



the smart H2O project

A European project on water sustainability

ABM SIMULATOR FOR WATER CONSUMPTION

A first prototype

SmartH2O

Project FP7-ICT-619172

Deliverable D3.3 WP3

Deliverable
Version 1.0 – 30 September 2015
Document. ref.:
D3.3.SUPSI.WP3.V1.0

Programme Name: ICT
Project Number: 619172
Project Title: Smarth2O
Partners: Coordinator: SUPSI
Contractors: POLMI, UoM, SETMOB, EIPCM,
TWUL, SES, MOONSUB

Document Number: smarth2o.D3.3.SUPSI.WP3.V1.0
Work-Package: WP3
Deliverable Type: Document
Contractual Date of Delivery: 30 September 2015
Actual Date of Delivery: 30 September 2015
Title of Document: ABM simulator for water consumption: a first
prototype
Author(s): Alessandro Facchini, Andrea Emilio Rizzoli,
Cristina Rottondi, Corrado Valeri, Andrea
Castelletti, Andrea Cominola, Matteo Giuliani,
Simona Denaro, Ahmed Alsahaf, J. Novak,
Paraskevi Lazaridou.

Approval of this report Approved by the Project Coordinator

Summary of this report: D3.3 A first prototype of ABM simulator for
water consumption

History: See document history

Keyword List: Agent-based modelling, water-consumption
prediction, simulation

Availability This report is restricted



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

This work is partially funded by the EU under grant ICT-FP7-619172

Document History

Version	Date	Reason	Revised by
0.1	4/09/2015	Initial table of content	A. Facchini
0.2	28/09/2015	First draft of the deliverable	A. Facchini, A.E. Rizzoli, C. Rottondi, C. Valeri
0.3	29/09/2015	First Review	A. Cominola, S. Denaro, A. Facchini
0.4	30/09/2015	Second Review	M. Giuliani, A. Alshahaf
1.0	30/09/2015	Final revision	A.E. Rizzoli

Disclaimer

This document contains confidential information in the form of the SmarH2O project findings, work and products and its use is strictly regulated by the SmarH2O Consortium Agreement and by Contract no. FP7- ICT-619172.

Neither the SmarH2O Consortium nor any of its officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7-ICT-2013-11) under grant agreement n° 619172.

The contents of this document are the sole responsibility of the SmarH2O consortium and can in no way be taken to reflect the views of the European Union.



Table of Contents

EXECUTIVE SUMMARY	1
1. INTRODUCTION	2
2. TOWARDS AN AGENT BASED MODEL FOR WATER CONSUMPTION: THE STATE OF THE ART	3
2.1 WHAT IS AN AGENT-BASED MODEL?	3
2.2 A SHORT REVIEW OF THE MAIN EXISTING AGENT-BASED MODELS FOR WATER CONSUMPTION	4
2.2.1 <i>The FIRMA Thames model</i>	4
2.2.2 <i>The FIRMABAR model: (1) the case of Barcelona</i>	4
2.2.3 <i>The FIRMABAR model: (2) the case of Valladolid</i>	5
2.2.4 <i>DANUBIA</i>	6
2.2.5 <i>“MEME” model</i>	7
2.2.6 <i>RWUM</i>	7
2.2.7 <i>DAWN</i>	8
2.2.8 <i>Single Household Model</i>	9
3. THE SH2O_D33 MODEL: A USER MANUAL	12
3.1 PRELIMINARY COMMENTS	12
3.2 DATA PRE-PROCESSING AND THE CHOICE OF ATTRIBUTES	12
3.3 MODEL STRUCTURE	14
3.4 CHARACTERISTICS	19
3.5 SIMULATION	20
3.6 VALIDATION	23
4. CONCLUSION AND FOLLOW-UP	26
5. REFERENCES	27

Executive Summary

This document accompanies **Deliverable D3.3, ABM Simulator for water consumption: A first prototype**, which is a **Software Deliverable**.

In this document, after reviewing the main existing ABM simulators, we present a first agent-based simulator for predicting household's water consumption based on past consumption data. In the present version social influence mechanisms are not taken into account. More specifically, we describe an ABM whose characteristics and parameters implement a model of user behaviour based on past water consumption data.

Because of the scarcity of the available measurements, and in particular of the absence of any behavioural data, the decision making process formulated in the current prototype model is constructed to implement only the empirical characteristics of the single water user's data. The water consumption behaviour is hence represented as a parameterized function depending on the empirical characteristics of each households, and does not take into account social phenomena or any influence mechanism.

Despite those limitations, we however verify that the current prototype is showing potential to effectively capture aggregated urban water demand variability with respect to meteorological conditions.

1. Introduction

In the WP2 deliverable D2.2 “Final requirements”, it has been underlined that an agent-based customer consumption simulator will provide utilities with the means to simulate customer behaviour”.

Such behaviour would provide water consumption estimations based on household’s and customers’ information and characteristics, past real-time consumption data, but also past data concerning responsiveness to different kinds of incentives/mimics mechanisms.

In the previous WP3 deliverable D3.2, a feature extraction method for identifying the main driving characteristics of household that are influencing water consumption, based on water consumption and various socio-psychographic and external factors, was proposed. Thence, a first model to predict the consumption profile of water users as a function of the selected variables resulting from solving the selection problem has been advanced.

At this stage, the water user was modelled as an isolated, autonomous agent. No interaction with other users, no external influence of consumption, like prices or other influence mechanism, are considered.

In this deliverable we make a first step in the aim of developing an agent-based simulator for predicting household’s water consumption based on past consumption data.

Because of the scarcity of the available measurements, and in particular of the absence of any behavioural data, the decision making process formulated in the current prototype model is constructed to implement only the empirical characteristics of the single water user’s data. It thence uses a water consumption function parameterised by the empirical characteristics of each households, and does not take into account social phenomena and any influence mechanism.

The deliverable is organised as follows.

in Section 2, after introducing agent-based modelling, we review the main existing agent-based models for simulating water consumption;

in Section 3 we describe our prototype, both its structure and its use. The model is constructed in three steps. First, a pre-processing of the available data is performed. Then a single user behavioural model outputs the characteristics and parameters to be used. Finally, the isolated attributes are incorporated in the prototype. Because of the scarcity of available data, only a small part of the Smart H2O dataset has been used to validate the proposed model.

2. Towards an agent based model for water consumption: the state of the art

2.1 What is an agent-based model?

Agent-based modelling is a “computational method that enables a research to create, analyse and experiment with models composed of agents that interact with an environment” [Gilb2008]. This means that an agent-based model (ABM for short)

- is a computational model, and as such, it is formulated as a computer program that, based on explicit set of rules, from a list of inputs it provides a list of outputs; an agent’s behaviour, and thence its interaction with other agents, is regulated by a set of rules specified by the researcher,
- it enables to treat a system in isolation and observe what happens thorough experiments,
- as a model, it represents some real existing phenomenon, could it be natural, physical, or social,
- it is composed of computational objects - agents - (formally: separated pieces of program), which by passing each other informational messages interact within an environment, i.e. a virtual worlds represented by some mathematical structure (e.g. a grid)

This characteristics make the ABM approach very appealing and suitable to study complex systems and the emergence of self-organisation and shared behaviours [Bali2011]. Agent based modelling has thence found numerous important applications in areas as diverse as physical, biological, social and management sciences (See e.g. [Gilb2008] and [Bali2011] for many references and examples).

From what above, an ABM is characterised by three main building blocks guiding the design of a model: a set of agents, an interaction module and an environment.

As we previously mentioned, an ABM starts with the specification of a set of objects, called *agents*, which, in analogy with game theory, can be seen as players in a multi-player game. An agent can represent a consumer, a voting citizen, a governmental institution, a distributor, a newspaper, etc. What matters, is that an agent is specified by a vector of attributes¹, or characteristics, and behaviours. To instantiate a model, a value is assigned to the agent for each attribute. The value of an attribute may change over time. We refer to the attribute values of an agent (at a moment t) as the agent’s state (at time t). Behaviours are instantiated by rules that tell the agents what to do. Behavioural rules depend on the current agent’s state. Analogously to the notion of strategy from game theory, they may also depend on the history of agent’s past states (history dependency) and take into account the potential behaviours of other agents.

A critical aspect of ABM is that it models agents’ relationships. This means that one has to define who is in relation with whom, and how this relation takes place, the dynamics of the interaction. That is, one has to define the so called *interaction module*. Those aspects are captured by specifying a topology of connectedness.

