Recent insights on uncertainties present in integrated catchment water quality modelling


Abstract
This paper aims to stimulate discussion based on the experiences derived from the QUICS project (Quantifying Uncertainty in Integrated Catchment Studies). First it briefly discusses the current state of knowledge on uncertainties in sub-models of integrated catchment models and the existing frameworks for analysing uncertainty. Furthermore, it compares the relative approaches of both building and calibrating fully integrated models or linking separate sub-models. It also discusses the implications of model linkage on overall uncertainty and how to define an acceptable level of model complexity. This discussion includes, whether we should shift our attention from uncertainties due to linkage, when using linked models, to uncertainties in model structure by necessary simplification or by using more parameters. This discussion attempts to address the question as to whether there is an increase in uncertainty by linking these models or if a compensation effect could take place and that overall uncertainty in key water quality parameters actually decreases. Finally, challenges in the application of uncertainty analysis in integrated catchment water quality modelling, as encountered in this project, are discussed and recommendations for future research areas are highlighted.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Deterministic integrated catchment water quality models are often the method of choice to predict surface water quality and assist in making decisions on waste water treatment requirements, sewer system upgrading or rural land use strategies, e.g. Crabtree et al. (2009), Benedetti et al. (2013) or Bach et al. (2014). In theory, such integrated models include both urban and rural catchment spatial and temporal scales, although in practice many integrated catchment models (ICM) often still focus on either urban areas, e.g. Rauch et al. (2002) and Freni and Mannina (2010), or rural areas, e.g. Pang et al. (2018). Studies considering contributions from both rural and urban areas within a single river catchment remain rare (e.g. Honti et al., 2017).

Deterministic integrated catchment models (ICMs) can simulate the interlinked dynamics of the catchment system, enable the assessment of a range of alternative mitigating responses (infrastructural/regulatory) and then allow the identification of an optimal response (i.e. the lowest cost or highest value) given that the beneficial impact of any response could be remote from its implementation location. Significant asset investment and detailed water management strategies are based on the outputs of such modelling studies. However, there is increasing concern that these deterministic models are leading to incorrect problem diagnosis and inefficient investment and management strategies (e.g. Schellart et al., 2010; Voinov and Shugart, 2013) because the simulation results are being used with insufficient consideration to the uncertainty contained within them.

Software used by practitioners to simulate water quality has been created by incorporating individual models often developed by academics, but generally without consideration of levels of predictive uncertainty (Schellart et al., 2010). Consequently, the degree of uncertainty in water quality predictions is currently often not quantified, and therefore cannot be considered in the investment decision-making process. The same level of predictive uncertainty may influence the decision making process differently, depending on the desired objective. For some modelling studies the predicted probability distributions for outcomes of interest are significantly wider than the differences between the expected values of the outcomes across different policy alternatives (Reichert and Borsuk, 2005). Even when applying a robust decision-making approach (Lempert et al., 2006) deep uncertainties can have a strong influence leading to different policy optima.

Models of integrated water systems include all aspects of uncertainty inherited from the modelled subsystems as well as uncertainty resulting from the linkage of these subsystems. Three dimensions of uncertainty can be distinguished: source, type and nature of uncertainty (Refsgaard et al., 2007; van der Keur et al., 2008; Walker et al., 2003). Sources of uncertainties in hydrology and water quality modelling, can be classified into uncertainties caused by input data, parameter and model structure uncertainty (Guzman et al., 2015). However, the definitions of these classifications tend to overlap or be loosely defined (Deletic et al., 2012; Tscheikner-Gratl et al., 2017). For example, hydraulic roughness can be seen as either a model input derived from pipe material specifications, or from a look-up table based on different river types, or it can be a model parameter that needs to be calibrated (Bellos et al., 2018b). This diversity of uncertainty sources in these models makes it nontrivial to deal with them in a rigorous way and avoid confusion between them. This results in a need for a consistent ontology for uncertainty assessment, as already advocated by Montanari (2007) for hydrology, but also the need to better communicate uncertainty throughout the whole duration of the modelling process. This becomes even more important when different catchment areas and models are integrated. One solution could be the use of a more philosophical basis (Nearing et al., 2016) or a more practical approach as suggested in the QUICS Framework (Tscheikner-Gratl et al., 2017).

Predictive uncertainty can become particularly large when interlinked hydraulic and water quality models of different spatial and temporal scales are coupled without balancing model complexity and model objectives. For example, if different impacts on receiving water bodies are to be modelled, varying time and spatial scales must be considered (see Fig. 1). Linking a complex, data-intensive model of a sewer network with a coarse river quality model may result in large unforeseen uncertainties in the prediction of water quality parameters in sensitive locations, so the benefit of integrating the models or choosing a very detailed description for one of the sub-models is lost (e.g. Willems, 2006; Schellart et al., 2010). Additionally, the interpolation techniques adopted when several sub-models are linked, in both spatial and temporal scale, may also create significant uncertainties.

