Article

Not peer-reviewed version

New Methodology for Evaluating Uncertainty in Mineral Resource Estimation

José Alberto Arias *, Alain Carballo, Elmidio Estévez, Reinaldo Rojas, Domingo A. Martín, Jorge L. Costafreda

Posted Date: 1 September 2025

doi: 10.20944/preprints202509.0114.v1

Keywords: mineral resources; uncertainty; gypsum; modelling; simple normal equation simulation; training image; prior probability maps

 doi:10.20944/preprints202509.0114.v1

Article

New Methodology for Evaluating Uncertainty in Mineral Resource Estimation

José Alberto Arias [1,2,*], Alain Carballo [2], Elmidio Estévez [3], Reinaldo Rojas [4], Domingo A. Martín [5] and Jorge L. Costafreda [5]

[1] Oficina Nacional de Recursos Minerales, Calzada 852 Esq.4, Municipio Plaza de la Revolución, Vedado, CP 10300 La Habana, Cuba
[2] Universidad de Moa, "Dr. Antonio Núñez Jiménez", Calixto García Iñiguez No 15, entre Ave 7 Diciembre y calle Reynaldo Laffita Rueda, Reparto Caribe, Moa, Holguín, Cuba
[3] Universidad de Pinar del Rio ,"Hermanos Saíz Montes de Oca", Calle Martí, No 300,Barrio Segundo Sur, entre 27 de noviembre y González Alcorta, Pinar del Rio, Cuba
[4] Centro de Investigaciones del Petróleo
[5] Escuela Técnica Superior de Ingenieros de Minas y Energía, Universidad Politécnica de Madrid
* Correspondence: j.aariasdt61@gmail.com; + 53 52540787 (J.A.A)

Abstract

Geological modelling is generally based on deterministic models, which provide a single representation of reality. Probabilistic modelling is more appropriate when quantifying or understanding the uncertainty associated with a parameter of interest as it considers several equally probable geological scenarios. The object of this study is to quantify the uncertainty in the estimation of the minerals in the Punta Alegre gypsum deposit, by applying a new method based on the simple normal equation geostatistical simulation technique. The Punta Alegre gypsum deposit is a sedimentary deposit of clastic origin, formed by the complex redeposition of salts, gypsum and other sediments. To carry out this research, 50 equiprobable scenarios were simulated, reproducing 1- overburden, 2- gypsum series (different types of gypsum) and 3-intercalated non-mineral lithologies (limestone and other rocks), in a network of nodes measuring 5 x 5 x 5 meters, using a training image, composites and prior probability maps as input data. As a result of scaling the previously simulated geological units, three-dimensional models of volume proportions and estimation error for gypsum were obtained for panels measuring 10 x 10 x 5 meters. The quantification of the uncertainty of the gypsum volume, determined by the root mean square error, established that the volume estimation error is small at a global scale (6.51%), given that there is no significant variation when comparing the deterministic model with the gypsum proportion model obtained from the 50 simulated scenarios. Conversely, at the local scale, there is a significant variation in gypsum volume of 42% in the 10 x 10 x 5 meter panels with a future impact on recoverable mining resources, given the uncertainty at a local scale, which will cause an increase in mining dilution due to the inclusion of non-mineral lithologies within the extracted mineral that will be sent to the processing plant. On the other hand, it will cause changes in the mining company's plan in areas where there are panels that were previously accounted for by the deterministic model as minerals and are not actually exploitable.

Keywords: mineral resources; uncertainty; gypsum; modelling; simple normal equation simulation; training image; prior probability maps

1. Introduction

Traditional geological modelling is deterministic, i.e. the geological model that is created constitutes a unique representation of reality, which presumably means perfect and absolute knowledge of the size and shape of geological units and spatial continuity [1], and is only valid when

 doi:10.20944/preprints202509.0114.v1

the available data is sufficient to create models with a high degree of certainty. Conversely, when the interpretation is not definitive and several equally valid scenarios could be inferred from the available information, deterministic modelling would lead to the creation of geological models with a high degree of uncertainty [2].

A set of geostatistical simulation techniques have now been developed for stochastic geological modelling which, in contrast to deterministic geological modelling, create multiple representations of the unknown reality, reflecting the spatial continuity of geological units and the ore of the mineral [3]. The series of models created in this way are the most common representation of what an uncertainty model is [4]. Due to this, geostatistical conditional simulation techniques have become one of the most widely used tools in the mining industry to quantify global or local uncertainty of a random nature [5], based on several simulated equiprobable scenarios [6].

Among the simulation techniques developed is SNESIM (Simple Normal Equation Simulation), which is based on the principle of multiple point simulation, since it does not require a prior variogram model to run it, but rather uses a training image [7,8] that represents the conceptual geological model of the mineral deposit. The training image itself is a repository that provides information on the spatial patterns and their probabilities found in the phenomenon from which the structures are taken directly. The result is a spatial distribution of lithotypes conditioned by real data from boreholes with geological information. The SNESIM method is particularly useful for creating stochastic images of a sparsely sampled geological phenomenon with complex features that cannot be captured and reproduced by two-point statistics [9].

Given that previous studies on evaporite deposits have not quantified the uncertainty of the estimates using geostatistical simulation techniques, this paper develops a novel methodology to quantify the uncertainty of global and local estimates of the recoverable mineral, by applying the multipoint geostatistical sequential simulation method known as SNESIM [10–15]. The main object is to reproduce the spatial continuity patterns, of the geological units beforehand such as the overburden, with multiple predefined equiprobable scenarios, the gypsum rock series as well as the non-mineral lithologies in the gypsum deposit series of Punta Alegre gypsum deposit, which is used as a case study in this research.

2. Theoretical Framework

2.1. Simple Normal Equation Simulation (SNESIM)

This simulation technique uses a single equation to model the probability of occurrence of a facies, lithotype or category in a particular node of the grid, using the well-known Bayes' rule (equation 1), which defines the conditional probability of an event d_n [10,14,16,17]:

$$\text{Prob}\{s(u)=(s_k|d_n)\}=\frac{Prob\{S_u=s_k\ y\ d_n\}}{Prob\{d_n\}}\approx\frac{C_k(d_n)}{C(d_n)} \quad (1)$$

Where $\text{Prob}\{d_n\}$ is the training replica as it has the same geometric configuration and lithotype and is calculated by scanning the training image and counting the number of training replicas $C(d_n)$, of the conditioning data event $C(d_n)$.

$\text{Prob}\{s_u = s_k\, y\, d_n\}$ is obtained by counting the number $c_k(d_n)$ of those replicas between the training replicas $C(d_n)$ with a central value of $S_u = S_k$.

In order to search for adjacent conditioning data in the training image, a 'search template' is used. The frequency of the training replicas obtained from the training image is stored in a search tree structure, allowing it to be easily retrieved during the simulation process (Figure 1).

 doi:10.20944/preprints202509.0114.v1

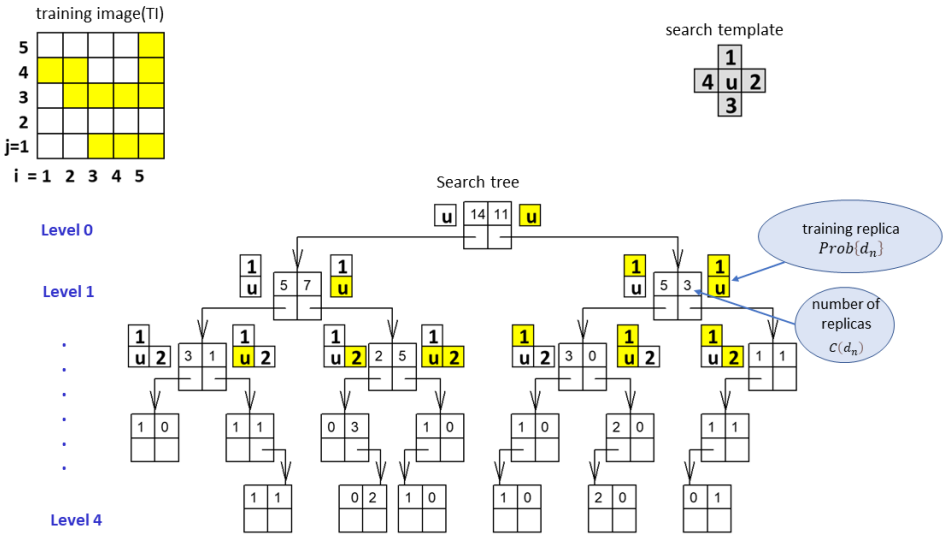

Figure 1. Training image for 2 categories, tree and search template, modified from [18].

