

Value of Thermostatic Loads in Future Low-Carbon Great Britain System

Vincenzo Trovato

Fei Teng

Goran Strbac

Department of Electrical and Electronic Engineering
Imperial College London
London, United Kingdom

Abstract—This paper quantifies the value of a large population of heterogeneous thermostatically controlled loads (TCLs). The TCL dynamics are regulated by means of an advanced demand side response model (DSRM). It optimally determines the flexible energy/power consumption and simultaneously allocates multiple ancillary services. This model explicitly incorporates the control of dynamics of the TCL recovery pattern after the provision of the selected services. The proposed framework is integrated in a mixed integer linear programming formulation for a multi-stage stochastic unit commitment. The scheduling routine considers inertia-dependent frequency response requirements to deal with the drastic reduction of system inertia under future low-carbon scenarios. Case studies focus on the system operation cost and CO₂ emissions reductions for individual TCLs for a) different future network scenarios, b) different frequency requirements, c) changes of TCL parameters (e.g. coefficient of performance, thermal insulation etc.).

Index Terms— Load management, Stochastic unit commitment, Power system economics, Frequency response, Renewable energy sources.

NOMENCLATURE

A. Constants

n	Node of the scenario tree.
$a(n)$	Parent node.
$\Delta\tau(n)$	Time interval corresponding to node n [h].
$\pi(n)$	Probability of reaching node n
G	Set of thermal generators.
C_{LL}	Value of lost load [£/MWh].
P_g^{max}	Maximum production of thermal unit g [MW].
$R_g^{P,max}$	Maximum primary response capability of thermal unit g [MW].
$R_g^{S,max}$	Maximum secondary response capability of thermal unit g [MW].
f_g^P, f_g^S	The proportion of the spinning headroom, which can contribute to primary and secondary response.
r_g^{max}	Maximum generator ramp rate [MW/min]
D	Load damping rate [%MW/Hz]
t_P, t_R	Primary response [s] and reserve [s] delivery time

$P_D(n)$ Total static demand at node n [MW].

B. Variables

$P_{LS}(n)$	Load shed at node n [MW].
$P_g(n)$	Power output of thermal unit g at node n [MW].
$P_T(n)$	TCLs power consumption at node n [MW].
$P_T^P(n)$	Primary response from TCLs at node n [MW].
$P_T^S(n)$	Secondary response from TCLs at node n [MW].
$S_T(n)$	TCL energy level at node n [MW].

C. Linear expression

$C_g(n)$	Operating cost of thermal unit g at node n .
$P_T^{ar}(n)$	Additional reserve requirement due to TCLs recovery at node n (MW).
$N_g^{up}(n)$	Operation status (0/1 for Offline/Online) of thermal unit g at node n .

I. INTRODUCTION

Demand Side Response (DSR) represents a valuable option for the cost-effective integration of large shares of Renewable Energy Sources (RES). The intrinsic flexibility associated to DSR technologies may facilitate the system operator to deal with the uncertainty and variability of RES [1]. In addition, DSR may facilitate the fulfillment of frequency response requirements, which are expected to grow due the current inability of RES to provide both inertial and governor response [2]. In effect, if conventional generators remain the only source for system control and flexibility, the cost of integrating RES will rise, eventually limiting their absorption [3]. An interesting subset of appliances for DSR encompasses thermostatically controlled loads (TCLs) e.g. refrigerators. Individual customers or entities like aggregators or system operator may employ the intrinsic flexibility of these devices and benefit from their provision of a portfolio of frequency response services or realizing energy arbitrage. The potential value of thermostatic loads in future low-carbon electricity systems has been demonstrated in previous works [2, 4, 5]. The studies in [2, 4] quantified the cost savings for TCLs supplying primary frequency response only, while [5] highlighted the

value of flexible energy consumption in response to wholesale energy price differences. Limitations in the adopted control strategies could not show the benefits of simultaneously optimizing the allocation of multiple services, as demonstrated in [6] for typical storage units. For same reasons, the effect on the TCL value of the energy payback, associated to the provision of response services, could not be quantified in [4] or relied on several approximations in [2]. In [7] the payback energy absorption was modelled as function of the previous power reduction through a generic constant parameter; however, it ignored the physical constraints of any actual demand-side technology.

Recently we developed a novel demand side response model (DSRM) for heterogeneous clusters of TCLs that solves both the challenges identified [8]. The DSRM optimally determines the flexible TCL energy/power consumption and simultaneously allocates multiple response services supplied by the controlled devices. The dynamics of the load recovery pattern are accurately included. The model is integrated in an advanced Stochastic Unit Commitment (SUC) based on the framework developed in [3]. The optimal solution minimizes the network operational cost also balancing the actual cost of committing extra generation reserve to assist the load recovery against the benefit of the TCL support. This SUC model also recognizes the impact of reduced inertia on the frequency response requirements.

This paper focuses on the device-level perspective. We apply the methodology proposed in [8] to different future GB system scenarios; in each of these, the installed generation mix is varied, reflecting potential future system developments. The case studies quantify the annual value of individual TCLs in term of system operation cost savings and CO₂ emissions reductions. In addition, a sensitivity analysis investigates the impact of different frequency response requirements (i.e. the maximum delivery time of primary response and the maximum rate of change of frequency), in line with National Grid proposals [9]. We assess the value of TCLs for different devices penetrations. The insulation capability and the coefficient of performance of TCLs are also varied to provide a guidance for devices manufactures.

