Transfer

Mark McLeod

I. Paper Section

A. Introduction

In order to reduce greenhouse gas emissions many governments are aiming to reduce their energy use, and to reduce the proportion of energy that is used that is sourced from unsustainable sources. This means a significant increase in renewable sources which provide a supply that is much less controllable, and much less reliable than traditional supply methods. One of the key changes required to implement these changes is the development of the smart grid, in which consumers and suppliers exchange information and modify their actions in order to ensure continuous power despite uncontrollable limitations in supply. Since the ability of humans to react to price variation is limited, and certainly not reasonably sufficient to find optimal strategies for complex problems such as scheduling heating under variable tariffs and external temperatures, the use of semi-autonomous devices that are able to react to changes in the network by altering their behaviour is required. These agents will be primarily aiming to achieve maximum utility for the consumer, so the field of incentive engineering is required, to ensure that the optimal course of action from the individual agent perspective leads to desirable behaviour in the entire system.

The current state of the electricity grid is that supply adapts to meet supply and the electricity tariff presented to domestic consumers is either constant, or a very simple time-of-use tariff such as economy-7. To ensure that demand is met a hierarchy of sources are used, with cheaper, more efficient but slower responding, supplies such as nuclear used wherever possible, and fast responding but inefficient and more expensive supplies such as gas turbines used to match variation on shorter time scales. Spinning reserves accommodate extremely fast variations and storage schemes such as pumped hydroelectric take up excess when required and feed it back at a later time. Supply companies buy and sell contracts to meet any predicted demand that cannot be met with their own sources, or sell excess supply in an advance market, then pay with hindsight for the difference between this and the energy actually consumed. In order to make improvement on this process it is necessary for the consumer to be made aware in some way of the availability of power and alter their behaviour by shifting their use to away from times when demand is high and must be met by inefficient sources, and towards times where there is excess capacity that would otherwise need to be stored and reintroduced with an associated loss. Previous schemes that have aimed to implement this are discussed in section REF. The proposed method here is to introduce a tariff that varies with time according to a parametrized generating function and is communicated in advance to customers. The customers then alter their behaviour to incur minimum cost under this tariff while still achieving their goals. It is then proposed that these parameters are optimized so that the load profile induced is the best from the suppliers point of view that is available. To achieve this optimization a model of the home and the agent controlling it is developed, so far only concerning heating, and an ensemble of draws for this model is used as a model of the consumer base. Evaluating the response of this model to a given tariff provides a prediction of the load and this can be used to optimize the parameters of the tariff. Results are presented showing that the optimum parametrization of the tariff provides an improvement of the induced load in terms of flatness.

B. The Consumer Problem

To form a realistic model of the response to variation in tariff a realistic model of the home is required. [?] divides domestic loads into these that are and are not suitable candidates for influence via pricing. The later are devices that are used directly by the residents such as lighting, cooking and entertainment, the former devices such as washing machines, refrigerators and heaters. This category is further split into shiftable static load (those that could start at a selection of times but once started will run a fixed profile) and thermal loads. A model for thermal loads is developed and it is these devices that are the focus of this paper. In 2012 $77.6TWh$ of electricity was used for domestic purposes, of which $16.5TWh$ was used for space heating [?] so thermal loads represent a significant fraction of electricity use. They also have the advantage over other devices that they are likely to be always on as a background process, and are already the subject of research to produce more intelligent versions [?] [?] with some currently available models having features such as individual room control [?], highly configurable schedules [?] and occupancy detection [?]

Multiple studies have

Thermostats are getting more intelligent, internet pricing publishing has already been proposed. RT pricing has already been proposed. Advance publishing allows the thermostats to plan ahead and has already been proposed, but as market price not as control signal[?].

The smart thermostat model presented here is based on the model proposed in and in that it uses a quadratic utility cost for deviation from a target temperature during times when the thermostat is active and sums this with the cost of the power used to form the objective function. The most notable difference is that rather than requiring the heating to be either on or off the heating control is allowed to be any value between zero and one. This assumes a pulse width modulation scheme is present in the controller so that the control input multiplied by the heating system power gives the expectation of load, rather than the true load. By doing this the duration of each control period can be increased, and so the number of degrees of freedom is reduced leading to a reduced computation time. It is also assumed here that the temperature cycle is periodic over 24 hours, this removes the initial value from the problem.

A simple exponential cooling model is used, the home is defined by a thermal mass, *cm*, and insulation value, $\frac{1}{k}$, and a heater power P If the external temperature is $T^e(t)$ then the temperature is governed by the first order differential equation

$$\frac{dT(t)}{dt} = -\frac{k}{cm}(T(t) - T^e(t)) + \frac{P\delta(t)}{cm} \tag{1}$$

where $\delta(t) \in (0,1)$ determines the power setting of the heater. For a sufficiently small timestep Δt this can be approximated as an update rule

$$T_{n+1} = T_n - \frac{\Delta tk}{cm}(T_n - T^e_n) + \frac{\Delta tP\delta_n}{cm} \tag{2}$$

