SIMULATION APPROACH SELECTION IN RESERVOIR MANAGEMENT

A Thesis

by

MAXIM ZUBAREV

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Michael J. King
Co-Chair of Committee,	Eduardo Gildin
Committee Members,	Akhil Datta-Gupta
	Richard Malak
Head of Department,	A. Daniel Hill

December 2016

Major Subject: Petroleum Engineering

Copyright 2016 Maxim Zubarev

ABSTRACT

Rapid evolution of technologies in petroleum industry in last decades has significantly improved our abilities in hydrocarbon reservoirs development. The number and complexity of tasks to be solved by reservoir engineers are gradually increasing, while the cost of field development projects is rising. In this conditions, optimal decision-making in reservoir management becomes critical since it might result in either significant benefit or financial loss to a production company. Although a significant improvement was made in project risk management to control project costs in the case of unfavorable outcome, reservoir evaluation still plays the important role and affect entire reservoir management and production process. Since the work of petroleum engineers actively involves reservoir simulation and target search for optimal solution of the particular reservoir assessment problems, selection of the most appropriate simulation approach in a timely manner is important. Successful search for suitable solution to a particular reservoir engineering problem is always a non-trivial task since it involves analysis and processing of large amounts of data and requires professional expertise in the subject area.

In this work we proposed an expert system, what provide flexible framework for the proper simulation approach selection and involves thorough data analysis, multiple constraints handling, expert knowledge utilization, and intelligent output requirements implementation. This expert system utilizes linguistic method of the pattern recognition theory for knowledge base design and inference engine implementation, what significantly simplifies procedures of the system design and provides it with tuning flexibility. This thesis elaborates on major aspects of the expert system design in close relation to data processing and recommended solution finding methods.

To validate the expert system's applicability, several tests were designed based on the synthetic Brugge field case and real petroleum reservoir data. These tests demonstrate functionality of the major expert system elements and advantages of selected implementation methods. Based on obtained results we can conclude successful development of the expert system for appropriate simulation approach selection.

DEDICATION

I dedicate this work to my beloved family. To my daughter Maria, my wife Varvara, my parents Igor and Nadezhda, and my brother Denis.

ACKNOWLEDGEMENTS

I am very grateful to my advisors, Dr. Michael J. King, Dr. Eduardo Gildin, Dr. Akhil Datta-Gupta, and Dr. Richard J. Malak for their wise guidance, tremendous support, and motivation throughout these years.

I would like to thanks Foundation CMG for financial support of this work and provided opportunity for me to be a part of innovative community.

I am also thankful to my friends, colleagues, and department faculty and staff for a nice time and great experience at Texas A&M University.

Finally, I am always appreciative to my family.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE		Page

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

Rapid evolution of technologies in petroleum industry during the last decades expanded capabilities in oil and gas reservoirs development. Along with the growing advances in exploration and production techniques, deployed in conventional and unconventional reservoirs, an increase in field development projects cost is also observed. Because of risk associated with project failure, it may therewith enlarge size of financial losses. Depending on the size of the losses for a certain company, actions that prevent similar failures in the future should be introduced. This section summarizes the importance of data evaluation approach and lay down the foundation to the development of an expert system in the thesis.

1.1 Importance of Data, Models, and Simulation Approach Selection in Reservoir Management

According to McVay and Dossary (2014), companies in petroleum industry continuously underperform compared to the project expectations. Authors suspect that while high oil prices of the last decade have overall improved industry performance, they also caused industry relaxation and worsening the quality of decision making. In general, this point of view does not exactly imply an irresponsible business management since entire decision-making process is very complex, situational, and involves multi-stage information handling with a high number of issues. However, it still requires improvements in the overall process workflows. For instance, in the third quarter of 2015 the Shell company reported “loss of $ 6.1 billion – net $ 8.2 billion of upstream write-downs and other charges primarily linked to its unsuccessful Arctic drilling off Alaska and Canadian Oil Sands project” (Smedley, 2015). As a major disappointing moment, drilling of a dry hole was mentioned in the report. This example shows that even a petroleum industry leader, fully equipped up with up-to-date technology, is not insured from failure. Hence, the use of the most advanced technologies in petroleum exploration and/or production cannot guarantee success until all the major parts of uncertainty are removed from data gathering and information processing, or full reservoir study is improved starting from the very basic level to a more complex.

It is important to point out that the main source of incorrect decisions is always related to the lack of required information or poor data assessment. Because entire reservoir management process directly involves field study as a starting point, reservoir evaluation plays a significant role affecting the output results. Taking in consideration a “cause and effect” concept, the visualization of poor outcomes in reservoir management is shown in **Figure 1.1**:

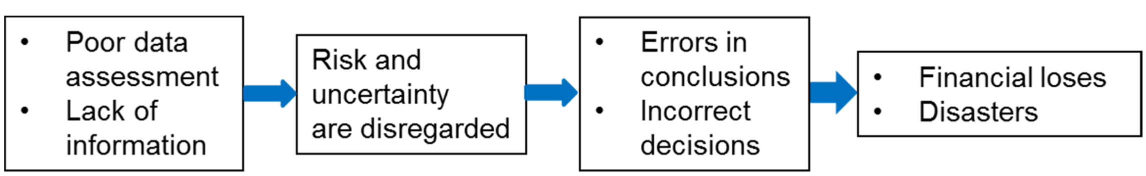

Figure 1.1 – Impact of unsatisfactory reservoir data on reservoir management outcomes

Figure 1.1 represents a basis for incorrect decision-making process. Initially misleading or insufficient data cause a situation when uncertainty either is evaluated incorrectly or remains unknown and not taken into account. Therefore, it becomes difficult or even impossible to assess data uncertainty ranges that further lead to improper decision-making risk evaluation. Consequently, errors in conclusions and incorrect decisions are inevitable, which can lead to financial loses or disasters. Hence, quantity and quality of reservoir data play a very important role in reservoir management, as they are key criteria that define what we exactly know and understand about the subsurface object.

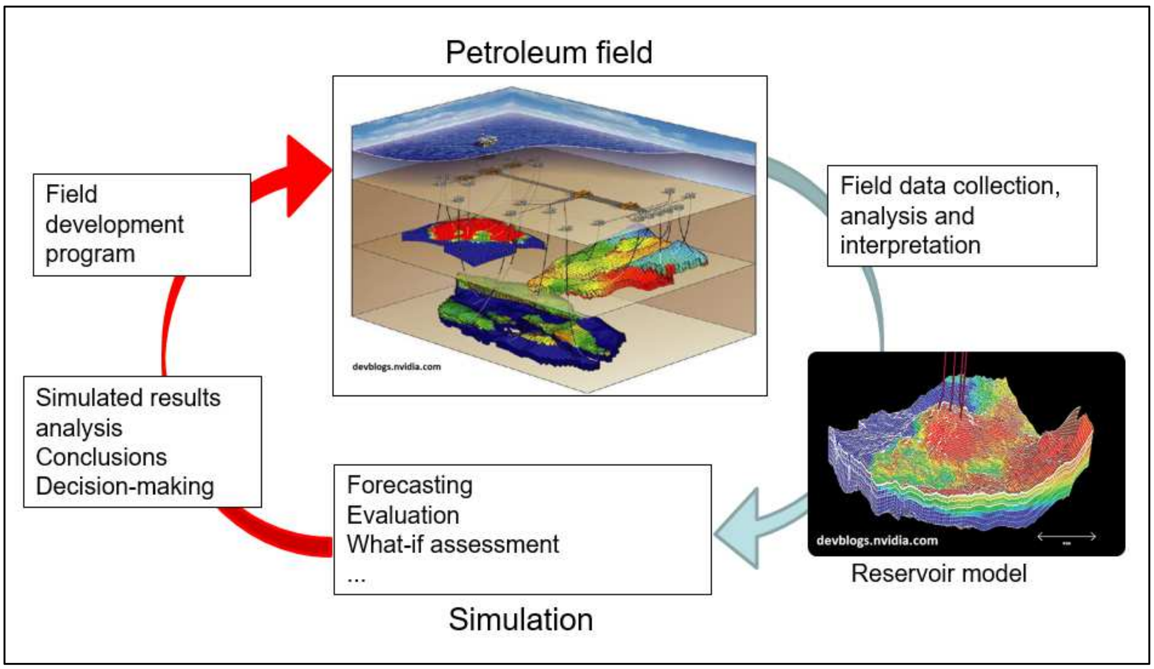

Figure 1.2 – The place of reservoir model design and simulation in field management process

Figure 1.2 shows data flow process during the life cycle of a petroleum field management. Collected field or reservoir data is used for the particular reservoir model design and rectification within all stages of a field life. This model is employed in simulation runs for variety of purposes that include reservoir performance studies under different development strategies called “what-if” scenarios. Nowadays, the use of reservoir simulation allows assessment of multiple “what-if” realizations and selection of the most optimal ones (Satter, Iqbal, and Buchwalter, 2008). Depending on the particular goal to be reached, the simulated outputs are further analyzed and implemented for conclusions and decision-making, which are then executed as the field management program.

In addition, the need of reservoir models in petroleum industry is dictated by long duration of the processes, control actions and object responses, that we need monitor and optimize. Therefore, we cannot experiment with reservoirs since by the time we see some response, it might be too late to take any actions to correct it. Hence, we need predictive models.

The study of a petroleum reservoir is truly a multidisciplinary effort. Specialists in Geology, Geostatistics, Geophysics, Geochemistry, Petrophysics, and Engineering contribute their joint work in an integrated reservoir model designs for a real reservoir performance understanding and forecasting (Satter, Iqbal, and Buchwalter, 2008). Results of various applied studies give us an information about subsurface objects, their history, properties, and features setting a basis of the model. Simulation model design, validation, and optimization processes is a separate wide topic that is not included into the scope of our work. Detailed description of the modelling methodology can be found in literature (Falkenhainer and Forbus, 1991; Levy, Iwasaki, and Fikes, 1997; Malak and Paeridis, 2007; Oberkampf and Roy, 2010).

Correct choice of an appropriate simulation approach to be used for the specific reservoir evaluation problem is a daunting task to be performed by petroleum engineers. Proper selection of the simulation method is critical since it determines the accuracy and applicability of simulated results and, consequently affects decision-making process. Depending on the exact goal, reservoir engineers select simulation approach based on the certain data availability, its quality, constraints existence, and methodology (Satter, Iqbal, and Buchwalter, 2008). In practice, determination of the most suitable simulation method at certain conditions is not easy since it is characterized by the following features:

- Different simulation approaches may give discrepant results at the same given conditions;
- Some of them can or cannot be used for a special task solving under number of constraints, data quality and insufficiency, and methodology in the basis;
- The most appropriate approach selection involves analysis and processing of large amount of data and requires professional expertise in the subject areas.

In addition, the sought-for proper simulation approach should provide:

- Sufficient accuracy, adequate complexity, and representation of available data;
- Robust and appropriate basis for realization of reservoir analysis objectives under existing constraints, among others.

Making a choice in such conditions is non-trivial while it also implies the use of sufficient amount of theoretical knowledge and practical experience that can be very limited.

1.2 Objectives and Scope of Work

Summing up, the quality of decision-making process in reservoir management is strongly dependent on a realistic understanding of the certain subsurface object, its parameters and features, and quality and sufficiency of data we build the model from. The choice of the most appropriate simulation method is very important, since it primarily defines the data requirements and accuracy and applicability of simulated output with respect to the particular reservoir engineering task. Correct selection of the proper simulation method implies an existence of extensive theoretical knowledge and practical experience. Their lack may result in reduction of reservoir evaluation quality.

The most suitable reservoir simulation approach selection under specific circumstances is a problem that has not yet been posed as a formal task in reservoir evaluation, and requires to be developed.[1] To contribute to a proper reservoir management improvement, we have posed the following goals and scope of our work:

- The primary objective is to formalize, design and test the reliable methodology to provide decision-making support in simulation approach selection;

[1] At least to the best of my knowledge, I am not aware of such proposed work developed elsewhere.

- This methodology should supply flexible framework, involve thorough data analysis, multiple constraints and limitations handling, expert knowledge utilization, and intelligent output requirements implementation;
- Use the linguistic method of the Pattern Recognition Theory to set a basis for the methodology realization procedures, such as data encoding, symbolic reasoning, search and recommending the most suitable simulation method, and suggesting on what should be additionally done for a specific engineering task solving;
- Implement the developed methodology in the knowledge-based expert system design, with specific structure and functionality, as a means for problem solving that requires expertise.

Realization of these objectives could significantly improve the simulation approach selection process, increase quality of data analysis and reduce the risk of errors to emerge.

Furthermore, the developed software can be also used for the guiding or coaching purposes. An ability of the expert system to generate explanations on output results could be useful for those users who experience a deficiency of qualification in the particular area of interest. Namely, users can get the information on:

- What reservoir simulation method is recommended to be used for particular problem solving as the most appropriate under number of constraints and limitations;

- What additional reservoir data should be obtained in order to execute other methods;
- What could be a predicted accuracy of calculated results;
- What should be the workflow when engineering problem solving require implementation of multiple simulation methods.

1.3 Thesis Organization

In the first section of this thesis we discussed motivation and objectives of our work and described proposed solution. I order to accomplish the main objectives proposed here, the following structure was organized.

In the second section, we introduce extensive literature review that enlightens application of five major simulation approaches with respect to particular reservoir evaluation tasks. The set of listed assumption, constraints, and limitations in the foundation of each method is used to create the basis of our methodology.

The third section provides overview of expert systems as a means of complex problem solving that requires professional expertise. The main concept and structure of the systems, their key features and functionality, history of development, and application in Petroleum Industry are described in this part of the thesis.

In the fourth section, we present the detailed framework depiction of the methodology and expert system design for the proper simulation method selection. Here, we explain the use of the linguistic method of the Pattern Recognition Theory as a means that determines data encoding algorithm for the alphabet of key parameters design, generation of scenarios as representation an expert knowledge, and solution search workflow. Additionally, we delineate the structure of the developed expert system, construction, and functionality of its components, such as Data Pre-processing, Scenario Generation, Knowledge Base, Inference Engine, and Decision Support modules. As a very important topic, the input data quality control and detection and dealing with constraints, which affect applicability of simulation method and restrict an accuracy of simulated results, are described as a part of Data Pre-processing module. Besides, we also present the well placement justification technique as an extension of the expert system functionality.

The fifth section represents the expert system validation and field application workflow using Brugge synthetic simulation model and offshore petroleum reservoir data. Obtained results are discussed in this section.

Finally, in the six section of the thesis we summarize obtained results and conclusions, discuss observations, and provide our vision of future work on the expert system improvement.

2. RESERVOIR SIMULATION METHODS

This section presents an overview of the major simulation approaches using in reservoir evaluation and areas of their application. We introduce the list of specific engineering tasks that can be solved using five major reservoir evaluation methods. Also, we mention assumptions, constraints, and limitations in the basis of each method which determine its applicability and accuracy of calculated results.

2.1 Discussion on Model Application in Reservoir Engineering

Reservoir simulation is based on the methodology put into a certain approach, where specific mathematical relationships describe ongoing physical processes in the reservoir (Odeh, 1969). These descriptions, i.e. models, define the certain set of engineering tasks that can be solved and parameters to be used. By its nature, any methodology is developed considering specific assumptions, stipulations, simplifications, and solving methods, which further establishes opportunities, requirements, and constraints for the use of the particular method (Satter, Iqbal, and Buchwalter, 2008).

Models can be based on the understanding of underlying physical processes that occur in the field, data collected from fields under development, or fields that show certain degree of similarity. To express the degree of similarity, term “analog” is often used. Analog means an object with properties so similar to the properties of the object under investigation, that a sufficient level of confidence exists in similarity of the reaction produced by these two objects if the similar actions are applied. In other words, for analogs we can extrapolate knowledge gained through observing one object to predict reaction of another object with certainty. In practice, establishing analogy between two reservoirs is a difficult task because of their originality. However, every reservoir is an analog for itself and this is actively used in petroleum industry since the early days and serves a basis for methodology of decline curve analysis. If extensive database of analogs exists, then correlations might be a viable approach to follow since they are simple, robust, and sufficiently accurate.

Oil and gas reservoirs are complex objects with multiple compartments, different drive mechanisms, and spatially variable rock and fluid properties. Some "simple" reservoirs might be represented as a single geobody with a simple geometry with relatively homogeneous rock properties and filled with a single fluid. "Complex" reservoirs can be comprised of multiple partially interconnected geobodies with a complex geometry, highly spatially varying rock properties, and containing fluids with highly varying properties as well. However, simplicity of reservoir is just a one axis that describes complexity of the case we are dealing with. Development scenario is another axis that controls the complexity of the case for our understanding. Even a simple reservoir with a complex development scenario might result in a case of a higher complexity than that of the complex reservoir with simple development plan. Therefore, complexity of the case depends not only on the nature of the object, but also on the kind of production process and scenario that we try to apply to it.

Overall, physics of the processes in conventional reservoirs behavior is well understood. Ability to understand the underlying processes allows creation of mathematical models that can be used to describe behavior of the object under certain conditions and control actions. Mathematical models that are solved analytically can describe a simple reservoir. Complex reservoirs require solution of more complex systems of equations that are solved numerically. However, overall complexity of the model that has to be solved depends on the application. Different model applications require different levels of model complexity. For example, a simple reservoir with complex pattern water-flood might require application of a numeric model. At the same time, complex reservoir that is developed by isolated producers under primary production might be sufficiently described by analytic models. Therefore, complexity of selected model depends not only on complexity of the object, but on analysis objectives (reservoir development scenario) it will be used for.

The nature of the objects that we study in reservoir engineering allows collection of a very limited amount of data. Scarcity of the data, multitude of scales and ambiguity of interpretation creates difficulties in proper characterization of the object. Knowledge of physics and numeric tools allow us to solve problems at a very fine scale. However, resolution and amount of data available limits the scale of object representation in the model. At the same time, scale at which we need model response might be much coarser that the one we can characterize the object at. Hence, while selecting the model for proper representation of the object we have to take into account scale expectations in addition to model complexity, analysis objectives and availability of analogous data.

Model scale is a sensitive topic in reservoir engineering. Intuitively, subsurface teams try to obtain model at the finest scale possible. However, application of fine scale models even at current level of computational hardware development is extremely time consuming. Certain models can run days and even weeks, which pushes reservoir engineering studies outside of the reasonable time frame. In practice, preference is given to the models that can run faster while still providing robust and accurate representation of the object. This, in turn, brings focus to selection of proper scale, proper simplifying assumptions and finally proper model representation. Proper model scale coarsening can be achieved with parameter upscaling techniques as long as proper model accuracy is supported.

To summarize, with respect to the particular simulation task, we are looking for the most appropriate simulation approach, where the simplest model (reduced) that provides sufficient accuracy (accurate), adequate complexity and representation of the available data. At the same time, it should provide robust and appropriate basis for realization of analysis objectives. In other words, making a choice of the proper simulator we should confidently understand the reservoir, thoroughly assess available data, clearly define simulation goal, and distinctly select appropriate methodology of problem solving (Satter, Iqbal, and Buchwalter, 2008).

Depending on the data required for reservoir evaluation and problem solving, all the simulation approaches that exist today can be combined in five major groups as it shown in **Figure 2.1**: Correlations, Proxy model based, Material Balance based, Streamlines, and Finite Difference (volume) numerical simulation.

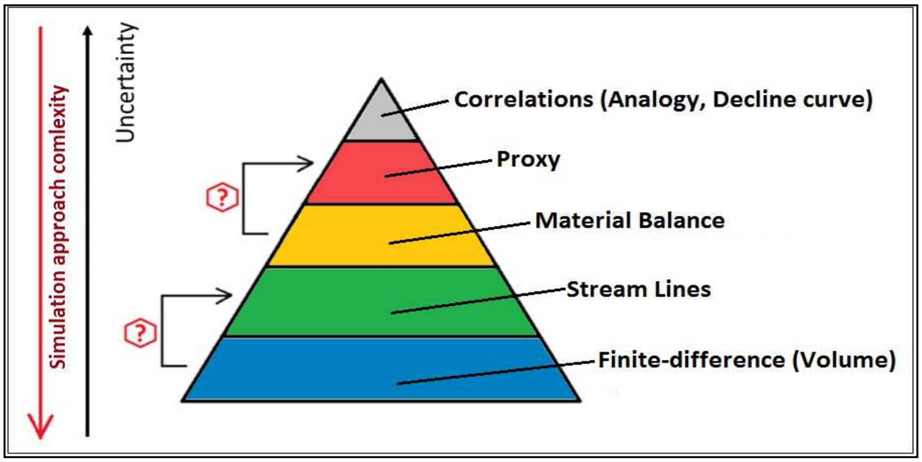

Figure 2.1 – Modelling methods (simulation approaches)

2.2 Correlations (Decline Curve Analysis)

In petroleum engineering correlation is a mathematical relationship that defines correspondence between an output variable and a set of input variables. Correlations are empiric relationships that depend on the availability of data. At the same time, they are based on analogy. Traditionally, they were actively implemented in the areas where physics of the process is not well defined or adequate mathematical models do not exist to represent physical processes. Namely, they are predominantly used when large uncertainty or lack of knowledge about the reservoir data do not allow application of more complex methods, such as numerical simulations.

Correlations were actively used in estimation of fluid properties (PVT), prediction of multiphase flow in the pipes, etc. Fundamentally, correlations are interpolation functions that are built on an extensive experimental dataset that covers possible combinations of input parameters and corresponding values of the output parameters. If all parameters significant for the estimation are taken into account and a proper interpolation function is developed, this function can be used to predict output values (i.e. PVT properties, flow regimes, etc.) for any combinations of input variables. However, one needs to make sure that parameters are selected within the interpolation region. This approach relies on an assumption that if multidimensional surface goes through some experimental points, it will give sufficiently accurate prediction for all points in between those experimental points. Therefore, use of these functions in extrapolation mode might not be appropriate. However, extrapolation might be appropriate if developed correlation relies not on the mathematical function that better fits the data, but has some resemblance of the physics as well. For example, decline curve analysis is based on observation made by Arps that production rate decline at the well can be described by a simple equation. So, Arps' equation adjusted to fit available data can successfully be used to extrapolate declining production rate into the future.

Decline curve analysis will be discussed further as the most typical correlation technique widely used in reservoir performance evaluation.

Historically, Aprs' observation that fluid production rate declines versus time exponentially (Arps, 1945) stimulated emergence and further development of decline curve analysis. Many researchers (Ershaghi and Omorigie, 1978; Blasingame, McCray, and Lee, 1991; Fetkovich's, 1996; Agarwal, Gardner, Kleinsteiber, et al., 1999, and many others) contributed in improvement and extension of this method applicability. Nowadays, decline curve analysis techniques are widespread because they are relatively simple and require a lesser data set than other methods. For instance, decline curve analysis is employed in reserves assessment for approximately "95% of the thousands of reservoirs in the United States" (Satter, Iqbal, and Buchwalter, 2008) since these reservoirs are not large in size, with studied recovery drive mechanisms, and do not require complex and expensive numerical simulation.

Fundamentally, decline curve method is based on the analysis of individual wells or field production rates, when sufficient data is available and production decline is established (**Figure 2.2**).

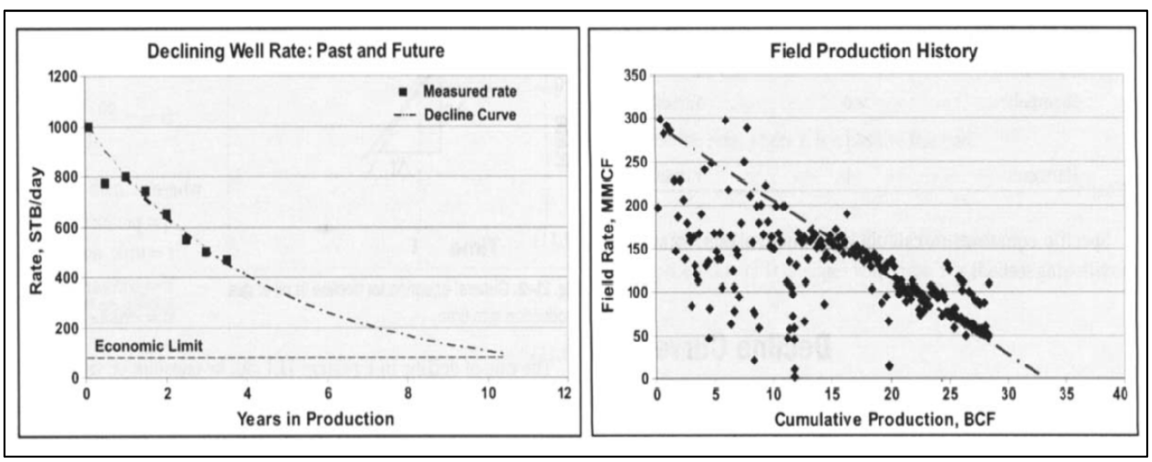

Figure 2.2 – Two views of decline (reprinted from Satter, Iqbal, and Buchwalter, 2008)

In **Figure 2.2**, the left plot represents measured oil production rate extrapolated until some economically reasonable value (abandonment limit). The extrapolation decline curve here, obtained by fitting measured early data, shows prediction of future well performance. Analogically, the right plot characterizes measured and predicted performance; in this case for entire gas field. The most important parameters that should be established by analysis and further used in extrapolation are decline rate and its exponent:

$$D = -\frac{dq/dt}{q} = Kq^{n} \ , \tag{2.1}$$

where: D – decline rate; q – production rate; t – time; K – constant; n – exponent.

Depending on combination of n and D and their characteristics, there are three main decline types can be identified in classical analysis (Satter, Iqbal, and Buchwalter, 2008):

1. Exponential – decline rate D is constant and exponent n = 0;
2. Hyperbolic – D varies with time and n = [0 < n < 1];
3. Harmonic – D varies with time and n = 0.

Once the decline type has been identified and average fluid and reservoir properties, such as reservoir thickness, rock and fluid compressibility, reservoir initial and bottomhole pressure, porosity, oil or gas saturation and formation volume factor were obtained, the following parameters can be calculated using specific equations related to the decline type and evaluation method:

- Predicted production rate;
- Ultimate recovery by summation of measured and expected (predicted) production;
- Remaining time of a well or reservoir production;
- Initial value of oil / gas in place;
- Recovery factor;
- Drainage area;
- Reservoir parameters, such as average permeability and skin factor.

However, the use of decline curve analysis must be taken with care in reservoir performance evaluation and prediction (Sun, 2015). This requirement emanates from assumptions and limitations put into methodology: each well produces from constant area, entire reservoir has no leaky boundary even though adjacent aquifer exists, depletion is the only drive mechanism, production data is sufficient for analysis, decline is established, and field operations will not consider future changes. Violation of any of these items immediately disturbs impracticability of the decline curve analysis. Additionally, the following factors influence production rates and decline curve performance, and should be taken into account:

- Early-time field life stage (exploration, appraisal, startup) do not allow to use this method since production data either do not exist or not sufficient for analysis;
- Early beginning of decline, when its trend is not confidently obvious, results in very low accuracy of calculations;

- Restricted production, bottomhole pressure changes, modification of production methods along with well treatments, workovers, implementation of enhanced recovery programs, water influx from aquifer and breakthrough are considered as intervention into a stabilized production regime. That interruption distorts decline trend and can make it impossible for using.

