Impact of Grid Refinement on a Non-Conventional Mesh Structure: A Case Study with INSIM-FT.

Dimary Moreno López*, João Teixeira *, Sinesio Pesco*, Abelardo B. B. Junior *,
José R. P. Rodrigues†, Regina Alves †, Rodrigo Gusmão Cavalcante † and Malú Grave‡.
* Pontifícia Universidade Católica do Rio de Janeiro, BR.
Email: dimarydelcmoreno@gmail.com, jopetesa@aluno.puc-rio.br, sinesio@puc-rio.br, abelardo-puc@puc-rio.br
†PETROBRAS R&D Center (CENPES), Rio de Janeiro, BR.
Email: jrprodrigues@petrobras.com.br, regina.costa@petrobras.com.br, rodrigogusmao@petrobras.com.br
‡ Universidade Federal Fluminense, Rio de Janeiro, BR.
Email: malugrave@id.uff.br

***Abstract*—To support decision-making, the oil and gas industry relies on flow simulators to evaluate multiple production scenarios. This demands simulation tools that are both accurate and computationally efficient. INSIM-FT *(Interwell Numeric Simulation Model with Front-Tracking)* is a hybrid flow simulator that combines physical modeling with data-driven components, using a non-conventional mesh built through Mitchell's Best Candidate algorithm and Delaunay Triangulation. Although finer meshes are expected to improve spatial representation, they also increase computational costs and may interfere with the performance of data-driven elements. This trade-off highlights the need to assess the influence of the number of imaginary nodes used in INSIM-FT. This work investigates how mesh configuration impacts the accuracy and performance of the simulator through a case study of a heterogeneous reservoir. Considering randomly generated models whose parameters are adjusted by the data-driven component, the results show that mesh refinement affects the spread of the adjusted prior models around the mean. Nevertheless, the mean of the simulations remains reasonably close to the observed data in all cases.**

I. INTRODUCTION

The exploration of oil and gas reservoirs is a highly relevant economic activity, whose financial return is directly influenced by the strategies adopted. As highlighted by Osunrinde *et. al.* in [1], well-informed strategic decisions can significantly enhance the economic value of a reservoir. Such decisions depend on evaluating multiple production scenarios, typically simulated using flow models, as discussed by Fraçois *et. al.* [2].

Conventional flow simulators usually discretize the entire reservoir using a large-scale grid, as exemplified by the models presented in [3], [4], [5], and [6]. While these models are generally considered accurate, they tend to be computationally expensive and demand detailed geological and fluid property data across different reservoir regions. Acquiring such data can be both costly and time-consuming, often requiring extensive laboratory testing.

The inherent challenges of conventional flow simulators, such as those mentioned above, have driven the development of more computationally efficient approaches based on more readily available data, such as injection and production rates, as explored by Heffer *et. al.* [7], Jansen and Kelkar [8], and Refunjol [9]. Among these approaches, models based on statistical correlations are noteworthy. Despite their efficiency, they may suffer from limited accuracy.

INSIM *(Interwell Numeric Simulation Model)*, proposed by Zhao *et. al.* [10] [11], was introduced as a hybrid alternative, aiming to balance accuracy and computational efficiency. It combines data-driven elements with physical modeling principles. INSIM discretizes the reservoir based on direct well-to-well connections, assuming one-dimensional flow between them. Saturation in these connections is determined using the Buckley-Leverett equation [12].

Several extensions of the INSIM framework have been proposed, including INSIM-FPT (INSIM Flow-Path Tracking) [13]. This method employs a Depth-First Search (DFS) algorithm to identify the main flow paths and define the mesh edges between injectors and producers. It simulates two-phase flow in 3D reservoirs. Another related development is INSIM-BHP, presented by Li and Onur [14], which incorporates bottom-hole pressure (BHP) data in addition to production rates for history matching.

Among the enhanced versions of INSIM, INSIM-FT, proposed by Guo in 2018 [15], [16], stands out. In this formulation, saturation profiles are computed using the Front-Tracking method. Furthermore, the method was extended to account for gravitational effects. One of the most significant innovations introduced in this version is the implementation of imaginary nodes in the computational mesh, a central focus of the present discussion. These nodes were conceived to provide a more flexible representation of reservoir heterogeneities and to increase the number of possible flow paths between wells, as outlined in the original formulation.

However, due to the hybrid nature of INSIM-FT, in which mesh refinement leads to an increase in the number of parameters to be calibrated, refinement not only raises computational costs, as expected for any simulator, but makes parameter tuning more challenging. Therefore, a detailed investigation of the mesh-related hyperparameters, particularly the number of imaginary nodes, is essential to ensure efficient and reliable simulations.

This study investigates the sensitivity of the INSIM-FT simulator to mesh refinement. As a foundation for this analysis, we first present the key features and underlying structure of the simulator.

II. INSIM-FT

The Interwell Numerical Simulation Model with Front-Tracking (INSIM-FT) is a flow simulator developed to describe the waterflooding process in oil reservoirs. As detailed by Zhao *et. al.* [10], Guo and Reynolds [15], [16], it is a hybrid model that combines physical principles with data-driven components, aiming to deliver a computationally efficient solution without compromising accuracy. In this section, we present the main features of the simulator based on the descriptions provided in the aforementioned references.

