; Fischer et al., 1979).

Empirical models were constructed based on each candidate numerical model and model resolutions. As per common practice, model construction process involved assembling data, model identification, model calibration and performance characterisation and model evaluation (Sun and Sun, 2015). The data for model construction consisted of the estimated dispersion coefficients obtained from model building BTCs in Chapter 4 and bulk flow characteristics. Model development data consisted of experiments 2, 3, 5, 6, 7, 8, 10, 11, and 12 from the Murray stream reach. The hydraulic characteristics considered were the flow rate, cross-sectional average velocity, the shear velocity and the stream flow depth. The flow depth was determined from the discharge and the wetted cross-sectional area, with the assumption that the wetted cross-sectional area was rectangular. The wetted cross-sectional area was determined from discharge and optimised values of velocity. The estimated parameter values together with hydraulic stream characteristics were later used to identify the functional form of the predictive model and subsequently the general model expression. Since the objective for developing models was for prediction purposes, it was necessary to assess the assumption that each set of observed dispersion was normally distributed. This was achieved using the Shapiro-Wilk hypothesis test.

According to the approach by several workers (e.g. Seo and Cheong, 1998; Falconer, Kashefipour and Falconer, 2002), the general functional form of empirical models were obtained using dimensional analysis. Identification of the functional form of the models was followed by developing a general regression model with unknown parameters. Unknown parameters were determined using regression analysis and calibration. Due to the influence of numerical methods and numerical properties on optimised solute transport parameters, unknown parameters were evaluated based on each numerical method and model resolution. Regression analysis was used to determine unknown regression parameters (coefficients and exponents) for smallest spatial resolution, and calibration was used to determine unknown parameters for the rest of space steps. During calibration, model exponents were fixed while the coefficients were adjusted until the best fit was achieved between calculated dispersion coefficients and those given by the predictive model. Several performance analysis metrics were used to assess the quality of calibrations. Performance analysis of the calibration process is discussed in the next chapter.

5.2 Methods

5.2.1 Testing normality of computed dispersion coefficients

The models were developed for purposes of predicting longitudinal dispersion coefficients over the range of flow rates based on computed dispersion coefficients. This implies that predictions by empirical models were conditional on computed dispersion coefficients. Therefore, it was prudent to check if computed dispersion coefficients were normally distributed. The results of no observance of normality assumption often lead to the use of substandard estimators, baseless inferential statements and incorrect conclusions (Montgomery and Runger, 2003).

The test involved normal scores (quantile-quantile; Q-Q) plot followed by a normality hypothesis test using the Shapiro-Wilk method. The method has been found to be a useful test that detects most departures from normality when the sample size is not more than 5000 (Ryan and Joiner, 1976; Analyse-it Software Ltd., 2009; Shapiro and Wilk, 2016). The test relates Shapiro-Wilk statistic (W), sample size (N) and confidence level (α).

The procedure for computation of the Shapiro-Wilk statistic is as follows:

- i. Rank order the sample data
- ii. Calculate, b , the weighted sum of the differences between the most extreme values
- iii. Divide the weighted sum by a multiple of standard deviation, and square the result to get the Shapiro-Wilk statistic (W), defined by:

$$W = \left(\frac{b}{s_y \sqrt{N-1}} \right), \quad (5.1)$$

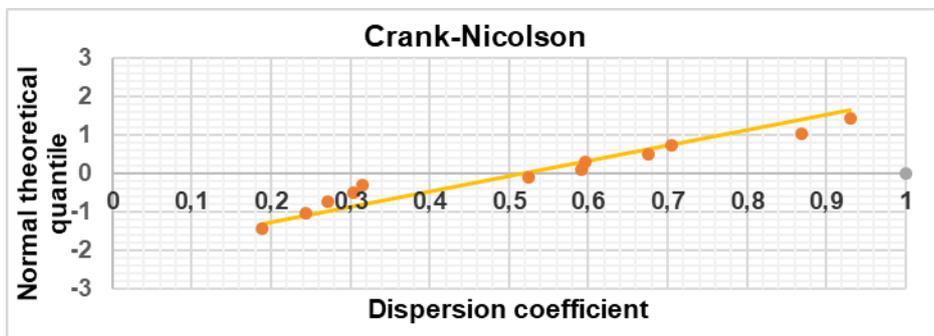
$$b = \sum_{j=1}^m c_{n-1+1} (\omega_{n-1+1} - y_j) = \sum_{j=1}^m b_j, \quad (5.2)$$

where s_y = the standard deviation of the sample, N = sample size, ω_j = the smallest order value of observed variable in the j^{th} pair of extremes, c_j = a coefficient which depends on sample size, and m = the greatest integer less than $N/2$. The hypothesis of normality is rejected at a specified level of significance when the Shapiro-Wilk statistic W is less than the applicable tabulated value (Ryan and Joiner, 1976; Hanusz et al., 2016). The highest critical value (for a confidence level of $\alpha=0.1$) of the Shapiro-Wilk test statistic for sample size $N=12$ was found to be equal to 0.7154 (Hanusz et al., 2016). Alternatively, a hypothesis test can be achieved by using P-value and specified level of significance (α -value) (Montgomery and Runger, 2003). If the P-value is greater than the level of significance then the null hypothesis can be accepted (Montgomery and Runger, 2003; Gogtay et al., 2016).

5.2.2 Results

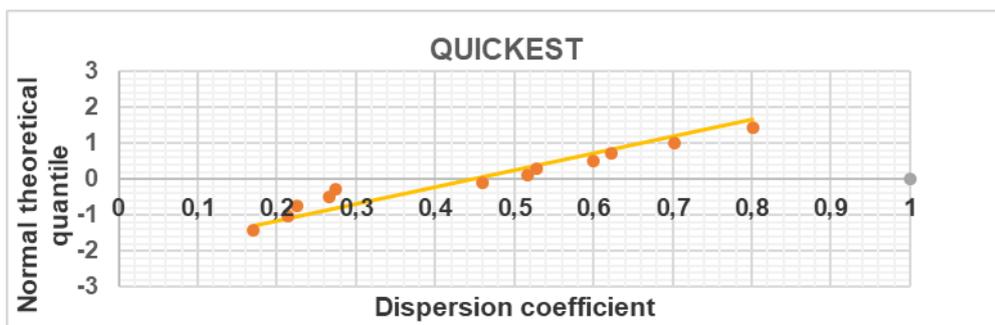
Figure 5-1 shows Q-Q plots and Shapiro-Wilk test reports for normality tests. The tests were carried on average values of estimated dispersion coefficients over the range of space steps given in Tables 4-5 to 4-7. The critical values of the Shapiro-Wilk tests (see Appendix B) are 0.4940 for $\alpha=0.01$, 0.6431 for $\alpha=0.05$ and 0.7154 for $\alpha=0.1$ (Hanusz et al., 2016). The calculated Shapiro-Wilk statistics (W) were greater than the critical Shapiro-Wilk statistics in all cases. Therefore, it was asserted that at the three significance levels the estimated dispersion coefficients were drawn from

a normal distribution. Additionally, P-values were greater than confidence levels in all cases. Therefore, it can further be stated that there is no significant departure from normality.



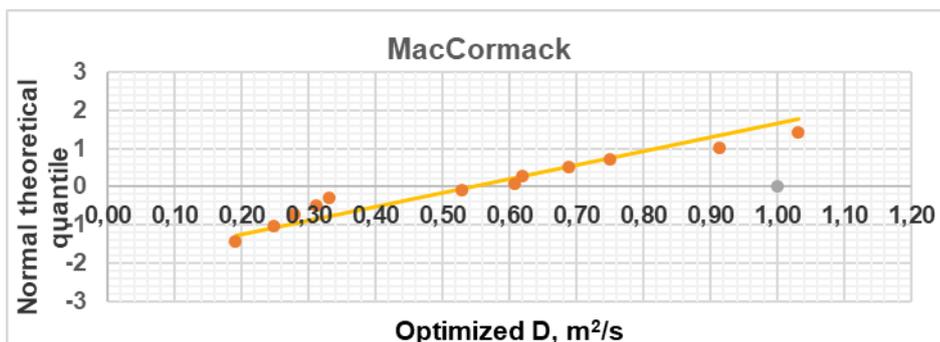
Shapiro-Wilk test

W statistic	0.93
p-value	0.3412



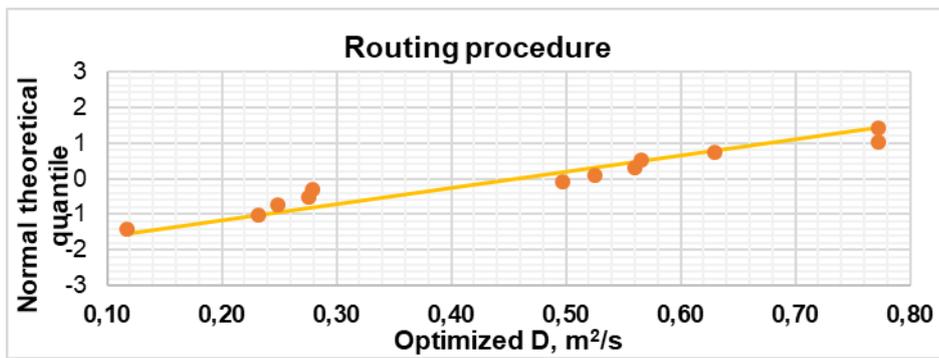
Shapiro-Wilk test:

W statistic	0.92
p-value	0.3251



Shapiro-Wilk test

W statistic	0.94
p-value	0.4436



Shapiro-Wilk test

W statistic	0.92
p-value	0.3184

Figure 5-1: Q-Q plots and normality test results of computed dispersion coefficients

5.2.3 Development of new models

Empirical models were developed based the Singh and Beck routing procedure and numerical methods, viz. Crank-Nicolson, MacCormack and QUICKEST. Also, for numerical methods, models were developed base on model resolution. Models were developed by correlating optimised dispersion coefficients with bulk hydraulic and channel characteristics.

In this context bulk characteristics that influence longitudinal dispersion coefficient can be classified as fluid properties, hydraulic characteristics and geometric forms (Rutherford, 1994; Falconer, Kashefipour and Falconer, 2002; Seo and Baek, 2004). Channel characteristics are channel width and channel depth, and hydraulic characteristics include shear velocity, flow velocity and channel flow rate. Geometric characteristics include bedforms such as dead zones, and channel sinuosity (Rutherford, 1994; Seo and Cheong, 1998; Kashefipour and Falconer, 2002). Fluid properties include fluid density and viscosity. Flow complexities such as sinuosity and shape factor represent characteristics that are not easily measured in natural channels, and the impact of these characteristics can be included in the friction term (Seo and Cheong, 1998).

Therefore, the longitudinal dispersion can functionally be expressed as

$$D = f_1(v, H, w, U_*, Q, \rho, \nu), \quad (5.3)$$

where v is the cross-sectional average flow velocity, H is the depth of flow, W is the top width of the flow, U_* is shear velocity, Q is flow rate, ρ is water density and ν is kinematic viscosity.

Shear velocity is given by $U_* = (gRs)^{0.5}$, where g is the gravitational acceleration and s is the channel slope and R is the hydraulic radius. According to the common practice (e.g. Seo and Cheong, 1998; Kashefipour and Falconer, 2002), using dimensional analysis, the functional relationship between dimensionless terms was established as

$$\frac{DH}{Q} = f_2 \left(\frac{vH}{\nu}, \frac{v}{U_*}, \frac{W}{H} \right). \quad (5.4)$$

For turbulent flows, the effect of Reynolds number may be neglected (Kashefipour and Falconer, 2002; Seo and Baek, 2004). Additionally, the channel width was not considered in the formulation of the empirical models as there was no detailed data on the variation of width with flow rate; only the average channel width was available. Therefore, the non-dimensional form relating longitudinal dispersion coefficient to bulk flow properties reduced to:

$$\frac{DH}{Q} = f_3 \left(\frac{v}{U_*} \right). \quad (5.5)$$

Two general expressions of predictive models for estimating dispersion coefficient were identified based on flow rate, cross-sectional average velocity, shear velocity and depth of flow as follows:

$$D = \sigma \frac{Q}{H} \left(\frac{v}{U_*} \right)^\lambda, \quad (5.6)$$

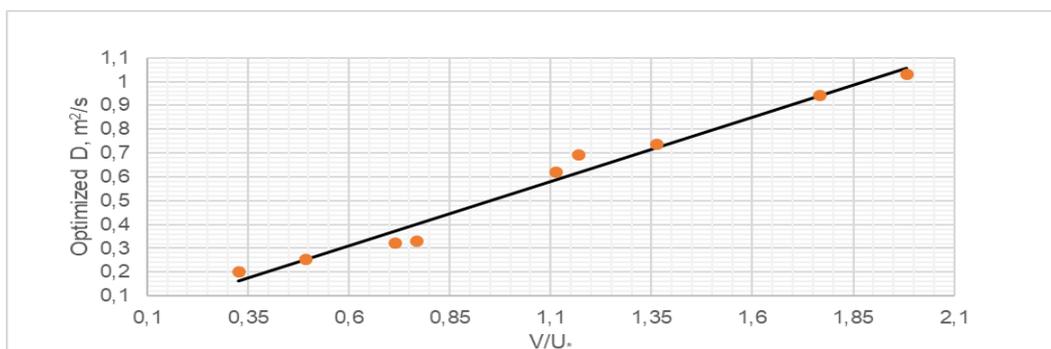
$$D = \alpha \left(\frac{Q}{H} \right)^\beta \left(\frac{v}{U_*} \right)^\gamma, \quad (5.7)$$

where σ and α are model coefficients, β , γ and λ are model exponents. These unknown quantities are collectively referred to as an unknown model or regression parameters. The bulk flow

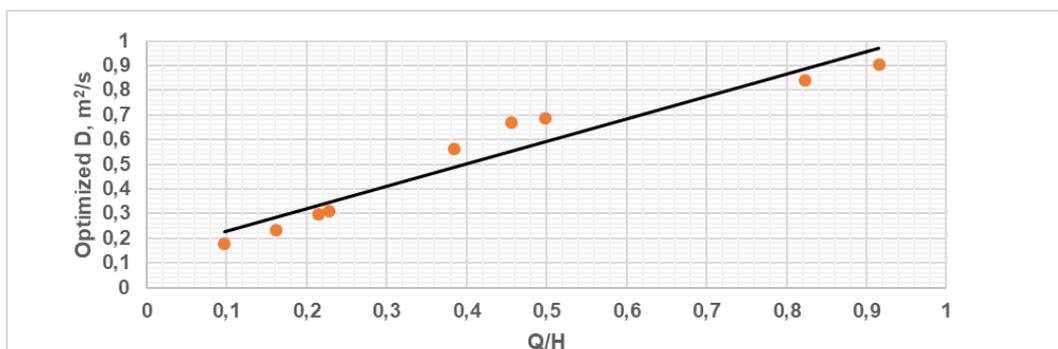
and channel characteristics appearing in the equations were calculated using computed flow rates and velocities. The wetted cross-sectional area was determined from cross-sectional average velocity and the stream flow rate, which was later used to estimate the depth of flow. The streamflow depths were estimated with the assumption that the channel was rectangular.

After preliminary assessment Equation (5.7) was found to be more accurate than Equation (5.6), although Equation 5.6 is dimensionally correct. Consequently, Equation (5.7) was considered for further analysis. Equation (5.7) has the disadvantage that the value of the exponent β is non-integer. Therefore, the equation is not dimensionally correct, and misgivings could be raised about the physical meaning of the model. However, this is similar to those proposed by Ani et al. (2009) and Whitehead et al. (1986), where the dispersion coefficient was correlated to channel flow rate.

To test the correlation between optimised dispersion coefficients and independent variables of Eq. (5.7), plots of computed dispersion coefficients against model variables were constructed shown in Figures 5-2 to 5-5.



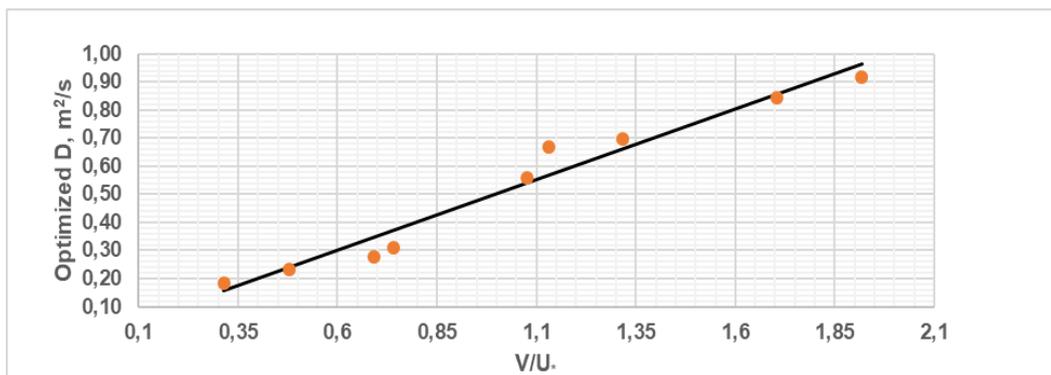
R^2	0.978
R^2 adjusted	0.974
RMSE	0.048993603



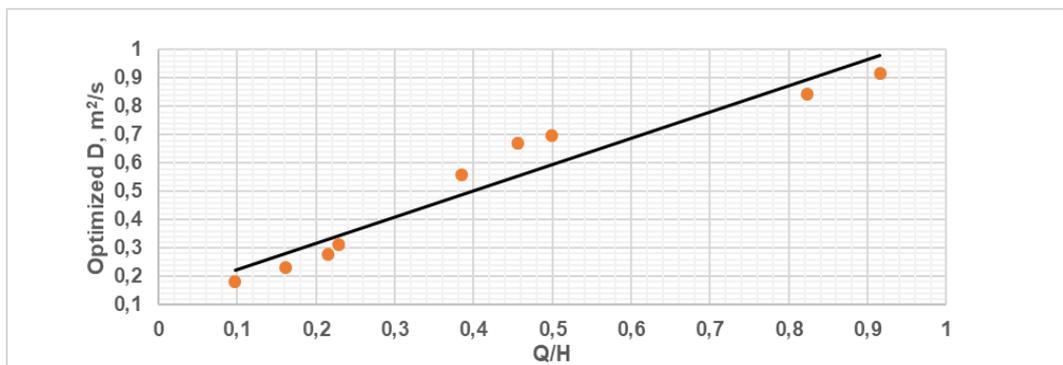
R^2	0.929
R^2 adjusted	0.919

RMSE | 0.077739153

Figure 5-2: Computed dispersion coefficients versus independent model variables (Crank-Nicolson)

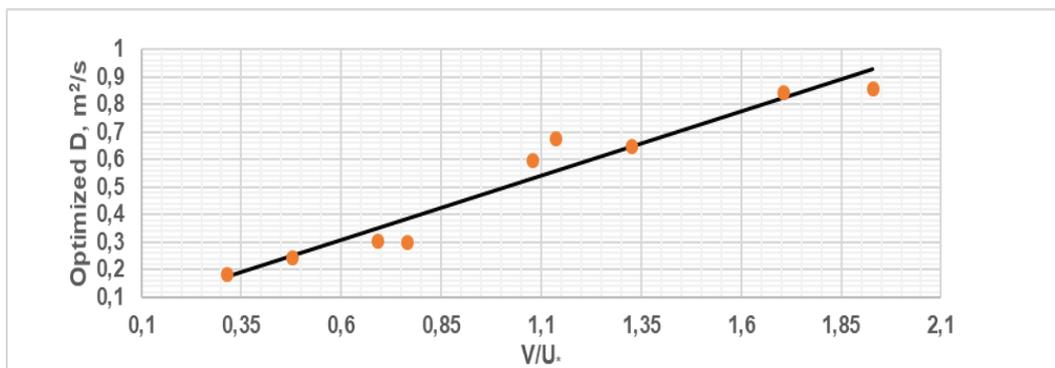


R ²	0.962
R ² adjusted	0.957
RMSE	0.057725058

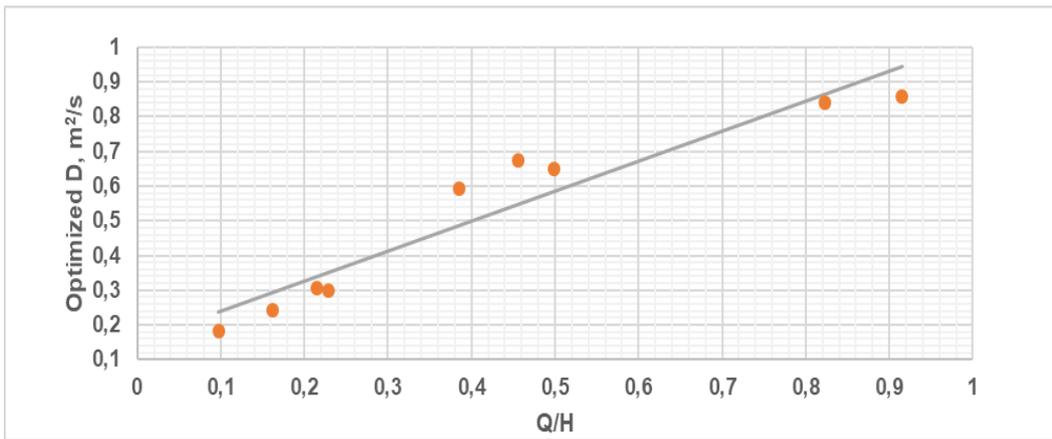


R ²	0.930
R ² adjusted	0.920
RMSE	0.078775741

Figure 5-3: Computed dispersion coefficient versus model independent variables (MacCormack)

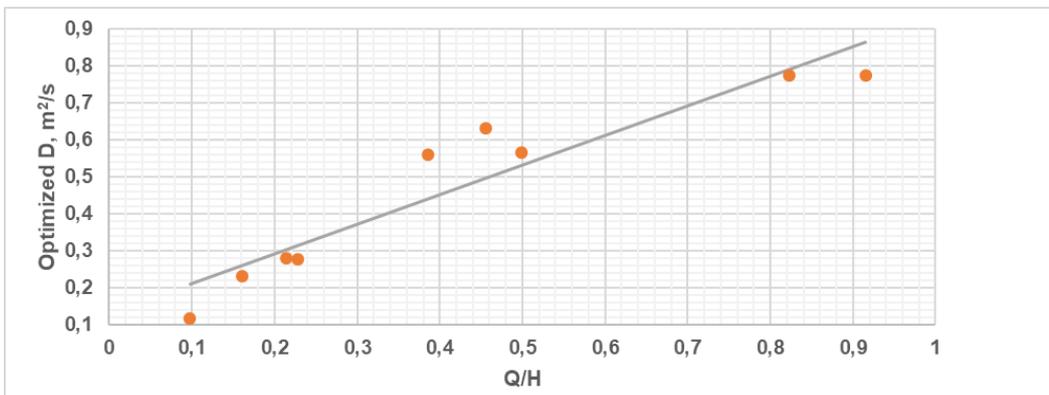


R ²	0.940
R ² adjusted	0.932
RMSE	0.068341229

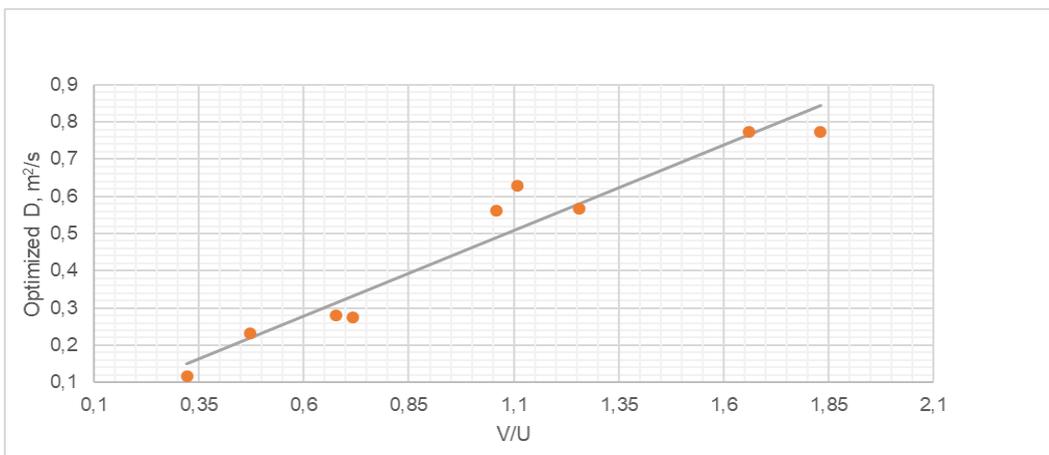


R ²	0.909
R ² adjusted	0.896
RMSE	0.084351317

Figure 5-4: Computed dispersion coefficients versus independent model variables (QUICKEST)



R ²	0.889
R ² adjusted	0.873
RMSE	0.087516836



R ²	0.938
R ² adjusted	0.929
RMSE	0.065336408

Figure 5-5: Computed dispersion coefficient versus the control variables (Singh and Beck routing procedure)

The plot of computed D versus v/U_* and computed D versus Q/H shown in Figures 5-4 to 5-6, respectively. The longitudinal dispersion coefficient increases as the friction term increases and the flow rate-to-flow depth increases. These figures demonstrate that the longitudinal dispersion coefficient depends on these two independent variables.

5.2.4 Evaluating model parameters and model calibration

The proposed model structure is a balance equation that relates the dispersion coefficient to components of the model; hydraulic and channel characteristics. The solution of the modified equation is generally determined by a linear least-squares procedure in which a residual sum of squares is minimised (Seo and Cheong, 1998; Montgomery and Runger, 2003; Chapra and Canale, 2008). For situations where errors are not normally distributed, estimates of model parameters reflect outliers, also known as leverage points, which influence the quality of the fit (Seo and Cheong, 1998). In this study, leverage points have not been observed in optimised dispersion coefficients which are the basis for the development of the model (see section 5.4.1).

The identified equation was a multiple nonlinear power equation. To enable the use of least-squares linear regression analysis to determine unknown parameters the equation required linearization. Typical non-linear multiple equations relating a dependent variable and q unknown control variables can be expressed in a general form as (Seo and Cheong, 1998; Chapra and Canale, 2008)

$$D = \alpha F_1^\beta F_2^\gamma \dots F_q^\lambda \delta, \quad (5.8)$$

where F are the control variables representing flow and channel characteristics, D is the dependent variable representing longitudinal dispersion coefficient, δ are independent random errors and $\alpha, \beta, \gamma, \dots, \lambda$ are unknown regression parameters. Equation (5.8) can be linearised by taking logarithms resulting in a multiple linear form as follows (Seo and Cheong, 1998; Chapra and Canale, 2008):

$$\ln D = \ln \alpha + \beta \ln F_1 + \gamma \ln F_2 + \dots + \lambda F_m + \ln \delta. \quad (5.9)$$

In this case in which there are three unknown regression parameters Eq. (5.9) can be as

$$\ln D = \ln \alpha + \beta \ln F_1 + \gamma \ln F_2 + \ln \delta. \quad (5.10)$$

Equation (5.10) can be rewritten as

$$\chi_j = \zeta + \beta Z_{1j} + \gamma Z_{2j} + e, \quad (5.11)$$

where, χ is the transformed dependent variable, ζ is the transformed model coefficient (α), Z_1 and Z_2 are transformed independent variables and e is the transformed error term. The optimal values of regression coefficients were obtained by setting up the residual sum of squares as,

$$RSS = \sum_{j=1}^n (\chi_j - \zeta - \beta z_{1j} - \gamma Z_{2j})^2. \quad (5.12)$$

Differentiating Eq. (5.12) with respect to unknown regression parameters (Montgomery and Runger, 2003; Chapra and Canale, 2008),

$$\frac{\partial RSS}{\partial \omega} = -2 \sum (\chi_j - \zeta - \beta Z_{1j} - \gamma Z_{2j}), \quad (5.13)$$

$$\frac{\partial RSS}{\partial \beta} = -2 \sum z_{1j} (\chi_j - \zeta - \beta Z_{1j} - \gamma Z_{2j}), \quad (5.14)$$

$$\frac{\partial RSS}{\partial \gamma} = -2 \sum z_{2j} (\chi_j - \zeta - \beta Z_{1j} - \gamma Z_{2j}), \quad (5.15)$$

The regression parameters that give the minimum residual sum of squares are determined by equating the partial derivatives to zero and writing the result in matrix notation as,

$$\begin{bmatrix} N & \sum Z_{1j} & \sum Z_{2j} \\ \sum Z_{1j} & \sum Z_{1j}^2 & \sum Z_{1j} Z_{2j} \\ \sum Z_{2j} & \sum Z_{1j} Z_{2j} & \sum Z_{2j}^2 \end{bmatrix} \begin{Bmatrix} \zeta \\ \beta \\ \gamma \end{Bmatrix} = \begin{Bmatrix} \sum \chi_j \\ \sum Z_{1j} \chi_j \\ \sum Z_{2j} \chi_j \end{Bmatrix}. \quad (5.16)$$

Equation (5.16) can be expressed succinctly in matrix notation as (Chapra and Canale, 2008):

$$[[Z]^T [Z]]\{C\} = \{[Z]^T \{x\}\}, \quad (5.17)$$

where $[Z]$ is the matrix of computed values of independent variables, $\{C\}$ is the column vector of unknown model parameters, $\{x\}$ is the column vector of observed values of the dependent variable and N is the number of observed values of the dependent variable.

Several methods are available to solve Equation (5.16) for unknown regression parameters, such as the Gauss elimination method, the matrix inversion method, etc. (Chapra and Canale, 2008). Using matrix inversion, the model parameters that give the minimum sum of the squares are obtained using the following expression (Chapra and Canale, 2008):

$$\{C\} = [[Z]^T [Z]]^{-1} \{[Z]^T \{x\}\}. \quad (5.18)$$

After evaluating regression model parameters, the regression model was transformed to its original state to be used for prediction of dispersion coefficients.

In numerical methods computation errors are dominated by truncation errors, which can be reduced by decreasing the step size of grid discretisation (Chapra and Canale, 2008). Therefore, the procedure of regression analysis was only used to determine all regression parameters for the smallest space step using model development data (i.e. space step of 2 m for the Crank-Nicolson and MacCormack based models and space step of 8 m for the QUICKEST based model). The model parameters for the rest of space steps were obtained by calibration, where values of the exponents were fixed, and the model coefficient α was adjusted until the best fit was obtained between the models' predictions and calculated dispersion coefficients. Having constant values of exponents and a varying coefficient resulted in more accurate predictions. This approach has the advantage that issues with system condition at several space steps are avoided and has the benefit of convenience when applying the models as only the appropriate value of the coefficient can be selected for an application. The adequacy of a solution of a system of the linear equations was dependent on the condition of the system.

A system of equations can be ill-conditioned if two or more of the equations are nearly identical or close to being singular. If a system is ill-conditioned, it is difficult to identify optimal parameter values of the system (Chapra and Canale, 2008). Consequently, depending on the values of variables obtained for a model resolution and solution method, the system of linear equations could result in an ill-conditioned system. Several ways are available to check if the system is ill-conditioned, which include the following methods (Chapra and Canale, 2008):

- a) Scaling the matrix of independent variables, such that the largest value of elements in each row is unity, inverting the scaled matrix and checking the order of magnitude of elements of the inverted scaled matrix. If there are elements that are much larger than unity, the system is probably ill-conditioned.
- b) Inverting the matrix of independent variables and multiplying the inverse by the original matrix of independent variables and checking whether the product is close to the unit matrix.
- c) Inverting the inverted matrix of independent variables and examine whether the outcome is satisfactorily close to the original matrix of independent variables.

While the main objective was to investigate the impact of numerical properties and numerical methods, system condition could also affect predicted results of empirical models. Therefore, calibration of the models by adjusting the coefficient α avoided the influence of system condition on the structure of models and results of the investigation.

5.2.5 Results and discussion: regression parameters

Four sets of model parameters were obtained based on four solution methods of AD-Model, viz. Crank-Nicolson, MacCormack, QUICKEST and the Singh and Beck routing procedure. Tables 5-1 to 5-3 presents space steps, and model (or regression) parameters for the Crank-Nicolson-based, the MacCormack-based and the QUICKEST-based empirical models, respectively. Table 5-4 lists model parameters for the empirical model based on the Singh and Beck routing procedure.

For models based on numerical methods, it can be observed that the model parameters vary with space step and numerical method used to compute dispersion coefficients for the development of

empirical models. Therefore, model parameters also vary with Peclet number, as the Peclet number increases with space step. However, the model based on the Singh and Beck routing procedure has a fixed set of parameters as the routing procedure does not involve spatial discretisation.

It can also be observed that the model coefficient α increases with space step (or Peclet number) for the Crank-Nicolson and the MacCormack based models, and the coefficient decreases with space step (or Peclet number) for model based on the QUICKEST method. This shows that empirical models based on the Crank-Nicolson and MacCormack predict higher values with increase in Peclet number while the empirical model based on the QUICKEST method predict lower values with increase in space step (or Peclet number). Variation of the models' coefficient with Peclet number is shown in Figure 5-6.

Table 5-1: Space step and model parameters for the Crank-Nicolson based model

Δx (m)	Coefficient, α	Exponent, β	Exponent, γ
2.00	0.607	0.233	0.779
4.00	0.611	0.233	0.779
4.60	0.611	0.233	0.779
5.75	0.612	0.233	0.779
7.36	0.613	0.233	0.779
8.00	0.614	0.233	0.779
8.36	0.615	0.233	0.779
9.20	0.616	0.233	0.779
11.50	0.598	0.233	0.779
12.27	0.624	0.233	0.779
13.14	0.627	0.233	0.779
14.15	0.631	0.233	0.779
16.73	0.642	0.233	0.779
18.40	0.649	0.233	0.779

Table 5-2: Space step and model parameters for the MacCormack-based model

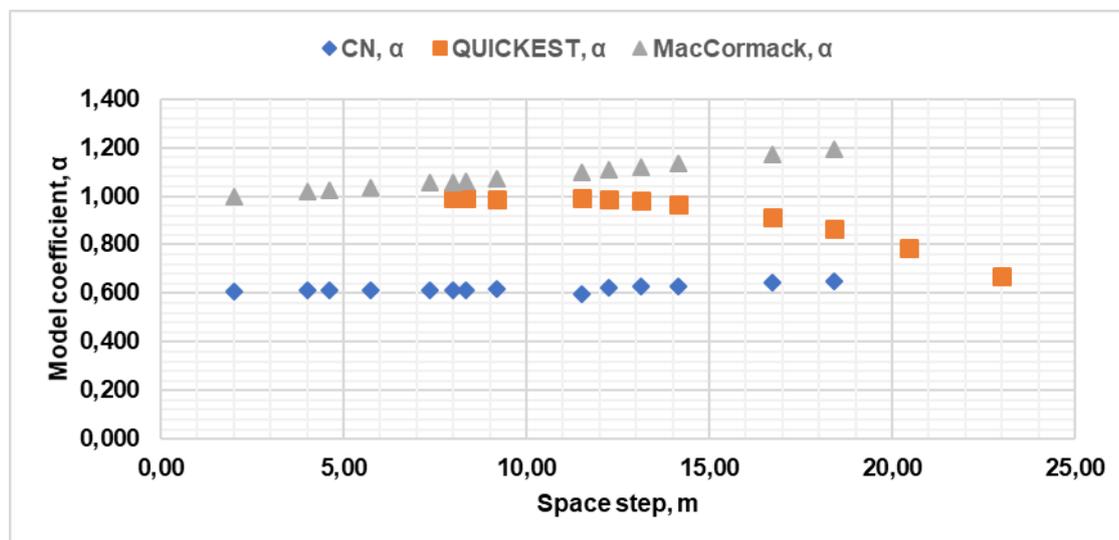
Δx (m)	Coefficient, α	Exponent, β	Exponent, γ
2.00	0.997	0.729	0.082
4.00	1.018	0.729	0.082
4.60	1.022	0.729	0.082
5.75	1.032	0.729	0.082
7.36	1.056	0.729	0.082
8.00	1.055	0.729	0.082
8.36	1.060	0.729	0.082
9.20	1.069	0.729	0.082
11.50	1.098	0.729	0.082
12.27	1.109	0.729	0.082
13.14	1.120	0.729	0.082
14.15	1.134	0.729	0.082
16.73	1.170	0.729	0.082
18.40	1.193	0.729	0.082

Table 5-3 Space step and model parameters for the QUICKEST-based model

Δx (m)	Coefficient, α	Exponent, β	Exponent, γ
8.00	0.990	0.727	0.040
8.36	0.991	0.727	0.040
9.20	0.989	0.727	0.040
11.50	0.990	0.727	0.040
12.27	0.987	0.727	0.040
13.14	0.980	0.727	0.040
14.15	0.967	0.727	0.040
15.33	0.947	0.727	0.040
16.73	0.915	0.727	0.040
18.40	0.865	0.727	0.040
20.44	0.789	0.727	0.040
23.00	0.670	0.727	0.040

Table 5-4: Regression parameters for the routing-based model

Coefficient, α	Exponent, β	Exponent, γ
0.509	0.117	0.950

Figure 5-6: Coefficients, α , of empirical models based on numerical methods versus space step

5.2.6 Results and discussion: calibration of models

Calibration of numerical-method based models involved adjusting the models' coefficient α to achieve an optimal fit between models' predictions and optimised dispersion coefficients. The process was repeated for space steps 2.00 m to 18.40 m the Crank-Nicolson and MacCormack based models and space steps 8.00 m to 23.00 m for the QUICKEST based model. Tables 5-5 to 5-7 list predicted dispersion coefficients of model based on Crank-Nicolson, MacCormack and QUICKEST numerical methods, respectively. Table 5-8 lists predicted dispersion coefficients

obtained by the empirical model based on the Singh and Beck routing procedure.

It can be observed that predicted dispersion coefficients vary with models and space step (or Peclet number). Like computed dispersion coefficients, the Crank-Nicolson and MacCormack based empirical models predict higher values of dispersion coefficients with an increase in space step (or Peclet number), while the QUICKEST based model predicts dispersion coefficient values that decrease with increase in space step (or Peclet number). However, all models predict dispersion coefficients that increase with flow rate as observed from previous studies (e.g. Rutherford, 1994).

Table 5-5: Flow rates, space steps and predicted dispersion coefficients (CN-based model)

Experiment	10	2	7	8	5	3	11	6	12
Flow rate (m³/s)	0.017	0.036	0.041	0.044	0.097	0.150	0.147	0.385	0.436
Space step, (m)									
2.00	0.144	0.223	0.317	0.340	0.514	0.638	0.555	0.876	0.984
4.00	0.145	0.225	0.319	0.342	0.518	0.643	0.559	0.883	0.991
4.60	0.145	0.225	0.319	0.342	0.518	0.644	0.560	0.884	0.992
5.75	0.145	0.225	0.320	0.343	0.519	0.645	0.561	0.886	0.995
7.36	0.146	0.226	0.322	0.345	0.522	0.648	0.563	0.890	0.999
8.00	0.146	0.227	0.323	0.346	0.523	0.650	0.565	0.892	1.002
8.36	0.146	0.227	0.323	0.346	0.524	0.651	0.566	0.894	1.003
9.20	0.147	0.228	0.324	0.348	0.526	0.653	0.568	0.897	1.007
11.50	0.143	0.222	0.316	0.339	0.512	0.637	0.553	0.875	0.938
12.27	0.150	0.232	0.331	0.355	0.535	0.665	0.578	0.915	1.027
13.14	0.151	0.234	0.333	0.357	0.539	0.670	0.582	0.921	1.033
14.15	0.152	0.236	0.336	0.360	0.544	0.676	0.587	0.929	1.042
16.73	0.155	0.241	0.344	0.369	0.556	0.691	0.601	0.951	1.067
18.40	0.158	0.245	0.349	0.375	0.565	0.703	0.611	0.967	1.085

Table 5-6: Flow rates, space steps and predicted dispersion coefficients (MacCormack-based model)

Experiment	10	2	7	8	5	3	11	6	12
Flow rate (m³/s)	0.017	0.036	0.041	0.044	0.097	0.150	0.147	0.385	0.436
Space step, (m)									
2.00	0.166	0.248	0.315	0.331	0.500	0.614	0.568	0.904	0.986
4.00	0.169	0.253	0.321	0.338	0.511	0.627	0.580	0.923	1.007
4.60	0.170	0.255	0.323	0.340	0.513	0.630	0.583	0.927	1.012
5.75	0.172	0.257	0.326	0.343	0.518	0.637	0.589	0.936	1.022
7.36	0.176	0.263	0.334	0.351	0.530	0.651	0.602	0.958	1.046
8.00	0.176	0.263	0.334	0.351	0.530	0.651	0.602	0.958	1.046
8.36	0.177	0.264	0.335	0.353	0.532	0.654	0.604	0.962	1.050
9.20	0.178	0.267	0.338	0.356	0.537	0.660	0.610	0.971	1.060
11.50	0.183	0.274	0.347	0.366	0.552	0.678	0.627	0.998	1.089
12.27	0.185	0.277	0.351	0.370	0.558	0.685	0.633	1.008	1.100
13.14	0.187	0.279	0.354	0.373	0.563	0.692	0.640	1.018	1.111
14.15	0.189	0.283	0.359	0.378	0.570	0.701	0.648	1.031	1.125
16.73	0.195	0.292	0.371	0.390	0.589	0.724	0.669	1.064	1.162
18.40	0.199	0.298	0.378	0.398	0.601	0.738	0.682	1.086	1.186

Table 5-7: Flow rates, space steps and predicted dispersion coefficients (QUICKEST-based model)

Experiment	10	2	7	8	5	3	11	6	12
Flow rate (m ³ /s)	0.017	0.036	0.041	0.044	0.097	0.150	0.147	0.385	0.436
Space step, (m)									
8.00	0.174	0.255	0.318	0.334	0.496	0.603	0.562	0.877	0.953
8.36	0.174	0.255	0.319	0.334	0.497	0.604	0.563	0.878	0.954
9.20	0.173	0.255	0.318	0.334	0.495	0.603	0.561	0.876	0.952
11.50	0.174	0.255	0.318	0.334	0.496	0.603	0.562	0.877	0.952
12.27	0.173	0.254	0.317	0.333	0.494	0.602	0.560	0.874	0.949
13.14	0.172	0.252	0.315	0.330	0.491	0.597	0.556	0.868	0.942
14.15	0.170	0.249	0.311	0.326	0.484	0.589	0.549	0.857	0.930
15.33	0.166	0.244	0.304	0.319	0.474	0.577	0.537	0.839	0.911
16.73	0.160	0.236	0.294	0.308	0.458	0.557	0.519	0.810	0.880
18.40	0.152	0.223	0.278	0.291	0.433	0.527	0.490	0.766	0.831
20.44	0.138	0.203	0.253	0.266	0.395	0.480	0.447	0.698	0.758
23.00	0.117	0.172	0.215	0.226	0.335	0.408	0.380	0.593	0.644

Table 5-8: Flow rates and predicted dispersion coefficients (routing method-based model)

Experiment	10	2	7	8	5	3	11	6	12
Flow rate Q, m ³ /s	0.017	0.036	0.041	0.044	0.097	0.15	0.147	0.385	0.436
Predicted D, m ² /s	0.132	0.201	0.293	0.312	0.480	0.582	0.512	0.805	0.894

Figures 5-7 and 5-8 demonstrate the variation of predicted dispersion coefficients with the Peclet number for experiments 10 and 12. Experiment 10 ($Q = 0.017 \text{ m}^3/\text{s}$), and experiment 12 ($Q = 0.436 \text{ m}^3/\text{s}$) were for the lowest and the highest flow rates, respectively, in the data set. It can be observed that the Crank-Nicolson and the MacCormack methods predict dispersion coefficients increasingly with an increase in Peclet number, while the model based on the QUICKEST method give predictions that decrease with increase in Peclet number. However, at low Peclet numbers predictions by models similar. This trend can be observed in Tables 5-5 to 5-7 for all experiments. This reflects the influence of numerical methods and numerical properties on predictions given by empirical models.

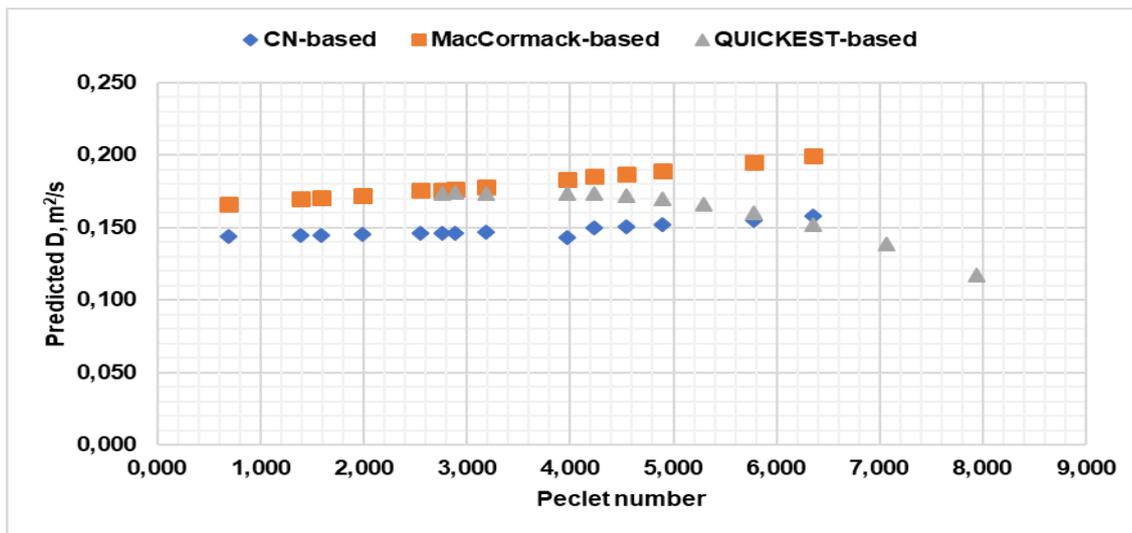


Figure 5-7: Plot of predicted dispersion coefficient versus the Peclet number (experiment 10)

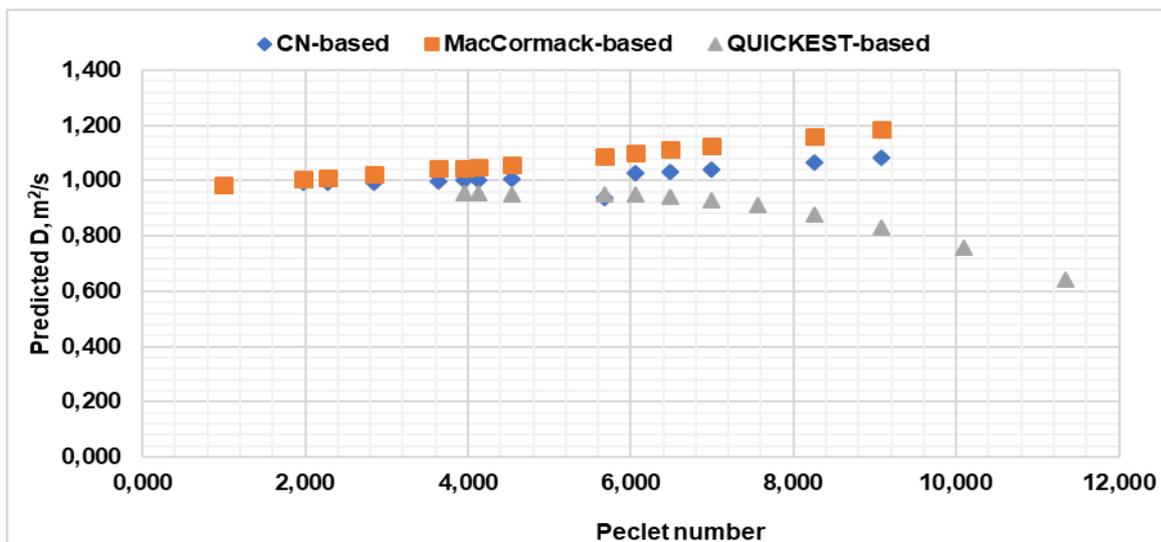


Figure 5-8: Plot of predicted dispersion coefficients versus the Peclet number (experiment 12)

Figure 5-9 and 5-10 show plots of predicted dispersion coefficient obtained by all empirical models versus flow rates. In Figure 5-9 predictions by empirical models based on numerical methods developed at small space steps are compared together with the model based on the routing procedure. In Figure 5.10 predictions by empirical models based on numerical methods developed at large space step are compared together with the model based on the routing procedure. In Figure 5-9 it can be observed that predicted dispersion coefficients by all methods are very similar over the whole range of flow rates. In contrast, Figure 5-10 shows that predicted dispersion coefficients differ

markedly with increase in flow rate, with the MacCormack based model giving highest predictions and the QUICKEST based model giving the lowest predictions with an increase in flow rate.

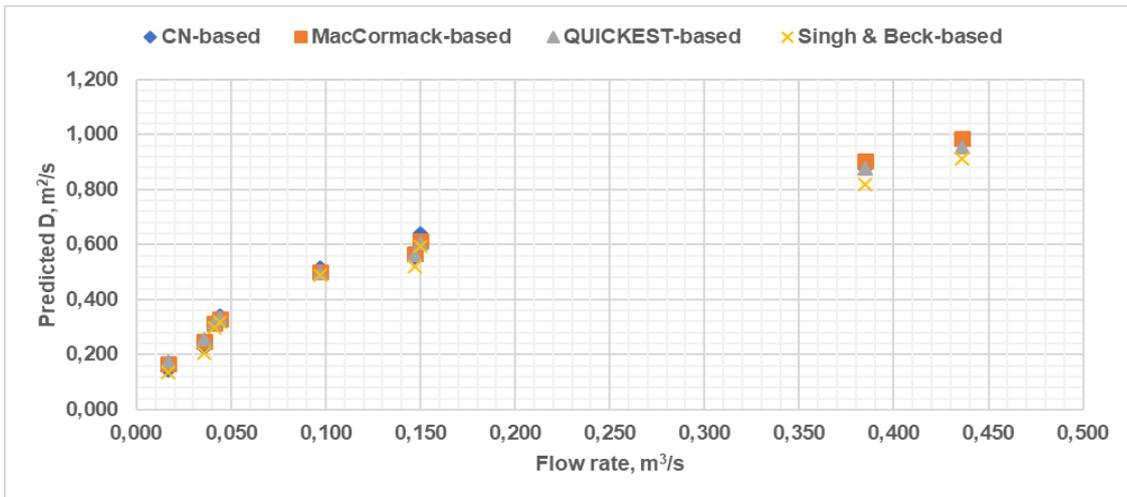


Figure 5-9: Plot of predicted dispersion coefficients against flow rate, $\Delta x = 2.00$ m for numerical method-based models.

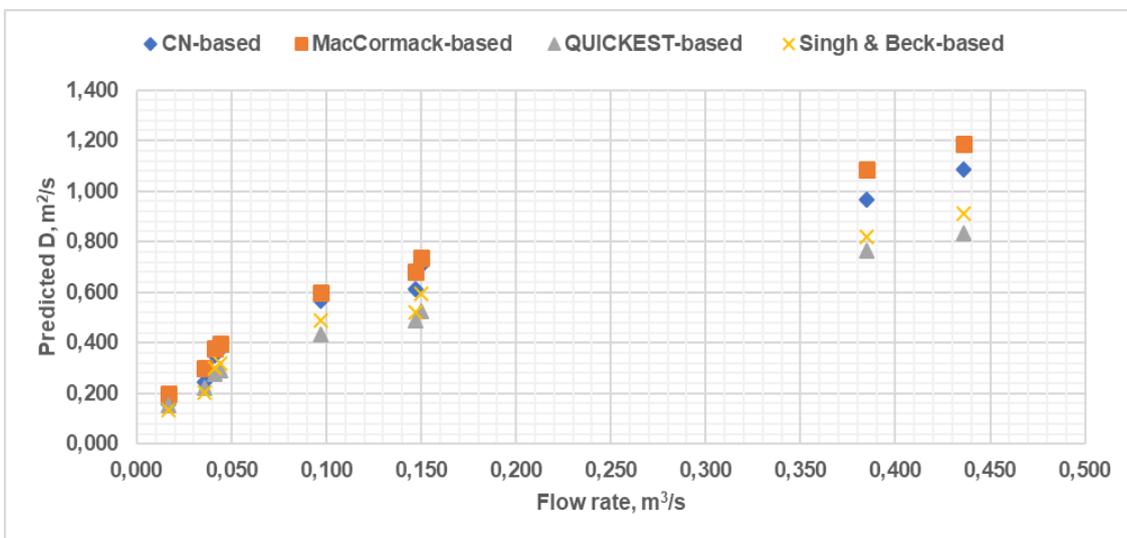


Figure 5-10: Plot of predicted dispersion coefficients against flow rate, $\Delta x = 18.40$ m for numerical method-based models.

5.3 Conclusion

One general functional form regression model with unknown regression parameters was developed. Three sets of values of regression parameters were determined for empirical models based on

numerical methods (i.e. Crank-Nicolson, MacCormack and QUICKEST) and Peclet number, and one set of regression parameters was determined for the empirical model based on the Singh and Beck routing procedure. Parameters for empirical models based on numerical models varied with the numerical method and Peclet number. Therefore, the structure of the model is influenced by the numerical method used to estimate parameters and numerical properties under which parameters were estimated.

Three sets of dispersion coefficients were obtained by empirical equations based on numerical methods and Peclet number, and one set of dispersion coefficients was obtained by the empirical model based on the routing procedure, from model building data. Values of dispersion coefficients obtained by empirical models based on numerical methods varied with model type and Peclet number. The Crank-Nicolson and the MacCormack based models gave higher values of dispersion coefficients with an increase in Peclet number, while the QUICKEST based model gave lower values with an increase in Peclet number. Predicted dispersion coefficients by models based on numerical methods reflect the variation of optimised dispersion coefficients with the numerical method and numerical properties. Therefore, predicted dispersion coefficients given by empirical models based on numerical methods are influenced by the numerical method used to estimate parameters and numerical properties under which parameters were observed for their development.

When predicted dispersion coefficients were plotted against flow rate, the trends were similar at low flow rates and low Peclet numbers, while marked differences in trends were observed at high Peclet numbers and flow rates. Therefore, at high Peclet numbers and flow rates predicted dispersion coefficients vary noticeably with the empirical model used.

6 PERFORMANCE ANALYSIS OF EMPIRICAL MODELS

6.1 Introduction

Calibration aims to adjust the parameters so that prediction optimally fits observations. The quality of the model calibration is assessed by performance measures (Chapra, 2008). For effective use of models for management purposes, it is important to demonstrate a level of assurance in their performance (Bennett et al., 2013). Environmental models commonly comprise numerous certain characteristics. If a performance analysis is based on one criterion, only certain features of performance are assessed (Bennett et al., 2013). Such a practice may lead to a preference of models that may not produce important properties of the system. Therefore, a system of measures addressing several characteristics may be required for a broad assessment of a model (Toprak and Cigizoglu, 2008). Quantitative methods allow comparison of models and give some measure of fairness in demonstrating advantages and disadvantages of a model. Quantitative assessment of models comprises computation of appropriate numerical metrics to identify model performance (Bennett et al., 2013).

Therefore, several workers recommend that a combination of methods be used to characterise the performance of models (e.g. Toprak and Cigizoglu, 2008; Chin, 2013). According to Toprak et al. (2004) models should be assessed using several descriptive statistics: such as maximum and minimum simulated values, standard deviation and variance, correlation coefficients and several error statistics; such as the mean square error, the standard error, normalised error etc. Chin (2013) recommends that graphical techniques, error statistics and residual error analysis be considered for model performance analysis. Chin (2013) further recommends that graphical methods are more appropriate for small data samples. Moriasi et al. (2015) recommend dimensionless metrics such as the Nash-Sutcliffe efficiency (NSE; E) (Nash and Sutcliffe, 1970), error indices such as the root mean square error (RMSE) and direct comparison (graphical) methods such as linear regression. Bennett et al. (2013) recommend a combination of graphical methods, error statistical methods and residual error analysis. Bennett et al. (2013) and Moriasi et al. (2015) recommend that the first step in assessing model performance is to use prescribed graphical performance metrics as they provide a

visual indication of the underlying behaviour of a model followed by computation of values of prescribed statistical performance measures. Therefore, quantitative performance analysis considered a combination of measures.

6.2 Methods

Based on recommendations of several workers (Bennett *et al.*, 2013; Chin, 2013; Moriasi *et al.*, 2015) a combination of performance measures was considered which included linear regression (graphical method), descriptive statistics, error indices and residual analysis. Firstly, graphical methods were used which included scatter plots, the slope of regression line and correlation coefficient. Descriptive statistics included minimum, maximum and variance of predicted values. Error indices included root mean square error (RMSE) and residual sum of squares (RSS). Lastly, residual error analysis methods were used which included coefficient of determination, Nash-Sutcliffe efficiency (NSE; E) and model adequacy assessment.

6.2.1 Graphical method

The graphical method involved scatter plots of predicted dispersion coefficient against optimised dispersion coefficients and linear regression. In this method, the slope of the regression line and the correlation coefficient were considered. The slope of the regression line and the correlation coefficient should be close to unity. The departure of the regression line from the unity-gradient line provides an observable appreciation of the underlying characteristics such as bias and systematic variance (Moriasi *et al.*, 2015). According to Chin (2013), when using linear regression for assessing model performance, the slope, the intercept and the correlation coefficient from the regression line may be calculated together before making any inference. However, the intercept was not considered in this study as values of the variables were not expected to be zero.

Correlation coefficient

The correlation coefficient is also called the Pearson's r is only used for bivariate normally distributed random variables that are correlated. The coefficient measures sample variance that is explained by the linear regression line (Montgomery and Runger, 2003; Ang and Tang, 2004) The correlation coefficient between two random variables, ω_j and ϖ_j , denoted as r is expressed as (Chin, 2013),

$$r = \frac{\sum_{j=1}^N (\varpi_j - \bar{\varpi})(\omega_j - \bar{\omega})}{\left[\sum_{j=1}^N (\varpi_j - \bar{\varpi})^2 \right]^{\frac{1}{2}} \left[\sum_{j=1}^N (\omega_j - \bar{\omega})^2 \right]^{\frac{1}{2}}}, \quad (6.1)$$

where ω_j are observed values, $\bar{\omega}$ is the mean of observed values, ϖ_j are the predicted values and $\bar{\varpi}$ is the mean of the predicted values. Values of the correlation coefficient can be anywhere in the range [-1.0 to 1.0]. Higher values of the coefficient are seen to be desirable for meaningful correlation (Chin, 2013).

6.2.2 Descriptive statistics

Toprak and Cigizoglu (2008) claim that in addition to error statistical measures of descriptive performance statistics should be used. There are several descriptive statistics, however minimum variable value, maximum variable value and variance of errors between observed and predicted values are considered in this study. The maximum and minimum predicted values show whether the model is over predicting or under predicting.

Variance of errors

The variance of the errors, $\hat{\sigma}^2$, is a measure of the dispersion of errors and is expressed as

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (\omega_i - \varpi_i)^2}{N-2} = \frac{RSS}{N-2}, \quad (6.2)$$

where RSS is the residual sum of squares.

6.2.3 Error statistics

Typical models used in environmental and water resources can be expressed in a general form as (Chin, 2013)

$$\bar{\mathbf{Y}}_i = f(\mathbf{Y}_i, \theta), \quad (6.3)$$

where $\hat{\mathbf{Y}}$ is the response vector, such that $\hat{\mathbf{Y}} = (\varpi_j; j = 1, 2, \dots, k)$ and $\mathbf{Y} = \{\omega_j; j = 1, 2, \dots, k\}$ is the input

vector, and $\theta = \{\theta_j; j = 1, 2, \dots, m\}$ are the model parameters. The minimisation of the error between the response vector and the input vector requires calibration which involves adjusting parameters values to obtain an optimal fit between observed and predicted data. According to the typical practice, calibration uses observed values to estimate the model response values, and the vector of residuals \mathbf{R} is evaluated as the mismatch between the model predictions and the observed values, thus,

$$\mathbf{R} = f(\bar{\mathbf{Y}}, \theta) - \bar{\mathbf{Y}}. \quad (6.4)$$

The estimated error vector and associated error statistics are, therefore, expressed in terms of the parameter vector (Chin, 2013). Several error analysis statistics are available in the literature. In this study, the *root-mean-square error* (RMSE) and *residual sum of squares* (RSS) were considered here.

Residual sum of squares

The residual sum of squares (RSS), also called the sum of squared errors (SSE), is expressed as

$$RSS = \sum_{j=1}^N (\omega_j - \varpi_j)^2, \quad (6.5)$$

where ω_j are observation random variables and ϖ are predicted random variables. The residual sum of squares measures the variance between observational data and model predictions.

Root mean square error:

The root mean square error (RMSE) is the square root of mean square error (also known as standard error) of the estimate in regression analysis (Moriassi *et al.*, 2015). The root mean square error evaluates the residual in terms of the units of the random variable and is expressed as

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (\omega_i - \varpi_i)}. \quad (6.6)$$

The RMSE has been observed to be the most important error statistic as it has same units as observational data, therefore, directly compares predicted data with observational data and is commonly used in model performance analysis (Singh *et al.*, 2005; Moriassi *et al.*, 2015).

6.2.4 Residual error analysis

The residual or error is the discrepancy between respective observed and simulated values. Preferably, the errors should be normally distributed and close to zero. Bennett et al. (2013) observed that error analysis was an important part for assessment of model performance. Residual error analysis helps to recognise if there is any bias. Quantitative measures of residual error analysis include coefficient of determination, model efficiency (Nash-Sutcliffe efficiency; E) and model adequacy assessment (Montgomery and Runger, 2003; Bennett et al., 2013; Chin, 2013).

Coefficient of determination

The coefficient of determination, R^2 , describes the part of the variance in the observed data that can be explained by the model or the amount of variance described by the regression line. Large values of R^2 do not certainly imply that the model will give reliable subsequent predictions (Montgomery, 2003). Its values range from 0.0 to 1.0, with high values showing good agreement between predictions and observations (Chin, 2013). The coefficient of determination is expressed as

$$R^2 = \left[\frac{\sum_{j=1}^N (\omega_j - \bar{\omega})(\varpi_j - \bar{\varpi})}{\sqrt{\sum_{j=1}^N (\omega_j - \bar{\omega})^2} \sqrt{\sum_{j=1}^N (\varpi_j - \bar{\varpi})^2}} \right]^2, \quad (6.7)$$

where, \bar{y} is the mean value of observations and $\bar{\hat{y}}$ is the mean value of predictions. The reliability of the coefficient of determination is diminished if model output is biased (McCuen et al., 2006).

Model efficiency

Model efficiency is a normalised statistic, and it measures the relative amount of error variance compared to the observational data variance. Model efficiency also called the Nash-Sutcliffe efficiency (NSE; E) (Sáez and Rittmann, 1992), or efficiency index (McCuen, Knight and Cutter, 2006) is a normalised statistic used to assess models outputs and is expressed as (Chin, 2013)

$$E = 1.0 - \frac{\sum_{j=1}^N (\omega_j - \varpi_j)^2}{\sum_{j=1}^N (\omega_j - \bar{\omega})^2}. \quad (6.8)$$

The statistic includes the ratio of two sums of squares, where the numerator evaluates the variation of data that is not defined by the model and the denominator evaluates the variation of data that can likely be explained by the model. Values of E range from $-\infty$ to 1.0, with higher values indicating more desirable agreement between the predicted values and observed values. If the value of model efficiency is more than zero, the model is considered a better estimator of system behaviour than the mean of the observed data (Chin, 2013). The E is considered a better measure than the coefficient of determination because the coefficient of determination is not sensitive to differences between model predictions and observational data (Harmel and Smith, 2007).

Model adequacy assessment

If a model is used to make inferences, it is important to check the assumptions underlying the analysis (Montgomery, 2003). The assumptions of a linear relationship between two correlated random variables and constancy of variance are inherent properties of two samples that are jointly normal. Therefore, the errors should also be normally distributed with constant variance. Additionally, to assume that the order of a model is appropriate, such as a linear model, is to assume that the event behaves linearly and that all the information on lack of fit is ascribed to the residuals (Montgomery and Runger, 2003). The normal probability plot of residuals is commonly used as a check for normality of small sample sizes (Montgomery and Runger, 2003). It is also helpful to plot residuals against the dependent variable or the control variable. The resulting graph will show whether the model is satisfactory or not. Alternatively, a plot of normalised residuals against predicted values is commonly used to check the adequacy of a model (Montgomery and Runger, 2003). Normalising residuals is expressed as

$$\hat{\partial}_j = \frac{\delta_j}{\sqrt{\hat{\sigma}^2}}, \quad j = 1, 2, \dots, N, \quad (6.9)$$

where $\hat{\partial}_j$ is normalised residual, $\delta_j = \omega_j - \bar{\omega}_j$ and $\hat{\sigma}^2$ is the variance of residuals.

The plot can be achieved by use of a statistical computer program such as Analyse-it for Microsoft Excel (2009). If the residuals are normally distributed, approximately 95 % of the standardised residuals should fall in the interval $-2, +2$ (Montgomery and Runger, 2003). There are three basic

scenarios, namely, a constant band of error variance with the response, an increase in error variance with the response and a systematic dependence of error variance on the response. A constant variance of residuals is an indication of satisfactory model adequacy. Therefore, depending on the pattern of the graph the variance may be satisfactory, may be dependent on the response or may indicate inequality of variance and indicate model inadequacy (Montgomery and Runger, 2003; Analyse-it Software Ltd., 2009).

6.3 Results and discussion

Tables 6-1 to 6-3 list various measures of model performance obtained during calibration for empirical models based on numerical methods. Table 6-4 lists performance measures for the routing method-based model, while Table 6-5 presents the average values of performance metrics for numerical methods-based models together with performance metrics for the routing method-based model. The performance presented in the tables are minimum and maximum optimised and predicted dispersion coefficients, the coefficient of determination (R^2), correlation coefficient (r), the slope of the regression line, variance, the residual sum of squares (RSS), root mean square error (RMSE) and Nash-Sutcliffe efficiency (E).

The graphical method included scatter plots of predicted dispersion coefficients versus optimised dispersion coefficients and linear regression. Plots of dispersion coefficients predicted by numerical methods-based models versus optimised dispersion coefficients are shown in Appendix C for space steps 2.00 m to 18.40 m for the Crank-Nicolson and the MacCormack based model, and space steps 8.00 m to 23.00 m for the QUICKEST based model and three of the plots are shown in Figures 6.1 to 6.3 for space step of 18.40 m. Figure 6.4 shows a plot of dispersion coefficient predicted by the routing method-based model versus optimised dispersion coefficient. It can be observed from data labels that predicted dispersion coefficients vary with empirical model. However, despite differences in model parameters good fits are achieved between predicted dispersion coefficients and optimised dispersion coefficients.

The values of correlation coefficients and slopes of regression lines are listed in Tables 6-1 to 6-4. It can be observed that both the values of slope and the correlation coefficient depend on the model

and model resolution. The values of the slope ranged from 0.947 to 1.039 for the Crank-Nicolson based model, 0.970 to 1.016 for the MacCormack based model and 0.940 to 1.016 for the QUICKEST based model. The values of the slope were generally close to the optimal value of 1.0, although showing bias which varies with model type and model resolution. The values of correlation coefficients range 0.955 to 0.989 for Crank-Nicolson based model, 0.977 to 0.991 for MacCormack based model and 0.899 to 0.976 for QUICKEST based model. This shows that the correlation between predicted and computed values of dispersion coefficients vary with model type and model resolution.

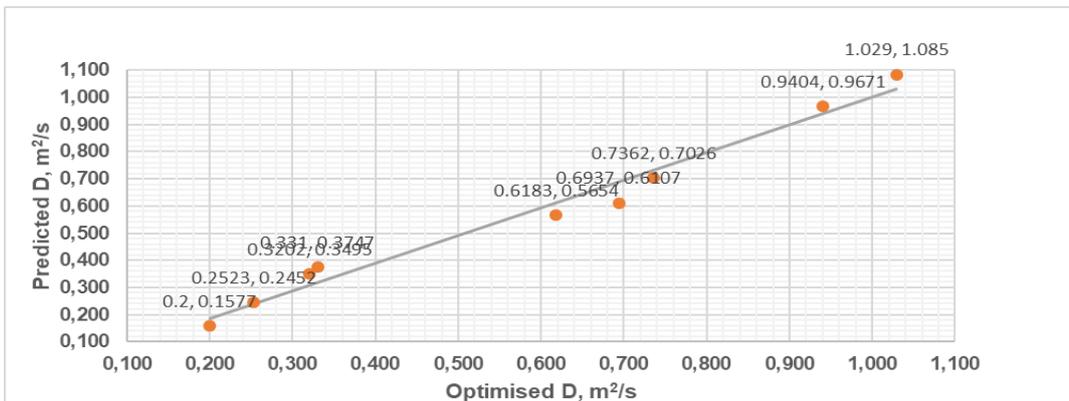


Figure 6-1: Predicted dispersion coefficients versus computed dispersion coefficients; $\Delta x = 18.40$ m (CN)

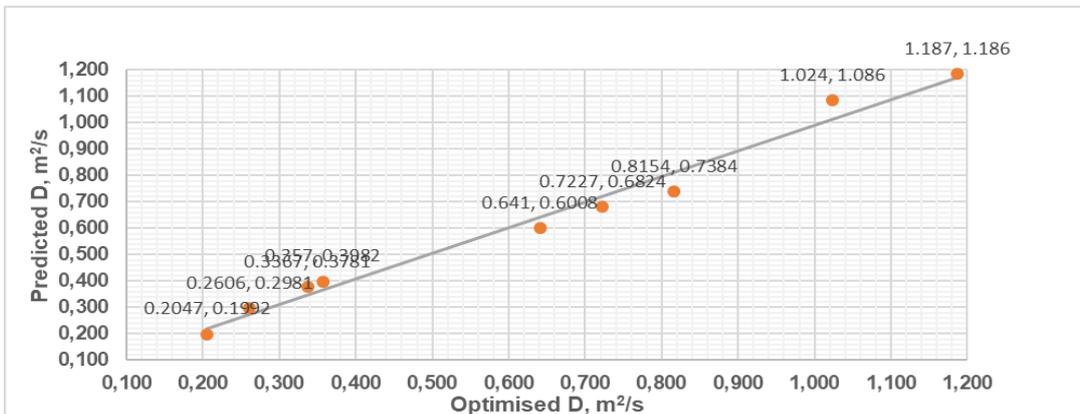


Figure 6-2: Predicted versus computed dispersion coefficients; $\Delta x = 18.40$ m (MacCormack)

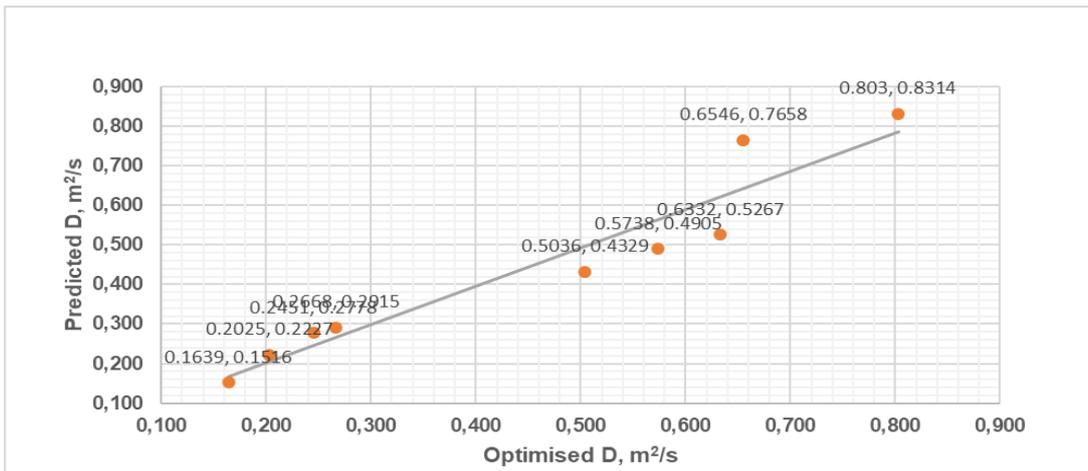


Figure 6-3: Predicted dispersion coefficient versus computed dispersion coefficient; $\Delta x = 18.40$ m (QUICKEST)

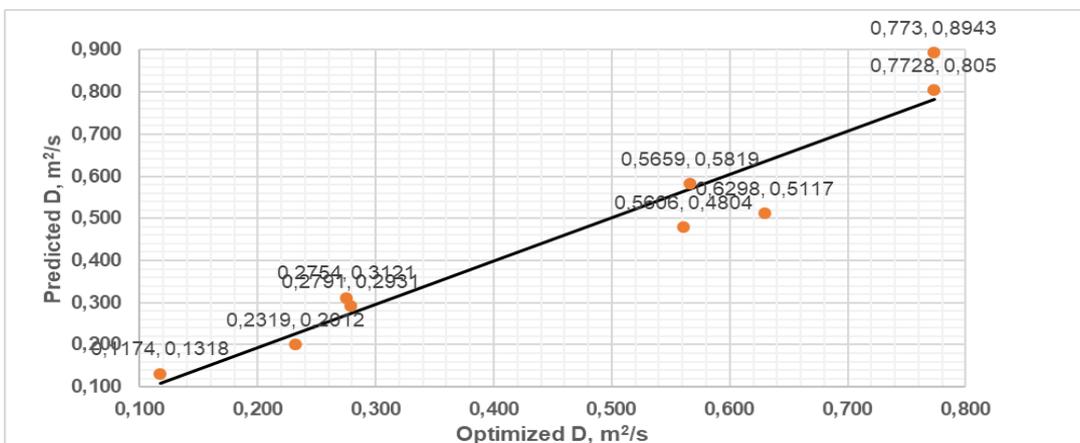


Figure 6-4: Predicted dispersion coefficient versus computed dispersion coefficient (routing method).

Descriptive performance measures included the minimum and the maximum predicted dispersion coefficients. It should be noted that the minimum values of dispersion coefficients are for the lowest flow rate, and the maximum values of dispersion coefficients are for the highest flow rate. For models based on numerical methods predicted values vary with the model type and space step. However, all the models overestimate for high flow rates and underestimate for low flow rates.

Error statistics included the root mean square error, the residual sum of squares and the variance. The values for RSS ranged from 0.018 to 0.055 for the Crank-Nicolson based model, 0.017 to 0.032 for the MacCormack based model and 0.028 to 0.051 for the QUICKEST based model. The values for RMSE are 0.051 to 0.088 for Crank-Nicolson based model, 0.049 to 0.068 for MacCormack based model and 0.063 to 0.086 for QUICKEST based model. The values of variance ranged from 0.003 to 0.008 for Crank-Nicolson based model, 0.002 to 0.005 for MacCormack based model and

0.004 – 0.007 for QUICKEST based model. Similarly, error statistics vary with model type and model resolution. However, the error statistics for all the models are comparable.

Residual error analysis statistics listed in tables are the coefficient of determination (R^2) and the model efficiency (E). The values for R^2 ranged from 0.911 to 0.977 for Crank-Nicolson based model, 0.953 to 0.982 for MacCormack based model and 0.809 to 0.953 for QUICKEST based model. Model efficiency values in all cases are very close to 1.0. The values for E ranged from 0.890 to 0.974 for the Crank-Nicolson based model, 0.947 to 0.982 for MacCormack based model and 0.782 to 0.951 for QUICKEST based model. Both the R^2 and the E were close to the optimal value of 1.0.

From Table 6-5 that the performance measures for model based on the Singh and Beck routing procedure are comparable to average values of performance measures for models based on numerical methods. Generally, the best performance metrics were obtained for the model based on the MacCormack numerical method.

Table 6-1: Space step, minimum and maximum computed and predicted dispersion coefficients (m^2/s) and performance metrics; CN-based model

Δx (meters)	Min. optimized D	Min. predicted D	Max. optimized D	Max. predicted D	R^2	r	Slope	RSS	Variance	RMSE	E
2.00	0.179	0.132	0.905	0.903	0.960	0.980	0.947	0.02225	0.00318	0.05637	0.953
4.00	0.184	0.145	0.907	0.991	0.956	0.978	1.039	0.02962	0.00423	0.06505	0.947
4.60	0.184	0.145	0.909	0.992	0.956	0.978	1.039	0.02951	0.00422	0.06493	0.947
5.75	0.184	0.145	0.911	0.995	0.957	0.978	1.039	0.02921	0.00417	0.06460	0.948
7.36	0.185	0.146	0.917	0.999	0.958	0.979	1.038	0.02856	0.00408	0.06387	0.949
8.00	0.185	0.146	0.920	1.002	0.959	0.979	1.038	0.02819	0.00403	0.06346	0.950
8.36	0.186	0.146	0.922	1.003	0.959	0.979	1.038	0.02799	0.00400	0.06323	0.951
9.20	0.186	0.147	0.927	1.007	0.960	0.980	1.037	0.02740	0.00391	0.06257	0.953
11.50	0.189	0.143	0.872	0.938	0.911	0.955	1.034	0.05481	0.00783	0.08849	0.890
12.27	0.190	0.150	0.953	1.027	0.966	0.983	1.034	0.02463	0.00352	0.05932	0.960
13.14	0.191	0.151	0.962	1.033	0.967	0.984	1.033	0.02372	0.00339	0.05821	0.962
14.15	0.193	0.152	0.974	1.042	0.969	0.985	1.031	0.02265	0.00324	0.05688	0.964
16.73	0.197	0.155	1.006	1.067	0.974	0.987	1.026	0.01996	0.00285	0.05340	0.971
18.40	0.200	0.158	1.029	1.085	0.977	0.989	1.023	0.01837	0.00262	0.05123	0.974
Average	0.188	0.147	0.937	1.006	0.959	0.979	1.028	0.02763	0.00395	0.06226	0.951

Table 6-2: Space step, minimum and maximum optimised and predicted dispersion coefficients (m^2/s) and performance metrics; MacCormack-based model

Δx (meters)	Min. optimized D	Min. predicted D	Max. optimized D	Max. predicted D	R ²	r	Slope	RSS	Variance	RMSE	E
2.00	0.182	0.168	0.917	0.998	0.953	0.976	1.015	0.03169	0.00453	0.06728	0.949
4.00	0.184	0.169	0.937	1.007	0.953	0.976	1.016	0.03228	0.00461	0.06791	0.947
4.60	0.185	0.170	0.944	1.012	0.954	0.977	1.014	0.03168	0.00453	0.06728	0.949
5.75	0.185	0.172	0.960	1.022	0.957	0.978	1.011	0.03042	0.00435	0.06592	0.953
7.36	0.187	0.176	0.995	1.046	0.962	0.981	1.004	0.02766	0.00395	0.06286	0.960
8.00	0.187	0.176	0.995	1.046	0.962	0.981	1.004	0.02766	0.00395	0.06286	0.960
8.36	0.188	0.177	1.001	1.050	0.963	0.981	1.003	0.02721	0.00389	0.06235	0.961
9.20	0.189	0.178	1.015	1.060	0.965	0.982	1.000	0.02617	0.00374	0.06115	0.964
11.50	0.192	0.183	1.057	1.089	0.971	0.985	0.993	0.02341	0.00334	0.05783	0.970
12.27	0.193	0.185	1.071	1.100	0.973	0.986	0.995	0.02211	0.00316	0.05620	0.972
13.14	0.195	0.187	1.088	1.111	0.974	0.987	0.987	0.02161	0.00309	0.05557	0.974
14.15	0.196	0.189	1.107	1.125	0.976	0.988	0.982	0.02075	0.00296	0.05444	0.976
16.73	0.201	0.195	1.156	1.162	0.980	0.990	0.975	0.01827	0.00261	0.05109	0.980
18.40	0.205	0.199	1.187	1.186	0.982	0.991	0.970	0.01702	0.00243	0.04930	0.982
Average	0.191	0.180	1.031	1.072	0.966	0.983	0.998	0.02557	0.00365	0.06015	0.964

Table 6-3: Space step, minimum and maximum computed and predicted dispersion coefficients (m^2/s) and performance metrics; QUICKEST-based empirical model

Δx (meters)	Min. Optimized D	Min. Predicted D	Max. Optimized D	Max. Predicted D	R ²	r	Slope	RSS	Variance	RMSE	E
8.00	0.183	0.174	0.858	0.953	0.938	0.969	1.003	0.03623	0.00518	0.07194	0.933
8.36	0.200	0.174	0.858	0.954	0.939	0.969	1.016	0.03608	0.00515	0.07179	0.931
9.20	0.182	0.173	0.858	0.952	0.940	0.970	1.006	0.03504	0.00501	0.07075	0.935
11.50	0.179	0.174	0.894	0.952	0.952	0.976	0.994	0.02796	0.00399	0.06320	0.950
12.27	0.178	0.173	0.896	0.949	0.953	0.976	0.996	0.02752	0.00393	0.06270	0.951
13.14	0.177	0.172	0.894	0.942	0.951	0.975	0.993	0.02792	0.00399	0.06316	0.949
14.15	0.175	0.170	0.888	0.930	0.948	0.974	0.989	0.02893	0.00413	0.06429	0.946
15.33	0.172	0.166	0.874	0.911	0.943	0.971	0.985	0.03073	0.00439	0.06625	0.941
16.73	0.169	0.160	0.848	0.880	0.933	0.966	0.980	0.03355	0.00479	0.06923	0.930
18.40	0.164	0.152	0.803	0.831	0.915	0.957	0.969	0.03792	0.00542	0.07361	0.911
20.44	0.157	0.138	0.729	0.758	0.884	0.940	0.965	0.04343	0.00620	0.07876	0.874
23.00	0.147	0.117	0.608	0.644	0.809	0.899	0.940	0.05138	0.00734	0.08567	0.782
Average	0.174	0.162	0.834	0.888	0.926	0.962	0.986	0.03472	0.00496	0.07011	0.919

Table 6-4: Minimum and maximum computed and predicted dispersion coefficients (m^2/s) and performance metrics for routing method-based model.

Min. optimized D	Min. predicted D	Max. optimized D	Max. predicted D	R ²	r	Slope	RSS	Variance	RMSE	E
0.117	0.132	0.773	0.894	0.930	0.96	1.030	0.03864	0.00552	0.07429	0.919

Table 6-5: Average values of performance metrics for numerical method-based models and for routing method-based model.

Empirical model	Min. Optimized D	Min. Predicted D	Max.Optimized D	Max.Predic	R ²	r	Slope	RSS	Variance	RMSE	E
Crank-Nicolson based	0.188	0.147	0.937	1.006	0.959	0.979	1.028	0.02763	0.00395	0.06226	0.951
MacCormack-based	0.191	0.180	1.031	1.072	0.966	0.983	0.998	0.02557	0.00365	0.06015	0.964
QUICKEST-based	0.174	0.162	0.834	0.888	0.926	0.962	0.986	0.03472	0.00496	0.07011	0.919
Routing procedure-based	0.117	0.132	0.773	0.894	0.930	0.964	1.030	0.03864	0.00552	0.07429	0.919

Plots of normalised errors versus predicted dispersion coefficients are presented in Appendix D for models based on numerical methods for all space steps (i.e. 2.00 m to 18.40 m for Crank-Nicolson and MacCormack based models, and 8.00 m to 23.00 m for the QUICKEST based model). The plots were obtained using Analyse-it for Microsoft Excel (2009) statistical software.

Figures 6-5 to 6-7 show plots of normalised residuals versus predicted dispersion coefficient for space step of 18.40 m. Figure 6-8 shows a plot of normalised errors versus the predicted dispersion coefficient for the model based on the routing procedure.

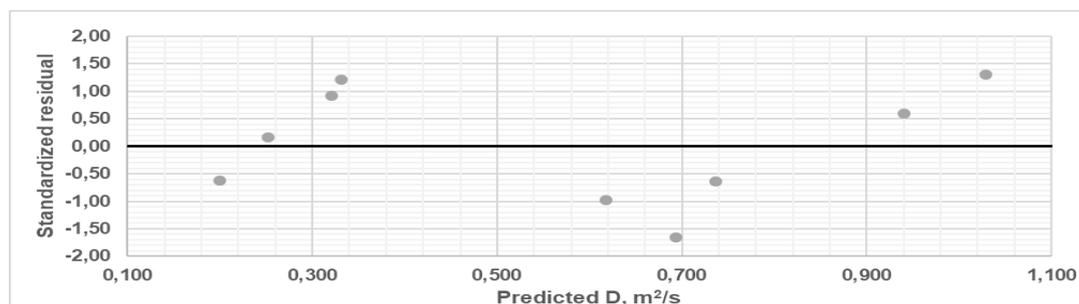


Figure 6-5: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; CN-based model.

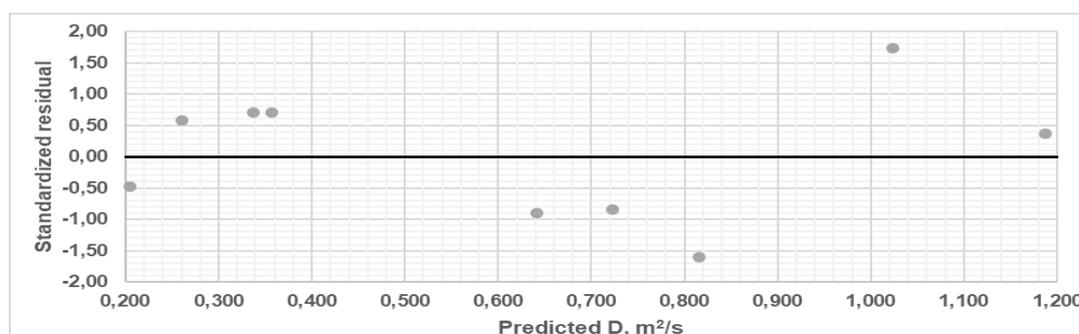


Figure 6-6: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; MacCormack based model.

It can be observed that for all models, normalised residuals fall in the interval $-2, +2$ and show a constant band of error variance with predicted values of dispersion coefficients. This illustrates that the errors are normally distributed and that predictions and observed dispersion coefficient are linearly related.

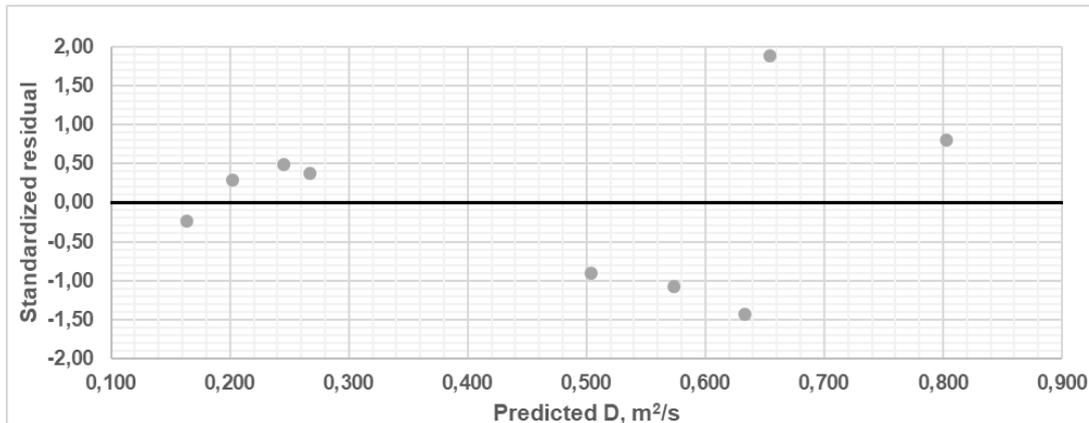


Figure 6-7: Normalised residuals versus predicted dispersion coefficient, $\Delta x = 18.40$ m; QUICKEST based model.

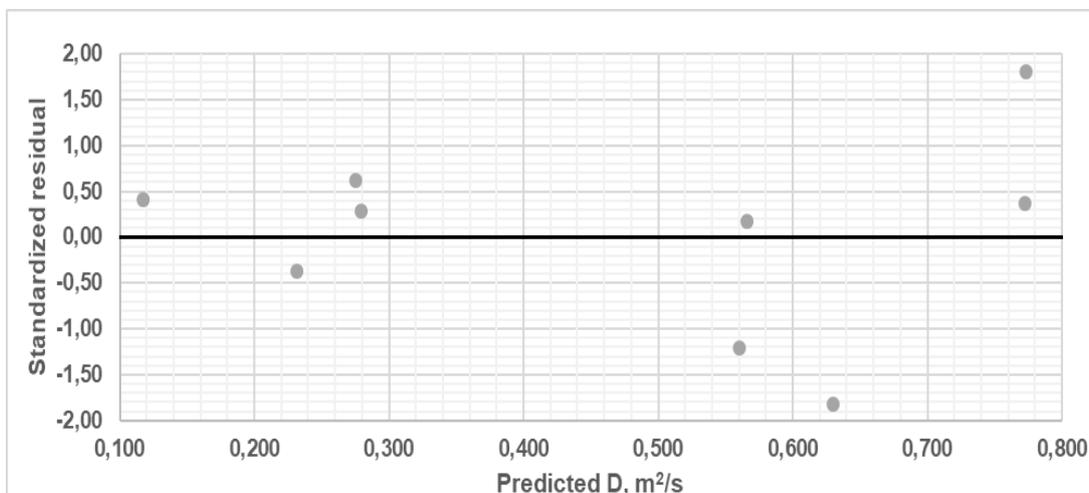


Figure 6-8: Normalised residuals versus predicted dispersion coefficient; routing method-based model.

6.4 Conclusion

Several performance metrics were quantified to characterise the performance of the models, namely, descriptive statistics, concurrent comparison, error statistics and residual analyses. Based on the slope of regression line and minimum and maximum predicted values it was observed that all models

are biased such that they overestimate large values and underestimate smaller values of dispersion coefficient. However, values of the slope of regression lines for all models were close to one, and that there was a high correlation between predicted and optimised dispersion coefficients.

Error indices given by the RMSE, RSS and variance between optimised and predictions showed small variations with model type and space step. However, the empirical model based on MacCormack numerical method gave the lowest values of error indices. These error statistics mean that the performance of models is comparable over the range of space steps as listed in Tables 6-1 to 6-3. Additionally, the model based on Singh and Beck routing method was also comparable with those based on numerical methods.

A residual error analysis showed that models behaved linearly as shown by the coefficients of determination since values were close to one; the optimal value of R^2 . The values of model efficiency, E , for all models was close to one, which is the optimal value (Krause et al., 2005; Moriasi et al., 2015). Plots of normalised residuals versus predicted values showed that for all models, errors are normally distributed with constant variance. Regardless of differences in values of dispersion coefficients used to develop the models, all empirical models were comparable and adequate based on performance measures considered, namely, descriptive statistics, graphical methods, error indices, and error analyses. Therefore, adequate empirical models can be developed based on different experimental methods of the AD-Model.

7 CONFIRMATION OF NEW MODELS AND COMPARISON WITH EXISTING MODELS

7.1 Introduction

Once a calibrated model has been established, all that is known is that the model fits a set of model development and calibration data. A model can be used with confidence if it is confirmed (Chapra, 2008). Model confirmation requires running the calibrated model for a separate set of observational data that reflect independent conditions from those defined by calibration data. The model is confirmed as a successful prediction method for the range of conditions determined by the calibration and confirmation data sets (Chapra, 2008; Sun and Sun, 2015). Model confirmation is the process of showing that a model demonstrates the behaviour of the real world, and it is intended to verify that the model accurately represents the responses of the study site. Unlike calibration, the parameters obtained during calibration are kept constant during the confirmation process (Chapra, 2008).

Performance evaluation and corresponding performance assessment criteria are necessary aspects of confirming water quality models (Bennett et al., 2013; Moriasi et al., 2015). The approach used for characterising model performance depends on the field of application, the type of model and specific objectives (Bennett et al., 2013). Every modelling undertaking has distinctive aims and challenges (Bennett et al., 2013; Moriasi et al., 2015; Kohavi, 2016). Consequently, there are no standard performance evaluation criteria that can apply to all models. However, it is possible to prescribe some general elements of performance measures that are worthwhile in modelling (Bennett et al., 2013; Moriasi et al., 2015). Acceptable initial choices for performance measures are those that have been used widely and can be easily communicated, such as the R^2 and the RMSE (Bennett et al., 2013). However, the RSS has been extensively used for such work (e.g. Ani et al., 2009; Wallis et al., 2013). Therefore, the RSS and the R^2 were considered for model confirmation.

Empirical models for predicting longitudinal dispersion coefficients in natural streams are based on correlating estimates of longitudinal dispersion from experimental methods to hydraulic and channel characteristics. The existence of such relationships implies that predicted dispersion coefficients by

such models are influenced by channel and hydraulic characteristics, the method used to estimate parameters and the conditions under which they were developed. However, several workers have attempted to develop a single model to be used for all cases (Wallis and Manson, 2004). Part of the thesis statement states that the reliability of empirical models for parameter estimation is influenced by the method used to estimate parameters and the conditions under which parameters were obtained for their development. Therefore, it was worthwhile to test the adequacy of some of the existing empirical models that were developed under different conditions for application to the Murray stream. Three widely used existing empirical models proposed by Seo and Cheong (1998), Fisher (Fischer et al., 1979) and Kashefipour and Falconer (2002) were compared with new models.

7.2 The method: confirmation of empirical models

Model confirmation (or evaluation) data consisted of experiments 4, 9 and 13 and contrary to calibration the model parameters (α , β and γ) determined during calibration, based on experimental methods (numerical methods and routing procedure) were kept constant. The developed empirical models were used to predict dispersion coefficient using estimated hydraulic and channel characteristics estimated from model confirmation data. The predicted dispersion coefficients were then used with numerical methods to simulate downstream concentration profiles of model confirmation data. Concentration predictions were achieved using the three numerical methods, viz. Crank-Nicolson, MacCormack and QUICKEST. The methods were applied for all space steps used in the calibration process, i.e. 2.00 m – 18.40 m for the Crank-Nicolson and the MacCormack methods, and 8.00 m – 23.00 m for QUICKEST. The stream reach was 184.00 m, and estimated flow rates were 0.084 m³/s, 0.037 m³/s and 0.148 m³/s for experiments 4, 9 and 13, respectively. The flow rates were determined by dilution gauging (Wallis, 2005). The channel characteristics were estimated using estimated flow rates and computed velocities. For each simulation, the performance of a numerical model was evaluated using the residual sum of squares (RSS) and the coefficient of determination (R^2) statistics as a measure of the discrepancy between the observed BTCs and output of numerical methods. Lower values of RSS between measured breakthrough curves (BTCs) and predicted profiles indicated better matching of observed profiles and predicted profiles (Wallis, 2013), while higher values of R^2 indicated better matching of measured BTCs and

predicted profiles.

7.2.1 Results and discussion: confirmation of empirical models

Three numerical models were applied to three data sets comprising experiment 4, 9, and 13. The numerical methods used predicted dispersion coefficients obtained by the developed empirical models, while the velocity values used in predictions were obtained by optimisation. Dispersion coefficients that were predicted by the empirical models were used by respective numerical methods and space steps, except for the empirical model which was developed using the Singh and Beck routing procedure. Dispersion coefficients that were predicted by the empirical model based on the Singh and Beck routing procedure were used with all three numerical methods and all space steps.

Predicted dispersion coefficients are presented in Appendix E-1 for all empirical models and confirmation experiments (experiments 4, 9 and 13). Figure 7-1 to 7-3 below show plots of predicted dispersion coefficients given by models based on numerical methods versus Peclet number for experiments 4, 9 and 13, respectively. It can be observed that predicted dispersion coefficients vary with the model type and Peclet number. Predictions by Crank-Nicolson and the MacCormack based models increase with Peclet number. Predictions by the QUICKEST based model generally decrease with Peclet number. Similar trends were observed with computed dispersion coefficients which are shown in Figures 4-9 to 4.11, although for different experiments. This illustrates that predictions by empirical models based on numerical methods of the AD-Model are influenced by numerical methods which were used to estimate parameters for their development and numerical properties. It should be noted that the empirical model based on the routing procedure predicts a single value of dispersion coefficient for each flow rate/experiment and the predicted dispersion coefficients for the three experiments are presented in Appendix E-1.