Overall, decline curve analysis is quite simple and efficient tool for the reservoir performance evaluation. At the condition, sufficient amount of analytical data is available and not any of the above listed limitations cause restrictions, this method gives fair results in solving particular tasks.

2.3 Proxy Models

Proxy models are fundamentally interpolation functions, but more sophisticated than the ones traditionally used. Basically, they are used as simplified models that are not based on physics, but closely resemble numeric models (i.e. can mimic their output for the same set of input parameters), "as a computationally cheap alternative to full numerical simulation" (Zubarev, 2009). Besides this, proxy models are also constructed by size reducing of an initial full physics model (Yang, Davidson, Fenter, et al., 2009). So, they can act as a proxy to a certain model, but cannot fully replace it.

As for interpolation method that is not restricted by the physics of the process, robustness of proxy models in extrapolation mode is not easy to justify. Therefore, to replace simulation models, engineers define a set of input parameters of interest (parameterize the problem) and create an, "experimental dataset," by running numeric simulation with selected parameter within defined ranges. Basic workflow of the proxy model design, shown in **Figure 2.3**, includes sensitivity analysis as mandatory and the most important step. Basically, this process is an evaluation of the effect of the input variables changes to the simulation model output. As a result of analysis, the input variables can be separated into the following groups:

- Variables that sufficiently affect simulation model response and should be used in dataset sampling;
- Insignificant variables that can be eliminated to reduce the model size.

Sampled datasets are used to create proxy models that can not only closely resemble outputs of the dataset, but can predict output values for different realization of input parameters. Depending on simulation model response, the proxy model is estimated separately with its quality validation. Namely, this model should reproduce the same results as a real model with required accuracy. Polynomial, artificial neural networks, genetic algorithms, kriging-based, and radial basis function based proxy models are the most widely used in the petroleum industry (Lophaven, S.N., Nielsen, H.B. and Sondergaard, J., 2002; Jurecka, 2007; Artun, Ertekin, Watson, et al., 2009).

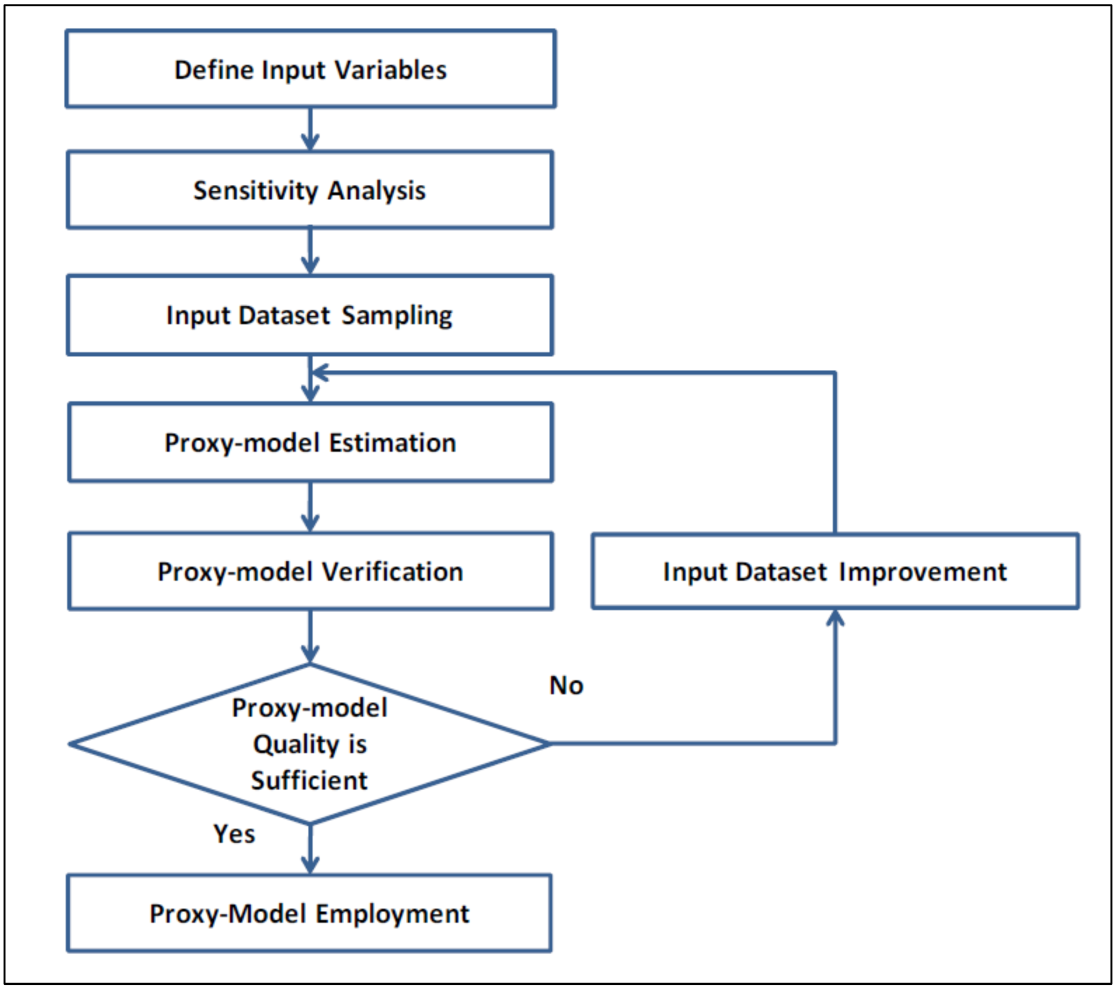

Figure 2.3 – Proxy-modelling workflow (reprinted from Zubarev, 2009)

Proxy models found a wide implementation in the petroleum industry. They are used not only as a substitution to numeric simulation models, but in virtual metering, well transient pressure data analysis, well test predictions, substitution for multiphase flow correlations, hydrocarbons initially in place calculation and much more. In general, typical application areas in reservoir simulation include (Zubarev, 2009):

- Sensitivity analysis of uncertainty variables (Yeten, Castellini, Guyaguler, et al., 2005; Junker, Dose, Plas, and Little, 2006; Slotte, and Smorgrav, 2008; Christie and Bazargan, 2012);
- Probabilistic forecasting and risk analysis (Kabir, Chawathe, Jenkins, et al., 2002; Osterloh, 2008);
- History matching (Cullick, Johnson, and Shi, 2006; Slotte, and Smorgrav, 2008; Christie and Bazargan, 2012);
- Reservoir connectivity evaluation, development modelling, screening, and production optimization (Pan and Horne, 1998; Guyaguler, Horne, Rogers, et al., 2000; Onwunalu, Litvak, Durlofsky, and Aziz, 2008; Artun, Ertekin, Watson, et al., 2009; Yang, Davidson, Fenter, et al., 2009; Pfeiffer, Reza, Schechter, McCain, and Mullins, 2011; Christie and Bazargan, 2012).

Proxy models provide certain advantages when used with problems of moderate non-linearity. However, with highly non-linear problems accuracy and robustness of proxy-models are actually questionable. The problem is not in their interpolating properties. All of the mentioned above types of proxy models are actually exact interpolators. Namely, in multidimensional parameter space they represent a surface that goes exactly through the experimental points. The problem with highly non-linear problems is in proper selection of these points. Obviously, if one runs an infinite number of simulations, a very accurate proxy model can be created. But the need for large number of simulation runs diminishes the effect of proxy models application for time saving.

Therefore, one has to come up with a small set of runs that allows creating a proxy-model properly representing model non-linearity and hence providing sufficient accuracy.

To date, different methods of Design of Experiments (DoE) are used to propose this set of models. They are generic “space filling” designs that randomly scatter sampling points over the parameters space insuring uniform coverage. This creates certain challenges.

First of all, proper representation of highly non-linear problems with small set of experiments requires a prior knowledge of the response surface complexity and effective sampling technique. Otherwise, a large set of experiments would be needed. Second, random nature of sampling combined with limited experimental sample makes accuracy and robustness of the approach questionable. Absence of “intelligent sampling” methodology makes creation of reliable model that can provide appropriate prediction accuracy across the parameter space a non-trivial task. At the same time, precision of the approach is not guaranteed because of random nature of sampling.

Overall, all proxy models are strongly dependent on real model complexity, sufficiency and quality of input data, and clear understanding of their constraints. Individually well-built proxy model can be a very convenient substitution of full numerical simulation model since it is capable of replicating the same output being exceedingly cheaper in computational time.

2.4 Material Balance Models

As it goes from the name, material balance model is based on the mass conservation law. The initial material balance equation, as a volume balance between cumulatively produced fluid and its expansion in a reservoir due to pressure drop, was presented by Schilthuis (Tracy, 1955; Dake, 1978). Many researchers, including Havlena and Odeh (1963, 1964), Tehrani (1972), Campbells (1978) and others sufficiently extended material balance analysis techniques and areas of application. This type of models focuses on volumetric characteristics and mass exchange between the reservoir and outer world (**Figure 2.4**). At the same time, it does not provide any insight into spatial saturation change and fluid movements within the reservoir due to single tank assumption. It is probably the most simplistic type of model that actually accounts for physical processes occurring within the reservoir during production.

In general, material balance equations for reservoir performance are expressed as follows (Satter, Iqbal, and Buchwalter, 2008):

- Oil reservoir:

$$F = N(E_o + mE_g + E_{fw}) + W_e \tag{2.2}$$

- Gas reservoir:

$$F = G(E_g + E_{fw}) + W_e \tag{2.3}$$

where: F – underground fluid withdrawal; N – original volume of oil in place; G – original volume of gas in place; E_o – expansion of oil and originally dissolved gas; E_g – expansion of gas cap gas; E_{fw} – expansion of connate water and reduction of pore volume; W_e – cumulative natural water influx; m – initial gas cap volume fraction.

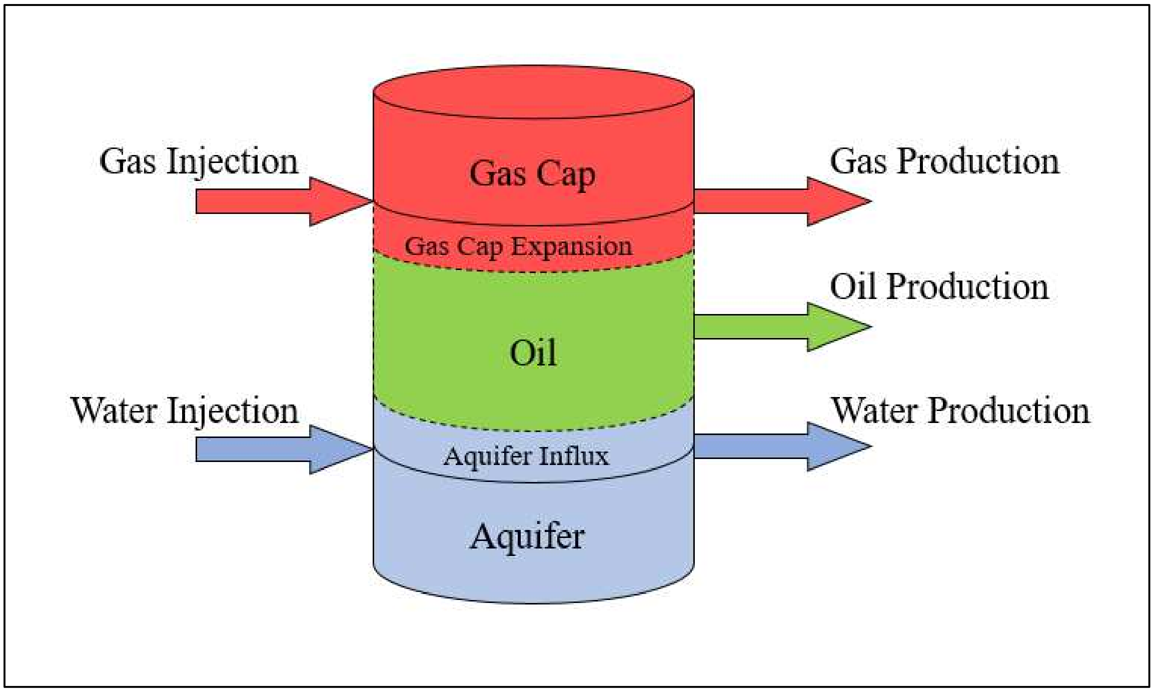

Figure 2.4 – Material balance tank model assumption (reworked from Dake, 1978)

As shown in **Figure 2.4**, material balance model represents reservoir as a tank of a certain volume filled with fluids and is a subject to fluid movement into and out of the tank due to the presence of sinks and sources. Fluid movement changes energy balance and impacts phase changes and PVT properties of the fluids which are also accounted for in the model.

Methodologically, the conceptual material balance model is based on the following assumptions:

- The reservoir is considered as a tank, homogeneous, with averaged rock and fluid properties (porosity, compressibility, permeabilities, and saturations) uniformly distributed within strata, as well as reservoir pressure;
- Fluid injection and production is assumed to be provided at certain areas of reservoir where these fluids are concentrated;
- All processes within the tank are considered as isothermal;
- Direction of fluids flow and distribution of wells in the reservoir are not taken into account.

These assumptions set a basis for creation of a simple reservoir model for the further analysis and generation of cogent results. Simplicity of the model and support of physical processes makes it a very popular tool that provides insight into reservoir performance. Nowadays, material balance methods allow to resolve the whole set of engineering tasks, such as:

- Assessment of oil and gas original volume in place;
- Determination of the presence, type, and size of aquifer and gas cap and depth of gas-oil, water-oil, and gas-water contacts;
- Forecasting production characteristics of the reservoir, such as pressure and future production, with respect to different recovery drive mechanisms, and recovery factor calculations;
- History matching of reservoir drive mechanisms.

Depending on the particular problem being solved and required governing equations for it's solution, the following parameters are used in material balance simulation (Satter, Iqbal, and Buchwalter, 2008):

- Reservoir geometry – area and thickness;
- Rock properties – average porosity and saturation, compressibility, and absolute and relative permeabilities;
- Fluid properties – oil, gas, and water compressibilities, solubilities, formation volume factors and viscosities related to pressure;
- Production and injection data - oil, gas, and water production and injection rates and pressures over the time, cumulative values of produced and injected fluids.

Although material balance method is quite simple and convenient tool in the reservoir characterization, its use must be taken with care in certain cases. This requirement emanates from methodological assumptions mentioned above. For example, estimation of fluid in place can be very inaccurate when significant heterogeneity of the reservoir exists. Moreover, it may give inadequate results in the study of fluid reinjection at the late reservoir life stage when fluid production involves water extraction from the aquifer.

2.5 Streamline Simulation

Streamline simulation methodology is essentially a simplification of the finite difference simulators where pressure change in reservoir is analyzed on a finite difference grid, but fluid movement and saturation change is analyzed along the flow lines (streamlines) that coincide with the fluid flow direction in the reservoir. Conceptually, the stream line simulation is a faster substitution of the finite-difference method even though they both use the same reservoir model with a similar set of variables.

Streamline simulation methods are based on the concept of particle tracking to design fluid flow path lines in the reservoir using time (time of flight) of tracer particle travel along these lines. The use of time of flight as a spatial coordinate variable allows to segregate mathematically a complex physics of flow transport from the reservoir heterogeneity, which is the key feature of this method. Another aspect of streamline simulation is that the time of flight coordinates can be used for the fluid flow visualization in three-dimensional space, which is extremely useful in solving such practical tasks as fluid front analysis, pattern balancing, and wells allocation. (Datta-Gupta and King, 2007).

Once streamlines designed, the convection-dominated spatial flow calculations in form of transport equations (saturation and concentration) are executed in 1-D along the individual streamlines and therefore can be performed faster. Further, these streamlines should not necessarily be rebuilt every simulation run, they can be used in multiple simulations until a change in well conditions occurs.

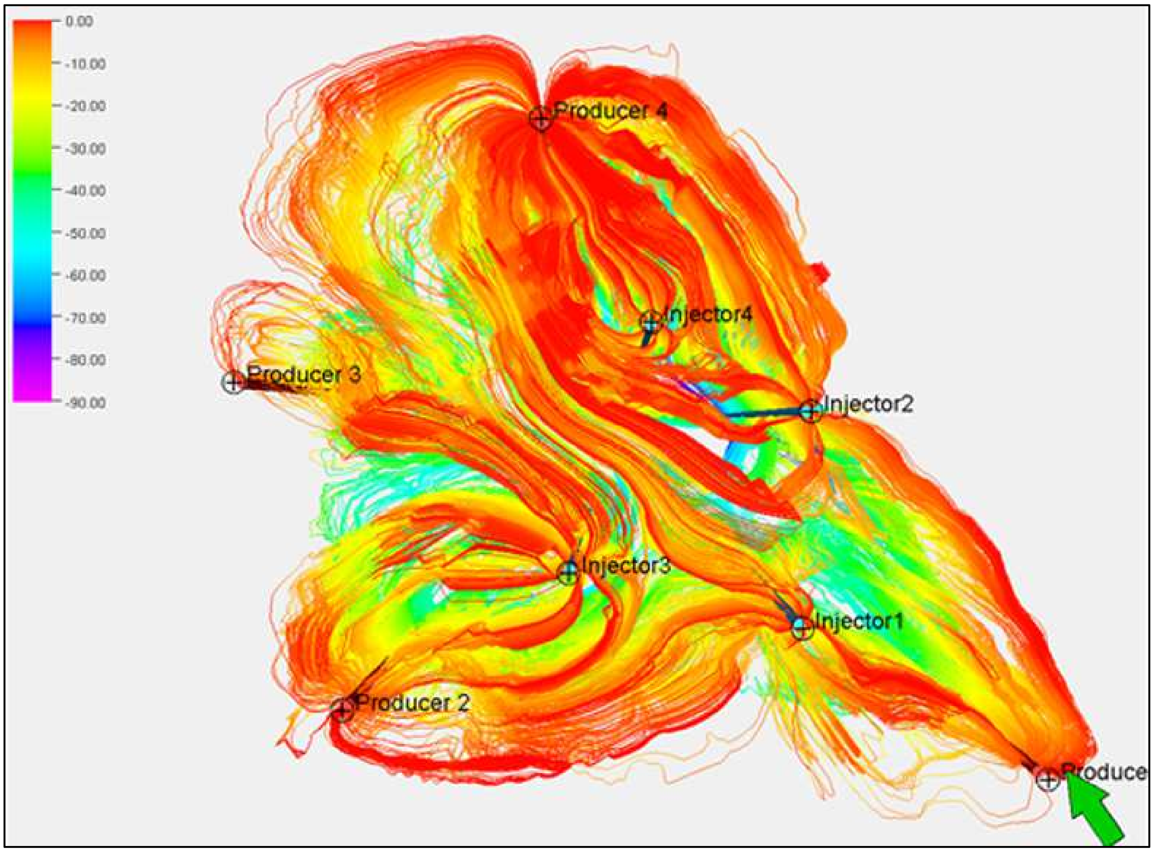

Figure 2.5 – Streamlines for Emerald 1380 synthetic case study

An example in **Figure 2.5** shows the visualization of the streamlines spatial distribution between injection and production wells in color scale with respect to depth. Such distribution provides an outstanding advantage in swept area and volume calculations that can be useful in flood optimization and pattern balancing.

Initial development of 3D two-phase streamline simulation techniques to model reservoir heterogeneity, changing well conditions, black oil displacement, and water flooding using numerical solutions along streamlines, gradually obtained such improvements as: ability to model dispersive transport flow, separate gravity and capillary terms from the convective ones, deal with capillary and gravity effects, perform simulation of dual-porosity and fractured reservoirs, and model compressible fluid flow, CO_2 injection and polymer flooding (Batycky, 1997; Jang, Lee, Choe, and Kang, 2002; Berenblyum, Shapiro, and Jessen, et al., 2003; Di Donato, Huang, and Blunt, 2003; Moreno, Kazemi, and Gilman, 2004; Cheng, Oyerinde, Datta-Gupta, and Milliken, 2006; Obi and Blunt, 2006; Thiele, Batycky, Pöllitzer, and Clemens, 2010).

In addition, streamline techniques were equipped with an ability to perform compositional fluid simulation (Thiele, Batycky, and Blunt, 1997; Crane, Bratvedt, Childs, et al., 2000; Jessen and Orr, 2004; Osako and Datta-Gupta, 2007; Tanaka, Datta-Gupta, and King, 2014).

As the further development of streamline simulation methods, the new trend arose recently. This is solving tasks related to the thermal simulation. Streamlines simulators extended to include thermal effects of temperature dependent parameters, such as viscosity and thermal expansivity, for hot water flooding and steam injection processes related to non-isothermal flow, physical diffusion of gravity, heat conduction, and energy and mass transfer (Pasarai and Arihara, 2005; Zhu, Gerritsen, and Thiele, 2010, 2011; Vicente, Priimenko, and Pires, 2014).

In general, all these techniques demonstrate a good accuracy of obtained results and an advantage in computational time comparing to finite-difference simulation. This is a kernel in the choice making between these two simulation approaches. However, such phenomena as existence of gravity effect, high compressibility, compositional representation of fluids, and complex physics still capable to cause some problems diminishing positive effect of streamline application (Datta-Gupta and King, 2007).

The streamline simulation approach is relatively young compared to other methods, and it is still in a development stage. Nevertheless, due to its computational speed and versatility the streamline simulation became very popular in the following reservoir engineering applications (Datta-Gupta and King, 2007):

- Sweep volume and efficiency calculations;
- Rate allocation and optimization;
- Pattern balancing and delineation of drainage zones;
- Modelling tracer flow, waterflooding, and well placement;
- Calculation of primary and enhanced hydrocarbon recovery;
- Uncertainty quantification, reservoir heterogeneity characterization, and ranking geostatistical models;
- Upgridding and upscaling of geological models;
- History matching with production data integration;
- Solvent flooding and compositional simulation;
- Reservoir management.

Overall, streamline simulators provide advantage of fast flow simulation which is critical when dealing with large models and multiple geologic realization. Flow path visualization and availability of properties such as “time of flight” provides basis for rate allocation, flood-front optimization, proper simulation mode upgridding, and solution to over problems that pose challenges for finite-difference simulation. However, advantages of the streamlines come with certain limitations such as introduction of material balance errors due to properties mapping between the grid and streamlines, limitations of the time step due to non-stationarity of the pressure solution and complexity in dealing with non-convective mechanisms such as gravity, capillarity and phase behavior.

2.6 Finite Difference Simulation

Finite difference simulator is the most versatile tool. Over the decades of use it was improved to account for variety of physical and chemical processes that can occur in reservoirs. This allows us to work with a range of models from very detailed to very coarse resolutions. At its extremes, the finite difference simulator can work with models at geologic scale and models that contain just a few cells and closely resemble material balance models and their functionality. It all depends on the resolution we need, data we have to construct the model, and objectives we are trying to achieve (Aziz and Settari, 1979; Mattax, and Dalton, 1990; Ertekin, Abou-Kassem, and King, 2001; Fanchi, 2006; Mustafiz and Islam, 2008; Islam, Moussavizadegan, Mustafiz, and Abou-Kassem, 2010). Simulation model construction is the most important and time consuming process. The quality of constructed model is critical since it directly defines accuracy and applicability of simulated results.

Engineers of different majors contribute their professional knowledge and experience doing teamwork in data gathering, processing, and integrated reservoir model design (Satter, Iqbal, and Buchwalter, 2008). The primary goal here is to build virtual representation of a real subsurface domain of interest fully described by:

- Three-dimensional reservoir geometry and connectivity;
- Spatial distribution of rock properties: pressure, compressibility, porosity, fluids absolute and relative permeability and initial and residual saturation;
- Types of reservoir fluids and their properties: compressibility, density, viscosity, formation volume factor, solubility, chemical composition, salinity and others;
- Presence and extend of fluid contact zones;
- Well allocation, completion, production and injection operating conditions.

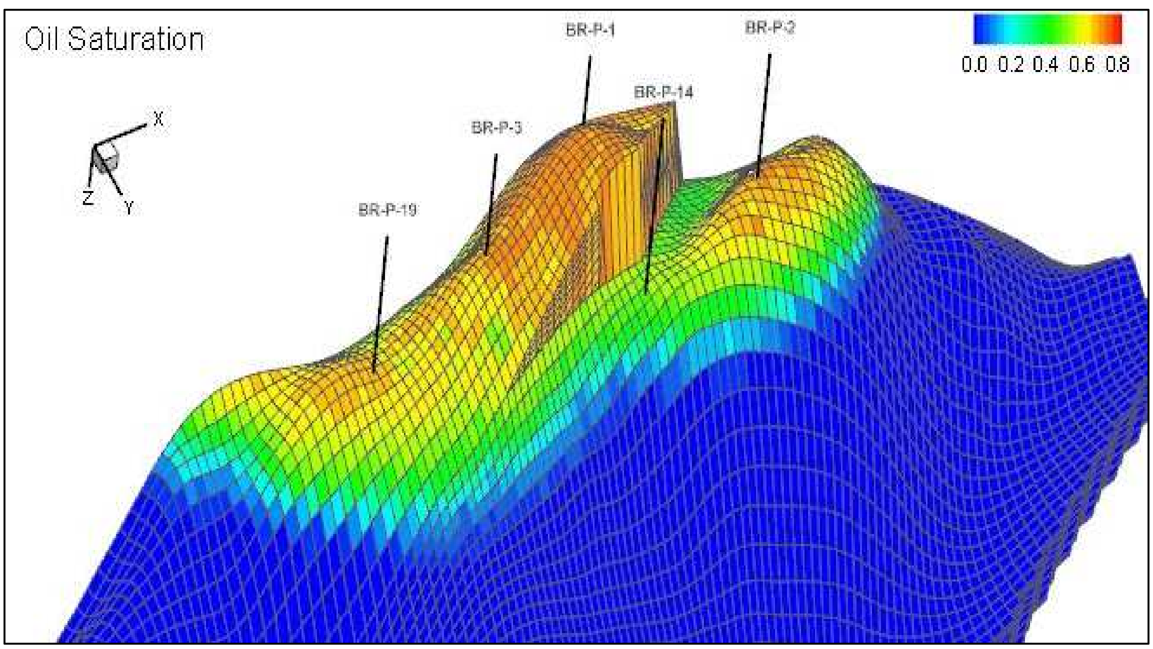

Figure 2.6 – Initial oil saturation for Brugge synthetic case study (adopted form Peters et al., 2009)

An example in **Figure 2.6** shows a typical representation of three-dimensional reservoir model that consists of certain number of grid blocks, where every cell is assigned with particular set of rock and fluid properties. This particular example visualizes a spatial distribution of the initial oil saturation within reservoir in color scale and locations of five production wells.

