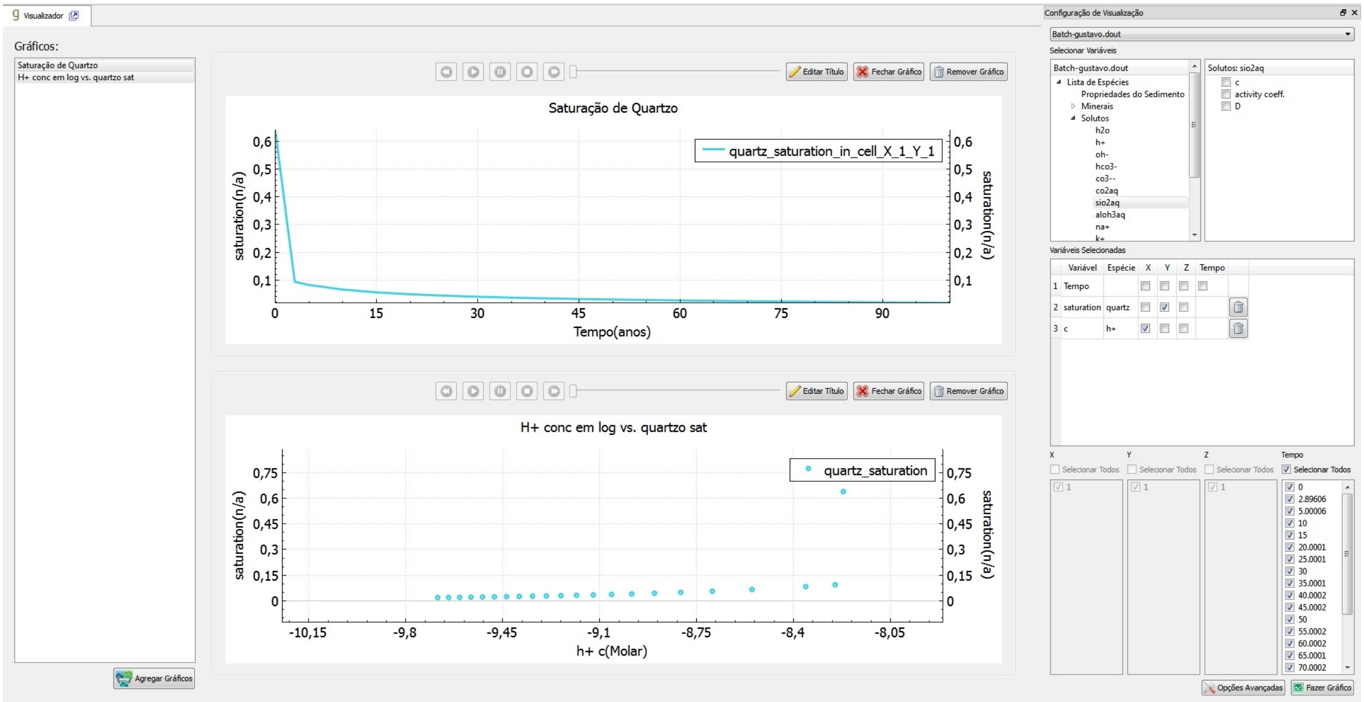
DiagenViz: Interactive Analysis of Simulation Results

Gustavo Jandt Feller[1], Marcos Antônio Klunk[2] and Carla Maria Dal Sasso Freitas[1]

[1]Institute of Informatics, Federal University of Rio Grande do Sul
[2]Institute of Geosciences, Federal University of Rio Grande do Sul



	Variável	Espécie	X	Y	Z	Tempo
1	Tempo		☐	☐	☐	☐
2	saturation	quartz	☐	☑	☐	
3	c	h+	☑	☐	☐	

Fig. 1. Overview of the DiagenViz interface. Left panel: list of plots shown in the central panel. Right panel: configuration options allowing selection of variables to be plotted, assignment of variables to X, Y, and Z axes, and selection of time steps. Center top: line plot of quartz saturation per time. Center bottom: scatterplot relating quartz saturation to hydrogen concentration.

