

Young Member Arena

Inaugural ISSMGE TC304 Student Contest on Data Analytics

Overview

ISSMGE TC304 on Engineering Practice of Risk Assessment and Management Student Contest on Data Analytics aims to encourage students to learn statistical and probabilistic methods (e.g., machine learning, big data, deep learning) and implement them in geotechnical data (e.g., 304dB complied by TC304, referred to http://140.112.12.21/issmge/Database_2010.htm). The inaugural TC304 Student Contest was held on August 18th 2018 in Harbin Institute of Technology in Harbin, China. A contest challenge question related to 304dB was released and given to all participants, and participants were given 3 months to address the question. Participants can work on the question alone or in team. They were asked to submit a full length paper in English and to present their works in conferences (10 minutes presentation plus 5 minutes Q&A). The ISSMGE TC304 will regularly hold TC304 Student Contest on Data Analytics in future conferences.

Contest question and logistics

The question for the first TC304 Student Contest on Data Analytics was released on May 1st 2018. It involves a site investigation dataset comprising of multivariate data of liquidity index, vertical effective stress, preconsolidation stress, remolded undrained shear strength and undrained shear strength obtained at different depths of a clay site, as shown in Fig. 1. The dataset was manipulated in the way that several numbers were replaced by outliers with various degrees of departure from the norm, which are indicated by solid circles in the figure. Participants were required to find the outliers using data analytics methods (e.g., statistics, probability, machine learning, data mining, etc.) and to present clearly the criteria, methods, and algorithms of detecting those outliers in their full paper. The full paper was due by August 3rd 2018, and 14 teams (36 participants, see Fig. 2) submitted their papers to the award committee by the date. All of these papers were presented by participants on August 18th 2018.

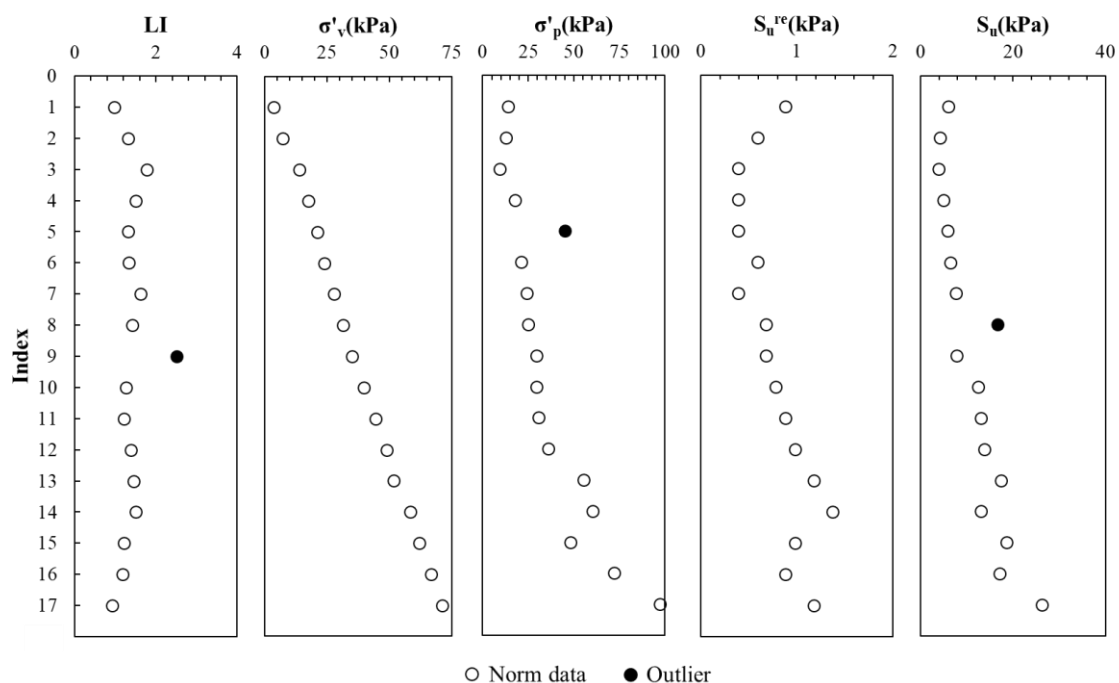


Figure 1. Dataset used in the Inaugural ISSMGE TC304 student contest on data analytics