In general, simulation model is a set of parameters that should be used by the simulator to achieve a particular simulation goal. The selection of these parameters and their properties is based on reasoning about application of the following (Aziz and Settari, 1979; Satter, Iqbal, and Buchwalter, 2008):

- Reservoir geometry model (one-, two-, or three-dimensional) and coordinate system (Cartesian, Cylindrical, or Spherical);
- Representation of fluid type as black oil (including dry gas, wet gas, heavy or volatile oil) or composition (in terms of moles of individual components) with number of phases;
- Description of a flow type in porous media by Darcy’s Law or its extension due to high-velocity effect, slippage effect, and other aberrations;
- Determination of mass and heat transfer mechanisms, such as immiscible fluid flow, phase composition flow, heat flow, mass transport due to dispersion, adsorption, and partitioning.

Once a simulation model is created, it is further sent to the simulator for processing. Finite difference simulator is a computer program that has the ability to solve a set of partial differential equations replaced with finite differences. The following simple example shows a typical isothermal simulator workflow, where finite differences are derived from Taylor's series (Fanchi, 2006):

1. The two-phase fluid flow equations are formulated as:

$$\frac{\partial}{\partial x}\left[\frac{Kk_r}{\mu B}\left(\frac{\partial P}{\partial x}\right)\right]+q_s\delta(x-x_o)=\frac{\partial}{\partial t}\left(\frac{\phi S}{B}\right) \tag{2.4}$$

where: K – absolute permeability of the fluid; k_r – relative permeability of the fluid; μ - fluid viscosity; B – fluid formation volume factor; P – pressure; q_s – fluid flow rate; φ – porosity; S – fluid saturation; x – coordinate along x-axis; t – time coordinate.

2. Derivatives are approximated with finite differences:

   a. Discretize region into grid blocks Δx :

$$\frac{\partial P}{\partial x}\approx\frac{P_{i+1}-P_i}{x_{i+1}-x_i}\equiv\frac{\Delta P}{\Delta x} \tag{2.5}$$

   b. Discretize time into time steps Δt :

$$\frac{\partial S}{\partial t}\approx\frac{S^{n+1}-S^n}{t^{n+1}-t^n}\equiv\frac{\Delta S}{\Delta t} \tag{2.6}$$

where: i – index labeling grid location along x-axis; n – index labeling the present time level, so that n+1 a future time level.

3. Numerically solve the resulting set of linear algebraic equations.

Once the finite difference analogs (2.5) and (2.6) of the partial differential equations obtained, they can be substituted into the flow equations (2.4). Further, the full set of flow equations is rearranged algebraically and solved using numerical methods. As a result of computation, the unknown primary variables, pressure and saturation, are calculated in spatiotemporal coordinates, what allows updating of the pressure-dependent (temperature-dependent for non-isothermal processes) parameters of the model. Iteratively, this process can be repeated many times.

Results of simulation represent the reservoir behavior in a time perspective under particular conditions. The model validation process is usually made by implementation of history-matching procedure, where observed or historical pressure, saturation, and productivity measurements are sequentially matched with simulated ones. In case, when there is no sufficient deviation observed the simulation model can be further used for the reservoir performance prediction including all life stages from exploration to abandonment. Otherwise, some key parameters should be revised and adjusted.

There is no doubt that this type of reservoir simulation is the most popular and powerful in the petroleum industry. It can assist in resolving most of the problems related to reservoir management, field development strategies design, performance prediction, primary and enhanced hydrocarbon recovery evaluation, and many others. However, computational speed is an issue especially for highly heterogeneous models consisting of more than one million grid blocks. Therefore, engineers constantly looking for alternative ways to do the work.

When we talk about the model to be used for finite difference simulation, speed is not the only criterion for selection. Every model comes along with certain simplifications and limitations that can make it a perfect or a bad candidate for use. Selected scale of uncertainty representation (number of components, gridblocks, etc.), objectives of the study, and minimum accuracy of the model can help us in selecting a good substitute for fine scale finite difference simulation.

2.7 Conclusions

Depending on a whole set of aspects of reservoir study, such as field-life stage, appraisal purpose, data and its source different simulation approaches can or cannot be used. They may give significantly different results even at the same given conditions. The sought-for result here implies finding of the proper simulation approach that provides sufficient accuracy, adequate complexity, and representation of the available data with respect to simulation objectives and existing constraints.

3. EXPERT SYSTEMS

The concept of an expert system, as a mean of complex problem solving that requires professional expertise, will be discussed in this section. The evolution of expert systems during last several decades resulted in a wide use of them in different areas including Petroleum Industry. The most common realizations will be discussed to formulate improvement in a decision-making support of simulation approach selection.

3.1 Definition of Expert System and Historical Review

Rapid development of computer technologies has given rise to emergence of a computer science's separate branch that is known as artificial intelligence systems. The term artificial intelligence combines a large set of procedures, principles, and algorithms that implement intelligent behavior based on conscious conclusions. In some ways, it is an attempt to replace the thought process of human by machine language formal logic. In most cases, it comes down to the analysis of a certain amount of information, its processing in accordance to the controlled rules, and the adoption of a final decision (Russel and Norvig, 2010).

The described above procedure suggests an existence of a very important feature that should be an integral part of any artificial intelligent system. This part is called as cognitive skills. Realization of human cognitive function became widespread within computer programs, which rather reason about problems than compute solution. Such approach stimulated emergence and implementation of artificial intelligence systems in a number of applied fields such as medicine, commerce, automation and control, manufacturing, navigation, aerospace, meteorology, and many others. Since 1956, the development of machine intelligence resulted in origin of the following major classes of artificial intelligent systems with respect to the solving tasks and methods used (Krishnamoorthy and Rajeev, 1996; Russel and Norvig, 2010):

- Problem solving and planning – setting goals, selection of the most important, and their hierarchical prioritization;
- Automated reasoning – generation of sensible inferences using accumulated information;
- Natural language processing – generation, analysis, recognition, translation, and grammatical and stylistic manipulation with text and speech;
- Learning – dealing with different types of machine learning to adapt them to new conditions;
- Computer vision – detection, perceiving, visualization, and analysis of objects;
- Robotics – dealing with robotics control;
- Neural networks – emulation of human learning and solution search by aggregating data classification, reasoning, and calculation;
- Genetic algorithms – implementation of adaptive algorithms in solution search, machine learning, and optimization processes;

- Expert systems - imitation of professional expertise in complex decision-making problems, including data classification and reasoning, by knowledge processing in specific area.

According to shown above classification, artificial neural networks and expert systems are more suitable tools for solving problems related to simulation of expert reasoning as a human with expertise. By definition, the expertise is the use of professional skill or knowledge in particular field of interest by a person, who has comprehensive and authoritative qualification. Thus, neural networks and expert systems are capable to determine relation between an input data set and output solution, which is called data classification: they can find an answer to the question whether the given data set belongs to the area of interest or not.

Even though these systems historically were elaborated to reach the same goals – implement machine intelligent behavior and emulate human cognitive ability, they are separated into different classes for several reasons. First, conceptually neural network and expert system are based on different organizational structure:

- neural network represents an array of interconnected elements, neurons, where knowledge is realized by elements connections adjusted by weights;
- expert system is formed by two distinctive modules, in which knowledge and solution search rules are separated.

Second, expert systems have strong advantages comparing to neural networks in dealing with certain tasks, where data classification and reasoning is not enough for solution. More precisely, Krishnamoorthy and Rajeev (1996) and Leibowitz (1997) provide two very important arguments:

1. The most significant weakness of neural networks is that they do not provide interpretation of why the certain inference they create, as that expert systems do. So, neural networks can emulate a human expert behavior limitedly.
2. Due to their structure, expert systems are more suitable in automation of decision-making and solution search in engineering problems solving. Namely, while neural network may require structural rebuilding and retraining in case of new tasks emerging, expert system may need only slight knowledge base and/or inference engine correction that is much faster in time and easier in effort.

Therefore, the necessity of solving issues that require expert judgement in the most approximate to the human expert extent, explanation of obtained conclusions, and flexibility in reconfiguration has created a separate large class within artificial intelligence systems called expert systems.

Giarratano and Riley (2004) proposed the following definition of an expert system as “a computer system that emulates the decision-making ability of a human expert.” In other words, the software tool substitutes the presence of an expert in some problem solving. It should be noted that expert systems have one major difference from other systems of artificial intelligence: they are not intended for solving some of the universal problems since they are designed to provide high quality solution of the certain problem in a specifically defined area.

Historically, the emergence and development of expert systems was associated with cognitive science. This is a study of human (expert) thinking process in problem solving. Since the late 1950-s, when Newell and Simon demonstrated that the most of human decision-making solutions are based on “IF-THEN” type production rules, the next several decades significantly contributed in expert systems evolution (Giarratano and Riley, 2004). The major stages in expert systems evolution are shown in **Figure 3.1**.

Starting with implementation of very simple programming algorithms, expert systems step-by-step obtained its personal language, complex logic, system shell, knowledge base, and inference and search engines. All these components, widely used in modern expert systems, resulted in conversion of initially quite simple computer programs to powerful software tools and applications (Giarratano and Riley, 2004). Badiru and Cheung (2002) pointed out that nowadays a new trend in expert systems design can be observed. Namely, expert systems are not created and used as independent software applications, but as constituent of software complex that may include more than one system. For instance, there are several commercial packages equipped with scilicet database and management, information management, statistical analysis, data analysis, and project management expert systems. The corresponding example of the modern expert system realization can be easily found on the Internet, which is Google or Yahoo search engines.

2000	Intelligent agents - Web, E-commerce, Search engines
1998	More than 12 500 expert systems (USA only)
1996	GUESS - Generically Used Expert Scheduling System
1994	Commercial Artificial Neural Systems
1993	Lumiere Project (Microsoft)
1990	Artificial Neural Systems
1985	CLIPS expert system tool (NASA)
1983	KEE expert system tool (IntelliCorp)
	Japan - Fifth Generation Project (intelligent computers)
	Hopfield Neural Net
1982	SMP math expert system
1980	LMI and symbolic founded for manufacturing LISP machines
	Inference Corp formed
	Artificial Intelligence goes commercial
1979	RETE Algorithm for fast pattern matching (Forgy)
	Meta-DENDRAL - metarules and rule induction (Buchanan)
1978	XCON/R1 (DEC computer system configuration) started (McDermott, DEC)
1977	OPS expert system shell for XCON/R1 (Forgy)
	Theory of Reasoning under Uncertainty (Dempster-Shafer)
	PROSPECTOR expert system started (Duda, Hart)
1976	Artificial Matematician (AM) discovery of math concepts (Lenat)
1975	Frames and knowledge representation (Minsky)
	HEARSAY II - cooperating experts blackboard model
	EMYCIN - expert system shell (Van Mele, Shortliffe, Buchanan)
	TEIRESIAS - explanation facility concept (Davis)
	GUIDON (intelligent tutoring) (Clancey)
1973	MYCIN expert system for medical diagnosis (Shortliffe)
	Human Problem Solving (Newell, Simon)
1971	HHEARSAY I (speech recognition)
1970	PROLOG work started(Colmerauer, Rousell)
	Perceptrons (Minsky and Papert)
1969	MACSYMA math expert system (Martin, Moses)
1968	Semantic nets, associative memory model (Quillian)
	DENDRAL started (Feigenbaum, Buchanan)
	Fuzzy logic (Zadeh)
1965	Automatic theorem Proving via Resolution Method (Robinson)
1962	Principles of Neurodynamics on perceptions (Rosenblatt)
1958	LISP artificial intelligence language invented (McCarthy)
1957	Perceptron invented. General Problem Solver started (Newell, Shaw, Simon)
1956	Artificial intelligence term is used. Logistic Theorist. Heuristic Search
1954	Markov Algorithm (control rule execution)
1943	Post production rules. The neuron model (McCulloch, Pitts)

Figure 3.1 – Milestones in the expert systems history (adopted from Noran; Giarratano and Riley, 2004)

3.2 Classification, Structure, and Design of Expert Systems

Depending on the specific tasks being solved by the expert systems, Hayes-Roth and Waterman (1983) proposed the following classification: interpreting, forecasting, diagnosing, designing, planning, monitoring, instruction, controlling, debugging, and repair systems. Since this classification allow overlapping and combining of specific tasks due to their inseparability, it was reworked by Clancey (1985) and is used nowadays – the following is the list of tasks where expert systems are effectively used:

- Classification – determination of an object belonging to particular area of interest (clustering) based on defined characteristics;
- Diagnosis – elicitation of nature and causes of the problem by examination of observed data;
- Monitoring – observation and checking the system progress or quality over a period of time to describe behavior of process;
- Process control – management by a process based on monitoring;
- Design – configuring an object in accordance to certain exposition;
- Scheduling and planning – design or modification of a workflow or actions depending on estimated conditions;
- Generation of options – creation of alternative decisions to a given task.

The presented list is not exhaustive because continuous evolution of expert systems engenders brand new tasks to be feasible. Nevertheless, this classification gives a clear idea about area of systems application. Considering the solving task of simulation approach selection, which is based on expert reasoning and should involve an explanation of made decisions, we can conclude that the two main goals of this project – data classification and generation of options – can be realized using expert system. On the next step of our search we should define a structure of the system, which will provide the optimal configuration to be developed in accordance to the project objectives.

Being a computer program, the expert system is called a “system”, not just a “program”, since it consists of several major components:

- a knowledge base that stores information required for a task solution;
- an inference engine;
- additionally, it may include an explanation module that provides description of how the system makes recommendation.

The knowledge base is the foundation of any expert system, which is compiled based on the professional expert knowledge. According to Engelmore and Feigenbaum (1993), the knowledge base is the set of factual and heuristic knowledge. The factual knowledge is widely shared in different sources, such as textbook, journals, and articles and have common implementation in the field of study. In contrast, the heuristic knowledge is more specific, individualistic, and based on experimental and practical performance of good judgement as well as very similar reasoning it the field. The expert cognition here is the combination of theoretical understanding of the certain problem and practical skills of its solving, which effectiveness is proven in a result of the practical work. Properly selected expert and successful formalization of their knowledge endow the expert system unique and valuable knowledge.

The inference engine usually represents a set of applied rules, such as match, select, execute etc. It is built as a set of algorithms, which provide suggestion about ways of posed problem solution based on knowledge base and input data set juxtaposition.

Structurally, all the diversity of expert systems is divided in two large groups based on their knowledge base construction principles: knowledge-based and rule-based systems. Although both groups have many common features, they are different.

The knowledge-based system (**Figure 3.2**) are used for creation of very powerful expert systems (Engelmore and Feigenbaum, 1993). Here, the knowledge base consists of set of various complex objects which characteristics and types have specific relationships. In other words, every object in the knowledge base is a combination of parameters, encoded in a certain manner, that describes a composition of data variables and cases of their use with respect to particular problems. It is a virtual representation of an expert judgment on the possibility to solve a particular problem with a specific input data set.

In the rule-based expert system (**Figure 3.3**), the knowledge base is represented by a set of production rules, where a group of simple "IF-THEN" statements represents knowledge (Engelmore and Feigenbaum, 1993). In general, the production rule consists of a condition (prerequisite) expressed by "IF" and conclusion (action) denoted by "THEN" (Giarratano and Riley, 2004). In the case of several dependent rules, they might be organized in the form of a decision-making tree.

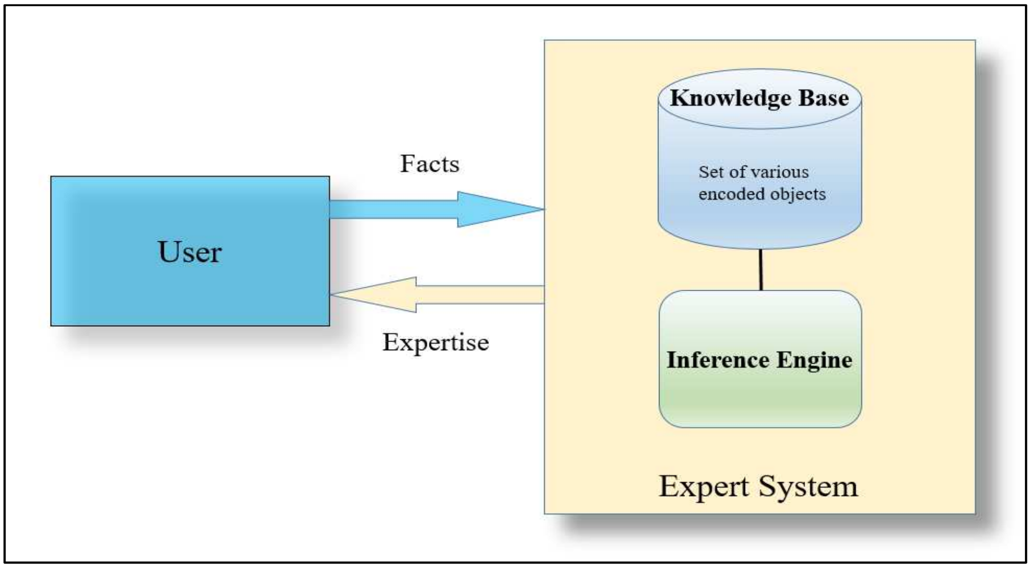

Figure 3.2 – The basic concept of a knowledge-based expert system (reworked from Giarratano and Riley, 2004)

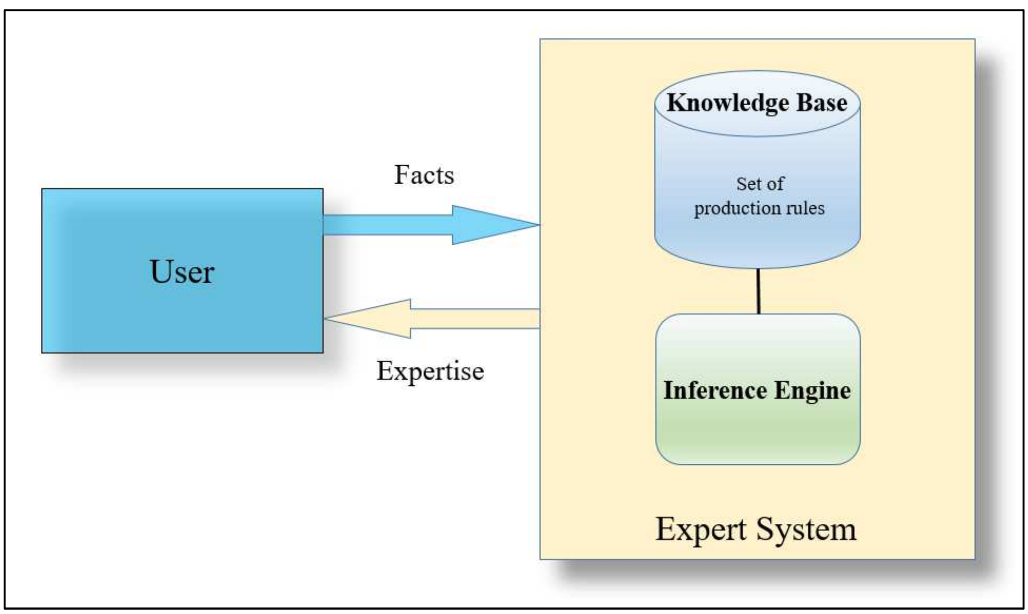

Figure 3.3 – The basic concept of a rule-based expert system (reworked from Giarratano and Riley, 2004)

As shown above, the rule-based systems are slightly different from the knowledge-based ones in structure, but sufficiently different in content of the knowledge base. In practice, this distinction affects expert system functionality and, consequently, area of applicability with respect to the resolving task. The summary of these distinctions is presented on **Table 3.1**.

Table 3.1: Summary of the key differences between rule-based and knowledge based systems (reworked from BizRules, 2006-2007)

	Can process	**Can output**	**Best for applications in**
Rule-Based System	Data, Rules	Information, Decisions, Real-Time Decisions	Decision-making, Compliance
Knowledge-Based System	Data, Rules, Knowledge	Information, Decisions, Real-Time Decisions, Expert Advice, Recommendations	Advising, Decision-making, Solution Selection, Recommending, Troubleshooting

As it mentioned above and can be inferred from the **Table 3.1**, knowledge-based systems have a very significant component comparing to the rule-based systems, what is knowledge. This “real” knowledge base seriously extends the expert system’s potential in output results obtaining and area of applicability

The above **Figures 3.2** and **3.3** represent the basic concept of interaction between a user and an expert system. The user gives some facts as input data to the expert system in order to obtain solution for particular task. Using the inference engine, the expert system processes the user's data, collating it with data set in the knowledge base, and making logical conclusions. The obtained solution returns to the user as a result of expertise, which can be either a solution for a given task or a conclusion about problem solvability and recommendation on what to do.

This example demonstrates one of two possible working modes, so-called "consulting regime", when the user applies to the expert system for problem solving. In this particular case, the user can be:

- non-professional in the area of interest, and he asks the expert system to find solution that he cannot get by himself;
- professional in the problematic area, but he uses the expert system as a part of routine work to speed up result finding.

Another working mode is called "teaching/training regime", when an expert works with the expert system instead of user. In this case, the expert describes problematic area with a set of facts and rules locating them in the knowledge base and inference engine. In other words, he fills out the expert system with knowledge that further allow solving the described problem independently of the expert. This mode is usually implemented during initial formation and filling of knowledge base and inference engine or when any correction of their content or structure is required.

In modern expert systems (Duggal and Chhabra, 2002; Kaimal et al, 2014), the training regime has a tendency to be automated by introduction of a learning engine in the system's interface, as is shown in **Figure 3.4**. Such extension of the system is usually made by the application of machine learning algorithms. This ability is especially valuable in changing conditions, when:

- the range of solving problems has tendency to expand;
- obtained results require correction of the knowledge base and/or inference engine;
- system adjustment procedures, such as modification, tuning, and training are too complex and require simplification.

In other words, these advantages provide possibility to increase the expert system's level of confidence.

Shown in **Figure 3.4** is an example that represents the following algorithm of self-learning regime. Initial data from user set is pre-processed in data interface and is then inputted into database and inference engine. Expert system finds solution for a given task and brings it to a graphical user interface. Simultaneously, the learning engine compares the system's output with other ones stored in database which have the same conditions with respect to input data set and solving problem. If the database response has good agreement with the system's output, then the learning engine perceives this situation as normal and does not require any additional action. Otherwise, the knowledge base is corrected by introduction of new or correction of existing rules (rule-based system) or objects (knowledge-based system) in the knowledge base via the use of specially designed algorithms.

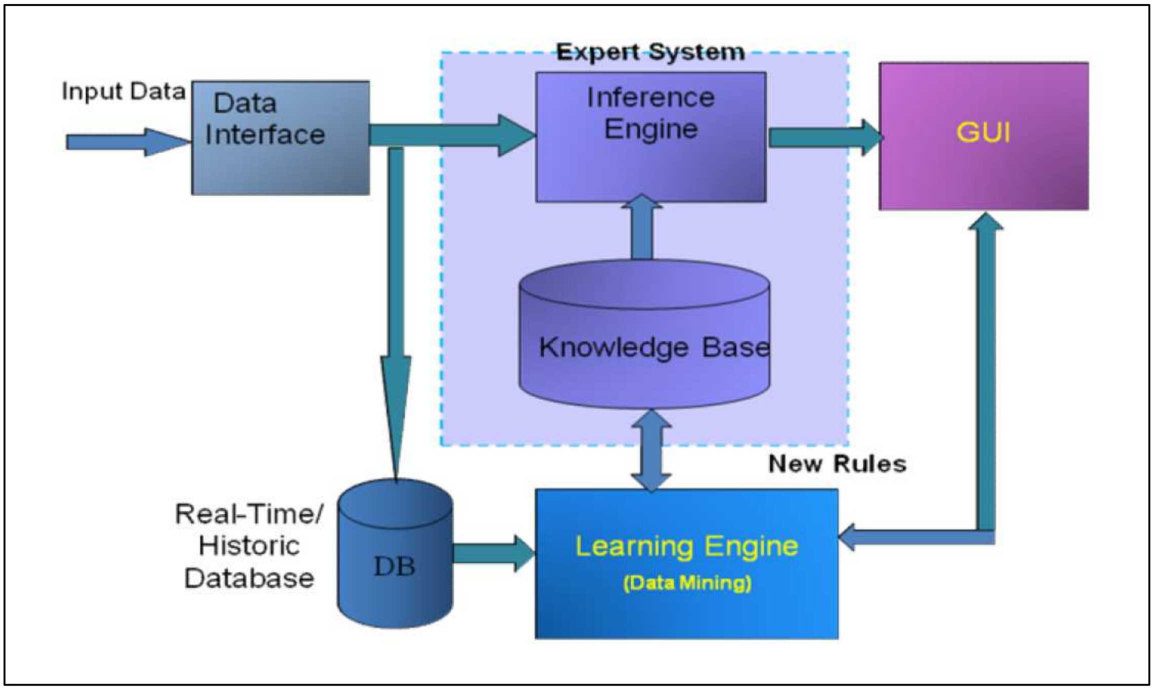

Figure 3.4 – Architecture diagram of expert system with learning engine (reprinted from Kaimal et al, 2014)

Thereby, an initially well-built expert system supplemented by a learning engine has the opportunity to educate itself on the problem solving via adding corrections into the knowledge base in conformity with obtained results, conclusions, and decisions.

Overall, an essential part of any expert system design is the development of knowledge base and inference engine (Giarratano and Riley, 2004). In general, this process implies multi-criterion data analysis and classification. Complex data structure, where parameters are represented not only by their measures, but also their connections, requests availability of certain procedures or algorithms to evaluate whether set of parameters is passable and satisfies the overall requirements or not. Therefore, it is very important to define initially the optimal method of data encoding that further allows the suitable solution search.

Addressing the issue of the data encoding method selection, Giarratano and Riley (2004) state that in contrast to some computer programs, which use just numerical calculations, the "expert systems are primarily designed for symbolic reasoning". Siller and Buckley (2005) note that "key to expert systems (and to artificial intelligence, for that matter) is the concept of reasoning with symbols." Many programming (procedural) languages, such as C, FORTRAN, and others can represent specific symbols in numerical or character strings data or even in complex objects. Nevertheless, for the purposes of expert systems design the more appropriate languages for symbolic reasoning (manipulation), than procedural languages, are LISP or PROLOG where "symbols can represent almost anything" (Siller and Buckley, 2005). In essence, these logic languages deal with syntactic structures, where:

- variables are denoted in string of letters;
- relations between them are defined by clauses;
- solution search logic is expressed by specific query over variables and relations.

Expressing data in the form of syntactic structure, as a natural language, has big advantage since it allows to construct logical representation or description of objects. It opens an opportunity to distinguish common or unique features for different objects, what simplifies their clustering for the purposes of further classification and inferring new facts about objects. Additionally, solution search logic is built in finding match of assigned criteria within object features.

