A generalized economic model for optimally selecting forecasted load profiles for measuring demand response in residential energy management system

Peter Jean-Paul[1], Tek-Tjing Lie[1], Timothy N. Anderson[1], Brice Vallès[2]
[1]School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, New Zealand
[2]Private Consultant, New Zealand
Email: peter.jeanpaulnz@gmail.com, tek.lie@aut.ac.nz, timothy.anderson@aut.ac.nz, brice.valles@gmail.co.nz

Abstract

To forecast the demand response of a residential customer, forecasted load profiles must first be calculated and the demand response produced by residential energy management devices determined. Finding an efficient method to select an optimal forecasted load profile to assess the demand response of these devices is a fundamental issue currently overlooked. Poorly forecasted load profiles can affect customer compensation for demand reduction, estimation of retail profits and allocation and reconciliation processes for electricity markets. This research resolved to develop an economic model for selecting an optimal forecasted load profile from which the demand response of residential energy management devices can be determined. Using two years of half-hourly electricity consumption data for 5379 households obtained from the “Low Carbon” London project (UK), a novel economic model was created to assess the economic advantage of nine different forecast methods. Results suggest that when creating forecasted load profiles to measure demand response of these devices the sociodemographic of customers, tariff type and the operational objectives of the device must be taken into account to maximize the economic benefit of demand response initiatives for the customer and retailer.

Keywords: residential energy management system, forecast method, load modelling, demand response

1. Introduction

1.1 Motivation

Residential energy management systems are demand response tools that shift, and curtail demand to improve the energy consumption and production profile of a dwelling on behalf of a consumer [1]. The amount of demand response obtained from these devices is measured as the difference between the load profile produced by the device during a given period and a forecasted baseline load profile. The forecasted baseline load profile is the load profile the customer is expected to have if the customer is not participating in a demand response program [2].

Nevertheless, since forecasted load profiles are estimates of the amount of electricity the customer is likely to use, they are inherently imperfect [3]; that is, they can never accurately measure what is going to occur. As a result, if a forecasted baseline load profile is not carefully chosen to measure customer demand response, it would be difficult to determine customer compensation for demand reduction, forecast of retail profits and engage residential customers in allocation and reconciliation processes on electricity markets. Yet, selecting an optimal baseline load profile for a given customer group has remained an outstanding issue [3]. Addressing this issue would contribute to making optimal decisions that are dependent on accurate determination of forecasted customer demand response.

Therefore, being an outstanding issue, the key problem is to select the best forecasted baseline load profile to measure the demand response of a given set of customers that have installed residential energy management systems as demand response tools. This baseline must also be economically advantageous to the retailer, improving retail profits and decreasing retail cost.

1.2 Literature Review

Most researchers who model residential energy management systems use historical data [4] as the baseline against which to measure the performance of these devices. There is consensus in the field of demand response measurement and verification that the raw historical data of electricity usage may be insufficient to accurately describe what a customer might do in the future [5]. This is why there has been much research in developing different methods to create more accurate baselines to measure the amount of electricity a customer saved during demand response [5]. Table 10, Appendix 1, provides a list of commonly used baseline methods [6] for measuring the demand response of residential customers. Examples of such methods include "Exponential Smoothing" method which weights past observations of electricity use with exponentially decreasing weights to forecast future electricity use [7]; the "High X of Y" method which uses the average of the X highest "daily kWh usage" days from a pool of Y most recent days as the baseline for the hourly electricity use for the current day [7].

Baselines created from each method are usually adjusted using a multiplicative adjustment technique [8]. In general, once the reduction in demand is calculated, the estimated baseline is adjusted to consider any special conditions (for example change in weather or unusual demand use) that may affect the amount of demand response forecasted. Nevertheless, because baselines are estimates of the amount of electricity the customer "would have used", baselines are inherently imperfect [3]; that is, they can never accurately measure what "would have occurred". Consequently, selecting an optimal baseline method to create a baseline load profile has remained an outstanding issue [3]. This issue is one of the major reasons why there is not a widespread use of demand response technologies such as residential energy management systems in households [9, 10]. Simply put, if an appropriate baseline is not used to measure the demand response of customers then there would be inaccuracies in determining the cost benefit of residential energy management systems. This in turn would affect the retailer's ability to accomplish tasks such as economic feasibility studies, revenue-risk modelling and justifying compensation to customers for demand response.

For this reason, researchers have aimed to develop novel baselines to measure customer demand response. For example [11], created a new method for calculating a demand response baseline using support vector regression. Whilst the method is novel it still suffers from not consistently ensuring a high level of accuracy and low level of bias across different types of customers. Moreover, the support vector machine method does not consider the varying economic advantage that it would have for different customer groups and the retailer. A more sophisticated novel baseline method was developed in [12] where a "stacked autoencoder" based customer baseline method was used to determine the demand response of a set of customers.

In [13], the authors explore the performance of different baseline methods including (1) averaging methods including "Y-day simple average", "Mid X of Y", and "Nearest X of Y"; (2) regression methods that relate the electric demand to weather and other relevant parameters; and (3) a simple linear interpolation method that uses least squares to fit a linear baseline to the fan power data over the 5-minute period just before the demand response event and the 5-minute period immediately after the settling time. The authors concluded that all methods examine had strengths and weaknesses in terms of accuracy and bias. The accuracy and bias of baseline methods for residential customers have even been compared with the aim of improving policy with respect of the energy industry in places such as South Korea and France [14].

Authors of [15] used a clustering method to place customers into groups and then selected the most accurate baseline for different customers groups. However, their research fell into the same weakness of having no rigorous method for selecting a particular baseline for a particular customer group. Their selection method was solely based on an overall error performance that was the weighted sum of the absolute value of accuracy and bias that a particular baseline produced when measuring demand response.

Whilst most research papers simply compare baseline methods in terms of accuracy and bias or select a specific baseline method and work to improve the performance of the method, every few papers have developed a rigorous approach to optimally selecting the best baseline method for a set of customers. Moreover, those papers that deal with selecting an optimal baseline use the same approach, clustering customers and then using a non-rigorous method for selecting a baseline for each customer cluster.

For example, [16] outlined the fact that demand response programs go through three basic steps of development; (1) establishing a baseline against which demand reduction is measured, (2) designing the payment scheme for agents who reduce their consumption from this baseline, and (3) the selection scheme for selecting the customers that would yield the maximum demand response. Researchers [17] went one step further and examined the merits of using different self-reported baselines to measure the performance of a demand response program. According to their research, the system operator recruits consumers for a demand response program. The recruited customers report their own baselines to the operator. The selected customers reduce their load during a demand response event and are rewarded for their services (or are penalized for deviating from their expected load reduction), based on the reported baselines. The weakness of this research is that self-reported baselines only considered the economic benefits of the customer and did not consider the direct and indirect ramifications of demand response for the retail portfolio of the electricity provider. Researchers [18] did extensive analyses to show that the selection of a baseline can and does affect the benefits that the retailer receives from a demand response program. Therefore, these baselines do not fully represent the benefits of all stakeholders involved in demand response programs.

As another example, [15] developed an “Overall Performance Index” using the ideas of accuracy and bias to measure the performance of a given residential customer baseline. This index is given in Equation (1).

$$\lambda|\alpha| + (1-\lambda)|\beta| \tag{1}$$

where	
λ	Weight indicating relative importance of baseline
α	Baseline accuracy
β	Baseline bias

Nevertheless, the weakness of [15] is that the impact that using a homogenous baseline on the benefits that different customers may receive from a demand response program was not considered.

Nevertheless, lessons learnt can be gleaned from the examples when considering a method for optimally selecting a baseline for a group of customers. Selecting a baseline necessarily involves examining the baseline accuracy and bias. It also involves clustering customers into groups to increase the baseline effectiveness in measuring demand response. Furthermore, demand response programs must be designed to maximize the social and economic welfare of all stakeholders; customers and retailer [19]. These general observations point to the fact that a basic economic model can be created to help select an appropriate baseline to measure the performance of residential energy management systems with consideration not only for customers but also for the retailer.

Conventional model development suggest that if a basic economic model is developed to help optimally select an appropriate baseline model complexity and model fit must be examined [20]; model complexity measures the number of parameters needed by the model to accurately capture patterns processes and relationships between variables in the data [20]. Model fit measured how well the model generalizes to other similar data [20]. Increasing model complexity decreases model fit and vice versa. It is for this reason that it is advisable to use metrics such as Bayes Information Criterion to measure model fit and complexity. The smaller the Bayes Information Criterion values the better the balance between model complexity and fit [21]. Therefore, in creating an economic model to select an optimal baseline to measure demand response, the complexity and fit of the model needs to be addressed. With a well-balanced model small disturbance in the inputs of the model does not produce large differenced in the optimal solution produced by the model [22]. This ensures that small errors in measurements of the forecasted load profile does not affect the selection of the most optimal forecasted load profile that should be used to measure the performance of the residential energy management system [22].

Developing a robust well-balanced model for selecting a baseline method and justifying this selection based on the economic advantage this baseline would provide to the customer and retailer would be a major contribution to the current state-of-the-art. Moreover, developing a model for residential energy management systems that consider the baseline used to measure its performance would add value to existing literature.

1.3 Major contributions

The overall purpose of this paper therefore was:

> *To create a method using the concepts of social welfare maximization and baseline accuracy and bias to optimally select an appropriate baseline to measure the demand response of residential energy management systems.*

In achieving the objective set out for this work, the following major contributions were made to the state-of-the-art:

1. Using the concepts of social welfare maximization, baseline accuracy-bias trade-off, and a cost benefit analysis done on these devices, a general economic model has been created to optimally select a baseline of measure the demand response of residential energy management systems. In the process of creating this model, a cost-benefit metric was

developed for the retailer based on the change in profits that they experience because of these devices and the correlation between customer diversity and retail risk.

2. With this economic model, the collective economic advantage of using a baseline was quantified. Based on this quantification it was clear that using a single baseline to measure the demand response of all customers would lead to a loss of social and economic welfare for both the customer and retailer. Instead, it was demonstrated that using customized baselines for different customer groups ensured both local maximization of benefits for the customer and a global, maximization of benefits for the retailer.

Recently published literature has mainly focused on creating new baselines [15]. These new baselines are supposedly more accurate or have less bias than existing baselines [17]. Unfortunately, creating new baselines would not significantly benefit the current state of the art as all baselines are inherently and inevitably imperfect [16]. The significance of the major contributions of this research rests in the fact that it takes a different approach; rather than creating a new baseline this research creates a model to optimally select a baseline from an existing set of baselines considering not just the bias and accuracy of the baseline but also the impact that the demand response technology can have on the social welfare of the major participants of a demand response program (customer cost-benefit and the impact on the retailer portfolio profit and risk.). Very few papers have considered optimally selecting a baseline from a given set of baselines. Those papers that have considered optimally selecting a baseline [15] have the limitation of only considering the accuracy and bias of the baseline itself and not the effect that the demand response technology can have on the customer and retailer portfolio.

The rest of this paper is devoted to providing a detailed explanation of how objectives were achieved and how this in turn gave rise to the development of the major contributions outlined in this section.

2. Development of mathematical models

To contextualize the research, the "Low Carbon" London smart-meter dataset containing hourly electricity use measurements for 5379 residential customers was procured from [23]. Half-hourly data for these households over a period of 2 years from January 2012 to December 2013 was used. These households were broadly divided into 2 groups based on the type of tariff that customers pay. The first group consisted of customers from 1093 households who paid a three-tier Time-of-Use tariff with High (67.20p/kWh), Low (3.99p/kWh) and normal (11.76p/kWh) price signals. The other 4286 households consisted of customers who paid a Flat tariff of 14.228 p/kWh.

For this research, it was assumed that customers had residential energy management systems installed in their homes. Since these households were clustered in central London (that is, in a single geographic region) it was not necessary to consider the spatial distribution of the households and its effects on electricity use when examining the operation of the residential energy management systems. Furthermore, spot price data was collected from the N2EX Great Britain power market [24] for the same period. This spot price data along with the tariff data enabled customer electricity bills and retail revenue to be calculated during demand response.

