Pergamon

Wat. Sci. Tech. Vol. 39, No. 9, pp. 1–8, 1999
© 1999 IAWQ
Published by Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0273–1223/99 $20.00 + 0.00

PII: S0273-1223(99)00210-3

FICTION AND REALITY IN THE MODELLING WORLD – BALANCE BETWEEN SIMPLICITY AND COMPLEXITY, CALIBRATION AND IDENTIFIABILITY, VERIFICATION AND FALSIFICATION

P. Harremoës* and H. Madsen**

** Department of Environmental Science and Engineering, Technical University of Denmark, 2800 Lyngby, Denmark*
*** Department of Mathematical Modelling, Technical University of Denmark, 2800 Lyngby, Denmark*

ABSTRACT

Where is the balance between simplicity and complexity in model prediction of urban drainage structures? The calibration/verification approach to testing of model performance gives an exaggerated sense of certainty. Frequently, the model structure and the parameters are not identifiable by calibration/verification on the basis of the data series available, which generates elements of sheer guessing – unless the universality of the model is be based on induction, i.e. experience from the sum of all previous investigations. There is a need to deal more explicitly with uncertainty and to incorporate that in the design, operation and control of urban drainage structures. © 1999 IAWQ Published by Elsevier Science Ltd. All rights reserved

KEYWORDS

Determinism; grey box models; modelling; stochasticity; urban drainage; uncertainty.

INTRODUCTION

The practise of urban storm drainage is changing from simple design rules to use of deterministic models for design. With the development of the modern computer and its ever-increasing calculation power, there has been a surge in the development of models. After the first enthusiasm, it is time to reflect on the purpose, the potential and the loopholes of our faith in and our application of ever more complicated models.

ENGINEERING APPROACHES

There are two approaches to engineering design and operation of technology:

- The **empirical iterative approach**, also called “trial and error”.
- The **deterministic predictive approach.**

Remarkable engineering accomplishments have been achieved by the **empirical iterative approach**: Roman aqueducts still standing, drainage in ancient Athens, middle age gothic cathedrals and many other famous structures. Many mistakes paved the road to these successes: the tower in Pisa is still (just) standing. The bridge over the St. Lawrence River in Quebec was designed by scaling up the bridge in Scotland across the river Forth, but columns cannot be scaled up proportionally and the bridge fell down! The empirical iterative approach is still a valid approach. Those favouring this approach tend to pride themselves by calling themselves: "practitioners".

In the other end of the scale is the **deterministic predictive approach**, which favours developing an understanding of all elements in the structure, so that the performance of the structure can be predicted. On that basis the structure can be designed to meet predetermined requirements. The enthusiasts of this approach will be called "theoretists".

There has always been a schism between the practitioners and the theoretists, frequently in relation to international conferences, at which it is difficult to satisfy both. The fact is that there is little justification for this schism, because there is need for both, e.g. expressed by the joke: "There is nothing more practical than a good theory".

The schism between the two approaches can also be illustrated by the development of engineering education, which has become still more theoretical – leaving the student to achieve know-how from practise after graduation. It is a valid question to pose, whether the tendency to educate by exemplified application of models is a good approach. Do the students get a basic understanding of theory and/or practise, or do they get entangled in the mechanisms of futile handling of hard- and software, which will be obsolete before the students leave the university?

DETERMINISM

Determinism is based on a **reductionist philosophy**. It is based on the concept that a full understanding can be achieved by identification and description of all underlying physical, chemical and biological laws of nature that govern the engineering application. On reflection, there is reason to question the validity of this basic axiom.

We have many indications that this is a very reasonable axiom, but philosophically we have indications that the validity is limited:

- Moving from macrophysics to microphysics, the Bohr-model of the atom is not deterministic and Heisenberg's uncertainty theorem quantifies the limitation.
- The theory of "chaos" showed us that there might not be unique solutions to even simple differential equations. In fact, at university I learned only about the soluble problems, which turns out to be a small fraction of the universe of problems.
- Neural networks illustrate a totally different set of learning and "understanding": pattern recognition – the holistic approach. Are we back to the empirical-iterative approach?

In my experience, the university graduated engineer goes into practise with an exaggerated faith in determinism, while the practitioners, who - after years in practise - has forgotten his theoretical background, defensively shows disrespect to theoretical approaches. The reality is somewhere in-between – but where?

