SUBMETERING DATA

M. Martinez

Image credit: Landis+Gyr

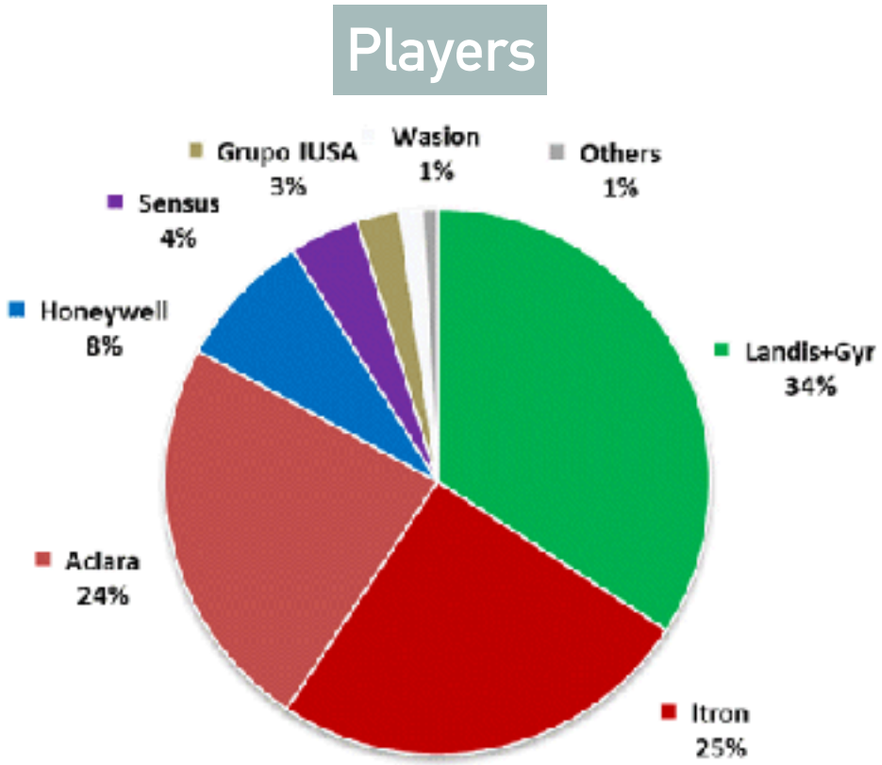
OUTLINE

- Definitions
- Objectives
- The data
- Recommendations

DEFINITIONS

Electrical Submeter are electrical devices for energy management. They measure and track building energy performance. Moreover, register and communicate the amount of electrical consumption.

- Allocate electrical usage costs among tenants.
- Allocate electrical usage costs among departments.
- Allow analyze power quality issues
- Detect maintenance issues based on energy consumption patterns.



Create an environment that encourages energy conservation and to improve energy reliability

OBJECTIVES

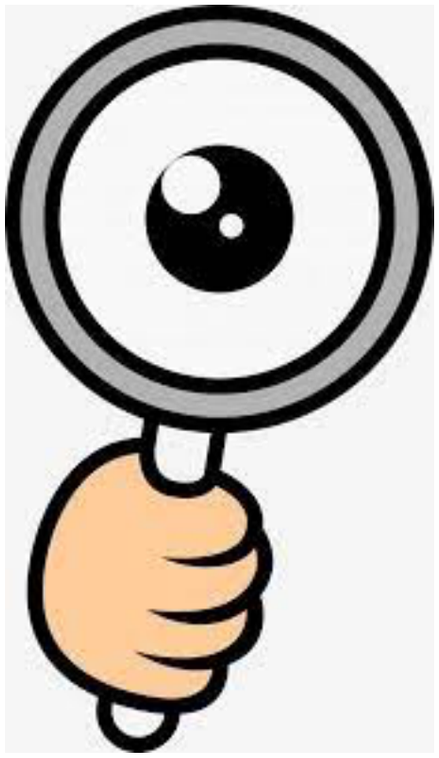
- Show evidence that support the advantages of sub-meters usage.
- Predict future energy consumption.

THE DATA

- The data provided contains 2,075,259 measurements gathered in a house located in Sceaux (7km of Paris, France) between December 2006 November 2010 (47 months).

Data inspection reveals that:

- In 2006 the data were obtained from December 16 to 31.
- In 2010 the data were obtained from January 01 to November 26.
- From 2007 to 2009 the data were gathered from January 01 to December 31.



THE DATA

The data contains:

- **Date**: YYYY-MM-DD
- **Time**: hh:mm:ss
- **Global_active_power** (in kilowatt-min): household net-transferred energy
- **Global_reactive_power** (in kilowatt-min): household NO net-transferred energy
- **Global_intensity** (in ampere-min): household magnitude of power (current) consumed.
- **Voltage** (in volt/min): the “pressure” that pushes electricity, higher voltages cause more electricity flow to an electronic device.
- **Sub_metering_1** (watt-hour): active energy (dishwasher, oven and microwave) ——> Kitchen
- **Sub_metering_2** (watt-hour): active energy (washing-machine, tumble-drier, refrigerator and light) ——> Laundry Room
- **Sub_metering_3** (watt-hour): active energy (electric water-heater and air-conditioner) ——> Water Heater & AC

GENERAL DATA EXPLORATION

- Active energy (watt-hour)

```
Sub_metering_1
Min.    : 0.000
1st Qu.: 0.000
Median : 0.000
Mean   : 1.122
3rd Qu.: 0.000
Max.    :88.000
```

```
Sub_metering_2
Min.    : 0.000
1st Qu.: 0.000
Median : 0.000
Mean   : 1.299
3rd Qu.: 1.000
Max.    :80.000
```