If the time period under consideration is discretized into N periods of Δt then T, T^e, δ can be represented as column vectors of size N and the update rule can be expressed as a matrix equation

$$\left(\mathbf{I}_{N+1\times N\times 1} - (1 - \frac{\Delta tk}{cm})\mathbf{U}\right)\underline{T} = \frac{\Delta tk}{cm}\mathbf{U}\underline{T}^e + \frac{\Delta tP}{cm}\underline{\delta} \tag{3}$$

which is an affine relation between temperature and input that can be expressed as

$$\underline{T} = \underline{\Psi} + \mathbf{\Phi}\underline{\delta} \tag{4}$$

where

$$\mathbf{\Phi} = \frac{\Delta tP}{cm}\left(\mathbf{I}_{N+1\times N\times 1} - (1 - \frac{\Delta tk}{cm})\mathbf{U}\right)^{-1} \tag{5}$$

$$\underline{\Psi} = \left(\mathbf{I}_{N+1\times N\times 1} - (1 - \frac{\Delta tk}{cm})\mathbf{U}\right)^{-1}\left(\frac{\Delta tk}{cm}\mathbf{U}\underline{T}^e\right) \tag{6}$$

and

$$\mathbf{U} = \left[\begin{array}{c|c} \underline{0}_{1\times N} & 1 \\ \hline \mathbf{I}_N & \underline{0}_{N\times 1} \end{array}\right] \tag{7}$$

The quantity that the home controller needs to minimise is the sum of the cost of deviation from the target temperature profile $\underline{T}^s$ and the cost of heating. The temperature deviation is only relevant when the home is occupied. This is encoded in the matrix $\mathbf{Q} = q \times diag(\underline{o})$ where $\underline{o}$ is a vector of size N with ones in positions where the home is occupied and zeros otherwise and q is a scalar that relates the utility of temperature deviation to monetary cost. The cost of heating is given by $P\underline{\Delta}^T\underline{\delta}$ where $\underline{\Delta} = [\lambda]$ is the vector of size N giving the unit cost of electricity over time. This gives an objective function

$$f = (\underline{T} - \underline{T}^s)^T\mathbf{Q}(\underline{T} - \underline{T}^s) + P\underline{\Delta}^T\underline{\delta} \tag{8}$$

So that the dimensionality of the problem can be reduced to a lower number than that defined by the resolution of the temperature update equation $\underline{\delta}$ is defined to be

$$\underline{\delta} = \mathbf{D}\underline{u} \tag{9}$$

where $\underline{u}$ is the M dimensional column vector that will be optimised over and D is the $M \times N$ matrix

$$\mathbf{D} = \mathbf{I}_M \otimes \underline{1}_{J\times 1}, \qquad N = MJ \tag{10}$$

.

Making these substitutions leads to the new objective function

$$f' = \underline{u}^T\mathbf{R}\underline{u} + \underline{S}^T\underline{u} \tag{11}$$

which differs from the original objective by a constant and where

$$\mathbf{R} = \mathbf{D}^T\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi}\mathbf{D} \tag{12}$$

and

$$\underline{S}^T = 2(\underline{\Psi} - \underline{T}^s)^T\mathbf{Q}\mathbf{\Phi}\mathbf{D} + P\underline{\Delta}^T\mathbf{D} \tag{13}$$

The only constraints strictly required are those that constrain the heater input to be set between zero and one.

$$0 \le \delta_n \le 1 \qquad \forall n \tag{14}$$

These are expressed in standard form as

$$\left[\begin{array}{c} \mathbf{I}_M \\ \hline -\mathbf{I}_M \end{array}\right]\underline{u} \le \left[\begin{array}{c} \underline{1}_{M\times 1} \\ \hline \underline{0}_{M\times 1} \end{array}\right] \tag{15}$$

In the following results further constraints are also placed on the absolute maximum and minimum temperature over all time, and that the temperature may not deviate by more than two degrees from the target temperature.

This specifies the procedure an individual smart thermostat will undertake to form an optimum heating plan. To perform this optimisation using 15 minute control periods and 1 minute cooling simulation requires approximately three seconds as a single threaded process on an Intel i7 processor. This leads to the set of 60 agents used in the evaluations below taking approximately 40 seconds to find the response to a given tariff. The results show that 60 agents still have significant differences in response between draws meaning greater numbers would be desirable. This is clearly therefore a sufficiently expensive to evaluate objective to justify the use of the techniques described below.

C. The Supplier Problem

In this work we consider an electricity supplier with a set of customers $c \in C$. Each of these customers has a load $l_c(t)$ so that the load to be met by the supplier is

$$L(t) = \sum_{c\in C} l_c(t) \tag{16}$$

There exists some utility functional for production $U_p(L(t))$ which specifies how much it costs the producer to supply this power. U will be based on the sources of power available to the supplier, and the market prices at which it must sell excess and buy shortfall. Each consumer is charged at a rate which we allow to be a function of time $\lambda(t)$. The balance of payments for the supplier is therefore

$$\Delta = \int_T L(t)\lambda(t)\mathrm{d}t - U_p(L(t)) \tag{17}$$

and the supplier naturally wishes to choose λ to maximize profit (subject to constraints such as not being so high that consumers will choose to use other supplies). To simplify the mathematics for this work we assume that the tariff will balance the component of U which is due to the average power consumption, so are only concerned with the shape of L rather than its absolute value. The following results use the integral of absolute difference from the mean power, but other options are of course available and are discussed in section REF. Further, the load is divided by end use into load that is and is not controllable via Λ and only that component which is controllable is considered.