Unfortunately, end users of ICMs often have neither the knowledge, nor the will (since there is no reward or reinforcement) nor the practical tools to estimate the levels of uncertainty associated with sub-models of different spatial and temporal resolution. Currently, there are also no practical tools available to describe how such uncertainties are propagated between sub-models when considering water quality prediction at a catchment scale. This lack of tools was the motivation of the QUICS project, a European consortium on Quantifying Uncertainty in Integrated Catchment Studies (www.quics.eu), This paper is an output of the project and aims to synthesize the learning developed in the project as researchers created new tools to quantify uncertainty across whole the catchment. In the context of this project, all scales were studied: from the rural scale of a big catchment using a hydrological rainfall-runoff model to the small scale of the flow into a gully or manhole using Computational Fluid Dynamics (CFD) models. Furthermore, several studies were performed crossing over the scales, both in time and space.

This paper will first briefly discuss the state of the art on methods and frameworks for quantifying uncertainties in sub-
models of and for ICMs in general, and on contributions on this subject delivered by the QUICS Project. Second, it will discuss the challenges of propagating uncertainties between different sub-models, such as the balance between creating uncertainty due to sub-model linkage, uncertainty caused by model structure simplification, and the implementation of more calibration parameters and if this additional calibration is desirable. Finally challenges and constraints restricting the application of uncertainty analysis in ICMs are discussed and future research areas are highlighted.

2. Uncertainties in the sub-models of integrated modelling

Integrated urban water modelling means the joint modelling of two or more systems that affect surface water bodies (Muschalla et al., 2009). This is accomplished by computationally linking the various elements of the system (Rauch et al., 2002). For integrated catchment water quality modelling we can classify five types of sub-models (where two or more of these can form part of an integrated catchment study, see Fig. 2):

- Rainfall–runoff and pollutant wash-off models (RRM and PWM respectively): they are implemented on the rural and urban (sub-)catchment scales. In the case of urban catchments, they usually transform directly the rainfall to runoff and pollutant concentration at the outlet of the catchment. In the case of rural runoff models, they usually feed into routing models.
- Urban Drainage models (UD): they are implemented on the sewer system scale and simulate the transport of the flow and pollutants, as well as the occurring biochemical transformation processes, through the urban drainage system.
- Rural runoff routing models (RRM): they are implemented in the river reach or landscape scale of the rural catchment and they simulate the routing of runoff and the transport of the pollutant over the rural catchment surface (and sometimes including the shallow subsurface, although groundwater flows did not form part of the QUICS study).
- River models (R): they are implemented in the river reach scale and simulate the transport of flow and pollutants including the transformation processes within receiving surface water bodies.
- Wastewater Treatment plant (WWTP) models: they are implemented at a single location and they simulate the processes included in the waste water treatment plant.

Fig. 2 shows an exemplary (but not exhaustive) structure of such an integrated catchment study including all five sub-models (boxes) and the linkage between them (arrows). It includes reference values of the spatial (S) and temporal (T) scale variability of the hydrological processes, starting with the input of rainfall data into a rainfall runoff model from which the flows (Q) as well as the concentrations (C) of pollutants are propagated through the entire integrated model.

2.1. Rainfall–runoff and pollutant wash-off models

Water quality models are generally driven by runoff/hydraulic models, but for example a pollutant wash-off model may be directly derived from rainfall data and does not necessarily need a rainfall runoff model as intermediary step. For each sub-model a certain parameter set is necessary as well as for each linkage a certain amount of uncertainty must be estimated. This uncertainty is highly scale-dependent and so when models are linked it should be ensured that the linking variables are appropriately up- or downscaled. For instance, rainfall predictions and uncertainties refer to a certain temporal (minutes, hourly, daily, weekly) and spatial support (point, m², hectare, km², catchment scale) and the linked models should be able to process these scales.

Precipitation is a key driver of integrated catchment models. Rainfall can be measured by different instruments such as rain gauges, disdrometers, microwave links, weather radars and satellite, and all have different challenges with either spatial coverage and/or accuracy of the measurement (Rico-Ramirez et al., 2015). Cristiano et al. (2017) highlighted that the uncertainty in the spatial and temporal variability of precipitation is an important source of error when modelling the hydrological processes in urban areas. As described by Cristiano et al. (2017), interactions between rainfall variability, urban catchment heterogeneity, and hydrological response at multiple urban scales remain poorly understood. Weather radars can provide spatial rainfall measurements suitable for urban applications, although radar rainfall measurements are prone to error (Cecinati et al., 2017a). Merging radar rainfall and rain gauge measurements can bring the benefits of both instruments, such as the measurement accuracy of point observations from rain gauges and better representation of the spatial distribution of precipitation from radar (Cecinati et al., 2017b; Delrieu et al., 2014; Wadoux et al., 2017), or the integration of radar and point data measurements with different accuracies (Cecinati et al., 2018).

Amongst others, Ochoa-Rodriguez et al. (2015) showed that the effect of the spatial resolution of precipitation on flow simulation in urban drainage models, in the order of magnitude of several km² areas, decreases significantly with the increase of catchment drainage area. Moreno-Rodenas et al. (2017b) described the simulation of dissolved oxygen (DO) in an approximately 800 km² large, highly urbanized, lowland river catchment using different spatial and temporal aggregation of rainfall inputs and an integrated catchment simulator. The results of these simulations show a negligible sensitivity to temporal aggregation of rainfall inputs (between 10 and 60 min accumulation) and a relevant impact of the spatial scale with a link to the storm characteristics to combined sewer overflow (CSO) and DO concentration in the receiving water body. These results however can only be generalised to similar
systems with an equivalent mechanistic relationship between urban areas, wastewater treatment plant (WWTP) and river. A study by Schellart et al. (2012), in an 11 km² hilly urban catchment, showed considerable reduction in flow peaks in the sewer system, when the rainfall time scale was changed from 5 to 60 min frequency, which would be expected to also significantly underestimate CSO spills.