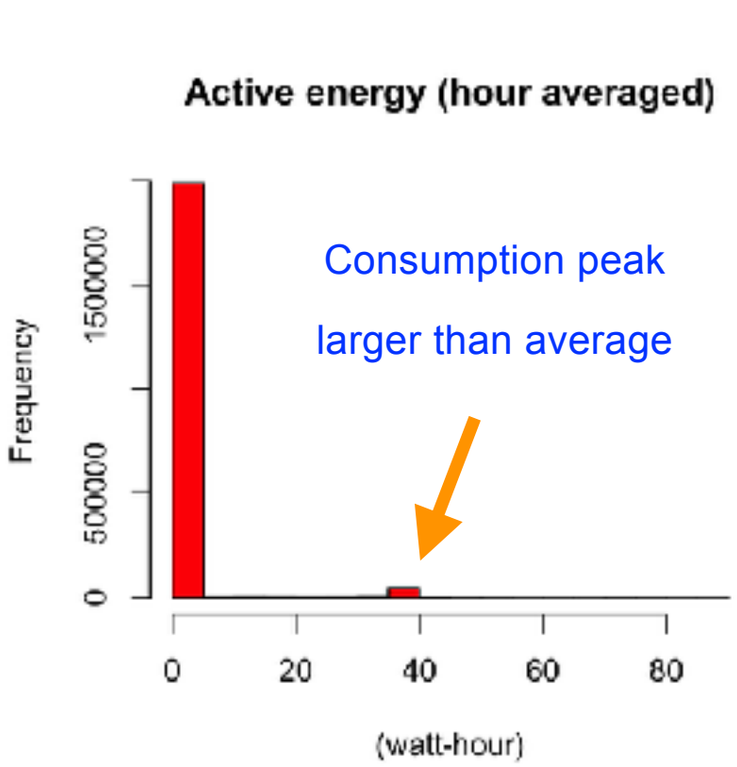
```
Sub_metering_3
Min.    : 0.000
1st Qu.: 0.000
Median : 1.000
Mean   : 6.458
3rd Qu.:17.000
Max.    :31.000
```

Sub-metering 1 is using the most energy, while sub-metering 3 the least.

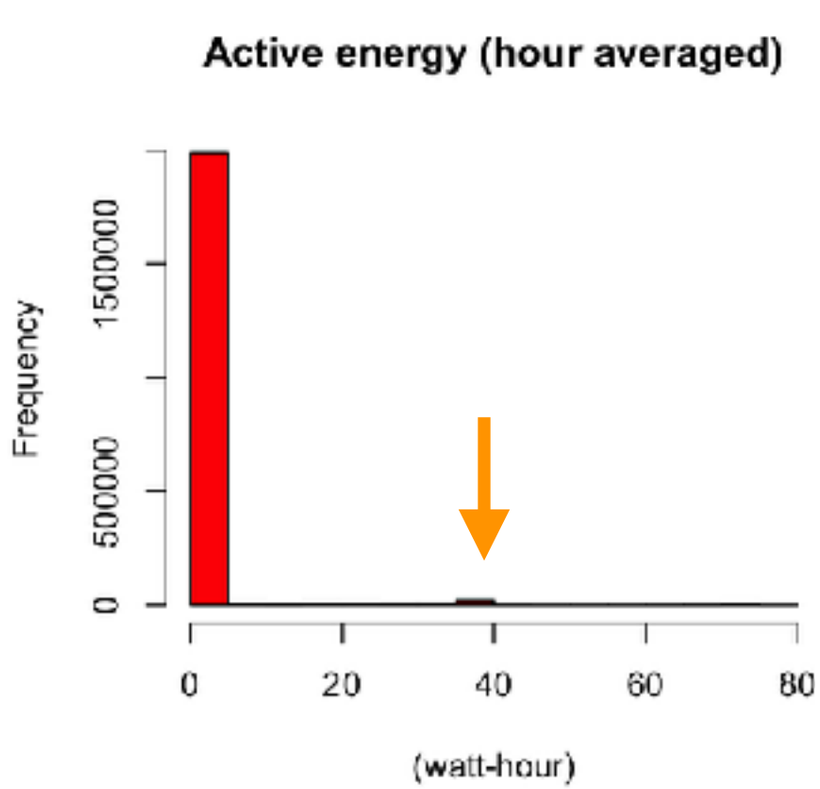
➤ What predicts the model? will be the same for the next two years?