For instance, there are several objects that can be described as “red car with four wheels is vehicle”, “blue bicycle with two wheels belongs to Mary”, and “yellow truck with eighteen wheels is long vehicle”. Relations between variables in these objects can be assigned as following: “IF wheels THEN car”, “IF wheels THEN bicycle”, IF wheels THEN truck”, and “IF wheels THEN vehicle”. Finding the answer to the question of whether bicycle is also vehicle or not, analysis of the given above common object features and assigned clauses leads to the next conclusion. Because bicycle has “wheels” as the other objects, defined as vehicle, it can be classified as “vehicle”.

This primitive example demonstrates a very simple case of symbolic reasoning used in expert systems. In practice, the objects structure and relations between their variables are more complex, what is directly depends on the required expert system functionality and the area of solving problems.

Vast majority of systems that are effectively employed today were built using languages of symbolic manipulation. For instance, PROLOG is implemented in all known operating systems and platforms, including Unix, Windows, Java and .NET.

3.3 The use of Expert Systems in Petroleum Industry

According to Waterman (1986) and Leibowitz (1997), the design and use of expert systems in Geology and Petroleum Industry began in late 1970s – early 1980s.

SRI International developed the very first system, named PROSPECTOR, in 1978. This system interprets geologic data in order to evaluate an existence of certain minerals in the region of interest. In 1981, Schlumberger-Doll Research Centre in association with Fairchild Labs for AI Research and MIT created DIPMETER Advisor. Interpreting dipmeter logs, this system shows information about geological structure around the well with respect to depth. Two years later, Schlumberger developed another expert system called LITHO. By using records of oil-well log data, this system issues description of the most plausible lithofacies detected in vertical lithological column. At the conference "Applications of Artificial Intelligence" (Denver, 1984), G. Khan and J. McDermott presented MUD expert system that was developed in collaboration of Carnegie Mellon University and NL Baroid. The main mission of MUD was to diagnose and remedy drilling problems via providing optimal properties of drilling mud.

Proving ability to solve complex engineering problems quickly and accurately and being easy to use, expert systems gained a lot of popularity. Starting 1986, there were several dozens of expert systems designed for use in various areas of Petroleum Engineering. The use of them allows solving a wide range of tasks, but only in highly specialized subject areas. Functionally, existing systems can be assigned to following groups:

- reservoir characterization (Erdle, Archer, Stiff, et al., 1986; Whittaker and Macpherson, 1986; Dharan, Turek, Vogel, 1989; Sanjay, Anuj, Sharma, 1989; Mabile, Hamelin, du Chaffaut, et al., 1989; Al-Kaabl, McVay, Lee, 1990; Kjell and Baleix, 1992; Surguchev, Zolotukhin, Bratvold, 1992; Garrouch, Malallah, AlEnizy, 2006; Nashawi and Malallah, 2009);
- drilling, completion, and production operations control (Martinez, 1992; Martinez, Moreno, Castillo, et al., 1993; d'Almeida, Silva, Ramos, 1997; Denney, 1999; Pandey, Osisanya, 2001; Al-yami, Schubert, 2012);
- drilling and workover operations design (Van Domelen, Ford, Chiu, 1992; Heinze, 1993; Kulakofsky, Wu, Onan, et al., 1993; Balch, Weiss, Ruan, et al., 2003; AlMousa, Ertekin, 2013);
- selection and optimization of enhanced oil recovery techniques (Guerillot, 1988; Khan, Pope, Sepehrnoori, 1993; Sheremetov, Cosultchi, Batyrshin, et al., 2007);
- reservoir performance prediction (Srinivasan, Ertekin, 2008; Moridis, Kuzma-Anderson, Reagan, et al., 2011; AlMousa, Ertekin, 2013; Siripatrachai, Rana, Bodipat, and Ertekin, 2014).

In general, these areas are well studied and provide clear strategy of decision-making.

Nowadays, the growth of expert systems quantity is diminishing. Developments in science and technology complicate the type of problems and approaches to their solution, what in turn significantly expands domain of required knowledge and experience. That is triggering natural constraints of expert systems:

- Transfer of deep knowledge about subject area to the expert system is not a trivial task due to the complexity of experts' heuristic knowledge formalization.
- Frequent involvement of software developers is required to support expert systems in actual condition, especially when problem-solving environment is changing. Without developers support systems quickly lose their relevance.

Nevertheless, despite all of these constraints expert systems have already proven its value and irreplaceability in some important applications.

3.4 Conclusions

The expert system is a good means for problem solving that requires expertise. The basis of any expert system is a complex of knowledge, which is structured in order to facilitate the decision-making process. Simultaneous application of input data analysis and expert knowledge and skills in making decisions, conclusions, predictions, and recommendations can be realized via knowledge base and inference engine creation.

Knowledge-based expert systems, comparing to rule-based ones, benefit in application development in which the use of composite functionality, including decision-making, solution search, and recommendations development, is required.

Overall, the knowledge-based expert system to be designed as a decision-making support in simulation approach selection. The symbolic (linguistic) data encoding and processing to be used in the system to make it effective and further designate the proper solution search.

4. DESIGN OF THE EXPERT SYSTEM FOR SIMULATION APPROACH SELECTION

This section discusses components of the expert system that have been developed and implemented in framework. We also describe the linguistic method of the Pattern Recognition Theory as a means that determines data encoding algorithm, knowledge base content, and solution search procedures with symbolic reasoning. In addition, we present the methodology for a new well placement justification as an extension of the expert system functionality.

4.1 Workflow Steps

As we mentioned earlier, the primary objective of the work is to formalize, design, and test the reliable methodology and software tool to provide decision-making support in simulation approach selection. This task is non-trivial since it requires emulation of a human cognitive ability in thorough data analysis and the appropriate simulation method selection. The complexity of this topic is caused by the need to design and implement algorithms of data processing and encoding, which forms the basis of the expert system's functionality. More precisely, the fulfillment of this task consist of the following stages:

- knowledge base design, including the alphabet and library creation;
- inference engine development as a set of data processing and matching procedures;

- assignment of the expert system functionality.

The basic concept of a knowledge-based expert system, shown in **Figure 3.2**, represents the conceptual image of the further framework of the system design. Additionally, it displays interaction between a user and the system. The user gives some facts as input data to the expert system in order to obtain a solution for a certain task. Using the inference engine, the expert system processes user's data collating it with another dataset in the knowledge base and making logical conclusions. The obtained solution returns to the user as a result of emulated human expertise.

4.2 Symbolic (Linguistic) Data Encoding

In the previous section we concluded that the use of symbolic (linguistic) data encoding and processing method makes the expert system effective and further designate the proper solution search. The LISP and PROLOG languages could be used for symbolic manipulation. Since these languages are too complex and cumbersome for use in this work, the alternative approach can be implemented and tested for the knowledge base and inference engine design. This opportunity is provided by the linguistic method of the Pattern Recognition Theory.

The linguistic approach is particularly useful dealing with objects which cannot be described by only numerical measurements or have complex structure as mixture of quantitative, qualitative, and perhaps structural or logical characteristics. This ability to encode, combine, and process data of different nature equips the linguistic method with indubitable advantage in solution search comparing to other ones, such as heuristic or mathematical (Tou and Gonzalez, 1974; Pearl, 1984; Chaban, 2004; Lepsky and Bronevich, 2009; Russel and Norvig, 2010; Martí and Reinelt, 2011).

The main goal of recognition procedure is the answer to the question whether the object, described with specified characteristics, is related to the certain category of interest, and if yes, to which one? (Chaban, 2004) In our work, the recognition process is a search for conformity between the specific simulation method and the problem to be solved with a given set of data (object).

In a very general case, any information model of an object, phenomena or process in the real or abstract world can be considered as a pattern (scenario). A distinctive feature of such model in the recognition task is the use of only exact objects characteristics subset which provides selection of one or several particular object type groups. A full set of the most informative features that fully describes an object is called an alphabet (Lepsky and Bronevich, 2009).

Any recognition algorithm can be expressed as the following abstract function (Chaban, 2004):

$$R = \{A, S, P\} \tag{4.1}$$

where in regard to the linguistic method: A – alphabet; the variety of uniquely encoded objects characteristics; S – scenarios; the variety of alphabet elements combined into possible patterns that uniquely describe object of interest; P – inferences; the variety of decision making rules.

In accordance to the expression (4.1), the further stages of the methodology and expert system design are reduced to the following steps:

- Alphabet design – selection of key parameters (A) involved into the particular reservoir evaluation problems solving; then, encoding them with unique symbolic names.
- Knowledge base (vocabulary) design – a set of scenarios generation (S); combining the alphabet elements into the particular sequences that define requirements to the data quality and sufficiency in the certain problem solving, accuracy of output results, and computational speed with respect to every simulation approach.
- Inference engine design – development of certain rules (P) that generate conclusions about which simulation approach should be used as optimal with a given input data set and/or provide suggestion on what should be additionally done to make other methods applicable.

Once the alphabet is created, then using its linguistic variables the composition of patterns (scenarios) is designed in a form of the parametric sets. As a result, every scenario uniquely represents an ability to use the exact simulation approach for the simulation goal achievement depending on the given set of field data and possible constraints. All the generated scenarios are put into the library that is called the knowledge base. Virtually, all these patterns are automatically combined into separate clusters, where each cluster represents the simulation approach that eliminates necessity of decisive function use in solution search (Lepsky and Bronevich, 2009).

Realization of the described above procedures turns the inference engine design into a quite simple task. In general, its implementation reduces to analyzing the match between patterns in the library with another one generated by the expert system through the user data processing. The user's data here is nothing more than the input data set of variables that he/she has, such as rock and fluid properties, production data, simulation objectives etc. In the case of full pattern match being obtained or part of the scenario being matched with a certain cluster from the vocabulary, decision regarding the simulation approach to be used is obvious. Otherwise, the library and/or vocabulary should be revised and adjusted by an expert because previously unknown/undescribed scenario has been met. That process is called training.

The principle of comparing with an etalon (scenarios in the knowledge base), as a match finding procedure, is used because it provides a tuning flexibility and possibility to create an adaptive regulation of decision-making (recognition) process. In addition, it allows the creation of an explanatory module that can generate comments on why the expert system made the certain decision and different recommendations on the problem solving workflow.

4.3 Knowledge Base Design

With regard to our work, the realization of the knowledge base involves several stages. The first one is the alphabet design, which includes:

- selection of parameters that are required to be used for each of simulation approaches;
- selection of constraints that provide some limitations in use of the certain simulation approach;
- selection of certain simulation goals that can be reached by the certain simulation approach;
- parametrization of selected data via encoding into the linguistic (symbolic) variables.

Depending on the methodology put to the basis of every simulation approach, described in section 2 of this work, the following major groups of required parameters were selected: simulation goals, reservoir rock properties, reservoir geometrical data, fluid properties, saturations and relative permeabilities, fluid types, initial volume of fluids, fluid contacts, production data, injection data, and number of production/injection wells. Additionally, there are several groups of parameters estimated that can affect the applicability and accuracy of the certain simulation method or can be considered as constrains: field maturity (life stage), reservoir heterogeneity level, source of rock and fluid properties, and set of special constraints that may be considered or ignored by user during data processing. All these parameters are coded by assigning them unique linguistic names. Thus, the alphabet is created, where 126 elements are combined in 16 groups. The example of parameters coding for two groups (fluid properties and specific constraints) is shown on **Table 4.1**. The full alphabet is presented on **Table A-1** (APPENDIX A).

Table 4.1: Fragment of the alphabet with encoded parameters' names for two groups (fluid properties and specific constraints)

Fluid properties - FP		**Constraints - CS**	
ODN	oil density [lb/cu.ft]	CTA	computational time advantage [G - good, P - poor]
WDN	water density [lb/cu.ft]	FSE	field scheduled events [affects predictability]
GDN	gas density [lb/cu.ft]	FLS	field life stage - goes from FM (field maturity)
OFR	oil formation volume factor related to pressure [rb/STB]	ACC	accuracy [L - limited, F - fair, G - good, B - the best]
WFR	water formation volume factor related to pressure [rb/STB]	CPH	complex physics
GFR	gas formation volume factor related to pressure [rb/SCF]	GRA	gravity effect is exist (ODN>=WDN at surface) (less than 10 yields constrain for stream-line)
GSR	gas solution in oil related to pressure [SCF/STB]	PWC	critical value of the water cut
OCM	oil compressibility [1/psi]	RDM	recovery drive mechanism (W - water, G - gas cap, S - solution gas drive, E - oil expansion drive)
WCM	water compressibility [1/psi]	PDE	production decline is established (Y - yes, N - no)
GCM	gas compressibility [1/psi]	NGD	number of grid blocks (more than 100000 for black oil - advantage in CTA for streamline vs FD)
OVS	oil viscosity [cP]	HTL	level of heterogeneity by Dykstra-Parson [0 … 1] (HTL>0.25 limits use of MBL; HTG > 0.5 advantage in use of streamline vs FD for black oil)
WVS	water viscosity [cP]		
GVS	gas viscosity [cP]		
GDF	gas deviation factor		

It must be noted that the created alphabet is not exhaustive. In case, when new elements have to be added or existing ones to be eliminated for some reason, the alphabet content can be revised and corrected.

In the next stage, scenarios that uniquely describe required set of parameters for solving the certain simulation problem with respect to exact simulation approach were generated. Basically, the scenario is a combination of parameters that methodologically are required for the particular solution. These parameters include available field data and constraints, which we described in section 2 of the thesis. Scenarios design is a very important part of the entire work since being a key to success of the knowledge base creation it directly determines the expert system level of confidence. Moreover, this procedure is exactly the process of theoretical knowledge and practical experience integration. In other words, generating each scenario we reproduce the same reasoning as a human expert on:

- What parameters are required to solve particular problem;
- Whether an amount of available data is sufficient or not;
- What is the accuracy of solution should be considering source of data, field-life stage, reservoir complexity, and constraints;
- Whether the certain simulation approach is applicable or not at the given conditions;
- Is it possible to obtain results using only one method or there several ones should be implemented as a multistage solution finding;
- How fast the sought-for results can be computed using the certain type of simulation;
- Overall, what simulation approach should be selected as the most appropriate in the given conditions, and/or what additionally should be done to make other methods applicable and improve quality of simulated results.

Specifically, for this procedure implementation, scenario generator was created with the following functionality. On the worksheet, shown in **Figure 4.1**, the names of alphabet parameters are located in upper part of each column. Depending on the necessity to introduce new variables or delete unused ones in the alphabet, the number of columns can change. There is no specific requirement for parameters ordering within a row and they can be organized in columns randomly. This is a very convenient feature because it allows an easy generator modification and flexibility in the scope expansion. Names of simulation task and related simulation approach are put in the first and second columns, respectively.

		RG	-reservoir		RP	-reservoir properties						- sat	RH	FT	FP				PD				WN	- wel	CS		
		.GBM	.ARE	.THC	.PER	.POR	.WST	.OST	.GST	.RCM	.RPI	.OSR	.HMG	.OIL	.ODN	.OFR	.OCM	.OVS	.PROF	.PWF	.PTM	.PDE	.PWN	.IWN	.CTA	.ACC	.FLS
DHP	CDC			1				1		1	1			1		1	1		1	1	1	O	1			F	
DHP	PRX																										
DHP	MBL		1	1		1		1		1	1		1	1	1	1	1		1	1			1	1		L	E
DHP	STL	1			1	1		1		1	1	1		1	1	1	1	1					1	1		G	
DHP	FDV	1			1	1		1		1	1	1		1	1	1	1	1					1	1		B	
DAE	CDC			1				1		1	1			1		1	1		1	1	1	O	1			F	
DAE	PRX																										
DAE	MBL																										
DAE	STL	1			1	1		1		1	1	1		1	1	1	1	1					1	1		G	
DAE	FDV	1			1	1		1		1	1	1		1	1	1	1	1					1	1		B	

Figure 4.1 – Scenario generator worksheet: an example of parameters distribution for hydrocarbons in place and drainage area estimation using five major simulation methods

Depending on the combination of simulation goal and simulation approach, an expert qualifies applicability of the certain parameter and constraints that may affect the output result. Thus, the transformation of the knowledge into parametric combination occurs. Once the worksheet is filled in, the unique combinations of data from rows and columns are automatically integrated as scenarios using Visual Basic program code. Further, these patterns are put together into the library (etalon) that represents the knowledge base. An example of the generated scenario for the oil in place estimation using correlation (decline curve) method is shown in **Figure 4.2**.

```
CDC    DHP.THC.OST.RCM.RPI.OIL.OFR.OCM.PROR.PWF.PTM.PDEO.PWN    .ACCF
```

Figure 4.2 – An example of generated scenario

In **Figure 4.2**, the following data is coded as scenario: simulation method (CDC – correlation, decline curve), simulation task (DHP – hydrocarbon in place estimation), required parameters as reservoir thickness (THC), oil saturation (OST), reservoir rock compressibility (RCM), reservoir initial pressure (RPI), fluid type (OIL), fluid formation volume factor (OFR), fluid compressibility (OCM), fluid production rate (PROR), well-bore flowing pressure (PWF), production time (PTM), number of production wells (PWN), an indicator of the method applicability (PDEO – oil production decline is established), predicted accuracy of method (ACCF – is fair).

4.4 Data Pre-processing and Scenario Generation Procedures

In order to improve the sought-for solution search, in this work we decided to additionally introduce the input data pre-processing procedures in the system scenario generation operation. The main goal here is to evaluate not only the existence of the certain parameter within user's input data, but also to assess some of them qualitatively and quantitatively, what is data quality control process. This approach significantly improves the input data analysis since it yields reasonable understanding of why one or another parameter, even if it exists within input data, was not included into system generated scenario, and why additional constraints were introduced in it. In other words, we equipped the Data Pre-processing module with specific procedures that qualify applicability of each parameter in the input data set and its ability to affect accuracy of output results.

Although the quantity of existing data control procedures and number of constraints and limitations for each simulation method are large, for the purposes of our work we selected only several of them to test. Following list of some data pre-processing tasks was formulated:

1. Evaluation of reservoir heterogeneity level by Dykstra-Parson coefficient. High heterogeneity restricts the use of material balance simulation. In combination with large number of grid-blocks it significantly reduces the computational speed of finite difference simulation.
2. Determination of production data availability and applicability. When required fluid production rate data does not exist, but can be obtained via certain data manipulation, the search for production data may be considered successful. For example, if user does have oil and liquid production rates, but does not have required water production rate or water cut, this information can be calculated

using available rates, and vice versa. For the correlation (decline curve) method an applicability of production data is critical. Thus, the presence of established decline for each of wells and stabilized production regime are also evaluated.

3. Appraisal of “complex physics”. The higher oil density with respect to water, compositional fluid representation, relatively high reservoir rock compressibility – all these criteria significantly restrict the use of streamline simulation, since it becomes less accurate and computationally slower comparing to the finite difference simulation.
4. Taking a field-life stage in consideration. In practice, data obtained during early stages is usually characterized with a higher level of uncertainty in comparison to the later ones, which may decrease an accuracy of output results. At the same time, late time production data, used by material balance technique, may give inadequate results in study of fluid reinjection at the late reservoir life stage when fluid production involves water extraction from the aquifer (Satter, Iqbal, and Buchwalter, 2008).

This list of the input data pre-processing procedures is not limited and can be further extended to improve quality of the expert system outputs. In fact, the necessity of new procedures introduction is dictated by evolution of particular problem solving techniques and methods, axillary software tools, and technologies. This progress determines emergence of new or changes in existing assumptions, constraints, and limitations that should be taken in consideration.

In order to realize the mentioned above principle of comparison with etalon, it is required to transform the input data set into the certain form comparable with knowledge base scenarios. Following the same approach logic presented in the section 4.3 and executing the data pre-processing procedures, the expert system generates new scenario from the input data. **Figure 4.3** shows the workflow of this process.

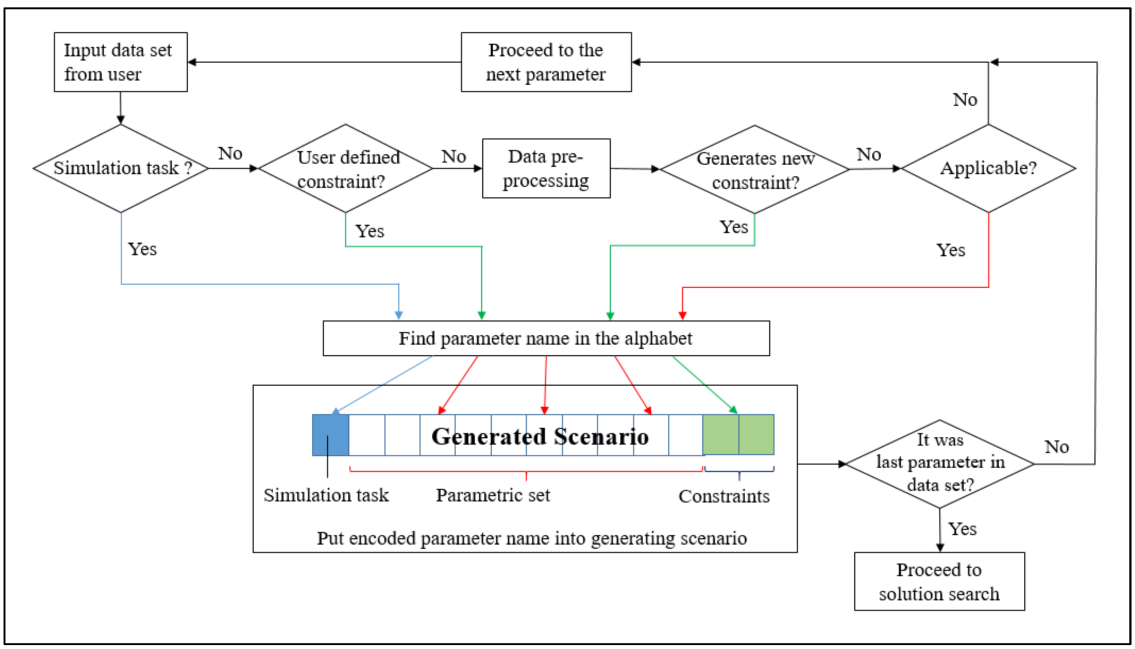

Figure 4.3 – The workflow diagram of the input data pre-processing and scenario generation: cycle for each parameter in the input data set

At the beginning, the user selects the simulation task to be solved and may additionally assign required accuracy and computational speed as “user defined constraints”. Then, the special algorithm takes corresponding alphabet elements and puts them into certain sequence, named system-generated scenario. During next steps, where their number is equal to the quantity of parameters in the input data set, the data pre-processing procedure evaluates each input parameter and qualifies its applicability. In the case of successful verification, the corresponding alphabet element is set into the "parametric set" section of the system-generated scenario. If the evaluating parameter value or quality generates additional constraint, then the certain alphabet element is added or corrected in the "constraints" section. If the input parameter fails verification, it is considered as inapplicable and rejected from consideration.

Overall, the system-generated scenario is the result of input data set analysis and processing. Each input parameter is evaluated with respect to its essence, applicability, and ability to generate additional constraints. Depending on the results, the parameter is either rejected or put into generating scenario with the specially assigned name. Once scenario is generated and not empty, the inference engine begins work.

4.5 Inference Engine Development

According to the existing methodology of the expert systems design and linguistic based pattern recognition, the inference engine can be realized as a procedure of match finding between the system generated scenario, based on the user's input data, and other scenarios from the library.

Once the scenario is generated based on the input data, it is addressed to the inference engine to evaluate the match with other scenarios within a library. The main goal here is to find answer to the questions:

- What simulation approach is recommended at the given conditions and will generate an appropriate expected result?
- Otherwise, what additional data is required for the certain task solving?

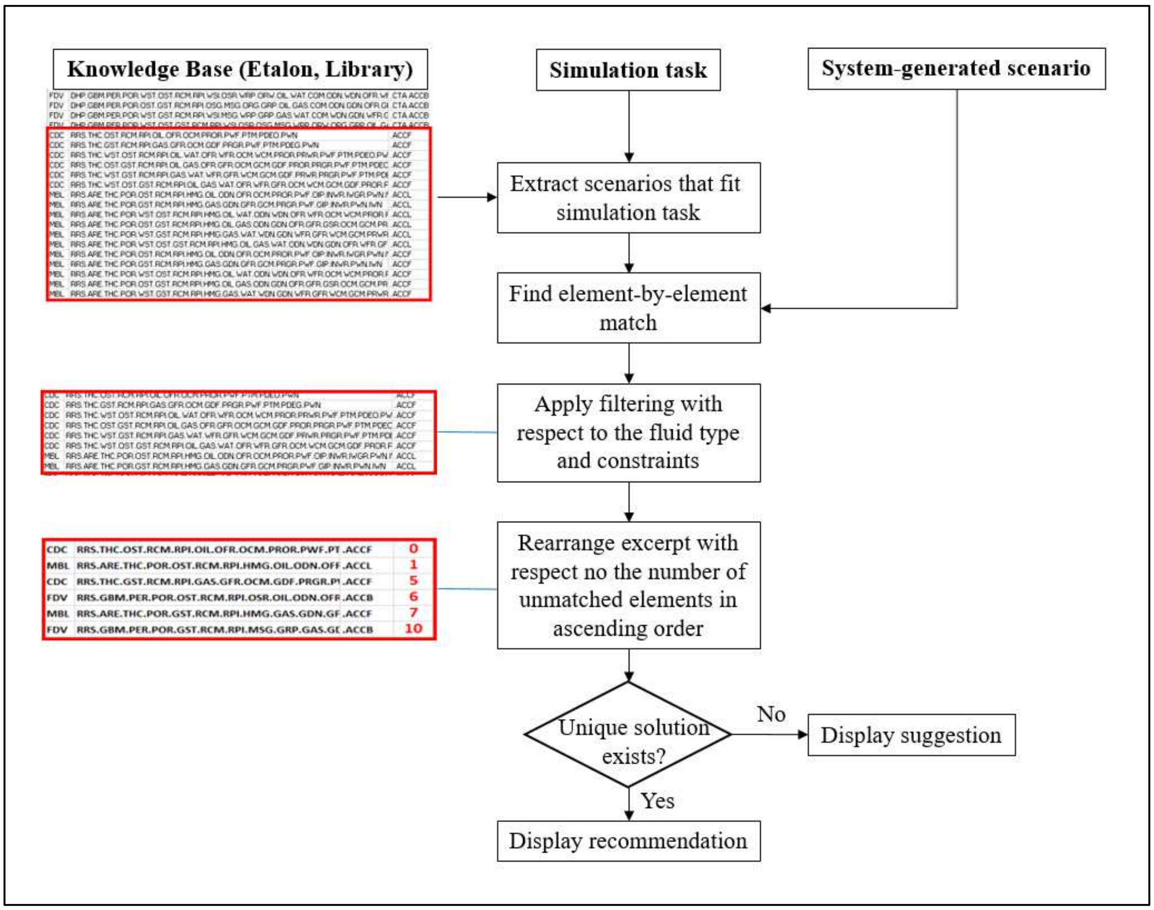

Figure 4.4 – The workflow diagram of the solution search process

As it is shown in **Figure 4.4**, the solution finding process is quite simple in case linguistic method is used. An existence of thoroughly prepared knowledge base and deliberately generated scenario reduces the inference engine to the set of ordinary procedures, such as “match”, “if-then”, and sorting.

