PROTOTYPING VISUALIZATIONS AS A SUPPORT FOR SELECTING REPRESENTATIVE MODELS OF PETROLEUM RESERVOIRS

Miguel Angelo Mena Póvoa, Samuel Oliveira da Silva and Celmar Guimarães da Silva
School of Technology, University of Campinas, Brazil

ABSTRACT

Petroleum engineers usually create a set of hundreds of models of a given oil reservoir under analysis to represent its uncertainties. Assisted optimization approaches may help engineers to select a subset of these models (a.k.a. representative models, or RMs for short), which is used in computational flow simulations in replacement of the original set, aiming to reduce the total simulation runtime without changing the quality of these results. Despite the power of visualization techniques to help people to understand multidimensional datasets like the ones provided in this scenario, we noted a few efforts that use these techniques to help petroleum engineers to assist the selection of RMs. In this context, our research aims to test the hypothesis that it is possible to improve how interactive visualization resources are currently used to aid decision-making regarding the selection of RMs, mainly in the presence of multiple sets of RMs or multiple variables (obtained from running model simulations). This work presents our first steps towards this goal: (a) literature review and (b) definition of visualization prototypes that aim to help the analysis of RMs regarding the values of the variables provided by simulation outputs, and the risk curves associated with these variables. As preliminary results, we present our proposed interactive visualizations and briefly point out the design rationale behind these prototypes.

KEYWORDS

Information Visualization, Oil Reservoir Model, Risk Curves, Cross Plots, Multivariate Data, Representative Models

1. INTRODUCTION

Oil reservoir development and management activities involve risks due to various physical, operational, and economic uncertainties (Schiozer et al., 2019), which are associated with the recoverable reserves of the oil wells and the flow characteristics, the availability of the production system, the oil price, investments, and operational expenses, among others. Besides, these activities also involve defining a production strategy, which consists of the sequencing of drilling, opening, injection and positioning of wells, among other activities. Due to these intrinsic uncertainties, defining this kind of strategy becomes a complex problem, whose solution embraces the analysis of numerous models that sample the possible combinations of the uncertain reservoir characteristics under analysis. Model reduction techniques are highly desirable, in order to reduce the time-expensive execution of flow simulations (up to days) for each possible production strategy and each model under analysis. This reduction should produce a viable set of models that is representative of the entire problem (Meira et al., 2016).

In this context, the RMFinder methodology and software (Meira et al., 2020) use an optimization-based approach proposed to automatically identify, according to the analysts' preferences, subsets of oilfield models, called Representative Models (RM), that represent a set of originally provided models. Although it is possible to select a set of RMs (a.k.a. "solution") in an automated way, the selection of RMs is an assisted optimization procedure, whose results must be inspected by petroleum analysts, which verify if the selected RMs appropriately represent the complete set of models. The RMFinder output sets of cross plots (which resemble scatterplot matrices) and charts with risk curves (a.k.a. complimentary cumulative distribution functions, or CCDFs) related to variables of the models, so that analysts could visually analyze the representativeness of a solution and make decisions such as accepting a given solution from a set of solutions, or rerunning RMFinder with distinct initialization parameters (e.g., with preselected models as RMs).

Despite these initial charts provide some support to analysts, we consider that there is a gap related to the almost absence of interactive visualization resources being used to support decision-making regarding the selection of RMs provided by the RMFinder technique. We focused our efforts to fill this gap in our current research, which aims to propose and improve the use of interactive visualizations to enhance this decision-making. Our focus is on the analysis of output variables of the models and their risk curves. We aim to help analysts to define if a given solution is really representative, and to compare solutions regarding this representativeness.

In this paper, we present our first steps regarding this goal. In Section 2, we briefly discuss the state-of-the-art on the use of visualizations for the selection of RMs. In Section 3, we describe our methodology, present our prototypes and point out their support to improve decision-making. The last section concludes the paper and points to future work.

2. RELATED WORK

We found some optimization-based approaches for the selection of RMs. On the other hand, we noted that few Information Visualization (InfoVis) techniques have been used in recent years as support for the analysis of one or more sets of RMs. Meira et al. (2017, 2020) used three visualization techniques to allow the user to check the quality of a solution: a set of scatter plots (cross plots) to allow visual analysis of the representativeness of RMs regarding each pair of output variables; superposed or juxtaposed risk curves of multiple variables and solutions; and bar charts to discuss the differences of the discretization of continuous attributes. Amaral (2021) added interactions to these views, such as zooming, brushing, selecting a specific model, or expanding a selected chart with model details.

