iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Research Needs in Water Modelling

Gustaf Olsson
Professor in Industrial Automation
Lund University, Sweden
Gustaf.olsson@iea.lth.se

Today's process engineering systems are of ever-increasing complexity. This is also true for water and wastewater systems. The processes need to be better understood in order to achieve an improved final product quality while ensuring a safe, economic and efficient plant operation. Advanced water and wastewater engineering systems are, inherently multidisciplinary. The processes are composed of interacting subsystems of parts from different engineering disciplines, requiring an integration of chemical, bio, mechanical, electrical, control, and process engineering. This also means that design of proper control laws for large systems almost always requires a well developed process model.

Many models are developed to summarize and encapsulate our current understanding of the processes. Well-known examples of this are the Activated Sludge Models (ASM) no 1, 2 and 3 and the Anaerobic Digestion Model (ADM), published by task groups in IWA. These models describe the kinetics and reactions taking place in the biological reactors in wastewater and sludge treatment systems. However, they do not describe the interaction between the plant unit processes, the hydraulics, the solids/liquid separation and many other phenomena in a complex plant.

Many models, including the ones mentioned, are used for design. Increasingly such models describe the dynamical response to load changes. It is obvious that these models cannot be verified until afterwards, since the plant does not yet exist. A reliable system operation also requires that faults and failures are detected automatically at an early stage. This requires other type of models, to support the detection process. Such models are part of the operator guidance and support system and should give a basis for manual decisions, where operating changes are performed by a human operator.

In most cases models are represented numerically as differential or difference equations. Increasingly we also see models based on rules – sometimes numerical and sometimes logical or linguistic - and they often aim at mimicing the operator or human behaviour.

Modelling other phenomena

Modelling water and wastewater processes and systems involve many kinds of models on top of the biological reaction models, mentioned above. The hydrodynamics is often assumed to be quite simple and mostly the biological reactors in wastewater treatment systems are represented as either a series of completely mixed reactors or as a plug flow tank. Incomplete mixing may often be a reason for significant modelling errors.

In solids/liquid separation there are other sources of incomplete knowledge and of modelling difficulties. One challenge has to do with the modelling of the relative settling velocity of the solids in the liquid. This is related to the floc forming process in the biological reactor. The other challenge is related to the hydrodynamics. Very often it is assumed that the solids/liquid separation is represented by an ideal separator or as a one-dimensional system, where hydrodynamics is not considered. In reality, however, the liquid forms significant flow patterns depending on the influent flow rate and on the boundary conditions of the tank.

Sludge bulking is a well-known phenomenon that causes large operational problems. Still we have quite an incomplete knowledge about the mechanisms. One is the formation of filamentous organisms. We know that this is related to the availability of oxygen and food. Still we lack reliable models of the floc formations in combination with the behaviour of the solids/liquid separation with both filamentous and floc forming organisms present.

Wastewater treatment presents interesting combinations of microscopic and macroscopic challenges. We can only manipulate the system with macroscopic methods, such as flow rates, air flow rates, recycles, dosing carbon sources and chemicals, etc. Still the consequences can often be seen at the microscopic level, where the organism behaviour is determined by our plant manipulations. As an example we can favour simultaneous nitrification-denitrification depending on the availability of dissolved oxygen and organic carbon.

A major challenge for the system wide control is to predict the flow rate and concentrations at important points in the sewer system some time ahead. Dynamic flow predictions should at the same time be used to minimise flooding problems of streets, basements, etc. This will also include the prediction of the influent load (including both flow rate and concentrations) to the wastewater treatment plant. Such a tool is still not generally available. However, some different tools are available and have been applied.

In a water distribution system the pressure in the water pipes is seldom constant. Instead, it is changing every time the consumption changes, or when a pump starts or stops or a valve opens or closes. Another important reason for pressure changes is small or large leakages. If there is a sudden burst due to a pipe break there will be a pressure wave propagating along the pipe. High frequency measurements of the pressure in combination with low frequency pressure and flow measurements can give significant information not only for the detection of a leakage but for the location of it, both in single pipes and in networks of pipes.

Models for detection, estimation and control

The tasks of an integrated supervision platform are many fold: plant-wide monitoring, database and knowledge base maintenance, fault detection, fault diagnosis and control. Failures or malfunctions of process components or instrumentation increase the operating costs of the plant. Gross failures, such as accidents, massive hydraulic overload or toxic influent in a wastewater treatment plant or large bursts and leakages or toxic spills in a water distribution system, are definitely more serious.

Modern process engineering requires prompt detection and classification of process anomalities, which would minimize the overall repair time by assisting operators and engineers in the diagnosis of system degradation. A specialized field called condition monitoring and diagnosis is devoted to determine the life span of operating devices. However, some devices tend to fail abruptly without showing any symptoms of warnings. Moreover, complex control laws often compensate for and conceal the faults. That is why real time process supervision assumes a central role in timely fault detection and isolation of complex systems.

