

Cephy: A stand-alone, free and open-source toolkit for hydrocarbon in place evaluation by ordinary kriging

Patient K. Mulekya^{a,b,c*}, Ilunga Kabasele^b, Kasongo Numbi^{b,c}, Princess C. Nwankwo^{a,d}, Tomisin B. Babalola^a, Dan Y. Mandjumba^e, Adolphe Kabongo Kayamba^f, Aurélie Nkayilu Waluvuvamu^b

^aPetroleum Geoscience Program, Pan African University, Life and Earth Sciences Institute (including Health and Agriculture), University of Ibadan, Ibadan, Nigeria

^bDepartment of Exploration and Production, Faculty of Oil, Gas and New Energies, University of Kinshasa, Kinshasa, DR Congo

^cLaboratory of Hydraulics Management and Energy Production Engineering, Faculty of Agronomic Sciences, University of Kinshasa, Kinshasa, DR Congo

^dDepartment of Petroleum Engineering, Faculty of Technology, University of Ibadan, Ibadan, Nigeria

^eDepartment of Refinery and Petrochemistry, Faculty of Oil, Gas and New Energies, University of Kinshasa, Kinshasa, DR Congo

^fDepartment of Mathematics and Computer Sciences, Faculty of Sciences, University of Kinshasa, Kinshasa, DR Congo

ABSTRACT

Cephy is a software toolkit written in Java that enables the user to perform spatial interpolation of subsurface data by ordinary kriging. Key features of the software include data exploration by map plot with symbol dimension proportional to the magnitude of the property at a given location; variogram modeling and interpolation of petrophysical properties; as well as oil in place estimation per area unit throughout the reservoir. Cephy enables the user to define the tolerance and the minimum number of pairs required to validate the experimental variogram of the corresponding lag distance when learning the relationship between geolocated data. The software is entirely made of a graphical user interface for adjusting the variogram modeling parameters and visualizing the subsequent interpolationally-generated map.

Cephy is applied on Kinkasi reservoir in Kinkasi field (Coastal basin of DR Congo) to demonstrate its functionalities. The result supports a better understanding of oil distribution throughout the reservoir and can be used as a baseline for the suggestion of new infill drilling locations.

Keywords: Geostatistics, Interpolation, Kriging, Oil in place, Variogram.

1. INTRODUCTION

The estimation of the volume of hydrocarbon in a reservoir is a core task in the upstream petroleum industry, and it provides the basis for a field development plan [1–3]. Two types of methods are used for evaluating the volume of hydrocarbon in place (HIP): volumetric methods and dynamic methods [4,5]. Dynamic methods are based on principles that depend on production data. They include the Material Balance Equation and simulations, while volumetric methods are static and rely on the volume of the net reservoir and petro physical properties, including the

area (A), the thickness (h), the net pay (he), gross thickness (ht), the porosity (Φ), the water saturation (Sw) and eventually the formation volume factor of hydrocarbon (FVF) in place^[2,6].

$$HIP = \frac{A \times h \times h_e / h_t \times \Phi \times (1 - S_w)}{FVF} \quad (1.1)$$

The two main issues with volumetric methods are the ignorance of the heterogeneity of reservoirs and the fact that samples are limited to the number of wells drilled in the reservoir. Therefore, the need for interpolation techniques comes in to estimate petrophysical properties at non-sampled locations [7].

Geostatistical interpolation techniques are more reliable since they take into account the dependence between observations when they are located in space. Several approaches have been developed, amongst which are the Inverse Distance Weighting [8], Splines [9], and Kriging [10]. Kriging, known as the best linear unbiased predictor, is a complex procedure for geospatial interpolation that relies on the variogram model depicted by a given spatial property [11–15].

In this article we present Cephy, a software toolkit that aims to simplify kriging for reservoir evaluation in a straightforward one-way flowchart while giving geoscientists the possibility to adjust variogram modeling parameters such as the minimum number of pairs and tolerance and visualize their fitting. The result of Cephy is not only the amount of oil and gas in place but also a 2D map of the distribution of oil quantity in the reservoir. The source code of Cephy is publicly available on GitHub (<https://github.com/patkambm2/cephy>). Cephy is written in Java and is designed to be operating system-independent – to be portable from one system to another (Windows, Linux, macOS, etc.).

The feature set of Cephy is divided into four parts:

- Input tab: it consists of importing data that should be in CSV format and assigning column heading to data.
- Data exploration tab: this panel displays the map of the wells in 2D, and it gives the possibility to assign a circular symbol whose radius is directly proportional to any property collected on the well in order for the user to depict any possible directional trend of data on the field.
- Variogram tab: it consists of five subtabs that display the default spherical variogram from each structural and petrophysical property that will be subjected to kriging allowing the user to change the tolerance and the minimum number of pairs of codistant wells to validate the variogram.
- Kriging tab: this last tab operates the kriging to finally calculate the quantity of oil in place (OIP) by volumetric method requiring the oil-water and gas-water contact, if any. It also visualizes the distribution of OIP over the 2D map of the oil field using a color bar scale.