II. AGGREGATE MODEL AND CONTROL OF TCLS

Recently it has been demonstrated in [10] that the aggregate flexibility of $N \gg 1$ heterogeneous TCLs is summarized by the (1-4).

$$\frac{dS(t)}{dt} = -\frac{1}{\hat{\tau}}S(t) + P(t) \quad (1)$$

$$P(t) = \hat{P}_0 \Pi(t) + O(N^{-1/2}) \quad (2)$$

$$\hat{P}_{min} = \max_a P_{min}^a \leq P(t) \leq \hat{P}_{max} = \min_a P_{max}^a \quad (3)$$

$$\hat{S}_{min} = \max_a S_{min}^a \leq S(t) \leq \hat{S}_{max} = \min_a S_{max}^a \quad (4)$$

Equation (1) describes the TCL cluster a single energy storage model with an evaporative term and it shows the aggregate energy ($S(t)$ [MWh]) and power ($P(t)$ [MW]) coupling. $\hat{\tau}$ [h] is the time constant for the population model

and it was selected by choosing the most restrictive effective time constant across all appliances; we refer the reader to Sec. 3.1.4 of [10] for the mathematical definition of $\hat{\tau}$. The dynamics of the leaky-storage unit are governed by a decentralized hybrid control strategy [11] and its extensions in [10]; under the considered strategy, each appliance individually targets the desired relative power curve $\Pi(t)$, so that the total power consumption equals (2). \hat{P}_0 [MW] is the average steady state population consumption and the superscript a is used for single device parameters. The controller also imposes constraints on the instantaneous power excursions; any signal, compatible with all the devices, must satisfy (3). Further energy bounds (4) prevents appliances to become too warm or too cold; the respect of this constraint makes this controller non-disruptive. It is worth pointing out that the respect of (3) and (4) is a sufficient condition to guarantee the device-level feasibility of the response (2), avoiding the need for per-device simulations. Some power profiles, although feasible, may envisage the energy consumption always equal to the lower/upper bound. Since it is important to safeguard the primary cooling/heating function of TCLs, we may also impose,

$$\frac{1}{w} \int_w S(t) dt = \hat{S}_0 \quad (5)$$

where $\hat{S}_0 = \hat{P}_0 \hat{\tau}$ is the steady state energy level and w the time window of interest.

III. DEMAND SIDE RESPONSE MODEL

The key-features of the DSRM are described in detail in [8]. In this section, we only recall the mathematical formulation of the DSRM. This model enables a flexible TCL energy/power consumption and the simultaneous provision of various frequency response services, so long as (3) and (4) are respected. Three frequency control services adopted in Great Britain by National Grid [12] to ensure a safely containment and restoration of the network frequency after an infrequent severe generator outage. The first service is primary response, provided by generators (power increase) and TCLs (power reduction) responding to a quick and large change in frequency within a few seconds. With secondary response, generators and active loads extend their power increase/reduction for several minutes, bringing frequency closer to the nominal value. Finally, fast reserve units, typically fast off-line units or spinning units, respond to the system operator's signals. These units take action several minutes after the generator loss and replace the support from previous units and TCL, completing the frequency restoration. As reserve represents a long-term commitment, it is not suitable for the TCL storage capacity of several classes of appliances [10].

A. Mathematical formulation.

The solution of (1) at generic step i across the interval $[0, t]$ of length Δt is given by (6); two alternative consumption profiles, constant power $P(t) = P_i$ (6a) or a linear power $P(t) = \rho_i \vartheta + q_i$ (6b), are implemented.

$$S_i = S_{i-1} \cdot e^{-\frac{\Delta t}{\hat{\tau}}} + \hat{\tau} P_i \cdot \left(1 - e^{-\frac{\Delta t}{\hat{\tau}}}\right) \quad (6a)$$

$$S_i = S_{i-1} e^{-\frac{\Delta t}{\tau}} + \hat{\tau} \rho_i \Delta t + \hat{\tau} \left(1 - e^{-\frac{\Delta t}{\tau}}\right) (q_i - \hat{\tau} \rho_i) \quad (6b)$$

where S_{i-1} and S_i [MWh] are the energy levels at the beginning and end of interval i , respectively. Considering the generic solutions (6), the DSRM (Fig.1) can be constructed to govern the energy consumption and the frequency response provision (with consequent payback) of TCLs at all the steps i of length Δt_1 [min] of the optimization horizon. This interval is further divided into three windows of length $\Delta t_2, \Delta t_3, \Delta t_4$ [min], respectively.

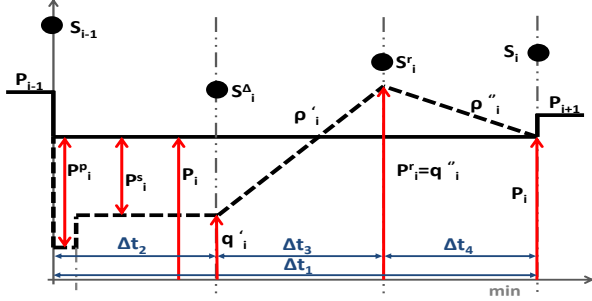


Figure 1. Demand side response model: actual energy/power and multiple services allocation.

