*Geosciences and Engineering, Vol. 9, No. 14 (2021), pp. 128–139.*
*https://doi.org/10.33030/geosciences.2021.14.128*

# PROXY MODEL FOR HYDROCARBON RECOVERY IN A SEVEN-SPOT WATERFLOODED WELL PATTERN

DÁNIEL BÁNKI[1] – ZOLTÁN TURZÓ[2, *]

[1, 2]*Petroleum and Natural Gas Institute, Faculty of Earth Science and Engineering,*
University of Miskolc
[2]*turzoz@uni-miskolc.hu*

**Abstract:** This article is about establishing an artificial modeling environment using real data, then creating special formulae in order to estimate the recovery factor of the depletion of the reservoir model. The key concept is that these aids in the form of formulae can be very useful, economic and fast methods to help reservoir simulation. Waterflooding was examined in previous years and a sensitivity analysis was conducted several times. First the parameters were screened for impact and importance, then the structure of the resulting proxy model was chosen, along with the accuracy, both based on the focus of investigation itself, which is the recovery factor. The artificial reservoir environment is a seven-spot water flooded well pattern in an initially undersaturated oil reservoir. The results in this particular study are three formulae which are capable of predicting the recovery factor with a satisfactory error margin after five, ten, and fifteen years of production as a function of the initial produced fluid rate target.

***Keywords:*** *reservoir simulation, proxy model, recovery factor estimation*

## 1. INTRODUCTION

The aim of analytical formulae, proxy models, and any mathematical short-term estimation tool is to help make complex calculations simple and to aid in fast decision making. The accuracy of these is always less than a normal modeling system's performance in this area, and often more application borders need to be set up in order to keep the formula working, but all these fall into the normal nature and behavior of the application of these methods. The benefit is more rapid decision making, due to the fact that for example in reservoir simulation no big simulation runs are needed to find out essential results of smaller modifications, thus saving both time and money.

This nature of descriptive science has always been part of reservoir engineering, geophysics, drilling engineering and so on [1]. One can call these rules of thumb, empirical formulae, or base functions, but the idea is the same: to provide an easier, more affordable tool for performing relatively small calculations, rough estimations etc. In order to have a better understanding of and justification over decisions made during the simulation and regression workflow, a short summary of previous studies is needed.

The first step in examining the behavior of complex hydrodynamic systems is to find a suitable environment for investigation [5]. This step may seem easy, but in the practical world of reservoir simulation this is the first challenges to overcome. The reason is quite simple: in this area of science, if somebody is familiar with the development, constraint and un-readiness of a field data reservoir model, it is not hard to conclude that if we would like to observe and describe something from the beginning, these environments are far from ideal. The magnitude of uncertainty and the size and reliability of these models are perfectly good for industrial use, but are less suitable base for an academic research project [1] [6].

The next step further is the introduction of artificial models, but this process brings up more problems to solve, even before any investigation is started. First and foremost, when using reservoir simulators, several iterational processes and flow equations are being solved in the background, and inside the solution chain there are also some empirical formulae. All in all, it is not sufficient to have artificial and real data sewn together; one has to have realistic artificial datasets. The pressure, composition, soluted content of gas within the oil phase and phase density according to this must be realistically paired up with the saturation functions, and then this set of data has to imported into an also realistic environment in terms of pressure, temperature, depth, contacts, etc. Even when using multiple real data sources, cautious attention has to be used to generate a hydrodynamic system which really could have been generated naturally.

As an example, if the PVT data are imported from real measurements of a heavy oil mixture with the lack of some intermediate hydrocarbons, then in theory the density and solution gas oil ratio can be imported from another real dataset, but in this case the density should be quite a bit higher than normal oil density, and the solution gas oil ratio should not be too big, hence the fact that this composition cannot really hold solution gas in the liquid phase without proper intermediate content. Usually, that is the reason why PVT datasets are almost always imported from the same source. Saturation functions (relative permeability, capillary pressure), can, in theory, be originated from different sources, but the same rules apply in their implementation.

After setting up a stable and realistic modeling environment, the well and production data is the next step. For this, a seven-spot pattern was used with peripheral waterflooding (six injectors) and one production well. If the dynamic variables are set, the sensitivity analysis is the following procedure. In this case the variable was the starting liquid rate target. The accurate rate changes throughout the production period due to natural behavior of production rate decline, but for convenience, later on, the desired starting rates will be referred as production rate. Once a desired rate is set, the production can only keep up that rate for a certain amount of time, and then as depletion follows, the oil production rate will decline.

