Predicting Agent Trustworthiness for Large-Scale Power Demand Shifting

Charilaos Akasiadis[1,2] and Georgios Chalkiadakis[1]

[1] School of Electronic and Computer Engineering, Technical University of Crete, Kounoupidiana Campus, 73100 Chania, Crete, Greece
[2] Institute of Informatics and Telecommunications, N.C.S.R. "Demokritos", 15310 Aghia Paraskevi, Greece
{akasiadi,gehalk}@intelligence.tuc.gr

Abstract. A variety of multiagent systems methods has been proposed for forming cooperatives of interconnected agents representing electricity producers or consumers in the Smart Grid. One major problem that arises in this domain is assessing participating agents' uncertainty, and correctly predicting their future behaviour regarding power consumption shifting actions. In this paper we adopt two stochastic filtering techniques, a Gaussian Process Filter and a Histogram Filter, and use these to effectively monitor the trustworthiness of agent statements regarding their final shifting actions. We incorporate these within a directly applicable scheme for providing electricity demand management services. Experiments were conducted on real-world consumption datasets from Kissamos, a municipality of Crete. Our results confirm that these techniques provide tangible benefits regarding enhanced consumption reduction performance, and increased financial gains for the cooperative.

1 Introduction

Smart Grid-related research has received much attention in the last few years. Its general objective is to create a more secure, reliable and efficient electricity networks infrastructure, with energy produced mostly by green sources, production costs minimized, and affordable electricity made easily available to the public [7]. Due to the scale and complexity of electrical networks management, artificial intelligence (AI) and multiagent systems (MAS) solutions are in high demand in the emerging markets involving business entities providing services in the Smart Grid [3,14]. Many such entities have already adopted a business model that pulls together the resources and abilities of multiple economically-minded individuals. For instance, the emergence of *Virtual Power Plants* or *cooperatives* of small-to-medium size electricity producers or consumers—represented by autonomous, selfish agents—has been hailed as a means to create large, efficient, trustworthy providers of renewable energy production or electricity consumption reduction (peak-trimming) services [1,3,5,15,14].

In order for agent cooperatives to be functional, efficient, and profitable, they need to take business decisions regarding which members to include in consumption shifting coalitions. These decisions naturally depend on the abilities (e.g., electricity production or consumption reduction capacities) of individual agents. These abilities need to be either monitored by some central cooperative-managing agent, or need to be truthfully and accurately communicated to it. However, it is clear that in the large and dynamically changing scene of the power Grid, trust between selfish agents is not implied, and must be guaranteed. *Mechanism design* and related approaches attempt to build trust among cooperative members, via providing them with the incentives to truthfully and accurately report their intended future actions, along with their corresponding uncertainty regarding those actions [1,5,11,15]. Unfortunately, even if participating agents are perfectly truthful regarding their abilities and corresponding uncertainty, their reports and estimates can still be highly inaccurate. This can be due to, for example, communication problems, malfunctioning equipment, or prejudiced beliefs and private assumptions—e.g., a truthful reporting agent might be overly pessimistic or optimistic.

As a result, monitoring the performance of individuals and correctly predicting their future contributing potential is of utmost importance to a cooperative or an organization relying on the services of selfish, distributed, autonomous agents. To this end, several approaches try to explicitly estimate agent electricity consumption and production amounts, by incorporating prediction models that rely on agent geographical location and weather forecasts, or the processing of macroeconomic data [10,12]. Although their results are promising, such methods cannot immediately predict the actual behaviour of a specific agent, which might be motivated by private knowledge or business concerns, neither do they account for errors due to equipment malfunction. In contrast, this paper proposes the application of generic prediction methods, which are nevertheless able to adapt to a specific agent's behaviour regarding the promised and final consumption shifting actions. In more detail, the techniques are incorporated within a recent, simple yet effective mechanism that promotes the formation of agent cooperatives for power consumption shifting (that of [1]).

Specifically, we propose the use of *stochastic filtering methods* to keep track of the parameters that best describe agent behaviour, and effectively estimate actual future agent performance. These techniques are able to not only fit the dynamics of the processes governing agent performance, but can also imbibe the potential errors of electricity metering or information transmission devices. In particular, we adopt the *Histogram Filter (HF)* [16] and the *Gaussian Process* filter (GP) [13] to predict the future actual actions of agents participating in cooperatives offering electricity demand management services.

These two methods are very generic, and have wide areas of application. Their employment in the power consumption shifting domain ensures that member agents can be ranked by the cooperative according to their perceived consumption shifting capacities; and thus untruthful or inaccurate agent statements regarding their capacity and corresponding uncertainty will not be able to jeopardize the stability and effectiveness of the overall mechanism governing the cooperative business decisions (e.g., which agents to select for consumption shifting at a given point in time). This is key for the economic viability of any such cooperative.