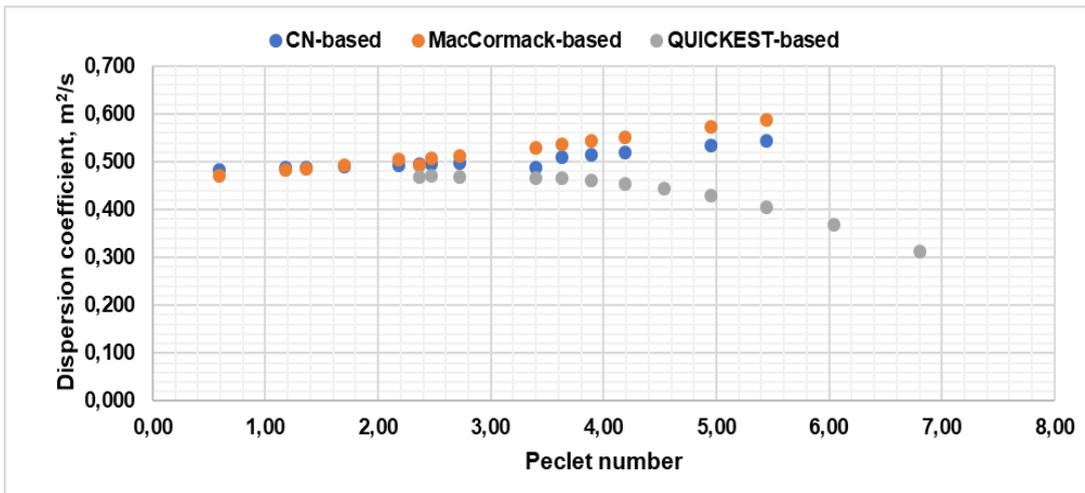


Figure 7-1: Predicted dispersion coefficient versus the Peclet number; experiment 4.

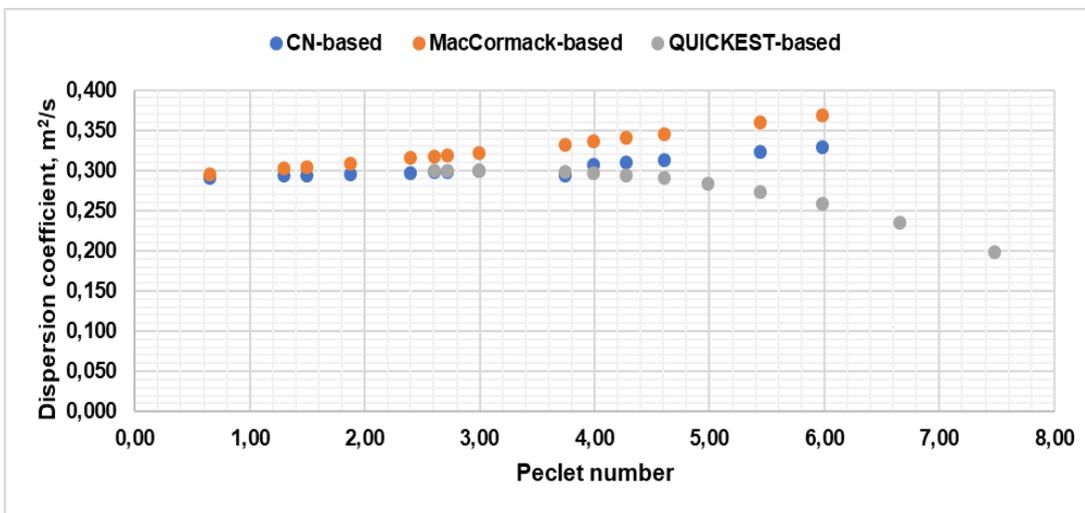


Figure 7-2: Predicted dispersion coefficient versus Peclet number; experiment 9.

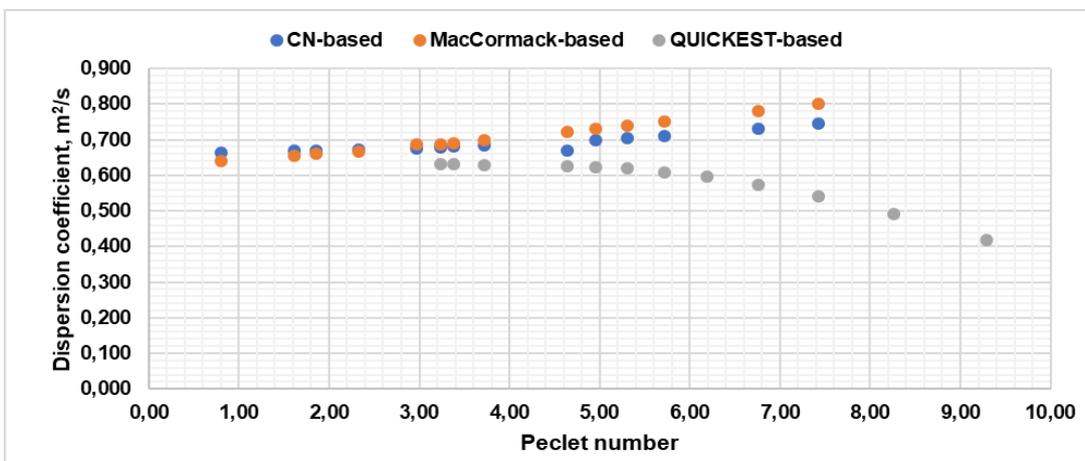


Figure 7-3: Predicted dispersion coefficients versus Peclet number; experiment 13.

Appendix E-2 presents the residual sum of squares between predicted concentration profiles and observed concentration profiles when predicted dispersion coefficients obtained by empirical models based on numerical methods were used with respective numerical methods.

Figures 7-4 to 7-6 show plots of the residual sums of squares versus Peclet number for experiments 4, 9 and 13, respectively. It can be observed that RSS vary with the model type and Peclet number. Similar trends were observed with optimised dispersion coefficients shown in Figures 4-15 to 4-17 above.

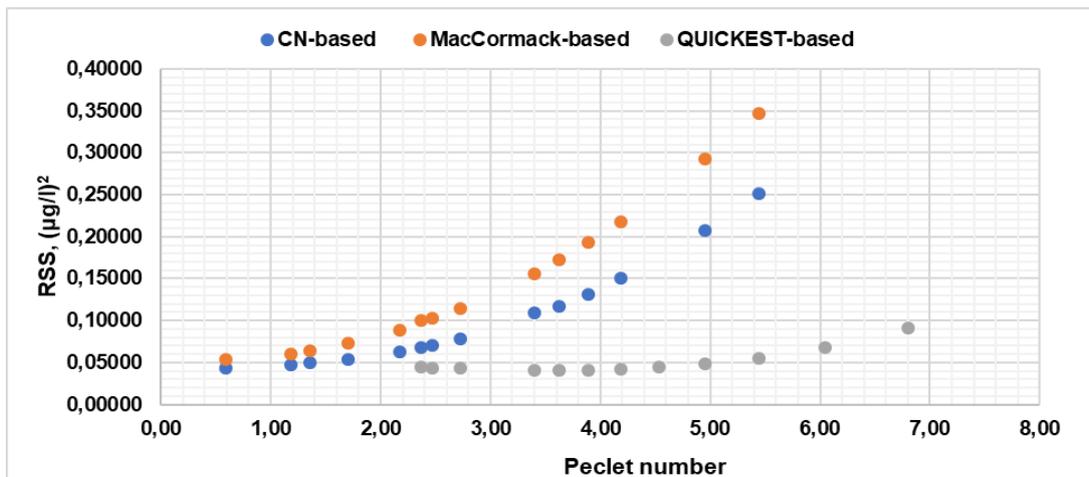


Figure 7-4: Plot of the residual sum of squares versus Peclet number; experiment 4

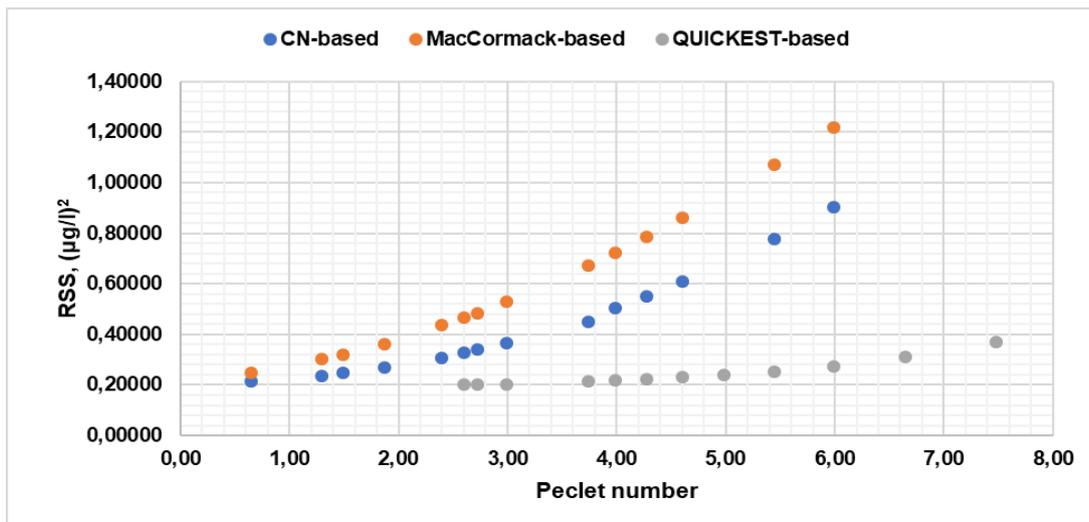


Figure 7-5: Plot of the residual sum of squares versus Peclet number; experiment 9

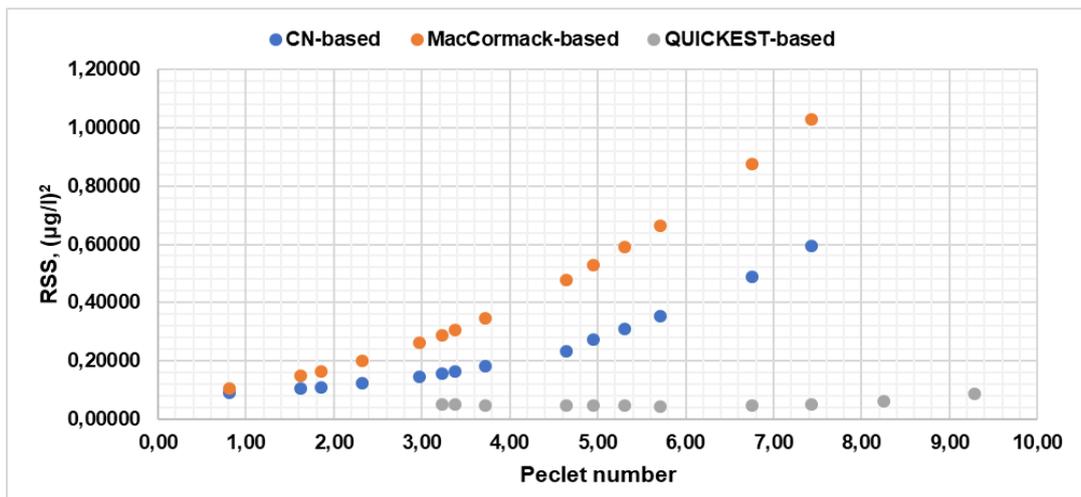


Figure 7-6: Residual sum of squares (RSS) versus Peclet number; experiment 13

Appendix E-3 presents coefficients of determination, which provides a measure of how well the downstream profiles are simulated by predicted concentration profiles, where empirical models based on numerical methods predicted dispersion coefficients.

Figures 7-7 to 7-9 show plots of coefficients of determination versus Peclet number for experiments 4, 9 and 13, respectively. It can be observed that the values of the coefficient of determination vary with the model type and Peclet number. Additionally, less accurate values of coefficients of determination are obtained with an increase in Peclet number.

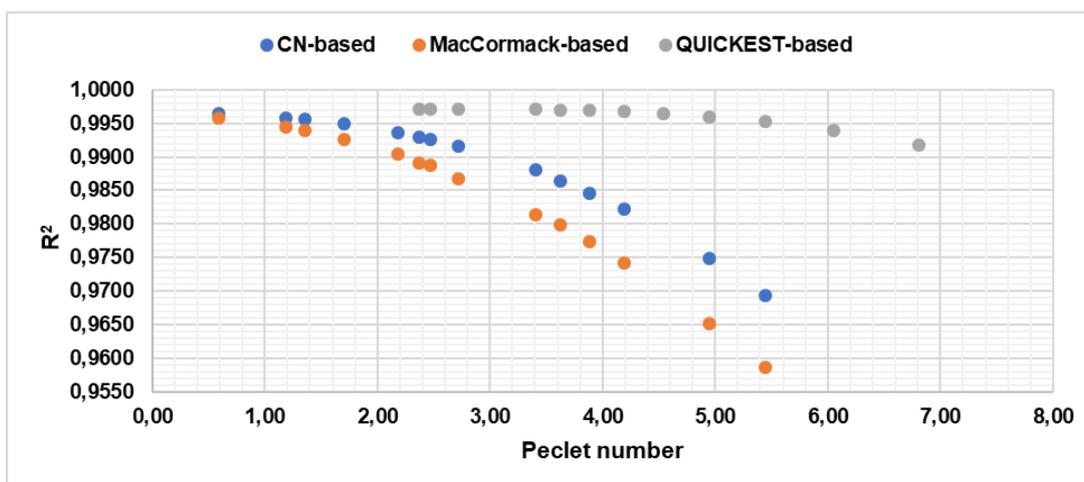


Figure 7-7: Coefficient of determination versus Peclet number; experiment 4

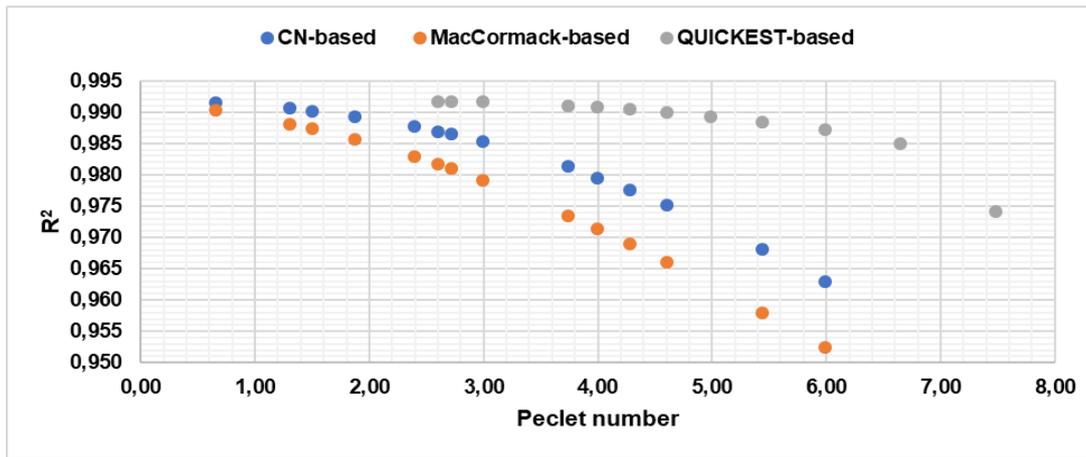


Figure 7-8: Coefficient of determination versus Peclet number; experiment 9

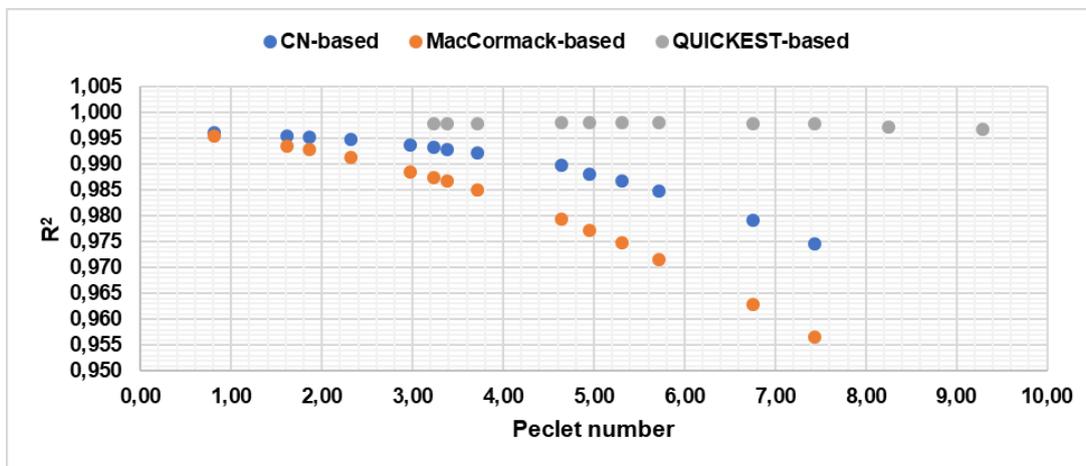


Figure 7-9: Coefficient of determination versus Peclet number; experiment 13

Appendix E-4 presents charts of concentration predictions using predicted dispersion coefficients obtained by empirical models based on numerical methods and respective numerical methods. It was observed in Section 5.4.5 (Figure 5-10) that marked differences in predictions can be observed at high flow rates and large Peclet numbers. Figures 7-4 to 7-6 show simulations of temporal concentration profiles given by numerical methods using predicted dispersion coefficients by empirical models based on numerical methods for experiment 13, which comprised the highest flow rate of the confirmation data. Each figure presents simulations for a small space step (2.00 m for Crank-Nicolson and MacCormack methods, and 8.00 m for QUICKEST method) and a large space step (18.40 m for Crank-Nicolson and MacCormack methods, and 23.00 m for QUICKEST method). It can be observed that more accurate simulations are achieved at small space steps than at large space steps. It can also be observed that the quality of simulations at large space steps varies with

the method, with the QUICKEST method showing the best fit than Crank-Nicolson and MacCormack methods.

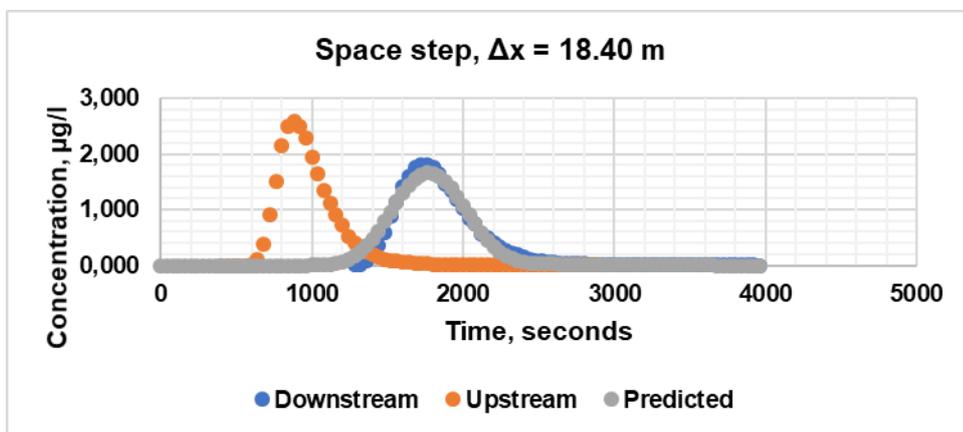
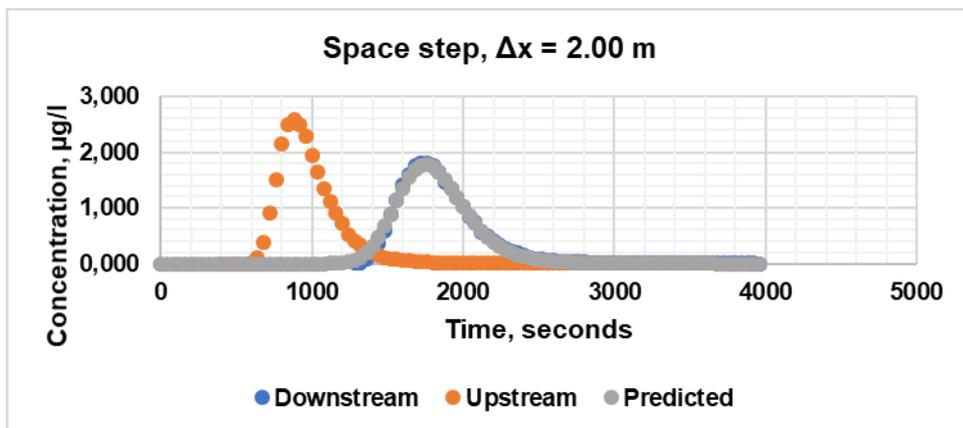
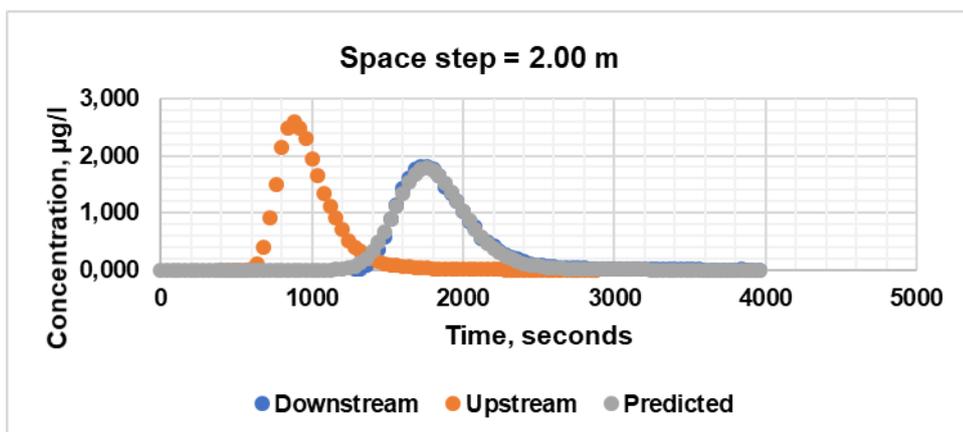


Figure 7-10: Fitting concentration profiles, dispersion coefficient predicted by CN-based model; experiment 13



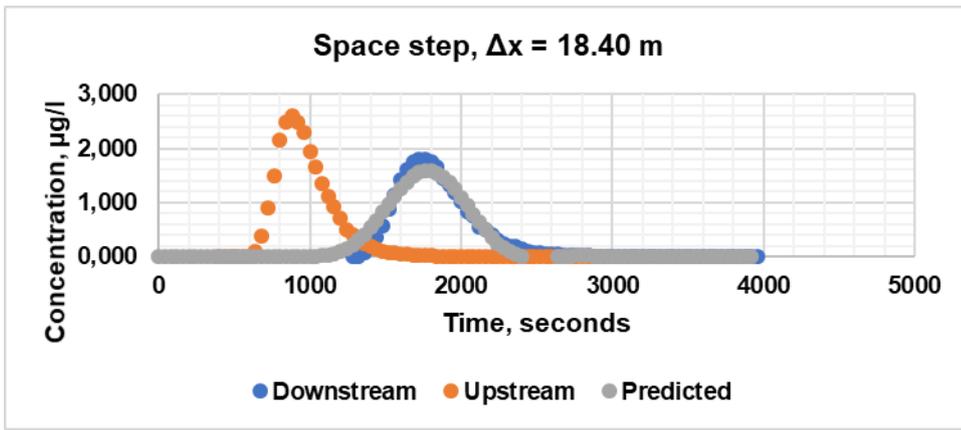


Figure 7-11: Fitting concentration profiles, dispersion coefficient predicted by MacCormack-based model; experiment 13

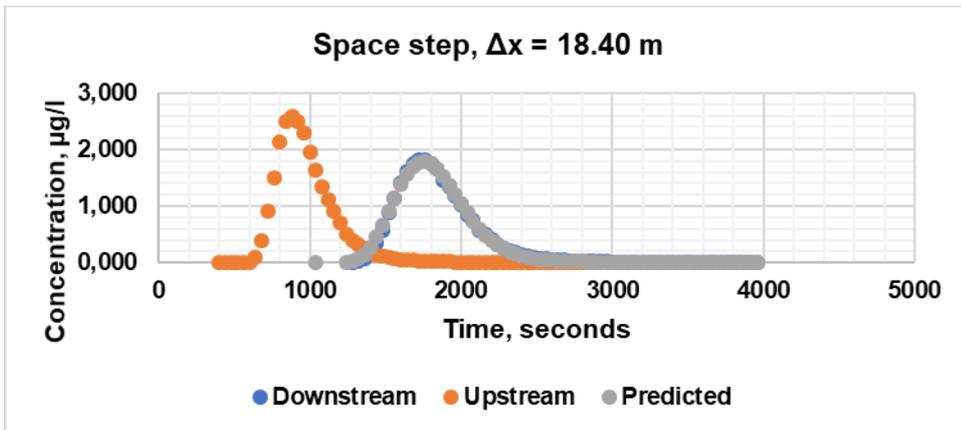
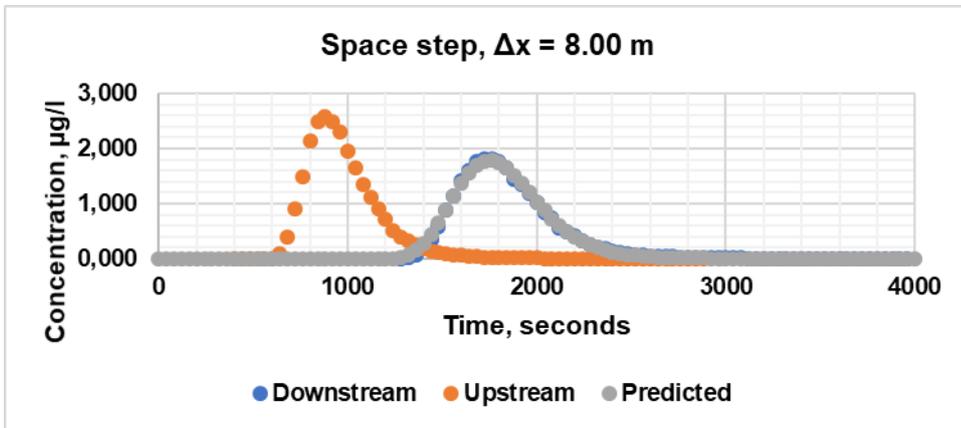


Figure 7-12: Fitting concentration profiles, dispersion coefficient predicted by QUICKEST-based model; experiment 13

Appendix F-3 presents charts of concentration predictions using predicted dispersion coefficients obtained by the empirical model based on the Singh and Beck routing procedure. The predicted dispersion coefficients were used with numerical methods to simulate downstream concentration

profiles of confirmation data.

Figures 7-13 to 7-15 show simulations for experiment 13 achieved by numerical methods using predicted dispersion coefficients obtained by the routing method-based empirical model. The figures show simulations that were obtained at small space steps (2.00 m for Crank-Nicolson and MacCormack methods, 8.00 m for QUICKEST method) and large space steps of 18.40 m for numerical methods. It can be observed that all methods at small space steps achieved accurate fits. In contrast, inaccurate simulations were obtained at large space steps by Crank-Nicolson and MacCormack methods. The results show that predictions given by the model based on the routing procedure achieve good simulations at small values of Peclet number, expected at small space steps.

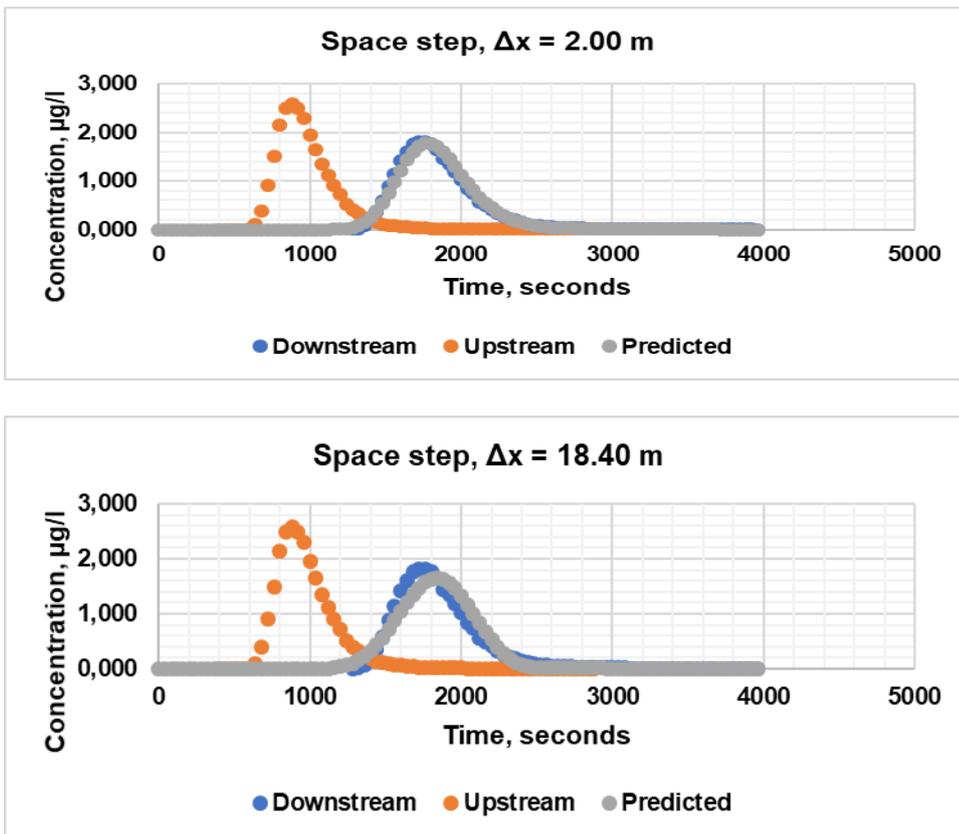


Figure 7-13: Fitting concentration profiles by Crank-Nicolson method, dispersion coefficient predicted by routing method-based model; experiment 13

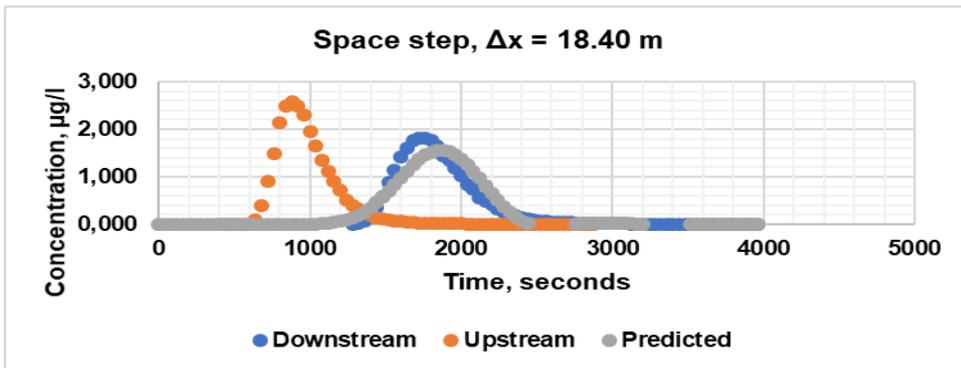
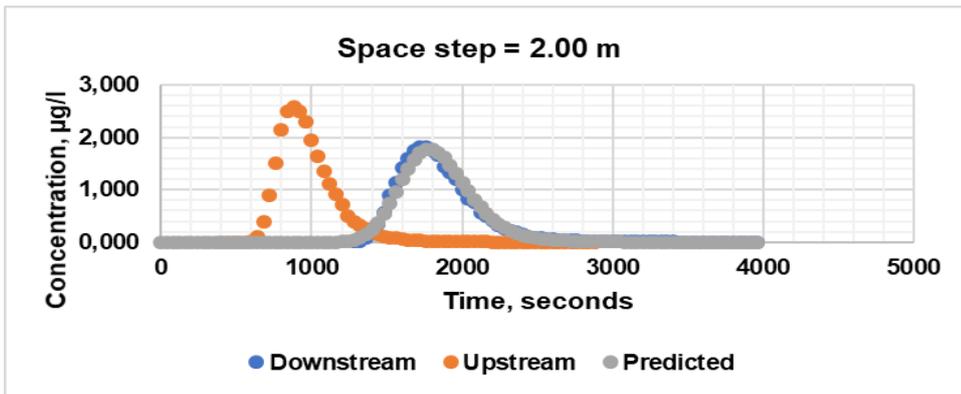


Figure 7-14: Fitting concentration profiles by MacCormack method, dispersion coefficient predicted by routing method-based model; experiment 13.

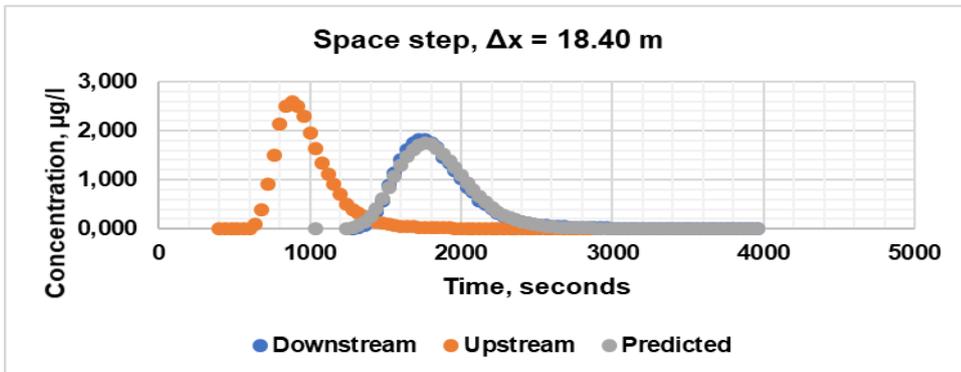
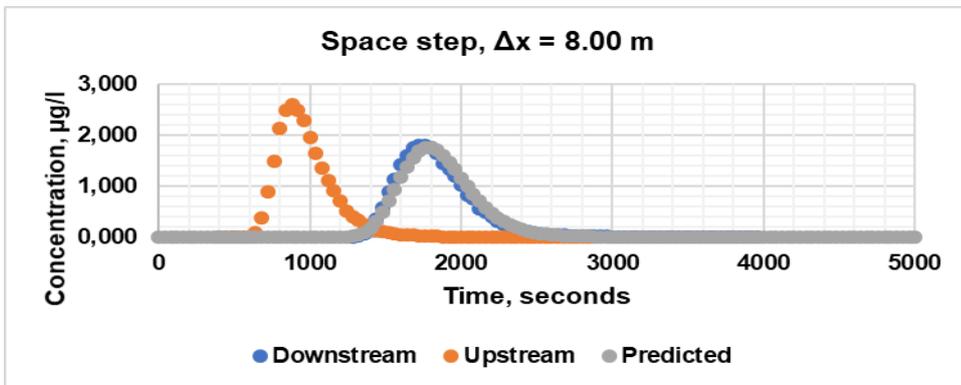


Figure 7-15: Fitting concentration profiles by QUICKEST method, dispersion coefficient predicted by routing method-based model, experiment 13.

Appendices F-1 and F-2 present values of the residual sum of squares (RSS) and coefficients of determination (R^2) for simulations using predicted dispersion coefficients obtained by the empirical model based on the routing procedure. The values of RSS and R^2 were obtained over the whole range of space steps for all sets of confirmation data.

Figure 7-16 compares the variation of residual sums of squares with the Peclet number, obtained after concentration prediction by numerical methods for experiment 13.

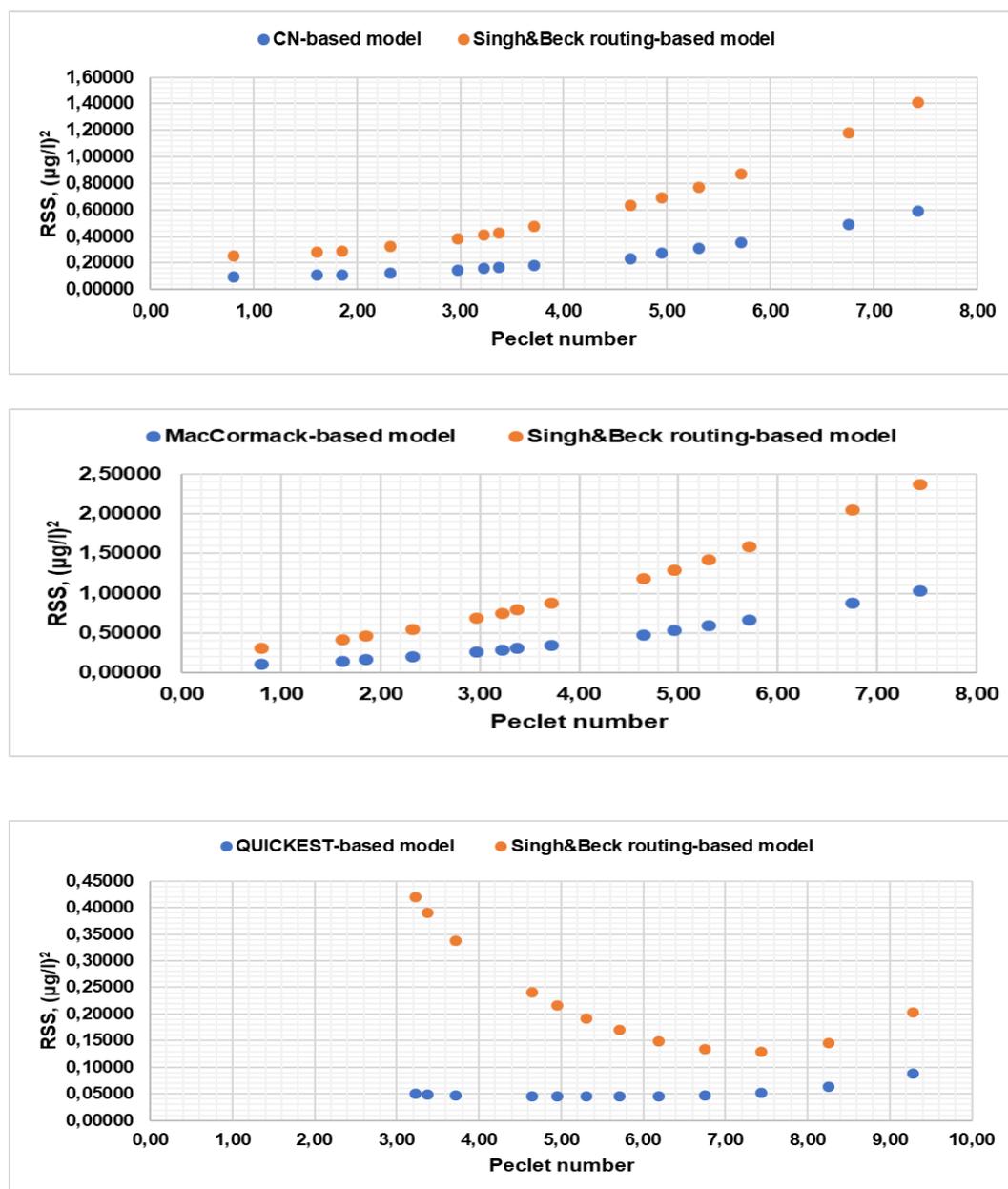
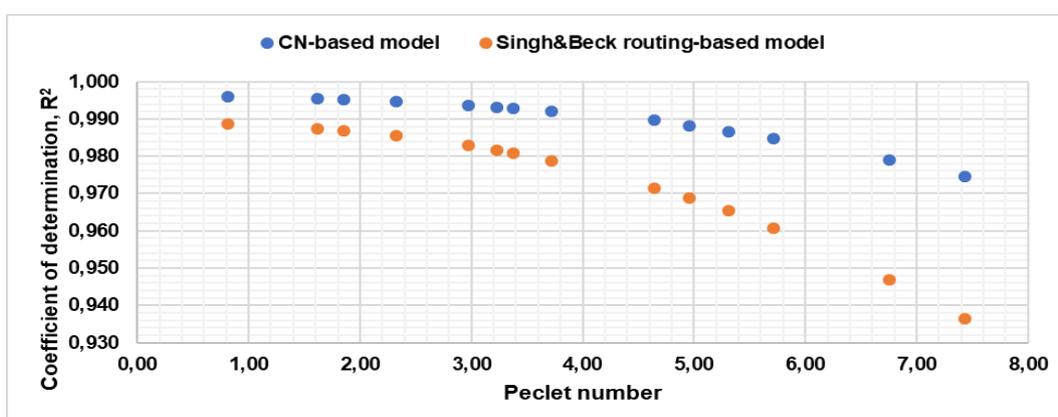


Figure 7-16: Residual sum of squares (RSS) versus Peclet number; experiment 13.

The residual sums of squares were also obtained by use of predicted dispersion coefficient given by empirical models based on each numerical method and the empirical model based on the Singh and Beck routing procedure for experiment 13. It can be observed that, generally, use of predicted dispersion coefficients obtained by the empirical model based on the routing procedure results in higher values of RSS which varies with the numerical method used for concentration prediction and Peclet number. Additionally, for Crank-Nicolson and MacCormack methods, use of predicted dispersion coefficient given by the model based on the routing procedure, values of RSS increase with Peclet number. For the QUICKEST method, RSS is high at low values, decreases with Peclet number until it reaches a minimum value between Peclet numbers 7.0 and 8.0, after that it increases with Peclet number.

Figure 7-17 compares the variation of coefficients of determination with Peclet number, obtained after concentration prediction by numerical methods. The coefficients of determination were obtained by use of dispersion coefficient predicted by empirical models based on numerical methods and the empirical model based on the Singh and Beck routing procedure for experiment 13. It can be observed that, generally, use of predicted dispersion coefficients obtained by the empirical model based on the routing procedure results in lower values of the coefficients of determination which varies with the numerical method used for concentration prediction and Peclet number. Additionally, for Crank-Nicolson and MacCormack methods, use of predicted dispersion coefficient given by the model based on the routing procedure, values of R^2 decrease with Peclet number at a higher rate than when the numerical method-based model is used.



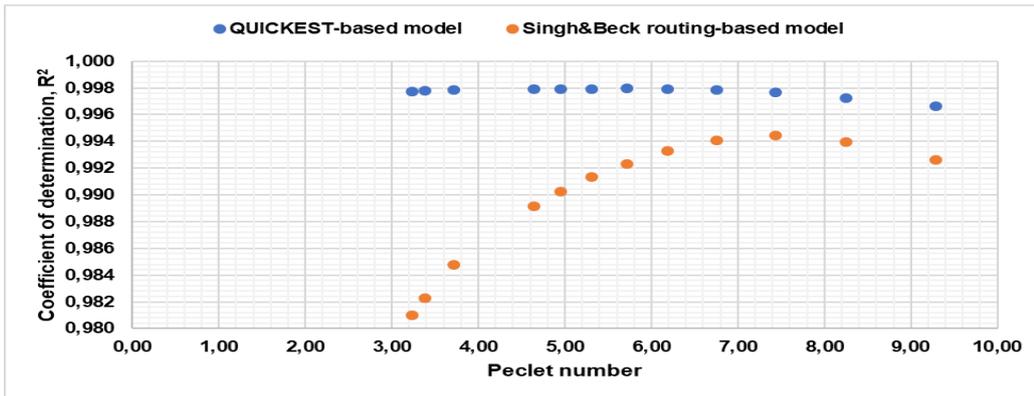
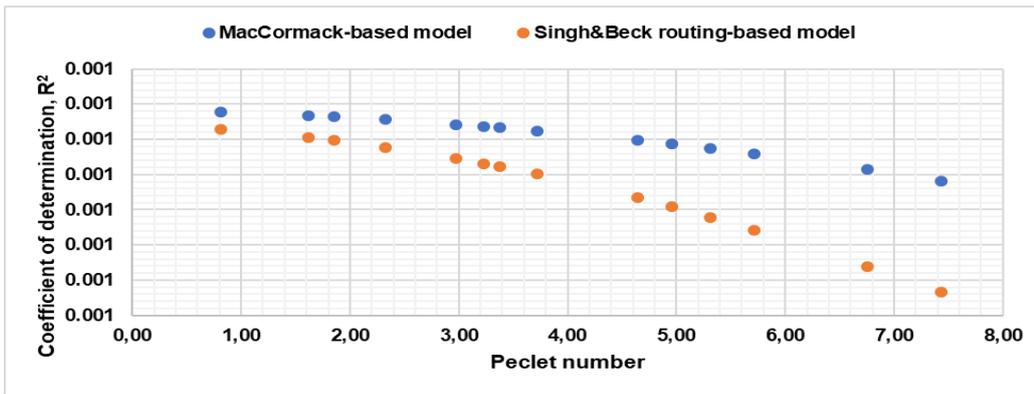


Figure 7-17: Coefficient of determination versus Peclet number; experiment 13

Figure 7-18 shows a comparison plot of the residual sum of squares obtained by use of dispersion coefficient predicted by routing-based empirical model versus Peclet number. It can be observed that at low Peclet numbers (less than 4.0), values of the residual sum of squares are comparable, while at high Peclet numbers, large values of RSS are obtained from the application of the Crank-Nicolson and MacCormack methods. Generally, values of RSS vary with the method used for concentration prediction.

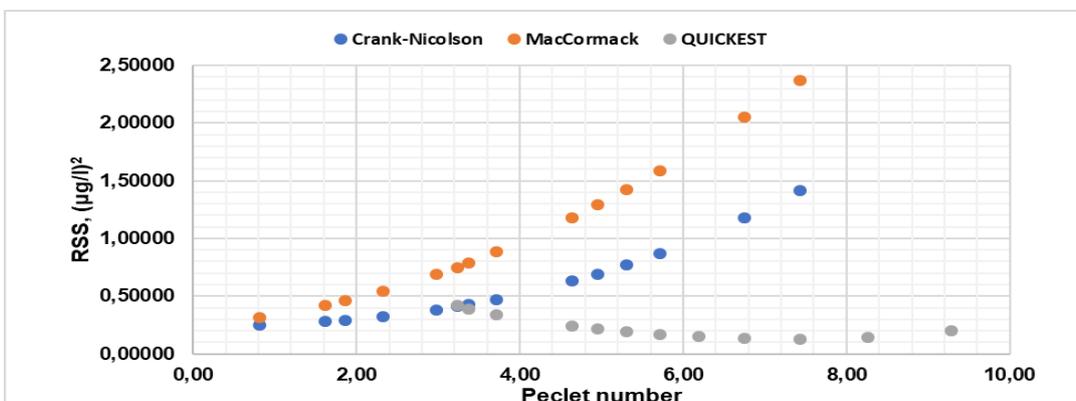


Figure 7-18: Analogy of RSS obtained by numerical methods, predicted dispersion coefficient given by the routing-based model; experiment 13.

Figure 7-19 shows a comparison plot of the coefficient of determination obtained by use of dispersion coefficient predicted by routing-based empirical model versus Peclet number. Similarly, as for RSS, it can be observed that at low Peclet numbers (less than 4.0), values of coefficients of determination achieved by the three numerical methods are comparable, while at high Peclet numbers, low values of coefficients of determination were achieved by Crank-Nicolson and MacCormack methods. Generally, values of R^2 vary with the method used for concentration prediction.

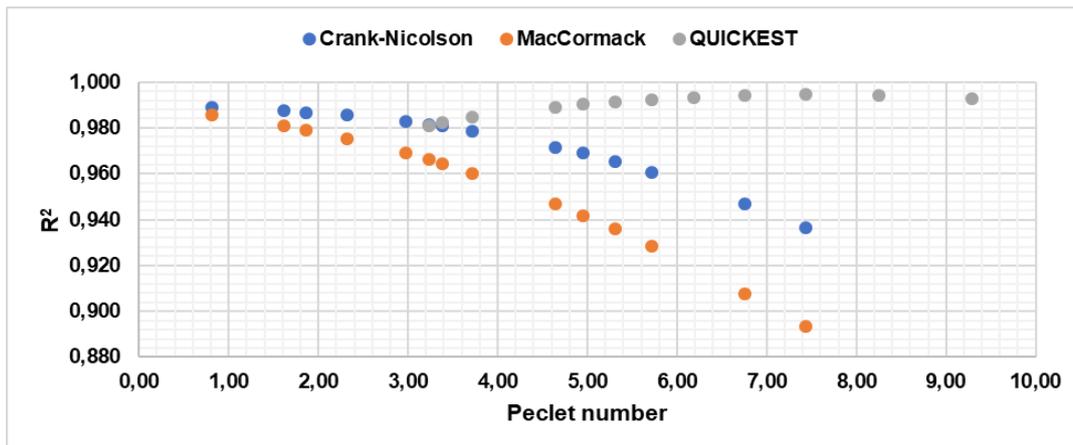


Figure 7-19: Analogy of R^2 achieved by numerical methods, predicted dispersion coefficient given by the routing-based model; experiment 13.

7.3 Comparison of new models with existing methods

Three existing empirical models widely used in literature together with the new models and the Singh and Beck routing procedure were applied to Murray stream data to predict values of longitudinal dispersion coefficients. Since longitudinal dispersion has been observed to be nonlinearly correlated with flow rate (Rutherford, 1994; Ani et al., 2009), predicted longitudinal dispersion coefficients were plotted against flow rate to assess the adequacy of models. The routing procedure was used as a reference benchmark in the comparison of estimated dispersion coefficients against empirical models.

The existing empirical models considered were those proposed by Fischer (1979), Seo and Cheong (1998) and Kashefipour and Falconer (2002). These empirical equations have been determined based on values of longitudinal dispersion coefficients observed from many rivers using the method

of moments and/or Fischer's routing method (Singh and Beck, 2003). Although longitudinal dispersion is influenced by discharge, one common feature of these methods is that flow rate is not included in their formulation.

Fischer (1979) empirical equation for predicting longitudinal dispersion coefficient is expressed as

$$D = 0.011 \frac{v^2 w^2}{HU_*}. \quad (7.1)$$

The equation proposed by Seo and Cheong (1998) developed using the one-step Huber regression method is expressed as:

$$D = 5.915 \left(\frac{w}{H} \right)^{0.620} \left(\frac{v}{U_*} \right)^{1.428} HU_*. \quad (7.2)$$

That proposed by Kashefipour and Falconer (2002) is expressed as

$$D = 10.612 H v \left(\frac{v}{U_*} \right). \quad (7.3)$$

The general form for the new models is expressed as

$$D = \alpha \left(\frac{Q}{H} \right)^\beta \left(\frac{v}{U_*} \right)^\gamma. \quad (7.4)$$

The parameters for new models that are based on numerical methods are given in Table 7-10 for space step of 18.40 m, and parameters for the model based on Singh and Beck routing procedure are given in Table 5-4. As observed above the accuracy of predictions is reduced with an increase in space step or Peclet number. Therefore, the largest common space step of 18.40 m was considered for comparison with existing models.

7.3.1 Results and discussion: comparison of models

Dispersion coefficients were determined by three well-known existing empirical models and four new

empirical models, and the Singh and Beck routing method from 11 experiments. The three existing empirical equations were applied to the Murray stream data to determine if they could make comparable predictions to both the routing procedure and the new models. The parameters (coefficients and exponents) for the new models are listed in Table 7-1.

Table 7-2 presents estimated dispersion coefficients estimated by the Singh and Beck routing procedure, three well known existing empirical models and the new models. The results were plotted against flow rate, as shown in Figure 7-20. It can be observed that trends of predictions given by the new models are like those obtained by the routing procedure. Kashefipour and Falconer's equation underestimates for low flows and overestimates for high flows. Seo and Cheong's equation consistently overestimates, while Fisher's model consistently underestimates.

Table 7-1: Parameters of new models at space step of 18.40 m

Model parameters	α	β	γ
CN-based model	0.649	0.233	0.779
MacCormack-based model	1.193	0.729	0.082
QUICKEST-based model	0.865	0.727	0.040

Table 7-2: Flow rates and dispersion coefficients estimated by new models, existing empirical models and Sing and Beck routing procedure.

Q (m ³ /s)	MacCormack-based	Kashefipour&Falconer	CN-based	Seo&Cheong	Fischer	QUICKEST-based	Routing-based	Singh&Beck's routing
0.017	0.203	0.024	0.159	0.131	0.005	0.148	0.141	0.117
0.037	0.369	0.101	0.329	0.358	0.017	0.259	0.285	0.249
0.041	0.399	0.123	0.362	0.410	0.020	0.279	0.315	0.279
0.044	0.426	0.140	0.391	0.449	0.022	0.296	0.335	0.275
0.084	0.587	0.369	0.544	0.864	0.039	0.406	0.489	0.497
0.097	0.630	0.454	0.583	0.991	0.043	0.435	0.526	0.561
0.147	0.803	0.824	0.745	1.472	0.057	0.543	0.652	0.566
0.148	0.800	0.854	0.746	1.514	0.060	0.544	0.671	0.525
0.150	0.718	0.728	0.626	1.319	0.040	0.495	0.569	0.630
0.385	1.154	2.826	1.005	3.249	0.077	0.775	0.941	0.773
0.436	1.286	3.528	1.139	3.794	0.093	0.855	1.048	0.773

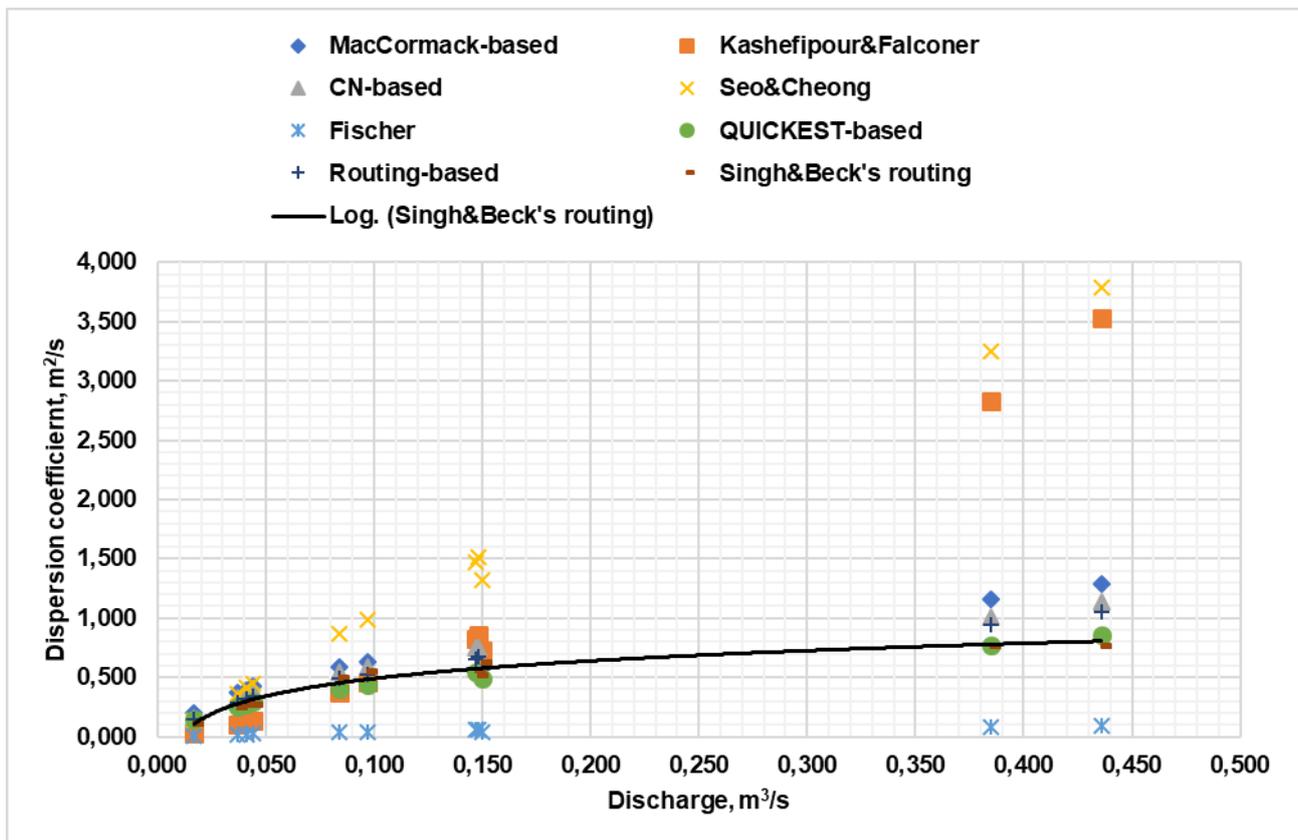


Figure 7-20: Plot of predicted dispersion coefficients obtained by several empirical methods versus flow rate.

7.4 Conclusion

Four new empirical models for predicting longitudinal dispersion coefficient for a stream reach on Murray stream confirmed using three sets of data, namely, experiments 4, 9 and 13. Three of the models were developed based on dispersion coefficients obtained by numerical methods (numerical method-based models), and one was developed based on dispersion coefficients obtained by Singh and Beck routing procedure (routing method-based model). The empirical models based on numerical methods were used to predict longitudinal dispersion coefficients under similar space steps as those covered during calibration. The empirical model based on the routing procedure predicted a single value for each of the confirmation data. The predicted dispersion coefficients were then used with numerical methods (Crank-Nicolson, MacCormack and QUICKEST) to simulate temporal concentration profiles of confirmation data. The quality of fits between predicted concentration-time profiles and observed concentration-time profiles were assessed by the coefficient of determination (R^2) and the residual sum of squares (RSS).

Furthermore, the new models were compared with existing models by plotting predicted dispersion coefficients against discharge. Predicted dispersion coefficients were further compared with optimised dispersion coefficients obtained by Singh and Beck routing procedure. Predicted dispersion coefficients obtained by numerical method-based models varied with model type and Peclet number for each of the confirmation data. Predicted dispersion coefficient obtained by Crank-Nicolson and MacCormack based models increased with increase in Peclet number, while those obtained by QUICKEST based model decreased with increase in Peclet number. Predicted dispersion coefficient obtained by the empirical models based on numerical methods were applied with respective numerical models, while predicted dispersion coefficients obtained by the empirical model based on the routing procedure were applied with all the three numerical methods and for all space step. Quality of fits varied with model type and Peclet number. Concentration predictions achieved by use of predicted dispersion coefficients obtained by numerical method-based models were more accurate than those obtained by use of dispersion coefficient predicted obtained by the routing method-based model, especially at high Peclet numbers. However, comparable accurate simulations were obtained by predicted dispersion coefficients obtained by both routing method-based and numerical method-based at low values of Peclet number.

Values of predicted dispersion coefficients obtained by existing models did not match those given by the routing procedure and new models. Existing empirical models either overestimated or underestimated dispersion coefficient with an increase in flow rate. Existing models may be well-founded within their calibrated hydraulic conditions and channel characteristics or may be used in a modified form as Ani et al. (2009) modified Fischer's model.

8 CONCLUSION AND RECOMMENDATIONS

8.1 Findings

The research has investigated the impact of numerical methods and numerical properties on estimated parameter values of the advection-dispersion model and construction of empirical models for estimating dispersion coefficient. Based on the results the following findings were observed from this research.

Appraisal of selected numerical methods of the advection-dispersion model

a) Although the order of a Eulerian numerical scheme is considered as a measure of accuracy, on a given grid (Ferziger and Peric, 2002), Eulerian schemes of the same order do not estimate the same dispersion coefficient for the same numerical properties. Therefore, when computing a dispersion coefficient, there is a risk of underestimating or overestimating dispersion coefficient if dispersion coefficients are to be estimated by optimising Eulerian numerical methods and the numerical method induces considerable numerical dispersion. However, different numerical methods achieve accurate concentration prediction with different parameter values.

Six different numerical methods were appraised, viz. Backward-Time/Centred-Space method, the Crank-Nicolson method, the MacCormack method, third-order upstream-differencing method, implicit scheme with QUICK and the QUICKEST method. The methods were applied using data in which solute transport parameters were known and were appropriate for application to Murray stream, based on previous studies. The methods were applied under the same range of Peclet number; 1.5 to 12.0. The best results were obtained by the Crank-Nicolson and MacCormack methods, with relative percentage errors ranging from 0.08 % to 10.25 % and 0.19 % to 14 %, respectively. The QUICKEST method was accurate only for a narrow range of Peclet number (2 -4). The Implicit Quick, Backward-Time/Centred-Space and third-order upstream-differencing schemes were quite inaccurate with relative errors ranging from 64 % to 97 %, 27 % to 70 % and 18 % to 63 %, respectively. The best performing numerical models for analysing breakthrough curves were found to be the Crank-Nicholson, MacCormack and QUICKEST methods. These were later used to develop new empirical models for the Murray Stream based

on bulk flow and channel characteristics.

Impact of numerical methods and numerical properties on stream estimated parameters

- b) Optimised parameter values are influenced by the numerical method used and the model resolution under which they are determined. However, optimised dispersion coefficients are more affected than optimised velocities.

Three methods were applied to observed temporal concentration profiles, viz. Crank-Nicolson, MacCormack and QUICKEST schemes. The methods were applied under similar space and time steps and optimised longitudinal dispersion coefficients and velocities were obtained. Optimised dispersion coefficients showed high variations with the method used and Peclet number, while optimised velocity values showed much lower variation with the method used and Peclet number. Optimised dispersion coefficients (m^2/s) obtained by Crank-Nicolson, MacCormack and QUICKEST methods ranged from 0.905 to 1.029, 0.917 to 1.031 and 0.858 - 0.608 respectively.

Development of empirical models based on numerical methods and numerical properties

- c) If a functional form of an empirical model is developed, values of model parameters (α, β, γ) vary with the numerical or experimental method used to develop the empirical model and non-dimensional numerical properties. Therefore, the structure of the model is influenced by the method used to estimate parameters and non-dimensional numerical properties under which parameters have been observed.

Three different sets of model parameter values were determined based on numerical methods (i.e. Crank-Nicolson, MacCormack and QUICKEST) and non-dimensional numerical properties, and one set of model parameters was determined based on the Singh and Beck routing procedure.

Impact of numerical methods and numerical properties on predicted dispersion coefficient by developed empirical models

- d) Predicted dispersion coefficients given by empirical models based on experimental methods are influenced by the numerical method used to estimate parameters and numerical properties

under which parameters have been observed for their development.

Three sets of dispersion coefficients were obtained by empirical equations based on numerical methods and Peclet number, and one set of dispersion coefficients was obtained by the empirical model based on the routing procedure. Values of dispersion coefficients obtained by empirical models based on numerical methods varied with model type and Peclet number. Additionally, predicted dispersion coefficients given by empirical models based numerical methods differed from predicted dispersion coefficients given by the model based on the routing procedure increasingly with an increase in the Peclet number.

Performance characterisation of empirical models based on numerical method and properties

- e) Adequate and comparable empirical models can be developed based on different experimental methods of the AD-Model, regardless of differences in parameter values used for their development.

Several performance metrics were quantified to characterise the performance of the models based on the Singh and Beck routing procedure and numerical methods. For numerical models, overall, values of the Peclet number ranged from 0.599 – 12.818. Performance measures included descriptive statistics, graphical methods, error statistics and residual error analyses. Regardless of differences in values of dispersion coefficients used to develop the models, all empirical models were comparable and adequate based on the performance measures considered. For example, average values of the coefficient of determination of 0.959, 0.966 and 0.926 were obtained for empirical models based on Crank-Nicolson, MacCormack and QUICKEST methods, respectively.

Impact of empirical model type and numerical properties on concentration predictions

- f) Concentration simulations by use of predicted dispersion coefficients obtained by empirical models based on the type of solution methods and numerical properties other than those that were used for its development result in poorer quality of concentration predictions, especially at high flow rates and Peclet numbers. For example, when predicted dispersion coefficients

given by routing method-based model were applied with numerical methods to simulate temporal concentration profiles of experiment 13 at space step of 18.40 m, values of RSS ($\mu\text{g}^2/\text{l}^2$) obtained by Crank-Nicolson, MacCormack and QUICKEST methods were 1.413, 2.371 and 0.129, respectively. In contrast, when predicted dispersion coefficients given by numerical method-based models were applied with numerical methods to simulate concentration profiles of experiment 13, values of RSS obtained by Crank-Nicolson, MacCormack and QUICKEST methods were 0.593, 1.027 and 0.052, respectively.

Comparison of new and existing empirical models

- g) Existing models which were developed under different conditions may not be reliably applied for new conditions. The new empirical predictive models were compared to generalised empirical models, proposed by Seo and Cheong (1998), Fischer (Fischer et al., 1979) and Kashefipour and Falconer (2002). It was found that a general application of such generalised models was not accurate for the Murray stream. This implies that empirical models should not be generally applied as is being commonly done currently and that existing models may be well-founded within their calibrated hydraulic conditions and channel characteristics.

8.2 Conclusion

The thesis read:

The reliability of an empirical model for parameter estimation of stream solute transport based on observed parameters by Eulerian numerical methods of the AD-Model is influenced by both the solution method of the AD-model and the numerical properties under which the parameters were obtained for its development, and the solution method and numerical properties for its application.

This statement was therefore supported by the results of the research as follows:

Optimised parameter values are influenced by the numerical method used, and the model resolution and that optimised dispersion coefficients are more affected than optimised velocities. Variability of optimised dispersion coefficients, therefore, affects the structure of developed empirical models such that model structure is dependent on the experimental method used to estimate parameters and

non-dimensional numerical properties under which parameters have been determined. Consequently, predicted dispersion coefficients given by empirical models vary with the structure of empirical models. Although adequate and comparable empirical models can be developed based on different experimental methods of the AD-Model, more accurate concentration simulations are achieved if predicted dispersion coefficients are applied with the AD-Model solution that was used for development of the predictor, especially at high flow rates and Peclet numbers. Thus, at low flow rates and Peclet numbers accurate concentration predictions can be achieved irrespective of the empirical model used to predict the dispersion coefficient.

8.3 Summary of contributions

Practical contribution

- a) A new approach was used in developing empirical mathematical models through consideration of data collection problem for model construction. In this case, solute transport parameters were obtained by several numerical methods over a range of flow rates and numerical properties. The approach was unique in that it involved several numerical methods and numerical properties. Furthermore, the study attempted to link numerical methods with empirical equations by considering the impact of model and numerical properties on developed empirical models.
- b) Since numerical methods would be required in most practical applications, this research was practically significant as the impact of numerical properties and numerical methods on constructed empirical models would enable practising engineers to interpret results given by empirical models.

Theoretical contribution

- a) The extension of the existing theory on the influence of numerical methods of the AD-Model and numerical properties on optimised parameter values of solute transport parameters by numerical methods and the development of empirical equations.
- b) The development of different empirical models that differed from those proposed in the literature, both in approach and structure. The empirical models were developed based on a routing procedure, several numerical methods and numerical properties.

- c) The need for consideration of type numerical method and numerical properties for the application with predicted values of dispersion coefficients estimated by empirical models. Reliability of concentration predictions using estimated dispersion coefficients is enhanced using numerical methods and model resolutions under which empirical models were developed.
- d) Development of adequate and comparable empirical models based on several numerical methods and over a range of numerical properties (or model resolutions). Adequate and comparable empirical models were developed having different parameters, appropriate for different numerical methods and model resolutions.

8.4 Recommendations

The research has the potential for further investigation which could not be probed within the current confines and time limit. These are categorized as experimental and computational.

Experimental

- a) The solute transport parameters that were determined for the Murray stream did not include transient and hyporheic storage. Therefore, it would be worthwhile to include transient storage as natural streams have a transient effect in the form of dead zones. However, this would require validation with physical measurements.
- b) Most predictive models include friction term and aspect ratio without considering other several factors which affect dispersion in streams and rivers, such as flow rate, sinuosity, bedforms, etc.. It would be worthwhile to include all the factors in predictive models. However, this would require methods of quantifying these factors.
- c) The Murray stream which was investigated is a small stream in which the maximum flow rate recorded in the period of data collection was $0.436 \text{ m}^3/\text{s}$. There is evidence that values of parameters for the stream influenced the choice of appropriate numerical methods for analysis of measured BTCs. Therefore, a similar study could be conducted on a big river to determine whether other numerical methods would be appropriate for parameter estimation.

Computational

- d) The modelling of observed BTCs was based on the assertion that the AD-Model through its numerical methods correctly models transport and dispersion characteristics of the stream. While all the efforts were made to eliminate inefficiencies in modelling to eliminate artificial dispersion the AD-Model has both physical and mathematical limitations. Further research may look at alternative models such as the transient storage model.
- e) The inclusion of transient storage and other channel characteristics in the development of an empirical model would result in a more complex model. Therefore, a more accurate approach would be required for analysis of important parameters using methods such as the standard regression coefficients approach (Saltelli et al., 2008).
- f) In this research numerical parameter estimation employed the steepest descent method. However, there are other several numerical parameter estimation methods, such as the Levenberg-Marquardt and Gauss-Newton methods (van den Bos, 2007). Therefore, future research could investigate the implications of using alternative tools for parameter estimation.
- g) Eulerian numerical methods have been observed to be inaccurate in advective-dominated mass transport problems because of artificial dispersion. In such cases, semi-Lagrangian methods have been observed to be more reliable (Sun and Sun, 2015). Also, mass transport problems characterised by sharp fronts would preferably require the use of total variation diminishing (TVD) schemes (Versteeg and Malalasekera, 2007). Therefore, further research would look at using the semi-Lagrangian and TVD schemes to develop simplified models appropriate for such cases.

BIBLIOGRAPHY

- Abbott, M.B. and Basco, D.R. 1989. *Computational fluid dynamics: an introduction for engineers*. Longman Scientific & Technical; Wiley.
- Analyse-it Ltd®. 2009. *Analyse-it statistical analysis and data visualisation software for Microsoft Excel*, Analyse-it, Ltd.
- Ang, A.H-S. and Tang, W.H. 2007. *Probability concepts in engineering: emphasis on application in civil engineering and environmental engineering*. Wiley, New York.
- Ani E-C, Wallis, S.G., Kraslawski, A. and Agachi, S. 2009. Development, calibration and evaluation of two mathematical models for pollutant transport in a small river. *Environmental Modelling and Software*. Elsevier Ltd,
- Arhonditsis, G.B. and Brett, M.T. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine ecology progress series*, 271:13-26
- Barnett, A.G. 1983. Exact and approximate solutions of the advection-dispersion equation. *Proc 20th IAHR Congress*, Moscow 3:180-190.
- Bencala, K.E. and Walters, R.A. 1983. Simulation of solute transport in a mountain pool-and-riffle stream: a transient storage model. *Water Resources Research*, 19:718-724
- Bennett, N.D., Croke, B.F., W.C., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D. and Andreassian, V. 2013. Characterising performance of environmental models. *Environmental Modelling and Software*, 40:1–20.
- Billo, E.J. 2007. *Excel for scientists and engineers: numerical methods*. Wiley, Hoboken, NJ, USA.
- Blasone, R., Madsen, H. and Rosbjerg, D. 2007. Parameter estimation in distributed hydrological modelling: comparison of global and local optimisation techniques. *Nordic Hydrology*, 38(4-5):451-476.
- Boxall, J.B. and Guymer, I. 2007. Longitudinal mixing in meandering channels: new experimental data set and verification of a predictive technique. *Water Research*, 41:341-354.
- Burke, N.A. 2002. *Travel time and flow characteristics of a small stream system*. PhD thesis, Heriot-Watt University, Edinburgh.
- Chanson, H. 2004. *Environmental hydraulics for open channel flows*. Oxford, Elsevier.
- Chapra, S.C. 2008. *Surface water-quality modeling*. Long Grove, Waveland Press.
- Chapra, S.C. and Canale, R.P. 2008. *Numerical methods for engineers*. New York, McGraw-Hill Education.
- Chin, D.A. 2013. *Water-quality engineering in natural systems: fate and transport processes in the water environment*. Hoboken,Wiley.
- Dattner, I. 2015. A model-based initial guess for estimating parameters in Systems of ordinary differential equations. *Biometrics* 71:1176-1184.
- Deng, Z.Q. and Jung, H.S. 2009. Scaling dispersion model for pollutant transport in rivers. *Environmental Modelling and Software* 24:627–631.

- Doherty, J. 2008. *PEST: Model-Independent Parameter Estimation User's manual*. Brisbane, Watermark Numerical Computing.
- Ferziger, J.H. and Peric, M. 2002. *Computational methods for fluid dynamics*. Berlin, Springer.
- Fischer, H.B., List, E.J. Koh, R.C.Y., Imberger, J. and Brooks, N.H. 1979. *Mixing in inland and coastal waters*. San Diego, Academic Press.
- Fischer, H.B. 1967. The mechanics of dispersion in natural streams. *American Society of Civil Engineers* 93:187-216.
- Fürst, J. and Furfánek, P. 2011. An implicit MacCormack scheme for unsteady flow calculations. *Computers and Fluids* 46:231–236.
- Fylstra, D., Lasdon, L., Watson, J. and Allan, W. 1998. Design and Use of the Microsoft Excel Solver. *Interfaces*, 28:29-55.
- Gallagher, M. and Doherty, J. 2007. Parameter estimation and uncertainty analysis for a watershed model. *Environmental Modelling & Software*, 22:1000-1020.
- Gogtay, N.J., Deshpande, S.P. and Thatte, U.M. 2016. Normal distribution “p” value and confidence intervals. *Journal of Association of Physicians of India*, 64:74–76.
- Hanusz, Z. and Tarasinska, J. 2016. Shapiro – Wilk test with known mean. *REVSTAT - Statistical Journal* 14(1):89–100.
- Harmel, R.D. and Smith, P.K. 2007. Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *Journal of Hydrology*, 337:326–336.
- Hayase, T., Humphrey, J.A.C. and Greif, R. 1992. A consistently formulated QUICK Scheme for fast and stable Convergence using finite-volume iterative calculation procedures. *J Comp Phys* 98:108-118.
- Heron, A.J. 2015. *Pollutant transport in rivers: estimating dispersion coefficients from tracer experiments*. MPhil Thesis, Heriot-Watt University, Edinburgh, UK.
- Hoffman, J.D. 2001. *Numerical Methods for Engineers and Scientists*. New York, Marcel Dekker.
- Jobson, H.E. 1997. Predicting travel time and dispersion in rivers and streams. *Journal of Hydraulic Engineering*, American Society of Civil Engineers, 123:971–978.
- Karahan, H. 2006. Implicit finite difference techniques for the advection-diffusion equation using spreadsheets. *Advances in Engineering Software* 37:601–608.
- Karahan, H. 2007. Unconditional stable explicit finite difference technique for the advection-diffusion equation using spreadsheets. *Advances in Engineering Software*, 38:80–86.
- Karahan, H. 2008. Solution of weighted finite difference techniques with the advection-diffusion equation using spreadsheets. *Computer Applications in Engineering Education*, 16:147–156.
- Kashefipour, S.M. and Falconer, R.A. 2002. Longitudinal dispersion coefficients in natural channels. *Water Research* 36:1596-1608.
- Kilpatrick, F.A., Wilson, J.F. and Hubbard, E.F. 1989. Measurement of time of travel in streams by dye tracing. *Techniques of water resources investigations of the United States Geological Survey*.

TWI3-A9. USGS.

Kohavi, R. 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence (IJCAI)*.

Kowalik, Z. and Murty, T.S. 1993. *Numerical Modeling of Ocean Dynamics*. Singapore, World Scientific.

Krause, P., Boyle, D.P. and Båse, F. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5:89-97.

Kumar, A., Jaiswal, D.K. and Kumar, N. 2009. Analytical solutions of one-dimensional advection-diffusion equation with variable coefficients. *J. Earth Syst.* 5:539–549.

Leonard, B.P. 1979. A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Computer methods in applied mechanics and engineering* 19:59–98.

Leonard, B.P. 1993. Comments on the policy statement on numerical accuracy. *Journal of Fluids Engineering*, 115:339-340.

Liu, H. 1977. Predicting dispersion coefficient of streams. *Journal of the Environmental Engineering Division*, American Society of Civil Engineers, 103:59-69.

MacCormack, R.W. 1982. A numerical method for solving the equations of compressible viscous flow. *AIAA Journal*, 20:1275–1281.

MacCormack, R.W. 1969. The effect of viscosity in hypervelocity impact cratering. *AIAA Journal*, 69-354.

Manson, J.R., Wallis, S.G. and Hope, D. 2001. A conservative semi-Lagrangian transport model for rivers with transient storage zones. *Water Resources Research*, 37:3321–3329.

Martin, J.E., Carr, M.L. and García, M.H. 2013. Riverine Transport, Mixing, and Dispersion. *Handbook of Environmental Fluid Dynamics*, CRC Press (2).

Martin, J.L. and McCutcheon, S.C. 1998. *Hydrodynamics and transport for water quality modeling*. CRC Press.

Matott, L.S. and Rabideau, A.J. 2008. Calibration of subsurface batch and reactive-transport models involving complex biogeochemical processes. *Advances in Water Resources*, 31(2):269-286.

McCuen, R.H. 1998. *Hydrologic design and analysis*. New Jersey, Prince Hall.

McCuen, R.H., Knight, Z. and Cutter, A.G. 2006. Evaluation of the Nash–Sutcliffe Efficiency Index. *Journal of Hydrologic Engineering* 11:597–602.

Montgomery, D.C. and Runger, G.C. 2003. *Applied statistics and Probability for Engineers*. Phoenix, Wiley,

Moriasi, D., Gitau, M.W., Pai, N. and Daggupati, P. (2015) Hydrologic and water quality models: performance measures and evaluation criteria. *American Society of Agricultural and Biological Engineers*, 58(6):1763-1785.

Nash, J.E. and Sutcliffe, J.V. 1970. River flow forecasting through conceptual Models Part1- A discussion of principles. *Journal of Hydrology* 10:282–290.

- Novak, P., Guinot, V., Jeffrey, A. and Reeve D.E. 2010. *Hydraulic Modelling - an introduction: principles, methods and applications*. New York, Spon Press.
- Ogata, A. and Banks, R.B. 1961. A solution of the differential equation of longitudinal dispersion in porous media. *U. S. Geological Survey Professional Paper 411-A*, Geological Survey, Washington.
- Rao, S.S. 2009. *Engineering Optimization: Theory and Practice*. Hoboken, New Jersey, Wiley.
- Razali, N.M. and Wah, Y.B. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2:21–33.
- Roache, P.J. 1997. Quantification of uncertainty in computational fluid dynamics. *Annual Review of Fluid Mechanics*, 29:123–160.
- Roache, P.J., Ghia, K.N. and White, F.M. 1986. Editorial policy statement on the control of numerical accuracy. *Journal of Fluids Engineering*.
- Runkel, R.L. 1998. One-dimensional Transport with Inflow and Storage: A solute transport model for streams and rivers. *U.S. Geological Survey, Water-Resources Investigation Report*, 98 – 4018.
- Rutherford, J.C. 1994. *River mixing*. Chichester, England, Wiley.
- Sáez, P. and Rittmann, B. 1992. Model-parameter estimation using least squares. *Water Research* 26:789–796.
- Saltelli, A., Chan, K., Scott, E.M. 2008. *Sensitivity analysis*. Chichester, John Wiley.
- Semuwemba, J. 2011. *Modelling Tracer Breakthrough Curves to Determine Stream Reaeration and Hydrodynamic Properties*. PhD Thesis, Queen's University, Belfast.
- Seo, I. and Cheong, T. 1998. Predicting longitudinal dispersion coefficient in natural streams. *Journal of Hydraulic Engineering*, 124:25-32.
- Shapiro, A.S.S. and Wilk, M.B. 1965. An Analysis of Variance Test for Normality. *Biometrika* 52(3): 591–611.
- Shucksmith, J., Boxall, J. and Guymer, I. 2007 Importance of advective zone in longitudinal mixing experiments. *Acta Geophysica*, 55(1):95-103.
- Singh, J., Knapp, H.V., Arnold, J.G. and Demissie, M. 2005. Hydrological modeling of the Iroquois River watershed using HSPF and SWAT. *Journal of the American Water Resources Association*, 41(2): 43–360.,
- Singh, S.K. and Beck, M.B. 2003. Dispersion Coefficient of Streams from Tracer Experiment Data. *Journal of Environmental Engineering*, 129(6):539–546.
- Skahill, B.E., Doherty, J. 2006. Efficient accommodation of local minima in watershed model calibration. *Journal of Hydrology*, 329(1-2):122-139.
- Sobey, R.J. 1984. Numerical alternatives in transient stream response. *Journal of Hydraulic Engineering*, American Society of Civil Engineers, 110(6):749–772.
- Sun, N.Z. and Sun, A. 2015. *Model Calibration and Parameter Estimation: for environment and water resource systems*. Springer, 53(9):1689–1699.

- Szymkiewicz, R. 2010. *Numerical modeling in open channels*. Dordrecht, Springer.
- Taylor, G. 1954. The dispersion of matter in turbulent flow through a pipe: *Mathematical, Physical and Engineering Sciences*, The Royal Society of London 446–468.
- Toprak, Z.F. and Cigizoglu, H.K. 2008 Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrological Processes* 22:4106-4129.
- Toprak, Z.F., Şen, Z. and Savci, M.E. 2004. Comment on Longitudinal dispersion coefficients in natural channels. *Water Research* 38:3139–3143.
- Vaghela, C.R. and Vaghela, A.R. 2014. Synthetic Flow Generation. *International journal of engineering research and application*, 4:66–71.
- Van den Bos, A. 2007. *Parameter Estimation for Scientists and Engineers*. Hoboken, New Jersey Wiley.
- Van Leer, B. 1974. Towards the ultimate conservative difference scheme: monotonicity and conservation combined in a second-order scheme. *Journal of Computational Physics*, 14:361–370.
- Versteeg, H.K. and Malalasekera, W. 2007. *An introduction to computational fluid dynamics: the finite volume method*. Harlow, Pearson Education.
- Wallis, S.G. 2005. *Experimental study of travel times in a small stream*. In: Czernuszenko W, Rowiński P (eds) *Water Quality Hazards and Dispersion of Pollutants*. New York, Springer.
- Wallis, S.G. 2007. The numerical solution of the advection-dispersion Equation : A review of some basic principles. *Acta Geophysica*, 55(1):85-94.
- Wallis, S.G. Manson, J.R. and Filippi, L. 1998. A conservative semi-Lagrangian algorithm for one-dimensional advection-diffusion. *Communications in Numerical Methods in Engineering* 14:671-679.
- Wallis, S.G., Bonardi, D. and Silavwe, D.D. 2013. Solute transport routing in a small stream. *Hydrological Sciences Journal*, 59:1894-1907.
- Wallis, S.G. and Manson, J.R. 1997. Accurate numerical simulation of advection using large time steps. *Numerical methods in fluids*, 24(2):127–139.
- Wallis, S.G. and Manson, J.R. 2004. Methods for predicting dispersion coefficients in rivers. *Proceedings of the ICE - Water Management*, 157(3):131–141.
- Wallis, S.G., Osuch, M., Manson, J.R., Romanowicz, R.J. and Demars, B.O.L 2013. On the Estimation of Solute Transport Parameters for Rivers. In: Rowiński P (ed.) *Experimental and Computational Solutions of Hydraulic Problems*. *GeoPlanet: Earth and Planetary Sciences*, 11:415–425.
- Whitehead, P.G., Williams, R.J. and Hornberger, G.M. 1986. On the identification of pollutant or tracer sources using dispersion theory. *Journal of Hydrology* 84:273-286.

APPENDICES

Appendix A: Analysed concentration-time data and charts; BTCs

Appendix A1: Concentration-time data for experiments 2 to 13.

