

# ANÁLISIS COMPARATIVO DE LA ESTIMACIÓN DE RECURSOS MINERALES MEDIANTE KRIGING Y REDES NEURONALES

*Se muestra un estudio comparativo de dos formas de estimar la concentración o ley mineral en lugares no muestreados. Una primera técnica geoestadística, muy popularizada en minería, llamada Kriging, la cual explota la variabilidad espacial como motor para la estimación, y otra segunda técnica, usando una Red Neuronal Artificial de dos tipos; perceptrón multicapa y de función de base radial. El estudio evalúa cuales son las ventajas de cada técnica y cuándo es viable combinarlas para mejorar la precisión de las estimaciones.*

Heber Hernández Guerra

Académicos Ingeniería Civil en Minas, FINARQ UCEN

Fred Camus Yeomans

Académicos Ingeniería Civil en Minas, FINARQ UCEN



Debido a limitaciones económicas y físicas, nuestra comprensión de los recursos minerales es limitada. Como resultado, la ley mineral, forma y cantidad del recurso permanecen desconocidas. Tradicionalmente, este problema se ha abordado utilizando el método geoestadístico de Kriging, donde la ley mineral se estima en ubicaciones sin mediciones utilizando valores de leyes conocidas cercanas. Básicamente, se asignan pesos a las leyes conocidas y sus productos se suman en una combinación lineal no sesgada que minimiza la varianza del error. La ventaja de este método radica en su cálculo de pesos a través de un modelo de variabilidad espacial conocido como semivariograma. Sin embargo, su desventaja es su dependencia de la suposición de aditividad, estacionariedad, linealidad y baja precisión en áreas escasamente muestreadas. Varios estudios han abordado el problema de la estimación de recursos minerales como un problema de regresión utilizando redes neuronales, que no están sujetas a las limitaciones del Kriging. La ley mineral se define como una función objetivo que depende de un conjunto de características de entrada, y el modelo es el resultado de un proceso de entrenamiento utilizando aprendizaje supervisado. La desventaja de este enfoque es que el modelo no reproduce el valor de la ley mineral y no considera su dependencia espacial. El propósito de este estudio es evaluar en qué escenarios de variabilidad espacial, muestreo y disponibilidad de información geológica es posible mejorar sinérgicamente el enfoque geoestadístico de Kriging combinándolo con una red neuronal. Para lograr esto, se simulan varios escenarios realistas y se realiza una comparación de resultados entre Kriging, una red neuronal de perceptrón multicapa, una red neuronal de función de base radial, un modelo combinado flexible y un modelo combinado rígido. Los resultados indican que un modelo combinado será una opción más adecuada que Kriging en tres casos: En situaciones estrictamente no estacionarias, cuando una red neuronal se entrena con el modelo de dominio de estimación geológica, o cuando se entrena con variables auxiliares que están más ampliamente muestreadas y correlacionadas significativamente con la ley mineral.

## Introducción

A lo largo de la historia, la humanidad ha obtenido beneficios aprovechando las riquezas de la Tierra. Los recursos minerales extraídos y utilizados por los humanos han tenido un impacto significativo en el desarrollo económico y social de las naciones y civilizaciones. Estos recursos minerales se emplean para satisfacer una amplia gama de necesidades humanas, que van desde la energía hasta la construcción, las industrias químicas y farmacéuticas, automotriz, electrónica e incluso la construcción de naves espaciales. Un recurso mineral se define comúnmente como una concentración de material sólido económicamente valioso en o sobre la corteza terrestre, en una forma, ley y cantidad que proporciona perspectivas razonables para una extracción económica eventual. La necesidad de estimación surge del hecho de que la información obtenida de la exploración minera a menudo es fragmentada o incompleta, lo que resulta en una parte significativa del depósito mineral con concentraciones de recursos desconocidas. La Estimación de Recursos Minerales es una de las etapas más importantes y críticas en la cadena de valor minera. Hace sesenta años, esta disciplina se basaba únicamente en herramientas estadísticas rudimentarias. El desarrollo de métodos de estimación geoestadística tuvo sus comienzos en la década de 1950. Sin embargo, fue en la siguiente década cuando se formalizó matemáticamente. Estos métodos consideran la variabilidad espacial de los recursos y su modelado a través de una función conocida como semivariograma, que mide la autocorrelación espacial de una variable regionalizada. Esta información se incorpora luego en una estimación lineal no sesgada y de mínima varianza. La introducción del interpolador espacial conocido como Kriging sentó las bases para la estimación espacial de los recursos minerales. Este término genérico se aplica a una variedad de métodos de estimación que buscan minimizar el error a través de un procedimiento de mínimos cuadrados. Desde entonces, ha habido un rápido progreso en los métodos geoestadísticos, impulsado por su capacidad para incorporar la variabilidad espacial del recurso y reducir la varianza de la estimación. Actualmente, existen varios tipos de Kriging, que difieren principalmente en las suposiciones que hacen sobre la estacionariedad, que se expresan como condiciones sobre el conjunto de pesos para cada punto de datos. Kriging ha sido discutido en varias referencias clásicas. La elección del tipo de Kriging depende del entorno geológico, la cantidad de información disponible y la función de distribución de probabilidad del recurso. El método de estimación más comúnmente utilizado es el Kriging Ordinario, debido a su versatilidad para dife-