Until now, rainfall variability at sub-kilometre scale and the relation between rainfall spatial and temporal resolution at such small scales has received limited attention, e.g. Ochoa-Rodriguez et al. (2015). Muthusamy et al. (2017) studied spatial and temporal variability of rainfall at sub-kilometre spatial scales, to understand the rainfall uncertainty due to upscaling and also to select an optimal temporal averaging interval for rainfall estimation of hydrologic and hydrodynamic modelling, especially for small urban catchments. Muthusamy (2018) used this information to examine the propagation of rainfall (input) uncertainty in urban pollutant wash-off modelling. In this study, it was observed that the level of propagated uncertainty in the predicted wash-off load can be smaller, similar or higher to the level of the rainfall uncertainty depending on the rainfall intensity range and the “first-flush” effect. Rico-Ramirez et al. (2015) studied the application of radar rainfall to simulate flow in sewer networks in an 11 km² catchment and showed that radar rainfall related uncertainties could explain the uncertainties observed in the simulated flow volumes in sewer networks in 55% of the observed rainfall events. For the remaining rainfall events this was not the case, hence additional uncertainty sources related to the urban drainage runoff model and sewer flow model structure, model parameters and measured sewer flows are also contributing to uncertainty in simulated flow volumes.

2.2. Rural runoff routing models

Multi-source analyses of uncertainty sources in rural hydrological models in the past years have begun to compare and analyse the contribution of several uncertainty sources in model application. This interest has been particularly fostered by projecting climate change impact on hydrology. The question being how much of the uncertainty inherent in future climate projections contributes to uncertainty in hydrological model outputs. Starting with ensembles of only model input data, more and more work has been put into the consideration of further uncertainty sources, by using a variety of input data sets (General Circulation Models - GCM) and the emission scenarios (Special Report on Emissions Scenarios - SRES, Representative Concentration Pathways - RCPs) of the Intergovernmental Panel on Climate Change (IPCC), hydrological models (from lumped to fully distributed ones) and model parameter sets. Samaniego et al. (2017) showed that the contribution of these uncertainty sources on model outputs are consistent and led by GCM followed by hydrological model uncertainty. Nevertheless, large differences between catchments exist. In other studies, hydrological model selection are equally or even more uncertain (Bastola et al., 2011; Exbrayat et al., 2014) and variable in time (Hattermann et al., 2018). A general conclusion from these studies indicates that model output uncertainty is significantly affected by catchment characteristics and boundary conditions. One observable tendency of today’s river catchment scale hydrological models is that input data resolution has less effect on model’s output and efficiency than often assumed. Considering three semi- and fully-distributed models, Bormann et al. (2009) showed that resolution of spatial input data on topography, land use and soils did not significantly differ for spatial aggregation levels of 25–300 m. In agreement with this, a recent comparison of high-resolution local input data compared to global data products showed that hydrological target values from the widely used Soil and Water Assessment Tool (SWAT) were only marginally affected by input data quality (Camargos et al., 2018). This conclusion, however, is challenged if not only hydrological fluxes but also hydro-chemical ones are considered, where scale effects for SWAT have been detected from 100 m resolution onwards (Chaubey et al., 2005).
2.3. Urban drainage models

Sewer water quality modelling, in contrast with sewer flow modelling, involves several types of additional uncertainties in pollution load inputs and sewer quality processes (Willems, 2008). Urban drainage systems comprise many different infrastructure elements. Buried infrastructure can be classified as the minor system, such as the piped sewer network which can either be combined (waste water and storm water) or separated (storm water only). The surface drainage network, such as channels or roads, used as channels in storm events, can then be classified as the major system. Finally, there are micro drainage systems, known in different countries as e.g. low impact development (LID), Sustainable Drainage Systems (SuDS) or Best Management Practices (BMPs). Piped systems come with many adjacent structures such as gullies, manholes, storage tanks, overflow and outflow structures. Micro drainage systems comprise many additional structures such as green roofs, infiltration trenches, swales, detention ponds or wetlands, to consider during simulation phase. Hence, river catchments that include urban areas, and urban drainage system catchments commonly have a large hydrological heterogeneity. The hydrological and hydraulic processes occurring in urban drainage strongly influence transport and dispersion of solute and particulate materials within the catchments. A study on uncertainty in sediment build-up in sewer systems (Schellart et al., 2010) showed that uncertainty in hydraulic roughness, particle size and uncertainty in coefficient in the sediment transport equation all contribute to uncertainty in predicted sediment build-up. Whereas a study on uncertainty in simulation of CSO volume (Sriwastava et al., 2018) showed that the main contributor was uncertainty in runoff coefficient, with limited contribution from uncertainty in hydraulic roughness and weir crest level. This could be explained, because sediment transport is characterized by significant non-linearities and rainfall-runoff is not, although neither of these two studies took the uncertainty of rainfall into account. Also, uncertainties connected with water quality tend to be higher than the ones associated with quantity modelling (Mannina and Viviani, 2010).

