QUICS: Quantifying Uncertainty in Integrated Catchment Studies

D 4.7 Tool to advise on using appropriate river pollutant transport model.

Lead Partner: University of Sheffield

v1.0: 30th May 2017

Report Details

Title: Tool to advise on using appropriate river pollutant transport model

Deliverable Number (If applicable): D 4.7

Author(s): Vivian Camacho Suarez, James Shucksmith, Alma Schellart

Dissemination Level: Public

Document History

Version	Date	Status	Submitted by	Checked by	Comment
0.1	04/04/2017	Draft First	Vivian Camacho Suarez	Alma Schellart	Need to check and update the references
0.5	15/05/2017	Second Draft	Alma Schellart	James Shucksmith	Some typos and changes in wording suggested
0.7	28/05/2017	Third draft	Alma Schellart	Will Shepherd and Simon Tait	Suggested tightening of focus of the report
1.0	30/05/3017	Final	Alma Schellart	Will Shepherd	

Acronyms and Abbreviations

ADE	Advection Dispersion Equation
ADZ	Aggregated Dead Zone
1D	1 Dimensional
2D	2 Dimensional

Acknowledgements

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000.

Executive Summary

This report describes the development and initial demonstration of a software tool developed to enable evaluation of structural and parameter uncertainty in models to describe pollutants in rivers. This tool can therefore assist in selecting appropriate river pollutant transport models.

The tool currently comprises of a flexible modelling framework which enables comparison of pollutant transport models of different complexity, as well as comparison of uncertainty in different longitudinal dispersion coefficients. The tool can generate concentrations for different combinations of inputs and models and calculates residuals by comparing model outputs (e.g. peak concentrations) to reference values. In the current tool, these reference values are the outputs from the more complex model (ADE 2D), the user can substitute this by observed field data. The tool was developed in Python and coded in a flexible way so that it can be adapted to suit the user's needs. Other parameters, such as for example hydraulic roughness, can be added to allow further investigation of its impact on uncertainty. The tool can also cope with input of probability distributions instead of single parameter values, in order to run Monte Carlo simulations. This deliverable report, however, only describes the application of the tool to compare different model structures, using rivers from the database described in Rutherford (1994).

Appendices to this report give an introduction to currently available river water quality models and software; a summary of the sources of uncertainty in river quality modelling and the definitions used throughout this report; and a description of the set-up and use of the modelling framework code. The code is available in the dissemination area of the QUICS project website www.quics.eu.

CONTENTS

1 Introduction

1.1 Partners Involved in Deliverable

University of Sheffield.

1.2 Deliverable Objectives

To prepare a flexible software tool that can be used to elucidate the relationship between different sources of uncertainty in surface water quality modelling for models of varying structure and complexity.

1.3 Background

Maintaining good surface water quality standards for different uses (e.g. drinking, recreation, ecological habitat) has become a challenging task due to stricter regulations, anthropogenic influences and climate change impacts. Environmental modelling assists in the assessment and improvement of surface water quality by simulating and predicting water quality conditions that may not be otherwise obtained from field monitoring.

Modelling the transport of solutes in rivers is important for both, point and non-point source pollution management. Point sources such as combined sewer overflows, industrial, domestic or wastewater discharges may have adverse effects on river health. Similarly, non-point source pollutants such as agricultural or urban runoff (for which quantification may be more difficult due to their diffuse nature) can alter the natural ecosystems in rivers leading to oxygen depletion and poor water quality conditions (Zheng et al., 2014). Therefore, to understand and take pro-active actions towards good water quality conditions, modelling offers a medium for water quality management. In order to model river water quality processes and quantify the associated uncertainties it is necessary to understand the natural river processes, modelling approaches, and the sources of uncertainty in the modelling approach.

Quantifying and communicating the accuracy of model predictions is also a key component for the proper management and decision making process (Refsgaard et al., 2006; van Griensven and Meixner, 2006). Moreover, the European Water Framework Directive suggests that water quality management should be addressed in an integrated manner so a good ecological status of the water bodies is reached (EC, 2000). In this integrated process, catchments are represented by models simplifying complex and non-linear processes. The scientific community (e.g. Dotto et al., 2012; Beck, 1987), recognize that modelling environmental systems can be highly uncertain, and there is a need to assess the accuracy of surface water quality models in order to be able to improve their predictions. Good quality high spatial and temporal resolution data is rare, which has made it difficult to predict these processes in detail.

River water quality modelling uncertainties are often inherited from the physical, mathematical and biochemical representation of pollutant transport processes and

transformation processes (structural uncertainty), quantification and selection of model parameters such as mixing and hydraulic parameters (parameter uncertainty), data inputs, and calibration (data input uncertainty).

This report will therefore describe the development of a modelling tool that can be used to enable evaluation of structural and model parameter uncertainty as well as input data uncertainty in models to describe pollutants in rivers. Appendices A, B, and C respectively provide descriptions of the river water quality models and available software; common sources of uncertainty; and the description of the modelling code developed for this deliverable.

1.4 The tool: a modelling framework for evaluating structural and parameter uncertainty in models to describe pollutant transport in rivers

This tool focuses on the model structures, model input data and parameters selected to describe pollutant transport in rivers since these have been less studied and contribute largely to the overall catchment uncertainty (Refsgaard et al., 2006; Zou et al., 2009). Understanding how parameter and structural uncertainties behave over different spatial and temporal scales provides a path towards uncertainty reduction in the modelling process. To determine the dominant uncertainties in pollutant transport modelling in rivers, the uncertainties due to the physical and biochemical model selection and parameter uncertainty at the various temporal and spatial scales are currently being studied.

The initial phase of this work, as described in this deliverable, investigates the effect of the model complexity on uncertainty over different spatial scales. By assessing how the uncertainties respond to changes in river aspect ratios, the behaviour of structural uncertainties can be studied.

A flexible modelling tool has been developed in the open source Python code (https://www.python.org/). The tool can generate concentrations for different combinations of input data and parameters and transport model structures and calculates residuals by comparing each model outputs (e.g. peak concentrations) to reference values. In the current framework, these reference values are the outputs from the more complex model (e.g. ADE 2D), the user can substitute this by observed field data.

2 The modelling framework

Flexible software code was developed according to the framework shown in Figure 1, to serve as a tool for uncertainty analysis considering a wide number of rivers with their corresponding hydraulic and geomorphological characteristics. The tool is flexible to implement several models and receive inputs in various forms, including probability distributions. The tool calculates soluble pollutant concentrations for various models and executes a structural and parameter uncertainty analysis. The framework contains data input, modelling and uncertainty analysis modules (Figure 1). The tool receives and stores the input parameters from a non-dimensional analysis for a specific river (e.g. narrow and slow or wide and fast). A setup in the tool creates a rectangular river reach with the given hydraulic and geomorphologic conditions. Consequently, the solute concentrations are calculated and plotted within the space and time domains. Similarly, the peak concentrations are calculated and plotted. The tool calculates the concentrations using the various pollutant transport models (with varying complexities) including the advection only, ADE and ADZ models. As part of future work, the tool will be extended to cover transient storage and transformation models. This will enable a comparison of concentration predictions from models with different complexities and parameter inputs.