rentes escenarios de muestreo espacial, adaptabilidad a diversas estructuras espaciales y la necesidad solo de datos medidos dentro del depósito mineral.

Aunque los métodos geoestadísticos son ampliamente estudiados y validados para, no se consideran infalibles y están sujetos a críticas. La suposición de estacionariedad y linealidad para su uso y la subjetividad en el proceso de modelado espacial, por ejemplo, son tres aspectos que podrían mejorarse. Es en esta oportunidad que surge el Aprendizaje Automático. Este se clasifica como un subcampo de la Inteligencia Artificial, que se define como la ciencia y la ingeniería de crear máquinas inteligentes. Su enfoque radica en el estudio y la aplicación de algoritmos informáticos que permiten a los sistemas mejorar automáticamente a través de la experiencia, sin ser programados explícitamente. Estos algoritmos tienen la capacidad de aprender y modelar patrones no lineales complejos en conjuntos de datos extensos, lo que les permite adaptarse a una amplia variedad de aplicaciones en las geociencias.

En las últimas dos décadas, los métodos basados en Machine Learning han ganado popularidad en la investigación de estimación de recursos. Varios estudios han destacado que estos métodos han surgido como modelos de predicción y alternativas importantes a la geoestadística. El algoritmo más comúnmente utilizado es el de Redes Neuronales. Estas redes son modelos computacionales que proporcionan una simplificación abstracta del cerebro humano. Conceptualmente, imitan el funcionamiento de las redes neuronales biológicas para reconocer relaciones en un conjunto de datos y predecir valores de salida para datos de entrada dados. Están compuestas por capas interconectadas de nodos que representan neuronas artificiales.

## Desarrollo

### Geoestadística

La geoestadística es una ciencia aplicada que se enfoca en modelar la continuidad espacial de los fenómenos naturales  $Z(x_i)$ . Esta información se utiliza para estimar valores en lugares no muestreados utilizando un método lineal óptimo no sesgado.

$$Z = \{Z(x_i), x_i \in D, i=1, 2, 3, \dots, n\}$$

La función aleatoria se caracteriza por su función de distribución de probabilidad n-variada.

$$F_{Z(x_1), Z(x_2), \dots, Z(x_n)}(z_1, z_2, \dots, z_n) = \Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2, \dots, Z(x_n) \leq z_n]$$

Y se resume prácticamente por los dos primeros momentos. La esperanza matemática de primer orden:

$$E[Z(x)] = m(x)$$

La varianza, covarianza y semivariograma como momentos de segundo orden.

$$\sigma^2(x) = \text{Var}[Z(x)] = E[\{Z(x) - m(x)\}^2]$$

$$C(x_i, x_u) = E[\{Z(x_i) - m(x_i)\}\{Z(x_u) - m(x_u)\}]$$

$$2\gamma(x_i, x_u) = \text{Var}[Z(x_i) - Z(x_u)]$$

Modelar la variabilidad espacial es un proceso estándar de los analistas de recursos minerales, donde durante los últimos sesenta años, el semivariograma ha sido la herramienta más aplicada en la minería.

