

Extending No-MASS: Multi-Agent Stochastic Simulation for Demand Response of residential appliances

A. Sancho-Tomás¹, J. Chapman², M. Sumner³, D. Robinson²

¹Department of Architecture and Built Environment, University of Nottingham, UK

²School of Architecture, University of Sheffield, UK

³Department of Electrical and Electronic Engineering, University of Nottingham, UK

Abstract

Demand Response (DR) has been proposed as an efficient and inexpensive solution to face the challenges of the evolving power system: to reduce peak demands and thus demands for additional generating capacity, and to improve localized energetic autonomy. The success of DR programs is highly dependent on the acceptance and reactions of end-users: their willingness to devolve control and/or to proactively adjust their energy using behaviours. However, experimental trials to identify the best DR configuration for each type of consumer is very costly. Here simulation has a valuable role to play, in identifying promising DR strategies, taking into account occupants characteristics.

This paper describes and evaluates a proof-of-concept extension of a multi-agent stochastic simulation platform that simulates occupants behaviours (No-MASS) to also simulate DR strategies. This is applied in the first instance to electrical devices in households: categorizing these as demand devices (e.g. large and small appliances, heating and hot water systems), supply devices (e.g. PV, cogeneration, wind turbines) and electrical storage devices (e.g. batteries, HEVs). The device-specific DR strategies are learned by our device-agents using machine (specifically reinforcement) learning, to maximize a function rewarding the utilization of locally generated energy. We close this paper by discussing our next steps to add further functionality to No-MASS for DR simulation.

Introduction

The power system at the distribution level is evolving towards decentralised monitoring, supply and control. Small-scale renewable energy generation and storage devices are increasingly incorporated into the network, which in turn needs intelligent management to fulfill local power demands. In this context, Demand Response (DR) strategies and technologies that modify consumption patterns in response to signals [e.g. price] is seen as a core and affordable technology to better allocate resources, without the need of costly upgrades to the current power distribution system.

Simulation software is needed that is capable of modelling and optimising Demand Response (DR) strate-

gies. To this end, we can articulate some key requirements R that a DR simulation platform should aim to satisfy. It should be capable of:

- R1. Simulating demands for four appliance archetypes: *i*) switched on, regulated by and switched off by the user (e.g. cooker), *ii*) switched on by the user and off when a programme is complete (e.g. washing machine), *iii*) switched on and off according to some programme or schedule (e.g. hot water system), *iv*) continuous cycling (e.g. refrigerator). The user-interaction should be stochastic.
- R2. Drawing power to satisfy demands from: *i*) local generation capacity, *ii*) local storage devices, *iii*) the local microgrid and/or the national grid; similarly of diverting locally generated power to either local demand and storage devices, or to the local/national grid.
- R3. Deciding, to satisfy some objective function, where power should be drawn from/diverted to; rescheduling demands (or the provision of energy related services), given some pre-defined constraints related to service delivery (e.g. a washing machine may be activated after having been enabled, but must complete the wash by a predefined time).
- R4. Presenting information to the user and emulating the users decision making rationale regarding the rescheduling of user-controlled devices.
- R5. Accounting for diversity in the extent to which users are willing to relinquish control and to actively engage in behavioural change. Beckman (2013) notes that user engagement is crucial to the success of DR strategies.
- R6. Facilitating the above for communities consisting of buildings with numerous demand devices and potentially numerous supply and storage devices which can communicate within and between buildings to achieve individual homeowners' requirements; potentially also those of the local low voltage network to which they are connected (e.g. in terms of network stability and safety).

Several software tools have been developed that address subsets of these requirements, simulating resi-

dential DR and accounting for a range of objective functions (Wijaya et al., 2013; Gudi et al., 2010), but none addresses the range of these requirements in a comprehensive way. Wijaya et al. (2013) present a simulator for DR, which mainly focus on R1 (demand profiles are based on correlated probabilities), R4 and R5 requirements. Although R4 and R5 are interestingly developed by modelling sensitivity in energy price, energy perception and communication, there is no attempt to consider supply agents (R2), or to involve network operation in any way; finally, the optimisation algorithm for shifting DR actions is left unspecified, claiming that "many candidates can be used". Another DR simulator by Faria et al. (2010), focuses on distributed generation issues (R2) and its effects on the electrical network and energy market elasticity (R6). Their objective is to simulate a variety of DR methods that minimize the costs of a generation shortage situation, implemented as a mixed-integer linear model (R3). However, energy use is considered with typical profiles, making this simulator incapable of dealing with R1, R4 and R5.

This paper describes the extension to an existing Multi-Agent Stochastic Simulation platform (No-MASS), to comprehensively address this shortfall in DR simulation capability.

No-MASS was initially developed to model the presence, activities and related behaviours of synthetic occupants (agents) of buildings that are co-simulated with EnergyPlus using the Functional Mockup Interface standard (Chapman, 2016). In this way it is straightforward to model occupants and their impacts on the energy performance and indoor comfort of the buildings they occupy.

Interestingly, the architecture of No-MASS is readily extensible to also consider Device-to-Device (D2D) communications. Thus creating a platform that has the capability to simulate:

- Occupant-agents behaviours and interactions between them (our prior No-MASS).
- Device-agents behaviours and interactions between them (D2D).
- Interactions between occupants and devices.