2. DESIGN AND IMPLEMENTATION

2.1. Theoretical and experimental variograms

According to Matheron^[10], nature is not entirely unpredictable. Two observations located close to each other should, on average, be more similar than two observations located far apart. The difference between the values taken by two random variables is $x(u) - x(u+L)$. It is also a random variable whose variance can be calculated. This variance should be minor when the points are close together (the values are more similar on average) and significant when the points are far apart. The half variance of this difference is called the semivariogram or variogram[16].

$$\gamma(\vec{L}) = \frac{1}{2} Var (x(\vec{u} + \vec{L}) - x(\vec{L})) \quad (2.1)$$

The variogram is defined as the variance of the difference between field values at two locations^[17]. It measures "spatial variability", i.e., the dissimilarity between values based on their separations^[18,19]. It describes the spatial continuity of the regionalized variable practically implemented as follows^[21].

$$\hat{\gamma}(\vec{L}) = \frac{1}{2n(\vec{L})} \sum_{i=1}^{n(\vec{L})} [x(\vec{u}_i) - x(\vec{u}_i + \vec{L})]^2 \tag{2.2}$$

$n(\vec{L})$ is the number of pairs with L the lag distance

$x(\vec{u}_i)$ the sampled property at \vec{u}_i location

$x(\vec{u}_i + \vec{L})$ the sampled property at $(\vec{u}_i + \vec{L})$ location

2.2. Getting sufficient pairs

The variogram can be computed over all directions (omnidirectional) for both regular and irregular sampling designs. For a regular grid, the initial nominal lag interval should be that of the grid spacing, whereas, for irregularly scattered data, the choice is wider because the observations may be separated by potentially unique lags in both distance and direction [22]. As defined, the variogram is a statistical property; we compute the average of the squared difference between two values located at a specific distance (lag distance) from each other. For this average to be truly representative of a certain lag distance, a certain number of data pairs are needed.

It is difficult to say that a precise number of pairs for all lag distances is sufficient. However, seven to ten pairs are needed for a reliable variogram estimate for a given lag distance, so the more pairs one has, the more accurate the variogram estimate will be [21].

There are no set rules for determining what lag size should be used. The researcher's skill, knowledge of the phenomenon being studied, and the reason(s) for modeling a variogram help to determine the appropriate lag size.

Eventually, as the lag distance corresponds to the exact distance between any two points, it may be that only one pair of data is available at this distance, on which one might not be able to obtain a reliable estimate of the variogram.

So, to avoid this problem, a rule of thumb found in the literature [22,23] is to use a maximum lag of no more than a third to half of the extent of the data in experimental variogram construction.

The other alternative to ensure a sufficient number of pairs for a given lag distance is to define the lag tolerance in both distance and direction [24]. In fact, the lag tolerance ΔL is a defined distance close to a given lag distance L which is set to validate pair of locations as at the lag distance L . To estimate the variogram at a lag distance L , we collect all pairs of lag distances $L \pm \Delta L$.

$$\hat{\gamma}(\vec{L} \pm \Delta\vec{L}) = \frac{1}{2n(\vec{L} \pm \Delta\vec{L})} \sum_{i=1}^{n(\vec{L})} [x(\vec{u}_i) - x(\vec{u}_i + (\vec{L} \pm \Delta\vec{L}))]^2 \tag{2.3}$$

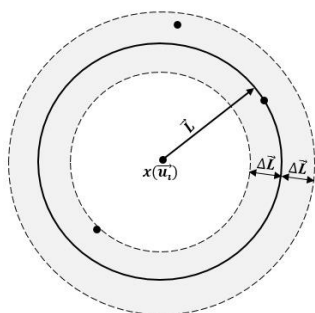
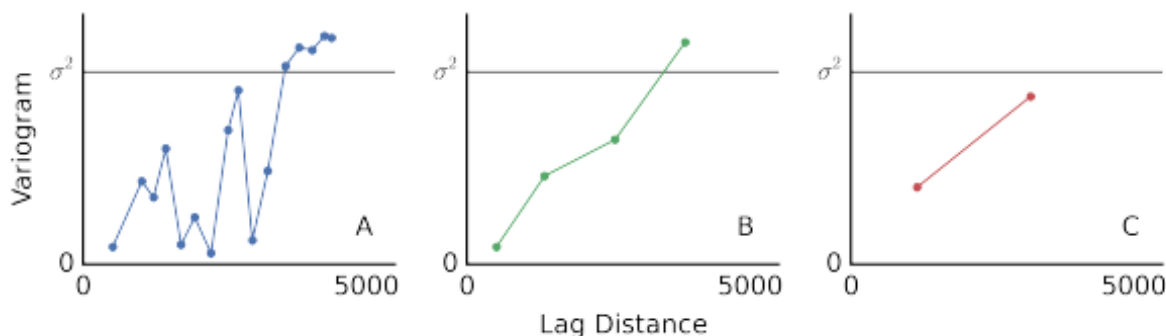


Figure 1. The tolerance ΔL in relation to the lag distance L

The appropriate tolerance in relation to distance and direction is a tough decision, and the key principle is to use an appropriate tolerance to obtain a clearly interpretable variogram structure [25].

In fact, according to Deutsch^[26] and Zhang et al. ^[25], tolerance parameters are required to infer an experimental variogram that is both precise and stable for subsequent modeling. The trade-off between stability and precision is illustrated in the experimental variograms shown in Figure 2. The variogram A is considered to be overly precise and highly unstable, with few pairs informing each point, and the variogram C to be too imprecise for accurate modeling. This study provides a starting point for experimental variogram tolerance parameter selection with the goal of



inferring a precise, stable variogram, such as variogram B.

Figure 2. Three experimental variograms calculated on the same data set with different lag separation distances and lag tolerances ^[26].

In general, spatial structure cannot be obtained by manipulating the tolerance and direction if it did not exist in the data; however, a wrong choice of tolerance could mask the existence of spatial structure, so a tolerance should be chosen in the sense of obtaining the most interpretable variogram possible ^[21].