GENERAL DATA EXPLORATION

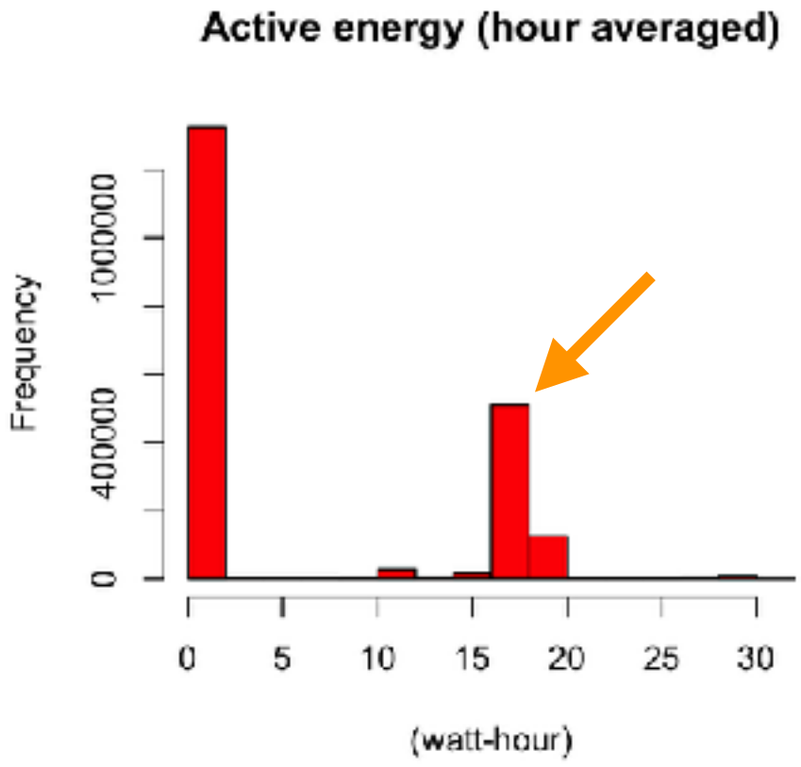
Sub-metering 1



Sub-metering 2



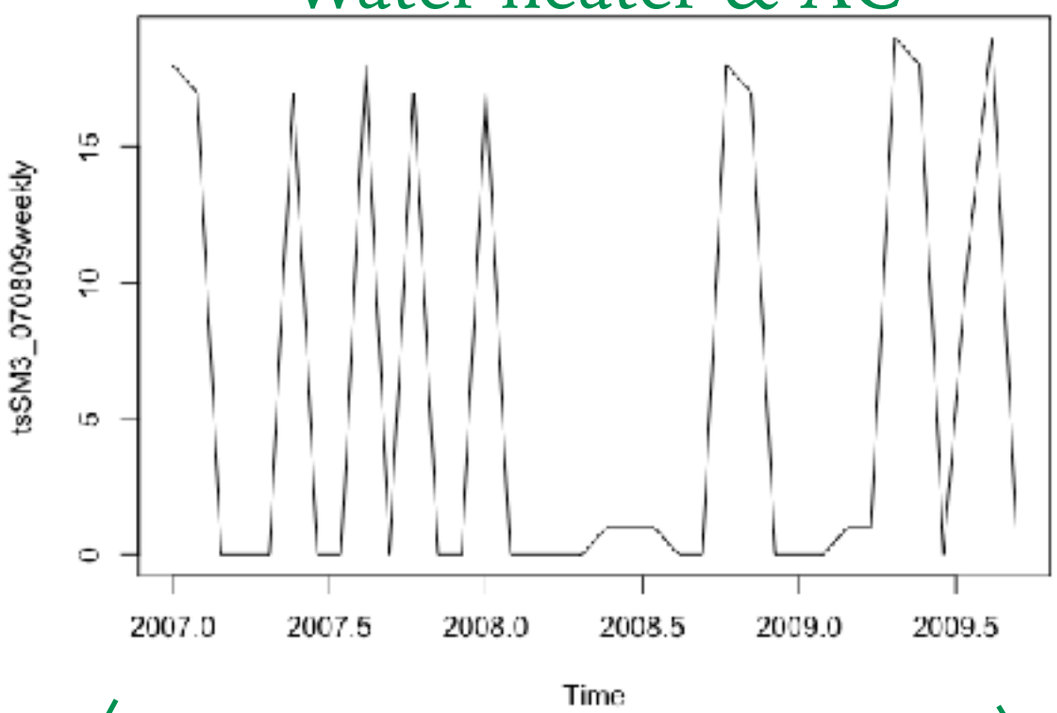
Sub-metering 3



- What is producing that peaks?