The amount of power actually absorbed by the TCL population is P_i [MW]. These quantities are limited by (2) and (4). The included services are primary (P_i^P [MW]) and secondary (P_i^S [MW]) response. After the deployment of primary response the total TCLs consumption cannot drop below \hat{P}_{min} ; adequate reserve for primary response allocations is enforced by (7).

$$0 \leq P_i^P \leq P_i - \hat{P}_{min} \quad (7)$$

The provision of this short-term service (30 s duration) is assumed to not impact the respect of energy limits. The same minimum power requirement is applied to secondary response:

$$0 \leq P_i^S \leq P_i - \hat{P}_{min} \quad (8)$$

The power decrease is sustained for Δt_2 (generally several minutes). Equation (9) ensures that the energy level after the secondary response provision will not exceed the lower energy bound. The energy storage level S_i^A therefore respects:

$$S_i^A = \gamma_2 S_{i-1} + \beta_2 P_i - \beta_2 P_i^S \geq \hat{S}_{min} \quad (9)$$

where β_2 and γ_2 are constant values. The mathematical derivation of (9-11) and the expressions of the constant coefficients used are included in Appendix. Afterwards the energy recovery phase starts and thus the power consumption increases with a fix slope ρ'_i [$\frac{MW}{min}$] from the intercept q'_i [MW]. In the end of the second time interval, of duration Δt_3 , the power consumption is P_i^r and greater than P_i ; again, P_i^r and S_i^r cannot exceed the maximum power and energy level.

$$P_i^r = L_1 S_{i-1} + L_2 P_i + L_3 P_i^S \leq \hat{P}_{max} \quad (10)$$

$$S_i^r = H_1 S_{i-1} + H_2 P_i + H_3 P_i^S \leq \hat{S}_{max} \quad (11)$$

Now, P_i^r and S_i^r are calculated such that, within Δt_4 , from the intercept q''_i with a slope ρ''_i , the power consumption and stored energy return to originally scheduled level P_i and S_i , respectively (see the appendix). Considering the expressions of

L_1 and L_2 in the appendix, it is easy to prove that these quantities are 0 and 1 respectively; this implies that the additional amount of power to add to the contingency reserve requirement, P_i^{ar} , only depends on the amount of secondary response allocated by means of L_3 , as stated in (12)

$$P_i^{ar} = P_i^r - P_i = (L_2 - 1)P_i + L_3 P_i^S = L_3 P_i^S \quad (12)$$

In particular, considering a time step Δt_1 and fixing the secondary response commitment (Δt_2), L_3 (and hence P_i^{ar}) only varies with Δt_3 ($\Delta t_3 + \Delta t_4 = \Delta t_1 - \Delta t_2 = \text{const}$).

IV. THE STOCHASTIC UNIT COMMITMENT MODEL

We recall the objective function and the fundamental constraints of the multi-stage stochastic unit commitment (SUC) with rolling planning used in [8] and largely based on [3]. The solution is obtained over a typical scenario tree. The SUC model optimizes the system operation by simultaneously scheduling energy production, inertia-dependent primary response, load-dependent secondary response and fast reserve, in the light of uncertainties associated with renewable production and generation outages. The objective function minimizes the balance between the expected generation operation cost and the expected cost of not meeting demand:

$$\sum_{n \in N} \pi(n) \left(\sum_{g \in G} C_g(n) + \Delta \tau(n) C_{LL} P_{LS}(n) \right) \quad (13)$$

subject to typical load balance constraint and local constraints for the thermal units. We refer the reader to [3] for details on these constraints and the equations describing generation costs. The primary $R_g^P(n)$ and secondary response $R_g^S(n)$ characteristics of the generating units are modelled according to the machine load level:

$$0 \leq R_g^P(n) \leq N_g^{up}(n) R_g^{P,max} \quad (14a)$$

$$R_g^P(n) \leq f_g^P N_g^{up}(n) (P_g^{max} - P_g(n)) \quad (14b)$$

$$0 \leq R_g^S(n) \leq N_g^{up}(n) R_g^{S,max} \quad (15a)$$

$$R_g^S(n) \leq f_g^S (N_g^{up}(n) P_g^{max} - P_g(n)) \quad (15b)$$

The contingency reserve characteristic $R_g^R(n)$ of generator is modelled as the power increase from a generator at the maximum unit's ramp rate within the predefined delivery time $t_R = \Delta t_3$, and it is also bounded by the spinning headroom of each generator:

$$0 \leq R_g^R(n) \leq N_g^{up}(n) * r_g^{max} * t_R \quad (16a)$$

$$R_g^R(n) \leq N_g^{up} \left[P_g^{max} - P_g(n) - \max \left(\frac{R_g^P(n)}{f_g^P}, \frac{R_g^S(n)}{f_g^S} \right) \right]; \quad (16b)$$

constraint (16b) requires that the spinning headroom scheduled for response should not to be used for reserve in order to allow the recovery of response provision for next time step.

