



# Consumption modeling based on Markov chains and Bayesian networks for a demand side management design of isolated microgrids

Tomislav Roje, Luis G. Marín, Doris Sáez<sup>\*,†</sup>, Marcos Orchard and Guillermo Jiménez-Estévez

Department of Electrical Engineering, University of Chile, Av. Tupper 2007, 8370451 Santiago, Chile

## SUMMARY

This paper proposes a novel simulator of energy consumption patterns that allows designing demand side management (DSM) strategies without economic incentives. The simulator emulates consumers' patterns with and without installed DSM interfaces, based on both actual consumption measurements and surveys applied to the inhabitants of an existing isolated microgrid (Huatacondo, Chile) that has a particular DSM strategy without economic incentives. The simulator uses Markov chains to generate data characterizing consumption patterns without DSM and Bayesian networks for cases in which the users respond to the DSM strategy. Data obtained from the simulator are used to derive a response model of the consumers to the DSM interface, which can be included for the energy management system design. Results show that the implemented strategy can be effective and can generate savings up to 4.45% in diesel consumption for an ideal case where all the dwellings have the interface installed. Copyright © 2016 John Wiley & Sons, Ltd.

## KEY WORDS

demand side management; microgrid; Markov chain; Bayesian network

## Correspondence

\*Doris Sáez, Department of Electrical Engineering, University of Chile, Av. Tupper 2007, 8370451 Santiago, Chile.

†E-mail: dsaez@ing.uchile.cl

Received 3 February 2016; Revised 17 June 2016; Accepted 27 June 2016

## 1. INTRODUCTION

Massive integration of distributed energy resources (DERs) arises as a feasible alternative to comply with power and energy requirements in developing countries. In this context, microgrids appear as the more suitable tool to incorporate the DER potential, contributing to the solution of challenges associated with rural electrification. Moreover, back in 2014, there were 79 electrically isolated communities with the potential to become isolated microgrids solely in Chile [1].

Microgrids are primarily considered as electrical distribution systems with DER that can be controlled and coordinated [2]. They may work either in a connected mode (where energy can be bought and sold to the external grid) or in a disconnected (or islanded) mode [3]. In the latter, the accurate management of limited resources is a challenge by itself. Microgrid operation and control obeys to a hierarchical control structure, where three main control levels are identified: primary, secondary, and tertiary control [4]. Primary control relates to short-term control actions to avoid situations in which either voltage or frequency values go beyond a safety range. Tertiary control, on the contrary, is associated to long-term actions that

frequently involve the interaction with the distribution network operator. Secondary control falls in between, and its main purpose is to (i) ensure that frequency and voltage specifications are met and (ii) to provide DER with set points guaranteeing both security and economic criteria. Tasks associated to the secondary control are performed by the energy management system (EMS) [3]. There are different aspects that may contribute to improve the EMS performance and, consequently, optimize the microgrid operation. One of these aspects is to incorporate a demand side management (DSM) strategy [5], which seeks to influence the users' energy consumption, helping to generate desired changes in the load profile [6].

According to Harper [7], DSM strategies can be divided in five categories according to their purpose. In the specific case of DSM in microgrids, incentive by price is the most reported class. Its subcategories are described by Han and Piette, Palensky and Dietrich [5,8], and Strbac [9] as direct load control, interruptible rates, time-of-use rates, and real-time pricing.

Less studied are DSM strategies that are relative to community involvement, either in terms of education about energy saving and its effects or information about the consumption associated with various appliances and microgrid

limitations. In fact, Allcott and Allcott and Mullainathan [10,11] show that behavioral interventions are effective, decreasing energy consumption in 2%, being comparable to price incentive initiatives. Also, Ayres and Raseman [12] studied the effect of a non-pecuniary strategy of energy conservation, comparing the peer consumptions, to reduce the energy consumption between 1.2 and 2%.

An important issue for DSM strategies is to model the consumers' response, allowing to determine the proper signal to be sent to the customer to generate the desired response. In Ref. [13], an accurate consumption model is required to decrease operational costs of the microgrid, but the biggest barrier to accomplish that is the consumers' random behavior toward the DSM strategy. In this context, different research works propose developing load profile simulators to characterize both microgrid operation and users' response to the DSM. Bottom-up approaches are the most frequently used to generate load profiles. In Ref. [14], devices were randomly added to the houses, following a given probability distribution and a characterization of their utilization. This probability distribution was obtained from vast (hourly) datasets as case studies. The generated profiles correlated well with the real data. A similar study is carried out in Ref. [15] for three households. The authors in [16] utilized probabilistic approaches for load utilization to generate synthetic consumption profiles of households, being characterized by socio-economic indicators.

Approaches with Markov chains (MC) were used in Refs. [17–22] to construct occupancy profiles for the residents of a household. The characterization of the MC considers that the number of active (in-home and not sleeping) occupants in the dwelling and their activities was registered. These profiles were utilized to generate lighting demand [18,19] and usage profile for different appliances. By associating related devices to these activities, a load profile was obtained [20,21]. These methods require a huge amount of detailed registry from the inhabitants' activities to build a load profile, information that is not usually available for the specific case studies and, furthermore, can be invasive to the population. On the other hand, to simulate and model users' responses to a DSM strategy, most efforts focus on economic incentives and the associated technologies. The authors [23–26] studied the response using technologies that apply direct load control, resulting in an effective approach to change the consumption profiles. In Refs. [27–30], the authors used the electricity price as DSM strategy, calculating the elasticity of the users' electricity demand. Finally, the authors [31,32] utilized an electricity price-based DSM but modeling the user as an agent looking to reduce their expenses by shifting the use of their appliances. These research efforts fail to consider the users' behavior and assume an ideal response.

In this regard, this paper proposes a new consumption model for the design of DSM strategies, aiming to estimate how the users respond to a specific DSM interface instead of simply guessing it. Contrary to other consumption

models, it does not incorporate economic incentives but considers dedicated and constant educational efforts on topics related to energy efficiency and the importance of the isolated microgrid for the community. This strategy was implemented in the microgrid of Huatacondo [33,34], described in detail in Section 2.2. Actually, this approach aims at solving one of the main disadvantages found in current EMS implementation: The system assumes that it is possible to vary the user's consumption by a constant amount during the whole day, instead of acknowledging possible modifications in the consumption profile on an hourly basis. Thus, it is important to generate community's consumption data both with and without implemented DSM strategies. For the generation of consumption profiles without DSM, MCs are utilized. These MCs are characterized by measured data of the village and certain dwellings. To generate the responses to a given DSM signal at specific time instant, Bayesian networks (BNs) are used. BN training is achieved through the use of surveys applied to the population, to characterize the use of certain electric devices in an hourly schedule. From the generated consumption profiles (with and without DSM), a load shifting factor is calculated as a ratio between both consumptions. The dataset of the load shifting factors is utilized to develop a response model to be included in the EMS as the expected users' response range, providing a realistic range to optimize the microgrid operation.