To model the response of the full set of customers multiple agents are created using a generative process to determine individual parameters and the sum of their load is used. That is is a single thermal agent from the set of agents in the ensemble $a \in A$ is characterized by a vector θ_a and the optimization process that produces a load profile $\underline{l}$ given a tariff profile $\underline{\Lambda}$ is denoted

$$\underline{l} = h(\underline{\Lambda} \mid \theta) \tag{18}$$

then the load $\underline{L}$ produced by the entire ensemble of agents is

$$\underline{L}(\underline{\Lambda}) = \sum_{a \in A} h(\underline{\Lambda} \mid \theta_a) \tag{19}$$

To provide an suitable target for optimization the load profile vector must be mapped to a scalar that conveys the utility to the supplier of that load. For the following results the objective used is the normalized integral of the deviation from the mean load.

$$U(\underline{L}) = \sum_{t=0}^{N} \frac{|\underline{L}[t] - \langle L \rangle|}{\langle L \rangle} \tag{20}$$

Various other possibilities are discussed in REF.

There is of course the possibility that a tariff might exist which exactly flattens the load and which can be found in a deterministic manner. For a single agent this would mean a constant power draw and if a value exists that satisfies the agent constraints then the tariff that induces this behaviour can be found by setting the derivative of equation 11 with respect to $\underline{u}$ to and rearranging to obtain

$$\underline{\Lambda} = -\frac{1}{P}\left(2\mathbf{R}\underline{u} + 2(\underline{\Psi} - \underline{T}^s)^T \mathbf{Q}\mathbf{\Phi}\mathbf{D}\right)^T \mathbf{D}^T (\mathbf{D}\mathbf{D}^T)^{-1} \tag{21}$$

However, this problem grows with the number of agents so will become intractable with larger numbers of agents, and since $\underline{\Lambda}$ must satisfy equation 21 for some constraint satisfying $\underline{u}$ for every agent present, and maintain the sum of all agent loads exactly flat there is no reason to believe that a solution will exist for non-trivial numbers of agents. Optimizing to move the load as close as possible to the desired shape is therefore the best that is likely to be achieved.

D. *thermal agent generative process*

In order for the model to provide useful information it must have similar characteristics to the true demand system. Creating a model with structure and parameters sufficiently well chosen to provide responses that are close to reality is discussed in section REF. For this work more simpler distributions, and nominal values have been used to provide a model that can be used for optimisation and is hopefully not too dissimilar to reality. The full set of parameters used are listed in table I

To determine the times that the home must be maintained at the disired temperature the work of [?] is used. They propose a Markov chain generative model based on survey data. The model provides transition matrices for the number of active occupants in a building every ten minutes throughout the day, with separate chains for buildings with between one and six occupants. For this it is required to specify the number of occupants living in the building. For this a discrete distribution is used, based on data collected in the CABE dwelling size survey [?] which lists the number of occupant for 200 English homes.

TABLE I
PARAMETERS OF HOME MODEL AND THEIR DISTRIBUTION

Parameter	Model	Source
Occupancy Number	Multinomial	[?]
Occupancy over time	Markov-Chain	[?]
Floor area (A)	Offset Gamma	[?]
Flat or House	Binomial	[?]
Insulation Value (House)	$3.6\sqrt{A} + 0.14A$	-
Insulation Value (Flat)	$3.3\sqrt{A}$	-
Electric Boiler Power	Uniform	-
Thermal Capacity	Log-Normal	[?]

For the thermal parameters of the home the floor area is drawn from an offset gamma distribution chosen to match the floor area results found by the CABE survey. Nominal values are used to convert floor area to surface area and from these values to derive insulation and thermal mass values which are drawn from log-normal distributions based on the nominal values given in The Governments Standard Assessment Procedure for Energy Rating of Dwellings [?]. Future improvements to this procedure are discussed in section REF

Analysis of the individual agent optimisation and QP solution.

Basic spec of the quadratic comfort cost and exponential cooling is a citation of [?]. Add a few more constraints, remove the asymmetry, make it cyclical to avoid initial conditions. Expectation of PWM rather than direct on-off input.

E. *Optimization*

The objective function we with to evaluate is

$$g = U(\underline{L}(\Lambda)) \tag{22}$$

Rather than search over the full N dimensional space of Λ we specify the tariff to be smoothly varying and defined by some parameterised function $\lambda(t \mid \theta)$. The search then takes place over the much smaller space defined by θ. For the following results we interpret θ as equally spaced support points of a cubic spline over the period, but with an additional constraint that the tariff is bounded above and below by some maximum and minimum value.