Connectedness can be provided in several ways. One can for instance specify a network structure, where nodes are agents and links are relationships. One can also specifically locate agents on an *environment*, e.g. a spatial grids or a GIS map. Then by choosing an appropriate metric, it is possible to define for each agent the neighbourhood within she can spatially interact with other agents.

The specification of an environment underlines the fact that in an ABM usually one of the attributes of an agent is her location. This, as for other dynamic attributes, can change over

¹ For instance, if we want to study the diffusion of a new durable product, for what concerns attributes, consumers can be characterised in terms of age, education, income, the sensitivity to a product features, price and promotion,

time, opening thence the possibility for an agent not only to interact within an environment, but also with the environment².

2.2 A short review of the main existing agent-based models for water consumption

In the WP7 Deliverable D7.1, we provided a list of previous studies where the ABM paradigm has been employed to model residential water demand, reported here in Table 1. They are briefly reviewed in the next subsections.

For each model to be discussed, we present its general structure and its characteristics.

2.2.1 The FIRMA Thames model

This ABM was developed within the European project on Freshwater Integrated A Resource Management with Agents (FIRMA). It concerns the south region of England, and the particular policy issue it addresses is the effectiveness of exhortation by the authorities in managing domestic water use during periods when drought prevail. It has been presented in [Bart2001,Down2000].

A. General model structure

There are two versions of the model, the latter being a fine-grained extension of the former based on the feedback received from stakeholders.

Both models implement:

- a hydrological model, whose purpose is to simulate the occurrence of drought conditions on the basis of real precipitation and temperature data
- a social model, representing the household agents and their social network
- a policy agent, representing the policy authorities and who determines restrictive exhortations when drought conditions prevail.

B. Characteristics

Agent households are distributed at random on a grid, with no more than one agent on a cell. Each agent had a horizon of four cells, and have access to the water consumptions of the agents accessible within her horizon (neighbourhood).

In the first model, there are two kinds of abstract water consumption events: private and public consumption events. Private events are only visible to the consuming agent. Public events are also visible to her neighbourhood. Except for that, there are no further distinctions between consumption events.

Each household determines its frequency of water consumption events and the quantity per event. This is done thorough a decision-making process based on endorsement schemes, meaning that an agent chooses between alternative behaviours according to social pressure, evidence and exhortation by public authorities.

The second model extends the first one by disaggregating the domestic water consumption and replacing random frequencies and volumes with actual statistical data provided by stakeholders about the micro-components of water consumption. This extension enabled the modellers in particular to take into account the effects of technological changes.

2.2.2 The FIRMABAR model: (1) the case of Barcelona

As the previous one, this model has been originally developed within the FIRMA project. It

² Consider for instance an ABM modelling the traffic of a city. The environment can clearly affect the space of actions at disposal to an agent.

has been applied to the Metropolitan region of Barcelona [Lope2005]. It is an integrated tool, meaning that it does not separate the social processes behind the concerned decision making behaviour with the economic, demographic, cultural and spatial sub-processes. It thence allows “to simulate and evaluate alternative supply and demand policies under different climatic and technological scenarios.”

A. General model structure

The model is based on the assumption that “urban dynamics play a central role in the collective behaviour”, and thence in the water consumption behaviour too. The urban space where agents live is represented as a grid, and the associated dynamics is based on cellular automata running over such a structure. The agent model involves the following actors:

- the families: they incorporate life styles and social attitudes. Each family is initially endowed with social attributes (e.g. social class, income) that can evolve during the simulation, and play their own life cycle (born, grow up and die). They compute their maximum expected water demand and real consumption.
- the real estate companies: they represent the actors determining the changes in the territorial model. They build new housing around the municipalities and act as intermediaries for the second hand market.

B. Characteristics

Families have to take two types of decisions: house movement and water consumption. The former decision depends on parameters such as the size of the house vs the size of the family, social class in the neighbourhood and the evolution of the prices for new houses vs the second-hand market. The latter depends essentially on the water price, the gross income, the maximum demanded water and the evolution of the reservoirs in the region. Besides these factors, it is also affected by the housing type and the appliances technology, the social attitude when the region is in drought or scarcity periods, the social class and the size of the family. Finally, the water consumption is adapted to the local habits by a mechanism of imitation within the closest neighbourhood.

2.2.3 The FIRMABAR model: (2) the case of Valladolid

This model is a variation of the previous version of the previous FIRMABAR model. It has been applied to the Metropolitan region of Valladolid [Gala2009]. In this case the tool integrates social sub-models with models of urban dynamics, water consumption, and technological and opinion diffusion in an agent-based model that is, in turn, linked with a geographic information system (GIS).

A. General model structure

Two types of entities are represented in the model:

- the environment: it is a computational entity imported from a vectorial GIS, where every block with dwellings in a geographical area is explicitly represented and characterised by their spatial and economic features. The second layer corresponds to the spatial distribution of the computational agents in the model.
- the agents, which correspond to the families.

B. Characteristics

The location and various other features of the families in the spatial environment are retrieved from socioeconomic geo-referenced databases of the region.

The variables influencing water consumption demand of families is captured by the following sub-models:

- The urban dynamics sub-model simulates the migratory movement of the families in the metropolitan space on the basis of their socioeconomic factors.
- a diffusion model simulates the process of embracing a certain consumption

behaviour because of the influence by the number of neighbouring adopters (hence giving rise to positive reinforcements).

- A Bass' model, coupled with a behaviour diffusion model, is finally used to reflect the influence for the water demand of the adoption and diffusion of new technological devices.
- Finally the consumption of the population has been characterised by a linear regression model derived from the databases of the supplier company in the region of study. The variables considered are nationality, average age of the family, education level, typology of residences and household density. The residences are classified in five typologies: 1, habitable with both commercial and domestic uses; 2, houses with gardens; 3, houses without garden; 4, old houses; and 5, apartments and flats. There is no statistical difference in consumption between types 3 and 5. Price has not been considered as a parameter in the study because water prices (fixed by the water supplier) in the region have been very stable in the past years.

2.2.4 DANUBIA

The integrated decision support system DANUBIA [Erns2005], developed within the interdisciplinary project GLOWA-Danube, simulates water resources management in the area of the Upper Danube Basin. The Actors component of the system comprises implementation of the Household, Demography, Economy, Water Supply and Tourism models. There, agent-based modelling plays a central role, and specifically within the Household actor model, which explores drinking water use of households under global warming conditions.

A. General model structure

The Household actor model incorporates survey data concerning habits of drinking water and statistical data about water demand and demographics. Those data are used to configure 25 household types, according to income, size of the family, use of water consuming devices, social class etc.

At first, a shallow model is constructed to implement the empirical characteristics of water use coming from survey data. It uses a simple computation of water demand, differentiated for each household type and does not take into account social phenomena. The shallow model is thence only capable of estimating future demand by extrapolating current trends. Next, the shallow model is replaced with a so-called deep model. The latter reconstructs the process aspects of decision making when adopting water saver technological innovations or changing one's habits. This is done by associating to each household types a core set of attributes, preferences, possible plans comprising different actions to respond to changing conditions (the set of plans is common to all types).

B. Characteristics

In the shallow model, the computation involves several factors, such as:

- consumptions for ten types of water use based on the empirical data seasonal variations,
- changes in water use due to price,
- influence of the population density, as suggested by the empirical data a baseline of water-saving technology introduction.

In the deep model, the deliberation process of households involves the following steps:

- perception of the physical, social and legal environment;
- selection of a plan of action, defining various water uses and their intensity;
- calculation of the subjective expected utility associated with the plan;
- execution of the plan and export of its consequences to the environment.

In the deep mode, a parameter is used to account for the different behaviours concerning the adoption of innovative water-saving technology. The diffusion of the technology is modelled as a communication and imitation process among peers in a network.

2.2.5 “MEME” model

In [Rixo2006], two models are presented. They both aim at exploring some aggregated properties relating to water use behaviour that emerge as effects of social interaction. The first model focuses on social networks and tariff structures. The second employs a mimetic framework capturing imitation of technology adoption and water use behaviour.

A. General model structure

In the first model, (aggregated) water consumption is controlled by means of tariff schemes: either fixed (flat daily cost for water, independent of use) or variable (charge dependent on the aggregated amount of water used by resident). The model has two types of agent:

- user (resident) agents: they are grouped in social networks (randomly generated or reconstructed from real data),
- a water utility agent, who imposes the water tariff schemes.

