# Annotation of Using Borehole Time-Lapse Gravity by Genetic Algorithm Inversion for Subsurface Modeling

**Indra Gunawan[1,2,*], Eko Januari Wahyudi[1,2], Susanti Alawiyah[1,2], Wawan Gunawan A. Kadir[1,2] & Umar Fauzi[3]**

[1]Geophysical Engineering Department, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Jalan Ganesa No.10, Bandung 40132, Indonesia
[2]Applied and Exploration Geophysics Research Group, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Jalan Ganesa No.10, Bandung 40132, Indonesia
[3]Physics of Earth and Complex Systems Research Group, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Jalan Ganesa No.10, Bandung 40132, Indonesia
[*]E-mail: gunawan@geoph.itb.ac.id

**Highlights:**

- The genetic algorithm (GA) inversion produced the best results when the borehole positions were placed in a state of symmetry towards the body object's mass.
- The addition of boreholes did not necessarily increase the success rate of the GA inversion.
- A number of boreholes in the right position proportional to the body object provided a better result of the GA inversion.

**Abstract.** We present the annotation to a genetic algorithm (GA) method for an inverse synthetic subsurface density model using surface and borehole time-lapse gravity data. The objective of the inversion is to find the boundaries of the object area and background, where one bit of the chromosome represents the densities. The model that was used in this paper was a simple homogeneous body anomaly and a simplified real water mass injection model in order to argue that the code is suitable for field modeling. We show the influences of the existence of borehole gravity data and location towards the inversion, where the result indicates that an additional good borehole location could increase the success rate up to 13.33% compared to without gravity borehole data for the simple model and up to 4.39% for the field model. The inversion produced the best results when the borehole positions were placed in a state of symmetry towards the body object's mass.

**Keywords**: *genetic algorithm; gravity inversion; subsurface modeling; surface and borehole gravity; time-lapse gravity.*

## 1 Introduction

Gravity inversion is one of the geophysical methods that can be used to model subsurface density. The time-lapse gravity modeling of subsurface condition

Received August 14th, 2019, Revised March 16th, 2020 , Accepted for publication February 26th, 2020.
 ISSN: 2337-5779,
DOI: 10.5614/j.eng.technol.sci.2020.52.2.2

changes is very efficient. It has promising potential for use in monitoring injection or fluid extraction in a reservoir, as shown by Kadir, *et al.* [1] and Santoso, *et al.* [2], both for broad and deep target areas. The gravity technique can even be used in volcanic environments, as shown by Battaglia, *et al.* [3]. This type of area is difficult to measure with seismic instruments because the irregular structure diversity scatters the active seismic wave signals.

Modeling of subsurface condition changes using gravity data is complicated since it has considerable ambiguity. The signal read by the gravimeter is the accumulation of all anomalies in the subsurface. Thus, it is challenging to distinguish the anomaly boundaries. The existence of borehole gravity technology can reduce the percentage of ambiguity, as reported by Jageler [4]. Surveyors have further developed borehole gravity technology for a long time since it was first reported by Smith [5]. In addition to the size of the instrument, its sensitivity to pressure also hinders the surveyor in placing the gravimeter at the desired stations. Recently, borehole gravimeters have been much improved, both in terms of capability and instrument size. The sensitivity of the latest borehole gravimeter instruments can resolve below five microgal, as reported by LaCoste [6], almost the same level as surface gravimeter instruments. Because of this, gravity borehole measurement is a prospect, especially in fluid injection monitoring, as shown by Brady, *et al.* [7].

The ability of data processing methods must keep up with the development of instrument technology. For this reason, it is necessary to develop a processing method that can increase the resolution of subsurface time-lapse anomalies horizontally and vertically through gravity measurements on the surface and in boreholes. There are many techniques in gravity inversion, but generally they fall into two categories. The first category is deterministic inversion. This kind of method constructs a model solution using a spatial position function. The application of deterministic inversion to both synthetic surfaces and borehole gravity data was presented by Jacob, *et al.* [8] to model carbon capture storage in reservoir monitoring, where the results showed that the existence of borehole gravity data could increase the success rate by up to 5%. The second category is stochastic inversion, which usually uses a random approach to obtain the solution. This method can be analyzed statistically due to its random nature. This behavior can be an advantage over the deterministic method since it has the flexibility to generate many different solutions to escape from local optima, while the deterministic approach is limited to the results from the initial model. By applying a model norm and using Tikhonov regularization, we can control the random characteristics so the solution approaches the desired results.