Furthermore, the conclusions are dependent on the typology of system (e.g. differences between gravity driven and pressure driven systems) making generalization a difficult task. Elements in urban drainage systems such as inlets, gullies and manholes, where flow is turbulent and should be studied as a 3D phenomenon, usually are simulated using simplified 1D models using calibration parameters to account for the true 3D behaviour (Lopes et al., 2017; Rubinato et al., 2018). For understanding uncertainty introduced by simplifying a 3D structure into a 1D model, such elements are being studied in detail (Beg et al., 2018; Martins et al., 2018) and it is envisaged that this information can be utilised to provide levels of uncertainty related to the use of calibration parameters that account for 3D behaviour of the flow.

2.4. Wastewater treatment plant models

Urban drainage systems can negatively impact water quality of the receiving water either directly through CSOs or, through the effluent of WWTPs. Modelling of WWTPs has become a standard in both industry and academia for a range of objectives, such as WWTP design, operation, and control. The dynamic simulators currently available combine Activated Sludge (reactor) Models (ASM) (Henze et al., 1999) with clarifier and settling models. The main weakness of WWTP models used in the simulators is the lack of balance between the hydraulic modelling, very often a simple CSTR (completely stirred reactor) tanks in series approach, and the more complex biokinetic modelling parts (Gujer, 2011). The increasing complexity with high number of model parameters and the high level of lumpedness of the WWTPs processes, has resulted in highly over-parameterised models. Consequently, automatic model calibration routines as used for e.g. the earlier mentioned rainfall-runoff models or hydrodynamic sewer models may result in the numerical best fit but fail to properly describe the relevant processes. To minimise this, strategies and protocols have been developed (Hulsbeek et al., 2002) for a structured model calibration aiming at minimising the uncertainties in the model output.

2.5. River models

Rivers are complex non-linear systems encompassing a wide range of physical, chemical and biological components and processes. Surface water quality models assist in understanding and predicting such river processes and providing scientific background for management decisions when evaluating and implementing management measures (e.g. Asfaw et al., 2018). Most modern surface water quality models are composed of hydraulic (including transport and dispersion), thermodynamic and water quality process sub-models (Thomann and Mueller, 1987). In most applications, these three components are simulated sequentially. There may exist subsidiary interactions between all the processes occurring, however, in many cases these subsidiary interactions are not perfectly understood and for the most part considered to have only a minor impact on water quality. However, Moreno-Rodenas et al. (2017a) compared the effect of using two different descriptions for the river hydrological processes. When calibrating for hydraulic flow, both models affected the dynamics of DO in a different manner and since hydraulic depth affects the reaerating pattern, this has a very relevant impact if left unchecked.

The focus in uncertainty analysis of hydraulic river models is on input data (hydrological, geometrical) and friction coefficient (parametric uncertainty). Parameterisation of friction is based on assumption of fully turbulent flow over a rough rigid boundary. Hence uncertainties can be introduced in the simulation of flows where these assumptions are invalid, for example in vegetated flows (Shucksmith et al., 2011). Studies such as Brandimarte and Woldeyes (2013), Dimitriadis et al. (2016) and Bellos et al. (2017) try to investigate uncertainty due to input data (constant inflow and Manning coefficient, using several model structures. However, the heterogeneity and variation of surface waters mean that dominant water quality and transport processes and associated uncertainties are site/case specific, being dependent on both the hydraulic and environmental conditions as well as determinants and time/length scales of interest (Lindenschmidt et al., 2007).
2.6. Uncertainty analysis methods and frameworks

To facilitate the analysis of uncertainty sources and their propagation in hydrological modelling, a large number of methods have been proposed in the 1980s (reviewed by Beck (1987)) and 1990s (e.g. Beven and Binley, 1992). Uncertainty analysis (UA) is the process of quantifying uncertainty in model outputs that result from uncertainty in model inputs, model parameters and model structure. UA can be extended into sensitivity analysis (SA), which aims to rank the various sources of uncertainty and apportion uncertainty contributions to parameters and inputs. Reviews and examples of UA methods are given in Beven and Binley (1992), Refsgaard et al. (2005), Brown et al. (2005), Vrugt and Robinson (2007), Blasone et al. (2008), Matott et al. (2009), Kleidorfer (2010) and Guzman et al. (2015), for reviews and examples of SA we refer to Ferretti et al. (2016), Guzman et al. (2015), Hall et al. (2009), Matott et al. (2009), Pianosi et al. (2016), Saitelli et al. (2006) and Shin et al. (2013).

For statistical uncertainties the selection of the method depends on the problem statement, data availability and computational expense for running the model (Tscheikner-Gratl et al., 2017). In practice, one frequently opts for Monte Carlo based methods because these are very flexible and easy to apply. The main problem is the computational complexity, but since the Monte Carlo method is well suited for parallel computing it may also be feasible for modestly complex integrative catchment models. Scenario analyses can be applied for cases in which uncertainties cannot easily be characterised by probability distributions (Börjeson et al., 2006; Herman et al., 2015; Kwakkel and Jaxa-Rozen, 2016) or for exploratory modelling (Kwakkel, 2017; Urich and Rauch, 2014). The identification of the most appropriate method for the problem at hand is always a trade-off between the need for a strong theory-based description of uncertainty, simplicity and computational efficiency (Dotto et al., 2012).