Some works use multidimensional projection and clustering techniques to support the selection of RMs. Anupam and Tewar (2021) use PCA, t-SNE and k-means to select RMs. Zheng et al. (2018) and Sahaf et al. (2018, 2019) apply MDS and k-means to represent the influence of production uncertainties and select RMs. Regarding risk curves, Silva et al. (2020) used reorderable heatmaps to show the dissimilarities of risk curve pairs (e.g., area between risk curves).

Besides, Silva et al. (2019) used a *small multiples* layout of spatial heatmaps of model maps clustered by similarity of spatial petrophysical variables. Highlighting RMs enable analysts to contrast these RMs and the defined clusters of model maps, given that RMFinder does not deal with spatial data to select RMs.

3. METHODOLOGY AND PRELIMINARY RESULTS

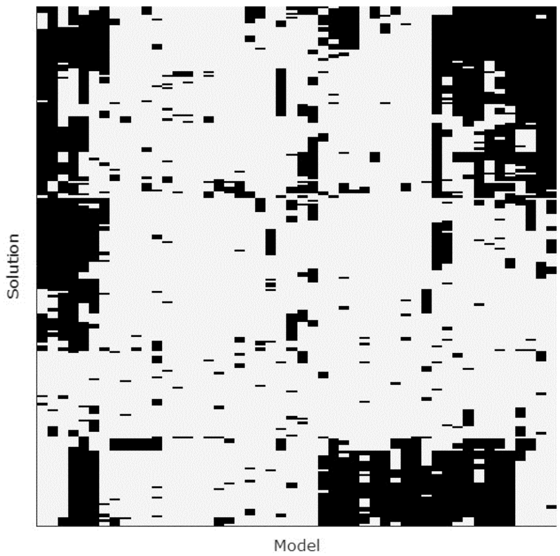
To develop this user-centered information visualization project, our methodology is inspired by the iterative development cycle presented by Kulyk et al. (2017). It comprises the following steps: users, data, and tasks characterization; designing a visualization method; implementation of visualization methods; and reviewing and testing the visualization methods (return to the first step if needed). We added a preliminary step, in which we hypothesized, with support of the experience of our research group, the characteristics of the RMFinder users, and the main tasks they want to accomplish to better select RMs or compare sets of RMs. We also prototyped visualizations that aim to provide insights to these analysts. Through interviews with users, we will validate our point of view regarding their tasks and the usefulness of our prototypes.

The analysts are engineers and students, within a broad spectrum of ages (approximately between 22 to 65 years old) and years of experience in reservoir engineering (2 to 20 years). We expect that they had previous contact with RMFinder and with the problem of selecting RMs. Their main tasks, as far as we aware, can be described as: (a) analyzing the representativeness of a solution; and (b) comparing a set of solutions to find the most representative one. Understanding the so-called representativeness is a challenge because it has explicit and tacit components. The explicit components are those already defined in the RMFinder optimization functions. The tacit components, as such, must be obtained by provoking reflections from the analysts – a task that we plan to accomplish when discussing our prototypes with them.