Experiment 2					
Time (min)	Upstream	Downstream	Time (min)	Upstream	Downstream
0	0.000		97	0.000	0.091
1	0.000		98	0.000	0.086
2	0.000		99	0.000	0.079
3	0.000		100		0.076
4	0.003		101		0.071
5	0.011		102		0.067
6	0.036		103		0.061
7	0.102		104		0.061
8	0.167		105		0.059
9	0.313		106		0.054
10	0.481		107		0.052
11	0.690		108		0.047
12	0.899		109		0.044
13	1.078		110		0.045
14	1.272		111		0.042
15	1.376		112		0.042
16	1.484		113		0.039
17	1.531		114		0.035
18	1.484		115		0.034
19	1.389		116		0.034
20	1.342	0.000	117		0.034
21	1.257	0.000	118		0.030
22	1.138	0.000	119		0.029
23	1.033	0.000			
24	0.899	0.000			
25	0.779	0.000			
26	0.690	0.000			
27	0.600	0.000			
28	0.511	0.000			
29	0.436	0.000			
30	0.384	0.000			
31	0.337	0.000			
32	0.294	0.000			
33	0.242	0.000			
34	0.214	0.000			
35	0.186	0.000			
36	0.167	0.004			
37	0.148	0.004			
38	0.129	0.005			
39	0.115	0.009			
40	0.100	0.015			
41	0.097	0.020			
42	0.084	0.037			
43	0.075	0.052			
44	0.070	0.071			
45	0.063	0.106			
46	0.057	0.134			
47	0.054	0.193			
48	0.054	0.241			
49	0.048	0.299			
50	0.048	0.362			
51	0.040	0.426			
52	0.040	0.490			
53	0.040	0.540			
54	0.035	0.607			
55	0.033	0.674			
56	0.030	0.725			
57	0.032	0.775			
58	0.029	0.809			
59	0.030	0.809			
60	0.030	0.842			
61	0.024	0.859			
62	0.024	0.876			
63	0.024	0.842			
64	0.023	0.842			
65	0.023	0.809			
66	0.021	0.792			
67	0.021	0.775			
68	0.020	0.725			
69	0.018	0.691			
70	0.018	0.658			
71	0.020	0.624			
72	0.017	0.591			
73	0.020	0.540			
74	0.015	0.507			
75	0.015	0.473			
76	0.015	0.440			
77	0.015	0.426			
78	0.014	0.373			
79	0.014	0.357			
80	0.012	0.309			
81	0.012	0.294			
82	0.012	0.283			
83	0.014	0.256			
84	0.011	0.235			
85	0.012	0.225			
86	0.011	0.198			
87	0.012	0.182			
88	0.009	0.172			
89	0.009	0.156			
90	0.009	0.145			
91	0.000	0.134			
92	0.000	0.124			
93	0.000	0.114			
94	0.000	0.108			
95	0.000	0.101			
96	0.000	0.092			

Experiment 3		
Time (min)	Upstream	Downstream
0	0.000	
1	0.000	
2	0.000	
3	0.152	
4	0.305	
5	2.367	
6	2.675	
7	2.110	
8	1.444	
9	0.879	
10	0.572	
11	0.305	
12	0.187	
13	0.110	
14	0.079	
15	0.045	0.094
16	0.034	0.279
17	0.029	0.672
18	0.023	1.022
19	0.018	1.324
20	0.015	1.415
21	0.013	1.365
22	0.011	1.164
23	0.010	0.975
24	0.008	0.784
25	0.008	0.593
26	0.005	0.434
27	0.008	0.299
28	0.008	0.239
29	0.005	0.173
30	0.005	0.118
31	0.003	0.091
32	0.005	0.067
33	0.003	0.052
34	0.005	0.041
35	0.003	0.033
36	0.002	0.027
37	0.005	0.025
38	0.003	0.021
39	0.002	0.021
40	0.002	0.016
41	0.003	0.014
42	0.002	0.011
43		0.008
44		0.008
45		0.008
46		0.008
47		0.008
48		0.005
49		0.005
50		0.008
51		0.006
52		0.006
53		0.005
54		0.005
55		0.005
56		0.005
57		0.005
58		0.003
59		0.002

Experiment 4						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0.000	0.000			48.500	0.000	0.025
0.500	0.000			49.000	0.000	0.023
1.000	0.000			49.500	0.000	0.022
1.500	0.000			50.000		0.021
2.000	0.015			50.500		0.021
2.500	0.059			51.000		0.021
3.000	0.152			51.500		0.021
3.500	0.342			52.000		0.020
4.000	0.600			52.500		0.017
4.500	0.838			53.000		0.015
5.000	1.123			53.500		0.015
5.500	1.335			54.000		0.015
6.000	1.486			54.500		0.014
6.500	1.486			55.000		0.013
7.000	1.486			55.500		0.012
7.500	1.436			56.000		0.012
8.000	1.335			56.500		0.012
8.500	1.187			57.000		0.012
9.000	1.044			57.500		0.011
9.500	0.917			58.000		0.010
10.000	0.758	0.002		58.500		0.009
10.500	0.663	0.001		59.000		0.008
11.000	0.552	0.000		59.500		0.009
11.500	0.457	0.001		60.000		0.010
12.000	0.383	0.002		60.500		0.009
12.500	0.317	0.001		61.000		0.008
13.000	0.267	0.000		61.500		0.008
13.500	0.222	0.000		62.000		0.008
14.000	0.182	0.000		62.500		0.009
14.500	0.152	0.000		63.000		0.010
15.000	0.127	0.000		63.500		0.008
15.500	0.110	0.001		64.000		0.007
16.000	0.091	0.002		64.500		0.007
16.500	0.080	0.004		65.000		0.007
17.000	0.067	0.005		65.500		0.007
17.500	0.059	0.008		66.000		0.007
18.000	0.050	0.012		66.500		0.007
18.500	0.043	0.027		67.000		0.007
19.000	0.038	0.042		67.500		0.008
19.500	0.034	0.076		68.000		0.008
20.000	0.030	0.110		68.500		0.007
20.500	0.030	0.158		69.000		0.005
21.000	0.026	0.205		69.500		0.005
21.500	0.026	0.272		70.000		0.005
22.000	0.023	0.338		70.500		0.005
22.500	0.021	0.401		71.000		0.005
23.000	0.019	0.465		71.500		0.004
23.500	0.018	0.545		72.000		0.004
24.000	0.016	0.626		72.500		0.005
24.500	0.013	0.666		73.000		0.007
25.000	0.015	0.707		73.500		0.006
25.500	0.013	0.739		74.000		0.005
26.000	0.011	0.771		74.500		0.005
26.500	0.015	0.795		75.000		0.005
27.000	0.011	0.819		75.500		0.005
27.500	0.010	0.795		76.000		0.005
28.000	0.010	0.771		76.500		0.004
28.500	0.010	0.755		77.000		0.004
29.000	0.011	0.739		77.500		0.004
29.500	0.010	0.707		78.000		0.005
30.000	0.008	0.674		78.500		0.004
30.500	0.008	0.634		79.000		0.004
31.000	0.008	0.594		79.500		0.004
31.500	0.008	0.545		80.000		0.004
32.000	0.007	0.497		80.500		0.004
32.500	0.008	0.457		81.000		0.004
33.000	0.008	0.416		81.500		0.004
33.500	0.007	0.382		82.000		0.004
34.000	0.008	0.348		82.500		0.004
34.500	0.007	0.328		83.000		0.004
35.000	0.005	0.307		83.500		0.003
35.500	0.005	0.272		84.000		0.002
36.000	0.005	0.236		84.500		0.003
36.500	0.004	0.218		85.000		0.004
37.000	0.004	0.200		85.500		0.004
37.500	0.007	0.177		86.000		0.004
38.000	0.002	0.154		86.500		0.004
38.500	0.004	0.137		87.000		0.004
39.000	0.005	0.119		87.500		0.003
39.500	0.005	0.114		88.000		0.002
40.000	0.005	0.108		88.500		0.003
40.500	0.004	0.099		89.000		0.004
41.000	0.004	0.089		89.500		0.004
41.500	0.004	0.080		90.000		0.004
42.000	0.004	0.071		90.500		0.003
42.500	0.002	0.064		91.000		0.002
43.000	0.005	0.057		91.500		0.002
43.500	0.004	0.053		92.000		0.002
44.000	0.004	0.049				
44.500	0.004	0.046				
45.000	0.005	0.044				
45.500	0.005	0.038				
46.000	0.004	0.033				
46.500	0.002	0.030				
47.000	0.002	0.028				
47.500	0.000	0.027				
48.000	0.000	0.026				

Experiment 5					
Time (min)	Upstream	Downstream	Time (min)	Upstream	Downstream
0.00	0.000		50.50		0.040
0.50	0.000		51.00		0.038
1.00	0.000		51.50		0.035
1.50	0.000		52.00		0.033
2.00	0.000		52.50		0.031
2.50	0.000		53.00		0.030
3.00	0.000		53.50		0.028
3.50	0.000		54.00		0.027
4.00	0.000		54.50		0.026
4.50	0.000		55.00		0.025
5.00	0.002		55.50		0.025
5.50	0.005		56.00		0.025
6.00	0.002		56.50		0.024
6.50	0.008		57.00		0.024
7.00	0.030		57.50		0.022
7.50	0.119		58.00		0.020
8.00	0.354		58.50		0.021
8.50	0.734		59.00		0.022
9.00	1.248		59.50		0.020
9.50	1.861		60.00		0.019
10.00	2.339		60.50		0.018
10.50	2.722		61.00		0.017
11.00	2.817	0.000	61.50		0.017
11.50	2.817	0.000	62.00		0.016
12.00	2.722	0.000	62.50		0.016
12.50	2.435	0.000	63.00		0.016
13.00	2.148	0.000	63.50		0.014
13.50	1.861	0.000	64.00		0.013
14.00	1.573	0.000	64.50		0.014
14.50	1.334	0.000	65.00		0.016
15.00	1.142	0.000	65.50		0.015
15.50	0.915	0.000	66.00		0.014
16.00	0.794	0.000	66.50		0.013
16.50	0.643	0.002	67.00		0.011
17.00	0.522	0.003	67.50		0.011
17.50	0.431	0.003	68.00		0.011
18.00	0.354	0.003	68.50		0.012
18.50	0.301	0.003	69.00		0.013
19.00	0.244	0.003	69.50		0.012
19.50	0.205	0.003	70.00		0.011
20.00	0.177	0.003	70.50		0.011
20.50	0.148	0.006	71.00		0.011
21.00	0.129	0.009	71.50		0.011
21.50	0.110	0.024	72.00		0.011
22.00	0.095	0.039	72.50		0.010
22.50	0.080	0.089	73.00		0.009
23.00	0.071	0.139	73.50		0.009
23.50	0.062	0.219	74.00		0.009
24.00	0.056	0.298	74.50		0.010
24.50	0.053	0.412	75.00		0.011
25.00	0.045	0.527	75.50		0.009
25.50	0.041	0.692	76.00		0.008
26.00	0.038	0.857	76.50		0.008
26.50	0.033	1.014	77.00		0.008
27.00	0.032	1.172	77.50		0.009
27.50	0.033	1.255	78.00		0.009
28.00	0.030	1.337	78.50		0.009
28.50	0.027	1.387	79.00		0.009
29.00	0.027	1.437	79.50		0.009
29.50	0.026	1.462	80.00		0.009
30.00	0.023	1.487	80.50		0.009
30.50	0.023	1.462	81.00		0.008
31.00	0.021	1.437	81.50		0.009
31.50	0.021	1.362	82.00		0.011
32.00	0.020	1.288	82.50		0.009
32.50	0.018	1.230	83.00		0.006
33.00	0.017	1.172	83.50		0.006
33.50	0.018	1.077	84.00		0.006
34.00	0.017	0.983	84.50		0.007
34.50	0.015	0.920	85.00		0.008
35.00	0.015	0.857	85.50		0.007
35.50	0.017	0.794	86.00		0.006
36.00	0.015	0.731	86.50		0.006
36.50	0.014	0.653	87.00		0.006
37.00	0.015	0.574	87.50		0.007
37.50	0.014	0.535	88.00		0.008
38.00	0.014	0.495	88.50		0.007
38.50	0.011	0.432	89.00		0.006
39.00	0.012	0.368	89.50		0.006
39.50	0.011	0.320	90.00		0.006
40.00	0.008	0.273	90.50		0.007
40.50	0.011	0.253	91.00		0.008
41.00	0.012	0.233	91.50		0.007
41.50	0.008	0.226	92.00		0.006
42.00	0.011	0.219	92.50		0.006
42.50	0.011	0.189	93.00		0.006
43.00	0.009	0.159	93.50		0.007
43.50	0.009	0.139	94.00		0.008
44.00	0.009	0.119	94.50		0.006
44.50	0.009	0.109	95.00		0.005
45.00	0.011	0.099	95.50		0.006
45.50	0.009	0.092	96.00		0.006
46.00	0.006	0.085	96.50		0.006
46.50	0.008	0.075	97.00		0.005
47.00	0.008	0.065	97.50		0.003
47.50	0.005	0.063	98.00		0.002
48.00	0.008	0.061			
48.50	0.008	0.058			
49.00	0.006	0.055			
49.50	0.006	0.049			
50.00		0.043			

Experiment 6						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0.00	0.000			32.33		0.007
0.33	0.000			32.67		0.007
0.67	0.000			33.00		0.004
1.00	0.000			33.33		0.006
1.33	0.000			33.67		0.006
1.67	0.000			34.00		0.004
2.00	0.000			34.33		0.005
2.33	0.000			34.67		0.005
2.67	0.000			35.00		0.003
3.00	0.000			35.33		0.003
3.33	0.000			35.67		0.003
3.67	0.000			36.00		0.003
4.00	0.000			36.33		0.005
4.33	0.009			36.67		0.005
4.67	0.074			37.00		0.004
5.00	0.315			37.33		0.003
5.33	0.519			37.67		0.004
5.67	0.723			38.00		0.006
6.00	1.324			38.33		0.006
6.33	1.863			38.67		0.005
6.67	2.234			39.00		0.003
7.00	2.512	0.000		39.33		0.006
7.33	2.512	0.000		39.67		0.007
7.67	2.373	0.000		40.00		0.006
8.00	2.095	0.000		40.33		0.006
8.33	1.817	0.000		40.67		0.006
8.67	1.493	0.001		41.00		0.006
9.00	1.250	0.006		41.33		0.004
9.33	1.045	0.003		41.67		0.003
9.67	0.811	0.001		42.00		0.003
10.00	0.649	0.001		42.33		0.004
10.33	0.503	0.001		42.67		0.003
10.67	0.385	0.002		43.00		0.001
11.00	0.310	0.003				
11.33	0.250	0.004				
11.67	0.180	0.009				
12.00	0.134	0.019				
12.33	0.116	0.063				
12.67	0.097	0.133				
13.00	0.074	0.230				
13.33	0.055	0.407				
13.67	0.053	0.619				
14.00	0.044	0.867				
14.33	0.035	1.085				
14.67	0.028	1.270				
15.00	0.025	1.420				
15.33	0.019	1.577				
15.67	0.020	1.671				
16.00	0.017	1.703				
16.33	0.016	1.671				
16.67	0.013	1.608				
17.00	0.010	1.514				
17.33	0.010	1.326				
17.67	0.010	1.180				
18.00	0.009	1.075				
18.33	0.010	0.956				
18.67	0.010	0.842				
19.00	0.006	0.733				
19.33	0.007	0.624				
19.67	0.007	0.515				
20.00	0.006	0.406				
20.33	0.004	0.358				
20.67	0.004	0.305				
21.00	0.004	0.249				
21.33	0.004	0.211				
21.67	0.003	0.178				
22.00	0.003	0.150				
22.33	0.006	0.125				
22.67	0.003	0.105				
23.00	0.001	0.089				
23.33	0.001	0.073				
23.67		0.062				
24.00		0.055				
24.33		0.048				
24.67		0.041				
25.00		0.035				
25.33		0.032				
25.67		0.029				
26.00		0.027				
26.33		0.023				
26.67		0.020				
27.00		0.019				
27.33		0.019				
27.67		0.018				
28.00		0.016				
28.33		0.015				
28.67		0.014				
29.00		0.012				
29.33		0.013				
29.67		0.012				
30.00		0.009				
30.33		0.009				
30.67		0.009				
31.00		0.010				
31.33		0.011				
31.67		0.009				
32.00		0.004				

Experiment 7						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0	0.000			97	0.000	0.025
1	0.000			98	0.000	0.023
2	0.000			99	0.000	0.021
3	0.000			100		0.021
4	0.000			101		0.020
5	0.000			102		0.016
6	0.000			103		0.018
7	0.000			104		0.016
8	0.000			105		0.014
9	0.050			106		0.014
10	0.260			107		0.013
11	0.855			108		0.013
12	1.717			109		0.014
13	2.992			110		0.013
14	3.960			111		0.013
15	4.572			112		0.011
16	4.623			113		0.011
17	4.317			114		0.013
18	3.706			115		0.011
19	3.043			116		0.009
20	2.380			117		0.009
21	1.819			118		0.009
22	1.361	0.000		119		0.009
23	1.032	0.000		120		0.009
24	0.742	0.000		121		0.009
25	0.581	0.000		122		0.007
26	0.419	0.000		123		0.007
27	0.341	0.000		124		0.007
28	0.260	0.000		125		0.007
29	0.209	0.000		126		0.007
30	0.163	0.000		127		0.006
31	0.137	0.000		128		0.004
32	0.112	0.000		129		0.006
33	0.097	0.004		130		0.004
34	0.079	0.011		131		0.004
35	0.071	0.026		132		0.004
36	0.063	0.054		133		0.006
37	0.055	0.124		134		0.006
38	0.050	0.214		135		0.004
39	0.046	0.401		136		0.006
40	0.041	0.627		137		0.002
41	0.037	0.835		138		0.004
42	0.034	1.079		139		0.002
43	0.031	1.340		140		0.004
44	0.028	1.670		141		0.002
45	0.026	1.908				
46	0.025	2.018				
47	0.021	2.183				
48	0.023	2.128				
49	0.020	2.183				
50	0.016	2.128				
51	0.016	2.073				
52	0.016	1.963				
53	0.015	1.798				
54	0.015	1.633				
55	0.012	1.523				
56	0.012	1.409				
57	0.012	1.253				
58	0.012	1.114				
59	0.010	1.009				
60	0.010	0.888				
61	0.010	0.783				
62	0.010	0.679				
63	0.007	0.574				
64	0.008	0.505				
65	0.007	0.470				
66	0.005	0.396				
67	0.005	0.357				
68	0.005	0.319				
69	0.007	0.269				
70	0.005	0.247				
71	0.004	0.214				
72	0.005	0.192				
73	0.005	0.165				
74	0.004	0.148				
75	0.004	0.132				
76	0.004	0.117				
77	0.004	0.110				
78	0.002	0.096				
79	0.004	0.086				
80	0.002	0.080				
81	0.002	0.072				
82	0.000	0.063				
83	0.000	0.056				
84	0.000	0.053				
85	0.000	0.051				
86	0.000	0.046				
87	0.000	0.044				
88	0.000	0.039				
89	0.000	0.035				
90	0.000	0.033				
91	0.000	0.033				
92	0.000	0.030				
93	0.000	0.028				
94	0.000	0.028				
95	0.000	0.025				
96	0.000	0.026				

Experiment 8		
Time (min)	Upstream	Downstream
0.0	0.000	
1.5	0.002	
3.0	0.003	
4.5	0.006	
6.0	0.003	
7.5	0.011	
9.0	0.179	
10.5	0.978	
12.0	2.316	
13.5	3.350	
15.0	3.491	
16.5	2.692	
18.0	1.894	
19.5	1.236	
21.0	0.785	
22.5	0.473	
24.0	0.296	
25.5	0.198	
27.0	0.132	
28.5	0.094	
30.0	0.072	0.000
31.5	0.055	0.001
33.0	0.051	0.023
34.5	0.042	0.082
36.0	0.033	0.263
37.5	0.029	0.522
39.0	0.026	0.919
40.5	0.024	1.200
42.0	0.018	1.399
43.5	0.020	1.634
45.0	0.018	1.634
46.5	0.018	1.581
48.0	0.015	1.515
49.5	0.014	1.349
51.0	0.012	1.134
52.5	0.012	0.969
54.0	0.011	0.820
55.5	0.011	0.671
57.0	0.011	0.522
58.5	0.009	0.423
60.0	0.009	0.357
61.5	0.008	0.279
63.0	0.006	0.222
64.5	0.008	0.180
66.0	0.005	0.148
67.5	0.005	0.122
69.0	0.006	0.092
70.5	0.006	0.086
72.0	0.003	0.066
73.5	0.005	0.058
75.0	0.005	0.049
76.5	0.006	0.043
78.0	0.002	0.038
79.5	0.005	0.033
81.0	0.005	0.029
82.5	0.006	0.026
84.0	0.003	0.023
85.5	0.003	0.020
87.0	0.003	0.020
88.5	0.003	0.018
90.0	0.002	0.016
91.5	0.002	0.015
93.0	0.002	0.015
94.5	0.003	0.011
96.0	0.002	0.010
97.5	0.002	0.011
99.0	0.002	0.008
100.5	0.003	0.008
102.0	0.002	0.008
103.5	0.003	0.006
105.0	0.002	0.006
106.5	0.002	0.008
108.0	0.002	0.006
109.5		0.006
111.0		0.005
112.5		0.006
114.0		0.005
115.5		0.005
117.0		0.005
118.5		0.005
120.0		0.005
121.5		0.005
123.0		0.005
124.5		0.003
126.0		0.005
127.5		0.003
129.0		0.001
130.5		0.003
132.0		0.003
133.5		0.003
135.0		0.001
136.5		0.001
138.0		0.001
139.5		0.001
141.0		0.001
142.5		0.001

Experiment 9						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0	0.035			97		0.012
1	0.147			98		0.009
2	0.493			99		0.009
3	1.092			100		0.009
4	1.803			101		0.009
5	2.418			102		0.009
6	2.675			103		0.009
7	2.675			104		0.009
8	2.418			105		0.007
9	2.111			106		0.007
10	1.649			107		0.007
11	1.291			108		0.007
12	0.995			109		0.007
13	0.736			110		0.009
14	0.541			111		0.007
15	0.395			112		0.007
16	0.296			113		0.007
17	0.229			114		0.007
18	0.173			115		0.006
19	0.132			116		0.005
20	0.105			117		0.005
21	0.082			118		0.005
22	0.073			119		0.005
23	0.057			120		0.005
24	0.048			121		0.005
25	0.042			122		0.005
26	0.037			123		0.004
27	0.034			124		0.003
28	0.029			125		0.004
29	0.026			126		0.005
30	0.023			127		0.004
31	0.021			128		0.003
32	0.016	0.188		129		0.003
33	0.016	0.277		130		0.003
34	0.016	0.361		131		0.003
35	0.014	0.482		132		0.003
36	0.013	0.641		133		0.003
37	0.011	0.835		134		0.003
38	0.008	0.870		135		0.003
39	0.010	1.063		136		0.003
40	0.011	1.081		137		0.002
41	0.010	1.152		138		0.002
42	0.008	1.152				
43	0.010	1.187				
44	0.008	1.169				
45	0.006	1.116				
46	0.006	1.046				
47	0.006	0.975				
48	0.005	0.940				
49	0.005	0.852				
50	0.005	0.764				
51	0.003	0.711				
52	0.005	0.623				
53	0.003	0.570				
54	0.005	0.535				
55	0.006	0.465				
56	0.003	0.422				
57	0.005	0.366				
58	0.003	0.333				
59	0.003	0.283				
60	0.003	0.255				
61	0.003	0.227				
62	0.003	0.199				
63	0.003	0.171				
64	0.003	0.149				
65	0.003	0.132				
66	0.003	0.127				
67	0.002	0.104				
68	0.001	0.097				
69	0.001	0.090				
70	0.001	0.083				
71	0.001	0.070				
72	0.001	0.067				
73		0.060				
74		0.054				
75		0.049				
76		0.046				
77		0.044				
78		0.037				
79		0.033				
80		0.032				
81		0.031				
82		0.028				
83		0.024				
84		0.024				
85		0.024				
86		0.021				
87		0.019				
88		0.019				
89		0.017				
90		0.017				
91		0.016				
92		0.014				
93		0.014				
94		0.014				
95		0.014				
96		0.012				

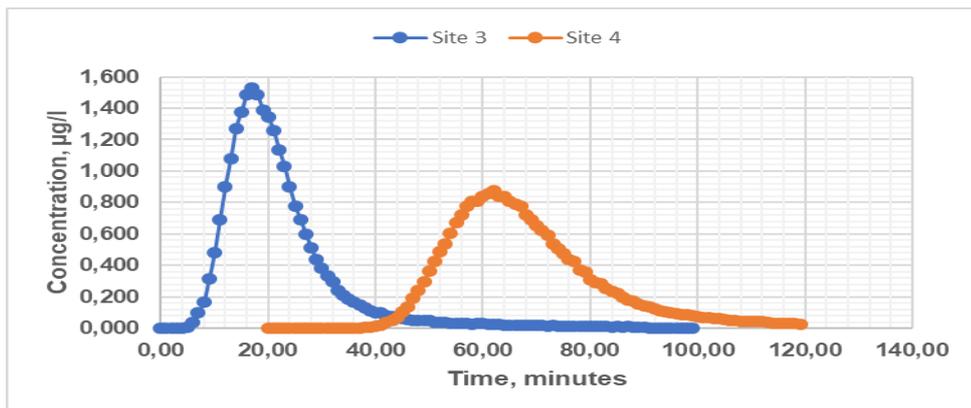
Experiment 10						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0	0.000			97	0.017	1.171
1	0.000			98	0.017	1.136
2	0.000			99	0.017	1.084
3	0.000			100	0.016	1.067
4	0.005			101	0.016	1.014
5	0.019			102	0.016	0.944
6	0.054			103	0.016	0.927
7	0.154			104	0.015	0.892
8	0.314			105	0.014	0.857
9	0.608			106	0.014	0.822
10	1.044			107	0.014	0.787
11	1.516			108	0.014	0.752
12	2.042			109	0.014	0.717
13	2.595			110	0.013	0.682
14	3.147			111	0.012	0.630
15	3.589			112	0.012	0.578
16	3.809			113	0.012	0.578
17	3.920			114	0.011	0.525
18	3.865			115	0.010	0.508
19	3.699			116	0.010	0.490
20	3.533			117	0.010	0.438
21	3.202			118	0.010	0.424
22	2.926			119	0.010	0.418
23	2.595			120	0.009	0.385
24	2.318			121	0.009	0.363
25	2.042			122	0.009	0.341
26	1.766			123	0.010	0.336
27	1.490			124	0.009	0.313
28	1.324			125	0.009	0.297
29	1.114			126	0.009	0.280
30	0.974			127	0.009	0.258
31	0.835			128	0.009	0.236
32	0.747			129	0.009	0.228
33	0.608			130	0.009	0.220
34	0.555			131	0.009	0.209
35	0.468			132	0.009	0.197
36	0.424			133	0.009	0.186
37	0.386			134	0.009	0.175
38	0.331			135	0.009	0.167
39	0.275			136	0.009	0.159
40	0.253			137	0.009	0.148
41	0.226			138		0.137
42	0.209			139		0.128
43	0.187			140		0.120
44	0.176	0.000		141		0.117
45	0.148	0.000		142		0.115
46	0.137	0.000		143		0.109
47	0.132	0.000		144		0.103
48	0.121	0.000		145		0.098
49	0.110	0.000		146		0.094
50	0.104	0.000		147		0.090
51	0.098	0.000		148		0.087
52	0.089	0.000		149		0.083
53	0.082	0.000		150		0.078
54	0.078	0.001		151		0.075
55	0.073	0.003		152		0.071
56	0.070	0.003		153		0.069
57	0.066	0.003		154		0.066
58	0.061	0.003		155		0.065
59	0.059	0.003		156		0.064
60	0.057	0.007		157		0.063
61	0.054	0.010		158		0.061
62	0.051	0.014		159		0.059
63	0.047	0.021		160		0.057
64	0.044	0.029		161		0.054
65	0.042	0.045		162		0.050
66	0.041	0.066				
67	0.040	0.090				
68	0.039	0.131				
69	0.038	0.148				
70	0.037	0.170				
71	0.035	0.247				
72	0.033	0.302				
73	0.031	0.380				
74	0.032	0.440				
75	0.033	0.490				
76	0.031	0.560				
77	0.030	0.647				
78	0.030	0.770				
79	0.030	0.805				
80	0.027	0.909				
81	0.024	0.962				
82	0.025	1.014				
83	0.026	1.119				
84	0.024	1.136				
85	0.023	1.189				
86	0.023	1.224				
87	0.023	1.224				
88	0.023	1.276				
89	0.023	1.276				
90	0.022	1.294				
91	0.021	1.276				
92	0.020	1.276				
93	0.019	1.276				
94	0.020	1.276				
95	0.021	1.241				
96	0.019	1.171				

Experiment 11						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0.0	0.000			48.5		0.013
0.5	0.000			49.0		0.014
1.0	0.000			49.5		0.014
1.5	0.000			50.0		0.011
2.0	0.000			50.5		0.011
2.5	0.000			51.0		0.010
3.0	0.000			51.5		0.010
3.5	0.000			52.0		0.010
4.0	0.000			52.5		0.008
4.5	0.011			53.0		0.007
5.0	0.081			53.5		0.007
5.5	0.336			54.0		0.007
6.0	0.845			54.5		0.007
6.5	1.568			55.0		0.008
7.0	2.241			55.5		0.008
7.5	2.723			56.0		0.008
8.0	2.963			56.5		0.007
8.5	2.963			57.0		0.007
9.0	2.819			57.5		0.006
9.5	2.530			58.0		0.005
10.0	2.290			58.5		0.004
10.5	1.953			59.0		0.003
11.0	1.664	0.000		59.5		0.004
11.5	1.375	0.000		60.0		0.005
12.0	1.195	0.000		60.5		0.004
12.5	0.906	0.000		61.0		0.003
13.0	0.799	0.000		61.5		0.003
13.5	0.632	0.000		62.0		0.002
14.0	0.540	0.000		62.5		0.002
14.5	0.434	0.000		63.0		0.002
15.0	0.360	0.000		63.5		0.001
15.5	0.288	0.000		64.0		0.000
16.0	0.240	0.002				
16.5	0.187	0.007				
17.0	0.163	0.018				
17.5	0.134	0.049				
18.0	0.110	0.114				
18.5	0.091	0.214				
19.0	0.076	0.349				
19.5	0.064	0.467				
20.0	0.057	0.673				
20.5	0.049	0.784				
21.0	0.044	1.037				
21.5	0.038	1.258				
22.0	0.037	1.353				
22.5	0.032	1.481				
23.0	0.028	1.581				
23.5	0.026	1.631				
24.0	0.023	1.681				
24.5	0.020	1.631				
25.0	0.020	1.631				
25.5	0.015	1.581				
26.0	0.017	1.481				
26.5	0.015	1.417				
27.0	0.012	1.353				
27.5	0.012	1.211				
28.0	0.014	1.132				
28.5	0.011	1.037				
29.0	0.011	0.926				
29.5	0.011	0.799				
30.0	0.011	0.736				
30.5	0.009	0.657				
31.0	0.009	0.594				
31.5	0.008	0.499				
32.0	0.008	0.435				
32.5	0.009	0.384				
33.0	0.008	0.349				
33.5	0.006	0.289				
34.0	0.008	0.249				
34.5	0.008	0.219				
35.0	0.006	0.209				
35.5	0.006	0.174				
36.0	0.005	0.154				
36.5	0.006	0.134				
37.0	0.005	0.109				
37.5	0.005	0.099				
38.0	0.005	0.089				
38.5	0.005	0.078				
39.0	0.005	0.078				
39.5	0.003	0.060				
40.0	0.003	0.059				
40.5	0.003	0.052				
41.0	0.005	0.048				
41.5	0.003	0.043				
42.0	0.003	0.041				
42.5	0.002	0.033				
43.0	0.002	0.033				
43.5	0.002	0.032				
44.0		0.024				
44.5		0.024				
45.0		0.022				
45.5		0.022				
46.0		0.021				
46.5		0.018				
47.0		0.018				
47.5		0.016				
48.0		0.016				

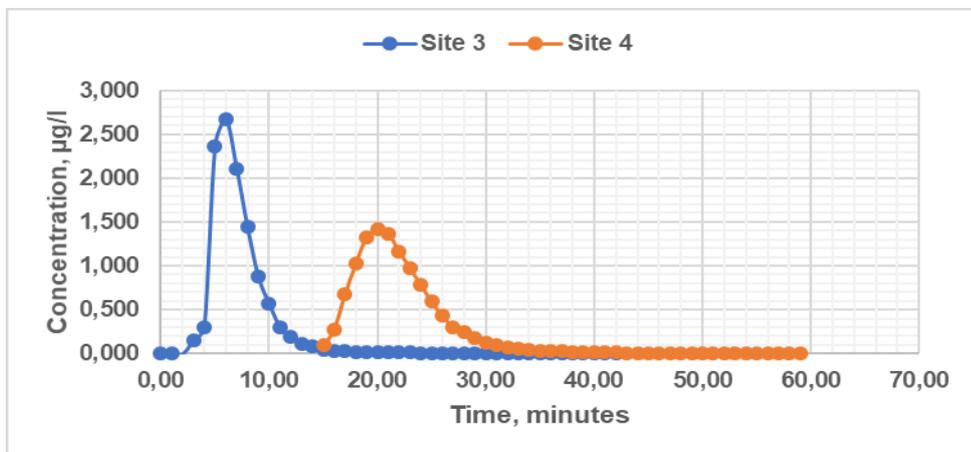
Experiment 12		
Time (min)	Upstream	Downstream
0.0		
0.5		
1.0		
1.5		
2.0		
2.5		
3.0		
3.5		
4.0		
4.5		
5.0		
5.5		
6.0	0.003	
6.5	0.106	
7.0	0.731	
7.5	1.865	
8.0	2.836	
8.5	3.347	
9.0	3.245	
9.5	2.836	
10.0	2.274	
10.5	1.660	
11.0	1.149	
11.5	0.844	
12.0	0.569	
12.5	0.407	0.010
13.0	0.265	0.051
13.5	0.183	0.186
14.0	0.122	0.562
14.5	0.091	1.071
15.0	0.061	1.376
15.5	0.048	2.000
16.0	0.032	2.312
16.5	0.024	2.468
17.0	0.021	2.364
17.5	0.016	2.156
18.0	0.010	1.792
18.5	0.010	1.532
19.0	0.008	1.121
19.5	0.008	0.923
20.0	0.006	0.710
20.5	0.003	0.513
21.0	0.003	0.398
21.5	0.006	0.300
22.0	0.003	0.217
22.5	0.003	0.155
23.0	0.003	0.108
23.5	0.003	0.098
24.0	0.002	0.067
24.5	0.002	0.052
25.0	0.002	0.039
25.5	0.002	0.033
26.0	0.002	0.023
26.5	0.002	0.021
27.0		0.018
27.5		0.011
28.0		0.015
28.5		0.013
29.0		0.010
29.5		0.010
30.0		0.010
30.5		0.008
31.0		0.006
31.5		0.006
32.0		0.006
32.5		0.006
33.0		0.005
33.5		0.003
34.0		0.003
34.5		0.003
35.0		0.003
35.5		0.002
36.0		0.003
36.5		0.003
37.0		0.002
37.5		0.002
38.0		0.003
38.5		0.002
39.0		0.002
39.5		0.002

Experiment 13						
Time (min)	Upstream	Downstream		Time (min)	Upstream	Downstream
0.00				64.67		0.003
0.67				65.33		0.005
1.33				66.00		0.000
2.00						
2.67						
3.33						
4.00						
4.67						
5.33						
6.00						
6.67	0.000					
7.33	0.000					
8.00	0.000					
8.67	0.000					
9.33	0.000					
10.00	0.011					
10.67	0.102					
11.33	0.392					
12.00	0.903					
12.67	1.500					
13.33	2.146					
14.00	2.493					
14.67	2.593					
15.33	2.493					
16.00	2.295					
16.67	1.947					
17.33	1.649					
18.00	1.351					
18.67	1.107					
19.33	0.918					
20.00	0.714					
20.67	0.510					
21.33	0.400	0.008				
22.00	0.328	0.022				
22.67	0.253	0.077				
23.33	0.189	0.211				
24.00	0.149	0.355				
24.67	0.114	0.577				
25.33	0.092	0.886				
26.00	0.073	1.146				
26.67	0.059	1.422				
27.33	0.050	1.604				
28.00	0.040	1.758				
28.67	0.033	1.810				
29.33	0.028	1.810				
30.00	0.025	1.758				
30.67	0.020	1.656				
31.33	0.017	1.450				
32.00	0.015	1.341				
32.67	0.014	1.195				
33.33	0.014	1.016				
34.00	0.011	0.837				
34.67	0.009	0.756				
35.33	0.008	0.561				
36.00	0.008	0.495				
36.67	0.006	0.414				
37.33	0.006	0.324				
38.00	0.004	0.257				
38.67	0.004	0.211				
39.33	0.004	0.190				
40.00	0.004	0.144				
40.67	0.004	0.113				
41.33	0.004	0.093				
42.00	0.003	0.079				
42.67	0.003	0.061				
43.33	0.003	0.055				
44.00	0.004	0.047				
44.67	0.003	0.044				
45.33	0.003	0.037				
46.00	0.003	0.032				
46.67	0.003	0.029				
47.33	0.003	0.024				
48.00	0.003	0.021				
48.67		0.019				
49.33		0.018				
50.00		0.014				
50.67		0.014				
51.33		0.013				
52.00		0.011				
52.67		0.011				
53.33		0.008				
54.00		0.009				
54.67		0.009				
55.33		0.009				
56.00		0.008				
56.67		0.008				
57.33		0.006				
58.00		0.006				
58.67		0.006				
59.33		0.006				
60.00		0.005				
60.67		0.005				
61.33		0.005				
62.00		0.005				
62.67		0.005				
63.33		0.005				
64.00		0.006				

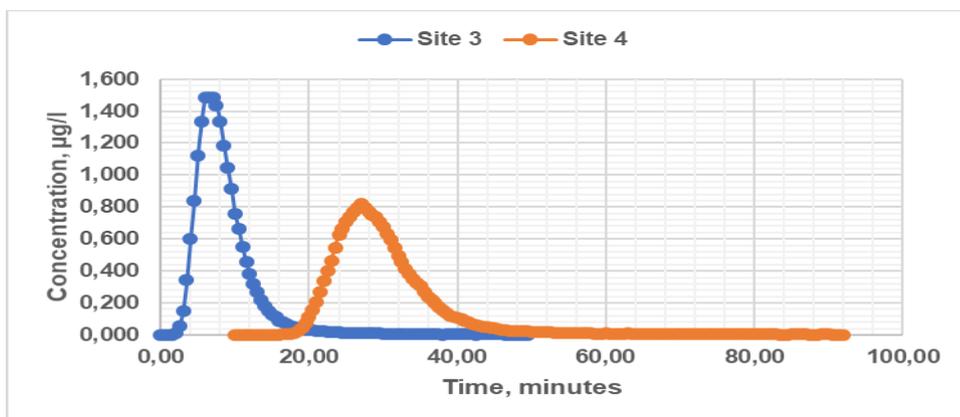
Appendix A2: Charts for analysed concentration-time data for experiments 2 – 13.



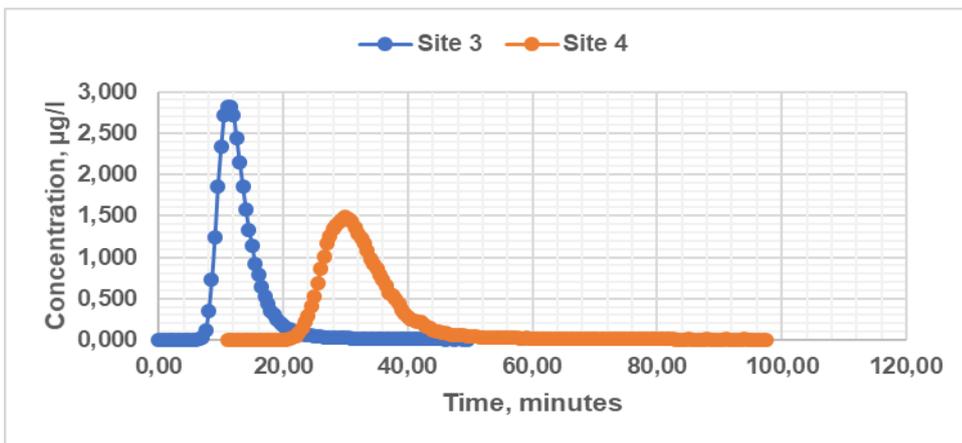
Experiment 2: Concentration-time profiles



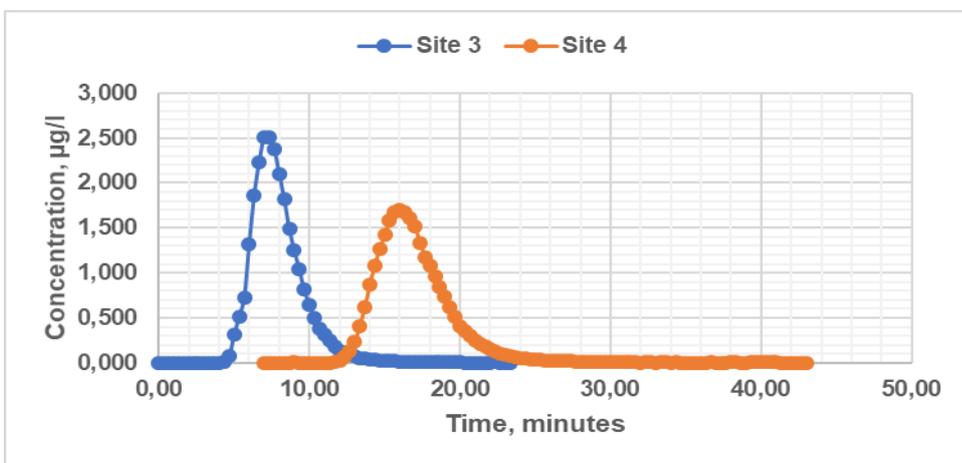
Experiment 3: Concentration-time profiles



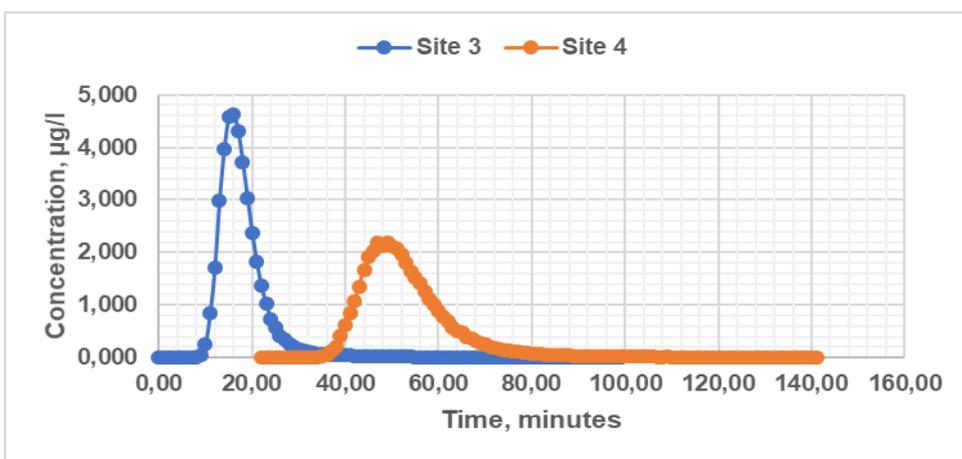
Experiment 4: Concentration-time profiles



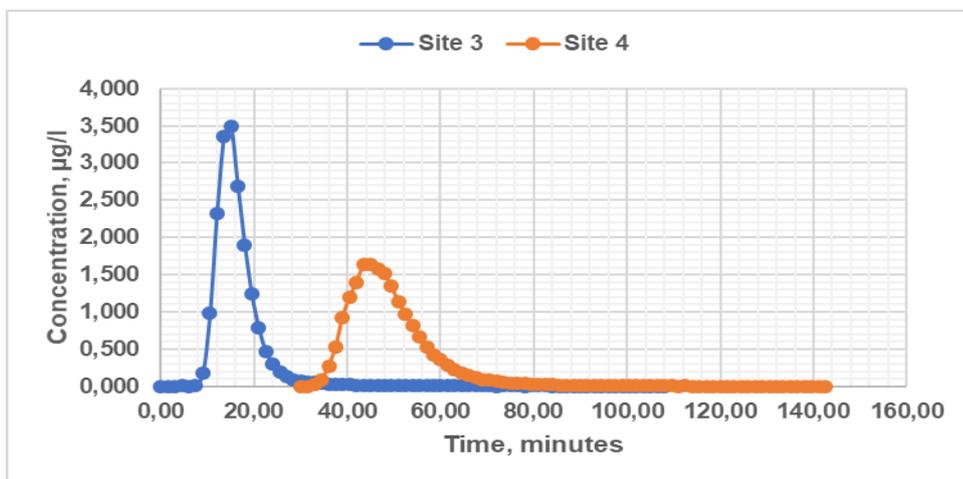
Experiment 5: Concentration-time profiles



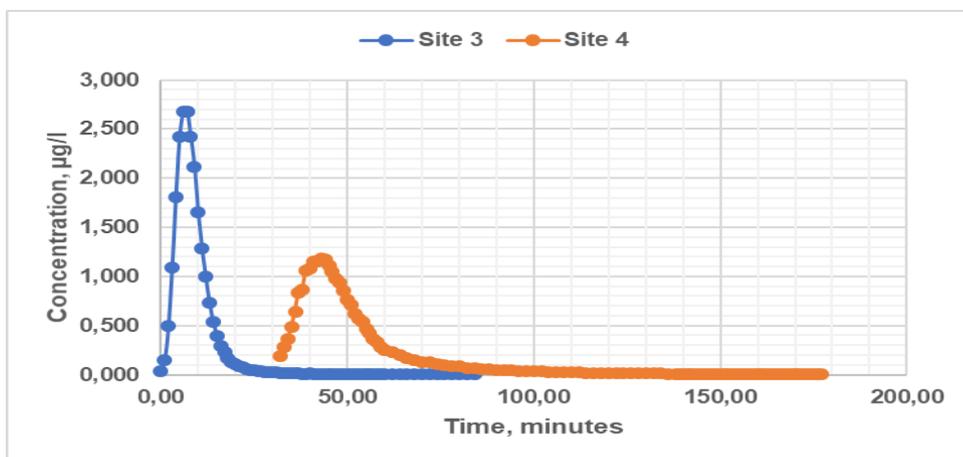
Experiment 6: Concentration-time profiles



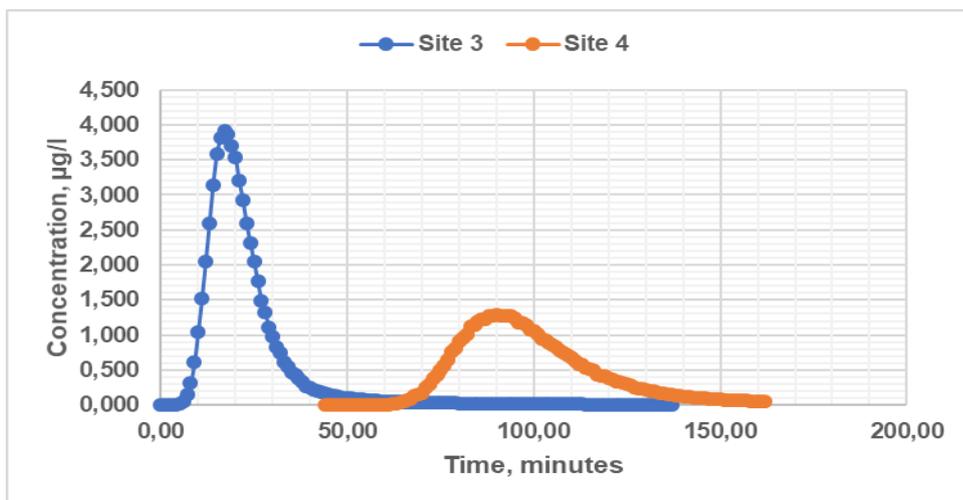
Experiment 7: Concentration-time profiles



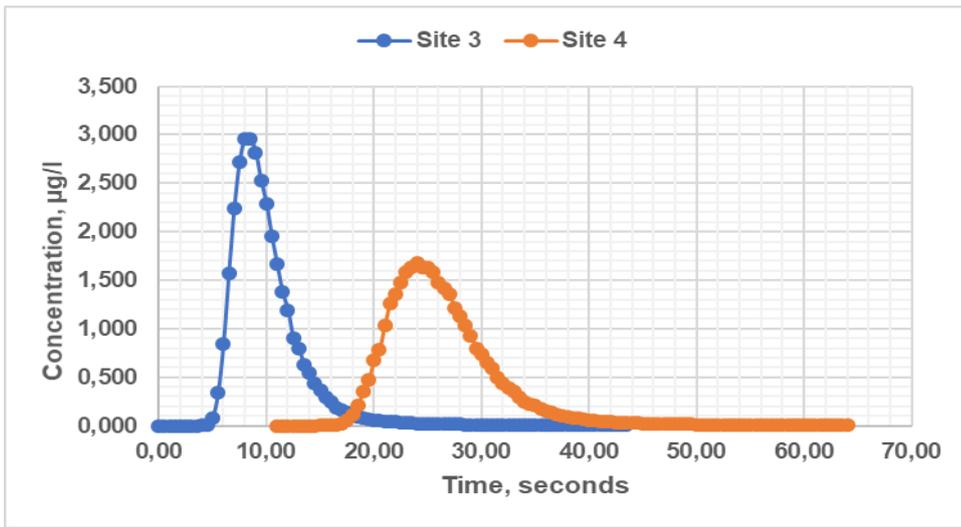
Experiment 8: Concentration-time profiles



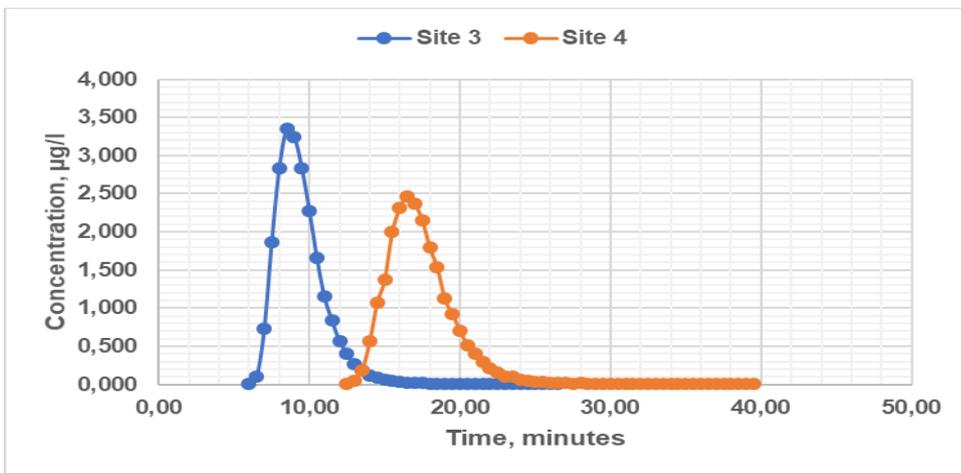
Experiment 9: Concentration-time profiles



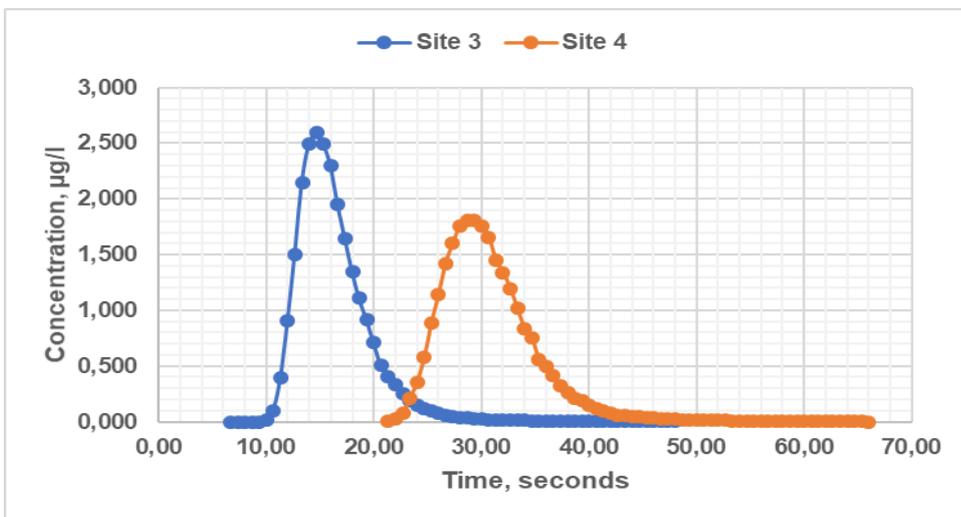
Experiment 10: Concentration-time profiles



Experiment 11: Concentration-time profiles



Experiment 12: Concentration-time profiles



Experiment 13: Concentration-time profiles

Appendix B: Critical values of Shapiro-Wilk normality test

Table of critical values of W_0 statistic for sample sizes n and significance level α (Hanusz and Tarasinska, 2016).

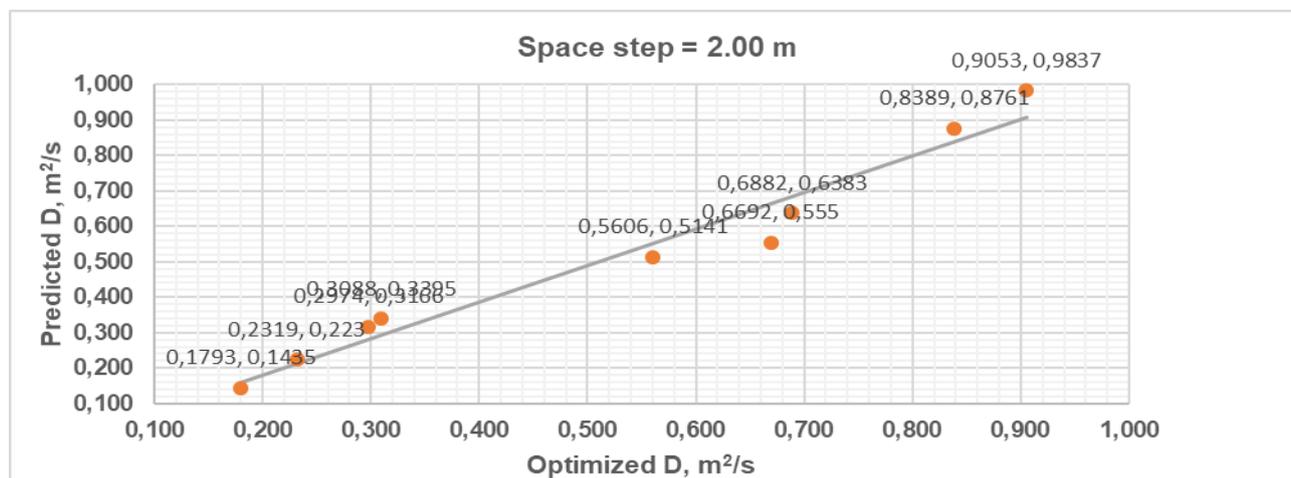
n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
3	0.0184	0.0881	0.1714
4	0.0721	0.2037	0.3127
5	0.1419	0.3086	0.4190
6	0.2090	0.3867	0.4952
7	0.2742	0.4525	0.5543
8	0.3299	0.5051	0.5998
9	0.3785	0.5493	0.6374
10	0.4233	0.5852	0.6682
11	0.4606	0.6165	0.6935
12	0.4940	0.6431	0.7154
13	0.5246	0.6661	0.7346
14	0.5494	0.6862	0.7504
15	0.5739	0.7038	0.7651
16	0.5954	0.7196	0.7778
17	0.6126	0.7337	0.7890
18	0.6319	0.7476	0.7998
19	0.6478	0.7590	0.8088
20	0.6626	0.7696	0.8176
21	0.6761	0.7792	0.8250
22	0.6876	0.7875	0.8319
23	0.7008	0.7965	0.8390
24	0.7104	0.8034	0.8446
25	0.7205	0.8103	0.8501
26	0.7296	0.8170	0.8553

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
27	0.7379	0.8232	0.8601
28	0.7463	0.8287	0.8645
29	0.7539	0.8340	0.8688
30	0.7611	0.8394	0.8730
31	0.7677	0.8437	0.8765
32	0.7746	0.8482	0.8800
33	0.7804	0.8524	0.8834
34	0.7871	0.8565	0.8863
35	0.7917	0.8602	0.8894
36	0.7969	0.8634	0.8921
37	0.8008	0.8670	0.8947
38	0.8063	0.8701	0.8972
39	0.8109	0.8731	0.8996
40	0.8145	0.8760	0.9018
41	0.8194	0.8787	0.9040
42	0.8227	0.8816	0.9061
43	0.8271	0.8839	0.9081
44	0.8301	0.8862	0.9100
45	0.8343	0.8887	0.9120
46	0.8374	0.8911	0.9138
47	0.8403	0.8931	0.9154
48	0.8433	0.8951	0.9169
49	0.8470	0.8974	0.9187
50	0.8491	0.8989	0.9200

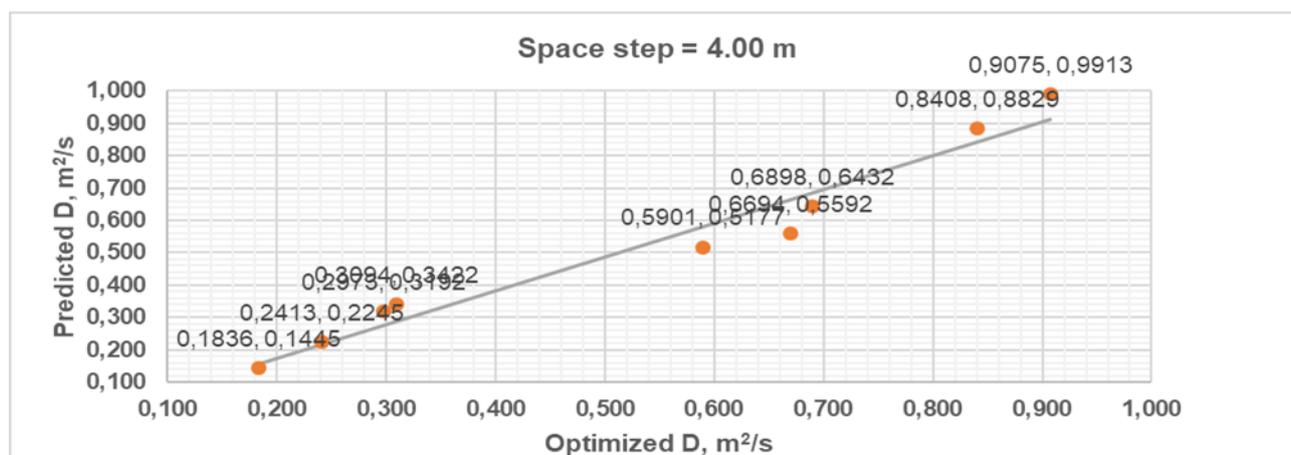
Appendix C: Results of graphical method of performance analysis

The figures show plots of predicted dispersion coefficients versus optimized dispersion coefficients, correlation coefficients, r , coefficients of determination, R^2 and root mean square error, RMSE.

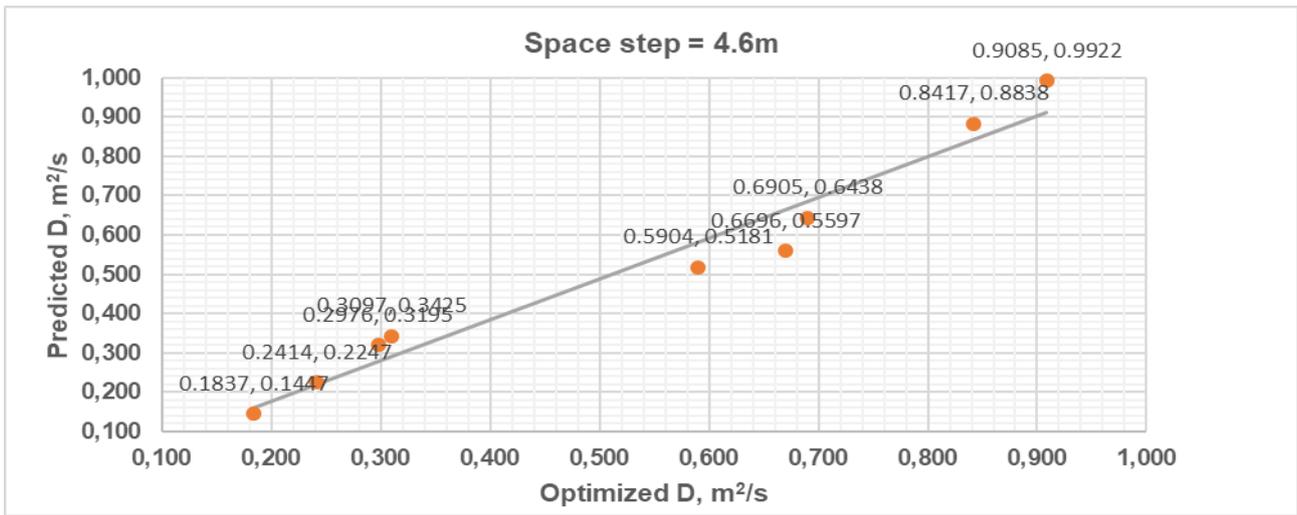
Appendix C-1: Crank-Nicolson based model



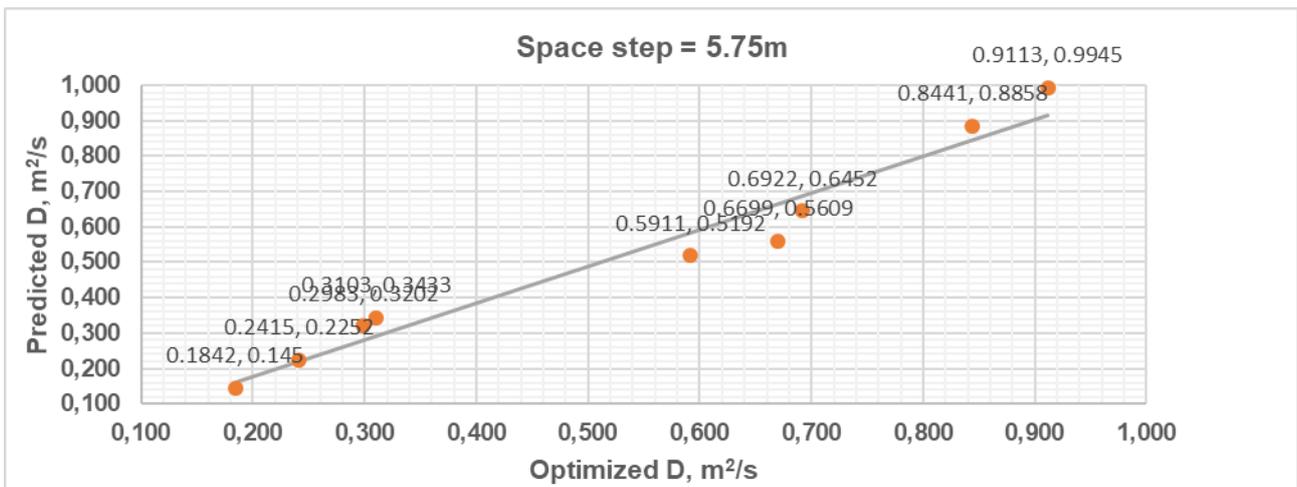
Pearson's r	Numerical model D	Predictive model, D
Numerical model D	-	0.980
Predictive model, D	0.980	-
R^2	0.960	
R^2 adjusted	0.954	
RMSE	0.061412816	



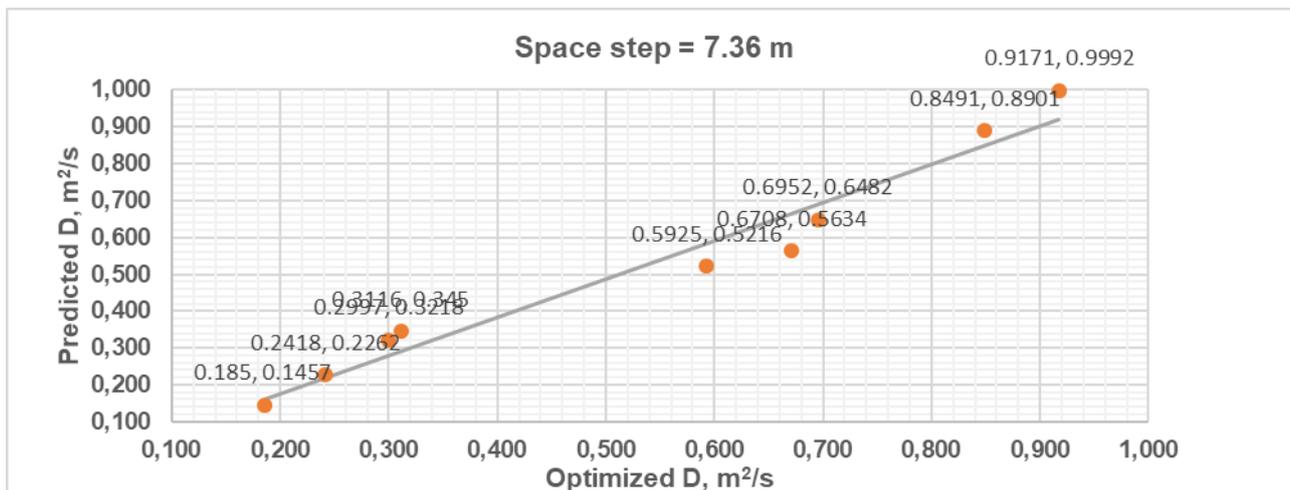
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.978
Predictive model, D	0.978	-
R^2	0.956	
R^2 adjusted	0.950	
RMSE	0.065052286	



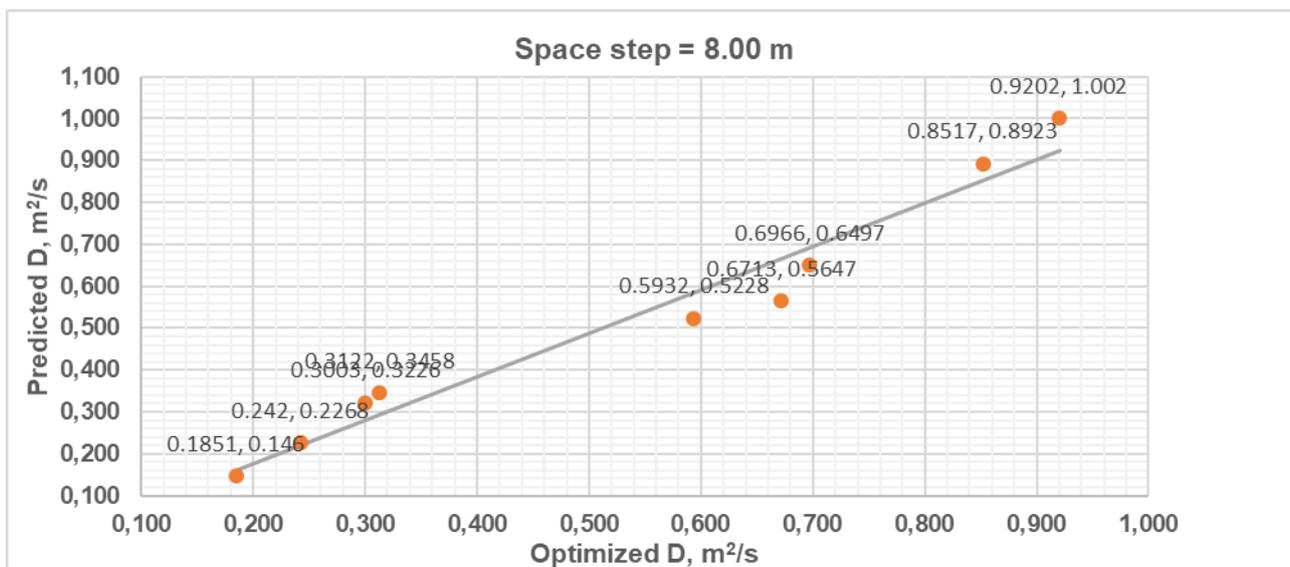
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.978
Predictive model, D	0.978	-
R ²	0.956	
R ² adjusted	0.950	
RMSE	0.064927544	



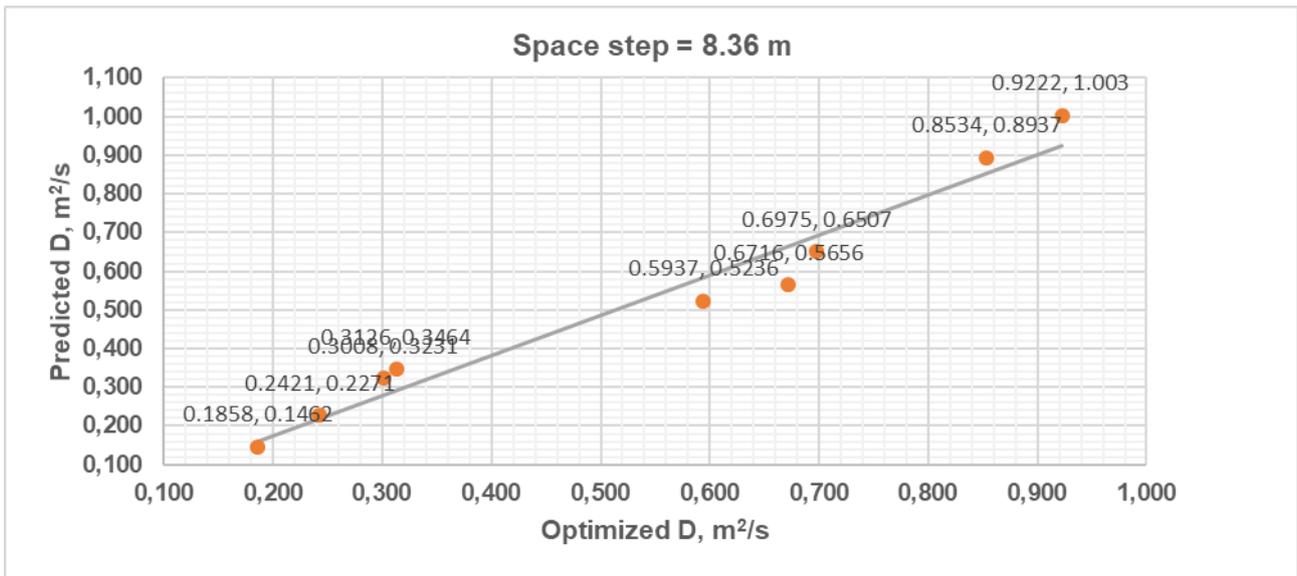
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.978
Predictive model, D	0.978	-
R ²	0.957	
R ² adjusted	0.951	
RMSE	0.064600055	



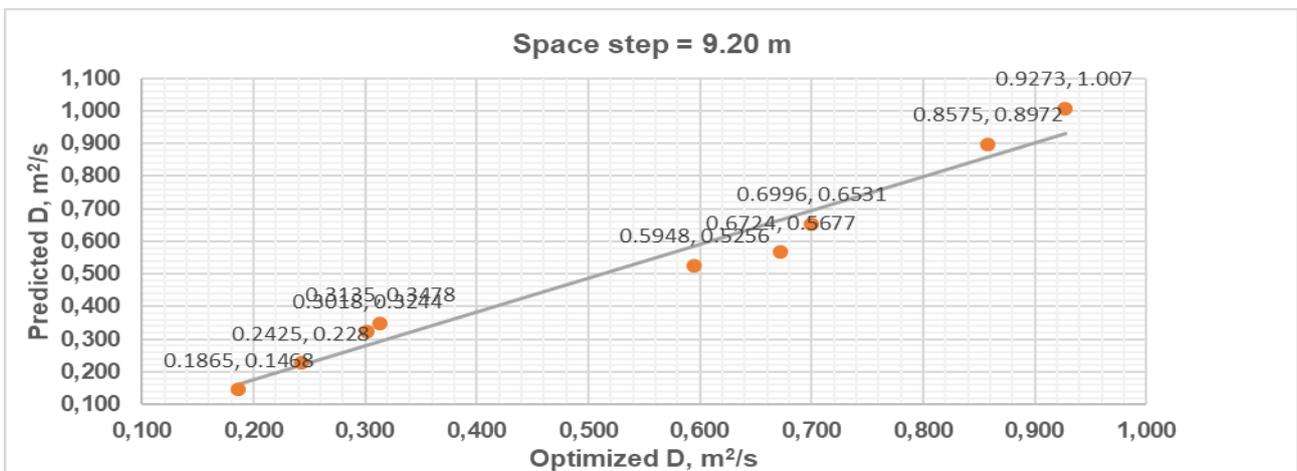
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.979
Predictive model, D	0.979	-
R ²	0.958	
R ² adjusted	0.952	
RMSE	0.063869546	



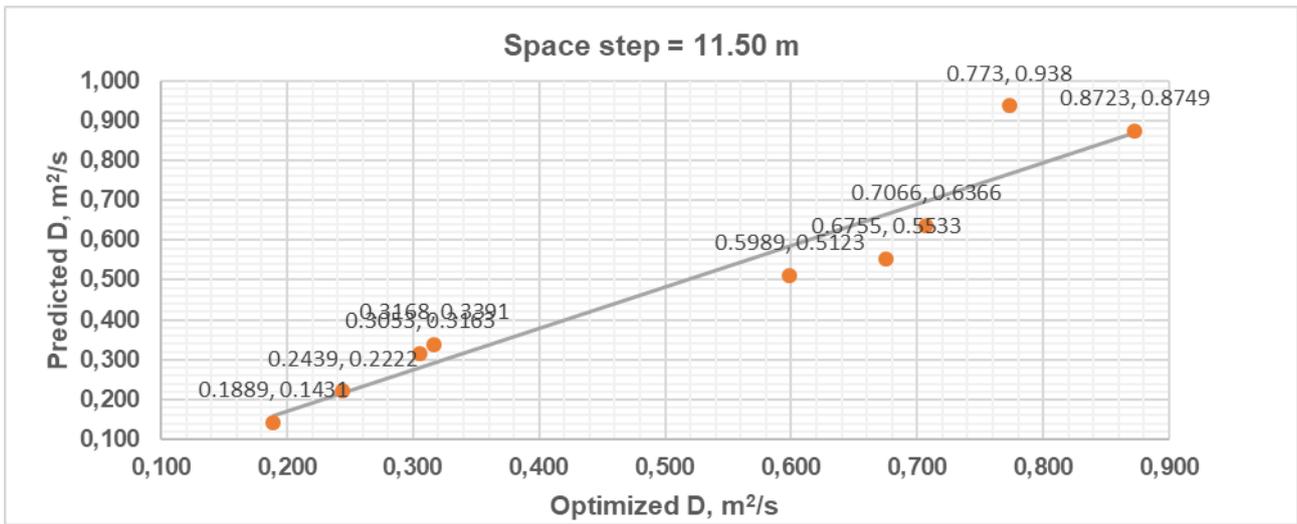
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.979
Predictive model, D	0.979	-
R ²	0.959	
R ² adjusted	0.953	
RMSE	0.063458472	



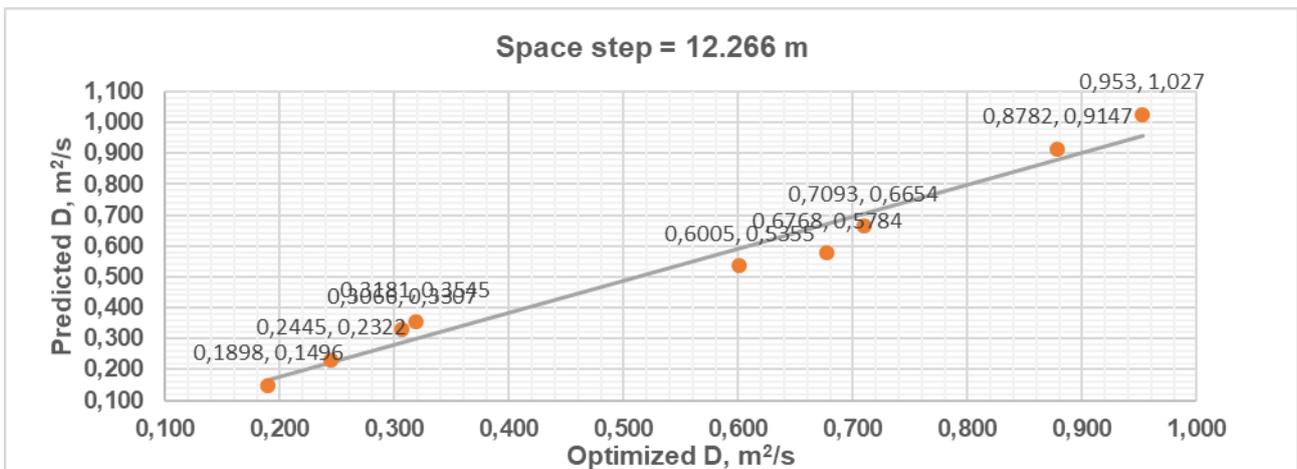
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.979
Predictive model, D	0.979	-
R ²	0.959	
R ² adjusted	0.953	
RMSE	0.063229217	



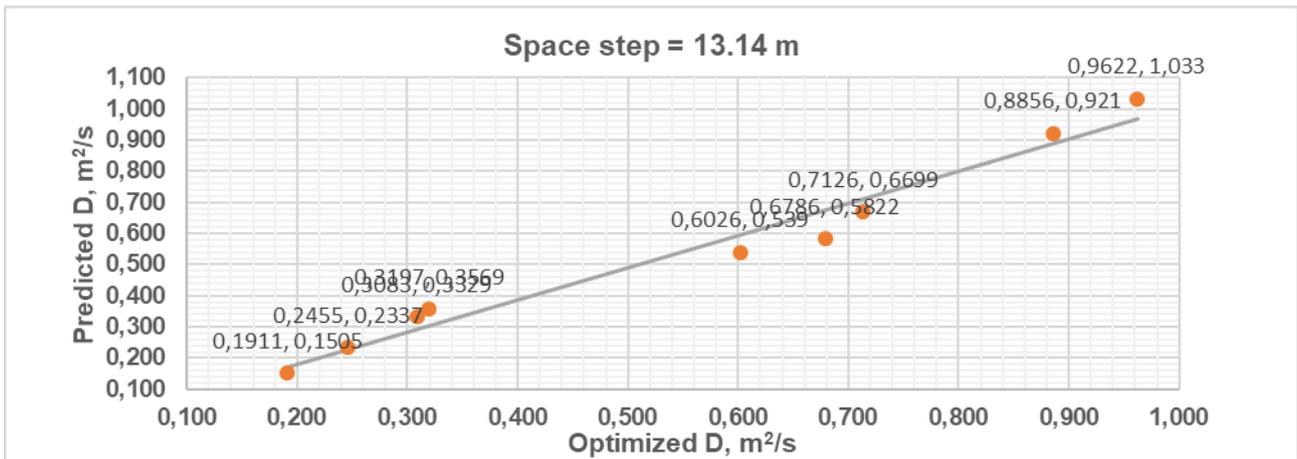
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.980
Predictive model, D	0.980	-
R ²	0.960	
R ² adjusted	0.955	
RMSE	0.062567672	



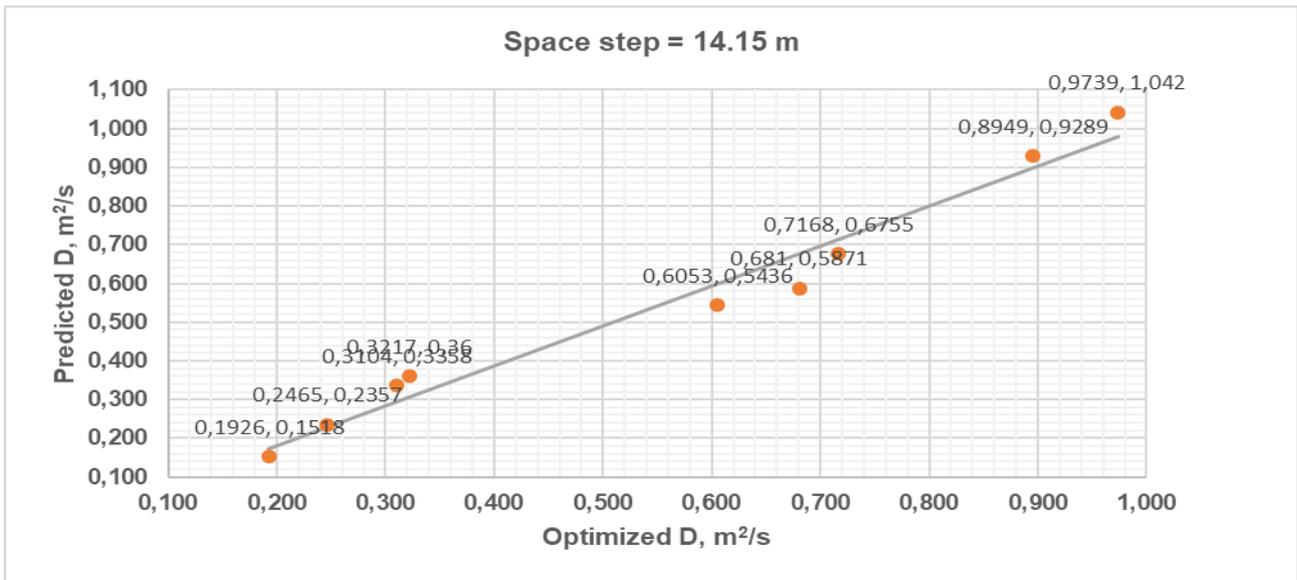
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.955
Predictive model, D	0.955	-
R ²	0.911	
R ² adjusted	0.899	
RMSE	0.088490507	



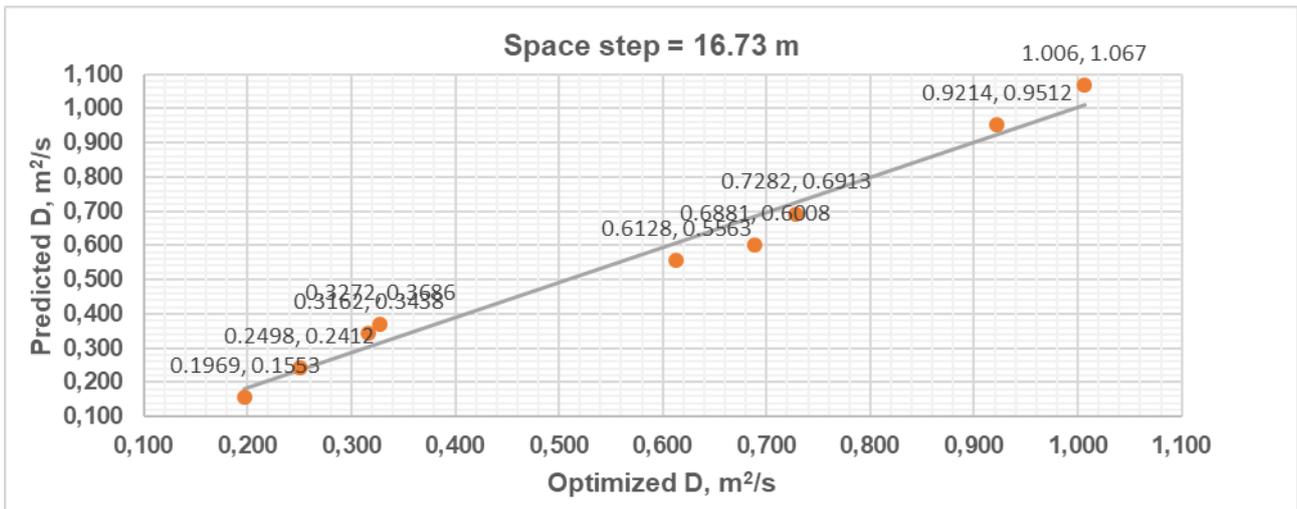
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.983
Predictive model, D	0.983	-
R ²	0.966	
R ² adjusted	0.961	
RMSE	0.059318940	



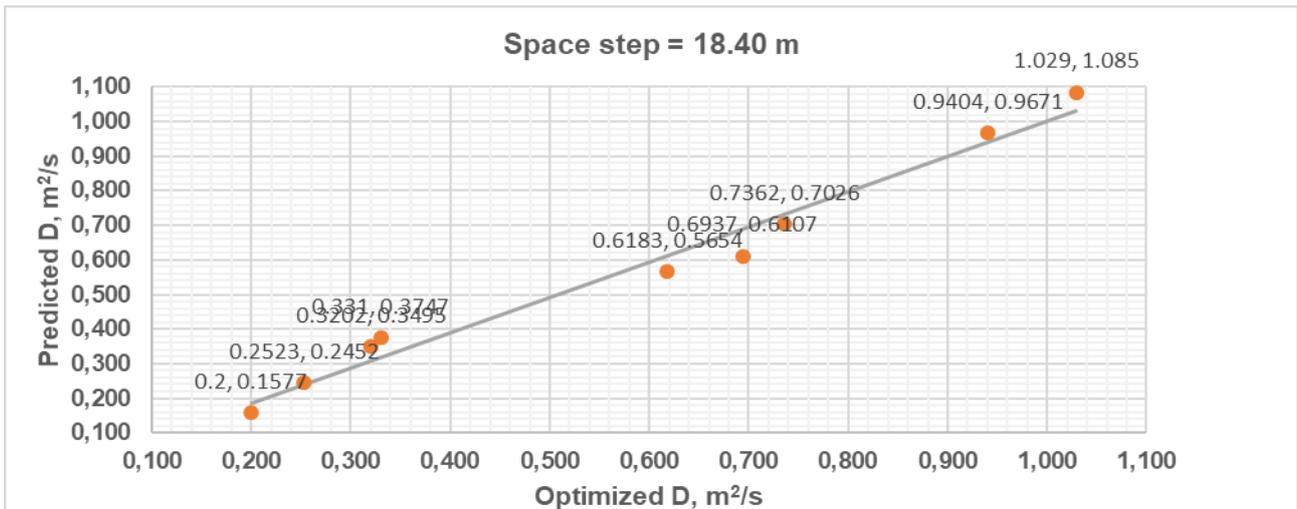
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.984
Predictive model, D	0.984	-
R ²	0.967	
R ² adjusted	0.963	
RMSE	0.058206904	



Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.985
Predictive model, D	0.985	-
R ²	0.969	
R ² adjusted	0.965	
RMSE	0.056883351	

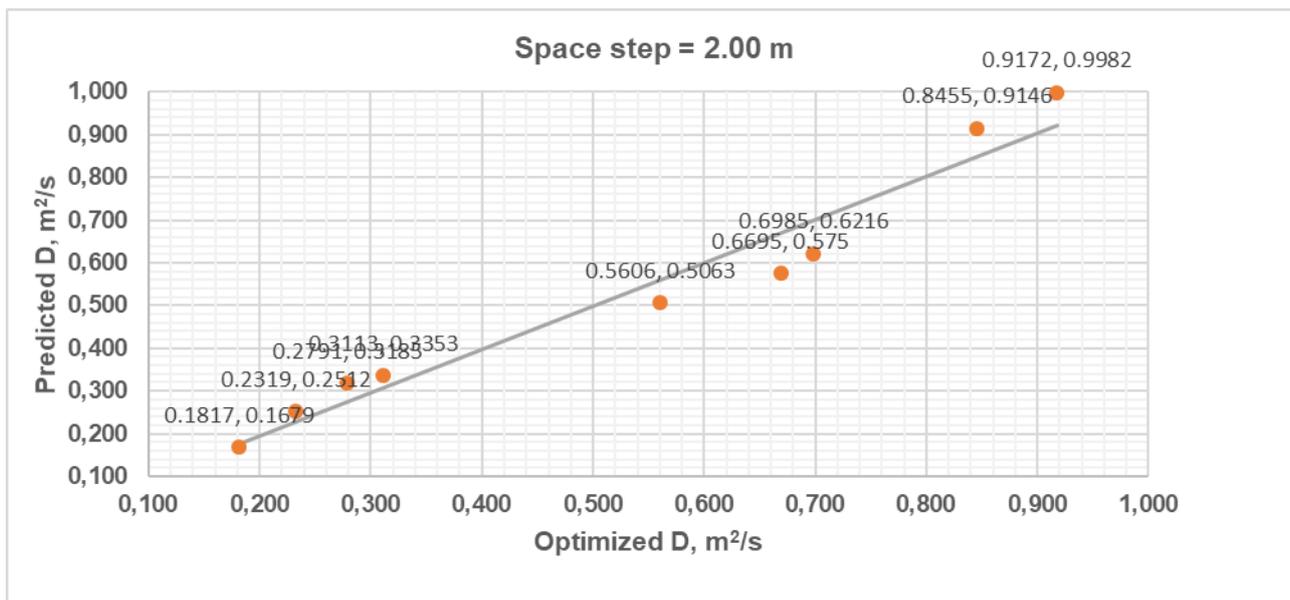


Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.987
Predictive model, D	0.987	-
R ²	0.974	
R ² adjusted	0.971	
RMSE	0.053403265	

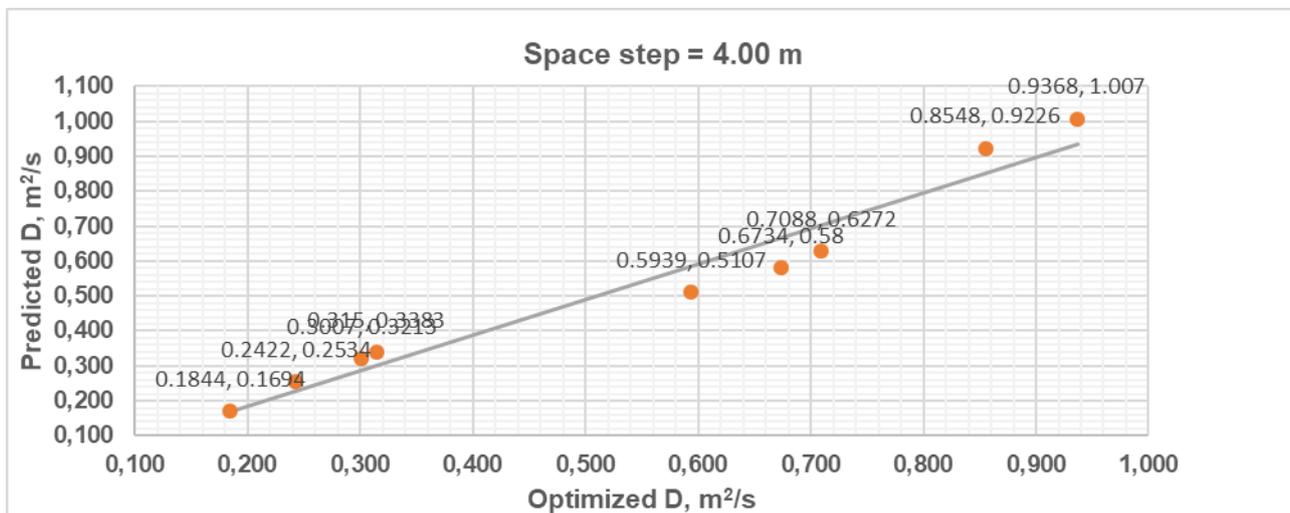


Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.989
Predictive model, D	0.989	-
R ²	0.977	
R ² adjusted	0.974	
RMSE	0.051225393	

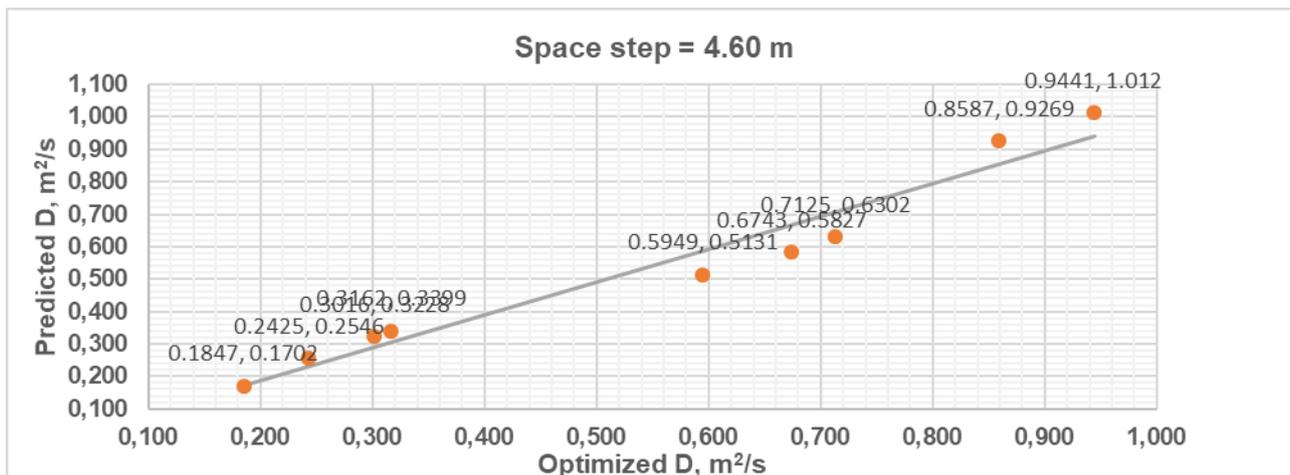
Appendix C-2: MacCormack based model



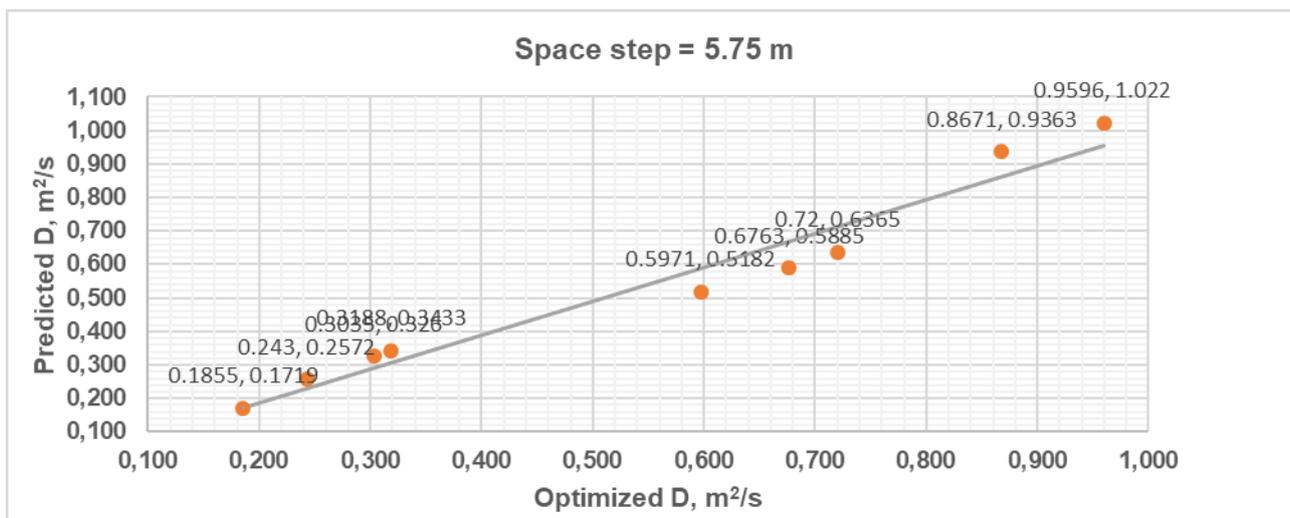
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.976
Predictive model, D	0.976	-
R ²	0.953	
R ² adjusted	0.946	
RMSE	0.067283860	



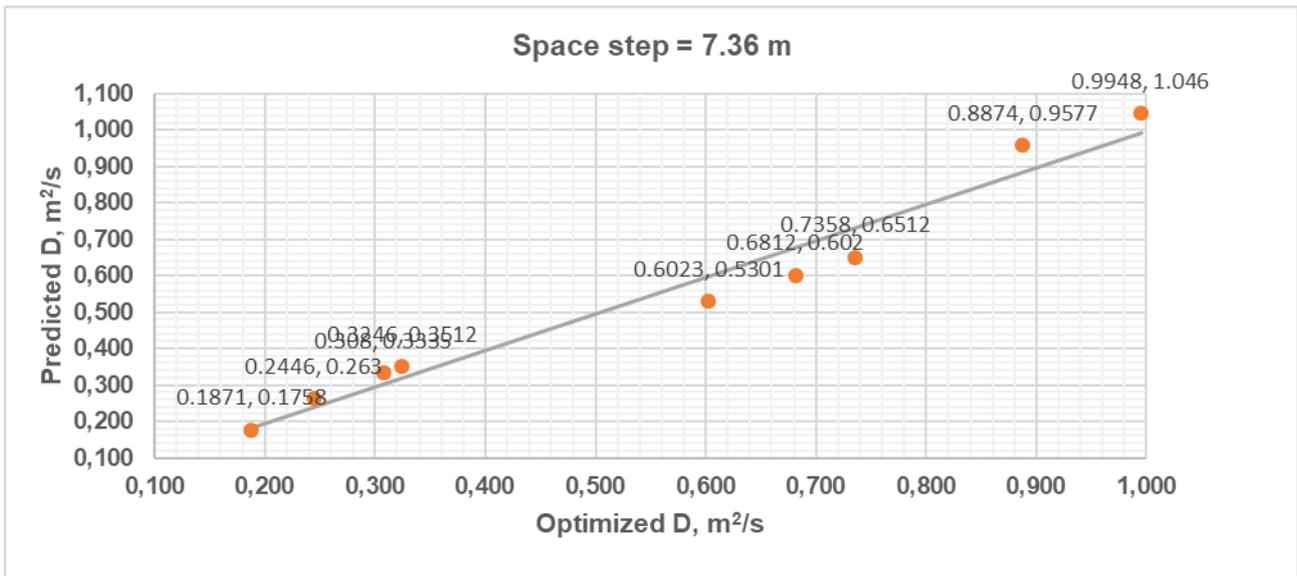
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.976
Predictive model, D	0.976	-
R ²	0.953	
R ² adjusted	0.946	
RMSE	0.067908734	



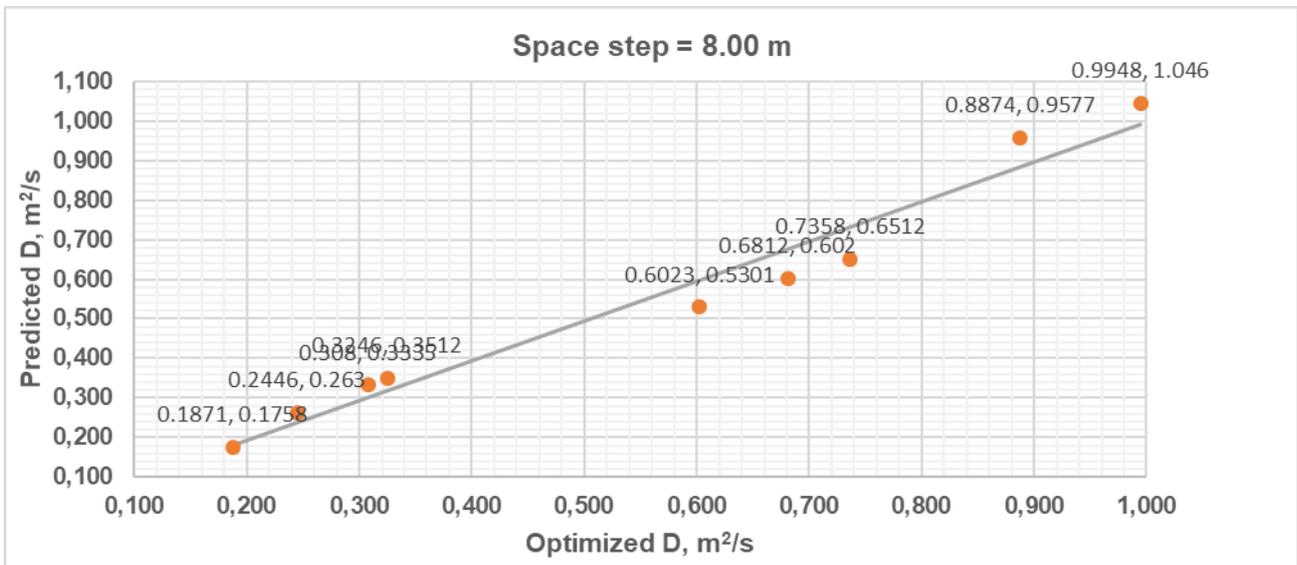
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.977
Predictive Model, D	0.977	-
R ²	0.954	
R ² adjusted	0.947	
RMSE	0.067275131	



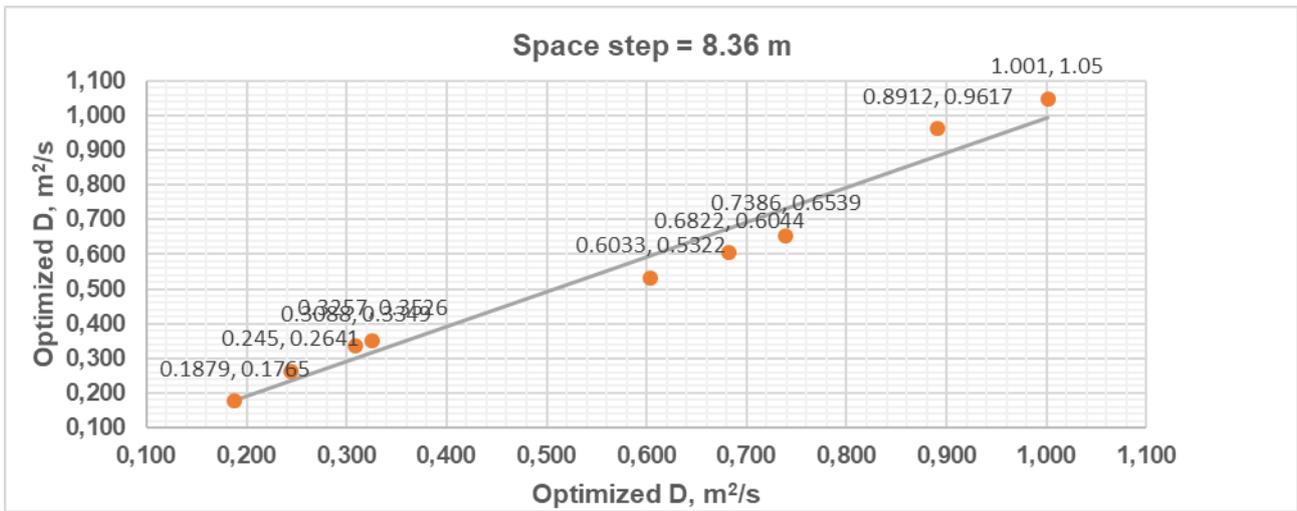
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.978
Predictive model, D	0.978	-
R ²	0.957	
R ² adjusted	0.950	
RMSE	0.065916878	



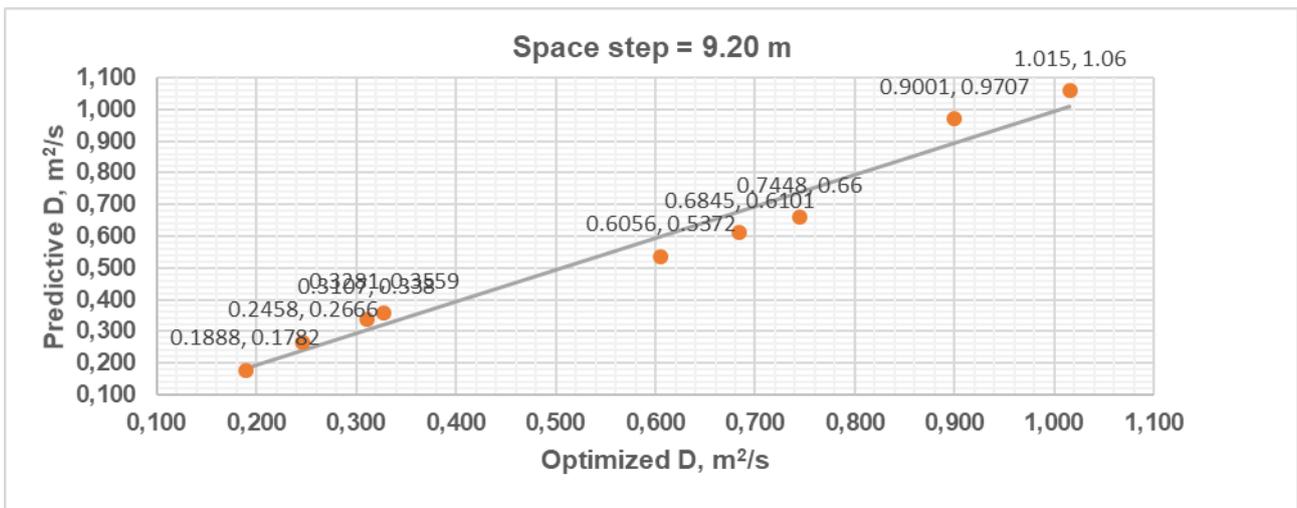
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.981
Predictive model, D	0.981	-
R ²	0.962	
R ² adjusted	0.957	
RMSE	0.062855824	



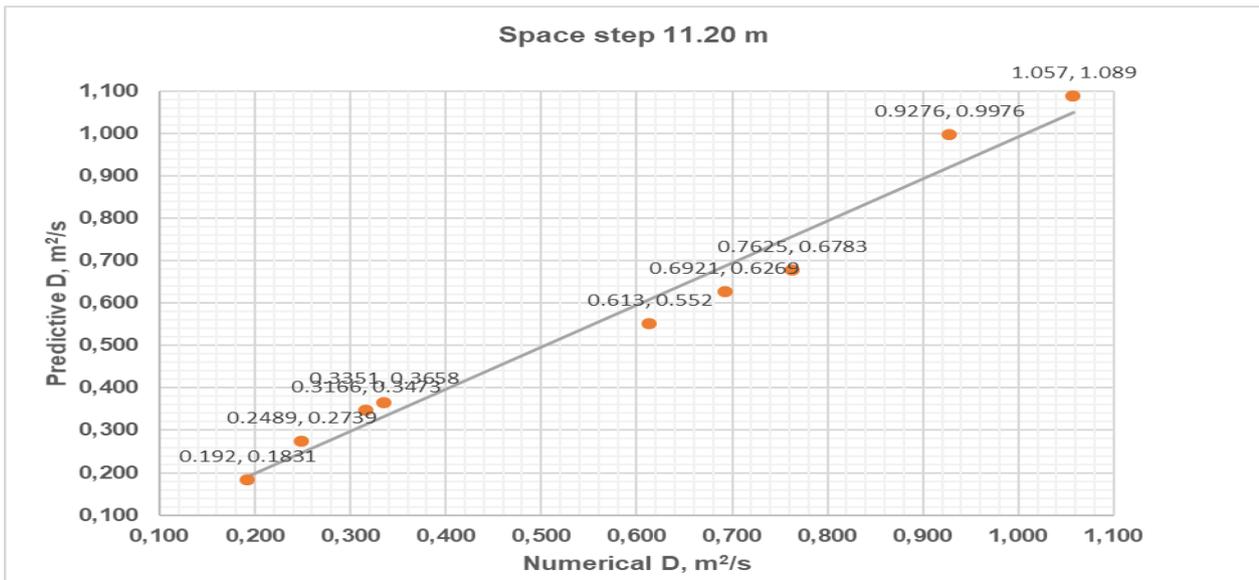
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.981
Predictive model, D	0.981	-
R ²	0.962	
R ² adjusted	0.957	
RMSE	0.062860856	



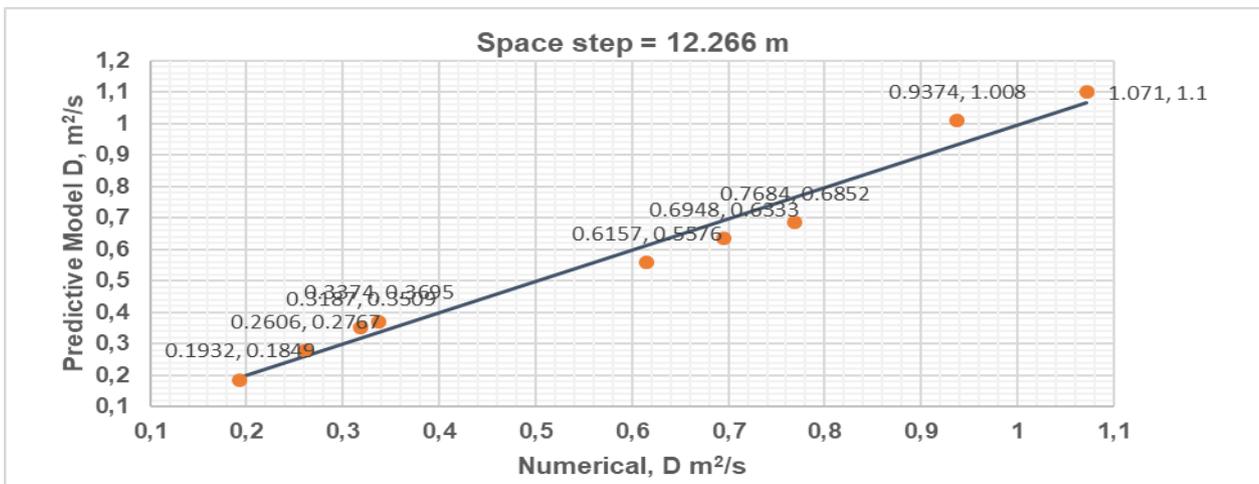
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.981
Predictive model, D	0.981	-
R ²	0.963	
R ² adjusted	0.958	
RMSE	0.062347443	