***Abstract*—Diagenesis comprehends chemical, physical and biological processes related to rock sediments in a reservoir. The simulation of these processes can tell the history of the reservoir, and can suggest the conditions that the oil went through, predicting the quality of the reservoir. This work shows a tool (DiagenViz), which aims at helping geologists to analyze results obtained from simulations of diagenetic processes. We discuss the data produced by the simulator as output and our approach of a graphical user interface to present and analyze the data. We also compare our tool with the commercial softwares GWB and PetraSim, which are used by most geologists due to the lack of specific diagenesis simulation software.**

I. INTRODUCTION

Diagenesis is defined as the set of chemical, physical and biological change processes through which the rocks sediments pass since its deposition, during and after lithification, and before the metamorphic conditions. The diagenetic processes are controlled by factors such as temperature, pressure, minerals, activity of the ions dissolved in water and organic systems [1]. These processes are active, and the sedimentary minerals react to restore equilibrium in an enviroment where pressure, temperature and chemical composition are changing. The reactions in the system can increase or decrease permeability and porosity [2]. All these processes correspond to the formation of the present rocks, and they occurred along millions of years. A geologist studying diagenesis usually wants to understand the processes that have occurred during that time, as well as factors that may have influenced the oil quality of a determined region. So, simulations are run to test hypotheses about how an oil reservoir formed in the past, and ultimately, to determine its quality.

The goal of this work is to develop an interactive analysis tool that implements a set of visualization techniques for supporting geologists in their analysis of simulations results.

To the best of our knowledge, there are no commercial simulators specific to simulate diagenetic processes. Simulators developed for other purposes are used for this task, and their results, although approximate, are accepted by geologists. There are some works in the literature describing models for diagenesis, as Boudreau [3] and Park and Ortoleva [4], but none of them treat the results visually.

In the next section, we present some background information to make the simulation process understandable, and also highlight relevant related works. Then, we present our tool for the visual analysis of simulation results (section III). We also report an evaluation based on the comparison of our tool with two commercial systems, GWB [10] and PetraSim [11] (Section IV). Finally, in Section V, we draw some conclusions.

II. BACKGROUND AND RELATED WORKS

A. *Geochemical Modeling*

Before addressing the simulation process itself, we need to introduce concepts from geochemical modeling. Geochemical modeling is only useful as a forecasting tool if there is possibility of validating the results. In real life, this is the goal that most often become unachievable because of the complexity of natural systems, insufficient field data and uncertainties related to how a system will change along time. A model must be treated as a simplification of reality, and its precision is dependent on how it is capable of estimating the probability of a forecast to be true or false [5].

A general geochemical simulation process is roughly divided into 3 major stages: data input, simulation core and data output.

Data input consists of collecting information relative to the geological medium of interest, through chemical analysis made in laboratory and through stratigraphic data of the sedimentary basin. These data are (i) water composition, (ii) mineral composition, (iii) kinetics and thermodynamics reactions, (iv) burial history (depth of rock formation, estimated time to occur lithology transformations, pressure, temperature) and (v) spatial domain (batch in an one-cell domain, one-, two-, and three-dimensional domain). Data can be input through a script or filling in a form in a graphical user interface (GUI).

Data is entered to the simulation core, which starts the simulation execution steps. In this stage, numerical methods are used to solve geochemical equations of fluid-rock interaction in the geological medium defined in the input data. As the simulation process executes, system state is updated for each simulation step, and partial simulation results are generated. This process goes on until the system reaches a steady state or a user-defined maximum simulation time.

Data output is the last part of each simulation step. The data generated by the simulation execution is stored in a file, usually text. Each simulator has its own standard for input and output files.