2.1 Modelling the energy management system and characterizing customer cost-benefit ($\boldsymbol{b_c}$)

Demand response amongst residential customers would not be possible without energy management systems. At this point it was imperative to develop a model for a generic residential energy management device before the central aim of this research could have been addressed. Auspiciously, a simple means of modelling residential energy management devices as well as a cost-benefit metric for customers employing these devices has been provided by [25]. The model is shown in Figure 1.

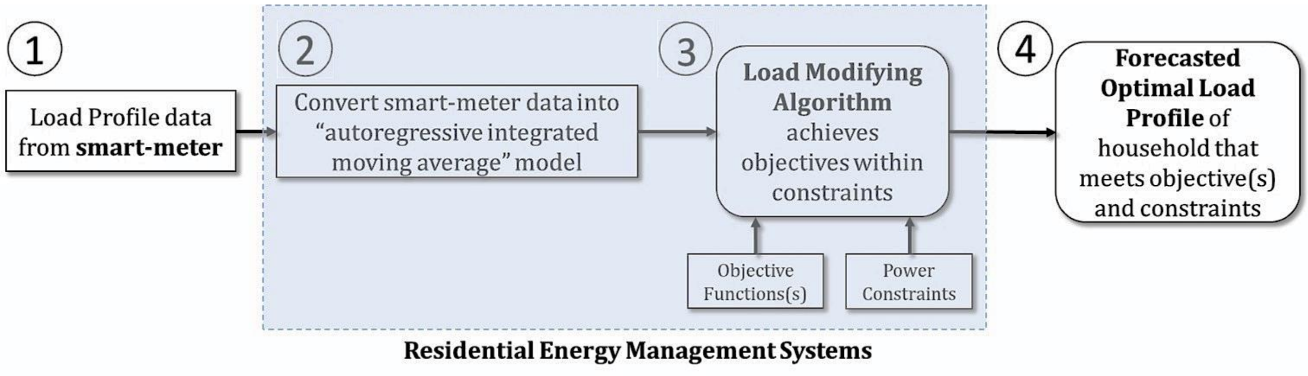

Figure 1. Generalized model characterising a residential energy management system

The authors converted the load profile data of a household into an autoregressive integrated moving average model and used that model to represent the reduction both average and peak-to-average in electricity use. An autoregressive integrated moving average model is a regression equation that uses past energy consumption data to forecast future electricity use [26]. The resulting expression developed by the authors was given in Equation (2).

$$d_t{}^* = k_1.\sum_{n=1}^{p} \alpha_l.A_{t-l} + k_1.k_2\sum_{n=1}^{p} \beta_n.Q_{t-n} + k_1.\sum_{m=1}^{q} \theta_n.e_{t-m} \tag{2}$$

Where	
$d_t{}^*$	Reduced electricity demand during hour "t"
k_1 and k_2	Percentage in average and peak-to-average electricity use remaining respectively after demand response by residential energy management system
p and q	Number of past electricity use values and past error values respectively needed in autoregressive integrated moving average model
α_l, β_n and θ_n	Regression coefficient for past average, peak-to-average and error electricity use values, respectively. Statistical software (MATLAB) was used to find these coefficients.
A_{t-l}	Past average electricity use values *l* time steps behind
Q_{t-n}	Past peak-to-average electricity use values *n* time steps behind
e_{t-m}	Past hourly forecast errors

From Equation (2), the total reduction in electricity consumed by a household with a residential energy management system was articulated by Equation (3).

$$D^* = \sum_{t=1}^{T} d_t{}^* \tag{3}$$

According to [25] the residential energy management system functioned according to two sets of objectives [1, 27]; objectives that minimize the cost incurred from having these devices [28], and those that maximize customer benefits [29]. These objectives are shown in Table 1.

Table 1. List of objectives used in this research and references for their formulas

Objective Functions	Formulas	Parameters	Reference
Maximize electricity bill savings (J_1): Electricity bills savings is the money the customer saves because of the energy management device reducing electricity demand.	$J_1 = \sum_{t=1}^{T}[(d_t - d_t^{*}) \times r_t]$		[30]
Maximize the emissions savings (J_2): Customers, in consuming less electricity cause power plants to produce less electricity resulting in less emissions into the environment. Emissions are the amount of greenhouse gasses (primarily carbon dioxide) released into the atmosphere.	$J_2 = \sum_{t=1}^{T}[(d_t - d_t^{*}) \times c \times c_r]$	$c = 70$ £ $c_r = 0.003\ kg/kWh$	[31]
Maximize tax savings (J_3): According to [32] reduction in taxes charged to the customer's electricity bill (τ) is generally given to customers that are part of a demand response program as incentive for reduce electricity use. Using the same scheme as [32] it was assumed that tax savings was derived from an addition taxation incentive scheme given by the European government. In such a scheme the ad valorem tax rate ($\tilde{\tau}$).	$J_3 = \sum_{t=1}^{T}(\tau . d_t - \tilde{\tau} . d_t^{*})$	$\tau = 5\%$ $\tilde{\tau} = \frac{\tau}{d_t}$	[32]
Maximize retail profit (J_4): The revenue obtained by the retailer is reflected in the electricity bill of the customer. Maximizing retail profit involves load shifting and this load shifting must satisfy one major condition: The profit per kWh made from period "*b*" is higher than the profit made per kWh during period "*a*". This implies that the retailer can make an increase in profit if load is shifted from period "*a*" to period "*b*".	$J_4 = \sum_{t=1}^{T}[(d_t - d_t^{*}) \times (r_t - s_t)]$		[33, 34]
Minimizing customer discomfort (J_5): One of the most common being the Taguchi loss function. This function gives a relationship between the reduction in consumption of a good (in this case electricity) and loss of satisfaction the consumer experiences [35].	$J_5 = \frac{X}{Z} \times Y$		[4]
Minimizing payback period (J_6): Simple payback period is the amount of time taken to recover the cost of the device. The payback period "n" (given in months) is simply given as the cost of the device divided by the savings per month produced by the energy management device.	$J_6 = \beta\left(p_p . J_1 - C_p\right)^2$	$C_p = 273.13$ £ $\beta = (1\ £)^{-1}$	[36, 37]

where
$X = \sum_{t=1}^{T}[d_t \times r_t]$
$Y = \sigma^2{}_t{}^{*} + \sum_{t=1}^{T}[(d_t{}^{*})^2]$
$Z = \sigma^2{}_t + \sum_{t=1}^{T}[d_t{}^2]$

The objective shown in Table 1 are combined into a single objective function that represents the cost-benefit (b_r) the customer derives from the device is shown in Expression (4). When maximized, the function leads to the optimal load profile of the household.

$$
\begin{aligned}
&Maximize \quad b_r = \frac{w_1.b_1 + w_2.b_2 + w_3.b_3 + w_4.b_4}{w_5.c_1 + w_6.c_2} \\
&Subject\ to \quad l_t \leq d_t^* \leq b_t \qquad for, \forall\ l_t, \forall\ d_t^*and\ \forall\ b_t
\end{aligned}
\tag{4}
$$

Where	
b_r	Optimal benefit-cost ratio customer derives from device
b_1, b_2, b_3, b_4	Objective function of energy savings (b_1), emissions savings (b_2), tax savings (b_3), retail portfolio savings (b_4) created using Equation (2) and is shown in Table 1.
c_1, c_2	Objective function of customer discomfort (c_1), payback period (c_2)
w_1, w_2, w_3, w_4, w_5, w_6	coefficients that take a discrete value either 0 or 1.
l_t, f_t	Hourly base load and forecasted load profile, respectively.

The binary coefficients (w_1 to w_6) in the objective function indicate whether an objective was considered in the operation of the device. If an objective was considered, then its coefficient had a value of 1 and if not, a value of 0. It is important to note that these weights do not indicate the relative importance of one objective over another and as such the sum of the weights do not add up to one.

In addition, as shown in Expression (4), in the process of reducing electricity use, the device must ensure that there is enough electricity for base load appliances and that electricity demand does not exceed the given forecasted load profile of the household (f_t). Hourly base load (l_t) is the persistent quantity of electricity necessary to keep household appliances such as fridges and water heaters running whilst residents are not actively using electricity [38].

The final phase of constructing the model involved specifying a how the objective function shown in Expression (4) modifies the load profile represented by Equation (2). A particle swarm optimization (PSO) was used to do this and is given in Table 2.

Table 2. Load modifying algorithm based on Particle Swarm Optimization

1	Initialize number of iterations (number of attempts to get a solution) to 150
2	Initialize number of particles (number of possible solutions to consider in each attempt) to 50
3	For each particle, the position (possible solution to $k_1$and k_2) and velocity (rate at which the solution will change) is randomly initialized
4	Measure the fitness (how good the solution is at maximizing the objective function and satisfying the constraints) for each particle
5	Store each particle best fitness in "pbest" and store the particle with the overall best fitness in "gbest"
6	For each particle update the position and velocity vectors according the update equation found at [39]
7	Repeat steps 3 to 6 until maximum number of iterations is reached.

Particle swarm optimization was chosen as the foundation for the algorithm because relative to other optimization techniques, it does not add to the complexity of the problem, it is excellent at escaping suboptimal solutions and finding (near) optimal solutions, and is simple to implement [40, 41]. The final output of the particle swarm optimization algorithm is given in Table 3.

Table 3. Output of residential energy management system

1.	optimal reduced load profile (D^*)
2.	% reduction (k_1) in average electricity use
3.	% reduction (k_2) in hourly peak-to-average electricity use
4.	benefit-cost ratio (b_c) for customer

2.2 Characterizing cost-benefit for retailer (b_r)

Aside from defining the customer cost-benefit for these devices, another preliminary step towards developing an economic model for selecting an optimal forecasted baseline load profile involved characterizing the retailer cost-benefit. As a starting point, according to the Modern Portfolio Theory [42] the performance of a retail profit can be expressed in terms of the profit obtained from the portfolio and the variance (or risk) experienced by the portfolio. When an asset is added to a portfolio it can change either the profit or variance in profit received from the portfolio. This causes a change in the portfolio performance metric. This change was considered the benefit received by the retailer from adding the asset to the portfolio. This benefit (b_r) was expressed by Equation (5).

$$b_r = \frac{\Delta\mu_p}{\Delta\sigma^2{}_p} \tag{5}$$

Where

$\Delta\mu_p$	Change in average profit earned by the retail portfolio
$\Delta{\sigma^2}_p$	Change in variance of profit gained from portfolio

Research has shown that residential energy management systems can influence the expected profit and financial risk experienced by the retailer [25]. Given these observations, the change in portfolio profit ΔP can simply be represented using Equation (6).

$$\Delta P = \mu_{p_2} - \mu_{p_1} \tag{6}$$

Where

μ_{p_2}	Final profit in portfolio after residential energy management systems have been deployed
μ_{p_1}	Initial profit in portfolio before residential energy management systems have been deployed

Characterizing the change in financial risk on the other hand was a bit more involved. Since it has been proven that these devices can influence financial risk directly by reducing variance in electricity use and indirectly by influencing the relationship between customer diversity and variance in electricity use [25], this direct and indirect effect needed to be incorporated into the change in risk ($\Delta{\sigma^2}_p$) experienced by the retailer.