To the best of our knowledge, this is the first work to use stochastic filtering and regression methods for assessing the performance of autonomous, economically-minded agents participating in Smart Grid cooperatives or other such entities. Our simulations demonstrate that employing any of the two filters tested leads to improved performance in the consumption shifting domain, when compared to that of a state-of-the-art, "baseline" mechanism that makes no use of such performance monitoring tools. Specifically, when using these enhancements, the cooperative achieves a higher overall electricity consumption reduction; and enjoys financial rewards that are higher than those generated by the baseline algorithm.

Both methods appear to be able to provide reasonable predictions regarding the actual performance of individual agents, given the agents' stated intended actions and related uncertainty. As such, the efficiency of these filtering methods is not restricted in the demand management and peak-trimming domains, but they can arguably be readily employed to monitor the trustworthiness of electricity producers' statements regarding their intended actions. In a nutshell, our results indicate that *both* filtering techniques examined in this paper are strong candidates for monitoring the trustworthiness of selfish agents in the Smart Grid. We thus believe they deserve to be further evaluated in this direction, since they can bring tangible benefits to business entities operating in this domain.

The rest of the paper is structured as follows. Section 2 describes past related work. Section 3 presents the consumption shifting problem, and explains how our filtering methods fit into that perspective. Section 4 describes the adopted filtering methods and their application to our problem in detail. Section 5 presents our simulations. Finally, Section 6 concludes this paper.

2 Related Work

Here we briefly review some work on filtering methods to predict future events based on past occurances, and techniques promoting trust. To begin, there exists much work on self-adapting systems and corresponding attempts to tackle related uncertainty. A detailed review of the types of uncertainty influencing the operation of a self-adaptive system, complete with techniques for uncertainty representation, can be found in [9].

One concrete example of a self-adaptive system is provided by the "RESIST" framework [6]. RESIST is a situated software system that monitors various information sources and reconfigures itself proactively. In that work, a *Hidden Markov Model (HMM)* is trained and solved. Contextual parameters are also used in order to make the system adaptive to environmental changes. However, the use of HMMs is not appropriate for our work here, as we are not trying to estimate each agent's exact state (although that could be done in a higher level analysis), but we instead wish to predict agent future performance based on past actions. A self-adaptive system that is closer to our requirements is the FeatUre-oriented Self-adaptatION framework [8]. The system operates using interconnected features, whose importance is constantly monitored and reconfigured accordingly.

Now, GPs have also been used recently by [4] to forecast electricity demand, and the predictions are tested in the electricity market simulation of PowerTAC. Drawing our attention to the energy domain, recent work has proposed the use of *Continuously Ranked Probability Score (CRPS)* scoring rule for evaluating power production or consumption predictions of agents participating in cooperatives operating in the Smart Grid [1,15]. CRPS provides a scoring function for evaluating the accuracy of a forecast, given its actual occurrence. When agent-stated forecasts are off the occurrences, contributors are "fined" proportionally to their CRPS score. While this *mechanism design* technique provides the agents with strong incentives to stay truthful (and, indeed, provides theoretical guarantees for statement truthfulness), it does not guarantee agent statements accuracy, as explained in Section 1.

In our work here, we adopt the general framework of [1] for demand curve trimming cooperatives, and then enhance it by proposing two distinct filtering techniques that can be used for system adaptation, via predicting the power consumption shifting efforts of participating agents. When an agent shifts some load, its actions are monitored and a corresponding model is induced. Then, instead of simply taking individual forecasts into account, the learned model is used to better predict agent and cooperative action quality, and improve monetary benefits and general performance.

3 Electricity Demand Shifting

In this section, we describe the approach of [1] for *collective power consumption shifting* provided to the Grid by electricity consumer cooperatives. Consumption shifting activities—from peak electricity demand intervals during a day, to low demand ones—can be key to bringing down electricity production costs and avoiding shortages. The cooperative (e.g., via some "central" agent) takes business decisions specifying the *coalitions* of cooperative members to provide shifting services at a particular time. For shifting coalitions to be effective, however, selected coalitions have to be as effective as possible. Thus, member agents are required to state to the cooperative *(a)* their estimated reduction capabilities, and *(b)* their respective confidence on the accuracy of that estimate. Member statements' truthfulness and accuracy is motivated via the employment of the *CRPS* scoring rule.

Specifically, [1] assumes the existence of a demand *threshold* τ, over which production is more expensive. Then, every *time interval* during which demand exceeds τ can be considered as *peak*, with the excessive ($> \tau$) load being termed *peak load*. Electricity price p_h for peak intervals is higher than the p_l price charged for consumption in non-peak ones. Shifting *peak load* to non-peak intervals can lead to the flattening-out of the demand curve (Figure 1). Naturally, there are constraints that need to be met for the consumption shifting operation to be possible, e.g. increasing at non-peak intervals up to a safety limit, and others, which are provided in [1].

