MODELING AND SIMULATION OF WASTEWATER TREATMENT PROCESSES

John F. Andrews*

Rice University, Houston, Texas, USA

ABSTRACT

Several different types of models which can be useful in describing the dynamic behavior of wastewater treatment plants are discussed. Included among these are (1) visual, (2) linguistic, (3) mental, (4) physical, (5) mathematical, and (6) fuzzy models. Some of the basic concepts of modeling with emphasis on the relationship of these to engineering research and practice are then presented, these being (1) differences in the use of, and expectations from, models by researchers and practitioners, (2) sources of information useful for model development, (3) the amount and type of testing needed for model validation, and (4) the required accuracy of models. Computer simulation is necessary for the solution of most dynamic models. Recent advances in both hardware and software for personal computers have resulted in inexpensive, user-friendly systems suitable for use in both large and small organizations as well as by individual engineers.

KEYWORDS

Dynamic modeling; computer simulation; wastewater treatment; model classification; model development; model testing; model accuracy; mathematical modeling.

INTRODUCTION

Neither modeling nor simulation are new and both have a long history of use in engineering and science. Scaled-down physical models have long been used for simulation in such diverse areas as astronomy (planetariums), hydraulic engineering (river models), architecture (building models), and chemical engineering (pilot plants). Biologists, chemists, and sociologists have studied model organisms (E. coli for example), model compounds, and model communities, respectively. Even the hypothesis which is formulated in applying the scientific method can be considered as a verbal model and simulation may be used to test the validity of a hypothesis when it is not feasible to test the real system.

This presentation will discuss modeling and simulation of wastewater treatment plants with emphasis on mathematical dynamic modeling and computer simulation. The literature on these topics is much too voluminous to review herein. For a more detailed description as well as some specific examples the reader is referred to Andrews (1992) and Patry and Chapman (1989). Discussion will be limited to some of the basic definitions and concepts with emphasis being placed on how these relate to both engineering research and practice.

The author strongly feels that progress in developing and applying dynamic models to wastewater treatment can be accelerated by practicing engineers learning how to develop their own dynamic models and conduct

* Address correspondence to John F. Andrews, 1719 Rayview, Fayetteville, AR 72703, USA
their own computer simulations. It is hoped that this paper will contribute to these goals. Rapid advances in both hardware and software for personal computers have resulted in inexpensive, user-friendly systems suitable for use in both large and small organizations as well as by individual engineers.

CLASSIFICATION OF MODELS

There are many types of models with the type selected being primarily dependent upon the purpose for which it is to be used. One possible classification is given below:

<table>
<thead>
<tr>
<th>Classification of Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
</tr>
<tr>
<td>Linguistic</td>
</tr>
<tr>
<td>Mental</td>
</tr>
<tr>
<td>Physical</td>
</tr>
<tr>
<td>Mathematical</td>
</tr>
<tr>
<td>Fuzzy</td>
</tr>
</tbody>
</table>

Each type of model can be further broken down into different categories. For example, visual models may consist of schematic diagrams such as flow diagrams for treatment plants. These diagrams are adequate for illustrating system components and qualitative descriptions of some of the interactions between components. However, they are not adequate for quantitative descriptions and a following step for design engineers is the preparation of plans (another type of pictorial model) to give dimensions of the components and distances between them.

There are also several different categories of linguistic models. An example is a procedural model consisting of a written and ordered list of tasks to be accomplished for a construction project. The addition of time to a procedural model results in a schedule (dynamic model) for completing the project. More recent examples of linguistic models are expert systems. These are a class of artificial intelligence (AI) computer programs usually intended to serve as computer assistants for decision making or for direct use in automatic control. In the context of this paper, they are usually written in the form of If-Then rules for describing dynamic behavior and specifying control actions. They overlap with mathematical models since they are based on the laws of logic. An example of a rule for a pumping station might be "IF the water level in the wet well is above 10 ft. THEN turn on pump No. 2." A starting reference for expert systems is the book by Bielawski and Lewand (1988). A discussion of the application of expert systems to the activated sludge process has been given by Barnett et al. (1992).

