Abstract

Developing proxy models has a long history in our industry. Proxy models provide fast approximated solutions that substitute large numerical simulation models. They serve specific useful purposes such as assisted history matching and production/injection optimization. Most common proxy models are either reduced models or response surfaces. While the former accomplishes the run-time speed by grossly approximating the problem the latter accomplishes it by grossly approximating the solution space. Nevertheless, they are routinely developed and used in order to generate fast solutions to changes in the input space. Regardless of the type of model simplifications that is used, these conventional proxy models can only provide, at best, responses at the well locations, i.e. pressure or rate profiles at the well.

In this paper we present application of a new approach to building proxy models. This method has one major difference with the traditional proxy models. It has the capability of replicating the results of the numerical simulation models, away from the wellbores. The method is called Grid-Based Surrogate Reservoir Model (SRM) since it is has the unique capability of being able to replicate the pressure and saturation distribution throughout the reservoir at the grid block level, and at each time step, with reasonable accuracy. Grid-Based SRM performs this task at high speed, when compared with conventional numerical simulators such as those currently in use (commercial and in-house) in our industry.

To demonstrate the capabilities of Grid-Based SRM, its application to three reservoir simulation models are presented. First is a giant oil field in the Middle East with a large number of producers, second, to a CO₂ sequestration project in Australia, and finally to a numerical simulation study of potential carbon storage site in the United States. The numerical reservoir simulation models are developed using two of the most commonly used commercial simulators¹. Two of the models presented in this manuscript are consisted of hundreds of thousands of grid blocks and one includes close to a million cells. The Grid-based SRM that learns and replicates the fluid flow through these reservoirs can open new doors in reservoir modeling by providing the means for extended study of reservoir behavior with minimal computational cost. Surrogate Reservoir Modeling (SRM) is

¹ Schlumberger’s ECLIPSE and CMG’s GEM
classified as an AI-Based reservoir model (Mohaghegh, 2011) referring to a process that accomplishes the task of proxy modeling by learning the specific behavior of a numerical reservoir simulation model through training on a uniquely developed spatio-temporal dataset. The spatio-temporal dataset is developed for each model using only a handful of simulation runs.

**Introduction**

Understanding the extent of changes in pressure and saturation throughout the reservoir, especially beyond the injection and production wells, is a key component in designing many reservoir engineering related operations. Examples where having access to dynamic behavior of pressure and saturation play a significant role in planning include, design of CO₂ sequestration projects where reach of the CO₂ plume is an important design parameter, design of EOR or water flooding projects, and design of multi-cluster, multi-stage hydraulic fractures in Shale formations for understanding the extend of stimulated volume. Numerical reservoir simulation models are the only tools that are capable of providing the design engineers with such information.

Because numerical reservoir simulation models solve the flow equation at the grid block level as they simulate the fluid flow throughout the reservoir; they calculate the pressure and saturation values at each time step, for each grid block. This strength of numerical reservoir simulation models comes with a hefty price. The price to pay for this information is time. Numerical reservoir simulation models solve a system of equations using iterative techniques. The convergence to final solution usually takes several iterations. The number of iteration required for convergence usually increases as the changes in pressure and saturation become significant. This is specifically true during the transient period or when large changes in pressure and saturation may result from highly conductive media. The final outcome is a pressure and saturation distribution map (volume) throughout the reservoir (at each grid block) for each time step.

In the past several decades proxy² models have become popular in our industry. Proxy models are used to fulfill many different purposes. They are used to assist in the field development planning, uncertainty analysis, optimization of operational design, and history matching. Current state of building proxy models, when it comes to representing numerical reservoir simulation models, leaves much to be desired. When successfully developed, proxy models can ultimately reproduce results of numerical reservoir simulation models at the well locations. These results are usually limited to pressure or production profiles. When it comes to changes in pressure and saturation throughout the reservoir; it is not practical and wasteful approach, no other viable solution exist, until now.

**Grid-Based Surrogate Reservoir Model (SRM)**

Surrogate Reservoir Model SRM was first introduced in 2006 (Mohaghegh, 2006a, Mohaghegh, 2006b, Mohaghegh, 2006c). The original SRM only dealt with pressure and saturation profiles at the well, and therefore was later names Well-Based SRM (Mohaghegh, 2011 and Mohaghegh 2012). The objective of Grid-based Surrogate Reservoir Model (SRM) is fast reproduction of the numerical simulation model's results, changes in pressure and saturation, as a function of time and space at the grid block level with high accuracy.