A. Incorporating the DSRM into the SUC

The energy level, at the end of each time period and referred to the generic node n , can be expressed by rearranging (6a) in the scenario tree framework as:

$$S_T(n) = S_T(a(n)) \cdot e^{-\frac{\Delta t_1}{\tau}} + P_T(n) \hat{t} \cdot \left(1 - e^{-\frac{\Delta t_1}{\tau}}\right) \quad (17)$$

where $a(n)$ identifies the node preceding node n in a generic scenario of the scenario tree. The TCL power levels are constrained by (2) and (4). The response services provision and the consequent payback are constrained by (7-11). In Sec. II-A we explained the need for the TCLs average energy to remain close to its steady state \hat{S}_0 . Let us first define $n_{\theta, \theta+i}^{(\sigma)}$ as the node at time step $t = \theta + i$ included in a generic scenario σ of the scenario tree and with root at time step $t = \theta$. The equality constraint (18) follows from inserting (1) in (5) and using (17) to eliminate $P_T(n)$ in each interval and it is in accordance with the notation introduced.

$$\frac{1}{w+p} \left[\varphi S(n_{\theta, \theta-p}^{(\sigma)}) + \sum_{i=1-p}^{w-1} S(n_{\theta, \theta+i}^{(\sigma)}) + \chi S(n_{\theta, \theta+w}^{(\sigma)}) \right] = \hat{S}_0 \quad (18)$$

Due to the application of rolling planning, the average energy constraint (18) also has to account for *past* energy levels p , already reached by TCLs at previous time steps. The constant values φ and χ are defined below and their sum equals 1.

$$\varphi = \left[\frac{\tau}{\Delta t_1} - \frac{e^{-\frac{\Delta t_1}{\tau}}}{\left(1 - e^{-\frac{\Delta t_1}{\tau}}\right)} \right]; \chi = \left[\frac{1}{\left(1 - e^{-\frac{\Delta t_1}{\tau}}\right)} - \frac{\tau}{\Delta t_1} \right] \quad (19)$$

B. Inertia-dependent Frequency Response Requirement

In this section we recall the frequency response requirements for the SUC presented in [8] and based on the analytical derivation in [3]. The time evolution of system frequency deviation is described by a first order ODE:

$$2H \frac{d\Delta f(t)}{dt} + DP_D \Delta f(t) = \sum_{g \in \mathcal{G}} R_g^P(t) + P_T^P(t) - \Delta P_L \quad (20)$$

The minimum required level of system inertia $H(n)$ to limit the absolute value of the rate of change of frequency (RoCoF) above the maximum value follows from (20):

$$H(n) = \frac{\sum_{g \in \mathcal{G}} H_g * P_g^{max} * N_g^{up}(n)}{f_0} \geq \left| \frac{\Delta P_L^{max}}{2RoCoF_{max}} \right| \quad (21)$$

where H_g is the inertia constant [s] of generator g , f_0 is nominal frequency (50Hz) and ΔP_L^{max} [MW] the amplitude of the maximum generation loss. The frequency nadir depends on system inertia, generators governor response and TCL response. The scheduled primary response of generators and TCLs is assumed to linearly increase with time by t_p [3] and, after this time, is constant. This choice can be actually implemented for TCLs by means of the control strategy adopted [11]. The aim is now ensuring that $|\Delta f_{nadir}(n)| \leq \Delta f_{max}$ at each node n ; following equivalent steps as in [3], the primary response requirement $P^P(n) = \sum_{g \in \mathcal{G}} R_g^P(n) + P_T^P(n)$, that satisfies the constraint on frequency nadir has to respect:

$$P^P(n) \geq \Delta P_L - DP_D(n) \Delta f_{max} \quad (22a)$$

$$P^P(n) * H(n) \geq k^*(n) \quad (22b)$$

We refer the reader to Sec. III-C of [3] for the mathematical proof of (22b); in particular, we demonstrated that for any given

value of D and ΔP_L^{max} , it exists only a unique value $k^*(n) = P^P(n) * H(n) \Rightarrow |\Delta f_{nadir}(n)| = \Delta f_{max}$. The bilinear constraint (24b) is then transferred to a mix integer linear formulation by applying standard reformulation method as in [3] (based on [13]). The provision of secondary response permits to stabilize the frequency deviation at least at maximum intermediate steady state value Δf_{max}^{iss} ; the combined action of secondary response and fast reserve brings frequency back to f_0 . The intermediate steady-state frequency deviation is obtained, by assuming in (20) that RoCoF is zero; hence the service requirement is expressed by,

$$P^S(n) \geq \Delta P_L^{max} - DP_D(n) \Delta f_{max}^{iss} \quad (23)$$

Finally, contingency reserve is required to assist the frequency recovery by taking over for frequency responsive plants and hence restoring their response capability. TCL energy recovery also affects the required reserve as the extra power absorbed by the devices is supplied by reserve generators. The following requirement is therefore applied:

$$P^R(n) \geq \Delta P_L^{max} + P^{ar}(n) \quad (24)$$

V. CASE STUDIES AND RESULTS

The Great Britain (GB) 2030 power system scenario [3] is adopted as reference case to quantify the annual economic and environmental value of individual TCLs. The maximum system demand is 60 GW and the total conventional generation capacity is 70 GW. Table I summarizes the characteristics of conventional plants.