A type of model which must precede either visual or linguistic models is a mental model. A common example of the interaction between mental, linguistic, and visual models in engineering is when a young engineer converts his mental model of an idea into a linguistic model in discussions with his coworkers. During the discussions, the key points of the model are often illustrated by sketches or visual models. A more recent example is the expert system computer program in which the mental models of skilled operating engineers are converted to If-Then rules. The emerging field of Knowledge Engineering encompasses the job of gathering and codifying (quantifying) the knowledge of human experts.

Physical models are well known and often used by engineers and scientists. In wastewater treatment plant engineering, three types of physical models are used in taking an idea (mental model) to application in the field. These are bench, pilot, and prototype. Bench scale experiments are designed to give "Yes or No" (qualitative) answers while pilot scale experiments give more quantitative answers. However, the ultimate test is whether the idea will work under field conditions so prototype testing is also an important step in physical modeling.

Mathematical models are used for quantitative descriptions and consist of one or more equations relating the important inputs, outputs, and characteristics of a system. To those not familiar with systems engineering
terminology, the term mathematical model is sometimes frightening since it may bring to mind large sets of complex equations and the use of sophisticated mathematical techniques. However, this need not be the case since most common engineering design formulae may also be called mathematical models. For example, as simple an expression as \( y = mx + b \) can be considered as a mathematical model where, in systems engineering terms, the system output (\( y \)) is related to the system input (\( x \)) by the system parameters (\( m, b \)).

Mathematical models can also be classified in many different ways, for example as steady state, dynamic, mechanistic, empirical, deterministic, or stochastic. For a more detailed discussion of these classifications, the reader is referred to Andrews (1992). A newer type of empirical model with which the reader may not be familiar is the neural network model. A good description of these models oriented toward their use on personal computers is given in the book by McCord-Nelson and Illingworth (1991)

Fuzzy models are a compromise between the vague statements which humans often use and the strict logic of rule based expert systems (Yes or No or True or False) and quantitative answers provided by other mathematical models. They permit improved communication between the computer and humans by converting (translating) human statements such as high, normal, and low into numbers which the computer can understand (defuzzification) and numbers generated by the computer into statements for humans (fuzzification). A common example of the use of a mental fuzzy model is the ordinary traffic signal in which a green light means \textit{Go} and a red light means \textit{Stop}. However, what does a yellow light mean? If we are very close to an intersection when the light turns from green to yellow we may \textit{Go} on yellow since the alternate could be being hit from behind if we brake too rapidly. However, if we are some distance from the intersection when the light turns from green to yellow we will most likely slow down and then \textit{Stop}. In other words, we have exerted fuzzy control over our automobile based on vague information (very close or some distance from the intersection).

The author is of the opinion that all of the different types of models are useful under the appropriate circumstances and that practical models for application will ultimately make use of all of the different types. An example of the combination of expert system (linguistic) and neural network (mathematical) models is that of Caudell (1991) which she calls "expert networks". Another example is the work of Barnett and Andrews (1992) on a combination of a mathematical model (differential equations) with an expert system and a fuzzy model for studying the dynamics and control of the anaerobic digester.

### SOME BASIC CONCEPTS

Modeling is widely used by many different disciplines as well as for many different purposes. It is therefore quite understandable that there would be disagreements between the different users and perhaps nowhere are these differences more obvious than those between engineering researchers and engineering practitioners. Some of the basic concepts of modeling, as listed below, will be discussed with emphasis on these differences. It is hoped that this discussion will lead to improved understanding of the different viewpoints and more joint efforts by engineering researchers and practitioners in applying modeling to both the design and operation of wastewater treatment plants. The discussion is heavily oriented toward mathematical modeling; however many of the issues addressed are applicable to other types of modeling.