The direct effect of the residential energy management system on the variance in revenue received from the customers was represented by "Cohen's d" [43] given in Equation (7). "Cohen's d" measures the size of the effect that residential energy management systems have on the variance in electricity use of the retail portfolio after deploying these devices amongst a diverse group of customers [43].

$$Cohen\ d = \frac{\mu_{\sigma_1} - \mu_{\sigma_2}}{\frac{1}{2}(\sigma_1 + \sigma_2)} \tag{7}$$

Where

μ_{σ_1}	Average variance in profit experienced by the retailer from households before the deployment of residential energy management systems
μ_{σ_2}	Average variance in profit experienced by the retailer from households after the deployment of residential energy management systems

σ_1	standard deviation in profit experienced by the retailer from households before the deployment of residential energy management systems
σ_2	standard deviation in profit experienced by the retailer from households before the deployment of residential energy management systems

Similar "Cohen's d" metric, Cohen's f^2metric was used to measure the indirect effect of these devices on retail risk through their influence on customer diversity [44]. According to Cohen's f^2 statistic if two variables are correlated then the effect that one variable has on another can be written in terms of their correlation as shown in Equation (8).

$$f^2 = \frac{\rho^2}{(1-\rho^2)} \tag{8}$$

Where

f^2	Effect size metric measuring the effect of diversity on variance
ρ^2	Correlation between diversity and variance

Incidentally, Cohen's f^2 is simply an extension of Cohen's d. These two values are related to each other as shown in Equation (9).

$$d^2 = 2k \times f^2 \tag{9}$$

The variable "k" in Equation (9) is the number of customer sets sampled to establish the relationship between variance and diversity [45]. By equating Equation (7) and Equation (8) the change in risk produced by the residential energy management system was expressed in terms of its direct and indirect effects on risk.

$$\frac{\left(\mu_{\sigma_1} - \mu_{\sigma_2}\right)^2}{\frac{1}{4}(\sigma_1 + \sigma_2)^2} = \frac{2k\rho^2}{(1-\rho^2)} \tag{10}$$

Rewriting Equation (10) gives Equation (11).

$$\Delta\sigma^2{}_p = \frac{k}{2} \times \frac{\rho(\sigma_1 + \sigma_2)}{\sqrt{(1-\rho^2)}} \tag{11}$$

Equation (11) expresses the change in financial variance experienced by a portfolio in terms of the influence of these devices on directly on retail risk and indirectly on customer diversity. Using Equation (6) and (11) the retail benefit (Equation (5)) was rewritten as Equation (12). Equation (12) represents the change in portfolio performance (or benefit) that the retailer can expect with the mass deployment of residential energy management systems.

$$b_r = \frac{2}{k} \times \frac{\Delta P\sqrt{(1-\rho^2)}}{\rho(\sigma_1 + \sigma_2)} \tag{12}$$

2.3 Derivation for Economic Advantage (E_A) Model

Having defined the cost-benefit ratio for the customer (b_c) and the retailer (b_r), the economic model for optimally selecting an appropriate forecasted load profile to measure the demand response of residential energy management systems was derived. Several authors model demand response programs by their ability to maximize the social welfare of the stakeholders of the program [19]; social welfare as the weighted sum of the benefits of the stakeholders of the demand response program. Since the customer and retailer are the focus of this research the weighted sum of their benefits was expressed by Equation (13) [19, 46].

$$U = \beta_1 . b_c + \beta_2 . b_r \tag{13}$$

Where	
U	Total social welfare from the demand response program
b_c, b_r	Optimal benefit or gain for the customer (b_c) and retailer (b_r)
β_1, β_2	Weights indicating relative importance of customer and retailer to the program. The weight values are generally arbitrarily selected and need not add up to one.

In optimally designing any aspect of a demand response program, one of the considerations that must be accounted is maximizing the total social welfare (U) derived from the program [47] (generally known as the social welfare maximization problem [48]). This is shown in Expression (14).

$$Maximize\ U = [\beta_1 . b_c + \beta_2 . b_r] \tag{14}$$

Subject to $B_L \leq {d_t}^* \leq B_U$

Where

B_L	Lower limit for demand response from participants of demand response program
${d_t}^*$	Hourly reduced demand from participants of demand response program after demand response event
B_U	Upper limit of demand from customers. This is often used as the baseline to measure demand response from participants.

As outlined by [15] another consideration that must be made is the accuracy and bias of a forecasted baseline load profile. The ensuring the highest accuracy from a forecast involves minimizing the error that is made to measure demand response [49]. Minimizing forecast error involves minimizing the bias and variance produced by the forecasted load profile. This was be represented by Expression (15) [49].

$$Minimize\ RMSE = \sqrt{bias^2 + variance} \tag{15}$$

Forecast bias is the tendency of the baseline method to make assumption about the customer demand that is not necessarily true [50]. Additionally, a forecasted baseline load profile is said to have high variance when there is a tendency of the profile to overfit data.

From Expression (15), RMSE represents the root mean square error associated with a forecast and is a measure of the accuracy of the forecasted load profile. Minimizing an equation is equivalent to maximizing it reciprocal. As such, Expression (15) can also be rewritten as a maximization problem shown in Expression (16).

$$\begin{gathered} Maximize\ \frac{1}{RMSE} \\ Subject\ to\ \ RMSE > 0 \end{gathered} \tag{16}$$

Selecting an optimal forecasted baseline load profile for residential energy management systems involved maximizing the social welfare of the stakeholders and the accuracy of the forecast. Therefore, Expression (14) and (16) was combined into a single maximization problem shown in Expression (17).

$$\begin{gathered} Maximize\ E_A = \frac{1}{RMSE} \times (\beta_1 . b_c + \beta_2 . b_r) \\ Subject\ to\ \ RMSE > 0 \quad \text{and} \quad B_L \le {d_t}^* \le B_U \end{gathered} \tag{17}$$

Where	
E_A	Economic advantage of selecting a baseline to measure the performance of energy management systems

From Expression (17), "E_A" represents the number of units of benefit (social welfare) the customer and retailer collectively gain from the demand response program considering the forecast accuracy. The higher the value of "E_A" the more economically advantageous for both the customer and retailer to select the given hourly baseline load profile.

2.4 Characterizing weights for cost-benefit values b_c and b_r

The coefficients ($w_1$and w_2) in Expression (17) simply indicate the relative important of the social welfare of the customer and retailer. These coefficients can take any arbitrary value and were set to the values shown in Table 4.

Table 4. Values of coefficients for cost-benefit variables b_c and b_r in Expression (17)

Suggested weight value	Reasons for selecting value
$w_1 = 1$	trivially given the value of 1
$w_2 = \frac{k\rho(\sigma_1 + \sigma_2)}{2}$	Selected to cancel the denominator in Equation (12). Reasons: • ρ (the correlation between diversity and quantity risk) can possibly be zero making Equation (12) undefined. • "k" (the number of groups made to get the correlation between diversity and quantity risk) was arbitrarily set and does not have any impact on the retail cost-benefit (b_r) • σ_1 and σ_2; these are the variances of the household before and after the operation of the residential energy management system can possibly be zero making Equation (12) undefined.

These weights resulted in the following formulation for the weighted cost benefits of the customer and retailer.

$$\beta_1 . b_c = b_c \qquad \beta_2 . b_r = \Delta P\sqrt{(1-\rho^2)} \tag{18}$$

2.5 Characterizing forecasted baseline load profile methods used to measure demand response

Forecasted baseline load profiles are generally named according to the procedures used to calculate them. Nine of the most used procedures were chosen for analysis in this research and a full description of each is provided in Table 10, Appendix 1. Each forecast procedure requires parameters to calculate their forecasts. The general parameters used in this research to calculate each forecast is given in Table 5. Typically, when modelling residential energy management systems, unaltered historical data is often used to determine demand reduction. Therefore, it was included in Table 5.

Table 5. Parameters for forecasted load profile procedures [5, 6]

Baseline Method	Parameterization*	
1. Historical Data		
2. Last Y days	$Y = 10\ days$	
3. Low X of Y days	$X = 4\ days,$	$Y = 10\ days$
4. Mid X of Y days	$X = 4\ days,$	$Y = 10\ days$
5. High X of Y days	$X = 4\ days,$	$Y = 10\ days$
6. Exponential Moving Average	$\alpha = 0.9,$	$Y = 10\ days$
7. Linear Regression	$X = 5\ days$	
8. Polynomial Interpolation	$X = 5\ days$	
9. Neural Network	$X = 5\ days$	

* the definitions given in Table 10 of

Appendix 1 explains how these parameters are used to create the different forecasted load profiles

For each of the two groups of households (customers paying Flat-rate and Time-Of-Use tariff), the methods listed in Table 5 coupled with the smart meter data obtained from the "Low Carbon" London project and was used to create 9 different possible forecasts for a typical day. To illustrate the forecasted load profiles that were produced, a comparison of the load profile of a typical day for customers on a Flat-rate tariff with the forecasts created is given in Appendix 2. Having developed a generalized economic model (Expression (17)) to determine which is the best forecasted load profile for a given customer it was also prudent to determine if the choice of forecasted baseline load profile made by the economic model was indeed the best choice. To do this the economic model's performance needed to be measured. This was done using model complexity, fit and convergence.

2.6 Characterizing model complexity, fit and convergence

Model complexity measures the number of parameters needed by the model to accurately capture patterns, processes and relationships between variables in the data [20]. Model fit measured how well the model generalizes to other similar data [20]. In the case of the economic model developed in Expression (17), examining model complexity and fit was important in providing evidence that the forecasted baseline load profile that produced an optimal economic advantage (E_A) is also the forecasted load profile that best models the future electricity demand of the customer. Model complexity and fit was measured using a single metric called the Bayesian Information Criterion (BIC) metric [51] given in Equation (19).

$$BIC = k \times ln(n) - 2 \times ln(\hat{L}) \tag{19}$$

Where	
k	Number of parameters in the model. This the case of the (E_A) model shown in Expression (17), there are 5 parameters.
$\hat{L}$	Likelihood function.
n	Number of data points used to calculate the model.

The lower the value of the Bayesian Information Criterion (BIC) metric the better the economic model at selecting an appropriate baseline load profile for the residential energy management system.

In addition to model complexity, model convergence was used to determine if the economic model produced a unique maximum solution [52]. The particle swam optimization technique was used to find the global optima for Expression (17) and it was necessary to check if the output of the algorithm converged to a constant value as the algorithm progressed through its iterations. If it did not, then the economic model would be inadequate to select an optimal forecasted baseline load profile to measure the demand response of residential energy management systems.

3. Simulation Studies

Having characterized the different variables used in the economic model (Expression (17)) and establish metrices to measure model complexity, fit, and convergence, optimally selecting the correct forecasted baseline load profile for each customer group was found. This was done using the flow-chart outlined in Figure 2.