TABLE I. THERMAL PLANTS CHARACTERISTICS – REFERENCE CASE

	Nuclear	Coal	CCGT	OCGT
<i>Number of units</i>	6	40	70	50
<i>Rated Power (MW)</i>	1800	500	500	200
<i>Min Stable Generation (MW)</i>	1800	250	250	50
<i>No-load cost (£/h)</i>	0	3364	7809	8000
<i>Marginal cost (£/MWh)</i>	10	72	51	110
<i>Start up cost (£)</i>	n/a	90000	32000	0
<i>Start up time (h)</i>	n/a	6	4	0
<i>Min down time (h)</i>	n/a	4	4	0
<i>Inertia Constant (s)</i>	6	5	5	5
<i>Max Primary Response (MW)</i>	0	50	50	20
<i>Primary Response Slope</i>	0	0.3	0.4	0.6
<i>Max Secondary Response (MW)</i>	0	50	50	20
<i>Secondary Response Slope</i>	0	0.3	0.4	0.6
<i>Ramp Rate (MW/min)</i>	0	15	25	35

The installed wind capacity is assumed to be 48 GW (45% wind penetration); wind farms do not provide inertial response and frequency services as current practice [2, 3]. C_{LL} is set at 30000 £/MW. The load damping rate D is 0.5/50Hz. The maximum frequency deviation Δf_{max} is set at -0.8 Hz and Δf_{max}^{iss} is -0.5 Hz. These values reflect the current National Grid (NG) settings, which are not expected to be changed [9]. The maximum RoCoF instead is expected to increase to deal with the lack of system inertia due to the integration of RES; we adopt 0.5 Hz/s [9]. The time to deliver primary response is $t_p=5s$, in line with future NG proposals [9]. The duration of each time step of the SUC is $\Delta t_1=30$ minutes; we set $\Delta t_2=10$ minutes, and finally we select $\Delta t_3=\Delta t_4=\Delta t_2$. Reference parameters for domestic fridge-freezers are taken from [10]; The parameters were then varied by $\pm 10\%$ to establish the data

for an heterogeneous set of 40 million [2] appliances; in particular, $\hat{\tau}=4.5$ h, $\hat{S}_{max}= 7.6$ GWh, $\hat{S}_{min}= 6.5$ GWh, $\hat{P}_{min}= 0.85$ GW, $\hat{P}_{max}=3.7$ GW and $\hat{P}_0= 1.6$ GW. The optimization was solved by using FICO Xpress 7.1, which was linked to a C++ simulation application via the BCL interface [14].

A. Value of TCLs under different future system scenarios.

This section assesses the value of individual thermostatic loads due to their flexible energy/power consumption and their provision of primary and secondary control. The annual savings for the reference case (central bar in Fig.2) reveal high revenues (14.61£/TCL/yr) for TCL owners; this represents an attractive option also for a long-term investment as the cost for installing a smart controller would be around 3-5 £/TCL [4] (based on manufactures analysis) and the typical lifetime of a domestic refrigerator is approximately 12-15 years [2].

Figure 2 also indicates that individual refrigerators, controlled with the DSRM, may annually reduce carbon emissions (red curve) from the electricity production by about 104 kg of CO₂, greatly helping the system operator towards the achievement of strict greenhouse reduction targets. This result comes from the TCL ability to reduce the volume of wind curtailment and part-loaded conventional plants. During those hours characterized by low net demand (system demand minus wind production), TCL can easily increase their consumption. In effect, they may not only exploit a relatively low system operation cost and but also be compensated by providing the large amounts of frequency response services (higher requirements due low inertia).

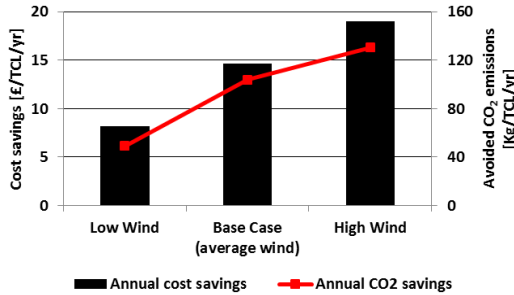


Figure 2. Annual cost and CO₂ emissions savings per individual TCL varying the installed wind capacity.

The reference case is compared against a *low wind* (32GW of wind – left bar) and a *high wind* scenario (64 GW of wind – right bar) to demonstrate the impact of wind penetration on the value of flexible TCLs. The economic value for TCLs varies between 8 £/TCL/yr (*low wind*) and 19 £/TCL/yr (*high wind*), while the carbon reduction varies between 50 kg of CO₂/TCL/yr (*low wind*) and 130 kg of CO₂/TCL/yr (*high wind*). This is a positive result as the installed wind capacity is going to constantly increase over the next decades [2].

Another study (Fig. 3) investigates the sensitivity of the value of TCLs to the thermal generators flexibility varying the installed CCGT and nuclear generation (i.e. *high nuclear*: 16.2 GW of nuclear and 30 GW of CCGT generation; *high CCGT*: no nuclear generators and 45 GW of CCGT generation). Coal, OCGT and wind generation are as in the reference case (Table I). The smart control of TCL dynamics reaches higher value under the *high nuclear* scenario. This results from the low level

of flexibility of the installed thermal plants, mainly due to the inability to vary the output and provide response services of nuclear plants.

