

Estimación de recursos minerales mediante un método convencional y Machine Learning (caso de estudio proyecto Carlés)

Pamela Chávez¹, Gilson Peñaloza¹

¹Universidad Central del Ecuador. Facultad de Ingeniería en Geología, Minas, Petróleos, y Ambiental

Resumen

Un recurso mineral es una concentración natural de algún elemento o compuesto de la corteza terrestre, que puede ser extraído o procesado con los medios tecnológicos disponibles. La estimación de recursos y reservas de un yacimiento tiene por objeto determinar la cantidad, calidad y posibilidad de explotación comercial del mineral que lo compone. En el presente documento se realiza la estimación de oro del proyecto Carles, mediante el uso de diferentes softwares, los mismos que conjuntamente con un análisis estadístico, se genera finalmente un modelo y las estimaciones correspondientes, no obstante, se aplica otra metodología mediante Machine Learning, tomando en cuenta los compósitos obtenidos de RecMin, y para el caso de estudio los dos cuerpos definidos en base a la litología de los sondeos, posteriormente mediante el uso de algoritmos y su posterior entrenamiento y evaluación se logra obtener la predicción en cuanto a las reservas minerales. Por tanto, el objetivo de este trabajo es determinar los recursos minerales y realizar el modelamiento correspondiente mediante las dos metodologías.

Palabras clave: *Recurso, Reserva, RecMin*

1. Introducción

El propósito de un proyecto minero es generar utilidad económica de a lo menos un determinado elemento metálico o no metálico de interés, el cual debe cumplir con ciertas características como estar espacialmente agrupado generando concentraciones anómalas, encontrarse en grandes volúmenes y tener factibilidad en su extracción (Hernández, 2017), para la determinación de los mismos se realiza el proceso de estimación de recursos minerales el cual es una etapa que se encarga de describir la zona de interés en cuanto a la concentración de los elementos presentes, dichas concentraciones generalmente son representadas por bloques de dimensiones determinadas, las mismas que son estimadas en la forma de objetos puntuales compuestos por la posición de

los sondeos originales y la información de estudios químicos o geológicos (Lara, 2020). El depósito de oro Carlés se encuentra ubicado en España específicamente en la provincia de Asturias, desde el punto de vista geológico, la mineralización de Carlés se enclava en la parte más occidental de la Zona Cantábrica y sobre el cinturón aurífero del río Narcea, la mineralización comprende una rica paragénesis de sulfuros y sulfosales metálicos acompañando al oro, entre los que sobresalen, por su presencia masiva, la calcopirita, la bornita y la arsenopirita (IGME, 1988).

El objetivo del presente estudio es realizar la estimación mineral del proyecto Carlés, mediante la información proporcionada por los sondeos a través del uso de los softwares especializados RECMIN y

SGEMS, y también mediante la aplicación de Machine Learning, con la finalidad de reconocer las áreas de mayor concentración mineral.

2. Estado del Arte

La estimación de recursos minerales es una actividad muy importante dentro del ámbito geológico - minero que se realiza con el objetivo de determinar una clasificación de recursos que pueda llegar a tener un depósito mineral y así establecer la factibilidad que representa el mismo, es un tema del cual se ha realizado diversas publicaciones tanto de manera convencional mediante el uso de herramientas como Recmin, Leaprog, Vulkan y otras más, así como la implementación en los últimos años de la IA para la estimación. Entre los diversos trabajos, se ha considerado cuatro principalmente, dos de ellos realizados mediante métodos convencionales, mientras que los dos restantes aplicando Machine Learning, dichos trabajos son los siguientes:

2.1. Aplicaciones de métodos convencionales

Manteca, J. (2015). Modelización de un yacimiento mineral polimetálico mediante el uso de Recmin, 122(5), 2-6.

- Realizó la modelización y evaluación mineral de un depósito VMS situado en la faja pirítica Ibérica a través del uso de la información de 18 sondajes, para la cual elaboró un modelo de bloques en el cual aplicó el método del inverso cuadrado de las distancias permitiéndolo así determinar la estimación de los recursos y reservas del depósito. Finalmente mediante los datos obtenidos realiza un estudio de viabilidad económica del mismo, con el objetivo de establecer su factibilidad.

Lara, R. (2020). Estimación de recursos minerales en dominios geometalúrgicos. Universidad de Concepción.

- En su trabajo realizó la estimación de recursos minerales, mediante el estudio de dos casos. En el primer caso, al cual denomina caso simulado crea 100 depósitos a los cuales divide en dominios geometalúrgicos, para posteriormente mediante la utilización de sondajes determinar la estimación de tales dominios y comparar sus resultados con la base de datos original, para el segundo caso al cual denomina caso real, realiza una estimación a través de la aplicación del modelo de bloques en los mismos escenarios. Finalmente indica que la estimación resulta ser mejor para el primer caso acercándose a un 99 % de precisión, con una disminución del RMSE para el Cu del 10.8 % en la técnica del inverso de la distancia, 10.4 % para el krigging ordinario, y 10.4 % para el co-krigging. .

2.2. Aplicaciones de ML

Galetakis, M. (2022). Estimation of Mineral Resource with Machine Learning Techniques.

- En su trabajo realizó la estimación de la ley y el cálculo de reservas de un depósito mineral de Cu, mediante la aplicación de técnicas de sistemas de inferencia difusa adaptativos o entre sus siglas denominados como ANFIS y también redes neuronales artificiales (NN), específicamente desarrolla una NN con retropropagación y dos ANFIS tipo sugeno para la estimación. Cabe indicar que para determinar la estimación, se basó en datos proporcionados por sondajes, una vez que realizó el análisis de los mismos, procedió a dividir los datos,

y aplicar un porcentaje de los mismos para el entrenamiento, validación, y el porcentaje restante para el test o prueba. Finalmente, los datos que obtuvo mediante la técnicas aplicadas fueron comparados con los obtenidos a través de métodos geoestadísticos (kriging), concluyendo que tales técnicas presentan potencial para ser utilizadas como estimadores de reservas de mineral.

Dutta et al.(2010). Machine Learning Algorithms and Their Application to Ore Reserve Estimation of Sparse and Imprecise Data. JILSA, 1(2010), 86-95.

