https://doi.org/10.33271/nvngu/2023-5/141

M. Mazari[*1,2],
orcid.org/0000-0002-7745-6845,
S. Chabou-Mostefai[1],
orcid.org/0000-0001-7024-7066,
A. Bali[3],
orcid.org/0000-0002-4699-849X,
K. Kouider[4,5],
orcid.org/0009-0005-0553-6116,
A. Benselhoub[6],
orcid.org/0000-0001-5891-2860,
S. Bellucci[7],
orcid.org/0000-0003-0326-6368

1 – Department of Mining Engineering, National Polytechnic School of Algiers, Algeria, Algeria
2 – Department of mines and geology, Faculty of Technology, Abderrahmene Mira University, Bejaia, Algeria
3 – Department of civil Engineering, Laboratory of Civil Engineering Materials and Environment "LMGCE", National Polytechnic School of Algiers, Algeria, Algeria
4 – Water Sciences Research Laboratory-National Polytechnic School of Algiers, Algeria, Algeria
5 – Faculty of Civil Engineering, University of Sciences and Technology Houari Boumediene Algiers, Algeria, Algeria
6 – Environment, Modeling and Climate Change Division, Environmental Research Center (C.R.E), Annaba, Algeria
7 – INFN-Frascti National Laboratories, Frascati, Rome, Italy
[*] Corresponding author e-mail: messaoud.mazari@g.enp.edu.dz

MINERAL RESOURCE ASSESSMENT THROUGH GEOSTATISTICAL ANALYSIS IN A PHOSPHATE DEPOSIT

Purpose. The selection of an appropriate variographic model is crucial in geostatistics to obtain accurate estimates of mineral reserves. The aim of this work is to develop a reserve estimation tool using a geostatistical approach.

Methodology. The geostatistical approach is based on selecting the most representative variographic models for the studied variables. The model selection is done by applying a cross-validation procedure leave-one-out (LOOCV). LOOCV is a resampling technique used in statistical analysis and machine learning to estimate the generalization error of a model and compare the performance of different models. The studied variables are then estimated using ordinary kriging.

Findings. The application of the proposed approach has resulted in satisfactory results in terms of dispersion of grades and thicknesses of mineralized layers in a phosphate deposit. To evaluate the quality of the adjustment models obtained, efficiency factors such as Nash-Sutcliffe, and RMSE (Root Mean Square Error), were employed. These factors provide quantitative measures of the agreement between the observed and predicted values. The NSE (Nash-Sutcliffe efficiency) and RMSE (root mean square error) values of 0.572 and 6.599, respectively, indicate a better fit and greater accuracy of the adjustment models. The accuracy and efficiency criteria of the studied variables have acceptable values, with a mean square error (MSE) of $1.54 \cdot 10^{-7}$.

Originality. The combination of the least squares and LOOCV methods in the geostatistical analysis leads to improved estimation precision, greater reliability in representing the spatial variability of the parameters, and enhanced confidence in the validity of the adjustment models.

Practical value. The development of a computer code for this geostatistical approach provides a practical tool for decision-makers to use in the management and exploitation of mining sites. Overall, this study has contributed to the advancement of geostatistical techniques and their application in the mining industry.

Keywords: *Bled el Hadba deposit, cross validation (LOOCV), geostatistics, kriging, mineral reserves*

Introduction. A mining project is a complex system involving any geological, geotechnical, metallurgical, mining, environmental, economic, legal and social variables. All these variables must be estimated in order to provide a database for evaluating the mining project [1].

The success of a mining project depends on accurately recognizing the subsoil, which involves managing the risks of geological uncertainty. However, accurately estimating the grades of minerals within a deposit is crucial in the mining industry and is used in several stages of mining, from exploration to exploitation.

Resource estimation is an essential step in feasibility studies and mine planning. Although advanced methodologies exist, they may not be suitable for every complex geological environment. Various researchers have proposed grade prediction models using techniques such as inverse distance weighing, kriging, and stochastic simulation [2]. Nevertheless, selecting the appropriate methodology for resource estimation in a mining project depends on various factors such as the complexity of the geological environment, data availability, and the level of accuracy required. In any case, the accuracy and reliability of the grade estimates are critical in determining the economic feasibility of a mining project, and therefore, the selection of the appropriate methodology must be carefully evaluated.

Detailed and extensive exploration operations are required to obtain geological models, which are used to accurately describe the ore body and estimate mineral reserves [3].

The assessment of mineral resources often requires the use of quantitative approaches, especially in creating a geological model. This model serves as the foundation for all mining activities such as mine planning, design, production scheduling, and development. It also plays a significant role in investment decisions [4].

In the present study, predictive data mining algorithms were applied to the Bled El Hadba deposit (Eastern Algeria) [5], to predict the probability of encountering ore in estimation maps, and to estimate the mineral reserves of the main layer of phosphate ore (the median).

Previous studies on the Bled El Hadba deposit showed significant differences in reserve estimation results. The first study estimated reserves of approximately 103 million tons, while the second study estimated reserves of around 133.6 million tons for the same site.

The aim of this article is to develop a reserve estimation tool using a geostatistical approach [6], to improve the quality of the estimation [7]. The tool includes an adjustment and a validation step of variographic models, which is a crucial phase integrated into the developed tool structure. A leave-one-out (LOOCV) cross-validation procedure [8] is used to systematize the treatment of variographic models and to select the most representative variograms.