The simulation using the SNESIM algorithm consists of two main parts. In the first part, a search tree is constructed, where all the proportions obtained from the training image are stored. The second part is the simulation itself, where these proportions are read and used to obtain the simulated values [19].

The SNESIM algorithm can be summarized as follows [10,11,19]: First, the training image (TI) is scanned to create the specific search tree from a search template τ_j.The original data (composite samples) are then relocated to the nodes closest to the network to be simulated and 'frozen' during the simulation process. A random path is then defined to visit each location only once; the conditional probability $\mathrm{Prob}\{s(u) = (s_k|d_n)\}$ of the event d_n integrating the proportions $C_k(d_n)/C(d_n)$ stored in the training image and the probabilities obtained from the secondary information ('soft data'). Finally, at random, a value of the possible category at location u_j from the estimated conditional probability, which is added to the original data to be used as conditioning data at other locations. The process is repeated until all nodes have been visited. Multiple realizations are obtained using different random paths.

The SNESIM algorithm was the geostatistical tool used in this study to simulate the lithotypes in the 'Punta Alegre' deposit.

2.2. Geological Setting

The area where the deposit is located and its surroundings have been extensively investigated by various authors and described both from a regional geological point of view [20] for the search for hydrocarbons [21–23] and locally for the exploration and exploitation of gypsum as a raw material for the cement industry [24,25].

The rocks that make up the 'Punta Alegre' deposit have been conditionally referred to as belonging to the Lower and Middle Jurassic interval and have been named the 'Punta Alegre' Formation. These are rocks of evaporitic origin, which generally include: salt, gypsum, anhydrite, dolomite and scarce limestones.

According to the stratigraphic lexicon of Cuba [26,27], their diagnostic lithology consists of gypsum breccias with clasts of limestone, slate, siltstone, sandstone and tuff from the Lower and Middle Jurassic periods.

From a structural point of view, these rocks have been attributed to the formation of a salt diapir [28,30]. The authors findings include within the gypsum mass 'exotic' rocks of varied make up and environment, attributing a tectonic origin to the breccia at Punta Alegre, interpreting the carbonates

 doi:10.20944/preprints202509.0114.v1

(brecciated calcareous clays) containing microfauna of foraminifera and ostracods from the Lower Miocene as strata included in the gypsum matrix, due to the spreading of the salt mass over the seabed in the Lower Miocene [20,22].

Recent studies [23] describe the Punta Alegre Formation geologically as a deposit of clastic sedimentary origin, formed by the complex successive redeposition of salts and gypsum by halokinetic processes together with other sediments in the post-orogenic Morón sub-basin, considered by the authors to be younger, i.e., Oligocene-Miocene. This formation mechanism explains the apparent contradiction between various authors regarding the age of the deposit, which ranges from Jurassic [21,22] to Miocene or younger [20,23–25].

Discordantly, on top of the rocks of the 'Punta Alegre' Formation (J_1-J_2) lie the Quaternary formations, mainly represented by massive biodetrital limestones, generally karstified ('Jaimanitas' Formation), sandy-clayey silts and silty clays with intercalations of fine gravel and carbonate concretions, and scattered gypsum crystals ('Camacho Formation'), alluvial, eluvial-colluvial-proluvial, marshy and silty-sandy deposits (Figure 2).

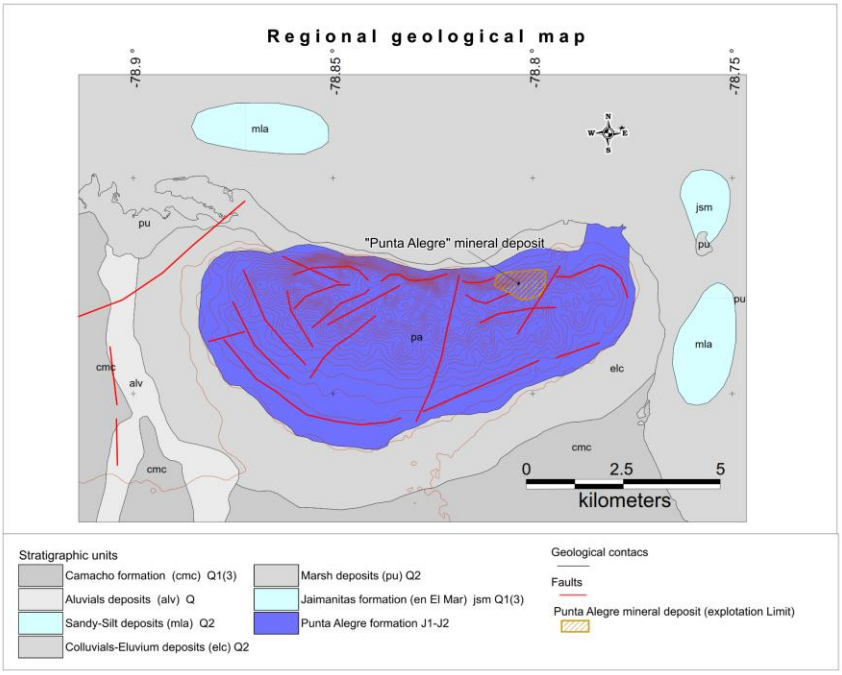

Figure 2. Regional geological map of the study area, modified from Instituto de Geología y Paleontología of Cuba (I.G.P.) [31].

According to [25] from a geological point of view, the 'Punta Alegre' deposit constitutes a thick series of rocks, with a thickness of between 150 and 450 meters, composed mainly of gypsum-bearing materials, in most cases in the form of breccia, of sedimentary origin (synsedimentary breccia), in blocks ranging from 2 to 15 meters, cemented by clayey material with intercalations of sometimes brecciated clays, ranging from the first decimeters to several meters (10-15 meters) in thickness. Blocks of limestone, slate and argillite are also present.

According to data provided by detailed exploration work carried out between 1964 and 1977, useful minerals are represented by varieties of gypsum which, depending on their purity, are: white and grey-white gypsum, pure, almost pure, yellowish, blue-green, reddish, grey-black, purple, with a brecciated texture.

 doi:10.20944/preprints202509.0114.v1

3. Materials and Methods

3.1. Materials

For this study, information was obtained from the records of 384 boreholes with 4,530 irregular intervals drilled approximately in 40 x 40 m, 20 x 20 m and 80 x 50 m grids (Figure 3), lithologically described and validated by consulting and checking the primary information in previous geological exploration reports on the deposit [25,32].

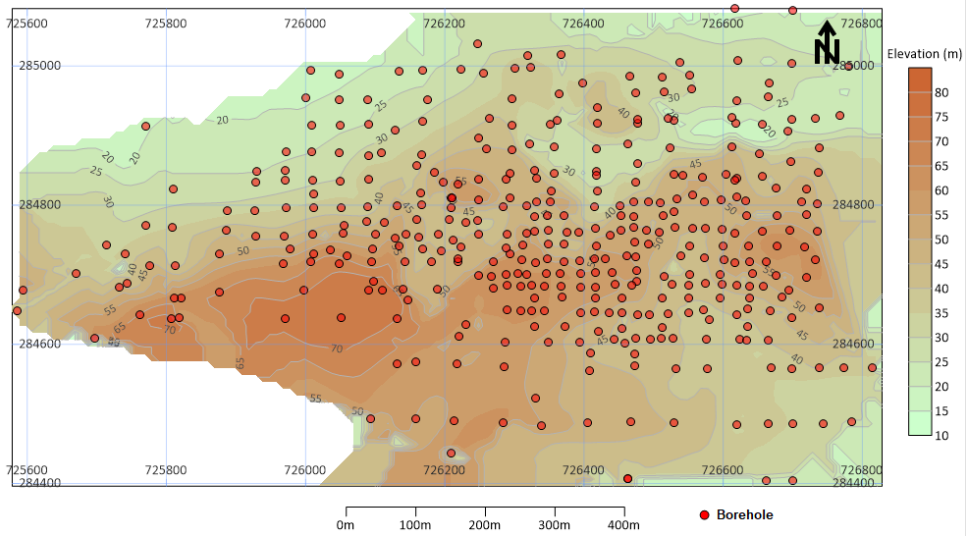

Figure 3. Location map of the exploration boreholes.

