ENAC / PROJET DE MASTER 2019-2020
SECTION DE GÉNIE CIVIL

# District scale prediction of subsurface waste heat flows

Autor : Thomas Wüthrich

Supervision : Prof. Dr. Lyesse Laloui [1] / Prof. Dr. Alessandro Rotta Loria[2] / ing. Dipl. Benoît Cousin[1]

[1] Soil Mechanics Laboratory (LMS), EPFL / [2] Mechanics and Energy Laboratory (MEL), NU Chicago

## Description

The human-made underground structures generate an anthropogenic heat flow which has a significant impact on the ground temperature and leads to the creation of urban underground heat islands with a high geothermal potential. This thesis is dedicated to propose a quick and accurate machine learning (ML) based approach for the evaluation of the underground waste heat flow of buildings with focus on the ground temperature change due to heat losses of basements. This method can be used as a supplement to the conventional finite element modeling (FEM) and are often suitable to avoid complex and time-consuming modelling. A comprehensive validation is presented for seven small scale (building scale) scenarios, and a district scale application for the downtown district 'Loop' in Chicago, USA.
To do so, we assume only a conductive heat flow mode and constant boundary conditions. As a result, 2D heat maps are created by discretizing the domain of interest in square elements. Each element serves as data point to which a temperature is assigned

## Machine Learning

ML is a subset of artificial intelligence that enables computers to learn without being explicitly programmed with predefined rules. The ML technique is subdivided in two main phases (Figure 1). In the **training phase** pairs of input and output data are known. The output data are denoted as labels. The input data, also known as features must be extracted from the available data in a previous phase. During the main process of the training phase, the goal is to find the relationship between the train labels and the features. Therefore, a Random Forest algorithm is used. The second phase, called the **prediction or validation phase** has the main goal to test and validate the algorithm with the determined prediction model. Therefore, an analysis of the testing error can be done to evaluate the accuracy of the model.

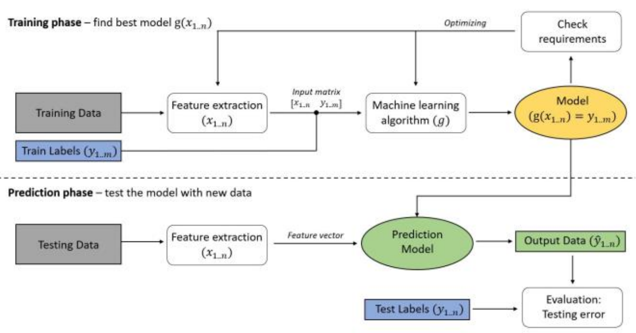


***Figure 1:*** Generalized machine learning principles

### *FEM and RF*

To perform numerical simulations which are necessaire to train and test the ML algorithm, the finite element software package COMSOL is used. The heat transfer is the governing equations to solve. The total heat transfer is expressed by the energy conservation equation. For any isolated domain, the considered equation expresses that the amount of energy remains constant.

The RF is an ensembles of decision trees. In these models a multiple of decision trees were combined to create a more powerful model. A simple decision tree uses a tree-like model of decisions (Figure 2). At each node, a series of binary splits separates the training data in two new children nodes. Often, the mean squared error or the standard deviation are used to find the best split at each node.