To accomplish this objective Grid-based SRM integrates reservoir engineering and reservoir modeling with machine learning and data mining. Grid-based SRM attempts to learn the mechanics of fluid flow in the porous media from the data generated by the numerical simulation model and reproduce it for all kinds of scenarios that it may or may not have seen during the training process. This unique characteristic of Grid-based SRM makes it accurate and quick. Grid-based SRM accurately replicates the pressure and saturation distribution throughout the reservoir (at every grid block) at a very fast speed. This feature of Grid-based SRM allows fast track analysis of complex reservoir simulation models as well as design and optimization of complex development scenarios in record time.

² The word proxy means "to act on behalf of another"
Mechanics of Grid-Based SRM

Design and development of Grid-based SRM is not trivial. It requires a reasonably deep understanding of data mining, machine learning reservoir modeling and reservoir engineering. The process of developing Grid-based SRM can be summarized as using data from a numerical reservoir simulation model to teach reservoir engineering to a machine, or for now, to a computer program. This may sound a bit unfamiliar at first, but let us take a few minutes and see what this phrase means and how it can be accomplished.

We teach reservoir engineering to our future engineers using mathematics and physics. We develop equations that govern fluid flow in porous media. Using physical principles such as Darcy's law and the diffusivity equation, we explain the role of changes in pressure, permeability and fluid viscosity in the amount of fluid that can move throughout the porous medium.

Since we cannot explain these physical phenomena to a machine (at least not yet), and since the language that machines understand is data, we have to try to explain (teach) these concepts to the computer, through the use of data in a process that is called training.

Machine learning works with input-output pairs of data. The above mentioned physical principles (fluid flow through porous media) should be summarized in the input-output pairs of data that is used for training purposes. This include the design of the input-output pairs of data such that it carries the information content that needs to be transferred to the computer in form of generalizable knowledge. This is the information that can be called upon for response to new situations. For the purposes of Grid-based SRM that is being presented in this article, the data is generated from the simulation runs and very much depends on the specific reservoir (or flow process) that is being modeled. This technology can be generalized such that it can eventually lead to a completely new way of performing reservoir simulation and modeling. This is the topic of a currently ongoing project within our research group at PEARL (Petroleum Engineering and Analytics Research Laboratory at West Virginia University) and will be discussed at the proper time.

Knowledge of reservoir engineering and reservoir modeling helps us in designing the input-output pairs of data such that they carry the required information content, while knowledge of machine learning and data mining helps us in designing the data structure and in identifying the most appropriate architecture and algorithms for training data driven models. The process of developing the dataset that is needed for the training process includes summarization, abstraction, and preparation.

Grid-based SRM offers a new way of utilizing numerical reservoir simulation models where changes in pressure and saturation throughout the reservoir can be incorporated into system models, optimization routines, history matching problems, uncertainty analysis and risk mitigation studies, all at the same resolution as the numerical model. Grid-based SRM may be coupled with Well-based SRM so that it can comprehensively represent all the functionalities of a numerical reservoir simulation models.

In order to successfully train a Grid-based SRM the problem should be treated as a “path function” where changes in pressure and saturation are tracked both in time and in space. Therefore, a data record that is consisted of an input-output pair should include the grid block being studied (that we will refer to as the active grid block) along with all the neighboring grid blocks that are in contact with it, as shown in Figure 1. All the static and dynamic information for these grid blocks must be included in each data record. The static information includes the active grid block’s location and its relative distance from dynamics of the reservoir such as all the inner and outer
boundaries that are defined during the modeling process. This is shown in Figure 2. Other static information may include reservoir characteristics for the active grid block being modeled and all the connected grid blocks (See Figure 1). No transmissibilities are explicitly calculated during the development of the Grid-based SRM. All transmissibilities are implicitly incorporated through the used of geometry and reservoir characteristics of all the involved grid blocks.

