Visual Analytics for Reservoir Analogues*

Emilio Vital Brazil[1], Vinícius Segura[1], Renato Cerqueira[1], Rogério de Paula[1], and Ulisses Mello[1]

Search and Discovery Article #70361 (2018)**
Posted October 1, 2018

*Adapted from extended abstract based on poster presentation given at AAPG 2018 Annual Convention & Exhibition, Salt Lake City, Utah, United States, May 20-23, 2018
 DOI:10.1306/70361Brazil2018

[1]IBM Research Brazil, Rio de Janeiro, Brazil (evital@br.ibm.com)

Abstract

In this work, we present a workflow that integrates the knowledge of the geoscientist with Artificial Intelligence (AI). It applies machine learning (ML) algorithms to go through extensive datasets of reservoir characterization, allowing and empowering experts to visually explore such datasets, to retrieve analogues, to estimate unknown parameters, and, thus, to make better-informed decisions. To prove the effectiveness of our workflow, we implemented a system that displays the reservoirs on a map enabling experts to visually inspect and explore data, giving them insights about it. RAVA (Reservoir Analogues Visual Analytics) is an interactive system that allows users to explore a reservoir database and to estimate unknown information of a target reservoir by automatically identifying analogues. Our workflow enables the experts to contribute with their knowledge in all critical steps and test what-if scenarios.

Introduction

Missing data is a common problem in many industries. This problem can occur by a malfunction of measurement tools or because the data were not collected or processed correctly. In the oil and gas industry, one of the areas that missing data is an issue is the reservoir studies. Acquiring geophysical data is a time-consuming and expensive activity. Moreover, depending on the exploration phase, there is little data available and it is not even possible to acquire new data (e.g., during the bidding phase for new areas to explore). Geoscientist experts must leverage the available geophysical data – albeit incomplete, low quality, and high uncertain – when making business-critical decisions, such as the “go / no go” of new acquisitions. One of the most common strategies in the industry is to use available known information from similar reservoirs (analogues) to better estimate the unknown information, drawing a more complete picture of the target area.

In the last two decades, the computer science community has developed a body of work to visualize and analyze data from different domains (Batch and Elmqvist (2018); Lam et al. (2018)). There is, however, a lack of analytics and visualization tools that facilitate the investigation of similarities and analogies specific for the oil and gas industry. Furthermore, for experts, facts that support their decisions are paramount for their daily work. Thus, systems that work with artificial intelligence (AI) should not only provide useful results, but evidences to help explain the similarity results. To approach these issues, we developed a workflow that integrates the geoscientist knowledge with AI, applying machine learning (ML) algorithms to go through extensive datasets of reservoir characterization, while also allowing and empowering experts to visually explore such datasets, retrieve analogues, estimate unknown parameters, and thus make better-informed decisions.

We implemented a system using the proposed workflow focused on the reservoir analogues (RA) problem. Perez-Valiente et al. [2014] described a method to complete missing data and then calculate a list of possible reservoir analogues. We expanded this process enabling the expert to inspect the dataset and the results using visual analytics tools. We implemented a system that displayed the reservoirs on a map allowing the experts to investigate and explore the data visually, giving them insights about it. RAVA (Reservoir Analogues Visual Analytics) is an interactive web application that allows users to explore a reservoir database and to estimate unknown information a target reservoir by automatically identifying analogues given a set of key parameters. Our workflow enables the experts to contribute with their knowledge in all critical steps and test what-if scenarios. The similarity function, which defines the set of analogues, can be determined according to the final goal. For instance, the set of similar reservoirs appropriated to gather information about EOR techniques is different from those suitable to study porosity curves. The system allows the expert to set the similarity function for each scenario. Additionally, we have created a method to examine the influence of each parameter on the analogues set. This technique enables the users to understand the reservoir's similarity characteristics more deeply.