Leave-one-out cross-validation (LOOCV) is a resampling technique used in statistical analysis and machine learning for model selection and validation [8]. The LOOCV method involves splitting the dataset into training and testing subsets, where in each iteration of the process, one sample from the dataset is selected as the testing subset, and the remaining samples are used as the training set. The model is trained on the training set and evaluated on the testing set, and the process is repeated for all samples in the dataset [8]. This results in a set of performance measures that can be averaged to give an estimate of the model's accuracy [9]. The LOOCV method is useful for estimating the generalization error of a model, which is the expected error when the model is applied to new, unseen data. It is also commonly used for comparing the performance of different models, as well as for selecting the optimal model parameters. In the case of reserve estimation for the Bled El Hadba phosphate deposit, LOOCV was used to select the best variographic models to be used for kriging. By systematically testing different variographic models using LOOCV, we were able to identify the most representative variograms for accurate reserve estimation.

Despite the many benefits of LOO, scaling this approach to large datasets can be challenging [10]. The naive approach of LOO involves computing n posteriors, which can become computationally expensive in situations where n is large [9]. Even the computation of a single posterior can be time-consuming, making it difficult to scale the approach to large datasets.

The current study holds both technical and economic significance by providing a sound evaluation of mining reserves, which in turn contributes to the enhancement of grade control, efficient management of extraction process, and effective ore processing.

Material and methods. ***Geological setting.*** The Bled El Hadba deposit is located 14 km south-east of Bir El Ater and 6 km from the Algerian-Tunisian border. The area of the zone recognized by exploratory boreholes is approximately 2.9 km^2.

Structurally, the Bled El Hadba area constitutes the western flank of the antiform structure of Jebel Zrega, whose crest line forms the Algerian-Tunisian border. This mining area is located symmetrically with respect to the southern flank of Jebel Onk (Djemi Djema and Kef Es Sennoun), whose phosphate layer is on average about 40 m thick [11].

The wall and roof structures of the phosphate bundle (Fig. 1) illustrate well the monoclinal dip, towards the West of the phosphate series. Several horizontal setbacks, NW-SE, are cartographically visible, but they do not cause significant changes in the phosphate layer geometry [5].

The phosphate deposit geology is relatively simple; it was described by Dussert (1924). The geological map of the region shows the phosphate layers of the Thanetian age, under the Ypresian flint limestones and the Miocene sands (Fig. 1), plunge in a monocline fashion and under a gentle slope of 6 to 10° towards the West and the North-West, this dip becomes more accentuated towards the South of the deposit (Fig. 2) [5].

The used approach description. In order to facilitate and improve the mineral reserves estimation, a geostatistical approach is used [7]. This approach involves several steps describing the kriging interpolation process, associated with an adjustment-validation coupling [8].

A MATLAB calculation code is implemented to automate the processing and improve the reserve calculation procedure performance.

Adjustment method. The least squares method is used for the adjustment by comparing the experimental data with a mathematical model supposed to describe these data [12].

In the case of nonlinear least squares it is about minimizing the function

$$\min_{x \in R^n} g(x) = \frac{1}{2} r(x)' r(x) = \frac{1}{2} \sum_{i=1}^{m} r_i(x)^2, \quad m \geq n, \tag{1}$$

with $r_i(x)$ $i = 1, ..., m$, is a nonlinear functional defined over R^n, where $r(x)$ is the vector of residuals dependent on the parameters x.

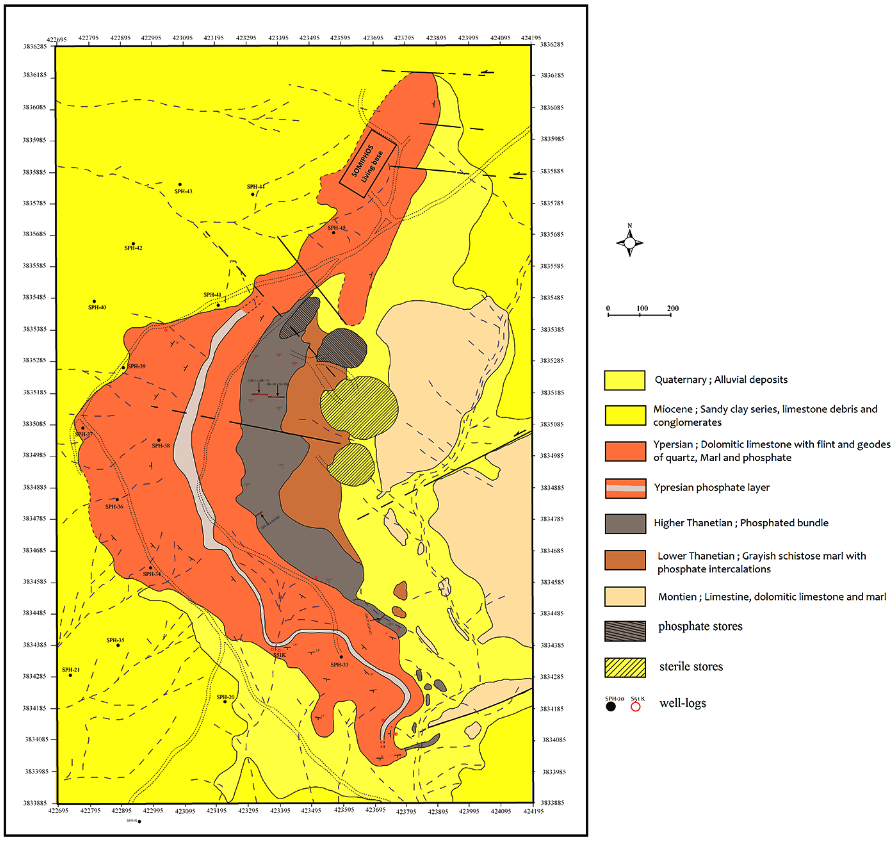

Fig. 1. Geological map of the structure of the Bled El Hadba phosphate deposit (Eastern Algeria) [5]