In this paper we focus on the first two of these capabilities. The paper is organised as follows. After reviewing DR methods and optimization algorithms, we describe No-MASS and the algorithms that have been incorporated to support load re-scheduling and battery operation. We then illustrate the application of this new prototypical platform to estimate the effects of appliance load shifting and electrical storage with the objective of maximising the use of locally generated renewable energy of a domestic building. We close the paper by outlining forthcoming work to address the third of the above capabilities and to evaluate this new platform using empirical data.

Centralised and decentralised DR

Looking at the control mechanism implemented for the DR schemes, two main groups arise: centralised or decentralised programs (Deng et al., 2015; Vardakas et al., 2015). In the former, decisions on load re-scheduling or power dispatch are made by a central controller, informed by the operation of each individual element in the network. Optimization methods used in these cases range from traditional calculus methods (Mohsenian-Rad and Leon-Garcia, 2010) to heuristic optimization (Gudi et al., 2010). However, these approaches become challenging when large and complex networks are considered. On the other hand, decentralised DR schemes incorporate the ability of distributed decision-making, assuming a certain degree of intelligence of the devices involved (such as smart meters and appliances, power electronics and so on), which ensures direct communication between the elements in the network.

DR and Multi-Agent Simulation

Multi-agent systems (MAS) have been proposed in different power engineering applications (McArthur et al., 2007), from marketplace simulations to operation and control methods of the power system. They have been exploited in different ways such as monitoring and diagnostics, distributed control (Vasirani et al., 2013), fault protection and modelling and simulation. Particular interests is posed in Multi-Agent systems being used for Demand Response (Iacovella et al., 2016; Wijaya et al., 2013; Gudi et al., 2010), which present a highly flexible and extensible modelling approach for simulating D2D communication between supply, storage and demand devices: systems are defined in terms of agents which fulfil individual and collective goals, being able to compete, collaborate, negotiate and learn behaviours.

No-MASS background

As noted earlier, No-MASS (Nottingham Multi-Agent Stochastic Simulation) is a platform that was originally developed to model the presence, activities and related behaviours of occupants of buildings, or groups of buildings, and the energy and indoor comfort consequences of these behaviours. To this end No-MASS employs four complementary modelling strategies:

1. **Data driven stochastic models** predicting occupants presence (Page et al., 2008) and associated metabolic heat gains, their activities whilst present (Jaboob, 2015) from which location is also inferred, and their use of lights (Reinhart, 2004), windows (Haldi and Robinson, 2009) and shading devices (Haldi and Robinson, 2010).
2. The processing of votes for multiply-occupied spaces to emulate agents **negotiation mechanisms** (e.g. agents influence-weighted votes to open [1] or close [0] a window; the outcome re-

