Time Series Analysis in Home Energy Monitoring

With potential applications to Google PowerMeter

Joe Kwiatkowski, 9^{th} May 2010

I discuss simple time series analysis as applied to energy data collected over the course of a month with a cheap, homemade, meter described in detail at joejk.com/conquan. I demonstrate the disaggregation of the oven from the total load, how models can be used to forecast consumption inform energy efficiency measures.

1 Non-intrusive load disaggregation

Figure 1 shows the spectral density collected from four different three hour windows in the data. The peak at 0.55 min^{-1} is uniquely caused by thermostatic cycle of the oven. As a result, it is trivial to find the times at which the oven is on (figure 2). Furthermore, the total oven load can be calculated by assuming:

- the oven has only two states, *on* and *off*.
- the fluctuations in other, simultaneous, loads are small.

The oven load is given in table 1.

Disaggregation, essentially a problem of multiple classification, is likely to be harder for other loads. However, other large loads should be identifiable by increasing the dimensionality of the analysis. For example, an air-conditioning load could potentially identified from its position in a multi-dimensional space spanned by:

- outdoor temperature.
- time of day, day of week, season of year.
- school in session?
- Δpower_{on}, Δpower_{off}.
- frequency of load cycle.
- $\cdots$

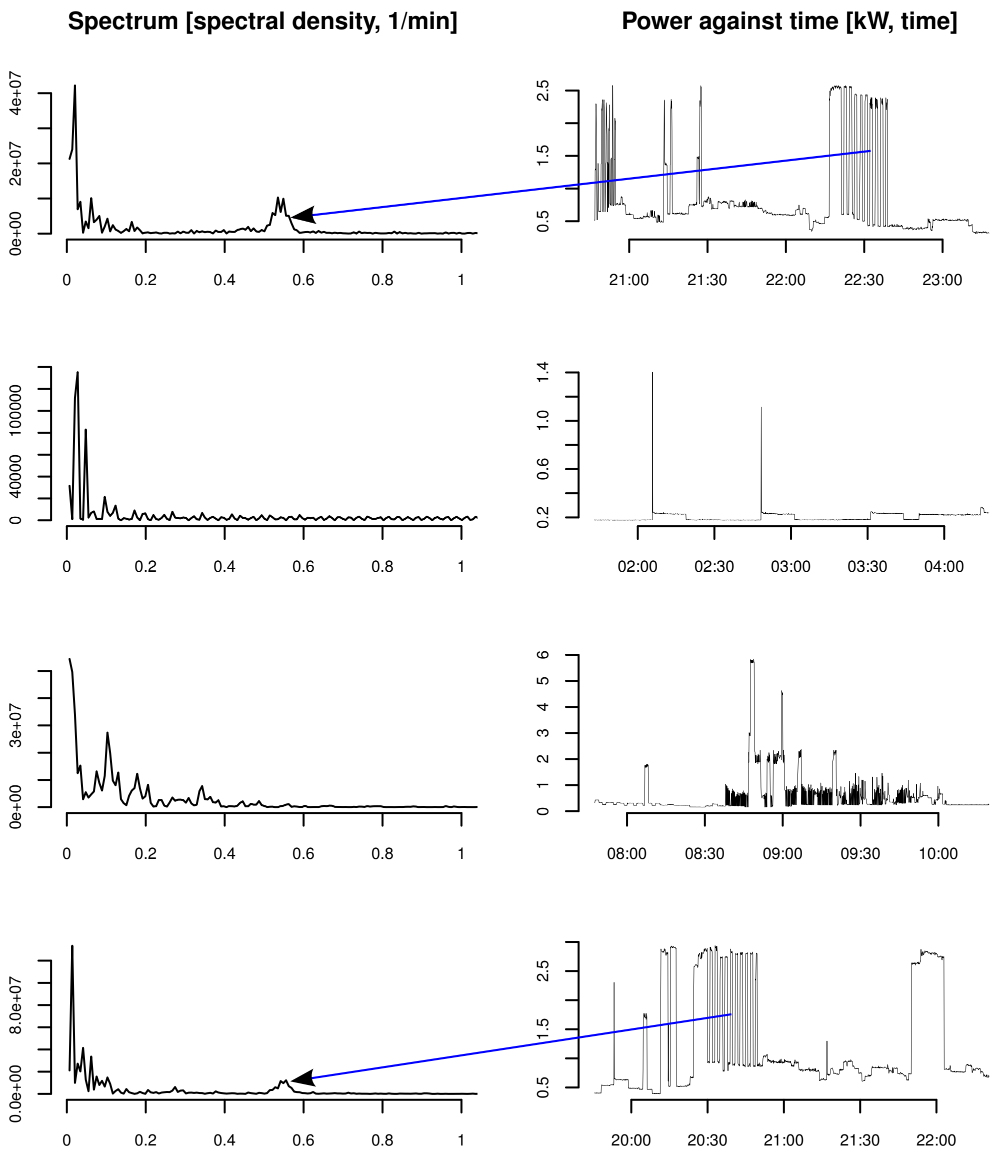