3.2. Methods

3.2.1. Study Area

The Punta Alegre gypsum deposit is in the province of Ciego de Ávila, municipality of Chambas, 2.0 km northwest of Central Máximo Gómez and 1.7 km southwest of the town of the same name (Figure 4), covering part of the so-called 'Loma de yeso' (Gypsum Hill), covering a total area of 35 km^2 [32].

 doi:10.20944/preprints202509.0114.v1

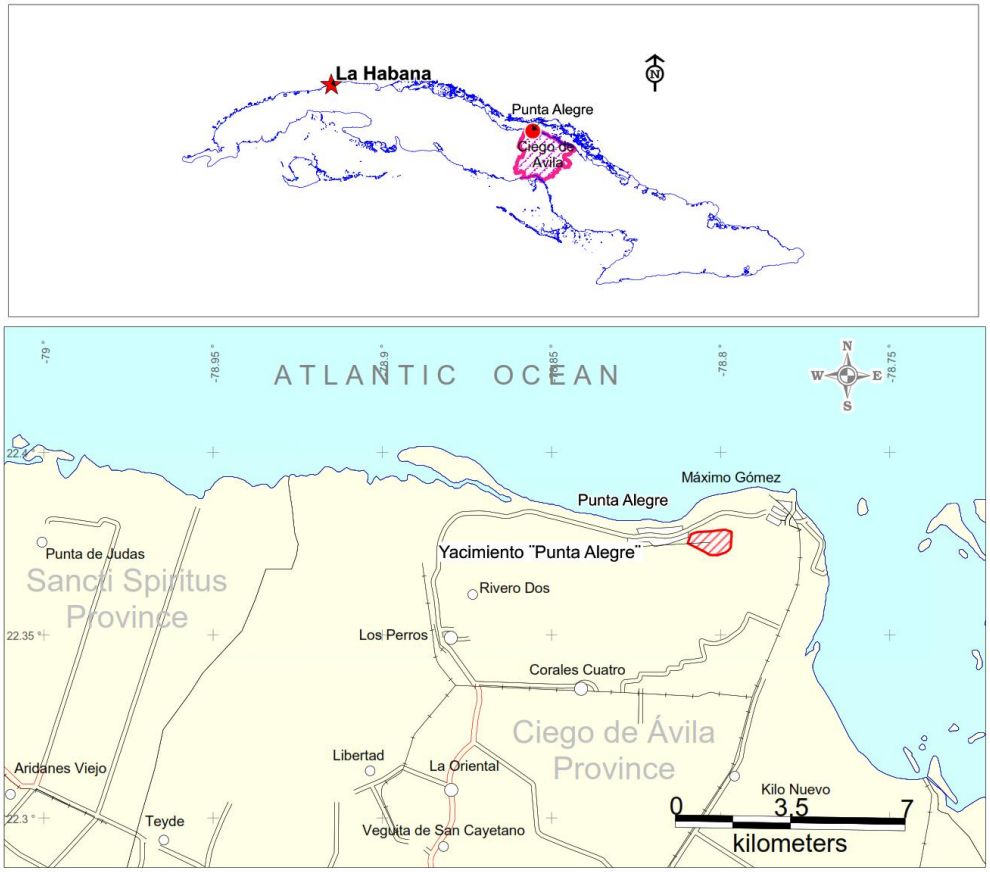

Figure 4. Geographical location of the 'Punta Alegre' gypsum deposit.

3.2.2. Geological Modelling of the Punta Alegre Deposit

For the modelling of the Punta Alegre gypsum deposit, the spatial boundary was defined as the volume of rock located between the topographic surface and the surface delimiting the base of the boreholes drilled for the exploration of the deposit.

The non-mineral envelopes were modelled in the space occupied by the of rocks that define the 'Punta Alegre' gypsum deposit, based on the manual digitization of the perimeters that encompass the non-mineral lithologies, such as inclusions of clayey rocks, limestones, dolomites and slates that appear in the section of the deposit forming the covering materials (overburden) as well as the inclusions of these lithologies within the useful mineral.

The perimeters were digitized in sections oriented in an east-west direction, based on the non-mineral intervals defined in the borehole. The wireframes of the non-mineral solids of the overburden were then created by joining the digitized perimeters with the tie lines that interconnect them. For non-mineral intercalations (inter-sterile), solids were constructed by extrusion from their perimeters to half the distance between the two adjacent sections.

Finally, the parts of the created non-mineral solids where mining had occurred were removed by a cut operation using the updated (current topography) digital terrain model. The general 3D model of the envelopes of the geological units, i.e., wireframes of the solids created, is shown in Figure 5.

 doi:10.20944/preprints202509.0114.v1

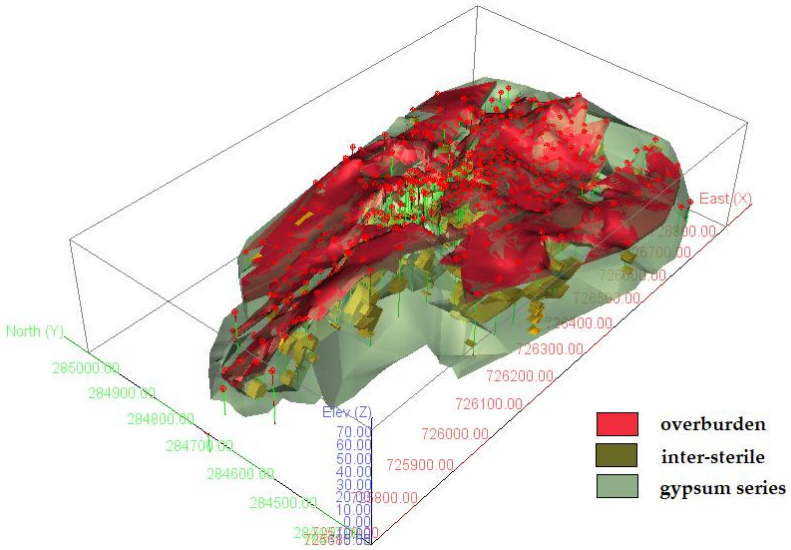

Figure 5. Boreholes and solids of the overburden (covering materials), inter-sterile (non-mineral intercalations) and the gypsum series (gypsum).

3.2.3. Identification of Non-Mineral Envelopes

Once the non-mineral solids ('wireframes') had been built they were identified and named as solids of the 'overburden' for the non-mineral lithologies located in the upper part of the section, which are more continuous and mainly composed of the topsoil and covering clays, and the non-mineral lithology solids intercalated in the section within the gypsum series (clays, limestones, dolomites and shales), which are less continuous and in the form of lenses and irregular bodies of smaller dimensions. The envelopes of these non-mineral lithologies intercalated within the gypsum series were called 'inter-sterile'.

3.2.4. Building the Training Image (TI)

The training image (TI) was built according to a network of 5 x 5 x 5 meter nodes, thus subdividing the 10 x 10 x 5 meter panels internally into smaller panels, according to a 2 x 2 x 1 unit discretization pattern for scaling.

First, all nodes in the 5 x 5 x 5 meter network were selected within the space delimited between the topography and the base of the boreholes (envelope of the deposit), assigning them a numerical code (code = 1). Then the nodes contained by the solids of the overburden (code = 0) and the inter-sterile (code = 2) were recorded. The training image (TI) created represents the conceptual model of the spatial distribution of the mineral and non-mineral units in the 'Punta Alegre' deposit (Figure 6).

 doi:10.20944/preprints202509.0114.v1

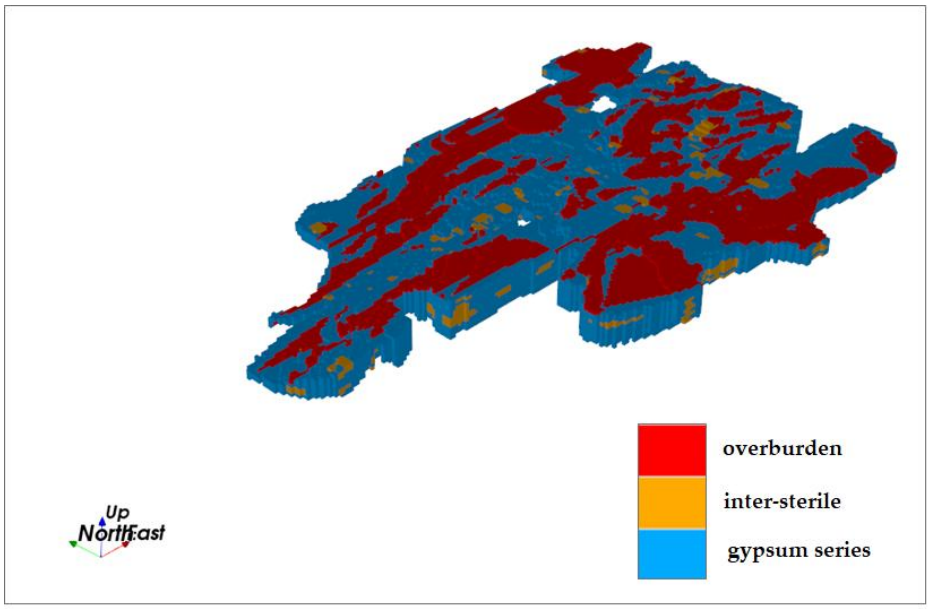

Figure 6. - 3D training image (TI) of the units: Overburden, Inter-sterile and gypsum, created according to a network of nodes measuring 5x5x5 meters.