In more detail, let $A = \{a_1, a_2, ..., a_n\}$ denote a set of agents that constitute the cooperative. In order for the cooperative to place a bid, each contributing agent i must state its reduction capacity, $\hat{r}_i^{t_h}$, at t_h high-consumption (peak) intervals, and corresponding shifting costs $c_i^{t_h \to t_l}$ for moving consumption to non-peak, t_l, intervals. Agents are also required to state their uncertainty over their *expected relative error* regarding their reduction capacity, in a form of a normal distribution $\mathcal{N}(\mu_i, \hat{\sigma}_i^2)$. Next, the cooperative assigns a *conservative* estimate of each agent's performance:

$$\tilde{r}_i^{t_h} = \hat{r}_i^{t_h} - \hat{\sigma}_i \hat{r}_i^{t_h} \tag{1}$$

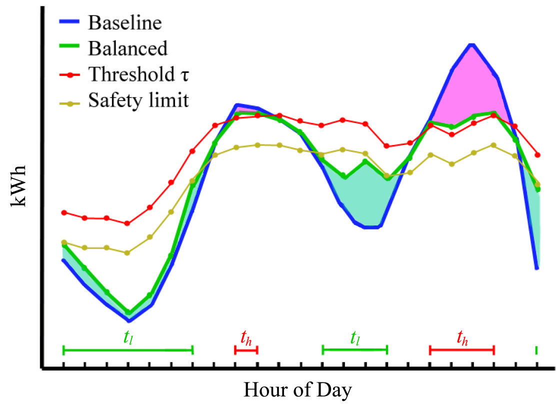


Fig. 1: Example shifting procedure. The purple-shaded load during t_h must be shifted to non peak intervals t_l.

Then, the agent's *reservation price*, $\hat{p}_i$ is calculated, that is, the highest price i is willing to pay for shifting consumption from t_h to t_l without suffering a monetary loss:

$$\hat{p}_i = p_h - c_i^{t_h \to t_l} \tag{2}$$

And so is the *contribution potential*, the product of the expected reduction and reservation price, $\tilde{r}_i^{t_h} \hat{p}_i$. The agents are then ranked by descending contribution potential, and *shifting coalitions* are formed by the number of top agents that meet the required constraints. Selected coalition agents are awarded low, *variable prices* for shifting to t_l, determined by a *group price* $p_g \leq p_l$ which is guaranteed by the Grid, and by monetary *transfers* that make it worthwhile for everyone selected to participate in the shift. As argued by the authors in [1], the scheme can be readily used by cooperatives offering electricity demand management services, as it is simple and requires no legislature changes whatsoever (in contrast to other approaches, such as *real-time pricing* [3]).

It is obvious by the discussion above that agent statements greatly affect co-operative decisions, and, if inaccurate, endanger the scheme's stability and effectiveness. This is why a *trusted index* r^*_{i,t_h} is needed—one not stated explicitly by i, but nevertheless revealing the distribution best describing future agent actions. This index can then be used instead of $\tilde{r}_i^{t_h}$ to calculate a more accurate *contribution potential* for i.

4 Method Description

To begin, we denote the *actual* amount of load reduced by $r_i^{t_h}$. In general, it can be assumed to be provided by a transformation of the stated $\hat{r}_i^{t_h}$ amount:

$$r_i^{t_h} = \alpha_i \hat{r}_i^{t_h} \tag{3}$$

with the (observed) "accuracy factor" α_i corresponding to a random variable characterizing the accuracy of the statement regarding the promised shifting amount. This variable follows some unknown probability distribution.

The objective of our work in this section is to build models of the agents' performance *by approximating the distributions that* α_i*s follow*. We can then sample such a distribution to obtain *a better* α_i *estimate*, denoted $\tilde{\alpha}_i$. We then use this estimate to obtain our trusted index r^*_{i,t_h} to replace $\tilde{r}_i^{t_h}$, as follows:

$$r^*_{i,t_h} = \tilde{\alpha}_i \hat{r}_i^{t_h} \tag{4}$$

As a result, more accurate predictions about individual agent and cooperative shifting abilities can be obtained.

Given all underlying uncertainty, an individual agent's final behaviour most likely corresponds to a complex, non-linear function of its past behaviour. Therefore, we chose to test two filtering approaches that are expected to fit such a function well: *(a)* a filter utilizing a *Gaussian Process (GP)* [13], and *(b)* the *Histogram Filter (HF)* [16], a non-parametric filtering technique.[3]

The GP uses historically observed pairs of agent statements and final actions values to fit normal distributions for the underlying random variables for each statement value. The HF ignores user stated uncertainty over the performance and takes into account only past observations for its predictions. On the other hand, GP assumes that final actions are functions of user stated uncertainty and performs accordingly. These methods are generic, with very wide areas of application. Moreover, not only are they able to fit the dynamics of the processes governing agent performance, but can also imbibe the potential errors of electricity metering or information transmission devices. Here we employ them to enhance the performance of our energy consumption shifting mechanism. By so doing, more accurate agent rankings can be obtained, and inaccurate statements will not be able to jeopardize the stability and effectiveness of the mechanism. We now describe the stochastic filtering methods and their application to our setting in detail.