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^n [Z(x_i) - Z(x_i+h)]^2 \quad \forall x_i, h, x_i+h \in D$$

$\gamma(h)$  representa el valor estimado del semivariograma para un desfase  $h$ , y se llama variograma experimental,  $n(h)$  es el número de pares de puntos separados por un desfase  $h$ ,  $Z(x_i)$  es el valor del atributo regionalizado en el punto  $x_i$ ,  $x_i$  y  $x_i+h$  son dos puntos en el dominio espacial  $D$  separados por un desfase  $h$ . Para modelar las dependencias espaciales en un conjunto de datos, es necesario ajustar un modelo matemático formalizado al semivariograma experimental. Los modelos teóricos comúnmente utilizados son esféricos, exponenciales y gaussianos. Todas estas funciones se caracterizan por dos parámetros principales: meseta ( $C$ ) y rango ( $a$ ).

$$\gamma(h)=\begin{cases} C \left( \frac{3}{2} \frac{h}{a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right) & h < a \\ C & h \geq a \end{cases}$$

Además, estos modelos pueden estar acompañados por un término aditivo llamado efecto pepita, el que describe la variabilidad observada entre muestras que están espaciadas muy cerca una de otra  $C_0$

$$\gamma(h)=\begin{cases} C_0+C \left( \frac{3}{2} \frac{h}{a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right) & h < a \\ C_0+C & h \geq a \end{cases}$$

Para la estimación se define la variable regionalizada como  $Z(x_i)$  de la cual se tienen  $n$  observaciones  $Z(x_1), Z(x_2), \dots, Z(x_n)$ . El Kriging consiste en predecir  $Z^*(x_0)$ , siendo  $x_0 \notin \{x_1, x_2, \dots, x_n\}$  una posición no muestreada. Esta técnica de interpolación está basada en algoritmos de regresión generalizados por mínimos cuadrados.

$$Z^*(x_0) - m(x_0) = \sum_{i=1}^n \lambda_i(x_0) \cdot (Z(x_i) - m(x_i)),$$

Siendo  $\lambda_i$  los valores de peso,  $m(x_0)$  y  $m(x_i)$  son los valores esperados por  $Z(x_0)$  y  $Z(x_i)$  respectivamente. El objetivo es minimizar la varianza del error  $\sigma_E^2(x)$ , que se da por la siguiente expresión:

$$\sigma_E^2(x) = \text{Var} [Z^*(x) - Z(x)],$$

Siendo  $Z^*(x)$  el valor estimado y  $Z(x)$  el valor real observado. La varianza del error se minimiza bajo la restricción de imparcialidad; es decir,  $E[Z^*(x) - Z(x)] = 0$ . La variable regionalizada se descompone en una componente residual  $e(x)$  y una determinística que define la tendencia de  $m(x)$ :

$$Z(x) = e(x) + m(x)$$

La componente residual se modela como una variable aleatoria estacionaria con media cero y covarianza  $C_e(h)$  siendo:

$$\begin{aligned} E[e(x)] &= 0 \\ \text{Cov}[e(x), e(x+h)] &= C_e(h) \end{aligned}$$

De esta manera, el valor esperado de la variable aleatoria  $Z$  en una cierta ubicación  $x$  está dado por el valor de la componente de tendencia en esa ubicación, es decir,  $E[Z(x)] = m(x)$ .

## Redes neuronales

Una Red Neuronal es una estructura computacional inspirada en el procesamiento neuronal biológico, capaz de percibir, reconocer patrones y realizar predicciones. El procesamiento de información ocurre en paralelo a través de neuronas interconectadas, lo que permite un mapeo eficiente no lineal de entrada-salida. Esta propiedad hace los modelos de red neuronal adecuados para la estimación de la ley mineral, ya que esta puede ser definida como una característica dependiente, entrenada con otras características relacionadas independientes. Las redes MLP (Multilayer Perceptron) pertenecen a la arquitectura de alimentación múltiple de capas, cuyo entrenamiento se realiza mediante un proceso supervisado. El flujo de información a través de la red es unidireccional, desde la capa de entrada hasta la capa de salida. En general, se puede dividir en tres partes: la capa de entrada, las capas ocultas ( $h=1, 2, \dots, n$ ), y la capa de salida. La capa de entrada recibe información del entorno externo, que generalmente se normalizan para mejorar la precisión numérica en las operaciones matemáticas de la red. Las capas ocultas están compuestas por neuronas que se encargan de extraer patrones asociados con el proceso o sistema analizado. Estas capas realizan la mayor parte del procesamiento interno de la red. La capa de salida es responsable de generar y presentar las salidas finales de la red (en el caso de estimación, una única neurona compone esta capa). Estas salidas se derivan del procesamiento realizado por las neuronas en las capas anteriores. Cada neurona en las capas ocultas es responsable de recibir información de la capa anterior, procesarla y propagarla hacia adelante para llegar a la capa de salida.