THE DETERMINISTIC MODEL

The deterministic model is a mathematical description of the physical, chemical and biological phenomena involved in the engineering problem of concern. They can be very simple and they can be very complicated. However, it must be emphasised as a philosophical fact that no model is anything but an approximation to the reality. It may be a good approximation or a poor one, but always an approximation.

In the engineering application the issue is not to develop an ever "better" description. That is frequently the model developers apparent goal. The issue is to identify and develop the model that fits the engineering problem to which it shall be applied. There is no one solution to this. To a multitude of engineering problems fit a multitude of models. Flexibility is a virtue.

From the last twenty years of model development, one gets the impression that "better" models are interpreted as more detailed models. The more that can be described, the better. That is not necessarily a virtue, as will be discussed in the following. Parsimony is a principle introduced by Ockham (eng. philosopher, 1290-1349). **Parsimony** means that the best approach is the simplest that fits the purpose of the application. Again, flexibility of model features is a virtue that fits with parsimony.

A good example of the need for parsimony is **model reduction** for application in real time control. Due to the limited time available for reaching a solution within a time step, the very complicated models are too time consuming for practical application. In a few years time, faster computers might overcome that. However, the complicated models may be difficult to keep stable. It may be difficult to analyse the reason for an error, because the model complexity complicates the understanding of performance.

The key point is that an engineering model has to model the essential features that are important to the resulting design or operation. All other details just obscure the picture and hamper engineering application. The slide-rule engineers of the past had to simplify down to the bare essentials, a few combinations of key parameters, and to incorporate the rest in safety factors. That is (was?) an engineering art. Models and computers have just shifted the level of comprehension; but the art of choosing a perceivable and manageable structure and a corresponding set of parameters is the same at another level. Is the art about to be forgotten?

RECEIVING WATER MODELLING

Combined sewer overflows and separate sewer outlets give rise to pollution. In order to abate the situation, logic demands that the engineering structures be designed and operated such that water quality objectives be complied with (Harremoës, 1989).

The following three points will be illustrated by models of the effect of urban drainage on receiving waters:

1. Models have to incorporate the important phenomena of the problem investigated. If that is not the case, no amount of detail in the description of the other phenomena will improve simulation.
2. Even simple cause-effect relationships may not have stable solutions, but may show chaotic behaviour.
3. The phenomena involved are so diverse and model structures become so complicated that neither structure nor parameters are identifiable with a reasonable effort – if at all possible.

The eutrofied lake

In the 1970's there was an enthusiasm with the development of eutrophication models. The attempt was to develop models that could be used as management tools in dealing with nutrient discharges from wastewater treatment plants and from urban drainage. The approach was typically reductionist: Algae concentration in the eutrofied lake is the measure of eutrophication (turbidity or Secchi disk reading). The postulate was: Algae concentrations are controlled by nutrient availability. The level of eutrophication is simulated by algae growth kinetics in balance with loss functions (respiration, washout, grazing). By comparison with field data it became obvious that the nutrient loading was an overall important driving force, but it is still possible to experience heavily loaded lakes which are clear in spite of the prediction of models based on algae growth. The bottom-up approach lacked important phenomena: **population dynamics**. It turned out that lakes could be managed by a top-down approach: Remove the fish and the daphnia will put so much predatory pressure on the algae that they fail to reach high numbers. The solar energy input to algae growth is converted to respiration and biomass at higher trophic levels. The bottom-up models are still being applied to lake modelling in relation to lake management!

Simple predator-prey relationships may have no simple solutions. They may show chaotic behaviour. That is emphasised in simple relationships, like: algae-zooplankton-fish. The chaotic nature of the (highly simplified) differential equations does not provide unique solutions. That is the nightmare of the design engineer. However, there are trends. It appears that "strange attractors" can be identified, around which the state will oscillate. At one particular nutrient loading there can be several such "attractors"(or local points of equilibrium). Accordingly, increase and decrease on nutrient loading may show "hysteresis" phenomena.

Finally, it is well known that these simple models are far from describing the complexity of the real system (involving rooted vegetation, different algae, zooplankton and fish species). Models gimmicking such complexity have been developed, but little success has been achieved. No experimental effort could make the structure nor the parameters identifiable.

The answer is to use simple models for conceptual understanding in combination with experience, but where to strike a balance between determinism and experience is still an open question.