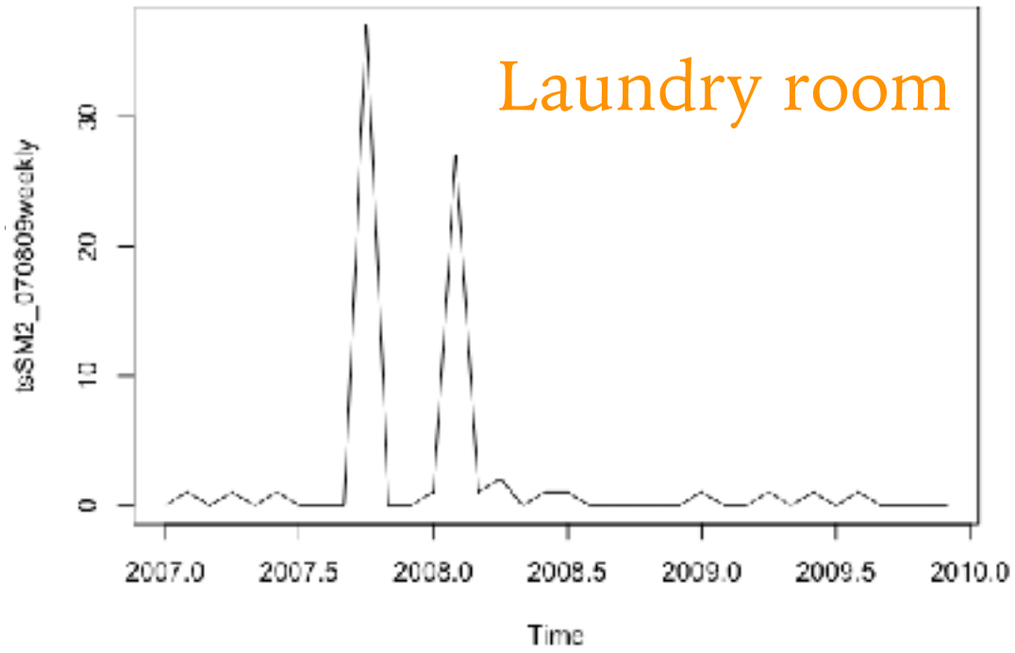
ANALYSIS–TIME SERIES

Water heater & AC

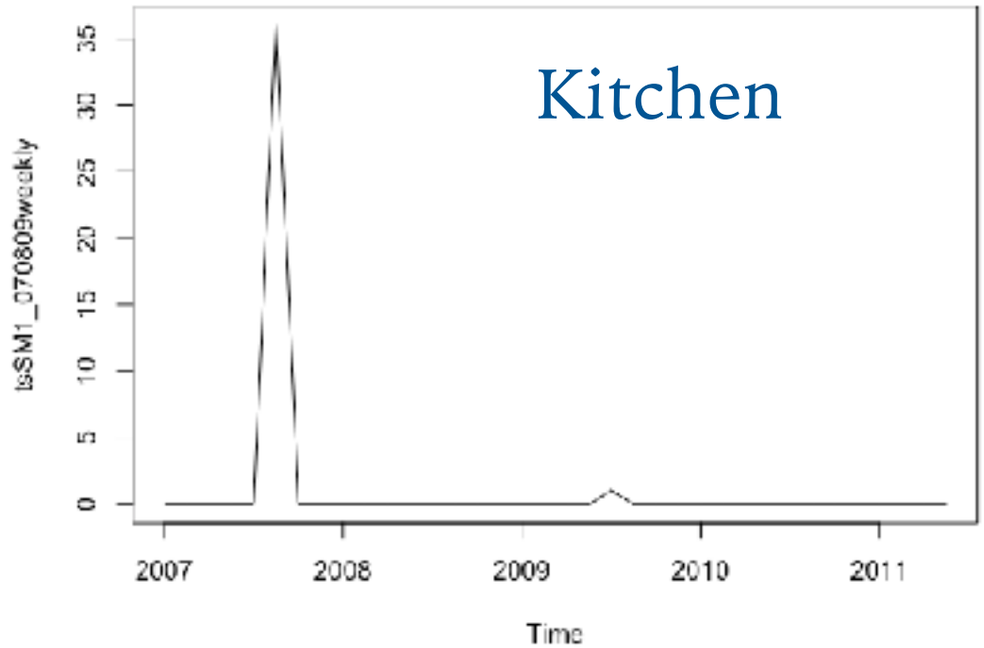


➤ There is a **constant** energy consumption (~6 watts-hour) from the water heater and AC.

Laundry room



Kitchen



➤ In general, there is a **low** energy consumption (~1 watts-hour) from the laundry room and kitchen, except for few months during the year around the summer.

ANALYSIS-FITTING

A linear model ***tslm*** is applied to the time series, including trend and season components.

```
#Fitting
fitSM3 <- tslm(tsSM3_070809weekly ~ trend + season)
summary(fitSM3)

fitSM2 <- tslm(tsSM2_070809weekly ~ trend + season)
summary(fitSM2)

fitSM1 <- tslm(tsSM1_070809weekly ~ trend + season)
summary(fitSM1)
```