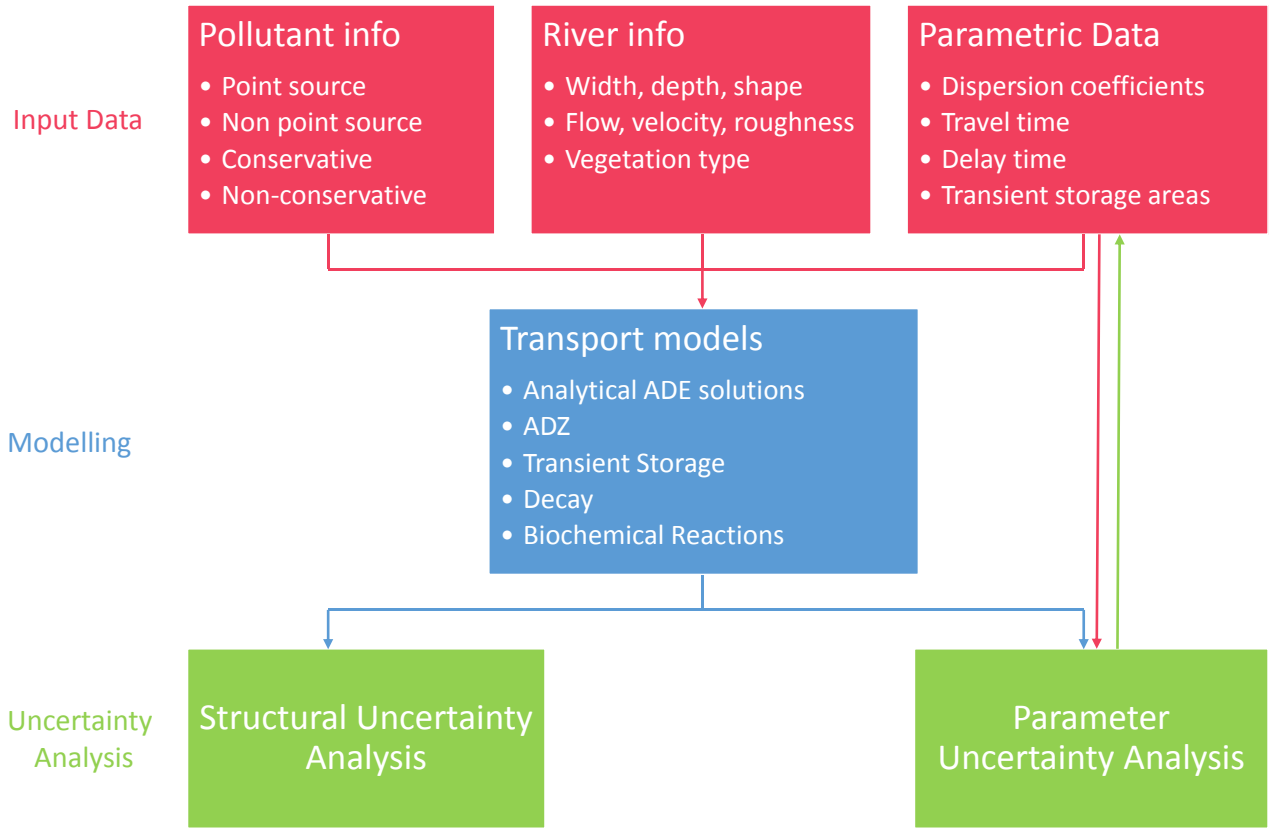

Figure 1. Modelling framework concept

Data input module

The data input module loads and stores the river, pollutant and mixing parameter variables required by the various models. These variables include the river width, depth, length, mean velocity, shear velocity, the pollutant mass, longitudinal and transverse dispersion coefficients, mean travel time and residence time. The module is flexible so that these can be implemented as deterministic values or as probability distributions. Both spatial and temporal domains can be selected for analysis. The data input module can calculate either the width or depth from the width to depth ratio and the mean velocity or shear velocity from the mean to shear velocity ratio depending on the user specification. If the shear velocity is not specified, the shear velocity is calculated using the relationship $U^* = \sqrt{g\ h\ S}$ where g is the gravitational constant, h is the depth and S is the channel slope. The tool currently only handles steady uniform flow. The data input module can also calculate the dispersion coefficients parameters from various methods as instructed by the user.

Modelling Module

To date, the module includes models describing the processes of (1) advection only in the longitudinal direction, (2) advection-dispersion in the longitudinal direction (ADE 1D), (3) advection and dispersion in both the longitudinal and transverse direction (ADE 2D), and (4) aggregated dead zones. These models are briefly explained below with their analytical solutions where applicable. In the future, the biochemical transformation models will also be included in this tool.

Advection only assumes a pollutant is cross-sectionally well mixed and is based on the river velocity as explained in Rutherford (1994) where the solute is transported only due to the longitudinal mean velocity. This is also termed plug-flow, i.e., there is no change in the concentration or spread of the pollutant.

Advection Dispersion Equation 1D (ADE 1D) – This equation is a depth and width averaged form simplified from the three-dimensional advection dispersion equation, assuming full instantaneous cross sectional mixing of a pollutant downstream of a release. The Fischer et al. (1979) analytical solution to the ADE 1D equation is used, where the pollutant concentration C(x,t) is:

Equation 1

$$C(x,t) = \frac{M}{A\sqrt{4\pi\ D_x t}} \exp\left[-\frac{(x - V_x t)^2}{4\ D_x t}\right]$$

Where M is the mass of the pollutant released at t=0 and x=0 and A is the cross sectional area of the channel, D_x is the longitudinal dispersion coefficient, V_x is the average velocity in the longitudinal direction, x is the spatial location and t is time. The advection dispersion equation is widely used to predict solute concentrations within water quality modelling tools (Kashefipour and Falconer, 2002).

Advection Dispersion Equation 2D (ADE 2D) **–** This equation is a depth averaged form of the simplified form of the three-dimensional advection dispersion equation, assuming full

instantaneous mixing of a pollutant over the flow depth downstream of a release. This model neglects transverse velocities (V_x>>V_y), but includes the effect of mixing in the transverse direction. Concentrations are estimated using the analytical solution suggested by Fischer et al (1979):

Equation 2

$$C(x,y,t) = \frac{M}{4\pi dt\sqrt{D_x D_y}}\exp\left[-\frac{(x-V_x t)^2}{4\,D_x t} - \frac{y^2}{4\,D_y t}\right]$$

Where the transverse dispersion coefficient D_y and transverse dimension y are introduced into the equation. The release of pollutant with mass M occurs at x=0, y=0 (at the middle section of the stream width) and time t=0.

Aggregated Dead Zone (ADZ) **–** The ADZ model assumes that dispersion occurs mainly due to dead zones. The model developed by (Beer and Young, 1983) estimates C using the discretization

Equation 3

$$C(x_2,t) = \alpha\, C(x_2\,, t-1) + \beta\, C(x, t-\delta)$$

Where

$$\alpha = -\exp\left(\frac{\Delta t}{T_R}\right) \qquad \beta = 1 + \alpha$$

$$\delta = \tau/\Delta t \qquad T_R = \bar{t} - \tau$$

T_R is the residence time, τ (Tau) is the time delay, Δt is the time step, and $\bar{t}$ (Tbar) is the average travel time.

Structural uncertainty analysis module

To determine the deficiencies in modelling due to the mathematical representations of the models, it was initially assumed that the most complex model was the most accurate representation of reality. Therefore, the ADE 2D model was used as baseline for comparison of the other models. The modelling module obtains the estimated concentrations from the advection only, ADE 1D and ADZ models and compares this against the ADE 2D (and the framework has the flexibility to include other benchmarks). The module then estimates the peak concentrations with respect to time. The peak concentrations are normalized against the maximum concentrations and the models can be compared. The uncertainty analysis module estimates the differences in each models' estimations of peak concentrations in the river as a function of time. The model differences are referred as residuals. Plots of the residuals against time or distance are created to observe the spatial and temporal behaviour of the model results.

Parameter uncertainty analysis module

The tool is coded so that parameter uncertainty can be analysed using Monte Carlo simulations, and the tool is flexible so that it can cope with input of both single values as well as probability distributions.

As initial part of this study, a selection of methods for determining the longitudinal dispersion coefficient (Dx) was obtained from: El Kadi Abderrezzak et al. (2015). These include: Elder (1959), Fischer et al. (1979), McQuivey and Keefer (1974), Liu (1977), Isawa and Aya (1991), Magazine et al. (1988), Koussis and Rodriguez Mirasol (1998), Seo and Cheong (1998), Deng et al. (2001), Kashefipour and Falconer (2002) for B/H >50 (B/H = width to depth ratio), and Kashefipour and Falconer (2002) for B/H <50. See Appendix B for the full set of equations. Most of the methods for quantifying the longitudinal coefficient are empirical relationships based on the aspect ratio (river width over depth) and hydraulics of the river channel (mean and shear velocities). The initial part of this study focusses on determining the impact of using the different longitudinal dispersion coefficients on pollutant peak concentrations by comparing model predictions using the different dispersion coefficient equations.

