1. Exploring a database (15%)

Adress of the database:

https://archive.ics.uci.edu/ml/datasets/Appliances+energy+prediction

Content of the database

The data set is at 10 min for about 4.5 months. The house temperature and humidity conditions were monitored with a ZigBee wireless sensor network. Each wireless node transmitted the temperature and humidity conditions around 3.3 min. Then, the wireless data was averaged for 10 minutes periods. The energy data was logged every 10 minutes with m-bus energy meters. Weather from the nearest airport weather station (Chievres Airport, Belgium) was downloaded from a public data set from Reliable Prognosis (rp5.ru) and merged together with the experimental data sets using the date and time column. Two random variables have been included in the data set for testing the regression models and to filter out non predictive attributes (parameters).

Size of the database

There are 19736 rows and 29 columns of data.

Attribute Information format and units.

Column1: ‘date’ includes year/month/day/hour/minute/second with an interval of ten minutes. Date format.

Column2: ‘Appliances’ includes energy use in Wh. Integer format.

Column3: ‘lights’ includes energy use of light fixtures in the house in Wh. Integer format.

Columns4,6,8,10,12,14,16,18,20: ‘T’ includes the temperature in different areas in the house, in Celsius. Float format.

Columns5,7,9,11,13,15,17,19,21: ‘RH\_’ includes the humidity in different areas in the house, in %. Float format.

Column22: ‘T\_out’ includes the temperature outside, in Celsius. Float format.

Column23: ‘Press\_mm\_hg’ includes atmospheric pressure, in mm-Hg. Float format.

Column24: ‘RH\_out’ includes Humidity outside, in %. Float format.

Column25: ‘Windspeed’ in m/s. Float format.

Column26: ’Visibility’ in km. Float format.

Column27: ‘Tdewpoint’, Â°C. Float format.

Column28: ’rv1’ Random variable 1, nondimensional. Float format.

Column29: ‘rv2’ Random variable 2, nondimensional. Float format.

Application

Aggregating energy use from households allow us to predict energy demand and optimize the distribution and energy storage use in a power producer point of view.

The data also allow consumers to visualize their use of energy and plan their electricity usage to lower cost of utility bills and energy consumption in general.

一張含有 文字, 窗戶, 螢幕擷取畫面 的圖片

自動產生的描述

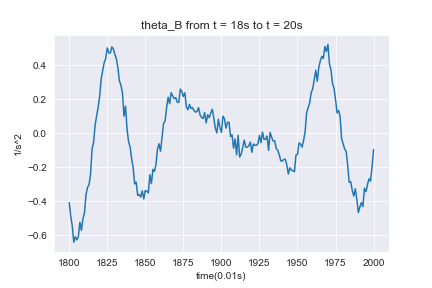
Source: Candanedo, Luis M., et al. “Data Driven Prediction Models of Energy Use of Appliances in a Low-Energy House.” Energy and Buildings, Elsevier, 31 Jan. 2017, https://www.sciencedirect.com/science/article/pii/S0378778816308970.

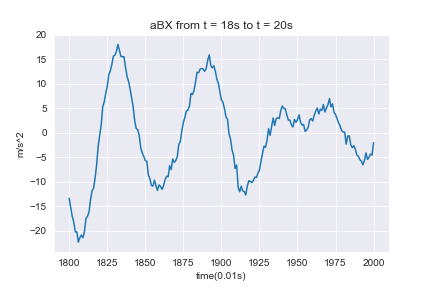
1. Dealing with Time-Series Data (60%)

(a)

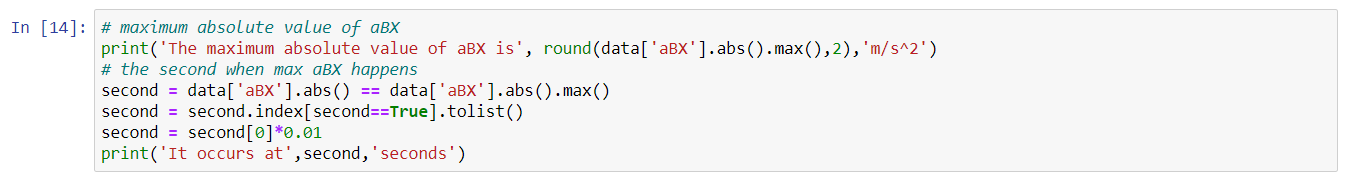
一張含有 文字 的圖片

自動產生的描述





(b)



The maximum absolute value of aBX is 22.36 m/s^2.

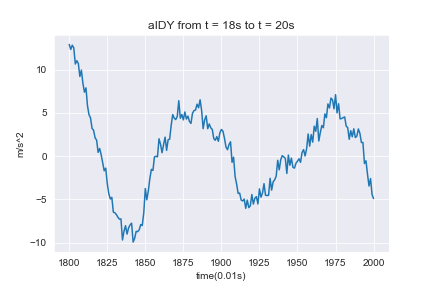
It occurs at 18.06 seconds.