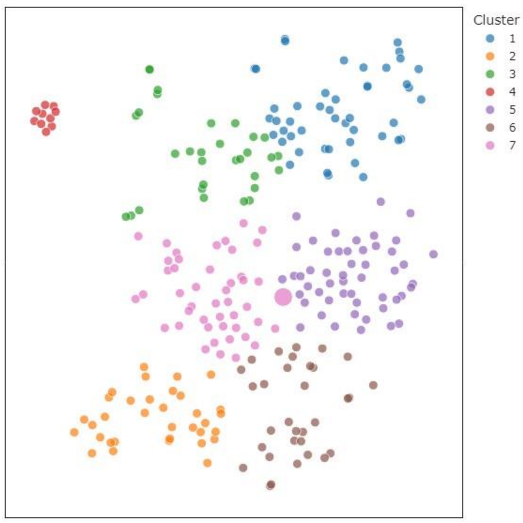
Our dataset may be defined as follows. The RMFinder's input dataset comprises a set M of models. Each model has values for a set V of variables, such as net present value (NPV), oil production (Np), water production (Wp), and oil recovery factor (ORF). The RMFinder defines one or more solutions to its optimization problem, where each solution S_i is a subset of M. For each S_i, the RMFinder tries to minimize a function that considers cross plot-based and risk curve-based representativeness. The *two variable-based representativeness* (2VR) of a solution is the sum of the quadratic distances between each RM and its nearest models in every possible cross plot (w_i, w_j), where $\{w_i , w_j\} \subset W$, and $W \subset V$. This concept reproduces the analysts' praxis when evaluating cross plots (i.e., the shorter the distances, the better the representativeness). The *risk curve-based representativeness* (RCR) of a solution S_i for each variable $v_k \in V$ is defined as a measure of the similarity between the risk curve for M and its risk curve for S_i (i.e., the closer the curves, the better the representativeness). RCR may be considered as the area between a pair of risk curves (as in the work of Meira et al. (2020)) or as the two-sample Kolmogorov-Smirnov statistic (a.k.a. K-S) of a pair of risk curves (Mahjour et al. (2020)). Also, it is worth providing a way to visually analyze the *variability* of a given set of solutions regarding the set of RMs they use.

Hereafter we briefly list a subset of our visualization prototypes (Figure 1 and 2):

- A binary, interactively reorderable heatmap of models used by each solution, so that the analyst may find relevant patterns in sets of solutions regarding their variability, such as common subsets of RMs (Figure 1a).
- A multidimensional projection of risk curves of the solutions in the set $\{M, S_1, \ldots, S_n\}$, regarding a distance measure based on the RCR related to a set of interactively filtered variables. Supported by clustering after projection, this visualization shows the similarity of the solutions and their representativeness regarding M (Figure 1b).
- A solutions-by-variables heatmap, such as the one from Silva et al. (2020), whose cells present RCR regarding the risk curve of a variable V_k for a solution S_i (e.g., the area between risk curves, or the K-S measure between risk curves) (Figure 2a).
- A plot of the risk curves (one per variable) of M enhanced with confidence bands that summarize sets of risk curves from multiple solutions, so that analysts can better analyze RCR and features of the set of risk curves (Figure 2b).
- An overview heatmap whose cells' colors represent the 2VR per solutions (Figure 2c). A second version of this heatmap uses dimensional stacking to detail the 2VR measure also per RMs (Figure 2d). These visualizations should enable analysts to quickly understand the quality of the presented solutions, as well as perceive possibly incorrect choices of RMs.
- Cross plots comparing models and RMs of a single solution regarding pairs of variables, enhance with convex hulls that outline each RM and the models next to it, aiming to provide visual cues about poorly represented models regarding 2VR (Figure 2e). The RMs are highlighted in bold.

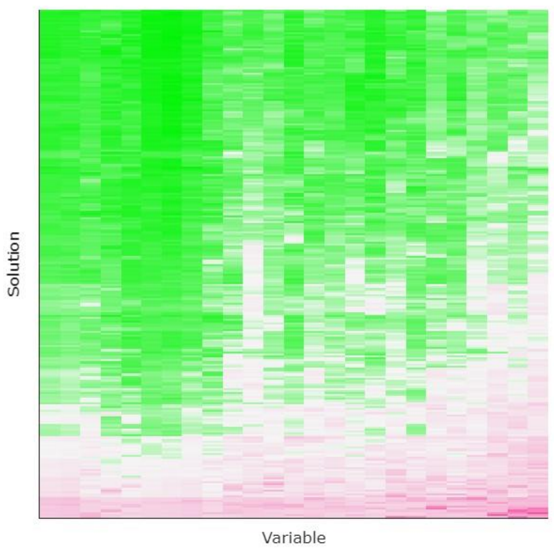


(a) Reorderable heatmap of models by solutions

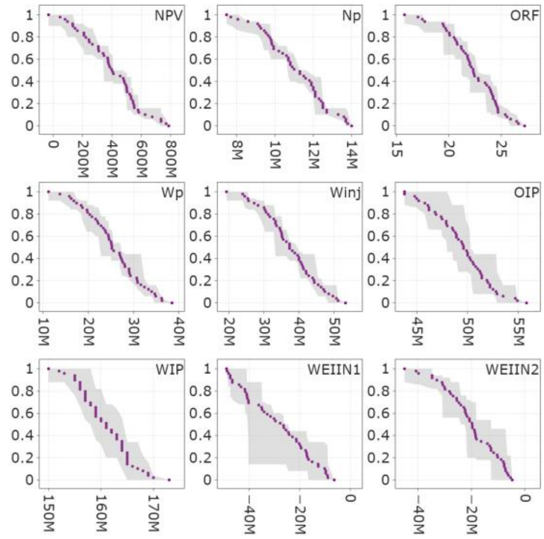


(b) Multidimensional projection of risk curves.