<table>
<thead>
<tr>
<th>Some Basic Concepts of Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of models</td>
</tr>
<tr>
<td>Information for model development</td>
</tr>
<tr>
<td>Testing of models</td>
</tr>
<tr>
<td>Required accuracy</td>
</tr>
</tbody>
</table>
Use of Models

The primary objective of most engineering researchers is to extend engineering knowledge and they see modeling as a tool for guiding research with possibly application of the model in practice at some future date. The researcher's initial models may be considered as mathematical versions of the verbal hypotheses used in the scientific method in which the possibility of being wrong is implicitly recognized. The researcher thus expects and is prepared for the fact that the model may give inaccurate or completely false predictions. This recognition of initial models as hypotheses makes it clear that the development of models is an iterative process (Fig. 1) in which models are formulated, predictions compared with observed results, and the model revised if necessary. This procedure is repeated until adequate agreement is obtained between predictions and known facts and can be a time consuming process since years are sometimes required before a research model is ready for practical application.

![Fig. 1. Steps in the development of mathematical models.](image)

The practitioner must usually operate on a much shorter time scale and is more concerned with immediate application of models to assist in design, construction, or operation of real systems with advancement of engineering knowledge being secondary. Practitioners usually recognize the need for collecting data to obtain numerical values for some of the parameters and expect that there will be some uncertainty in model predictions; however, they normally expect the basic structure of the model to be such that it will give satisfactory predictions without being modified. The practitioner cannot afford highly inaccurate predictions since the consequences (loss of life, reputation, money) can be severe. The practitioner may also not be able to take the time to go through several iterations of the modeling process. These differences in the use of and expectations from models by engineering researchers and practitioners can lead to disagreements as to the value of models unless their intended use is clearly defined.

Information for Model Development

Other points of possible misunderstanding between engineering researchers and practitioners, as well as between scientific and engineering researchers, are the sources of information used in developing mathematical models. Figure 1, which might also be called a block diagram of the scientific method, illustrates the use of four sources of information, these being: (1) existing knowledge, (2) observation of plant behavior, (3) model predictions, and (4) the results of planned experiments. The emphasis placed on each of these different sources of information is very much dependent upon the basic discipline and background of the developer or user of the model. For example, design engineers naturally tend to stress sources of design information whereas operating engineers will emphasize their experience, and that of others, in observing plant behavior. Some scientific researchers go by the maxim that "the experiment is
everything" and thus emphasize the conduct of physical experiments while some engineering researchers emphasize modeling and simulation techniques and may be saying that "the model is everything".

The author is of the opinion that many different sources of information have potential value for the development of models. This is especially so for wastewater treatment plant models which require knowledge from a wide variety of scientific and engineering disciplines as well as practical knowledge from plant designers and operators. It is the synthesis of existing knowledge from other disciplines into models that may be one of the distinguishing characteristics between engineering and scientific research. Engineering researchers do consider new applications of existing knowledge from other branches of science or engineering as original research whereas only the discovery of new knowledge is considered as original research by some scientists.

Testing of Models

The amount of testing needed will be a function of the purpose for which the model is to be used and the blend of existing knowledge and new hypotheses incorporated in the model. If the model is entirely based on well established principles from the sciences or other engineering disciplines, the structure of the model should be adequate and all that may be required is determination of numerical values for the parameters in the model. Moreover, first estimates of some of these may either be available or calculable from data available in the literature. An example might be a coupling of hydraulic models for individual processes for simulation of the dynamics of fluid flow in an entire plant as reported by Olsson and Stephenson (1986). Such models could well have been established in previous years since most of the basic knowledge (fluid mechanics) needed to set up the equations comprising the model has been available for some time. However, they may not exist for a reason previously given, this being that most engineers are practical people and inclined to say "Why set up equations for which it is not currently possible to obtain solutions?" In other words, advances in solution techniques may result in a renewed interest in old problems.

In the above case it might be more appropriate to say that the structure of the model requires no testing and the purpose of testing is to determine numerical values for the parameters. Such models are no longer in the realm of engineering research but are instead ready for engineering application.

At the other extreme are research models containing primarily new hypotheses for which only qualitative observations are available to support the hypotheses. Such models need extensive: testing to prove or disprove the hypotheses and, if these are proven correct, to structure the model so that it can be applied in practice. Both modeling and simulation can play a significant role in this testing. Simulations using the model can be used to design physical experiments with first experiments usually being at bench scale to prove or disprove hypotheses and obtain rough estimates of parameters. The model is then modified to incorporate these increases in knowledge and then used in simulations to design experiments for pilot and/or full-scale testing.

