An ML analysis have been conducted on settlements and rotations of foundations in a plane strain FE calculations. Two different models have been applied

* **Linear regression model** –Based on a linear relationship between the features. Applies least squares method to optimize the fit.
* **Random forest regression** – A model that optimizes predictions by creating multiple decision trees, each trained on a random subset of the data and a random subset of features. The result is the average of the predictions from all the trees in the forest.

The training models are based on a stiff plate element in a plane strain model, exposed to a fixed placed with a various eccentricity. The subgrade consists of random soil layers per 0.5 m. The subgrade has been modelled as a Mohr-Coulomb soil model, with a high failure criterion. The intention is not to evaluate the plastic behaviour of the soil, merely to allow for zero stresses at a foundation edge under high eccentricity. Below is given a list of feature details.

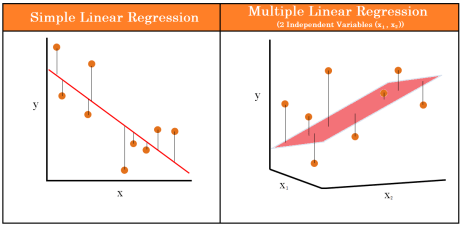
* Stiff plate element, width varies between 1-4 m with 0.5 m intervals.
* Constant load per plate width (100 kN/m/m \* foundation width).
* Load placed with a random eccentricity between 0 and 0.3 \* foundation width.
* Model side BC’s are located 2\*foundation width from the foundation center. Hence, a total model width of 4 times foundation width.
* Model bottom BC is located 2\*foundation width under the foundation.
* Subgrade with random E-modulus between 10-100 MPa are modelled per 0.5 m under the foundation down to bottom BC. An angle of friction of 40° and c’=300 kPa has been applied.

The model has been trained on around 5000 Plaxis 2D calculations. The test samples are just below 600.

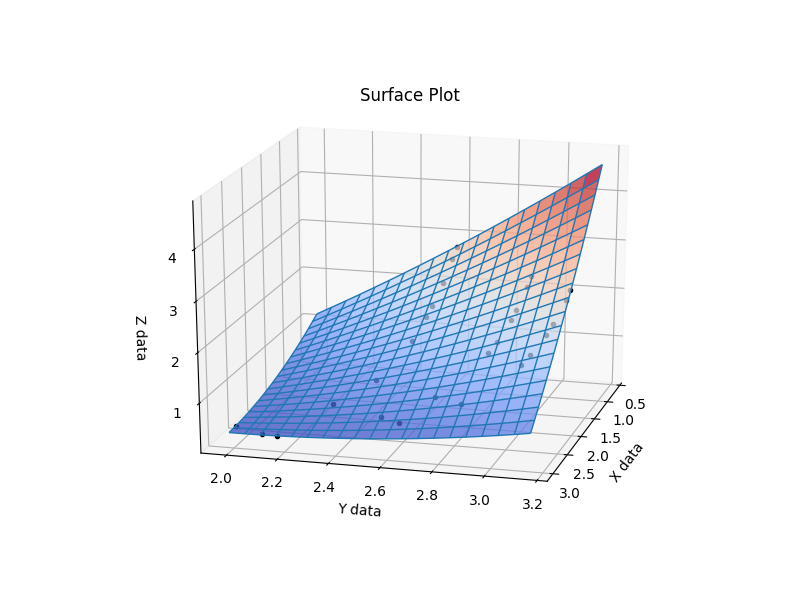
# Models used

## Linear Regression model

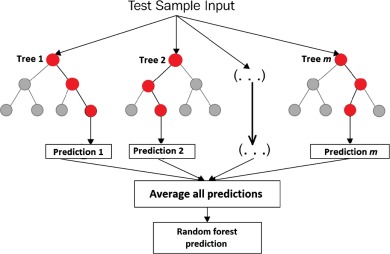
For each point i, for each feature j



## Polynomial regression model



## Random forest model



In the following, the performance of three models have been evaluated. It should be noted that, for simplicity, the hyperparameters of each model have been kept at scikit-learns’ standard settings.

Models tested are linear regression, random forest and polynomial regression. As we do expect the settlements to have a somewhat linear correlation to the features, we do expect both linear and polynomial regression to perform pretty well. The random forest model has been added as a reference, however we do expect it to perform quite well in general.