The second model has the same type of agents: user (resident) agents and the water utility agent, the main focus being on the former. Resident agents are assumed to consume water according to a set of water memes, some of which are water saving memes and some other are not and suggest indifferent behaviour.

B. Characteristics

In the first model, when a resident agent is unable to get the desired amount of water due to high prices, she becomes water stressed. This water stress leads the agent to seek for agents in the network who are using more water and to place peer pressure on them to reduce water use. On the other hand, the water utility agent modifies water tariffs (they are increased or decreased) depending on the acceptable degree of water stress within the resident population, and on the water levels, i.e. whether or not they have reached critical levels.

For what concerns the MEME model, each of the water memes represents a direct mapping between the type of device, the frequency of a water consumption event and the amount of water used. Moreover, it has a cost associated with imitation, implemented as a certain required degree of belief in water saving in order for a user agent to implement (adopt) a certain meme.

User agents are born with a degree of belief in water saving. This is thence updated dynamically according to the percentage of water saving memes in the agent's current meme set. A degree of water worry is also defined as an inverse function of the current dam level's capacity. Water worry is scaled by the resident agent's degree of belief in water saving. This mirrors the fact that users tend to care more about the state of water reserves if they have a strong belief in water saving.

Water memes are copied based on social interaction. Such interactions follow some rules:

- users within the same household imitate each other,
- water worried agents seek out water savers in their friendship networks,
- any agent (disregarding the associated beliefs) tends to become more like the agents in her networks, independently of whether the latter are water savers or water utilitarians.

In the MEME model, the role of the water utility agent is simply to monitor the current water levels and, when they reach critical levels, to send out a random water-saving meme with the bill. Users who are water worried are thence able to copy and implement this meme.

2.2.6 RWUM

In [Chu2009], an agent-based Residential Water Use Model (RWUM) is developed and calibrated against residential water use data in Beijing. By disaggregating total water

demands down to constituent end-uses, this model evaluated heterogeneous consumer responses on water, taking into account the factors of market penetration of water-saving technologies, regulatory policies, economic development, as well as social consciousness and preferences.

A. General model structure

The RWUM model consists of three agents:

- the regulator agent: it is responsible for establishing the structure and level of water prices (including water resources, water supply and wastewater treatment) based on factors such as costs of water supply, historical water pricing and affordability for house-holds. It can also promote the practice of water conservation and water efficiency, e.g. by providing new, and more efficient, water devices to certain households free-of-charge;
- the water appliance market agent: in addition to physically providing households with their chosen water device, it collects and synthesises information for households, such as water use efficiency, costs, life-span and availability of each type of water appliance
- the household agent: it consists in a set of attributes and a water use behaviour, which may sometimes also change during the simulation

B. Characteristics

The implement process goes as follows:

- in the initial year, agents of households, regulator and water appliance market are created and initialised;
- the ownership and technical properties of water use devices and water use habits among households are done such that those match a real distribution in Beijing in that year. The market penetration rate is distributed to all households in a random way. Water price of the initial year is set by the regulator agent.
- Households use water according their water use devices and habits, and the output is the amount of water total households use.
- In the cycle year, new household agents are created and initialised with population increase in the city. The water price of each year is decided by the regulator agent. Water appliance market agent provides the information on water devices in the market for households to choose.
- Households use water according to their device characteristics, income and price level and other information to make decisions on water device purchase, accelerated replace, water use frequency change or wastewater adoption and form total household water use in the new year.

2.2.7 DAWN

DAWN [Atha2005] (also used in [Tseg2009]) is a simulation tool for evaluating water-pricing policies in urban areas. More specifically, it aims at estimating the water consumption under different scenarios of pricing policies, explicitly taking into account social interaction, the propagation of water conservation signals among individual consumers, and responsiveness to water conservation policies.

It is a hybrid approach, in the sense that it integrates a conventional econometric model with an agent-based social simulation model.

A. General model structure

The agent-based model simulates the water demand-supply cycle. It includes the following agent types:

- a Simulator Agent (SA): it is used to moderate and synchronise the simulation

procedure (and is also responsible for capturing a user-defined scenario);

- a Water Supplier Agent (WSA): it represents the water utility,
- the Consumer Agents (CA), each representing a single consumer or a consumer group having common needs,
- a Meteorologist Agent (MOA): it supplies the system with meteorological conditions.

The water-demand cycle involves the following sequence of interactions.

1. First of all, WSA initiates a water pricing policy, according to a user-defined scenario. Thence:
MOA informs WSA on meteorological conditions.
2. WSA asks each CA to determine its own water demand. CAs are also informed about the current meteorological conditions and water price.
3. In response to this, each CA reconsiders its water consumption with respect to water price, social influence, weather conditions, and other fixed parameters.
4. WSA collects all individual demands for water, calculates the total consumption and, if necessary, revises the water-pricing policy.

B. Characteristics

Consumers estimate their consumption using an econometric model, augmented with a social activity model. Social interaction simulates the propagation of water conservation signals within the consumer community. The influence diffusion mechanism works as follows. Consumer agents are distributed randomly on a grid, with no more than one agent on a cell. Each agent had a spatial horizon of four cells, her neighbourhood. Consumers are partitioned within four types (each representing both a certain ability to comprehend and to communicate water conservation signals). According to the type she belongs, each agent is assigned

- a social weight, representing her power of persuasion,
- a diffraction function, that represents her ability to comprehend water conservation signals

By a word-of-mouth mechanism, neighbouring agents communicate each other their social weights. Each CA accumulates all social weights received from its neighbours and adjusts the sum according to her diffraction function. This variable is then incorporated into the generic econometric model used by the agent to estimate her personal consumption for each time interval.

2.2.8 Single Household Model

An agent based model for single household water use has been developed in [Link2013].

The model aims to address the dynamic behaviour of occupants. It simulates hourly water-using activities of household members within a technological and behavioural context. It thence generates data on domestic water consumption pattern both at the individual and household levels.

A. General model structure

The agent-based model simulates the water consumption system within a single household. Agents are thence the occupants of the household. They are associate with a list of attributes supposed to affect their decision making process, like employment, age, being home, water use specific practice characteristics (personal hygiene, cooking and drinking, doing laundry, toilet and washing dishes, etc.). The environment is determined by the technological water infrastructure in place. To each agent is also associated a function encoding her intention to save water. This function depends on her care for the environment, but also on the presence of water saving campaigns and on water meter levels.

B. Characteristics

Once the model is initialised, all agents execute their water use habits. An agent's behaviour is specified by a set of rules determining which actions are executed. For each action (e.g. "need to clean body", "need to eat or drink") there are behavioural plans for which then agent can choose. According to the frequency with which the agent performs each plan, she forms her strategy for that particular practice. Agent's needs and the frequency with which she perform each plan are stochastic.

Table 1: Benchmarks for agent based simulation for modelling residential water demand.

Reference	Title	Region	ABM Model	Data	Simulation period	Validation
<i>Chu2009</i>	Agent-based residential water use behaviour simulation and policy implications: A case-study in Beijing city	Beijing	Residential Water Use Model (RWUM)	Disaggregated (at the end-use level)	1985-2030	$R^2=0.818$, (accuracy of water use capita per day)
<i>Down2000</i>	Understanding Climate Policy Using Participatory Agent-Based Social Simulation	Thames region of England	FIRMA Thames (1st version)	Aggregated (at the end-use level)	1970-1996	no
<i>Bart2001</i>	Policy modelling with ABSS: The case of water demand management	Thames region of England	FIRMA Thames (2nd version)	Disaggregated (at the end-use level)	NA	stakeholders validation
<i>Rixo2006</i>	Exploring Water Conservation Behaviour through Participatory Agent-Based Modelling	-	" MEME model"	Aggregated (at the end-use level)	730 days	no
<i>Atha2005</i>	A Hybrid Agent-Based Model for Estimating Residential Water Demand	Thessaloniki (GR)	DAWN	Aggregated (at the end-use level)	1994-2000	no
<i>Tseg2009</i>	Agent-Based Modeling to Estimate Residential Water Demand and to Explore Optimal Demand Side Water Management strategies	-	DAWN inspired	Aggregated (at the end-use level)	60 months	no

<i>Lope2005</i>	Urban Water management with artificial societies agents: the FIRMABAR simulator	Barcelona (ES)	FIRMABAR (1)	Aggregated	10 years	stakeholders validation
<i>Gala2009</i>	An agent-based model for domestic water management in valladolid metropolitan area	Valladolid (ES)	FIRMABAR (2)	Aggregated	40 quarters	stakeholders validation
<i>Erns2005</i>	Shallow and deep modeling of water use in large, spatially explicit, coupled simulation system	Upper Danube basin	DANUBIA	Disaggregated	2011-2040	no
<i>Link2013</i>	An Agent Based Model of Household Water Use	USA / NL	"Single household model"	Disaggregated	1 year	5-15% aggregated data. much worst when taking into account disaggregated data.