A comparison of the model predictions due to the inherent uncertainty in using deterministic empirical dispersion coefficient equations by using Monte Carlo simulations and probability distributions derived to describe uncertainty in Longitudinal Dispersion Coefficients is currently being carried out (Camacho Suarez et al., In preparation).

3 Example demonstration of the modelling framework tool

Structural uncertainty analysis

Camacho et al. (2015) carried out an analysis to understand the impacts of differences in structures of river mixing models on the estimated pollutant concentrations for rivers of different characteristics. The Rutherford (1994) river database to classify rivers was used to demonstrate the use of the tool for structural uncertainty analysis. Table 1 shows the minimum, maximum and quartiles of the mean to shear velocity ratios and the aspect ratios in the database. A wide river was defined as a river that had a large width with respect to its depth, thus a large aspect ratio. Similarly, a fast river was defined as a river with a large mean velocity and small shear stress resulting in a high mean to shear velocity ratio. Peak concentrations were estimated for each model using the longitudinal mixing coefficient from Elder (1959) and the transverse mixing coefficient from Fischer et al. (1979). The parameters tau and mean travel time for the aggregated dead zone model were obtained from the relationship given in Lees et al. (2000). For the uncertainty analysis, the residuals for each model were calculated. The residuals were the result of the differences in peak concentrations from the studied model and the complex model (ADE 2D). The assumption that the ADE 2D model was the most accurate model was made. With the classification presented in Table 1, the modelling tool was used to calculate the concentrations and the differences in models for a 'wide' versus a 'narrow' river, and a 'fast' versus 'slow' river (see Table 2).

Table 1. Distribution of mean to shear velocity and width to depth ratios obtained from Rutherford's river characteristics database

	Vx/U*	***B/H***
MIN	1.1	2.2
Q1	4.5	21.7
MEDIAN	6.1	37.9
Q3	9.0	60.0
MAX	21.0	174.0

Table 2 and Figure 2 show the differences between a narrow river (aspect ratio = 2.2) and a wide river (aspect ratio = 173.8) which have similar mean to shear velocity ratios. The modelling outputs (Figure 2) show the calculated concentrations of the models 50 m downstream of the pollutant release at the centre of the cross section and the residuals from the differences in peak concentrations and the more complex ADE 2D model. Figure 2 shows that in the narrow river, the advection only model over predicts the concentrations while in the wide river, the advection only model under predicts when comparing it to the ADE 2D. This over prediction of the advection only model in a narrow river is expected because the advection only model assumes instantaneous mixing over the cross section, and diluting the mass of the pollutant over a smaller cross section results in a higher concentration as opposed to a wider cross section where the dilution occurs over a larger

area leading to a smaller concentration. This behaviour of over prediction in the narrow river and under prediction in the wide river of the advection only is observed even at longer time scales (observed in the residuals of the concentrations at the channel centreline as shown in Figure 2). As opposed to the advection only model, the ADE 1D model has a closer prediction to the ADE 2D model for the narrow river than the wide river where the concentrations are under predicted. This indicates that for a narrow river, the ADE 1D and the ADE 2D model predictions potentially converge faster than in wide rivers. This is confirmed by looking at the residuals, where it can be seen that for the residuals decrease with time and the models' predictions converge albeit within 60 seconds in the narrow river and 600 seconds in the wide river. This is expected because as the tracer travels downstream in the ADE 2D model, transverse mixing occurs until the complete cross section is mixed and hence conditions become similar to the ADE 1D model where complete cross-sectional mixing is assumed at the pollution source. Once both models are cross-sectionally well mixed, the models result in similar predictions and the structural uncertainty decreases. The ADZ model under predicts the concentrations for both cases, the wide and narrow river.

Table 2. Hydraulic and geomorphological properties of Rivers Yuma Mesa ('narrow') and Mississippi ('wide')

	Yuma Mesa	**Mississippi**
Depth (m)	3.45	3.1
Width (m)	7.6	530
Aspect ratio (-)	2.2	173.8
Mean velocity (ms^{-1})	0.68	0.08
Shear velocity (ms^{-1})	0.047	0.0056
Mean/shear velocity ratio (-)	14.47	14.2
Slope (-)	0.000065	0.000001
Dx (m^2s^{-1})	0.961	0.098
Dy (m^2s^{-1})	0.024	0.002
Pollutant mass (kg)	5.0	5.0
Tau (s)	125.9	1103.1
Tbar (s)	151.4	1311.8

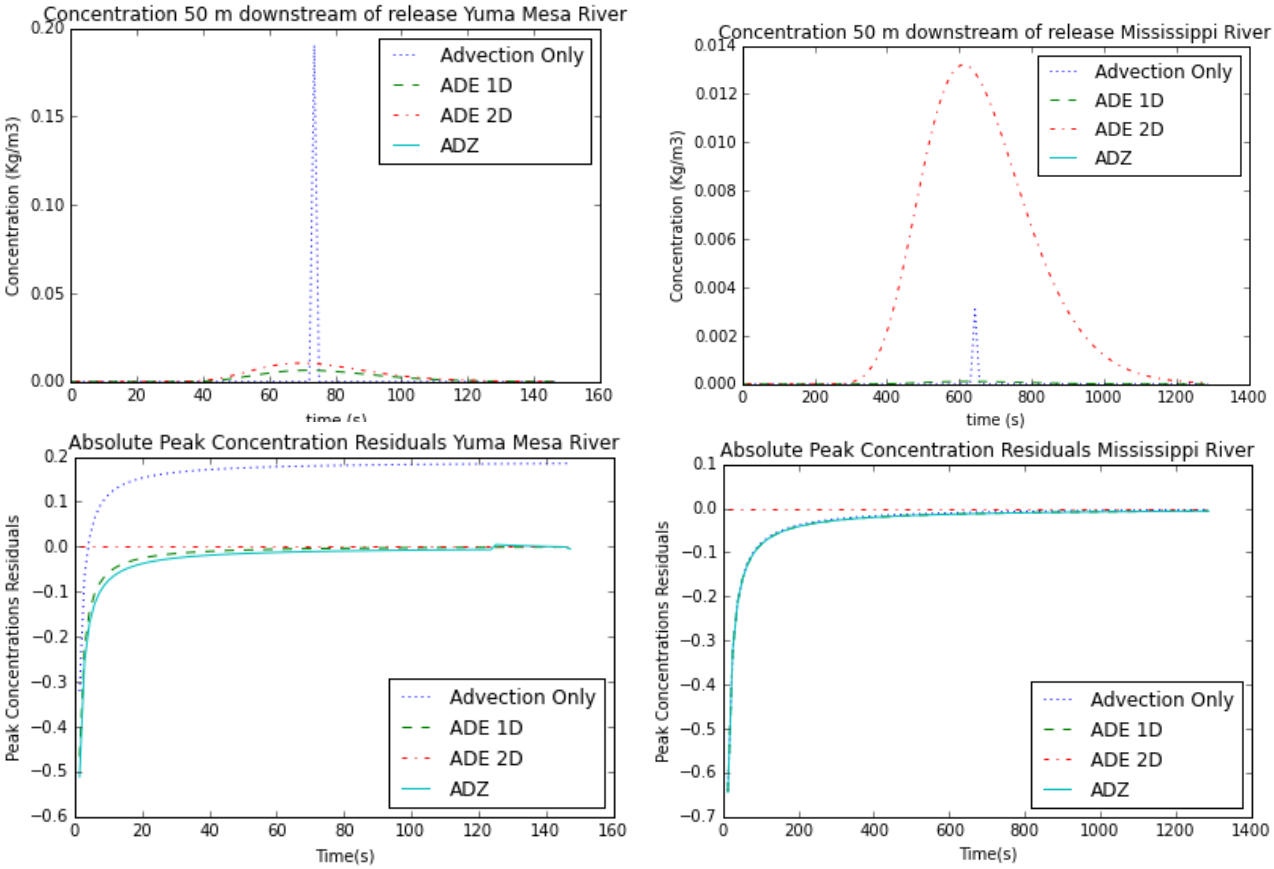

Figure 2. Uncertainty tool outputs for Yuma Mesa and Mississippi Rivers: The Yuma Mesa River is a narrow river where the advection-only model overestimates the tracer concentrations while the Mississippi River is a wide river where the advection-only model under-predicts the concentrations. NB the plots show concentrations at the channel centre line.