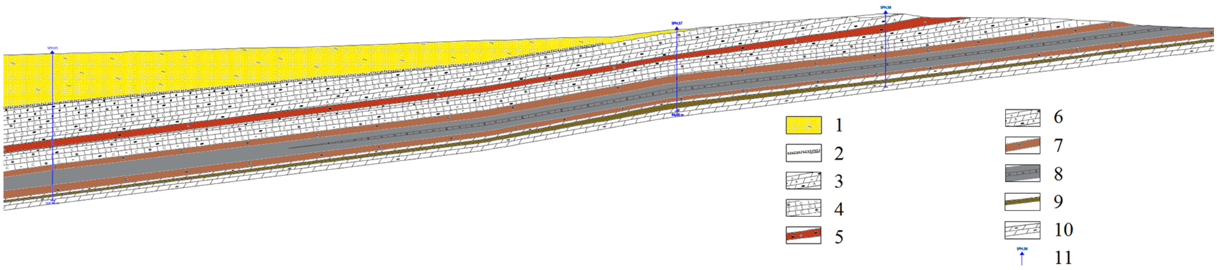

Fig. 2. Geological Profile oriented east–west [5]:

1 – Miocene detrial deposits (sands and clays); 2 – Miocene basal conglomerate; 3 – Whitish organic limestone, strongly gypsum; 4 – Lumichellic limestone, residually phosphate, with flint nodules and quartz geodes; 5 – Phosphate with nodules of flint, and debris of limestone and marl; 6 – Gypsum marl, whitish, with flint nodules, residually phosphate; 7 – phosphated layer of the higher thanetian (CS + CB: lumachellic phosphate with quartz geodes); 8 – CM phosphated layer of the higher thanetian; 9 – Pelitic marl bedded, residually phosphate; 10 – Coprolithic phosphate, marly cement; 11 – Well-logs (exploratory boreholes)

With

$$r_i(x) = y_i - f(t_i, x), \quad i = 1, \ldots, m,$$

where $f(t_i, x)$ is a nonlinear function (the model) with ti the independent variables and $x \in R^n$ the parameter vector to be estimated.

In order to write the quadratic model for the minimization of equation (1) we need the first and second derivatives of $g(x)$.

The first derivative is written

$$\nabla g(x) = \sum_{i=1}^{m} r_i(x) \cdot \nabla r_i(x) = r(x)' r(x), \tag{2}$$

with

$$\nabla r(x) = \begin{bmatrix} \frac{\partial r_1(x)}{\partial x_1} & \ldots & \frac{\partial r_1(x)}{\partial x_n} \\ \vdots & \ldots & \vdots \\ \frac{\partial r_m(x)}{\partial x_1} & \ldots & \frac{\partial r_m(x)}{\partial x_n} \end{bmatrix},$$

where $\nabla r(x)$ is the Jacobian matrix.

The vector $\nabla r_i(x) = \begin{bmatrix} \frac{\partial r_i(x)}{\partial x_1} \\ \vdots \\ \frac{\partial r_i(x)}{\partial x_n} \end{bmatrix}$ corresponds to the line i of the Jacobian matrix.

The second derivative is written as

$$\begin{aligned} \nabla^2 g(x) &= \sum_{i=1} (\nabla r_i(x) \cdot \nabla r_i(x)' + r_i(x) \cdot \nabla^2 r_i(x)) = \\ &= \nabla r(x)' \nabla r(x) + S(x), \end{aligned} \tag{3}$$

with $S(x) = \sum_{i=1}^{m} r_i(x) \cdot \nabla^2 r_i(x)$.

The Gauss-Newton method uses an approximation of the matrix of second derivatives (3) omitting the term $S(x)$. As $S(x)$ is composed of a sum of terms $r_i(x)\nabla^2 r_i(x)$, this simplification is justified in a situation where the residuals $r_i(x)$ are small.

Cross validation LOOCV. The Leave-one-out Cross-Validation (LOOCV) method [8, 9] is based on dividing the data set under study into two parts, an observation pair for validation, i.e. (x_1, y_1) and the rest of the samples $(x_2, y_2), \ldots, (x_n, y_n)$ as learning samples. Note that the pair (x_1, y_1) is not used for model fitting. The value x_1 and the estimation function $\hat{f}_1$ are used to find the value of $\hat{y}_1$. Then, the mean-square error (MSE) of the test for this observation couple is calculated as follows

$$MSE_1 = (y_1 - \hat{y}_1)^2.$$

This process is repeated for each pair (x_i, y_i), $i = 1, \ldots, n$.

The test error is calculated for each i

$$MSE_i = (y_i - \hat{y}_i)^2. \tag{4}$$

The test error for the LOOCV method is the average of the errors

$$CV = \frac{1}{n} \sum_{i=1}^{n} MSE_i. \tag{5}$$

Models evaluating criteria. The model to be used is chosen according to the minimum value of MSE. The existing agreement between experimental and theoretical variogram values is assessed using both the efficiency indicator (Nash-Sutcliffe efficiency NSE) [13] and the precision with the root-mean-square error (RMSE) indicator, and RMSE-observations standard deviation ratio (RSR)[14].

The Nash-Sutcliffe efficiency (NSE) criterion is calculated as follows [15]

$$NSE = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}. \tag{6}$$

The NSE range is between $-\infty$ and 1 (perfect fit) [13].