To define the simulator, a computational mesh is first constructed to discretize the study area within the reservoir. In order to adequately represent the physical characteristics of the porous medium, several parameters must be assigned to the mesh's nodes and edges. These parameters can be estimated through a history matching process, which constitutes the data-driven component of the model. This process calibrates the parameters so that the simulator can accurately reproduce the observed data.

At each time step, the simulator performs two sequential stages. In the first stage, the mass balance equation is discretized over the mesh to obtain an estimate of the pressure at mesh nodes. This pressure field is then used in the second stage, where saturation is computed using a Front-Tracking procedure. Finally, the simulator can also be employed as a tool for production optimization.

A. Mesh

In INSIM-FT, the reservoir is represented by a set of control points (nodes), between which connective flow volumes are defined. Each control point is assigned a fixed volume that remains constant over time, as illustrated in Figure 1 by the red dashed circles.

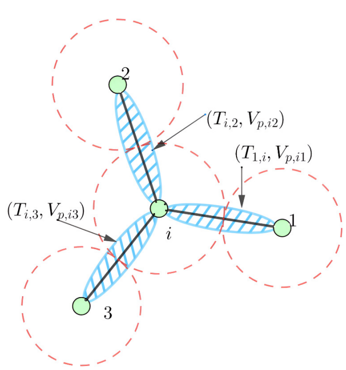

Fig. 1. Connective units between wells of INSIM-FT.

Associated with each control volume, the pore volume of each cell is defined and denoted by $V_{p,i}$ for each node i. This quantity may vary over time as a function of pressure and the compressibility of the porous formation. Similarly, for the connections between nodes i and j, an associated pore volume $V_{p,ij}$ is defined, along with the transmissibility T_{ij}. It is important to note that, as illustrated in Figure 1, not all nodes are directly connected to each other.

Initially, nodes are placed in the mesh at locations corresponding to well completions. To enhance the representation of other reservoir regions, particularly where property variations may occur, additional nodes and connections can be introduced. These extra nodes, referred to as imaginary nodes, are incorporated to create additional flow paths between injectors and producers.

Regarding their placement, the mesh generation process employs Mitchell's Best Candidate Algorithm to determine the locations of these additional nodes, combined with a modified Delaunay Triangulation Algorithm to establish the connections between them.

Mitchell's Best Candidate Algorithm [17] is a sequential sampling technique widely used to generate uniformly distributed points. In this context, the set C of mesh nodes is initially defined as the set of well completions, which constitutes the mesh in the traditional INSIM formulation. Then, the number of imaginary nodes, N_I, to be added to the mesh is specified. At each step of the algorithm, a fixed number of random candidate points are generated within a user-defined domain representing the physical extent of the reservoir. Among these candidates, the point that maximizes the distance to the current set C is selected and added to the mesh. The set C is then updated, and the process is repeated N_I times until all imaginary nodes are generated. The optimal selection of the number of imaginary nodes, N_I, depends on a more detailed analysis of the node configuration and is a subject of investigation in this work. Nevertheless, a common heuristic in INSIM-FT is to set N_I to be one to two times the number of nodes associated with well completions [16].

Subsequently, the Delaunay Triangulation Algorithm [18] can be used to define the possible connections between nodes. In INSIM-FT, a `MATLAB` implementation of this algorithm is employed to determine connectivity. In this procedure, connections are established so that the circumcircle of any triangle formed by three connected nodes contains no other node within it. This criterion tends to prevent the formation of triangles with extremely small or large angles; in particular, the smallest angle in the triangulation is guaranteed to be at least as large as that of any other possible triangulation [18]. Moreover, when nodes are uniformly distributed, as in meshes generated using Mitchell's Best Candidate Algorithm, the resulting triangulation naturally avoids connections between distant nodes.

Within the set of candidate edges generated by the triangulation algorithm, a post-processing procedure is required to eliminate edges that do not provide meaningful contributions to the model's accuracy or physical representativity. Specifically, edges that form triangles with an opposing angle exceeding 120°are systematically remove, as described by Guo

[16]. Additionally, to ensure the physical consistency of the simulation, edges connecting nodes classified both as injectors or both as producers are discarded, as well as edges linking nodes associated with completions of the same well.

In the two-dimensional case, simpler algorithms can be employed, such as the one proposed by Kaviani *et. al.* [19], which was used in the original version of INSIM. This algorithm establishes connections between nodes that lie within a fixed distance. In this context, a post-processing step is also performed, where connections are removed if the opposing angle in the formed triangle is less than 15°, as well as connections between the extreme nodes when three nodes are nearly collinear. This filtering is applied because direct interaction between such nodes is expected to be insignificant. An illustration of this scenario can be seen in Figure 2, where the connection between nodes $P1$ and $P3$ would be removed.

Fig. 2. Example of three nearly colinear nodes in the INSIM mesh. In this configuration, the edge connecting the outer nodes may be removed during post-processing to simplify the network structure.

Finally, since there is no geological model associated with the mesh, information regarding volumes and areas, typically present in simulation meshes, is unavailable. Similarly, spatially distributed data on porosity and permeability, which are essential for determining pore volumes and associated transmissibilities, are lacking. Therefore, these missing properties are inferred through a history matching process.

