# Building of machine learning model

The dataset in “**2022\_Test\_ML.csv**” consists of 4 features (s\_mt, s\_mq, d, h\_p) and of 2 objectives (QW, DP). The goal of the task is to build the model which approximates and generalizes QW and DP for arbitrary values of s\_mt, s\_mq, d, h\_p within the range of their variation (Table 1).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | s\_mt | s\_mq | d | h\_p |
| Min | 0.8 | 0.8 | 1.0 | 4.0 |
| Max | 2.7 | 2.1 | 3.0 | 10.0 |

Table : Ranges of variations for considered features

# Requirements for the test task:

1. The task should be solved with the use of Python in the framework of Jupyter notebook
   1. Analyze data to understand the dataset quality and properties of the target functions
   2. Building the model for target functions. The choice of the regression model could be arbitrary but should be justified
   3. Assess the approximation and generalization errors of the obtained model. Use the coefficient of determination (R2) for the error metrics
   4. Show the dependency of the model approximation and generalization errors from the total number of points in the dataset
2. The results should be summarized in the PowerPoint presentation containing the test case description and results