Since the function, g, we with to optimize involves the solution to a constrained multidimensional quadratic for each agent in the ensemble it requires significant computing time to evaluate. It is therefore worthwhile to invest some additional computing time in selecting the

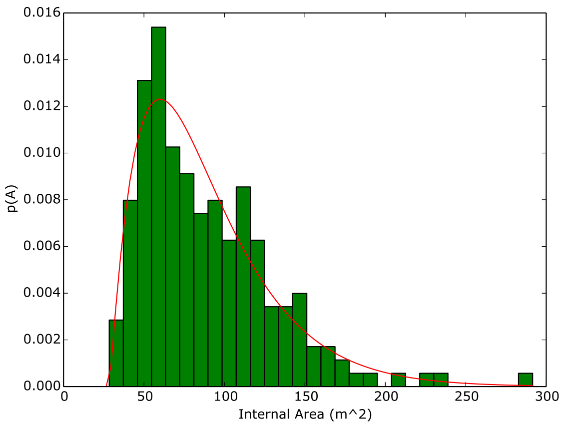

Fig. 1. The internal floor area of a home. The histogram shows data obtained from the CABE study. This has been fitted to an offset Gamma distribution with $shape = 2.10$, $scale = 28.70$, and $offset = 28.39$ shown in red. The mean and variance of the histogram and the fitted Gamma distribution are equal, the offset it that which minimised the squared error between the Gamma distribution and the histogram at the bin centres.

next point to evaluate, rather than following grid-search or multi-start line-search algorithms. The method used is Gaussian Process Global Optimization (GPGO)[?] in which given a mean and kernel function a value for the expectation and variance can be obtained for any point in closed form given the points that have been evaluated. Specifically, when evaluating a function $y = f(\underline{x})$ it is assumed that the value of y has a prior given by $\mu(\underline{x})$ and is jointly Gaussian at all points with the covariance between two locations x and x' defined by a valid kernel unction $k(\underline{x}, \underline{x}')$ given a set of points $\mathbf{x}_0$ that have been evaluated as $\underline{y}_0$ then the posterior, $p(y_s)$ at a point of interest $\underline{x}_s$ is normally distributed

$$y_s \mid \underline{x}_0, y_0 \sim \mathcal{N}\left(\langle y_s \rangle, \mathrm{var}(y_s)\right) \tag{23}$$

where

$$\langle y_s \rangle = K_{sx} K_{xx}^{-1} y_0 \tag{24}$$

$$\mathrm{var}\left(y_s\right) = K_{ss} - K_{sx} K_{xx}^{-1} K_{sx}^{T} \tag{25}$$

$$K_{ss} = k(\underline{x}_s, \underline{x}_s) \tag{26}$$

$$K_{sx} = \left[k\left(\mathbf{x}_0[i], \underline{x}_s\right)\right] \tag{27}$$

$$K_{xx} = \left[k\left(\mathbf{x}_0[i], \mathbf{x}_0[j]\right)\right] \tag{28}$$

In implementation the Cholesky decomposition of K_{xx} is used rather than the inverse in order to reduce computing time and avoid numerical problems. The expectation and variance of the posterior of the function can be combined to provide some metric for the utility of an evaluation at that point. The utility used is here is the expected improvement (EI) over the minimum so far observed, which is defined as

$$EI(\underline{x}_s \mid \mathbf{x}_0, \underline{y}_0, k) = \int_{-\infty}^{y_{opt}} (y_{opt} - y_s) p(y_s) \mathrm{d}y_s \tag{29}$$

where

$$y_{opt} = \min(\underline{y}_0) \tag{30}$$

The choice of mean and kernel and utility functions and the hyperparameters of the kernel function and their prior of course have a major influence on the effectiveness of the optimization. In the results presented below a zero mean function and EI utility function have been used. For the kernel a squared exponential function has been used with independent Gaussian priors over the hyperparameter values. The hyperparameters are updated under to the maximum posterior values given the observed points between evaluations. Searching for the maximum EI location, and for the hyperparameter values is done using the DIRECT search algorithm [?]. Alternatives to these choices are discussed in section REF.

F. Results

The tariff is chosen to be a generated by a cubic spline passing through equally spaced support points. The support points are constrained to lie between five and 40 pence per kWh. Since when adjacent support points have a large difference this leads to a curve that can go considerable outside the range the tariff is also clipped so that it remains between five and 40 pence per kWh. The tariff is tested on a set of 60 models of a home created from a generative model with parameters drawn independently according to the distributions shown in table I. These agents optimize their load profile while still fulfilling their constraints using a 15 minute resolution for the control input, and a one minute resolution for cooling simulation, that is in equation 10 $N = 1440$ and $M = 96$. The utility of the induced load is the negative of the sum of the difference from the mean load over the day, normalised against the mean load. Intuitively this would mean that while a continuous unit load would have a utility of zero a load of two units for half the time and zero for the rest would have a utility of -1.

A squared exponential kernel is used in the Gaussian process. Inputs are expressed in pounds per kWh so the prior over input scales are chosen to be lognormal distributions centered on 10^-1 with standard deviations of one decade. The prior over the output scale is also lognormal, with a mean of 10^0 and a decade standard deviation. The optimization is permitted to run for 250 evaluations of the objective. The first twelve points are selected with at random to initialize the GP, with independent uniform distribution over the support points. The covariance is then updated to the MAP estimate after the initialization. Since the evaluation of the kernel likelihood required a cholesky decomposition the time required grows rapidly with the number of points evaluated. The hyperparameters are therefore only optimized once for every ten evaluations of the objective up to the first 80 evaluations, and every 15 thereafter. In both searching for hyperparameters and for expected improvement the DIRECT algorithm is used.