B. *History Matching*

History matching is based on a constrained optimization process of a function that depends on the parameters to be estimated and the observed data, which are assumed to be known, as discussed by Ren *et. al.* [20]. This step constitutes the data-driven component of the model.

The parameters to be estimated include the pore volumes of the connections, the transmissibilities, and the parameters that define the relative permeability curves for each edge, such as the Corey exponents for each phase and the relative permeability of water at maximum saturation. In particular, these relative permeability curves can be assumed to differ across edges, significantly increasing the number of parameters to be estimated in models with more complex meshes. While these choices enhance the representativeness of the artificial model to the real reservoir, the increase in parameter count can substantially complicate the history matching process. Therefore, determining an appropriate level of mesh complexity becomes a critical task.

It is worth noting that certain physical constraints must be enforced to ensure model consistency. These include the non-negativity of transmissibilities and pore volumes, as well as the requirement that the total reservoir volume equals the sum of the connection volumes, among other conditions.

Given the high dimensionality of the parameter space, the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) method is employed for optimization, following the formulation proposed by Emerick *et. al.* [21], [22], [23], [24].

All parameters are defined with the purpose of constructing a simulator capable of reproducing known production estimates over a specific time period, ultimately enabling reliable production forecasts. In the following section, we describe how the simulator is developed and highlight its main features.

C. *The Simulator*

In INSIM-FT, a sequential two-step approach is used to determine the flow at each time step. In the first stage, pressure is estimated by solving the mass balance equation defined at each node of the mesh. Subsequently, in the second stage, an estimate of saturation along each edge is computed, based on both the data from the previous time step and the pressure estimates obtained in the first stage of the simulator. The saturation calculation is performed independently on each edge, where the flow is assumed to be one-dimensional.

Saturation on an edge is defined by the water mass balance equation, with the initial condition given by the saturation profile at the previous time step. This is, in general, a Cauchy problem with a non-uniform initial condition. To solve it, the initial condition can be approximated by a step function with a finite number of discontinuities, as illustrated in Figure 3.

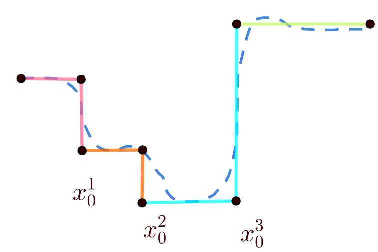

Fig. 3. Piecewise constant approximation of a continuous function.

The solution of this problem is determined using the *Front-Tracking* method, a strategy explored by Dafermos [25] and Langseth *et. al.* [26]. To this end, the Cauchy problem is divided into several Riemann subproblems, each defined at a discontinuity point. These subproblems are solved using the convex hull method.

In INSIM-FT, this is performed using the *Graham's Scan* algorithm , discussed by De Berg [27]. The convex hull method consists in approximating the fractional flow function f by a piecewise linear function, $f_{w,pl}(S)$. Then, the convex envelope of the graph of $f_{w,pl}(S)$ is determined. It can be shown that the solution is piecewise constant, and the saturation values are separated in the xt-plane by straight lines with slopes related to the local linear approximation of f (including the convex hull in the corresponding intervals). Therefore, the saturation solutions near a discontinuity point x_0 are as illustrated in Figure 4. This particular structure of the solution gives rise to the term *"Riemann fans"*, which is commonly used to refer to them.

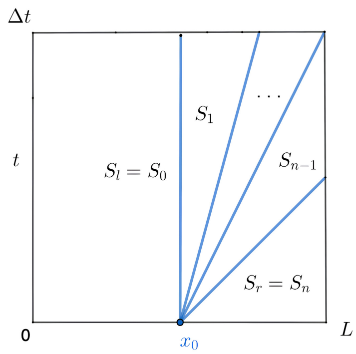

Fig. 4. Solution of the Riemann problem: Riemann fan

Once the solutions of the Riemann subproblems have been determined, the solution of the initial value problem, whose initial condition is a step function, is obtained by combining the subproblem solutions at the initial time, as illustrated in Figure 5.

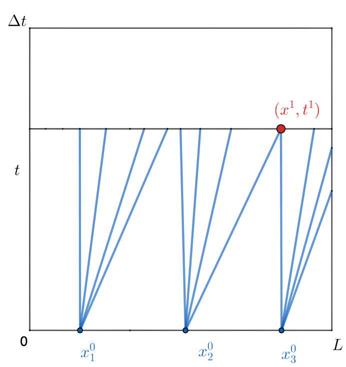

Fig. 5. Solution of the Cauchy problem by combining Riemann subproblem solutions

The solutions are well defined up to the point where saturation values collide. At the first time such a collision occurs, a new Riemann problem is generated, which determines how the solution will propagate from that point onward (see Figure 6). If another collision occurs, the procedure must be repeated until the next collision takes place outside the time domain under consideration.