GA is a stochastic algorithm that is quite popular for solving optimization problems and is widely used in diverse applications. In the field of seismic

studies, GA has excellent performance modeling the subsurface using waveform data, as shown by Yamanaka and Ishida [9], Pezeshk and Zarrabi [10], Sajeva, *et al.* [11] and Salamanca, *et al.* [12]. In the field of geomagnetic studies, GA has shown satisfying performance in modeling the magnetic structure, as reported by Yamamoto and Seama [13].

Inspired by the theory of natural evolution, the GA method uses heuristic search. The objective model of the inversion is based on minimization and the density of the subsurface model can be discretized, which makes GA a suitable method to separate the background from the object in gravity inversion. GA variables (i.e. chromosomes and genes) can translate into the model solution in gravity modeling, usually related to shaped objects (i.e. prism, cylinder, or point of mass) with different densities. Even though the GA approach is known as a computationally expensive method, it is very convenient to invert an object boundary when the densities are quite discrete from the background, especially when we can apply a small bit to discretize the density difference.

The GA inversion of gravity data has been presented by Krahenbuhl and Li [14] to image salt bodies. The method has applied by Wahyudi, *et al.* [15] to time-lapse density anomalies that accommodate two types of density changes. Since it was able to get sufficiently good results with surface data in both cases, we expected that the existence of borehole data could contribute to enhancing the inversion results. In this paper, we demonstrate GA inversion applied to both surface and borehole gravity data to represent a real model. We implemented several boreholes to find out the influence of their location on the results of the inversion. Borehole gravity measurement is expensive. The simulation calculation to obtain the estimated parameters with the proposed method before the survey may reduce project costs.

## 2 Methods

In a gravity inversion, the objective model is to minimize the misfit function ($\varphi$), which consists of misfit data ($\varphi_D(\rho)$) and the model norm ($\varphi_M(\rho)$). It is formulated as follows:

$$\varphi = \varphi_D(\rho) + B\varphi_M(\rho) \rightarrow \min, \tag{1}$$

where $B$ is the variable to control the balance in the Tikhonov regularization curve. The best value of $B$ is manually obtained by plotting the position of the misfit data versus the model norm for some calculations and then choosing the value that has the best balance. The misfit data are defined as the average difference of the inverted gravity ($d^{calc}$) and the observation data ($d^{obs}$) for each iteration (Eq. (2)),

$$\varphi_D(\rho) = \sqrt{\frac{1}{N}\sum_{a=1}^{N}\left\|d_a^{calc} - d_a^{obs}\right\|^2}, \quad (2)$$

where $N$ is the number of the station, and $a$ is the station index. The gravity measurement in the field obtains the observation data and the inverted gravity is the forward modeling of the model as it is formulated by equations from Okabe [16] and Li and Chouteau [17]. Furthermore, the model norm is defined as a smoothing function for lateral distributions. The model norm is the average of the difference of all nearest neighbor densities (Eq. (3)) in horizontal positions (Figure 1). We neglect the calculation in the edging prisms (white area) because it does not have the same number of neighbors as the others (gray area). Thus, we begin the calculation between the prism index $m = 2$ and $n = 2$ to $m = s - 1$ and $n = t - 1$. Where $s$ and $t$ are the last index of the prism in the X and Y-axis direction in the model, respectively.

$$\varphi_M(\rho) = \sum_{m=2,n=2}^{m=k-1,n=l-1}\left(\frac{1}{8}\left\{\begin{array}{c}\left[\rho_{m,n} - \rho_{m,n-1}\right]^2 + \left[\rho_{m,n} - \rho_{m,n+1}\right]^2 \\ +\left[\rho_{m,n} - \rho_{m-1,n}\right]^2 + \left[\rho_{m,n} - \rho_{m+1,n}\right]^2 \\ +\left[\rho_{m,n} - \rho_{m-1,n-1}\right]^2 + \left[\rho_{m,n} - \rho_{m-1,n+1}\right]^2 \\ +\left[\rho_{m,n} - \rho_{m+1,n-1}\right]^2 + \left[\rho_{m,n} - \rho_{m+1,n+1}\right]^2\end{array}\right\}\right) \quad (3)$$

