

# A Scalable Low-Cost Solution to Provide Personalised Home Heating Advice to Households

Alex Rogers, Siddhartha Ghosh, Reuben Wilcock and Nicholas R. Jennings  
Electronics and Computer Science  
University of Southampton, Southampton, SO17 1BJ, UK  
{acr,sg2,rw3,nrj}@ecs.soton.ac.uk

## ABSTRACT

In this paper, we present a deployed prototype of a scalable low-cost solution providing personalised home heating advice to households. Our solution, named MyJoulo ([www.myjoulo.com](http://www.myjoulo.com)), uses intelligent algorithms to analyse data collected from a specially designed USB temperature logger, placed on top of the thermostat, in order to build a thermal model of the home and to infer the operational settings of the heating system. This model is then used to calculate the impact, in terms of percentage reduction in heating costs, of various interventions (such as reducing the thermostat set-point temperature or adjusting timer settings); providing specific actionable advice to the household. The system was launched in beta form in December 2012 and registered over 750 users in its three months of operation.

## Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

## General Terms

Design, Experimentation

## Keywords

home energy, heating, temperature logger

## 1. INTRODUCTION

Recent years have seen significant research into providing households with realtime feedback on their energy consumption in an effort to reduce the carbon emissions from the domestic sector. Much of this work has focused on low cost solutions that use a small number of easily deployed sensors in conjunction with artificial intelligence algorithms to infer what cannot be measured directly. A mature example of this approach is non-intrusive appliance load monitoring (NIALM) that attempts to disaggregate total electricity consumption, using a single clamp-on sensor on the main feed to the

home, into individual appliance level consumption [3]. These systems typically model the operational state of individual appliances within the home using a hidden Markov model, and then infer the operational state of each appliance at any time, using both the instantaneous total consumption of the home, and the changes in consumption that occur when appliances change state [6, 5]. A number of commercial entities already provide such services to households, focusing on a small number of high consumption appliances (see for example [www.plotwatt.com](http://www.plotwatt.com) and [www.bidgely.com](http://www.bidgely.com)).

However, electrical appliances are not the only consumers of energy within the home, and other forms of consumption can be equally or indeed more significant. In particular, home heating typically accounts for 60-70% of domestic energy consumption in cooler climates, and yet, has received less attention in the research literature. Recent work has tended to focus on smart heating controls that use some modelling of the home to provide feedback to users and to optimally control heating [8, 10], and innovative products such as the Nest thermostat (see [www.nest.com](http://www.nest.com)) have brought these approaches to the market. To date though, no system offers the type of low-cost, minimally intrusive approach, aimed at providing advice to householders, that has been successfully applied to electricity consumption. There appear to be two reasons for this. Firstly, heating in many countries uses gas-fired boilers, and non-invasive metering of gas consumption is not straightforward. Furthermore, there is little quantitative data on how households use their existing heating systems [4], and thus, developing algorithms to provide actionable advice in this context is problematic.

To address both these issues, in this paper we present a deployed prototype system, named MyJoulo, that both provides personalised home heating advice to households, and is collecting data on typical heating use across a large number of UK homes. The system was launched in beta form in December 2012 and registered over 750 users in its three months of operation. Rather than attempting to directly measure gas or electricity consumption of the heating system, we use a specially designed low-cost USB temperature logger, placed on top of the thermostat (the single point of control of most heating systems), to infer the operation of the heating system indirectly. The temperature logger is sent to users by mail, and returned after use, providing a cost effective solution at scale. Our solution then uses external temperature data, from the internet, to build a thermal model of the home, describing the rate at which heat is gained and lost, which is then used to infer the operational settings (e.g. the timer settings and the thermostat set-point) of the heating system. This model, and the inferred operational settings, is then used to calculate the impact (in terms of percentage reduction in heating costs) of interventions such as reducing the thermostat set-point temperature or adjusting timer settings; providing specific actionable advice to the household.

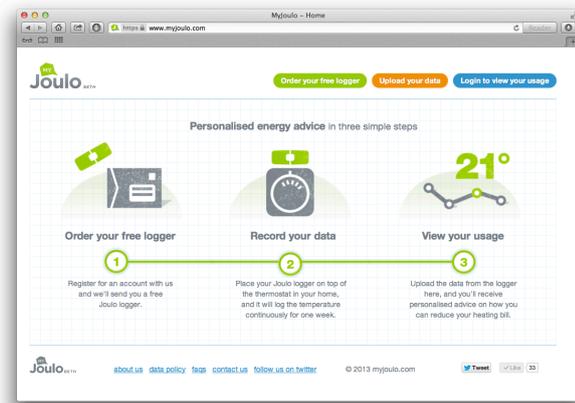


Figure 1: MyJoulo website ([www.myjoulo.com](http://www.myjoulo.com)).



Figure 2: USB temperature logger in packaging as sent to the 750 users who registered for the initial trial.

The MyJoulo system consists of three components: (i) a website through which households can sign up to the service and see the final analysis and feedback, (ii) a low-cost USB temperature logger which is sent to the user once they have registered on the website, and (iii) intelligent algorithms that model and calculate the impact of interventions. In the rest of this paper, we describe each in more detail, and then discuss the trial deployment, the lessons learned and the challenges to be addressed in future work.