The rest of the paper is organized as follows: In Section 2, a brief description of the isolated microgrid is provided. Section 3 proposes the structure and method utilized to develop the consumer's simulator. Section 4 specifies the response model obtained with data, to be included in the offline EMS with the specific DSM strategy. Section 5 presents the results of implement the users' response model to the offline EMS. Finally, Section 6 provides the main conclusions and contributions of this work.

## 2. ISOLATED MICROGRID DESCRIPTION

This work is developed using the isolated microgrid installed in Huatacondo village (20°55'36.37"S69°3'8.71"W), located at the Atacama Desert, Chile [35]. This community has 25 dwellings permanently occupied with nearly 70 inhabitants.

Prior to the installation of the microgrid, the village was supplied with only 10 h a day of electricity by a diesel generator. Nowadays, it has 24-h energy supply, thanks to the use of renewable energies. The microgrid is composed of photovoltaic generation (23 kW), wind generation (3 kW), an energy storage system (ESS; 140 kWh), and a diesel generator (150 kVA).

### 2.1. Energy management system

The Huatacondo microgrid includes an EMS that coordinates and optimizes the operation of generation units and

loads, minimizing the operating costs, non-delivered power, losses, and non-delivered water, as shown in the following expression [35]:

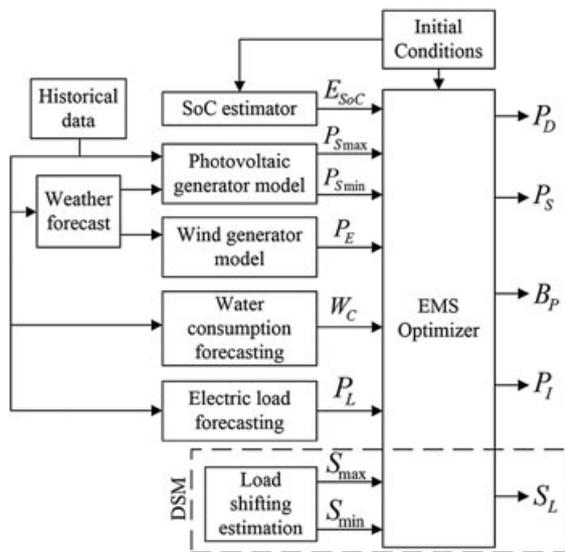
$$J = \delta_k \sum_{k=1}^T C(k) + \sum_{k=1}^T C_S(k) + C_{US} \delta_k \sum_{k=1}^T P_{US}(k) + C_{Tr} \sum_{k=1}^T V_{Tr}(k) + C_H(T) , \quad (1)$$

where  $T$  is the prediction horizon,  $\delta_k$  is the duration of time period,  $C(k)$  is the operational cost of diesel generator,  $C_S(k)$  is the start-up cost of diesel generator,  $C_{US}$  is the price for unserved energy,  $P_{US}(k)$  is the unserved power in the system,  $C_{Tr}$  is the cost of unserved water,  $V_{Tr}(k)$  is the volume of unserved water, and  $C_H(T)$  is the cost of using the ESS. The mixed integer programming optimization problem is solved using the commercial package CPLEX. [33]

From Figure 1, the inputs to the EMS are the state of charge ( $E_{SoC}$ ) estimator, minimum and maximum attainable solar power ( $P_{S_{min}}$  and  $P_{S_{max}}$ ), wind power forecasting ( $P_E$ ), water consumption forecasting ( $W_C$ ), load forecasting ( $P_L$ ), as well as minimum and maximum expected load shifting ( $S_{min}$  and  $S_{max}$ , respectively) as part of DSM.

Outputs from the EMS are the diesel set point power ( $P_D$ ), solar power ( $P_S$ ), a signal to the water pump ( $B_P$ ), the ESS inverter power ( $P_I$ ), and the desired load shifting factor ( $S_L$ ).

The solar power is controlled by the east-to-west inclination angle of the panels ( $\alpha$ ). The maximum power  $P_{S_{max}}$  is obtained by the optimal orientation of the PV panels. Sometimes, it is not feasible to use the maximum solar power because, for example, the ESS is completely charged. To prevent this problem, the minimum power  $P_{S_{min}}$  is obtained when the photovoltaic panels are oriented



**Figure 1.** Block diagram for the energy management system installed in Huatacoondo village.

to the minimum irradiance subject to the panel structure physical constraints. Therefore, the solar power  $P_S$  is obtained from the EMS, where it limits to a value between  $P_{S_{min}}$  and  $P_{S_{max}}$ .

### 2.2. Demand side management

The load shifting coefficients  $S_L(k)$ , provided by the EMS optimizer, are the desired variation of the consumer power consumption. Therefore, the expected load for the EMS optimizer is [35]:

$$\tilde{P}_L(k) = S_L(k)P_L(k) , \quad (2)$$

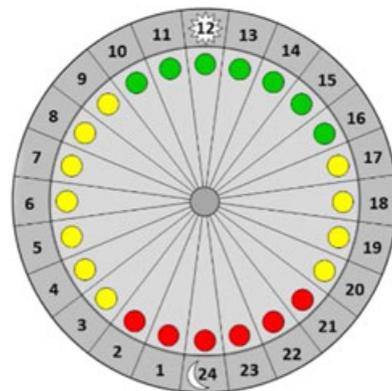
where  $P_L(k)$  is the load forecasting input.

The DSM strategy applied in Huatacoondo does not have economic incentives. It uses an interface (Figure 2) that shows light colors to the consumers, at each hourly division, based on the shifting coefficients  $S_L$ . These coefficients are calculated by the EMS in the optimization process, indicating users how to modify their electric consumption to adjust to the available resources.

To limit the expected consumers' increase or decrease response to the signals given by the interface, the shifting factor is bounded with limits for the consumption increase or decrease:

$$S_{min}(k) \leq S_L(k) \leq S_{max}(k) . \quad (3)$$

These values of  $S_{min}$  and  $S_{max}$  are arbitrarily fixed in the actual implementation at Huatacoondo microgrid, for example,  $[S_{min}, S_{max}] = [0.95, 1.05]$ , indicating that it can be expected up to 5% of variation over the forecasted consumption. Also, it is assumed that the daily energetic consumption with DSM remains constant for the whole optimization period, with respect to the base case [33]:



**Figure 2.** Implemented demand side management interface in Huatacoondo village.

$$\sum_{k=T_1}^{T_2} P_L(k) = \sum_{k=T_1}^{T_2} \tilde{P}_L(k), \quad (4)$$

where  $T_1$  and  $T_2$  define the optimization period, with a sampling time of 15 min, given by the EMS operation.

