Qeios

Peer Review

Review of: "Uncertainty-Driven Modeling of Microporosity and Permeability in Clastic Reservoirs Using Random Forest"

Mohammad Reza Rasaei[1]

1. School of Chemical Engineering, Institute of Petroleum Engineering, University of Tehran, Iran, Islamic Republic of

The topic is of scientific and practical importance. The manuscript needs major improvements in its different aspects and sections before it can be published:

1- At the end of the Introduction section, the research gap and the proposed solution should be clearly explained to justify the current research despite all the previous ones.

2- Microporosity measurement and its distinction from total effective porosity are not clearly addressed in the Data Collection section. It is important to explain the methodology and reliability of the collected data in this section. How does it compare with more advanced techniques such as 3D micro-CT scans?

3- A major challenge in tight rocks is permeability measurement. Again, this important aspect is not fully explored. How did you make sure non-Darcy effects are accounted for in this section?

4- The methodology of Uncertainty Analysis is not clear to me. Based on the collected data, you will get a specific probability distribution for each parameter over which the machine learning model will be trained. How could you identify the uncertainty level of these parameters? What is your methodology to consider the associated uncertainty for each parameter? What do you mean by the resampling process that increased your data point measurements from 308 to 61300?

5- Why is 90% of the data used for model training instead of the routine figure of 70%? Discuss the impact and model performance.

6- Model evaluation is performed by investigating the correct prediction of porosity in two categories and permeability in three classes. While this is a good step, it is not addressing the main challenge of capturing the porosity-permeability correlation. From a reservoir characterizing point of view, we need

 doi.org/10.32388/8IGCBR

to categorize and predict the main rock types with distinct porosity-permeability correlations. I would like to see how the model can do this critical job.

7- Your argument about improving model performance by incorporating data uncertainty is not robust. You need to show that increasing the number of measurements in the data set from 308 to more than 61000 is not the main reason!

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/8IGCBR