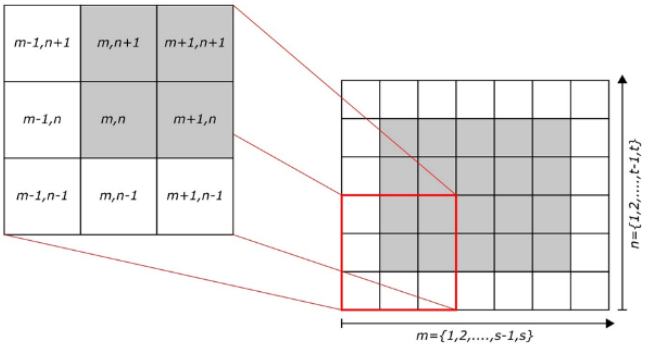


**Figure 1** Illustration of model block positions.

The focus of this research was on the one bit of the genetic algorithm inversion for the time-lapse gravity. There are only two possible values of each variable (i.e. either true or false), so the object densities are discretized into two values that represent the background and the object of the density anomalies. In this way, we can quickly determine the density anomaly object boundaries. The genetic algorithm is performed using a Boolean data type as the object container so that the code is lighter and can run faster to save computing resources and time.

The GA method is based on the processes of recombination, mutation, and selection to produce the best offspring for each iteration. In the process, we modified the function to get better results to make it convenient for our method. In our research, we used uniform crossover, which was first introduced by Syswerda [18], for the recombination function and then compared some modified randomly flipping cell mutation techniques to find the most suitable parameters for our calculation (Figure 2). The additional GA parameters were taken from Herrera, *et al.* [19].

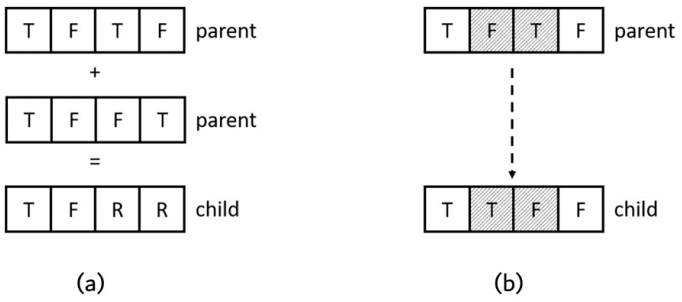


**Figure 2** Illustration of crossover (a) and mutation (b) in the recombination section. T, F, and R mean true, false, and random (i.e. true or false) Boolean values, respectively. The grey areas (b) indicate the chromosomes doing mutation.

The uniform crossover, mutation, and ranking selection in our code follow these rules:

1. *Uniform crossover*. Combining two parents produces a child. The children will have either the first or the second parent's properties. The parents and the child will gather together to join the population. All parents recombine, but their partners are randomly selected.
2. *Mutation.* A parent produces a child. For every *n* number of the chromosome, each gene mutates into the opposite value. The user defines *n*, but the location of the chromosomes that are mutated is randomly selected. Both parents and the child join the population. The user defines the number of parents that mutate. In our case, it was 20 percent of the parent population. The selection of parents that mutate is randomly chosen.
3. *Ranking selection.* New offspring are selected from the population. The best individuals that have a better weight will be chosen to become new offspring. All other individuals are killed. The user defines the number of new offspring.

To avoid homogeneous offspring that have recently appeared in the genetic algorithm for a large number of generations, we applied an additional mutation when we found similar individuals in the parent population.

## 3 Results and Discussion

### 3.1 Simple Model (SM)

We built the data from a synthetic model that illustrates fluid injection in a 1-km deep reservoir with an area of 2 km east-west and 2 km north-south. The aquifer's thickness was 150 m, with neglected porosity. The injection took place in the center of the reservoir. After time *t* there was a density-contrast object shaped like a circle from a 2D top view. By calculating the differences in the gravity response at the beginning of the injection and after injection at time *t*, we got the time-lapse gravity response. We assumed there was no added noise due to deformation, terrain changes, surveying process, or other interference. There were two kinds of density in the reservoir: the background density and the object density. We also assumed that the injection fluid had uniform density, i.e. 500 kg/m$^3$ in all areas. Since there was no density change in the background, we defined the background density as 0 kg/m$^3$. The dimensions of the block prisms in the reservoir were 100 x 100 x 150 m; it consisted of 400 body prisms (Figure 3).