Figure 1 can also be used to illustrate the difference in attitude toward the testing of models by engineering researchers working on "natural" systems (streams, lakes, estuaries, etc.) and those working on "man-made" systems such as treatment plants. Experimentation is frequently not possible for natural systems and the usual model testing approach for these systems is to collect data over a long period of time and then analyze the data to examine the "fit" between model predictions and the data. However, what if the dynamic phenomena of interest (a "spill" of a toxic chemical, for example) does not occur during this period of time? Or if the frequency of the measurements is inadequate to quantify the dynamic response?

With man-made systems such as treatment plants, experimentation is often possible and an integral part of model testing should be the design and conduct of carefully planned experiments including appropriate input disturbances and frequency of measurements so that data on the phenomena of interest can be more rapidly and accurately obtained. This type of information may be difficult to obtain from normal operating records since the control exerted in normal operation can remove much of the dynamic information of interest. In other words, the signal to noise ratio is too low. During times of crisis, such as process upsets or pending
process failure, significant dynamic phenomena are occurring. However, during such times operating engineers are naturally more concerned with restoring good performance than they are with collecting data for the testing of models! This points out the need for the conduct of carefully planned experiments to evaluate such phenomena.

The testing of models also brings up the issue of measurements (observations) and thus illustrates another difference between the development and use of models by engineering researchers and practitioners. Models for research do tend to be more complex and include more terms than those used for design, some of which may be difficult to measure. This can be an acceptable situation in research since the investigators usually have a good command of the basic theory involved and the ability to use the more sophisticated instruments needed for measurements. However, models for practical applications should be as simple as possible and focus on the inclusion of parameters or variables which are relatively easy to either measure or specify. A model which is too complex to understand or requires complicated measurements for implementation is subject to either "misuse" or "disuse".

**Required Accuracy**

How accurate must be the prediction of a model? Is a "Yes" or "No" sufficient or must the answer be more quantitative? If it must be more quantitative, how closely must the predicted and observed behavior agree? The answers to these questions depend upon the purpose for which the model is to be used.

In the early stages of model development, a "Yes" or "No" answer may be quite adequate. For example, in the author's early work on the dynamics of anaerobic digestion (Andrews, 1968) it became apparent that the model must be able to predict that process failure can be caused by organic overloading since this phenomena was known to occur in the field and existing models could only predict failure by hydraulic overloading. Once the occurrence of the phenomena could be predicted on a "Yes" or "No" basis, the next step was to expand the model (Andrews, 1969; Andrews and Graef, 1971) so that it could predict the time-dependent behavior of the key variables normally monitored for evaluating process condition. The term "semi-quantitative", which is deliberately vague, was coined to indicate that predictions were in the appropriate direction (up or down) and of the right order of magnitude. Confirmation of these semi-quantitative predictions (model testing) could be obtained from the literature and discussions with experienced operating engineers. However, more quantitative model testing necessitates physical experimentation.

The differences in agreement between model predictions and observed results is due to uncertainty in the inputs to the model, the model itself, and the information available for testing the model. In design, uncertainty is taken into account by the use of safety factors which often result in increases in size. For example, there is usually considerable uncertainty involved in predicting the population of a city at some time in the future (the design period) so the design engineer must use his best judgment to make this prediction. Increases in size to accommodate the uncertainty in these predictions should not be called overdesign, as they are sometimes labeled, but are instead attempts to protect against uncertainty. Since there is always uncertainty in attempts to predict the future, the possibility exists that the load on the plant at the end of the design period can be either larger or smaller than that predicted.