Practically, the variogram estimation should start with a small tolerance, and if the variogram does not clearly show a defined structure or shows strong fluctuations, the tolerance should be increased; the smallest tolerance from which an interpretable structure can be observed should be used for the variogram estimation. This technique applies to determine whether the tolerance is defined with respect to distance or direction. If the tolerance is so large, the variogram estimate may not be precise, and some detail of the structure may be lost.

Especially if the spatial structure indicates anisotropy, a high tolerance with respect to direction could camouflage this anisotropic structure. It is, however, preferable to start with a variogram of an isotropic structure before looking for the presence of the anisotropic one; the isotropic structure of the variogram shows that a variogram structure is a function of the distance and not of the direction.

2.3. Modeling the variogram

The experimental variogram is not defined everywhere, especially at distances h for which there are no pairs of measurement points. Therefore, a mathematical function called the variogram model is fitted to it. The spherical model is the most widely used of the existing models ^[27,28]. Its equation is written as follows:

$$M_{Sa}(\vec{L}) = \gamma(\vec{L}) = Co \left[\frac{3}{2} \left(\frac{|\vec{L}|}{a} \right) - \frac{1}{2} \left(\frac{|\vec{L}|}{a} \right)^3 \right] \quad \text{for } |\vec{L}| < a \quad (2.4)$$

$$\text{and } M_{Sa}(\vec{L}) = Co \quad \text{for } |\vec{L}| \geq a \quad (2.5)$$

with $M_{Sa}(\vec{L})$ the spherical model which has range a . An equivalent covariance model can be written as follows:

$$C(\vec{L}) = C_0 \left[1 - \frac{3}{2} \left(\frac{|\vec{L}|}{a} \right) - \frac{1}{2} \left(\frac{|\vec{L}|}{a} \right)^3 \right] \quad \text{for } |\vec{L}| < a \quad (2.6)$$

$$\text{and } C(|\vec{L}|) = 0 \quad \text{for } |\vec{L}| \geq a \quad (2.7)$$

2.4. Ordinary kriging

Kriging is a linear combination of neighboring regionalized variables $X(\vec{u}_i)$ variables, each of which is assigned a weighting λ_i in order to estimate variable $X^*(\vec{u}_0)$ at an unsampled location as follows [16,29]:

$$X^*(\vec{u}_0) = \sum_{i=1}^n \lambda_i X(\vec{u}_i) \quad (2.8)$$

Forcing $\sum_{i=1}^n \lambda_i = 1$ and we add Lagrange parameter μ in the sum of variances.

The matrix is then written with the constraint $\sum_{i=1}^n \lambda_i = 1$

$$\begin{pmatrix} c(\vec{u}_1, \vec{u}_1) & \dots & c(\vec{u}_1, \vec{u}_n) & 1 \\ \vdots & & \vdots & \vdots \\ c(\vec{u}_n, \vec{u}_1) & \dots & c(\vec{u}_n, \vec{u}_n) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} c(\vec{u}_1, \vec{u}_0) \\ \vdots \\ c(\vec{u}_n, \vec{u}_0) \\ 1 \end{pmatrix} \quad (2.9)$$

As usual, for calculating λ_i , we write $|\Lambda| = |C|^{-1}|c|$

As λ_i is calculated, we get $X^*(\vec{u}_0)$. And we can also estimate the variance of the error such that

$$\hat{\sigma}_E^2 = c(\vec{u}_0, \vec{u}_0) + \sum_{i=1}^n \lambda_i c(\vec{u}_i, \vec{u}_0) - \mu \quad (2.10)$$

2.5. Block Kriging

Block kriging allows the estimation of the value of unsampled sites over a certain area (2D) or volume (3D)^[30]. As for Cephy, block kriging is especially useful for generating cell properties for subsurface mapping and for estimating the quantity of hydrocarbons in place based on the estimated reservoir properties.

In principle, block kriging calculates the arithmetic mean of several points located in the same cell. Its interest is that it considerably decreases the computational cost of estimating the values on several points within the cell by just calculating the covariances with a certain number of points in the cell and then deducing the arithmetic mean, and it is this mean covariance that is used to characterize the whole cell - so only one matrix to invert per cell.

As for mapping, there is often little difference between the estimates from punctual and block kriging, but the block estimates may seem more reliable because as the size of the block increases, the estimation variances decrease^[31].

3. CASE STUDY OF KINKASI RESERVOIR

3.1. Available data

The available data were from 42 wells spread across Kinkasi field in the DR Congo Coastal Basin and penetrated six Turono-Coniacian layers of the net reservoir that are subject to the present evaluation. For each well and layer, the top and base depths, the mean porosity, and saturation were available.