## 2. MYJOULO WEBSITE

The website (see [www.myjoulo.com](http://www.myjoulo.com) and Figure 2) allows households to request a temperature logger be mailed to them, upload data recorded by the logger, and view the resulting energy saving advice. Since the temperature logger is returned after use, the marginal cost of providing the service is very low (just postage charges and packaging replacement), and the service can operate at scale at low cost.

## 3. USB TEMPERATURE LOGGER

In our initial evaluation, commercially available temperature loggers failed to provide the necessary accuracy and ease of use (typically requiring installed software or additional hardware to access the logger data, both of which precluded their easy use by our users), and thus, a specially designed temperature logger, based around an Atmel AT90USB162 micro-controller with a Texas Instruments TMP275 temperature sensor, was developed (see Figure 3 for a schematic of the logger, and Figure 4 for a photograph of

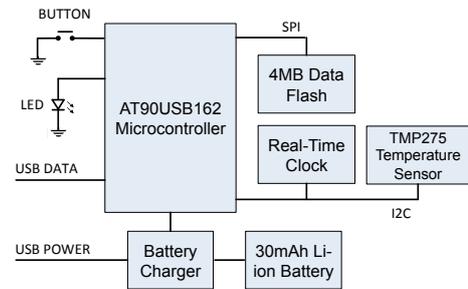


Figure 3: USB temperature logger schematic.



Figure 4: Deployed USB temperature logger.

the logger correctly positioned on top of a home heating thermostat). This provides better than  $\pm 0.2^{\circ}\text{C}$  measurement accuracy without requiring additional calibration. The logger is triggered by the single button, and then records temperature at 2 minute intervals for 7 days; collecting a total of 5041 measurements. The logger firmware supports both USB HID and Mass Storage protocols, such that it can be configured prior to dispatch (setting sampling rates and serial number), and then appear as a conventional flash drive (with the recorded data in a 'DATA.TXT' file), allowing the householder to upload the recorded data to the website where it is analysed in real time.

## 4. ANALYSIS ALGORITHMS

The analysis algorithms form the core of the MyJoulo system; inferring the operational settings of the heating system from the raw temperature measurements collected by the USB temperature logger, and calculating the energy savings from various interventions which might then be recommended to the user. The first stage of the analysis consists of retrieving external temperature data from the internet. We use data from the closest live weather station to the postal address that the logger was sent to (collected automatically from [www.weatherunderground.com](http://www.weatherunderground.com)). However, since these temperature measurements are not time aligned with the internal temperature measurements, and often exhibit periods of missing data, we use Gaussian process regression to interpolate

the temperature readings and complete any missing periods. We then use a simple thermal model of the home to infer the parameters that describe the heat output of the heating system, the rate at which heat leaks away from the interior of the home, and the operational settings of the heating system (specifically, the times at which the heating system starts and stops each day, and the thermostat set-point). Finally, using seasonal average temperature patterns, we calculate the energy savings that would result from making specific interventions in the use of the heating system; specifically, calculating the impact from reducing the set-point either by one degree, or to the recommended value of 19°C. These insights are then presented to the user through visually appealing infographics.

## 4.1 Interpolating External Weather Readings

In order to time align the internal and external temperature measurements, and to handle missing data, we use Gaussian process (GP) regression [7]. Our USB temperature logger collects  $n$  temperature measurements,  $\mathbf{T}_{log}$ , at times,  $\mathbf{t}$ . For the same time period, we download from the internet,  $m$  external temperature measurements,  $\mathbf{T}_{ext}$ , made at times,  $\boldsymbol{\tau}$ . We then wish to infer the external temperatures,  $\hat{\mathbf{T}}_{ext}$ , at times  $\mathbf{t}$ . Using the standard result of GP regression, this is given by:

$$\hat{\mathbf{T}}_{ext} = K(\boldsymbol{\tau}, \mathbf{t})^T K(\boldsymbol{\tau}, \boldsymbol{\tau})^{-1} \mathbf{T}_{ext} \quad (1)$$

where the elements of the matrices,  $K(\boldsymbol{\tau}, \mathbf{t})$  and  $K(\boldsymbol{\tau}, \boldsymbol{\tau})$ , is defined by a covariance function,  $k(t, t')$ , that encodes our prior knowledge of how external temperature measurements at a time  $t$  correlate to measurements at another time,  $t'$ . The function we use takes the form:

$$k(t, t') = k_{SE}(t, t') + k_P(t, t') + k_N(t, t') \quad (2)$$

The first term encodes our belief that the external temperature will vary smoothly, and is given by a standard *squared exponential* function:

$$k_{SE}(t, t') = \sigma_{SE}^2 \exp\left(-\frac{(t-t')^2}{2\ell_{SE}^2}\right) \quad (3)$$

where  $\sigma_{SE}^2$  is the amplitude of the process and  $\ell_{SE}$  is the characteristic length-scale that determines how rapidly the correlation between temperature measurements decreases as  $t$  and  $t'$  diverge. The second term encodes the observation that the external temperature profiles often exhibit diurnal patterns (cold at night, and warmer throughout the day), and thus, we also use a squared exponential periodic covariance function with unit periodicity given by:

$$k_P(t, t') = \sigma_P^2 \exp\left(-\frac{\sin^2 \pi(t-t')}{2\ell_P^2}\right) \quad (4)$$

where  $\sigma_P^2$  is again the amplitude of the process and  $\ell_P$  scales the change in correlation across the day. This ensures a reasonable temperature profile is inferred even when a whole day of external temperature measurements is missing. Finally, we also consider the noise in measurements by applying an additive Gaussian noise function given by:

$$k_N(t, t') = \sigma_N^2 \delta_{t,t'} \quad (5)$$

where  $\delta_{t,t'}$  is the Kronecker delta and  $\sigma_N^2$  is the noise variance. We find the values of the set of hyper-parameters,  $\boldsymbol{\theta}_{GP} = \{\sigma_{SE}^2, \sigma_P^2, \sigma_N^2, \ell_{SE}, \ell_P\}$  by maximum likelihood, given that the log-likelihood of the observed data given the hyper-parameters,  $\ln p(\mathbf{T}_{ext} | \boldsymbol{\theta}_{GP})$ , is described by:

$$-\frac{1}{2} \ln |K(\boldsymbol{\tau}, \boldsymbol{\tau})| - \frac{1}{2} \mathbf{T}_{ext}^T K(\boldsymbol{\tau}, \boldsymbol{\tau})^{-1} \mathbf{T}_{ext} - \frac{m}{2} \ln(2\pi) \quad (6)$$

---

### Algorithm 1 Algorithm to estimate $\hat{\mathbf{T}}_{int}$

---

```

for  $i = 1 : n - 1$  do
   $r_h^i = 0$ 
  if  $(t^i \geq s_1$  and  $t^i \leq e_2)$  then
    if  $\hat{T}_{int}^i < T_{set}$  then  $r_h^i = 1$ 
    if  $m = 2$  and  $(t^i \geq e_1$  and  $t^i \leq s_2)$  then  $r_h^i = 0$ 
  end if
  Compute  $\hat{T}_{int}^{i+1}$  as per Equation 7
end for

```

---

Since Gaussian process regression is relatively computationally expensive and takes approximately 30 seconds to infer external temperature at the 5041 sample times (using 7 days of external temperature data sampled at 10 minute intervals, and implementing the Gaussian process regression using the Python pyGPs library — [github.com/marionmari/pyGPs](https://github.com/marionmari/pyGPs) — on a standard server), we use a simpler spline approach (which does not represent periodicity in the temperature measurements) to perform an initial regression such that we can present data to the user immediately after they have uploaded their data. We then replace this interpolation with the Gaussian process version once the background processing has completed.

## 4.2 Thermal Modelling

In order to provide feedback to householders on how they are operating their heating system, and the energy savings that they might achieve by making specific actionable changes, it is necessary to model both the thermal properties of the home, and also to infer the operational settings of the heating system. Since the USB temperature logger is placed on top of the home's thermostat, we are not required to model the flow of heat throughout the home, but rather can focus on the heat flows in the immediate vicinity of the thermostat, which, being the control point of the heating system, will ultimately determine the operation time of the boiler, and thus, the heating costs incurred. To do so, we use a standard building thermal model, where heat leaks from the home at a rate proportional to the difference between internal and external temperatures [1]. This model can be expressed as a discrete stochastic difference equation given by:

$$\hat{T}_{int}^{i+1} = \hat{T}_{int}^i + \left[ r_p \times r_h^i - \phi \left( \hat{T}_{in}^i - T_{ext}^i \right) \right] \Delta t + \epsilon^i \quad (7)$$