Gravity forward modeling generated the time-lapse gravity response due to the anomalous object model. On the surface, the gravity response had 482 stations that were mostly placed at 100 m distance from each other over the anomalous object. Furthermore, six borehole locations were applied to observe the influence of the existence of boreholes and their positions on the inverted results. The stations of the boreholes were placed at every 100 m from the surface until 2 km depth. Thus there were 20 additional gravity data for each borehole. There were six scenarios in the inversion: invert the model from the surface data only (S1); invert the model from the surface and the BH01 data (S2); invert the model from the surface and the BH02 data (S3); invert the model from the surface and the BH03 data (S4); invert the model from the surface and the BH01, BH02, BH04 data (S5); and invert the model from the surface and the BH01, BH02, BH04, BH05, BH06 (S6).

We followed the method from Jacob, *et al.* [8] to measure the error rate of the inversion. Missed volume and false alarm were calculated by comparing the model volume with the inverted volume. Missed volume was defined as the area that should be inside the model but appears outside the inverted volume. Vice versa, false alarm was defined as the area that should be outside of the model but appeared inside the inverted volume. We calculated each scenario for each model three times and then calculated the average to ensure that the results fell within the convergence.

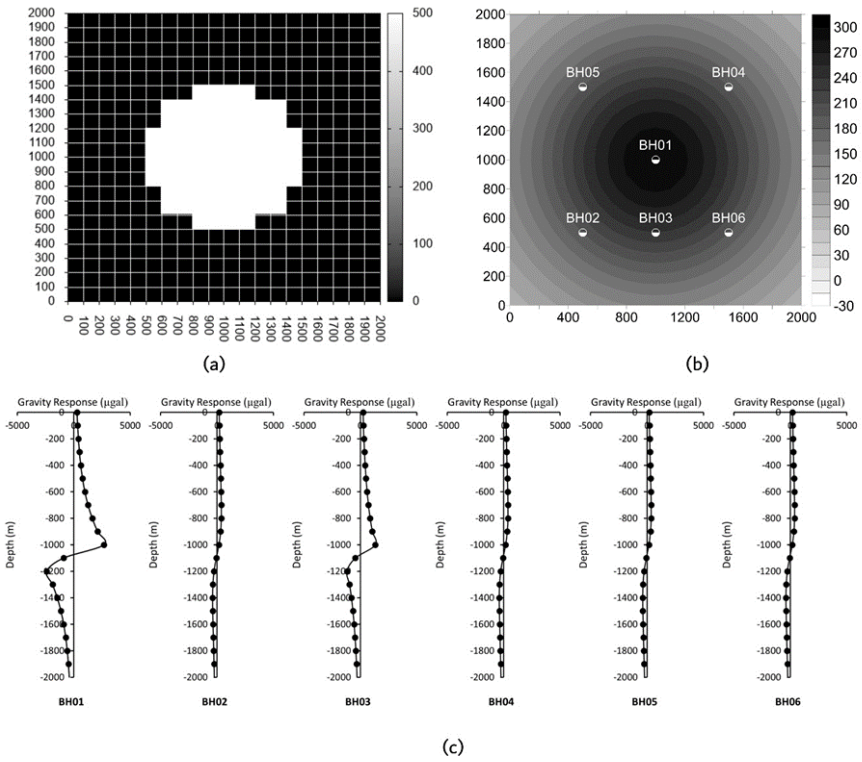


**Figure 3** Illustration of synthetic time-lapse subsurface density model in kg/m$^3$ (a), the gravity response at the surface (b), and the boreholes (c) in microgal.

The inverted results for the SM from the surface data (S1) had an average error rate of 20.83%, while the average missed volume error was 10% and the average false alarm error was 10.83%. We could decrease the average error rate by 7.5% by adding one borehole data (S2), placed in the center of the anomalous gravity (i.e. BH01). However, when its position was changed to the outer areas (i.e. BH02 or BH03), it led to worse results (i.e. 23.83% and 26.25% average error rate). The results were even worse than for S1 (i.e. when the inverted results were produced only from the surface data). It shows that in the case of S3 and S4, the existence of borehole data becomes interference for the inversion.