![Diagram](image)

**Figure 2.** Location of the active block with respect to dynamics of the reservoir, shown in two dimensions.

Dynamic information that should be included are the constraints that are imposed on all the active wells in the field, as well as the pressure and saturation values of all involved grid blocks (active and neighboring blocks) at one time step behind.

The spatio-temporal data set that is assimilated using the above information must have all that one wishes to teach an SRM. This data set should be analyzed carefully so that it has just the right amount of redundancy as well as all the required information. It needs to be noted that given the sheer number of grid blocks that are included in a reservoir simulation models (sometimes multi-million grid blocks) can make this efforts quite a challenging one.

It is important to realize that since during the development of the Grid-based SRM we try to teach reservoir engineering to a computer and that we try to accomplish this task using data, we need examples that are not the same (redundant). Therefore, one can see that Grid-based SRM thrives on heterogeneity of the reservoir. In other words, when the reservoir is homogenous, there is not much that we can teach the Grid-based SRM and our work gets much harder. When modeling academic or toy problems, we tend to assume homogeneity in order to be able to understand the problem better and by eliminating much of the real world’s complexity we try to minimize variability during much of the study. Although such an approach has plenty of educational benefits, it makes development of a Grid-based SRM very difficult. It is strongly recommended not to test capabilities of Grid-based SRM using toy problems.
**Examples of Grid-Based SRM**

Grid-base SRM has been field tested multiple times to demonstrate its capabilities. In this article we briefly demonstrate three examples of using this technology to replicate pressure and saturation changes in numerical reservoir simulation models.

**Case 1. Matton CO₂ Sequestration Project, USA**

Mattoon field is located in the eastern three quarters of section 8 of Mattoon Township, Coles County, IL (Figure 3). The CO₂ injection well is close to the center of the Mattoon site. The field has an area of 444.2 acres. The location of the injection well will be at latitude 39.5 N and longitude 88.4 W. Primary saline formation in Mattoon is Mt. Simon. There exists a thick regional seal above the primary saline formation (primary target), the Mt. Simon. This seal is 500-700ft thick and consists of low permeability siltstones and shales of the Eau Claire Formation, underlain by Precambrian granite basement. Beyond the regional seal are two secondary seals. Pennsylvanian cyclic shales, limestones and sandstones provide almost 3000ft of protective barrier between the uppermost secondary seal and the deepest underground sources of drinking water. Currently there is no oil or gas production in the Mt. Simon.

Given that there is no wells that are penetrating the Mt. Simon depth at Mattoon site, information from a nearby well is used and mapped to help the modeling process of Mattoon site. Information available from a log in Weaber-Horn well located in Louden field close to Mattoon is used to assign the build a 53 layer static model for Mattoon with porosity and thicknesses taken from this log and permeability is calculated based on a correlation with porosity.

To model the sequestration process and determine if the reservoir has the necessary capacity to sequester the CO₂ a base model was built based on the parameters gathered using the documents from DOE’s FutureGen program. Using the information in these documents a numerical reservoir simulation model was built as shown in Figure 4.

The top of the formation is considered to be at 6950ft with the thicknesses of the simulate layers ranging from to 300 ft. The fact that porosity and permeability in each of the layers have assumed to be homogenous (basically for lack data) has made the development of the Grid-base SRM complicated. Simplification of the problems by making the unrealistic assumptions that formations are homogenous, as was mentioned before, has adverse impact on the training, calibration and validation of Grid-base SRM.

![Figure 3. Mattoon CO₂ injection site.](image)

![Figure 4. Dynamic model for the Mattoon CO₂ injection site.](image)
This project has recently been started and the results presented here are preliminary. Four layers of the reservoir (2, 4, 5, and 6) are considered to be the most porous and most permeable layers. These are chosen as the injection targets in the simulation models. Figure 5 shows the injection schedules that have been studied.

![CMG - GEM](image1.jpg) ![SRM](image2.jpg)

**Figure 5.** Comparison of pressure distribution throughout the reservoir for one of the layers during the middle of the injection period for the Mattoon CO$_2$ injection site.

Figure 5 shows the comparison of pressure distribution throughout the reservoir for one of the layers during the middle of the injection period for the Mattoon CO$_2$ injection site. The results shown in this figure can and will be enhanced as the project continues.