A simple way to obtain the model accuracy consists of calculating RMSE (root-mean-square error) [15]. The RMSE represents the differences between the values predicted by a model and the values actually observed.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}. \tag{7}$$

The RSR is calculated by the ratio between the RMSE and the standard deviation of the measured data [14]

$$RSR = \frac{\sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}. \tag{8}$$

The RSR varies from the optimal value of 0, which indicates zero RMSE or a residual variation and therefore a perfect model simulation, to a high positive value. The lower is the RSR, the lower is the RMSE indicating a better match between the observed and simulated data.

Principal Features of the variogram model. The variogram is a fundamental tool in geostatistics for describing the spatial structure of regionalized variables. It serves as the foundation for prediction and simulation algorithms and provides valuable insights into the properties of these variables. It is characterized by parameters such as the range, nugget effect, sill, and variograph (Fig. 3).

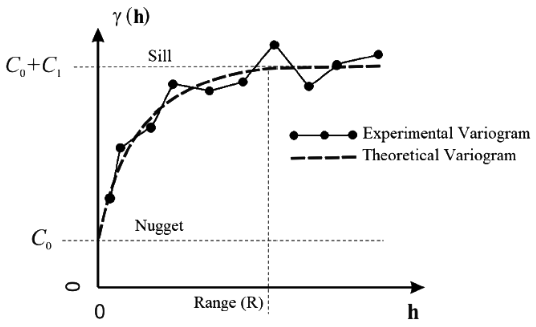

Fig. 3. Variogram parameters [16]:

a — Range; C_0 — Nugget effect; C_0 + C — Sill variance; γ(h) — Semi-variance; h — Distance (m)

The variogram is constructed by plotting the semivariance of the attribute against the distance between data points. This resulting graph is then fitted with a mathematical function, like spherical, exponential, or Gaussian models, to characterize spatial correlation.

Estimation method. The variographic model chosen at the outcome of the cross-validation is used to interpolate by kriging [4], the studied variables, where they were not sampled, and subsequently to evaluate the mine reserves by the geostatistical method (ordinary kriging)[7].

The elaborated tool structure. The calculation code is developed as nested loops dealing with different phases of a geostatistical calculation. The observations outcoming from the exploration campaigns are the input data of the developed model.

The calculations are structured as follows:

1. Calculation of experimental variograms from survey data introduced, followed by the theoretical models adjustment using the least squares method (iterative calculation).

2. Identification of the best model of the adjusted variogram using the cross validation method (LOOCV): The experimental variogram is established from the $n - 1$ samples, the remaining sample is used to compare the estimated variables to the sampled variables. This procedure is repeated n times corresponding to the number of samples.

3. Discretization of the study area and establishment of the mesh matrix according to the chosen space step.

4. Establishment of the kriging matrices corresponding to different mesh points, thus making it possible to estimate the variables and their estimation variances.

5. Display of estimation results in graphical and tabular forms.

The geostatistical approach application. *Application case.* After having implemented the geostatistical approach, it is necessary to implement it on a real case, in order to test the followed approach performance.

The Bled El-Hadba deposit (North-East of Algeria) was chosen to study the spatial variability of P_2O_5 contents and thicknesses. We are interested in the study of the deposit main layer, which corresponds to the middle layer.

The geochemical data (61 holes) outcoming from the deposit exploration campaigns are used. These samples are distributed in an irregular manner; the inter-sample distance is quite large, reaching up to 500 m in some places.

Results. The geostatistical approach application to the Bled-El-Hadba deposit allowed us to estimate phosphate reserves, based on available geochemical data.

The calculation results of the phosphate grades and thicknesses of the study area are presented below in the form of variographic models and estimation maps.

Phosphate layer thicknesses estimation. Theoretical and experimental variograms derived from the LOO cross-validation are presented in the following figure.

Experimental variograms showed the influence of some samples in the variance calculation (the difference can reach 15). The theoretical variograms adjustment to the experimental data shows a good fit following to the exponential model used (Fig. 4).

After calculating the variograms and applying the validation procedure, it is a question of choosing the most representative model (Fig. 5).

This variogram reaches a sill of $C(0) = 35.6677$ for a maximum range of $a = 1380$ m. The chosen model shows a continuity near the origin illustrated by a nugget effect ($C_0 = 0$).

The adequacy between theoretical and experimental models is quantified by the evaluating criteria of the fit quality (MSE, RMSE, NSE, RSR). The results are shown in Table 1.

The concordance between theoretical and experimental models is evaluated by the RMSE, NSE and RSR criteria, which display allowable values (Table 1).

The mean squared error (MSE) is vital in the choice of the best validation model. As a result, the chosen model has a minimum MSE of 0.036.

The theoretical model chosen is used in an Ordinary Kriging procedure in order to evaluate the dispersion of the phosphate layer thicknesses in the study area (Fig. 6).

This map shows the thicknesses distribution in the deposit, which vary between 8 and 25 meters

In order to evaluate the quality of the estimation results, it is necessary to calculate the estimation variance (Fig. 7).

