



Chapter XII

Applications of Data-Driven Modeling and Machine Learning in Control of Water Resources

D.P. Solomatine

International Institute for Infrastructural, Hydraulic and Environmental Engineering
(IHE-Delft), The Netherlands

ABSTRACT

Traditionally, management and control of water resources is based on behavior-driven or physically based models based on equations describing the behavior of water bodies. Since recently models built on the basis of large amounts of collected data are gaining popularity. This modeling approach we will call data-driven modeling; it borrows methods from various areas related to computational intelligence—machine learning, data mining, soft computing, etc. The chapter gives an overview of successful applications of several data-driven techniques in the problems of water resources management and control. The list of such applications includes: using decision trees in classifying flood conditions and water levels in the coastal zone depending on the hydrometeorological data, using artificial neural networks (ANN) and fuzzy rule-based systems for building controllers for real-time control of water resources, using ANNs and M5 model trees in flood control, using chaos theory in predicting water levels for ship guidance, etc. Conclusions are drawn on the applicability of the mentioned methods and the future role of computational intelligence in modeling and control of water resources.

INTRODUCTION

A *model* can be defined as a simplified representation of reality with an objective of its explanation or prediction. In engineering, the term *model* is used traditionally in one of two senses:

- (a) a mathematical model based on the description of behaviour (often physics, or first-order principles) of a phenomenon or system under study, referred to later as behavioural (also process, or physically based) models;
- (b) a model built of material components or objects, which is often referred to as scale (or physical) model (since it is usually smaller than the real system).

These views of a model are widely adopted and taught. Understandingly, in social and economical studies, scale modelling would be a difficult undertaking, but behavioural models based on mathematical descriptions of processes are widely spread and used.

Traditionally, management and control of water resources was based on good understanding of the underlying processes and use so-called “physically based” (or “knowledge-driven,” behavioral) models. These could be for example models based on Navier-Stokes' equation describing behavior of water in particular circumstances. Examples are surface (river) water 1D models, coastal 2D models, groundwater models, etc. Equations are solved using finite-difference, finite-element or other schemes, and results—normally water levels, discharges—are presented to decision makers. Often such models are called *simulation models*. Knowledge-driven models can be also “social,” “economic,” etc.

On the contrary, a “*data-driven*” model of a system is defined as a model connecting the system state variables (input, internal and output variables) with only a limited knowledge of the details about the “physical” behavior of the system. “Hybrid models” combine both types of models.

It should be stated that the process of modeling includes studying the system, posing the problem, data preparation, building the model (normally a machine learning model), testing the model, using the model, interpreting the results and, possibly, reiterating. In this chapter we will consider only the techniques for data-driven modeling proper.

Techniques used in data-driven modeling originate in various areas (often overlapping):

- machine learning (decision trees, Bayesian and instance-based methods, neural networks, reinforcement learning);
- soft computing, and in particular fuzzy rule-base systems induced from data;
- data mining (uses methods of machine learning and statistics);
- methods of non-linear dynamics and chaos theory (often considered as part of time series analysis, but which are actually oriented towards the analysis of large data sets).

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/applications-data-driven-modelling-machine/6839

Related Content

Agent-Based Middleware for Advanced Communication Services in a Ubiquitous Computing Environment

Takuo Suganuma, Hideyuki Takahashi and Norio Shiratori (2010). *International Journal of Software Science and Computational Intelligence* (pp. 1-23).

www.irma-international.org/article/agent-based-middleware-advanced-communication/39102

Developing Concept Enriched Models for Big Data Processing Within the Medical Domain

Akhil Gudivada, James Philips and Nasseh Tabrizi (2020). *International Journal of Software Science and Computational Intelligence* (pp. 55-71).

www.irma-international.org/article/developing-concept-enriched-models-for-big-data-processing-within-the-medical-domain/258866

Machine-Learning-Based White-Hat Worm Launcher in Botnet Defense System

Xiangnan Pan, Shingo Yamaguchi, Taku Kageyama and Mohd Hafizuddin Bin Kamilin (2022). *International Journal of Software Science and Computational Intelligence* (pp. 1-14).

www.irma-international.org/article/machine-learning-based-white-hat-worm-launcher-in-botnet-defense-system/291713

Human Identification Using Gait Skeletal Joint Distance Features

Md Wasiur Rahman and Marina L. Gavrilova (2017). *International Journal of Software Science and Computational Intelligence* (pp. 19-33).

www.irma-international.org/article/human-identification-using-gait-skeletal-joint-distance-features/197783

On Visual Semantic Algebra (VSA): A Denotational Mathematical Structure for Modeling and Manipulating Visual Objects and Patterns

Yingxu Wang (2012). *Software and Intelligent Sciences: New Transdisciplinary Findings* (pp. 68-81).

www.irma-international.org/chapter/visual-semantic-algebra-vsa/65123