- En su trabajo implementa los algoritmos de Machine Learning principalmente Redes Neuronales, y Support Vector Machine (SVM) para la estimación de reservas minerales, dicha estimación la realiza a través del uso de datos de perforación que presenta las propiedades típicas de escasez e imprecisión que podrían asociarse a un conjunto de datos mineros. Una vez obtenidos los datos de los resultados mediante los algoritmos de Redes Neuronales y SVM, compara los mismos con resultados obtenidos mediante el método geoestadístico de Krigging Ordinario (OK). Finalmente para determinar el rendimiento que tuvieron los modelos utiliza los criterios del Error Medio (ME), el Error Medio Absoluto (MAE), el Error Cuadrático Medio (RMSE), y el Coeficiente de determinación R².

La estimación de recursos minerales es una actividad de suma importancia puesto que mediante su aplicación se puede determinar la factibilidad de un proyecto minero. Mediante el presente documento se pretende realizar la estimación de recursos minerales a través del modelo de bloques mediante

la aplicación del software RecMin y Machine Learning, con la finalidad de establecer las zonas más enriquecidas del depósito mineral Carlés.

3. Metodología

La mineralización aurífera de Carles se localiza en el flanco SW del sinclinal de Cornellana, en el término municipal de Salas (Asturias), a la altura del PK 15 de la vía AS-15, entre las localidades de La Vega y Soto de los Infantes, en la margen izquierda (norte) del río Narcea a 80 m (Figura 1).



Figura 1: Mapa de ubicación

La metodología planteada para la realización del presente proyecto se encuentra en base al siguiente diagrama de flujo (Figura 2).

El diagrama se encuentra clasificado en una serie de colores, el color gris comprende la preparación de los datos, el color verde se relaciona con la parte estadística, principalmente en lo que corresponde al análisis exploratorio de datos, el color celeste se refiere al análisis estructural y finalmente el color naranja tiene que ver con la parte correspondiente a la estimación.

En el desarrollo de la metodología se evidenciará el uso de diferentes softwares y su importancia en la evolución y tratamiento de la información.

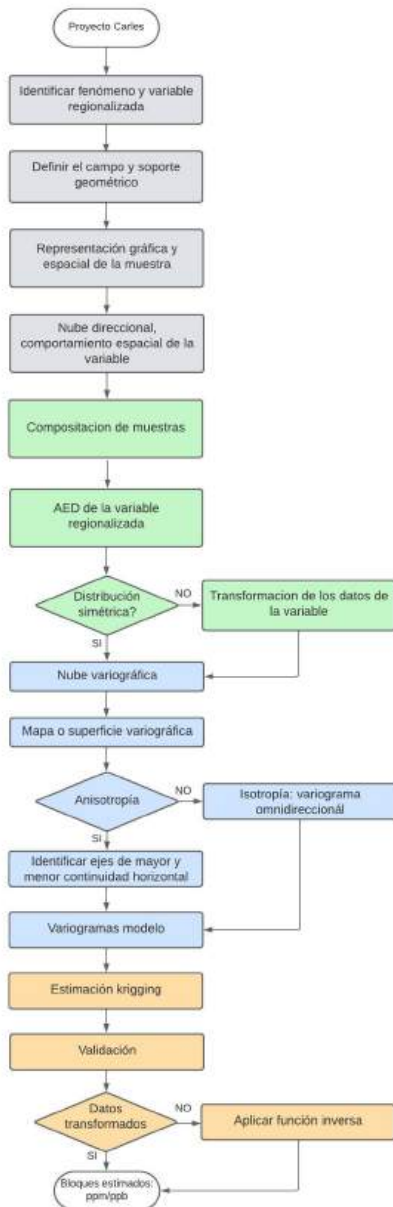


Figura 2: Diagrama de flujo de la metodología usada para el cálculo de recursos minerales del proyecto Carlés

La base de datos se encuentra conformada por 87 sondajes, la información de los mismos es importada a RecMin como son: datos de sondeo, desviaciones, litología, muestras. Al momento de cargar las ultimas, se debe agregar las que no se tenga en la sección de índice de elementos y posteriormente en control de muestras, posteriormente, activar los ca-

silleros agregados, para que se muestren al momento de importar los sondajes. En la sección dibujo, se representa la información y se procedió a realizar las correcciones de los errores correspondientes en cada sondaje.

REPRESENTACIÓN GRÁFICA.- En este caso de estudio, se pretende realizar la modelización de las reservas de Au, basados en los datos correspondientes a 87 sondajes obtenidos del Proyecto Carlés y procesados en diferentes softwares (Figura 3).

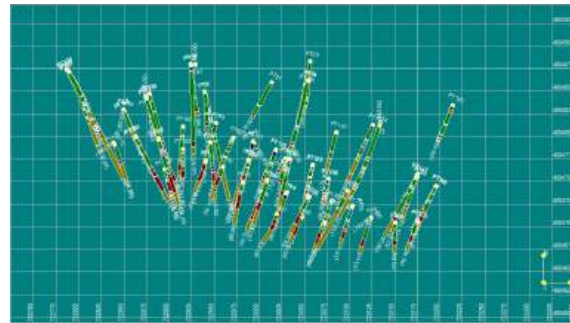


Figura 3: Representación de los sondeos del proyecto Carlés en RecMin

REALIZAR LAS SECCIONES Y DEFINIR CUERPOS.-En la sección de dibujo se carga los sondajes y litologías correspondientes y se procede a realizar secciones delimitando los sondajes y tomando en cuenta como se encuentran dispuestos. Para el caso de estudio se realizarán dos cuerpos. Se asigna un color para el cuerpo superior e inferior, y se dibuja las secciones necesarias y posteriormente verificar los vértices y forma de cada una y toda esta información se guarda como fichero, con extensión (.lin). Uno de los criterios que se puede usar, es realizarlo con el cuerpo, y ver hasta que litología es el cuerpo superior, establecer en el índice de zonas tanto para el superior como para el inferior. Para generar los cuerpos sólidos, se selecciona las secciones de cada uno de los cuerpos correspondientes en

lista de objetos y se procede a unir con T3 auto. En los extremos del cuerpo, se selecciona editar superficies, dar clic a cualquier vértice y seleccionar rellenar con T3 para completar la estructura. Para el cálculo del volumen de cada cuerpo, se une las secciones activas en lista de mallados para posteriormente dar clic en Sup-Lin-T3 y de esta manera calcular el volumen del T3 cerrado.

