Loading data

Once the main R script has been launched it will start by loading observations and related data from the tables containing raw data from Energy Key. Only observations which have come in since last time the process ran will be loaded. For each building in the system, the below operations are performed to get consistent time series for each supported consumption type:

- Only observations of types water (m3), electricity (kwH), and heat energy (GJ) are loaded as those are the only ones currently supported by the website.
- Only observations originating from *billing* meters are loaded because the readings from a *distribution meter* can not be added to the total consumption.
- Observations which do not represent a one hour timeslot of consumption are removed because they do not originate from the automatic meters and are as explained earlier too inconsistent to deal with.
- Observations across different *billing* meters which are of the same type and measure the same time slot are summed together in order to get a time serie representing the total consumption within each consumption type.
- A number of projections are done in order to get all the new raw observations and previously processed observations from the W_consumptions table merged into a single vector of values representing the consumption each hour since a specified epoch date which is loaded from the W_settings.
- A linear interpolation is performed in order to sort out missing values. The max gap for the linear interpolation is 5. This means if there is no data for more than 5 consecutive hours in a row the values will just be left as missing values.

Now once the loading of data has finished the system holds for each building and consumption type vectors of equal length, representing the hourly consumption since a specified epoch date until now. At each index *i* will be the amount consumed between *i* hours after *epoch* until *i + 1* hours after epoch. This vector is stored as an array, along with the building id and consumption type into the *W_consumptions* table where it can be accessed by the website.

illustration