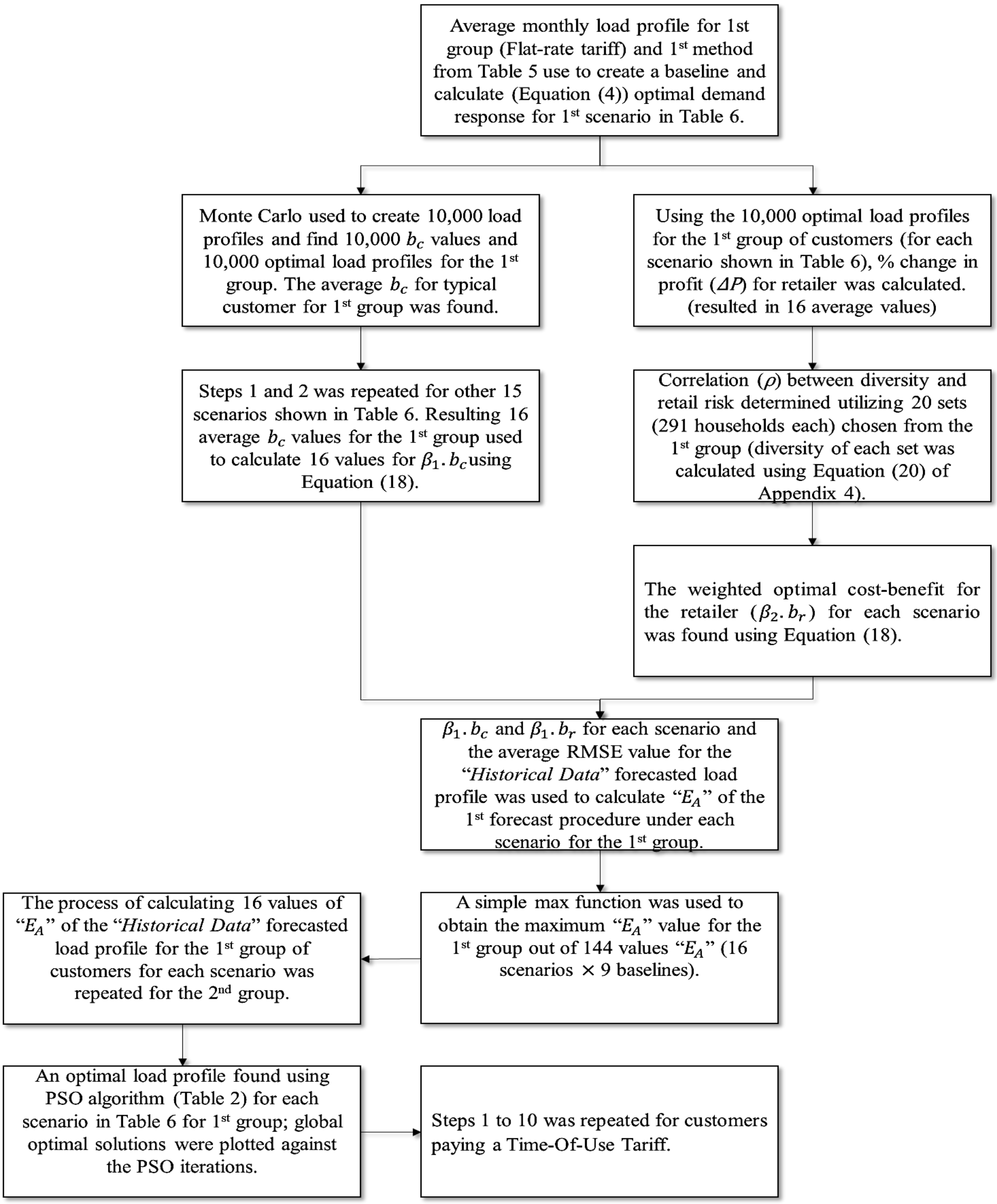

Figure 2. Flow chart for procedure used to calculate E_A for each baseline

In accordance with Figure 2, the average monthly load profile for the first subgroup of households (Lavish Lifestyles customers paying a Flat-rate tariff) was found and used to represent the electricity use of a typical customer within that subgroup. Using the average monthly load profile, the first forecasting procedure from Table 5, "*Historical Data*", was selected to create a baseline monthly load profile. The baseline load profile was coupled with the residential energy management system model given in Equation (4) to simulate and calculate the optimal demand response of the device for different combinations of the device's objectives. This was done by first taking the coefficients $\boldsymbol{w_1}$ **to** $\boldsymbol{w_4}$ given in Equation (4) and setting them to "1" or "0" to produce the combinations given in Table 6.

Table 6. Possible combination of objectives for a residential energy management system

Scenarios	w_1	w_2	w_3	w_4	Scenarios	w_1	w_2	w_3	w_4
1	0	0	0	0	9	0	0	0	1
2	1	0	0	0	10	1	0	0	1
3	0	1	0	0	11	0	1	0	1
4	1	1	0	0	12	1	1	0	1
5	0	0	1	0	13	0	0	1	1
6	1	0	1	0	14	1	0	1	1
7	0	1	1	0	15	0	1	1	1
8	1	1	1	0	16	1	1	1	1

The coefficients w_5, w_6 remained constant with a value of 1. It is assumed that the customer always experiences discomfort (because of reduced electricity use) and spent money over time for the installation of the device.

Each combination represented a different scenario in which the residential energy management systems operated with a particular set of objectives. The coefficients w_5 to w_6 was set to a constant value of 1. This indicated that customer discomfort and cost of the device was always considered in the operation of the device. For example, in "scenario 15" the device was operating to achieve emissions and tax savings for the customer and increase retail profit for the electricity retailer whilst minimizing the discomfort caused by the device and the cost of the device.

Each scenario was examined for the first subgroup of customers using the algorithm given in Table 7.

Table 7. Algorithm for generating optimal benefit cost for each subgroup under each tariff for the possible cases outlined in Table 6

1	Find the representative average monthly load profile for current subgroup.
2	Perform a Monte Carlo generator [25] to generate 10000 load profiles from the average load profile.
3	Convert each Monte Carlo Profile into an ARIMA model (see Equation (2)).
4	Find the optimal cost-benefit for customer using Equation (4). Store results.
5	Repeat lines 1 to 4 for each of the 16 cases
6	Find the average optimal benefit-cost ratio for the subgroup.
7	Repeat lines 1 to 6 for each of the 2 subgroups
8	Repeat lines 1 to 7 for each of the 9 baselines

In accordance with the algorithm, using average monthly load profile created for the first subgroup, a Monte Carlo generator given by [25] was used to generate 10,000 monthly electricity use profiles. The 10,000 load profiles were each compared to the *"Historical Data"* forecast for the household and the root-mean-square error value for each generated profile was found. This resulted in 10,000 root-mean-square error values which was used to find an average root-mean-square error value. This average represented the root-mean-square error for the group of customers if they used *"Historical Data"* forecasted load profile. The variance of the root-mean-square errors ($\sigma_{\varepsilon}{}^{2}$) was also found and used to compute the Bayesian Information Criterion (BIC) value for the "*Historical Data*" forecasted load profile. To compute the BIC value using Equation (19), the variables "*k*" was set to 5 and "*n*" was set to 720 (there are 720 hours in a 30-day month). This value represented the complexity and fit of the economic model when selecting a forecasted load profile created using the "*Historical Data*" forecast procedure.

For the first scenario shown in Table 6, the 10,000 load profiles generated were fed into the residential energy management system model to find 10,000 values for the optimal cost-benefit ratio for the customer (b_c) and 10,000 optimal load profiles for the subgroup. The 10,000 results for cost-benefit ratio (b_c) was used to find an average cost-benefit for a typical customer for the first subgroup. This was repeated for the other 15 scenarios shown in Table 6.

This resulted in 16 average values for cost-benefit ratio (b_c) for the first subgroup (one for each scenario shown in Table 6). These 16 values were used to calculate 16 values for the weighted customer cost-benefit ($\beta_1 . b_c$) using Equation (18). Having determined weighted customer cost-benefit ($\beta_1 . b_c$) for the *"Historical Data"* forecasted load profile for each scenario shown in Table 6, it was necessary to determine the weighted retail cost-benefit ($\beta_2 . b_r$). Using the 10,000 optimal load profiles for the first subgroup of customers (for each scenario shown in Table 6), the percentage change in aggregated profit (ΔP) for the retailer was calculated. This resulted in 16 average values for change in aggregated profit (ΔP) for the first subgroup (one for each scenario shown in Table 6).

After change in profit (ΔP) was dealt with, the correlation (ρ) between customer diversity and retail risk was determined for the first subgroup. To do this, 20 sets of households were created from the households belonging to the first subgroups. Each set comprised 291 households chosen from the subgroup using a "Random Sampling with Replacement" [53] technique. The number of sets created was determined arbitrarily. However, the size of each set was determined using a statistical tables [54]. According to the statistical table [54], to determine the correlation between customer diversity and risk at a 95% confidence level, with a 5% margin of error, each set must have at least 291 households. Confidence level specifies how certain the correlation results are expected to be [54].

The diversity of each of the 20 sets was calculated using Equation (20) of Appendix 3. The average aggregated month load profile for each of the set was also found. For each scenario of operational objectives shown in Table 6, the sets' average monthly load profiles were feed to residential energy management system model and the optimal load profiles was obtained. This resulted in 20 optimal load profiles for each scenario shown in Table 6. The risk (variance in profit) from these optimal load profiles were calculated and recorded. Using the results, the correlation (ρ) between the diversity and retail risk for each scenario was found. This resulted in 16 correlation values (one for each scenario shown in Table 6). Authors of [55] provide an interpretation of various correlation (ρ) values.

Having found change in profit (ΔP) and correlation (ρ) each scenario of operational objectives shown in Table 6 for the first subgroup, the weighted optimal cost-benefit for the retailer ($\beta_2 . b_r$) for each scenario was found using Equation (20).

Once this was done, the weighted benefit-cost ratio for the customer ($\beta_1 . b_c$) and the retailer ($\beta_1 . b_r$) for each scenario and the average root-mean-square-error value for the "*Historical Data*" forecasted load profile was used to calculate the economic advantage "E_A" of the "*Historical Data*" forecast procedure under each scenario for the first subgroup. This resulted in 16 values of the economic advantage "E_A" of the "*Historical Data*" forecasted load profile for the first subgroup.

The process of calculating 16 values of the economic advantage "E_A" of the "*Historical Data*" forecasted load profile for the first subgroup of customers for each scenario was repeated for the second subgroup. The result was 32 values of "E_A" (16 for each subgroup). This permitted the economic advantage of the baseline created from the "*Historical Data*" forecast procedure to be compared across different tariff types and different scenarios for the operational objectives of the residential energy management system. The process of getting the 32 values of "E_A" (16 for each subgroup) was repeated for each of the other load profile forecast procedures shown in Table 5.

After obtaining 16 values of "E_A" per baseline for the first subgroup, there were 144 values from which the maximum value was to be obtained. A simple max function was used to do this. Based on the maximum "E_A" value found, the optimal baseline and scenario producing this value was recorded. This was also done for the second subgroup of customers.

Finally, the value of the economic advantage of the baselines depended on the optimal load profiles obtained by the particle swarm algorithm. To show model convergence, the typical load profile created for the first subgroup of customers was coupled with the residential energy management system model, and an optimal load profile was found using the particle swarm optimization algorithm (given in Table 2) for each scenario in Table 6. The particle swarm optimization algorithm run for 150 iterations and after each iteration the global optima was recorded. This was done for each scenario shown in Table 6. The global optimal solutions were plotted against the iterations. This was repeated for the Time-Of-Use subgroup of customers. The graphs produced were checked for convergence; that is, if the graphs settled on a constant value within the 150 iterations.

4. Results

The optimal economic advantage "E_A" calculated for each subgroup of customers and for each of the 16 scenarios (given in Table 6) was presented in Table 8 and Table 9. Every entry in these tables shows the economic advantage of selecting a load profile forecast procedure for a given operating scenario of the residential energy management system. In Table 8, for example, when selecting a "*High X of Y days*" forecast procedure for creating a forecasted load profile for Flat-rate customers in "scenario 8" (where the device is operated to produce electricity bill, emissions and tax savings for the customer and an increase retail profit) the optimal economic advantage "E_A" was found to be "2.335". This means that the customer and retailer would make a total of "2.335" units of benefit for every unit of cost they experience from the "*High X of Y days*" forecast.

Clearly economic advantage of the load profile forecast procedure is influenced by the operating behaviour of the residential energy management system and the tariff paid by the customer. The last column in Table 8 shows that creating a baseline using the "*Mid X of Y*" procedure provides the highest average economic advantage (-0.766) for both the customer (paying a Flat-rate tariff) and the retailer irrespective of the operating objectives of the residential energy management system. However, the last row of Table 8 shows that the economic advantage when a single forecast procedure is used for all customers is low, ranging from (-16.090 to 0.959). This suggests that using a single forecasted load profile to measure the demand response performance of all customers may not be appropriate; that is, a one-size-fits-all approach that is traditionally used for demand response programs may not be the most optimal approach to measuring demand response. A better approach would be to allow the device to select a forecast procedure that maximizes the economic advantage of its demand response. The device can then report to the retailer the demand response expected for the household. In effect this would ensure that, locally, the customer's welfare is maximized and globally, the retailer's welfare is also maximized.

Negative economic advantage values indicate that there would be an overall lost in benefits for all stakeholders. An overall lost does not imply that all stakeholders would simultaneously not benefit from the system but that the benefit that anyone stakeholder would obtain is greatly outweighed by the lost experienced by other stakeholders. For example, in Scenario 2 of Table 8 for customers playing a Flat-rate tariff, all baseline methods for creating a forecasted baseline load profile led to a negative economic advantage. However, using an "Exponential Moving Average" baseline method resulted in a worse economic advantage (-18.622) than the Last Y days method (-6.833). This means that the number of customers who benefited from the energy management devices increased then the Last Y days method was used to create a forecasted baseline to measure the demand response of the energy management devices. Nevertheless, in both cases there was still an overall lost in economic advantage of using both baseline methods.