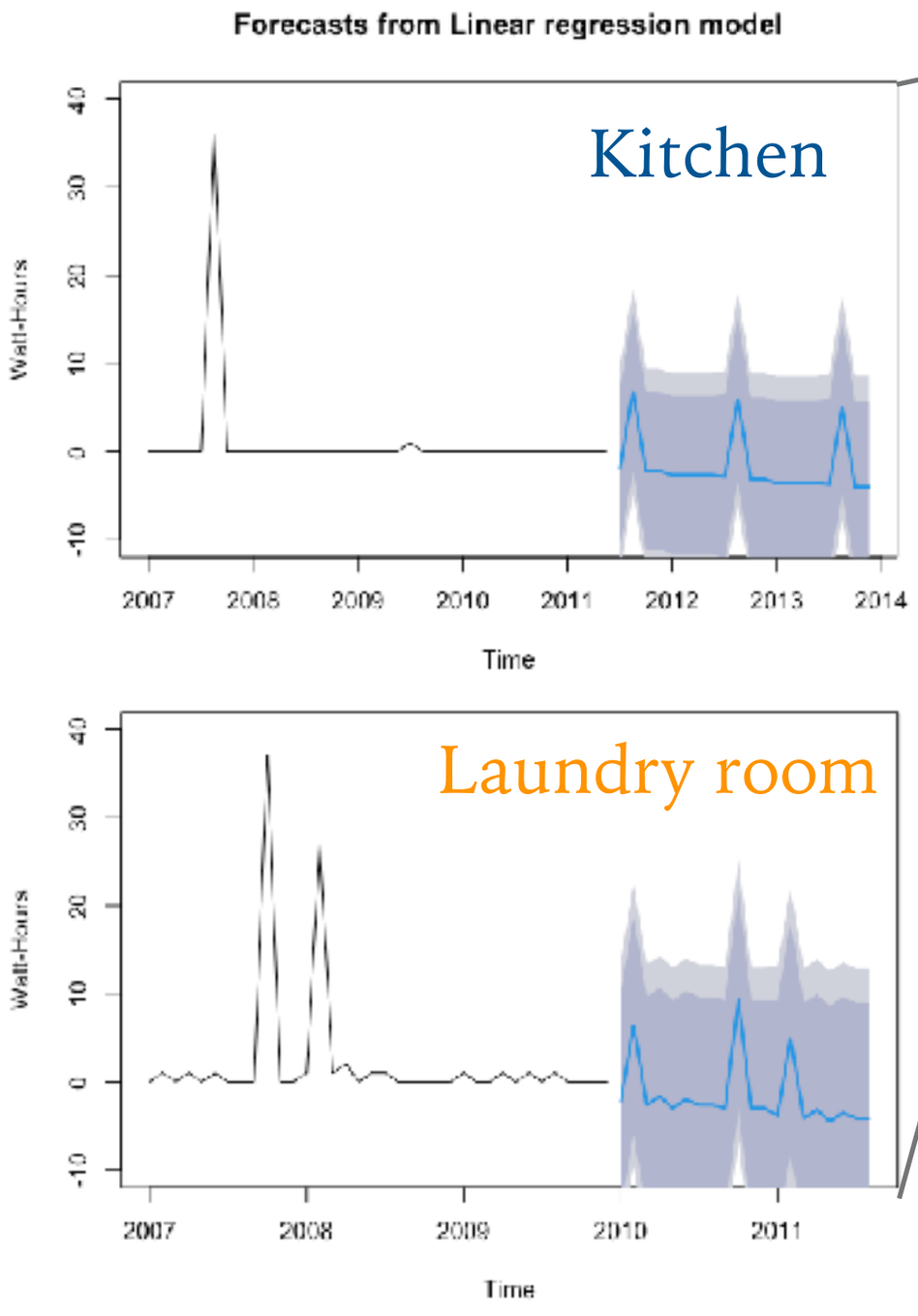
Time series

Components

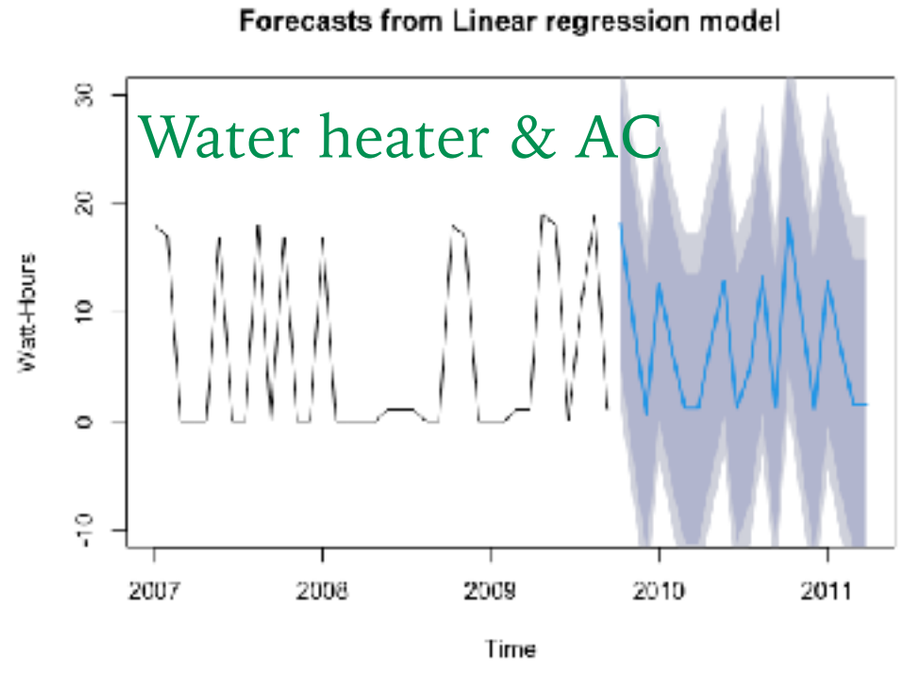
Fitting results	R2	Adjusted R2	p-value
Sub-metering 1	0.27	0.04	0.33
Sub-metering 2	0.32	-0.04	0.57
Sub-metering 3	0.46	0.14	0.22

➤ The most confident prediction is for the sub-metering 3.