$$y' = f\left(\sum_j x_j w_j + \theta\right),$$

donde  $x_j$  son las entradas a la unidad, los  $w_j$  son los pesos,  $\theta$  es el término de sesgo,  $f()$  es la función de activación no lineal, y  $y$  es la activación de la unidad. Extendiendo este concepto a múltiples capas:

$$\hat{y}_k(p) = f\left(\sum_{j=1}^m x_{jk}(p) \cdot w_{jk} + \theta_k\right)$$

Donde  $m$  es el número de entradas para la neurona  $k$  desde la capa de salida. La función de activación de cada neurona también es un hiper-parámetro de la red neuronal y puede ser dinámico. Un tipo de red utilizado en este estudio (Red de Función de Base Radial) es una simplificación de la presentada anteriormente (Perceptrón Multicapa), donde solo utiliza una única capa oculta y funciones de activación de base radial.

La propagación hacia adelante de la red produce una magnitud de error. Para disminuir este error, es necesario modificar los pesos de la red hasta que alcance un punto mínimo, lo cual se logra con la retropropagación ajustada con un algoritmo de optimización como el descenso por gradiente o alguna de sus variantes.

## Modelos combinados

Las redes neuronales presentan limitaciones tales como no reproducir el valor de la ley mineral en los lugares muestreados, y no incorporar la correlación espacial de forma explícita. Dos rasgos que sí consideran los métodos geostatísticos.

Una solución a las limitaciones de ambos enfoques consiste en combinar las redes neuronales con Kriging utilizando la varianza de estimación de Kriging como factor de ponderación. Entendiendo que en ubicaciones donde se midió la ley del mineral, su ponderación será cero. Entonces, si la varianza de estimación de Kriging es cero, todo el peso se asigna a la estimación de Kriging, y viceversa, en el caso de que la varianza sea uno, todo el peso se asigna a la estimación del regresor de aprendizaje automático.

$$\begin{aligned} Z^*_{\text{hibrido}}(x) &= Z^*_{\text{RN}} \cdot w(x) + (1 - w(x)) \cdot Z^*_{\text{KR}}(x) \\ w(x) &= (\sigma_E^2(x))^b \end{aligned}$$

Este estudio emplea dos métodos para determinar el peso en el modelo híbrido. El primero se llama modelo combinado flexible (FCM), donde el peso se distribuye según la varianza, y el segundo se denomina modelo combinado rígido (RCM) y asigna todo el peso a Kriging o todo el peso a la red neuronal basado en un valor de corte  $H$  para la varianza del error de Kriging estandarizado.

$$\begin{aligned} \text{Si } (\sigma_E^2(x) > H); Z^*_{\text{hibrido}}(x) &= Z^*_{\text{RN}}(x) \\ \text{Si } (\sigma_E^2(x) \leq H); Z^*_{\text{hibrido}}(x) &= Z^*_{\text{KR}}(x) \end{aligned}$$