Every column in Table 8 shows that as the operating conditions of the residential energy management system changes, the baseline method that gives the highest economic advantage to the customer and retailer also changes. For example, for scenario 7 the optimal forecast procedure is the "*Linear Regression*" procedure which will give a total economic advantage "E_A" of "5.230" units of benefit for Flat-rate customers. But for scenario 8, for the same set of customers, the optimal forecast procedure is the "*High X of Y*" procedure which will give a total economic advantage "E_A" of "2.335" units. This implies that when selecting a forecasting procedure to measure the performance of a demand response program involving residential energy management systems the operational objectives of the systems should be considered.

The same analysis was done on Table 9 which shows the economic advantage of the selected load profile forecast procedures for customers paying a Time-Of-Use tariff and for the retailer. The last column of Table 9 shows that the forecast procedure with the highest economic advantage irrespective of the operating scenario of the residential energy management system is "*Neural Network*" procedure whose "E_A" is –0.253 units, a loss of benefit. This enforces the idea that in the process of installing residential energy management systems, the customer and retailer must be very specific about the operational objectives used to run the device (when considering the baseline used to measure its performance). Like Table 8, every column in Table 9 showed that as the operational conditions of the residential energy management systems changes, the forecast procedure that gives the highest economic advantage to the customer and retailer also changes; ranging from –16.337 units to 2.215 units (a slightly larger ranger than for Flat-rate customers). This result supports the idea that a single forecasted baseline load profile to measure the demand response performance of all customers may not be appropriate.

Overall, the economic advantage associated with the forecasted load profile irrespective of customer type and operational objective was higher for Time-Of-Use (overall average "E_A" of -0.865 units) than for customers that for Flat-rate customers ("E_A" of -2.132 units).

This result agrees with the well-known notion that Time-Of-Use tariff is generally better for customers that the Flat-rate tariff. However, the variance in economic advantage for Flat-rate customers (28.443 units^2) is less then with Time-Of-Use customers (145.472 units^2). This implies greater stability of benefits for Flat-rate customers participating in demand response programs.

It is noteworthy to mention that the generalized economic model presented in Expression (17), only consider the economic advantage of using a particular load profile forecast procedure considering only the customer and retailer. The model can be extended to include other stakeholders like the distribution system operator.

Furthermore, using the economic model provided in this research can help identify not just the optimal forecasted load profile, but also the operating conditions of the technology that would produce the highest cost-benefit for all stakeholders involved.

Table 8. Economic Advantage E_A of selected forecast procedures for customers paying a Flat-rate tariff under different operating scenarios for energy management system

Baseline Methods	Scenarios																
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Mean**
Historical Data	0.000	-11.360	-14.927	-18.487	-1.919	3.738	1.562	0.894	-0.026	-0.027	-0.028	-0.028	-0.029	-0.028	-0.029	-0.026	-2.545
Last Y days	0.000	-6.833	-4.533	-36.233	-3.515	0.262	0.990	0.827	-0.029	-0.028	-0.025	-0.026	-0.030	-0.029	-0.029	-0.028	-3.079
Low X of Y days	0.000	-12.030	-19.940	10.279	-15.990	1.675	1.275	0.293	-0.023	-0.025	-0.022	-0.024	-0.024	-0.024	-0.026	-0.023	-2.164
Mid X of Y days	0.000	-10.896	-16.035	1.353	6.497	1.686	4.518	0.824	-0.024	-0.026	-0.025	-0.027	-0.027	-0.025	-0.024	-0.025	-0.766
High X of Y days	0.000	-10.606	-43.909	-14.895	0.491	0.979	0.555	2.335	-0.023	-0.022	-0.020	-0.022	-0.022	-0.019	-0.023	-0.019	-4.076
Exponential Moving Average	0.000	-18.622	-19.234	-1.010	2.675	-5.421	1.834	-0.447	-0.033	-0.034	-0.029	-0.031	-0.032	-0.032	-0.030	-0.033	-2.530
Linear Regression	0.000	-10.376	-13.628	0.905	-0.791	1.703	5.230	0.190	-0.025	-0.027	-0.024	-0.024	-0.026	-0.025	-0.028	-0.024	-1.061
Polynomial Interpolation	0.000	-11.644	-15.803	0.523	-1.884	1.034	1.248	-1.923	-0.023	-0.025	-0.023	-0.024	-0.024	-0.024	-0.022	-0.023	-1.790
Neural Network	0.000	-10.708	3.201	-0.465	0.001	2.974	-14.687	1.068	-0.025	-0.024	-0.023	-0.024	-0.026	-0.026	-0.025	-0.025	-1.176
Mean	0.000	-11.453	-16.090	-6.448	-1.604	0.959	0.280	0.451	-0.026	-0.027	-0.024	-0.025	-0.027	-0.026	-0.026	-0.025	-2.132
Variance	0.000	9.482	163.976	201.514	37.794	6.822	34.130	1.368	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	28.443

Table 9. Economic Advantage E_A of selected forecast procedures for customers paying a Time-Of-Use tariff under different operating scenarios for energy management system

Baseline Methods	Scenarios																
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Mean**
Historical Data	0.000	-7.125	-8.346	4.305	1.723	-0.080	0.609	0.334	0.013	0.013	0.012	0.014	0.014	0.013	0.013	0.014	-0.530
Last Y days	0.000	-6.107	-7.106	-7.629	1.854	1.968	1.134	0.413	0.013	0.012	0.012	0.012	0.011	0.012	0.012	0.015	-0.961
Low X of Y days	0.000	-5.853	-5.553	-0.154	-0.653	0.764	0.929	0.284	0.011	0.010	0.009	0.011	0.011	0.010	0.011	0.012	-0.634
Mid X of Y days	0.000	-11.657	-4.254	-3.457	2.071	-2.505	1.157	0.593	0.013	0.012	0.010	0.011	0.012	0.012	0.012	0.014	-1.122
High X of Y days	0.000	-6.770	-1.940	-27.721	2.773	1.092	-0.663	0.395	0.011	0.010	0.009	0.010	0.011	0.009	0.009	0.012	-2.047
Exponential Moving Average	0.000	-12.713	-9.233	-6.618	3.033	1.805	1.726	0.517	0.017	0.016	0.016	0.020	0.019	0.020	0.015	0.021	-1.334
Linear Regression	0.000	-6.248	-4.251	4.825	1.084	-3.497	-0.829	0.477	0.014	0.011	0.012	0.013	0.014	0.014	0.012	0.016	-0.521
Polynomial Interpolation	0.000	-5.431	-102.900	97.260	2.559	0.983	0.908	0.417	0.013	0.013	0.013	0.012	0.013	0.013	0.011	0.014	-0.381
Neural Network	0.000	-7.964	-3.454	-3.070	5.492	3.760	0.374	0.724	0.010	0.010	0.009	0.011	0.010	0.010	0.009	0.014	-0.253
Mean	0.000	-7.763	-16.337	6.416	2.215	0.477	0.594	0.462	0.013	0.012	0.011	0.013	0.013	0.013	0.012	0.015	-0.865
Variance	0.000	6.903	1059.290	1252.840	2.717	5.060	0.719	0.018	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	145.472

To support the observations made in Table 8 and Table 9, the Bayesian Information Criterion values was calculated for the different forecasts. The result is shown in Figure

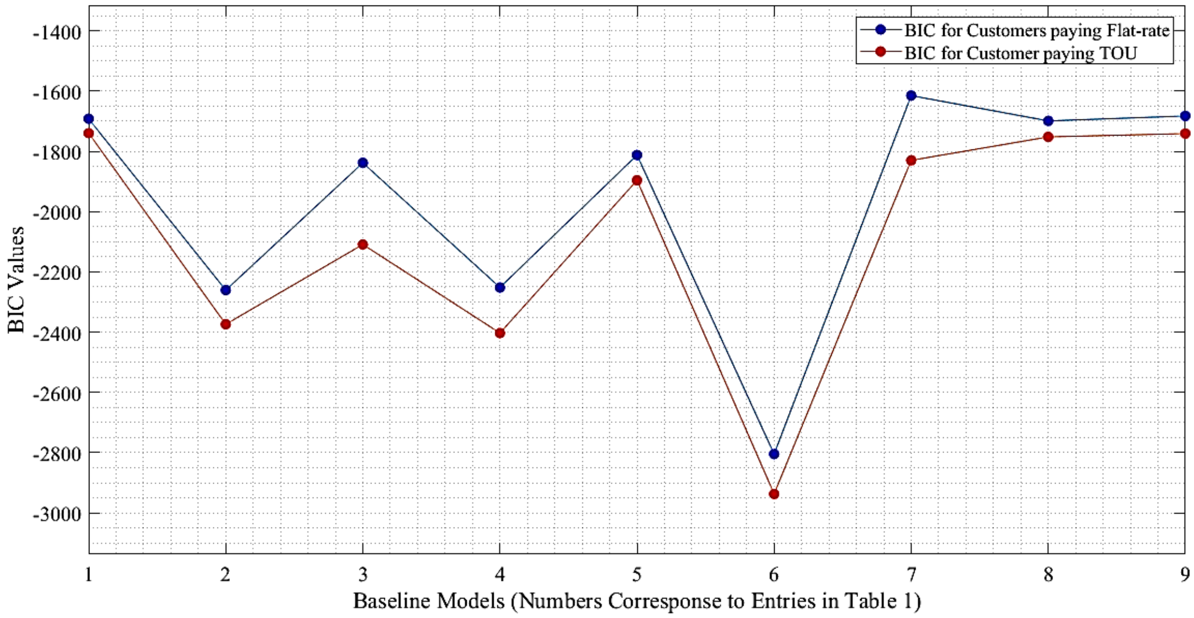

Figure 3. BIC values for different baselines used to measure demand response for customer groups

Lower BIC values indicate better model fit and complexity. Therefore, it was clear that "*Exponential Moving Average*" load profile forecast had a better model fit and complexity to measure the demand response of customers paying a Flat-rate and Time-Of-Use tariff. This result shows that even though "*Mid X of Y days*" had a highest economic advantage for the Flat-rate customers (as seen in Table 8) it may not be the best fitting model for measuring demand response. This result signifies that it is left to demand response program participants to decide which is more important economic advantage or model complexity and fit. Best fitting models would generally have less errors associated with them.

Finally, it was necessary to show that as the particle swarm optimization algorithm progresses through its iterations, the solutions for the optimal load profile from which the economic advantage of a given baseline is calculated converges to a steady value. This involved examining every forecasted load profile, customer group and scenario (given in Table 6) and graphing the solutions of the particle swarm optimization for every iteration against the number of iterations. For brevity only two graphs are shown in this paper. These are shown in Figure 4 and Figure 5. These graphs show the solutions developed by the particle swarm algorithm for scenario 4 (Table 6) for a typical customers paying a Flat-rate tariff (Figure 4) and a typical customer paying a Time-Of-Use tariff (Figure 5) when measuring the demand response using the "*Historical data*" forecast. These graphs clearly show that as the optimization algorithm progresses through its iterations, the solutions for the optimal load profile from which the economic advantage of a given forecasted load profile is calculated converges to a steady value for both the Flat-rate tariff and Time-Of-Use tariff. All other graphs examining convergence had similar outcomes. This confirms the integrity of the process to calculate the economic advantage E_A and was sufficient to prove that the function to calculate economic advantage a maximum solution.