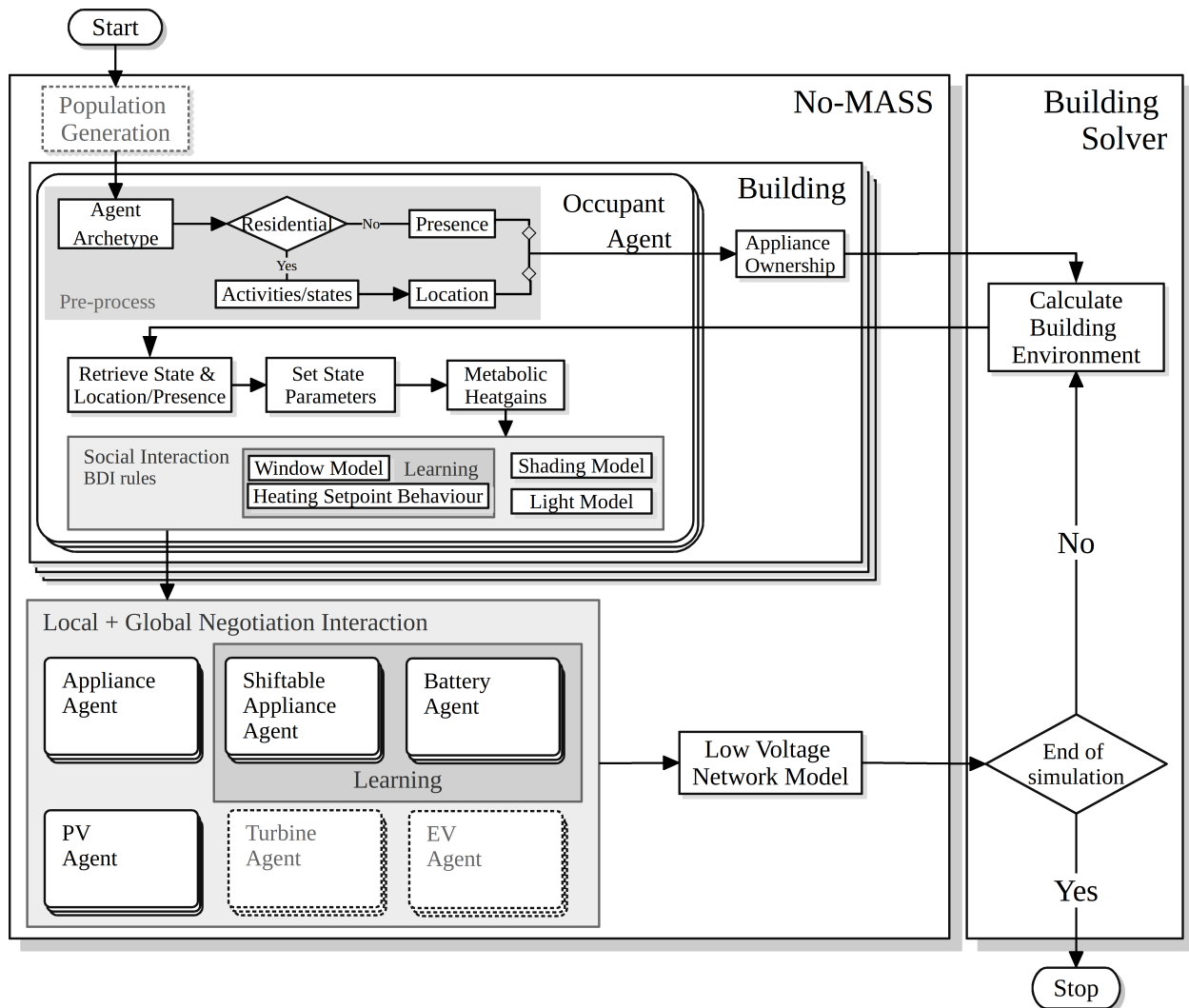


Figure 1: No-MASS flowchart.

ceiving the greatest number of votes, weighted by influence, being effected).

3. Use of a **Belief-Desire-Intent** (BDI) framework to emulate agents behaviours that are simple in character (e.g. closing curtains when it is dark, the closing of a window whilst bathing) but for which data is scarce.
4. **Agent learning** mechanisms to emulate agents behaviours that are simple in character (e.g. choice of heating set-points) and for which data is also scarce.

The architecture of No-MASS (and its coupling with EnergyPlus), the data-driven (Strategy 1) models and their application to the simulation of both domestic and non-domestic buildings are described in Chapman (2016). The extension of No-MASS to model agents negotiations and the data scarce modelling of both simple and more complex behaviours (Strategies 2-4) are described in (Chapman, 2016).

To facilitate the extension of No-MASS to handle DR

simulation and optimisation (and indeed the more complete modelling of the impacts of occupants behaviours), our Strategy 1 models have been complemented with models of occupants ownership and use of large (Jaboob, 2015) and small (Sancho-Tomás et al., 2017) electrical appliances. These models have themselves been complemented with models and data of electrical storage (an electric battery) and supply or conversion (a photovoltaic panel) devices; and of a low-voltage network model.

Workflow

Figure 1 illustrates the newly extended No-MASS architecture. First (upper left) a population generator creates a household of a size and demographic composition that is suited to the house being simulated (Le Thellier et al., 2017). Next the conventional No-MASS modelling tasks are executed.

Household member agents are assigned archetypal properties (room associated with activities, clothing and activity characteristics which randomly assign

unique behaviour probabilistic models). Activities are then computed for the length of the simulation in a pre-process (for activities are not dependent upon environmental inputs); likewise electrical appliances are assigned to the household.

A loop then commences in which indoor/outdoor environment parameters are parsed from EnergyPlus to No-MASS for the present time step. From the pre-processed activities, agents' location, activity and clothing level are set, from which metabolic heat gains are calculated.

A series of models predicting interactions with windows, shading devices, heating systems and lights are then called. For heating interactions, agent learning is employed to determine the transient setpoints that minimise a cost function that combines heating costs and discomfort costs. In the case of multiply-occupied spaces, social interactions are considered at this stage through the vote processor to determine the negotiated outcome. Finally, BDI rules are called for straightforward interactions for which data is scarce. The workflow proceeds with the prediction of demands for small and large appliance, supply from energy conversion systems (only PV being enabled at this stage) and storage (only from batteries at present). Appliances are also shifted at this stage, again using agent learning, to maximise a cost reward function (of which more later). Although an LV network model has been developed and integrated with No-MASS, this is not currently utilised. It will be enabled in a future version in which No-MASS is configured to model multiple buildings simultaneously. The calculation then proceeds to the next time-step, exiting this loop at the end of the simulation period. The corresponding physical models (for D2D modelling) are described below.

Implementation: No-MASS for DR

The extension of No-MASS will, once complete (for this paper describes a partial proof of concept), address the range of requirements outlined earlier. In its present form No-MASS addresses requirements R1 (demand models), R2 (various power flows) and R3 (DR optimisation), further explained in this section. The strategies for achieving this are as follows:

- Develop mechanisms for D2D communications between energy conversion, storage and demand devices represented using the models described above.
- Implement strategies for load-shifting and optimal charge/discharge of a battery using agent learning algorithms. This will reflect whether or not multi-agent simulation and machine learning are effective approaches to test different DSM strategies on different socio-demographic groups.
- Extract or quantify differences when simulating residential electrical self-consumption/autonomy

(or other DSM objectives) for different socio-demographic groups.

If successful, then No-MASS could potentially be used to evaluate scenarios involving human interaction and behaviour change (R4 and R5) and to support tariff design. For instance, we could assess to what extent active engagement of users increases electrical energy autonomy, or to what extent price signal impacts on behaviour change.

In its present form No-MASS is employed to simulate a single-house that follows a price-based DR scheme influencing the operation of two *smart* appliances and a battery storage system, with the objective of maximizing self-consumption from a rooftop PV panel. The following sections, explore in more detail the methods and algorithms used for this purpose.

Electricity demand models (R1)

The demand forecast models developed for No-MASS are based on stochastic methods. Devices have been classified in large appliances (high-load and commonly owned) and small appliances (range of low-load appliances), and follow two different modelling methodologies.

Large Appliances

Large appliances include the cooker, TV, microwave, washing machine and dishwasher. They have been modelled as a three-step process (Jaboob, 2015). First, the probability of switching on is predicted using a time-dependent Bernoulli process. Second, the duration for which devices will remain on is predicted using survival analysis. Finally, transitions between categories of fractional power demand (the fraction of maximum possible) are predicted as a Markov process, at 10 minute resolution. Probabilities of switching on are depicted in Figure 2.