Based on the assigned simulation goal, the inference engine extracts all related scenarios from the library. Then, the generated scenario is compared element-by-element with the other ones within excerpt. In general, the sought-for solution is based on: number of matched elements, critical (expert system determined) and user-defined constraints.

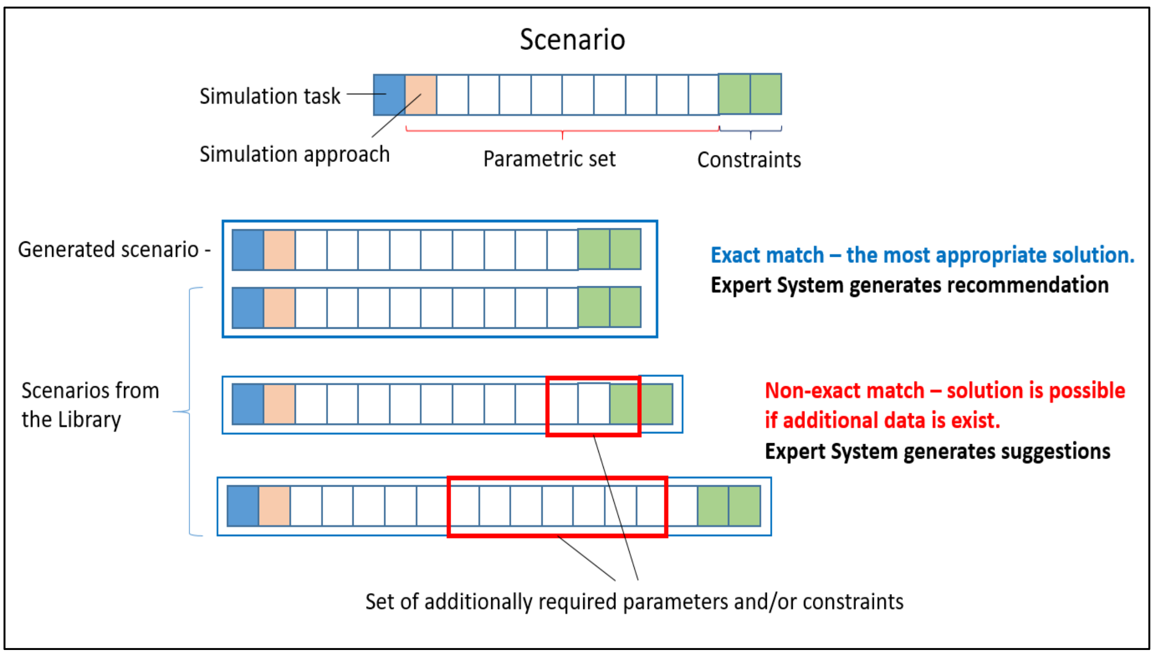

Figure 4.5 – Representation of match-finding process

Figure 4.5 represents the element-by-element match-finding process. Match is "exact" when number of unmatched elements is zero and all constraints are satisfied. The expert system provides recommendation to use this certain simulation approach as the most appropriate. When match is "non-exact", the expert system counts the number of unmatched elements (marked by red color in **Figure 4.4**) and provides suggestion about what should be additionally done to solve assigned problem. For instance, what input parameters are additionally required to make simulation approaches applicable.

After the match-finding process is over, the excerpt is filtered with respect to the fluid type and critical constraints. Namely, the quantity of possible outcomes is reduced in selection by eliminating the unreasonable ones. The excerpt is then rearranged with respect to the number of unmatched elements in ascending order.

Finally, the expert system displays results of the data analysis and processing. In the case, when only one simulation approach has zero number of unmatched elements, solution is unique and recommended to the user. Furthermore, if there is more than one simulation method that does not contain unmatched positions, then either all of them can be recommended for user to choose or some of them can be eliminated implementing additional user-defined constraints, such as combination of accuracy and computational speed. Otherwise, the list of feasible simulation methods will be displayed to the user with recommendation about what additionally required parameters should be obtained for each method.

4.6 Well Placement Justification

As an extension of the expert system functionality, the well placement justification technique, using decline curve analysis, was implemented in order to enable a preliminary assessment of the need for new wells placement within limitation of initially available data. Generally, this approach should give answers to the questions:

- Are additional wells required?
- If yes, then should they be producers and/or injectors?
- If injectors, where they should be placed?

This method provides an initial guess as to the necessity and number of production/injection wells and their placement zones. The exact locations should be further optimized using specific well placement techniques that are not in the scope of this work.

The main concept of the method is based on evaluation of internal reservoir energy that support hydrocarbon production. Direct energy quantification and its sufficiency assessment are very challenging tasks that by definition go out of the scope of traditional expert systems. In order to simplify this problem solution and incorporate it within our expert system, we have implemented the method of indirect reservoir energy appraisal using production data and specific criteria. Namely, using a production data analysis technique it is possible to assess whether the desired amount of hydrocarbon can be produced within assigned period of time or not. If sought-for answer is yes, then the conclusion is that reservoir energy is sufficient to support production. Otherwise, it is necessary to introduce new wells. Presented below is the methodology description in case of oil production.

The following input data is used for the method implementation:

- Oil production rate q_{oil} and water production rate q_{water} (if water is produced) per each production well;
- Reservoir geometry, initial oil saturation S_{oi} , rock properties, and oil formation volume factor B_o to calculate the value of total reservoir stock-tank oil in place $STOIP_t$. Otherwise, user is asked about the value of total $STOIP_t$;
- Areal well zonation, initial oil saturation S_{oi}, rock properties, and formation volume factor B_o to calculate the value of initial stock-tank oil in place $STOIP_i$ for each well. Otherwise, user is asked about the value of initial $STOIP_i$ for each well. In addition, well location within each zone is needed (see **Figure 4.6**);
- Value of the minimum economically acceptable production rate q_e (user-defined variable);
- The critical value of water cut WC (user-defined variable);
- The value of remaining recovery factor RF_L (user-defined variable) within time interval $[t_o, t_{wc}] < t_d$. Parameter t_d here is production time, upon which maximum amount of oil should be produced according to the field development plan (user-defined variable).

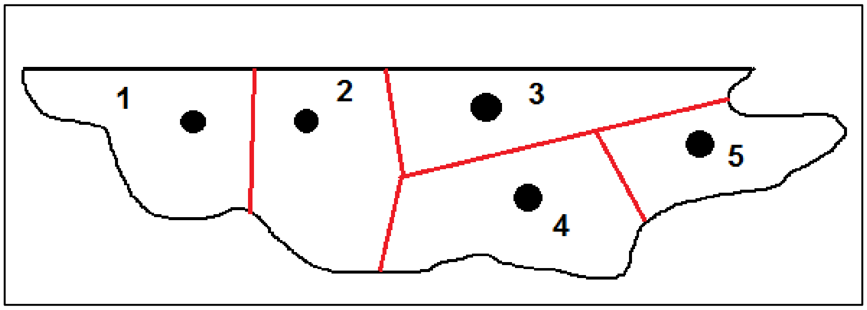

Figure 4.6 – Areal well zonation and wells locations (for case with producers only)

Figure 4.6 shows distribution of production wells, numbered and marked as black dots, within areal extent of reservoir. This is a plain view representation of the Brugge simulation model we used for the expert system validation, which described in details in the next section of the work. Red lines here display borders that separate areas related to the certain wells.

In accordance to the methodology, the following steps are involved into data calculation and analysis:

1. Water flood justification for the whole reservoir. On this step, the input data pre-processing procedure determines whether production decline regime is established or not. If it is, then using wells (preferably) or field production rates system calculates the amount of cumulatively produced oil upon the beginning of forecast Q_o and oil to be produced Q_f, as it shown in **Figure 4.7**:

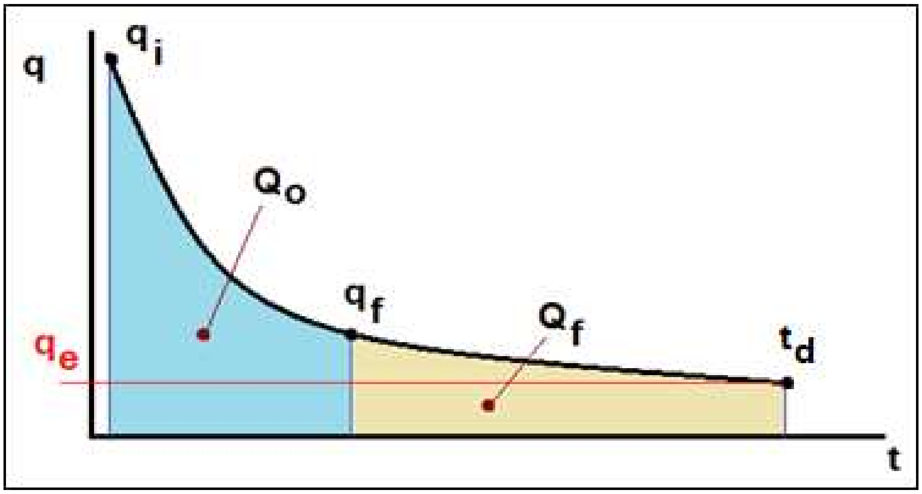

Figure 4.7 – Estimation of EUR using decline curve method (Q_o – cumulatively produced oil upon the beginning of forecast, Q_f – forecasted amount of oil to be produced)

Then, the values of oil EUR for each well are computed with respect to the q_e according to the following equation (for exponential decline):

$$EUR_i = (Q_{Oi} + Q_{Fi}) = Q_{Oi} + \frac{q_f - q_e}{a} \tag{4.1}$$

where: q_f – production rate at the beginning of forecast [STB/day]; a – exponential decline rate; i – number of wells.

In case of hyperbolic or harmonic production decline regime, the second term in the right-hand side of equation (4.1) is replaced with the appropriate one.

Further summation of EUR values and comparing them with the magnitude of $STOIP_t$ yields the estimation of whole reservoir recovery factor RF_F:

$$RF_F = \frac{\sum_i EUR_i}{STOIIP_t} \tag{4.2}$$

If RF_F < 30 % (average recovery factor with water flooding, user-defined value), then water flood is required. In other words, this is a confirmation of fact that internal energy is not sufficient to support the desired level of production.

2. In this step, it is necessary to determine whether each well is flowing optimally-normal or improvement is required. For those wells where production decline regime is established, system calculates the following parameters:

   - Time t_o, when oil production rate will reach q_e. Equation (4.3) is used for the exponential decline regime. In case of hyperbolic or harmonic decline, the denominator is replaced with the appropriate one:

   $$t_O = \frac{\ln\left(\frac{q_f}{q_e}\right)}{a} \tag{4.3}$$

   - Generates the water cut profile with respect to time. This calculation is implemented if water production data is available in accordance to the equation:

   $$WC(t) = \frac{q_{WATER}(t)}{q_{OIL}(t) + q_{WATER}(t)} \tag{4.4}$$

   Then, system finds the time t_{wc} at which water cut reaches the assigned critical value of WC.

   - Value of EUR_i using equation (4.1) and then value of the well recovery factor RF_{Wi} as:

$$RF_{Wi} = \frac{EUR_i}{STOIP_i} \tag{4.5}$$

Case A. If RF_{Wi} > 30% (user defined value), then drainage improvement is most likely not required. Namely, this situation means that the particular well production is sufficiently supported by the reservoir energy, and the expected oil recovery level exceeds assigned threshold value when the water flood should be implemented. Recommendation: do nothing or an additional production well may be introduced into the related reservoir zone by user choice as a result of further investigation.

If RF_{Wi} < 30% (user defined value), then system determines the following:

$t = \min\{t_d, t_o, t_{wc}\}$ - evaluation criteria.

Case B. If $t = t_d$, as presented in **Figure 4.8**, oil reserves are too large for this one well to be produced. In other words, the desired amount of oil will not be produced by the time t_d, since oil production rate is still high and does not reach minimum level q_e prior to the end of production time t_d. Very significant amount of oil may remain unproduced. Recommendation: an additional producer is required in this particular area.

Case C. If $t = t_{wc}$, as shown in **Figure 4.9**, the well reaches the critical value of water cut WC earlier than assigned end of production time t_d. In such particular situation, the oil production is no longer economically reasonable. Recommendation: do nothing or consider to transform this well to an injector.

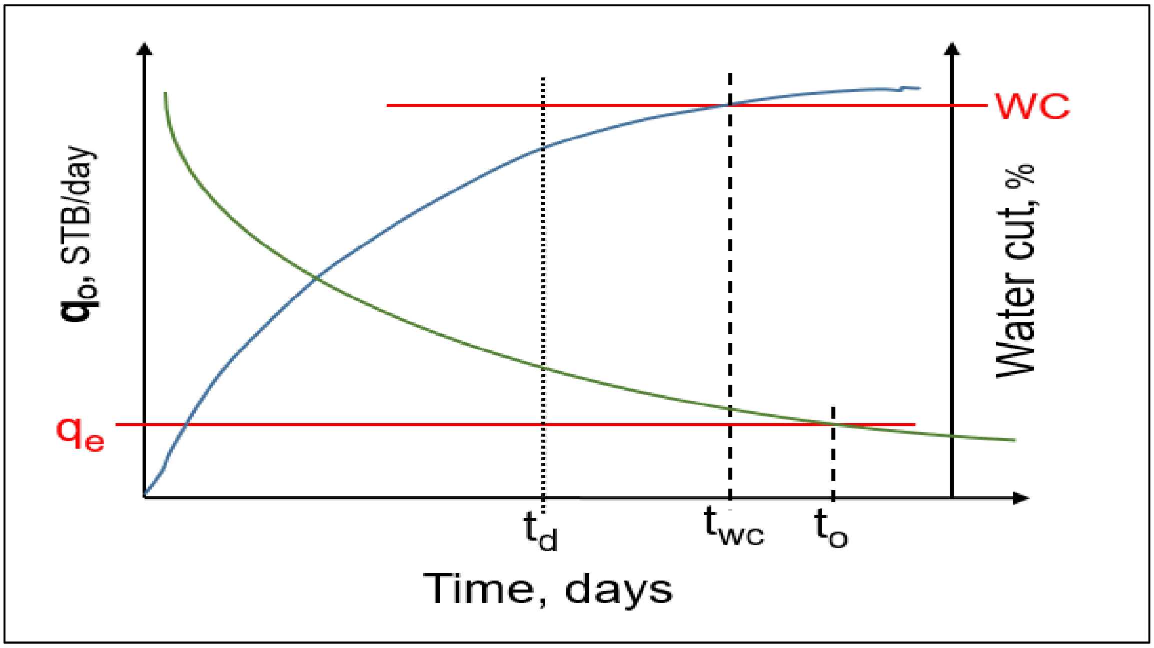

Figure 4.8 – Cross-plot of oil production rate and water cut versus time for Case B

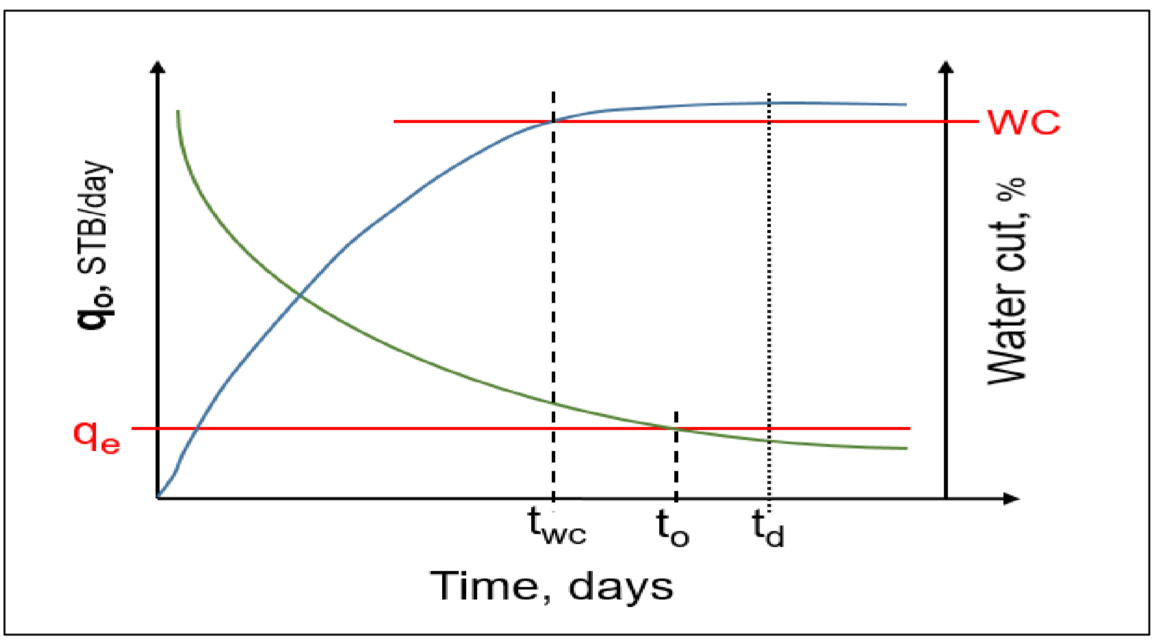

Figure 4.9 – Cross-plot of oil production rate and water cut versus time for Case C

If $t = t_o$, then there are two additional options are possible:

- Case D. If $t_d < t_{wc}$, as depicted in **Figure 4.10**, production is poorly supported by the reservoir energy. Namely, oil production decline reaches minimum level q_e prior to the end of production time t_d, and certain amount of economically profitable oil may remain unproduced. Moreover, the water production should not affect the oil production because the achievement of critical water cut level WC, at the given conditions, is supposed to be the latest in time. Recommendation: an additional injector is required for this particular area since the reservoir energy is not sufficient.

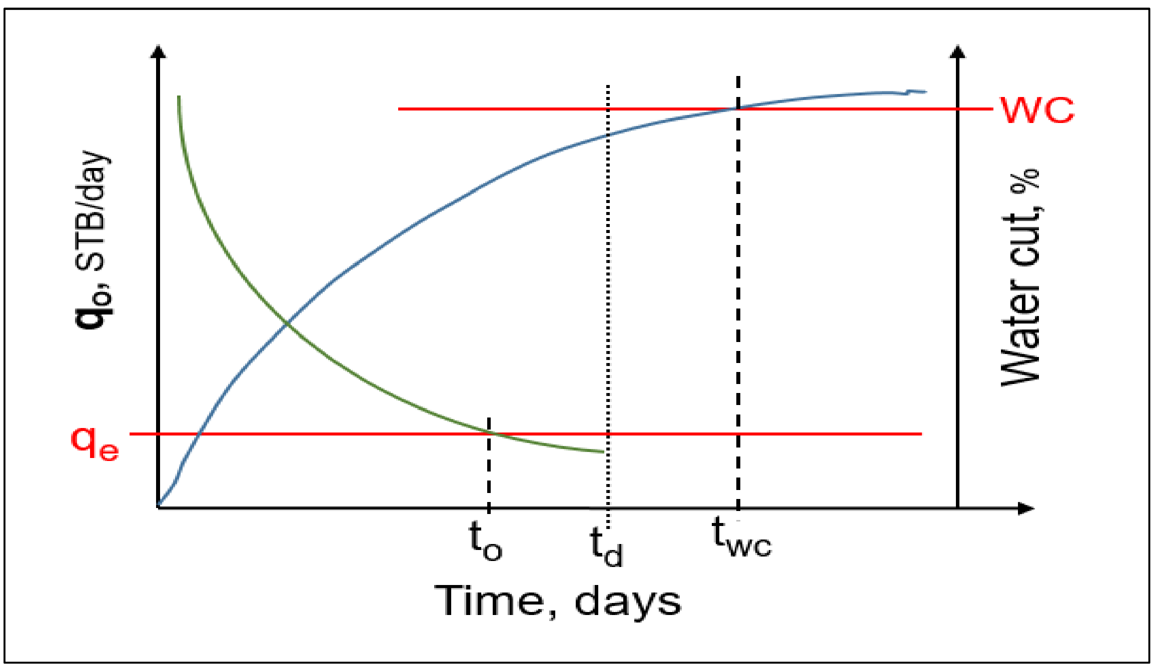

Figure 4.10 – Cross-plot of oil production rate and water cut versus time for Case D

- Case E. If $t_d > t_{wc}$, as presented in **Figure 4.11**, that is doubtful case. At first look, lack of reservoir energy to support production is obvious, and analogically to Case D an additional injection well is required. At the same time, there is a high risk exist that the water cut may reach critical value of WC very soon. In such conditions, an implementation of water flood might be not beneficial. Hence, an additional calculation of leftover recovery factor RF_L is needed using equation (4.6) for exponential decline. In case of hyperbolic or harmonic decline, the denominator is replaced with the appropriate one:

$$RF_L = \frac{(q_{t_o} - q_{t_{WC}})}{a} / STOIP_i \qquad (4.6)$$

where: q_{to} – value of production oil rate at time t_o [STB/day]; q_{twc} – value of production oil rate at time t_{wc} [STB/day].

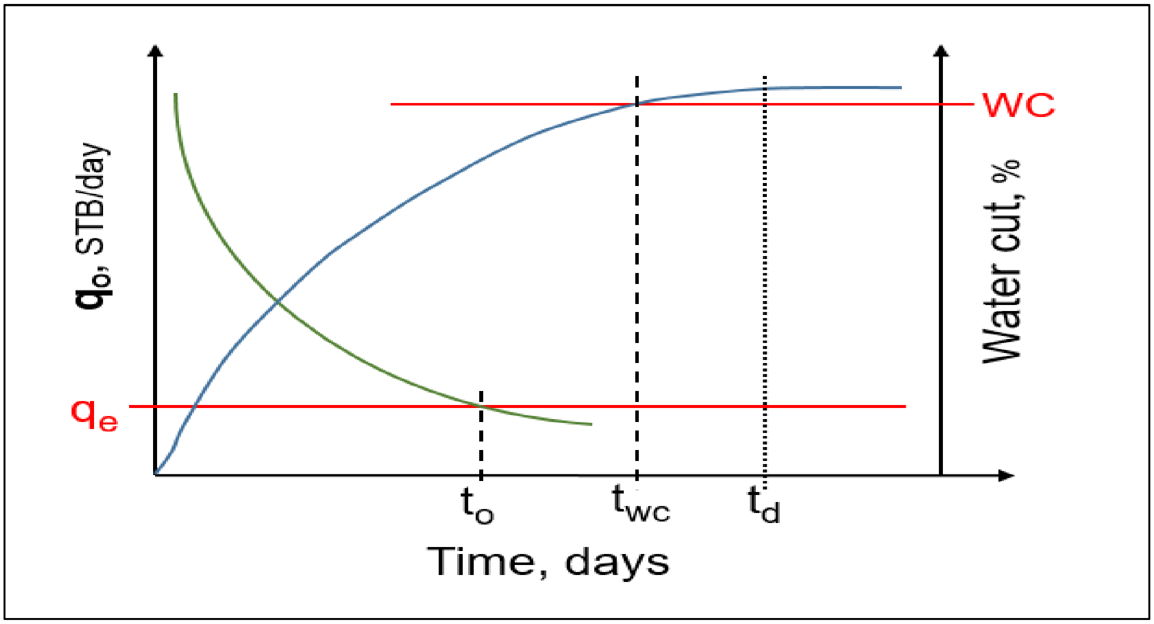

Figure 4.11 – Cross-plot of oil production rate and water cut versus time for Case E

If calculated value of RF_L is larger than it specified by user, then an additional injector is required since producible amount of oil is profitable;

If RF_L is less than it specified by user - highly detailed economic analysis is required for decision making (or consider this well as injector).

3. Depending on the obtained results, all well areal zones can be sorted as:

   - Injector requiring;
   - Producer requiring.

As mentioned above, the exact positions of new wells are the subject for discussion and further study that is out of our work scope. But general criterion, as initial guess for well position optimization, can be described as the following: new injection well should be located between producers that require injection and/or on the flank of reservoir.

4.7 Expert System Functionality

Conceptually, the designed linguistic method based pattern recognition expert system consist of two major blocks with following functionality, as shown in **Figure 4.12**:

1. Data Processing Block performs the input data analysis, data processing, and simulation approach selection. The input data set here is a collection of the reservoir parameters, including rock and fluid properties, production data, and other information available to a user and is to be used for solving a particular problem. Passing through the Data Pre-processing and Scenario Generation modules, the initial input data set is analyzed and certain scenario is generated. Then, Inference Engine compares the obtained

result with other ones from the Knowledge Base. If a solution is found, corresponding decision support, such as recommendation and suggestion, is proposed to a user as the expert system output. Otherwise, the system concludes that there is an undescribed case and the Knowledge Base should be revised and supplemented (extended) using scenario generation tool described in the section 4.3 of this work.

2. Analytical Block provides possibility of the simulated results quality assessment. Depending on the simulation goal, simulated data might be evaluated quantitatively or/and qualitatively. For example, quantitative evaluation could be made by comparing simulated data with some etalon solution available for comparing: physically measured as pressure or flow rate, obtained by finite difference (volume) simulator or from another trustful source. As a result, the magnitude of data deviation, its accuracy, and model predictability are assessed and could be used for error analysis. The qualitative evaluation is proposed to be provided by user and contain such evaluation criteria as: uncertainty, bias, CPU timing, computational costs, overall model applicability, quality of initial data source etc. If simulated results do satisfy entire quality requirements, then it is considered that no correction is needed. Otherwise, an additional error analysis is provided and recommendations about how to improve the simulated results are designed in dependence of the estimated source of error: incorrect simulation approach selection, uncalibrated simulation model, and/or doubtful input data.

Since realization of the Analytical Block implies the design of self-teaching option, this task is not included into the current work and taken out to the further system development.