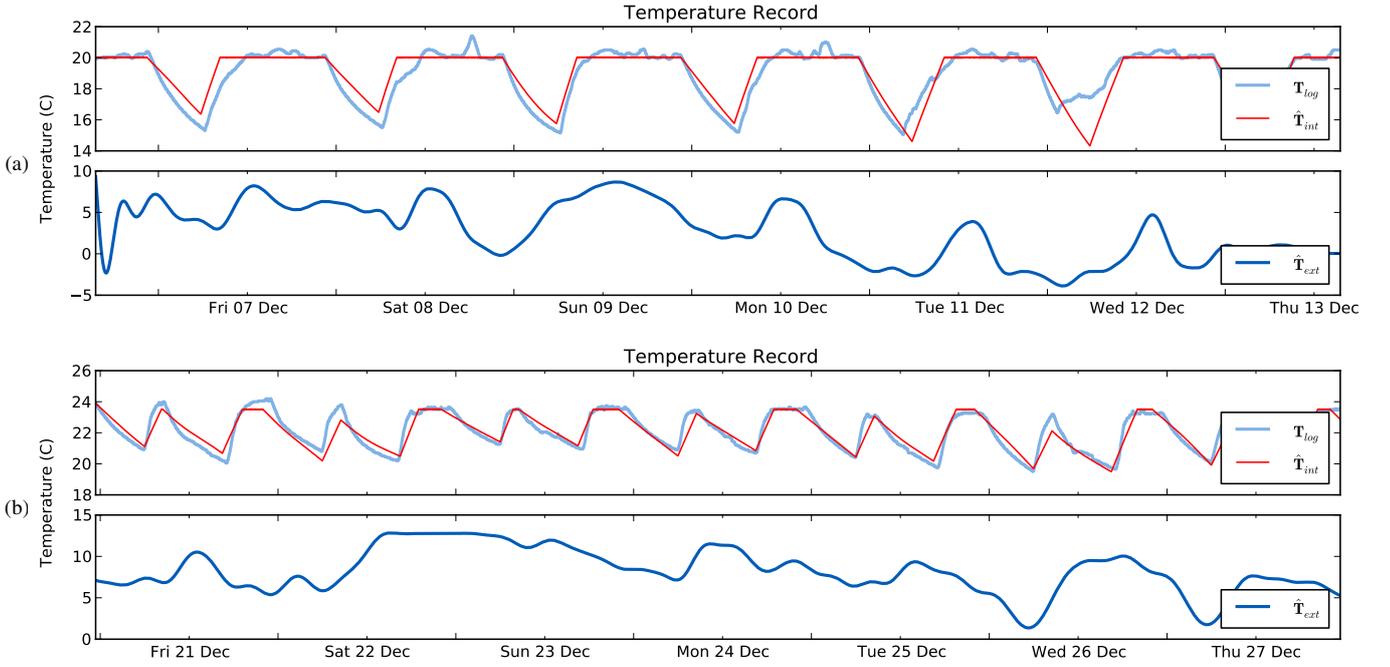
where at time  $t_i$ ,  $\hat{T}_{in}^i$  is the estimated internal temperature (°C),  $\hat{T}_{ext}^i$  is the interpolated external temperatures described above (°C),  $r_h^i$  is a binary variable which indicates whether the heating system is on or off,  $r_p$  is the heater output (°C/hr),  $\phi$  is the leakage rate of the home (1/hr),  $\Delta t$  is the time interval (1/30 hr in this case) and  $\epsilon^i$  is Gaussian noise capturing un-modelled effects.

We model this heating system by assuming that the thermostat has a single set-point throughout the day, and that the boiler is controlled by a separate timer which allows either one or two heating periods per day. This is a common configuration in UK homes with a single zone heating system, and we discuss later how we address data that does not seem to conform to this model.

We perform inference of the thermal properties of the home, and the operational settings of the heating system, by defining the parameter vector,  $\boldsymbol{\theta}$ , given by:

$$\boldsymbol{\theta} = [r_p, \phi, T_{set}, s_1, e_1, s_2, e_2, m] \quad (8)$$

where  $r_p$  and  $\phi$  are the heater output and the leakage rate from the thermal model,  $s_i$  and  $e_i$  are the start and end times of each timer period, and  $m \in \{1, 2\}$  is the operational mode of the heating sys-



**Figure 5: Temperature logger records from (a) a home with one timer period and set-point temperature of 20°C and (b) a home with two timer periods and a set-point temperature of 23.5°C.**

tem. If  $m = 1$ , then just one heating period is applied each day from  $s_1$  until  $e_2$ , and if  $m = 2$ , two heating periods are applied from  $s_1$  until  $e_1$ , and from  $s_2$  until  $e_2$ . We then use these parameters to estimate the internal temperature over the logging period, by initialising our first estimate of the internal temperature to be equal to the first temperature logger measurement,  $\hat{T}_{int}^1 = T_{log}^1$ , and using Algorithm 1 to iteratively propagate the thermal model forward, updating the internal temperature estimate given the thermal performance of the home, and the real-time control policy of the heating system (using on the thermostat setting, the internal temperature and the timer setting to determine  $r_h^i$  in each time period). The optimal parameter vector,  $\theta^*$ , is that which minimises the squared error between the estimated internal temperature measurements, and those actually recorded by the logger, and is given by:

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^n (\hat{T}_{int}^i - T_{log}^i)^2 \quad (9)$$

This objective function is convex and quadratic with added constraints that,  $r_h^i \in \{0, 1\}$  and  $m \in \{1, 2\}$ , and thus, we solve this constrained convex optimisation problem using the *interior-point algorithm* with conjugate gradient steps [2].

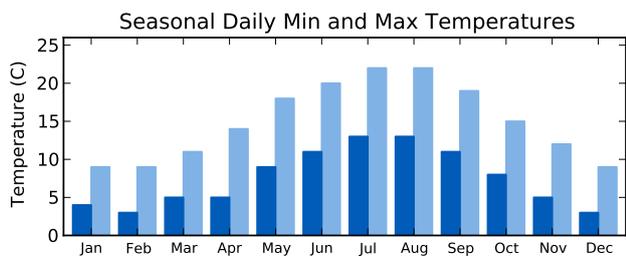
Figure 5 shows this process applied to two example datasets collected within the 2013 live trial. Figure 5a shows a home with one timer period and set-point temperature of 20°C and Figure 5(b) shows a home with two timer periods and a set-point temperature of 23.5°C. The figures show the internal temperature measurements made by the USB temperature logger placed on the homes' thermostats (light blue line), the external temperature taken from the internet (dark blue line), and the best fit thermal model derived from the parameters values found in equation 8 (red line). Note that the fit is generally very good, with the only significant deviation occurring in the first time series on the night of 11th December due

to additional un-modelled heating from the anti-frost mechanism switching on the boiler as the external temperature drops below 0°C, which causes the leakage rate to be slightly underestimated overall.