ANALYSIS-FORECASTS

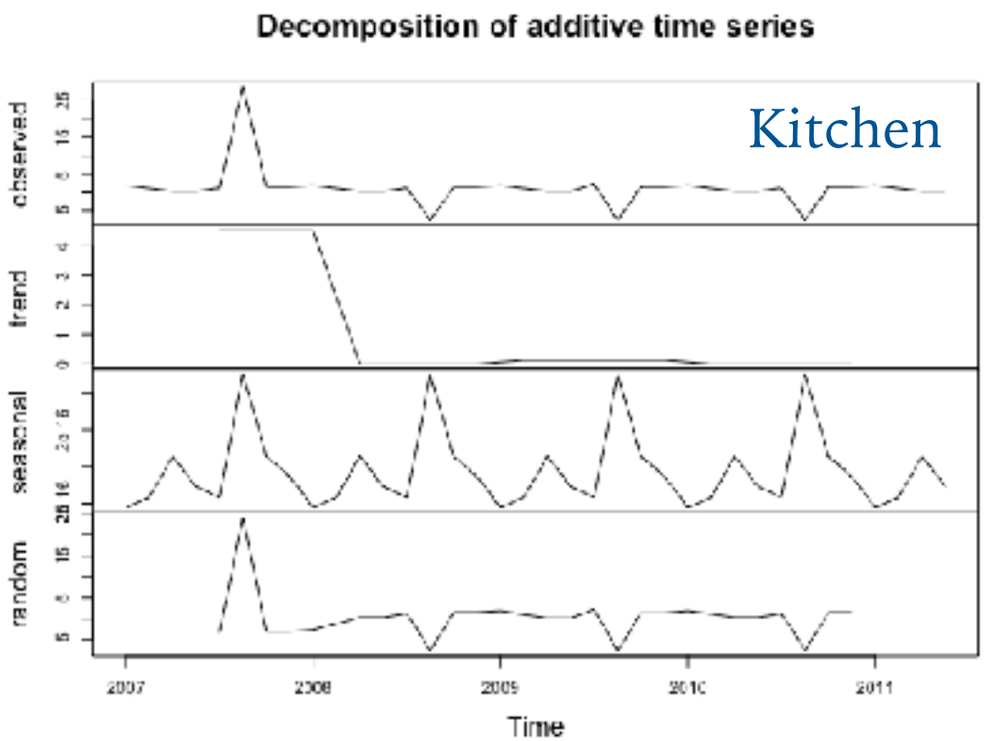


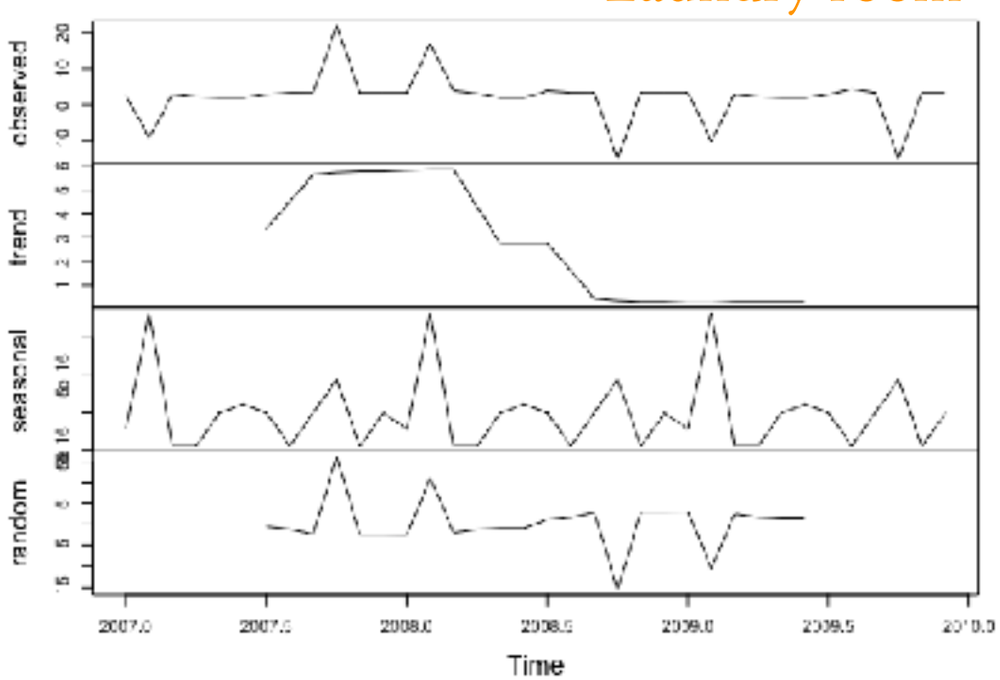
➤ The model predicts that the energy consumption will be nearly zero for the next two years with confidence of ~<67%. Use this prediction with caution.



➤ Our model predict that the energy consumption will be nearly the same for the following next two years with a confidence of ~78%.