The Workflow

Our workflow (Figure 1) starts with a set of tools to inspect a dataset visually, and it should empower the expert with visual cues that enables the creation of hypotheses about the data and specific tasks. For instance, having in mind the task of deciding the best EOR method for a given reservoir, the expert must know which parameters are essential to define analogues for this task, which ones are missing and if the dataset has enough data to estimate those. After that, the system runs the ML algorithms and gives the opportunity to the expert to visually investigate the results and validate them. It is another critical step that the knowledge of the expert is crucial in the investigation. Based on the ML output, the expert can decide if she/he should change the values or not, choose weights for the similarity algorithm, or re-start the process. After that, the similarity algorithm outputs a list of analogues, and the expert can validate that. All expert interactions and algorithm results are stored in an experiment knowledge base that can be used to improve the outcomes of other interactions.

To aid the geoscientist in finding analogues, we propose the RA strategy, summarized in Figure 2. We start with an incomplete database (i.e., containing lots of missing values) of reservoirs' parameter values. The first step is "data preparation", by choosing a target reservoir (the reservoir that we want to find analogues to), a parameter subset – the key parameters – and the key parameters' weight that will be used as criteria to find analogues. Given this configuration, we proceed with the "data estimation" step. Given the incomplete database, we use machine learning techniques to better estimate the missing values, thus expanding the number of analogues that can be found. With a more complete database (not every missing value may be estimated), we continue to the "analogues finding" step: based on the key parameters and their weights, we establish a similarity function to calculate the similarity between the target reservoir and every other reservoir in the database. This allows us to retrieve an ordered list of analogues, from the most similar one to the least similar one. Given a subset of analogues, we can follow with the "parameter estimation", generating a probability distribution function based on the chosen analogues.

The RAVA system

RAVA was created to allow the exploration of the reservoir database and the RA algorithm leveraging visual analytics techniques. It is comprised of three main pages: (i) the map view, (ii) the experiment configuration, and (iii) the experiment results view. Figure 3 summarizes each page features and the possible workflow amongst them. The map view (Figure 4) works as the main landing page. It highlights the geographical distribution of reservoir and properties, by placing each reservoir on its geographical position (latitude/longitude) and coloring it according to the current selected parameter. This allows the user to detect patterns regarding the parameter distribution behavior, providing insights to fill the missing data. For example, in Figure 4, it is easy to notice the “Longitude” parameter behavior, going from lighter colors on the left to darker colors on the right.

The left sidebar shows additional information regarding the information on the database. It changes its content based on the current selection, allowing the user to explore the database from different abstraction levels as summarized in Figure 5. When nothing else is selected, the user explores the database in the “continent” abstraction level. The “Overall” tab use box plots to display the distribution of all parameters in the database considering all available data. The “Parameters” tab shows the distribution of the current selected parameter (in Figure 5, “Area (Km2)”) grouped by continents. In the “Papers” tab, the user may select a continent to see academic literature related to the selected continent. By choosing a basin, the content is updated. The “Overall” tab adds the parameters distribution of the basin, allowing the comparison between the basin and the whole database. The “Parameters” and the “Papers” tabs keep their functionality.

If the user selects a reservoir, the sidebar is updated again. The “Overall” tab shows the selected reservoir’s parameter values compared to the distribution of the basin where it is located. The “Parameters” tab shows the parameters’ values current saved on the database in a table format. The “Papers” tab keeps its functionality, displaying literature related to the reservoir geographical position. Finally, the “Experiments” tab (only useful when a reservoir is selected) presents the list of RA experiments that were already executed using the reservoir as target. Moreover, if a single experiment is selected, the map shows the target reservoir connected to the experiment’s found analogues.

Given a selected reservoir, the user may choose to start a new experiment, navigating to the experiment configuration page (shown in Figure 6). The focus of this page is the target reservoir and the key parameters that will be used in the experiment. The user selects one of the pre-defined key parameter templates on the left sidebar and the weight of each parameter. The central part of the screen shows the target reservoir parameter values for the selected key parameters, showing the numerical ones on the radial bar chart and the categorical ones on a list on the right. Finally, the user should choose the experiment’s minimum similarity and maximum number of analogues, to limit the number of analogues found.