3.2. Input module

It consists of the following tools:

- Importing data in the file menu
- Linking column headings to boxes at the bottom of the screen

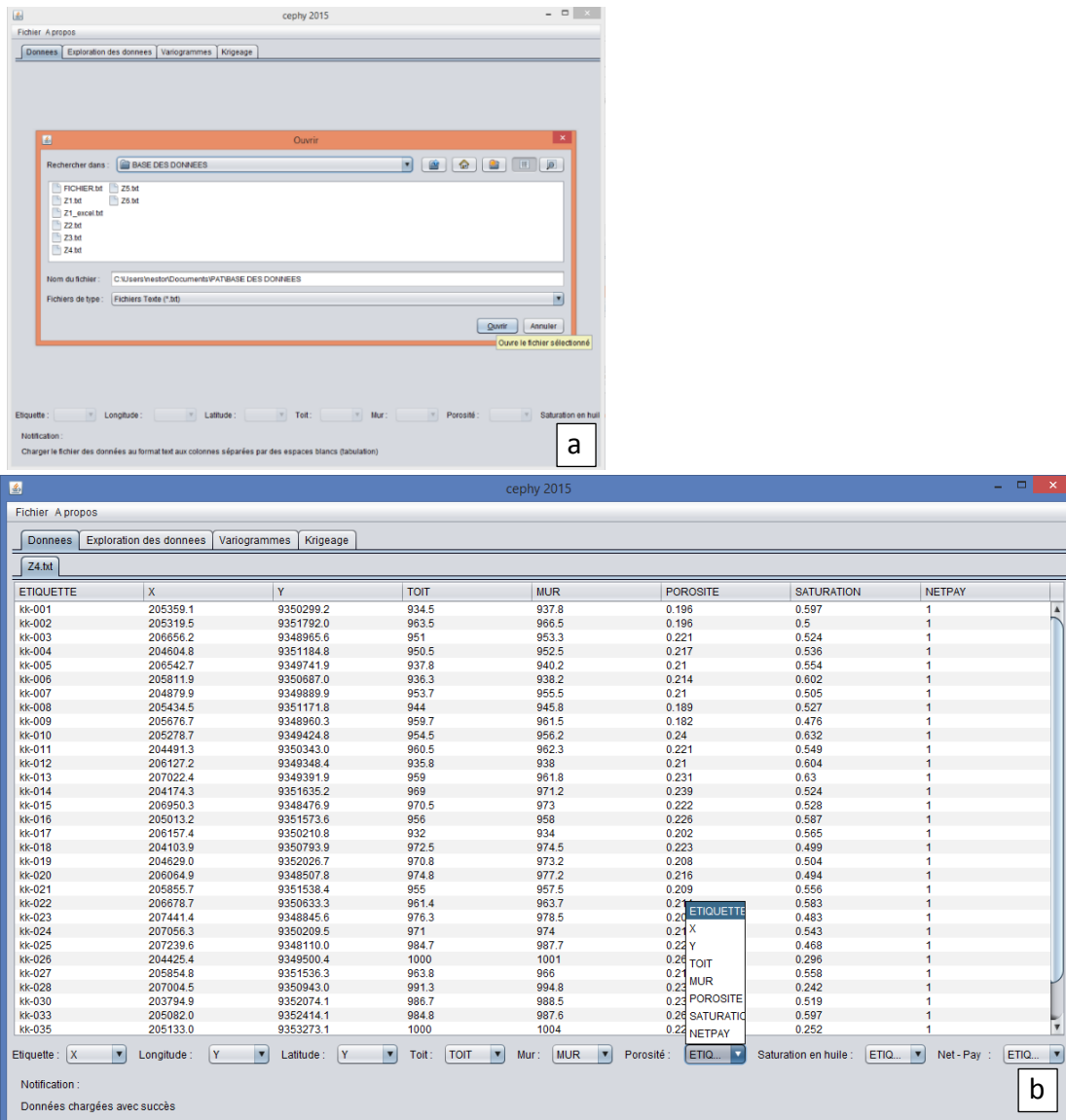


Figure 3. Graphical interface of Cephy: (a) dialog box for data import (b) matching titles to columns.

3.3. Exploration Module

It consists of the following tools shown in Figure 4:

- This tab allows us to visualize the distribution of wells on a left-hand map bordered by UTM coordinates with the possibility of labeling wells by petrophysical parameter values in order to study trends in values.
- The map on the right labels the wells with circular symbols such that the radius of the circle is proportional to the values corresponding to the wells respectively, depending on whether the user wants to view one or another parameter in the drop-down menu on the far right.

