DESIGN PROJECT – SIE 2022

An API for the short-term simulation of heating demand in a district heating network with machine learning models

Students: Lucas Palumbo
Julia Wälti

EPFL supervisor: Prof. François Golay

External supervisor: Jerome Kaempf, Giuseppe Peronato
Idiap Research Institute

Context

In Monthey and Collombey-Muraz, the heating demand of 8,000 households is satisfied by the Satom waste incineration plant, which is equipped with an energy recovery system and connected to a 73 km heat network. In order to ensure stable supply and reduce the losses, the heat production should be matched to the demand as closely as possible. Indeed, heat is difficult to store so any extra heat created above the demand is wasted. The project is a part of the Eguzki project which is a collaboration between, Idiap, RWB, Altis and OIKEN.

Objectives

The goal of this project is to use artificial intelligence to create a machine learning model allowing short-term simulations of the heating demand based on building characteristics and weather data. The optimization of the supply requires an accurate prediction of the heating demand. The heat produced and injected in the network could be dynamically controlled using the hourly predictions of the demand.

Methodology

Data extraction
3D - custom CityGML 2.0 parser function

→

Data transformation
2D – each hour is a tuple for each building

→

Train/Test split
Separation 70-30%

→

Base model fitting on training set
Input: see list of features
Output: energy demand at time t

→

Test of the base datasets
On 4 chosen days: 21/03, 21/06, 21/09, 21/12 – absolute error/m²

→

Optimization of the best base dataset
- Addition of previous timesteps data
- Removal of features

→

TPOT

→

Final optimi-zation

→

Deploy-ment
With an API REST

232 training buildings | **8760** training hours | **100** test buildings | **24** hour prediction

Base models
- Linear regression with only the outside temperature
- Linear regression
- Random Forest Regressor
- Multilayer Perceptron Regressor

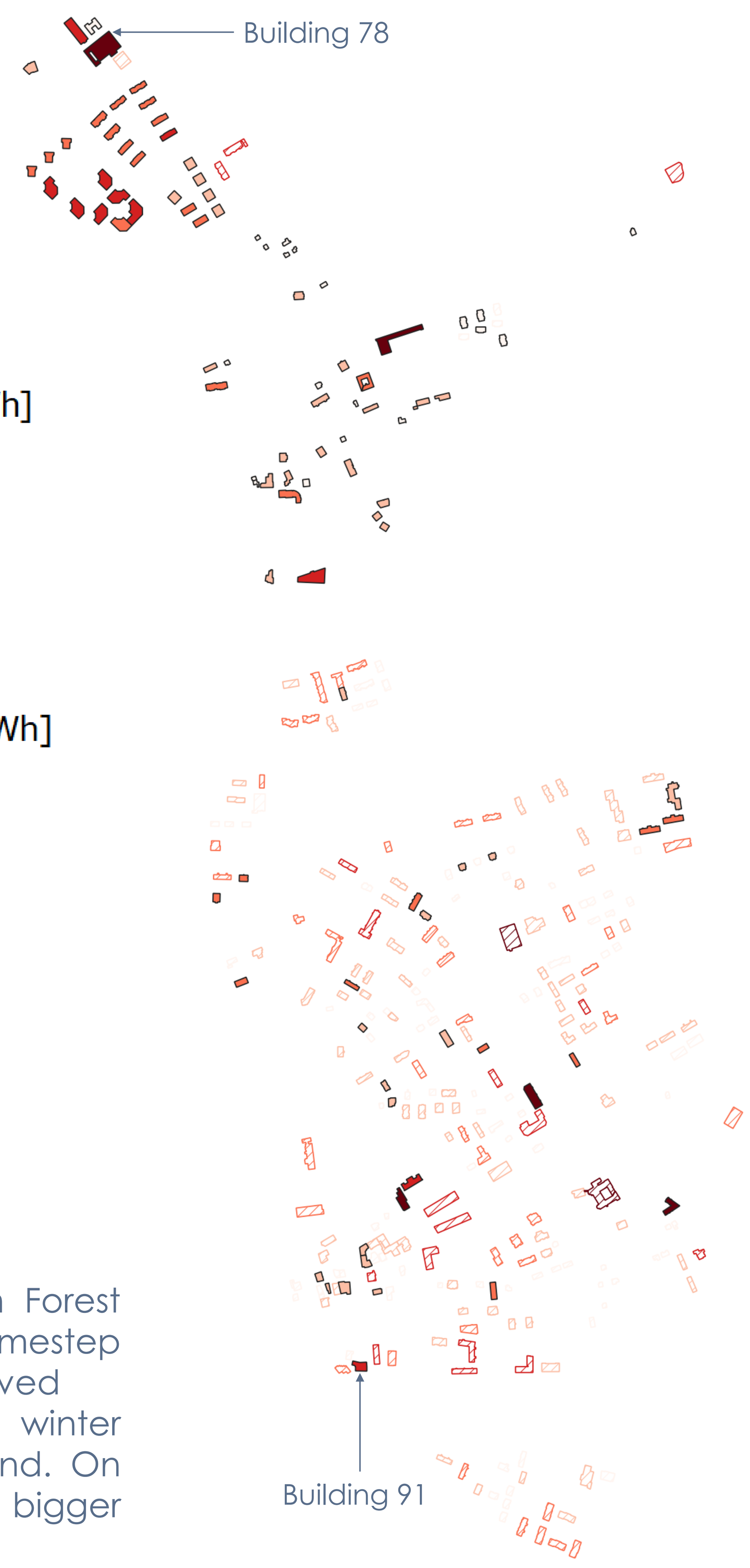
Comments
- Each optimization step requires to compare the errors
- The method followed is a trial and error approach
- TPOT is an AutoML tool that performs an intelligent search over machine learning pipelines with cross validation