Similarly, the effect of the hydraulic conditions was analysed by comparing a slow to a fast river. The slow river had a mean to shear velocity ratio of 1.24 while the fast river had a mean to shear velocity ratio of 16.5, both rivers had similar aspect ratios (Table 3). Figure 3 shows that in both cases the advection only model over predicts when compared to the ADE 2D in both slow and fast rivers. However, in the fast river, the differences between the peak concentrations are smaller. This is expected since in faster and more turbulent rivers, mixing in the transverse direction occurs more quickly.

Table 3. Hydraulic and geomorphological properties of Rivers Punehu (slow) and Coachella (fast)

	Punehu	**Coachella**
Depth (m)	0.28	1.56
Width (m)	5.0	24.0
Aspect ratio (-)	17.9	15.4
Mean velocity (ms^{-1})	0.26	0.71
Shear velocity (ms^{-1})	0.2098	0.0428
Mean/Shear velocity (-)	1.24	16.5
Slope (-)	0.0160	0.0001
Dx (m^2s^{-1})	0.395	0.395
Dy (m^2s^{-1})	0.008	0.010
Pollutant mass (kg)	5.0	5.0
Tau (s)	330.2	127.8
Tbar (s)	394.7	142.9

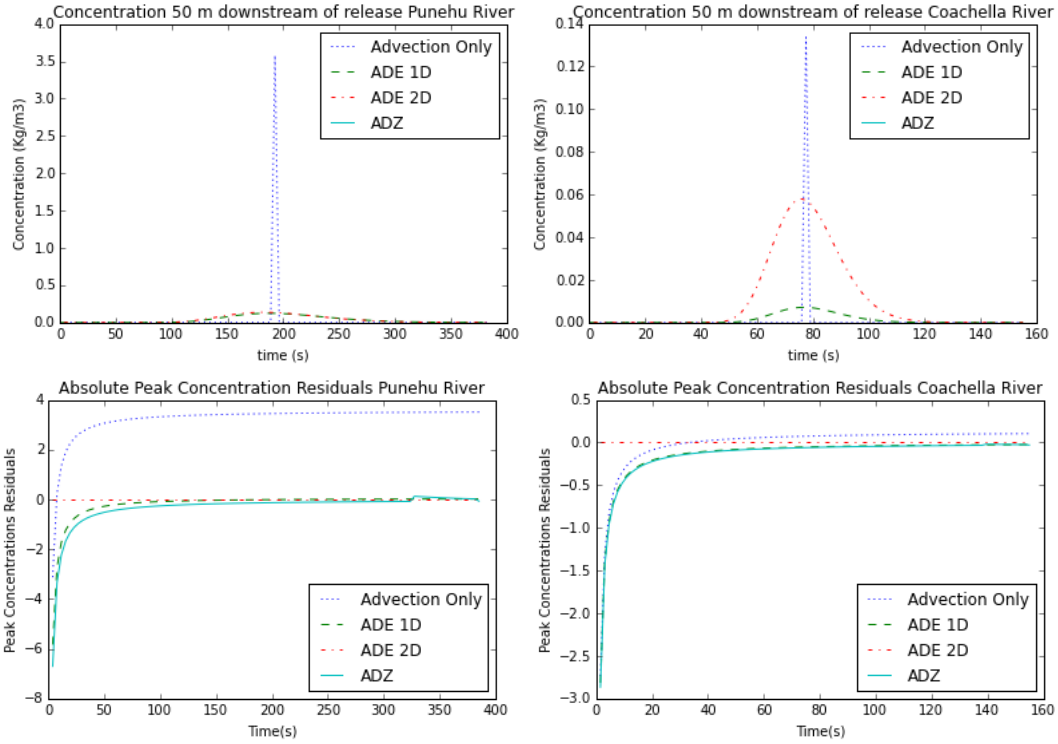

Figure 3. Modelling framework outputs for the Punehu (slower) River and Coachella (faster) River: The Punehu River is a slower River than the Coachella River. In both rivers, the advection-only over predicts the concentrations. NB the plots show concentrations at the channel centre line.

Initial work on Longitudinal Dispersion Coefficient Uncertainty Analysis

To understand the parameter uncertainty inherited from the longitudinal dispersion coefficient (Dx) on the one-dimensional advection dispersion model (ADE 1D), ten equations for calculating Dx were retrieved from El Kadi Abderrezzak et al. (2015). These equations are listed in Table 4. One river was selected from the database (Table 5), and

the estimations for the ADE 1D model were computed using the different longitudinal dispersion coefficients. Figure 4 shows the range of peak concentrations against time obtained when using various dispersion coefficients. Elder's longitudinal dispersion coefficient (0.6 m^2s^{-1}) predicts higher peak concentrations than the other relationships of dispersion coefficients. Elder's equation estimates higher coefficients because it is proportional to the multiplication of depth and shear velocity while some of the other equations are inversely related to the shear velocity. In addition, Elder's equation does not include the transverse variations in velocity (El Kadi Abderrezzak et al., 2015). This leads to a lower dispersion rate. On the cases of Seo and Cheong (1998), Deng et al. (2001), and Kashefipour and Falconer (2002) equations, these have been obtained by regression of larger data sets, resulting in more variables such as the shear to mean velocity ratio and slope properties.

Table 4. Equations selected for initial work on longitudinal dispersion coefficient.

Model developed by	Equation
Elder (1959)	D_x = 5.93*h*Vs
Fischer (1975)	D_x = 0.011 *(V/Vs)^2*(B/h)^2 *h*Vs
McQuivey and Keefer (1974)	D_x = 0.058 *h*V/Se
Liu (1977)	D_x = 0.18 * (V/Vs)^0.05 *(B/h)^2 *h*Vs
Isawa and Aya (1991)	D_x = 2*h*Vs*(B/h)^1.5
Magazine et al. (1988)	D_x = 75.86*(0.4*V/Vs)^-1.632 *Rh*V
Koussis and Rodriguez Mirasol (1998)	D_x = 0.6*h*Vs*(B/h)^2
Seo and Cheong (1998)	D_x = 5.92 * (V/Vs) ^1.43 *(B/h) ^0.62 *h*Vs
Deng et al. (2001)	D_x = (0.15/ (8 * Et)) *(V/Vs)^2 *(B/h) ^1.67 *h* Vs
Kashefipour and Falconer (2002) for B/h >50	D_x = 10.612 * (V/Vs)* h* Vs
Kashefipour and Falconer (2002) for B/h <50	D_x = [7.428 +1.775 *(B/h) ^0.62 *(V/Vs)^ 0.572]* (V/Vs)* h* Vs

Table 5. River and pollutant characteristics for longitudinal dispersion analysis

	Value
Depth (m)	0.85
Width (m)	18.0
Aspect ratio (-)	21.2
Mean velocity (ms^{-1})	0.6
Shear velocity (ms^{-1})	0.10
Mean/Shear velocity (-)	6.0
Slope (-)	0.00120
Dy (m^2s^{-1})	0.0127
Pollutant mass (kg)	5.0

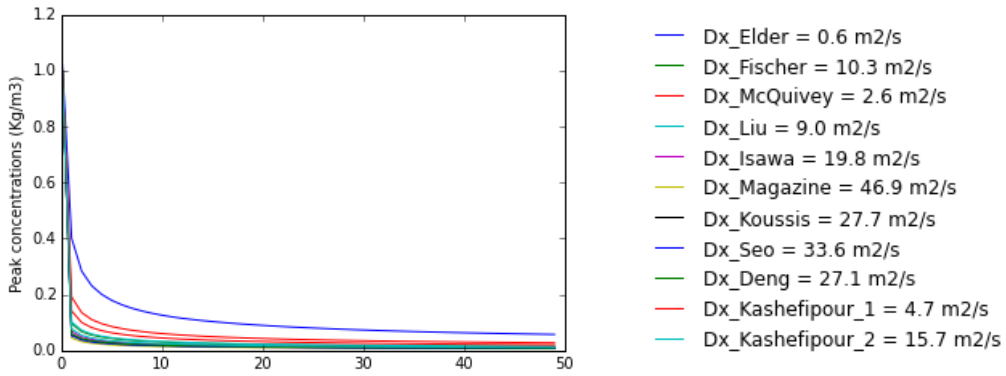

Figure 4. Peak concentrations for pollutant release using ADE 1D model and several longitudinal dispersion coefficients (Dx). Dispersion coefficients are in (m^2s^{-1})

The parameter uncertainty analysis will be expanded to include additional coefficients derived from other studies, and to include the uncertainty inherent in using a deterministic empirically derived Longitudinal Dispersion coefficient D_x (Camacho et al., in preparation). Also the comparison with field collected dispersion coefficients will be carried out and analysis of variables that affect the dispersion coefficient such as the Manning roughness coefficient and the varying discharge flows.