After an experiment is complete, the user may see its results, as shown in Figure 7. On the left, once again, there are the parameter values for the target reservoir. This time, however, missing parameters have been estimated using the RA algorithm. For example, the “Average Water Saturation (%)” (the highlighted line) has been estimated with a 17.82 value (compare with the image at the last row and second column in Figure 5, in which it is displayed the missing value “-”). Just below the table showing the parameter values, there is a map showing the geographical position of the analogues related to the target reservoir (the one with the marker). This is important, so the user may validate the findings by comparing to his/her knowledge. For example, the map is showing that the target reservoir – located in Africa's west coast – has many analogues in South America's east coast, which makes sense from a geological historical formation standpoint.

On the right side, the list of analogues is displayed using the same visualization as of the target reservoir, which occupies the central part of the screen (in Figure 7 it is being hidden by the probability distribution chart, which will be later explained). By selecting a parameter, the slice corresponding to that parameter is highlighted in every visualization and the value is displayed right below the reservoir's name. Moreover, the user may also see the probability distribution function amongst the analogues list, to aid in the estimation of the missing values. The chart is dynamically updated: if the user unselects an analogue (by unchecking the checkbox to the left of its name), it is redrawn to not consider the analogue. This way the user may have a fine control of the analogues list.

Conclusion

In this work, we presented a workflow to help geoscientists in finding analogues that mixes ML methods and visual analytics techniques. This combination enabled us to create a system that supports the expert gaining a better understanding of the reservoir dataset to help then formulate hypotheses about the data for specific tasks. As future work, we plan to research other visualization methods and their integration with ML techniques. We will also investigate the parameterization of the system taking into account the task, expert, and dataset enabling a suggestion system based on previous interactions. Finally, we plan to study the similarity function as one of the main research directions of this work. There are many possible similarity functions, and we would like to incorporate them and explore the different outcomes.

References Cited

Batch, A., and N. Elmqvist, 2018, The Interactive Visualization Gap in Initial Exploratory Data Analysis: in IEEE Transactions on Visualization and Computer Graphics, v. 24/1, p. 278-287.

Lam, H., M. Tory, and T. Munzner, 2018, Bridging from Goals to Tasks with Design Study Analysis Reports: in IEEE Transactions on Visualization and Computer Graphics, v. 24/1, p. 435-445.

Perez-Valiente, M.L., H. Martin Rodriguez, C.N. Santos, M.R. Vieira, and S.M. Embid, 2014, Identification of Reservoir Analogues in the Presence of Uncertainty: in SPE Intelligent Energy Conference & Exhibition, Utrecht, The Netherlands. Society of Petroleum Engineers, SPE-167811-MS, April 2014.