The expected consumer response should be to increase, maintain, or decrease the consumption when a green, yellow, or red light is shown, respectively.

To determine the interface light color from  $S_L(k)$  calculated by the EMS, the values of  $S_{min}$  and  $S_{max}$  are used as follows:

$$\text{Light}(S_L(k)) = \begin{cases} \text{Green} & \text{if } S_L(k) \geq 1 + \frac{S_{max} - 1}{2} \\ \text{Yellow} & \text{if } S_{min} + \frac{1 - S_{min}}{2} \leq S_L(k) < 1 + \frac{S_{max} - 1}{2} \\ \text{Red} & \text{if } S_L(k) < S_{min} + \frac{1 - S_{min}}{2}. \end{cases} \quad (5)$$

### 2.3. Problem statement

Currently, the consumer response to the DSM strategy is implemented, considering that the bounds for the shifting factors  $S_{min}(k)$  and  $S_{max}(k)$  are arbitrarily fixed and constants over the optimization period (a day). Effectively, these values should be selected according to the consumer behavior and therefore should be modeled as time-variant signals.

This work presents the development of a consumption simulator, based on measurements and surveys conducted at homes in Huatacondo, which allows us to obtain the associated response model to be included as an input to the EMS. This simulator uses MC and BN to characterize the consumer's behavior according to the DSM interface, which in turn provides a more realistic response for the load shifting factors at every sampling time.

## 3. PROPOSED CONSUMPTION SIMULATOR

The proposed simulator allows generating consumption data with, and without, considering the DSM strategy. For the generation of base consumption patterns (without response), the use of MCs is proposed. The generation of responses to the DSM interface is performed using BNs.

The structure of the simulator of consumption patterns is shown in Figure 3. On the one hand, the inputs are the time  $k$  (sampling time of 15 min) and the associated color  $c$  of the DSM interface. On the other hand, the outputs are the consumption for the entire village, considering two different scenarios: with and without response to the DSM, respectively. The base consumption  $P_L^d(k)$  for each dwelling  $d$  is given by the realizations of the MC, using the information stored in state transition matrices. Similarly, the power consumption variation  $\Delta P_L^d(k, c)$ , given the hour and color, is generated with BNs for each dwelling, considering for this purpose the information stored in

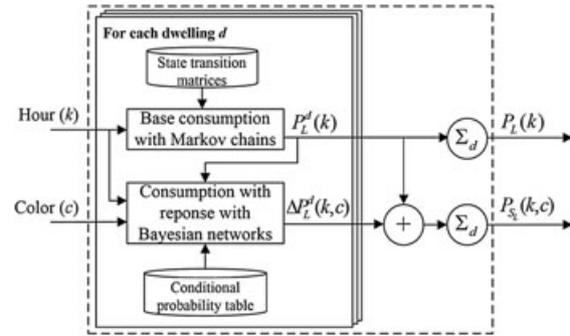


Figure 3. Outline structure of the simulator.

conditional probability tables. Adding up all the base consumptions per dwelling, the base consumption for the village  $P_L(k)$  is obtained. If all variations  $\Delta P_L^d(k, c)$  are added to the previous base consumption, then the total consumption for the village with demand response  $P_{S_L}(k, c)$  is obtained.

### 3.1. Available data

Data used to develop the simulator correspond to energy consumption observations ([Wh], each 15 min) measured at 20 dwellings (i.e., 96 samples a day) for 69 days. Based on these data, the average power consumption (measured in  $W$ ) is obtained at each sample time. In addition, the total load of the entire community is available for the same period.

To model the response to the DSM interface, 41 surveys were carried out to 14 dwellings. In these surveys, the users were asked for their typical consumption without DSM, and after that, their consumption with four sequences of DSM interface were presented. Two of the sequences represent typical sequences given by the optimizer for a sunny and a cloudy day, while the other two represent sequences for atypical days. Knowing the behavior of the users in typical and atypical days allows estimating the change in consumption. The users responded how they would use their devices if they had to follow the DSM interface signals at each hourly block between 08:00 and 24:00h, as shown in Table I. Comparing the behavior between a case without DSM and a case with DSM, a difference in the response can be calculated.

Table I. Consumption survey.

Hour	Fridge	TV	Electric oven	-	Iron
08:00	x	-	-	-	-
09:00	x	x	-	-	-
-	-	-	-	-	-
23:00	x	x	-	-	x
00:00	x	x	-	-	-

### 3.2. Generation of consumption based on Markov chains

The implementation of the proposed simulator requires the generation of data that could effectively be used to characterize future consumption patterns for the community. In this regard, we want to avoid simplistic procedures that merely use subsequent replications of measured data in time. Instead, a more sophisticated (and more precise) approach is chosen in this work that incorporates stochastic characterization of those consumption patterns. After exploring different alternatives, the discrete-time Markov chains (DTMCs) [36] have been chosen for this purpose. In these MCs, each state is associated to a specific consumption level, and thus, the uncertainty associated with consumption time series is characterized through the definition of appropriate transition probabilities between those states. As DTMCs can only model stationary processes, and considering the notorious differences observed in the acquired data in terms of average power consumption during extended (and regular) lapses throughout the day (Figure 4), a collection of different DTMCs has been chosen to explain variations around those average power consumption values.

Considering all of those in the preceding texts, additional assumptions considered for the implementation of this method are:

- Each realization of the DTMC can be used to generate data that will statistically represent future consumption patterns.
- For all practical purposes, no seasonal variations are considered, making the simulator valid for just one season.

After a careful analysis of average power levels exhibited throughout the day, three DTMCs are finally included in our simulator. Indeed, the day can be divided in three groups, limited by the times  $k_1$ ,  $k_2$ , and  $k_3$ , with possible

values between 1 and 96 (given by the number of samples a day), and represents roughly low, medium, and high consumption levels. Figure 4 illustrates the proposed method to generate base consumption patterns with MCs.

The definition of appropriate parameters for each MC is not an easy task. Firstly, the most adequate number of states to be used must be found. Secondly, transition probabilities must be computed. As the precision related to the estimation procedure is a function of the number of available data samples, it is important to avoid an excessive number of parameters (which translates into poor estimation results).

As shown in Ref. [37], the maximum likelihood estimator of the transition probability from state  $i$  to state  $j$  is equal to:

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^m n_{ij}}, \tag{6}$$

where  $n_{ij}$  is the number of measured data that make a transition from state  $i$  to state  $j$ , and  $m$  is the total number of states in the chain.