Figure 2 shows the performance improvement using the best tariff found at various degrees of freedom for both test and training data. A tariff with four degrees of freedom provides a 6.7% improvement of the objective function compared to a flat tariff, seven degrees of freedom allow the greatest observed improvement in the training set of 12.6% improvement. It is clear that the use of the variable tariff has caused the induced load to be flattened, and this flattening is increased with more degrees of freedom in the tariff. It is worth noting that it is quite

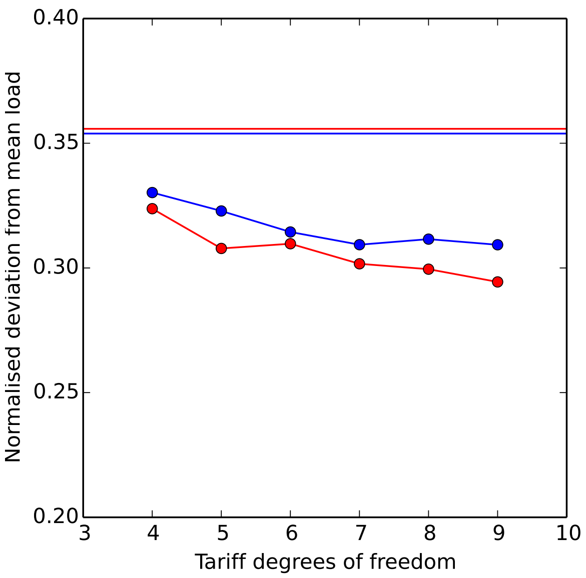

Fig. 2. The performance in terms of deviation from the mean load of the best cubic spline tariffs obtained with varying degrees of freedom on the training set of 60 agents (blue) and on a test set also of 60 agents (red). The tariff is constrained between 5 and 40 pence per kWh. Four is the minimum number of support points required to define a cubic. The horizontal lines shows the response to a flat tariff, the value taken by the flat tariff affects the response by only 1% over the full range of 5 to 40 pence per kWh.

likely that there is further improvement available with further searching about 7 degrees of freedom where the curve begins to flatten, as for these searches the expected improvement available at the final point evaluated was orders of magnitude greater than the final evaluation of the searches in lower degrees of freedom. The 4 DOF search actually terminated before 250 evaluations since the expected improvements at all points tested by the DIRECT search was less than the numerical precision and evaluated as zero. This trend is also present on a test set of equal size, which demonstrates that the optimization is not over fitting the tariff to the individual agents in the test data.

Figure [**?**] shows the load curves over the day. It can be seen that the load with a flat tariff begins at about $1MW$ until roughly 7am where there is a significant spike of roughly $4.5MW$ following which it returns to a reasonable flat $2.5MW$ for the rest of the day. The optimized tariffs make significant reductions in the peak of the spike in the morning by shifting this load earlier.

Normally it would be appropriate to compare the price paid by the consumers under the optimized tariff to the price under the flat tariff to determine whether there is an incentive to switch. However, the response of the objective function is primarily to the shape of the tariff rather than to the value. This means that the utility of the response under a flat tariff at the maximum value of $40 pence per kWh$, which would of course lead to all consumers making savings, is only 1% different from the flat response at at the minimum tariff of $5 pence per kWh$, in which case all consumers would pay more. It is therefore clearly possible to ensure that customers would prefer to select a variable tariff by appropriate selection of the value of the flat-tariff alternative. To find the truly optimal tariff it is necessary to attach a realistic to the coefficient to the distributor of non-flatness and sum this with the tariff collected from the consumers to find the real cost of a given load. This is discussed in section REF.

G. Conclusion

A probabilistic model for the thermal behavior and requirements of a domestic home, and an intelligent agent that can calculate optimal behavior for the home given a continuously varying tariff have been defined. Using a training set drawn from this model it has been demonstrated that the shape of the electrical load responds to variations in the shape of the tariff. Using a parameterized tariff an expensive global optimization method has been used to find the parameters that maximize the flatness of the induced load for varying degrees of freedom, and it was found that the improvements obtained increased with increased degrees of freedom in both training and evaluation data. It is therefore concluded that searching for variable tariffs is a feasible method for influencing the shape of the load, and that it is worthwhile to develop both the model and the search technique with the aim of achieving a model that is sufficiently similar to real consumer behavior, and searching that is sufficiently efficient that tariffs can be proposed for real application in reasonable time.