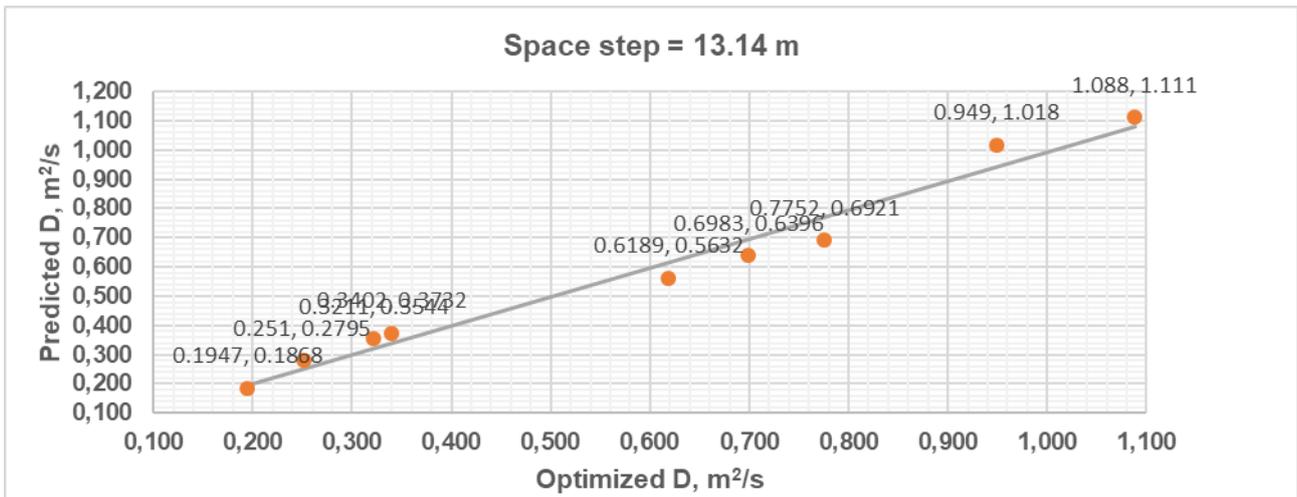
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.982
Predictive model, D	0.982	-
R ²	0.965	
R ² adjusted	0.960	
RMSE	0.061145490	



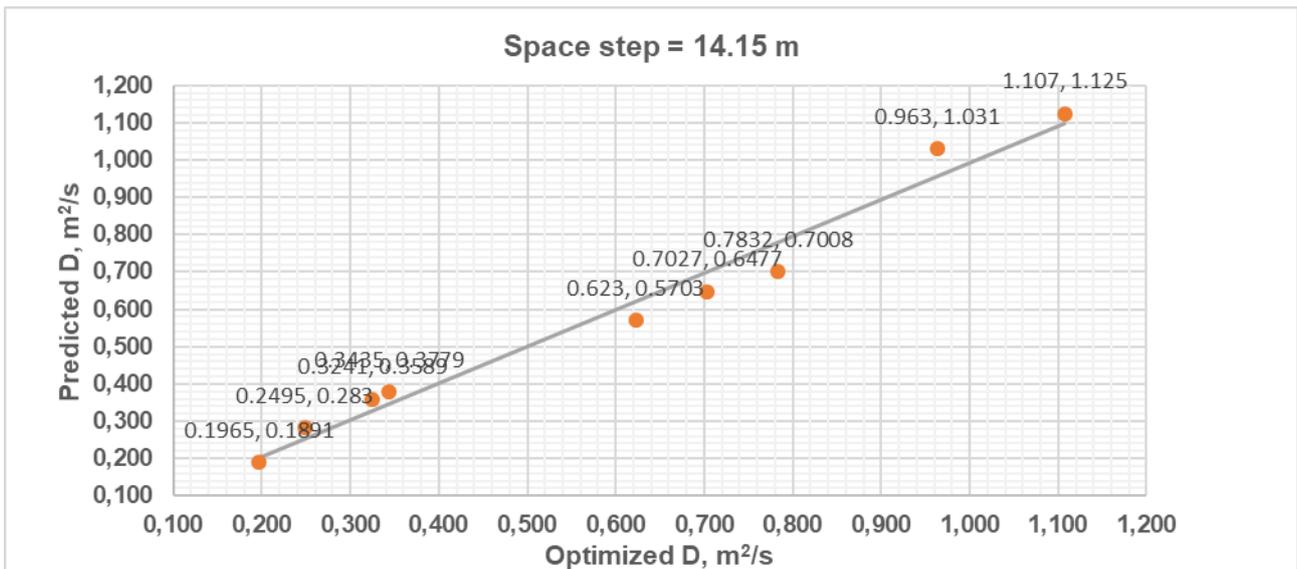
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.985
Predictive model, D	0.985	-
R ²	0.971	
R ² adjusted	0.966	
RMSE	0.057833000	



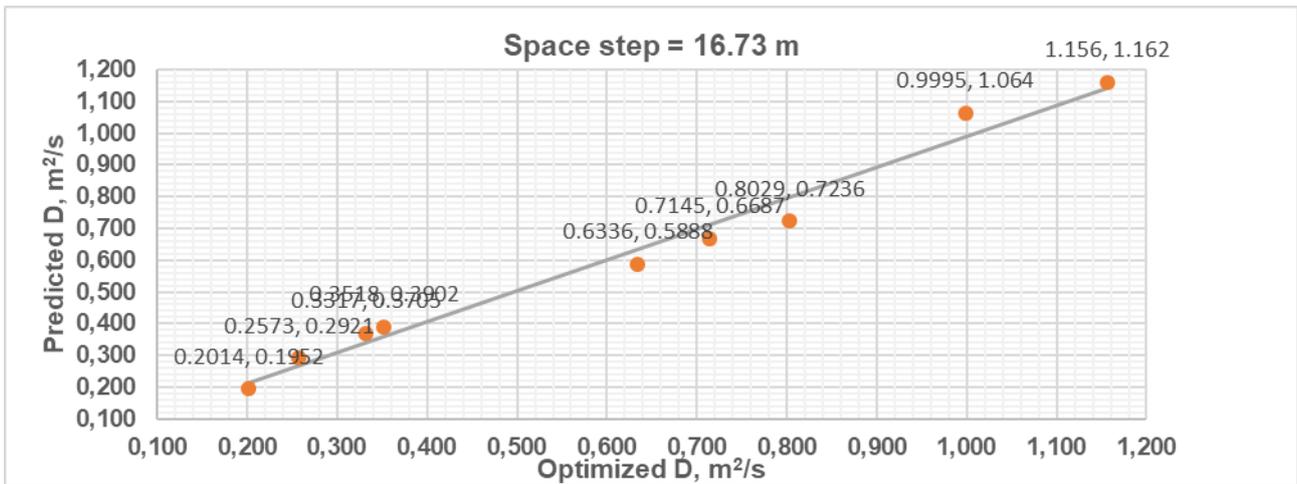
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.986
Predictive model, D	0.986	-
R ²	0.973	
R ² adjusted	0.969	
RMSE	0.056204335	



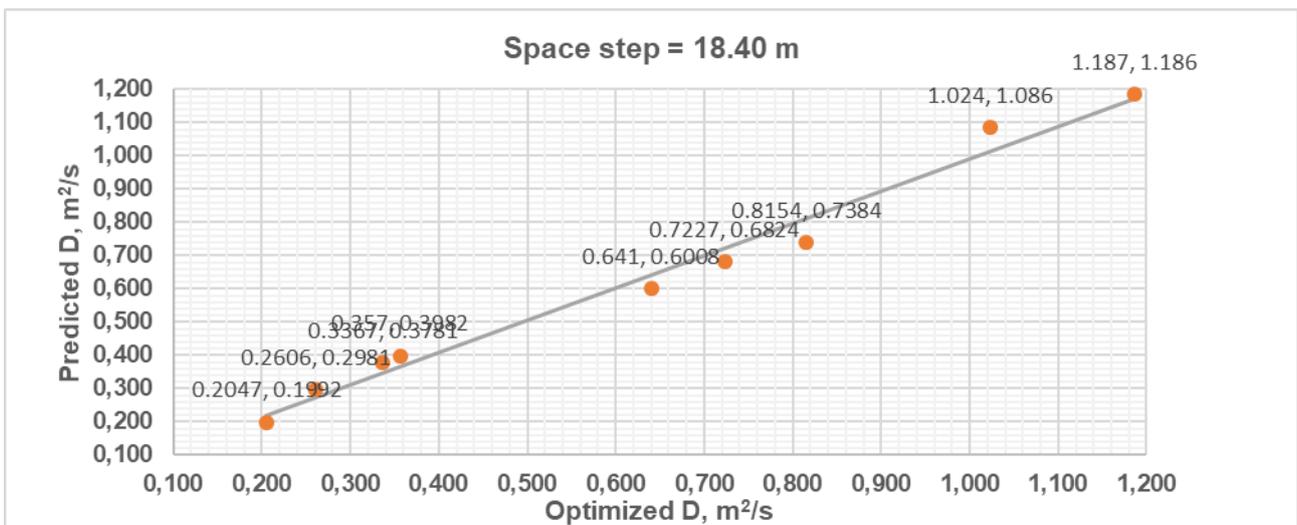
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.987
Predictive model, D	0.987	-
R ²	0.974	
R ² adjusted	0.970	
RMSE	0.055565980	



Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.988
Predictive model, D	0.988	-
R ²	0.976	
R ² adjusted	0.972	
RMSE	0.054443078	

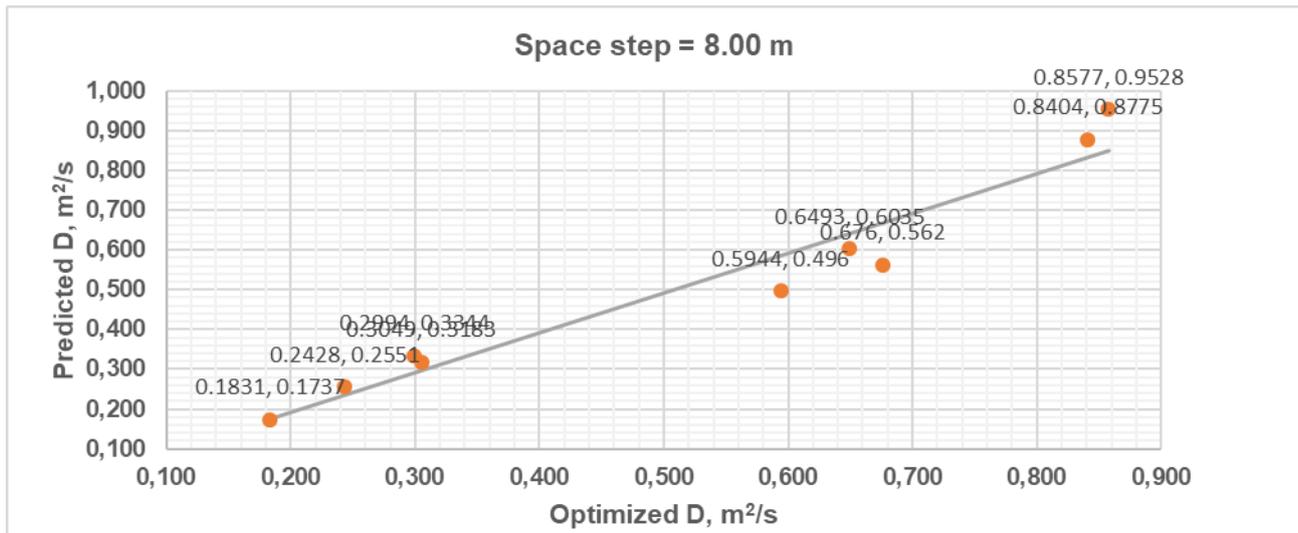


Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.990
Predictive model, D	0.990	-
R ²	0.980	
R ² adjusted	0.977	
RMSE	0.051093970	

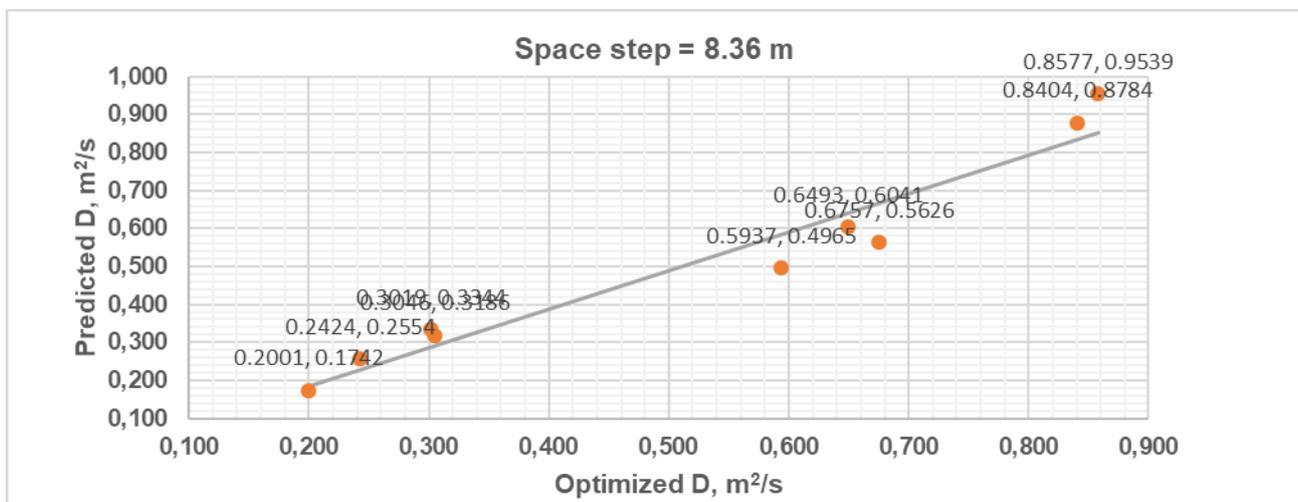


Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.991
Predictive model, D	0.991	-
R ²	0.982	
R ² adjusted	0.979	
RMSE	0.049304336	

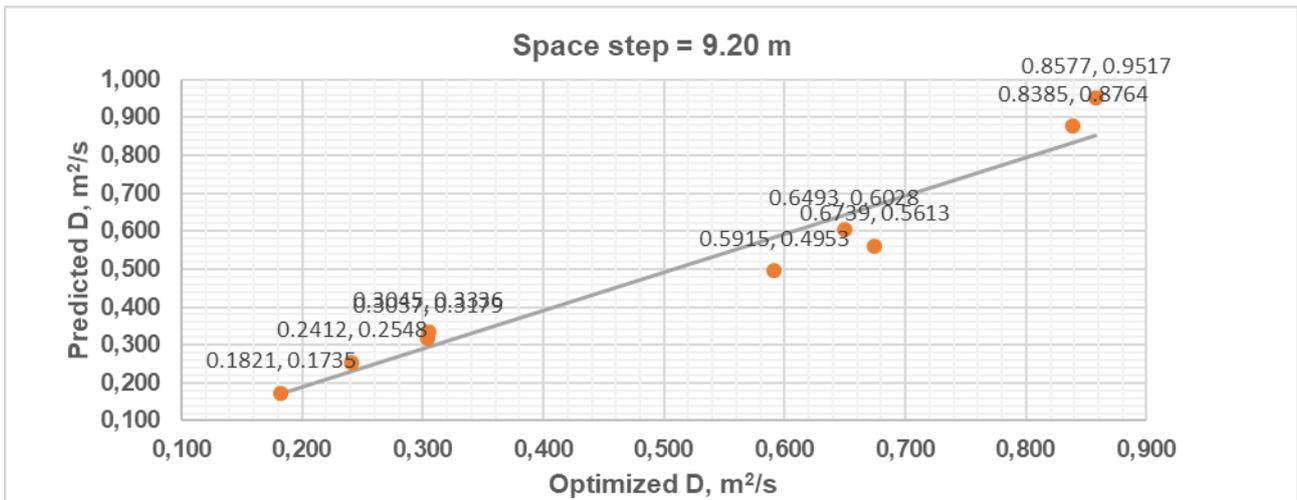
Appendix C-3: QUICKEST based model



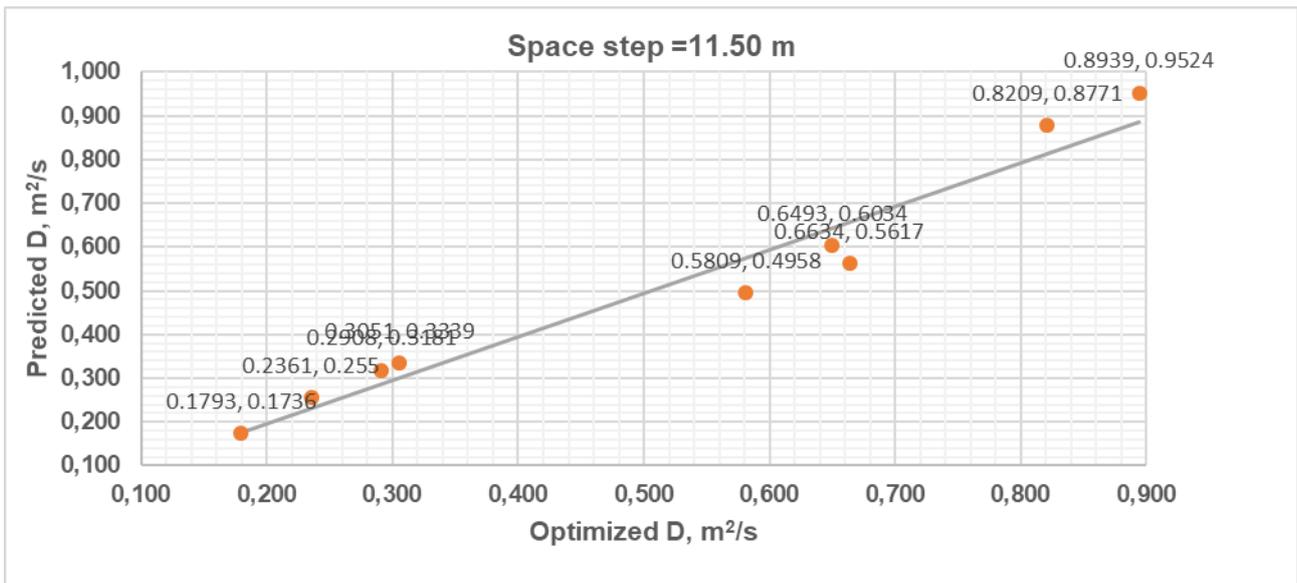
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.969
Predictive model, D	0.969	-
R ²	0.938	
R ² adjusted	0.930	
RMSE	0.071944120	



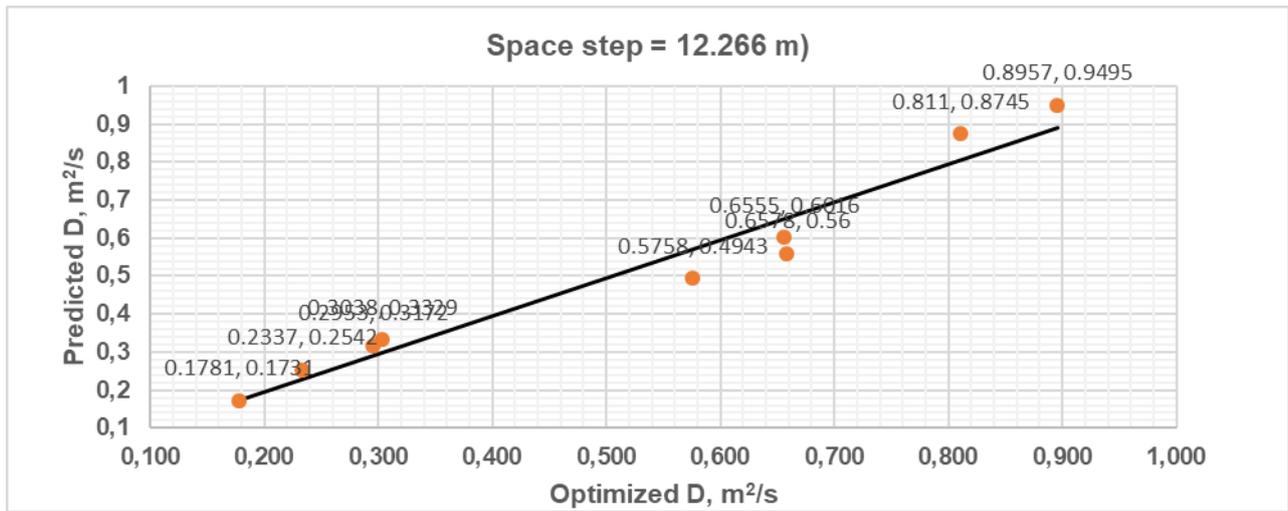
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.969
Predictive model, D	0.969	-
R ²	0.939	
R ² adjusted	0.930	
RMSE	0.071794283	



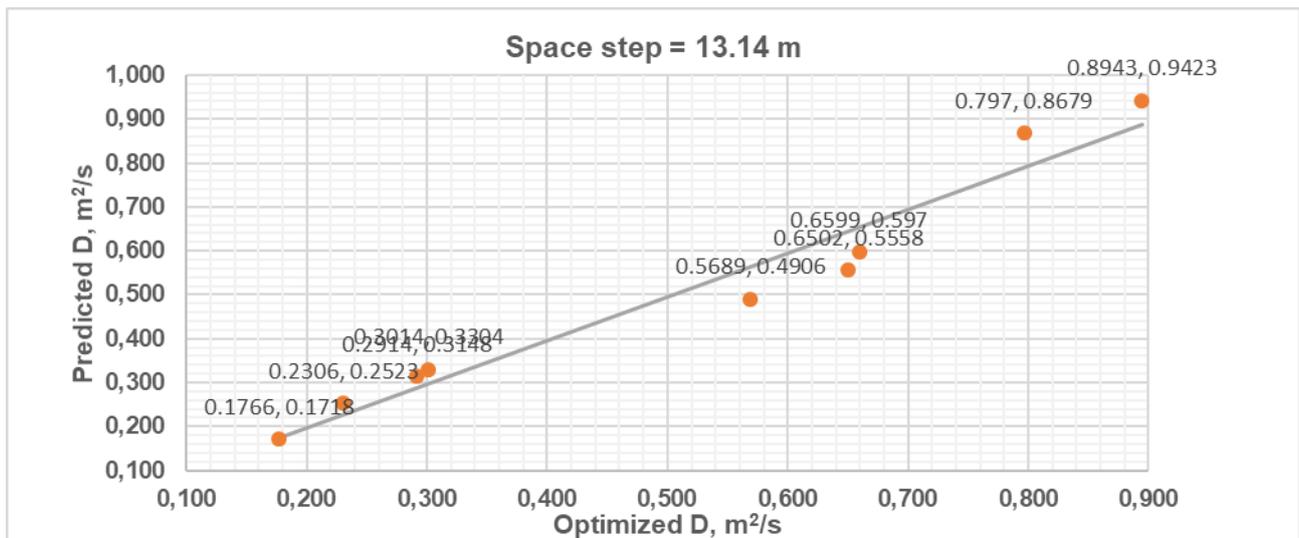
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.970
Predictive model, D	0.970	-
R ²	0.940	
R ² adjusted	0.932	
RMSE	0.070748822	



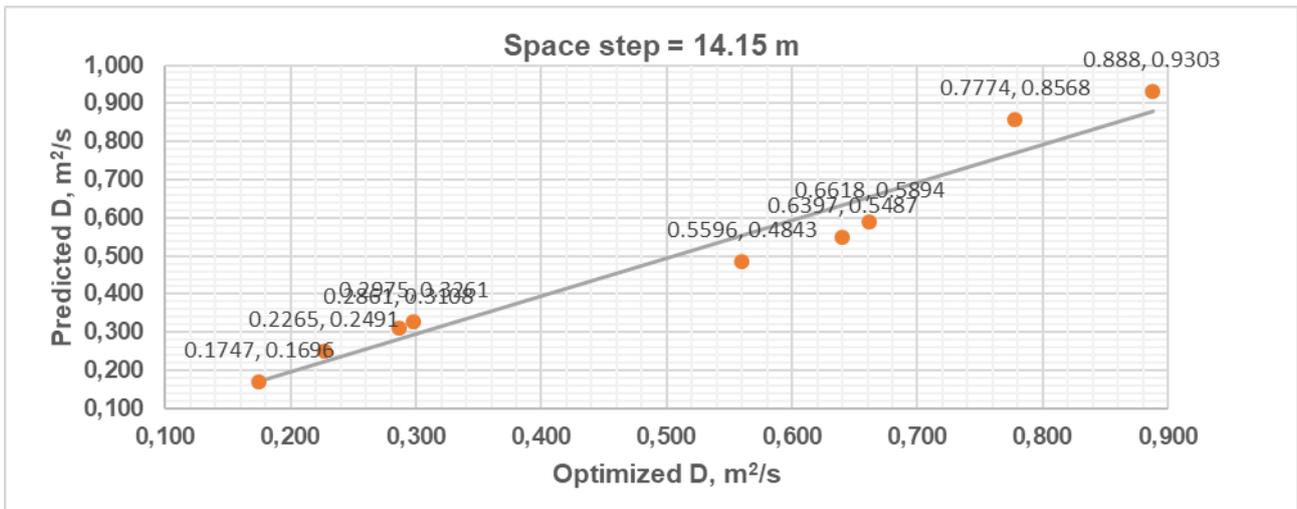
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.976
Predictive model, D	0.976	-
R ²	0.952	
R ² adjusted	0.946	
RMSE	0.063195714	



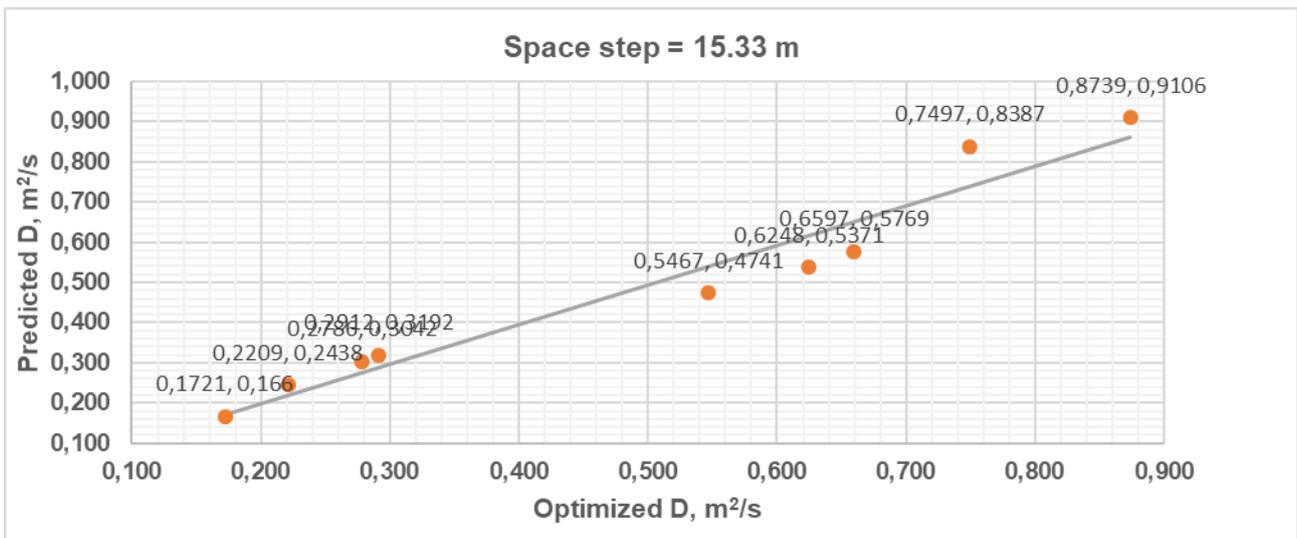
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.976
Predictive model, D	0.976	-
R ²	0.953	
R ² adjusted	0.946	
RMSE	0.062702446	



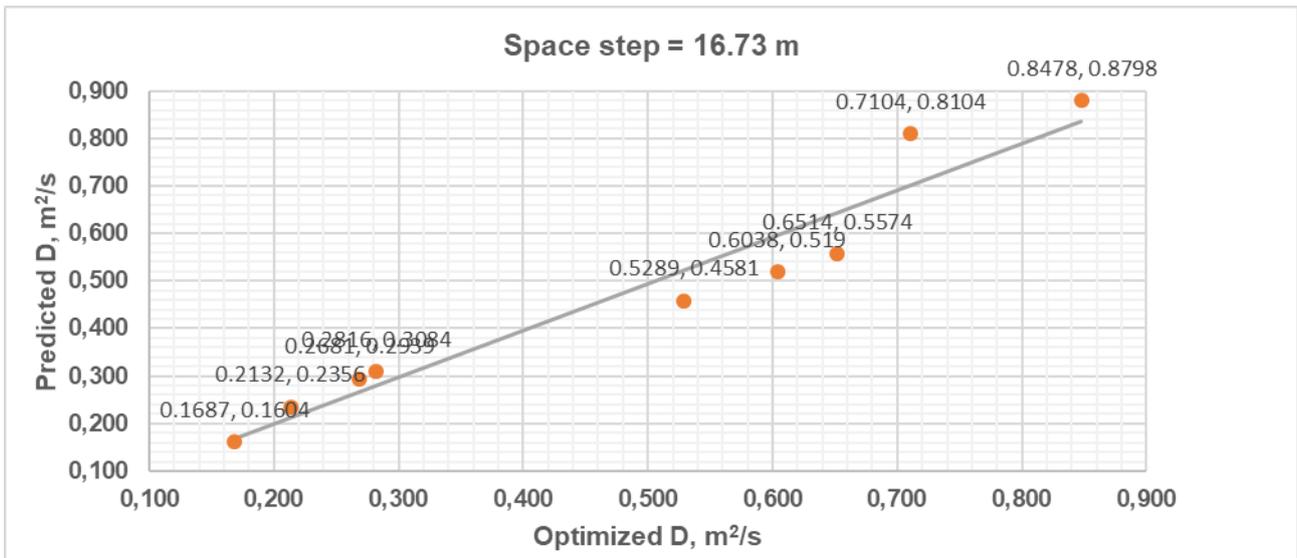
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.975
Predicted model, D	0.975	-
R ²	0.951	
R ² adjusted	0.945	
RMSE	0.063158140	



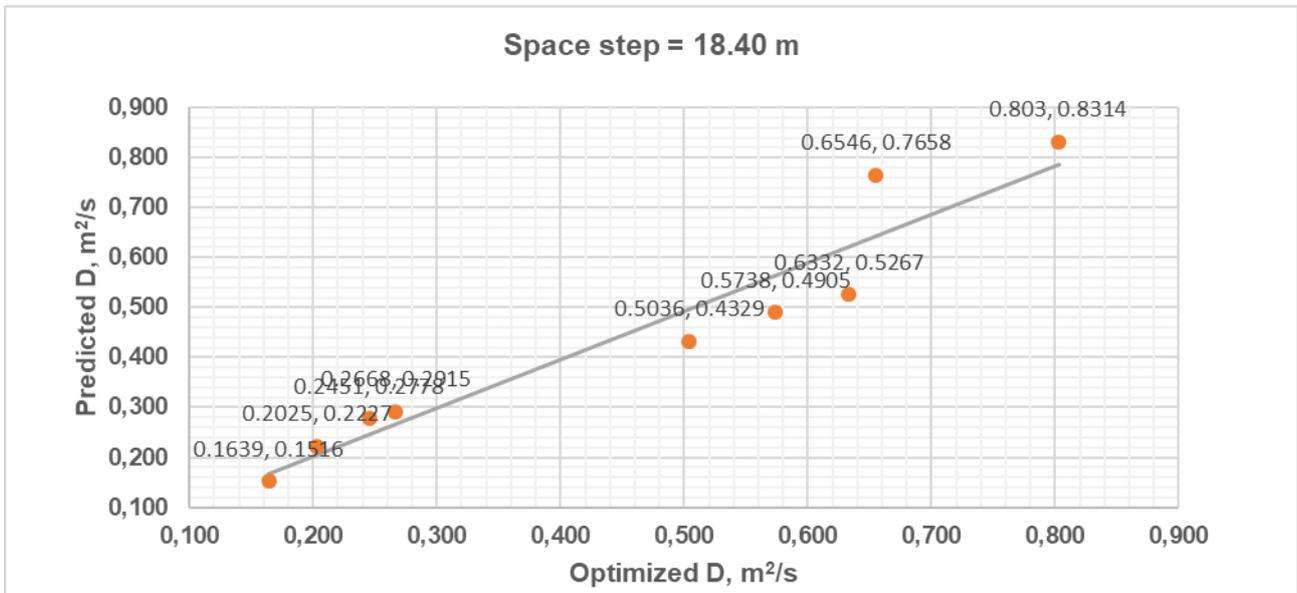
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.974
Predictive model, D	0.974	-
R ²	0.948	
R ² adjusted	0.941	
RMSE	0.064290507	



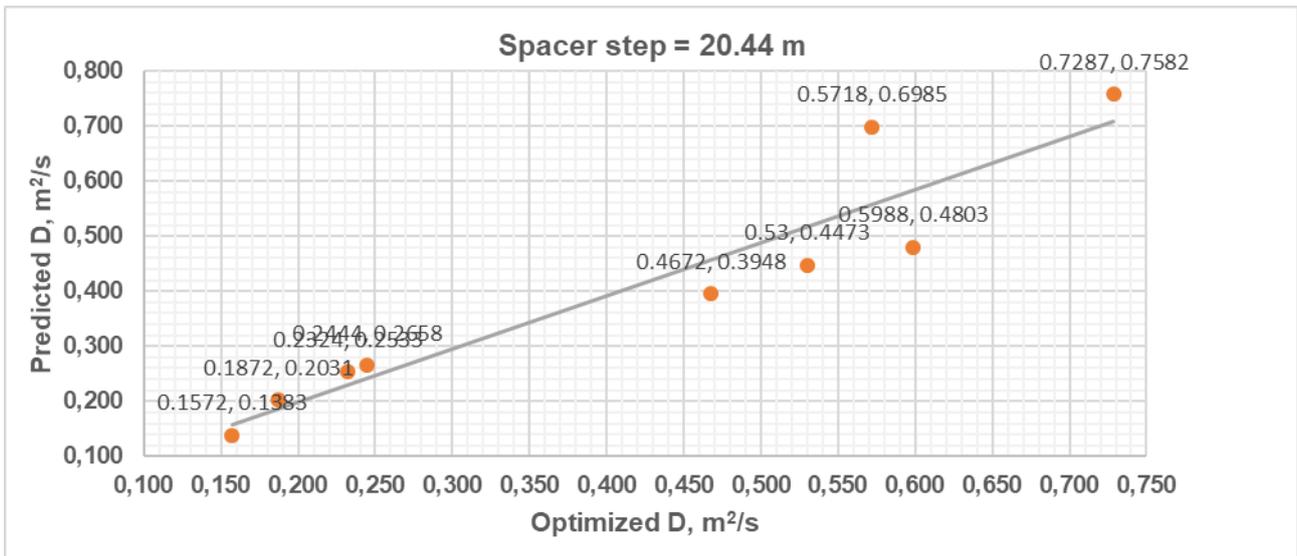
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.971
Predictive model, D	0.971	-
R ²	0.943	
R ² adjusted	0.935	
RMSE	0.066254868	



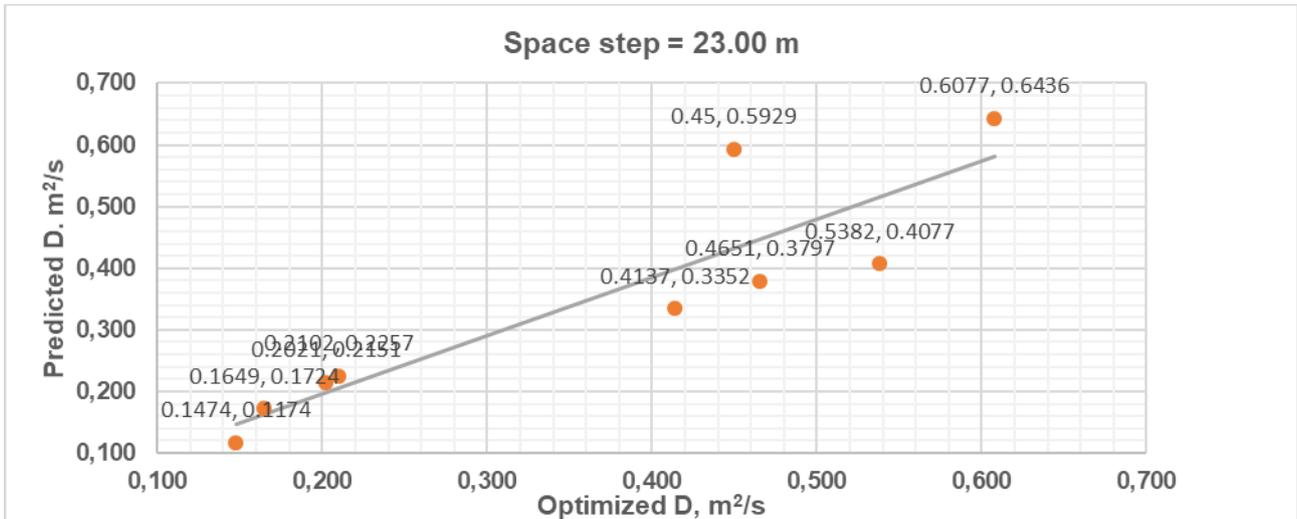
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.966
Predictive model, D	0.966	-
R ²	0.933	
R ² adjusted	0.924	
RMSE	0.06923	



Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.957
Predictive model, D	0.957	-
R ²	0.915	
R ² adjusted	0.903	
RMSE	0.073605980	



Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.940
Predictive model, D	0.940	-
R ²	0.884	
R ² adjusted	0.867	
RMSE	0.078763102	

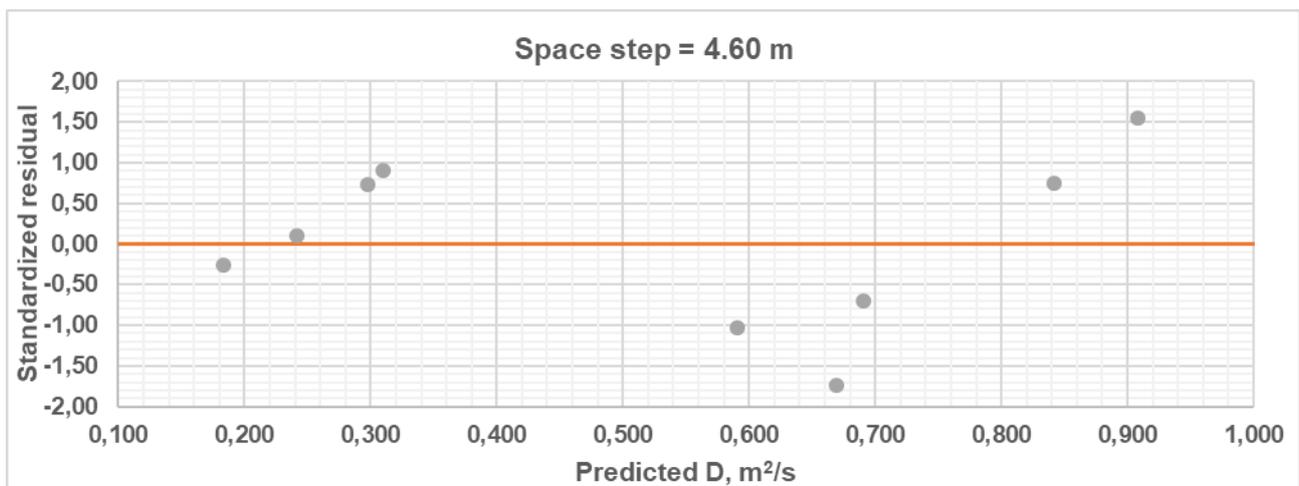
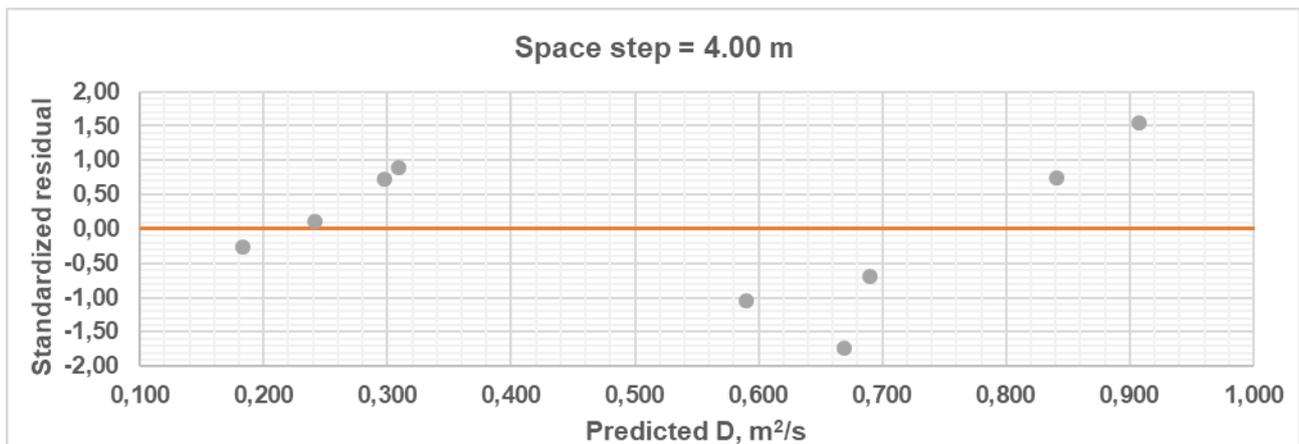
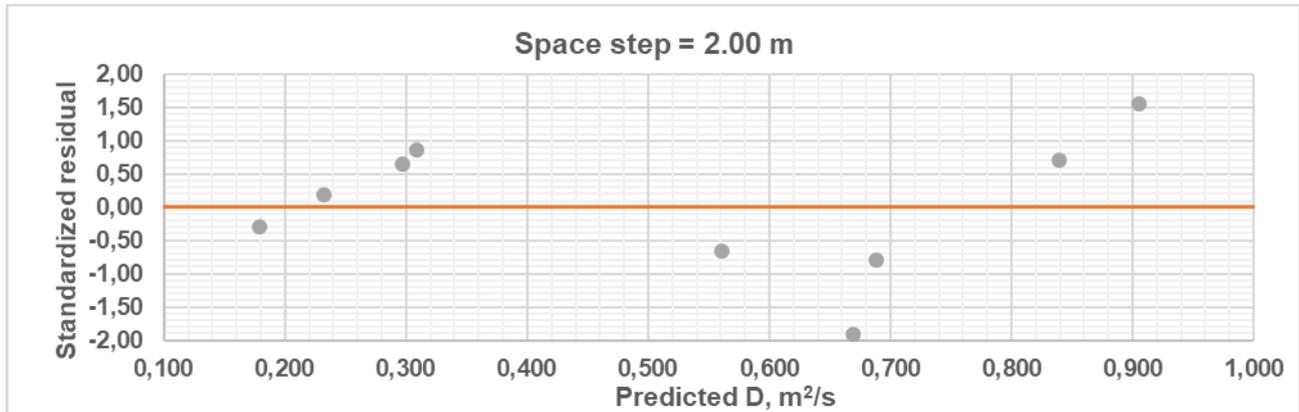


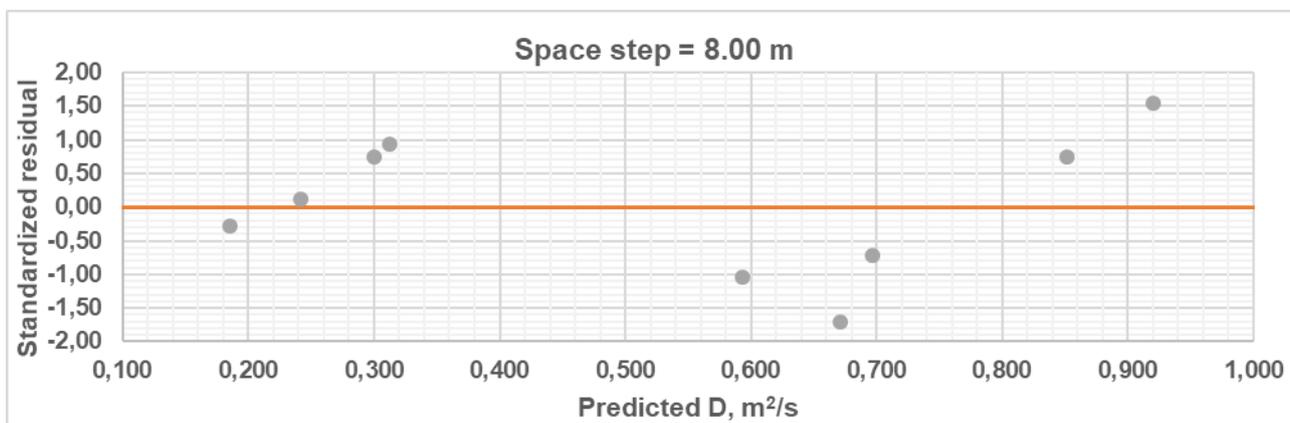
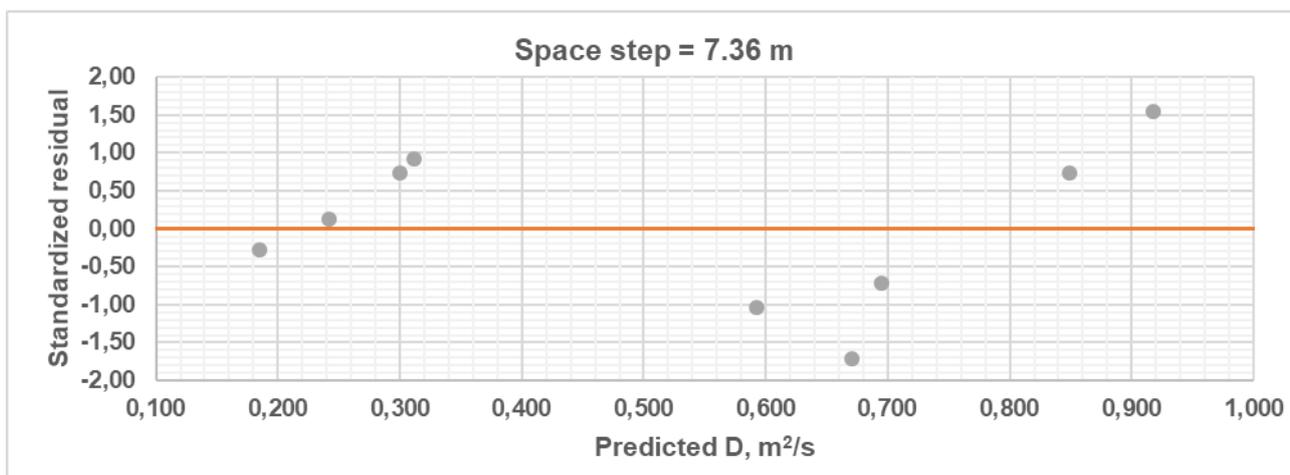
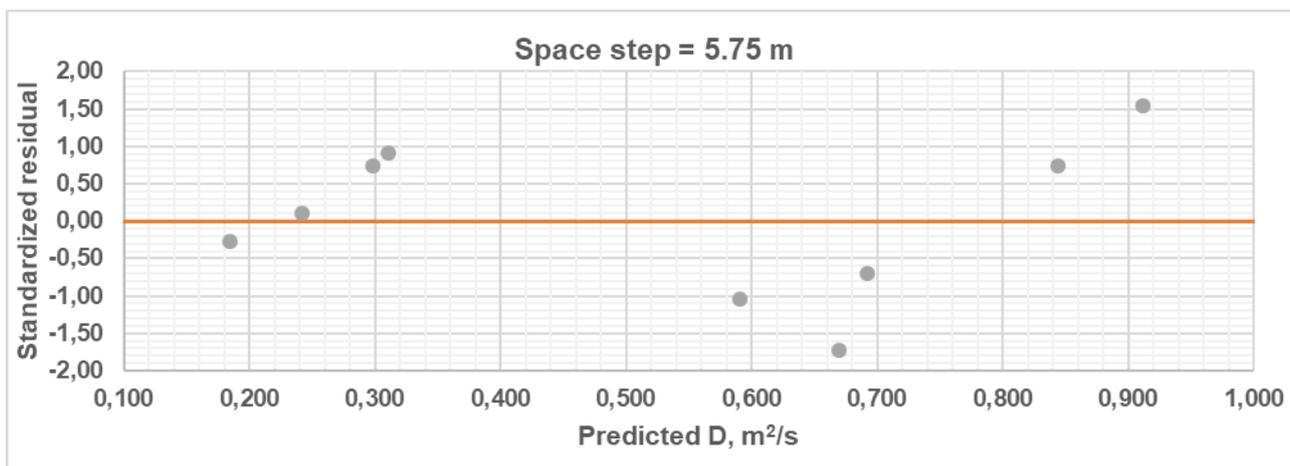
Pearson's r	Numerical model, D	Predictive model, D
Numerical model, D	-	0.899
Predictive model, D	0.899	-
R ²	0.809	
R ² adjusted	0.781	
RMSE	0.085674293	

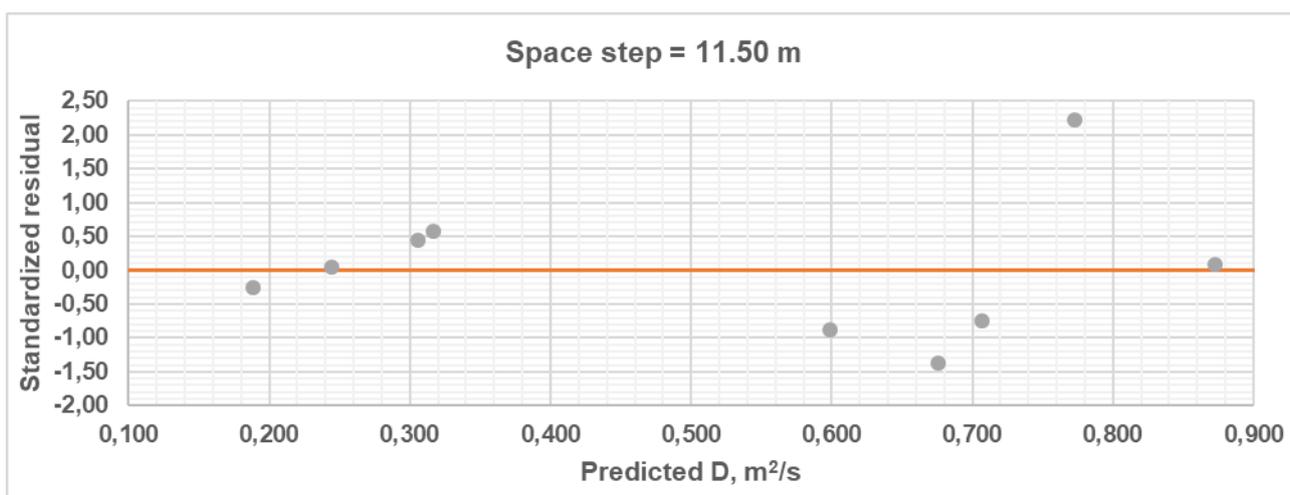
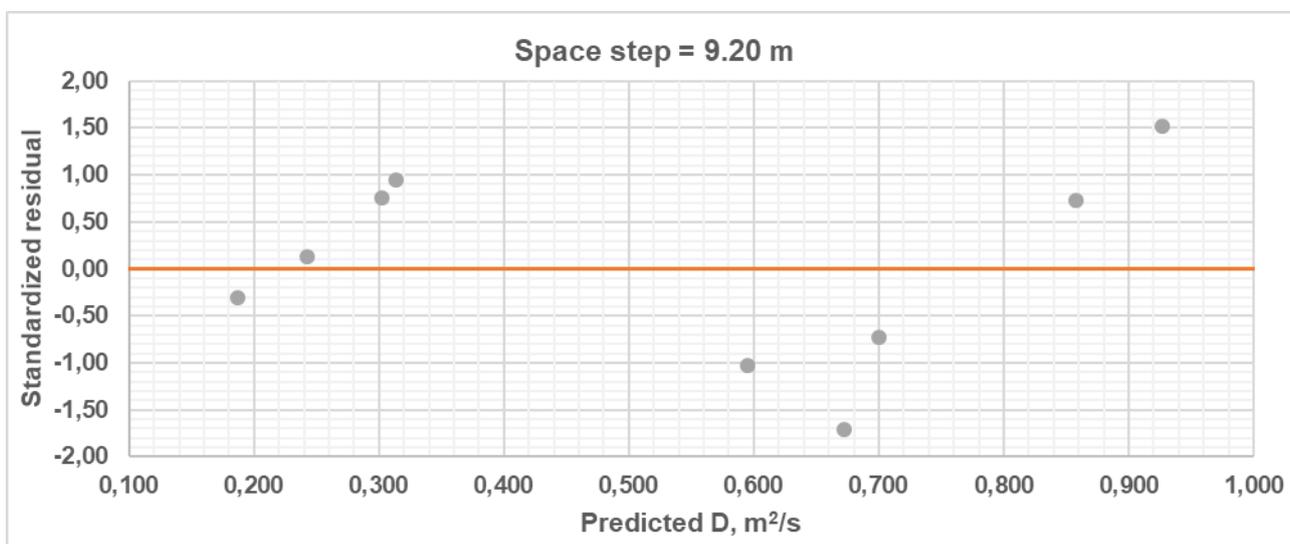
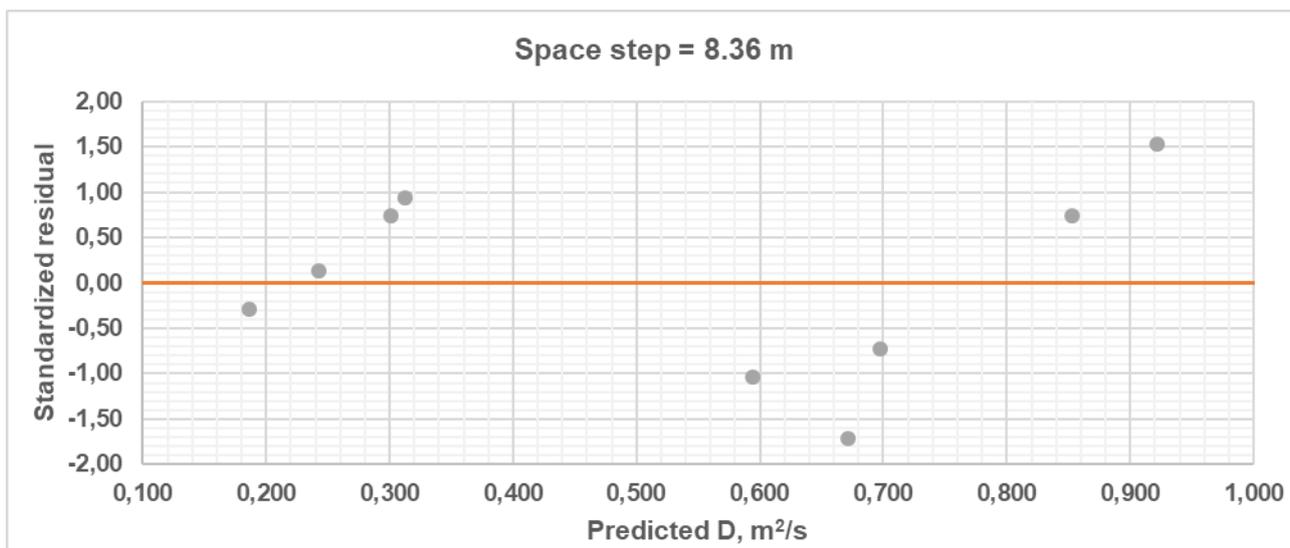
Appendix D: Model adequacy assessment

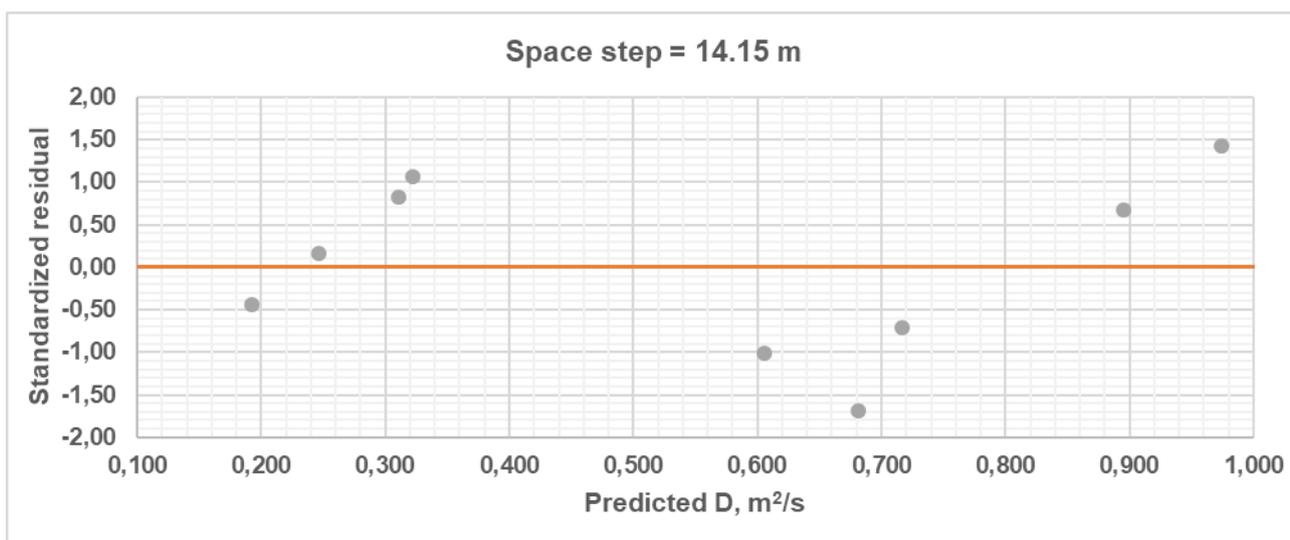
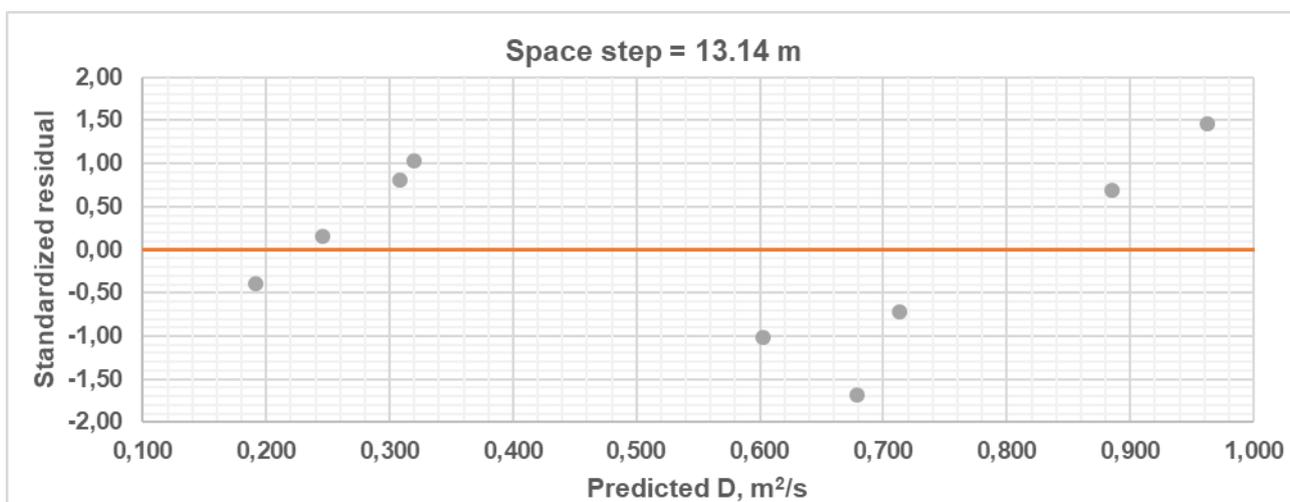
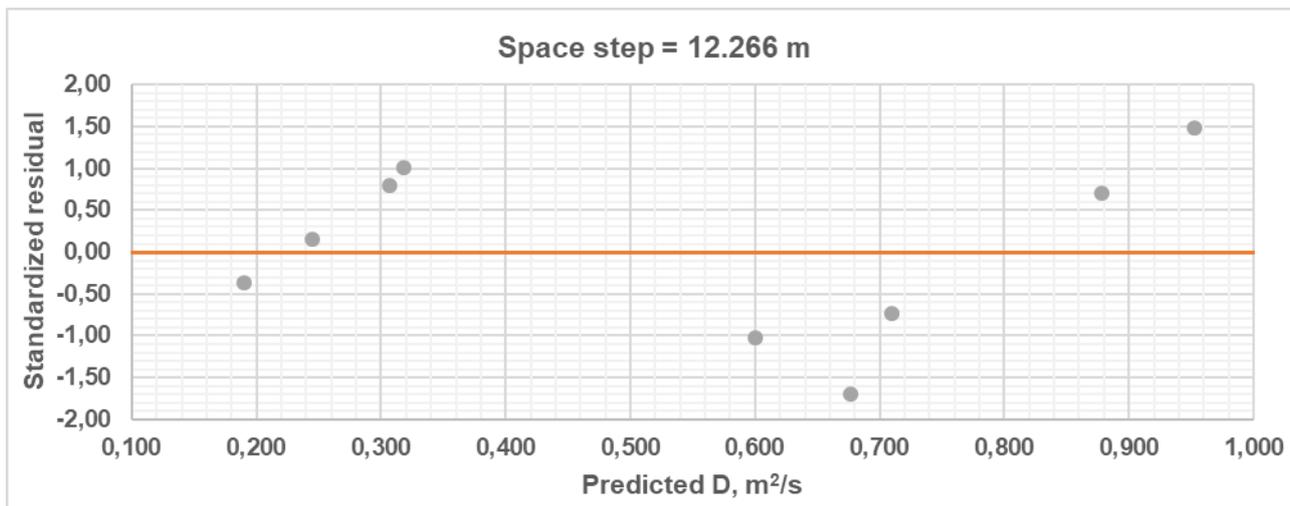
The models were checked for adequacy using plots of normalized residuals versus predicted longitudinal dispersion coefficients.

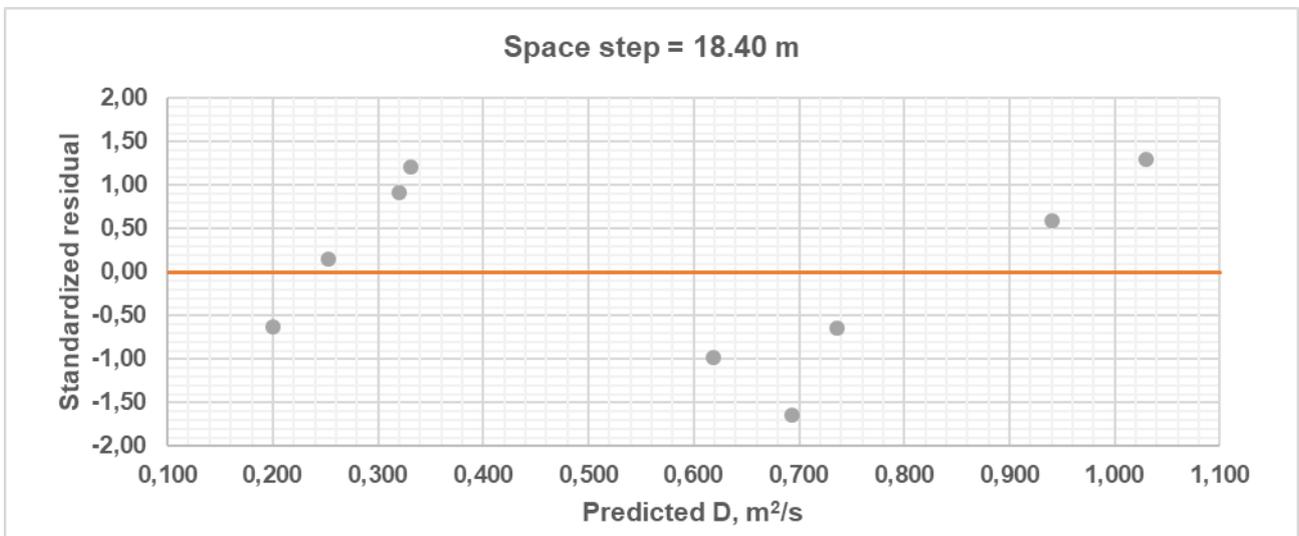
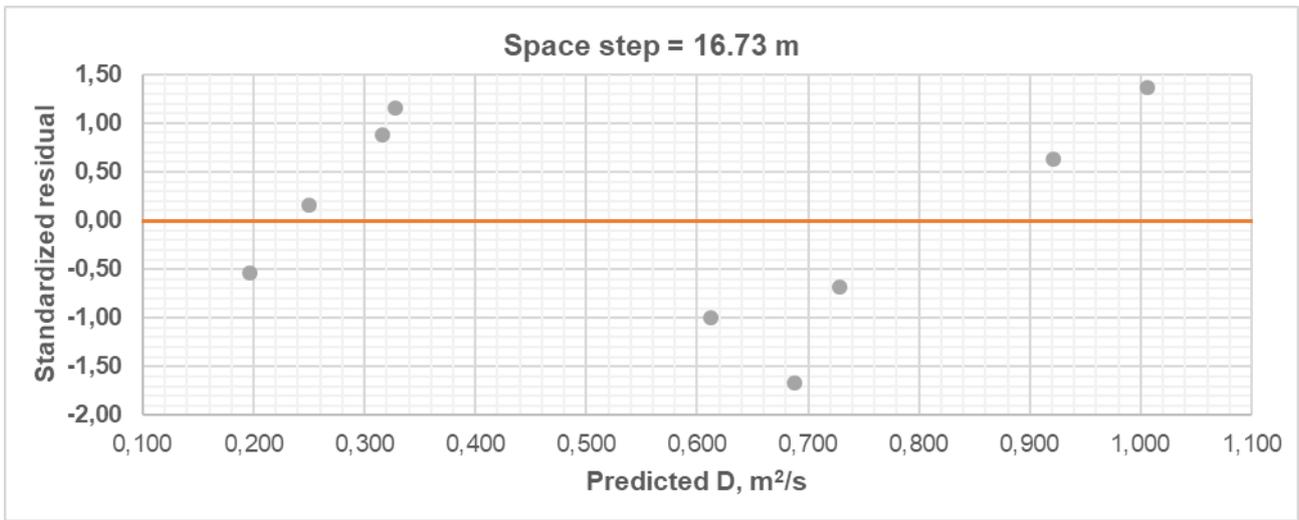
Appendix D-1: Crank-Nicolson based model



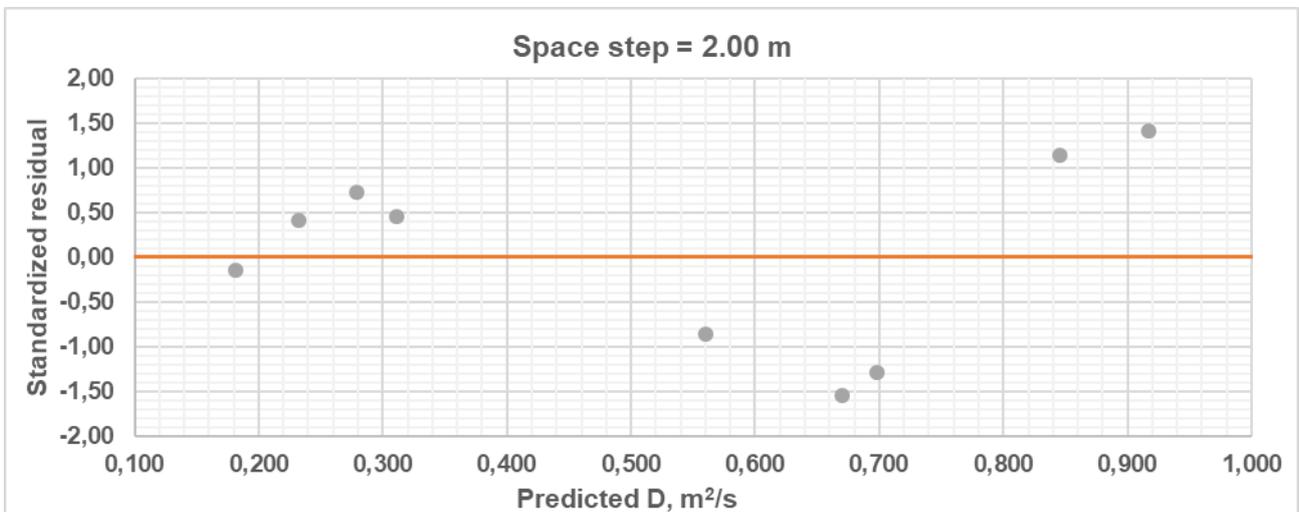


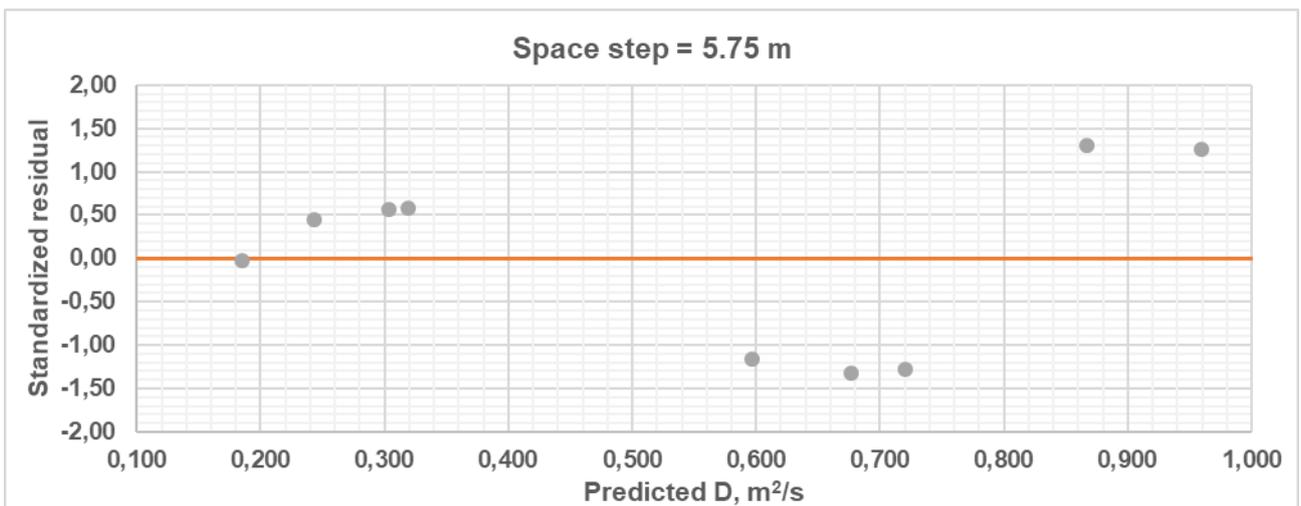
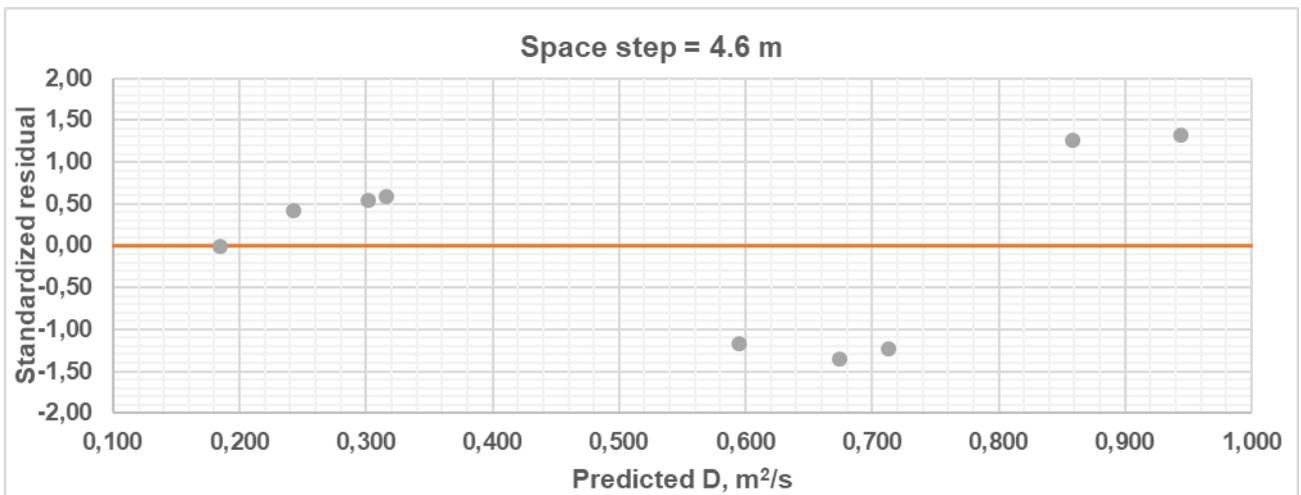
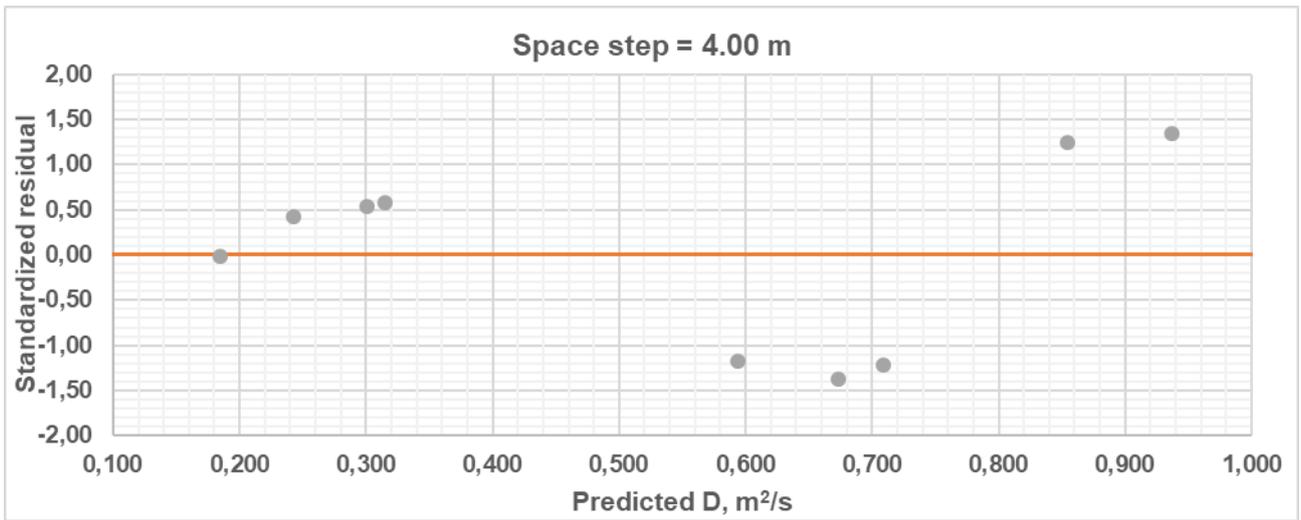


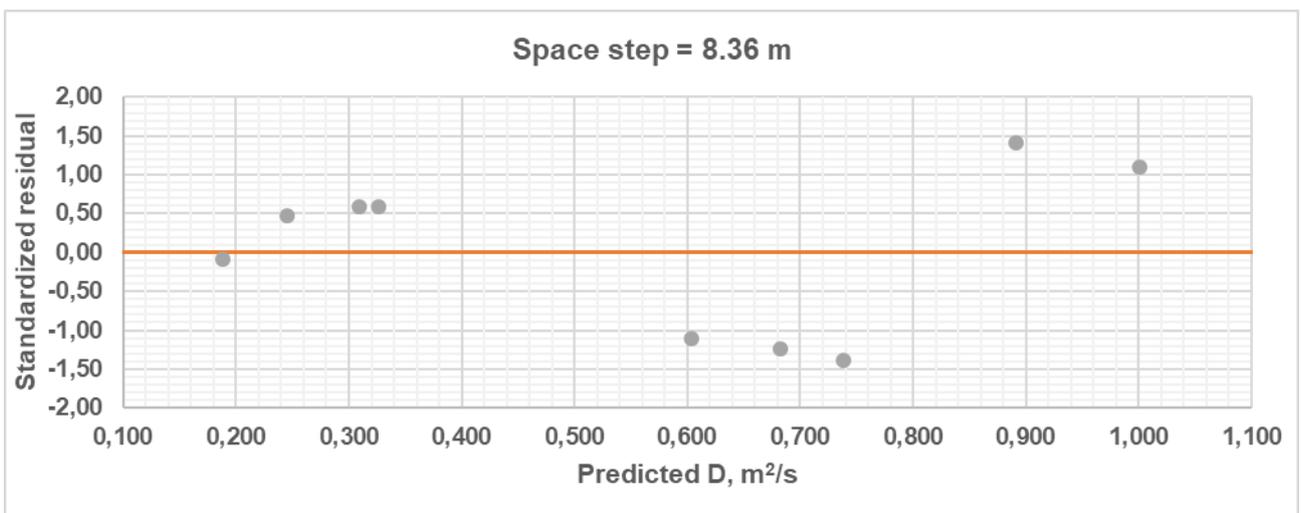
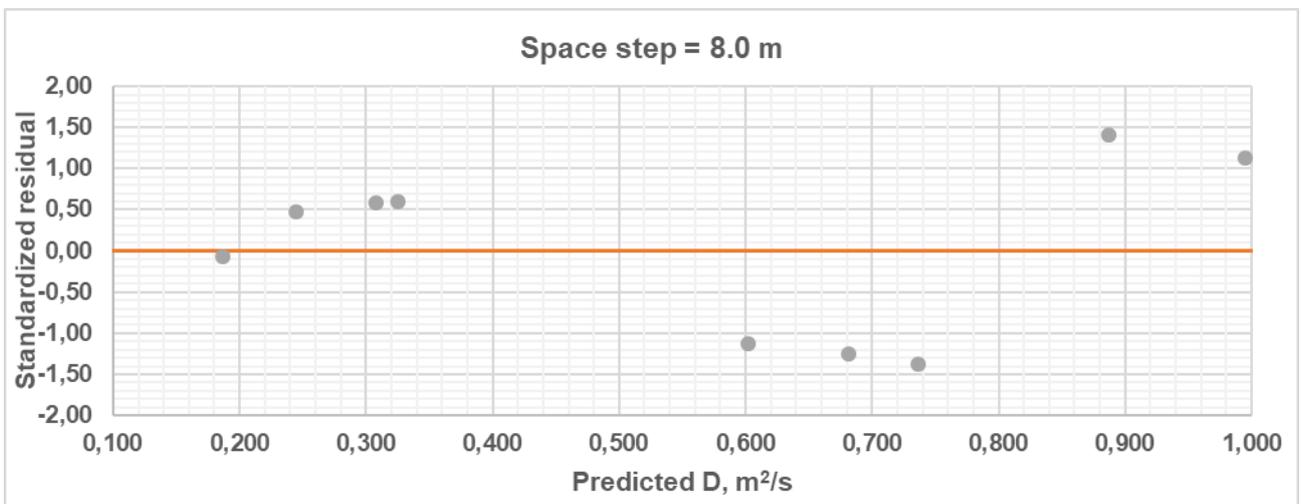
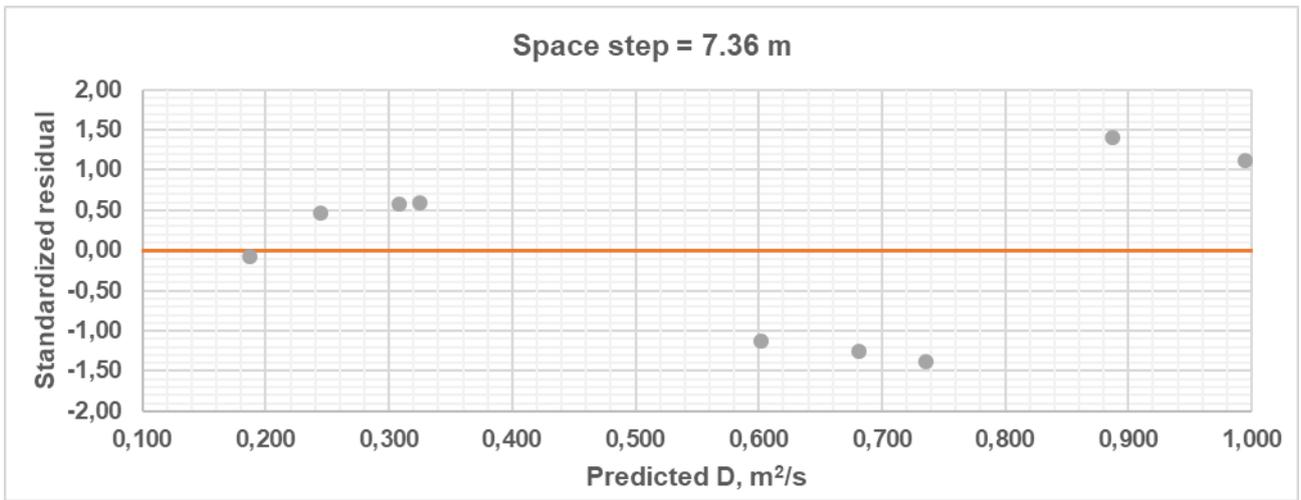


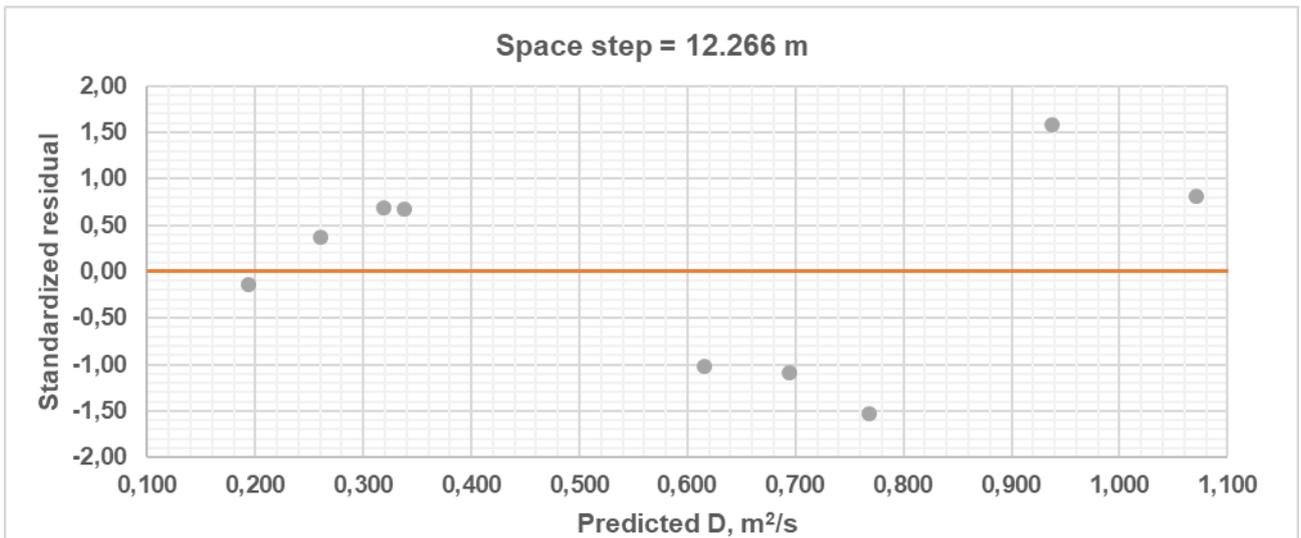
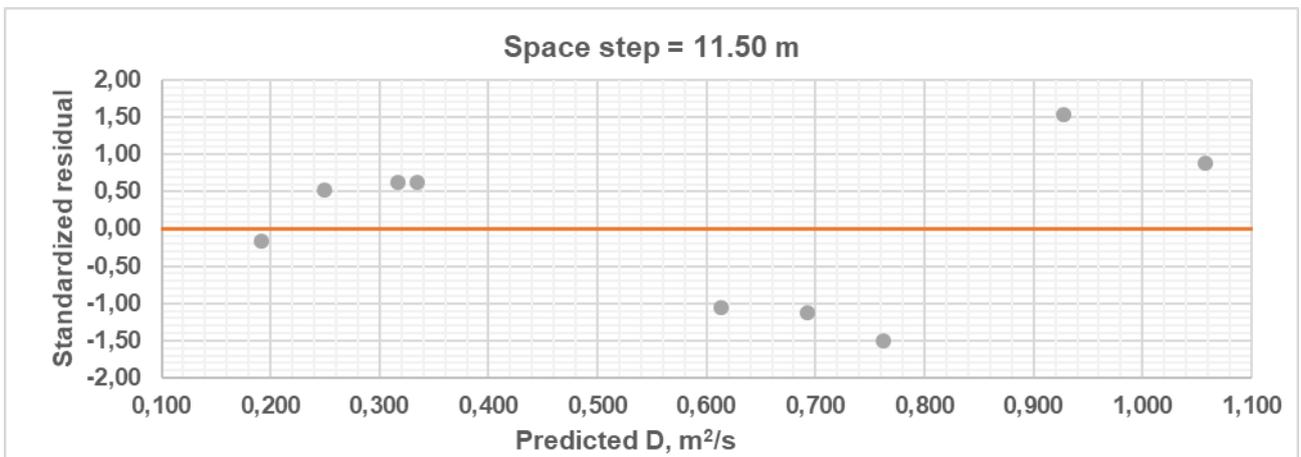
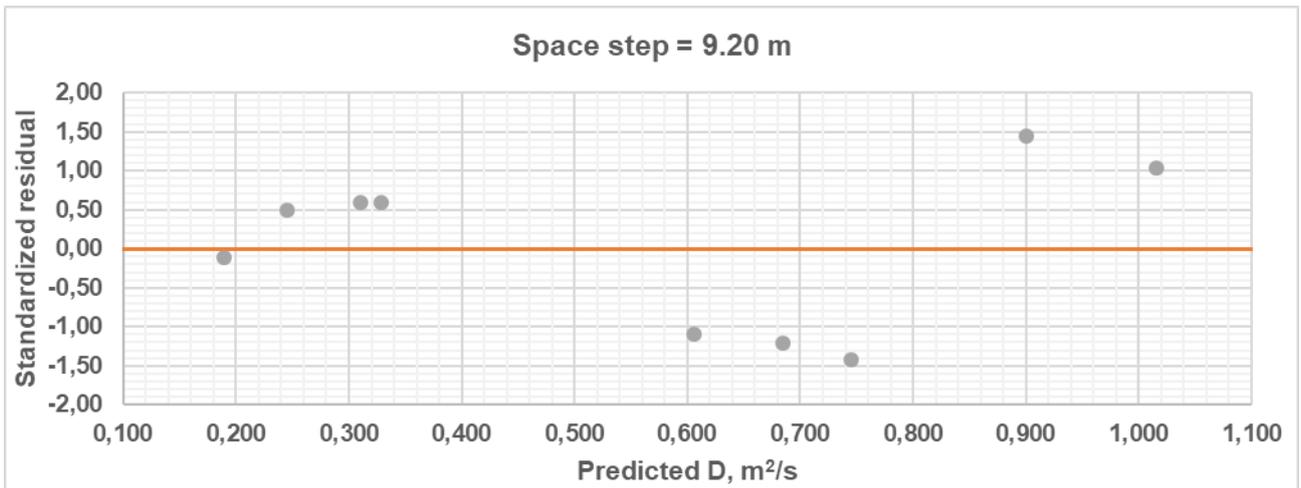


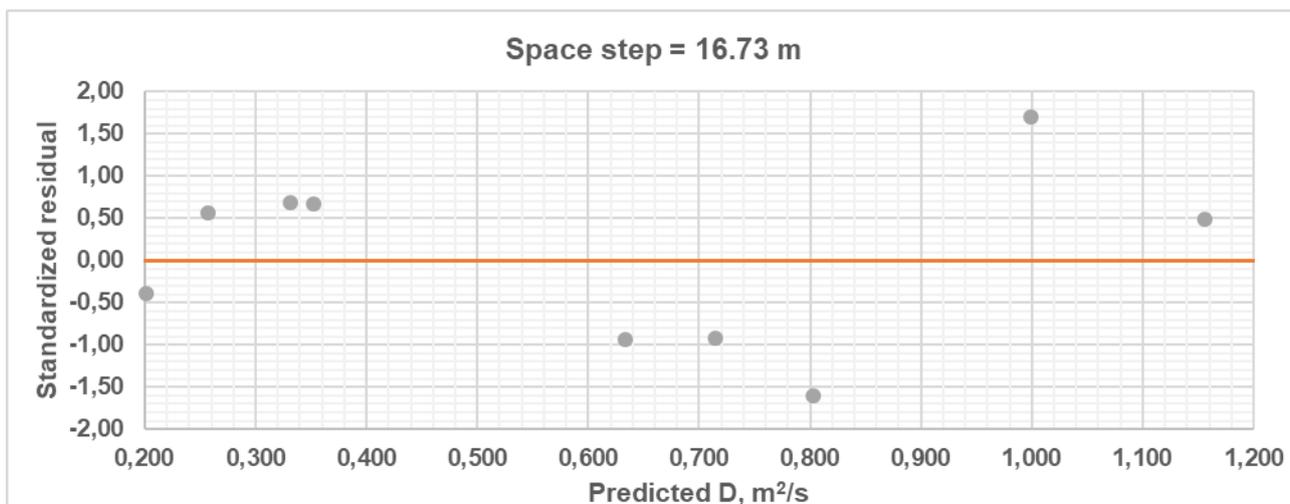
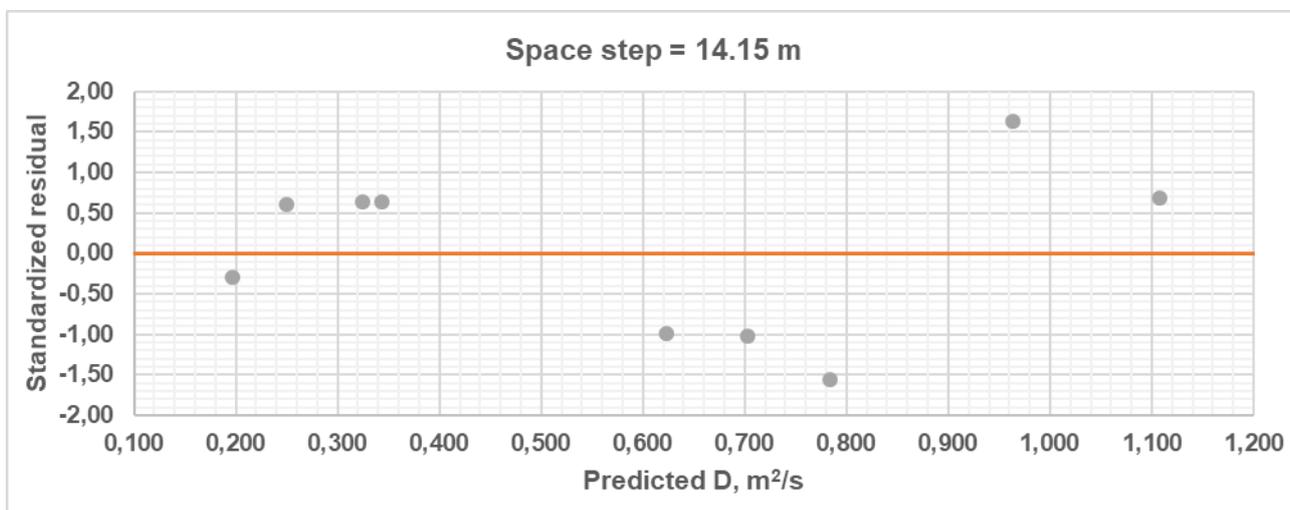
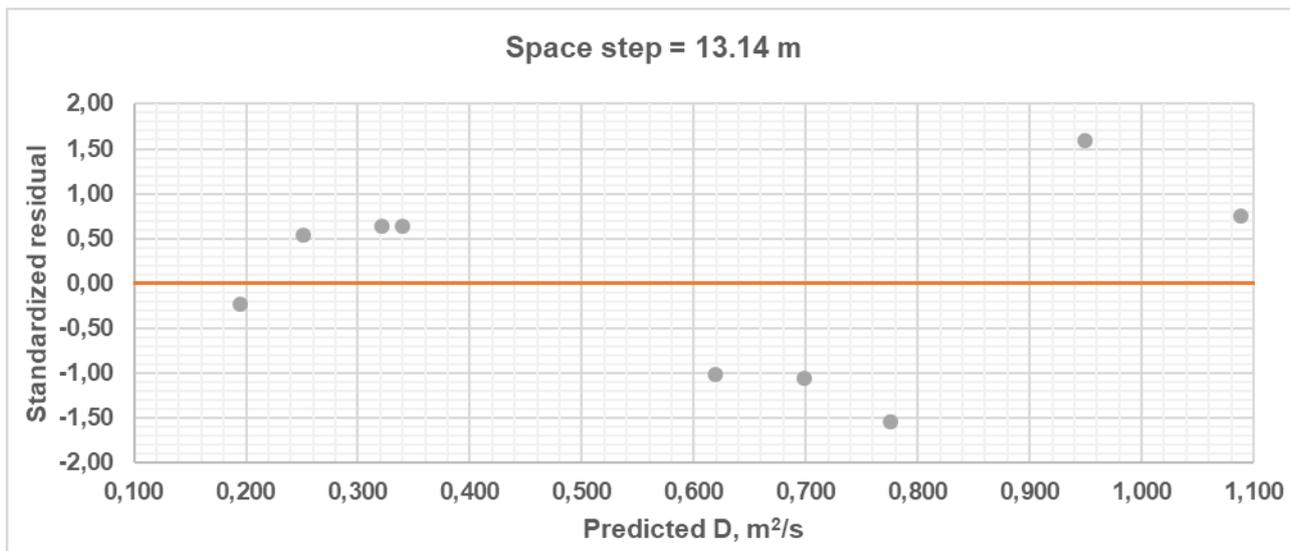
Appendix D-2: MacCormack based model