4 Conclusions

A flexible tool was developed in Python, consisting of a framework in which different mixing models and different model parameters can be included and the residuals compared.

The framework has demonstrated the capability to run various models and estimate the concentrations and the differences in model predictions. Therefore, the main tool for this analysis has been developed and will continue to grow to accommodate additional physical and biochemical models with their corresponding uncertainty analysis.

5 REFERENCES

Ani, E.-C., Wallis, S., Kraslawski, A., and Agachi, P. S.: Development, calibration and evaluation of two mathematical models for pollutant transport in a small river, Environmental Modelling & Software, 24, 1139-1152, http://dx.doi.org/10.1016/j.envsoft.2009.03.008, 2009.

Beck, M. B.: Water quality modeling: A review of the analysis of uncertainty, Water Resources Research, 23, 1393-1442, 1987.

Beer, T., and Young, P. C.: Longitudinal Dispersion in Natural Streams, Journal of Environmental Engineering, 109, 1049-1067, doi:10.1061/(ASCE)0733-9372(1983)109:5(1049), 1983.

Blumensaat, F., Seydel, J., Krebs, P., and Vanrolleghem, P. A.: Model structure sensitivity of river water quality models for urban drainage impact assessment, Proceedings - 7th International Congress on Environmental Modelling and Software: Bold Visions for Environmental Modeling, iEMSs 2014, 2014, 609-618.

Camacho Suarez, V., Shucksmith, J.D., Schellart, A.N.A. (In preparation). Quantifying the impact of uncertainty in the dispersion coefficient on water quality modelling in rivers.

Camacho, V., Norris, T., Shucksmith, J., Schellart, A., Analysis of Structural Uncertainty in the Analytical Solution of ADE in River Impacted By CSOs, 10th International Conference on Urban Drainage Modelling, Quebec, Canada, 203-207, 2015.

Choi, H. G., and Han, K. Y. P.: Development and applicability assessment of 1-D water quality model in nakdong river, KSCE Journal of Civil Engineering, 18, 2234-2243, 10.1007/s12205-014-0457-7, 2014.

Cox, B. A.: A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers, Science of The Total Environment, 314–316, 335-377, http://dx.doi.org/10.1016/S0048-9697(03)00063-9, 2003.

Deng, Z. Q., Singh, V. P., and Bengtsson, L.: Longitudinal dispersion coefficient in straight rivers, Journal of Hydraulic Engineering, 127, 919-927, 10.1061/(ASCE)0733-9429(2001)127:11(919), 2001.

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and Deletic, A.: Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling, Water Research, 46, 2545-2558, http://dx.doi.org/10.1016/j.watres.2012.02.009, 2012.

EC: European Commision. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community actions in the field of water policy, Official Journal of the European Communities, 327, 2000.

El Kadi Abderrezzak, K., Ata, R., and Zaoui, F.: One-dimensional numerical modelling of solute transport in streams: The role of longitudinal dispersion coefficient, Journal of Hydrology, 527, 978-989, 10.1016/j.jhydrol.2015.05.061, 2015.

Elder, J. W.: The dispersion of marked fluid in turbulent shear flow, Journal of Fluid Mechanics, 5, 544-560, doi:10.1017/S0022112059000374, 1959.

Fischer, H. B., List, E.J., Koh, R.C.Y., Imberger, J, Brooks, N.H.: Mixing in Inland and Coastal Waters, Academic Press, 1979.

Freni, G., and Mannina, G.: Bayesian approach for uncertainty quantification in water quality modelling: The influence of prior distribution, Journal of Hydrology, 392, 31-39, http://dx.doi.org/10.1016/j.jhydrol.2010.07.043, 2010.

Freni, G., Mannina, G., and Viviani, G.: Assessment of the integrated urban water quality model complexity through identifiability analysis, Water Research, 45, 37-50, http://dx.doi.org/10.1016/j.watres.2010.08.004, 2011.

González-Pinzõn, R., Haggerty, R., and Dentz, M.: Scaling and predicting solute transport processes in streams, Water Resources Research, 49, 4071-4088, 2013.

Iwasa, Y., and Aya, S.: Predicting longitudinal dispersion coefficient in open-channel flows, International Symposium on Environmental Hydraulics, Hong Kong, 505 – 510,1991.

Kannel, P.R., Kanel, S.R., Lee, S., Lee, Y.S., Gan, T.Y. A review of public domain water quality models for simulating dissolved oxygen in rivers and streams. Environmental Modelling and Assessment, 16, 2, 183-204, 2011.

Kashefipour, S. M., and Falconer, R. A.: Longitudinal dispersion coefficients in natural channels, Water Research, 36, 1596-1608, http://dx.doi.org/10.1016/S0043-1354(01)00351-7, 2002.

Koussis, A. D., and Rodríguez-Mirasol, J.: Hydraulic estimation of dispersion coefficient for streams, Journal of Hydraulic Engineering, 124, 317-320, 1998.

Lees, M. J., Camacho, L. A., and Chapra, S.: On the relationship of transient storage and aggregated dead zone models of longitudinal solute transport in streams, Water Resources Research, 36, 213-224, 10.1029/1999WR900265, 2000.

Lindenschmidt, K.-E., Fleischbein, K., and Baborowski, M.: Structural uncertainty in a river water quality modelling system, Ecological Modelling, 204, 289-300, http://dx.doi.org/10.1016/j.ecolmodel.2007.01.004, 2007.

Liu, H.: Predicting dispersion coefficient of streams, ASCE J Environ Eng Div, 103, 59-69, 1977.

Magazine, M. K., Pathak, S. K., and Pande, P. K.: Effect of bed and side roughness on dispersion in open channels, Journal of Hydraulic Engineering, 114, 766-782, 10.1061/(ASCE)0733-9429(1988)114:7(766), 1988.

Mannina, G., and Viviani, G.: A hydrodynamic water quality model for propagation of pollutants in rivers, Water Science and Technology, 62, 288-299, 2010.

McQuivey, R. S., and Keefer, T. N.: Simple method for predicting dispersion in streams, Journal of Environmental Engineering Division ASCE, 100, 997-1011, 1974.

Pelletier GJ, Chapra SC (2005) QUAL2Kw theory and documentation (version 5.1), a modeling framework for simulating river and stream water quality. Washington: Department of ecology

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., and van der Keur, P.: A framework for dealing with uncertainty due to model structure error, Advances in Water Resources, 29, 1586-1597, 10.1016/j.advwatres.2005.11.013, 2006.

Runkel, R. L., and Bencala, K. E.: Transport of reacting solutes in rivers and streams, in: Environmental Hydrology, edited by: Singh, V. P., Kluwer Academic Netherlands, 137-164, 1995.

Rutherford, J. C.: River mixing, Wiley, 1994.

Seo, I. W., and Cheong, T. S.: Predicting longitudinal dispersion coefficient in natural streams, Journal of Hydraulic Engineering, 124, 25-32, 1998.

Sharma, D. and Kansal, A. Assessment of river quality models: a review. Rev. Environ. Sci. Biotechnol., 12, 285-311, 2013. DOI 10.1007/s11157-012-9285-8

Taylor, G.: The Dispersion of Matter in Turbulent Flow through a Pipe, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 223, 446-468, 10.1098/rspa.1954.0130, 1954.