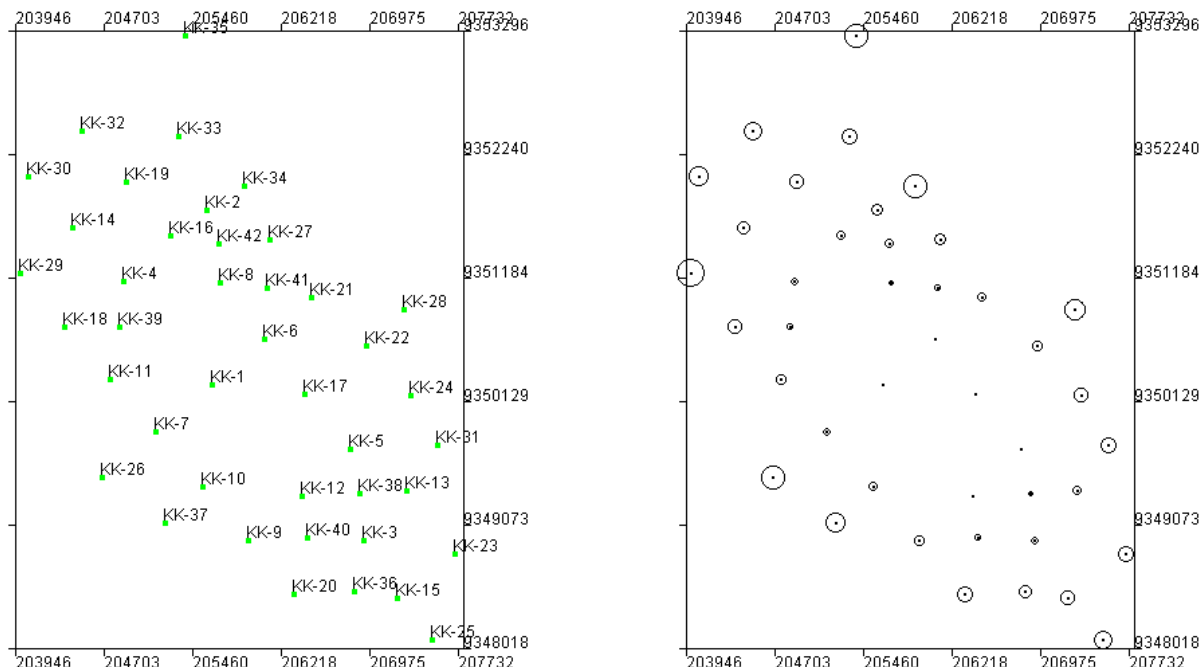


Figure 4. Mapping of all data with symbol width proportional to the magnitude of the sampled property.

This right-hand circular label map allows us to detect whether the data have a variation related to a particular direction in order to exclude the assumption of isotropy of the regionalized variable.

At least we see that no particular direction is followed by the depths of the top; rather, depths are increasing from the center to the reservoir periphery.

3.4. Variogram module

It consists of five sub-tabs (Top, Base, Porosity, Saturation, Net-Pay) which play the same role for different parameters respectively. As an example, we will only show in this text the functioning of the top sub-tab.

By clicking on the display button of this sub-tab, it displays the experimental variogram as well as the curve of the corresponding spherical model (in blue) with the indication of the sill and the range in red (figure 5).

The experimental variogram consists of the points (γ, h) calculated on the basis of the following relation:

$$\hat{\gamma}(\vec{L} \pm \vec{t}) = \frac{1}{2n(\vec{L})} \sum_{i=1}^{n(\vec{L})} [x(u) - (u + (\vec{L} \pm \vec{t}))]^2 \quad (3.1)$$

With \vec{t} being the tolerance.

This tolerance can be adjusted in the drop-down menu on the right. It is by default 0.02% of the maximum value separating all the points and can be adjusted by the user to 0.04%, 0.06%, 0.08% to 0.1%.

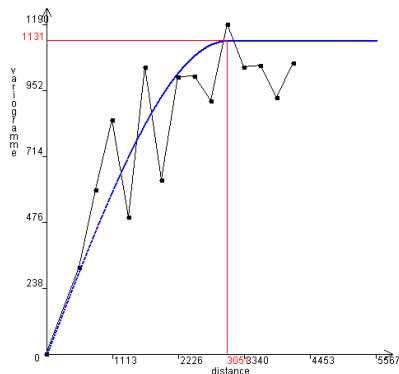


Figure 5. Display of the experimental variogram and visual fitting of the spherical model

The same applies to the minimum number of pairs of sampled stations to take into account a point (γ, h) .

The curve of the spherical model is obtained by linear regression (least square) of the points found by the experimental variogram. After linearization of the spherical model (eq. 2.4) as follows:

$$\frac{\gamma(\bar{L})}{|\bar{L}|} = C_0 \left[\frac{3}{2} \left(\frac{1}{a} \right) - \frac{1}{2} \frac{|\bar{L}|^2}{a^3} \right] \quad (3.2)$$

By posing $\frac{\gamma(\bar{L})}{|\bar{L}|} = Y$; $|\bar{L}|^2 = X$; $\frac{3}{2} \left(\frac{1}{a} \right) C_0 = b$; $\frac{C_0}{2a^3} = a$

We have the following spherical model in the linear form: $Y = aX + b$

At a tolerance of 58 meters of distance with a minimal number of 7 pairs of stations per variogram, we note that our experimental variogram has an interpretable aspect by the spherical model, which is the only model that we implemented in this first version of Cephy.

3.5. Kriging module

This tab has three main features illustrated in figures 6 and 7:

a) Calculate

The calculate button requests the insertion of the Oil-Water Contact (WOC), Gas-Oil Contact (GOC) depths, and the Oil Formation Factor (Boi) to precede the reserve calculation. If the WOC and GOC box is not filled or an incomplete value is entered, the program notifies that it is assumed that OWC and GOC have not been met and proceeds with the calculation in the following manner, while the insertion of Boi is mandatory:

- It retrieves the variogram equation according to the distance corresponding to the last adjustments made in the previous tab (variogram)
- It takes the map set up in the "Data Exploration" tab and subdivides it into small square areas. This mesh is made so that the longitude is subdivided into 225 blocks(pixel).
- For each block, it looks for 4 nearest stations with which it will proceed to the estimation of the value of the block according to the process of the kriging of block aforementioned in section 2.5.
- Then for each block whose estimated top or base values include or are between the spaces in the oil zone, it calculates the amount of oil in place included in that small portion of the block volume by the general oil accumulation formula $OIP = AH_i \phi S_h / B_{oi}$ with A the area.
- Finally, it displays the sum of these small amounts of hydrocarbons obtained in each cell in the OIIP box, and thus the main objective of this study is achieved.

It should be remembered that the kriging algorithms involve the inversion of matrices to obtain the weights λ_i described in equation 2.27. As the number of selected neighboring stations increases, the size of the matrix increases. When the size of the matrix becomes larger, the memory and computation time also becomes more demanding. For example, doubling the number of samples requires four times more memory and eight times more computing time (processor)^[21].

