

# **Using Data Analytics and Artificial Intelligence-based techniques to promote water-conscious behaviour through personalised recommendations**

**by Md Shamsur Rahim**

Thesis submitted in fulfilment of the requirements for  
the degree of

**Doctor of Philosophy**

under the supervision of Professor Michael Blumenstein &  
Professor Damien Giurco

University of Technology Sydney  
Faculty of Engineering and Information Technology

June 2021

## **CERTIFICATE OF ORIGINAL AUTHORSHIP**

I, Md Shamsur Rahim declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Information Technology, in the School of Computer Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

**Signature:** Production Note:  
Signature removed prior to publication.

**Date:** 18/06/2021

## Abstract

Water scarcity has become an imminent problem throughout the world due to population growth, prolonged drought, climate change, poor water management, lack of infrastructures and so on. In addition, many metropolitan water utilities are suffering from low water pressure during the hours of peak demand. Several approaches have been taken to overcome these challenges. Rolling out digital or intelligent or smart water meters was one of these approaches. Recent advances in machine learning and data analytics technologies provide excellent opportunities to utilise the vast amount of data generated from these meters.

This study aims to generate new insights from the water end-use data to provide a better understanding of water consumption behaviour and investigate the scope of personalised recommendations for the water industry that will contribute to promoting water conscious behaviours during water restrictions. To achieve this goal, the contribution of this research can be divided into five categories.

First, a systematic literature review was conducted to identify the state-of-the-art and research gaps. The findings from the literature review suggested that highly personalised recommendations through recommender systems (RSs) could play a vital role in promoting water-conscious behaviours. However, before developing an RS for the water industry, some preliminary research issues needed to be addressed. These issues were: advanced household profiling, clustering consumers based on new insights from water consumption patterns, and finally, the integration of RSs in behaviour change interventions.

Second, to achieve personalisation at the household level and identify water consumption habits/behaviour patterns, an advanced household profiling approach is proposed. The proposed approach uses the water end-use data as input and produces two outputs: the time of use and weighted probability of use (TUWPU) data set and the extracted/engineered features (EF) data set. The TUWPU data set provides the probability of an event occurring at a particular time by calculating the weighted probability distribution based on the end-use data using a proposed algorithm. On the other hand, the EF data set is generated by extracting features from the end-use data. The proposed technique is able to identify and address changes in water consumption behaviour which is essential for promoting and sustaining long-term water conservation behaviours.

Next, to understand the water consumption patterns of different groups of households, a clustering approach is presented. Extensive experiments were conducted to gain a better understanding of consumption patterns and determine the optimal number of clusters in terms of clustering quality and performance. Two different data sets which were generated from the aforementioned profiling approach were used in this experiment. For the EF data set, the

optimal number of clusters was determined using the elbow method, and several clustering techniques were applied to determine the most suitable technique. However, for the TUWPU data set, the most suitable profiling interval and linkage method was identified in terms of cluster quality metrics. Moreover, water consumption patterns are also discussed based on the findings from the approach.

Later, a generic model is proposed to integrate RSs in behaviour change intervention studies with an application in water conservation. Although the applications of RSs in behaviour change interventions are increasingly common in recent years, there is no model that can be generalised. Since RSs and behaviour change interventions are two different disciplines, a generic model can reduce the time and effort required to design such a study for students, PhD researchers, and practitioners. The proposed model addresses this research gap. Based on the proposed model, a prototype of an RS was developed, and attitudes towards the prototype of the personalised recommendations were evaluated through a survey questionnaire.

Finally, based on the findings from the research, future research directions are suggested, and recommendations are made that have implications for intelligent water conservation programs.

## Acknowledgements

I would like to thank my primary supervisor Professor Michael Blumenstein and co-supervisor Professor Damien Giurco for their continuous support, patience, trust, and feedback. Thank you for giving me the opportunity to work with you and for believing in me and in my work. No words are enough to express my gratitude.

I would also like to thank the partner investigators Professor Rodney Stewart and Dr Khoi Nguyen from Griffith University for their valuable input, support, feedback, and encouragement.