4.1 Histogram filter

Histogram filters decompose a continuous state space to a finite set of areas or bins:

$$dom(X) = \mathbf{x}_1 \cup \mathbf{x}_2 \cup .. \cup \mathbf{x}_K$$

The HF uses a histogram to map a probability p_k to each of the bins $\mathbf{x}_k$. The value of each p_k depends on the frequency of the observations in the range of bin k.

With this approach, agent forecasts $\hat{\sigma}$ are completely ignored and only past observations of α_i are taken into account. Every time an agent participates in a consumption shifting coalition, its actions are monitored and stored. A histogram is calculated over the set of available observations. Then, according to each bin's height, a colored roulette

[3] All methods require an adequate amount of historical data collected, in order to form an initial model that can be sampled. To this purpose, the *conservative estimate* can be employed at first, and then get replaced by one of the proposed methods once enough data is available.

wheel is constructed that can be sampled to obtain the most probable ranges of α_i, i.e. the more frequent values appear in a bin the more probable it's range is selected. So, we sample the corresponding roulette wheel, and come up with a specific bin $\mathbf{x}_k$. The final $\tilde{\alpha}_i$ estimate is another sample from a uniform distribution normalized to have range equal to that of the bin obtained —i.e., for $\mathbf{x}_k$ with lower limit $\mathbf{x}_k^-$, and upper $\mathbf{x}_k^+$ we have:

$$\tilde{\alpha}_i \sim \mathcal{U}(\mathbf{x}_k^-, \mathbf{x}_k^+) \tag{5}$$

The advantage of HF is that it requires no prior knowledge about the form of the distribution that α_i follows, and adapts effectively to all kinds of non-linearities [16]. On the other hand, it needs a number of measurements before it starts working accurately and performance might be unacceptable in initial stages with no actual measurements. Another drawback is that if the distribution changes over time, the length of a history window must be re-set, in order to get rid of expired measurements interference. Also, it does not take into account the error variance $\hat{\sigma}$ that is, the agent-stated confidence.

4.2 Gaussian Process filter

Here we describe the application of Gaussian processes for probabilistic regression, to construct a filter that is going to be used for monitoring and prediction purposes.[4] For a set of training samples, $\mathcal{D} = \{(\mathbf{x}_j, y_j), j = 1, ..., n\}$ ($\mathbf{x}_j$ inputs and y_j noisy outputs) we need to predict the distribution of the noisy output at some test locations $\mathbf{x}_*$. We assume the following model:

$$y_j = f(\mathbf{x}_j) + \epsilon_j, \text{ where } \epsilon_j \sim \mathcal{N}(0, \sigma^2_{noise})$$

with σ^2_{noise} the variance noise.

GP regression is a Bayesian approach that assumes a priori that function values follow: $p(\mathbf{f}|\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) = \mathcal{N}(\mathbf{0}, K)$ where $\mathbf{f} = [f_1, f_2, ..., f_n]^T$ is the vector of latent function values, $f_j = f(\mathbf{x}_j)$ and K is the covariance matrix that is computed by a covariance function $K_{jk} = k(\mathbf{x}_j, \mathbf{x}_k)$. The joint GP prior and the independent likelihood are both Gaussian with mean and variance as follows:

$$GP_\mu(\mathbf{x}_*, \mathcal{D}) = K_{*,f}(K_{f,f} + \sigma^2_{noise} I)^{-1}\mathbf{y} \tag{6a}$$

$$GP_\sigma(\mathbf{x}_*, \mathcal{D}) = K_{*,*} - K_{*,f}(K_{f,f} + \sigma^2_{noise} I)^{-1}K_{f,*} \tag{6b}$$

GPs also require value assignments to the vector $\theta = [\mathbf{W}\ \sigma_f\ \sigma_{noise}]$ that contains the hyperparameters, with $\mathbf{W}$ holding the distance measure of each input in its diagonal, σ_f being the variance of the input and σ_{noise} the variance of the process noise. We can find the optimal values for θ by maximizing the log likelihood:

$$\theta_{max} = \arg\max_\theta \{\log(p(\mathbf{y}|\mathbf{X}, \theta))\} \tag{7}$$

[4] Note that in a previous short paper [2], we also explored the use of *GP-UKF*, an *unscented Kalman filter combined with Gaussian process regression*. However, in order to exploit the full power of that technique, in reality one needs to have access to a realistic model of the stochastic dependencies among the past $\hat{\sigma}_i$ agent statements. Without such a model, one cannot observe significant differences between using *GP-UKF* and *GP* alone.

In our application setting, $\mathbf{x}_*$ is the reported agent confidence $\hat{\sigma}_i$ for the interval under examination, and $\mathcal{D} = \{(\hat{\sigma}_j, \alpha_j), j = 1, ..., n\}$ the pairs of past agent values. Finally, the estimate of future agent behavior can be calculated by:

$$\tilde{\alpha}_i = GP_\mu(\hat{\sigma}_i, \mathcal{D}) + u_\tau \tag{8}$$

with noise u_τ following $\mathcal{N}(0, GP_\sigma(\hat{\sigma}_i, \mathcal{D}))$. Note that this approach takes into account agent statements regarding their uncertainty, $\hat{\sigma}_i$, thus incorporates more prior information than the *HF* approach.