Van Griensven, A. and Meixner, T.: Methods to quantify and identify the sources of uncertainty for river basin water quality models, Water Sci. Technol., 53, 51–59, 2006

Wool TA, Ambrose RB, Martin JL, Comer EA (2001) Water quality analysis simulation program, WASP6. Part A: model documentation, U.S. Environmental Protection Agency, Center for exposure assessment modelling, Athens, GA

Wynne, B.: Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm, Global Environmental Change, 2, 111-127, http://dx.doi.org/10.1016/0959-3780(92)90017-2, 1992.

Young, P. C., and Wallis, S. G.: Aggregated Dead Zone (ADZ) model for dispersion in rivers, 1986, 421-433,

Zheng, Y., Han, F., Tian, Y., Wu, B., and Lin, Z.: Chapter 5 - Addressing the Uncertainty in Modeling Watershed Nonpoint Source Pollution, in: Developments in Environmental Modelling, edited by: Sven Erik Jørgensen, N.-B. C., and Fu-Liu, X., Elsevier, 113-159, 2014.

Zou, R., Lung, W.-S., and Wu, J.: Multiple-pattern parameter identification and uncertainty analysis approach for water quality modeling, Ecological Modelling, 220, 621-629, http://dx.doi.org/10.1016/j.ecolmodel.2008.11.021, 2009.

APPENDIX A. River water quality models and available software

River water quality modelling is mainly based on empirically calibrated mathematical relationships derived from theoretical and empirical understanding of the key processes of advection, dispersion and biochemical reactions.

Physical pollutant behaviour

In 1855, Adolf Eugen Fick made the analogy that salt diffuses in a water as heat diffuses in a metal road. As a result, Fick's law (Equation A1) stated that the net flux of tracer concentration was proportional to the concentration gradient:

Equation A1

$$J_x = -e_m \frac{\partial c}{\partial x}$$

Where J_x is the mass flux in (MT^{-1}), $-e_m$ is the molecular diffusion coefficient (L^2T^{-1}) and c is the tracer concentration in (ML^{-3}). Taylor (1954) showed that in pipe flow, there is a point downstream (after the pollutant has been released) where the velocity shear and diffusion reach equilibrium as discussed in the dispersion section (1.3.1). At this point, the variance of the concentration profile increases linearly, the skewness reduces and the concentrations tend towards a Gaussian distribution and hence can be described as a Fickian process. Figure A1 shows the advective and equilibrium zones of a concentration profile. The advective zone is the region closest to the source where the velocity distribution plays a crucial role in mixing of the solute. The equilibrium zone is the zone where the variance of the concentration profile becomes linear, the skewness reduces (Rutherford, 1994) and the pollutant cloud can be modelled as a Fickian process. Figure A1 can be expressed against either distance or time.

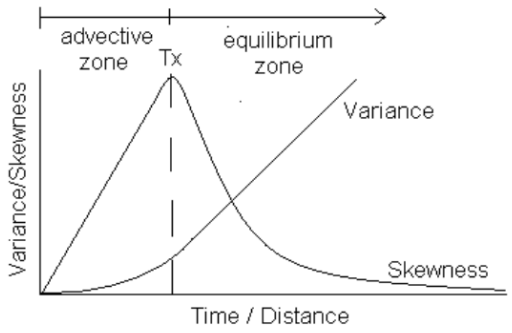

Figure A1. Variance and skewness vs. time of a change of concentration profile according to Fickian model predictions (Rutherford, 1994)

Advection-Dispersion Equation (ADE)

The advection-dispersion equation is derived from the conservation of mass in a unit volume where the accumulation of mass equals the mass input minus the mass output. Assuming that the flow and cross section are constant, the advective flux is characterized as the river velocity times the solute concentration.

The flux out of the volume is equal to the influx plus the change in flux in the control volume. Then, Fick's law (Equation A2) is used to describe the dispersive flux where the mixing coefficient D is proportional to the concentration gradient. Both, advective and dispersive fluxes are then placed in the conservation of mass equation leading to the 3D ADE equation:

Equation A2

$$\frac{\partial c}{\partial t} + u_x \frac{\partial c}{\partial x} + u_y \frac{\partial c}{\partial y} + u_z \frac{\partial c}{\partial z} = \frac{\partial}{\partial x}\left(D_x \frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_y \frac{\partial c}{\partial y}\right) + \frac{\partial}{\partial z}\left(D_z \frac{\partial c}{\partial z}\right) + R(c, P)$$

Where c is the solute concentration in (ML^{-3}), $x\ y\ z$ are the longitudinal, transverse and vertical directions in (L), $u_x\, u_y\, u_z$ are the velocities (LT^{-1}), $D_x\, D_y D_z$ are the mixing coefficients in (L^2T^{-1}) and R represents the biochemical transformations as a function of concentration and other parameters (P). The mixing coefficient represents mixing due to both turbulence and diffusion, but they are usually combined into a single coefficient.

The three dimensional form of the ADE is the more complex method for estimating concentrations. The ADE can be simplified assuming that the pollutant mixes instantaneously across the depth, or over the width of the channel. In these cases, the terms in the vertical or transverse dimensions can be eliminated. These simplifications are reasonable in shallow rivers or rivers of narrow width. (Runkel and Bencala, 1995).

Aggregated Dead Zone (ADZ) model

The ADZ model is an alternative to the ADE. Developed by Young and Wallis (1986), the ADZ considers the river an imperfectly mixed system (Figure A2) where advection takes place first and then dispersion occurs in a mixing zone (Lees et al., 2000). Equation A3 shows the ADZ model where the change in concentration (c) with respect to time (t) is proportional to the discharge Q in ($L^3\ T^{-1}$) over the volume (V) of water in the reach in (L^3) and the difference in upstream (c_U) and downstream (c_D) concentrations in (ML^{-3}).

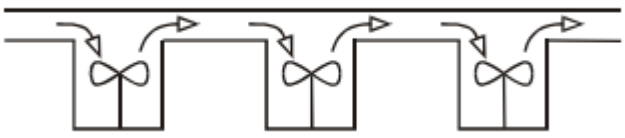

Figure A2. Representation of mixing zones in aggregated dead zone model (after Young and Wallis, 1986)

The ADZ model parameters include: (1) an average travel time which is the total time the solute is advected and dispersed in the river, (2) a dispersive fraction that indicates the ratio of mixed volume over the total volume, (3) a residence time that indicates the overall time of travel associated with dispersion, and (4) a time delay due to the dispersive effect. These parameters are presented under Equation A3:

Equation A3

$$\frac{\partial c(t)}{\partial t} = \frac{Q}{V}[c_U(t) - c_D(t)]$$

Software for Water Quality Modelling

A variety of software for water quality modelling is available such as Simulation Catchment (SIMCAT), QUAL2E, MIKE11, QUASAR, Flood Modeller Pro, SWAT, WASP, PEST, etc.

Most of the software available uses the one dimensional advection-dispersion equation assuming full instantaneous mixing across the river cross-section. A synthesis of software models reviewed is shown in Table A1.