3.2.5. Composites

Composites were created to ensure that all samples had the same weight [33], to standardize the lithological information in the database of the Punta Alegre deposit, from the exploration boreholes at various stages [25,32].

The composites were created within the intervals resulting from the intersection of the boreholes and the solids created, using an average length for the composite of 0.5 meters, which corresponds to the most frequent length found in the intervals in the 'lithology' table. For this purpose, code 1 (gypsum) was first assigned to all composites located within the general envelope of the deposit, and they were subsequently recorded with the appropriate code according to whether they belonged to the overburden or the 'inter-sterile' (code = 0 and code = 2, respectively). A total of 8,730 composites were created.

3.2.6. Prior Probability Maps of the Overburden, the Valuable Mineral (Gypsum) and the Intercalated Sterile (Inter-Sterile')

For the building of the prior probability maps of the units represented by the overburden, the valuable mineral (gypsum) and the intercalated sterile ('inter-sterile'), each composite interval was previously coded using binary indicators [0,1], which indicate the presence or absence (zero or one, respectively) of any of the aforementioned types, according to the condition, i(u,k)=1 if Z(u)=K ;0 or not. Finally, the prior probability maps were created by averaging the values of the indicators [0,1] in the 5 x 5 x 5 meter node network using the moving window algorithm [34] in a 150 x 150 x 5 meter search area to obtain probability maps with adequate smoothing (Figure 7).

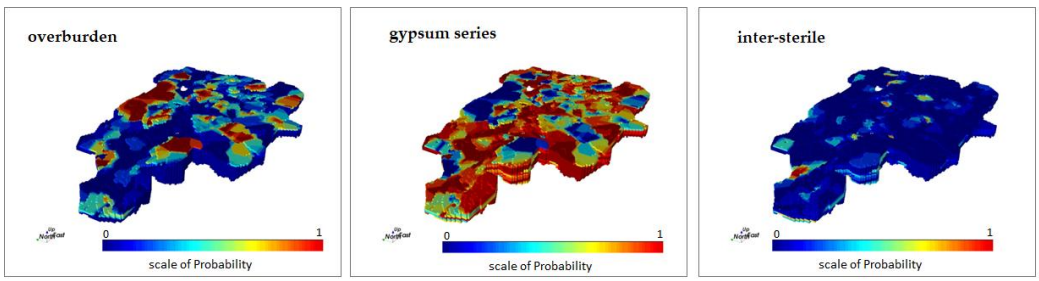

Figure 7. Prior probability maps (5x5x5 meter nodes).

 doi:10.20944/preprints202509.0114.v1

4. Results and Discussion

4.1. Results

The simulation was carried out for 50 equally probable scenarios using the SNESIM sequential simulation algorithm, implemented in the geostatistical software 'SGeMS' [11], using the training image (TI) and composites as hard data. As secondary information, the prior probability maps created beforehand by the moving average algorithm for the aforementioned units were used, with a search template of 150 x 150 x 5 meters and a maximum of 80 conditioning data points. The simulated units for the overburden, inter-sterile and gypsum series (gypsum) for scenarios 1, 5, 18, 23, 34 and the training image (TI) are shown below, the latter being used as input for the simulation (Figure 8).

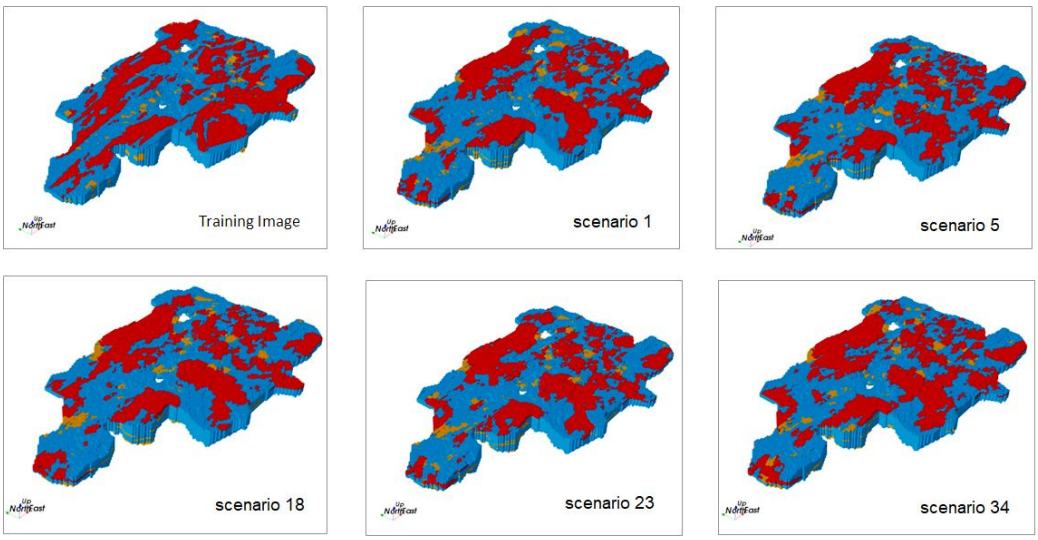

Figure 8. Training image (TI) versus simulated scenarios.

To validate the simulated models, various criteria were used: a) reproduction of the overall statistics and spatial correlation, and b) visual verification of the correlation between the input data and the simulated models.

Judging by the overall proportions obtained for overburden gypsum (mineral) and intercalated non-mineral lithologies (sterile) obtained for 10 randomly selected scenarios out of a total of 50 simulated scenarios, the images obtained satisfactorily reproduce the global proportions of the training image and the composites, as shown in Table 1 and Figure 9.

Table 1. Actual global proportions versus simulated global proportions.

	Composites	**Training image (TI)**	**Simulated scenarios**									
			1	**4**	**13**	**19**	**25**	**34**	**39**	**41**	**46**	**50**
Overburden	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Gypsum series	0.84	0.82	0.81	0.82	0.82	0.82	0.82	0.81	0.81	0.81	0.82	0.81
Inter-sterile	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.11	0.10	0.10	0.10	0.10

 doi:10.20944/preprints202509.0114.v1

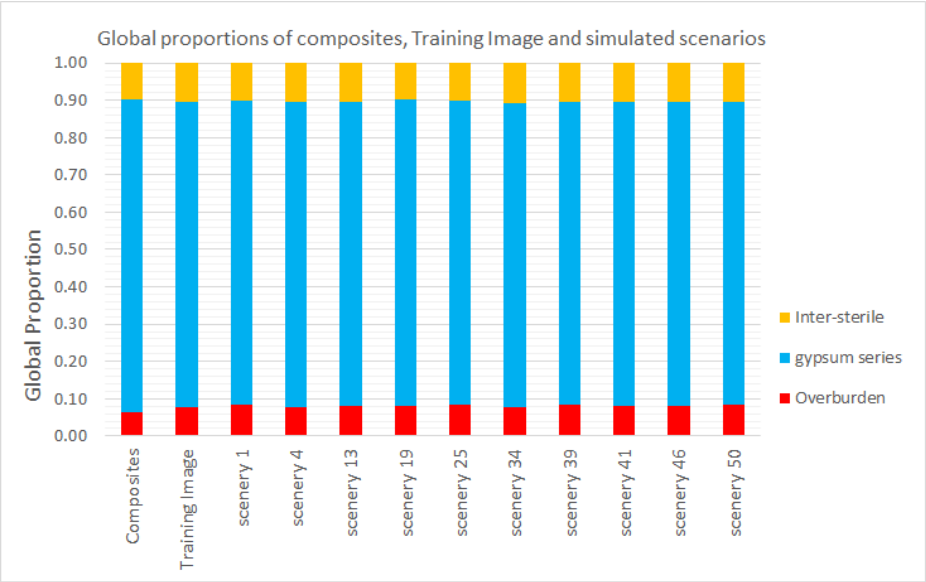

Figure 9. Comparative graph of proportions for 10 equally probable scenarios of the simulated units.

Simulated semi-variogram indicators in the North-South and East-West directions of the overburden, the Inter-sterile and gypsum, were also reproduced satisfactorily in all units for the 50 simulated scenarios, verifying that the spatial correlation model (covariance) of the statistics between two points in the simulated units is correct (Figure 10).