BLOQUES.- Definir el área que comprende los sondeos y la profundidad mayor de los sondeos, posterior a esto crear una nueva base de datos de bloques y nos dirigimos a nueva tabla de bloques la misma que es nombrada como (BLKbloques). Luego se debe definir el tamaño de los bloques que se van a generar, en este caso 5x5x5 y las coordenadas de la sección inferior izquierda. Se genera los archivos para el SGEMS, o archivos txt, filtrado, por zonas, se seleccionaba cada uno de los cuerpos, y se debe guardar con extensión txt para Excel. En un principio se delimito el área para poder hacer los bloques, en base a coordenadas, con los T3 se desactivo bloques fuera de la línea existente para poder ver los bloques. Generación de los bloques: Se establece los cuatro vértices que cubren las áreas de los sondeos, mediante las coordenadas se cubre el área con bloques, estos bloques se delimitaban dentro de los T3 y a estos se los exportaba como bloques vacíos. Anteriormente generado la base de datos de bloques y la tabla de bloques

COMPÓSITOS.- Antes de hacer los compósitos hay que ir al módulo de yacimientos, seleccionamos al proyecto, y se realiza un índice de zonas, para poder discriminar los dos cuerpos poniendo los compósitos en cada zona. Con densidad de 2.8 Volver al modulo de dibujo, poner los sondeos, poner las secciones, poner vista adecuada, seleccionar el sondeo, con clic derecho vamos a ir a ver datos de muestras y empezamos a discriminar.

SGEMS.- Cargar los objetos, luego se carga los archivos exportados de RecMin, e irle acortando los bloques con la finalidad de que se quede al ras con el sondeo final. Se realizo una tabla en Excel con todas las coordenadas de los vértices, y se empieza a definir la z inferior, y se empieza a definir el tamaño del bloque y, se filtra el dato menor porque se quiere ver la cota en profundidad. Se utiliza la cota de superficie, restado para la cota en profundidad del bloque y eso se lo obtiene de SGEMS.

RSTUDIO.- Se realizó el análisis de datos exploratorio, estadística descriptiva, con la finalidad de observar los valores atípicos, el sesgo. Para esto se aplicó normalización mediante logaritmo en base 10. Y el mismo proceso se realizó en Excel y se guardó en formato csv. (Figura 4)

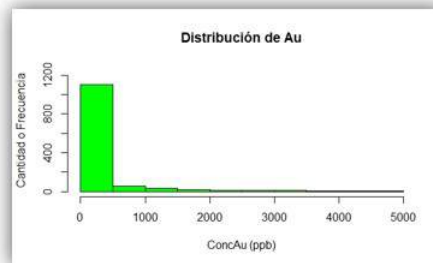


Figura 4: Representación en histograma de la distribución del Au

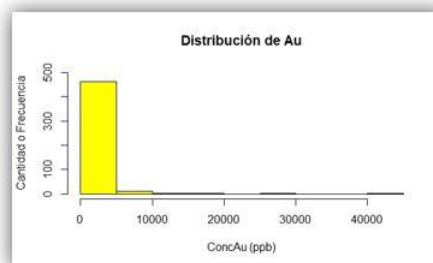


Figura 5: Representación en histograma de la distribución del Au

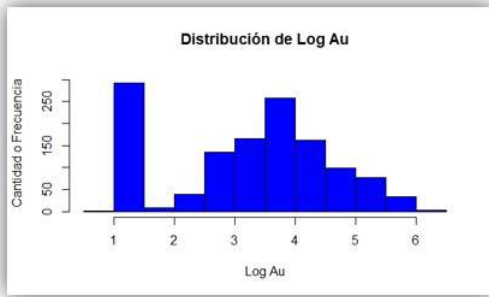


Figura 6: Representación en histograma normalizado de la distribución del Au

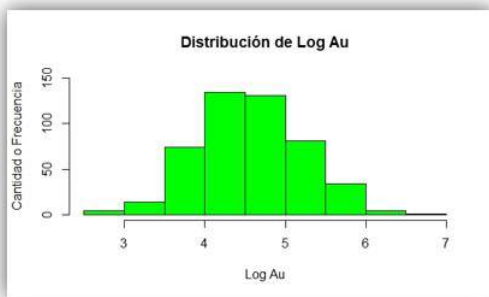


Figura 7: Representación en histograma normalizado de la distribución del Au

SAGAGIS.- Se generan las superficies variográficas, para así definir el acimut y bz del variograma, posteriormente se realizó el histograma, para poder observar la varianza en el variograma. Nuestra área en total tenía 500, buscar un valor de número de lags para que nos pueda dar un adecuado variograma,, definir la tolerancia el ancho de banda, etc

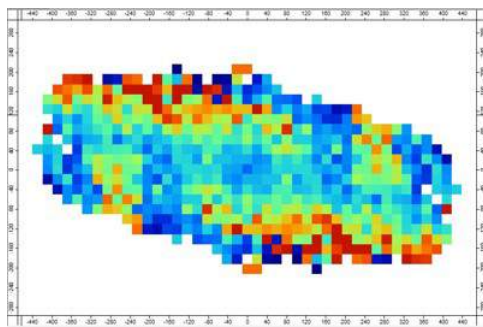


Figura 8: Superficie variográfica cuerpo superior

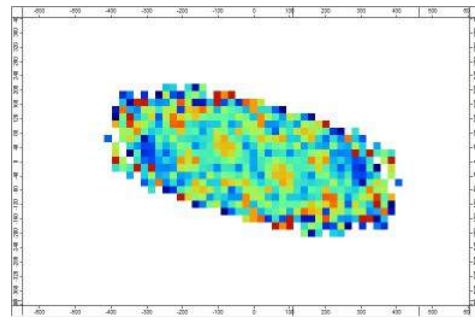


Figura 9: Superficie variográfica cuerpo inferior

4. Metodología-Machine Learning

Este método, se realizó siguiendo el diagrama de flujo presentado a continuación.