The polluted river

The classical model of river pollution is the Streeter-Phelps formulation from the 1920's. That is the starting point of all river modelling, to which an array of other oxygen-depleting phenomena have been added. These models were advocated in the 1970's, but the fact is that they never caught on with the biologists of the river authorities. First of all, they predicted an improvement with the introduction of biological wastewater treatment that did not materialise. The failure was due to lack of attention to diffuse sources and due to the fact that the models never build the bridge between the mechanistic cause-effect formulation of loading versus chemical water quality on the one hand and the biological state of the river on the other hand. The latter is the ultimate determinant and that is managed by the empirical-iterative approach, i.e. by experience.

Building that bridge between physical, chemical water quality and biological river quality ought to be a top research priority; but it is hardly receiving attention. Why?

CATEGORIES OF MODELLING

Models should be flexible to fit the identified purpose. They are listed below, according to increasing complexity:

1. Planning models
   The planning model is suited for analysis of a practical, hypothetical situation of the future, where the input is very uncertain. Simple models are justified.
2. Models for conceptual understanding
   Simple model formulations intended for analysis of fundamental interrelationships.
3. Design models
   These models are used for explicit design, like the rational method
4. Models for real time control
   Models for real time control have to be simple in order to accomplish calculations on one time step, in order to maintain stable control and in order to be identifiable. However, the reality of the model can be achieved by on-line adaptation.
5. Models for analysis
   The analysis of the performance of a particular problem in a sewer system, treatment plant or receiving water may call for detailed, complex models, but the monitoring programme must be sufficiently comprehensive to justify the complexity
6. Detailed research models
   Models are excellent research tools. In such cases, model complexity has to fit the issue addressed. That may be complex, but as stated above, that does not necessarily call for complex models – in fact it may be the other way around.

IDENTIFIABILITY, VERIFICATION AND FALSIFICATION

The universality of a hypothesis or a postulate in an open system can never be proven, but confidence in its universality can be improved by induction, i.e. experience. On the other hand, a hypothesis or postulate can be falsified. It takes just one example to prove it wrong. These philosophical facts are worth being reminded of in a situation where the reality of models is routinely postulated on the basis of calibration and verification. The practise is to use half of an available time series for calibration (fitting the parameters of the model) and to use the other half of the time series for verification (showing that the model performed well). There are several points to be made in this context:

- The universality of the calibrated/verified model does not go beyond the universality of the data series used for the analysis in question – unless the model structure and the parameters, that were not calibrated/verified by the time series in question, have a universality indicated by induction, i.e. experience from other time series.

This point is very relevant to urban drainage structures, which are frequently designed on the basis of extreme events, e.g. flooding or oxygen depletion in rivers. Data series seldom incorporate extreme events – and if they do, very few data are of an extreme character. The calibration/verification of models is rarely related to the extreme events to which the models are applied. The reality of the extrapolation is based entirely on theoretical considerations based on induction (experience) from rare sets of data – if any. This can be exemplified by the simple question: To what extent do permeable urban areas contribute to runoff during rare intense rains (Arnbjerg-Nielsen and Harremoes, 1996). There is only one theoretical solution: Use our best theoretical knowledge of the basic hydrological mechanisms, extrapolate on that basis and hope that a universality applies to the case in question – but do not claim that the model has been calibrated/verified, please.

- Calibration/verification of a model for use in practise to a particular urban area seldom provide data series of such comprehensiveness that it permits calibration/verification of other than a few local characteristics. In fact, many of the parameters of the model are not identifiable for the simple reason that the time series in question does not contain the information required to determine the parameters

The loophole of calibration/validation in relation to complex models is that a good fit can be achieved with different sets of parameters – leaving the "calibrator" with a false sense of certainty. That mistake has been made repeatedly. In some cases, it is inherited in the system that the model incorporates certain combinations of parameters that can only be calibrated as a combination – unless very elaborate scientific investigations are undertaken. The key question is: Does this reality call for simple models or complicated models in which certain features are taken for granted and not subject to calibration/verification? So far, the experience is that the approach should be tailored to the situation in question.

These issues should be given more emphasis – otherwise the modelling approach will loose credibility.

GREY-BOX MODELS

Deterministic models of a system are characterised by the fact that the future states of the system can be predicted exactly, no matter the prediction horizon. Hence it is obvious that most environmental systems can not be sufficiently modelled using deterministic models. Such models act as approximations on shorter time scales, and moreover the degree of approximation is not described by the model.