Generally speaking we need models as a basis for controller design. However, for many processes in the water and wastewater industry a control engineer have used the knowledge of the process and its underlying physics to design very simple controllers, such as PID controllers. Analysis methods can be combined with an empirical approach of the problem to design controllers with an acceptable level of performance. During the last decades more powerful control techniques have been developed and used in various industrial sectors in order to increase their productivity. A typical example is that of model-based predictive control (MPC) in the chemical industry, where improving the pureness of some products by a few percent can yield very important profits. An interesting aspect of MPC is that it was born more than 30 years ago and developed in the industrial world, based on a very pragmatic approach of what optimal industrial control could be. It only began to arouse the interest of the scientific community much later.

Nearly all modern optimal control design techniques rely on the use of a model of the system to be controlled. As the performance specifications became more stringent with the advent of new technologies, the need for precise models from which complex controllers could be designed became a major issue, resulting in the theory of system identification. The theorists of system identification quickly oriented their research towards the computation of the 'best' estimate of a system, from which an optimal controller could be designed, using the so called certainty equivalence principle: the controller was designed as if the model was an exact representation of the system.

However, a model is never perfect, with the result that the controller designed to achieve some performance with this model may fail to meet the minimum requirements when applied to the true system. Robust control is an answer to this problem. It uses bounds on the modelling error to design controllers with guaranteed stability and/or performance. During the last twenty years many efforts have been accomplished in order to bridge the gap between robust control and system identification and a new sub-discipline has appeared, called identification for control. This became a very active research area during the 1990s. It resulted in iterative schemes where steps of closed loop identification – using the latest controller designed – alternated with steps of control design, using the last developed model.

Verifying models

Verifying a model to the physical reality is far from trivial. First of all one has to define the criteria for the basis of the decision. Traditionally this has been determined based on the residuals, i.e. the different between the model output and the true plant output. Some statistics can be evaluated based on these residuals or the prediction errors.

The activated sludge models are represented by both organism and various substrate concentrations; each one of them is a state in the model. The organism concentrations can not be measured on-line and still it is very desirable to know them. In principle this can be accomplished by state estimation, where on-line measurements are combined with a process model to find out the current values of the unknown states. A lot of attempts have been made and many estimation methods are available. Most of them are variants of the famous Kalman filter. A major problem that is often overlooked is the model accuracy. A model error in the filter will cause estimation errors.

Another challenge is to identify model parameters, given on-line (and noisy) measurements. Typically the activated sludge models have many parameters and not all of them can be estimated from plant measurements. Given a model having "perfect" measurements it is possible to find out the identifiability of some of these parameters. However, with more noisy measurements the identifiability is often in doubt.

To track the current process operational state via the instrumentation is called monitoring. In a sophisticated treatment plant there is a huge data flow from the process. More instrumentation and new instrumentation development will further provide more data. Unlike humans, computers are infinitely attentive and can detect abnormal patterns in plant data. Information technology can be used more to encapsulate process knowledge, i.e. knowledge about how the process works and how to best operate it.

Model complexity vs. accuracy

When characterizing models we talk about both complicated and complex models. In many processes there are complicated ingredients, like nonlinearities, time varying parameters and time delays. When modelling large, integrated water systems the complexity is of another kind. They include several phenomena of biological, physical and chemical nature. They include a large number of states. The systems are often stiff, i.e. the ratio between the fastest and the slowest response times are several orders of magnitude. Measurements are often scarce, which means that the identifiability and model verification is difficult. Furthermore, the systems are often widespread geographically.

A lot of research is being spent on integrated modelling of systems. Already during the 1970s and 1980s large and comprehensive water models were developed. Relatively few measurements were gathered to fit the models to data and often the model was fit to the data. There are many examples to show that such models perform only marginally better, or sometimes worse, than simple models. The lesson to learn is that the accuracy of a model does not always increase if the number of states increases. If the purpose of the model is to find an accurate relationship between manipulated variables and measurements, then a small dimension model may be satisfactory. The parameter identifiability problem is always present in large systems. This means that several sets of parameters may satisfy the data set. Therefore we may never be confident that the model will predict correctly, once we are outside the operating range of the data set. However, often we do not need to know the exact answers in terms of amplitudes and response times. We wish to understand the various interactions between phenomena and illustrate qualitative behaviour. The human has often difficulties to understand couplings, especially if there are a lot of internal feedback loops in the system.

Large systems are not only a collection of equations for different mass balances. The crucial question is often to formulate the purpose of the model, the relevant time scale, and identifying manipulated variables. In particular, for a plant wide model we need to find a systematic and structured way to coordinate all control actions.

In the full presentation more examples and references will be given about research results and needs.