Several frameworks were developed to provide a common uncertainty language and repository of methods. Sriwastava and Moreno-Rodenas (2017) give an extensive overview of these frameworks and their applications. Most notable are the frameworks of Refsgaard et al. (2007) and the Global Assessment of Modelling Uncertainties (GAMU) framework of Deletic et al. (2012). While these frameworks have provided an excellent structure to analyse and understand uncertainty, their application remains a challenge in practice:

- The current frameworks mainly focus on quantifying the total uncertainties in the output, without investigating the decomposition of uncertainty contributions into different sources, although research focussing on quantifying contributions from parameters, input and structural uncertainties in predictions has been done (Reichert and Mieletiner, 2009; Willems, 2012, 2008; Yang et al., 2018).
- In many applications the uncertainty analysis is still often considered as a standalone task and not an integral part of the modelling workflow directed to update and improve model conceptualisations and further data acquisition.
- Proposed methods are seldom applicable to full-scale catchment water quality models. Reasons are the increased computational burden, local interpretation of environmental legislation or accepted best-practice guides, favouring for example the use of specific types of deterministic models and design rainfall that is not spatially varied.

In spite of many methods being available, these methods are generally not utilised by practitioners, with few exceptions (Langeveld et al., 2013). Experience from the QUICS network indicates that it is mainly the lack of incentive from local regulators, and a culture of deterministic models that are ‘accepted’ by regulators, that prevents uptake of uncertainty analysis methods. There is furthermore a lack of practical demonstration case studies that show the benefits of uncertainty analysis. Those benefits can translate into e.g. lower investment costs or lower risk of failure of programmes of measures. Another reason may be that mature guidance for practitioners on methods and applications does not exist to a sufficient extent (Pappenberger and Beven, 2006). In this context a framework (Tscheikner-Gratl et al., 2017) and code of practice (Bellos et al., 2018a) were developed to address those challenges.

This lack of case studies extends also to available literature about uncertainty analysis in ICMs. Radwan et al. (2004) presented a variance decomposition scheme for the modelling of dissolved oxygen in a water quality model for a stream Belgium. However, they did not use a full integrated model, but used effluent data from WWTP and urban areas as input for a river model, showing that input rural and urban pollution loads were responsible for most of the DO uncertainty. Schellart et al. (2010) estimated DO-NH4 failure probabilities in an expert elicited forward propagation scheme for an impact based water quality model, integrating urban and WWTP dynamics. Due to the “Laffer-Curve” shape, they computed a forward uncertainty analysis scheme using the two most sensitive parameters (soil moisture depth and particle size). Freni and Mannina (2010, 2012) used a full ICM (WWTP, urban drainage and river) in a small catchment in Sicily. They showed that urban drainage is the most dominant source of uncertainty in their system.

3. Linkage or how much integration is too much

In integrated modelling, typically one wishes to simulate a range of systems and associated processes on a spectrum of time and space dimensions together. Including an increasing number of sub-systems and processes tends to dramatically increase the need for input data (on geometry, boundary conditions and process parameters). However, a distinction must be made between complexity of processes and complexity induced by linkage. Although linking models and complexity often go together this has not always to be the case. It seems possible to have a model of a single system which is overly complex, likewise you can have a linked model that is too simple for a specified task. Still, the thought that adding more and more detail into a model leads to better and more accurate results is paramount to this urge for more integration, but the question remains if linking together different models can always deliver enhanced modelling results and how long it takes before we have an “Integronster” (Voinov and Shugart, 2013) or a “random number generator” (Willems, 2006). The opposite trend to this drive for more integration in integrated hydrological and water quality models can be observed in the field of structural mechanics, where significant effort is made in models based on Finite Elements Method to ensure that the behaviour of material relationships at interfaces is stable and smooth and moving the focus away from ever more detailed approaches.

3.1. The sensible level of detail

In a sense, a similarity between the level of detail strived for in integrated modelling and optimal tax theory is observed, which is often described by the “Laffer-Curve”. “When the tax rate is 0%, no tax revenue is generated, while when the tax rate is 100% no tax revenue is generated as well”. This latter observation is a strong simplification of the very complicated and long lasting discussions on optimal tax theory (e.g. Mirrlees, 1971). There should be an
optimum (at least seen from the perspective of the tax collector) between these two extremes, but what this exactly is, is hard, if not impossible to determine and depends largely on subjective preferences and/or political viewpoints. A similar parallel can be observed, in terms of indeterminism, in the application of one of the most successful theories in modern physics: quantum mechanics. Using quantum mechanical theory, the behaviour of elementary particles, atoms and to a certain extent molecules, can be described in detail and with an unprecedented accuracy. However, modelling the behaviour of something that can be found to impact on water quality in a catchment, for example a cow, using the same theory, seems impossible due to the prohibitive calculation efforts needed for such an enterprise.

Transferring these concepts to integrated catchment modelling, implies that the usability of the results is zero when no model is applied, while when taking everything imaginable into account, the usability is zero as well due to a possible explosion of propagated uncertainties in the end and/or the calculation effort needed. Now the question arises “How to determine a sensible level of detail of an integrated model?” This refers to level of detail of process descriptions, the information needed on initial and boundary conditions and on the geometry and structure of a given problem. Posing the question is easier than formulating a generic answer, as it is in optimal tax theory, nevertheless some elements of an answer will be addressed in the following.