The sampled area has an acceptable estimation quality (variance < 35.66), where the best quality is found in the cen-

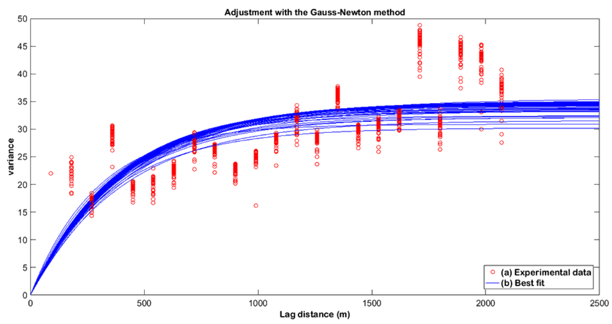

Fig. 4. Phosphate layer thicknesses validation variograms:

a — Experimental data; b — Best fit

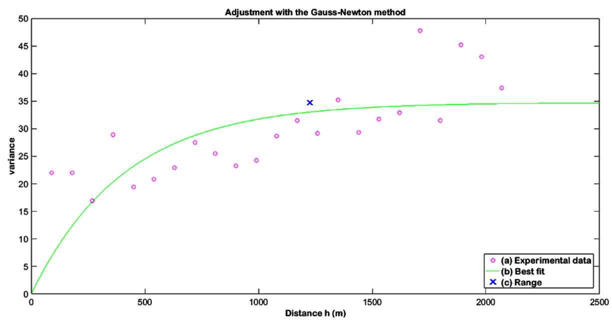

Fig. 5. Phosphate layer thicknesses variographic model selected:

a — Experimental data; b — Best fit; c — Range

Table 1

Quality criteria of fitted models for thicknesses of the phosphate layer

		MSE	RMSE	NSE	RSR
Theoretical models	MIN	0.036	6.226	0.318	0.514
	MAX	182.335	8.851	0.694	0.845
	MEAN	23.037	8.205	0.570	0.646
Retained model	LOOCV	0.036	6.599	0.572	0.642

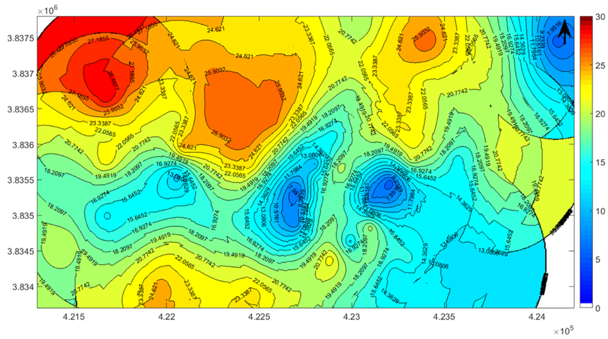

Fig. 6. Phosphate layer thickness dispersion map

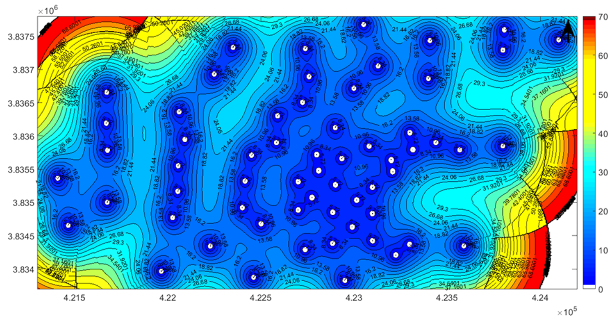

Fig. 7. Phosphate layer thicknesses estimating variance map

tral deposit zone (variance < 12), characterized by a higher sampling density. A proportionality relationship is established between the estimation variance and the samples number.

The phosphate content estimation. The geostatistical approach application in the spatial variability study on the grades gave similar results to those of the thicknesses, corresponding to cross-validation variograms (Fig. 8), variographic model retained (Fig. 9), content dispersion map (Fig. 10), content variance map (Fig. 11).

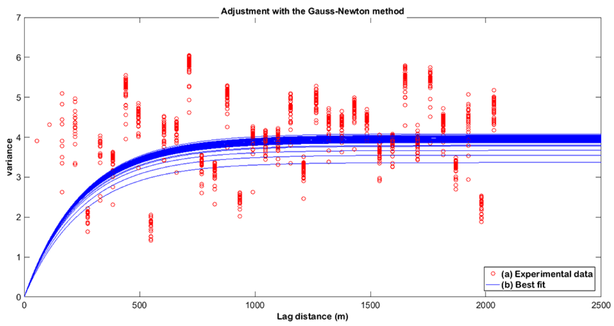

Fig. 8. Phosphate content Cross-validation Variograms: a — Experimental data; b — Best fit

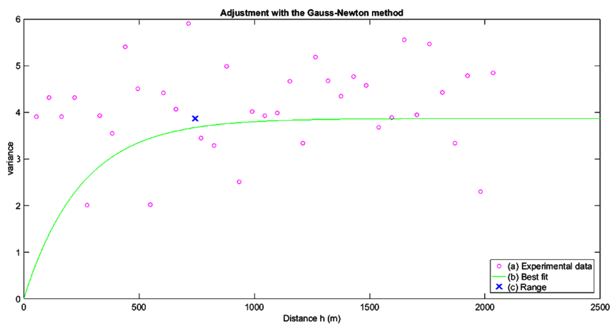

Fig. 9. Phosphate content retained variographic model: a — Experimental data; b — Best fit; c — Range