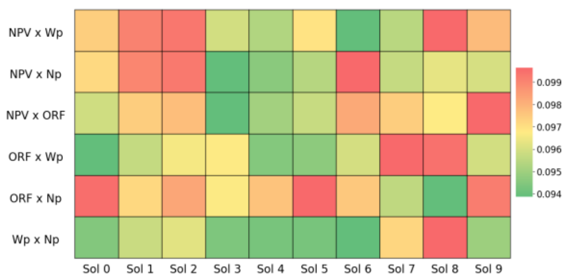
Figure 1. Visualization prototypes using data from the Olympus (Fonseca et al., 2020) and UNISIM-I-D (Gaspar et al., 2015) case studies



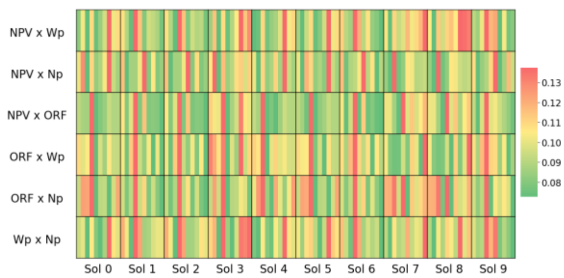
(a) Reorderable heatmap solutions by variables



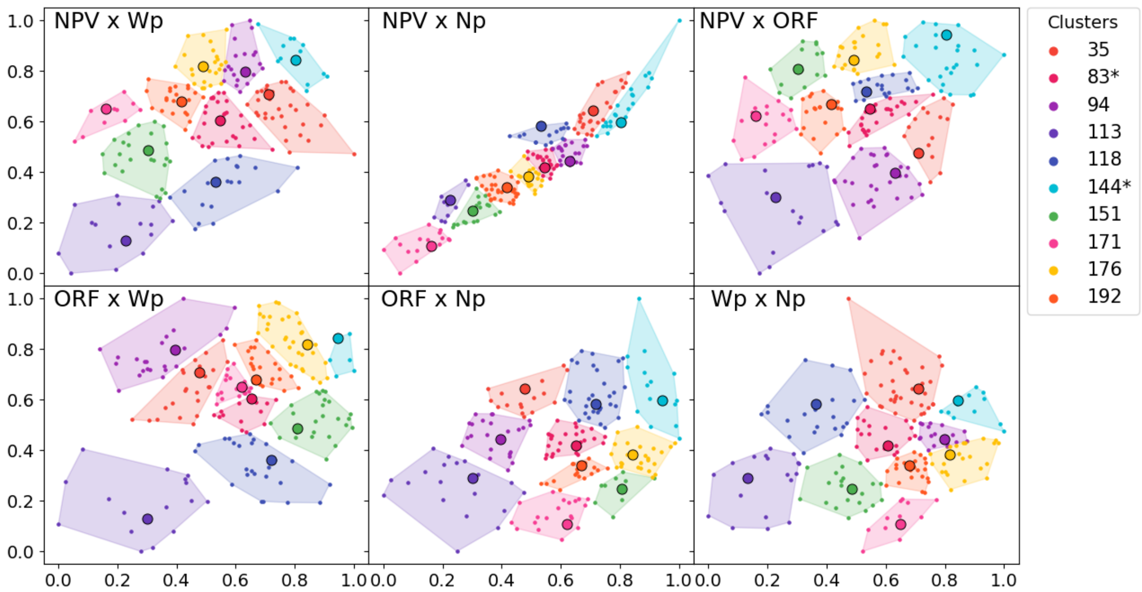
(b) Risk curves with confidence bands.



(c) Representativeness heatmap by solutions and pairs of variables.



(d) Representativeness heatmap by solutions, pairs of heatmaps and RMs.



(e) Cross plots of pairs of variables with convex hulls of clusters for each RM.