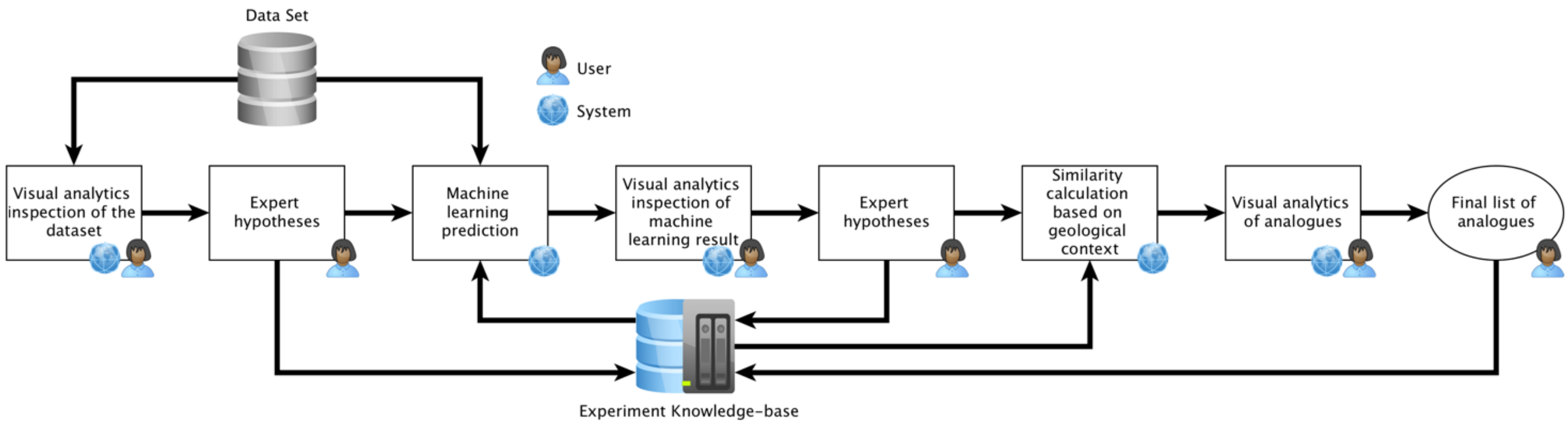

Figure 1. General workflow for analogues discovery.

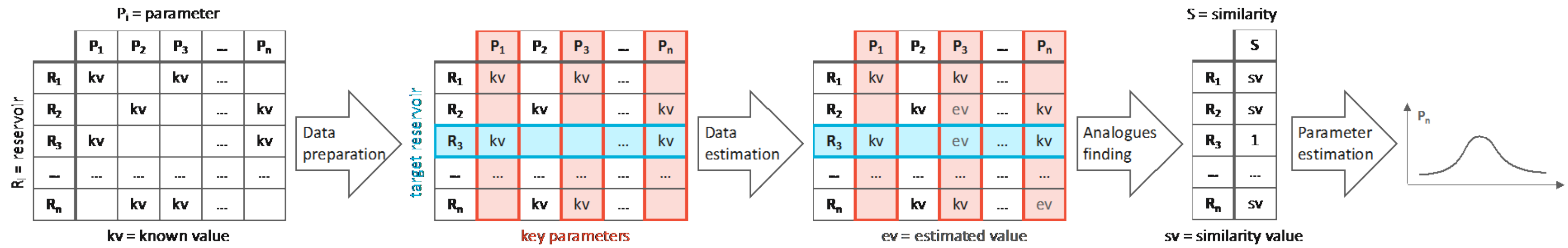

Figure 2. RA algorithm strategy.

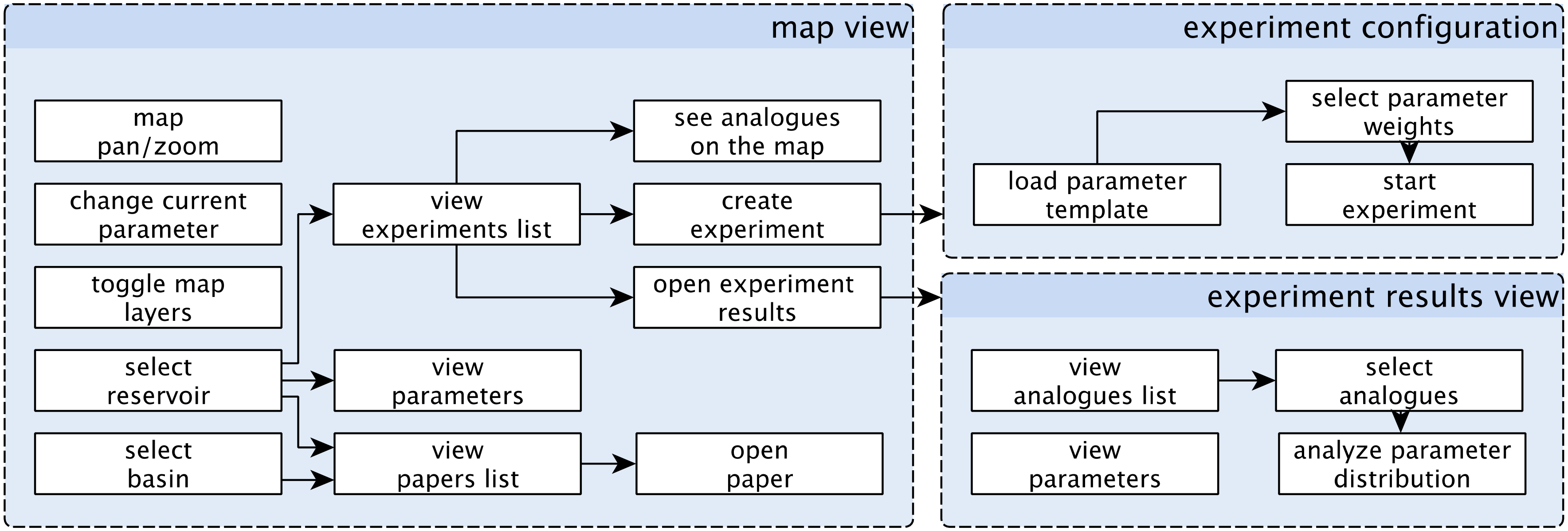

Figure 3. RAVA's main features and workflow.