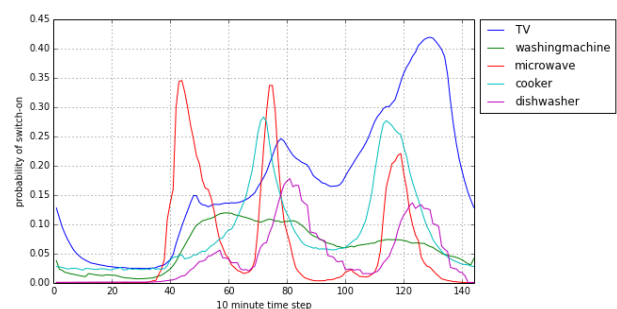


Figure 2: Switch on probabilities of large appliances.

Small Appliances

Small appliances are modelled as aggregations of appliances, following four categories: small appliances in the kitchen, audio-visual appliances, computing appliances and other appliances (Sancho-Tomás et al., 2017). A multi-state survival model is used: eleven fractional energy (ratio to maximum energy) states are defined $f = \{0; 0-0.1; 0.1-0.2; \dots; 0.9-1\}$.

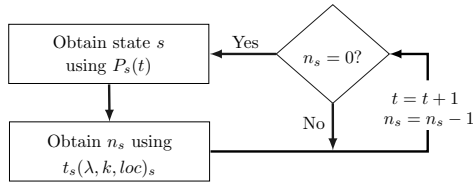


Figure 3: Small appliances simulation flowchart.

The survival time that the appliances remain in one of those states is calculated using equation a Weibull distribution. A simplified flowchart of this process is represented in Figure 3.

Heating

Electrical demand for heating can be obtained thanks to the coupling of No-MASS with Energy Plus.

Supply and power flow (R2)

Photovoltaic data

However, for the current purpose of demonstrating the proof of concept of our proposed modelling approach, we are using measured performance data that characterise PV systems' output in-lieu of a predictive model. In the future we will of course integrate a suite of predictive energy conversion system models.

Electric Battery Storage System

A linear model has been implemented for the charge and discharge of the battery (Pholboon et al., 2015). For the moment, conversion losses of the battery have been assumed to be constant with the State Of Charge (SOC) and equal to 2%. In the case study, an energy storage system of 2kWh is taken into account, with a power limit of 1kW.

Low Voltage network model

A forward/backward sweep solver¹ has been used for power-flow analysis. It has been implemented as a recursive² algorithm. In the case of a distribution network, recursion constitutes a method to efficiently solve branched networks. During the forward sweep, the currents and voltages at each node depend on the currents and voltages at the child nodes connected to them (one if linear network, several if branched); whereas during the backward sweep, currents and voltages depend on the parent node.

Although the network model is implemented and available in No-MASS, its demonstration is out of the scope of this paper.

Agent representation

In the initial methodology of No-MASS, household occupants were identified as software agents. In ex-

¹Background on iterative power-flow methods can be found in Kersting (2012).

²In Computer Science, *recursion* is a method employed when the solution to a problem requires the solution of smaller instances of the same problem. In practice, it is implemented by defining functions that are allowed to call themselves.

tending No-MASS to handle D2D interactions, electrical appliances and energy systems are represented as an agent sub-class. Each agent has an ID, a peak power (either for demand or supply) and a priority of service. Also, the operation starting time and profile can be calculated (type of prediction depends on the type of device).

During each time step, each device communicates the amount of energy (power integrated over the duration of the time step) to be requested or delivered. First, energy from our PV panel is allocated to appliances based on their priority. Any shortage is provided (perfectly) by the upscale grid. When an electrical storage device is unavailable, any surplus is exported to the grid. When it is available, the battery can be charged using excess PV energy, and can be later discharged to run electric appliances when price conditions favour this strategy.

Agent learning for DR optimisation (R3)

In this particular work, we focus on load shifting and optimal electrical battery operation as a means to maximise locally generated energy. To achieve this, a reinforcement learning algorithm is implemented in two types of the agents: *a)shiftable* appliance agents, that can be regulated autonomously: washing machines and dishwashers, and *b) battery* agents, that learn optimised charge and discharge operation.

Load Shifting Optimisation

The core idea for appliance-reschedule is depicted in Figure 4. For each time step, the switch-on model is run. When a switch-on event is predicted, the profile of use and the new starting time are calculated. The temporal window within which appliances may be shifted has a maximum extent of 24h, but this may be reduced to meet pre-defined delivery time constraints.

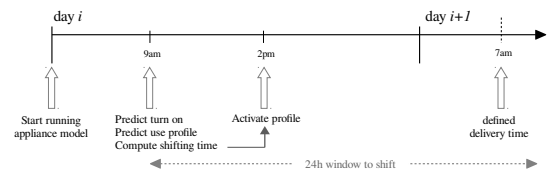


Figure 4: Appliance re-schedule diagram.

Battery charge and discharge

Our battery has been implemented to charge whenever there is a surplus of PV power available (supply > demand), before it is exported to the grid (whilst storage capacity is surpassed). For the discharge process, the battery has been configured to relieve peaks of high demand that require electricity imports. The battery learns when the highest hourly peak demand $P_{grid_import_MAX}$ and whilst imports from the grid exceed some threshold (currently 70%) of this de-

mand: $P_{grid_import} > 0.7 P_{grid_import_MAX}$. Otherwise, it learns to turn itself off in that hour, storing the remaining energy for peak periods.