With the results for the recovery factor after 5, 10, 15 years of production, the datasets for regression were basically gained, but before fitting a formula onto the

dataset, the validity borders and even the extents of the data ranges have to be investigated and post processed. Once the "valuable" data remain, the next step is regression in order to find out the best correlation between the starter rates and the recovery factor. In particular, three third-degree polynomial formulae were found and 2D and 3D response surfaces were also generated.

## 2. MODEL SETUP AND PROPERTIES

In this chapter, the main characteristics of the reservoir modeling environment will be discussed. In order to establish dynamic flow modeling, it is necessary to start with the static environment, the grid, and the static reservoir properties themselves. Then the fluid, rock and saturation parameters are added, such as compressibilities, reference depths and pressures, relative permeability and capillary pressure datasets, densities etc. As a final configuration, well placement, completion data, wellbore sizes, and skin factors are added, and in the last section production rates and limits and injection controls are added for customized depletion. The datafile ends with the timestep controls, and in the last section there is also possibility to tune the solution controls, for example the maximum number of non-Newtonian iterations, etc.

### 2.1 Grid, geometry and reservoir parameters

The main aim was to generate an artificial reservoir model depleted with waterflooding, and for this, a seven-spot well pattern was chosen. The inclusive grid consists of 13 × 9 × 10 cell blocks, each block is 360 × 360 feet, with a thickness of 15 feet. This is the main frame, from which, using cell deactivation, the hexagonal shape of the well-pattern was carved out, leaving 77 cells out of the original 117 in each layer. Cell layers are defined for modeling, but for regions, lithological layers should be established. That means in this case, that out of the 10 cell layers, 5 "lithological" units were formed, each having different porosity and permeability values. The environment mimicked is a sandstone reservoir, so the porosity-permeabilities values fit into this trend and scale. The depth of the cell layer tops and the regarding poro-perm data can be found in *Table 1*.

***Table 1***

*Regions, porosities and permeabilities of the reservoir*

| Layer | Region | Top of structure [ft] | Porosity [%] | Permeability [mD] |
|---|---|---|---|---|
| 1 | 1 | 7,800 | 15 | 170 |
| 2 | | 7,815 | | |
| 3 | | 7,830 | | |
| 4 | 2 | 7,845 | 20 | 600 |
| 5 | | 7,860 | | |
| 6 | 3 | 7,875 | 15 | 170 |
| 7 | 4 | 7,890 | 18 | 200 |
| 8 | | 7,905 | | |
| 9 | 5 | 7,920 | 20 | 600 |
| 10 | | 7,935 | | |

The top of the reservoir is at 7,800 ft in depth, with the 10 ×15 feet cell thickness yielding a total reservoir thickness of 150 feet. The oil-water contact is set to 8,200 feet, acting as a semi-closed reservoir with a passive aquifer, hence the transformation of the pore volume higher than the oil-water-contact (OWC). The initial average reservoir pressure is 3,316 psi. The pressure and the solution gas-oil ratio were set to yield an undersaturated oil reservoir at the start of the simulation. [3] A 3D picture of the grid with the active cells, major extents and the wells is shown in *Figure 1*.

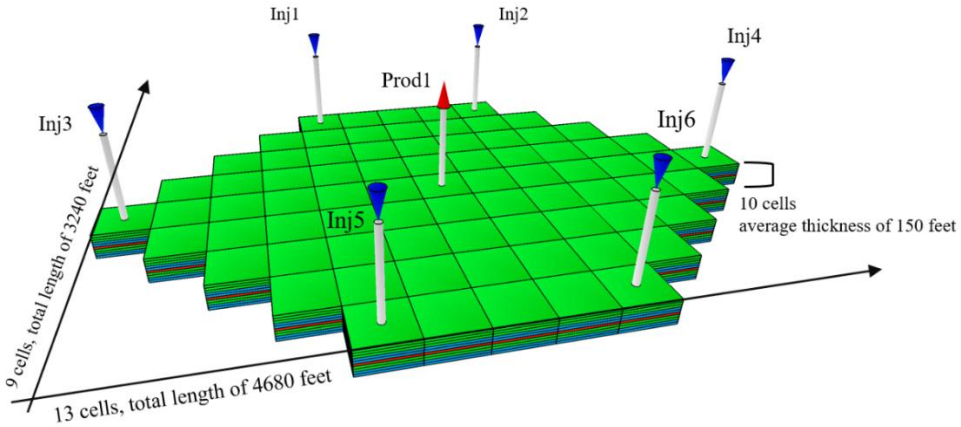


***Figure 1***

*Geometry, extent and injection and production well positioning of the reservoir*