Figure 1: *Four different 3 hour windows in the power time series, shown in the frequency and time domain. The peak at 0.55 min^{-1} in the spectral density is uniquely related to the oven load.*

Load	Monthly consumption [kWh]	Percentage of total consumption
Oven	**15.5**	**4.8**
'Always On'	**82.3**	**25.6**

Table 1: *Oven consumption, relative to the 'Always On' load and total load.*

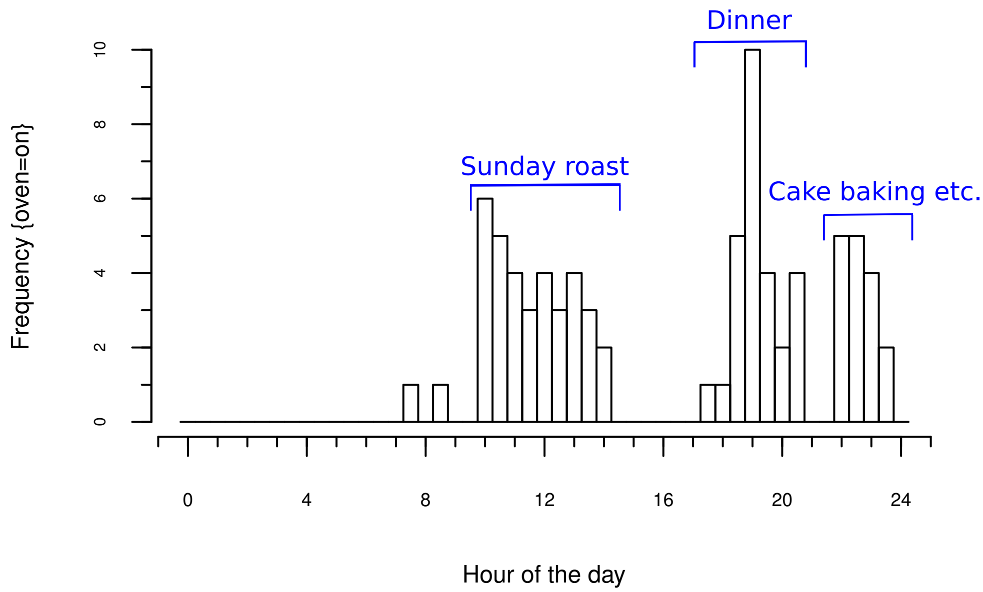

Figure 2: *The frequency of oven use against time of the day.*

2 Modeling and forecasting power consumption

A general time series can be decomposed into four components: the 'AlwaysOn', trend component(s), periodic component(s), and random fluctuations (figure 3).

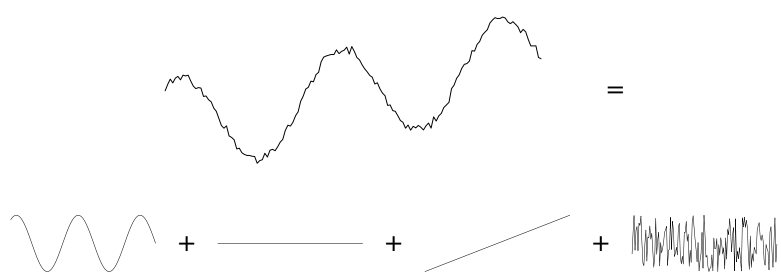

Figure 3: *The decomposition of a time series into a periodic component, an 'always on' component, a trend component, and random fluctuations.*

Intuitively, one might expect the periodicity of these data to be days, weeks, or a combination of both. These alternatives can be tested by looking at the autocorrelation within data that has been corrected for their periodicity. For example, given a dataset d of hourly consumption, the daily and weekly differenced datasets are:

$$\begin{aligned} D^i_{daily} &= d_{i+24} - d_i \\ D^i_{weekly} &= d_{i+168} - d_i \end{aligned} \tag{1}$$

Figure 4 shows the partial autocorrelation functions for D_{daily} and D_{weekly} respectively. Whereas data remains autocorrelated under the assumption of daily periods (particularly at 24 hour lags), weekly periods remove almost all autocorrelation beyond lag = 1 hour.