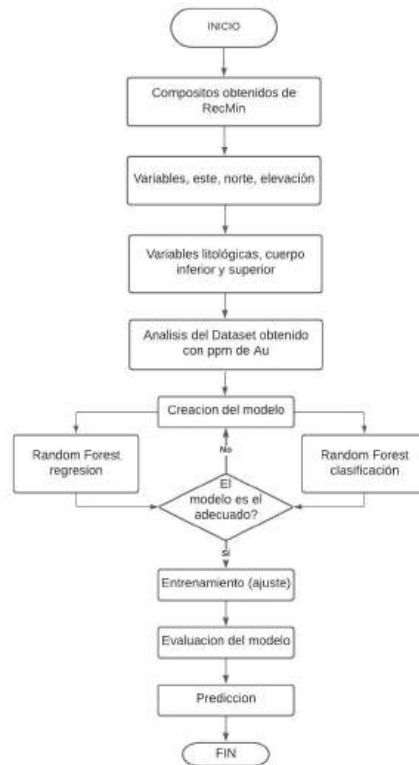


Figura 10: Diagrama de flujo de la metodología usada para el cálculo de reservas usando Machine Learning

La información base, fue obtenida a partir del

dataset, en este caso los compósitos generados en RecMin. Mediante el software RecMin se obtiene un dataset con las variables correspondientes a los sondeos, su litología, representada en un cuerpo superior e inferior y el elemento de interés, en este caso el oro. A continuación se realizó un pre-procesamiento del Dataset y posteriormente una normalización de las variables.

1. Importación del Dataset.

2. Pre-procesamiento del Dataset.

Las variables definidas en el dataset fueron: este, norte y elevación obtenidas mediante la tabla de compósitos de RecMin. La variable independiente Au, que en este caso es el mineral de interés.

3. Creación del modelo Random Forest

Es un método de aprendizaje conjunto para la clasificación, la regresión y tareas que operan mediante la construcción de una variedad de árboles de decisión en el momento del entrenamiento. Este procedimiento de arranque conduce a un mejor rendimiento del modelo porque disminuye la varianza del modelo, sin aumentar el sesgo. Esto significa que mientras que las predicciones de un solo árbol son muy sensibles al ruido en su conjunto de entrenamiento, el promedio de muchos árboles no lo es, siempre que los árboles no estén correlacionados. Simplemente entrenar muchos árboles en un solo conjunto de entrenamiento daría árboles fuertemente correlacionados.

4. Entrenamiento (training) - ajuste

En esta etapa se establecieron los parámetros e hiperparámetros con los que se va a realizar el entrenamiento, el cual consiste en optimizar al modelo, para ello a medida que se va entrenando se utiliza el conjunto de datos de validación para verificar la optimización.

5. Evaluación del Modelo

En este ítem se realiza una evaluación previa, conjuntamente con los datos de entrenamiento y test para conocer el funcionamiento del modelo. Para ello se utiliza la métrica denominada como Accuracy mediante la cual se obtuvo una valoración de 100.

6. Predicción

Mediante el dataset se asigna datos a cada variable, a la vez se utilizan comandos para que los mismos sean leídos por el modelo. Posteriormente se realiza la normalización de los datos observando que los valores estén dentro de los rangos que se establecieron. Finalmente se aplica el predict llamando al arreglo numpy.

5. Resultados

La estimación por modelo de bloques obtenidos mediante la representación de los datos en el software Recmin, reflejan los siguientes resultados. Se seleccionaron 3 métodos de clasificación de datos en función de los resultados obtenidos, seleccionando el método que más se ajusta al modelo.

5.1. Resultados convencionales Cuerpo Superior

5.1.1. Clasificación Intervalo igual

Consiste en dividir la amplitud máxima existente entre los datos, entre el número de clases que se haya elegido $(n_f - n_i) / n$. Las figuras 11 y 12 representan las vistas superior y 3D del cuerpo superior, se observa de acuerdo a la vista superior que las mayores concentraciones de Au se encuentran ubicadas hacia la parte SW del cuerpo, pudiendo determinar que la mayoría de concentraciones de acuerdo a la presente clasificación se encuentran en rangos bajos y dominantes en el cuerpo.

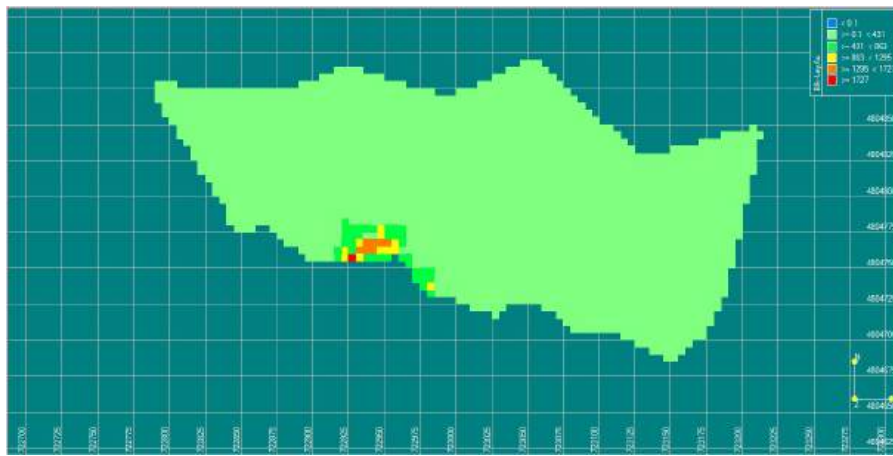


Figura 11: Vista superior modelo de estimación mediante clasificación Intervalo Igual