3. The SH2O_D33 model: a user manual

In this section we provide a detailed description of the characteristics and functioning of the first prototype, called SH2O_D33.

3.1 Preliminary comments

The SH2O_D33 model is implemented within the commercial simulation software AnyLogic (www.anylogic.com).

This first prototype is essentially an input-output model. It aims at capturing the influence of household's characteristics on domestic water demand. More specifically, it predicts future water consumption based solely on the specific values associated to the attributes of an agent, its current level of water consumption and the meteorological sensibility. The prototype thence resembles to the DANUBIA shallow model [Erns2005].

At the current stage of the modelling process, social phenomena and price sensibilities are not taken into account and thence have not been incorporated in the prototype.

The reason for the absence of a developed interaction module capturing social influence and other similar interactive behaviours is because, currently, behavioural data have still to be collected and therefore their analysis is not allowed. This is due to the fact that the SmartH2O app is not yet equipped with interactive features between user's groups. Inspired by the approaches adopted in the DAWN model [Atha2005] and in the MEME model [Rixo2006], as well as by the agent-based study of the mechanism underlying the emergence of social norms presented in [Helb2014], we only dispose of some theoretical hypothesis on social interaction and social influence that will be implemented in the next stage of the agent based model.

The submitted prototype can be seen as an ABM implementation of the single-user consumption behaviour model at the household level developed in D3.2. Indeed, recall that the latter is structured as a two-step procedure. Firstly, the most relevant determinants of users' consumption profiles are identified, and secondly a model learning procedure relates the extracted subset of determinants to the water consumption level. In the ABM prototype, the consumption level classifications obtained thorough the model learning procedure is implemented within the Household agents as an explicit attribute and determine the characteristics of the water consumption function.

Notice that, since the water users' database of Tegna does not currently contains explicit values for users' psychographics and houses' characteristics, at the current stage of the ABM's development, the Smart Meter ID and the historical real time consumption measurements are the only data that can be retrieved from the database. Subsequently the consumption level classifications obtained thorough a learning procedure uniquely determine the water consumption characteristics of agents representing the households in the ABM.

3.2 Data pre-processing and the choice of attributes

The SmartH2O database includes water consumption measurements gathered from 255 smart meters installed in Tegna during a period that spans from 16/10/2014 to 31/07/2015. Notice that at the moment the database does not include any additional information related to the households (e.g. number of dwellers, presence of garden/swimming pool, etc.). Therefore, our analysis is based exclusively on the knowledge of the Smart meter IDs and the hourly water consumption measurements.

The following data pre-processing steps have been carried out to determine the relevant

attributes to be used for the calibration of the ABM model:

- records of users showing data inconsistency (e.g. negative consumption measurements) or missing data were removed from the dataset;
- association of meteorological data (temperature, irradiation) and type of the day (week-end vs working day) with each daily profile. The meteorological data concern the period from 16/10/2014 to 31/07/2015 and the region of Locarno (where Tegna is located). They were retrieved from the official website of the *Osservatorio ambientale per la Svizzera Italiana* (www.ti.ch/oasi) ;
- data clustering by means of the k-means algorithm with 10-fold cross validation over daily profiles using the aforementioned meteorological attributes. Determination of five meteorological classes parameterised by minimum and maximum value of temperature, and minimum and maximum value of solar irradiation as described in Table 2.

Table 2: Meteo classes and corresponding parameters.

class ID	min temperature	max temperature	min irradiation	max irradiation
0	15	35	0	240
1	-20	14.9	0	160
2	15	23	241	400
3	-20	14.9	161	400
4	23.1	35	241	400

- daily profiles are divided according to the ten combinations generated by considering the five meteorological classes from Table 2 and the two types of the day (weekday vs weekend).
- for each combination, data clustering by means of a k-means algorithm with 10-fold cross validation was applied to determine 3 daily water consumption classes. For each of the 10 classes we therefore have:
 - a class of low consumption in weekdays
 - a class of low consumption in weekends
 - a class of medium consumption in weekdays
 - a class of medium consumption in weekends
 - a class of high consumption in weekdays
 - a class of high consumption in weekends
- each water consumption class is associated to a probability distribution estimated from the available data to characterise the water consumption levels within each class. This is done by providing a partition of the consumption spectrum of the class in 100 bins, each characterised by a minimum value, a maximum value, and a probability (such values can be consulted at the page “consumption classes” in xls file *results_meteo_classes* in the folder of the prototype).
- For each meter ID, meteorological class and type of day, the probability that the daily consumption of given meter falls in the low/medium/high consumption class has been estimated based on the available data. To do so, only smart meters monitored for a sufficient number of days falling within each meteorological class have been considered.

Therefore, the above described pre-processing defines:

1. the relevant attributes of the households to be specified during the initialisation process: function assigning to each pair (class ID, type of day) the probability for the agent to belong to one of the corresponding three consumption classes;
2. the parameters of the model: the probability distributions of water consumption classes,
3. the structural form of the water consumption function to be specified in the model for the household's agent, as described in the next subsection.

3.3 Model structure

The AnyLogic file SH2O_D33.alp of the model is located in the folder SH2O_D33. The folder also contains:

- *meteo_table.xlsx*: an excel file containing the chosen meteorological data that are going to be used during the simulation,
- *results_meteo_classes.xlsx*: an excel file containing the results on the available data described in the previous subsection, that is: (a) the properties of the obtained 5 meteorological classes ("meteo classes" page), (b) the obtained probability distributions for each of the 30 obtained consumption classes ("consumption classes" page), (c) the obtained single user behavioural model, that is, for each household, the probability distribution associated to each meteorological class ("user classification" page). In addition to that, a page "*results validation*" contains results discussed in section 3.6 about validation.
- Tegna-Google *Maps-Cartina.png*: a png map of the Tegna area,
- Tegna-Google *Maps-Terreno.png*: a satellite png map of the Tegna area
- a folder *x3d*: it contains the households' image files used as presentation shapes during the simulation;
- a folder containing the AnyLogic's database files used in the implementation of the model.

Start AnyLogic 7.2. If the Welcome page displays, close it.

Open the SH2O_D33 model by selecting: *File > Open* and then selecting SH2O_D33.alp located in the SH2O_D33 folder. The model opens.

See box[1]³ in Figure 1: in addition of the *Main* agent acting as the environment for all other agents, the SH2O_D33 model consists of two main agent's types:

- the *Supplier* type,
- the *Household* type,

and other types instrumental for the behaviour of the model: the *ConsumptionRequest* agent type, the *ConsumptionType* agent type and its extensions *ConsumptionTypeHigh*, *ConsumptionTypeMedium* and *ConsumptionTypeLow*.

³ With '[X]' we refer to the fact that what follows is going to describe what is denoted with the number X in a snapshot from the Anylogic view of the model.

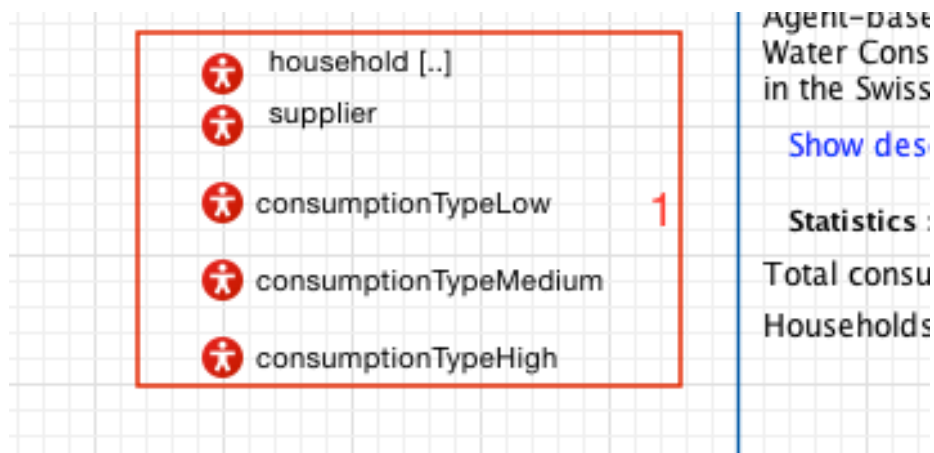


Figure 1. Agent types.