B. *Related Works*

Related to our context, there are two well known simulators for geochemical modeling: Geochemist's Workbench (GWB)[10] and Toughreact [11].

In GWB, the user sets an initial geochemical system to be taken to thermodynamic equilibrium. The software automatically inserts a known volume of water in the system (1 kg). Then, the user sets the amounts of solutes present in that water. GWB starts the calculations and the necessary iterations that lead to a speciation model. When GWB finishes the simulation, output data is generated. Data contained in the output file are temperature, pressure, pH, ionic strength, water activity, mass of solvent, dissolved solids, solution density and mass of the rock. A list of aqueous species is also output with all solutes present in the simulation. An important indicator is the "Saturation Index – SI" of the fluid, which informs: (i) mineral and solution are in equilibrium; (ii) solution is supersaturated; and (iii) solution is under-saturated.

Toughreact can be used in one-, two-, or three-dimensional geological domains in heterogeneous physical and chemical environments, i.e., a wide range of conditions. Input files are provided through a GUI called PetraSim. Firstly, the user selects the solutes that will compose the aqueous phase, and then selects the lithology of interest composing the geological environment. Kinetics and thermodynamic parameters are adjusted after the user builds the interaction model. Once all requirements are satisfied, the software starts the simulation. Toughreact output data is generated basically to provide plots of the quantity of solute and volume variation versus simulation time. If the user wants to visualize saturation index, Toughreact generates text files that need to be exported to spreadsheets like EXCEL.

As for visualization, GWB provides the tool named Gtplot that allows users to display simulation results with 2D visualization techniques, such as line plots, pie charts, color maps, contour plots, vector plots and star plots. However, regarding diagenesis, only line plots are used. On the other hand, PetraSim [11] provides more visualization techniques like line plots, 3D iso-surface visualization, vector and contour 3D plots.

Nevertheless, considering specific diagenetic processes simulators, visualization is still an open problem. Early works like the model of carbon and nutrient diagenesis in aquatic enviroments [3] and the numerical model of sedimentary early diagenetic processes presented by Soetaert et al. [15], and relatively recent ones such as WRIS.TEQ [4], a simulator of diagenetic alterations of sediments composed of complex mineralogy and heterogeneity, still miss both a GUI to parameterize the model and visual analytics tools to better support geologists studies.

III. DIAGENVIZ

DiagenViz is implemented in C++, using the Qt Framework [6] for the GUI and two external libraries for visualization: QCustomPlot [7] for the 2D plots, and QwtPlot3D [8] for 3D plots. In this section, we firstly describe the data, then we explain how the possible visualizations are configured, and finally we describe the implemented visualization techniques.

A. *Data Description*

The data the tool analyzes comes from two output files generated by a simulator developed in an ongoing project at UFRGS. A header file contains the description of the variables in the output, and for each step there is a data file containing the values of the variables defined in the header file. Each variable in the header file is defined as a composition of textual information (Fig. 2).

[index] VAR [variable name] SPC [specie name] UNIT [unit] TYPE [type of the variable] IDX [index of the specie] LABEL [label name]

Fig. 2. A standard variable description in the header file.

There are four types of variables depending on to what entities they refer to: (i) sediment, (ii) element, (iii) solute and (iv) solid. The *sediment* type variables are related to the sediment itself as porosity, temperature, water velocity, and so on, being the only species of the type. *Element* variables are about some information of the quantity of each chemical element present in the related cell. *Solute* variables represent solute concentration in the water and activity in reactions, and each solute present in the system is defined in the data input, as H^+, HCO_3^-, etc. *Solid* variables are related to minerals, e.g. Quartz and Calcite, precipitation and dissolution, saturation, volume fraction and mineralization rate.

