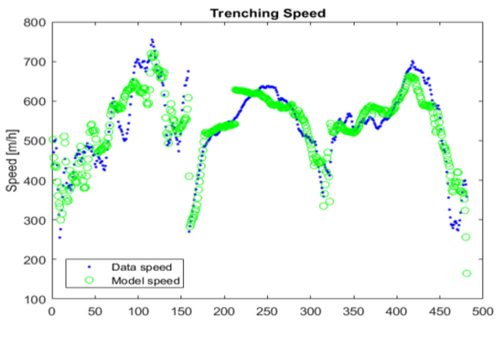
Machine Learning models to predict subsea trenching speed

Project summary



Client:	SAIPEM - Sonsub
Partner:	-
Year:	2017-2018
Service:	Data Analysis & Machine Learning
Sector:	Geomechanical Engineering
Project in numbers:	800+ lines of code
Coding language:	Matlab
Other details:	-

Abstract

Predicting the trenching speed is a task of paramount importance, in particular to correctly estimate the time (and hence the costs) required for the subsea operations.

The project aimed at developing Machine Learning models to predict subsea trenching activities. The models have been trained and tested on real operational data, acquired during an off-shore campaign in the Mediterranean see. The project mainly focused on predicting the trenching speed, as a function of the trencher configuration and soil characteristics. The results of the project show how the ML models outperform the results given by the available analytical models, provided that input data are consistent with the range where the ML models have been trained.

Project description

The project focused on developing data-driven Machine Learning models, based on data acquired during the trenching activities conducted with the PL3 plough in the Zohr field, during the period May-July 2017. Collected data were provided by several sensors applied directly on the trencher, which logged the position, the applied forces, the speed, etc., and by the geotechnical surveys, which provided the principal geomechanical characteristics of the soil.

Initially, the available data have been preprocessed, i.e. filtered and cleaned, in order to obtain a dataset with the input variables (*X*) and the output variable (*Y*) to be predicted, namely the speed of the trencher. In particular, two tabular data have been created, separating frictional from cohesive soils.
Subsequently, several Machine Learning models (belonging to the class of multivariable linear regression models) have been developed, considering separately the two types of soil. Each model has been validated with K-fold crossvalidation techniques, with 70%-30% train/test split.

M^{3}E srl, Spin-off dell'Università degli Studi di Padova
via Giambellino 7, 35129 – Padova, Italy
Partita IVA: 04573020288

R.E.A.: n° 400761
E-mail: info@m3eweb.it
Website: www.m3eweb.it

More then 60 models have been generated, in order to identify the best hyperparameters configuration and avoid overfitting problems.

Machine Learning models were able to provide an $R^2 > 0.9$, which is a promising result considering that an ideal exact model will give an $R^2 = 1$. Considering the *Root Mean Squared Error* (which is a measure of the difference between the *predicted* and the *real measured* values) we obtained up to a relative $RMSE^{rel} \simeq 0.2$ and an absolute $RMSE^{abs} \simeq 40$ *m/h,* which again are promising results to use data-driven models to predict trenching operations.

The above-mentioned results were all obtained avoiding overfitting in the models, i.e. excluding all those models where the metrics in the test set where significantly worse than those in the test set.

Finally, a preliminary comparison between the data driven models and the analytical model (namely the Cathie model) has been conducted, to gain sensibility on the predicted speed by the two different approaches. The models have been compared within the range of available data and hypothesizing different feasible scenario.

The projects showed how data-driven models and Machine Learning approach are a very powerful approach to predict trenching activities.

Further developments of the project will involve testing more advanced Machine Learning models, calibrating gray box models (combination of white and black box models), extending the Machine Learning approach to other trenchers.

Project outcome

The outcomes of the project have been:

- a 60-pages report containing the workflow to preprocess and analyze the data, a descriptive analytics section with several data visualization contents (scatter plots, box plots, correlation matrix, etc.) and the results of the models.
- a Matlab suite of codes implementing the complete workflow, from the raw data import of **.csv* files to the generation and testing the Machine Learning models. The suite of code is completely parametric in order to be used to train more models if new additional data are available.

References

State of the Art Ploughability Assessment, Indrasenan Thusyanthan, (2016), The 26th International Ocean and Polar Engineering Conference.

Pipeline trenching using plows: performance and geotechnical hazards, Cathie, David N., and Jean François Wintgens, (2001), Offshore Technology Conference. Offshore Technology Conference.

Geotechnics for subsea pipelines, White, D. J., and D. N. Cathie., (2010) Proceedings of the 2nd International Symposium on Frontiers in Offshore Geotechnics, Perth.