## 2.2 Fluid properties and saturation functions

The initial calculation of the modeling process is called equilibration, often regarded as time zero calculation. Based on the built-up grid, the rest of the cells should also be filled with fluid parameters. From the reference pressures, depths, hydrostatics with the densities, capillary pressure and relative permeability curves equilibrium is calculated in order to yield the initial fluid in place data before even the first timestep of dynamic modeling is reached [4].

For the success of this equilibration, the next process is to implement all the fluid properties needed. Density of the phases for oil, water and gas were 52.9989, 66.7593 and 0.0651035 in lb/ft$^3$, respectively. For both water and reservoir rock, a reference pressure and compressibility data pair were also used. For PVT data, real fluid measurement was used, where datasets included bubble point pressure, formation volume factor and viscosity in terms of gas-oil ratio for oil, and formation volume factor with viscosity for the gas phase. The reason not including these here is the length and size of the datasets: only for the oil phase, one set of PVT data is more than three pages. As an example, a graph for the viscosity vs pressure vs solution gas oil ratio for the oil phase is displayed in *Figure 2*.

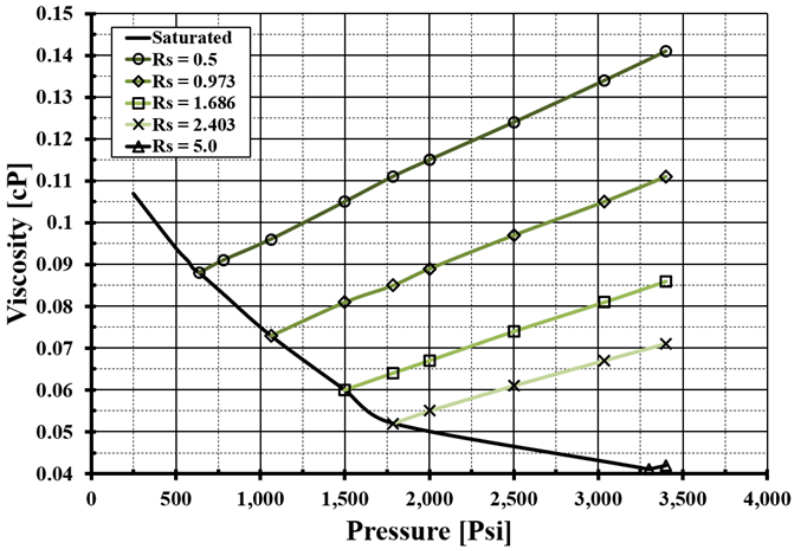


***Figure 2.*** *Effect of different solution gas-oil ratios on viscosity*

In order to have any quantitative result on fluid flow within the cells, the content of the cells is calculated with the usage of saturation functions. Under this expression, a set of water-oil and gas-oil relative permeability curves and a matching water-oil capillary pressure curve are defined. [2] The input is in tabular format, and the simulator will yield the non-existing data ranges later on by using spline interpolation. These datasets have to be treated as one unity; even in artificial model establishment, the saturation function should come from one specific fluid system. The values for relative permeabilities in the water-oil system are used as an example in *Figure 3*.