ANALYSIS-DECOMPOSITION



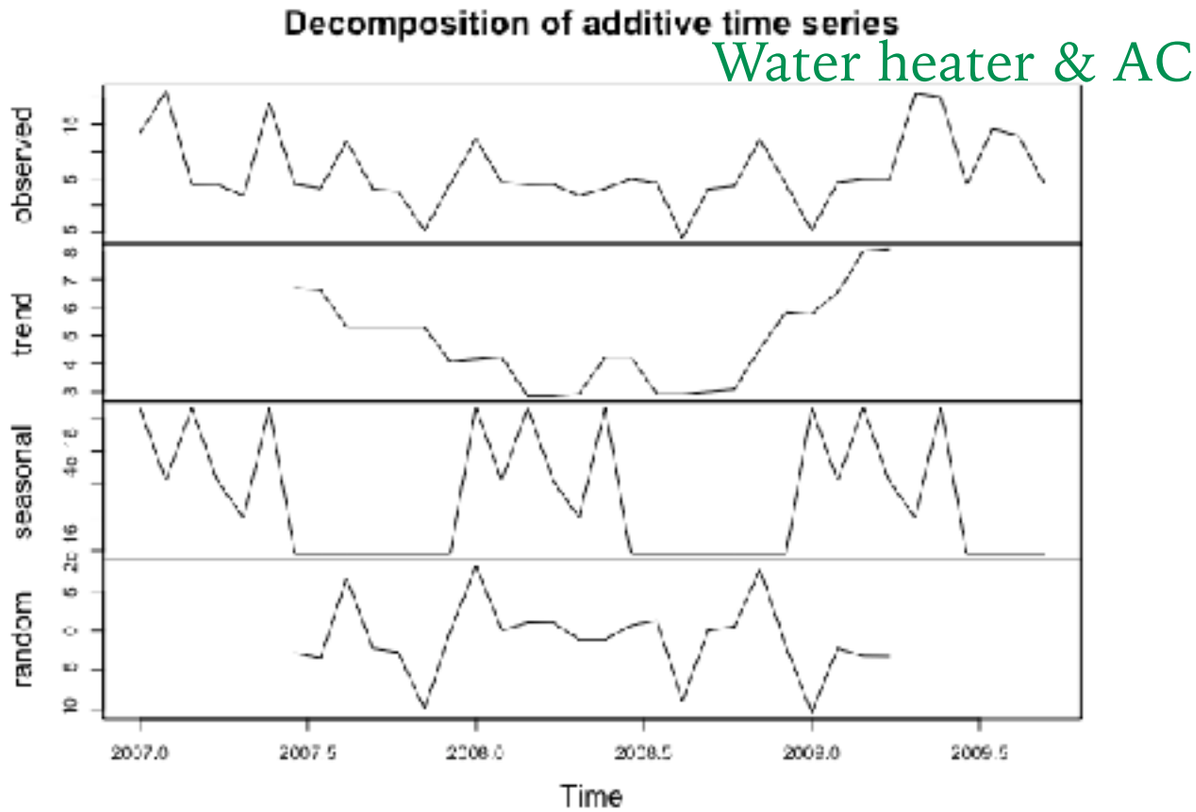


Components

Trend -> it describes a visible change of time values for higher or lower values over a prolonged period.

Seasonal -> these are periodic temporal fluctuations.

Random -> these are irregular and unpredictable fluctuations.



ANALYSIS–HOLT-WINTERS EXPONENTIAL SMOOTHING

In this technique, forecasting is based on all previous average weighted values, assuming that recent values are much more important than older ones.

- Alpha close to zero indicates that little weight is placed on the most recent observations.
- A lower SSE means a more accurate forecast.

Statistic results

Laundry room

```
> tsSM2_HW070809
Holt-Winters exponential smoothing without
 trend and without seasonal component.

Call:
HoltWinters(x = tsSM2_070809Adjusted, beta
 = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 6.610696e-05
 beta : FALSE
 gamma: FALSE

Coefficients:
      [,1]
a 2.343261
```

SSE=1516

Water heater & AC

```
> tsSM3_HW070809
Holt-Winters exponential smoothing without
 trend and without seasonal component.

Call:
HoltWinters(x = tsSM3_070809Adjusted, beta
 = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.2354651
 beta : FALSE
 gamma: FALSE

Coefficients:
      [,1]
a 9.273611
```

The least accurate

SSE=1706

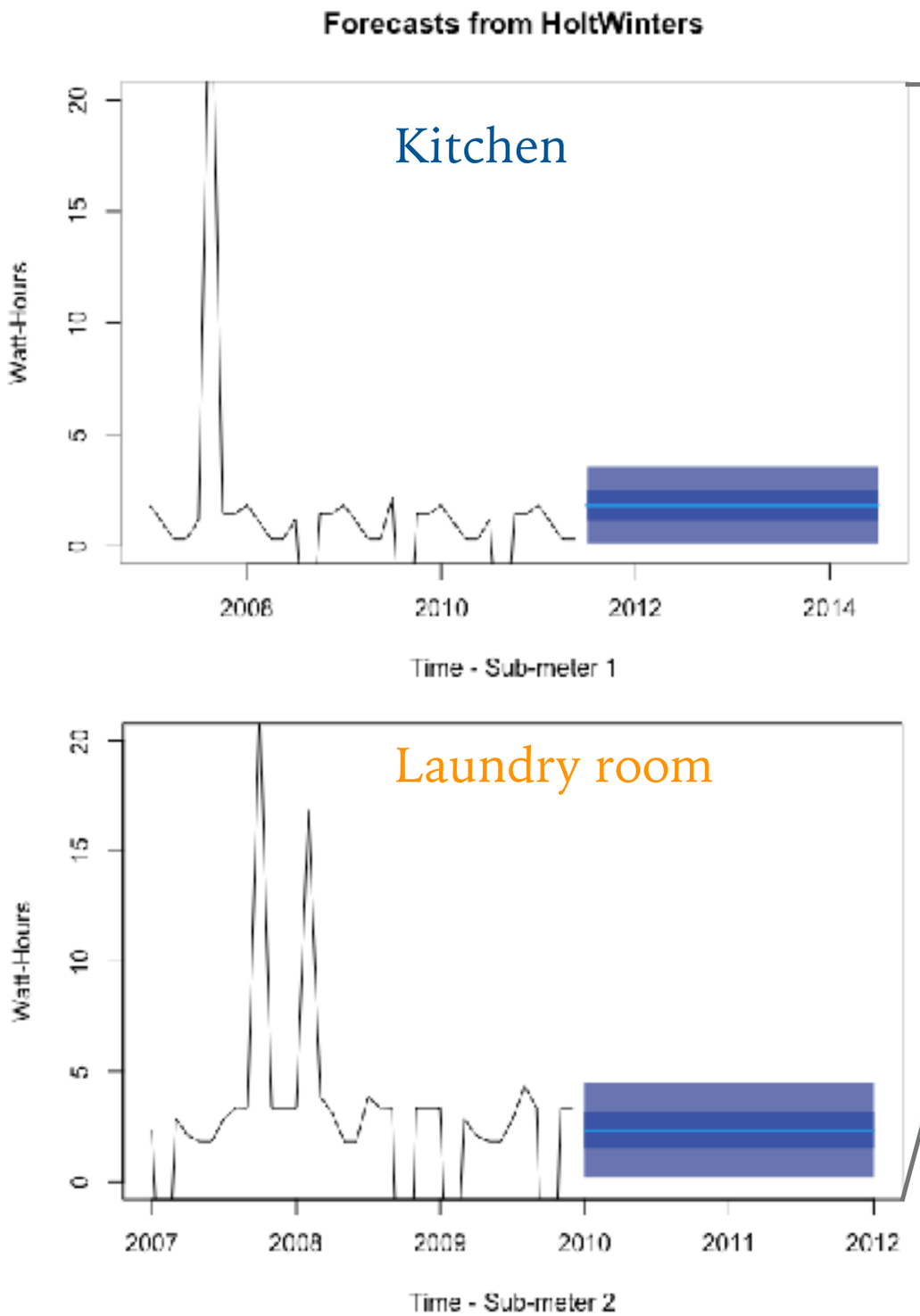
Kitchen

```
> tsSM1_HW070809
Holt-Winters exponential smoothing without
 trend and without seasonal component.

Call:
HoltWinters(x = tsSM1_070809Adjusted, beta
 = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 6.610696e-05
 beta : FALSE
 gamma: FALSE

Coefficients:
      [,1]
a 1.816061
```