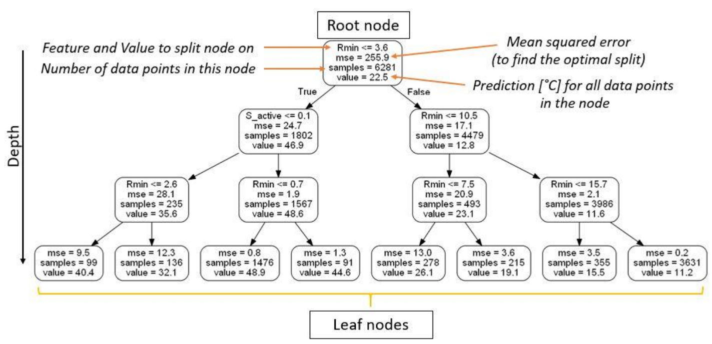


***Figure 2:*** Illustration of a single decision tree

## Results "Building scale"

For the small scale application, a domain of $16\,x\,16\,m^2$ is chosen. There are various heat sources in the domain which contribute to the change in ground temperature. For this study different heat source geometries, locations, number and thermal loads are studied. Furthermore, the ML algorithm has been applied on different depths, simulation time periods and grid sizes.
The first three scenarios (heat source geometry, location and number) have shown that the error remained constant early on and a larger training set did not lead to a strong decrease in error. In all three cases only 3 to 4 FE models were needed to adequately predict most of the test models. In the study of different heat source temperatures, however, a larger training set led to an even greater reduction in error. A further study has shown that the error is independent of the meshing size. The finer the mesh the more realistic the prediction, but also more time consuming as more grid cells must be predicted. The error increases as soon as the heat map is created for a depth below all heat sources. In a last study it could be shown that the error is in function of time. During the time dependent phase, the error increases with increasing simulation time, but then remains constant when the steady state phase is reached. From the analysis of the seven different small scale scenarios we can conclude that the testing error is always in the same range and therefore the RF algorithm works equally well for all cases. We obtain a mean absolute error always smaller than 4 to 8%, a root mean squared error of always less than 5 to 10% and a maximum error of around 20%. Furthermore, it is very important that the train set is always representative of all possible situations so that the RF algorithm can better generalize all problems. Figure 3 illustrates the heat maps for three selected scenarios. In each case the results are represented for the FEM and the machine learning based RF algorithm.

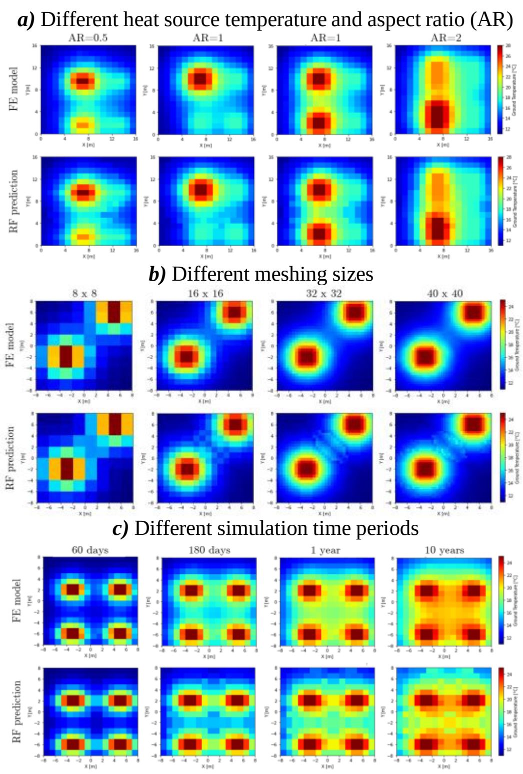


***Figure 3:*** Heat map of three studied small scale scenarios, generated by FEM (top) and with the RF algorithm (bottom)

## Results "District scale"

It was noted, that 25% of the grid cell data from the FE model are sufficient for training the ML model to predict the whole district. Figure 4 shows the results of a 50-year simulation period of FEM and RF. Some changes of the district (omission of buildings) can still be predicted with the same ML algorithm, i.e. with the data of the same FE model. Obviously, the error increases, but it remains small. The mean absolute error is never higher than 3.5%, the root mean squared error is always less than 5.5% and less than 9% of prediction data have a maximum error greater than 5%. If 50% of the data points are in the train set the error can be further reduced. However, with more data the learning time increases by almost twice as much. Figure 4 shows the results of a 50-year simulation period of FEM and the RF prediction.

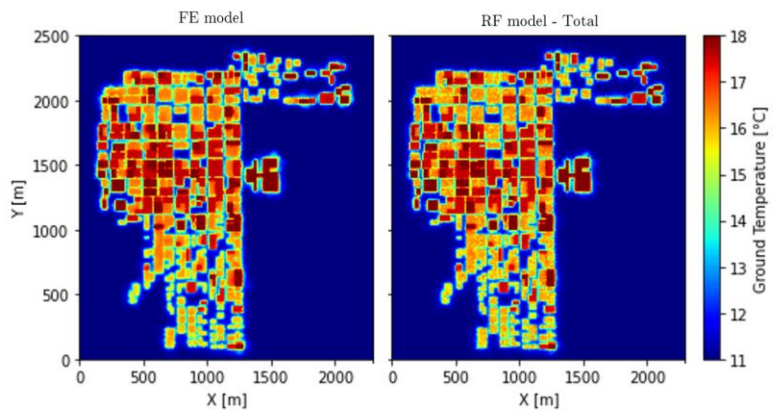


***Figure 4:*** FEM and RF model for the whole Loop district, trained with 25% of the FEM grid cell data