Double-Click on the agent type *Supplier* in the project's tree structure to visualise its characteristics.

The role of the Supplier agent is to be responsible for informing the households at each step of the simulation process about the current meteorological conditions.

Referring to Figure 2:

[2] Information about the meteorological situation (temperature and solar irradiation) are encoded in the database table *meteo_table* visible by clicking the Database node of the project tree, and are accessed by the *Supplier* agent by using the *getDayMeteoClass* method [3].

[4] The *ConsumptionRequest* agent type is linked to the *Supplier* (it "lives" within the *Supplier* agent).

[5] The interaction phase is initiated at each simulation step by the event *StartProcess* specified in the *Supplier's* agent diagram. At the end of the interaction phase, the *Main* agent "recollects" the consumption data from all the households and add them to the statistics.

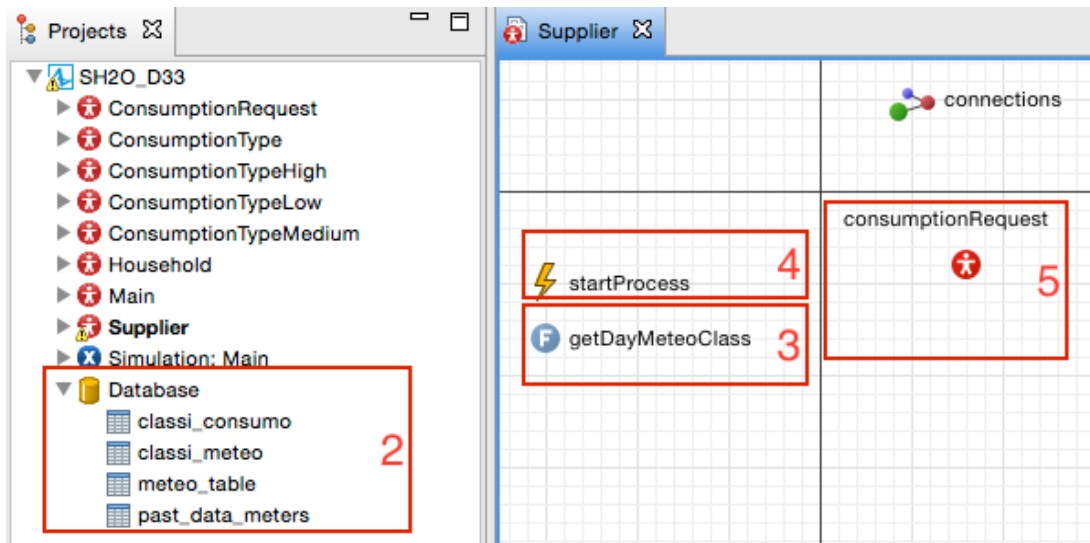


Figure 2. Database connection.

Double-Click on the agent type *ConsumptionRequest* in the project's tree structure to visualise its characteristics. The role of this agent is to be the body of the message containing the current meteorological information, which is sent by the supplier to each household during the interaction phase of the simulation. Formally, this is done by using the variable *meteoClass*, which will store the meteorological class determined by the *getDayMeteoClass* method. The method outputs the meteorological class based on the characteristics corresponding to current time step in the simulation: type of the day (weekend vs weekday), and corresponding temperature and solar irradiation levels taken from the meteorological database. In Figure 3 [6] by clicking on the method, the corresponding *Properties* view open. One can thence visualise the body of the function.

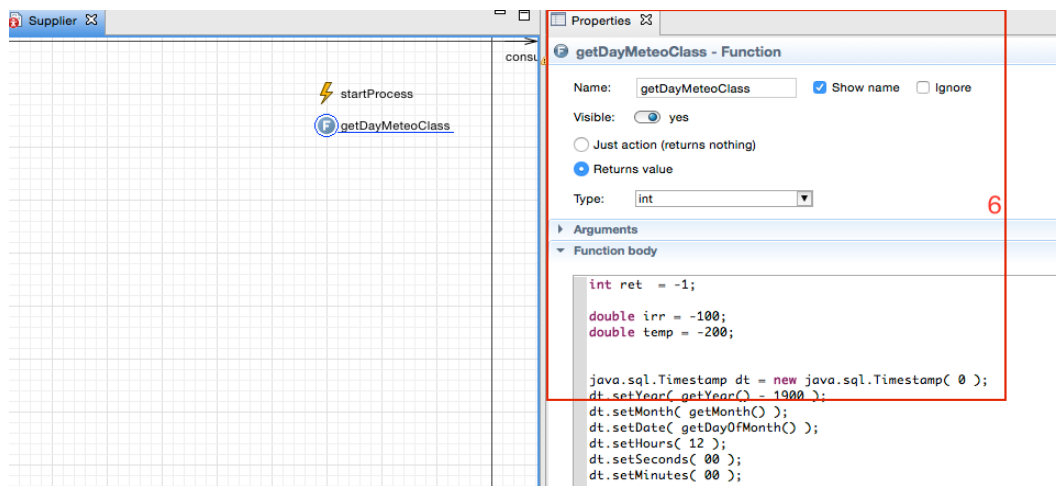


Figure 3. Properties of the getDayMeteoClass function.

Double-Click on the agent type *Household* in the project's tree structure to visualise its

characteristics.

In Figure 4 are displayed the main features of the *Household* agent type in the current model:

- a list of attributes corresponding to information about households, listed in the “information on household” box [7]. Ideally those parameters should correspond to the ones introduced and discussed in the WP3 deliverables D3.1 and D3.2. Since those information have still to be specified within the Smarth2O water users’ database in Tegna, the only available attributes is the Smart meter ID;
- additional parameters, variables and methods are used to store at the moment the model is initialised past average daily consumption characteristics of the users coming from the data pre-processing procedure [8] listed in the “Past consumption data characteristics” box and the consumption’s level and status of the households at each time step of the simulation (the variables *statsConsumption* and *consumptionState*);
- [9] a consumption function *waterConsumption*, whose current definition is discussed in the next subsection;
- a data histogram *dailyConsumption* to collect the consumption levels outputs during the simulation process. The collected results are then displayed in the two graphics in the “Statistic” box: the first graphic displays the distribution among the different consumption levels, while the second elucidate the daily consumption behaviour during the simulation [10];
- [11] a state-chart *computeWaterUse* defining the consumption behaviour at each time step of the simulation (click on each of its components to inspect its properties). This construct is at the core of the interaction module, which is given by the simple message passing mechanism between the supplier and households described in the following subsection;
- [12] the additional variable *meteoDayClass* is used to store the meteorological information received by the Supplier agent;
- [13] three objects represent stylised houses. They are the households’ animation shapes, one for each colour that will indicate the current consumption level (green=low, yellow=medium, red=high). They can also be visualised in the *Main* agent diagram.