This fairly primitive (dis-)charging strategy is defined in Figure 5. We plan to refine this strategy to also consider (dis-)charging from the grid (dashed lines in Figure 5); again while price conditions are favourable.

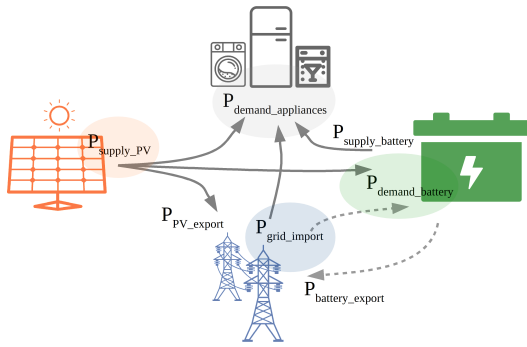


Figure 5: Power flows.

Q-learning algorithm

There are multiple algorithms available for reinforcement learning. Q-learning algorithms Watkins (1989) are one of the most widely used, with their major advantage being their great simplicity. They require a minimal amount of computation and, on its basic formulation, they can be expressed by single equations Sutton and Barto (1998), and easily implemented in computer programs. Q-learning algorithms (Watkins, 1989) allow agents to learn a response from a reward, to an action. This allows agents to develop an understanding of their preferences over time. An agent learns the best action in a given state by trying every action in a state and updating the expected reward with the actual reward for that action. This is particularly useful, compared to other machine learning methods, for modelling appliance shifting, as the appliances can test different strategies for maximising their reward where there is a clear link between an action and a driving stimulus. For example, does a chosen action (turn on at a later time) reduce peak power demand over a time-period. This is a quick and effective methodology that would be difficult to model through rules due to the complexities involved, especially when considering multiple shifting appliances. Each appliance would need its own set of rules to ensure they would not turn on at the same time, whereas Q-learning allows them to learn their own preferences, considering other appliance demands. In Q-learning algorithms, an agent chooses an action at a given state based on a Q-quantity, which is a weighted reward based on the expected highest long term reward (Watkins, 1989). The Q-quantity

is defined for each state-action combination, creating a Q-table. The values in the Q-table are updated each time an agent selects an action. Let s_t express the agent's state at time step t , and a_t a chosen action. Using this information, the Q-value for the corresponding combination of (s_t, a_t) is updated:

$$Q_t(s_t, a_t) = Q_t(s_t, a_t) + \alpha [r + \gamma * \max_a(Q(s_{t+1}, a)) - Q_t(s_t, a_t)], \quad (1)$$

where r is the reward function, $\alpha \in (0, 1]$ is the learning rate and $\gamma \in (0, 1]$ is the discount factor. The discount factor specifies how soon the agent cares about the reward: near terms goals when $\gamma \sim 0$ (myopic agent), otherwise long term rewards when $\gamma \sim 1$.

Moreover, the selection of an action is not completely deterministic, but uses an epsilon-greedy approach: the best action is selected with $1 - \epsilon$ probability, and a random action is selected otherwise. For instance, if $\epsilon = 0.1$, the (currently) best action will be adopted 90% of the time. This randomness is introduced so that the agent explores more in order to discover the best action over the whole period of time.

For an appliance agent, we map the state (hour of the day) to an action (future hour of day to initiate the appliance). This creates a mapping of 24 hours to 24 hours making the Q-table space of 576 combinations in size. An example of a profile shift might involve calculating whether at state s_t the appliance is required to be turned on (e.g. the dishwasher is loaded and ready). If so the appliance demand profile should be calculated using the large appliance model. However this should not yet be initiated. Instead the Q-table should be used to retrieve a time (action) at which the appliance programme should be initiated (see Figure 4).

The effectiveness of the learning process is highly dependent on the selected parameters α , γ and ϵ . Moreover, the reward function r can consider different variables of interest, such as cost, autonomy or voltage stability, allowing to use the same methodology to explore a range of objectives. Convergence of the learning process is achieved when the Q-table no longer changes its best values between simulations.

Reward function. For the purpose of this paper, the reward function consists of two components:

1. The inverse of a cost signal (effectively then an income signal), based on indicative prices for energy imports. A Time-Of-Use (TOU) tariff is tested, with three different pricing periods: on-peak (between 7-9h and between 16-21h), off-peak (at night between 23-7h), and mid-peak for the rest of the day. In Figure 6, the TOU signal is related to an averaged Real Time Price (RTP) signal³. Values of the tariffs have been normalized

³Data available from http://bmreports.com/bsp/bsp_home.htm

$[0, 1]$, since only relative differences are useful for the learning algorithm.

2. In the case of load shifting, punishment when the service is not satisfied on time (the washing cycle is not complete within the delivery-time-constrained window).

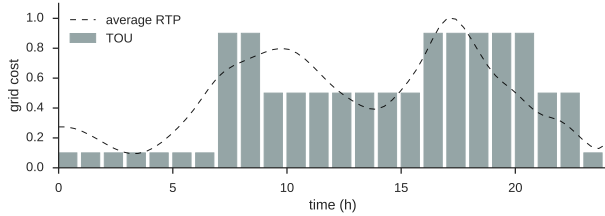


Figure 6: Reward function uses price signals.