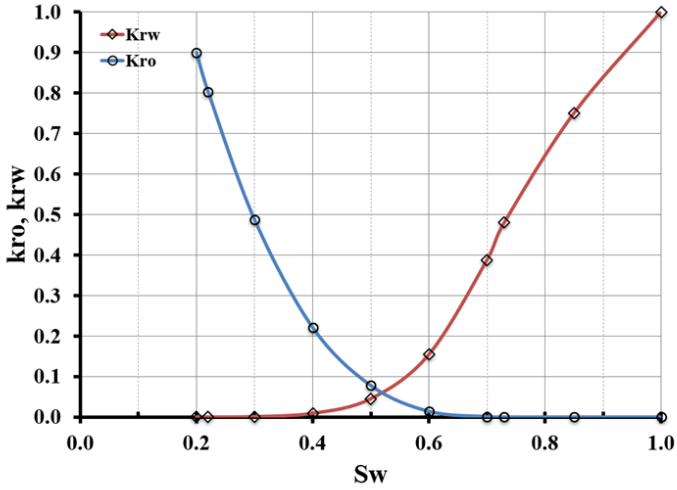


***Figure 3.*** *Water-oil relative permeability data*

### 2.3 Well specifications and production criteria

For a seven-spot pattern type, the layout of the wells in the waterflooded reservoir model has six injector wells at the peripheral corners of the hexagonal shape and a single producer in the middle of the grid. The diameter of the tubing for each one is 0.625 feet, and all near wellbore vicinity zones have a unified skin factor of 7.5. According to completion all wells are open in every cell layer. These values and options may be too idealized, but the aim of the study is to examine recovery factor based on initial rate only. Of course in a real reservoir management scenario there is a big step of optimizing depletion, but in this case that change would also alter the outcome and focus of the study.

The reservoir is depleted using waterflooding for pressure maintenance and displacement towards the production well. In order to keep the liquid rate controlled, a specific liquid rate of 2,000 STB/day primary liquid production rate is set up. To aid displacement, reservoir voidage replacement was used with a multiplier of 0.95 for the injector wells, which were under group control based on the actual liquid outtake of the producer well. To prevent the production of gas and the formation of a secondary gas cap, a flowing bottomhole pressure limit was also set for the producer. Checking the solution GOR of the initial reservoir conditions and matching it with the corresponding bubble point pressure from the PVT datasets, a pressure limit of 1,400 psi was set to ensure that this value is always higher than the actual bubble point, and as since the average reservoir pressure is even higher than that, the formation of a secondary gas cap was prevented.

### 2.4 Time variables, base case characteristics

The size of the timesteps was 30 days, and based on the number of the timesteps (185), total simulation time was around 15 years. The values described in this chapter are considered as the base case values for the sensitivity analysis. Based on this, 35 other production rate-based scenarios yielded the data for regression in the sensitivety analysis. The graphs in *Figure 4* show the main characteristics and behavior of the base case scenario. For this display, flowing bottomhole (BHP) and average reservoir pressures (FPR), oil production rate (FOPR), gas-oil ratio (GOR), watercut (FWCT), and recovery factor (FOE) were selected.

During the depletion period, the reservoir pressure is decreasing mildly due to the controlled voidage replacement. GOR, as a constant, is a good indication of no free gas in the system. Regarding flowing bottomhole pressure, the line moves together with reservoir pressure up the point when the desired oil and liquid rate cannot be managed. The liquid rate is insufficient due to pressure decrease, and the oil rate cannot be held because – as the watercut curve indicates – there is a water breakthrough in roughly the tenth year of production. The recovery factor increases monotonously as long as the oil production rate is constant, but after reaching the extents of this depletion, the steepness is reduced. If the production period were longer, the recovery factor curve would converge to a possible maximum value.