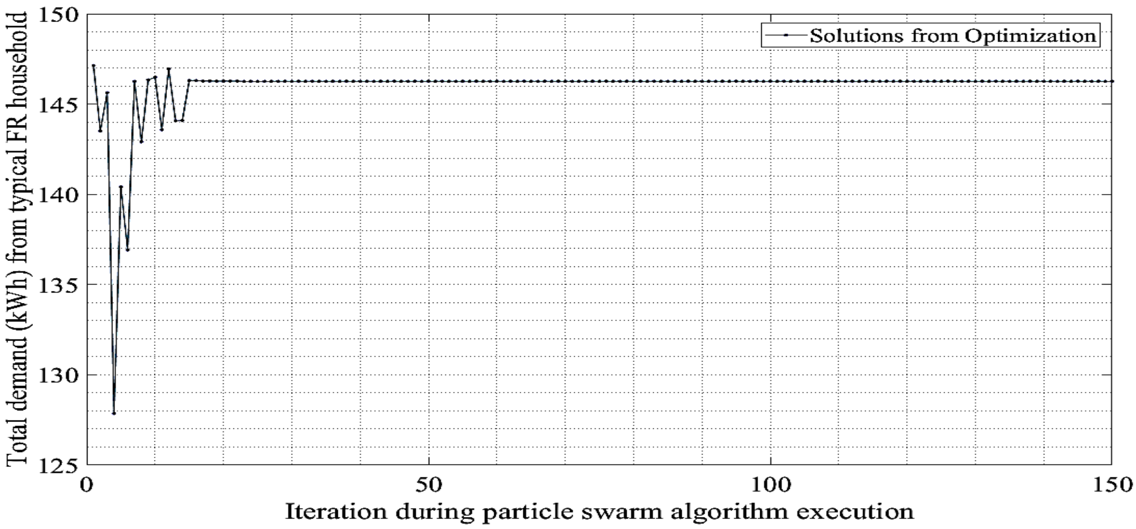

Figure 4. Model convergence for particle swarm optimization for typical Flat-rate (FR) customer

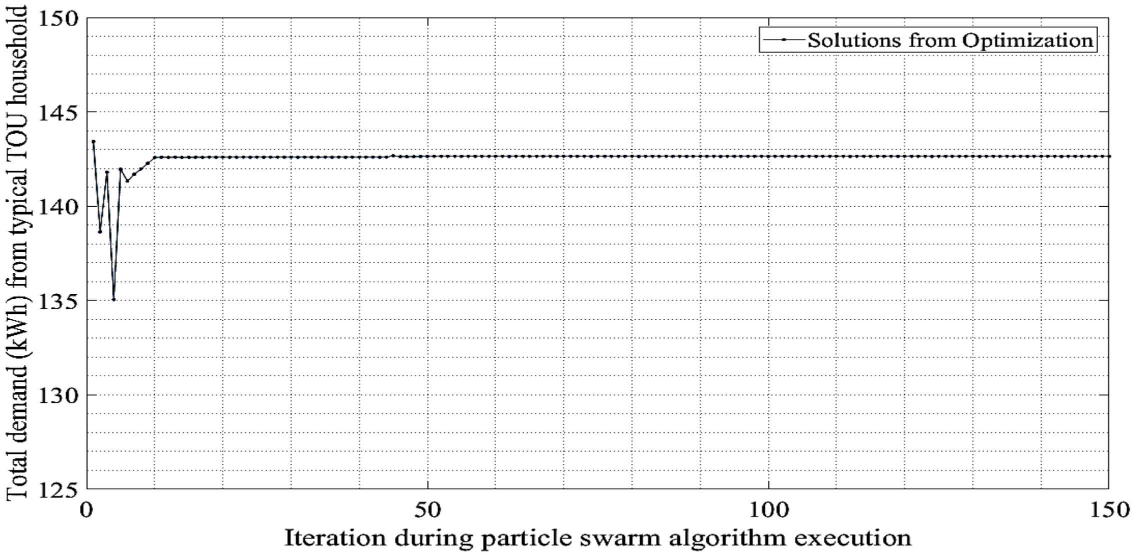

Figure 5. Model convergence for particle swarm optimization Time-Of-Use (TOU) customer

5. Discussion

Even though it was not explicitly indicated, the problem of finding the optimal baseline to measure the demand response of the customer was decomposed into a multi-objective problem. This approach was taken because multiple stakeholders would have different objectives when using residential energy management systems. In addition, the retailer may have their own objective for deploying these systems. It is unlikely in a practical setting that all customers and the retail would collaborate on which objectives shown be fulfilled to maximize the overall benefit of all stakeholders. Therefore, in solving the problem it was necessary to look at multiple objectives, assume that all these objectives were optimized separately (as would be the case in a noncollaborative environment) and then consider the best baseline that would ensure that as many of these objectives as possible were optimally satisfied.

In addition, there are some recommendations to be made with regards to refining the current model. Most of the research done on baselines to measure demand response involves creating new baseline methods. However, this paper takes the different theoretical approach and chooses to select the best baseline method using the benefits that the customer and retailer derived from the demand response produced by demand response devices. This approach acknowledges that no one baseline would ever perfectly measure the demand response of a customer; however, there are baselines that are more appropriate than others in a given situation.

Practically the results suggest that residential energy management systems can be designed to measure economic advantage and adjust its measurement parameters to ensure that the customer and retailer always benefits from the demand response produced by these devices. As such, Figure 1 can be augmented with a baseline selection component to produce a new iteration of the model as shown in Figure 6.

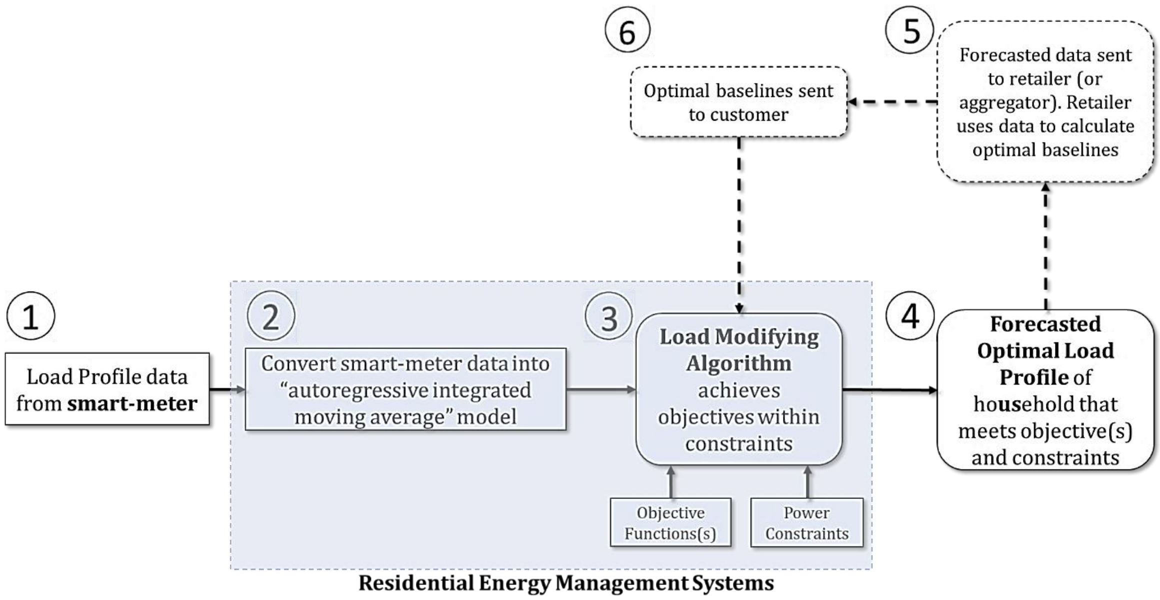

Figure 6. Updated model generalized model characterising residential energy management system

This research stressed the importance of optimally selecting a baseline from which to measure the model's performance. Nevertheless, a demand response baseline only deals with the upper constraint of the generalised model presented in Equation (2). The lower constraint was assumed to be the baseload for the household. Consequently, the lower constraint should be defined in a more comprehensive manner.

The lower constraint of the model indicates that the customer is willing to reduce load to the base load of the household. However, the customer may not reduce load to the point that only the basic needs of the household is satisfied; the customer may have other needs or may behave irrationally. More research needs to be done to specifically address the optimal lower demand response limit of the economic model.

The economic advantage metric created in this paper is based on the cost-benefit definitions for each stakeholder (b_c and b_r). These definitions need to be standardised for comparing demand response programs across the world. Research specifically focusing on standardising definitions of cost-benefits for different stakeholders and the economic advantage of different baselines presents a huge area of research that is yet to be explored. It is recommended that focus be placed in that area to enhance and add value to the current state of-the-art.

Finally, the customer and the retailer are not the only stakeholders in a demand response program. Therefore, cost-benefit metrices for other stakeholders such as the system operator can also be considered and added to the to the economic advantage formula given in Equation (17) to determine the economic benefit of these stakeholders as well.

6. Conclusion

The forecasted load profile selected to measure the performance of residential energy management system in demand response programs plays a critical role in determining the economic benefits gain from these devices. Nevertheless, there is no efficient method for selecting an optimal forecast procedure to measure the performance of these devices. To overcome this issue, this research resolved to develop an economic model for selecting an optimal forecast to measure the performance of residential energy management systems.

Using hourly electricity data was collected from 5379 households that participated in the Low Carbon London project, this research developed an economic metric to evaluate the economic advantage of 9 baseline methods across two groups of customers (customers paying a Flat-rate tariff and Time-Of-Use tariff).

This research demonstrated that the optimal forecast to use in a demand response program depends on the operating objectives of the residential energy management systems and the type of tariff that the customer pays. Using the traditional one-size-fits-all approach to measure the demand response performance may not be the most appropriate to maximize the economic advantage of demand response programs involving these devices. A better approach would be to allow the residential energy management systems to select a load profile forecast that maximizes the economic advantage of its demand response for both the customer and retailer. The device can then report to the retailer the demand and demand response expected for the household. In effect the device would ensure that, locally, the customer's welfare is maximized and globally, the retailer's welfare is also maximized.

Strategically, these results can be used to inform economic feasibility studies on demand response program involving residential energy management systems. In the case where using a single baseline may result in showing these devices may be uneconomical, it may simply be a matter of allowing these devices to select their own forecast procedure to produce a more feasible outcome. Beyond that, the economic model created here can be used to investigate possible effects that large-scale deployment of these devices can have on the retailer's portfolio and cost-benefit outcomes for the customer outcomes.

Finally, the customer and the retailer are not the only stakeholders in a demand response program. Therefore, cost-benefit metrices for other stakeholders such as the system operator can also be considered and added to the to the economic advantage formula given in Equation (17) to determine the economic benefit of these stakeholders as well. In addition, the cost benefit definitions for the customer and retailer needs to be standardised for comparing demand response programs across the world. Research specifically focusing on standardising definitions of cost-benefits for different stakeholders and the economic advantage of different baselines presents a huge area of research that is yet to be explored.

References

[1] H. Shareef, M. S. Ahmed, A. Mohamed, and E. Al Hassan, "Review on Home Energy Management System Considering Demand Responses, Smart Technologies, and Intelligent Controllers," *IEEE Access,* vol. 6, pp. 24498-24509, 2018.

[2] R. Sharifi, S. Fathi, and V. Vahidinasab, "Customer baseline load models for residential sector in a smart-grid environment," *Energy Reports,* vol. 2, pp. 74-81, 2016.

[3] S. Mohajeryami, M. Doostan, A. Asadinejad, and P. Schwarz, "Error analysis of customer baseline load (CBL) calculation methods for residential customers," *IEEE Transactions on Industry Applications,* vol. 53, no. 1, pp. 5-14, 2016.

[4] G. Lin, Y. Yang, F. Pan, S. Zhang, F. Wang, and S. Fan, "An Optimal Energy-Saving Strategy for Home Energy Management Systems with Bounded Customer Rationality," *Future Internet,* vol. 11, no. 4, p. 88, 2019.

[5] J. Jazaeri, T. Alpcan, R. Gordon, M. Brandao, T. Hoban, and C. Seeling, "Baseline methodologies for small scale residential demand response," in *2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*, 2016, pp. 747-752: IEEE.