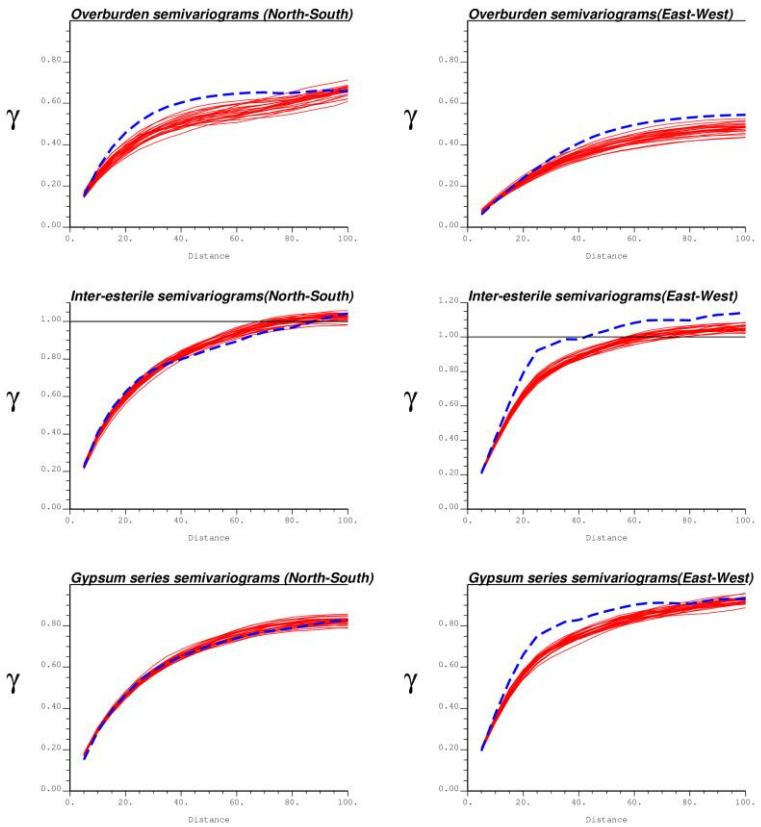

Figure 10. Simulated indicator semi-variograms (solid lines) versus semi-variograms of the TI training image (dotted line) for the units, the overburden, Inter-sterile and Gypsum.

 doi:10.20944/preprints202509.0114.v1

Visual verification of the simulated unit models shows that they are spatially consistent with the composite units used as conditioning data in the simulation process. In this case, simulated scenario #1 is shown as a reference. This shows a good correspondence between the composites and the simulated units (Figure 11).

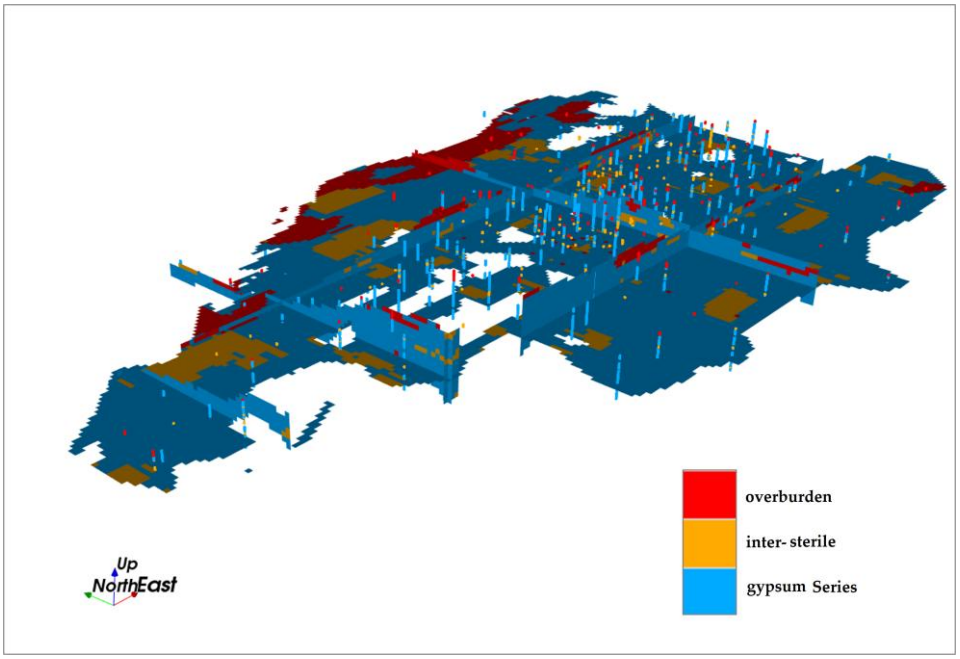

Figure 11. Visual validation for simulated scenario # 1. Comparison between the simulated units of the 'overburden', the non-mineral intercalations (Inter-sterile) and the gypsum series (Gypsum), and the composites used in the simulation.

4.1.1. Deterministic Model for Gypsum in 10 x 10 x 5 Meter Panels

The deterministic model for gypsum in 10 x 10 x 5 meter panels was created by scaling the nodes of the training image (TI), arranged in a 5x5x5 meter grid to 10x10x5 meter panels, calculating the proportion of each unit within the panel and assigning it to the unit with the highest proportion.

Each panel was then classified and coded with a value of one if the panel belonged to the gypsum series or with a value of zero if the panel was non-mineral, i.e. if it belonged to the overburden or inter-sterile, respectively (Figure 12).

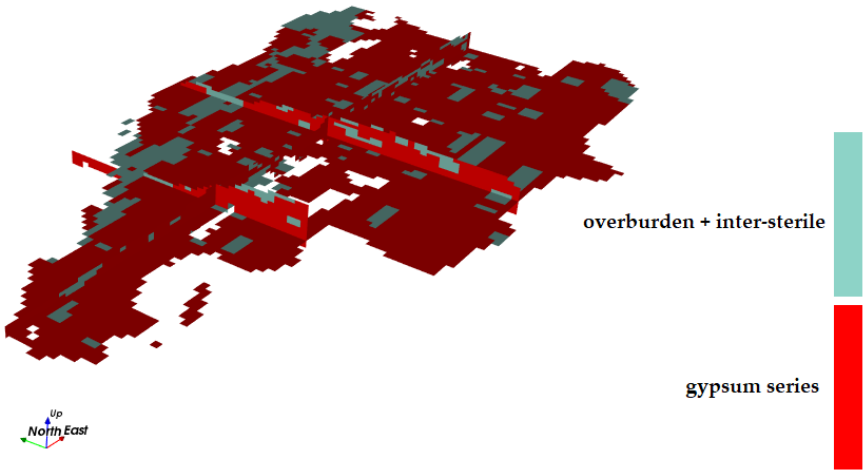

Figure 12. Deterministic model; gypsum series (gypsum) and non-mineral (overburden + inter-sterile) in the 10 x 10 x 5 meter panels.

 doi:10.20944/preprints202509.0114.v1

4.1.2. Model of the Average Proportions of Gypsum in the 10 x 10 x 5 Meter Panels

The model of average gypsum proportions in 10 x 10 x 5 meter panels was created in two stages using a procedure like that used to create the deterministic model of the gypsum series (gypsum).

In the first stage, the proportion of gypsum in the volume of each 10 x 10 x 5 meter panel was calculated for each of the 50 simulated scenarios using a support scaling, that is, by calculating the number of nodes within the 10 x 10 x 5 meter panel that resulted in gypsum according to the simulation, divided by the number of sub-blocks (4) into which the panel is discretized (Figure 13).

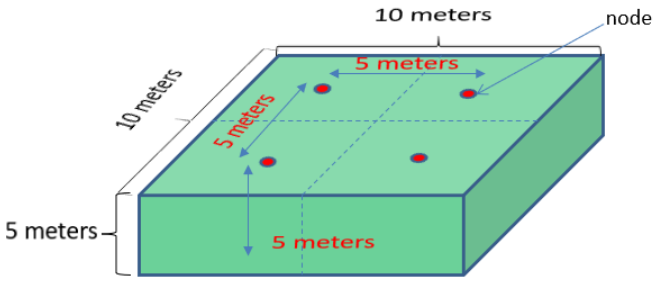

Figure 13. - Panel measuring 10 x 10 x 5 meters, discretized according to a 2 x 2 x 1 pattern.

The second stage corresponds to the calculation of the average P_B in which gypsum occupies the volume of the 10 x 10 x 5 meter panel, obtained by averaging the individual proportions calculated for gypsum in every 10 x 10 x 5 meter block of each of the 50 simulated scenarios.

The average gypsum values range between the extreme values of zero and one; zero means the total absence of gypsum in the panel volume, and the value of one means that the panel is completely made up of gypsum (Figure 14).