Once the saturation profile has been determined along each edge, it is possible to compute the saturation at the mesh nodes. For nodes that belong to a single edge, the saturation is directly given by the saturation on that edge. For nodes that are connected to multiple edges, the estimated saturation

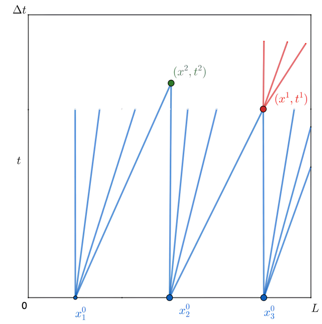

Fig. 6. Collision of the solutions of the Riemann subproblems.

from each edge may differ. Therefore, a reconciling step is required. In this step, the node saturation is estimated based on the inflow from each of the edges to which the node is connected. Further details are discussed by Guo [16].

III. Example applications

To investigate the influence of the number of imaginary nodes on the ability of INSIM-FT to model flow in heterogeneous reservoirs, we consider the "channel" scenario, which represents a reservoir composed of two connected regions with contrasting properties. Similar models have been widely studied in the literature due to their representative features , as in the works of Emerick [28], Grave *et. al.* [29], and Zhao [30].

The reservoir heterogeneity is clearly defined by the presence of two distinct regions. Figure 7 shows a grid discretization of the domain. Inside the channel, the permeability is set to $5000\,mD$, while outside the channel it is reduced to $500\,mD$.

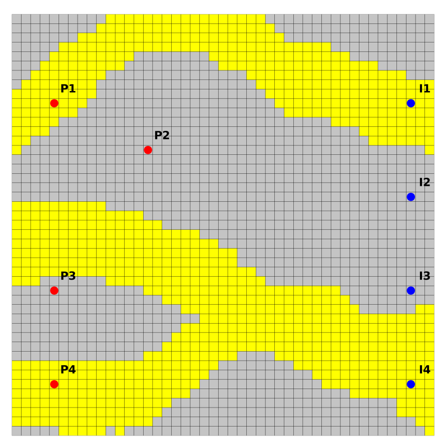

Fig. 7. Channelized reservoir in IMEX.

The rock compressibility is estimated as $2.17551 \times 10^{-5}\,1/Pa$, and the porosity is assumed to be 20%. The irreducible water saturation is set to 0.25%, and the residual oil saturation to 0.15%. Water has a compressibility of $3.77905 \times 10^{-5}\,1/(kgf/cm^2)$, viscosity of $0.337\,cp$ an a density of $1024.77kg/m^3$, while oil has a compressibility of $3.57 \times 10^{-4} 1/(kgf/cm^2)$, a higher viscosity of $1.58\,cp$ and a lower density of $833.33\,kg/m^3$.

The simulation setup includes four injection wells and four production wells, with half of each group located inside the high-permeability channel.

As a reference, we use production data generated by CMG IMEX for a reservoir discretized with a 45×45 grid of blocks, each measuring $20\,m \times 20\,m \times 20\,m$, as shown in Figure 7. The simulation spans 11 years, during which all injectors operate at a constant rate of $156\,m^3/day$ while each producer withdraws $150\,m^3/day$.

From the 11 years of IMEX simulation, the first 5 years of production data are used as observed data for history matching in INSIM-FT. The remaining 6 years are used to assess the predictive performance of the calibrated models.

In each case, 100 prior models are generated with randomly sampled parameters. The relative permeability curves are assumed to be the same across all edges. The water relative permeability at maximum water saturation is drawn from a normal distribution with mean of 0.3 and a standard deviation of 0.5. The Corey exponents for oil and water follow normal distributions with means of 3.0 and 1.0, respectively, and a standard deviation of 0.1. Transmissibilities are sampled from a normal distribution with 20% standard deviation, and spatial correlation is introduced via a spherical variogram with correlation lengths $L_x = L_y = L_z = 20\,m$ and azimuthal angle $\theta = 0$. The mean permeability used to calculate transmissibility is assumed to be $2500\,mD$. The parameter β_0, which controls the allocation of total pore volume to each edge, is sampled from a distribution with mean proportional to the transmissibility of the edge and a standard deviation of 10%. The history matching process is performed using five ES-MDA assimilation steps, which is a common choice for problems of this type.

To evaluate the influence of the number of imaginary nodes in the mesh, we consider configurations with 0, 10, 50, 100, 150, 200, and 250 imaginary nodes, in addition to the nodes corresponding to well completions. As comparative metrics, we use the mean squared error (MSE) and the mean absolute percentage error (MAPE) of oil production, computed for each individual well as well as for the total oil production. These metrics are widely adopted in reservoir engineering and data assimilation studies for evaluating model performance and forecast accuracy [31].

Tables I and III present the estimated errors between the average production of the posterior models and the reference production data obtained from IMEX, using the MSE and MAPE metrics, respectively.

To further assess model performance beyond average behavior, we select, for each mesh configuration, the single posterior model that achieved the best match to the observed data during the training period and compare its performance in the forecasting phase. The error metrics for this latter case are presented in in Tables II and IV.

The impact of the number of imaginary nodes on oil production forecast accuracy was inconclusive, both when considering the average production estimates from all posterior models and when focusing on the single model that best matched the observed data during the history matching period.

In particular, the estimated total oil production was notably accurate across all models. Figures 9, 10, 11 and 12 show the estimated total oil production for meshes with 0, 10, 50, and 100 imaginary nodes, respectively. More refined meshes do not exhibit any significant difference compared to the model with 100 imaginary nodes.