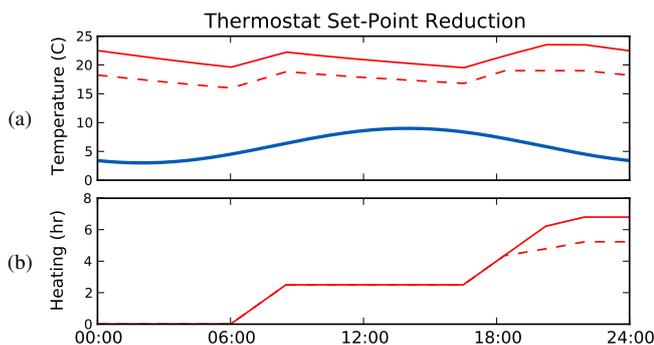
### 4.3 Calculating the Impact of Intervention

Having inferred the thermal properties of the home, and the operational settings of the heating system, we can calculate the impact that any intervention will have. In the 2013 live deployment of MyJoulo, we focused on the thermostat set-point value, and provided feedback on the savings that would be made if it were reduced to a recommended value of 19°C. In order to do so, however, we can not simply consider the energy consumption over the week during which the temperature logger was deployed, since any particular week may exhibit unseasonably low or high temperatures. Thus, we use monthly daily maximum and minimum temperatures, for the postal address that the logger was sent to (again collected automatically from [www.worldweatheronline.com](http://www.worldweatheronline.com)), to calculate the expected external temperature profile over the course of any particular day (using a sinusoid with a minimum temperature at 02:00 and a maximum temperature at 14:00). Having calculated the expected external temperature, we can then use it for  $\hat{T}_{ext}$  in Algorithm 1 to calculate the resulting internal temperature profile on any given day for the optimal parameters found in the last section; iterating until  $\hat{T}_{int}^1 = \hat{T}_{int}^{721}$  since we do not know the initial temperature. For any such day, we assume that the total energy is proportional to total time over which the heating system was applying heat, and this can be found directly by calculating,  $\sum_{i=1}^{720} r_h^i$  (where 720 is the number of 2 minute intervals in the day). To calculate the saving, we repeat the process using the reduced thermostat set-point within Algorithm 1, and calculate the percentage reduction that results.

Figure 6 shows the seasonal daily minimum and maximum tem-



**Figure 6: Seasonal daily minimum and maximum temperatures for the location of the dataset show in Figure 5b.**



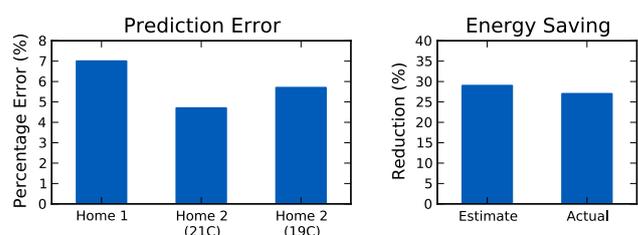
**Figure 7: Internal temperature, and cumulative heat input, for the dataset show in Figure 5b when the thermostat set-point is reduced from 23.5°C (solid) to 19°C (dashed).**

peratures for the location for the dataset shown in Figure 5b, and Figure 7 shows the result of using Algorithm 1 to calculate the corresponding internal temperature for an example day in December, when the thermostat set-point is set at the original 23.5°C, and at the reduced setting of 19°C. Figure 7a shows the internal and external temperatures, and Figure 7b shows the cumulative period over which the heating system was delivering heat in both cases. Note that when the thermostat set-point is reduced, the internal temperature exhibits a similar reduction in temperature (being approximately 4.5°C lower everywhere). However, the temperature profile is not completely unchanged. In this example, during the first timer period (from 06:00 to 08:30), the internal temperature does not reach the thermostat set-point in either case, and thus the cumulative heater on time is 2.5 hrs over this period for both thermostat settings. However, in the second timer period (from 16:30 to 22:00), the internal temperature reaches the reduced set-point temperature more rapidly, and then acts to maintain the temperature at this point by alternative periods of heating and not heating. This results in an overall reduction in the total heat input into the home, and a resulting reduction in heating costs.

Performing this calculation over a 7 month heating period from 1st October to 30th April, yields an energy reduction in the case of the dataset shown in Figure 5a when the thermostat is reduced from 20°C to 19°C, and a 27% saving for the dataset shown in Figure 5b when the thermostat is reduced from 23.5°C to 19°C. These figures conform to the often quoted benchmark in the UK that reducing the thermostat by 1°C results in a 10% reduction in heating costs.