To discretize the power in states, represented as  $1 \dots m_i$  in Figure 4, the use of the k-mean clustering is proposed, where each clusters centroid corresponds to a state of the MC. As a result, each power measure is associated with one specific cluster and thus with a state of the MC. This information is then used to determine the moments in which the system makes a transition between those states. Finally, maximum likelihood estimate of transition probabilities are computed, considering the total number of transitions between states.

Regarding the determination of the number of states for the MC, Navarrete [38] presents a method to calculate an upper bound for maximum number of states, as a function of design parameters  $p^*$  and  $t$ . Parameter  $p^*$  represents the maximum probability that can be accepted for errors bigger than  $t$  for the maximum likelihood estimator of  $p_{ij}$ . Being  $n_i$  the total number of measured data that makes a transition

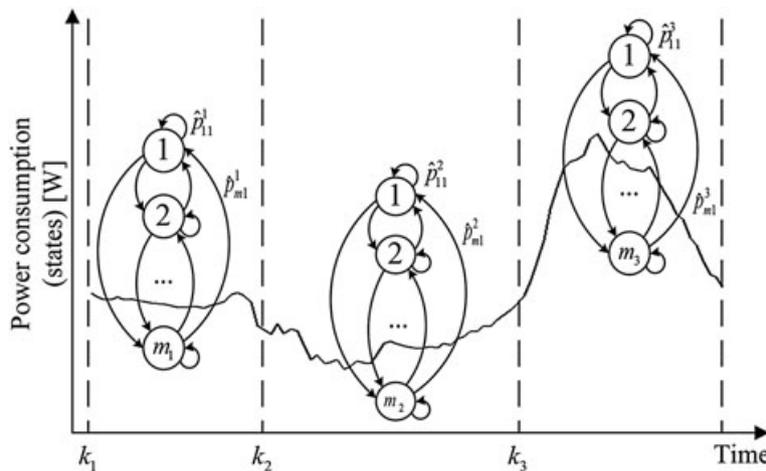


Figure 4. Markov chain structure to the generation of load profiles without demand response.

from state  $i$  to any state  $j$ , the upper bound to the probability  $p^*$ ,  $c(n_i)$ , is calculated, as shown in Equation (7).

$$c(n_i) = \min \left\{ 1, 2e^{-2n_i t^2}, \frac{1}{4n_i t^2} \right\}. \quad (7)$$

Through an iterative process, starting with a high number of states, it is tested if  $c(n_i) \leq p^*$  for the transitions from state  $i$ , for every state  $i$ ; in the case that this is not true for at least one  $i$  state, the number of states must be reduced by one. When  $c(n_i) \leq p^*$  is true for every state  $i$ , the process stops, and the maximum number of states  $m^*$  is obtained.

Given the fact that there is only few data available, the previous method can result in  $m^* = 1$  as the maximum number of states. To avoid this, it is convenient to group consumption data that belong to a particular time interval and have similar characteristics to characterize the MCs, as shown in Figure 4. The determined number of groups is 3, after the observational verification of the data, where high, medium, and low levels of consumption were observed.

To define these groups, given by the times  $k_1$ ,  $k_2$ , and  $k_3$ , the vector of characteristics  $x(k)$  is calculated at each sampling time  $k$  and corresponds to the minimum, the maximum, and the mean of the consumption at the time  $k$ . The hourly range is determined by minimizing the intragroup differences, looking to determine where the divisions between groups are, and to group similar data, to be able to be represented by the limited number of states of the MC.

Taking  $S_s$  as the group  $s$  of the data, with  $s = \{1, 2, 3\}$ , the vector of characteristics  $x(k)$  satisfies that  $x(k) \in S_s \Leftrightarrow k \in [k_s, k_{s+1} - 1]$ . In the case of  $S_3$ ,  $x(k) \in S_3 \Leftrightarrow k \in [k_3, 96] \cup [1, k_1 - 1]$ , because when the time is  $k = 97$ , the interval returns to  $k = 1$  (next day):

$$\bar{\mu}_s = \frac{1}{N_s} \sum_{x(k) \in S_s} x(k), \quad (8)$$

$$\operatorname{argmin}_{k_1, k_2, k_3} J = \sum_{j=1}^3 \sum_{x(k) \in S_s} \|x(k) - \bar{\mu}_s\|^2, \quad (9)$$

where  $N_s$  is the number of elements  $x(k)$  in the group  $s$ , and  $\bar{\mu}_s$  is the mean of all the  $x(k)$  that belong to the same group  $S_s$ .

The minimization of  $J$  is performed with genetic algorithms to avoid local minima. The chromosomes represent the parameters  $k_s$ .

The initial probability distribution of the MC states is obtained as a probability proportional to the number of data that belong to each state in the initial time value  $k_s$  of the range that defines the group  $S_s$ . This probability is calculated as:

$$p_i = \frac{n_i}{\sum_{j=1}^m n_{ij}}, \quad (10)$$

where  $p_i$  is the probability that the initial state of the chain

is the state  $i$ ,  $m$  is the number of states, and  $n_i$  is the number of measures that belongs to the state  $i$  in the time value  $k_s$ .

As a verification measure, the root-mean-square error (RMSE) of the means at each measure time for the generated and real consumption is used.

Over these base consumptions, the users' response, given the hour and color of the DSM interface generated using BNs, is applied.

### 3.3. Generation of consumption with response to the DSM based on Bayesian networks

Once the surveys are conducted, it is necessary to use the collected information to simulate the consumers' response to the DSM interface.

Each device indicated in the surveys has a typical mean power over a 1-h interval that can be seen in Table II. Comparing the base consumption from a survey with the consumption with DSM, the variation of consumption, given the hour and color for each load, can be obtained.

Bayesian networks are utilized to include the uncertainty of the users' behavior and to extend the responses to other hour and color combinations. The assumptions considered for this implementation are:

- The surveys represent faithfully the village behavior to a particular DSM strategy.
- Each device has its own pattern of use.
- The factors that alter the decision of the users are the hour of the day and the color displayed in the interface.
- The users' answers do not vary between weekdays and weekends.

For this application, each device has its own BN, under the supposition that each one has a different disposition to be turned on or off. Each network is obtained using all the responses given by the surveyed consumers for each device, and a reasoning is applied, where the causes are the evidences, and the consequences are generated, given the causes and the conditional probabilities [39].

**Table II.** Devices' typical consumption.

Load	Power consumption [W]	Load	Power consumption [W]
Refrigerator	81.9	Microwave	320
Electric oven	1300	Bulb	23
Boiler	72	Hair iron	75
TV	100	Freezer	90
Iron	600	Computer	300
Washing machine	182	Kitchen tool	200
Radio	60		

The resulting graph considers, as evidences, two multinomial nodes:  $k$  – representing the hour from 08:00 to 24:00 – and  $c$ , representing the color *green*, *yellow*, and *red*. As query, there are two binomial nodes:  $i$  – corresponding to the initial state (*on* or *off*) – and  $v$ , indicating if there is variation over the base state (*true* or *false*). It is supposed that the base state of a device depends solely on the hour, while the variation with respect to this base state depends on the hour, the color, and the base state itself. This is because a device can be turned on only if it was initially off and vice versa.