Simulation: maximise self-consumption

In this section we present the results of simulating a household with a single professional resident occupying a detached house using No-MASS. Our objective is to maximise the utilisation of energy converted by a 3.8kW-peak PV panel installed on the roof by a set of large⁴ and small appliances⁵, including electrical heating, connected to a 2kWh electric storage system. Note that we aim to present a proof-of-concept of the integration of DR techniques with No-MASS; for that, in the results presented below there is not building simulation involved.

To this end we use price-incentive based DR control strategies to: *a)* re-schedule autonomously controlled washing machine and dishwasher appliances; *b)* charge and discharge the battery. In both cases, device operation can be tuned using a tariff signal that fosters the switch on of autonomous appliances whilst sunshine is likely, and the charging of the battery whilst there is excess PV generation. This operation is aimed to maximise self-consumption over time, while reducing on-peak demand (as defined in Figure 6).

Self consumption (SC), also referred to as a load matching index (Salom et al., 2011), can be defined as the ratio of power use from on-site generation to the total power used, as expressed in equation 2. It is inversely related to the amount of PV power exported to the grid.

$$\text{Self Consumption}(\%) = \frac{P_{\text{demand from PV}}}{P_{\text{demand}}} \quad (2)$$

Although maximising self-consumption is our goal, this is not formally optimised by our Q-learning algorithm (eq. 1), as it is not explicitly expressed in

⁴Cooker, dishwasher, washing machine, fridge and TV.

⁵Set of audio-visual appliances, small kitchen appliances, computing appliances and other appliances. The set size depends on the simulation, in which they are allocated following a Monte Carlo sampling.

our reward function. But this is indirectly achieved through the tariff signal, (Figure 6), through which we pay a low price whilst using off-peak centrally-generated energy and close to zero for locally generated energy. In this way, smart appliances learn to turn on during low-cost periods.

Results

Three scenarios are presented for comparison: first, the base case, where the models and systems are run without shifting demand and without considering a battery. Second, appliance re-scheduling capabilities are added to the base case. Third, a battery system is also considered. These three scenarios are simulated for winter and summer (heating not necessary).

Each simulation runs for a period of one week. To account for the stochasticity in the calculations, results are presented as a distribution from a set of replicates; 200 for each scenario. Where agent learning is involved, the Q-learning algorithm requires training to populate the Q-tables. For the work presented here, a learning period of 300 weeks was necessary. This may seem long, but it is understandable, given the nature of the events: a turn-on of the washing machine or dishwasher will seldom occur more than once a day, and very rarely more than twice. Efficient learning needs a large sample of events, and will thus need a relatively long learning period. However, this period could be reduced by defining fewer states and actions, thus reducing the dimensionality of the Q-table. A graphical example of a 24×24 Q-table is presented in 7, in which each pixel is false coloured according to the q-value of an action from its state s_j to s_{j+1} . Shifting to peak times (7-9h and 16-21h) has a smaller reward (darker pixels) than shifting to sunshine hours, or off-peak (23-7h) times.

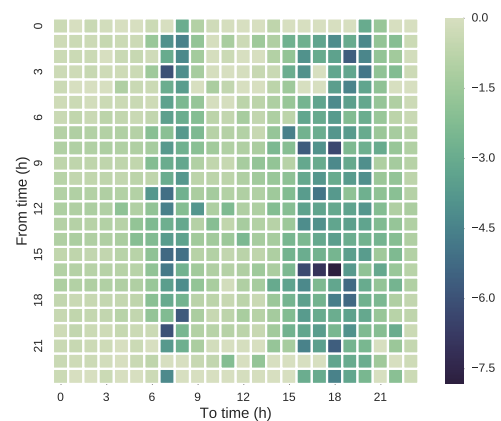


Figure 7: 24×24 Q-table. The units of the legend correspond to reward values, derived (as explained above) from the tariff unit price.

One day of winter simulation is presented in Figure 8. The electrical demand consists of large appliances, small appliances and heating (delivered by an electric heating system). Demand coming from shifted

devices (washing machine and dishwasher) and heating demands are indicated by shading. It can be seen that the total load is comparable to the amount of energy generated on-site. However, occupants schedules and their main use of electrical devices does not always occur during sunshine hours. Nevertheless, it can be seen that part of the flexible load has been shifted to the middle of the day where it can operate using locally converted power.

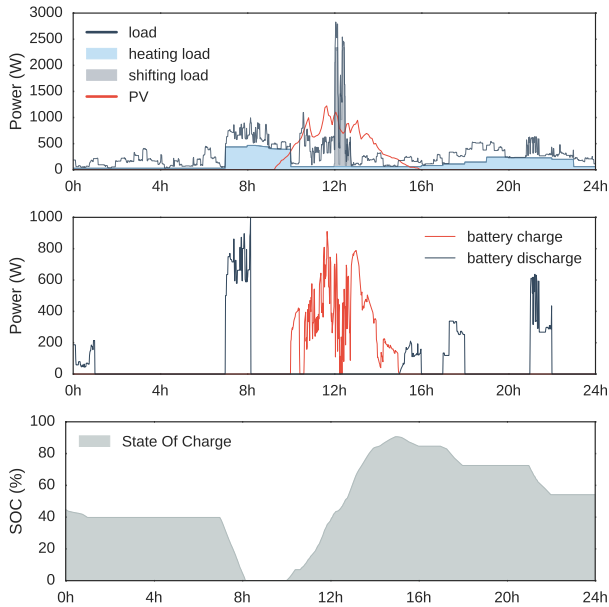


Figure 8: 24h of simulation for a winter day. PV profile and electrical demand on top; battery charge and discharge in the middle; SOC level at the bottom.