#### 4.4 Verifying Energy Savings

Whilst we can not verify the predicted savings, as calculated above,



**Figure 8: Average prediction error in boiler on times in Home 1 and Home 2, and estimated and actual boiler on time reduction yielded in Home 2.**

directly since any individual heating season may deviate from the statistical average, we can verify the accuracy of our thermal model in predicting the boiler on time in any individual home. To do so we instrumented two homes owned by the University of Southampton, to collect live internal temperature, thermostat set point, and the actual boiler firing state at 1 minute intervals. In addition, we collected external air temperature data from a deployed temperature logger. Then across a two week period, one house (Home 1) maintained a constant thermostat set point of 21°C, while the other (Home 2) maintained a set point of 21°C for the first week, and then reduced it to 19°C for the second week (the exact values decided through negotiation with the residents of the homes to balance the aims of the experiment against their comfort preferences).

Fitting the thermal model exactly as described above across the first week in each home yields the average error in the predicted boiler on times as shown in Figure 8. Note that the average error in Home 1 where the set point was held constant is 7%, and in Home 2 the average errors are 4.7% and 5.7% respectively, corresponding to the periods when the set point is first maintained at 21°C and then reduced to 19°C. These results indicate a very good fit between our simple thermal model and reality. We then predict the average percentage reduction in boiler on time (assuming that this will be proportional to heating costs) in the second week in Home 2 from reducing the set point temperature by 2°C. However, here rather than using seasonal average temperatures, we use the actual external temperature that was observed. Figure 8 shows the results of this comparison where the predicted reduction was 29%, and the actual realised reduction was 27%, indicating the very good predictive properties of the relatively simple thermal model<sup>1</sup>.

#### 4.5 More Complex Temperature Profiles

While the approach described above works when the internal temperature exhibits regular heating patterns, more complex temperature profiles can arise when there are additional heating sources within the home (such as wood burning stoves), when a more sophisticated programmable timer is used with many changes of set-point, or when the thermostat is frequently manually changed. These are all frequent enough occurrences that any deployed system must be able to detect and handle them. To do the former, we can simply observe the deviation between the modelled internal temperature and that measured by the temperature logger, as calculated in Equation 9, and use it to indicate when the existing model does a poor job of describing the temperature measurements made by the logger. To do the later, we must both detect the set-point temperature

<sup>1</sup>Note that the saving here is larger than that in Figure 5 since this was a relatively warm week, and thus the percentage change in the difference between the internal and external temperatures was greater than average.



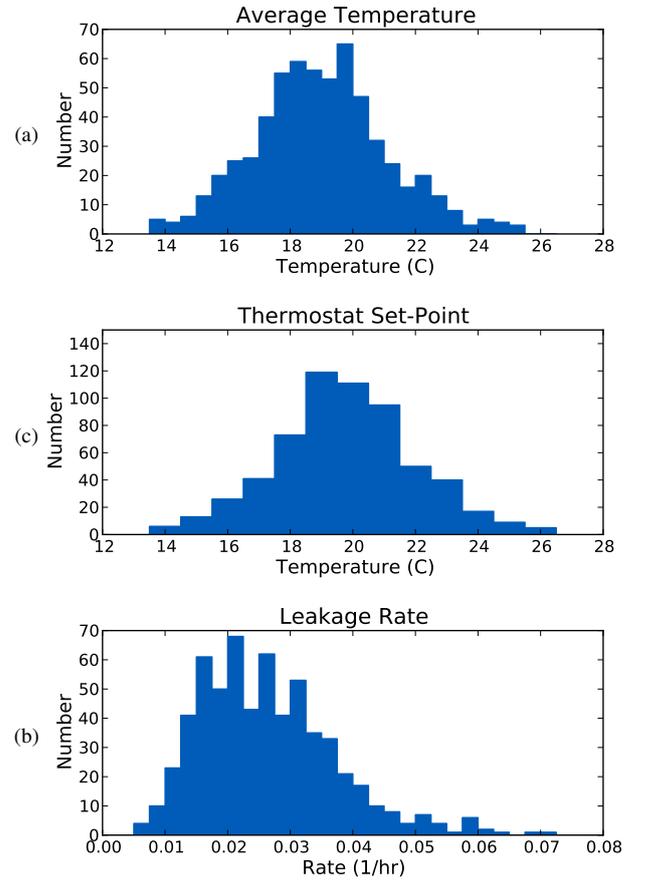
**Figure 9: Example analysis page presented to the user of the MyJoulo website showing the 9% saving, for the dataset shown in Figure 5a, when the thermostat set-point is reduced from 20°C to 19°C.**

and model the impact of the intervention. Thus, we use a statistical approach and take the mode of the upper 50% percentile of the temperature logger measurements as a representative set-point temperature,  $\hat{T}_{set}$ . Then, rather than using a parameterised model of the operational settings of the heating system to calculate the internal temperature profile and the total heat input, we rearrange Equation 7, and sum over the temperature record, to calculate the total energy input over the course of a week:

$$(T_{log}^n - T_{log}^1) + \phi \times \Delta t \times \sum_{i=1}^n (\hat{T}_{log}^i - \hat{T}_{ext}^i) \quad (10)$$

where  $\hat{T}_{ext}^i$  is the external temperature calculated as above for any particular week using the seasonal daily minimum and maximum temperatures, and  $\phi$  is the leakage rate of the home, for which we use a standard value of 0.3 1/hr. Performing this calculation for one week for each month between 1st October to 30th April, and scaling appropriately by the number of days in the month, allows us to estimate the total heat input over a heating season.