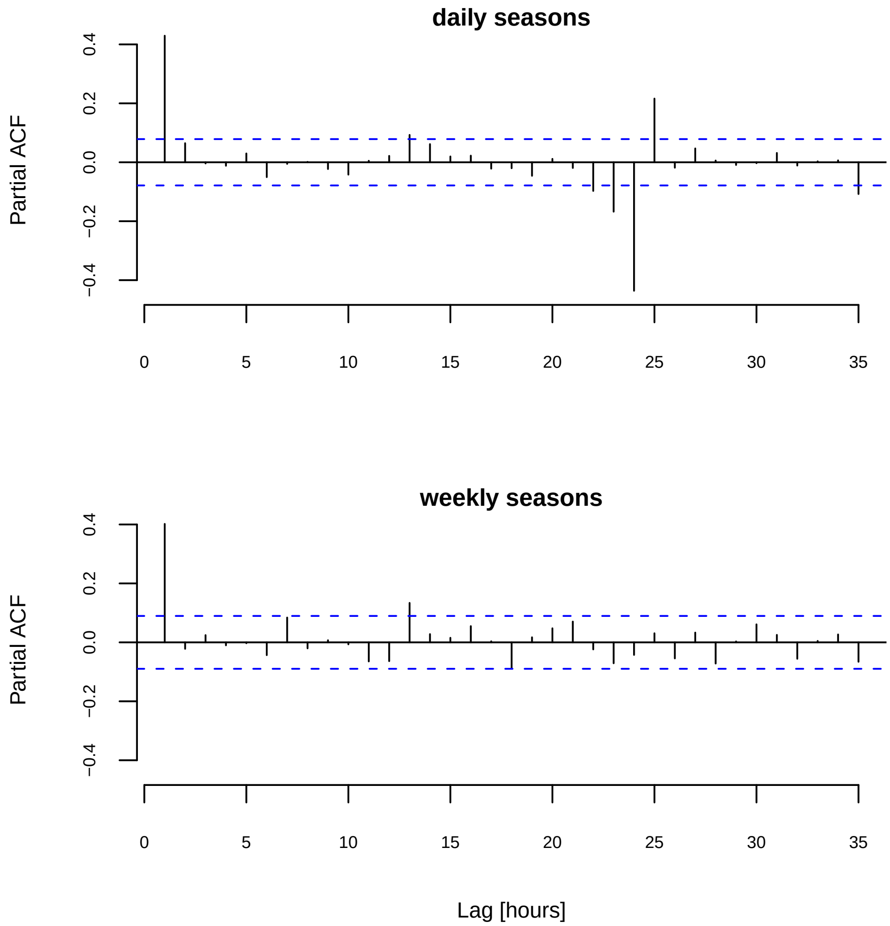

Figure 4: *Partial autocorrelation functions for hourly aggregated data, assuming daily and weekly periods. Extensive autocorrelation remains when daily periods are assumed, but not when weekly periods are assumed. The dashed lines show 95 % confidence limits.*

Furthermore, this latter chart suggests that there is no trend in the series, and that the data is best modeled as a first order autoregressive series where the power at time t is composed of an AlwaysOn component, an average load of that hour within the week, some positive correlation to the previous hour's consumption, and a random fluctuation:

$$\text{Power}_t = \text{AlwaysOn} \; + \; \text{HourWithinWeek}_{t\,\%\,168} \; + \; \phi \cdot \text{Power}_{t-1} \; + \; \epsilon_t \tag{2}$$

This equation can be used to forecast power consumption (figure 5).