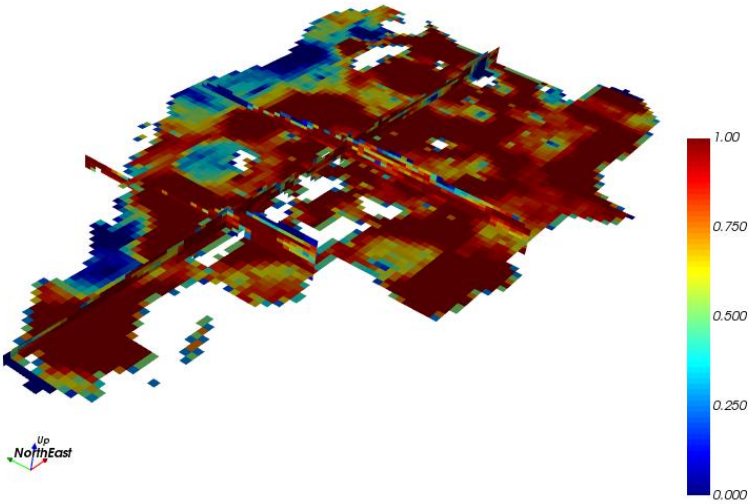

Figure 14. Model of proportions for gypsum in 10 x 10 x 5 meter panels at the Punta Alegre deposit according to 50 simulated scenarios.

Subsequently, the estimated gypsum volume (V_E) in each 10 x 10 x 5 meter panel was obtained by multiplying the mean estimated proportions of gypsum for the panel according to 50 simulated equiprobable scenarios by the unit of the panel, equal to 500 cubic meters.

 doi:10.20944/preprints202509.0114.v1

Finally, the spatial distribution model of the ε_B of the gypsum volume in the 10 x 10 x 5 meter panels (Figure 15) was obtained by the panel by panel difference between the gypsum volume of the deterministic mode (V_B) and the estimated gypsum volume (V_E) obtained from the proportion model for gypsum in the 10 x 10 x 5 meter panels.

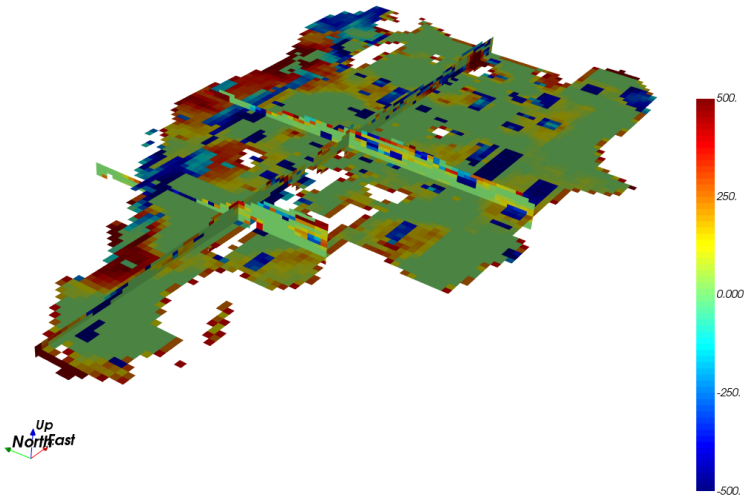

Figure 15. Spatial distribution model of the estimation error of the gypsum volume ε_B in the 10 x 10 x 5 meter panels at the Punta Alegre deposit. The scale is given in cubic meters.

4.2. Discussion

The quantification of the uncertainty of the useful mineral volume at the Punta Alegre deposit was carried out through comparative statistical processing of the differences between the gypsum volumes of the 10 x 10 x 5 meter panels according to the deterministic rock model and the gypsum volumes obtained for the same panels, using the estimated gypsum proportion model (Figure 17). The latter model was created by scaling the simulated values in the 5 x 5 x 5 meter grid for 50 equiprobable scenarios obtained through stochastic simulation.

The metrics used to quantify uncertainty were: Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Standard Deviation. Root Mean Square Error (RMSE) which measures the magnitude of the error between two sets of data. The latter statistic has been widely used to measure model performance in some studies [35–38], in this case it was used to measure the differences between the estimated gypsum volume of the panels according to the rock model (deterministic) V_B, and the volume obtained from the model of the estimated gypsum proportions in the panels and the volume V_E.

As stated in [39],many mineral deposits, given their complex geological nature, involve inherent uncertainty, which leads to underestimation or overestimation of mineral resources, such is the case of the Punta Alegre deposit, where the results obtained support this result, given that the frequency histogram of the differences between the volumes of V_B and V_E., shows a distribution of errors by class, arranged in negative and positive values, which represent the differences by default and by excess, respectively, relative to the variation in the volume of gypsum in the panel according to the deterministic rock model. This result is interpreted as evidence of underestimation and overestimation of the volume at the scale of the 10 x 10 x 5 meter panels.

These differences fall between the values of -500 and +500 cubic meters of panel volume, corresponding to the minimum and maximum values, respectively (Figure 16).

 doi:10.20944/preprints202509.0114.v1

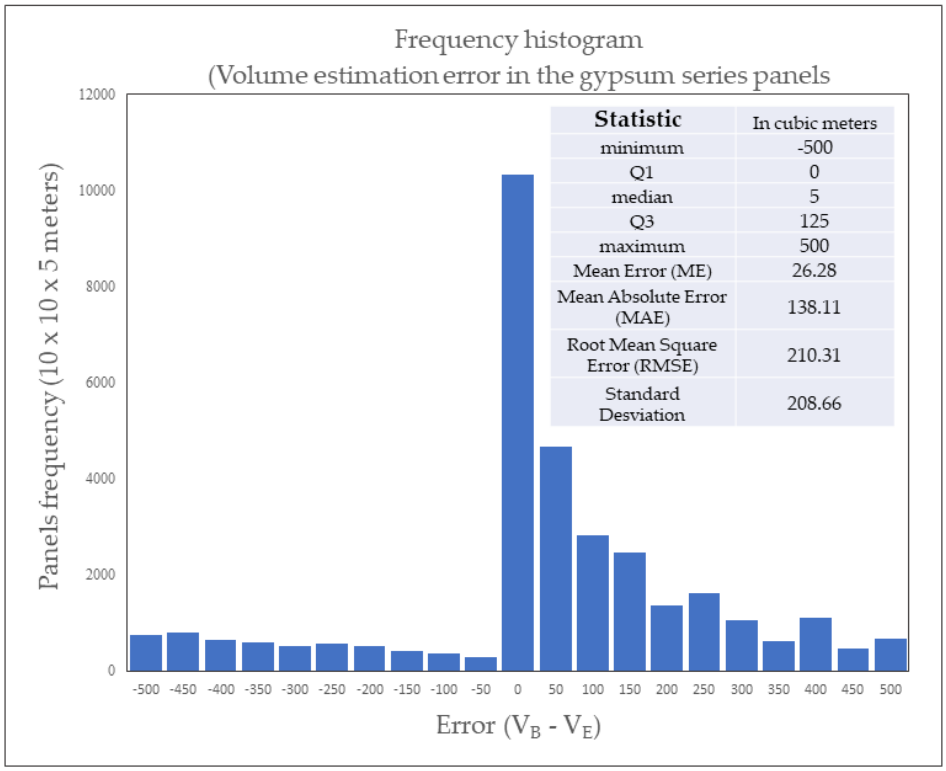

Figure 16. Frequency histogram of estimation errors ε_B of the gypsum volume of the deterministic model (V_B) and the volume estimated from the gypsum proportion model V_E in the 10x10x5 meter panels.

The fact that there are a greater number of 10x10x5 meter panels with a positive value for the estimation error ε_B of the gypsum volume is interpreted as meaning that there is a systematic overestimation of the gypsum volume according to the deterministic model (V_B .) therefore there is a dominant positive bias with respect to the volume estimated according to the model obtained from the average gypsum proportions (V_E .), showing a greater number of panels with a frequency where ε_B is greater than or equal to zero. Similarly, there is an underestimation of the gypsum volume for the 10x10x5 meter panels. All the above is a consequence of small-scale variations in volume, especially in panels located near the contacts of the gypsum series with non-mineral units, both in the overburden and near non-mineral intercalations, where spatial heterogeneity affects interpolation.

The mean estimation error, ε_B determined by the difference between the volume of gypsum in the panel according to the deterministic model and the volume obtained from the proportions estimated for the panel according to 50 scenarios, by scaling to the 10x10x5 meter support from the network of 5 x 5 x 5 meter nodes; is 26.28 cubic meters, this error considers the sign of the deviations from the central value, which is equal to zero; however, the absolute average error obtained for these same differences amounts to 138.11 cubic meters, which represents approximately 27.6 percent of the volume of the panel.