Furthermore, the additional borehole data were also problematic because they did not decrease the error rate. When we applied three borehole data to the calculation, in the case of S5, the inversion only generated an average error rate of 11.67%. Although this was better than S1, it looks like a waste of resources since it did not do better than using fewer boreholes (S2). This trend continued when we applied more borehole data to the calculation, where five boreholes (S6)

even increased the level of the inversion error rate (15%). Figure 4 shows a summary (error rate) of the results of all inverted models.

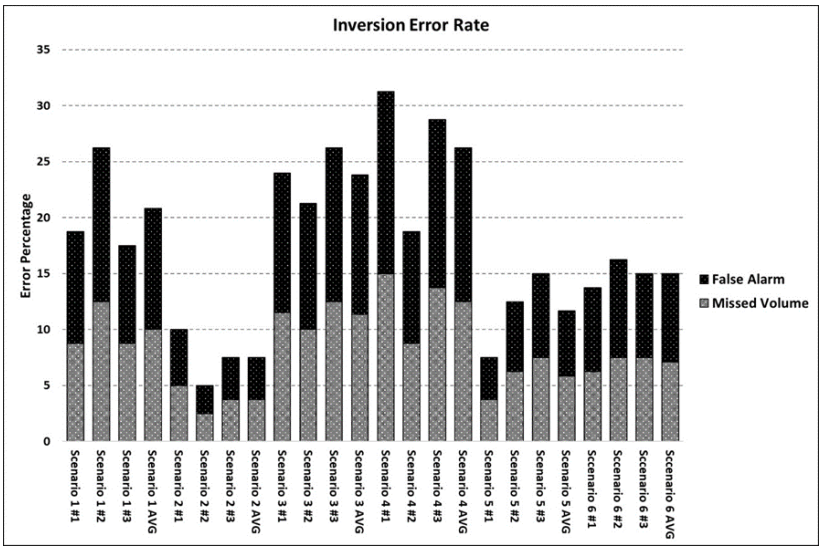


**Figure 4** Summary of the error rates of the inverted models for SM. AVG stands for average.

## 3.2 Field Model (FM)

The field data were taken from the Hare, *et al.* model [20]. This time-lapse model illustrates the waterflood surveillance at Prudhoe Bay, Alaska. The area of investigation was a 28 km east-west and 31 km north-south area, where the reservoir was at 2.5 km depth with 150 m of thickness, represented by 868 cells with a size of 1 km x 1 km x 150 m of each prism. The reservoir had 20% porosity. Since our model used only two density differences, we simplified the model into two values: the delta density due to the time-lapse was 60 kg/m$^3$, uniform for all areas of the object anomaly, while the background time-lapse density was 0 kg/m$^3$.

The observation stations on the surface had 928 points. Boreholes were located in three different places (i.e. BH01, BH02, and BH03). Each borehole had 30 additional gravity observation stations that were set at every 100 m from the surface until 3 km depth. The surface and borehole gravity data were generated by doing forward modeling on density models following Reference [20]. We added ten microgal synthetic standard deviations of Gaussian noise to both the surface and the borehole gravity response (~15% of noise). Figure 5 illustrates the object body and gravity responses. We applied the inversion to four scenarios:

invert the model from the surface data only (S1); invert the model from the surface and BH01 data (S2); invert the model from the surface and BH02 data (S3); and invert the model from the surface and BH02 and BH03 data (S4).