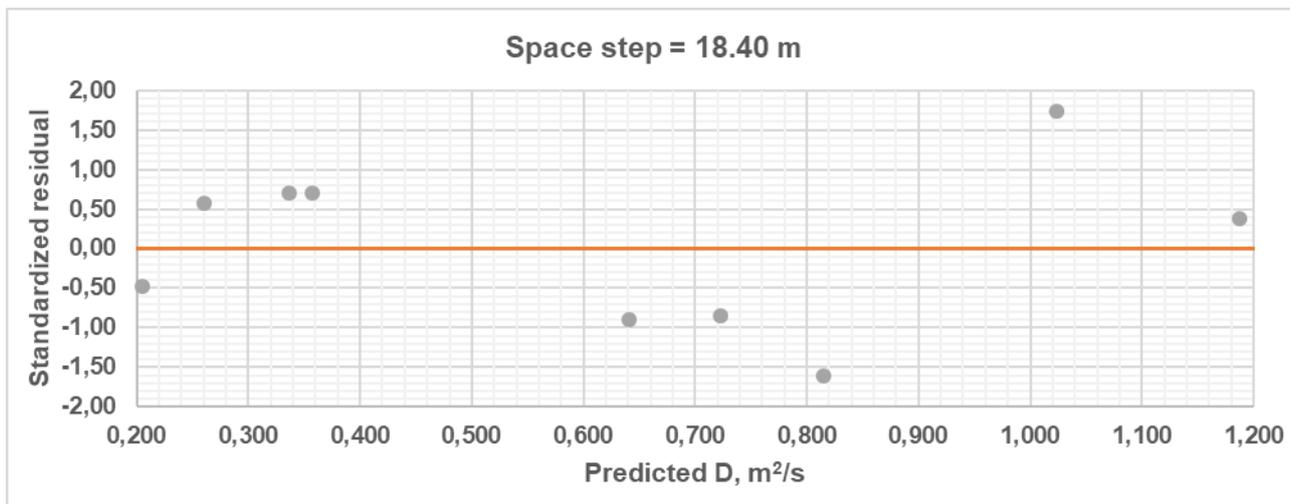




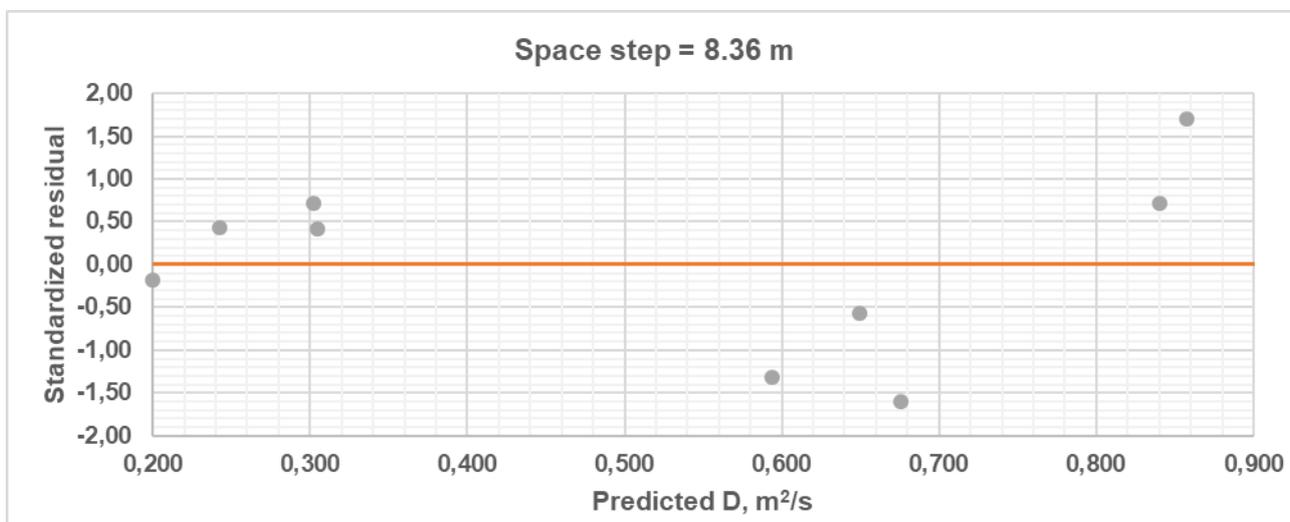
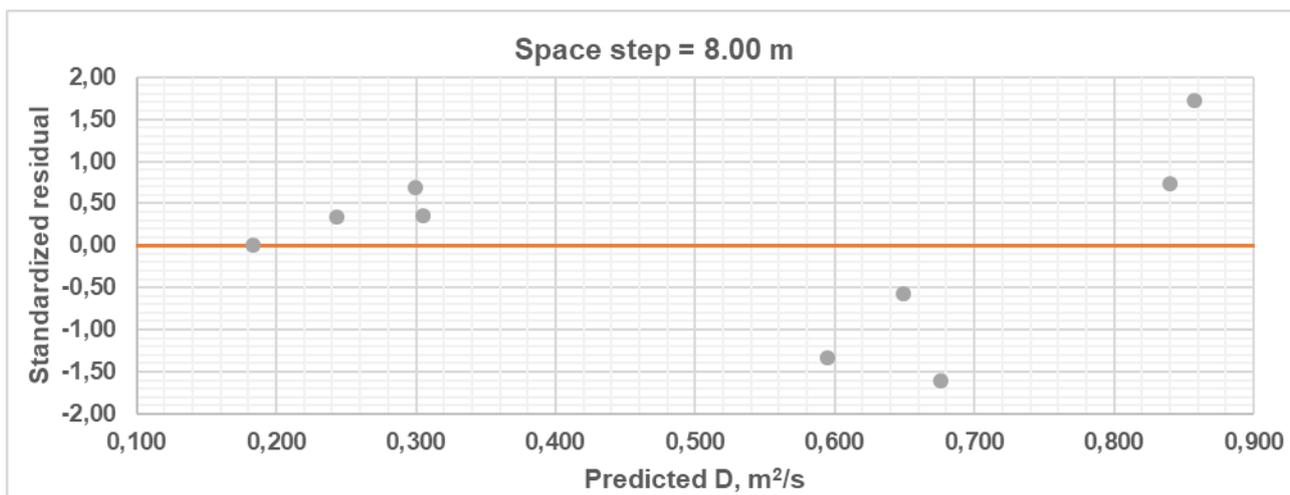


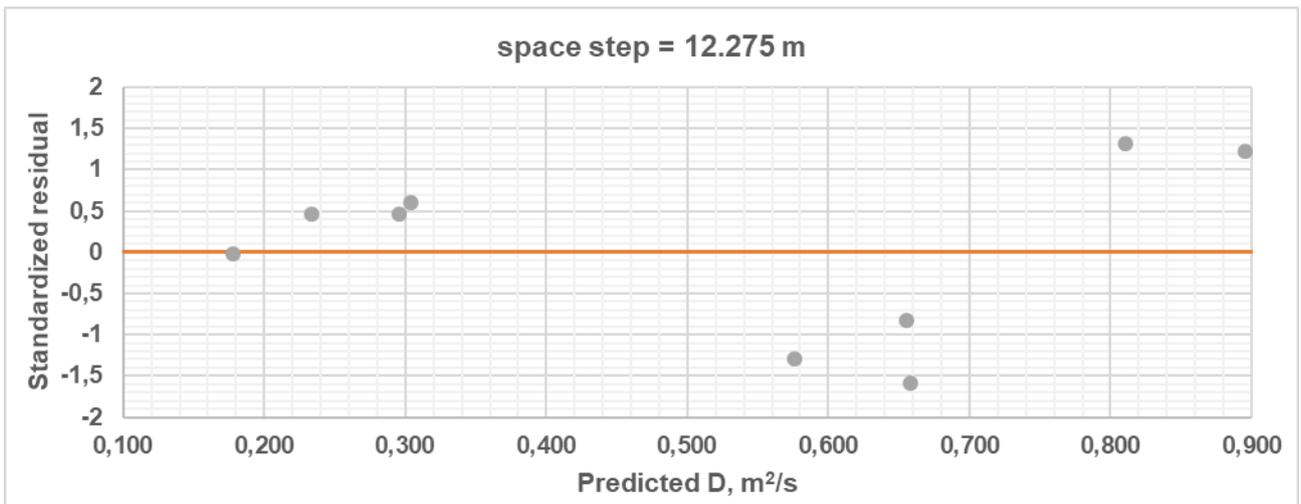
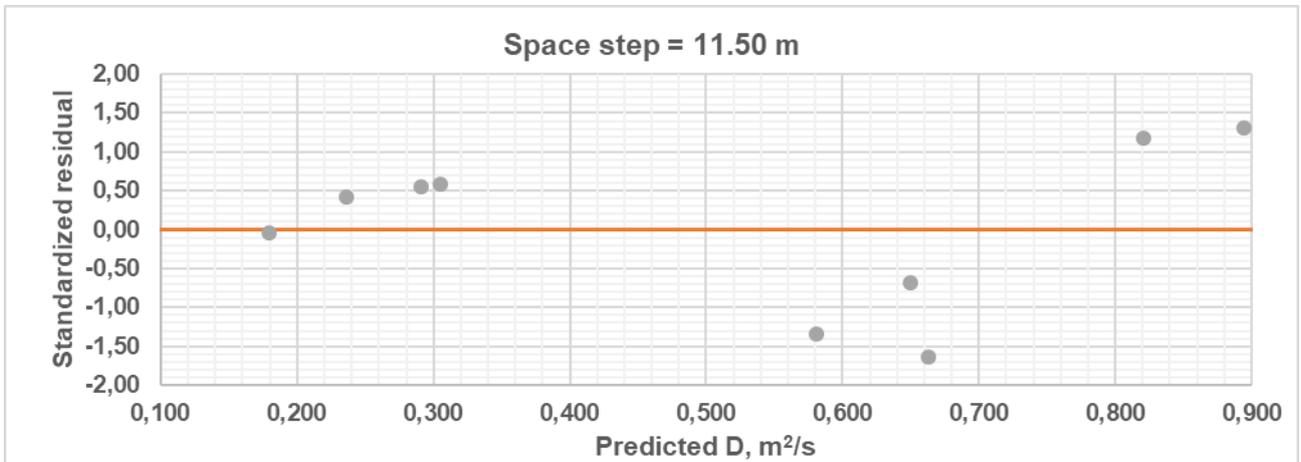
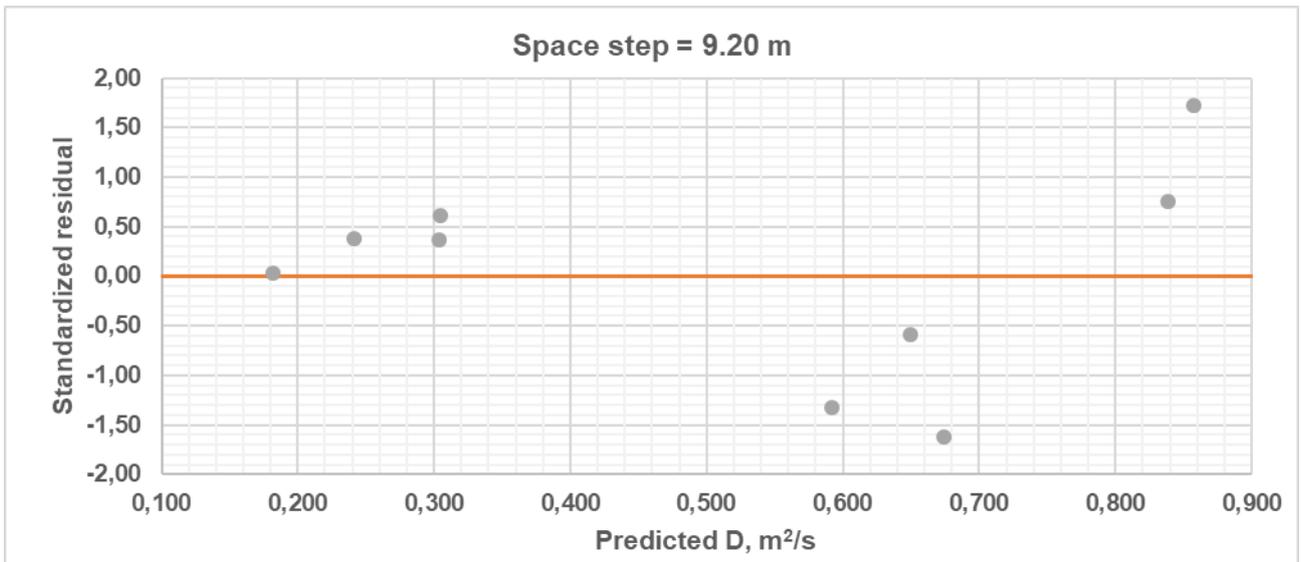


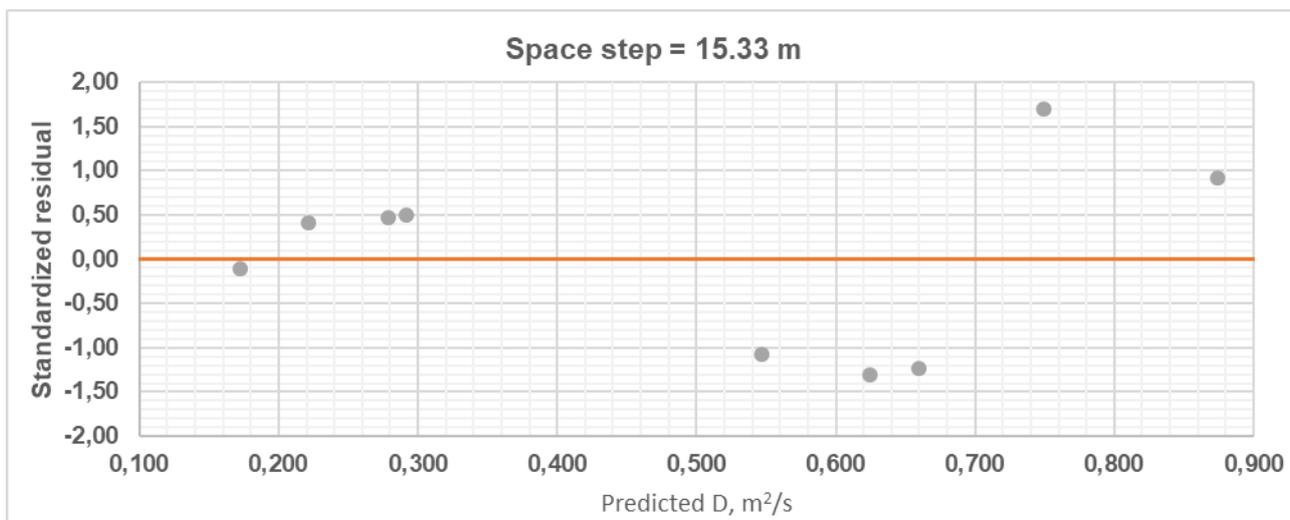
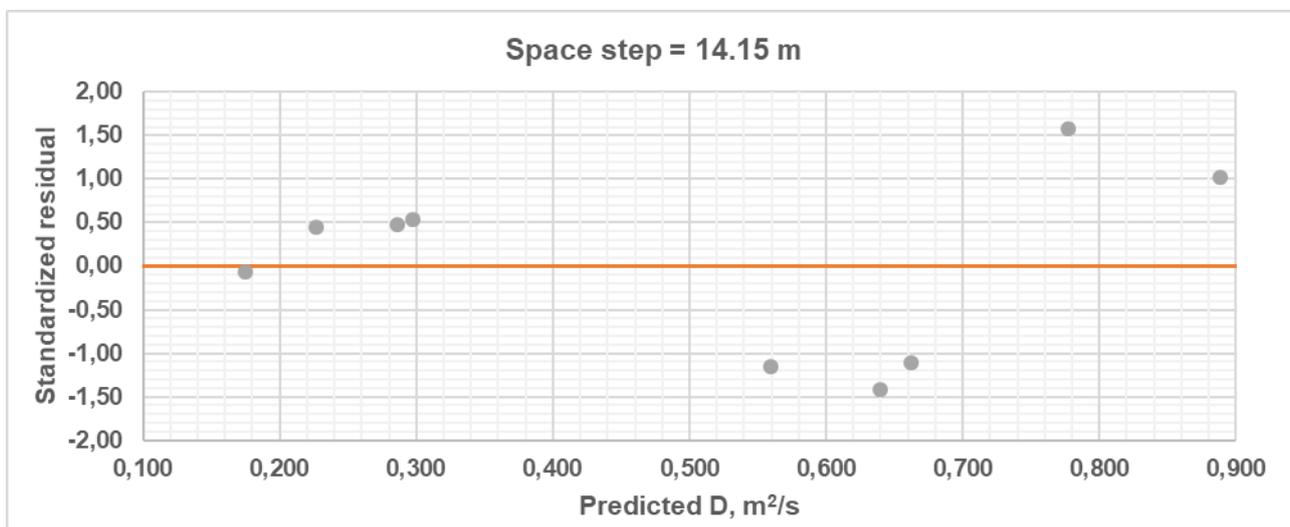
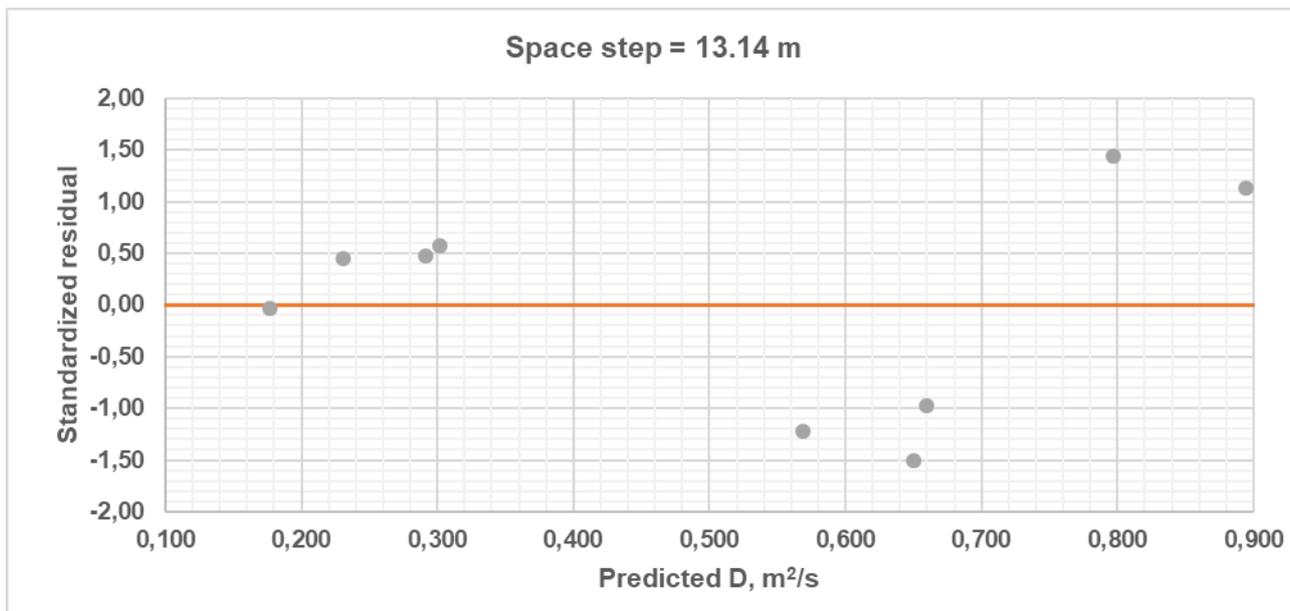


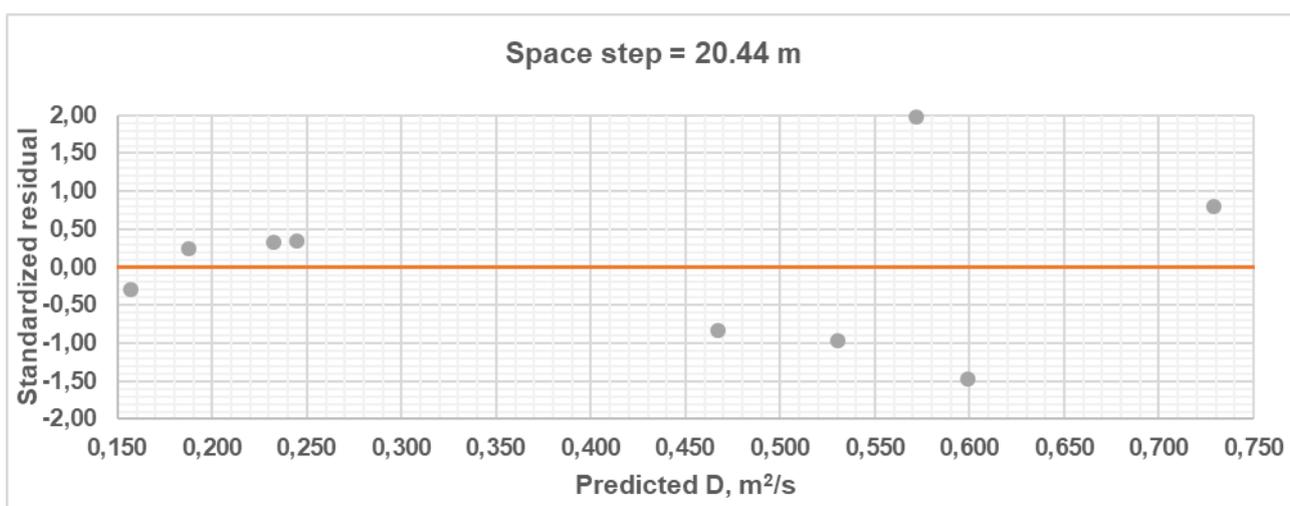
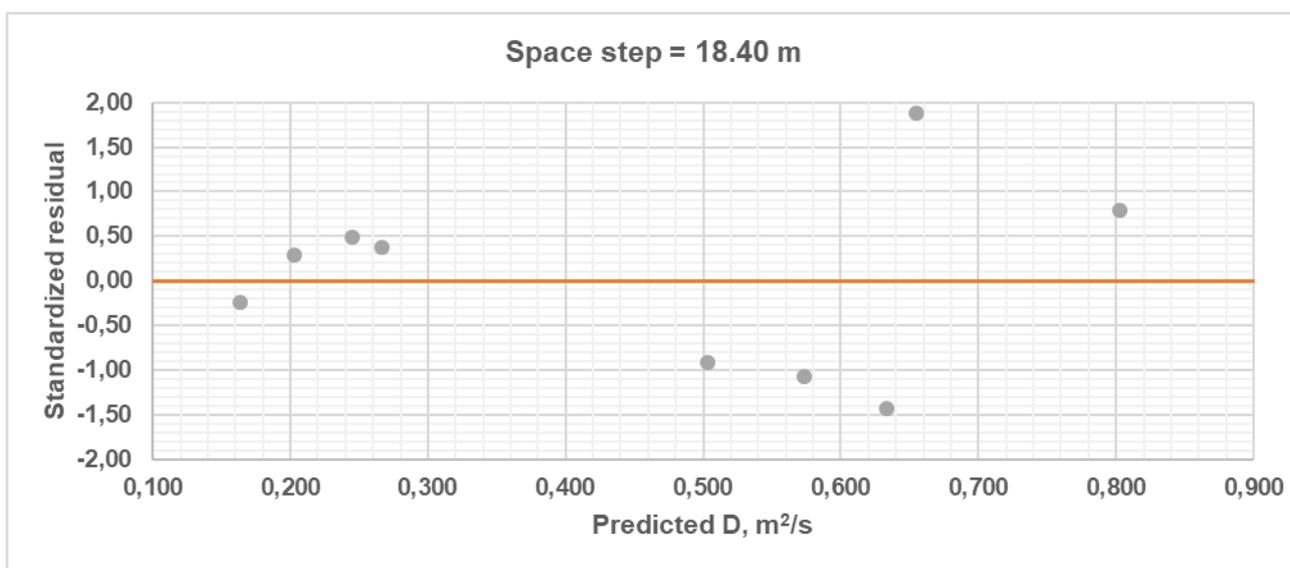
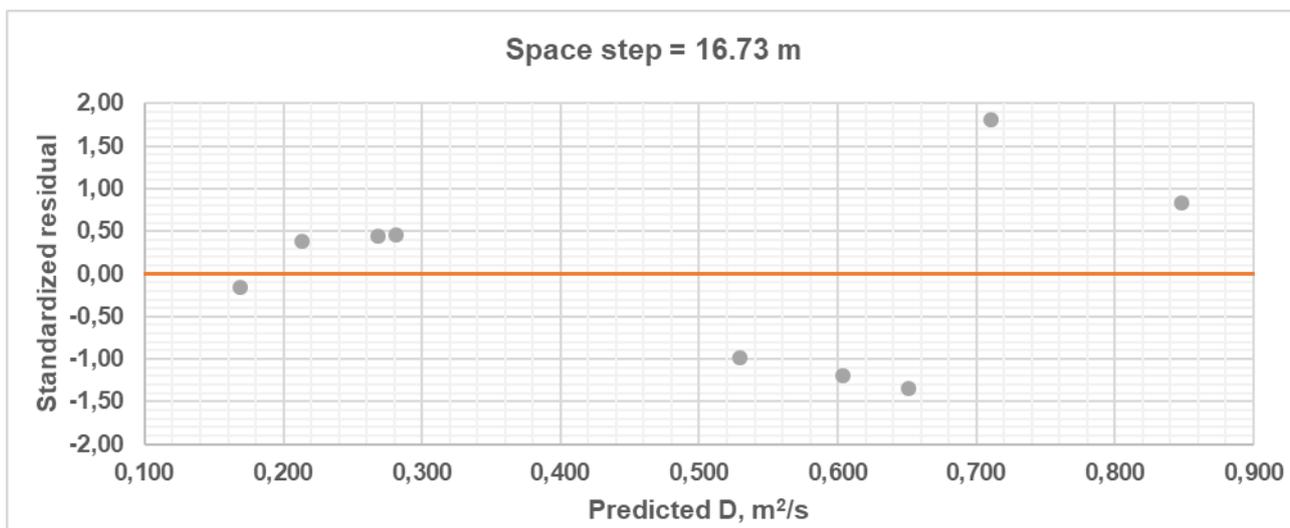


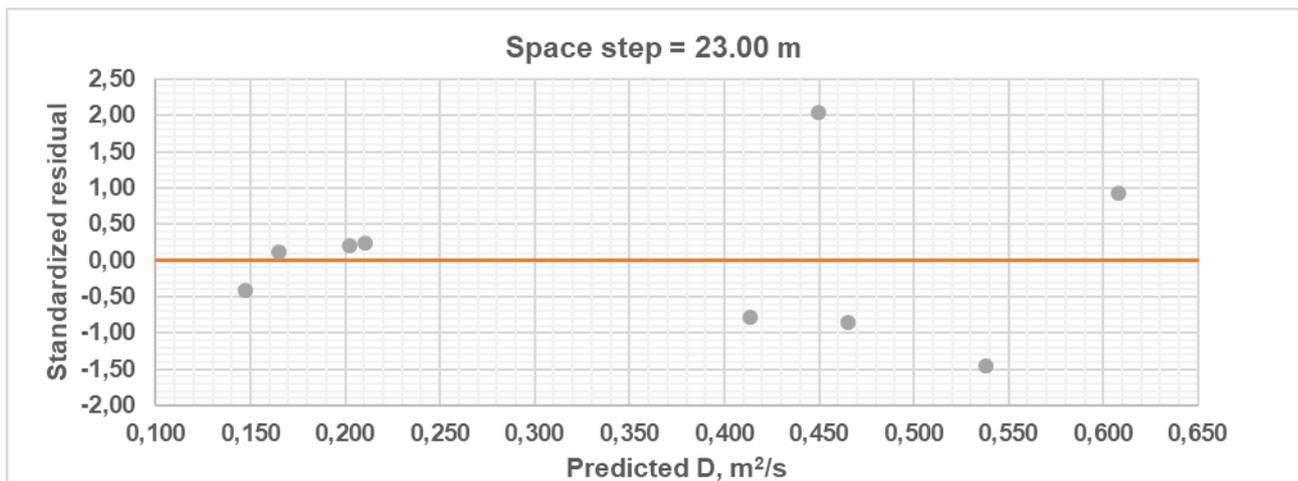
Appendix D-3: QUICKEST based model











Appendix E: Confirmation of numerical methods-based empirical models

Empirical models based on numerical methods were used to predict dispersion coefficients using evaluation data. The models used model (or regression) parameters (α , β & γ) determined during calibration to predict dispersion coefficients. The predicted dispersion coefficient were later used with respective numerical methods to simulate downstream concentration profile. The quality of simulation was measured by residual sum of squares (RSS) and the coefficient of determination (R^2). Evaluation data consisted of experiments 4, 9 and 13.

Appendix E-1: Tables showing predicted dispersion coefficients given by empirical models based on numerical methods.

Peclet number, model parameters and predicted dispersion coefficients for experiment 4 (CN-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
2.00	0.599	0.607	0.233	0.779	0.483
4.00	1.147	0.611	0.233	0.779	0.487
4.60	1.320	0.611	0.233	0.779	0.488
5.75	1.652	0.612	0.233	0.779	0.489
7.36	2.117	0.613	0.233	0.779	0.493
8.00	2.302	0.614	0.233	0.779	0.494
8.36	2.407	0.615	0.233	0.779	0.495
9.20	2.648	0.616	0.233	0.779	0.498
11.50	3.309	0.598	0.233	0.779	0.487
12.27	3.526	0.624	0.233	0.779	0.510
13.14	3.774	0.627	0.233	0.779	0.514
14.15	4.057	0.631	0.233	0.779	0.519
16.73	4.766	0.642	0.233	0.779	0.534
18.40	5.218	0.649	0.233	0.779	0.545

Peclet number, model coefficients and predicted dispersion coefficient for experiment 4 (MacCormack-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
2.00	0.599	0.997	0.729	0.082	0.472
4.00	1.146	1.018	0.729	0.082	0.483
4.60	1.318	1.022	0.729	0.082	0.486
5.75	1.647	1.032	0.729	0.082	0.491
7.36	2.106	1.056	0.729	0.082	0.504
8.00	2.362	1.055	0.729	0.082	0.493
8.36	2.390	1.060	0.729	0.082	0.507
9.20	2.626	1.069	0.729	0.082	0.513
11.50	3.263	1.098	0.729	0.082	0.530
12.27	3.473	1.109	0.729	0.082	0.536
13.14	3.710	1.120	0.729	0.082	0.543
14.15	3.980	1.134	0.729	0.082	0.551
16.73	4.656	1.170	0.729	0.082	0.573
18.40	5.086	1.193	0.729	0.082	0.587

Peclet number, model parameters and predicted dispersion coefficients for experiment 4 (QUICKEST-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
8.00	2.281	0.990	0.727	0.040	0.469
8.36	2.386	0.991	0.727	0.040	0.470
9.20	2.632	0.989	0.727	0.040	0.468
11.50	3.345	0.990	0.727	0.040	0.467
12.27	3.599	0.987	0.727	0.040	0.465
13.14	3.903	0.980	0.727	0.040	0.461
14.15	4.274	0.967	0.727	0.040	0.455
15.33	4.741	0.947	0.727	0.040	0.444
16.73	5.353	0.915	0.727	0.040	0.429
18.40	6.197	0.865	0.727	0.040	0.404
20.44	7.453	0.789	0.727	0.040	0.368
23.00	9.549	0.670	0.727	0.040	0.312

Peclet number, model parameters and predicted dispersion coefficient for experiment 9 (CN-based model)

Δx (m)	Pe	Coefficient α	Exponent β	Exponent γ	Predicted D, m ² /s
2.00	0.628	0.607	0.233	0.779	0.291
4.00	1.254	0.611	0.233	0.779	0.293
4.60	1.441	0.611	0.233	0.779	0.294
5.75	1.799	0.612	0.233	0.779	0.295
7.36	2.296	0.613	0.233	0.779	0.297
8.00	2.492	0.614	0.233	0.779	0.298
8.36	2.603	0.615	0.233	0.779	0.299
9.20	2.857	0.616	0.233	0.779	0.300
11.50	3.544	0.598	0.233	0.779	0.294
12.27	3.770	0.624	0.233	0.779	0.308
13.14	4.025	0.627	0.233	0.779	0.310
14.15	4.317	0.631	0.233	0.779	0.314
16.73	5.046	0.642	0.233	0.779	0.323
18.40	5.510	0.649	0.233	0.779	0.329

Peclet number, model parameters and predicted dispersion coefficients for experiment 9 (MacCormack-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
2.00	0.666	0.997	0.729	0.082	0.296
4.00	1.244	1.018	0.729	0.082	0.303
4.60	1.427	1.022	0.729	0.082	0.304
5.75	1.777	1.032	0.729	0.082	0.308
7.36	2.258	1.056	0.729	0.082	0.316
8.00	2.447	1.055	0.729	0.082	0.317
8.36	2.554	1.060	0.729	0.082	0.318
9.20	2.797	1.069	0.729	0.082	0.322
11.50	3.455	1.098	0.729	0.082	0.333
12.27	3.668	1.109	0.729	0.082	0.337
13.14	3.911	1.120	0.729	0.082	0.341
14.15	4.188	1.134	0.729	0.082	0.346
16.73	4.879	1.170	0.729	0.082	0.360
18.40	5.318	1.193	0.729	0.082	0.369

Peclet number, model parameters and predicted dispersion coefficients for experiment 9 (QUICKEST-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
8.00	2.51	0.990	0.727	0.040	0.300
8.36	2.63	0.991	0.727	0.040	0.300
9.20	2.90	0.989	0.727	0.040	0.299
11.50	3.70	0.990	0.727	0.040	0.298
12.27	3.99	0.987	0.727	0.040	0.297
13.14	4.34	0.980	0.727	0.040	0.294
14.15	4.77	0.967	0.727	0.040	0.290
15.33	5.32	0.947	0.727	0.040	0.284
16.73	6.06	0.915	0.727	0.040	0.274
18.40	7.10	0.865	0.727	0.040	0.258
20.44	8.70	0.789	0.727	0.040	0.235
23.00	11.52	0.670	0.727	0.040	0.199

Peclet number, model parameters and predicted dispersion coefficients for experiment 13 (CN-based model)

Δx (m)	Pe	Coefficient α	Exponent β	Exponent γ	Predicted D, m ² /s
2.00	0.752	0.607	0.233	0.779	0.664
4.00	1.505	0.611	0.233	0.779	0.670
4.60	1.731	0.611	0.233	0.779	0.670
5.75	2.164	0.612	0.233	0.779	0.672
7.36	2.768	0.613	0.233	0.779	0.676
8.00	3.008	0.614	0.233	0.779	0.678
8.36	3.144	0.615	0.233	0.779	0.680
9.20	3.455	0.616	0.233	0.779	0.683
11.50	4.467	0.598	0.233	0.779	0.670
12.27	4.578	0.624	0.233	0.779	0.699
13.14	4.891	0.627	0.233	0.779	0.705
14.15	5.248	0.631	0.233	0.779	0.712
16.73	6.129	0.642	0.233	0.779	0.732
18.40	6.680	0.649	0.233	0.779	0.747

Peclet number, model parameters and predicted dispersion coefficient for experiment 13 (MacCormack-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m ² /s
2.00	0.749	0.997	0.729	0.082	0.641
4.00	1.495	1.018	0.729	0.082	0.656
4.60	1.716	1.022	0.729	0.082	0.660
5.75	2.136	1.032	0.729	0.082	0.668
7.36	2.716	1.056	0.729	0.082	0.686
8.00	2.942	1.055	0.729	0.082	0.687
8.36	3.070	1.060	0.729	0.082	0.690
9.20	3.361	1.069	0.729	0.082	0.698
11.50	4.137	1.098	0.729	0.082	0.722
12.27	4.387	1.109	0.729	0.082	0.730
13.14	4.669	1.120	0.729	0.082	0.739
14.15	4.988	1.134	0.729	0.082	0.751
16.73	5.771	1.170	0.729	0.082	0.781
18.40	6.260	1.193	0.729	0.082	0.801

Peclet number, model parameters and predicted dispersion coefficient for experiment 13 (QUICKEST-based model)

Δx (m)	Pe	Coefficient, α	Exponent, β	Exponent, γ	Predicted D, m^2/s
8.00	3.066	0.990	0.727	0.040	0.633
8.36	3.189	0.991	0.727	0.040	0.633
9.20	3.480	0.989	0.727	0.040	0.630
11.50	4.340	0.990	0.727	0.040	0.627
12.27	4.648	0.987	0.727	0.040	0.624
13.14	5.019	0.980	0.727	0.040	0.619
14.15	5.477	0.967	0.727	0.040	0.610
15.33	6.061	0.947	0.727	0.040	0.596
16.73	6.841	0.915	0.727	0.040	0.574
18.40	7.951	0.865	0.727	0.040	0.542
20.44	9.684	0.789	0.727	0.040	0.493
23.00	12.818	0.670	0.727	0.040	0.417

Flow rates, model parameters and predicted dispersion coefficient (Routing method-based model)

Experiment number	Flow rate (m^3/s)	Coefficient α	Exponent β	Exponent γ	Predicted D (m^2/s)
9	0.037	0.509	0.117	0.950	0.266
4	0.084	0.509	0.117	0.950	0.448
13	0.148	0.509	0.117	0.950	0.606

Appendix E-2: Residual sum of squares between predicted and observed concentration profiles. Numerical methods were applied with predicted dispersion coefficients to predict concentration profiles. The tables show space steps, predicted dispersion coefficients, nondimensional parameters and residual sum of squares (RSS).

Crank-Nicolson based model (experiment 4)

Δx , meters	Δt , seconds	D, m^2/s	V, m/s	c	d	Pe	RSS, $(\mu g/l)^2$
2.00	30	0.483	0.149	2.233	3.626	0.616	0.04354
4.00	30	0.487	0.149	1.118	0.913	1.224	0.04741
4.60	30	0.488	0.149	0.972	0.692	1.406	0.04929
5.75	30	0.489	0.149	0.779	0.444	1.756	0.05420
7.36	30	0.493	0.150	0.610	0.273	2.236	0.06334
8.00	30	0.494	0.150	0.562	0.232	2.428	0.06813
8.36	30	0.495	0.150	0.538	0.212	2.536	0.07107
9.20	30	0.498	0.150	0.490	0.177	2.777	0.07853
11.50	30	0.487	0.151	0.394	0.110	3.571	0.10962
12.27	30	0.510	0.152	0.371	0.102	3.644	0.11704
13.14	30	0.514	0.152	0.347	0.089	3.884	0.13150
14.15	30	0.519	0.152	0.323	0.078	4.155	0.15012
16.73	30	0.534	0.154	0.276	0.057	4.812	0.20713
18.40	30	0.545	0.154	0.252	0.048	5.215	0.25141

Crank-Nicolson based model (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.00	60	0.291	0.083	2.475	4.365	0.567	0.21574
4.00	60	0.293	0.083	1.240	1.099	1.129	0.23757
4.60	60	0.294	0.083	1.079	0.834	1.295	0.24741
5.75	60	0.295	0.083	0.865	0.535	1.616	0.26854
7.36	60	0.297	0.083	0.678	0.329	2.060	0.30776
8.00	60	0.298	0.083	0.624	0.279	2.235	0.32665
8.36	60	0.299	0.083	0.598	0.256	2.331	0.33881
9.20	60	0.300	0.083	0.545	0.213	2.560	0.36704
11.50	60	0.294	0.084	0.438	0.133	3.286	0.44948
12.27	60	0.308	0.084	0.412	0.123	3.353	0.50374
13.14	60	0.310	0.084	0.385	0.108	3.578	0.54991
14.15	60	0.314	0.085	0.359	0.094	3.817	0.60982
16.73	60	0.323	0.085	0.306	0.069	4.422	0.77895
18.40	60	0.329	0.086	0.280	0.058	4.802	0.90180

Crank-Nicolson based model (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.00	40	0.664	0.218	4.370	6.640	0.658	0.09267
4.00	40	0.670	0.219	2.188	1.675	1.306	0.10611
4.60	40	0.670	0.219	1.904	1.267	1.503	0.11063
5.75	40	0.672	0.219	1.525	0.813	1.876	0.12267
7.36	40	0.676	0.220	1.195	0.499	2.393	0.14583
8.00	40	0.678	0.220	1.100	0.424	2.597	0.15733
8.36	40	0.680	0.220	1.053	0.389	2.709	0.16483
9.20	40	0.683	0.221	0.959	0.323	2.971	0.18319
11.50	40	0.670	0.222	0.773	0.203	3.815	0.23221
12.27	40	0.699	0.222	0.725	0.186	3.899	0.27537
13.14	40	0.705	0.223	0.678	0.163	4.152	0.30972
14.15	40	0.712	0.223	0.631	0.142	4.440	0.35356
16.73	40	0.732	0.225	0.538	0.105	5.144	0.48795
18.40	40	0.747	0.226	0.492	0.088	5.575	0.59306

MacCormack based model (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.00	30	0.472	0.149	2.234	3.540	0.631	0.05313
4.00	30	0.483	0.150	1.122	0.906	1.238	0.06067
4.60	30	0.486	0.150	0.976	0.689	1.417	0.06436
5.75	30	0.491	0.150	0.783	0.446	1.758	0.07334
7.36	30	0.504	0.151	0.614	0.279	2.201	0.08894
8.00	30	0.493	0.147	0.550	0.231	2.381	0.09973
8.36	30	0.507	0.151	0.542	0.217	2.493	0.10225
9.20	30	0.513	0.151	0.494	0.182	2.716	0.11443
11.50	30	0.530	0.153	0.398	0.120	3.309	0.15599
12.27	30	0.536	0.153	0.374	0.107	3.498	0.17244
13.14	30	0.543	0.153	0.350	0.094	3.710	0.19283
14.15	30	0.551	0.154	0.326	0.083	3.952	0.21837
16.73	30	0.573	0.155	0.278	0.061	4.530	0.29272
18.40	30	0.587	0.156	0.254	0.052	4.893	0.34752

MacCormack based model (experiment 9)

Δx , mete	Δt , secur	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.00	60	0.296	0.083	2.482	4.440	0.559	0.24652
4.00	60	0.303	0.083	1.246	1.136	1.096	0.30160
4.60	60	0.304	0.083	1.085	0.862	1.258	0.31958
5.75	60	0.308	0.083	0.870	0.559	1.557	0.36217
7.36	60	0.316	0.084	0.683	0.350	1.951	0.43631
8.00	60	0.317	0.084	0.629	0.297	2.118	0.46473
8.36	60	0.318	0.084	0.603	0.273	2.210	0.48247
9.20	60	0.322	0.084	0.549	0.228	2.406	0.52894
11.50	60	0.333	0.085	0.443	0.151	2.930	0.67173
12.27	60	0.337	0.085	0.416	0.134	3.095	0.72469
13.14	60	0.341	0.085	0.389	0.118	3.287	0.78663
14.15	60	0.346	0.086	0.363	0.104	3.501	0.86203
16.73	60	0.360	0.086	0.310	0.077	4.012	1.07108
18.40	60	0.369	0.087	0.283	0.065	4.331	1.21594

MacCormack based model (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.00	40	0.641	0.219	4.385	6.410	0.684	0.10440
4.00	40	0.656	0.220	2.202	1.640	1.342	0.14916
4.60	40	0.660	0.220	1.917	1.248	1.537	0.16545
5.75	40	0.668	0.221	1.538	0.808	1.904	0.20103
7.36	40	0.686	0.222	1.207	0.507	2.384	0.26433
8.00	40	0.687	0.223	1.113	0.429	2.592	0.28961
8.36	40	0.690	0.223	1.066	0.395	2.701	0.30611
9.20	40	0.698	0.223	0.971	0.330	2.945	0.34723
11.50	40	0.722	0.225	0.783	0.218	3.586	0.47919
12.27	40	0.730	0.226	0.736	0.194	3.793	0.52900
13.14	40	0.739	0.226	0.689	0.171	4.026	0.58925
14.15	40	0.751	0.227	0.642	0.150	4.282	0.66412
16.73	40	0.781	0.229	0.548	0.112	4.911	0.87500
18.40	40	0.801	0.231	0.502	0.095	5.300	1.02683

QUICKEST based model (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	30	0.469	0.149	0.560	0.220	2.545	0.04464
8.36	30	0.470	0.149	0.535	0.202	2.654	0.04392
9.20	30	0.468	0.149	0.485	0.166	2.926	0.04372
11.50	30	0.467	0.148	0.387	0.106	3.650	0.04113
12.27	30	0.465	0.148	0.362	0.093	3.904	0.04083
13.14	30	0.461	0.148	0.337	0.080	4.213	0.04107
14.15	30	0.455	0.148	0.313	0.068	4.591	0.04193
15.33	30	0.444	0.147	0.288	0.057	5.087	0.04417
16.73	30	0.429	0.147	0.264	0.046	5.732	0.04800
18.40	30	0.404	0.147	0.239	0.036	6.681	0.05533
20.44	30	0.368	0.146	0.215	0.026	8.130	0.06799
23.00	30	0.312	0.146	0.190	0.018	10.762	0.09099

QUICKEST based model (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	60	0.300	0.083	0.621	0.281	2.207	0.20333
8.36	60	0.300	0.083	0.593	0.257	2.305	0.20379
9.20	60	0.299	0.083	0.538	0.212	2.539	0.20392
11.5	60	0.298	0.082	0.429	0.135	3.170	0.21387
12.27	60	0.297	0.082	0.401	0.118	3.388	0.21884
13.14	60	0.294	0.082	0.374	0.102	3.661	0.22297
14.15	60	0.290	0.082	0.347	0.087	3.990	0.22968
15.33	60	0.284	0.082	0.319	0.072	4.405	0.23973
16.73	60	0.274	0.081	0.292	0.059	4.970	0.25312
18.40	60	0.258	0.081	0.265	0.046	5.791	0.27304
20.44	60	0.235	0.081	0.238	0.034	7.045	0.30921
23.00	60	0.199	0.081	0.211	0.023	9.333	0.37161

QUICKEST based model (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	40	0.633	0.221	1.106	0.396	2.796	0.05014
8.36	40	0.633	0.221	1.056	0.362	2.918	0.04956
9.20	40	0.630	0.220	0.957	0.298	3.215	0.04713
11.50	40	0.627	0.219	0.761	0.190	4.011	0.04589
12.27	40	0.624	0.218	0.712	0.166	4.291	0.04576
13.14	40	0.619	0.218	0.663	0.143	4.626	0.04571
14.15	40	0.610	0.217	0.614	0.122	5.045	0.04557
15.33	40	0.596	0.217	0.566	0.101	5.580	0.04592
16.73	40	0.574	0.216	0.517	0.082	6.303	0.04730
18.40	40	0.542	0.216	0.469	0.064	7.321	0.05186
20.44	40	0.493	0.215	0.421	0.047	8.912	0.06283
23.00	40	0.417	0.214	0.372	0.032	11.811	0.08774

Appendix E-2: Prediction of downstream concentration profile measured by coefficients of determination. Numerical methods were applied with predicted dispersion coefficients given by numerical methods-based models. The tables show space steps, predicted dispersion coefficients, nondimensional parameters and coefficients of determination (R^2).

Crank-Nicolson method (Experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	30	0.483	0.149	2.233	3.626	0.616	0.996
4.00	30	0.487	0.149	1.118	0.913	1.224	0.996
4.60	30	0.488	0.149	0.972	0.692	1.406	0.996
5.75	30	0.489	0.149	0.779	0.444	1.756	0.995
7.36	30	0.493	0.150	0.610	0.273	2.236	0.994
8.00	30	0.494	0.150	0.562	0.232	2.428	0.993
8.36	30	0.495	0.150	0.538	0.212	2.536	0.993
9.20	30	0.498	0.150	0.490	0.177	2.777	0.992
11.50	30	0.487	0.151	0.394	0.110	3.571	0.988
12.27	30	0.510	0.152	0.371	0.102	3.644	0.986
13.14	30	0.514	0.152	0.347	0.089	3.884	0.985
14.15	30	0.519	0.152	0.323	0.078	4.155	0.982
16.73	30	0.534	0.154	0.276	0.057	4.812	0.975
18.40	30	0.545	0.154	0.252	0.048	5.215	0.969

Crank-Nicolson Experiment 9

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	60	0.291	0.083	2.475	4.365	0.567	0.991
4.00	60	0.293	0.083	1.240	1.099	1.129	0.991
4.60	60	0.294	0.083	1.079	0.834	1.295	0.990
5.75	60	0.295	0.083	0.865	0.535	1.616	0.989
7.36	60	0.297	0.083	0.678	0.329	2.060	0.988
8.00	60	0.298	0.083	0.624	0.279	2.235	0.987
8.36	60	0.299	0.083	0.598	0.256	2.331	0.986
9.20	60	0.300	0.083	0.545	0.213	2.560	0.985
11.50	60	0.294	0.084	0.438	0.133	3.286	0.981
12.27	60	0.308	0.084	0.412	0.123	3.353	0.980
13.14	60	0.310	0.084	0.385	0.108	3.578	0.978
14.15	60	0.314	0.085	0.359	0.094	3.817	0.975
16.73	60	0.323	0.085	0.306	0.069	4.422	0.968
18.40	60	0.329	0.086	0.280	0.058	4.802	0.963

Crank-Nicolson method Experiment 13

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	40	0.664	0.218	4.370	6.640	0.658	0.996
4.00	40	0.670	0.219	2.188	1.675	1.306	0.995
4.60	40	0.670	0.219	1.904	1.267	1.503	0.995
5.75	40	0.672	0.219	1.525	0.813	1.876	0.995
7.36	40	0.676	0.220	1.195	0.499	2.393	0.994
8.00	40	0.678	0.220	1.100	0.424	2.597	0.993
8.36	40	0.680	0.220	1.053	0.389	2.709	0.993
9.20	40	0.683	0.221	0.959	0.323	2.971	0.992
11.50	40	0.670	0.222	0.773	0.203	3.815	0.990
12.27	40	0.699	0.222	0.725	0.186	3.899	0.988
13.14	40	0.705	0.223	0.678	0.163	4.152	0.987
14.15	40	0.712	0.223	0.631	0.142	4.440	0.985
16.73	40	0.732	0.225	0.538	0.105	5.144	0.979
18.40	40	0.747	0.226	0.492	0.088	5.575	0.975

MacCormack method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	30	0.472	0.149	2.234	3.540	0.631	0.996
4.00	30	0.483	0.150	1.122	0.906	1.238	0.994
4.60	30	0.486	0.150	0.976	0.689	1.417	0.994
5.75	30	0.491	0.150	0.783	0.446	1.758	0.993
7.36	30	0.504	0.151	0.614	0.279	2.201	0.990
8.00	30	0.493	0.147	0.550	0.231	2.381	0.989
8.36	30	0.507	0.151	0.542	0.217	2.493	0.989
9.20	30	0.513	0.151	0.494	0.182	2.716	0.987
11.50	30	0.530	0.153	0.398	0.120	3.309	0.981
12.27	30	0.536	0.153	0.374	0.107	3.498	0.980
13.14	30	0.543	0.153	0.350	0.094	3.710	0.977
14.15	30	0.551	0.154	0.326	0.083	3.952	0.974
16.73	30	0.573	0.155	0.278	0.061	4.530	0.965
18.40	30	0.587	0.156	0.254	0.052	4.893	0.959

MacCormack method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	60	0.296	0.083	2.482	4.440	0.559	0.990
4.00	60	0.303	0.083	1.246	1.136	1.096	0.988
4.60	60	0.304	0.083	1.085	0.862	1.258	0.987
5.75	60	0.308	0.083	0.870	0.559	1.557	0.986
7.36	60	0.316	0.084	0.683	0.350	1.951	0.983
8.00	60	0.317	0.084	0.629	0.297	2.118	0.982
8.36	60	0.318	0.084	0.603	0.273	2.210	0.981
9.20	60	0.322	0.084	0.549	0.228	2.406	0.979
11.50	60	0.333	0.085	0.443	0.151	2.930	0.973
12.27	60	0.337	0.085	0.416	0.134	3.095	0.971
13.14	60	0.341	0.085	0.389	0.118	3.287	0.969
14.15	60	0.346	0.086	0.363	0.104	3.501	0.966
16.73	60	0.360	0.086	0.310	0.077	4.012	0.958
18.40	60	0.369	0.087	0.283	0.065	4.331	0.952

MacCormack method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	40	0.641	0.219	4.385	6.410	0.684	0.995
4.00	40	0.656	0.220	2.202	1.640	1.342	0.993
4.60	40	0.660	0.220	1.917	1.248	1.537	0.993
5.75	40	0.668	0.221	1.538	0.808	1.904	0.991
7.36	40	0.686	0.222	1.207	0.507	2.384	0.989
8.00	40	0.687	0.223	1.113	0.429	2.592	0.987
8.36	40	0.690	0.223	1.066	0.395	2.701	0.987
9.20	40	0.698	0.223	0.971	0.330	2.945	0.985
11.50	40	0.722	0.225	0.783	0.218	3.586	0.979
12.27	40	0.730	0.226	0.736	0.194	3.793	0.977
13.14	40	0.739	0.226	0.689	0.171	4.026	0.975
14.15	40	0.751	0.227	0.642	0.150	4.282	0.972
16.73	40	0.781	0.229	0.548	0.112	4.911	0.963
18.40	40	0.801	0.231	0.502	0.095	5.300	0.957

QUICKEST method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
8.00	30	0.469	0.149	0.560	0.220	2.545	0.997
8.36	30	0.470	0.149	0.535	0.202	2.654	0.997
9.20	30	0.468	0.149	0.485	0.166	2.926	0.997
11.50	30	0.467	0.148	0.387	0.106	3.650	0.997
12.27	30	0.465	0.148	0.362	0.093	3.904	0.997
13.14	30	0.461	0.148	0.337	0.080	4.213	0.997
14.15	30	0.455	0.148	0.313	0.068	4.591	0.997
15.33	30	0.444	0.147	0.288	0.057	5.087	0.996
16.73	30	0.429	0.147	0.264	0.046	5.732	0.996
18.40	30	0.404	0.147	0.239	0.036	6.681	0.995
20.44	30	0.368	0.146	0.215	0.026	8.130	0.994
23.00	30	0.312	0.146	0.190	0.018	10.762	0.992

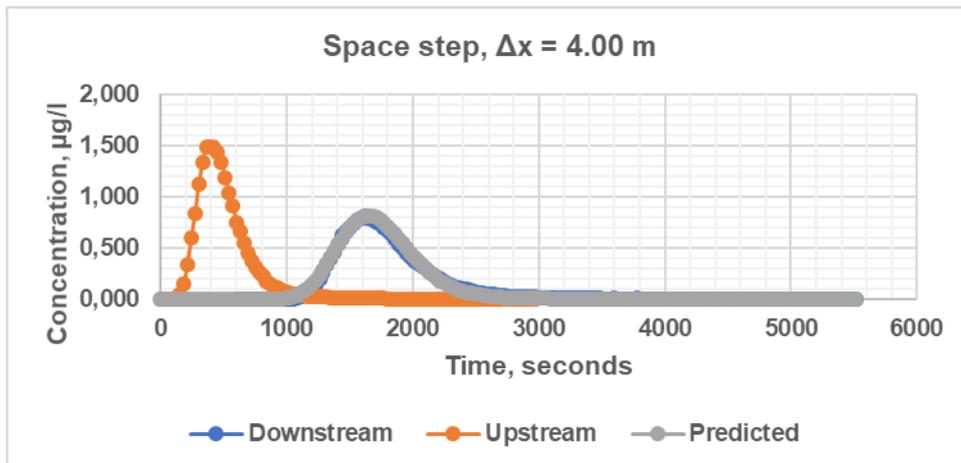
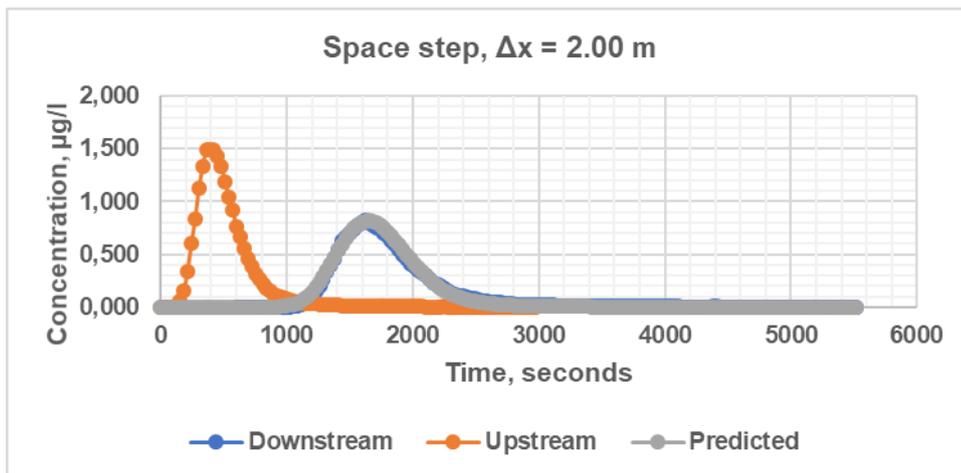
QUICKEST method (experiment 9)

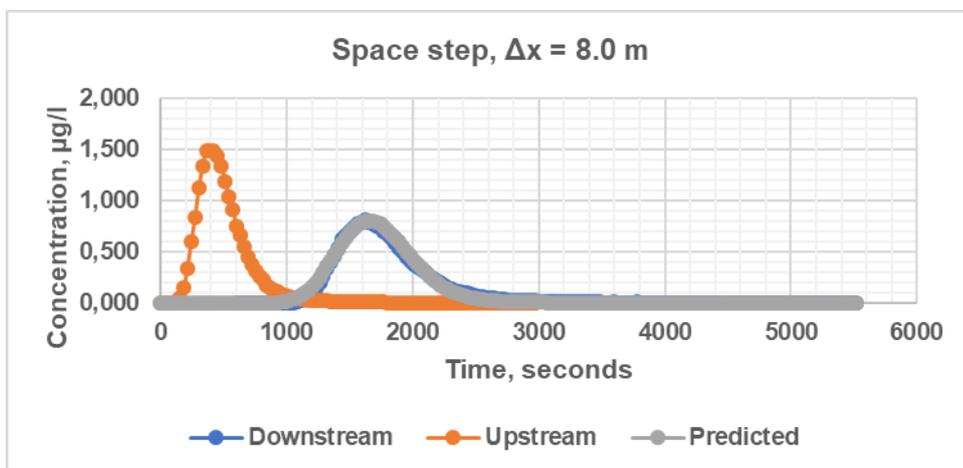
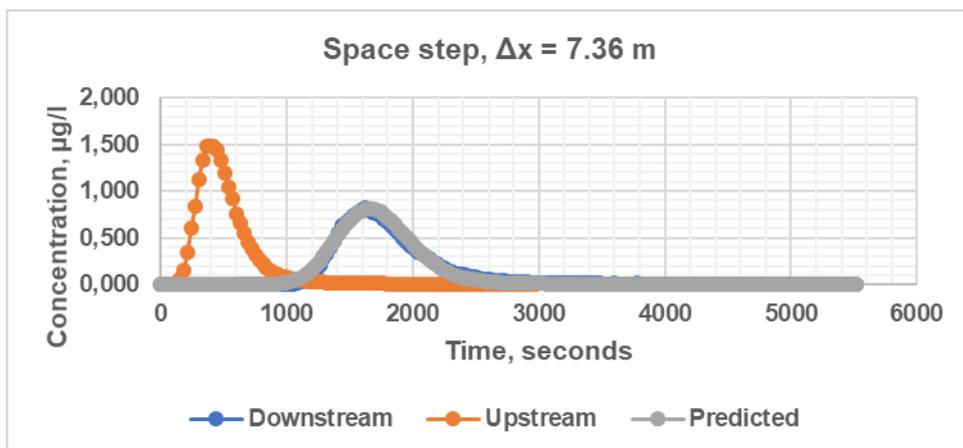
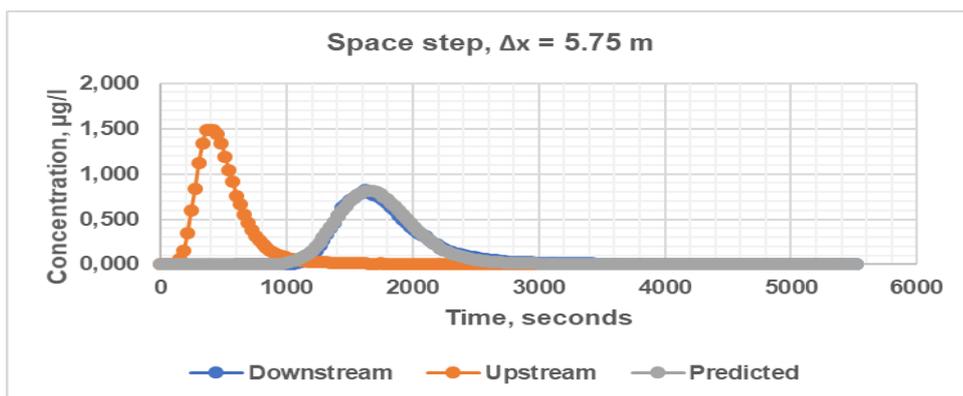
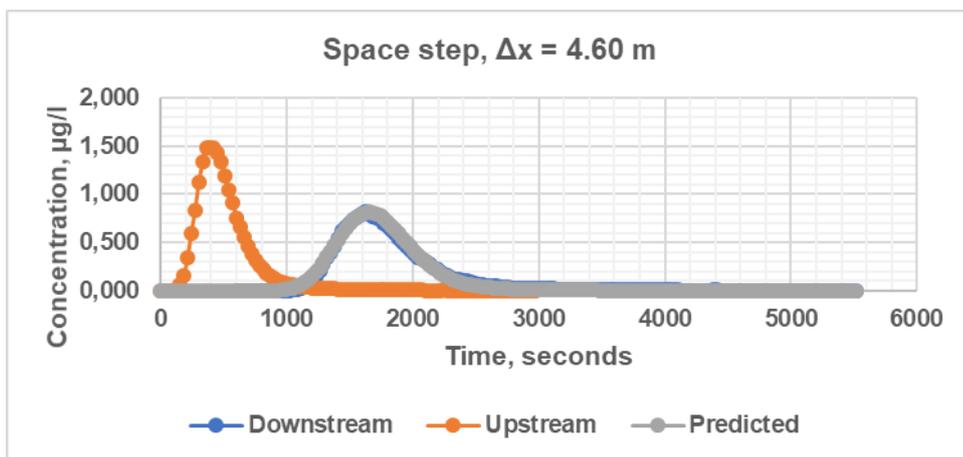
Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
8.00	60	0.300	0.083	0.621	0.281	2.207	0.992
8.36	60	0.300	0.083	0.593	0.257	2.305	0.992
9.20	60	0.299	0.083	0.538	0.212	2.539	0.992
11.50	60	0.298	0.082	0.429	0.135	3.170	0.991
12.27	60	0.297	0.082	0.401	0.118	3.388	0.991
13.14	60	0.294	0.082	0.374	0.102	3.661	0.990
14.15	60	0.290	0.082	0.347	0.087	3.990	0.990
15.33	60	0.284	0.082	0.319	0.072	4.405	0.989
16.73	60	0.274	0.081	0.292	0.059	4.970	0.988
18.40	60	0.258	0.081	0.265	0.046	5.791	0.987
20.44	60	0.235	0.081	0.238	0.034	7.045	0.985
23.00	60	0.199	0.081	0.211	0.023	9.333	0.974

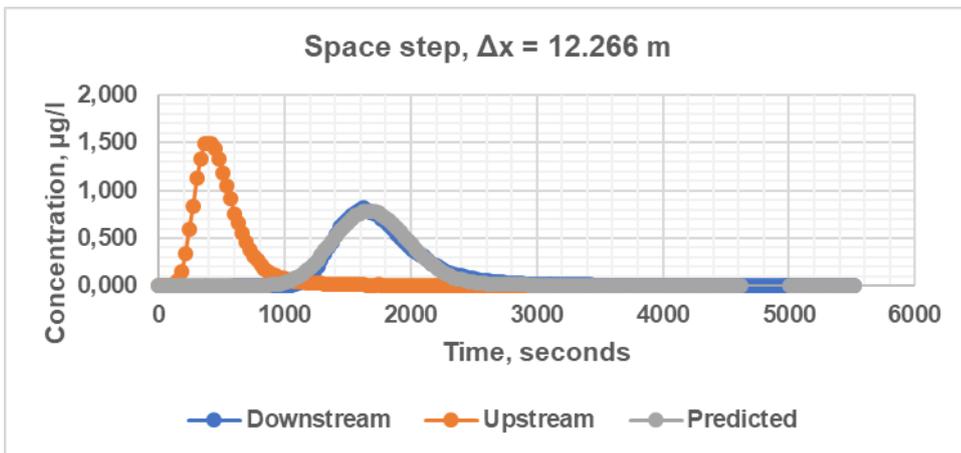
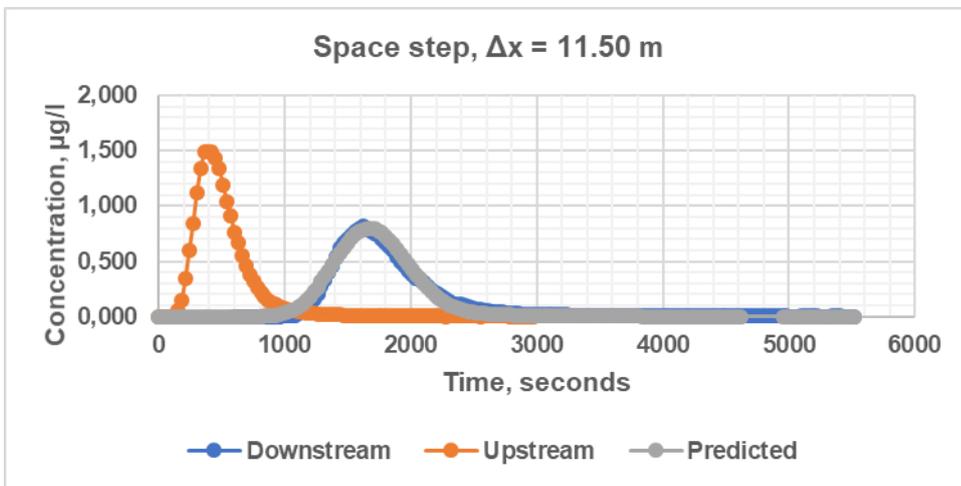
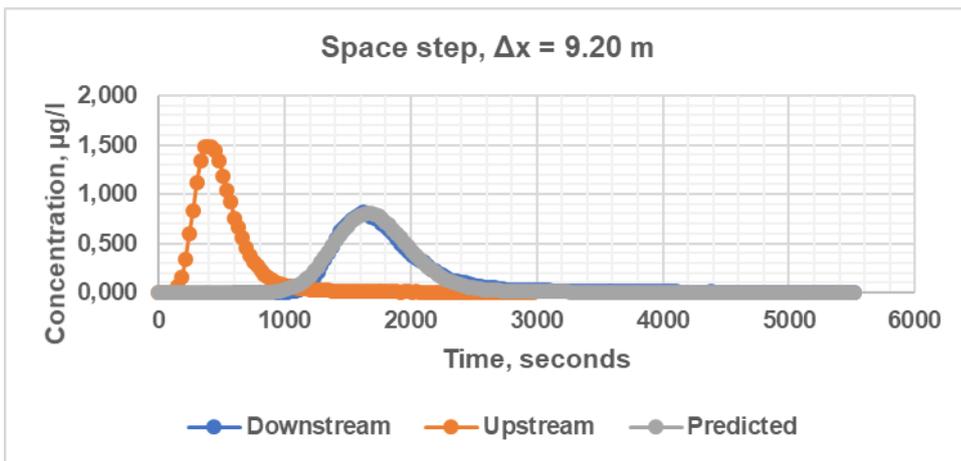
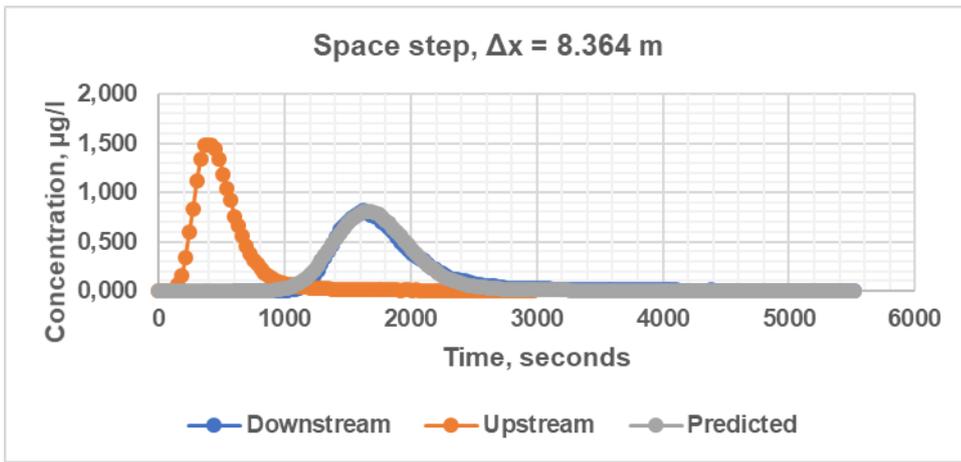
QUICKEST (experiment 13)

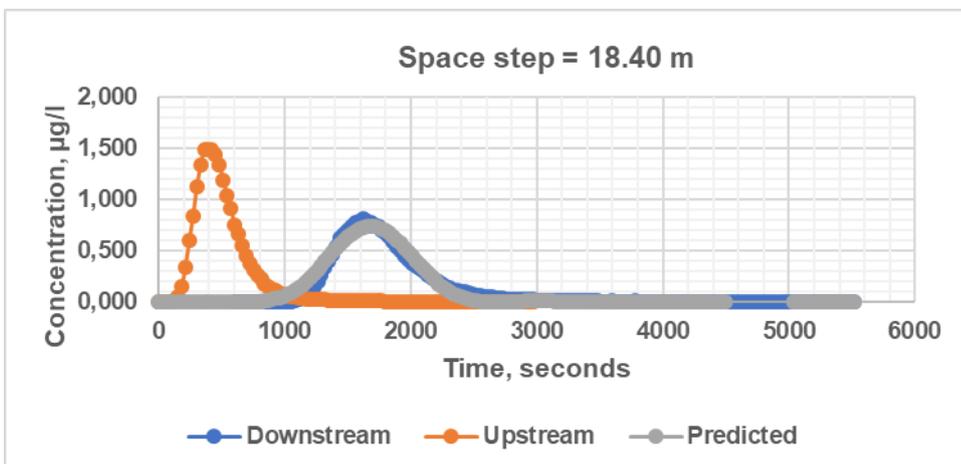
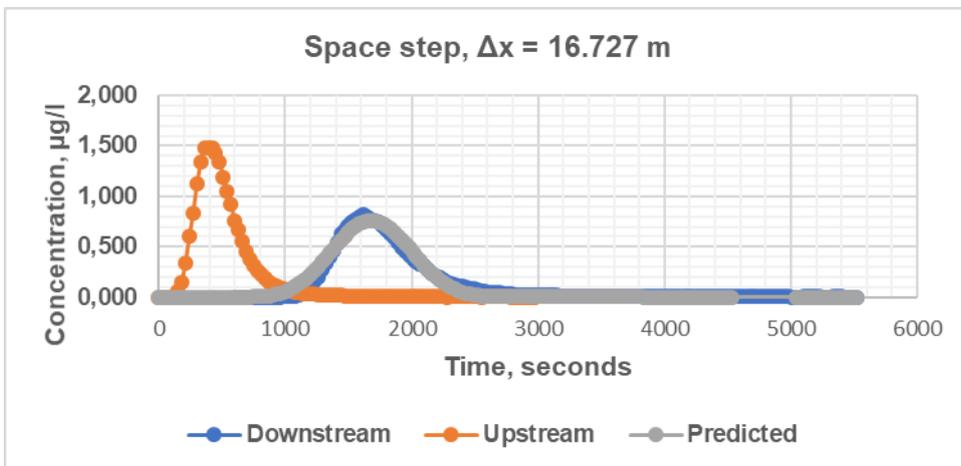
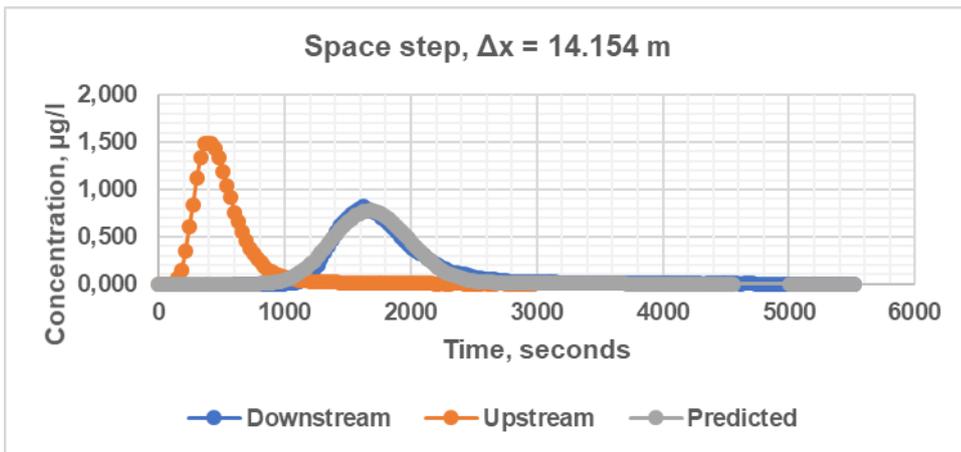
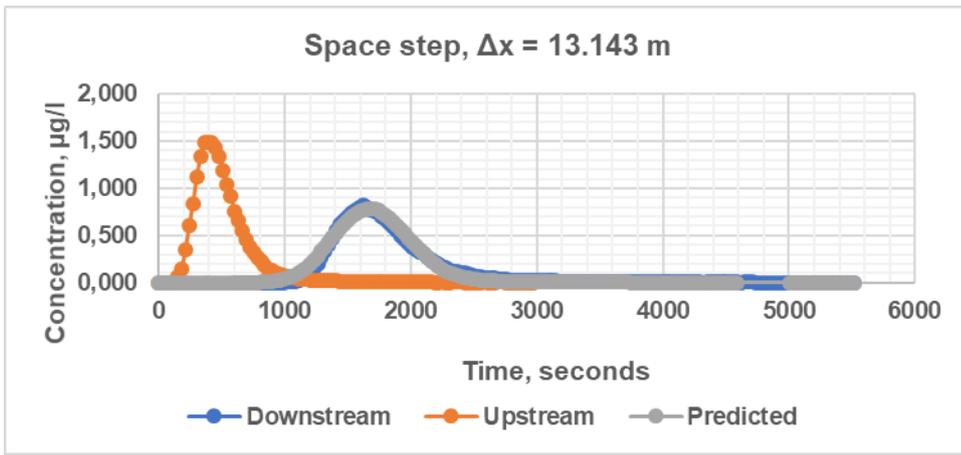
Δx , meters	Δt , seconds	D , m ² /s	V , m/s	c	d	Pe	R^2
8.00	40	0.633	0.221	1.106	0.396	2.796	0.998
8.36	40	0.633	0.221	1.056	0.362	2.918	0.998
9.20	40	0.630	0.220	0.957	0.298	3.215	0.998
11.50	40	0.627	0.219	0.761	0.190	4.011	0.998
12.27	40	0.624	0.218	0.712	0.166	4.291	0.998
13.14	40	0.619	0.218	0.663	0.143	4.626	0.998
14.15	40	0.610	0.217	0.614	0.122	5.045	0.998
15.33	40	0.596	0.217	0.566	0.101	5.580	0.998
16.73	40	0.574	0.216	0.517	0.082	6.303	0.998
18.40	40	0.542	0.216	0.469	0.064	7.321	0.998
20.44	40	0.493	0.215	0.421	0.047	8.912	0.997
23.00	40	0.417	0.214	0.372	0.032	11.811	0.997

Appendix E-3: Charts showing simulations of downstream concentration-time profiles using predicted dispersion coefficients given in Appendix E-1

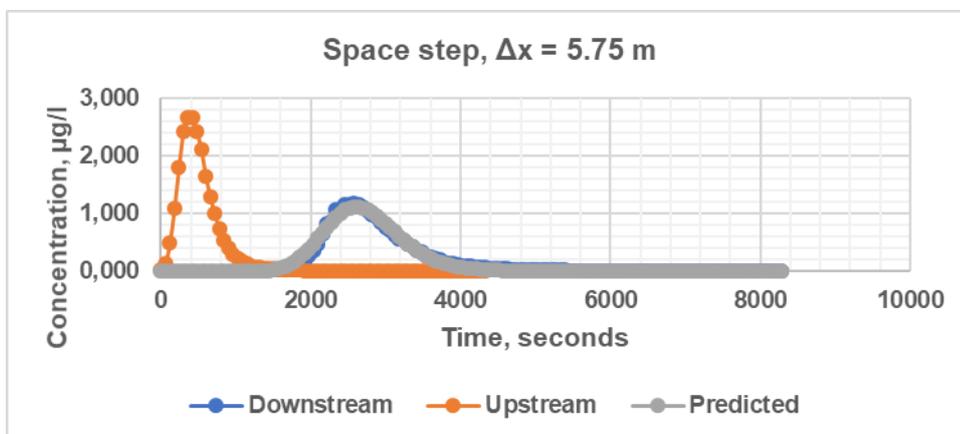
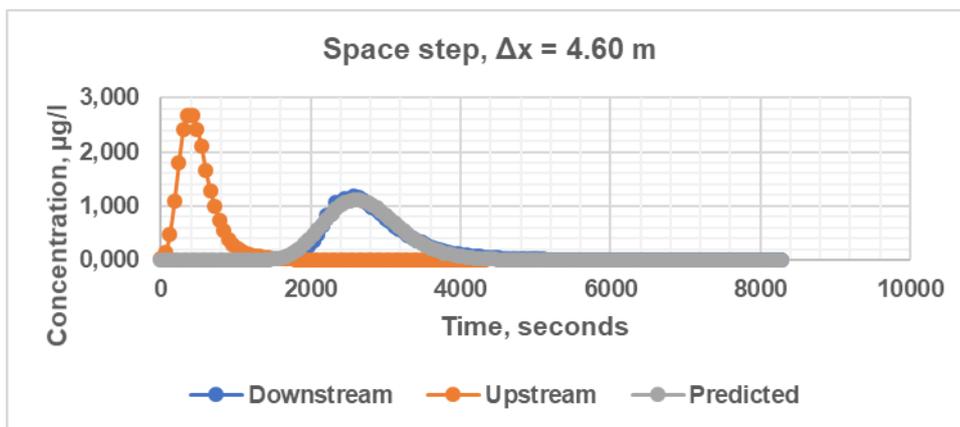
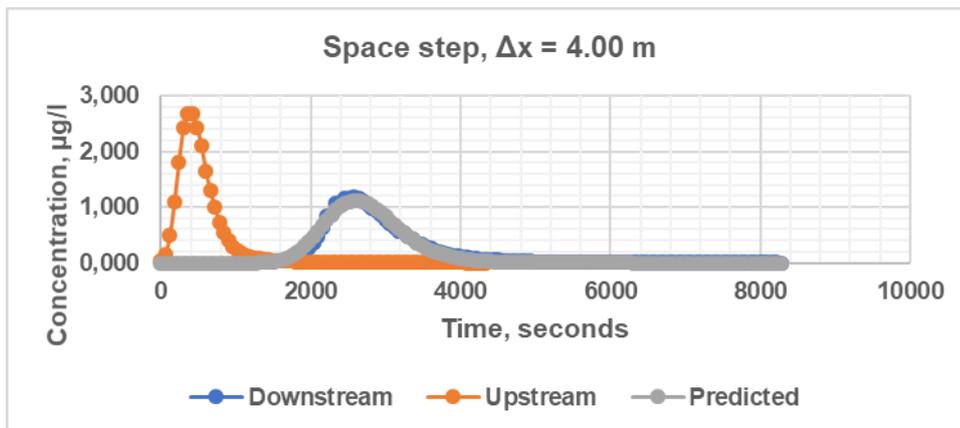
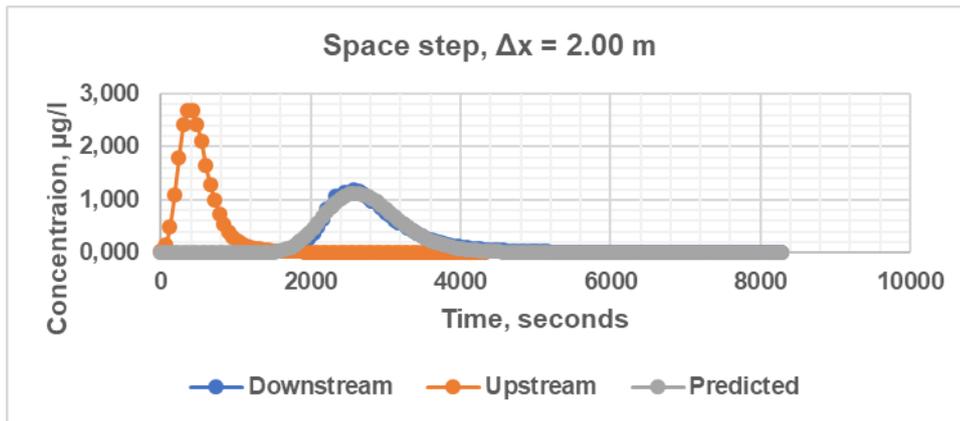
Crank-Nicolson method (experiment 4)

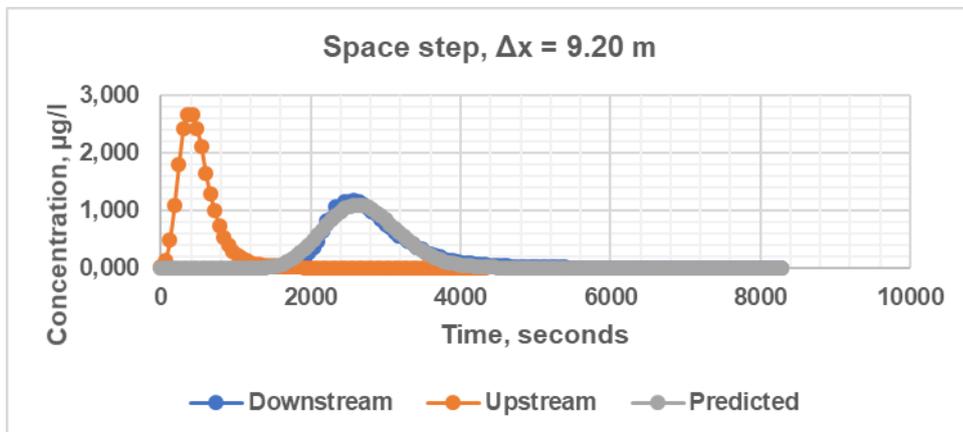
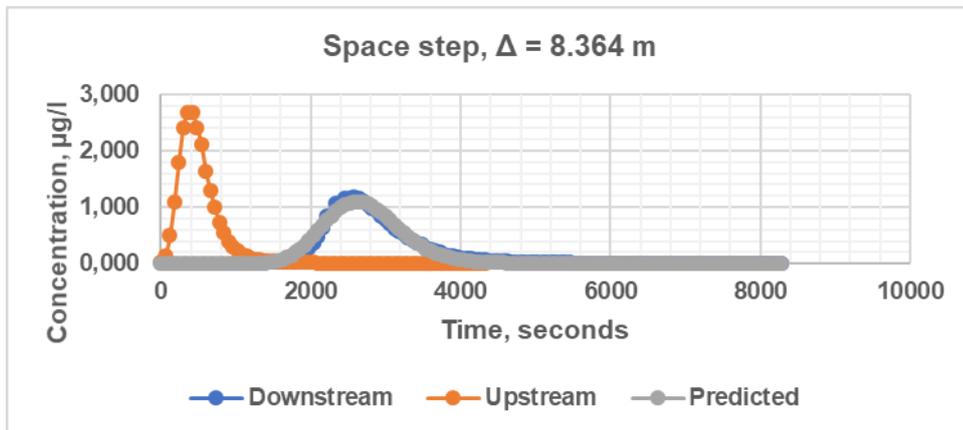
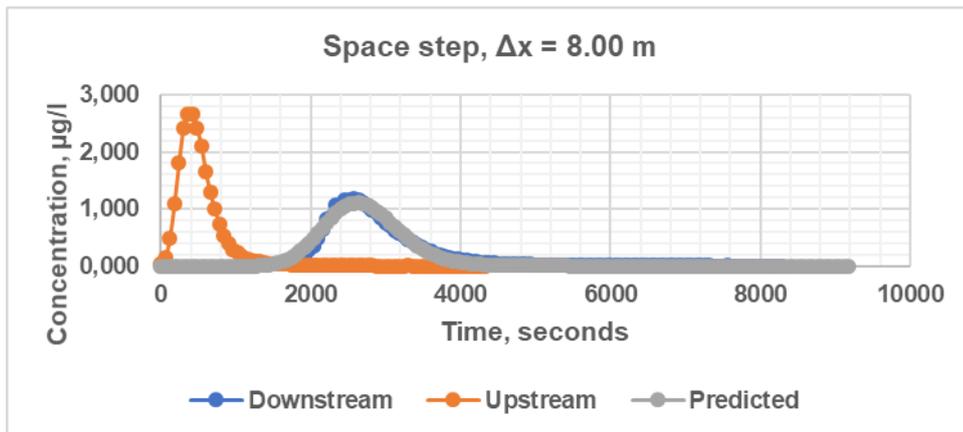
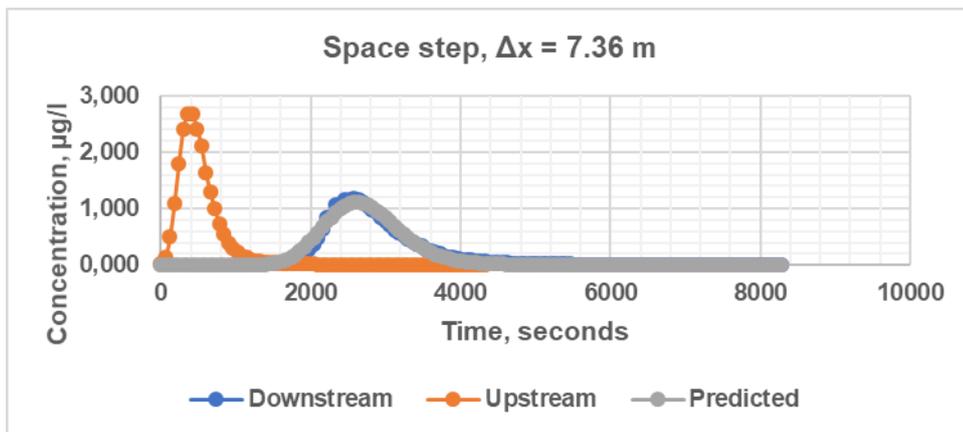


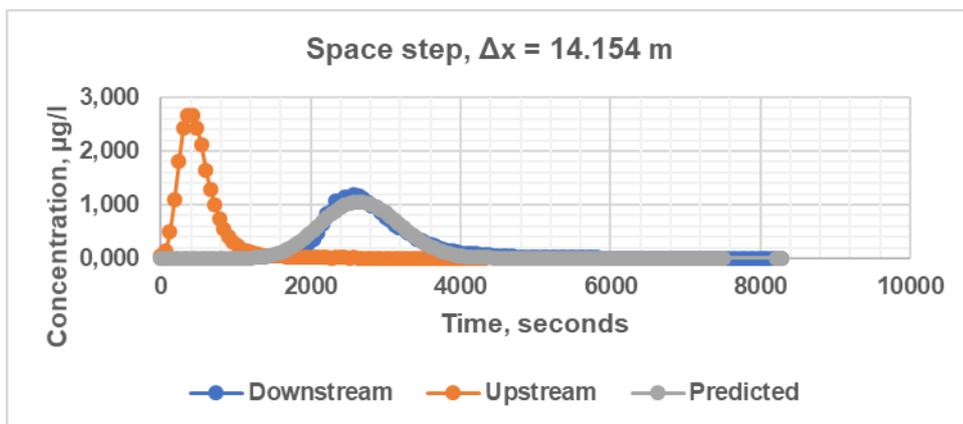
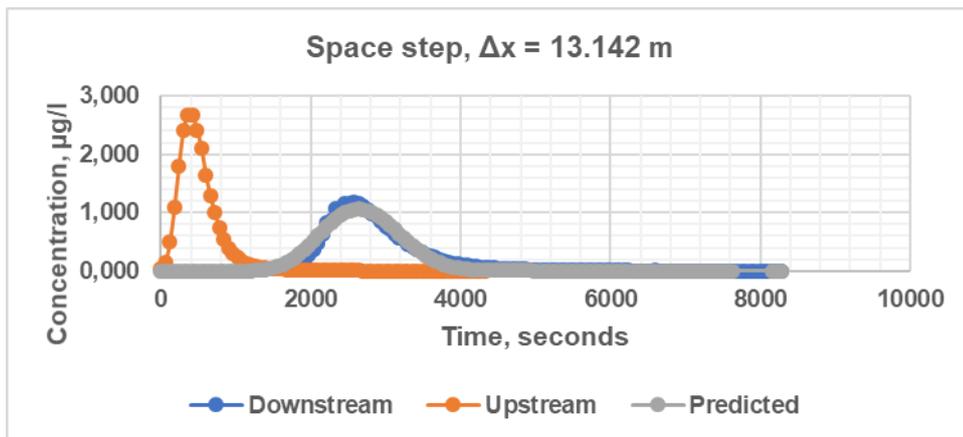
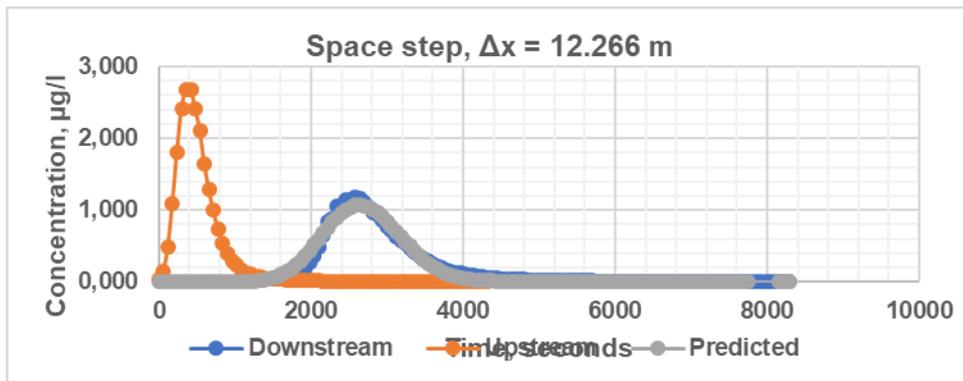
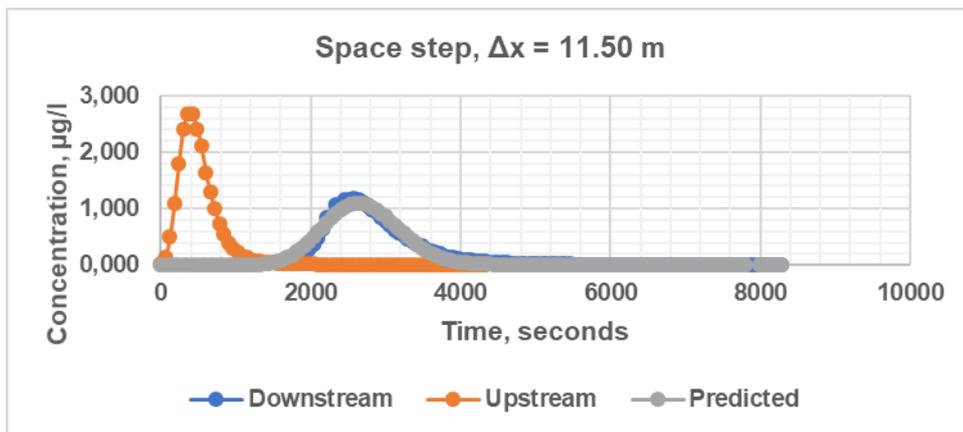


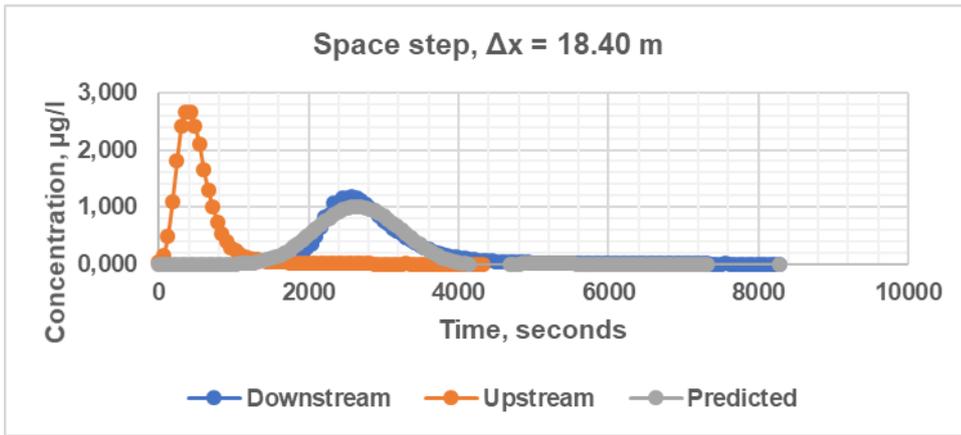
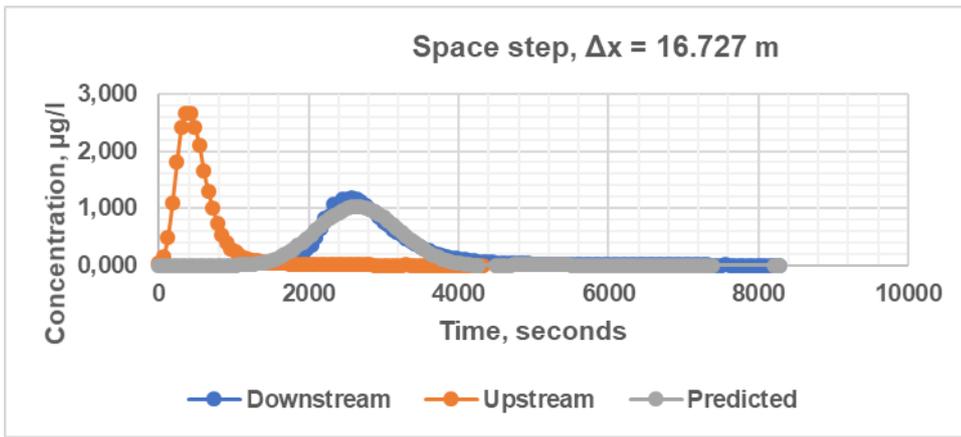


Crank-Nicolson simulations (experiment 9)

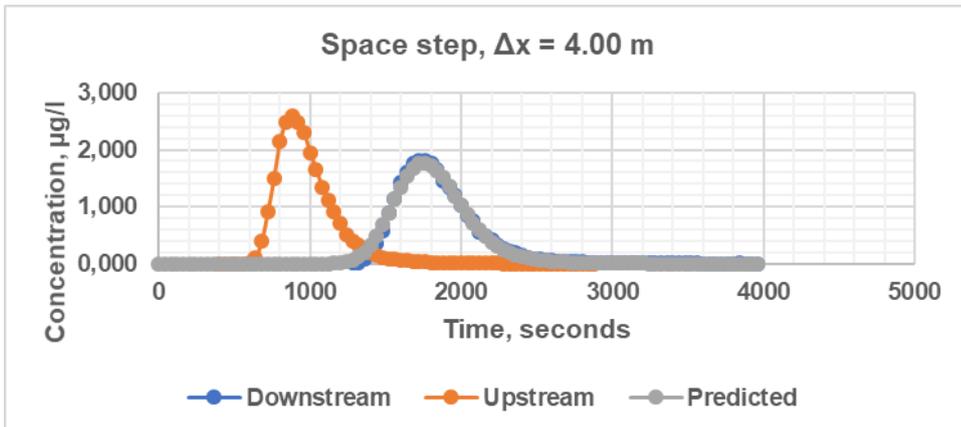
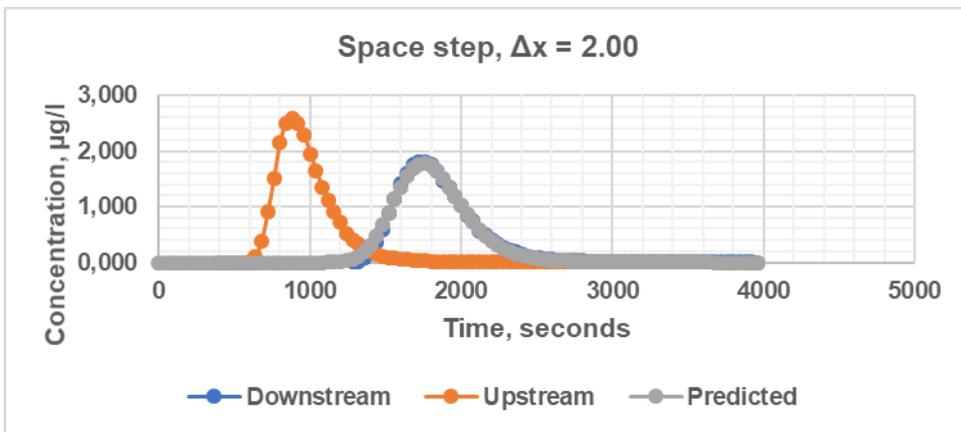