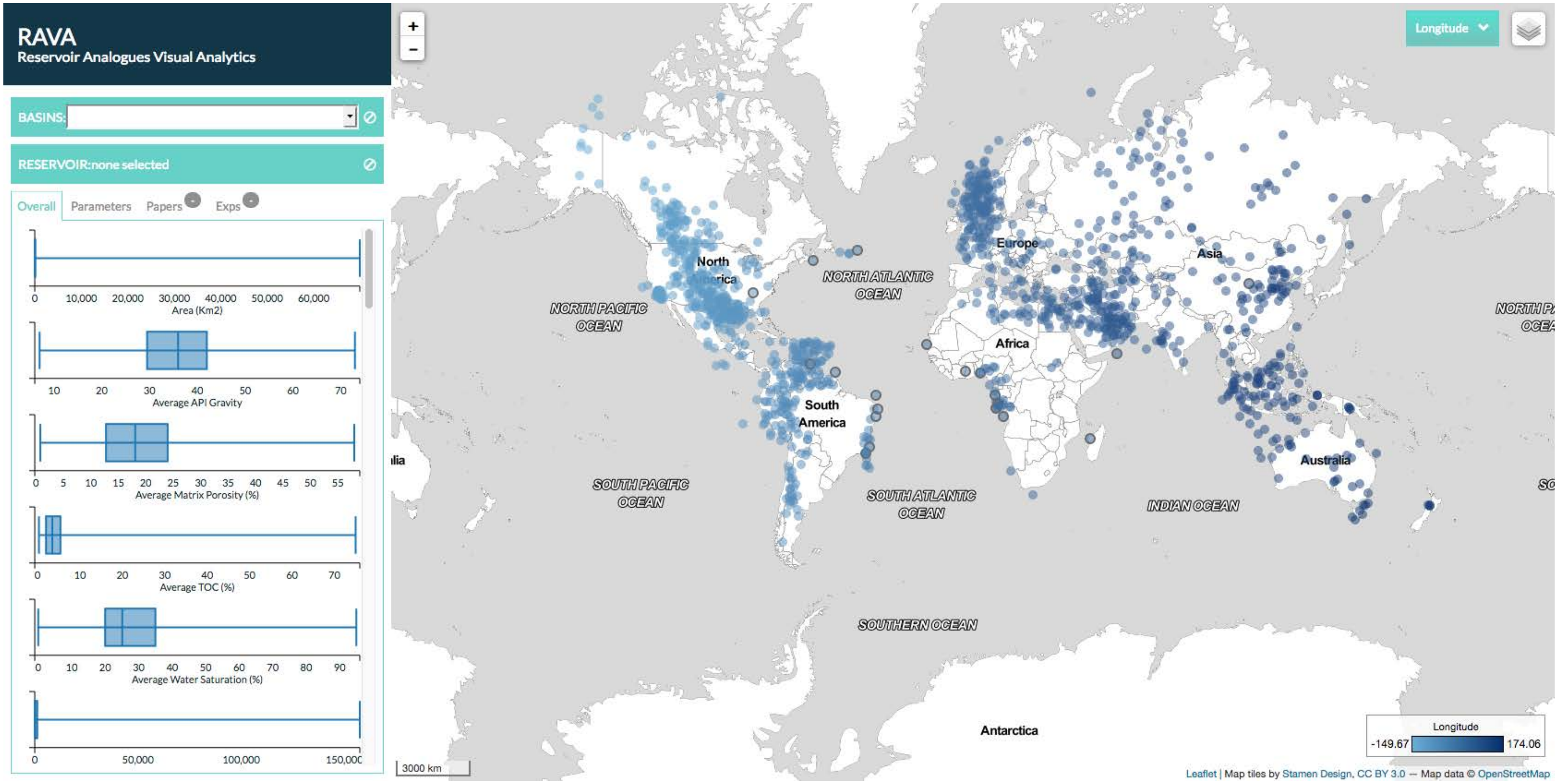

Figure 4. The map view showing the “Longitude” parameter.