**Case 2. Otway CO$_2$ Sequestration Project, Australia**

This project is underway in south-western Victoria to demonstrate that carbon capture and storage (CCS) is a technically and environmentally safe way to make deep cuts into Australia’s greenhouse gas emissions. Figure 6 shows the location of the project.

The Otway Project is the Australia’s first demonstration of the deep geological storage or geosequestration of carbon dioxide. The project provides technical information on geosequestration processes, technologies and monitoring and verification regimes that will help inform public policy and industry decision-makers while also providing assurance to the community.

The objective of the Otway Project is to demonstrate that carbon capture and storage (CCS) is technically and environmentally safe and meets the expectations of government and the community. The geology of this site comprises of a series of thick layers of porous sandstones that are suitable for the storage of carbon dioxide. It is overlain by caprock of mudstone that prevents the leakage of the carbon dioxide to the layer above it.

Figure 7 shows the numerical simulation model that was developed for this project. This numerical reservoir simulation model is consisted of 100,000 grid blocks. It has been history matched with about one year of gas production. The objective is to use the history matched model and use it to design and analyze injection of large quantities of CO2. Figure 7 shows the complex structure of this site identifying the locations of the gas production and CO$_2$ injection wells. As it is demonstrated in Figure 7, the reservoir structure is distinguished by a fault. The outerboundaries of this reservoir include both no-flow boundaries and active aquifers. The active aquifer is at the southern edge of the reservoir while the other parts of the outer boundary are no-flow.
Upon completion of the numerical simulation model that included history matching the gas production and the CO$_2$ injection process, it is desired to optimize the CO$_2$ injection process in this site, identifying the extent of the CO$_2$ plume as a function of time and amount of CO$_2$ being injected. It is clear that the extent of the CO$_2$ plume will be a function of reservoir characteristics and that these characteristics are quite uncertain. Quantification of these uncertainties requires large number of simulation runs, so much that it can make the whole process impractical. Grid-base SRM can be very helpful in providing the means to perform such analyses, quite quickly.

It needs to be mentioned that most of the CO$_2$ injection projects are planned in sites where static characteristics of the reservoir are poorly understood. This is due to the fact that not many wells have been drilled in these sites to begin with. On the other hand, extend of the CO$_2$ plume is a major design objective that is very much a function of distribution of reservoir characteristics. Therefore, being able to perform uncertainty analysis (quantify the uncertainties associated with the static model) as far as the CO$_2$ plume extent is concerned, becomes a very important task in design and planning of CO$_2$ injection projects.

Given the fact that numerical reservoir simulation models are the only viable design tool for this purpose and that these models are very slow (takes hours for a single run on clusters of CPUs), role of Grid-bases SRM as a tool that is capable of providing fast track analysis becomes very important. Although this project is also yet to be completed, some preliminary results (as presented in this article) show that Grid-base SRM can play an important role in accomplishing the objectives of this project.

**Figure 6.** Location of Otway CO$_2$ sequestration project.
Figures 8 and 9 show the comparison of the results generated by the numerical reservoir simulation model with Grid-based SRM for pressure and saturation in one of the injection layers. These figures show these results for three randomly selected months. It is evident from these figures that Grid-based SRM is quite powerful in accurately replicating the results of the numerical simulation model. The speed at which the Grid-based SRM replicates the numerical simulation model’s results is basically limited by the speed of the computer being used and the speed of its graphic card since calculating and displaying the results of the Grid-based SRM is pretty much instantaneous. To put the results shown in Figures 8 and 9 in perspectives it is important to note the following:

1. The Otway model includes more than 100,000 grid blocks. Only a small fraction of these grid blocks were used to train the Grid-based SRM. A large percentage of the grid blocks were essentially blind to the Grid-based SRM.

2. The Grid-based SRM is validated in two steps. First, by the performance of the Grid-based SRM on the blind dataset that includes a large portion of the grid blocks as explained in bullet item 1. Second, by completely blind simulation runs where injection scenario is modified and/or the distribution of reservoir characteristics is modified.