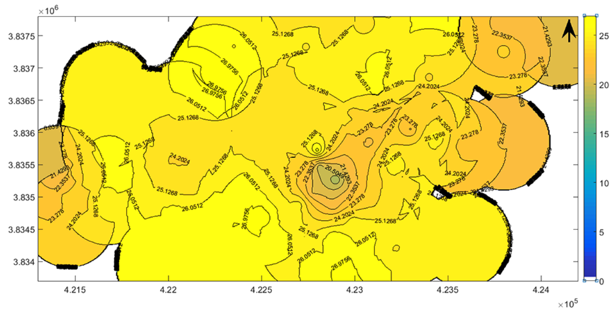

Fig. 10. The phosphate content dispersion map

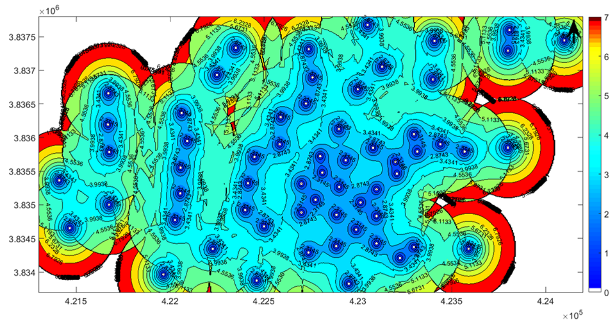

Fig. 11. Estimating the phosphate content variance map

The range value corresponding to the retained phosphate contents variogram has considerably decreased (742 m) compared to that of the thicknesses variogram. The level of this variogram reaches a sill of $C_0 = 3.87$.

The assessing criteria of the content fit quality leads to allowable values (Table 2).

Reserves estimation. The phosphate reserves of the deposit middle layer (Table 3) are calculated from the thicknesses and contents. These are estimated by applying a geostatistical method associated with a cross-validation procedure allowing improving accuracy by choosing the best variographic model.

Discussion. The application of a geostatistical approach allows improving the quality of the reserve estimation, by studying the variographic model choice. This approach is developed as a calculation code under Matlab, associating the various adjustment steps of validation and kriging.

The models used for kriging are chosen using a cross-validation procedure (LOOCV), which improves the quality of ordinary kriging results (Figs. 6, 7, 10 and 11). The leave-one-out cross-validation application, allows the generation of a large number of variograms (Figs. 4 and 8) according to the number of samples used, thus allowing us to study the spatial variability grade and thicknesses of the mineralized layer.

Table 2

Quality criteria of fitted models for phosphate content

		MSE	RMSE	NSE	RSR
Theoretical models	MIN	$1.54 \cdot 10^{-7}$	34.791	0.084	0.883
	MAX	214.554	35.883	0.281	1.076
	MEAN	23.053	34.900	0.199	0.990
Retained model	LOOCV	$1.54 \cdot 10^{-7}$	34.953	0.248	0.989

Table 3

The phosphate reserves estimation

The mineralized rock reserves (million m^3)	Phosphate reserves (million m^3)	Phosphate Content
81.185	20.255	%

The use of different criteria such as NSE, RMSE, and RSR in the adjustment phase of theoretical models to the experimental variograms allows for the evaluation of the fit quality and estimation error. The NSE criterion provides information about the goodness of fit, while RMSE and RSR quantify the estimation error. By combining these criteria, an overall assessment of the quality of the model can be obtained.

Adapting the NSE criterion for quality improvement not only provides engineering information about the system, but also offers a comprehensive solution for quality engineering [17]. This means that the NSE criterion can be used to evaluate the quality of the models used in this study and provide information about the system being modeled. Furthermore, the NSE criterion can be used as a tool for quality improvement in the mining industry, providing a comprehensive solution for quality engineering.

According to the results presented in Tables 1 and 2, the models used in this study provide reliable results for both the thicknesses variograms and phosphate grades variograms, with median NSE values of 0.570 and 0.199, respectively, and RSR values of 0.646 and 0.990, respectively. The choice of the best variogram is based on the quantification of the minimum square error of the thicknesses and grade variograms.

The efficiency and precision criteria also display allowable values, indicating satisfactory adjustment quality of the experimental data to the theoretical models. Therefore, leave-one-out cross validation can be used to estimate and compare the performance of the models.

The above statement indicates that the criteria used to evaluate the quality of the adjustment of experimental data to theoretical models are satisfactory. The mean-square error (MSE) of the thickness and grade variograms is 0.036 and $1.54 \cdot 10^{-7}$, respectively, indicating that the models are well-fitted to the data (Figs. 2 and 6). Additionally, the NSE (Nash-Sutcliffe efficiency) and RMSE (root mean square error) values of 0.572 and 6.599, respectively, indicate that the models provide reasonable estimates of the observed data. Overall, the results suggest that the theoretical models provide a good representation of the observed data and can be used to make accurate predictions.

The use of leave-one-out cross-validation (LOOCV) in this study is appropriate because the dataset used is relatively small, with only 61 samples. LOOCV is a powerful method for model selection because it is model agnostic and provides an unbiased estimate of model performance [18]. However, it can be computationally expensive for large datasets due to the need to refit the model for each sample repeatedly [19]. According to the literature cited, LOOCV is favored for a limited number of data, and in this case, it offers a better solution with minimal time cost. Overall, the use of LOOCV in this study is justified and can provide reliable estimates of model performance.

The results obtained from the ordinary kriging phase are considered representative and are supported by the geostatistical approach and LOOCV procedure (as seen in Figs. 3, 4, 7 and 8). The use of these methods provides confidence in the accuracy and reliability of the obtained results.

This validation can be seen from two perspectives. The first is used in the computer science community, it consists in improving the quality of the adjusted model. The second perspective is from a practical standpoint, where the validated model can be used to make better decisions in resource management and exploitation. By having a more accurate estimate of the mineral reserves, mining companies can optimize their extraction plans, reduce waste and minimize costs. This can lead to significant improvements in profitability and sustainability of the mining operations. Additionally, the use of geostatistics and cross-validation methods can increase the reliability and credibility of the resource estimates, which is important for investors, regulatory bodies and other stakeholders involved in the mining industry. Overall, the validation of the variograms models using the LOOCV method provides both technical and practical benefits, making it an important tool for the mining industry.