5 Experimental Evaluation

In this section we conduct extensive experimental simulations of our mechanism on real consumption patterns. We first present our real world dataset, and explain how we augmented it to also contain information which is still unavailable and cannot be obtained, but is nevertheless required in order to illustrate the methods' performance. Next, we compare the results of the collective consumption shifting procedure, when employing the proposed filtering techniques for monitoring and prediction.

5.1 The simulations dataset

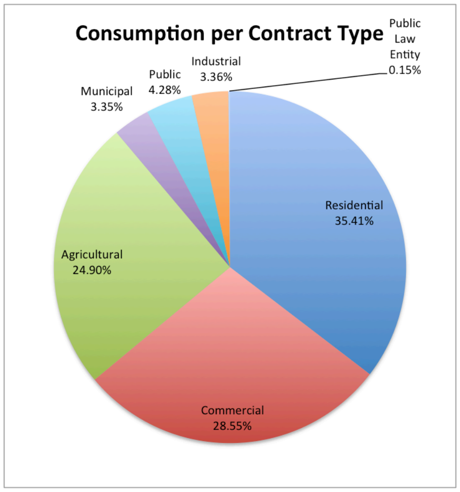
To experimentally evaluate our proposed methods we created a simulations dataset based on real electricity consumption data,[5] from Kissamos, a municipality at the greek island of Crete. The dataset contains consumption values for the year 2012, as well as contract types and geographical locations, and a summary of its contents appears in Table 1.

Table 1: Kissamos 2012: Number of consumers and corresponding individual average consumption for each consumption contract type.

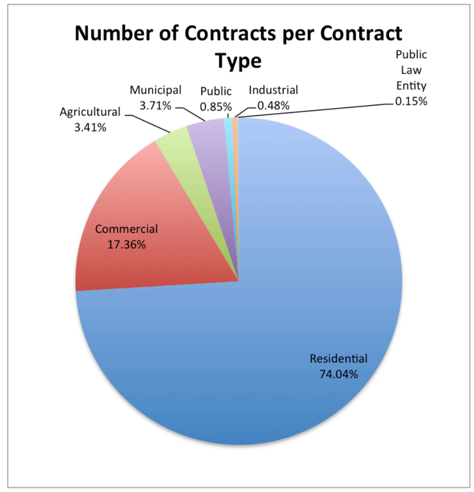
Type	Count	Avg Daily Individ. Cons. (kWh)	Avg Daily/ Type (kWh)	% Of total
Residential	5889	7.2944	42956.7216	35.4100
Commercial	1381	25.0804	34636.0324	28.5508
Agricultural	271	111.4737	30209.3727	24.9018
Municipal	295	13.7762	4063.979	3.3499
Public	68	76.3616	5192.5888	4.2803
Industrial	38	107.2575	4075.785	3.3597
Public Law	12	14.9211	179.0532	0.1475
Total	7954	-	121313.532	100

Then, in Figure 2, we present a visualization of electricity usage with respect to the contracts found in our dataset. Specifically, Figure 2a reports the percentage of total electricity usage per contract type, and Figure 2b shows the partition of the consumers with respect to their contract types.

[5] The dataset was provided by the Hellenic-Public Power Company (PPC, `www.dei.gr`).



(a) Electricity by end-use



(b) Contract types

Fig. 2: Kissamos 2012: Electricity contracts and usage

Agent consumption patterns and corresponding electricity bills We employed our dataset in order to retrieve the $\bar{D}^d_{i,x}$ average daily consumption of each consumer i of type x, calculated over the entire 2012 year—and in order to estimate its $D^t_{i,x}$ consumption per each t time-interval. In Table 2 we present the average electricity bill amount paid by a participant of a particular type for consumption over 100 days. The highest

Table 2: Average electricity bills for each contract type, for a 100 days period.

Type	Bill cost (€)
Residential	61.66
Commercial	213.20
Agricultural	933.17
Municipal	113.30
Public	645.36
Industrial	903.74
Public Law Entity	120.13

bill amount is paid by agricultural customers—whom, however, we do not include in the simulated shifting cooperative, due to their practically non-existing load-shifting ability (mainly pumps which constitute a non-shiftable load). The second highest amount is charged to industrial customers, and the lowest to the residential ones.