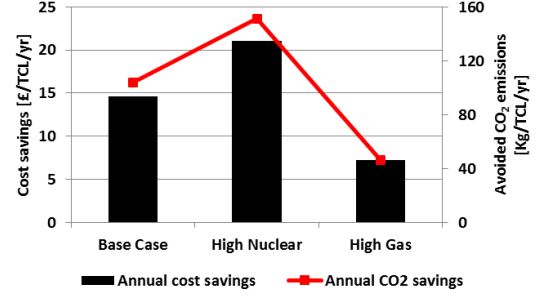


Figure 3. Annual cost and CO₂ emissions savings per individual TCL varying the mix of thermal plants.

The value of the devices significantly reduces for the *high CCGT* scenario as the system presents sufficient flexibility in form of dispatchable CCGT generators (high ramp rates and response slopes). Similar trends are identified considering the annual avoided CO₂ emissions realized by individual TCLs (red curve).

B. Value of TCL under different post-fault frequency requirements.

Embedded generators are equipped with RoCoF-sensitive protection schemes. In order to prevent dangerous cascading disconnection of such distributed generation due to high RoCoF conditions, the system operator may review and increase the current RoCoF threshold (0.125 Hz/s) to deal with a shortage of system inertia. Figure 4 investigates the impact of varying this threshold from 0.2 Hz/s up to 0.5 Hz/s (reference case).

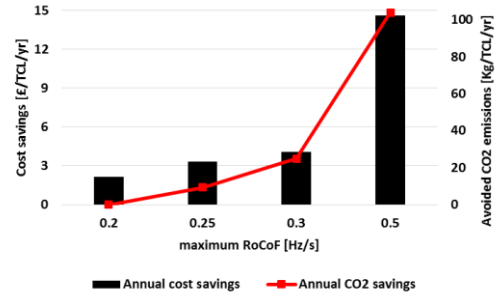


Figure 4. Annual cost and CO₂ emissions savings per individual TCL varying the RoCoF threshold in (23).

Keeping a low value of the maximum RoCoF setting would lead to poor annual cost savings and emissions reductions. In effect, in order to satisfy (21), many conventional generators would be committed to meet the system inertial response requirement; however, this reduces the advantage to schedule refrigerators for the fulfillment of response service requirements (22-23). The system operator should avoid this condition as it leads to high operational cost and large volumes of wind curtailment. It is worth pointing out that moving from 0.3 Hz/s to 0.5 Hz/s, the value of TCLs rises by a factor three.

Due to an overall shortage of inertia, the transient frequency evolution will achieve the nadir much faster than in current GB operation (around 10s [9]). This will force conventional plant to provide their governor response more rapidly. Currently, the

maximum time to deliver primary response (t_p) is 10s but NG has envisaged to reduce it up to 5s (reference setting in this work), for future low carbon scenarios [9] (response capabilities and response slopes are assumed to remain the same [9]). Figure 5 shows the impact on value of TCLs if this change will be actually implemented (reference case) or not.

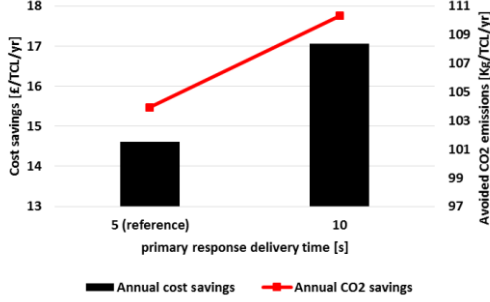


Figure 5. Annual cost and CO2 emissions savings per individual TCL varying t_p in (24).

The system operation cost savings and CO₂ emission reductions are actually higher for $t_p=10$ s. In this case, generators provide a slower response and therefore the primary frequency response requirement increases especially for those system conditions with low net demand. The high need for TCL support in the fulfillment of this requirement makes their contribution more valuable.

C. Value of TCLs under different devices penetration and parameters.

The economic and environmental value for TCL is now assessed varying devices parameters. In Fig. 6 the cost and carbon emission savings are function of the percentage of the appliances, within the whole population, that is controllable.

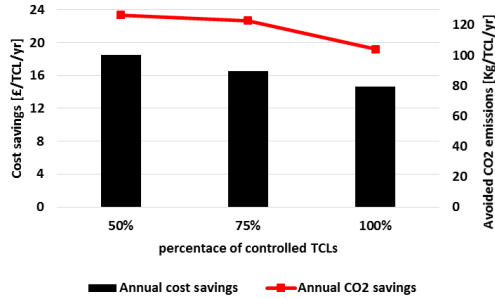


Figure 6. Annual cost and CO2 emissions savings per individual TCL varying the percentage of controllable devices.

For instance, if 75% of the TCLs are dispatchable (with the DSRM), the remaining 25% does not provide response services and it always absorbs the nominal power over the optimization horizon. Higher savings are achieved for lower TCL penetrations, although the differences are mild. The results in Fig. 7 indicates how the value of TCLs is affected by improvements in the refrigerators insulation system (e.g. new materials, or thicker insulation). In particular, we decreased the thermal conductance of the refrigerators by 5, 10 and 20%. This mainly changes the thermal time constant $\hat{\tau}$, the average power consumption \hat{P}_0 and the power controller thresholds \hat{P}_{max} and \hat{P}_{min} .