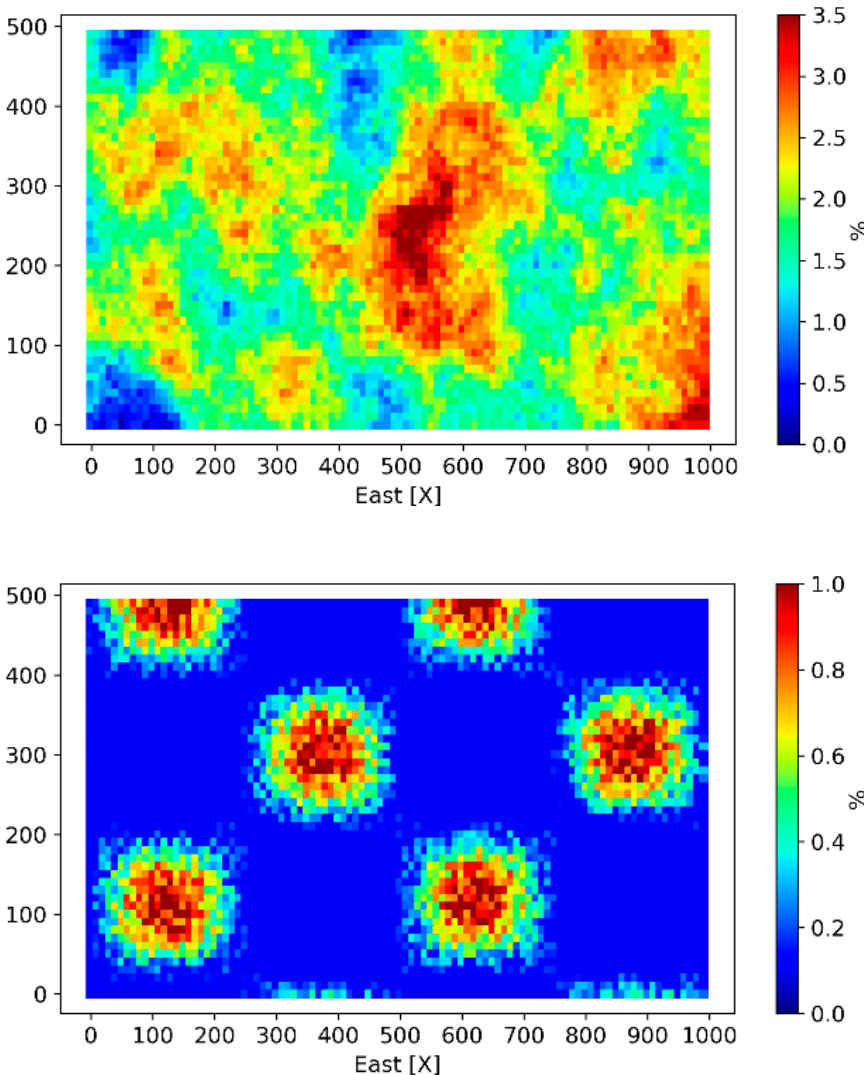
**Tabla 1.** Ventajas y limitaciones entre redes neuronales y geoestadística.

| Condiciones                                     | Red Neuronal | Geoestadística |
|-------------------------------------------------|--------------|----------------|
| Reproduce los datos en su ubicación             | No           | Sí             |
| Considera la autocorrelación espacial           | No           | Sí             |
| Requiere estacionariedad                        | No           | Sí             |
| Requiere linealidad                             | No           | Sí             |
| Requiere variograma                             | No           | Sí             |
| Funciona bien con datos de alta dimensionalidad | Sí           | No             |

### Análisis

Se simula la distribución espacial de una ley mineral hipotética en un dominio 2D, en el primer caso asumiendo estacionariedad de segundo orden y estadísticas de simetría. Posteriormente, se aplica una transformación afín correctiva a la distribución de la variable corrigiendo la media y la desviación estándar. En el segundo caso, se construye un escenario espacialmente no estacionario a través de la función:

$$Z(X,Y) = \left( \sin \sin \left( \frac{x}{80} \right) \cdot \cos \cos \left( \frac{y}{120} \right) + \text{uniforme}(-0.2,0.2), 0.1 \right)$$



**Figura 1**

Ley mineral simulada caso 1 (izquierda), caso 2 (derecha).

En las etapas iniciales de la exploración minera, a menudo se realiza un muestreo regular, donde las mediciones están dispersas espacialmente y equidistantes entre sí. Más adelante, a medida que avanza la exploración, el área se delimita y el muestreo se vuelve irregular, con conglomerados en áreas de mayor ley mineral. Esta última consideración se asume para el muestreo de los escenarios respectivos discutidos a continuación.

Para el primer caso (Figura 1, izquierda), hay tres escenarios de muestreo, cada uno con un nivel diferente de información asociada con las coordenadas espaciales.