- Clearly define the type of results sought for in terms of parameters, time and space scales.
- Identify the sub models needed and their data need.
- Identity the set of unknown data in the collection of sub-models.
- Evaluate whether enough measuring data of the right quality are available for calibration of the sub models.
- Consider how the interfaces of the sub models are described and if suitable interpolation procedures are in place to transfer information from one sub model to the other.
- Identify which component of the integrated model is responsible for the largest contribution to the uncertainty and re-evaluate the results of the integrated model on the usability for the original goal set when this element is left out.

In this manner the level of detail in terms of processes, interpolation procedures, geometrical data and model calibration is tuned on the usability of the results obtained.

3.2. Does linking of sub-models result in an explosion of uncertainty?

Estimating the global uncertainty of ICMs is still limited by appropriate methods to estimate the various uncertainty sources and the propagation of uncertainty. Nevertheless, simply calibrating and investigating the uncertainty of sub-modules and then only further considering the best sub-model parameterization in the ICM, is insufficient as well. Multi-criteria assessment of ICMs, selecting criteria depending on the modelling objective, can at least provide valuable insights to the behaviour of complex, coupled models. Houksa et al. (2017) investigated the performances of coupled hydrological-biogeochemical models and evaluated parameter sets that simulated well single criteria and those parameter sets that performed well for all target criteria. In their Monte Carlo based study, they needed to reduce their acceptable parameter space by 99.9%, discarding the majority of model setups that performed well for single criteria.

Another topic concerning the linkage is the directional flow of information. The flow of information is a relevant factor when designing the architecture of integrated catchment modelling studies. Models (and software) can be directly linked in an output-input basis only if there is an upstream to downstream unidirectional flow of information (i.e. no feedback). This is insufficient when control systems are used which propagate information from downstream state-variables to actuate on upstream ones. This is an existing practice in some water systems (e.g. linking WWTP states with the control of operations in the urban drainage system). It is also reasonable that in a foreseeable future, extensive sensor networks (e.g. Internet of Things) will play an increasing role in water management (Chen and Han, 2018). This will allow for assimilating an increasing amount of data in the system (Wani et al., 2017) and possibly to control the operation of urban systems accounting for the status of the receiving water system and treatment capacity. In such cases, seamless model integration is necessary to account for the multi-directional information flow. Current numerical solver schemes are highly tailored for the individual sub-model characteristics. Nevertheless, commercial software is progressively adapting to the environment of integrated catchment modelling, for instance linking 1D and 2D hydrodynamic models for flood prediction (Leandro et al., 2009) or linking a simplified ordinary differential equation (ODE) based integrated system to represent WWTP, urban drainage and river dynamics (Achleitner et al., 2007; Solbi, 2007). Further development of robust multi-scale solvers and software, which allow for the integration of simplified and physically based processes, is required.

Given that an integrated model has a larger number of uncertainty sources than each of its sub-models, it is tempting to think that its output uncertainty will also be larger. This impression is reinforced by the belief that ‘uncertainty propagation’ is synonymous to ‘uncertainty amplification’, suggesting that in a chain of models the uncertainty can only grow. However, this is not necessarily true. In contrast, there are several cases in which the output uncertainty will decrease when models are coupled because of a “compensation effect” (an analogy can be found e.g. for the rainfall scaling effect (Ciach and Krajewski, 2006)). We illustrate this concept with some simplified examples.

Consider a sewer system node where n pipes join and the sewage fluxes merge and flow into a single, larger pipe. Let the uncertainty of the chemical oxygen demand (COD) concentration of the effluent be equal for all pipes and quantified by a standard deviation $\sigma$ (mg/l). Then, the standard deviation of the COD concentration of the larger pipe will be some value between $\frac{n}{n}$ and $\sigma$, thus either smaller or equal to that of the individual pipes. In fact, it will only be equal to $\sigma$ if the COD uncertainties for all pipes are perfectly correlated, which is not realistic. Thus, uncertainty decreases.

A similar effect occurs when models are coupled in catchment modelling. Consider a case where farmers apply a pesticide to their land to protect their crops. Part of the pesticide will reach the local ditches through sub-surface and surface flow. This can be modelled using a soil hydrological and chemical model. Next, the pesticide is transported to local streams and channels as modelled by a surface hydrological model. Finally, the pesticide reaches the river and sea as predicted by a hydraulic model. While the uncertainty about the pesticide concentration in the local ditch may be extremely high, it will be small in the river and sea. Again, averaging-out effects cause uncertainty to decrease. In addition, subsystems and consequently sub-models may act as low pass filters in terms of event frequencies, as e.g. CSOs only start spilling after the entire sewer system volume has been filled. In the Netherlands, with on average 8 mm in sewer storage, this results in a CSO frequency of only 5–6 spills per year. This means that the uncertainty in the runoff routing due to uncertain initial conditions of the sewer catchment, which is relatively high for smaller storms, does not strongly affect the quality of river DO simulations.
These examples show that uncertainty often decreases when sub-models are integrated. Of course, there are also cases where the opposite occurs. A simple example is when a water quantity and a water quality model are coupled to calculate the pollution load (kg/s) of a stream or sewer pipe. The load is the product of flux (m$^2$/s) and concentration (kg/m$^3$), and if both have a relative error of 10%, then the relative error of the load will increase to a value between 14% and 20%, depending on the degree of correlation between the two uncertainty sources and assuming that the correlation is non-negative (if it is negative, which is not unlikely, then the relative error will be smaller than 14%). In chaotic systems, small deviations in variables can have huge consequences and hence in such systems it may occur that model coupling leads to an ‘explosion’ of uncertainty. In the case of integrated catchment water quality modelling, this might happen when an intervention is based on an uncertain system variable. For instance, if authorities impose restrictive measures on industry and farmers based on whether a water quality index is above or below a threshold, then a small uncertainty in the water quality index may have dramatic consequences if the index is close to its threshold value.