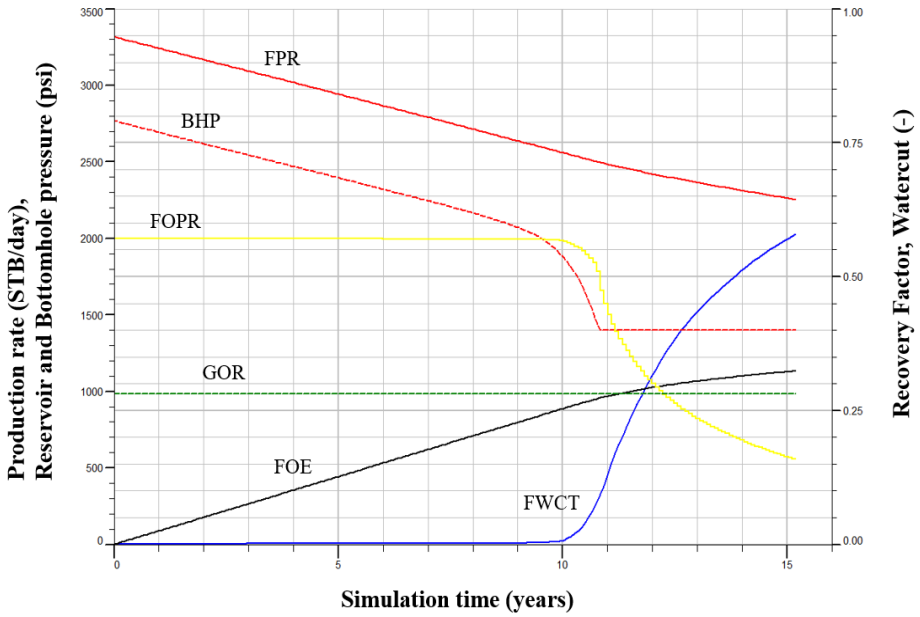


***Figure 4.*** *Main characteristics of the base case scenario*

## 3. SENSITIVITY ANALYSIS AND REGRESSION

After establishing the main model, running and understanding the base-case scenario, the next step of investigation was the sensitivity analysis, which in this study focuses on checking the recovery factor at 5, 10 and 15 years of production if the initial liquid target changes. For this, a total set of 36 scenarios was established. The results of the analysis, the production start rates in stb/day and the corresponding recovery factor values can be found in *Table 2*.

From these, minor adjustments should be made to increase later accuracy of the formula and decrease unwanted error in the regression. At selected points, the predefined initial values were set, though at high preferred initial values, the reservoir cannot even bear that rate even at the start. That means that after this critical rate, basically only the numbers were changed in the scenarios, but the outcome was the same. A critical rate of 7,300 STB/day was found to be the first rate that cannot be sustained by the reservoir from the start date. For the sake of objectivity, in *Table 2*, all the results are included, but the values shown in italics were removed due to these reasons. In other words, if for regression the rate is changing but the outcome is the same, that will fail any attempt to find correlation at that part of the data, and will generate significant error in the fit, as the regression will try to use best fit, and shift the fitted curve towards these false datasets, which is undesirable.

***Table 2***

*Recovery factors based on initial rate after 5, 10 and 15 years of depletion*

| Rate | RF 5 | RF 10 | RF 15 |
|---|---|---|---|
| 100 | 0.0064 | 0.0127 | 0.0191 |
| 200 | 0.0127 | 0.0254 | 0.0381 |
| 300 | 0.0191 | 0.0381 | 0.0572 |
| 400 | 0.0254 | 0.0508 | 0.0762 |
| 500 | 0.0318 | 0.0635 | 0.0953 |
| 600 | 0.0381 | 0.0762 | 0.1143 |
| 700 | 0.0445 | 0.0889 | 0.1334 |
| 800 | 0.0508 | 0.1016 | 0.1524 |
| 900 | 0.0572 | 0.1143 | 0.1715 |
| 1000 | 0.0635 | 0.1270 | 0.1905 |
| 1200 | 0.0762 | 0.1524 | 0.2286 |
| 1400 | 0.0889 | 0.1778 | 0.2661 |
| 1600 | 0.1016 | 0.2032 | 0.2959 |
| 1800 | 0.1143 | 0.2286 | 0.3126 |
| 2000 | 0.1270 | 0.2539 | 0.3226 |
| 2200 | 0.1397 | 0.2771 | 0.3294 |
| 2400 | 0.1524 | 0.2908 | 0.3346 |
| 2600 | 0.1651 | 0.2998 | 0.3386 |
| 2800 | 0.1778 | 0.3063 | 0.3418 |
| 3000 | 0.1905 | 0.3113 | 0.3444 |
| 3200 | 0.2032 | 0.3152 | 0.3465 |
| 3400 | 0.2159 | 0.3183 | 0.3483 |
| 3600 | 0.2286 | 0.3209 | 0.3498 |
| 3800 | 0.2413 | 0.3231 | 0.3511 |
| 4000 | 0.2537 | 0.3249 | *0.3522* |
| 4500 | 0.2720 | 0.3284 | *0.3544* |
| 5000 | 0.2800 | 0.3308 | *0.3558* |
| *5500* | *0.2844* | *0.3323* | *0.3568* |
| *6000* | *0.2869* | *0.3331* | *0.3573* |
| *7000* | *0.2888* | *0.3338* | *0.3577* |
| *8000* | *0.2888* | *0.3338* | *0.3577* |
| *10000* | *0.2888* | *0.3338* | *0.3577* |
| *12000* | *0.2888* | *0.3338* | *0.3577* |
| *16000* | *0.2888* | *0.3338* | *0.3577* |
| *20000* | *0.2888* | *0.3338* | *0.3577* |