This is where we ask for more patience for users who have to wait for results for up to twenty minutes depending on how good their PC is, but 5 to 10 minutes is enough for a 2.4 GHz processor computer.

b) Show top

This button then allows each block(pixel) to be painted with a corresponding color proportional to the estimated top depth for that block relative to the color scale displayed to the right.

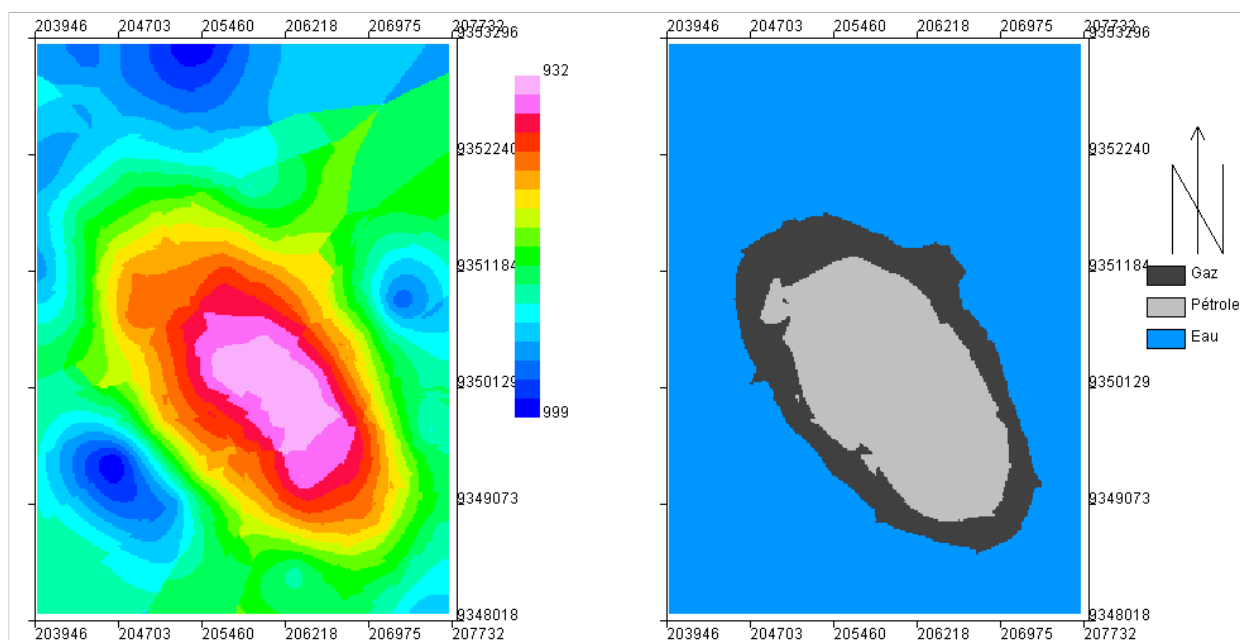


Figure 6. On the left is the bathymetric map of the top of the reservoir; on the right, the 2D oil and gas mapping; and on the extreme right panel, the oil in place quantity.

It also displays in 2D (seen from above) the distribution of water, oil, and possibly gas in the reservoir.

As we could already predict from the data exploration, the structure of our reservoir reflects the shape of an ellipsoidal anticline whose major axis lies in the NW-SE direction through its pinkish colouring in the centre, signifying the shallow depth of the reservoir top and less warm colours towards the peripheries.

c) View OIIP

This button allows you to map the hydrocarbon-impregnated blocks with a color legend bar corresponding to the amount of hydrocarbon estimated on this block.