Dr Tanvir Ahmed, Dr Mohammad Saqib, Dr Nabin Sharma, and Md Rafiqul Islam, thank you for your suggestions which made this journey easier.

My sincere gratitude to my teachers from American International University-Bangladesh especially Mr Mashiour Rahman, Dr Mohammad Saiedur Rahaman, Mr A.G.M. Zaman, and Mr Al Imran for helping me to complete my study during a difficult time and giving me life-changing advice. Also, I would like to express my gratitude to Dr Tabin Hasan, Dr Saddam Hossain and Dr Dip Nandi for motivating me to pursue research.

I was lucky to work and collaborate with some outstanding colleagues who always motivated me to do something better. Thank you AZM Ehtesham Chowdhury, Shovra Das, Hasib Hasan, Dr Andrew Levula, and A B M Mehedi Hasan. Also, thanks to Michele Mooney for proofreading this thesis.

I also would like to thank my friends, Irfan Alam Nabil, Amna Islam Badhan, Hridoy Ahmed, Tahmina Akhter, and Umayer Reza for helping me, inspiring me and for their constant belief in me.

Finally, and most importantly, I would like to thank my family for their sacrifice, patience, and continuous encouragement. Special thanks to my wife Ananya Rahman for supporting me in every step. I am indebted to my mother and brother as I was not able to fully shoulder my responsibilities due to my study. Thank you for your patience and for believing in me.

## List of Publications

This thesis comprises a series of published and to-be-published articles together with an exegesis. The list of the publications is as follows:

- I. **Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. 2020, 'Machine Learning and Data Analytic Techniques in Digital Water Metering: A Review', *Water*, vol. 12, no. 1.
- II. **Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. 2021, 'Advanced household profiling using digital water meters', *Journal of Environmental Management*, vol. 288, p. 112377.
- III. **Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Ahmed, T., Giurco, D. & Blumenstein, M. 2021 (in-press), 'A Clustering Solution for Analyzing Residential Water Consumption Patterns', *Knowledge-Based Systems*, Elsevier.
- IV. **Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. ' Recommender systems in Behaviour Change Interventions: A generalized model with application in water conservation' [Draft in preparation. Will be submitted to the journal *Computers in Human Behavior*, Elsevier]

Other relevant publications developed but not included in this thesis include:

- I. **Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. 2019, 'Predicting Household Water Consumption Events: Towards a Personalised Recommender System to Encourage Water-conscious Behaviour', pp. 1-8. [peer reviewed conference paper]
- II. Nguyen, K.A., Stewart, R., Zhang, H., Jones, C., Siriwardene, N., Brown, A., Radion, A., Crook, J., Stevens, M., Smith, N., Giurco, D., Blumenstein, M. & **Rahim, M.S.** 2019, 'Developing a next generation machine learning system for enhanced urban water management: Autoflow', paper presented to the Ozwater19, Melbourne, Australia, 7-9 May 2019. [peer reviewed conference paper]

## Table of Contents

Abstract.....	iii
Acknowledgements.....	v
List of Publications .....	vi
Co-authors' statements.....	viii
1. Introduction.....	1
1.1 Research background .....	1
1.1.1 Water scarcity and demand management .....	1
1.1.2 Digital water meters and residential water end-use .....	2
1.1.3 Digital Water Metering: past, present, and future.....	3
1.2 Research aim .....	7
1.3 Research questions .....	8
1.4 Thesis structure .....	9
2. Overview of the publications .....	10
Block A: Paper I.....	11
I. Machine Learning and Data Analytic Techniques in Digital Water Metering: A Review .	11
Block B: Paper II .....	14
II. Advanced household profiling using digital water meters .....	14
Block C: Paper III .....	16
III. 'A Clustering Solution for Analyzing Residential Water Consumption Patterns' .....	16
Block D: Paper IV.....	19
IV. Recommender systems in Behaviour Change Interventions: A generalized model with application in water conservation.....	19
3. Discussion of research findings .....	21
3.1 Conclusions and Recommendations .....	21
4. Publications .....	26
4.1 Paper I (Block A).....	27
4.2 Paper II (Block B).....	57
4.3 Paper III (Block C).....	80
4.4 Paper IV (Block D) .....	108
Exegesis References.....	130
Appendix A Information sheet and consent form for online survey.....	134
Appendix B Survey Questionnaire.....	136
Appendix C Copyright release for published material .....	136

Co-authors' statements

**Publication I**

I, Md Shamsur Rahim, contributed 75% to the paper entitled:

**Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. 2020, 'Machine Learning and Data Analytic Techniques in Digital Water Metering: A Review', Water, vol. 12, no. 1.