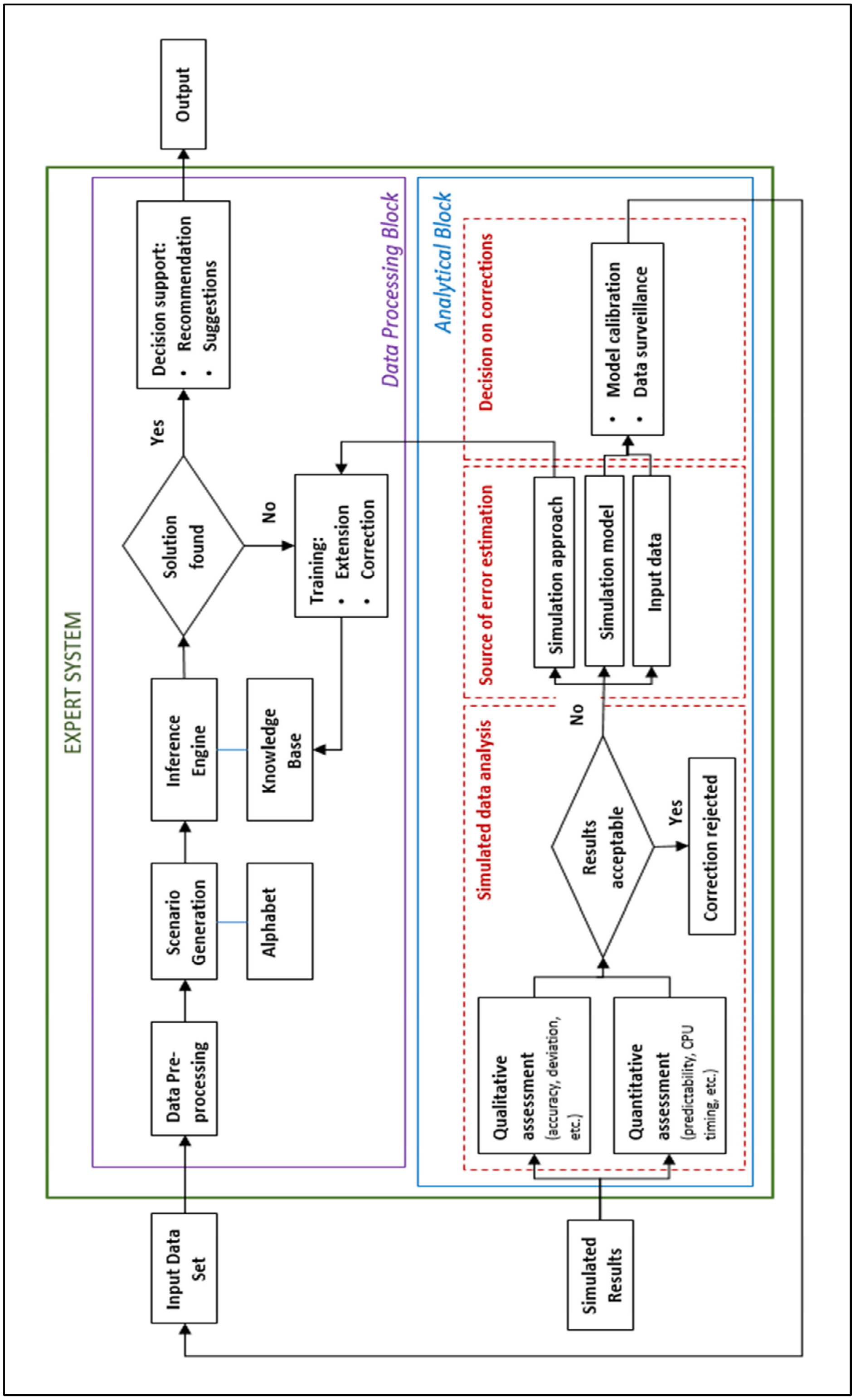

Figure 4.12 – The expert system functionality, conceptual diagram

4.8 Conclusions

The use of linguistic method of the Pattern Recognition Theory allowed creation of the alphabet as a set of 126 selected key parameters combined in 16 major groups. Initially created in MS Excel, the alphabet was transferred to MatLab database. It is used by the Data Pre-processing, Inference Engine, and Decision Support modules of the expert system.

The set of 522 scenarios, describing requirements for solving of 16 reservoir evaluation problems with respect to 5 simulation methods, was created and represents the Knowledge Base of the expert system. It was constructed in MS Excel using the alphabet elements and then transferred to MatLab database, where it is accessed by the Inference Engine for the suitable simulation method selection.

Any corrections of the alphabet or scenarios does not change the methodology.

The Data Pre-processing module was designed for execution of the input data quality control, constraints handling, and construction of the system-generated scenario.

The Decision Support module was introduced in the expert system to realize explanatory function. Depending on the Inference Engine output, it generates results of expertise as recommendation and suggestions.

5. EXPERT SYSTEM VALIDATION AND FIELD APPLICATION

This section presents the expert system validation and field application workflow. It was performed in order to examine correctness of ideas and methodology in the basis of the expert system and its capability to provide reservoir data expertise in simulation approach selection. Two data sets, synthetic Brugge model and offshore petroleum reservoir model, were processed and evaluated by the expert system for several test problems. Obtained results will be discussed in this section.

5.1 Brugge Simulation Model

For the purposes of the expert system validation, we have used the complete synthetic Brugge simulation model by TNO company (Peters et al., 2009). This model consists of 60 000 grid block with detailed set of rock and fluid properties. Representing initial reservoir development with five equally spaced producers (**Figure 5.1**), Brugge model was run to obtain the test production dataset.

Initial 4 years of simulated data was used to introduce actual field production and majority of the expert system input to solve multiple tasks such as estimation of:

- recoverable field resources (EUR);
- stock-tank oil initially in place (STOIIP);
- drainage area;
- recovery factor;

- investigation of water-flood feasibility.

Finite difference model simulation results are used as reference for another methods and to validate appropriateness of solution suggested by the expert system.

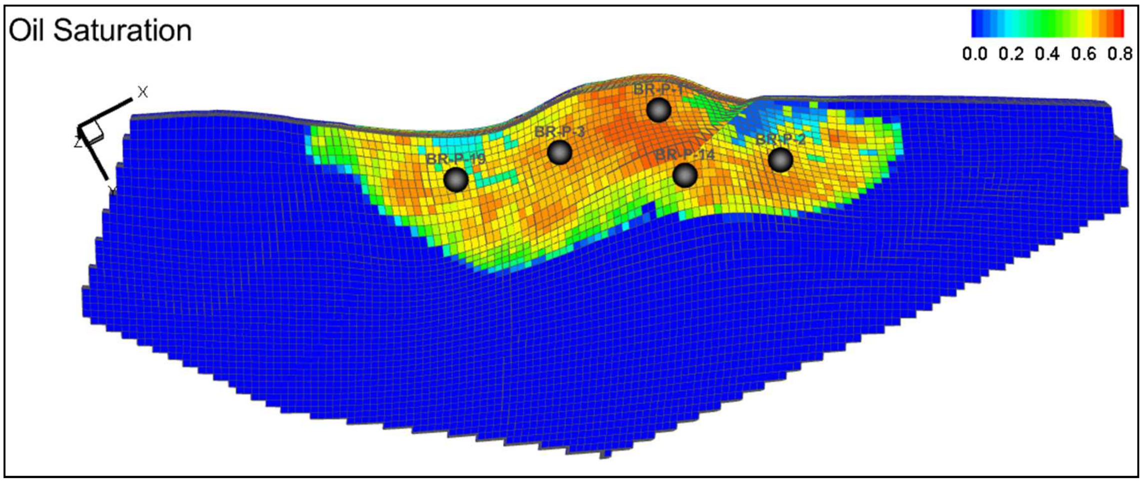

Figure 5.1 –Brugge simulation model with five equally spaced producers. Initial oil saturation is shown in color scale (adopted form Peters et al., 2009)

For each task, the control results were obtained using decline curve analysis technique, material balance simulation (MBAL software by Petroleum Experts), and finite difference simulation (Eclipse 100 software by Schlumberger). List of the used input parameters and estimated results for each method are shown in **Figures 5.2 – 5.4**. These control results, derived via mentioned methods, were assessed in accuracy and computational speed as evaluation criteria. Namely, we discovered capability of each simulation approach to solve particular tasks with regard to the output exactness and time expenses. This information is very important, since it is used in the expert system predictability validation.

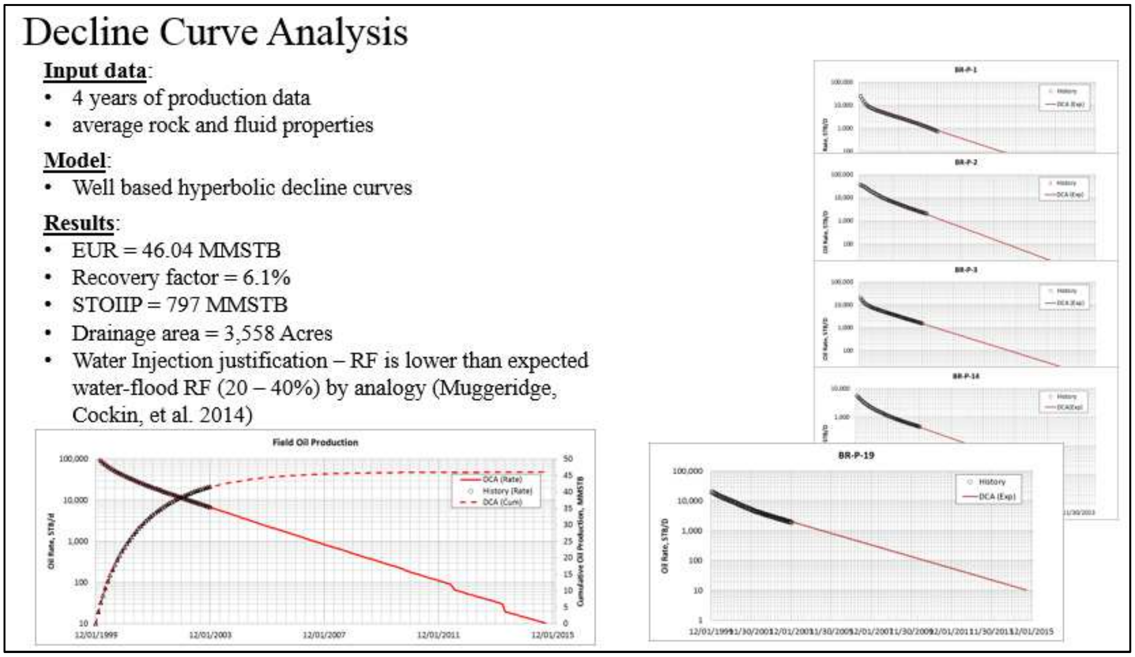

Figure 5.2 – Input parameters and calculated data using decline curve analysis (Brugge simulation model)

Figure 5.2 represents the use of the decline curve analysis for the test problems solving. All required results were calculated with respect to each of five producing wells and then combined to represent full field data. Since this method does not allow direct assessment of water flood justification, we made required conclusion by comparison of calculated recovery factor value with its average values in case of primary oil recovery without water-flood (Muggeridge, Cockin et al., 2014).

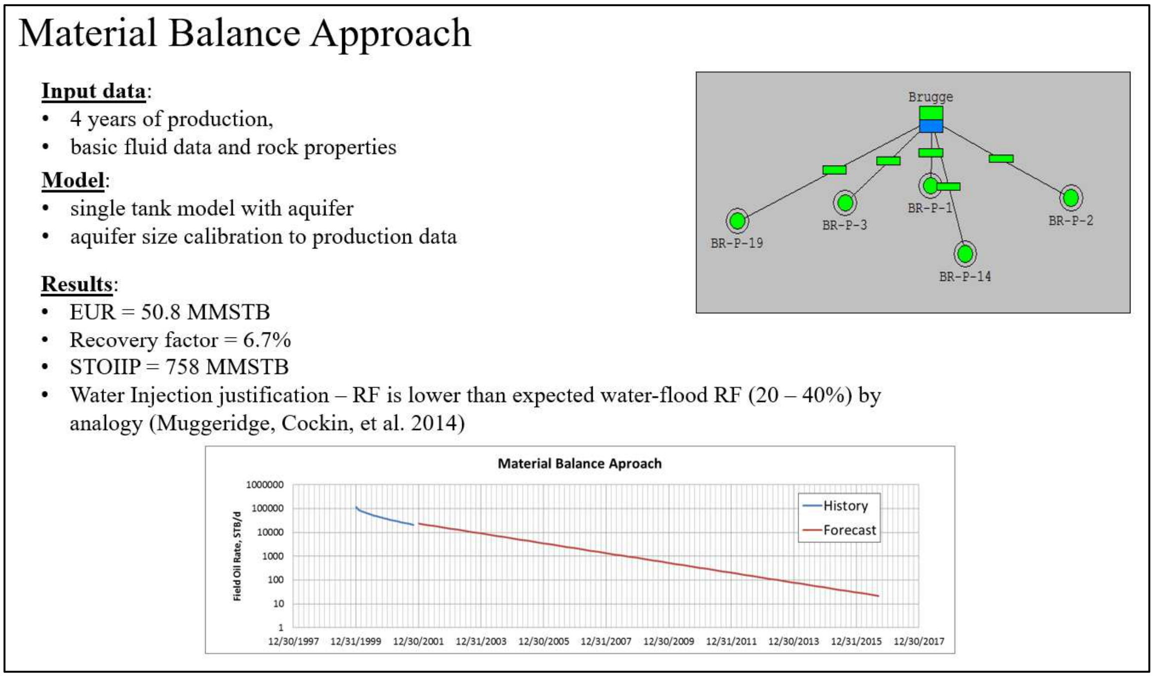

Figure 5.3 – Input parameters and calculated data using material balance approach (MBAL software, Brugge simulation model)

Application of the material balance method in assigned problems solving is shown in **Figure 5.3**. The single tank model with an aquifer and five production wells was created and used for simulation purposes. Additionally, the aquifer size was calibrated to match previously generated production data. Analogically to the described above decline curve analysis results, the water injection justification has been made indirectly. We have made the same logical conclusion because the calculated value of recovery factor is lower than it could be when the water-flood is implemented.

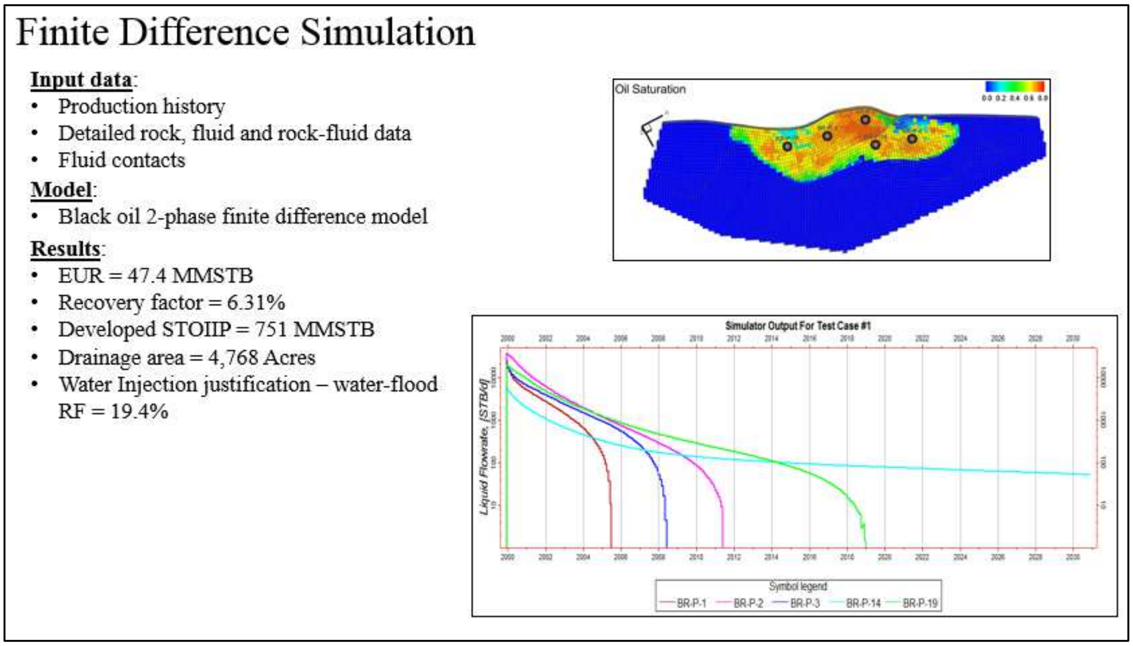

Figure 5.4 – Input parameters and calculated data using finite difference simulation (Eclipse 100 software, Brugge simulation model)

Figure 5.4 depicts details of the finite difference simulation method implementation. As a result of simulation, the required solution for each test problem was calculated and further used as reference data. In order to investigate the water injection justification, we supplemented initial model with an additional water injection well and executed one more simulation run. Consequently, we have received extra confirmation that introduction of water injection can improve the oil production.

Table 5.1: Summary of results calculated by different simulation methods using Brugge simulation model

Engineering Task	Units	Decline Curve Analysis	Material Balance	Finite Difference Simulation (REFERENCE)
Estimation of recoverable field resources (EUR)	MMSTB	46.04	50.8	47.4
Estimation of recovery factor	%	6.1	6.7	6.31
Estimation of STOIIP	MMSTB	797	758	751
Estimation of drainage area	Acres	3,558	N/A	4,768
Investigation of water flood feasibility	---	Indirect, by analogy (water flood is required)	Indirect, by analogy (water flood is required)	19.4 % recovery factor (water flood is required)

Summary of the calculated results for each of the selected test problems is presented on **Table 5.1**. Comparison of obtained numbers and conclusions additionally supports our initial statement that the most appropriate selection of simulation approach to be used for particular problem solving is non-trivial. As it shown on the **Table 5.1** and discussed in the section 2, different methods can generate variety of output results for each task although the same input data is used. Emergence of such situations is defined by not only assumptions and limitations in the basis of each method, but also by the quality and sufficiency of input data. Thus, certain cautions should be taken into account when accuracy of simulated results is critical. Suitability of Decline Curve Analysis and

Material Balance method in terms of accuracy of calculated results is highlighted in colors, considering Brugge model correction explained below. Comparing to the reference data, green color marks the recommended method to be used, yellow marked approach might be used at the discretion of a user, and red color suggests to avoid this method.

Time expenses for each test problem solution are almost the same with respect to the particular simulation approach, since previously generated reservoir production data and full simulation model with relatively small number of grid blocks were used.

In order to improve quality of the expert system validation process, we have complicated the input data set. Namely, end point of the residual oil saturation and several nearby ones were eliminated from the initial Brugge simulation model. Therefore, we have artificially reduced applicability of the finite difference simulation in the test problems solving, while decline curve analysis and material balance methods still can be used. In other words, we designed experiment that allows to conduct study and evaluate quality of the expert system's modules performance. Namely, we tested an ability of the procedures and algorithms in the basis of the data pre-processing, scenario generation, knowledge base, and inference engine to deal with the quality and sufficiency of input data and generate correct conclusions.

The corrected, as described above, Brugge simulation model and generated production data was processed by the expert system. For the testing purposes, we transferred simulation model and production data variables and their values to MatLab databases as the input data set for the expert system. These input files contain detailed information about:

- reservoir grid geometry;
- detailed rock properties – spatial porosity and permeabilities distribution, compressibility;
- detailed reservoir fluids properties – densities, compressibilities, formation volume factors, viscosities, relative permeabilities with respect to saturations (except deleted oil residual saturation), initial saturations, gas oil ratio and deviation factor;
- reservoir fluid types – oil and water;
- fluid contacts – water-oil contact depth;
- pressures – initial, capillary with respect to water saturation;
- well data – number and locations of production wells;
- production data – fluids production rates, wellbore flowing pressures, production time.

To distinguish variables, the same Eclipse keywords were used as they appear in ASCII files of Brugge simulation model.

For each task, the system has generated expert recommendations. Examples of solution search workflow and outputs are shown in **Figures 5.5 – 5.9** and on **Table 5.2**. Validity of these recommendation was tested through comparison with previously calculated ones (**Table 5.1**) by decline curve analysis, material balance simulation, and finite difference simulation.

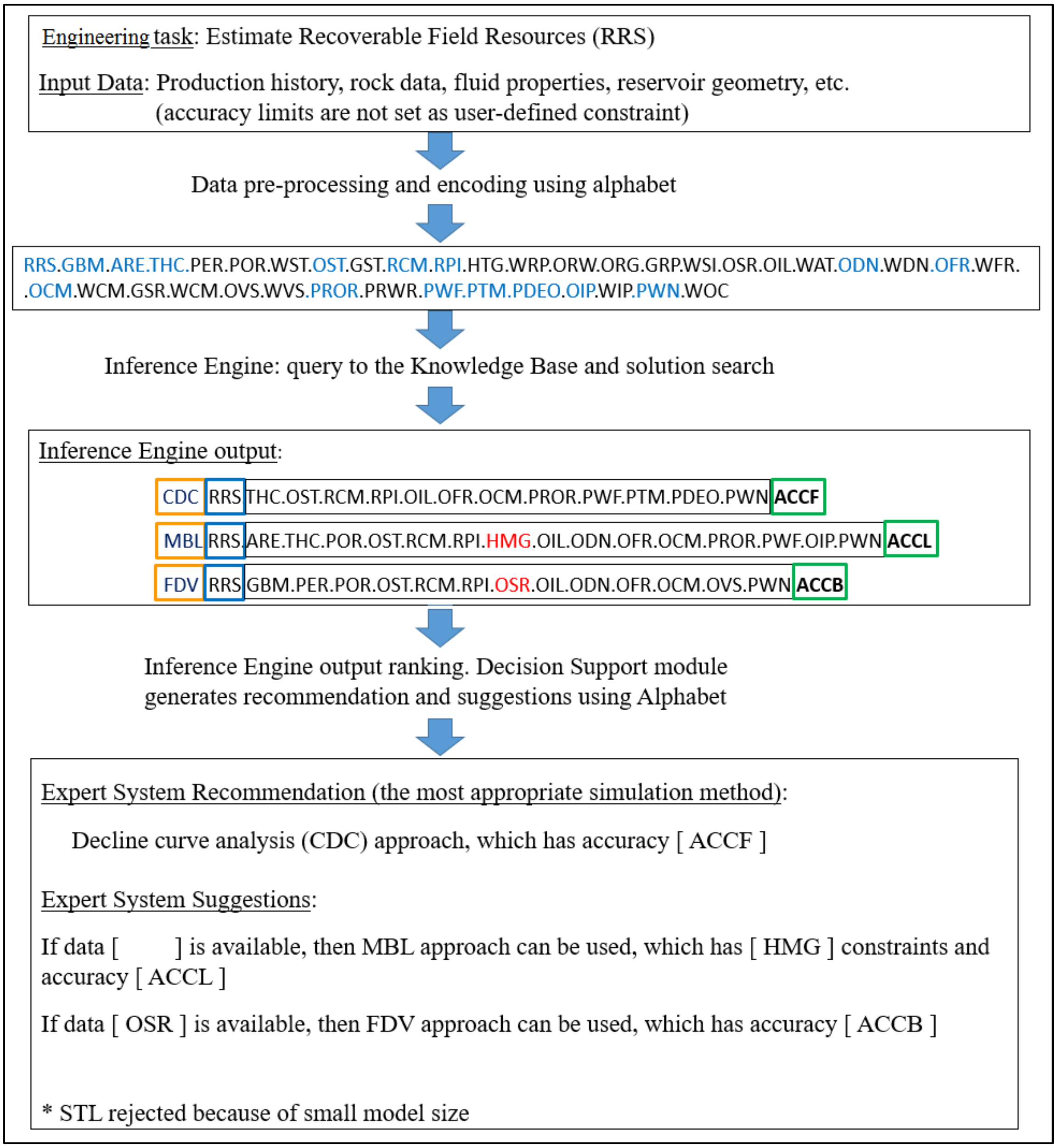

Figure 5.5 – Expert system workflow for estimation of recoverable field resources (corrected Brugge simulation model)

Figure 5.5 represents an example of the expert system reasoning and solution search structure, where from top to bottom:

- The upper box shows main input for the system as selection of certain simulation goal and description of the input data set;
- The second from above box displays results of the input data pre-processing as the system-generated scenario;
- The third box represents selected simulation approaches ranged in applicability as the “Inference Engine output” results. Yellow boxes highlight simulation methods, blue ones mark the name of solving problem, and green rectangles shows predicted accuracy of calculated results;
- The lower box demonstrates recommendation and suggestion on applicability of the selected simulation approaches with some explanations and predictions.

At the beginning, user defines the particular simulation problem to be solved and desired level of accuracy and computational time (user-defined constraints). In this example, the simulation task is encoded as “RRS” and put into the beginning of parametric set of the line that represents the system-generated scenario. Here and further, the meaning of encoded parameters can be found on the **Table A-1** (APPENDIX A).

Then, the input data set is processed by the Data Pre-processing module in accordance to the methodology described in the section 4.4 of this work. Once system-generating scenario created and is not empty, the inference engine evaluates its match with other scenarios, which related to the same simulation goal, within the knowledge base. In **Figure 5.5**, the matched elements in system-generated scenario are marked in blue color.

These elements are common for all three scenarios that the Inference Engine selected, ranged, and put in rows as the "Inference Engine output". If element (marked in red color), required for implementation in the particular scenario, does not exist in the system-generated scenario or is considered as system-determined constraint, then the system generates explanation and suggestion for a user. For instance, in this particular case the expert system selected three applicable scenarios, where:

- The first one, related to decline curve analysis (CDC), does not contain unmatched elements in the parametric set and applicability of this approach has been verified because required production data is characterized by established decline regime (PDEO). Predicted accuracy of calculation is supposed to be fair (ACCF).
- Next one, associated with material balance (MBL), contains only one red-colored unmatched element (HMG) that is homogeneity. Because calculated value of Dykstra-Parson's coefficient for the given model is equal to 0.35, the Data Pre-processing module considered this situation as system-determined constraint for the material balance method applicability. Hence, in case of that approach implementation the accuracy of the calculated result is expected to be low (ACCL).
- The last one, linked with finite difference simulation (FDV), also contains only one unmatched element (OSR) which is residual oil saturation to water. Since the initial Brugge simulation model was corrected as it mentioned above, the Data Pre-processing module has not detected this parameter within the input

data set. Thus, the expert system considered that case as system-determined constraint for application of the finite difference simulation, since simulation model is incomplete and additional time is required to obtain required data. Expected accuracy of simulated results is predicted as the best (ACCB).

Finally, the expert system has generated output for a user as “Expert System recommendation” and “Expert System suggestions”. The system recommends to use the decline curve analysis as the most appropriate method for the recoverable reserves calculation because all requirements and constraints are satisfied, predicted accuracy is fair, and computation is fast. It additionally suggests that the material balance method can also be used at the discretion of a user. The input data set is sufficient for that, but estimated level of heterogeneity can negatively affect the accuracy of solution. Moreover, the expert system has verified applicability of the finite difference simulation and what information should be additionally obtained for it execution. Besides, the streamline simulation method was rejected by the expert system from consideration due to small simulation model size, what does not provide any advantage comparing to the finite difference simulation.