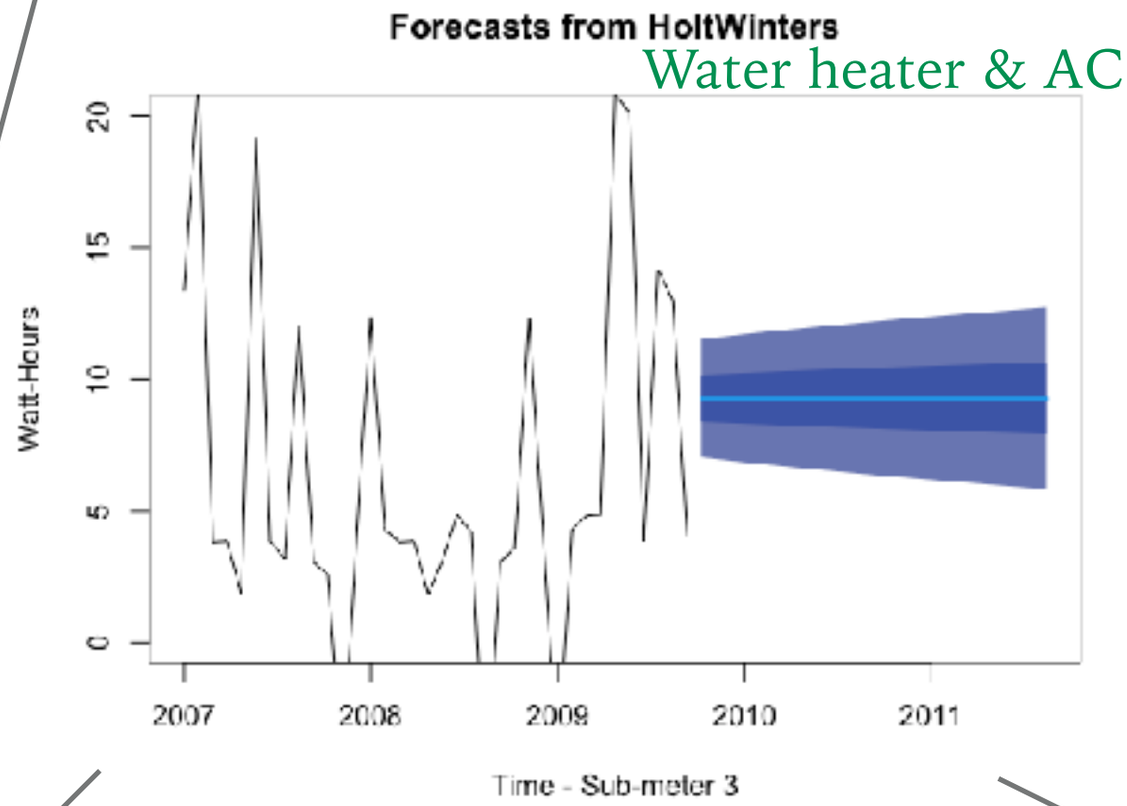
The most accurate

SSE=1000

TIME SERIES FORECASTING SOLUTIONS



➤ The energy consumption will be < 5 watt-hours.



➤ The energy consumption will be, on average, ~10 watt-hour.

SUMMARY AND RECOMMENDATIONS

- The sub-meters are very useful to keep proper energy consumption in our house or company.
- Time series machine learning is a powerful tool that allows predicting energy consumption.
- A good practice to avoid fires (due to overcharge), supervise the energy consumption, and save money should include using sub-meters and predictive analysis of the data obtained.
- In big houses, it is recommendable to have more than one single sub-meter.
- However, small houses with many electrical devices, it is always recommendable the use of several sub-meters.

Proper energy consumption allows having a more healthy environment for everyone.

LESSONS LEARNED

- It is important to understand the data structure before creating the time series dataset, for example, to know with which frequency the data was obtained.
- Creating different time-steps time series is beneficial for a better exploration of the possible components.
- It is very important to make a carefully fitting and forecasting processes and identify the subjacent components like the season, trend and/or random.
- It is fundamental to put together the results on the fitting statistical result, predictions, plots, common sense, and empathy to build useful recommendations that can derive into actions.