Signature of Candidate:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
-------------------------	--	------------------

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Co-Author 1. Khoi Nguyen

Signature:	Production Note: Signature removed prior to publication.	Date: 15/06/2021
------------	--	------------------

Co-Author 2. Rodney Stewart

Signature:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
------------	--	------------------

Co-Author 3. Damien Giurco

Signature:	Production Note: Signature removed prior to publication.	Date: 15/6/21
------------	--	---------------

Co-Author 4. Michael Blumenstein

Signature:	Production Note: Signature removed prior to publication.	Date: 14/06/21
------------	--	----------------



## Publication II

I, Md Shamsur Rahim, contributed 75% to the paper entitled:

**Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. 2021, 'Advanced household profiling using digital water meters', Journal of Environmental Management, vol. 288, p. 112377.

Signature of Candidate:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
-------------------------	--	------------------

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Co-Author 1. Khoi Nguyen

Signature:	Production Note: Signature removed prior to publication.	Date: 15/06/2021
------------	--	------------------

Co-Author 2. Rodney Stewart

Signature:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
------------	--	------------------

Co-Author 3. Damien Giurco

Signature:	Production Note: Signature removed prior to publication.	Date: 15/6/21
------------	--	---------------

Co-Author 4. Michael Blumenstein

Signature:	Production Note: Signature removed prior to publication.	Date: 14/06/21
------------	--	----------------

### Publication III

I, Md Shamsur Rahim, contributed 70% to the paper entitled:

**Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Ahmed, T., Giurco, D. & Blumenstein, M. 2021 (under revision), 'A Clustering Solution for Analyzing Residential Water Consumption Patterns', Knowledge-Based Systems, Elsevier.

Signature of Candidate:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
-------------------------	--	------------------

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Co-Author 1. Khoi Nguyen

Signature:	Production Note: Signature removed prior to publication.	Date: 15/06/2021
------------	--	------------------

Co-Author 2. Rodney Stewart

Signature:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
------------	--	------------------

Co-Author 3. Tanvir Ahmed

Signature:	Production Note: Signature removed prior to publication.	Date: 15/06/2021
------------	--	------------------

Co-Author 4. Damien Giurco

Signature:	Production Note: Signature removed prior to publication.	Date: 15/6/21
------------	--	---------------

Co-Author 5. Michael Blumenstein

Signature:	Production Note: Signature removed prior to publication.	Date: 14/06/21
------------	--	----------------

## Publication IV

I, Md Shamsur Rahim, contributed 75% to the paper entitled:

**Rahim, M.S.**, Nguyen, K.A., Stewart, R.A., Giurco, D. & Blumenstein, M. ' Recommender systems in Behaviour Change Interventions: A generalized model with application in water conservation' [Draft in preparation. Will be submitted to the journal Computers in Human Behavior, Elsevier].

Signature of Candidate:	Production Note: Signature removed prior to publication.	Date: 03/06/2021
-------------------------	--	------------------

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Co-Author 1. Khoi Nguyen

Date: 15/06/2021

Signature:	Production Note: Signature removed prior to publication.
------------	--

Co-Author 2. Rodney Stewart

Date: 03/06/2021

Signature:	Production Note: Signature removed prior to publication.
------------	--

Co-Author 3. Damien Giurco

Date: 15/6/21

Signature:	Production Note: Signature removed prior to publication.
------------	--

Co-Author 4. Michael Blumenstein

Date: 14/06/21

Signature:	Production Note: Signature removed prior to publication.
------------	--