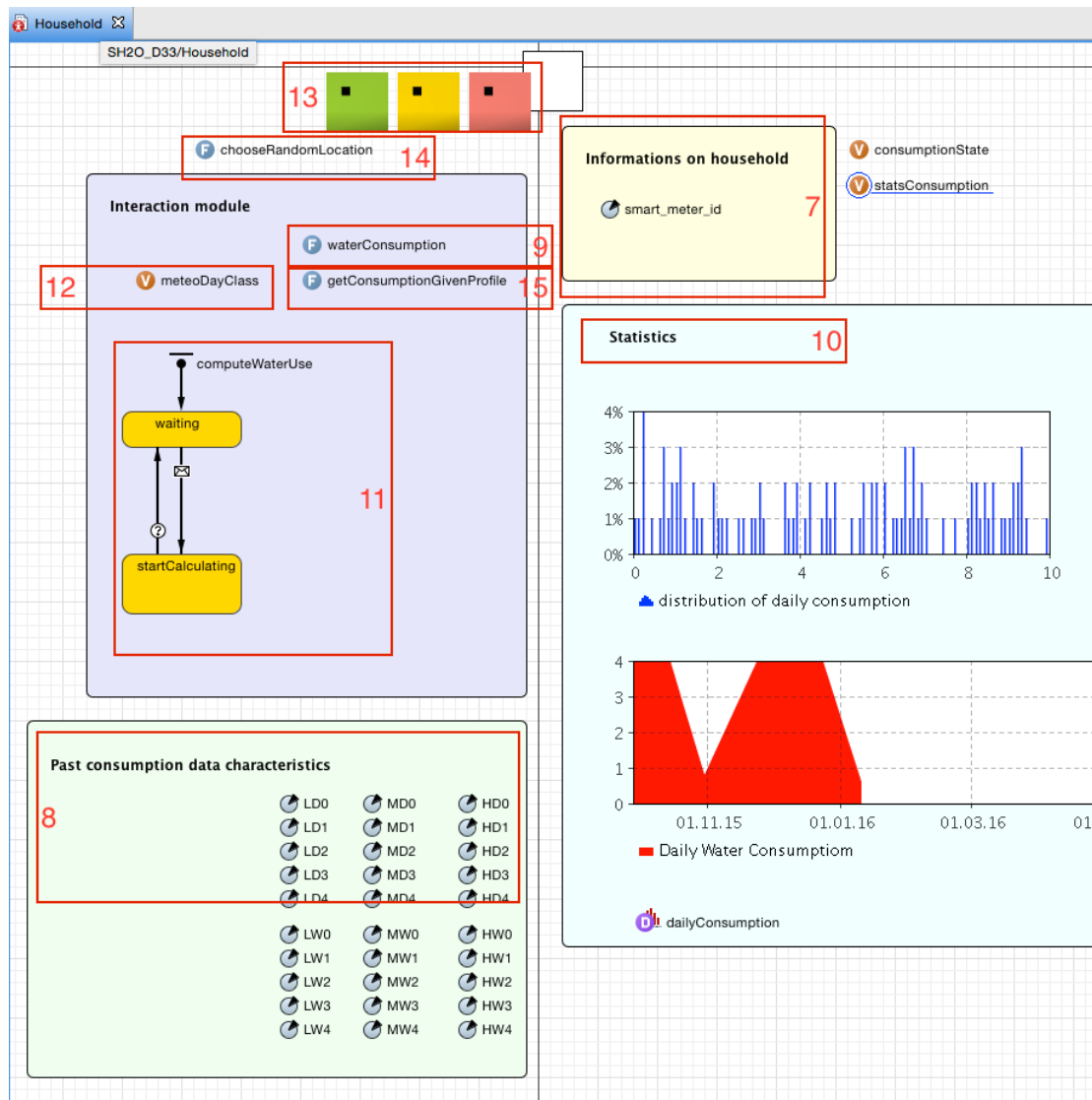


Figure 4. The household agent.

Double-Click on the agent type *ConsumptionType* in the project's tree structure to visualise its characteristics. These consist in the variables *minThreshold* and *maxThreshold*. For each of the three extensions (*ConsumptionTypeHigh*, *ConsumptionTypeMedium* and *ConsumptionTypeLow*), their values can be specified either at the initialisation of the model or manually. They correspond to minimum and maximum of the interval for the chosen average daily consumption classes:

- Low if its value lies between 0 [m³/day] to 0.5 [m³/day],
- Medium if its value lies between 0.5 [m³/day] and 1 [m³/day],
- High otherwise.

The specific environment of the model where households will be located can be visualised in the *Main* agent diagram. It is constituted by a map of Tegna. Households are generated randomly on it at the beginning of the simulation through the *chooseRandomLocation* function defined in the Household agent's diagram [14]. Since there is no interaction between households, neither a topology of connectedness nor a network structure are specified.

On the same diagram, it is possible to notice elements used to visualise statistics collected during the simulation:

- the total consumption of households, and total consumption of the households belonging to a specific consumption level (both measured in m3),
- below the previous data, the chart displaying the percentage of consumption levels among the population,
- under the map of the city of Tegna, a box “Simulation results and statistics”. There are two graphics, two bar charts and a Histogram data. The two graphics represent the behaviour of the daily average water consumption over the total households’ population and the aggregated daily consumption (m3). The data histogram *totAvrgDailyConsumption* collects the aggregated statistics. The two bar charts display the actual average and total water consumptions at the current moment of the simulation process.

On the left parts of the *Main* diagram we have some variables used for the simulation animation (“Animation” box) and to store data for statistics (“Statistics” box), but more important in the “Meteo classes” box we have parameters for each meteorological class from Table 2.

3.4 Characteristics

The simulation is performed as follows:

- at the initial step of the simulation, the supplier and households are created and initialised according to the data included in the database *past_data_meters* (See [2] in Figure 2).
- Households are randomly distributed over the space representing the commune of Tegna the *chooseRandomLocation* method,
- at each simulation step, the time of the model being measured in days, the supplier agent sends a message to all households containing the information about the current meteorological data. This process is initialised by the event *startProcess*
- Based on the received data and on its consumption attributes, each household calculate its own water consumption. This mechanism, evaluated statistically with a random choice of the consumption class and using the empirical probability distributions of the consumption level for each class, takes place thorough the state chart *computeWaterUse*. The water consumption function *waterConsumption* is based on the method *getConsumptionGivenProfile* (see [15] in Figure 4). The underlying logic is the following (to inspect the structure of one of the two functions - *waterConsumption* or *getConsumptionGivenProfile*-, click on the method label to open its *Properties* view):
 - the Household has received the meteorological datum stored in the variable *meteoClass*. Such datum correspond to the meteorological class linked to the current day in the simulation process and is registered in the variable *meteoDayClass*,
 - given the meteorological class stored in *meteoDayClass*, the Household determines thorough the *waterConsumption* method its daily consumption level. Firstly it determines its consumption class. Once the consumption class has been chosen, a consumption level is randomly determined according to the empirical distribution from table *classi_consumo* in the database by calling the *getConsumptionGivenProfile* method.
- The output is added to the statistics (see e.g. Subsection 3.5.C).

3.5 Simulation

A. Settings

Double-click on the SH2O_33 root of the project tree. In the corresponding *Properties* view, set the *Model time unit* as “days”.

Double-click on the Simulation node of the project tree. In the corresponding *Properties* view, it is possible to select the length and period of the simulation under the *Model time* section. Here, we decide to run simulations from 1/1/2014 until 21/9/2015 to cover the period of the available meteorological data (see *meteo_table* in the model database). Any desired period between these two dates can be chosen.

For what concerns the *randomness*, definable under the Randomness section, choose the *Random seed* option.

B. Launching the simulation

To start the simulation, proceeds as follows:

- In Figure 5 [16] locate the Run button and choose SH2O_D33 / Simulation from the list

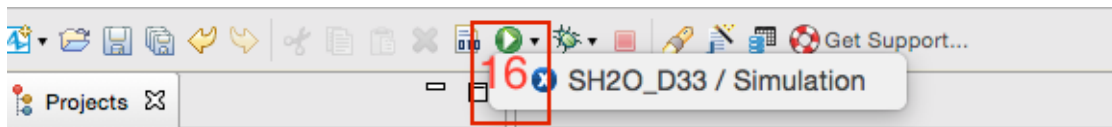


Figure 5. Starting the simulation.

- after you start the model, the presentation window displays the launches experiment *Simulation*.
- Click the “Run” button to run the model (see [17] in Figure 6).

SmartH2O model for water consumption: Tegna (TI - CH)

This is an agent-based water consumption model of the small Swiss city of Tegna. Households are the visible agent displayed on the map of the city. Each household is characterized by a series of parameters based on its past consumption. Such parameters, joint with the meteorological situation, determine the household's water consumption. You can view a household consumption statistics by clicking its image. You can also pick a random household to view its individual state.



Figure 6. The simulation output control panel.

