

EGU21-13899

<https://doi.org/10.5194/egusphere-egu21-13899>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Application of tree modeling techniques using Weka for the analysis of records of an erosion experiment on cohesive soils carried out by the Federal Highway Administration

Ernesto Garcia Rugerio¹, Rabindranarth Romero López¹, and Gerarld Corzo Pérez²

¹Universidad Veracruzana, Facultad de Ingeniería, Posgrado en Ingeniería, Xalapa, Ver, México (ernesto_ob@hotmail.com)

²IHE Delft Institute for Water Education, DELFT, Netherlands

The methodologies applied in the analysis of scour in cohesive soils that exist have been evaluated based on linear or potential regressions of the results of experiments carried out in laboratories, however these procedures do not allow to clearly identify the weight of each variable in the explanation of the response variable, they also do not have the ability to carry out regionalizations of the analyzed data universe so that a better coupling of the resulting equations can be done.

Every day data mining techniques are more usefull for analysis of different problems, in the present case study, the use of these techniques is evaluated in the analysis of results of an erosion experiment in cohesive soils carried out by the Federal Highway Administration (FHWA), these results were published in technical report No. FHWA-HRT-15-033 dated May 2015.

The geotechnical and hydraulic variables and the erosion results obtained during the execution of the experimentation were used, with which it was analyzed using the WEKA software (Waikato Environment for Knowledge Analysis) of the University of Waikato in New Zealand, which uses data mining techniques based on different rules and types of information classification such as decision trees.

Through the application of the tree section, various tests were carried out, this with the intention of determining the most important factors that describe the phenomenon of erosion, on the other hand, a series of classifications and equations were obtained through the M5P model that describe the phenomenon . As a result, it was obtained that the variables that describe the erosion phenomenon better according to the analysis of the M5P model are the shear stress, the plasticity index, the unconfined compression stress of the samples and the content of humidity. The result is a tree with 6 rules that zoning and regressing each zone obtaining a correlation coefficient of 0.9246 with an absolute relative error of 33.5874% and a root of the relative square error of

38.0878%. It is mentioned that with the adjustment through potential regressions obtained by the FHWA, a coefficient of determination (R^2) of 0.73 was obtained.

The application of this type of techniques allows a deeper knowledge of the erosion phenomenon by classifying and regionalizing the explanatory variables, as well as carrying out regressions within these classifications, explaining the behavior of soils with content of cohesive material as a function of its variables. The implementation of these data mining techniques has more advantages than simple linear or potential regressions, being of great help in research and experimentation in the field of geotechnics and river hydraulics.