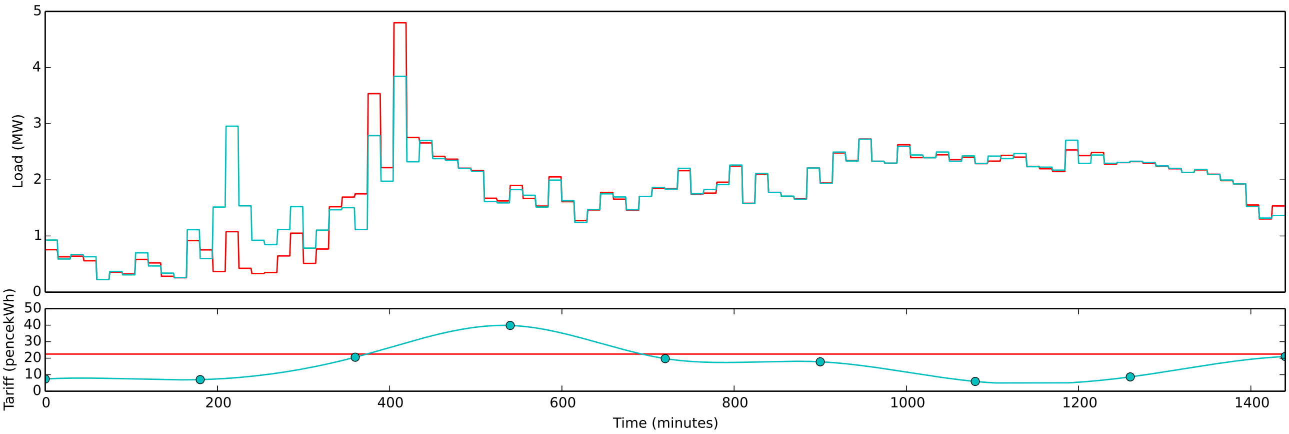

Fig. 3. Load response of an ensemble of 60 agents to variable tariffs. The upper axis shows the load in MW while the lower axis shows the tariff in pence per kWh with support points shown as filled circles. Red is a flat tariff for reference, green is the best tariff found using four support points, blue is the best tariff found using seven support points.

II. Review/Proposal

A. Review-optimisation

Various methods are available for global optimization. Optimization over functions that are easy to evaluate can be done using local search techniques with multiple start points to ensure that all local minima are found, by simple gridsearch with a spacing at the required accuracy or more intelligent but grid-like deterministic methods such as DIRECT. ***branch and bound, trust region, pattern search***.

However, in the case of functions that incur significant cost to evaluate it is necessary to invest significant effort into determining the next point to evaluate. Surrogate surfaces techniques are commonly used in this situation, in these an approximation to the objective function is made under some model based on the points that have been observed so far. This approximation, known as the surrogate surface, is cheap and fast to evaluate so the prediction about the true objective can be exploded in order to carefully choose the next point to evaluate.

[**?**] provides an overview of some more common surrogate surface techniques. Regression methods with finite degrees of freedom are quickly as they do not pass through all the evaluated points and fail even a simple example function. These methods are only useful if a model that is known to be a very good fit for the objective is available. CITE next consider methods which interpolate between the evaluated points using weighted sums of fixed basis functions centred at the evaluated points, such as splines or polynomials. Evaluating at the minimum predicted by these methods provides better results than finite DOF models but can still fail to find the global minimum. The most promising methods are those that have a basis in probability, viewing the objective function as an unknown variable. Rather than evaluating at the minimum of the surrogate surface they form a measure of the utility for evaluation at a point using a combination of both the prediction and its uncertainty at that point. By doing this it can be ensured that all points in the space will eventually be visited, which is a required condition for guaranteeing that the global minimum will be found CITE. This is the basis of Gaussian process global optimisation, which is the method used here, although CITE name it kriging.

The mathematical basis of GPGO by which the function is modelled as a draw from a space of functions that are jointly Gaussian between all points (with the correlation between points defined by a kernel function) and expectation and covariance at a point given existing points is given above

Surrogate surface techniques. GPs. different maximization options, EI, PI, Entropy. multistep lookahead. Sing the S

[**?**] lists three common acquisition functions, the expected improvement, the probability of improvement and global confidence bounds. They also propose a per-second variation of expected improvement for objectives which have variable evaluation times. They suggest that expected improvement is the preferred method, and show that it does provide better performance than the confidence bound metric on some test functions. For this reason expected improvement is the method used here.

Other methods for selecting points based on a Gaussian process model of the objective are suggested by [**?**] and CITEsoo. Hennig rather than aiming to make improvements on the observed value of the function aims to improve an entropy based measure of knowledge about the location of the global minimum. CITE uses a tree exploration structure to avoid the requirement to search over the acquisition surface, but combines this with a Gaussian process confidence bound method for deciding which elements of the tree to explore. They show that this method performs similarly to GPGO in problems of low dimensionality, but is able to converge faster for higher dimensional problems.

In all methods involving it is of course important to ensure that

B. utility of loading

Various possibilities exist for evaluating the utility of a proposed load profile. Since a value proportional to the total energy used over the day will be recovered in tariff payments the utility of a profile must be in some way related to its variations from the mean. The simplest model, the one used above, is arrived at by assuming that any variation will incur a cost proportional to it's magnitude as that difference must be bought or sold on the electricity market. Alternatively by assuming that a cheap but relatively constant power source such as nuclear or hydroelectric is available to provide the mean value and any excess must be accommodated by faster responding, but more expensive, sources such as oil or gas turbines. The utility is therefore the sum of absolute deviation from the mean, possibly with an asymmetric gradient for values above or below the mean, possibly with varying weightings through the day according to variations in market price. Other possibilities are the load factor (the ratio of maximum to mean load) which is used by [?], or the absolute maximum value, or the ratio or difference of maximum to minimum loads. In reality the utility of a loading would be a complex function related to availability and response times of all the available power sources and the daily variations in wholesale price. This functional is likely to vary daily if a significant proportion of green energy makes up the supply. Wind and solar power will vary significantly from day to day while tidal power will have a gradual trend according to the lunar cycle and ideally the demand would be shifted to match the change in supply. Therefore the aim is develop and optimization method to cater for any arbitrary functional mapping a load vector to a utility scalar.