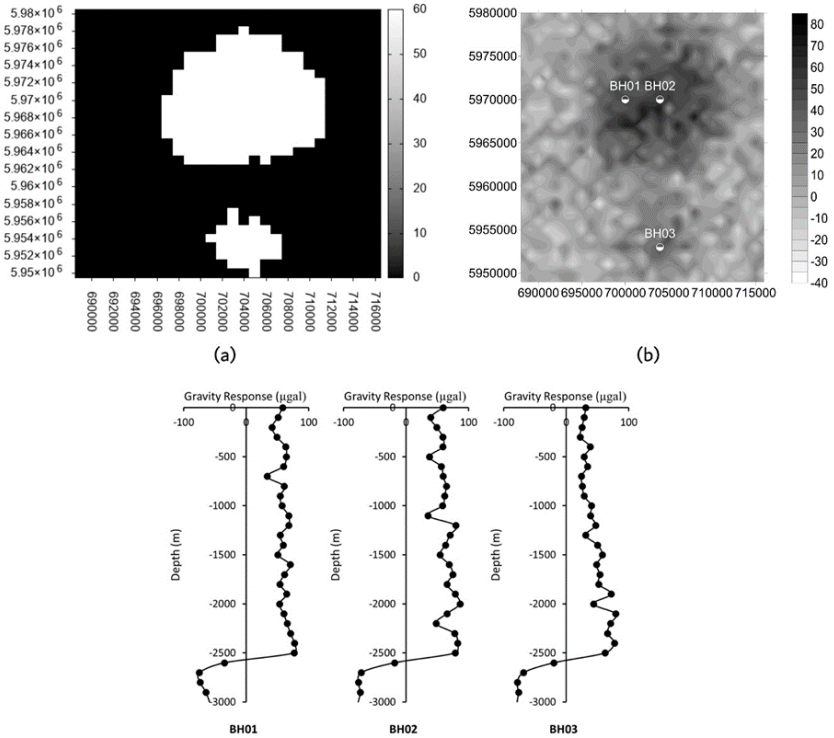


**Figure 5** The synthetic time-lapse subsurface density model in kg/m$^3$ (left), the gravity response at the surface (middle), and the boreholes (right) with ten microgal Gaussian noise. Coordinates in UTM.

The inversion was executed three times for each scenario, after which the average was calculated. Then the error ratio for each calculation was determined. A summary of the results is presented in a graph (Figure 6) and illustrated in Figure 7.

It was found that S4 was the best ratio. It could invert the model up to an average error rate of 19.41%, while S1, S2, and S3 produced 23.79%, 24.11%, and 23.01%, respectively. Our stochastic method of inversion showed that it was in close agreement with Jacob, *et al.* [8], who used a deterministic approach, i.e. the borehole only decreased the error rate by about 5%. Furthermore, the result of S2 shows that the existence of borehole data did not reduce the error rate of the inversion, while S3 only cut a small amount from the error rate.