The next adjustment is due to the shape, in other word the mathematical structure of the recovery factor curve itself. As discussed previously, after the breakthrough, or after several years of high depletion rate, the curve itself will tend towards the possible maximum value, and become flat afterwards. This, as generally having the same effect as being above the critical rate, will also affect the fit. In order to eliminate this, the datasets were only included to the point where there is at least 0.1–0.5% change in the recovery factor between the two neighboring scenarios. These are also in italics in the table. For 5 and 10 years, the last scenario is 5,000 stb/day, and for 15 years, due to being at the maximum for a long period, is 3,800 stb/day. The recovery factor curves are shown in *Figure 5*.

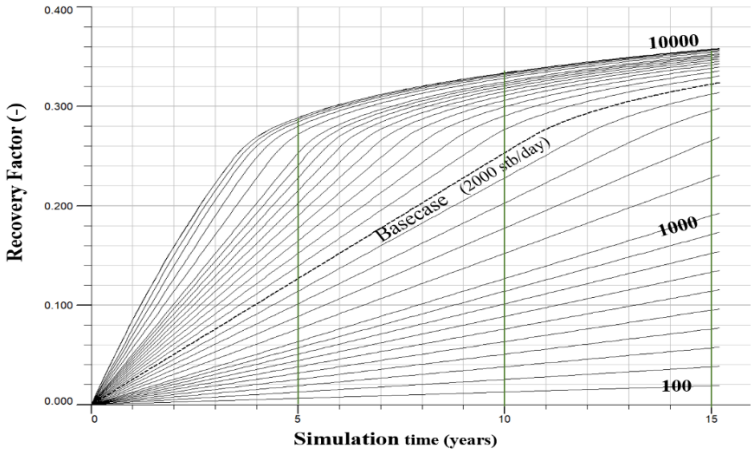


***Figure 5.*** *Recovery factor curves of the sensitivity analysis*

Regression was applied after the datasets were chosen. Based on the behavior of the Recovery factor curve, the idea to split the time period into three discrete formulas was described in [8]. To sum it up, as the recovery factor curve has very distinct parts, the mathematical functions to fit those parts are specific, too. Therefore, having one formula for the entire timespan will not lead to the desired accuracy. In addition, companies have production plans for discrete time periods, so the idea can still be applied when the desired timespan is known. As a result, three distinct datasets were fitted with functions. In previous work examining water breakthrough, the logarithmic fit proved to be successful, but in a form of a log scale straight line equation [8]. In this study the resulting best fit after the ln-ln plot was a third-degree polynomial. The maximum allowed difference in estimated and simulated recovery factor was set to 1.5%. The plotted datasets with the fitted curves are displayed in *Figure 6.*

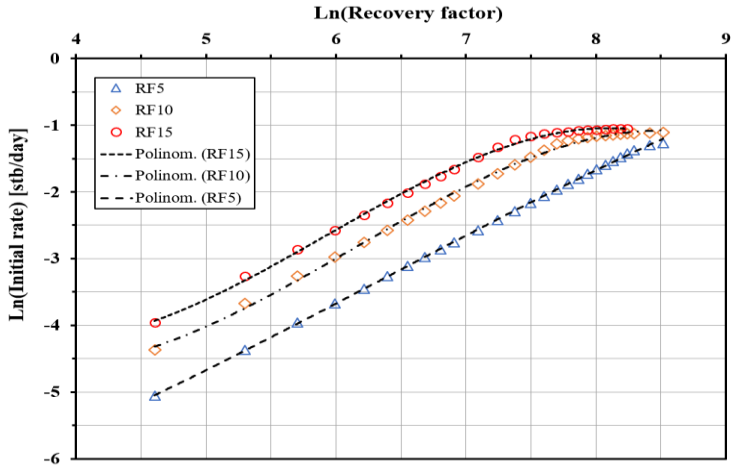


***Figure 6.*** *Dataset and fitted functions for the recovery factors*

The parametric form of the fitted function is described in *Equation 1*, and the coefficients for the three regressions are listed in *Table 3*.