We then use the insights of the last section to calculate how the heat input would be reduced if the thermostat set-point were reduced. To do so, we subtract from  $T_{log}$  the difference between  $\hat{T}_{set}$  and the recommended value of 19°C, such that the reduced internal temperature is given by  $T_{log}^i - (\hat{T}_{set} - 19)$ . For example, if we estimate the set-point to be 20°C, we reduce all the mea-



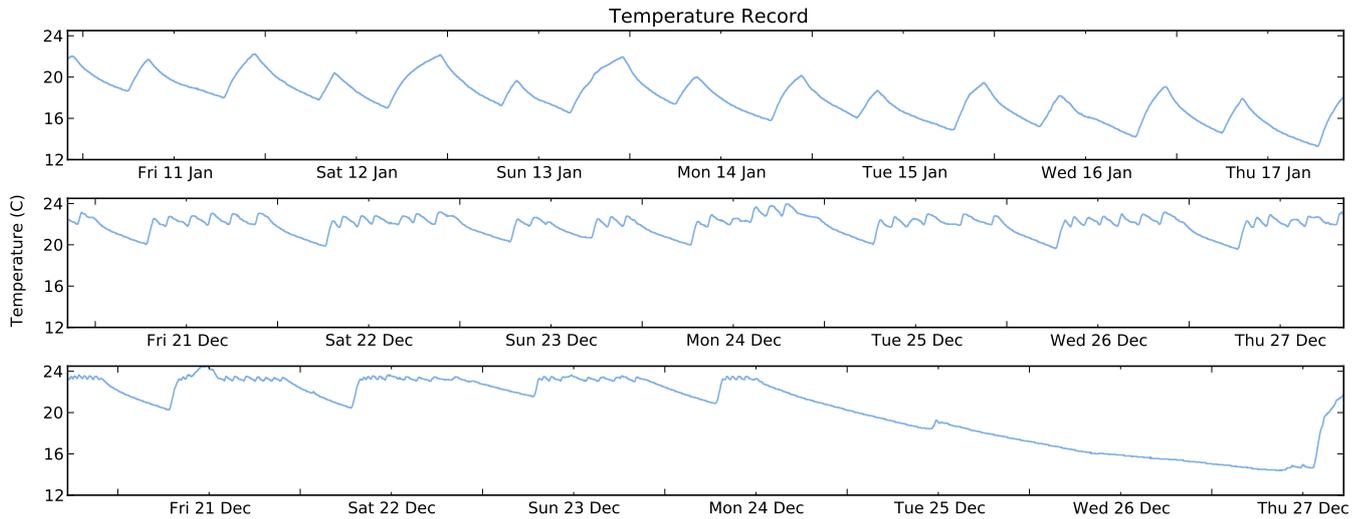
**Figure 10: Aggregate analysis of the collected dataset.**

sured logger temperatures by 1°C. Repeating the calculation over the same time period as above allows us to calculate the percentage savings in heating costs that will result.

Using this approach implicitly assumes that the leakage rate observed in  $T_{log}$  does not depend upon the difference between the external and internal temperatures. We know that this is not the case. However, if the difference between  $\hat{T}_{ext}$  and typical seasonal values is small, and  $\hat{T}_{set}$  is close to 19°C, then the error is small. Indeed, for the two example datasets, this alternative calculation yields set-points of 20°C and 23°C, and reductions of 9% and 30%, compared to 9% and 27% as calculated above, indicating that it provides reasonable agreement. More problematic, is that this approach precludes us providing more sophisticated feedback on operational changes (such as timer settings).

## 4.6 Providing Feedback

We provide feedback to MyJoulo users through a visually appealing infographic (shown in Figure 9) that shows a comparison of the user's thermostat set-point against a typical range, a comparison of the minimum, average and maximum temperature recorded by the logger, the temperature profile recorded by the logger, and the saving that would result from reducing the set-point to the recommended value of 19°C. Taking inspiration from use of social norms in energy feedback to prevent users who are already under-consuming from reverting to the norm, we also inform users whose set-point is already below 19°C how much they have already saved, and inform them how much they would save from a further reduc-



**Figure 11: Example temperature profiles showing a home that fails to reach the set point (top), a mechanical thermostat with a wide hysteresis band (middle), and a week including days when the heating was turned off (bottom).**

tion of  $1^{\circ}\text{C}$  [9]. Note that since we do not know the size of the home, we can not calculate the actual monetary cost saving automatically. However, we provide the user with the option of entering their annual heating bill, and on doing so, convert the percentage reductions to monetary savings. Given an average UK heating bill of  $\pounds 1000$ , this results in significant annual savings of  $\pounds 90$  and  $\pounds 270$  for the two example datasets considered here.