**A.** Muestreo irregular donde solo se conoce la ley mineral (Figura 2, I.A).

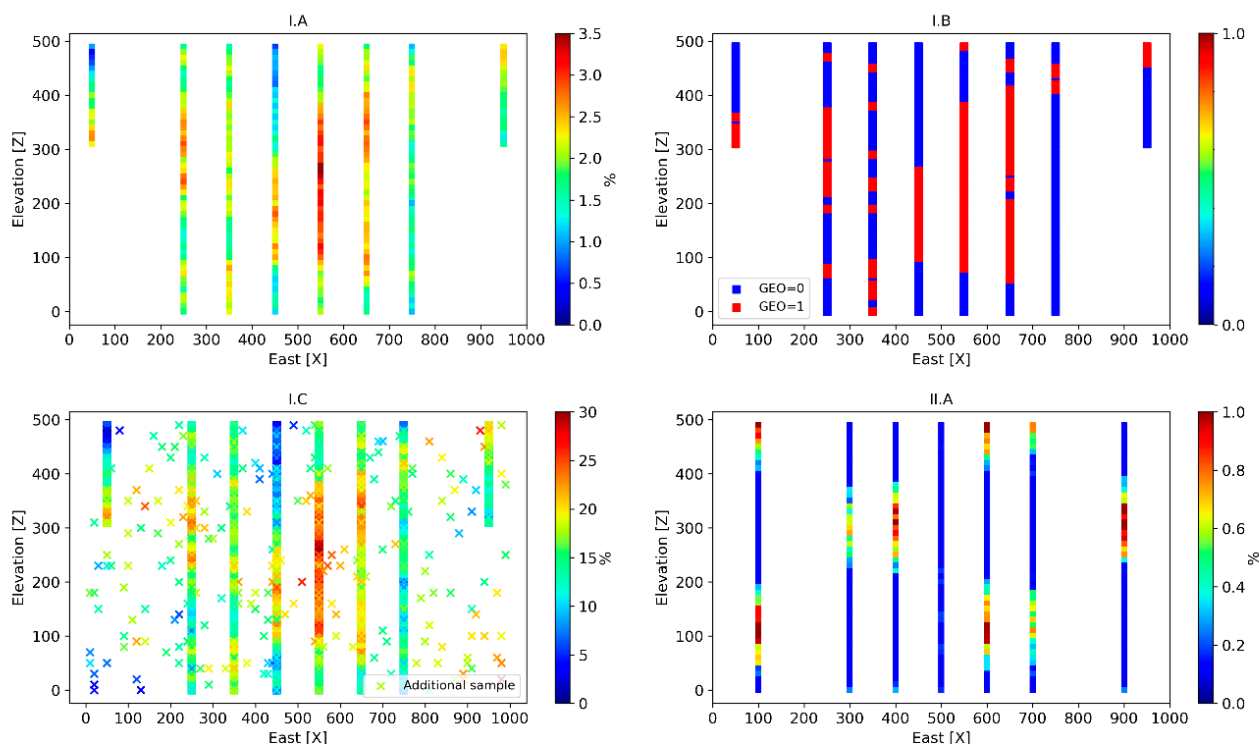
**B.** Muestreo irregular homotópico donde tanto la ley mineral como el control geológico (variable categórica) son conocidos (Figura 2, I.B).

**C.** Muestreo irregular parcialmente heterotópico donde se conocen la ley mineral y una variable auxiliar muestreada en su mayoría (variable numérica continua) (Figura 2, I.C).

Para el segundo caso no estacionario (Figura 1, derecha), donde la ley mineral presenta menos continuidad y más complejidad, solo habrá un escenario de muestreo irregular donde la ley es conocida como la variable exclusiva (Figura 2, II.A). Esto se basa en la hipótesis de que las redes neuronales en tales casos producen mejores resultados que el Kriging bajo cualquier escenario de muestreo.

### Escenario de muestreo caso I-A

Bajo este primer escenario de muestreo, resulta que el uso de un modelo combinado no mejora significativamente los resultados del enfoque geoestadístico, lo que significa que la mejora no justifica el costo en tiempo del proceso. Esto se atribuye al hecho de que la red neuronal solo se entrena con coordenadas espaciales etiquetadas para la ley mineral, proporcionando muy poco complemento al Kriging en áreas fuera del condicionamiento.



**Figura 2**

Muestreo caso 1 (I.A, I.B, I.C) y caso 2 (II.A).