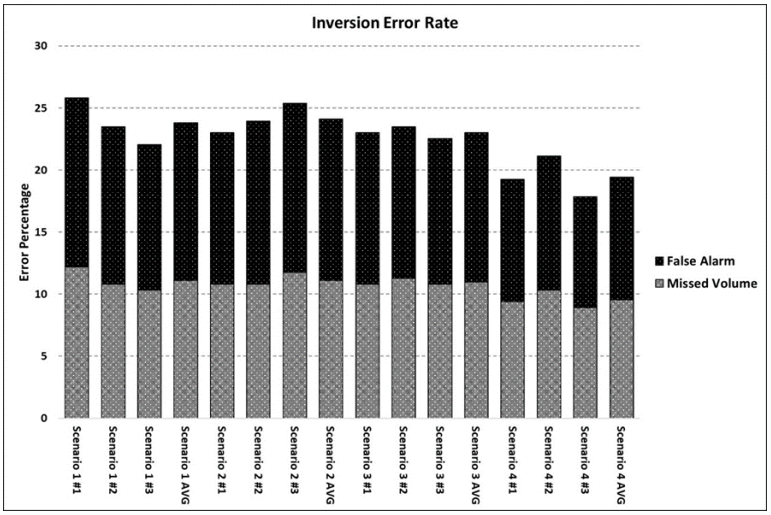


**Figure 6** Summary of the error rate of the inverted model for FM. AVG stands for average.

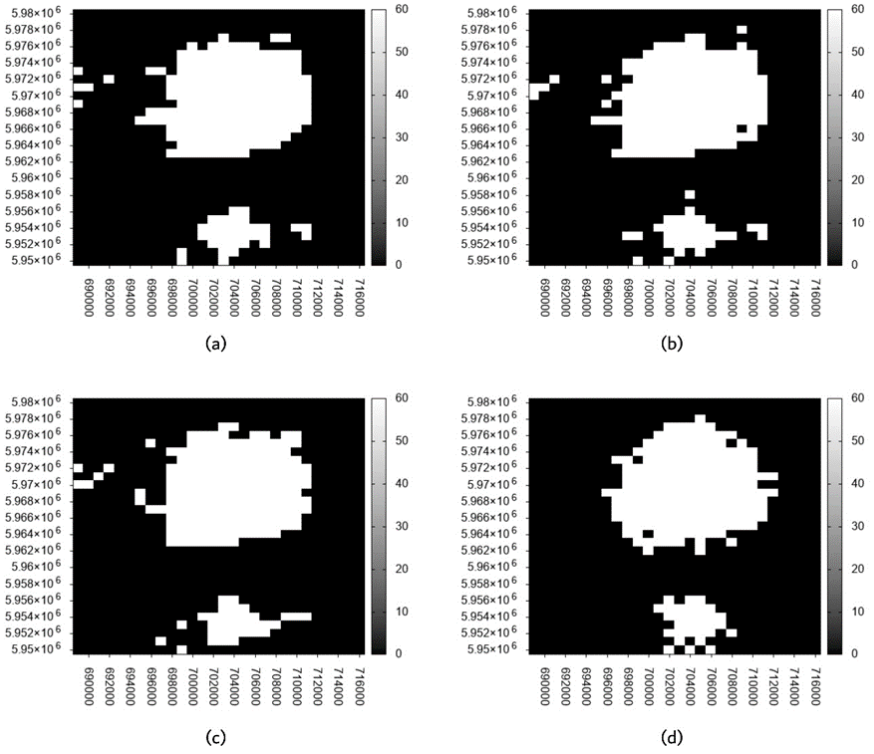


**Figure 7** Map of the inverted densities of S1 (a), S2 (b), S3 (c), and S4 (d) on the FM. Coordinates in UTM.

### 3.3 Discussion

By comparing the position of the boreholes (i.e. S2, S3 to S4 in the case of SM and S2 to S3 in the case of FM), we found that the position of the boreholes is essential in the GA inversion process, where the proper borehole location determines whether the data will be useful or not. The other results (i.e. S5 and S6 in SM and S4 in FM) also showed that the addition of boreholes does not necessarily decrease the error rate of the inversion result if it is not proportional to the number of object bodies. This behavior occurred in the GA inversion due to the distance between the object model and the observation stations, which was defined inversely proportional to the gravity values. This causes the calculation of the data misfit (Eq. (2)) at the minimum distance to be dominant in the inversion process. Stations in the boreholes adjacent to the object model give a larger gravity response, which leads to a more significant misfit effect when the inversion has not fallen into the optimal solution than when a station that is farther removed from the object model.

This behavior leads to the conclusion that the position of the borehole observation station that gives the most optimal inversion results must be in a place that provides perfect symmetry of the mass distribution of the object model to be inverted. By doing an example calculation in the case of SM that uses a perfectly circular object (symmetry in the XY plane), we can see that the position of the borehole observation station that had the most symmetric condition was in S2, not in S3 or S4. In S5 and S6, even though there were many borehole observation stations, symmetry was not better than in S2, thus the inverted model produced was not better. In the case of the FM model, we can look at S2 and S3, which proved such behavior. However, because the FM model had two body density objects, the addition of a borehole that is proportional to the body object number will undoubtedly provide better results (S4), of course, with the provision of a position of the borehole that is symmetric towards the body objects, as discussed above.

## 4 Conclusions

An application of genetic algorithm inversion for gravity modeling was developed that works for 1-bit resolution. It was applied to a model that was depicted as a collection of rectangular prisms. The GA method was successful in performing stochastic inversion on surface and borehole gravity data to determine the object boundaries. The results indicate that additional good borehole locations could increase the success rate up to 13.33% compared to without gravity borehole data for the simple model and up to 4.39% for the field model. The inversion also showed that the existence and location of the borehole gravity data affected the error rate of the inverted results to a certain degree. The inversion

produced the best results when the borehole positions were symmetric towards the body object's mass; this was the annotation to the GA inversion that was applied to the joint surface and borehole gravity data.

## References

[1] Kadir, W.G.A., Santoso, D., Alawiyah, S., Setianingsih, Widianto, E., Sardjito, E. & Waluyo, Microgravity *Monitoring for Sandstone and Carbonate Reservoirs Along Water Injection Activity Case Study of Talang Jimar and Tambun oil fields, Indonesia,* The Contribution of Geosciences to Human Security, Institut Teknologi Bandung, Kyoto University, Kyoto University GCOE Program of HSE, eds., Logos Verlag Berlin, pp. 127-140, 2011.

[2] Santoso, D., Kadir, W.G.A., Alawiyah, S., Setianingsih, Wahyudi, E.J., Sarkowi, M. & Minardi, S., *Understanding the Time-Lapse Microgravity Response due to Subsidence and Groundwater Level Lowering*, The Contribution of Geosciences to Human Security, Institut Teknologi Bandung, Kyoto University, Kyoto University GCOE Program of HSE, eds., Logos Verlag Berlin, pp. 27-28, 2011.