C.

Choice of covariance function is one of the most important factors in the effectiveness of the optimizations, since it determines how well the Gaussian process is able to make predictions about the objective, and how certain those predictions are considered to be. Many kernel functions are expressed in terms of a weighted euclidean distance between points, where the weights of each axis, θ_d, constitute most of the hyper-parameters of the kernel.

$$r^2(\underline{x},\underline{x}') = \sum_{d \in D} \frac{1}{\theta_d}(\underline{x}_d - \underline{x}'_d) \tag{31}$$

among these are some of the most commonly used functions such as the squared exponential and the Matern 5/2. In the tariff setting problem where the aim is to improve the flatness of the induced load gradient may be of greater impact on the output than absolute value, so applying covariance functions not to the tariff support points, but to a linear transform of the support points that switches to differences between adjacent points and a single mean value may provide useful results.

D. Reviewtariffsetting

[?] proposes a method of homeostatic control, in which the consumer is sent a signal containing both a carbon intensity prediction, and a target for consumption relative to the same time the previous day. The agent then adapts its schedule for charging and discharging local storage to achieve a minimum cost load profile. Under this scheme the carbon emission of the system is reduced by 25% and the consumer costs by up to 14.5%. Partly this is not desirable since it relies on the consumer being willing to subscribe to a scheme which aims to reduce cost weighted against with carbon emission rather than the absolutely cheapest method. Since 10.4% of English households were considered to be in fuel poverty in 2012 [**?**] this scheme is unlikely to have complete uptake. The homeostatic method also relies on local home storage capacity as the mechanism by which the load the consumers present to the grid can be altered. Local storage may become prevalent in future, particularly if electric cars become common, but is not available at present. Furthermore this means that the homeostatic method is doing nothing to alter the actually pattern of use in the home which is the aim of the method proposed here.

[**?**] uses the model of the consumer on which the one used here is based. Loads are classified as unmoving, shiftable-static or thermal and the agent optimizes its scheduling given a tariff schedule to provide maximum utility at minimum cost. However, they use an adapted real time price mechanism in which the consumer is exposed to the predicted wholesale market price of electricity. This method causes the optimal behavior of agents to tend to create new peaks at the minimum price instead of smoothing the load towards a flat profile. To remedy this they propose that agents use a learning mechanism in which they gradually shift their scheduling from the initial maximum utility times towards the optimal time. This method prevents the formation of new peaks and using it they show that the demand converges to a calculated optimum. This again requires the agent to not be purely cost oriented. They use a learning rate of 0.05 to produce convergence, while an agent purely concerned with cost would of course learn move immediately to its optimum schedule, an effective learning rate of one. The method proposed here manipulates the tariff that the agent observed, such that the optimum schedule for the agent is also the optimum schedule for the supplier. Cooperation by the consumer is not required.

Smartgrid overview. Other proposals for load management. This method is new because it both publishes ahead of time (giving agents time to plan ahead and be more flexible) and is not a direct exposure of market prices so is a proper control signal. some use carbon value, not really very good.

E. Proposalmodel

In an implementation of the proposed scheme individual thermostats would learn the thermal characteristics of the home they were controlling and the occupancy habits of the owners. [**?**] also uses a quadratic method to derive an optimal heating schedule given either a tariff or a carbon intensity schedule, but also combines this with a Gaussian process technique for learning the thermal parameters of the home and combining these with local weather forecasts to provide an accurate prediction of heating requirements.

The above work used only the sum of customer heating loads for optimization and aimed to flatten the profile as much as possible. This is based on the assumption that the supplier has cheap access to their own generating capacity which is roughly equal to the total demand, but that deviations from the mean must be purchased or sold at a greater cost. Work in the immediate future will combine the tariff-influenced heating loads with uncontrollable loads to produce a more accurate model of the demand, and will aim to shift the load profile to fulfil more complex objective that is flexible enough to account for a varying portfolio of owned sources, all with their own cost profiles, and varying prices in the wholesale market. For example the supplier may wish to deliberately cause a peak in demand to match the short term availability of power from a tidal hydroelectric facility. It may also be useful to include optimisation of large-scale storage capacity, such as pumped storage to produce a three stage process for determining the utility of a tariff. For a given tariff first the demand in computed, then the optimal storage strategy is found, then the running costs, tariffs collected and cost of sale and purchase of imbalance are combined to give the final utility of the tariff.