Stochastic models are most often rather simple models; but such a model is able to express how accurate the states of the system can be predicted, and this information is very useful for many practical applications, like prediction and design of controllers. Since the stochastic models contain elements which describe the uncertainty in terms of distribution functions or by some of the moments of the random variation (like the auto covariance function) statistical approaches can be used for identifying the models structure as well as for estimation of the model parameters. Furthermore, a large number of statistical approaches for validation of the model exists. Historically the most widely used stochastic models for environmental modelling are regression models (generalised linear models) and transfer function models (like the ARMAX- and Box-Jenkins models). Stochastic models of this type are very often called black-box models. A serious drawback of using black-box models is that, in general, it is difficult to provide a physical interpretation of the estimated parameters.

The major advantages of using deterministic models is that most prior physical, chemical or biological knowledge about the system exists, in terms of deterministic models of subprocesses in the system. However, most practical environmental systems represent a very complex structure of coupled subprocesses of which each subprocess represents a detailed discipline in its own right, and the subprocesses interacts across a wide spectrum of space and time in a complicated manner which ultimately determines the dynamics of the complete system. If detailed models of the dynamics of each of the subprocesses of the environmental system existed, and the necessary computing power were available, one could try to obtain a total model simply by coupling the individual sub-models together into a very comprehensive numerical model. However, it is questionable whether such a deductive, reductionistic approach would be successful. Even if the unavoidable idealisations and simplifications, introduced into the models for the individual subprocesses, are verified individually, it does not necessarily follow that these idealisations and simplifications are valid for the whole system too.

Grey-box models combine the deterministic and stochastic approach such that the model structure and the parameterisation are based on prior knowledge about the system, and the parameters can be directly physically interpreted. The models are stochastic and most adequately formulated as a system of stochastic differential equations. Statistical procedures can be used for evaluating the model structure, estimating the model parameters and for a validation the model. The relevant statistical procedures are, however, more involved than the procedures known for the above mentioned black-box models.

The grey-box approach makes it possible to combine information in terms of physical, chemical or biological knowledge with information embedded in any available data. Grey-box modelling is thus a combination of the deductive and the inductive approach. Once a grey-box model is obtained for the most important relations in a system, it can be iteratively improved by a combination of more detailed comparisons with data and use of more detailed physical facts.

Grey-box modelling was first proposed in Bohlin (1984). Today this approach is used in many disciplines, and for instance to model the variation of the oxygen content in a small river (Jacobsen, *et al.*, 1995) and to model the accumulation of organic material in sewer system (Bechman *et al.*, 1998a, 1998b). The mathematical and statistical details are described in (Madsen and Holst, 1997), and some methods for validation of grey-box models are found in Holst *et al.* (1992).

NOT-KNOWING

We live in a period concerned with the "ultimate" perspectives. Sustainability incorporates the concern for the well being of the generations to come. Climate change is an example of concern that reaches far into the basic assumption on which we base decisions. How does that relate to urban drainage? The uncertainty of the basis for our decisions becomes of paramount importance. Still, the engineering profession, not least in relation to the application of models, talks of "uncertainty" as if it were a quantifiable concept. The following categorisation can be used (Wynne, 1992):

- **Determinism**. It is assumed that all physical, chemical and biological cause-effect relationships are known and that future performance of a structure can be predicted with certainty. That is an ideal never reached in practise.
- **Risk**. We understand the errors and the uncertainty in a quantifiable way. Results can be given as a mean with a statistical distribution. Decision-making can take the likelihood of failure into account.
- **Uncertainty**. The risk of failure cannot be calculated, but we do have a sense of orientation and scale with respect to evaluation of uncertainty. The risk is not quantifiable, but common sense based on experience does provide guidance in decision making.
- **Ignorance**. It has to be realised that we may be completely ignorant with respect to the cause-effect relationship. The history of the last 30 years of environmental concern tells many stories of complete ignorance with respect to effects: The toxicity of organo-mercury compound produced in sediments was unknown (Minnamata, 1963), the effect of PCB on bird egg hatching, the percolation of biocides to aquifers for water supply, the pollutant transport with rain discharge from separate sewer systems. All these and many more examples of serious decision were made in complete ignorance of what the effects were.

- **Inderterminacy**. It can be argued that in many cases there is no way in which we can determine the cause-effect of the future. The sustainability concept is quite idealised. How can the well being of future generations be taken into account when the circumstances and the individual and communal values can be predicted only by extrapolation from our own circumstances and values. Just imagine a prediction of our circumstances and values at the turn of the century 1999-2000 that would have been made by fortune-tellers at the turn of the last century 1899-1900.