When a battery is available, the excess of solar energy is stored. Thus, the charging power (limited to 1kW) is shown at the bottom of Figure 8.

As a consequence of the introduction of the TOU tariff, we can simulate how the DR algorithm adjusts when electricity is consumed from the grid. For our three-case scenario, results of the weekly energy grid import at the different tariffs, as well as PV export, are represented for summer and winter in Figure 9 (the bars represent average values for a set of 200 simulations). Our first observation is that the introduction of the tariff scheme is efficient in reducing the on-peak grid imports, in both summer and winter periods. Furthermore, a mean of over 13kWh of solar energy is exported in winter when there is no storage system available. A dramatic reduction close to 90% is obtained when the battery model is introduced. 70% reduction is obtained for the summer simulations. This shows that storage systems can play a significant role when considering the electrical autonomy of distribution networks. Also, although the algorithm employed is relatively simple in its formulation, it is nevertheless very efficient in reducing on-peak and (modestly) mid-peak grid imports, whilst maintaining cheaper off-peak imports

and maximising the use of solar energy. The increase in self-consumption (equation 2), by 82% and 86% with respect to its initial value in winter and summer, respectively, is more clearly presented in Figure 10.

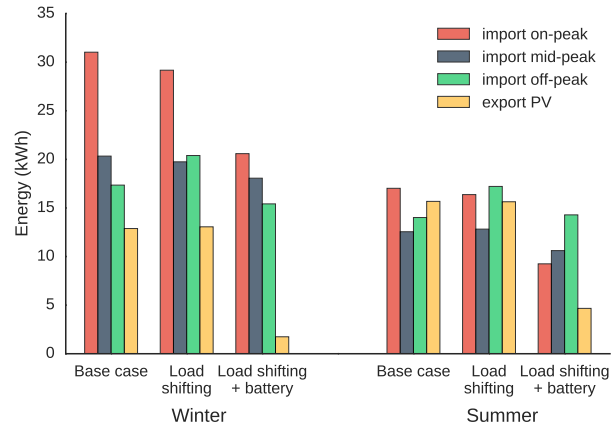


Figure 9: Import and export for the three scenarios, for summer and winter.

However, the reduction obtained only with the shiftable appliances is very modest. This is partly because the two appliances considered represent a small proportion of the total energy demand; suggesting that further shiftable autonomous devices should be considered to better understand the potential for this strategy, and/or that new tariff structures are explored that encourage this behaviour.

At present we are simply evaluating the utility of this platform for the case of a single building. But if we were to simulate an entire district of buildings connected to an LV network downstream of a substation, doubling their cumulative on-site energy use from 15% to almost 30% would have a considerable impact on upstream demands.

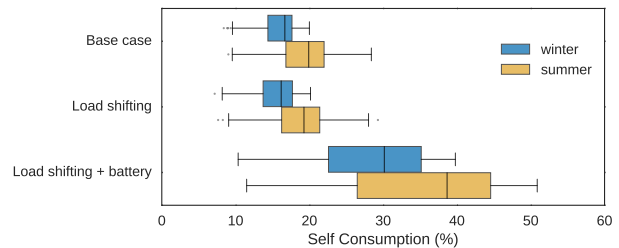


Figure 10: Self consumption index for the scenarios, for summer and winter.

As the number of buildings (and consequently agents) increases, so does the complexity of the multi-agent approach. An extra layer of coordination needs to be added to the architecture, to make the agent learning efficient by limiting the number of communication channels between agents to those strictly necessary. Multi-building simulation using No-MASS will be studied in the future. Preliminary tests suggest

that No-MASS/DR is able to simulate at least 100 buildings.

The way that No-MASS/DR has been developed allows to a flexible configuration of the systems and models used, through a XML file. In the proof-of-concept presented here only those models related to electrical energy use and DR were selected, but the general version of No-MASS contain models to simulate person-to-person interaction, as well as person-to-device.

Conclusion and future work

In this paper, we have described and evaluated an extension to the multi-agent stochastic simulation platform No-MASS to support the simulation of demand response strategies.

The prior focus of No-MASS was on the integration of models of occupants behaviours with building simulation software, in particular with EnergyPlus.

Our hypothesis was that the No-MASS's underlying software architecture, and many of the modelling techniques already utilised within it, are readily extendable to handle DR simulation (simulating device-agents in addition to occupant-agents). Although only for a simple case of a single building, we believe that we have successfully tested this hypothesis.

In particular we have demonstrated the integration of price-based Demand Response strategies to optimise for load-shifting and the charging and discharging of a battery, by incorporating learning abilities into our device agents. Q-learning has been proven to be a successful candidate for that task, although other methods should be tested in the future.

More specifically, when applied to our individual house case study that utilises time of use tariff structure, this learning algorithm has effectively: shifted demand to from on- and mid- peak periods to off-peak periods, particularly when combining shiftable demand devices with battery storage; doubled the self-consumption percentage (the fraction of energy demand that it satisfied by on-site generation).

Although we believe that we have successfully demonstrated this proof-of-concept there is considerable scope for improvement to this framework. In particular with respect to (and still only considering requirements R1 to R3):

1. The battery model should be improved to model state-of-charge dependent losses.
2. In this work, battery and heating demand have been combined to reduce on-peak grid imports. In a more realistic situation, the heating system can be considered as an additional learning device on its own, that contributes to the global goals. More sophisticated reward functions for charge and discharge of the battery need to be tested for this.
3. For islanded scenarios in which power cannot be drawn upstream of the mains network, a for-

mal bidding mechanism needs to be integrated, favouring devices of relatively high priority when reserves are limited (e.g. emergency lighting, or freezers that have undergone a long delay since the compressor was last enabled).