## 5. DEPLOYMENT CHALLENGES

The system as described above was launched in beta form in December 2012 in order to test the effectiveness of the methodology employed. In its three months of operation, 750 users registered and were sent USB temperature loggers, and 600 completed the process of recording temperature for a week and uploading the data to the website. Figure 10 shows the distributions across of this dataset of three key measures of each home: the average temperatures across the week, the identified thermostat set-point, and the thermal leakage rate,  $\phi$ . Note that the later is calculated even for homes where the full thermal model, including timer operational settings, do not fit well, by identifying the long periods of decreasing temperature that often occur at the end of each day when the heating is turned off. Taking each in order, the average temperature plot shows a clear peak around  $18\text{--}20^{\circ}\text{C}$ , with a significant spread of both cold and very warm homes. The thermostat set point distribution shows a peak at  $19^{\circ}\text{C}$ , the temperature that we chose as the reference point for providing feedback to the householders, and again shows a number of homes with high settings. Finally, the leakage rate shows a cluster between  $0.01$  and  $0.04$   $1/\text{hr}$ , and a significant number of outliers with high leakage rates.

While our peer comparison to date has focused on the thermostat set point, in future work we plan to exploit these additional comparisons to identify particular classes of homes. For example, a home with low average temperature, a normal thermostat setting and a high leakage rate, may indicate a home which is poorly insulated, where the installation of roof or cavity wall insulation would yield significant energy savings, and comfort improvements.

Further information and feedback could also be generated by more complex analysis of the temperature profiles. For example, Figure 11 shows three examples of common temperature profiles

that challenge our existing analysis. The first is an example of a case where the heating system is undersized, and the home never reaches the thermostat set point. In this case, there is no single best fit thermostat set point derived from the thermal model fitting, but rather, the model will indicate that above a certain value, all thermostat set points fit the data equally well. This is important, since in this case, the household heating bill is determined by the timer settings, and a reduction of thermostat set point may not yield any saving at all.

The second example shows the ability to classify the type of thermostat installed within the home with this plot showing the temperature profile of a mechanical thermostat with large hysteresis band which gives relatively poor temperature control, and drifts over the course of the day slightly, compared to the two examples of electronic thermostats with time proportional integral (TPI) control shown in Figure 5.

Finally, the third example shows an example of a week where the heating was switched off for two days (in this case over Christmas Day and Boxing Day). While such periods can be addressed by including a further 7 binary parameters indicating whether the heating is on or off each day, doing so fails to accurately capture this long-term temperature decay, which rather than asymptoting to the external air temperature, settles at an intermediate temperature of the envelope or structure of the home. Adding this additional complexity to fully model this decay, is something we intend to address in future work, since the decay rate here is highly indicative of the thermal properties of the home [1].

## 6. CONCLUSIONS

In this paper, we presented a deployed prototype of a scalable low-cost solution to provide personalised home heating advice to households, and discussed the results of an initial deployment of the system that saw 750 users order a temperature logger, and 600 of these users upload the logged dataset the MyJoulo website. As discussed above, our ongoing future work is aimed at broadening the analysis provided to users. In particular, we would like to provide energy saving analysis and recommendations on the timing of heater use and also the thermal leakage rate of the home itself; identifying homes which are particularly leaky where adding insulation will

be most effective. In order to do so it is necessary to broaden the range of heating profiles for which we can use the model-based approach, recognising and handling cases where the programmable thermostats are used or where householders make frequent manual changes to their thermostat and timer settings.

## 7. ACKNOWLEDGMENTS

The work presented in this paper is supported by the EPSRC funded ORCHID ([www.orchid.ac.uk](http://www.orchid.ac.uk)) and Intelligent Agents for Home Energy Management ([www.homeenergyagents.info](http://www.homeenergyagents.info)) projects.

## 8. REFERENCES

- [1] P. Bacher and H. Madsen. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7):1511–1522, 2011.
- [2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Mar. 2004.
- [3] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.
- [4] Kempton. Two theories of home heat control. *Cognitive Science*, 10(1):75–90, 1986.
- [5] J. Z. Kolter and T. Jaakkola. Approximate inference in additive factorial hmms with application to energy disaggregation. *Journal of Machine Learning Research - Proceedings Track*, pages 1472–1482, 2012.
- [6] O. Parson, S. Ghosh, M. Weal, and A. Rogers. Non-intrusive load monitoring using prior models of general appliance types. In *Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*, pages 356–362, Toronto, Canada, 2012.
- [7] C. E. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [8] A. Rogers, S. Maleki, S. Ghosh, and J. Nicholas R. Adaptive home heating control through gaussian process prediction and mathematical programming. In *Second International Workshop on Agent Technology for Energy Systems (ATES 2011)*, pages 71–78, May 2011.
- [9] P. W. Schultz, J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius. The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5):429 – 434, 2007.
- [10] J. Scott, A. Bernheim Brush, J. Krumm, B. Meyers, M. Hazas, S. Hodges, and N. Villar. Preheat: controlling home heating using occupancy prediction. In *Proceedings of the 13th international conference on Ubiquitous computing, UbiComp '11*, pages 281–290, New York, NY, USA, 2011. ACM.