**Tabla 2.** Métricas resultantes caso I-A

|                      | NN   | OK   | FCM  | RCM  |
|----------------------|------|------|------|------|
| <b>R</b>             | 0.64 | 0.64 | 0.66 | 0.67 |
| <b>R<sup>2</sup></b> | 0.42 | 0.39 | 0.42 | 0.43 |
| <b>RMSE</b>          | 0.43 | 0.43 | 0.42 | 0.42 |
| <b>MSE</b>           | 0.19 | 0.18 | 0.17 | 0.17 |
| <b>MAE</b>           | 0.33 | 0.31 | 0.31 | 0.31 |

#### Escenario de muestreo caso I-B

En general, al entrenar la red neuronal con un dominio geológico que ejerce un fuerte control sobre la ley mineral, esta información adicional suele conducir a resultados significativamente mejores en comparación con el Kriging solo.

**Tabla 3.** Métricas resultantes caso I-B

|                      | NN   | OK   | FCM  | RCM  |
|----------------------|------|------|------|------|
| <b>R</b>             | 0.65 | 0.64 | 0.76 | 0.74 |
| <b>R<sup>2</sup></b> | 0.41 | 0.39 | 0.57 | 0.54 |
| <b>RMSE</b>          | 0.42 | 0.43 | 0.36 | 0.37 |
| <b>MSE</b>           | 0.18 | 0.18 | 0.13 | 0.14 |
| <b>MAE</b>           | 0.31 | 0.31 | 0.27 | 0.27 |

#### Escenario de muestreo caso I-C

Se puede incluir en el entrenamiento de la red neuronal una variable auxiliar numérica, mayormente muestreada, que esté correlacionada significativamente con la ley mineral (95% en este ejemplo). Esta adición mejora la precisión de las predicciones en áreas fuera de la zona de condicionamiento del Kriging, lo que resulta en modelos combinados más efectivos.

**Tabla 4.** Métricas resultantes caso I-C

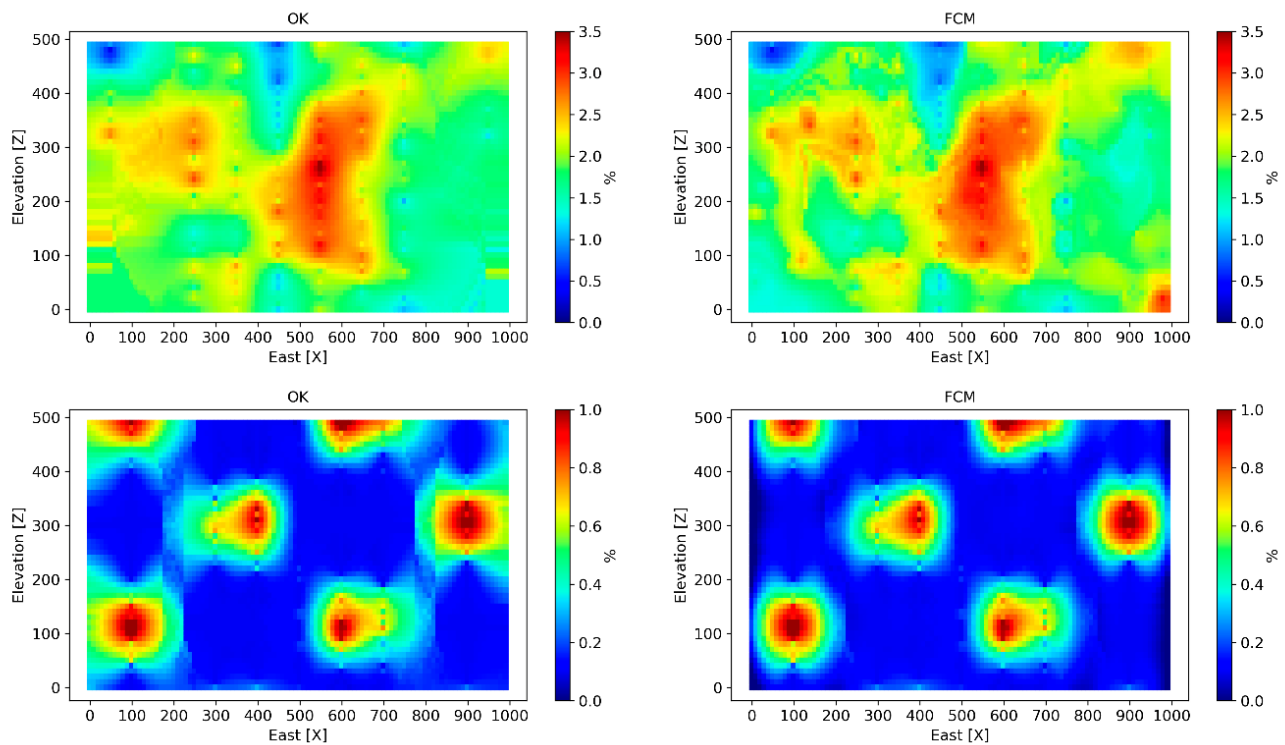
|                      | NN   | OK   | FCM  | RCM  |
|----------------------|------|------|------|------|
| <b>R</b>             | 0.78 | 0.64 | 0.79 | 0.76 |
| <b>R<sup>2</sup></b> | 0.60 | 0.39 | 0.63 | 0.58 |
| <b>RMSE</b>          | 0.35 | 0.43 | 0.33 | 0.36 |
| <b>MSE</b>           | 0.12 | 0.18 | 0.11 | 0.13 |
| <b>MAE</b>           | 0.27 | 0.31 | 0.25 | 0.27 |