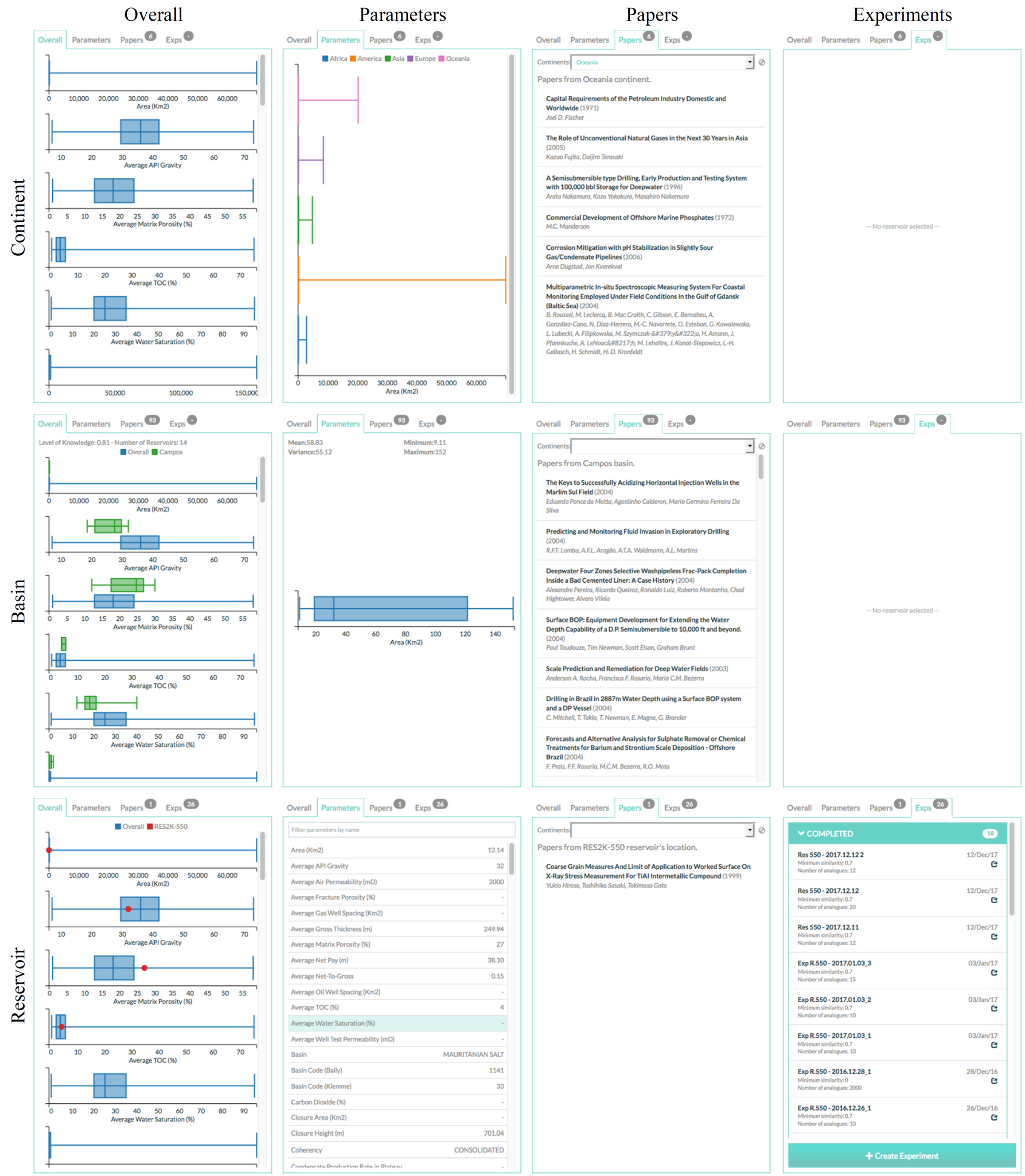

Figure 5. Sidebar information for the different selections (rows) and different tabs (columns).