It is an interesting observation that the engineering profession during the last half century has been preoccupied with development of deterministic descriptions in still greater detail, and that the calculated risk analysis (as defined above) has been introduced, but is being applied in special cases only. The other items: non-quantifiable uncertainty, ignorance and indeterminancy, are left entirely to the political arena. However, these issues may be much more important than the development of the next detail in an already complex deterministic or stochastic model structure. Is it true that the engineering profession tends to find the solution to the problems that can be solved and leaves little to no attention to the problems that do not have a "solution"?

From a water resource conservation point of view, a century of experience with rain runoff through pipe systems is exchanged by a concept of local infiltration of rain. Have the implications been anticipated? In the same vein, the centralised water supply and sewerage has been governed by the principle of complete separation between the clean and the dirty water. No "dirty" water (that means: not up to drinking water standard) is allowed to re-enter the house. Again, from a water resource conservation point of view, a century of successful experience is exchanged with use of roof runoff and grey wastewater in households. Have the implications been adequately predicted before introduction? Does the profession opportunistically pursue the latest gospel of political concern without regard for non-quantifiable uncertainty, ignorance and indeterminancy, or even risk? Or, would it be untimely conservatism to demand better predictability before implementation? The conclusion is that the profession is not in the driver's seat with respect to new concepts to be anticipated and analysed before they become a political reality. The profession prefers to indulge in further refinements of determinism and risk analysis using the deterministic, predictive approach within the well-known framework of concepts. That is the basis for the introduction of the precautionary principle. New initiatives come from outside and are dominated by the empirical iterative (trial and error) approach. Why? The profession has chosen to ignore non-quantifiable uncertainty, ignorance and indeterminancy, and even risk. It is time for a new paradigm (Harremoës, 1996). The profession must deal realistically with these forms of uncertainty and be more open about the philosophical aspects of not knowing.

CONCLUSIONS

Deterministic modelling has reached the stage where it is time to reflect on the direction of further development. Purely physical models have shown great ability to simulate reality, though significant uncertainty is inherent in lack of knowledge about model structure, input functions and parameter estimation. This uncertainty is much greater, when modelling incorporates chemical and biological phenomena. Model complexity is not a virtue by itself. The best model is a model that provides a suitable simulation of reality with the least complexity and an appropriate set of data for calibration. Model application must include estimation of uncertainty of prediction.

REFERENCES

Arnbjerg-Nielsen, K. and Harremoës, P. (1996). Prediction of hydrological reduction factors and initial loss in urban surface runoff from small ungauged catchments. *Atmospheric Res.*, **42**, 137-147.

Bechman, H., Nielsen, M. K., Madsen, H. and Poulsen, N. K. (1998a). Control of sewer systems and wastewater treatment using pollutant concentration profiles. *Wat. Sci. Tech.*, **37**(12), 87-93.

Bechman, H., Madsen, H., Poulsen, N. K. and Nielsen, M. K. (1998b). Grey Box Modelling of First Flush and Incoming Wastewater at a Wastewater Treatment Plant. Submitted.

Bohlin, T. (1984). Computer-aided grey-box validation. Tech. Rep. TRITA-REG 8403, Dept. of Automatic Control, Royal Institute of Technology, Stockholm.

Harremoës, P. (1989). Overflow quantity, quality and receiving water impact. In: *Urban discharges and receiving water quality impacts* (*Adv. Wat. Pollut. Control* no. 7). Pergamon Press, Oxford, UK.

Harremoës, P. (1996). Dilemmas in ethics: Towards a sustainable society. *Ambio*, **25**, 390-395.

Holst, J., Holst, U., Madsen, H. and Melgaard, H. (1992). Validation of Grey-Box Models, IFAC International Symposium on Adaptive Systems in Control and Signal Processing, Genoble, p. 407-414.

Jacobsen, J. L., Madsen, H. and Harremoës, P. (1995). Modelling the transient impact of rain events on the oxygen content of a small Creek. *Wat. Sci. Tech.*, 33(2), 177-185.

Madsen, H. and Holst, J. (1997). Modelling Non-linear and Non-stationary Time Series, IMM, DTU, 259 pp.

Scheffer, M. (1998). *Ecology of Shallow Lakes*. Chapman & Hall, ISBN 0 412 74920 3.

Wynne, B. (1992). *Uncertainty and Environmental Learning, Global Environmental Change*. Butterworth-Heinemann Ltd.