Limitations and directions of research development. It is important to note that the choice of cross-validation method depends on the number and distribution of samples. When dealing with limited sampling, the LOOCV method is recommended, as it involves eliminating a sample in turn for verification [18]. This method ensures that the spatial continuity of the parameters being studied is not affected by using (n – 1) samples for experimental variogram calculations.

In the future, it is important to explore methods that offer more robust evaluations while considering simplicity of application and minimal computational cost. By seeking such methods, you can strike a balance between accuracy and computational efficiency. However, it is crucial to select the validation method based on the specific characteristics of the problem and the available data. Different methods may be more suitable for different scenarios, and careful consideration should be given to ensure the chosen method aligns with the requirements and limitations of the study.

Conclusion. The geostatistical approach developed in this study has demonstrated its effectiveness in improving the estimation of mineral reserves. The use of the least squares and LOOCV methods in the variogram model selection and estimation phases, respectively, has resulted in a more accurate representation of the spatial variability of the studied parameters. The evaluation of efficiency factors, such as Nash-Sutcliffe, RMSE, and RSR, has confirmed the quality of the adjustment models obtained.

However, the sensitivity of the least squares method to initial data highlights the need for caution in its use, and the applicability of the LOOCV method may be limited to sites with a smaller number of data points. The development of a computer code for this geostatistical approach provides a practical tool for decision-makers to use in the management and exploitation of mining sites. Overall, this study has contributed to the advancement of geostatistical techniques and their application in the mining industry.

We suggest a generalization of the accuracy assessment method using Leave-One-Out Cross-Validation, a model validation technique widely used in estimation fields like geostatistics. This method is particularly suitable for small datasets, as commonly encountered in the mining industry.

References.

1. Ajak, A. D., Lilford, E., & Topal, E. (2018). Application of predictive data mining to create mine plan flexibility in the face of geological uncertainty. *Resources Policy, 55*, 62-79. https://doi.org/10.1016/j.resourpol.2017.10.016.

2. Li, Z., Zhang, X., Zhu, R., Zhang, Z., & Weng, Z. (2020). Integrating data-to-data correlation into inverse distance weighting. *Computational Geosciences, 24*, 203-216. https://doi.org/10.1007/s10596-019-09913-9.

3. Choi, Y., Baek, J., & Park, S. (2020). Review of GIS-based applications for mining: Planning, operation, and environmental management. *Applied Sciences, 10*(7), 2266. https://doi.org/10.3390/app10072266.

4. Jalloh, A. B., Kyuro, S., Jalloh, Y., & Barrie, A. K. (2016). Integrating artificial neural networks and geostatistics for optimum 3D geological block modeling in mineral reserve estimation: A case study. *International Journal of Mining Science and Technology, 26*(4), 581-585. https://doi.org/10.1016/j.ijmst.2016.05.008.

5. Rebbah, R., Duarte, J., Djezairi, O., Fredj, M., & Baptista, J. S. (2021). A Tunnel under an In-Pit Mine Waste Dump to Improve Environmental and Landscape Recovery of the Site. *Minerals, 11*(6). https://doi.org/10.3390/min11060566.

6. Zerzour, O., Gadri, L., Hadji, R., Mebrouk, F., & Hamed, Y. (2021). Geostatistics-Based Method for Irregular Mineral Resource Estimation, in Ouenza Iron Mine, Northeastern Algeria. *Geotechnical*

and Geological Engineering, 39(5), 3337-3346. https://doi.org/10.1007/s10706-021-01695-1.
7. Afeni, T. B., Akeju, V. O., & Aladejare, A. E. (2021). A comparative study of geometric and geostatistical methods for qualitative reserve estimation of limestone deposit. *Geoscience Frontiers, 12*(1), 243-253. https://doi.org/10.1016/j.gsf.2020.02.019.
8. Berrar, D. (2019). Cross-Validation. *Encyclopedia of Bioinformatics and Computational Biology,* (1), 542-545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X.
9. Vehtari, A., Simpson, D. P., Yao, Y., & Gelman, A. (2018). Limitations of "Limitations of Bayesian Leave-one-out Cross-Validation for Model Selection". *Computational Brain & Behavior, 2*(1), 22-27. https://doi.org/10.1007/s42113-018-0020-6.
10. Giordano, R., Stephenson, W., Liu, R., Jordan, M., & Broderick, T. (2019). A Swiss Army Infinitesimal Jackknife. In C. Kamalika, & S. Masashi (Eds.). *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, PMLR: Proceedings of Machine Learning Research, (pp. 1139-1147). Retrieved from http://proceedings.mlr.press/v89/giordano19a/giordano19a.pdf.
11. Kechiched, R., Laouar, R., Bruguier, O., Salmi-Laouar, S., Kocsis, L., Bosch, D., ..., & Larit, H. (2018). Glauconite-bearing sedimentary phosphorites from the Tébessa region (eastern Algeria): Evidence of REE enrichment and geochemical constraints on their origin. *Journal of African Earth Sciences, 145*, 190-200. https://doi.org/10.1016/j.jafrearsci.2018.05.018.
12. Zhongda, T., Shujiang, L., Yanhong, W., & Xiangdong, W. (2018). Mixed-kernel least square support vector machine predictive control based on improved free search algorithm for nonlinear systems. *Transactions of the Institute of Measurement and Control, 40*(16), 4382-4396. https://doi.org/10.1177/0142331217748193.
13. Duc, L., & Sawada, Y. (2022). *A signal processing-based interpretation of the Nash-Sutcliffe efficiency.* EGUsphere. https://doi.org/10.5194/egusphere-2022-955.
14. Chen, C., Zhang, Q., Kashani, M. H., Jun, C., Bateni, S. M., Band, S. S., Dash, S. S., & Chau, K.-W. (2022). Forecast of rainfall distribution based on fixed sliding window long short-term memory. *Engineering Applications of Computational Fluid Mechanics, 16*(1), 248-261. https://doi.org/10.1080/19942060.2021.2009374.
15. Zhong, X., & Dutta, U. (2015). Engaging Nash-Sutcliffe efficiency and model efficiency factor indicators in selecting and validating effective light rail system operation and maintenance cost models. *Journal of Traffic and Transportation Engineering, 3*, 255-265. https://doi.org/10.17265/2328-2142/2015.05.001.
16. Yasojima, C., Protázio, J., Meiguins, B., Neto, N., & Morais, J. (2019). A new methodology for automatic cluster-based kriging using K-nearest neighbor and genetic algorithms. *Information, 10*(11), 357. https://doi.org/10.3390/info10110357.
17. Zeybek, M. (2018). Nash-sutcliffe efficiency approach for quality improvement. *Journal of Applied Mathematics and Computation, 2*(11), 496-503. https://doi.org/10.26855/jamc.2018.11.001.
18. Magnusson, M., Vehtari, A., Jonasson, J., & Andersen, M. (2020). Leave-One-Out Cross-Validation for Bayesian Model Comparison in Large Data. In C. Silvia, & C. Roberto (Eds.). *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. PMLR: Proceedings of Machine Learning Research,* (pp. 341-351). Retrieved from http://proceedings.mlr.press/v108/magnusson20a/magnusson20a.pdf.
19. Gronau, Q. F., & Wagenmakers, E.-J. (2019). Limitations of Bayesian leave-one-out cross-validation for model selection. *Computational brain & behavior,* (2), 1-11. https://doi.org/10.1007/s42113-018-0011-7.