The size of the mean square error obtained from the differences between the volume of gypsum calculated according to the deterministic model for gypsum, V_B .and the volume obtained from the model of the estimated gypsum proportions, V_E , is 201.31 cubic meters, which represents a 42 percent variation with respect to the total volume of the panels of 500 cubic meters, reflecting a high geological variability, which is very close to the value of REMC, to the value of the standard deviation obtained (208.66 cubic meters) between these two data series, which suggests that the errors are not concentrated in extreme values, but widely distributed. In the old geological exploration reports, which exist about the Punta Alegre [25] and more recent [23] deposits, which were referred to above, the complex geological constitution of the gypsum series of the deposit is highlighted, due to its genesis by redeposition of salts, gypsum and other rocks to the deposition basin during its formation.

 doi:10.20944/preprints202509.0114.v1

This point supports the results obtained in the research regarding the marked uncertainty in the estimation of the gypsum volume at a local scale.

Regarding the limitations to this study, it should be mentioned that the resolution used to discretize the block of 10 x 10 x 5 according to a pattern of dimensions of 5x5x5 meters to make the change of support by scaling, was insufficient to provide statistically more robust mean probability values. On the other hand, the number of simulated scenarios, while valid, could be increased to reduce the standard error.

There is no precedent in the country for the application of geostatistical simulation to determine uncertainty in these types of deposits, whose practical implication derived from this study justifies the use of stochastic models for heterogeneous areas, especially in complex geological contacts such as those existing in the Punta Alegre deposit.

5. Conclusions

The total volume of gypsum ore in the Punta Alegre deposit according to the reference deterministic rock model, amounts to 13.1 x 106 cubic meters (13.1 million cubic meters), on the other hand, the total volume obtained using the estimated proportions of gypsum amounts to 12.3 x 106 tons (12.3 million cubic meters); therefore, although there are differences between the two volumes, these are not significant and tend to compensate at the scale of the deposit, given that the global value of these differences only represents 6.51 % of the total volume of the useful mineral (gypsum).

However, if these volume differences are analyzed at a local scale, such as for the panels, the results achieved in this research show that there is a predominance of the systematic overestimation of the volume of gypsum in the panels, judging by the histogram of frequencies of classes of the estimation error; as a result of the effect on the boundaries, that is, on the contacts of the gypsum series with the non-mineral units represented by the overburden and "Inter-sterile", there being a significant variation in the volume of gypsum of the panels, judging by the figure obtained from the REMC, which is 201.31 cubic meters, denoting a high geological variability existing in the deposit.

When comparing the panel volume differences of the reference deterministic rock model with the panel volume of the model obtained using the estimated gypsum proportions, the difference represents a percentage change in gypsum volume of 42 percent.

The results achieved in this study suggest that the variations obtained at a local scale will have an indisputable influence on mining, given the differences in gypsum amounts in the panels at that scale. This should be considered for planning mining activities, especially in those panels close to the areas in contact with the intercalations of the lithologies that have been referred to in this study as non-mineral or sterile, such as limestones, dolomites, slates, clays, etc.

Although the results obtained positively validate the volume of gypsum resources that are estimated on a global scale according to the reference model; at the local scale the panels shown that there is a significant variation of this parameter, which will impact the gypsum resources that could be recovered by mining, given the existing uncertainty, which will cause an increase in mining dilution both internally and externally, due to the inclusion of non-mineral lithologies within the extracted mineral mass (gypsum) and that will be sent to the processing plant. On the other hand, it will cause deviations from the mining company's plan, in areas where there are panels where the volumes were previously accounted for by the deterministic model as mineral and are not truly exploitable.

Author Contributions: "Conceptualization, J.A.A., A.C. and E.E.; methodology, J.A.A.; software, J.A.A.; validation, J.A.A.; formal analysis, J.A.A. and E.E.; data curation, J.A.A; writing—original draft preparation, J.A.A. and R.R.; writing—review and editing, J.A.A., J.L.C, and D.A.M.; supervision, J.A.A. All authors have read and agreed to the published version of the manuscript.".

Funding: This research received no external funding.

Data Availability Statement: The data is unavailable due to privacy.

 doi:10.20944/preprints202509.0114.v1

Acknowledgments: The authors wish to thank the National Office of Mineral Resources of the Republic of Cuba for providing the information to realize and publishing of this work; to Opengeostat Consulting of Vancouver, Canada (http://opengeostat.com/) for the reviewing this manuscript and the valuable comments and observations made. The authors also thank the University of Moa (Cuba) and the Universidad Politécnica de Madrid (Spain) for their invaluable help. The authors have reviewed and edited the output and take full responsibility for the content of this publication.".

Conflicts of Interest: "The authors declare no conflicts of interest.".

References

1. Rossi, M.E., Deutsch, C.V. Mineral Resource Estimation. *Springer*. *Dordrecht Heidelberg New York London*; 2014. pp. 337. ISBN 978-1-4020-5716-8. DOI 10.1007/978-1-4020-5717-5.
2. Van Der Grijp, Y.M. Application of simulation techniques for modelling uncertainty associated with gold mineralization. A project report submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Science in Engineering by advanced coursework and research, 2014. pp. 70. Available online: https://core.ac.uk/download/pdf/39676262.pdf (accessed on 11 April 2025).
3. Deutsch, C.V. The Place of Geostatistical Simulation through the Life Cycle of a Mineral Deposit. *Minerals* 2023, 13, 1400. https://doi.org/10.3390/min13111400
4. Caers, J. Modeling Spatial Uncertainty. *Modeling Uncertainty in the Earth Sciences*. *John Wiley & Sons, Ltd*; Oxford, UK 2011. p.p. 93-106. ISBN:9781119992639. DOI:10.1002/9781119995920.
5. Dowd, P. Quantifying the Impacts of Uncertainty. Chapter 18. *In: Handbook of Mathematical Geosciences: Fifty Years of IAMG*. Springer Open, 2018. pp. 349-73. ISBN 978-3-319-78998-9. https://doi.org/10.1007/978-3-319-78999-6
6. Júnior, S., da Silva, A.A. Simulação de litotipos de depósito de minério de ferro com geoestatística de múltiplos pontos. *Universidade Federal Do Rio Grande Do Sul*. *Programa de Pós-Graduação em Engenharia de Minas*, 2013. pp. 102. Available online: https://lume.ufrgs.br/handle/10183/79828 (accessed on 16 february 2024)
7. Gómez-Hernández, J.J., Srivastava, R.M. One Step at a Time: The Origins of Sequential Simulation and Beyond. *Math Geosci*, 53, 2021. pp. 193–209. https://doi.org/10.1007/s11004-021-09926-0
8. Abulkhair, S., Madani, N. Stochastic modeling of iron in coal seams using two-point and multiple-point geostatistics: A case study. *Mining, Metallurgy & Exploration 39,* 2022. pp. *1313–1331.* https://doi.org/10.1007/s42461-022-00586-0.
9. Guardiano, F.B., Srivastava, R.M. Multivariate Geostatistics: Beyond Bivariate Moments. In: Soares, A. (eds) Geostatistics Tróia '92. Quantitative Geology and Geostatistics, vol 5, 1993. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-1739-5_12
10. Strebelle, S. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. Mathematical Geology 34, 2002. pp. 1–21. https://doi.org/10.1023/A:1014009426274
11. Remy, N., Boucher, A., Wu, J. Applied Geostatistics with SGeMS: A User's Guide. *Mathematical Geosciences. Cambirdge University press,* 2009;41:353-6. ISBN: 978-0-521-51414-9.
12. Tahmasebi, Pejman. Multiple point statistics: a review. In: Handbook of mathematical geosciences: Fifty years of IAMG. Springer Open, 2018. pp. 613-643. ISBN 978-3-319-78998-9. https://doi.org/10.1007/978-3-319-78999-6
13. Talesh Hosseini, S., Asghari, O., Torabi, S.A., Abedi, M. An Optimum Selection of Simulated Geological Models by Multi-Point Geostatistics and Multi-Criteria Decision-Making Approaches; a Case Study in Sungun Porphyry-Cu deposit, Iran. *Journal of Mining and Environment*, 11, 2, 2020. pp. 481-503. DOI: https://doi.org/10.22044/jme.2020.8710.1757
14. Avalos, S., Ortiz, J.M. Multiple-point statistics: tools and methods. Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report, 2020, paper 2020-03. pp. 33-60. Availablew online: https://qspace.library.queensu.ca/handle/1974/28556 (accessed on 11 february 2014).
15. Strebelle, S. Multiple-Point Statistics Simulation Models: Pretty Pictures or Decision-Making Tools? *Math Geosci* 53, 2021. pp.267–278. DOI: https://doi.org/10.1007/s11004-020-09908-8