Figure 6. Experiment configuration page.

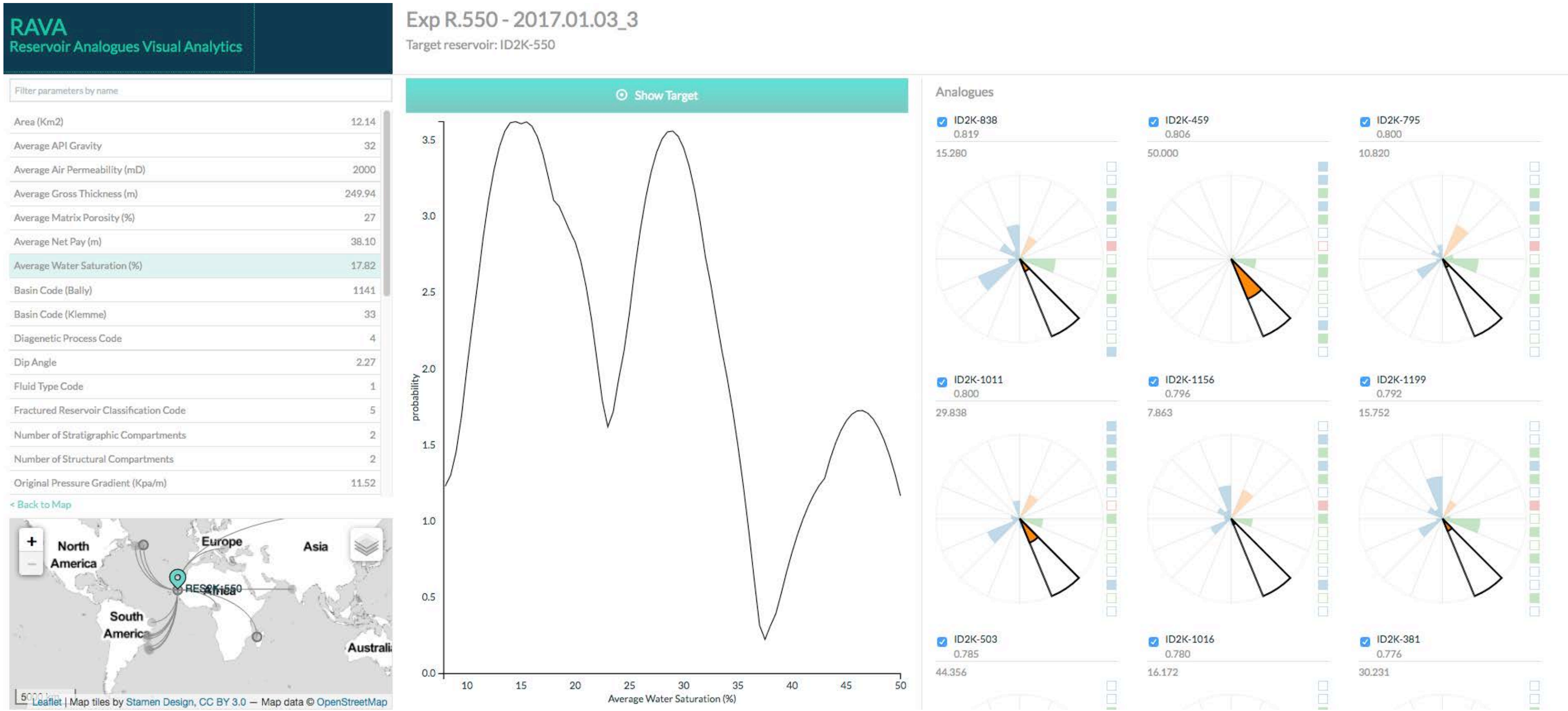

Figure 7. The experiment results page.