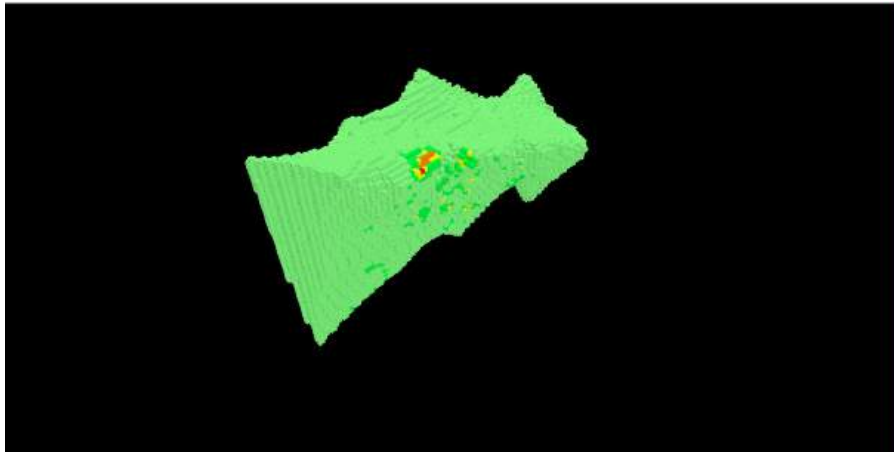


Figura 12: Vista 3D modelo de estimación mediante clasificación Intervalo Igual

5.1.2. Clasificación Cuantiles

En una clasificación de cuantiles, cada clase contiene el mismo número de entidades. La clasificación de cuantiles es adecuada para los datos distribuidos linealmente. El cuantil asigna el mismo número de valores de datos a cada clase. Las figuras 13 y 14 representan las vistas superior y 3D del cuerpo superior clasificado mediante cuantiles, en las mismas se puede evidenciar que las concentraciones se

encuentran de cierto modo un poco más dispersas obteniendo así de acuerdo a la vista superior concentraciones de Au hacia la parte N y S del cuerpo, sin embargo, mediante la vista 3D se observa que las concentraciones más altas se encuentra en la parte frontal del cuerpo siendo dominantes, lo cual puede ser resultado de una clasificación errónea debido a que visualmente se observa un exceso en lo que respecta a las concentraciones altas.

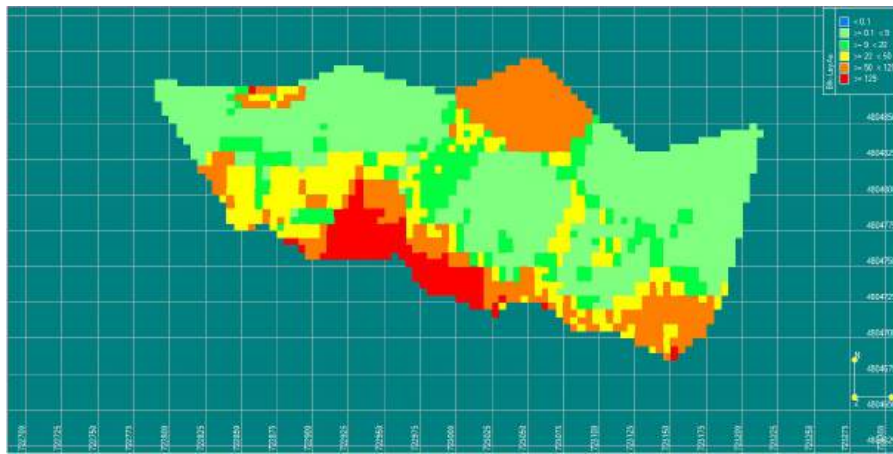


Figura 13: Vista superior modelo de estimación mediante clasificación Cuantiles

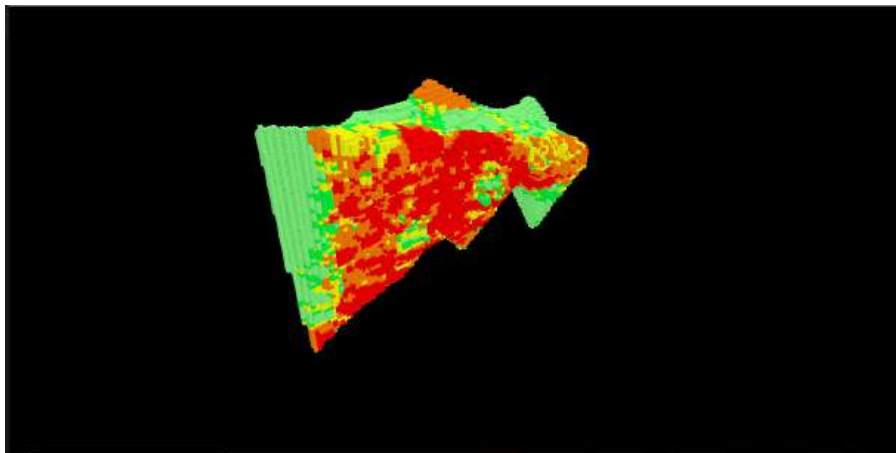


Figura 14: Vista 3D modelo de estimación mediante clasificación Cuantiles

5.1.3. Clasificación Natural breaks (Jenks)

Con una clasificación de cortes naturales (Jenks), las clases se basan en las agrupaciones naturales inherentes a los datos. Las rupturas de clase se crean de manera que los valores similares se agrupan mejor y se maximizan las diferencias entre clases. Las figuras 15 y 16 muestran que las mayores concentraciones de acuerdo a la clasificación

por natural breaks se encuentran en la parte SW de la vista superior y en la parte frontal de la vista 3D, observándose que las concentraciones bajas son las que se encuentran mayormente en el cuerpo, se evidencia que entre la clasificación por natural breaks y la clasificación por intervalo igual muestra coincidencia en cuanto a la dominancia de las concentraciones bajas.