You will see the model's presentation, as defined for *Main* agent

C. Reading the results of the simulation

Once the simulation has finished (or you choose to pause it), the results can be observed as follows:

- the obtained aggregated results are displayed on the root:Main page of the simulation frame in the box "Simulation results and statistics". There are two graphics, two bar charts and a Histogram data. The two graphics represent the behaviour of the daily average water consumption over the total households' population and the aggregated daily consumption (m3). By clicking the data histogram *totAvgDailyConsumption* it is possible to display the aggregated statistics. The two bar charts represent the actual average and total water consumptions at the moment the simulation stops. The latter results can be stored by clicking on the button "Store obtained values". Such values are stored and displayed in the two charts "Total consumption" and "Average consumption" in the main presentation window of the experiment *Simulation* [18]. By starting a new run, one can thus store additional experimental results to be subsequently analysed. The obtained results can be copied by clicking on the corresponding *Copy Results* button [19].
- There are two possibilities to visualise the results for a single households. Firstly, one can simply click on a stylised house figure to visualise the associated basic information: smart meter ID, daily average consumption and corresponding current consumption class (see [20] in Figure 7).

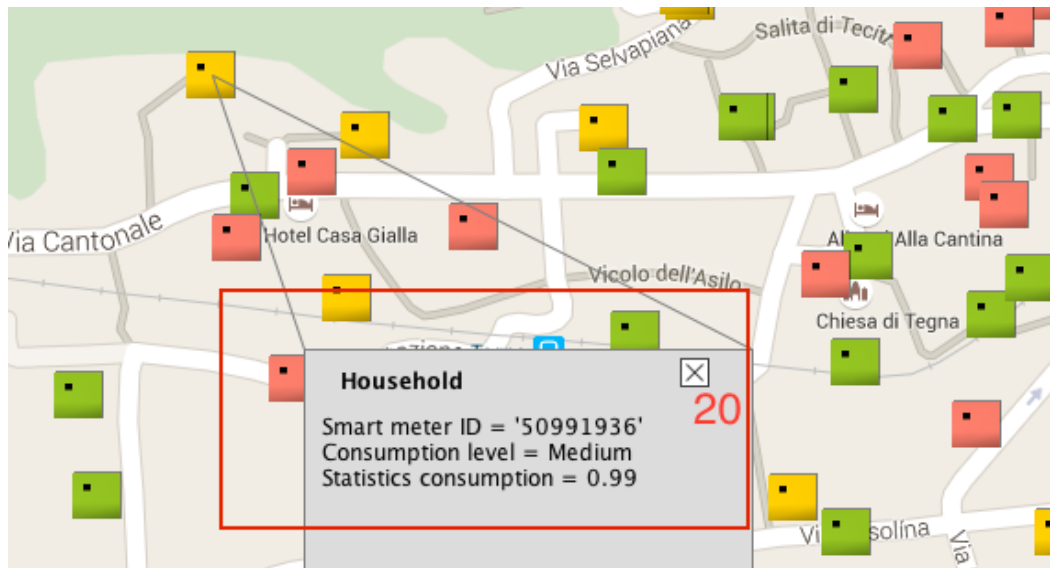


Figure 7. Graphical simulation output.

Secondly, one can access the statistics for a specific households as follows:

- select the node “household” [21] in Figure 8.

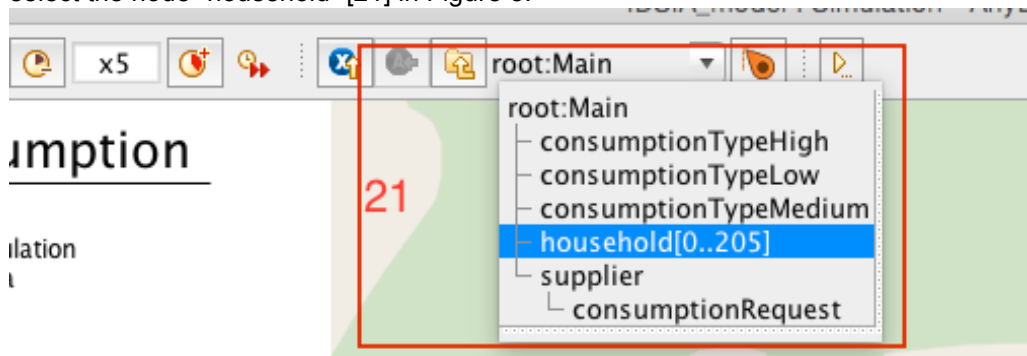


Figure 8. Accessing the simulation data of a specific household.

- choose the household agent instance you want to visualise by writing its number in the text box (see [22] in Figure 9). The correspondence between household agent number and Smart Meter ID is given by the database *past_data_meters*;

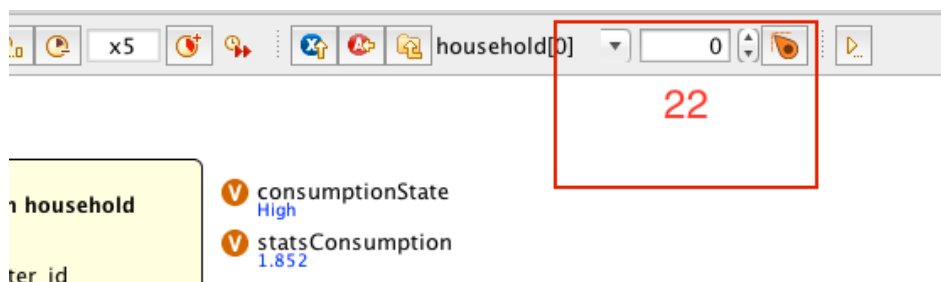


Figure 9. Selecting the household agent.

- the statistics box in the displayed household["chosen number"] page presents two graphic: a graphic representing the behaviour of the daily average water consumption of the agent, and the distribution of its daily water consumption. By clicking the data histogram *dailyConsumption* it is possible to display the statistics [23] in Figure 10.

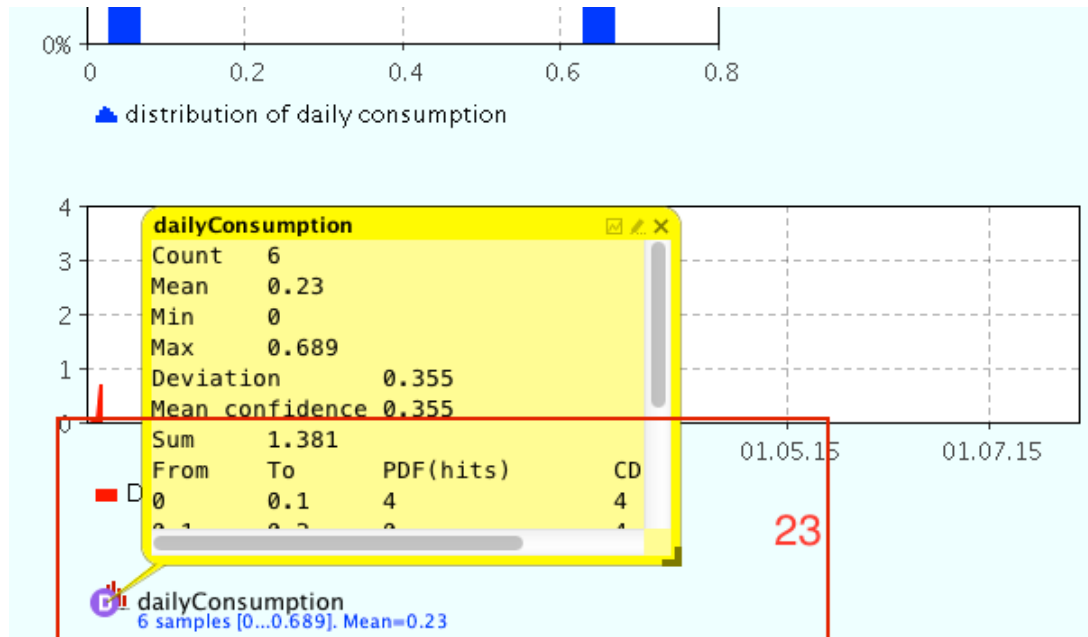


Figure 10. Displaying the household statistics.

3.6 Validation

We aim at identifying an agent based model that, given as input the extracted characteristics based on the past consumption measurements, returns as output the aggregated predicted water consumption of all households.

The available do not include any psychographic characteristics as the Smarth2O platform has been launched in its basic version only in June 2015. Moreover, since the Smarth2O app is not yet equipped with interactive features between users groups, behavioural data also have still to be collected. The first prototype can therefore be seen as an input-output model based on the implementation of a single user model capturing the influence of meteorological characteristics on water consumption.