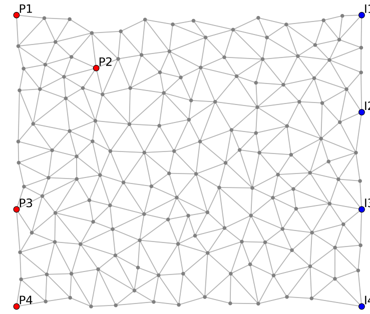

Fig. 8. INSIM-FT grid with 150 imaginary nodes

TABLE I
MEAN SQUARED ERROR (MSE) BETWEEN THE AVERAGE OIL PRODUCTION PREDICTED BY THE POSTERIOR MODELS AND THE REFERENCE PRODUCTION FROM IMEX.

Well/nIm	P1	P2	P3	P4	Total
0	9.81	41.64	7.67	2.66	33.42
10	4.23	19.91	3.47	1.29	16.63
50	12.48	29.01	10.84	3.85	33.67
100	10.50	29.46	7.52	1.97	50.26
150	9.8	38.03	4.85	8.21	37.09
200	9.41	42.04	12.85	9.3	53.6
250	6.54	71.27	6.53	10.77	68.42

TABLE II
MEAN SQUARED ERROR (MSE) OF THE POSTERIOR MODEL THAT ACHIEVED THE BEST FIT TO THE OBSERVED PRODUCTION DATA DURING THE HISTORY MATCHING PERIOD.

Well/nIm	P1	P2	P3	P4	Total
0	10.27	46.96	7.13	3.4	42.36
10	7.25	20.05	8.7	7.29	69.93
50	6.67	46.52	19.91	4.67	104.14
100	4.31	37.84	7.84	4.26	59.54
150	12.10	10.38	9.13	5.83	21.65
200	11.96	29.35	6.91	3.77	44.92
250	14.20	104.79	15.48	3.44	103.64

TABLE III
MEAN ABSOLUTE PERCENTAGE ERROR (MAPE) BETWEEN THE AVERAGE OIL PRODUCTION PREDICTED BY THE POSTERIOR MODELS AND THE REFERENCE PRODUCTION FROM IMEX.

Well/nIm	P1	P2	P3	P4	Total
0	9.25	14.39	6.39	3.41	3.66
10	3.13	6.00	4.30	3.37	1.88
50	8.00	8.66	9.13	3.43	2.55
100	5.44	11.46	7.58	1.13	4.75
150	7.18	12.0	5.32	3.74	3.53
200	7.04	12.55	9.71	5.08	5.01
250	5.13	16.67	7.6	3.44	5.98

TABLE IV
MEAN ABSOLUTE PERCENTAGE ERROR (MAPE) OF THE POSTERIOR MODEL THAT ACHIEVED THE BEST FIT TO THE OBSERVED PRODUCTION DATA DURING THE HISTORY MATCHING PERIOD.

Well/nIm	P1	P2	P3	P4	Total
0	9.11	15.13	6.38	4.11	3.94
10	3.20	9.19	7.33	2.76	4.64
50	5.57	15.09	12.58	2.64	6.87
100	4.8	13.61	7.36	2.07	5.08
150	10.16	6.64	5.21	6.5	1.18
200	7.41	11.28	6.78	2.24	3.81
250	10.92	22.36	11.56	2.62	7.3

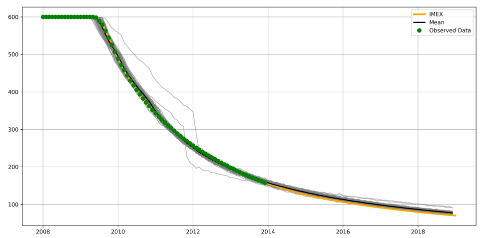

Fig. 9. Comparison of simulated and observed total oil production over time using a mesh with 0 imaginary nodes.

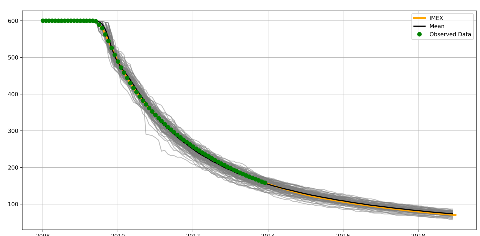

Fig. 10. Comparison of simulated and observed total oil production over time using a mesh with 10 imaginary nodes.

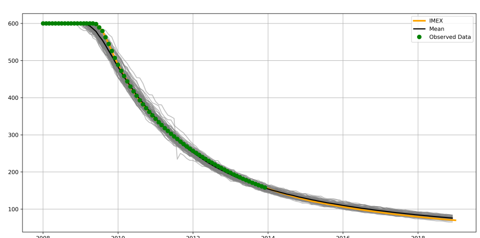

Fig. 11. Comparison of simulated and observed total oil production over time using a mesh with 50 imaginary nodes.

Despite the overall good performance in total production forecasts, individual well matching remains more challenging, mainly due to their relative positions within the reservoir, as evidenced by this case study. Producer well 2 posed the greatest challenge in terms of history matching accuracy. Figures 13, 14, 15 and 16 present the production estimates for this well using meshes with 0, 10, 50, and 200 imaginary nodes.