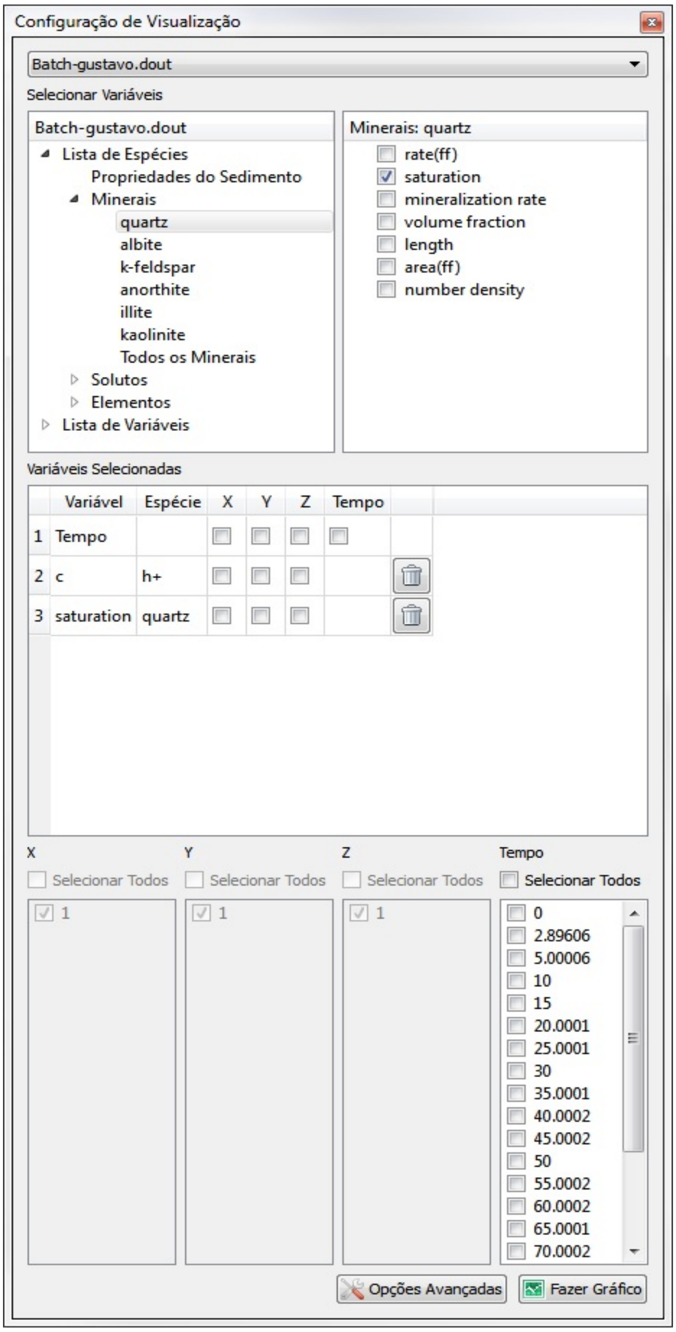
Using Munzner classification [9], the output data is a Field, because the variables (described in the header file) are associated to each cell at each time step. Variables associated with distance or time are considered continuous data, and variables associated to another variable is not continuous.

B. *Variable Selection*

In DiagenViz, the GUI is divided in two panels: the visualization panel (discussed in the next subsection) and the variable selection panel. After all data is loaded, the variable selection panel, which can be seen at the right, in Fig. 1, and in Fig. 3) is displayed. It is divided in three main parts : (A) Variable selection itself, (B) axis selection and (C) time and/or cell selection.

In (A), the user select the variable that he wants to analyze. A tree widget is used to select a species or a variable to be analyzed, depending on what the user wants to focus. In (B) the user selects the variable to be represented in each axis. By default, the plot will have at least one variable (time), but it may also have distances, depending of the dimension of the domain. Also, when the user selects one of the default variables for one of the axis, its list will fade out from (B), because it will be plotted. In (C) the user defines which time step or cells in the domain she/he wants to visualize, and this combines all selected cells and time steps the user has selected.