#### Escenario de muestreo caso II

En casos de variabilidad compleja, como en este caso, el Kriging puede enfrentar desafíos para modelar un semivariograma representativo. En tales situaciones, es recomendable utilizar un rango corto únicamente con el fin de incorporar la reproducción de la ley mineral en un modelo combinado, donde la influencia más significativa se logra mediante el uso de una red neuronal.

**Tabla 5.** Métricas resultantes caso II

|                      | NN   | OK   | FCM  | RCM  |
|----------------------|------|------|------|------|
| <b>R</b>             | 0.89 | 0.86 | 0.90 | 0.89 |
| <b>R<sup>2</sup></b> | 0.75 | 0.71 | 0.80 | 0.76 |
| <b>RMSE</b>          | 0.13 | 0.14 | 0.11 | 0.13 |
| <b>MSE</b>           | 0.02 | 0.02 | 0.01 | 0.02 |
| <b>MAE</b>           | 0.09 | 0.09 | 0.07 | 0.09 |



**Figura 3**

Mapa de estimación de Kriging y FCM en caso I-C (arriba), caso II (abajo).

## Conclusiones

Los métodos geoestadísticos fueron diseñados para estudiar la variabilidad espacial e incorporarla en la estimación de recursos minerales a través de modelos matemáticos. Sin embargo, su uso está sujeto a condiciones desafiantes y ciertas críticas, como la aditividad, estacionariedad, linealidad, modelado subjetivo de semivariogramas, requisitos de expertos, extenso preprocesamiento de datos y complejidad en casos multidimensionales. Por otro lado, las redes neuronales no requieren las mismas condiciones que los métodos geoestadísticos. Ofrecen ventajas en términos de eficiencia temporal, reducción de preprocesamiento de datos, requisitos menos intensivos en mano de obra y excelente rendimiento en casos multidimensionales. Esto permite la fácil incorporación de varios dominios geológicos y otras variables para mejorar la etiquetación de ley mineral para el entrenamiento. Sin embargo, las redes neuronales omiten dos propiedades cruciales: la dependencia espacial de la ley mineral y la necesidad de reproducirla en las ubicaciones medidas. En conclusión, ambos enfoques tienen sus propias condiciones y requisitos que seguirán manteniéndose, lo que resultará en un aumento tanto en el tiempo como en los costos computacionales. Sin embargo, también ofrecen beneficios que, cuando se combinan, mejoran significativamente la estimación de recursos minerales. Un modelo combinado se convierte en una alternativa justificable en escenarios estrictamente no estacionarios, cuando hay información detallada disponible sobre el modelo geológico, o existe una variable auxiliar mayormente muestreada que mantiene relación con la ley mineral.

## Referencias

- [1] Erdogan Erten, G., Nat Resour Res 31, 191–213 (2022).
- [2] Galetakis, M. et al., Mater. Proc. 5, 122 (2021) p216-234.
- [3] Samanta, B., Math Geol 38, 175–197 (2006).
- [4] Samanta, B., Exploration and Mining Geology, 11(1-4), 69–76 (2002).
- [5] Tsae, N.B. et al., Minerals 13, 658 (2023).