[6] S. Mohajeryami, M. Doostan, and P. Schwarz, "The impact of Customer Baseline Load (CBL) calculation methods on Peak Time Rebate program offered to residential customers," *Electric Power Systems Research,* vol. 137, pp. 59-65, 2016.

[7] T. Song *et al.*, "A Cluster-Based Baseline Load Calculation Approach for Individual Industrial and Commercial Customer," *Energies,* vol. 12, no. 1, pp. 1-17, 2018.

[8] S. Mohajeryami, M. Doostan, and A. Asadinejad, "An investigation of the relationship between accuracy of customer baseline calculation and efficiency of peak time rebate program," in *2016 IEEE Power and Energy Conference at Illinois (PECI)*, 2016, pp. 1-8: IEEE.

[9] M. Weck, J. Van Hooff, and W. van Sark, "Review of barriers to the introduction of residential demand response: a case study in the Netherlands," *International journal of energy research,* vol. 41, no. 6, pp. 790-816, 2017.

[10] S. Nolan and M. O'Malley, "Challenges and barriers to demand response deployment and evaluation," *Applied Energy,* vol. 152, pp. 1-10, 2015.

[11] Y. Chen *et al.*, "Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings," *Applied Energy,* vol. 195, pp. 659-670, 2017.

[12] X. Wang, Y. Wang, J. Wang, and D. Shi, "Residential customer baseline load estimation using stacked autoencoder with pseudo-load selection," *IEEE Journal on Selected Areas in Communications,* vol. 38, no. 1, pp. 61-70, 2019.

[13] S. Lei, J. MATHIEU, and R. Jain, "Performance of existing baseline models in quantifying the effects of short-term load shifting of campus buildings," *SLAC National Accelerator Laboratory, Tech. Rep., SLAC,* 2019.

[14] S. LEEa, "Comparing methods for customer baseline load estimation for residential demand response in South Korea and France: Predictive Power and Policy implications," *working paper, Chaire European electricity markets,* pp. 1-51, 2019.

[15] P. Schwarz, S. Mohajeryami, and V. Cecchi, "Building a better baseline for residential demand response programs: Mitigating the effects of customer heterogeneity and random variations," *Electronics,* vol. 9, no. 4, p. 570, 2020.

[16] D. Muthirayan, D. Kalathil, K. Poolla, and P. Varaiya, "Baseline Estimation and Scheduling for Demand Response," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 4857-4862: IEEE.

[17] D. Muthirayan, E. Baeyens, P. Chakraborty, K. Poolla, and P. P. Khargonekar, "A minimal incentive-based demand response program with self reported baseline mechanism," *IEEE Transactions on Smart Grid,* vol. 11, no. 3, pp. 2195-2207, 2019.

[18] S. Mohajeryami, "An investigation of the impact of customer baseline (CBL) calculation on peak time rebate program," Master's thesis, Dept. Elec. & Comp. Eng., UNC Charlotte, Charlotte, NC, 2016.

[19] H. Takano, A. Kudo, H. Taoka, and A. Ohara, "A basic study on incentive pricing for demand response programs based on social welfare maximization," *Journal of International Council on Electrical Engineering,* vol. 8, no. 1, pp. 136-144, 2018.
[20] J. J. Dziak, D. L. Coffman, S. T. Lanza, R. Li, and L. S. Jermiin, "Sensitivity and specificity of information criteria," *Briefings in bioinformatics,* vol. 21, no. 2, pp. 553-565, 2020.
[21] J. Lorah and A. Womack, "Value of sample size for computation of the Bayesian information criterion (BIC) in multilevel modeling," *Behavior research methods,* vol. 51, no. 1, pp. 440-450, 2019.
[22] D. Varma and P. Chandrashekar, "A second-order, discretely well-balanced finite volume scheme for euler equations with gravity," *Computers & Fluids,* vol. 181, pp. 292-313, 2019.
[23] L. D. Store, "SmartMeter Energy Consumption Data in London Households," *[Online]. Avaliable: https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households [Accessed Apr. 23 2018].*
[24] "Nord Pool.," *[Online]. Avaliable at: https://www.nordpoolgroup.com/Market-data1/#/nordic/table. [Retreived 19 April 2020],* Nord Pool AS (2020).
[25] P. Jean-Paul, Tek-Tjing Lie, Timothy N. Anderson, and Brice Vallès, "The impact of residential energy management systems on electricity retail portfolios " *Proceedings of the 2020 International Conference on Smart Grids and Energy Systems (SGES 2020)*, 2020.
[26] Y. Kim, H.-g. Son, and S. Kim, "Short term electricity load forecasting for institutional buildings," *Energy Reports,* vol. 5, pp. 1270-1280, 2019.
[27] M. Beaudin and H. Zareipour, "Home energy management systems: A review of modelling and complexity," *Renewable and sustainable energy reviews,* vol. 45, pp. 318-335, 2015.
[28] Y. Ma and B. Li, "Hybridized Intelligent Home Renewable Energy Management System for Smart Grids," *Sustainability,* vol. 12, no. 5, p. 2117, 2020.
[29] H. M. Hussain, N. Javaid, S. Iqbal, Q. U. Hasan, K. Aurangzeb, and M. Alhussein, "An efficient demand side management system with a new optimized home energy management controller in smart grid," *Energies,* vol. 11, no. 1, p. 190, 2018.
[30] A. Jamil *et al.*, "An Innovative Home Energy Management Model with Coordination among Appliances using Game Theory," *Sustainability,* vol. 11, no. 22, p. 6287, 2019.
[31] N. Hajibandeh, M. Shafie-Khah, G. J. Osório, J. Aghaei, and J. P. Catalão, "A heuristic multi-objective multi-criteria demand response planning in a system with high penetration of wind power generators," *Applied energy,* vol. 212, pp. 721-732, 2018.
[32] N. Voulis, M. J. van Etten, É. J. Chappin, M. Warnier, and F. M. Brazier, "Rethinking European energy taxation to incentivise consumer demand response participation," *Energy policy,* vol. 124, pp. 156-168, 2019.
[33] A. Dehdarian, "Scenario-based system dynamics modeling for the cost recovery of new energy technology deployment: The case of smart metering roll-out," *Journal of cleaner production,* vol. 178, pp. 791-803, 2018.
[34] A. S. Gazafroudi, J. Soares, M. A. F. Ghazvini, T. Pinto, Z. Vale, and J. M. Corchado, "Stochastic interval-based optimal offering model for residential energy management systems by household owners," *International Journal of Electrical Power & Energy Systems,* vol. 105, pp. 201-219, 2019.
[35] F. Xhafa, L. Barolli, and M. Greguš, *Advances in Intelligent Networking and Collaborative Systems: The 10th International Conference on Intelligent Networking and Collaborative Systems (INCoS-2018)*. Springer, 2018.
[36] "Marco Matteini & Giorgia Pasqualetto. Cost-benefit analysis of energy management systems implementation at enterprise and programme level," *Industrial Energy Efficiency Division, Department of Energy, UNIDO. (2018)*.
[37] T. Terlouw, T. AlSkaif, C. Bauer, and W. van Sark, "Optimal energy management in all-electric residential energy systems with heat and electricity storage," *Applied Energy,* vol. 254, p. 113580, 2019.
[38] J. Khoury, R. Mbayed, G. Salloum, and E. Monmasson, "Predictive demand side management of a residential house under intermittent primary energy source conditions," *Energy and Buildings,* vol. 112, pp. 110-120, 2016.
[39] A. Ahmad *et al.*, "An optimized home energy management system with integrated renewable energy and storage resources," *Energies,* vol. 10, no. 4, p. 549, 2017.
[40] G. Xu and G. Yu, "Reprint of: On convergence analysis of particle swarm optimization algorithm," *Journal of Computational and Applied Mathematics,* vol. 340, pp. 709-717, 2018.
[41] M. A. Hossain, H. R. Pota, S. Squartini, F. Zaman, and J. M. Guerrero, "Energy scheduling of community microgrid with battery cost using particle swarm optimisation," *Applied Energy,* vol. 254, p. 113723, 2019.

[42] G. Arici, M. Dalai, R. Leonardi, A. J. J. o. R. Spalvieri, and F. Management, "A communication theoretic interpretation of modern portfolio theory including short sales, leverage and transaction costs," vol. 12, no. 1, pp. 1-11, 2018.
[43] C. Albers and D. J. J. o. e. s. p. Lakens, "When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias," vol. 74, pp. 187-195, 2018.
[44] S. Yigit and M. J. R. S. J. Mendes, "Which effect size measure is appropriate for one-way and two-way ANOVA models? A Monte Carlo simulation study," vol. 16, pp. 295-313, 2018.
[45] G. M. Huebner, I. Hamilton, Z. Chalabi, D. Shipworth, and T. Oreszczyn, "Explaining domestic energy consumption–the comparative contribution of building factors, socio-demographics, behaviours and attitudes," *Applied energy,* vol. 159, pp. 589-600, 2015.
[46] L. C. Corchón and R. J. Torregrosa, "Two remarks on consumer surplus," *Available at SSRN 3305094,* 2019.
[47] T. Hirotaka, T. Naoto, K. Shou, and O. Atsumi, "A Design Method for Incentive-based Demand Response Programs Based on a Framework of Social Welfare Maximization," *IFAC-PapersOnLine,* vol. 51, no. 28, pp. 374-379, 2018.
[48] M. Nijhuis, M. Babar, M. Gibescu, and S. Cobben, "Demand response: Social welfare maximization in an unbundled energy market case study for the low-voltage networks of a distribution network operator in the netherlands," *IEEE Transactions on Industry Applications,* vol. 53, no. 1, pp. 32-38, 2016.
[49] K. Li, B. Wang, Z. Wang, F. Wang, Z. Mi, and Z. Zhen, "A baseline load estimation approach for residential customer based on load pattern clustering," *Energy Procedia,* vol. 142, pp. 2042-2049, 2017.
[50] L. Ljung, T. Chen, and B. Mu, "A shift in paradigm for system identification," *International Journal of Control,* vol. 93, no. 2, pp. 173-180, 2020.
[51] I. Choi and H. Jeong, "Model selection for factor analysis: Some new criteria and performance comparisons," *Econometric Reviews,* vol. 38, no. 6, pp. 577-596, 2019.
[52] M. Ren, X. Huang, X. Zhu, and L. Shao, "Optimized PSO algorithm based on the simplicial algorithm of fixed point theory," *Applied Intelligence,* pp. 1-16, 2020.
[53] S. A. Mostafa and I. A. Ahmad, "Kernel Density Estimation Based on the Distinct Units in Sampling with Replacement," *Sankhya B,* pp. 1-41, 2020.
[54] Q. Yang, M. Su, Y. Li, and R. Wang, "Revisiting the Relationship Between Correlation Coefficient, Confidence Level, and Sample Size," *Journal of chemical information and modeling,* vol. 59, no. 11, pp. 4602-4612, 2019.
[55] Y. Liu and J. Wu, "Research on new energy and traditional energy based on SPSS," *International Journal of Economics and Finance,* vol. 9, no. 2, pp. 189-195, 2017.
[56] A. K. Thukral, R. Bhardwaj, V. Kumar, and A. Sharma, "New indices regarding the dominance and diversity of communities, derived from sample variance and standard deviation," *Heliyon,* vol. 5, no. 10, p. e02606, 2019.