As an option, after setting (C), the user can define if the data on a specific axis is in linear scale or logarithmic scale. Also, the user can filter the data he/she wants to visualize, selecting the interval of values to be plotted.



☐ rate(ff)
☑ saturation
☐ mineralization rate
☐ volume fraction
☐ length
☐ area(ff)
☐ number density

	Variável	Espécie	X	Y	Z	Tempo
1	Tempo		☐	☐	☐	☐
2	c	h+	☐	☐	☐	
3	saturation	quartz	☐	☐	☐	

X: ☑ 1
Y: ☑ 1
Z: ☑ 1
Tempo: ☐ 0, ☐ 2.89606, ☐ 5.00006, ☐ 10, ☐ 15, ☐ 20.0001, ☐ 25.0001, ☐ 30, ☐ 35.0001, ☐ 40.0002, ☐ 45.0002, ☐ 50, ☐ 55.0002, ☐ 60.0002, ☐ 65.0001, ☐ 70.0002

Fig. 3. Configuration panel. Top: list of variables present in the output data file. Center: list of selected variables for axis assignment. Bottom: time step or domain selection.

C. *Visualization Techniques*

The visualization techniques provided by the tool were chosen based on user preference. Our users are the geologists that interact constantly with the simulator development team. The techniques are quite simple in terms of visualization, but they are based on those techniques that geologists are used to: (i) line plots, (ii) line plots with two Y axes, (iii) scatterplots and (iv) 3D surface plots. Line plots are used when one of the variables is continuous, i.e time or distance.

Line plots with two Y axes (Fig. 4) are used when the user selects two different types of variables for the Y axis, e.g saturation and volume fraction, and X axis depicts a continuous variable. Scatterplots (see Fig. 1) are used when the user wants to compare two non-continuous variables, to analyze their relation, e.g concentration of Ca^{++} and Calcite saturation. 3D surface plots (Fig. 5) are used when the user wants to analyze the variation of one variable per two others, as for example, Quartz saturation per time and distance. For all the techniques, we also implemented animation to allow display of the plots along the simulation time (the animation controls are at the top of each plot).

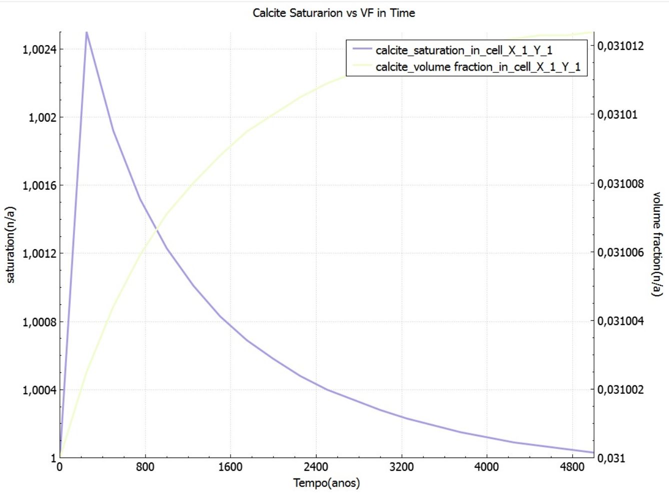


Fig. 4. Example of a two Y axes plot, comparing the saturation and volume fraction of Calcite during a simulation.

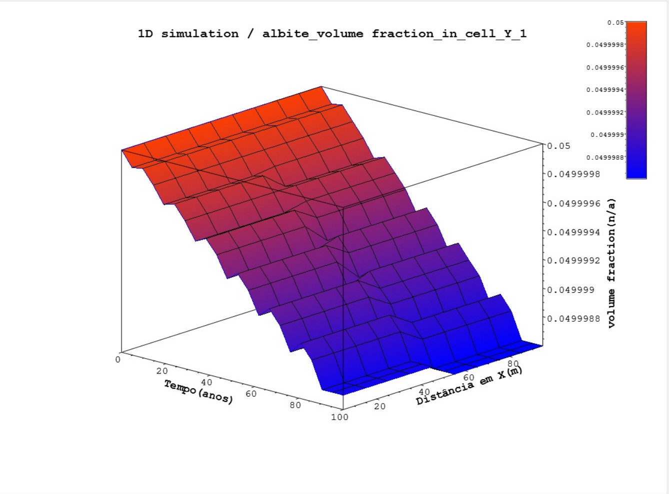


Fig. 5. Example of a three-dimensional plot, showing Albite volume fraction by time and distance.