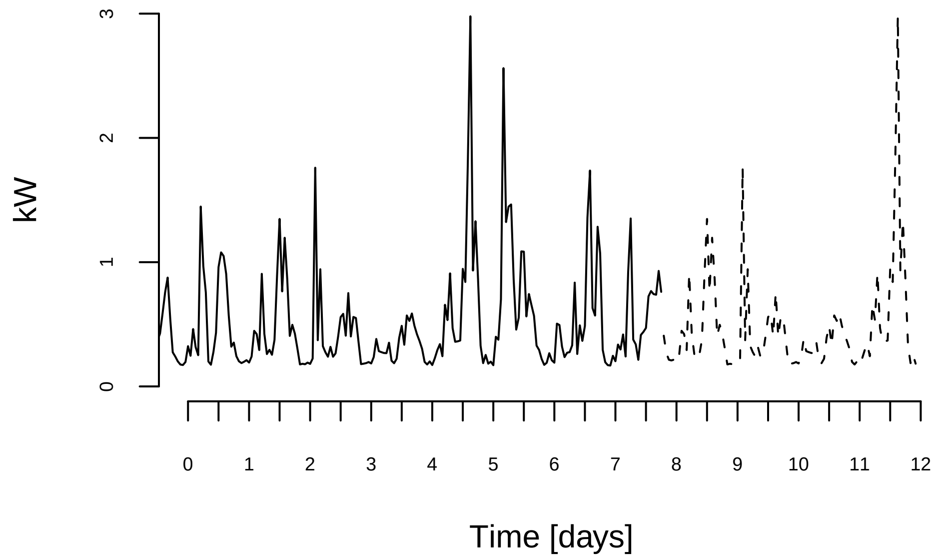

Figure 5: *Historical and forecast data. Forecasts are made by fitting equation 2 to historical data.*

3 Informing demand reduction strategies

Given a forecast of future energy consumption, energy efficiency strategies can be suggested. One way to do this is in terms of the components of the forecasted data: always on, periodic, and trends (table 2). Additionally, spurious events could be found from exceptionally high model residuals. These strategies could potentially be optimised by wrapping this analysis in a Monte Carlo code.

Component	**Possible causes**	$\mathbf{E_{tot} \rightarrow E_{tot}^{95\%}}$
'Always on'	Computers, transformers, ...	19.5 % reduction
Periodic	Varying usage with day of week.	Reduce weekend load by 15.7 %
Trends	New purchases, replacing small devices with large, ...	No tend identified
Spurious events	Large thermal loads left on, ...	No spurious events identified

Table 2: *Various components of the forecast load, possible causes, and potential strategies to reduce total energy consumption by 5 %.*

‘Seasonal subseries’ plots might be helpful to inform efficiency measures (figure 6). This format could be further combined with analyses discussed above. For example, the difference between consumption on Friday and Sunday is largely explained by the additional 2.9 kWh of oven use on the latter.

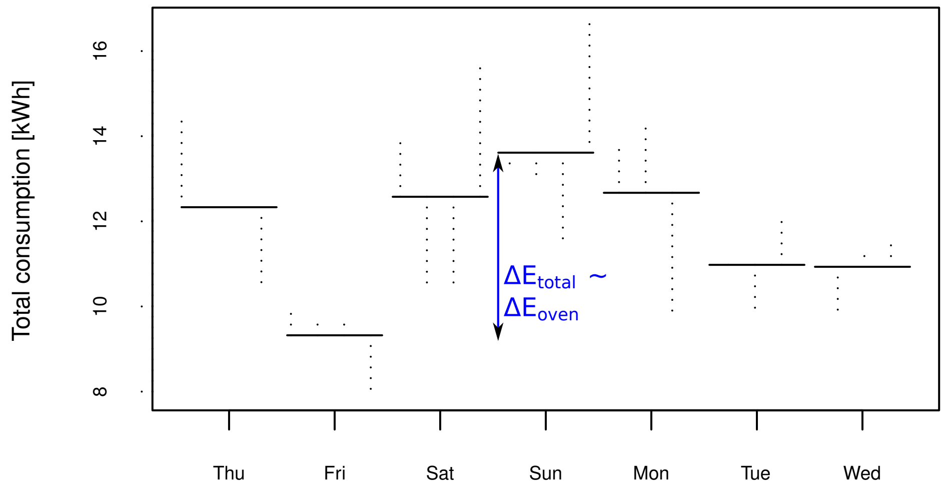

Figure 6: *A ‘seasonal subseries’ plot showing the consumption throughout the week, averaged over a month (horizontal bars), and for individual days (vertical lines). Large variance (e.g. Saturdays) may indicate opportunities for savings based on behavioural changes. The difference in oven usage largely explains the difference between total consumption on Friday and Sunday.*