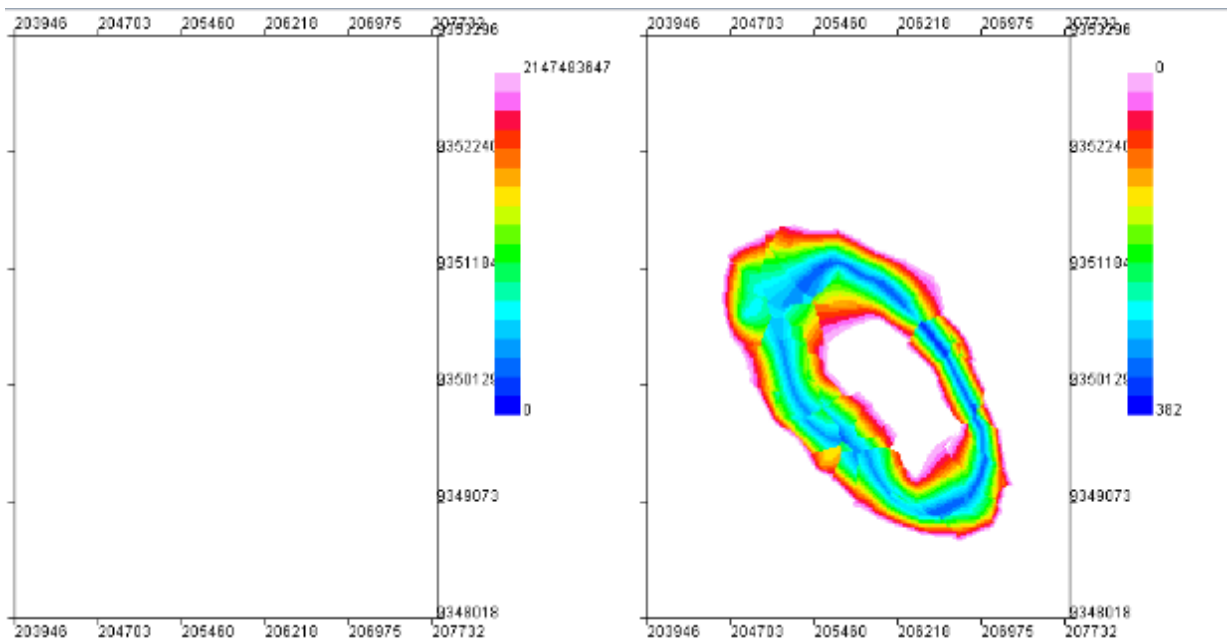


Figure 7. Oil in place distribution in the reservoir

The blank map on the left is for the interpolated net-to-gross ratio map, but unfortunately, the data were not available. The distribution of net pay in the field would serve as a guide for locating highly porous and permeable places to drill.

The map on the right shows the distribution of oil quantity in the field in 2D (top view) per unit of area.

4. RESULTS AND DISCUSSION

A significant amount of oil was found in layers Z2 to Z5, while Z1 contained essentially gas and Z6 water. Table 1. shows the amount of oil generated from interpolation of porosities, saturations, top depth, and base depth in reservoir conditions since the FVF was not available.

Table 1. Amount of oil in place by area

Zones	Accumulations (m ³)	Accumulations (rb)
Z2	417 621	2 631 012
Z3	2 590 267	16 318 682
Z4	1 300 872	8 195 493
Z5	3 300 943	20 795 940
TOTAL	7 663 703	48 281 328

The maps of tops obtained by our software perfectly match the literature on Kinkasi reservoir in Kinkasi field (Coastal Basin of DR Congo), which states that it consists of an ellipsoidal anticline with a long axis of about 7.5 km running NW-SE and a short axis running NE-SW of at least 4 km.

5. CONCLUSION

The focus of this work was to develop a software for automatic evaluation of oil accumulations in a given terrain using the solely well data by ordinary kriging interpolation over the reservoir.

The calculation of the quantity of hydrocarbons in place was carried out by an optimal probabilistic method (kriging) which made it possible to calculate not only the volume of the impregnated reservoir rock but also the petrophysical parameters all over the reservoir by taking into account the spatial variation relationship.

We have opted for ordinary kriging with the change of support (point to surface) to calculate the quantities of hydrocarbons contained in thousands of small virtual blocks representing the ground.

The software is capable of visualizing the results of the variogram analysis as well as maps of the subsurface showing the 2D distribution of liquid and gaseous hydrocarbons and possibly water while giving the geoscientist the possibility to adjust and visualize the variogram model, and depict eventual outlier, clustering or directional trends in data from the exploration tool on the map.

An application of the Cephy tools on the Kinkasi reservoir (Coastal Basin of DR Congo) demonstrated its usefulness by simplifying kriging in one way, straight forward steps. The use of the exploration tool facilitated understanding of the distribution of data in the field and helped to depict an outlier. This tool is a framework for further modeling the processes of the subsurface by the scientific community, public administration, and the private sector for a wide range of reservoir modeling, namely petrophysical properties distribution throughout the reservoir.

The software is proposed in its first version, and further development will address 3D modeling with the integration of dynamic data such as permeability and production history in order to make a micro numerical reservoir simulator that is simple, light, free, and fast.

DECLARATION

Competing Interests: On behalf of all the co-authors, the corresponding author states that there is no conflict of interest.

REFERENCES

1. Masoudi, P., Ali, M., Farshid, A., & SeyedMohammad, Z. (2011). Estimation of in place hydrocarbon volume in multilayered reservoirs using deterministic and probabilistic approaches. *ENERGY EXPLORATION & EXPLOITATION* , 29, 543–558.
2. Ibrahim, M. M., Abdulaziz, A. M., & Fattah, K. A. (2018). STOIP validation for a heterogeneous multi-layered reservoir of a mature field using an integrated 3D geo-cellular dynamic model. *Egyptian Journal of Petroleum*, 27(4), 887–896. <https://doi.org/10.1016/J.EJPE.2018.01.004>
3. Ochoma, U., Ochu, E. D., & Horsfall, O. I. (2020). Deterministic Hydrocarbon Volume Estimation in the Onshore Fuba Field Niger Delta, Nigeria. *IOSR Journal of Applied Geology and Geophysics*, 8, 34–40. <https://doi.org/10.9790/0990-0801033440>
4. Dake, L. P. (1983). *Fundamentals of Reservoir Engineering, Volume 8 1st Edition*. 464. <https://www.elsevier.com/books/fundamentals-of-reservoir-engineering/dake/978-0-444-41830-2>
5. Ahmed, T. (2019). Reservoir Engineering Handbook-Tarek Ahmed-5th Edition. *Oxford: Elsevier Inc.*, 5(4), 53.
6. Fanchi, J. R. (2010). Geology. *Integrated Reservoir Asset Management*, 33–48. <https://doi.org/10.1016/B978-0-12-382088-4.00003-7>
7. Sharma, V., Sharma, M., Pandita, S., Kour, J., & Sharma, N. (2021). Application of geographic information system and remote sensing in heavy metal assessment. *Heavy Metals in the Environment*, 191–204. <https://doi.org/10.1016/B978-0-12-821656-9.00011-0>
8. Shepard, D. (1968). A two-dimensional interpolation for irregularly-spaced data function. *In Proceedings of the 1968 23rd ACM National Conference, ACM*, 517–524.