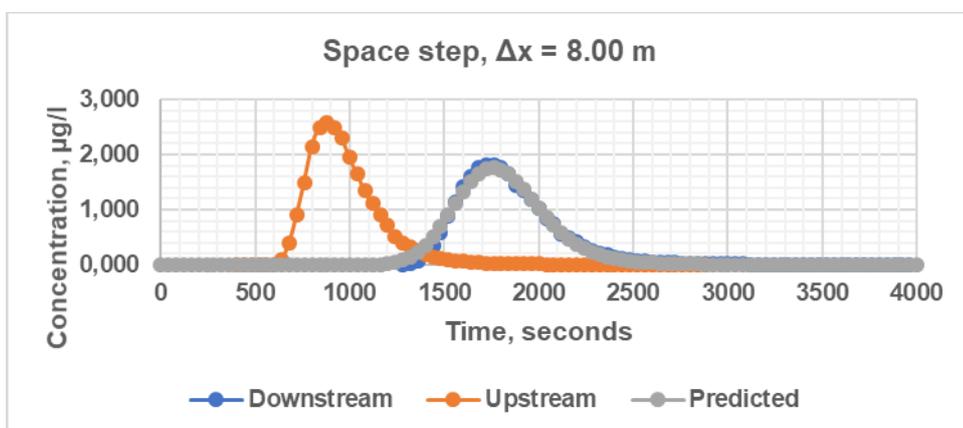
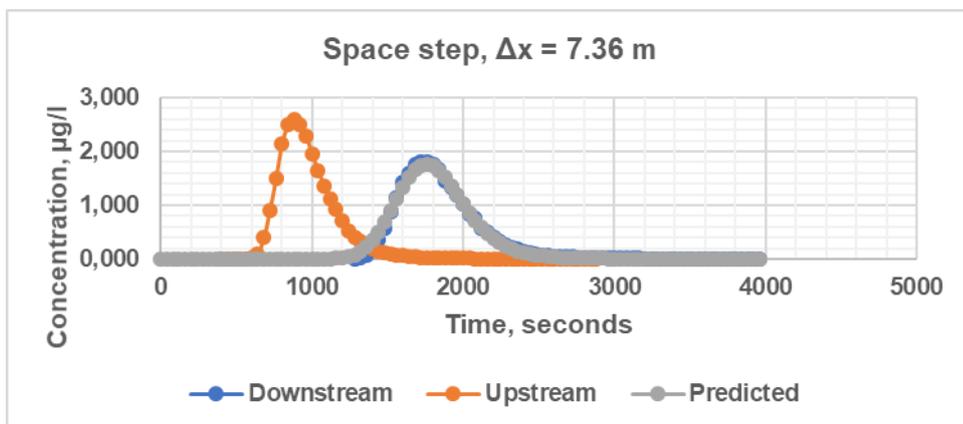
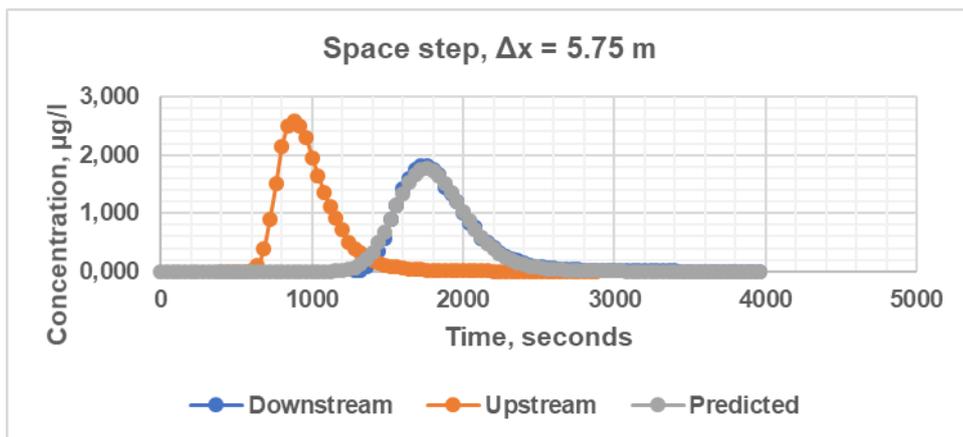
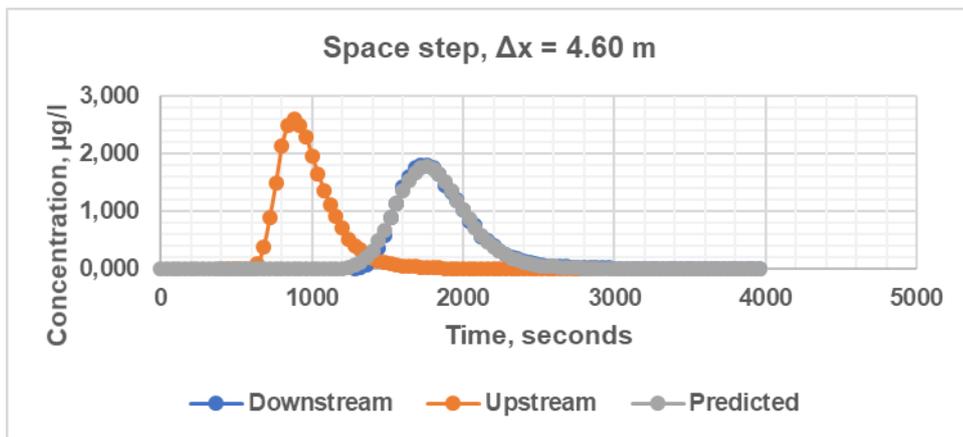


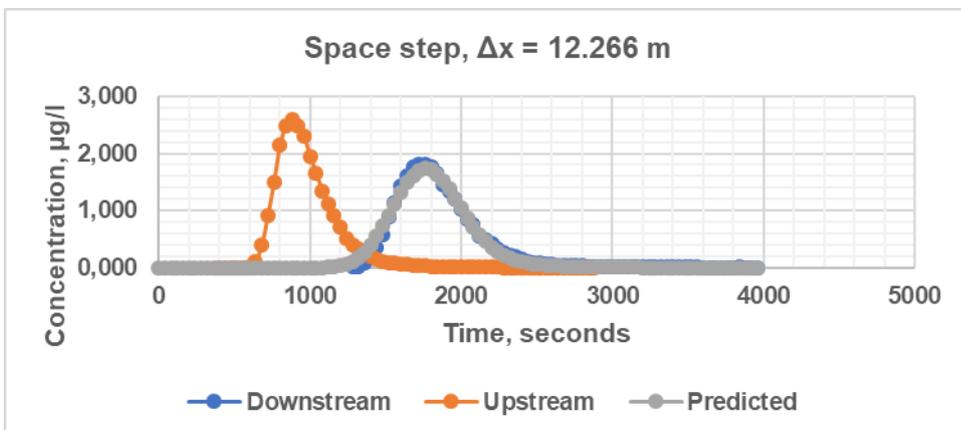
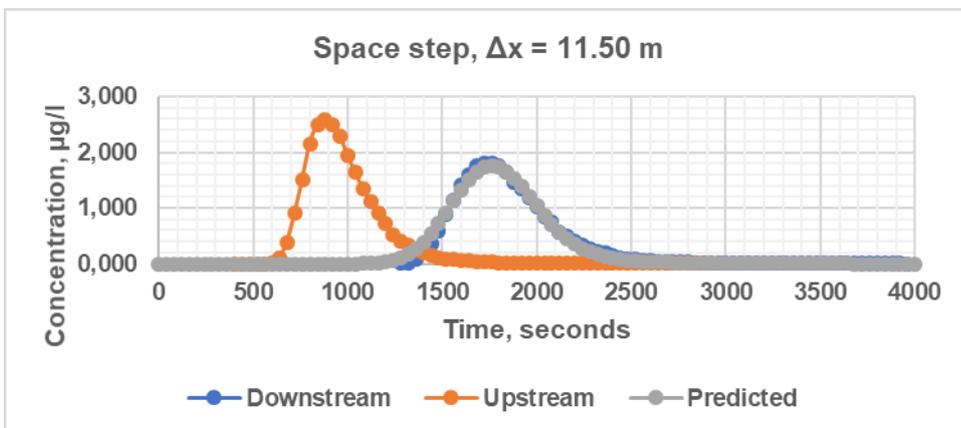
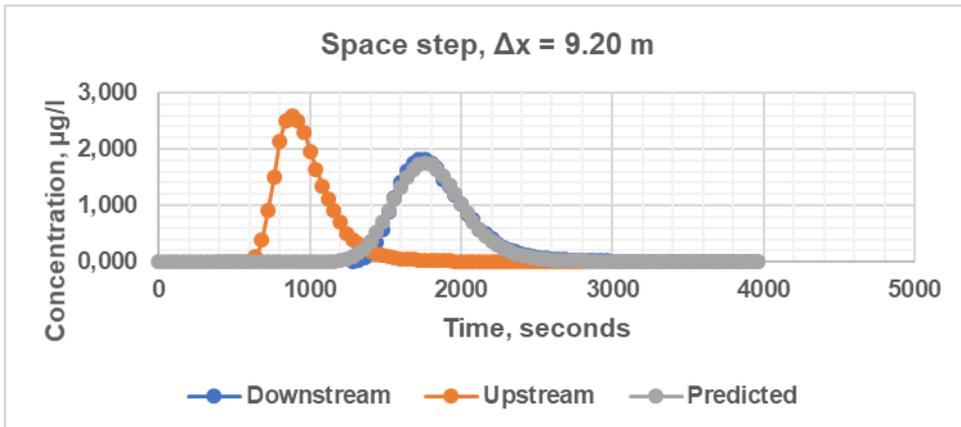
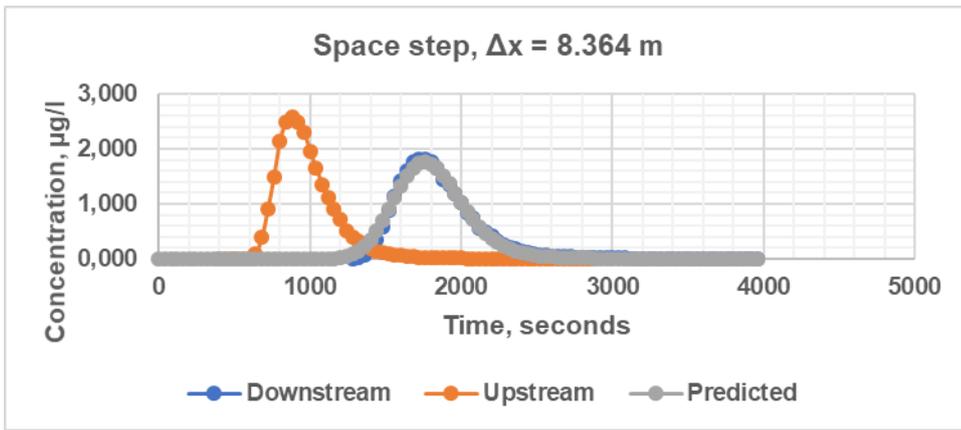


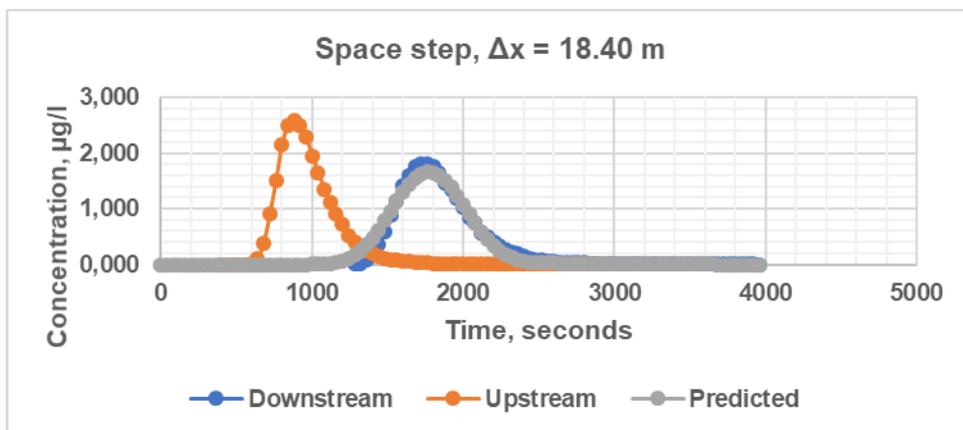
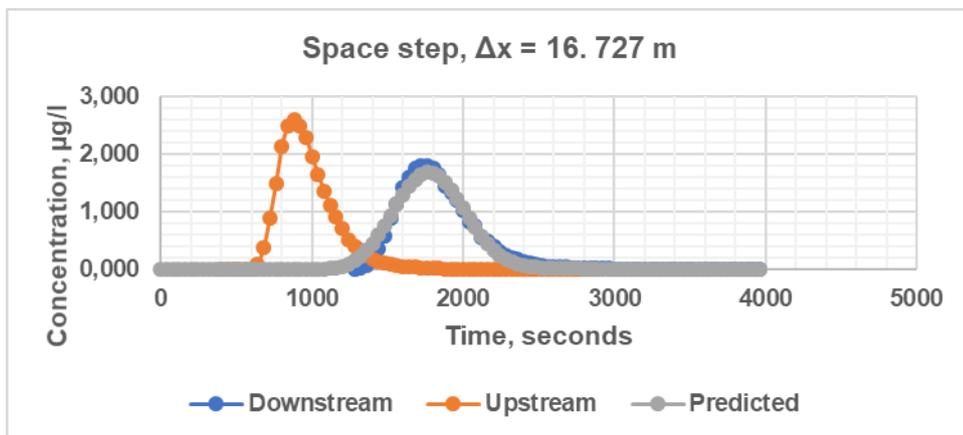
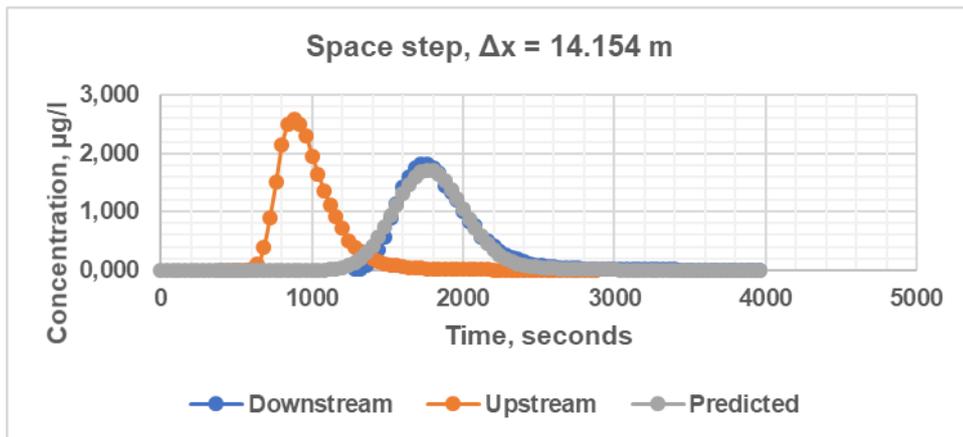
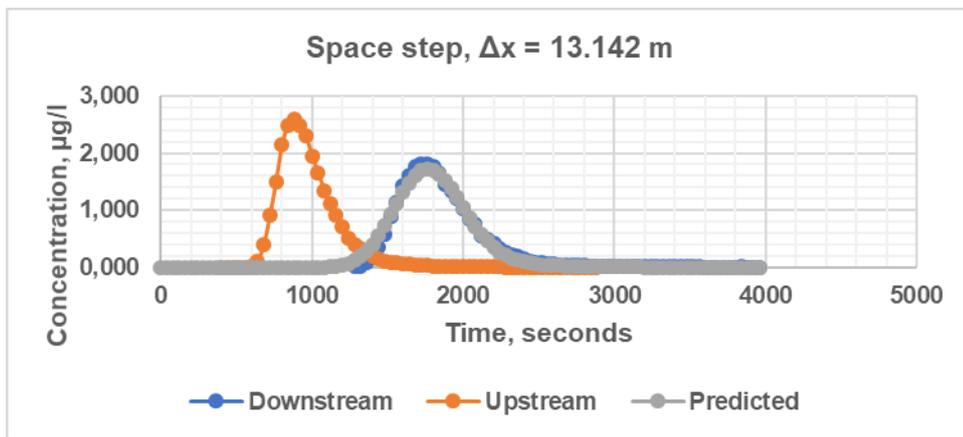


Crank-Nicolson Simulations (experiment 13)

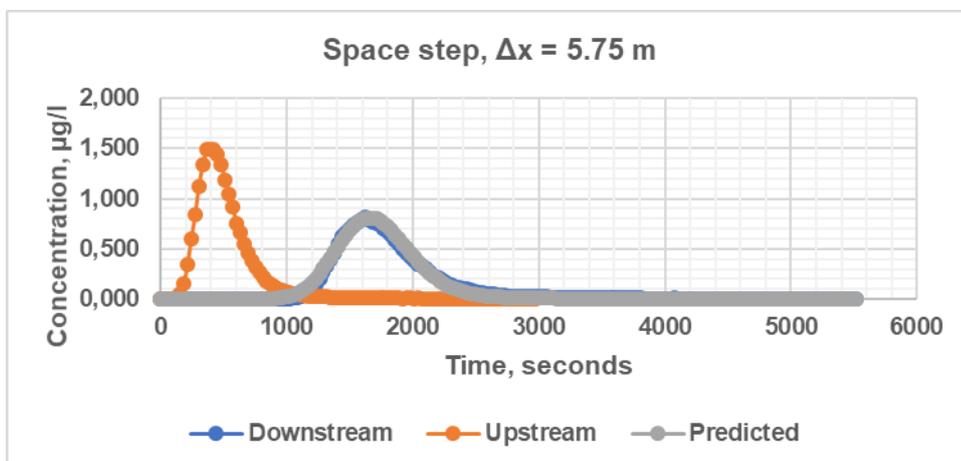
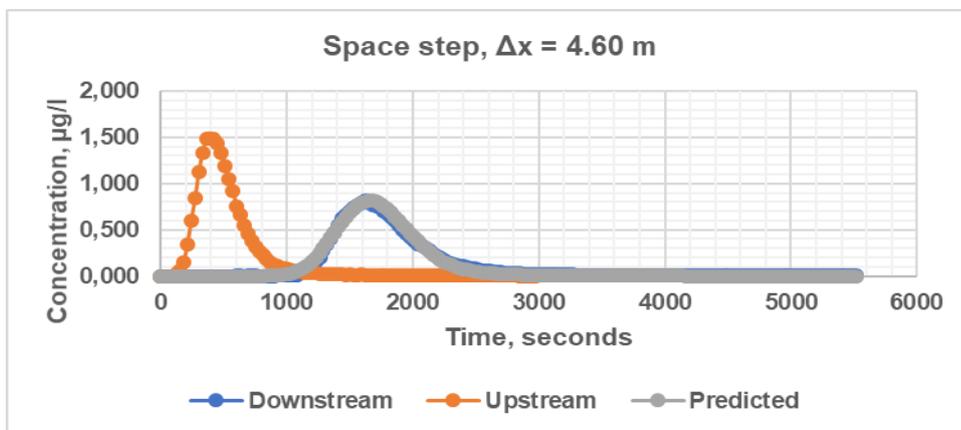
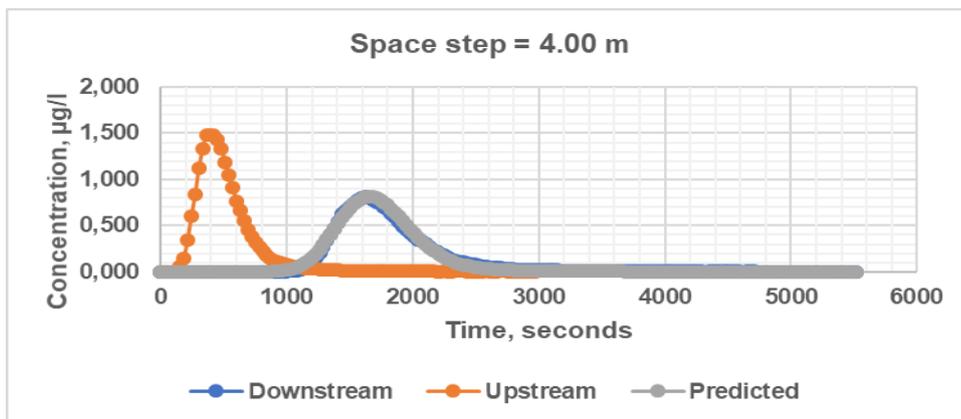
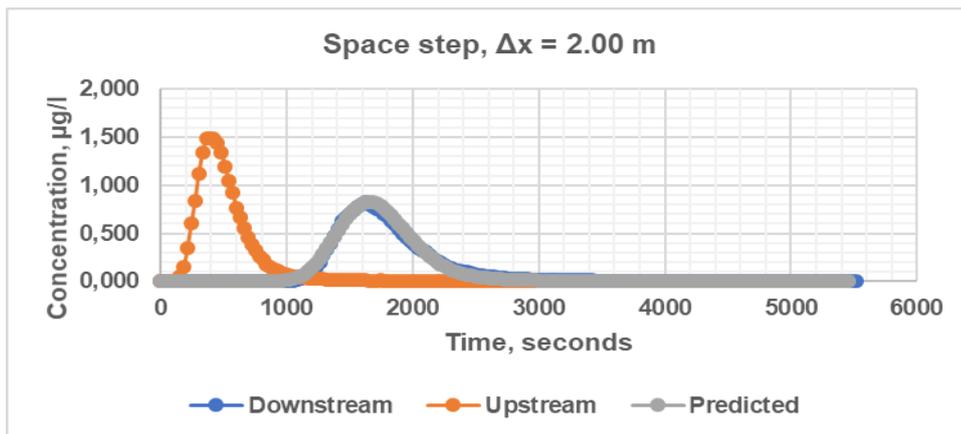


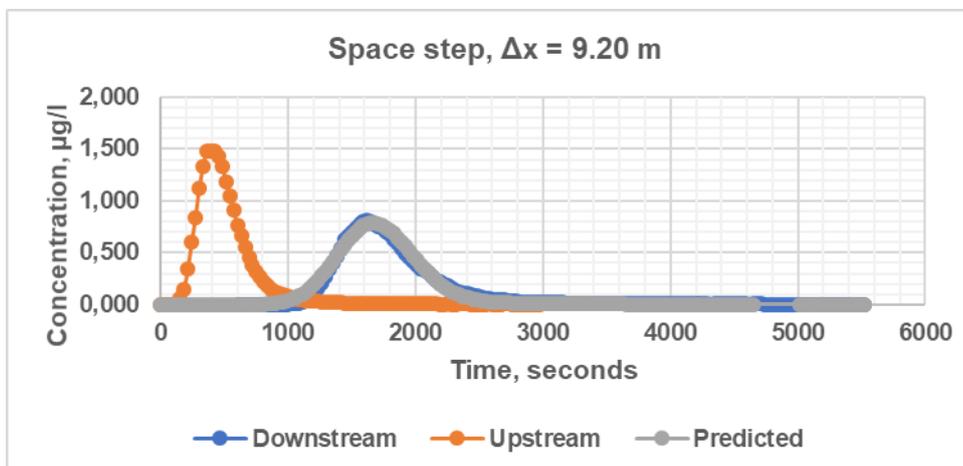
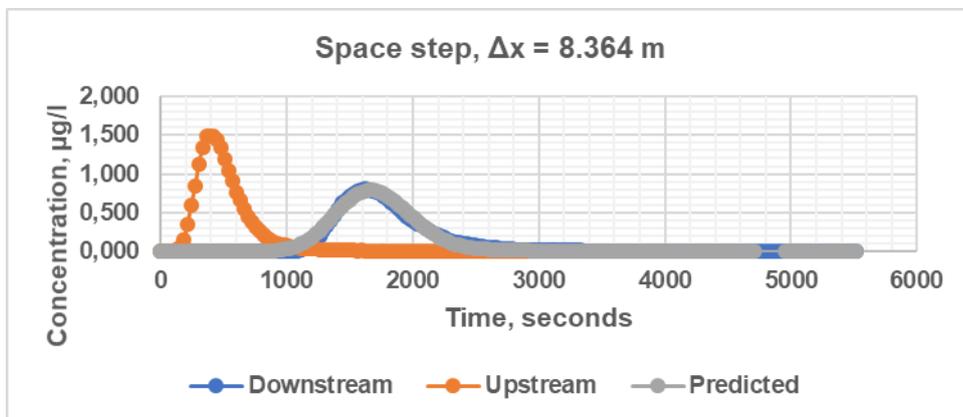
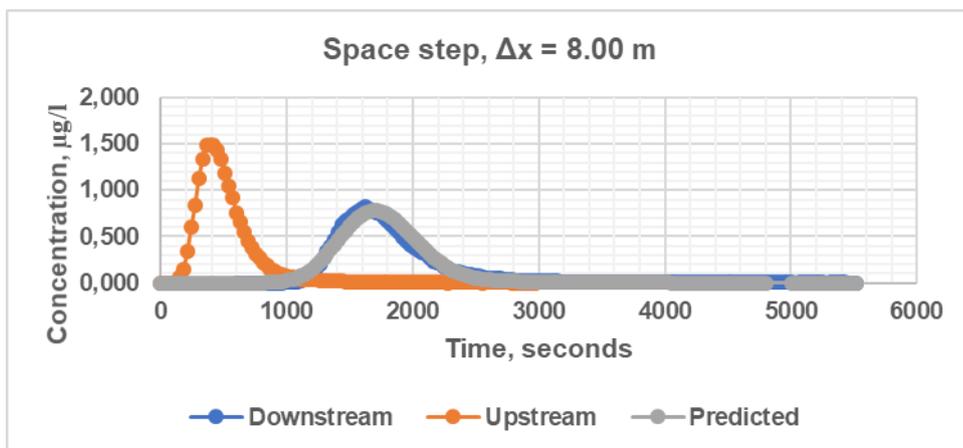
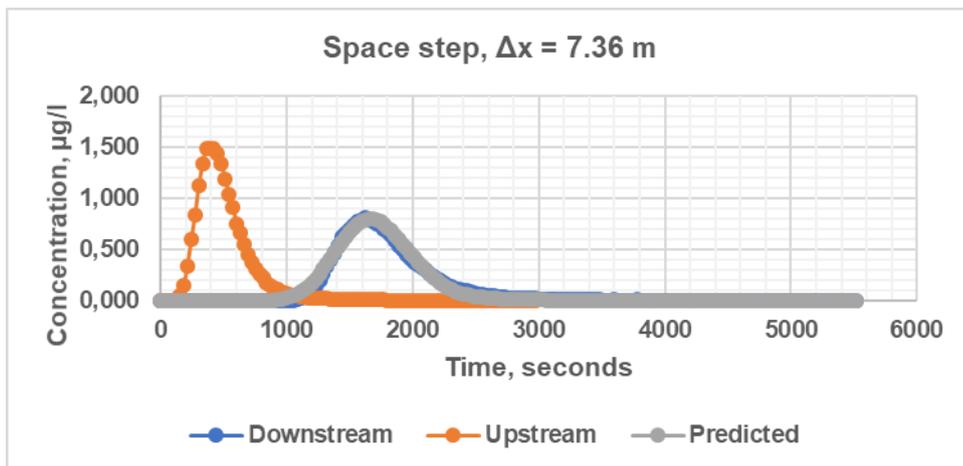


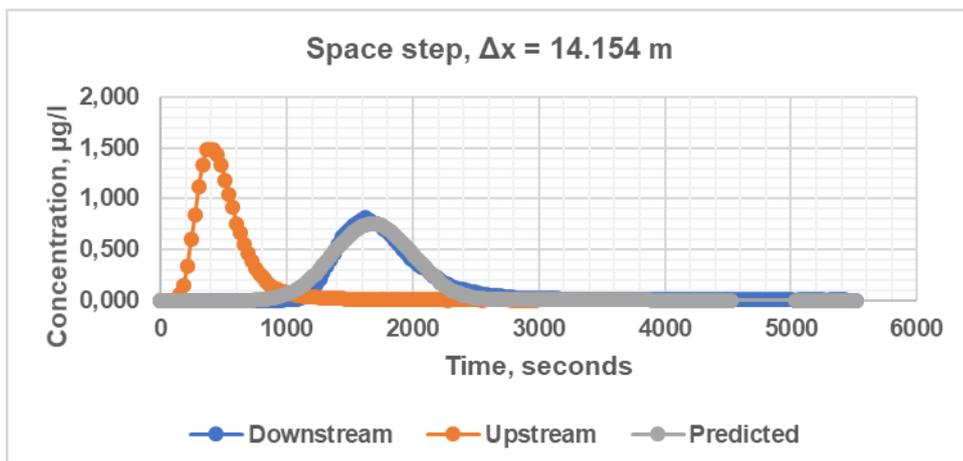
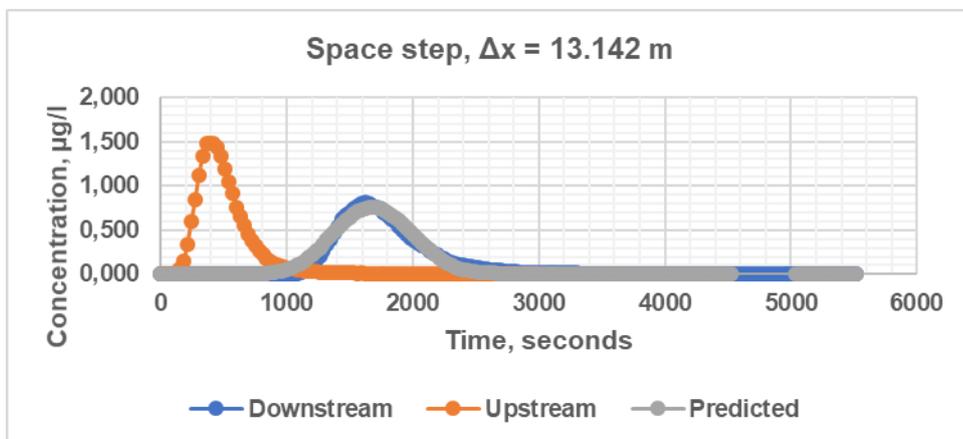
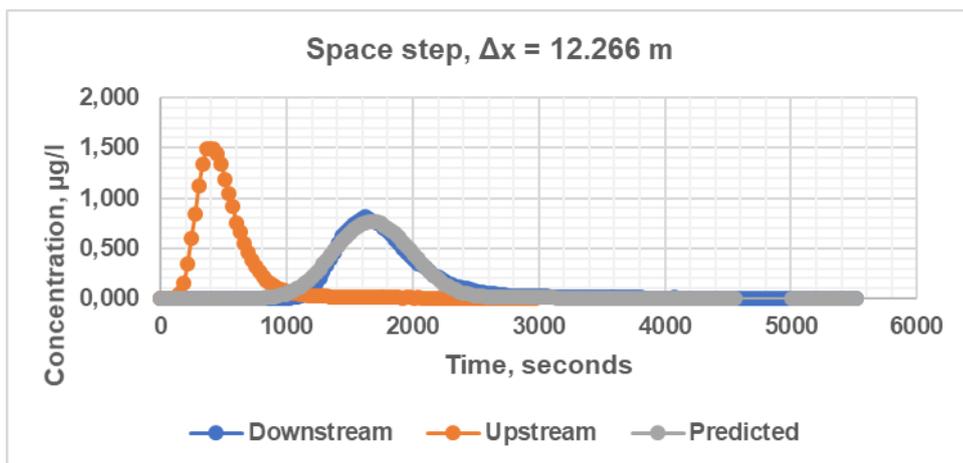
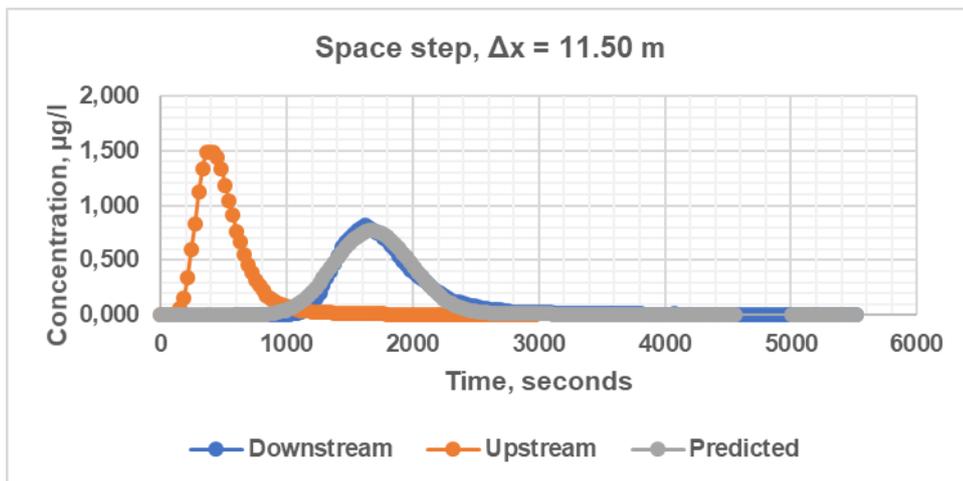


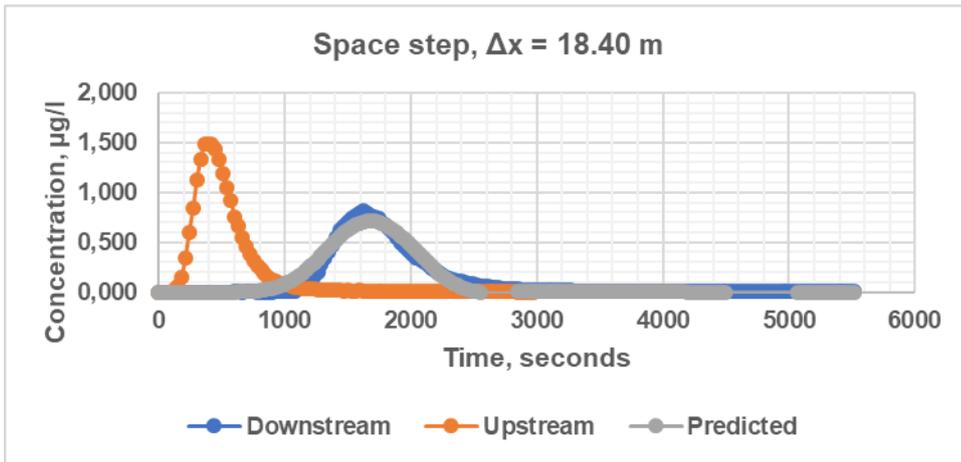
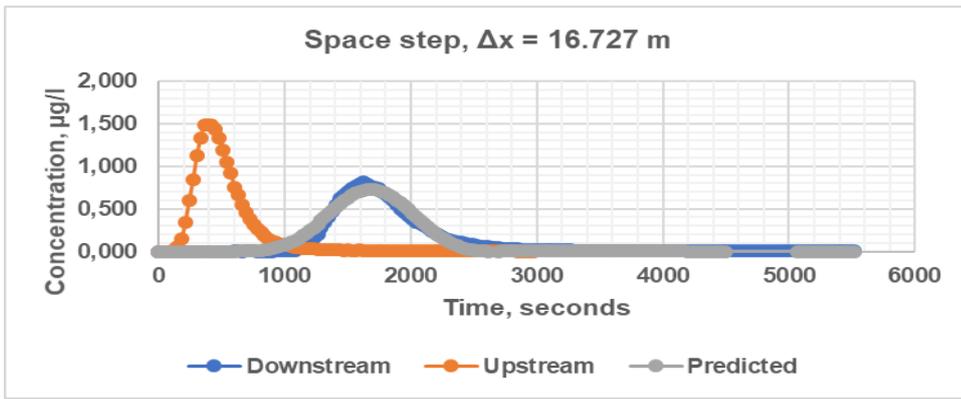


MacCormack simulations (experiment 4)

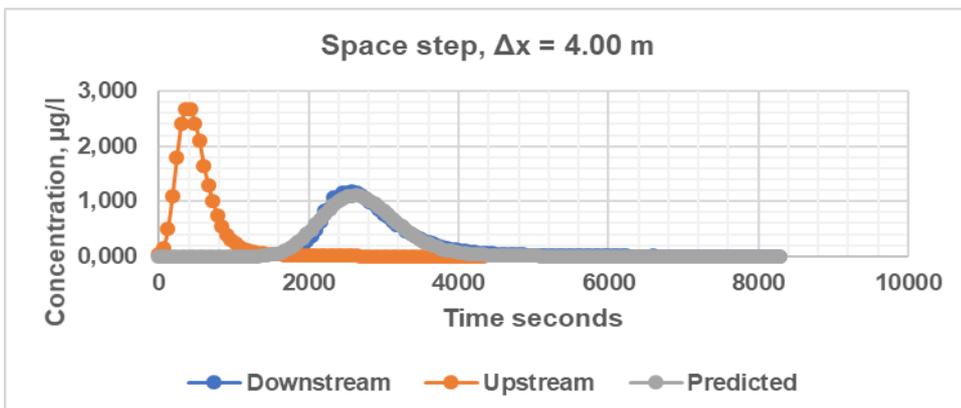
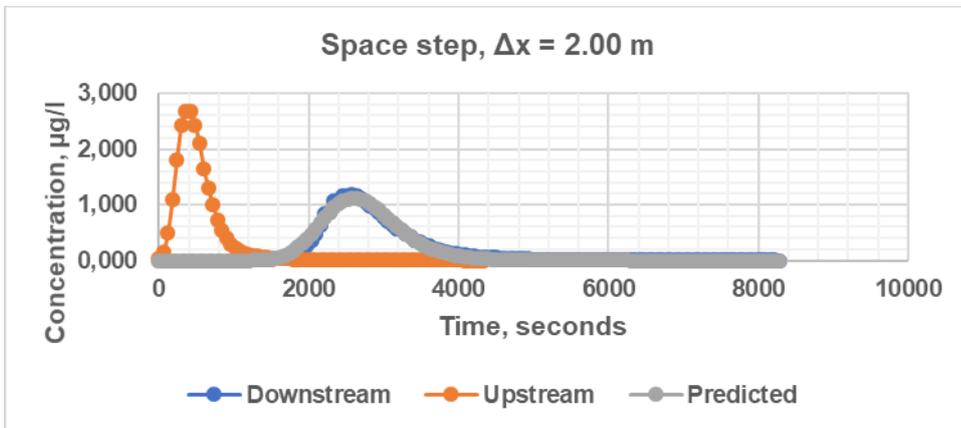


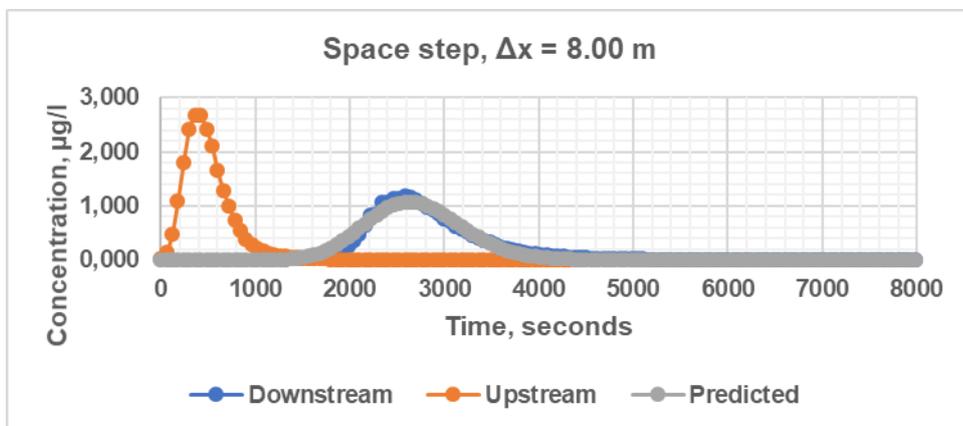
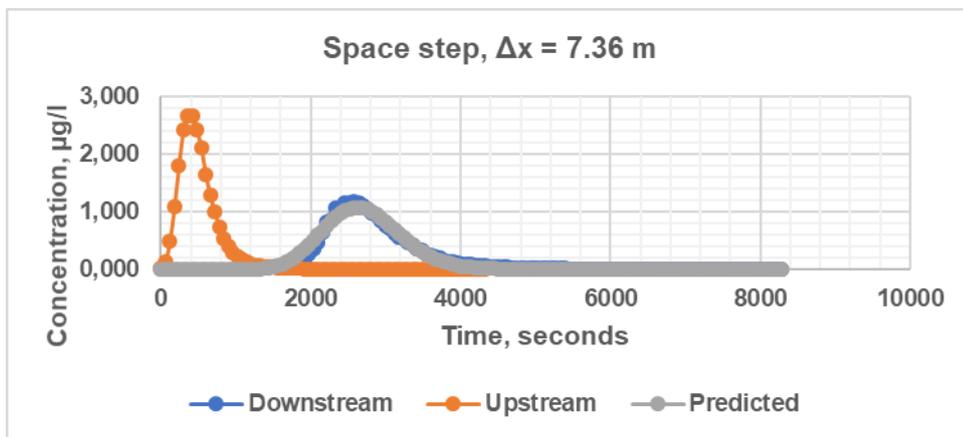
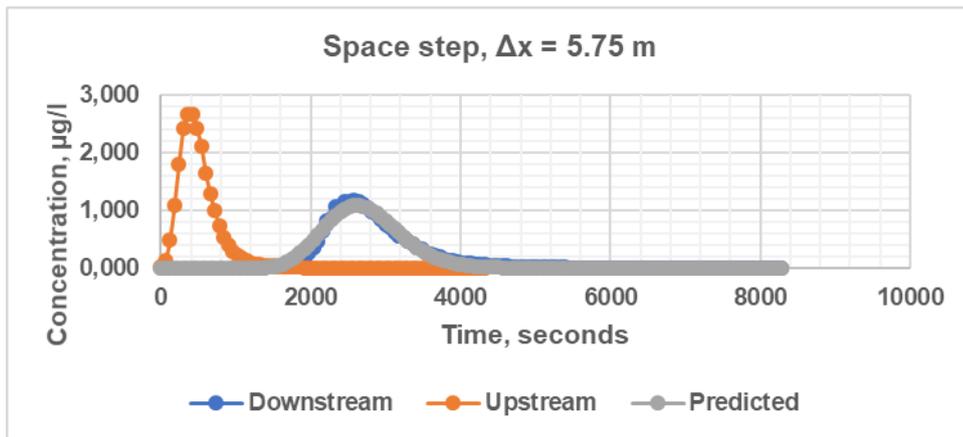
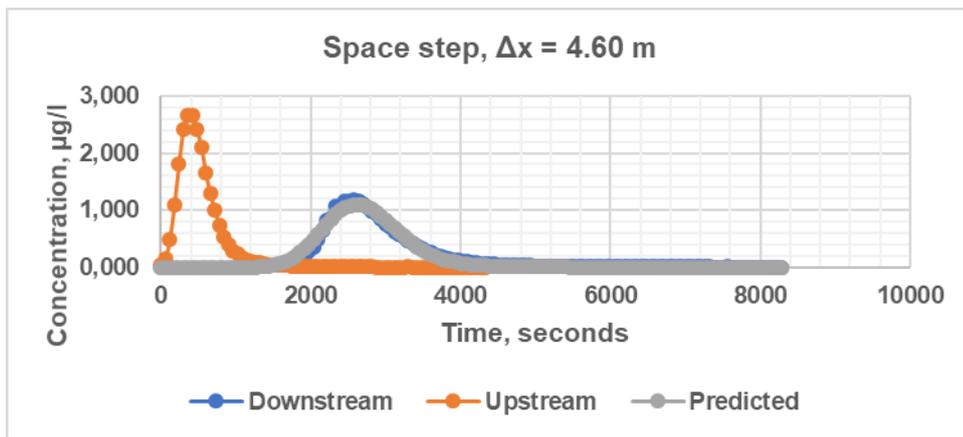


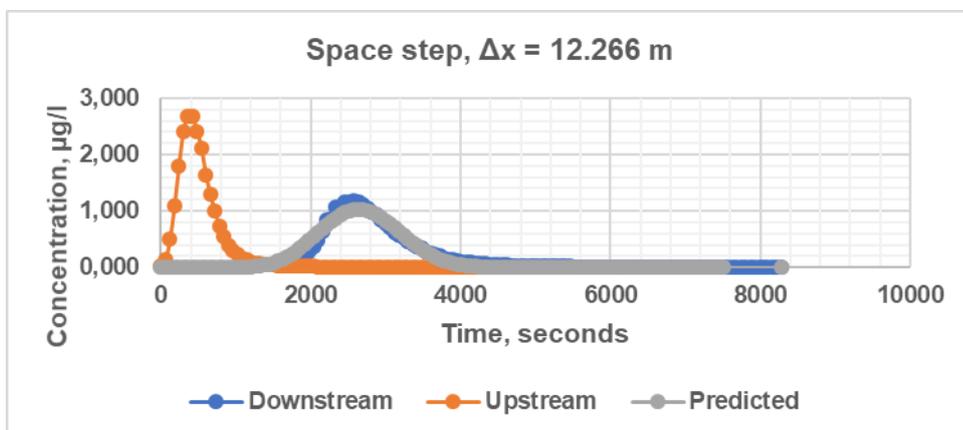
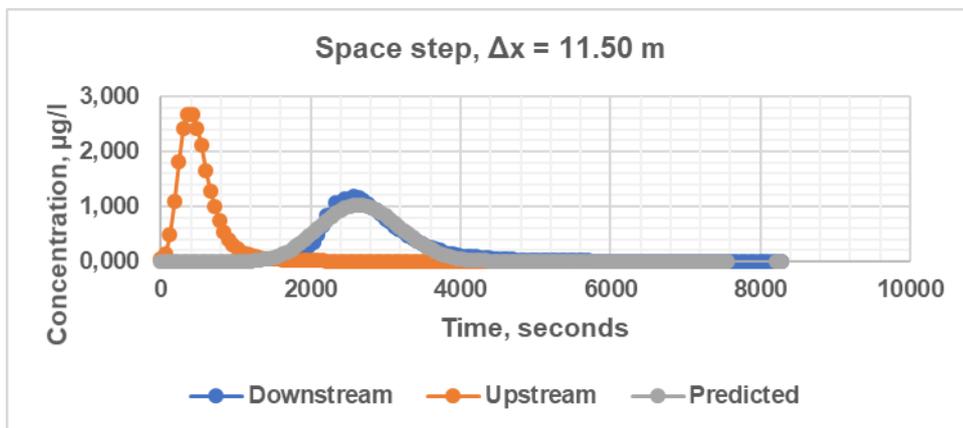
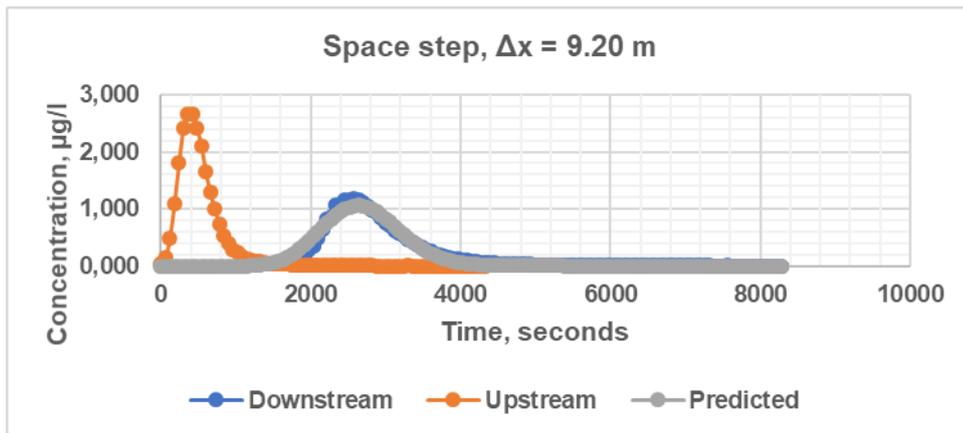
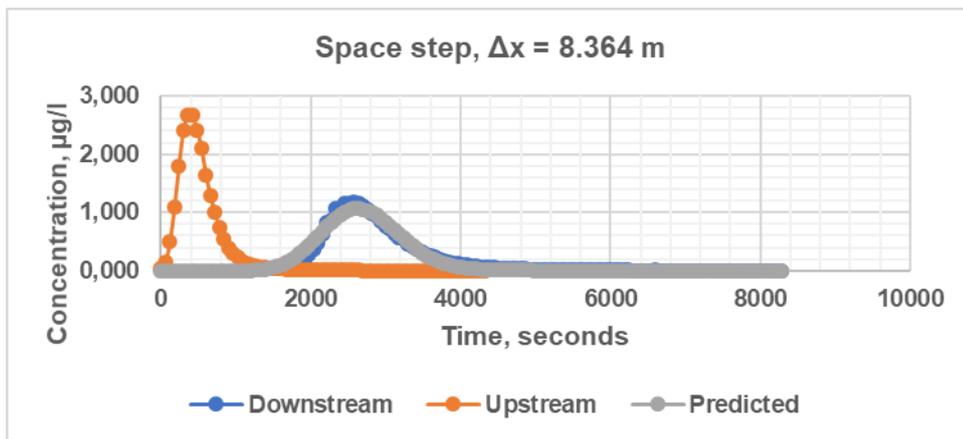


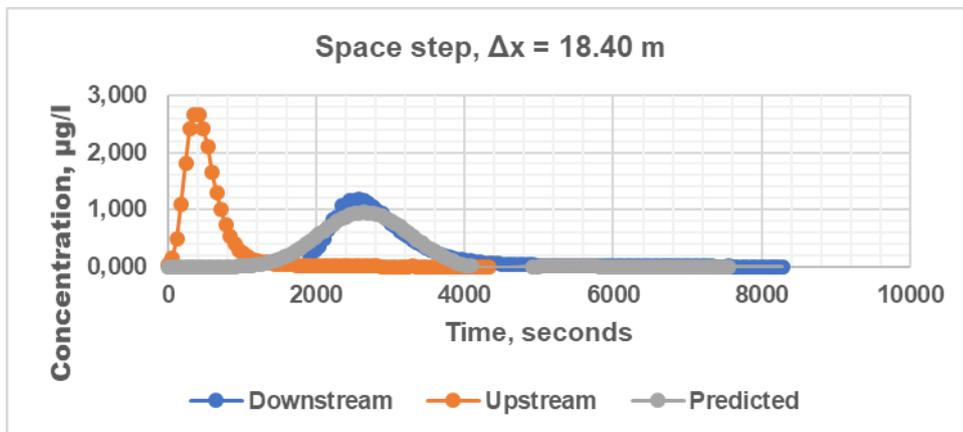
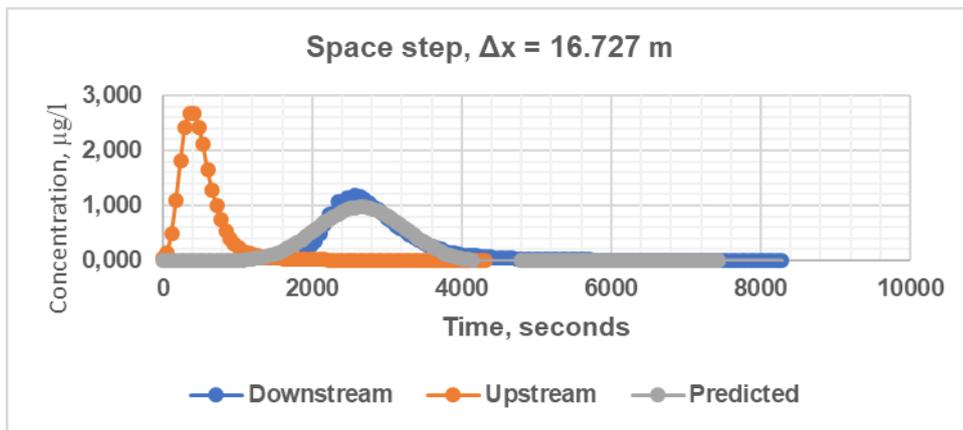
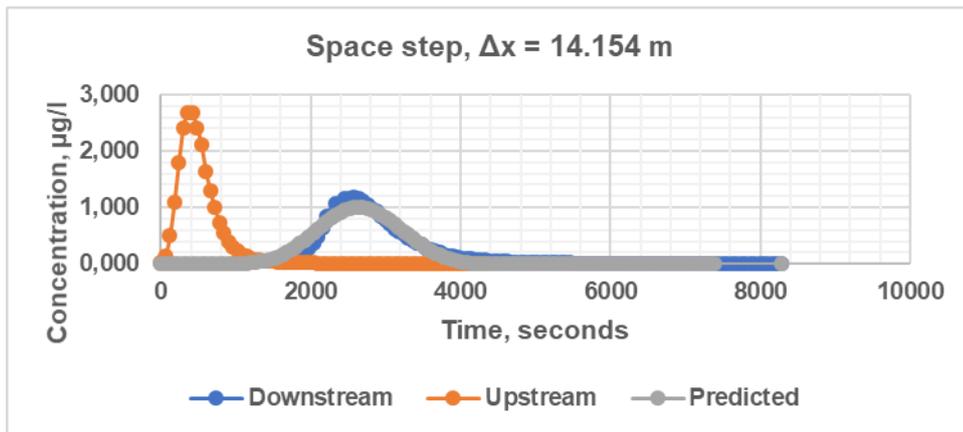
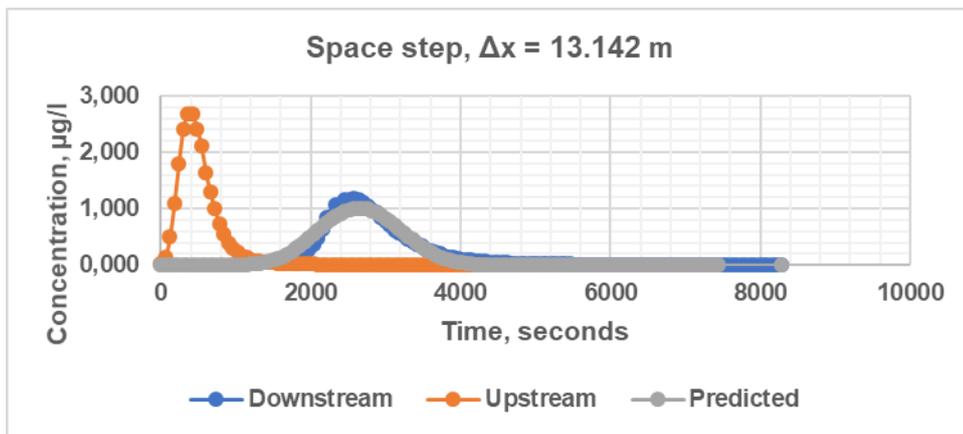


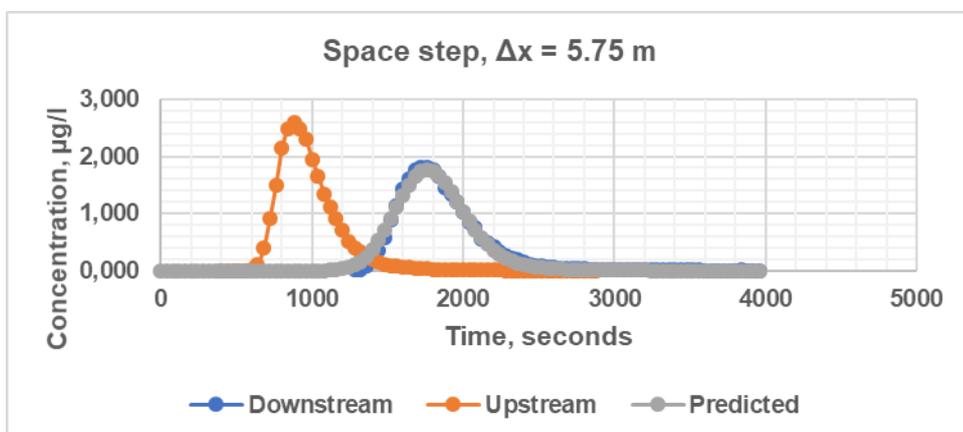
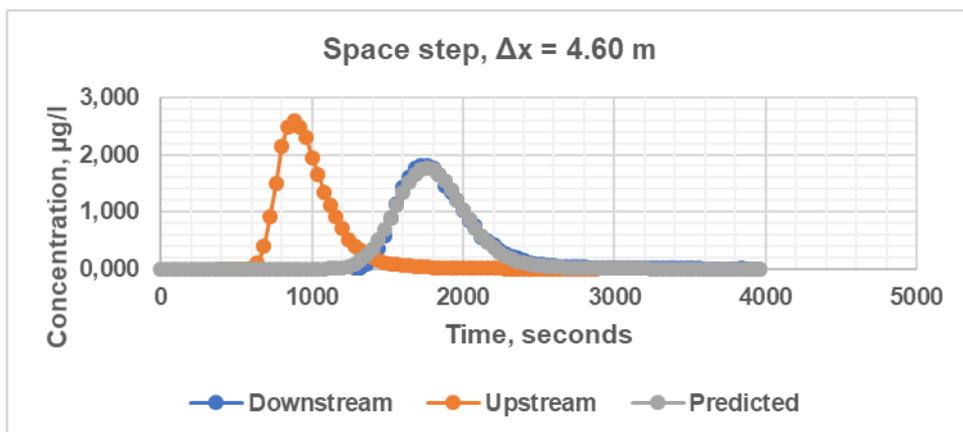
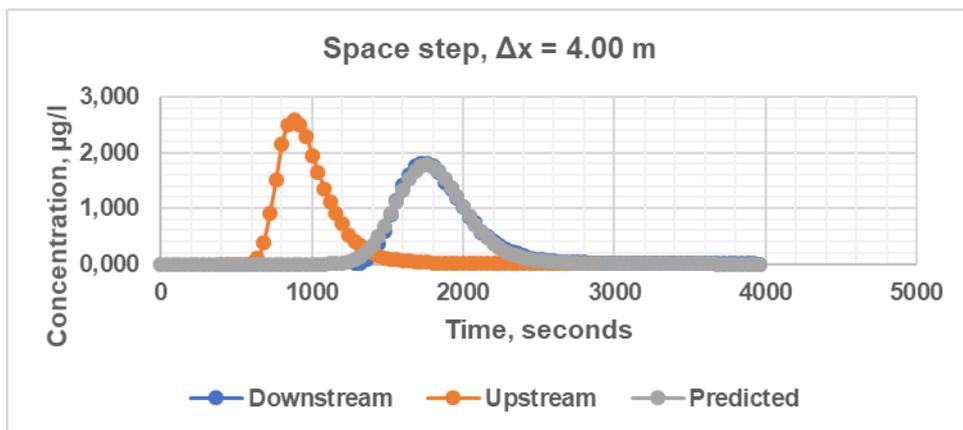
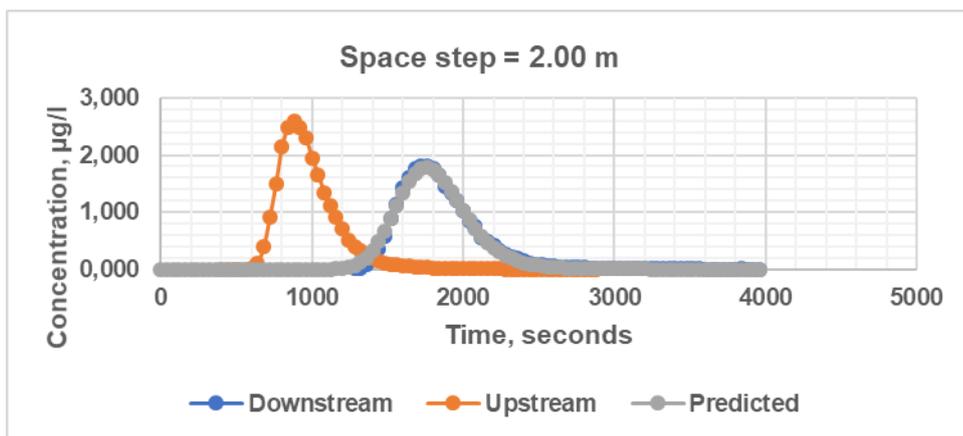
MacCormack simulations (experiment 9)

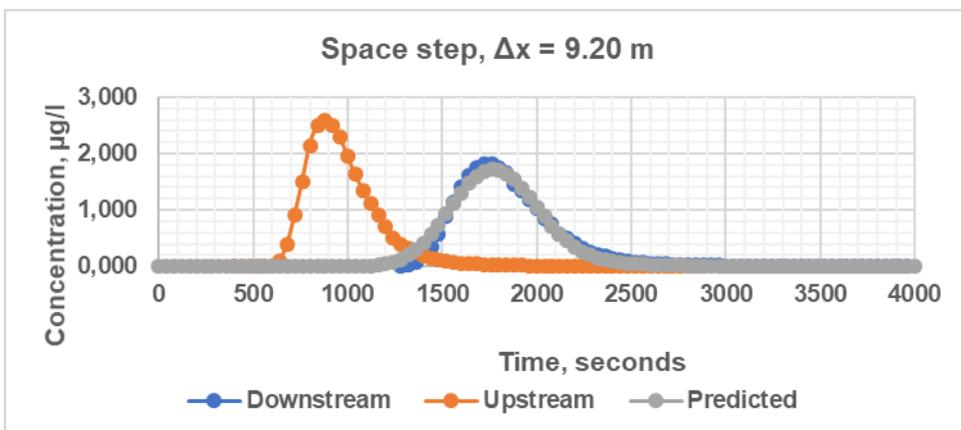
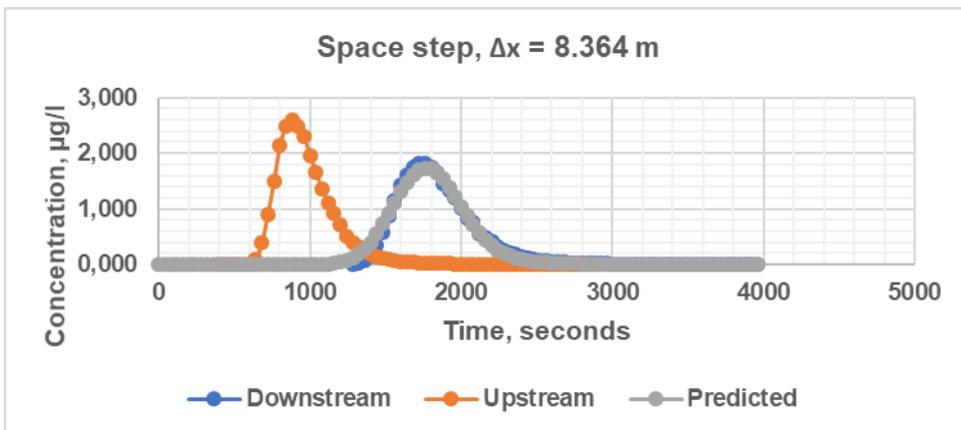
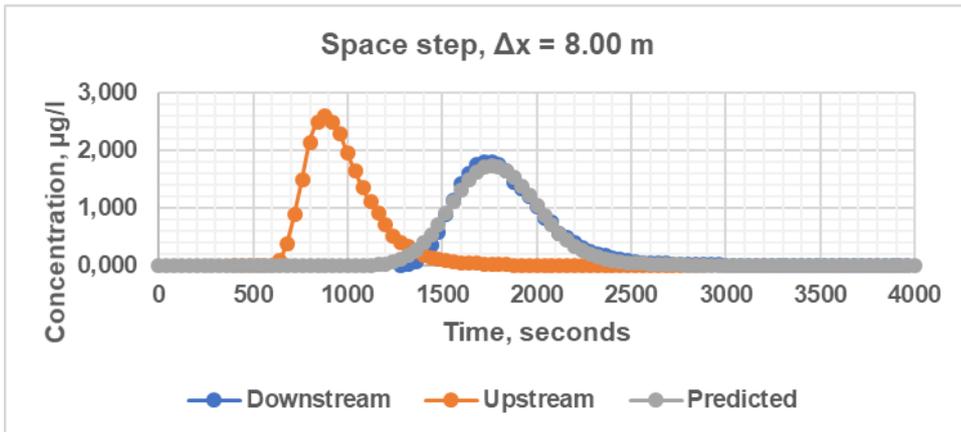
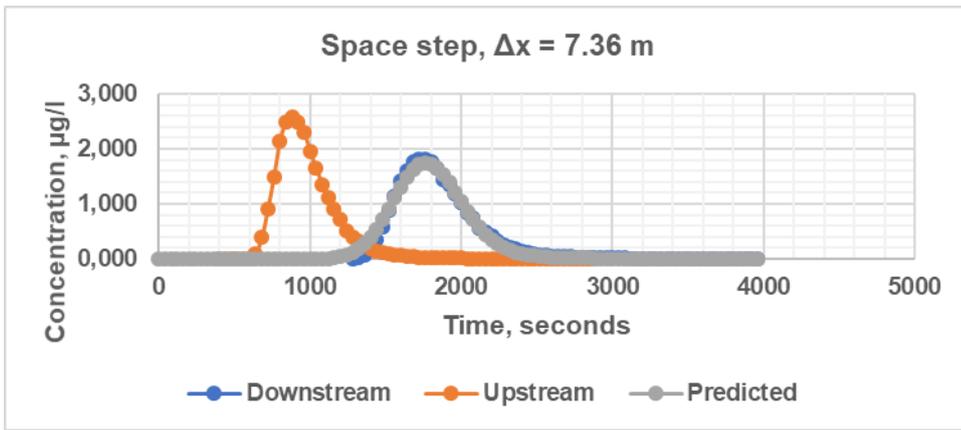


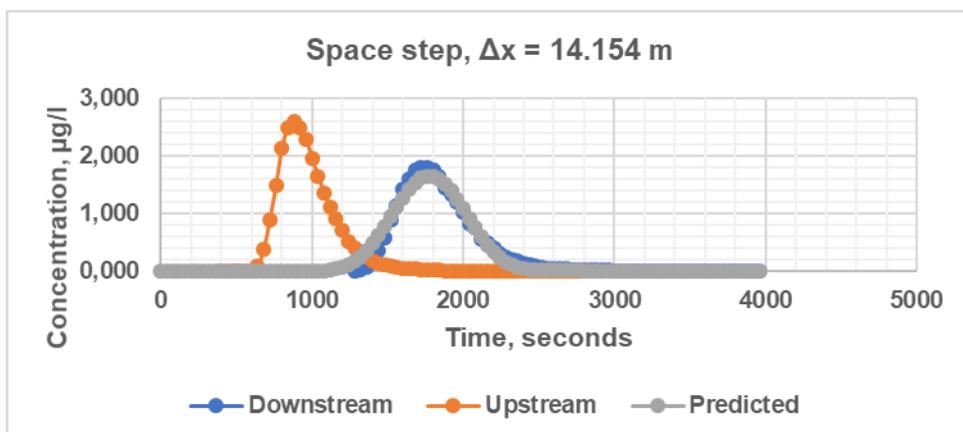
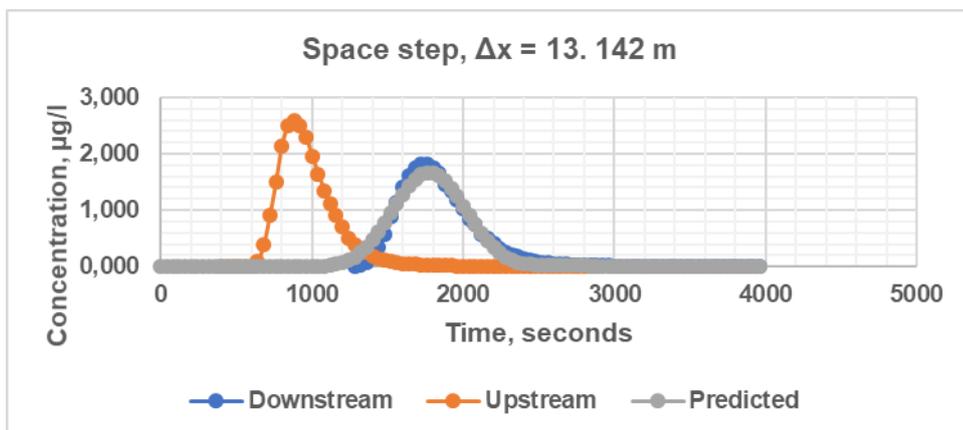
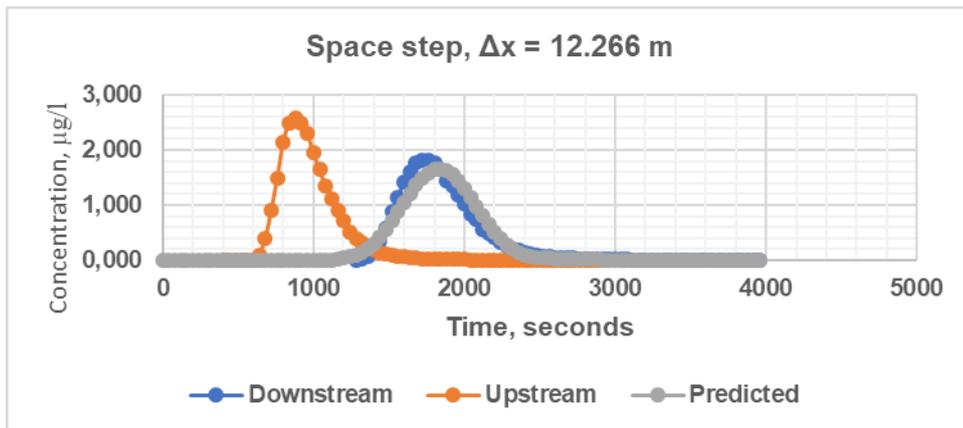
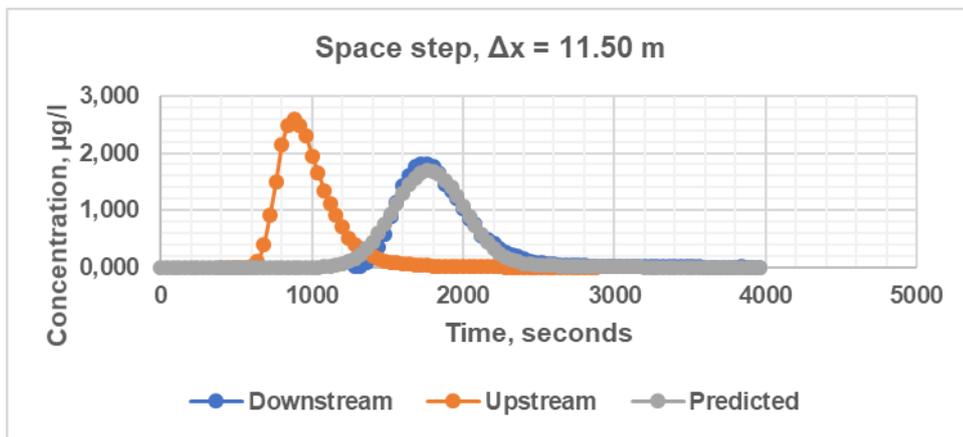


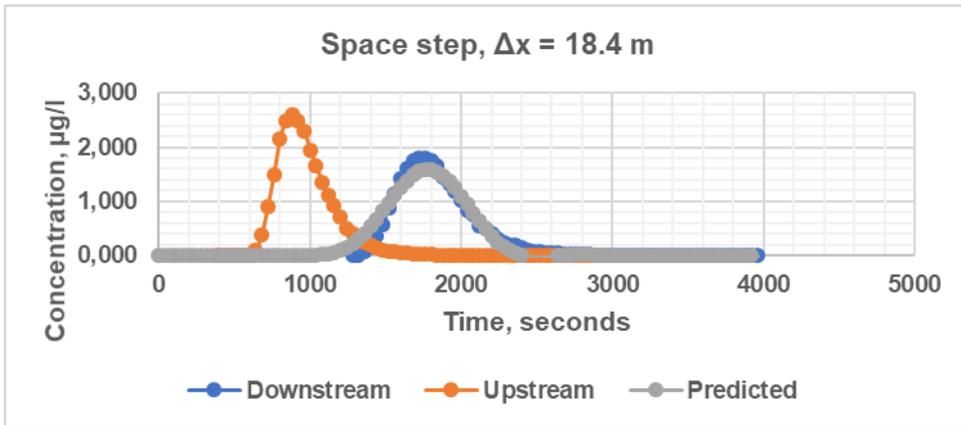
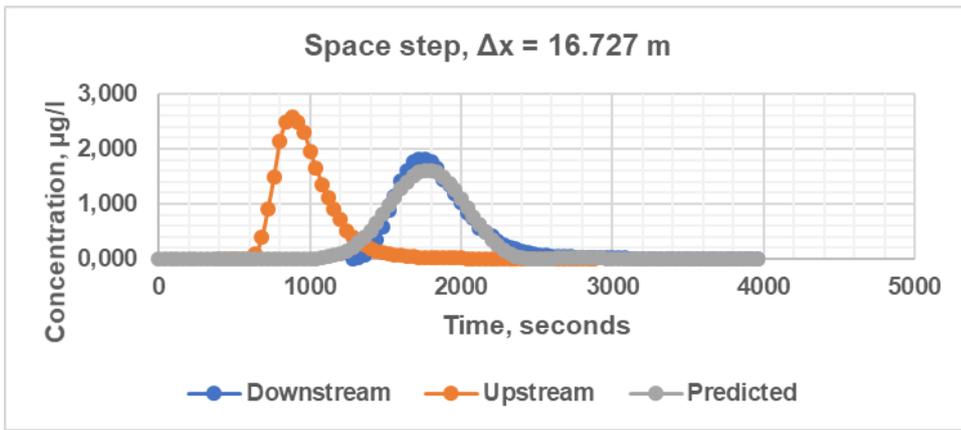




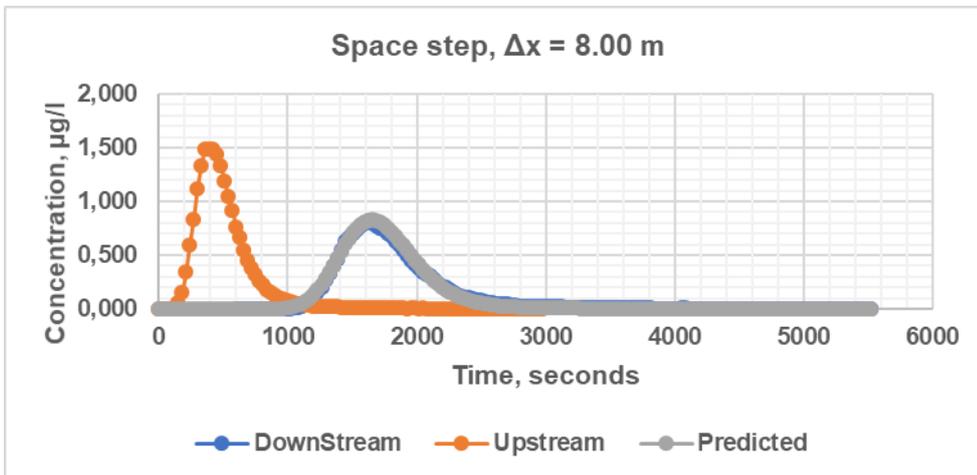


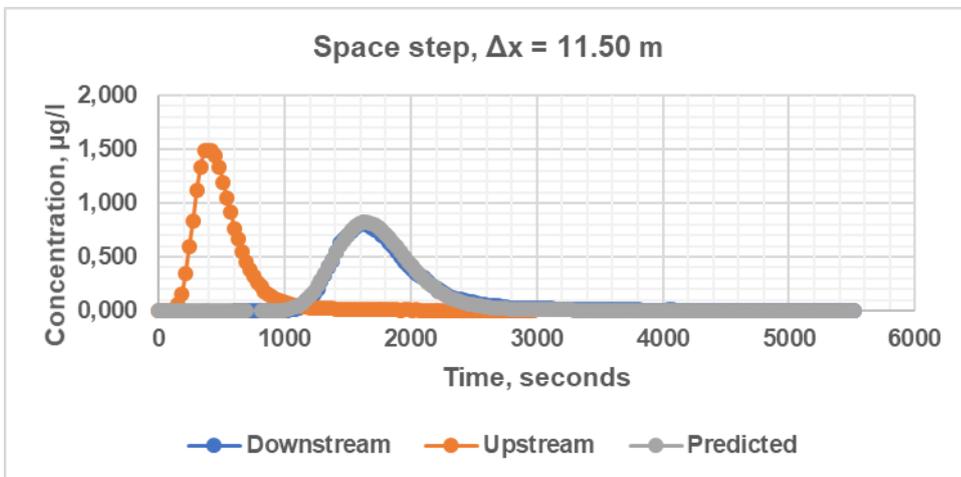
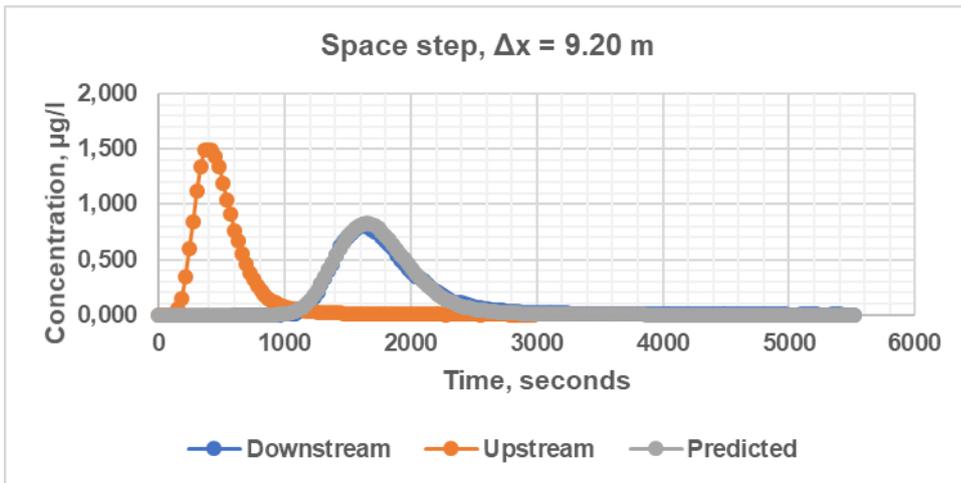
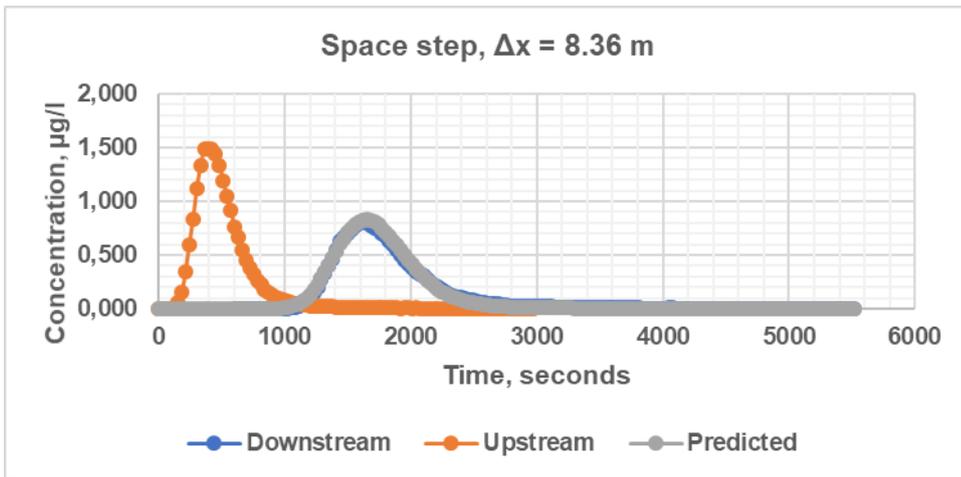


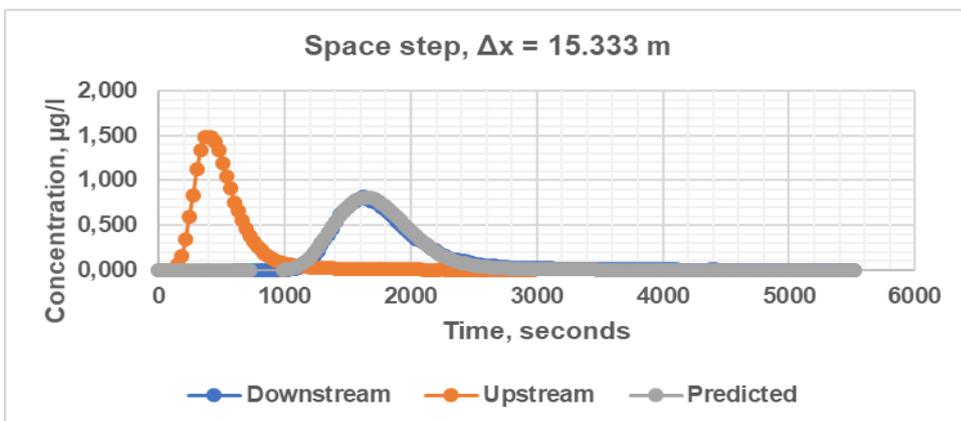
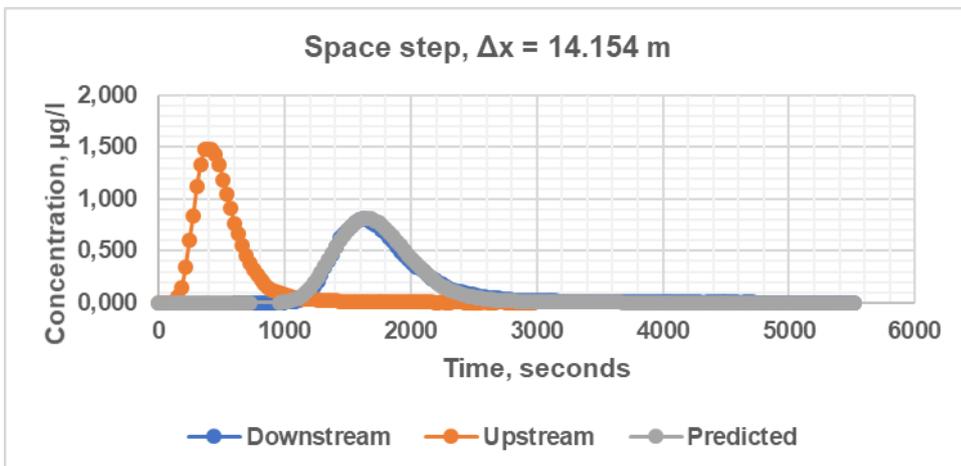
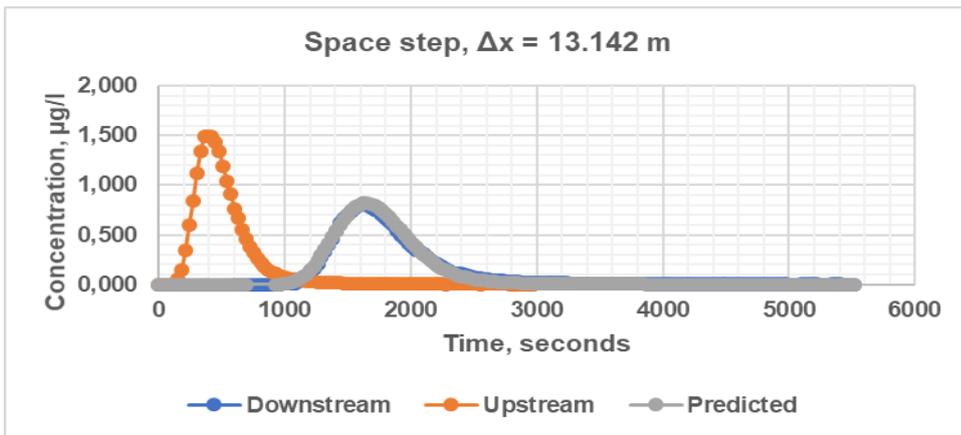
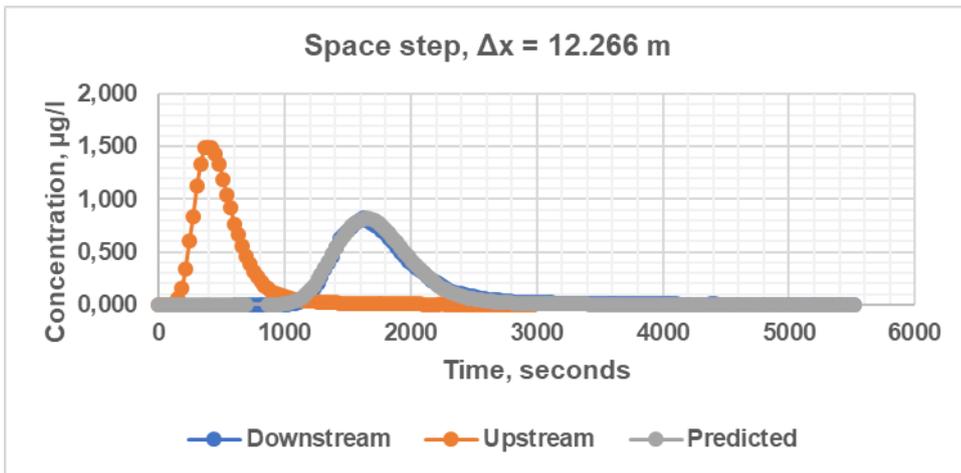


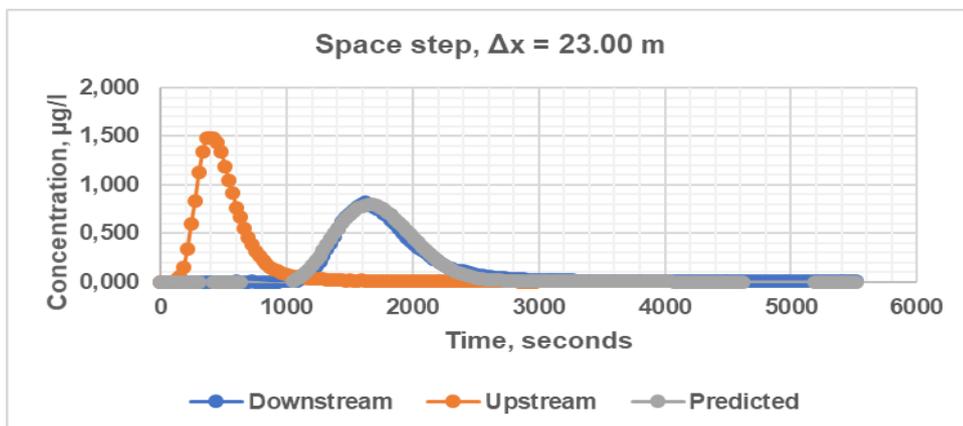
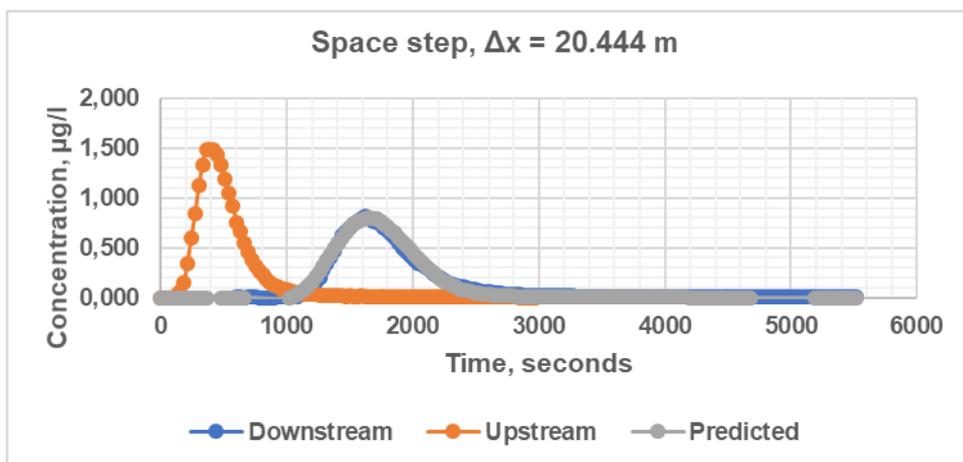
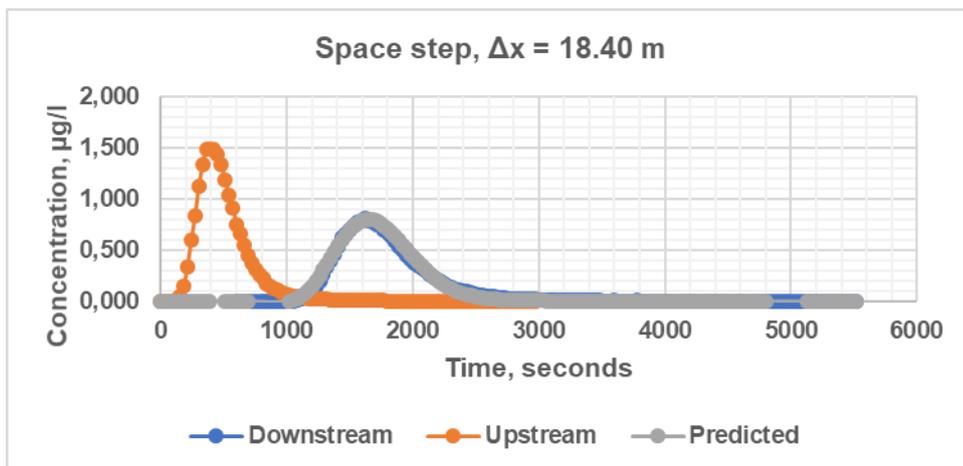
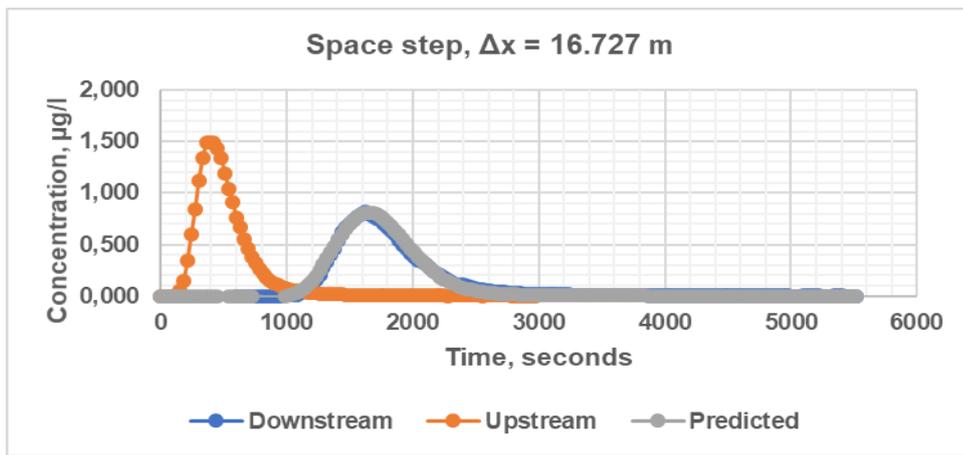


QUICKEST simulations (experiment 4)

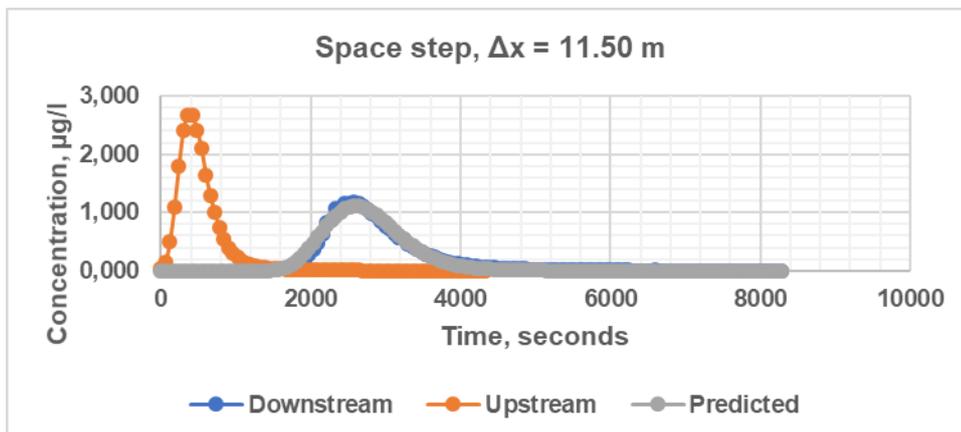
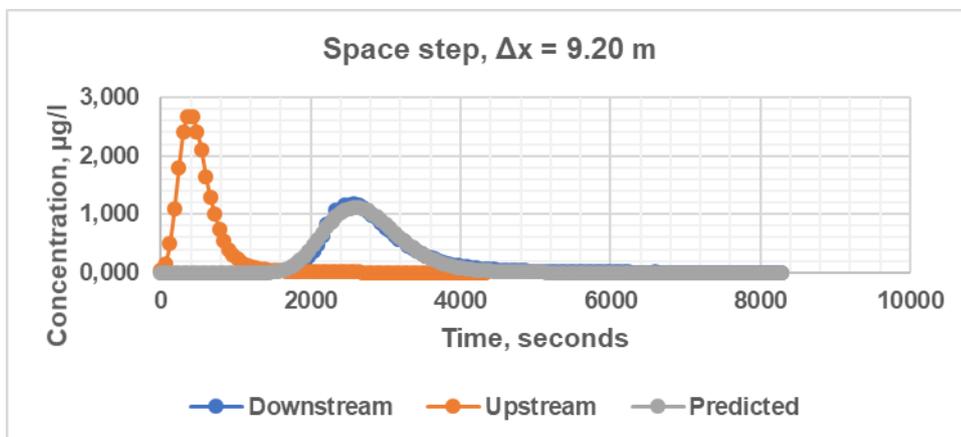
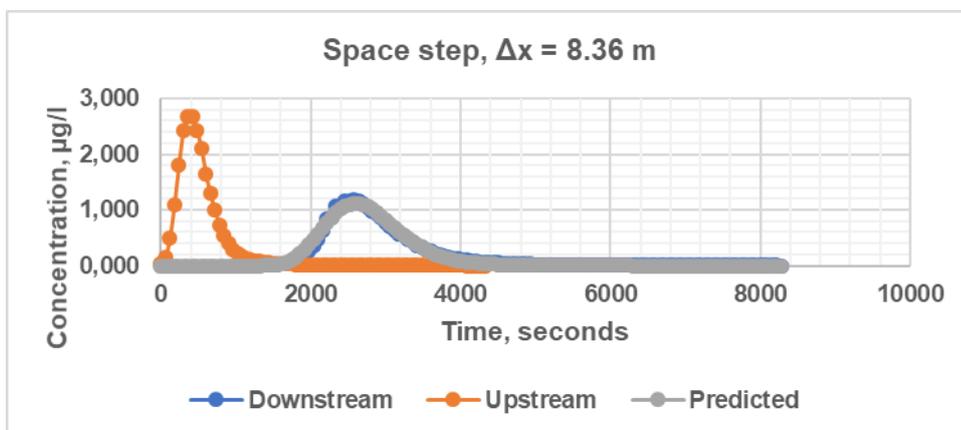
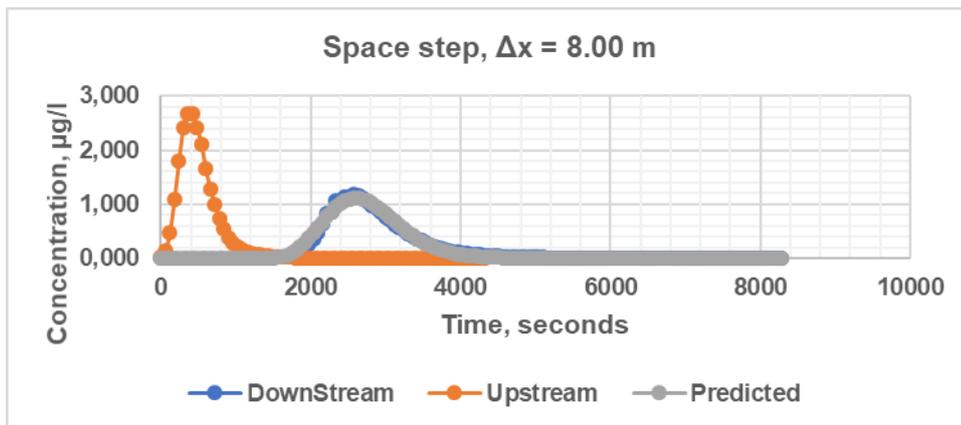