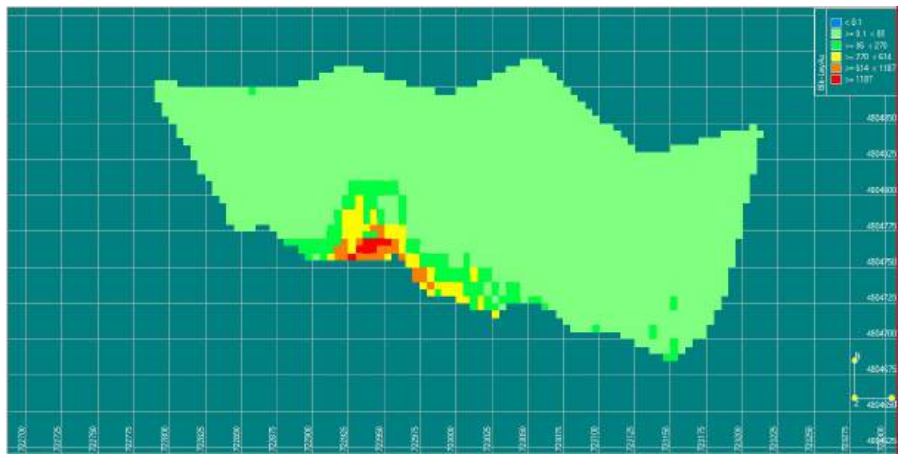


Figura 15: Vista superior modelo de estimación mediante clasificación Natural Breaks (Jenks)

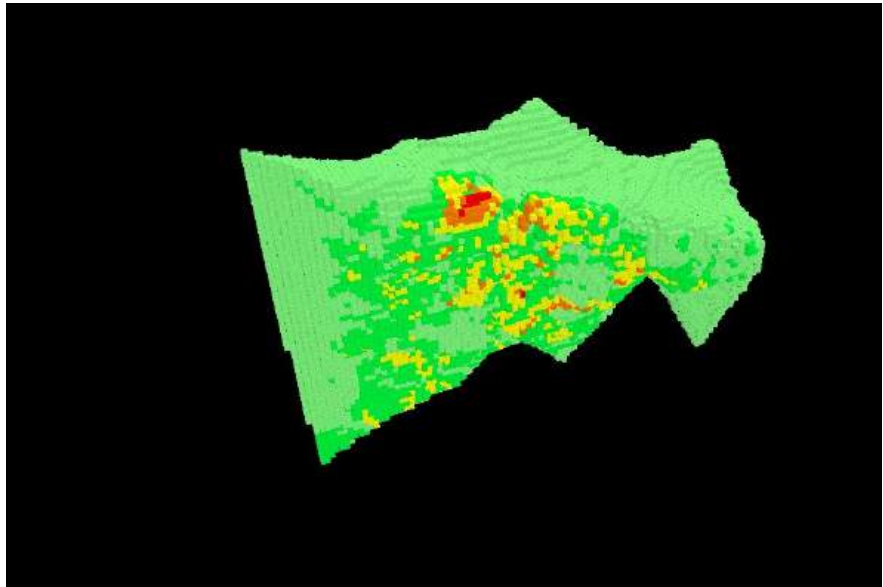


Figura 16: Vista 3D modelo de estimación mediante clasificación Natural Breaks (Jenks)

5.2. Resultados convencionales Cuerpo Inferior

5.2.1. Clasificación Intervalo igual

Las figuras 17 y 18 representan al cuerpo inferior categorizando sus leyes de acuerdo a la clasificación de intervalo igual. En primera instancia se puede evidenciar que las concentraciones bajas y modera-

das se encuentran más distribuidas y dominantes en el cuerpo, mientras que las concentraciones altas se presentan de manera puntual indicando su escasa presencia, no obstante, es importante recalcar que al igual que el cuerpo superior se observa un marcado control de las muestras con leyes de Au muy bajas.

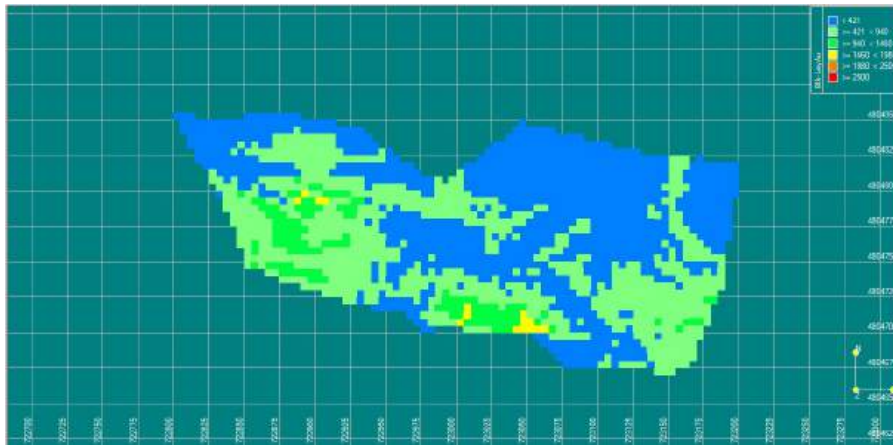


Figura 17: Representación vista superior modelo de estimación mediante clasificación Intervalo Igual

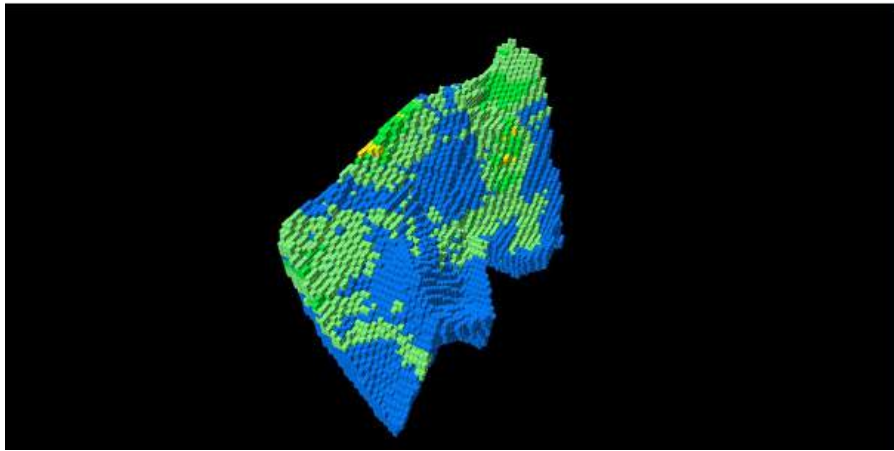


Figura 18: Vista 3D modelo de estimación mediante clasificación Intervalo Igual

5.2.2. Clasificación Cuantiles

En las figuras 19 y 20 se evidencia que las concentraciones de leyes de Au bajas se encuentran mayormente distribuidas y dominando en el cuerpo, sin

embargo, se evidencia una distribución tanto en la parte occidental como oriental del cuerpo de las concentraciones de leyes moderadas a altas, indicando que de acuerdo a esta clasificación existe una mejor deistribución o agrupación de las concentraciones.