IV. EVALUATION AND DISCUSSION

In this section, we compare our GUI and visualizations to those provided by two commercial systems, GWB [10] and PetraSim [11], briefly described before.

A. Evaluation Procedure

For the evaluation, we have one persona, which is the specialist in geochemistry that wants to study the rock formation in a certain basin. Due to the expertise requirement, we had only two users performing the evaluation. *User1* is a young specialist in geochemistry that frequently uses GWB and Petrasim/Toughreact for his research. *User2* is a senior specialist in geochemistry that does not use any of these tools, since he has developed his own simulator. Since User1 is acquainted to GWB and Petrasim, he used DiagenViz and answered the questionnaire. User2 used all the three tools (GWB, Petrasim/Toughreact and DiagenVIz), and after that answered the same questionnaire. So, User1 provided us the impressions of a well-trained user in the tools, while User2 told us about his first impressions.

The questionnaire was conceived with 7 questions for each tool, evaluating several aspects like, for example, the selection of which variables to plot and the interactive tasks that can be performed with the plots. Six questions are in the form of affirmative statements to which the users have to agree/not agree in a 5-point Likert scale [14], while in the last one, the users were asked to provide free observations. The six statements are presented in Table I.

Statement	Question/statement
S1	Variable selection is intuitive
S2	Variable selection is fast
S3	The tool allows me to select any variable for visualization
S4	I can easily set the intervals of any variable to obtain new plots
S5	I can easily understand the behavior of variables looking at the plots
S6	I can easily interact with the plots to explore data

TABLE I
DEFINITION OF STATEMENTS

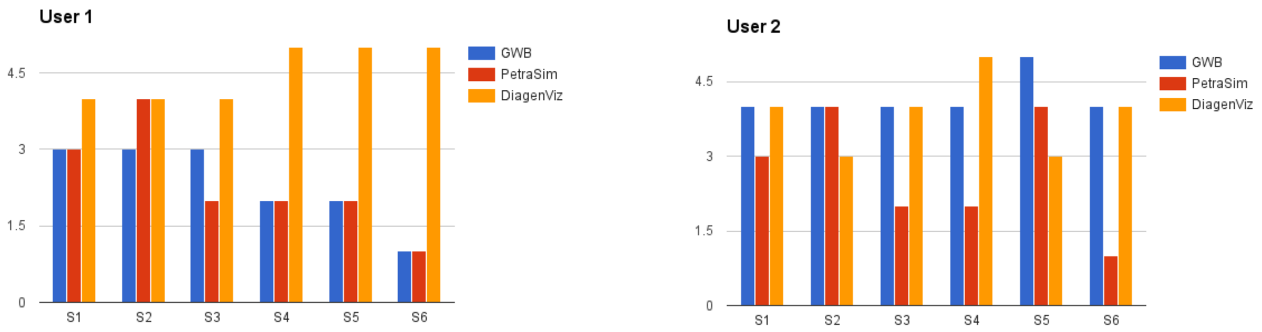
Another consideration is that, although our visual analysis tool covers 2D simulations, we have used only Batch and 1D simulations for evaluation.

B. Evaluation Results

Considering the results in Fig. 6, we can observe that User1 and User2 responded only slightly differently about the three systems regarding S1, S2 and S3 statements, while for S4, S5 and S6 their opinions are very different.