3. Since the Grid-based SRM replicates the results of the numerical reservoir simulation in both space and time, it uses the results of its own predictions as input, one time step behind. This process is referred to as cascading.
Figure 8. Comparison of SRM with the numerical reservoir simulation model of the Otway CO2 injection project. Pressure distribution in one of the injection layers is shown for three months.
**Figure 9.** Comparison of SRM with the numerical reservoir simulation model of the Otway CO2 injection project. Saturation distribution in one of the injection layers is shown for three months.
**Case 3. A Giant Oilfield in the Middle East**

Figure 10 shows the numerical reservoir simulation model that represents a giant oilfield in the Middle East. The total daily oil production from this field is capped at 250,000 barrels per day. Furthermore, each well is capped at 1,500 barrels of liquid per day. Since water is being injected as part of a pressure maintenance program, and since water cut has been a problem in some wells, the production cap is imposed in order to avoid bypassing oil and creating hard to produce oil banks that are left behind. On the other hand it was suspected that several wells in the field might be capable of producing more oil without the threat of high water cut and a carefully planned rate relaxation program was desired and became the main objective of this project.

The operation in this field includes water injection into some of the layers for pressure maintenance and sweep purposes. Gas injection is also taking place in some areas of the field. The reservoir includes many major and minor faults that have been detected by geoscientists and are part of the geological model that has been used to build the full field numerical simulation model. Several rock types have been identified in this reservoir and have played an important role in developing the geological and later the dynamic model. The dynamic model has been developed using ECLIPSE™ and includes about one million grid blocks. A single run of the version of the dynamic model used for this study took about 10 hours on a cluster of twelve 3.2 GHz Intel Xenon CPUs. This reservoir simulation model includes more than 800,000 grid blocks.

Changes in water saturation were one of the results from this numerical simulation model that was of importance to the asset management. Since water is being injected for pressure maintenance and oil displacement, and given the fact that increase in water cut was being observed in several of the wells, it was important to be able to estimate the movement of the water through the reservoir. Uncertainties associated with the static model, which

---

The reservoir simulation model has been developed by the NOC. All runs were made by the asset management team at the NOC and only the results of the runs were shared with those involved in the development of the Grid-based SRM.
is an important issue that all numerical reservoir simulation models must to deal with, complicates such estimations. Therefore, it is important to quantify the uncertainties associated with water distribution throughout the reservoir. This requires access to high speed interaction with the model that would be impractical, given the slow nature of the numerical reservoir simulation model.

**Figure 11.** Comparison of SRM with the numerical reservoir simulation model of the giant oilfield in the Middle East. Saturation distribution in four 2D slices are shown at a given time step.
A Grid-based Surrogate Reservoir Model was developed to assist in this effort. The Grid-base Surrogate Reservoir Model was trained and validated using a handful of simulation runs and was able to accurately replicate changes of pressure and saturation throughout the reservoir. The trained Grid-base Surrogate Reservoir Model is an extremely fast tool (as was mentioned in the previous sections) and can be a vital tool for reservoir management.

Figure 11 shows the comparison of water saturation distribution generated by the numerical reservoir simulation model with those generated by the Grid-base SRM of a blind run for several 2D slices of the reservoir. The orientation of these layers is identified in Figure 10.

**Conclusions**

A new reservoir management and analysis tool is introduced. This new tool is called Grid-based Surrogate Reservoir Model. The Grid-based SRM is capable of fast and accurate replication of numerical reservoir simulation model’s results, such as pressure and saturation changes throughout the reservoir (between wells), at the grid block level. To develop (train, calibrate and validate) a Grid-based SRM only a few (about 10) simulation runs are required. Since Grid-based SRM are an integration of reservoir engineering and reservoir modeling with data mining and machine learning, their development require certain degree of understanding of all these disciplines. Grid-based SRM thrive on heterogeneity of the reservoir since they learn from multiple examples of fluid flow through porouse media and as such are not appropriate for academic or toy problems.

**Acknowledgement**

Authors wish to acknowledge Computer Modeling Group and Schlumberger for providing the reservoir Simulation models GEM™ and ECLIPSE™ for the first two case studies that were conducted at West Virginia University and Intelligent Solutions, Inc. for providing IDEA™ that was used to develop the Grid-base SRM. Authors also would like to acknowledge DOE-NETL’s NRAP (National Risk Analysis Partnership) program for funding one of the case studies that was presented in this paper.

**References**