$$ln[RF] = a_1[ln(lrate)]^3 + a_2[ln(lrate)]^2 + a_3[ln(lrate)] + c \quad (1)$$

***Table 3***

*Coefficients, $R^2$ values and average differences of the fitted functions*

| Timespan | a1 | a2 | a3 | c | $R^2$ | avg. difference in Rf |
|---|---|---|---|---|---|---|
| Rf 5 year | –0.0098 | 0.1855 | –0.1535 | –7.3219 | 0.9997 | –0.26% |
| Rf 10 year | –0.0667 | 1.2285 | –6.4207 | 5.7121 | 0.9980 | +0.14% |
| Rf 15 year | –0.0793 | 1.4124 | –7.2716 | 7.3504 | 0.9986 | +0.32% |

Near the end of the three periods, at the highest rates for each, the recovery of the reservoir model approaches the possible maximum of the system. If we take only the linear part on the ln-ln fit, the accuracy is better, but of course the validity range is smaller. Because the correlation between the initial liquid rate and the recovery was changing shape at the end of the 5-, 10- and 15-year depletion periods, a polynomial fit seemed to be better, with controlled degree. The end result is a third-degree polynomial with the right curvature at the end of the datasets. In terms of accuracy, there is a strictly linear part, which will have some errors when approached with a polynomial. The concept of accuracy was to define the time period and give a solution with the lowest average error [7]. If the production time is even longer, or the breakthrough is postponed, then a further split of the fit is suggested, to have both a linear part and a polynomial part of the fitted curve. Another possible approach for the short-term forecast to have high accuracy is to only take the linear part into consideration when applying the fit. The error of the formulae is displayed in *Figure 7*.

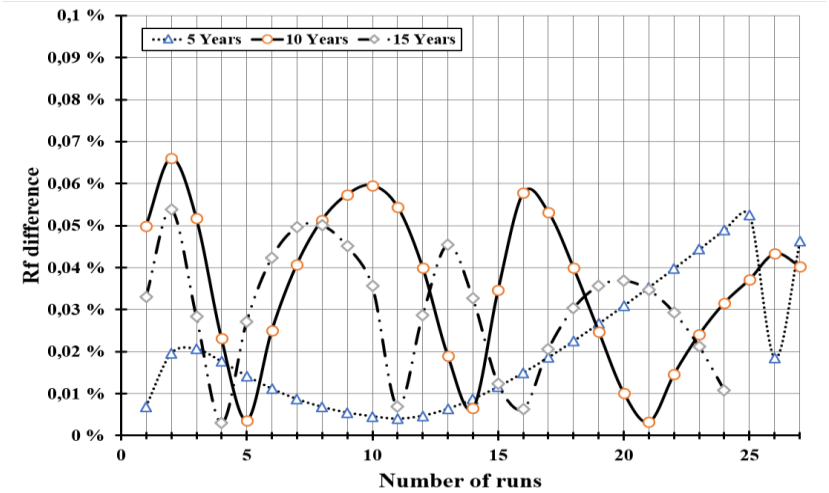


***Figure 7.*** *Accuracy of the fitted functions*

## 4. RESULTS AND CONCLUSIONS

The aim of the study was to generate a proxy model – in other words, analytical formulae – in order to estimate the recovery of an artificial reservoir model. In this article the examined reservoir was an undersaturated oil reservoir depleted by water-flooding, using a seven-spot well pattern with peripheral water injection and central oil production. The produced fluid was set to be only oil and water with the introduction of a minimum flowing bottomhole pressure limit in order to keep the entire flowing system in the reservoir and at the sandface single phase. For this purpose, first the modeling grid and reservoir properties were implemented in simulation software, then the fluid and well properties were added, to create a base-case scenario. After creating a stable base case, sensitivity analysis was conducted to gain datasets of recovery factors of 5, 10 and 15 years of production, respectively, as a function of the initial liquid production rate stated in the datafile. The results in order of increasing recovery factor can be found in *Figure 8*.