Results for producer well 1 using meshes with 0, 10, 50, and 100 imaginary nodes are shown in Figures 17, 18, 19 and 20. The results for producer wells 3 and 4 exhibit a similar

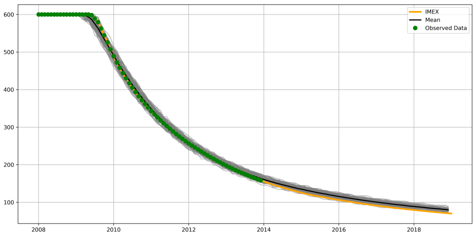

Fig. 12. Comparison of simulated and observed total oil production over time using a mesh with 100 imaginary nodes.

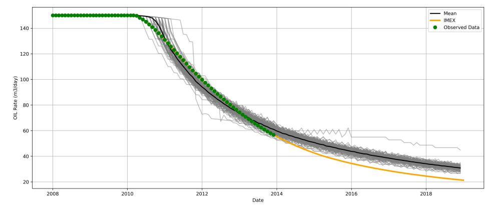

Fig. 13. Simulated vs Observed Data Over Time for P2. Mesh with 0 Imaginary Nodes.

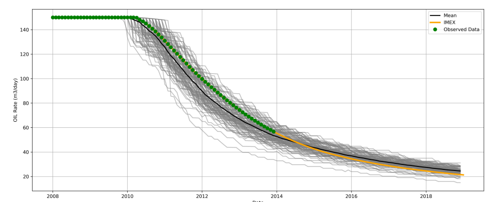

Fig. 14. Simulated vs Observed Data Over Time for P2. Mesh with 10 Imaginary Nodes.

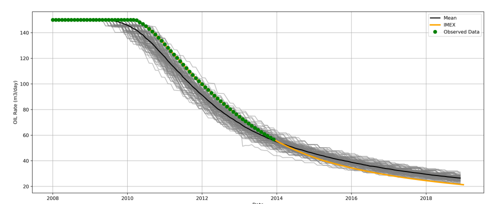

Fig. 15. Simulated vs Observed Data Over Time for P2. Mesh with 50 Imaginary Nodes.

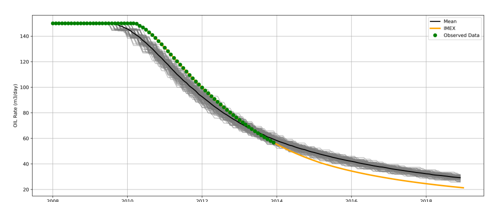

Fig. 16. Simulated vs Observed Data Over Time for P2. Mesh with 200 Imaginary Nodes.

overall behavior to that of well 1, with a good match to the observed data.

For all cases analyzed, models with more than 100

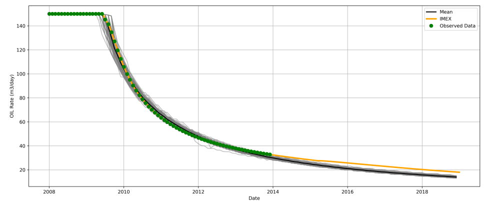

Fig. 17. Simulated vs Observed Data Over Time for P1. Mesh with 0 Imaginary Nodes.

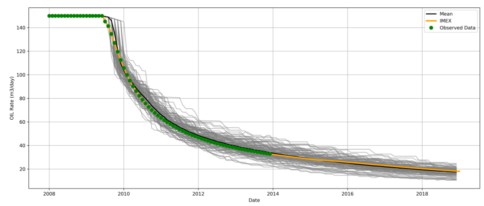

Fig. 18. Simulated vs Observed Data Over Time for P1. Mesh with 10 Imaginary Nodes.

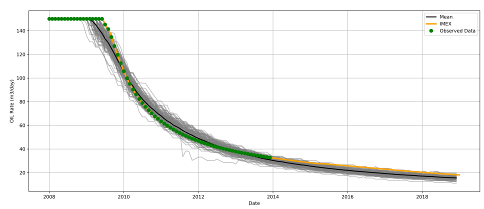

Fig. 19. Simulated vs Observed Data Over Time for P1. Mesh with 50 Imaginary Nodes.

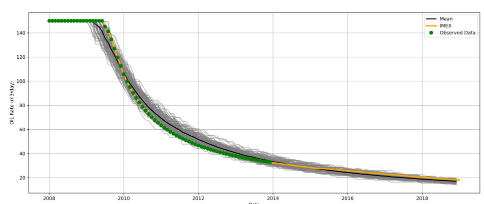

Fig. 20. Simulated vs Observed Data Over Time for P1. Mesh with 100 Imaginary Nodes.

imaginary nodes exhibited behavior similar to the 100-node configuration. As such, the corresponding plots are omitted for brevity.

Although it was not possible to establish a clear relationship between model accuracy and mesh refinement, certain phenomena directly related to the number of imaginary nodes were observed.

In tests with fewer imaginary nodes, there was a greater difficulty in parameter adjustment. This is illustrated by the production estimates of well 2 for the models with meshes containing 0 and 10 imaginary nodes, shown in Figures 13 and 14, respectively.