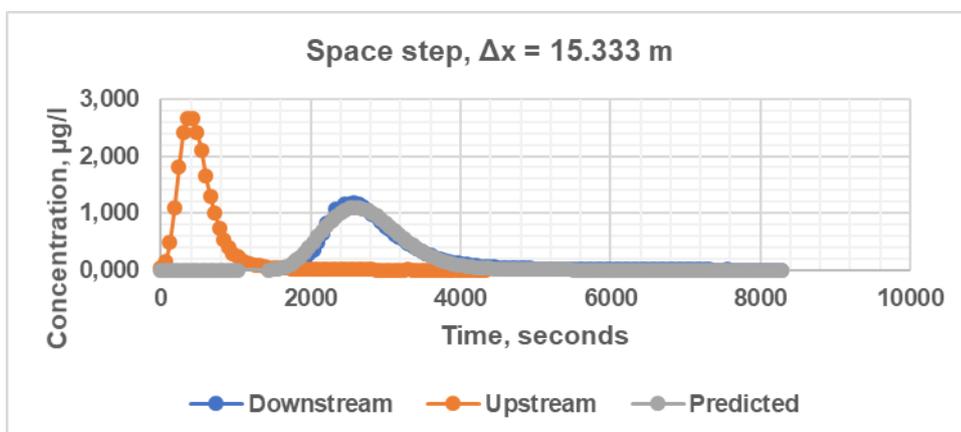
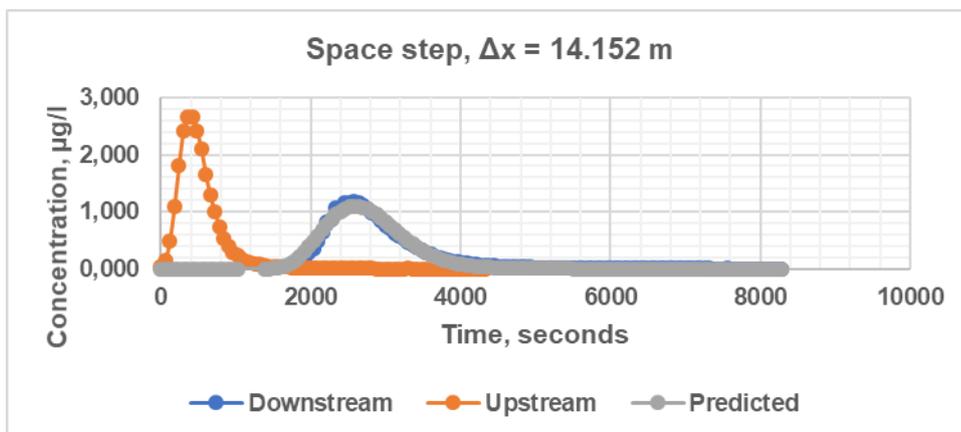
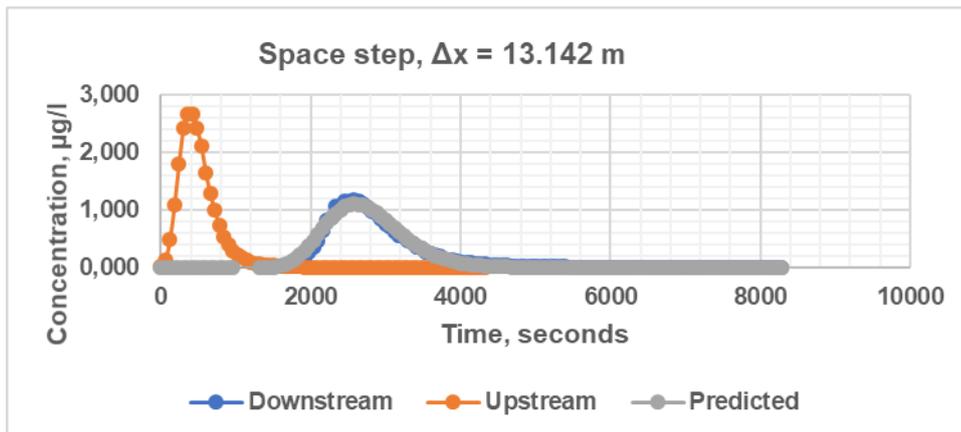
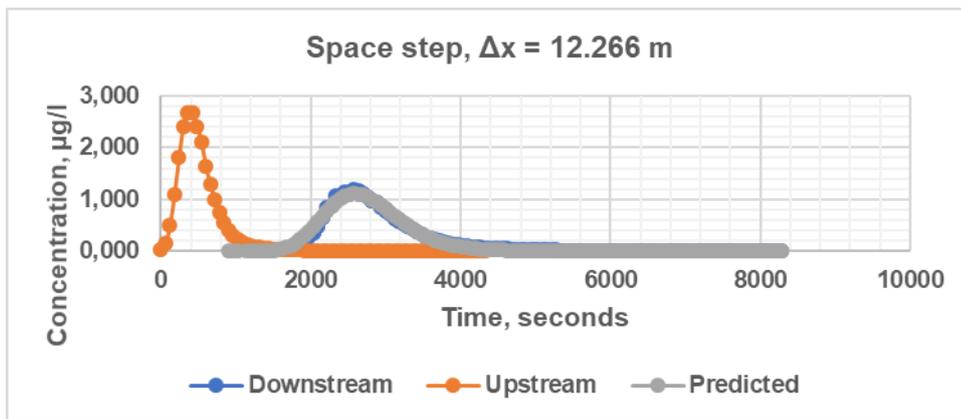


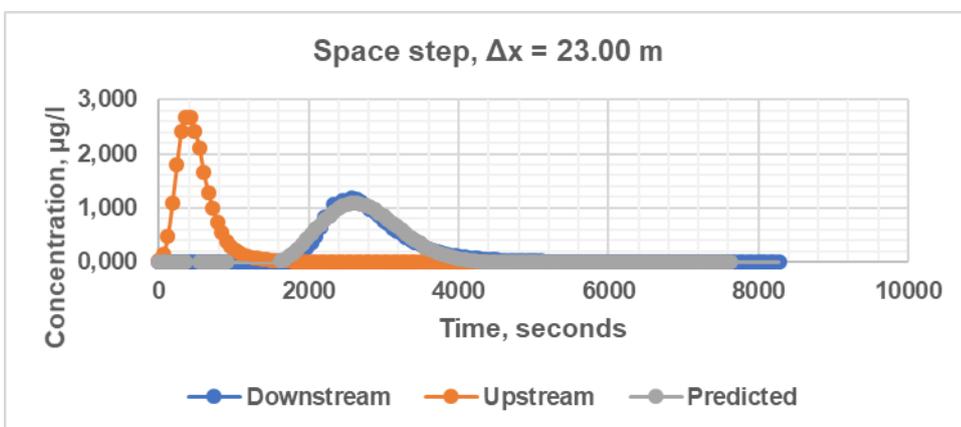
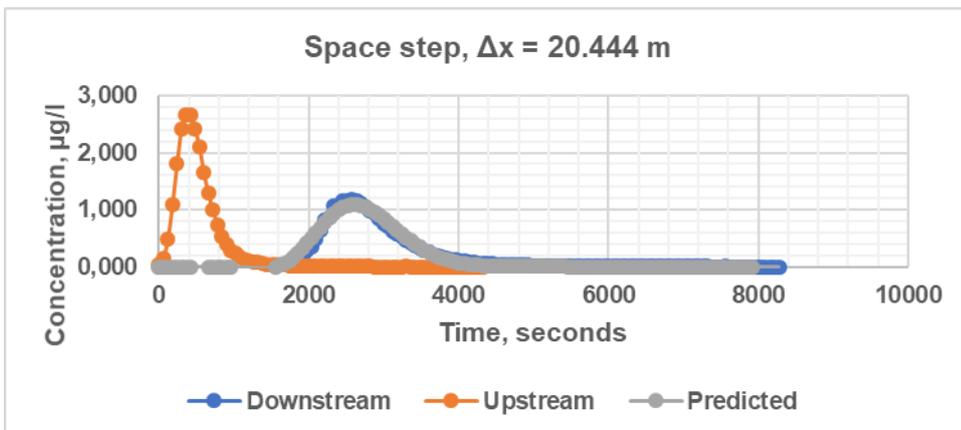
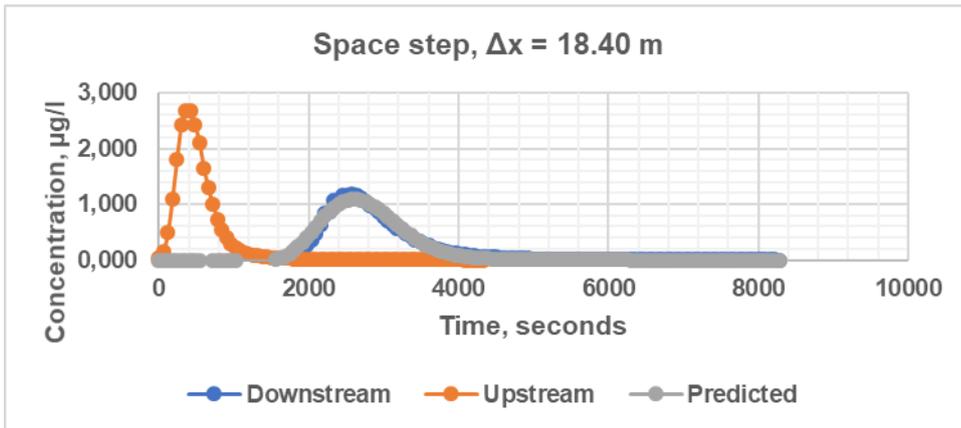
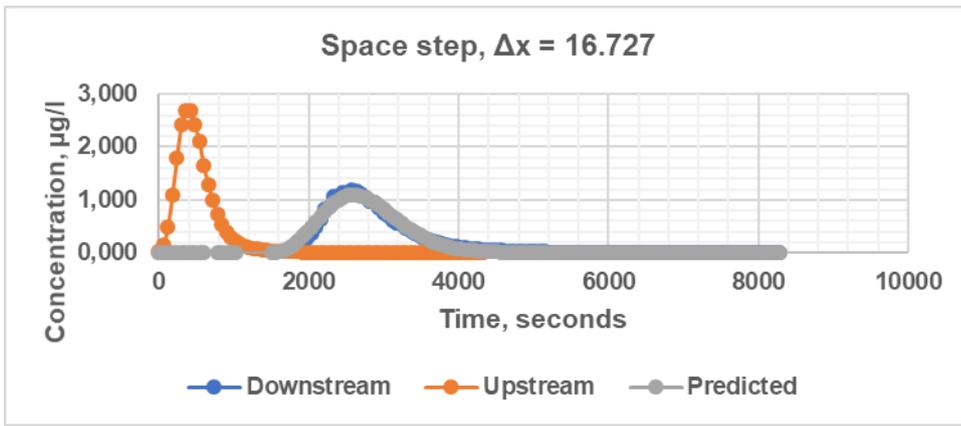




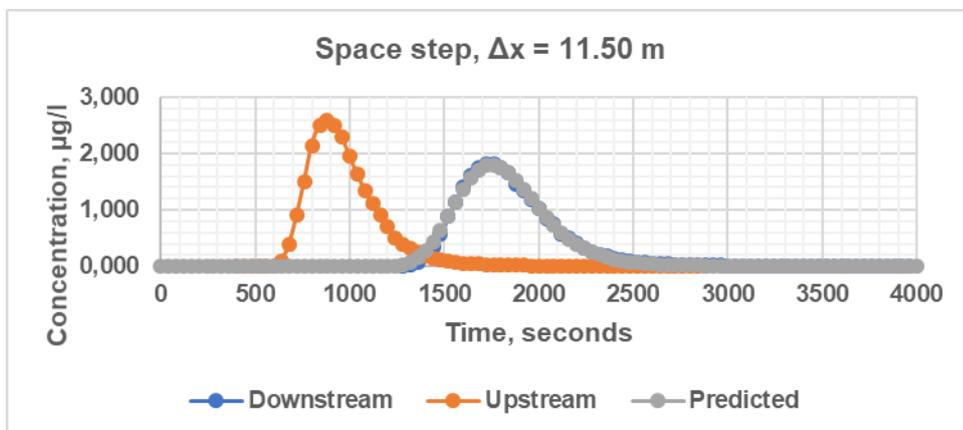
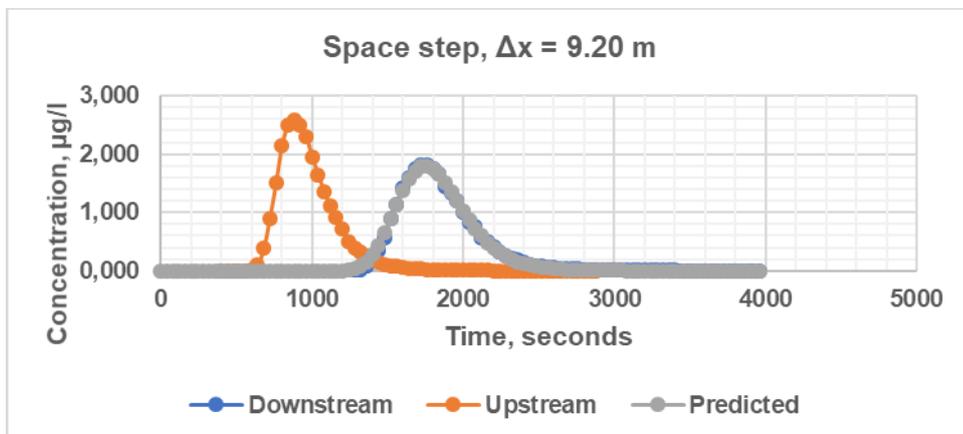
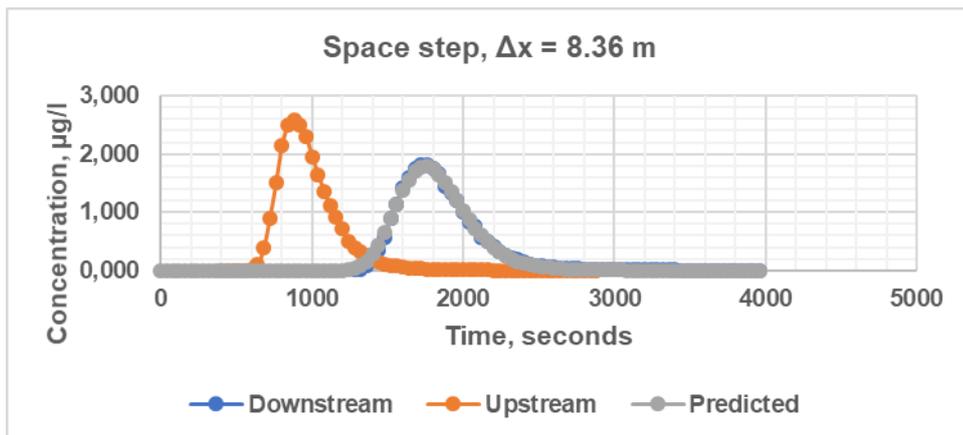
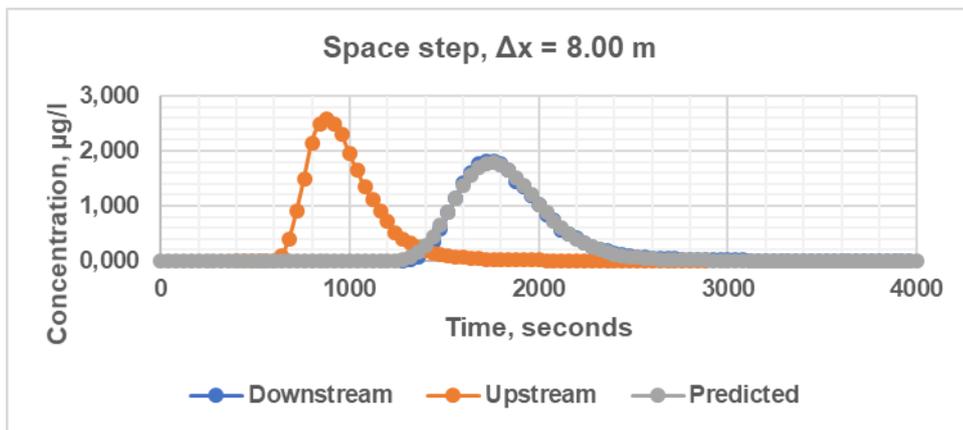
QUICKEST simulations (experiment 9)

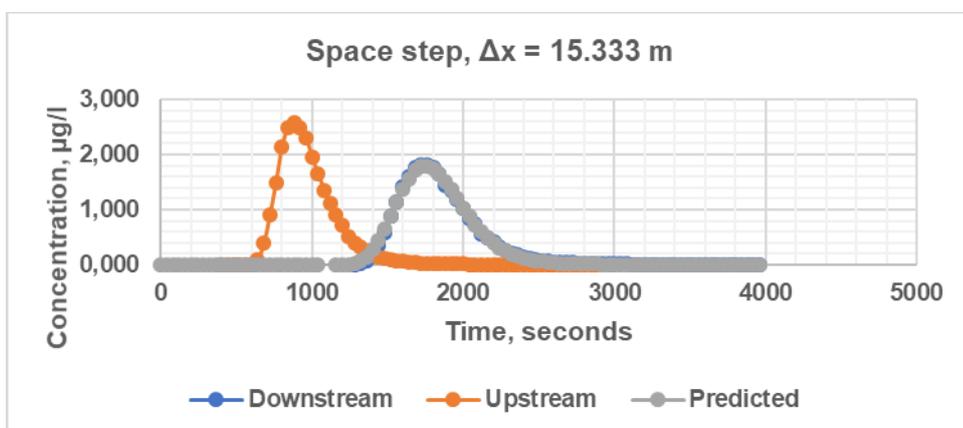
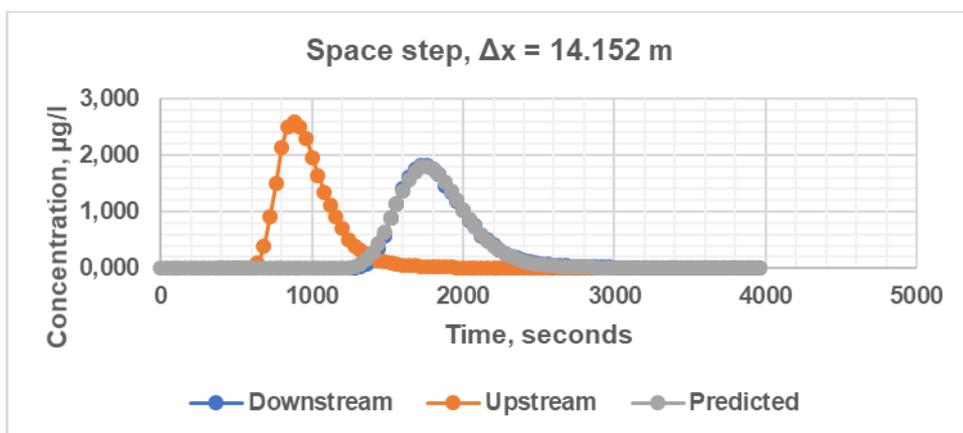
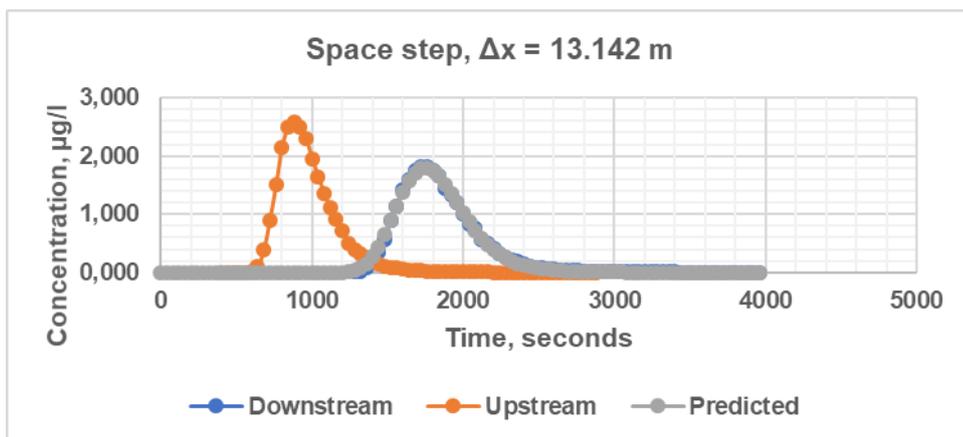
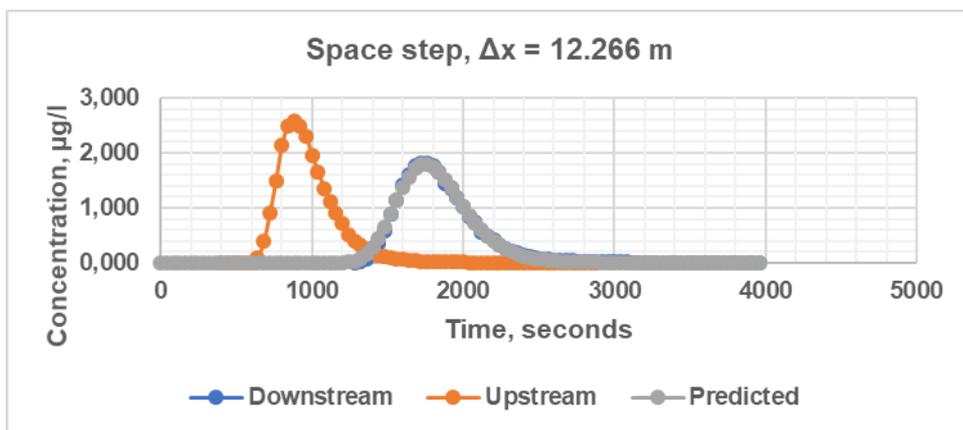


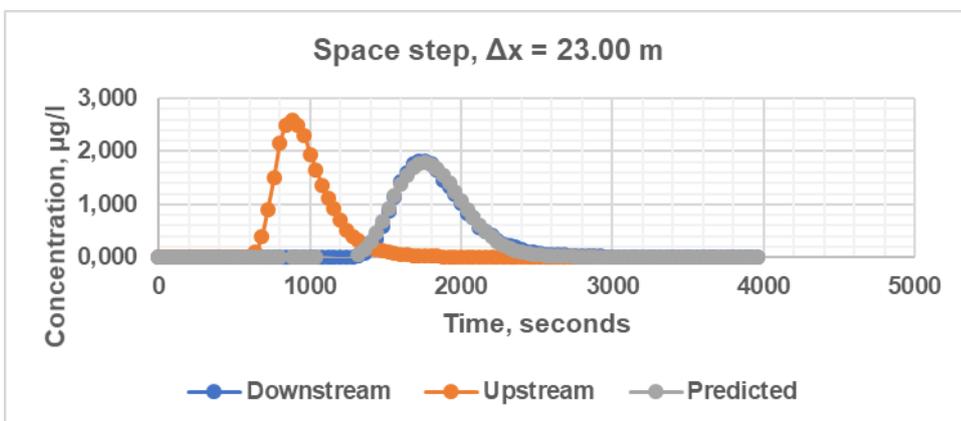
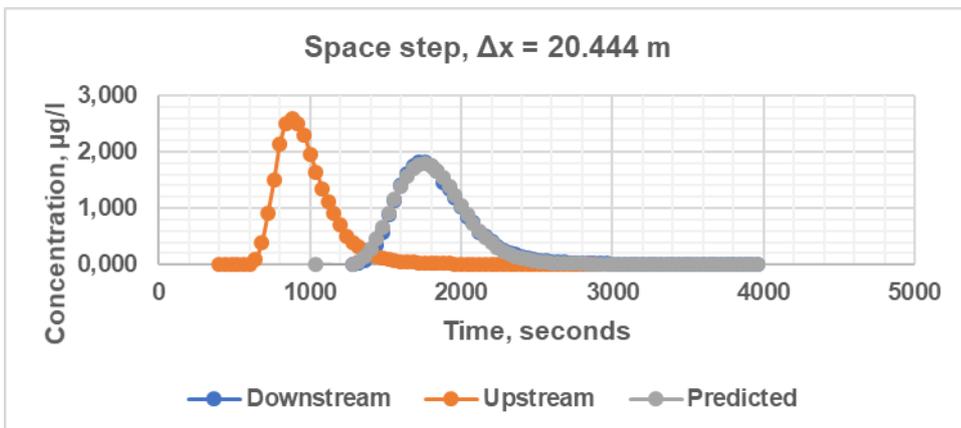
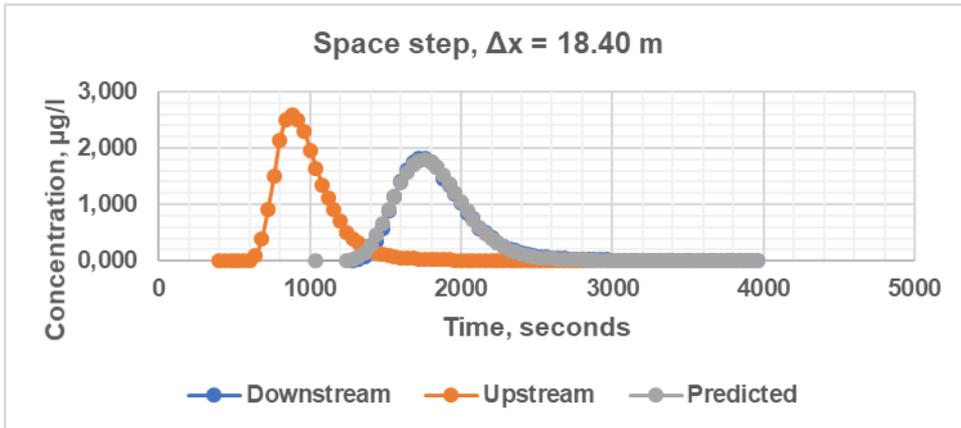
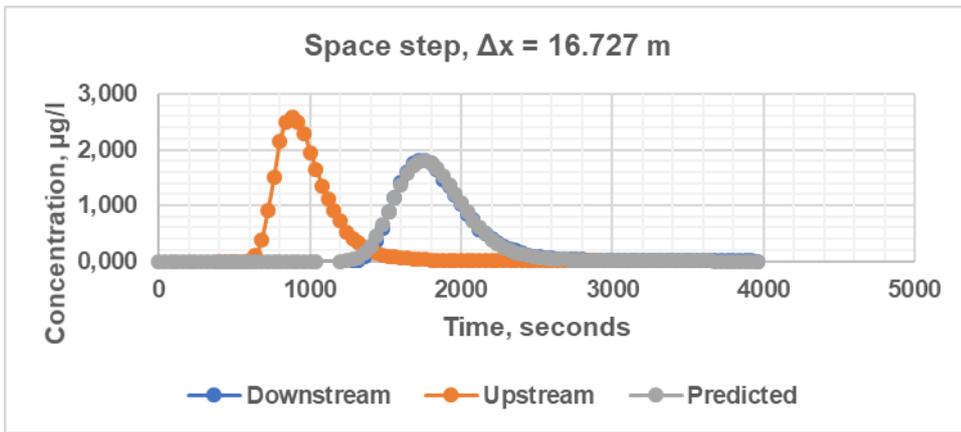




QUICKEST simulations (experiment 13)







Appendix F: Confirmation of routing method-based model

The routing method-based model was used to predict dispersion coefficients using evaluation data. The predicted dispersion coefficient was then applied with numerical methods to simulate downstream concentration profile. The quality of simulation was measured by residual sum of squares (RSS) and the coefficient of determination (R^2). Evaluation data consisted of experiments 4, 9 and 13.

Appendix F-1: Residual sum of squares between predicted and observed concentration profiles. Numerical methods were applied with predicted dispersion coefficients to predict concentration profiles. The tables show space steps, predicted dispersion coefficients, nondimensional parameters

and residual sum of squares (RSS).

Crank-Nicolson method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, (μg/l) ²
2.00	30	0.448	0.147	2.201	3.361	0.655	0.08369
4.00	30	0.448	0.147	1.101	0.840	1.310	0.09392
4.60	30	0.448	0.147	0.957	0.635	1.506	0.09860
5.75	30	0.448	0.147	0.766	0.407	1.883	0.10998
7.36	30	0.448	0.147	0.598	0.248	2.410	0.13204
8.00	30	0.448	0.147	0.550	0.210	2.620	0.14307
8.36	30	0.448	0.147	0.526	0.192	2.739	0.15021
9.20	30	0.448	0.147	0.478	0.159	3.013	0.16762
11.50	30	0.448	0.147	0.383	0.102	3.766	0.23003
12.27	30	0.448	0.147	0.359	0.089	4.017	0.25541
13.14	30	0.448	0.147	0.335	0.078	4.304	0.28823
14.15	30	0.448	0.147	0.311	0.067	4.635	0.32952
16.73	30	0.448	0.147	0.263	0.048	5.477	0.45272
18.40	30	0.448	0.147	0.239	0.040	6.025	0.54593

Crank-Nicolson method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, (μg/l) ²
2.00	60	0.266	0.081	2.436	3.985	0.611	0.23999
4.00	60	0.266	0.081	1.218	0.996	1.223	0.27179
4.60	60	0.266	0.081	1.059	0.753	1.406	0.28602
5.75	60	0.266	0.081	0.847	0.482	1.758	0.31999
7.36	60	0.266	0.081	0.662	0.294	2.250	0.38368
8.00	60	0.266	0.081	0.609	0.249	2.446	0.41461
8.36	60	0.266	0.081	0.583	0.228	2.557	0.43410
9.20	60	0.266	0.081	0.530	0.188	2.812	0.48170
11.50	60	0.266	0.081	0.424	0.121	3.515	0.64436
12.27	60	0.266	0.081	0.397	0.106	3.750	0.70779
13.14	60	0.266	0.081	0.371	0.092	4.017	0.78680
14.15	60	0.266	0.081	0.344	0.080	4.327	0.88681
16.73	60	0.266	0.081	0.291	0.057	5.113	1.17098
18.40	60	0.266	0.081	0.265	0.047	5.625	1.37764

Crank-Nicolson method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, (μg/l) ²
2.00	40	0.606	0.212	4.248	6.056	0.701	0.25028
4.00	40	0.606	0.212	2.124	1.514	1.403	0.27945
4.60	40	0.606	0.212	1.847	1.145	1.613	0.29245
5.75	40	0.606	0.212	1.478	0.733	2.017	0.32343
7.36	40	0.606	0.212	1.154	0.447	2.581	0.38164
8.00	40	0.606	0.212	1.062	0.378	2.806	0.41008
8.36	40	0.606	0.212	1.016	0.346	2.933	0.42842
9.20	40	0.606	0.212	0.923	0.286	3.227	0.47236
11.50	40	0.606	0.212	0.739	0.183	4.033	0.63344
12.27	40	0.606	0.212	0.693	0.161	4.302	0.68960
13.14	40	0.606	0.212	0.646	0.140	4.609	0.76873
14.15	40	0.606	0.212	0.600	0.121	4.964	0.87189
16.73	40	0.606	0.212	0.508	0.087	5.866	1.17743
18.40	40	0.606	0.212	0.462	0.072	6.453	1.41247

MacCormack method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.000	30	0.448	0.147	2.201	3.361	0.655	0.09566
4.000	30	0.448	0.147	1.101	0.840	1.310	0.12083
4.600	30	0.448	0.147	0.957	0.635	1.506	0.13075
5.750	30	0.448	0.147	0.766	0.407	1.883	0.15324
7.360	30	0.448	0.147	0.598	0.248	2.410	0.19316
8.000	30	0.448	0.147	0.550	0.210	2.620	0.21195
8.364	30	0.448	0.147	0.526	0.192	2.739	0.22369
9.200	30	0.448	0.147	0.478	0.159	3.013	0.25186
11.500	30	0.448	0.147	0.383	0.102	3.766	0.34557
12.266	30	0.448	0.147	0.359	0.089	4.017	0.38123
13.142	30	0.448	0.147	0.335	0.078	4.304	0.42527
14.154	30	0.448	0.147	0.311	0.067	4.635	0.48052
16.727	30	0.448	0.147	0.263	0.048	5.477	0.63435
18.400	30	0.448	0.147	0.239	0.040	6.025	0.74448

MacCormack method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.000	60	0.266	0.081	2.436	3.985	0.611	0.28017
4.000	60	0.266	0.081	1.218	0.996	1.223	0.35892
4.600	60	0.266	0.081	1.059	0.753	1.406	0.38876
5.750	60	0.266	0.081	0.847	0.482	1.758	0.45448
7.360	60	0.266	0.081	0.662	0.294	2.250	0.56587
8.000	60	0.266	0.081	0.609	0.249	2.446	0.61645
8.364	60	0.266	0.081	0.583	0.228	2.557	0.64734
9.200	60	0.266	0.081	0.530	0.188	2.812	0.72081
11.500	60	0.266	0.081	0.424	0.121	3.515	0.95350
12.266	60	0.266	0.081	0.397	0.106	3.750	1.03908
13.142	60	0.266	0.081	0.371	0.092	4.017	1.14228
14.154	60	0.266	0.081	0.344	0.080	4.327	1.26878
16.727	60	0.266	0.081	0.291	0.057	5.113	1.60961
18.400	60	0.266	0.081	0.265	0.047	5.625	1.84469

MacCormack method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
2.000	40	0.606	0.212	4.248	6.056	0.701	0.31522
4.000	40	0.606	0.212	2.124	1.514	1.403	0.42130
4.600	40	0.606	0.212	1.847	1.145	1.613	0.46018
5.750	40	0.606	0.212	1.478	0.733	2.017	0.54461
7.360	40	0.606	0.212	1.154	0.447	2.581	0.68592
8.000	40	0.606	0.212	1.062	0.378	2.806	0.74989
8.364	40	0.606	0.212	1.016	0.346	2.933	0.78920
9.200	40	0.606	0.212	0.923	0.286	3.227	0.88203
11.500	40	0.606	0.212	0.739	0.183	4.033	1.17948
12.266	40	0.606	0.212	0.693	0.161	4.302	1.28964
13.142	40	0.606	0.212	0.646	0.140	4.609	1.42407
14.154	40	0.606	0.212	0.600	0.121	4.964	1.59081
16.727	40	0.606	0.212	0.508	0.087	5.866	2.04791
18.400	40	0.606	0.212	0.462	0.072	6.453	2.37102

QUICKEST method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	30	0.448	0.147	0.550	0.210	2.620	0.08621
8.36	30	0.448	0.147	0.526	0.192	2.739	0.08315
9.20	30	0.448	0.147	0.478	0.159	3.013	0.07601
11.50	30	0.448	0.147	0.383	0.102	3.766	0.05969
12.27	30	0.448	0.147	0.359	0.089	4.017	0.05513
13.14	30	0.448	0.147	0.335	0.078	4.304	0.05087
14.15	30	0.448	0.147	0.311	0.067	4.635	0.04736
15.33	30	0.448	0.147	0.287	0.057	5.021	0.04518
16.73	30	0.448	0.147	0.263	0.048	5.477	0.04613
18.40	30	0.448	0.147	0.239	0.040	6.025	0.05307
20.44	30	0.448	0.147	0.215	0.032	6.695	0.07124
23.00	30	0.448	0.147	0.191	0.025	7.532	0.10994

QUICKEST method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	60	0.266	0.081	0.609	0.249	2.446	0.23454
8.36	60	0.266	0.081	0.583	0.228	2.557	0.22891
9.20	60	0.266	0.081	0.530	0.188	2.812	0.21673
11.50	60	0.266	0.081	0.424	0.121	3.515	0.19571
12.27	60	0.266	0.081	0.397	0.106	3.750	0.19279
13.14	60	0.266	0.081	0.371	0.092	4.017	0.19288
14.15	60	0.266	0.081	0.344	0.080	4.327	0.19795
15.33	60	0.266	0.081	0.318	0.068	4.687	0.21102
16.73	60	0.266	0.081	0.291	0.057	5.113	0.23784
18.40	60	0.266	0.081	0.265	0.047	5.625	0.28726
20.44	60	0.266	0.081	0.238	0.038	6.250	0.37377
23.00	60	0.266	0.081	0.212	0.030	7.031	0.52011

QUICKEST method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	RSS, ($\mu\text{g/l}$) ²
8.00	40	0.606	0.212	1.062	0.378	2.806	0.42036
8.36	40	0.606	0.212	1.016	0.346	2.933	0.39112
9.2	40	0.606	0.212	0.923	0.286	3.227	0.33698
11.50	40	0.606	0.212	0.739	0.183	4.033	0.24036
12.27	40	0.606	0.212	0.693	0.161	4.302	0.21600
13.14	40	0.606	0.212	0.646	0.140	4.609	0.19248
14.15	40	0.606	0.212	0.600	0.121	4.964	0.17066
15.33	40	0.606	0.212	0.554	0.103	5.378	0.14978
16.73	40	0.606	0.212	0.508	0.087	5.866	0.13417
18.40	40	0.606	0.212	0.462	0.072	6.453	0.12907
20.44	40	0.606	0.212	0.416	0.058	7.170	0.14493
23.00	40	0.606	0.212	0.369	0.046	8.066	0.20250

Appendix F-2: Prediction of downstream concentration profiles measured by coefficients of determination. Numerical methods were applied with predicted dispersion coefficients, given by the routing method-based model, to predict downstream concentration profiles. The tables show space steps, predicted dispersion coefficients, nondimensional parameters and coefficients of determination (R^2).

Crank-Nicolson method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	30	0.448	0.147	2.201	3.361	0.655	0.992
4.00	30	0.448	0.147	1.101	0.840	1.310	0.991
4.60	30	0.448	0.147	0.957	0.635	1.506	0.990
5.75	30	0.448	0.147	0.766	0.407	1.883	0.989
7.36	30	0.448	0.147	0.598	0.248	2.410	0.986
8.00	30	0.448	0.147	0.550	0.210	2.620	0.985
8.36	30	0.448	0.147	0.526	0.192	2.739	0.984
9.20	30	0.448	0.147	0.478	0.159	3.013	0.981
11.50	30	0.448	0.147	0.383	0.102	3.766	0.974
12.27	30	0.448	0.147	0.359	0.089	4.017	0.970
13.14	30	0.448	0.147	0.335	0.078	4.304	0.966
14.15	30	0.448	0.147	0.311	0.067	4.635	0.961
16.73	30	0.448	0.147	0.263	0.048	5.477	0.946
18.40	30	0.448	0.147	0.239	0.040	6.025	0.934

Crank-Nicolson method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	60	0.266	0.081	2.436	3.985	0.611	0.987
4.00	60	0.266	0.081	1.218	0.996	1.223	0.985
4.60	60	0.266	0.081	1.059	0.753	1.406	0.984
5.75	60	0.266	0.081	0.847	0.482	1.758	0.981
7.36	60	0.266	0.081	0.662	0.294	2.250	0.977
8.00	60	0.266	0.081	0.609	0.249	2.446	0.975
8.36	60	0.266	0.081	0.583	0.228	2.557	0.974
9.20	60	0.266	0.081	0.530	0.188	2.812	0.970
11.50	60	0.266	0.081	0.424	0.121	3.515	0.959
12.27	60	0.266	0.081	0.397	0.106	3.750	0.955
13.14	60	0.266	0.081	0.371	0.092	4.017	0.950
14.15	60	0.266	0.081	0.344	0.080	4.327	0.943
16.73	60	0.266	0.081	0.291	0.057	5.113	0.923
18.40	60	0.266	0.081	0.265	0.047	5.625	0.909

Crank-Nicolson method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	40	0.606	0.212	4.248	6.056	0.701	0.989
4.00	40	0.606	0.212	2.124	1.514	1.403	0.987
4.60	40	0.606	0.212	1.847	1.145	1.613	0.987
5.75	40	0.606	0.212	1.478	0.733	2.017	0.985
7.36	40	0.606	0.212	1.154	0.447	2.581	0.983
8.00	40	0.606	0.212	1.062	0.378	2.806	0.982
8.36	40	0.606	0.212	1.016	0.346	2.933	0.981
9.20	40	0.606	0.212	0.923	0.286	3.227	0.979
11.50	40	0.606	0.212	0.739	0.183	4.033	0.971
12.27	40	0.606	0.212	0.693	0.161	4.302	0.969
13.14	40	0.606	0.212	0.646	0.140	4.609	0.965
14.15	40	0.606	0.212	0.600	0.121	4.964	0.961
16.73	40	0.606	0.212	0.508	0.087	5.866	0.947
18.40	40	0.606	0.212	0.462	0.072	6.453	0.936

MacCormack method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	30	0.448	0.147	2.201	3.361	0.655	0.991
4.00	30	0.448	0.147	1.101	0.840	1.310	0.988
4.60	30	0.448	0.147	0.957	0.635	1.506	0.987
5.75	30	0.448	0.147	0.766	0.407	1.883	0.984
7.36	30	0.448	0.147	0.598	0.248	2.410	0.979
8.00	30	0.448	0.147	0.550	0.210	2.620	0.976
8.36	30	0.448	0.147	0.526	0.192	2.739	0.975
9.20	30	0.448	0.147	0.478	0.159	3.013	0.971
11.50	30	0.448	0.147	0.383	0.102	3.766	0.959
12.27	30	0.448	0.147	0.359	0.089	4.017	0.956
13.14	30	0.448	0.147	0.335	0.078	4.304	0.951
14.15	30	0.448	0.147	0.311	0.067	4.635	0.944
16.73	30	0.448	0.147	0.263	0.048	5.477	0.926
18.40	30	0.448	0.147	0.239	0.040	6.025	0.913

MacCormack method (experiment 9)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	60	0.266	0.081	2.436	3.985	0.611	0.984
4.00	60	0.266	0.081	1.218	0.996	1.223	0.979
4.60	60	0.266	0.081	1.059	0.753	1.406	0.977
5.75	60	0.266	0.081	0.847	0.482	1.758	0.972
7.36	60	0.266	0.081	0.662	0.294	2.250	0.965
8.00	60	0.266	0.081	0.609	0.249	2.446	0.961
8.36	60	0.266	0.081	0.583	0.228	2.557	0.959
9.20	60	0.266	0.081	0.530	0.188	2.812	0.954
11.50	60	0.266	0.081	0.424	0.121	3.515	0.938
12.27	60	0.266	0.081	0.397	0.106	3.750	0.932
13.14	60	0.266	0.081	0.371	0.092	4.017	0.925
14.15	60	0.266	0.081	0.344	0.080	4.327	0.916
16.73	60	0.266	0.081	0.291	0.057	5.113	0.892
18.40	60	0.266	0.081	0.265	0.047	5.625	0.876

MacCormack method (experiment 13)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
2.00	40	0.606	0.212	4.248	6.056	0.701	0.986
4.00	40	0.606	0.212	2.124	1.514	1.403	0.981
4.60	40	0.606	0.212	1.847	1.145	1.613	0.979
5.75	40	0.606	0.212	1.478	0.733	2.017	0.975
7.36	40	0.606	0.212	1.154	0.447	2.581	0.969
8.00	40	0.606	0.212	1.062	0.378	2.806	0.966
8.36	40	0.606	0.212	1.016	0.346	2.933	0.964
9.20	40	0.606	0.212	0.923	0.286	3.227	0.960
11.50	40	0.606	0.212	0.739	0.183	4.033	0.947
12.27	40	0.606	0.212	0.693	0.161	4.302	0.942
13.14	40	0.606	0.212	0.646	0.140	4.609	0.936
14.15	40	0.606	0.212	0.600	0.121	4.964	0.928
16.73	40	0.606	0.212	0.508	0.087	5.866	0.908
18.40	40	0.606	0.212	0.462	0.072	6.453	0.893

QUICKEST method (experiment 4)

Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
8.00	30	0.448	0.147	0.550	0.210	2.620	0.992
8.36	30	0.448	0.147	0.526	0.192	2.739	0.992
9.20	30	0.448	0.147	0.478	0.159	3.013	0.993
11.50	30	0.448	0.147	0.383	0.102	3.766	0.995
12.27	30	0.448	0.147	0.359	0.089	4.017	0.995
13.14	30	0.448	0.147	0.335	0.078	4.304	0.996
14.15	30	0.448	0.147	0.311	0.067	4.635	0.996
15.33	30	0.448	0.147	0.287	0.057	5.021	0.996
16.73	30	0.448	0.147	0.263	0.048	5.477	0.996
18.40	30	0.448	0.147	0.239	0.040	6.025	0.994
20.44	30	0.448	0.147	0.215	0.032	6.695	0.992
23.00	30	0.448	0.147	0.191	0.025	7.532	0.987

QUICKEST method (experiment 9)

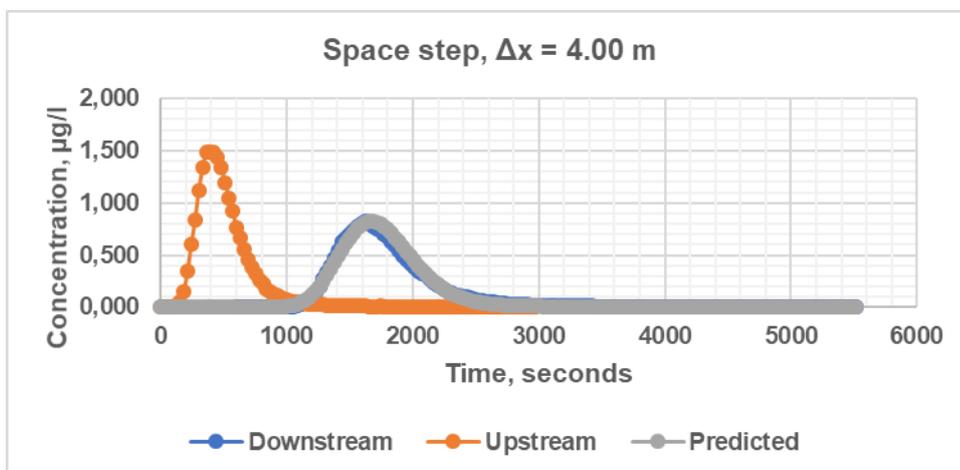
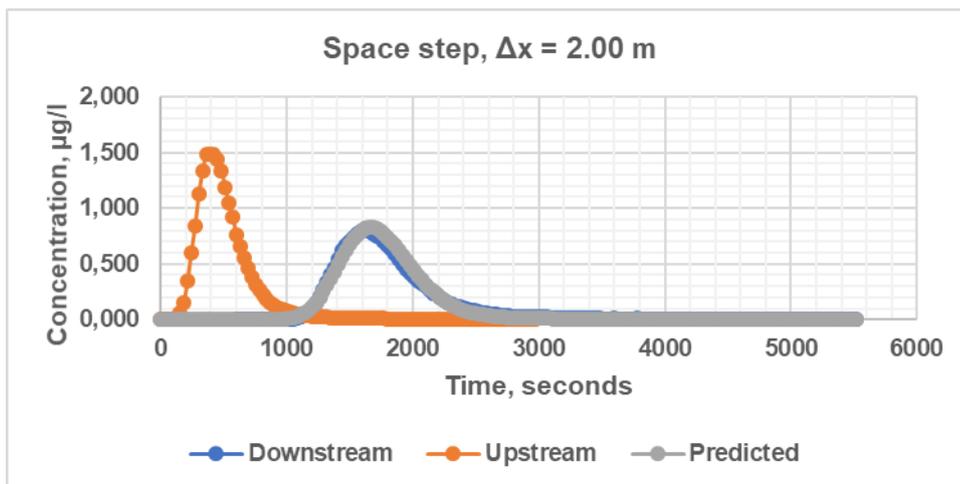
Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
8.00	60	0.266	0.081	0.609	0.249	2.446	0.986
8.36	60	0.266	0.081	0.583	0.228	2.557	0.987
9.20	60	0.266	0.081	0.530	0.188	2.812	0.988
11.50	60	0.266	0.081	0.424	0.121	3.515	0.990
12.27	60	0.266	0.081	0.397	0.106	3.750	0.990
13.14	60	0.266	0.081	0.371	0.092	4.017	0.990
14.15	60	0.266	0.081	0.344	0.080	4.327	0.990
15.33	60	0.266	0.081	0.318	0.068	4.687	0.990
16.73	60	0.266	0.081	0.291	0.057	5.113	0.989
18.40	60	0.266	0.081	0.265	0.047	5.625	0.987
20.44	60	0.266	0.081	0.238	0.038	6.250	0.983
23.00	60	0.266	0.081	0.212	0.030	7.031	0.964

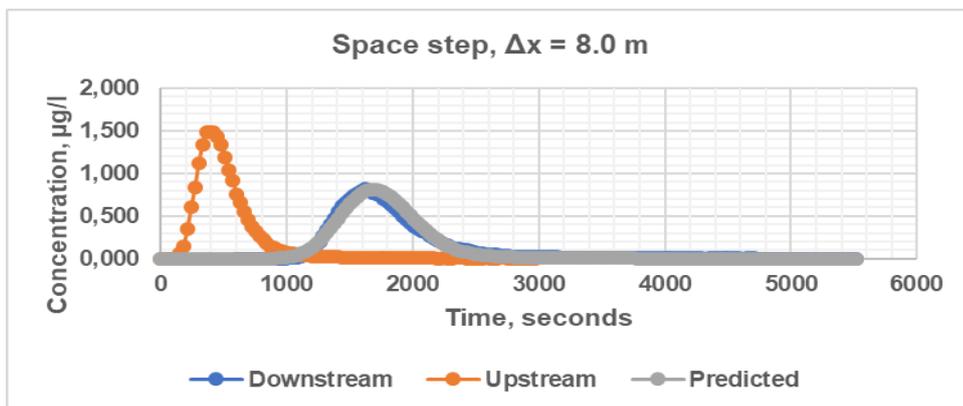
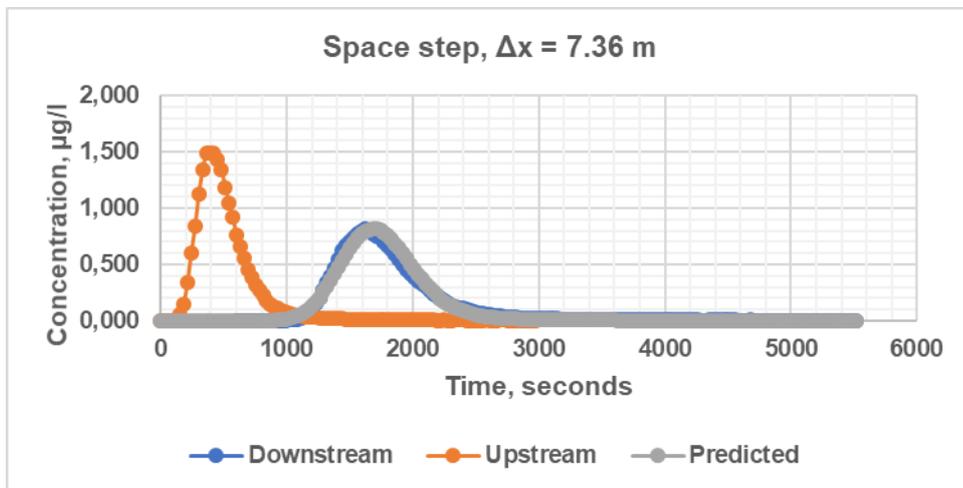
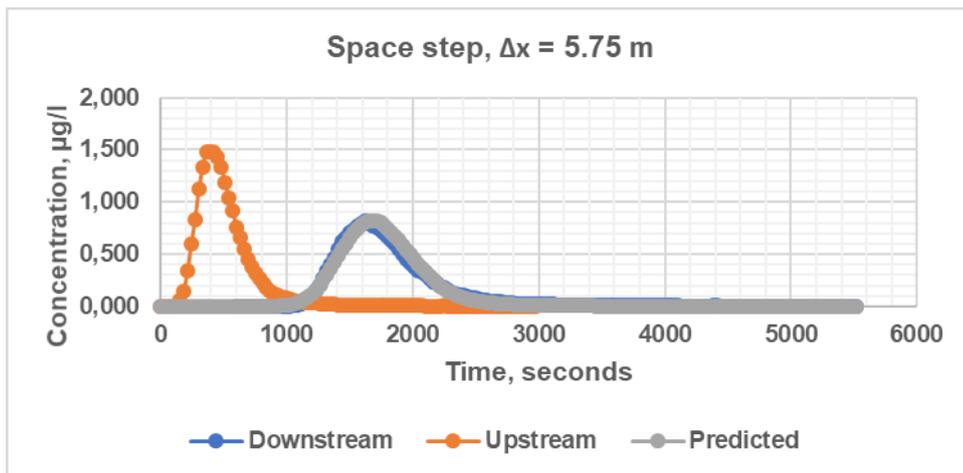
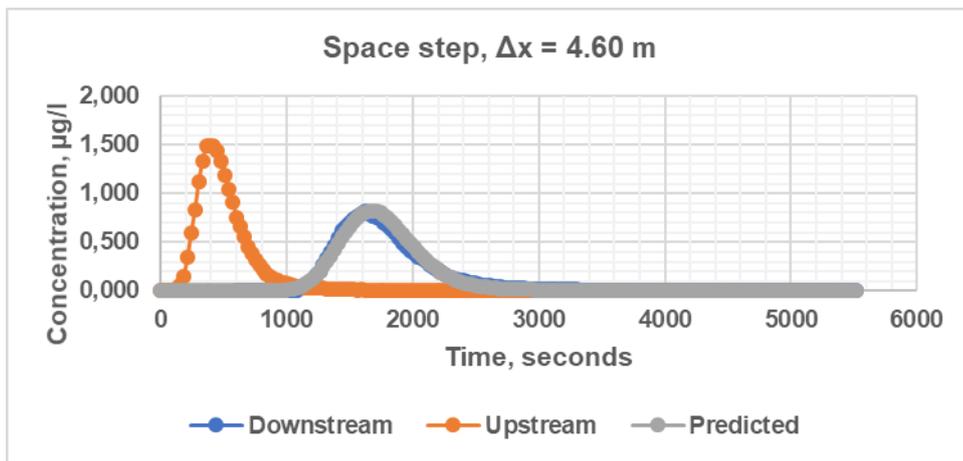
QUICKEST method (experiment 13)

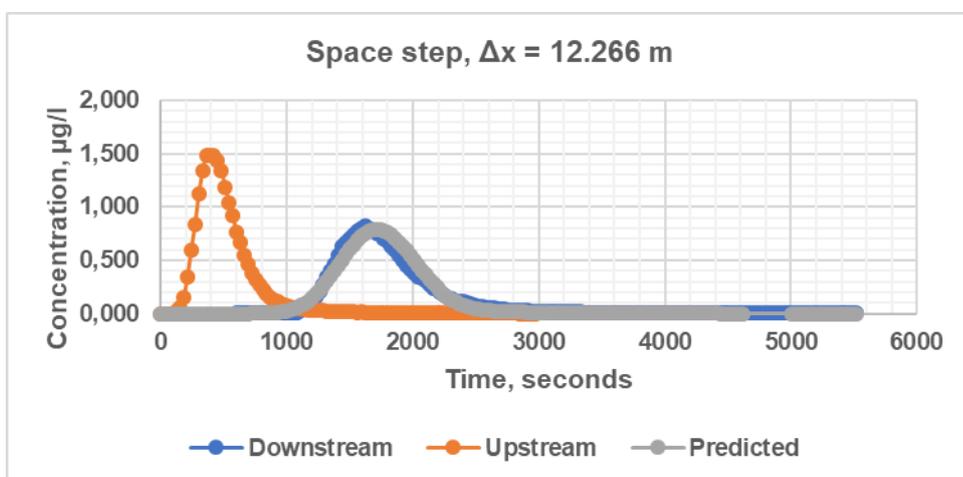
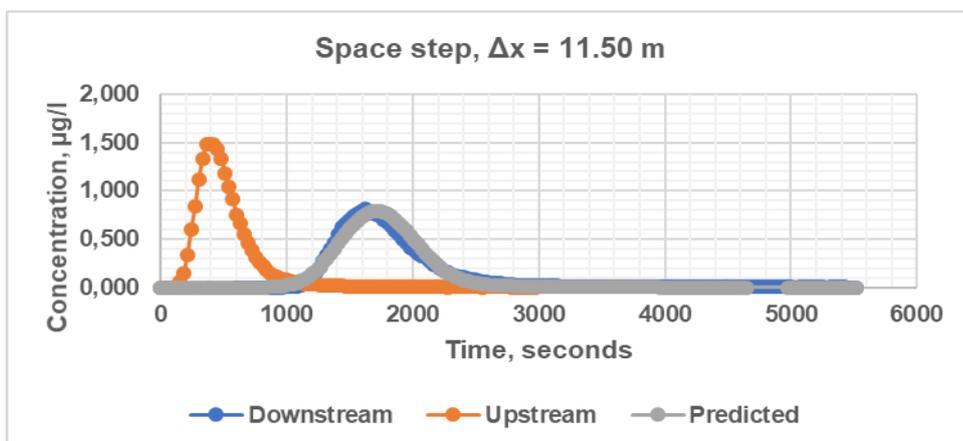
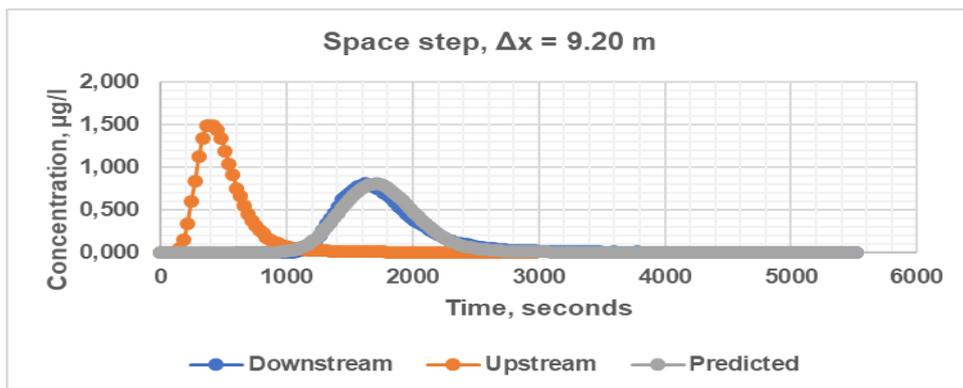
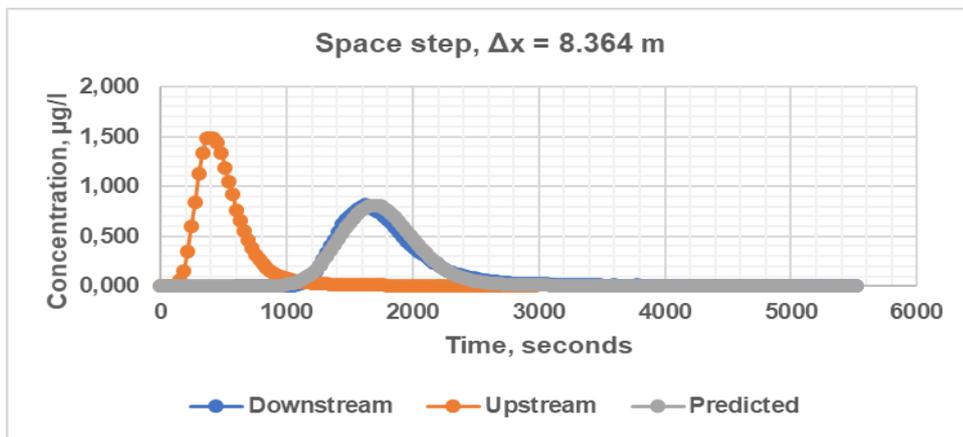
Δx , meters	Δt , seconds	D, m ² /s	V, m/s	c	d	Pe	R ²
8.00	40	0.606	0.212	1.062	0.378	2.806	0.981
8.36	40	0.606	0.212	1.016	0.346	2.933	0.982
9.20	40	0.606	0.212	0.923	0.286	3.227	0.985
11.50	40	0.606	0.212	0.739	0.183	4.033	0.989
12.27	40	0.606	0.212	0.693	0.161	4.302	0.990
13.14	40	0.606	0.212	0.646	0.140	4.609	0.991
14.15	40	0.606	0.212	0.600	0.121	4.964	0.992
15.33	40	0.606	0.212	0.554	0.103	5.378	0.993
16.73	40	0.606	0.212	0.508	0.087	5.866	0.994
18.40	40	0.606	0.212	0.462	0.072	6.453	0.994
20.44	40	0.606	0.212	0.416	0.058	7.170	0.994
23.00	40	0.606	0.212	0.369	0.046	8.066	0.993

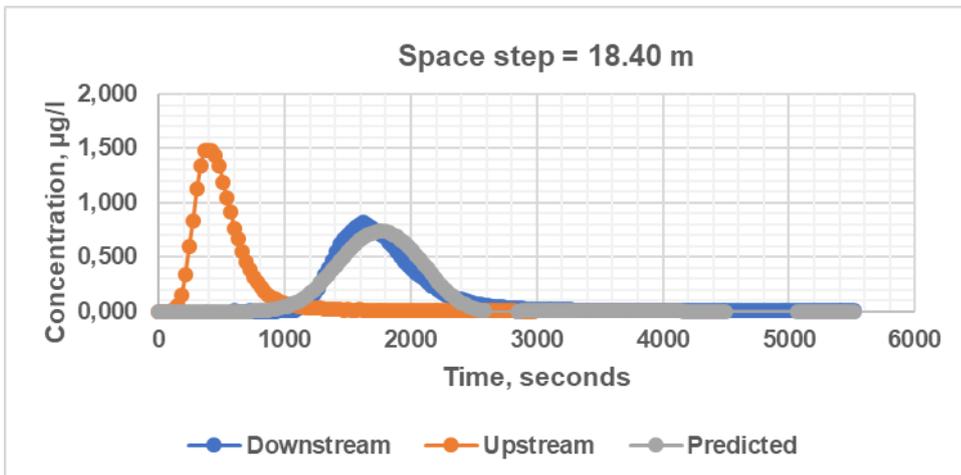
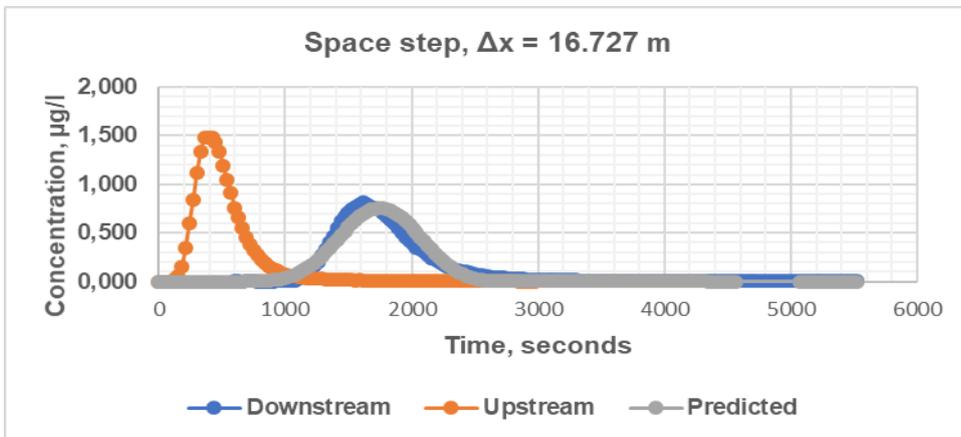
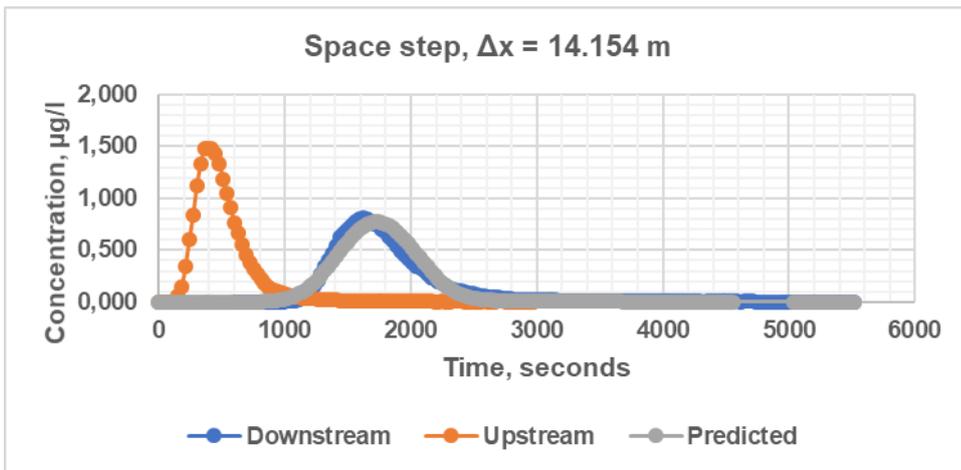
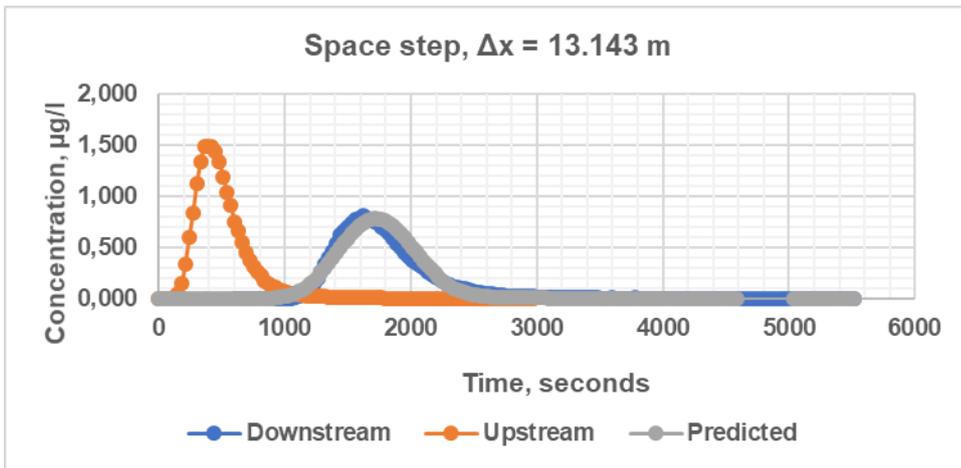
Appendix F-3: Charts showing simulations of downstream concentration-time profiles using dispersion coefficients predicted by routing method-based model.

Crank-Nicolson method (experiment 4)

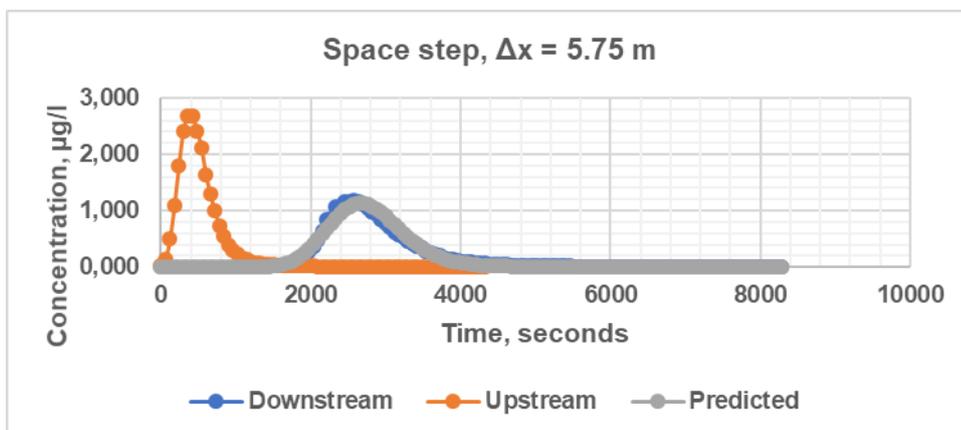
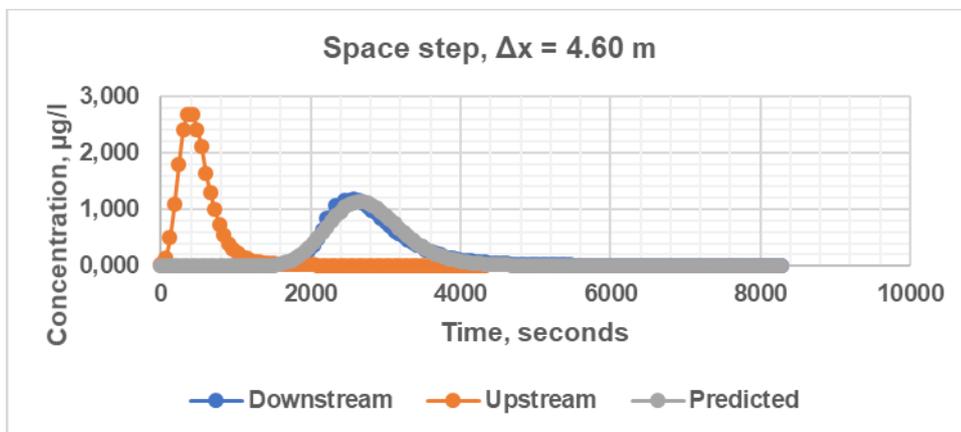
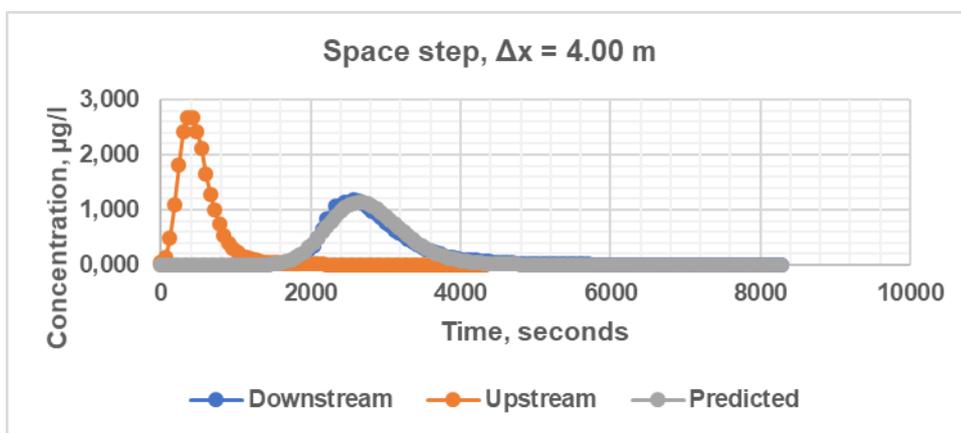
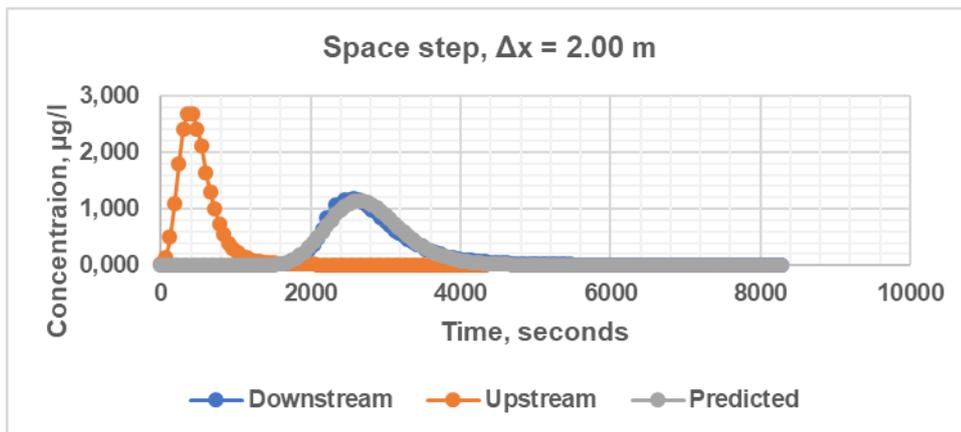


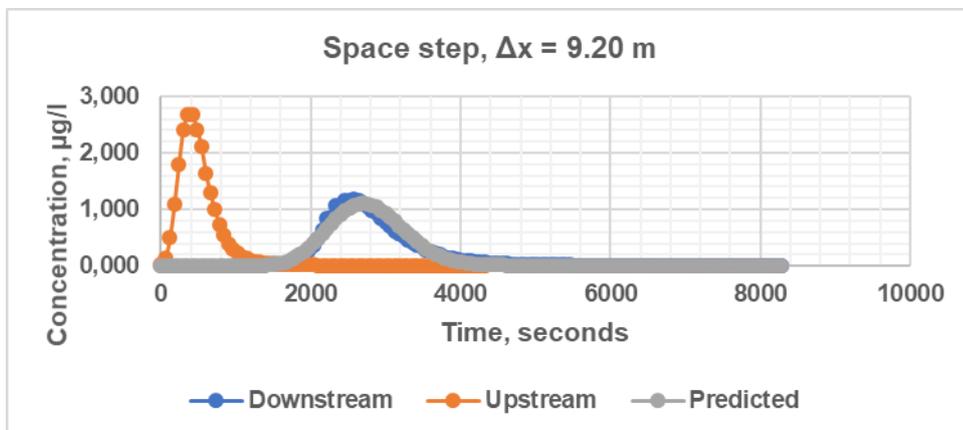
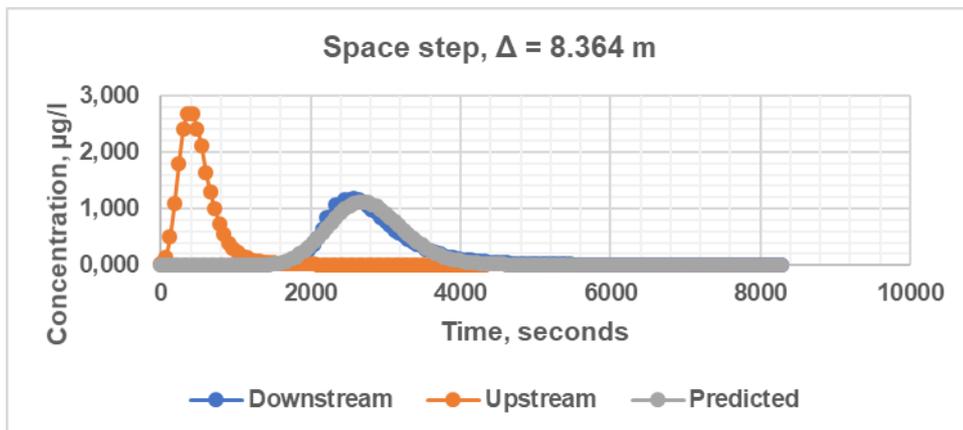
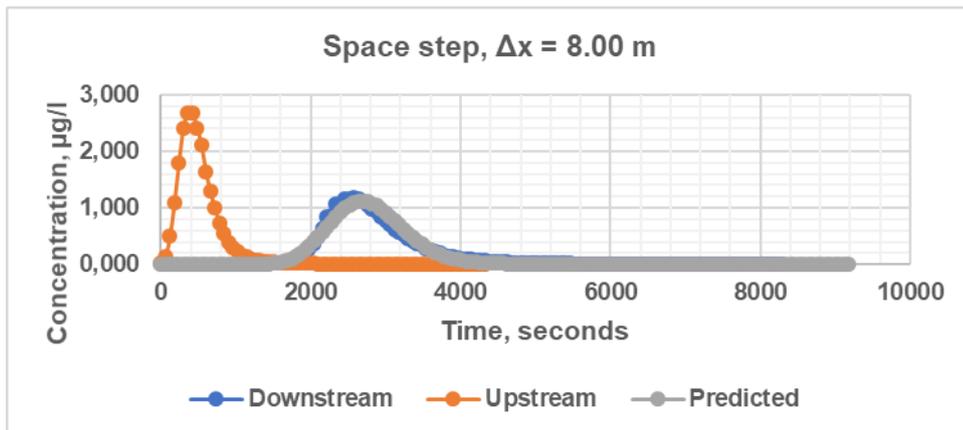
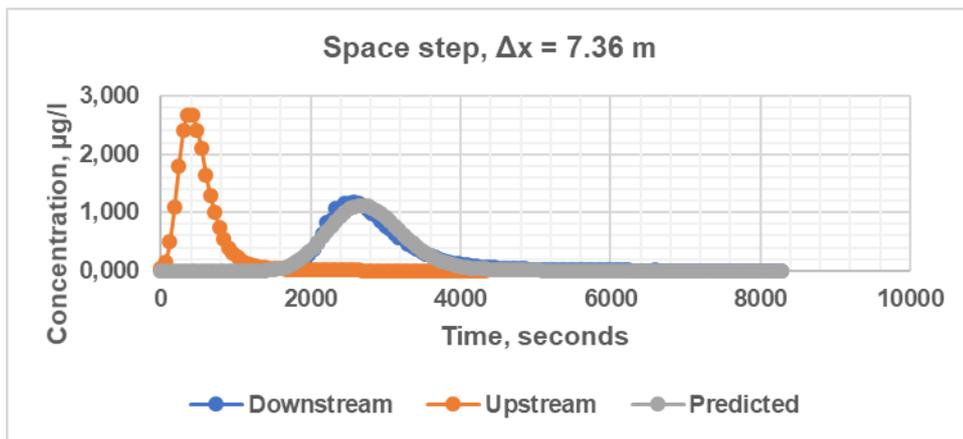


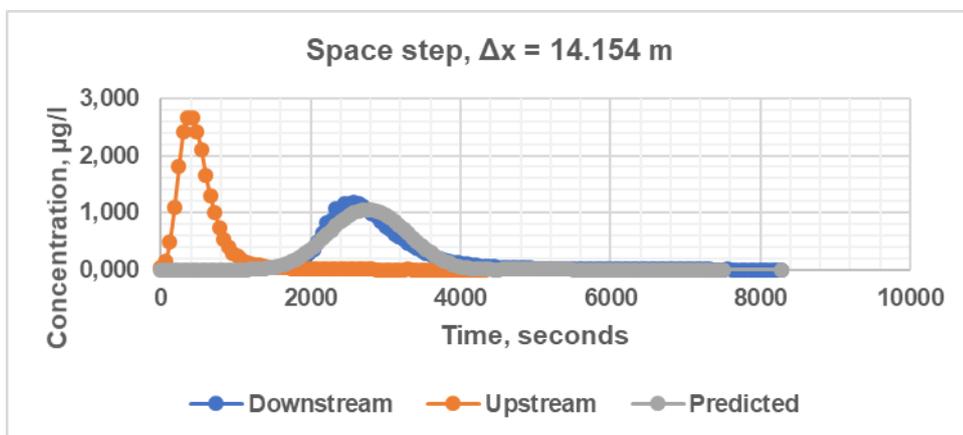
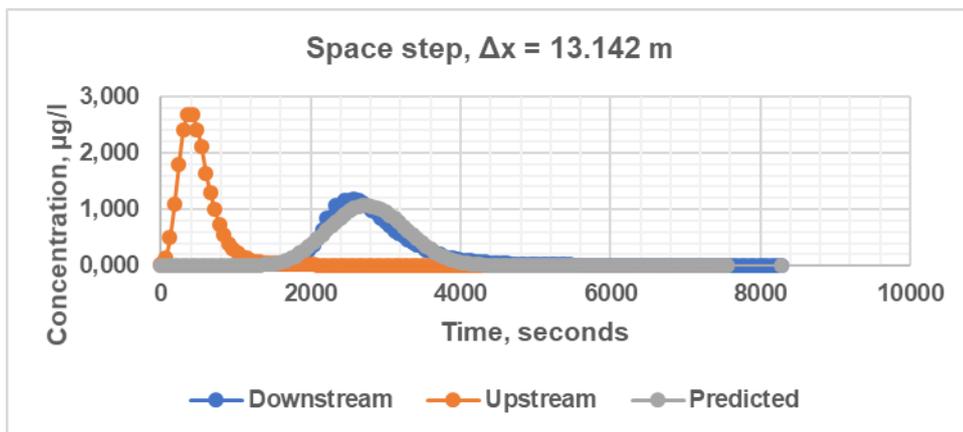
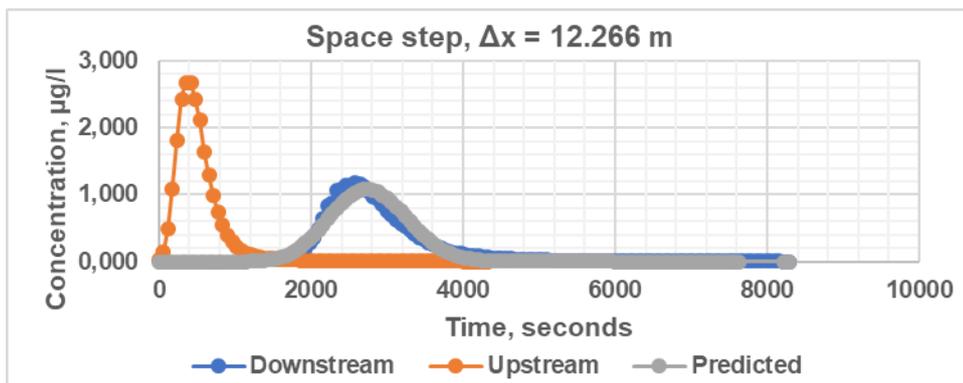
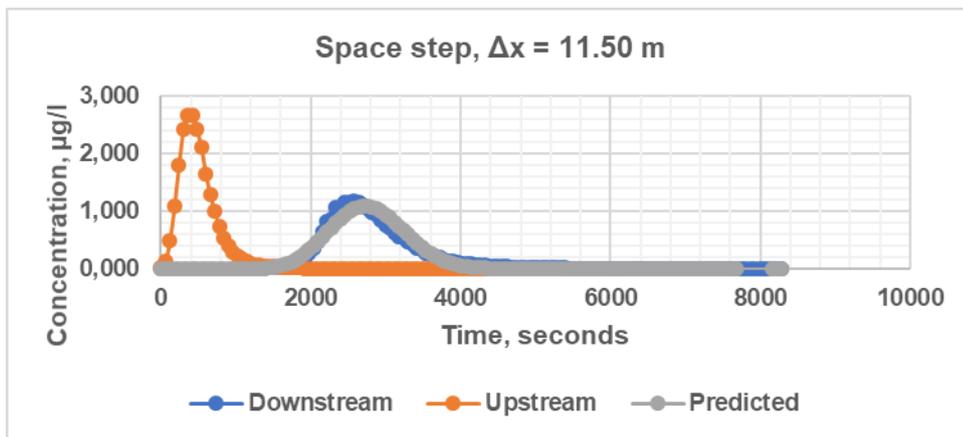


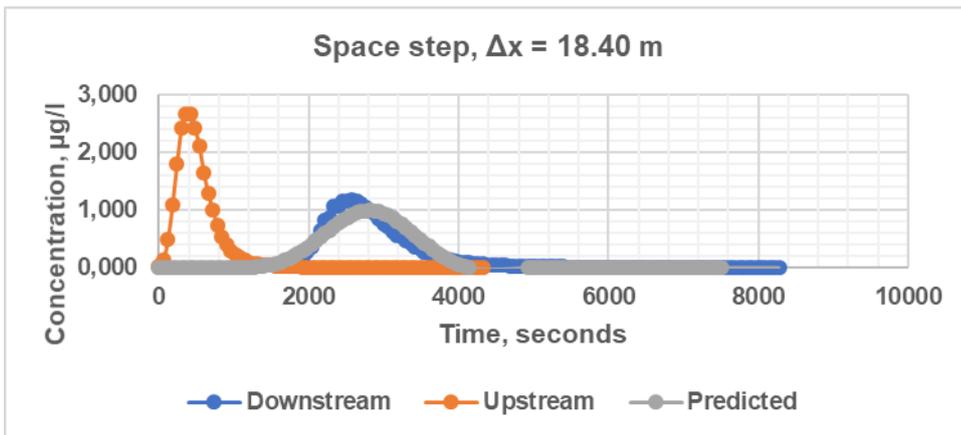
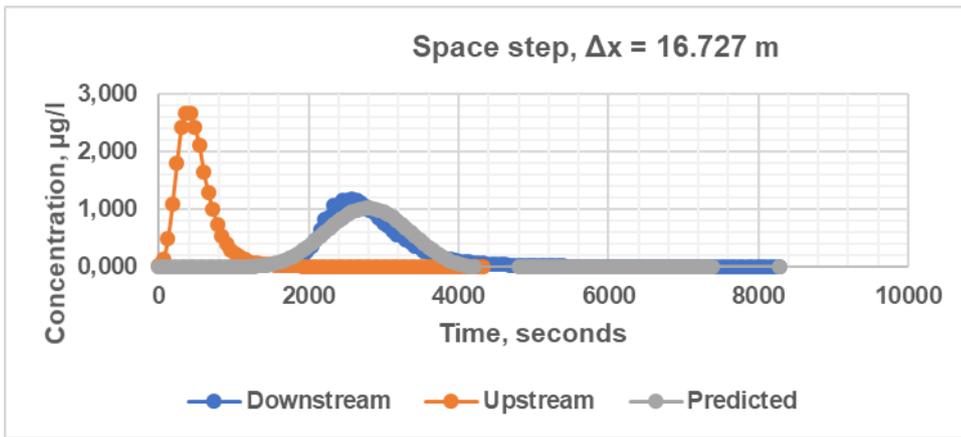


Crank-Nicolson method (experiment 9)

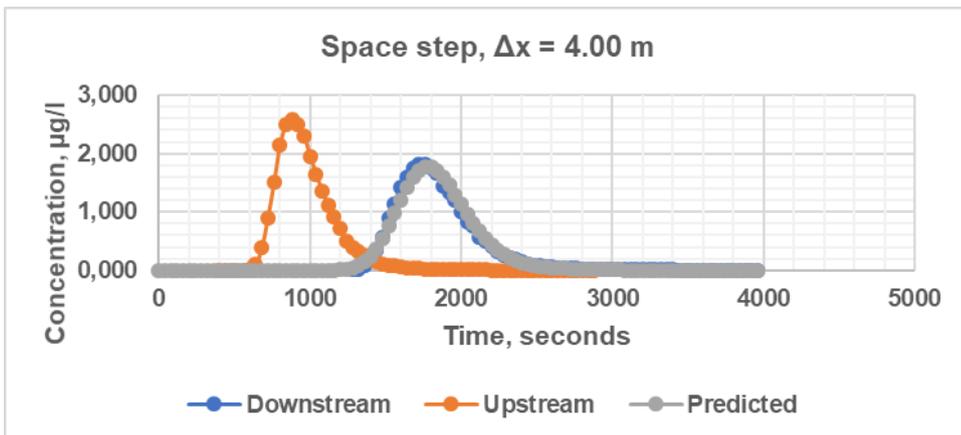
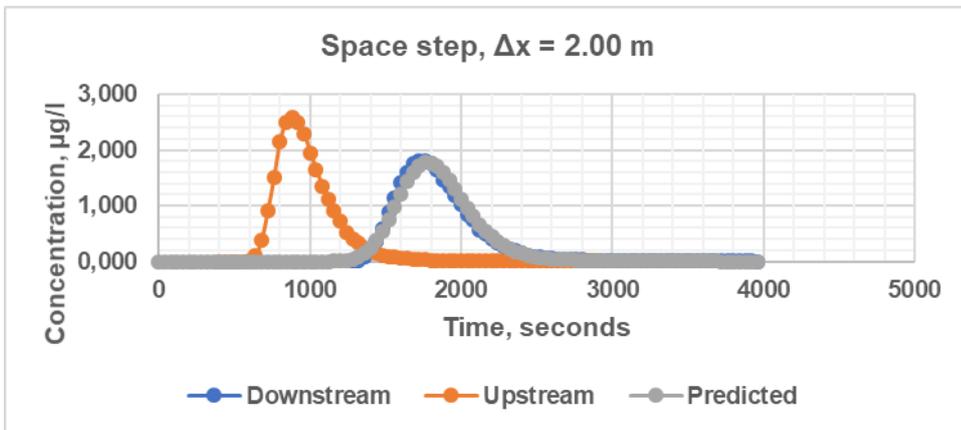


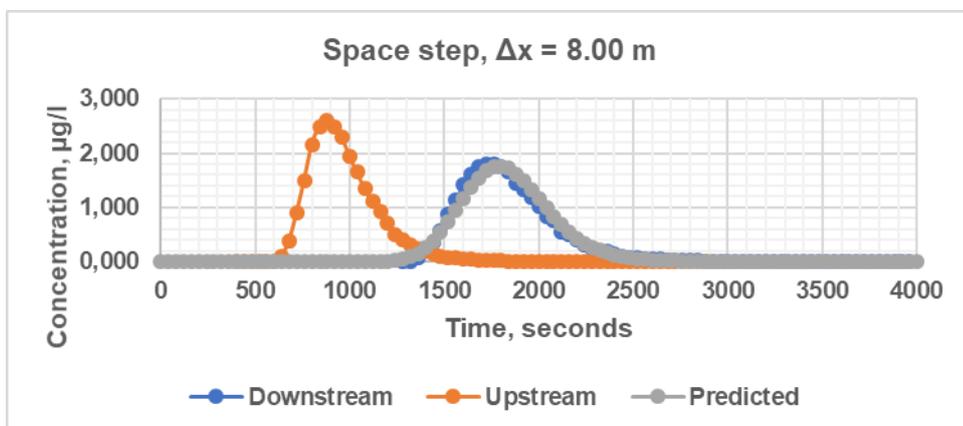
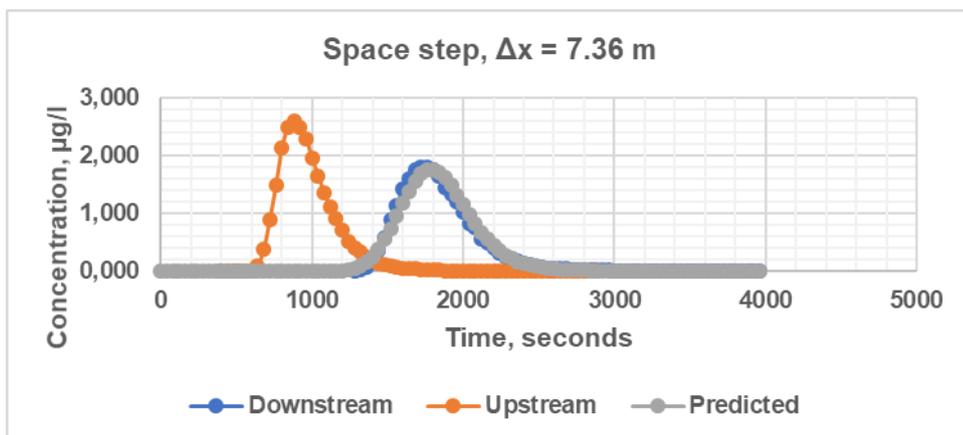
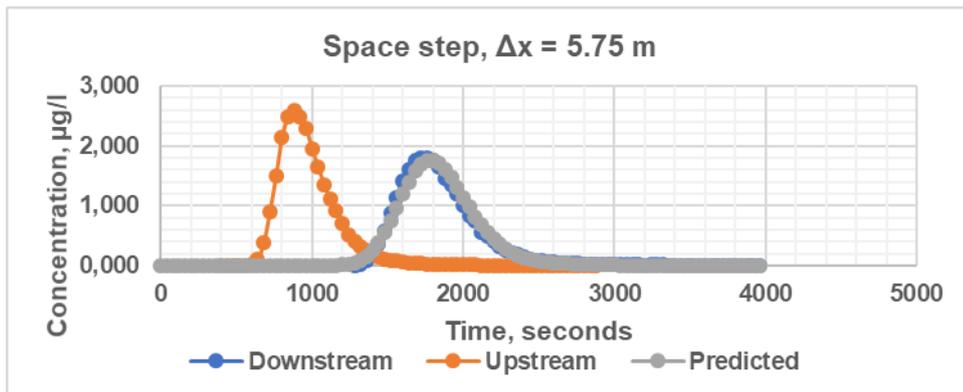
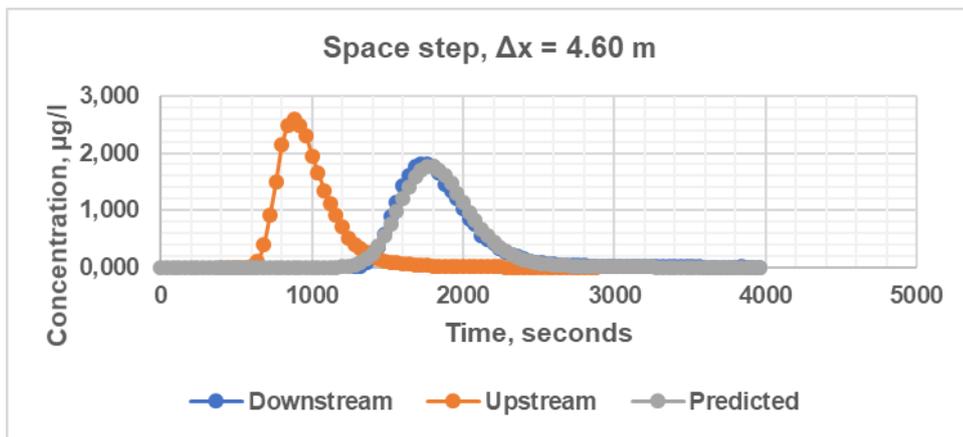


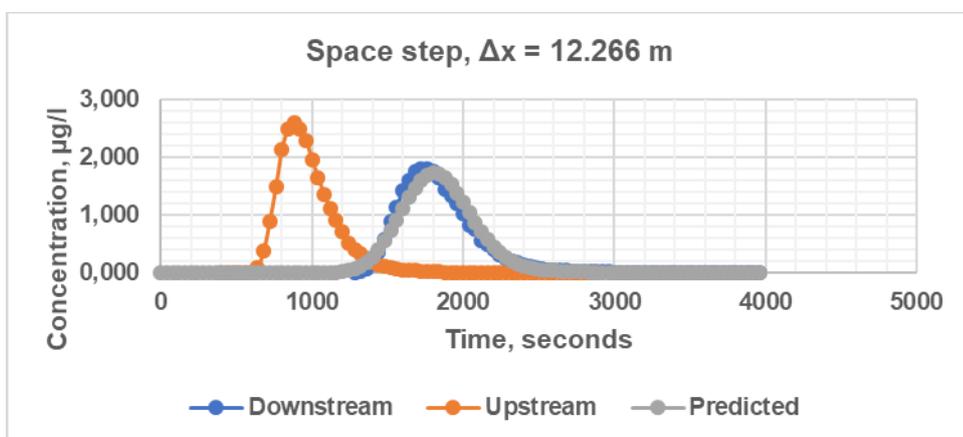
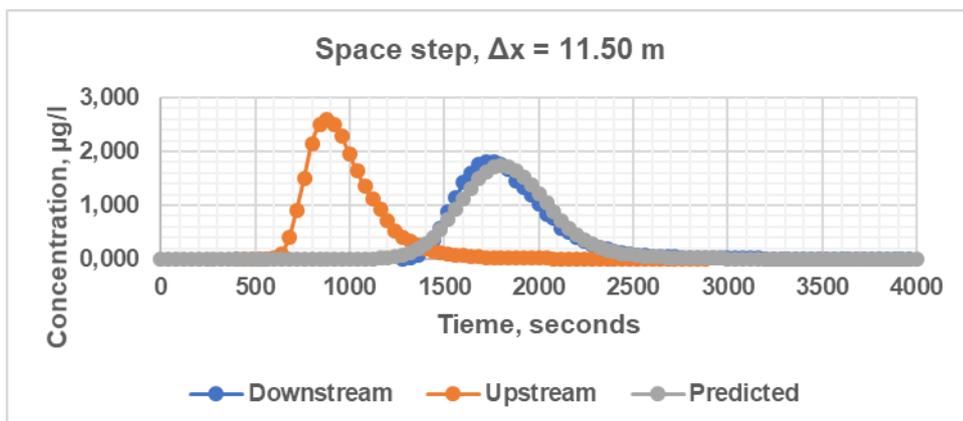
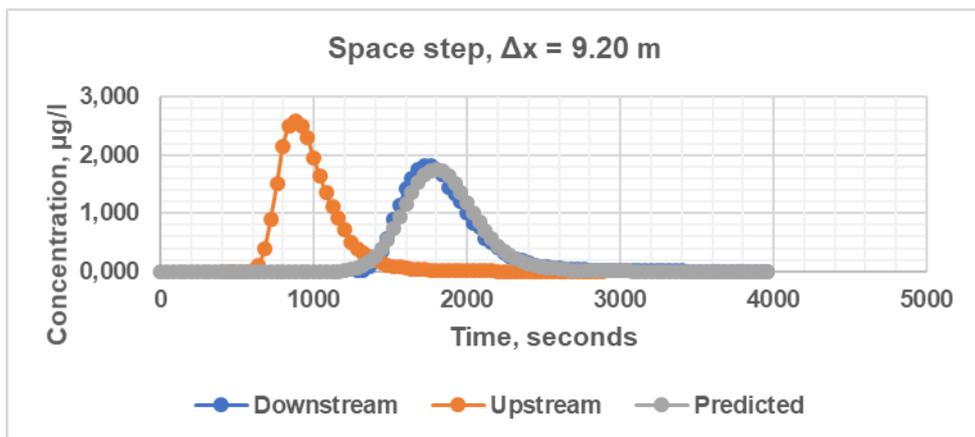
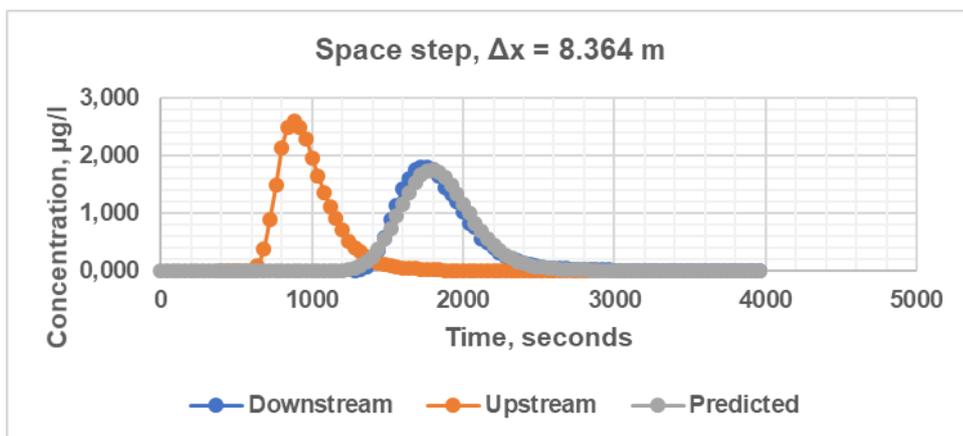