 doi:10.20944/preprints202509.0114.v1

16. Osterholt, V., Dimitrakopoulos, R. Simulation of Orebody Geology with Multiple-Point Geostatistics—Application at Yandi Channel Iron Ore Deposit, WA, and Implications for Resource Uncertainty. *In: Dimitrakopoulos, R. (eds)* Advances in Applied Strategic Mine Planning. Springer, 2018. pp. pp 335–352. DOI: https://doi.org/10.1007/978-3-319-69320-0_22
17. Paithankar, A., Chatterjee, S. Grade and Tonnage Uncertainty Analysis of an African Copper Deposit Using Multiple-Point Geostatistics and Sequential Gaussian Simulation. Nat Resour Res 27, 2018, vol. 27. pp. 419–436. DOI: https://doi.org/10.1007/s11053-017-9364-1
18. Strebelle, S. Sequential simulation drawing structures from training images. *Stanford Center for Reservoir Forecasting. Ph.D. Thesis, Stanford University*, 2000. pp. 24. UMI Number: 3000105. Available on: https://github.com/SCRFpublic/snesim-standalone/blob/master/snesimtheory.ppt (accessed on 10 April 2024).
19. Amarante, F.A.N., Rolo, R.M., Costa, J.F.C.L. Assessing geologic model uncertainty - a case study comparing methods. REM - *International Engineering Journal. Mining REM, 2019. Int. Eng. J. 72 (4)*. Available on: https://www.scielo.br/j/remi/a/L3vMZxBqGrtf5x4qtB74dxL/?lang=en (accessed on 10 May 2024)
20. Iturralde-Vinent, M., Marrero, F. Nuevos datos sobre las estructuras diapíricas de Punta Alegre y Turiguanó, en la Provincia Ciego de Avila. *VIII Jornada Cientffica del Instituto de Geologia y Paleontologia*. Revista Ciencias de la Tierra y del Espacio, 1982, vol.9. pp. 47-55. Available on: http://www.redciencia.cu/geobiblio/paper/1982_Iturralde_roque_marrero.pdf (accessed on 10 May 2024)
21. Meyerhoff, A.A., Hatten, C.W. Diapiric Structures in Central Cuba1. *In: Braunstein J, O'Brien GD, editores*. Diapirism and Diapirs: a symposium. *American Association of Petroleum Geologists,* 1968. pp. 315-57. Available on: http://www.redciencia.cu/geobiblio/paper/1982_Iturralde_roque_marrero.pdf (accessed on 16 february 2024)
22. Sánchez Arango, J.R. Las evaporitas en Cuba y su relación con escenarios exploratorios de hidrocarburos. Memorias, Trabajos y Resúmenes V Convención Cubana de Ciencias de la Tierra (Geociencias' 2013) Sociedad Cubana de Geología, 2013. pp. 12. PETRO1-O2. SSN-2307-499X. Available on: http://www.redciencia.cu/geobiblio/paper/2013_Sanchez_Arango_PETRO1-O2.pdf (accessed on 16 february 2024)
23. Rojas-Consuegra, R., Tamayo Castellanos, Y., Torres Díaz, M., Pérez Peña, M.V. Reevalución estratigráfica Integral de bloques para la exploración de hidrocarburos. Internal Archive; Unpublished Report; National Office of Mineral Resources of the Republic of Cuba: La Habana, Cuba, 2020. Available on: https://www.minem.gob.cu/es/actividades/geologia/informes-concluidos (accessed on 16 february 2024)
24. Lukač, M. Estratigrafía y génesis de la sal gema en Punta Alegre y en Loma Cunagua, Provincia de Camagüey. Revista Tecnológica. 1969; 20-42. Available on: http://www.redciencia.cu/geobiblio/geobiblio.html#L (accessed on 16 february 2024)
25. Kazak, Y., Setién, C., Hurtado, J. Informe general sobre los trabajos detallados de exploración geológica en yeso, en los años 1964-1965,1967-1970 y 1976-1977, yacimiento Punta Alegre. Internal Archive. Oficina Nacional de Recursos Minerales; 1981 p. 255.
26. Instituto de Geología y Paleontología of Cuba. Léxico Estratigráfico de Cuba. 2014. Book. ISBN: 978-959-7117-57-5. Available on: https://isbn.cloud/9789597117575/lexico-estratigrafico-de-cuba-2013/(accessed on 12 february 2024)
27. Cala, E.L., Delgado, D.E.G. Stratigraphy of Cuba. In: Pardo Echarte, M.E. (eds) Geology of Cuba. Regional Geology Reviews. Springer nature, Cham., 2021. pp 143–188. DOI: https://doi.org/10.1007/978-3-030-67798-5_4
28. Ducloz, C. Apuntes sobre el yeso del Valle de Yumurí Matanzas. Apuntes sobre el yeso del valle de Yumuri Matanzas. In: Memorias de la Sociedad cubana de historia natural 'Felipe Poey', 1960, vol. 25, n° 1, p. 1–9. Available on: https://archive-ouverte.unige.ch/unige:156815 (accessed on 16 february 2024)
29. Meyerhoff, A.A., Hatten, C.W. Diapiric Structures in Central Cuba1. In: Braunstein J, O'Brien GD, editores. Diapirism and Diapirs: a symposium. American Association of Petroleum Geologists; 1968 [citado 16 de febrero de 2024]. p. 315-57. DOI: https://doi.org/10.1306/M8361C21 (accessed on 16 february 2024)
30. Sánchez Arango, J.R. Las evaporitas en Cuba y su relación con escenarios exploratorios de hidrocarburos. V Convención Cubana de Ciencias de la Tierra. IV Congreso Cubano de Petróleo y Gas *(PETROGAS´2013)*,

 doi:10.20944/preprints202509.0114.v1

Exploración de Petróleo y Gas, 2013; PETRO1-O2. pp. 12. Available on: http://www.redciencia.cu/geobiblio/paper/2013_Sanchez_Arango_PETRO1-O2.pdf

31. Instituto de Geología y Paleontología of Cuba. Mapa Geológico Hoja Punta Alegre a escala 1:100000, 2011.a ed. La Habana: Instituto de Geología y Paleontología (Cuba). Internal Report; Recovey from internal website: https://www.igp.minem.cu/
32. Ortíz, E. Re-estimación de recursos para la concesión del yacimiento «Yeso Punta Alegre». CEPRONIQUEL; 2018. La Habana (Cuba). pp. 79. Internal Report No.: 3281.
33. Palmer, L.W. Compositing and regularization of drillhole data for geostatistical resource estimation. Journal of the Southern African Institute of Mining and Metallurgy. Johannesburg, 2024. pp.331-8. ISSN 2411-9717. DOI: https://doi.org/10.17159/2411-9717/258/2024
34. Marwanza, I., Nas, C., Azizi, M.A., Simamora, J.H. Comparison between moving windows statistical method and kriging method in coal resource estimation. Journal of Physics: Conference Series. *Journal of Physics: Conference Series*, Vol. 1402, Issue 3. DOI:10.1088/1742-6596/1402/3/033016. Available on: https://iopscience.iop.org/article/10.1088/1742-6596/1402/3/033016/pdf
35. Anggara, B., Marwanza, I., Ahmad Azizi, M., Dahani, W. Subandrio. Fitting the variogram model of nickel laterite using root means square error in Morowali, Central Sulawesi. IOP Conference Series: Earth and Environmental Science. *IOP Conf. Ser.: Earth Environ. Sci.* 882 012042. DOI 10.1088/1755-1315/882/1/012042. Available on: https://iopscience.iop.org/article/10.1088/1755-1315/882/1/012042/pdf
36. Garnier-Villarreal, M. Introducción al análisis geoestadístico de datos en geociencias: teoría y aplicación. Revista geológica de América Central, ISSN 0256-7024, vol. 67, 2022. pp. 26-48. 2022. DOI: https://doi.org/10.15517/rgac.v67i0.51474. Available on: https://archivo.revistas.ucr.ac.cr/index.php/geologica/article/view/51474
37. Hodson, T. O.: Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev., 15. pp. 5481–5487. DOI: https://doi.org/10.5194/gmd-15-5481-2022, 2022.
38. Nimmo, M.J. ML and AI for resource estimation – what could possibly go wrong? Nothing! Everything! Mineral Resource Estimation Conference 2023. Perth, Australia; 2023.
39. Lindi, O.T., Aladejare, A.E., Ozoji, T.M., Ranta, J.P. Uncertainty Quantification in Mineral Resource Estimation. *Nat Resour Res,* 2024, vol. 33. pp. 2503–2526 (2024). DOI: https://doi.org/10.1007/s11053-024-10394-6