Оцінка мінеральних ресурсів за допомогою геостатистичного аналізу на фосфатному родовищі

М. Мазарі[*1,2], *С. Чабу-Мостефай*[1], *А. Балі*[3], *К. Куйдер*[4,5], *А. Бенселгуб*[6], *С. Беллуччі*[7]

1 – Кафедра гірничої інженерії, Національна політехнічна школа Алжиру, м. Алжир, Алжир
2 – Кафедра гірничої справи та геології, Технологічний факультет, Університет Абдеррахмена Міра, м. Беджая, Алжир
3 – Кафедра цивільного будівництва, Лабораторія будівельних матеріалів та екології, Національна політехнічна школа Алжиру, м. Алжир, Алжир
4 – Науково-дослідна лабораторія водних наук, Національна політехнічна школа Алжиру, м. Алжир, Алжир
5 – Факультет цивільного будівництва, Університет наук і технологій імені Хуарі Бумедьєна, м. Алжир, Алжир
6 – Відділ навколишнього середовища, моделювання та зміни клімату, Науково-дослідний центр з охорони довкілля, м. Аннаба, Алжир
7 – Національні лабораторії Фраскаті, Фраскаті, м. Рим, Італія
[*] Автор-кореспондент e-mail: messaoud.mazari@g.enp.edu.dz

Мета. Вибір відповідної варіографічної моделі має вирішальне значення в геостатистиці для отримання точних оцінок запасів корисних копалин. Метою цієї роботи є розробка інструменту оцінки запасів із використанням геостатистичного підходу.

Методика. Геостатистичний підхід базується на виборі найбільш репрезентативних варіографічних моделей для змінних, що досліджуються. Відбір моделей здійснюється шляхом застосування процедури перехресного затверджування з виключенням по одному (LOOCV). LOOCV – це метод повторної вибірки, що використовується у статистичному аналізі й машинному навчанні для оцінки помилки узагальнення моделі та порівняння ефективності різних моделей. Після цього змінні, що досліджуються, оцінюються за допомогою звичайного кригінгу.

Результати. Застосування запропонованого підходу дозволило отримати задовільні результати з точки зору дисперсії сортів і товщини мінералізованих шарів на фосфатному родовищі. Для оцінки якості отриманих моделей коригування були використані коефіцієнти ефективності, такі як коефіцієнт Неша-Саткліффа та коренева середньоквадратична похибка (RMSE). Ці фактори надають кількісну оцінку узгодженості між значеннями, що спостерігаються та прогнозуються. Значення ефективності Неша-Саткліффа та середньоквадратичної похибки 0,572 та 6,599 відповідно, свідчать про кращу відповідність і більшу точність моделей коригування. Критерії точності та ефективності досліджуваних змінних мають прийнятні значення із середньоквадратичною похибкою $1{,}54 \cdot 10^{-7}$.

Наукова новизна. Поєднання методів найменших квадратів і LOOCV у геостатистичному аналізі призводить до підвищення точності оцінок, більшої надійності у представленні просторової мінливості параметрів і більшої впевненості в достовірності моделей коригування.

Практична значимість. Розробка комп'ютерного коду для цього геостатистичного підходу забезпечує практичний інструмент для осіб, які приймають рішення, для використання в управлінні та експлуатації гірничодобувних об'єктів. Загалом, це дослідження сприяло розвитку геостатистичних методів і їх застосуванню в гірничодобувній промисловості.

Ключові слова: *родовище Блед-Ель-Хадба, перехресне затверджування (LOOCV), геостатистика, кригінг, запаси корисних копалин*

The manuscript was submitted 10.03.23.