Table A1: Summary of water quality software models

WATER QUALITY MODEL	DESCRIPTION	INPUTS	GOVERNING EQUATIONS	DIMENSIONS	REVIEW FROM	DEVELOPED BY
Simulation Catchment (SIMCAT)	Can model up to 600 reaches and BOD, COD, DO, ammonia concentrations	Flow and quality data upstream, discharges and abstractions (descriptions of statistical distributions - means and deviations)	Does not include ADE, it assumes full instantaneous mixing throughout each reach, solutes moving with same river velocity, uses first order decay to calculate concentration that will enter next reach. Flow: mass balance including tributary, effluent discharges, abstractions, and upstream conditions. Solute: for conservative pollutants - only advection, for non-conservative - first order decay (BOD and NH3), DO uses decay, temperature and reaeration	1D	Kannel et al (2011), B. A Cox (2003)	Anglian Water
Temporal overall model for catchments (TOMCAT)	Simulates current conditions and also changes to improve water quality. Can model storm events by diverting effluent discharges	Physical parameters, flow and quality data		1D		Thames Water
QUAL2EU	It can model waste loads on in-stream water quality and non-point source waste loads	River, global variables, and forcing functions	ADE	1D	Kannel et al (2011)	US EPA
QUASAR	To assess the environmental impact of pollutants on river water quality, real time control, dynamics between the flow and water quality, point and non-point sources. pH, e Coli, algae, BOD, DO, Conservative pollutants.	river map, boundary conditions, observed data	mass balance	1D		developed as part of the Bedford Ouse Study
MIKE-11	flow and quality in rivers	cross section, hydrodynamic advection dispersion and water quality parameters along the river, flow and WQ time series	St Venant equations (diffusive wave and kinematic wave)		B.A. Cox (2003)	

WATER QUALITY MODEL	DESCRIPTION	INPUTS	GOVERNING EQUATIONS	DIMENSIONS	REVIEW FROM	DEVELOPED BY
Flood modeller Pro (Former ISIS)	flow and quality	boundary conditions, time series of flow and water quality at upstream and tributaries. Also at downstream boundary, cross section geometry, abstractions, discharges	St Venant equations, finite difference approximation to 1D- ADE, transformation equations		B.A. Cox (2003)	
AQUATOX	Nutrients, organic chemicals, suspended and bedded sediments, DO fluctuations, toxicity from low oxygen and ammonia	Hydraulic, geometric, data, abiotic and biotic, physical, biota, remineralisation and eco-toxicology	Mass balance of nutrients	1D	Sharma, D. & Kansal, A. (2013)	
One dimensional Riverine Hydrodynamic and Water Quality Model (EPD-RIV1)	16 variables including water temperature, N, P, DO, CBOD, algae, Fe, MN, Coliform bacteria, macrophytes, varying point and non-point source pollution, cycling of nutrients, and fate, effect of toxic materials	Geometric data, initial conditions, model forcing data, hydraulic and control parameters and calibration data	1D advection - dispersion with decay and sinks	1D	Sharma, D. & Kansal, A. (2013)	
QUAL2Kw	can model pathogens as a function of temperature, light settling velocity, temp, pH, conductivity, inorganic suspended solids, DO, CBOD, N, Ammonia, P, biomass, algae, alkalinity,	flow and concentrations for headwater, discharges and withdrawals, reach segment lengths, elevations, hydraulic geometry, weather data	In general first order decay, mass balance	1D	Sharma, D. & Kansal, A. (2013); Kannel et al 2011	Pelletier & Chapra (2005)
WASP 7	Fate and transport of pollutants in surface waters including DO, N, P, C, Temp, salinity, bacteria, silica, sediments, heavy metals, mercury, inorganic loads	Model segmentation, boundary conditions, point and non-point source loads, kinetic parameters, flow, initial concentrations, numerical integration control options, weather data	ADE and kinetic transformation	1D, 2D, 3D (Wool et al. 2001)	Sharma, D. & Kansal, A. (2013)	
Water Quality for River - Reservoirs Systems	water quality conditions in rivers and reservoirs	geometry, meteorology, initial conditions, hydraulic and kinetic parameters	Conservation of heat and mass spatially and temporal, hydrologic routing, kinematic routing, steady flow, or full St Venat equations	1D	Sharma, D. & Kansal, A. (2013)	
Branched Lagrangian Transport Model (BLTM)		stream/river parameters, global variable forcing functions	1st order decays, 1D advective dispersion equation (Langrangian reference frame)	1D	Sharma, D. & Kansal, A. (2013)	

APPENDIX B. Sources of uncertainty in river quality modelling

Many different definitions of sources of uncertainty in models exist, however, this report will use the terminology and focus on types of uncertainty described in Dotto et al. (2012), as these are quantifiable sources of uncertainty in environmental models. The sources considered are input data, model parameters, calibration data, selection of objective functions and model structure uncertainties. Other sources of uncertainty such as 'ignorance' as described by Wynne (1992) will not be the included in this tool. Below, the definitions of structural and parameter uncertainties that will be used throughout this toolkit are described.

Structural uncertainty

Model structure uncertainty is usually referred to as the uncertainty associated with the deficiencies in matching the model to the real processes of interest (Refsgaard et al., 2006). Frequently, in order to simplify complex processes, key components are not considered or undergo scaling problems (Blumensaat et al., 2014). Model structure uncertainty is also associated with the mathematical expressions chosen to represent reality. Although widely accepted as a major source of uncertainty, structural uncertainty has been often neglected (Refsgaard et al., 2006;Freni and Mannina, 2010;Lindenschmidt et al., 2007).

Several sources of structural uncertainty are identified within the pollutant transport and mixing models. For instance, the advection dispersion model does not represent the skewness typically observed in tracer concentration profiles from field data. Explanations including the effect of trapping areas (frequently called dead zones) or the effect of shear velocities have been attributed to this skewness effect (e.g. data collected in the advective zone). Other assumptions in the analytical solutions of the ADE 1D such as an instantaneous fully mixed cross section of the pollutant mass downstream of the source, or the steady uniform flow also lead to inaccuracies in the predictions.

Another contribution to structural uncertainty is the reduction from the three-dimensional space to one-dimensional space. Most solute transport studies involving uncertainty analysis have been carried out in one dimension as observed in Mannina and Viviani (2010) and Choi and Han (2014) and Ani et al. (2009). However, the question of the level of uncertainty from dimensionality reduction still remains. It is important to note that the inclusion of a second or third dimension involve the inclusion of other parameters which will add to the parameter uncertainty (Ani et al., 2009).

Parameter uncertainty

Parameter uncertainty is associated with the process of selection of the parameters used in the model (Freni et al., 2011). The main parameters in pollutant transport and mixing are the dispersion coefficients in the ADE model and time delay and residence time in the ADZ model. These parameters are difficult and can be costly to quantify since they often require field data collection for calibration.

ADE Dispersion Coefficients

The concept of longitudinal dispersion was first introduced by Taylor (1954) in a circular pipe. Then, Elder (1959) derived a more theoretical dispersion equation for an infinitely wide channel. Elder's equation has been scrutinized for underestimating the natural dispersion in rivers (Fischer et al, 1979). Since then, more empirical equations for estimating the dispersion coefficients have been developed by McQuivey and Keefer (1974), Liu (1977), Magazine et al. (1988), Seo and Cheong (1998) and Kashefipour and Falconer (2002) amongst others. Through tracer experiments, dispersion coefficients have also been estimated. But although these experiments are more site specific and may give better parameter estimations, they are also associated with field collection errors and are specific to the time of year they were carried out, i.e. during high/low flow and summer/winter vegetation.

ADZ time delay and residence time

The time delay and residence time are obtained from the pollutant cloud travelling times. The time delay is the advective time it takes for the cloud to move only due to advection of the bulk flow while the residence time is a lumped parameter that describes the travel time associated with dispersion (González-Pinzõn et al., 2013). These parameters can also be obtained from experimental studies. Relationships between the ADZ parameters and hydraulic conditions have been studied for instance by Lees et al. (2000).

APPENDIX C. Software Tool Code Description

Overview

The code is developed in Python 2.7.1 (https://www.python.org) using the Spyder 2.3 platform (https://pypi.python.org/pypi/spyder/2.3.0), it is available from the dissemination pages of the QUICS project website, www.quics.eu. On a higher level, the code is separated into Module classes (Figure C1). Objects must be created. Therefore, there are two ways to create the main module (CMain) which is the core of the modelling framework. The first option is using the '1_framework' file, and the second option is using the '3_interface' file. These are the wrappers to the CMain class. The 1_framework allows the user to understand the code implementation. On the other hand, the 3_interface wrapper gives the user a menu that can be used to run the code directly. It also allows the user to load into the memory a list of analyses that have already been saved. This way, the user can re-plot calculated data without having to recalculate that data first. Both 1_framework and 3_interface are simply different entry points to the same underlying code and its functionality.

The main module is the blueprint of the framework. It defines the:

- CInputModule: responsible for all input actions.
- CCalculateConcentrationsModule: responsible for all calculations of concentrations.
- CPostPostprocessingModule: responsible for post-processing actions, such as those made on the already-defined input parameters, or concentration that may have been calculated.
- COutpuTModule: responsible for all output actions.