Key parameters of the shifting scheme We assume a threshold τ for our model, to the 93% of the maximum demand value among all time intervals, and is fixed for all of them—though, it could also be variable across time intervals. The safety limit is 97.5% of τ. The p_l, p_h price levels correspond to the day/night prices provided by PPC, the greek public power company, i.e. $p_l = 0.0785$ €/kWh and $p_h = 0.094$ €/kWh. The p_g group price rate ranges from $p_g^{max} = 0.05625$ to $p_g^{min} = 0.0214$ €/ kWh, depending on the reduction size q:

$$p_g(q) = \frac{0.0214 - 0.05625}{Q^{t_h}_{max} - q^{t_h}_{min}}(q - q^{t_h}_{min}) + 0.05625 \tag{9}$$

with q ranging from $Q^{t_h}_{max}$, that is the amount of load above τ, to some minimum $q^{t_h}_{min}$, which we set to $0.3Q^{t_h}_{max}$.

Agent statements and final shifting actions We need a model to describe how *(i)* the agent statements on their uncertainty regarding shifting capacities at t_h occur; and *(ii)* their actual, final shifting actions occur. To this end, we define two main agent classes; the first one, denoted as *BB*, describes the realistic case where agents are mainly *confident* about their statements, and also have a high probability to deliver what they promised. In the *BB* class, the stated error standard deviation $\hat{\sigma}$ and the observed α_i accuracy factor follow two *Beta* distributions, $\mathcal{B}(1,5)$ and $\mathcal{B}(4,2)$ respectively, which are depicted in Fig. 3. We choose *Beta* distributions primarily because they very good at representing and updating probabilistic beliefs regarding potential behaviours, as also manifested by their widespread use for simulating behaviours and uncertainties related to real world scenarios (see, e.g., [11]). The use of the particular two aforementioned *Betas* corresponds to error statements having a low mean of $= \frac{1}{6}$ (i.e., to agents stating high certainty), whereas the final observed accuracy factor is closer to 1 (implying increased actual performance effectiveness). The second agent class is the *uncertain predictors*, *UP*, where consumers might or might not follow stated forecasts, so α_i and $\hat{\sigma}_i$ both follow the same normal distribution with a slightly raised variance—i.e., the $\mathcal{N}(0.5, 0.15)$ distribution depicted in Fig. 3. Table 3 summarizes the parameters and distributions associated with *BB* and *UP* agents behaviour.

We believe that *BB* and *UP* sufficiently capture two realistic, highly plausible scenarios of agent behaviour. Of course, the underlying models of agents behaviour could follow any other distribution as well.[6] About 50% of the participants in our setting belong to the *BB* class, with the rest being *UP* agents (since agents are assigned to a specific class with 50% probability).

5.2 Evaluation of monitoring techniques

To test-evaluate the performance of our two monitoring techniques, we first applied them on a single agent of the *BB* class, trained over 1000 past value couples that were

[6] For instance, in [2], the classes of *accurate* and *inaccurate* predictors were also introduced. However, these behavioural classes are less interesting since their behaviour is easier to predict; and, moreover, they are likely in practice, since *(i)* in realistic settings, errors do occur, while *(ii)* if predictions are highly inaccurate, the agents would most probably be acting upon them already, to avoid penalties.

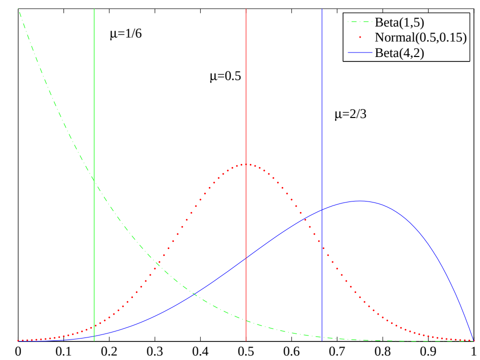


Fig. 3: Probability density functions of agent behaviors

Table 3: Behavioural classes of the scheme participants.

Class	Parameter	Distribution
BB	$\hat{\sigma}$	$\mathcal{B}(1,5)$
	α_i	$\mathcal{B}(4,2)$
UP	$\hat{\sigma}$	$\mathcal{N}(0.5,0.15)$
	α_i	$\mathcal{N}(0.5,0.15)$

generated using the same distributions as in the simulation. Figures 4 and 5 show the outcomes for the *HF* and *GP*, respectively. We can observe that although the *HF* does not take into account agent $\hat{\sigma}_i$ statements, it fits well to the real underlying distribution.[7]

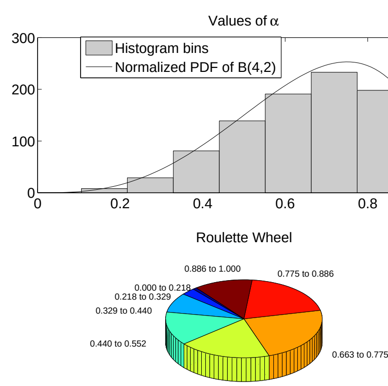
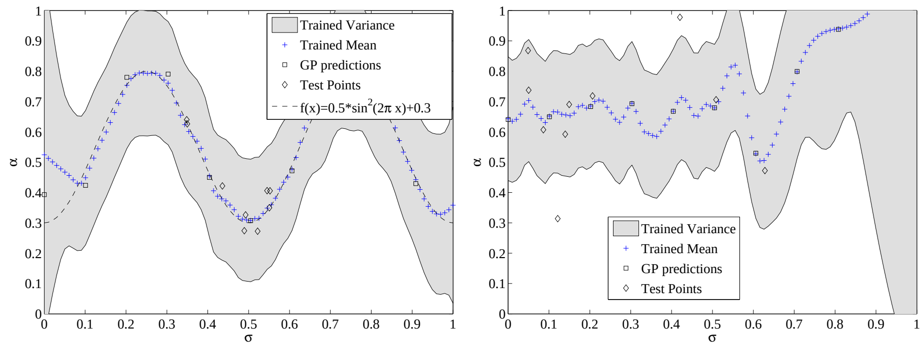