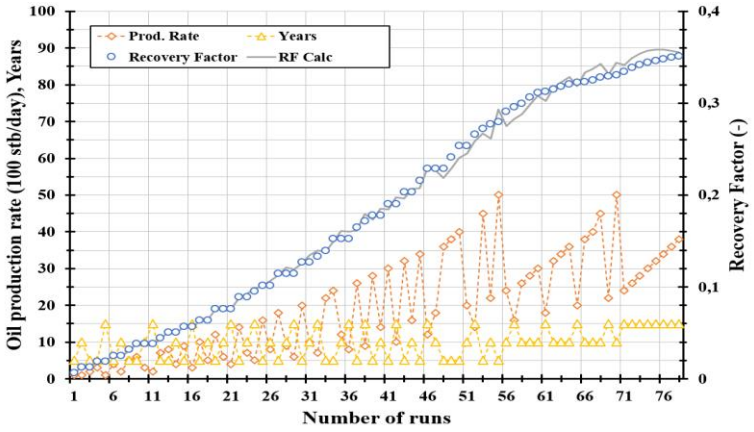


***Figure 8.*** *Results of the calculation with increasing recovery factor*

After the removal of duplicate and false datasets, a regression was performed on the remaining datasets in each case, resulting in third-degree polynomial formulae that are capable of estimating the recovery factor with a maximum error margin of 0.07% (in terms of RF difference). The results are displayed in *Equations (2)–(4).*

$$ln[RF\,5] = -0.0098[ln(lrate)]^3 + 0.1855[ln(lrate)]^2 - 0.1535[ln(lrate)] - 7.3219 \quad (2)$$

$$ln[RF\,10] = -0.0667[ln(lrate)]^3 + 1.2285[ln(lrate)]^2 - 6.4207[ln(lrate)] - 5.7121 \quad (3)$$

$$ln[RF\,15] = -0.0793[ln(lrate)]^3 + 1.4124[ln(lrate)]^2 - 7.2716[ln(lrate)] - 7.3504 \quad (4)$$

For further application, these types of formulae can also be used to generate charts of the solution, and it is also possible to establish a three-dimensional response surface, such as that seen in *Figure 9*. Note that in this particular case, hence the three

distinct equations, the points between the functions in the figure are results of simple linear interpolation for display purposes only.

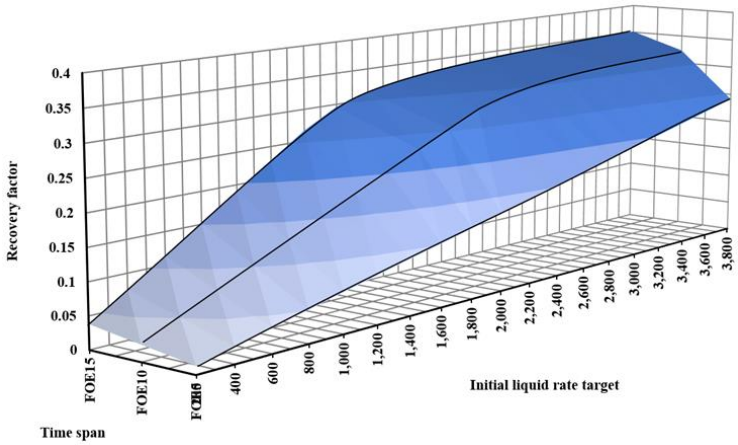


***Figure 9.*** *3D surface generated from the formulae*

## REFERENCES

[1] Gokhale, S.S. (1998). *Analytical Modeling, Encyclopedia of Distributed Systems*. Kluwer Academic Publishers, Amsterdam.

[2] Papay, J. (2003). *Development of Petroleum Reservoirs.* Budapest, Akadémiai Kiadó.

[3] Schlumberger Information Solutions (2008). *ECLIPSE Blackoil Reservoir Simulation*. Houston, Schlumberger.

[4] Schlumberger (2012). *ECLIPSE Reference Manual.* Houston, Schlumberger.

[5] Robinson, S. (2014). *Simulation: The Practice of Model Development and Use*. 2nd ed., London, Palgrave Mcmillan.

[6] Bánki D., Jobbik A: (2016). Dynamic modeling of hydrocarbon reservoirs using finite difference methods and derivation of analytic formulae for easier forecasting, *Proc. ISCAME Conference*, Debrecen, October 13–15, 2016.

[7] Bánki, D., Turzó, Z. (2020). Aiding Reservoir Simulation and Maintenance Using analytical Formulae, *Műszaki Földtudományi Közlemények*, Miskolc, Volume 89, pp. 207–215.

[8] Bánki, D., Turzó, Z. (2020). Application of analytical formulae in recovery factor estimation of hydrocarbon reservoirs. *Proc. XX. Ph.D. Conference proceedings*, Budapest PEME Egyesület, November 12, 2020.