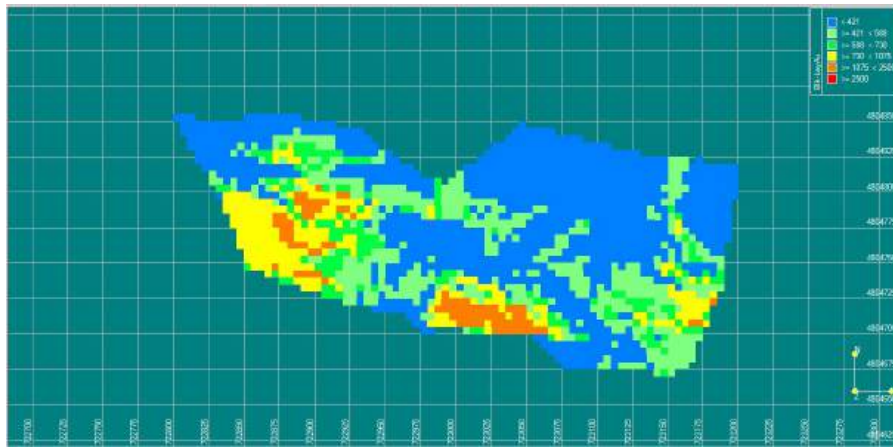


Figura 19: Vista superior modelo de estimación mediante clasificación Cuantiles

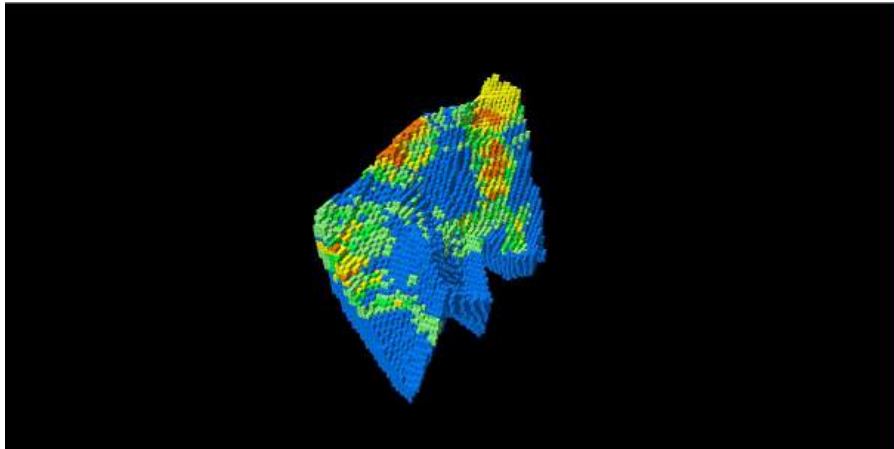


Figura 20: Vista 3D modelo de estimación mediante clasificación Cuantiles

5.2.3. Clasificación Natural breaks (Jenks)

Las figuras 21 y 22 muestran al igual que en las anteriores representaciones, que las concentraciones bajas dominan en el cuerpo, sin embargo, se evidencia que existe similitud con la clasificación de

cuantiles, con la única diferencia que las concentraciones moderadas presentan mayor relevancia que las concentraciones altas mismas que se encuentran distribuidas de manera puntual hacia la parte S y W de la vista superior.

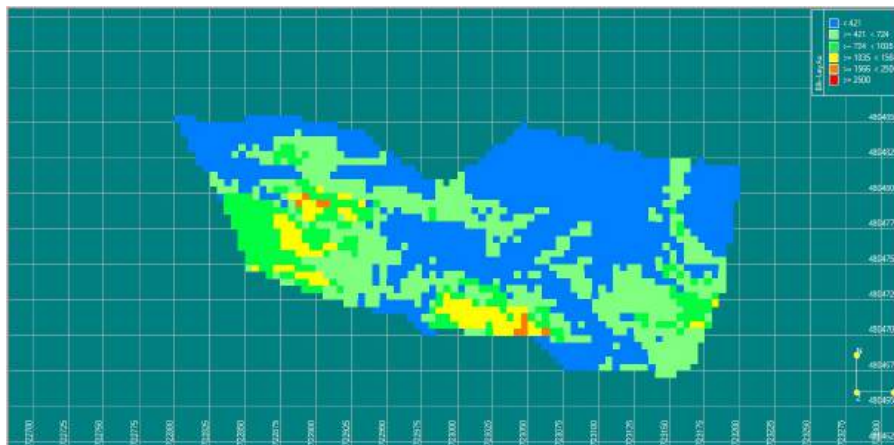


Figura 21: Vista superior modelo de estimación mediante clasificación Natural Breaks (Jenks)

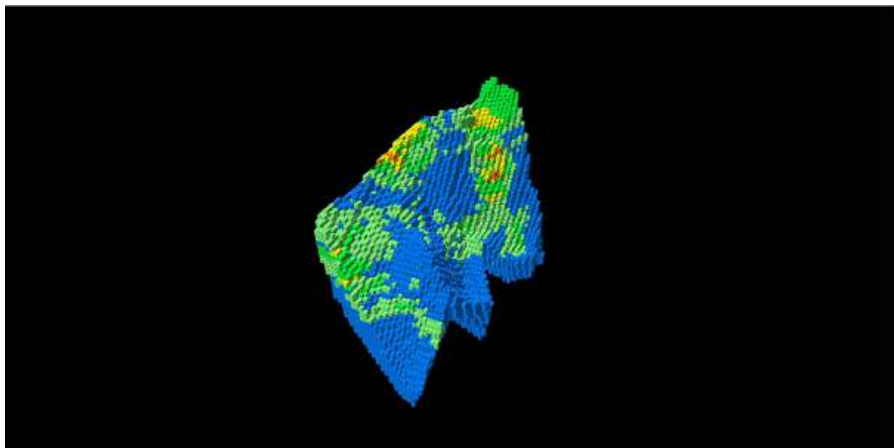


Figura 22: Vista 3D modelo de estimación mediante clasificación Natural Breaks (Jenks)

5.3. Resultado mediante Machine Learning

Las figuras 23 y 24 muestran la distribución de las concentraciones de Au de acuerdo a las predicciones realizadas mediante el algoritmo Random

Forest, se evidencia que la distribución de las mismas presentan similitud con los resultados convencionales, no obstante, se puede observar que existe una mayor distribución hacia la parte occidental del cuerpo, y una mayor propagación hacia la parte inferior del cuerpo total.