List of features

Wall area [m²]	Hour of the day [-]
Floor area [m²]	Week end? [bool]
Roof area [m²]	Month of the year [-]
Open area [m²]	Wind speed [m/s]
Mean transmittance [-]	Wind direction [°]
Mean u value [W/(m²K)]	Diffuse horizontal irradiation [W/m²]
Building gross volume [m³]	Normal beam irradiation [W/m²]
Infiltration rate [1/h]	Relative humidity [%]
Number of occupants [-]	Precipitation [mm]
Occupation rate [%]	Air temperature [°C]
Total heat [W]	Soil temperature [°C]
Convective heat [W]	Zenith angle [°]
Radiant heat [W]	Azimuth angle [°]
Electrical appliance [Wh]	

Results

Fig 1 - QGIS representation of the results for 21/03 at 7am



Predictions for test buildings [Wh]
- 3539 - 19778
- 19778 - 42962
- 42962 - 66804
- 66804 - 94512
- 94512 - 143160

Real values for train buildings [Wh]
- 3304 - 30071
- 30071 - 55015
- 55015 - 89034
- 89034 - 157714
- 157714 - 246751

Comments
- Selected model: Random Forest Regressor, no previous timestep data and no feature removed
- Absolute error larger in winter because of higher demand. On the contrary, relative error bigger in summer
- No major impact of the building characteristics on the error
- The model is deployed thanks to an API REST. The user uploads the input files and downloads the output files thanks to a Python script executed in the command prompt
- The input and output of the model are GeoJSON files containing the optimized dataset plus the desired predictions for the output