Figure 2. Visualization prototypes (continuation of Figure 1)

4. CONCLUSIONS

This paper presented our first efforts to aid reservoir engineers using visualizations. We aim to aid the engineers in better selecting representative models based on representativeness and variability criteria applied to single or multiple sets of RMs. We briefly explained and exemplified our proposed multidimensional visualizations for analyzing problem variables and their respective risk curves. Future works include discussing with users how these prototypes fit their needs regarding these tasks and evolving them in that direction.

ACKNOWLEDGEMENT

We gratefully acknowledge the support of Energy Production Innovation Center (EPIC), hosted by the University of Campinas (UNICAMP) and sponsored by Equinor Brazil and FAPESP – São Paulo Research Foundation (2017/15736-3). We acknowledge the support of ANP (Brazil's National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation. Acknowledgements are extended to the Center for Petroleum and Energy Studies (CEPETRO), School of Mechanical Engineering (FEM), and School of Technology (FT). This study was partially financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

REFERENCES

Amaral, L. D. B. *Visualização de informação aplicada à análise de resultados de processo de seleção de modelos representativos de reservatórios de petróleo*. Undergraduate final project dissertation. University of Campinas. 2021.

Anupam, A.; Tewari, R. Easy Steps in Probabilistic Dynamic Modelling in FDP. In: *Int. Petroleum Technology Conf*. 2021.

Fonseca, R. M. et al. Introduction to the special issue: Overview of OLYMPUS Optimization Benchmark Challenge. *Computational Geosciences*, 24(6), 1933-1941. 2020.

Gaspar, A. T. F. S. et al. UNISIM-I-D: Benchmark Studies for Oil Field Development and Production Strategy Selection. *Int. Journal of Modeling and Simulation for the Petroleum Industry*, v. 9, n. 1, p. 21-30, 2015.

Kulik, O. et al. Human-Centered Aspects. In: Kerren, A. et al. (Ed.). *Human-Centered Visualization Environments*. Berlin: Springer, 2017. p. 13–75.

Sahaf, Z. et al. A visual analytics framework for exploring uncertainties in reservoir models. In*: VISIGRAPP 2018-Proceedings of the 13th Int. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications*. Funchal, Madeira, Portugal: SciTePress, 2018. v. 3, p. 74–84.

Sahaf, Z. et al. An ROI visual-analytical approach for exploring uncertainty in reservoir models. In*: Communications in Computer and Information Science*. Funchal, Madeira, Portugal: Springer Verlag, 2019. v. 997, p. 121–142.

Silva, C. G. et al. Visualization of Ensembles of Oil Reservoir Models Based on Pixelization, Small Multiples and Reservoir Similarities. In: *Offshore Technology Conference Brasil*. Rio de Janeiro, BR: Offshore Technology Conference, 2019.

Silva, C. G. et al. Using Reorderable Matrices to Compare Risk Curves of Representative Models in Oil Reservoir Development and Management Activities. In: *Proc. 14th Int. Conf. on Computer Graphics, Visualization, Computer Vision and Image Processing. IADIS*, 2020. p. 19–26.

Schiozer, D. J. et al. Model-based decision analysis applied to petroleum eld development and management. *Oil and Gas Science and Technology*, Editions Technip, v. 74, 2019.

Mahjour, S. K. et al. Developing a workflow to select representative reservoir models combining distance-based clustering and data assimilation for decision making process. *Journal of Petroleum Sci. and Eng.*, Elsevier B.V., v. 190, 2020.

Meira, L. A. A. et al. Selection of Representative Models for Decision Analysis Under Uncertainty. *Computers and Geosciences*. Elsevier B.V., v. 88, p. 67-82, 2016.

Meira, L. A. A. et al. RMFinder 2.0: An Improved Interactive Multi-Criteria Scenario Reduction Methodology. In*: SPE Latin America and Caribbean Petroleum Eng. Conference. Society of Petroleum Engineers*. 2017.

Meira, L. A. A. et al. Improving representativeness in a scenario reduction process to aid decision making in petroleum elds. *Journal of Petroleum Science and Engineering*, Elsevier B.V., v. 184, p. 1–19, 2020.

Zheng, J. et al. A Cluster-Based Approach for Visualizing and Quantifying the Uncertainty in the Impacts of Uncertain Shale Barrier Configurations on SAGD Production. In: *SPE Canada Heavy Oil Conference*. 2018.