

## ***Mind the Gap – Overcoming Data Issues in Process Modelling***

Wastewater utilities are increasingly investing in process models to support operational or capital planning decisions. However, growth of this application is slow. The WEF Modelling for Plant Operations Task Group performed a survey of 22 U.S. utilities and identified time availability for modeling tasks and data management issues as two major obstacles hindering model application at the utility level. An improved pipeline from existing SCADA and laboratory information management systems (LIMS) to validated data sets for modelling purposes is necessary. Data quality evaluation and management is considered an essential element of such a pipeline. Indeed, modelers are frequently faced with the need to use available plant data for modelling and simulation, in part to reduce the expense of resources and time associated with dedicated measurement campaigns, tracer tests, assessment of daily dynamics, biodegradability tests, and batch tests for kinetic model identification. As commonly available data is not routinely checked and controlled in view of such purposes, a modeler typically spends a majority of her time on data screening and collecting additional measurements with little time left for model development and evaluation. This workshop strives to answer the question: How can we reduce the benefit-to-time-cost ratio of process modeling at the utility level?

### **Expected discussions and results**

How will the workshop go about answering the posed question?

- The workshop is split into two sessions, a morning session and an afternoon session. In the morning, we particularly focus on identifying existing challenges in data treatment for modelling. In the afternoon, we focus on available solution and seek to find the most urgent challenges which have not been solved to satisfaction yet.

Who will be presenting in the workshop, and how will they contribute to the discussions?

- We have presentations from experts in the field which have taken on academic, consultancy, software developer, and utility management roles. This representation of all business roles in the water sector will guarantee a balanced discussion. Contributions by all attendees is ensured by breaking up the attendees into smaller groups in break-out sessions. The smaller size of this group allows everyone to speak so that all voices are heard.

How will you summarize results for larger WWTmod group?

- The first action includes a summary presentation during the conference. In addition, we hope to gather sufficient interest to develop and disseminate the generated ideas further, either by means of white paper or by means of a new task group.

What are plans for workshop subject after WWTmod? i.e. white paper, publications, other?

- This remains to be decided at the end of the workshop. A white paper or publication is considered a likely scenario. In addition, a task group is also considered.

## Workshop set-up

The workshop agenda is distributed over two sessions. The morning session focuses on ensuring consensus on the most important obstacles and problem definition. During the afternoon session, brainstorming is organized to deliver solutions. In each of the sessions, attendees are split into smaller groups to facilitate discussion. These groups then report back for a general discussion. The main outcomes for this workshop are as follows:

1. Increased awareness of the necessity to plan and design for high data quality sets and the available tools to achieve this
2. Consensus building among academic and practice-oriented efforts in data-hungry model calibration and simulation efforts
3. A list of identified bottlenecks, issues, problems in the collection and preparation of data for modelling
4. A list of currently available solutions for data management
5. Identification of the largest gaps between available tools and identified challenges
6. An opinion paper or task group, pending further interest

## Chair/Co-chair

Chairs	<i>Adrienne Menniti</i> (Clean Water Services, USA) <i>Kris Villez</i> (Eawag, Switzerland)
Co-chairs	<i>John Copp</i> (Primodal, Canada) <i>Oliver Schraa</i> (inCTRL Solutions Inc., Canada)

## Speakers / Moderators

*Kurt Carson* (Denver Metro, USA)  
*John Copp* (Primodal, Canada)  
*Adrienne Menniti* (Clean Water Services, USA)  
*Oliver Schraa* (inCTRL Solutions Inc., Canada)  
*Kris Villez* (Eawag, Switzerland)  
*Tom Wambecq* (Aquafin, Aartselaar, Belgium)  
*Stefan Weijers* (Waterschap De Dommel, Boxtel, The Netherlands)

## Target Participants

Do you deal often with incomplete, unchecked, or otherwise low-quality data?

Then this is the workshop for you!

Due to insufficiently checked data, a modeler typically spends a majority of her time on data screening and collecting additional measurements with little time left for model development and evaluation. This workshop strives to answer the question: *How can we reduce the benefit-to-time-cost ratio of process modeling?*

## Programme

Time	Topic	Presenter/Moderator
09:45 - 10:00	<b>Setting the scene:</b> Introductory presentation by the group leader (Motivation, scope, and objectives; Present workshop structure, participants, etc.)	Adrienne Menniti
10:00 - 10:45	<b>Break-out session 1:</b> Attendees are split into groups of according to their business role: (i) utility companies (ii) consultancy and software developers (iii) academic institutes As a primer, the following questions are raised: (i) What are your most common challenges concerning data? (ii) How you do handle these challenges? (iii) What's missing to make life easier?	Adrienne Menniti
10:45 - 11:15	Coffee break	
11:15 - 11:30	<b>Prepare summary presentation of break-out session 1</b>	Kris Villez
11:30 - 11:45	<b>Present summary presentations of break-out session 1</b>	Kris Villez
11:45 - 12:45	<b>General discussion</b> Leading question: Which problems <ul style="list-style-type: none"> <li>• ... are common for all business roles?</li> <li>• ... are unique to single business roles?</li> </ul>	Kris Villez
12:45 - 13:45	Lunch break	
13:45 - 14:15	<b>Overview of current methods and data requirements, to deal with:</b> <ul style="list-style-type: none"> <li>• Historical data</li> <li>• Real-time data</li> <li>• Preparing for high quality data</li> </ul>	John Copp Oliver Schraa Kris Villez
14:15 - 15:00	<b>Break-out session 2:</b> Attendees are split into groups with mixed business roles and come up with favoured elements for the data-to-model pipeline. Utility representatives spike the discussion with a presentation on their applied method. Priming question: What kind of data treatment tool <ul style="list-style-type: none"> <li>• ... has worked for you?</li> <li>• ... did not work for you?</li> </ul>	John Copp (moderator) Kurt Carson Tom Wambecq Stefan Weijers
15:00 - 15:30	Coffee break	
15:30 - 15:45	<b>Prepare summary presentation of break-out session 1</b>	Oliver Schraa
15:45 - 16:00	<b>Present summary presentations of break-out session 1</b>	Oliver Schraa
16:00 - 16:45	<b>General discussion,</b> Leading questions: <ul style="list-style-type: none"> <li>• Which problems are common for all business roles?</li> <li>• Which problems are unique to a single business role?</li> </ul>	Oliver Schraa
16:45 - 17:15	Wrap-up, composing summary, report and presentation	Adrienne Menniti Kris Villez