Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Viktor Duplyakov

PhD Program: Petroleum Engineering

Title of Thesis: Machine learning on field data for hydraulic fracturing design optimization

Supervisor: Professor Andrei Osiptsov
Co-supervisor: Professor Evgeny Burnaev

Name of the Reviewer: Egor Dontsov

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: 07-09-2023**

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

This dissertation employs machine learning approach to address the problem of hydraulic fracture design optimization. Typically engineers employ hydraulic fracturing simulators to optimize the design. This involves running to code several times to find the best possible scenario. Unfortunately, this is a high effort task that is often very time consuming. Therefore, alternatives are always welcome. One such alternative is presented in this dissertation. In particular, data from thousands of wells is collected and analyzed from the statistics point of view. Then, machine learning is used to construct a quick model, which is in turn used to find optimal designs for a series of test wells. The results are compared with actual data and show an overall good degree of consistency.

I think that this dissertation provides an excellent contribution to the field. I have a couple of bigger comments and quite a few very minor ones.

1. It is quite apparent that most of the effort is spent on data mining, which is a crucial step to success. However, from reader's perspective, the emphasis should be on the results. I found that the first part of the thesis is written better that the second one that deals with the results. It is harder to follow and the key points are not well emphasized. One reason, perhaps, is that figure captions are too cryptic. In other words, it is practically impossible to understand what is plotted on the figures by just reading captions. The reader needs to go back and forth to the text in order to have a full picture. To sum up, the first major comment is to make chapter 4 more clear and to better highlight the main results that are relevant to field applications.

2. The second issue, which is partly discussed in future work, is the coupling with economics. I think that there should be at least a qualitative discussion on how economics affects the result. Because otherwise, the answer is the bigger the better. The more fracs are out there and the bigger the fracs are, the more oil they are going to drain. There are no constraints. And economics provides these constrains. Here is an example how this is done in ResFrac: https://www.resfrac.com/blog/resfracs-automated-economic-optimization-tool. Once you add economics, then typically there is a clear maximum or optimal value.

Minor comments:

1. List of your publications and conference proceedings: use the same format for all entrees.

2. "Pipeline presented in this study" – consider rewording as "The approach presented in this study".

3. "The activation of the natural fractures network by hydraulic fracturing is a key issue in the commercial production of shale reservoirs." - this was a line of thought a few years ago and it was based mostly on microseismic results. This applies to US at least. Right now, there is a lot of evidence from fiber optic measurements in the offset wells that fractures are predominantly planar in shales. That's why nowadays people rarely discuss stimulation of natural fractures.

4. “special dimensionality reduction” – did you mean “spatial”?

5. “yields an maximum” – typo.

6. “good“ – check all quotation marks for correct formatting.

7. “euclidean distance” – I think “e” should be capitalized, at least in english version. Fix throughout the whole thesis.

8. Figure 4-2. Update caption so that readers can understand what is actually plotted. Is it prediction of production for various parameters for two models? Please check all figure captions to make sure that it is possible to understand what is plotted by just reading the caption.

9. Regarding the field test described in 4.2.5. How different were the parameters from the training set? Can they be plotted somehow on the parametric diagram?

11. Idea for future work – multiple stacked wells. This is at least where things are in the US.

12. Make sure that all references at the end of the dissertation are formatted to the same style. Also fix “booktitle=SPE Oil others”.

Provisional Recommendation

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate’s thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*