In summary, coupling models does not automatically lead to an increase of uncertainty. It very much depends on the correlation between the processes and the scale of spatial and temporal averaging. To be certain whether uncertainties amplify or cancel out, it is imperative that a sound and case-specific uncertainty propagation analysis is conducted.

4. Challenges and bottlenecks in application of uncertainty analysis in integrated water quality models

Despite considerable uncertainty, integrated models are important for effective decision support in major investment decisions for water utilities:

- Model outputs can be justifi ed to a regulator in a transparent, comprehensible and repeatable way, especially when industrial modelling practice guidance is used in the creation of (sub-) models.
- Relative comparison between solutions is still useful to rank alternatives.
- They are often cheaper than performing extensive and long-term measurements.
- There is a capability to simulate extreme events, although calibration based on regular events may decrease the validity of the results.

Lee Jr. (1973) defined seven sins for large-scale models for urban planning and we would like to revisit them for integrated water quality models in the context of uncertainty, to highlight remaining challenges in application. Although defined 45 years ago some of these points may still ring true in the ears of integrated modellers.

Hyper-comprehensiveness (1), defined as an overly complex model structure, and complicatedness (2) in terms of interactions between the model components are connected to the question of linking different sub-models. These two points lead to the rule that not the most complex model should be selected, but, following Ockham’s razor, the least complex that answers the asked question reliably, in a comprehensible and verifiable way (Rauch et al., 2002). Also, grossness (3), which means the level of detail for model results used for predictions may be too coarse for effective decision making, relates to this aspect. The objective of the modelling endeavour should be clarified and with it the scale and level of detail of the necessary results. This adds to the challenge of linking models that represent processes that act on different space and time scales. Blöschl and Sivapalan (1995) distinguish between a process, observation and modelling scale. Under the best scenario, those scales should match, but in integrated catchment studies this is generally not the case, as for example pollutant wash-off, or a combined sewer overflow happens at small spatial and temporal scale, but the effects can be found in the receiving water at larger spatial and temporal scales. Transformations based on downscaling and upscaling techniques are generally necessary to obtain the required match between scales (Cristiano et al., 2017). The unwrathfulness (4) for data is connected to this point in the way that an adequate amount of data is essential to define the model setup and to identify the model parameters (Muschalla et al., 2009). Different level of model integration also demands different amount and quality of data for modelling and decision-making (Eggimann et al., 2017). Furthermore, the lack of data in environmental studies and a common non-sharing policy (Camargos et al., 2018) and the need to satisfy local regulators (Sriwastava et al., 2018) remain major problems for performing a complete uncertainty analysis.

Mechanicalness (5), defined as errors caused by the computational representation of the model, could in times of increasing computational resources be reinterpreted as limitations in terms of computational power availability, accessibility to source code and ability to adapt model settings and computational cost. A typical uncertainty analysis study requires a significant amount of simulations which cannot be performed computationally. Therefore, the modeller should find ways to automatize this process, though most of the commercial software does not provide this capability. Even if such a capability exists, sufficient documentation to guide the end user is sparse. It is also common practice that several parameters (e.g. the time step, the space step, the tolerance in iteration loops) in commercial software are considered as default values and adaptation and changes to these parameters might result in simulation malfunctions. There are several types of models which are characterised by significant computational cost for each run (in the magnitude of hours/days), especially if they are used for real world case studies. Therefore, a typical uncertainty analysis is often not feasible. One way to cope with that, is the use of adaptive or informed samplers (Goodman and Weare, 2010; Hoffman and Gelman, 2014; Laloy and Vrugt, 2012) and surrogate models or emulators (e.g. for sewer hydraulics (Carbajal et al., 2017), for hydrological models (Yang et al., 2018), or for rainfall dynamics in 2D physically-based flow models (Moreno-Rodenas et al., 2018).

The problem of wrongheadedness (6) can be explained by the gap between the behaviour of the model which it was built for and what it is used for. Models often represent rather the data available than focussing on the objectives. This can lead to focusing on aspects that might not matter and forgetting about those that do and that we simply do not have data about. This connects to the observation that perception of uncertainty on different inputs and parameters for existing model does not scale when used for several objectives, which may change the temporal and spatial extent of the project (see Fig. 3). The calculated level of uncertainty, although the numbers do not change, will be perceived differently depending on the nature of the objective, being either small scale measures (e.g. the design of a CSO using only design rainfall events) or strategic decisions (e.g. water quality considerations of a whole river basin). For integrated models that are used for decision making based on acute effects, such as ammonia toxicity or oxygen depletion, the focus is on bigger events, as the sewer and WWTP act as high pass filters for smaller events and the impacts only occur when the assimilative capacity of the river is exceeded. For those events, the relative uncertainty due to e.g. initial losses decreases rapidly. Therefore, for different objectives the same calculated level of uncertainty will result in a different objective specific level of uncertainty. This difference between the calculated and objective specific uncertainty is called objective shift. For example, the usage
of literature values for initial losses on the surface (e.g. 4 mm) are much more sensitive for the design of a CSO if a 10 mm rainfall event is used than for a 50 mm one, the same holds true for the calibration of models on these events.