Note that for the mesh with 0 imaginary nodes, the mean production estimate tends to overestimate the observed values. Furthermore, some prior models deviate significantly from the IMEX estimates, even during the history matching period.

On the other hand, for the mesh with 10 imaginary nodes, the mean predictions are closer to reality. However, despite the mean being relatively accurate, the spread of the models around it is larger. This phenomenon is not observed in more refined meshes, where the model predictions exhibit a narrower convergence range, as illustrated in Figure 15, which shows the predictions for a mesh with 50 imaginary nodes.

This behavior becomes more apparent as the number of imaginary nodes increases: the ensemble of randomly generated models exhibits reduced dispersion, tending to cluster around the mean. However, this does not necessarily imply a better approximation of the true model, as illustrated in Figure 16, which shows the estimated production of well 2 for models with a mesh containing 200 imaginary nodes.

To evaluate the increase in computational complexity as a function of mesh complexity, Table V presents the average time (in seconds) required to compute the model behavior during the forecast stage. It is important to note that this time does not include the processing performed during the training phase. The table also reports the number of parameters estimated during the history matching, as well as the number of edges defined in the mesh.

TABLE V
IMPACT OF THE NUMBER OF IMAGINARY NODES (NIM) ON MESH COMPLEXITY, INCLUDING THE NUMBER OF EDGES, ESTIMATED PARAMETERS DURING HISTORY MATCHING, AND TOTAL FORECAST SIMULATION TIME (IN SECONDS).

nIm	Edges	History-Matched Parameters	Forecast Total Time (s)
0	7	18	54.96
10	35	74	40.41
50	147	298	130.73
100	289	582	210.92
150	427	858	295.61
200	575	1154	383.31
250	718	1440	472.80

As expected, simulation time increases with mesh complexity in most cases. However, for simpler meshes, such as those with 0 or 10 imaginary nodes, the estimated simulation time does not vary significantly.

IV. CONCLUSION

The tests conducted allowed for the identification of relevant phenomena associated with mesh refinement in the INSIM-FT simulator, which should be considered when selecting the model configuration. One of the main findings was that excessive mesh refinement did not yield significant improvements in oil production forecast accuracy. For the case study analyzed, more refined meshes did not provide notable advantages compared to the configuration with 100 imaginary nodes.

On the other hand, computational complexity increases substantially with the number of imaginary nodes. Specifically, the number of parameters to be estimated during the history matching process grows almost linearly with theses nodes. This observation holds for the case in which the same relative permeability curves are assumed for all mesh edges.

In more complex versions of the simulator, where distinct relative permeability curves can be assigned to each edge, the number of parameters increases significantly, further impacting computational feasibility.

Another relevant point concerns model stability. In coarser meshes, such as those with 0 or 10 imaginary nodes, some models generated from random parameters remain far from the observed data even after history matching. Moreover, in these cases, the spread of the adjusted prior models around the mean is wider. Nevertheless, the mean of the simulations remains reasonably close to the observed data, suggesting that the ensemble average may be robust, even if model dispersion increases.

Finally, it was observed that, in heterogeneous reservoirs, certain wells may present greater challenges for matching simulated and observed data, even during the history matching period. Still, the total field production is well estimated, even when using coarse meshes, which reinforces the simulator's ability to capture the overall system behavior despite limitations in reproducing specific local responses.

Acknowledgment

The authors would like to thank CAPES for financial support, and Petrobras for their financial contribution as well as for the support provided during the research.

References

[1] T. O. Osunrinde, I. Chiamaka, and Y. Ahmad, "Field developmental plan analysis: a case study of 'x'reservoir," *Journal of Petroleum Exploration and Production Technology*, vol. 9, no. 3, pp. 2185–2203, 2019.

[2] A. François, F. Maubeuge, E. Tourte, P. Montagnier, and O. Dubrule, "High-performance computing: Towards a safer, faster and sharper exploration & production," in *World Petroleum Congress*. WPC, 2014, p. D033S005R002.

[3] C. Maschio, C. P. V. de Carvalho, and D. J. Schiozer, "A new methodology to reduce uncertainties in reservoir simulation models using observed data and sampling techniques," *Journal of Petroleum Science and Engineering*, vol. 72, no. 1-2, pp. 110–119, 2010.

[4] S. Afshari, B. Aminshahidy, and M. R. Pishvaie, "Application of an improved harmony search algorithm in well placement optimization using streamline simulation," *Journal of Petroleum Science and Engineering*, vol. 78, no. 3-4, pp. 664–678, 2011.

[5] P. Sarma, K. Aziz, and L. J. Durlofsky, "Implementation of adjoint solution for optimal control of smart wells," in *SPE Reservoir Simulation Conference*. SPE, 2005, pp. SPE–92 864.

[6] G. Van Essen, M. Zandvliet, P. Van den Hof, O. Bosgra, and J.-D. Jansen, "Robust waterflooding optimization of multiple geological scenarios," *Spe Journal*, vol. 14, no. 01, pp. 202–210, 2009.

[7] K. J. Heffer, R. J. Fox, C. A. McGill, and N. C. Koutsabeloulis, "Novel techniques show links between reservoir flow directionality, earth stress, fault structure and geomechanical changes in mature waterfloods." *SPE Journal*, vol. 2, no. 02, pp. 91–98, 1997.