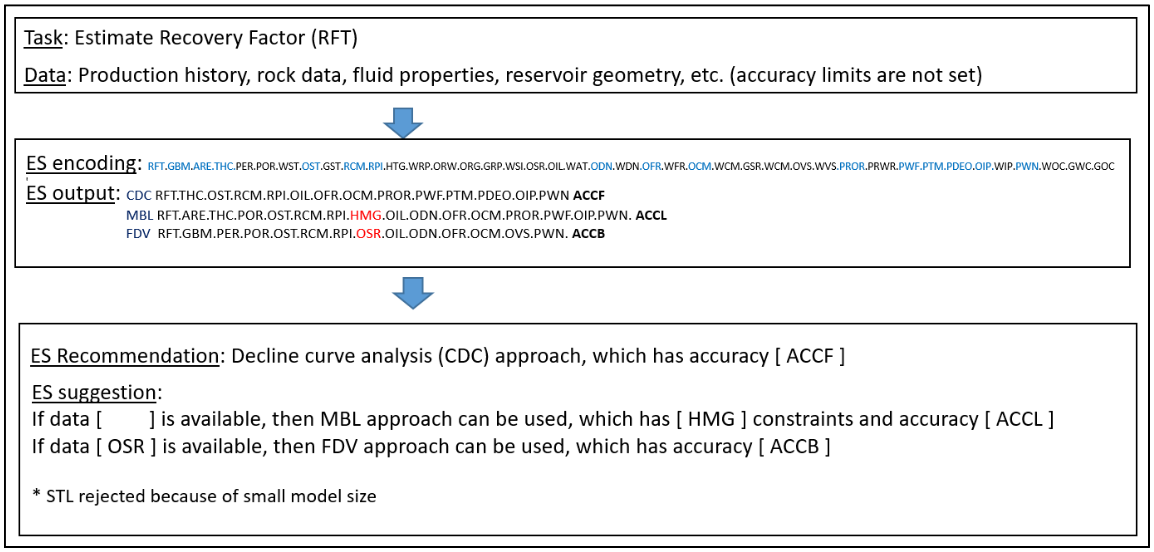

Figure 5.6 – Expert system workflow for estimation of recovery factor (corrected Brugge simulation model)

Analogically to the previous one, **Figure 5.6** represents an example of the expert system reasoning and solution search structure. As a result of the input data set pre-processing, the system has generated scenario and found matching scenarios in the knowledge base, which correspond to the decline curve analysis (CDC), material balance method (MBL), and finite difference simulation (FDV). Predicted accuracy of recovery factor estimation for each of these approaches is supposed to be fair (ACCF), low (ACCL), and as the best (ACCB), respectively. Finally, the expert system recommends to use the decline curve analysis as the most suitable and suggests implementation of other two approaches considering particular constraints, such as heterogeneity and lack of required data.

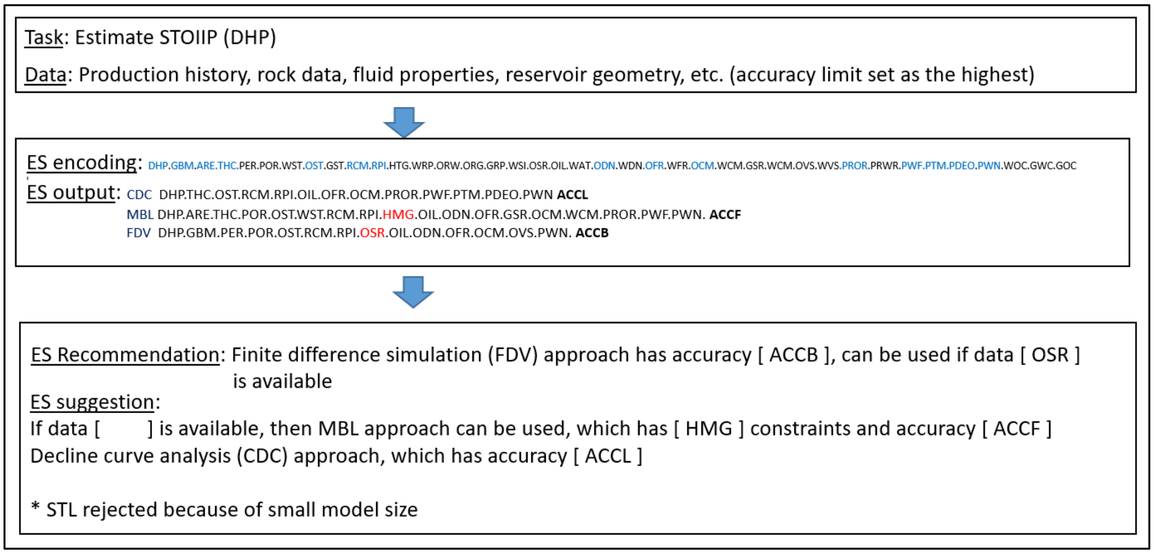

Figure 5.7 – Expert system workflow for estimation of stock-tank oil initially in place (corrected Brugge simulation model)

For the task of STOIIP estimation, required accuracy of simulated result was assigned as the highest; it is user-defined constraint. In such condition, the entire workflow of the expert system reasoning and solution search, shown in **Figure 5.7**, differs from previous cases. Although the system found scenario with zero unmatched elements, which corresponds to the decline curve analysis, this scenario is not recommended as the optimal. Because the assigned accuracy requirement is considered by the expert system as a critical constraint, the "ES output" excerpt was additionally rearranged in descending order with respect to predicted accuracy. Therefore, the expert system generates recommendation to obtain missing data and use the finite difference simulation. As alternative, the system suggests two more approaches with lower predicted accuracy of simulated result at the discretion of a user.

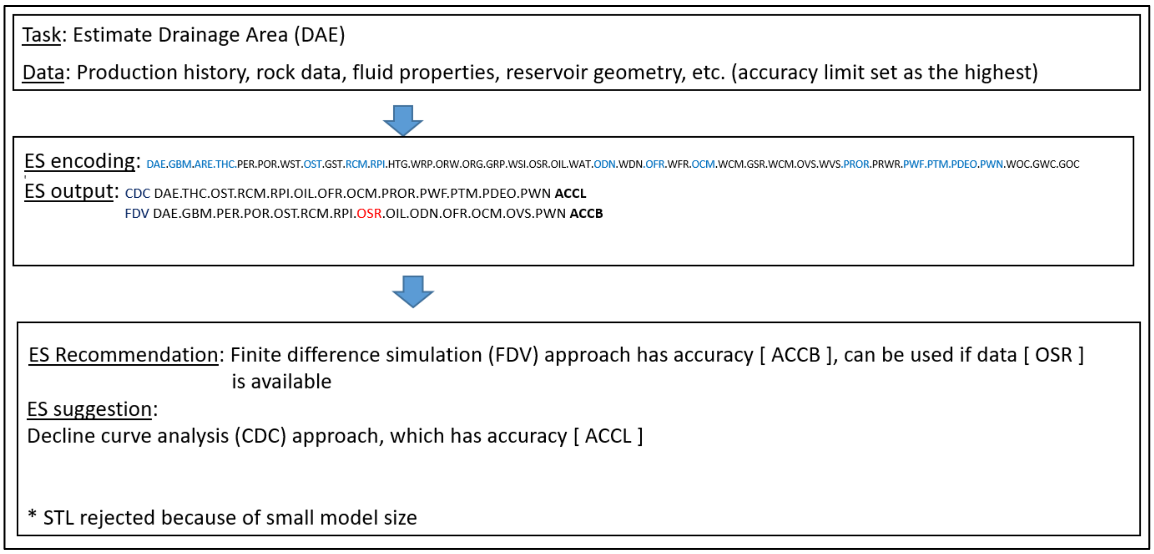

Figure 5.8 – Expert system workflow for estimation of drainage area (corrected Brugge simulation model)

Figure 5.8 represents an example of the expert system workflow for estimation of drainage area. As a result of the input data processing and solution search, there only two approaches were found by the system to be applicable. Material balance approach has been rejected from consideration as unsuitable due to limitations in the basis of this method. Because the accuracy limit has been set, the expert system recommends to use the finite difference simulation as optimal although lack of required data was detected. In addition, the system suggests implementation of the decline curve analysis, taking into account that predicted accuracy is low.

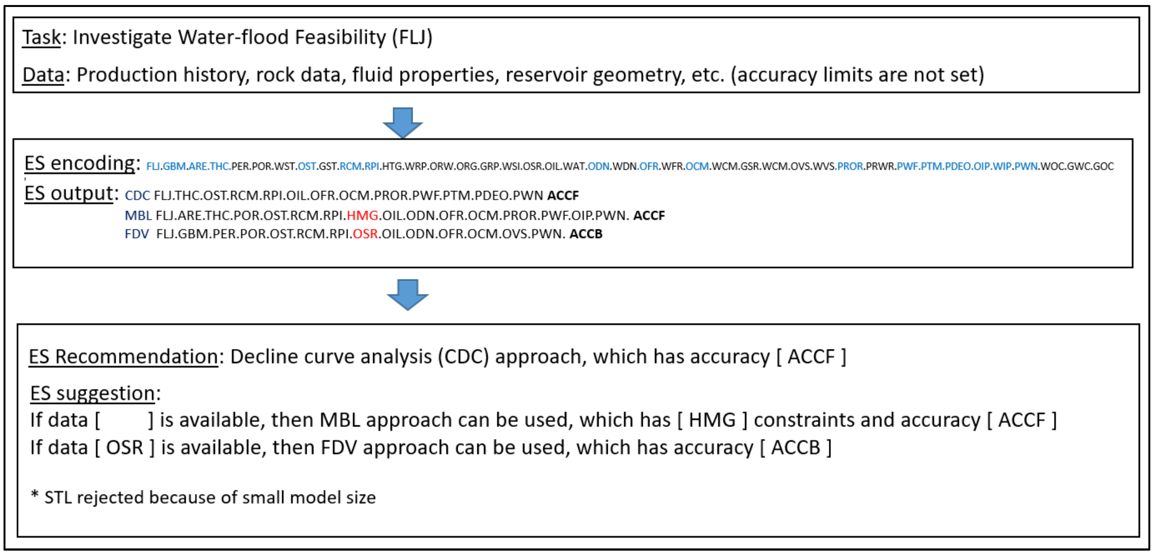

Figure 5.9 – Expert system workflow for investigation of water-flood feasibility (corrected Brugge simulation model)

Investigation of water-flood feasibility is closely related to the recovery factor estimation and requires the same data set. The expert system workflow and output are shown in **Figure 5.9**. As inference, the expert system recommendation is to use the decline curve analysis as appropriate. In addition, the system suggestion is that other two approaches, material balance and finite difference simulation, could be used considering detected constraints.

Executing methodology described in the section 4.6, we have tested an ability of the expert system to assess necessity of new wells and their types (well placement justification problem). The described method was implemented in the expert system as a separate procedure. For the testing purposes, we made the following assumptions for the values of user-defined variables:

- critical value of water cut WC = 95%;
- average recovery factor with water flooding = 30 %;
- production time, upon which maximum amount of oil should be produced t_d = 30 years;
- value of the minimum economically acceptable production rate q_e = 50 STB/day;
- total reservoir stock-tank oil in place $STOIP_t$ = 751 MMSTB;
- initial stock-tank oil in place for each well $STOIP_i$ = 150.2 MMSTB.

Initial 4 years of simulated production data were used to evaluate the necessity of production and injection wells. The wells spatial distribution within reservoir is the same as it schematically shown in **Figure 4.6**.

As a result of the production data analysis, the data pre-processing module detected that decline regime is established for each well; hence, the decline curve analysis can be used in recovery factor calculation. Further, the system calculated amount of oil that can be produced (EUR) by the time t_o when the value of production rate reaches q_e. Then, the corresponding values of t_o for each well have been also estimated. Summary of the test results is presented on **Table 5.2**.

Table 5.2: Evaluation of additionally required wells with respect to existing producers

Parameter	For field	For well 1	For well 2	For well 3	For well 4	For well 5
STOIP, MMSTB	751.0	150.2	150.2	150.2	150.2	150.2
EUR, MMSTB	45.85	11.53	8.88	6.79	2.65	16.00
RF_F , %	7.52<30 %	---	---	---	---	---
RF_W , %	---	7.68<30 %	5.91<30 %	4.52<30 %	1.76<30 %	10.65<30 %
t_o , years	---	12.5	12.0	8.5	9.0	10.0
$t = \min\{t_d, t_o, t_{wc}\}$	---	t_o	t_o	t_o	t_o	t_o
Expert System Conclusion	Reservoir energy is not enough to support production	Reservoir energy is not enough to support production	Reservoir energy is not enough to support production	Reservoir energy is not enough to support production	Reservoir energy is not enough to support production	Reservoir energy is not enough to support production
Expert System Recommendation	New wells are required	Injector is required	Injector is required	Injector is required	Injector is required	Injector is required

The expert system calculated values of whole reservoir recovery factor RF_F and recovery factor for each well RF_W, shown on **Table 5.2**, are much lower than they could be in case of water flooding. In addition, estimated values of time t_o, when oil production rate reaches q_e, also smaller then assigned production time t_d. Considering these two facts, the expert system makes conclusion that reservoir internal energy is not enough to support the desired level of production. Namely, all the wells suffer of low reservoir energy, and significant amount of oil remains unproduced until the established time t_d. Hence, the expert system recommends introducing of new injection wells to improve oil extraction. Since areal wells zonation is not available for more precise determination of the new wells locations, the system generated an ordinary recommendation to put injectors somewhere between producers and on the flank of reservoir. This recommendation is very general and can be used as initial guess for the further implementation of well placement optimization techniques.

Summarizing the expert system validation with the data set of Brugge simulation model, the comparison of previously obtained control results with the system outputs is presented on **Table 5.3**. Here, the following simulation approaches ranking, recommended by the system, highlighted in color:

- green – is recommended for use as the most appropriate;
- yellow – may be used at the discretion of user;
- red – is not recommended.

As expected, decline curve analysis provides fast solution framework, but some of the results might be inaccurate if drive mechanism is different from the fluid expansion dominated. In contrast to that, material balance provides reasonably fast solution framework for different drive mechanisms. However, spatial metrics is not well supported by the method. Finite difference simulation provides most accurate results, but requires more time for data analysis, model construction and simulation. Streamline simulation was rejected by the system because of small model size that is not critical for the use of finite difference simulation with respect to the tested engineering tasks.

Table 5.3: The qualitative comparison of calculated results with the system-predicted ones (Brugge simulation model).

Engineering Task	Comparative criteria	Decline Curve Analysis		Material Balance		Finite Difference Simulation	
		Predicted	Actual	Predicted	Actual	Predicted	Actual
Estimation of recoverable field resources (EUR)	Speed / Accuracy	Fast / Accurate	Fast / Accurate	Fast / Acceptable	Fast / Acceptable	Slow* / Accurate	Slow / Accurate
Estimation of recovery factor	Speed / Accuracy	Fast / Accurate	Fast / Accurate	Fast / Acceptable	Fast / Acceptable	Slow* / Accurate	Slow / Accurate
Estimation of STOIIP	Speed / Accuracy	Fast / Inaccurate	Fast / Inaccurate	Fast / Acceptable	Fast / Acceptable	Slow* / Accurate	Slow / Accurate
Estimation of drainage area	Speed / Accuracy	Fast / Inaccurate	Fast / Inaccurate	Rejected	N/A	Slow* / Accurate	Slow / Accurate
Investigation of water flood feasibility	Speed / Accuracy	Fast / Acceptable	Fast / Acceptable	Fast / Acceptable	Fast / Acceptable	Slow* / Accurate	Slow / Accurate

Slow computational speed was assumed by the expert system for the finite difference simulation approach because the input data set does not contain all the required parameters. Hence, the system considers that simulation model is not prepared yet, and additional time is required besides the computational time.

We can conclude that obtained control results are consistent with the expert system recommendations based on the input data and decision criteria provided.

5.2 Offshore Petroleum Reservoir Model

In order to evaluate applicability of the expert system in real field conditions, we did an additional system test using an offshore reservoir model. This offshore reservoir is characterized as heterogeneous with strong aquifer and free gas cap. The given simulation model consists of 140 000 grid blocks with detailed rock and fluid properties Three years of recorded production data is very noisy, affected by multiple well shut in and gas lift, where decline regime is established for only four last months of production. Current value of recovery factor without water flooding is equal to 48.6 %.

In analogy to the previous section, we have tested the expert system with the same set of simulation tasks. The control results were also obtained using decline curve analysis technique, material balance simulation (MBAL software by Petroleum Experts), and finite difference simulation (Eclipse 100 software by Schlumberger). Finite difference simulation results are used as etalon (reference) solution, since this model is complete, calibrated, and provides reasonable realistic output.

Because the reservoir location and its parameters is non-public information, we do not present the control results calculation details and the expert system reasoning and solution search workflow. Moreover, the control results, shown on **Table 5.4**, were normalized but still give insight into a state of affairs.

Table 5.4: Summary of results calculated by different simulation methods using real petroleum reservoir model

Simulation Task	Units	Decline Curve Analysis	Material Balance	Finite Difference Simulation (REFERENCE)
Estimation of recoverable field resources (EUR), normalized	Ratio to reference	0.99	1.03	1
Estimation of recovery factor, normalized	Ratio to reference	0.99	1.03	1
Estimation of STOIIP, normalized	Ratio to reference	0.95	1	1
Estimation of drainage area, normalized	Ratio to reference	1.39	N/A	1
Investigation of water flood feasibility	---	Indirect, by analogy (water flood is not required)	Indirect, by analogy (water flood is not required)	1 % RF (water flood is not required)

Comparison of control results with the system outputs is presented on **Table 5.5**, where the color scheme of simulation approaches ranking, recommended by the system, is the same as in the previous section. Streamline simulation again was rejected by the expert system because of small model size.

As expected, different simulation techniques provide variation of calculated results with respect to the particular task. Nevertheless, the expert system is capable to predict these outputs and generate relatively correct recommendations.

Table 5.5: The qualitative comparison of calculated results with the system-predicted ones (real petroleum reservoir model)

Simulation Task	**Comparative criteria**	**Decline Curve Analysis**		**Material Balance**		**Finite Difference Simulation (REFERENCE)**	
		Predicted	Actual	Predicted	Actual	Predicted	Actual
Estimation of recoverable field resources (EUR)	Speed / Accuracy	Fast / Accurate	Fast / Accurate	Fast / Acceptable	Fast / Acceptable	Fast / Accurate	Fast / Accurate
Estimation of recovery factor	Speed / Accuracy	Fast / Accurate	Fast / Accurate	Fast / Acceptable	Fast / Acceptable	Fast / Accurate	Fast / Accurate
Estimation of STOIIP	Speed / Accuracy	Fast / Inaccurate	Fast / Inaccurate	Fast / Accurate	Fast / Accurate	Fast / Accurate	Fast / Accurate
Estimation of drainage area	Speed / Accuracy	Fast / Inaccurate	Fast / Inaccurate	Rejected	N/A	Fast / Accurate	Fast / Accurate
Investigation of water flood feasibility	Speed / Accuracy	Fast / Acceptable	Fast / Acceptable	Fast / Acceptable	Fast / Acceptable	Fast / Accurate	Fast / Accurate

Overall, we can conclude that obtained results are also consistent with the expert system recommendations, what confirms correctness of ideas and methodology in the basis of the expert system.

In order to explain why the expert system outputs are slightly different for both Brugge synthetic and offshore petroleum reservoir cases, the comparison of their key features is shown on **Table 5.6**.

Table 5.6: Comparison of Brugge and offshore reservoir outputs

Brugge Synthetic Case	Offshore Petroleum Reservoir Case
Decline Curve Analysis	
4 years of production data with established smooth decline regime.	3 years of production data, very noisy, affected by multiple well shut in and gas lift, decline regime is established for only four last months of production.
Material Balance Method	
Reservoir is heterogeneous, calculated Dykstra-Parson's coefficient is equal to 0.35. Calibrated weak aquifer.	Reservoir is heterogeneous, calculated Dykstra-Parson's coefficient is equal to 0.41. Strong aquifer and free gas cap.
Finite Difference Simulation	
Implemented simulation model does not contain all required parameters to execute simulation. Additional time is required to complete simulation model. Number of grid blocks (60 k) is small, so computational speed is considered as high.	Implemented simulation model does contain all required parameters to execute simulation. Since number of grid blocks (140 k) is small, computational speed is considered as high.

This comparison shows, the expert system is capable to detect correctly these key features that affect accuracy and applicability of particular methods. It is additionally supports previous conclusions that the system predicted outputs are reasonable. Hence, it also proves the correctness of ideas in the basis of the developed methodology.

6. SUMMARY AND FUTURE WORK

This section summarizes previously made conclusions and obtained results. Our vision of the future work and recommendations on the expert system improvement will also be discussed.

In petroleum industry, we are always dealing with processes that we need control and optimize. We have to come up with “educated” actions and decisions in a timely manner to make sure that processes flow in an optimal way. Efficiency of decision-making in reservoir management is strongly dependent on quantity and quality of knowledge about particular subsurface object. Successful search for optimal solution to a particular reservoir engineering problem is always a non-trivial task since it involves analysis and processing of large amounts of data and requires professional expertise in the subject area. Depending on a whole set of aspects of reservoir study, different simulation approaches can or cannot be used because they may give significantly different results even at the same given conditions. The sought-for result here implies finding of the most appropriate simulation approach that provides sufficient accuracy, adequate complexity, and representation of the available data with respect to simulation objectives and existing constraints.

Based on our previous discussions, we have made an attempt to improve selection of the most appropriate simulation method as a part of reservoir management workflow. Summarizing results of extensive literature review and practical work, we decided to design the knowledge-based expert system as a good means for problem solving that requires expertise.

Thorough analysis of existing reservoir evaluation methods and techniques resulted in selection of the set of key parameters. In general, these parameters are quantitative and qualitative variables, which are used for reservoir description and involved into particular problems solving. On the one hand, certain combinations of these parameters are determined by methodology in a basis of the simulation approach to be used for resolving of assigned task. On another hand, specific values of these parameters strongly affect applicability of reservoir evaluation methods and may generate constraints and limitations. Therefore, such dualism is a subject for expertise and establishes basis for the expert system functionality.

Using linguistic method of the Pattern Recognition Theory, the selected set of key parameters was encoded with unique symbolic names and put into alphabet; then, we brought encoded elements into particular sets named scenarios. Scenario is a description of the required set of parameters for solving the certain simulation problem with respect to exact simulation approach. To make this process little easier, the scenario generator was specially created. The full collection of generated scenarios is a core of the expert system called Knowledge Base. Along with the Inference Engine, this base is used to execute technical expertise of the reservoir data and simulation approach selection. We found out that the symbolic (linguistic) data encoding and processing makes the expert system effective, allows further improvement of the proper solution search and realization of an explanatory module. In addition, it enables easy adjustment of the system scope and functionality extension. To wit, once correction of the alphabet content and size or scenarios adjustment are required due to some reasons, it can be easily done. At the same time, these corrections do not affect the developed methodology. The designed workflow is general and remains the same.

In order to enhance the expert system level of confidence, the workflow was equipped with the input data pre-processing module. This module performs data quality control procedures, evaluates applicability of each parameter in the input data set and its ability to produce constraints, and creates the system-generated scenario. If this scenario is not empty, it is further used by the inference engine in match finding with other scenarios within the knowledge base. As a result of the search, the inference engine creates an excerpt, where selected from the knowledge base scenarios are filtered with respect to user- or system-defined constraints and arranged in number of unmatched elements in ascending order.

In case of finding of one (or more) scenario with zero unmatched elements, the expert system determines respective simulation approach and recommends it to a user as the most appropriate in the given conditions. Appropriate means that this particular approach allows to solve assigned problem using available input data, under existing constraints and limitations, and reach the desired level of accuracy and computational time expenses. Otherwise, the expert system displays a list of feasible simulation methods to a user with suggestion about what additionally required parameters should be obtained for each method, and what expected accuracy of solution should be then. Such useful explanatory function of the expert system was realized based on the use of symbolic encoding of scenario elements. Since the exact unmatched elements are detected by the inference engine, their meanings are easily decoded with the alphabet and utilized to generate advices for a user.

The obtained results of the expert system validation using Brugge simulation model and the application with offshore petroleum reservoir data confirm correctness of the tested ideas and methodology in the basis of the system. Simultaneous application of input data analysis and expert reasoning in suitable simulation approach selection, making conclusions, predictions, and recommendations resulted in creation of convenient software tool which can improve quality of reservoir engineering work. The expert system validation and field application tests show simulated results are consistent with the expert system predictions and recommendations. Practically, the developed expert system can be further used as a separate or integrated tool for solving reservoir evaluation problems and for personnel coaching.

The key feature that determines applicability of any expert system is its level of confidence. In other words, it is a capability of the system to provide sufficiently accurate or trustful response to user's request. With regard to the designed expert system, we can define two main ways of the future work related to the system improvement.

First, the input data pre-processing module should be extended by introduction of additional data quality control procedures. For instance, the same methods that are usually applied for quality check of measured PVT-properties and special core analysis data, uncertainty quantification, and other ones could be incorporated in this module, if applicable. It may significantly improve the input data analysis and detection of possible constraints that affect the accuracy and applicability of simulated results, what undoubtedly leads to increase of the expert system efficiency.

Second, we assume realization of the Analytical Block, as it described in the section 4.7 of this work. Integration of this module into the expert system, as we expect, will significantly enhance its level of confidence. Moreover, implementation of that module sets a basis for the learning engine design, what we would like to test. Supplemented by the learning engine, the expert system might educate and train itself. Successful realization of this idea will help us to eliminate the main weakness of expert systems, which is: frequent involvement of software developers is required to support expert systems in actual condition, especially when problem-solving environment is changing.

REFERENCES

Agarwal, R. G., Gardner, D. C., Kleinsteiber, S. W., et al. (1999, October 1). Analyzing Well Production Data Using Combined-Type-Curve and Decline-Curve Analysis Concepts. Society of Petroleum Engineers. doi:10.2118/57916-PA

Al-Kaabl, A.-A. U., McVay, D. A., and Lee, J. W. 1990. Using an Expert System to Identify a Well-Test-Interpretation Model. Society of Petroleum Engineers. doi:10.2118/18158-PA

Al-yami A.S., Schubert, J. 2012. Underbalanced Drilling Expert System Development. Society of Petroleum Engineers. doi:10.2118/152101-MS

AlMousa, T. s., Ertekin, T. 2013. Development and Utilization of Integrated Artificial Expert Systems for Designing Multi-Lateral Well Configurations, Estimating Reservoir Properties and Forecasting Reservoir Performance. Society of Petroleum Engineers. doi:10.2118/167462-MS

Arps, J. J. (1945, December 1). Analysis of Decline Curves. Society of Petroleum Engineers. doi:10.2118/945228-G

Artun, E., Ertekin, T., Watson, R. W., et al. (2009, January 1). Development of Universal Proxy Models for Screening and Optimization of Cyclic Pressure Pulsing in Naturally Fractured Reservoirs. International Petroleum Technology Conference. doi:10.2523/IPTC-13663-MS

Aziz, K., Settari, A. 1979. Petroleum reservoir simulation, Blitzprint Ltd., Calgary, Alberta, Canada. ISBN: 0-97360614-0-5

Badiru A.B., Cheung J., 2002. Fuzzy Engineering Expert Systems with Neural Network Applications, John Wiley & Sons, Inc., New York, NY, USA. ISBN: 0471293318

Balch, R. S., Weiss, W. W., and Ruan, T. et al. 2003. Simulated Expert Interpretation of Regional Data to Predict Drilling Risk. Society of Petroleum Engineers. doi:10.2118/84067-MS

Batycky, R. P. 1997. A Three-dimensional Two-phase Scale Streamline Simulator, PhD dissertation, The Department of Petroleum Engineering, Stanford University

Berenblyum, R. A., Shapiro, A. A., Jessen, K., Stenby, E. H., and Orr, F. M. (2003, January 1). Black Oil Streamline Simulator With Capillary Effects. Society of Petroleum Engineers. doi:10.2118/84037-MS

BizRules, 2006-2007. What is the difference between data-based, rule-based, and knowledge-based systems? http://www.bizrules.info/page/art_dbvsrbvskb.htm

Blasingame, T. A., McCray, T. L., and Lee, W. J. (1991, January 1). Decline Curve Analysis for Variable Pressure Drop/Variable Flowrate Systems. Society of Petroleum Engineers. doi:10.2118/21513-MS

Campbell, R.A. and Campbell, J.M. 1978. Mineral Property Economics, Volume 3, Petroleum Property Evaluation , Campbell Petroleum Series, Norman, OK, USA.