Fig 2 - Absolute error for each day tested and the median values

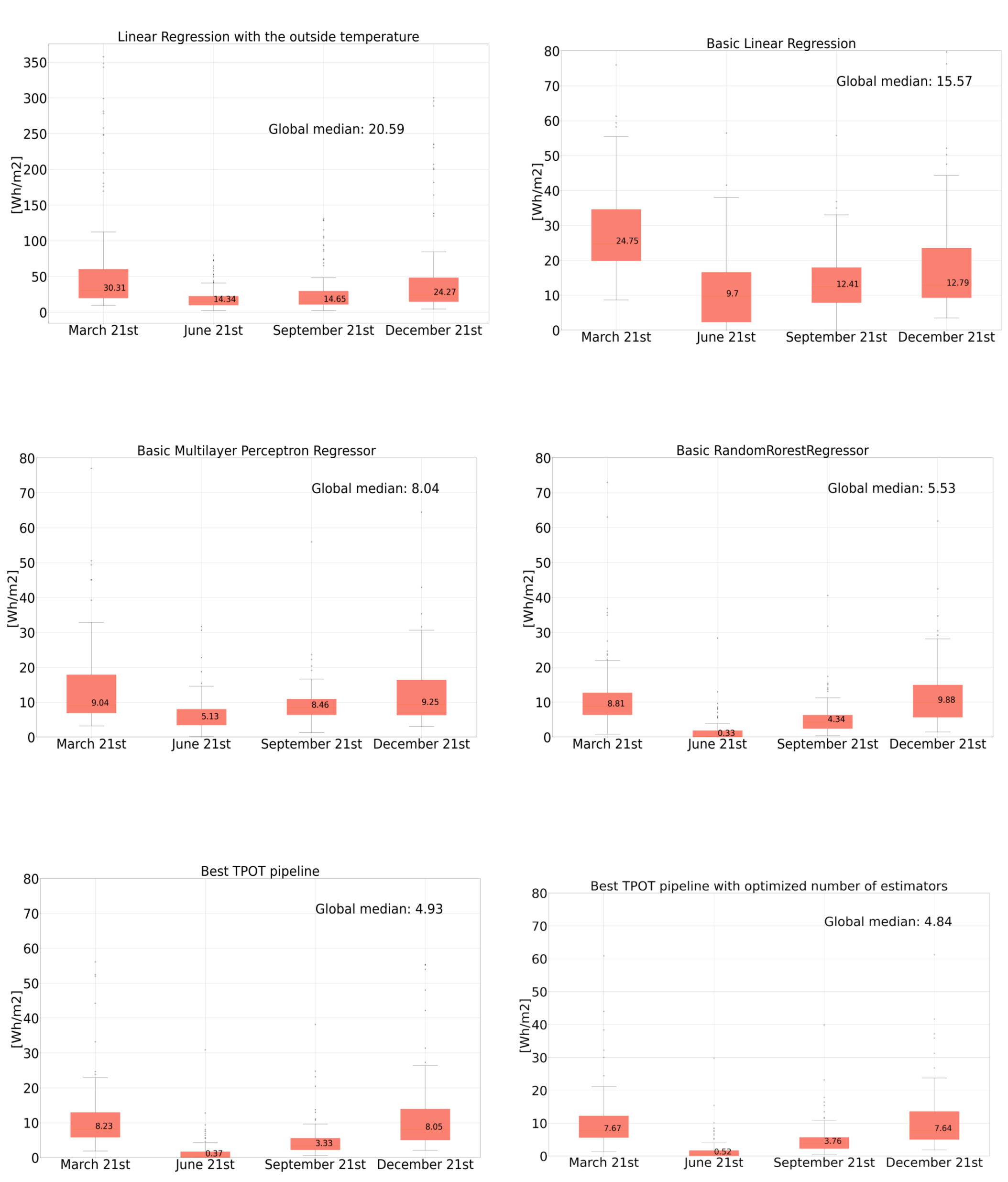
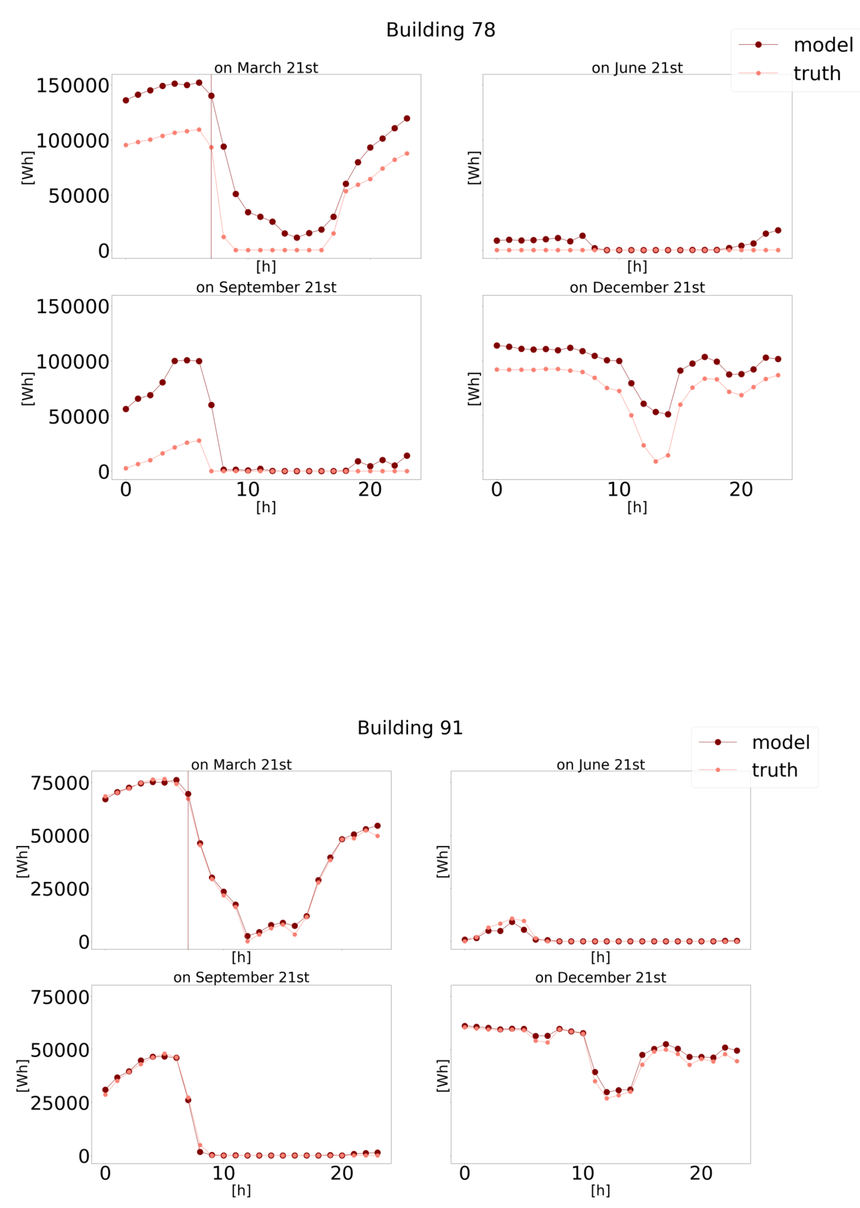


Fig 3 - Prediction for the best and the worst performing test buildings



6 Wh/m2/h mean error

39 Wh/m2/h mean consumption

36 % mean relative error

Conclusion

In this project, a CityGML 2.0 parser function has been developed to create a functioning dataset. With this data and the weather information, a model has been developed to predict the short-term heating energy demand of buildings. The best model has been defined using a trial and error approach. The model has a good performance, although it could be improved for some buildings and in the summer.
The final model has been deployed via an API and the results can be visualized in QGIS.