[8] F. Jansen and M. Kelkar, "Non-stationary estimation of reservoir properties using production data," in *SPE Annual Technical Conference and Exhibition*. SPE, 1997, pp. SPE–38 729.

[9] B. T. Refunjol, "Reservoir characterization of north buck draw field based on tracer response and production/injection analysis," Ph.D. dissertation, The University of Texas at Austin, 1996.

[10] H. Zhao, Z. Kang, X. Zhang, H. Sun, L. Cao, and A. C. Reynolds, "Insim: A data-driven model for history matching and prediction for waterflooding monitoring and management with a field application," in *SPE Reservoir Simulation Conference*. SPE, 2015, p. D021S007R004.

[11] ——, "A physics-based data-driven numerical model for reservoir history matching and prediction with a field application," *SPE Journal*, vol. 21, no. 06, pp. 2175–2194, 2016.

[12] S. E. Buckley and M. Leverett, "Mechanism of fluid displacement in sands," *Transactions of the AIME*, vol. 146, no. 01, pp. 107–116, 1942.

[13] H. Zhao, W. Liu, X. Rao, G. Sheng, H. A. Li, Z. Guo, D. Liu, and L. Cao, "Insim-fpt-3d: A data-driven model for history matching, water-breakthrough prediction and well-connectivity characterization in three-dimensional reservoirs," in *SPE Reservoir Simulation Conference*. SPE, 2021, p. D011S006R004.

[14] Y. Li and M. Onur, "Insim-bhp: A physics-based data-driven reservoir model for history matching and forecasting with bottomhole pressure and production rate data under waterflooding," *Journal of Computational Physics*, vol. 473, p. 111714, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002199912200777X

[15] Z. Guo and A. C. Reynolds, "Insim-ft in three-dimensions with gravity," *Journal of Computational Physics*, vol. 380, pp. 143–169, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002199911830812X

[16] Z. Guo, *History Matching, Prediction and Production Optimization with a Physics-Based Data-Driven Model*. The University of Tulsa, 2018.

[17] D. P. Mitchell, "Spectrally optimal sampling for distribution ray tracing," in *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, 1991, pp. 157–164.

[18] H. Edelsbrunner, T. S. Tan, and R. Waupotitsch, "An o (n 2log n) time algorithm for the minmax angle triangulation," in *Proceedings of the sixth annual symposium on Computational geometry*, 1990, pp. 44–52.

[19] D. Kaviani, P. P. Valkó, and J. L. Jensen, "Application of the multiwell productivity index-based method to evaluate interwell connectivity," in *SPE Improved Oil Recovery Conference*. SPE, 2010, pp. SPE–129 965.

[20] G. Ren, Z. Wang, Y. Lin, T. Onishi, X. Guan, and X.-H. Wen, "A fast history-matching and optimization tool and its application to a full field with more than 1,000 wells," *SPE Journal*, vol. 29, no. 10, pp. 5538–5558, 10 2024. [Online]. Available: https://doi.org/10.2118/212188-PA

[21] A. A. Emerick and A. C. Reynolds, "History matching time-lapse seismic data using the ensemble kalman filter with multiple data assimilations," *Computational Geosciences*, vol. 16, pp. 639–659, 2012.

[22] ——, "Ensemble smoother with multiple data assimilation," *Computers & Geosciences*, vol. 55, pp. 3–15, 2013.

[23] ——, "History-matching production and seismic data in a real field case using the ensemble smoother with multiple data assimilation," in *SPE Reservoir Simulation Conference*. SPE, 2013, pp. SPE–163 675.

[24] A. Emerick, *Ensemble Data Assimilation Applied to Geological Reservoir Models*, 05 2025.

[25] C. M. Dafermos, "Polygonal approximations of solutions of the initial value problem for a conservation law," *Journal of Mathematical Analysis and Applications*, vol. 38, no. 1, pp. 33–41, 1972. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022247X7290114X

[26] J. O. Langseth, N. H. Risebro, and A. Tveito, "A conservative front tracking scheme for 1d hyperbolic conservation laws," in *Nonlinear Hyperbolic Problems: Theoretical, Applied, and Computational Aspects: Proceedings of the Fourth International Conference on Hyperbolic Problems, Taormina, Italy, April 3 to 8, 1992*. Springer, 1993, pp. 385–392.

[27] M. De Berg, *Computational geometry: algorithms and applications*. Springer Science & Business Media, 2000.

[28] A. A. Emerick, "Investigation on principal component analysis parameterizations for history matching channelized facies models with ensemble-based data assimilation," *Mathematical Geosciences*, vol. 49, no. 1, pp. 85–120, 2017.

[29] M. Grave, E. da Silva Castro, S. Pesco, A. B. B. Junior, L. da Silva Gasparini, R. G. Cavalcante, R. R. da Costa Alves, and J. R. P. Rodrigues, "A mass conservative insim-ft-3d physics-based data-driven oil-water reservoir simulator," *Geoenergy Science and Engineering*, vol. 241, p. 213102, 2024.

[30] Y. Zhao, *Parameterization and Data Assimilation for Complex Geological Reservoirs*. The University of Tulsa, 2019.

[31] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.