Appendix 1. Description of customer load profile forecast procedures

Table 10. Description of Nine method generally used for creating residential customer load profile forecasts [6]

Type of Baseline	Definition
Historical Data	This is raw data collected from the smart meter
Last Y days	Historical smart-meter data is obtained, and the average demand use for every hour of the last Y most recent days is used to determine the hourly demand baseline for the current historical day.
Low X of Y days	Historical data is obtained, and the average demand use for every hour of X days with the lowest total demand out of the last Y most recent days is used to determine the hourly demand baseline for the current historical day.
Mid X of Y days	Historical data is obtained for the Y most recent days. These days are sorted in order of increasing total daily demand. The average demand use for every hour of the middle X out of the sorted Y most recent days is used to determine the hourly demand baseline for the current historical day.
High X of Y days	Historical data is obtained, and the average demand use for every hour of X days with the highest total demand out of the last Y most recent days is used to determine the hourly demand baseline for the current historical day.
Exponential Moving Average	Historical data is obtained and the moving average of the customer's load for every hour is calculated. The baseline for the current historical day is the weighted sum of the moving average and the hourly demand for the current day. The weighting for the moving average is α and the weighting for the current day is $(1-\alpha)$; where α is in the interval [0,1].
Linear Regression	A linear regression equation is constructed from variables that may affect the hourly electricity use of the current historical day. A simple construction would be to construct a linear regression equation using the historical hourly demand of the last Y days to predict the baseline hourly demand of the current day.
Polynomial Interpolation	With this method for every hour historical data of the past recent Y days and the future most recent Y days are used collected and a polynomial (usually of degree 4) is fitted to the data. The polynomial is then used to determine the value of demand for every hour of the current day. This method usually involves 24 polynomials (one for every hour).
Neural Network	A simple feedforward neural network is trained with input being historical data from the collection of variables that may affect the customer demand for the historical current day. In the simplest case historical hourly demand of the first Y days in the data set is used to train the network. The network is then used to predict hourly demand for the next Y historical days.

*Forecasts created from each method was adjusted using a multiplicative adjustment technique [8]. In general, once the reduction in demand is calculated the estimated baseline is adjusted to take into account the conditions any special conditions (for example change in weather or unusual demand use) at the time of the demand response period.

Appendix 2. Comparison between "*Historical data*" all other forecasted load profiles used in this research for a typical day for customers paying a Flat-rate tariff

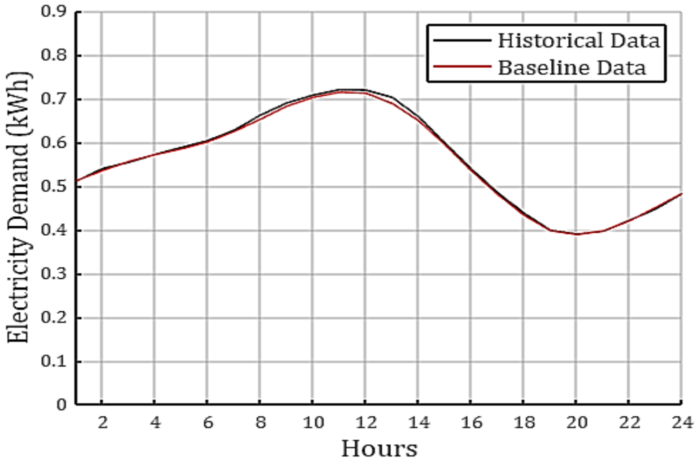

Figure 7. Comparison between "*Historical data"* and "*Last Y days"* forecasts for single day

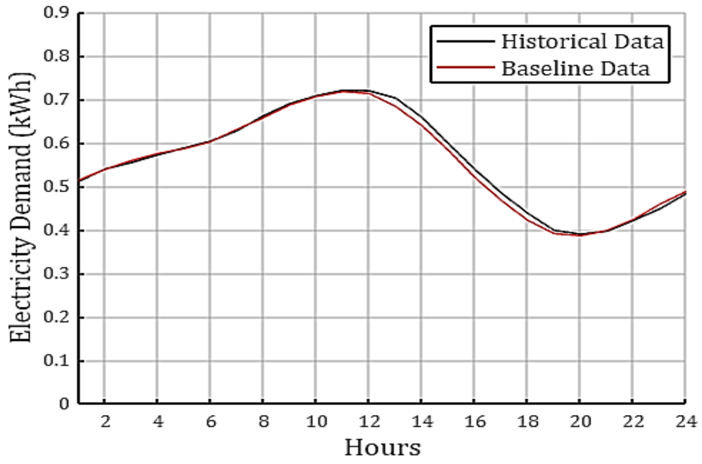

Figure 8. Comparison between "*Historical data"* and "*Low X of Y days"* forecasts for single day

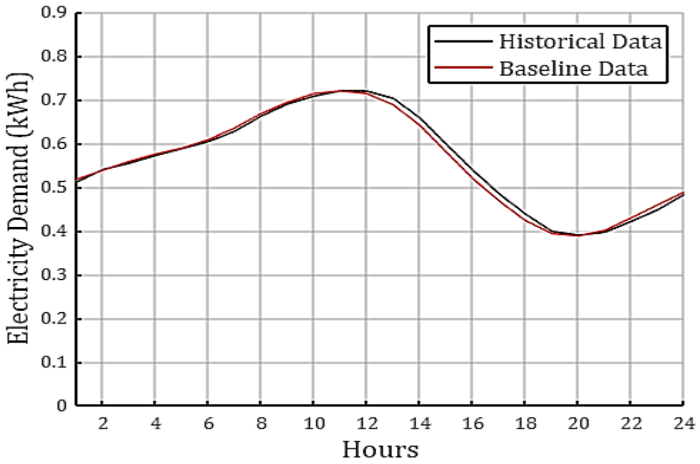

Figure 9. Comparison between "Historical data" and "Mid X of Y days" forecasts for single day

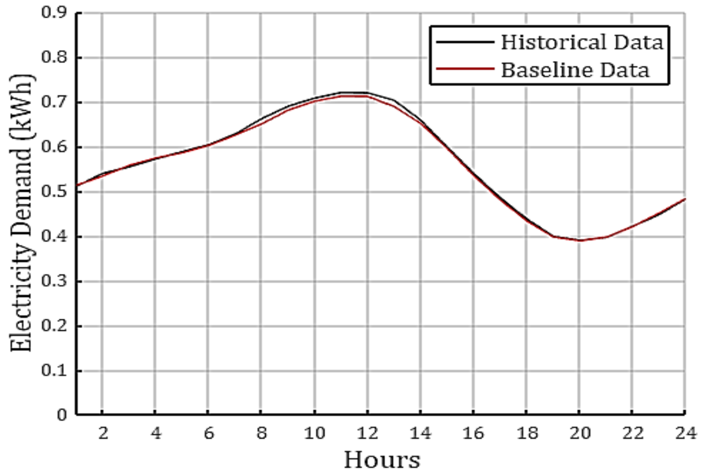

Figure 10. Comparison between "*Historical data*" and "*High X of Y days*" forecasts for single day

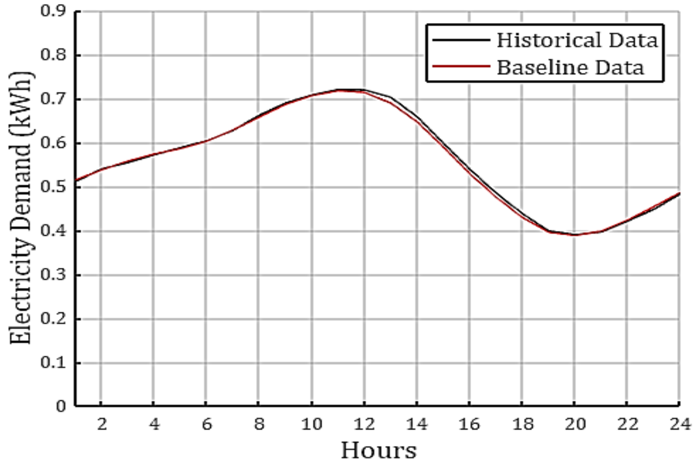

Figure 11. Comparison between "*Historical data*" and "*Exponential Moving Average*" forecasts for single day

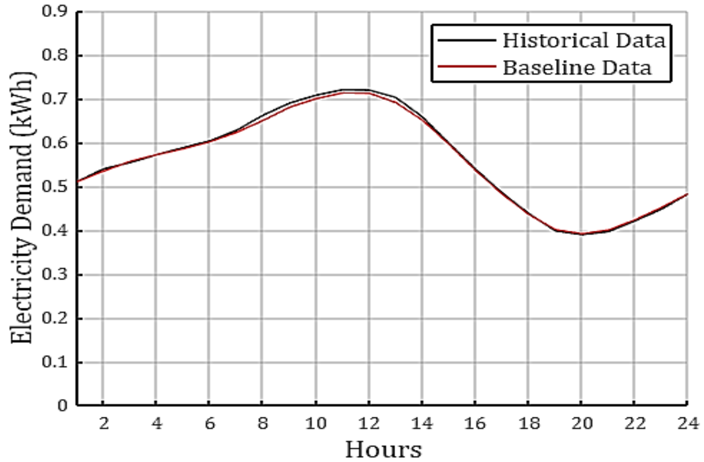

Figure 12. Comparison between "*Historical data*" and "*Linear Regression*" forecasts for single day

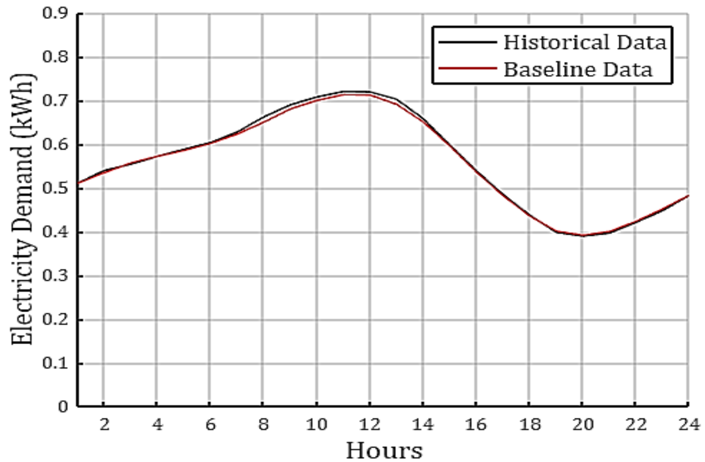

Figure 13. Comparison between "Historical data" and "Polynomial Interpolation" forecasts for single day

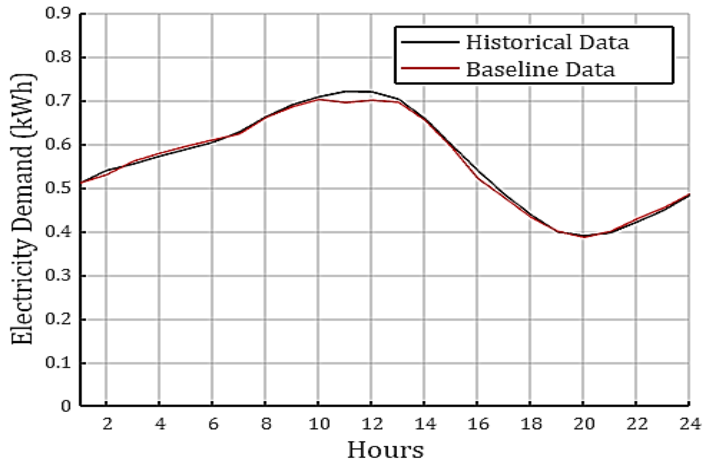

Figure 14. Comparison between "*Historical data*" and "*Neural Network*" forecasts for single day

Appendix 3. Characterizing Diversity

Customer diversity was characterized using the Simpsons Index [56]. Simpson Index was chosen because it is widely used, easy to calculate and can be adapted to different situations where diversity needs to be calculated [56]. Simpsons Index considers the proportional contribution of each member of the group for the characteristic being measured. In the case of residential customers, the characteristic being measured is the amount of electricity consumed by each type of customer. The Simpson's Diversity index (S) as shown in Equation (20).

$$S_i = 1 - \sum_{i=1}^{N} \frac{s_h(s_h - 1)}{s_N(s_N - 1)} \tag{20}$$

where	
S_i	Customer diversity during hour i
s_h	Demand of household h during hour i
s_N	Aggregated demand of households during hour i
N	Total number of households

here S_i is the customer diversity during hour "i". Once the customer diversity is known for every hour under consideration, the average customer diversity for a given period can be found.