As asking for every possible hour and color combination is unpractical; thus, the structure of the BNs is given, being trained with incomplete data. To train the proposed BN, the *expectation maximization* algorithm described in Ref. [40] is utilized. This iterative algorithm converges to a *local maximum* of the likelihood function. Thus, it is necessary to provide an initial guess of the parameters, in this case, probabilities. These a priori probabilities are obtained by using the responses given by all the surveyed users to the survey (as in Table I). Once the parameters of the network are calculated, the obtained probabilities  $\mathbb{P}(on|k)$  and  $\mathbb{P}(var|k, c, i)$  are employed to make realizations of the network, to get generated data indicating the device behavior.

To allocate generation of data to the users' responses to the survey, concerning the number of devices that are expected to be turned on or off, the following procedure is proposed:

- Match the consumption given by the MCs with the base consumption given by the BN. To accomplish this, turn off the devices from the network until the consumptions match.
- To determine from the surveys how many devices can be responsive at the same time, let, at maximum, these  $n$  devices to be responsive.

By executing the previous steps, the response per device over a base case is calculated: After multiplying these responses by the typical power for the devices, and then adding up these results, a power consumption variation  $\Delta P_L^c(k, c)$  at each sample time  $k$  and color  $c$ , for the dwelling  $d$ , can be obtained.

## 4. PROPOSED CONSUMPTION MODEL FOR AN EMS DESIGN

A user response model is obtained by using the generated data from the developed simulator. This model of response is included in the EMS optimization process, by providing as inputs the expected response range.

### 4.1. Data generated by the simulator

With the consumptions with  $(P_{S_L}(k, c))$  and without  $(P_L(k))$  DSM, a load shifting factor that expresses how

much the consumption varies can be calculated as the ratio:

$$\rho_L(k, c) = \frac{P_{S_L}(k, c)}{P_L(k)}. \quad (11)$$

This factor depends on the time  $k$  and the color  $c$ . With that information, a dataset can be obtained from the data generated by the simulator to derive a model.

### 4.2. Consumption response model

The proposed model consists of a lookup table of size  $96 \times 3$ , where each cell indicates the expected response at a certain time and color, so it contains every time and color combination. The values of the lookup table are obtained by calculating the mean values at every hour and color combination, as shown in Equation (12):

$$\bar{\rho}_L(k_i, c_i) = \text{mean} \left( \sum_{k=k_i} \sum_{c=c_i} \rho_L(k, c) \right), \quad (12)$$

where  $\rho_L(k, c)$  is the simulator output,  $k_i = \{1 \dots 96\}$  is the sample time, and  $c_i = \{1, 2, 3\}$  is the color indicated by the DSM interface.

### 4.3. EMS based on a new model

As the EMS described in Section 2 considers as inputs values the maximum expected consumption increase or decrease for the optimization period, it is desirable to utilize the model to obtain these input vectors. The proposed method is straightforward and consists of getting the minimum and maximum estimated value of the calculated shifting factor at each time for the three possible colors of the light:

$$S_{\min}(k) = \min_{l=1,2,3} \bar{\rho}_L(k, c), S_{\max}(k) = \max_{l=1,2,3} \bar{\rho}_L(k, c). \quad (13)$$

Repeating the procedure for every sample, we obtain the input vectors to the EMS, simultaneously limiting the desired shifting factor, as shown in Equation (3).

## 5. RESULTS

The following results were obtained by applying the method described in Sections 3 and 4 for the community of Huatacocondo. The tests were performed in an offline EMS simulator, with real consumption, solar, and wind power data.

### 5.1. Simulator performance

The main objective of the simulator is to generate data with enough variability to calculate the values of the described model. To do that, it is necessary to obtain results from both the base consumption and the response to the DSM.

### 5.1.1. Base consumption with Markov chains

The available data are grouped as described in Section 3, allowing the characterization of three MCs for each dwelling that generates the base consumption.

For the number of state selection, the chosen design values are  $t=0.1$  and  $p^*=0.1$ . These values were selected with the intention of not having a high exigency for the number of states, given the restricted amount of data. Figure 5 shows the mean of 69 days of power consumption measured and generated by the proposed simulator for the community.

The performance metric to use for the base consumption evaluation is the RMSE, defined as:

$$RMSE = \frac{1}{N} \sum_{k=1}^N \left( \sqrt{(P_L(k) - P_g(k))^2} \right), \quad (14)$$

where  $P_L$  is the mean of the measured data, and  $P_g$  is the mean of the generated data with MCs;  $N$  is the number of data points (=96).

The RMSE between both consumptions is 0.829 kW, with a maximum mean consumption of the real data of 13.9 kW. However, these differences should not be considered negatively, because the purpose of the simulator is to generate data with enough variability to consider different cases. In both the real and generated consumptions, three levels of consumption can be identified – low, medium, and high – that coincide in time.

### 5.1.2. Response to the DSM with Bayesian networks

The dataset used to obtain the BNs correspond to the survey's results, where each data sample corresponds to the hour, the color, the initial state of the device in the base case, and the variation in the use. Two cases are studied: Case no. 1 = DSM interface installed in four dwellings. Case no. 2 = DSM interface installed in 25 dwellings.

The first case corresponds to the current case, while the second one considers the ideal case, where the dwellings that are occupied all the year have the DSM interface installed.

Figure 6 illustrates a realization of the simulator for both cases compared with the base case for a random sequence of the DSM interface in the range of hours where it was assumed that there is a user response. As reference, each month (30 days) of simulation takes 70 s.

As it was expected, the response magnitude in the ideal case is greater than the response in the current case. In general, the users' response follows the desired logic, where the green light stimulates the consumption increase, while the red light stimulates the decrease.

It can be seen that between the 13th and 14th hour of operation, the response is not intuitive, increasing the consumption when there is a yellow light indicated.