To make a useful model rather than a proof of concept it will be necessary to make improvements to model of the home, and to the generative process used to create sample houses. In the work above all parameters have been drawn independently when they will of course have correlations, in many cases quite strong. For instance floor area and number of occupants is almost certainly correlated. Insulation values, floor area, boiler power and inclination to allow deviation from the desired temperature are all quite likely to be linked to the household income, as is whether the building as a flat or house. For future work including external temperature variations the external conditions will of course be geographically linked, and location on a large scale can be linked to income, and if considering larger countries with multiple time zones the times that he heating will be running.

To create such a model it will be of course be necessary to obtain sufficient data. Occupancy and income are gathered in the national census and can be linked to location through this source. External temperature is of course well documented over time and location by the MET office. Standard assessments of thermal characteristics are made on all new buildings, and have been made on a number of existing buildings under schemes such as the Warm Front Scheme[**?**]. If a variable tariff and smart-thermostat scheme were to be implemented in reality it would of course be possible to implement anonymous data gathering of parameters from the meters themselves, so a model need only be sufficiently accurate to demonstrate that worthwhile savings are possible and to act as a prior in the startup period while real data was scarce. In a real implementation weekly and monthly patterns, local and national holidays, weather forecasts would all be included in the next days demand model in order to obtain the best possible tariff.

F. Proposal-optimisation

One of the drawbacks of standard GPGO using the expected improvement metric is that the choice of the next point to evaluate under this basis is not actually what is desired, except for the very final evaluation. What is truly desired is the point that will lead to eventually finding the minimum many steps later.

One method proposed to improve on this situation is to consider the expected improvement two or more steps ahead conditioned on evaluation at the proposed next point. This would ideally be extended to consider the expected improvement at the final point to be evaluated, given all permutations of intermediate points between termination and the present, but this is not feasible in any simple implementations as the computational load grows very rapidly with the number of steps taken. Two methods for achieving faster computation are proposed as potential avenues of investigation. First, since the expected improvement multiple steps ahead for evaluation at a proposed point has become in itself an expensive to evaluate function it is possible that it is worthwhile to use GPGO with a lower lookahead window to minimize the acquisition function. That is a GP will be fitted to the expected improvement in the surrogate surface and used to choose the next point to evaluate. The aim will be to determine whether the additional cost of maintaining a GP as an inner process is worth the reduced time spent on acquisition function evaluations. The second proposed opportunity to achieve faster optimization is one

The main avenue of research proposed in the area of optimization is to attempt to create a principled approach to the case where the noise on the evaluation of the objective function is a variable, with increased accuracy coming at increased cost. Specifically in addition to choosing the next point to evaluate the accuracy of that evaluation must be selected, with some budget on evaluations causing less future evaluations to be possible if greater accuracy is chosen. This is applicable in finite element simulations, where the results become more accurate with smaller mesh sizes; in hyperparameter selection for classifiers, where larger training sets prevent over-fitting; and in simulations with particle filter or ensemble models such as the home heating model described here, where the behavior tends towards the continuous case as the number tens to infinity. The entropy approach to optimization proposed by CITEhennig seems to be the most suitable for development in this area given its focus on knowledge about the location of the minimum rather than greedy improvements to the current best.

One immediately apparent method of implementing such a process is to propose some function that represents a prior belief of how much information about the location of the global minimum can be gained within a given budget. Specifically if the function is as usual over $\underline{x} \in \mathbb{R}^d$ and the points $\mathbf{X}_0$ have already been evaluated then using information I about the location of the global minimum as the reciprocal of entropy used by CITEHennig we define three information quantities:

$$I(\mathbf{X}) \tag{32}$$

the information conferred by the evaluation of the points in $\mathbf{X}$,

$$\langle \Delta I(\underline{x}_p, b \mid \mathbf{X}) \rangle \tag{33}$$

is the expected gain in information by evaluating at the point $\underline{x}_p$ with the accuracy available to us using a computational budget b given the points $\mathbf{X}$ that have already been evaluated and

$$\langle I_g(\beta) \rangle \tag{34}$$

the information expected to be gained by expending a computational budget β on evaluating points. The point to be evaluated and the computational budget given an available budget B to use is therefore

$$\underset{\underline{x}, b \in \mathbb{R}^d \times [0,B]}{\arg\min} \quad \langle \Delta I(\underline{x}, b \mid \mathbf{X}) \rangle + \langle I_g(B-b) \rangle \tag{35}$$

This formulation is analagous to a reinforcement learning problem. $\mathbf{X}, B$ define the state, $\langle I_g(B) \rangle$ is the state value, $\underline{x}_p, b$ is the action and $\langle \Delta I(\underline{x}_p, b \mid \mathbf{X}) \rangle$ is the action value, which can be evaluated using the procedure defined in CITEhenngi.

The crucial missing element in this formulation is the state value function which encodes how much information is expected to be gained from a given computational budget, which must be reasonably accurate to provide useful results. In the case of problems where a lot of similar optimizations will be taking place it can be learned, using reinforcement learning techniques for continuous space. This method could be applied to a tariff setting problem where the consumer model takes account of external environmental factors. The objective will be broadly similar each day as the consumer requirements have not changed but the differing predicted weather conditions each day cause minor changes to the response. In this case the information gain from previous days can be used as a state value function. It may be feasible to construct a more general approach to determining the state value function using no or only a small number of parameters and improving this estimate online, investigation into this possibility would be a major theme of research.