Chaban, L.N., 2004. Theory and Algorithms of Pattern Recognition. Textbook Series, Russian Federation Ministry of Education and Science, Moscow State University of Geodesy and Cartography, Moscow, Russia.

Cheng, H., Oyerinde, A. S., Datta-Gupta, A., and Milliken, W. J. (2006, January 1). Compressible Streamlines and Three-Phase History Matching. Society of Petroleum Engineers. doi:10.2118/99465-MS

Christie, M. A., and Bazargan, H. (2012, January 1). Efficient Polynomial Chaos Proxy-based History Matching and Uncertainty Quantification for Complex Geological Structures. Society of Petroleum Engineers. doi:10.2118/163282-MS

Clancey, W.J. Heuristic classification, Artificial Intelligence, Volume 27, Issue 3, 1985, Pages 289-350, ISSN 0004-3702, http://dx.doi.org/10.1016/0004-3702(85)90016-5.

Crane, M., Bratvedt, F., Bratvedt, K., Childs, P., and Olufsen, R. (2000, January 1). A Fully Compositional Streamline Simulator. Society of Petroleum Engineers. doi:10.2118/63156-MS

Cullick, A. S., Johnson, W. D., and Shi, G. (2006, January 1). Improved and More Rapid History Matching with a Nonlinear Proxy and Global Optimization. Society of Petroleum Engineers. doi:10.2118/101933-MS

d'Almeida, A. L., Silva, I. M., and Ramos, J. S. 1997. A Fishing Operations Expert System. Society of Petroleum Engineers. doi:10.2118/37590-MS

Dake, L.P. 1978. Fundamentals of Reservoir Engineering, Elsevier, Amsterdam, The Netherland. ISBN: 978-0-444-41830-2

Datta-Gupta, A., King, M. J. 2007. Streamline Simulation: Theory and Practice, Society of Petroleum Engineers, Richardson, TX, USA. ISBN: 978-1-55563-11-6

Denney, D. 1999. Management of Well Production With Real-Time Expert Systems. Society of Petroleum Engineers. doi:10.2118/1099-0034-JPT

Dharan, M. B., Turek, E. A., and Vogel, J. L. (1989, January 1). The Fluid Properties Measurement Expert System. Society of Petroleum Engineers. doi:10.2118/19134-MS

Di Donato, G., Huang, W., and Blunt, M. (2003, January 1). Streamline-Based Dual Porosity Simulation of Fractured Reservoirs. Society of Petroleum Engineers. doi:10.2118/84036-MS

Duggal, S.M., Chhabra, R. Learning Systems and Their Applications: Future of Strategic Expert Systems, International Association for Computer Information Systems, Volume 3, March 2002, Pages 179-185. http://www.iacis.org/iis/iis_articles.php?volume=3

Engelmore, R. S., Feigenbaum, E. 1993. Knowledge-Based Systems in Japan. Japanese Technology Evaluation Center. Available electronically at http://www.wtec.org/loyola/kb/c1_s1.htm

Erdle, J. C., Archer, D. A., Stiff, T. J., and Callihan, M. C. (1986, January 1). Well Test Engineering Software With Built-In Expert Advice. Society of Petroleum Engineers. doi:10.2118/15309-MS

Ershaghi, I., Omorigie, O. (1978, February 1). A Method for Extrapolation of Cut vs. Recovery Curves. Society of Petroleum Engineers. doi:10.2118/6977-PA

Ertekin, T., Abou-Kassem, J. H., and King, R. G. 2001. Basic Applied Reservoir Simulation, Society of Petroleum Engineers, Richardson, TX, USA. ISBN: 1-55563-089-8

Falkenhainer, B. and Forbus, K. D., 1991. Compositional modelling: finding the right model for the job, Artificial Intelligence, Volume 51, Issues 1–3, October 1991, Pages 95-143, ISSN 0004-3702, http://dx.doi.org/10.1016/0004-3702(91)90109-W

Fanchi, J. R. 2006. Principles of Applied Reservoir Simulation, Third edition, Gulf Professional Publishing, Burlington, MA, USA. ISBN: 978-0-7506-7933-6. http://dx.doi.org/10.1016/B978-075067933-6/50000-3

Fetkovich, M. J., Fetkovich, E. J., and Fetkovich, M. D. (1996, February 1). Useful Concepts for Decline Curve Forecasting, Reserve Estimation, and Analysis. Society of Petroleum Engineers. doi:10.2118/28628-PA

Garrouch, A. A., Malallah, A., and AlEnizy, M. 2006. A Comprehensive Expert System for Diagnosing and Assessing Formation Damage. Society of Petroleum Engineers. doi:10.2118/99277-MS

Giarratano J.C., Riley G. 2004. Expert Systems: Principles and Programming, Fourth Edition, MA: Thomas / PSW Publishing Co., Boston, MA, USA. ISBN: 0534384471

Guerillot, D. R. 1988. EOR Screening With an Expert System. Society of Petroleum Engineers. doi:10.2118/17791-MS

Guyaguler, B., Horne, R. N., Rogers, L., et al. (2000, January 1). Optimization of Well Placement in a Gulf of Mexico Waterflooding Project. Society of Petroleum Engineers. doi:10.2118/63221-MS

Havlena, D., and Odeh, A.S. The Material Balance as an Equation of Straight Line, Journal of Petroleum Technology, August 1963, Pages 896-900.

Havlena, D., and Odeh, A.S. The Material Balance as an Equation of Straight Line – Part II, Field Cases, Journal of Petroleum Technology, July 1964, Pages 815-822.

Hayes-Roth, F., Waterman, D. A, et al. 1983. Building Expert Systems. Addison-Wesley Publishing Company Inc., Reading, MA, USA. ISBN: 0201106868

Heinze, L. R. 1993. CHES: Casing Hydraulic Expert System. Society of Petroleum Engineers. doi:10.2118/24420-PA

Islam, M. R., Moussavizadegan, S. H., Mustafiz, S., and Abou-Kassem, J.H. 2010. Advanced Petroleum Reservoir Simulation, Scrivener Publishing LLC, Salem, MA, USA. Online ISBN: 9780470650684. DOI: 10.1002/9780470650684

Jang, M., Lee, J., Choe, J., and Kang, J. M. (2002, January 1). An Analysis of Advection-Dispersion Ratio to Incorporate Dispersive Transport into Streamline Simulation. Society of Petroleum Engineers. doi:10.2118/77380-MS

Jessen, K., and Orr, F. M. (2004, January 1). Gravity Segregation and Compositional Streamline Simulation. Society of Petroleum Engineers. doi:10.2118/89448-MS

Junker, H., Dose, T., Plas, L., and Little, A. J. H. (2006, January 1). Modern Approach to Estimation of Uncertainty of Predictions With Dynamic Reservoir Simulation - A

Case Study of a German Rotliegend Gasfield. Society of Petroleum Engineers. doi:10.2118/103340-MS

Jurecka, F. 2007. Robust Design Optimization Based on Metamodeling Techniques. PhD thesis, Fakultät für Bauingenieur und Vermessungswesen der Technischen Universität München

Kabir, C. S., Chawathe, A., Jenkins, S. D., et al. (2002, January 1). Developing New Fields Using Probabilistic Reservoir Forecasting. Society of Petroleum Engineers. doi:10.2118/77564-MS

Kaimal, L. B., Metkar, A. R., and Rakesh, G. Self-Learning Real Time Expert Systems, International Journal on Soft Computing, Artificial Intelligence and Applications, Volume 3, No. 2, May 2014, Pages 13-25. doi: 10.5121/csit.2013.3641

Khan, S. A., Pope, G. A., and Sepehrnoori, K. 1993. An Expert System for Miscible Gasflooding. Society of Petroleum Engineers. doi:10.2118/24431-PA

Kjell, T., Baleix, J. M. 1992. Application of a Pressure Analysis Expert System During Exploration Drilling. Society of Petroleum Engineers. doi:10.2118/25000-MS

Krishnamoorthy, C.S., Rajeev, S. 1996. Artificial intelligence and expert systems for engineers. CRC Press Inc., Boca Raton, FL, USA. ISBN: 0-8493-9125-3

Kulakofsky, D. S., Wu, F.-L., Onan, D. D. et al. 1993. Development and Application of a Knowledge-Based Expert System for Cement Slurry Design. Society of Petroleum Engineers. doi:10.2118/24417-PA

Leibowitz, J., 1997. The Handbook of Applied Expert Systems, CRC Press LLC, Boca Raton, FL, USA. ISBN: 0-8493-3106-4

Lepsky, A.E., Bronevich, A.G. 2009. Mathematical Methods of Pattern Recognition: Lectures course. Textbook, Russian Federation Ministry of Education and Science, Technological Institute of the South Federal University, Taganrog, Russia. ISBN: 978-5-8327-0306-0

Levy, A. Y., Iwasaki, Y., Fikes, R., 1997. Automated model selection for simulation based on relevance reasoning, Artificial Intelligence, Volume 96, Issue 2, November 1997, Pages 351-394, ISSN 0004-3702, http://dx.doi.org/10.1016/S0004-3702(97)00056-8

Lophaven, S.N., Nielsen, H.B. and Sondergaard, J., 2002. DACE — a Matlab Kriging Toolbox, Version 2, Informatics and mathematical modelling. Technical University of Denmark, Technical Report No. IMM-TR-2002-12.

Mabile, C. M., Hamelin, J.-P. A., du Chaffaut, B. A. et al. 1989. An Expert System Helps in Formation Recognition. Society of Petroleum Engineers. doi:10.2118/19132-MS

Malak, R. J. and Paredis, C. J. J., 2007. Validating behavioral models for reuse, Research in Engineering Design, Volume 18, Issue 3, August 2007, Pages 111 – 128, ISSN 1435-6066, http://dx.doi.org/10.1007/s00163-007-0031-0

Martí, R., Reinelt, G. 2011. The Linear Ordering Problem: Exact and Heuristic Methods in Combinatorial Optimization. Springer-Verlag Berlin Heidelberg, Berlin, Germany. ISBN 978-3-642-16728-7; DOI 10.1007/978-3-642-16729-4

Martinez, E. 1992. Directional Drilling Expert System. Society of Petroleum Engineers. doi:10.2118/23664-MS

Martinez, E. R., Moreno, W. J., Castillo, V. J. et al. 1993. Rod Pumping Expert System. Society of Petroleum Engineers. doi:10.2118/26246-MS

Mattax, C. C., and Dalton, R. L. (1990, June 1). Reservoir Simulation (includes associated papers 21606 and 21620). Society of Petroleum Engineers. doi:10.2118/20399-PA

McVay, D.A., and Dossary, M.N. 2014. The Value of Assessing Uncertainty. SPE Economics & Management, April 2014. SPE-160189. doi:10.2118/160189-MS

Moreno, J., Kazemi, H., and Gilman, J. R. (2004, January 1). Streamline Simulation of Countercurrent Water-Oil and Gas-Oil Flow in Naturally Fractured Dual-Porosity Reservoirs. Society of Petroleum Engineers. doi:10.2118/89880-MS

Moridis, G. J., Kuzma-Anderson, H., Reagan, M. T. et al. 2011. A Self-Teaching Expert System for the Analysis, Design, and Prediction of Gas Production from Unconventional Gas Resources. Society of Petroleum Engineers. doi:10.2118/149485-MS

Muggeridge A, Cockin A, Webb K, Frampton H, Collins I, Moulds T, Salino P. 2014 Recovery rates, enhanced oil recovery and technological limits. Phil. Trans. R. Soc. A 372: 20120320. http://dx.doi.org/10.1098/rsta.2012.0320

Mustafiz, S., and Islam, M. R. 2008. State-of-the-art Petroleum Reservoir Simulation, Petroleum Science and Technology, 26:10-11, Pages 1303-1329. doi: 10.1080/10916460701834036

Nashawi, I. S., Malallah, A. 2009. Improved Electrofacies Characterization And Permeability Predictions In Sandstone Reservoirs Using a Data Mining And Expert System Approach. Society of Petrophysicists and Well-Log Analysts.

Noran O.S., The Evolution of Expert Systems, available electronically at http://cit.gu.edu.au/~noran

Oberkampf, W. L. and Roy, C. J., 2010. Verification and Validation in Scientific Computing, Cambridge University Press, New York, USA. ISBN: 978-0-521-11360-1

Obi, E.-O. I., Blunt, M. J. 2006. Streamline-based simulation of carbon dioxide storage in a North Sea aquifer, Water Recourses, Res., 42, W03414, doi:10.1029/2004WR003347.

Odeh, A. S. 1969. Reservoir Simulation ...What Is It. Society of Petroleum Engineers. doi:10.2118/2790-PA

Onwunalu, J. E., Litvak, M. L., Durlofsky, L. J., and Aziz, K. (2008, January 1). Application of Statistical Proxies to Speed Up Field Development Optimization Procedures. Society of Petroleum Engineers. doi:10.2118/117323-MS

Osako, I., and Datta-Gupta, A. (2007, January 1). A Compositional Streamline Formulation With Compressibility Effects. Society of Petroleum Engineers. doi:10.2118/106148-MS

Osterloh, W. T. (2008, January 1). Use of Multiple-Response Optimization to Assist Reservoir Simulation Probabilistic Forecasting and History Matching. Society of Petroleum Engineers. doi:10.2118/116196-MS

Pan, Y., and Horne, R. N. (1998, January 1). Improved Methods for Multivariate Optimization of Field Development Scheduling and Well Placement Design. Society of Petroleum Engineers. doi:10.2118/49055-MS

Pandey, V., Osisanya, S. O. 2001. Development of an Expert System for Solids Control in Drilling Fluids. Petroleum Society of Canada. doi:10.2118/01-09-05

Pasarai, U., and Arihara, N. (2005, January 1). Application of Streamline Method to Hot Water-Flooding Simulation for Heavy Oil Recovery. Society of Petroleum Engineers. doi:10.2118/93149-MS

Pearl, J. 1984. Heuristics: Intelligent Search Strategies for Computer Problem Solving, Addision-Wesley Longman Publishing Co., Inc. Boston, MA, USA. ISBN: 0-201-05594-5

Peters, L., Arts, R., Brouwer, G., and Geel, C. (2009, January 1). Results of the Brugge Benchmark Study for Flooding Optimisation and History Matching. Society of Petroleum Engineers. doi:10.2118/119094-MS

Pfeiffer, T., Reza, Z., Schechter, D., McCain, W. D., and Mullins, O. C. (2011, January 1). Fluid Composition Equilibrium; a Proxy for Reservoir Connectivity. Society of Petroleum Engineers. doi:10.2118/145703-MS

Russel, S., Norvig, P. 2010. Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River, NJ, USA. ISBN: 0-13-604259-7

Sanjay, V., Anuj, G., and Sharma, M. M. 1989. A Rule-Based System for Estimating Clay Distribution, Morphology, and Formation Damage In Reservoir Rocks. Society of Petroleum Engineers. doi:10.2118/16870-PA

Satter, A., Iqbal, G. M., Buchwalter, J. L. 2008. Practical Enhanced Reservoir Engineering: Assisted with Simulation Software, PennWell Corporation, Tulsa, OK, USA. ISBN: 1-59370-056-3

Sheremetov, L. B., Cosultchi, A., Batyrshin, I. et al. 2007. New Insights and Applications of Soft Computing on Analysis of Water Production from Oil Reservoirs. Society of Petroleum Engineers. doi:10.2118/108702-MS

Siller W., Buckley J.J., 2005. Fuzzy Expert Systems and Fuzzy Reasoning, John Wiley & Sons, Inc., Hoboken, NJ, USA. ISBN: 0-471-38859-9

Siripatrachai, N., Rana, S., Bodipat, K., and Ertekin, T. 2014. An Effective Coupling of Type Curves and Expert Systems for Evaluating Multi-stage Hydraulically Fractured Horizontal Wells in Composite Dual-porosity Shale Gas Reservoirs. Society of Petroleum Engineers. doi:10.2118/170963-MS

Slotte, P. A., and Smorgrav, E. (2008, January 1). Response Surface Methodology Approach for History Matching and Uncertainty Assessment of Reservoir Simulation Models. Society of Petroleum Engineers. doi:10.2118/113390-MS

Smedley, M., 2015. Shell Loses $ 6 Billion on Q3 Writedowns. Petroleum Economist. 16 November, 2015.

Srinivasan, K., Ertekin, T. 2008. Development and Testing of an Expert System for Coalbed Methane Reservoirs Using Artificial Neural Networks. Society of Petroleum Engineers. doi:10.2118/119935-MS

Sun, H. 2015. Advanced Production Decline Analysis and Application, Gulf Professional Publishing, Waltham, MA, USA. ISBN: 978-0-12-802411-9

Surguchev, L. M., Zolotukhin, A. B., and Bratvold, R. B. 1992. Reservoir Characterization and Modeling Supported by Expert Knowledge Systems. Society of Petroleum Engineers. doi:10.2118/24279-MS

Tanaka, S., Datta-Gupta, A., and King, M. J. (2014, April 12). Compositional Streamline Simulation of CO2 Injection Accounting for Gravity and Capillary Effects Using Orthogonal Projection. Society of Petroleum Engineers. doi:10.2118/169066-MS

Tehrani, D. H. (1972, January 1). An Analysis Of Volumetric Balance Equation For Calculation Of Oil-In-Place And Water Influx. Society of Petroleum Engineers. doi:10.2118/3960-MS

Thiele, M. R., Batycky, R. P., and Blunt, M. J. (1997, January 1). A Streamline-Based 3D Field-Scale Compositional Reservoir Simulator. Society of Petroleum Engineers. doi:10.2118/38889-MS

Thiele, M., Batycky, R., Pöllitzer, S., and Clemens, T. (2010, April 1). Polymer-Flood Modeling Using Streamlines. Society of Petroleum Engineers. doi:10.2118/115545-PA

Tou, J.T., Gonzalez, R.C. 1974. Pattern Recognition Principles. Addision-Wesley Longman Publishing Co., Inc. Boston, MA, USA.

Tracy, G. W. (1955, January 1). Simplified Form of the Material Balance Equation. Society of Petroleum Engineers.

Van Domelen, M. S., Ford, W. G. F., and Chiu, T. J. 1992. An Expert System for Matrix Acidizing Treatment Design. Society of Petroleum Engineers. doi:10.2118/24779-MS

Vicente, B. J., Priimenko, V. I., and Pires, A. P. (2014, May 21). Numerical Simulation of 2D Hot Waterflooding Using Analytical Solutions Along Streamlines. Society of Petroleum Engineers. doi:10.2118/169361-MS

Waterman D.A., 1986. A guide to expert systems, First edition, Reading, Mass.: Addison-Wesley Publishing Company, Lebanon, IN, USA. ISBN: 0-201-08313-2

Whittaker, A. H., Macpherson, J. D. (1986, January 1). Assistant Expert Log Data Analysis System: Applications Of Artificial Intelligence Programming Techniques In Formation Evaluation. Society of Petrophysicists and Well-Log Analysts.

Yang, Y., Davidson, J. E., Fenter, D. J., et al. (2009, January 1). Reservoir Development Modeling Using Full Physics And Proxy Simulations. International Petroleum Technology Conference. doi:10.2523/IPTC-13536-MS

Yeten, B., Castellini, A., Guyaguler, B., et al. (2005, January 1). A Comparison Study on Experimental Design and Response Surface Methodologies. Society of Petroleum Engineers. doi:10.2118/93347-MS

Zhu, Z., Gerritsen, M., and Thiele, M. (2010, June 1). Thermal Streamline Simulation for Hot Waterflooding. Society of Petroleum Engineers. doi:10.2118/119200-PA

Zhu, Z., Thiele, M. R., & Gerritsen, M. G. (2011, January 1). Thermal Streamline Simulation: Steam Floods. Society of Petroleum Engineers. doi:10.2118/139501-MS

Zubarev, D. I. (2009, January 1). Pros and Cons of Applying Proxy-models as a Substitute for Full Reservoir Simulations. Society of Petroleum Engineers. doi:10.2118/124815-MS

APPENDIX A

TABLES

Table A-1: Alphabet – list of encoded key parameters combined into groups

Simulation approach - SA		**Reservoir geometry data - RG**	
CDC	correlation (decline curve)	GBM	grid block mesh [number of grid blocks]
PRX	proxy modelling	ARE	reservoir areal extend [acre, sc.ft]
MBL	material balance	THC	reservoir thickness [ft]
STL	stream line		
FDV	finite difference (volume) numerical	**Reservoir rock properties - RP**	
		POR	porosity
Field maturity (life stage) - FM		RCM	rock compressibility
EXP	exploration	RPI	reservoir initial pressure [psi]
APP	appraisal	RCP	capillary pressure distribution
DEV	development		
PLT	plateau	**Saturations and Relative permeabilities - PR**	
DCL	decline	WSI	Irreducible water saturation
		OSR	residual oil saturation to water
Simulation task (goal) - ST		OSG	residual oil saturation to gas
DHP	hydrocarbon in place	MSG	minimum gas saturation
RRS	recoverable resource / EUR	WRP	water relative permeability
RFT	recovery factor calculation	ORW	oil relative permeability to water
DAE	drainage area estimation	ORG	oil relative permeability to gas
PRC	production rate calculation	GRP	gas relative permeability
RPC	reservoir pressure calculation	PER	absolute permeability
HIM	history matching	WST	average water saturation
RDP	reservoir drainage zones delineation	OST	average oil saturation
IWJ	in-fill well justification	GST	average gas saturation
WLP	well placement justification		
RSO	reservoir sweep optimization	**Reservoir heterogeneity - RH**	
FLJ	flood feasibility	HMG	homogeneous
FLO	flood optimization	HTG	heterogeneous
UGM	upscaling of geological model		
UCF	uncertainty quantification		
DMD	drive mechanism determination		

Table A-1 Continued

Fluid type - FT		Number of wells - NW	
OIL	oil	PWN	production wells number
GAS	gas	IWN	injection wells number
WAT	water		
COM	composite fluid (multicomponent)	**Production data - PD**	
		PROR	production oil rate (well/field) [STB/day]
Fluid properties - FP		PRWR	production water rate (well/field) [STB/day]
ODN	oil density [lb/cu.ft]	PRGR	production water rate (well/field) [SCF/day]
WDN	water density [lb/cu.ft]	PWLR	well liquid production rate [STB/day]
GDN	gas density [lb/cu.ft]	PWLC	well liquid cumulative production [STB]
OFR	oil formation volume factor related to pressure [rb/STB]	PFLR	field liquid production rate [STB/day]
WFR	water formation volume factor related to pressure [rb/STB]	PFLC	field liquid cumulative production [STB]
GFR	gas formation volume factor related to pressure [rb/SCF]	PWOR	well oil production rate [STB/day]
GSR	gas solution in oil related to pressure [SCF/STB]	PWOC	well oil cumulative production [STB]
OCM	oil compressibility [1/psi]	PFOR	field oil production rate [STB/day]
WCM	water compressibility [1/psi]	PFOC	field oil cumulative production [STB]
GCM	gas compressibility [1/psi]	PWWR	well water production rate [STB/day]
OVS	oil viscosity [cP]	PWWC	well water cumulative production [STB]
WVS	water viscosity [cP]	PFWR	field water production rate [STB/day]
GVS	gas viscosity [cP]	PFWC	field water cumulative production [STB]
GDF	gas deviation factor	PWGR	well gas production rate [SCF/day]
		PWGC	well gas cumulative production [SCF]
Initial volume of fluid in place - IV		PFGR	field gas production rate [SCF/day]
OIP	oil initially in place	PFGC	field gas cumulative production [SCF]
WIP	water initially in place	PWF	production well bottom hole pressure [psi]
GIP	gas initially in place	PTM	production time [hours, days, months]
		PTH	production well head pressure [psi]

Table A-1 Continued

Injection data - ID		**Constraints - CS**	
INWR	injection water rate (well/field) [STB/day]	CTA	computational time advantage [G - good, P - poor]
INGR	injection water rate (well/field) [SCF/day]	FSE	field scheduled events [for predictability]
IWWR	well water injection rate [STB/day]	FLS	field life stage goes from FM (field maturity) (affects accuracy ACC)
IWWC	well water cumulative injection [STB]	ACC	accuracy [L - limited, F - fair, G - good, B - the best]
IFWR	field water injection rate [STB/day]	CPH	complex physics
IFWC	field water cumulative injection [STB]	GRA	gravity effect is exist (ODN>=WDN at surface) (less than 10 yields constrain for streamline)
IWGR	well gas injection rate [SCF/day]	PWC	critical value of the water cut
IWGC	well gas cumulative injection [SCF]	RDM	recovery drive mechanism (W - water, G - gas, S - solution gas drive, E - oil expansion)
IFGR	field gas injection rate [SCF/day]	PDE	production decline is established (Y - yes, N - n)
IFGC	field gas cumulative injection [SCF]	NGD	number of grid blocks (more than 100000 for black oil advantage in CTA for streamline vs FD) (more than 28000 for compositional advantage in CTA for streamline vs FD)
IWF	injection well bottomhole pressure [psi]		
Fluid contacts - FC			
WOC	oil-water contact / aquifer		
GOC	gas-oil contact	HTL	level of heterogeneity by Dykstra-Parson [0 … 1] (HTL>0.25 limits use of MBL; HTG > 0.5 advantage in use of streamline vs FD for black oil)
GWC	gas-water contact		
FEC	fluid expansion when GOC and WOC are absent		
Source of rock and fluid properties - SP			
GSI	geologic and seismic interpretation		
ANL	analogue		
WLA	well logging		
WTA	well test		
CRA	laboratory core analysis		
PVA	laboratory fluid analysis		
IWI	inter-well interpolation		