## 5.2. Model results for consumption response

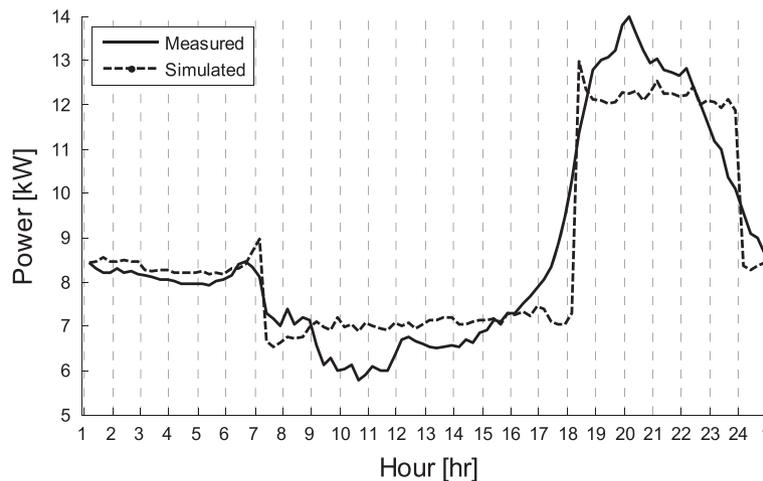
Having the data from the simulator, the lookup table can be calculated for both cases. The performance metric in this case is the RMSE, defined as:

$$RMSE = \frac{1}{N} \sum_{k=1}^N \left( \sqrt{(\rho_L(k, c) - \bar{\rho}_L(k, c))^2} \right), \quad (15)$$

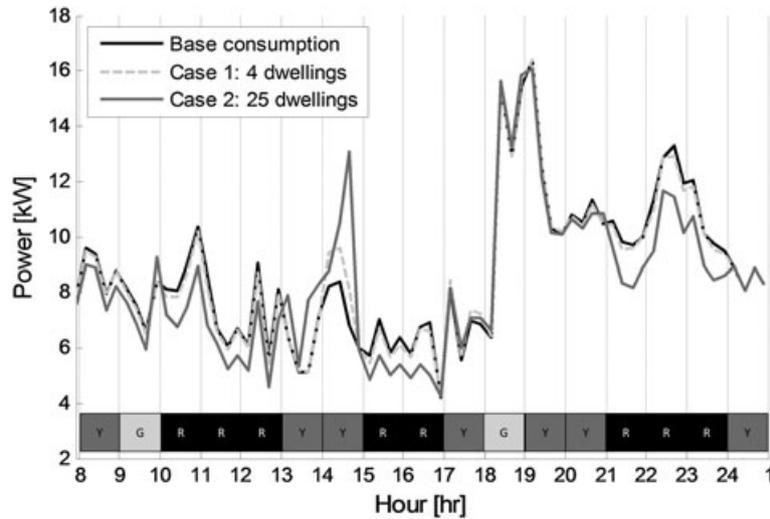
where  $\rho_L$  is the shifting factor obtained by simulations, and  $\bar{\rho}_L$  is the estimated shifting factor computed by lookup table (Equation (12));  $N$  is the number of samples.

The RMSE between the target output and the load shifting output (dimensionless) of this model is calculated for both cases: 0.0311 for Case no. 1 (four dwellings) and 0.0823 for Case no. 2 (25 dwellings).

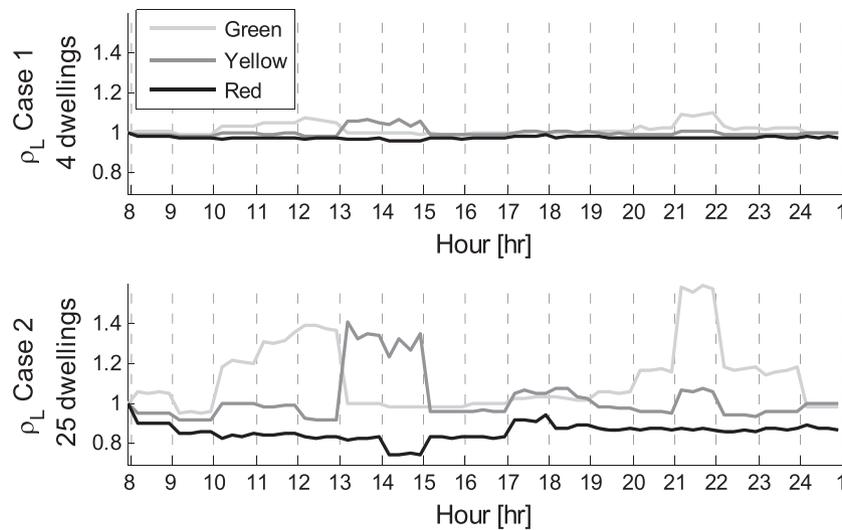
Figure 7 shows the expected response given by the model for every color and sample time.



**Figure 5.** Average base consumption measured and simulated for the village. Root-mean-square error = 0.829 kW and maximum real power 13.9 kW.



**Figure 6.** Village’s power consumption with demand side management in a base case, actual case, and ideal case (G, green light; Y, yellow light; and R, red light).



**Figure 7.** Responses for each color during a day from the model simulation. Green line, shifting factor for green light; yellow line, shifting factor for yellow light; red line, shifting factor for red light.

As expected, the users’ load shifting factor is noticeably higher in the ideal case, where all the usually occupied dwellings have the DSM interface installed versus the current case. In general, the users’ responses make sense, increasing, maintaining, or decreasing their consumption when the light is green, yellow, or red, respectively.

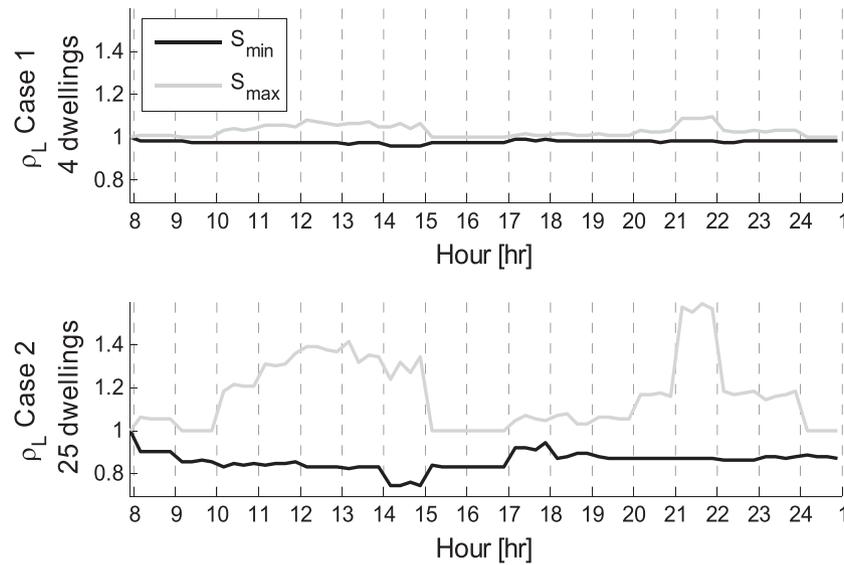
Unexpectedly, in Case no. 2, slightly between 17–19 h and more notoriously between 13–15 h, there is an increase in the consumption for the yellow light, which in fact is even higher than the response observed for the green light. This response reflects the answers given in the surveys, so it is supposed that not all the users understand completely the functioning of the DSM interface, which aims to shift the loads from red to green lights and not to yellow lights. That indicates that further educational work is needed.

