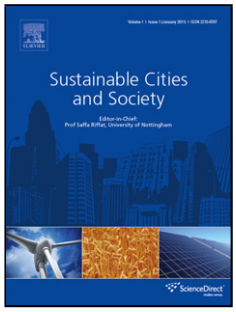
Accepted Manuscript



Title: Enhancing water system models by integrating big data

Author: M. Ehsan Shafiee Zachary Barker Amin Rasekh

PII: S2210-6707(17)30384-0
DOI: https://doi.org/doi:10.1016/j.scs.2017.11.042
Reference: SCS 872

To appear in:

Received date: 10-4-2017
Revised date: 28-11-2017
Accepted date: 28-11-2017

Please cite this article as: M. Ehsan Shafiee, Zachary Barker, Amin Rasekh, Enhancing water system models by integrating big data, *<![CDATA[Sustainable Cities and Society]]>* (2017), https://doi.org/10.1016/j.scs.2017.11.042

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Enhancing water system models by integrating big data

M. Ehsan Shafiee[a], Zachary Barker [b] , Amin Rasekh [c]

[a] *Corresponding author, Hydraulic Engineer/ Data Scientist, Sensus USA Inc. ehsan.shafiee@xyleminc.com*
[b] *Hydraulic Engineer/ Data Scientist, Sensus USA Inc., zachary.barker@xyleminc.com*
[c] *Hydraulic Engineer/ Data Scientist, Sensus USA Inc., amin.rasekh@xyleminc.com*

Abstract

The past quarter century has witnessed development of advanced modeling approaches, such as stochastic and agent-based modeling, to sustainably manage water systems in the presence of deep uncertainty and complexity. However, all too often data inputs for these powerful models are sparse and outdated, yielding unreliable results. Advancements in sensor and communication technologies have allowed for the ubiquitous deployment of sensors in water resources systems and beyond, providing high-frequency data. Processing the large amount of heterogeneous data collected is non-trivial and exceeds the capacity of traditional data warehousing and processing approaches. In the past decade, significant advances have been made in the storage, distribution, querying, and analysis of big data. Many tools have been developed by computer and data scientists to facilitate the manipulation of large datasets and create pipelines to transmit the data from data warehouses to computational analytic tools. A generic framework is presented to complete the data cycle for a water system. The data cycle presents an approach for integrating high-frequency data into existing water-related models and analyses, while highlighting some of the more helpful data management tools. The data tools are helpful to make sustainable decisions, which satisfy the objectives of a society. Data analytics distribution tool Spark is introduced through the illustrative application of coupling high-frequency demand metering data with a water distribution model. By updating the model in near real-time, the analysis is more accurate and can expose serious misinterpretations.

Keywords:
water systems, modeling, big data, automation, Hadoop, Apache Spark, cloud computing

1. Introduction

The water resources community relies on computer models to conceptualize and reproduce behavior of systems, aiding in planning, design, and analysis.

Accepted Manuscript

Download English Version:

https://daneshyari.com/en/article/6775653

Download Persian Version:

https://daneshyari.com/article/6775653

Daneshyari.com