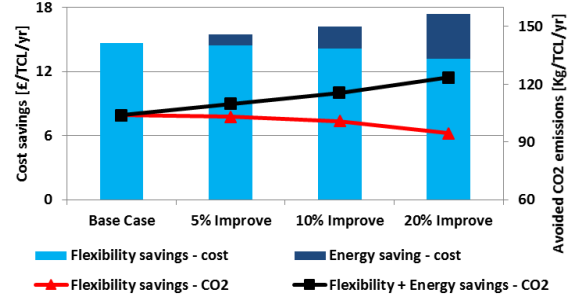


Figure 7. Annual cost and CO2 emissions savings per individual TCL improving the TCL insulation.

Overall, the economic and environmental benefits for individual TCL grow with a better insulation system. We identify two reasons that build this increment, characterized by opposite trends. A better insulated TCL absorbs less power (smaller duty cycle) and thus it is more efficient. For this reason, it realizes energy savings (blue areas of the bars) compared to the reference devices. On the other side, the value of TCLs, associated to the flexible power/energy consumption and provision of response services, decreases if the insulation level grows (light blue areas). This is mainly because of their lower average power consumption, which is reflected in a reduced amount of response services to supply. Moreover, TCL may not be able to absorb high power consumptions (to compensate the response service provision) as they may hit the upper energy threshold more rapidly due to their better insulation. Similar trends are identified for the carbon emissions savings (only due to flexibility savings – red curve, overall savings – black curve).

The final case study (Fig. 8) shows the effect of changing the coefficient of performance (COP) of the refrigerators [15].

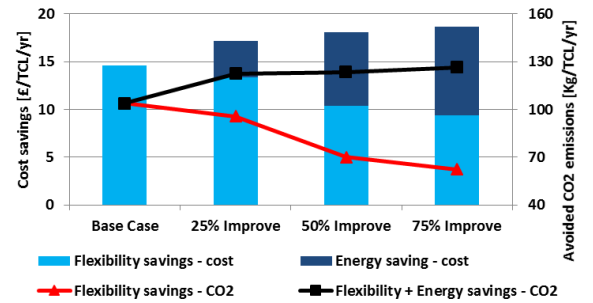


Figure 8. Annual cost and CO2 emissions savings per individual TCL varying the COP of the devices.

This parameter represents the ratio of the heat extracted from the refrigerated compartment over the electrical energy consumed to do this; it is a quick measure of the refrigerator efficiency. The findings previously discussed regarding the thermal insulation can be extended to this case. TCLs are more efficient and would consume less energy, increasing the value of TCL; however, the increased efficiency also reduces the ability to contribute to the system frequency control (flexibility savings). The last two case studies revealed the increasing overall savings for improved thermal insulation and higher COPs; this represent a positive input for TCL manufactures to invest in higher device efficiency.

VI. CONCLUSIONS

This paper built on a novel and accurate model (DSRM) for optimal flexible scheduling of TCL energy/power consumption and the simultaneous provision of a portfolio of frequency response services [10]. The methodology developed in [10] has been applied to several case studies that quantified the value for individual TCLs in term of annual system operation cost savings and annual CO₂ emissions reduction. The results demonstrated the large benefits for individual TCLs due to a smart energy operation. In particular, high penetrations of wind generation and other inflexible technologies (e.g. nuclear) are key drivers of the value of TCLs. Moreover, the settings of the frequency response requirements show significant impact on the need and benefits of smart operation of TCLs. Initial but effective input are offered to devices manufactures; the overall system operation cost savings and CO₂ reductions increase by improving the TCL efficiency (e.g. thermal insulation and coefficient of performance), although these may decrease the flexibility provided by TCLs. This study demonstrated the value of TCLs by means of a cost-based analysis. However, additional work is required to develop a detailed business model that recognizes this value into a market-based framework (i.e. TCLs may operate individually, or be managed by entities such as aggregators or the system operator). The current regulatory framework may need to change to facilitate TCL participation. Further studies will run a complete investment analysis for the smart operation of TCLs considering the entire lifetime of the appliances and the accurate cost for the control equipment.

APPENDIX

We now proceed to validate (9-11) and derive the constant coefficients introduced in those equations.

$$\gamma_i = e^{-\frac{\Delta t_i}{\tau}} \quad (25a) \quad \beta_i = \hat{\tau} \cdot (1 - \gamma_i) \quad (25b)$$

Considering (6b), we calculate S_i^r as the energy at the end the second sub-interval in Fig.1.

$$S_i^r = \gamma_3 S_i^A + \hat{\tau} \Delta t_3 \rho_i' + \beta_2 q_i' - \beta_2 \hat{\tau} \rho_i' \quad (26)$$

$$\rho_i' = \frac{P_i^r - (P_i - P_i^s)}{\Delta t_3} \quad q_i' = P_i - P_i^s \quad (27)$$

By inserting (30) in (29) and considering (9) we obtain

$$S_i^r = R_1 S_{i-1} + R_2 P_i - R_2 P_i^s + R_3 P_i^r \quad (28)$$

$$R_1 = \gamma_2 \gamma_3; \quad R_2 = \gamma_3 \beta_2 - \hat{\tau} + \beta_3 + \frac{\beta_3 \hat{\tau}}{\Delta t_3}; \quad R_3 = \hat{\tau} - \frac{\beta_3 \hat{\tau}}{\Delta t_3} \quad (29)$$