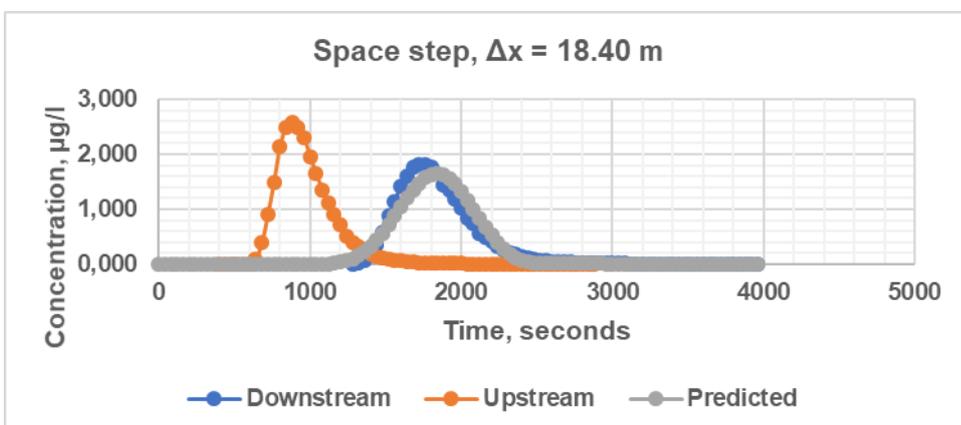
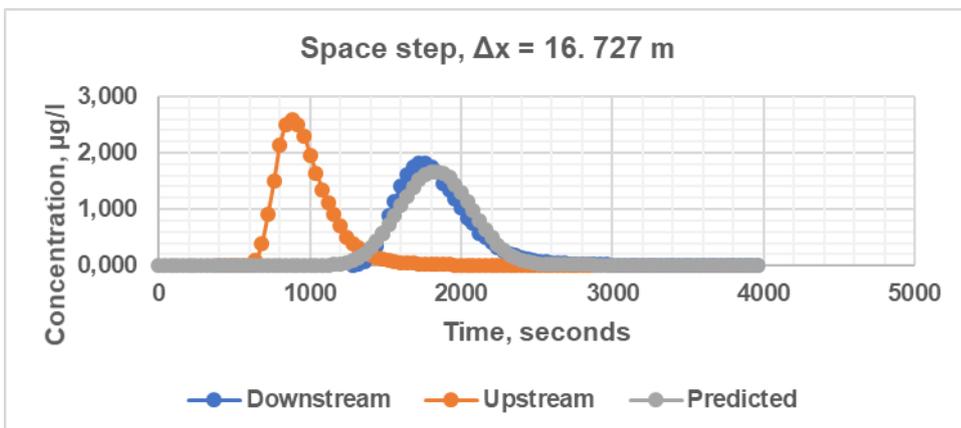
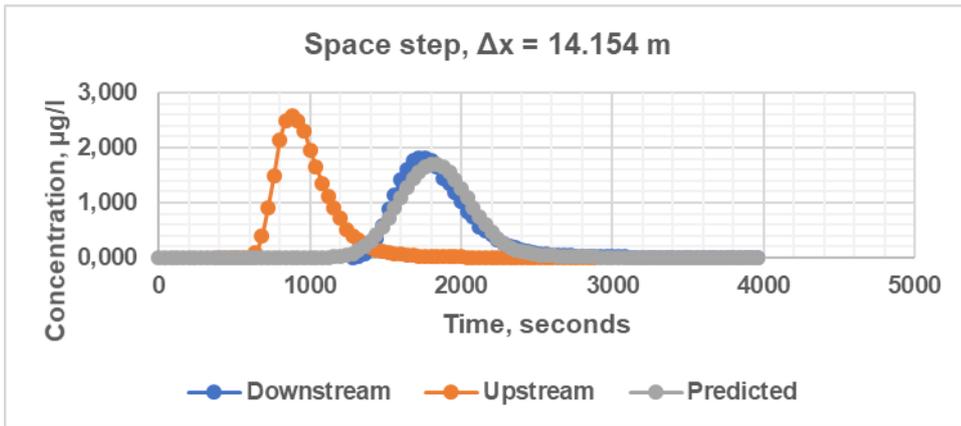
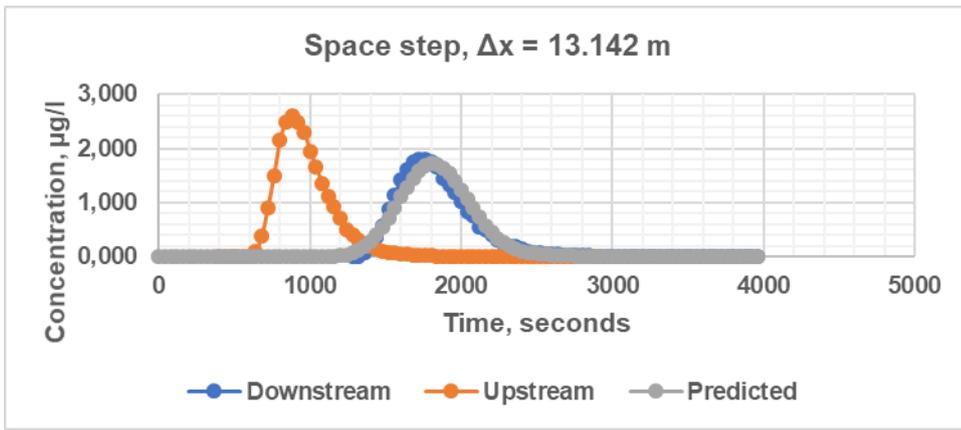


Crank-Nicolson method (experiment 13)

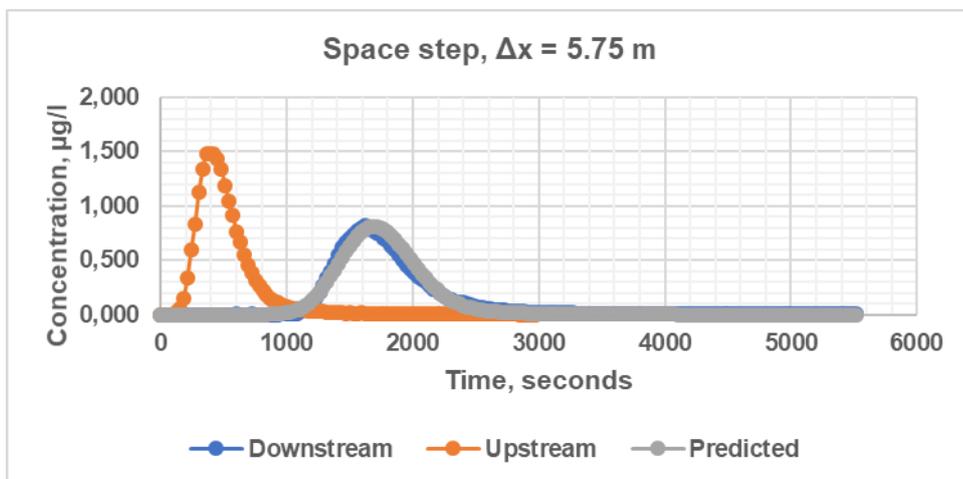
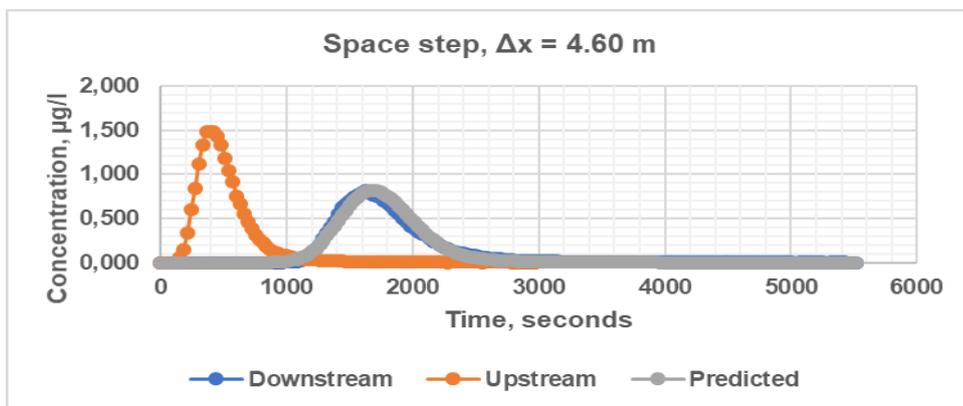
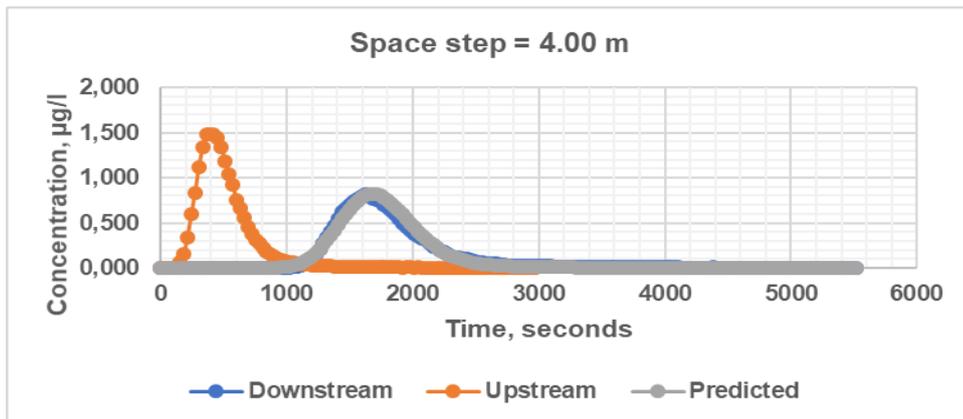
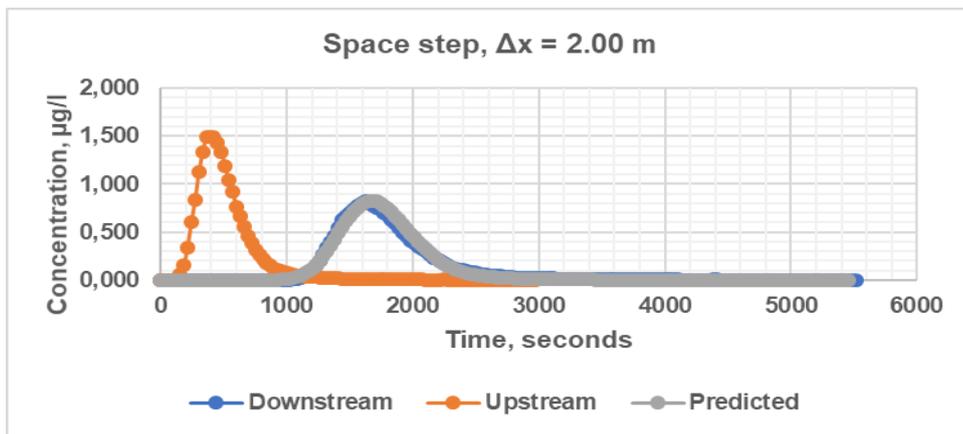


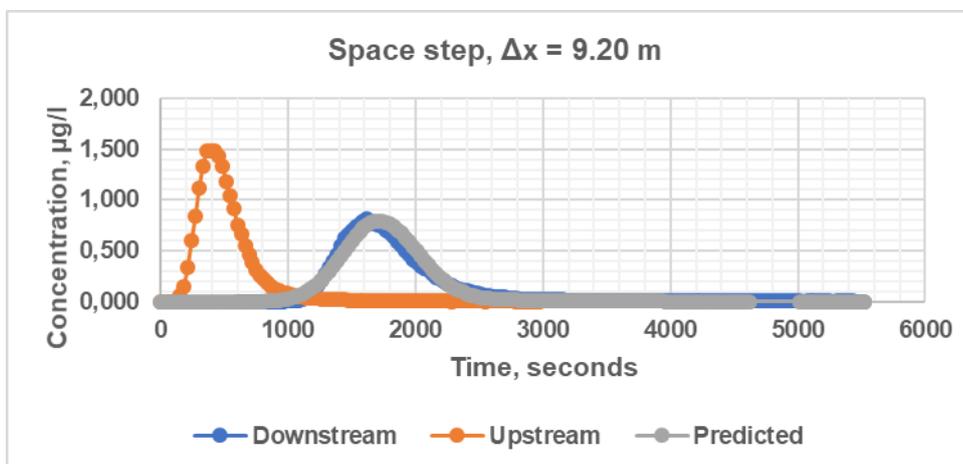
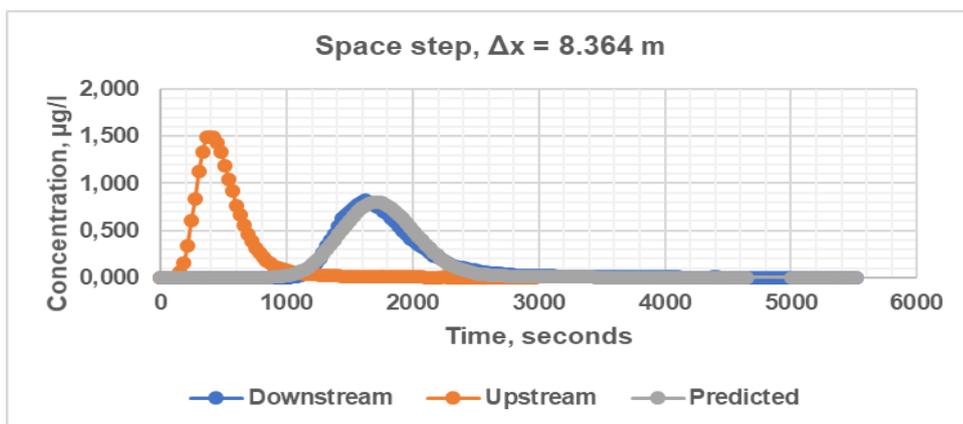
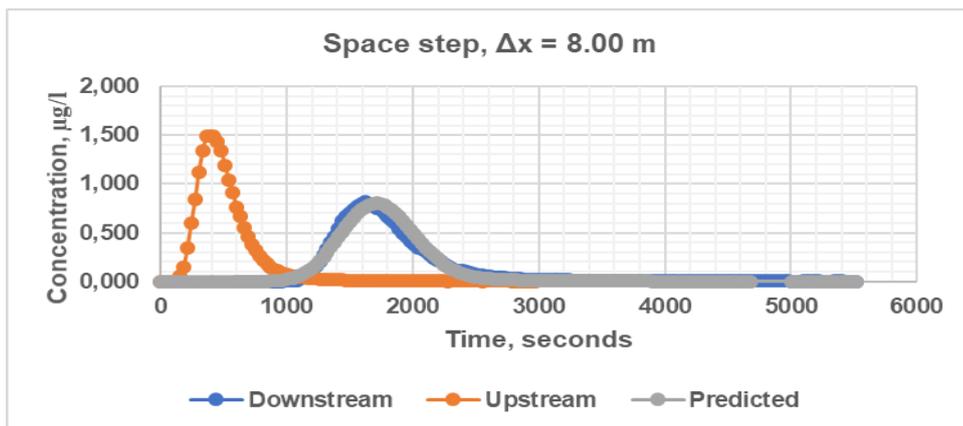
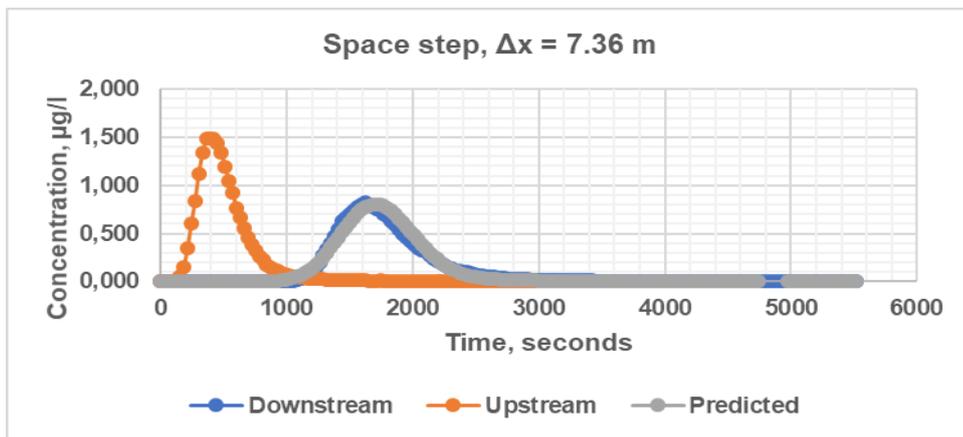


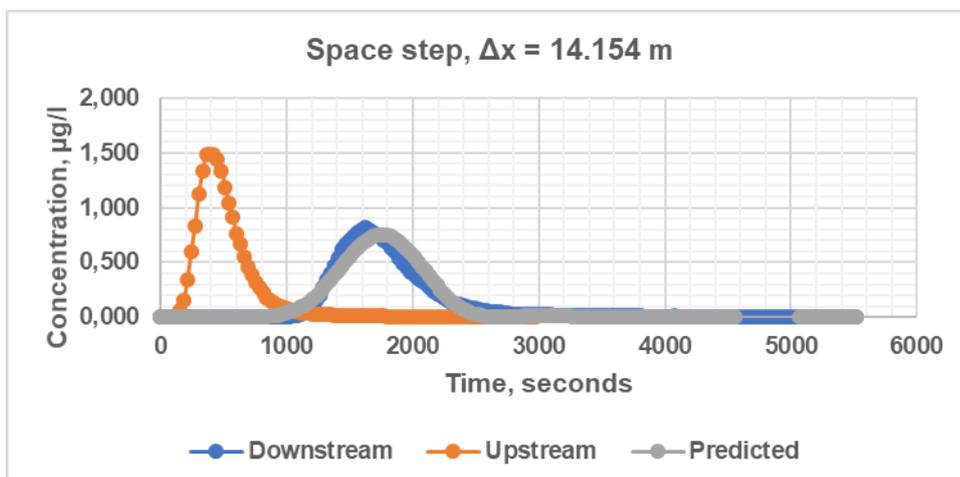
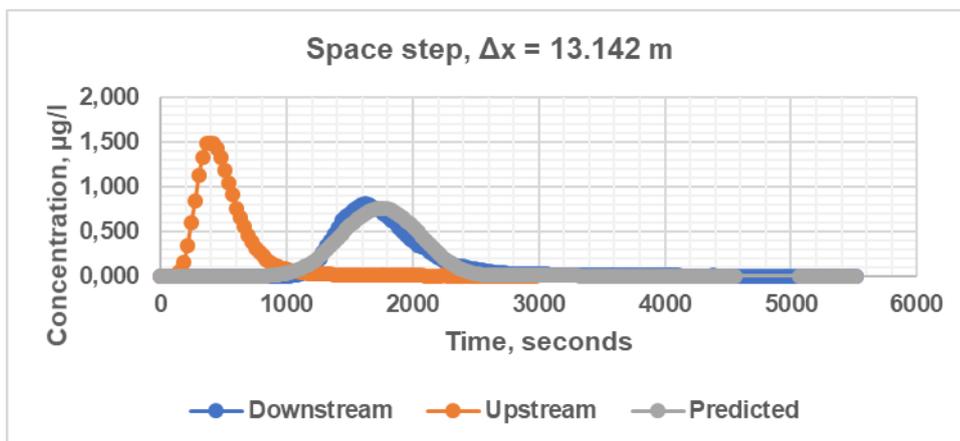
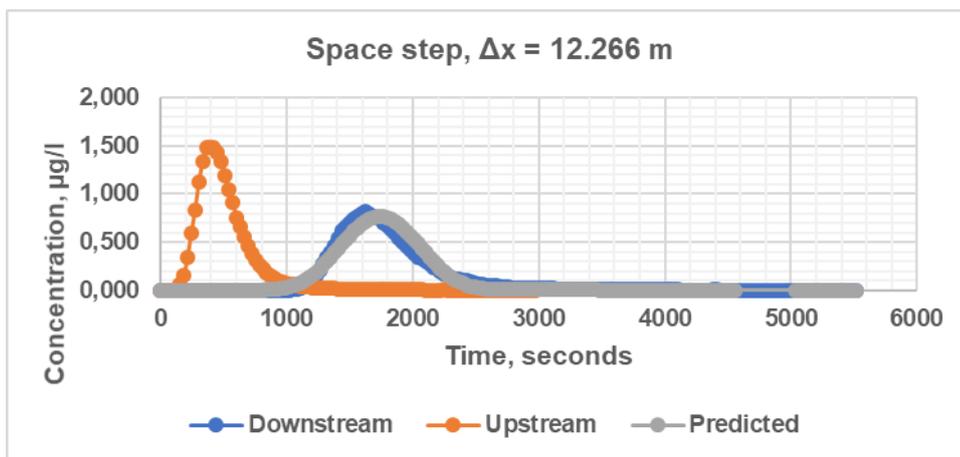
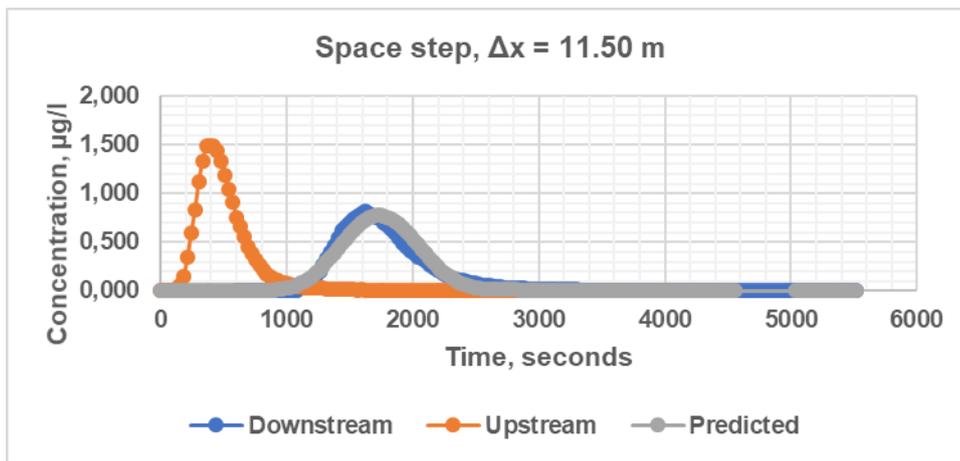


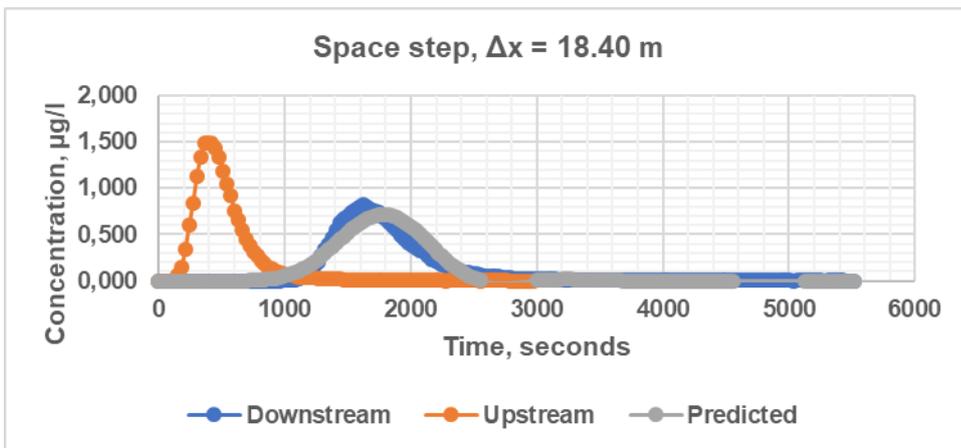
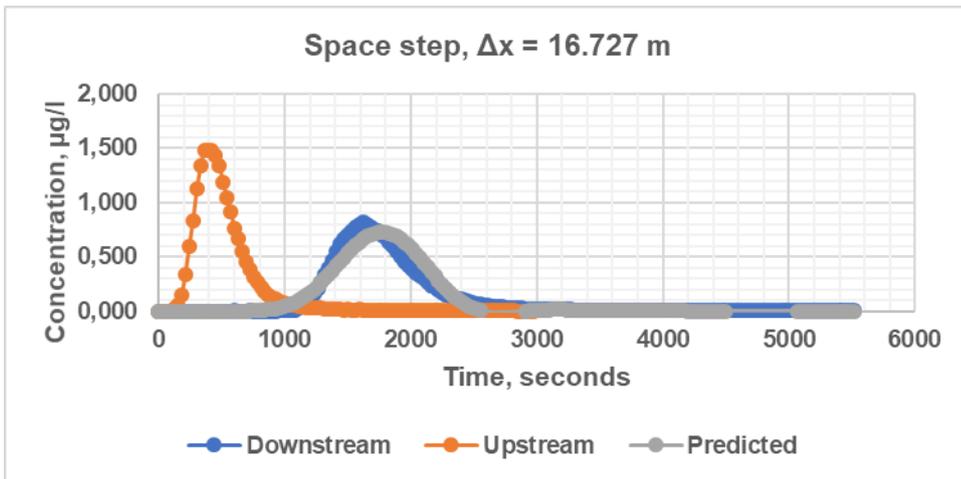


MacCormack method (experiment 4)

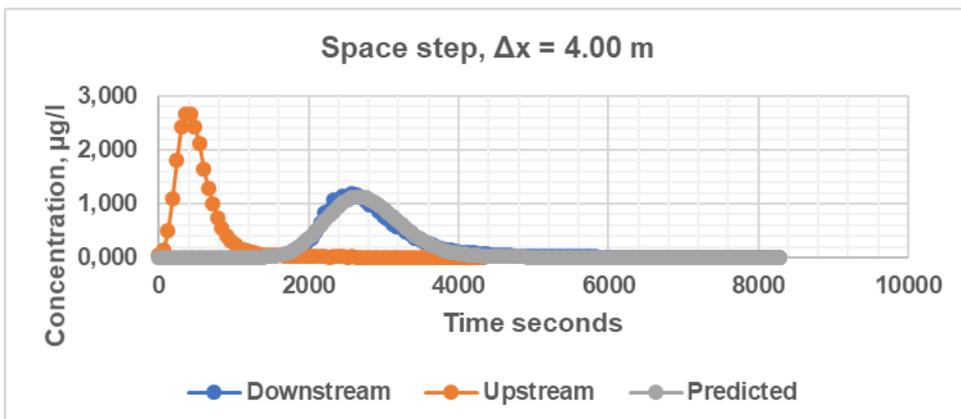
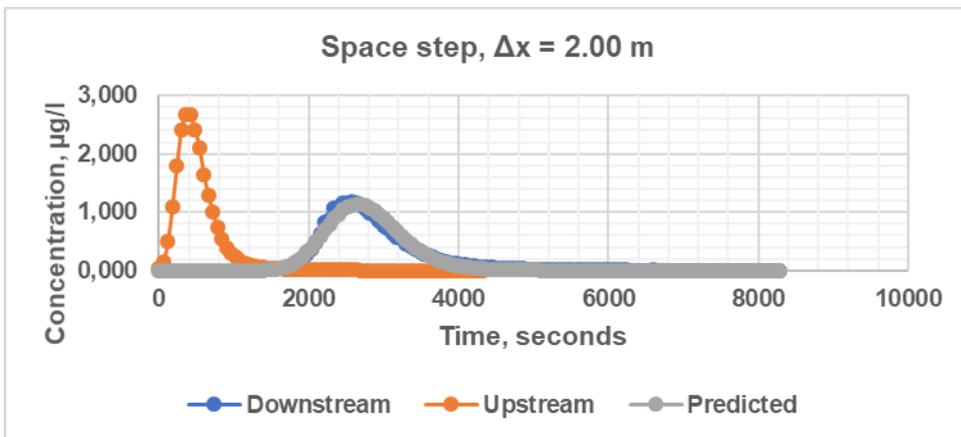


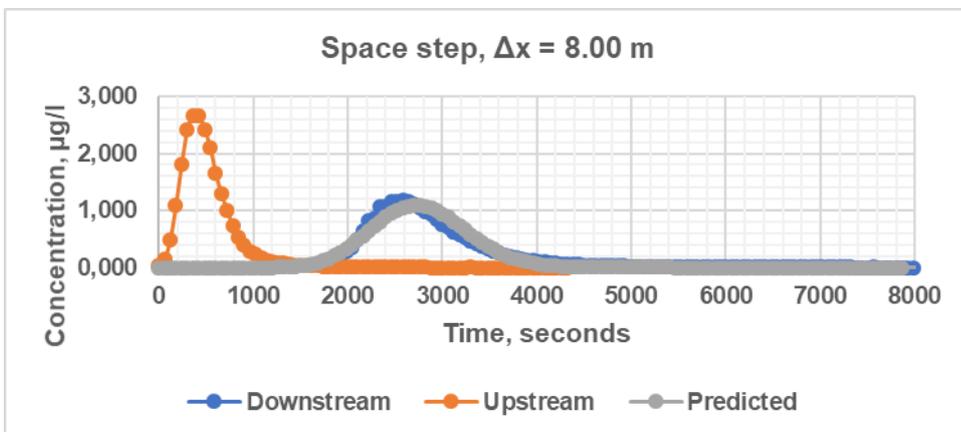
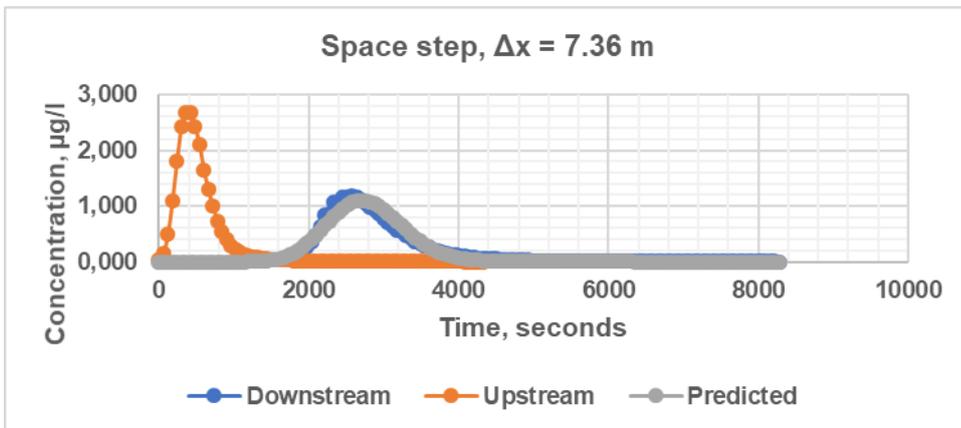
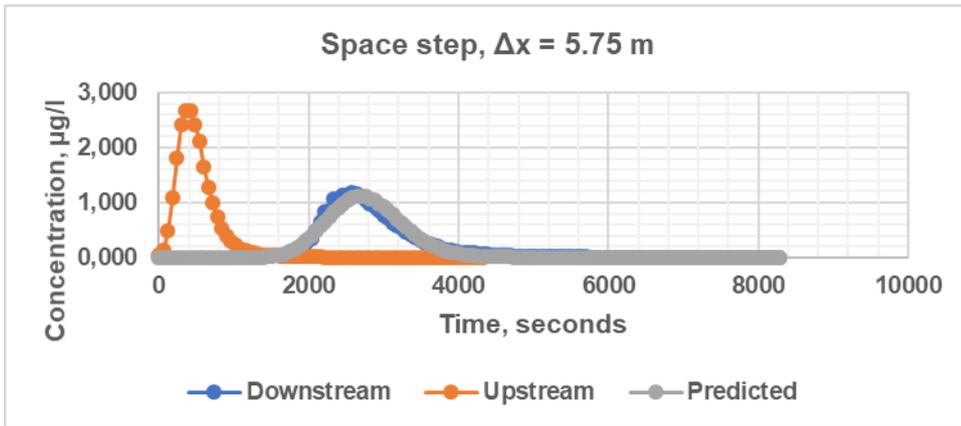
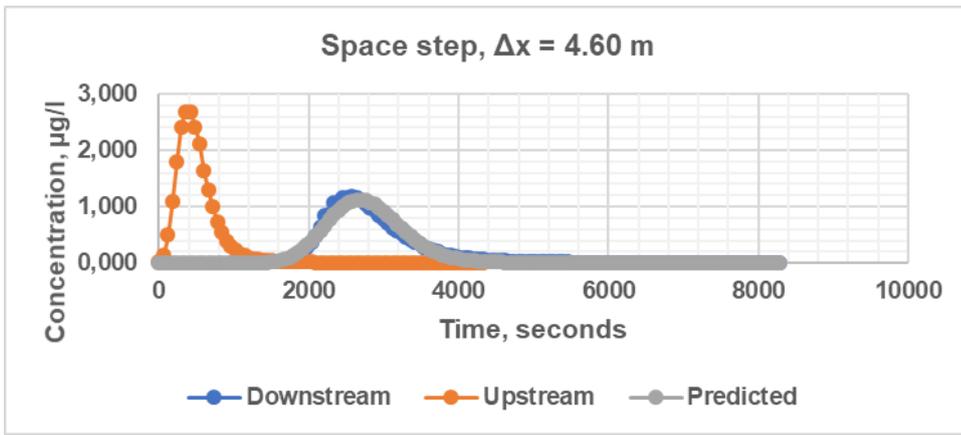


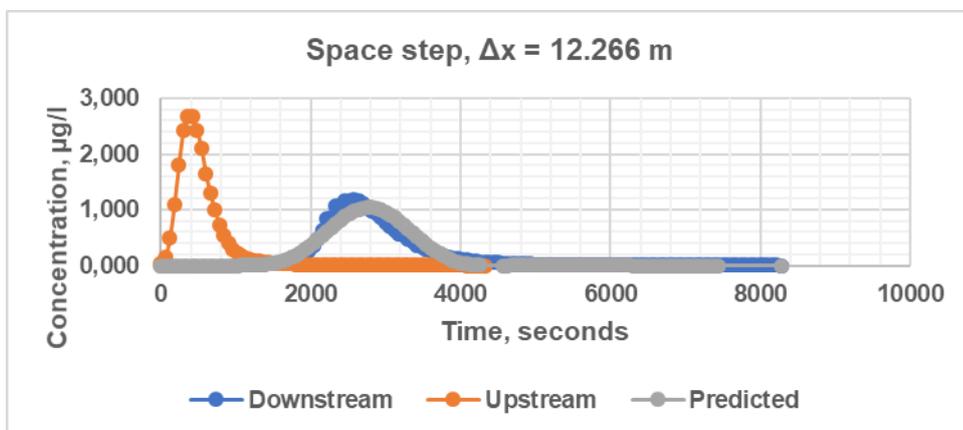
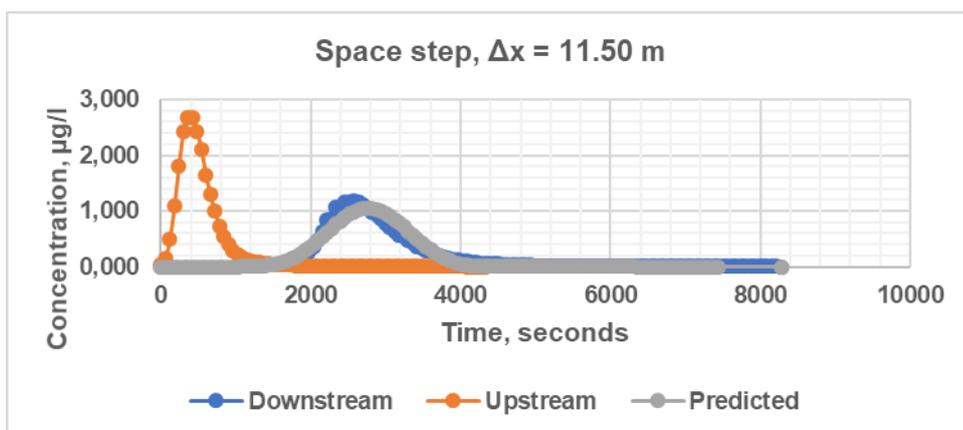
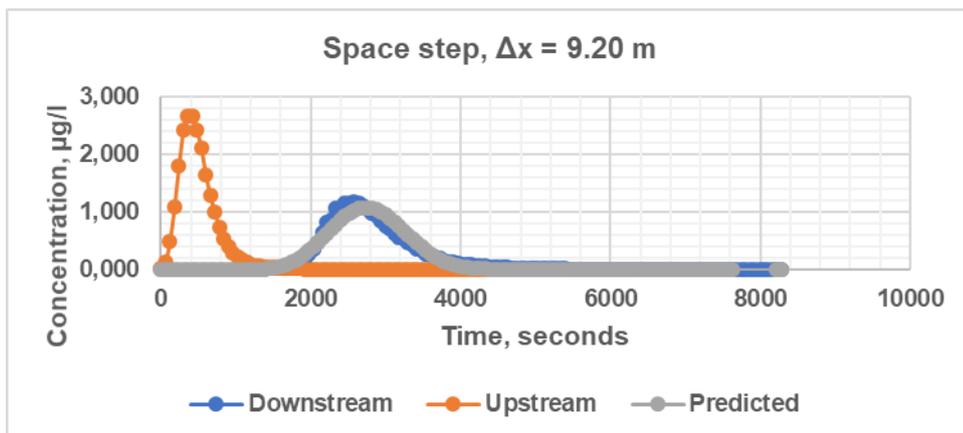
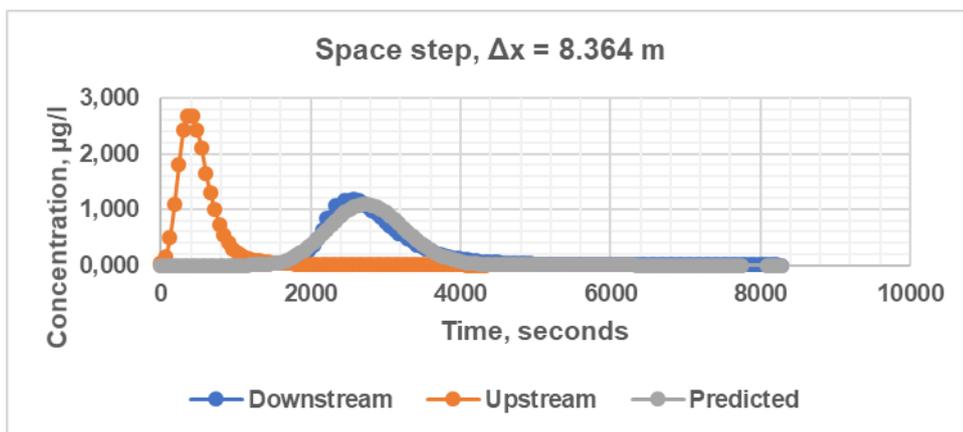


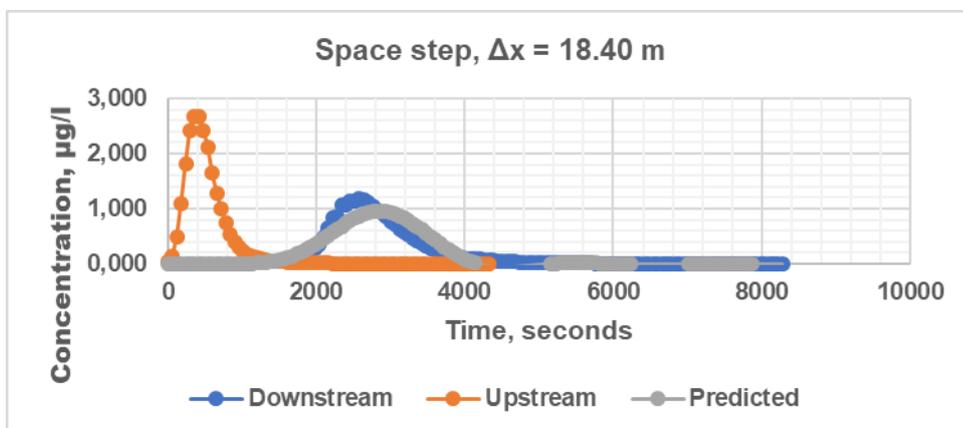
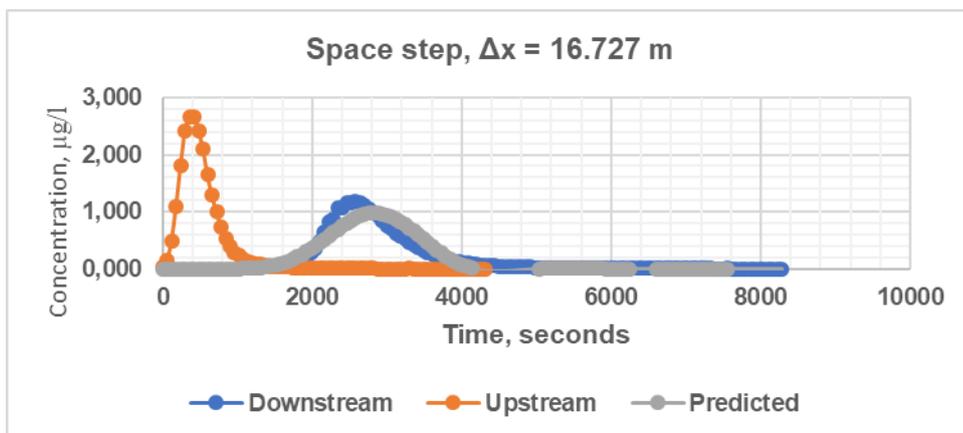
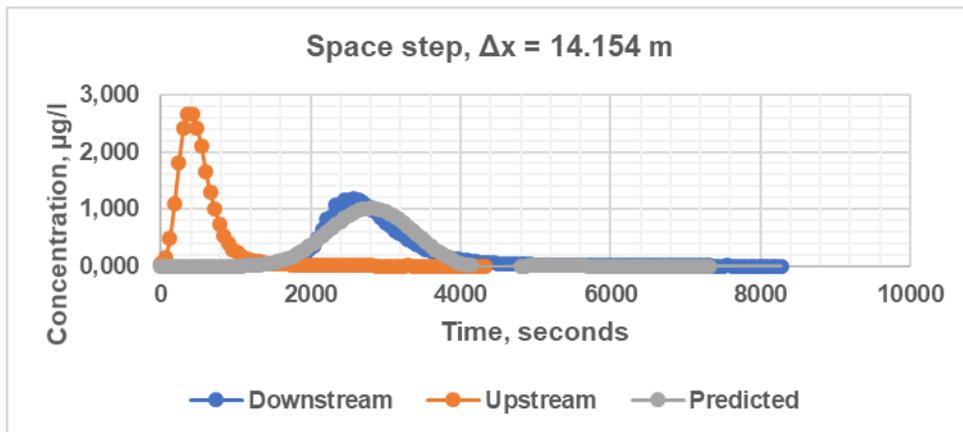
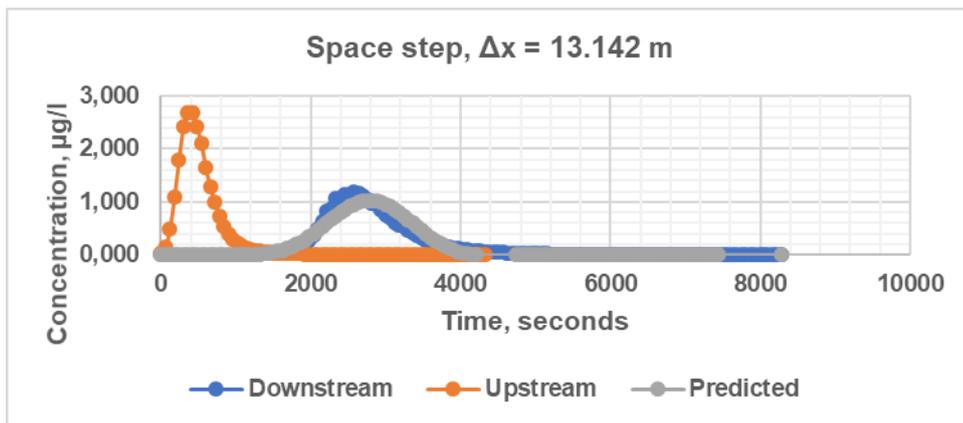


MacCormack method (experiment 9)

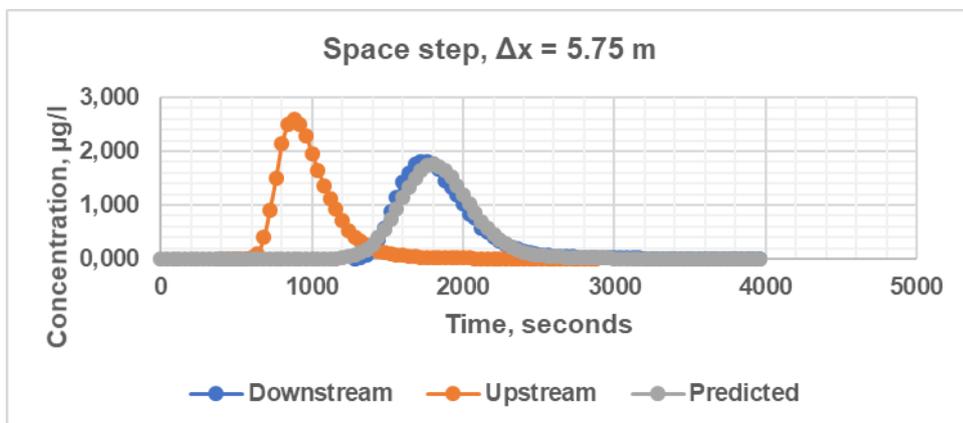
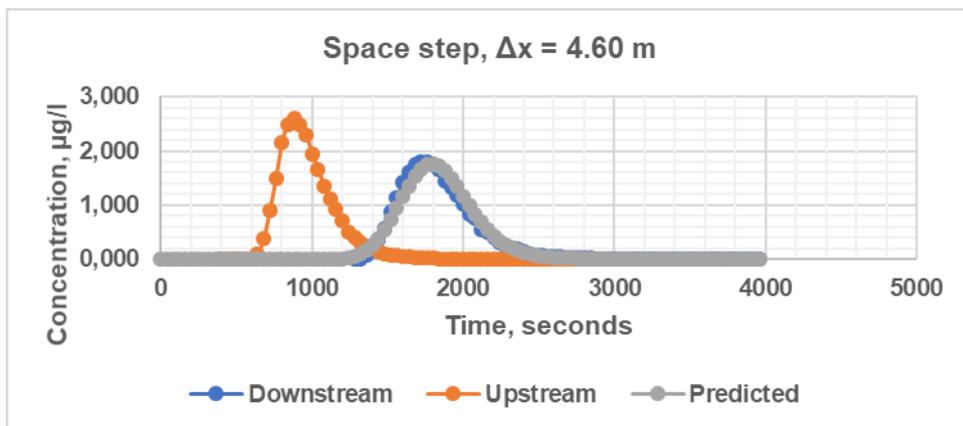
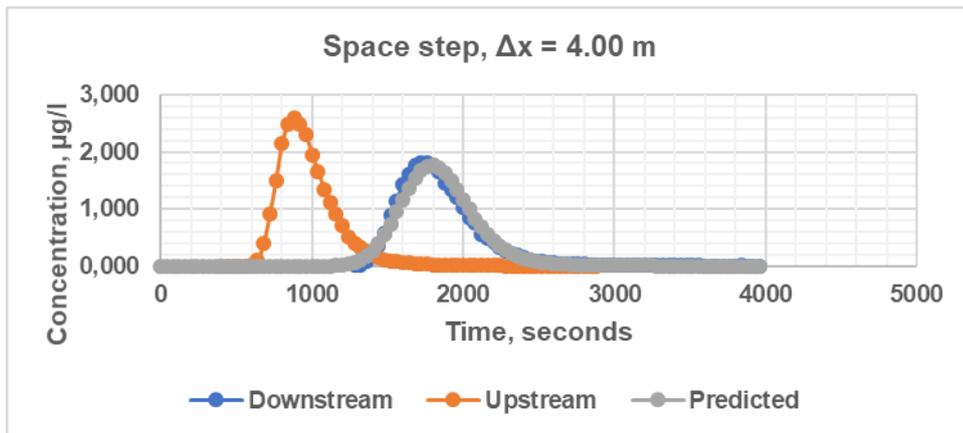
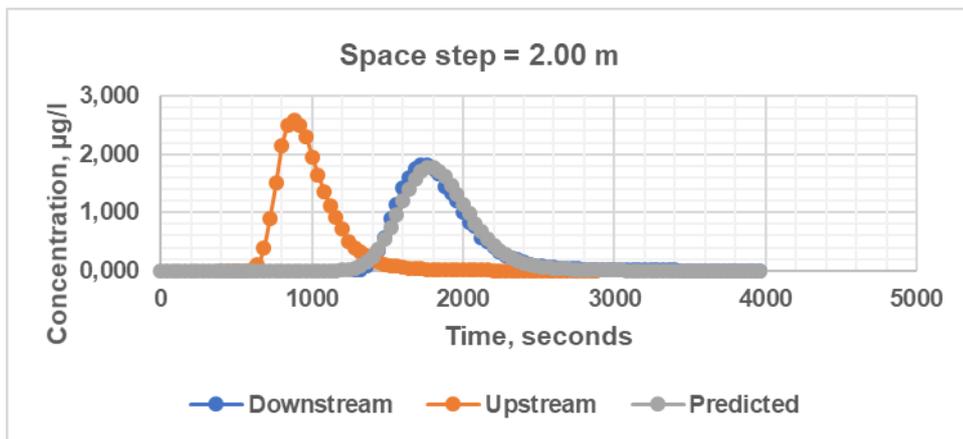


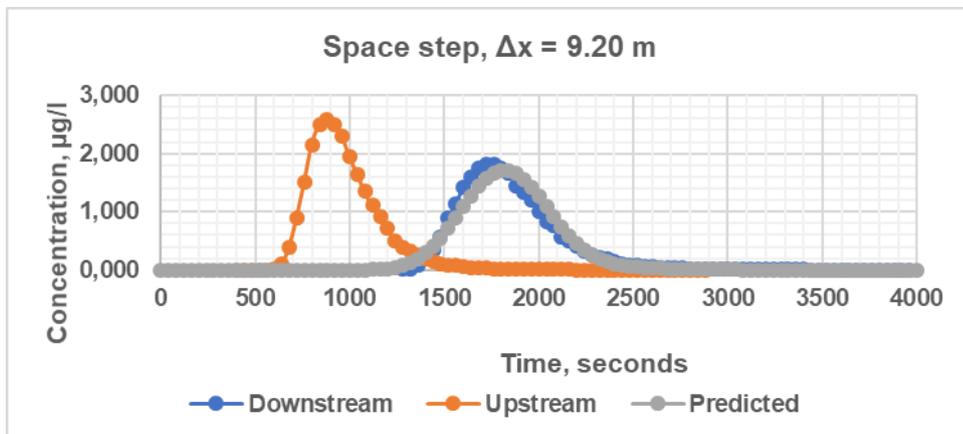
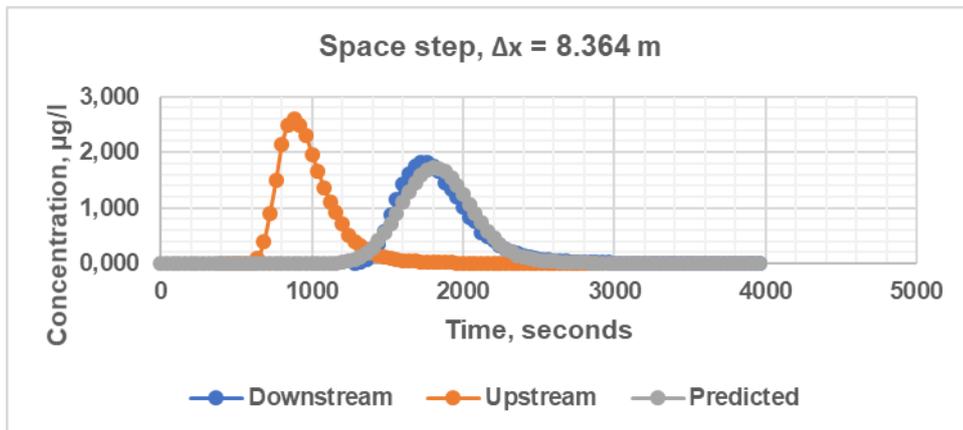
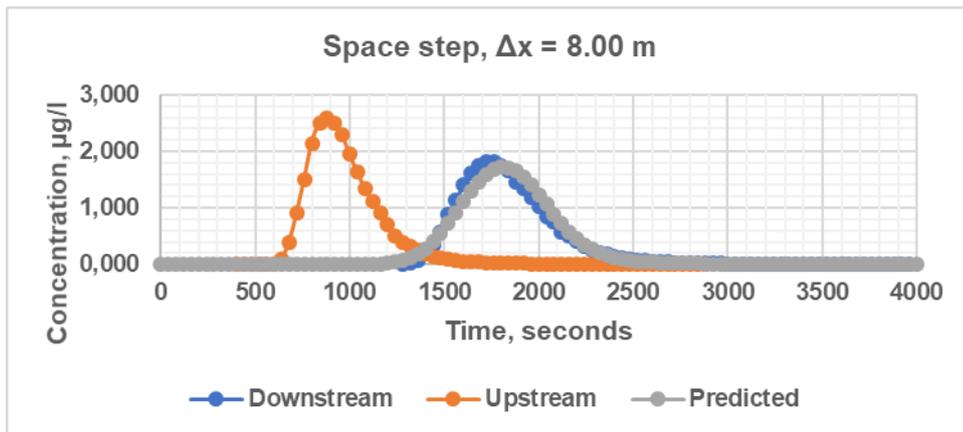
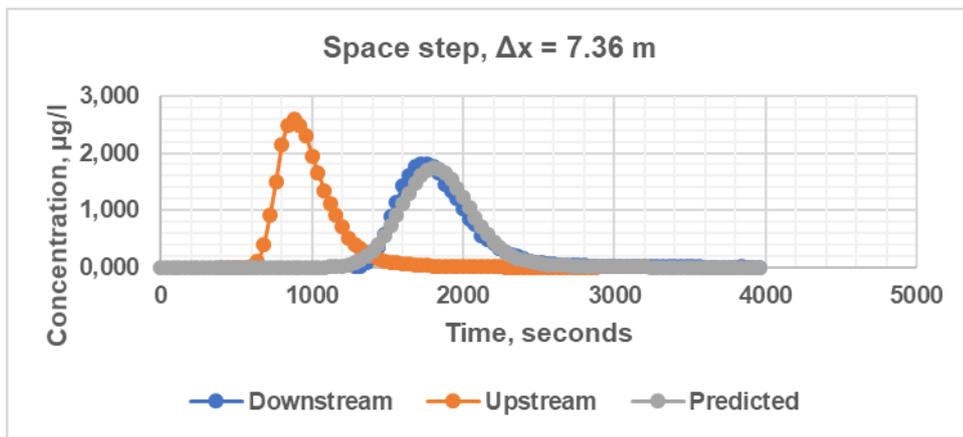


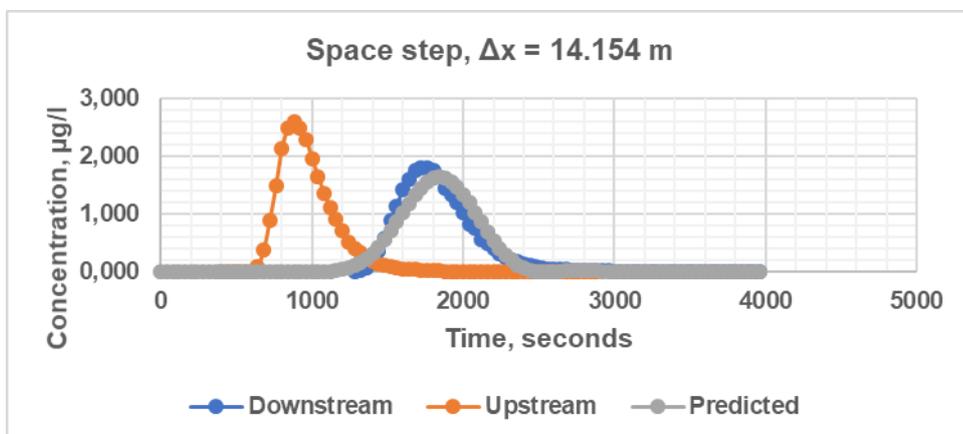
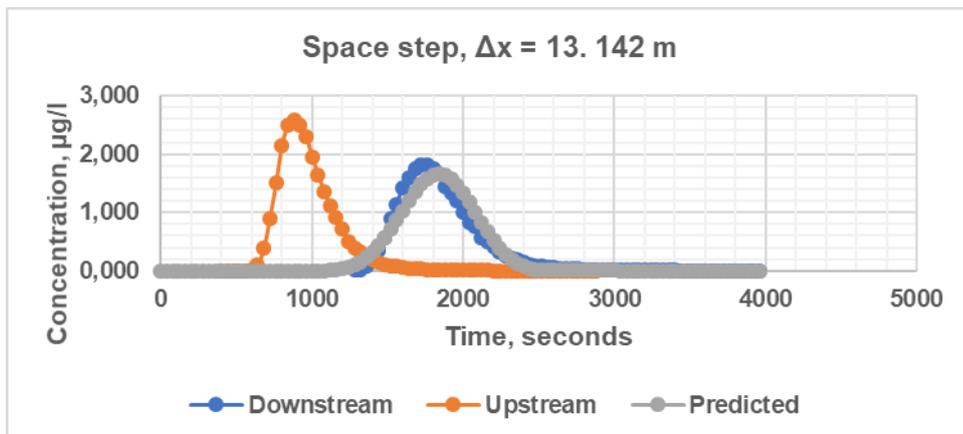
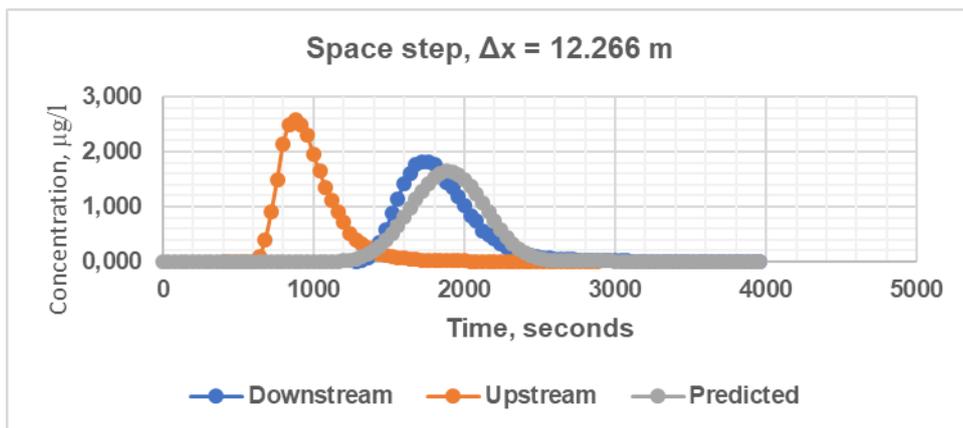
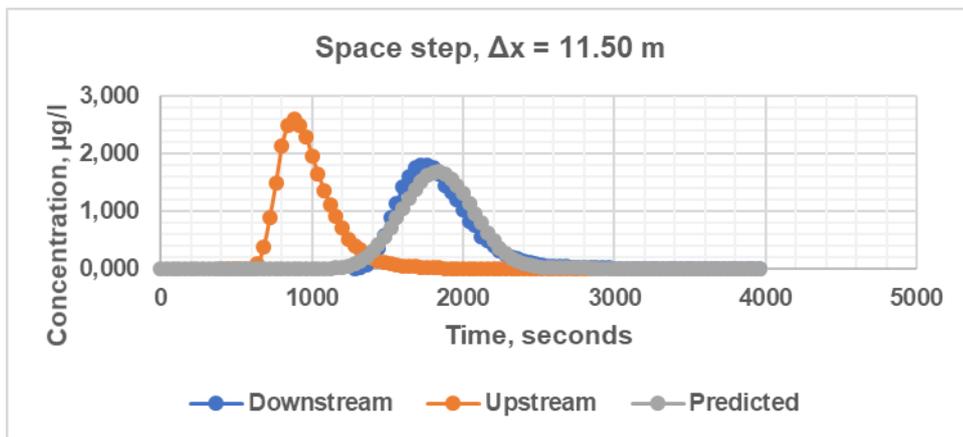


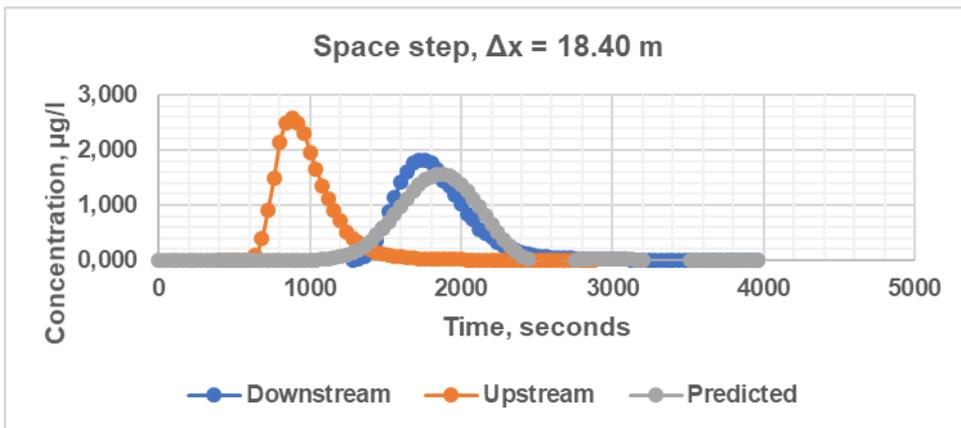
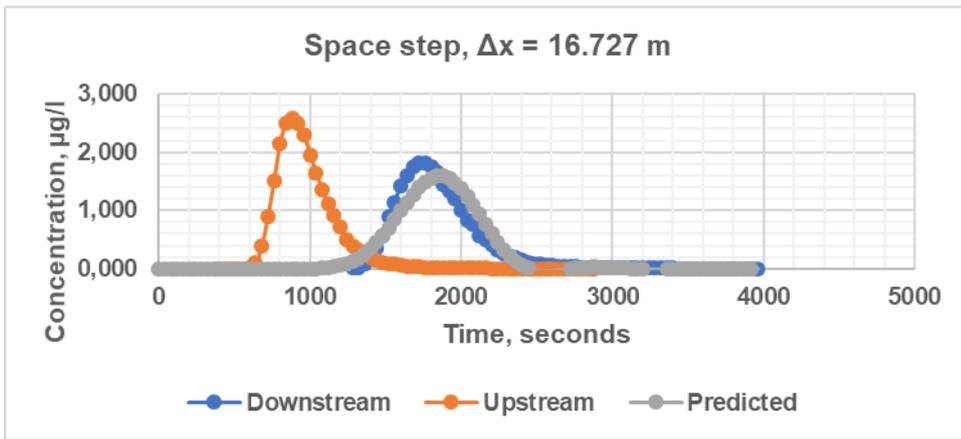


MacCormack method (experiment 13)

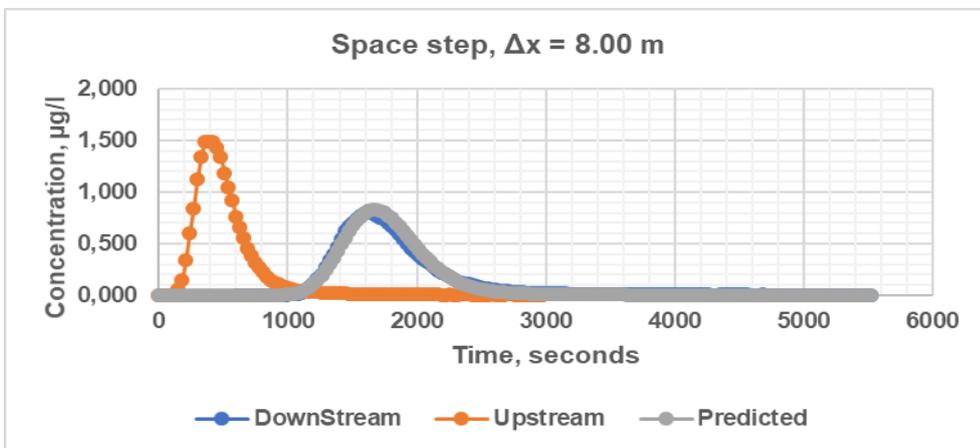


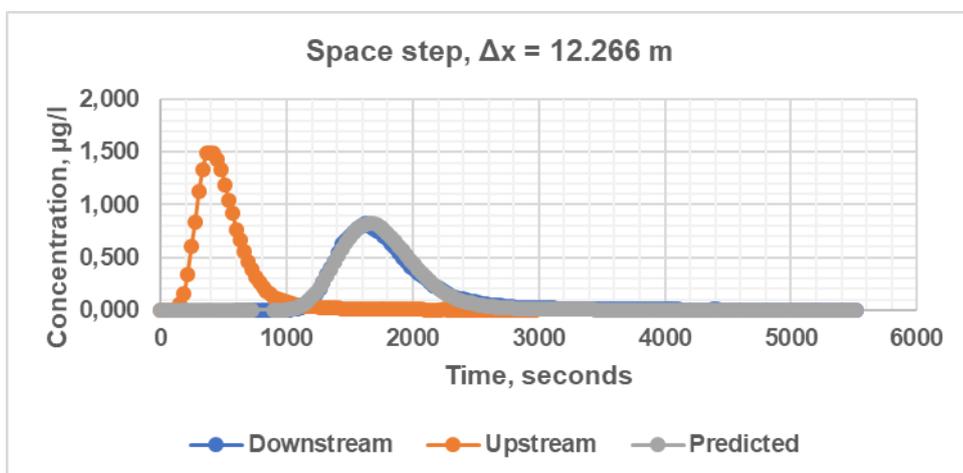
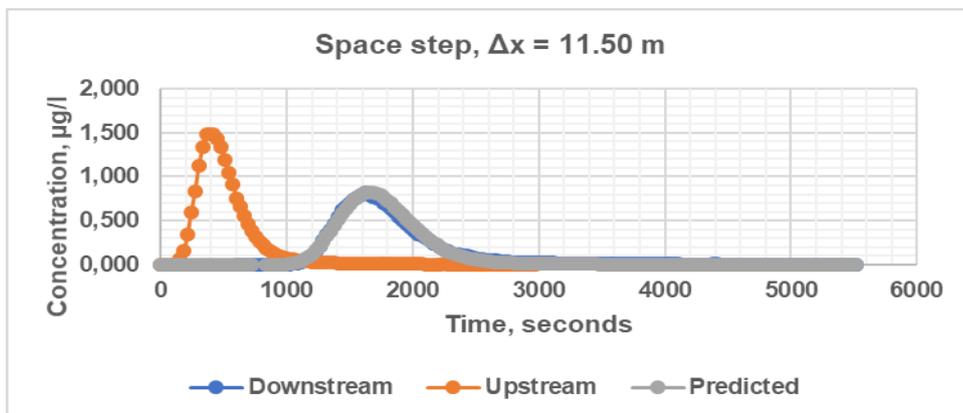
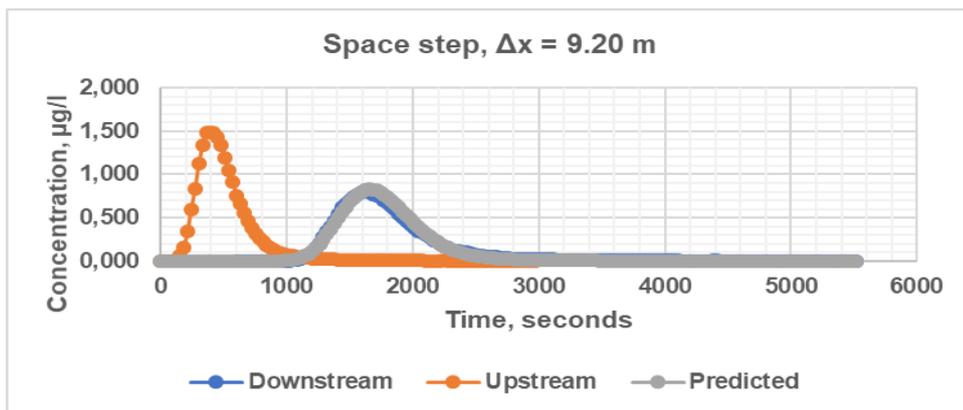
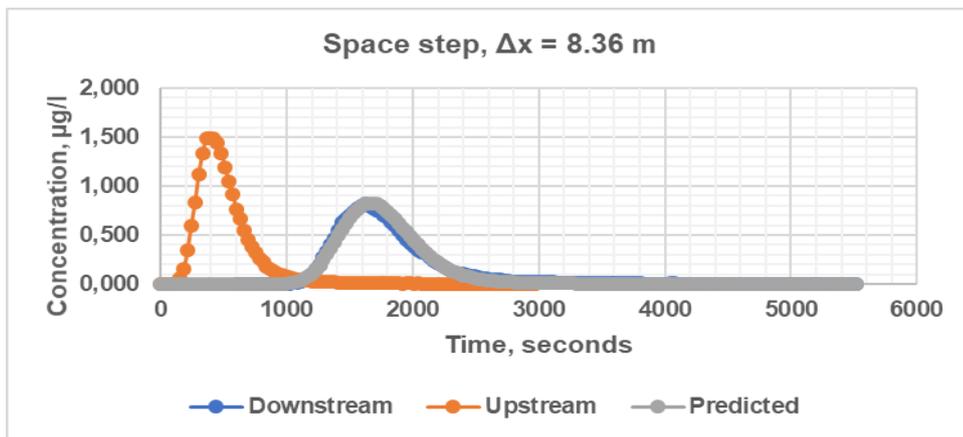


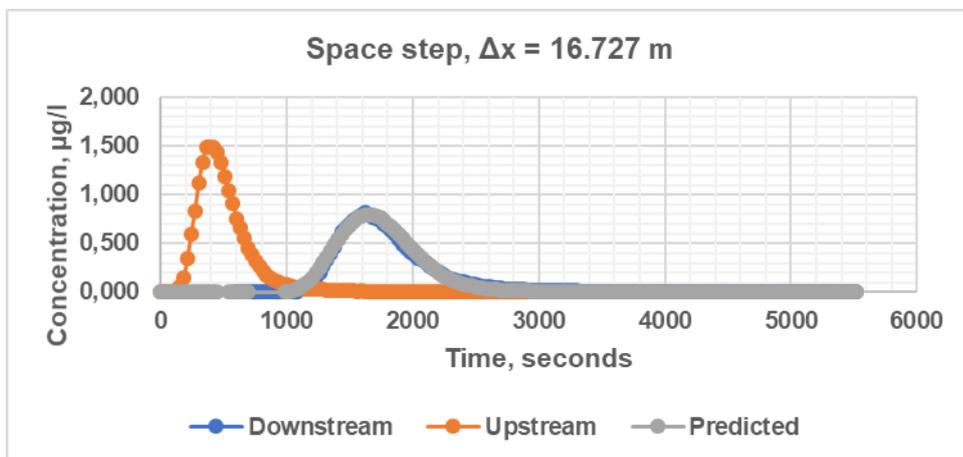
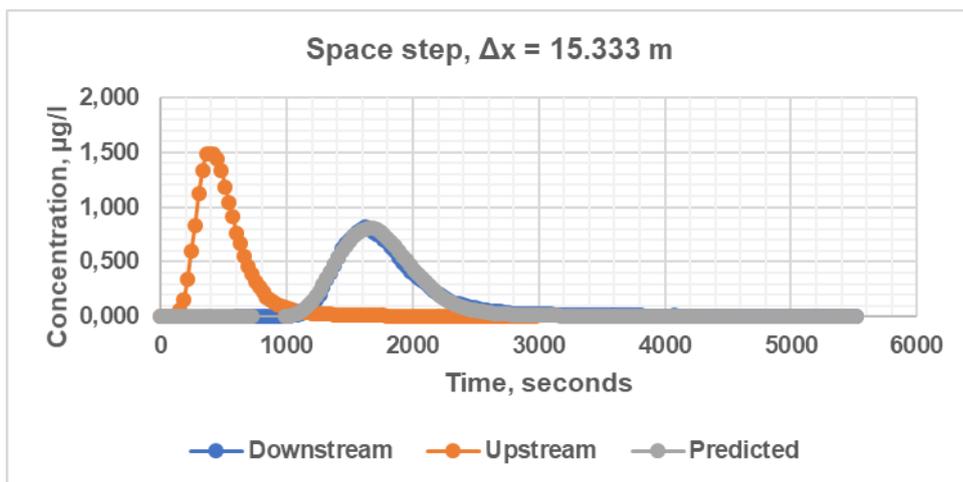
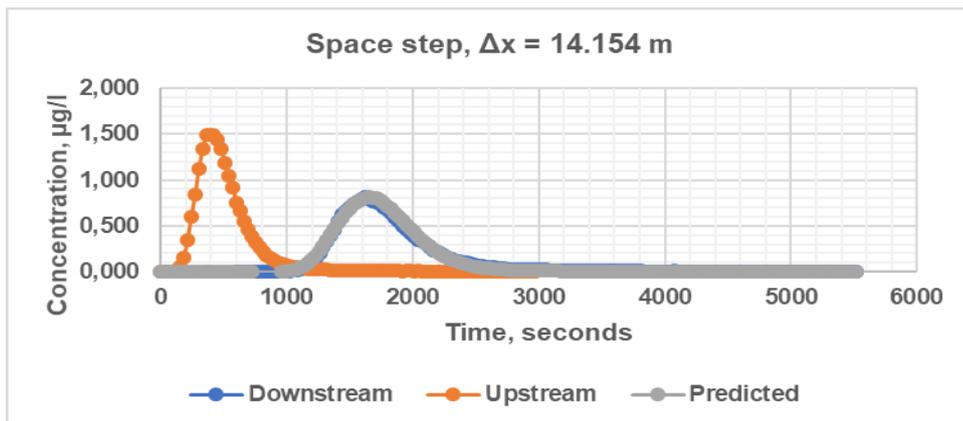
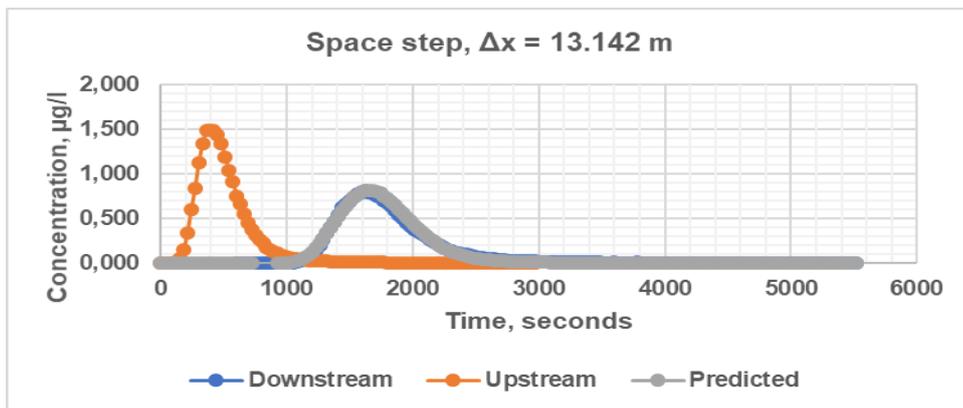


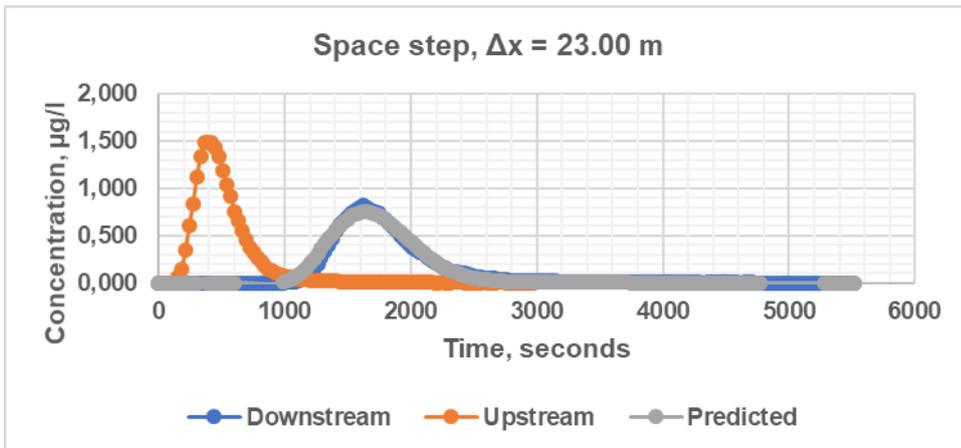
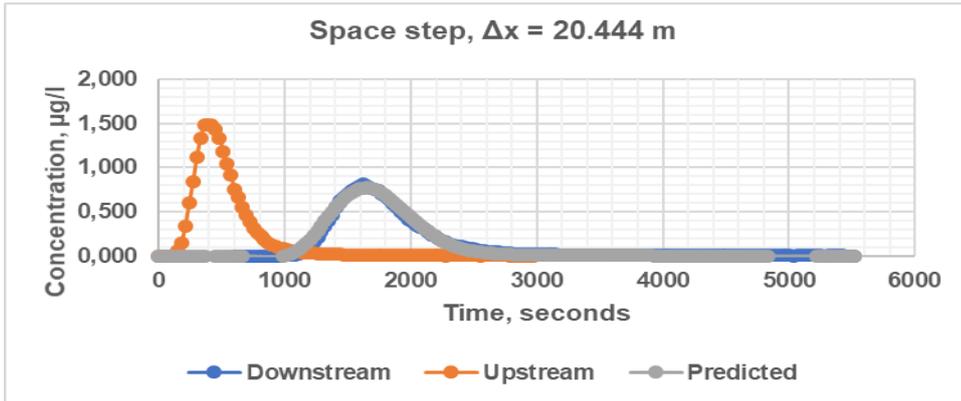
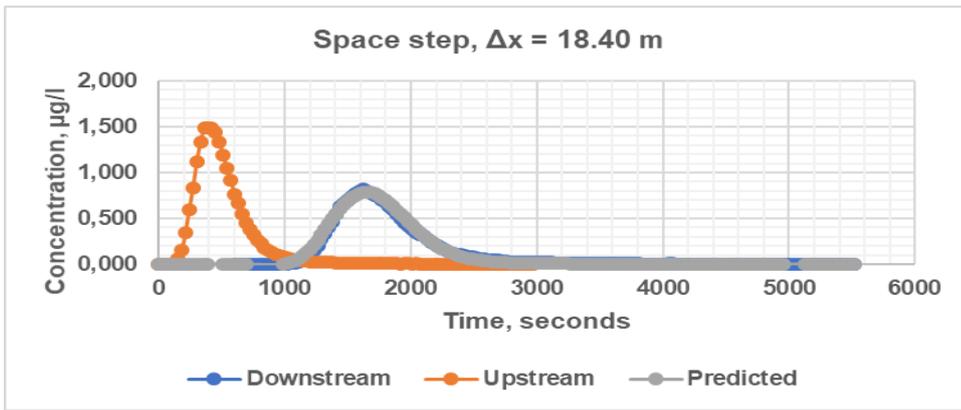


QUICKEST method (experiment 4)

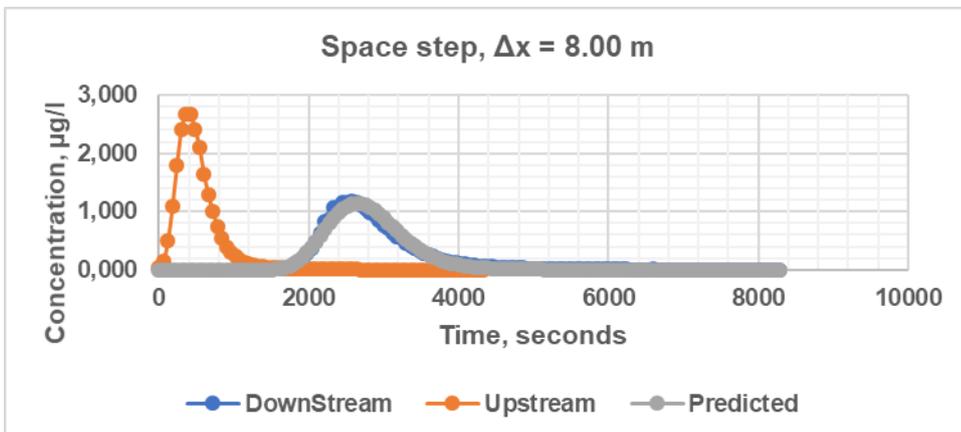


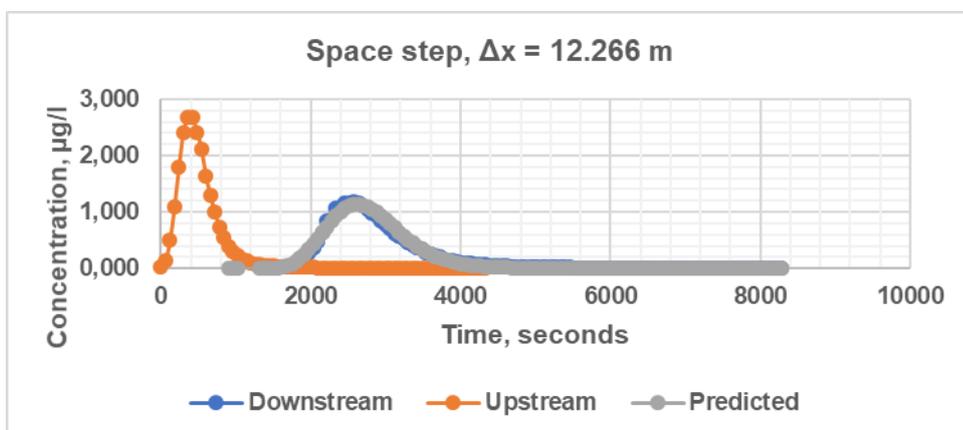
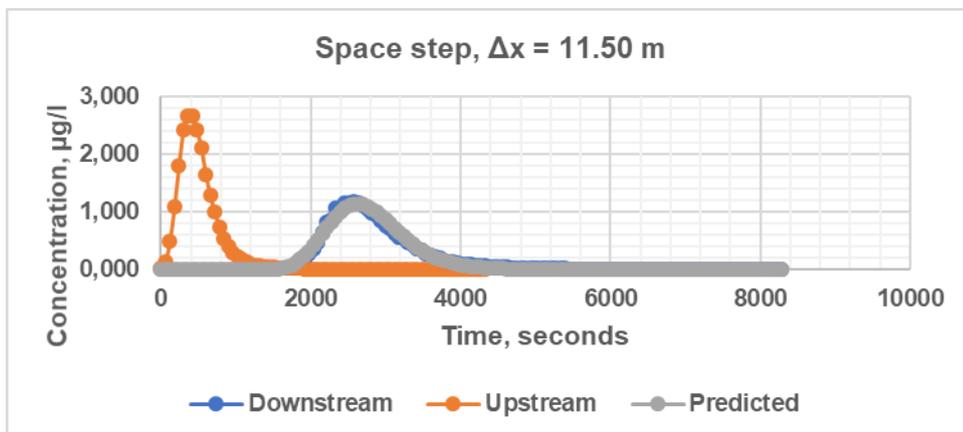
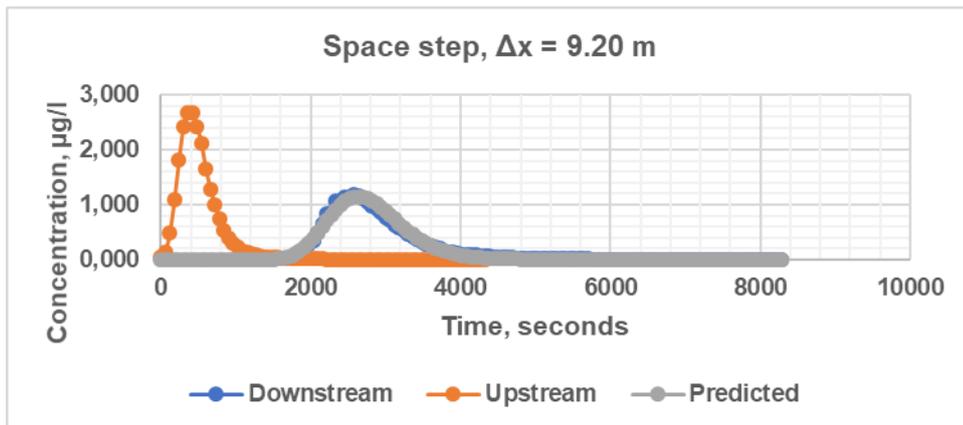
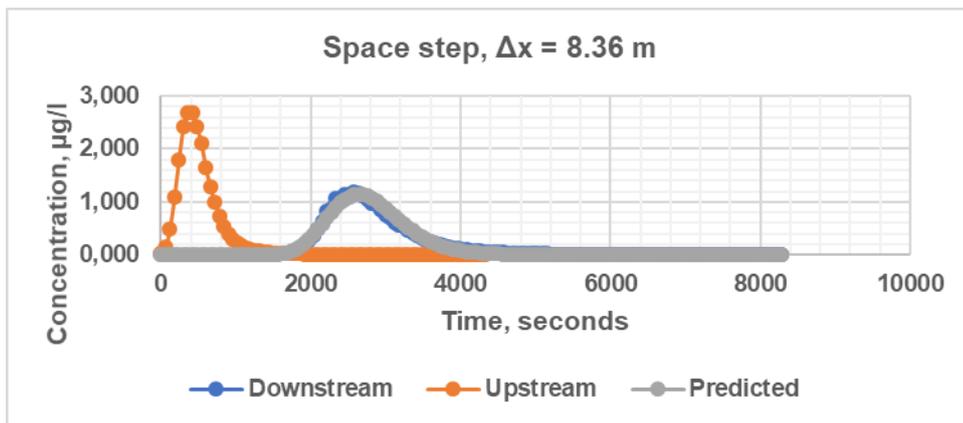


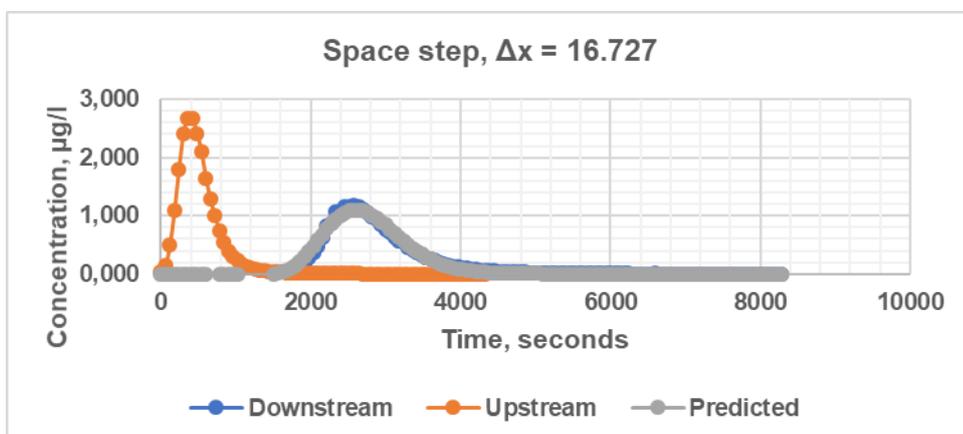
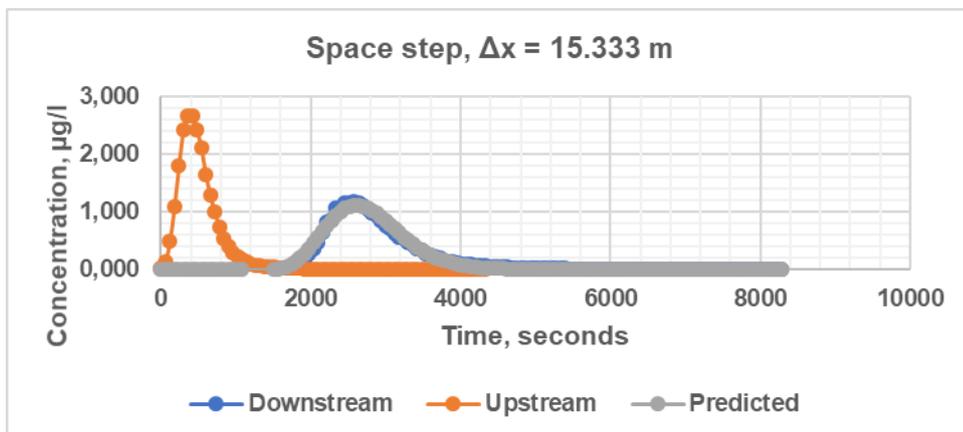
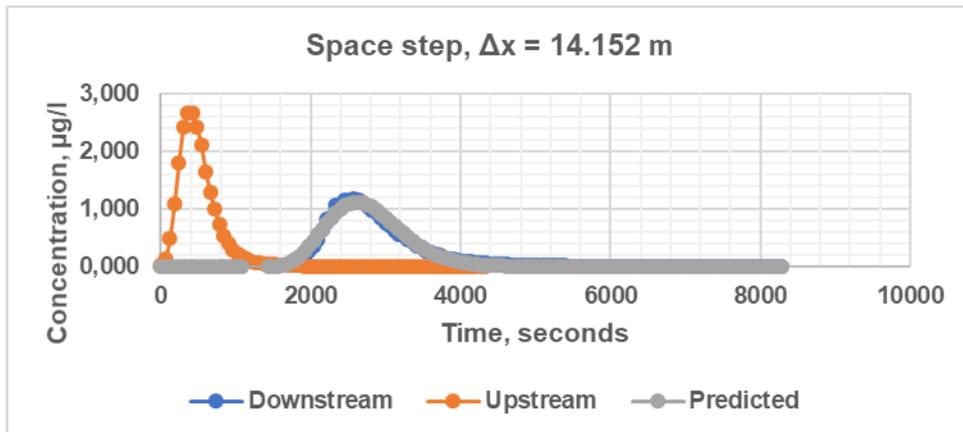
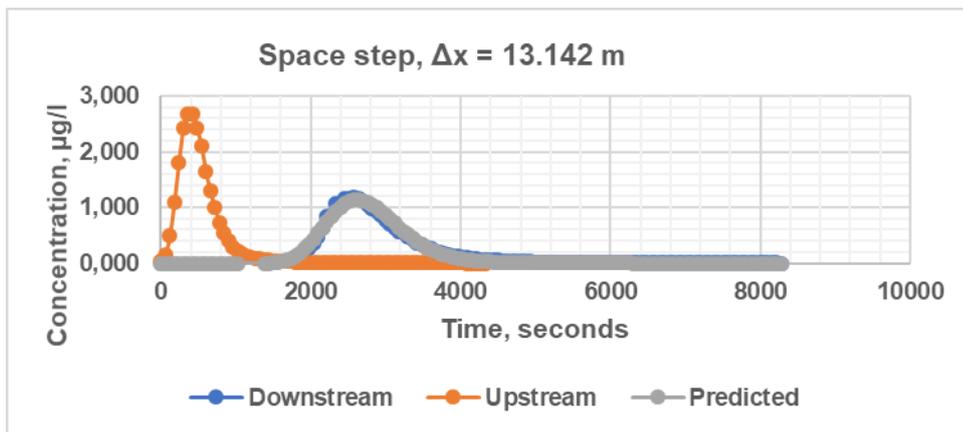


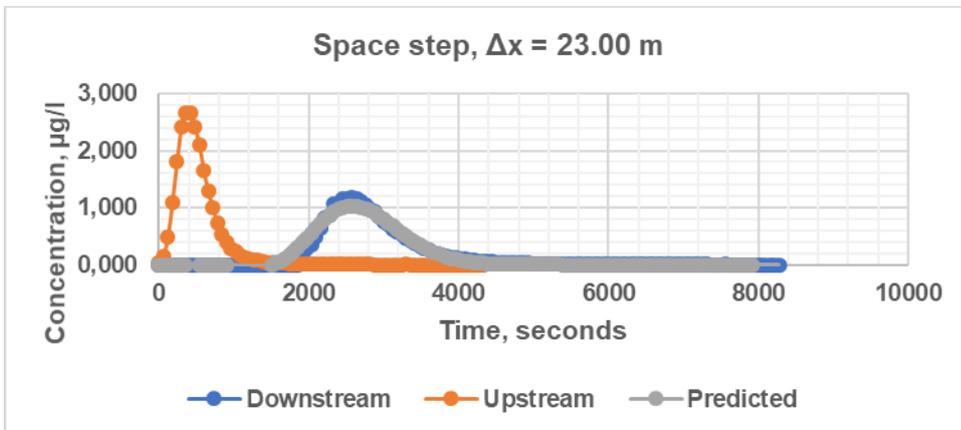
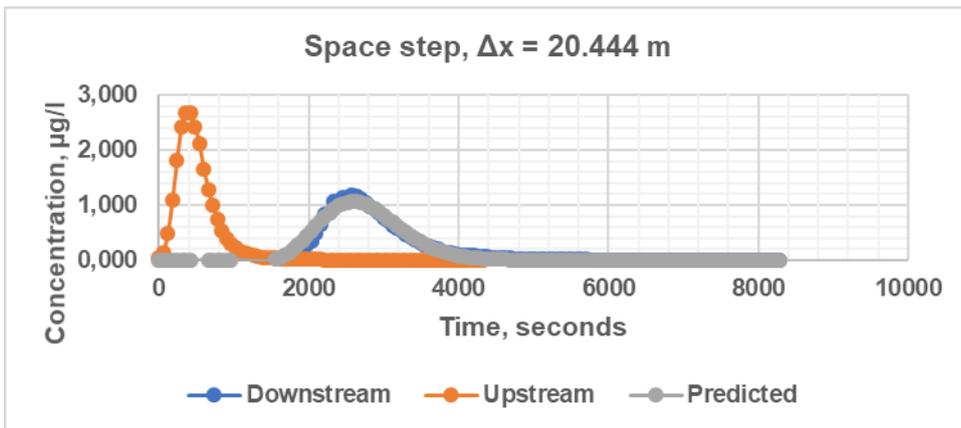
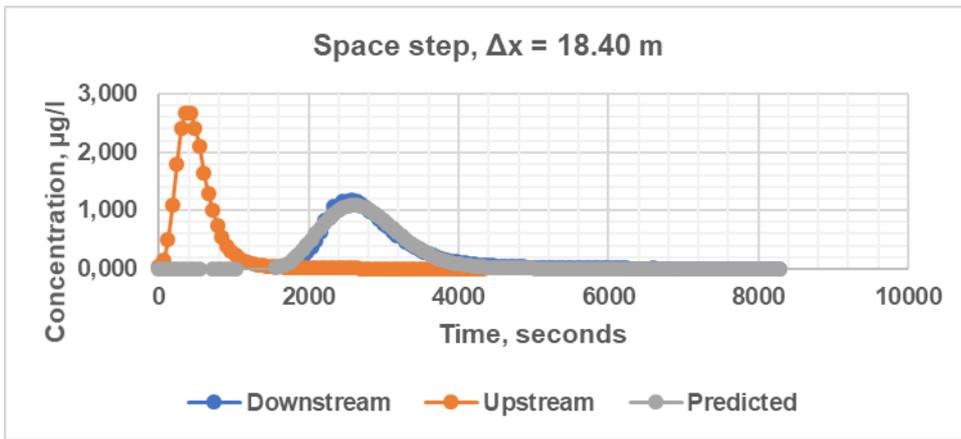


QUICKEST method (experiment 9)









QUICKEST method (experiment 13)