[3] Battaglia, M., Gottsmann, J., Carbone, D. & Fernandez, J., *4D Volcano Gravimetry*, Geophysics, **73**(6), pp. WA3-WA18. 2008.

[4] Jageler, A.H., *Improved Hydrocarbon Reservoir Evaluation through Use of Borehole-Gravimeter Data*, Journal of Petroleum Technology, **28**, pp. 709-718, 1976.

[5] Smith, N.J., *The Case for Gravity Data from Boreholes,* Geophysics, **15**(4), pp. 605-636, 1950.

[6] Micro-g LaCoste, *Borehole Gravimeter Test Report,* University of Texas, Austin Campus, 2004.

[7] Brady, J.L., Bill, M.L., Nind, C., Pfutzner, H., Legendre, F., Doshier, R.R, MacQueen, J.D. & Beyer, L.A, *Performance and Analysis of a Borehole Gravity Log in an Alaska North Slope Grind-and-Inject Well,* Society of Petroleum Engineers Annual Technical Conference and Exhibition, pp. 1324-1338, 2013.

[8] Jacob, T., Rohmer, J. & Manceau, J.C., *Using Surface and Borehole Timelapse Gravity to Monitor Co2 in Saline Aquifers: A Numerical Feasibility Study,* Greenhouse Gases: Science and Technology, **6**(1), pp. 34-54, 2016

[9] Yamanaka, H. & Ishida, H., *Application of Genetic Algorithms to an Inversion of Surface-Wave Dispersion Data,* Bulletin of the Seismological Society of America, **86**(2), pp. 436-444, 1996.

[10] Pezeshk, S. & Zarrabi, M., *A New Inversion Procedure for Spectral Analysis of Surface Waves Using a Genetic Algorithm*, Bulletin of the Seismological Society of America, **95**(5), pp. 1801-1808, 2005.

[11] Sajeva, A., Aleardi, M. & Mazzotti, A., *Genetic Algorithm Full-Waveform Inversion: Uncertainty Estimation and Validation of the Results*, Bollettino di Geofisica Teorica ed Applicata, **58**(4), pp. 395-414, 2017.
[12] Salamanca, A., Gutiérrez, E. & Montes, L., *Optimization of a Seismic Inversion Genetic Algorithm*, SEG International Exposition and 87th Annual Meeting, pp. 823-827, 2017.
[13] Yamamoto, M. & Seama, N., *Genetic Algorithm Inversion of Geomagnetic Vector Data Using A 2.5-Dimensional Magnetic Structure Model*, Earth Planets Space, **56**, pp. 217-227, 2004.
[14] Krahenbuhl, R. & Li, Y., *Inversion of Gravity Data Using a Binary Formulation*, Geophysical Journal International, **167**(2), pp. 543-556, 2006.
[15] Wahyudi, E.J., Santoso, D., Kadir, W.G.A. & Alawiyah, S., *Designing A Genetic Algorithm for Efficient Calculation in Time-Lapse Gravity Inversion*, Journal of Engineering and Technological Sciences, **46**(1), pp. 58-77, 2014.
[16] Okabe, M., *Analytical Expressions for Gravity Anomalies Due to Homogeneous Polyhedral Bodies and Translations into Magnetic Anomalies*, Geophysics, **44**(4), pp. 730-741, 1979.
[17] Li, X. & Chouteau, M., *Three-Dimensional Gravity Modeling in All Space*, Surveys in Geophysics, **19**(4), pp. 339-368, 1998.
[18] Syswerda, G., *Uniform Crossover in Genetic Algorithms*, 3rd International Conference on Genetic Algorithms, pp. 2-9, 1989.
[19] Herrera, F., Lozano, M. & Verdegay, J., *Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis*, Artificial Intelligence Review, **12**(4), pp. 265-319, 1998.
[20] Hare, J., Ferguson, J. & Brady, J., *The 4D microgravity Method for Waterflood Surveillance: Part Iv - Modeling and Interpretation of Early Epoch 4D Gravity Surveys at Prudhoe Bay, Alaska*, Geophysics, **73**(6), pp. WA173-WA180, 2008.