### 5.3. EMS performance based on a new model

Once the model is obtained, the inputs to the EMS are calculated as the minimum/maximum expected response at each sample time, for the three possible lights, as in Equation (13). Figure 8 shows the obtained range to be provided as inputs to the EMS.

Providing the obtained  $S_{min}$  and  $S_{max}$  vectors to the offline EMS optimizer for both cases, tests are performed, consisting in simulate the EMS microgrid, with actual data of wind and solar power, and forecasted consumption data, with the obtained load shifting factors as inputs.

Wind, solar power, and forecasted consumption data are used as inputs for the microgrid simulator, including  $S_{min}$



**Figure 8.** Response range ( $S_{\min}$  and  $S_{\max}$ ) expected for both actual and ideal cases.

and  $S_{\max}$  factors at each hourly interval. To evaluate the performance of the EMS with the new model, two microgrid topologies were used. The first one considers the microgrid with photovoltaic, wind, and diesel generation, while the second one considers only photovoltaic and diesel generation. This choice is based on the fact that the wind energy converter is not always working in the Huatacondo microgrid due to technical issues. The analysis of actual data during 30 days of operation for each topology helped to characterize diesel consumption variations for different  $S_{\min}$  and  $S_{\max}$  factors.

Table III shows the average costs of diesel consumption in a case without DSM strategy implemented and for two cases of the installed interface with four dwellings (Case no. 1) and 25 dwellings (Case no. 2).

The savings in diesel costs are respect to a base case without a DSM strategy. Comparisons with cases in which the  $[S_{\min}, S_{\max}]$  range is constant for every measure time and equal to  $[0.95, 1.05]$  and  $[0.9, 1.1]$  for the 5 and 10% of maximum response cases, respectively, are considered.

As the optimizer considers that the actual power consumption is the expressed in Equation (2), these results assume that the desired shifting factor can be achieved by the community.

**Table III.** Average diesel costs (CLP).

	Solar/wind/ diesel	Average savings	Solar/diesel	Average savings
Base	\$16 487	–	\$20 909	–
5%	\$16 098	2.36%	\$20 527	1.83%
10%	\$15 505	5.96%	\$20 199	3.40%
Case 1	\$16 367	0.72%	\$20 830	0.38%
Case 2	\$15 753	4.45%	\$20 163	3.57%

The results show that a saving up to 4.45% in the diesel costs can be expected in the ideal case where all the occupied houses respond to the DSM interface, and there is wind power available. In the microgrid topology without wind power, up to 3.57% of savings are expected in the ideal case. The results show that in an ideal case, where all the dwellings have the DSM interface installed, and wind and solar power are available, we can expect up to 4.45% of savings in diesel consumption. In contrast, up to 0.72% of savings are expected in the current case, with the DSM interface installed in four dwellings. These savings are lower than the case in which a constant 10% of variation is assumed: That difference depends highly in the availability of the energy over the day and the solution space for the optimization problem given by the  $[S_{\min}, S_{\max}]$  range, letting the desired  $S_L$  factor to take values that minimize the operational costs. The situation is similar with the second microgrid topology (without available wind power), except that the ideal case has a higher expected saving compared with the case that considers 10% of variation, also highly dependent of the energy availability and the solution space given.

Although the results with a constant shifting factor range generate higher savings in the diesel costs, these factors do not incorporate a study of how the users would respond and would therefore be unrealistic. In addition, these results show that a DSM without economic incentives can lead to effective saving, helping to match the generation with the consumption, always taking care of the educational work with the communities.

## 6. CONCLUSIONS

This paper presents a simulator for the generation of load profiles in both cases where DSM strategies have not been

implemented and cases where a particular DSM strategy without economic incentives is installed. The simulator was developed by using MCs, maintaining the consumption characteristics of the Huatacondo community. To represent changes in consumption patterns when using a DSM interface, BNs were used under the assumption that faithfully represents the users' expected behavior. The resulting dataset allows to estimate the degree of response of the community, showing that users' responses are generally reasonable, having a tendency to decrease their consumption with a red light and increase it with green lights. Nevertheless, in some cases, the users did not understand completely the purpose of the DSM interface, even increasing their consumption if a yellow light is shown. This behavior should be corrected with more educational work for the good performance of the DSM strategy.

Due to the structure inputs to the EMS optimizer, a model consisting in a lookup table can be calculated from the generated dataset, providing the required vector inputs to the optimizer.

The results show that the interface can generate response in the users, allowing to save up to 4.45% in diesel consumption in the ideal case. This in contrast with the cases with fixed expected response, with up to 5.96% of savings, but with unrealistic ranges and without a study of the users' response.

These results allow the estimation of how the users would respond in an ideal case, improving the DSM strategy, sending requisite signals to the users, and obtaining the desired response in the community. Also, the results show that a community involvement-based DSM strategy can change the users' pattern consumption in a community educated energetically.

As future work, the most promising feature is to explicitly include a response model into the optimization process, to obtain the colors to be shown by the DSM interface. Another interesting improvement is to consider the  $S_L$  factor as a discrete rather than a continuous optimization variable, due to the fact that the  $S_L$  factor values are associated to the discrete signals shown by the DSM interface, which cannot take intermediate light colors to generate every desired response in the continuum range  $[S_{min}, S_{max}]$ . Also, it must be noted that the inclusion of information related to previous color transitions can improve the characterization of the users' response. However, the inclusion of these transitions in our model requires the application of new surveys, as well as modifications to the EMS. This task is currently considered part of of future research work.

## ACKNOWLEDGEMENTS

This research was partially funded by the Complex Engineering Systems Institute, ISCI (ICM-FIC: P05-004-F, CONICYT: FB0816) FONDECYT project 1140775, and CONICYT/FONDAP 15110019. Luis G. Marín has been supported by a PhD scholarship from CONICYT-PCHA/Doctorado Nacional para extranjeros/2014-63140093.