Hence we derive again S_i^r (34) as the energy at the beginning the third sub-interval in Fig.1

$$S_i^r = \frac{1}{\gamma_4} (S_i - \hat{\tau} \Delta t_4 \rho_i'' - \beta_4 q_i'' + \beta_4 \hat{\tau} \rho_i'') \quad (30)$$

$$S_i = \gamma_1 S_{i-1} + \beta_1 P_i \quad (31)$$

$$\rho_i'' = \frac{P_i - P_i^r}{\Delta t_4} \quad q_i'' = P_i^r \quad (32)$$

$$G_1 = \frac{\gamma_1}{\gamma_4}; \quad G_2 = \frac{\beta_1 - \hat{\tau}}{\gamma_4} + \frac{\beta_4 \hat{\tau}}{\gamma_4 \Delta t_4}; \quad G_3 = \frac{\hat{\tau} - \beta_4}{\gamma_4} - \frac{\beta_4 \hat{\tau}}{\gamma_4 \Delta t_4}; \quad (33)$$

$$S_i^r = G_1 S_{i-1} + G_2 P_i + G_3 P_i^r \quad (34)$$

We impose (28) equals (34) to derive the expression of P_i^r in (10), with

$$L_1 = \frac{G_1 - R_1}{R_3 - G_3}; \quad L_2 = \frac{G_2 - R_2}{R_3 - G_3}; \quad L_3 = \frac{R_2}{R_3 - G_3} \quad (35)$$

Finally we substitute (10) in (34) to formulate (11) with

$$H_1 = G_1 + G_3 L_1; \quad H_2 = G_2 + G_3 L_2; \quad H_3 = G_3 L_3 \quad (36)$$

ACKNOWLEDGMENTS

Fei Teng is supported by the UK Engineering and Physical Sciences Research Council under Grant EP/L014351/1.

REFERENCES

- [1] Ofgem, "Creating the right environment for demand-side response," Jun. 2013. [Online]. Available: <https://www.ofgem.gov.uk/ofgem-publications/75245/creatingtherightenvironmentfordemand-sideresponse.pdf>.
- [2] M.Aunedi et al., "Economic and Environmental Benefits of Dynamic Demand in Providing Frequency Regulation," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2036-2048, 2013.
- [3] F.Teng, V.Trovato and G.Strbac, "Stochastic scheduling with inertia-dependent frequency regulation," *IEEE Trans. Power Syst.*, accepted, in press - 2015.
- [4] J. Short, D. Infield and L. Freris, n"Stabilization of Grid Frequency Demand in Providing Dynamic Demand Control," *IEEE Trans. on Power Syst.*, vol. 22, no. 3, pp. 1284 - 1293, 2007.
- [5] J.L. Mathieu, M. Kamgarpour, J. Lygeros and D.S. Callaway, "Energy arbitrage with thermostatically controlled loads," in *European Control Conference (ECC)*, Zurich, 2013.
- [6] R.Moreno, R.Moreira and G. Strbac, "A MILP model for optimising multi-service portfolios of distributed energy storage," *Applied Energy - Elsevier*, 2014.
- [7] E.Karangelos and F. Bouffard, "Towards Full Integration of Demand-Side Resources in Joint Forward Energy/Reserve Electricity Markets," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 280-289, Feb. 2012.
- [8] V. Trovato, F. Teng and G. Strbac, "Stochastic scheduling with thermostatically controlled loads," *IEEE Transactions on Smart Grids*, 2015 - under review.
- [9] National Grid, "Frequency response technical sub-group report," 2011.
- [10] V. Trovato, S. Tindemans and G. Strbac, "The Leaky Storage Model for Optimal Multi-service Allocation of Thermostatic Loads," *IET Generation, Transmission & Distribution*, accepted, inpress - 2015.
- [11] S.Tindemans, V.Trovato and G.Strbac, "Decentralised control of thermostatic loads for flexible demand response," *IEEE Trans. on Control Syst. Technol.*, vol. 23, no. 5, pp. 1685-1700, 2015.
- [12] National grid, "Balancing services," [Online]. Available: <http://www2.nationalgrid.com/uk/services/balancing-services/>.
- [13] W. Qianfan, J-P. Watson and Y. Guan, "Two-Stage Robust Optimization for-Contingency-Constrained Unit Commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2366-2375, 2013.
- [14] Fico Xpress iptimization suite, [Online]. Available: <http://www.fico.com/en/products/fico-xpress-optimization-suite>.
- [15] Coefficient of performance, [Online]. Available: http://www.industrialheatpumps.nl/en/how_it_works/cop_heat_pump/.

Vincenzo Trovato (M'15, S'M'12) is Research Associate at Imperial College London where he received a Ph.D. degree in electrical engineering in 2015. His research interests include demand side response and its impact on the power system control and economics.

Fei Teng (M'15) is Research Associate at Imperial College London. His research interests include power system operation, integration of renewable energy.

Goran Strbac (M'95) is Professor of Electrical Energy Systems with Imperial College London. His research interests also are in modeling and optimization of electricity system operation and investment.