Fig. 4: Agent population vs. average agent shifting cost

[7] In our setting, the $\hat{\sigma}_i, \alpha_i$ values are statistically independent.



(a) Illustrating the *GP* capabilities: fitting an arbitrary non-linear function

(b) Fitting of the behavioural function of the Kissamos dataset

Fig. 5: *GP* fits. Illustration of *GP* training for two non-linear cases

The *GP*, on the contrary, does take agent commitments into account. In Figure 5a we can see the *GP* fit an arbitrary non-linear scenario, proving that non-linearities can be handled. Fig. 5b presents the trained variances and means for a typical agent participating in the cooperative shifting process for the setting where $\hat{\sigma}$ and α_i follow $\mathcal{B}(1,5)$ and $\mathcal{B}(4,2)$ respectively. In this case, there is no function of a specific known form for the Gaussian Process to approximate, as the points are both random variables following different distributions. Despite that, *GP* has converged to some relationship between input and output values. We can infer that this estimated complex function is meaningful, by the fact that most "test points" fall within the shaded area representing the *GP* output variance; the "test points" plotted in this case are random $\hat{\sigma}_i, \alpha_i$ values, sampled by the $\mathcal{B}(1,5)$ and $\mathcal{B}(4,2)$ respectively. Thus, *GP* is apparently able to produce meaningful predictions, even when the relationships between variables are governed by some highly complex function. We can observe that the main mass of the observations is in the upper left quadrant ($E[\hat{\sigma}_i] \approx 0.16$ and $E[\alpha_i] \approx 0.66$) and that the trained means are close to that area. Note that because $\mathcal{B}(1,5)$ gives very low to zero probability for $\hat{\sigma}$ values between 0.7 and 1, the number of corresponding training points is very low, so uncertainty in that region is very high. This is not an issue though, as *GP* is not likely to be asked to provide predictions in that range.

Next, we employ these techniques in the cooperative consumption shifting simulation scenario [1]. For training, we used 100 $(\hat{\sigma}_i, \alpha_i)$ couples generated from the same distributions; these were considered to be the historical values. We compare the gains produced when using the two stochastic filtering techniques to those accrued when using "conservative" estimates for performance prediction. The *conservative*, *HF*, and *GP* estimates are calculated as described in Section 3 and 4.

The numerical results from a 100 days simulation are presented in Table 4. All three estimators generate *expected* gain and reduction values that are pretty close to each other, though estimates differ among the two coalition formation cases. However, when one examines the *final outcomes* (actual cooperative gain), it becomes obvious that the *HF* and *GP* filtering methods perform much better than the *conservative* one. The *HF* and *GP* performance is comparable, with *GP* generating slightly better actual gains and better precision for individual forecasts, as indicated by its lower "unallocated" cooperative surplus.

Table 4: Average results from an 100 days simulation.

		Conserv.	*HF*	*GP*
Expected Coop. Gain (€/day)	μ	55.79	54.87	54.61
	σ	20.94	22.36	22.53
Actual Coop. Gain (€/day)	μ	35.38	47.04	51.10
	σ	14.16	20.04	21.46
Coop. "Surplus" (€/day)	μ	0.19	0.56	0.12
	σ	0.18	2.31	0.13
Expected Reduction (kWh/day)	μ	1239.135	1219.125	1211.522
	σ	466.138	501.049	504.137
Final Reduction (kWh/day)	μ	1038.433	1155.718	1204.058
	σ	395.930	483.849	504.177
Accuracy (%)	μ	83.80	94.79	99.38
Peak (Demand $\geq \tau$) Trimmed (%)	μ	74.70	82.88	85.99
	σ	19.95	24.43	25.89
Avg. Reducing Coalition Size	μ	207.14	272.99	284.59
	σ	83.34	112.97	120.59

Returning to the numerical results of Table 4, we observe that the final performance of the shifting coalitions reaches prediction accuracy values that exceed 94% for the two filtering techniques. Moreover, *GP* is able to achieve 99.38% peak trimming performance.