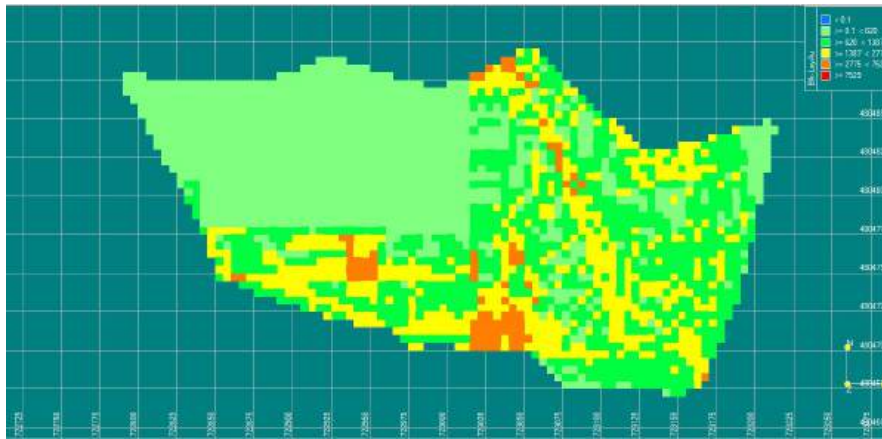


Figura 23: Vista superior modelo de estimación mediante clasificación Natural Breaks (Jenks)

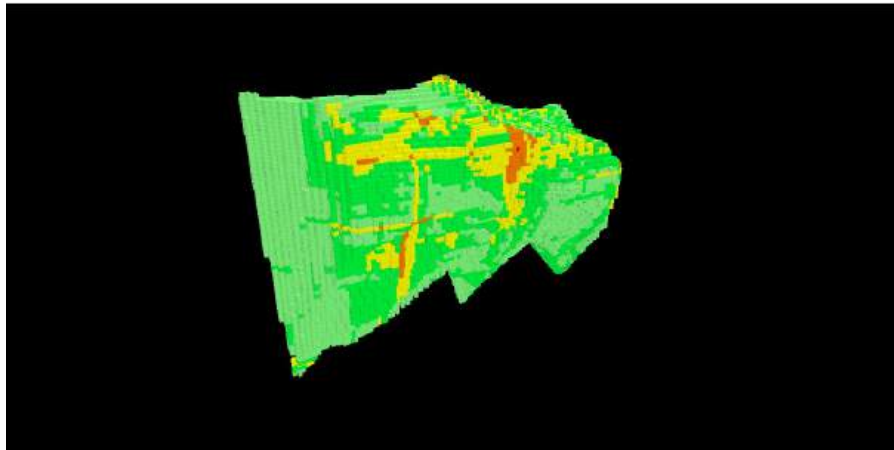


Figura 24: Vista 3D modelo de estimación mediante clasificación Natural Breaks (Jenks)

6. Conclusiones

En los resultados obtenidos mediante la metodología convencional para el cuerpo superior, se evidencia que existe similitud entre la clasificación por intervalo igual y la realizada por natural breaks, ambas divergen notoriamente respecto a la clasificación por cuantiles la cual debido a estas diferencias es descartada como la más adecuada, sin embargo, a pesar de la similitud entre la de intervalo igual y la de natural breaks, se ha optado por establecer como la representación más apropiada a la de natural breaks debido al criterio de maximizar las diferencias entre grupos que utiliza para la clasi-

ficación. Además, la dominancia de concentraciones bajas en el cuerpo inferior, se debe principalmente a la frecuencia de estos en la base de datos.

En los resultados obtenidos para el cuerpo inferior a través de la metodología convencional, se puede evidenciar de primera mano que a diferencia del cuerpo superior existe un menor control de las concentraciones bajas, por otro lado, se observa que de manera similar al primer caso existe similitud entre la clasificación de los cuantiles y la de natural breaks, estableciendo como la más apropiada a la clasificación por natural breaks puesto que la metodología de clasificación de los cuantiles no es la más apropiada para el establecimiento de grupos

o clases.

Finalmente los resultados obtenidos mediante el algoritmo Random Forest muestran una distribución no muy similar a las obtenidas mediante la metodología convencional, al igual que los resultados previos la clasificación por natural breaks debido a su distribución y a su criterio de clasificación ha sido seleccionada como la más adecuada para representar la estimación.

Se determina que la metodología convencional es aquella que representa los mejores resultados debido a que la metodología por Machine Learning presenta un error del 23 %, dicho error se le atribuye a la poca cantidad de datos con la cual se entrena y evalúa al modelo lo cual es una limitante para el algoritmo Random Forest que por definición trabaja mejor con una gran cantidad de datos, adicionalmente la cantidad de variables es muy baja lo cual nos permite determinar que el modelo posiblemente posee underfitting. Por tanto, debido a

las limitaciones que se exponen se considera que los resultados obtenidos por el algoritmo no son representativos para este caso.

7. Recomendaciones

Se debe prestar mucha atención a los resultados que se obtienen de la estadística descriptiva, debido a que son de gran importancia para una adecuada normalización de la información disponible. Además es importante conocer las características físico-químicas de los elementos que se quiera hacer el calculo de reservas, debido a que su distribución va a estar relacionada con el modelado.

En general, la elección de un método de estimación de recursos naturales depende de la naturaleza de los datos disponibles y del tipo de recurso que se está evaluando. La combinación de métodos convencionales y de Machine Learning puede proporcionar una estimación más precisa y detallada de los recursos naturales.

Referencias

- Hernández. (2017). Gestión de riesgos y desastres socioambientales. el caso de la mina buenavista del cobre de cananea. *Investigaciones geográficas*(93), 0–0.
- IGME. (1988). El yacimiento aurífero de carlés (asturias). *Mineralogía*, 11(1), 47–53.
- Lara. (2020). Estimación de recursos minerales en dominios geometalúrgicos.