(c)

一張含有 文字 的圖片

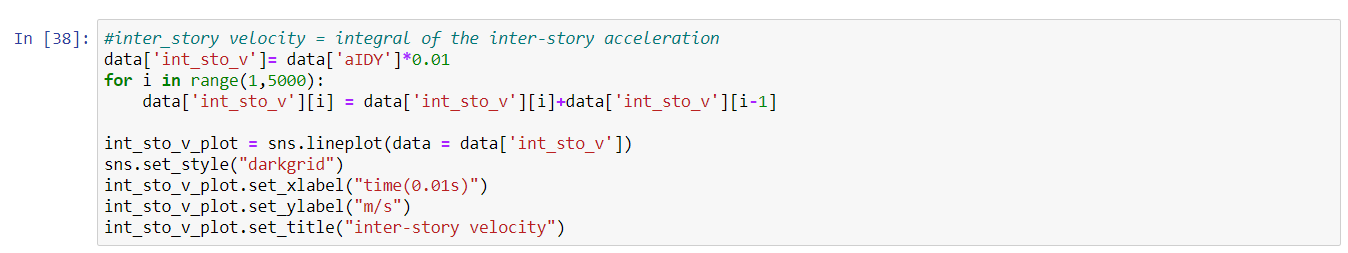
自動產生的描述





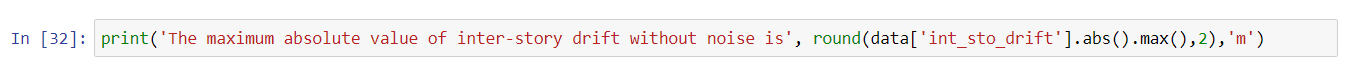
The maximum absolute value of aIDY is 13.81 m/s^2.

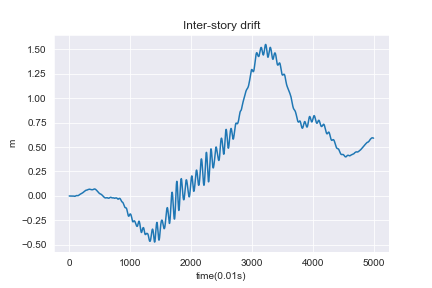
(d)



一張含有 文字 的圖片

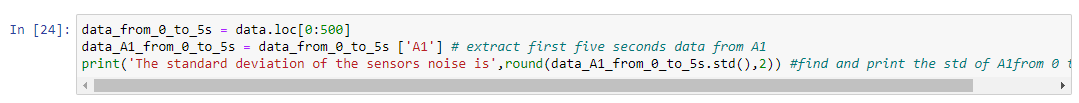
自動產生的描述





The maximum absolute value of inter-story drift is 1.55 m.

(e)



The standard deviation of a1 sensor’s noise is 0.61.

(f)

If we consider all sensor has the same noise as a1,

The formula to derive for is

If we calculate the noise of all sensors individually,

the standard deviation of a2 noise is 0.59

the standard deviation of a5 noise is 0.63

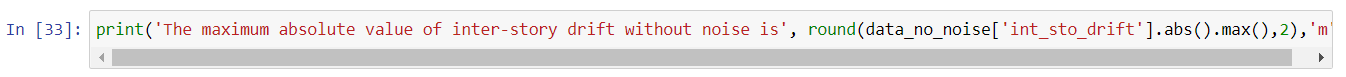
the standard deviation of a6 noise is 0.64

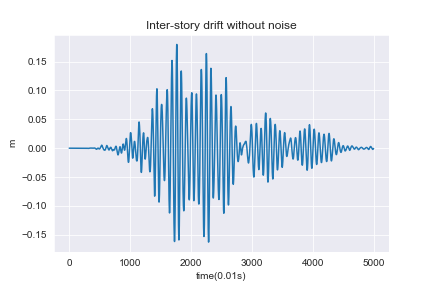
The formula to derive for is

(g)

一張含有 文字 的圖片

自動產生的描述





The maximum absolute value of inter-story drift without noise is 0.18 m.

The signal from this plot is without noise, therefore, the true value of the of inter-story drift can be obtained.

In (d), calculating the inter-story drift requires four sensor data, when all four sensors include noise, the real value of the inter-story drift is distorted.

(h)

In (c) we are first calculating inter-story acceleration for every time instance, then we find the time instance with the highest absolute value. The second method in the prompt grabs the maximum of aBY and aAY from different time instances and subtract to find the difference. We will only get the same result from the two methods when the two maximum happens in the same time instance. The more consistent method would be method provided by (c).

(i)

Yes, we can obtain building movement from sensor a4 and a8. For example, to calculate and  we can use this formula instead.

No, I don’t think we can obtain a more accurate result, because when involving more sensors that record noise into our calculations, we will bring about more uncertainty into our results.

1. Uncertainty Propagation (25%)