In S1, User1 agreed with variable selection intuitiveness in our tool, while he neither agreed nor disagreed with GWB and PetraSim, but User2 agreed with GWB and DiagenViz intuitiveness and neither agreed nor disagreed with PetraSim. User2 could have been influenced by the limitations of PetraSim in plotting different variables, while GWB and DiagenViz provide more freedom.

Regarding S2, they both agreed that PetraSim is fast in selecting variables (considering that it does not have much in selecting variables), while they had different opinions when it comes to GWB and DiagenViz. User1 found DiagenViz



(a) User1 Evaluation

(b) User2 Evaluation

Fig. 6. Evaluation results: plots indicate the user's level of agreement with the statements listed in Table I.

faster than GWB, while User2 thought the opposite. This could be due to the difference in experience in using the two systems. Since User1 is trained, the DiagenViz approach to select variables is more "automatic" than in GWB, while first time users like User2 need to think more about how selecting a desired variable.

As for S3, they disagreed with the freedom of selecting variables of PetraSim, probably because PetraSim does not have really a variable selection interface, as GWB and DiagenViz. They both agreed with DiagenViz freedom, but, for GWB, User1 neither agreed nor disagreed, while User2 agreed. Considering the training that User1 had, probably he found situations that GWB limited him in some aspect that DiagenViz did not.

From S4, we start to have the most different opinions. Both users strongly agreed with the scaling of DiagenViz, and they both disagreed for PetraSim, but, for GWB, User1 disagreed and User2 agreed. One hypothesis that we have is that User1 considered the possibility of zooming in DiagenViz, and he didn't considered the scaling in GWB helpful, while User2 considered the scaling in GWB helpful. So, for a better conclusion, we need more tests.

In S5, a larger opinion difference was found: User1 disagreed with the understandability of plots in GWB and PetraSim, and strongly agreed for DiagenViz. However, User2 strongly agreed in GWB, agreed in PetraSim and neither agreed nor disagreed in DiagenViz. User2 probably had some difficulty to understand the two Y axes plot, because it is different from the plots geologists use.

Regarding S6, other larger difference: User1 strongly agreed with the interactivity of the plots in DiagenViz, and stronly disagreed for GWB and PetraSim, but User2 agreed for DiagenViz and GWB and strongly disagreed for PetraSim. Considering that DiagenViz has zooming and panning interactions in the plots, but GWB and PetraSim plots are static when trying to interact with them, this was evident for User1. Probably, User2 could think that, in many cases, this kind of interaction is not necessary.

In general, PetraSim got the worst results in most of the aspects. Both users complained about the limitation of the plots in PetraSim, because they could only visualize solute concentration and mineral volume fraction, but not saturation index (as described in section II), which is one of the most important variables for diagenesis studies. Two good points noticed by User2 about PetraSim are the speed and simplicity in plotting the available variables, even it is not complete. Another drawback was catched through statement S6, since both users answered that PetraSim does not give to users ways to explore data. This may result from the lack of variables to be shown. Considering the agreement levels related to all statements, we can suggest that a tool of diagenesis processes simulation needs to allow the user to visualize all variables in the output files.

In comparing GWB and DiagenViz, we did not find a consensus as in PetraSim. We can notice that User1 prefers DiagenViz rather than GWB, but as for User2, his answers were almost equal, with GWB better than DiagenViz in 2 statements (S2 and S5), and DiagenViz better than GWB in 1 statement only (S4).

As User1 is more trained than User2 in the use of these tools, he may had considered cases where he needed features that GWB does not have, but DiagenViz has. User2 had more a first impression of these tools. So, we can consider that, for a first time user, both tools are good, but GWB is simpler than DiagenViz, since one can set the variables to plot faster and in a more understandable way, while DiagenViz is better in interactivity.

These results allowed us to come up with a new hypothesis to prove: is DiagenViz better for trained users than other tools? Since the simulator behind DiagenViz is still under development, this is left for future work.