# Brief explanation of the models

## Linear regression

Heps

## Polynomial regression

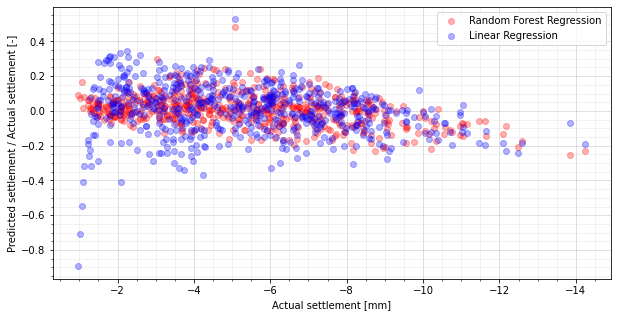
Tjep

## Random forest

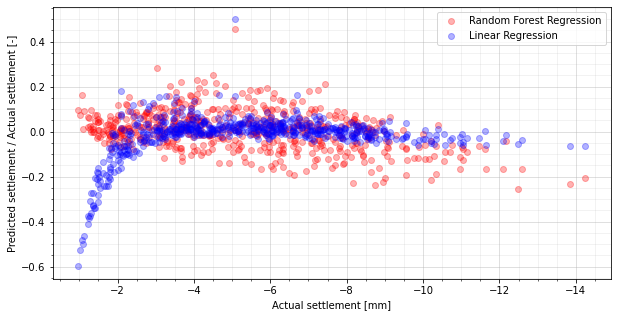
Tjop

# Results

A result with models trained on all features as defined in Plaxis.

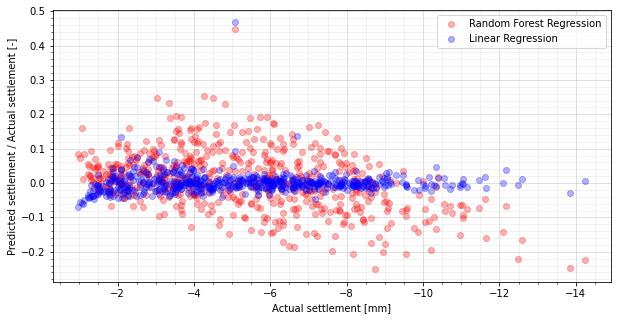


With E-modulus inversed. That is, because the settlements are not linearly dependent on the E-modulus. S=sigma/E. S is however linearly dependent on E inversed. Therefore, to better capture the dependency of settlements with a linear regression, we could apply the feature E inversed instead of E for each layer. This gives a significantly better result than out first calculation.

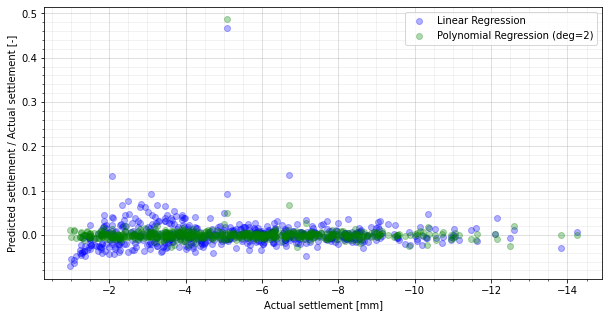


It is known that foundation size has a significant impact on the subgrade layers that are activated. However, in this model, we are fitting all layers with one coefficient independent on foundation size. As the change of foundation size has the biggest influence on small foundations, it can be seen that our model are significantly off for the smallest foundation sizes (leading to the smallest settlements). As the model has been fitted on all calculations, the results with smallest and largest foundation sizes are expected to be most off the test result.

This effect, however can to some extend be captured by adding one new feature per soil layer, which is the soil layer inverse stiffness multiplied by foundation size. This way, it allows the model to adjust the foundation size’ influence on the soil layer coefficient – however only with a first degree.



As we can see, the model has a significantly better fit. Of course, we are still applying linear presumptions on the relations in the model. Applying a second degree polynomial regression, we can capture smaller non-linearities.



Looking at the feature coefficients of each soil layer, it can here be seen that soil layer 0 has the highest coefficient, and that soil layer 0 with foundation interaction is not as high. That is because soil layer 0 always have a significant impact on the settlements. Therefore, the sole soil layer coefficient is higher with a lower influence of foundation size. Looking at soil layer 3, however, it has a quite smaller effect on itself, but as the foundation sizes increases, the impact is rising. As we go further down the soil layers, both the soil layer itself, and the foundation interaction coefficient decreases, as we expect that the influence on the soil layer decreases in general as we get down in the deeper layers.