## REFERENCES

1. Fundación Chile (2014) Almacenamiento de energía para comunidades aisladas [December 28th 2015]; Available from: <http://www.fch.cl/comunidades-aisladas-podran-acceder-a-energia-de-bajo-costo-y-autogestionada/>.
2. Ubilla K, Jimenez-Estevez GA, Hernandez R, *et al.* Smart microgrids as a solution for rural electrification: ensuring long-term sustainability through cadastre and business models. *IEEE Transactions on Sustainable Energy* 2014; **5**(4):1310–1318.
3. Lasseter RH. MicroGrids. *Power Engineering Society, IEEE Winter Meeting* 2002; **1**:305–308.
4. Olivares DE, Mehrizi-Sani A, Etemadi AH, *et al.* Trends in microgrid control. *IEEE Transactions on Smart Grid* 2014; **5**(4):1905–1919.
5. Han J, Piette M a. Solutions for summer electric power shortages : demand response and its applications in air conditioning and refrigerating systems. *Refrigeration Air Conditioning Electrical Power Machines* 2008; **29**(1):1–4.
6. Gellings CW. The concept of demand-side management for electric utilities. *Proceedings of the IEEE* 1985; **73**(10):1468–1470.
7. Harper, M. (2013) Review of strategies and technologies for demand-side management on isolated mini-grids. (March).
8. Palensky P, Dietrich D. Demand side management: demand response, intelligent energy systems, and smart loads. *IEEE Transactionson Industrial Informatics* 2011; **7**(3):381–388.
9. Strbac G. Demand side management: benefits and challenges. *Energy Policy* 2008; **36**:4419–4426.
10. Allcott H. Social norms and energy conservation. *Journal of Public Economics* 2011; **95**(9-10):1082–1095.
11. Allcott H, Mullainathan S. Behavior and energy policy. *Science* 2010; **327**(5970):1204–1205.
12. Ayres I, Raseman S. Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. 5th Annu. Conf. Empir. Leg. Stud. Pap, 2009.
13. Khan AR, Mahmood A, Safdar A, *et al.* Load forecasting, dynamic pricing and DSM in smart grid: a review. *Renewable and Sustainable Energy Reviews* 2016; **54**:1311–1322.
14. Paatero JV, Lund PD. A model for generating household electricity load profiles. *International Journal of Energy Research* 2006; **30**(April 2005):273–290.
15. Armstrong MM, Swinton MC, Ribberink H, *et al.* Synthetically derived profiles for representing occupant-driven electric loads in Canadian housing. *Journal of Building Performance Simulation* 2009; **2**(2014):15–30.

16. Fischer D, Härtl A, Wille-Haussmann B. Model for electric load profiles with high time resolution for German households. *Energy and Buildings* 2015; **92**:170–179.
17. Richardson I, Thomson M, Infield D. A high-resolution domestic building occupancy model for energy demand simulations. *Energy and Buildings* 2008; **40**:1560–1566.
18. Richardson I, Thomson M, Infield D, Delahunty A. Domestic lighting: a high-resolution energy demand model. *Energy and Buildings* 2009; **41**:781–789.
19. Widén J, Nilsson AM, Wäckelgård E. A combined Markov-chain and bottom-up approach to modelling of domestic lighting demand. *Energy and Buildings* 2009; **41**:1001–1012.
20. Richardson I, Thomson M, Infield D, Clifford C. Domestic electricity use: a high-resolution energy demand model. *Energy and Buildings* 2010; **42**(10): 1878–1887.
21. Widén J, Wäckelgård E. A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy* 2010; **87**(6):1880–1892.
22. Widén J, Lundh M, Vassileva I, *et al.* Constructing load profiles for household electricity and hot water from time-use data—modelling approach and validation. *Energy and Buildings* 2009; **41**:753–768.
23. Ruiz N, Cobelo I, Oyarzabal J. A direct load control model for virtual power plant management. *IEEE Transactions on Power Apparatus and Systems* 2009; **24**(2):959–966.
24. Zaidi AA, Zia T, Kupzog F. Automated demand side management in microgrids using load recognition. *IEEE Int. Conf. Ind. Informatics*, 2010; 774–779.
25. Kallel R, Boukettaya G, Krichen L. Demand side management of household appliances in stand-alone hybrid photovoltaic system. *Renewable Energy* 2015; **81**:123–135.
26. Zhu L, Yan Z, Lee W-J, *et al.* Direct load control in microgrids to enhance the performance of integrated resources planning. *IEEE Transactions on Industry Applications* 2015; **9994**(c):1–1.
27. David AK, Li YZ. Consumer rationality assumptions in the real-time pricing of electricity. *IEE Proc. C Gener. Transm. Distrib.*, 1992; 139 (November), 315.
28. Kirschen DS, Strbac G, Cumperayot P, De Mendes DP. Factoring the elasticity of demand in electricity prices. *IEEE Transactions on Power Apparatus and Systems* 2000; **15**(2):612–617.
29. Venkatesan N, Solanki J, Solanki SK. Market optimization for microgrid with demand response model. *NAPS 2011 — 43rd North Am. Power Symp* 2011.
30. Basu aK, Panigrahi TK, Chowdhury S *et al.* Key energy management issues of setting market clearing price (MCP) in micro-grid scenario. *Proc. Univ. Power Eng. Conf.*, 2007; (1): 854–860.
31. Nunna HSVSK, Member S, Doolla S. Demand response in smart distribution system with multiple microgrids. *IEEE Transactions on Smart Grid* 2012; **3**(4):1641–1649.
32. Sesetti A, Battola S, Nunna HSVSK, Doolla S. Intelligent agent framework for demand response aggregation in smart microgrids. *IECON 2014 — 40th Annu. Conf. IEEE Ind. Electron. Soc.* 2014; 3549–3555.
33. Palma-Behnke R, Benavides C., Aranda E, *et al.* Energy management system for a renewable based microgrid with a demand side management mechanism. *IEEE SSCI 2011 — Symp. Ser. Comput. Intell. — CIASG 2011 2011 IEEE Symp. Comput. Intell. Appl. Smart Grid*, 2011; 131–138.
34. Davito B, Tai H, Uhlaner R. The smart grid and the promise of demand-side management. *McKinsey Smart Grid* 2010:38–44.
35. Palma-Behnke R, Benavides C, Lanas F, *et al.* A microgrid energy management system based on the rolling horizon strategy. *IEEE Transactions on Smart Grid* 2013; **4**(2):996–1006.
36. Taylor HM, Karlin S. Markov chains: introduction in *An introduction to stochastic modeling*, 1984; 95–99.
37. Anderson TW, Goodman LA. Statistical inference about Markov chains. *Annals of Mathematical Statistics* 1957:89–110.
38. Navarrete H. Caracterización estadística del perfil de uso de baterías para el pronóstico del estado de carga. Undergraduate Thesis Project, Electrical Engineering, 2014.
39. Ben-Gal I. Bayesian Networks. In *Encyclopedia of Statistics in Quality and Reliability*, Ruggeri F, Faltin F, Kennet R (eds.). Wiley & Sons: Chichester, UK, 2007.
40. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 1977; **39**(1):1–38.