Many of these classes have methods to calculate values rather than storing them. This is designed to avoid duplication. For example, the cross-sectional area is defined as the result of width multiplied by depth. The area of a river is not stored as this duplicates river data and if we should change either the depth or width, then we either have to remember to update the area too, or risk getting our variables out of sync with each other.

However, the modules do not contain the data. The data is contained in the 'analyses' object which is a list of analyses. Each analysis holds all the information to define that analysis uniquely. Each permutation of input parameters constitutes a new analysis. So if the same river is analysed twice, only with different spatial parameters such as xStep = 0.1 and then xStep = 0.2, there will two separate analysis objects.

This way, the framework simply iterates through each river in the input file (which is defined in the CInputModule), and tells the main module to create a new analysis object for each river and append it to its list of analyses. It continues to run through all the steps in turn to calculate and process more data (e.g. concentrations) by

various means and plot the data in various forms. At the end, there is an option to save the analyses to file.

In addition there is 2_TestFramework, which also provides an additional entry point to run the same code, but not with the same purpose. 1_framework and 3_interface provided different ways that the user can interact with the code. 2_TestFrameworkis simply to test individual parts of the code to check that it has been implemented in such a way that we satisfy unit tests.

Modules

CMainModule: Creates the 'input', 'output', 'post-processing', and 'calculate concentrations' modules with their corresponding functions.

CInputModule: Requires the user to enter the following:

- Database of rivers containing the river names, depths (m), widths (m), mean velocities (ms^{-1}), and slopes.
- Longitudinal distance along the river (river_length) of interest (m).
- Mass of the pollutant released in the river (kg).
- Longitudinal and transverse dispersion coefficients (m^2s^{-1}). At the moment, the code uses (Elder, 1959) coefficients, but this can be changed by entering the dispersion coefficients in lines 73 and 74 of the input module.

CCalculateConcentrationsModule: Calculates the concentrations using the following functions:

- Advection Only
- ADE_Longitudinal_Dispersion (ADE 1D)
- ADE_Transverse_Mixing (ADE 2D)
- ADZ

Then the concentrations are saved into the analysis object

COutputModule:

- Plots the concentrations vs time or distance.
- Calculates peak concentrations.
- Calculates normalised concentrations and plots them against the Peclet number.
- Calculates peak normalised concentrations and plots them against the Peclet number.

- It can calculate if the concentration of a water quality parameter has been exceeded – This requires changing the default exceedance concentration value of 2.0 by the value of interest.

CPostPostprocessingModule:

- Calculates the residuals. At the moment, the ADE 2D is considered the most accurate model, therefore, the residuals are calculated using the ADE 2D as the reference value, but this can be changed in the PeakConcentrationsOverTimeResiduals function in the CModel class.

Analysis Object

The analysis object contains the river properties (CRiver), time and space domains (CDomain), the calculated concentrations for the Advection only, ADE 1D, ADE 2D and ADZ models (Cmodels) and the pollutant data (Cpollutant)

CRiver:

- The river class stores the river name, depth, width, flow, mean velocity, and slope.
- It calculates the shear velocity, mean to shear velocity ratio, width to depth ratio and cross-sectional area.
- It prints both, the river input and calculated properties to the console.

CDomain:

- Linearly discretises the x, y space and time domains.
- Prints the x, y and time steps to the console.

CPollutant:

- Stores the initial pollutant mass or the pollutant concentration.
- Prints the pollutant properties to the console.

CMixingParameters:

- Sets the default dispersion longitudinal and transverse dispersion coefficients.
- Prints the dispersion coefficients to the console.

Cmodels:

- Stores the calculated concentrations from the advection only, ADE 1D, ADE 2D and ADZ functions.
- Stores peak and normalised concentrations.

CSpatialParameters

- Stores the x, y time steps.

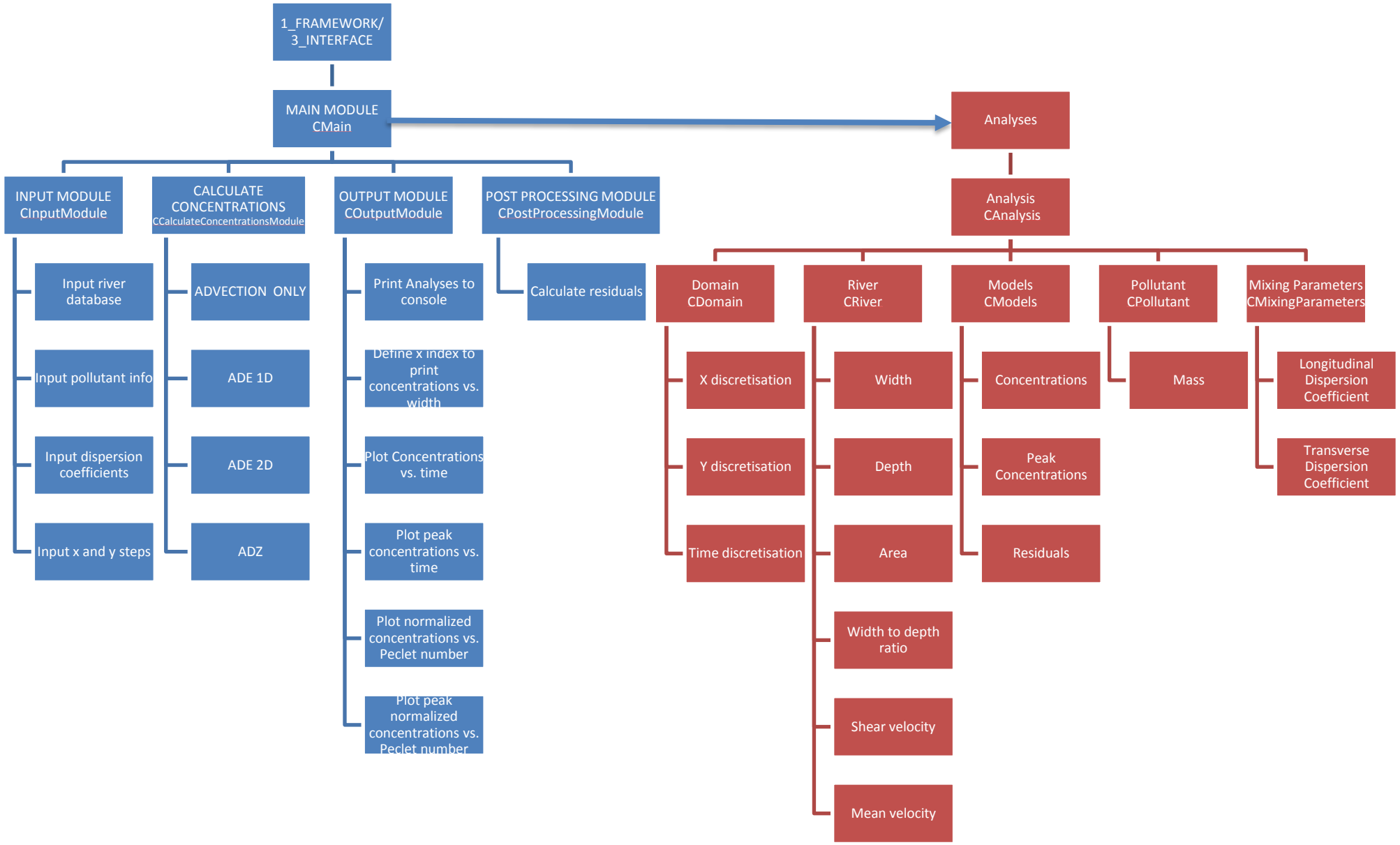

Figure C1. Pollutant Transport Modelling Framework Structure. The blue section is the blueprint for carrying out the calculations, while the r section stores the data inputs, calculations and outputs

How to run the code:

1) Define inputs in the CInputModule:
    a. Input river database in the form of River Name, Depth, Width, Shear Velocity and Slope.
    b. Define longitudinal distance or river length (river_length)
    c. Define pollutant mass.
    d. Define x and y steps.
    e. Define longitudinal (x) and transverse (y) dispersion coefficients.
2) Run 1_framework or 3_interface.
3) Save the results if needed for later analysis.
4) Run 3_interface to load previous saved analysis.