V. FINAL COMMENTS

Considering the goal of this work, that was developing an interactive analysis tool for geologists testing their hypothesis, we can conclude that we quite achieved that. According to the results shown in Fig. 6, our tool obtained better scores than one of the most used tools by geologists, while it is comparable to other well-known one, considering that GWB and PetraSim conquered the simulation market satisfactorily. The differences between DiagenViz and the other two tools are the better interaction features provided to users, the possibility of plotting and presenting multiple charts at the same time, and the possibility (that still need to be tested) of having better options to analyze 2D simulations.

As an improvement, comparing to the other tools, we need to implement a faster variable selection method, since ours requires many steps. We need to automate some steps, considering what the user may want to visualize, probably separating users in categories and assuming which variables they would need to visualize more often, or use a machine learning approach to suggest variables that they may want to visualize. Other aspect to improve is the selection of items for the axes, which was reported as confusing by one of the users.

Other possibility that we discussed for improving DiagenViz is the display of the simulation domain associated with the variables plots.

ACKNOWLEDGMENTS

We thank our users for the patience in helping us to improve the tool during the development, and for providing new insights during the evaluation phase. We also acknowledge the support from CNPq, PFRH-PETROBRAS and UFRGS.

REFERENCES

[1] L. F. De Ros, Composition controls on sandstones diagenesis: compr. summ. Uppsala Diss. Facul. Sci. Tech. 198:1-24. 1996.

[2] R. H. Worden, S. D. Burley, "Sandstone Diagenesis: The Evolution of Sand to Stone," in Burley SD and Worden RH (eds): Sandstone Diagenesis: Recent and Ancient. Malden, Massachusetss, USA: Wiley-Blackwall Publishing, International Association of Sedimentologist Reprint Series, vol. 4: 3-44. 2003.

[3] B. P. Boudreau. A Method-of-lines code for carbon and nutrient diagenesis in aquatic sediments. Computers & Geosciences, Great Britain, Vol. 22, No. 5, p 479-496. 1996.

[4] A. J. Park, P. J. Ortoleva. WRIS.TEQ: multi-mineralic water–rock interaction, mass-transfer and textural dynamics simulator. Computers & Geosciences, Vol. 29, pp 277-290. 2003

[5] D. K. Nordstrom, On the evaluation and application of geochemical models, Appendix 2 in Proceedings of the 5th CEC Natural Analogue Working Group and Alligator Rivers Analogue Project, an international workshop held in Toledo, Spain, October 5-19, 1992, EUR 15176 EN, pp. 375-385. 1994.

[6] Qt Framework [online]. Available: http://www.qt.io/ (Accessed: 10 March 2015)

[7] QCustomPlot [online]. Available: http://qcustomplot.com/ (Accessed: 8 April 2014)

[8] QwtPlot3D [online]. Available: http://qwtplot3d.sourceforge.net/ (Accessed: 25 August 2014)

[9] T. Munzner, Visualization Analysis and Design. A K Peters Visualization Series. CRC Press. 2014.

[10] Aqueous Solutions LLC, The Geochemist's Workbench [online]. Available: http://www.gwb.com/ (Accessed: 15 October 2014).

[11] Thunderhead Engineering Consultants, PETRASIM 5 [online]. Available: http://www.thunderheadeng.com/petrasim/ (Accessed: 17 October 2014).

[12] C. M. Bethke, GWB Reaction Modeling Guide – A User's Guide to React and Gtplot, University of Illinois, 74p. 2008.

[13] K. Pruess et al., TOUGH2 User's Guide, Version 2.0, Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, Calif., 1999.

[14] R. Likert, A Technique for the Measurement of Attitudes. Archives of Psychology 140: 1–55. 1932.

[15] Soetaert, K.E.R.; Herman, P.M.J.; Middelburg, J.J., A model of early diagenetic processes from the shelf to abyssal depths. Geochimica et Cosmochimica Acta, Vol. 60, No. 6, p. 1019-1040, 1996.