Similar to the selection of different input data (Tscheikner-Gratl et al., 2016; Vonach et al., 2018) the calibration on different objectives (e.g. different water quality parameters or water quantity) influences the model behaviour. The objectives of the modelling effort very much determines the characterisation of a model (Bennett et al., 2013). It is therefore advisable, although difficult in practice, to apply the model in consequence only for the objective it is calibrated to. Due to the fact, that not all sub-models can be calibrated and every sub-model will always be influenced by the calibration of the ‘upstream’ models used as input, distortion is unavoidable in practice with linked models and could only be avoided by building an integrated model from scratch (Tscheikner-Gratl et al., 2017). Furthermore, the question arises if the importance of still statistical graspable uncertainties dwindles in comparison to deep uncertainties, when the objective scale changes. Because if you consider a whole river Basin over a period of 100 years, then things such as climate change, population in- or decrease and land use changes could be the cause of huge uncertainties while other uncertainties that would dominate at smaller scales could diminish in importance. This objective shift is one reason why the application of one model for several objectives without adapting and scaling the input data and the model could lead to poor predictions with high levels of uncertainty. Also, uncertainty quantification is seldom scalable and therefore there exists no one-fits-all solution.

If we translate all the limitations mentioned to a common metric, we end up at the sin of *expensiveness* (7). In practice addressing most of these issues requires monetary investments (e.g. for software, sensors, experienced modellers) as well as time resources. These costs must be covered by the modeller and in consequence need to be billable to the decision maker (in case that decision maker and modeller are different entities). And although reducing uncertainty is valuable (Reckhow, 1994), it is difficult to communicate this. This influences also the point of data availability due to costs for data acquisition (often data transmission) and costs for data quality control (which is often neglected). Cost minimization schemes may also lead to the application of existing models for different objectives without the necessary and costly adaptions. This encompasses also a lack of performance assessment, not only if the measures had the expected effect but also to collect evidence about the degree of uncertainty of model predictions. There are, however, opportunities to use information about uncertainty in models to better understand trade-offs between risks of failing environmental standards and investment costs. Communicating uncertainty in models as probability of failure of environmental standards, as well as impact of uncertainty on investment costs (Sriwastava et al., 2018) tends to gain interest among practitioners.

5. Conclusion

In the course of the QUICS project we found several key points that we want to highlight here:

1. Uncertainty analysis of integrated catchment water quality modelling should be a continuous process performed in parallel to the modelling exercise rather than being an activity that is carried out after the main modelling activities. This inclusion starts with the definition of the modelling...
goals, and should be present through the model build, simulation and interpretation phases of an integrated catchment modelling study.

(2) Linking together different models is a non-trivial task and requires careful and detailed handling. However, coupling models does not automatically lead to an increase of uncertainty. Results and ongoing studies from and based on the QUICS project [highlighted with an asterisk in the references] indicate that uncertainty in water quantity and quality predictions does not necessarily increase across an ICM. The important issue is not the scale of a model, but the integration of different models developed at different spatial and temporal scales. It is often at these interfaces that the modelling approaches radically changes. This issue is particularly important when modelling water quality processes; often the scale and temporal resolution of the hydraulic model that drives the water quality model has been developed at a different scale from the water quality process model.

(3) Further research in uncertainty decomposition and model acceleration, especially the generalisation of input-based emulation into water quality dynamics, is still required, to allow for the implementation of uncertainty analysis frameworks in practice. Uncertainty analysis can lead to potentially less expensive solutions with a better understanding of risk water quality compliance failure. Simplifying approaches such as the use of emulators and new upsampling/downscaling techniques can be applied successfully in ICM studies, without a significant loss of accuracy in determining the magnitude of model uncertainties. There is the potential for simplified and computationally efficient approaches for water quality uncertainty in ICM studies to be developed and then used by end users that are faced with decisions on investment.

(4) Understanding the outcomes and the inherent uncertainties of ICMs poses a challenge in practical application. Each submodel reflects the knowledge of some specialism (e.g. hydraulic, water-quality, WWTP performance) and persons that master all these different subjects entirely are very rare. This implies, that in practice, one needs a team of experts to understand, apply and communicate the results correctly for the purpose the model was assembled for. Consequently, more research should also be carried out on how to involve local environmental regulators and organisations in order that they can become aware and then appreciate the role of uncertainty analysis in determining what investment actions are required to meet regulatory requirements. In the longer term, regulators need to be explicit as to how they would incorporate uncertainty analysis into their decision making processes. Finally, although modelling activity is highly tailored for each specific scenario, further discussion on details such as globally acceptable uncertainty degrees or sub-model linkage strategies is needed.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was conducted as part of the Marie Curie ITN – Quantifying Uncertainty in Integrated Catchment Studies (QUICS) project. This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 607000. The authors would like to thank all QUICS fellows, partners and associate partners, and especially Jörg Rieckermann from EAWAG and Lisa Scholten from TU Delft for their input and discussion points.

References