Young Member Arena

Inaugural ISSMGE TC304 Student Contest on Data Analytics

Awards

Depending on how this contest question was addressed, one TC304 Student Contest Award and three Encouragement Awards were selected and given by the award committee. The award committee members for the inaugural contest includes Prof. Kok Kwang Phoon (Chair of the award committee, National University of Singapore), Prof. Jianye Ching (National Taiwan University), and Prof. Limin Zhang (Hong Kong University of Science and Technology). The inaugural TC304 Student Contest Award was given to a team comprised of graduate students (Shuo Zheng, Qinxuan Deng, and Yuxin Zhu, see Fig. 3) from Wuhan University, who proposed a novel, but fairly simple, method for outlier detection of multivariate data and applied the proposed method to the given contest dataset for illustration and validation. Three other teams from Tongji University (Jingjian Hu, Jinzheng Hu, Tao Cheng), Wuhan University (Te Xiao and Jian He), Harbin Institute of Technology (Jingze Gao, Guoliang Wang, Renming Deng, and Laiming Deng) won the three Encouragement Awards (see Fig. 4). The TC304 Student Contest Award paper and Encouragement Award papers can be found on ISSMGE TC304 webpage via the following link: <http://140.112.12.21/issmge/tc304.htm?=10>.



Figure 2. Group photo of TC304 student contest participants



Figure 3. Winners of TC304 Student Contest Award (from left to right: Shuo Zheng, Qinxuan Deng, and Yuxin Zhu)



Figure 4. Winners of Encouragement Awards

Young Member Arena (Con't)

Inaugural ISSMGE TC304 Student Contest on Data Analytics

As a follow-up, a three question survey were sent to winners (Shuo Zheng, Qinxuan Deng, and Yuxin Zhu) to get their feedback about their experiences on the contest. The questions and answers from the winners are provided in the appendix of this article. Besides, they also indicated that “It is of our great honor to participate in this student contest. We cannot deny that the contest is a hard-fought campaign for us, especially considering the long-time span in three months and many other excellent teams. We are glad that our team performed well and were solidly united during the contest. We worked hard and exerted our all-out intelligence and compatibility to the limit. For sure, the process helped us a lot in terms of widening our vision and expanding our knowledge boundary. We are also grateful for contest organization committee and award committee for their efforts on organizing this contest and are thankful for knowing friends from other teams during the contest.”

Appendix:

Q1. What did you learn from the contest?

Winners: First of all, the contest made us realize the importance of data analytics and its value in dealing with data from real world, such as 304dB complied by TC304. Besides the technical parts, this contest taught us the importance of courage and confidence in conducting research. By independently working on the contest question, we start thinking about the process of research when confronted with a new topic we did not work on. That is a great training process for us.

Q2. How will this affect your approach to uncertainty in geotechnical engineering?

Winners: It is well-recognized that geotechnical data may fluctuate due to various reasons. We hope that the method we proposed can screen out the outliers in advance so as to obtain more meaningful statistical analysis results from geotechnical data. This may help understand the information provided by geotechnical data in a more precise way.

Q3. What is your opinion of machine learning in engineering?

Winners: Machine learning and its increasing applications in engineering may change a lot in future. Machine learning gives a more general framework to analyze various types of data obtained in engineering by providing a plenty of advanced probabilistic models and efficient statistical inference algorithms. This allows us to acquire more useful information and values from data, and may change our decision-making process in design and retrofitting of engineering structures.