9. Burrough, P. A., & McDonnell, R. (1998). *Principles of geographical information systems*. Oxford University Press.
https://books.google.com/books/about/Principles_of_Geographical_Information_S.html?hl=fr&id=shHznQEACAAJ
10. Matheron, G. (1962). *Traité de géostatistique appliquée*. Technip.
11. Dauphiné, A. (2017). Models of Basic Structures: Points and Fields. *Geographical Models with Mathematica*, 163–197. <https://doi.org/10.1016/B978-1-78548-225-0.50010-5>
12. Magesh, N. S., & Elango, L. (2019). Spatio-temporal variations of fluoride in the groundwater of Dindigul District, Tamil Nadu, India: A comparative assessment using two interpolation techniques. *GIS and Geostatistical Techniques for Groundwater Science*, 283–296. <https://doi.org/10.1016/B978-0-12-815413-7.00020-1>
13. Bai, T., & Tahmasebi, P. (2021). Accelerating geostatistical modeling using geostatistics-informed machine Learning. *Computers and Geosciences*, 146. <https://doi.org/10.1016/j.cageo.2020.104663>
14. Fanchi, J. R. (2002). Modern Reservoir Characterization Techniques. *Shared Earth Modeling*, 182–198. <https://doi.org/10.1016/B978-075067522-2/50011-2>
15. Ramezanali, A. K., Feizi, F., Jafarirad, A., & Lotfi, M. (2020). Geochemical Anomaly and Mineral Prospectivity Mapping for Vein-Type Copper Mineralization, Kuhshah-e-Urmak Area, Iran: Application of Sequential Gaussian Simulation and Multivariate Regression Analysis. *Natural Resources Research*, 29(1), 41–70. <https://doi.org/10.1007/s11053-019-09565-7>
16. Matheron, G. (1970). La théorie des variables régionalisées et ses applications. *Les Cahiers Du Centre de Morphologie Mathématique de Fontainebleau*.
17. Shekhar, S., Li, Y., Ali, R. Y., Eftelioglu, E., Tang, X., & Jiang, Z. (2018). Spatial and Spatiotemporal Data Mining. *Comprehensive Geographic Information Systems*, 3, 264–286. <https://doi.org/10.1016/B978-0-12-409548-9.09594-4>
18. Kigobe, M., & Kizza, M. (2006). Dealing With Spatial Variability Under Limited Hydrogeological Data. Case Study: Hydrological Parameter Estimation in Mpigi-Wakiso. *Proceedings from the International Conference on Advances in Engineering and Technology*, 211–220. <https://doi.org/10.1016/B978-008045312-5/50024-8>
19. Hosseini, S. M. (2017). Parameter estimation of aquifer transmissivity variogram using WLS, GA and PSO algorithms. *International Journal of Hydrology Science and Technology*, 7(2), 103–123. <https://doi.org/10.1504/IJHST.2017.084138>
20. Calder, C. A., & Cressie, N. (2009). Kriging and Variogram Models. *International Encyclopedia of Human Geography*, 49–55. <https://doi.org/10.1016/B978-008044910-4.00461-2>
21. Kelkar, M., & Perez, G. (2002). *Applied Geostatistics for Reservoir Characterization* (SPE).
22. Oliver, M. A., & Webster, R. (2015). *Basic Steps in Geostatistics: The Variogram and Kriging* (SPRINGER). SPRINGER BRIEFS IN AGRICULTURE. <http://www.springer.com/series/10183>
23. Pyrcz, M. J., & Deutsch, C. v. (2014). *Geostatistical reservoir modeling: Vol. 2nd ediion* (Oxford). Oxford University Press.
24. Drumond, D. A., Rolo, R. M., & Costa, J. F. C. L. (2019). Using Mahalanobis Distance to Detect and Remove Outliers in Experimental Covariograms. *Natural Resources Research*, 28(1), 145–152. <https://doi.org/10.1007/s11053-018-9399-y>
25. Zhang, S. E., Nwaila, G. T., Tolmay, L., Frimmel, H. E., & Bourdeau, J. E. (2021). Integration of Machine Learning Algorithms with Gompertz Curves and Kriging to Estimate Resources in Gold Deposits. *Natural Resources Research*, 30(1), 39–56. <https://doi.org/10.1007/s11053-020-09750-z>

26. Deutsch, J. L. (2015). *Experimental Variogram Tolerance Parameters Learning Objectives* (J. Deutsch, Ed.). Geostatistics Lessons. <http://geostatisticslessons.com/lessons/variogramparameters>
27. Sarma, D. D. (2009). *Geostatistics with Applications: Vol. 2nd Edition* (Springer).
28. Isaaks, E. H., & Srivastava, R. M. (1989). *Applied Geostatistics*.
29. Kholghi, M., & Hosseini, S. M. (2006). Estimation of Aquifer Transmissivity using Kriging, Artificial Neural Network, and Neuro-Fuzzy models. *Journal of Spatial Hydrology*, 6(7).
30. Kang, J., Jin, R., Li, X., & Zhang, Y. (2017). Block kriging with measurement errors: A case study of the spatial prediction of soil moisture in the middle reaches of Heihe river basin. *IEEE Geoscience and Remote Sensing Letters*, 14(1), 87–91. <https://doi.org/10.1109/LGRS.2016.2628767>
31. Oliver, M. A., & Webster, R. (1990). Kriging: A method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, 4(3), 313–332. <https://doi.org/10.1080/02693799008941549>

*Corresponding author E-mail: patient.mulekya@unikin.ac.cd