The performance of the model presented in Sections 3.3 and 3.4 has been evaluated by considering a period of 52 days ranging from 1/8/2015 to 21/9/2015. We selected 193 users who had been monitored continuously during such period and for a sufficiently long time interval within each of the five meteorological classes, in order to ensure the statistical significance of the consumption profiles obtained by the model.

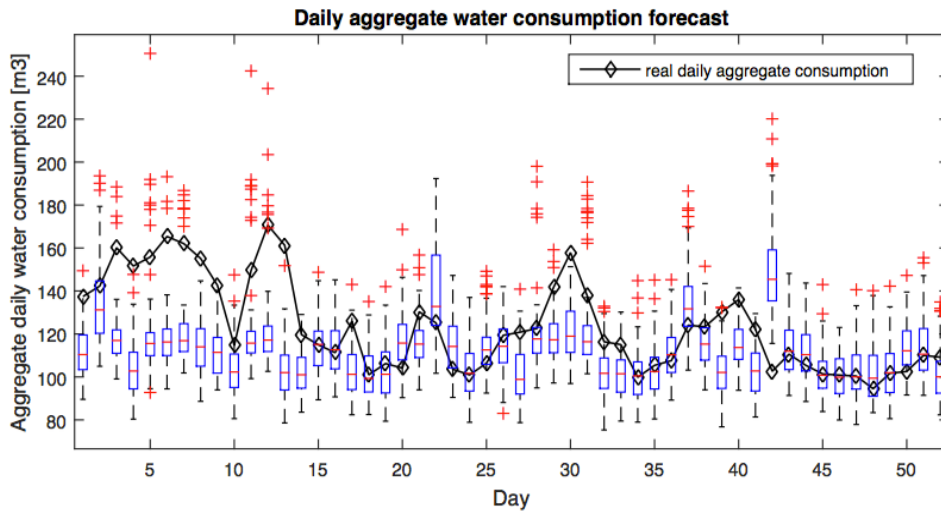


Figure 11. Comparison of the aggregate daily water consumption forecasted by our model (boxplots) with the actual consumption values (black line)

The meteorological forecasts provided by the website *Osservatorio Ambientale della Svizzera Italiana's* website (ti.ch/oasi) were used to identify the meteorological class for each of the 52 days, according to the temperature and irradiation ranges reported in Table 2. Then, the daily water consumption has been forecasted according to the consumption profiles computed over the training data (see Subsection 3.2). More in detail, for each day, the consumption class (low, medium or high) of each user has been randomly chosen according to the probabilities characterising the user's consumption profile. Once the consumption class has been selected, the daily consumption has been generated according to the empirical probability distribution characterising the consumption class.

The whole procedure has been repeated 100 times and the following metrics have been computed:

- aggregate daily water consumption of the 193 users
- aggregate consumption over 3 days
- aggregate consumption over 5 days.

The forecasted consumption amounts have then been compared to the real values extracted from the meter readings dataset. Results summarised in table 3 show that the consumption forecasts provided by the model are quite close to the real water usage (average gap around -7%, mostly due to a slight consumption underestimation during hot sunny days).

Table 3: Forecasting consumption results

	average % gap	max % gap	min % gap
daily aggregate	-7.05	114.89	-51.29
3 days aggregate	-7.63	34.17	-37.47
5 days aggregate	-7.92	27.62	-31.98

In particular, Figure 11 compares the aggregate daily water consumption forecasted by our model (boxplots) to the actual consumption values (black line). As the boxplot pictorial view shows, the daily forecasts exhibit high variance. This is mainly due to the difficulty of predicting the consumption of large consumers, whose water usage pattern may vary consistently from day to day. However, such variance diminishes when forecasts are aggregated over a few days (see the values of maximum/minimum gaps reported Table 3, columns 3 and 4, which decrease as the aggregation time span increases).

The proposed prototype shows therefore potential to effectively capture aggregated urban water demand variability with respect to meteorological sensibility. It thence tell us that we are on the right path to develop an agent-based model to help water utilities and urban planners in their decision making process.

The proposed model is clearly missing several aspects. Firstly it has to be strengthened by accessing more data and by incorporating users' psychographics and the house characteristics. Secondly, in parallel with the analysis of behavioural data, we should add a fully developed interaction module to capture influence mechanisms and social interaction.

4. Conclusion and follow-up

In this deliverable, we presented a first agent-based simulator for predicting household's water consumption based on past real-time consumption data. More specifically, we described an ABM whose characteristics and parameters implement a model of user behaviour based on past consumption data.

The agent-based model is planned to be implemented within the SmartH2O platform and being able to directly access and analyse past consumption measurements.

In doing so, by accessing the model through the SmartH2O platform, stakeholders will be able to foresee consumers' behavioural patterns in front of exogenous variables, like meteorological conditions, but also of the adoption of social and price policies influence mechanisms.

In view of the last points, the next steps within WP3 will be:

- to continue testing the model against more water consumption data and compare the obtained results against the benchmarks described in Section 2;
- once users' psychographics and the house characteristics data will be available, to fully incorporate the single user behavioural model developed according to the lines described in the WP3 deliverable D3.2
- to further develop the agent based model in such a way that it takes into account the social interaction among consumers, and the different influence mechanisms, such as water price and rewards changes.

5. References

- [Atha2005] Athanasiadis, I.N., A.K. Mentis, P.A. Mitkas., and Y.A. Mylopoulos (2005). A Hybrid Agent-Based Model for Estimating Residential Water Demand. *Simulation*, 81, 3: 175-187.
- [Bart2001] Barthelemy, O., S. Moss, T. Downing, and J. Rouchier (2001). Policy modelling with ABSS: The case of water demand management. Centre for Policy Modelling, Manchester Metropolitan University, Manchester, CPM Report.
- [Bali2011] Baliotti S., Helbing D. (2011), How to Do Agent-Based Simulations in the Future: From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design. Santa Fe Institute Working Paper 2011-06-024. Reprinted as Chapter 2 in: Helbing D. (2012), *Social Self-Organization, Understanding Complex Systems*, Springer Verlag.
- [Blok2010] Blokker, E., J. Vreeburg, and J. van Dijk (2010), Simulating residential water demand with a stochastic end-use model, *Journal of Water Resources Planning and Management*, 136(1), 19-26.
- [Chu2009] Chu, J., C. Wang, J. Chen, and H. Wang (2009), Agent-based residential water use behaviour simulation and policy implications: A case-study in Beijing city, *Water resources management*, 23(15): 3267-3295.
- [Down2000] Downing, T.E., S. Moss, and C. Pahl-Wostl (2000). Understanding Climate Policy Using Participatory Agent-Based Social Simulation. *MABS*: 198-213.
- [Erns2005] Ernst, A., C. Schulz, N. Schwarz, and S. Janisch (2005). Shallow and deep modeling of water use in large, spatially explicit, coupled simulation system. *ESSA*: 158-164
- [Helb2014] Helbing D, Yu W, Opp K-D, Rauhut H (2014). Conditions for the Emergence of Shared Norms in Populations with Incompatible Preferences. *PLoS ONE* 9(8): e104207. doi:10.1371/journal.pone.0104207
- [Gala2009] Galan, J., A. Lopez-Paredes, and R. Del Olmo (2009), An agent-based model for domestic water management in Valladolid metropolitan area, *Water Resources Research*, 45(5).
- [Gilb2008] Gilbert, N. (2008). *Agent-Based Models*. Sage Publications, Inc.
- [Link2013] Linkola, L., C.J. Andrews, and T. Schuetze (2013). An Agent Based Model of Household Water Use. *Water* (2013) 5, 3: 1082-1100.
- [Lope2005] Lopez-Paredes, A., D. Sauri, and J.M. Galan (2005). Urban Water management with artificial societies agents: the FIRMABAR simulator. *Simulation*, 81(3): 189-199.
- [Maca2010] Macal C.M., North M.J. (2010), Tutorial on agent-based modelling and simulation. *Journal of Simulation* 4: 151-162
- [Rixo2006] Rixon, A., M. Moglia, and S. Burn (2007). Exploring Water Conservation Behaviour through Participatory Agent-Based Modelling. In: Andrea Castelletti, Rodolfo Soncini-Sessa (ed.). *Topics on System Analysis and Integrated Water Resources Management*, Elsevier: 73-96.
- [Tseg2009] Tsegaye, S., and K. Vairavamoorthy (2009). Agent-Based Modeling to Estimate Residential Water Demand and to Explore Optimal Demand Side Water Management strategies. SWITCH Project Report.