4. Alternative tariff structures should be explored that better favour self-consumption.

Looking further to the future (R4 to R6):

4. No-MASS should be generalised to solve for multiple buildings (both occupant-agents and device-agents) simultaneously; these buildings being inter-connected via our LV network model.
5. The above tariffs might also consider local network integrity, when the LV model is utilised as intended.
6. Empirical evidence from DR field trials should be used to predict the extent to which users are willing to: *a)* devolve control to autonomous devices, and *b)* adjust their behaviours in response to information (e.g. tariff and/or CO₂ emissions) feedback.

Finally, field trials to test the validity of this new more general multi-agent stochastic simulation framework are necessary to prove and compare real-world demand response scenarios with the results produced by our simulation platform.

Acknowledgement

Authors gratefully acknowledge the European Commission for providing financial support during the conduct of research under FP7-PEOPLE-2013 Marie Curie Initial Training Network "CI-ENERGY" project with Grant Agreement Number 606851.

References

- Chapman, J. (2016). *Multi-agent stochastic simulation of occupants in buildings*. Ph. D. thesis, University of Nottingham.
- Deng, R., Z. Yang, M.-Y. Chow, and J. Chen (2015). A survey on demand response in smart grids: Mathematical models and approaches. *IEEE Transactions on Industrial Informatics* 11(3), 570–582.
- Faria, P., Z. A. Vale, and J. Ferreira (2010). Demsia demand response simulator in the context of intensive use of distributed generation. In *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*, pp. 2025–2032. IEEE.
- Gudi, N., L. Wang, V. Devabhaktuni, and S. Depuru (2010). Demand response simulation implementing heuristic optimization for home energy management. In *North American Power Symposium (NAPS), 2010*, pp. 1–6. IEEE.

- Haldi, F. and D. Robinson (2009). Interactions with window openings by office occupants. *Building and Environment* 44(12), 2378–2395.
- Haldi, F. and D. Robinson (2010). Adaptive actions on shading devices in response to local visual stimuli. *Journal of Building Performance Simulation* 3(2), 135–153.
- Iacovella, S., P. Vingerhoets, G. Deconinck, N. Honeth, and L. Nordstrom (2016). Multi-agent platform for grid and communication impact analysis of rapidly deployed demand response algorithms.
- Jaboob, S. (2015). *Stochastic Modelling of Occupants' Activities and Related Behaviours*. Ph. D. thesis, University of Nottingham.
- Kersting, W. H. (2012). *Distribution system modeling and analysis*. CRC press.
- Le Thellier, M., A. Sancho-Tomás, G. Sousa-Ramirez, and D. Robinson (2017). Creation of synthetic households for multi-agent stochastic simulation. In progress.
- McArthur, S. D. J., E. M. Davidson, V. M. Catterson, A. L. Dimeas, N. D. Hatziargyriou, F. Ponci, and T. Funabashi (2007). Multi-agent systems for power engineering applications - Part I: Concepts, approaches, and technical challenges. *IEEE Transactions on Power Systems* 22(4), 1743–1752.
- Mohsenian-Rad, A.-H. and A. Leon-Garcia (2010). Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE transactions on Smart Grid* 1(2), 120–133.
- Page, J., D. Robinson, N. Morel, and J.-L. Scartezzini (2008). A generalised stochastic model for the simulation of occupant presence. *Energy and buildings* 40(2), 83–98.
- Pholboon, S., M. Sumner, E. Christopher, and S. A. Norman (2015). Real-time battery management algorithm for peak demand shaving in small energy communities. In *Innovative Smart Grid Technologies Latin America (ISGT LATAM), 2015 IEEE PES*, pp. 19–24. IEEE.
- Reinhart, C. F. (2004). Lightswitch-2002: a model for manual and automated control of electric lighting and blinds. *Solar Energy* 77(1), 15–28.
- Salom, J., J. Widén, J. Candanedo, I. Sartori, K. Voss, and A. Marszal (2011). Understanding net zero energy buildings: evaluation of load matching and grid interaction indicators. In *proceedings of building simulation*, Volume 6, pp. 2514–2521.
- Sancho-Tomás, A., M. Sumner, and D. Robinson (2017). A generalised model of electrical energy demand from small household appliances. *Energy and Buildings* 135, 350 – 366.
- Sutton, R. S. and A. G. Barto (1998). *Reinforcement learning: An introduction*, Volume 1. MIT press Cambridge.
- Vardakas, J. S., N. Zorba, and C. V. Verikoukis (2015). A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms. *IEEE Communications Surveys & Tutorials* 17(1), 152–178.
- Vasirani, M., R. Kota, R. L. Cavalcante, S. Ossowski, and N. R. Jennings (2013). An agent-based approach to virtual power plants of wind power generators and electric vehicles. *IEEE Transactions on Smart Grid* 4(3), 1314–1322.
- Watkins, C. J. C. H. (1989). *Learning from delayed rewards*. Ph. D. thesis, University of Cambridge England.
- Wijaya, T. K., D. Banerjee, T. Ganu, D. Chakraborty, S. Battacharya, T. Papaioannou, D. P. Seetharam, and K. Aberer (2013). Drsim: A cyber physical simulator for demand response systems. In *Smart Grid Communications (Smart-GridComm), 2013 IEEE International Conference on*, pp. 217–222. IEEE.