Process control based on model predictions is also subject to uncertainty. However, unlike design, it is possible in many cases to correct for this uncertainty by using automatic feedback control in which the amount of control exerted depends upon the difference between the desired and observed values of performance (the error signal). This means that a reasonable amount of uncertainty can be tolerated in a dynamic model for process control and models with considerable error in the predictions (semi-quantitative) can also be useful. It should be noted that feedback is also involved in process design in that when the design engineer becomes aware of deficiencies in his design, he uses these to correct future designs. However, the feedback is not automatic and is on a much longer time scale (years and sometimes decades) than feedback in process control. It may therefore not be very useful to the operating engineer who has to operate in a much shorter time scale with existing facilities.
COMPUTER SIMULATION

After a dynamic mathematical model has been developed for a process, the equations which comprise the model must be solved in order to predict process behavior with respect to time. This is known as simulation and can be defined as the use of a model to explore the effects of changing conditions on the real system. Obviously, the model must be a reasonable representation of the real system in order for the results to be meaningful since simulation results can be no better than the model and data on which they are based.

Prior to the advent of computers, a computational bottleneck (Fig. 2) existed and efforts at mathematical modeling were frequently of little practical value since the equations comprising the model could not be solved. Engineers, being practical people would usually say, “Why set up the equations when it is impossible to solve them?” The ready availability of computers and simulation languages has largely eliminated this bottleneck with the current bottleneck being primarily the development of realistic models.

![Fig. 2. Difficulties in developing mathematical models before and after the availability of computer simulation (after Franks, 1967).](http://iwaponline.com/wst/article-pdf/28/11-12/141/24370/141.pdf)

**History of Computer Simulation.**

Early computers used for simulation were analog in nature and involved substituting relationships between voltage, current and resistance for those of other physical variables. Some analog computers were general purpose while others were designed to simulate specific systems. The general purpose analog computer could be used to simulate any system for which a deterministic mathematical model could be developed. An example of the use of an analog computer for simulation of a specific system was the McIlroy analyzer in which vacuum tubes with special characteristics were used to simulate the steady state relationships between fluid flow (current) and pressure drop (voltage) in pipe networks for water distribution systems, with the characteristics of the pipes (coefficient of friction, length, diameter) being proportional to the resistance of the tubes.

Although useful in their time, analog computers were not easily programmed by many practicing engineers since they required "scaling" of electrical variables (voltage, current, resistance) to represent numerical values of the variables of interest and "patching" of a computer board to electrically connect the different analog elements. They have been largely replaced by simulation languages for use on digital computers. In these simulation languages, the differential equations comprising the mathematical model are solved (simulated) using numerical techniques to approximate continuous functions by discrete functions. Neither "scaling" nor "patching" is required since the digital computer can handle a wide range of numerical values and can automatically sort the equations to be solved so that variables are calculated in the appropriate order. However, users of simulation languages should always keep in mind that numerical solution of differential
equations is an approximation and sometimes may give incorrect answers. It is especially important to be aware of this when studying process stability and one must always ask the question "Is the instability really due to process instability or is it due to instability of the numerical technique used for approximate integration?" Users who plan to make extensive use of computer simulation are urged to learn more about numerical analysis so that they will be aware of the strengths and weaknesses of the different techniques available.

The early use of digital computers for simulation was primarily restricted to specialists since a detailed knowledge of numerical techniques was required and a considerable amount of time was needed to learn programming languages such as FORTRAN. These problems were overcome by the development of simulation languages which were heavily user oriented ("user friendly") thus permitting the engineer or scientist to concentrate on model development and interpretation of simulation results instead of programming details. Most simulation languages could also be learned in a relatively short period of time (hours instead of days or weeks) thus satisfying the engineering desire to "get on with the solution of the problem." Franks (1967) has presented a history of the early simulation languages.

The early simulation languages were batch oriented since most digital computers of that era (1960–1975) were operated as batch systems using relatively expensive computers. This type of operation impeded a more widespread use of computer simulation since the user needed access to a computer center and was not able to interact directly with the computer. These obstacles were removed by low-cost personal computers (PCs) which became available in the early 1980s. A wide variety of simulation languages are now available for use on PCs. The Society for Computer Simulation (1991) publishes an annual directory of simulation software. The 1991 edition of this directory lists 116 general and special purpose simulation software packages with 77 of these being suitable for use on PCs.