Summarizing, the evaluation demonstrates that *GP* is potentially the most effective of the prediction techniques examined. This is best illustrated by its better performance with respect to prediction accuracy, and its nearly perfect effectiveness in terms of peak load trimmed. Recall that for the *HF*, agent forecasts are not taken into account, so potentially important information is ignored. Intuitively, *GP* can effectively learn and adapt to the underlying model that relates agent forecasts and final actions, thus enabling the cooperative to choose reducing coalitions that often deliver what they promised. *HF*, however, exhibits a strong performance also. Thus, in a nutshell, our results indicate that both filtering techniques examined are strong candidates for monitoring the accuracy of selfish agent statements in Smart Grid consumer cooperatives.

Finally, we also report that when using *GP*, the most active participants achieve higher gains (approximately 0.04 €/*kWh* shifted), as compared to those achieved when using the conservative technique for estimating future agent performance. In addition, participants with 15 participations or more within a month, receive bills that are reduced by 2.4%.

6 Conclusions and Future Work

In this paper we presented two different methods for monitoring and predicting agent actions in a power consumption shifting scheme, a Histogram Filter and a Gaussian Processes based filter to recognize possible underlying relationships between agent forecasts and final actions in a cooperative electricity consumption shifting scheme. The methods outperform the formerly used method *wrt.* prediction accuracy —and financial gains generated. Our techniques are generic and can be integrated within different types of systems (e.g., they can be used for monitoring the accuracy of electricity production statements). This is the first time these methods are applied in this domain; and the potential value of this work to any real-world enterprise operating in the Smart Grid can be very high.

Future plans include testing different classes of agent behavior, that might change over time. We also intend to devise methods for distributed, combined performance monitoring, so that predictions are effectively cross-validated. Furthermore, we will investigate the adoption of cryptocurrency approaches, which can allow the creation of job-markets for prosumers. Finally, we aim to further test the proposed methods by using smart meters in real homes and large enterprises.

References

1. C. Akasiadis and G. Chalkiadakis. Agent Cooperatives for Effective Power Consumption Shifting. In *In Proc. of the 27th AAAI Conference on Artificial Intelligence, (AAAI-2013)*, pages 1263–1269, 2013.
2. C. Akasiadis and G. Chalkiadakis. Stochastic filtering methods for predicting agent performance in the smart grid. In *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)*, pages 1205–1206, 2014.
3. P. Asmus. Microgrids, virtual power plants and our distributed energy future. *The Electricity Journal*, 23(10):72 – 82, 2010.
4. M. Blum and M. Riedmiller. Electricity Demand Forecasting using Gaussian Processes. In *Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.
5. G. Chalkiadakis, V. Robu, R. Kota, A. Rogers, and N. R. Jennings. Cooperatives of distributed energy resources for efficient virtual power plants. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2011)*, pages 787–794, May 2011.
6. D. Cooray, R. Malek, S.and Roshandel, and D. Kilgore. RESISTing reliability degradation through proactive reconfiguration. In *Proc. of the IEEE/ACM International Conference on Automated Software Engineering*, ASE '10, 2010.
7. B. Davito, H. Tai, and R. Uhlaner. The smart grid and the promise of demand-side management. *McKinsey*, 2010.
8. A. Elkhodary, N. Esfahani, and S. Malek. FUSION: a Framework for Engineering Self-Tuning Self-Adaptive Software Systems. In *Proc. of the 18th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, FSE '10, pages 7–16, New York, NY, USA, 2010. ACM.
9. N. Esfahani and S. Malek. Uncertainty in Self-Adaptive Software Systems. In *Software Engineering for Self-Adaptive Systems II*, volume 7475 of *LNCS*, pages 214–238. 2013.

10. J. B. Jorgensen and F. Joutz. Modelling and Forecasting Residential Electricity Consumption in the U.S. Mountain Region. Technical Report 2012-003, The George Washington University, Dept. of Economics, January 2012.
11. R. Kota, G. Chalkiadakis, V. Robu, A. Rogers, and N. R. Jennings. Cooperatives for demand side management. In *The Seventh Conference on Prestigious Applications of Intelligent Systems (PAIS @ ECAI)*, pages 969–974, August 2012.
12. A. A. Panagopoulos, G. Chalkiadakis, and E. Koutroulis. Predicting the power output of distributed renewable energy resources within a broad geographical region. In *ECAI-2012/PAIS-2012: 20th European Conference on Artificial Intelligence, Prestigious Applications of Intelligent Systems Track*, August 2012.
13. J. Quiñonero Candela and C. E. Rasmussen. A Unifying View of Sparse Approximate Gaussian Process Regression. *J. Mach. Learn. Res.*, 6:1939–1959, December 2005.
14. S. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings. Putting the "Smarts" into the Smart Grid: A Grand Challenge for Artificial Intelligence. *Communications of the ACM*, 2012.
15. V. Robu, R. Kota, G. Chalkiadakis, A. Rogers, and N. R. Jennings. Cooperative Virtual Power Plant Formation using Scoring Rules. In *Proc. of the 26th Conference on Artificial Intelligence (AAAI-12)*, pages 370–376, August 2012.
16. S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, 2005.