The author has used computer simulation since the mid-1960s and during this time has used general purpose analog computers, simulation languages (PACTOLUS and CSMP) for main frame computers and now uses SIMNON (1993) on a personal computer. The PC version of SIMNON can handle the simultaneous solution of up to 300 differential equations. More details, with some simple examples, on the use of SIMNON can be found in Andrews (1992). An example of a more comprehensive simulation system is the general purpose simulator (GPS) which has been described by Patry and Takacs (1992). Their paper also describes the use of GPS in a study of the Hamilton-Wentworth plant in Canada. This simulator, which is designed for use on a Sun workstation, contains a variety of primarily deterministic mathematical models for wastewater treatment processes which can be easily linked together to form a plant model using object oriented concepts and an interactive screen oriented interface. Another simulation package oriented toward the dynamic modeling and control of wastewater treatment plants is SACCESS (Statistical analysis and control of environmental systems) which has been described by Hiroaka and Fujiwara (1992). This package, available in both workstation and PC versions, is based on time series analysis and thus represents modeling using both deterministic and stochastic components. The application of SACCESS to modeling and control of the Kawamata plant in Japan has been described by Hiroaka and Tsumura (1992).

Although space does not permit a detailed review, it should be noted that there are many general purpose PC software packages available for simulation using neural networks, expert systems, and fuzzy logic. The term "general purpose" is used herein to indicate that these packages are not specific to any one application, such as wastewater treatment, but can instead be used for many different types of applications. Special computer chips are also available for neural networks and fuzzy logic so these may be implemented in hardware as well as software. Mention should also be made of the use of commercially available spreadsheets (Excel, Lotus 1-2-3, etc). Most readers will be familiar with the use of spreadsheets for solving explicit algebraic equations and perhaps also their use for iterative ("trial and error") solution of implicit equations. However, they may not know that spreadsheets can also be used for dynamic modeling (solution of differential equations) as discussed by Julian (1989) and EI Shayal (1990).

Advantages and Disadvantages of Computer Simulation

Computer simulation has many of the same advantages, and disadvantages, as physical simulation.
Considerable knowledge can be gained about a process through the development of a mathematical model and subsequent computer simulations using the model. Sensitivity analysis, or the response of the model to changes in specific variables, can be used for model improvement by indicating those variables which are most significant. Sensitivity analysis can also be used for model simplification by indicating those variables which have little effect on system performance. Simulation permits examination of large systems, such as river basins, where physical experimentation on the full scale system may not be possible. Time can be compressed on the computer with simulations being conducted in seconds or minutes. This is especially important for the biological processes used in wastewater treatment where rates are relatively slow and physical experimentation may require weeks and even months.

There are, of course, disadvantages to computer simulation. The results of the simulations are no better than the mathematical model and data on which they are based. Poor models and/or poor data can give inaccurate or even completely misleading results. Also, some of the advantages of computer simulation can be disadvantages. Computer simulation can be easier, cheaper, and faster than physical experimentation. In the case of wastewater treatment, it can also be a much cleaner job! These factors can lead to one becoming overly enamored with the tools (modeling and computer simulation) and consequently neglecting physical experimentation and/or the collection of data.

SUMMARY

Several different types of models can be useful in describing the dynamic behavior of wastewater treatment plants. Included among these are; (1), visual, (2), linguistic, (3) mental, (4) physical, (5) mathematical, and (6) fuzzy models.

A discussion of some of the basic concepts of modeling with emphasis on the relationship of these to engineering research and practice is presented. These concepts are:

- differences in the use of, and expectations from, models by researchers and practitioners
- sources of information useful for model development
- the amount and type of testing needed for model validation
- the required accuracy of models

Dynamic mathematical models are solved using computer simulation to predict plant behavior with respect to time. The early use of digital computers for simulation was primarily restricted to those with a knowledge of numerical techniques and computer languages such as FORTRAN. These problems were overcome by "user friendly" simulation languages which permitted concentration on model development instead of programming details. However, the early simulation languages were batch oriented for use on relatively expensive computers controlled by computer centers. The availability of interactive simulation languages for use on low-cost personal computers (PCs) has removed these obstacles.

REFERENCES


